KAPLAN & SADOCK'S

COMPREHENSIVE TEXTBOOK OF PSYCHIATRY

TENTH EDITION

Benjamin James Sadock, M.D.
Virginia Alcott Sadock, M.D.
Pedro Ruiz, M.D.

Wolters Kluwer

VOLUME ONE

KAPLAN & SADOCK'S COMPREHENSIVE TEXTBOOK OF PSYCHIATRY

TENTH EDITION

Benjamin James Sadock, M.D.
Virginia Alcott Sadock, M.D.
Pedro Ruiz, M.D.

Wolters Kluwer

VOLUME ONE

Contributing Editors

Hagop S. Akiskal, M.D.
Professor, Department of Psychiatry, Director, International Mood Center, University of California San Diego School of Medicine, La Jolla, California; Chief, Mood Disorders, VA San Diego Healthcare System, San Diego, California.

Dilip V. Jeste, M.D.
Senior Associate Dean for Healthy Aging and Senior Care, Estelle and Edgar Levi Chair in Aging, Distinguished Professor of Psychiatry and Neurosciences, Director, Sam and Rose Stein Institute for Research on Aging, University of California San Diego School of Medicine, La Jolla, California.

John H. Krystal, M.D.
Robert L. McNeil, Jr., Professor of Translational Research; Chair of Psychiatry, Yale University School of Medicine; Chief of Psychiatry Services, Yale-New Haven Hospital; Director, Clinical Neuroscience Division, VA National Center for PTSD; Director, NIAAA Center for the Translational Neuroscience of Alcoholism, New Haven, Connecticut.

Roberto Lewis-Fernández, M.D.
Professor of Psychiatry, Columbia University College of Physicians and Surgeons; Director, New York State Center of Excellence for Cultural Competence, Director, Hispanic Treatment Program, New York State Psychiatric Institute, New York, New York.

Kathleen Ries Merikangas, Ph.D.
Senior investigator, Chief, Genetic Epidemiology, Research Branch, Division of Intramural Research Programs at the National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland.

Charles B. Nemeroff, M.D., Ph.D.
Professor and Chair, Department of Psychiatry and Behavioral Science, University of Miami Leonard M. Miller School of Medicine, Miami, Florida.

Karin Neufeld, M.D., M.P.H.
Associate Professor, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine; Clinical Director of Psychiatry, Johns Hopkins Bayview Medical Center, Baltimore, Maryland.

Caroly Sue Pataki, M.D.
Clinical Professor of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.

Robert G. Robinson, M.D.
Professor, Department of Psychiatry, University of Iowa, Iowa City, Iowa.

Eric Cameron Strain, M.D.
Professor, Director, Johns Hopkins Center for Substance Abuse Treatment and Research, Medical Director, Behavioral Pharmacology Research Unit, Johns Hopkins University School of Medicine, Baltimore, Maryland.

Robert A. Sweet, M.D.
Professor of Psychiatry and Neurology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

Carol A. Tamminga, M.D.
Professor and Chairman, Department of Psychiatry, University of Texas Southwestern Medical School, Dallas, Texas.

KAPLAN & SADOCK'S

COMPREHENSIVE TEXTBOOK OF PSYCHIATRY

VOLUME I/II

TENTH EDITION

EDITORS

Benjamin J. Sadock, M.D.

Menas S. Gregory Professor of Psychiatry,
Department of Psychiatry, New York University School of Medicine;
Attending Psychiatrist, Tisch Hospital;
Attending Psychiatrist, Bellevue Hospital Center;
New York, New York

Virginia A. Sadock, M.D.

Professor of Psychiatry, Department of Psychiatry, New York University School of Medicine;
Attending Psychiatrist, Tisch Hospital; Attending Psychiatrist, Bellevue Hospital Center, New York, New York

Pedro Ruiz, M.D.

Clinical Professor, Menninger Department of Psychiatry and Behavioral Sciences
Baylor College of Medicine Houston, Texas

Wolters Kluwer

Philadelphia • Baltimore • New York • London
Buenos Aires • Hong Kong • Sydney • Tokyo

Acquisitions Editor: Chris Teja
Product Development Editor: Andrea Vosburgh, Franny Murphy
Editorial Assistant: Brian Convery
Marketing Manager: Rachel Mante Leung
Production Project Manager: Bridgett Dougherty
Design Coordinator: Holly McLaughlin
Manufacturing Coordinator: Beth Welsh
Prepress Vendor: Aptara, Inc.

10th edition

Copyright © 2017 Wolters Kluwer.

"Kaplan Sadock Psychiatry" with the pyramid logo is a trademark of Lippincott Williams & Wilkins.

Copyright © 2009 by Lippincott Williams & Wilkins. All rights reserved. This book is protected by copyright. No part of this book may be reproduced or transmitted in any form or by any means, including as photocopies or scanned-in or other electronic copies, or utilized by any information storage and retrieval system without written permission from the copyright owner, except for brief quotations embodied in critical articles and reviews. Materials appearing in this book prepared by individuals as part of their official duties as U.S. government employees are not covered by the above-mentioned copyright. To request permission, please contact Wolters Kluwer at Two Commerce Square, 2001 Market Street, Philadelphia, PA 19103, via email at permissions@lww.com, or via our website at lww.com (products and services).

9 8 7 6 5 4 3 2 1

Printed in China

978-1-4511-0047-1
1-4511-0047-7
Library of Congress Cataloging-in-Publication Data
available upon request

This work is provided "as is," and the publisher disclaims any and all warranties, express or implied, including any warranties as to accuracy, comprehensiveness, or currency of the content of this work.

This work is no substitute for individual patient assessment based upon healthcare professionals' examination of each patient and consideration of, among other things, age, weight, gender, current or prior medical conditions, medication history, laboratory data and other factors unique to the patient. The publisher does not provide medical advice or guidance and this work is merely a reference tool. Healthcare professionals, and not the publisher, are solely responsible for the use of this work including all medical judgments and for any resulting diagnosis and treatments.

Given continuous, rapid advances in medical science and health information, independent professional verification of medical diagnoses, indications, appropriate pharmaceutical

selections and dosages, and treatment options should be made and healthcare professionals should consult a variety of sources. When prescribing medication, healthcare professionals are advised to consult the product information sheet (the manufacturer's package insert) accompanying each drug to verify, among other things, conditions of use, warnings and side effects and identify any changes in dosage schedule or contraindications, particularly if the medication to be administered is new, infrequently used or has a narrow therapeutic range. To the maximum extent permitted under applicable law, no responsibility is assumed by the publisher for any injury and/or damage to persons or property, as a matter of products liability, negligence law or otherwise, or from any reference to or use by any person of this work.

LWW.com

Dedicated to all persons who work with and care for the mentally ill

About the Editors

BENJAMIN J. SADOCK, M.D.

Benjamin James Sadock, M.D., is the Menas S. Gregory Professor of Psychiatry in the Department of Psychiatry at the New York University (NYU) School of Medicine. He is a graduate of Union College, received his M.D. degree from New York Medical College, and completed his internship at Albany Hospital. He completed his residency at Bellevue Psychiatric Hospital and then entered military service as Captain US Air force, where he served as Acting Chief of Neuropsychiatry at Sheppard Air Force Base in Texas. He has held faculty and teaching appointments at Southwestern Medical School and Parkland Hospital in Dallas and at New York Medical College, St. Luke's Hospital, the New York State Psychiatric Institute, and Metropolitan Hospital in New York City. Dr. Sadock joined the faculty of the NYU School of Medicine in 1980 and served in various positions: Director of Medical Student Education in Psychiatry, Co-Director of the Residency Training Program in Psychiatry, and Director of Graduate Medical Education. Currently, Dr. Sadock is Co-Director of Student Mental Health Services, Psychiatric Consultant to the Admissions Committee, and Co-Director of Continuing Education in Psychiatry at the NYU School of Medicine. He is on the staff of Bellevue Hospital and Tisch Hospital. Dr. Sadock is a Diplomate of the American Board of Psychiatry and Neurology and served as an Associate Examiner for the Board for more than a decade. He is a Distinguished Life Fellow of the American Psychiatric Association, a Fellow of the American College of Physicians, a Fellow of the New York Academy of Medicine, and a member of Alpha Omega Alpha Honor Society. He is active in numerous psychiatric organizations and was president and founder of the NYU-Bellevue Psychiatric Society. Dr. Sadock was a member of the National Committee in Continuing Education in Psychiatry of the American Psychiatric Association, served on the Ad Hoc Committee on Sex Therapy Clinics of the American Medical Association, was a Delegate to the Conference on Recertification of the American Board of Medical Specialists, and was a representative of the American Psychiatric Association Task Force on the National Board of Medical Examiners and the American Board of Psychiatry and Neurology. In 1985, he received the Academic Achievement Award from New York Medical College and was appointed Faculty Scholar at NYU School of Medicine in 2000. He is the author or editor of more than 100 publications (including 50 books), a reviewer for psychiatric journals, and lectures on a broad range of topics in general psychiatry. Dr. Sadock maintains a private practice for diagnostic consultations and psychiatric treatment. He has been married to Virginia Alcott Sadock, M.D., Professor of Psychiatry at NYU School of Medicine, since completing his residency. Dr. Sadock enjoys opera, golf, traveling, and is an enthusiastic fly fisherman.

VIRGINIA A. SADOCK, M.D.

Virginia Alcott Sadock, M.D., joined the faculty of the New York University (NYU) School of Medicine in 1980, where she is currently Professor of Psychiatry and Attending Psychiatrist at the Tisch Hospital and Bellevue Hospital. She is Director of the Program in Human Sexuality at the NYU Langone Medical Center, one of the largest treatment and training programs of its kind in the United States. She is the author of more than 50 articles and chapters on sexual behavior and was the developmental editor of *The Sexual Experience*, one of the first major textbooks on human sexuality, published by Williams & Wilkins. She serves as a referee and book reviewer for

several medical journals, including the *American Journal of Psychiatry* and the *Journal of the American Medical Association*. She has long been interested in the role of women in medicine and psychiatry and was a founder of the Committee on Women in Psychiatry of the New York County District Branch of the American Psychiatric Association. She is active in academic matters, served as an Assistant and Associate Examiner for the American Board of Psychiatry and Neurology for more than 20 years, and was also a member of the Test Committee in Psychiatry for both the American Board of Psychiatry and the Psychiatric Knowledge and Self-Assessment Program (PKSAP) of the American Psychiatric Association. She has chaired the Committee on Public Relations of the New York County District Branch of the American Psychiatric Association, has been a regional council member of the American Association of Sex Education Counselors and Therapists, a founding member of The Society of Sex Therapy and Research, and is President of the NYU Alumni Association of Sex Therapists. She has participated in the National Medical Television Network series *Women in Medicine* and the Emmy Award–winning PBS television documentary *Women and Depression* and currently hosts the radio program *Sexual Health and Well-being* (Sirius-XM) at NYU Langone Medical Center. She lectures extensively both in this country and abroad on sexual dysfunction, relational problems, and depression and anxiety disorders. She is a Distinguished Fellow of the American Psychiatric Association, a Fellow of the New York Academy of Medicine, and a Diplomate of the American Board of Psychiatry and Neurology. Dr. Sadock is a graduate of Bennington College, received her M.D. degree from New York Medical College, and trained in psychiatry at Metropolitan Hospital. She lives in Manhattan with her husband, Dr. Benjamin Sadock, where she maintains an active practice that includes individual psychotherapy, couples and marital therapy, sex therapy, psychiatric consultation, and pharmacotherapy. She and her husband have two children, James and Victoria, both emergency physicians, and three grandchildren, Oliver, Emily and Celia. In her leisure time, Dr. Sadock enjoys theater, film, golf, reading fiction, and travel.

PEDRO RUIZ, M.D.

Pedro Ruiz, M.D., is currently Clinical Professor at the Menninger Department of Psychiatry & Behavioral Sciences of Baylor College of Medicine in Houston, Texas and also Professor Emeritus in the Department of Psychiatry & Behavioral Sciences of Albert Einstein College of Medicine of Yeshiva University in New York City, New York. He graduated from medical school at the University of Paris in France. He conducted his residency training in psychiatry at the University of Miami Leonard M. Miller School of Medicine in Florida. He has held faculty appointments at a professorial level at Albert Einstein College of Medicine in New York City, Baylor College of Medicine in Houston, the University of Texas Medical School at Houston, and the University of Miami Leonard M. Miller School of Medicine in Miami. He has served in various academic positions during his career; among them: Director of the Lincoln Hospital Community Mental Health Center, Director of the Bronx Psychiatric Center, Assistant Dean and Vice Chair of the Department of Psychiatry at the Albert Einstein College of Medicine in New York City; Chief of Psychiatric Service at Ben Taub General Hospital and Vice Chair of the Department of Psychiatry at Baylor College of Medicine in Houston, Texas; Medical Director of the University of Texas Mental Sciences Institute, Vice Chair and Interim Chair of the Department of Psychiatry at the University of Texas Medical School at Houston, Texas; and Executive Vice Chair of the Department of Psychiatry and Behavioral Sciences of the University of Miami Leonard M. Miller School of Medicine in Miami, Florida. Dr. Ruiz is also a Distinguished Life Fellow of the American Psychiatric Association, a Fellow of the American College of Psychiatrists, the American Association for Social Psychiatry, the Benjamin Rush Society and the American Group Psychotherapy Association, and an Honorary Fellow of the World Psychiatric Association. He is also a member of the American Academy of Addiction Psychiatry, the Group for the Advancement of Psychiatry, the American Association of Community Psychiatrists and the American Association of Psychiatric Administrators. Dr. Ruiz has also been President of the American College of Psychiatrists (2000–2001), the American Association for Social Psychiatry (2000–2002), the American Board of Psychiatry and Neurology (2002–2003), the American Psychiatric Association (2006–2007), the World Psychiatric Association (2011–2014), and the World Association on Dual Disorders (2015–2018). Dr. Ruiz has also served in more than 45 Editorial Boards, among them: *The American Journal of Psychiatry, Psychiatric Services, The*

American Journal on Addictions, Addictive Disorders and Their Treatment, and *The Journal of Nervous and Mental Disease*. Dr. Ruiz has also received over 65 awards and honors and has published over 600 publications, including over 20 books. Dr. Ruiz has also delivered over 250 grand rounds and invited lectures, and has made over 500 worldwide scientific presentations. Dr. Ruiz and his wife Angela have two children: Pedro Pablo and Angela Maria, and four wonderful grandchildren: Francisco Antonio, Pedro Pablo Jr., Omar Joseph, III, and Pablo Antonio. Dr. Ruiz also enjoys reading literary novels, theater, traveling, and fishing.

Contents

About the Editors

Contributors

Preface

Foreword

Herbert Pardes, M.D.

Volume I

1 NEURAL SCIENCES

1.1 The Neuroscience of Psychiatry

John H. Krystal, M.D.

1.2 Functional Neuroanatomy

Darlene Susan Melchitzky, M.S., and David A. Lewis, M.D.

1.3 Neural Development and Neurogenesis

Emanuel DiCicco-Bloom, M.D., and Maryann Obiorah, Ph.D.

1.4 Biogenic Amine Neurotransmitters

John D. Elsworth, Ph.D., and Robert H. Roth, Ph.D.

1.5 Amino Acid Neurotransmitters

Joseph T. Coyle, M.D.

1.6 Neuropeptides: Biology, Regulation, and Role in Neuropsychiatric Disorders

Shaun P. Brothers, Ph.D., Larry James Young, Ph.D., and Charles B. Nemeroff, M.D., Ph.D.

1.7 Neurotrophic Factors

Francis S. Lee, M.D., Ph.D., and Moses V. Chao, Ph.D.

1.8 Novel Neurotransmitters

Thomas W. Sedlak, M.D., Ph.D., and Adam I. Kaplin, M.D., Ph.D.

1.9 Intraneuronal Signaling

John A. Gray, M.D., Ph.D., and Bryan L. Roth, M.D., Ph.D.

1.10 Cellular and Synaptic Basis of Neural Signaling
Charles F. Zorumski, M.D., and Steven Mennerick, Ph.D.

1.11 Genome, Transcriptome, and Proteome: The Molecular Genetics and Biochemistry Underlying the Neurobiology of Mental Disorders
Marjan Bozinoski, M.S., and Christopher E. Mason, Ph.D.

1.12 Psychoneuroendocrinology
Debra S. Harris, M.D., Owen M. Wolkowitz, M.D., and Victor I. Reus, M.D.

1.13 Immune System and Central Nervous System Interactions
Charles L. Raison, M.D., Rebecca G. Reed, Ph.D., Graham A.W. Rook, B.Chir., M.D., and Andrew H. Miller, M.D.

1.14 Chronobiology
Megan A. Miller, M.S., Kathryn A. Roecklein, Ph.D., and Ignacio Provencio, Ph.D.

1.15 Applied Electrophysiology
Jason K. Johannesen, Ph.D., and Patrick D. Skosnik, Ph.D.

1.16 Nuclear Magnetic Resonance Imaging and Spectroscopy: Basic Principles and Recent Findings in Neuropsychiatric Disorders
Graeme F. Mason, Ph.D., John H. Krystal, M.D., and Gerard Sanacora, M.D., Ph.D.

1.17 Radiotracer Imaging with Positron Emission Tomography and Single Photon Emission Computed Tomography: Fundamental Principles, Methodology, and Role in Neuropsychiatric Research
Mark Slifstein, Ph.D., John H. Krystal, M.D., and Anissa Abi-Dargham, M.D.

1.18 Gene Mapping Investigations of Psychiatric Disorders
Scott C. Fears, M.D., Ph.D.

1.19 Animal Models in Psychiatry
Sietse Jonkman, Ph.D., and Paul J. Kenny, Ph.D.

1.20 Pain Systems: Interface with Affective and Motivational Mechanisms
Jon-Kar Zubieta, M.D., Ph.D.

1.21 Basic Science of Self
Debra A. Gusnard, M.D.

1.22 Basic Science of Sleep
Ruth M. Benca, M.D., Ph.D., Chiara Cirelli, M.D., Ph.D., and Giulio Tononi, M.D., Ph.D.

1.23 Basic Science of Appetite
Nori Geary, Ph.D., and Timothy H. Moran, Ph.D.

1.24 Neuroscience of Substance Use Disorders
Elena H. Chartoff, Ph.D., William A. Carlezon, Jr., Ph.D., and Hilary S. Connery, M.D., Ph.D.

1.25 **Epigenetics in Psychiatry: The Promise for New Biomarkers and Treatments**
Jennifer B. Dwyer, M.D., Ph.D., and Arie Kaffman, M.D., Ph.D.

1.26 **Basic Systems Neuroscience**
André Antonio Fenton, Ph.D., and Dino Dvorak, Ph.D.

1.27 **Learning Theory and Psychosis, Anxiety, and Addiction**
Philip R. Corlett, Ph.D.

1.28 **Functional Brain Connectivity and Psychopathology**
Deanna M. Barch, Ph.D.

1.29 **Computational Modeling Approaches to Psychiatry**
John D. Murray, Ph.D., and Alan Anticevic, Ph.D.

1.30 **Functional Genomics of Human Brain Development**
Sirisha Pochareddy, Ph.D., John C. Silbereis, Ph.D., and Nenad Sestan, M.D., Ph.D.

2 **NEUROPSYCHIATRY AND BEHAVIORAL NEUROLOGY**

2.1 **The Neuropsychiatric Approach to the Patient**
Fred Ovsiew, M.D.

2.2 **Neuropsychiatric Aspects of Cerebrovascular Disorders**
Sarah T. McCutcheon, M.D., and Robert G. Robinson, M.D.

2.3 **The Neuropsychiatry of Brain Tumors**
Alasdair G. Rooney, M.D.

2.4 **Neuropsychiatric Aspects of Epilepsy**
Pongsatorn Paholpak, M.D., Joanna Juan Gan, M.D., and Mario F. Mendez, M.D., Ph.D.

2.5 **Neuropsychiatric Consequences of Traumatic Brain Injury**
Joanne Alonso Byars, M.D., and Ricardo E. Jorge, M.D.

2.6 **Neuropsychiatric Aspects of Movement Disorders**
Laura Marsh, M.D., and Russell L. Margolis, M.D.

2.7 **Neuropsychiatric Aspects of Multiple Sclerosis and Other Demyelinating Disorders**
Russell T. Joffe, M.D.

2.8 **Neuropsychiatric Aspects of HIV Infection and AIDS**
Glenn Jordan Treisman, M.D., Ph.D., Jeffrey Hsu, M.D., Heidi E. Hutton, Ph.D., and Andrew F. Angelino, M.D.

2.9 **Neuropsychiatric Aspects of Other Infectious Diseases (Non-HIV)**
Brian A. Fallon, M.D., and Shannon L. Delaney, M.D.

2.10 **Neuropsychiatric Aspects of Prion Disease**
Alireza Minagar, M.D., Nadejda Alekseeva, M.D., Paul Shapshak, Ph.D., and Francisco Fernandez, M.D.

2.11 **Neuropsychiatric Aspects of Headache**

Kathleen Ries Merikangas, Ph.D., Tarannum Lateef, M.D., and James R. Merikangas, M.D.

2.12 **Neuropsychiatric Aspects of Neuromuscular Diseases**
Timothy P. Wuerz, D.O., Rajeet Shrestha, M.D., and Alan J. Lerner, M.D.

2.13 **Psychiatric Aspects of Child Neurology**
Jeffrey E. Max, M.B.B.Ch., Sarah N. Mattson, Ph.D., Yvonne Eleanor Vaucher, M.D., Sharon L. Nichols, Ph.D., and Mark P. Nespeca, M.D.

2.14 **Neuropsychiatry of Neurometabolic and Neuroendocrine Disorders**
Mark Walterfang, Ph.D., Ramon Mocellin, M.B.B.S., M.Sc., M.Psych. (Med), F.R.A.N.Z.C.P., Olivier Bonnot, M.D., Ph.D., and Dennis Velakoulis, DMedSci

3 **CONTRIBUTIONS OF THE PSYCHOLOGICAL SCIENCES**

3.1 **Contributions of the Psychological Sciences: Sensation, Perception, and Cognition**
Louis J. Cozolino, Ph.D., and Daniel J. Siegel, M.D.

3.2 **Piaget and Cognitive Development: History, Theory, Extensions and Modifications**
Stanley I. Greenspan, M.D., and John F. Curry, Ph.D.

3.3 **Learning Theory**
Mark E. Bouton, Ph.D.

3.4 **Biology of Memory**
Joel L. Voss, Ph.D., and Ken A. Paller, Ph.D.

3.5 **Consciousness and Dreaming from a Pathophysiological Perspective: The Thalamocortical Dysrhythmia Syndrome**
Rodolfo R. Llinás, M.D., Ph.D.

3.6 **Normality and Mental Health**
George E. Vaillant, M.D.

4 **CONTRIBUTIONS OF THE SOCIAL SCIENCES**

4.1 **Evolutionary Foundations for Psychiatric Research and Practice**
Randolph M. Nesse, M.D.

4.2 **Neurocentrism: Implications for Addiction and the Courtroom**
Sally L. Satel, M.D., and Scott O. Lilienfeld, Ph.D.

4.3 **Transcultural Psychiatry**
Robert Kohn, M.D., Ronald M. Wintrob, M.D., and Renato D. Alarcón, M.D.

5 **QUANTITATIVE AND EXPERIMENTAL METHODS IN PSYCHIATRY**

5.1 **Quantitative and Experimental Methods in Psychiatry: Epidemiology**
William E. Narrow, M.D., and Diana E. Clarke, Ph.D.

5.2 **Statistics and Experimental Design**
Eugene M. Laska, Ph.D., Carole E. Siegel, Ph.D., Morris Meisner, Ph.D., and Isaac R. Galatzer-Levy, Ph.D.

6 THEORIES OF PERSONALITY AND PSYCHOPATHOLOGY

6.1 Classical Psychoanalysis
William W. Meissner, S.J., M.D.

6.2 Erik H. Erikson
Dorian S. Newton, Ph.D.

6.3 Other Psychodynamic Schools
Adam M. Brenner, M.D., and Laura S. Howe-Martin, Ph.D.

6.4 Approaches Derived from Philosophy and Psychology
Paul T. Costa, Jr., Ph.D., and Robert R. McCrae, Ph.D.

7 DIAGNOSIS AND PSYCHIATRY: EXAMINATION OF THE PSYCHIATRIC PATIENT

7.1 Psychiatric Interview, History, and Mental Status Examination of the Adult Patient
Rory P. Houghtalen, M.D., and John S. McIntyre, M.D.

7.2 Outline for a Psychiatric Examination
Benjamin J. Sadock, M.D.

7.3 Practice Guidelines in Psychiatry
Laura J. Fochtmann, M.D., and John S. McIntyre, M.D.

7.4 Clinical Neuropsychology and Intellectual Assessment of Adults
Rex M. Swanda, Ph.D., and Kathleen Y. Haaland, Ph.D.

7.5 Personality Assessment of Adults
Russell L. Adams, Ph.D., and David L. Adams, Psy.D.

7.6 Medical Assessment and Laboratory Testing in Psychiatry
Jennifer L. Kruse, M.D., Martha L. James, M.D., J.D., and Barry H. Guze, M.D.

7.7 Principles and Applications of Quantitative Electroencephalography in Psychiatry
Leslie S. Prichep, Ph.D., and E. Roy John, Ph.D.

7.8 Psychiatric Rating Scales
Samuel Justin Sinclair, Ph.D., Mark A. Blais, Psy.D., and Deborah Blacker, M.D., Sc.D.

7.9 Electronic Media in Psychiatry
Zebulon Taintor, M.D.

7.10 Cultural Competency in the Organization and Delivery of Care
Carole E. Siegel, Ph.D., Gary Haugland, M.A., and Helena B. Hansen, M.D., Ph.D.

7.11 Medical Error
Nora C. Oberfield, M.D., and Benjamin J. Sadock, M.D.

8 CLINICAL MANIFESTATIONS OF PSYCHIATRIC DISORDERS
Anu A. Matorin, M.D., Asim A. Shah, M.D., and Pedro Ruiz, M.D.

9 CLASSIFICATION IN PSYCHIATRY

9.1 Present and Future of Classification Systems for Mental Disorders
Javier I. Escobar, M.D., and Humberto Marin, M.D.

9.2 The Classification of Mental Disorders in the International Classification of Diseases (ICD-11)
Javier I. Escobar, M.D., and Carlos López-Jaramillo, M.D., Ph.D.

10 NEUROCOGNITIVE DISORDERS

10.1 Cognitive Disorders: Introduction
Robert A. Sweet, M.D.

10.2 Delirium
Tanya J. Fabian, Pharm.D., Ph.D., and LalithKumar K. Solai, M.D.

10.3 Dementia
Julie A. Graziane, M.D., and Robert A. Sweet, M.D.

10.4 Amnestic Disorders and Mild Cognitive Impairment
Carmen Andreescu, M.D., and Daniel Varon, M.D.

10.5 Other Cognitive and Mental Disorders due to Another Medical Condition
Jennifer Clare Prins, M.D., and Ellen M. Whyte, M.D.

11 SUBSTANCE-RELATED DISORDERS

11.1 Substance Use Disorders: Introduction
Eric Cameron Strain, M.D.

11.2 Alcohol-Related Disorders
Marc A. Schuckit, M.D.

11.3 Stimulant-Related Disorders
Rocco A. Iannucci, M.D., and Roger D. Weiss, M.D.

11.4 Caffeine-Related Disorders
Laura M. Juliano, Ph.D., and Roland R. Griffiths, Ph.D.

11.5 Cannabis-Related Disorders
Wayne D. Hall, Ph.D., and Louisa Degenhardt, Ph.D.

11.6 Hallucinogen-Related Disorders
Michael P. Bogenschutz, M.D., and Stephen Ross, M.D.

11.7 Inhalant-Related Disorders
Matthew O. Howard, Ph.D., Scott E. Bowen, Ph.D., and Eric L. Garland, Ph.D.

11.8 Tobacco-Related Disorders
Douglas M. Ziedonis, M.D., Makenzie E. Tonelli, B.A., and Smita Das, M.D., Ph.d.

11.9 Opioid Use and Related Disorders: From Neuroscience to Treatment
Sean X. Luo, M.D., Ph.D., and Adam Bisaga, M.D.

11.10 Sedative-, Hypnotic-, or Anxiolytic-Related Disorders

John J. Mariani, M.D.

11.11 Anabolic-Androgenic Steroid-Related Disorders

Harrison G. Pope, Jr., M.D., and Gen Kanayama, M.D., Ph.D.

12 SCHIZOPHRENIA AND OTHER PSYCHOTIC DISORDERS

12.1 Introduction and Overview

Carol A. Tamminga, M.D.

12.2 Phenomenology of Schizophrenia

Stephen F. Lewis, M.D., Rodrigo Escalona, M.D., and Samuel J. Keith, M.D.

12.3 Worldwide Burden of Schizophrenia

Assen Jablensky, M.D.

12.4 Genetics of Schizophrenia

Kimberley M. Kendall, B.Sc. (Hons.), M.B.B.Ch., M.R.C.Psych. George Kirov, M.D., Ph.D., and Michael J. Owen, M.D., Ph.D.

12.5 The Clinical Epidemiology of Schizophrenia

Jim van Os, Ph.D., and Ulrich Reininghaus, Ph.D.

12.6 Cellular and Molecular Neuropathology of Schizophrenia

Jessica Marie Perez, M.S., and Subroto Ghose, M.D., Ph.D.

12.7 Structural Brain Imaging in Schizophrenia

Amanda E. Lyall, Ph.D., Marek Kubicki, M.D., Ph.D., and Martha E. Shenton, Ph.D.

12.8 Functional Brain Imaging in Schizophrenia

Raquel E. Gur, M.D., Ph.D., and Ruben C. Gur, Ph.D.

12.9 Molecular Brain Imaging in Schizophrenia

Dean F. Wong, M.D., Ph.D., James Robert Brašic´, M.D., Andrew Horti, Ph.D., and Gerhard Gründer, M.D.

12.10 Neurocognition in Schizophrenia

Philip D. Harvey, Ph.D., Richard S.E. Keefe, Ph.D., and Charles E. Eesley, Ph.D.

12.11 Phenotypes of Psychosis

Brett A. Clementz, Ph.D., Matcheri S. Keshavan, M.D., Godfrey D. Pearlson, M.D., John A. Sweeney, Ph.D., and Carol A. Tamminga, M.D.

12.12 Schizophrenia: Pharmacological Treatment

John M. Kane, M.D., and Christoph U. Correll, M.D.

12.13 Psychiatric Rehabilitation

Robert E. Drake, M.D., Ph.D., Gary R. Bond, Ph.D., and Dror Ben-Zeev, Ph.D.

12.14 Medical Health in Schizophrenia

John W. Newcomer, M.D., Martin T. Strassnig, M.D., Peter A. Fahnestock, M.D., and Michele L. Pergadia, Ph.D.

12.15 Recovery in Schizophrenia

Dina Hooshyar, M.D., Mark Ragins, M.D., and Frederick J. Frese, Ph.D.

12.16 Psychosis as a Defining Dimension in Schizophrenia
Elena I. Ivleva, M.D., Ph.D., and Carol A. Tamminga, M.D.

12.17 Other Psychotic Disorders
Ramin Mojtabai, M.D., Ph.D., Laura J. Fochtmann, M.D., and Evelyn J. Bromet, Ph.D.

13 MOOD DISORDERS

13.1 Mood Disorders: Historical Introduction and Conceptual Overview
Hagop S. Akiskal, M.D.

13.2 Mood Disorders: Epidemiology
Kathleen Ries Merikangas, Ph.D., and Zoltán Rihmer, M.D., Ph.D., D.Sc.

13.3 Mood Disorders: Genetics
John R. Kelsoe, M.D., and Tiffany A. Greenwood, Ph.D.

13.4 Mood Disorders: Clinical Features
Hagop S. Akiskal, M.D.

13.5 Mood Disorders: Intrapsychic and Interpersonal Aspects
John C. Markowitz, M.D., and Barbara L. Milrod, M.D.

13.6 Mood Disorders: Suicidal Behavior
Zoltán Rihmer, M.D., Ph.D., D.Sc., and Maurizio Pompili, M.D., Ph.D.

13.7 Mood Disorders: Pharmacological Treatment of Depression and Bipolar Disorders
Lakshmi N. Yatham, M.B.B.S., M.R.C.Psych. (UK), and Sidney H. Kennedy, M.D.

13.8 Mood Disorders: Psychotherapy
John R. McQuaid, Ph.D., and Michael E. Thase, M.D.

13.9 Mood Disorders: Neurobiology
Michael E. Thase, M.D.

14 ANXIETY DISORDERS

14.1 Anxiety Disorders: Introduction and Overview
Kathleen Ries Merikangas, Ph.D.

14.2 Clinical Features of the Anxiety Disorders
Ryan J. Kimmel, M.D., and Peter Roy-Byrne, M.D.

14.3 Epidemiology of Anxiety Disorders
Kathleen Ries Merikangas, Ph.D., and John David Eun, B.A.

14.4 Genetics of Anxiety Disorders
John M. Hettema, M.D., Ph.D., and Takeshi Otowa, M.D., Ph.D.

14.5 Anxiety Disorders: Neurobiology and Neuroscience
Adriana Feder, M.D., Sara Costi, M.D., Brian M. Iacoviello, Ph.D., James W. Murrough, M.D., and Dennis S. Charney, M.D.

14.6 Neuroimaging and the Neuroanatomical Circuits Implicated in Anxiety, Fear, and Stress-Related Disorders
Dylan G. Gee, Ph.D., and B.J. Casey, Ph.D.

14.7 Anxiety Disorders: Cognitive-Behavioral Therapy
Jonathan D. Huppert, Ph.D., and Edna B. Foa, Ph.D.

14.8 Somatic Therapies for Anxiety Disorders
Lakshmi Ravindran, M.D., and Murray B. Stein, M.D.

15 OBSESSIVE-COMPULSIVE AND RELATED DISORDERS
Dan J. Stein, M.D., Ph.D., and Christine Lochner, Ph.D.

16 PATHOLOGICAL AND PROBLEM GAMBLING: GAMBLING DISORDER
Harvey Roy Greenberg, M.D.

17 POSTTRAUMATIC STRESS DISORDER
Arieh Y. Shalev, M.D., and Charles R. Marmar, M.D.

18 SOMATIC SYMPTOM AND RELATED DISORDERS
Javier I. Escobar, M.D., and Joel E. Dimsdale, M.D.

19 FACTITIOUS DISORDER
Davin K. Quinn, M.D., Dora Wang, M.D., Seth Powsner, M.D., and Stuart J. Eisendrath, M.D.

20 DISSOCIATIVE DISORDERS
Richard J. Loewenstein, M.D., Paul Frewen, Ph.D., and Roberto Lewis-Fernández, M.D.

21 NORMAL SEXUALITY AND SEXUAL DISORDERS

21.1 Normal Human Sexuality and Sexual Dysfunctions
Virginia A. Sadock, M.D.

21.2 Homosexuality, Gay and Lesbian Identities, and Homosexual Behavior
Jack Drescher, M.D., and William M. Byne, M.D., Ph.D.

21.3 Paraphilic Disorders
Reneé M. Sorrentino, M.D.

21.4 Gender Identity, Gender Variance, and Gender Dysphoria
Jack Drescher, M.D., and William M. Byne, M.D., Ph.D.

21.5 Sexual Addiction
Aviel Goodman, M.D.

22 FEEDING AND EATING DISORDERS
Christine C. Call, A.B., Evelyn Attia, M.D., and B. Timothy Walsh, M.D.

23 SLEEP DISORDERS

Max Hirshkowitz, Ph.D., and Amir Sharafkhaneh, M.D., Ph.D.

24 IMPULSE-CONTROL DISORDERS

24.1 Intermittent Explosive Disorder

Divya Ramesh, Ph.D., Sameer Hassamal, M.D., and Frederick Gerard Moeller, M.D.

24.2 Pyromania

Sameer Hassamal, M.D., Divya Ramesh, Ph.D., and Frederick Gerard Moeller, M.D.

24.3 Kleptomania

Sameer Hassamal, M.D., Divya Ramesh, Ph.D., and Frederick Gerard Moeller, M.D.

25 ADJUSTMENT DISORDERS

Jeffrey W. Katzman, M.D., and Cynthia M. A. Geppert, M.D., Ph.D.

26 PERSONALITY DISORDERS

C. Robert Cloninger, M.D., Ph.D., and Dragan M. Svrakic, M.D., Ph.D.

Volume II

27 PSYCHOSOMATIC MEDICINE

27.1 History and Current Trends

Carol L. Alter, M.D., and Steven A. Epstein, M.D.

27.2 Cardiovascular Disorders

Peter A. Shapiro, M.D., and Adam R. Critchfield, M.D.

27.3 Gastrointestinal Disorders

Claude Andrew Botha, Ph.D., James K. Ruffle, B.Sc., M.B.B.S., and Adam D. Farmer, Ph.D.

27.4 Psychosomatic Medicine: Obesity

Kathleen R. Ashton, Ph.D., and Megan E. Lavery, Psy.D.

27.5 Respiratory Disorders

Thomas Ritz, Ph.D., Chelsey A. Werchan, M.A., and E. Sherwood Brown, M.D., Ph.D.

27.6 Diabetes: Psychosocial Issues and Psychiatric Disorders

Wayne Katon, M.D., and Paul Ciechanowski, M.D.

27.7 Endocrine and Metabolic Disorders

Katherine E. Williams, M.D., and Natalie L. Rasgon, M.D., Ph.D.

27.8 Psycho-Oncology

Sean P. Heffernan, M.D., Yesne Alici, M.D., William S. Breitbart, M.D., and

Marguerite S. Lederberg, M.D.

27.9 End-of-Life and Palliative Care

Marguerite S. Lederberg, M.D., Xiomara Rocha-Cadman, M.D., and Mindy Greenstein, Ph.D.

27.10 Stress and Psychiatry

Paul S. Nestadt, M.D., Traci J. Speed, M.D., Ph.D., Francis J. Keefe, Ph.D., and Joel E. Dimsdale, M.D.

27.11 Psychiatric Morbidity Following Critical Illness

O. Joseph Bienvenu, M.D., Ph.D., and Dimitry S. Davydow, M.D.

27.12 Psychocutaneous Disorders

Adarsh K. Gupta, M.D., and Mohammad Jafferany, M.D.

27.13 Organ Transplantation

Andrea Frances DiMartini, M.D., Mary Amanda Dew, Ph.D., and Catherine Chang Crone, M.D.

27.14 Psychiatric Care of the Burned Patient

Martha C. Gamboa, M.D., Julianne K. Suojanen, D.O., Ejine Okoroafor, M.D., and Michael Blumenfield, M.D.

27.15 Management of Chronic Pain

J. Gregory Hobelmann, M.D., and Michael R. Clark, M.D.

28 ADDITIONAL CONDITIONS THAT MAY BE A FOCUS OF CLINICAL ATTENTION

28.1 Malingering

Frank John Ninivaggi, M.D.

28.2 Adult Antisocial Behavior, Criminality, and Violence

Dorothy Otnow Lewis, M.D.

28.3 Borderline Intellectual Functioning and Academic or Educational Problem

Frank John Ninivaggi, M.D.

29 CULTURAL CONCEPTS OF DISTRESS AND ASSESSMENT

29.1 Cultural Concepts of Distress

Roberto Lewis-Fernández, M.D., Laurence J. Kirmayer, M.D., Peter J. Guarnaccia, Ph.D., and Pedro Ruiz, M.D.

29.2 Cultural Assessments in Psychiatry

Roberto Lewis-Fernández, M.D., and Neil Krishan Aggarwal, M.D.

30 THE NEUROPSYCHIATRY OF HUMAN AGGRESSION

Jeff Victoroff, M.D.

31 SPECIAL AREAS OF INTEREST

31.1 Psychiatry and Reproductive Medicine

Sarah L. Berga, M.D., Barbara L. Parry, M.D., and Eydie L. Moses-Kolko, M.D.

31.2 Genetic Counseling for Psychiatric Conditions
Holly L. Peay, Ph.D., Barbara Bowles Biesecker, Ph.D., and Jehannine C. Austin, Ph.D.

31.3 Physical and Sexual Abuse of Adults
Shannon F. Stromberg, M.D.

31.4 Complementary, Alternative, and Integrative Approaches in Mental Health Care
James H. Lake, M.D.

31.5 Disaster Psychiatry: Disasters, Terrorism, and War
David M. Benedek, M.D., Joshua Chaim Morganstein, M.D., Harry C. Holloway, M.D., and Robert J. Ursano, M.D.

31.6 Psychiatry and Spirituality
Armando R. Favazza, M.D.

31.7 Physician and Medical Student Mental Health
Khleber Chapman Attwell, M.D., and Graciana Lapetina, M.D.

32 PSYCHIATRIC EMERGENCIES

32.1 Suicide: Overview and Epidemiology
Danuta Wasserman, M.D.

32.2 Suicide Treatment
Howard S. Sudak, M.D.

32.3 Other Psychiatric Emergencies
David A. Baron, D.O., Robert T. Cobb, M.D., and Grace Mitu Juarez, M.D.

33 PSYCHOTHERAPIES

33.1 Psychoanalysis and Psychoanalytic Psychotherapy
Sylvia R. Karasu, M.D.

33.2 Psychoanalytic Treatment of Anxiety Disorders, Obsessive-Compulsive and Trauma-Related Disorders
Eric M. Plakun, M.D.

33.3 Behavior Therapy
Derek R. Hopko, Ph.D., Caroline Gray Clark, B.A., M.S.S.W., and Rebecca Shorter, M.A.

33.4 Hypnosis
Allan David Axelrad, M.D., Daniel P. Brown, Ph.D., and Harold J. Wain, Ph.D.

33.5 Group Psychotherapy
Henry I. Spitz, M.D.

33.6 Family and Couples Therapy
Henry I. Spitz, M.D., and Susan T. Spitz, A.C.S.W.

33.7 Cognitive Therapy
Judith S. Beck, Ph.D., and Robert Hindman, Ph.D.

33.8 Interpersonal Psychotherapy
Robert W. Guynn, M.D.

33.9 Dialectical Behavior Therapy
M. Zachary Rosenthal, Ph.D., and Marcus A. Rodriguez, M.A., M.S.

33.10 Intensive Short-Term Psychodynamic Psychotherapies
Manuel Trujillo, M.D.

33.11 Other Methods of Psychotherapy
Adam M. Brenner, M.D., and E. Will Clark, M.D.

33.12 Narrative Psychiatry
Bradley Lewis, M.D., Ph.D.

33.13 Positive Psychology
Nansook Park, Ph.D.

33.14 Psychodrama
Edward J. Schreiber, T.E.P., D.D.

33.15 Evaluation of Psychotherapy
Megan L. Petrik, Ph.D., and Stephen M. Saunders, Ph.D.

33.16 Individual Psychodynamic Psychotherapy and Cognitive Behavioral Therapy for Psychosis
Michael D. Garrett, M.D.

33.17 Hallucinogen-Assisted Psychotherapy
Stephen Ross, M.D., and Michael P. Bogenschutz, M.D.

33.18 Mentalization-Based Treatment
Anthony W. Bateman, F.R.C.Psych., and Peter Fonagy, Ph.D.

34 BIOLOGICAL THERAPIES

34.1 General Principles of Psychopharmacology
Olusola Ajilore, M.D., Ph.D., and Mark M. Rasenick, Ph.D.

34.2 Drug Development and Approval Process in the United States
Amir Kalali, M.D., David Akers Hosford, M.D., Ph.D., and Sarah J. Richerson, Ph.D.

34.3 Medication-Induced Movement Disorders
Philip G. Janicak, M.D., and Kumail Hussain, M.D.

34.4 α2-Adrenergic Receptor Agonists: Clonidine, Clonidine Extended Release, Guanfacine, and Guanfacine Extended Release
Eric Hollander, M.D., and Jennifer N. Petras, M.D.

34.5 β-Adrenergic Receptor Antagonists
Roger S. McIntyre, M.D.

34.6 Anticholinergics and Amantadine
Samoon Ahmad, M.D.

34.7 Anticonvulsants: Gabapentin, Topiramate, Tiagabine, Levetiracetam, Zonisamide, and Pregabalin
Po W. Wang, M.D., and Terence Arthur Ketter, M.D.

34.8 Antihistamines
Roger S. McIntyre, M.D.

34.9 Barbiturates and Similarly Acting Substances
Steven L. Dubovsky, M.D.

34.10 Benzodiazepine Receptor Agonists and Antagonists
Steven L. Dubovsky, M.D.

34.11 Bupropion
Charles DeBattista, M.D., and Alan F. Schatzberg, M.D.

34.12 Buspirone
Molly M. McCarthy, D.O., and Regina M. Carney, M.D.

34.13 Calcium Channel Inhibitors
Steven L. Dubovsky, M.D.

34.14 Carbamazepine
Robert M. Post, M.D., and Mark A. Frye, M.D.

34.15 Cholinesterase Inhibitors and Memantine
Elizabeth A. Crocco, M.D., and Samir A. Sabbag, M.D.

34.16 Disulfiram and Acamprosate
Allison K. Ungar, M.D., Katherine S. Pier, M.D., and Iliyan Ivanov, M.D.

34.17 First-Generation Antipsychotics
Martin T. Strassnig, M.D., and Philip D. Harvey, Ph.D.

34.18 Lamotrigine
Po W. Wang, M.D., and Terence Arthur Ketter, M.D.

34.19 Lithium
Charles L. Bowden, M.D.

34.20 Melatonin Receptor Agonists
Samantha B. Saltz, M.D., Karen M. Padron, M.D., Ph.D., Brett Lloyd, Jr., M.D., Ph.D., and Jeffrey Newport, M.D.

34.21 Mirtazapine
Radu V. Saveanu, M.D., Mousa Sobhy Azer Botros, M.D., and Michael E. Thase, M.D.

34.22 Monoamine Oxidase Inhibitors
Ethan F. Karsen, M.D., and Paul E. Holtzheimer, M.D.

34.23 Nefazodone
Amir A. Khan, M.D., and Susan G. Kornstein, M.D.

34.24 Opioid Maintenance Treatments: Methadone and Buprenorphine
Andrew J. Saxon, M.D., Aimee L. McRae-Clark, Pharm.D., and Kathleen T.

Brady, M.D., Ph.D.

34.25 Opioid Receptor Antagonists: Naltrexone and Nalmefene
Srinivas B. Muvvala, M.D., Suchitra Krishnan-Sarin, Ph.D., and Stephanie S. O'Malley, Ph.D.

34.26 Selective Serotonin-Norepinephrine Reuptake Inhibitors
Michael E. Thase, M.D., and Pierre Blier, M.D., Ph.D.

34.27 Selective Serotonin Reuptake Inhibitors
Norman Sussman, M.D.

34.28 Second-Generation Antipsychotics
Stephen R. Marder, M.D., and Michael C. Davis, M.D., Ph.D.

34.29 Central Nervous System Sympathomimetic Compounds
Manpreet Kaur Singh, M.D., and Yuriy Dobry, M.D.

34.30 Dopamine Receptor Agonists
Carlos A. Zarate, Jr., M.D., and Lawrence T. Park, M.D.

34.31 Thyroid Hormones
Russell T. Joffe, M.D.

34.32 Trazodone
John T. Vernon, M.D., and Susan G. Kornstein, M.D.

34.33 Tricyclics and Tetracyclics
J. Craig Nelson, M.D.

34.34 Valproate
Robert M. Post, M.D., and Mark A. Frye, M.D.

34.35 Electroconvulsive Therapy
Joan Prudic, M.D., and Yuanli Duan, M.D., Ph.D.

34.36 Brain Stimulation Methods
Stefan B. Rowny, M.D., and Melanie Lucas, Ph.D.

34.37 Neurosurgical Treatments
Benjamin D. Greenberg, M.D., Ph.D.

34.38 Combination Pharmacotherapy
Charles DeBattista, M.D., and Alan F. Schatzberg, M.D.

34.39 Reproductive Hormonal Therapy: Theory and Practice
David R. Rubinow, M.D., and Peter J. Schmidt, M.D.

34.40 Drug Treatment of Erectile Disorder
Caridad Ponce Martinez, M.D., Fadi Haddad, M.D., and Anita H. Clayton, M.D.

34.41 Vilazodone
Lourdes Illa-Sanchez, M.D.

35 CHILD PSYCHIATRY

35.1 Introduction and Overview

Caroly Sue Pataki, M.D.

35.2 Normal Child Development

Gabrielle Shapiro, M.D., Robert J. Jaffe, M.D., and Alexander Kolevzon, M.D.

35.3 Adolescent Development

Ardis C. Martin, M.D., and Caroly Sue Pataki, M.D.

36 CHILD PSYCHIATRIC EXAMINATION

36.1 Psychiatric Examination of the Infant, Child, and Adolescent

Robert A. King, M.D., Prudence W. Fisher, Ph.D., Michael H. Bloch, M.D., Armin Paul Thies, Ph.D., and Mary E. Schwab-Stone, M.D.

36.2 Psychiatric Assessment of Preschool Children

Helen Link Egger, M.D., and Caroly Sue Pataki, M.D.

36.3 Neuropsychological and Cognitive Assessment of Children

Martha Bates Jura, Ph.D., and Lorie A. Humphrey, Ph.D.

36.4 Psychiatric Emergencies in Children and Adolescents

Ruth S. Gerson, M.D., and Fadi Haddad, M.D.

37 GENETICS IN CHILD PSYCHIATRY

Erika L. Nurmi, M.D., Ph.D., Meredith A. Hannan, M.D., Ph.D., and James T. McCracken, M.D.

38 NEUROIMAGING IN PSYCHIATRIC DISORDERS OF CHILDHOOD

Frank P. MacMaster, Ph.D., Mariko Sembo, B.A., and David R. Rosenberg, M.D.

39 TEMPERAMENT: RISK AND PROTECTIVE FACTORS FOR CHILD PSYCHIATRIC DISORDERS

David C. Rettew, M.D.

40 INTELLECTUAL DISABILITY

Bryan H. King, M.D., Karen E. Toth, Ph.D., Nina De Lacy, M.D., and Dan Doherty, M.D., Ph.D.

41 SPECIFIC LEARNING DISORDER

Rosemary Tannock, Ph.D.

42 DEVELOPMENTAL COORDINATION DISORDER

Caroly Sue Pataki, M.D., and Wendy G. Mitchell, M.D.

43 COMMUNICATION DISORDERS

43.1 Language Disorder

Emiko Koyama, Ph.D., and Joseph H. Beitchman, M.D.

43.2 Social (Pragmatic) Communication Disorder

Emiko Koyama, Ph.D., and Joseph H. Beitchman, M.D.

43.3 Speech Sound Disorder
Emiko Koyama, Ph.D., and Joseph H. Beitchman, M.D.

43.4 Stuttering
Robert M. Kroll, Ph.D., and Joseph H. Beitchman, M.D.

43.5 Unspecified Communication Disorder
Tim Bressmann, Ph.D., and Joseph H. Beitchman, M.D.

44 AUTISM SPECTRUM DISORDER AND SOCIAL COMMUNICATION DISORDER
Fred Robert Volkmar, M.D., Ami Klin, Ph.D., Robert T. Schultz, Ph.D., and Matthew W. State, M.D., Ph.D.

45 ATTENTION-DEFICIT DISORDERS

45.1 Attention-Deficit/Hyperactivity Disorder
Alma M. Spaniardi, M.D., Laurence L. Greenhill, M.D., and Lily I. Hechtman, M.D.

45.2 Adult Manifestations of Attention-Deficit/Hyperactivity Disorder
James J. McGough, M.D., M.S.

46 DISRUPTIVE BEHAVIOR DISORDERS IN CHILDREN AND ADOLESCENTS
Daniel F. Connor, M.D.

47 FEEDING AND EATING DISORDERS OF INFANCY AND EARLY CHILDHOOD
Irene Chatoor, M.D., and Loredana Lucarelli, Psy.D.

48 TIC DISORDERS
Rahil Jummani, M.D., and Barbara J. Coffey, M.D.

49 ELIMINATION DISORDERS
Edwin Jens Mikkelsen, M.D.

50 OTHER DISORDERS OF INFANCY, CHILDHOOD, AND ADOLESCENCE

50.1 Reactive Attachment Disorder and Disinhibited Social Engagement Disorder
Margaret J. Hauck, Ph.D., Mary Margaret Gleason, M.D., and Charles H. Zeanah, M.D.

50.2 Stereotypic Movement Disorders in Children
Robert Lloyd Doyle, M.D., D.D.S.

51 MOOD DISORDERS IN CHILDREN AND ADOLESCENTS

51.1 Depressive Disorders and Suicide
Karen Dineen Wagner, M.D., Ph.D., and David A. Brent, M.D.

51.2 Early-Onset Bipolar Disorder
Gabrielle A. Carlson, M.D., Stephanie E. Meyer, Ph.D., and Caroly Sue Pataki,

M.D.

52 ANXIETY DISORDERS IN CHILDREN

52.1 Obsessive-Compulsive Disorder in Childhood
Adam B. Lewin, Ph.D., and John Piacentini, Ph.D.

52.2 Posttraumatic Stress Disorder in Children and Adolescents
Judith A. Cohen, M.D.

52.3 Separation Anxiety, Generalized Anxiety, and Social Anxiety
Courtney P. Keeton, Ph.D., Jessica Lee Schleider, M.A., and John T. Walkup, M.D.

52.4 Selective Mutism
R. Lindsey Bergman, Ph.D., and Araceli Gonzalez, Ph.D.

53 EARLY-ONSET PSYCHOTIC DISORDERS
Linmarie Sikich, M.D., and Tara Chandrasekhar, M.D.

54 CHILD PSYCHIATRY: PSYCHIATRIC TREATMENT

54.1 Individual Psychodynamic Psychotherapy (Child Psychiatry)
David L. Kaye, M.D., and Zhanna Elberg, M.D.

54.2 Brief Psychotherapies for Childhood and Adolescence
Anthony L. Rostain, M.D., and Martin Franklin, Ph.D.

54.3 Cognitive-Behavioral Psychotherapy for Children and Adolescents
Anne Marie Albano, Ph.D., and Sandra S. Pimentel, Ph.D.

54.4 Group Psychotherapy
Margo L. Thienemann, M.D., Sonal Rana, M.D., and Nicole Quiterio, M.D.

54.5 Family Therapy
John Sargent, M.D., and Neha Sharma, D.O.

54.6 Pediatric Psychopharmacology
Timothy E. Wilens, M.D., and Christopher J. Kratochvil, M.D.

54.7 Inpatient Psychiatric, Partial Hospital, and Residential Treatment for Children and Adolescents
Dana Kober, M.D., and Andrés Martin, M.D.

54.8 Community-Based Treatment
Andrés J. Pumariega, M.D.

54.9 The Treatment of Adolescents
Steven C. Schlozman, M.D., and Eugene V. Beresin, M.D.

55 CHILD PSYCHIATRY: SPECIAL AREAS OF INTEREST

55.1 Adoption and Foster Care
Sandra B. Sexson, M.D., and Dale Aaron Peeples, M.D.

55.2 Child Maltreatment
William Bernet, M.D., and Bradley W. Freeman, M.D.

55.3 Children's Reaction to Illness and Hospitalization
Susan Beckwitt Turkel, M.D., Julienne Jacobson, M.D., and Maryland Pao, M.D.

55.4 Psychiatric Sequelae of HIV and AIDS
Mark DeAntonio, M.D.

55.5 Adolescent Substance Use Disorders
Sean R. LeNoue, M.D., and Paula D. Riggs, M.D.

55.6 Forensic Child and Adolescent Psychiatry
Peter Ash, M.D., and Elissa P. Benedek, M.D.

55.7 Ethical Issues in Child and Adolescent Psychiatry
Adrian Natan Sondheimer, M.D., and Arden D. Dingle, M.D.

55.8 School Consultation
Jeff Q. Bostic, M.D., Ed.D., Sharon Hoover Stephan, Ph.D., and Alexa Bagnell, M.D.

55.9 Prevention of Psychiatric Disorders in Children and Adolescents
William R. Beardslee, M.D., David A. Mrazek, M.D., and Patricia J. Mrazek, Ph.D.

55.10 Child Mental Health Services Research
Bonnie T. Zima, M.D., and Regina Bussing, M.D.

55.11 Impact of Terrorism on Children
Wanda P. Fremont, M.D., and James S. Demer, M.D.

55.12 Impact on Parents of Raising a Child with Psychiatric Illness and/or Developmental Disability
Alice R. Mao, M.D., Diane E. Treadwell-Deering, M.D., Matthew Neil Brams, M.D., and Pieter Joost van Wattum, M.D.

55.13 Pediatric Sleep Disorders
Jess P. Shatkin, M.D., and Anna Ivanenko, M.D., Ph.D.

56 ADULTHOOD
Calvin A. Colarusso, M.D.

57 GERIATRIC PSYCHIATRY

57.1 Overview

57.1a Introduction to Geriatric Psychiatry
Dilip V. Jeste, M.D.

57.1b Epidemiology of Psychiatric Disorders
Celia F. Hybels, Ph.D., and Dan German Blazer, II, M.D., Ph.D.

57.2 Assessment

57.2a Psychiatric Assessment of the Older Patient

Davangere P. Devanand, M.D., and Gregory H. Pelton, M.D.

57.2b Complementary, Alternative, and Integrative Medicine in Geriatric Psychiatry
Harris A. Eyre, M.B.B.S. (Hons.), Bernhard T. Baune, M.D., Ph.D., and Helen Lavretsky, M.D.

57.2c The Aging Brain
Douglas R. Galasko, M.D.

57.2d Psychological Changes with Normal Aging
Jennifer J. Dunkin, Ph.D., and Kimberly Ann Corp, M.S., M.F.T.

57.2e Neuropsychological Evaluation
Barton W. Palmer, Ph.D., and Gauri N. Savla, Ph.D.

57.2f Neuroimaging
Lisa T. Eyler, Ph.D., and Zvinka Z. Zlatar, Ph.D.

57.2g Genetics of Late-Life Neurodegenerative Disorders
Seethalakshmi Ramanathan, M.B.B.S., D.P.M., and Stephen J. Glatt, Ph.D.

57.3 Psychiatric Disorders of Late Life

57.3a Assessment of Functioning
David J. Moore, Ph.D., and Pariya L. Fazeli, Ph.D.

57.3b Psychiatric Problems in the Medically Ill Geriatric Patient
Anna Borisovskaya, M.D., Beth Ann LaBardi, M.D., Matthew Alan Schreiber, M.D., Ph.D., and Marcella Pascualy, M.D.

57.3c Sleep–Wake Disorders in Older Adults
Jeanne E. Maglione, M.D., Ph.D., Jana R. Cooke, M.D., and Sonia Ancoli-Israel, Ph.D.

57.3d Anxiety Disorders
Carolyn I. Rodriguez, M.D., Ph.D., and Murray B. Stein, M.D.

57.3e Geriatric Mood Disorders
Jennifer Rose Gatchel, M.D., Ph.D., James M. Wilkins, M.D., Brent P. Forester, M.D., Robert Emmett Kelly, Jr., M.D., and George S. Alexopoulos, M.D.

57.3f Alzheimer Disease and Other Neurocognitive Disorders
Gary W. Small, M.D., and Stephen T. Chen, M.D.

57.3g Delirium
Susan W. Lehmann, M.D., and Benjamin Liptzin, M.D.

57.3h Schizophrenia and Delusional Disorders
Ipsit V. Vahia, M.D., and Carl I. Cohen, M.D.

57.3i Personality Disorders
Marc E. Agronin, M.D.

57.3j Alcohol and Substance Abuse in Older Adults

Steve Koh, M.D., Michael C. Hann, M.D., and David W. Oslin, M.D.

57.4 Treatment of Psychiatric Disorders

57.4a Treatment of Psychiatric Disorders: General Principles

Bruce G. Pollock, M.D., Ph.D.

57.4b Antidepressants and Mood Stabilizers

Carl Salzman, M.D.

57.4c Antianxiety Drugs

Alastair J. Flint, M.B.

57.4d Antipsychotic Drugs

Kelly C. Lee, Pharm.D., M.A.S., B.C.P.P., and Vivian Hook, Ph.D.

57.4e Antidementia Drugs

Lon S. Schneider, M.D.

57.4f Electroconvulsive Therapy and Other Neurostimulation Treatments

Mustafa M. Husain, M.D., Shawn M. McClintock, Ph.D., and Paul E. Croarkin, D.O.

57.4g Psychosocial Factors in Psychotherapy of the Elderly

Joel I. Sadavoy, M.D., and Virginia Wesson, M.D.

57.4h Individual Psychotherapy

Joel I. Sadavoy, M.D., and Lawrence W. Lazarus, M.D.

57.4i Cognitive-Behavioral Therapy

Eric Granholm, Ph.D., John R. McQuaid, Ph.D., and Patricia A. Areán, Ph.D.

57.4j Family Intervention and Therapy with Older Adults

Carol Ann Podgorski, Ph.D., Lauren N. DeCaporale-Ryan, Ph.D., Ann Elizabeth Cornell, Psy.D., and Deborah A. King, Ph.D.

57.4k Group Therapy

Kenneth M. Schwartz, M.D., and Molyn Leszcz, M.D.

57.4l Counseling and Support Needs of Dementia Caregivers

Liat Ayalon, Ph.D., and Patricia A. Areán, Ph.D.

57.5 Financial Issues in the Delivery of Geriatric Psychiatric Care

Steve Koh, M.D., and Helen H. Kyomen, M.D.

57.6 Special Areas of Interest

57.6a Psychiatric Aspects of Long-term Care

Abhilash K. Desai, M.D., and George T. Grossberg, M.D.

57.6b Forensic Aspects

David Naimark, M.D., Keith A. Davidson, J.D., and Ansar M. Haroun, M.D.

57.6c Ethical Issues

Nathan Fairman, M.D., Scott A. Irwin, M.D., Ph.D., and Barton W.

Palmer, Ph.D.

57.6d Gender Issues

Helen H. Kyomen, M.D., and Ma'ayan E. Lieberman, M.D.

57.6e The Legal and Ethical Analysis of Elder Abuse

Paul Richard Greenwood, L.L.B. (Hons.), Ansar M. Haroun, M.D., and David Naimark, M.D.

57.6f Sexuality and Aging

Abesh Kumar Bhattacharjee, M.D., Ph.D., and Daniel D. Sewell, M.D.

57.6g HIV and Aging

Raeanne C. Moore, Ph.D., and María J. Marquine, Ph.D.

57.6h Technology for Seniors

Rujvi Kamat, Ph.D., Colin A. Depp, Ph.D., and Ipsit V. Vahia, M.D.

57.6i Positive Psychiatry of Aging

Dilip V. Jeste, M.D., and Barton W. Palmer, Ph.D.

58 PUBLIC PSYCHIATRY

58.1 Public and Community Psychiatry

Leighton Y. Huey, M.D., Robert F. Cole, Ph.D., and Julian D. Ford, Ph.D.

58.2 Reforming Health Care

Leighton Y. Huey, M.D., Julian D. Ford, Ph.D., and Robert F. Cole, Ph.D.

58.3 The Role of the Hospital in the Care of Persons with Mental Illness

Jeffrey L. Geller, M.D.

58.4 Mental Health Services Research

Cassis L. Henry, M.D., and Lisa B. Dixon, M.D., Ph.D.

58.5 The Psychiatric Hospitalist

Barry H. Guze, M.D., and Angel Cienfuegos, M.D.

58.6 Psychiatric Rehabilitation

Radu V. Saveanu, M.D., Rosie E. Curiel Cid, Psy.D., and David Loewenstein, Ph.D.

58.7 A Sociocultural Framework for Mental Health and Substance Abuse Service Disparities

Margarita Alegría, Ph.D., Kiara Alvarez, Ph.D., Bernice A. Pescosolido, Ph.D., and Glorisa Canino, Ph.D.

58.8 Criminalization of Persons with Serious Mental Illness

H. Richard Lamb, M.D., and Linda Edna Weinberger, Ph.D.

59 PSYCHIATRIC EDUCATION

59.1 Graduate Psychiatric Education

George Alan Keepers, M.D.

59.2 Examining Psychiatrists and Other Mental Health Professionals

James Morrison, M.D., and Rodrigo A. Muñoz, M.D.

60 ETHICS AND FORENSIC PSYCHIATRY

60.1 Clinical-Legal Issues in Psychiatry

Clarence Watson, J.D., M.D., Spencer Eth, M.D., and Gregory B. Leong, M.D.

60.2 Ethics in Psychiatry

Laura Weiss Roberts, M.D.

60.3 Correctional Psychiatry

Henry C. Weinstein, M.D., and Joseph V. Penn, M.D.

60.4 Neuroimaging, Psychiatry, and the Law

Joseph R. Simpson, M.D., Ph.D., and Jonathan D. Brodie, M.D., Ph.D.

61 HISTORY OF PSYCHIATRY

Edward Shorter, Ph.D.

62 WORLD ASPECTS OF PSYCHIATRY

Dinesh Bhugra, Ph.D., Mario Maj, M.D., Ph.D., Antonio Ventriglio, M.D., Ph.D., and Pedro Ruiz, M.D.

APPENDIX A: HISTORY OF PSYCHIATRY TABLE

APPENDIX B: GLOSSARY OF PSYCHIATRY and Psychology Terms

Index

Contributors

Anissa Abi-Dargham, M.D. Professor of Psychiatry, Vice Chair of Research, Department of Psychiatry, Stony Brook University Health Sciences Center School of Medicine, Stony Brook, New York.
1.17 Radiotracer Imaging with Positron Emission Tomography and Single Photon Emission Computed Tomography: Fundamental Principles, Methodology, and Role in Neuropsychiatric Research

David L. Adams, Psy.D. Adjunct Faculty, University of California San Diego School of Medicine, La Jolla, California; Clinical Psychologist and Counselor, Meier Clinics, Laguna Hills, California.
7.5 Personality Assessment of Adults

Russell L. Adams, Ph.D. David Ross Boyd Professor, Clinical Neuropsychologist, Department of Psychiatry and Behavioral Sciences, University of Oklahoma College of Medicine, Oklahoma City, Oklahoma.
7.5 Personality Assessment of Adults

Neil Krishan Aggarwal, M.D. Assistant Professor of Clinical Psychiatry, Columbia University College of Physicians and Surgeons; Research Psychiatrist, New York State Psychiatric Institute, New York, New York.
29.2 Cultural Assessments in Psychiatry

Marc E. Agronin, M.D. Affiliate Associate Professor of Psychiatry and Neurology, University of Miami Leonard M. Miller School of Medicine; Vice President for Behavioral Health and Clinical Research, Miami Jewish Health Systems, Miami, Florida.
57.3i Personality Disorders

Samoon Ahmad, M.D. Clinical Associate Professor, Department of Psychiatry, New York University School of Medicine; Unit Chief and Attending Psychiatrist, Department of Psychiatry, Bellevue Hospital, New York, New York.
34.6 Anticholinergics and Amantadine

Olusola Ajilore, M.D., Ph.D. Assistant Professor, Associate Director, Residency Training and Education; Co-Director, Adult/Neuroscience Research Track, Department of Psychiatry, University of Illinois College of Medicine; Chicago, Illinois.
34.1 General Principles of Psychopharmacology

Hagop S. Akiskal, M.D. Professor, Department of Psychiatry, Director, International Mood Center, University of California San Diego School of Medicine, La Jolla, California; Chief, Mood Disorders, VA San Diego Healthcare System, San Diego, California.

13.1 Mood Disorders: Historical Introduction and Conceptual Overview, 13.4 Mood Disorders: Clinical Features

Renato D. Alarcón, M.D. Emeritus Professor of Psychiatry, Mayo Clinic College of Medicine, Rochester, Minnesota; Honorio Delgado Chair, Universidad Peruana Cayetano Heredia Facultad de Medicina Alberto Hurtado, Lima, Peru.
4.3 Transcultural Psychiatry

Anne Marie Albano, Ph.D. Professor of Medical Psychology in Psychiatry, Division of Child and Adolescent Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York.
54.3 Cognitive-Behavioral Psychotherapy for Children and Adolescents

Margarita Alegría, Ph.D. Professor, Departments of Medicine and Psychiatry, Harvard Medical School; Chief, Disparities Research Unit, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts.
58.7 A Sociocultural Framework for Mental Health and Substance Abuse Service Disparities

Nadejda Alekseeva, M.D. Staff Psychiatrist, Overton Brooks VA Medical Center, Shreveport, Louisiana.
2.10 Neuropsychiatric Aspects of Prion Disease

George S. Alexopoulos, M.D. S.P. Tobin and AM Cooper Professor, Department of Psychiatry, Weill Cornell Medical College; Director, Weill Cornell Institute of Geriatric Psychiatry, New York Presbyterian Hospital, White Plains, New York.
57.3e Geriatric Mood Disorders

Yesne Alici, M.D. Assistant Professor of Psychiatry, Weill Cornell Medical College; Attending Psychiatrist, Department of Psychiatry and Behavioral Sciences, Memorial Sloan Kettering Cancer Center, New York, New York.
27.8 Psycho-Oncology

Carol L. Alter, M.D. Adjunct Associate Professor, Department of Psychiatry, Georgetown University; Chief Medical Officer, Mindoula Health, Silver Spring, Maryland.
27.1 History and Current Trends

Kiara Alvarez, Ph.D. Instructor in Medicine, Harvard Medical School; Instructor, Disparities Research Unit, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts.
58.7 A Sociocultural Framework for Mental Health and Substance Abuse Service Disparities

Sonia Ancoli-Israel, Ph.D. Professor Emeritus of Psychiatry and Medicine, Departments of Psychiatry and Medicine, University of California San Diego School of Medicine, La Jolla, California.
57.3c Sleep-Wake Disorders in Older Adults

Carmen Andreescu, M.D. Assistant Professor of Psychiatry, University of Pittsburgh School of Medicine, Western Psychiatric Institute and Clinic, Pittsburgh, Pennsylvania.
10.4 Amnestic Disorders and Mild Cognitive Impairment

Andrew F. Angelino, M.D. Associate Professor, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine; Chair of Psychiatry, Johns Hopkins Howard County General Hospital, Columbia, Maryland.

2.8 Neuropsychiatric Aspects of HIV Infection and AIDS

Alan Anticevic, Ph.D. Assistant Professor, Co-Director, Division of Neurocognition, Neurocomputation, and Neurogenetics (N3), Departments of Psychiatry and Psychology, Yale School of Medicine, New Haven, Connecticut.
1.29 Computational Modeling Approaches to Psychiatry

Patricia A. Areán, Ph.D. Professor in Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington.
57.4i Cognitive-Behavioral Therapy, 57.4l Counseling and Support Needs of Dementia Caregivers

Peter Ash, M.D. Director, Psychiatry and Law Service, Emory University School of Medicine; Professor, Department of Psychiatry and Behavioral Sciences, Atlanta, Georgia.
55.6 Forensic Child and Adolescent Psychiatry

Kathleen R. Ashton, Ph.D. Psychologist, Bariatric and Metabolic Institute, Cleveland Clinic, Cleveland, Ohio.
27.4 Psychosomatic Medicine: Obesity

Evelyn Attia, M.D. Professor of Psychiatry, Columbia University College of Physicians and Surgeons, New York State Psychiatric Institute, Weill Cornell Medical College, New York, New York.
22 Feeding and Eating Disorders

Khleber Chapman Attwell, M.D. Clinical Assistant Professor, Department of Psychiatry, New York University School of Medicine, New York, New York.
31.7 Physician and Medical Student Mental Health

Jehannine C. Austin, Ph.D. Associate Professor and Canada Research Chair, Departments of Psychiatry and Medical Genetics, University of British Columbia Faculty of Medicine, Research Scientist, British Columbia Mental Health and Addiction Research Institute, Vancouver, British Columbia, Canada.
31.2 Genetic Counseling for Psychiatric Conditions

Allan David Axelrad, M.D. Clinical Associate Professor of Psychiatry, Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, Texas.
33.4 Hypnosis

Liat Ayalon, Ph.D. Professor, School of Social Work, Bar-Ilan University Faculty of Medicine in the Galilee, Ramat Gan, Israel.
57.4l Counseling and Support Needs of Dementia Caregivers

Alexa Bagnell, M.D. Associate Professor, Department of Psychiatry, Dalhousie University Faculty of Medicine; Interim Chief of Psychiatry, Izaak Walton Killam Health Centre, Halifax, Nova Scotia, Canada.
55.8 School Consultation

Deanna M. Barch, Ph.D. Gregory B. Couch Professor of Psychiatry, Chair, Department of Psychological & Brain Sciences, Psychiatry, and Radiology, Washington University School of Medicine, St. Louis, Missouri.
1.28 Functional Brain Connectivity and Psychopathology

David A. Baron, D.O. Professor, Departments of Psychiatry and Neurosurgery; Assistant Dean for

International Relations, Keck School of Medicine of the University of Southern California, Los Angeles, California.
32.3 Other Psychiatric Emergencies

Anthony W. Bateman, F.R.C.Psych. Visiting Professor, University College London School of Medicine; Honorary Professor of Psychotherapy, University of Copenhagen Faculty of Health and Medical Sciences, Copenhagen, Denmark; MBT Training Coordinator, Anna Freud Centre, London, United Kingdom.
33.18 Mentalization-Based Treatment

Bernhard T. Baune, M.D., Ph.D. Professor and Chair of Psychiatry, Head of Discipline of Psychiatry, School of Medicine, Royal Adelaide Hospital, University of Adelaide Faculty of Health Sciences, Adelaide, Australia.
57.2b Complementary, Alternative, and Integrative Medicine in Geriatric Psychiatry

William R. Beardslee, M.D. Distinguished Gardner/Monks Professor of Psychiatry, Harvard Medical School; Director, Baer Prevention Initiatives, Department of Psychiatry, Boston Children's Hospital, Boston, Massachusetts.
55.9 Prevention of Psychiatric Disorders in Children and Adolescents

Judith S. Beck, Ph.D. Clinical Professor of Psychology in Psychiatry, Perelman School of Medicine at University of Pennsylvania, Philadelphia, Pennsylvania.
33.7 Cognitive Therapy

Joseph H. Beitchman, M.D. Professor, Department of Psychiatry, University of Toronto Faculty of Medicine; Staff Psychiatrist, Center for Addiction and Mental Health, Toronto, Ontario, Canada.
43.1 Language Disorder, 43.2 Social (Pragmatic) Communication Disorder, 43.3 Speech Sound Disorder, 43.4 Stuttering, 43.5 Unspecified Communication Disorder

Ruth M. Benca, M.D., Ph.D. Professor and Chair, Department of Psychiatry and Human Behavior, University of California Irvine School of Medicine, Irvine, Orange, California.
1.22 Basic Science of Sleep

David M. Benedek, M.D. Professor/Deputy Chair, Department of Psychiatry, F. Edward Hébert School of Medicine, Uniformed Services University of the Health Sciences, Chief of Forensics and Director, Center for Forensic Behavioral Sciences, Walter Reed National Military Medical Center, Bethesda, Maryland.
31.5 Disaster Psychiatry: Disasters, Terrorism, and War

Elissa P. Benedek, M.D. Adjunct Professor of Psychiatry, University of Michigan Medical School, Private Child Adolescent Forensic Psychiatry, Ann Arbor, Michigan.
55.6 Forensic Child and Adolescent Psychiatry

Dror Ben-Zeev, Ph.D. Professor, Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington.
12.13 Psychiatric Rehabilitation

Eugene V. Beresin, M.D. Professor of Psychiatry, Harvard Medical School; Executive Director, The MGH Clay Center for Young Healthy Minds; Senior Educator in Child and Adolescent Psychiatry, Massachusetts General Hospital, Boston, Massachusetts.

54.9 The Treatment of Adolescents

Sarah L. Berga, M.D. Frank R. Lock Professor and Chair, Department of Obstetrics and Gynecology, Section of Reproductive Endocrinology and Infertility, Wake Forest School of Medicine; Attending Physician and Executive Director, Women's Health Services; Wake Forest Baptist Medical Center, Winston Salem, North Carolina.
31.1 Psychiatry and Reproductive Medicine

R. Lindsey Bergman, Ph.D. Associate Clinical Professor, Department of Psychiatry and Biobehavioral Sciences; Jane and Terry Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
52.4 Selective Mutism

William Bernet, M.D. Professor Emeritus, Department of Psychiatry, Vanderbilt University School of Medicine, Nashville, Tennessee.
55.2 Child Maltreatment

Abesh Kumar Bhattacharjee, M.D., Ph.D. Staff Psychiatrist, Geriatric Psychiatry Clinic and BHIP Clinic, VA San Diego Healthcare System, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
57.6f Sexuality and Aging

Dinesh Bhugra, Ph.D. Emeritus Professor of Mental Health and Cultural Diversity, Health Service and Population Research Department, Institute of Psychiatry, Psychology and Neuroscience, King's College London; Former Honorary Consultant, Maudsley Hospital, London, United Kingdom.
62 World Aspects of Psychiatry

O. Joseph Bienvenu, M.D., Ph.D. Associate Professor and Director of Anxiety Disorders Clinic, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine; Director, Inpatient General Hospital Psychiatry Consultation-Liaison Service and Residents' Outpatient Continuity Clinic, Johns Hopkins Hospital, Baltimore, Maryland.
27.11 Psychiatric Morbidity Following Critical Illness

Barbara Bowles Biesecker, Ph.D. Adjunct Associate Professor, Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health; Program Director and Associate Investigator, Social and Behavioral Research Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland.
31.2 Genetic Counseling for Psychiatric Conditions

Adam Bisaga, M.D. Professor of Psychiatry, Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York.
11.9 Opioid Use and Related Disorders: From Neuroscience to Treatment

Deborah Blacker, M.D., Sc.D. Professor of Psychiatry, Harvard Medical School; Deputy Chair and Professor in Epidemiology, Harvard T.H. Chan School of Public Health; Associate Chief for Research, Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts.
7.8 Psychiatric Rating Scales

Mark A. Blais, Psy.D. Associate Professor of Psychology in Psychiatry, Harvard Medical School; Director, Psychological Evaluation and Research Laboratory, Massachusetts General Hospital, Boston, Massachusetts.
7.8 Psychiatric Rating Scales

Dan German Blazer, II, M.D., Ph.D. JP Gibbons Professor Emeritus, Department of Psychiatry and Behavioral Sciences; Professor, Department of Community and Family Medicine; Duke University School of Medicine, Durham; Adjunct Professor, Department of Epidemiology, University of North Carolina School of Medicine, Chapel Hill, North Carolina.
57.1b Epidemiology of Psychiatric Disorders

Pierre Blier, M.D., Ph.D. Professor, Department of Psychiatry, University of Ottawa Faculty of Medicine; Director, Mood Disorders Research, Royal Ottawa Mental Health Center, Ottawa, Ontario, Canada.
34.26 Selective Serotonin-Norepinephrine Reuptake Inhibitors

Michael H. Bloch, M.D. Assistant Professor, Yale Child Study Center, Yale School of Medicine, New Haven, Connecticut.
36.1 Psychiatric Examination of the Infant, Child, and Adolescent

Michael Blumenfield, M.D. Sidney E. Frank Distinguished Professor Emeritus of Psychiatry and Behavioral Sciences, New York Medical College, Valhalla, New York; Private Practice, Los Angeles, California.
27.14 Psychiatric Care of the Burned Patient

Michael P. Bogenschutz, M.D. Professor, Department of Psychiatry, University of New Mexico School of Medicine, Albuquerque, New Mexico.
11.6 Hallucinogen-Related Disorders, 33.17 Hallucinogen-Assisted Psychotherapy

Gary R. Bond, Ph.D. Senior Research Associate, Individual Placement and Support Employment Center, Westat, Lebanon, New Hampshire.
12.13 Psychiatric Rehabilitation

Olivier Bonnot, M.D., Ph.D. Professor, Department of Child and Adolescent Psychiatry, Faculté de Médecine, Université de Nantes, Nantes, France.
2.14 Neuropsychiatry of Neurometabolic and Neuroendocrine Disorders

Anna Borisovskaya, M.D. Assistant Professor, Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine; ECT Director, Staff Psychiatrist, VA Puget Sound Healthcare System, Seattle, Washington.
57.3b Psychiatric Problems in the Medically Ill Geriatric Patient

Jeff Q. Bostic, M.D., Ed.D. Professor of Child and Adolescent Psychiatry, Georgetown University School of Medicine; Division of Child Psychiatry, MedStar Georgetown University Hospital, Washington, D.C.
55.8 School Consultation

Claude Andrew Botha, Ph.D. Clinical Research Fellow, Centre for Digestive Diseases, Blizard Institute of Cell and Molecular Science, Wingate Institute of Neurogastroenterology, Barts and the London School of Medicine and Dentistry, Queen Mary University of London; Consultant Psychiatrist, Department of Gastroenterology, Barts Health NHS Trust, London, United Kingdom.

27.3 Gastrointestinal Disorders

Mousa Sobhy Azer Botros, M.D. Instructor, Department of Psychiatry and Behavioral Sciences, University of Miami Leonard M. Miller School of Medicine, Miami, Florida.
34.21 Mirtazapine

Mark E. Bouton, Ph.D. Lawson Green and Gold Professor of Psychology, Department of Psychological Sciences, University of Vermont College of Medicine, Burlington, Vermont.
3.3 Learning Theory

Charles L. Bowden, M.D. Professor of Psychiatry, Nancy U. Karren Chair of Psychiatry, Department of Psychiatry, University of Texas Medical School at San Antonio, San Antonio, Texas.
34.19 Lithium

Scott E. Bowen, Ph.D. Associate Professor, Chair, Behavioral and Cognitive Neuroscience, Department of Psychology, Wayne State University School of Medicine, Detroit, Michigan.
11.7 Inhalant-Related Disorders

Marjan Bozinoski, M.S. Research Scientist, Department of Pharmacology, Weill Cornell Medical College, New York, New York.
1.11 Genome, Transcriptome, and Proteome: The Molecular Genetics and Biochemistry Underlying the Neurobiology of Mental Disorders

Kathleen T. Brady, M.D., Ph.D. Distinguished University Professor and Vice President for Research, Medical University of South Carolina, Charleston, South Carolina.
34.24 Opioid Maintenance Treatments: Methadone and Buprenorphine

Matthew Neil Brams, M.D. Clinical Assistant Professor of Psychiatry, Baylor College of Medicine, Houston, Texas.
55.12 Impact on Parents of Raising a Child with Psychiatric Illness and/or Developmental Disability

James Robert Brašic´, M.D. Assistant Professor, Department of Radiology and Radiological Science, Division of Nuclear Medicine and Molecular Imaging, The Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, Maryland.
12.9 Molecular Brain Imaging in Schizophrenia

William S. Breitbart, M.D. Professor of Clinical Psychiatry, Vice Chairman, Department of Psychiatry, Weill Cornell Medical College; Chairman, Jimmie C. Holland Chair in Psychiatric Oncology, Chief, Psychiatry Service, Department of Psychiatry and Behavioral Sciences, Memorial Sloan Kettering Cancer Center, New York, New York.
27.8 Psycho-Oncology

Adam M. Brenner, M.D. Associate Professor of Psychiatry, Distinguished Teaching Professor, Vice Chair for Education, Director of Residency Training in Psychiatry, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas.
6.3 Other Psychodynamic Schools, 33.11 Other Methods of Psychotherapy

David A. Brent, M.D. Professor of Psychiatry, Pediatrics, and Epidemiology, University of Pittsburgh School of Medicine; Academic Chief, Child and Adolescent Psychiatry, UPMC Endowed Chair in Suicide

Studies, Pittsburgh, Pennsylvania.
51.1 Depressive Disorders and Suicide

Tim Bressmann, Ph.D. Associate Professor, Department of Speech-Language Pathology, University of Toronto Faculty of Medicine, Toronto, Ontario, Canada.
43.5 Unspecified Communication Disorder

Jonathan D. Brodie, M.D., Ph.D. Professor, Department of Psychiatry, New York University School of Medicine, New York, New York.
60.4 Neuroimaging, Psychiatry, and the Law

Evelyn J. Bromet, Ph.D. Distinguished Professor of Psychiatry and Preventive Medicine, Stony Brook University School of Medicine, Stony Brook, New York.
12.17 Other Psychotic Disorders

Shaun P. Brothers, Ph.D. Assistant Professor, Center for Therapeutic Innovation, Department of Psychiatry and Behavioral Sciences, University of Miami Leonard M. Miller School of Medicine, Miami, Florida.
1.6 Neuropeptides: Biology, Regulation, and Role in Neuropsychiatric Disorders

Daniel P. Brown, Ph.D. Associate Clinical Professor in Psychology, Department of Psychiatry, Harvard Medical School, Boston, Massachusetts.
33.4 Hypnosis

E. Sherwood Brown, M.D., Ph.D. Professor and Vice Chairman for Clinical Research, Aradine S. Ard Chair in Brain Science, Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas.
27.5 Respiratory Disorders

Regina Bussing, M.D. Donald R. Dizney Chair in Psychiatry, Professor and Interim Chair, Department of Psychiatry, University of Florida College of Medicine, Gainesville, Florida.
55.10 Child Mental Health Services Research

Joanne Alonso Byars, M.D. Assistant Professor, Beth K. and Stuart C. Yudofsky Division of Neuropsychiatry, Department of Psychiatry, Baylor College of Medicine, TIRR Memorial Hermann, Houston, Texas.
2.5 Neuropsychiatric Consequences of Traumatic Brain Injury

William M. Byne, M.D., Ph.D. Associate Professor, Department of Psychiatry, Icahn School of Medicine at Mount Sinai, Mental Illness Research, Education and Clinical Center, James J Peters VA Medical Center, Bronx, New York, New York.
21.2 Homosexuality, Gay and Lesbian Identities, and Homosexual Behavior, 21.4 Gender Identity, Gender Variance, and Gender Dysphoria

Christine C. Call, A.B. Doctoral Student in Clinical Psychology, Department of Psychology, Drexel University College of Medicine, Philadelphia, Pennsylvania.
22 Feeding and Eating Disorders

Glorisa Canino, Ph.D. Professor, Department of Pediatrics, University of Puerto Rico School of Medicine; Director, Behavioral Sciences Research Institute, University of Puerto Rico School of Medicine, Medical Sciences, San Juan, Puerto Rico.
58.7 A Sociocultural Framework for Mental Health and Substance Abuse Service Disparities

William A. Carlezon, Jr., Ph.D. Professor of Psychiatry and Neuroscience, Harvard Medical School; Chief, Basic Neuroscience Division; Director, Behavioral Genetics Laboratory, McLean Hospital, Belmont, Massachusetts.
1.24 Neuroscience of Substance Use Disorders

Gabrielle A. Carlson, M.D. Professor of Psychiatry and Pediatrics, Department of Psychiatry; Director Emerita, Division of Child and Adolescent Psychiatry, Distinguished Consulting Psychiatrist, Department of Psychiatry, Stony Brook University School of Medicine, Stony Brook, New York.
51.2 Early-Onset Bipolar Disorder

Regina M. Carney, M.D. Assistant Clinical Professor, Department of Psychiatry and Behavioral Sciences, University of Miami Leonard M. Miller School of Medicine, Miami, Florida.
34.12 Buspirone

B.J. Casey, Ph.D. Professor of Psychology, Yale School of Medicine, New Haven, Connecticut.
14.6 Neuroimaging and the Neuroanatomical Circuits Implicated in Anxiety, Fear, and Stress-Related Disorders

Tara Chandrasekhar, M.D. Medical Instructor, Division of Child and Family Mental Health and Developmental Neuroscience, Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, North Carolina.
53 Early-Onset Psychotic Disorders

Moses V. Chao, Ph.D. Professor of Cell Biology, Neuroscience and Physiology and Psychiatry, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, New York, New York.
1.7 Neurotrophic Factors

Dennis S. Charney, M.D. Professor, Departments of Psychiatry, Neuroscience, and Pharmacological Sciences; Anne and Joel Ehrenkranz Dean, Icahn School of Medicine at Mount Sinai; President for Academic Affairs, Mount Sinai Health System, New York, New York.
14.5 Anxiety Disorders: Neurobiology and Neuroscience

Elena H. Chartoff, Ph.D. Assistant Professor of Psychiatry and Neuroscience, Harvard Medical School, Boston, Massachusetts; Director, Neurobiology of Motivated Behavior Laboratory, McLean Hospital, Belmont, Massachusetts.
1.24 Neuroscience of Substance Use Disorders

Irene Chatoor, M.D. Professor of Psychiatry and Pediatrics, George Washington University School of Medicine and Health Sciences; Vice Chair, Department of Psychiatry and Behavioral Sciences, Children's National Medical Center, Washington, D.C.
47 Feeding and Eating Disorders of Infancy and Early Childhood

Stephen T. Chen, M.D. Associate Clinical Professor of Health Sciences, Department of Psychiatry and

Biobehavioral Sciences, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
57.3f Alzheimer Disease and Other Neurocognitive Disorders

Paul Ciechanowski, M.D. Clinical Associate Professor, Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington.
27.6 Diabetes: Psychosocial Issues and Psychiatric Disorders

Angel Cienfuegos, M.D. Associate Clinical Professor of Health Sciences, David Geffen School of Medicine at University of California Los Angeles; Mental Health Psychiatrist, Greater Los Angeles VA Health Care System, Los Angeles, California.
58.5 The Psychiatric Hospitalist

Chiara Cirelli, M.D., Ph.D. Professor, Department of Psychiatry, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin.
1.22 Basic Science of Sleep

Caroline Gray Clark, B.A., M.S.S.W. Researcher, Behavioral Activation Therapy and Theory Lab for Emotion, University of Tennessee Health Science Center College of Medicine, Memphis, Tennessee.
33.3 Behavior Therapy

E. Will Clark, M.D. Assistant Professor, Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas; Attending Psychiatrist and Assistant Medical Director, Zale Lipshy University Hospital – Inpatient Psychiatry, Dallas, Texas.
33.11 Other Methods of Psychotherapy

Michael R. Clark, M.D. Vice Chair, Clinical Affairs; Director, Pain Treatment Programs, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland.
27.15 Management of Chronic Pain

Diana E. Clarke, Ph.D. Research Statistician/Epidemiologist, American Psychiatric Association, Arlington, Virginia.
5.1 Quantitative and Experimental Methods in Psychiatry: Epidemiology

Anita H. Clayton, M.D. David C. Wilson Professor and Chair, Department of Psychiatry and Neurobehavioral Sciences; Professor of Clinical Obstetrics & Gynecology, University of Virginia School of Medicine, Charlottesville, Virginia.
34.40 Treatment of Erectile Disorder

Brett A. Clementz, Ph.D. Professor of Psychology and Neuroscience, Bioimaging Research Center, Medical College of Georgia, Augusta, Georgia.
12.11 Phenotypes of Psychosis

C. Robert Cloninger, M.D., Ph.D. Department of Psychiatry, Washington University School of Medicine in St. Louis, St. Louis, Missouri.
26 Personality Disorders

Robert T. Cobb, M.D. Assistant Professor of Clinical Psychiatry, Department of Psychiatry and Behavioral

Sciences, Keck School of Medicine of the University of Southern California; Attending Staff, Psychiatric Emergency Services, Los Angeles County and University of Southern California Medical Center; Attending Staff, Department of Psychiatry, Keck Hospital of the University of Southern California, Los Angeles, California.
32.3 Other Psychiatric Emergencies

Barbara J. Coffey, M.D. Chief, Tics and Tourette's Clinical and Research Program, Division of Tics, OCD, and Related Problems, Icahn School of Medicine at Mount Sinai; Professor, Department of Psychiatry, Mount Sinai Behavioral Science Unit, New York, New York.
48 Tic Disorders

Carl I. Cohen, M.D. Distinguished Service Professor, Division of Geriatric Psychiatry, State University of New York Downstate Medical Center, Brooklyn, New York.
57.3h Schizophrenia and Delusional Disorders

Judith A. Cohen, M.D. Professor of Psychiatry, Drexel University College of Medicine, Philadelphia, Pennsylvania; Professor and Medical Director, Center for Traumatic Stress in Children and Adolescents, Department of Psychiatry, Allegheny General Hospital, Pittsburgh, Pennsylvania.
52.2 Posttraumatic Stress Disorder in Children and Adolescents

Calvin A. Colarusso, M.D. Clinical Professor of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
56 Adulthood

Robert F. Cole, Ph.D. Adjunct Assistant Professor, Psychiatry, University of Connecticut School of Medicine, Farmington, Connecticut.
58.1 Public and Community Psychiatry, 58.2 Reforming Health Care

Hilary S. Connery, M.D., Ph.D. Assistant Professor, Department of Psychiatry, Harvard Medical School, Boston; Clinical Director, Division of Alcohol and Drug Abuse, McLean Hospital, Belmont, Massachusetts.
1.24 Neuroscience of Substance Use Disorders

Daniel F. Connor, M.D. Lockean Distinguished Chair in Mental Health Education, Research and Clinical Improvement, Professor of Psychiatry, Chief, Division of Child and Adolescent Psychiatry, University of Connecticut School of Medicine, Farmington, Connecticut.
46 Disruptive Behavior Disorders in Children and Adolescents

Jana R. Cooke, M.D. Pulmonary and Sleep Medicine, The Everett Clinic, Everett, Washington, D.C.
57.3c Sleep-Wake Disorders in Older Adults

Philip R. Corlett, Ph.D. Assistant Professor, Department of Psychiatry, Yale University School of Medicine; Ribicoff Research Facilities, Connecticut Mental Health Center, New Haven, Connecticut.
1.27 Learning Theory and Psychosis, Anxiety, and Addiction

Ann Elizabeth Cornell, Psy.D. Assistant Professor of Psychiatry, School of Medicine and Dentistry, University of Rochester School of Medicine and Dentistry, Rochester, New York.
57.4j Family Intervention and Therapy with Older Adults

Kimberly Ann Corp, M.S., M.F.T. Doctoral Candidate, Alliant International University, San Diego, California.
57.2d Psychological Changes with Normal Aging

Christoph U. Correll, M.D. Professor of Psychiatry and Molecular Medicine, Department of Psychiatry and Molecular Medicine, Hofstra North Shore-LIJ School of Medicine at Hofstra University; Medical Director, Recognition and Prevention (RAP) Program, Department of Psychiatry, Zucker Hillside Hospital, Hempstead, New York.
12.12 Schizophrenia: Pharmacological Treatment

Paul T. Costa, Jr., Ph.D. Adjunct Professor, Department of Psychiatry and Behavioral Science, Duke Medical Center, Durham, North Carolina; Professor of Behavioral Biology, Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine; Professor, Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore Maryland.
6.4 Approaches Derived from Philosophy and Psychology

Sara Costi, M.D. Research Fellow, Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York.
14.5 Anxiety Disorders: Neurobiology and Neuroscience

Joseph T. Coyle, M.D. Eben S. Draper Professor of Psychiatry and Neuroscience, Department of Psychiatry, Harvard Medical School, Boston, Massachusetts; McLean Hospital, Belmont, Massachusetts.
1.5 Amino Acid Neurotransmitters

Louis J. Cozolino, Ph.D. Professor of Psychology, Pepperdine University, Malibu, California; Adjunct Clinical Professor of Psychiatry, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
3.1 Contributions of the Psychological Sciences: Sensation, Perception, and Cognition

Adam R. Critchfield, M.D. Instructor in Psychiatry, Department of Consultation-Liaison Psychiatry, Columbia University College of Physicians and Surgeons; Assistant Attending Psychiatrist, Department of Consultation-Liaison Psychiatry, Columbia University Medical Center, New York, New York.
27.2 Cardiovascular Disorders

Paul E. Croarkin, D.O. Assistant Professor of Psychiatry, Mayo Clinic, Rochester, Minnesota.
57.4f Electroconvulsive Therapy and Other Neurostimulation Treatments

Elizabeth A. Crocco, M.D. Associate Clinical Professor, Department of Psychiatry and Behavioral Science, University of Miami Leonard M. Miller School of Medicine, Miami, Florida.
34.15 Cholinesterase Inhibitors and Memantine

Catherine Chang Crone, M.D. Associate Professor of Psychiatry, George Washington University School of Medicine and Health Sciences, Washington, D.C.; Clinical Professor of Psychiatry, Virginia Commonwealth University School of Medicine, Richmond; Vice Chair, Department of Psychiatry, Inova Fairfax Hospital, Falls Church, Virginia.
27.13 Organ Transplantation

Rosie E. Curiel Cid, Psy.D. Assistant Professor, Department of Psychiatry and Behavioral Sciences,

University of Miami Leonard M. Miller School of Medicine, University of Miami Hospital & Jackson Memorial Medical Center, Miami, Florida.
58.6 Psychiatric Rehabilitation

John F. Curry, Ph.D. Professor, Department of Psychiatry and Behavioral Sciences and Department of Psychology and Neuroscience, Duke University School of Medicine, Durham, North Carolina.
3.2 Piaget and Cognitive Development: History, Theory, Extensions and Modifications

Smita Das, M.D., Ph.D. Clinical Assistant Professor (Affiliated), Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine; Director of Addiction Treatment Services, Department of Psychiatry, Veterans Administration Palo Alto Health System Stanford, California.
11.8 Tobacco-Related Disorders

Keith A. Davidson, J.D. Partner, Albertson & Davidson, LLP, Ontario, California.
57.6b Forensic Aspects

***Michael C. Davis, M.D., Ph.D.** Medical Officer, Division of Psychiatry Products, Office of New Drugs, Center for Drug Evaluation and Research, U.S. Food and Drug Administration.
34.28 Second-Generation Antipsychotics

Dimitry S. Davydow, M.D. Associate Professor, Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington.
27.11 Psychiatric Morbidity Following Critical Illness

Nina De Lacy, M.D. Acting Assistant Professor, Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine; Co-Director, Developmental Neuropsychiatry, Seattle Children's Hospital, Seattle, Washington.
40 Intellectual Disability

Mark DeAntonio, M.D. Professor of Clinical Psychiatry, Jane and Terry Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine at University of California Los Angeles; Director, Inpatient Child and Adolescent Service, Resnick Neuropsychiatric Hospital, Los Angeles, California.
55.4 Psychiatric Sequelae of HIV and AIDS

Charles DeBattista, M.D. Professor of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California.
34.11 Bupropion, 34.38 Combination Pharmacotherapy

Lauren N. DeCaporale-Ryan, Ph.D. Assistant Professor of Psychiatry (Psychology), Medicine, & Surgery, University of Rochester School of Medicine and Dentistry, Rochester, New York.
57.4j Family Intervention and Therapy with Older Adults

Louisa Degenhardt, Ph.D. Professor of Epidemiology, National Drug and Alcohol Research Centre, University of New South Wales Faculty of Medicine, Sydney, New South Wales, Australia.
11.5 Cannabis-Related Disorders

Shannon L. Delaney, M.D. Child and Adolescent Psychiatry Division, Columbia University College of Physicians and Surgeons, New York, New York.

2.9 Neuropsychiatric Aspects of Other Infectious Diseases (Non-HIV)

James S. Demer, M.D. Assistant Professor, Department of Psychiatry and Behavioral Sciences, State University of New York Upstate Medical University, Syracuse, New York.
55.11 Impact of Terrorism on Children

Colin A. Depp, Ph.D. Associate Professor, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
57.6h Technology for Seniors

Abhilash K. Desai, M.D. Adjunct Associate Professor, Department of Psychiatry, Division of Geriatric Psychiatry, Saint Louis University School of Medicine, St. Louis, Missouri; Adjunct Associate Professor, Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington.
57.6a Psychiatric Aspects of Long-term Care

Davangere P. Devanand, M.D. Professor of Psychiatry and Neurology, Director of Geriatric Psychiatry, Columbia University College of Physicians and Surgeons; Attending Psychiatrist, New York Presbyterian Hospital New York, New York.
57.2a Psychiatric Assessment of the Older Patient

Mary Amanda Dew, Ph.D. Professor of Psychiatry, Psychology, Epidemiology, Biostatistics, Clinical and Translational Science, Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.
27.13 Organ Transplantation

Emanuel DiCicco-Bloom, M.D. Professor, Department of Neuroscience and Cell Biology, Department of Pediatrics (Division of Child Neurology and Neurodevelopmental Disabilities), Robert Wood Johnson Medical School of Rutgers University, New Brunswick, New Jersey.
1.3 Neural Development and Neurogenesis

Andrea Frances DiMartini, M.D. Professor of Psychiatry and Surgery, Medical Director of Transplant Psychiatry, University of Pittsburgh School of Medicine, Starzl Transplant Institute, Pittsburgh, Pennsylvania.
27.13 Organ Transplantation

Joel E. Dimsdale, M.D. Distinguished Professor of Psychiatry Emeritus and Research Professor, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
18 Somatic Symptom and Related Disorders, 27.10 Stress and Psychiatry

Arden D. Dingle, M.D. Professor, Department of Psychiatry, University of Texas Rio Grande Valley School of Medicine, Brownsville, Texas.
55.7 Ethical Issues in Child and Adolescent Psychiatry

Lisa B. Dixon, M.D., Ph.D. Professor of Psychiatry, Columbia University College of Physicians and Surgeons; Director, Division of Behavioral Health Services and Policy Research, New York State Psychiatric Institute, New York, New York.
58.4 Mental Health Services Research

Yuriy Dobry, M.D. Assistant Professor, Division of Child and Adolescent Psychiatry, Department of Psychiatry, University of California San Francisco School of Medicine, San Francisco, California.
34.29 Central Nervous System Sympathomimetic Compounds

Dan Doherty, M.D., Ph.D. Associate Professor of Pediatrics, Divisions of Genetic and Developmental Medicine, University of Washington School of Medicine, Seattle Children's Hospital, Seattle, Washington.
40 Intellectual Disability

Robert Lloyd Doyle, M.D., D.D.S. Instructor in Psychiatry, Harvard Medical School; Staff Psychiatrist, Child and Adolescent Psychiatry, Massachusetts General Hospital, Boston, Massachusetts.
50.2 Stereotypic Movement Disorders in Children

Robert E. Drake, M.D., Ph.D. Professor of Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth; Vice President, Westat Corporation, Lebanon, New Hampshire.
12.13 Psychiatric Rehabilitation

Jack Drescher, M.D. Clinical Professor of Psychiatry, Department of Psychiatry, Columbia University College of Physicians and Surgeons; Adjunct Professor, Postdoctoral Program in Psychotherapy and Psychoanalysis, New York University School of Medicine, New York, New York.
21.2 Homosexuality, Gay and Lesbian Identities, and Homosexual Behavior, 21.4 Gender Identity, Gender Variance, and Gender Dysphoria

Yuanli Duan, M.D., Ph.D. Attending Psychiatrist, Department of Psychiatry & Behavioral Sciences, Nassau University Medical Center, East Meadow, New York.
34.35 Electroconvulsive Therapy

Steven L. Dubovsky, M.D. Professor and Chair, Department of Psychiatry, Jacobs School of Medicine and Biomedical Sciences University at Buffalo, Buffalo, New York; Adjoint Professor of Psychiatry and Medicine, University of Colorado School of Medicine, Denver, Colorado; University Chair, Psychiatry, Erie County Medical Center, Buffalo, New York.
34.9 Barbiturates and Similarly Acting Substances,
34.10 Benzodiazepine Receptor Agonists and Antagonists,
34.13 Calcium Channel Inhibitors

Jennifer J. Dunkin, Ph.D. Clinical Professor (Voluntary), Department of Psychiatry and Behavioral Sciences, University of California San Diego School of Medicine, La Jolla, California; Staff Psychologist, San Diego Veterans Administration Healthcare System, San Diego, California.
57.2d Psychological Changes with Normal Aging

Dino Dvorak, Ph.D. Postdoctoral Associate, Neurobiology of Cognition Laboratory, Center for Neural Science, New York University School of Medicine, New York, New York.
1.26 Basic Systems Neuroscience

Jennifer B. Dwyer, M.D., Ph.D. Fellow, Albert J. Solnit Integrated Training Program in Adult and Child Psychiatry, Yale Child Study Center, Yale School of Medicine, New Haven, Connecticut.
1.25 Epigenetics in Psychiatry: The Promise for New Biomarkers and Treatments

Charles E. Eesley, Ph.D. Assistant Professor and Morgenthaler Faculty Fellow, Department of

Management Science and Engineering, School of Engineering, Stanford University School of Engineering, Stanford, California.
12.10 Neurocognition in Schizophrenia

Helen Link Egger, M.D. Assistant Professor of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, North Carolina.
36.2 Psychiatric Assessment of Preschool Children

Stuart J. Eisendrath, M.D. Professor of Psychiatry, University of California San Francisco School of Medicine, San Francisco, California.
19 Factitious Disorder

Zhanna Elberg, M.D. Assistant Professor of Clinical Psychiatry, Director, Psychotherapy Training, Department of Psychiatry, Jacobs School of Medicine and Biomedical Sciences University at Buffalo; Department of Psychiatry, Attending Psychiatrist/Director of Outpatient Residency Training, Department of Psychiatry, Erie County Medical Center Outpatient behavioral health, Buffalo, New York.
54.1 Individual Psychodynamic Psychotherapy (Child Psychiatry)

John D. Elsworth, Ph.D. Senior Research Scientist, Department of Psychiatry, Neuropsychopharmacology Research Unit, Yale University School of Medicine, New Haven, Connecticut.
1.4 Biogenic Amine Neurotransmitters

Steven A. Epstein, M.D. Professor and Chair, Department of Psychiatry, Georgetown University School of Medicine; Chief of Service, Department of Psychiatry, MedStar Georgetown University Hospital, Washington, D.C.
27.1 History and Current Trends

Rodrigo Escalona, M.D. Professor of Psychiatry, Vice Chair for Veterans Affairs, University of New Mexico School of Medicine; Director, Behavioral Health Care Line; Chief of Psychiatry, New Mexico VA Health Care System, Albuquerque, New Mexico.
12.2 Phenomenology of Schizophrenia

Javier I. Escobar, M.D. Professor of Psychiatry and Family Medicine, Associate Dean for Global Health, Robert Wood Johnson Medical School of Rutgers University, New Brunswick, New Jersey.
9.1 Present and Future of Classification Systems for Mental Disorders, 9.2 The Classification of Mental Disorders in the International Classification of Diseases (ICD-11), 18 Somatic Symptom and Related Disorders

Spencer Eth, M.D. Professor and Director, Forensic Psychiatry Fellowship Program, University of Miami Leonard M. Miller School of Medicine; Chief of Mental Health, Miami VA Healthcare System, Miami, Florida.
60.1 Clinical-Legal Issues in Psychiatry

John David Eun, B.A. Postbaccalaureate IRTA Fellow, Genetic Epidemiology Research Branch, National Institutes of Mental Health, Bethesda, Maryland.
14.3 Epidemiology of Anxiety Disorders

Lisa T. Eyler, Ph.D. Professor, Department of Psychiatry, University of California San Diego School of

Medicine, La Jolla, California; Clinical Research Psychologist, Desert Pacific Mental Illness Research Education, and Clinical Center, VA San Diego Healthcare System, San Diego, California.
57.2f Neuroimaging

Harris A. Eyre, M.B.B.S. (Hons.) Visiting Fulbright Graduate Researcher, Jane and Terry Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California; Ph.D. Candidate, Discipline of Psychiatry, University of Adelaide Faculty of Health Sciences, Adelaide, South Australia, Australia; Affiliate Research Fellow, College of Medicine and Dentistry, James Cook University College of Medicine and Dentistry, Townsville, Queensland, Australia.
57.2b Complementary, Alternative, and Integrative Medicine in Geriatric Psychiatry

Tanya J. Fabian, Pharm.D., Ph.D. Assistant Professor of Pharmacy and Therapeutics and Psychiatry, University of Pittsburgh School of Medicine; Director of Pharmacy Research and Pharmacy Services, Western Psychiatric Institute and Clinic of University of Pittsburg Medical Center, Pittsburgh, Pennsylvania.
10.2 Delirium

***Peter A. Fahnestock, M.D.** Affiliated Assistant Professor of Psychiatry, University of Central Florida College of Medicine, Orlando, Florida; Psychiatrist, Mental Health and Behavioral Service, Department of Veterans Affairs, Bay Pines Veterans Affairs Healthcare System, Bay Pines, Florida.
12.14 Medical Health in Schizophrenia

Nathan Fairman, M.D. Assistant Clinical Professor, Department of Psychiatry and Behavioral Sciences, University of California Davis School of Medicine, Sacramento, California.
57.6c Ethical Issues

Brian A. Fallon, M.D. Professor of Psychiatry, Columbia University College of Physicians and Surgeons, New York State Psychiatric Institute, New York, New York.
2.9 Neuropsychiatric Aspects of Other Infectious Diseases (Non-HIV)

Adam D. Farmer, Ph.D. Neurogastroenterology Group, Blizard Institute of Cell and Molecular Science, Wingate Institute of Neurogastroenterology, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom; Department of Gastroenterology, University Hospitals of North Midlands, Stoke on Trent, United Kingdom.
27.3 Gastrointestinal Disorders

Armando R. Favazza, M.D. Emeritus Professor, Department of Psychiatry, University of Missouri School of Medicine, Columbia, Missouri.
31.6 Psychiatry and Spirituality

Pariya L. Fazeli, Ph.D. Assistant Professor, Department of Family, Community, and Health Systems, School of Nursing, University of Alabama School of Medicine, Birmingham, Alabama.
57.3a Assessment of Functioning

Scott C. Fears, M.D., Ph.D. Associate Professor, Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
1.18 Gene Mapping Investigations of Psychiatric Disorders

Adriana Feder, M.D. Associate Professor, Psychiatry Department; Associate Director for Research, World Trade Center Mental Health Program, Icahn School of Medicine at Mount Sinai, New York, New York.
14.5 Anxiety Disorders: Neurobiology and Neuroscience

André Antonio Fenton, Ph.D. Professor of Neural Science, Department of Psychology, Center for Neural Science, New York University School of Medicine, New York, New York.
1.26 Basic Systems Neuroscience

Francisco Fernandez, M.D. Inaugural Dean, School of Medicine, Vice President, Medical Affairs, University of Texas Rio Grande Valley School of Medicine, Brownsville, Texas.
2.10 Neuropsychiatric Aspects of Prion Disease

Prudence W. Fisher, Ph.D. Associate Professor of Clinical Psychiatric Social Work (in Psychiatry), Columbia University College of Physicians and Surgeons; Adjunct Associate Professor, Columbia School of Social Work; Research Scientist, Division of Child and Adolescent Psychiatry, New York State Psychiatric Institute, New York, New York.
36.1 Psychiatric Examination of the Infant, Child, and Adolescent

Alastair J. Flint, M.B. Professor and Vice Chair Research, Department of Psychiatry, University of Toronto Faculty of Medicine, Psychiatrist, Centre for Mental Health, University Health Network, Toronto, Ontario, Canada.
57.4c Antianxiety Drugs

Edna B. Foa, Ph.D. Professor of Clinical Psychology in Psychiatry, Department of Psychiatry, Perelman School of Medicine at University of Pennsylvania, Philadelphia, Pennsylvania.
14.7 Anxiety Disorders: Cognitive-Behavioral Therapy

Laura J. Fochtmann, M.D. Distinguished Service Professor, Department of Psychiatry, Department of Pharmacological Sciences, Department of Biomedical Informatics, Stony Brook University School of Medicine, Stony Brook, New York.
7.3 Practice Guidelines in Psychiatry, 12.17 Other Psychotic Disorders

Peter Fonagy, Ph.D. Professor and Head, Research Department of Clinical, Educational, and Health Psychology, University College London School of Medicine, London, United Kingdom.
33.18 Mentalization-Based Treatment

Julian D. Ford, Ph.D. Professor of Psychiatry and Law; Director, Center for Trauma Recovery and Juvenile Justice, University of Connecticut School of Medicine, Farmington, Connecticut.
58.1 Public and Community Psychiatry, 58.2 Reforming Health Care

Brent P. Forester, M.D. Assistant Professor of Psychiatry, Harvard Medical School, Boston, Massachusetts; Director, Division of Geriatric Psychiatry, Mclean Hospital, Belmont; Medical Director – Behavioral Health, Center for Population Health, Partners HealthCare, Boston, Massachusetts.
57.3e Geriatric Mood Disorders

Martin Franklin, Ph.D. Associate Professor of Clinical Psychology in Psychiatry, Department of Psychiatry, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania.
54.2 Brief Psychotherapies for Childhood and Adolescence

Bradley W. Freeman, M.D. Assistant Professor of Clinical Psychiatry, Department of Psychiatry and Behavioral Sciences, Vanderbilt University School of Medicine, Nashville, Tennessee.
55.2 Child Maltreatment

Wanda P. Fremont, M.D. Professor, Department of Psychiatry; Chief, Division of Child and Adolescent Psychiatry, State University of New York Upstate Medical University College of Medicine, Syracuse, New York.
55.11 Impact of Terrorism on Children

Frederick J. Frese, Ph.D. Professor, Department of Psychiatry, Northeast Ohio Medical University, Rootstown, Ohio; Clinical Assistant Professor of Psychology, Department of Psychiatry, Case Western Reserve University School of Medicine, Cleveland, Ohio.
12.15 Recovery in Schizophrenia

Paul Frewen, Ph.D. Associate Professor, Department of Psychiatry, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada.
20 Dissociative Disorders

Mark A. Frye, M.D. Professor and Chair, Department of Psychiatry and Psychology, Mayo Clinic; Director, Mayo Clinic Depression Center, Rochester, Minnesota.
34.14 Carbamazepine, 34.34 Valproate

Douglas R. Galasko, M.D. Professor, Department of Neurosciences, University of California San Diego School of Medicine, La Jolla, California; Attending Neurologist, Department of Neurosciences, University of California San Diego Medical Center; Staff Physician, Neurology Service, Veterans Affairs Medical Center, San Diego, California.
57.2c The Aging Brain

Isaac R. Galatzer-Levy, Ph.D. Research Assistant Professor, Department of Psychiatry, New York University School of Medicine, New York, New York.
5.2 Statistics and Experimental Design

Martha C. Gamboa, M.D. Assistant Professor, Department of Psychiatry, New York Medical College, Valhalla, New York.
27.14 Psychiatric Care of the Burned Patient

Joanna Juan Gan, M.D. Instructor, Department of Neurology, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
2.4 Neuropsychiatric Aspects of Epilepsy

Eric L. Garland, Ph.D. Associate Dean for Research College of Social Work, University of Utah School of Medicine, Salt Lake City, Utah.
11.7 Inhalant-Related Disorders

Michael D. Garrett, M.D. Professor of Clinical Psychiatry, Director of Psychotherapy Training, State University of New York Downstate Medical Center College of Medicine, Brooklyn, New York.
33.16 Individual Psychodynamic Psychotherapy and Cognitive Behavioral Therapy for Psychosis

Jennifer Rose Gatchel, M.D., Ph.D. Instructor in Psychiatry, Department of Psychiatry, Harvard Medical School; Assistant Psychiatrist, Department of Psychiatry, Massachusetts General Hospital, Boston; McLean Hospital, Belmont, Massachusetts.
57.3e Geriatric Mood Disorders

Nori Geary, Ph.D. Department of Psychiatry, Weill Cornell Medical College (Retired), New York, New York.
1.23 Basic Science of Appetite

Dylan G. Gee, Ph.D. Assistant Professor, Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.
14.6 Neuroimaging and the Neuroanatomical Circuits Implicated in Anxiety, Fear, and Stress-Related Disorders

Jeffrey L. Geller, M.D. Professor of Psychiatry, University of Massachusetts Medical School, Worcester, Massachusetts.
58.3 The Role of the Hospital in the Care of Persons with Mental Illness

Cynthia M. A. Geppert, M.D., Ph.D. Professor, Department of Psychiatry, University of New Mexico School of Medicine; Chief, Consultation Psychiatry and Ethics, New Mexico Veterans Affairs Health Care System, Albuquerque, New Mexico.
25 Adjustment Disorders

Ruth S. Gerson, M.D. Assistant Professor (Clinical), Department of Child and Adolescent Psychiatry, New York University School of Medicine; Director, Children's Comprehensive Psychiatric Emergency Program, Bellevue Hospital Center, New York, New York.
36.4 Psychiatric Emergencies in Children and Adolescents

Subroto Ghose, M.D., Ph.D. Associate Professor, Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas.
12.6 Cellular and Molecular Neuropathology of Schizophrenia

Stephen J. Glatt, Ph.D. Associate Professor, Department of Psychiatry and Behavioral Sciences; Associate Professor, Department of Public Health and Preventive Medicine; Associate Professor, Department of Neuroscience and Physiology; Director, Psychiatric Genetic Epidemiology & Neurobiology Laboratory, State University of New York Upstate Medical University College of Medicine, Syracuse, New York.
57.2g Genetics of Late-Life Neurodegenerative Disorders

Mary Margaret Gleason, M.D. Associate Professor, Department of Psychiatry and Behavioral Sciences, Tulane University School of Medicine, New Orleans, Louisiana.
50.1 Reactive Attachment Disorder and Disinhibited Social Engagement Disorder

Araceli Gonzalez, Ph.D. Assistant Professor, Department of Psychology, California State University, Long Beach, California.
52.4 Selective Mutism

Aviel Goodman, M.D. Director, Minnesota Institute of Psychiatry, St. Paul, Minnesota.
21.5 Sexual Addiction

Eric Granholm, Ph.D. Professor, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California; Veteran Affairs San Diego Healthcare System, San Deigo, California.
57.4i Cognitive-Behavioral Therapy

John A. Gray, M.D., Ph.D. Assistant Professor, Department of Neurology, Center for Neuroscience, University of California Davis School of Medicine, Sacramento, California.
1.9 Intraneuronal Signaling

Julie A. Graziane, M.D. Assistant Professor, Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.
10.3 Dementia

Benjamin D. Greenberg, M.D., Ph.D. Professor, Department of Psychiatry and Human Behavior, Warren Alpert Medical School of Brown University; Butler Hospital, Providence, Rhode Island.
34.37 Neurosurgical Treatments

Harvey Roy Greenberg, M.D. Clinical Professor of Psychiatry, Albert Einstein College of Medicine at Yeshiva University, Bronx, New York.
16 Pathological and Problem Gambling: Gambling Disorder

Laurence L. Greenhill, M.D. Ruane Professor of Child Psychiatry, Division of Child and Adolescent Psychiatry, Columbia University College of Physicians and Surgeons; Research Psychiatrist, New York State Psychiatric Institute; Attending in Psychiatry, Department of Psychiatry, New York Presbyterian Hospital, New York, New York.
45.1 Attention-Deficit/Hyperactivity Disorder

†**Stanley I. Greenspan, M.D.** Clinical Professor of Psychiatry and Behavioral Science and Pediatrics, George Washington University School of Medicine and Health Sciences; Supervising Child Psychoanalyst, Washington Psychoanalytic Institute, Washington, D.C.
3.2 Piaget and Cognitive Development: History, Theory, Extensions and Modifications

Mindy Greenstein, Ph.D. Consultant, Department of Psychiatry and Behavioral Science, Memorial Sloan Kettering Cancer Center, New York, New York.
27.9 End-of-Life and Palliative Care

Paul Richard Greenwood, L.L.B. (Hons.) Deputy District Attorney, San Diego County and Head of Elder Abuse Prosecutions, San Diego, California.
57.6e The Legal and Ethical Analysis of Elder Abuse

Tiffany A. Greenwood, Ph.D. Assistant Professor, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
13.3 Mood Disorders: Genetics

Roland R. Griffiths, Ph.D. Department of Psychiatry and Behavioral Sciences, Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland.
11.4 Caffeine-Related Disorders

George T. Grossberg, M.D. Samuel W. Fordyce Professor, Director, Division of Geriatric Psychiatry,

Department of Psychiatry and Behavioral Neuroscience, St. Louis University School of Medicine, St. Louis, Missouri.
57.6a Psychiatric Aspects of Long-term Care

Gerhard Gründer, M.D. Professor of Psychiatry, Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Aachen, Germany.
12.9 Molecular Brain Imaging in Schizophrenia

Peter J. Guarnaccia, Ph.D. Professor and Investigator, Institute for Health, Health Care Policy, and Aging Research, Rutgers University, New Brunswick, New Jersey.
29.1 Cultural Concepts of Distress

Adarsh K. Gupta, M.D. Assistant Clinical Professor of Psychiatry, Albert Einstein College of Medicine of Yeshiva University, Bronx; Voluntary Attending Psychiatrist, Department of Psychiatry, Long Island Jewish Medical Center, Northshore University Hospital, New York, New York.
27.12 Psychocutaneous Disorders

Raquel E. Gur, M.D., Ph.D. The Karl and Linda Rickels Professor, Department of Psychiatry, Perelman School of Medicine at University of Pennsylvania, Philadelphia, Pennsylvania.
12.8 Functional Brain Imaging in Schizophrenia

Ruben C. Gur, Ph.D. Professor, Department of Psychiatry, Director, Brain Behavior Laboratory and the Center for Neuroimaging in Psychiatry Hospital, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania.
12.8 Functional Brain Imaging in Schizophrenia

Debra A. Gusnard, M.D. Retired, Departments of Radiology and Psychiatry, Washington University School of Medicine, St. Louis, Missouri.
1.21 Basic Science of Self

Robert W. Guynn, M.D. Professor Emeritus, Department of Psychiatry and Behavioral Science, John P. and Kathrine G. McGovern Medical School, Houston, Texas.
33.8 Interpersonal Psychotherapy

Barry H. Guze, M.D. Professor, Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at University of California Los Angeles; Chief of Psychiatry, Greater Los Angeles VA Medical Center, Los Angeles, California.
7.6 Medical Assessment and Laboratory Testing in Psychiatry, 58.5 The Psychiatric Hospitalist

Kathleen Y. Haaland, Ph.D. Professor, Department of Psychiatry and Behavioral Sciences, Department of Neurology, University of New Mexico School of Medicine, Albuquerque, New Mexico.
7.4 Clinical Neuropsychology and Intellectual Assessment of Adults

Fadi Haddad, M.D. Clinical Assistant Professor, Department of Child and Adolescent Psychiatry, New York University School of Medicine; Medical Director, Center for Family Development NYC Division, New York, New York.
34.40 Treatment of Erectile Disorder, 36.4 Psychiatric Emergencies in Children and Adolescents

Wayne D. Hall, Ph.D. Professor, Center for Youth Substance Abuse Research, University of Queensland School of Medicine, Herston, Queensland, Australia; Professor, National Addiction Centre, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London, England.
11.5 Cannabis-Related Disorders

Michael C. Hann, M.D. Directorate for Mental Health, Naval Medical Center San Diego, San Diego, California.
57.3j Alcohol and Substance Abuse in Older Adults

Meredith A. Hannan, M.D., Ph.D. MIRECC Advanced Fellow, Department of Psychiatry, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
37 Genetics in Child Psychiatry

Helena B. Hansen, M.D., Ph.D. Assistant Professor, Departments of Psychiatry and Anthropology, New York University School of Medicine, New York, New York; Nathan Kline Institute for Psychiatric Research, Orangeburg, New York.
7.10 Cultural Competency in the Organization and Delivery of Care

Ansar M. Haroun, M.D. Clinical Professor of Psychiatry, University of California San Diego School of Medicine, La Jolla, California; Adjunct Professor of Law, University of San Diego School of Law, San Diego, California.
57.6b Forensic Aspects, 57.6e The Legal and Ethical Analysis of Elder Abuse

Debra S. Harris, M.D. Staff Psychiatrist, Department of Veterans Affairs, Monterey Mental Health Clinic, Seaside, California.
1.12 Psychoneuroendocrinology

Philip D. Harvey, Ph.D. Leonard M. Miller Professor of Psychiatry and Behavioral Sciences, University of Miami Leonard M. Miller School of Medicine, Senior Health Scientist, Research Service, Bruce W. Carter VA Medical Center, Miami, Florida.
12.10 Neurocognition in Schizophrenia, 34.17 First-Generation Antipsychotics

Sameer Hassamal, M.D. Addiction Psychiatry Fellow, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California; Addiction Psychiatry Fellow, Kern Medical Center, Bakersfield, California.
24.1 Intermittent Explosive Disorder, 24.2 Pyromania, 24.3 Kleptomania

Margaret J. Hauck, Ph.D. Child Psychiatry Faculty, Department of Psychiatry and Behavioral Sciences, Tulane University School of Medicine, New Orleans, Louisiana.
50.1 Reactive Attachment Disorder and Disinhibited Social Engagement Disorder

Gary Haugland, M.A. Co-Director, Center of Excellence in Culturally Competent Mental Health, Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York.
7.10 Cultural Competency in the Organization and Delivery of Care

Lily I. Hechtman, M.D. Professor of Psychiatry and Pediatrics, Director of Research, Division of Child Psychiatry, McGill University Faculty of Medicine; Montreal Children's Hospital, Montreal, Quebec, Canada.

45.1 Attention-Deficit/Hyperactivity Disorder

Sean P. Heffernan, M.D. Schweitzer Fellow in Affective Disorders, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland.
27.8 Psycho-Oncology

Cassis L. Henry, M.D. Psychiatrist, Behavioral Health, Boston Health Care for the Homeless Program, Boston, Massachusetts.
58.4 Mental Health Services Research

John M. Hettema, M.D., Ph.D. Associate Professor of Psychiatry, Virginia Institute for Psychiatric and Behavioral Genetics, Attending Psychiatrist, VCUHS Anxiety Disorders Specialty Clinic, Department of Psychiatry, Virginia Commonwealth University School of Medicine, Richmond, Virginia.
14.4 Genetics of Anxiety Disorders

Robert Hindman, Ph.D. Clinical Faculty, Beck Institute for Cognitive Behavior Therapy, Bala Cynwyd, Pennsylvania.
33.7 Cognitive Therapy

Max Hirshkowitz, Ph.D. Sleep Disorders and Research Center, Michael E. DeBakey VA Medical Center, Houston, Texas.
23 Sleep Disorders

J. Gregory Hobelmann, M.D. Part time Faculty, Department of Psychiatry and Behavioral Medicine, Johns Hopkins University School of Medicine; Psychiatrist and Liaison, Pain Program, Ashley Addiction Treatment, Baltimore, Maryland.
27.15 Management of Chronic Pain

Eric Hollander, M.D. Clinical Professor of Psychiatry and Behavioral Sciences Albert Einstein College of Medicine of Yeshiva University; Director of Autism and Obsessive Compulsive Spectrum Program and Director of Anxiety and Depression Program, Montefiore Medical Center, Bronx, New York.
34.4 α_2-Adrenergic Receptor Agonists: Clonidine, Clonidine Extended Release, Guanfacine, and Guanfacine Extended Release

Harry C. Holloway, M.D. Professor of Psychiatry, Professor of Neuroscience, Department of Psychiatry, F. Edward Hebert School of Medicine, Uniformed Services University of the Health Services, Bethesda, Maryland.
31.5 Disaster Psychiatry: Disasters, Terrorism, and War

Paul E. Holtzheimer, M.D. Associate Professor, Departments of Psychiatry and Surgery; Director, Mood Disorders Service, Geisel School of Medicine at Dartmouth, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire.
34.22 Monoamine Oxidase Inhibitors

Vivian Hook, Ph.D. Professor, Departments of Neurosciences and Pharmacology, Skaggs School of Pharmacy and Pharmaceutical Sciences, School of Medicine, University of California San Diego School of Medicine, La Jolla, California.
57.4d Antipsychotic Drugs

Dina Hooshyar, M.D. Assistant Professor, Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas; Medical Director, VA North Texas Health Care System Comprehensive Homeless Center Programs, Dallas, Texas.
12.15 Recovery in Schizophrenia

Derek R. Hopko, Ph.D. Professor, Department of Psychology, University of Tennessee College of Medicine, Knoxville, Tennessee.
33.3 Behavior Therapy

Andrew Horti, Ph.D. Professor of Radiology, Johns Hopkins University School of Medicine, Baltimore, Maryland.
12.9 Molecular Brain Imaging in Schizophrenia

David Akers Hosford, M.D., Ph.D. Adjunct Associate Professor, Department of Neurology, Duke University School of Medicine, Durham, North Carolina; U.S. Food and Drug Administration, Division of Neurology Products (CDER), Food and Drug Administration, Silver Springs, Maryland.
34.2 Drug Development and Approval Process in the United States

Rory P. Houghtalen, M.D. Medical Director, Education and Adult Mental Health Ambulatory Services, Rochester Regional Health System, Unity Behavioral Health, Rochester, New York.
7.1 Psychiatric Interview, History, and Mental Status Examination of the Adult Patient

Matthew O. Howard, Ph.D. Daniels Distinguished Professor, Associate Dean for Doctoral Education, Editor, Journal of Addictive Diseases, University of North Carolina School of Medicine, Chapel Hill, North Carolina.
11.7 Inhalant-Related Disorders

Laura S. Howe-Martin, Ph.D. Associate Professor, Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas; Licensed Psychologist, Department of Psychiatry, UT Southwestern Medical Center, Dallas, Texas.
6.3 Other Psychodynamic Schools

Jeffrey Hsu, M.D. Assistant Professor, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland.
2.8 Neuropsychiatric Aspects of HIV Infection and AIDS

Leighton Y. Huey, M.D. Birnbaum/Blum Professor, Department of Psychiatry, University of Connecticut School of Medicine, Farmington, Connecticut.
58.1 Public and Community Psychiatry, 58.2 Reforming Health Care

Lorie A. Humphrey, Ph.D. Associate Clinical Professor (Voluntary), Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
36.3 Neuropsychological and Cognitive Assessment of Children

Jonathan D. Huppert, Ph.D. Professor and Chair, Department of Psychology, Hadassah Medical School, Jerusalem, Israel.
14.7 Anxiety Disorders: Cognitive-Behavioral Therapy

Mustafa M. Husain, M.D. Adjunct Professor of Psychiatry and Neurology, Member, Duke Institute of Brain Sciences (DIBS), Professor of Psychiatry, Neurology, and Medicine; Vice Chair, Department of Psychiatry, Duke University School of Medicine, Durham, North Carolina; Lydia Bryant Test Distinguished Professorship in Psychiatry Research; Director, Neuromodulation Research & Therapeutics Program, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas.
57.4f Electroconvulsive Therapy and Other Neurostimulation Treatments

Kumail Hussain, M.D. Instructor, Department of Psychiatry, Rush Medical College, Chicago, Illinois.
34.3 Medication-Induced Movement Disorders

Heidi E. Hutton, Ph.D. Associate Professor, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland.
2.8 Neuropsychiatric Aspects of HIV Infection and AIDS

Celia F. Hybels, Ph.D. Associate Professor Emeritus, Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, North Carolina.
57.1b Epidemiology of Psychiatric Disorders

Brian M. Iacoviello, Ph.D. Assistant Professor, Department of Psychiatry; Director of Psychotherapy Research, Mood and Anxiety Disorders Program, Icahn School of Medicine at Mount Sinai; Attending Psychologist, Department of Psychiatry, Mount Sinai Hospital, New York, New York.
14.5 Anxiety Disorders: Neurobiology and Neuroscience

Rocco A. Iannucci, M.D. Instructor in Psychiatry, Harvard Medical School, Boston, Massachusetts; Director, McLean Fernside, Division of Alcohol and Drug Abuse, McLean Hospital, Belmont, Massachusetts.
11.3 Stimulant-Related Disorders

Lourdes Illa-Sanchez, M.D. Associate Professor, Department of Psychiatry and Behavioral Sciences, University of Miami Leonard M. Miller School of Medicine, Miami, Florida.
34.41 Vilazodone

Scott A. Irwin, M.D., Ph.D. Associate Professor of Clinical Psychiatry; Director, Psychiatry and Psychosocial Services, Patient and Family Support Services, San Diego Moores Cancer Center; Director, Palliative Care Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
57.6c Ethical Issues

Anna Ivanenko, M.D., Ph.D. Professor of Clinical Psychiatry and Behavioral Sciences, Feinberg School of Medicine, Northwestern University, Division of Child and Adolescent Psychiatry, Ann and Robert H. Lurie Children's Hospital of Chicago, Chicago, Illinois.
55.13 Pediatric Sleep Disorders

Iliyan Ivanov, M.D. Associate Professor, Division of Child and Adolescent Psychiatry, Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York.
34.16 Disulfiram and Acamprosate

Elena I. Ivleva, M.D., Ph.D. Assistant Professor, Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas.

12.16 Psychosis as a Defining Dimension in Schizophrenia

Assen Jablensky, M.D. Winthrop Professor of Psychiatry; Director, Centre for Clinical Research in Neuropsychiatry (CCRN), Division of Psychiatry, University of Western Australia Faculty of Medicine, Dentistry and Health Sciences; Consultant Psychiatrist, Royal Perth Hospital, Perth, Western Australia, Australia.
12.3 Worldwide Burden of Schizophrenia

Julienne Jacobson, M.D. Assistant Professor of Clinical Psychiatry and Pediatrics, Children's Hospital Los Angeles, Keck School of Medicine of the University of Southern California, Los Angeles, California.
55.3 Children's Reaction to Illness and Hospitalization

Robert J. Jaffe, M.D. Assistant Professor of Psychiatry and Pediatrics, Icahn School of Medicine at Mount Sinai, New York, New York.
35.2 Normal Child Development

Mohammad Jafferany, M.D. Clinical Associate Professor, Department of Psychiatry, Central Michigan University College of Medicine, Mount Pleasant; Medical Director, Jafferany Psychiatric and JPS Psychological Services, Saginaw, Michigan.
27.12 Psychocutaneous Disorders

Martha L. James, M.D., J.D. Los Angeles, California.
7.6 Medical Assessment and Laboratory Testing in Psychiatry

Philip G. Janicak, M.D. Adjunct Professor, Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago, Illinois.
34.3 Medication-Induced Movement Disorders

Dilip V. Jeste, M.D. Senior Associate Dean for Healthy Aging and Senior Care, Estelle and Edgar Levi Chair in Aging; Distinguished Professor of Psychiatry and Neurosciences; Director, Sam and Rose Stein Institute for Research on Aging, University of California San Diego School of Medicine, La Jolla, California.
57.1a Introduction to Geriatric Psychiatry, 57.6i Positive Psychiatry of Aging

Russell T. Joffe, M.D. Professor of Psychiatry, Department of Psychiatry, Hofstra Northwell School of Medicine at Hofstra University; Chairman, Department of Psychiatry and Behavioral Sciences, North Shore Staten Island University Hospital, Staten Island, New York.
2.7 Neuropsychiatric Aspects of Multiple Sclerosis and Other Demyelinating Disorders, 34.31 Thyroid Hormones

Jason K. Johannesen, Ph.D. Assistant Professor of Psychiatry, Yale University School of Medicine, New Haven, Connecticut; Staff Psychologist, VA Connecticut Healthcare System, West Haven, Connecticut.
1.15 Applied Electrophysiology

†**E. Roy John, Ph.D.** Professor of Psychiatry, Department of Psychiatry, New York University School of Medicine, New York; Former Director, Brain Research Laboratories, New York, New York.
7.7 Principles and Applications of Quantitative Electroencephalography in Psychiatry

Sietse Jonkman, Ph.D. Research Associate, Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, New York, New York.

1.19 Animal Models in Psychiatry

Ricardo E. Jorge, M.D. Professor of Psychiatry and Behavioral Sciences, Michael E. DeBakey VA Medical Center, Baylor College of Medicine, Houston, Texas.
2.5 Neuropsychiatric Consequences of Traumatic Brain Injury

Grace Mitu Juarez, M.D. Clinical Psychiatrist, Department of Psychiatry and Behavioral Sciences, Los Angeles and University of Southern California Medical Center, Los Angeles, California.
32.3 Other Psychiatric Emergencies

Laura M. Juliano, Ph.D. Professor, Department of Psychology, American University, Washington, D.C.
11.4 Caffeine-Related Disorders

Rahil Jummani, M.D. Clinical Assistant Professor, Medical Director, Long Island Campus, Department of Child and Adolescent Psychiatry, The Child Study Center at NYU Langone Medical Center, Lake Success, New York.
48 Tic Disorders

Martha Bates Jura, Ph.D. Clinical Professor (Voluntary), Department of Psychiatry and Biobehavioral Sciences; Attending Psychologist, Resnick Neuropsychiatric Hospital, Jane and Terry Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
36.3 Neuropsychological and Cognitive Assessment of Children

Arie Kaffman, M.D., Ph.D. Assistant Clinical Professor, Departments of Psychiatry and the Child Study Center, Yale University School of Medicine, New Haven; Psychiatrist, Post Traumatic Stress Disorder firm, Veterans Affairs Medical Center, Newington, Connecticut.
1.25 Epigenetics in Psychiatry: The Promise for New Biomarkers and Treatments

Amir Kalali, M.D. Clinical Professor of Psychiatry, University of California San Diego School of Medicine, La Jolla; Global Head, Neuroscience Center of Excellence, Quintiles, Inc., California.
34.2 Drug Development and Approval Process in the United States

Rujvi Kamat, Ph.D. Assistant Professor, Neuropsychologist, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
57.6h Technology for Seniors

Gen Kanayama, M.D., Ph.D. Instructor, Department of Psychiatry, Harvard Medical School, Boston, Massachusetts; Associate Director for Substance Abuse Research, Biological Psychiatry Laboratory, McLean Hospital, Belmont, Massachusetts.
11.11 Anabolic-Androgenic Steroid-Related Disorders

John M. Kane, M.D. Professor and Chairman, Department of Psychiatry, Hofstra North Shore-LIJ School of Medicine at Hofstra University, Hempstead; Chairman, Department of Psychiatry, The Zucker Hillside Hospital, Glen Oaks, New York.
12.12 Schizophrenia: Pharmacological Treatment

Adam I. Kaplin, M.D., Ph.D. Assistant Professor, Department of Psychiatry and Behavioral Science, Johns Hopkins University School of Medicine, Baltimore, Maryland.

1.8 Novel Neurotransmitters

Sylvia R. Karasu, M.D. Clinical Professor of Psychiatry, Weill Cornell Medical College; Attending Psychiatrist, New York Presbyterian Hospital; Member, Institutional Review Board, Rockefeller University, New York, New York.
33.1 Psychoanalysis and Psychoanalytic Psychotherapy

Ethan F. Karsen, M.D. Instructor in Psychiatry, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire.
34.22 Monoamine Oxidase Inhibitors

†**Wayne Katon, M.D.** Professor, Department of Psychiatry and Behavioral Sciences, University of Washington Medical School, Seattle, Washington.
27.6 Diabetes: Psychosocial Issues and Psychiatric Disorders

Jeffrey W. Katzman, M.D. Professor and Vice Chair, Department of Psychiatry, Clinical and Academic Services, University of New Mexico School of Medicine; Executive Medical Director, Behavioral Health Care, University of New Mexico Hospital, Albuquerque, New Mexico.
25 Adjustment Disorders

David L. Kaye, M.D. Professor of Psychiatry, Vice Chair for Academic Affairs (Psychiatry), Medical Director, CAP PC, Dept. of Psychiatry, Erie County Medical Center, Jacobs School of Medicine and Biomedical Sciences University at Buffalo, Buffalo, New York.
54.1 Individual Psychodynamic Psychotherapy (Child Psychiatry)

Francis J. Keefe, Ph.D. Professor, Departments of Psychiatry and Behavioral Sciences, Psychology and Neuroscience, Anesthesiology, and Medicine; Director, Pain Prevention and Treatment Research Program, Duke University College of Medicine, Durham, North Carolina.
27.10 Stress and Psychiatry

Richard S.E. Keefe, Ph.D. Professor of Psychiatry, Behavioral Sciences, Psychology and Neuroscience, Duke University School of Medicine, Durham, North Carolina.
12.10 Neurocognition in Schizophrenia

George Alan Keepers, M.D. Carruthers Professor and Chair, Department of Psychiatry, Oregon Health and Science University School of Medicine, Portland, Oregon.
59.1 Graduate Psychiatric Education

Courtney P. Keeton, Ph.D. Assistant Professor, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland.
52.3 Separation Anxiety, Generalized Anxiety, and Social Anxiety

Samuel J. Keith, M.D. Milton Rosenbaum Professor of Psychiatry and Psychology, Emeritus, Chairman, Department of Psychiatry, University of New Mexico School of Medicine, Albuquerque, New Mexico.
12.2 Phenomenology of Schizophrenia

Robert Emmett Kelly, Jr., M.D. Assistant Professor of Psychiatry, Weill Cornell Medical College, New York, New York.

57.3e Geriatric Mood Disorders

John R. Kelsoe, M.D. Professor, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
13.3 Mood Disorders: Genetics

Kimberley M. Kendall, B.Sc. (Hons.), M.B.B.Ch., M.R.C.Psych. Wellcome Trust Clinical Research Fellow, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff, United Kingdom.
12.4 Genetics of Schizophrenia

Sidney H. Kennedy, M.D. Professor of Psychiatry, University of Toronto Faculty of Medicine; Arthur Sommer Rotenberg Chair in Depression and Suicide Studies, St. Michael's Hospital, Toronto, Ontario, Canada.
13.7 Mood Disorders: Pharmacological Treatment of Depression and Bipolar Disorders

Paul J. Kenny, Ph.D. Professor, Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, New York, New York.
1.19 Animal Models in Psychiatry

Matcheri S. Keshavan, M.D. Stanley Cobb Professor of Psychiatry, Harvard Medical School; Vice Chair of Public Psychiatry, Beth Israel Deaconess Medical Center, Boston, Massachusetts.
12.11 Phenotypes of Psychosis

Terence Arthur Ketter, M.D. Professor, Department of Psychiatry and Behavioral Sciences, Stanford School of Medicine, Stanford, California.
34.7 Anticonvulsants: Gabapentin, Topiramate, Tiagabine, Levetiracetam, Zonisamide, and Pregabalin, 34.18 Lamotrigine

Amir A. Khan, M.D. Assistant Clinical Professor of Psychiatry, Warren Alpert Medical School of Brown University, Providence VA Hospital, Providence, Rhode Island.
34.23 Nefazodone

Ryan J. Kimmel, M.D. Associate Professor, Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine; Medical Director, Hospital Psychiatry, University of Washington Medical Center, Seattle, Washington.
14.2 Clinical Features of the Anxiety Disorders

Bryan H. King, M.D. Professor and Vice Chair of Psychiatry and Behavioral Sciences, Director, Seattle Children's Autism Center; Director of Child and Adolescent Psychiatry, University of Washington School of Medicine, Seattle Children's Hospital, Seattle, Washington.
40 Intellectual Disability

Deborah A. King, Ph.D. Professor of Psychiatry (Psychology); Associate Chair for Education, Department of Psychiatry, University of Rochester School of Medicine and Dentistry, Rochester, New York; VA VISN 2 Center of Excellence for Suicide Prevention, Canandaigua, New York.
57.4j Family Intervention and Therapy with Older Adults

Robert A. King, M.D. Yale Child Study Center, Yale University School of Medicine, New Haven,

Connecticut.
36.1 Psychiatric Examination of the Infant, Child, and Adolescent

Laurence J. Kirmayer, M.D. James McGill Professor and Director, Division of Social and Transcultural Psychiatry, McGill University Faculty of Medicine; Director, Culture & Mental Health Research Unit, Department of Psychiatry, Jewish General Hospital, Montreal, Quebec, Canada.
29.1 Cultural Concepts of Distress

George Kirov, M.D., Ph.D. Professor in Psychiatry, Cardiff University School of Medicine; MRC Centre for Neuropsychiatric Genetics & Genomics, Cardiff, United Kingdom.
12.4 Genetics of Schizophrenia

Ami Klin, Ph.D. Georgia Research Alliance Eminent Scholar Professor and Chief, Division Of Autism & Related Disorders, Department Of Pediatrics, Emory University School Of Medicine; Director, Marcus Autism Center, Children's Healthcare Of Atlanta, Atlanta, Georgia.
44 Autism Spectrum Disorder and Social Communication Disorder

Dana Kober, M.D. Clinical Assistant Professor, Baylor College of Medicine; Psychiatrist, Legacy Community Health Services, Houston, Texas.
54.7 Inpatient Psychiatric, Partial Hospital, and Residential Treatment for Children and Adolescents

Steve Koh, M.D. Assistant Professor; Director, Outpatient Psychiatric Services, Hillcrest; Director, Community Psychiatry Program; Director, TeleMentalHealth Program, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
57.3j Alcohol and Substance Abuse in Older Adults, 57.5 Financial Issues in the Delivery of Geriatric Psychiatric Care

Robert Kohn, M.D. Professor, Department of Psychiatry and Human Behavior, Warren Alpert Medical School of Brown University, Providence, Rhode Island.
4.3 Transcultural Psychiatry

Alexander Kolevzon, M.D. Professor of Psychiatry and Pediatrics, Child Behavioral Health and Science Center, Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York.
35.2 Normal Child Development

Susan G. Kornstein, M.D. Professor of Psychiatry and Obstetrics-Gynecology; Director, Institute for Women's Health, Virginia Commonwealth University School of Medicine, Richmond, Virginia.
34.23 Nefazodone, 34.32 Trazodone

Emiko Koyama, Ph.D. Margaret and Wallace McCain Centre for Child, Youth & Family Mental Health, Underserved Populations Program, Centre for Addiction and Mental Health, Toronto, Ontario, Canada.
43.1 Language Disorder, 43.2 Social (Pragmatic) Communication Disorder, 43.3 Speech Sound Disorder

Christopher J. Kratochvil, M.D. Associate Vice Chancellor for Clinical Research, University of Nebraska College of Medicine; Vice President for Research, Nebraska Medicine, Chief Medical Officer, UNeHealth, Omaha, Nebraska.
54.6 Pediatric Psychopharmacology

Suchitra Krishnan-Sarin, Ph.D. Professor of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.
34.25 Opioid Receptor Antagonists: Naltrexone and Nalmefene

Robert M. Kroll, Ph.D. Adjunct Professor, Department of Speech-Language Pathology, University of Toronto Faculty of Medicine; Executive Director, The Speech and Stuttering Institute, Toronto, Ontario, Canada.
43.4 Stuttering

Jennifer L. Kruse, M.D. Assistant Professor, Associate Director, Consultation and Liaison Psychiatry Service, Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
7.6 Medical Assessment and Laboratory Testing in Psychiatry

John H. Krystal, M.D. Robert L. McNeil, Jr., Professor of Translational Research; Chair of Psychiatry, Yale University School of Medicine; Chief of Psychiatry Services, Yale-New Haven Hospital; Director, Clinical Neuroscience Division, VA National Center for PTSD; Director, NIAAA Center for the Translational Neuroscience of Alcoholism, New Haven, Connecticut.
1.1 The Neuroscience of Psychiatry, 1.16 Nuclear Magnetic Resonance Imaging and Spectroscopy: Basic Principles and Recent Findings in Neuropsychiatric Disorders, 1.17 Radiotracer Imaging with Positron Emission Tomography and Single Photon Emission Computed Tomography: Fundamental Principles, Methodology, and Role in Neuropsychiatric Research

Marek Kubicki, M.D., Ph.D. Associate Professor, Departments of Psychiatry and Radiology, Harvard Medical School; Research Scientist, Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts.
12.7 Structural Brain Imaging in Schizophrenia

Helen H. Kyomen, M.D. Assistant Professor, Department of Psychiatry, Boston University School of Medicine; Lecturer, Department of Psychiatry, Harvard Medical School, Boston, Massachusetts Medical Director of the Geriatric Psychiatry Program, Department of Psychiatry, St. Elizabeth's Medical Center; Clinical Affiliate, Department of Psychiatry, McLean Hospital, Greater Boston area, Massachusetts.
57.5 Financial Issues in the Delivery of Geriatric Psychiatric Care, 57.6d Gender Issues

Beth Ann LaBardi, M.D. Acting Assistant Professor, Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington.
57.3b Psychiatric Problems in the Medically Ill Geriatric Patient

James H. Lake, M.D. Clinical Assistant Professor, Department of Medicine, Arizona Center for Integrative Medicine, University of Arizona College of Medicine, Tucson, Arizona.
31.4 Complementary, Alternative, and Integrative Approaches in Mental Health Care

H. Richard Lamb, M.D. Professor, Department of Psychiatry and the Behavioral Sciences, Keck School of Medicine of University of Southern California, Los Angeles, California.
58.8 Criminalization of Persons with Serious Mental Illness

Graciana Lapetina, M.D. Clinical Assistant Professor, Department of Psychiatry, New York University School of Medicine; Attending Psychiatrist, Department of Psychiatry, Bellevue Hospital Center, New

York, New York.
31.7 Physician and Medical Student Mental Health

Eugene M. Laska, Ph.D. Professor, Department of Psychiatry, New York University School of Medicine, New York, New York.
5.2 Statistics and Experimental Design

Tarannum Lateef, M.D. Assistant Professor, Departments of Neurology and Pediatrics, Children's National Health System, Washington, D.C.; Pediatric Specialists of Virginia, Woodbridge, Virginia.
2.11 Neuropsychiatric Aspects of Headache

Megan E. Lavery, Psy.D. Clinical Health Psychologist, Bariatric and Metabolic Institute, Sleep Disorders Center, Cleveland Clinic, Cleveland, Ohio.
27.4 Psychosomatic Medicine: Obesity

Helen Lavretsky, M.D. Professor, Jane and Terry Semel Institute for Neuroscience and Human Behavior, Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
57.2b Complementary, Alternative, and Integrative Medicine in Geriatric Psychiatry

Lawrence W. Lazarus, M.D. Assistant Professor of Psychiatry, Rush Medical College Chicago, Illinois; Former Director, Geropsychiatric Fellowship Training Program; Private Practice, Santa Fe, New Mexico.
57.4h Individual Psychotherapy

Marguerite S. Lederberg, M.D. Psychiatry Attending Emeritus, Department of Psychiatry and Behavioral Sciences, Memorial Sloan Kettering Cancer Center, New York, New York.
27.8 Psycho-Oncology, 27.9 End-of-Life and Palliative Care

Francis S. Lee, M.D., Ph.D. Mortimer D. Sackler Professor, Sackler Institute for Developmental Psychobiology, Department of Psychiatry and Pharmacology, Weill Cornell Medical College, New York, New York.
1.7 Neurotrophic Factors

Kelly C. Lee, Pharm.D., M.A.S., B.C.P.P. Associate Professor of Clinical Pharmacy, Skaggs School of Pharmacy and Pharmaceutical Sciences; Psychiatric Pharmacy Specialist, Department of Pharmacy, University of California San Diego School of Medicine, La Jolla, California.
57.4d Antipsychotic Drugs

Susan W. Lehmann, M.D. Associate Professor, Department of Psychiatry and Behavioral Sciences; Clinical Director, Division of Geriatric Psychiatry and Neuropsychiatry, Johns Hopkins University School of Medicine, Baltimore, Maryland.
57.3g Delirium

Sean R. LeNoue, M.D. Addiction Psychiatry Fellow, Department of Psychiatry, University of Colorado School of Medicine; Child and Adolescent Psychiatrist, Children's Hospital Colorado and Denver Health Medical Center, Aurora, Colorado.
55.5 Adolescent Substance Use Disorders

Gregory B. Leong, M.D. Clinical Professor, Department of Psychiatry and Behavioral Sciences, Keck School of Medicine of University of Southern California, Los Angeles, California; Clinical Professor, Department of Psychiatry and Human Behavior, University of California, Irvine, School of Medicine, Irvine, California.
60.1 Clinical-Legal Issues in Psychiatry

Alan J. Lerner, M.D. Professor of Neurology, Case Western Reserve University School of Medicine; Director, Brain Health and Memory Center, Neurological Institute, University Hospitals Cleveland Medical Center, Cleveland, Ohio.
2.12 Neuropsychiatric Aspects of Neuromuscular Diseases

Molyn Leszcz, M.D. Professor and Vice Chair, Clinical, Department of Psychiatry, University of Toronto Faculty of Medicine; Psychiatrist-in-Chief, Mount Sinai Hospital, Toronto, Ontario, Canada.
57.4k Group Therapy

Adam B. Lewin, Ph.D. Associate Professor and Program Director, Departments of Pediatrics, University of South Florida Health Morsani College of Medicine, Tampa, Florida.
52.1 Obsessive-Compulsive Disorder in Childhood

Bradley Lewis, M.D., Ph.D. Associate Professor, Gallatin School of Individualized Study; Affiliated Appointments in the Department of Psychiatry and the Department of Cultural Analysis, New York University School of Medicine, New York, New York.
33.12 Narrative Psychiatry

David A. Lewis, M.D. Professor and Chair; Distinguished Professor of Psychiatry Neuroscience and Thomas Detre Professor and Chair, Department of Psychiatry; Medical Director, Director of Research Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.
1.2 Functional Neuroanatomy

Dorothy Otnow Lewis, M.D. Clinical Professor (Retired), Yale Child Study Center, Yale University School of Medicine, New Haven, Connecticut.
28.2 Adult Antisocial Behavior, Criminality, and Violence

Stephen F. Lewis, M.D. Professor, Department of Psychiatry, University of New Mexico School of Medicine, Albuquerque, New Mexico.
12.2 Phenomenology of Schizophrenia

Roberto Lewis-Fernández, M.D. Professor of Psychiatry, Columbia University College of Physicians and Surgeons; Director, New York State Center of Excellence for Cultural Competence; Director, Hispanic Treatment Program, New York State Psychiatric Institute, New York, New York.
20 Dissociative Disorders, 29.1 Cultural Concepts of Distress, 29.2 Cultural Assessments in Psychiatry

Ma'ayan E. Lieberman, M.D. Clinical Associate, Tufts University School of Medicine, St. Elizabeth's Medical Center, Boston, Massachusetts.
57.6d Gender Issues

Scott O. Lilienfeld, Ph.D. Samuel Candler Dobbs Professor, Department of Psychology, Emory University

School of Medicine, Atlanta, Georgia.
4.2 Neurocentrism: Implications for Addiction and the Courtroom

Benjamin Liptzin, M.D. Professor of Psychiatry and Tufts Chair at Baystate, Tufts University School of Medicine, Boston, Massachusetts; Chairman, Department of Psychiatry, Baystate Medical Center, Springfield, Massachusetts.
57.3g Delirium

Rodolfo R. Llinás, M.D., Ph.D. University Professor, Professor of Physiology and Neuroscience, New York University School of Medicine, New York, New York.
3.5 Consciousness and Dreaming from a Pathophysiological Perspective: The Thalamocortical Dysrhythmia Syndrome

Brett Lloyd, Jr., M.D., Ph.D. Assistant Professor, Associate Residency Program Director, Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine; McGaw Medical Center, Chicago, Illinois.
34.20 Melatonin Receptor Agonists

Christine Lochner, Ph.D. Associate Professor, SU/UCT Medical Research Council Unit on Anxiety and Stress Disorders, Department of Psychiatry, University of Stellenbosch Faculty of Medicine and Health Sciences, Cape Town, South Africa.
15 Obsessive-Compulsive and Related Disorders

David Loewenstein, Ph.D. Professor of Psychiatry and Behavioral Sciences, University of Miami Lenonard M. Miller School of Medicine, Miami, Florida.
58.6 Psychiatric Rehabilitation

Richard J. Loewenstein, M.D. Clinical Professor, Department of Psychiatry, University of Maryland School of Medicine; Medical Director, The Trauma Disorders Program, Sheppard Pratt Health System, Baltimore, Maryland.
20 Dissociative Disorders

Carlos López-Jaramillo, M.D., Ph.D. Faculty of Medicine, Chairman, Coordinator Research Group in Psychiatry GIPSI, Department of Psychiatry, Universidad de Antioquia Facultad de Medicina; Director, Mood Disorders Program, Hospital Universitario de San Vicente Fundación, Medellín, Antioquia, Colombia.
9.2 The Classification of Mental Disorders in the International Classification of Diseases (ICD-11)

Loredana Lucarelli, Psy.D. Professor of Dynamic Psychology, Department of Pedagogy, Psychology, Philosophy, Università degli Studi di Cagliari Facoltà di Medicina e Chirurgia, Monserrato; Tutor Professorm, Department of Dynamic and Clinical Psychology, Università degli Studi di Roma 'La Sapienza' Facoltà di Medicina e Psicologia, Rome, Italy.
47 Feeding and Eating Disorders of Infancy and Early Childhood

Melanie Lucas, Ph.D. Student, Clinical Psychology with Health Emphasis, Ferkauf Graduate School of Psychology, Albert Einstein College of Medicine of Yeshiva University, Bronx, New York; Division of Experimental Therapeutics, New York State Psychiatric Institute, New York, New York.
34.36 Brain Stimulation Methods

Sean X. Luo, M.D., Ph.D. Leon Levy Fellow, Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York.
11.9 Opioid Use and Related Disorders: From Neuroscience to Treatment

Amanda E. Lyall, Ph.D. Instructor, Department of Psychiatry, Harvard Medical School; Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts.
12.7 Structural Brain Imaging in Schizophrenia

Frank P. MacMaster, Ph.D. Associate Professor, Departments of Psychiatry and Pediatrics, University of Calgary; Scientific Director, Strategic Clinical Network for Addictions and Mental Health, Alberta Health Services, Calgary, Alberta.
38 Neuroimaging in Psychiatric Disorders of Childhood

Jeanne E. Maglione, M.D., Ph.D. Assistant Clinical Professor of Psychiatry, University of California San Diego School of Medicine, La Jolla, California; Primary Care Mental Health Integration Psychiatrist, VA San Diego Healthcare System, San Deigo, California.
57.3c Sleep-Wake Disorders in Older Adults

Mario Maj, M.D., Ph.D. Professor of Psychiatry; Chairman, Department of Psychiatry, Università degli Studi di Napoli Federico II Scuola di Medicina e Chirurgia, Naples, Italy.
62 World Aspects of Psychiatry

Alice R. Mao, M.D. Professor of Psychiatry and Behavioral Sciences, Texas Children's Hospital Child Psychiatry Autism and Developmental Disorders Clinic, Baylor College of Medicine, Houston, Texas.
55.12 Impact on Parents of Raising a Child with Psychiatric Illness and/or Developmental Disability

Stephen R. Marder, M.D. Daniel X. Freedman Professor of Psychiatry; Vice Chair for Education, Section on Psychosis, Jane and Terry Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine at University of California Los Angeles; Director, Veterens Affairs Desert Pacific Mental Illness Research, Education, and Clinical Center, Los Angeles, California.
34.28 Second-Generation Antipsychotics

Russell L. Margolis, M.D. Professor of Psychiatry and Neurology, Department of Psychiatry, Johns Hopkins University School of Medicine; Attending Physicians, Johns Hopkins Hospital and Johns Hopkins Bayview Hospital Baltimore, Maryland.
2.6 Neuropsychiatric Aspects of Movement Disorders

John J. Mariani, M.D. Associate Professor of Clinical Psychiatry, Division of Substance Abuse, Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York.
11.10 Sedative-, Hypnotic-, or Anxiolytic-Related Disorders

Humberto Marin, M.D. Professor of Psychiatry, Rutgers Robert Wood Johnson Medical School of Rutgers University, New Brunswick, Piscataway, New Jersey.
9.1 Present and Future of Classification Systems for Mental Disorders

John C. Markowitz, M.D. Professor of Clinical Psychiatry, Columbia University College of Physicians and Surgeons; Research Psychiatrist, New York State Psychiatric Institute, New York, New York.
13.5 Mood Disorders: Intrapsychic and Interpersonal Aspects

Charles R. Marmar, M.D. Lucius R. Littauer Professor; Chair, Department of Psychiatry, New York University School of Medicine; Director, Steven and Alexandra Cohen Center for the Study of Posttraumatic Stress and Traumatic Brain Injury, New York, New York.
17 Posttraumatic Stress Disorder

María J. Marquine, Ph.D. Assistant Professor, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
57.6g HIV and Aging

Laura Marsh, M.D. Professor, Departments of Psychiatry and Neurology, Baylor College of Medicine; Director, Mental Health Care Line, Michael E. DeBakey Veterans Affairs Medical Center, Houston, Texas.
2.6 Neuropsychiatric Aspects of Movement Disorders

Andrés Martin, M.D. Riva Ariella Ritvo Professor, Child Study Center, Yale University School of Medicine, New Haven, Connecticut.
54.7 Inpatient Psychiatric, Partial Hospital, and Residential Treatment for Children and Adolescents

Ardis C. Martin, M.D. Private Practice, Walnut Creek, California.
35.3 Adolescent Development

Caridad Ponce Martinez, M.D. Assistant Professor, Department of Psychiatry, University of Massachusetts Medical School, Worcester; Massachusetts Assistant Professor Adjunct, Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.
34.40 Treatment of Erectile Disorder

Christopher E. Mason, Ph.D. Associate Professor, Department of Physiology and Biophysics, Weill Cornell Medical College, New York, New York.
1.11 Genome, Transcriptome, and Proteome: The Molecular Genetics and Biochemistry Underlying the Neurobiology of Mental Disorders

Graeme F. Mason, Ph.D. Professor of Radiology & Biomedical Imaging and Psychiatry, Yale University School of Medicine, New Haven, Connecticut.
1.16 Nuclear Magnetic Resonance Imaging and Spectroscopy: Basic Principles and Recent Findings in Neuropsychiatric Disorders

Anu A. Matorin, M.D. Associate Professor, Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine; Medical Director, Psychiatry Emergency Center; Deputy Chief Psychiatry (Psychiatry ER, Psychiatry Inpatient and Psychiatry Consult/Liason Service), Department of Psychiatry, Ben Taub Hospital, Houston, Texas.
8 Clinical Manifestations of Psychiatric Disorders

Sarah N. Mattson, Ph.D. Professor, Department of Psychology; University of California San Diego School of Medicine, La Jolla, California; Director of Clinical Research, Center for Behavioral Teratology, Co-Director, Center for Clinical and Cognitive Neuroscience, San Diego, California.
2.13 Psychiatric Aspects of Child Neurology

Jeffrey E. Max, M.B.B.Ch. Professor, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla California; Director, Neuropsychiatric Research, Rady Children's Hospital, San Diego,

California.
2.13 Psychiatric Aspects of Child Neurology

Molly M. McCarthy, D.O. Forensic Psychiatry Fellow, Department of Psychiatry, University of Miami Leonard M. Miller School of Medicine/Jackson Hospital, Miami, Florida.
34.12 Buspirone

Shawn M. McClintock, Ph.D. Adjunct Associate Professor, Duke University School of Medicine, Durham, North Carolina; Associate Professor, Department of Psychiatry and Behavioral Sciences, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas.
57.4f Electroconvulsive Therapy and Other Neurostimulation Treatments

James T. McCracken, M.D. Joseph Campbell Professor of Child Psychiatry; Director of the Division of Child and Adolescent Psychiatry, Department of Psychiatry and Biobehavioral Sciences, Jane and Terry Semel Institute of Neuroscience and Human Behavior, David Geffen School of Medicine at University of California Los Angeles, California.
37 Genetics in Child Psychiatry

Robert R. McCrae, Ph.D. Editorial Board, Journal of Personality, Gloucester, Massachusetts.
6.4 Approaches Derived from Philosophy and Psychology

Sarah T. McCutcheon, M.D. Assistant Professor, Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, Iowa.
2.2 Neuropsychiatric Aspects of Cerebrovascular Disorders

James J. McGough, M.D., M.S. Professor of Clinical Psychiatry and Biobehavioral Sciences, Jane and Terry Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine at University of California Los Angeles; Attending Physician, Resnick Neuropsychiatric Hospital at UCLA, Los Angeles, California.
45.2 Adult Manifestations of Attention-Deficit/Hyperactivity Disorder

John S. McIntyre, M.D. Clinical Professor of Psychiatry, University of Rochester School of Medicine and Dentistry; Past Chair, Department of Psychiatry and Behavioral Health, Unity Health System, Rochester, New York.
7.1 Psychiatric Interview, History, and Mental Status Examination of the Adult Patient, 7.3 Practice Guidelines in Psychiatry

Roger S. McIntyre, M.D. Professor of Psychiatry and Pharmacology, University of Toronto Faculty of Medicine; Executive Director, Brain and Cognition Discovery Foundation, Head, Mood Disorders Psychopharmacology Unit, University Health Network, Toronto, Ontario, Canada.
34.5 β-Adrenergic Receptor Antagonists, 34.8 Antihistamines

John R. McQuaid, Ph.D. Professor, Department of Psychiatry, University of California San Francisco School of Medicine; San Francisco VA Medical Center, San Francisco, California.
13.8 Mood Disorders: Psychotherapy, 57.4i Cognitive-Behavioral Therapy

Aimee L. McRae-Clark, Pharm.D. Professor of Psychiatry and Behavioral Sciences, Professor of Neuroscience, Medical University of South Carolina, Charleston, South Carolina.

34.24 Opioid Maintenance Treatments: Methadone and Buprenorphine

Morris Meisner, Ph.D. Research Associate Professor, Department of Psychiatry, New York University School of Medicine, New York, New York; Research Scientist, Statistics and Services Research, Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York.
5.2 Statistics and Experimental Design

†**William W. Meissner, S.J., M.D.** University Professor of Psychoanalysis, Boston College, Training and Supervising Analyst Emeritus, Psychoanalytic Society of New England East, Inc., Boston, Massachusetts.
6.1 Classical Psychoanalysis

Darlene Susan Melchitzky, M.S. Laboratory Director and Research Associate, Department of Biology, Mercyhurst University, Erie, Pennsylvania; Research Principal, Translational Neuroscience Program, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.
1.2 Functional Neuroanatomy

Mario F. Mendez, M.D., Ph.D. Professor of Neurology and Psychiatry and Biobehavioral Sciences, Director, Behavioral Neurology Program, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
2.4 Neuropsychiatric Aspects of Epilepsy

Steven Mennerick, Ph.D. Professor, Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri.
1.10 Cellular and Synaptic Basis of Neural Signaling

James R. Merikangas, M.D. Clinical Professor of Psychiatry and Behavioral Sciences, George Washington University School of Medicine and Health Sciences, Washington, D.C.
2.11 Neuropsychiatric Aspects of Headache

Kathleen Ries Merikangas, Ph.D. Senior Investigator, Chief, Genetic Epidemiology, Research Branch, Division of Intramural Research Programs at the National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland.
2.11 Neuropsychiatric Aspects of Headache, 13.2 Mood Disorders: Epidemiology, 14.1 Anxiety Disorders: Introduction and Overview, 14.3 Epidemiology of Anxiety Disorders

Stephanie E. Meyer, Ph.D. Child Psychologist, Inc., Los Angeles, California.
51.2 Early-Onset Bipolar Disorder

Edwin Jens Mikkelsen, M.D. Associate Professor of Psychiatry, Harvard Medical School, Boston, Massachusetts.
49 Elimination Disorders

Andrew H. Miller, M.D. William P. Timmie Professor of Psychiatry and Behavioral Sciences, Vice Chair of Research, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine; Co-Leader, Cancer Prevention and Control, Winship Cancer Institute, Atlanta, Georgia.
1.13 Immune System and Central Nervous System Interactions

Megan A. Miller, M.S. Graduate Student, Department of Psychology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

1.14 Chronobiology

Barbara L. Milrod, M.D. Professor of Psychiatry, Weill Cornell Medical College; Attending Psychiatrist, New York Presbyterian Hospital, New York, New York.
13.5 Mood Disorders: Intrapsychic and Interpersonal Aspects

Alireza Minagar, M.D. Magale Professor and Chairman, Department of Neurology, Louisiana State University Health Sciences Center, Shreveport, Louisiana.
2.10 Neuropsychiatric Aspects of Prion Disease

Wendy G. Mitchell, M.D. Professor of Clinical Neurology, Department of Neurology, Keck School of Medicine of University of Southern California; Attending Physician, Division of Neurology, Children's Hospital Los Angeles, Los Angeles, California.
42 Developmental Coordination Disorder

Ramon Mocellin, M.B.B.S., M.Sc., M.Psych. (Med), F.R.A.N.Z.C.P. Consultant, Neuropsychiatry Unit, The Royal Melbourne Hospital, Parkville, Australia.
2.14 Neuropsychiatry of Neurometabolic and Neuroendocrine Disorders

Frederick Gerard Moeller, M.D. Professor of Psychiatry, Pharmacology and Toxicology, and Neurology; Director, C. Kenneth and Dianne Wright Center for Clinical and Translational Research and Institute for Drug and Alcohol Studies; Associate Vice President for Clinical Research, Virginia Commonwealth University School of Medicine, Richmond, Virginia.
24.1 Intermittent Explosive Disorder, 24.2 Pyromania, 24.3 Kleptomania

Ramin Mojtabai, M.D., Ph.D. Professor, Department of Mental Health, Bloomberg School of Public Health and Department of Psychiatry, Johns Hopkins University School of Medicine; Attending Psychiatrist, Community Psychiatry Program, Johns Hopkins Hospital, Baltimore, Maryland.
12.17 Other Psychotic Disorders

David J. Moore, Ph.D. Associate Professor, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
57.3a Assessment of Functioning

Raeanne C. Moore, Ph.D. Assistant Professor, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California, Research Health Science Specialist, VA San Diego Healthcare System, San Diego, California.
57.6g HIV and Aging

Timothy H. Moran, Ph.D. Paul R. McHugh Professor, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland.
1.23 Basic Science of Appetite

Joshua Chaim Morganstein, M.D. Associate Professor and Assistant Chair, Department of Psychiatry, F. Edward Hébert School of Medicine of Uniformed Services University of the Health Sciences; Staff Attending, Psychiatry Consultation Liaison Service, Walter Reed National Military Medical Center Bethesda, Maryland.
31.5 Disaster Psychiatry: Disasters, Terrorism, and War

James Morrison, M.D. Affiliate Professor, Department of Psychiatry, Oregon Health and Science University School of Medicine, Portland, Oregon.
59.2 Examining Psychiatrists and Other Mental Health Professionals

Eydie L. Moses-Kolko, M.D. Associate Professor, Department of Psychiatry, University of Pittsburgh School of Medicine, Western Psychiatric Institute and Clinic, Pittsburgh, Pennsylvania.
31.1 Psychiatry and Reproductive Medicine

†**David A. Mrazek, M.D.** Professor of Psychiatry and Pediatrics, Mayo Clinic College of Medicine; Chair, Psychiatry and Psychology, Mayo Clinic, Rochester, Minnesota.
55.9 Prevention of Psychiatric Disorders in Children and Adolescents

Patricia J. Mrazek, Ph.D. Consultant, Mayo Clinic, Rochester, Minnesota.
55.9 Prevention of Psychiatric Disorders in Children and Adolescents

Rodrigo A. Muñoz, M.D. Clinical Professor of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
59.2 Examining Psychiatrists and Other Mental Health Professionals

John D. Murray, Ph.D. Assistant Professor, Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.
1.29 Computational Modeling Approaches to Psychiatry

James W. Murrough, M.D. Assistant Professor of Psychiatry, Departments of Psychiatry and Neuroscience, Icahn School of Medicine at Mount Sinai; Attending Psychiatrist, Mount Sinai Hospital, New York, New York.
14.5 Anxiety Disorders: Neurobiology and Neuroscience

Srinivas B. Muvvala, M.D. Assistant Professor of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.
34.25 Opioid Receptor Antagonists: Naltrexone and Nalmefene

David Naimark, M.D. Clinical Professor of Psychiatry, University of California San Diego School of Medicine, La Jolla, California; Supervising Psychiatrist, Adult Forensics, County of San Diego, San Diego, California.
57.6b Forensic Aspects, 57.6e The Legal and Ethical Analysis of Elder Abuse

William E. Narrow, M.D. Associate Professor, Department of Psychiatry and Behavioral Sciences, University of New Mexico School of Medicine; Attending Psychiatrist and Medical Director of Community-Based Services, University of New Mexico Psychiatric Center, Albuquerque, New Mexico.
5.1 Quantitative and Experimental Methods in Psychiatry: Epidemiology

J. Craig Nelson, M.D. Leon J. Epstein Professor of Psychiatry, University of California San Francisco School of Medicine; Director of Geriatric Psychiatry, Langley Porter Psychiatry and the UCSF Weill Institute for Neurosciences, San Francisco, California.
34.33 Tricyclics and Tetracyclics

Charles B. Nemeroff, M.D., Ph.D. Leonard M. Miller Professor and Chairman, Director, Department of

Psychiatry and Behavioral Science, University of Miami Leonard M. Miller School of Medicine, Miami, Florida.
1.6 Neuropeptides: Biology, Regulation, and Role in Neuropsychiatric Disorders

Mark P. Nespeca, M.D. Clinical Professor of Neurosciences, University of California San Diego School of Medicine, La Jolla, California; Medical Director, Comprehensive Epilepsy Center and EEG Laboratory, Rady Children's Hospital, San Deigo, California.
2.13 Psychiatric Aspects of Child Neurology

Randolph M. Nesse, M.D. Professor of Life Sciences, Director, Center for Evolution and Medicine, Arizona State University, Tempe, Arizona; Professor Emeritus of Psychiatry, Professor Emeritus of Psychology, University of Michigan Medical School, Ann Arbor, Michigan.
4.1 Evolutionary Foundations for Psychiatric Research and Practice

Paul S. Nestadt, M.D. Psychiatric Epidemiology Fellow, Department of Mental Health, Johns Hopkins Bloomberg School of Public Health; Supervising Psychiatrist, Anxiety Clinic, Department of Psychiatry and Behavioral Sciences, Johns Hopkins Medical Institution, Baltimore, Maryland.
27.10 Stress and Psychiatry

John W. Newcomer, M.D. Professor of Integrated Medical Science and Vice Dean for Research and Innovation, Charles E. Schmidt College of Medicine at Florida Atlantic University, Boca Raton, Florida.
12.14 Medical Health in Schizophrenia

Jeffrey Newport, M.D. Professor of Psychiatry & Behavioral Sciences and Obstetrics & Gynecology, Women's Reproductive Mental Health, University of Miami Leonard M. Miller School of Medicine, Director, Women's Reproductive Mental Health; Medical Director, Health & Recovery Center, Jackson Behavioral Health Hospital Miami, Florida.
34.20 Melatonin Receptor Agonists

Dorian S. Newton, Ph.D. Private Practice, Psychologist and Psychoanalyst, Retired from Mills College, Oakland, California.
6.2 Erik H. Erikson

Sharon L. Nichols, Ph.D. Department of Neurosciences, University of California San Diego School of Medicine, La Jolla, California.
2.13 Psychiatric Aspects of Child Neurology

Frank John Ninivaggi, M.D. Assistant Clinical Professor, Department of Child Psychiatry, Yale School of Medicine; Associate Attending Physician, Department of Pediatric Psychiatry, Yale-New Haven Hospital, New Haven, Connecticut.
28.1 Malingering, 28.3 Borderline Intellectual Functioning and Academic or Educational Problem

Erika L. Nurmi, M.D., Ph.D. Assistant Professor of Psychiatry, Division of Child and Adolescent Psychiatry, Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at University of California Los Angeles, Jane and Terry Semel Institute of Neuroscience and Human Behavior; Acting Deputy Chief of Mental Health, Department of Psychiatry, Greater Los Angeles VA Hospital, Los Angeles, California.
37 Genetics in Child Psychiatry

Nora C. Oberfield, M.D. Child and Adolescent Psychiatry Instructor, New York-Presbyterian Hospital, Child and Adolescent Psychiatry Residency Training Program of Columbia and Cornell Universities, New York, New York.
7.11 Medical Error

Maryann Obiorah, Ph.D. Sedical Writer, Strategic Scientific Communications, Vaniam Group.
1.3 Neural Development and Neurogenesis

Ejine Okoroafor, M.D. Psychosomatic Medicine Fellow, New York Medical College, Westchester Medical Center, Valhalla, New York.
27.14 Psychiatric Care of the Burned Patient

Stephanie S. O'Malley, Ph.D. Professor and Deputy Chair for Clinical Research, Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.
34.25 Opioid Receptor Antagonists: Naltrexone and Nalmefene

David W. Oslin, M.D. Professor of Psychiatry, Perelman School of Medicine at University of Pennsylvania; Corporal Michael J. Crescenz VA Medical Center, Philadelphia, Pennsylvania.
57.3j Alcohol and Substance Abuse in Older Adults

Takeshi Otowa, M.D., Ph.D. Professor, Graduate School of Clinical Psychology, Teikyo Heisei University, Tokyo, Japan.
14.4 Genetics of Anxiety Disorders

Fred Ovsiew, M.D. Professor of Clinical Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Attending Physician, Northwestern Memorial Hospital, Chicago, Illinois.
2.1 The Neuropsychiatric Approach to the Patient

Michael J. Owen, M.D., Ph.D. Director, Division of Psychological Medicine and Clinical Neurosciences and MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff, Wales, United Kingdom.
12.4 Genetics of Schizophrenia

Karen M. Padron, M.D., Ph.D. Adjunct Assistant Professor, Department of Psychiatry and Behavioral Science, Emory University School of Medicine, Atlanta, Georgia.
34.20 Melatonin Receptor Agonists

Pongsatorn Paholpak, M.D. Lecturer/Instructor, Department of Psychiatry, Khon Kaen University Faculty of Medicine, Kohn Kaen, Thailand.
2.4 Neuropsychiatric Aspects of Epilepsy

Ken A. Paller, Ph.D. Professor and Director of the Cognitive Neuroscience Program; Fellow, Cognitive Neurology and Alzheimer's Disease Center; Member, Center for Sleep and Circadian Biology; Member, Center for Circadian and Sleep Medicine, Northwestern University Feinberg School of Medicine; Chicago, Illinois.
3.4 Biology of Memory

Barton W. Palmer, Ph.D. Professor, Department of Psychiatry, University of California San Diego School

of Medicine, La Jolla, California; Clinical Psychologist, Psychology Service, Veterans Affairs San Diego Healthcare System, San Deigo, California.
57.2e Neuropsychological Evaluation, 57.6c Ethical Issues, 57.6i Positive Psychiatry of Aging

Maryland Pao, M.D. Clinical Director and Deputy Scientific Director, Department of Health and Human Services, National Institute of Mental Health, National Institutes of Health, U.S. Public Health Service, Bethesda, Maryland.
55.3 Children's Reaction to Illness and Hospitalization

Nansook Park, Ph.D. Professor, Department of Psychology, Director, Michigan Positive Psychology Center, University of Michigan School of Medicine, Ann Arbor, Michigan; Distinguished Visiting Professor, University of Johannesberg, Johannesberg, South Africa.
33.13 Positive Psychology

Lawrence T. Park, M.D. Medical Director, Clinical Research Unit, Experimental Therapeutics and Pathophysiology Branch, Division of Intramural Research Programs, National Institute of Mental Health, Bethesda, Maryland.
34.30 Dopamine Receptor Agonists

Barbara L. Parry, M.D. Professor of Psychiatry, University of California San Diego School of Medicine, La Jolla, California; Director, Department of Psychiatry, Women's Mood Disorders Clinic, California.
31.1 Psychiatry and Reproductive Medicine

Marcella Pascualy, M.D. Associate Professor, Director, Psychosomatic Medicine Fellowship, University of Washington School of Medicine, Seattle, Washington.
57.3b Psychiatric Problems in the Medically Ill Geriatric Patient

Caroly Sue Pataki, M.D. Clinical Professor of Health Sciences, Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at University of California Los Angeles; Attending Psychiatrist, UCLA Division of Child and Adolescent Psychiatry, Los Angeles, California.
35.1 Introduction and Overview, 35.3 Adolescent Development, 36.2 Psychiatric Assessment of Preschool Children, 42 Developmental Coordination Disorder, 51.2 Early-Onset Bipolar Disorder

Godfrey D. Pearlson, M.D. Professor of Psychiatry and Neuroscience, Yale University School of Medicine, New Haven, Connecticut; Director, Olin Neuropsychiatry Center, Hartford, Connecticut.
12.11 Phenotypes of Psychosis

Holly L. Peay, Ph.D. Research Scientist, Social, Statistical and Environmental Sciences, RTI International, Research Triangle Park, North Carolina.
31.2 Genetic Counseling for Psychiatric Conditions

Dale Aaron Peeples, M.D. Associate Professor of Child and Adolescent Psychiatry, Medical College of Georgia at Augusta University; Assistant Training Director of Child and Adolescent Psychiatry, Augusta University Medical Center and Children's Hospital of Georgia, Augusta, Georgia.
55.1 Adoption and Foster Care

Gregory H. Pelton, M.D. Assistant Professor of Psychiatry and Neurology, Department of Psychiatry, Columbia University College of Physicians and Surgeons; Attending, Department of Geriatric Psychiatry,

New York State Psychiatric Institute, New York, New York.
57.2a Psychiatric Assessment of the Older Patient

Joseph V. Penn, M.D. Clinical Professor, Department of Psychiatry, University of Texas Medical Branch School of Medicine, Galveston; Director, Mental Health Services, University of Texas Medical Branch Correctional Managed Care, Conroe, Texas.
60.3 Correctional Psychiatry

Jessica Marie Perez, M.S. Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas.
12.6 Cellular and Molecular Neuropathology of Schizophrenia

Michele L. Pergadia, Ph.D. Associate Professor of Integrated Medical Science, Charles E. Schmidt College of Medicine of Florida Atlantic University, Boca Raton, Florida.
12.14 Medical Health in Schizophrenia

Bernice A. Pescosolido, Ph.D. Distinguished Professor, Department of Sociology, Indiana University School of Medicine, Indianapolis, Indiana.
58.7 A Sociocultural Framework for Mental Health and Substance Abuse Service Disparities

Jennifer N. Petras, M.D. Adjunct Faculty, Department of Child and Adolescent Psychiatry, Westchester Medical Center, Valhalla, New York.
34.4 α2-Adrenergic Receptor Agonists: Clonidine, Clonidine Extended Release, Guanfacine, and Guanfacine Extended Release

Megan L. Petrik, Ph.D. Assistant Professor, Department of Medicine, Division of General Internal Medicine, University of Minnesota Medical School; Clinical Health Psychologist, University of Minnesota Medical Center, Minneapolis, Minnesota.
33.15 Evaluation of Psychotherapy

John Piacentini, Ph.D. Professor of Psychiatry and Biobehavioral Sciences; Director, Child OCD, Anxiety, and Tic Disorders Program, Division of Child and Adolescent Psychiatry, Jane and Terry Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
52.1 Obsessive-Compulsive Disorder in Childhood

Katherine S. Pier, M.D. Instructor, Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York.
34.16 Disulfiram and Acamprosate

Sandra S. Pimentel, Ph.D. Assistant Professor, Department of Psychiatry and Behavioral Sciences, Albert Einstein College of Medicine of Yeshiva University; Chief, Child and Adolescent Psychology; Director, Anxiety and Mood Program; Associate Director of Psychology Training, Montefiore Medical Center, Bronx, New York.
54.3 Cognitive-Behavioral Psychotherapy for Children and Adolescents

Eric M. Plakun, M.D. Associate Medical Director, Director of Biopsychosocial Advocacy, Austen Riggs Center, Stockbridge, Massachusetts.

33.2 Psychoanalytic Treatment of Anxiety Disorders, Obsessive-Compulsive and Trauma-Related Disorders

Sirisha Pochareddy, Ph.D. Associate Research Scientist, Department of Neuroscience, Yale University School of Medicine, New Haven, Connecticut.
1.30 Functional Genomics of Human Brain Development

Carol Ann Podgorski, Ph.D. Associate Professor of Psychiatry, University of Rochester School of Medicine and Dentistry; Marriage and Family Therapist, Psychiatry, University of Rochester Medical Center, Rochester, New York.
57.4j Family Intervention and Therapy with Older Adults

Bruce G. Pollock, M.D., Ph.D. Professor of Psychiatry and Pharmacology, University of Toronto Faculty of Medicine; Vice President, Research Director, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario, Canada.
57.4a Treatment of Psychiatric Disorders: General Principles

Maurizio Pompili, M.D., Ph.D. Associate Professor of Psychiatry, Director, Psychiatry Residency Training Program, Department of Neurosciences, Mental Health, and Sensory Organs; Director, Suicide Prevention Center, Sant'Andrea Hospital, Università degli Studi di Roma 'La Sapienza' Facoltà di Medicina e Psicologia, Rome, Italy.
13.6 Mood Disorders: Suicidal Behavior

Harrison G. Pope, Jr., M.D. Professor of Psychiatry, Harvard Medical School, Boston, Massachusetts; Director, Biological Psychiatry Laboratory, McLean Hospital, Belmont, Massachusetts.
11.11 Anabolic-Androgenic Steroid-Related Disorders

Robert M. Post, M.D. Clinical Professor of Psychiatry, George Washington University School of Medicine and Health Sciences, Washington, D.C.
34.14 Carbamazepine, 34.34 Valproate

Seth Powsner, M.D. Professor, Departments of Psychiatry and Emergency Medicine, Yale University School of Medicine, New Haven, Connecticut.
19 Factitious Disorder

Leslie S. Prichep, Ph.D. Adjunct Professor, Department of Psychiatry, New York University School of Medicine, New York, New York.
7.7 Principles and Applications of Quantitative Electroencephalography in Psychiatry

Jennifer Clare Prins, M.D. Graduate, Department of Psychiatry, University of Pittsburgh School of Medicine; Board Certified Adult Psychiatrist, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.
10.5 Other Cognitive and Mental Disorders due to Another Medical Condition

Ignacio Provencio, Ph.D. Professor, Department of Biology, University of Virginia School of Medicine, Charlottesville, Virginia.
1.14 Chronobiology

Joan Prudic, M.D. Professor of Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York.

34.35 Electroconvulsive Therapy

Andrés J. Pumariega, M.D. Professor and Chair, Department of Psychiatry, Cooper Medical School of Rowan University; Chief, Department of Psychiatry, Cooper University Health, Camden, New Jersey.
54.8 Community-Based Treatment

Davin K. Quinn, M.D. Associate Professor, Department of Psychiatry, University of New Mexico School of Medicine; Chief, Division of Behavioral Health Consultation and Integration, Director, University of New Mexico VA Psychosomatic Medicine Fellowship, Albuquerque, New Mexico.
19 Factitious Disorder

Nicole Quiterio, M.D. Medical Director, Bay Area Children's Association; The Chester M. Pierce MD Division of Global Psychiatry, Oakland, California.
54.4 Group Psychotherapy

Mark Ragins, M.D. Psychiatrist, Counseling and Psychological Services/Student Health Services, California State University, Long Beach, California.
12.15 Recovery in Schizophrenia

Charles L. Raison, M.D. Mary Sue and Mike Shannon Chair for Healthy Minds, Wisconsin Department of Children and Families; Professor, Human Development and Family Studies, School of Human Ecology and Department of Psychiatry, School of Medicine and Public Health, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin.
1.13 Immune System and Central Nervous System Interactions

Seethalakshmi Ramanathan, M.B.B.S., D.P.M. Clinical Assistant Professor, State University of New York Upstate Medical University; Staff Psychiatrist, Hutchings Psychiatric Center, Syracuse, New York.
57.2g Genetics of Late-Life Neurodegenerative Disorders

Divya Ramesh, Ph.D. Research Associate, Center for Advancement in Managing Pain, University of Connecticut School of Nursing, Farmington, Connecticut.
24.1 Intermittent Explosive Disorder, 24.2 Pyromania, 24.3 Kleptomania

Sonal Rana, M.D. Private Practice, Houston, Texas.
54.4 Group Psychotherapy

Mark M. Rasenick, Ph.D. Distinguished Professor of Physiology & Biophysics and Psychiatry, Director, Biomedical Neuroscience Training Program, University of Illinois College of Medicine; Research Service, Jesse Brown VAMC, Chicago, Illinois.
34.1 General Principles of Psychopharmacology

Natalie L. Rasgon, M.D., Ph.D. Professor, Psychiatry and Behavioral Sciences; Director of the Center for Neuroscience in Women's Health, Stanford University School of Medicine, Stanford, California.
27.7 Endocrine and Metabolic Disorders

Lakshmi Ravindran, M.D. Assistant Professor, Department of Psychiatry, Faculty of Medicine, Staff Psychiatrists, Mood and Anxiety Ambulatory Services, Centre for Addiction and Mental Health, University of Toronto Faculty of Medicine, Toronto, Ontario, Canada.

14.8 Somatic Therapies for Anxiety Disorders

Rebecca G. Reed, Ph.D. Postdoctoral Fellow, Department of Psychology, University of Kentucky College of Medicine, Lexington, Kentucky.
1.13 Immune System and Central Nervous System Interactions

Ulrich Reininghaus, Ph.D. Associate Professor, Department of Psychiatry and Psychology, School for Mental Health and Neuroscience, Maastricht University, Maastricht, Netherlands; Visiting Lecturer, Department of Health Service and Population Research, Centre for Epidemiology and Public Health, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK.
12.5 The Clinical Epidemiology of Schizophrenia

David C. Rettew, M.D. Associate Professor of Psychiatry and Pediatrics, University of Vermont Larner College of Medicine; Program Director, Child and Adolescent Psychiatry Fellowship; Director, Pediatric Psychiatry Clinic, University of Vermont Medical Center, Burlington, Vermont.
39 Temperament: Risk and Protective Factors for Child Psychiatric Disorders

Victor I. Reus, M.D. Professor, Department of Psychiatry, University of California San Francisco School of Medicine; Attending Physician, Langley Porter Hospital and Clinics, San Francisco, California.
1.12 Psychoneuroendocrinology

Sarah J. Richerson, Ph.D. Deputy Head, Neuroscience Center of Excellence, QuintilesIMS, Chicago, Illinois.
34.2 Drug Development and Approval Process in the United States

Paula D. Riggs, M.D. Professor and Director, Division of Substance Dependence, Department of Psychiatry, University of Colorado School of Medicine, Aurora, Colorado.
55.5 Adolescent Substance Use Disorders

Zoltán Rihmer, M.D., Ph.D., D.Sc. Professor of Psychiatry, Department of Clinical and Theoretical Mental Health, Semmelweis University Faculty of Medicine; Director, Laboratory for Suicide Research and Prevention, National Institute of Psychiatry and Addictions, Budapest, Hungary.
13.2 Mood Disorders: Epidemiology, 13.6 Mood Disorders: Suicidal Behavior

Thomas Ritz, Ph.D. Professor, Department of Psychology, Dedman College of Humanities and Sciences at Southern Methodist University, Dallas, Texas.
27.5 Respiratory Disorders

Laura Weiss Roberts, M.D. Katharine Dexter McCormick and Stanley McCormick Memorial Professor and Chairman, Department of Psychiatry and Behavioral Sciences, Stanford School of Medicine; Chief, Psychiatry Service; Attending Physician, Stanford Hospital and Clinics Stanford, California.
60.2 Ethics in Psychiatry

Robert G. Robinson, M.D. Professor, Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, Iowa.
2.2 Neuropsychiatric Aspects of Cerebrovascular Disorders

Xiomara Rocha-Cadman, M.D. Instructor in Psychiatry, Department of Psychiatry, Weill Cornell Medical College; Assistant Attending, Memorial Sloan Kettering Cancer Center, New York, New York.

27.9 End-of-Life and Palliative Care

Carolyn I. Rodriguez, M.D., Ph.D. Assistant Professor, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine; Staff Psychiatrist, Department of Psychiatry, VA Palo Alto Health Care System, Palo Alto, California.
57.3d Anxiety Disorders

Marcus A. Rodriguez, M.A., M.S. Doctoral Candidate, Department of Psychology and Neuroscience, Duke University School of Medicine, Durham, North Carolina.
33.9 Dialectical Behavior Therapy

Kathryn A. Roecklein, Ph.D. Assistant Professor, Department of Psychology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.
1.14 Chronobiology

Graham A.W. Rook, B.Chir., M.D. Emeritus Professor, of Medical Microbiology, Centre for Clinical Microbiology, Department of Infection, University College London School of Medicine, London, United Kingdom.
1.13 Immune System and Central Nervous System Interactions

Alasdair G. Rooney, M.D. MRC/MRF PsySTAR Fellow, Division of Psychiatry, University of Edinburgh, Specialty Registrar (Psychiatry), Royal Edinburgh Hospital, NHS Lothian, Edinburgh, Scotland.
2.3 The Neuropsychiatry of Brain Tumors

David R. Rosenberg, M.D. Professor and Chair, Department of Psychiatry and Behavioral Neurosciences, Wayne State University School of Medicine; Chief, Department of Child Psychiatry and Psychology, Children's Hospital of Michigan; Psychiatrist-in-Chief, Detroit Medical Center, Detroit, Michigan.
38 Neuroimaging in Psychiatric Disorders of Childhood

M. Zachary Rosenthal, Ph.D. Associate Professor, Director, Adult Outpatient Psychiatric and Behavioral Services, Department of Psychiatry and Behavioral Sciences, Department of Psychology and Neuroscience, Duke University School of Medicine, Durham, North Carolina.
33.9 Dialectical Behavior Therapy

Stephen Ross, M.D. Associate Professor of Psychiatry, Child and Adolescent Psychiatry, Oral & Maxillofacial Pathology, Radiology and Medicine, New York University School of Medicine; Senior Director, Department of Psychiatry, Division of Alcoholism & Drug Abuse, Director, Bellevue Hospital Opioid Overdose Prevention Program, Bellevue Hospital Center; Director, Addiction Psychiatry, Tisch Hospital; Director, Addiction Psychiatry Fellowship; Director, Experimental Therapeutics Research Laboratory; Director, Psychedelic Research Group; Principal Investigator, Psilocybin Cancer Project; Principal Investigator, Religious Professionals Study; Co-Principal Investigator, Psilocybin Alcoholism Project, NYU Langone Medical Center New York, New York.
11.6 Hallucinogen-Related Disorders, 33.17 Hallucinogen-Assisted Psychotherapy

Anthony L. Rostain, M.D. Professor, Psychiatry and Pediatrics, Perelman School of Medicine at University of Pennsylvania, Philadelphia, Pennsylvania.
54.2 Brief Psychotherapies for Childhood and Adolescence

Bryan L. Roth, M.D., Ph.D. Michael Hooker Distinguished Professor of Pharmacology, University of North Carolina School of Medicine, Chapel Hill, North Carolina.
1.9 Intraneuronal Signaling

Robert H. Roth, Ph.D. Professor of Pharmacology and Psychiatry, Department of Psychiatry, Neuropsychopharmacology Research Unit, Yale School of Medicine, New Haven, Connecticut.
1.4 Biogenic Amine Neurotransmitters

Stefan B. Rowny, M.D. Assistant Professor of Clinical Psychiatry, Department of Psychiatry, Columbia University College of Physicians and Surgeons; Division of Experimental Therapeutics, Department of Psychiatry, New York State Psychiatric Institute, New York, New York.
34.36 Brain Stimulation Methods

Peter Roy-Byrne, M.D. Professor of Psychiatry, University of Washington School of Medicine, Seattle, Washington.
14.2 Clinical Features of the Anxiety Disorders

David R. Rubinow, M.D. Assad Meymandi Distinguished Professor and Chair, Department of Psychiatry, University of North Carolina School of Medicine, University of North Carolina Hospitals, Chapel Hill, North Carolina.
34.39 Reproductive Hormonal Therapy: Theory and Practice

James K. Ruffle, B.Sc., M.B.B.S. Honorary Clinical Research Fellow, Centre for Neuroscience and Trauma, Blizard Institute, Wingate Institute of Neurogastroenterology, Barts and the London School of Medicine and Dentistry, Queen Mary University of London; Academic Foundation Programme Doctor, Acute Assessment Unit, Whipps Cross University Hospital, Barts Health NHS Trust, London, United Kingdom.
27.3 Gastrointestinal Disorders

Pedro Ruiz, M.D. Clinical Professor, Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine Houston, Texas.
8 Clinical Manifestations of Psychiatric Disorders, 29.1 Cultural Concepts of Distress, 62 World Aspects of Psychiatry

Samir A. Sabbag, M.D. Assistant Professor of Clinical Psychiatry Department of Psychiatry and Behavioral Sciences; Associate Program Director, Psychiatry Residency Training Program; Director Psychiatry Clerkship, University of Miami Leonard M. Miller School of Medicine, Miami, Florida.
34.15 Cholinesterase Inhibitors and Memantine

Joel I. Sadavoy, M.D. Professor of Psychiatry, Department of Psychiatry, University of Toronto Faculty of Medicine; Founder, Geriatric Psychiatry Program; Pencer Family Chair in Applied General Psychiatry; Head, Geriatric and Community Psychiatry; Director, Cyril and Dorothy, Joel And Jill Reitman Centre for Alzheimer's Support and Training, Mount Sinai Hospital, Toronto, Ontario, Canada.
57.4g Psychosocial Factors in Psychotherapy of the Elderly, 57.4h Individual Psychotherapy

Benjamin J. Sadock, M.D. Menas S. Gregory Professor of Psychiatry, Department of Psychiatry, New York University School of Medicine; Attending Psychiatrist, Tisch Hospital; Attending Psychiatrist, Bellevue Hospital Center, New York, New York.
7.2 Outline for a Psychiatric Examination, 7.11 Medical Error

Virginia A. Sadock, M.D. Professor of Psychiatry, New York University School of Medicine; Attending Psychiatrist, Tisch Hospital; Attending Psychiatrist, Bellevue Hospital Center, New York, New York.
21.1 Normal Human Sexuality and Sexual Dysfunctions

Samantha B. Saltz, M.D. Instructor, Child and Adolescent Psychiatry Chief Fellow, Department of Psychiatry and Behavioral Sciences, University of Miami Leonard M. Miller School of Medicine, Miami, Florida.
34.20 Melatonin Receptor Agonists

Carl Salzman, M.D. Professor of Psychiatry, Harvard Medical School; Massachusetts Mental Health Center; Beth Israel Deaconess Medical Center, Boston, Massachusetts.
57.4b Antidepressants and Mood Stabilizers

Gerard Sanacora, M.D., Ph.D. Professor of Psychiatry, Yale University School of Medicine; Director, Yale Depression Research Program; Co-Director, Yale-New Haven Hospital Interventional Psychiatry Service, New Haven, Connecticut.
1.16 Nuclear Magnetic Resonance Imaging and Spectroscopy: Basic Principles and Recent Findings in Neuropsychiatric Disorders

John Sargent, M.D. Professor of Psychiatry and Pediatrics, Tufts University School of Medicine; Director, Division of Child and Adolescent Psychiatry, Tufts Medical Center, Boston, Massachusetts.
54.5 Family Therapy

Sally L. Satel, M.D. Lecturer, Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.
4.2 Neurocentrism: Implications for Addiction and the Courtroom

Stephen M. Saunders, Ph.D. Professor of Psychology, Department of Psychology, Medical College of Wisconsin, Milwaukee, Wisconsin.
33.15 Evaluation of Psychotherapy

Radu V. Saveanu, M.D. Professor and Executive Vice Chair for Education, Department of Psychiatry and Behavioral Sciences, University of Miami Leonard M. Miller School of Medicine, Miami, Florida.
34.21 Mirtazapine, 58.6 Psychiatric Rehabilitation

Gauri N. Savla, Ph.D. Licensed Clinical Psychologist, Encinitas, California.
57.2e Neuropsychological Evaluation

Andrew J. Saxon, M.D. Professor, Director, Addiction Psychiatry Residency Program, University of Washington School of Medicine; Director, Center of Excellence in Substance Abuse Treatment and Education, VA Puget Sound Healthcare System, Seattle, Washington.
34.24 Opioid Maintenance Treatments: Methadone and Buprenorphine

Alan F. Schatzberg, M.D. Kenneth T. Norris, Jr. Professor of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California.
34.11 Bupropion, 34.38 Combination Pharmacotherapy

Jessica Lee Schleider, M.A. Ph.D. Candidate, Department of Psychology, Harvard University, Cambridge,

Massachusetts.
52.3 Separation Anxiety, Generalized Anxiety, and Social Anxiety

Steven C. Schlozman, M.D. Assistant Professor of Psychiatry, Harvard Medical School; Course Director, Psychiatry and Psychopathology, Health Science And Technology Program at Harvard Medical School and MIT; Staff Child Psychiatrist, Associate Director, Clay Center for Young Healthy Minds; Staff Psychiatrist and Consultant, Pediatric Solid Organ Transplant Team at Massachusetts General Hospital, Boston, Massachusetts.
54.9 The Treatment of Adolescents

Peter J. Schmidt, M.D. Chief, Section on Behavioral Endocrinology, Division of Intramural Research Programs at the National Institute of Mental Health, National Institutes of Health, National Institute of Health Center, Bethesda, Maryland.
34.39 Reproductive Hormonal Therapy: Theory and Practice

Lon S. Schneider, M.D. Professor of Psychiatry, Neurology, and Gerontology, Davis School of Gerontology, Keck School of Medicine of the University of Southern California, Los Angeles, California.
57.4e Antidementia Drugs

Edward J. Schreiber, T.E.P., D.D. Executive Director, Zerka T. Moreno Foundation, Easthampton, Massachusetts.
33.14 Psychodrama

Matthew Alan Schreiber, M.D., Ph.D. Acting Assistant Professor, Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington.
57.3b Psychiatric Problems in the Medically Ill Geriatric Patient

Marc A. Schuckit, M.D. Distinguished Professor, Department of Psychiatry, University of California San Diego School of Medicine; Director, Alcohol Medical Scholars Program; Field Editor, Journal of Studies on Alcohol and Drugs, La Jolla, California.
11.2 Alcohol-Related Disorders

Robert T. Schultz, Ph.D. RAC Professor of Psychology, Departments of Pediatrics and Psychiatry, Perelman School of Medicine at the University of Pennsylvania; Director, Center for Autism Research, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania.
44 Autism Spectrum Disorder and Social Communication Disorder

Mary E. Schwab-Stone, M.D. Associate Professor (Retired), Community Faculty, Yale Child Study Center, Yale University School of Medicine, New Haven, Connecticut.
36.1 Psychiatric Examination of the Infant, Child, and Adolescent

Kenneth M. Schwartz, M.D. Assistant Professor, Department of Psychiatry, University of Toronto Faculty of Medicine; Geriatric Psychiatrist, Baycrest Department of Psychiatry, Toronto, Ontario, Canada.
57.4k Group Therapy

Thomas W. Sedlak, M.D., Ph.D. Assistant Professor, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland.
1.8 Novel Neurotransmitters

Mariko Sembo, B.A. Research Assistant, MacMaster Lab, Department of Psychiatry, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada.
38 Neuroimaging In Psychiatric Disorders of Childhood

Nenad Sestan, M.D., Ph.D. Professor, Department of Neuroscience Genetics and Psychiatry, Section of Comparative Medicine, Program in Cellular Neuroscience, Neurodegeneration and Repair, Yale Child Study Center, and Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, Connecticut.
1.30 Functional Genomics of Human Brain Development

Daniel D. Sewell, M.D. Clinical Professor, Professor of Clinical Psychiatry, Geriatric Psychiatry Program; Director Emeritus, Department of Psychiatry, Division of Geriatric Psychiatry, Department of Psychiatry, University of California, San Diego; Medical Director and Director, Senior Behavioral Health Program, Department of Psychiatry, UC San Diego Health, San Diego, California.
57.6f Sexuality and Aging

Sandra B. Sexson, M.D. Professor and Chief, Division of Child, Adolescent, and Family Psychiatry, Department of Psychiatry and Health Behavior, Medical College of Georgia, Augusta, Georgia.
55.1 Adoption and Foster Care

Asim A. Shah, M.D. Professor and Vice Chair for Community Psychiatry, Department of Psychiatry, Family & Community Medicine; Site Director, Psychiatric Residency Education, Baylor College of Medicine; Chief of Psychiatry, Ben Taub Hospital/Harris Health System; Director, Mood Disorder Research Program at Ben Taub Hospital; Director, Community Behavioral Health Program; Executive Director, Psychotherapy Services, Neuropsychiatric Center, Houston, Texas.
8 Clinical Manifestations of Psychiatric Disorders

Arieh Y. Shalev, M.D. Barbara Wilson Professor of Psychiatry, Department of Psychiatry, New York University School of Medicine, New York, New York.
17 Posttraumatic Stress Disorder

Gabrielle Shapiro, M.D. Clinical Professor of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York.
35.2 Normal Child Development

Peter A. Shapiro, M.D. Professor of Psychiatry, Columbia University College of Physicians and Surgeons; Director, Fellowship Training Program in Psychosomatic Medicine; Director, Consultation-Liaison Psychiatry Service, New York-Presbyterian Hospital, Columbia University Medical Center, New York, New York.
27.2 Cardiovascular Disorders

Paul Shapshak, Ph.D. Adjunct Professor, Division of Infectious Diseases and International Medicine, Department of Internal Medicine, University of South Florida Health Morsani College of Medicine, Tampa, Florida.
2.10 Neuropsychiatric Aspects of Prion Disease

Amir Sharafkhaneh, M.D., Ph.D. Professor of Medicine, Section of Pulmonary, Critical Care and Sleep Medicine, Baylor College of Medicine; Staff Physician, Michael E. DeBakey VAMC, Houston, Texas.

23 Sleep Disorders

Neha Sharma, D.O. Assistant Professor, Department of Psychiatry, Tufts University School of Medicine; Training Director, Department of Psychiatry, Division of Child and Adolescent Psychiatry, Tufts Medical Center, Boston, Massachusetts.
54.5 Family Therapy

Jess P. Shatkin, M.D. Professor of Child & Adolescent Psychiatry and Pediatrics; Vice Chair for Education, New York University School of Medicine, Child Study Center at New York University Langone Medical Center, Bellevue Hospital Center, Rockland Children's Psychiatric Center, New York, New York.
55.13 Pediatric Sleep Disorders

Martha E. Shenton, Ph.D. Professor, Departments of Psychiatry and Radiology, Harvard Medical School, Boston, Massachusetts; Senior Scientist, Departments of Psychiatry and Radiology, Brigham and Women's Hospital Boston, Massachusetts; Health Scientist, VA Boston Healthcare System, Brockton, Massachusetts.
12.7 Structural Brain Imaging in Schizophrenia

Edward Shorter, Ph.D. Jason A. Hannah Professor of the History of Medicine, Professor of Psychiatry, Faculty of Medicine, University of Toronto Faculty of Medicine, Toronto, Ontario, Canada.
61 History of Psychiatry

Rebecca Shorter, M.A. Clinical Psychology Doctoral Student, Graduate Research Assistant, Department of Psychology, University of Tennessee-Knoxville, Knoxville, Tennessee.
33.3 Behavior Therapy

Rajeet Shrestha, M.D. Assistant Professor, Department of Psychiatry, Case Western Reserve University School of Medicine, Department of Psychiatry, University Hospitals Cleveland Medical Center, Cleveland, Ohio.
2.12 Neuropsychiatric Aspects of Neuromuscular Diseases

Carole E. Siegel, Ph.D. Research Professor, Department of Psychiatry, New York University School of Medicine, New York, New York.
5.2 Statistics and Experimental Design, 7.10 Cultural Competency in the Organization and Delivery of Care

Daniel J. Siegel, M.D. Clinical Professor, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California; Executive Director, Mindsight Institute, Santa Monica, California.
3.1 Contributions of the Psychological Sciences: Sensation, Perception, and Cognition

Linmarie Sikich, M.D. Associate Professor, Department of Psychiatry and Behavioral Sciences; Associate Director, Duke Center for Autism and Brain Development, Duke University School of Medicine, Durham, North Carolina.
53 Early-Onset Psychotic Disorders

John C. Silbereis, Ph.D. Postdoctoral Fellow, Department of Neuroscience, Yale University School of Medicine, New Haven, Connecticut.
1.30 Functional Genomics of Human Brain Development

Joseph R. Simpson, M.D., Ph.D. Associate Clinical Professor, Department of Psychiatry and Behavioral

Sciences, Keck School of Medicine of the University of Southern California, Los Angeles, California.
60.4 Neuroimaging, Psychiatry, and the Law

Samuel Justin Sinclair, Ph.D. Assistant Professor, Clinical Psychology Program, William James College, Newton, Massachusetts.
7.8 Psychiatric Rating Scales

Manpreet Kaur Singh, M.D. Assistant Professor of Psychiatry and Behavioral Sciences, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine; Director, Pediatric Mood Disorders Program, Department of Psychiatry and Behavioral Sciences, Lucile Packard Children's Hospital, Stanford School of Medicine, Stanford, California.
34.29 Central Nervous System Sympathomimetic Compounds

Patrick D. Skosnik, Ph.D. Assistant Professor, Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut; VA Connecticut Healthcare System, West Haven, Connecticut.
1.15 Applied Electrophysiology

Mark Slifstein, Ph.D. Associate Professor of Neurobiology (in Psychiatry), Department of Psychiatry, Columbia University College of Physicians and Surgeons; New York State Psychiatric Institute, New York, New York.
1.17 Radiotracer Imaging with Positron Emission Tomography and Single Photon Emission Computed Tomography: Fundamental Principles, Methodology, and Role in Neuropsychiatric Research

Gary W. Small, M.D. Professor of Psychiatry and Biobehavioral Science, Parlow-Solomon Professor on Aging, Director, Division of Geriatric Psychiatry; Director, UCLA Longevity Center, Jane and Terry Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
57.3f Alzheimer Disease and Other Neurocognitive Disorders

LalithKumar K. Solai, M.D. Assistant Professor of Psychiatry, University of Pittsburgh School of Medicine; Medical Director and Chief, Geriatric Psychiatry Service, Medical Director, Electroconvulsive Therapy Services, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.
10.2 Delirium

Adrian Natan Sondheimer, M.D. Clinical Assistant Professor of Psychiatry, Jacobs School of Medicine and Biomedical Sciences University at Buffalo, Buffalo, New York.
55.7 Ethical Issues in Child and Adolescent Psychiatry

Reneé M. Sorrentino, M.D. Instructor in Psychiatry, Harvard Medical School; Clinical Assistant Professor, Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts.
21.3 Paraphilic Disorders

Alma M. Spaniardi, M.D. Child and Adolescent Psychiatry Fellow, Residency Training Program of Columbia and Cornell Universities, New York Presbyterian Hospital, New York, New York.
45.1 Attention-Deficit/Hyperactivity Disorder

Traci J. Speed, M.D., Ph.D. Research Fellow, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland.

27.10 Stress and Psychiatry

Henry I. Spitz, M.D. Clinical Professor of Psychiatry, Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York.
33.5 Group Psychotherapy

Susan T. Spitz, A.C.S.W. Clinical Instructor in Psychiatry, Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York.
33.6 Family and Couples Therapy

Matthew W. State, M.D., Ph.D. Oberndorf Family Distinguished Professor of Psychiatry; Chair, Department of Psychiatry, Director, Langley Porter Psychiatric Institute, Weill Institute for Neurosciences, University of California San Francisco School of Medicine, San Francisco, California.
44 Autism Spectrum Disorder and Social Communication Disorder

Dan J. Stein, M.D., Ph.D. Professor and Chair, Department of Psychiatry and Mental Health, University of Cape Town Faculty of Health Sciences, Cape Town, South Africa.
15 Obsessive-Compulsive and Related Disorders

Murray B. Stein, M.D. Distinguished Professor of Psychiatry and Family and Preventive Medicine, Vice Chair for Clinical Research in Psychiatry, University of California San Diego School of Medicine, La Jolla, California; Staff Psychiatrist, VA San Diego Health System, San Diego, California.
14.8 Somatic Therapies for Anxiety Disorders, 57.3d Anxiety Disorders

Sharon Hoover Stephan, Ph.D. Associate Professor of Psychiatry, Co-Director, Center for School Mental Health, Division of Child and Adolescent Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland.
55.8 School Consultation

Eric Cameron Strain, M.D. Professor, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine; Director, Johns Hopkins Center for Substance Abuse Treatment and Research; Executive Vice Chair, Department of Psychiatry and Behavioral Sciences, Director, Behavioral Pharmacology Research Unit, Attending Psychiatrist, Department of Psychiatry, Johns Hopkins Bayview Medical Center, Baltimore, Maryland.
11.1 Substance Use Disorders: Introduction

Martin T. Strassnig, M.D. Associate Professor, Department of Integrated Medical Science; Associate Director, Clinical and Translational Research Unit, Charles E. Schmidt College of Medicine at Florida Atlantic University, Boca Raton, Florida.
12.14 Medical Health in Schizophrenia, 34.17 First-Generation Antipsychotics

Shannon F. Stromberg, M.D. Associate Professor, Department of Psychiatry, University of New Mexico School of Medicine, Albuquerque, New Mexico.
31.3 Physical and Sexual Abuse of Adults

Howard S. Sudak, M.D. Clinical Professor (Retired), Department of Psychiatry, Perelman School of Medicine at University of Pennsylvania, Philadelphia, Pennsylvania.
32.2 Suicide Treatment

Julianne K. Suojanen, D.O. Attending Psychiatrist, Montefiore Medical Center, Bronx, New York.
27.14 Psychiatric Care of the Burned Patient

Norman Sussman, M.D. Professor of Psychiatry, New York University School of Medicine, New York, New York.
34.27 Selective Serotonin Reuptake Inhibitors

Dragan M. Svrakic, M.D., Ph.D. Professor of Psychiatry, Washington University School of Medicine, St. Louis, Missouri.
26 Personality Disorders

Rex M. Swanda, Ph.D. Clinical Assistant Professor, Department of Psychiatry, University of New Mexico School of Medicine; Clinical Neuropsychologist, Director of Clinical Neuropsychology, Behavioral Healthcare Line, New Mexico VA Healthcare System, Albuquerque, New Mexico.
7.4 Clinical Neuropsychology and Intellectual Assessment of Adults

John A. Sweeney, Ph.D. Professor, Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati College of Medicine, Cincinnati, Ohio.
12.11 Phenotypes of Psychosis

Robert A. Sweet, M.D. Professor of Psychiatry and Neurology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.
10.1 Cognitive Disorders: Introduction, 10.3 Dementia

Zebulon Taintor, M.D. Adjunct Professor of Psychiatry, New York University School of Medicine, New York, New York.
7.9 Electronic Media in Psychiatry

Carol A. Tamminga, M.D. Professor and Chairman, Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas.
12.1 Introduction and Overview, 12.11 Phenotypes of Psychosis, 12.16 Psychosis as a Defining Dimension in Schizophrenia

Rosemary Tannock, Ph.D. Professor Emerita, Applied Psychology and Human Development, University of Toronto Faculty of Medicine; Senior Scientist, Neurosciences and Mental Health Research Program, The Hospital for Sick Children, Toronto, Ontario, Canada.
41 Specific Learning Disorder

Michael E. Thase, M.D. Professor of Psychiatry; Director of Mood and Anxiety Section, Department of Psychiatry, Perelman School of Medicine at the University of Pennsylvania, Corporal Michael J. Crescenz Veterans Affairs Medical Center, Philadelphia, Pennsylvania.
13.8 Mood Disorders: Psychotherapy, 13.9 Mood Disorders: Neurobiology, 34.21 Mirtazapine, 34.26 Selective Serotonin-Norepinephrine Reuptake Inhibitors

Margo L. Thienemann, M.D. Clinical Professor of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California.
54.4 Group Psychotherapy

Armin Paul Thies, Ph.D. Associate Clinical Professor (Retired), Yale Child Study Center, Yale University School of Medicine, New Haven, Connecticut.
36.1 Psychiatric Examination of the Infant, Child, and Adolescent

Makenzie E. Tonelli, B.A. Project Coordinator, Department of Psychiatry, University of Massachusetts Medical School, Worcester, Massachusetts.
11.8 Tobacco-Related Disorders

Giulio Tononi, M.D., Ph.D. Professor of Psychiatry, David P. White Chair in Sleep Medicine, Distinguished Professor of Consciousness Science, University of Wisconsin-Madison School of Medicine and Public Health; Director of the Wisconsin Institute for Sleep and Consciousness, Madison, Wisconsin.
1.22 Basic Science of Sleep

Karen E. Toth, Ph.D. Research Scientist, Seattle Children's Research Institute, Center for Child Health, Behavior, and Development, Seattle Children's Hospital; Department of Psychiatry and Behavioral Medicine, Seattle Children's Autism Center, Seattle, Washington.
40 Intellectual Disability

Diane E. Treadwell-Deering, M.D. Associate Professor, Departments of Psychiatry and Behavioral Sciences and Pediatrics, Center for Medical Ethics and Health Policy, Baylor College of Medicine; Co-Chief, Fragile X Clinic; Psychiatrist, Autism Center, Texas Children's Hospital, Houston, Texas.
55.12 Impact on Parents of Raising a Child with Psychiatric Illness and/or Developmental Disability

Glenn Jordan Treisman, M.D., Ph.D. Professor of Psychiatry and Behavioral Sciences, Professor of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland.
2.8 Neuropsychiatric Aspects of HIV Infection and AIDS

Manuel Trujillo, M.D. Clinical Professor of Psychiatry, New York University School of Medicine, New York, New York.
33.10 Intensive Short-Term Psychodynamic Psychotherapies

Susan Beckwitt Turkel, M.D. Associate Professor Departments of Psychiatry, Pathology, and Pediatrics, Keck School of Medicine of University of Southern California; Chief, Division of Psychiatry, Department of Pediatrics, Children's Hospital Los Angeles, Los Angeles, California.
55.3 Children's Reaction to Illness and Hospitalization

Allison K. Ungar, M.D. Clinical Assistant Professor of Psychiatry, New York University School of Medicine; Staff Psychiatrist, VA NY Harbor Healthcare System New York, New York.
34.16 Disulfiram and Acamprosate

Robert J. Ursano, M.D. Professor and Chair, Department of Psychiatry, Director, Center for the Study of Traumatic Stress, F. Edward Hébert School of Medicine, Uniformed Services University of the Health Sciences, Walter Reed National Military Medical Center, Bethesda, Maryland.
31.5 Disaster Psychiatry: Disasters, Terrorism, and War

Ipsit V. Vahia, M.D. Assistant Professor of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
57.3h Schizophrenia and Delusional Disorders, 57.6h Technology for Seniors

George E. Vaillant, M.D. Professor of Psychiatry, Harvard Medical School; Consultant, Department of Psychiatry, Massachusetts General Hospital Boston, Massachusetts.
3.6 Normality and Mental Health

Jim van Os, Ph.D. Professor and Chairman, Department of Psychiatry and Psychology, Maastricht University Medical Center, Maastricht, Netherlands.
12.5 The Clinical Epidemiology of Schizophrenia

Pieter Joost van Wattum, M.D. Associate Clinical Professor of Child Psychiatry and Psychiatry, Yale University School of Medicine; Chief Medical Officer, Clifford Beers Guidance Clinic, New Haven, Connecticut.
55.12 Impact on Parents of Raising a Child with Psychiatric Illness and/or Developmental Disability

Daniel Varon, M.D. Assistant Professor of Psychiatry; Associate Director of Geriatric Psychiatry Fellowship Program, University of Pittsburgh School of Medicine; Geriatric Inpatient Attending, Western Psychiatric Institute and Clinic, Pittsburgh, Pennsylvania.
10.4 Amnestic Disorders and Mild Cognitive Impairment

Yvonne Eleanor Vaucher, M.D. Clinical Professor of Pediatrics, University of California San Diego School of Medicine, La Jolla, California.
2.13 Psychiatric Aspects of Child Neurology

Dennis Velakoulis, DMedSci Professor, Neuropsychiatry Unit, Royal Melbourne Hospital, Melbourne Neuropsychiatry Centre, University of Melbourne Faculty of Medicine, Dentistry and Health Sciences, Parkville, Australia.
2.14 Neuropsychiatry of Neurometabolic and Neuroendocrine Disorders

Antonio Ventriglio, M.D., Ph.D. Department of Clinical and Experimental Medicine, Università degli Studi di Foggia Facoltà di Medicina e Chirurgia, Foggia, Italy.
62 World Aspects of Psychiatry

John T. Vernon, M.D. Assistant Professor, Department of Psychiatry, Virginia Commonwealth University School of Medicine, Richmond, Virginia.
34.32 Trazodone

Jeff Victoroff, M.D. Associate Professor of Clinical Neurology and Psychiatry, Keck School of Medicine of University of Southern California, Los Angeles, California.
30 The Neuropsychiatry of Human Aggression

Fred Robert Volkmar, M.D. Irving B. Harris Professor of Child Psychiatry, Pediatrics, and Psychology, Yale Child Study Center, Yale University School of Medicine, New Haven, Connecticut.
44 Autism Spectrum Disorder and Social Communication Disorder

Joel L. Voss, Ph.D. Assistant Professor, Departments of Medical Social Sciences, Neurology and Psychiatry, Northwestern University Feinberg School of Medicine, Chicago, Illinois.
3.4 Biology of Memory

Karen Dineen Wagner, M.D., Ph.D. Professor and Chair, Department of Psychiatry and Behavioral

Sciences, Titus Harris Chair, Harry K. Davis Professor, University of Texas Medical Branch, Galveston, Texas.
51.1 Depressive Disorders and Suicide

Harold J. Wain, Ph.D. Professor, Department of Psychiatry, F. Edward Hébert School of Medicine, Uniformed Services University of Health Sciences; Chief, Psychiatry Consultation Liaison Service, Department of Psychiatry, Walter Reed National Military Medical Center, Bethesda, Maryland.
33.4 Hypnosis

John T. Walkup, M.D. Adjunct Professor, Johns Hopkins School of Medicine, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; Professor and Vice Chair, Department of Psychiatry, Weill Cornell Medical College; Attending Psychiatrist and Director, Division of Child and Adolescent Psychiatry, Department of Psychiatry, New York-Presbyterian, New York, New York.
52.3 Separation Anxiety, Generalized Anxiety, and Social Anxiety

B. Timothy Walsh, M.D. Ruane Professor of Psychiatry, Columbia University College of Physicians and Surgeons; Director, Division of Clinical Therapeutics, New York State Psychiatric Institute, New York, New York.
22 Feeding and Eating Disorders

Mark Walterfang, Ph.D. Neuropsychiatry Unit, Royal Melbourne Hospital, Melbourne Neuropsychiatry Centre, University of Melbourne Faculty of Medicine, Dentistry & Health Sciences, Melbourne, Australia.
2.14 Neuropsychiatry of Neurometabolic and Neuroendocrine Disorders

Dora Wang, M.D. Clinical Associate Professor and Historian for the School of Medicine, Department of Family and Community Medicine, University of New Mexico Health Sciences Center; Telepsychiatrist, Butte County Department of Mental Health, Albuquerque, New Mexico.
19 Factitious Disorder

Po W. Wang, M.D. Clinical Professor, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California.
34.7 Anticonvulsants: Gabapentin, Topiramate, Tiagabine, Levetiracetam, Zonisamide, and Pregabalin, 34.18 Lamotrigine

Danuta Wasserman, M.D. Professor of Psychiatry and Suicidology, Karolinska Institute Faculty of Medicine, Department of Learning, Informatics, Medical Management and Ethics, National Centre for Suicide Research and Prevention of Mental Ill-Health, WHO Collaborating Centre for Research, Methods Development and Training in Suicide Prevention, Stockholm County Hospital, Centre for Health Economics and Informatics, Stockholm, Sweden.
32.1 Suicide: Overview and Epidemiology

Clarence Watson, J.D., M.D. Clinical Associate Professor in Psychiatry; Director, Forensic Psychiatry Fellowship Program, Department of Psychiatry, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania.
60.1 Clinical-Legal Issues in Psychiatry

Linda Edna Weinberger, Ph.D. Professor of Clinical Psychiatry and the Behavioral Sciences, Keck School of Medicine of University of Southern California; Chief Psychologist, University of Southern California

Institute of Psychiatry, Law, and Behavioral Science, Los Angeles, California.
58.8 Criminalization of Persons with Serious Mental Illness

Henry C. Weinstein, M.D. Clinical Professor, Department of Psychiatry, New York University School of Medicine, New York, New York.
60.3 Correctional Psychiatry

Roger D. Weiss, M.D. Professor of Psychiatry, Harvard Medical School, Boston, Massachusetts; Chief, Division of Alcohol and Drug Abuse, McLean Hospital, Belmont, Massachusetts.
11.3 Stimulant-Related Disorders

Chelsey A. Werchan, M.A. Graduate Research Assistant, Department of Psychology, Dedman College of Humanities and Sciences at Southern Methodist University, Dallas, Texas.
27.5 Respiratory Disorders

Virginia Wesson, M.D. Assistant Professor, Department of Psychiatry, University of Toronto Faculty of Medicine; Staff Psychiatrist, Sinai Health System, Ontario, Toronto, Canada.
57.4g Psychosocial Factors in Psychotherapy of the Elderly

Ellen M. Whyte, M.D. Assistant Professor, Department of Psychiatry, University of Pittsburgh School of Medicine; Medical Director, Psychiatric Services, Benedum Geriatric Center; Medical Director, Adult Integrated Behavioral Health Service, Western Psychiatric Institute and Clinic, Pittsburgh, Pennsylvania.
10.5 Other Cognitive and Mental Disorders due to Another Medical Condition

Timothy E. Wilens, M.D. Associate Professor of Psychiatry, Harvard Medical School; Chief, Division of Child and Adolescent Psychiatry; Co-Director, Center for Addiction Medicine, Massachusetts General Hospital, Boston, Massachusetts.
54.6 Pediatric Psychopharmacology

James M. Wilkins, M.D. Geriatric Psychiatry Fellow, McLean Hospital, Division of Geriatric Psychiatry, Belmont, Massachusetts.
57.3e Geriatric Mood Disorders

Katherine E. Williams, M.D. Clinical Associate Professor of Psychiatry and Behavioral Sciences, Center for Neuroscience in Women's Health, Stanford University School of Medicine, Stanford, California.
27.7 Endocrine and Metabolic Disorders

Ronald M. Wintrob, M.D. Clinical Professor of Psychiatry and Human Behavior, Warren Alpert Medical School of Brown University, Providence, Rhode Island.
4.3 Transcultural Psychiatry

Owen M. Wolkowitz, M.D. Professor of Psychiatry, University of California San Francisco School of Medicine, San Francisco, California.
1.12 Psychoneuroendocrinology

Dean F. Wong, M.D., Ph.D. Professor of Radiology, Psychiatry, Neurology, Neuroscience, and Environmental Health Sciences and Carey Business School, Johns Hopkins University School of Medicine, Baltimore, Maryland; Adjunct Professor of Neuroimaging, University of Copenhagen Faculty of Health

and Medical Sciences, Copenhagen, Denmark; Director, Section of High Resolution Brain PET Imaging, Nuclear Medicine, Johns Hopkins Medical Institutions, Baltimore Maryland.
12.9 Molecular Brain Imaging in Schizophrenia

Timothy P. Wuerz, D.O. Neuropsychiatry and Behavioral Neurology Fellow at University Hospitals, Department of Neurology, Neurological Institute Endowed Chair in Memory and Cognition, Case Western Reserve University School of Medicine; Beachwood, Ohio.
2.12 Neuropsychiatric Aspects of Neuromuscular Diseases

Lakshmi N. Yatham, M.B.B.S., M.R.C.Psych. (UK) Professor of Psychiatry, University of British Columbia Faculty of Medicine; Regional Head of Psychiatry, Vancouver Coastal Health and Providence Healthcare, Vancouver, British Columbia, Canada.
13.7 Mood Disorders: Pharmacological Treatment of Depression and Bipolar Disorders

Larry James Young, Ph.D. Professor, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia.
1.6 Neuropeptides: Biology, Regulation, and Role in Neuropsychiatric Disorders

Carlos A. Zarate, Jr., M.D. Chief, Experimental Therapeutics and Pathophysiology Branch, Section on the Neurobiology and Treatment of Mood Disorders, Division of Intramural Research Program, National Institute of Mental Health, Bethesda, Maryland.
34.30 Dopamine Receptor Agonists

Charles H. Zeanah, M.D. Professor of Psychiatry and Pediatrics; Mary Peters Sellars Polchow Chair in Psychiatry; Vice Chair for Child and Adolescent Psychiatry, Tulane University School of Medicine, New Orleans, Louisiana.
50.1 Reactive Attachment Disorder and Disinhibited Social Engagement Disorder

Douglas M. Ziedonis, M.D. Professor and Chair, Department of Psychiatry; Professor, Graduate School of Nursing & Graduate School of Biomedical Sciences; President, UMass Memorial Health Care, University of Massachusetts Medical School; Worcester, Massachusetts.
11.8 Tobacco-Related Disorders

Bonnie T. Zima, M.D. Professor-in-Residence, Department of Child and Adolescent Psychiatry, Associate Director, Jane and Terry Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, California.
55.10 Child Mental Health Services Research

Zvinka Z. Zlatar, Ph.D. Assistant Adjunct Professor, Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California.
57.2f Neuroimaging

Charles F. Zorumski, M.D. Samuel B. Guze Professor and Head of Psychiatry, Professor of Neuroscience Director of the Taylor Family Institute for Innovative Psychiatric Research, Director of the Center for Brain Research in Mood Disorders, Washington University School of Medicine; Chief of Psychiatry, Barnes-Jewish Hospital, St. Louis, Missouri.
1.10 Cellular and Synaptic Basis of Neural Signaling

Jon-Kar Zubieta, M.D., Ph.D. William H. and Edna Stimson Presidential Endowed Chair, Professor and Chair, Department of Psychiatry University of Utah School of Medicine; Psychiatrist-in-Chief, University Neuropsychiatric Institute, University of Utah Health Sciences Center, Salt Lake City, Utah.
1.20 Pain Systems: Interface with Affective and Motivational Mechanisms

---

*The views expressed in the chapter do not necessarily represent the views of the FDA or the United States.

*The contents of this publication do not represent the views of the Department of Veterans Affairs or the United States Government.

†Deceased

†Deceased

†Deceased

†Deceased

†Deceased

Preface

This marks the 50th anniversary since the first edition of the *Comprehensive Textbook of Psychiatry* was published in 1967. That edition consisted of one volume of about 600 pages with less than 100 contributors. The current 10th edition has over 5,000 pages in two volumes with over 650 contributors. This expansion represents the growth of psychiatry marked by an explosion of research and knowledge in neural science and in basic biological and psychological science.

On this 50th anniversary, it bears noting the role of Harold Kaplan, M.D., who recognized that psychiatry needed a textbook in the tradition of other great textbooks in medicine. He worked hard to achieve that goal, which he did well before his death in 1998. In recognition of that accomplishment this textbook still carries his name. We should also take note of our long and productive relationship with our publishers, beginning with Williams & Wilkins, Lippincott Williams & Wilkins, and Wolters Kluwer. Certain key persons stand out without whom the continuity of this textbook could not have occurred nor endured. These include Charles Reville, James Gallagher, Sarah Finnegan, and Charley Mitchell. There are others of course; but these four represent the very best in the world of publishing. They served as counselors, friends, and advisors, and at times, stern task masters. We take this opportunity to thank them once again.

Psychiatry, like all of medicine, is both art and science; but in the case of psychiatry, perhaps more art than science. Over the course of 50 years, however, the trajectory in psychiatry has veered toward more research and science, the results of which have been remarkable and the advances of knowledge robust. But psychiatry remains a "frontier science" and like all frontiers presents new and exciting areas yet to be explored.

The role of psychiatry in medicine has never been more important. Mental illness has increased and our understanding of its etiology and our ability to treat is not yet where we want it to be, even though both have advanced tremendously over this last half century. This observation was expressed by the Editors in the Preface to the first edition, part of which is reproduced below. It was prescient in many ways and as the reader will note, remains relevant today:

The underpinnings of clinical psychiatry have increased both in number and in strength; practitioners now have at their disposal large amounts of new knowledge in three major areas of behavioral science which must be incorporated within the existing theoretical and therapeutic knowledge in psychiatry. These three areas are the basic sciences such as neurophysiology and neuropharmacology; those social sciences which deal with individuals; and those which deal with the behavior of groups and of systems. Knowledge in those fields is not to be viewed as a kind of intellectual veneer. It is in fact a prerequisite for clinical competence. The perplexing array of drugs, for example, not only invites comprehension of their mode of action on the central nervous system, but also it demands such understanding to obviate the mechanical dispensing of preparations in the manner of a tradesman.

Similarly, the increasing awareness of the contributions of social factors to the development as well as the maintenance of mental illness calls for systematic training of the student in relevant aspects of the social sciences. The interplay between individual and social structure is especially important in disorders such as addiction, alcoholism, delinquency, and educational disabilities, problem areas that psychiatry can no longer avoid.

The *Comprehensive Textbook* is a "university without walls" whose aim is to educate all those who work with the mentally ill—psychiatrists and other physicians, psychologists, psychiatric social workers, psychiatric nurses, and mental health professionals from all fields. Its goal remains unchanged: to foster professional competence and to ensure the highest quality of care. The textbook has earned the reputation of being a thoroughly up-to-date encyclopedic compendium of psychiatric knowledge. As editors, we are extremely gratified by its wide acceptance and use both in this country and abroad. No other major textbook in psychiatry can lay claim to having such a long, consistent, and enriched publication history.

COMPREHENSIVE TEACHING SYSTEM

The textbook forms one part of a comprehensive system developed by us to facilitate the teaching of psychiatry and the behavioral sciences. At the head of the system is the *Comprehensive Textbook of Psychiatry*, which is global in depth and scope, designed for and used by psychiatrists, behavioral scientists, and all workers in the mental health field. *Synopsis of Psychiatry* is a relatively compact, highly modified, original, and current text useful for medical students, psychiatric residents, practicing psychiatrists, and mental health professionals. Two special texts, derived from *Synopsis*, are the *Concise Textbook of Clinical Psychiatry* and the *Concise Textbook of Child and Adolescent Psychiatry*. The former covers descriptions of all psychiatric disorders, including their diagnosis and treatment and the latter limits it to disorders of children and adolescents. Both books are useful for clinical clerks and psychiatric residents who need a succinct overview of the management of clinical problems. Another part of the system, *Study Guide and Self-Examination Review of Psychiatry*, consists of multiple-choice questions and answers; it is designed for students of psychiatry and for clinical psychiatrists who require a review of the behavioral sciences and general psychiatry in preparation for a variety of examinations. The questions are modeled after and consistent with the

format used by the American Board of Psychiatry and Neurology (ABPN), the National Board of Medical Examiners (NBME), and the United States Medical Licensing Examination (USMLE). Other parts of the system are the pocket handbooks: *Pocket Handbook of Clinical Psychiatry, Pocket Handbook of Psychiatric Drug Treatment, Pocket Handbook of Emergency Psychiatric Medicine,* and *Pocket Handbook of Primary Care Psychiatry*. These books cover the diagnosis and treatment of psychiatric disorders, psychopharmacology, psychiatric emergencies, and primary care psychiatry, respectively, and are designed and written to be carried in the pocket by clinical clerks and practicing physicians, whatever their specialty, to provide a quick reference. Finally, the *Comprehensive Glossary of Psychiatry and Psychology* provides simply written definitions for psychiatrists and other physicians, psychologists, students, other mental health professionals, and the public. Together, these 10 books create a multifaceted approach to the teaching, study, and learning of psychiatry.

Changes to the Tenth Edition

Adding new contributors and new sections to each edition is a hallmark of the *Comprehensive Textbook,* and this edition is no exception. New authors ensure a fresh approach to each topic and keep the textbook vital and current. The editors are deeply grateful to the more than 2,000 psychiatrists and behavioral scientists who contributed to previous editions, all of whom maintained the highest standards of scholarship. Many of those sections remain classics in the field and are accessible to the interested reader.

The editors especially wish to thank Herbert Pardes, M.D., for writing the *Foreword* to this textbook in which he comments on important issues facing the field, both now and in the future. Dr. Pardes served as Vice President for Health Sciences and Dean of the Faculty of Medicine at Columbia University College of Physicians and Surgeons and is one of this country's most distinguished psychiatrists and educators.

Classification of Mental Disorders. A fifth edition of the *American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders* was published in 2013, titled DSM-5. It contains the official nomenclature used by psychiatrists and other mental health professionals in the United States; the psychiatric disorders discussed in the textbook are consistent with and follow that nosology. Every section dealing with clinical disorders has been updated thoroughly and completely to include the revisions contained in DSM-5. The reader also will find DSM-5 tables and descriptions for every major mental disorders listed in that manual as it has been in each of our editions.

The DSM is the "law of the land" and, as mentioned previously, is the nomenclature used throughout this textbook; however, some clinicians and

researchers have reservations about various aspects of the DSM, which readers will find mentioned herein. As future editions of the DSM appear, this textbook, as always, will allow room for dissent before and especially after every new version appears. It will continue to provide a forum for discussion, evaluation, criticism, and disagreement, while duly acknowledging the official nomenclature.

ICD-10. Readers also should be aware of a parallel classification system developed by the World Health Organization (WHO) called the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10). There are textual differences between DSM and ICD, but according to treaties between the United States and the WHO, the diagnostic code numbers must be identical to ensure uniform reporting of national and international psychiatric statistics. ICD diagnoses and numerical codes are accepted by Medicare, Medicaid, and private insurance companies for reimbursement purposes in the United States.

Neural Sciences. The many advances in the neural sciences have been so vast that the reader will hardly recognize this area of the book since the last edition of this textbook was published. Almost every section has been updated and revised by new contributors and many new sections have been added. Several new sections deserve special mention: *Epigenetics*, which deals with the interaction of genes and the environment; *Microbiome*, which explores the importance of the microbiome in psychiatry; *Systems Neuroscience* which covers how neuroscience has broadened its base; *Human Connectome*, which examines the exciting new area of neuronal mapping and connectivity; *Computational Neuroscience and Psychiatry*, which explores how technology has expanded our knowledge of the brain; and *Neurodevelopment*, which examines how the brain and nervous system evolve and develop. We thank John Kristal, M.D., who served as contributing editor of this section. He is a renowned neuroscientist who brought his expertise to bear in organizing this area of the textbook.

Schizophrenia. As with the neural science section, the sections on schizophrenia were extensively updated to provide the reader with information about latest advances. The rapidly growing field of neuroimaging in schizophrenia is presented in three sections: *Structural Brain Imaging, Functional Brain Imaging,* and *Molecular Brain Imaging*. Three additional sections, *Schizophrenia: Phenotypic Manifestation, Phenomenology of Schizophrenia,* and *Psychosis as a Defining Dimension in Schizophrenia,* describe schizophrenia in a unique way and provide a humanistic understanding of what it means to suffer from psychosis. The section, *Medical Health in Schizophrenia,* provides an in-depth coverage of the medical care required to thoroughly manage this disorder. *The Concept*

of Recovery in Schizophrenia describes the complexities and varieties of outcomes and prognoses of this illness. The reader will find the most extensive survey and overview of schizophrenia to be found in any modern textbook of psychiatry. We thank Carol A. Tamminga, M.D., contributing editor, for her scientific and creative abilities in organizing this section.

Mood Disorders. The chapters covering the entire range of mood disorders provide the most up-to-date exposition of these disorders available in any textbook. New advances and new concepts in bipolar disorder are represented in two important chapters: *Psychoeducation for Bipolar Disorder* and *Brain Circuits in Major Depressive and Bipolar Disorders*. The first increases our therapeutic understanding and the second increases our scientific understanding of these complex disorders. A new section, *Suicidal Behavior: Epidemiology and Overview,* recognizes the increase in suicide both in this country and around the world and examines new methods of prevention and treatment. We wish to thank Hagop Akiskal, M.D., for his work as contributing editor for mood disorders in this and previous editions of the textbook.

Anxiety Disorders. This section has been thoroughly rewritten and updated to provide the reader with a comprehensive overview of the topic. The reader will find two updated sections of great interest: *Genetics of Anxiety Disorders* and *Neuroimaging of Anxiety Disorders,* which represent the latest advances in the field and the cutting edge of new approaches in understanding the etiology of these disorders. We wish to thank Kathleen Ries Merikangas, Ph.D., section editor, for her superb help in organizing this section.

Psychosomatic Medicine. The chapter on *Psychosomatic Medicine* includes three important sections, *Diabetes, Transplantation,* and *Burns,* all of which represent areas in which psychiatry has made significant contributions. Even though not part of the DSM-5 nomenclature, the editors have included a chapter, *Obesity*, which is a major public health problem in which psychiatry has an important role to play both in understanding the disorder and its treatment. A new section, *Management of Chronic Pain,* was written especially for this edition to call attention to this ancient malady in which the modern day psychiatrist has much to contribute. Another new section in this edition, *Psychiatric Sequelae of Critical Illness,* covers the need for psychiatrists to engage with seriously ill patients and participate in their care. The *Psychosomatic* section is one of the most comprehensive to be found in any textbook. Karin Neufeld, M.D., was contributing editor for this section, and we extend our sincere thanks to her.

Public and Global Psychiatry. As described by the authors of *Public and Community Psychiatry*, public psychiatry includes medical and psychiatric services directed "for the public good," which are comprehensively described. The reader will find an extensive overview of psychosocial needs and services around the world, in the section, *World Aspects of Psychiatry*. Other updated sections in this area include *The Psychiatric Hospitalist*, *A Socio-Cultural Framework for Mental Health and Substance Abuse Disparities*, and *Criminalization of Persons with Severe Mental Illness*. One of the editors, Pedro Ruiz, M.D., played a crucial role in organizing this section of the textbook.

Geriatric Psychiatry. The section on geriatric psychiatry has been expanded to include three new sections: the first, *HIV and Aging*, examines what was once a fatal disease to one that is now chronic and found in older adults as a result. The second, *Technology for Seniors,* covers the effects of the digital and technological revolution on the mental and physical health of the geriatric population. Finally, the new chapter, *Positive Psychiatry of Aging,* focuses on how persons who age can maintain a quality of life that allows them to reap the benefits of old age, which, instead of being seen as a disease, is and should be a period of contentment and meaningfulness. The editors wish to thank Dilip V. Jeste, M.D., for his role as contributing editor to geriatric psychiatry, which he carried out with outstanding ability and judgment in this and previous editions.

Biological Therapies

The psychopharmacological revolution in psychiatry began in the middle of the last century. In the first edition of this textbook, there were only about 50 pages devoted to this area. In this edition, there are close to 1,000 pages. Such is the tremendous growth and importance of this branch of psychiatric therapeutics. To help the reader understand this modality better, we were the first to recognize that drugs used to treat mental disorders should be classified according to their mechanisms of action rather than using such broad categories as antidepressants, antipsychotics, anxiolytics, and mood stabilizers, which are overinclusive and do not scientifically reflect the clinical use of psychopharmacological agents. For example, many antidepressant drugs are used to treat anxiety disorders; some anxiolytics are used to treat depression and bipolar disorder; and drugs from all categories are used to treat other clinical problems such as eating disorders, panic disorders, and impulse control disorders. Furthermore, there are many other drugs used in the treatment of mental disorders that do not fall neatly into any broad classification. Information about all pharmacological agents is comprehensive and includes data about indications, pharmacokinetics, dosages, adverse effects, and drug–drug interactions.

Psychotherapies

Despite the dramatic rise in pharmacologic treatment of mental disorders, psychotherapy continues to play a major role in the care of the mentally ill. *Psychotherapy of the Psychoses* is an important new addition to this section. We hope this inspires a new generation of psychiatrists and mental health professionals to revisit what psychotherapy has to offer to those with psychoses, especially those suffering from schizophrenia. A major advance in psychotherapy is to be found in the chapter, *Mentalization Based Therapy*, written by its originators, a new type of psychotherapy effective in a broad range of mental disorders. And finally, on the cutting edge, is the section, *The Use of Hallucinogenic Drugs in Psychotherapy,* which describes an exciting new area that may benefit those with mental illness and those with physical illness as well.

Child Psychiatry

This section of the book is "a text within a text." It is thoroughly comprehensive and each section represents an important new advance in diagnosis and treatment. *Neuroimaging in Child and Adolescent Psychiatry* describes in detail how imaging techniques are advancing the field of child psychiatry. The section, *Assessment of Preschoolers,* describes the special approaches to diagnosis for this unique developmental period. New advances in genetics are covered in the section, *Genetics in Child Psychiatry*. A thorough discussion of sleep problems in children is covered in the new section, *Pediatric Sleep Disorders*. Finally, a new section, *Impact on Parents of Raising a Psychiatrically Disabled Child,* deals with the difficult problems involved in managing this special patient population.

We wish to thank Caroly Sue Pataki, M.D., for her outstanding efforts as section editor. She has served in this capacity for several editions, and we owe her a debt of gratitude for her prodigious efforts.

Neurocognitive Disorders

This section of the textbook has been reorganized to reflect the new diagnostic nomenclature of DSM-5. The sections that cover delirium and dementia were completely updated and revised. All sections were rewritten by new contributors to provide a fresh approach to these brain disorders. We thank Robert A. Sweet, M.D., for his extraordinary help in organizing this section.

Case Histories

Throughout time, the teaching of psychiatry depended on the discussion and analysis of case histories, which still play an important part in psychiatric education. Case descriptions are used extensively in the text to add clarity and bring life to the clinical disorders described. They represent

the experiences of the many contributors to this textbook who share their clinical expertise and wisdom with our readers. Cases appear in shaded type so that they can be found easily. We also direct the reader to *Famous Named Cases in Psychiatry,* in Appendix A, which chronicles important psychiatric case histories from the 16th through the 21st century. We wish to acknowledge the author, David Davis, M.D., who is deceased but whose work lives on in this astounding compilation.

Citations

The style of this textbook is similar to other great textbooks of medicine: no internal citations are used. This requires contributors to evaluate the extensive and sometimes conflicting literature to create evidence-based conclusions for the benefit of the reader. That is often a difficult task, but as experts in their respective fields, they do it well. Contributors were also asked to limit references to 30 to 40 major books, monographs, and review articles and to include current references; thus, some citation lists are not as long as some of the authors would have wished. Contributors were also asked to note the five most important references with asterisks. References are as up-to-date as possible. The editors are also mindful that modern-day readers consult internet databases to stay abreast of the most current literature and they encourage that trend.

Acknowledgments

As mentioned earlier, over 2,000 psychiatrists and behavioral scientists have contributed to this text over the past half century. Obviously, we cannot thank them all. Many are long gone; but their work lives on in the various editions of this book which can be accessed in this digital era. This can be said of them all: we received their best work which is the highest compliment we can give them as editors. Many of their works are considered classics in the field.

The support staff in New York and Miami was of tremendous help. In New York we thank Nitza Jones Sepulveda who worked as senior project editor on this and previous editions of the *Comprehensive Textbook of Psychiatry*. Others who helped were Hailey Weinberg and Heidianne Grech. We also thank Dorice Viera, Associate Curator of the Frederick L. Ehrman Medical Library at the NYU School of Medicine for her valuable assistance in the preparation of this and previous editions in which she was so very helpful. In Miami, we wish to thank Gladys Robles who was of immense help, especially to Dr. Pedro Ruiz.

Special mention should be made of Norman Sussman, M.D. and Samoon Ahmad, M.D. They are not only contributors to this edition but also helped the editors in the production of this book. Dr. Sussman was contributing editor in biological psychiatry for several past editions and Dr.

Ahmad was contributing editor in psychopharmacology to many of our *Pocket Handbooks*. We value them not only as colleagues but also as close friends. We also wish to acknowledge James Sadock, M.D., and Victoria Sadock Gregg, M.D., for their assistance in their areas of expertise, emergency adult and emergency pediatric medicine, respectively. We want to take this opportunity to thank those who have translated this and other books into foreign languages, including Bulgarian, Chinese, Croatian, French, German, Greek, Indonesian, Japanese, Polish, Portuguese, Romanian, Russian, Spanish, and Turkish, in addition to a special Asian and international student edition.

We mentioned above our relationship with our publishers and those working there which was marked, by an unprecedented history of close collaboration and mutual respect. There have been so many persons over those 50 years without whose help and support these textbooks could not have been produced. We cannot possibly list them all; but in this edition there are several who deserve mention. We thank Andrea Vosburgh, Development Editor, who worked with great enthusiasm and who helped us immensely in every aspect of production. We also thank Franny Murphy, Senior Development Editor, who handled each manuscript with extreme care and interacted with almost every contributor offering them both guidance and support. We also wish to thank Chris Miller of Aptara who worked on many of our books over the years and whose expertise in the production of this book was invaluable. At Wolters Kluwer we thank Lisa McAllister, Vice-President, Medicine and Advance Practice Publishing and Rebecca Gaertner, Executive Editor, Ashley Fischer, Production Developmental Editor, Bridgett Dougherty, Production Project Manager, and Chris Teja, Acquisitions Editor for their support as well.

CONTRIBUTING EDITORS

In the early editions of this book, there were no contributing editors. Psychiatry was a field in which there was not as much specialization and subspecialization as there is today. The various experts in their respective fields were few and as editors we knew who they were. But as the field grew, we knew that to keep the book up-to-date and current we had to enlist the aid of experts who knew other experts in their respective fields. So began the long list of contributing editors whose help writing this edition and those that came before it possible. These men and women, experts in their respective fields, kept us informed of not only the latest advances in their respective fields but also provided us with the names of contributors most knowledgeable in a particular area of psychiatry and the behavioral sciences. In this edition they include the following experts who deserve special recognition: Robert Robinson, M.D., who organized the section on *Neuropsychiatry* ; John Kristal, M.D., who organized the section on *Neural Science* ; Roberto Lewis-Fernández, M.D., who organized the

section on the *Psychological Sciences* ; Robert A. Sweet, M.D., who was responsible for the section on *Neurocognitive Disorder* ; Eric Strain, M.D., who organized the section of *Substance Related Disorders* ; Charles Nemeroff, M.D., who organized the section on *Biological Psychiatry* ; Carol A. Tamminga, M.D., who organized the sections on *Schizophrenia* ; Hagop S. Akiskal, M.D., who organized the section on *Mood Disorders* ; Kathleen Merinkangas, M.D., who organized the section on *Anxiety Disorders* ; Karen Neufeld, M.D., who organized the *Psychosomatic Medicine* section; Caroly Sue Pataki, M.D., who organized the section on *Child and Adolescent Psychiatry* ; and Dilip V. Jeste, M.D., who organized the *Geriatric Psychiatry* section. The editors thank them for their prodigious efforts. The editors and the field of psychiatry are in their debt.

We especially wish to thank those at our respective Universities who enabled us to pursue our work as faculty scholars. We have been helped especially by Charles R. Marmar, M.D., Professor and Chairman of the Department of Psychiatry at the NYU School of Medicine who supported our work throughout.

Finally, we thank each of our contributors who cooperated in every aspect of this textbook.

Benjamin J. Sadock, M.D.
Virginia A. Sadock, M.D.
Pedro Ruiz, M.D.

Foreword

HERBERT PARDES, M.D.

Psychiatry has a long and complex history with substantial impact on today's society. The extent and complexity is covered by this remarkable book. Psychiatry has come a great distance from the time when people thought psychiatric conditions may be due to demons in the brain or some other exotic force. Bleuler and Kraeplin worked, studied, and categorized psychiatric illness. Freud introduced psychodynamics and psychoanalytic theory, the backdrop for psychotherapeutic approaches beginning in the late 19th and early 20th centuries.

Globally countries are faced with large populations dysfunctional because of psychiatric problems. In the early 20th century, most patients were hospitalized in chronic state psychiatric hospitals. Treatment was limited. Electroconvulsive and other convulsive therapies were used; psychotherapies were used for people less blatantly psychotic. It was not until the 1950s that major psychiatric medications started to be introduced to address symptomatically many of the symptoms that afflict people with such illnesses. At the high point there were close to 600,000 people in state hospitals around the country in the 1950s. Most care was provided in in-patient settings.

And then, many factors—new medications, patient's rights, budgetary implications of maintaining large numbers of psychiatric institutions, criticism that they were nontherapeutic—combined into a movement to change psychiatric care. Some heroes have tried to humanize care over the years with mixed success.

Patients with more modest symptomatology were treated in out-patient offices by psychiatrists and other mental health professionals. Stigma was universal and destructive.

President Kennedy fostered legislation making mental health care more visible, accessible, and patient centered. This generated the Community Mental Health program in the 1960s and 1970s. This federal plan involved bringing mental health to the community with the intended construction of 1,500 community mental health centers. Assuming 200,000 patients would be assigned to each center, the projection was 300 million people

could be served.

The intention to move patients out of the chronic hospitals led to the deinstitutionalization era of psychiatric patients launched in the 1960s and continuing for decades. Some argued patients were discharged from hospitals only to find themselves in other institutions like nursing homes, prisons, or as part of the homeless community.

In the evolution of mental health programs, World War II was important. Recruitment by the armed forces revealed the large numbers that had psychiatric illness. This excluded them from service. Beyond that, the large numbers of soldiers who during their war service had personal psychiatric challenges referred to as "battle fatigue," contributed to the rallying cry to have a national focus on mental health. In the late 1940s, the federal government established the National Institute of Mental Health (NIMH) charged with conducting research, clinical care through the community mental health centers, and education to generate a national clinical mental health workforce. The latter included psychiatrists, clinical psychologists, nurses, and social workers, as well as paraprofessionals.

Medicines such as Thorazine, antidepressants, and lithium became available. Enthusiasm regarding discharging patients from chronic hospitals and putting them out in the community strengthened.

Given the promise and need for the workforce charged to care for psychiatric patients, there was great interest among physician trainees for psychiatry as a career. This was sustained through the early 1960s. Possibly because of the diminishing assessment of the therapeutic impact of psychodynamic approaches, there was a drop in recruits to psychiatry from about 10% in the early 1960s to about 2 to 3 percent of graduating medical students by the beginning of the 1970s.

Mental health was a particularly low priority during the republican national administrations in the late 1960s and early 1970s. Money for support of the programs at NIMH was sharply curtailed, while other areas such as cancer and heart disease were increased. Lack of reimbursement rose as a major insidious issue for those trying to promote mental health services.

Important in the evolution of mental health and of psychiatry was the Carter Commission of the late 1970s. This Commission recommended major changes in mental health care. The "Mental Health Systems Act" passed in 1980 was intended to increase federal involvement in mental health and focus on underserved populations such as children, minorities, elderly, chronically mentally ill, and psychiatric patients in nonpsychiatric medical settings. But after working several years and actually passing that legislation, the Carter Administration was replaced by the Reagan Administration. It challenged sharply the budgets of the NIMH, and the National Institute on Alcohol Abuse and Alcoholism, and National Institute on Drug Abuse conducting programs that dealt with alcohol and drug

abuse, respectively.

The late 1970s witnessed an improvement in the congressional view regarding NIMH. The Reagan Administration was impressed by findings such as the demonstration by Lou Sokoloff of the value of 2-deoxyglucose used in PET scanning for visualization of activity in the brain. This was enhanced by the Nobel Prize awarded to Roger Sperry, an NIMH research psychologist grantee which led to the understanding that there are major differences in the functions of the two halves (left and right) of the brain.

These advances from biological research were accompanied by a vigorous effort by the mental health community leading to revaluing of mental health programs. Productive studies emanated from Seymour Kety regarding genetic contributions to psychiatric disorders. Others produced increasingly systematic studies of brain and behavior. These developments improved the regard in which mental health programs were held during the 1980s. The Reagan Administration began recommending increases in the NIMH budget more comparable to the size of increases allocated to other federal medical areas—heart disease, cancer, among others.

Important too was the development of more sophisticated epidemiological programs. These came about in the late 1970s and 1980s. These highlighted by the Epidemiologic Catchment Program generated by the NIMH were conducted in a more rigorous manner and confirmed the large amounts of mental illness.

Further, in the enhancement of mental health programs was the explosive rise of activity in neuroscience in the 1980s continuing today. It became clear owing to the flourishing of research that psychiatric causative factors were of a mixed and complicated nature brought about by multiple factors. In the psychiatric community, there was increased tension that arose owing to disagreement regarding the importance of the biological, congenital, and genetic factors as opposed to environmental factors.

One unfortunate longstanding hypothesis that was repudiated was that the “schizophrenogenic parents” bore responsibility for psychiatrically ill people. The excitement mounting in the 1980s and 1990s, and the early 20th century in genetics lead to wide scale searches for genetic factors.

The fact that close to 20 percent of the population suffer in any given year from psychiatric illness began to have an impact on the general public view of the importance of psychiatric programs. Psychiatry had primarily been treated as a low priority and consequently suffered from a paucity of resources. This was compounded by the fact that there were few if any procedures useful in psychiatric diagnosis and care in contrast to other medical disciplines. Those procedures are often invaluable in the generation of resources available to healthcare professionals working on orthopedic, cardiac, cancer disorders, among others.

While brain sciences evoked considerable excitement, activity in the psychosocial realm in that period failed to elicit comparable enthusiasm.

It was important to achieve consistency regarding psychiatric diagnosis. It had been demonstrated through US/UK studies that the percentages of psychiatric diagnoses varied considerably from country to country. This led to attempts stimulated by the NIMH and the American Psychiatric Association to gain international agreement on the criteria for psychiatric disorders to achieve more consistency. As a result, a major conference was held in Copenhagen Denmark in the late 1970s bringing worldwide psychiatric experts together. This facilitated the thrust for consistency in criteria for psychiatric diagnosis. The DSM (Diagnostic and Statistical Manual) for psychiatric illness became a major effort of organized psychiatry.

Several DSM efforts sharpened further the observation that in making psychiatric diagnoses, the Achilles heel facing psychiatric diagnosis was that it did not have hard and fast criteria for diagnosis and evaluation as contrasted to other medical disciplines—for instance, the use of blood pressure readings in hypertension, EKGs in heart disease, and glucose readings in diabetes.

From the 1980s through today, there have also been broad efforts to amplify and sharpen the interventions used for psychiatric care. Multiple kinds of psychotherapy including brief psychotherapy, cognitive-behavioral therapy, and interpersonal psychotherapy became available changing what had been a predominant use of psychodynamic psychotherapy and psychoanalysis. Electroconvulsive therapy (ECT) was refined to make the side effects of the treatment more effective. With many improvements ECT focused more on specific conditions for which it could be of greatest help, for example, major unresponsive depression.

Meanwhile there was a proliferation of psychiatric drugs. One goal was to reduce the side effects disturbing to patients and another was to increase options for doctors treating patients for psychiatric illness. Multiple available psychopharmacological treatments meant that doctors could switch from one drug that failed to another to find the right medication for a given patient.

Over time, it was recognized that many psychiatric conditions which might look manifestly similar were brought about by different factors. Thus, since in practice the same treatment was given to many people, the medications had uneven results.

This resulted in disappointment at the modest percentages of people who had definitive and sustained help from psychopharmacological treatments. Of related interest, mental retardation had at one point been thought of as related to a limited set of factors. Over the years, this changed and mental retardation is today recognized as being the result of varied combinations of factors. The fact that the etiological factors vary also suggested that if one could match specific treatments to patients' needs, one might improve the overall outcome statistics of patients using

psychiatric interventions.

When considering mental health and psychiatry care today, it is noteworthy that a number of causes, some positive and some negative, have created a unique situation. Psychiatric programs and psychiatry is receiving much more visibility. There are many reasons:

1. The unfortunate reality that large numbers of Veterans return from war with psychiatric problems including PTSD, drug abuse, etc.
2. The increase in suicide in both the civilian and veteran population.
3. The recognition that psychiatric care so frequently coexists with nonpsychiatric medical situations, that is, comorbidity. Not only has this been recognized as an increasing phenomenon, but as economists search for the costly components of healthcare requiring maximum resources, it is apparent patients with comorbidities such as diabetes and depression or heart disease and depression need more care. Consequently, care for these patients involves higher costs.
4. The above factors are important; their impact is dramatically augmented by the rash of violent mass killings frequently perpetrated indiscriminately by patients with psychiatric illnesses on large crowds in settings around the world. This greatly enhanced public attention.

The debate regarding where society should put its effort in reaction to these terrifying and destructive mass killing is highly contentious. While it is recognized that generally psychiatric patients are not prone to be perpetrators of violence, they are often the culprits in episodes of mass killings by individuals mostly with guns.

The realization that attention to policies both regarding gun availability and also expanded access to psychiatric treatment as components of a joint strategy may be optimal has been thwarted because the issue is highly politicized.

This multitude of factors, that is, veterans problems, increased suicide, increased cost of comorbid psychiatric and nonpsychiatric conditions, mass violence along with recognition of the overwhelming numbers of individuals who both need care yet are winding up in prisons and jails, has given mental health an unprecedented visibility.

CITIZENS AND ADVOCATES

A welcome development began in the late 1970s. Families of psychiatric patients created an organization, the National Alliance for the Mentally Ill (NAMI) to advocate for better programs for the mentally ill.

In a noteworthy meeting in 1979, parents of schizophrenic people came together in Madison, Wisconsin and agreed to launch a citizen advocacy group particularly focused on mental illness. While organizations already existed, some groups like the NAMI prioritized strengthening programs in research and care for seriously ill psychiatric patients.

That led to a merger in which NAMI and a schizophrenia foundation from Kentucky agreed to establish a private not-for-profit support organization, the National Alliance for Research on Schizophrenia and Depression (NARSAD) (recently renamed Brain and Behavior Research Foundation [BBRF]). This became the primary private organization worldwide supporting psychiatric research. It focused on basic and clinical research to understand the causes of major psychiatric illness and generate more and better treatment. From its launch in the mid-1980s, BBRF has provided $346 million to thousands of scientists of diverse disciplines around the world to stimulate more psychiatric research.

NAMI became the lead national family advocacy effort working to increase NIMH support and create a more constructive alliance between citizens, families, and psychiatric health professionals—both in clinical and research activity.

The NIMH has the largest pool of money by far for psychiatric programs with its intramural program at Bethesda and its extramural program around the world. A second government organization, Substance Abuse and Mental Health Services Administration (SAMSHA), was created in the 1980s to conduct the block grant program for mental health care that resulted from the conversion from the earlier federal grants in mental health, alcohol, and drugs to block grants for states. NIMH and SAMSHA became the two primary federal organizations that support mental health work.

The intent of the Reagan Administration was to transfer back to the states money and responsibility for these programs. This undid some of what the Carter Administration had recommended or accomplished.

CURRENT PSYCHIATRIC RESEARCH

The result of public and private support of research has led to many exciting developments.

Once clear that genetic factors play a role in psychiatric illness, a multitude of studies over the last several decades have identified genes as contributory factors in some of the main psychiatric illnesses. This work continues to determine the full extent of genetic involvement in psychiatric illness.

Also invaluable has been brain imaging. The search is for changes in the brain, early signs of developing psychiatric problems, and determination of the linkage between observed brain changes and individual behavioral functioning, and illness.

One particularly unique new research approach is Optogenetics. This has provided an acclaimed approach for examining the functions of nerve cells by using light to help determine how they are influenced and modified.

The range of research possibilities is exhaustive. There is ongoing

considerable effort designed to understand the way the treatments work.

In suicide, many questions range around identifying warning signs and clarifying how one can intercept a developing effort to complete a suicide. Deep brain stimulation therapy has become another area of intense interest for depression and transcranial magnetic stimulation for unresponsive depression.

Unfortunately, pharmaceutical industry has decreased its research on potential medications for psychiatric illness. The funding overall for individuals doing psychiatric science has been adversely influenced by this lessening of pharmaceutical support. On the other hand, there has been benefit from more equitable increases in NIMH support and increased private support.

EDUCATION

In education the increased role of biological factors has influenced heavily the training of medical students and psychiatric residents. Whereas several decades ago psychodynamic, psychotherapeutic education, supervision, and monitoring dominated training and education, this has decreased substantially.

In many ways the position that Gerry Klerman articulated when he was head of the Alcohol, Drug Abuse, and Mental Health Administration stressing the value of combined psychological and pharmacological therapies in the treatment of a particular individual has become more mainstream.

PRIMARY CARE AND MENTAL HEALTH CARE

One of the other major challenges is the need for greater manpower. This has led to increased use of people other than mental health professionals especially, primary care clinicians. They are seen as a valuable resource with which to respond the vast amount of psychiatric need. Various programs are being developed that integrate primary and mental health care. Fueling this trend is the belief that by such integration, cost will go down, and treatment will be better. The program led by Jürgen Unutzer in Seattle Washington is a good example of the effectiveness of this perspective.

At New York-Presbyterian Hospital and Columbia University Medical Center, the finding that one can fully integrate psychiatric faculty into the team on a given internal medicine in-patient unit and as a result achieve better outcomes and reduced length of stay should result in a wide embrace of this type program.

In psychiatry it is increasingly recognized that in facilitating the careers of young people, mentoring and modeling is important. Repeatedly, young people point to revered or respected senior faculty member who were critical in their maturation. The mentors are appreciated as helpers, guides,

and models to emulate.

EARLY INTERVENTION

What also has become apparent is that early intervention is important in the treatment of psychiatric conditions. Therapeutic assistance in a timely way means that rather than having a patient suffering from an illness, one can avoid the effects of the illness itself compounding the original disorder. This potential exacerbation of their dysfunction is countered by early identification. This is central to prevention. Devising programs which match the patient's needs is most valuable.

SUMMARY

There have been vast changes in psychiatric care, psychiatric research, and psychiatry over the past 150 years. Still the elusive nature of the etiology of psychiatric conditions remains perplexing. There have been Nobel prizes won by Drs. Eric Kandel, Arvid Carson, and Paul Greengard, and Lasker Awards by Drs. Lou Sokoloff, Nancy Wexler, Eric Kandel and others, but the core complexity of psychiatric illness remains. New developments like the launch of the Lieber Institute at Johns Hopkins by Steve and Connie Lieber, two of the most generous philanthropists in psychiatric research, and the federal launch of a major effort to explicate the functions of the brain are welcome developments. The stigma that for decades has inhibited philanthropic support for psychiatric care has been diminishing; there has been more support as illustrated by the Graham Boeckh Foundation in Canada.

The world oversees an enormous population of psychiatric need. The resources available are uneven around the world. And yet as readers will see illustrated in this most informative edition of the latest *Kaplan & Sadock's Comprehensive Textbook of Psychiatry*, the global range of potential scientists and programs working on these problems is encouraging. Also, global collaboration is increasing, allowing exploitation of the collective increased resources of the scientific enterprise.

Psychiatric patients are certainly treated better now than in the past. But given criminalization and resource weakness, there is still a long way to go. Whether the new forces mentioned earlier will induce greater attention and better programming for psychiatric problems remains to be seen, but certainly is a goal for which the broad mental community is striving.

1

Neural Sciences

▲ 1.1 The Neuroscience of Psychiatry

JOHN H. KRYSTAL, M.D.

INTRODUCTION

This section of the *Textbook* contains a remarkable collection of chapters that provide an overview of the landscape of the current understanding of the neuroscience foundations of psychiatry. If the brain is the principal organ of psychiatry, then an understanding of brain science is the starting point for a comprehensive knowledge of the field. These chapters bear careful reading as they reflect the first indications of the emergence of explanatory science within psychiatry. This progress arises from the application of technologies that have arisen since the prior edition of the textbook, techniques such as optogenetics, DREADDs, and CRISPR, for example. It also arises from the maturation of fields, such as cognitive neuroimaging and psychiatric genetics; transforming them from intellectual curiosity to powerful tools to answer important questions about the human behavior and mental illness.

What are signs of this progress? There are many inspiring stories. With respect to etiology, advances in genome sequencing have enabled the discovery of rare gene variants that contribute significantly to illness risk and account for approximately 1 percent of the cases of autism. Psychologists are beginning to understand where in the brain these genes are preferentially expressed and when in development these genes are critically engaged. These studies set the stage for disease-modifying interventions. Similar progress is emerging for schizophrenia, a disorder that appears to share some risk genes with autism. With respect to pathophysiology, optogenetic studies in animals and neuroimaging studies in humans have provided deep new insights into how depression might be expressed at the level of neural circuits. These "raw data" provide the basis for computational neuroscience studies that rigorously describe for the first time the relationships between the activity of neurons and networks or

networks and behavior. In turn, these insights are providing a basis for a growing array of neurostimulation treatments for mood disorders. At the same time, some transformative advances in treatments, such as the discovery of the rapid antidepressant effects of ketamine, have stimulated basic science advances that shed new light on the neurobiology and may lead to new classes of antidepressant medications.

Why do psychiatrists need to understand neuroscience to practice their art? Psychiatry is moving away from viewing individual psychiatric patients solely as representative of static diagnostic categories, as might be reflected in the *Diagnostic and Statistical Manual*. The bold leadership of Thomas Insel in advancing the NIMH Research Domain Criteria initiative has pushed the field to encompass an understanding of the neural circuitry of behavior and the dimensional features of the pathophysiology and phenomenology of psychiatric patients that cross traditional diagnostic boundaries. One might predict that as psychiatrists become increasingly comfortable incorporating an understanding of the neurobiology of psychiatric disorders into their clinical formulations of individual patients, it will enrich their understanding of their patients and suggest new opportunities to enhance treatment.

The author will illustrate this point by referring to a case conference in which he participated, which he first presented in a blog to the American College of Neuropsychopharmacology in January 2012. The Yale Psychiatry Residents invited the author to interview a patient and then to present his formulation to their group. The author came to the session without any information about the patient and had 30 minutes to interview the patient.

The patient was a young man who suffered with long-standing psychotic symptoms. He met diagnostic criteria for schizophrenia, exhibiting many domains of preserved intellectual capacity but noted problems with sustained attention and memory. He also exhibited low levels of affective blunting but impairments in many social and functional domains. Upon interview, he also described vividly a childhood pattern of accentuated attention to nonsocial environmental details and he both described and displayed perseverative tendencies and mild irritability when challenged. He also mentioned that both the diagnoses of schizophrenia and bipolar disorder ran in his family. With regard to treatment, he felt that clozapine was the most effective medication prescribed for him. In addition, he was treated with the anticonvulsants, lamotrigine, and valproate.

How could psychiatrists understand the appearance of features of autism spectrum disorder (ASD) and his family history of both schizophrenia and bipolar disorder? This patient presented opportunities to reflect on challenges in our diagnostic schema. Bleuler suggested that autistic thinking was a core feature of schizophrenia, but this feature is no longer included in the diagnosis of schizophrenia. Yet, a growing number of rare gene variants with relatively large effects on the risk for schizophrenia

also contribute to the risk for ASD. Overlap of schizophrenia and ASD symptoms might be an illustration of pleiotropy. Similarly, recent large-scale GWAS studies report regions of the genome that appear to convey relatively modest increases in the risk for both schizophrenia and bipolar disorder. Thus, genetics is providing some clues that may help to guide thinking about the features of the presentation of individual patients that conflict with the diagnostic categories of DSM-IV.

Why would someone add the antipsychotics lamotrigine and valproate to enhance clozapine treatment of schizophrenia? Currently, there is a small evidence base to guide the pharmacotherapy of clozapine-resistant schizophrenia symptoms. The author was reminded of the evidence in animals and humans, that lamotrigine, by itself or in combination with clozapine, attenuates some of the effects of psychotigenic NMDA receptor antagonists. Subsequently, lamotrigine was not found generally effective as an antipsychotic augmentation strategy unless patients had symptoms that had been highly resistant to treatments including clozapine. Further, anticonvulsants, like lamotrigine, might protect against the seizure risk associated with pharmacotherapy with higher clozapine doses. Thus, there was a rationale for adding lamotrigine to clozapine in this case. However, the same could not be said for valproate. Psychiatrists discussed the pharmacoepidemiologic data describing common co-prescription of valproate and antipsychotic medications for schizophrenia despite the lack of compelling efficacy data for valproate in this adjunctive role.

As noted in the blog, the invitation from the residents to interview a patient in front of them and to present a "biopsychosocial" formulation was embedded in an unspoken question. Does it make a difference to a practicing psychiatrist to be a student of genetics and translational neuroscience? The answer turned out to be, absolutely! In fact, it seemed that the author's background in neuroscience informed nearly every aspect of the session, the questions that he asked, the hypotheses that he generated about the patient, and his thoughts about the treatment plan. While it was not yet clear to him how his knowledge of the genetics of schizophrenia would guide the treatment of this patient, it was not hard to see how keeping abreast of developments in neuroscience might someday make this possible.

REVIEW OF THE NEUROSCIENCE CHAPTERS

To facilitate this type interplay between neuroscience and clinical psychiatry, the authors structured this section of the *Textbook* to cover the waterfront of progress in neuroscience that was most relevant to psychiatry. Experts in the field were asked to provide chapters that would be rigorous but accessible to clinicians. The section builds from elemental questions, that is, genetics, the properties of neurons, the basics of chemical neurotransmission, cortical networks and network functions, and

the pathophysiology of psychiatric disorders. The topics covered are broad, but some topics were not included. Similarly, the reviews provide depth, but hopefully not at the expense of accessibility.

The first subsection begins with two chapters that describe the organization and development of the brain. Chapter 1.2 provides a wonderful introduction to the neuroanatomical organization of the brain. Chapter 1.3 describes the processes through which the complex organization of the brain emerge through development. This chapter also introduces the reader to adult neurogenesis, the process through which new neurons emerge in the hippocampus in adulthood.

The next eleven chapters provide an overview of the basic signaling mechanisms that underlie communication and plasticity within the brain. Five chapters review the principal classes of chemical neurotransmission in the brain including monoamine neurotransmitters (Chapter 1.4), the amino acid neurotransmitters glutamate and GABA (Chapter 1.5), neuropeptides (Chapter 1.6), and novel neurotransmitters (Chapter 1.8). Chapter 1.7 covers the neurotrophic factors. These substances are increasingly implicated in neuropsychiatric pathophysiology and treatment, particularly the activity of both traditional and rapidly acting antidepressant medications. Intraneuronal signaling, the mechanisms through which the effects of transmitters and drugs are transduced chemically within neurons, are reviewed in Chapter 1.9. Neurophysiologic functions are covered in Chapter 1.10. Two chapters address the molecular biology of psychiatric disorders. Next, there is an update on "omic" technologies (genomics, transcriptomics, proteomics) in Chapter 1.11. Chapter 1.13 covers neural–immune interactions, with a particular focus on neuroinflammation, a key mediator of stress effects on the brain. This subsection concludes with a chapter on chronobiology. Chapter 1.14 covers the genes, circuitry and mechanisms that govern circadian and other rhythms that have profound effects on energy metabolism, sleep, behavior, and mood.

The third subsection consists of four chapters related to technologies that are increasingly informing psychiatrists' understanding the biology of psychiatric disorder. Chapter 1.15 addresses advances in electroencephalography including studies of event-related potentials and cortical oscillations. Chapter 1.16 addresses advances in magnetic resonance imaging and magnetic resonance spectroscopy, a central approach for noninvasively studying brain structure, function, and chemistry. Chapter 1.17 reviews the other principal noninvasive approaches for studying brain chemistry, particularly for quantifying the levels of specific brain proteins that are present in very low quantities in the brain, such as the receptors for neurotransmitters. Chapter 1.18 focuses on advances in gene mapping approaches that are contributing to the emerging understanding of the genetics of psychiatric disorders.

The next subsection includes eight chapters that cover areas of focus in

psychiatry research. Chapter 1.19 covers the important topic of animal models for psychiatry and addiction. Animal models have provided fundamental insights into the neurobiology and treatment of psychiatric disorders. Nonetheless, these models have shown important limitations in their ability to predict new treatments for psychiatric disorders, a factor that has contributed to the paucity of new treatment mechanisms in psychiatry over the past 20 years. There are chapters in this subsection on pain (Chapter 1.20), sense of self (Chapter 1.21), sleep (Chapter 1.22), and appetite (Chapter 1.23). These are fundamental features of human experience that have been studied intensively. The last two sections cover relatively new areas that are playing growing roles in the field of psychiatry. Epigenetics, that is, the mechanisms through which environmental factors have lasting and even transgenerational effects, is reviewed in Chapter 1.25. Finally, Heath reviews the implications of microbiome research for human behavior.

The final subsection on the neuroscience of psychiatry covers special topics in systems, cognitive, and behavioral neuroscience. Chapter 1.26 covers issues in systems neuroscience. Chapter 1.27 addresses the broad impact of learning theory. Barch, a central figure in the Human Connectome Project, addresses the important area of human cortical functional connectivity in Chapter 1.28. Chapter 1.29 addresses the growing area of computational neuroscience, that is, the effort to develop computational models for cortical microcircuit and macrocircuit functions. The systems neuroscience of cortical development is covered in Chapter 1.30. Finally, Salloum addresses the evolving topic of social cognition.

CONCLUSIONS

In the final analysis, psychiatrists must be humble with regard to the reach of their knowledge. The brain is unimaginably complicated and fundamental aspects of the pathophysiology of neuropsychiatric disorders remain beyond grasp. Psychiatrists are prone to overestimate their knowledge and to underestimate the challenges faced in advancing their understanding of the etiology, pathophysiology, treatment, and prevention of neuropsychiatric disorders. Nonetheless, psychiatrists have a clinical, scientific, and humanistic imperative to advance their understanding of the neurobiology of psychiatric disorders and to see that this knowledge informs the clinical practice of psychiatry.

REFERENCES

Anand A, Charney DS, Oren DA, et al. Attenuation of the neuropsychiatric effects of ketamine with lamotrigine: support for hyperglutamatergic effects of N-methyl-D-aspartate receptor antagonists. *Arch Gen Psychiatry*. 2000;57:270–276.

Association AP. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Press; 2013.

Brody SA, Geyer MA, Large CH. Lamotrigine prevents ketamine but not amphetamine-induced deficits in prepulse inhibition in mice. *Psychopharmacology (Berl)*. 2003;169:240–246.

Chen H, Deshpande AD, Jiang R, Martin BC. An epidemiological investigation of off-label anticonvulsant drug use in the Georgia Medicaid population. *Pharmacoepidemiol Drug Saf*. 2005;14:629–638.

Duman RS, Aghajanian GK, Sanacora G, Krystal JH. Synaptic plasticity and depression: new insights from stress and rapid-acting antidepressants. *Nat Med*. 2016;22:238–249.

Farber NB, Jiang XP, Heinkel C, Nemmers B. Antiepileptic drugs and agents that inhibit voltage-gated sodium channels prevent NMDA antagonist neurotoxicity. *Mol Psychiatry*. 2002;7:726–733.

Goff DC, Keefe R, Citrome L, et al. Lamotrigine as add-on therapy in schizophrenia: results of 2 placebo-controlled trials. *J Clin Psychopharmacol*. 2007;27:582–589.

Gulsuner S, Walsh T, Watts AC, et al. Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell*. 2013;154:518–529.

Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010;167:748–751.

Krystal JH, State MW. Psychiatric disorders: diagnosis to therapy. *Cell*. 2014;157:201–214.

Moskowitz A, Heim G. Eugen Bleuler's Dementia praecox or the group of schizophrenias (1911): a centenary appreciation and reconsideration. *Schizophr Bull*. 2011;37:471–479.

Ripke S, Sanders AR, Kendler KS, et al. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet*. 2011;43:969–976.

Schwarz C, Volz A, Li C, Leucht S. Valproate for schizophrenia. *Cochrane Database Syst Rev*. 2008:CD004028.

State MW, Levitt P. The conundrums of understanding genetic risks for autism spectrum disorders. *Nat Neurosci*. 2011;14(12):1499–1506.

Tiihonen J, Wahlbeck K, Kiviniemi V. The efficacy of lamotrigine in clozapine-resistant schizophrenia: a systematic review and meta-analysis. *Schizophr Res*. 2009;109:10–14.

Treadway MT, Pizzagalli DA. Imaging the pathophysiology of major depressive disorder—from localist models to circuit-based analysis. *Biol Mood Anxiety Disord*. 2014;4:5.

Tye KM, Mirzabekov JJ, Warden MR, et al. Dopamine neurons modulate neural encoding and expression of depression-related behaviour. *Nature*. 2013;493:537–541.

Wang XJ, Krystal JH. Computational psychiatry. *Neuron*. 2014;84:638–654.

Williams HJ, Zamzow CR, Robertson H, Dursun SM. Effects of clozapine plus lamotrigine on phencyclidine-induced hyperactivity. *Prog Neuropsychopharmacol Biol Psychiatry*. 2006;30(2):239–243.

Willsey AJ, Sanders SJ, Li M, et al. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell*. 155:997–1007.

▲ 1.2 Functional Neuroanatomy

DARLENE SUSAN MELCHITZKY, M.S., AND DAVID A. LEWIS, M.D.

The broad range of affective, cognitive, and behavioral characteristics of humans arises as a consequence of specific patterns of activation in networks of neurons that are distributed across the central nervous system (CNS). These patterns of activation are mediated by the connections among specific brain structures. Consequently, understanding the neurobiologic bases for the disturbances in affective, cognitive, and behavioral processes present in psychiatric disorders requires an appreciation of the major principles governing the functional organization of these structures and their connections in the human brain. This section reviews some of these anatomic principles and illustrates them in the functional circuitry of several neural systems. These neural systems—the thalamocortical, basal ganglia, and limbic systems—were selected because of their particular relevance for psychiatric disorders.

PRINCIPLES OF BRAIN ORGANIZATION

Cells

The human brain contains approximately 10^{11} nerve cells, or *neurons*. In general, neurons are composed of four morphologically identified regions (Fig. 1.2–1): (1) the cell body, or *soma,* which contains the nucleus and can be considered the metabolic center of the neuron; (2) the *dendrites,* processes that arise from the cell body, branch extensively, and serve as the major recipient zones of input from other neurons; (3) the *axon,* a single process that arises from a specialized portion of the cell body (the *axon hillock*) and conveys information to other neurons; and (4) the *axon terminals,* fine branches near the end of the axon that form contacts (*synapses*) generally with the dendrites or the cell bodies of other neurons, release neurotransmitters, and provide a mechanism for interneuronal communication.

Most neurons in the human brain are considered to be multipolar in that they give rise to a single axon and several dendritic processes. Although there are numerous classification schemes for neurons in different brain regions, almost all neurons can be considered to be either projection or local circuit neurons. *Projection neurons* have long axons and convey information from the periphery to the brain (sensory neurons), from one brain region to another, or from the brain to effector organs (motor neurons). In contrast, *local circuit neurons* or *interneurons* have short axons and process information within distinct regions of the brain.

Neurons can also be classified according to the neurotransmitters they contain (for example, the dopamine neurons of the substantia nigra). Identification of neurons by their neurotransmitter content in anatomic studies provides a means for correlating the structure of a neuron with certain aspects of its function. However, neurotransmitters have defined effects on the activity of neurons, whereas complex brain functions, such as

those disturbed in psychiatric disorders, are mediated by the coordinated activity of ensembles of neurons. Thus, the effects of neurotransmitters (or of pharmacologic agents that mimic or antagonize the action of neurotransmitters) on behavioral, emotional, or cognitive states must be viewed within the context of the neural circuits that they influence.

In addition to neurons, the brain contains several types of *glial* cells (Table 1.2–1), which are at least 10 times more numerous than neurons. *Astrocytes* are the most numerous class of glial cells and are classified as either protoplasmic or fibrous. Protoplasmic astrocytes are large cells located exclusively in the gray matter with many fine and elaborate processes. In contrast, fibrous astrocytes are smaller with less complex processes and reside exclusively in the white matter. Protoplasmic astrocytes have been demonstrated to serve a number of functions, including participation in the formation of the blood–brain barrier (BBB), removal of glutamate and γ-aminobutyric acid (GABA) from the synaptic cleft, and buffering of the extracellular potassium (K^+) concentration.

The close contact of their processes with neurons and blood vessels suggest that protoplasmic astrocytes may help support the energy requirements of neurons. For example, distinctive processes of protoplasmic astrocytes called endfeet line the basal lamina of blood vessels. These endfeet appear to extract glucose from the blood and convert it to lactate which is then released for neurons to uptake and use as an energy substrate. The endfeet may also play a role in the regulation of cerebral blood flow (CBF) leading protoplasmic astrocytes to be referred to as a neurovascular bridge. In addition to these regulatory functions, protoplasmic astrocytes are involved in synaptic neurotransmission in two distinct ways. First, perisynaptic astrocytic processes express a variety of neurotransmitter receptors, and these receptors are stimulated by the release of neurotransmitters from presynaptic axon terminals. The activated glial cell then releases gliotransmitters that can stimulate the postsynaptic neuron. Thus, the perisynaptic astrocyte is an active partner in synaptic transmission, creating what has been termed the tripartite synapse. Second, neurotransmission is regulated by the structural network of astrocytes, which has been shown to comprise nonoverlapping domains of individual astrocytes. In the hippocampus and the cortex, individual astrocytes possess their own domain, with only their most distal processes interdigitating with processes of neighboring astrocytes. This pattern of organization creates distinct astrocytic domains, in which individual astrocytes modulate the activity of neurons and synapses.

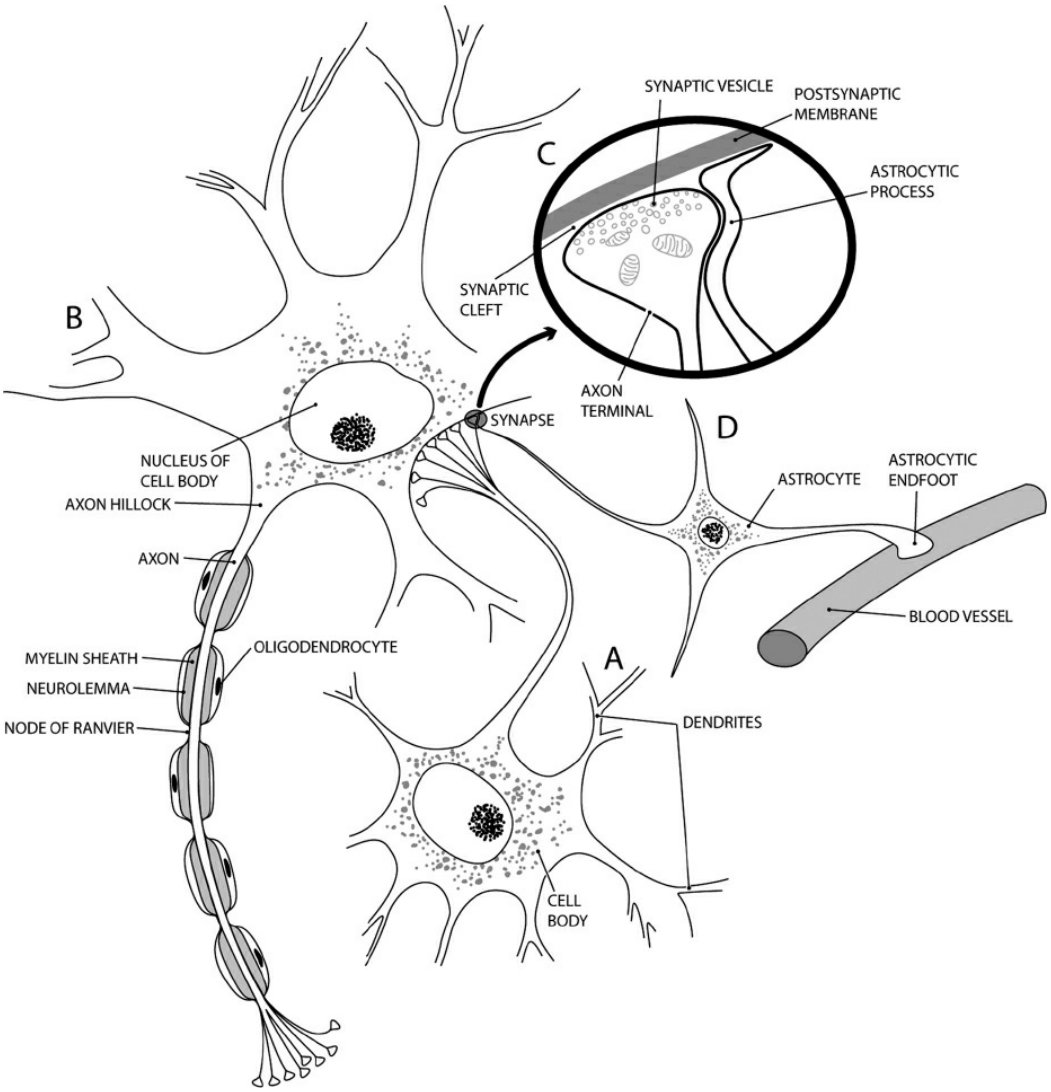

FIGURE 1.2–1. Drawing of the major features of a typical neuron, oligodendrocyte, and astrocyte. **A** and **B:** The main features of neurons include a cell body, dendrites, a principal axon, and axon terminals. The principal axon is myelinated by neighboring oligodendrocytes. **C:** Axon terminals branch off axons and form synapses with postsynaptic neurons via the release of neurotransmitters stored in synaptic vesicles. **D:** Fine astrocytic processes surround synapses and regulate neurotransmission, while astrocytic endfeet contact blood vessels and regulate blood flow. (Adapted from Gilman S, Winans-Newman S. *Manter and Gatz's Essentials of Clinical Neuroanatomy and Neurophysiology*, 10th ed. Philadelphia, PA: FA Davis Co; 2003:2.)

Table 1.2–1. Glial Cells

	Fibrous Astrocytes	Protoplasmic Astrocytes	Oligodendrocytes	Schwann Cells	Microglia
Location	Contact neuronal cell bodies, dendrites, and axons in gray matter; form a complete lining around the	Organized along white matter tracts; endfeet contact blood	Myelinating oligodendrocytes form myelin sheaths around CNS axons; satellite oligodendrocytes	Form myelin sheaths around myelinated axons and ensheath unmyelinated	Gray and white matter of CNS

	external surfaces of the CNS and around CNS blood vessels	vessels, ependymal, axons, and nodes of Ranvier	surround CNS neuronal cell bodies	axons	
Function	Maintenance of extracellular ionic environment; secretion of gliotransmitters and growth factors; structural and metabolic support of neurons; regulation of blood flow	Support of axons in white matter	Myelinating oligodendrocytes—myelination; satellite oligodendrocytes—unknown	Myelination; biochemical and structural support of myelinated and unmyelinated axons	Scavenging and phagocytosis of debris after cell injury and death; secretion of cytokines

Modified from Haines DE. *Fundamental Neuroscience for Basic and Clinical Applications*, 3rd ed. Philadelphia, PA: Elsevier; 2006:25.

The cell bodies of fibrous astrocytes are evenly distributed within the white matter, and unlike protoplasmic astrocytes, their processes overlap. Some processes form endfeet that contact blood vessels, ependymal cells, axons, and nodes of Ranvier. Fibrous astrocytes appear to play the traditional supportive role of neuroglia.

Oligodendrocytes and *Schwann cells,* found in the CNS and peripheral nervous system (PNS), respectively, are small cells that wrap their membranous processes around axons in a tight spiral. The resulting *myelin sheath* facilitates the conduction of action potentials along the axon. *Satellite oligodendrocytes* are found preferentially in gray matter and their perineuronal location suggests that they may be involved in the regulation of the environment surrounding neurons.

The third class of glial cells, the *microglia,* is derived from macrophages and historically was thought to function solely as scavengers, eliminating the debris resulting from neuronal death and injury. However, studies on microglial function over the last decade have demonstrated that microglia also regulate neurogenesis and synaptic function.

Alterations in glial cells may contribute to the pathophysiology of psychiatric disorders. For example, stereologic studies of postmortem human brain tissue have revealed decreases in glial cell number in the dorsal prefrontal cortex of individuals with schizophrenia and in the subgenual medial prefrontal cortex (MPFC) in depressed individuals. Moreover, the decrease in CBF in the prefrontal cortex of individuals with depression may be linked to a reduction in the amount of blood vessel coverage by astrocytic endfeet. Although some putative susceptibility genes

for schizophrenia are selectively expressed in glial cells, it is still unclear whether the alterations in glial cell number reflect the disease process or are a consequence of treating the disease.

Architecture

Neurons and their processes form groupings in many different ways, and these patterns of organization, or architecture, can be evaluated by several approaches. The pattern of distribution of nerve cell bodies, called *cytoarchitecture,* is revealed by aniline dyes called Nissl stains that stain ribonucleotides in the nuclei and the cytoplasm of neuronal cell bodies. The Nissl stains show the relative size and packing density of the neurons and, consequently, reveal the organization of the neurons into the different layers of the cerebral cortex. In certain pathologic states, such as Alzheimer disease (referred to as Major Neurocognitive Disorder due to Alzheimer type in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* [DSM-5]), neuronal degeneration and loss result in striking changes in the cytoarchitecture of some brain regions (Fig. 1.2–2).

Other types of histologic techniques, such as silver stains, selectively label the myelin coating of axons and consequently reveal the *myeloarchitecture* of the brain. For example, certain regions of the cerebral cortex—such as area MT, a portion of the temporal cortex involved in processing visual information—can be identified by a characteristic pattern of heavy myelination in the deep cortical layers. The progression of myelination is highly region-specific, may not be complete for years after birth, and may be a useful anatomic indicator of the functional maturation of brain regions.

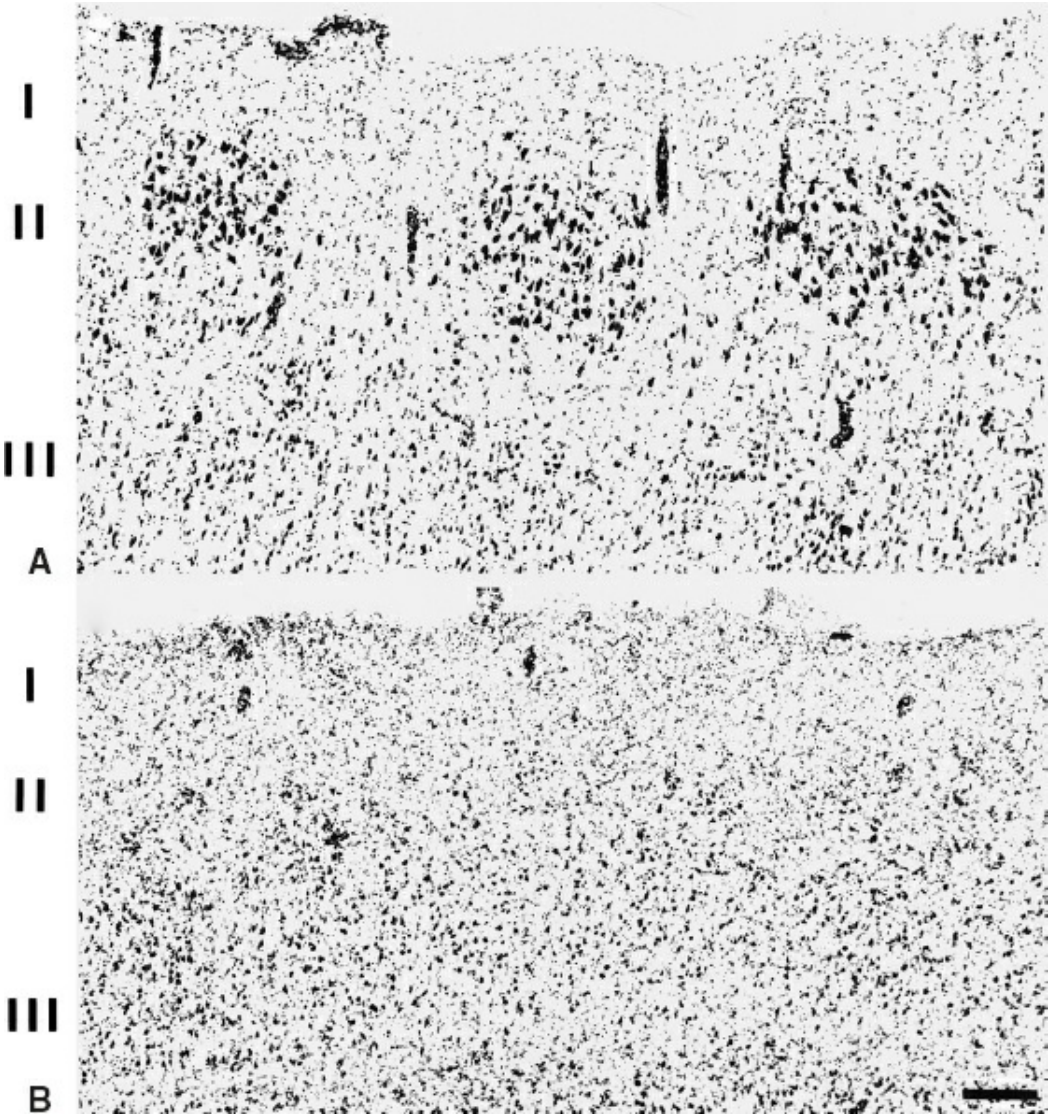

FIGURE 1.2–2. Nissl-stained sections of the superficial layers of the intermediate region of human entorhinal cortex. **A:** In the control brain, layer II contains clusters or islands of large, intensely stained neurons. **B:** In Alzheimer disease, these layer II neurons are particularly vulnerable to degeneration, and their loss produces a marked change in the cytoarchitecture of the region. Roman numerals indicate the location of the cortical layers. Calibration bar (200 µm) applies to **A** and **B.**

Immunohistochemical and other related techniques—which identify the location of neurotransmitters, their synthetic enzymes, or other molecules within neurons—can be used to determine the *chemoarchitecture* of the brain (Fig. 1.2–3B). In some cases, these techniques reveal striking regional differences in the chemoarchitecture of the brain that are difficult to detect in cytoarchitecture.

Connections

Every function of the human brain is a consequence of the activity of specific neural circuits. The circuits form as a result of several developmental processes. First, each neuron extends an axon, either after it

has migrated to its final location or, in some cases, before. The growth of an axon along distinct pathways is guided by molecular cues from its environment and eventually leads to the formation of synapses with specific target neurons. Although the projection of axons is quite precise, some axons initially produce an excessive number of axon branches, or *collaterals,* and contact a broader set of targets than are present in the adult brain. During later adolescence, the connections of particular neurons are focused by the pruning or elimination of axonal projections to inappropriate targets. The developmental timing of synaptic and axonal elimination seems to be highly specific across regions of the brain.

Within the adult brain, the connections among neurons or neural circuits follow several important principles of organization. First, many connections between brain regions are *reciprocal,* that is, each region tends to receive input from the regions to which it sends axonal projections. In some cases, the axons arising from one region may directly innervate the reciprocating projection neurons in another region; in other cases, local circuit interneurons are interposed between the incoming axons and the projection neurons that furnish the reciprocal connections. For some projections, the reciprocating connection is indirect, passing through one or more additional brain regions and synapses before innervating the initial brain region. In addition, connections within brain regions also display reciprocity. For example, in monkey prefrontal cortex, tract-tracing studies have shown that the axons and cell bodies of pyramidal neurons in layers II and III are arranged in a series of discrete stripes (Fig. 1.2–4). Reciprocity in this system is represented by the coregistration of anterogradely labeled axons and retrogradely labeled neurons within individual intrinsic and associational stripes. In addition, anterogradely labeled axon terminals form asymmetric synapses onto retrogradely labeled dendritic spines within individual stripes, providing further evidence of reciprocity in these connections.

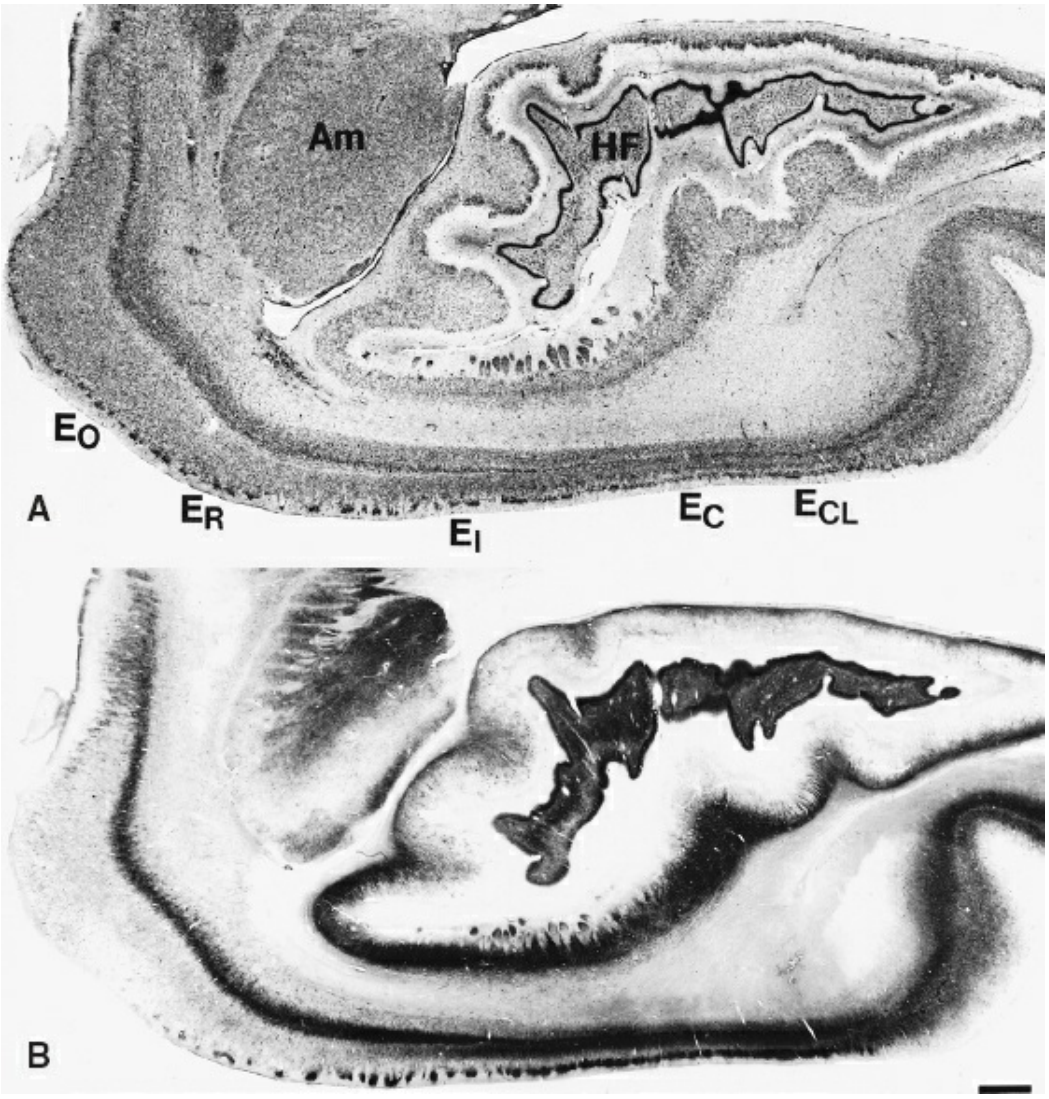

FIGURE 1.2–3. Adjacent sagittal sections through the medial temporal lobe of the human brain labeled to reveal the cytoarchitecture (**A**—Nissl stain) and chemoarchitecture (**B**—nonphosphorylated neurofilament protein immunoreactivity) of the entorhinal cortex. *Letters* indicate some of its subdivisions. Am, amygdala; HF, hippocampal formation. Calibration bar (2 mm) applies to both panels. (From Beall MJ, Lewis DA. Heterogeneity of layer II neurons in human entorhinal cortex. *J Comp Neurol*. 1992;321:241. Used with permission.)

Second, many neuronal connections are either divergent or convergent in nature. A *divergent* system involves the conduct of information from one neuron or a discrete group of neurons to a much larger number of neurons that may be located in diverse portions of the brain. The *locus coeruleus,* a small group of norepinephrine-containing neurons in the brainstem that sends axonal projections to the entire cerebral cortex and other brain regions, is an example of a highly divergent system. In contrast, the output of multiple brain regions may be directed toward a single area, forming a *convergent* system. Projection from multiple association areas of the cerebral cortex to the entorhinal region of the medial temporal lobe (MTL) is an example of a convergent system. Connections within brain regions

also display divergence and convergence (Fig. 1.2–4). For example, in monkey prefrontal cortex, pyramidal neurons within an individual stripe have axons that project to several other stripes (divergence), and individual stripes receive input from more than one stripe (convergence). The divergence in this system may provide an anatomic substrate that would allow a spatially restricted input to recruit a group of neurons whose coordinated activation is necessary to generate a particular response. Convergence in this system could allow information from different modalities present in the array of stripes to be relayed to a single location, facilitating the integration of their information content.

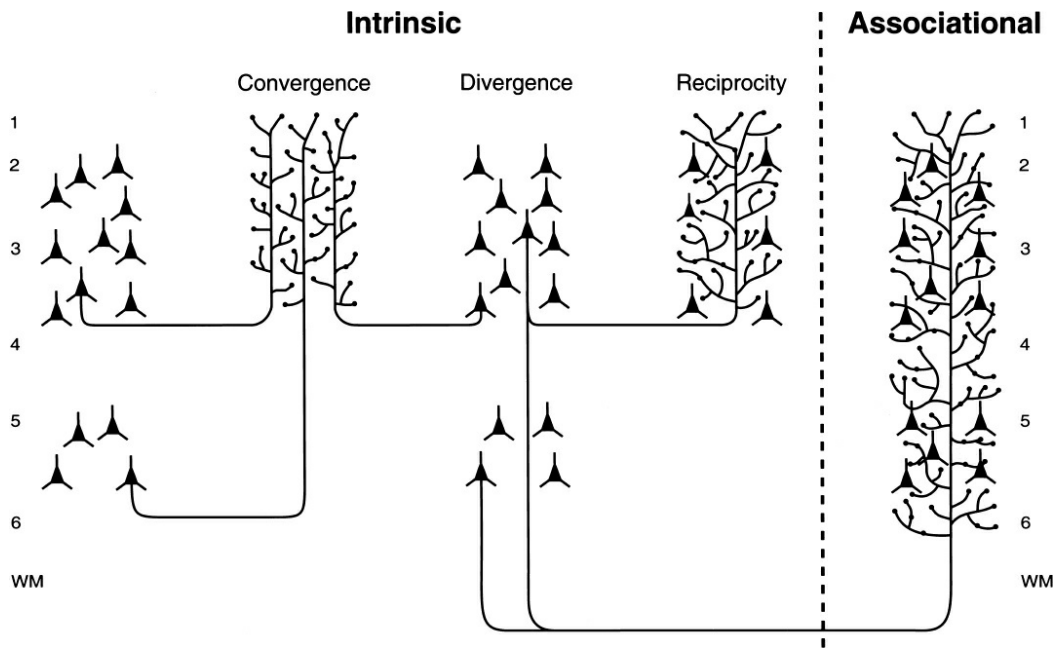

FIGURE 1.2–4. Schematic drawing comparing the spatial organization of intrinsic and associational connections and the reciprocity, convergence, and divergence in these connections in monkey prefrontal cortex. (Adapted from Pucak ML, Levitt JB, Lund JS, Lewis DA. Patterns of intrinsic and associational circuitry in monkey prefrontal cortex. *J Comp Neurol*. 1996;376:614.)

Third, the connections among regions may be organized in a hierarchical or parallel fashion, or both. Visual input is conveyed in a serial or hierarchical fashion through several populations of neurons in the retina to the lateral geniculate nucleus, to the primary visual cortex, and then, progressively, to the multiple visual association areas of the cerebral cortex. Within the hierarchical scheme, different types of visual information (e.g., motion and form) may be processed in a parallel fashion through different portions of the visual system.

Finally, regions of the brain are specialized for different functions. Lesions of the left inferior frontal gyrus (Broca area) (Fig. 1.2–5) produce a characteristic impairment in speech production. However, speech is a complex faculty that depends not only on the integrity of Broca area but

also on the distributed processing of information across numerous brain regions through divergent and convergent, serial and parallel interconnections. Thus, the role of any particular brain region or group of neurons in the production of specific behaviors or in the pathophysiology of a given neuropsychiatric disorder cannot be viewed in isolation, but must be considered within the context of the neural circuits connecting the neurons with other brain regions.

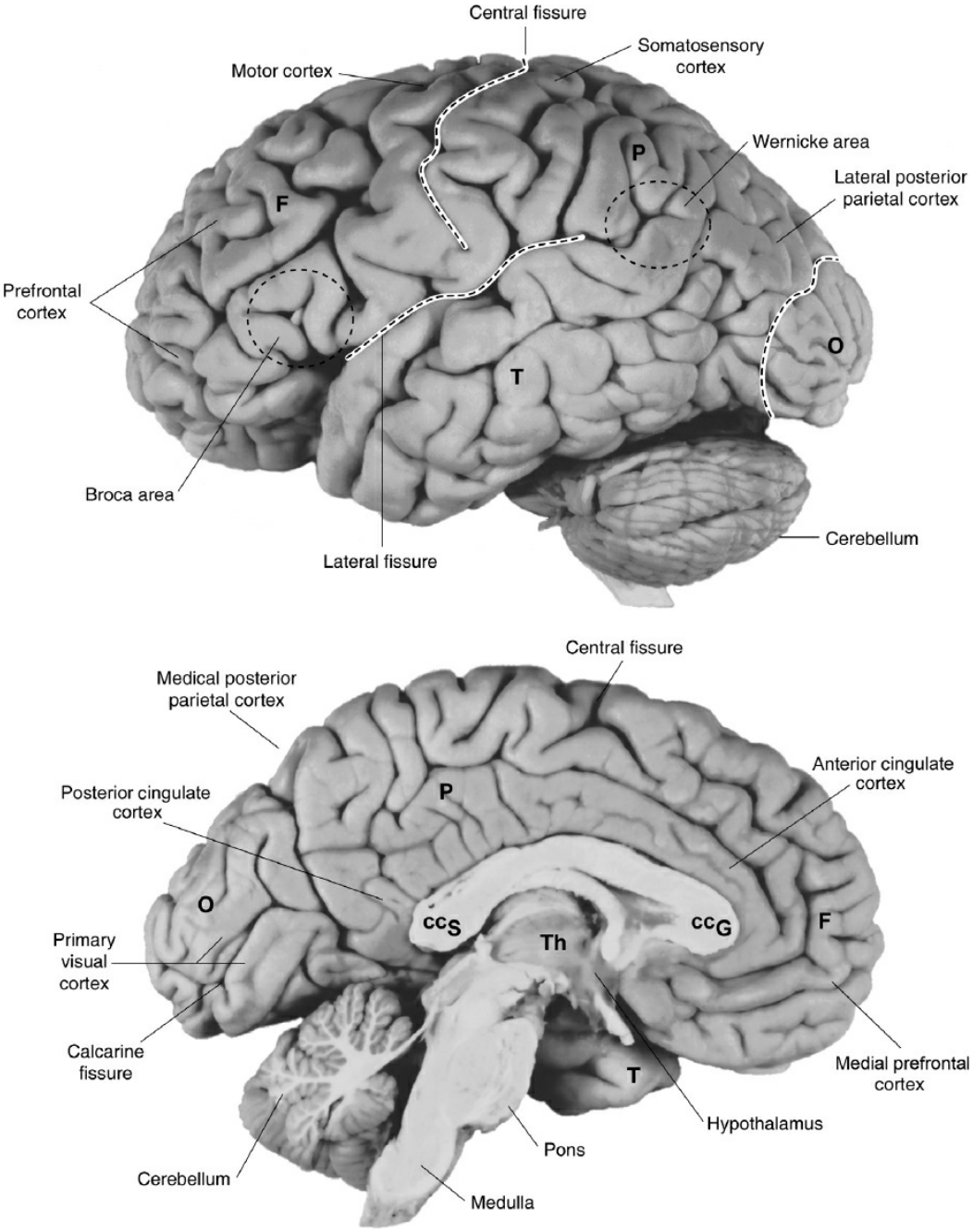

FIGURE 1.2–5. Photographs of the lateral (*top*) and medial (*bottom*) aspects of the left hemisphere of a human brain indicating the location of major surface landmarks. F, frontal lobe; O, occipital lobe; P, parietal lobe; T, temporal lobe; Th, thalamus; cc_G, genu of the corpus callosum; cc_S, splenium of the corpus callosum.

Distinctiveness of the Human Brain

Compared with the brains of other primate species, the human brain is substantially greater in size, with certain areas expanded disproportionately. The prefrontal cortex has been estimated to occupy only 3.5 percent of the total cortical volume in cats and 11.5 percent in monkeys, but close to 30 percent of the much larger cortical volume of the human brain. Conversely, the relative representation of other regions is decreased in the human brain; for example, the primary visual cortex accounts for only 1.5 percent of the total area of the cerebral cortex in humans, but in monkeys a much greater proportion (17 percent) of the cerebral cortex is devoted to this region. Thus, the distinctiveness of the human brain is attributable to its size and to the differential expansion of certain regions, particularly the areas of the cerebral cortex devoted to higher cognitive functions.

In addition, the expansion and differentiation of the human brain are associated with substantial differences in the organization of certain elements of neural circuitry. For example, compared with rodents, the dopaminergic innervation of the human cerebral cortex is much more widespread and regionally specific. The primary motor cortex and certain posterior parietal regions receive a dense dopamine innervation in monkeys and humans, but these areas receive little dopamine input in rats. The human brain is also distinctive at the cellular level, with greater relative percentages of certain neuronal and glial cells in the cerebral cortex. For example, the relative proportion of GABAergic interneurons is ~15 percent in rodent cerebral cortex but in humans, nearly 25 percent of all neurons in most cortical areas are GABAergic. Intralaminar and polarized astrocytes are found exclusively in humans and human protoplasmic astrocytes are larger and have more processes than their rodent counterparts. Furthermore, the astrocyte to neuron ratio is fivefold greater in humans compared to rodents, leading to greater astrocytic coverage of neurons and synapses and thus possibly enhanced regulation of neurotransmission in human cerebral cortex. These findings suggest a greater contribution of astrocytes in human, relative to rodent, brain function.

These types of species differences indicate that there are limits to the accuracy of generalizations made concerning human brain function when using studies in rodents or even nonhuman primates (NHPs) as the basis for the inference. Direct investigation of the organization of the human brain, however, is obviously restricted and complicated by numerous factors.

As indicated earlier, the expansion of the human brain is associated with the appearance of additional regions of the cerebral cortex. For example, the entorhinal cortex of the MTL in humans is sometimes

considered to be a single cortical region, but the cytoarchitecture and chemoarchitecture of this cortex differ substantially along its rostral–caudal extent (Fig. 1.2–3). It is tempting to identify these regions by their location relative to other structures, but sufficient interindividual variability exists in the human brain to make such a topologic definition unreliable. In the case of the entorhinal cortex, the location of its different subdivisions relative to adjacent structures, such as the amygdala and the hippocampus, varies across human brains. Therefore, in all studies, particularly studies using the human brain, areas of interest must be defined in a manner (for example, using cyto-, chemo-, or myeloarchitectural features) that allows investigators to accurately identify the same region in all cases.

An additional limitation to the study of the human brain concerns the changes in morphology and biochemistry that can occur during the interval between the time of death and the freezing or fixation of brain specimens. In addition to the influence of the known postmortem interval, such changes may begin to occur during the agonal state preceding death. When comparing aspects of the organization of the human brain with that of other species, the researcher must try to account for changes that may have occurred in the human brain as a result of postmortem delay or agonal state. In the study of disease states, appropriate controls must be used because differences in neurotransmitter content or other characteristics among cases could be a result of factors other than the disease state, such as methods of tissue preparation. Studies of the human brain in vivo—using such imaging techniques as positron emission tomography (PET), magnetic resonance imaging (MRI), magnetic resonance spectroscopy (MRS), and diffusion tensor imaging (DTI)—circumvent many of these problems, but are limited by insufficient resolution for the study of many aspects of human brain organization.

STRUCTURAL COMPONENTS

Major Brain Structures

In the early stages of human brain development, three primary vesicles can be identified in the neural tube: the *prosencephalon,* the *mesencephalon,* and the *rhombencephalon* (Fig. 1.2–6). Subsequently, the prosencephalon divides to become the *telencephalon* and the *diencephalon*. The telencephalon gives rise to the cerebral cortex, the hippocampal formation, the amygdala, and some components of the basal ganglia. The diencephalon becomes the thalamus, the hypothalamus, and several other related structures. The mesencephalon gives rise to the midbrain structures of the adult brain. The rhombencephalon divides into the *metencephalon* and the *myelencephalon*. The metencephalon gives rise to the pons and the cerebellum; the medulla is the derivative of the myelencephalon.

The cerebral cortex of each hemisphere is divided into four major regions: the *frontal, parietal, temporal,* and *occipital* lobes (Fig. 1.2–5). The frontal lobe is located anterior to the central sulcus and consists of the primary motor, premotor, and prefrontal regions (Fig. 1.2–7). The prefrontal cortex can be divided into dorsolateral and ventrolateral regions, with each of these regions having different functional properties. For example, the dorsolateral prefrontal cortex (DLPFC) seems to be more involved in the manipulation of data during working memory tasks than the ventrolateral prefrontal cortex (VLPFC), which seems to be more involved with pure maintenance of information during working memory. The primary somatosensory cortex is located in the anterior parietal lobe; in addition, other cortical regions related to complex visual and somatosensory functions are located in the posterior parietal lobe. The superior portion of the temporal lobe contains the primary auditory cortex and other auditory regions; the inferior portion contains regions devoted to complex visual functions. In addition, some regions within the superior temporal sulcus receive a convergence of input from the visual, somatosensory, and auditory sensory areas. The occipital lobe consists of the primary visual cortex and other visual association areas.

Beneath the outer mantle of the cerebral cortex are many other major brain structures, such as the caudate nucleus, the putamen, and the globus pallidus (Fig. 1.2–8). These structures are components of the basal ganglia, a system involved in the control of movement and certain cognitive processes. The hippocampus and the amygdala, components of the limbic system, are located deep in the medial temporal lobe (Figs. 1.2–9, 1.2–10, and 1.2–11). In addition, the derivatives of the diencephalon, such as the thalamus and the hypothalamus, are prominent internal structures; the thalamus is a relatively large structure composed of numerous nuclei that have distinct patterns of connectivity with the cerebral cortex (Figs. 1.2–9, 1.2–10, and 1.2–11). In contrast, the hypothalamus is a much smaller structure involved in autonomic and endocrine functions.

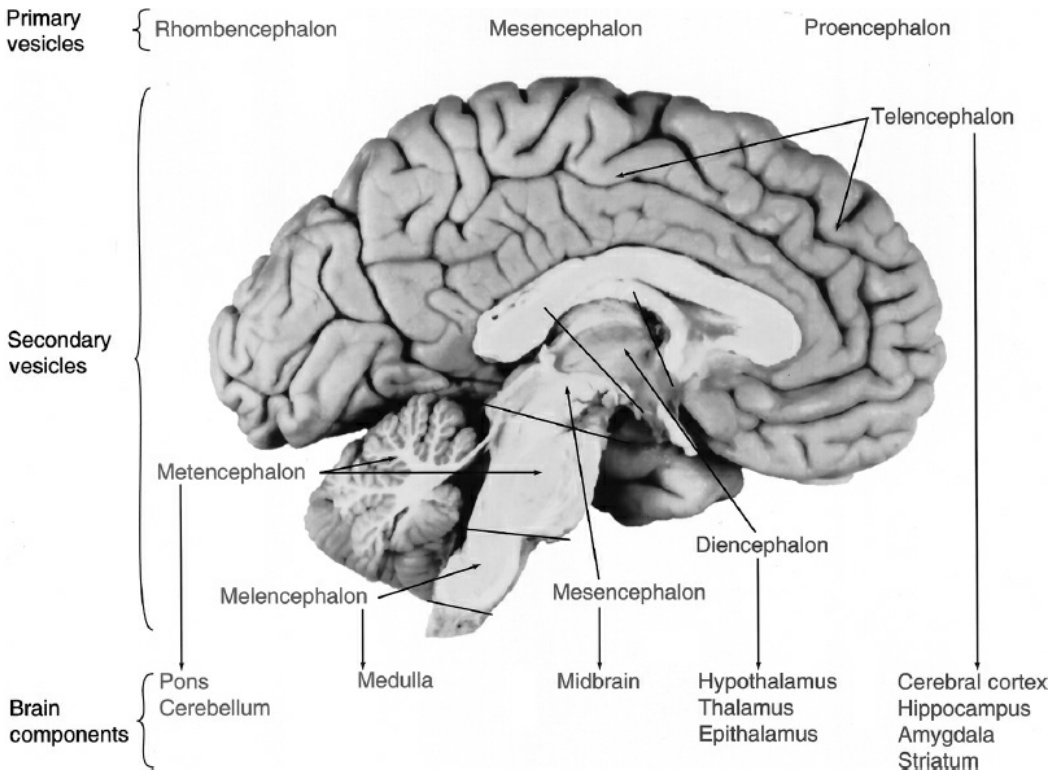

FIGURE 1.2–6. Schematic representation of the primary vesicles of the neural tube and their derivatives.

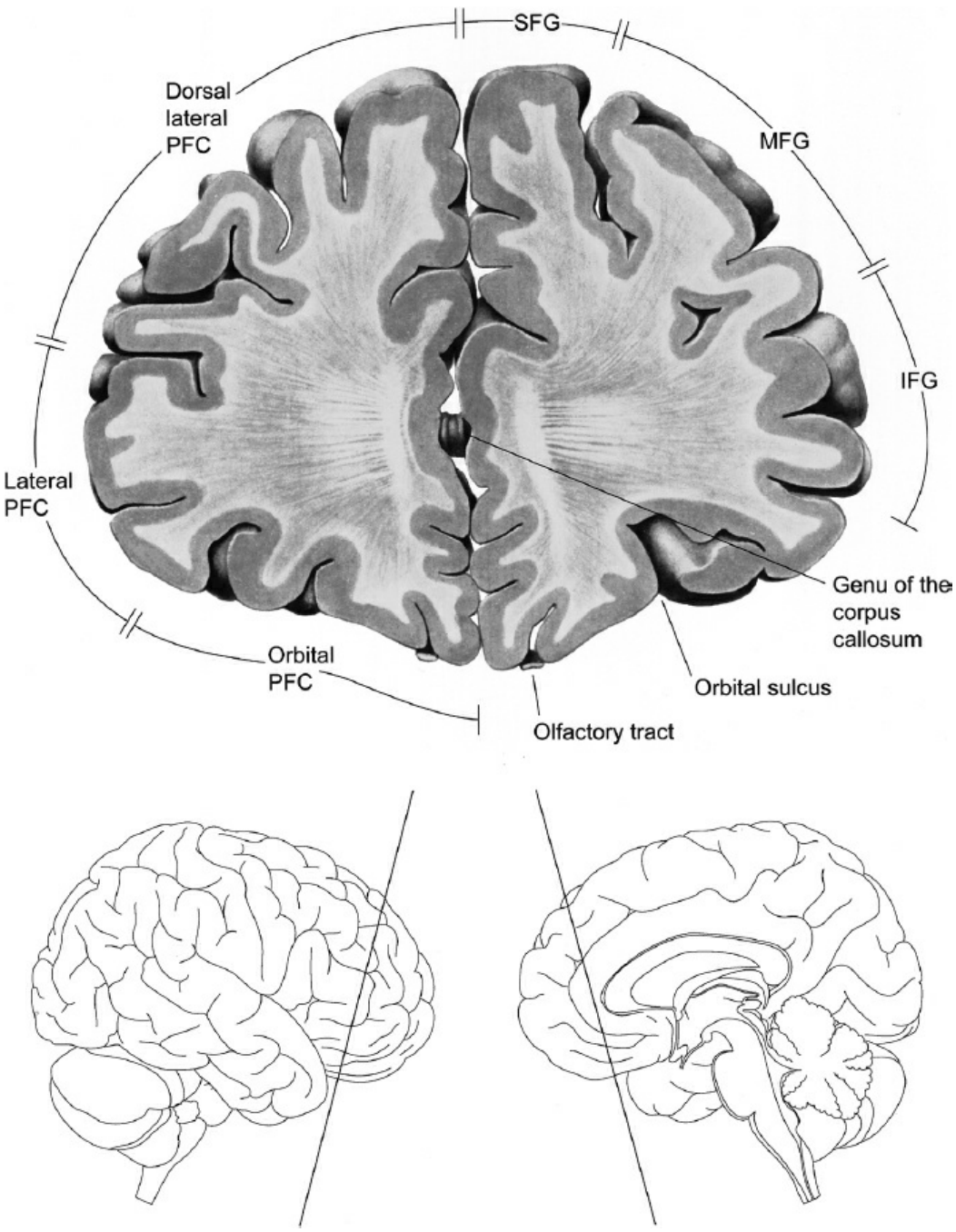

FIGURE 1.2–7. Drawing of a coronal section just anterior to the genu of the corpus callosum of a human brain. The *inset* below indicates the level of the section. IFG, inferior frontal gyrus; MFG, middle frontal gyrus; PFC, prefrontal cortex; SFG, superior frontal gyrus. (Adapted from Nieuwenhuys R, Voogd J, van Huijzen C. *The Human Central Nervous System: A Synopsis and Atlas*, 3rd ed. New York: Springer; 1988:68.)

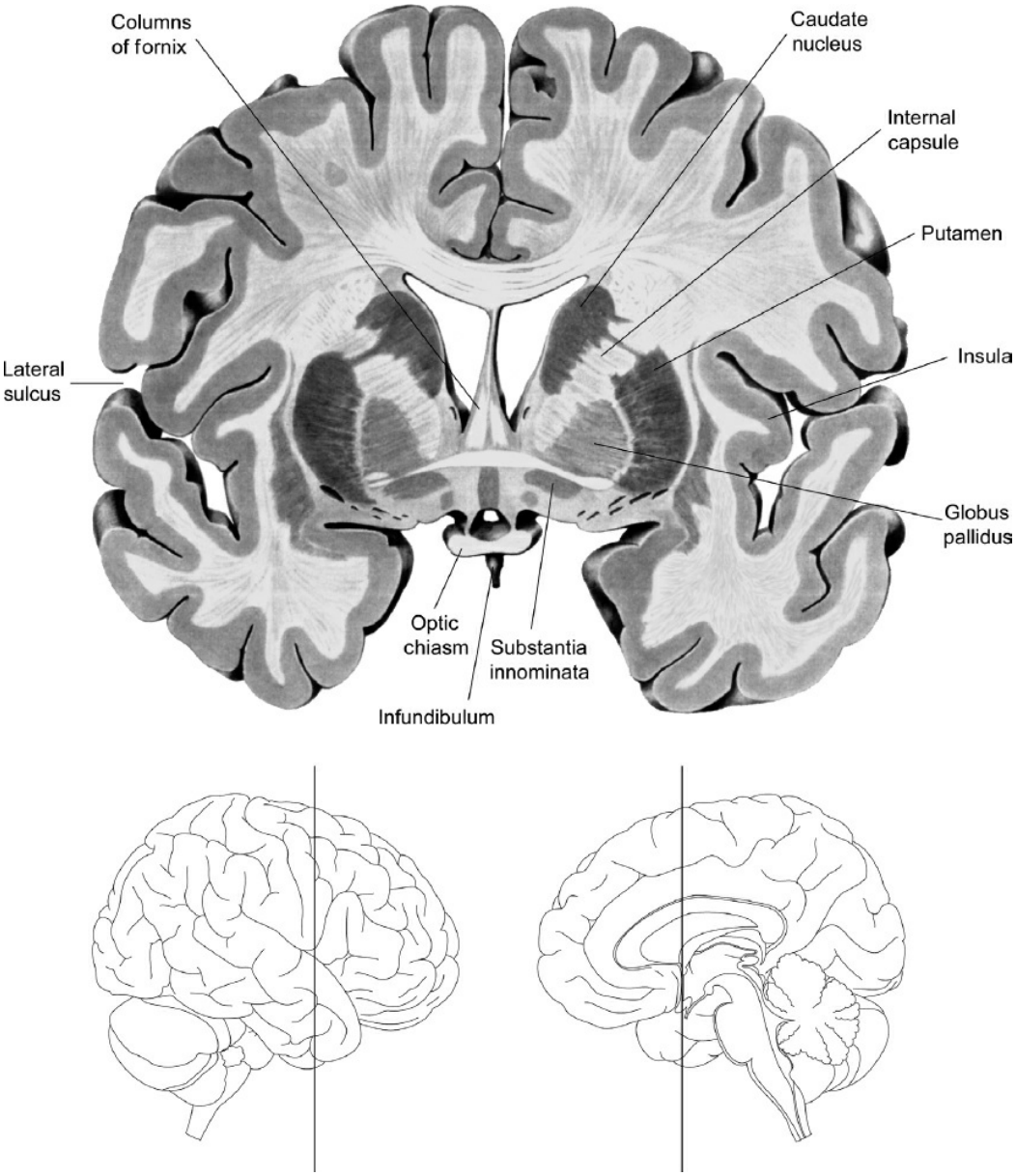

FIGURE 1.2–8. Drawing of a coronal section through the optic chiasm of a human brain. The *inset* below indicates the level of the section. (Adapted from Nieuwenhuys R, Voogd J, van Huijzen C. *The Human Central Nervous System: A Synopsis and Atlas*, 3rd ed. New York: Springer; 1988:70.)

White Matter Tracts

The cerebral hemispheres contain billions of myelinated axons or fibers, giving the white matter its characteristic color, which carry information to and from the cerebral cortex. These axons are bundled into white matter tracts that include projection, commissural, and associational fibers.

Projection Fibers. Two of the major projection fiber systems comprise fibers that originate in the cerebral cortex and project to subcortical targets (corticofugal) and fibers that originate outside of the telencephalon and project to the cerebral cortex (corticopetal). Examples of these are the

corticothalamic and thalamocortical projections, respectively. These projection fibers travel through the *internal capsule,* a compact bundle of fibers that is structurally associated with the thalamus and lenticular nucleus (i.e., the putamen and globus pallidus considered as one structure). In each cerebral hemisphere, the internal capsule is bordered laterally by the lenticular nucleus and medially by the thalamus and head of the caudate (Fig. 1.2–12). Other fiber systems, such as the corticopontine, corticospinal, and corticobulbar tracts, descend from the cortex through the internal capsule and cerebral peduncle to reach their destinations in the pons, spinal cord, and brainstem. All the fibers traveling through the internal capsule form the *corona radiata,* a fan-like structure that sits just above the internal capsule.

The internal capsule has been divided into five regions with the name and location of each region based on its relationship with the lenticular nucleus. In addition, each region of the internal capsule contains different fiber systems. The *anterior limb* lies between the lenticular nucleus and the head of the caudate and carries frontopontine fibers and fibers interconnecting the thalamus and the frontal cortex. The *posterior limb,* the largest component, is located between the lenticular nucleus and the thalamus and conveys corticospinal fibers. The *genu* is the intersection of the anterior and posterior limbs and carries corticobulbar fibers. The *retrolenticular limb* lies behind or posterior to the lenticular nucleus, and fibers in this region of the internal capsule form the bulk of the *optic radiation,* the large group of fibers projecting from the lateral geniculate thalamic nucleus to the primary visual cortex in the occipital lobe. The *sublenticular limb* lies inferior to the lenticular nucleus and contains fibers of the *auditory radiation,* a collection of fibers connecting the medial geniculate thalamic nucleus with primary auditory cortex in the temporal lobe.

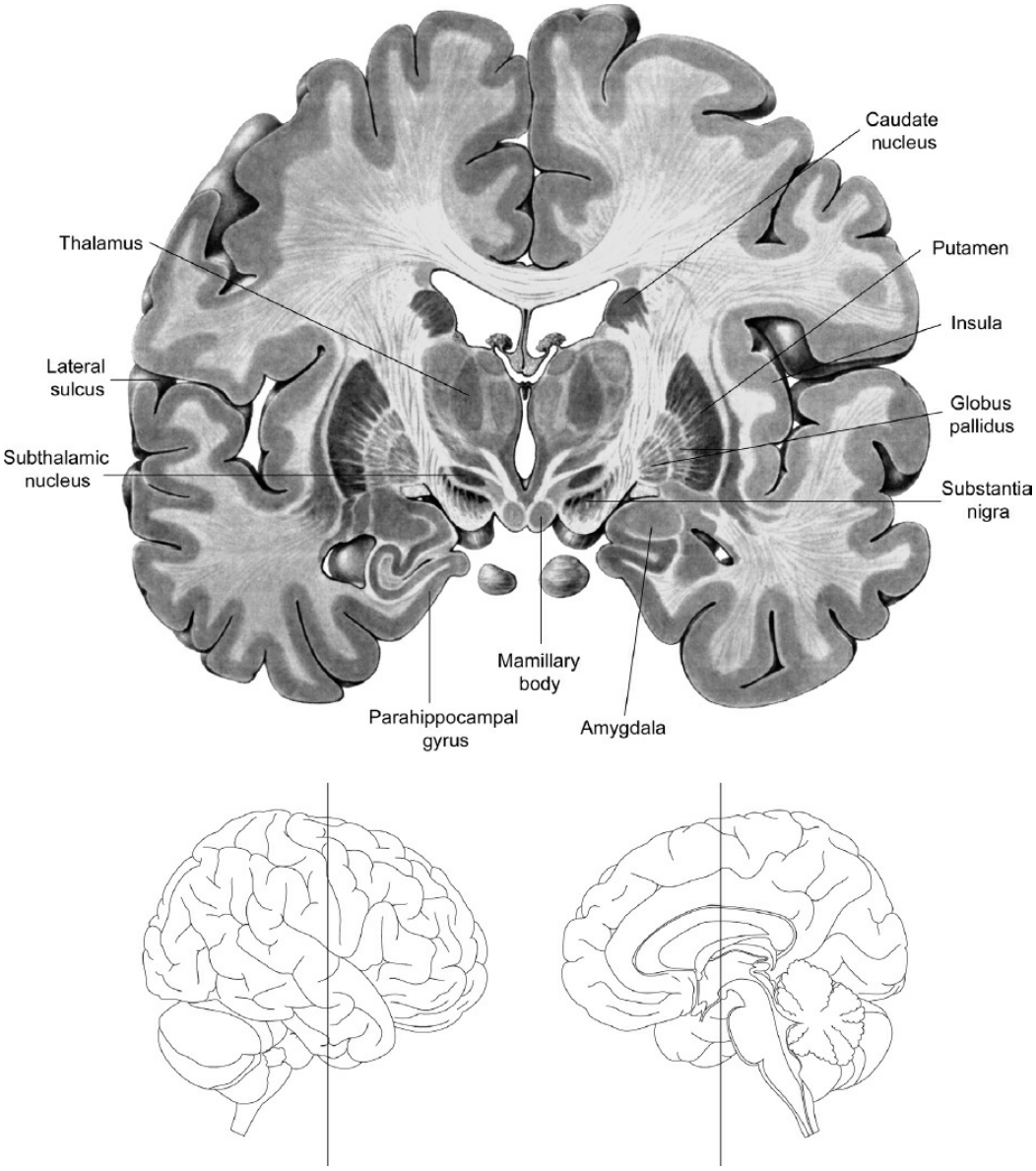

FIGURE 1.2–9. Drawing of a coronal section at the level of the mammillary bodies. The *inset* below indicates the level of the section. (Adapted from Nieuwenhuys R, Voogd J, van Huijzen C. *The Human Central Nervous System: A Synopsis and Atlas*, 3rd ed. New York: Springer; 1988:72.)

Commissural Fibers. Commissural fibers interconnect areas in the two cerebral hemispheres with each other. The two main commissural fiber systems are the corpus callosum and the anterior commissure. The *corpus callosum* is the largest fiber bundle in the brain, containing roughly 300 million axons. Most of these axons interconnect cortical regions in one lobe with homotopic (i.e., similarly placed) regions in the opposite lobe. However, heterotopic connections (i.e., those that link dissimilar cortical regions) are also carried in the corpus callosum. Almost all cortical regions are connected via the corpus callosum with the notable exceptions of the hand area of the motor and somatosensory cortices and all of the primary visual cortex except the portion representing areas adjacent to the vertical

midline.

The corpus callosum consists of five parts (Fig. 1.2–13). The corpus callosum starts at the *rostrum* and then curves anteriorly and dorsally to form the *genu*. The *body* of the corpus callosum is the largest part and gives way to the *splenium,* the enlarged, rounded posterior end. Sometimes the narrow portion of the corpus callosum between the body and the splenium is referred to as the *isthmus*. Axons carrying higher-order cognitive and sensory information from the prefrontal, temporal, and parietal cortices primarily travel through the genu and splenium, whereas visual, auditory, and somatosensory information is carried predominantly in the body and isthmus of the corpus callosum.

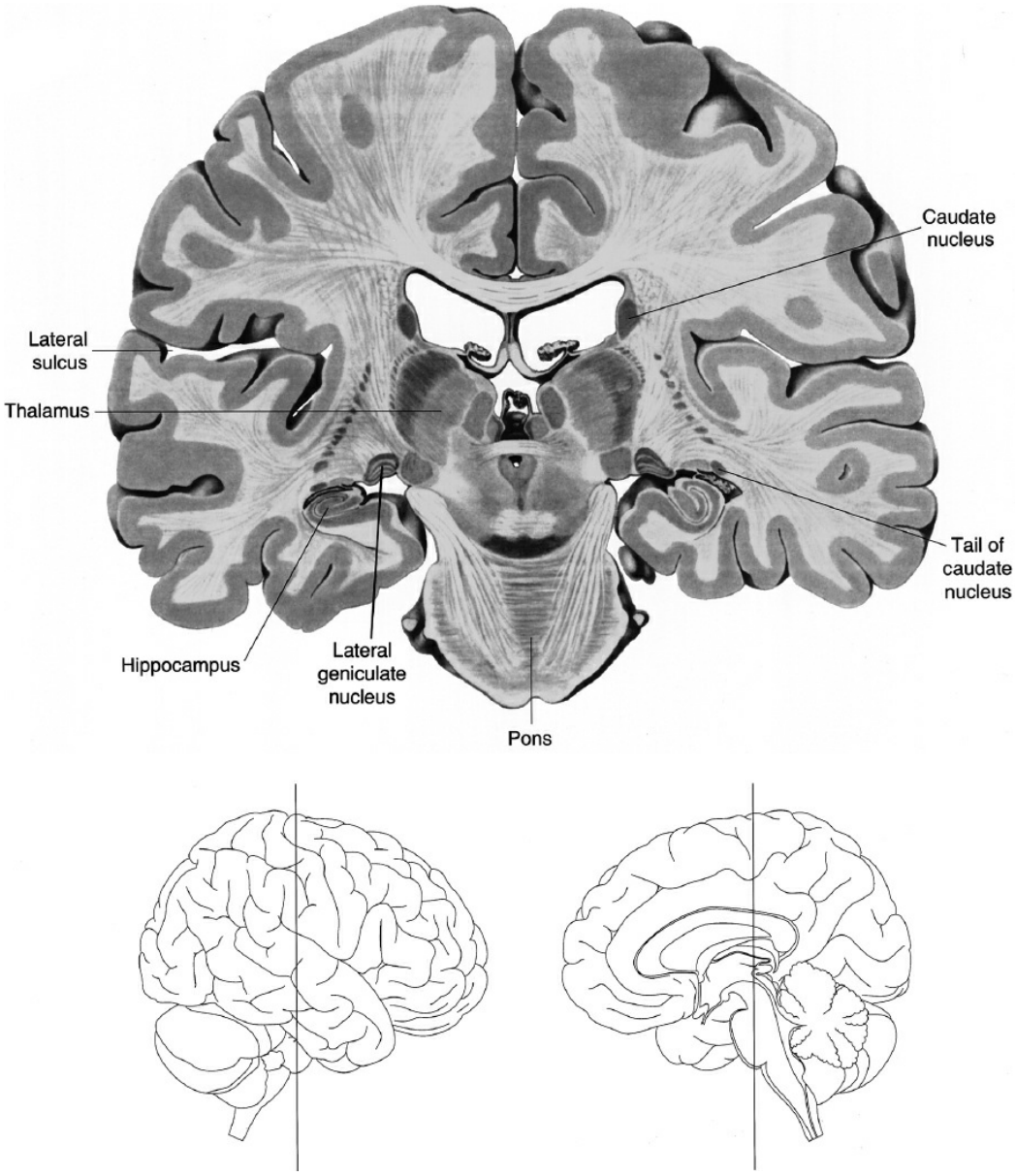

FIGURE 1.2–10. Drawing of a coronal section through the posterior thalamus. The *inset* below indicates the level of the section. (Adapted from Nieuwenhuys R, Voogd J, van Huijzen C. *The Human Central Nervous System: A Synopsis and Atlas*, 3rd ed. New York: Springer; 1988:74.)

The anterior commissure is a compact bundle of fibers that is caudal to the corpus callosum and crosses the midline in front of the fornix (Fig. 1.2–8). The anterior commissure interconnects areas in the two temporal lobes and fibers from the anterior olfactory nucleus. Smaller commissural fiber tracts include the posterior commissure, which connects caudal portions of the diencephalon, and the hippocampal commissure, which interconnects the two hippocampal formations.

Associational Fibers. Associational fibers connect cortical areas within a hemisphere and range in size from very short fibers that connect areas within the same lobe to longer fibers that connect areas within different lobes. *Short association fibers* connect adjacent gyri and are often called *U fibers* because they form a U connecting one gyrus to another gyrus (Fig. 1.2–14). There are five major tracts of *long association fibers* that connect distant cortical areas within the same hemisphere. The *superior longitudinal fasciculus* is located laterally within the hemisphere above the insula and connects frontal, parietal, and occipital cortices. The *arcuate fasciculus* interconnects the frontal and temporal lobes. The *uncinate fasciculus* is a curved fiber bundle that connects the orbital portion of the frontal lobe with the anterior region of the temporal lobe. As its name implies, the *inferior occipitofrontal fasciculus* connects the occipital and frontal lobe in a bundle of fibers that courses ventrally and laterally within the hemisphere. The *cingulum* lies within the white matter under the cingulate gyrus and connects this gyrus with the parahippocampal gyrus. The *inferior longitudinal fasciculus* connects the temporal and occipital lobes. These fiber bundles are not discrete, point-to-point pathways between cortical regions, but are continuous pathways with fibers entering and leaving all along their course. Other associational fiber bundles include the *external capsule,* which is sandwiched between the claustrum and the putamen, and the *extreme capsule,* which lies between the claustrum and the insular cortex (Fig. 1.2–8).

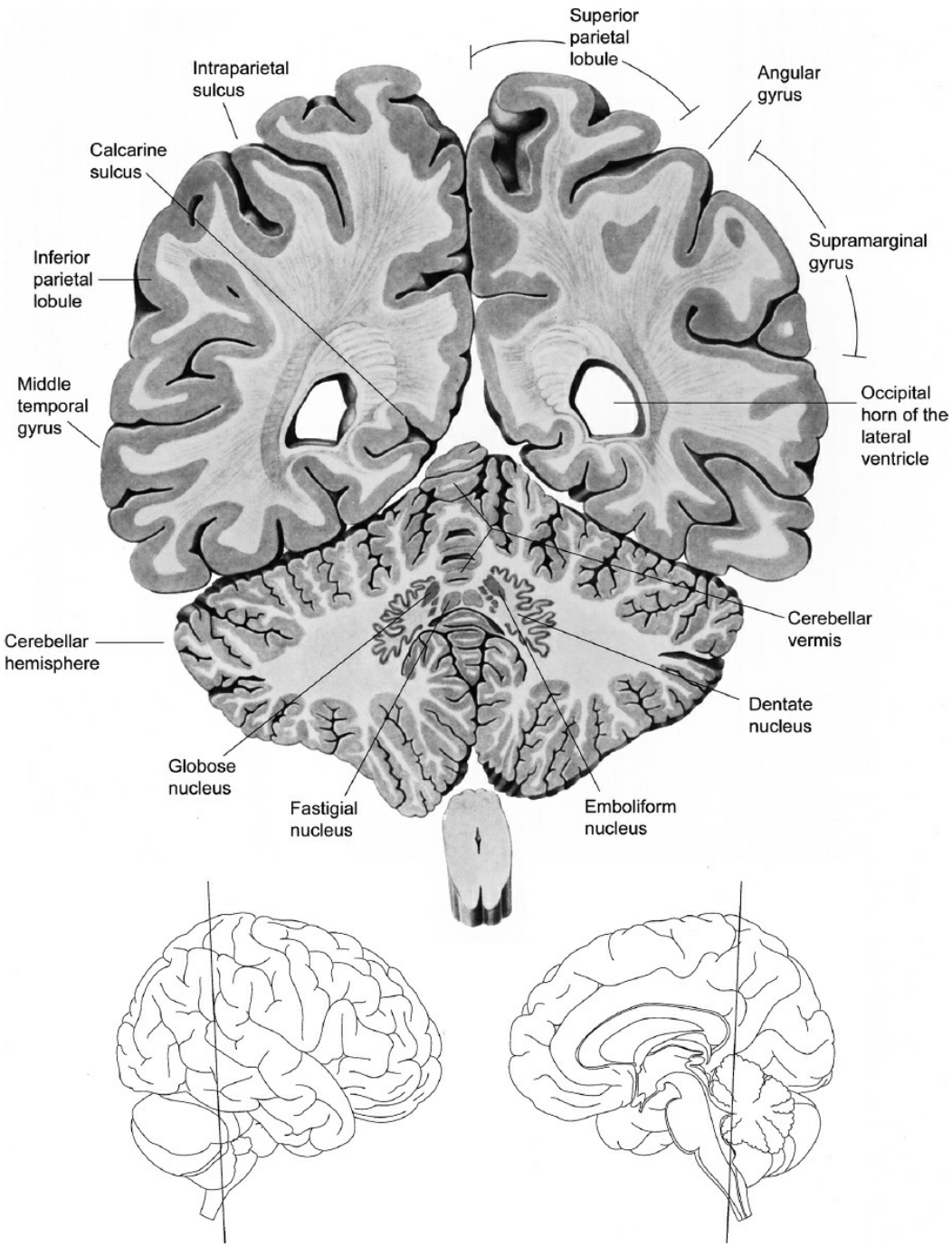

FIGURE 1.2–11. Drawing of a coronal section through the cerebral hemispheres just posterior to the splenium of the corpus callosum and through the deep nuclei of the cerebellum. The *inset* below indicates the level of the section. (Adapted from Nieuwenhuys R, Voogd J, van Huijzen C. *The Human Central Nervous System: A Synopsis and Atlas*, 3rd ed. New York: Springer; 1988:77.)

Disturbances in the connectivity within and between hemispheres have been implicated in the pathophysiology of schizophrenia. For example, MRI studies of individuals with schizophrenia have revealed decreases in white matter density in the corpus callosum, internal capsule, and anterior commissure. In addition, studies using DTI, which provides information on the organization and microstructure of tissue, have shown abnormalities in

almost every major white matter tract, including the corpus callosum, internal capsule, cingulum bundle, uncinate fasciculus, and fornix, in patients with schizophrenia, as well as altered microstructure in the corpus callosum and uncinate fasciculus in patients with depression. Abnormalities in white matter tracts have also been reported in other neuropsychiatric disorders. For example, MRI studies have revealed a reduction in the cross-sectional area of the corpus callosum in individuals with Alzheimer disease as well as in children with autism.

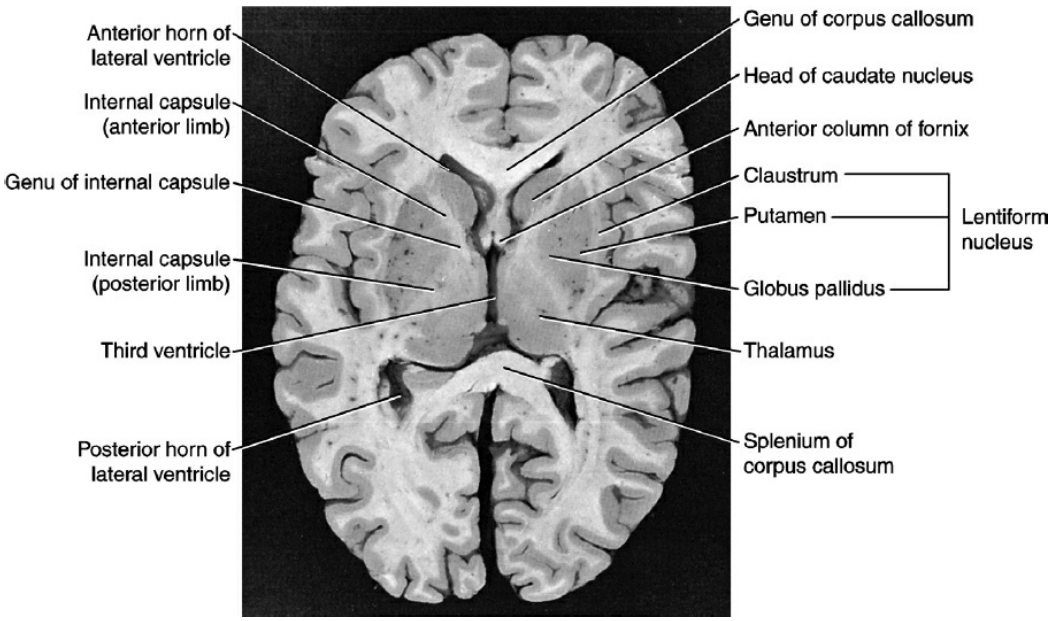

FIGURE 1.2–12. A horizontal section through the cerebrum shows the location of the internal capsule fibers in relation to components of the basal ganglia and the corpus callosum. (From Snell RS. *Clinical Neuroanatomy*, 7th ed. Philadelphia, PA: Wolters Kluwer; 2010:xv.)

Ventricular System

As the neural tube fuses during development, the cavity of the neural tube becomes the ventricular system of the brain. It is composed of two C-shaped *lateral ventricles* in the cerebral hemispheres that can be divided further into five parts: the anterior horn (which is located in the frontal lobe), the body of the ventricle, the inferior or temporal horn in the temporal lobe, the posterior or occipital horn in the occipital lobe, and the atrium (Fig. 1.2–15). The *foramina of Monro* (interventricular foramina) are the two apertures that connect the two lateral ventricles with the *third ventricle,* which is found on the midline of the diencephalon. The *cerebral aqueduct* connects the third ventricle with the *fourth ventricle* in the pons and the medulla.

The ventricular system is filled with cerebrospinal fluid (CSF), a colorless liquid containing low concentrations of protein, glucose, and potassium and relatively high concentrations of sodium and chloride. Most

(70 percent) of the CSF is produced at the choroid plexus located in the walls of the lateral ventricles and in the roof of the third and fourth ventricles. The *choroid plexus* is a complex of ependyma, pia, and capillaries that invaginate the ventricle. In contrast to other parts of the brain, the capillaries in the choroid plexus are fenestrated, which allows substances to pass out of the capillaries and through the pia mater. The ependymal or choroid epithelial cells, however, have tight junctions between cells to prevent the leakage of substances into the CSF; this provides what is sometimes referred to as the *blood–CSF barrier*. In other parts of the brain, the endothelial cells of the capillaries exhibit tight junctions that prevent the movement of substances from the blood to the brain; this is referred to as the *BBB*.

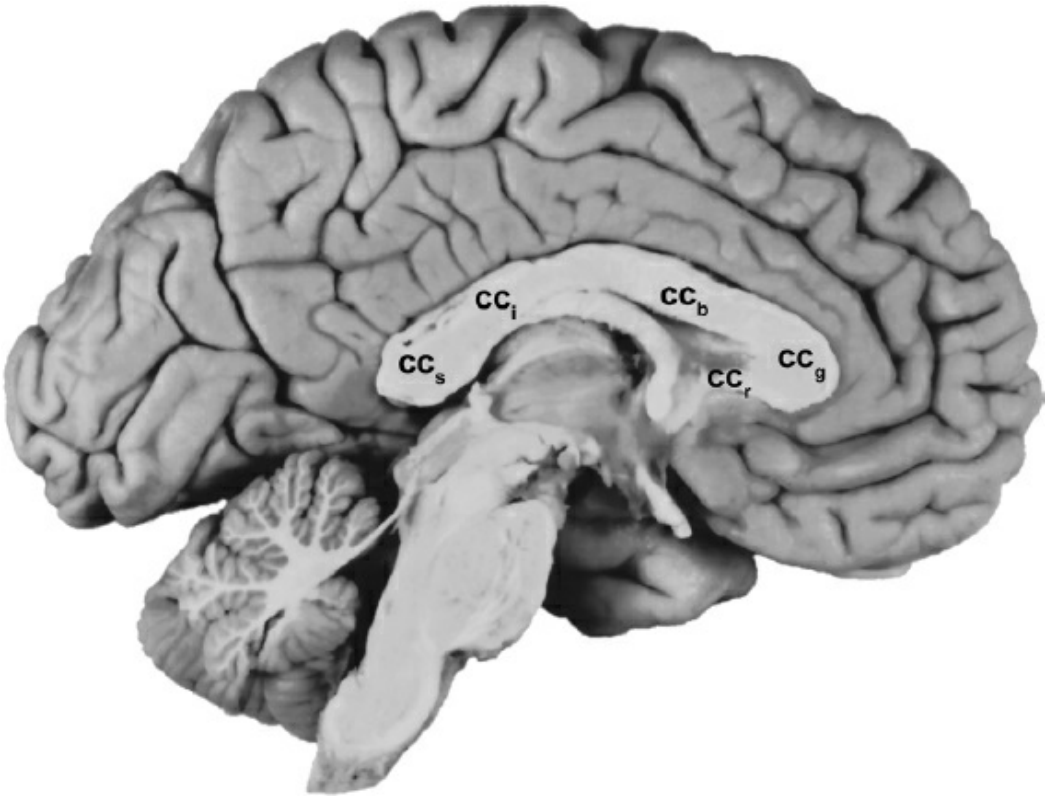

FIGURE 1.2–13. Photograph of the medial view of the right cerebral hemisphere of a human brain illustrating the different parts of the corpus callosum. cc_r, rostrum; cc_g, genu; cc_b, body; cc_i, isthmus; cc_s, splenium.

The CSF is constantly produced and circulates through the lateral ventricles to the third ventricle and then to the fourth ventricle. The CSF then flows through the medial and lateral apertures to the cisterna magna and pontine cistern and, finally, travels over the cerebral hemispheres to be absorbed by the arachnoid villi and released into the superior sagittal sinus. Disruptions in the flow of the CSF usually cause some form of hydrocephalus; for example, if an intraventricular foramen is occluded, the associated lateral ventricle becomes enlarged, but the remaining

components of the ventricular system remain normal.

Several functions are attributed to the CSF: it serves to cushion the brain against trauma, to maintain and control the extracellular environment, and to spread endocrine hormones. Because the CSF bathes the brain and is in direct communication with extracellular fluid, it is possible to measure the amount of certain compounds in the CSF as a correlate of the amount of that substance in the brain. For example, levels of homovanillic acid (HVA), a metabolite of the neurotransmitter dopamine, are thought to reflect the functional activity of that neurotransmitter. The concentration of HVA in samples of the CSF taken in a lumbar puncture may provide a picture of brain dopaminergic function. Because the CSF bathes the entire brain, however, the CSF levels of HVA may not be a valid indicator of the activity of dopamine neurons in any particular brain area. Consequently, caution must be exercised in interpreting the findings of investigations that rely on CSF measurements as indicators of neurotransmitter activity.

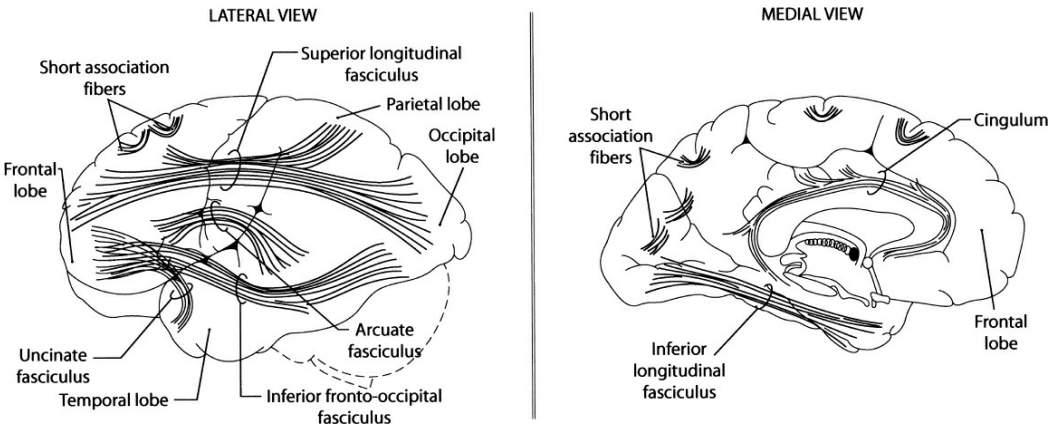

FIGURE 1.2–14. Drawings illustrating the main associational fiber tracts as visualized from lateral (*left panel*) and medial (*right panel*) aspects of the left hemisphere. (Adapted from Haines DE. *Fundamental Neuroscience for Basic and Clinical Applications*, 3rd ed. Philadelphia, PA: Churchill Livingstone; 2006:253.)

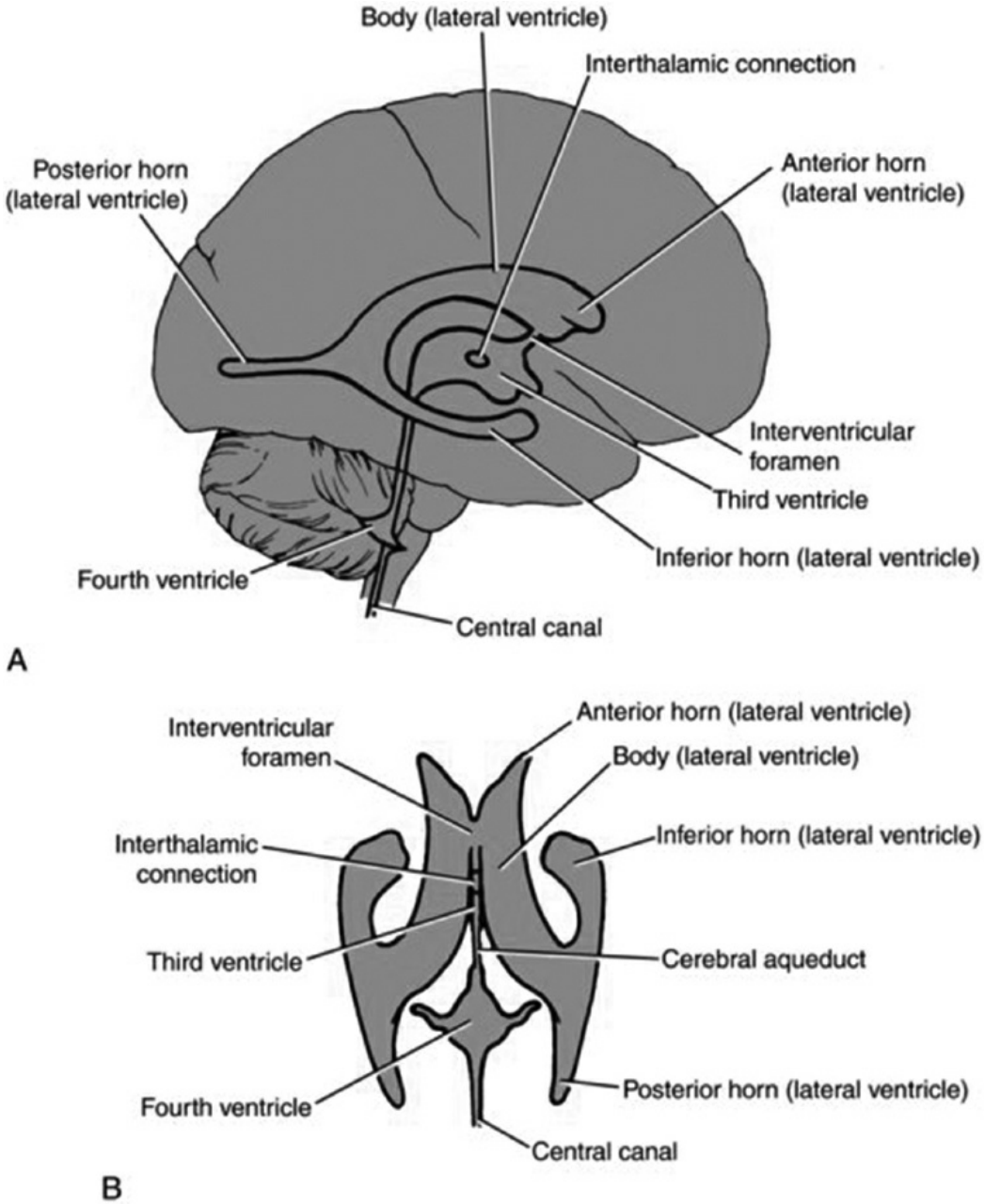

Figure 1.2–15. **A:** Diagram of the ventricles of the brain and central canal of the spinal cord. **B:** A three-dimensional representation of the ventricles of the brain. (From Snell RS. *Clinical Neuroanatomy*, 7th ed. Philadelphia, PA: Wolters Kluwer; 2010:469.)

FUNCTIONAL BRAIN SYSTEMS

The relationships between the organizational principles and the structural components of the human brain are illustrated in three functional systems: the thalamocortical, basal ganglia, and limbic systems.

Cerebral Cortex

The *cerebral cortex* is a laminated sheet of neurons, several millimeters thick, that covers the cerebral hemispheres. It consists of approximately 22.5 billion neurons communicating via approximately 165 trillion

synapses. These neurons have approximately 12 million kilometers of dendrites, and the cerebral cortex and subcortical regions are interconnected by approximately 100,000 km of axons. More than 90 percent of the total cortical area consists of the *neocortex,* which has a six-layered structure (at least at some point during development). The remainder of the cerebral cortex is referred to as the *allocortex* and consists of the *paleocortex* and the *archicortex,* regions that are restricted to the base of the telencephalon and the hippocampal formation, respectively.

Within the neocortex, the two major neuronal cell types are the pyramidal and stellate, or nonpyramidal, neurons (Fig. 1.2–16). *Pyramidal neurons,* which account for approximately 70 percent of all neocortical neurons, usually have a characteristically shaped cell body that gives rise to a single apical dendrite that ascends vertically toward the cortical surface. In addition, the neurons have an array of short dendrites that spread laterally from the base of the cell. The dendrites of pyramidal neurons are coated with short protrusions, called *spines,* which are the sites of most of the excitatory synapses to these neurons (Fig. 1.2–17). Most pyramidal cells are projection neurons that are thought to use excitatory amino acids as neurotransmitters. Interestingly, in postmortem studies, subjects with schizophrenia appear to have fewer spines on the dendrites located at the base of pyramidal neurons in deep layer III of the prefrontal cortex.

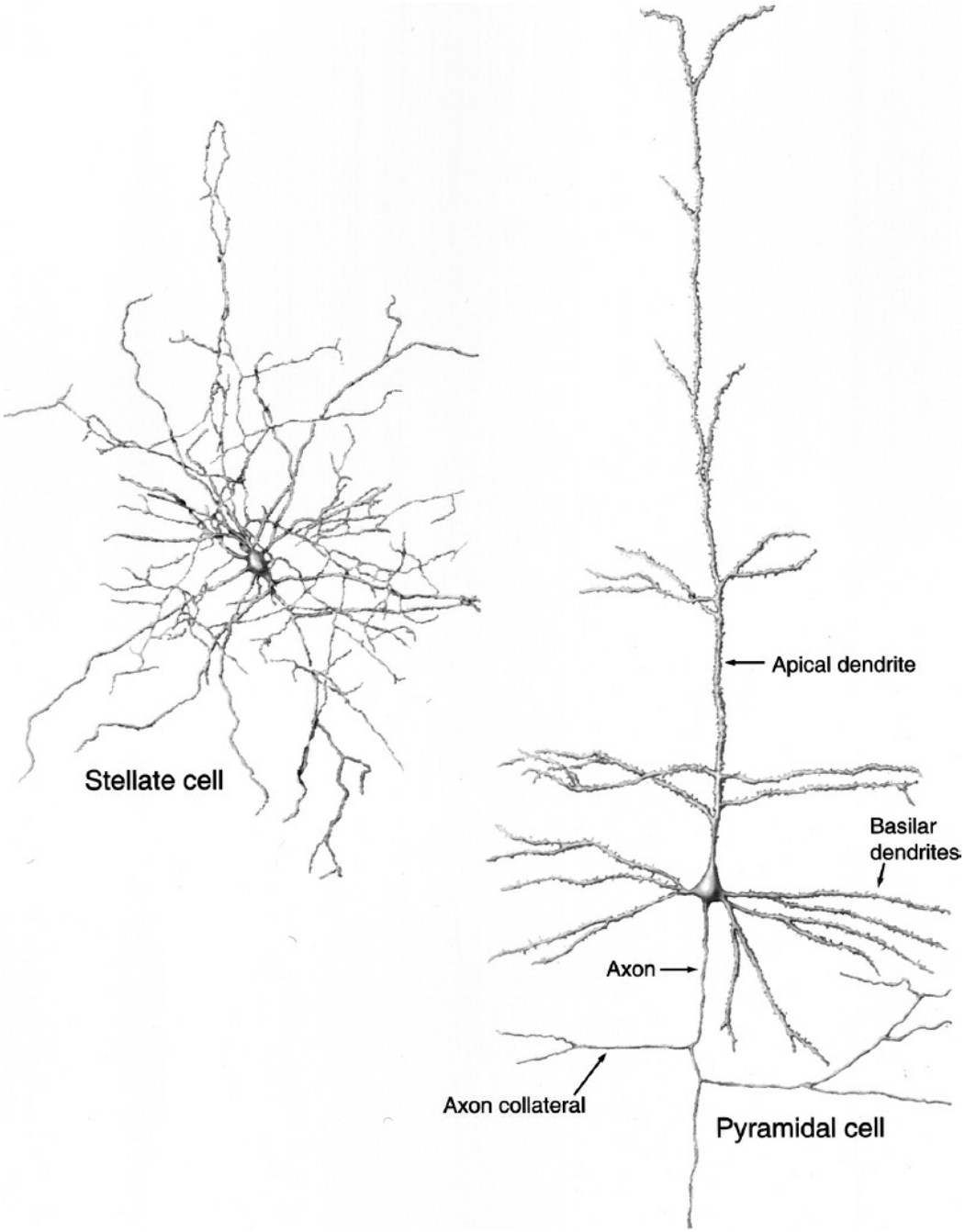

FIGURE 1.2–16. Drawings of a stellate neuron (*left*) and a pyramidal neuron (*right*). Note the difference in the morphology of these two types of neurons. The soma of stellate cells tends to be round or ovoid, whereas that of pyramidal neurons generally appears triangular from a two-dimensional perspective. Also, note the difference in the dendritic and axonal arbors between the two cells. The processes arising from the stellate cell appear to branch in multiple directions, whereas the pyramidal neuron has prominent, well-defined apical and basilar dendrites. Note the small protuberances visible on the apical and basilar dendrites; these are dendritic spines. (Adapted from Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain*. Philadelphia, PA: Lippincott Williams & Wilkins; 2001:45.)

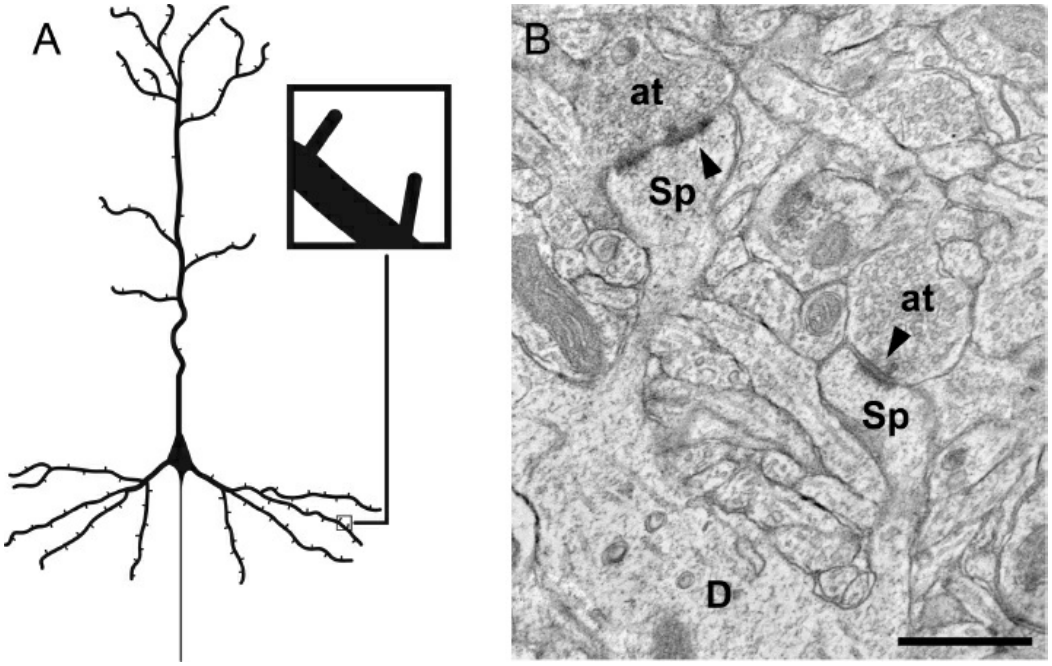

FIGURE 1.2–17. Dendritic spines in the cerebral cortex. **A:** Diagram of a cortical pyramidal neuron illustrating the location of dendritic spines. **B:** Electron micrograph from monkey prefrontal cortex showing two dendritic spines (*Sp*) emanating from a single dendrite (*D*), both of which receive an asymmetric synapse from an axon terminal (*at*). Calibration bar = 200 nm.

In contrast, *nonpyramidal cells* are generally small, local circuit neurons, many of which use the inhibitory neurotransmitter GABA (Fig. 1.2–18). Also known as interneurons, the axons of cortical GABA cells arborize within the gray matter and do not project out of the cortical region in which they reside. Twelve different subtypes of GABA neurons can be found in the cortex, and these can be distinguished biochemically, electrophysiologically, and morphologically. For example, subpopulations of GABA cells can be distinguished by the presence of certain neuropeptides or calcium-binding proteins. In addition, the organization of the axonal arbor and synaptic targets of the axon terminals differ greatly across these different subtypes. As depicted in Figure 1.2–18, the chandelier class of GABA cell expresses the calcium-binding protein parvalbumin and has axon terminals that are arrayed as distinct vertical structures termed cartridges (Fig. 1.2–19). These axon terminals form inhibitory or symmetric synapses exclusively with the axon initial segments of pyramidal cells. Parvalbumin-containing basket neurons form symmetric synapses onto the cell bodies and dendrites of pyramidal neurons. Parvalbumin-containing neurons are predominantly located in layers III and IV. In contrast, the Martinotti class of GABA neurons contain the neuropeptide somatostatin and form symmetric synapses onto the tuft dendrites of pyramidal neurons. Some double-bouquet GABA neurons have radially oriented axonal arbors, contain somatostatin and the calcium-binding protein calbindin, and form symmetric synapses onto the distal

dendritic shafts and spines of pyramidal neurons. The neuropeptide cholecystokinin (CCK) is present in a separate population of basket cells which also express the brain cannabinoid receptor 1 (CB1). These CCK/CB1 basket cells predominantly synapse onto the dendritic shafts and cell bodies of pyramidal neurons. In contrast, the calcium-binding protein, calretinin-containing double-bouquet cells form symmetric synapses predominantly onto the dendritic shafts of other GABA neurons. Calretinin-containing Cajal–Retzius cells reside solely in layer I and target the tuft dendrites of pyramidal neurons.

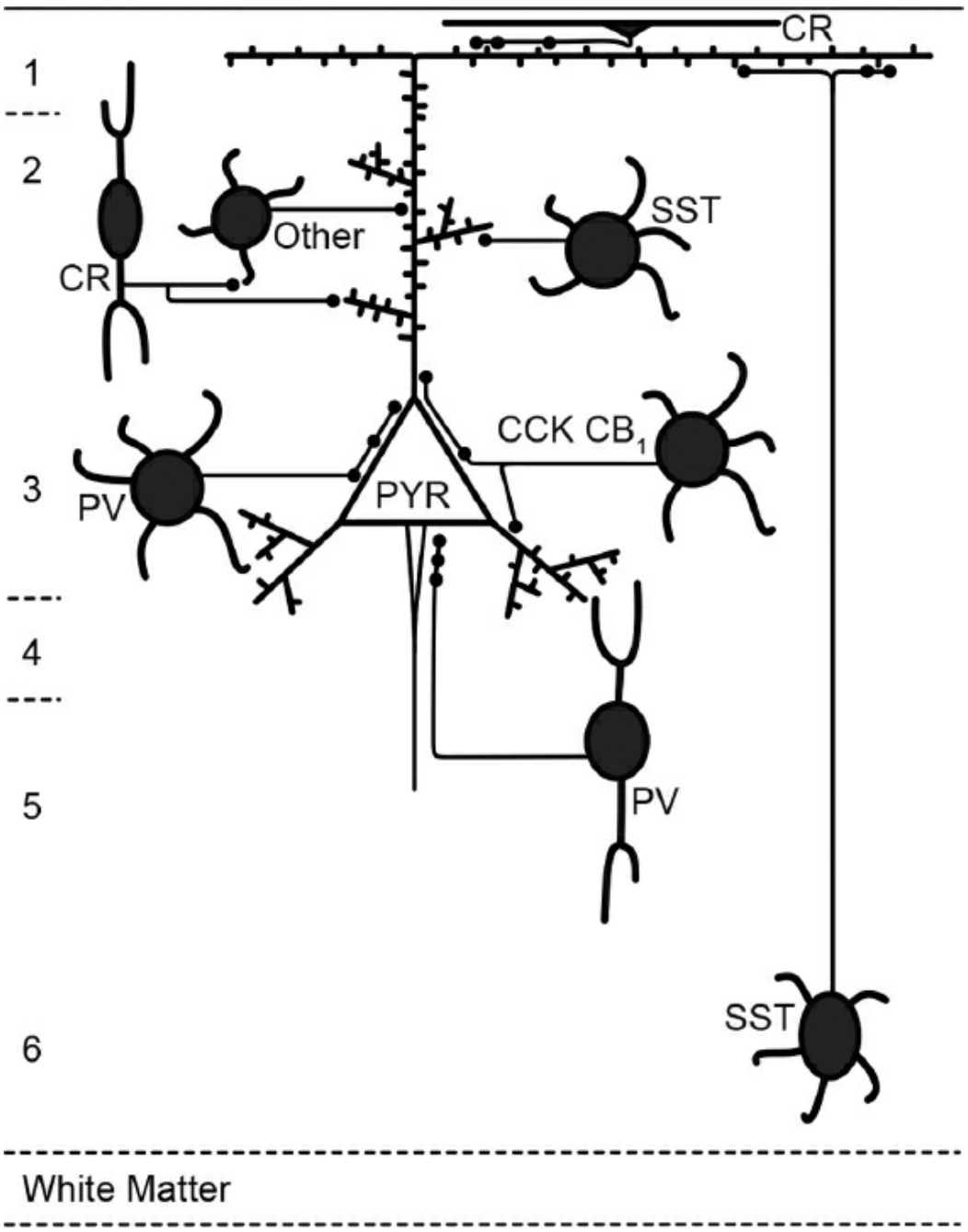

FIGURE 1.2–18. Schematic drawing of different morphologic subclasses of GABA-containing local circuit neurons in the primate prefrontal cortex. The axons of neurons in these subclasses selectively target different portions of pyramidal neurons. (Adapted from

Gonzalez-Burgos G, Hashimoto T, Lewis DA. Inhibition and timing in cortical neural circuits. *Am J Psychiatry*. 2007;164:12.)

Many studies have revealed abnormalities in populations of nonpyramidal neurons in a variety of neurological and psychiatric disorders. For example, reduced numbers of calretinin neurons and morphological changes in the axonal arbors of parvalbumin neurons have been reported in epileptic tissue. In schizophrenia, basket cells express lower levels of parvalbumin and the 67-kDa form of the biosynthetic enzyme glutamic acid decarboxylase (GAD67). The axon terminals of chandelier neurons have reduced levels of the membrane GABA transporter 1 (GAT1) and greater levels of $GABA_A$ α_2-receptors in postsynaptic pyramidal neurons (Fig. 1.2–19C). In concert, these findings reflect dysfunctional prefrontal cortical circuitry which may contribute to the cognitive symptoms observed in schizophrenia.

Neocortical neurons are distributed across six layers of the neocortex; these layers are distinguished by the relative size and packing density of their neurons (Fig. 1.2–20). Each cortical layer tends to receive particular types of inputs and furnish characteristic projections. For example, afferents from thalamic relay nuclei terminate primarily in deep layer III and layer IV, whereas corticothalamic projections originate mainly from layer VI pyramidal neurons (Fig. 1.2–21). These laminar distinctions provide important clues for dissecting possible pathophysiologic mechanisms in psychiatric disorders. Reports of decreased somal size and diminished spine density on deep layer III pyramidal neurons in the prefrontal cortex of schizophrenic patients suggest that these changes may be related to abnormalities in afferent projections from the medial dorsal thalamic nucleus. Consistent with this interpretation, the number of neurons in the medial dorsal nucleus has been reported to be decreased in schizophrenic patients.

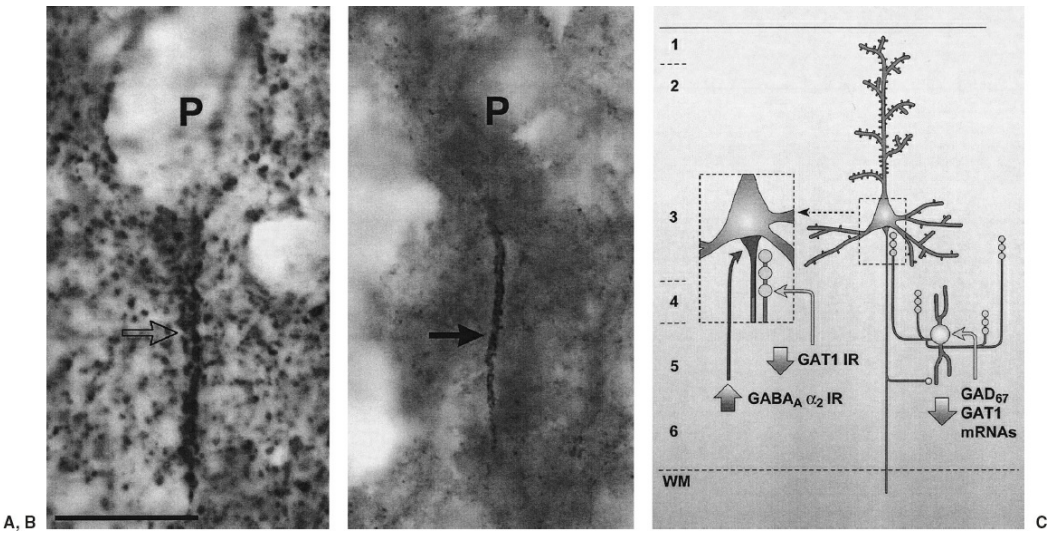

FIGURE 1.2–19. **A:** Brightfield photomicrograph of chandelier neuron axon terminal (*arrow*) immunostained for the GABA transporter-type 1 (GAT-1). **B:** Brightfield photomicrograph of a pyramidal neuron (*P*) axon initial segment (the site of action potential generation) immunostained for the α_2 subunit of the $GABA_A$ receptor. **C:** Schematic diagram of the synaptic relationship between chandelier and pyramidal neurons illustrating the pre- and postsynaptic changes in schizophrenia. (Adapted from Volk DW, Pierri JN, Fritschy J-N, Auh S, Sampson AR, Lewis DA. Reciprocal alterations in pre- and postsynaptic inhibitory markers at chandelier cell inputs to pyramidal neurons in schizophrenia. *Cereb Cortex*. 2002;12:1063.)

In addition to the horizontal laminar structure, many aspects of cortical organization have a vertical or columnar characteristic. For example, the apical dendrites of pyramidal neurons and the axons of some local circuit neurons have a prominent vertical orientation, indicating that these neural elements may sample the input to, or regulate the function of, neurons in multiple layers, respectively. Afferent inputs to the neocortex from other cortical regions also tend to be distributed across cortical layers in a columnar fashion. Finally, physiological studies in the somatosensory and visual cortices have shown that neurons in a given column respond to stimuli with particular characteristics, whereas neurons in adjacent columns respond to stimuli with different features.

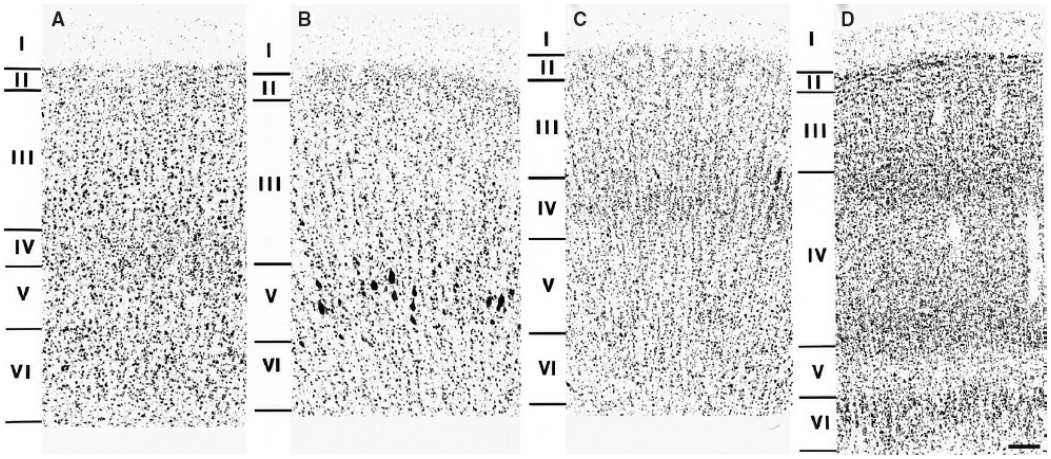

FIGURE 1.2–20. Nissl-stained sections of Brodmann area 46 (dorsolateral prefrontal cortex) (**A**), area 4 (primary motor cortex) (**B**), area 41 (primary auditory cortex) (**C**), and area 17 (primary visual cortex) (**D**) from a control human brain. Note the marked differences in the size and laminar organization of neurons across areas. The large neurons in *panel B* are Betz cells, which extend their axons to the spinal cord. Roman numerals indicate the cortical layers. Calibration bar (200 µm) applies to **A–D.**

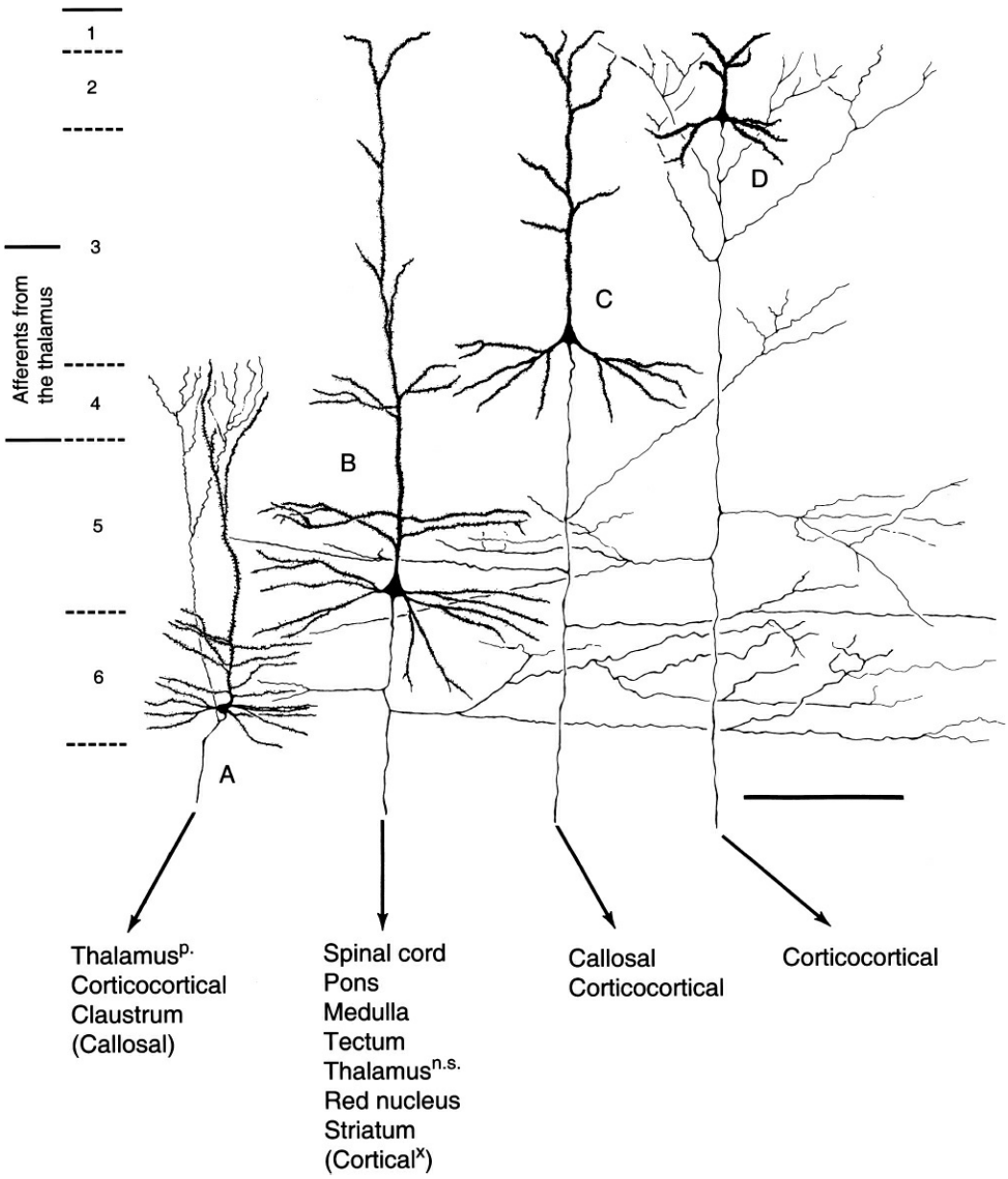

FIGURE 1.2–21. Schematic diagram of the laminar origins of efferent projections from the cerebral cortex. These data are mainly derived from the study of monkeys via tract-tracing studies. Parentheses indicate projections that may not arise from the identified layer in all species or in all cortical areas. Note that afferents from the thalamus project mainly to the lower half of layers III and IV. (Adapted from Jones EG. Laminar distribution of cortical efferent cells. In: Peters A, Jones EG, eds. *Cerebral Cortex: Cellular Components of the Cerebral Cortex*. Vol 1. New York: Plenum Press; 1984:535.)

Although best studied in sensory cortices, this pattern of organization is also present in association cortices. More recent studies in monkeys using tract-tracing techniques have shown that clusters of prefrontal cortical neurons are organized into reciprocally connected, discrete modular stripes that appear to be the analog of columns identified in the visual cortex (Fig. 1.2–4). It has been hypothesized that this organization may subserve prefrontal working memory and executive functions.

The neocortex can be divided into two general types of regions. Regions

with a readily identifiable six-layer appearance are known as the *homotypical cortex,* and are found in association regions of the frontal, temporal, and parietal lobes. In contrast, some regions of the neocortex do not have a six-layer appearance. These regions, referred to as the *heterotypical cortex,* include the primary motor cortex, which lacks a defined layer IV, and primary sensory regions, which exhibit an expanded layer IV. The neocortex can be divided further into discrete areas, each area having a distinctive architecture, certain set of connections, and role in particular brain functions. Most subdivisions of the human neocortex have been based on cytoarchitectural features; that is, subdivisions differ in the size, packing density, and arrangement of neurons across layers (Fig. 1.2–20). The most widely used system is that of Korbinian Brodmann (Fig. 1.2–22), who divided the cortex of each hemisphere into 44 numbered areas. Some of these numbered regions correspond closely to functionally distinct areas, such as area 4 (primary motor cortex in the precentral gyrus) and area 17 (primary visual cortex in the occipital lobe). In contrast, other Brodmann areas appear to encompass several cortical zones that differ in their functional attributes. Although Brodmann brain map has been used extensively in postmortem studies of psychiatric disorders, many of the distinctions among regions are subtle, and the locations of the boundaries between regions vary across individuals.

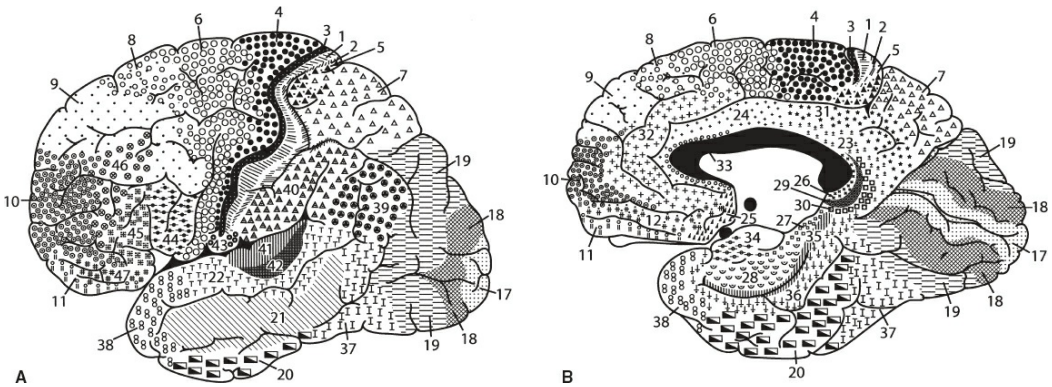

FIGURE 1.2–22. Drawing of the lateral view (**A**) and medial view (**B**) of the cytoarchitectonic subdivisions of the human brain as determined by Brodmann.

Although a given cortical area may receive other inputs, it is heavily innervated by projections from particular thalamic nuclei and from certain other cortical regions either in the same hemisphere (*associational fibers*) or the opposite hemisphere (*commissural fibers*). The patterns of connectivity make it possible to classify cortical regions into different types. *Primary sensory areas* are dominated by inputs from specific thalamic relay nuclei and are characterized by a topographic representation of visual space, the body surface, or the range of audible frequencies on the cortical surface of the primary visual, primary somatosensory, and primary auditory cortices. These regions project to nearby *unimodal association*

regions, which are also devoted to processing information from a particular sensory modality. Output from these regions converges in *multimodal association areas,* such as the prefrontal cortex or the temporoparietal cortical regions. Neurons in these regions respond to complex stimuli and are thought to be mediators of higher cognitive functions. Finally, these regions influence the activity of the *motor areas* of the cerebral cortex that control behavioral responses.

Although this classification scheme of cortical regions is accurate in many respects, it fails to account for some of the known complexities of cortical information processing. For example, somatosensory input from the thalamus projects to several distinct topographically organized maps in the cerebral cortex. In addition, information flow within the cortex is not confined to the serial processing route implied in the classification scheme, but also involves parallel processing streams, such as sensory input from the thalamus to the primary and the association areas.

Although this discussion has not distinguished between the cerebral hemispheres, certain brain functions, such as language, are localized to one hemisphere. The structural bases for the lateralization of function have not been determined, but some anatomical differences between the cerebral hemispheres have been observed. For example, a portion of the superior temporal cortex, called the *planum temporale,* is generally larger in the left hemisphere than in the right hemisphere. That cortical area, which is located close to the primary auditory cortex and includes the region known as *Wernicke area* (Fig. 1.2–5), seems to be involved in receptive language functions that are localized to the left hemisphere. In addition, Brodmann area 44 in the left inferior frontal cortex (Broca area) (Fig. 1.2–5) contains larger pyramidal neurons than the homotopic region of the right hemisphere, a difference that may contribute to the specialization of Broca area for motor speech function. A lesion in Broca area causes broken speech, whereas a lesion in Wernicke area causes wordy speech that does not make sense.

Functional Circuitry

CORTICOCORTICAL CONNECTIONS. Corticocortical connections comprise the majority of connections within the cerebral cortex and can be divided into three basic categories: (1) intrinsic—connections within one cortical area, (2) associational—connections between cortical areas within the same hemisphere, and (3) callosal—connections between cortical areas in different hemispheres. Intrinsic connections are formed by the axon collaterals of pyramidal neurons, whereas associational and callosal connections are formed by the principal axons of pyramidal neurons.

Corticocortical connections are responsible for the coordinated activity of large groups of neurons distributed within and across cortical areas. The

synchronization of this activity appears to be involved in the majority of perceptual and cognitive functions ascribed to the cerebral cortex.

The synchronous neural activity or oscillations of these connected groups of neurons generates characteristic patterns of electrical activity that can be recorded with an electroencephalogram (EEG). The greater the number of neurons oscillating together at a particular frequency, the more the amplitude of the EEG signal increases above background noise. These oscillations are grouped into different frequency bands such as θ (4 to 8 Hz), α (8 to 12 Hz), and γ (30 to 80 Hz). Oscillatory activity in all frequency bands is thought to arise from networks of inhibitory, GABAergic neurons whose output modulates the activity of pyramidal neurons.

High-frequency oscillations have been suggested as a possible mechanism for the integration of distributed neuronal networks involved in a variety of mental processes such as perception, attention, and cognition. γ-Band oscillations in the DLPFC are thought to be a key neural substrate for cognition. Moreover, the synaptic connections between the parvalbumin basket subclass of GABA neurons and pyramidal neurons (Fig. 1.2–18) in the dorsolateral prefrontal appear to generate this oscillatory activity. Interestingly, individuals with schizophrenia exhibit lower power of γ-oscillations recorded from the DLPFC when performing cognitive tasks, perhaps a consequence of the abnormalities in the GABAergic system, including PV neurons, observed in postmortem DLPFC from individuals with schizophrenia.

NETWORKS OF CORTICOCORTICAL CIRCUITRY. The study of corticocortical connections had once only been possible in animals with the use of tract-tracing methods. Although improvements in these methods have refined and deepened our knowledge of the projections between different cortical areas, the connections revealed by this methodology can only be inferred to be the same in humans. Over the last 20 years neuroimaging studies have vastly expanded our knowledge of corticocortical connectivity in humans. However, it is important to note that most of these neuroimaging studies have identified functional connectivity (i.e., brain areas that are coactivated during a specific task) and not physical connections between regions. Evidence of direct anatomical corticocortical connections in the human brain is sparse, although newer imaging techniques such as DTI do provide some structural detail. Large-scale cortical networks linking distributed areas that subserve particular functions, such as attention or memory, have long been suggested and are now more readily studied with the aforementioned neuroimaging techniques. Numerous networks have been described, but three particular networks have received the greatest focus.

The central executive network includes the DLPFC (Fig. 1.2–7) and multiple areas of lateral posterior parietal cortex (1.2–5) and is involved in certain cognitive functions such as attention, working memory, and

decision making (Table 1.2–2). The salience network, comprised of the anterior cingulate cortex (ACC) (Fig. 1.2–5), the ventromedial prefrontal cortex (VMPFC) (Fig. 1.2–7), and the insula (Figs. 1.2–8 and 1.2–9), responds to the degree of salience of incoming cognitive, homeostatic, or emotional information (Table 1.2–2). Perhaps the most intriguing network that has been described is the default mode network, which is formed by the medial prefrontal cortex (Fig. 1.2–5), posterior cingulate cortex (Fig. 1.2–5), and medial areas of posterior parietal cortex (Fig. 1.2–5). The brain regions of the default mode network are active when a person is at rest and not engaged in an externally cued task, leading researchers to hypothesize that this network is involved in internal modes of cognition including recall of personal experiences, envisioning the future and theory of mind (Table 1.2–2). Recent studies suggest that these three networks interact, with the salience network acting as a switch between the task-positive central executive network and the task-negative default mode network. The frontoinsular region of the salience network appears to play a key role in switching between the default mode and central executive networks. Interestingly, the human-specific von Economo class of neurons is found exclusively in the frontoinsular cortex and it has been suggested that these neurons are responsible for the switching capability of the salience network.

Table 1.2–2.
Functional Networks in Cerebral Cortex

Network	Brain Areas Involved	Function
Central executive	Dorsolateral prefrontal cortex Lateral posterior parietal cortex	Attention Working memory Decision making
Salience	Anterior cingulate cortex Ventromedial prefrontal cortex Insula	Detection of relevant incoming stimuli
Default	Medial prefrontal cortex Posterior cingulate cortex Medial posterior parietal cortex	Internal cognition Theory of mind

Given that these networks are involved in the processing of cognitive and emotional information, it is not surprising that their dysfunction has been implicated in a variety of neuropsychiatric disorders. For example, individuals with depression have abnormally high functional connectivity within the default mode network and reduced connectivity among regions involved in the central executive network. In schizophrenia, the functional connectivity across the default mode, central executive and salience networks is less organized. In particular, salience network dysfunction, perhaps driven by gray matter reduction in the insula and/or anterior cingulate, results in an imbalance between the default mode and central

executive networks. This breakdown in network connectivity may lead to a disruption in the capability to distinguish between the external world and internal thoughts and feelings, manifested as psychotic symptoms such as delusions and hallucinations.

Thalamus

The largest portion of the diencephalon consists of the *thalamus,* a group of nuclei located medial to the basal ganglia that serves as the major synaptic relay station for the information reaching the cerebral cortex. On an anatomic basis, the thalamic nuclei can be divided into six groups: anterior, medial, lateral, reticular, intralaminar, and midline nuclei (Fig. 1.2–23). A thin Y-shaped sheet of myelinated fibers, the *internal medullary lamina,* delimits the anterior, medial, and lateral groups of nuclei. In the human thalamus, the anterior and medial groups each contain a single large nucleus, the *anterior* and *medial dorsal nuclei.* The lateral group of nuclei can be subdivided further into dorsal and ventral tiers. The dorsal tier is composed of the *lateral dorsal,* the *lateral posterior,* and the *pulvinar nuclei;* the ventral tier consists of the *ventral anterior,* the *ventral lateral,* the *ventral posterior lateral,* and the *ventral posterior medial nuclei.* The lateral group of nuclei is covered by the *external medullary lamina,* another sheet of myelinated fibers. Interposed between these fibers and the internal capsule is a thin group of neurons forming the *reticular nucleus* of the thalamus. The *intralaminar nuclei,* the largest of which is the *central median nucleus,* are located within the internal medullary lamina. The final group of thalamic nuclei, the *midline nuclei,* covers portions of the medial surface of the thalamus. The midline nuclei of each hemisphere may fuse to form the interthalamic adhesion, which is variably present.

Thalamic nuclei also can be classified into several groups based on the pattern and information content of their connections (Table 1.2–3). *Relay nuclei* project to and receive input from specific regions of the cerebral cortex. These reciprocal connections apparently allow the cerebral cortex to modulate the thalamic input it receives. *Specific relay nuclei* process input either from a single sensory modality or from a distinct part of the motor system. For example, the lateral geniculate nucleus receives visual input from the optic tract and projects to the primary visual area of the occipital cortex. As summarized in Figure 1.2–24, neurons of the thalamic relay nuclei furnish topographically organized projections to specific regions of the cerebral cortex, although some cortical regions receive input from more than one nucleus.

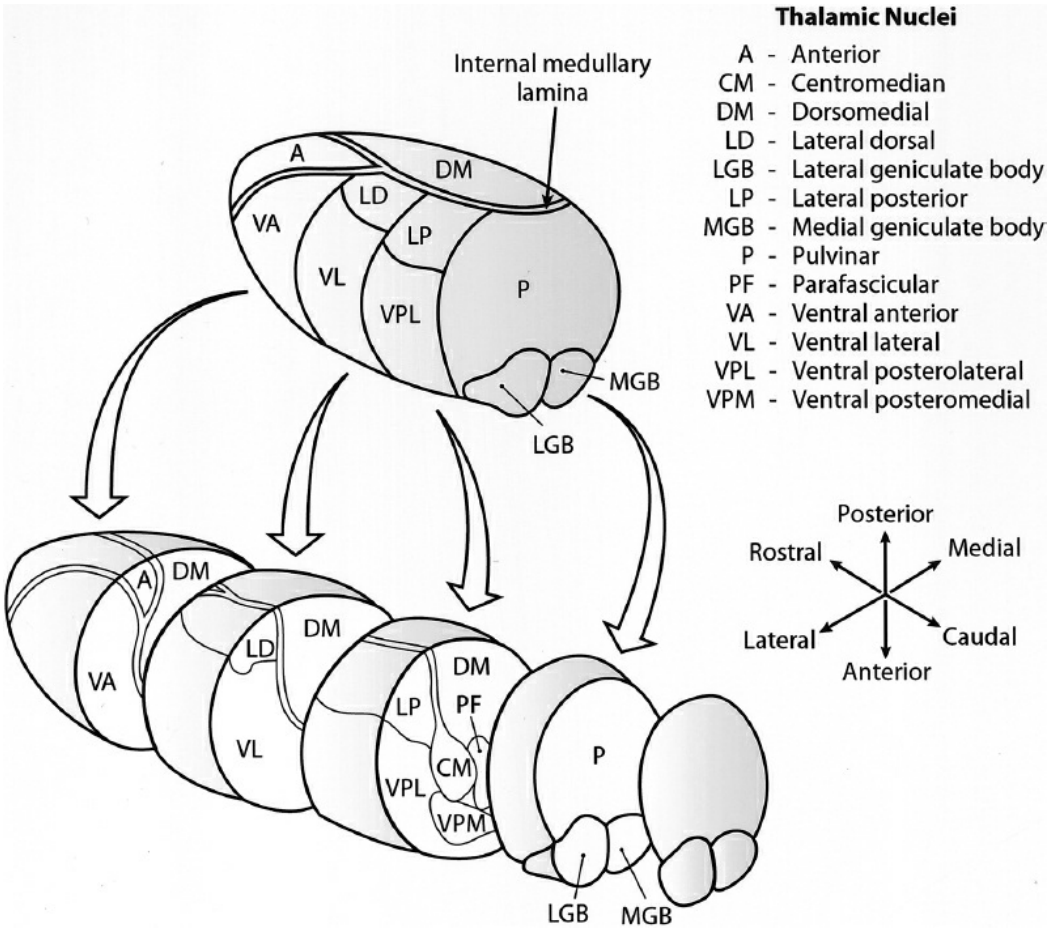

FIGURE 1.2–23. Exploded view of the dorsal thalamus illustrating the organization of thalamic nuclei. (Reprinted from Haines DE. *Fundamental Neuroscience for Basic and Clinical Applications*, 3rd ed. Philadelphia, PA: Churchill Livingstone; 2006:236.)

In contrast, *association relay nuclei* receive highly processed input from more than one source and project to larger areas of the association cortex. For example, the medial dorsal thalamic nucleus receives input from the hypothalamus and the amygdala and is reciprocally interconnected with the prefrontal cortex and certain premotor and temporal cortical regions (Fig. 1.2–25). In contrast to relay nuclei, *diffuse-projection nuclei* receive input from diverse sources and project to widespread areas of the cerebral cortex and to the thalamus. The divergent nature of the cortical connections of these nuclei indicates that they may be involved in regulating the level of cortical excitability and arousal. Finally, the reticular nucleus is unique in that it contains inhibitory neurons that receive input from collaterals of the axons that reciprocally connect other thalamic nuclei and the cerebral cortex. Each portion of the reticular nucleus then projects to the thalamic nucleus from which it receives input. The pattern of connectivity indicates that the reticular nucleus samples cortical afferent and efferent activity and then uses that information to regulate thalamic function.

Table 1.2–3.
Connections of Thalamic Nuclei[a]

Type	Nuclei	Principal Afferent Inputs	Major Projection Sites
Specific relay	Anterior	Mammillary body of hypothalamus	Cingulate cortex
	Ventral anterior	Globus pallidus	Premotor cortex
	Ventral lateral	Dentate nucleus of cerebellum	Motor, premotor cortices
	Ventral posterior lateral	Medial lemniscal and spinothalamic pathways	Somatosensory cortex
	Ventral posterior medial	Sensory nuclei of trigeminal nerve	Somatosensory cortex
	Medial geniculate	Inferior colliculus	Auditory cortex
	Lateral geniculate	Optic tract	Visual cortex
Association relay	Lateral dorsal	Unknown	Cingulate cortex
	Lateral posterior	Superior colliculus	Parietal cortex
	Pulvinar	Superior colliculus	Temporal, parietal, occipital cortices
	Medial dorsal	Amygdala and hypothalamus	Prefrontal cortex
Diffuse-projection	Midline	Reticular formation, hypothalamus	Basal forebrain, cortex
	Intralaminar	Reticular formation, spinothalamic tract, globus pallidus	Basal ganglia, cortex
	Reticular	Cerebral cortex, thalamus	Thalamus

[a]This table does not include the cortical inputs to each thalamic nucleus.

Modified from Kelly JP. The neutral basis of perception and movement. In: Kandel ER, Schwartz JH, Jessell TM, eds. *Principles of Neural Science*, 3rd ed. New York: Elsevier; 1991:291.

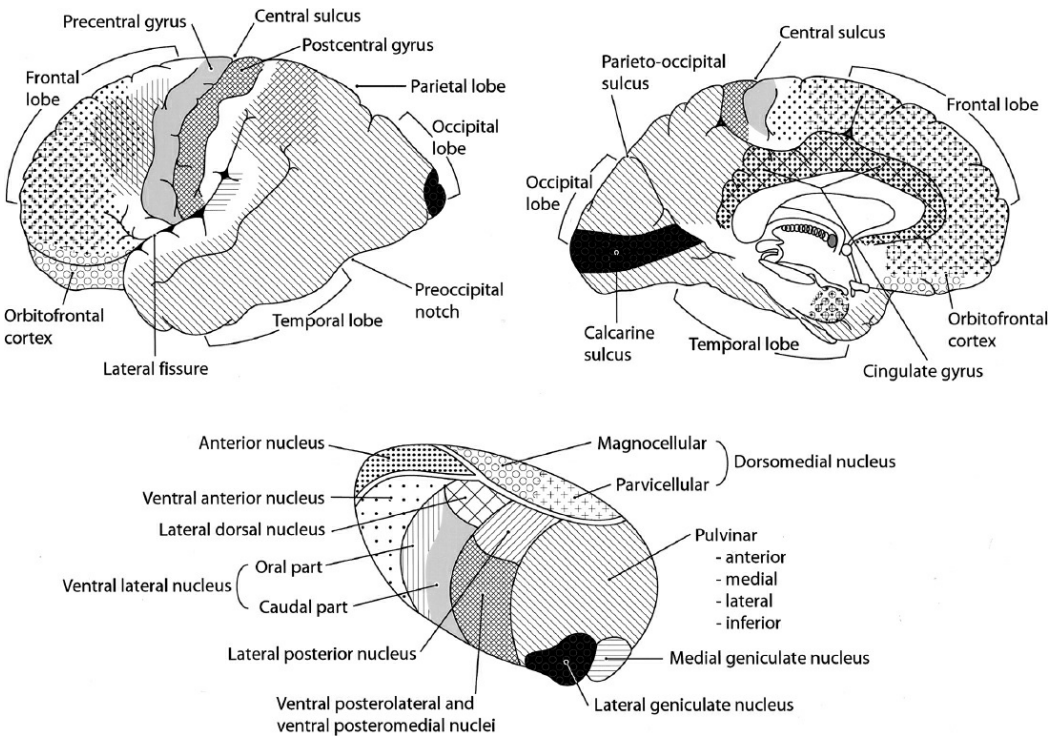

FIGURE 1.2–24. Schematic drawings of the lateral *(upper left)* and medial *(upper right)* surfaces of the right cerebral hemisphere and the right thalamus *(lower).* Each thalamic nucleus is patterned coded to match its target area in the cerebral cortex. (Adapted from Haines DE. *Fundamental Neuroscience for Basic and Clinical Applications,* 3rd ed. Philadelphia, PA: Churchill Livingstone; 2006:237.)

The classification of thalamic nuclei into different categories has been refined over the last 20 years to reflect our greater understanding of thalamic circuitry. Although classification into different nuclei is useful, focus has shifted to the type of inputs to the various nuclei. For example, thalamic afferents can be defined as either drivers or modulators. Drivers convey information from the senses and periphery that are then relayed to cerebral cortex. These inputs determine the nature of the thalamic relay. For example, the lateral geniculate nucleus is a visual relay because it transmits retinal input to primary visual cortex. The driver inputs for association and diffuse-projection nuclei are more difficult to determine, but these inputs most likely arise from layer V pyramidal neurons in the cerebral cortex. Modulators include all other inputs to the thalamus, such as monoaminergic and cholinergic projections from brainstem nuclei. Modulator inputs also include the feedback projections from layer VI pyramidal neurons in the cerebral cortex. Thalamic circuits can then be distinguished based on the source of driver inputs. First-order relays have drivers that arise from ascending pathways and transmit information that has not yet been processed in the cortex. In contrast, higher-order relays have drivers that arise from layer V pyramidal neurons, allowing for transmission of information from one cortical area to another via the thalamus. Thus, it is possible that a given thalamic nucleus contains both

first-order and higher-order relays. Interestingly, over half of the circuits in the thalamic nuclei of primates are considered to be of higher order, suggesting a more complex role of the thalamus in cortico-thalamo-cortical function.

Functional Circuitry. The connections between the thalamus, the cortex, and certain related brain structures constitute three types of thalamocortical systems, each with different patterns of functional circuitry. These three systems—sensory, motor, and association systems—are described separately here, but are heavily interconnected.

THALAMOCORTICAL SENSORY SYSTEMS. Several general principles govern the organization of the thalamocortical sensory systems. First, sensory receptors transduce certain stimuli in the external environment to neural impulses. The impulses ascend, often through intermediate nuclei in the spinal cord and the medulla, and ultimately synapse in specific relay nuclei of the thalamus.

Second, projections from peripheral sensory receptors to the thalamus and the cortex exhibit topography, that is, a particular portion of the external world is mapped onto a particular region of the brain. For example, in the somatosensory system, axons carrying information regarding a distinct part of the body synapse in a discrete part of the ventral posterior nucleus of the thalamus. Specifically, the ventral posterior medial nucleus receives inputs regarding the head, and the ventral posterior lateral nucleus receives inputs regarding the remainder of the body. The nuclei project topographically to the primary somatosensory cortex, where several representations of the contralateral half of the body can be found. These representations are distorted; regions heavily innervated by sensory receptors, such as the fingers, are disproportionately represented in the primary somatosensory cortex.

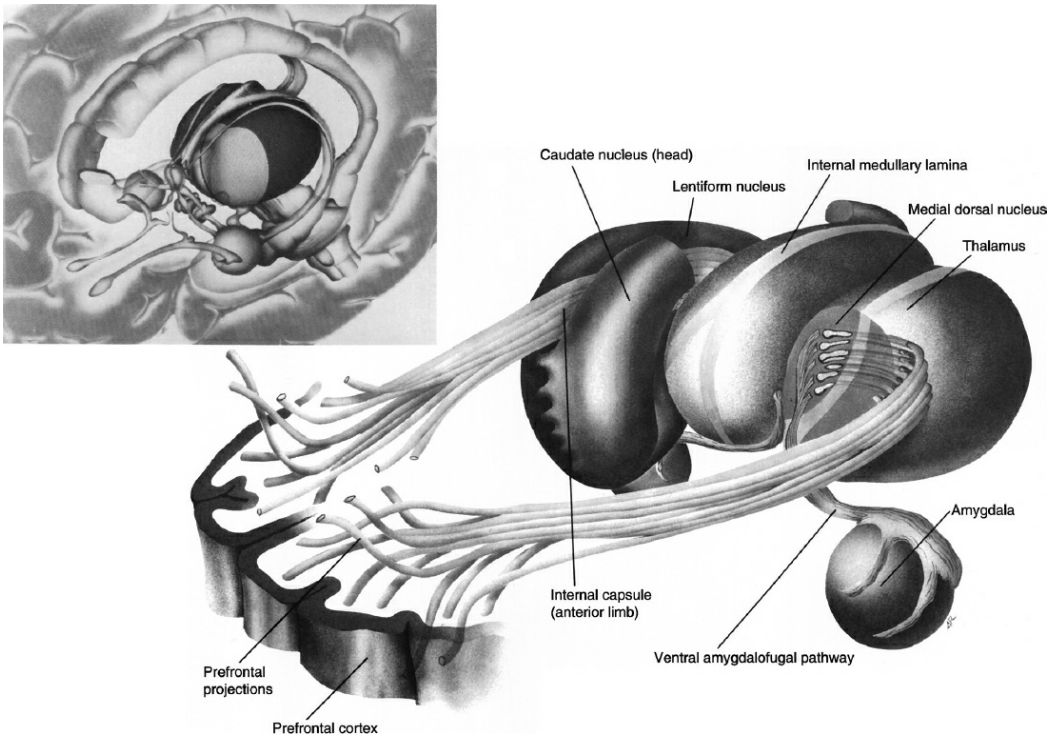

FIGURE 1.2–25. Drawing of the thalamus showing the pathway of projections from the mediodorsal nucleus through lateral thalamic nuclei to the prefrontal cortex. Also shown are afferents from the amygdala to the medial dorsal nucleus. The *inset* shows the thalamus embedded in the limbic system of which it is a key component. (Adapted from Hendelman WJ. *Student's Atlas of Neuroanatomy*. Philadelphia, PA: WB Saunders; 1994:199.)

Third, in some cases, sensory inputs travel to the thalamus in a segregated manner according to the submodality of the information conveyed. The inputs are processed in a parallel fashion; particular pathways may be devoted exclusively to processing a submodality. An example of such segregation is evident in the somatosensory system (Fig. 1.2–26), where most fibers carrying tactile and proprioceptive information travel in the medial lemniscus, whereas fibers carrying pain and temperature information travel in the spinothalamic tract to the ventral posterior thalamic nuclei. Although some tactile information is carried in the spinothalamic tract, the submodalities of pain and temperature are largely segregated from tactile and proprioceptive inputs as they ascend to the thalamus.

Finally, sensory pathways exhibit convergence, that is, primary sensory areas process sensory information and project to unimodal association areas. Subsequently, the unimodal areas project to and converge in multimodal associational areas. Convergence in sensory pathways is illustrated in the somatosensory system. The primary somatosensory cortex, located in the anterior parietal lobe, has been divided into four regions on the basis of cytoarchitecture. Each of the cytoarchitectonic regions—numbered 1, 2, 3a, and 3b by Brodmann—contains a

topographical representation of the body. The regions are heavily interconnected, and all project to the next level of somatosensory processing in area S-II (Brodmann areas 40 and 43). This type of projection, from one level of processing to a more advanced level, is termed a *feedforward* projection. The reciprocal connection, from the more advanced processing level back to the simpler level, is called a *feedback* projection. Both projections have distinct patterns of laminar termination: feedforward projections originate in the superficial layers of cortex (layer III) and terminate in layer IV; feedback projections originate in layers III, V, and VI, and terminate outside layer IV. Further processing of somatosensory information occurs in higher-order somatosensory areas, such as area 7b of the posterior parietal cortex, which receives feedforward projections from S-II. Lesions of the posterior parietal cortex reflect the complexity of the information processed there; after a person has sustained a posterior parietal lesion, the ability to understand the significance of sensory stimuli is impaired, and extreme cases result in contralateral sensory neglect and inattention.

THALAMOCORTICAL MOTOR SYSTEMS. The thalamocortical motor systems exhibit some unique organizational principles, but also share many of the features present in the sensory systems. First, in contrast to sensory systems, which primarily ascend from sensory receptor to cortical association areas, motor systems descend from association and motor regions of the cortex to the brainstem and the spinal cord. The corticospinal tract originates in the large Betz cells of layer V in the premotor and primary motor cortices (Fig. 1.2–20B) of the frontal lobe and terminates in the spinal cord to influence motor behavior.

Second, motor systems exhibit strong topography at the thalamic and cortical levels. The corticospinal tract is organized so that a topographical representation of the contralateral half of the body is evident in the primary motor and premotor cortices. The representation of the body is disproportionate, with large regions of the motor cortex devoted to areas of the body involved in fine movement, such as the face and the hands.

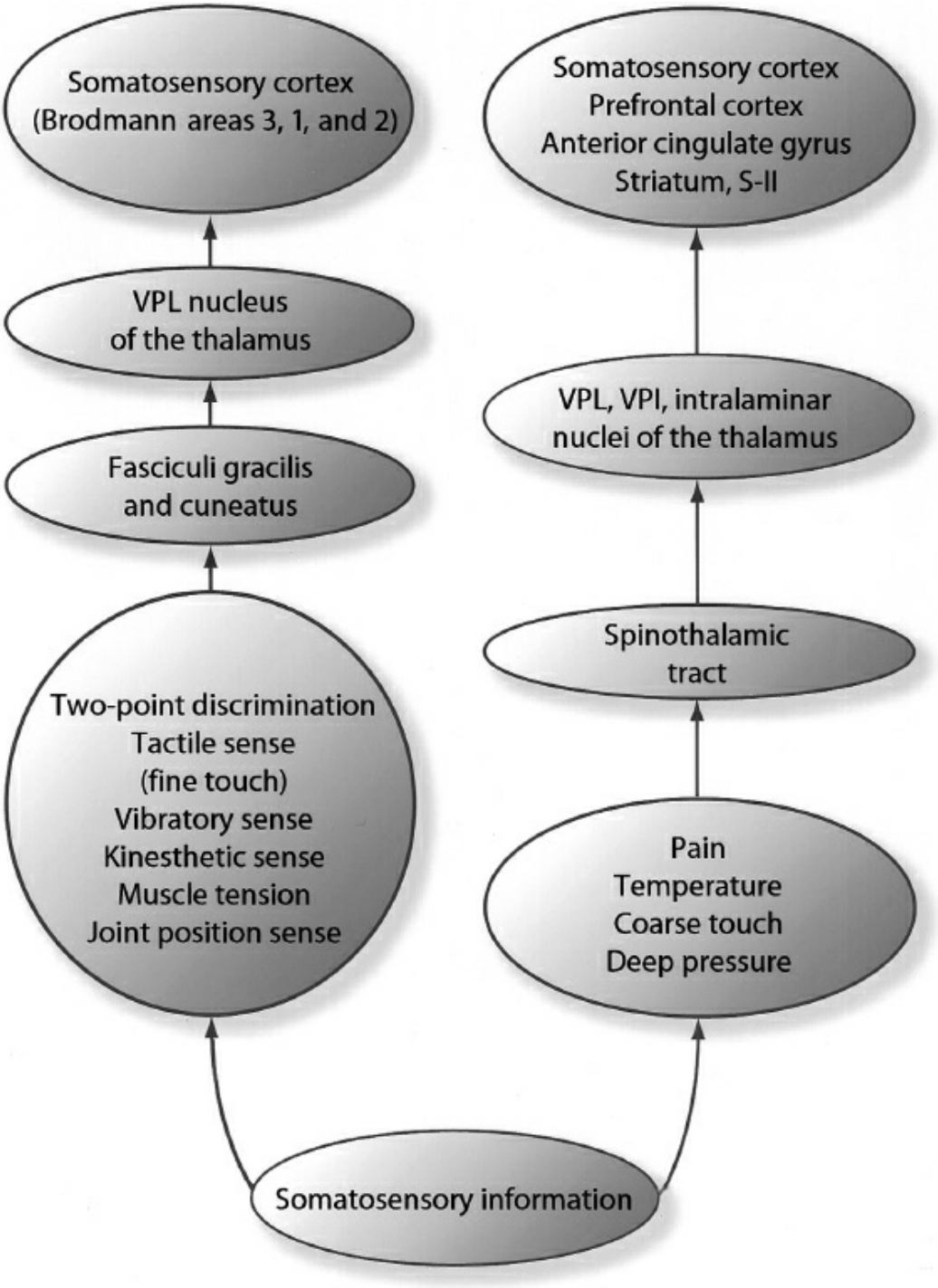

FIGURE 1.2–26. Pathway of somatosensory information processing. (Adapted from Patestas MA, Gartner LP. *A Textbook of Neuroanatomy*. Malden, MA: Blackwell; 2006:149.)

Finally, there is a convergence of projections from several sensory association regions to the motor regions of the frontal cortex. The premotor cortex receives a convergence of afferents from higher-order somatosensory and visual areas of the posterior parietal cortex, whereas afferents from the primary somatosensory cortex converge on the primary motor cortex. In addition to cortical input, the primary motor cortex receives afferents from the ventral lateral nucleus of the thalamus; this

nucleus receives afferents predominantly from the cerebellum. The premotor cortex receives input from the ventral anterior thalamic nucleus, which receives much of its input from the globus pallidus.

THALAMOCORTICAL ASSOCIATION SYSTEMS. The multimodal association areas of the cortex are organized according to several general principles. First, association regions receive a convergence of input from a variety of sources, including unimodal and multimodal association regions of the cortex, association nuclei of the thalamus, and other structures. The prefrontal cortex receives afferents from higher-order sensory cortices of the parietal and temporal lobes, the contralateral prefrontal cortex, the cingulate cortex of the limbic system, the medial dorsal nucleus of the thalamus (an association relay nucleus), and portions of the amygdala. The medial dorsal nucleus receives highly processed inputs from many sources, including some regions, such as the amygdala, hippocampus, and hypothalamus, that project directly to the prefrontal cortex. The redundant (direct and indirect) projections may serve to attach additional significance to certain inputs received by the prefrontal cortex. The significance of these inputs may also be influenced by their temporal and spatial coincidence with modulatory inputs from brainstem nuclei that use the monoamine neurotransmitters dopamine, norepinephrine, or serotonin. These monoamine systems project broadly to the cerebral cortex, although with substantial regional differences in density (Fig. 1.2–27). In addition, the innervation density in the cerebral cortex is typically much lower than in some subcortical areas.

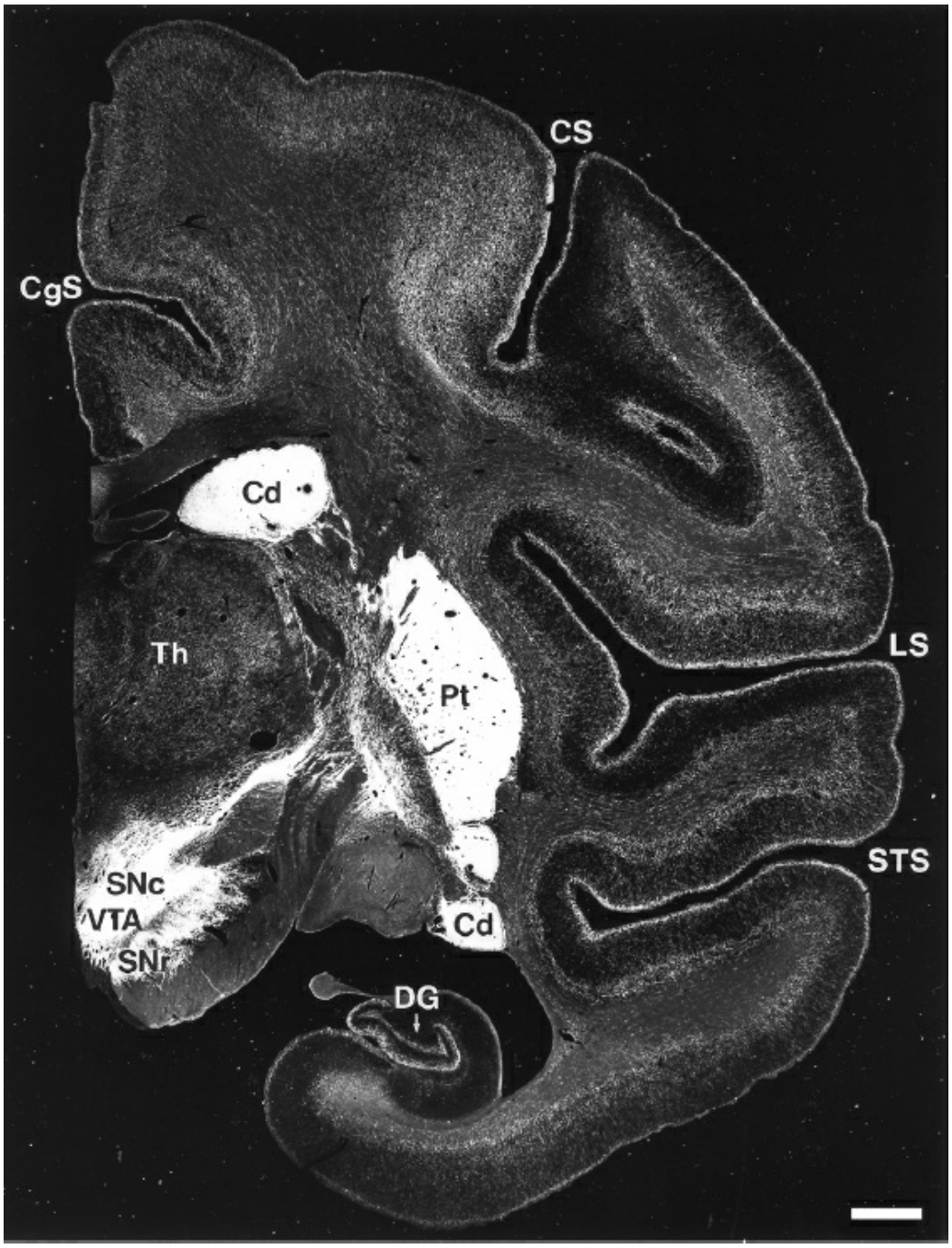

FIGURE 1.2–27. Darkfield photomicrograph of a coronal section through a hemisphere of a macaque monkey immunolabeled for the dopamine transporter. This image illustrates the differential distribution of dopamine-containing axons in different regions of the brain. The brighter the image, the greater the quantity of dopamine-containing axons. Dopamine-rich areas such as the caudate *(Cd),* putamen *(Pt),* ventral tegmental area (*VTA*), and the substantia nigra (*SNc* and *SNr*) appear white, whereas dopamine innervation of the cortex and thalamus, although clearly seen, is less dense and varies by the specific cortical and thalamic region. CgS, cingulate sulcus; CS, central sulcus; DG, dentate gyrus; LS, lateral sulcus; STS, superior temporal sulcus; Th, thalamus. Calibration bar = 2 mm. (From Lewis DA, Melchitzky DS, Sesack SR, Whitehead RE, Auh S, Sampson A. Dopamine transporter immunoreactivity in monkey cerebral cortex: Regional, laminar and ultrastructural localization. *J Comp Neurol*. 2001;432(1):119. Used with permission.)

The second way that the projections are organized is according to

topography. The projections that terminate in multimodal association regions exhibit a topographic organization. Different cytoarchitectonic regions of the medial dorsal nucleus project to discrete regions of the prefrontal cortex. In addition, some cortical afferents received by the prefrontal cortex are topographically organized; certain regions of the prefrontal cortex predominantly receive highly processed information from one modality.

The patterns of connectivity are clearly related to some of the functional characteristics attributed to the prefrontal cortex. For example, in monkeys, lesions of the DLPFC consistently produce impairments in a monkey's ability to perform spatial delayed-response tasks. These tasks require that monkeys maintain a spatial representation of the location of an object during a delay period in which the object is out of sight; it has been suggested that the prefrontal cortex plays a role in maintaining the spatial representation of the object. Such a function would require that the prefrontal cortex receive information regarding the location of objects in space, and the DLPFC is innervated by afferents from association regions of the parietal cortex that convey such information. Although the DLPFC is necessary for the performance of delayed-response tasks in monkeys, it is insufficient for the performance of the task. For example, lesions of the medial dorsal nucleus in monkeys result in similar impairments on the performance of spatial delayed-response tasks. The functions attributed to the prefrontal cortex are a result of the neural circuitry involving the region.

Knowledge of the integration of afferent inputs into the neural circuitry of certain prefrontal regions may also be important for understanding the nature of prefrontal cortical dysfunction in schizophrenia. Individuals with schizophrenia perform poorly on tasks that are known to be mediated by the prefrontal cortex. These findings have been correlated with other measures to suggest, albeit indirectly, that the dopamine projections to the prefrontal cortex are impaired in schizophrenia. Studies in NHPs have shown that performance of delayed-response tasks that require working memory, the same type of behaviors that are impaired in subjects with schizophrenia, requires an appropriate level of dopamine input to the DLPFC.

During visuospatial delayed-response tasks the DLPFC guides decisions about how to sample and process complex visual information and it does this through circuitry involving the frontal eye fields, the mediodorsal thalamic nucleus, and the superior colliculus. The frontal eye fields are located in the frontal cortex, situated between the more anterior prefrontal cortex and the more caudal motor cortex, and control saccadic eye movements which are an integral component of visuospatial delayed-response activity. Interestingly, the mediodorsal nucleus has recently been described as containing both first-order and higher-order relays, both of

which appear to be involved in visuospatial delayed-response tasks. First-order brainstem afferents from the superior colliculus send eye movement–related information to the mediodorsal nucleus, which then transmits this information onto the frontal eye fields. In addition to direct corticocortical connections between the DLPFC and the frontal eye fields, information from the frontal eye fields concerning the executed saccadic eye movement is transmitted to prefrontal cortex via higher-order relays in the mediodorsal nucleus. This transthalamic circuitry suggests that the mediodorsal nucleus plays a significant role in contributing to the control of behavior executed by the DLPFC.

CEREBELLOTHALAMOCORTICAL SYSTEMS. The cerebellum traditionally has been considered to be involved solely with motor control, regulating posture, gait, and voluntary movements. However, more recent studies indicate that the cerebellum may also play an important role in the mediation of certain cognitive abilities through inputs to portions of the thalamus that project to association regions of the cerebral cortex.

The cerebellum is located in the posterior cranial fossa, inferior to the occipital lobes (Figs. 1.2–5 and 1.2–11). The external surface of the cerebellum, the cerebellar cortex, is composed of small folds, termed *folia,* separated by sulci. Viewed from the dorsal surface, the cerebellum contains a raised central portion, called the *vermis,* and lateral portions called the *cerebellar hemispheres* (Fig. 1.2–11). Located within the cerebellum are the deep cerebellar nuclei, which are arranged as follows: the fastigial nucleus is located next to the midline; the globose and emboliform nuclei are slightly more lateral; and the largest nucleus, the dentate, occupies the most lateral position. Generally, the cerebellar cortex can be considered to process the inputs to the cerebellum, and the deep nuclei to process the outputs. Although many portions of the cerebellum are interconnected with brain regions that regulate motor actions, the circuitry of the cerebellum involved in cognitive functions may be of greatest interest from the standpoint of psychiatric illness. For example, the lateral cerebellar cortex and the dentate nucleus are markedly expanded in the primate brain. It has been suggested that these changes are associated with an increase in the size of cortical areas (especially the prefrontal regions) influenced by cerebellar output and an expanded role of the cerebellum in cognitive functions. More recent studies in NHPs have shown that the DLPFC receives inputs from two ipsilateral thalamic nuclei (medial dorsal and ventral lateral), which receive inputs from the contralateral cerebellar dentate nucleus. The cells of the dentate nucleus involved in these connections are distinct from the cells that influence the motor and premotor regions of the cerebral cortex. Interestingly, functional imaging studies in schizophrenic subjects have revealed abnormal patterns of activation in the cerebellum, thalamus, and prefrontal cortex, suggesting

that dysfunction of this circuitry might be associated with the disturbances in cognitive processes exhibited by these patients.

Basal Ganglia System

The basal ganglia are a collection of nuclei that have been grouped together on the basis of their interconnections. These nuclei play an important role in regulating movement and in certain disorders of movement (dyskinesias), which include jerky movements (chorea), writhing movements (athetosis), and rhythmic movements (tremors). In addition, more recent studies have shown that certain components of the basal ganglia play an important role in many cognitive functions.

Major Structures. The basal ganglia are generally considered to include the *caudate nucleus,* the *putamen,* the *globus pallidus* (referred to as the *paleostriatum* or *pallidum*), the *subthalamic nucleus,* and the *substantia nigra* (Fig. 1.2–28). The term *striatum* refers to the caudate nucleus and the putamen together; the term *corpus striatum* refers to the caudate nucleus, the putamen, and the globus pallidus; and the term *lentiform nucleus* refers to the putamen and the globus pallidus together.

Although these nuclei are generally agreed to belong to the basal ganglia, some controversy exists concerning whether other nuclei should be included in the definition of the basal ganglia. Some investigators believe that additional regions of the brain have anatomic connections that are similar to other components of the basal ganglia and should, therefore, be included in the term. These additional regions are usually termed the *ventral striatum* and the *ventral pallidum*. The ventral striatum includes the nucleus accumbens (Fig. 1.2–29), which is the region where the putamen and the head of the caudate nucleus fuse, and the olfactory tubercle. The ventral pallidum is a region that receives afferents from the ventral striatum and includes, but is not limited to, a group of neurons termed the *substantia innominata* (Fig. 1.2–8). This section focuses on the structures generally accepted as belonging to the basal ganglia, but also discusses additional structures when relevant to the functional anatomy of the system.

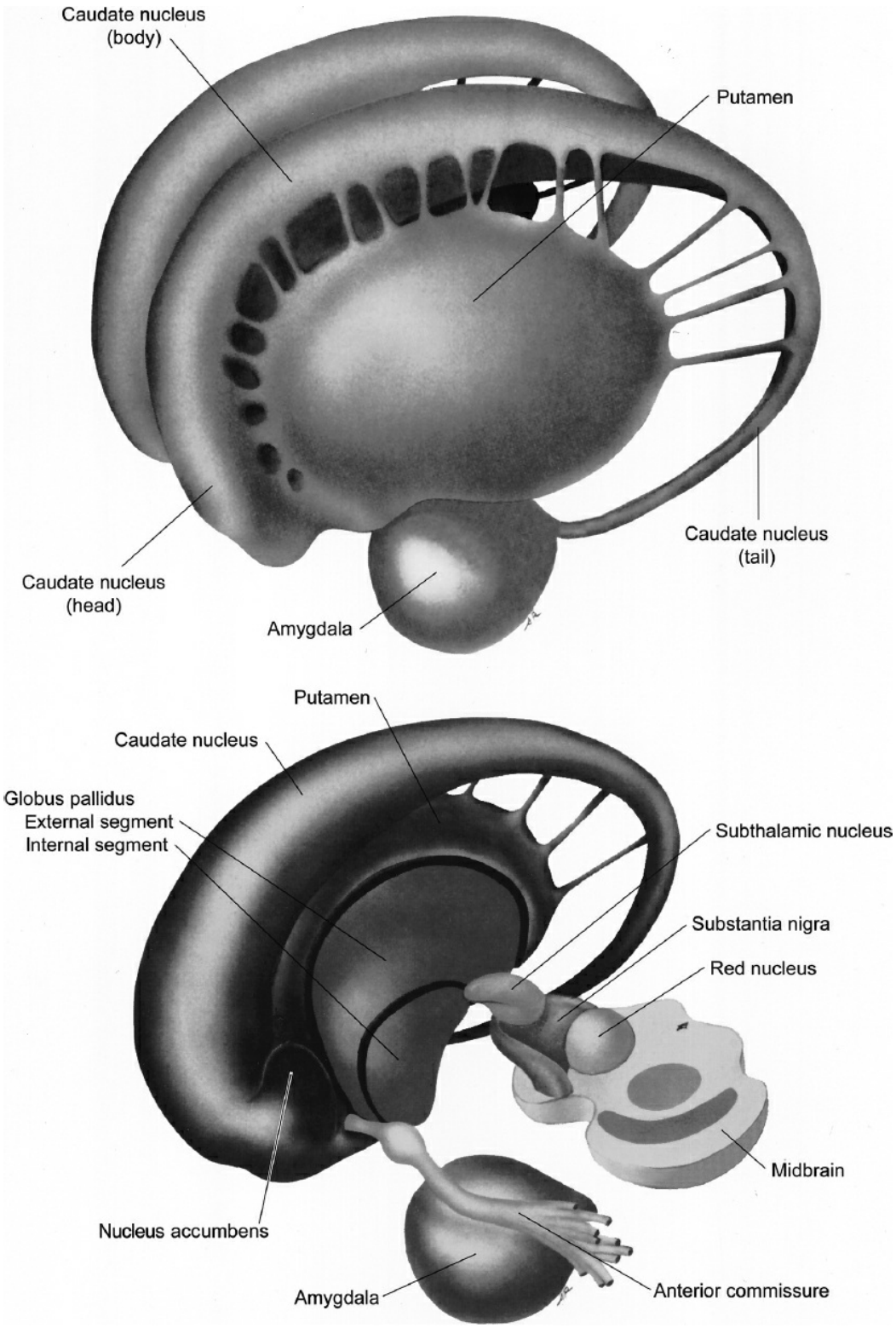

FIGURE 1.2–28. Schematic drawing of the isolated basal ganglia as seen from the dorsolateral perspective, so that the caudate nucleus is apparent bilaterally. In the bottom panel, the basal ganglia from the left hemisphere has been removed, exposing the medial surface of the right putamen and globus pallidus, and the subthalamic nucleus and substantia nigra. (Adapted from Hendelman WJ. *Student's Atlas of Neuroanatomy*. Philadelphia, PA: WB Saunders; 1994:37.)

CAUDATE NUCLEUS. The caudate nucleus is a C-shaped structure that is

divided into three general regions. The anterior portion of the structure is referred to as the *head,* the posterior region is the *tail,* and the intervening region is the *body* (Fig. 1.2–28). The caudate nucleus is associated with the contour of the lateral ventricles: the head lies against the frontal horn of the lateral ventricle, and the tail lies against the temporal horn (Figs. 1.2–8, 1.2–9, and 1.2–10). The head of the caudate nucleus is continuous with the putamen; the tail terminates in the amygdala of the temporal lobe.

PUTAMEN. The putamen lies in the brain medial to the insula and is bounded laterally by the fibers of the external capsule and medially by the globus pallidus (Figs. 1.2–8 and 1.2–9). As noted earlier, the putamen is continuous with the head of the caudate nucleus (Fig. 1.2–28). Although bridges of neurons between the caudate nucleus and the putamen show the continuity of the nuclei, the two structures are separated by fibers of the anterior limb of the internal capsule (Fig. 1.2–29).

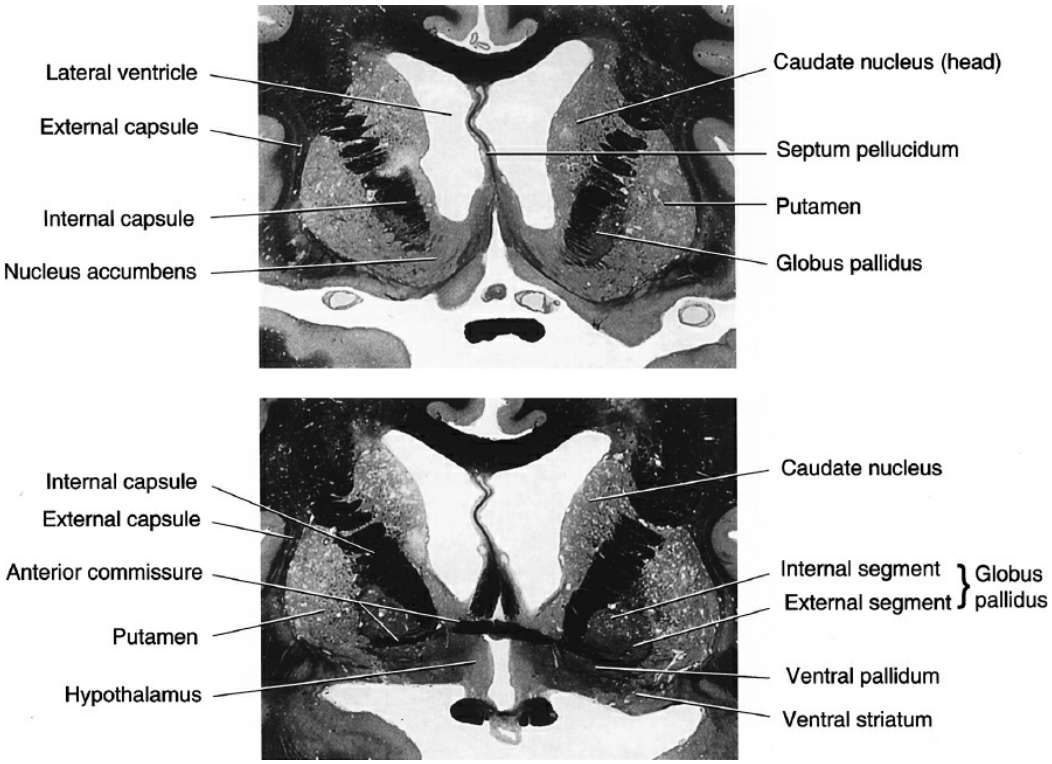

FIGURE 1.2–29. Photographs of cross-sections of human brain showing basal ganglia nuclei and related structures. (Reprinted from Haines DE. *Fundamental Neuroscience for Basic and Clinical Applications*, 3rd ed. Philadelphia, PA: Churchill Livingstone; 2006:416.)

GLOBUS PALLIDUS. In contrast to the caudate nucleus and the putamen, which are telencephalic in origin, the globus pallidus is derived from the diencephalon. The globus pallidus constitutes the inner component of the lentiform nucleus (Fig. 1.2–28, bottom panel); with the putamen, it forms a cone-like structure, with its tip directed medially (Figs. 1.2–8 and 1.2–9). The posterior limb of the internal capsule bounds the globus pallidus

medially and separates it from the thalamus; the putamen borders the globus pallidus laterally. In humans, the medial medullary lamina divides the globus pallidus into external (lateral) and internal (medial) segments (Fig. 1.2–29).

SUBTHALAMIC NUCLEUS. The subthalamic nucleus (of Luys) is also derived from the diencephalon. The large-celled nucleus lies dorsomedial to the posterior limb of the internal capsule and dorsal to the substantia nigra (Figs. 1.2–9 and 1.2–28). Discrete lesions of the subthalamic nucleus in humans lead to hemiballism, a syndrome characterized by violent, forceful choreiform movements that occur on the side contralateral to the lesion.

SUBSTANTIA NIGRA. The substantia nigra is present in the midbrain between the tegmentum and the basis pedunculi and is mesencephalic in origin (Fig. 1.2–9). The substantia nigra consists of two components: a dorsal cell–rich portion referred to as the *pars compacta* and a ventral cell–sparse portion denoted the *pars reticulata*. Most of the neurons in the pars compacta of the substantia nigra in humans are pigmented because of the presence of neuromelanin; these cells contain the neurotransmitter dopamine (Fig. 1.2–27). The dendrites of the pars compacta neurons frequently extend into the pars reticulata, where they receive synapses from the neurons of the pars reticulata that use the inhibitory neurotransmitter GABA.

In rodents, the dopamine-containing neurons of the substantia nigra (A9 region) have been distinguished from the neurons located in the ventral tegmental area VTA (A10 region) (Fig. 1.2–30) and the retrorubral field (A8 region), but more recent studies in monkeys and humans suggest that dopamine neurons can be more meaningfully organized at a functional level into dorsal and ventral tiers (Fig. 1.2–30). The dorsal tier is formed by a medially and laterally oriented band of neurons that includes the dopamine-containing cells that are (1) located in the medial ventral mesencephalon, (2) scattered dorsal to the dense cell clusters in the substantia nigra, and (3) distributed lateral and caudal to the red nucleus. The ventral tier comprises the dopamine neurons that are densely packed in the substantia nigra and the cell columns that penetrate into the substantia nigra pars reticulata. Dorsal tier dopamine neurons express relatively low levels of messenger RNA (mRNA) for the dopamine transporter (DAT) and the dopamine type 2 receptor (D_2), contain the calcium-binding protein calbindin, and send axonal projections to the regions of the striatum that are dominated by input from limbic-related structures and association regions of the cerebral cortex. In contrast, ventral-tier neurons contain high levels of mRNA for the DAT and the D_2 dopamine receptor, typically lack calbindin, and send axonal projections to the sensorimotor regions of the striatum. Each of these features may

contribute to the greater vulnerability of ventral tier neurons to the pathology of Parkinson disease, whereas dorsal tier neurons may be more likely to be involved in the pathophysiology of schizophrenia.

Internal Organization. The caudate nucleus and the putamen are frequently referred to together because of their common characteristics. In rodents, these nuclei are a continuous structure, and, in all mammals, they are composed of histologically identical cells. Most neurons in the striatum are medium-sized cells (10 to 20 μm in diameter) that possess spines on their dendrites; these so-called *medium spiny neurons* are known to send their axons out of the striatum. In addition to medium spiny neurons, medium-sized cells without spines (medium aspiny neurons) are present, as are large cells with and without spines (large spiny neurons and large aspiny neurons). With the exception of the medium and large spiny cells, most other striatal neurons are local circuit neurons.

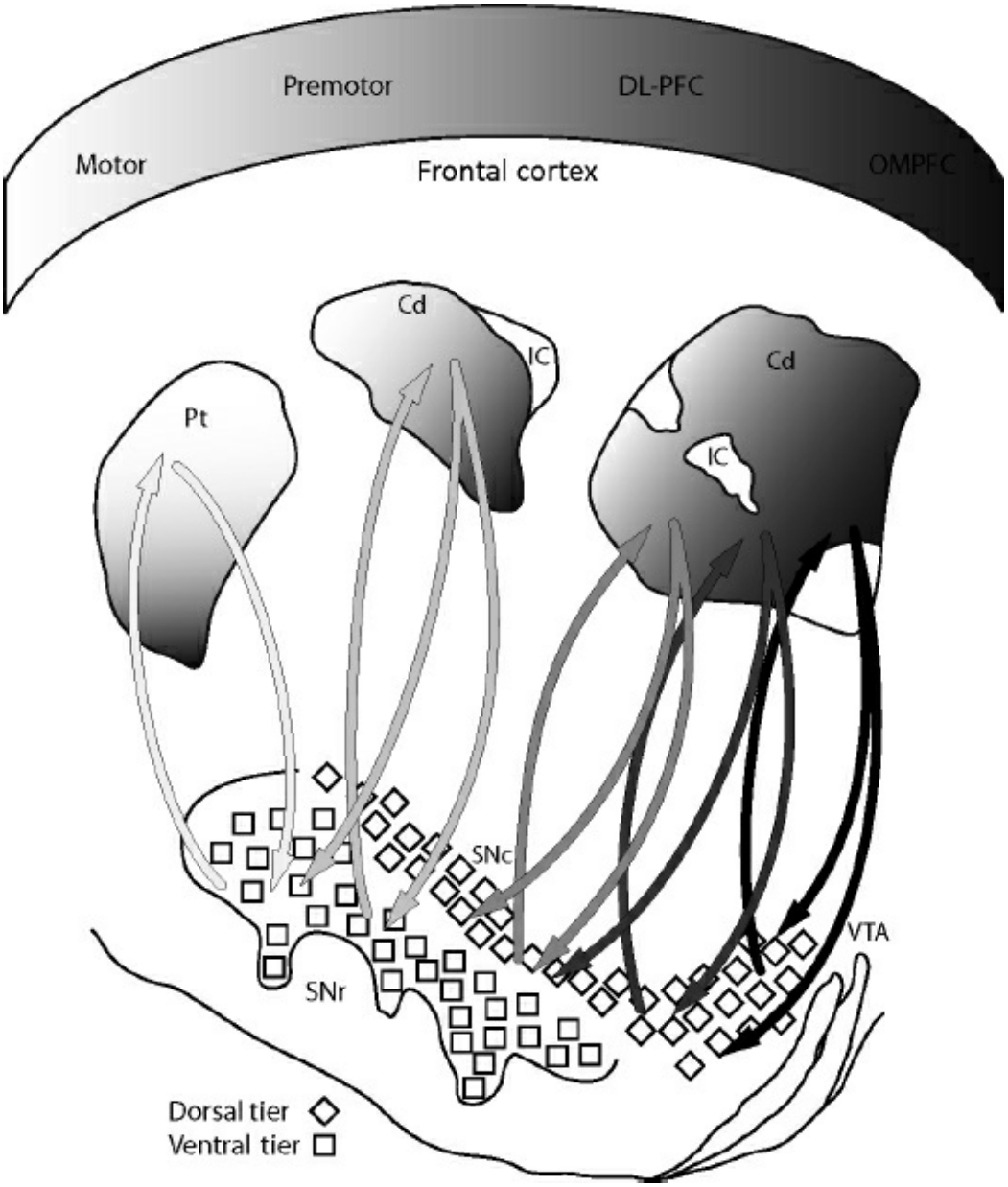

FIGURE 1.2–30. Diagram of the organization of the striatonigrostriatal and corticostriatal projections in monkeys. DL-PFC, dorsolateral prefrontal cortex; IC, internal capsule; OMPFC, orbital and medial prefrontal cortex; S, shell; SNC, substantia nigra pars compacta; SNr, substantia nigra pars reticulata; VTA, ventral tegmental area. (Modified from Haber SN, Fudge JH, McFarland NR. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci*. 2000;20:2369.)

Immunohistochemical and receptor-binding studies have shown a discontinuity in the distribution of certain neurotransmitter-related substances that form the functional circuitry of the basal ganglia. For example, in the striatum, zones that contain a low density of acetylcholinesterase (AChE) enzymatic activity are surrounded by regions rich in AChE activity. The AChE-rich regions are referred to as the *matrix,* and the AChE-poor zones are termed either *striosomes* in primates or *patches* in rodents. The organization of several neuropeptide systems follows this organization. For example, the distributions of enkephalin,

substance P, and somatostatin immunoreactivity are organized in a similar manner as the AChE-rich and the AChE-poor areas in the striatum. In addition, in rodents, certain subtypes of dopamine receptors are present predominantly in one compartment compared with the other. In addition, the distribution of some afferent systems terminating in the striatum follows the striosome matrix organization. For example, afferents from the thalamus terminate preferentially in the matrix, rather than in the striosome.

Functional Circuitry. Projections into, within, and out of the basal ganglia are topographically organized and maintain this topography throughout the processing circuits of the basal ganglia. The existence of such patterns of connectivity has resulted in the hypothesis that parallel independent circuits in the basal ganglia process information from different regions of the brain and subserve separate complex functions. For example and as illustrated in Figure 1.2–30, there is an inverse dorsal–ventral topographic organization to the projection from the dorsal and ventral dopamine neurons to the striatum. Dorsally and medially located dopamine neurons project to the ventral and medial parts of the striatum, whereas ventrally and laterally located dopamine neurons project to dorsal and lateral parts of the striatum. Another prominent input to the striatum comes from the cerebral cortex, and this projection has a topographic organization related to that of the striatonigrostriatal pathway. Orbital and medial prefrontal cortex projects to the ventral striatum, the DLPFC projects to the central striatum, and premotor and motor cortices project to the dorsolateral striatum. These topographies create limbic, associative, and motor pathways within the corticostriatal and striatonigrostriatal projections.

INPUTS TO THE BASAL GANGLIA. The striatum is the major recipient of the inputs to the basal ganglia. Three major afferent systems are known to terminate in the striatum: the corticostriatal, nigrostriatal, and thalamostriatal afferents (Fig. 1.2–31). The corticostriatal projection originates from all regions of the neocortex, arising primarily from the pyramidal cells of layers V and VI, which use the excitatory neurotransmitter glutamate. A topography governing corticostriatal projections has been found in monkeys. Afferents from the sensorimotor cortex terminate predominantly in the putamen; association regions of the cortex terminate preferentially in the caudate nucleus. The prefrontal regions, in particular, provide a heavy input to the head of the caudate nucleus. In addition, afferents from the limbic cortical areas, the hippocampus, and the amygdala terminate in the ventral striatum. The second major class of afferents uses the neurotransmitter dopamine. In Figure 1.2–31, these projections are shown arising from the substantia

nigra pars compacta, but, as noted earlier (Fig. 1.2–30), different portions of the striatum receive input from the dorsal-tier or ventral-tier dopamine-containing neurons of the ventral mesencephalon. Electron microscopy studies have shown that many of the synapses formed by dopamine axon terminals on medium spiny neuron dendrites are immediately adjacent to the synapses provided by corticostriatal axons, suggesting that dopamine may play an important role in modulating the excitatory influence of cortical projections on striatal neurons. The third afferent system originates in the thalamus. The thalamic nuclei providing the projections are the intralaminar nuclei, particularly the central median nucleus.

Disruption of the input pathways of the basal ganglia has been associated with some movement disorders, such as Parkinson disease, which is characterized by muscular rigidity, fine tremor, shuffling gait, and bradykinesia. The most consistent neuropathological feature of Parkinson disease is a degeneration of the dopamine neurons in the substantia nigra pars compacta, accompanied by a loss of dopamine terminals in the striatum. The compound levodopa (Larodopa, Dopar), a precursor in the biosynthesis of dopamine, is used as a treatment for Parkinson disease because of its ability to augment the release of dopamine from the remaining terminals. Conversely, the administration of typical antipsychotic agents in the treatment of schizophrenia is frequently associated with parkinsonian features and other motor system abnormalities; the fact that these agents are D_2 dopamine receptor antagonists is thought to explain their movement-related side effects.

INTERNAL PROCESSING. The major processing pathways within the basal ganglia are summarized in Figure 1.2–32. As noted earlier, the striatum receives a major projection from the cerebral cortex. Within the striatum, the subclass of medium spiny neurons that contain the neuropeptide substance P sends inhibitory projections to the internal segment of the globus pallidus in what is termed the *direct pathway*. In contrast, the subpopulation of medium spiny neurons that contain the neuropeptide enkephalin provides inhibitory projections to the external segment of the globus pallidus, which sends inhibitory afferents to the internal segment of the globus pallidus in what is termed the *indirect pathway*. The globus pallidus external segment also projects to the pars reticulata of the substantia nigra. The topography found in the afferent projections to the striatum appears to be maintained in that processing pathway. For example, the sensorimotor territories of the striatum project most heavily to the ventral portion of the globus pallidus, whereas association territories project to the dorsal regions of the globus pallidus.

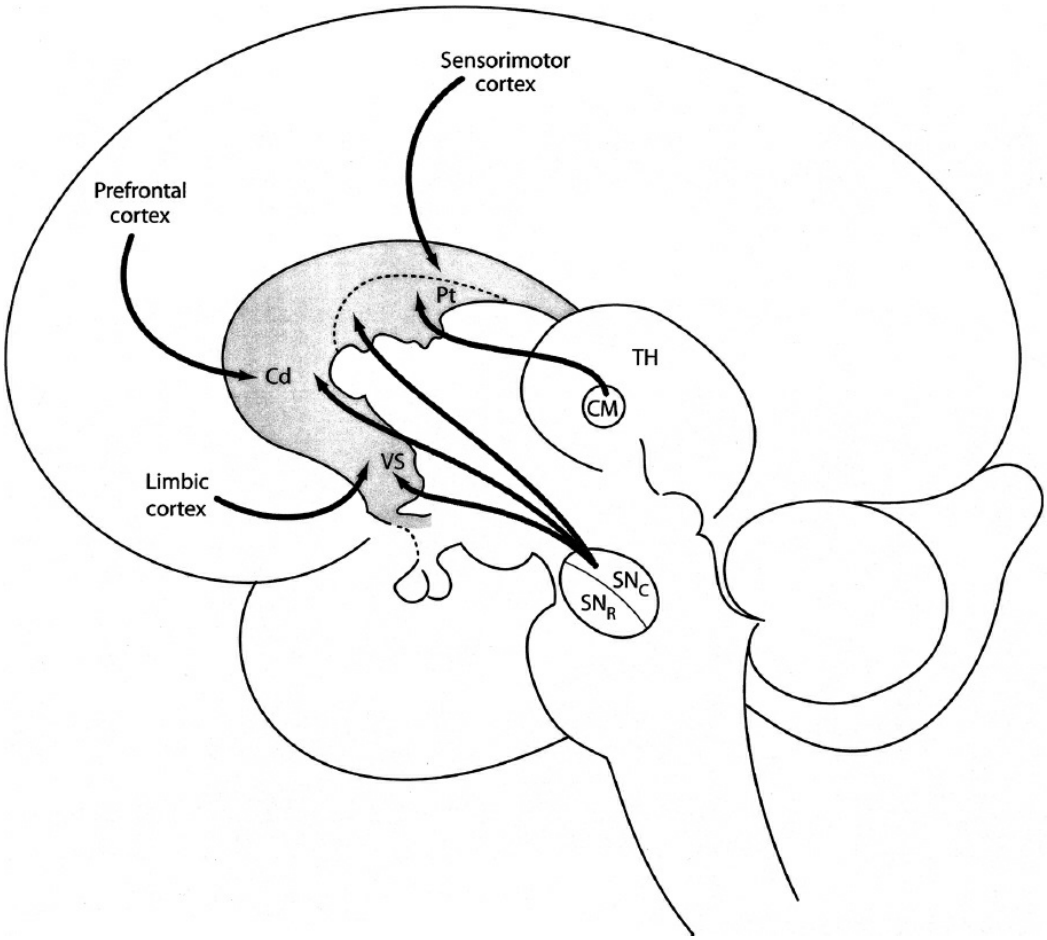

FIGURE 1.2–31. Diagram of the inputs to the basal ganglia system. Three major afferent systems have been identified: the corticostriatal, thalamostriatal, and nigrostriatal pathways.

The external segment of the globus pallidus also gives rise to an inhibitory projection that terminates in the subthalamic nucleus. In contrast, neurons in the subthalamic nucleus provide excitatory projections that terminate in both segments of the globus pallidus and in the pars reticulata. Although most connections within the basal ganglia are unidirectional, a reciprocal projection is found between the external segment of the globus pallidus and the subthalamic nucleus. The intrinsic circuitry of the basal ganglia is disrupted by a severe loss of neurons in the striatum in Huntington disease. This autosomal dominant disorder is characterized by progressive chorea and dementia. Although the Huntington disease gene has been identified, how the excessive number of trinucleotide repeats in this gene leads to the selective degeneration of striatal cells is currently a matter of intense investigation. More recent studies indicate that cortical neurons are also subject to degeneration in Huntington disease.

OUTPUT OF BASAL GANGLIA. The internal segment of the globus pallidus is the source of much of the output of the basal ganglia (Fig. 1.2–33). This

segment of the globus pallidus provides a projection to the ventral lateral and ventral anterior nuclei of the thalamus and to the intralaminar thalamic nuclei, particularly the central median nucleus. The pars reticulata of the substantia nigra also provides a projection to the ventral anterior and ventral lateral thalamic nuclei. These portions of the ventral lateral and ventral anterior thalamic nuclei project to the premotor and prefrontal cortices. As a result of the projections of the premotor and prefrontal cortices to the primary motor cortex, the basal ganglia are able to influence indirectly the output of the primary motor cortex. In addition, the cortical output of the basal ganglia exhibits marked convergence; although the striatum receives afferents from all regions of the neocortex, the eventual output of the globus pallidus and pars reticulata is largely conveyed through the thalamus to a much smaller portion of the neocortex —the premotor and prefrontal regions.

The functional consequences of the neural circuitry of the basal ganglia can also be considered in the context of some of the neurotransmitters used (Figs. 1.2–32 and 1.2–33). Because the afferents from the cortex are thought to use glutamate, which is an excitatory neurotransmitter, cortical afferents presumably excite the structures of the basal ganglia in which they terminate. Many of the processing pathways within the basal ganglia use the inhibitory neurotransmitter GABA. Finally, the output pathways of the basal ganglia—the globus pallidus and the substantia nigra pars reticulata—use GABA as well. Excitation from cortical afferents eventually disinhibits the target structures of the basal ganglia because of the back-to-back inhibitory pathways of the basal ganglia.

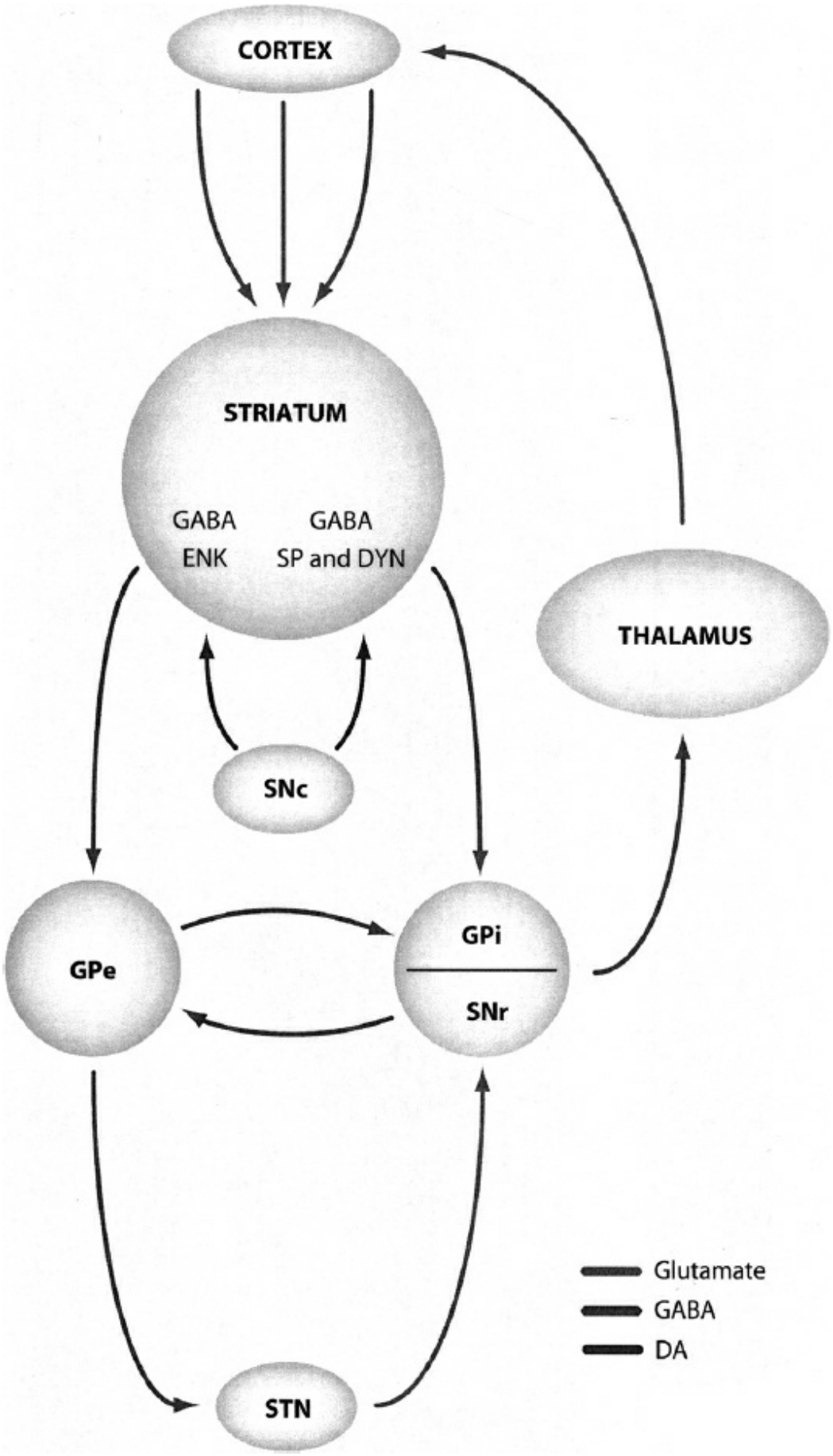
CORTEX
STRIATUM
GABA
ENK
GABA
SP and DYN
THALAMUS
SNc
GPe
GPi
SNr
STN
Glutamate
GABA
DA

FIGURE 1.2–32. Diagram of the intrinsic circuitry of the basal ganglia. Substance P (SP)–containing striatal neurons send an inhibitory projection directly to the internal segment of the globus pallidus, whereas neurons containing enkephalin provide an inhibitory projection to GABA neurons in the external segment of the globus pallidus, which project to the internal segment of the globus pallidus. The subthalamic nucleus receives a projection from the external segment of the globus pallidus and projects back to both segments. Finally, the subthalamic nucleus and globus pallidus external project to the substantia nigra pars reticulata.

Historically, motor systems have been divided into pyramidal (corticospinal) and extrapyramidal (basal ganglia) components; this division is based on clinical findings suggesting that lesions of each system result in distinct motor syndromes. For example, lesions of the extrapyramidal system result in involuntary movements, changes in muscle tone, and slowness of movement; lesions of the pyramidal system lead to spasticity and paralysis. Because of these findings, the pyramidal and extrapyramidal systems were thought to control voluntary and involuntary movement independently. However, this division is no longer accurate for several reasons. First, other structures of the brain outside the traditional pyramidal and extrapyramidal systems, such as the cerebellum, are involved in the control of movement. Second, the pyramidal and extrapyramidal systems are not independent; the neural circuits of these systems are interconnected. For example, the basal ganglia influence motor behavior through certain regions of the cerebral cortex, which then directly (through the corticospinal tract) or indirectly (through specific brainstem nuclei) produce motor activity.

Finally, although the basal ganglia are important in the control of movement, this neural system also seems to be involved in other functions of the brain. More recent studies of the connections of the basal ganglia in NHPs also support a role for these structures in cognitive functions. The DLPFC has been shown to receive inputs from portions of the thalamus that are the targets of projections from specific locations within the internal segment of the globus pallidus, providing evidence for a distinct pallidothalamocortical pathway. In addition to linking association regions of the cerebral cortex, such as the prefrontal and posterior parietal areas, with the control of motor activity in the primary motor cortex, some of the output of the basal ganglia seems to be directed back to regions of the prefrontal cortex. These findings suggest that “closed” loops are present between the prefrontal cortex and basal ganglia, which presumably have a cognitive rather than a motor function.

Limbic System

The concept of the limbic system as a neural substrate for emotional experience and expression has a rich but controversial history. More than 100 years ago, Paul Broca applied the term *limbic* (from the Latin *limbus,*

meaning "border") to the curved rim of the cortex, including the cingulate and the parahippocampal gyri, located at the junction of the diencephalon and the cerebral hemispheres (Fig. 1.2–34). In 1937, James Papez postulated, primarily on the basis of anatomic data, that these cortical regions were linked to the hippocampus, mammillary body, and anterior thalamus in a circuit that mediated emotional behavior (Fig. 1.2–35). This concept was supported by the work of Heinrich Klüver and Paul Bucy, who showed that temporal lobe lesions, which disrupt components of the circuit, alter affective responses in NHPs. In 1952, Paul MacLean coined the term *limbic system* to describe Broca limbic lobe and related subcortical nuclei as the neural substrate for emotion. However, over the last 40 years, it has become clear that some limbic structures (e.g., the hippocampus) are also involved in other complex brain processes, such as memory. In addition, expanding knowledge of the connectivity of traditional limbic structures has made it increasingly difficult to define the boundaries of the limbic system. Despite these limitations, the concept of a limbic system may still be a useful way to describe the circuitry that relates certain telencephalic structures and their cognitive processes with the hypothalamus and its output pathways that control autonomic, somatic, and endocrine functions.

Major Structures. As suggested earlier, no unanimity exists on the brain structures that constitute the limbic system. This section includes the brain regions that are most commonly listed as components of the limbic system: the cingulate and parahippocampal gyri (limbic cortex), the hippocampal formation, the amygdala, the septal area, the hypothalamus, and related thalamic and cortical areas.

LIMBIC CORTEX. The limbic cortex is composed of two general regions, the cingulate gyrus and the parahippocampal gyrus (Fig. 1.2–34). The *cingulate gyrus,* a C-shaped stretch of cortex that lies dorsal to and follows the corpus callosum, includes several cortical regions that are defined by their relationship to the corpus callosum (Fig. 1.2–36). For example, the subgenual anterior cingulate lies ventral to the corpus callosum, extending from the anterior boundary of the genu to the rostrum (Fig. 1.2–36). The subgenual anterior cingulate is composed of Brodmann area 25 and subcallosal portions of Brodmann areas of 32 and 24. The subgenual anterior cingulate is connected to other structures of the limbic system, such as the amygdala and hypothalamus, brain regions also often associated with depression. Recent imaging studies reveal that the subgenual anterior cingulate is overactive in people with depression and that a variety of treatments, such as antidepressants, electroconvulsive therapy (ECT), and repetitive transcranial magnetic stimulation (rTMS), result in a decrease of activity in this area. Moreover, in preliminary studies

of patients with treatment resistant depression, deep brain stimulation (DBS) of the subgenual anterior cingulate and underlying white matter has proven effective in the alleviation of symptoms. It is important to note that the response documented after DBS stimulation likely results from not only direct activation of the subgenual anterior cingulate, but also the activation of brain regions directly connected to it, including the amygdala, hypothalamus, and hippocampus. These areas are connected to the subgenual anterior cingulate via the callosum bundle, the uncinate fasciculus, and the subcortical fascicle and stimulation of these white matter tracts is necessary for the antidepressant response to DBS. Moreover, transsynaptic activation of areas not directly connected to the subgenual anterior cingulated, such as the DLPFC and the insula, might also play a role in the antidepressant effect of DBS.

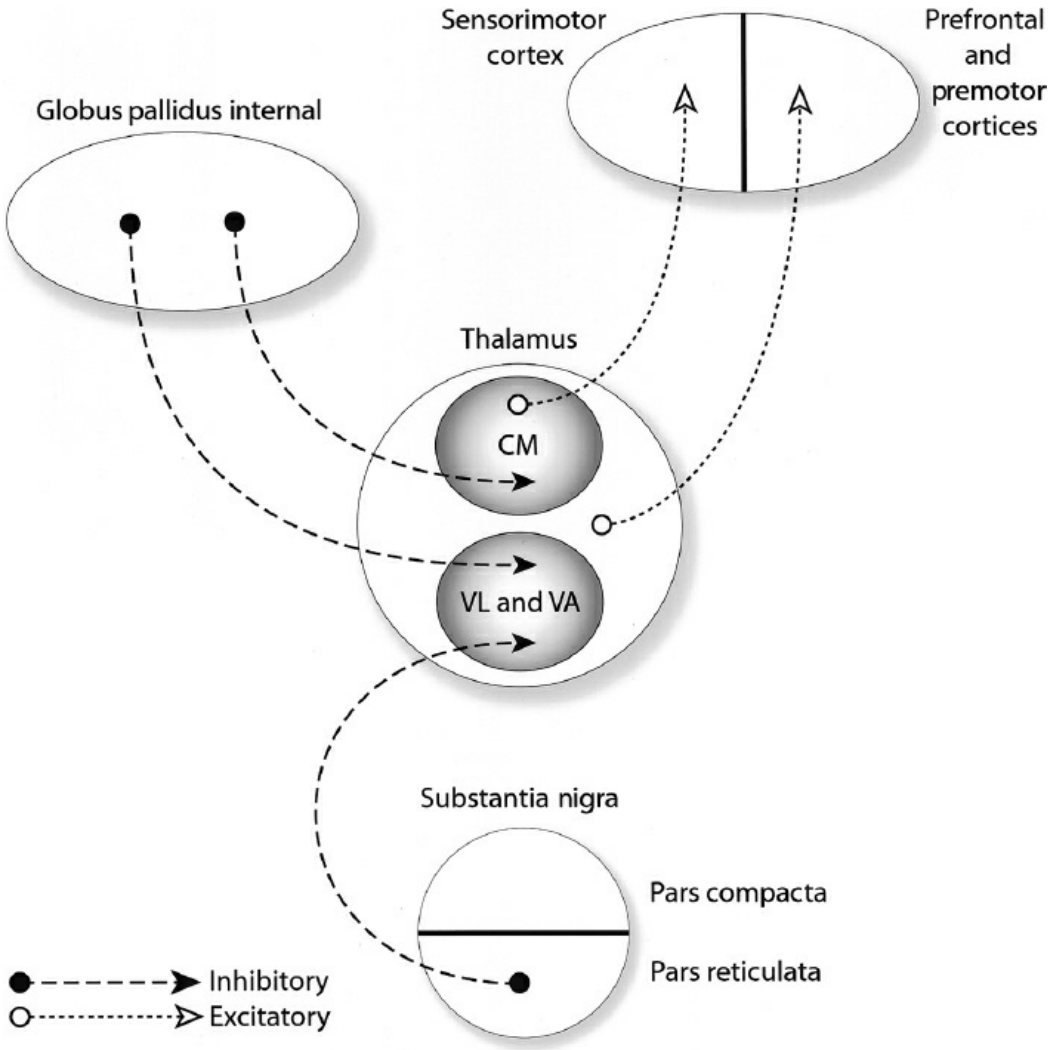

FIGURE 1.2–33. Diagram of the output of the basal ganglia system. The internal segment of the globus pallidus projects to the central median (*CM*), ventral lateral (*VL*), and ventral anterior (*VA*) nuclei of the thalamus. Those nuclei then project to sensorimotor, prefrontal, and premotor cortices. The substantia nigra pars reticulata also projects to the VL and VA nuclei.

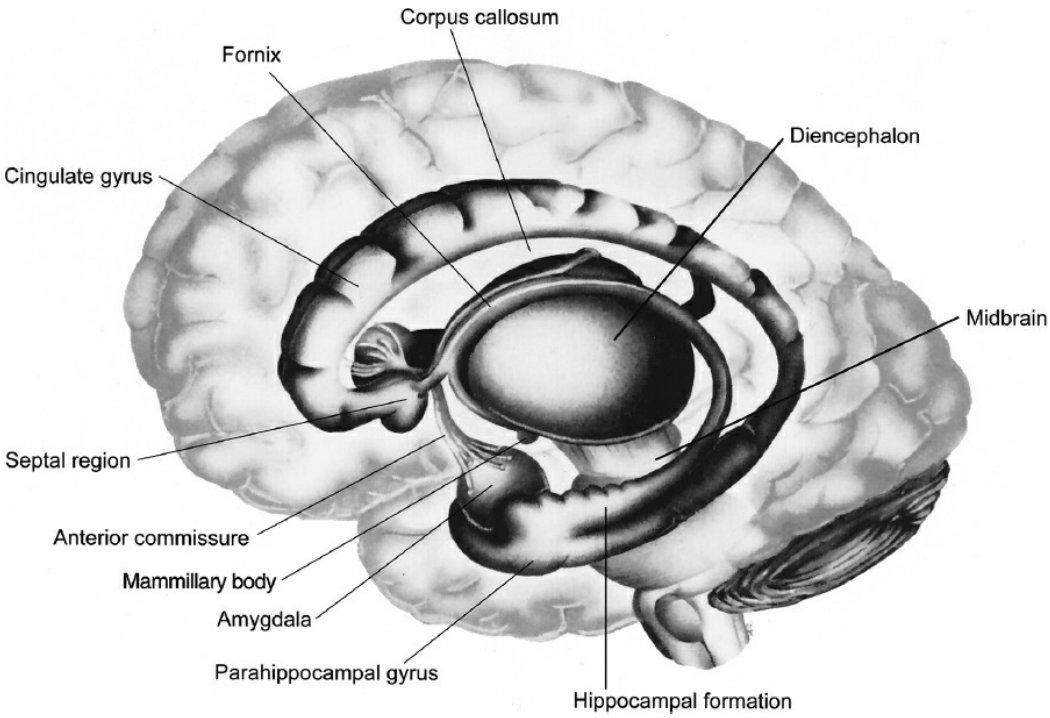

FIGURE 1.2–34. Schematic drawing of the major anatomic structures of the limbic system. The cingulate and parahippocampal gyri form the "limbic lobe," a rim of tissue located along the junction of the diencephalon and the cerebral hemispheres. (Adapted from Hendelman WJ. *Student's Atlas of Neuroanatomy*. Philadelphia, PA: WB Saunders; 1994:179.)

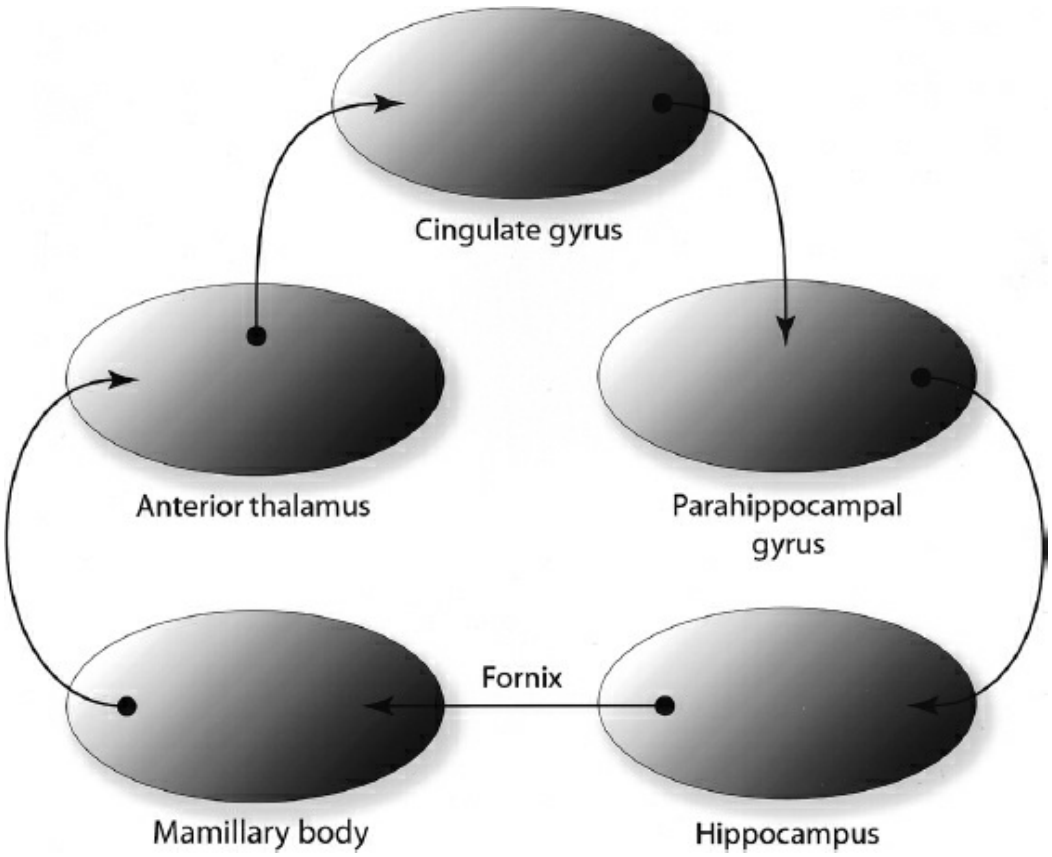

FIGURE 1.2–35. Diagram of the neural circuit for emotion as originally proposed by James Papez.

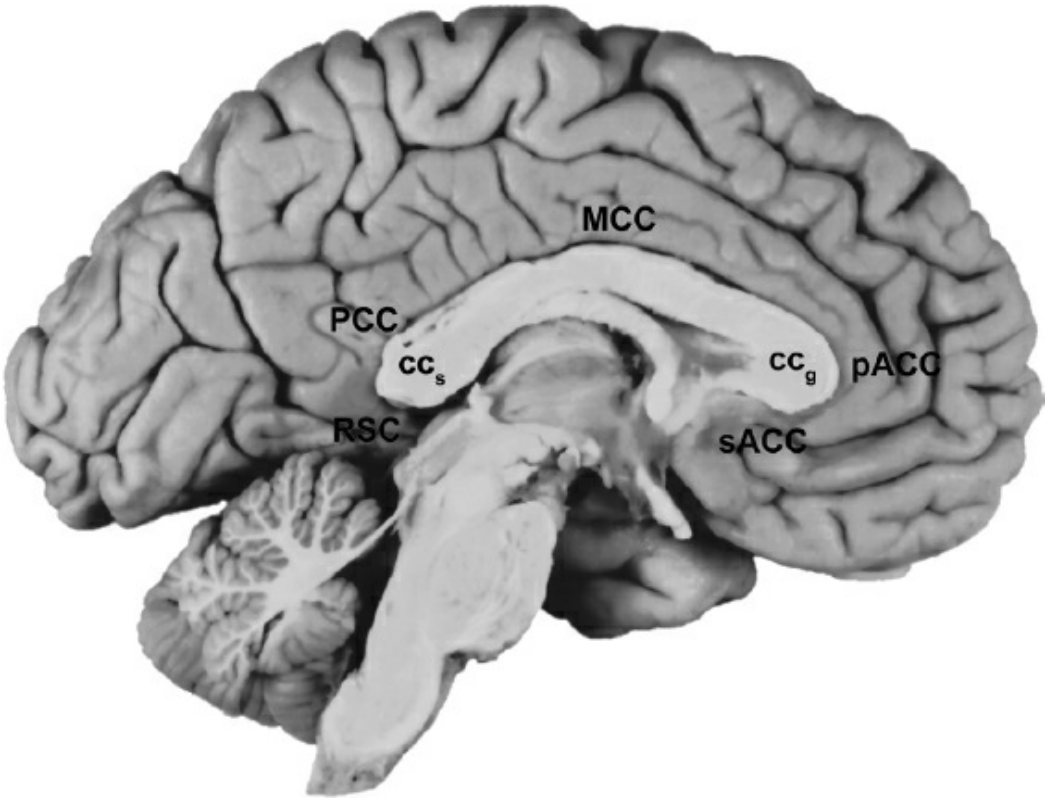

FIGURE 1.2–36. Photograph of the medial aspect of the left hemisphere of a human brain indicating the different subdivisons of the cingulate gyrus. sACC, subgenual anterior cingulate cortex; pACC, pregenual anterior cingulate cortex; MCC, midcingulate cortex; PCC, posterior cingulate cortex; RSC, retrosplenial cingulate cortex; cc_g, genu of the corpus callosum; ccs, splenium of the corpus callosum.

The habenula, although technically part of the epithalamus, has significant connections within the limbic system and is divided into medial and lateral subdivisions. The lateral habenula serves as a hub connecting limbic, basal ganglia, and brainstem circuits. These links, specifically those involving projections from the lateral habenula to serotonergic neurons of the raphe nuclei and dopaminergic cells of the VTA, suggest that the lateral habenula may be involved in mood regulation and disorders, such as depression. Preliminary studies targeting the lateral habenula for DBS have been successful and open the door to further study of this brain region in the pathophysiology and treatment of depression.

As illustrated in Figure 1.2–36, the cingulate gyrus is extensive and as the cingulate gyrus travels posteriorly, it becomes continuous (via the cingulum bundle of fibers in the white matter) with the *parahippocampal gyrus,* located in the MTL, which contains several distinct cytoarchitectonic regions. One of the most important of these regions is the *entorhinal cortex,* which not only funnels highly processed cortical information to the hippocampal formation, but also is a major output pathway from the hippocampal formation.

HIPPOCAMPAL FORMATION. The hippocampal formation comprises three

distinct zones—the dentate gyrus, the hippocampus, and the subicular complex—and is located in the floor of the temporal horn of the lateral ventricle (Fig. 1.2–10). These zones are composed of adjacent strips of cortical tissue that run in a rostral–caudal direction, but fold over each other mediolaterally in a spiral fashion, resulting in a C-shaped appearance. The *dentate gyrus* comprises three layers: an outer, acellular molecular layer, which faces the subarachnoid space of the hippocampal fissure; a middle layer composed of granule cells; and an inner polymorphic layer (Fig. 1.2–37). The granule cells extend their dendritic trees into the molecular layer and give rise to axons that form the mossy fiber projection to the hippocampus.

The *hippocampus* is also a trilaminate structure composed of molecular and polymorphic layers and a middle layer that contains pyramidal neurons. On the basis of differences in the cytoarchitecture and connectivity, the hippocampus can be divided into three distinct fields, which have been labeled CA3, CA2, and CA1 (Fig. 1.2–37). (*CA* is derived from the term *cornu ammonis* after the Egyptian deity Ammon, who was depicted with ram's horns, which some early investigators thought described the shape of the hippocampus.) Some disagreement exists regarding the so-called CA4 region. This terminology has been applied to the portion of the hippocampus adjacent to CA3 and within the "C" created by the granule cell layer of the dentate gyrus. Connectional studies have revealed, however, that this area is more closely related to the dentate gyrus and should be referred to more appropriately as the hilar region or hilus of the dentate gyrus (Fig. 1.2–37). The white matter adjacent to the polymorphic layer of the hippocampus is known as the *alveus*. The axons in this structure contribute to the fimbria, which, at the caudal end of the hippocampus, becomes the crus of the *fornix*. These bilateral structures converge to form the body of the fornix, which travels anteriorly and then turns inferiorly to form the columns of the fornix, which pass through the hypothalamus into the mammillary bodies (Fig. 1.2–38). The *subicular complex* is generally considered to have three components—the presubiculum, the parasubiculum, and the subiculum—that together serve as transition regions between the hippocampus and the parahippocampal gyrus.

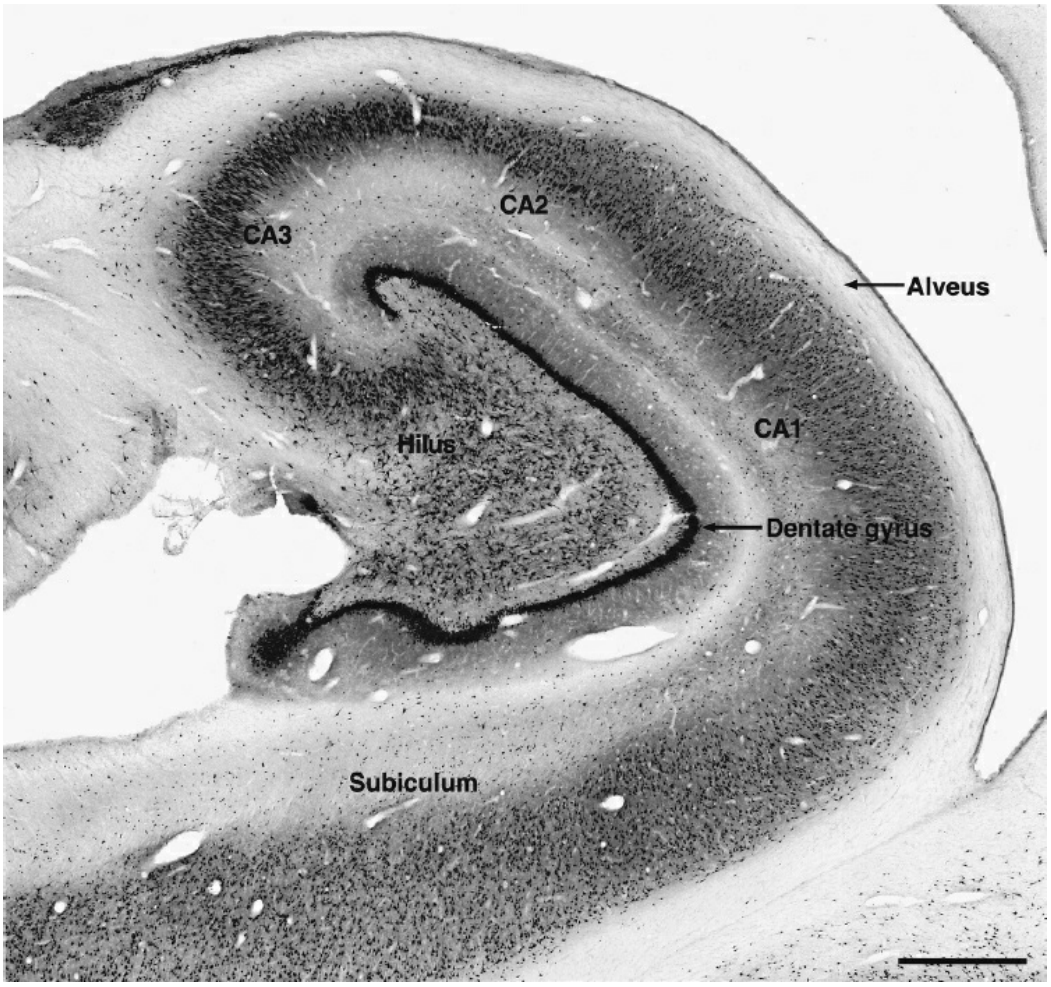

FIGURE 1.2–37. Photomicrograph of neurons immunoreactive for neuron-specific nuclear protein in the human hippocampal formation. The immunostaining illustrates the major components of the hippocampal formation, such as the dentate gyrus. Scale bar = 1 mm.

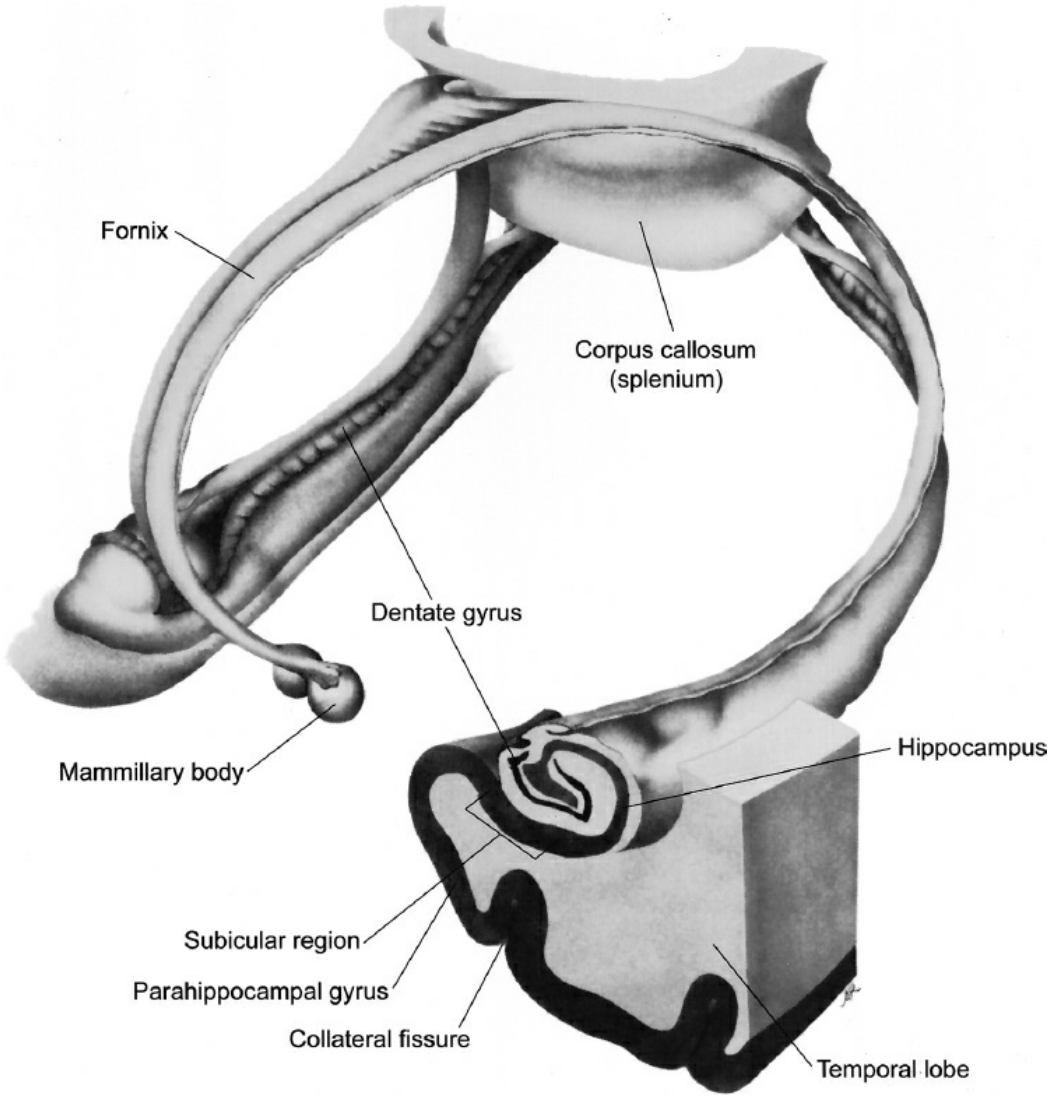

FIGURE 1.2–38. Schematic drawing of a cross-sectional view of the hippocampal formation and the path of the fornix running between that structure and the mammillary bodies. (Adapted from Hendelman WJ. *Student's Atlas of Neuroanatomy*. Philadelphia, PA: WB Saunders; 1994:189.)

The components of the hippocampal formation have a distinct pattern of intrinsic connectivity that is largely unidirectional and provides for a specific flow of information. The major input to the hippocampal formation arises from neurons in layers II and III of the entorhinal cortex that project through the *perforant path* (i.e., through the subiculum and the hippocampus) to the outer two-thirds of the molecular layer of the dentate gyrus, where they synapse on the dendrites of granule cells. The mossy fiber axons of the granule cells provide a projection to the pyramidal neurons of the CA3 field of the hippocampus. Axon collaterals from CA3 pyramidal neurons project within CA3 and, through the so-called Schäffer collaterals, to the CA1 field of the hippocampus. This region projects to the subicular complex, which provides output to the entorhinal cortex, completing the circuit.

AMYGDALA. Located in the MTL just anterior to the hippocampal formation are a group of nuclei referred to as the *amygdala* (Fig. 1.2–9). These nuclei form several distinct clusters: the basolateral complex, the centromedial amygdaloid group, and the olfactory group, which includes the cortical amygdaloid nuclei. These nuclei are usually delineated using cytoarchitectural features revealed by Nissl stains. However, the chemoarchitecture of cannabinoid CB_1 receptor–containing axons also clearly demarcates these nuclei (Fig. 1.2–39). CB_1 receptor immunoreactivity is found within the basolateral nuclei, the largest of the three groups, whereas the central and medial nuclei are devoid of CB_1 receptor labeling. The *basolateral complex* differs from the remaining amygdaloid nuclei in many respects. Although the basolateral complex is not a laminated structure, its connectivity and some other anatomic characteristics are more similar to cortical regions than to the remaining amygdaloid nuclei. For example, the basolateral nuclei are directly and reciprocally connected with the temporal, insular, and prefrontal cortices. In addition, similar to some cortical regions, the basolateral complex shares bidirectional connections with the medial dorsal thalamic nucleus and receives projections from the midline and intralaminar thalamic nuclei. Finally, neurons of the basolateral complex with a pyramidal-like morphology seem to furnish projections to the striatum that use excitatory amino acids as neurotransmitters. On the basis of these anatomic characteristics, one may hypothesize that the basolateral complex actually functions similar to a multimodal cortical region.

In contrast, the *centromedial amygdala* appears to be part of a larger structure that is continuous through the sublenticular substantia innominata with the bed nucleus of the stria terminalis (BNST). This larger structure, which has been termed the *extended amygdala,* consists of two major subdivisions. The central subdivision of the extended amygdala includes the central amygdaloid nucleus and the lateral portion of the BNST. This subdivision is reciprocally connected with brainstem viscerosensory and visceromotor regions and with the lateral hypothalamus. In addition, it receives afferents from cortical limbic regions and the basolateral amygdaloid complex. In contrast, the medial subdivision of the extended amygdala, composed of the medial amygdaloid nucleus and its extension into the medial part of the BNST, is distinguished by reciprocal connections with the medial or endocrine portions of the hypothalamus.

SEPTAL AREA. The *septal area* is a gray matter structure located immediately above the anterior commissure (Fig. 1.2–40). The septal nuclei are reciprocally connected with the hippocampus, the amygdala, and the hypothalamus, and project to numerous structures in the brainstem.

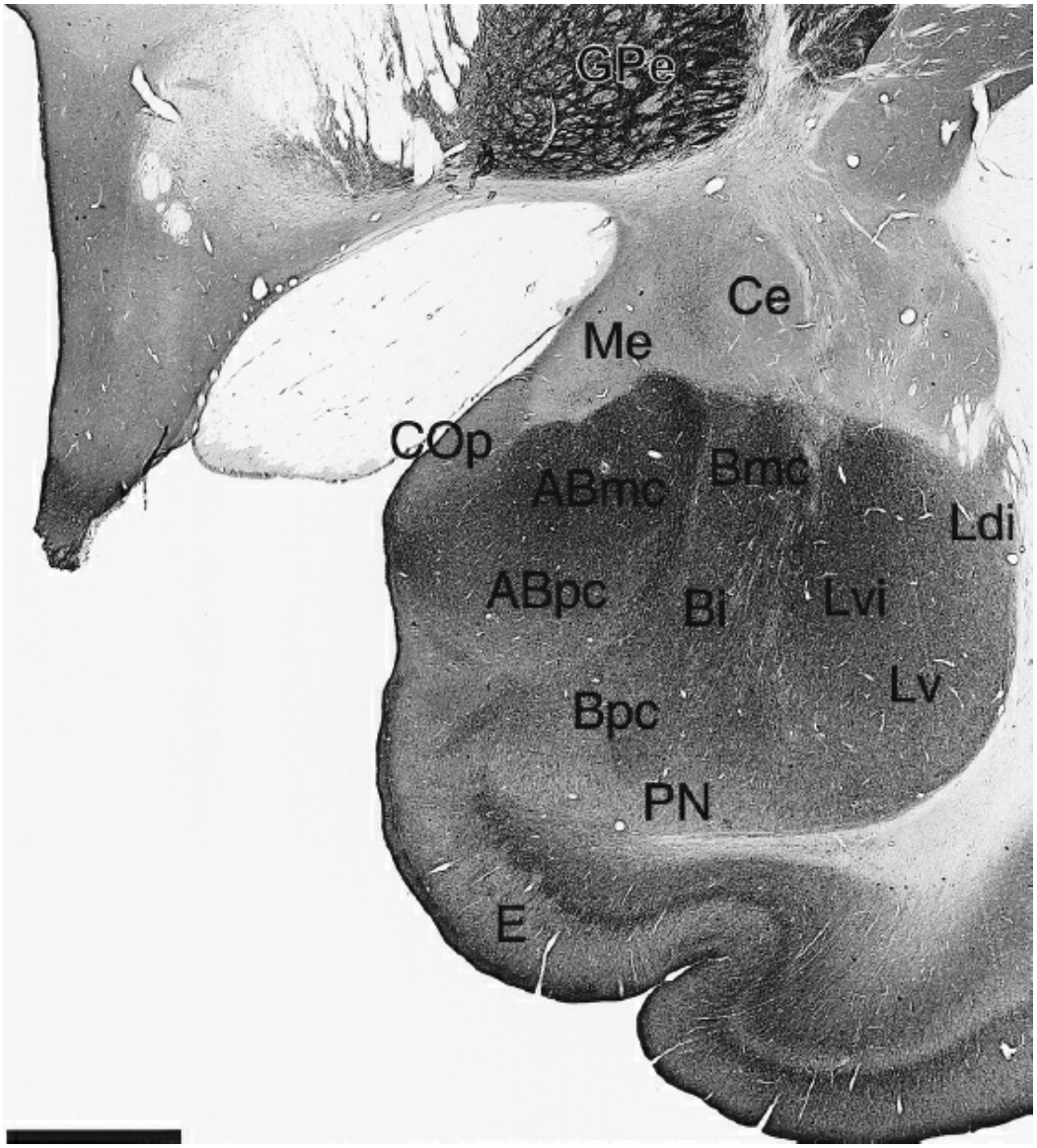

FIGURE 1.2–39. Photomicrograph of a coronal section through macaque monkey brain illustrating the distribution of the cannabinoid CB_1-immunoreactive axons in the amygdala. The density of labeled axons is high in the cortical-like basolateral nuclei (*ABmc, ABpc, Bi, Bmc, Bpc, Ldi, Lv, Lvi*), whereas the striatal-like central (*Ce*) and medial (*Me*) nuclei are devoid of CB_1-immunoreactive axons. ABmc, accessory basal nucleus, magnocellular division; ABpc, accessory basal nucleus, parvicellular division; Bi, basal nucleus, intermediate division; Bmc, basal nucleus, magnocellular division; Bpc, basal nucleus, parvicellular division; Ce, central amygdaloid nucleus; Cop, posterior cortical nucleus; E, entorhinal cortex; GPe, globus pallidus, external; Ldi, lateral nucleus, dorsal intermediate division; Lv, lateral nucleus, ventral division; Lvi, lateral nucleus, ventral intermediate division; Me, medial amygdaloid nucleus; PN, paralaminar nucleus. Calibration bar = 2 mm. (Adapted from Eggan SM, Lewis DA. Immunocytochemical distribution of the cannabinoid CB1 receptor in the primate neocortex: A regional and laminar analysis. *Cereb Cortex*. 2007;17:175.)

HYPOTHALAMUS. The hypothalamus, a small structure within the diencephalon, is a crucial component of the neural circuitry regulating not only emotions, but also autonomic, endocrine, and some somatic functions. In addition to its relationships with other components of the limbic system,

it is interconnected with various visceral and somatic nuclei of the brainstem and the spinal cord and provides an output that regulates the function of the pituitary gland. On its inferior surface, the hypothalamus is bounded rostrally by the optic chiasm and caudally by the posterior edge of the mammillary bodies. The area of the hypothalamus between these two structures, called the *tuber cinereum,* gives rise to the median eminence, which is continuous with the infundibular stalk and then the posterior lobe of the pituitary gland (Fig. 1.2–41). On the basis of these features, the hypothalamus is subdivided from anterior to posterior into three zones: the supraoptic region, the infundibular region, and the mammillary region. (In addition, the preoptic area, a telencephalic structure located immediately anterior to the supraoptic region, is usually considered part of the hypothalamus.) These three zones also are divided on each side into medial and lateral areas by the fornix as it travels through the body of the hypothalamus to the mammillary bodies. As shown in Table 1.2–4, the six parts of the hypothalamus contain different nuclei.

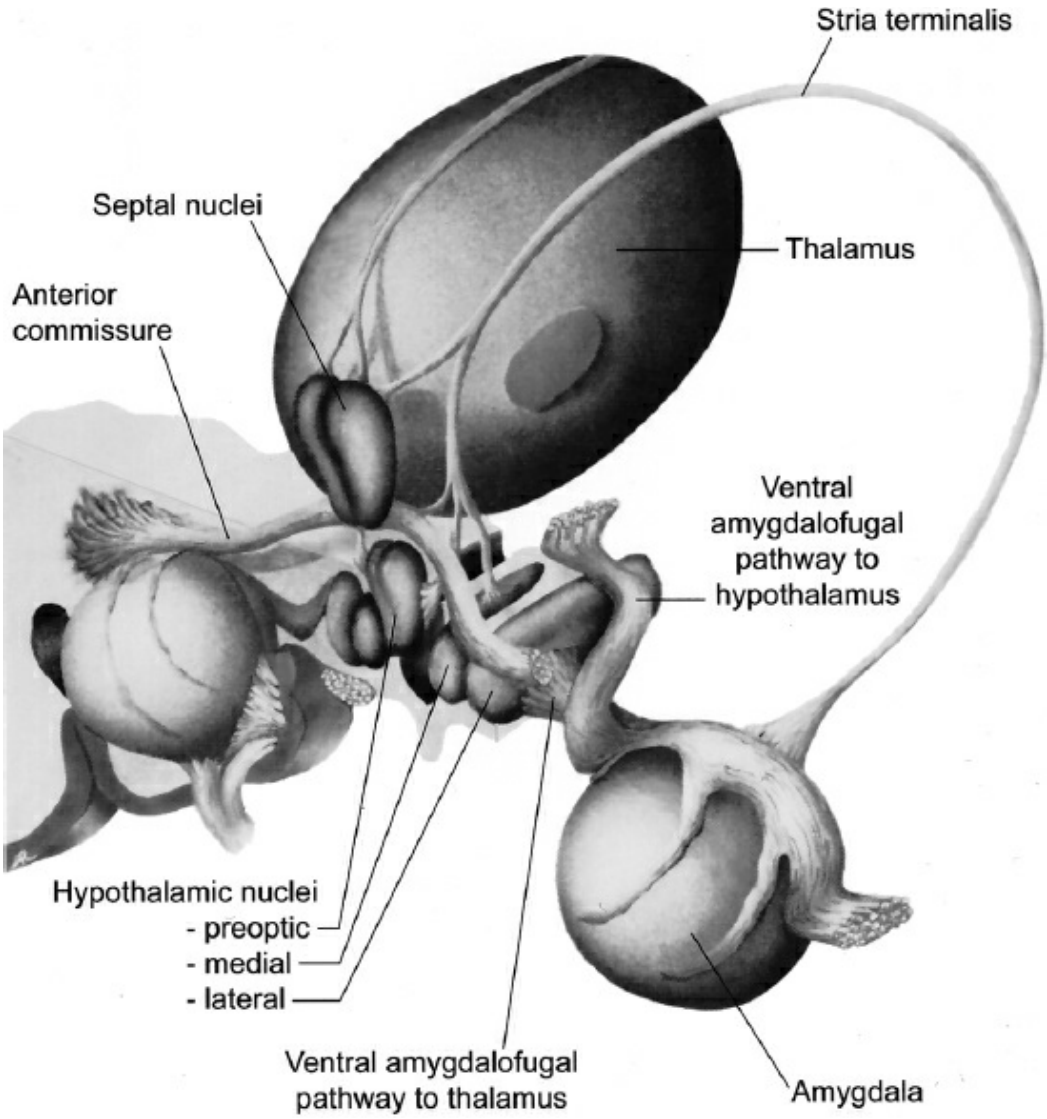

FIGURE 1.2–40. Schematic drawing of some components of the limbic system showing the major output pathways of the amygdala, the stria terminalis, and the ventral amygdalofugal pathway. (Adapted from Hendelman WJ. *Student's Atlas of Neuroanatomy*. Philadelphia, PA: WB Saunders; 1994:183.)

These different nuclei subserve the diverse functions of the hypothalamus. The *suprachiasmatic nucleus* (*SCN*) receives direct and indirect projections from the retina and seems to be important in the regulation of diurnal rhythms. The *supraoptic* and *paraventricular nuclei* contain large cells (magnocellular neurons) that send oxytocin-containing and vasopressin-containing fibers to the posterior neural lobe of the pituitary. In addition, some neurons of the paraventricular nucleus (PVN) project to the median eminence, where they release neuropeptides, such as corticotropin-releasing factor (CRF), into the portal blood system. These neuropeptides control the synthesis and release of anterior pituitary hormones. The PVN also gives rise to descending projections that regulate the sympathetic and parasympathetic autonomic areas of the medulla and the spinal cord.

Within the medial tuberal region of the hypothalamus, the *ventromedial* and *arcuate nuclei* also participate in the regulation of the anterior pituitary function. In addition, the ventromedial nucleus may play an important role in reproductive and ingestive behavior. The medial posterior section of the hypothalamus contains the *posterior nucleus* and the *mammillary bodies*. Within the mammillary bodies, the lateral and medial mammillary nuclei receive hippocampal input through the fornix (Fig. 1.2–38) and project to the anterior nuclei of the thalamus. The posterior nucleus shares reciprocal connections with the extended amygdala. This nucleus appears to be relatively more developed in primates than in rodents, suggesting that it plays an important but still-to-be-clarified role in the human brain.

The lateral portions of the hypothalamus contain a low density of neurons scattered among longitudinally running fibers of the medial forebrain bundle. This region is interconnected with multiple regions of the forebrain, the brainstem, and the spinal cord. The lateral hypothalamic area (LHA) also contains a population of neurons that express the orexin neuropeptides, orexin A and orexin B (also known as hypocretin A and hypocretin B), which seem to be involved in sleep and wakefulness. The approximately 70,000 orexin-producing neurons in the human brain project throughout the brain, with the exception of the cerebellum (Fig. 1.2–42). Orexin neurons project to most of the monoaminergic (i.e., substantia nigra, locus coeruleus, dorsal raphe) and cholinergic (i.e., medial septum, pedunculopontine, laterodorsal tegmental [LDT]) nuclei. Orexin neurons also have widespread projections throughout the cerebral cortex. Areas containing high densities of orexin axons include the

paraventricular thalamic nucleus, the arcuate nucleus of the hypothalamus, the locus coeruleus, and the dorsal raphe nucleus. The projections of orexin neurons to neuronal systems involved in sleep and wakefulness (i.e., locus coeruleus, raphe nuclei, and laterodorsal/pedunculopontine tegmental [PPT] nuclei) suggest that orexin neurons participate in these functions. Numerous studies in animals and humans show that an orexin deficiency is the main cause of narcolepsy. For example, mice that lack the orexin gene exhibit physiologic symptoms similar to human narcolepsy, and postmortem examination of the brains of narcolepsy patients have revealed an 85 to 95 percent reduction in the number of orexin-immunoreactive neurons.

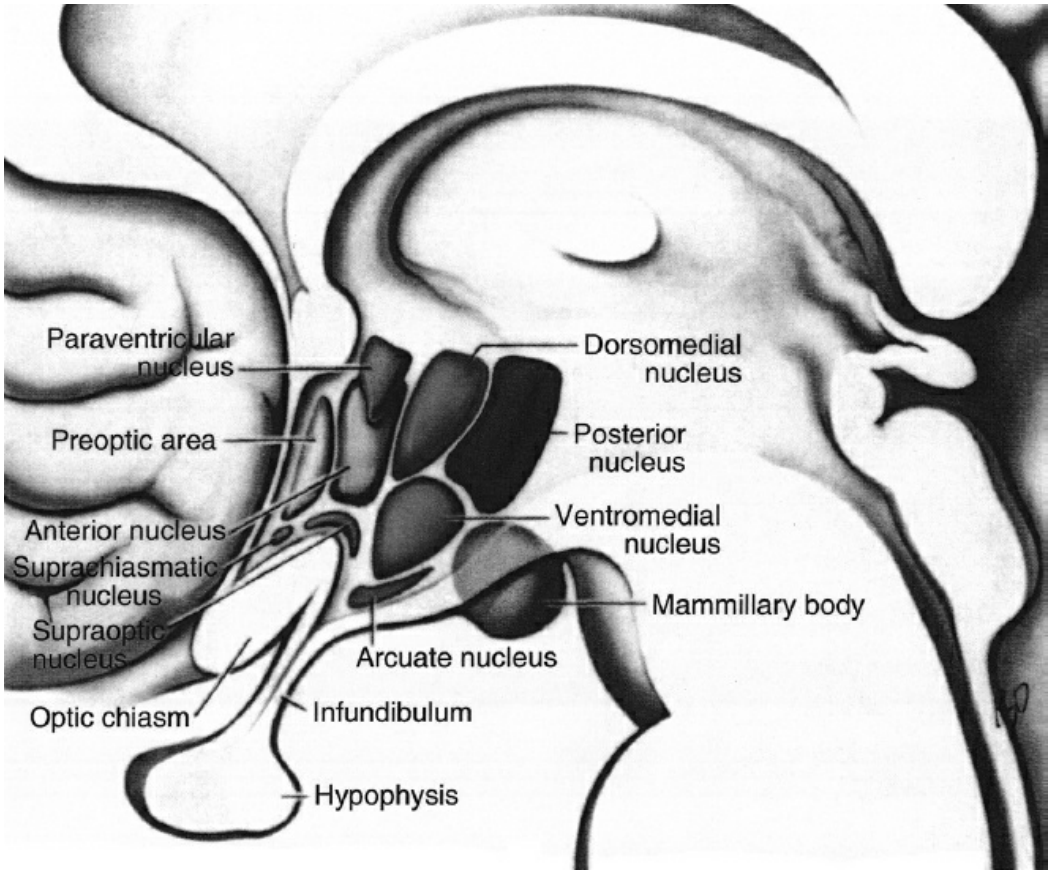

FIGURE 1.2–41. Schematic drawing of the nuclei in the medial hypothalamus. (Modified from Parent A. *Carpenter's Human Neuroanatomy*, 9th ed. Media, PA: Williams & Wilkins; 1996:707.)

Table 1.2–4
Hypothalamic Nuclei

Hypothalamic Regions		Periventricular Zone	Medial Zone	Lateral Zone
Anterior	Preoptic	Preoptic nucleus Periventricular nuclei	Medial preoptic nucleus	Lateral preoptic nucleus
	Supraoptic	Suprachiasmatic nucleus	Anterior hypothalamic nucleus	Lateral hypothalamic nucleus

		Periventricular nuclei	Paraventricular nucleus	
			Supraoptic nucleus	
Middle	Infundibular	Arcuate nucleus	Dorsomedial nucleus	Lateral tuberal nuclei
			Ventromedial nucleus	Lateral hypothalamic nucleus
Posterior	Mammillary		Mammillary nuclei	Lateral hypothalamic nucleus
			Posterior hypothalamic nuclei	

Modified from Patestas MA, Gartner LP. *A Textbook of Neuroanatomy*. Malden, MA: Blackwell; 2006:363.

Functional Circuitry. The major structures of the limbic system are interconnected with each other and with other components of the nervous system in various ways. Several major output pathways of the limbic system are clearly defined. In one pathway (Fig. 1.2–43), highly processed sensory information from the cingulate, the orbital and temporal cortices, and the amygdala is transmitted to the entorhinal cortex of the parahippocampal gyrus and then to the hippocampal formation. After traversing the intrinsic circuitry of the hippocampal formation, information is projected through the fornix either to the anterior thalamus, which projects to the limbic cortex, or to the septal area and the hypothalamus. These latter two regions provide feedback to the hippocampal formation through the fornix. In addition, the mammillary bodies of the hypothalamus project to the anterior thalamus. Finally, the hypothalamus and the septal area project to the brainstem and the spinal cord.

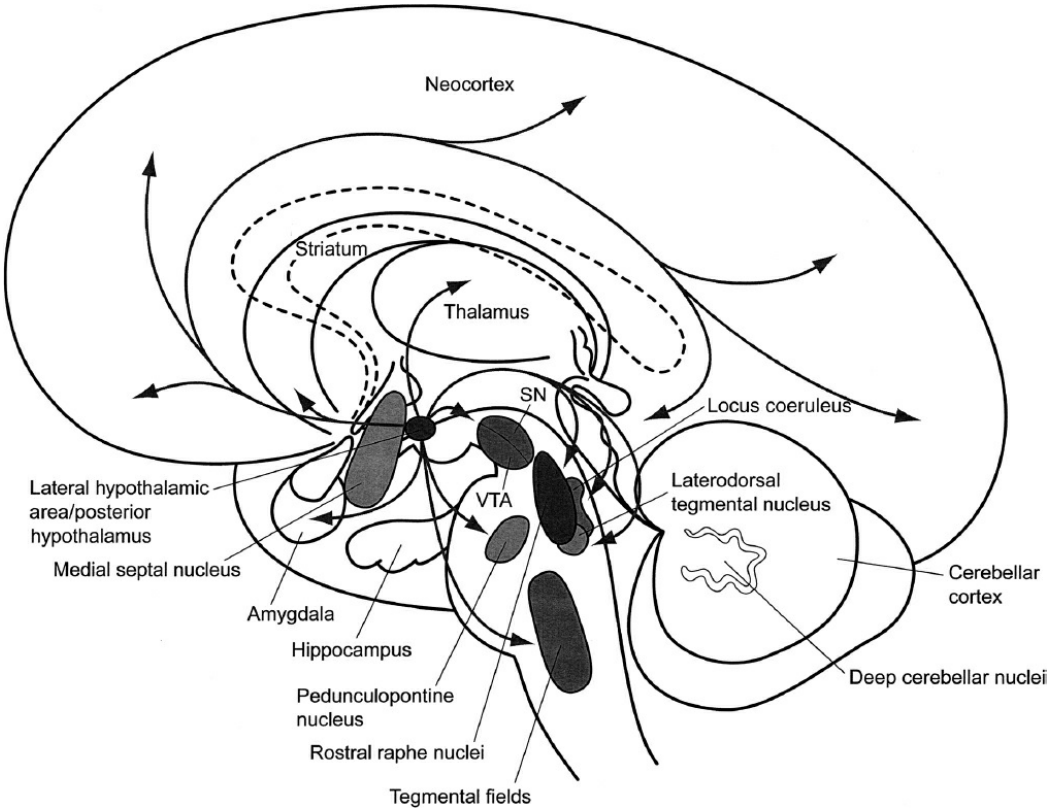

FIGURE 1.2–42. Schematic drawing illustrating the circuitry of orexin neurons.

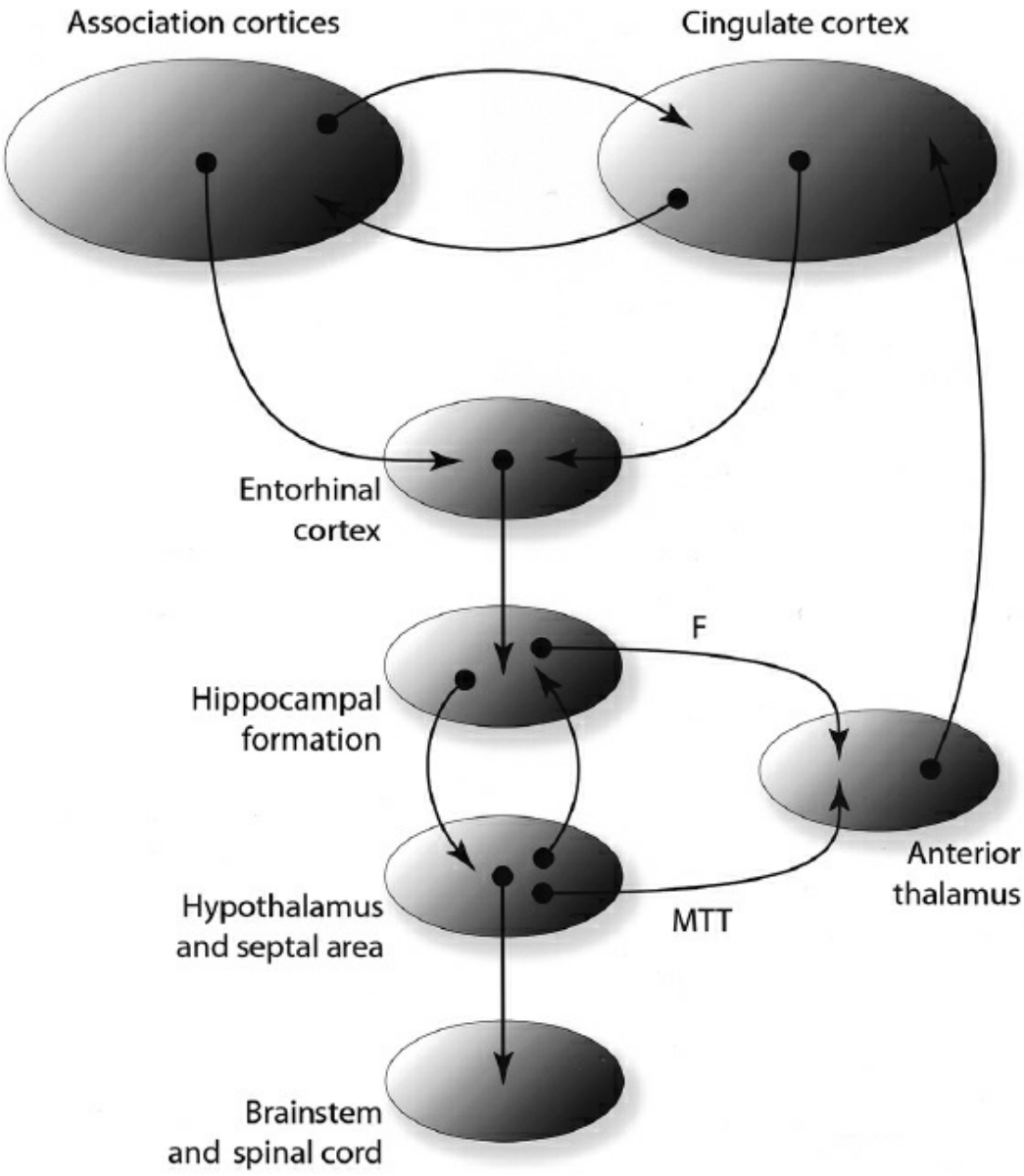

FIGURE 1.2–43. Functional neural circuitry of the limbic system. This diagram illustrates the manner in which the hippocampal formation and the anterior thalamus provide a mechanism for the integration of information between the cerebral cortex and the hypothalamus. F, fornix; MTT, mammillothalamic tract. (Adapted from Nolte J. *The Human Brain: An Introduction to Its Functional Anatomy*, 3rd ed. Mosby, St. Louis: Mosby; 1993:399.)

Another major pathway within the limbic system centers on output from the amygdala (Fig. 1.2–44). Highly sensory information, primarily from the association regions of the prefrontal and temporal cortices, projects to the amygdala. Output from the amygdala is conducted through two main pathways (Fig. 1.2–40). A dorsal route, the *stria terminalis,* accompanies the caudate nucleus in an arch around the temporal lobe and contains axons that project primarily to the septal area and the hypothalamus. The second major output route, the *ventral amygdalofugal pathway,* passes below the lenticular nucleus and contains fibers that terminate in many regions, including the septal area, the hypothalamus,

and the medial dorsal thalamic nucleus. The medial dorsal nucleus projects heavily to prefrontal and some temporal cortical regions.

Both of these pathways reveal how the limbic system is able to integrate the highly processed sensory and cognitive information content of the cerebral cortical circuitry with the hypothalamic pathways that control autonomic and endocrine systems. In addition, the limbic system interacts with components of the basal ganglia system (Fig. 1.2–45). For example, the ventral amygdalofugal pathway also projects to the nucleus accumbens (ventral striatum), the area where the head of the caudate nucleus fuses with the putamen (Figs. 1.2–28 and 1.2–29). This region sends efferents to the ventral pallidum, an extension of the globus pallidus. This area projects to the medial dorsal thalamic nucleus. The pathway indicates that the functions of the basal ganglia extend beyond the regulation of motor activities and shows the necessity of considering the function or dysfunction of particular brain regions in the context of all aspects of their circuitry.

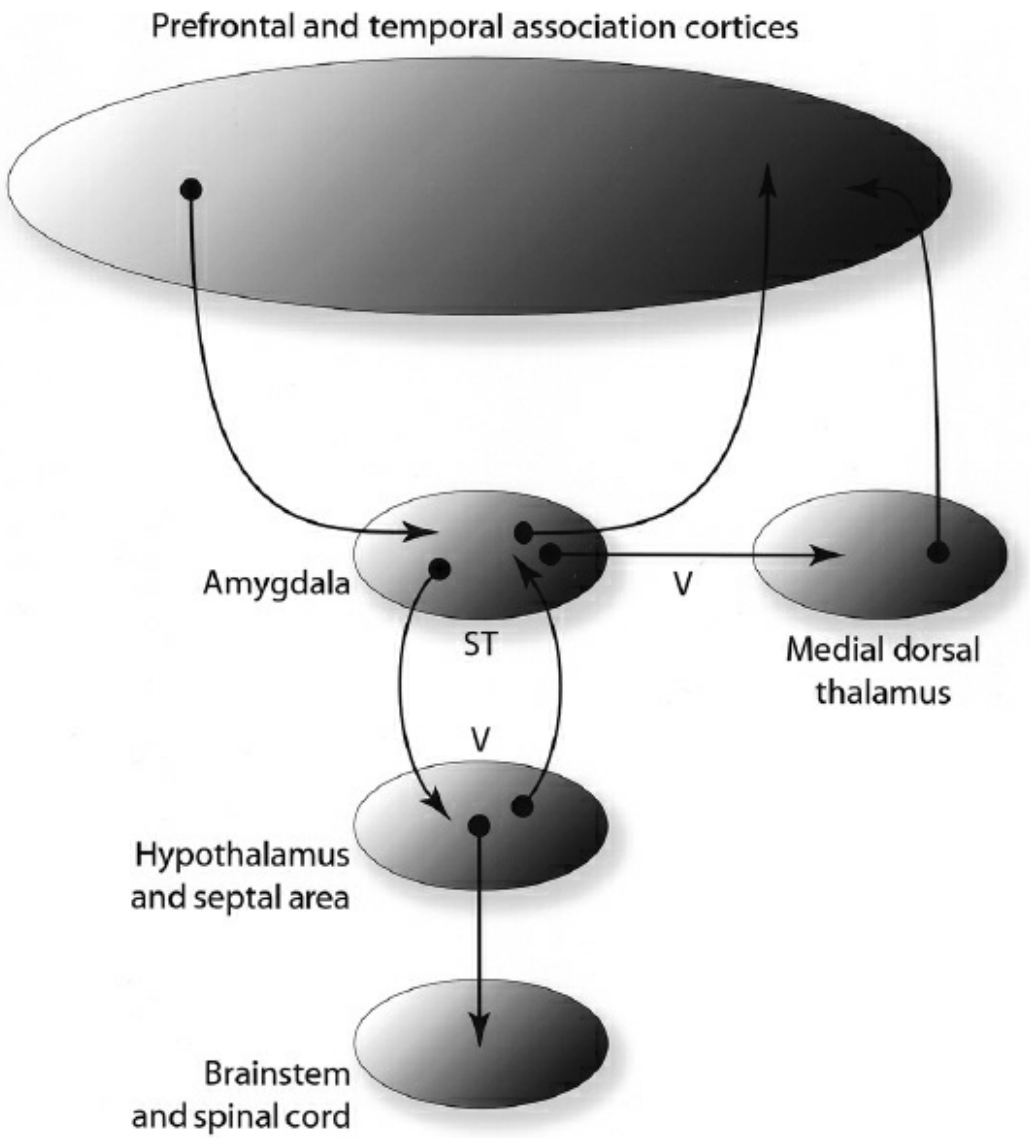

FIGURE 1.2–44. Functional neural circuitry of the limbic system. This diagram illustrates how the amygdala and the medial dorsal thalamus serve to integrate information processing between prefrontal and temporal association cortices and the hypothalamus. V, ventral amygdalofugal pathway; ST, stria terminalis. (Adapted from Nolte J. *The Human Brain: An Introduction to Its Functional Anatomy*, 3rd ed. Mosby, St. Louis: Mosby; 1993:399.)

IMPLICATIONS FOR BIOLOGICALLY BASED DIAGNOSTIC SYSTEMS

The integrity of the neuroanatomic features described in this chapter can be assessed in individuals with psychiatric disorders at different levels of resolution. Disease-associated changes in neuron number, neuron size, or connections among neurons may be reflected in gross structural alterations detected by in vivo imaging techniques. It remains unclear, however, if the resolution afforded by these imaging techniques would be able to discriminate among different disease processes in a manner that could inform clinical diagnosis. For example, a disease-related difference in the volume of a given brain region could be due to fewer neurons, smaller neurons, or fewer neuronal connections; the same abnormality evident by structural imaging could represent very different underlying disease processes. Although postmortem studies provide the level of resolution needed to distinguish among such possibilities, the diagnostic value of findings from these investigations (as is the case for imaging studies) requires the capacity to distinguish among the following four "C's": (1) *cause,* an upstream factor related to the disease pathogenesis; (2) *consequence,* a deleterious effect of a cause; (3) *compensation,* the brain's response to either a cause or a consequence that helps restore homeostasis; or (4) *confound,* a product of factors frequently associated with, but not a part of, the disease process, or an artifact of the approach used to obtain the measure of interest. The future incorporation of anatomic data into a diagnostic schema for psychiatric disorders will depend on the ability to make measurements at the appropriate level of resolution and to determine which "C" category a given observation represents.

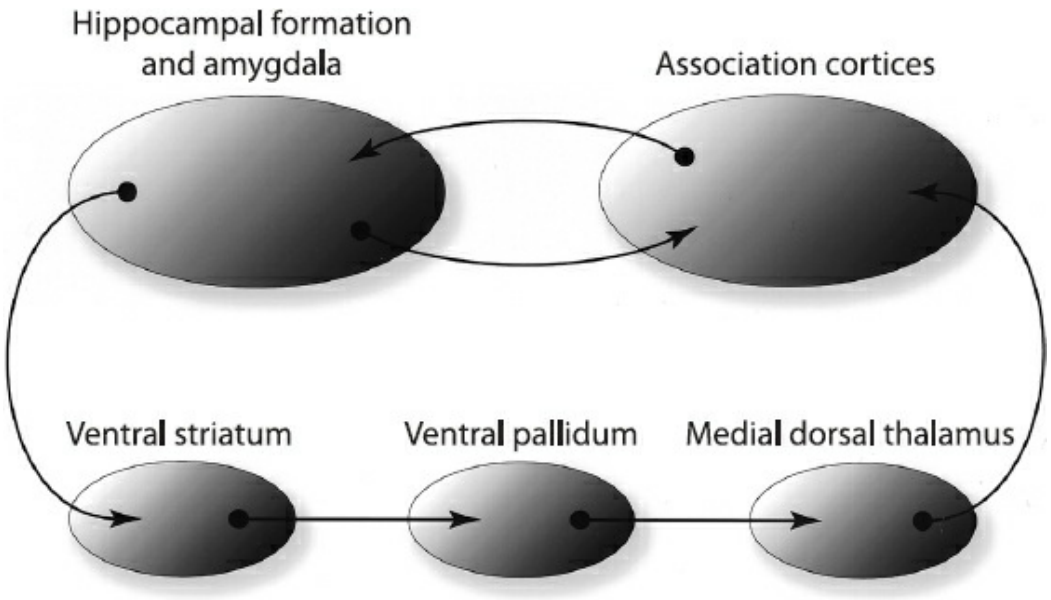

FIGURE 1.2–45. Functional neural circuitry of the limbic system. This drawing illustrates the interaction between the limbic system and certain components of the basal ganglia. (Adapted from Nolte J. *The Human Brain: An Introduction to Its Functional Anatomy*, 3rd ed. Mosby, St. Louis: Mosby; 1993:412.)

REFERENCES

Barbas H. General cortical and special prefrontal connections: principles from structure to function. *Ann Rev Neurosci.* 2015;38:269–289.

Beneyto M, Sibille E, Lewis DA. *Human postmortem brain research in mental illness syndromes. Neurobiology of Mental Illness,* 3rd ed. New York: Oxford University Press; 2009.

*Björklund A, Dunnett SB. Dopamine neuron systems in the brain: an update. *Trends Neurosci.* 2007;30:194–202.

Braak H, Del Tredici K. Cortico-basal ganglia-cortical circuitry in Parkinson's disease reconsidered. *Exp Neurol.* 2008;212(1):226–229.

Chen AC, Oathes DS, Chang C, et al. Causal interactions between fronto-parietal central executive and default-mode networks in humans. *PNAS.* 2013;110:19944–19949.

Chudasama Y, Robbins TW. Functions of frontostriatal systems in cognition: Comparative neuropsychopharmacological studies in rats, monkeys and humans. *Biol Psychol.* 2006;73:19–38.

Crowell AL, Garlow SJ, Riva-Posse P, Mayberg HS. Characterizing the therapeutic response to deep brain stimulation for treatment-resistant depression: a single center long-term perspective. *Front Integrative Neurosci.* 2015;9:41.

DeFelipe J. The evolution of the brain, the human nature of cortical circuits, and intellectual creativity. *Front Neuroanat.* 2011;5:29.

DeLong MR, Wichmann T. Circuits and circuit disorders of the basal ganglia. *Arch Neurol.* 2007;64:20–24.

Eggan SN, Lewis DA. Immunocytochemical distribution of the cannabinoid CB1 receptor in the primate neocortex: a regional and laminar analysis. *Cereb Cortex.* 2007;17:175–191.

Field RD. White matter in learning, cognition and psychiatric disorders. *Trends Neurosci.*

2008;31(7):361–370.

*Fuster JM.*The Prefrontal Cortex*, 5th ed. San Diego, CA: Academic Press; 2015.

Glausier JR, Lewis DA. Dendritic spine pathology in schizophrenia. *Neurosci.* 2013;251:90.

Gonzalez-Burgos G, Cho RY, Lewis DA. Alterations in Cortical Network Oscillations and Parvalbumin Neurons in Schizophrenia. *Biol Psychiatry*. 2015;77:1031–1040.

Haines DE. *Fundamental Neuroscience for Basic and Clinical Applications,* 4th ed. Philadelphia, PA: Churchill Livingstone; 2006.

*Lewis DA, Curley AA, Glausier JR, Volk DW. Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. *Trends Neurosci.* 2012;35;57–67

Lewis DA, Melchitzky DS, Gonzalez-Burgos G. Specificity in the functional architecture of primate prefrontal cortex. *J Neurocytol.* 2002;31:265–276.

*Nolte J. *The Human Brain: an Introduction to Its Functional Anatomy,* 6th ed. St. Louis, MO: Mosby; 2008.

*Oberheim NA, Wang X, Goldman S, Nedergaard M. Astrocytic complexity distinguishes the human brain. *Trends Neurosci.* 2006;29:547–553.

Patestas MA, Gartner LP. *A Textbook of Neuroanatomy*. Malden, MA: Blackwell; 2006.

Purves D, Augustine GJ, Fitzpatrick D, Hall WC, LaMantia A-S. *Neuroscience,* 5th ed. Sunderland, MA: Sinaver Associates, Inc.; 2012.

Ramnani N. The primate cortico-cerebellar system: anatomy and function. *Nat Rev Neurosci.* 2006;7:511–522.

*Sherman SM. Thalamus plays a central role in ongoing cortical functioning. *Nat Neurosci.* 2016;19:533–541.

Sillito AM, Cudeiro J, Jones HE. Always returning: Feedback and sensory processing in visual cortex and thalamus. *Trends Neurosci.* 2006;29:307–316.

Squire LR, Bloom FE, McConnell SK, Roberts JL, Spitzer NC. *Fundamental Neuroscience,* 4th ed. San Diego, CA: Academic Press; 2012.

Steriade M. Grouping of brain rhythms in corticothalamic systems. *Neuroscience.* 2007;137:1087–1106.

The Petilla Interneuron Nomenclature Group (PING). New insights into the class and nomenclature of cortical GABAergic interneurons. *Nat Rev Neurosci.* 2013;14:202–216.

Tsujino N, Sakuri T. Orexin/Hypocretin: A neuropeptide at the interface of sleep, energy homeostasis, and reward system. *Pharmacol. Rev.* 2009;61:162–176.

Wolff M, Alcaraz F, Marchand AR, Coutureau E. Functional heterogeneity of the limbic thalamus: From hippocampal to cortical functions. *Neurosci Biobehav Rev.* 2015;54:120–130.

▲ 1.3 Neural Development and Neurogenesis

EMANUEL DICICCO-BLOOM, M.D., AND MARYANN OBIORAH, PH.D.

NEURAL DEVELOPMENT AND NEUROGENESIS

The human brain is a structurally and functionally complex system that exhibits ongoing modification in response to both experience and disease. The anatomical and neurochemical systems that underlie the cognitive, social, emotional, and sensorimotor functions of the mature nervous system emerge from neuronal and glial cell populations that arise during the earliest periods of development. Indeed, the nervous system starts forming immediately after the primitive gut (archenteron) invaginates the embryonic ball of cells known as the blastula. This chapter describes the molecular and genetic mechanisms that regulate the generation and movements of cells required to elaborate region-specific populations whose interconnections form functional networks. General developmental principles are highlighted, and the current understanding of the roles of adult neurogenesis, environmental toxicants, and drugs of abuse in brain function and possibly as factors contributing to neuropsychiatric disorders are described.

An understanding of molecular and cellular mechanisms mediating nervous system development is critical in psychiatry it is now we now known that abnormalities of developmental processes contribute to many brain disorders. While a developmental basis may not be surprising in early childhood disorders, such as autism, fragile X mental retardation, and Rett syndrome, even mature diseases including schizophrenia and depression reflect ontogenetic factors. For example, there are reductions in forebrain region volumes, neuron and glial cell numbers, and some classes of interneurons in schizophrenia that are apparent at the time of diagnosis. Similarly, in autism, early brain growth is abnormally increased and abnormalities of cellular organization are observed that reflect disturbances in developmental cell proliferation and migration. When there is abnormal regulation of early brain development, we lay down a foundation of altered neuron populations that may differ in cell types, numbers, and positions or may elaborate abnormal connections, with consequences for interacting glial populations. With progressive postnatal development, the maturing brain systems call upon component neurons to achieve increasing levels of complex information processing, which may be deficient should initial conditions be disturbed. New neural properties emerge during maturation as neuron populations elaborate additional functional networks based upon and modified by ongoing experience. Given the brain's dynamic character, it may be expected that

developmental abnormalities in neural populations and systems, caused by genetic as well as environmental factors, will manifest at diverse times in a person's life.

OVERVIEW OF NERVOUS SYSTEM MORPHOLOGICAL DEVELOPMENT

In considering brain development, several overarching principles may serve to guide psychiatrists' understanding. First, different brain regions and neuron populations are generated at distinct times of development, and exhibit specific temporal schedules. This has implications for the consequences of specific developmental insults, such as the production of autism following fetal exposure to the drug thalidomide only during days 20 to 24 of gestation. Second, the sequence of cellular processes comprising ontogeny predicts that abnormalities in early events necessarily lead to differences in subsequent stages, though not all abnormalities may be accessible to psychiatrists' clinical tools. For example, a deficit in the number of neurons will likely lead to reductions in axonal processes and ensheathing white matter in the mature brain. However, at the clinical level, since glial cells outnumber neurons eight to one, psychiatrists may only appreciate changes in the majority glial cell population, the oligodendrocytes and their myelin, which will appear as altered white matter on neuroimaging with little evidence of a neuronal disturbance. Third, it is clear that specific molecular signals, such as extracellular growth factors and cognate receptors or transcription factors, play roles at multiple developmental stages of the cell. For example, both insulin-like growth factor I (IGF-I) and brain-derived neurotrophic factor (BDNF) regulate multiple cellular processes during the developmental generation and mature function of neurons, including cell proliferation, survival promotion, neuron migration, process outgrowth, and the momentary synaptic modifications (plasticity) underlying learning and memory. Thus changes in expression or regulation of a ligand or its receptor, by experience, environmental insults, or genetic mechanisms, will have effects on multiple developmental and mature processes. Finally, the evolution of the newer mammalian neocortex, where higher mental functioning occurs, highlights the emergence of cortical folding to accommodate larger cortical surface area. As a result, higher functioning mammals possess a higher capacity for neurogenesis by way of an additional proliferative zone. This chapter considers these principles as it examines cellular and molecular systems regulating development, and discusses implications for psychiatric disease.

The Neural Plate and Neurulation

The nervous system of the human embryo first appears between 2 and 1/2 and 4 weeks of gestation. During development, emergence of new cell

types, including neurons, results from interactions between neighboring layers of cells. On gestational day 13, the embryo consists of a sheet of cells. Following gastrulation (days 14 and 15), which forms a two-cell layered embryo consisting of ectoderm and endoderm, the neural plate region of the ectoderm is delineated by the underlying mesoderm, which appears on day 16. The mesoderm forms by cells entering a midline cleft in the ectoderm called the primitive streak. After migration, the mesodermal layer lies between ectoderm and endoderm and induces overlying ectoderm to become neural plate. Induction usually involves release of soluble growth factors from one group of cells, which in turn bind receptors on neighboring cells, eliciting changes in nuclear transcription factors which control downstream gene expression. In some cases, cell–cell contact mediated mechanisms are involved. In the gene patterning section below, the important roles of soluble growth factors and transcription factor expression will be described.

The neural plate, whose induction is complete by 18 days, is a sheet of columnar epithelium, and is surrounded by ectodermal epithelium. After formation, the edges of the neural plate elevate, forming the neural ridges. Subsequently, changes in intracellular cytoskeleton and cell–extracellular matrix attachment cause the ridges to merge in the midline and fuse, a process termed neurulation, forming the neural tube, with a central cavity presaging the ventricular system (Fig. 1.3–1). Fusion begins in the cervical region at the hindbrain level (medulla and pons) and continues rostrally and caudally. Neurulation occurs at 3 to 4 weeks of gestation in humans, and its failure results in anencephaly rostrally and spina bifida caudally. Neurulation defects are well known following exposure to retinoic acid in dermatological preparations and anticonvulsants, especially valproic acid (VPA), as well as diets deficient in folic acid.

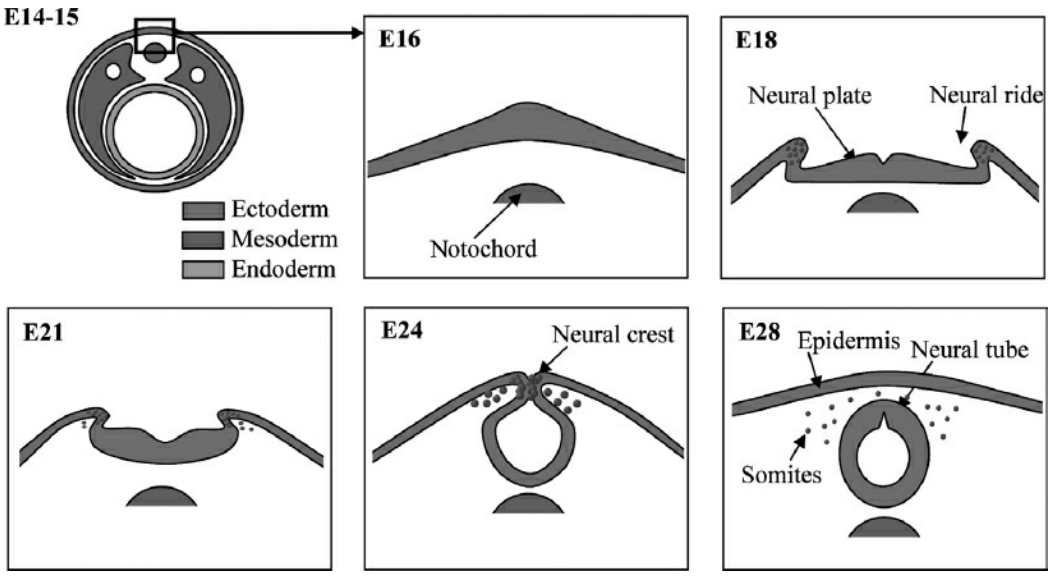

FIGURE 1.3–1. Mechanisms of neurulation. Neurulation begins with the formation of a

neural plate in response to soluble growth factors released by the underlying notochord. The neural plate originates as a thickening of the ectoderm that results from cuboidal epithelial cells becoming columnar in shape. With further changes in cell shape and adhesion, the edges of the plate fold and rise, meeting in the midline to form a tube. Cells at the tips of the neural folds come to lie between the neural tube and overlying epidermis, forming the neural crest that gives rise to the peripheral nervous system and other structures.

Another product of neurulation is the neural crest, whose cells derive from the edges of the neural plate and dorsal neural tube. From this position, neural crest cells migrate dorsolaterally, under the skin to form melanocytes, and ventromedially to form dorsal root sensory ganglia and sympathetic chains of the PNS and ganglia of the enteric nervous system. However, neural crest gives rise to diverse tissues including cells of neuroendocrine, cardiac, mesenchymal, and skeletal systems, forming the basis of many congenital syndromes involving brain and other organs. The neural crest origin at the border of neural and epidermal ectoderm and its generation of melanocytes forms the basis of the neurocutaneous disorders, including tuberous sclerosis and neurofibromatosis. Finally, another non-neuronal structure of mesodermal origin formed during neurulation is the notochord found on the ventral side of the neural tube. As shown below, the notochord plays a critical role during neural tube differentiation, since it is a signaling source of soluble growth factors, such as sonic hedgehog (Shh), which impact gene patterning and cell determination.

Regional Differentiation of the Embryonic Nervous System

After closure, the neural tube expands differentially to form major morphological subdivisions that precede the major functional divisions of the brain. These subdivisions are important developmentally since different regions are generated according to specific schedules of proliferation and subsequent migration and differentiation. The neural tube can be described in three dimensions, including longitudinal, circumferential, and radial. The longitudinal dimension reflects the rostrocaudal (anterior–posterior) organization, which most simply consists of brain and spinal cord. Organization in the circumferential dimension, tangential to the surface, represents two major axes: in the dorsoventral axis, cell groups are uniquely positioned from top to bottom. On the other hand, in the medial to lateral axis, there is mirror image symmetry, consistent with right–left symmetry of the body. Finally, the radial dimension represents organization from the inner most cell layer adjacent to the ventricles to the outer most surface, and exhibits region-specific cell layering. At 4 weeks, the human brain is divided longitudinally into the prosencephalon (forebrain), mesencephalon (midbrain), and rhombencephalon (hindbrain). These three subdivisions or “vesicles” divide further into five divisions by 5 weeks, consisting of the

prosencephalon which forms the telencephalon (including cortex, hippocampus, and basal ganglia) and diencephalon (thalamus and hypothalamus), the mesencephalon, (midbrain), and the rhombencephalon yielding metencephalon (pons and cerebellum) and myelencephalon (medulla). Morphological transformation into five vesicles depends on region-specific proliferation of precursor cells adjacent to the ventricles, the so-called ventricular zones (VZs). As discussed below, proliferation intimately depends on soluble growth factors made by proliferating cells themselves or released from regional signaling centers. In turn, growth factor production and cognate receptor expression also depend on region-specific patterning genes. It is now known that VZ precursors, which appear morphologically homogeneous, express a checkerboard array of molecular genetic determinants that control the generation of specific types of neurons in each domain (Fig. 1.3–2).

In the circumferential dimension, organization begins very early and extends over many rostrocaudal subdivisions. In spinal cord, the majority of tissue comprises the lateral plates, which later divide into dorsal or alar plates, composed of sensory interneurons, and motor or basal plates, consisting of ventral motor neurons. Two other diminutive plates, termed the roof plate and floor plate, which are virtually absent in maturity, however, play critical regulatory roles as growth factor signaling centers in the embryo. Indeed, the floor plate, in response to Shh from the ventrally located notochord, produces its own Shh, which in turn induces neighboring cells in ventral spinal cord and brainstem to express region-specific transcription factors that specify cell phenotype and function. For example, in combination with other factors, floor plate Shh induces midbrain precursors to differentiate into dopamine secreting neurons of the substantia nigra. Similarly, the roof plate secretes growth factors, such as BMPs, which induce dorsal sensory neuron cell fate in spinal cord. In the absence of roof plate, dorsal structures fail to form, such as cerebellum, and midline hippocampal structures are missing. Finally, in the radial dimension, the organization of layers is subdivision-specific, produced by differential proliferation of VZ precursors and cell migration, described below.

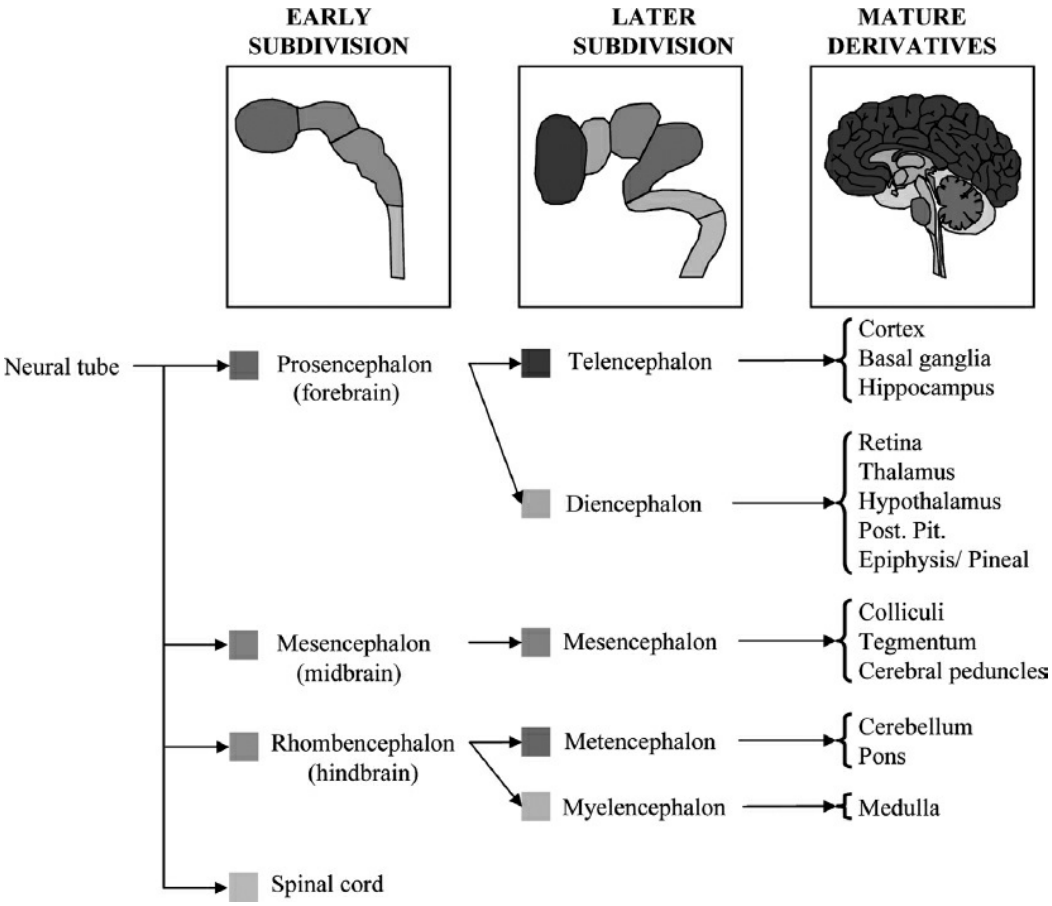

FIGURE 1.3–2. Progression of brain regional differentiation. Early after neurulation, the neural tube differentiates into four regions (forebrain, midbrain, hindbrain, and spinal cord) that gives rise following later divisions and maturation to the different brain structures.

The Ventricular and Subventricular Proliferative Zones

The distinct patterns of precursor proliferation and migration in different regions generate the radial organization of the nervous system. In each longitudinal subdivision, the final population size of a brain region depends on the interplay of regulated neurogenesis with programmed cell death (see below). Traditional concepts had suggested that there was excess cell production everywhere, and that final cell number regulation was achieved primarily after neurogenesis through selective cell death mediated by target-derived survival (trophic) factors. It is now known that the patterning genes discussed below play major roles in directing regional precursor proliferation that is coordinated with final structural requirements, and that programmed cell death occurs at multiple stages. Consequently in diseases characterized by brain regions smaller than normal, such as schizophrenia, there may be a failure to generate neurons initially, as opposed to normal generation with subsequent cell loss.

The generation of specific cell types involves proliferation of undifferentiated precursor cells (or progenitors), followed by cessation of proliferation (exit from the cell cycle) and expression of specific phenotypic

characters, such as neurofilaments and neurotransmitter systems. Precursor proliferation occurs primarily in two densely packed regions during development. The primary site is the VZ lining the walls of the entire ventricular system, which contributes to all brain regions in the rostrocaudal dimension. For select regions, however, including the cerebral cortex, hippocampus, and cerebellar cortex, precursors from the VZ migrate out to secondary zones where they generate a more restricted range of cell types.

In the early embryo, neural tube VZ progenitors are arranged as a one-cell layer thick, pseudostratified neuroepithelium. The bipolar VZ precursors have cytoplasmic processes that span from the ventricular to the pial surface. During the cell cycle, the VZ appears multi-layered, or stratified, because cell nuclei undergo movements, called interkinetic nuclear migration. New cells are produced through the cell cycle which comprises four stages, including mitosis (M), when nuclei and cells divide, G1 when cells grow in size before dividing again, S-phase, when cells synthesize DNA and replicate chromosomes, and a brief G2 period followed by M phase. Precursor cell division (M phase) occurs at the ventricular margin, producing two new cells (Fig. 1.3–3). The progeny then reenter G1 as they move outward toward the pia. Under the influence of extracellular signals these cells become committed to another round of division, marked by entry into S-phase, which occurs near the upper VZ margin. After replicating DNA, nuclei move back down during G2 to the ventricular surface where they undergo mitosis and divide. The role of nuclear migration is not known, though it may allow nuclei access to environmental cues produced by postmitotic cells that effect subsequent proliferation and gene expression. Several human genetic mutations interfere with interkinetic nuclear movement and cell migration, producing heterotopic neurons and epilepsy syndromes (see below).

At the earliest stages, VZ cells divide to increase the pool of progenitors before producing postmitotic neurons. Then, during the prolonged period of neurogenesis, with each cell cycle on average, a cell divides giving rise to both a postmitotic neuron and another dividing precursor. At the end of neurogenesis, precursor division gives rise to two postmitotic neurons only, greatly increasing neuron production and depleting the precursor pool. The newly born neurons do not remain in the VZ, but instead migrate out to their final destinations, such as the cerebral cortical plate, traveling along the processes of radial glial cells (RGCs) (Fig. 1.3–4A). These bipolar VZ precursors described above are called radial glia; they have one process associated with the ventricular surface and the other contacts the pial surface (see below). The association between the newborn neurons and the radial glial process allows cells generated within localized VZ domains, known to express distinct patterning genes (see below), to migrate to specific cortical functional areas, such as visual or motor cortex. This

suggests that VZ precursors already have their phenotypic fate specified at the genetic level prior to ceasing cell division and beginning migration. However, although there is clearly a role for early expressed VZ genes in maturing neurons, the ingrowing thalamic afferents also impact cortical neuronal fate and function. While in rodents neurons are generated prior to birth and glia are produced after, in the human brain, neuron production generally occurs for the first 4 months of gestation, whereas from then on until birth, neurons undergo migration, whereas glial precursors proliferate, migrate, and produce myelin.

In addition to this general plan of neurogenesis in the VZ, secondary proliferative zones produce specific neuron populations in particular regions. For example, in cerebral cortex and thalamus, the subventricular zone (SVZ) produces intermediate precursor cells, that can generate oligodendrocytes, diverse astrocytes, and neurons. In hippocampus, the hilus and later the subgranular zone (SGZ) produce dentate gyrus (DG) granule neurons, a lifelong process of adult neurogenesis (see below). Finally, in newborn cerebellum, the overlying external germinal layer (EGL) generates granule neurons for several weeks in rodents and for 7 to 20 months in humans, a population likely affected by medical treatments administered in the neonatal intensive care unit (ICU). In contrast to the VZ, secondary zone cells do not exhibit nuclear movements, suggesting distinct mechanisms of regulation. After neurogenesis is complete, the VZ differentiates into ciliated epithelial cells of the ependymal lining. Underlying the ependyma, undifferentiated cells of the SVZ, referred to as subependyma, have been identified as a neural stem cell (NSC) population, capable of proliferating and generating neurons and glia throughout life. In humans, the final destinations of these newly generated neuroblasts are unknown, but recent evidence suggests that lateral ventricle-derived neuroblasts migrate to the nearby striatum during adulthood to differentiate into interneurons. Furthermore, their depletion is observed during the advanced stages of Huntington disease.

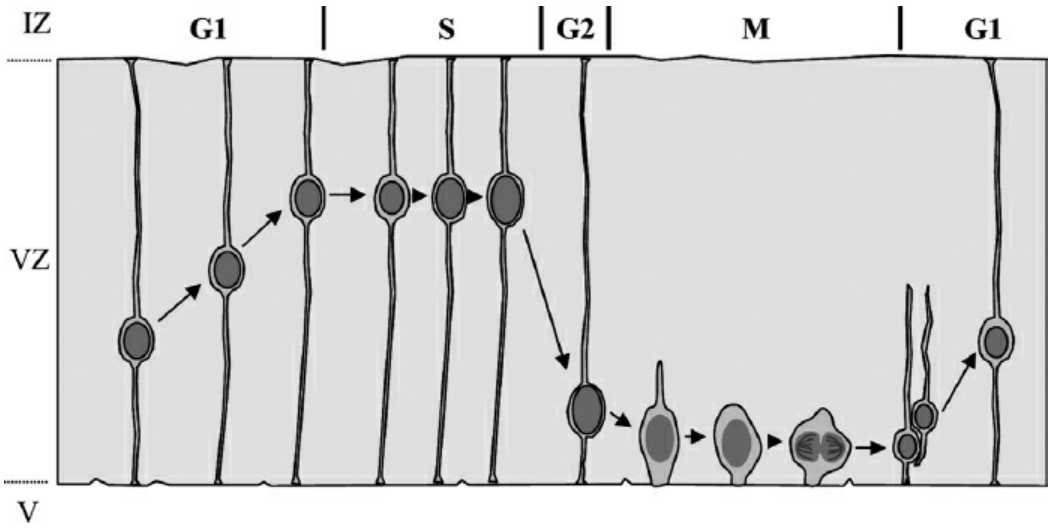

FIGURE 1.3–3. Interkinetic nuclear migration in the ventricular zone. During each cell cycle, nuclei move from the ventricular surface at G1 to the border of the ventricular zone where it enters S phase. Nuclei move down during G2, and reach the ventricular surface where they undergo mitosis. Asymmetric division leads to generation of a postmitotic cell that leaves the ventricular zone to produce a neocortical neuron, while the remaining stem cell continues to divide. IZ, intermediate zone; VZ, ventricular zone; V, ventricle. (Modified from Jacobson M. *Developmental Neurobiology*, 3rd ed. New York: Plenum Press; 1991.)

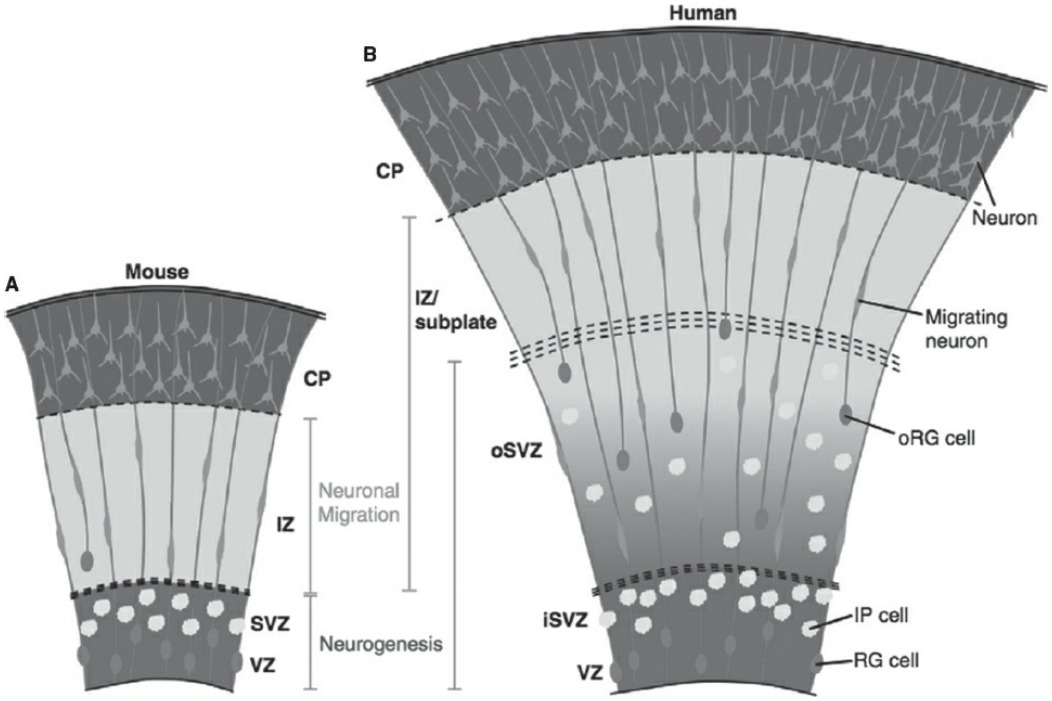

FIGURE 1.3–4. Comparison of the developing mouse and human cortex. **A:** Model of mouse neocortical development. The VZ and SVZ are the primary zones of neurogenesis (*dark gray*) where RG cells and IP cells reside, respectively. Neurons born in the VZ or SVZ migrate along the RG radial fiber scaffold found in the intermediate zone (IZ), the primary zone of neuronal migration (*light gray*) in the mouse, to reach the CP. oRG cells are infrequent and not located in a distinct progenitor zone. **B:** Model of human neocortical development. RG cells, IP cells, and oRG cells are found in neurogenic zones (*dark gray*), which are the VZ, the inner SVZ (iSVZ), and the outer SVZ (oSVZ). Neurons migrate through the OSVZ and IZ/subplate (zones of neuronal migration, *light gray*) to populate the CP. Neurons must navigate a larger distance than in the mouse, and a radial fiber scaffold of increased complexity, to reach the CP. (From LaMonica BE, Lui JH, Wang X, Kriegstein AR. OSVZ progenitors in the human cortex: an updated perspective on neurodevelopmental disease. *Curr Opin Neurobiol*. 2012;5:747.)

Similar to the SVZ, humans and NHPs have an additional proliferative region in the cortex called the outer subventricular zone (OSVZ), whose extensive proliferation accounts for the massive evolutionary increase in the cerebral cortex. This region first appears during the beginning of the second trimester in humans. By mid-trimester it becomes the main site of neurogenesis, surpassing the neurogenic capacity of the SVZ by almost fourfold. Recent studies have characterized the cell types and neurogenic contributions of this region. The OSVZ contains radial glia-like (oRG) cells that have only one long basal process that contacts the pia, unlike radial

glia (RG) of the VZ, which have both apical (touching the ventricle) and basal processes (Fig. 1.3–4B). The origin of these cells is unknown, but one hypothesis is that they originated in the VZ and migrated to the OSVZ via somal translocation. oRG cells have self-renewing capacity and undergo mitosis during somal translocation in the OSVZ, but do not translocate to the ventricle. They give rise to intermediate progenitor daughter cells that also have proliferative capacity and after migrating up the basal process, differentiate into excitatory neurons. The OSVZ might have an evolutionary role in the formation of gyrencephalic (highly folded) brain development, where cortical folding allows the much larger cortical surface area found in higher functioning vertebrates to fit into a limited cranial space. Recent evidence shows that the fan-like orientation of oRG processes favor the formation of gyri, thereby increasing cortical surface area. These areas also tend to have higher rates of intermediate progenitor proliferation. Therefore, the OSVZ appears to have evolved in higher functioning organisms to increase neurogenic capacity in the cortex, allowing for more complex neuronal functioning.

Radial and Tangential Patterns of Neurogenesis and Migration

There are three well-recognized spatio-temporal patterns of neurogenesis which underlie regional brain formation. The following examples illustrate common principles concerning relationships of cell cycle exit (cell birthday) to final cell position, the roles of radial glia in migration, and the distinct capacities of secondary proliferative zones. There are two radial patterns of cell migration from the VZ, referred to as inside-to-outside and outside-to-inside. The third involves nonradial or tangential migration of cells, some of which originate in secondary proliferative zones. Experimentally, these patterns are defined in animals by marking mitotic cells using nuclear incorporation of labeled DNA precursors, either tritiated (^{3}H)-thymidine or bromodeoxyuridine (BrdU), to identify the last day a precursor is in S phase (its birthday), after which it exits the cell cycle, differentiates, and migrates to its final position.

The two radial patterns of neurogenesis reflect whether a structure is phylogenetically older, such as spinal cord, tectum, and hippocampal dentate gyrus, or more recently evolved, like cerebral cortex. In older structures, early generated cells are positioned on the outside, with later born cells residing inside, closer to the VZ. In the second pattern relevant to cerebral cortex, early born cells are located on the inside, with later born cells migrating past earlier ones to take up position outside. This inside-to-outside gradient requires a more complex mechanism of cell localization. Finally, the specific character of a region may be altered by nonradial inward migration of cells generated in other locations, relevant to GABA interneurons in cortex and hippocampus, or granule neurons in cerebellum, hippocampal dentate gyrus, and olfactory bulb.

Of interest to psychiatry, the cerebral cortex is the paradigmatic model of inside-to-outside neurogenesis. Derived from the embryonic forebrain telencephalic vesicles, the characteristic six-cell layers represent a common cytoarchitectural and physiological basis for neocortical function. Within each layer, neurons exhibit related axodendritic morphologies, use common neurotransmitters, and establish similar afferent and efferent connections. In general, pyramidal neurons in layer III establish synapses within and between cortical hemispheres whereas deeper layer V/VI neurons project primarily to subcortical nuclei, including thalamus, brainstem, and spinal cord. The majority of cortical neurons originate from the forebrain VZ. At earliest stages, the first postmitotic cells migrate outward from the VZ to establish a superficial layer termed the preplate. Two important cell types comprise the preplate, Cajal–Retzius cells that form outermost layer I or marginal zone, and subplate neurons, which lay beneath future layer VI. These distinct regions form when later born cortical plate neurons migrate within and divide the preplate in two (Fig. 1.3–5).

After preplate formation, the cortical VZ generates in inside-to-outside fashion first layer V/VI neurons, and then more superficial layers in temporal sequence. Thus the day on which a precursor exits the cell cycle in the VZ, its birthday, essentially predicts the kind and localization of the neuron generated. Molecular mechanisms mediating this correlation are currently being defined (see below), including specific stimulatory and inhibitory proliferative signals and genetic determinants. Notably, the cortical VZ is the primary source of excitatory pyramidal neurons that secrete glutamate.

The embryonic preplate harbors Cajal–Retzius cells that produce the extracellular glycoprotein reelin, an important signal for neuronal migration. When reelin gene is genetically deleted in mice, cortical neuron migration is inverted. That is, the usual inside-to-outside gradient of cell generation and laminar position becomes inverted, yielding an outside-to-inside pattern. Thus, early born neurons appear furthermost from the VZ, and latest born cells remain closest to the ventricles. Abnormal levels of reelin protein and mRNA have been observed in several diseases, including major depression, bipolar disorder, schizophrenia, autism, and most recently Alzheimer disease. Human reelin mutation is associated with lissencephaly (smooth brain), a gyral patterning malformation with loss of gyri and sulci, and abnormalities in cerebellum (see below). On the other hand, the subplate neurons, which persist only until early postnatal development in rodents, play a critical role as temporary targets for thalamic axon terminals on their way to cortex. After pyramidal neurons settle into correct layers in cortical plate, thalamic processes migrate further to reach layer IV targets, and subplate neurons undergo programmed cell death.

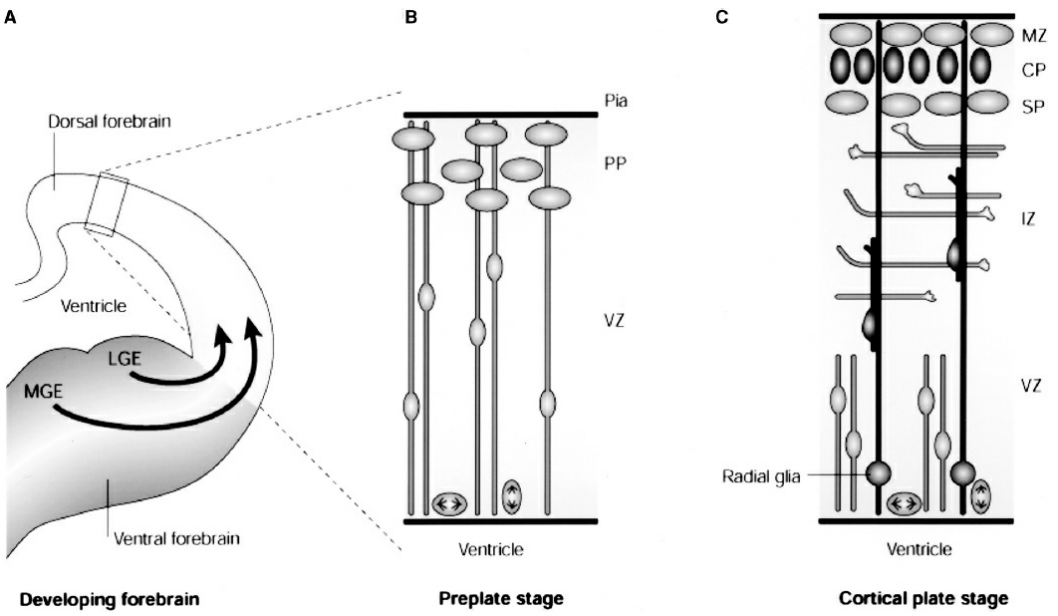

FIGURE 1.3–5. Schematic drawing of radial and tangential migration during cerebral cortex development. **A:** A coronal section of one-half of the developing rat forebrain. The dorsal forebrain gives rise to the cerebral cortex. Medial ganglionic eminences (MGEs) and lateral ganglionic eminences (LGEs) of the ventral forebrain generate neurons of the basal ganglia and the cortical interneurons. The arrows indicate the tangential migration route for γ-aminobutyric acid interneruons to the cortex. The boxed area (enlarged in **B** and **C**) shows the developing cortex at early and late stages. **B:** In the dorsal forebrain, the first cohort of postmitotic neurons migrate out from the ventricular zone (VZ) and create a preplate (PP) below the pial surface. **C:** Subsequent postmitotic neurons will migrate along radial glia through the intermediate zone (IZ) and take position in the middle of the preplate, creating a cortical plate (CP) between the outer marginal zone (MZ) and inner subplate (SP). Ultimately, the CP will be composed of six layers that are born sequentially, migrating in an inside-to-outside pattern. Horizontal processes in the IZ represent axon terminals of thalamic afferents. (From Nadarajah B, Parnavelas JG. Modes of neuronal migration in the developing cerebral cortex. *Nat Neurosci*. 2002;3:423, with permission.)

The discovery that the neocortex is also populated by neurons generated in the ventral forebrain has changed the view of the origins of cortical neuron populations involved in human brain disease (Fig. 1.3–5). Molecular studies of patterning genes, especially *Dlx,* strongly support this model (see below). In contrast to excitatory pyramidal neurons, the inhibitory GABA secreting interneurons originate from mitotic precursors of the ganglionic eminences that generate the neurons of the basal ganglia. Subsets of interneurons also secrete neuropeptides, such as NPY and somatostatin, and express NO generating enzyme, NOS. Not associated with cortical VZ radial glia, these GABA interneurons reach the cortical plate by migrating tangentially, in either the superficial marginal zone or in a deep position above the VZ, the subplate region where thalamic afferents are also growing. Significantly, in brains from schizophrenic patients, the prefrontal cortex exhibits a reduced density of interneurons in layer II. In addition, there is upregulation of GABAa receptor binding, a potential functional compensation, as well as a relative deficiency of NOS expressing neurons. These observations have led to the hypothesis that schizophrenia

is due to reduced GABAergic activity, though recent models suggest this defect may reflect abnormalities in excitatory glutamate neuron input. The origin of GABA interneurons from the ganglionic eminences, and their association with specific patterning genes (*Dlx,* see below), raises new genetic models of disease causation, and possible strategies for disease intervention. Thus, more broadly, normal cortical development depends on a balance of two principal patterns of neurogenesis and migration, consisting of radial migration of excitatory neurons from the dorsal forebrain VZ and tangential migration of inhibitory neurons from the ventral forebrain.

In contrast to inside-to-outside neurogenesis observed in cortex, phylogenetically older hypothalamus, spinal cord, and hippocampal DG exhibit the reverse order of cell generation. First formed postmitotic neurons lie superficially and last generated cells localize toward the center. While this outside-to-inside pattern might reflect passive cell displacement, radial glia and specific migration signaling molecules are clearly involved. Furthermore, cells do not always lie in direct extension from their locus of VZ generation. Rather, some groups of cells migrate to specific locations, as observed for neurons of the inferior olivary nuclei.

Of prime importance in psychiatry, the hippocampus demonstrates both radial and nonradial patterns of neurogenesis and migration. The pyramidal cell layer, Ammon Horn CA 1 to 3 neurons, is generated in a typical outside-to-inside fashion in the dorsomedial forebrain for a discrete period, from 7 to 15 weeks of gestation, and exhibits complex migration patterns. In contrast, the other major population, DG granule neurons, starts appearing at 18 weeks, and exhibits prolonged postnatal neurogenesis, originating from several migrating secondary proliferative zones. In rat, for instance, granule neurogenesis starts at E16 with proliferation in the forebrain VZ. At E18, an aggregate of precursors migrates along a subpial route into the DG itself where they generate granule neurons in situ. After birth, there is another migration, localizing proliferative precursors to the dentate hilus, which persists until 1 month of life. Thereafter, granule precursors move to a layer just under the dentate gyrus, termed the SGZ, which produces neurons throughout life in adult rats, primates, and humans. In rodents, SGZ precursors proliferate in response to cerebral ischemia, tissue injury and seizures, as well as growth factors (see below). Finally, the diminished hippocampal volumes reported in schizophrenia and depression raise the possibility that disordered neurogenesis plays a role in pathogenesis, either as a basis for dysfunction, or a consequence of brain injuries, consistent with associations of gestational infections with disease manifestation.

Finally, a different combination of radial and nonradial migration is observed in cerebellum, a region that contributes to both motor and nonmotor circuits, with particular significance for ASDs. Except for granule

cells, the other major neurons, including Purkinje and deep nuclei, originate from the primary VZ of the fourth ventricle, coincident with other brainstem neurons. In rat, this occurs at E13–E15 and in humans, 5 to 7 weeks of gestation. The granule neurons, as well as basket and stellate interneurons, originate in the secondary proliferative zone, the EGL, which covers newborn cerebellum at birth. EGL precursors originate in the fourth ventricle VZ, and migrate dorsally through the brainstem to reach this superficial position. The rat EGL proliferates for 3 weeks, generating more neurons than in any other structure, while in humans, EGL precursors exist for at least 7 weeks and up to 2 years. When an EGL precursor stops proliferating, the cell body sinks below the surface, grows bilateral processes that extend transversely in the molecular layer, and then the soma migrates further down into the internal granule layer (IGL). Cells reach the IGL along specialized Bergmann glia, which serve guidance functions similar to the radial glia. However, in this case, cells originate from a secondary proliferative zone that generates neurons exclusively of the granule cell lineage, indicating a restricted neural fate. Clinically, this postnatal population in infants makes cerebellar granule neurogenesis vulnerable to infectious insults of early childhood, and an undesirable target of several therapeutic drugs, such as steroids, well known to inhibit cell proliferation. In addition, proliferative control of this stem cell population is lost in the common childhood brain tumor, medulloblastoma (Fig. 1.3–6).

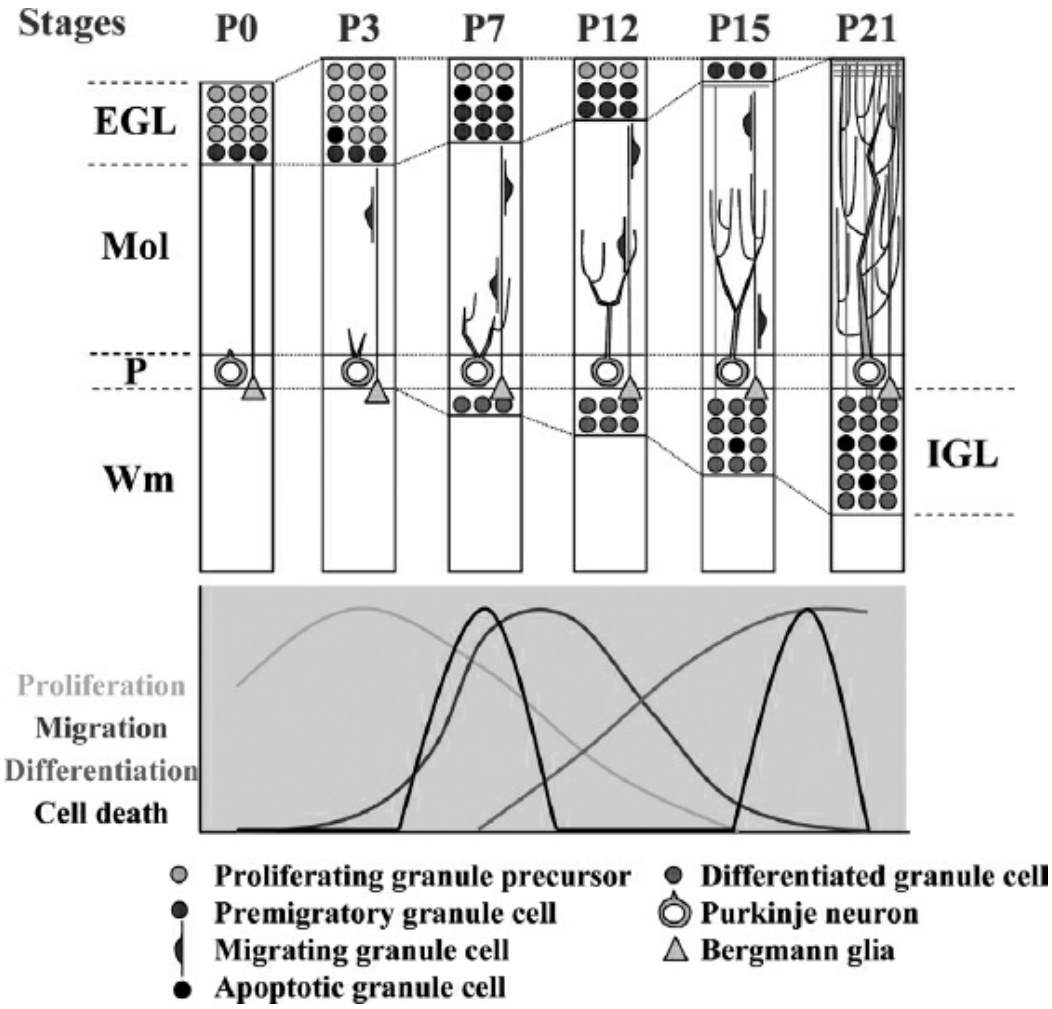

FIGURE 1.3–6. Neurogenesis, migration and differentiation of granule cells during cerebellar development. Granule cell precursors proliferate in the external germinal layer. After exiting the cell cycle, they migrate through the molecular layer and past the Purkinje neurons to reach the internal granule layer where they differentiate and make synapses. Neurons that do not migrate properly or that do not establish proper synaptic connections undergo apoptotis. EGL, external germinal cell layer; Mol, molecular layer; P, Purkinje cell layer; IGL, internal granule cell layer; Wm, white matter.

Developmental Cell Death

During nervous system development, cell elimination is required to coordinate the proportions of interacting neural cells. Developmental cell death is a reproducible, spatially and temporally restricted death of cells that occurs during the organism's development. Three types of developmental cell death have been described: (i) Phylogenetic cell death that removes structures in one species that served evolutionarily earlier ones, such as the tail or the vomeronasal nerves; (ii) Morphogenetic cell death, which sculpts the fingers from the embryonic paddle and is required to form the optic vesicles, as well as the caudal neural tube; and (iii) Histogenetic cell death, a widespread process that allows removal of select cells during development of specific brain regions. Depending on brain region, histogenetic cell death can impact between 20 and 80 percent of

neurons. A major role for developmental cell death was proposed in the 1980s based on the paradigm of nerve growth factor (NGF), suggesting that following neurogenesis, neurons enter in competition for trophic factors. In this model, survival of differentiating neurons depended absolutely on establishing axonal connections to the correct targets in order to obtain survival promoting (trophic) growth factors, such as the neurotrophins. Otherwise, they would be eliminated by programmed cell death. This competitive process was thought to ensure proper matching of new neuronal populations with the size of its target field. Although such interactions are involved in controlling cell degeneration, this model is overly simplistic: developmental cell death also occurs in neural precursors and immature neurons, before any synaptic contacts are established. Based on morphological criteria, three types of programmed cell death have been described. The first type, "apoptotic cell death," is the most common, and is characterized by chromatin condensation and membrane blebbing, followed by nuclear fragmentation and cell shrinkage. "Autophagic degeneration" involves contiguous groups of degenerating cells and features autophagic vacuoles and pyknotic nuclei. Much less common are "nonlysosomal disintegration" and "cytoplasmic-type cell death," forms that exhibit similarities to necrosis.

As apoptotic cell death, or apoptosis, is the major type of developmental cell degeneration, underlying molecular mechanisms have been extensively examined. Apoptosis involves molecules that possess enzymatic activities such as **c**ysteine-containing **asp**artate-specific prote**ases**, also called caspases, which participate in complex intracellular mechanisms. A large number of signals (both pro- and antiapoptotic) converge to regulate common signaling pathways. Both developmental as well as pathological cell death involve many of the same signaling cascades. A failure to inhibit apoptosis is involved in cancers and autoimmune diseases (multiple sclerosis), while excess stimulation of apoptosis is observed in neurodegenerative diseases during development (Huntington disease, lysosomal diseases, leukodystrophy) and aging (Alzheimer's, Parkinson's). Massive apoptosis is also observed during developmental brain injuries such as hypoxia-ischemia, fetal alcohol syndrome (FAS), or exposure to ionizing radiations and neurotoxicants. Thus dysregulation of apoptotic cell death during development can lead to severe brain abnormalities, which may only become manifest later as mature functional impairments.

Mechanisms of programmed cell death are divided into three phases: First, a regulatory phase, termed "initiation," involves numerous extracellular and intracellular factors that promote survival or death. These signals converge on common components, such as initiator caspases, which serve as a switch to initiate (or not) cell degeneration. Then, in the case of cell death ignition, the second phase called "execution" begins. During the execution phase, effector enzymes such as caspases 3 and 7 are activated

and cleave specific substrates leading to the last and irreversible step of apoptosis. During this final programmed degeneration, caspase-activated DNAse (CAD) cleaves exposed chromosomal DNA located between the nucleosomes, cytoskeletal components are disassembled, and plasma membranes swell into vesicles termed apoptotic bodies. The cell is then dismantled and phagocytosed without any release of its contents, which would otherwise induce a damaging inflammatory response.

In mammals, regulation of programmed cell death is highly complex (Fig. 1.3–7). Historically, two main pathways were described: (i) The "extrinsic pathway," which mediates effects of death factors such as tumor necrosis factor α (TNF-α) and Fas ligand and involves recruitment of caspase-8; (ii) The "intrinsic pathway" that involves release of mitochondrial factors, such as cytochrome C, and activation of caspase-9. It is now clear that the concept of distinct separation of these two pathways is overly simplistic: in most cases the cell death decision results from the interaction between multiple factors, involving multiple pro- or anti-apoptotic signaling molecules, exerting positive or negative regulation of one another (Fig. 1.3–7).

Programmed cell death is a necessary process during neurodevelopment, as genetic deletion of caspases in embryonic mice produces enlarged and disorganized brains with marked regional specificity. Programmed cell death occurs at multiple stages of nervous system development, interacting with neurogenesis and differentiation with precise and complex mechanisms. As many neuropathologies also involve dysregulation of apoptosis, future studies hold promise for elucidation and treatment of neurological diseases.

THE CONCEPT OF NEURAL PATTERNING

Principles of Function

The morphological development of the nervous system is controlled by interactions between extracellular factors and intrinsic genetic programs. In many cases, extracellular signals are soluble growth factors secreted from regional signaling centers (i.e., notochord, floor or roof plates, surrounding mesenchymal tissues). The precursor's ability to respond (competence) depends on cognate receptor expression, which is determined by patterning genes whose proteins regulate gene transcription. Furthermore, the embryonic telencephalon is subdivided based on differences in morphology, connectivity, and neurochemical profiles, and distinct patterns of gene expression. Classical models had suggested that the cerebral cortex developed initially as a fairly homogeneous structure, with individual functional areas specified only after cortical layer formation, by the ingrowth of afferent axons from the thalamus. In contrast, studies now indicate that proliferative VZ precursors

display regional molecular determinants, a "protomap," which the postmitotic neurons carry with them as they migrate along radial glia to the cortical plate. Consequently, innervating thalamic afferents may only regulate intrinsic molecular determinants of the protomap. In two different genetic mutants, *Gbx2* and *Mash1,* thalamocortical innervation is disrupted, but the expression of cortical patterning genes is not altered. On the other hand, thalamic afferent growth may be directed by patterning genes, and subsequently plays roles in modulating regional expression patterns. Thus, experience-dependent processes may contribute less to cortical specialization than originally postulated.

Patterning genes express proteins that control transcription of other genes whose products include other transcription factors or proteins involved in cellular processes (i.e., proliferation, migration, differentiation). Transcription factors are proteins that contain two principal domains: a DNA-binding domain that binds promoter or enhancer regions of genes, and a trans-activating domain, which interacts with proteins that can activate or repress gene transcription. Transcription factors can form multimeric protein complexes to control gene activation, which enables a single transcription factor to play diverse roles in multiple cell types and processes. This combinatorial nature of gene promoter regulation leads to a diversity of functional outcomes when a single patterning gene is altered. Furthermore, since protein interactions depend on protein–protein affinities, there may be complex changes when a single factor's expression level is altered. This may be one important mechanism of human variation and disease susceptibility, since polymorphisms in gene promoters—known to be associated with human disease—can alter levels of gene protein products. A transcription factor may associate primarily with one partner at a low concentration, but with another at a higher titer. The multimeric nature of regulatory complexes allows a single factor to stimulate one process while simultaneously inhibiting another. During development, a patterning gene may thus promote one event by stimulating one gene promoter, while simultaneously sequestering another factor from a different promoter whose activity is required for an alternative phenotype. Finally, the factors frequently exhibit cross-regulatory activity, where one factor negatively regulates expression of another. This activity leads to the establishment of tissue boundaries, allowing the formation of regional subdivisions, such as basal ganglia and cerebral cortex in the forebrain (see below).

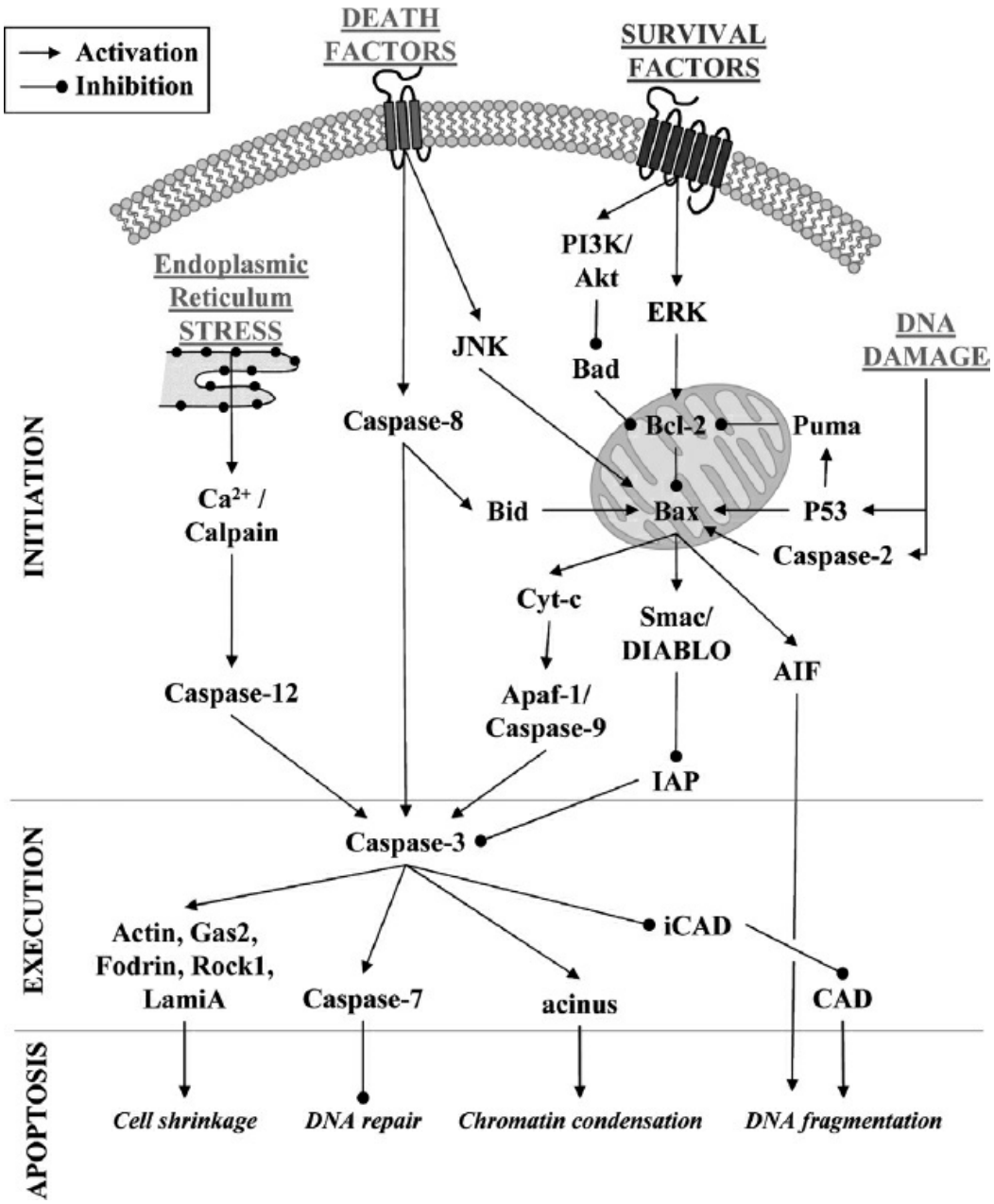

FIGURE 1.3–7. Regulation of apoptosis. Various positive and negative signals are integrated to trigger caspase activation. Caspases are present in cells as inactive zymogens and are converted into their active form through cleavage of the proenzyme. Each caspase cleaves its substrates at specific aspartate residues, thus initiator caspases cleave effector caspases leading to their activation. AIF, apoptosis inducing factor; Apaf-1, Apoptotic protease activating factor-1; Bax, Bcl-2–associated protein; Bcl-2, B cell lymphoma-2; CAD, caspase-activated DNAse; Cyt-c, cytochrome-c; ERK, extracellular regulated protein kinase; IAP, inhibitor of apoptosis proteins; JNK, c-jun N-terminal kinase; PI3K, phosphatidyl inositol triphosphate kinase; Smac, second mitochondria-derived activator of caspases = DIABLO.

In addition to combinatorial interactions, patterning genes act in hierarchical fashion by exhibiting distinct temporal sequences of expression and function. To determine functional hierarchies, researchers use genetic approaches (either deleting a gene [loss of function] or over-/ectopically expressing it [gain of function]) and observe the developmental

consequences. Generally, genetic analyses indicate that regionally restricted patterning genes participate in specifying the identity, and therefore function, of cells in which they are expressed. Subdivisions of the brain (specifically cerebral cortex) are identified by regionalized gene expression in the proliferative VZ of the neural tube, which leads to subsequent differentiation of distinct types of neurons in each mature (postmitotic) region. Thus the protomap of the embryonic VZ apparently predicts the cortical regions it will generate, and may instruct the hierarchical temporal sequence of patterning gene expression. It appears that the different genes underlie multiple stages of brain development including: (1) determining that ectoderm will give rise to nervous system (as opposed to skin), (2) defining the dimensional character of a region, such as positional identity in dorsoventral or rostrocaudal axes, (3) specifying cell class, such as neuron or glia, (4) defining when proliferation ceases and differentiation begins, (5) determining specific cell subtype, such as GABA interneuron, as well as projection pattern, and (6) defining laminar position in the region, such as cerebral cortex. Even though combinations of genes act in hierarchical fashion to control the temporal sequences of development, any single transcription factor will likely contribute to multiple stages in the developmental life of a cell, yielding complex outcomes in genetic loss/gain of function studies and human disease.

The transcription factors that pattern the embryonic nervous system belong to protein families that have been highly conserved throughout evolution. Many factors important for brain development were discovered initially in Drosophila, where they mediate body and organ segmentation and morphogenesis, or regulate neural development. In mammals, the Hox family mainly determines positional identity by delineating the anterior–posterior axis from tail to midbrain, playing major roles in defining segments of the hindbrain (rhombomeres) and its cranial nerves. The basic helix-loop-helix (bHLH) family, binding DNA and proteins through the basic and helix regions respectively, regulates multiple stages sequentially from neural plate to neurogenesis. Other transcription factor families bear names reflecting protein interaction domains, including LIM homeodomain (*Lhx*), zinc finger, paired domain (*Pax*), winged helix (*BF1* = *Foxg1, Hnf3β*), and *Pou*. It should be noted that restricting the discussion of developmental patterning to transcription factors only is arbitrary for purposes of simplicity, since downstream target genes and proteins, such as cadherins and ephrins that are important in cell migration and axon pathfinding (see below), similarly localize to specific regions.

Finally, patterning gene expression in nervous system subdivisions can be regulated by environmental signals, such as growth factors released from regional signaling centers. These growth factors establish tissue gradients that provide positional information (dorsal or ventral), impart

cell specification, and/or control regional growth. Some examples of signals include the BMPs, the Wingless-int proteins (Wnts), Shh, fibroblast growth factors (FGFs), and epidermal growth factors (EGFs). Cortical development depends on these types of signals that set up developmental domains. These are characterized by the expression of specific transcription factors that in turn, control further regional gene transcription and developmental processes. In light of the temporal and combinatorial principles discussed above, brain development can be viewed as a complex and evolving interaction of extrinsic and intrinsic information.

Specific Inductive Signals and Patterning Genes in Development

Induction of the CNS begins at the neural plate stage when the underlying notochord and mesenchyme, and surrounding epidermal ectoderm produce signaling molecules that affect the identity of neighboring cells. Perhaps surprisingly, neural differentiation is the default state of early embryonic tissues unless it is inhibited. To prevent its own neural differentiation, the neighboring ectoderm produces BMPs that instead promote and maintain epidermal differentiation. To elicit neural differentiation, Hensen node, a signaling center at the rostral end of the primitive streak, secretes inhibitory proteins that block the BMP-mediated epidermis-inducing activity. Once the neural tube closes, the roof plate and floor plate become new signaling centers, organizing dorsal and ventral neural tube, respectively. As a principle stated earlier, the same ligand/receptor system is used sequentially for multiple functions during development. BMPs are a case in point, since they prevent neural development at neural plate stage, while after neurulation, the factors are produced by the dorsal neural tube itself to induce sensory neuron fates.

The Cerebral Cortex

Forebrain development depends on interactions between inductive signals and patterning genes. The dorsal forebrain structures include the hippocampus medially, the cerebral cortex dorsolaterally, and the entorhinal cortex ventrolaterally, whereas in basal forebrain, the globus pallidus lies medially and the striatum laterally. In the embryo, the forebrain is divided into a checkerboard-like grid-pattern of transcription factors arranged in longitudinal columns and transverse segments that produce unique combinations of factors within each subdivision. Many of the genes, including *Hnf3β, Emx2, Pax6,* and *Dlx2,* are first expressed even before neurulation in the neural plate and are then maintained, providing the "protomap" determinants of the VZ described above. Initial forebrain gene expression is influenced by an array of signaling center soluble factors, including Shh, BMP, and retinoic acid. As the telencephalic vesicles form, signaling centers localize to the edges of the cortex. For example,

dorsally, there is the anterior neural ridge, an anterior cranial mesenchyme secreting FGF8, the roof plate, and, at the junction of the roof plate with the telencephalic vesicle, the cortical hem (Fig. 1.3–8). Other factors originate laterally from the junction of the dorsal and ventral forebrain regions.

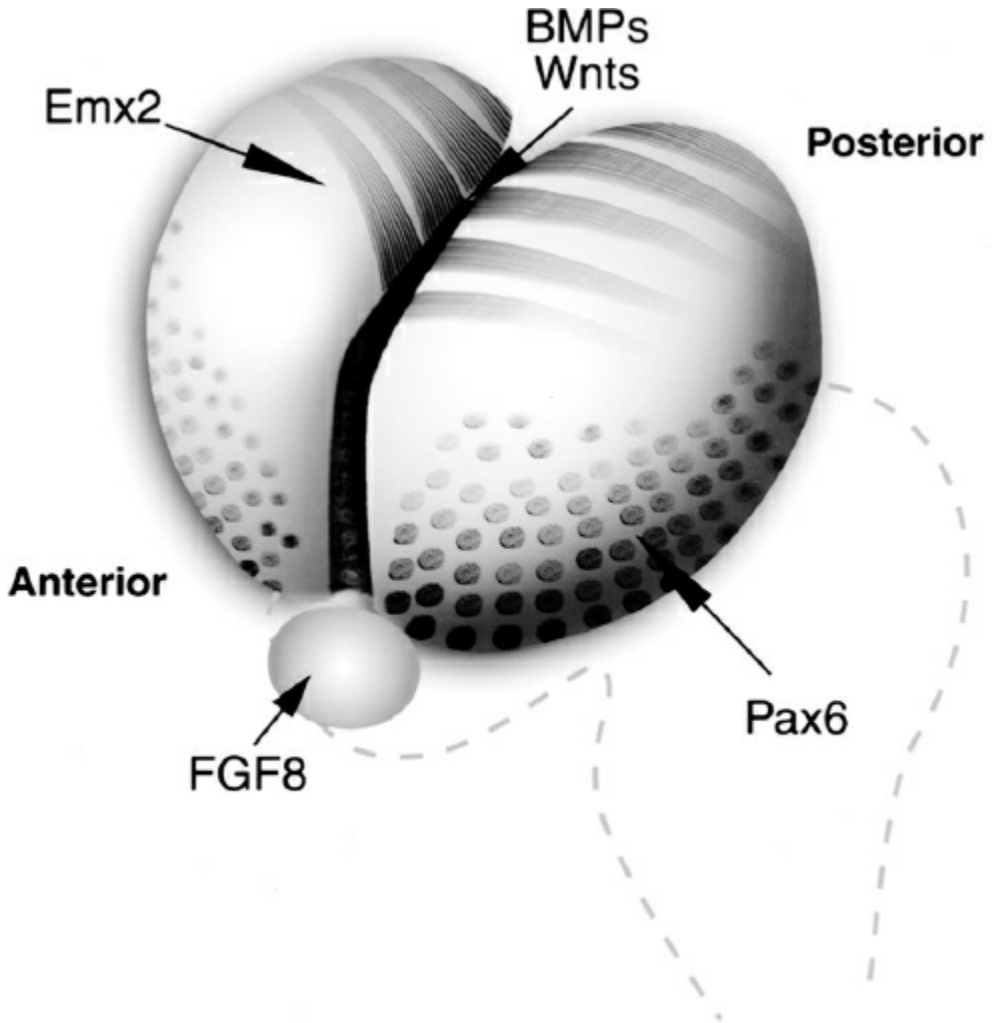

FIGURE 1.3–8. Patterning genes and signaling centers in the developing cerebral cortex. This schematic diagram shows a lateral-superior view of the two cerebral hemispheres of the embryonic mouse, sitting above the midbrain and hindbrain (*broken lines*). Circles indicate the anterior-lateral extent of Pax6 gene expression. The posterior-medial domain of Emx2 expression is indicated by stripes. The genes exhibit continuous gradients of expression that decrease as they extend to opposite poles. The signaling factor fibroblast growth factor 8 (FGF8) is produced by and released from mesenchymal tissue in the anterior neural ridge, which regulates Pax6 and Emx2 expression. In the midline, bone morphogenetic proteins (BMPs) and Wingless-Int proteins (Wnts) are secreted from other signaling centers, including the roof plate and the cortical hems. (Courtesy of E. DiCicco-Bloom and K. Forgash.)

Initial forebrain development starts with formation of two telencephalic vesicles from the rostral-most neural tube, the prosencephalon. Secreted signaling molecules, such as FGF8 and Shh, from the anterior neural ridge, the roof plate, the cortical hem, and cells of the meninges and skin,

influence this process. Disruption of roof plate signaling causes holoprosencephaly (HPE), characterized by a single forebrain ventricle with a continuous cerebral cortex across the midline. Human HPE is linked to genetic mutations in several components of the inductive cascade, including Shh, its Patched (Ptc) receptor, and several transcription factors, including *Six3, Zic2,* and *TGIF,* a component of the TGF-*β* family. Shh and *Six3* are co-expressed in the anterior neural ridge and the ventral midline, whereas *Zic2* is expressed in the dorsal roof plate. Furthermore, HPE is seen in some cases of Smith–Lemli–Opitz syndrome, a defect in biosynthesis of cholesterol, which is necessary for full activity of Shh. Thus, forebrain morphogenesis requires normal signaling center activity. On the other hand, over-expression of dorsal signal BMP elicits cyclopia in the embryo, indicating the balance of inductive signal interactions in forebrain development. Finally, when the anterior neural ridge source of FGF8 is obliterated, there is no induction of *BF1* (*Foxg1*) and an almost complete absence of cerebral cortex results.

As stated above, regional patterning genes frequently exhibit cross-regulatory activity, where one factor negatively regulates expression of its neighbors. This feature is well exemplified in cerebral cortex development, where loss of one factor leads to changes (usually expansions) in the boundaries of other regulators, with well-known consequences for human brain malformations. After telencephalic vesicles form, opposing gradients of patterning genes are critical in specifying the cortical regions. For example, the rostral/lateral cortex expresses high levels of homeodomain gene *Pax6,* whereas caudal/medial cortex exhibits *Emx2, Lhx2,* and *Lhx5*. A prediction would be that altering gene expression should cause a change in cortical areas, especially the proportions of motor to sensory cortex. Indeed, expression of motor cortex markers is markedly diminished in mice mutant for *Pax6,* and its downstream bHLH transcription factor, *Ngn2,* which it regulates. In addition, there is a proportionate increase in caudal sensory cortex. Moreover, there are changes in the dorsoventral dimension: Genes of the ventral striatum and pallidum, *Gsh* and *Dlx,* are now expressed ectopically in dorsal territory. A similar dorsal shift occurs with deletion of other dorsal transcription factors, *Ngn1/2* and *Gli3,* yielding loss of the cerebral cortex. These observations indicate that patterning genes exert reciprocal inhibitory functions in several dimensions. Significantly, in human development, PAX6 deletion results in abnormalities of the eyes (cataracts, aniridia, anophthalmia) and the olfactory epithelium and bulbs. Furthermore, the cerebral cortex is hypoplastic, exhibiting nodules of poorly differentiated cells adjacent to proliferative zones, an absence of the marginal zone (layer I) and schizencephaly, a disorder characterized by full thickness clefts through the cerebral hemispheres.

Conversely, loss of *Emx2* in mice results in a small and mispatterned

cortex, with caudomedial areas lost and expansion of anterior cortex into the vacated posterior area. Similarly, in human development, homozygous mutations of EMX2 produce schizencephaly, whereas heterozygotes exhibit less severe lesions. The gene dosage effects of human EMX2 mutations suggest that more moderate changes in EMX2 expression during development, say from promoter polymorphisms, could have more subtle yet widespread effects on cortical cell composition and function.

The impact of soluble signaling molecules on patterning genes that specify cortical areas has been elegantly demonstrated in genetic experiments that alter levels of FGF8. Over-expression of FGF8 in its normal anterior neural ridge location causes a posterior shift of cortical areas, increasing motor and reducing sensory cortices. Conversely, reducing endogenous FGF8 by over-expressing a soluble receptor fragment shifts borders anteriorly. These results indicate that FGF8 signaling alters the ratio of *Pax6* to *Emx2* in the cortical neuroepithelium, and thereby respecifies the rostrocaudal character that emerges. In addition to FGF8, Wnt and BMP signaling also regulate *Emx2* transcription, indicating that combinations of extracellular signals act on patterning gene expression, and consequent cortical development. More generally, gradients of patterning genes and soluble growth factors regulate the nature of cortical areas in all three dimensions. Downstream of these regional patterning genes are molecular targets that mediate cell–cell interactions, such as adhesive cadherins, membrane bound ephrins and their Eph receptors, and members of the immunoglobulin superfamily that play roles in cell differentiation, cell migration, and neuronal process outgrowth and pathfinding.

Given the substantial roles of early patterning genes and soluble growth factors in specifying cortical regions, do the ingrowing thalamic afferents contribute to cortical development? Do these molecular studies identify how different cortical regions interact with thalamic neurons to establish specific functional modalities, such as vision and sensation? And once regional identity is established, can it be modified by later developmental events? The abundant molecular evidence suggests that intrinsic differences are established early in the neuroepithelium by molecular determinants that regulate areal specification (the “protomap”), including the targeting of thalamic axons. However, despite the evidence of early molecular regionalization, the initial targeting of thalamic axons to the cortex is independent of these molecular differences. However, once they reach the cortex soon after birth, the thalamic axons interact with local regional cues and modify their initial outgrowth to establish proper connections that conform to cortical area molecular identities. Furthermore, the developing cortex exhibits a remarkable level of flexibility in modifying its specific functions: in the ferret, elimination of the lateral geniculate nucleus to which visual retinal ganglion axons project causes

them to reroute to the auditory cortex, which then successfully mediates vision! Thus the animal's visual information is effectively processed by their auditory cortex.

Layer-Specific Generation of Neurons during Cortical Development. As described earlier, neurons within specific cortical layers establish connections with distinct targets including upper layer neurons projecting within and between cerebral hemispheres, whereas lower layer neurons project to subcortical structures, including thalamus, basal ganglia, hindbrain, and spinal cord. Cortical lamination is an intricate process that is under regulation by both intrinsic genetic mechanisms and extracellular/environmental cues. Abnormalities in these processes can result in malformations of cortical development (MCDs), which are major contributors to mental retardation, seizures, and psychiatric disorders. Examples of MCDs include lissencephaly, focal cortical dysplasias, and polymicrogyria. MCDs can be caused by genetic mutations that effect cell proliferation, apoptosis, neuronal migration, cortical organization, and to a lesser extent, by hypoxia-ischemia and viral infection. Since most MCDs are caused by multiple gene mutations, they are typically classified according to brain morphology and genetic abnormalities. For example, lissencephaly can be caused by single mutations in four different genes (DCX, RELN, LIS1, ARX) and a partial chromosomal deletion (Chr 17p—Miller–Dieker syndrome). Recent studies of the six-layer cortical formation seek to identify layer-specific molecular markers, which can provide clues for diagnosing MCDs as well as elucidate the molecular mechanisms that control normal cortical development.

There are numerous molecular markers that are expressed during cortical development, and since they appear in a strict temporal pattern, they are convenient tools to distinguish between cortical layers and predict progenitor cell fate. During development these markers are predominantly layer-specific, but they become less so with maturation because the types of neurons found within layers are more diverse. Therefore, the term "layer-specific" marker may only label a subset of projection neurons in one or multiple layers. Through numerous studies, a variety of markers have been identified in mice, consisting of transcription factors (Ctip2, Foxp2, Tbr1—deep layer; Satb2, Brn1, Svet1, Cux1, Cux2—upper layer), hormone receptors (Ror β—deep layer), enzymes (Cyp39a1—upper), and secreted proteins (Reelin—upper layer). Cortical markers found in humans and NHPs include neurofilaments (nonphosphorylated neurofilament heavy chain—deep layer), calcium-binding proteins (calretinin—upper layer), and enhancer proteins (TLE4—deep layer).

Recent research has shown that proteins that are active during specific stages of neurogenesis can predict the laminar fate of neurons. For example, during the cortical progenitor stage in mice, high levels of the

signaling molecule β-catenin are found in precursors that later differentiate into deep-layer neurons, whereas low levels (or inhibition of β-catenin) result in increased upper-layer neuron production. Thus, altering β-catenin expression can alter the proportions of deep-layer neurons. Similarly, during the postmitotic neuron phase, the transcription factor SIP1 has been found to modulate laminar fate. Deletion of SIP1 in mice prematurely increases the production of upper-layer neurons (layers II–VI), leading to a diminished generation of lower-layer neurons. This precocious shift to upper-layer neuron generation was not caused by altered cell cycle rate or premature cell cycle exit. Thus, SIP1 might have a direct role in determining cell fate.

Do cortical progenitors have a genetic internal clock that determines when a neuron will be born? Evidence supporting an intrinsic program in cell fate determination derives from studies of cortical progenitors grown in vitro where they sequentially generate a variety of cortical neurons. The timing of cell cycle exit, when a postmitotic neuron is born, is thought to be a key determinant of laminar position, thus regulation of the cell cycle has been highly studied. Recent studies implicate proteins that inhibit progress of the cell cycle, specifically the cyclin-dependent kinase inhibitors (CKIs). These include $p57^{KIP2}$, a gene within the Beckwith–Wiedemann syndrome gene interval on chromosome 11, whose loss results in organomegaly, and structurally related $p27^{KIP1}$. $p57^{KIP2}$ is a strong modulator of the G1/S phase transition in intermediate progenitor cells (IPCs). In mice, deletion of $p57^{KIP2}$ leads to macroencephaly associated with cortical hyperplasia, which is caused by increased proliferation of both RGCs and IPCs. While the neurogenic phase of cortical development, when cells exit the cycle to become postmitotic neurons, was also delayed, the majority of new neurons produced by excess proliferation were localized to layers—V and VI, expressing lower layer markers Ctip 2 and Tbr1s. Conversely, deletion of the related CKI, $p27^{KIP1}$, promoted proliferation specifically in IPCs during E16.5, which led to increased neuron production in layers II to V that expressed marker Cux1.

This genetic internal clock may also be affected by epigenetic regulation, where changes in transcriptional potential can be maintained through multiple generations. Histone modification (acetylation, methylation) and RNA-mediated regulation are some examples of epigenetic mechanisms that modulate activation or repression of gene transcription. Studies have shown that histone methylation plays a role in controlling the proportions of upper and lower layer neurons during cortical development. For example, deletion of the histone methyltransferase gene Ezh2 (in the polycomb repressive complex 2) in mice caused a twofold reduction in the numbers of upper-layer Cux1-expressing neurons, without affecting the deep-layer neuron population. The deletion led to premature differentiation of early cortical progenitors into neurons, thereby reducing

the later pool of Pax6-expressing VZ progenitors that produce upper-layer neurons. Studies regarding histone acetylation also implicate a mechanism in modulating upper-layer neuron specification. For example, mouse embryonic stem cells (mESCs) are commonly used as an in vitro model to recapitulate cortical development. However, the system is not perfect since only early corticogenesis is reproduced, causing a lack of upper-layer neurons. VPA, an anticonvulsant and mood stabilizer, is a known inhibitor of histone deacetylases (HDACs). VPA treatment of mESC-derived precursors increased the proportion of Cux1-expressing upper-layer neurons and reduced the generation of Ctip2-expressing deep-layer neurons. This suggests that histone acetylation plays an important role in upper-layer neuron specification during late corticogenesis.

The Hippocampus

As a region of major importance in schizophrenia, depression, autism, and other disorders, defining mechanisms regulating hippocampal formation may provide clues to their developmental bases. In mouse, the hippocampus is located in the medial wall of the telencephalic vesicle. Where it joins the roof plate, the future roof of the third ventricle, there is a dorsal signaling center, the cortical hem, which secretes BMPs, Wnts, and FGFs. Genetic experiments have defined patterning genes localized to the cortical hem and hippocampal primordia, whose deletions result in a variety of morphogenetic defects. In mice lacking Wnt3a expressed in the cortical hem, the hippocampus is either completely missing or greatly reduced, while neighboring cerebral cortex is preserved. A similar phenotype is produced by deleting an intracellular factor downstream of Wnt receptor activation, the *Lef1* gene, suggesting the Wnt3a-*Lef1* pathway is required for hippocampal cell specification and/or proliferation. When another cortical hem gene, *Lhx5,* is deleted, mice lack both the hem and neighboring choroid plexus, both sources of growth factors. However, in this case, cortical hem cells proliferate in excess, and the hippocampal primordia becomes disorganized due to abnormalities in proliferation, migration, and differentiation. A related abnormality is observed with *Lhx2* mutation. Finally, a sequence of bHLH transcription factors plays roles in hippocampal neurogenesis: DG differentiation is defective in *NeuroD* and *Mash1* mutants. Significantly, expression of all these hippocampal-patterning genes is regulated by factors secreted by the anterior neural ridge, the roof plate, and the cortical hem, including FGF8, Shh, BMPs, and Wnts. These various signals and genes serve as candidates for disruption in human diseases of the hippocampus.

The Basal Ganglia

In addition to motor and cognitive functions, the basal ganglia take on new

importance in neocortical function, since they are the embryonic origin of GABA interneurons that reach the neocortex through tangential migration (Fig. 1.3–5). The expression of several transcription factors in precursors of the ventral forebrain ganglionic eminences allowed tracking of interneurons as they migrated dorsally into the cortical layers. Conversely, transcription factor deletion mutants, including *Pax6, Gsh2,* and *Nkx2.1,* exhibited diminished or absent interneurons. These transcription factors establish boundaries between different precursor zones in the ventral forebrain VZ, through mechanisms involving mutual repression. As a simplified model, the medial ganglionic eminence (MGE) expresses primarily *Nkx2.1* and gives rise to most GABA interneurons of the cortex and hippocampus, whereas the lateral ganglionic eminence (LGE) expresses *Gsh2* and generates GABA interneurons of the SVZ and olfactory bulb. The boundary between ventral and dorsal forebrain then depends on LGE interaction with the dorsal neocortex, which expresses *Pax6*. When *Nkx2.1* is deleted, LGE transcription factor expression spreads ventrally into the MGE territory, and there is a 50 percent reduction in neocortical and striatal GABA interneurons. In contrast, deletion of *Gsh2* leads to ventral expansion of the dorsal cortical molecular markers and concomitant decreases in olfactory interneurons. Finally, *Pax6* mutation causes both MGE and LGE to spread laterally and into dorsal cortical areas, yielding increased interneuron migration. The final phenotypic changes are complex, as these factors exhibit unique and overlapping expression, and interact to control cell fate.

Other transcription factors expressed in the MGE and LGE, including *Mash1, Dlx1, Dlx2, Dlx5, Dlx6, Lhx6,* and *Lhz7,* regulate both the timing of differentiation as well as the type of interneuron generated. *Mash1* is expressed in early born cells, whereas *Dlx1/Dlx2* appears in later maturing neurons, having as targets other family members, *Dlx5/Dlx6*. In the *Dlx1/Dlx2* double knock out, there is a 75 percent reduction in neocortical interneurons, and complete absence in the hippocampus, while olfactory neurons are preserved. A regulatory cascade has been suggested since *Mash1* can regulate *Dlx* expression, while *Dlx2* can induce expression of the GABA synthetic enzyme, glutamic acid decarboxylase (GAD). Consistent with this model, the *Mash1* deletion mutant exhibits reduced cortical GABA interneurons and striatal cholinergic interneurons. Similarly, *Nkx2.1* loss also alters neuron subpopulations, leading to complete absence of all cortical interneurons expressing NPY, somatostatin, and NOS. These studies suggest that transcription factors play roles at multiple stages in neuronal production including generic neuronal fate specification, as well as neuron subtype determination.

Neuronal Specification

As described for basal ganglia, throughout the nervous system

transcription factors regulate decisions at multiple levels, including determining the generic neural cell, neuron or glial cell, as well as neuron subtypes. *Mash1* can promote a neuronal fate over a glial fate, as well as induce the GABA interneuron phenotype. However, another bHLH factor, *Olig1/2,* can promote oligodendrocyte development, whereas it promotes motor neuron differentiation elsewhere, indicating that the variety of factors expressed in a specific cell leads to combinatorial effects and thus diverse outcomes for cell differentiation. The bHLH inhibitory factor, *Id,* is expressed at the transition from somatosensory to motor cortex, implying roles of family members in areal characteristics. In the hippocampus, granule neuron fate is dependent on *NeuroD* and *Math1,* with deficient cell numbers when either one is deleted. The role of specific factors in cortical cell layer determination remains an area of active investigation, but likely includes *Tbr1, Otx1,* and *Pax6*.

REGULATION OF GENE EXPRESSION BY MICRO RNAS

Over the last decade a new mechanism for regulating messenger RNA (mRNA) has been explored in simple to complex organisms that involves microRNAs (miRNAs). It is now known that miRNAs contribute not only to normal development and brain function but also to brain disorders, such as Parkinson and Alzheimer disease, tauopathies, and brain cancer. miRNAs can affect the regulation of RNA transcription, alternative splicing, molecular modifications, or RNA translation. miRNAs are 21- to 23-nucleotide long single strand RNA molecules. Unlike mRNAs that encode the instructions for ribosome complex translation into proteins, miRNAs are noncoding RNAs that are not translated, but are instead processed to form loop structures. miRNAs exhibit a sequence that is partially complementary to one but usually several cellular mRNAs. By binding to target mRNA transcripts, the miRNAs serve to interfere with their function, thereby downregulating expression of these gene products. This gene silencing involves a complex mechanism: the larger miRNA primary transcript is first processed by the Microprocessor, an enzymatic complex consisting of the nuclease Drosha and the double-stranded RNA binding protein Pasha. The mature miRNA binds to its complementary RNA and then interacts with the endonuclease Dicer, that is part of the RNA-induced silencing complex (RISC), resulting in the cleavage of the target mRNA and gene silencing.

Currently, 475 miRNAs have been identified in humans, and their total number is estimated between 600 and 3,441. Potentially, up to 30 percent of all genes might be regulated by miRNAs, a whole new layer of molecular complexity. Importantly, many miRNAs are encoded by genomic sequences outside the ~20,000 structural genes, emphasizing the importance of studying the whole genome, as opposed to only exomes, which account for only ~2 percent of the chromosome. A connection between miRNAs and

several brain diseases has already been made. For example, miR-133b, which is specifically expressed in midbrain dopaminergic neurons, is deficient in midbrain tissue from patients with Parkinson disease. Further, the miRNAs encoding miR-9, miR-124a, miR-125b, miR-128, miR-132, and miR-219 are abundantly represented in fetal hippocampus, are differentially regulated in the aged brain, and are altered in Alzheimer disease hippocampus. Similar RNA species termed small interfering RNAs (siRNAs) have been discovered in plants where they prevent the transcription of viral RNA. The mechanisms involved in these effects are closely related to that of miRNA. Thus siRNAs are now being used in both basic and clinical research to downregulate specific cellular gene products, advancing the study of pathways involved in neurodevelopment and providing new selective tools to regulate disease causing genes or therapeutic molecular targets.

REGULATION OF NEURODEVELOPMENT BY EXTRACELLULAR FACTORS

Extracellular factors that regulate cell proliferation, migration, differentiation, and survival interact with intrinsic genetic determinants to control region-specific neurogenesis. This section focuses on precursor proliferation as one model process. Patterning genes control the expression of growth factor receptors and the molecular machinery of the cell division cycle. These factors are known to stimulate or inhibit proliferation of VZ precursors and are termed autocrine, paracrine, and endocrine, meaning that the signals originate from the cells themselves, neighboring cells/tissues, and from the general circulation, respectively. Although defined initially in cell culture, a number of mitogenic growth factors are now well characterized in vivo. There are those that stimulate proliferation, such as basic fibroblast growth factor (bFGF), EGF, IGF-I, Shh, and signals that inhibit cell division, such as pituitary adenylate cyclase activating polypeptide (PACAP), GABA, glutamate, and members of the TGF-β superfamily. In addition to acting as mitogens, which stimulate reentry of cells into the cell cycle, extracellular signals also act as trophic factors, which enhance proliferation by promoting survival of the mitotic population. Stimulation of both pathways is necessary to produce maximal cell numbers. Several of the neurotrophins, especially BDNF and neurotrophin-3 (NT3), promote survival of mitotic precursors as well as the newly generated progeny.

The expression of extracellular mitogens and their receptors during development is essential for proper neurogenesis to occur. Therefore, altering their expression and activity has profound and permanent consequences. For example, administering growth factors to developing embryos or pups induces changes in proliferation in the prenatal cortical VZ, postnatal cerebellar EGL, and hippocampal DG that produce lifelong modifications in brain region population size and cell composition. In

addition, environmental factors (hypoxia, maternal/fetal infection, drug or toxicant exposure) can also dysregulate mitogenic pathways during development, which will likely produce subtle changes in the size and composition of the developing cortex. Such changes may be relevant to structural differences observed in neuropsychiatric disorders, such as depression, schizophrenia, and autism. Indeed, abnormalities in the expression of the FGF ligand and receptor systems have been found consistently in human brain samples from those with major depressive disorders (MDDs), as well as numerous animal models of stress and depressive behaviors.

Similar to cerebral cortex, later generated populations of granule neurons, such as those in cerebellum and hippocampal dentate gyrus, are also sensitive to growth factor manipulation, especially relevant to therapies administered intravenously to premature and newborn infants in the neonatal nursery. Like the human, cerebellar granule neurons are produced postnatally in rat, but for only three weeks, whereas in both species, hippocampal DG neurons are produced throughout life. Proliferative signals regulating cerebellar granule neurogenesis include bFGF, Shh and PACAP, whose disruption contributes to human medulloblastoma, whereas in hippocampus, the Wnt family may be involved.

Therapies that cause alterations in the effects of growth factors in newborns have several clinical implications. First, the long-term consequences of administering therapeutic agents, which might have neurogenetic effects on the brain, may need to be investigated. Second, since bFGF is as effective in stimulating adult neurogenesis (see below) as in newborns because of specific transport across the mature BBB, it is possible that other protein growth factors preferentially enter the brain and alter ongoing neurogenesis. Indeed, in rats, IGF-I also stimulates mature hippocampal DG neurogenesis. Third, other therapeutics cross the BBB efficiently due to their lipid solubility, such as steroids, which inhibit neurogenesis across the age spectrum. Steroids are frequently used perinatally to promote lung maturation and treat infections and trauma, but effects on human brain formation have not been examined. Fourth, it is well known that neurological development may be delayed in children experiencing serious systemic illness that is associated with numerous inflammatory cytokines, and one may wonder to what degree this reflects interference with neurogenesis and concomitant processes, potentially producing long-term differences in cognitive and motor functional development. Finally, maternal infection during pregnancy is a known risk factor for schizophrenia, and cytokines that cross the placental barrier may directly affect fetal brain cell proliferation and differentiation as well as cell migration, axon target selection, and synapse maturation as shown in animal models, eventually leading to multiple brain and behavioral

abnormalities in the adult offspring.

Cell Migration

During development, new neurons migrate radially (cortex), tangentially (inhibitory GABA interneurons from MGEs), or through a migratory stream (granule interneurons of the olfactory bulb), to their final destinations. Disrupting neuronal migration can result in abnormal cell localization and function. Animal models have defined molecular pathways involved in neuronal migration. Cell movement requires signals to start and stop migration, adhesion molecules to guide migration, and functional cytoskeleton to mediate cell translocation. The best-characterized mouse model of aberrant neuronal migration is reeler, a spontaneous mutant in which cortical neuron laminar position is inverted, being generated in outside-to-inside fashion. Reelin is a large, secreted extracellular glycoprotein produced embryonically by the earliest neurons in the cortical preplate, Cajal–Retzius cells, and in hippocampus and cerebellum. Molecular and genetic analysis has established a signaling sequence in reelin activity which includes at least two receptors, the very-low-density lipoprotein receptor (VLDLR) and the apoprotein E receptor 2 (ApoER2), and the intracellular adapter protein, disabled 1 (Dab1), initially identified in the scrambler mutant mouse, a reelin phenocopy. Current thoughts consider the reelin system as one mediator of radial glial-guided neuron migration, though specific functions in starting or stopping migration remain controversial. The roles of the VLDL and ApoE2 receptors are intriguing for their possible contributions to Alzheimer disease risk, based on many genetic studies. Other studies have found human reelin gene (RELN) mutations associated with autosomal recessive lissencephaly with cerebellar hypoplasia, exhibiting a markedly thickened cortex with pachygyria, abnormal hippocampal formations, and severe cerebellar hypoplasia with absent folia. Additional studies suggest that reelin polymorphisms may contribute to ASD risk as well.

With regard to cytoskeletal proteins, studies of the filamentous fungus, *Aspergillus nidulans,* surprisingly provided insights into molecular machinery underlying the human migration disorder, Miller–Dieker syndrome, a lissencephaly associated with abnormal chromosome 17q13.3. Lissencephaly is a diverse disorder characterized by a smooth cortical surface lacking in gyri and sulci, with markedly reduced brain surface area. The absence of convolutions results from a migration defect, the majority of neurons failing to reach final destinations. In classical lissencephaly (Type I), cerebral cortex is thick and usually four-layered, while in cobblestone lissencephaly (Type II) the cortex is chaotically organized with a partly smooth and partly pebbled surface and deficient lamination. The most severely affected parts of the brain are the cerebral cortex and hippocampus, with cerebellum less affected. In fungus, the gene *NudF* was

found essential for intracellular nuclear distribution, a translocation process also involved in mammalian cell migration. The human homologue of *NudF* is LIS-1 or PAFAH1B1, the mutation of which accounts for up to 60 percent of lissencephaly cases, of type I pathology. The LIS-1 gene product interacts with microtubules, and related motor components, dynein and dynactin, as well as doublecortin (DCX), which regulate microtubule stability. Mutations in DCX gene result in X-linked lissencephaly in males and bands of heterotopic neurons in white matter in females, appearing as a "double cortex" on imaging studies, producing severe mental retardation and epilepsy. Other migratory defects occur when proteins associated with the actin cytoskeleton are affected, such as mutation in filamin 1 gene responsible for periventricular nodular heterotopias in humans, and mutations of a regulatory phosphokinase enzyme, the CDK5/p35 complex.

Cell migration also depends on molecules mediating cellular interactions, which provide cell adhesion to establish neuron–neuron and neuron–glial relationships, or induce attraction or repulsion. Astrotactin is a major glial protein involved in neuronal migration on radial glial processes, whereas neuregulins and their receptors, ErbB2–4, play roles in neuronal–glial migratory interactions. Studies associate neuregulin polymorphisms with schizophrenia, suggesting that this developmental disease may depend on altered oligodendrocyte numbers and activities and synaptic functions. Furthermore, some work suggests that early appearing neurotransmitters themselves, GABA and glutamate, and platelet-derived growth factor (PDGF), regulate migration speed. In contrast to radial migration from cortical VZ, GABA interneurons generated in ganglionic eminences employ different mechanisms to leave the ventral forebrain and migrate dorsally into the cerebral cortex. Several signaling systems have been identified, including the Slit protein and Robo receptor, the semaphorins and their neuropilin receptors, and hepatocyte growth factor and its c-Met receptor, all of which appear to repel GABA interneurons from basal forebrain, promoting tangential migration into cortex (Fig. 1.3–5). Significantly, the c-Met receptor is associated with ASDs, suggesting that altered GABA interneuron migration into cortex and deficits in inhibitory signaling may contribute to the phenotype including seizures and abnormal cognitive processing. Finally, several human forms of congenital muscular dystrophy with severe brain and eye migration defects result from gene mutations in enzymes that transfer mannose sugars to serine/threonine -OH groups in glycoproteins, interrupting interactions with several extracellular matrix molecules, producing type II cobblestone lissencephalies.

Differentiation and Neuronal Process Outgrowth

After newly produced neurons and glial cells reach their final destinations, they differentiate into mature cells. For neurons, this involves outgrowth of

dendrites and extension of axonal processes, formation of synapses, and production of neurotransmitter systems, including receptors and selective reuptake sites. Most axons will become insulated by myelin sheaths produced by oligodendroglial cells. Many of these events occur with a peak period from 5 months of gestation onward. During the first several years of life, many neuronal systems exhibit exuberant process growth and branching, which is later decreased by selective "pruning" of axons and synapses dependent on experience, while myelination continues for several years after birth and into adulthood.

While there is tremendous synapse plasticity in adult brain, a fundamental feature of the nervous system is the point-to-point or topographic mapping of one neuron population to another. During development, neurons extend axons to innervate diverse distant targets, such as cortex and spinal cord. The structure that recognizes and responds to cues in the environment is the growth cone, located at the axon tip. The axonal process is structurally supported by microtubules that are regulated by numerous microtubule-associated proteins (MAPs), whereas the terminal growth cone exhibits a transition to actin-containing microfilaments (Fig. 1.3–9). The growth cone has rod-like extensions called filopodia that bear receptors for specific guidance cues present on cell surfaces and in extracellular matrix. Interactions between filopodial receptors and environmental cues cause growth cones to move forward, turn, or retract. Abnormalities in actin modulating proteins and kinases involved in rapid growth cone movements may contribute to language and cognitive disorders, such as LIMK kinase whose mutation causes Williams syndrome. The region-specific expression of extracellular guidance molecules, such as cadherins, regulated by patterning genes *Pax6* and *Emx2,* results in highly directed outgrowth of axons, termed axonal pathfinding. These molecules affect the direction, speed, and fasciculation of axons, acting through either positive or negative regulation. Guidance molecules may be soluble extracellular factors, or alternatively, may be bound to extracellular matrix or cell membranes, such as the transmembrane ephrin proteins. Playing major roles in topographic mapping between neuron populations and their targets, ephrins act via the largest known family of tyrosine kinase receptors in brain, Eph receptors. Ephrins frequently serve as chemorepellent cues, negatively regulating growth by preventing developing axons from entering incorrect target fields. For example, the optic tectum (superior colliculus) expresses ephrins A2 and A5 in decreasing gradients from posterior to anterior, whereas innervating retinal ganglion cells express a gradient of Eph receptors. Ganglion cell axons from posterior retina, which possess high Eph A3 receptor levels, preferentially innervate the anterior tectum because the low-level ephrin expression does not activate the Eph kinase that causes retraction. In the category of soluble molecules, netrins serve primarily as

chemoattractant proteins secreted, for instance, by the spinal cord floor plate to stimulate spinothalamic sensory interneurons to grow into the anterior commissure, whereas Slit is a secreted chemorepulsive factor which through its roundabout (Robo) receptor regulates midline crossing, and axonal fasciculation and pathfinding.

In neocortex, layers V and VI axons exit the hemisphere laterally via the internal capsule to reach subcortical destinations, whereas layer III axons extend medially through corpus callosum to innervate the opposite hemisphere. The internal capsule carries bidirectional axons, from cortex to thalamus and beyond, as well as thalamocortical processes, exhibiting precise connections between individual thalamic nuclei and distinct cortical domains. During development, thalamic axons must travel a complex route, passing through lateral ventral thalamus, turning to enter the internal capsule and turning dorsally to reach cortical targets. However, thalamic axons reach the developing neocortex before their target neurons have reached the appropriate layers. Instead, the early-generated subplate neurons lying below layer VI serve as temporary targets for thalamic axons. The subplate neurons express the chemoattractant netrin 1 and chemorepellent ephrin-A5, which is complemented by Eph receptor expression on thalamic axon growth cones. After cortical neurons complete laminar migration, thalamic axons leave subplate neurons, which apparently undergo apoptosis, and extend into proper cortical layers guided by a number of cues, including chondroitin sulphate proteoglycans, ephrins, and cadherins under patterning gene regulation. Similar principles regulate axonal targeting throughout the brain and misrouted axon terminals have been shown in animal models when ephrin/Eph expression is altered.

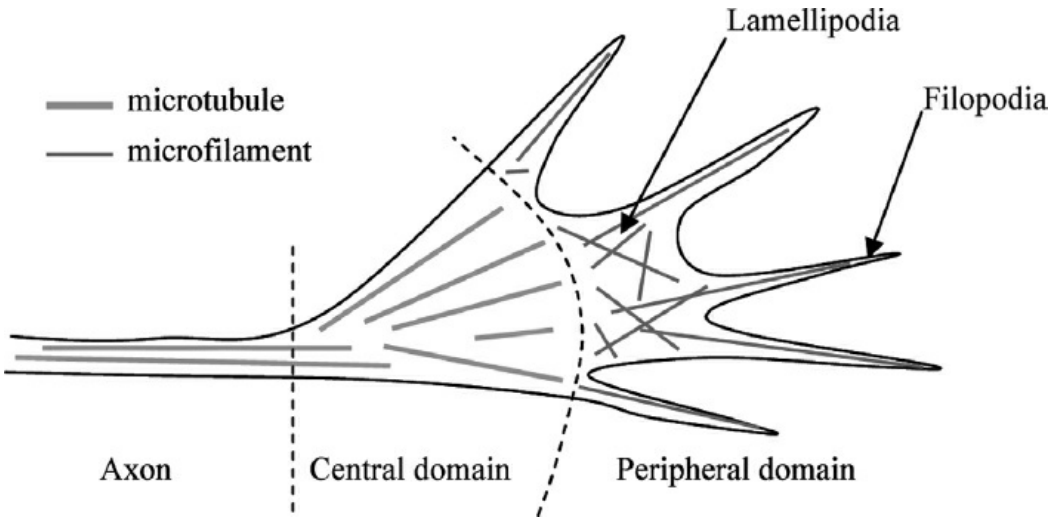

FIGURE 1.3–9. Structure of the growth cone. The cone is subdivided into two domains: the central domain, which contains mitochondria and microtubules, and the peripheral domain, containing veil-like lamellipodia and spike-like filopodia. In the lamellipodia, microfilaments consisting of actin form a meshwork, while in filopodia, they have the same

orientation. Cell surface receptors on growth cone processes sense extracellular guidance cues to control navigation.

GENETIC AND ENVIRONMENTAL FACTORS IN NEUROPSYCHIATRIC DISEASE

An increasing number of neuropsychiatric conditions such as schizophrenia, depression, autism, and attention deficit hyperactivity disorder (ADHD) are considered to originate during brain development. In addition, ongoing research suggests that environmental factors such as neurotoxicants and drugs of abuse contribute to the development of neuropsychiatric disease. Defining when a condition begins helps direct attention to underlying pathogenetic mechanisms. The term neurodevelopmental suggests that the brain is abnormally formed from the very beginning due to disruption of fundamental processes, in contrast to a normally formed brain that is injured secondarily or that undergoes degenerative changes. However, the value of the term neurodevelopmental needs to be reconsidered, because of different usage by clinicians and pathologists. In addition, given that the same molecular signals function during development and maturation, altering an early ontogenetic process (i.e., by changes in growth factor signaling) probably means that other adult functions will also exhibit ongoing dysregulation. For example, clinical researchers of schizophrenia consider the disorder neurodevelopmental because at the time of onset and diagnosis, the prefrontal cortex and hippocampus are smaller and ventricles enlarged already at adolescent presentation. In contrast, the neuropathologist uses the term neurodevelopmental for certain morphological changes in neurons. If a brain region exhibits a normal cytoarchitecture but with neurons of smaller than normal diameter, reminiscent of “immature” stages, then this may be considered an arrest of development. However, if inflammatory signs, such as gliosis and white blood cell infiltrate, accompany the same cellular changes this is termed neurodegeneration. These morphological and cellular changes may no longer be adequate to distinguish disorders that originate from development versus adulthood, especially given the roles of glial cells, including astrocytes, oligodendrocytes, and microglia, as sources of neurotrophic support during both periods of life. Thus, abnormalities in glial cells may occur in both epochs to promote disease, or act as mechanisms of repair. Many neurodegenerations are associated with microglial cells such as Alzheimer and Parkinson diseases. On the other hand neuronal dysfunction in adulthood associated with cell shrinkage may occur without inflammatory changes. In animal models, interrupting BDNF neurotrophic signaling in adult brain results in neuronal and dendrite atrophy in cerebral cortex without eliciting glial cell proliferation. Thus finding small neurons without gliosis in the brains of schizophrenic and autistic patients does not necessarily mean the condition is only or primarily developmental in

origin. In turn, several etiological assumptions about clinical brain conditions may require reexamination.

Because the same processes that mediate development, including neurogenesis, gliogenesis, axonal growth and retraction, synaptogenesis, and cell death, also function during adulthood, a new synthesis has been proposed by Mattson. All of these processes, though perhaps in more subtle form, contribute to adaptive and pathological processes (Fig. 1.3–10). Successful aging of the nervous system may require precise regulation of these processes, allowing the brain to adapt properly and counteract the numerous intrinsic and extrinsic events that could potentially lead to neuropathology. For example, adult neurogenesis (see below) and synaptic plasticity are necessary to maintain neuronal circuitry and ensure proper cognitive functions. Programmed cell death is crucial to prevent tumorigenesis that can occur as cells accumulate mutations throughout life. Thus, dysregulation of these ontogenetic processes in adulthood will lead to disruption of brain homeostasis, expressing itself as various neuropsychiatric diseases (Fig. 1.3–10).

Schizophrenia

As schizophrenia and its causes is the subject of many chapters in this textbook, this chapter's discussion is restricted to several disease manifestations that may involve neurodevelopmental mechanisms. The neurodevelopmental hypothesis of schizophrenia postulates that etiologic and pathogenetic factors occurring before the formal onset of the illness (during gestation) disrupt the course of normal development. These subtle early alterations in specific neurons, glia, and circuits confer vulnerability to other later developmental factors, ultimately leading to malfunctions. Schizophrenia is clearly a multifactorial disorder, which is caused by both genetic and environmental factors. Clinical studies using risk assessment have identified some relevant factors, including prenatal and birth complications (hypoxia, infection, substance, and toxicant exposure), family history, body dysmorphia, especially structures of neural crest origin, and presence of mild premorbid deficits in social, motor, and cognitive functions. These risk factors may impact ongoing developmental processes such as experience-dependent axonal and dendritic production, programmed cell death, myelination, and synaptic pruning. An intriguing animal model using human influenza-induced pneumonia of pregnant mice shows that the inflammatory cytokine response produced by the mother may directly affect the offspring's brain development, with no evidence of the virus in the fetus or placenta.

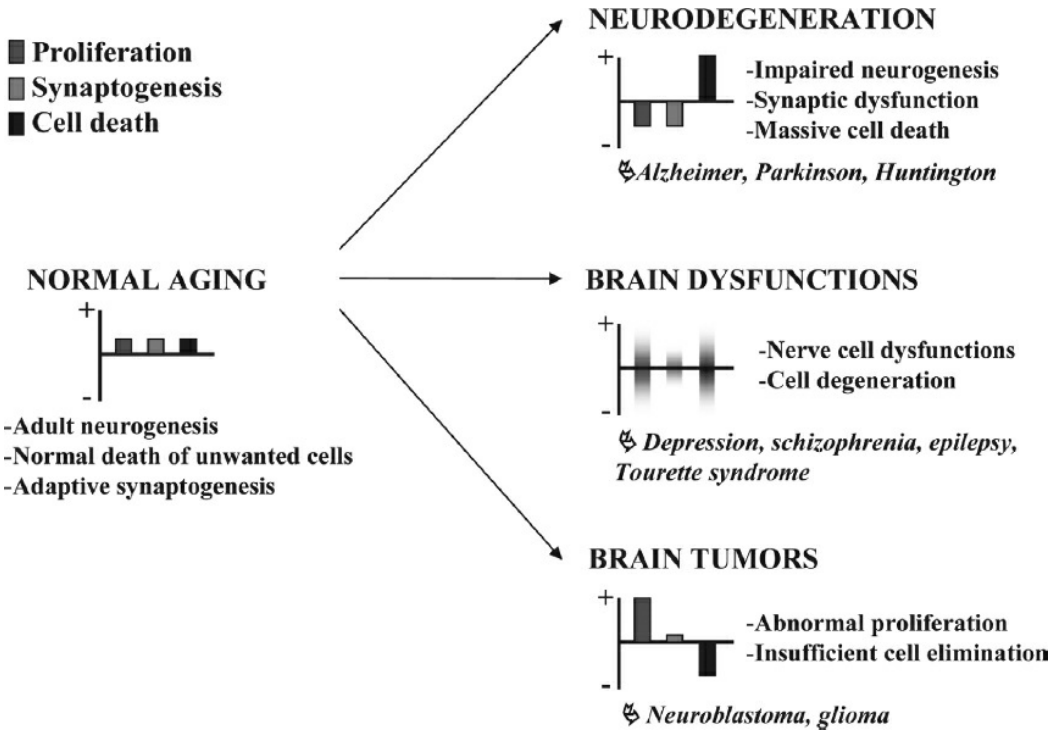

FIGURE 1.3–10. Dysregulation of neurodevelopmental processes during aging. Successful aging requires a balance between adult neurogenesis, death of unwanted cells (tumoral cells), and adaptive synaptogenesis. These processes involve mechanisms similar (but not identical) to those observed during neurodevelopment: cell proliferation, synaptogenesis, cell death. Dysregulation of these mechanisms can lead to neurodegeneration, brain tumors, or various brain dysfunctions.

Neuroimaging and pathology studies suggest abnormal development because there are structural abnormalities at disease presentation, including smaller prefrontal cortex and hippocampus and enlarged ventricles. More severely affected patients exhibit a greater number of affected regions with larger changes. In some cases, ventricular enlargement and cortical grey matter atrophy increase with time. These ongoing progressive changes should lead one to reconsider the potential role of active degeneration in schizophrenia, whether due to the disease, or its consequences, such as stress or drug treatment. However, classic signs of neurodegeneration with inflammatory cells are not present.

Structural neuroimaging strongly supports the conclusion that the hippocampus in schizophrenia is about 5 percent smaller. In turn, brain morphology has been used to assess etiological contributions of genetic and environmental factors. Comparisons of concordance for schizophrenia in monozygotic and dizygotic twins support roles for both factors. Among monozygotic twins, only 40 to 50 percent of both twins have the illness, indicating genetic constitution alone does not assure disease, and suggesting the embryonic environment also contributes. Neuroimaging, pharmacological, and pathological studies suggest that some genetic factors confer susceptibility, and that secondary insults, such as birth trauma or perinatal viral infection, also contribute. This model is consistent with

imaging studies showing small hippocampus in both affected and unaffected monozygotic twins. Moreover, healthy individuals with genetic susceptibility show hippocampal volumes (smaller) more similar to affected probands than normal controls. Thus hippocampal volume reduction is not pathognomonic of schizophrenia, but rather may represent a biological marker of genetic susceptibility. It is not difficult to envision roles for altered developmental regulators in producing a smaller hippocampus, which in turn limits functional capacity. A smaller hippocampus may result from subtle differences in the levels of transcription factors, such as *NeuroD, Math1,* or *Lhx,* signaling by Wnt3a and downstream mediator *Lef1,* or proliferative control mediated by bFGF, whose family members exhibit altered expression levels in schizophrenia brain samples. Such genetic limitations may only become manifest following another developmental challenge, such as gestational infection, stressors, or toxicant exposure.

A regional locus of schizophrenia pathology remains uncertain but may include hippocampus, entorhinal cortex, multimodal association cortex, limbic system, amygdala, cingulate cortex, thalamus, and MTL. Despite size reductions in specific regions, attempts to define changes in cell numbers have been unrewarding, since most studies do not quantify the entire cell population, but only assess regional cell density. Without assessing a region's total volume, cell density measures alone are limited in revealing population size. Most studies have found no changes in cell density in diverse regions. A single study successfully examining total cell number in hippocampus found normal neuron density and a 5 percent volume reduction on the left and 2 percent on the right, yielding no significant change in total cell number.

In contrast to total neuron numbers, studies using neuronal cell type-specific markers have found a decreased density of GABA interneurons in cortex and hippocampus. In particular, parvalbumin-expressing interneurons are reduced, whereas calretinin containing cells are normal, suggesting a deficiency of an interneuron subtype. These morphometric data are supported by molecular evidence for decreased GABA neurons, including reduced mRNA and protein levels of the GABA synthesizing enzyme, GAD67, in cortex and hippocampus. Another product of the adult GABA secreting neurons, reelin, which initially appears in Cajal–Retzius cells in embryonic brain, is reduced 30 to 50 percent in schizophrenia and bipolar disorder with psychotic symptoms. Such a deficiency, leading to diminished GABA signaling, may underlie a potential compensatory increase in GABAa receptor binding detected in hippocampal CA2 to 4 fields by both pyramidal and nonpyramidal neurons, apparently selective since benzodiazepine binding is unchanged. More generally, deficiency in a subpopulation of GABA interneurons raises intriguing new possibilities for schizophrenia etiology. As indicated in the gene patterning section above,

different subpopulations of GABA interneurons originate from distinct precursors located in the embryonic basal forebrain. Specifically, cortical and hippocampal GABA interneurons derive from the MGE under control of patterning gene *Nkx2.1*. Further, the timing and sequence of GABA interneuron generation depend on a regulatory network including *Mash1, Dlx1/2,* and *Dlx5/6,* all gene candidates for schizophrenia risk. Indeed, *DLX1* gene expression is reduced in the thalamus of patients with psychosis. Thus abnormal regulation of these factors may diminish selectively GABA interneuron formation, which in turn may represent a genetically determined vulnerability, and may contribute to diminished regional brain size and/or function.

The most compelling neuropathological evidence for a developmental basis is the finding of aberrantly localized or clustered neurons especially in lamina II of the entorhinal cortex and in the white matter underlying prefrontal cortex, and temporal and parahippocampal regions. These abnormalities represent alterations of developmental neuronal migration, survival and connectivity. In addition, in hippocampus and neocortex, pyramidal neurons appear smaller in many studies, exhibiting fewer dendritic arborizations and spines with reduced neuropil, findings that are associated with reductions in neuronal molecules including MAP2, spinophilin, synaptophysin, and SNAP25.

Autism Spectrum Disorders

Another condition that is clearly neurodevelopmental in origin is ASD, a complex and heterogeneous group of disorders characterized by abnormalities in social interaction and communication, and the presence of restricted/repetitive interests and activities. Based on the DSM-5 classification, the ASD now replaces the previously separate disorders of autistic disorder, Asperger syndrome, and pervasive developmental disorder-not otherwise specified. These three disorders are grouped together because of their common occurrence in families, indicating related genetic factors, and shared signs and symptoms which exhibit a wide range of symptom severity. Nonetheless, recent conceptualizations of ASD propose that there are multiple “autisms” differing in underlying pathogenetic mechanisms and manifestations. It is likely that the different core symptom domains (or other endophenotypes) will be more heritable than the syndromic diagnosis, which was constructed to be inclusive. The large diversity of ASD signs and symptoms reflects the multiplicity of abnormalities observed in pathological and functional studies, and include both forebrain and hindbrain regions. Forebrain neurons in the cerebral cortex and limbic system play critical roles in social interaction, communication, learning, and memory. For example, the amygdala, which connects to prefrontal and temporal cortex and fusiform gyrus, plays a prominent role in social and emotional cognition. In ASD, the amygdala

and fusiform gyrus abnormally activate during facial recognition and emotional attribution tasks. Some investigators hypothesize that ASD reflects dysfunctions in specific neural networks, such as the social network. On the other hand, neurophysiological tests of evoked cortical potentials and oculomotor and auditory responses indicate normal perception of primary sensory information, but disturbed higher cognitive processing. The functional impairments in higher-order cognitive processing and neocortical circuitry suggest that this developmental disorder involves synaptic organization more broadly. This mechanism may be uniformly present throughout the brain, which is a model in distinct contrast to abnormalities of specific neural networks. Examples of such widely expressed molecules include those involved in pre- and postsynaptic adhesion and signaling including neurexins and neuroligins, as well as postsynaptic density (PSD) machinery, such as SHANK3 and mTOR.

One of the most important discoveries in ASD pathogenesis is the widely reported and replicated brain growth phenotype. Starting with probably normal size at birth, the brain exhibits an accelerated increase in volume by the end of the first year compared to the typically developing child based on neuroimaging and head circumference measures, and this process continues from ages 2 to 4 years. It is not known whether this reflects an acceleration of normal developmental processes, or alternatively, a disease specific aberration in postnatal development, including changes in cell numbers, neuronal processes, synapse formation and pruning, or glial cell dysfunction, to name a few. The most prominent differences are observed in frontal and parietal cortex, cerebellar hemispheres, and the amygdala. These findings are also consistent with reports of macrocephaly in up to ~20 percent of ASD cases.

Functional neuroimaging studies indicate broad forebrain but also cerebellar dysfunctions in ASD, and classical pathological studies suggested abnormalities that were restricted to limbic and cerebellar structures. However, classical studies were hampered by small sample size, poor control for co-morbidities like epilepsy and mental retardation, and use of tissue cell density measures as opposed to unbiased stereological methods to estimate neuron numbers. While previous studies described increased densities of small neurons in limbic nuclei, including CA fields, septum, mammillary bodies, and amygdala, other laboratories have not replicated these results. In contrast, the most consistent neuropathology has been observed in the cerebellum (21 of 29 brains), showing reductions in the number of Purkinje neurons without signs of acquired postnatal lesions, suggesting prenatal origins.

A number of studies now identify widespread and nonuniform abnormalities, suggesting dysregulation of many processes, including neuron proliferation, migration, survival, organization, and programmed

cell death. In cerebral cortex, neuropathology reveals thickened or diminished grey matter, disorganized laminar patterns, misoriented pyramidal neurons, ectopic neurons in both superficial and deep white matter, and increased or decreased neuron densities. This evidence of abnormal cortical neurogenesis and migration accords well with the deficits in cognitive functions. In brainstem, neuronal disorganization appeared as discontinuous and malpositioned neurons in olivary and dentate nuclei, ectopic neurons in medulla and cerebellar peduncles, and aberrant fiber tracts. There were widespread patchy or diffuse decreases of Purkinje neurons, sometimes associated with increased Bergmann glia, or ectopic Purkinje neurons in the molecular layer. Hippocampal neuronal atrophy was not observed, and quantitative stereology found no consistent change in neuron density or number. More recent studies define region-specific decreases in neuron numbers and/or neuron size in circuits involved in language and social communication, including the fusiform gyrus and inferior frontal cortex. While seemingly incompatible, these various data support a model of developmental abnormalities occurring at different times, altering regions according to specific schedules of neurogenesis and differentiation. Importantly, a similar range of abnormalities were found in classical studies but were originally discounted because they did not occur in every brain examined.

Current analyses suggest that poorly defined environmental factors and multiple genetic factors interact to produce ASD. An early but rare environmental factor was the anti-nausea drug, thalidomide, which was found to be teratogenic. In 15 children exposed to thalidomide during days 20 to 24 of gestation, when cranial and Purkinje neurogenesis occurs in brainstem, four cases exhibited autism, suggesting that autism is associated with brain insults at 3 weeks. Another environmental factor is gestational exposure to VPA (Depakote). This anticonvulsant is given for seizures, mood disorders, and migraine prevention, but produces major congenital malformations and later life neurodevelopmental disorders. New clinical studies suggest that gestational insults, prematurity, and perinatal cerebellar injuries are major factors in ASD pathogenesis, an area that is under active study.

Genetic factors play major roles in ASD etiology, and there are numerous available reviews. The genetic factors include changes that directly affect the ~20,000 of genes as well as others that alter noncoding DNA regions (~98 percent of the chromosomal material), that contain a wide variety of nuclear RNAs involved in regulating gene expression, such as miRNAs. Of the genes, it is now thought that between 500 and 1,000 genes may contribute to ASD risk, with these many genes acting on far fewer molecular pathways that control all aspects of development and function. For example, many genes fall into the category of developmental transcription factors, as described above in patterning, whereas others

control synaptic structure and function, or regulate histone proteins that control gene expression. Some genetic factors are rare variants whose mutation occurs in the egg, sperm, or zygote, so-called de novo mutations, accounting for ~15 percent of ASD cases. These rare events produce major neuropsychiatric phenotypes in all persons who carry them. On the other hand, all persons carry common variants that produce small effects on phenotypes, a mechanism affecting many multifactorial diseases like diabetes and heart disease. It may be that ~40 to 50 percent of ASD risk depends on a collection of common variants to produce brain dysfunction, while they produce no disease when acting alone. However, the landscape of ASD genetics seems to be constantly changing, and current thoughts about mechanisms that were defined in relatively small populations (several hundreds to thousands) will surely change as studies employ tens of thousands of genetic samples. Furthermore, there are a number of monogenic syndromes that account for ~10 percent of cases, and these include tuberous sclerosis, Rett syndrome, fragile X mental retardation, neurofibromatosis, and Smith–Lemli–Opitz syndrome. While the genes underlying these monogenic disorders do not contribute to general idiopathic ASD, most of them point to one major locus for disease production: the synapse, where they control protein synthesis in response to synaptic activity, acting via the PI3K/Akt signaling pathway and the mTOR complex. This knowledge has led to a number of clinical trials in ASD that focus on altering synaptic activity.

Methylmercury

During development, the brain is sensitive to environmental toxicants that can act as teratogens, substances that injure the brain in reproducible patterns that reflect more the developmental period of exposure than the specific nature of the compound. Methylmercury (MeHg), a pervasive contaminant of fish and shellfish, is one of numerous environmental agents that alter developmental neurogenesis, and various epidemiological studies in frequent fish-eating populations show an association between prenatal exposure and cognitive deficits (learning and memory). Given the effects of MeHg exposure on learning and memory and the role of the hippocampus in these processes, it is reasonable to focus on neurogenesis that occurs in the hippocampal DG throughout development into adulthood. Extensive studies on the acute and long-term effects of MeHg on rat hippocampal neurogenesis reveal that there is a developmental period of vulnerability for proliferative cell types of the neurogenic cascade. In addition, these early insults can have lasting effects on later neurogenesis and can lead to the development of learning and memory deficits.

Studies using postnatal day 7 (P7) rats, which is the peak of postnatal neurogenesis and corresponds to the third trimester of human brain development, assessed the effects of a single subcutaneous MeHg injection

on hippocampal NSCs (Table 1.3–1). Twenty-four hours after exposure, biochemical and immunohistochemical analysis of the hippocampus revealed that there was a reduction of cells engaged in S-phase (BrdU), and that NSCs identified by expression of Sox2 and Nestin, underwent cell cycle arrest during the G1/S phase transition (due to cyclin E degradation). Subsequently, there was activation of the programmed cell death cascade including mitochondrial-dependent apoptosis (increased Bax, cytosolic cytochrome c, cleaved caspases 3 and 9). Consequently, 2 weeks later (P21), the volume of the DG was reduced, and many cells were diminished including the total number of granule neurons, the mitotic S-phase NSCs (particularly intermediate Tbr2+ progenitors), and the newly born DCX+ neurons. Furthermore, these early exposed rats developed hippocampal-dependent spatial learning and memory deficits (Morris water maze) in adolescence (P35). These findings demonstrate that NSCs appear to be the most vulnerable population in the hippocampus to MeHg during the first postnatal week, and that early toxicological insults have long-term effects on hippocampal structure and function.

Similar studies performed in older rats using acute exposures (24 hours) also demonstrated that the developmental window of vulnerability to MeHg in the hippocampus ends between prepubescence (P14) and adolescence (P21). This is significant because children have higher susceptibility to neurotoxicants (such as organophosphate pesticides) than adults, thus it is important to define the period of developmental vulnerability to MeHg exposure in order to guide public policy. Similar to the P7 rat, acute exposure at P14 decreased the number of S-phase cells and NSCs after 24 hours, though the magnitude of change was smaller. However, there was no sign of vulnerability in the P21 hippocampus at either 24 or 48 hours postexposure. What is the nature of this changed vulnerability? The smaller magnitude of toxic insult at P14 compared to P7 might be due to unique properties of the juvenile BBB. During fetal development, tight junctions that provide a structural barrier against xenobiotics are formed and many transporters involved in influx and efflux mechanisms are expressed. However, there is evidence that expression levels of the neutral amino acid transporters, which are a known conduit for MeHg, are higher in the developing BBB than in the adult. This might suggest that older rodents absorb less MeHg into the brain, which results in milder insult upon exposure. Alternatively, regional expression of endogenous antioxidants such as glutathione might confer additional protection in the brain. For example, astrocytes (a rich source of glutathione precursors) isolated from the cerebellum have lower glutathione capacity and are therefore more sensitive to MeHg exposure than astrocytes from the cerebral cortex. Further, the increasing presence of glia that are generated during the postnatal period might also affect the MeHg burden to NSCs. Furthermore, microglia and astrocytes residing in

the NSC niche secrete neurotrophic factors, which might provide protection against a MeHg insult. When cultured with neurons, astrocytes increase their uptake of MeHg, leading to reduced neuronal MeHg uptake and mitochondrial dysfunction. In summary, these results provide new knowledge of an expanded period of neurodevelopmental vulnerability that might inform national guidelines that regulate fish consumption during gestation and development.

Drugs of Abuse

Worldwide, substance abuse during pregnancy has been and remains a persistent public health problem. The majority of women who are addicted to drugs in the United States are of childbearing age, between 15 and 44. For decades, researchers have studied the effects of drug abuse on the developing brain. The consensus is that these children have a higher risk of developing cognitive deficits, as well as behavioral, social, emotional, and psychiatric problems. This is due to the fact that many drugs of abuse easily penetrate the placenta and fetal BBB to enter the brain, where they mimic and/or alter neurotransmitter and metabolic systems. Here, some of the most commonly used drugs during pregnancy are discussed: nicotine, cannabis (marijuana), and cocaine.

Despite public awareness campaigns, nicotine is still one of the most commonly used drugs of abuse during pregnancy. According to a CDC report (Trends in Smoking Before, During, and After Pregnancy—Pregnancy Risk Assessment Monitoring System, United States, 40 Sites, 2000–2010), about 10 percent of women admitted to smoking tobacco during the third trimester. Smoking during pregnancy increases the mother's chance of miscarriage, premature birth, low birth weight, and sudden infant death syndrome. It is also associated with cognitive, language, and attention deficits, behavioral and psychiatric disorders, and an increased likelihood of drug addiction later in life. Nicotine, the primary addictive substance in tobacco smoke, acts as an agonist on nicotinic cholinergic receptors (nAChRs). Since nAChRs are expressed by brainstem dopamine neurons, exposure to nicotine during development might lead to permanent alterations in the mesolimbic pathway that mediates reward.

Table 1.3–1.
Neurodevelopmental Vulnerability to Methylmercury in P7 Rats

Time post-exposure	24 hours	2 weeks	4 weeks
	Neural stem cells undergo cell cycle arrest at G1/S transition Neural stem cells undergo cell	Smaller DG volume and total cell number Reduction of intermediate progenitor cells in S-phase	Hippocampal-dependent spatial learning and memory deficits

	death via intrinsic mitochondrial apoptosis	and newborn neurons

Short- and long-term effects of P7 MeHg exposure on hippocampal neurogenesis and function in the rat. In the short-term (24 h), MeHg causes cell cycle arrest and apoptosis of hippocampal neural stem cells. The cell deficit resulting from this early insult is not fully rectified by adolescence (P21); evidenced by a smaller dentate gyrus (due to lower cell number), and fewer S-phase intermediate progenitors and newborn neurons. By 4 weeks (P35), animals display deficits in hippocampal-dependent spatial learning and memory tasks.

Cannabis (from the *Cannabis sativa* plant) is the most commonly used illicit drug among pregnant women. Though the main psychoactive ingredient, δ-9-tetrahydrocannabinol (THC), readily crosses the placenta, less appears to enter fetal blood compared to other drugs. In addition, studies show that in many species, fetal THC concentrations tend to be lower than maternal concentrations. However, cannabis exposure does impair blood flow to the fetus, which in turn can have deleterious effects on brain development. THC mimics the endocannabinoid neurotransmitter, anandamide, by targeting cannabinoid G-protein–coupled receptors (GPCR) type 1 and 2 (CB_1R, CB_2R). CB_1R in particular is expressed as early as the pre-implantation period of gestation. During development, CB_1R activity and synthesis (and secretion) of another endocannabinoid 2-arachidonoylglycerol (2-AG) occurs in the growth cones of developing neurons. When the growth cone reaches its final destination for synapse formation, 2-AG is naturally degraded by monoacylglycerol lipase (MAGL). However, MAGL cannot degrade THC, thus exposure to THC during neuronal development can disturb axon growth and synapse formation. Longitudinal epidemiological and animal studies have found that prenatal cannabis exposure is associated with lower birth weight, learning and memory deficits, and increased anxiety.

Cocaine is the third most commonly used illicit drug in the United States. It easily crosses the placenta and BBB and produces severe deleterious effects on the developing brain. Like METH, it primarily affects monoamine neurotransmitter systems, by serving as a triple monoamine (dopamine, serotonin, norepinephrine) reuptake inhibitor. This has a severe effect on cortical neuronal development that depends on the neurotrophic functions of monoamines, and the mesolimbic reward pathway. Brain regions that regulate executive function and attention appear to be the most vulnerable to prenatal cocaine exposure. Exposed children are born with lower IQ, learning, language (mild), and memory deficits.

POSTNATAL NEUROGENESIS IN MAMMALS: SUBVENTRICULAR ZONE AND HIPPOCAMPUS

For the past few decades, there has been a fundamental shift in paradigm

regarding the limits of neurogenesis in the brain, with important implications for neural plasticity, mechanisms of disease etiology and therapy, and possibilities of repair. For most of the 20th century, the notion was that humans do not produce new neurons in the brain after birth (or soon thereafter, considering cerebellar EGL), thus brain plasticity and repair depends on modifications of a numerically static neural network. The contrary principle (with strong evidence from numerous studies) that new neurons are generated throughout life in certain regions is now accepted. This concept is well documented across the phylogenetic tree, including birds, rodents, primates, and humans. As an area of intense interest and investigation, psychiatrists may expect rapid progress over the next two decades, likely altering models described herein.

In this context, neurogenesis is the generation of new neurons from neural precursors. Scientists who observed ^{3}H-thymidine labeled neurons (a common marker for cell proliferation) in tissue reported the first evidence of mammalian neurogenesis in the adult hippocampus in the 1960s. Cell proliferation is defined as an absolute increase in cell number, which occurs only if cell production is not balanced by cell death. More recent studies of neurogenesis employ the more convenient thymidine analogs BrdU and 5-ethynyl-2′-deoxyuridine (EdU), which can be injected into living animals and then detected by immunohistochemistry.

During embryonic development, neurons are produced from almost all regions of the ventricular neuroepithelium. However, postnatal (from birth to puberty) and adult neurogenesis are largely restricted to two regions: the SVZ lining the lateral ventricles, and a narrow proliferative zone underlying the DG granule layer SGZ in the hippocampus. In mice, rodents, and monkeys, newly produced neurons migrate from the SVZ in an anterior direction toward the olfactory bulb to become GABA (and dopamine)--containing interneurons. The process has been elegantly characterized at both ultrastructural and molecular levels (Table 1.3–1). In the SVZ, the neuroblasts (A cells) on their way to the olfactory bulb create chains of cells, and migrate through a scaffold of glial cells supplied by slowly dividing astrocytes (B cells). Within this network of cell chains, there are groups of rapidly dividing neural precursors (C cells). Evidence suggests that the B cells give rise to the C cells which later develop into the A cells, the future olfactory bulb interneurons. The existence of a sequence of precursors with progressively restricted abilities to generate diverse neural cell types makes defining mechanisms regulating adult neurogenesis in vivo a great challenge.

Studies in rodents show that adult neurogenesis is regulated by extracellular signals (such as EGF, bFGF, IGF-I, BDNF, LIF/CNTF) that control precursor proliferation and survival. In turn, precursor proliferation and survival also modulate expression and activity of these factors. While the hallmark of NSCs includes the capacity to generate

neurons, astrocytes, and oligodendroglia, termed multi-potentiality, specific signals appear to produce relatively different profiles of cells that may migrate to distinct sites. Intraventricular infusion of EGF promotes primarily gliogenesis in the SVZ, with cells migrating to olfactory bulb, striatum, and corpus callosum, whereas bFGF favors generation of neurons destined for the olfactory bulb. Both factors appear to stimulate mitosis directly, with differential effects on the cell lineage produced. In contrast, BDNF may increase neuron formation in SVZ as well as striatum and hypothalamus, though effects may be primarily through promoting survival of newly generated neurons that otherwise undergo cell death. Finally, CNTF and related LIF may promote gliogenesis, or alternatively, support self-renewal of adult stem cells rather than enhancing a specific cell category.

In addition to direct intraventricular infusions, peripheral levels of growth factors, hormones, and neuropeptides also affect adult neurogenesis. Peripheral administration of both bFGF and IGF-I stimulate neurogenesis, selectively increasing mitotic labeling in the SVZ and hippocampal SGZ, respectively, suggesting there are specific mechanisms for factor transport across the BBB. Interestingly, elevated prolactin levels, induced by peripheral injection or natural pregnancy, stimulate proliferation of progenitors in the mouse SVZ (Fig. 1.3–11), leading to increased olfactory bulb interneurons, potentially playing roles in maternal learning of the new infant's scents. This may be relevant to changes in prolactin seen in psychiatric disease. Conversely, in behavioral paradigms of social stress, such as territorial challenge by male intruders, activation of the hypothalamic–pituitary–adrenal (HPA) axis with increased glucocorticoids leads to reduced neurogenesis in the hippocampus. Inhibition is also observed after peripheral opiate administration, a model for substance abuse. Thus, neurogenesis may be one target process affected by changes of hormones and neuropeptides associated with several psychiatric conditions.

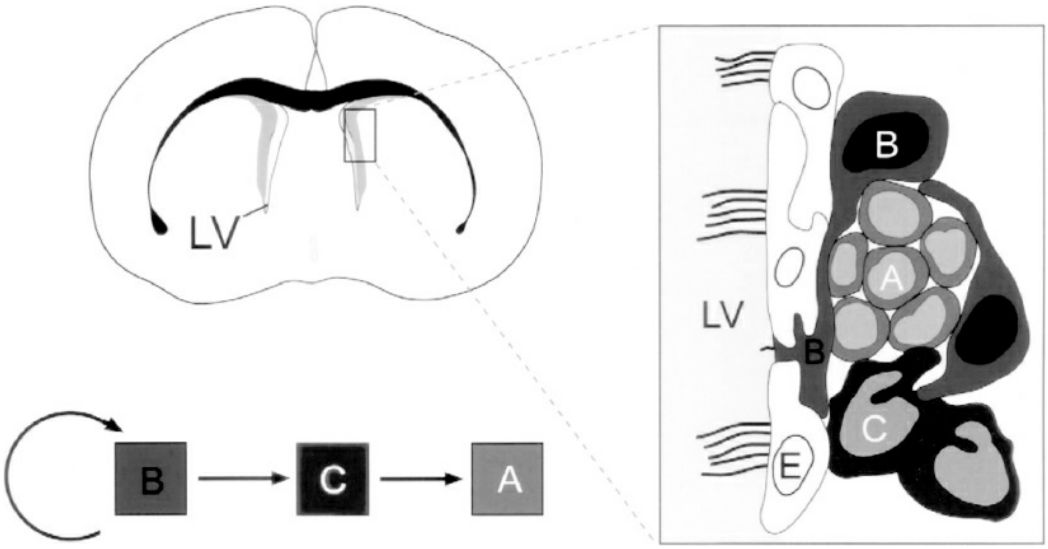

FIGURE 1.3–11. Adult neural stem cells localize to the lateral ventricular wall. This drawing shows a cross-section of the adult mouse brain with the boxed area representing an enlargement of the subventricular zone, based on electron microscopic ultrastructural studies. Ciliated ependymal cells (E) line the lateral ventricles (LVs), and behind this lining, astrocytes (B) can be found. These glial cells give rise to dividing precursor cells (C), which in turn generate the neuroblasts (A). The neuroblasts migrate to the olfactory bulb by forming chains of cells within glial tunnels composed of astrocytes. The B cell is considered a stem cell that renews itself on each division, indicated by the circular arrow, as well as gives rise to dividing precursors fated to become neurons. (From Alvarez-Buylla A, Seri B, Doetsch F. Identification of neural stem cells in the adult vertebrate brain. *Brain Res Bull*. 2002;57:751, with permission.)

The discovery of adult neurogenesis naturally leads to questions about whether new neurons can integrate into the complex cytoarchitecture of the mature brain, and the functional significance. In rodents, primates, and humans, new neurons are generated in the DG of the hippocampus, an area important for learning and memory. Some adult-generated neurons in humans have been shown to survive for at least 2 years. Further, newly generated cells in adult mouse hippocampus have extensive dendritic and axonal arborizations appropriate to the neural circuit, and display functional synaptic inputs and action potentials. From a functional perspective, the generation and/or survival of new neurons correlate strongly with multiple instances of behavioral learning and experience. For example, survival of newly generated neurons is markedly enhanced by hippocampal-dependent learning tasks and by an enriched, behaviorally complex environment. Of perhaps greater importance, a reduction in DG neurogenesis impairs the formation of trace memories, that is, when an animal must associate stimuli that are separated in time, a hippocampal-dependent task. Finally, in songbirds, neurogenesis is activity-dependent, and is increased by foraging for food and learning new song, whether it occurs seasonally or is induced by steroid hormone administration.

From clinical and therapeutic perspectives, scientists have observed that changes in neurogenesis can contribute to disease or help the brain

recover from injury. A neurogenetic response has been shown for multiple pathological conditions in the adult, including brain trauma, stroke, and epilepsy. For instance, ischemic stroke in the striatum stimulates adjacent SVZ neurogenesis (Fig. 1.3–11), with neurons migrating to the injury site. Furthermore, in a highly selective paradigm not involving local tissue damage, degeneration of cerebral cortical layer III pyramidal neurons elicited SVZ neurogenesis and cell replacement. These studies raise the possibility that newly produced neurons normally participate in recovery, and may be stimulated as a novel therapeutic strategy. However, neurogenesis can also play roles in pathogenesis. In a kindling model of epilepsy, newly generated neurons migrated to incorrect positions and formed aberrant neuronal circuits, reenforcing the epileptic state. Conversely, reductions in neurogenesis can contribute to several conditions that implicate dysfunction or degeneration of the hippocampal formation. DG neurogenesis is inhibited by increased glucocorticoids levels observed in aged rats and is reversed by steroid antagonists and adrenalectomy. These observations are potentially relevant to the correlation of elevated human cortisol levels with reduced hippocampal volumes and the presence of memory deficits. Similarly, stress-induced increases in human glucocorticoids may contribute to decreased hippocampal volumes seen in schizophrenia, depression, and posttraumatic stress disorder (PTSD).

Studies in the area of postnatal and adult human neurogenesis are still nascent, but recent findings suggest that the magnitude of new neuron generation in the human DG is comparable to that in rodents. In adult rats, monkeys, and humans, the number of new hippocampal neurons generated daily amounts to approximately 0.2 percent, 0.02 percent, and 0.004 percent of the total granule cell population, respectively. However, there are some differences between rodent and human neurogenesis. Unlike rodents, humans do not appear to have evident neurogenesis in the olfactory bulb, and age-dependent decline of neurogenesis in humans does not follow the sharp decline that is normally seen in rodents.

Hippocampal neurogenesis seems to provide an avenue for incorporating new memories and information. Thus, a balance is made between adding new information and retaining previously stored information. In the hippocampus, new granule neurons integrate into the existing neuronal network, forming new connections that might replace or cooperate with older, established synaptic connections. If the established connections are replaced, memories are forgotten because the previously stored information is lost. Sigmund Freud coined the term “infantile amnesia” to describe the phenomenon of the inability of adult humans to recall declarative (facts and events) memories from their first few years of life, followed by patchy memory recall until the age of 7. Recent studies are testing the neurogenic hypothesis of infantile amnesia, which states that high levels of postnatal hippocampal neurogenesis lead to the replacement

of preexisting synaptic connections, which inhibits the ability to form long-lasting memories. The age-related decline in neurogenesis rate results in lower turnover of established synaptic connections, and is therefore accompanied by a higher ability to form long-lasting memories. Overall, findings in rodents suggest that increasing neurogenesis after the acquisition of memories causes forgetting, while genetically or pharmacologically reducing neurogenesis reduces forgetting.

Defining the effects of hippocampal postnatal neurogenesis on memory can also provide clues as to why memory formation and retention changes with age. Although the human brain is fully developed by birth, the DG of the hippocampus undergoes significant development postnatally. Thus, it is not surprising that the ability to retain declarative memories can be affected by postnatal neurogenesis. Precocial animals such as the degu and guinea pig tend to have low levels of postnatal neurogenesis compared to nonprecocial animals, because the majority of granule cells are generated prenatally. Thus, nearly complete DG development in precocial animals by birth allows for superior memory retention during infancy. Studies comparing memory retention in precocial and nonprecocial animals (i.e., rat) have shown that retention of memories (acquired via learning a place discrimination task) in guinea pigs was not age-dependent, whereas in rats, older animals remembered the task better than those trained at a younger age. However, infantile amnesia can be induced experimentally in precocial animals by genetically or pharmacologically increasing levels of hippocampal neurogenesis. In humans, the age-dependent decline in hippocampal neurogenesis might reflect an intricate balance between encoding and forgetting. In early life, the high neurogenesis rate favors forgetting in order to retain only the most important memories in a highly demanding sensory environment.

New hippocampal neurons (and stem and progenitor cells) that are generated during adulthood are implicated in response to treatments for mood disorders such as depression. Indeed, so much research on the effects to adult neurogenesis of antidepressant treatment has occurred in the past decade that the neurogenic hypothesis of depression was formed to link the research together. It states that antidepressant efficacy and mood control depend on the continued generation of new neurons in the adult hippocampus. So far, there have been no human studies that have tested the contribution of adult hippocampal neurogenesis to antidepressant efficacy, but recent research reveals similarities with findings from rodent depression model studies that show correlations between antidepressant treatment and enhanced hippocampal neurogenesis. In postmortem and MRI studies in patients with depression, researchers observed a smaller DG and lower granule cell numbers compared to healthy controls, which may be as a result of reduced neurogenesis. Furthermore, correlative increases in anxiety and depression have been observed in patients

undergoing cancer treatments such as irradiation, which is known to reduce hippocampal neurogenesis. Conversely, antidepressant treatment in patients appears to cause the hippocampus to grow larger (defined through postmortem and MRI studies), leading to increased DG size and granule cell number. Similar findings have been observed in NHPs, where antidepressant treatment restored normal neurogenesis levels in stressed monkeys that exhibited clinical responses. It is thought that mechanisms of antidepressant actions are similar to those mediating stress-coping responses.

Current studies demonstrate that adult neurogenesis in the SVZ has potential links to the development of neurodegenerative and psychiatric disease. Recently, the fate of neuroblasts generated from NSCs of the human SVZ (from studies in postmortem human brain tissue) was discovered. In infant brains, researchers found migrating immature neurons in the SVZ and rostral migratory stream to the olfactory bulb. Proliferating cells were also found in the human SVZ, but their numbers (as well as the migrating neuroblasts) decreased drastically within the first 18 months of life. By adulthood, very few migrating neuroblasts were found in these same regions. The mechanisms guiding this drastic decline are not known, but it can be concluded that neurogenesis in the human SVZ undergoes dramatic changes throughout the life span. More recently, the medial migratory system (MMS) was discovered only in 4- to 6-month human brain tissue, where many neuroblasts (some expressing interneuron markers tyrosine hydroxylase and calretinin) migrate to the VMPFC. This discovery of SVZ neurogenesis to the prefrontal cortex might provide clues to the etiology of mental disorders such as schizophrenia, especially if neurogenesis is altered. What is not yet known is the function of these neurons in the prefrontal cortex. A recent study observed neural precursors (which generate interneurons) in the SVZ that borders the striatum, and their loss is implicated in the development of Huntington disease. Techniques such as nuclear proton nuclear magnetic resonance spectroscopy (1H-NMR and 1H-MRS) are currently being used to visualize human neural precursor cells (hNPCs) live and in vivo by identifying hNPC-specific metabolites. In one study, a 1.28-ppm biomarker on the 1H-NMR spectra was established to be NPC-specific. Thus, peaks at this spectrum were detected live in neurogenic human brain regions such as the hippocampus but not in the cortex. Not surprisingly, the amplitude of these peaks decreased with the age (prepubescent, adolescent, adult) of the brain. Though quality control considerations are yet to be ironed out, this technology shows great promise for elucidating alterations in neurogenesis for patients with neurodevelopmental, psychiatric, and neurodegenerative disorders.

More generally, the discovery of adult neurogenesis has led to major changes in psychiatrists' perspectives on the regenerative capacities of the

human brain.

REFERENCES

Aimone JB, Li Y, Lee SW, Clemenson GD, Deng W, Gage FH. Regulation and function of adult neurogenesis: From genes to cognition. *Physiological Reviews*. 2014;94:991–1026.

Alvarez-Buylla A, Seri B, Doetsch F. Identification of neural stem cells in the adult vertebrate brain. *Brain Res Bull*. 2002;57:751–758.

Amaral DG, Schumann CM, Nordahl CW. Neuroanatomy of autism. *Trends in Neurosci*. 2008;31:137–145.

Behnke M, Smith VC, Committee on Substance Abuse; Committee on Fetus and Newborn. Prenatal substance abuse: Short- and long-term effects on the exposed fetus. *Pediatrics*. 2013;131:e1009–e1024.

Bishop KM, Goudreau G, O'Leary DD. Regulation of area identity in the mammalian neocortex by Emx2 and Pax6. *Science*. 2000;288:344–349.

Cheng Y, Black IB, DiCicco-Bloom E. Hippocampal granule neuron production and population size are regulated by levels of bFGF. *Eur J Neurosci*. 2002;15:3–12.

DiCicco-Bloom E, Lord C, Zwaigenbaum L, et al. The developmental neurobiology of autism spectrum disorder. *J Neurosci*. 2006;26:6897–6906.

Eriksson PS, Perfilieva E, Björk-Eriksson T, et al. Neurogenesis in adult human hippocampus. *Nat Med*. 1998;4:1313–1317.

Ernst A, Alkass K, Bernard S, et al. Neurogenesis in the striatum of the adult human brain. *Cell*. 2014;156:1072–1083.

Falluel-Morel A, Sokolowski K, Sisti HM, et al. Developmental mercury exposure elicits acute hippocampal cell death, reductions in neurogenesis, and severe learning deficits during puberty. *Journal Neurochem*. 2007;103:1968–1981.

Fukuchi-Shimogori T, Grove EA. Neocortex patterning by the secreted signaling molecule FGF8. *Science*. 2001;294:1071–1074.

Guillemot F. Cell fate specification in the mammalian telencephalon. *Prog Neurobiol*. 2007;83:37–52.

Hansen DV, Lui JH, Parker PR, Kriegstein AR. Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature*. 2010;464:554–561.

Heckers S, Konradi C. Hippocampal neurons in schizophrenia. *J Neural Transm*. 2002;109:891–905.

Hevner RF. Layer-specific markers as probes for neuron type identity in human neocortex and malformations of cortical development. *J Neuropathol Exp Neurol*. 2007;66:101–109.

Josselyn SA, Frankland PW. Infantile amnesia: a neurogenic hypothesis. *Cold Spring Harbor Laboratory Press*. 2012;19:423–433.

Kempermann G, Gage FH. Neurogenesis in the adult hippocampus. *Novartis Found Symp*. 2000;231:220–235.

Kohlmeier KA. Nicotine during pregnancy: changes induced in neurotransmission, which could heighten proclivity to addict and induce maladaptive control of attention. *Journal of Dev Origins of Health and Disease*. 2014;11:1.

LaMonica BE, Lui JH, Wang X, Kriegstein AR. OSVZ progenitors in the human cortex: an

updated perspective on neurodevelopmental disease. *Curr Opin Neurobiol.* 2012;5:747–753.

Lewis DA. Inhibitory neurons in human cortical circuits: substrate for cognitive dysfunction in schizophrenia. *Current Opinion in Neurobiol.* 2014;26:22–26.

Lewitus E, Kelava I, Huttner WB. Conical expansion of the outer subventricular zone and the role of neocortical folding in evolution and development. *Frontiers in Human Neurosci.* 2013;7:424.

Mairet-Coello G, Tury A, Van Buskirk E, Robinson K, Genestine M, DiCicco-Bloom E. $p57^{KIP2}$ regulates radial glia and intermediate precursor cell cycle dynamics and lower layer neurogenesis in developing cerebral cortex. *Development.* 2012;139:475.

Miller BR, Hen R. The current state of the neurogenic theory of depression and anxiety. *Current Opinion in Neurobiol.* 2015;30:51–58.

Monuki ES, Walsh CA. Mechanisms of cerebral cortical patterning in mice and humans. *Nat Neurosci.* 2001;4:1199–1206.

MuhChyi C, Juliandi B, Matsuda T, Nakashima K. Epigenetic regulation of neural stem cell fate during corticogenesis. *International Journal of Dev Neurosci.* 2013;31:424–433.

Nadarajah B, Parnavelas JG. Modes of neuronal migration in the developing cerebral cortex. *Nat Neurosci.* 2002;3:423–432.

Noctor SC, Flint AC, Weissman TA, Dammerman RS, Kriegstein AR. Neurons derived from radial glial cells establish radial units in neocortex. *Nature.* 2001;409:714–720.

O'Leary DD, Sahara S. Genetic regulation of arealization of the neocortex. *Curr Opin Neurobiol.* 2008;1:90–100.

Reynolds BA, Weiss S. Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science.* 1992;255:1707–1710.

Ross CA, Margolis RL, Reading SAJ, Plentikof M, Coyle JT. Neurobiology of schizophrenia. *Neuron.* 2006;52:139–153.

Shors TJ, Miesegaes G, Beylin A, Zhao M, Rydel T, Gould E. Neurogenesis in the adult is involved in the formation of trace memories. *Nature.* 2001;410:372–376.

Sokolowski K, Obiorah M, Robinson K, McCandlish E, Buckley B, DiCicco-Bloom E. Neural stem cell apoptosis after low methylmercury (MeHg) exposures in postnatal hippocampus produce persistent cell loss and adolescent memory deficits. *Dev Neurobiol.* 2013;73:936–949.

Vaccarino FM, Schwartz ML, Raballo R, et al. Changes in cerebral cortex size are governed by fibroblast growth factor during embryogenesis. *Nat Neurosci.* 1999;2:848.

van Praag H, Schinder AF, Christie BR, Toni N, Palmer TD, Gage FH. Functional neurogenesis in the adult hippocampus. *Nature.* 2002;415:1030.

▲ 1.4 Biogenic Amine Neurotransmitters

JOHN D. ELSWORTH, PH.D., AND ROBERT H. ROTH, PH.D.

INTRODUCTION

The biogenic amine neurotransmitters are the catecholamines (dopamine, norepinephrine, and epinephrine) and serotonin (5-hydroxytryptamine,

5HT), histamine, and acetylcholine. The importance of biogenic amines is demonstrated by their implication in the pathophysiology and treatment of a wide variety of psychiatric and neurological disorders. These neurotransmitters have an unusual but functionally significant organization in the CNS. Their cell bodies are restricted to a small number of nuclei in the midbrain, brain stem, basal forebrain, and hypothalamus; yet their axons project throughout the whole CNS, some more topographically than others. It is well known that these systems play critical roles in cognition, memory, motivation, and motor function and arousal.

This widely projecting influence permits each of these neurotransmitters to modulate activity in diverse circuits throughout the CNS, sometimes in a coordinated fashion. This complex anatomical organization poses a major challenge to understand the precise cellular, molecular, and system level pathways by which these neurotransmitters impact neuropsychiatric disorders. It is also important to realize that a pharmacological treatment can be clinically effective even though it may not directly target the specific biological deficit causing the symptoms.

Recent advances in human genetics and genomics, as well as experimental neuroscience, have revealed some insights into the complexity of these neurotransmitter systems. Molecular cloning has identified a large number of genes that regulate monoaminergic neurotransmission, such as the enzymes, receptors, and transporters that mediate the synthesis, cellular actions, and cellular reuptake of these neurotransmitters, respectively. Human genetics studies have provided evidence of tantalizing links between allelic variants in specific monoamine-related genes and psychiatric disorders and trait abnormalities, while the ability to modify gene function and cellular activity in experimental animals has clarified the roles of specific genes and neural pathways in mediating behavioral processes.

In this chapter, highlight for each of the biogenic amines the details of its life cycle (incorporating synthetic and metabolic enzymes and transporters) are highlighted together with their receptors. Ultimately, the effects of biogenic amines on CNS function and behavior depend upon their interactions with receptor molecules. The binding of biogenic amines to these plasma membrane proteins initiates a series of intracellular events that modulate neuronal excitability. Unlike the transporters, multiple receptor subtypes exist for each biogenic amine transmitter. Neurotransmitter receptors produce intracellular effects by one of two basic mechanisms: (1) by interactions with G proteins that couple receptors to intracellular effector systems and (2) by providing channels through which ions flow when transmitters bind (ligand-gated ion channels). The neurotransmitters receptors covered in this chapter with the exception of the serotonin 5-HT_3 receptor subtype and the cholinergic nicotinic

receptors all belong to the superfamily of GPCRs. However, within each neurotransmitter receptor group, the subtypes are heterogeneous with regard to the G proteins with which they interact and to the second messenger effects that they produce. Receptors are also diverse in their regional patterns of expression within the brain, their neurotransmitter binding affinities, and their synaptic localization. Whereas many receptor subtypes are located exclusively in postsynaptic membranes, others are located presynaptically and some both pre- and postsynaptically. Some receptors on the presynaptic terminal respond to a transmitter that is released by that neuron; these presynaptic "*autoreceptors*" often act to inhibit neurotransmitter release. When a receptor for a particular transmitter is located on terminals of a different transmitter it is a "*heteroreceptor,*" and its activation can increase or decrease release.

Much recent effort has been focused on determining the functional roles of individual neurotransmitter receptor subtypes. Limited availability of selective agonist and antagonist drugs hampers this effort, but recent advances in gene manipulation and targeted gene editing is advancing the fields of neuroscience and pharmacology considerably. Such studies in animals are providing clues to receptor function and to the contributions of each receptor to the actions of nonspecific drugs. It is anticipated that the next generation of subtype-selective compounds will lead to novel therapeutic agents that alter biogenic amine transmission in a more refined manner.

DOPAMINE AND NOREPINEPHRINE

Dopamine Anatomy and Functions

The majority of dopaminergic innervation in brain is provided by long projection neurons arising from two clusters of cell bodies in the midbrain (mesencephalon), the substantia nigra, and the VTA. Neurons from the substantia nigra provide a dense input to the dorsal striatum (caudate nucleus and putamen), while projections from the VTA principally target limbic regions (e.g., nucleus accumbens and septum) in addition to cortical regions (e.g., prefrontal cortex and cingulate cortex). It should also be noted though, that it is not possible to completely segregate the motor, motivational, and cognitive behaviors that are often attributed to the striatal, limbic, and cortical dopamine systems, respectively.

Dysfunction in striatal dopamine transmission is strongly associated with movement disorders such as Parkinson disease or tardive dyskinesia. The hallmark signs of Parkinson disease originate from a loss of striatal dopaminergic innervation. Tardive dyskinesia can occur following high-dose or long-term exposure to certain antipsychotic drugs that block dopamine receptors. It is apparent that nigrostriatal dopamine neurons do not merely allow motor behavior to occur, but that they also play an

important role in the selection and initiation of actions and establishing motor skills and habits. Nigrostriatal dopamine neurons display a uniquely complex arborization of terminals. This large architecture together with the unmyelinated nature of its axons puts the nigrostriatal dopamine neuron under a phenomenal energy demand and probably is a critical factor in their particular susceptibility to a variety of biochemical and pharmacological insults.

Dopamine neurons arising in the VTA and terminating in the frontal cortex play an important role in attention, working memory. Dopamine neurons innervating limbic regions such as nucleus accumbens are strongly implicated in goal-oriented (motivated) behaviors, in addition to reward, attention, and pharmacologically induced locomotion. Enhancement of dopamine transmission in mesolimbic system has been linked with the addicting, reinforcing, and sensitizing effects of repeated exposure to psychostimulant drugs of abuse. Tourette syndrome and ADHD have been linked to a dysfunction of dopaminergic neurotransmission in limbic and cortical regions, based mainly on the success of dopaminergic drugs in treatment of their symptoms. The oldest and most established hypothesis of schizophrenia posits that the disease is related to excessive central dopamine activity in striatal and limbic regions. A revision to the hypothesis is the concept of reduced cortical dopamine transmission together with overactivity in subcortical dopamine systems.

Anterior to the mesencephalic substantia nigra and VTA, are several other relatively small groups of dopaminergic cell bodies. The most prominent of these are the tuberoinfundibular dopamine neurons, which are located in arcuate nucleus of the hypothalamus and project to the median eminence. Dopamine released from these neurons enters the circulation and is transported to the anterior pituitary where it inhibits the release of prolactin from the anterior lobe of the pituitary. Antipsychotic drugs that block dopamine receptors in the pituitary induce an increase in blood prolactin levels.

Electrophysiological studies have shown that dopamine cells fire in two different modes, either single spiking or burst firing, presumably as a result of a change in balance between excitatory and inhibitory inputs. As cells are able to switch between these levels, the transition in activity may be a mechanism for altering the impact of dopamine neurotransmission on receptive cells. Bursting activity elicits more terminal dopamine release per action potential than pacemaker activity, and bursting appears to elicit release of peptide co-transmitters more readily than nonbursting patterns. The different rates of firing of dopamine neurons probably code different responses in the brain. At the fastest scale, with burst firing, dopamine neurons can signal a reward, or to indicate to what extent a reward occurs differently than predicted. At the slowest time scale, the pacemaker activity of dopamine neurons has been linked to a tonic enabling function on a

variety of motor, cognitive, and motivational processes, which are deficient in Parkinson disease and which can be restored by dopamine replacement therapy.

In addition to the DA systems in the brain and pituitary, dopamine appears to have a restricted role in the periphery. The best-defined role for peripheral dopamine neurons is in renal regulation of sodium homeostasis, and impairment of this control may underlie development of hypertension in some cases. Another important function of dopamine outside the brain is in various retinal processes, including light adaptation and contrast sensitivity.

Norepinephrine Anatomy and Functions

The widespread norepinephrine ("*noradrenergic*") innervation in brain mainly derives from a very small cluster of neurons located in the locus coeruleus (LC, part of the pons) in addition to a more diffuse grouping in the lateral tegmental area (part of the midbrain). These two groups provide input to virtually all brain regions, with LC neurons primarily targeting the forebrain, thalamus, cerebellum, and spinal cord and the lateral tegmental neurons principally sending fibers to the hypothalamus and basal forebrain. A single norepinephrine neuron typically innervates several brain regions. LC norepinephrine neurons exhibit different activity modes: tonic firing with short periods of burst firing. Alterations in the tonic mode are associated with the sleep–waking cycle, whereas the phasic mode occurs in response to arousing stimuli. There is good evidence for an important role of LC norepinephrine neurons in vigilance, attention, and working memory.

Epinephrine Anatomy and Functions

In the sympathetic nervous system (ANS) and the adrenal medulla, epinephrine shares with norepinephrine the role of final neurotransmitter, and the proportion of this sharing is species-dependent and hormonally modified. Relatively few epinephrine cells exist in brain and are found in the medulla, providing innervation to subregions of hypothalamus.

Life Cycle of Dopamine, Norepinephrine and Epinephrine

Synthesis. These three catecholamine neurotransmitters share a common synthetic pathway (Fig. 1.4–1). Blood-borne tyrosine, derived from dietary proteins and from phenylalanine metabolism, enters the brain by the large neutral amino acid (LNAA) transport system. Tyrosine in brain extracellular fluid is taken up into catecholamine neurons by a high affinity amino acid transporter. In the first reaction tyrosine is converted to L-dopa by the enzyme tyrosine hydroxylase in the cytosol. Because this step is rate limiting, it sets the pace for the conversion of tyrosine to dopamine, which

makes this step the most susceptible to physiologic regulation and pharmacologic manipulation. Tyrosine hydroxylase requires tetrahydrobiopterin as a cofactor, and the synthesis of tetrahydrobiopterin is itself dependent on the activity of another enzyme, GTP-cyclohydrolase-1. The next reaction is the decarboxylation of L-dopa by aromatic amino acid decarboxylase (AADC) in the cytosol to form dopamine. AADC decarboxylates L-dopa so avidly that the level of this amino acid in brain is very low under normal conditions. L-dopa is the principal drug used to treat Parkinson disease as it restores the deficient striatal dopamine level that occurs in the disorder. Supplying exogenous L-dopa to the neuron ensures efficient conversion to dopamine because it bypasses tyrosine hydroxylase, the rate-limiting step in dopamine biosynthesis. An inhibitor of AADC that does not cross the BBB (e.g., carbidopa or benserazide) is routinely administered with levodopa to patients with Parkinson disease to inhibit the peripheral decarboxylation of L-dopa to dopamine, which reduces peripheral side effects resulting from abnormally high circulating levels of dopamine. One prominent side-effect of excessive circulating dopamine is emesis due to the activation of dopamine receptors in the area postrema, an area outside the BBB. In fact, D_2 antagonists (e.g., promethazine [Phenergan]) are a widely used class of antiemetic agents. However, because they are prone to induce parkinsonian-like side effects, alternatives such as $5HT_3$ antagonists (e.g., ondansetron [Zofran]) and neurokinin-1 receptor antagonists (e.g., aprepitant [Emend]) are increasingly used to avoid these problems associated with D_2 antagonism.

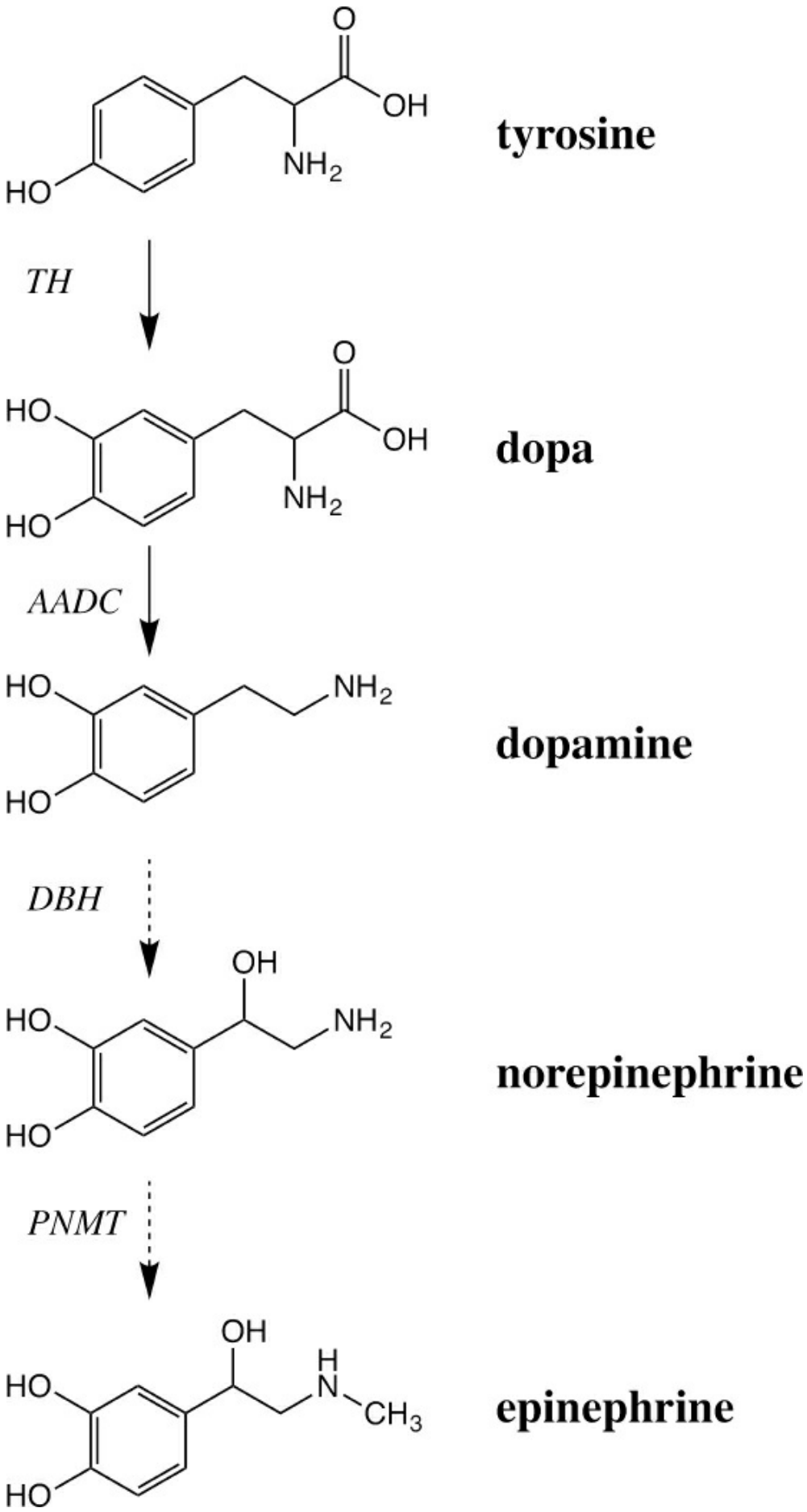
O
OH
NH2
HO
tyrosine
TH
O
HO
OH
NH2
HO
dopa
AADC
HO
NH2
HO
dopamine
DBH
OH
HO
NH2
HO
norepinephrine
PNMT
OH
H
N
CH3
HO
HO
epinephrine

FIGURE 1.4–1. Biosynthesis of dopamine, norepinephrine, and epinephrine. The catecholamine neurotransmitters are produced by a multistep synthesis that starts with the dietary amino acid, tyrosine. Dopamine neurons contain the enzymes TH and AADC, whereas norepinephrine neurons also express DBH. PNMT is an additional enzyme that is present only in epinephrine neurons. AADC, aromatic amino acid decarboxylase; DBH, dopamine-β-hydroxylase; PNMT, phenol-N-methyltransferase.

In norepinephrine neurons, as well as in chromaffin cells of the adrenal medulla, an additional enzyme, dopamine-beta-hydroxylase (DBH) is present that enables the conversion of dopamine to norepinephrine. This enzyme is concentrated inside the storage vesicles, so this step of norepinephrine synthesis does not occur in the cytoplasm. Disulfiram (Antabuse), a drug used to treat alcohol addiction, is an inhibitor of DBH. However, the principal relevant property of disulfiram in treating alcoholism is to block acetaldehyde dehydrogenase, the enzyme that metabolizes acetaldehyde formed from ethanol, and this results in elevated acetaldehyde levels that causes moderate to severe adverse effects after consumption of ethanol.

In central epinephrine neurons and in adrenal medulla and heart, norepinephrine is N-methylated by the enzyme phenylethanolamine-N-methyltransferase (PNMT) to form epinephrine. As PNMT is located in the cytoplasm, norepinephrine has to exit its storage site in vesicles to be converted to epinephrine before being transported back into vesicles. This reaction presumably takes place during spontaneous leakage of norepinephrine from vesicles in cytoplasm, or after norepinephrine has been released into the synapse and taken back up into the cytoplasm and exposed to PNMT.

Phenylketonuria (PKU) is a rare genetic disease caused by a mutation in the enzyme phenylalanine hydroxylase, which hampers synthesis of tyrosine as well as disturbing other metabolic pathways of phenylalanine. The elevated concentrations of phenylalanine seen in PKU has devastating effects on the developing brain typically leading to severe mental retardation unless steps are taken to limit dietary intake of phenylalanine. Newborns identified with such mutation can avoid symptoms of PKU by eliminating phenylalanine from their diet.

Vesicular Transporter. Vesicles present in the presynaptic terminals are specialized for the uptake and storage of catecholamines, thus protecting the transmitter from degradation by the enzyme monoamine oxidase (MAO) found in mitochondria. The accumulation and concentration of transmitter inside the vesicles is accomplished by vesicular monoamine transporter (VMAT). There are two forms of VMAT and it is VMAT-2 that is expressed in central neurons.

Dopamine Plasma Membrane Transporter. A high-affinity DAT serves to

recapture released dopamine and limits the concentration and actions of dopamine in the synapse (Fig. 1.4–2). Importantly, dopamine neurons in striatum and nucleus accumbens are equipped with a high density of DATs and so are especially susceptible to drugs that target this site. For example, blockade of dopamine re-uptake by cocaine in the nucleus accumbens, prolongs and intensifies dopamine's actions in this region and is key to the euphoric effects of the drug. Similarly, amphetamine and methamphetamine amplify dopamine signaling in the nucleus accumbens by blocking dopamine reuptake and increasing release of dopamine. The VTA-nucleus accumbens pathway is part of the brain's reward pathway stimulated by all types of reinforcing stimuli such as food, sex, and many drugs of abuse. While stimulants can be addictive, some (e.g., methylphenidate, amphetamine) are used at lower doses to treat ADHD, and probably achieve their beneficial effects by actions on the dopamine and norepinephrine transporters (NETs). Lisdexamfetamine (Vyvanse) is a long-acting oral prodrug that is hydrolyzed in the blood to active D-amphetamine, and it has been approved for treating ADHD. The DAT is also the pharmacological target for the parkinsonian neurotoxins, 6-hydroxydopamine and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP). In fact the toxicity of MPTP is dependent on the combined function of MAO-B, DAT, and VMAT-2. Once inside the brain, MPTP is metabolized by MAO-B to the actual toxic species, MPP+, which is a substrate for DAT. When inside the cytoplasm MPP+ acts as a mitochondrial poison, although its potency is restrained by VMAT2 which can sequester MPP+ inside vesicles away from mitochondria. The tuberoinfundibular, tuberohypophysial, olfactory bulb, mesoprefrontal, and mesoamygdala dopamine neurons appear to differ from the other dopamine systems described by the absence, or diminished expression, of DAT.

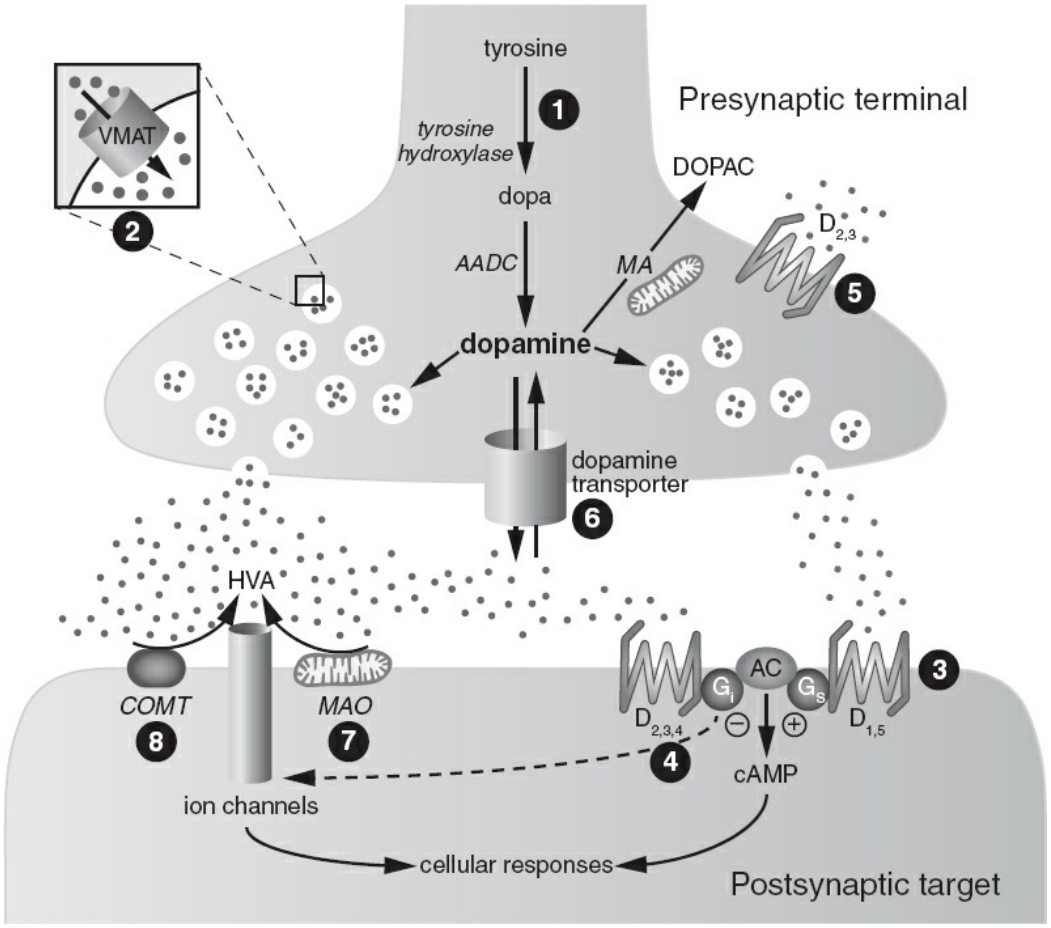

FIGURE 1.4–2. Schematic model of a dopaminergic synapse with sites of drug action. Site 1: Tyrosine hydroxylase reaction blocked by the competitive inhibitor, α-methyltyrosine. Administration of L-dopa augments dopamine levels and is the mainstay treatment of Parkinson disease. Carbidopa is an inhibitor of AADC that does not cross the blood–brain barrier, and is used in conjunction with L-dopa in the treatment of Parkinson disease to reduce the peripheral side effects of L-dopa. Site 2: VMAT transports cytoplasmic DA into storage vesicles, decreasing its cytoplasmic concentration and preventing metabolism by monoamine oxidase. Reserpine inhibits VMAT and depletes dopamine stores. Tetrabenazine also interferes with the uptake–storage mechanism and has been used for treatment of hyperkinetic movement disorders. Sites 3 and 4: Once released, dopamine can interact with postsynaptic D_1-like receptors (site 3) and D_2-like receptors (site 4). Bromocriptine (Parlodel) and apomorphine (Apokyn) are D_2 agonists used in the treatment of Parkinson disease. Pramipexole (Mirapex) is a D_3 agonist used in the treatment of Parkinson disease and restless legs syndrome. Aripiprazole (Abilify) is a partial D_2 agonist used as an antidepressant. All antipsychotic drugs block D_2 receptors to some degree. Site 5: Interaction of dopamine with autoreceptors regulates synthesis and release of dopamine. Low doses of apomorphine have been used experimentally as an autoreceptor agonist. Site 6: The plasma membrane transporter recaptures synaptic dopamine. The drug of abuse, cocaine, inhibits the dopamine transporter, and amphetamine increases dopamine release via the transporter. Site 7: MAO present outside the presynaptic terminal can metabolize dopamine that has been released. Selegiline (Eldepryl) is an MAO-B inhibitor used in the treatment of Parkinson disease. Site 8: Dopamine is a substrate for COMT, an enzyme that is inhibited by tolcapone and entacapone, which are used in treatment of Parkinson disease. AADC, aromatic amino acid decarboxylase; AC, adenylate cyclase; cAMP, cyclic AMP; COMT, catechol-O-methyltransferase; $D_{1–5}$, dopamine receptor subtypes; DOPAC, dihydroxyphenylacetic acid; G, G protein; HVA, homovanillic acid; MAO, monoamine oxidase; VMAT, vesicular monoamine transporter. (Modified from Nestler E, Hyman S, Holtzman D, Malenka R.

Molecular Neuropharmacology, McGraw-Hill Companies, Inc; 2015.)

Norepinephrine Plasma Membrane Transporter. The NET is located on all norepinephrine neurons and efficiently recaptures most of the transmitter that is released into the synapse (Fig. 1.4–3). In addition to the impact of the psychostimulants, such as cocaine and the amphetamines, on DAT, these drugs also interact with NET and intensify the synaptic actions of NE. The mechanism of certain antidepressant drugs includes blockade of NET, and these include the tricyclic antidepressants (TCAs) and the more specific serotonin–norepinephrine reuptake inhibitors (SNRIs) such as duloxetine (Cymbalta) and venlafaxine (Effexor), norepinephrine–dopamine reuptake inhibitors (NDRIs) such as bupropion (Wellbutrin), and the selective norepinephrine reuptake inhibitors (NRIs), such as maprotiline and atomoxetine (Strattera), which is also used in the treatment of ADHD.

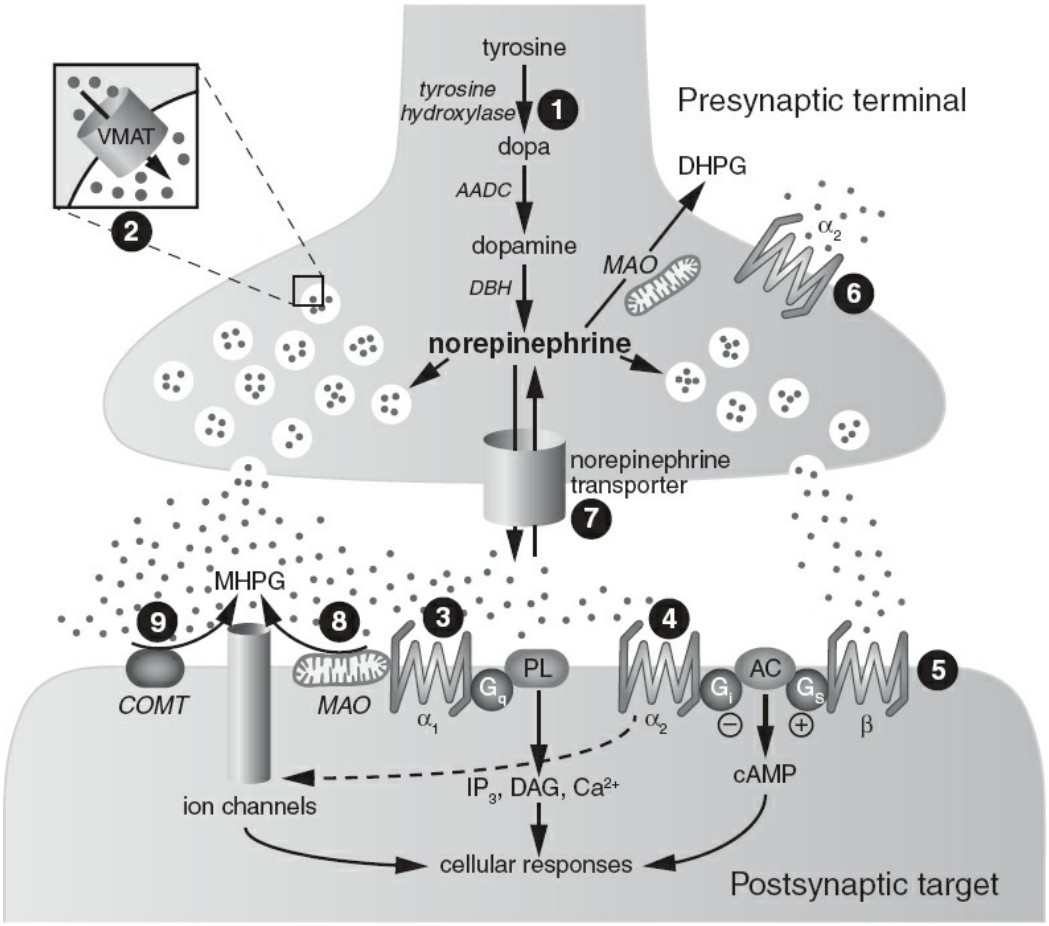

FIGURE 1.4–3. Schematic model of a central norepinephrine synapse with sites of drug action. Site 1: Tyrosine is transported into the presynaptic terminal by an active uptake mechanism and converted to NE by a series of enzymatic steps. The competitive inhibitor, α-methyltyrosine blocks the tyrosine hydroxylase reaction. DBH reaction is prevented by disulfiram. Site 2: Norepinephrine is taken up from the cytoplasm and transported into storage vesicles by the VMAT-2. Reserpine inhibits VMAT and depletes norepinephrine stores; the drug was once widely prescribed for its antihypertensive actions. Tetrabenazine also interferes with the uptake–storage mechanism and has been used for treatment of

hyperkinetic movement disorders. Sites 3, 4, and 5: After release, norepinephrine can interact with two categories of G-protein–coupled adrenergic receptors, α_1 and α_2 (sites 3 and 4) and β (site 5). The majority of adrenergic receptors are localized postsynaptically where they mediate the cellular responses of the postsynaptic neuron. Propranolol (Inderal) is an antagonist of β-adrenergic receptor, and is used to treat performance anxiety and hypertension. α_1-Receptor (site 3) is inhibited prazosin (Minipress), used as an antihypertensive, but which also has anxiolytic effects used in the treatment PTSD. Site 6: α_2-Autoreceptors localized in the nerve terminals modulate the synthesis and release of norepinephrine, while cell body and dendritic α_2-autoreceptors modulate impulse flow. Low doses of clonidine (Catapres) and guanfacine (Intuniv) are agonists at these sites. α_2-Agonists are used to treat opioid withdrawal, Tourette syndrome, and hypertension. Yohimbine is an antagonist at this receptor site. Site 7: The synaptic action of norepinephrine is terminated by reuptake into the presynaptic terminal by norepinephrine plasma membrane transporter. The tricyclic antidepressants, such as desipramine (Norpramin), are inhibitors of this uptake mechanism. Some newer antidepressants are more selective for NET, such as atomoxetine (Strattera), and by SNRIs such as duloxetine (Cymbalta). Site 8: Norepinephrine is metabolized by MAO and it is a preferential substrate for MAO-A. Inhibitors of MAO-A block metabolism of norepinephrine, and this action may contribute to their antidepressant property. Examples of irreversible nonselective inhibitors are isocarboxazid, phenelzine, and tranylcypromine. Moclobemide inhibits MAO-A reversibly. Site 9: Norepinephrine is also metabolized by COMT, which is inhibited by tolcapone (Tasmar) and entacapone (Comtan), drugs that are used in treatment of Parkinson disease. The combined actions of MAO and COMT produces MHPG, a major brain norepinephrine metabolite. α and β, subtypes of norepinephrine receptor; AADC, aromatic amino acid decarboxylase; AC, adenylate cyclase; cAMP, cyclic AMP; COMT, catechol-O-methyltransferase; DAG, diacylglycerol; DBH, dopamine β-hydroxylase; DHPG, 3,4-dihydroxyphenylglycol; G, G protein; IP_3, inositol triphosphate; MAO, monoamine oxidase; MHPG, 3-methoxy-4-hydroxyphenylglycol; SNRI, serotonin–norepinephrine reuptake inhibitor; VMAT, vesicular monoamine transporter. (Modified from Nestler E, Hyman S, Holtzman D, Malenka R. *Molecular Neuropharmacology*, McGraw-Hill Companies, Inc; 2015.)

Since the original studies on dopamine uptake, it has been known that dopamine is a good substrate for the NET. However, not until relatively recently has the importance of this been appreciated. In regions, such as striatum and nucleus accumbens, where dopamine neurons express a high density of DAT, there is a relatively sparse norepinephrine innervation, and dopamine uptake depends primarily on DAT. However, in regions such as the prefrontal cortex, in which dopamine neurons express low levels of DAT and there is a rich norepinephrine innervation, dopamine uptake depends strongly on NET. In fact, it has been suggested that dopamine can be co-released with norepinephrine by norepinephrine neurons, as a result of nonspecific uptake of dopamine by the NET and/or because dopamine is a precursor in the biosynthesis of norepinephrine.

Epinephrine Plasma Membrane Transporter. There is no evidence for a specific transporter for epinephrine and no drug that selectively blocks epinephrine reuptake has been described. All of the pharmacological agents that block the NET selectively are also effective inhibitors of epinephrine uptake. Thus, it appears unlikely that a selective transporter for epinephrine exists.

Metabolism. The major mammalian enzymes of importance in the metabolic degradation of catecholamines are MAO and catechol O-methyltransferase (COMT), and either of these enzymes can catalyze the first step in catecholamine metabolism. MAO converts its substrates to their corresponding aldehydes, and these aldehyde intermediates are rapidly metabolized either by the aldehyde dehydrogenase or aldehyde reductase. Thus, multiple metabolites of catecholamines are formed. In primate brain the major end metabolites for dopamine is homovanillic acid (HVA), whereas in rodent brain 3,4-dihydroxyphenylacetic acid (DOPAC) and DOPAC sulfate predominate. For norepinephrine, 3-methoxy-4-hydroxyphenylglycol (MHPG) is the main end metabolite in primate brain, but in rodent brain both 3,4-dihydroxyphenylglycol (DHPG) and MHPG exist in high concentrations and both metabolites are mainly sulfate conjugated. In the periphery, the major metabolite of dopamine is HVA and for norepinephrine it is vanillylmandelic acid (VMA) with a significant proportion of both compounds being conjugated to either sulfate or to glucuronide.

It is well documented that, in the animals, short-term fluctuations in the levels of dopamine metabolites, DOPAC and HVA, in the striatum, nucleus accumbens, and prefrontal cortex provide a useful index, respectively, of alterations in impulse flow in the nigrostriatal, mesolimbic, and mesocortical dopamine pathways. Changes in dopamine metabolite concentration in CSF have also been associated with alterations in central dopaminergic activity, or integrity of the central dopamine systems. Similarly alterations in the concentration of MPHG have been associated with changes in norepinephrine function.

MAO. This enzyme is located on the outer membranes of mitochondria. Separate genes encode two isoforms of MAO (types A and B), which can be distinguished by substrate specificity and sensitivity to the selective inhibitors. In brain, MAO-A is preferentially located in dopaminergic and noradrenergic neurons, while MAO-B is the major form present in serotonergic neurons and glia. Dopamine and norepinephrine are effective substrates for both forms, whereas serotonin is a preferential substrate for MAO-A. The prototypical inhibitors of the two forms are clorgyline for MAO-A and selegiline (formerly known as deprenyl) for MAO-B. Selegiline (Eldepryl) or rasagiline (Azilect) are irreversible MAO-B inhibitors that are used in the treatment of Parkinson disease, usually in combination with L-dopa to prolong or enhance the effect of dopamine formed from L-dopa.

COMT. There are two isoforms of COMT, a membrane-bound form and a soluble form. Membrane-bound COMT is the major form found in the CNS, has a higher affinity for catecholamines, and it is located principally in neurons. The soluble form has a lower affinity for catecholamines and is the major form expressed in the periphery, but is present also in CNS glia. Two COMT inhibitors, tolcapone and entacapone,

are used as adjuncts to L-dopa and treatment in Parkinson disease to slow L-dopa metabolism and thereby prolong its effects and reduce L-dopa requirements.

Accumulating evidence from clinical and preclinical studies shows that COMT plays a more significant role in dopamine metabolism in the prefrontal cortex than in other regions such as striatum. Another interesting facet to the role of COMT in frontal cortex is the finding that the COMT gene contains a polymorphism (Val158Met) that affects the in vivo activity of the enzyme. Met158 homozygotes have approximately one-third less COMT enzyme activity in prefrontal cortex than Val158 homozygotes. Consistent with its role in modulating prefrontal cortex dopamine levels, the Val158Met polymorphism is associated with performance on tests of working memory and executive function, which depend on prefrontal cortex function. Thus, in healthy volunteers the high-activity Val158 allele is linked with relatively poorer performance on such tasks, relative to the Met158 allele, presumably as a result of increased dopamine metabolism and lower synaptic dopamine levels. However, this relationship is not constant across patient populations and drug-treated individuals as an inverted "U"-shaped curve describes the relationship between dopaminergic transmission in prefrontal cortex and cognitive performance, with both suboptimal and supraoptimal dopamine activity impairing cognitive performance. The COMT polymorphism has been implicated in a number of neuropsychiatric phenotypes including evidence to support an association between COMT allele frequency and the genetic risk of schizophrenia.

Dopamine Receptors

The physiological actions of dopamine are mediated by five distinct but closely related GPCRs. On the basis of their structure, pharmacology, and primary effector mechanisms, they are grouped as D_1-like (D_1 and D_5) and D_2-like (D_2, D_3, D_4) dopamine receptors (Fig. 1.4–2). The D_1-like dopamine receptors activate the Gs family of G proteins, increase adenylate cyclase activity and are found postsynaptically on dopamine-receptive cell. The D_2-like dopamine receptors (D_2, D_3, and D_4) couple to the Gi family, inhibit adenylate cyclase activity, and are expressed both postsynaptically on dopamine-receptive cells and presynaptically on dopaminergic neurons.

In brain D_1 dopamine receptors are expressed in particularly high density in the nigrostriatal, mesolimbic, and mesocortical dopamine systems, such as the striatum, nucleus accumbens, and frontal cortex. In contrast, D_2 receptors are more evenly distributed in brain, with the highest levels in striatum and nucleus accumbens. Dopamine transmission through D_1 and D_2 receptors in striatum and prefrontal cortex plays an essential role in many aspects of motor control, reward systems, and

cognition. The alternative splicing of the D_2 gene results in generation of two major variants, termed D_{2S} (D_2-short) and D_{2L} (D_2-long). There is some evidence that D_{2S} is mostly expressed presynaptically and is involved in autoreceptor functions, with D_{2L} being predominantly a postsynaptic isoform. The D_3 dopamine receptor has a more limited pattern of distribution, with the highest level of expression being observed in the limbic areas, such as in the shell subdivision of the nucleus accumbens. Some data suggest that D_3 receptors can operate as autoreceptors, to complement the D_2 autoreceptor role in regulating the neuronal firing rate, synthesis, and release of dopamine. Among the dopamine receptors, the D_4 subtype has the lowest level of expression in the brain, with documented expression in frontal cortex, amygdala, hippocampus, and hypothalamus. D_5 dopamine receptors are expressed at low levels in multiple brain regions, including frontal cortex, hypothalamus, and hippocampus. D_1, D_2, and D_4 dopamine receptors have also been observed in the retina. High expression of D_2 dopamine receptors also occurs in the pituitary gland. In the periphery, all subtypes of dopamine receptors have been observed in varying proportions in the kidney, adrenal glands, sympathetic ganglia, gastrointestinal tract, blood vessels, and heart.

Some of the older, so-called "*typical,*" antipsychotic drugs, such as chlorpromazine, block both D_1 and D_2 receptors, while others like haloperidol, are selective for D_2 receptors. All "*atypical*" antipsychotics effectively block D_2 dopamine receptors with little or no impact on D_1 receptors. As the full beneficial effects of antipsychotic drugs can take weeks to develop, it is apparent that the mechanism involved is more than just blockade of the D_2 receptor site. In addition, all antipsychotic drugs have appreciable affinities for non-DA receptors, and atypical antipsychotic drugs tend to have substantial potency at serotonin receptors, although it is not clear to what extent the interaction with serotonin receptors improves the antipsychotic effect or reduces some of the unwanted effects of dopamine receptor blockade. Typical antipsychotic drugs are prone to induce extrapyramidal side effects, which include parkinsonism, dystonia, and choreiform movements (tardive dyskinesia). An additional side effect is elevated serum prolactin, resulting from a disturbance of dopamine-mediated regulation of prolactin secretion. The consequences of hyperprolactinemia include menstrual disturbances, galactorrhea, sexual dysfunction, gynecomastia, infertility, decreased bone mineral density, and possibly breast cancer. Atypical antipsychotics with especially high affinity for D_2 receptors, such as risperidone (Risperdal), are more prone to induce hyperprolactinemia, whereas this side effect is rare with those drugs that bind to D_2 less avidly, such as aripiprazole (Abilify), clozapine (Clozaril),

olanzapine (Zyprexa), quetiapine (Seroquel), and ziprasidone (Geodon). Although clozapine was discovered in the 1960s, it still has an important role in combatting treatment-resistant schizophrenia with low likelihood of extrapyramidal side effects. However, approximately 1 percent of patients treated with clozapine develop serious agranulocytosis, while it also shares the frequent side effects of other antipsychotic drugs such as weight gain. The success of clozapine was initially linked to its antagonism of D_4 receptors, which distinguishes it from other antipsychotic drugs. However, current thinking is that clozapine's combined actions at several receptors may be responsible for its therapeutic properties. Of the newer antipsychotic drugs, olanzapine has the closest receptor profile to clozapine. Atypical antipsychotic drugs, including clozapine, are generally superior to typical neuroleptics with regard to treating negative symptoms and cognitive dysfunction associated with schizophrenia.

D_2-like agonists have an important use in the symptomatic treatment of Parkinson disease. Such drugs include apomorphine, a D_2 agonist, and also pramipexole (Mirapex) and ropinirole (Requip), which are D_2-like agonists with greatest affinity for the D_3 receptor. Nonselective dopamine agonists (e.g., rotigotine) have also been used in Parkinson disease, in addition to restless legs syndrome. D_2 autoreceptors are generally activated by a lower concentration of dopamine agonists than necessary to activate postsynaptic D_2 receptors, so it possible in experimental models to demonstrate that the same dopamine agonist induces a biphasic effect depending on its dose. However, no drugs are used clinically to preferentially target D_2 autoreceptors.

Historically, GPCRs are believed to function as monomeric units; however, it now apparent in some instances dopamine receptors may directly interact with members of the same receptor family and with structurally divergent families of receptors to form heteromers with unique functional and pharmacological properties, although this knowledge has yet to lead to novel pharmacological agents.

Adrenergic Receptors

For many years, receptors for norepinephrine and epinephrine (adrenergic receptors) were divided into just three distinct classes, α_1, α_2, and β. A postsynaptic location is typical for α_1- and β-receptors, whereas α_2-receptors exist both pre- and postsynaptically. Currently, three α_1 subtypes (α_{1a}, α_{1b}, α_{1d}), four α_2-receptor subtypes (α_{2A-D}), and three β-receptor subtypes (β_{1-3}) are recognized (Fig. 1.4–3). All β-receptor are Gs-coupled, whereas most α_1-receptors are Gq-coupled and α_2-receptors are typically Gi-coupled.

As norepinephrine is the transmitter used by most postganglionic

neurons of the ANS, agonists and antagonists at various adrenergic receptors exert profound effects on the functioning of many peripheral organs by mimicking or blocking respectively the sympathetic innervation. Agonists are referred to as "*sympathomimetics*" and antagonist as "*sympatholytics*." In brain, the locus coeruleus noradrenergic system exerts a widespread influence on neuronal circuits that are essential substrates of arousal and cognition and several cognitive and affective disorders have been posited to involve a dysregulation of noradrenergic neurotransmission. In addition, by sending direct projections to the sympathetic preganglionic neurons in the spinal cord and parasympathetic preganglionic neurons in the brainstem and spinal cord, the noradrenergic locus coeruleus plays central role regulating autonomic activity.

The predominant peripheral effect of α_1-adrenergic agonists is the constriction of arterial smooth muscle, which in turn causes an increase in blood pressure, and drugs such as phenylephrine and methoxamine are used occasionally in the clinic to increase blood pressure during a hypotensive crisis. α_1-Adrenergic agonists are also widely used as nasal decongestants, and their decongestant activity is mediated by constriction of blood flow to nasal mucosa and by a reduction in airway secretions. Antagonists of α_1-adrenergic receptors inhibit smooth muscle contraction and reduce arteriolar resistance by dilating arteries. Their major uses are as an adjuvant treatment for hypertension and for treating symptomatic benign prostatic hypertrophy. α_1-Antagonists such as prazosin (Minipress) have also been shown to be useful in treating some of the symptoms of PTSD.

α_2-Adrenergic agonists, like clonidine (Catapres) and guanfacine (Intuniv) have been used to treat hypertension and the mechanism appears to be a reduction in sympathetic tone elicited by activation of α_2-receptors in the anterior hypothalamus. α_2-Adrenergic agonists also have a role in the treatment of certain CNS disorders, especially those involving prefrontal cortex dysfunction. Damage to the catecholaminergic innervation of the prefrontal cortex impairs performance in tests of executive function, and although dopamine has a prominent role in mediating the effects of these lesions, substantial evidence indicates that norepinephrine also exerts a potent modulatory influence on prefrontal cortex functions. Early pharmacological studies on the role of α_2-receptors in prefrontal cortex function focused on clonidine, although clonidine produces mixed effects, presumably due to competing pre- versus postsynaptic effects and limitations from sedative and hypotensive side effects. A superior profile has been observed with the more selective α_{2A}-agonist, guanfacine, which has been shown to improve working memory in animals and in some human populations, possibly by binding to

postsynaptic α_{2A}-receptors and increasing noradrenergic tone. Guanfacine is now being used to treat a variety of disorders with prominent prefrontal cortex dysfunction, including Tourette syndrome and ADHD. Clonidine is also used to manage the acute phase of opioid withdrawal, an effect achieved by normalizing the hyperactivity of noradrenergic locus coeruleus neurons. The α_2-antagonist, yohimbine, which increases activity of locus coeruleus noradrenergic activity by blocking autoreceptors, can induce symptoms of fear and anxiety in animals and humans.

These adrenergic agonists are used in the treatment of certain cases of hypertension.

β-Adrenergic agonists have profound effects on the cardiac and pulmonary systems and are one of the primary agents used to treat asthma and other chronic obstructive pulmonary diseases. Antagonists of β-adrenergic receptors are mainstays in the treatment of ischemic heart disease and hypertension. β-Adrenergic antagonists, such as propranolol, are also used to reduce signs of increased sympathetic activity in individuals with social anxiety (stage fright), such as hand tremor, sweating, and rapid heartbeat.

SEROTONIN

Anatomy and Functions

Mammals employ serotonin as a neurotransmitter within both the central and PNS. Serotonin is also found in many cells that are not neurons, such as platelets, mast cells, and the enterochromaffin cells. More than 80 percent of the serotonin found in the body is found in the gastrointestinal system where it modulates motility and digestive functions. In brain only one in a million neurons produces serotonin, yet this evolutionary ancient signaling molecule influences virtually all aspects of CNS function due to the extensive innervation pattern of the neurons. Due to this widespread distribution in brain, serotonin has been implicated in a vast array of physiological and pathological processes.

The cell bodies of serotonergic neurons are clustered in multiple midline raphe nuclei of the brainstem. The caudal raphe nuclei send descending projections into the medulla, cerebellum, and spinal cord, which are involved in motor activity, pain control, and regulation of autonomic processes. The more rostral located dorsal and median raphe nuclei send ascending axonal projections to the rest of the brain. The median raphe provides most of the serotonergic fibers that innervate the limbic system, while the dorsal raphe nucleus provides most of the serotonergic fibers that innervate the cerebral cortex, thalamus, and striatum, in addition to the dopaminergic neurons residing in the substantia nigra and VTA. These rostral serotonergic systems have a modulatory role in many functions including mood, anxiety, aggression,

cognition, feeding, the sleep–wake cycle, and sexual behavior. The link between serotonin and depression is particularly compelling, and the clinical efficacy of antidepressant drugs is strongly associated with an enhancement of serotonergic function.

Interestingly, the projections from the dorsal and median raphe can be distinguished morphologically. Dorsal raphe serotonergic fibers are very fine, with small vesicle-coated swellings called varicosities that are fusiform or granular in appearance, while median raphe fibers are thicker with large spherical or beaded varicosities. The two systems coexist in most areas of the brain, although the ratio varies considerably. The striatum, thalamus, and dopaminergic cell bodies in the midbrain receive innervation predominantly by the neurons originating in the dorsal raphe, while the hippocampus, septum, and hypothalamus receive a particularly high proportion of fibers arising in the median raphe. In cerebral cortex, there is extensive overlap between dorsal and median raphe projections. The two types of fibers exhibit differential sensitivity to the neurotoxic effects of the amphetamine analog 3,4-methylenedioxy-methamphetamine (MDMA, "ecstasy"), which causes long-lasting or permanent damage to the fine axons of the dorsal raphe while sparing the thick beaded axons of the median raphe.

Life Cycle

Synthesis. In brain the first step in serotonin synthesis is active uptake of the dietary amino acid tryptophan across the BBB and then into neurons (Fig. 1.4–4). Brain tryptophan is utilized in all brain cells for protein synthesis, in addition to serving as a substrate for transmitter synthesis in serotonin neurons. The carrier processes that transport tryptophan from plasma into brain and from extracellular fluid into neurons is open to competition from LNAAs, including other aromatic amino acids (e.g., tyrosine and phenylalanine), branched-chain amino acids (e.g., leucine, isoleucine, and valine), and others (e.g., methionine and histidine). The competitive nature of the LNAA carrier means that brain levels of tryptophan will be determined not only by the plasma concentration of tryptophan but also by the plasma concentration of competing neutral amino acids. The rate limiting enzyme for the synthesis of serotonin in brain, tryptophan hydroxylase, is not saturated under normal physiological conditions. Therefore, any decrease or increase in brain tryptophan levels results in a reduction or a stimulation of serotonin synthesis, respectively. Thus, dietary protein and carbohydrate content can specifically influence brain tryptophan and serotonin levels by effects on plasma amino acid patterns. Because plasma tryptophan has a daily rhythmic variation in its concentration, this concentration variation can also influence the rate and synthesis of brain serotonin. However, it is unclear to what extent presynaptic and postsynaptic homeostatic mechanisms prevent

tryptophan-induced alterations in serotonin synthesis from being translated into changes in neurotransmission.

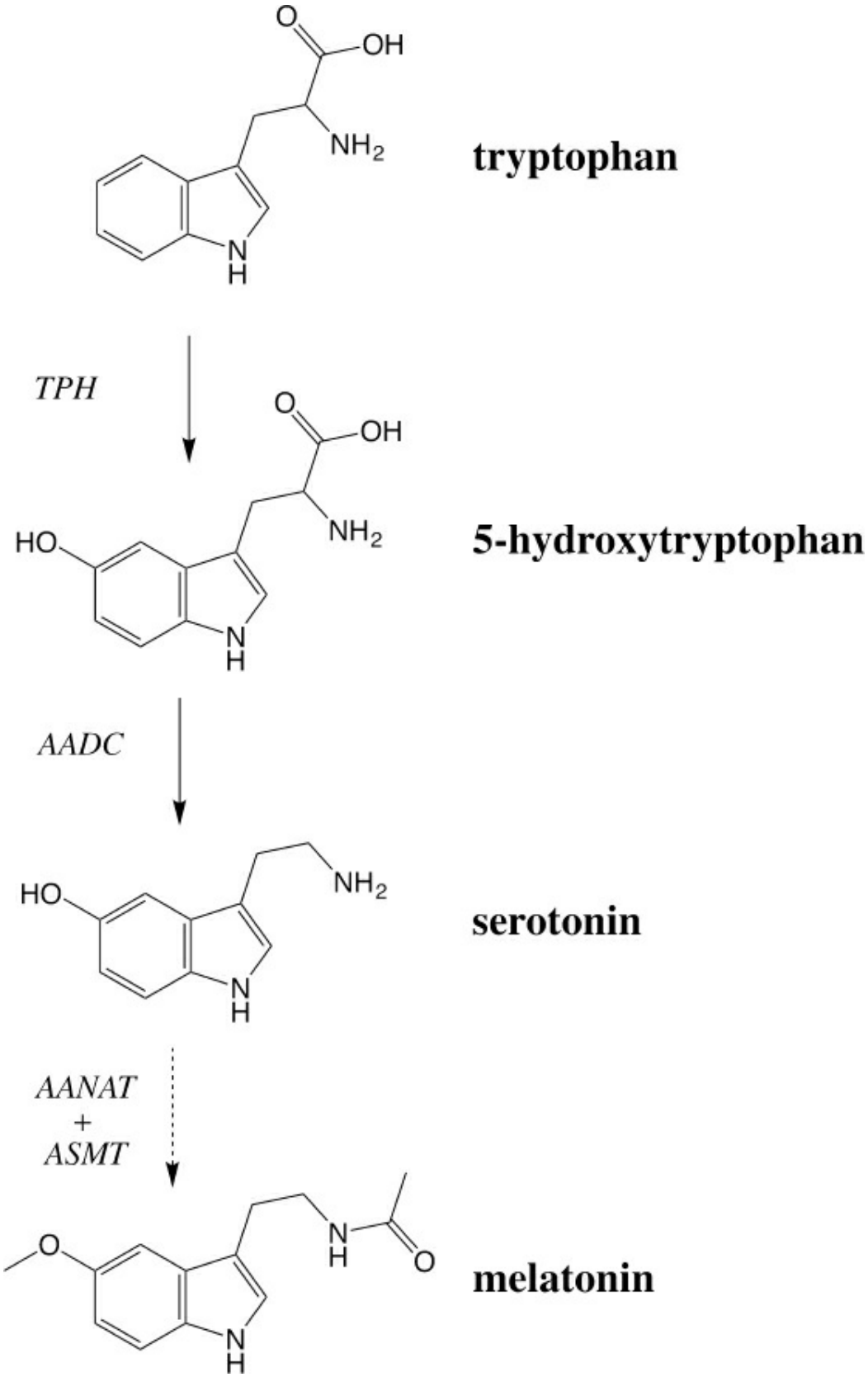

FIGURE 1.4–4. Biosynthesis of serotonin and melatonin. Serotonin (5-

hydroxytryyptamine, 5HT) and melatonin are produced by a multistep synthesis that starts with the dietary amino acid, tryptophan. Cells in the pineal gland are the major site of melatonin synthesis, and they contain the enzymes AANAT and ASMT that are necessary to convert serotonin to melatonin. AADC, aromatic amino acid decarboxylase; AANAT, aralkylamine *N*-acetyltransferase (also known as serotonin acetylase); ASMT, acetylserotonin O-methyltransferase (also known as hydroxyindole O-methyltransferase); TPH, tryptophan hydroxylase.

Tryptophan hydroxylase catalyzes the formation of 5-hydroxytryptophan from tryptophan, and this step in the synthesis of serotonin can be specifically blocked by p-chlorophenylalanine, which competes directly with tryptophan and binds irreversibly to the enzyme. Once synthesized from tryptophan, 5-hydroxytryptophan is almost immediately decarboxylated to yield serotonin. The enzyme responsible for this conversion is AADC, and is identical to the enzyme that decarboxylates L-dopa in dopamine and norepinephrine neurons. AADC is widespread in distribution; it is found in the peripheral and CNS associated with catecholamine- and serotonin-containing neurons and in the adrenal and pineal glands. It is also found in the kidney, liver, and various other tissues in which little or no monoamine transmitter is normally produced. It is possible to increase serotonin formation experimentally by administering 5-hydroxytryptophan and bypassing the rate-limiting tryptophan hydroxylase step. Thus, unlike tryptophan administration, which can result in a selective increase in serotonin in serotonin-containing neurons, 5-hydroxytryptophan administration will result in the nonspecific formation of serotonin at all sites containing AADC, including the catecholamine-containing neurons.

In the pineal gland, serotonin can be further metabolized by acetylation and O-methylation to form melatonin. Melatonin has a major role in regulation of cyclic bodily activities. The SCN of the anterior hypothalamus is the master clock controlling circadian rhythms in mammals and it controls melatonin secretion from the pineal by an indirect pathway in response to the environmental light/dark cycle. Peak melatonin secretion occurs during darkness. Melatonin signals both the time of the day ("clock" function) and season of the year ("calendar" function) to many tissues of the body. Melatonin is available for over-the counter use as a hypnotic agent to induce sleep, but has limited effectiveness in healthy subjects at night when endogenous production of melatonin is high. Administration of melatonin may be of use to treat jet lag or to help shift-workers adjust to new work hours. Melatonin agonists have been developed and recently marketed, such as ramelteon (Rozerem).

Transporters. After synthesis of serotonin in the cytoplasm, it is actively sequestered and concentrated in synaptic vesicles to await release (Fig. 1.4–5). Uptake of serotonin into vesicles is accomplished by VMAT-2 in brain, and is inhibited by drugs such as reserpine and tetrabenazine.

Once released into the synapse, serotonin can interact with presynaptic and postsynaptic serotonin receptors. The synaptic effects of serotonin are limited by reuptake into the presynaptic terminal and by enzymatic conversion of serotonin to inactive metabolites. Reuptake serves as the major mechanism for the termination of the action of synaptic serotonin. The primary targets for the TCAs, such as amitriptyline (Elavil), clomipramine (Anafranil), desipramine (Norpramin), imipramine (Tofranil), and nortriptyline (Pamelor) is blockade of serotonin and NETs, which prolongs the action of the transmitters in the synapse, although this class of drug possesses several other pharmacological actions. Newer SSRIs are now the most widely prescribed antidepressants and antianxiety medications, such as citalopram (Celexa), fluoxetine (Prozac), fluvoxamine (Luvox), sertraline (Zoloft), and paroxetine (Paxil). More recently, SNRIs, such as duloxetine (Cymbalta) and venlafaxine (Effexor) have been introduced for treatment of depression disorders. SSRIs are also acknowledged as the first-line pharmacological treatment of choice for obsessive-compulsive disorder (OCD). A polymorphism occurs in the promoter region of the serotonin transporter gene that affects its expression, and this finding provoked considerable controversy over the possible functional consequences of a particular variant (5-HTTLPR). It now appears that there is a small but real effect of this polymorphism on antidepressant response and on depressive symptoms following adverse life events.

Psychomotor stimulants like amphetamine enter monoamine nerve terminals by their transporters (DAT, NET, and SERT) and stimulate release of all three monoamines into the synapse by rapid reverse transport. Halogenated amphetamine derivatives (e.g., fenfluramine) are selective for SERT compared with DAT and NET. There is some evidence that the appetite suppressing action of fenfluramine is mediated by release of serotonin and activation of $5HT_{2C}$ receptors in the medial hypothalamus. Fenfluramine alone or in combination with phentermine (called "fen-phen") was formerly prescribed as an appetite suppressant. However, fenfluramine was withdrawn from the market when it was found to be associated with a high incidence of pulmonary hypertension and valvular heart disease. This toxic effect is believed to be related either to the peripheral release of serotonin from platelets or more likely as a result of activation of cardiac $5HT_{2B}$ receptors by norfenfluramine, the main metabolite of fenfluramine.

Metabolism. Serotonin is enzymatically degraded by MAO, and the product of this reaction, 5-hydroxyindoleacetaldehyde, is oxidized to 5-hydroxyindoleacetic acid (5-HIAA). Although serotonin is deaminated more effectively by MAO-A compared to MAO-B, it is actually the B-form that is located inside serotonergic neurons. Some research has shown that

people with a particular loss-of-function variant of the MAO-A gene are predisposed to violence and other antisocial behaviors, but this research is quite controversial. Inhibitors of MAO were the first drugs to be marketed as antidepressants, and the first generation of monoamine oxidase inhibitors (MAOIs) were not selective for the A or B isoforms and inhibited the enzyme irreversibly, such as tranylcypromine (Parnate) and phenelzine (Nardil). However, the propensity of tyramine-rich foods (such as cheese) to provoke dangerous hypertensive reactions in patients taking these MAOIs was soon recognized and forced their use to be severely curtailed. As the intestinal mucosa contains predominantly MAO-A, it is an important site for metabolizing ingested MAO substrates like tyramine. When intestinal MAO-A is inhibited, dangerous amounts of tyramine can enter the circulation and release norepinephrine from sympathetic neurons and epinephrine from adrenal medulla, leading to a pressor response. Selective MAO-B inhibitors were shown to be safer with regard to this "cheese reaction," but unfortunately lacked antidepressant efficacy. Nevertheless, the selective MAO-B inhibitor, selegiline is marketed as an antidepressant, but has to be used at high enough doses to inhibit MAO-A. The delivery of selegiline transdermally (Emsam) is designed to reduce the risk of tyramine-provoked reactions by minimizing inhibition of gastrointestinal and hepatic MAO-A activity, while allowing inhibition of central MAO-A and MAO-B. Reversible inhibitors of MAO-A (RIMA) (e.g., moclobemide) are efficacious in depressive disorders and appear relatively safe with respect to tyramine sensitivity.

Receptors

Seven distinct families of 5HT receptors have been identified ($5HT_1$ to $5HT_7$), and subpopulations have been described for several of these (Fig. 1.4–5). All of them except the ionotropic $5HT_3$ receptor are G protein coupled. For serotonin GPCRs, three main types of primary coupling to G proteins have been described. The $5HT_{1A}$ receptors activate Gi/Go proteins, the $5HT_{2A}$ receptors activate Gq/11, and the $5HT_4$, $5HT_6$, and $5HT_7$ activate Gs receptors.

$5HT_1$ Receptors. $5HT_1$ family comprises five different gene products: $5HT_{1A}$, $5HT_{1B}$, $5HT_{1D}$, $5HT_{1E}$, and $5HT_{1F}$ receptors. The $5HT_{1A}$ receptor is the most widely distributed of all 5HT receptors, and they are expressed on the soma and dendrites of all serotonin neurons where they serve as autoreceptors to regulate cell excitability, and the synthesis and release of 5HT. $5HT_{1A}$ receptors are also located postsynaptically and on astrocytes and other glia. Partial $5HT_{1A}$ agonists, such as buspirone (Buspar), are clinically useful anxiolytic drugs.

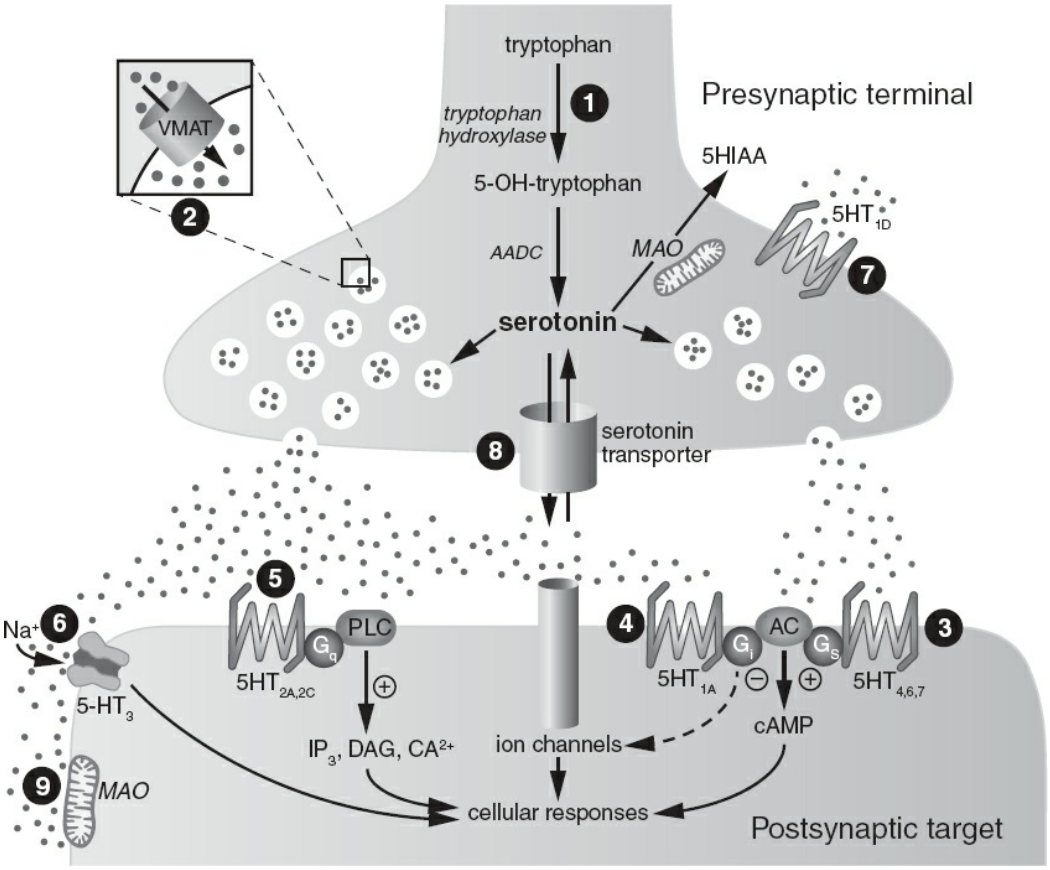

FIGURE 1.4–5. Schematic model of a serotonin synapse with sites of drug action. Site 1: Tryptophan is taken up into the neuron by an active transport mechanism and converted to 5-OH-tryptophan by the rate-limiting enzyme tryptophan hydroxylase; this enzyme can be effectively inhibited by p-chlorophenylalanine. AADC converts 5-OH-tryptophan to serotonin. Site 2: Serotonin is taken up and stored in vesicles by VMAT-2. Reserpine and tetrabenazine interfere with the uptake-storage mechanism of the amine granules, causing a marked depletion of serotonin. Sites 3, 4, and 5: Once released, serotonin can interact with as many as 14 different receptors. A number of G-protein–coupled serotonin receptors are located postsynaptically. Lurasidone (Latuda), a novel antipsychotic drug with possible procognitive properties, has high affinity for the $5HT_7$ receptor (site 3). The anxiolytics, buspirone (Buspar) and gepirone (Ariza), are $5HT_{1A}$ receptors (site 4). Many antipsychotic drugs inhibit $5HT_{2A}$ receptors (site 5). The hallucinogenic effects of drugs such as LSD, mescaline, and psilocybin are mediated by agonist interactions at the $5HT_{2A}$ receptor. Site 6: In contrast to the other serotonin receptors, the $5HT_3$ receptor is a ligand-gated ion channel. Ondansetron is a $5HT_3$ antagonist that is an effective antiemetic and lacks parkinsonian side effects inherent in D_2 antagonists like metoclopramide (Reglan) or prochlorperazine (Compazine). Site 7: Serotonin autoreceptors (5-HT_{1B} rodent and 5-HT_{1D} human) modulate the stimulus-induced release of serotonin and are located on the presynaptic terminals. $5HT_{1A}$ autoreceptors are located on the serotonin cell bodies and dendrites where they modulate impulse flow. Anti-migraine triptan drugs, such as sumatriptan (Imitrex), are $5HT_{1D/1B}$ agonists. Site 8: Termination of the synaptic actions of serotonin is accomplished by its reuptake into the neuron by the serotonin transporter. SSRIs and SNRIs are classes of antidepressant that inhibit the serotonin transporter and increase the synaptic actions of serotonin. The halogenated amphetamine, fenfluramine, causes release of serotonin by rapid reverse transport. Site 9: Synaptic actions of serotonin are also attenuated by its metabolism. Inhibitors of MAO-A block metabolism of serotonin and are used as antidepressants. Examples of irreversible inhibitors are isocarboxazid (Marplan), phenelzine (Nardil), and tranylcypromine (Parnate), whereas moclobemide inhibits MAO-A reversibly. 5HIAA, 5-hydroxyindoleacetic acid; $5HT_{1-7}$, subtypes of

serotonin receptor; 5-OH-tryptophan, 5-hydroxytryptophan; AADC, aromatic amino acid decarboxylase; AC, adenylate cyclase; cAMP, cyclic AMP; DAG, diacylglycerol; G, G protein; IP_3, inositol triphosphate; MAO, monoamine oxidase; SSRI, selective serotonin reuptake inhibitor; VMAT, vesicular monoamine transporter. (Modified from Nestler E, Hyman S, Holtzman D, Malenka R. *Molecular Neuropharmacology*, McGraw-Hill Companies, Inc; 2015.)

$5HT_{1B}$ and $5HT_{1D}$ receptors resemble each other in structure and distribution. They are predominantly located on the presynaptic terminals of serotonin neurons where they function as autoreceptors to control local synthesis and release of 5HT. $5HT_{1B}$ receptors are also present on axon terminals of nonserotonergic neurons where they act as heteroreceptors to control the presynaptic function of other transmitters. The serotonergic system has long been linked to migraine and $5HT_{1B/D}$ agonists, such as sumatriptan (Imitrex) and other triptan drugs are effective treatments for these types of headaches. The mechanism is not clear however, and may be due either to vasoconstriction induced by activation of $5HT_{1B}$ and $5HT_{1D}$ receptors on cranial blood vessels or to modulation of activity in trigeminal afferents that carry pain information from the meninges.

$5HT_{1E}$ is a highly expressed receptor in human brain, and is particularly enriched in hippocampus, making it a possible drug target for memory disorders, such as Alzheimer disease or temporal lobe epilepsy. Interestingly, this receptor subtype is absent in rats and mice. Research into the functions of the $5HT_{1E}$ receptor has been hampered by the lack of selective pharmacological tools and specific antibodies.

Expression of $5HT_{1F}$ has been reported in various brain regions, in addition to the trigeminal ganglia, and cerebral blood vessels. Triptans, which are used to treat migraine, have appreciable affinity for $5HT_{1F}$ receptors, and the development of $5HT_{1F}$ agonists may provide promising new treatments for migraine.

Although there was originally a receptor designated as $5HT_{1C}$, it has been reclassified as $5HT_{2C}$ receptor due to its high sequence homology with other members of the $5HT_2$ subclass.

$5HT_2$ Receptors. The $5HT_2$ family has three members, $5HT_{2A}$, $5HT_{2B}$, and $5HT_{2C}$, which share similarities in amino acid sequence and signaling properties. They all are coupled to phospholipase C (PLC) by Gq/11.

$5HT_{2A}$ is expressed widely throughout the brain, with especially high abundance in the cortex, where it is present in both interneurons and pyramidal neurons. The disturbance in sensory perception induced by hallucinogenic (psychedelic) drugs (e.g., D-lysergic acid diethylamide [LSD], mescaline, psilocybin, and 2,5-dimethoxy-4-methylamphetamine), is mediated by full or partial agonism at the $5HT_{2A}$ receptor. Second-

generation antipsychotic drugs are characterized by their relatively higher potency as antagonists at the $5HT_{2A}$ site compared to D_2 receptors. There is some evidence that downregulation $5HT_{2A}$ receptor and/or $5HT_{2C}$ receptors may be an adaptive mechanism contributing to the antidepressant effect of SSRIs, and in fact blockade of $5HT_{2A}$ or $5HT_{2C}$ receptors enhances antidepressant-like effects of SSRIs in animal models. In agreement with this hypothesis is evidence of the effectiveness of antipsychotic drugs given with SSRIs in treatment-resistant depression. One useful effect of $5HT_{2A}$ receptor activation is a reduction in intraocular pressure, and this has led to the development of compounds that reduce pressure inside the eyes, but without crossing the BBB and producing hallucinogenic side effects.

$5HT_{2B}$ receptors have a restricted expression in the CNS, which contrast with their abundance in the periphery. Research on $5HT_{2B}$ has been limited due to the cardiotoxicity of $5HT_{2B}$ agonists, yet recent evidence in animal models suggest that $5HT_{2B}$ receptors positively modulates serotonergic activity may be required for the therapeutic actions of SSRIs. Consequently, the $5HT_{2B}$ receptor may become a new target in antidepressant treatments.

$5HT_{2C}$ receptor has a widespread and heterogeneous distribution in brain. $5HT_{2C}$ receptors have been implicated in the mechanism underlying weight gain resulting from chronic administration of antipsychotic drugs, as those with the potential to produce weight gain, such as clozapine (Clozaril) and olanzapine (Zyprexa), have greater affinity for the $5HT_{2C}$, compared with $5HT_{2A}$ receptor. Antipsychotic drugs that are more selective antagonists at the $5HT_{2A}$ subtype have less liability for weight gain. Agonists at the $5HT_{2C}$ receptor are being investigated as anti-obesity agents. Interestingly, $5HT_{2C}$ in humans is located at the X chromosome, whereas all other 5HT receptors are autosomal. Thus, it is possible that polymorphisms at this receptor would affect the two sexes to a differing extent. $5HT_{2C}$ polymorphism appears to determine levels of circulating leptin, providing a potential mechanism underlying the genetic association of the $5HT_{2C}$ receptor with weight gain.

$5HT_{2C}$ receptors are closely linked with the midbrain dopamine neurons, being able to regulate dopamine release in its terminal regions. For example, the firing rate of VTA dopamine neurons is inhibited or increased by $5HT_{2C}$ agonists or antagonists, respectively, such that $5HT_{2C}$ receptors regulate the release of dopamine in nucleus accumbens. The basis for this interaction is the location of $5HT_{2C}$ receptors in the VTA on both dopamine and GABA neurons that innervate the nucleus accumbens. Because of the involvement of $5HT_{2C}$ receptors with the VTA-nucleus accumbens pathway, specific ligands for this receptor subtype are under

development for the treatment of schizophrenia and substance abuse disorders.

$5HT_3$ Receptors. As mentioned above, $5HT_3$ receptor is ligand-gated ion channel, so that it differs markedly in structure and mechanism from the other 5HT receptor subtypes. $5HT_3$ receptors are located in many brain areas, with highest densities in the medulla and spinal cord. Some $5HT_3$ receptors can be located postsynaptically, but in many regions they also are frequently expressed presynaptically where their activation modulates the release of a variety of neurotransmitters. $5HT_3$ antagonists, such as ondansetron (Zofran), granisetron (Kytril), dolasetron (Anzemet), and palonosetron (Aloxi) are used clinically in the treatment of emesis associated with cancer chemotherapy and radiotherapy and to combat symptoms of irritable bowel syndrome. Owing to the unfavorable effects of $5HT_3$ agonists (e.g., nausea and anxiety) no clinical use of these is likely in the near future. In the periphery, $5HT_3$ receptors are prominently expressed in a variety of peripheral ganglia, and they have been proposed to modulate nociception and pain responses, as well as enteric and cardiovascular reflexes.

$5HT_4$ Receptors. These receptors are present in brain and periphery. The $5HT_4$ agonist cisapride (Propulsid) is in clinical use as a gastroprokinetic agent to increase motility in the upper GI tract.

$5HT_5$ and $5HT_6$ Receptors. Although $5HT_5$ receptors have been detected in brain, so far no drugs with selectivity for this subtypes is used clinically. The $5HT_6$ receptor is expressed almost exclusively in the brain, and its abundance in limbic and cortical regions has stimulated considerable research interest. There is recent evidence that an antagonist of $5HT_6$ receptors (Idalopirdine) may improve cognitive performance in patients with Alzheimer disease. Although the mechanism is not clear, the beneficial effect is thought to result from enhancements of cholinergic, glutamatergic, noradrenergic, and dopaminergic neurotransmission.

$5HT_7$ Receptors. The $5HT_7$ receptor, the most recently discovered member of the 5HT receptor family, has been identified in many brain regions, in addition to its presence in the periphery. There is some evidence for existence of both presynaptic and postsynaptic $5HT_7$ receptors. The function of this receptor has been most firmly linked with control of circadian rhythm and thermoregulation. However, a major recent interest in $5HT_7$ receptors derives from their possible relevant role in cognitive processes and the presence of $5HT_7$ receptor in hippocampus and

prefrontal cortex supports this. Furthermore, it is relevant that several antipsychotics and antidepressants have high affinity for the $5HT_7$ receptor, which may contribute to their beneficial effects. $5HTR_7$ has also been implicated in the modulation of structural plasticity in adolescent and mature brain circuits. Lurasidone (Latuda) is an atypical antipsychotic drug that is a particularly potent $5HT_7$ antagonist and is used to treat schizophrenia and depressive episodes associated with bipolar 1 disorder. Studies have even indicated that lurasidone can have procognitive properties in normal monkeys.

ACETYLCHOLINE

Anatomy and Function

In brain the axonal processes of cholinergic neurons either project to distant targets or contact cells locally within the same region (interneurons). Widely projecting cholinergic neurons originate from several clusters of cells within the basal forebrain complex and mesopontine complex in the upper brainstem. The basal forebrain population of cholinergic neurons comprises locations in medial septal nucleus, the diagonal band and the nucleus basalis of Meynert, and these neurons provide the vast majority of the cholinergic innervation to the telencephalon. Within the mesopontine complex, cholinergic neurons are located in pedunculopontine nucleus and in the LDT nucleus, and innervate the thalamus, and midbrain regions (including the substantia nigra and VTA), in addition to descending innervations to regions such as the locus coeruleus and dorsal raphe. While long projecting cholinergic neurons do innervate the striatum, the principal source of striatal acetylcholine is the cholinergic interneuron, which comprises only about 1 to 2 percent of all striatal cells, yet sends dense arbors of projections throughout the striatum. As would be predicted from their widespread distribution in brain, central cholinergic neurons are involved in a wide variety of processes such as cognition, learning and memory, arousal, motor control, and CBF. In Alzheimer disease there is significant degeneration of neurons in the nucleus basalis, leading to substantial reduction in cortical cholinergic innervation. The extent of neuronal loss correlates with degree of dementia, and the cholinergic deficit may contribute to the cognitive decline in this disease, consistent with the beneficial effects of drugs that promote acetylcholine signaling in this disorder.

In the autonomic nervous system (ANS), acetylcholine is responsible for preganglionic transmission for both parasympathetic and sympathetic neurons and also at postganglionic synapses for the parasympathetic system. However, there are exceptions to this organization, such as the sympathetic innervation of adrenal glands and sweat glands that is

provided by acetylcholine.

Life Cycle

The intraneuronal synthesis of acetylcholine involves only a single-step reaction (Fig. 1.4–6). The substrate choline is transported into cholinergic neurons by both high-affinity and low-affinity mechanisms, and acetylcholine is formed when the enzyme choline acetyltransferase catalyzes the transfer of an acetyl group from cytoplasmic acetyl coenzyme A to choline. The rate-limiting step in acetylcholine synthesis is the high affinity uptake of choline and while it is not clear how this is regulated, it is known that increased neuronal activity enhances choline entry into neurons. Once synthesized, acetylcholine is taken up into vesicles by the vesicular cholinergic transporter, a process that can be inhibited by vesamicol.

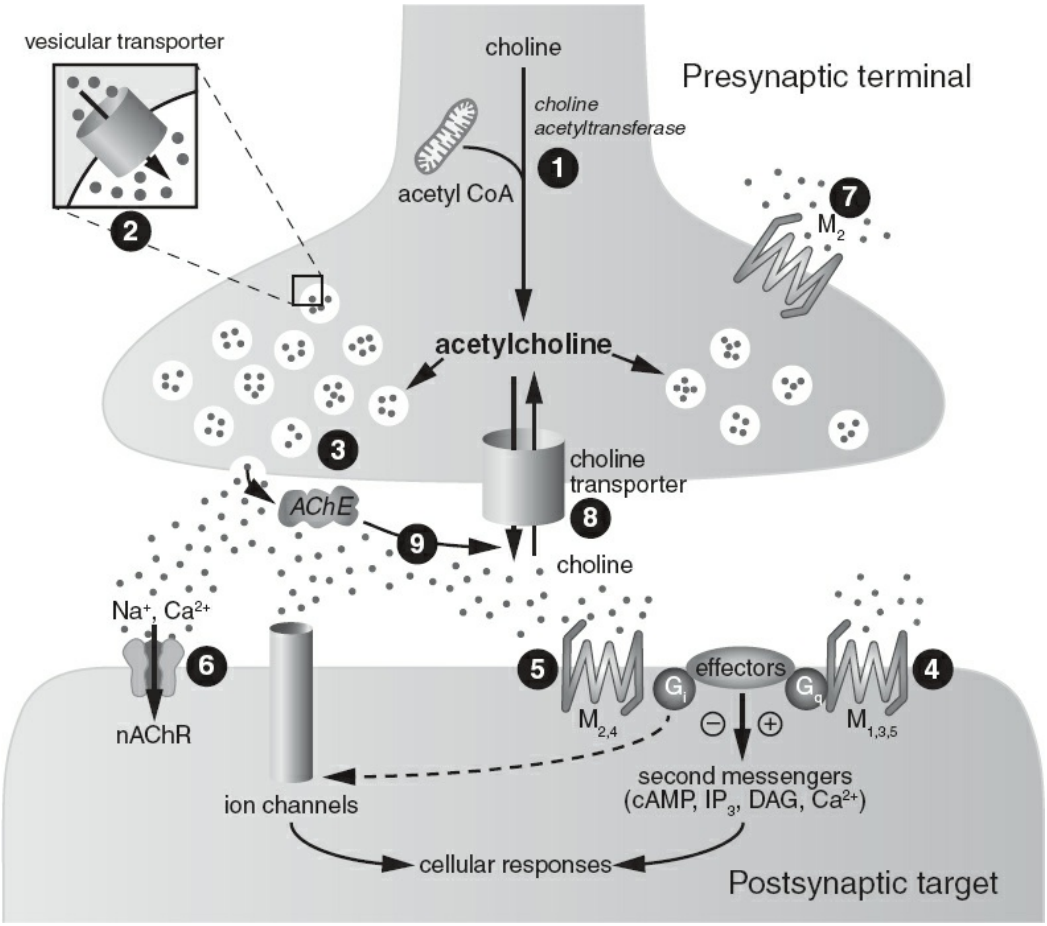

FIGURE 1.4–6. Schematic model of an acetylcholine synapse with sites of drug action. Site 1: Choline is transported into the presynaptic terminal by an active uptake mechanism and converted to acetylcholine by a single enzymatic step. Acetyl coenzyme A required for acetylcholine synthesis is provided by presynaptic mitochondria. Site 2: Acetylcholine is taken up and stored inside vesicles. The vesicular transporter in cholinergic neurons is distinct from VMAT-2 present in other biogenic amine neurons and is blocked by vesamicol. Site 3: Release is promoted by black widow spider venom and blocked by botulinum toxin (peripheral synapse) and tetanus toxin (central synapse). Sites 4 and 5: After release,

acetylcholine may interact with both G-protein–coupled muscarinic receptors (sites 4 and 5) on the postsynaptic cell. There are no muscarinic receptor subtype selective drugs, although some agents with partial selectivity are used experimentally. A clinically relevant nonselective muscarinic agonist is carbachol (Miostat), which is used in ophthalmological procedures. A typical antagonist is atropine, which is also in ophthalmology and to reduce salivary secretions during surgery. A subtype of nicotinic receptors is activated by varenicline, which is used to treat nicotine dependence. Nicotinic receptors are blocked by the pharmacological tools, α-bungarotoxin and hexamethonium. Site 6: Acetylcholine released from presynaptic terminals may also bind to nicotinic ligand-gated ion channels. A subtype of nicotinic receptors is activated by varenicline (Chantix), which is used to treat nicotine dependence. Nicotinic receptors are blocked by the pharmacological tools, α-bungarotoxin and hexamethonium. Site 7: Acetylcholine interacts with presynaptic muscarinic receptors to modulate subsequent release of transmitter. Site 8: In contrast to monoamine synapses, the plasma membrane transporter at a cholinergic synapse does not return the neurotransmitter to the presynaptic terminal but rather recycles precursor choline liberated after metabolism of acetylcholine. Hemicholinium-3 is a drug used to block choline transport experimentally. Site 9: Acetylcholine is metabolized by AChE, which present on both pre- and postsynaptic membranes, and serves to terminate its action. Examples of AChE inhibitors are physostigmine, which is used in eye drops to treat glaucoma, and donepezil (Aricept), which improves symptoms of Alzheimer disease. AChE, acetylcholinesterase; cAMP, cyclic AMP; DAG, diacylglycerol; G, G protein; IP_3, inositol triphosphate; $M_{1–5}$, subtypes of muscarinic receptor; MAO, monoamine oxidase; nAChR, nicotinic acetylcholine receptor; VMAT, vesicular monoamine transporter. (Modified from Nestler E, Hyman S, Holtzman D, Malenka R. *Molecular Neuropharmacology*, McGraw-Hill Companies, Inc; 2015.)

Several potent toxins affect release of acetylcholine from vesicles into the synapse. Botulinum toxin A and tetanus toxin are derived from bacteria and prevent release of acetylcholine resulting in paralysis. However, if it is injected in small amounts, botulinum toxin can weaken a muscle locally and this effect has been utilized in the treatment of spasms and dystonias, in addition as a cosmetic application in smoothing of facial wrinkles. Once released from the presynaptic neuron into the synapse, acetylcholine is not recaptured by an uptake mechanism, as occurs with the other biogenic amines. Instead the termination of cholinergic transmission is primarily dependent on the breakdown of acetylcholine to choline and acetate, a reaction catalyzed by AChE, an enzyme that exists both in soluble and membrane bound forms. Inhibition of AChE activity by drugs enhances the stimulation of cholinergic receptors by acetylcholine. Neostigmine (Prostigmin) is a reversible inhibitor that does not cross the BBB and is used to improve muscle tone in cases of myasthenia gravis. As AChE inhibitors cause decreased fluid pressure in the eye, they are used to treat certain types of glaucoma. Centrally acting reversible inhibitors of AChE that penetrate the brain well after oral administration, such as donepezil (Aricept) and rivastigmine (Exelon), are used to improve the symptoms of Alzheimer disease and dementia associated with Parkinson disease. Irreversible inhibitors of AChE permanently inactivate the enzyme, such that restoration of activity requires synthesis of new protein, and such compounds are employed as insecticides and have been used as nerve gases (e.g., sarin, soman) in chemical warfare.

Receptors

Acetylcholine receptors fall into two major classes, muscarinic and nicotinic receptors, named for their respective naturally occurring agonists, muscarine and nicotine (Fig. 1.4–6).

Muscarinic Receptors. Five subtypes of muscarinic receptors (M_1 to M_5) had been cloned and all are present in brain. They are coupled to G proteins, and either act directly on ion channels or are linked to a variety of second-messenger systems. M_1, M_3, and M_5 are Gq/11-coupled whereas M_2 and M_4 are Gi-coupled. M_1, M_3, and M_4 are primarily located in cerebral cortex and hippocampus, and mediate some of the effects of acetylcholine on learning and memory. High densities of M_1 and M_4 are located in striatum and mediate cholinergic signaling in extrapyramidal motor circuits and response to rewards. M_2 receptors are concentrated in the basal forebrain, where they function as autoreceptors to control synthesis and release of acetylcholine. The M_5 receptor is expressed at low levels in brain.

Nonspecific muscarinic agonists include some naturally occurring alkaloids such as pilocarpine and arecoline and synthetic compounds such as carbachol, oxotremorine, and bethanechol. Such drugs are not used to treat CNS disorders but are employed to reduce symptoms of glaucoma and nonobstructive urinary retention, and to treat the dry mouth that occurs in Sjogren syndrome. Arecoline is the active ingredient in the areca nut and betel leaf that are chewed in many Asian cultures for the stimulating effect they induce. The most prominent adverse effects of muscarinic cholinergic agents (bradykinesia, gastrointestinal distress, salivation, and sweating) are mediated by activation of peripheral M_1, M_2, and M_3 receptors. There are no subtype selective muscarinic agonists in clinical use, although they are being developed for potential use in treating cognitive disorders and in regulating the parasympathetic nervous system.

The prototypical muscarinic antagonists are atropine and scopolamine, which occur naturally in the plant deadly nightshade (Atropa belladonna). The translation of “belladonna” from Italian is “beautiful woman,” and this name was given because some women used to apply belladonna berry juice to their eyes in order dilate the pupils, giving them a striking appearance. Ingestion of muscarinic antagonists leads to other effects resulting from inhibition of the parasympathetic nervous system, such as dry mouth, constipation, and urinary retention, in addition to central effects such as cognitive impairment and in extreme cases, delirium. Antagonists of muscarinic antagonists have some therapeutic uses. For example scopolamine has anti-emetic activity and is used to treat nausea and vomiting caused by motion sickness. In addition, muscarinic antagonists

such as benztropine (Cogentin) and trihexyphenidyl (Artane) are employed to treat the motor symptoms of Parkinson disease, and are also sometimes helpful in ameliorating tardive dystonia, a possible side effect induced by long-term use of "typical" antipsychotic drugs. Many psychotropic drugs possess unwanted muscarinic antagonistic properties, including some TCAs and antipsychotic drugs. Some first-generation H_1 antagonist antihistamines are sold in over-the-counter hypnotics and anti-allergy medications and these drugs also have significant muscarinic antagonist actions that are responsible for side effects such as dry mouth, constipation, urinary retention, and even delirium to which geriatric patients are more susceptible.

Nicotinic Receptors. These are ligand-gated ion channels located at the neuromuscular junction, autonomic ganglia, adrenal medulla, brain, and spinal cord. Nicotinic receptors that are expressed in both neurons and muscle mediate rapid, excitatory signaling but differ in their subunit composition. Every nicotinic receptor consists of five subunit molecules arranged around a central channel that opens to admit ions when the receptor is activated. The subunit combination affects the pharmacological and functional properties of an individual complex, such as cation permeability and the rate of desensitization. These various nicotinic acetylcholine receptor subunits are categorized into three general functional classes: (1) skeletal muscle subunits (α_1, β_1, δ, and ε), (2) standard neuronal subunits (α_2–α_6 and β_2–β_4), and (3) subunits capable of forming homomeric receptors (α_7–α_9) where all five subunits are the same. Neuronal nicotinic acetylcholine receptors are widely distributed in brain, but they do not appear to function in the classical postsynaptic, directly excitatory manner of their muscle counterparts. Indeed, a large proportion of nicotinic binding sites are found on pre-, peri- and extrasynaptic sites in addition to the classic postsynaptic location.

Nicotinic receptors have been implicated in several pathological conditions. For example, mutations in one of the nicotinic receptor subunit genes is responsible for a specific form of epilepsy, autosomal dominant frontal lobe epilepsy (ADNFLE). In addition, a nicotinic receptor subunit gene mutation may be among the factors predisposing individuals to schizophrenia, and α_7-nicotinic receptors have emerged as a potential therapeutic target for the treatment of neurocognitive dysfunctions in schizophrenia. In Alzheimer disease and Parkinson disease, there is a significant loss of high affinity nicotinic receptor sites, which may contribute to the pathophysiological changes underlying the cognitive decline observed in these conditions. Because of the evidence for their involvement in a number of CNS disorders, there is great interest in the future development of selective nicotinic receptor agonists as therapeutics.

The most obvious role of neuronal nicotinic acetylcholine receptors is mediation of reinforcing and addicting properties of nicotine in chronic tobacco and e-cigarette users, in addition to the symptoms of withdrawal experienced upon cessation of use. The reinforcing effects of nicotine depend on its interaction with nicotinic receptors in the brain's reward system, notably in releasing dopamine from neurons that originate in the VTA and innervate the nucleus accumbens.

Although drugs such as hexamethonium and mecamylamine are useful pharmacological tools used to block nicotinic receptors, there are relatively few nicotinic agonists or antagonists that are used therapeutically. One example is succinylcholine, a weak partial agonist at nicotinic receptors, which is used clinically during anesthesia as a paralytic agent. Varenicline (Chantix) is a partial agonist with selectivity at the α_4/β_2 nicotinic receptor, which is now widely marketed as medication to treat nicotine dependence and as an aid in cessation of smoking. Galantamine (Razadyne) is an allosteric potentiator of several subtypes of neuronal nicotinic receptor and a weak competitive reversible inhibitor of AChE, which is used in the treatment of Alzheimer disease.

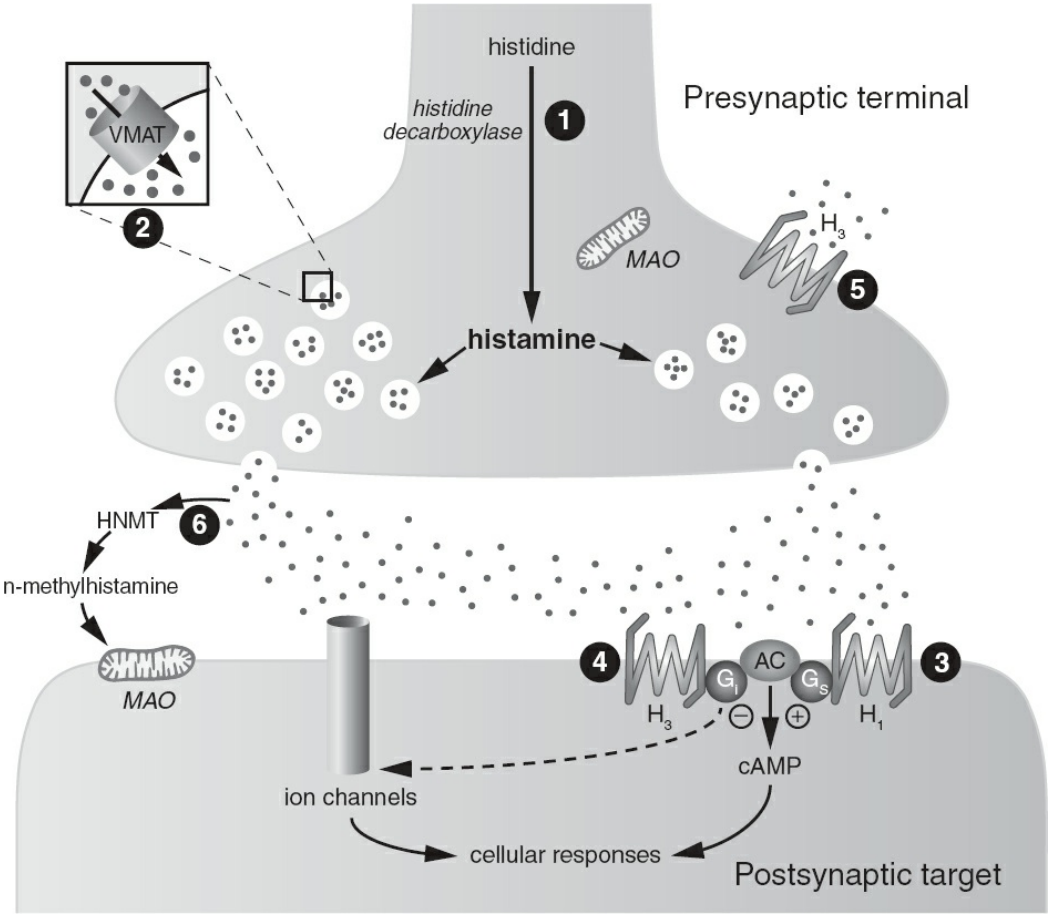

FIGURE 1.4–7. Schematic model of a central histaminergic synapse with site of drug action. Site 1: Histidine is taken up into histamine-containing neurons and converted to histamine by histidine decarboxylase. This enzyme is distinct from AADC that is used in the synthesis of other biogenic amines. No inhibitors of the histidine decarboxylase are used

clinically. Site 2: Histamine is accumulated in storage vesicles by the action of VMAT. Reserpine interferes with vesicular storage of histamine. Sites 3 and 4: After release histamine interacts with postsynaptic receptors. Chlorpheniramine and diphenhydramine possess H_1 antagonist properties and are used as anti-allergy medications. H_1 antagonists, such as meclizine (Bonine) and dimenhydrinate (Dramamine) have proved useful in the treatment of motion sickness and vestibular disturbances. Site 5: Activation of autoreceptors serves to regulate subsequent release of histamine. Pitolisant is a selective inverse agonist at the H_3 receptor that increases histamine release, and is used to treat excessive diurnal sleepiness. There are no therapeutic products that act as selective antagonists for H_3 receptors. Site 6: There is no recapture of synaptic histamine, and its actions are terminated by metabolism by histamine n-methyltransferase. No selective drugs are used therapeutically to inhibit this enzyme. The metabolite, n-methylhistamine is a substrate for MAO. AADC, aromatic amino acid decarboxylase; AC, adenylate cyclase; cAMP, cyclic AMP; H_{1-3}, subtypes of histamine receptor; HNMT, histamine N-methyltransferase; G, G protein; MAO, monoamine oxidase; VMAT, vesicular monoamine transporter. (Modified from Nestler E, Hyman S, Holtzman D Malenka R: *Molecular Neuropharmacology*, McGraw-Hill Companies, Inc; 2015.)

HISTAMINE

Anatomy and Functions

In brain, histamine-containing neuronal cell bodies are restricted to the tuberomammillary complex of the posterior hypothalamus. Like other biogenic amine neurons, the histaminergic system comprises long, highly divergent, mostly unmyelinated, varicose fibers that provide diffuse projections to many telencephalic, mesencephalic, and cerebellar structures. Individual neurons can innervate widely divergent areas but make few typical synaptic contacts.

Functions closely associated with histaminergic projections are the regulation of sleep and arousal, and control of feeding and appetite. In the periphery, histamine released from mast cells plays an important role in the allergic reactions, and histamine released from enterochromaffin-like cells and mast cells stimulates acid secretion in the stomach.

Life Cycle

Histamine is formed in brain primarily by decarboxylation of histidine by the enzyme histidine decarboxylase (Fig. 1.4–7), although histidine is also a substrate for AADC, which is also involved in the synthesis of the catecholamines and serotonin. The transporter protein responsible for concentrating histamine in synaptic vesicles has not been definitively established, although both reserpine and tetrabenazine, which interfere with VMAT-2 function, induce release of brain histamine. In brain, histamine is principally metabolized by histamine methyltransferase to the n-methylhistamine (tele-methylhistamine). In the periphery, histamine is also metabolized by diamine oxidase to imidazole acetaldehyde, an intermediate that is rapidly oxidized to imidazoleacetic acid. Unlike the other biogenic amines, there is no evidence for a specific plasma membrane

transporter to recapture and terminate the effects of released histamine in brain.

Receptors

Four subtypes of histamine receptor have been identified (H_1 to H_4), and all belong to the G-protein–coupled superfamily of receptors (Fig. 1.4–7). In brain, H_1 and H_2 receptors are expressed both in neurons postsynaptic to histamine terminals and also in glia. Blockade of H_1 receptors results in sedation, a side-effect inherent in some antipsychotic drugs, such as clozapine (Clozaril), antidepressant drugs, such as doxepin (Deptran) and allergy medications, such as diphenhydramine (Benadryl) that inhibit H_1 receptors. In fact, the sedative property of doxepin is used therapeutically to alleviate the symptoms of primary insomnia. To overcome the sedating properties of H_1 antagonists in treating allergic reactions, newer nonsedating antihistamines, such as loratadine (Claritin) have been designed that do not cross the BBB. H_1 antagonists, such as meclizine (Antivert) and dimenhydrinate (Dramamine) are also effective in the treatment of motion sickness and vestibular disturbances such as Meniere disease.

H_2 receptors are expressed in neurons postsynaptic to histamine terminals. Compared with H_1 receptors, much less is known of the function of H_2 receptors, partly due to the inadequacy of specific agonists and antagonists that penetrate the BBB. H_2 antagonists that do not cross the BBB, such as cimetidine (Tagamet) and ranitidine (Zantac) are used to lower secretion of gastric acid, and so reduce the pain and hasten the healing of gastric ulcers.

H_3 receptors are located presynaptically on histamine neurons where they serve as autoreceptors to inhibit the release of histamine. In addition, H_3 receptors are present as heteroreceptors on terminals of other neurotransmitters to regulate their function. Recently a selective inverse agonist at the H_3 receptor, pitolisant, has been shown to be effective at countering excessive diurnal sleepiness in narcolepsy, Parkinson disease, and obstructive sleep apnea. Antagonists and inverse agonists for the H_3 receptor have potential as treatments for cognitive disorders.

H_4 receptors are present at low levels in brain, but are expressed in the periphery at several sites, including spleen, bone marrow, and leucocytes and are linked with regulation of the immune system. Increasing attention has been directed toward therapeutic modification of H_4 receptor signaling in disorders such as allergy, asthma, chronic pruritus, and autoimmune diseases.

REFERENCES

Artigas F. Serotonin receptors involved in antidepressant effects. *Pharmacol Ther*. 2013;137:119–131.

*Beaulieu JM, Gainetdinov RR. The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacol Rev*. 2011;63:182–217.

Bernstein AI, Stout KA, Miller GW. The vesicular monoamine transporter 2: an underexplored pharmacological target. *Neurochem Int*. 2014;73:89–97.

Black SA, Rylett RJ. Choline transporter CHT regulation and function in cholinergic neurons. *Cent Nerv Syst Agents Med Chem*. 2012;12:114–121.

Buckley JS, Salpeter SR. A risk-benefit assessment of dementia medications: systematic review of the evidence. *Drugs Aging*. 2015;32:453–467.

Carboni E, Silvagni A. Dopamine reuptake by norepinephrine neurons: exception or rule?. *Crit Rev Neurobiol*. 2004;16:121–128.

Cools R, D'Esposito M. Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biol Psychiatry*. 2011;69:e113–e125.

Deutch A, Roth R. Pharmacology and biochemistry of synaptic transmission. From molecules to networks. J. Bryne, R. Heidelberger, M. Waxham. New York: Academic Press; 207–237.

Dickinson D, Elvevag B. Genes, cognition and brain through a COMT lens. *Neuroscience*. 2009;164:72–87.

Eisenhofer G, Kopin IJ, Goldstein DS. Catecholamine metabolism: a contemporary view with implications for physiology and medicine. *Pharmacol Rev*. 2004;56:331–349.

Fernstrom JD. Large neutral amino acids: dietary effects on brain neurochemistry and function. *Amino Acids*. 2013;45:419–430.

Freedman R. Alpha7-nicotinic acetylcholine receptor agonists for cognitive enhancement in schizophrenia. *Annu Rev Med*. 2014;65:245–261.

Gasbarri A, Pompili A. Serotonergic 5-HT7 receptors and cognition. *Rev Neurosci*. 2014;25:311–323.

Gowrishankar R, Hahn MK, Blakely RD. Good riddance to dopamine: roles for the dopamine transporter in synaptic function and dopamine-associated brain disorders. *Neurochem Int*. 2014;73:42–48.

*Iversen LL, Iversen SD, Bloom FE, Roth RH. *Introduction to Neuropsychopharmacology*. New York: Oxford; 2009.

McGuffin P, Alsabban S, Uher R. The truth about genetic variation in the serotonin transporter gene and response to stress and medication. *Br J Psychiatry*. 2011;198:424–427.

*Meltzer HY, Massey BW, Horiguchi M. Serotonin receptors as targets for drugs useful to treat psychosis and cognitive impairment in schizophrenia. *Curr Pharm Biotechnol*. 2012;13:1572–1586.

Naoi M, Riederer P, Maruyama W. Modulation of monoamine oxidase (MAO) expression in neuropsychiatric disorders: genetic and environmental factors involved in type A MAO expression. *J Neural Transm*. 2015.

*Nestler E, Hyman S, Holtzman D, Malenka R. *Molecular Neuropharmacology,* McGraw-Hill Companies, Inc; 2015.

*Olivier B. Serotonin: a never-ending story. *Eur J Pharmacol.* 2015;753:2–18.

Panula P, Nuutinen S. The histaminergic network in the brain: basic organization and role in disease. *Nat Rev Neurosci.* 2013;14:472–487.

Picciotto MR, Higley MJ, Mineur YS. Acetylcholine as a neuromodulator: cholinergic signaling shapes nervous system function and behavior. *Neuron.* 2012;76:116–129.

Roeper J. Dissecting the diversity of midbrain dopamine neurons. *Trends Neurosci.* 2013;36:336–342.

Sara SJ. The locus coeruleus and noradrenergic modulation of cognition. *Nat Rev Neurosci.* 2009;10:211–223.

Seamans JK, Yang CR. The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol.* 2004;74:1–58.

Shan L, Dauvilliers Y, Siegel JM. Interactions of the histamine and hypocretin systems in CNS disorders. *Nat Rev Neurol.* 2015;11:401–413.

Tohen M, Case M, Trivedi MH, Thase ME, Burke SJ, Durell TM. Olanzapine/fluoxetine combination in patients with treatment-resistant depression: rapid onset of therapeutic response and its predictive value for subsequent overall response in a pooled analysis of 5 studies. *J Clin Psychiatry*. 2010;71:451–462.

Witte AV, Floel A. Effects of COMT polymorphisms on brain function and behavior in health and disease. *Brain Res Bull.* 2012;88:418–428.

Zimmerman G, Soreq H. Termination and beyond: acetylcholinesterase as a modulator of synaptic transmission. *Cell Tissue Res.* 2006;326:655–669.

▲ 1.5 Amino Acid Neurotransmitters

JOSEPH T. COYLE, M.D.

For nearly 70 years, biogenic amines dominated thinking about the role of neurotransmitters in the pathophysiology of psychiatric disorders. However, over the past 15 years, evidence has accumulated from postmortem, brain imaging, and genetic studies that the amino acid neurotransmitters, in particular glutamic acid and GABA, play an important, if not central, role in the pathophysiology of a broad range of psychiatric disorders including schizophrenia, bipolar disorder, major depression, Alzheimer disease, and anxiety disorders. This is not entirely surprising given the fact that virtually every neuron in the CNS is innervated by GABAergic and glutamatergic neurons. Consistent with this, the concentrations of synaptic GABA and glutamate are in the millimolar range whereas biogenic amine and peptide neurotransmitters are in the micromolar range or lower. The purpose of this chapter is to review the current understanding of amino acid neurotransmitters and to address their potential involvement in psychiatric disorders.

GLUTAMIC ACID

Glutamate mediates fast excitatory neurotransmission in the brain and is the neurotransmitter for approximately 80 percent of brain synapses, particularly those associated with dendritic spines. The repolarization of neuronal membranes that have been depolarized by glutamatergic neurotransmission may account for as much as 80 percent of the energy expenditure in the brain. The concentration of glutamate in brain is 10 mM, the highest of all amino acids, of which approximately 20 percent represents the neurotransmitter pool of glutamate.

The postsynaptic effects of glutamate are mediated by two families of receptors. The first are the glutamate-gated cation channels that are responsible for fast neurotransmission. The second type of glutamate receptor is the metabotropic glutamate receptor (mGluR), which are GPCRs like α-adrenergic receptors and dopamine receptors. The mGluRs primarily modulate glutamatergic neurotransmission.

Synthesis of Glutamate

Given the excitatory effects of glutamate, it is not surprising that it is excluded from the brain by the BBB. Thus, glutamate in the brain must be synthesized *de novo* from glucose through the tricarboxylic acid cycle, which generates α-ketoglutarate. The α-ketoglutarate receives an amino group via a transaminase reaction, converting it to glutamic acid.

Glutamate is in equilibrium with α-ketoglutarate, and virtually all glucose entering the brain is cycled through glutamic acid. The portion of glutamate dedicated to neurotransmission (approximately 20 percent) is actively sequestered in storage vessels by the vesicular glutamate transporter.

A second metabolic pathway is particularly important for replenishing synaptic glutamate. This pathway exploits the intimate relationship between the glutamatergic synapse (i.e., the synaptic bouton and the postsynaptic spine) and the astrocytic endfeet that envelop the synapse. It is the astrocyte and not the glutamatergic terminal that expresses glutamate transporters (EAAT1 and 2) that remove glutamate from the synapse, thereby terminating its action. Within the astrocyte, glutamine synthetase, a cytosolic adenosine triphosphate (ATP)-dependent enzyme, catalyzes the conversion of glutamate to generate glutamine. Glutamine synthetase is expressed in glia but not neurons. Glutamine is then released by the astrocyte and taken up by the glutamatergic terminal where it is converted back to glutamate by phosphate-activated glutaminase, a mitochondrial enzyme. This process is known as the "glutamine cycle" and accounts for approximately 40 percent of glutamate turnover.

Major Glutamatergic Pathways in the Brain

All primary sensory afferent systems appear to use glutamate as their neurotransmitter including retinal ganglion cells, cochlear cells, trigeminal nerve, and spinal afferents. The thalamocortical projections that distribute afferent information broadly to the cortex are glutamatergic. The pyramidal neurons of the corticolimbic regions, the major source of intrinsic, associational, and efferent excitatory projections from the cortex are glutamatergic. A temporal lobe circuit that figures importantly in the development of new memories is a series of four glutamatergic synapses: The perforant path innervates the hippocampal granule cells that innervate CA3 pyramidal cells that innervate CA1 pyramidal cells. The climbing fibers innervating the cerebellar cortex are glutamatergic as well as the corticospinal tracks.

Ionotropic Glutamate Receptors

Three families of ionotropic glutamate receptors have been identified on the basis of selective activation by conformationally restricted or synthetic analogues of glutamate. These include α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA), kainic acid (KA), and *N*-methyl-D-aspartic acid (NMDA) receptors (Fig. 1.5–1). Subsequent cloning revealed 16 mammalian genes that encode structurally related proteins, which represent subunits that assemble into functional receptors. Glutamate-gated ion channel receptors appear to be tetramers, and subunit

composition affects both the pharmacologic and the biophysical features of the receptor.

AMPA receptors mediate the depolarizing postsynaptic currents primarily responsible for excitatory neurotransmission and are broadly distributed in the CNS (Fig. 1.5–2). The AMPA receptor family consists of four subunits of GluR1 to GluR4. However, additional complexities affect AMPA receptor function. Alternative splicing of a 115 bp cassette for GluR1 to GluR4 messenger ribonucleic acid (mRNA) results in two forms (*flip* and *flop*) that give rise to receptors that differ in desensitization rate and regional distribution in the brain. In the second transmembrane domain, GluR1, 3, and 4 have a glutamine (Q) residue that results in high Ca^{2+} conductance whereas GluR2 has an arginine (R) in this position that severely restricts Ca^{2+} passage and conducts only Na^{+}, mRNA editing by adenosine deaminase, which converts the codon for GluR2 from an arginine to a glutamine. This radically increases the channel permeability to Ca^{2+} of AMPA receptors containing the edited GluR2 subunit. Similar mRNA editing mechanisms have been described for kainate receptors.

The kainite receptor family consists of five subunits. GluR5 to GluR7 represents subunits that form glutamate-gated cation channels. KA1 and KA2 exhibit negligible channel activity but associate with GluR5 to GluR7 to form high-affinity kainate receptors. The role of kainate receptors is less clearly defined than that of AMPA receptors, but with their presynaptic localization on glutamatergic terminals, kainite receptors reduce glutamatergic neurotransmission when activated.

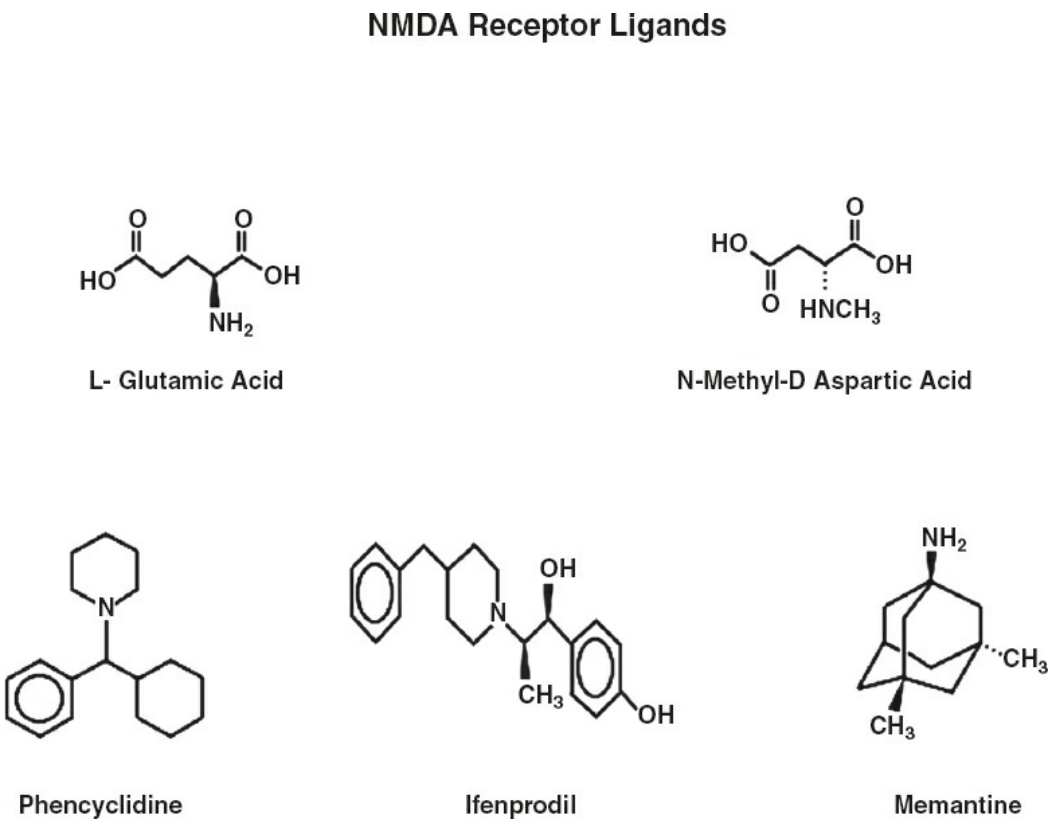

FIGURE 1.5–1. NMDA receptor ligands.

Seven genes encode subunits that comprise the NMDA receptor family. The NMDA receptor has several unique features (Fig. 1.5–3). First, the channel is blocked by magnesium (Mg^{2+}) at resting membrane potential. Thus, NMDA receptors are "silent" until activated AMPA receptors have sufficiently depolarized the neuronal membrane to relieve the Mg^{2+} blockade. The NMDA receptor requires the simultaneous binding of two ligands to two separate recognition sites in order for the channel to open. On the NR1 subunit, which forms the channel, is a binding site termed the *glycine modulatory site,* for which glycine and D-serine are the endogenous ligands. On the second nonchannel subunit (NR2A-D) is the binding site for the neurotransmitter glutamate. Unless the glycine modulatory site is occupied, glutamate cannot open the channel. The NMDA receptor has been described as a coincidence detector because these three events must occur simultaneously for the channel to open. Thus, it becomes functional only when a sufficient amount of presynaptic glutamate has been released to saturate the receptor, glycine and/or D-serine are released, and the synaptic membrane is sufficiently depolarized to remove the Mg^{2+} blockade. Typically, for this to occur, multiple, converging glutamatergic afferents fire simultaneously.

The NMDA receptor channel is sufficiently large that Ca^{2+} readily passes through it. Intracellularly, Ca^{2+} activates a number of kinases that ultimately affect gene expression in the neuron. For example, the

immediate-early gene cFos is a sensitive surrogate for NMDA receptor activation. The type of NR2 subunit affects the pharmacology and biophysics of the NMDA receptor. For example, NR2B-containing receptors are much more Ca^{2+} permeable than NR2A-containing receptors. NR2A-containing receptors are expressed primarily in corticolimbic regions in the mature brain. NR2B is expressed at high levels in the immature cortex and decreases with maturation. NR2C is expressed primarily in the cerebellum, and NR2D is localized predominately to the cerebellum and midbrain brainstem.

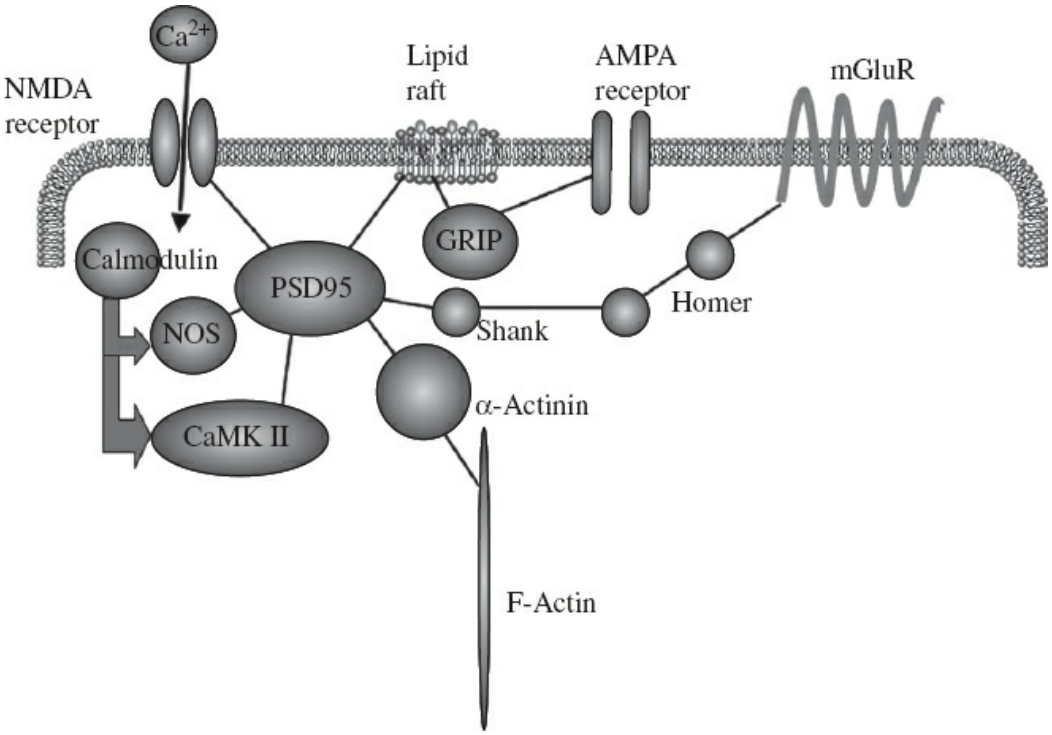

FIGURE 1.5–2. The postsynaptic density of the excitatory synapse. The NMDA receptor is bound to the principal organizing protein, PSD-95. Effector enzymes such as nitric oxide synthase (NOS) and calmodulin-activated protein kinase II (CaMKII) are also bound to PSD-95, keeping them in close proximity to the Ca^{2+} permeable NMDA receptor channel. PSD-95 is connected to the intracellular cytoskeletal protein, F-action by α-actinin. PSD-95 also is attached to neuronal membrane lipid rafts, which indirectly links it to the AMPA receptor through glutamate receptor interacting protein (GRIP). Finally, postsynaptic group I mGluR receptors are linked to PSD-95 by the scaffolding proteins Shank and Homer.

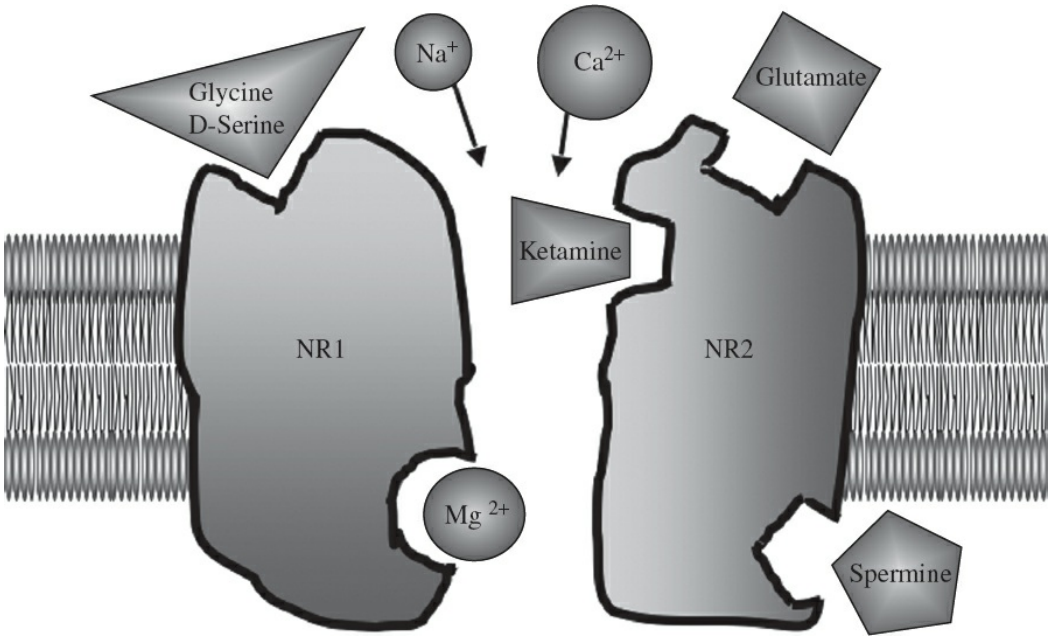

FIGURE 1.5–3. Schematic representation of the NMDA receptor. The NMDA receptor is a heterotetramer composed of the NR1 subunit that comprises the channel and the NR2 subunit, which contains the ligand binding site for the agonist, glutamate. The channel is blocked at the resting membrane potential by Mg^{2+}. The glycine modulatory site on the NR1 subunit must be occupied by the endogenous agonists, D-serine or glycine, for the channel to open. The channel accommodates both Na^{+} and Ca^{2+}. A polyamine site positively modulates the receptor. Within the channel is the binding site for dissociative anesthetics, which are use-dependent, noncompetitive inhibitors of the NMDA receptor.

Given their prominent role in learning and in excitotoxicity (see below), it is not surprising that NMDA receptors are among the most tightly regulated of neurotransmitter receptors. As described above, two separate ligands must be bound to two distinct subunits on the NMDA receptor for it to function. In addition, there are binding sites for Zn^{2+} and H^{+} that inhibit ion flux. A polyamine site enhances channel opening. Furthermore, the channel is sensitive to its redox state, which also affects ion currents. Subunit composition affects responses to modulators. For example, the NR2A subunit is much more sensitive to inhibition by Zn^{2+} whereas the NR2B is differentially more sensitive to the polyamine site antagonist ifenprodil. The influx of Ca^{2+} via the NMDA receptor activates calmodulin, which then binds to the C-terminus of NR1 and reduces channel opening frequency and duration.

Metabotropic Glutamate Receptors

These receptors are so designated because their effects are mediated by G-proteins. All mGluRs are activated by glutamate although their sensitivities vary remarkably. To date, eight mGluRs have been cloned. These genes encode seven membrane-spanning proteins that are members of the superfamily of GPCRs. They are sub-grouped into three classes based upon

amino acid sequence homology, agonist pharmacology, and signal transduction pathway utilized. Group I mGluRs, which includes mGluR1 and 5, activate PLC, presumably through G_Q, group II includes mGluR2 and 3, and group III includes mGluR4, 6, 7, and 8. Group II and III mGluRs inhibit adenylyl cyclase through G_i protein. The abundant neuropeptide *N*-acetylaspartylglutamate is a specific agonist at mGluR3.

mGluRs located postsynaptically modulate a number of channels and receptors. All three groups inhibit L-type voltage-dependent calcium channels, and groups I and II inhibit N-type calcium channels. In addition, mGluRs are reported to close voltage-dependent K^+ channels, thereby slowing depolarization and reducing excitability. Presynaptic mGluRs on both GABAergic and glutamatergic terminals inhibit neurotransmitter release, possibly by inhibiting the P/Q-type calcium channel.

Postsynaptic Density

Considerable advances have been made in characterizing the organization and dynamics of glutamate receptors at the postsynaptic density. The PSD is a multiprotein complex that contains scaffolding proteins, cell adhesion molecules, and proteins for intercellular signaling pathways. A major scaffolding protein is PSD-95 (an acronym for *PSD* protein with a molecular weight of 95 kDa). PSD-95 contains several regions that bind other proteins. There are three PDZ domains (an acronym for *PSD-95/disc large/zona occludens-1*). The PDZ domains contain approximately 90 amino acids that bind the C-termini of proteins with complementary amino acid sequences. Neuroligin binds to the PDZ and extends into the synaptic cleft to bind to β-neurexin, which is anchored to the presynaptic component of the synapse. This arrangement stabilizes the synapse by connecting pre- and postsynaptic components. Two N-terminal cysteines of PSD-95 bind palmitic acid, which links the protein to lipid rafts in the plasma membrane. The NR2 subunit of the NMDA receptor binds to the PDZ domain. PSD-95 also binds α-actin, which is bound to filamentous actin, an important component of the cytoskeletal complex in the dendritic spine.

In contrast to NMDA receptors, AMPA receptors do not appear to bind directly to PSD-95 but are associated with it indirectly by binding to intermediary proteins that bind to PSD-95. These include *glutamate receptor interacting protein* (GRIP), *protein interacting with C-kinase1* (PICK1), and *synapse associated protein of 97 kDa* (SAP-97). In addition, there are the *transmembrane AMPA receptor regulatory proteins* (TARPs), which are involved in transporting AMPA receptors to and intercalating them within the PSD. Whereas the number of NMDA receptors in mature synapses tend to be relatively constant, the number of AMPA receptors in the postsynaptic densities varies tremendously. In fact,

there are some postsynaptic densities that contain no AMPA receptors but do contain NMDA receptors. They are referred to as "silent synapses" because glutamate has no excitatory effects since the NMDA receptors are inactivated at the resting membrane potential. With the exception of mGluR7, postsynaptic mGluRs are tethered to the periphery of the PSD by binding to two scaffolding proteins *homer* and *shank,* the latter of which binds to PSD-95. Notably, mutations in proteins that comprise the PSD including *neurexin, neuroligin,* and *shank* have been implicated in autism.

The Role of Astrocytes

Specialized endfeet of the astrocyte surround glutamatergic synapses. The astrocyte expresses the two Na^+-dependent glutamate transporters that play the primary role in removing glutamate from the synapse, thereby terminating its action: EAAT1 and EAAT2 (*excitatory amino acid transporter*). The neuronal glutamate transporter, EAAT3, is expressed in upper motor neurons whereas EAAT4 is expressed primarily in cerebellar Purkinje cells and EAAT5 in retina. Mice homozygous for null mutations of either EAAT1 or EAAT2 exhibit elevated extracellular glutamate and excitotoxic neurodegeneration. Notably, several studies have described the loss of EAAT2 protein and transport activity in the ventral horn in amyotrophic lateral sclerosis (ALS).

The astrocytes express AMPA receptors so that they can monitor synaptic glutamate release. GlyT1, which maintains subsaturating concentrations of glycine in the synapse, is expressed on the astrocyte plasma membrane. GlyT1 transports three Na^+ out for each molecule of glycine transported into the astrocyte. This stoichiometry results in a robust reversal of the direction of transport when glutamate released in the synapse activates the AMPA receptors on the astrocyte, thus depolarizing the astrocyte. Thus, glycine release in the synapse by GlyT1 is coordinated with glutamatergic neurotransmission. D-serine, which is found primarily in forebrain, is synthesized by serine racemase (SR), which is expressed primarily in pyramidal neurons and to a lesser extent GABAergic neurons. D-Serine levels are also determined by D-amino acid oxidase (DAAO) with low D-serine levels in the cerebellum and brainstem where DAAO expression is high, and high D-serine levels are found in corticolimbic brain regions where DAAO expression is quite low. In contrast, the expression of GlyT1 is highest in the cerebellum and brainstem. This distribution suggests that D-serine is the primary modulator of the NMDA receptor in the forebrain whereas glycine is more prominent in the brainstem and cerebellum. Nevertheless, GlyT1 inhibitors have pro-cognitive effects in experimental animals and are being tested clinically for the treatment of cognitive impairment and negative symptoms in schizophrenia.

Plasticity in Glutamatergic Neurotransmission

Hebb postulated that learning and memory involved use-dependent changes in synaptic efficacy. Neurophysiological studies, originally exploiting the identifiable glutamatergic Schaffer collateral synapse on the hippocampal CA1 pyramidal cell, showed that a brief period of intense stimulation of the Schaffer collateral (100 Hz) resulted in a subsequent persistent increase in the efficacy of synaptic neurotransmission at these synapses. This phenomenon is known as *long-term potentiation* (LTP) and is quite widespread with regard to glutamatergic synapses. In contrast, a period of stable low-frequency stimulation of the glutamatergic axons results in a persistent reduced efficacy of synaptic neurotransmission, a phenomenon known as *long-term depression* (LTD). LTP in these hippocampal studies requires the activation of NMDA receptors as demonstrated by the fact that it is blocked by NMDA receptor antagonists such as the dissociative anesthetics ketamine and phencyclidine (PCP). Conditions resulting in blockade of LTP in the hippocampus are associated with impairments in the acquisition of new memories. The nature of these plastic changes has been the focus of intense research. The persistent changes in synaptic efficacy in LTP and LTD result from the insertion (LTP) or removal (LTD) of AMPA receptors from the postsynaptic densities of affected synapses. Thus, in contrast to the NMDA receptors, the AMPA receptors are quite dynamic, and their synaptic function is controlled through trafficking.

The extinction of conditioned fear has been shown to be an active process mediated by the activation of NMDA receptors in the amygdala. Treatment of rats with NMDA receptor antagonists prevents the extinction of conditioned fear whereas treatment with the glycine modulatory site partial agonist D-cycloserine facilitates the extinction of conditioned fear. (D-Cycloserine is an antibiotic used to treat tuberculosis that has 50 percent of the efficacy of glycine at the NMDA receptor.) To determine whether the phenomenon generalizes to humans, patients with acrophobia were administered either placebo or a single dose of D-cycloserine along with cognitive behavioral therapy (CBT). D-Cycloserine plus CBT resulted in a highly significant reduction in acrophobic symptoms that persisted for at least 3 months as compared to placebo plus CBT. Other placebo-controlled clinical trials support the notion that D-cycloserine is a robust enhancer of CBT, suggesting that pharmacologically augmenting neural plasticity may be used to bolster psychological interventions.

Glutamate-mediated synaptic plasticity is not only functional it is also structural. Dendritic spines are dynamic appendages. Such dynamics have cleverly been demonstrated in real time in vitro and in vivo in the brains of mice in which a gene from jellyfish that encodes a green fluorescent protein is inserted into the mouse genome in a manner so that only glutamatergic

neurons become fluorescent. Persistent activation of NMDA receptors results in spine maturation from long and skinny to fat and stubby and even the elaboration of new spines. This is mediated in part by the influx of calcium through the activated NMDA receptors. Such rapid structural changes in the synaptic spine reflect the fact that protein synthesis (i.e., translation) occurs within individual spines.

Fragile X mental retardation protein (FMRP), which is deficient in individuals with fragile X syndrome, appears to be synthesized locally within the spine during times of NMDA receptor activation and also plays a role in transporting specific mRNAs to the spine for translation. Notably, mice in which the FMRP gene has been inactivated through a null mutation as well as patients with Fragile X syndrome have fewer dendritic spines, the preponderance of which have an immature morphology. Loss of FMRP exaggerates responses of mGluR5, which stimulates dendritic protein synthesis, and treatment with an mGluR5 antagonist reverses the fragile-X-like phenotype in mice with the FMRP gene inactivated.

Excitotoxicity

In the early 1970s, it was shown that the systemic administration of large amounts of monosodium glutamate to immature animals resulted in the degeneration of neurons in brain regions where the blood–brain barrier was deficient. Subsequent studies showed that the direct injection of ionotropic glutamate receptor agonists such as kainic acid, ibotenic acid, and NMDA caused a pattern of neuronal degeneration that affected neurons with their cell bodies in proximity to the injection site but spared axons passing through the area arising from distant neurons. Persistent and overwhelming activation of AMPA/kainate receptors and NMDA receptors causes a tremendous influx of Na^+ and Ca^{2+} and a secondary influx of H_2O. The resulting acute cellular edema causes a necrotic cell death. At sites more distant from the injection, the persistent elevation of Ca^{2+} disrupts the mitochondria, which release cytochrome C and activate caspases, resulting in programmed cell death (apoptosis).

The neuronal degeneration occurring after ischemic stroke is the result of excitotoxicity. Local hypoxia due to ischemia results in a cessation of ATP production, causing the collapse of the sodium gradient across the neuronal membrane and astroglial membrane. As a consequence, the vector of the sodium-dependent glutamate transporters is reversed, causing a massive release of glutamate. Although a number of drugs that block the ionotropic glutamate receptors or the downstream events caused by their overstimulation have proved effective in reducing the amount of neuronal damage in animal models of stroke, no agent has yet proved effective in clinical trials.

Excitotoxicity has also been implicated in the proximate cause of

neuronal degeneration in Alzheimer disease. Most evidence points to the toxic consequences of aggregates of β-amyloid, especially β-amyloid$_{1-42}$. The β-amyloid fibrils depolarize neurons, resulting in loss of the Mg^{2+} block and enhanced NMDA receptor sensitivity to glutamate. The fibrils also impair glutamate transport into astrocytes, thereby increasing the extracellular concentration of glutamate. β-Amyloid directly promotes oxidative stress through inflammation that further contributes to neuronal vulnerability to glutamate. Thus, several mechanisms contribute to neuronal vulnerability to NMDA-receptor-mediated excitotoxicity in Alzheimer disease. Memantine, a recently approved treatment for mild to moderate Alzheimer disease, is a weak noncompetitive inhibitor of NMDA receptors. It reduces tonic sensitivity of NMDA receptors to excitotoxicity but does not interfere with "phasic" neurotransmission, thereby attenuating neuronal degeneration in Alzheimer disease.

Inhibitory Amino Acids: GABA

GABA is the major inhibitory neurotransmitter in the brain where it is broadly distributed and occurs in millimolar concentrations. In view of its physiological effects and distributions, it is not surprising that the dysfunction of GABAergic neurotransmission has been implicated in a broad range of neuropsychiatric disorders including anxiety disorders, schizophrenia, alcohol dependence, and seizure disorders. Chemically, GABA differs from glutamic acid, the major excitatory neurotransmitter, simply by the removal of a single carboxy group from the latter.

GABA is synthesized from glutamic acid by GAD, which catalyzes the removal of the α-carboxyl group. In the CNS, the expression of GAD appears to be restricted to GABAergic neurons although in the periphery it is expressed in pancreatic islet cells. Two distinct but related genes encode GAD. GAD65 is localized to nerve terminals where it is responsible for synthesizing GABA that is concentrated in the synaptic vesicles. Consistent with its role in fast inhibitory neurotransmission, mice homozygous for a null mutation of GAD65 have an elevated risk for seizures. GAD67 appears to be the primary source for neuronal GABA because mice homozygous for a null mutation of GAD67 die at birth, have a cleft pallet, and exhibit major reductions in brain GABA.

GABA is catabolized by GABA transaminase (GABA-T) to yield succinic semialdehyde. Transamination generally occurs when the parent compound, α-ketoglutarate, is present to receive the amino group, thereby regenerating glutamic acid. Succinic semialdehyde is oxidized by succinic semialdehyde dehydrogenase (SSADH) into succinic acid, which re-enters the Krebs cycle. GABA-T is a cell surface, membrane-bound enzyme, which is oriented toward the extracellular compartment and is expressed by both neurons and glia. As would be anticipated, drugs that inhibit the catabolism

of GABA have anticonvulsant properties. One of the mechanisms of action of VPA is the competitive inhibition of GABA-T. γ-Vinyl-GABA is a suicide substrate inhibitor of GABA-T that is used as an anticonvulsant in Europe (vigabatrin [Sabril]).

The synaptic action of GABA is also terminated by high-affinity transport back into the presynaptic terminal as well as into astrocytes. Four genetically distinct GABA high-affinity transporters have been identified with differing kinetic and pharmacological characteristics. They all share homology with other neurotransmitter transporters with the characteristic of 12 membrane-spanning domains. The active transport is driven by the sodium gradient so that upon depolarization, transportation of GABA out of the neuron is favored. GABA transported into astrocytes is catabolized by GABA-T and ultimately converted to glutamic acid and then to glutamine, which is transported back into the presynaptic terminal for GABA synthesis. Tiagabine (Gabitril) is a potent GABA transport inhibitor that is used to treat epilepsy.

Anatomy of GABAergic Systems

In the corticolimbic regions of the brain GABA is localized to the intrinsic (i.e., local circuit) neurons. In the columnar organization of the cerebral cortex, the GABAergic neurons provide the outer boundaries of the column with inwardly directed axons. While the GABAergic interneurons comprise a minority of cortical neurons (15 to 25 percent), they exert a profound degree of inhibition of the activity of the glutamatergic pyramidal cells. The remarkable efficacy of inhibition reflects two neuroanatomical features of GABAergic synapses, which are concentrated on the shafts of spines to mitigate glutamatergic depolarization and on the neuronal cell body and proximal axon to restrict the generation of action potentials. In the cortex the GABAergic interneurons are the primary site of colocalization of neuropeptides. These include CCK, dynorphin, neuropeptide Y (NPY), somatostatin, substance P, and vasoactive intestinal peptide (VIP).

In the striatum, GABAergic neurons project directly to the *substantia nigra pars reticulata,* which regulates dopaminergic neuronal activity. In addition, there are striatal GABAergic neurons that project to the *globus pallidus* to synapse on pallidal–subthalamic GABAergic neurons that regulate the excitatory output from the subthalamic nucleus. In the cerebellum, GABAergic Purkinje cells are its main efferent system.

$GABA_A$ Receptors

$GABA_A$ receptors are distributed throughout the brain. The $GABA_A$ receptor complex, when activated, mediates an increase in membrane conductance with an equilibrium potential near the resting membrane potential of −70 mV (Fig. 1.5–4). In the mature neuron, this typically

results in an influx of Cl^-, causing membrane hyperpolarization. Hyperpolarization is inhibitory because it increases the threshold for generating an action potential. In immature neurons, which have unusually high levels of intracellular Cl^-, activating the $GABA_A$ receptor can counterintuitively cause depolarization. For this reason, anticonvulsants that act by enhancing $GABA_A$ receptor activity may actually exacerbate seizures in the neonatal period.

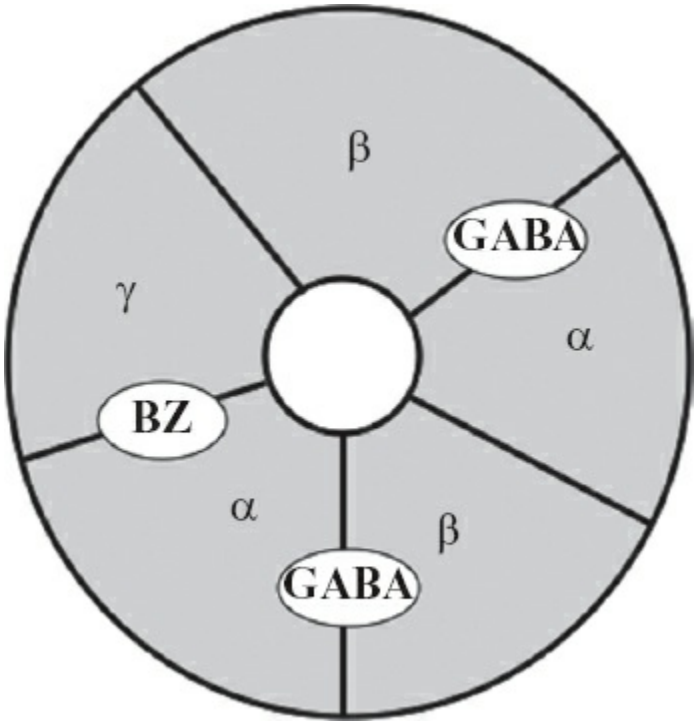

FIGURE 1.5–4. Schematic representation of the $GABA_A$ receptor. The receptor–channel complex is a heteropentamer. The GABA binding site is at the interface of the α and β subunits. The benzodiazepine binding site is at the interface between the γ and α subunits.

The $GABA_A$ receptor subunits exhibit sequence homology to a larger family of ligand-gated channels including the nicotinic acetylcholine receptor and the glycine receptor. At least 19 distinct but closely related genes have been identified that encode $GABA_A$ receptor subunits. Each subunit contains four α-helical membrane-spanning domains, the sequences of which are highly conserved among the subunits. The receptor complex is a heteropentamer. The ligand binding site is formed at the interface between the α- and the β-subunits. The subunit composition affects the biophysical and pharmacological characteristics of the receptor. Different subunit-containing $GABA_A$ receptor complexes are expressed at different stages of development as well as in different regions of the brain.

The pharmacology of $GABA_A$ receptors is particularly rich. A component of the extracts of the psychoactive mushroom *Amanita muscaria* is muscimol, which is a direct agonist at the $GABA_A$ receptor. The

prototypical $GABA_A$ antagonist is bicuculline, which acts by decreasing the frequency and duration of channel opening and has proconvulsant effects. Picrotoxin, another proconvulsant, acts by blocking the chloride channel.

Reminiscent of the NMDA receptor, the $GABA_A$ receptor complex is noteworthy for multiple allosteric modulatory interactions. These include benzodiazepines, barbiturates, general anesthetics, ethanol, and neurosteroids. Benzodiazepines bind to a distinct site in the $GABA_A$ receptor complex and allosterically increase the frequency of channel opening in response to GABA. Therefore, benzodiazepines do not directly activate the receptor, but they enhance the phasic responses to synaptically released GABA. This indirect mechanism of action and the localization of benzodiazepine-sensitive receptors account for the lower risk of respiratory suppression for benzodiazepines as compared to that of barbiturates. The benzodiazepine site is allosterically linked to the binding sites of other modulatory ligands, which can contribute to toxic interactions. Notably, antagonists and inverse agonists for the benzodiazepine receptor have been developed that demonstrate anxiogenic effects.

Barbiturates such as phenobarbital and pentobarbital are noted for their sedative and anticonvulsant activities. Barbiturates allosterically increase the affinities of the binding sites for GABA and benzodiazepines at concentrations that are pharmacologically relevant. Barbiturates also affect channel dynamics by markedly increasing the long open state and reducing the short open state, thereby increasing Cl^- conductance and inhibition. Chemically modified analogs of progesterone and corticosterone have been shown in behavioral studies to have sedative and anxiolytic effects through their interaction with the $GABA_A$ receptor complex. They share features with barbiturates although they act at a distinctly different site. Thus, they allosterically enhance agonist binding to the receptor and increase the duration of Cl^- channel opening. A variety of behavioral effects associated with steroid administration or fluctuation of endogenous steroids and sex-specific affects of GABAergic drugs have been linked to the action of endogenous neurosteroids.

With regard to $GABA_A$ receptor antagonists, picrotoxin, like the barbiturates, alters channel dynamics but in the opposite direction by reducing long open states and favoring the briefest open state. The proconvulsant pentylenetetrazol also acts by reducing Cl^- channel permeability. Penicillin, which at high concentrations is proconvulsant, binds to the positively charged residues in the channel, thereby occluding it. As a general class, anesthetics including barbiturates, steroids, and volatile anesthetics increase chloride conductance, thereby inhibiting neurotransmissions. Amino acids in the membrane-spanning domain of the GABA receptor subunits confer sensitivity to anesthetics. The precise mechanism whereby ethanol enhances $GABA_A$ receptor function remains

unclear due to inconsistent results, suggesting that subunit composition may be important. However, recent studies suggest that ethanol increases the response of the tonic GABA-activated currents, which contain the δ-subunit and exhibit remarkably high affinity for GABA.

Recombinant DNA strategies exploiting site-directed mutagenesis have permitted the identification of sites on the specific subunits that mediate the pharmacological action of drugs such as the benzodiazepines. Removal of the binding ability for benzodiazepines has established that the α_1-subunit plays a major role in the sedative and amnestic effects of benzodiazepines whereas inactivating the benzodiazepine site on the α_2-subunit eliminates the anxiolytic effect of benzodiazepines.

$GABA_B$ Receptors

The $GABA_B$ receptors are distinguished pharmacologically from $GABA_A$ receptors by the fact that they are insensitive to the canonical $GABA_A$ receptor antagonist bicuculline and that they are potently activated by baclofen [β-(4-chlorophenyl)-γ-aminobutyric acid], which is inactive at $GABA_A$ receptors. They are members of the G-protein–coupled superfamily of receptors but are highly unusual as they are made of a dimer of two seven transmembrane-spanning subunits. $GABA_B$ receptors are widely distributed throughout the nervous system and are localized both pre- and postsynaptically. The postsynaptic $GABA_B$ receptors cause a long-lasting hyperpolarization by activating potassium channels. Presynaptically, they act as auto- and heteroreceptors to inhibit neurotransmitter release.

All $GABA_B$ receptors in the vertebrate brain are the sole products of the $GABA_B(1)$ and $GABA_B(2)$ genes (Fig. 1.5–5). Pharmacological heterogeneity among $GABA_B$ receptors reflects different isoforms resulting from splice variants. The most common variants that are conserved across species are $GABA_B(1A)$ and $GABA_B(1B)$. They exhibit different regional distributions in the brain where, for example, $GABA_B(1A)$ transcripts are confined to the granule cell layer whereas $GABA_B(1B)$ transcripts are expressed primarily in Purkinje cells. All $GABA_B$ agonists and competitive antagonists bind to the extracellular domain of the $GABA_B(1)$ subunit. The $GABA_B(2)$ subunit also has a large extracellular domain, which may be the site for allosteric modulation of the $GABA_B$ receptor. Interestingly, there is a binding site for Ca^{2+} in the ligand-binding pocket of the $GABA_B$ (1) subunit that increases the affinity for GABA. This Ca^{2+} binding site is typically saturated at physiological concentrations of Ca^{2+}.

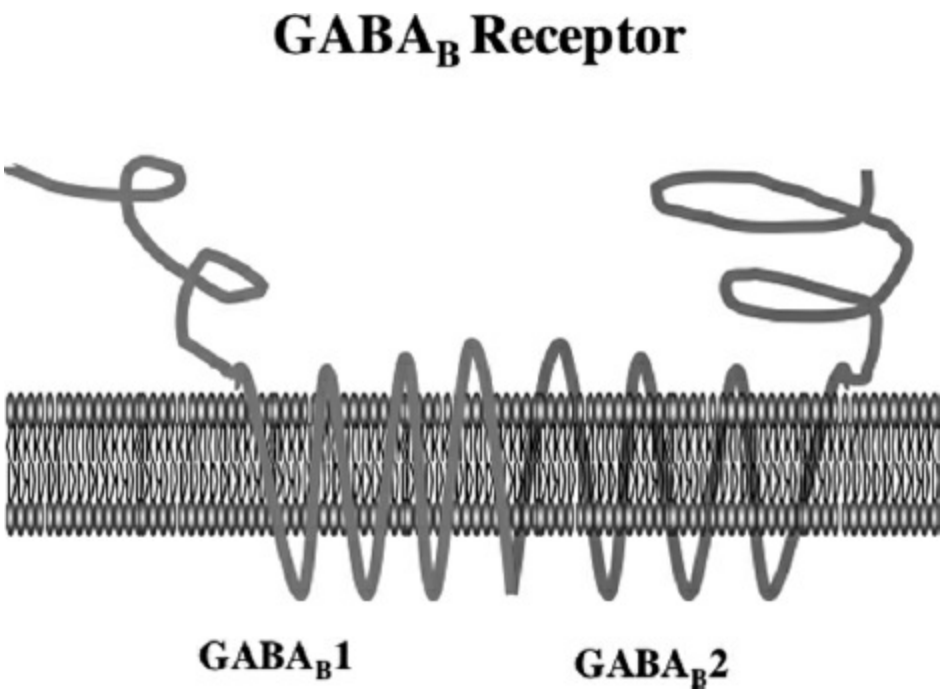

FIGURE 1.5–5. Schematic representation of the GABA$_B$ receptor. This G-protein–coupled receptor GPCR is a heterodimer of two seven-membrane-spanning GPCRs. The extensive extracellular domains have the binding sites for GABA and allosteric modulators.

γ-Hydroxybutyrate (GHB), which is approved for the treatment of narcolepsy, has been misused as a "date rape" drug because it rapidly induces deep sleep. Although high-affinity binding sites for GHB have been described in the brain, the sedative and hypnotic effects of exogenous GHB can be prevented by GABA$_B$ receptor antagonists. Furthermore, administration of GHB to GABA$_B$−/− mice resulted in no behavioral effects of GHB.

Glycine as a Neurotransmitter

Glycine is an inhibitory neurotransmitter primarily in the brainstem and spinal cord, although the expression of glycine receptor subunits in the thalamus, cortex, and hippocampus suggest a broader role. Glycine is a nonessential amino acid that is synthesized in the brain from L-serine by serine hydroxymethyltransferase. Glycine is concentrated within synaptic vesicles by H^+-dependent vesicular inhibitory amino acid transporter (VIAAT or VGAT), which also transports GABA. Termination of the synaptic action of glycine is through reuptake into the presynaptic terminal by the glycine transporter II (GlyT2), which is quite distinct from GlyT1 that is expressed in astrocytes and modulates NMDA receptor function.

The inhibitory effects of glycine are mediated by a ligand-gated chloride channel, which can also respond to β-alanine, taurine, L-alanine, L-serine, and proline but not to GABA. The canonical antagonist for the glycine receptor is the plant alkaloid strychnine. The receptor was first identified

through the specific binding of [3H]strychnine. [3H]Glycine binds to two sites: One that is displaceable by strychnine and represents the glycine A receptor and a second that is insensitive to strychnine and is designated the glycine B receptor, representing the glycine modulatory site on the NMDA receptor.

The glycine A receptor is a macromolecular complex of approximately 250 kDa that is comprised of five subunits surrounding a central pore. There are two subunits with a high degree of homology: The 48-kDa α-subunit and the 58-kDa β-subunit. The subunits display a structural similarity to other members of this ion channel family with four hydrophobic domains that span the lipid bilayer in α helices. The binding site for both glycine and strychnine is located on the α-subunit. There are four genes that encode α-subunits, but only one gene encodes for the β-subunit. Interestingly, the β-subunit is expressed fairly abundantly in rostral brain regions that exhibit no [3H]strychnine binding or α-subunit expression.

Hyperekplexia is a disorder due to mutations in genes encoding components of the glycinergic synapse. It is characterized by stiffness and excessive startle in infancy that subsides with maturation. Mutations causing hyperekplexia have been described in the α-subunit (*GLRA1*) and in the β-subunit (*GLRB*) of the glycine receptor but also in GlyT2 (*SLC6A5*).

Neuropsychiatric Implications of Amino Acid Transmitters

Schizophrenia. Evidence accumulating from postmortem, pharmacological, and genetic studies is shifting the focus of the pathophysiology of schizophrenia from dopamine to glutamate and GABA. Indeed, after the use of dopamine D_2 receptor antagonists as the sole treatment of schizophrenia for the last 50 years, over two-thirds of the treated patients remain substantially disabled. Early postmortem studies indicated a reduction in the activity of GAD in the cortex in patients with schizophrenia as compared to suitable controls. With the advent of immunocytochemistry and gene expression techniques, it has been possible to more precisely define the GABAergic deficit in schizophrenia. It appears that the parvalbumin-positive GABAergic interneurons in the intermediate layers of the cortex bear the brunt of the pathology, which includes reduced expression of GAD67, parvalbumin, and the GABA transporter (GAT). The finding that $GABA_A$ receptors are upregulated, as measured by autoradiography or with antibodies, supports the theory that these changes reflect hypofunction of the presynaptic GABAergic neurons. These particular GABAergic interneurons, which include the chandelier cells, play an important role in negative feedback inhibition to the pyramidal cells in the cortex. In spite of this highly reproducible neuropathology, genes

related to GABAergic function have not figured prominently in genome-wide searches, suggesting that GABAergic deficits may be a downstream consequence of some more proximal genetic defects.

The theory that hypofunction of NMDA receptors is an etiologic factor in schizophrenia initially arose from the observation that PCP and related dissociative anesthetics that block NMDA receptors produce a syndrome that can be indistinguishable from schizophrenia (Fig. 1.5–6). Dissociative anesthetics are so named because they prevent the acquisition of new memories without diminishing the apparent level of consciousness. In fact, under laboratory conditions, low-dose infusion of ketamine can produce the positive symptoms, negative symptoms, and specific cognitive deficits associated with schizophrenia in clear consciousness. Subsequent studies indicated that low-dose ketamine can also cause enhanced amphetamine-induced subcortical dopamine release as is observed in schizophrenia as well as abnormal cortical event-related potentials (ERPs) and disruption of prepulse inhibition in experimental animals.

Early gene studies in schizophrenia were limited by biased selection of "candidate" genes and small numbers of subjects, resulting in contradictory results. Recently, a large genome wide association study comparing nearly 40,000 subjects with schizophrenia to a hundred thousand controls identified 108 sites on the genome, which attained the stringent level of genome-wide significance (5×10^{-8}). Notably, none contributed more than 5 percent to the ascribable risk for schizophrenia, indicating that schizophrenia was a disorder of complex genetics like diabetes with multiple risk genes of modest effects interacting with environment to produce the phenotype. Three major categories of risk genes were identified: major histocompatibility complex (MHC) genes, genes related to neurotransmission and genes involved in development. Nearly a dozen genes are within 2 degrees of separation from the NMDA receptor including SR. Genetic inactivation of SR in mice results in a phenotype that shares many features with the neurochemical, neuronal structural and cognitive features of schizophrenia.

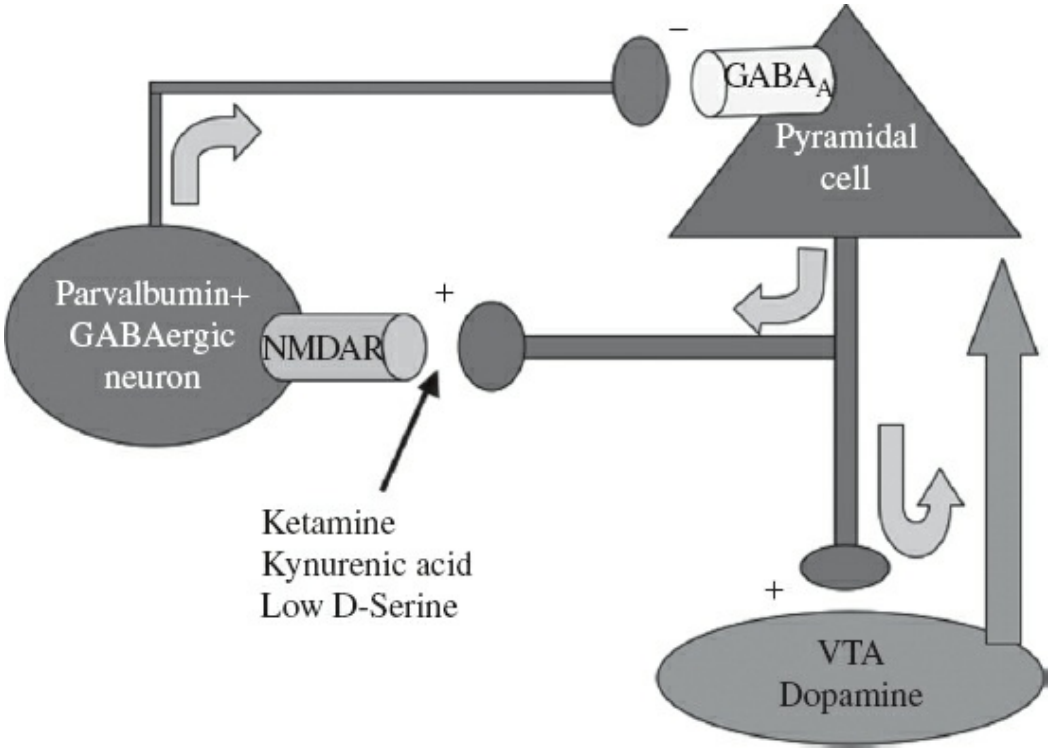

FIGURE 1.5–6. Pathological circuit in schizophrenia. The NMDA receptors on the rapidly firing parvalbumin (PV) expressing GABAergic interneurons in the intermediate layers of the cortex are disproportionately sensitive to antagonists or loss of the co-agonist, D-serine. NMDA receptor hypofunction causes reduced expression of PV, GAD67, and the GABA transporter and upregulation of $GABA_A$ receptors on pyramidal neurons. Disinhibition of the pyramidal neurons causes cognitive dysfunction and negative symptoms and drives excessive subcortical dopamine release resulting in psychosis.

Several clinical trials have been carried out with agonists at the glycine modulatory site on the NMDA receptor on patients receiving concurrent treatment with antipsychotic medications. The hypothesis being tested was that enhancing NMDA receptor function would reduce negative symptoms and improve cognition, aspects of the disorder unaffected by antipsychotic drugs. A meta-analysis of the results of these placebo-controlled trials of agonists at the NMDA receptor glycine modulatory site revealed highly significant reductions in depression, negative symptoms, cognitive deficits, and even positive symptoms.

Recent findings have provided a link between the GABAergic neuropathology and NMDA receptor hypofunction. Chronic treatment of rats with NMDA receptor antagonists causes a downregulation of GAD67, parvalbumin, and GAT. The sensitive subpopulation of GABAergic neurons is the rapidly firing interneurons that provide the perisomatic and axon hillock innervation of the pyramidal cells. Their NMDA receptors appear to be much more sensitive to antagonists than those less active GABAergic neurons and pyramidal cells. The subtly reduced GABAergic inhibition results in a disinhibition of the glutamatergic pyramidal output. This degradation of the inhibitory feedback could account for the cognitive

deficits and negative symptoms in schizophrenia, and the disinhibited output also results in elevated subcortical dopamine release and psychosis. Thus, psychosis would be considered a downstream event resulting from a disruption in critical glutamatergic–GABAergic synaptic function in the cerebral cortex.

Anxiety and Depression

GABAergic dysfunction has been associated with anxiety disorders, especially panic disorder, as well as with MDD. Clinically, there is considerable co-morbidity between anxiety and affective disorders. Decreased levels of the $GABA_A$ receptor modulators, the three α-reduced neuroactive steroids, have been found both in plasma and in CSF in MDD. Effective treatment with selective serotonin reuptake inhibitor (SSRI) increases the neurosteroid levels. In contrast, in patients suffering from panic disorder, the plasma neurosteroid levels were significantly elevated, perhaps as a compensatory mechanism. MRS has disclosed significant reductions in GABA levels in the anterior cingulate and in the basal ganglia of medicated patients with panic disorder. PET scanning reveals a highly selective reduction in benzodiazepine receptor sites bilaterally in the insular cortex in panic disorder. A genome-wide screen has shown significant linkage at 15q in a region containing $GABA_A$ receptor subunit genes and panic disorder. MRS reveals significant reductions in both GABA and glutamate/glutamine (Glx) in the prefrontal cortex in MDD. Postmortem studies indicate up-regulation of the $GABA_A$ receptor α_1- and β_3-subunits in the cerebral cortices of depressed patients who committed suicide, consistent with a reduction in GABAergic neurotransmission. The reduced levels of GABA in the occipital cortex in episodes of MDD normalized with effective treatment with SSRI or with ECT.

Glutamatergic dysfunction has also been implicated in depression. NMDA receptor antagonists have antidepressant effects in several animal models of depression including forced swim, tail suspension, and learned helplessness. A single injection of ketamine provides protection from the induction of behavioral despair in rats for up to 10 days. Chronic treatment with antidepressants alters the expression of NMDA receptor subunits and decreases glycine receptor B binding. Several placebo-controlled clinical trials have shown that a single dose of ketamine can produce a rapid, substantial, and persistent reduction in symptoms in patients with MDD, especially those who have not responded to SSRI/SNRI treatment.

Alcoholism

Ethanol at concentrations associated with intoxication has a dual action of enhancing GABAergic receptor function and attenuating NMDA receptor function. The GABA receptor effects may be associated with the anxiolytic

effects of ethanol. Persistent abuse and dependency on ethanol result in a downregulation of $GABA_A$ receptors and an upregulation of NMDA receptors such that acute discontinuation of ethanol results in a hyperexcitable state characterized by *delirium tremens*. Furthermore, supersensitive NMDA receptors in the context of thiamine deficiency may contribute to the excitotoxic neuron degeneration of Wernicke–Korsakoff syndrome.

Acamprosate is a derivative of homotaurine that was developed as an agent to reduce alcohol consumption, craving, and relapse in alcoholic patients, for which it exhibits moderate efficacy in clinical trials. Because of taurine's resemblance to GABA, it was thought that acamprosate acted via $GABA_A$ receptors, but electrophysiological studies found little evidence to support this hypothesis. Subsequent studies demonstrated that it inhibited NMDA receptor responses in cortical slices and recombinant NMDA receptors. The precise mechanism whereby acamprosate alters NMDA receptor function, however, remains unclear.

FAS is the most common preventable cause of mental retardation. Convincing evidence has been developed that the microcephaly associated with fetal alcohol exposure results from inhibition of NMDA receptor function, resulting in widespread neuronal apoptosis in the immature cortex. NMDA receptor activation is essential for immature neuronal survival and differentiation.

REFERENCES

Balu DT, Coyle JT. The NMDA receptor "glycine modulatory site" in schizophrenia: D-serine, glycine, and beyond. *Curr Opin Pharmacol.* 2015;20:109–115.

Bode A, Lynch JW. The impact of human hyperekplexia mutations on glycine receptor structure and function. *Mol Brain.* 2014;7:2.

Coyle JT, Basu A, Benneyworth M, Balu D, Konopaske G. Glutamatergic synaptic dysregulation in schizophrenia: therapeutic implications. *Handb Exp Pharmacol.* 2012;213:267–295.

Dutertre S, Becker CM, Betz H. Inhibitory glycine receptors: an update. *J Biol Chem.* 2012;287:40216–40223.

Gonzalez-Burgos G, Cho RY, Lewis DA. Alterations in cortical network oscillations and parvalbumin neurons in schizophrenia. *Biol Psychiatry.* 2015;77(12):1031–1040.

Granger AJ, Nicoll RA. Expression mechanisms underlying long-term potentiation: a postsynaptic view, 10 years on. *Philos Trans R Soc Lond B Biol Sci.* 2013;369:2013–2036.

Hirano T. Long-term depression and other synaptic plasticity in the cerebellum. *Proc Jpn Acad Ser B Phys Biol Sci.* 2013;89:183–195.

Jensen AA, Fahlke C, Bjørn-Yoshimoto WE, Bunch L. Excitatory amino acid transporters: recent insights into molecular mechanisms, novel modes of modulation and new therapeutic possibilities. *Curr Opin Pharmacol.* 2015;20:116–123.

Lisman JE, Coyle JT, Green RW, et al. Circuit-based framework for understanding neurotransmitter and risk gene interactions in schizophrenia. *Trends Neurosci.* 2008;31:234–

242.
Mathews DC, Zarate CA Jr. Current status of ketamine and related compounds for depression. *J Clin Psychiatry*. 2013;74:516–517.
Michalon A, Bruns A, Risterucci C, et al. Chronic metabotropic glutamate receptor 5 inhibition corrects local alterations of brain activity and improves cognitive performance in fragile X mice. *Biol Psychiatry*. 2014;75:189–197.
Moghaddam B, Krystal JH. Capturing the angel in "angel dust": twenty years of translational neuroscience studies of NMDA receptor antagonists in animals and humans. *Schizophr Bull*. 2012;38:942–949.
Price RB, Shungu DC, Mao X, et al. Amino acid neurotransmitters assessed by proton magnetic resonance spectroscopy: relationship to treatment resistance in major depressive disorder. *Biol Psychiatry*. 2009;65:792–800.
Proctor WR, Diao L, Freund RK, Browning MD, Wu PH. Synaptic GABAergic and glutamatergic mechanisms underlying alcohol sensitivity in mouse hippocampal neurons. *J Physiol*. 2006;575:145–159.
Rudolph U, Knoflach F. Beyond classical benzodiazepines: novel therapeutic potential of GABAA receptor subtypes. *Nat Rev Drug Discov*. 2011;10:685–697.
Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511:421–427.
Senter RK, Ghoshal A, Walker AG, et al. The role of mGlu receptors in hippocampal plasticity deficits in neurological and psychiatric disorders: implications for allosteric modulators as novel therapeutic strategies. *Curr Neuropharmacol*. 2016;14:455–473.
Singewald N, Schmuckermair C, Whittle N, et al. Pharmacology of cognitive enhancers for exposure-based therapy of fear, anxiety and trauma-related disorders. *Pharmacol Ther*. 2015;149:150–190.
Straub C, Tomita S. The regulation of glutamate receptor trafficking and function by TARPs and other transmembrane auxiliary subunits. *Curr Opin Neurobiol*. 2012;22:488–495
Swick TJ. Swick TJ. Treatment paradigms for cataplexy in narcolepsy: past, present, and future. *Nat Sci Sleep*. 2015;7:159–169.
Zorumski CF, Paul SM, Izumi Y, Covey DF, Mennerick S. Neurosteroids, stress and depression: potential therapeutic opportunities. *Neurosci Biobehav Rev*. 2013;37:109–122.

▲ 1.6 Neuropeptides: Biology, Regulation, and Role in Neuropsychiatric Disorders

SHAUN P. BROTHERS, PH.D., LARRY JAMES YOUNG, PH.D., AND CHARLES B. NEMEROFF, M.D., PH.D.

INTRODUCTION

Neuropeptides are the most diverse class of signaling molecules in the CNS. Initially discovered for their role in the hypothalamic regulation of pituitary hormone secretion, peptides have been found to play a complex role in brain function over the last 35 years. Neuropeptides and their receptors are

widely distributed throughout the CNS, where they exert an extraordinary array of direct or neuromodulatory effects, ranging from modulating neurotransmitter release and neuronal firing patterns to the regulation of emotionality and complex behaviors. Over 100 unique biologically active neuropeptides have been identified in the brain, a subset of which is presented in Table 1.6–1. Adding to the complexity of neuropeptide systems in the CNS, the actions of many peptides are mediated via multiple receptor subtypes localized in different brain regions. In fact, the discovery of new peptides and receptor subtypes has outpaced the current understanding of the roles of these peptides in normal or aberrant CNS function. Pharmacological, molecular, and genetic approaches are now leading the way in the understanding of the contribution of neuropeptide systems in psychiatric disorders. The first half of this chapter provides a general overview of neuropeptide biology, including their biosynthesis, distribution, and signaling. The second half of the chapter provides specific, detailed information for a select group of neuropeptides, with special relevance to behavior and psychiatric disorders. By definition, a neuropeptide is a chain of two or more amino acids linked by peptide bonds and differs from other proteins only in the length of the amino acid chain. Neuropeptides range in length from two (e.g., carnosine and anserine) to over 40 amino acids (e.g., CRF and urocortin). By convention, peptides greater than 90 amino acids in length (about 10,000 molecular weight) are considered proteins. The neuropeptides highlighted in detail in this chapter include thyrotropin-releasing hormone (TRH), CRF, oxytocin (OT), arginine vasopressin (AVP), neurotensin (NT), and NPY. The structures of these neuropeptides are illustrated in Table 1.6–2 and are written using the single-letter amino acid code by convention from the amino terminus (NH_2-) beginning on the left to the carboxyl terminus (-COOH) on the right. Of course, there are many other examples of neuropeptides of relevance to psychiatric disorders, and a brief discussion of some additional peptides of particular interest is also presented at the end of the chapter. A detailed discussion of all neuropeptide systems of potential relevance to psychiatry is beyond the scope of this chapter. TRH and CRF are hypothalamic hypophysiotropic hormones, which stimulate the release of thyroid-stimulating hormone (TSH) and adrenocorticotropic hormone (ACTH), respectively, from the anterior pituitary (adenohypophysis) into the bloodstream. OT and AVP are neurohypophysial peptides that are released directly into the bloodstream from the posterior pituitary (neurohypophysis) under specific physiological conditions. However, all of these above-mentioned peptides, including NT, also function in the CNS as neurotransmitters, neuromodulators, or neurohormones in ways that are often quite distinct and independent from their effects on peripheral endocrine systems. Neuropeptides have been implicated in the regulation of a variety of behavioral and physiological

processes, including thermoregulation, food and water consumption, sex, sleep, locomotion, learning and memory, responses to stress and pain, emotion, and social cognition, to name a few. Involvement in such behavioral processes suggests that it is likely that neuropeptidergic systems contribute to the symptoms and behaviors exhibited in the major psychiatric disorders, including autism, schizophrenia, mood disorders, and anxiety.

Table 1.6–1.
Selected Neuropeptide Transmitters

Adrenocorticotropin hormone (ACTH)
Angiotensin
Apelin
Atrial natriuretic peptide
Bombesin
Calcitonin
Calcitonin gene–related peptide (CGRP)
Cocaine and amphetamine regulated transcript (CART)
Cholecystokinin (CCK)
Corticotropin-releasing factor (CRF)
Dynorphin
β-Endorphin
Leu-enkephalin
Met-enkephalin
Galanin
Gastrin
Ghrelin
Gonadotropin-releasing hormone (GnRH)
Growth hormone
Growth hormone–releasing hormone (GHRH; GRF)
Insulin
Kisspeptin
Motilin
Neuropeptide S
Neuropeptide Y (NPY)
Neurotensin
Neuromedin-N
Nociceptin
Orexin
Oxytocin
Pancreatic polypeptide
Prolactin
Secretin
Somatostatin (SS; SRIF)
Substance K
Substance P
Thyrotropin-releasing hormone (TRH)
Urocortin (1, 2 and 3)
Vasoactive intestinal polypeptide (VIP)
Vasopressin (AVP; ADH)

Table 1.6–2.
Selected Neuropeptide Structures

Name	Amino Acid Sequence
Thyrotropin-releasing hormone (TRH)	pE-H-P-NH_2
Corticotropin-releasing factor (CRF)	S-E-E-P-P-I-S-L-D-L-T-F-H-L-L-R-E-V-L-E-M-A-R-A-E-Q-L-A-Q-Q-A-H-S-N-R-K-L-M-E-I-I-NH_2
Arginine Vasopressin (AVP)	C-Y-I-Q-N-C-P-L-G-NH_2
Oxytocin (OT)	C-Y-F-Q-N-C-P-R-G-NH_2
Neurotensin (NT)	pE-L-Y-E-N-K-P-R-R-P-Y-I-L-OH
Neuropeptide Y	Y-P-S-K-P-D-N-P-G-E-D-A-P-A-E-D-M-A-R-Y-Y-S-A-L-R-H-Y-I-N-L-I-T-R-Q-R-Y-NH_2

Note the cyclized glutamines at the N-terminals of TRH and NT indicate by pE-, the cysteine–cysteine disulfide bonds of AVP and OT, and the amidated carboxy-terminals of TRH, CRF, AVP, OT and NPY.

Investigating Neuropeptide Function

The roles of neuropeptides in CNS function and behavior have been examined using a multitude of experimental techniques. The levels of analysis include the following: molecular structure and biosynthesis of the peptide and its receptor(s), the neuroanatomical localization of the peptide and its receptor(s), the regulation of the expression and release of the peptide, and finally the behavioral effects of the peptide. The vast majority of information on neuropeptide biology is derived from laboratory animal studies; however, there is a growing database on the localization, activity, and potential psychiatric relevance of several neuropeptides systems in humans.

Most neuropeptide structures have been identified based on chemical analysis of purified biologically active peptides, leading ultimately to the cloning and characterization of the genes encoding them. Characterization of the gene structure of peptides and their receptors has allowed for understanding of the molecular regulation of these systems, and their chromosomal localization is useful in genetic studies examining the potential roles of these genes in the pathogenesis of psychiatric disorders. Structural characterization permits the production of immunological and molecular probes that are useful in determining peptide distribution and regulation in the brain. Quantitative enzyme-linked immunosorbent assay (ELISA) on microdissected brain regions or immunocytochemistry on brain sections is typically used to characterize the regional distribution of peptides throughout the brain. Both techniques use specific antibodies generated against the neuropeptide to detect its presence. Immunocytochemistry allows researchers to visualize the precise cellular localization of peptide-synthesizing cells as well as their projections throughout the brain, though the technique is generally not quantitative.

Using molecular probes homologous to the mRNA encoding the peptides or receptor, in situ hybridization can be used to localize and quantify gene expression in brain sections. This is a powerful technique for examining the molecular regulation of neuropeptide synthesis with precise neuroanatomical resolution, which is impossible for other classes of nonpeptide neurotransmitters that are not derived directly from the translation of mRNAs, such as dopamine, serotonin, and norepinephrine. In addition to immunocytochemistry and in situ hybridization, receptor autoradiography on brain sections (see Fig. 1.6–4) or "grind and bind" receptor-binding assays on microdissected brain tissue is frequently used to localize and quantify neuropeptide receptors in specific regions of the brain. Receptor autoradiography involves allowing a radiolabeled ligand to bind the receptor on a thin slice of tissue and then detecting the bound by visualizing it on x-ray film or other means. Other molecular techniques, such as Northern blot analysis, ribonuclease protection assay, and quantitative polymerase chain reaction (PCR), are also commonly used to measure neuropeptide and receptor expression and regulation by quantifying the mRNAs encoding the peptide or receptor. However, quantification of neuropeptide gene expression or immunoreactivity within a cell or tissue homogenate does not provide information on neuropeptide release. In vivo microdialysis, in which peptide concentrated in the extracellular fluid is collected at sequential time intervals using dialysis probes implanted into specific brain regions, may be used to quantify neuropeptide release under defined physiological or behavioral circumstances.

Generally, the behavioral effects of neuropeptides are initially investigated by infusions of the peptide directly into the brain. Unlike many nonpeptide neurotransmitters, most neuropeptides do not penetrate the BBB in amounts sufficient enough to produce CNS effects. Furthermore, serum and tissue enzymes tend to degrade the peptides before they reach their target sites. The degradation is usually the result of the cleavage of specific amino acid sequences targeted by a specific peptidase designed for that purpose. Thus, intracerebroventricular (ICV) or site-specific infusions of peptides in animal models are generally required to probe for behavioral effects of peptides. However, there are some examples of delivery of neuropeptides via intranasal infusions in human subjects, which in some cases appears to permit access of the peptide to the brain.

In many cases, the interpretation of neuropeptide infusion studies is complicated because of the considerable cross-talk between specific neuropeptides and several heterologous receptors. For example, OT and vasopressin differ by only two of nine amino acids, and both peptides cross-react to some degree with both receptor types. In some cases, highly selective peptide-receptor synthetic agonists or antagonists have been developed that allow researchers to examine the role of specific

neuropeptide receptors in the regulation of behavior or physiological processes. In addition, transgenic and knockout mouse approaches are becoming more and more commonly used approaches to investigate neuropeptide function. For example, mutant mouse strains with null mutations (knockouts) in either the peptide gene or the corresponding receptor have been developed and have proven quite useful for exploring the role of neuropeptides in behavioral processes. siRNA or short hairpin RNA (shRNA) techniques, which lead to selective degradation of the targeted mRNA in specific brain regions, have been used to examine the function of specific neuropeptide-producing neuronal populations. The CRISPR/Cas system has been used for gene editing and holds potential to begin to separate temporal effects of up/downregulating neuropeptides and genes. Similar efforts with viruses to selectively overexpress or knock down neuropeptide or receptor RNA in specific microinjected brain regions have been pursued as well. Most recently, there have been developments in optogenetics that, while still in the early stages of use, offer the potential to probe neuronal activity in living organisms. With optogenetics, LASER light is used to activate neuronal firing, which can provide special and temporal resolution of real-time neuronal activation in different models, and potentially in combination with the modulators and technologies described above.

As noted above, one of the greatest impediments for exploring the role and potential therapeutic value of neuropeptides is the inability of the peptides or their peptidergic agonists/antagonists to penetrate the BBB. Thus, the behavioral effects of most peptides in humans are largely uninvestigated, with the exception of a few studies utilizing intranasal delivery. However, in some instances, small molecule nonpeptide agonists/antagonists have been developed that can be administered peripherally and permeate the BBB in sufficient quantities to affect receptor activation.

Humans are less-than-ideal subjects for neuropeptide research for several reasons. First, although blood samples to determine plasma hormone concentrations are relatively easy to obtain, the independent regulation of peripheral and CNS peptide release, the high concentration of plasma peptidases, and the BBB make it virtually impossible to infer CNS peptide physiology from plasma hormone concentrations. Also, the use of biopsy to directly assess tissue peptide concentrations is not ideal because it is not routinely repeatable, is limited to superficial structures, and suffers from potential morbidity. In contrast, however, CSF has been shown to reflect extracellular fluid concentrations of several transmitter substances, is in direct contact with the CNS, is screened from peripheral serum sources by the BBB, and may be sampled across time. The limitations of human CSF studies include a lack of information about the regional CNS source of any changes in peptide concentration detected, the use of lumbar

CSF, which is somewhat removed from higher forebrain CNS sources of peptides and subject to spinal cord peptide contributions, the potentially confounding effects of previous drug treatments or disease episodes, and the reluctance of research subjects to consent to this procedure. Postmortem tissue studies of neuropeptide concentration changes in psychiatric disease have been informative in many cases, but interpretation must include consideration of postmortem delay, previous drug treatment, and coexisting illnesses. Most of the data on alterations in CSF or tissue concentrations of neurotransmitters have been derived from comparisons between diagnostically defined psychiatric groups and control groups. However, the controls may be so-called "neurologic or psychiatric controls," not healthy volunteers, and the accuracy and consistency of the diagnoses may be less than optimal. In addition, the etiology of a syndromal diagnosis may differ among subjects in the same diagnostic group. Even after matching for age, gender, or other demographic variables, heterogeneity among human research populations results in individual variations of absolute peptide values that are often quite wide. Such variances severely reduce the power of group comparisons to detect alterations in peptide concentrations. The use of pre- and posttreatment CSF samples, or of samples obtained during the active disease state versus when the patient is in remission, addresses some of the serious limitations in study design. For such progressive diseases as schizophrenia or Alzheimer disease, serial CSF samples may be a valuable indicator of disease progression or response to treatment. Even with these constraints, significant progress has been made in describing the effects of various psychiatric disease states on neuropeptide systems in the CNS.

BIOSYNTHESIS

Unlike other neurotransmitters, the biosynthesis of a neuropeptide involves transcription of an mRNA from a specific gene, translation of a polypeptide preprohormone encoded by that mRNA, and then posttranslational processing involving proteolytic cleavage of the preprohormone to yield the active neuropeptide. Over the past 30 years, the gene structures and biosynthetic pathways of many neuropeptides have been elucidated. The gene structure of selected neuropeptides is illustrated in Figure 1.6–1. Neuropeptide genes are generally composed of multiple exons that encode a protein preprohormone. The N-terminus of the preprohormone contains a signal peptide sequence, which guides the growing polypeptide to the rough endoplasmic reticulum (RER) membrane. The single preprohormone molecule often contains the sequences of multiple peptides that are subsequently separated by proteolytic cleavage by specific enzymes. For example, translation of the gene encoding NT yields a preprohormone, which upon enzymatic cleavage produces both NT and neuromedin-N. Other neuropeptide genes, such as

the TRH gene, encode multiple copies of the peptide sequence or, as in the case of OT and vasopressin, also encode other proteins essential in the posttranslational processing and transport of the neuropeptide. The neuroanatomical localization and abundance of neuropeptides is determined primarily by the region-specific expression and regulation of its gene. Each neuropeptide gene is expressed in well-defined populations of neurons within the brain. The precise neuroanatomical pattern of peptide hormone gene expression is determined by regulatory DNA sequences surrounding the gene. This has been elegantly demonstrated for the OT gene. OT is expressed in a subset of magnocellular neurons in the PVN of the hypothalamus. Transgenic mice carrying the rat OT gene with the surrounding regulatory sequences expressed the rat OT transgene specifically in the mouse magnocellular oxytocinergic neurons. Smaller constructs lacking these regulatory regions did not result in the correct expression patterns in the brain. Transcription factor binding sites (TFBSs) located in the promoter of the gene are also involved in the physiological regulation of peptide gene expression. Analysis of promoter sequences of peptide genes has provided insights into the molecular regulation of peptide biosynthesis.

The mRNA encoding the preprohormone is translated by ribosomes associated with the RER and the growing polypeptide is translated into the cisternae of the RER, with the signal peptide anchored in the RER membrane. Once translated, the signal peptide of the preprohormone is cleaved by a signal endopeptidase, freeing the prohormone polypeptide. The prohormone is then shuttled to the Golgi apparatus, where packaging into granules or vesicles occurs. Proteolytic cleavage of the prohormone into the biologically active neuropeptide begins in the Golgi and continues in the granules. Production of biologically active neuropeptides from prohormones begins with cleavage at specific sites adjacent to the neuropeptide sequence by specific endopeptidases known as prohormone convertases. Prohormone convertases cleave generally at pairs of basic amino acids (e.g., Lys-Arg, Lys-Lys and Arg-Arg) flanking the neuropeptide sequence. There are at least seven prohormone convertases, each with unique properties including substrate specificity and neuroendocrine distribution. Prohormone convertases are copackaged with the prohormones in the granules at the Golgi apparatus. The substrate specificity and differential distribution of the prohormone convertases provide a mechanism by which different neuropeptides encoded by a single prohormone can be differentially produced in an active form. Following endopeptidase cleavage, the peptide fragments are subjected to exoproteolysis by carboxypeptidases and/or aminopeptidases in order to remove the residual basic residues on the carboxy- or amino-terminus of the peptide fragments. The synthesis and processing of neuropeptides is illustrated in Figure 1.6–2.

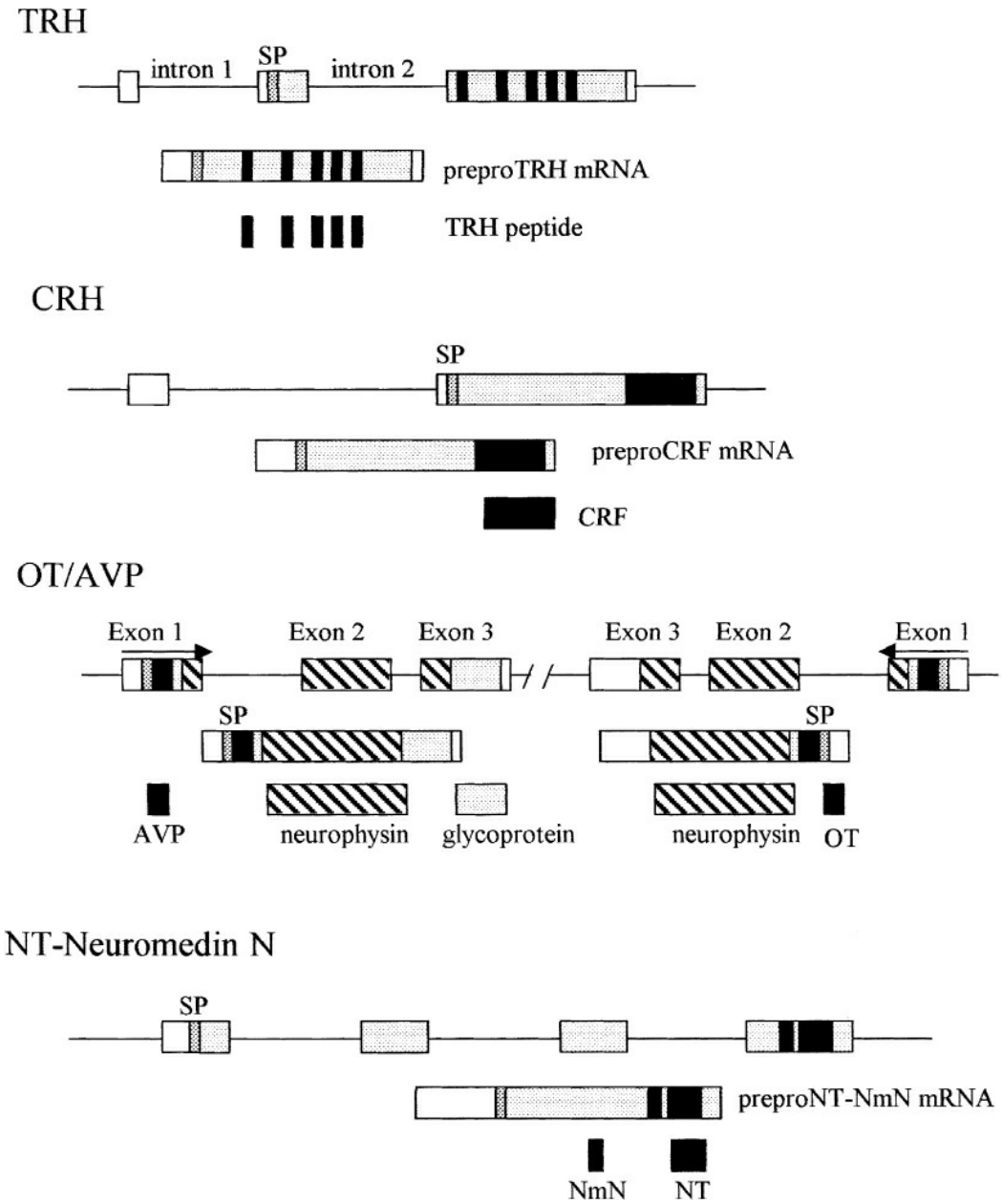

FIGURE 1.6–1. Schematics illustrating the gene structure, preprohormone mRNA, and processed neuropeptides of TRH, CRF, OT, AVP, and NT. Boxed regions indicate the locations of the exons in the respective genes. Shaded or hatched regions indicate coding regions. Each preprohormone begins with a signal peptide sequence (SP). Black boxes indicate the location of the sequences encoding the neuropeptide.

Although many known peptides are complete and biologically active when cleaved from the prohormone, many others are subjected to additional posttranslational processing. Certain peptides have a metabolically blocked carboxy terminus that is often amidated. A glycine residue in the prohormone sequence often acts as the amide donor, and in the case of TRH, is cleaved by a monooxygenase that is contained in secretory granules. TRH is further processed on the N-terminus, where glutamine is cyclized by a glutamylcyclase to yield a pyroglutamyl moiety. These alterations are usually effective in reducing susceptibility to

degradation and are often required for biological activity, as is the case of TRH, which is rendered inactive when the C-terminal amide is removed by proline endopeptidase to generate the free-acid structure. Other posttranslational processing events for active peptides include glycosylation, phosphorylation, and the formation of disulfide bonds, which are often required for either biological activity or transport. Several neuropeptides, including OT and vasopressin, contain a cysteine–cysteine disulfide bond, resulting in cyclic peptide structures (see Table 1.6–2).

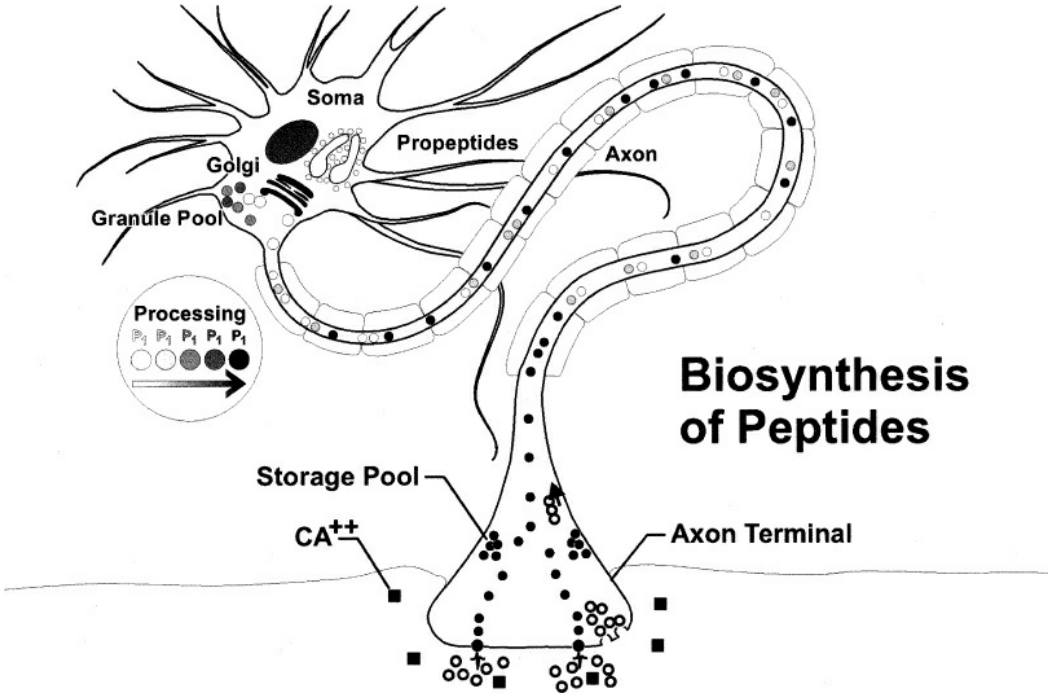

FIGURE 1.6–2. The peptidergic neuron. The figure shows the main steps in the chain of events from the information stored in the DNA molecule to the peripherally detected peptide fragments. The DNA sequence in the nucleus is transcribed to the mRNA molecule for further transport to the endoplasmic reticulum, where a translation takes place to form a large precursor protein (preproprotein). That protein is prepared for axonal transport by packaging into neurosecretory vesicles, or granules within the Golgi complex. During transport, the precursor protein is processed by specific cleavage enzymes into active and inactive peptide fragments. After release, the peptides are further degraded into smaller peptide fragments or constituent amino acids.

DISTRIBUTION AND REGULATION

Although many neuropeptides were originally isolated from pituitary and peripheral tissues, the majority of neuropeptides were subsequently found to be widely distributed throughout the brain. The peptides involved in regulating pituitary secretion are concentrated in the hypothalamus. Hypothalamic releasing and release-inhibiting factors are produced in neurosecretory neurons adjacent to the third ventricle that send projections to the median eminence, where they release peptides into the hypothalamo-hypophyseal portal circulatory system. Peptides produced in these neurons are often subject to feedback regulation by the peripheral

hormones that they regulate. For example, TRH regulates activity of the hypothalamic–pituitary–thyroid (HPT) axis and thyroid hormones negatively feedback on TRH gene expression. However, neuropeptide-expressing neurons and their projections are found in many other brain regions, including limbic structures, cerebral cortex, midbrain, hindbrain, and spinal cord. Neuropeptides are often colocalized and released with other neuropeptide or nonpeptide neurotransmitters, refuting the "one neuron–one transmitter" tenet. The colocalization of neuropeptides within classical neurotransmitter circuits suggests an interaction between these systems, and the modulation of monoamine neurotransmitter (e.g., norepinephrine) function by neuropeptides (e.g., NPY) is common and well established. Indeed, many neuropeptides are cotransmitted with nonpeptide neurotransmitters and serve to modulate neurotransmission of both the pre- and postsynaptic cells. These interactions have stimulated speculation concerning the involvement of neuropeptides in the underlying pathophysiology of psychiatric disorders.

NEUROPEPTIDE SIGNALING

Neuropeptides may act as neurotransmitters, neuromodulators, or neurohormones. Neuropeptide neurotransmitters are typically released from axonal terminals into a synapse, where binding to neuropeptide receptors changes the postsynaptic membrane potential, either depolarizing or hyperpolarizing the cell. For classical neurotransmitters, this often involves direct modulation of voltage-gated ion channels. In contrast, neuropeptide neuromodulators and neurohormones do not directly affect the firing of the target cell itself, but may alter the response of the cell to other neurotransmitters through modulation of second messenger pathways. Neuropeptide release is not restricted to synapses or axon terminals but may occur throughout the axon or even from dendrites. Neuropeptides may also diffuse quite a distance from the release site to the target cell, which possesses the neuropeptide receptor, where it acts as a neurohormone. Such paracrine neuropeptide release can occur even from nonneuronal cell populations in the brain, such as glial cells. In fact, there are numerous examples of a mismatch between neuropeptide and neuropeptide receptor distribution in the brain. Neuropeptides are released by exocytosis of the granules in response to electrical or hormonal stimulation of the neuron containing the neuron. Stimulation results in an increase in intracellular calcium concentrations, which leads to the fusion of the peptidergic granules with the plasma membrane and expulsion of the peptide into the extracellular space. The cellular signaling of neuropeptides is mediated by specific neuropeptide receptors. Thus, understanding the neuropeptide receptor function is essential for understanding neuropeptide biology. Neuropeptide receptors have undergone the same process of discovery and characterization that receptors for other neurotransmitters

have enjoyed. The vast majority of neuropeptide receptors are G-protein–coupled, seven transmembrane domain receptors (GPCR) belonging to the same family of proteins as the monoamine receptors. Each neuropeptide receptor is usually, but not always exclusively, coupled to one type of G protein (e.g., G_s, G_i, G_q). Receptors can also homo- and heterdimerize allowing multiple modes of action via a single receptor population, though this process is not well understood. Depending on the subtype of G protein with which the receptor interacts, receptor activation may result in the stimulation or inhibition of specific second messenger pathways. The most common types of receptor signaling pathways involve the activated G protein modulating the activity of either adenylate cyclase (e.g., G_s or G_i) or PLC (e.g., G_q). Stimulation of adenylate cyclase results in increased cyclic adenosine monophosphate (cAMP) production in the cell, while stimulation of PLC results in increased diacylglycerol (DAG) and inositol triphosphate (IP_3). These so-called "second messengers" then effect increases in intracellular calcium concentrations, activation of protein kinases, and ultimately a host of cellular responses including altered gene expression, altered membrane potential, or cellular degranulation.

Many neuropeptides exert their effects through multiple different receptor subtypes, which have different affinities for the peptides and activate different second-messenger pathways. These receptor subtypes are typically differentially distributed throughout the brain. Furthermore, many receptors may be modulated by more than one neuropeptide. For example, there are three subtypes of the vasopressin receptor, the V1a, V1b, and V2 subtypes, with the V1a and V1b predominating in the brain, while the V2 is localized in the kidney. Each of these receptor subtypes exhibits a unique tissue distribution, interacts with different G proteins, and activates different second-messenger systems. Within the brain, there are remarkable species differences as well as individual variation in OT and V1a receptor distribution, which has been associated with intra- and interspecies variation in social behaviors. In addition, OT may stimulate vasopressin receptor subtypes, and vasopressin may stimulate the OT receptor. Likewise, the main two CRF receptors are differentially localized within the brain, and both receptors can be modulated by both CRF and urocortin I, making it difficult to ascertain the relative role of each receptor in CRF functioning.

Many, if not perhaps all, neurotransmitters appear to be coreleased with at least one other neurotransmitter at the synapse. ATP, for example, can apparently be coreleased from all neuronal types, though the action of the molecule will vary according to biology of the postjunctional cell. This cotransmission certainly also occurs for neuropeptides. One of the first to be delineated, NPY is coreleased with norepinephrine from the same vesicles within cells to act simultaneously on the target cell. NPY can also

act via feedback inhibition on the releasing cell to reduce signaling tone. In this instance, NPY appears to play a neuromodulatory role at the synapse, by controlling signal strength and acting as a shutoff switch.

Other types of cotransmission can occur including differential release of transmitters and cotransmitters (e.g., via frequency modulated release), postjunctional and/or prejunctional targeting cotransmitters, synergistic versus inhibitory cotransmitters, and others. It is important to consider neuropeptides as cotransmitters in such systems because cotransmission has the potential to mediate or modulate the underlying disease pathophysiology. Using NPY as an example again, NPY is also cotransmitted with NE in the adrenals within the HPA axis. NPY appears to play an overall suppressive role in NE signaling within the adrenals, thus suppressing the stress response. It has been widely theorized that activation of NPY can ameliorate some of the stress associated with chronic stress disorders and drug abuse/alcoholism, where there is prolonged activation of the HPA axis and where stress is viewed as playing a major role in relapse and drug-seeking behaviors.

Most, if not all, genes for neuropeptide receptors and ligands have been cloned and characterized, though some receptors called “orphan receptors” have as of yet unknown endogenous ligands. Similarly, some ligands, for example, “cocaine and amphetamine regulated transcript” (CART), are thought to signal via GPCRs, which remain to be identified. It is likely, however, that all known GPCRs and ligands will be “de-orphanized” within the next decade, as numerous efforts are underway to do so. Once the cDNA of the receptor has been isolated, it can be used to produce purified receptor protein for structural and functional studies. By mutating specific amino acids in the receptor structure and determining relative binding affinities of peptides with various amino acid substitutions, it is possible to elucidate the nature of the ligand–receptor interaction and to inform crystal structures of the receptors. This information facilitates the development of compounds that specifically modulate receptor function, including nonpeptide drugs, leading to the ability to manipulate peptide systems in ways that are currently enjoyed by the more classic neurotransmitters. The availability of cDNAs encoding the receptor also permits the neuroanatomical mapping of the receptor-producing cells in the brain by use of probes that target the mRNA of the receptor specifically, which is critical for understanding the neural circuits modulated by the peptide. Finally, with the cloned receptor in hand, it is possible to use transgenic techniques, such as targeted gene overexpression, or gene knockouts to further elucidate receptor function. siRNA, shRNA, and viral techniques now allow the targeted synthesis disruption of specific receptor populations, allowing researchers to examine the role of these receptor populations on physiology and behavior.

The three factors that determine the biological role of a neuropeptide

hormone are (i) the temporal-anatomical release of the peptide, (ii) functional coupling of the neuropeptide receptor to intracellular signaling pathways, and (iii) the cell type and circuits in which the receptor is expressed. Genetic studies have demonstrated that regulatory sequences flanking the receptor-coding region determine the expression pattern of the receptor, and thus the physiological and behavioral response to the neuropeptide. For example, mice and voles differ in the localization of AVP receptors in the brain, and they also differ in their behavioral response to AVP. However, when transgenic mice were created carrying the vole AVP receptor gene with the flanking regulatory sequences, the mice expressed the receptor in a pattern similar to that of the vole and then displayed behavioral responses to AVP similar to that of voles. This study suggests that polymorphisms in the regulatory region of a neuropeptide receptor gene could result in significant differences in neuropeptide function, and thus could potentially be relevant to psychiatric disorders. Many receptor genes have now been localized to specific chromosomal loci and are being examined in genetic studies for associations with psychiatric disorders.

Historically, the inability to block specific neuropeptide signals pharmacologically has severely hindered research into the roles of the endogenous peptides in various behaviors and physiological effects. However, for many neuropeptide receptors, selective agonists and antagonists are now available, which have been extremely informative in preclinical studies to examine receptor function. As mentioned above, most of these compounds are derivatives of the peptide itself and therefore do not pass the BBB. More recently, a number of pharmaceutical companies, and even academic labs, have synthesized nonpeptidergic, lipophilic compounds that can permeate the BBB and may act as neuropeptide agonists or antagonists. The development of these types of compounds is essential for understanding the role of neuropeptide receptor function in human behavior and may also be useful in the development of PET radioligands to study receptor distribution in living human subjects. These compounds also hold promise as therapeutic agents in the treatment of certain psychiatric disorders.

Peptidases

Unlike monoamine neurotransmitters, peptides are not inactivated by being actively taken up by presynaptic nerve terminals. Rather, released peptides are degraded into smaller fragments, and eventually into single amino acids, by specific enzymes termed peptidases. The enzymes may be found bound to post- or presynaptic neural membranes or in solution in the cytoplasm and extracellular fluid, and they are distributed widely in peripheral organs, and serum, as well as in the CNS. As a result, neuropeptides generally have half-lives on the order of minutes once released. There are several general classes of peptidases, with several

distinct enzymes in each class. Those classes include the serine endopeptidases containing such enzymes as trypsin and chymotrypsin; thiol peptidases, such as pyroglutamate amino peptidase and cathepsin B and C; acid proteases, such as pepsin and renin; metalloendopeptidases, such as neural endopeptidase and angiotensin-converting enzymes (ACEs); and metalloexopeptidases, such as aminopeptidases and carboxypeptidases such as enkephalin-convertase and carboxypeptidase A and B. These degradative enzymes are often the same as those used in processing but have different subcellular locations. An example is carboxypeptidase B, which cleaves the dibasic amino acid residues flanking the active peptide sequence in the prohormone during processing, or reduces activity at the receptor if the peptide contains dibasic amino acids in the active sequence, such as NT. Peptidases have pH and temperature optimums for activity and can be inhibited by various chemicals or chelators, or by amino acid substitution at vulnerable points in the peptide chain. Alterations in peptidase activity or concentration can contribute to alterations in the synaptic availability of a peptide, and the regulation of peptidase levels may be as exquisitely controlled as receptor number, and peptide synthesis and release. Cleavage of the actively released form of the peptide usually ends or significantly reduces biological activity, but examples abound of partial or complete receptor activation by partially metabolized peptides or their fragment. Peptidases offer yet another potential opportunity for the integration and regulation of neuropeptide transmitter actions and synaptic availability. As the present peptidase inhibitors are relatively nonspecific in their ability to inhibit various peptidases, there have been few attempts to influence peptide concentrations by pharmacological blockade of their associated peptidases. ACE inhibitors such as captopril and lisinopril are one exception to that generality. It is expected that second- and third-generation peptidase inhibitors, with discrete peptidase and possibly regional specificity, will be developed that eventually may allow the truly elegant manipulation of endogenous neuropeptide concentrations.

SPECIFIC NEUROPEPTIDES AS PROTOTYPES OF NEUROPEPTIDE BIOLOGY

Thyrotropin-Releasing Hormone

In 1969, TRH, a pyroglutamyl histidyl prolinamide tripeptide (Table 1.6–2), became the first of the hypothalamic releasing hormones to be isolated and characterized, work that, in part, resulted in the Nobel prize for Drs. Andrew Schally and Roger Guillemin in 1977. The discovery of the structure of this hormone led to the conclusive demonstration that peptide hormones secreted from the hypothalamus regulate secretion of hormones from the anterior pituitary. The gene for TRH in humans resides on chromosome 3q13.3-q21. In rats, it consists of three exons (coding regions) separated by

two introns (noncoding regions) (Fig. 1.6–1). The first exon contains the 5′ untranslated region (UTR) of the mRNA encoding the TRH preprohormone, the second exon contains the signal peptide (SP) sequence and much of the remaining amino terminal-end of the precursor peptide, and the third contains the remainder of the sequence, including five copies of the TRH precursor sequence, the carboxy terminal region, and the 3′ UTR. The 5′ flanking region of the gene, or promoter, contains sequences homologous to the glucocorticoid receptor and the thyroid hormone receptor DNA-binding sites, providing a mechanism for the regulation of this gene by cortisol and negative feedback by thyroid hormone. Enzymatic processing of TRH begins with excision of the progenitor peptides by carboxypeptidases, amidation of the C-terminal proline and cyclization of the N-terminal glutamine, to yield five TRH molecules per prohormone molecule. TRH is widely distributed in the CNS, with TRH immunoreactive neurons being located in the olfactory bulbs, entorhinal cortices, hippocampus, extended amygdala, hypothalamus, and midbrain structures. As is the case for most neuropeptides, the TRH receptor is also a member of the seven-transmembrane domain, GPCR family.

Hypothalamic TRH neurons project nerve terminals to the median eminence, where they release TRH into the hypothalamo-hypophyseal portal system, where it is transported to the adenohypophysis, causing the release of TSH into the systemic circulation. TSH subsequently stimulates the release of the thyroid hormones, triiodothyronine (T_3) and thyroxine (T_4), from the thyroid gland. TRH neurons in the PVN contain thyroid hormone receptors and respond to increases in thyroid hormone secretion with a decrease in TRH gene expression and synthesis. This negative feedback of thyroid hormones onto the TRH-synthesizing neurons was first demonstrated by a decrease in TRH content in the median eminence, but not in the PVN of the hypothalamus, after thyroidectomy. This effect can be reversed with exogenous thyroid hormone treatment. The treatment of normal rats with exogenous thyroid hormone decreases TRH concentration in the PVN and the posterior nucleus of the hypothalamus. Using a probe against the TRH preprohormone mRNA, in situ hybridization studies have demonstrated that TRH mRNA is increased in the PVN 14 days after thyroidectomy. The ability of thyroid hormones to regulate TRH mRNA can be superseded by other stimuli that activate the HPT axis. In that regard, repeated exposure to cold (which releases TRH from the median eminence) induces increases in the levels of TRH mRNA in the PVN despite concomitantly elevated concentrations of thyroid hormones. Further evidence of the different levels of communication of the HPT axis are seen in TRH's ability to regulate the production of mRNA for the pituitary TRH receptor and for TRH concentrations to regulate the mRNA coding for both the α and β subunits of the thyrotropin (TSH) molecule. In addition, TRH-containing synaptic boutons have been observed in contact with TRH-

containing cell bodies in the medial and periventricular subdivisions of the PVN, thus providing anatomical evidence for ultrashort feedback regulation of TRH release. Negative feedback by thyroid hormones may be limited to the hypothalamic TRH neurons because negative feedback on TRH synthesis by thyroid hormones has not been found in extrahypothalamic TRH neurons.

The early availability of adequate tools to assess HPT axis function (i.e., radioimmunoassays, synthetic peptides), coupled with observations that primary hypothyroidism is associated with depressive symptomatology, ensured extensive investigation of the involvement of this axis in affective disorders. Early studies established the hypothalamic and extrahypothalamic distribution of TRH. This extrahypothalamic presence of TRH quickly led to speculation that TRH might function as a neurotransmitter or neuromodulator. Indeed, a large body of evidence supports such a role for TRH. Within the CNS, TRH is known to modulate several different neurotransmitters, including dopamine, serotonin, acetylcholine, and the opioids. TRH has been shown to arouse hibernating animals, and counteracts the behavioral response and hypothermia produced by a variety of CNS depressants including barbiturates and ethanol.

Interest in putative CNS actions of TRH was stimulated by studies of the HPT axis and depression by Prange and colleagues. Three decades ago, it was hypothesized that thyroid function was integral to the pathogenesis of, and recovery from, affective disorders due to the numerous interactions among thyroid hormones, catecholamines, and adrenergic receptors in the CNS. Overall, these studies suggested a role for thyroid dysfunction in refractory depression and are consonant with clinical studies, suggesting the existence of an increased rate of hypothyroidism among patients with refractory depression. The use of TRH as a provocative agent for the assessment of HPT axis function evolved rapidly after its isolation and synthesis. Clinical use of a standardized TRH stimulation test, which measures negative feedback responses, revealed blunting of the TSH response in approximately 25 percent of euthyroid patients with major depression. These data have been widely confirmed. The observed TSH blunting in depressed patients does not appear to be the result of excessive negative feedback due to hyperthyroidism because thyroid measures such as basal plasma concentrations of TSH and thyroid hormones are generally in the normal range in these patients. It is possible that TSH blunting is a reflection of pituitary TRH receptor downregulation as a result of median eminence hypersecretion of endogenous TRH. Indeed, the observation that CSF TRH concentrations are elevated in depressed patients as compared to controls supports the hypothesis of TRH hypersecretion but does not elucidate the regional CNS origin of this tripeptide. In fact, TRH mRNA expression in the PVN of the hypothalamus is decreased in patients with

major depression. However, it is not clear whether the altered HPT axis represents a causal mechanism underlying the symptoms of depression or simply a secondary effect of depression-associated alterations in other neural systems.

Corticotropin-Releasing Factor and Urocortins

In the 1950s, it was observed that pituitary extracts contained a factor, referred to as CRF, which could stimulate the release of ACTH from anterior pituitary cells in vivo. After a search spanning nearly three decades, Vale and colleagues isolated and characterized CRF as a 41-amino acid peptide in 1981. The gene for CRF in humans is located on chromosome 8q13 and is composed of 2 exons, with the CRF preprohormone being encoded entirely on exon 2 (Fig. 1.6–1). More recently, the related neuropeptides, urocortin 1, urocortin 2, and urocortin 3, have been identified and share similar gene structures. CRF is the primary hypothalamic ACTH secretagogue in most species, and it also functions as an extrahypothalamic neurotransmitter/neuromodulator in a CNS network which, along with the urocortins, globally coordinates responses to stressors. There is convincing evidence to support the hypothesis that CRF and the urocortins play a complex role in integrating the endocrine, autonomic, immunological, and behavioral responses of an organism to stress.

Although it was originally isolated because of its functions in regulating the HPA axis, CRF is widely distributed throughout the brain. The PVN of the hypothalamus is the major site of CRF-containing cell bodies that influence anterior pituitary ACTH. These neurons originate in the parvocellular region of the PVN and send axon terminals to the median eminence, where CRF is released into the hypothalamic-hypophyseal portal system in response to stressful stimuli. A small group of PVN neurons also project to the brainstem and spinal cord, where they regulate autonomic aspects of the stress response. CRF-containing neurons are also found in other hypothalamic nuclei, the neocortex, the extended amygdala, brainstem, and spinal cord. Central CRF infusion into laboratory animals produces physiological changes and behavioral effects similar to those observed following stress, including increased locomotor activity, increased responsiveness to an acoustic startle, and decreased exploratory behavior in an open field.

In a manner similar to that described for TRH and thyroid hormones, CRF gene expression and content in the PVN are negatively influenced by glucocorticoids (cortisol) and positively regulated by a wide variety of stressors. Adrenalectomy results in an increase in CRF mRNA expression in the PVN, and glucocorticoid replacement decreases CRF mRNA expression in a dose-dependent manner. In contrast to their effects in the PVN, glucocorticoids increase CRF mRNA content in the amygdala rather

than decreasing it.

CRF is also found in the raphe nuclei and the locus coeruleus (LC), the origin of the major serotonergic and noradrenergic projections to the forebrain, respectively, circuits long postulated to play a role in the pathophysiology of depression and anxiety. Increased anxiety observed after direct CNS administration of CRF has been hypothesized to be associated, in part, with increased noradrenergic activity. Stress has been shown to produce an increase in CRF content in the LC and a decrease in CRF concentrations in the median eminence (consistent with increased release). Other studies have shown that CRF-containing nerve terminals impinge upon noradrenergic neurons of the LC and that exogenous CRF applied to those neurons alters their firing rate. Some of the noradrenergic LC neurons, in turn, project to the hypothalamic PVN, where their input increases CRF synthesis and release. As CRF injection into the LC elicits fearful or anxious behavior, one could postulate that stress activates the CRF neurons terminating on the LC noradrenergic neurons, which then may, acting along with other inputs to the PVN, stimulate the stress-induced increased release of CRF from the median eminence. Another hypothesis that has been pursued in recent years is severe stress causes nucleus accumbens dopamine neurons to become insensitive to CRF, resulting in reduced dopamine release and increased anhedonia. Interestingly, adult animals exposed to maternal separation early in life, an animal model for early adverse childhood experiences and a very severe form of stress, exhibit elevated CRF concentrations in the LC, increased CRF mRNA expression, and exaggerated HPA axis responses to stress.

The physiological and behavioral roles of the urocortins are less understood, but several studies suggest that urocortin 2 and 3 are anxiolytic and may dampen the stress response. This has led to the hypothesis that CRF and the urocortins act in opposition, but this is likely an oversimplification. Urocortin 1 is primarily synthesized in the Edinger–Westphal nucleus, lateral olivary nucleus, and supraoptic hypothalamic nucleus. Urocortin 2 is synthesized primarily in the hypothalamus, while urocortin 3 cell bodies are found more broadly in the extended amygdala, perifornical area, and preoptic area.

The CRF system is further complicated by the fact that the effects of CRF and the urocortins are mediated by at least two receptor subtypes, CRF_1- and CRF_2-receptor (Fig. 1.6–3). The CRF_1 receptor is abundantly expressed in the cerebral cortex, cerebellum, medial septum, and anterior pituitary, whereas the CRF_2 receptor is predominantly found in the lateral septum, ventromedial hypothalamus (VMH), and choroid plexus of rodents, but has considerable expression in the human cortex. The CRF_1 receptor appears to be the predominant receptor mediating the effects of CRF in the stress response. The CRF_1 receptor has 4- to 10-fold higher

affinity for CRF than for urocortin 1, with very low affinity for the other urocortins. In contrast, the CRF_2 receptor has a 40-fold higher affinity for the urocortins relative to CRF. Thus, the urocortins have been proposed to be the endogenous ligands for the CRF_2 receptor, but little is known regarding their physiological role. As expected, CRF_1 receptor knockout mice display decreased anxiety-like behavior, have an impaired stress response, and exhibit elevated CRF mRNA expression in the PVN due to a lack of glucocorticoid negative feedback. In contrast, CRF_2 receptor knockout mice display increased anxiety-like behavior and are hypersensitive to stress. Hyperactivity of the HPA axis in major depression remains one of the most consistent findings in biological psychiatry. The reported HPA axis alterations in major depression include hypercortisolemia, resistance to dexamethasone suppression of cortisol secretion (a measure of negative feedback), blunted ACTH responses to intravenous CRF challenge, increased cortisol responses in the combined dexamethasone/CRF test, and elevated CSF CRF concentrations. The exact pathological mechanism(s) underlying HPA axis dysregulation in major depression and other affective disorders remains to be elucidated.

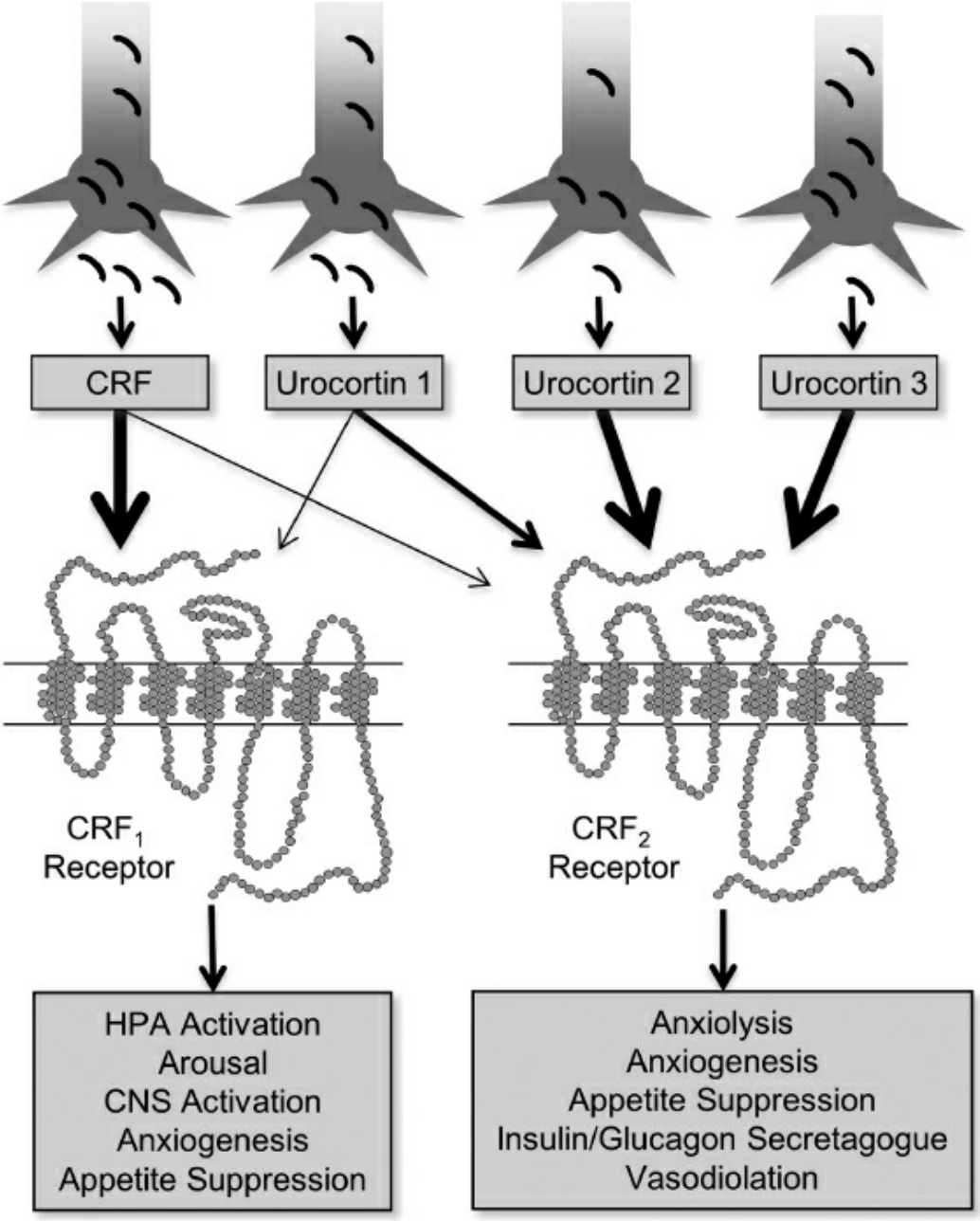

FIGURE 1.6–3. Ligands and receptors of the corticotropin-releasing factor (CRF) signaling network and their putative roles. The figure illustrates the complexity of the CRF system with 4 different ligands modulating two different receptors, each of which regulates divergent physiological processes. The thickness of the arrows represents the relative affinity of each ligand for the respective receptors. (Adapted from Nemeroff CB, Vale WW. The neurobiology of depression: inroads to treatment and new drug discovery. *J Clin Psychiatry*. 2005;66(Suppl 7):5–13.)

Once the phenomenon of HPA axis hyperactivity in patients with major depression was established, many research groups utilized various provocative neuroendocrine challenge tests as a "window into the brain" in attempts to elucidate pathophysiological mechanisms. In normal subjects, the CRF stimulation test, using either rat/human or ovine CRF, yields robust ACTH, β-endorphin, β-lipotropin, and cortisol responses following

intravenous or subcutaneous administration. However, in patients with major depression, blunting of ACTH or β-endorphin secretion with a normal cortisol response has been repeatedly reported. Patients with PTSD, 50 percent of whom also fulfill criteria for major depression, also show blunted ACTH secretion in response to a CRF challenge. Importantly, researchers have reported normalization of the ACTH response to CRF following clinical recovery from depression, suggesting that the blunted ACTH response, like dexamethasone nonsuppression, may be a state-marker for depression. Early-life stress apparently sensitizes the HPA axis and leads to a greater risk of developing depression later in life. Depressed women who were victims of childhood abuse exhibit an exaggerated ACTH and cortisol response to a psychosocial stressor, presumably due to CRF hypersecretion. Depressed men with a history of childhood abuse exhibit marked HPA axis hyperactivity in the combined dexamethasone/CRF test.

Mechanistically, two hypotheses have been advanced to account for the ACTH blunting following exogenous CRF administration. The first hypothesis suggests that pituitary CRF receptor downregulation occurs as a result of hypothalamic CRF hypersecretion. The second hypothesis postulates altered sensitivity of the pituitary to glucocorticoid negative feedback. Substantial support has accumulated favoring the first hypothesis. However, neuroendocrine studies represent a secondary measure of CNS activity; the pituitary ACTH responses principally reflect the activity of hypothalamic CRF rather than that of the corticolimbic CRF circuits. The latter of the two are more likely to be involved in the pathophysiology of depression. A potentially more direct method for the evaluation of extrahypothalamic CRF tone may be obtained from measurements of CRF concentrations in the CSF. A marked dissociation between CSF and plasma neuropeptide concentrations has been described, thus indicating that neuropeptides are secreted directly into CSF from brain tissue as opposed to being derived from plasma-to-CSF transfer. Evidence that CSF CRF concentrations originate from extrahypothalamic CRF neurons has been obtained from studies in which CSF CRF concentrations were repeatedly measured over the course of the day. CSF CRF concentrations in rhesus monkeys do not appear to be entrained with pituitary–adrenal activity. The proximity of corticolimbic, brainstem, and spinal CRF neurons to the ventricular system of the brain suggests that these areas make substantial contributions to the CSF CRF pool.

In a series of studies, significant elevations of CSF CRF concentrations in drug-free patients with major depression or following suicide have been demonstrated. Additionally, severity of depression appears to correlate significantly with CSF CRF concentrations in patients with anorexia nervosa, multiple sclerosis, and Huntington disease. The elevation of CSF CRF concentrations in patients with anorexia nervosa reverts to the normal range as these patients approach normal body weight. No alterations of

CSF CRF concentrations have been reported in other psychiatric disorders including mania, panic disorder, and somatization disorders as compared to controls. It is now clear that patients with early-life trauma in the form of child abuse or neglect exhibit increased CSF CRF concentrations, as has now been demonstrated in patients with major depression, PTSD, and antisocial personality disorder. Of particular interest is the demonstration that the elevated CSF CRF concentrations in drug-free depressed patients are significantly decreased after successful treatment with ECT, indicating that CSF CRF concentrations, like hypercortisolemia, represent a state, rather than a trait, marker. Other studies have confirmed this normalization of CSF CRF concentrations following successful treatment with fluoxetine. One group demonstrated a significant reduction of elevated CSF CRF concentrations in 15 female patients with major depression, who remained depression-free for at least 6 months following antidepressant treatment as compared to little significant treatment effect on CSF CRF concentrations in 9 patients who relapsed in this 6-month period. This suggests that elevated or increasing CSF CRF concentrations during antidepressant treatment may be the harbinger of a poor response in major depression despite early symptomatic improvement. Interestingly, treatment of normal subjects with desipramine or, as noted above, of depressives with fluoxetine is associated with a reduction in CSF CRF concentrations.

In preclinical studies, CRF hypersecretion is associated with CRF-receptor downregulation. Depression is a major determinant of suicide, with more than 50 percent of completed suicides accomplished by patients with major depression. If CRF hypersecretion is a characteristic of depression, evidence of related CRF-receptor downregulation should be evident in the CNS of depressed suicide victims. Indeed, in two studies, a marked decrease in the density of CRF receptors in the frontal cortex of suicide victims as compared to matched control samples was observed.

If CRF hypersecretion is a factor in the pathophysiology of depression, then reducing or interfering with CRF neurotransmission might be an effective strategy to alleviate depressive symptoms. Over the past several years, a number of pharmaceutical companies have committed considerable effort to the development of small molecule CRF_1 receptor antagonists that can effectively penetrate the BBB. Several compounds have been produced with reportedly promising drug-like characteristics.

More than 20 clinical trials have been conducted or are currently underway on the effects of CRF receptor antagonists on mood and anxiety disorders. These clinical trials have been disappointing because of the lack of efficacy for CRF_1 receptor antagonists in patients with major depression or generalized anxiety disorder. It is important to note that these studies have not used narrowly defined depressive patient subtypes such as psychotic, anxious, or atypical depression but instead have utilized a much

broader DSM inclusion criterion. Considering the fact that there is overwhelming evidence that CRF is hypersecreted in depressed patients with a history of childhood abuse and neglect, a clinical trial of CRF receptor antagonists in this discrete subset of patients with major depression would be of considerable interest. Moreover, two studies have documented elevated CSF CRF concentrations in PTSD.

Oxytocin and Vasopressin

The vasopressor effects of posterior pituitary extracts were first described in 1895 and the potent extracts were named vasopressin. In 1953, OT became the first peptide hormone to have its structure elucidated and the first to be chemically synthesized, leading to the Nobel Prize in chemistry being award to Vincent du Vingeaud in 1955. The human OT and AVP genes are situated tandemly in a head-to-tail fashion on chromosome 20p13 separated by a several kilobase intergenic sequence (Fig. 1.6–1). Both peptides are cyclical nonapeptides containing a cysteine–cysteine disulfide bond and differ at only two amino acid residues (Table 1.6–2). Like the sequence homology of the peptides themselves, the genes for OT and AVP share a common structure, suggesting that the two hormones are derived from a single ancestral hormone as a result of a gene duplication event early in vertebrate evolution. The two genes are organized in a tail-to-tail orientation and the OT and AVP mRNA are transcribed from opposite DNA strands toward each other. Each gene consists of 3 exons, with the first exon encoding the 5′ UTR, and the translation initiation codon followed by the signal peptide sequence and the peptide hormone portion of the preprohormone. Exons 2 and 3 encode the neurophysin portion of the prohormone molecule. The AVP prohormone also contains a glycoprotein, copeptin, whose function is unclear but that has been suggested as a biomarker in a number of diseases. Neurophysin is thought to play a role in the posttranslational processing and transport of the peptides, and copeptin is thought to do the same.

Oxytocin and vasopressin mRNAs are among the most abundant messages in the hypothalamus, being heavily concentrated in the magnocellular neurons of the paraventricular and the supraoptic nuclei of the hypothalamus, which send axonal projections to the neurohypophysis. These neurons produce all of the OT and AVP that is released into bloodstream, where these peptides act as hormones on peripheral targets. OT and AVP are generally synthesized in separate neurons within the hypothalamus. OT released from the pituitary is most often associated with functions associated with female reproduction, such as regulating uterine contractions during parturition and the milk ejection reflex during lactation. AVP, also known as antidiuretic hormone, regulates water retention in the kidney and vasoconstriction through interactions with vasopressin V2 and V1a receptor subtypes, respectively. AVP is released

into the bloodstream from the neurohypophysis following a variety of stimuli including plasma osmolality, hypovolemia, hypertension, and hypoglycemia. The actions of OT are mediated via a single receptor subtype (OTR), which is distributed in the periphery and within the limbic CNS. In contrast to the OTR, there are three AVP receptor subtypes, V1a, V1b, and V2 receptors, each of which is G-protein–coupled, seven-transmembrane domain receptors. The V2 receptor is localized in the kidney and is not found in the brain. The V1a receptor is distributed widely in the CNS and is thought to mediate most of the behavioral effects of AVP. The V1b receptor is concentrated in the anterior pituitary and in the hippocampus. Although its function in the brain is largely unknown, V1b knockout mice display decreased aggression and altered social behaviors.

Some parvocellular neurons in the PVN of the hypothalamus also project to the median eminence, where AVP is released into the portal system and delivered to the anterior pituitary. Through interactions with V1b receptors located on corticotrophs in the adenohypophysis, AVP acts to potentiate the effects of CRF on ACTH secretion. AVP is colocalized with CRF in the parvocellular neurons of the PVN. Given the link between HPA axis dysregulation and depression, recent attention has been given to the possible relationship between AVP secretion and psychiatric disorders. Although alterations in CSF AVP concentrations have been reported in patients with major depression, bipolar disorder, schizophrenia, anorexia, and Alzheimer disease, the findings are not as consistent as for CRF and many discrepant reports have appeared. In a postmortem study, an increase in the number of paraventricular AVP neurons colocalized with CRF cells has been reported in depressed patients compared to controls. A selective, nonpeptide V1b receptor antagonist, SSR149415, was in development, but was discontinued in 2008 due to apparent lack of clinical efficacy. Microdialysis experiments have demonstrated that AVP is released within the CNS in response to stressful stimuli.

In addition to the hypophyseal OT and AVP systems, magno- and parvocellular hypothalamic and extrahypothalamic neurons produce OT and AVP, and send projections to the forebrain and brainstem. The release of peptide from these neurons is generally independent of neurohypophysial release and it should be noted that OT and AVP released into the bloodstream do not re-enter the brain due to their inability to traverse the BBB. OT and AVP projections from the PVN to the brainstem regulate a host of autonomic functions. However, in the forebrain, these peptides are now known to regulate a number of processes, ranging from anxiety, learning, and memory to complex social behaviors.

Central OT produces clear anxiolytic effects in animal models. This is particularly evident during lactation in rats, when OT results in a blunted behavioral and ACTH response to a stressful acoustic stimulus. In contrast, central AVP appears to exert anxiogenic effects. In animal models, OT has

been most intensively studied for its role in facilitating specific, complex social behaviors. OT has been reported to facilitate female sexual behavior, increase social interest, and facilitate the onset of maternal behavior. For example, in parturient rats, the onset of maternal behavior is blocked by OT antagonists, whereas maternal behavior can be observed in virgin females after infusion of OT directly into the brain. Likewise, in sheep, mother–infant bonding is facilitated by OT infusions. Studies with OT knockout mice suggest that this peptide plays a specific role in the processing of socially salient stimuli. For example, OT knockout mice have normal nonsocial cognitive abilities but have a specific deficit in the ability to recognize previously encountered individuals, even though olfactory processing is intact. Studies in highly social, monogamous rodents suggest that OT is also involved in the formation of selective social attachments between mates. Furthermore, species differences in OT receptor expression pattern appear to correlate with species differences in social behavior in rodents. For example, monogamous prairie voles have high densities of OT receptors in the striatum, while nonmonogamous species do not (Fig. 1.6–4). Behavioral pharmacological studies demonstrate that these ventral striatal receptors are critical for social bond formation. Ventral striatal OT receptors mediate social reward in mice. Studies in humans suggest that OT plays a similar role in regulating human social cognition. Polymorphisms in the human OT receptor gene predict face recognition abilities, and intranasal OT delivery causes men to rate their partner as more attractive and enhances ventral striatal activity in response to viewing the partner's face. All of these findings have led to the hypothesis that OT is involved in the regulation of the social brain, suggesting that dysregulation of this peptide could potentially explain social deficits in certain psychiatric disorders such as autism. Several studies using intranasal delivery of OT now confirm that this neuropeptide modulates brain function and cognition in humans. For example, intranasal OT enhances trust in economic games and enhances the ability to infer the internal states of others for subtle affective facial expressions. Imaging studies reveal that intranasal OT reduced amygdala activation and reduced coupling of the amygdala to brainstem regions implicated in autonomic and behavioral manifestations of fear in response to fear-inducing visual stimuli. There is evidence that early-life experiences also alter the OT system because women with a history of childhood abuse or neglect exhibit reduced CSF OT concentrations and polymorphisms of the OTR have recently been shown to mediate vulnerability to depression and anxiety. OT dysfunction has also been implicated in ASDs. One study has reported decreased plasma OT concentrations in autistic patients and further suggested that this deficit may be due to alterations in the activities of the prohormone convertases responsible for cleaving OT into its active form. However, this observation must be interpreted cautiously because plasma OT levels are

not necessarily an index of CNS concentrations. OT concentrations in the CSF of autistic patients have not yet been measured. Several studies have suggested that intranasal OT can improve several aspects of social cognition, including emotion recognition and social reciprocity, as well as enhance activation of the reward system in autistic subjects. The OT system is currently viewed as the most promising target for improving social cognition in psychiatric disorders; however, the efficacy of activating the OT system through intranasal OT administration is hindered by the necessity of delivering the poorly brain-penetrant OT peptide to the brain due to a lack of useful small molecule brain-penetrant OT activators. Extrahypothalamic AVP-producing neurons in the extended amygdala are sexually dimorphic, with males having many more AVP-expressing neurons than females. These neurons project through the ventral forebrain to the lateral septum where they form a dense plexus of AVP-containing fibers in males, much more so than in females. Castration diminishes this sex difference and androgen treatment re-establishes the sexually dimorphic pattern. Thus, extrahypothalamic AVP is predicted to be involved in the regulation of sex-specific behaviors in males. Vasopressin has been reported to modulate a variety of behaviors in males, including anxiety, aggression, affiliation, and social attachment, in several animal models. For example, infusion of AVP into the hamster brain stimulates territorial and aggressive behaviors within minutes of administration. Extending this observation to humans, one study reported that individuals with a history of violent tendencies have elevated levels of AVP in the CSF compared to nonviolent controls.

One of the most intriguing features of the AVP system is the species specificity in the behavioral effects of AVP. Consistent with this observation, the neuroanatomical localization of V1a AVP receptors is highly species specific, often with little overlap between even closely related species. In fact, the specific behavioral role of AVP seems to be correlated with the localization of V1a receptors in specific brain regions. For example, AVP facilitates affiliation and social attachment in monogamous mammals. In the prairie vole, a monogamous rodent, AVP has been identified as the neurochemical trigger that stimulates pair bonding between the male and its mate. Comparisons among monogamous rodents and closely related nonmonogamous species have revealed that species differences in social organization are associated with species differences in receptor distribution within the brain. In several monogamous species, including the prairie vole, the vasopressin V1a receptor subtype is abundant in the mesolimbic dopamine reward pathway. In contrast, this region has few V1a receptors in the nonmonogamous, asocial montane vole (Fig. 1.6–4). Infusion of a V1a receptor antagonist directly into the ventral pallidum of the prairie vole completely blocks pair bonding. Thus, AVP released during mating facilitates social bonding by modulating the mesolimbic dopamine pathway

in prairie voles, but cannot do so in nonmonogamous species because of the lack of receptors in that region. Molecular analysis of the V1a receptor genes of these different species has revealed DNA sequences in the promoter of the gene that may be responsible for the differential distribution of the receptor in the brain and thus the differences in behavioral patterns. This variability in distribution across species, along with the association between expression patterns and behavior, has led to the hypothesis that individual differences in receptor expression, due to individual variation in gene promoter elements, could potentially have important behavioral consequences in humans. In fact, three separate genetic association studies now have reported associations between polymorphisms in the V1aR promoter and ASD. Thus, dysregulation of AVP and/or its receptor may represent a risk factor that contributes to the social cognition deficits in autism.

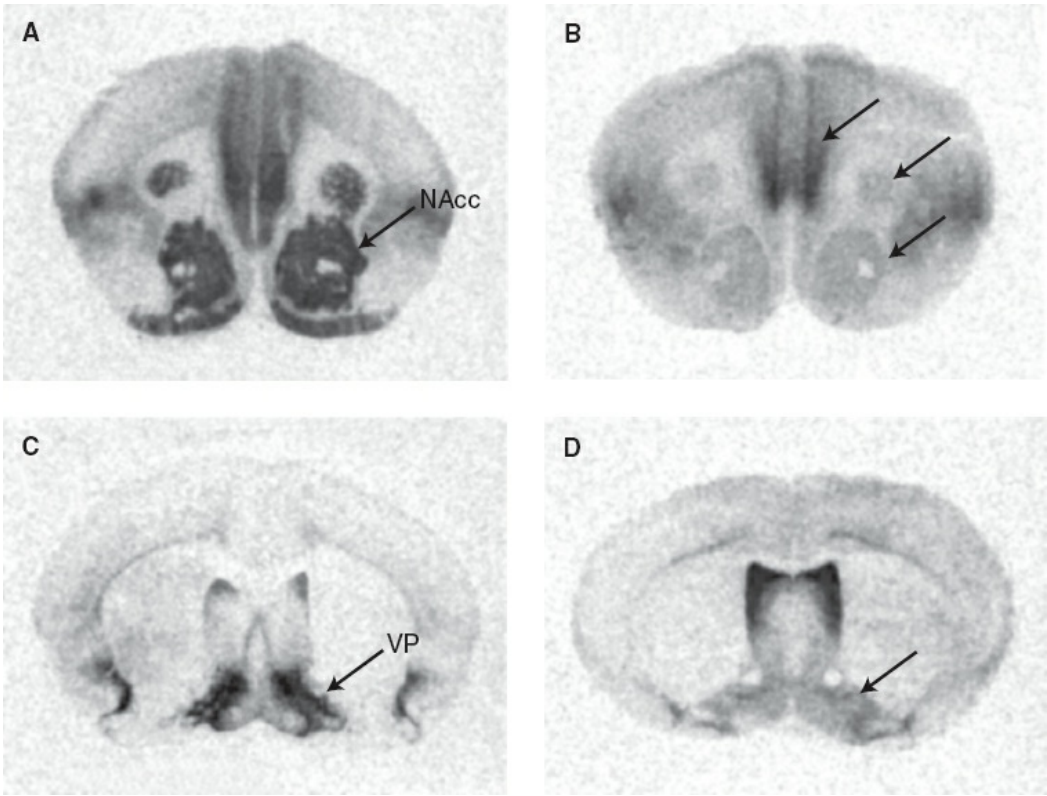

FIGURE 1.6–4. Oxytocin and vasopressin receptor distribution patterns in the brain are associated with social behavior. The upper panels depict receptor autoradiograms illustrating the localization of oxytocin receptor binding in the highly social and monogamous prairie vole (**A**) and the asocial montane vole (**B**). The lower panels illustrate vasopressin V1a binding in the monogamous prairie vole (**C**) and nonmonogamous montane vole (**D**). Note the high density of oxytocin receptor in the nucleus accumbens (NAcc) and V1a receptor binding in the ventral pallidum (VP) of the prairie vole but not the montane vole. These receptor populations are critical for social attachment in monogamous rodents. (Adapted from Young LJ and Wang ZX. The neurobiology of the social bond. *Nature Neuroscience*. 2004;7:1048–1054.)

Neurotensin

NT was isolated, based on its hypotensive properties, from bovine hypothalamus in 1973. The neurotensin-neuromedin N gene was originally cloned from canine ileal mucosa, and complementary deoxyribonucleic acid (cDNA) probes constructed against this form were used to clone the rat gene. The rat gene contains four exon sequences separated by three introns and spans approximately 10.2 kilobases (Fig. 1.6–1). In rats, the neurotensin-neuromedin N sequence is contained in the fourth exon and the single copy of each peptide sequence is bounded and separated by Lys-Arg basic amino pairs. The human NT gene has been localized to chromosome 12 (12q21). In pheochromocytoma (PC-12) neurons in culture, the NT-neuromedin N gene is regulated by lithium, NGF, cAMP activators, and dexamethasone through their effects on a 5′ flanking promoter region. The distribution of the NT-neuromedin N mRNA is generally the same as described for NT-containing neuronal cell bodies, except in the hippocampus and subiculum, where few neurons stain immunohistochemically for NT and yet an abundance of the NT-neuromedin N mRNA is found. NT-producing cells are found in the midbrain (VTA and to a lesser extent the substantia nigra [SN]), ventral striatum, extended amygdala, lateral septum, and arcuate nucleus. The actions of NT are mediated by 3 receptors, the NT_1, NT_2, and NT_3 receptor subtypes. The NT_1 and NT_2 receptors are seven-transmembrane, GPCRs, while NT_3 is a Type I amino acid receptor with a single transmembrane domain and is located intracellularly.

Although NT is found in a number of brain regions, it has been most thoroughly investigated in terms of its association with other neurotransmitter systems, particularly the mesolimbic dopamine system, and has gained interest in research on the pathophysiology of schizophrenia. There are several lines of evidence suggesting that NT and its receptors should be considered as potential targets for pharmacological intervention in this disorder. First, the NT system is positioned anatomically to modulate the neural circuits implicated in schizophrenia. Second, peripheral administration of antipsychotic drugs has been shown to consistently modulate NT systems. Third, there is evidence that central NT systems are altered in schizophrenic patients.

Although likely other neurotransmitter systems are involved, one prevalent model of the pathophysiology of schizophrenia is over-activity of the mesolimbic dopamine system. Within the midbrain, NT-producing neurons are found in the VTA and SN. Within the VTA, NT is found in dense-core vesicles only in tyrosine-hydroxylase–positive staining cell bodies, indicating colocalization with dopamine. These NT-dopamine cells project to the prefrontal cortex, striatum, amygdala, and lateral septum. A subset of those NT-dopamine cells projecting from the VTA to the prefrontal cortex also produces CCK. In contrast to the VTA, the NT-

producing cells in the SN are tyrosine hydroxylase negative. In addition to the NT-producing cells, dense fibers in the VTA, staining positive for NT and originating from projections from the forebrain, do not contain tyrosine hydroxylase. The midbrain also expresses NT receptors, with the vast majority of NT-receptor-containing neurons in the VTA being dopamine-positive neurons. NT producing cells, fibers, and NT receptors are also located in the ventral striatum. Thus, NT is colocalized with DA in the mesolimbic dopamine system, and this system is, in turn, sensitive to NT modulation due to the presence of the NT receptors.

NT was first shown to interact with dopamine systems while undergoing characterization of its potent hypothermic and sedative-potentiating activity. Subsequent work indicated that NT possessed many properties that were also shared by antipsychotic drugs, including the ability to inhibit avoidance, but not escape, responding in a conditioned active avoidance task; the ability to block the effects of indirect dopamine agonists or endogenous dopamine in the production of locomotor behavior; and the ability to elicit increases in dopamine release and turnover. Perhaps most importantly, both antipsychotic drugs and NT neurotransmission enhance sensorimotor gating. Sensorimotor gating is the ability to screen or filter relevant sensory input, deficits that may lead to an involuntary flooding of indifferent sensory data. Increasing evidence suggests that deficits in sensorimotor gating are a cardinal feature of schizophrenia. Both dopamine agonists and NT antagonists disrupt performance on tasks designed to gauge sensorimotor gating. Unlike antipsychotic drugs, NT is not able to displace dopamine from its receptor. As noted above, NT is colocalized in certain subsets of dopamine neurons and is coreleased with dopamine in the mesolimbic and medial prefrontal cortical dopamine terminal regions that are implicated as the site of dopamine dysregulation in schizophrenia. Antipsychotic drugs that act at dopamine D_2 and D_4 receptors increase the synthesis, concentration, and release of NT in those dopamine terminal regions but not in others. That effect of antipsychotic drugs in increasing NT concentrations persists after months of treatment and is accompanied by the expected increase in NT mRNA concentrations, as well as expression of the “immediate early gene” c-fos within hours of initial drug treatment. The altered regulation of NT expression by antipsychotic drugs apparently extends to the peptidases that degrade the peptide because reports have revealed decreased NT metabolism in rat brain slices 24 hours after the acute administration of haloperidol.

While NT is certainly important in schizophrenia, there are numerous other neuropeptides implicated as well. For example, cortical neuropeptides that localize to parvalbumin positive cells, such as somatostatin and CCK, may play a role as well. These cells express parvalbumin calcium binding protein and coordinate cortical GABA release. Coordinated GABA release is central to the generation of gamma

wave rhythmicity. In schizophrenics, gamma oscillations appear to be impaired owing to dysfunctional inhibitory GABA circuits. Neuropeptides such as somatostatin and CCK are also released from GABAergic Martinotti and basket cells, respectively, and contribute to gamma oscillation frequency. Both peptides are also reduced in schizophrenia, and likely contribute to the disease state and represent attractive targets for potential therapies for this disease.

With regard to schizophrenia, decreased CSF NT concentrations have been reported in several populations of patients when compared to controls or other psychiatric disorders. Although treatment with antipsychotic drugs has been observed to increase NT concentrations in the CSF, it is not known whether this increase is causal or merely accompanies the decrease in psychotic symptoms seen with successful treatment. Postmortem studies have shown an increase in NT concentrations in the dopamine-rich Brodmann's area 32 of the frontal cortex, but that result may have been confounded by premortem antipsychotic treatment. Other researchers have found no postmortem alterations in NT concentrations of a wide sampling of subcortical regions. Decreases in NT receptor densities in the entorhinal cortex have been reported in postmortem samples of schizophrenic patients. A critical test of the hypothesis that NT acts as an endogenous antipsychotic-like substance awaits the development of an NT receptor agonist that can penetrate the BBB.

Other Neuropeptides

A number of other neuropeptides have been implicated in the pathophysiology of psychiatric disorders. These include CCK, substance P (SP), and NPY, though numerous others are implicated, far too many to cover in this chapter. A brief overview of the potential involvement of CCK, SP, and NPY neuropeptides in psychiatric disorders are provided as exemplars.

CCK was originally discovered in the gastrointestinal tract and its receptor are found in areas of the brain associated with emotion, motivation, and sensory processing (e.g., cortex, striatum, hypothalamus, hippocampus, and amygdala). CCK is often colocalized with dopamine in the VTA neurons that comprise the mesolimbic and mesocortical dopamine circuits. Like NT, CCK decreases dopamine release. Infusions of a CCK fragment have been reported to induce panic in healthy individuals, and patients with panic disorder exhibit increased sensitivity to the CCK fragment compared to normal controls. Pentagastrin, a synthetic CCK agonist, dose-dependently produced increased blood pressure, pulse, HPA activation, and physical symptoms of panic. Recently, a CCK receptor gene polymorphism has been associated with panic disorder.

The undecapeptide substance P is localized in the amygdala, hypothalamus, periaqueductal gray, LC, and parabrachial nucleus, and is

colocalized with norepinephrine and serotonin. Substance P serves as a pain neurotransmitter, and administration to animals elicits behavioral and cardiovascular effects resembling the stress response. More recent data suggest a role for substance P in major depression and PTSD. Both depressed and PTSD patients exhibit elevated CSF substance P concentrations. Furthermore, in PTSD patients, marked increases in CSF substance P were detected following precipitation of PTSD symptoms. One study has indicated that a substance P receptor (termed the neurokinin 1 or NK1 receptor) antagonist capable of passing the BBB is more effective than placebo, and as effective as paroxetine in patients with major depression with moderate-to-severe symptom severity, though subsequent studies failed to confirm these findings, including in patients with PTSD and/or comorbid alcohol use.

NPY is a 36 amino acid peptide found in the hypothalamus, brain stem, spinal cord, and several limbic structures, and is involved in the regulation of appetite, reward, anxiety, and energy balance. NPY is colocalized with serotonergic and noradrenergic neurons, and is thought to facilitate the containment of negative effects following exposure to stress. Suicide victims with a diagnosis of major depression have been reported to exhibit a pronounced reduction in NPY levels in the frontal cortex and caudate nucleus. Furthermore, CSF NPY levels are decreased in depressed patients. Chronic administration of antidepressant drugs increases NPY concentrations in the neocortex and hippocampus in rats. Plasma NPY levels were found to be elevated in soldiers subjected to the "uncontrollable stress" of interrogation, and NPY levels were correlated with feeling of dominance and confidence during stress. Additionally, low NPY response to stress has been associated with increased vulnerability to depression and PTSD. Remarkably, only a few scattered efforts have focused on developing NPY receptor modulators for psychiatric purposes. Although NPY itself is an anxiolytic-like substance, a synthetic compound that blocks the Y2 receptor will likely also be an anxiolytic because the NPY-Y2 receptor (Y2R) is located presynaptically and negatively regulates NPY release in the brain. Consistent with this hypothesis, Y2R null mice have reduced anxiety-like behavior. There are currently only two small molecule NPY Y2 receptor antagonists, and both are in preclinical development, thus this hypothesis remains to be tested in humans.

FUTURE DIRECTIONS

The current understanding of the role of neuropeptide systems in psychiatric disorders is derived primarily from correlational studies in human samples (e.g., CSF peptide concentrations or postmortem analyses), which preclude inferences of causality, or from animal models, which may or may not accurately reflect psychopathology. The inability to directly modulate CNS neuropeptide receptor activity in human subjects is a major

impediment to the direct examination of the role of neuropeptide systems in human psychopathology. Considerable effort is being devoted to the development of small molecule nonpeptide drugs that readily pass the BBB and selectively modulate CNS peptide receptor activity. Small molecule agonists or antagonists for several neuropeptide systems, including CRF, OT, AVP, orexin and substance P, are the subject of preclinical and clinical investigations. Over the next decade, these new pharmacological tools will likely contribute significantly to the general understanding of the roles of these peptides in both normal human behavior and various psychopathologies. Small molecule drugs targeting neuropeptide receptors will undoubtedly lead to novel pharmacotherapy approaches for the treatment of psychiatric disorders such as anxiety disorders, depression, and ASDs. Small molecule agonists or antagonists will also likely lead to the development of novel PET ligands, allowing visualization of peptide receptors in the CNS of human subjects, a great unmet need.

In addition to drug development and novel brain imaging tools, advances in psychiatric genetics are likely to reveal novel relationships between neuropeptide systems and psychopathology over the next few years. Polymorphisms in several neuropeptide receptor systems have already been implicated as risk factors in psychiatric disorders. Combining brain imaging techniques with genetic analyses will aid in understanding how these polymorphisms affect brain functioning. The plethora of recent GWAS studies in patients with mood and anxiety disorders, and schizophrenia will likely identify novel neuropeptides associated with disease vulnerability. Finally, psychopharmacogenomics, which examines how genotype influences clinical responses to drugs, may lead to individualized therapies targeting peptide systems based on the patient's genotype. This burgeoning field has been helped along with technology to be able to fully sequence a patient genome in a few days time, but still needs to be broadly informed by examining patient cohorts in order to identify relevant genotypes.

Clearly, the complexity of the brain's rich neuropeptide systems and their contribution to mental health are just beginning to be understood. This area of research will continue to provide novel insights into the biological basis of psychopathology over the next few decades and will likely produce the next generation of pharmacological interventions for psychiatric disorders.

References

Alhajji L, Nemeroff CB. Personalized medicine and mood disorders. *Psychiatr Clin North Am*. 2015;38(3):395–403.

Andari E, Duhamel JR, Zalla T, Herbrecht E, Leboyer M, Sirigu A. Promoting social behavior with oxytocin in high-functioning autism spectrum disorders. *Proc Natl Acad Sci*. 2010;107(9):4389–4394.

Binder EB, Nemeroff CB. The CRF system, stress, depression and anxiety-insights from human genetic studies. *Mol Psychiatry*. 2010;15(6):574–588.

Block SG, Nemeroff CB. Emerging antidepressants to treat major depressive disorder. *Asian J Psychiatr*. 2014;12:7–16.

Boules MM, Fredrickson P, Muehlmann AM, Richelson E. Elucidating the role of neurotensin in the pathophysiology and management of major mental disorders. *Behav Sci (Basel)*. 2014;4(2):125–153.

Brothers SP, Saldanha SA, Spicer TP, et al. Selective and brain penetrant neuropeptide y y2 receptor antagonists discovered by whole-cell high-throughput screening. *Mol Pharmacol*. 2010;77(1):46–57.

Brothers SP, Wahlestedt C. Therapeutic potential of neuropeptide Y (NPY) receptor ligands. *EMBO Mol Med*. 2010;2(11):429–439.

Bunevicius R. Thyroid disorders in mental patients. *Curr Opin Psychiatry*. 2009;22(4):391–395.

Caceda R, Kinkead B, Nemeroff CB. Involvement of neuropeptide systems in schizophrenia: human studies. *Int Rev Neurobiol*. 2007;78:327–376.

Coccaro EF, Lee R, Owens MJ, Kinkead B, Nemeroff CB. Cerebrospinal fluid substance P-like immunoreactivity correlates with aggression in personality disordered subjects. *Biol Psychiatry*. 2012;72(3):238–243.

Donaldson ZR, Young LJ. Oxytocin, vasopressin, and the neurogenetics of sociality. *Science*. 2008;322(5903):900–904.

Fliers E, Kalsbeek A, Boelen A. Beyond the fixed setpoint of the hypothalamus-pituitary-thyroid axis. *Eur J Endocrinol*. 2014;171(5):R197–208.

Gordon I, Vander Wyk BC, Bennett RH, et al. Oxytocin enhances brain function in children with autism. *Proc Natl Acad Sci*. 2013;110(52):20953–20958.

Green JJ, Hollander E. Autism and oxytocin: new developments in translational approaches to therapeutics. *Neurotherapeutics*. 2010;7(3):250–257.

Griebel G, Holsboer F. Neuropeptide receptor ligands as drugs for psychiatric diseases: the end of the beginning? *Nat Rev Drug Discov*. 2012;11(6):462–478.

Johnson ZV, Young LJ. Neurobiological mechanisms of social attachment and pair bonding. *Curr Opin Behav Sci*. 2015:3:38–44.

Kamath J, Yarbrough GG, Prange AJ, Jr, Winokur A. The thyrotropin-releasing hormone (TRH)-immune system homeostatic hypothesis. *Pharmacol Ther*. 2009;121(1):20–28.

Lemos JC, Wanat MJ, Smith JS, et al. Severe stress switches CRF action in the nucleus accumbens from appetitive to aversive. *Nature*. 2012;490(7420):402–406.

Myers AJ, Williams L, Gatt JM, et al. Variation in the oxytocin receptor gene is associated with increased risk for anxiety, stress and depression in individuals with a history of exposure to early life stress. *J Psychiatr Res*. 2014;59:93–100.

Neumann ID, Landgraf R. Balance of brain oxytocin and vasopressin: implications for anxiety, depression, and social behaviors. *Trends Neurosci*. 2012;35(11):649–659.

Rilling JK, Young LJ. The biology of mammalian parenting and its effect on offspring social development. *Science*. 2014;345(6198):771–776.

Ross HE, Young LJ. Oxytocin and the neural mechanisms regulating social cognition and affiliative behavior. *Front Neuroendocrinol*. 2009;30(4):534–547.

Scheele D, Wille A, Kendrick KM, et al. Oxytocin enhances brain reward system responses in men viewing the face of their female partner. *Proc Natl Acad Sci U S A*. 2013;110(50):20308–20313.

Skuse DH, Lori A, Cubells JF, et al. Common polymorphism in the oxytocin receptor gene (OXTR) is associated with human social recognition skills. *Proc Natl Acad Sci*. 2014;111(5):1987–1992.

Strand FL. Neuropeptides: general characteristics and neuropharmaceutical potential in treating CNS disorders. *Prog Drug Res*. 2003;61:1–37.

Watanabe T, Abe O, Kuwabara H, et al. Mitigation of sociocommunicational deficits of autism through oxytocin-induced recovery of medial prefrontal activity: a randomized trial. *JAMA Psychiatry*. 2014;71(2):166–175.

Young LJ. Oxytocin, social cognition and psychiatry. *Neuropsychopharmacology*. 2015;40(1):243–244.

Young LJ, Barrett CE. Neuroscience. Can oxytocin treat autism? *Science*. 2015;347(6224):825–826.

▲ 1.7 Neurotrophic Factors

FRANCIS S. LEE, M.D., PH.D., AND MOSES V. CHAO, PH.D.

Neurotrophins are a unique family of polypeptide growth factors that influence the proliferation, differentiation, survival, and death of neuronal and nonneuronal cells. The effects of neurotrophins depend upon their levels of availability, their affinity of binding to transmembrane receptors, and the downstream signaling cascades that are stimulated after receptor activation. While the biological roles for neurotrophins were initially characterized during development of the nervous system, it is now clear that they have multiple roles in the adult nervous system, such as regulating synaptic connections, synapse structure, neurotransmitter release, LTP, mechanosensation, pain, and synaptic plasticity. Alterations in neurotrophin levels have been implicated in neurodegenerative disorders such as Alzheimer disease and Huntington disease (HD), and as well as psychiatric disorders such as anxiety disorders, depression, and substance abuse. These recent insights provide new approaches to understanding the etiology and treatment of psychiatric disorders.

THE NEUROTROPHIN FAMILY

A large number of polypeptide factors affect the survival, growth, and differentiation of the nervous system. The neurotrophins, comprising of NGF, BDNF, NT-3, neurotrophin-4 (NT-4), are best understood and most widely expressed in the nervous system. The neurotrophins are initially synthesized as precursors or proneurotrophins that are cleaved to release the mature, active proteins. The mature proteins, approximately 12 to 14

kDa in size, form stable, noncovalent dimers and are normally expressed at very low levels during development. Proneurotrophins are cleaved intracellularly by furin or proconvertases utilizing a highly conserved dibasic amino acid cleavage site to release carboxy-terminal mature proteins or extracellularly by plasmin or by matrix metalloproteases.

The mature proteins mediate neurotrophin actions by selectively binding to members of the tropomyosin-related kinase (Trk) family of receptor tyrosine kinases to regulate neuronal survival, differentiation, and synaptic plasticity. In addition, all mature neurotrophins interact with p75 neurotrophin receptor ($p75^{NTR}$), which can modulate the affinity of Trk neurotrophin associations. NGF was the first identified neurotrophic factor, and has a restricted distribution. In the PNS, it acts on sympathetic neurons, as well as sensory neurons involved in nociception and temperature sensation. In the CNS, NGF promotes the survival and functioning of cholinergic neurons in the basal forebrain. These neurons project to the hippocampus and are believed to be important for memory processes, which are particularly affected in Alzheimer disease. The other neurotrophins are more widely expressed in the CNS. BDNF and NT-3 are highly expressed in cortical and hippocampal structures, and have been linked to the survival and functioning of multiple neuronal populations.

NEUROTROPHIN RECEPTORS

Neurotrophins are unique in exerting their cellular effects through the actions of two different receptors: the Trk receptor tyrosine kinase and $p75^{NTR}$, a member of the tumor necrosis factor (TNF) receptor superfamily. Trk receptors consist of an extracellular ligand-binding region, a single transmembrane domain, and a highly conserved intracellular tyrosine kinase domain. The $p75^{NTR}$ receptor consists of an extracellular ligand-binding region, a single transmembrane domain, and an intracellular portion containing a protein-association region termed the death domain (Fig. 1.7–1). All neurotrophins bind to the p75 receptor. There are three vertebrate *trk* receptor genes: *trkA, trkB,* and *trkC.* All Trk receptors exhibit high conservation in their intracellular domains, including the catalytic tyrosine kinase domain and the juxtamembrane domain. The Trk receptors also exhibit a number of truncated isoforms. There are no sequence similarities between Trk and p75 receptors in their either ligand-binding or cytoplasmic domains.

Neurotrophins bind as dimers to Trk family members, leading to receptor dimerization and activation of the catalytic tyrosine protein kinase domains. The dimerized Trk receptors autophosphorylate several key intracellular tyrosine residues, which rapidly initiate intracellular signaling cascades. This is accomplished by the phosphorylated tyrosines on the receptor acting as recognition sites for the binding of specific adaptor proteins that contain phosphotyrosine-binding motifs such as Src

homology domain 2 (SH2). In particular, the Shc adaptor protein links the activated Trk receptor to two separate intracellular signaling pathways that mediate the majority of the biological effects of neurotrophins.

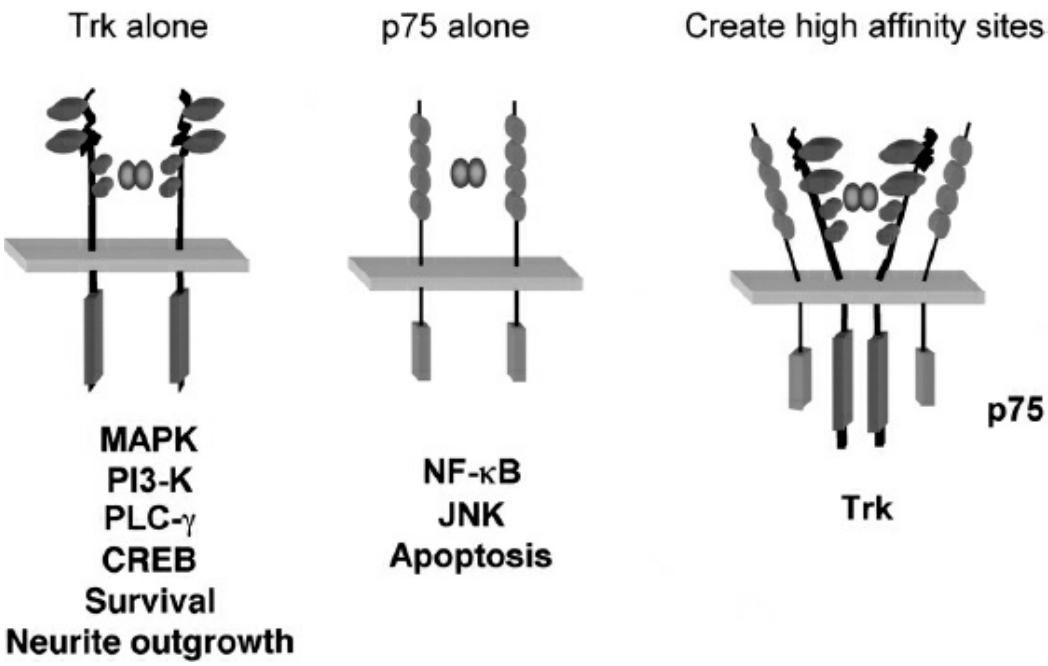

FIGURE 1.7–1. Neurotrophin receptor signaling. Neurotrophins bind to Trk tyrosine kinase receptors (*right*) and p75 neurotrophin receptors ($p75^{NTR}$) (*middle*). Trk receptors mediate differentiation and survival signaling through mitogen-activated protein kinase (MAPK), phosphatidylinositol-3-kinase (PI3-K), and phospholipase C-γ (PLC-γ) pathways, which lead to effects on transcription factors, such as the cyclic adenosine monophosphate response element–binding protein (CREB). Trk receptors contain IgG domains for ligand binding and a catalytic tyrosine kinase sequence (*left*) in the intracellular domain. $p75^{NTR}$ mediates apoptotic and cell migration responses through nuclear factor κB (NF-κB) and c-Jun N-terminal kinase (JNK) pathways. The extracellular part of $p75^{NTR}$ contains four cysteine-rich repeats; the intracellular domain contains a death domain (*middle*). Interactions between Trk and $p75^{NTR}$ receptors can lead to changes in binding affinity for neurotrophin (*right*).

The primary survival pathway involves Shc, linking Trk receptor activation to increases in phosphotidylinositol-3-kinase (PI3-K) activity. This in turn activates another protein kinase, Akt (protein kinase B), which has multiple effects on the cell's apoptotic pathways. Also, Shc phosphorylation by Trk receptor activation leads to increases in Ras and MAP kinase activities. These events in turn influence transcriptional events such as the induction of the cyclic adenosine monophosphate response element–binding protein (CREB) transcription factor. CREB produces a multitude of effects on the cell cycle, neurite outgrowth, and synaptic plasticity. In addition, phospholipase-C-γ (PLC-γ) binds to activated Trk receptors and initiates an intracellular signaling cascade release of inositol phosphates and activation of protein kinase C (PKC). Trk receptor activation leads to a multitude of downstream signaling events, leading to

changes in transcriptional programs.

NGF binds most specifically to TrkA, BDNF and NT-4 to TrkB, and NT-3 to TrkC receptors. The $p75^{NTR}$ receptor can bind to each neurotrophin but has the additional capability of regulating a Trk's affinity for its cognate ligand. Trk and $p75^{NTR}$ receptors have been referred to as high- and low-affinity receptors, respectively. However, this is not correct since TrkA and TrkB actually bind mature neurotrophins with an affinity of 10^{-9} to 10^{-10} M, which is lower than the high-affinity site ($K_d = 10^{-11}$ M). Also, the precursor form of NGF displays high-affinity binding to $p75^{NTR}$. Trk-mediated responsiveness to low concentrations of NGF is dependent upon the relative levels of $p75^{NTR}$ and TrkA receptors and their combined ability to form high-affinity sites. This is important since the ratio of receptors can determine responsiveness and ultimately neuronal cell numbers.

Although p75 and Trk receptors do not bind to each other directly, there is evidence that complexes form between the two receptors. Perhaps as a result of these interactions, increased ligand selectivity can be conferred onto Trk receptors by the p75 receptor. One way of generating specificity is by imparting greater discrimination of ligands for the Trk receptors (Fig. 1.7–2). For example, BDNF, NT-3, and NT-4/5 can each bind to the TrkB receptor, but in the presence of p75, only BDNF provides a functional response. Likewise, NGF and NT-3 both can bind to TrkA, but p75 restricts the signaling of TrkA to NGF and not to NT-3 (Fig. 1.7–2). Hence, p75 and Trk receptors interact in order to provide greater discrimination among different neurotrophins.

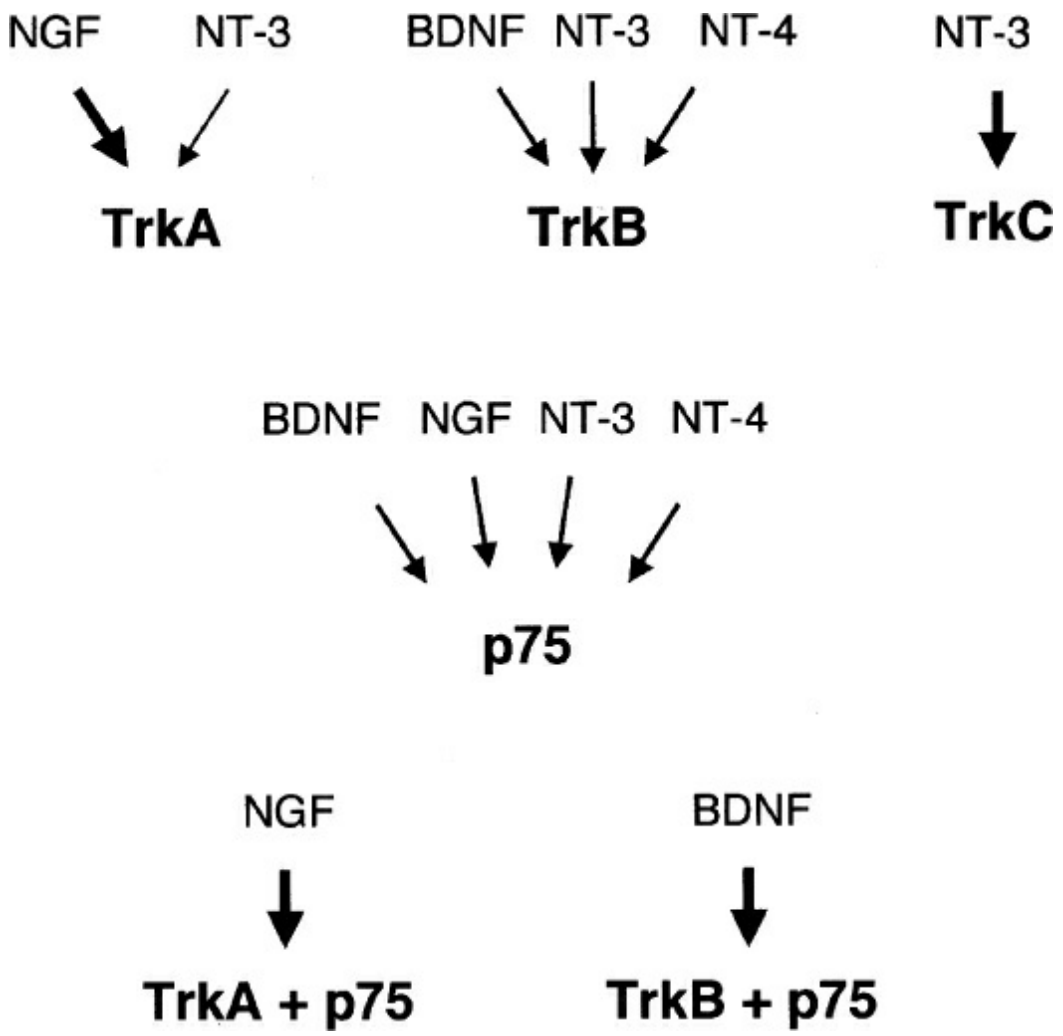

FIGURE 1.7–2. Neurotrophin-binding specifities. All neurotrophins bind to p75 neurotrophin receptors ($p75^{NTR}$). Neurotrophins bind selectively to specific tropomyosin-related kinase (Trk) receptors, and this specificity can be altered by $p75^{NTR}$. Several neurotrophins, neurotrophin-3 (NT-3) and neurotrophin-4 (NT-4), can bind to multiple Trk receptors. BDNF, brain-derived neurotrophic factor; NGF, nerve growth factor.

NEUROTROPHIC FACTORS AND DEVELOPMENT

The formation of the vertebrate nervous system is characterized by widespread programmed cell death, which determines cell number and appropriate target innervation during development. Neurotrophins are highly expressed during early development and have been shown to be essential for the survival of selective populations of neurons during different developmental periods. The neurotrophic hypothesis provides a functional explanation for the role of neurotrophic factors in the development of the nervous system (Fig. 1.7–3). During development, neurons approaching the same final target vie for limited amounts of target-derived neurotrophic factors. In this way, the nervous system molds itself to maintain only the most competitive and appropriate connections. Competition among neurons for limiting amounts of neurotrophin molecules produced by target cells accounts for selective cell survival (Fig.

1.7–3). Two predictions emanate from this hypothesis. First, the efficacy of neuronal survival will depend upon the amounts of trophic factors produced during development. Second, specific receptor expression in responsive cell populations will dictate neuronal responsiveness.

On one level, neurotrophins fit well with the neurotrophic hypothesis, as many peripheral neuronal subpopulations depend on a specific neurotrophin during the period of naturally occurring cell death. In the CNS, the overlapping expression of multiple neurotrophin receptors and their cognate ligands allows for more diverse connectivity, which extends well into adulthood. In addition, it is clear now that neurons can release neurotrophins that act on themselves (autocrine transmission) or can be anterogradely transported down axons and act on neighboring neurons. Also, glial cells can release neurotrophins that act upon neurons in a paracrine fashion. In the periphery, neurotrophin retrograde signaling occurs through a pathway that must efficiently transmit information over long distances, at times over a meter.

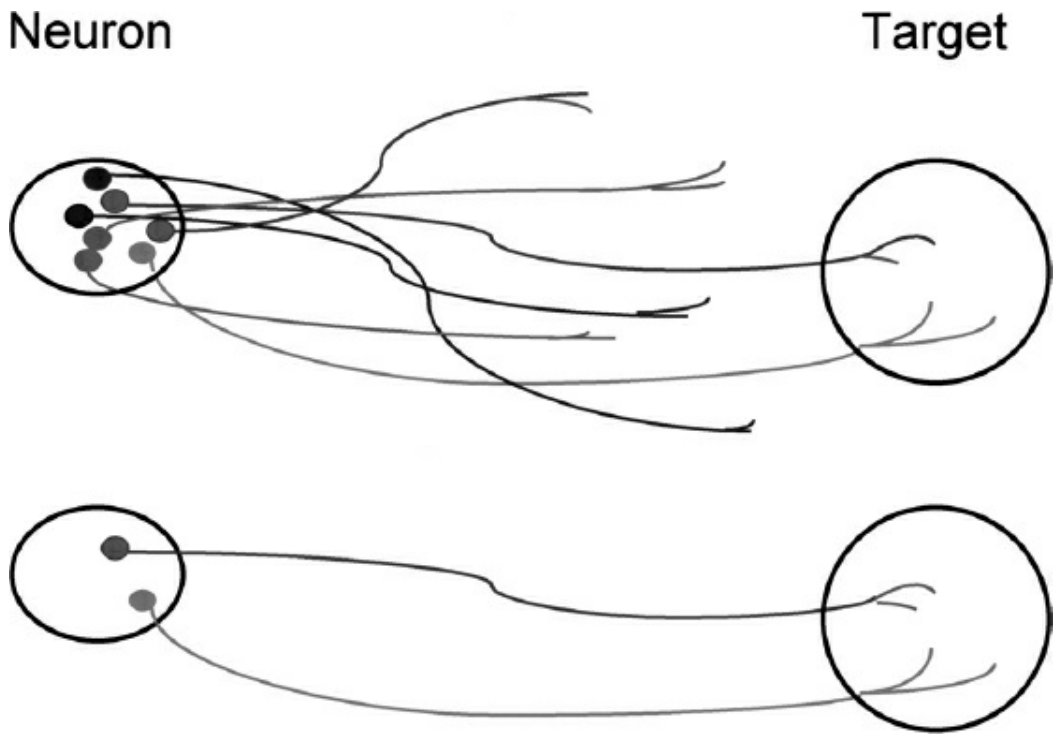

FIGURE 1.7–3. The neurotrophin hypothesis. Neurons compete for limited quantities of neurotrophins in target regions, which leads to selective neuronal survival. Levels of target-derived neurotrophins and neurotrophin receptors will determine efficacy of survival and responsiveness of the neurons. The ability to form high–affinity-binding sites allows for greater responsiveness under limiting quantities of trophic factors. Lack of trophic support or incorrect targeting of axons to the wrong target results in programmed cell death.

Neurotrophins promote cell survival and differentiation during neural development. Paradoxically, they can also induce cell death. $p75^{NTR}$ serves as a proapoptotic receptor during developmental cell death and after injury to the nervous system (Fig. 1.7–1). Increases in $p75^{NTR}$ expression are

responsible for apoptosis in embryonic retinas and sympathetic neurons during the period of naturally occurring neuronal death. Whereas BDNF binding to $p75^{NTR}$ in sympathetic neurons causes rapid cell death, NGF binding to the TrkA receptor on the same neurons provides a survival signal. In the context of neurotrophin processing, proneurotrophins are more effective than mature NGF in inducing $p75^{NTR}$-dependent apoptosis. These results suggest that the biological action of the neurotrophins can be regulated by proteolytic cleavage, with proforms preferentially activating $p75^{NTR}$ to mediate apoptosis and mature forms selectively activating Trk receptors to promote survival.

Neurotrophins may induce apoptosis in order to prune neurons efficiently during periods of developmental cell death. In addition to competing for trophic support from the target, neurons must establish connections with the proper target. In the event of mistargeting, neurons may undergo apoptosis if the appropriate set of trophic factors or receptors is not encountered. For example, BDNF can cause sympathetic cell death by binding to $p75^{NTR}$ in the absence of TrkB. However, NT-4 causes $p75^{NTR}$-mediated cell death in BDNF-dependent trigeminal neurons, due presumably to preferential $p75^{NTR}$ stimulation over TrkB. Therefore, Trk and $p75^{NTR}$ receptors can give opposite outcomes in the same cells. Cell death mediated by $p75^{NTR}$ may be important for the refinement of correct target innervation during development.

AXONAL TRANSPORT

Specificity of the biological effects of neurotrophins can also be modulated by the intracellular location of the neurotrophin ligand receptor complex. During development, neurotrophins are produced and released from the target tissues and become internalized into vesicles, which are then transported to the cell body. Interestingly, the biological effects of neurotrophins require that signals are conveyed over long distances from the nerve terminal to the cell body. Therefore, a central theme of the neurotrophic hypothesis is that neuronal survival and differentiation depend upon the retrograde signaling of trophic factors produced at the target tissue.

Each neurotrophin binds to transmembrane receptors and undergoes internalization and transport from axon terminals to neuronal cell bodies. Measurements of ^{125}I-NGF transport from distal axons to the cell body in compartment chambers indicate a rate from 3 to 10 mm per hour. Both Trk and $p75^{NTR}$ receptors undergo retrograde transport. The term “signaling endosome” has been coined to describe membrane vesicles that carry Trk, $p75^{NTR}$, and NGF.

A complex of NGF–TrkA has been found in clathrin-coated vesicles and endosomes, giving rise to the model that NGF and Trk are components of the retrograde signal. Several tyrosine-phosphorylated proteins are

associated with the TrkA receptor during transport, suggesting that signaling by neurotrophins persists following internalization of their receptors. Internalization of NGF from axon terminals is necessary for phosphorylation and activation of the CREB transcription factor, which leads to changes in gene expression and increased neuronal cell survival. These events likely require the internalization and transport of activated Trk receptors and result in a survival response.

Neurotrophins and Synaptic Plasticity

Recent studies have established that neurotrophic factors play significant roles in influencing synaptic plasticity in the adult brain. Many neuronal populations are not only dependent upon these neurotrophins for their survival but also for modulating neuronal activity. Developmental regulation of synaptic plasticity in the visual system is illustrated by the formation of ocular dominance columns in layer 4 of the cortex, which can be strongly influenced by exogenous neurotrophins such as BDNF. Also, the effects upon the visual system can be observed using blocking antibodies for the neurotrophins as well as neurotrophin antagonists (TrkB–IgG fusion proteins that bind neurotrophins), indicating that an alteration in the levels of endogenous neurotrophins has dramatic consequences on the formation of critical periods.

Modulation of synaptic plasticity in the differentiated adult brain has also been demonstrated in the hippocampus in a series of studies. BDNF promoted the induction of a synaptic strengthening, termed LTP, in hippocampal slices, while blocking reagents such as the TrkB–IgG fusion protein interfered with the induction of LTP. In addition, hippocampal preparations containing little or no BDNF gave rise to the same reduction in LTP, suggesting that there was a minimal quantity of BDNF required for the modulation of LTP. Subsequent addition of extra BDNF or adenoviral expression of BDNF to these preparations from mutant mice restored LTP. Neurotrophins have also been shown to evoke other forms of synaptic transmission. Exogenous BDNF or NT-3 has been shown to induce enhanced evoked responses in both hippocampal preparations and neuromuscular junctions. Thus, neurotrophins can modulate synaptic strengthening and neurotransmission as well as promote cell survival and axonal and dendritic growth. Conversely, the precursor of BDNF, proBDNF has been shown to facilitate LTD in hippocampal slices by activation of $p75^{NTR}$ and exogenous proBDNF induces synaptic depression at the neuromuscular synapse, suggesting that proBDNF opposes the actions of mature BDNF on synaptic plasticity.

Neurotrophins and Behavior

A recent series of studies on genetically modified mice with reduced levels

of BDNF have indicated striking effects upon adult brain function and behavior. These studies are important as earlier neurotrophin knockout mice studies were limited due to embryonic lethality or early postnatal death. However, heterozygous $BDNF^{+/-}$ mice in which BDNF levels are reduced by approximately one-half are viable and display a number of behaviors suggestive of impulse control abnormalities. In the absence of normal levels of BDNF, mice exhibit enhanced aggressiveness, hyperactivity, and hyperphagia. ICV infusion of BDNF or NT-4 led to a striking reversal of the feeding phenotype. In these heterozygous $BDNF^{+/-}$ mice, serotonergic neuronal functioning was abnormal in the forebrain, cortex, hippocampus, and hypothalamus. Most strikingly, administration of fluoxetine, an SSRI, ameliorated the aggressive behavior, hyperphagia, and hyperlocomotor activity. In addition, a region-specific conditional deletion of BDNF in the brains of postnatal mice also led to hyperphagia, hyperactivity, as well as higher levels of anxiety as measured by a light/dark exploration test. This study and other conditional BDNF mice demonstrated that the feeding phenotype and the other behavioral abnormalities were mediated by the functioning of BDNF in the CNS as compared to any peripheral actions of the neurotrophin.

Lack of BDNF also created defects in memory tasks, consistent with defects in LTP found in the hippocampal slice studies. Haploinsufficient $BDNF^{+/-}$ mice had impairments in spatial memory tasks such as the Morris water maze. Abnormal behaviors elicited by partial deletion of BDNF indicate a significant role for this neurotrophin in higher-order behaviors, which have clinical correlates to psychiatric disorders, especially those associated with alteration in central serotonergic functioning.

OTHER NEUROTROPHIC FACTORS

Several prominent neurotrophic factor families carry out similar functions as the neurotrophins. Glial-derived neurotrophic factor (GDNF) is an 18-kDa protein, originally isolated from an astrocyte cell line and it is produced by many types of neurons. It represents one of the most potent trophic factors for dopaminergic neurons. In both in vitro and in vivo studies, GDNF has been shown to maintain the survival of dopaminergic neurons in the midbrain as well as neurons in the myenteric plexus in the gut. Due to its trophic effects on dopaminergic neurons, it has been considered a potential therapeutic agent for Parkinson disease.

GDNF binds to a protein, GFRα1, which is anchored to the plasma membrane by a glycophospholipid. Other ligands have also been discovered, namely, artemin, neurturin, and persephin, which recognize specific GFRα receptors. This ligand–receptor complex then associates with Ret, a receptor tyrosine kinase, which, like the Trk receptors, undergoes dimerization and becomes catalytically active. Phosphotyrosine-

binding adaptor proteins such as Shc then bind to the Ret receptor and mediate downstream signaling cascades such as the MAP kinase pathway. Mutations in the Ret receptor and GFRα1 have been associated with Hirschsprung disease, a disorder caused by the lack of development of myenteric plexus neurons, leading to abnormal gut motility.

Ciliary neurotrophic factor (CNTF) belongs to a family of cytokines, including leukemia inhibitory factor (LIF) and interleukin-6, which maintain the survival of ciliary neurons as well as motor neurons. Due to its ability to rescue motor neurons after axotomy in animal studies, CNTF has been investigated as a therapeutic agent for motor neuron diseases such as ALS. These factors utilize a receptor complex consisting of a plasma-membrane–bound CNTF-binding protein (CNTFα), a glycoprotein (gp130), and a LIF receptor (LIFR) to transduce signals. Upon formation of this complex, a soluble tyrosine kinase, the Janus kinase (JAK), is activated and leads to the activation of a specific family of transcription factors termed Signal Transducer and Activator of Transcription (STAT) proteins.

Therefore, trophic factors exemplified by NGF, CNTF, and GDNF and their family members all utilize intracellular tyrosine phosphorylation to mediate neuronal cell survival. CNTF acts through a complex of a CNTF receptor, gp130, and LIFR subunits that are linked to the JAK/STAT signaling molecules, whereas the GDNF receptor consists of the c-Ret receptor tyrosine kinase and a separate α-binding protein.

CLINICAL CORRELATES

Neurotrophic factors regulate numerous neuronal functions in development and adult life and in response to injury. As a result, neurotrophins have been implicated in the pathophysiology of a wide variety of neurodegenerative and psychiatric disorders and have been considered as a therapeutic strategy for many neuropsychiatric disorders. It should be emphasized though that few human diseases affecting the nervous system have been shown to be caused by a defect in the neurotrophins or their receptors. Still, the finding that neurotrophic factors modulate neuronal survival and axonal growth was the initial rationale for potential clinical correlates to neurodegenerative disorders and neuronal injury such as Alzheimer disease, Parkinson disease, HD, and ALS as well as spinal cord injury. The additional effects of neurotrophic factors on synaptic connections, synaptic plasticity, and neurotransmission have formed the basis for their association with psychiatric disorders such as depression and substance abuse. In these conditions, the response to acute and chronic environmental changes leads to alterations in neuronal function.

The hypothesis underlying these clinical correlations as well as development of therapeutic strategies using neurotrophic factors assumes that these disease states result in either (1) decreased availability of

neurotrophins for the affected neurons, (2) a decreased number of neurotrophin receptors on the affected neurons, and/or (3) decreased neuronal survival. These deficits can be ameliorated by the addition of neurotrophic factors.

An example of how decreased availability of neurotrophins can account for a neuropsychiatric clinical problem is HD, which is characterized by progressive death of medium spiny neurons in the striatum. One mechanism to explain the pathogenic effects of mutant huntingtin protein is through a decrease in BDNF levels brought on by lower anterograde transport from the cortex to the striatum. Other reported effects have been on transcription of the BDNF gene by mutant huntingtin. In the case of Alzheimer disease, addition of NGF or BDNF to vulnerable regions of the brain, such as the entorhinal cortex or hippocampus, can increase spatial learning and memory retention.

The assumption has been that exogenous neurotrophic factors would provide symptomatic treatment for the disease state rather than a cure for the core pathophysiology of these nervous system disorders. A more complete cell expression profile of the neurotrophins and their receptors in well-defined circuits will provide strategies to target degenerative disorders of different origins.

Neurodegenerative Disorders

The initial clinical correlation to Alzheimer disease was made in the 1980s based on studies on aged animals that showed that cholinergic neurons in the basal forebrain could be rescued with ICV NGF, resulting in concomitant improvements in memory function. Subsequent animal studies of impaired motor neuron populations demonstrated that other neurotrophins, BDNF, NT-3, NT-4, and CNTF, could rescue those neurons in an axotomized facial nerve and sciatic nerve. In addition, mutant mouse models of motor neuron disease (progressive motor neuron disease, *wobbler*), in which there was motor neuron degeneration, demonstrated that BDNF and CNTF could increase the number of motor neurons and improve motor performance. These studies led to the therapeutic strategy to attempt to treat degenerative diseases affecting motor neurons with neurotrophins.

In the 1990s, great effort was focused on studying whether neurotrophic factors could be used as a treatment strategy for ALS, a progressive neurodegenerative disorder that specifically affects motor neurons and leads to death due to respiratory failure. With the development of recombinant forms of the neurotrophic factors, namely, BDNF, clinical trials have taken place on patients with ALS. Subcutaneous or intrathecal delivered BDNF had minimal beneficial effect and was associated with side effects such as pain and gastrointestinal symptoms. It was due to these side effects that decreased doses were used as compared to the doses in the

animal studies. Similarly, use of another neurotrophic factor, CNTF, also led to even more significant side effects such as fever, pain, and anorexia, which also limited the doses used. These multisite clinical trials highlighted the challenges of delivery of large quantities of these proteins to CNS and PNS neurons. Similar clinical studies using NGF for the treatment of patients with Alzheimer disease and diabetic neuropathy encountered similar hurdles involving problems of delivery and uncertain pharmacokinetics of the proteins.

Although these clinical trials have been disappointing, there is growing evidence that several specific neurodegenerative diseases would benefit from increasing the levels of neurotrophins. HD is caused by a polyQ expansion in the huntingtin protein, which results in abnormal motor movements, personality changes, cognitive decline, and early death. Many studies have indicated that BDNF is a major target of mutant huntingtin protein. Decreased BDNF levels in the striatum have been detected in human HD subjects and mouse models of HD. A transgenic animal model in which BDNF has been specifically reduced in the cortex resulted in early dendritic changes, later loss of striatal medium spinal neurons, and early onset of clasping behavior. Moreover, gene expression profiling indicates that the depletion of BDNF in the cortex most closely resembles early grade human HD. These results suggest that striatal-specific atrophy in HD may be a consequence of a decrease in cortical BDNF by mutant huntingtin.

Correlates to Psychiatric Disorders

The profound effects of neurotrophic factors on synaptic connections, synaptic plasticity, and neurotransmission have formed the basis for their association with psychiatric disorders, such as anxiety disorders, depression, and substance abuse. In these conditions, the response to acute and chronic environmental changes leads to alterations in neuronal function. These additional functions have made neurotrophins attractive molecular intermediates that may be involved in the pathophysiology of psychiatric disorders in which environmental inputs can presumably lead to alterations in neuronal circuitry and ultimately behavior. In particular, it has become clear that neurotrophins can produce long-term changes by regulating transcriptional programs on the functioning of adult neurons. This could explain the long delay in therapeutic action of many psychiatric treatments. Again the clinical correlation is based on the assumption that there is a deficit in access or responsiveness to neurotrophic factors contributing to the phenotype of the disease state.

Major Depressive Disorder

The strongest evidence for a role of neurotrophins has come from the pathophysiology of depression, especially those associated with stress. For

depression, it is believed that there is a fundamental dysregulation of synaptic plasticity and neuronal survival in regions of the brain such as the hippocampus. There are several lines of evidence suggesting a role of neurotrophins in depression. First, in animal models, restraint stress leads to decreased expression of BDNF in the hippocampus. In addition, chronic physical or psychosocial stress leads to atrophy and death of hippocampal neurons especially in the CA3 region in rodents and primates. Also, MRI studies have shown that patients with depressive or PTSDs exhibit a small decrease in hippocampal volume. It is unclear though whether the atrophy and/or death of these neurons is directly related to the decreased availability of BDNF. In addition, not all forms of depression are associated with stress. However, if structural remodeling and synaptic plasticity are involved in the cellular pathophysiology of depression, then BDNF is an attractive candidate molecule to mediate these alterations.

Exogenously administered BDNF in the hippocampus had antidepressant effects in two animal models of depression (i.e., the forced swim and learned helplessness paradigms) comparable to those of chronic treatment with pharmacological antidepressants. In addition, BDNF has also been shown to have trophic effects on serotonergic and noradrenergic neurons in vitro and in vivo. Mutant mice with decreased levels of BDNF have been shown to have a selective decrement in serotonergic neuron function and corresponding behavioral dysfunction consistent with serotonergic abnormalities.

Also, serotonin and norepinephrine reuptake inhibitor antidepressants upregulate CREB, a cAMP-dependent transcription factor, and BDNF in a time course that corresponds to therapeutic action (10 to 20 days). The CREB transcription factor is involved in the induction of BDNF gene expression in neurons. This effect on the cAMP pathway provides a link between monoamine antidepressants and neurotrophin actions. These antidepressant treatments also lead to increases in expression of TrkB receptors in the hippocampus in a time course that also parallels the long time course of therapeutic action of these treatments. The effect of prolonged serotonin and norepinephrine reuptake inhibitor treatment involves enhancing neurotrophin signaling. Two other antidepressant treatments, MAOIs and ECT, also upregulate BDNF transcription. In rodents, long-term ECT has been shown to elicit the sprouting of hippocampal neurons that was attenuated in mutant mice that express lower levels of BDNF. Finally, recent studies have demonstrated that ketamine, a noncompetitive N-methyl-D-aspartate (NMDA) receptor antagonist, could mediate a rapid antidepressant response through a BDNF-dependent mechanism. In particular, after NMDA receptor blockade, there is a desuppression of protein translation resulting in an upregulation of BDNF that triggers enhanced synaptic plasticity.

Conversely, exogenously administered BDNF in the mesolimbic

dopamine system appears to have an opposite effect—increasing depression-like behavior. In addition, removal of BDNF in this dopamine circuit appears to have antidepressant effects on a social defeat paradigm. These findings emphasize the complexity of BDNF's role in mediating aspects of behavior related to depression. Together, these studies provide a framework to examine further the neurotrophin system as a potential therapeutic target for the treatment of depression.

NEUROTROPHINS AND GENETICS

Until recently, few genetic associations have been found between any neurotrophin and a human neurological or psychiatric disorder. A recent series of studies has linked one polymorphism in the BDNF gene with depression, bipolar disorder, and schizophrenia. This polymorphism identified from a single nucleotide polymorphism (SNP) screen leads to a single amino acid change from valine (Val) to methionine (Met) at position 66 in the pro region of the BDNF protein. This region is believed to be important in proper folding and intracellular sorting of the BDNF. Interestingly, proforms of neurotrophins have recently been shown to act as selective ligands for the $p75^{NTR}$. The mechanisms that contribute to altered $BDNF_{Met}$ function have been studied in neuronal culture systems. The distribution of $BDNF_{Met}$ to neuronal dendrites and its activity-dependent secretion are decreased. These trafficking abnormalities are likely to reflect impaired binding of $BDNF_{Met}$ to a sorting protein, sortilin, which interacts with BDNF in the prodomain region that encompasses the Met substitution.

This polymorphism is common in human populations with an allele frequency of 20 to 30 percent in Caucasian populations. This alteration in a neurotrophin gene correlates with reproducible alterations in human carriers. Humans heterozygous for the Met allele have smaller hippocampal volumes and perform poorly on hippocampal-dependent memory tasks. Using batteries of neuropsychological tests, carriers of the Met allele performed worse on tasks that involved recalling places and events but did not differ from Val/Val individuals on tasks that have been classically shown to be less hippocampal-dependent, such as word learning and planning tasks.

However, genetic association studies for psychiatric disorders have presented a more complex picture. In patients with bipolar disorder, the Val allele appears to confer greater risk for the disease, while in patients with schizophrenia, depression, and anxiety disorders, there is little consensus as to whether the allele confers altered susceptibility, though recent studies suggest that the Met allele may have impact on treatment response in PTSD. Inconsistency across genetic studies may be attributable to sampling and measurement issues, genetic heterogeneity due to

differential sampling of populations, or a low frequency of homozygous Met carriers, which may lessen the effect size of any particular association. It may also relate to a failure to take into account relevant gene-by-gene and gene-by-environment interactions. This point is highlighted by a recent study of BDNF "knock-in" in mice ($BDNF^{Met/Met}$). The knock-in mice reproduced the phenotypic hallmarks related to hippocampal function that are seen in humans with this BDNF SNP. Subsequent analyses of these mice elucidated phenotypes that had not been initially identified in human carriers: increased anxiety and decreased capacity for fear extinction learning. In addition, when stressed, $BDNF^{Met/Met}$ mice display increased anxiety-related behaviors, suggesting that environmental factors are likely required to elicit symptoms related to psychiatric disorders.

THERAPEUTIC POTENTIAL OF NEUROTROPHINS

The recent clinical trials have provided limits in designing therapeutic strategies to use neurotrophic factors for neurodegenerative and psychiatric disorders. First, it has become clear that the physical delivery of sufficient quantities to target neurons is a major obstacle. Development of small molecules that readily cross the BBB to activate neurotrophin receptors or potentiate the actions of neurotrophins is an approach that is in its infancy.

Second, because neurotrophins have multiple effects on neuronal activity, indiscriminate "flooding" of the CNS with neurotrophic factors will likely lead to untoward side effects such as epileptic activity. In addition, it had been noted in the clinical trials with BDNF that downregulation of the TrkB receptors after unregulated application of BDNF may have also contributed to the minimal therapeutic effects. New strategies are being studied that include more local and regulated application of neurotrophins through stereotactic injection of viral vectors or engineered progenitor cells. In particular, this approach is currently being applied to diseases such as Alzheimer disease where there is a defined neuronal population such as basal forebrain cholinergic neurons that undergoes degeneration and is dependent on one neurotrophin such as NGF.

The activation of the neurotrophin system through other receptor signaling systems offers an alternative strategy. For example, antidepressant agents acting via monoamine GPCRs can lead to increased expression of both neurotrophins and neurotrophin receptors. Importantly, only the neurons that express the monoamine GPCRs will have enhanced production of the neurotrophin or Trk receptor. Recently, it has also been shown that other GPCRs, the purine adenosine 2A receptor, and pituitary adenylate-cyclase–activating peptide (PACAP) neuropeptide receptor can transactivate Trk neurotrophin receptors in the absence of neurotrophins in hippocampal neurons in vitro. Therefore, small molecules can activate Trk receptors in the absence of neurotrophins. These results raise the

possibility that small molecules may be used to elicit neurotrophic effects for the treatment of neurodegenerative diseases by selective targeting of neurons that express specific GPCRs and Trk receptors.

Other potent ways of increasing the level of neurotrophins involve using specific classes of antidepressant drugs. As many antidepressants are clinically effective only after several weeks of administration, newer approaches have examined the effects of ketamine, a fast acting, noncompetitive NMDA antagonist, in the treatment of MDD. The antidepressant effects are believed to be due, in part, to the actions of BDNF. In a similar vein, physical exercise robustly increases the expression of brain BDNF, which is associated with improvements in cognition function, neurogenesis and anxiolytic effects.

It should be emphasized that all these strategies to utilize neurotrophic factors are based on an assumption of symptomatic treatment of impaired neurons. This impairment implies not only cell survival but also proper physiological functioning and connectivity of these neurons. With greater understanding of the signal transduction pathways that are activated by neurotrophins, alternate strategies can be devised to manipulate these pathways through new drug development. A completely different approach is to block neurotrophins, such as NGF, which elicits hyperalgesia after tissue damage or inflammation. Development of neutralizing antibodies against NGF and their administration to reduce pain thresholds has proven to be efficacious in many clinical pain conditions. Therefore, further understanding of the core pathophysiological mechanism for neurodegenerative and psychiatric disorders will benefit the development of rational therapies that involve engaging the neurotrophin system.

References

Berton O, McClung CA, Dileone RJ, et al. Essential role of BDNF in the mesolimbic dopamine pathway in social defeat stress. *Science*. 2006;311:864–868.

Bespalov MM, Saarma M. GDNF family receptor complexes are emerging drug targets. *Trends Pharmacol Sci*. 2007;28:68–74.

Cabelli RJ, Hohn A, Shatz CJ. Inhibition of ocular dominance column formation by infusion of NT4/5 or BDNF. *Science*. 1995;267:1662–1666.

Casey BJ, Glatt CE, Lee FS. Treating the developing versus developed brain: translating preclinical mouse and human studies. *Neuron*. 2015;86:1358–1368.

Chao MV. Neurotrophins and their receptors: a convergence point for many signaling pathways. *Nat Rev Neurosci*. 2003;4:299–309.

Chao MV, Hempstead BL. p75 and Trk: a two-receptor system. *Trends Neurosci*. 1995;18:321–326.

*Chen ZY, Jing DQ, Bath KG, et al. Genetic variant BDNF (Val66Met) polymorphism alters anxiety-related behavior. *Science*. 2006;314:140–143.

*Duman RS, Heninger GR, Nestler EJ. A molecular and cellular theory of depression. *Arch Gen*

Psychiatry. 1997;54:597–606.

Felmingham KL, Dobson-Stone C, Schofield PR, Quirk GJ, Bryant RA. The brain-derived neurotrophic factor Val66Met polymorphism predicts response to exposure therapy in posttraumatic stress disorder. *Biol Psychiatry*. 2013;73(11):1059–1063.

*Hempstead BL. The many faces of $p75^{NTR}$. *Curr Opin Neurobiol*. 2002;12:260–267.

Ibanez CF, Andressoo JO. Biology of GDNF and its receptors - relevance for disorders of the central nervous system. *Neurobiol Dis*. 2016. (epub).

Kaplan DR, Miller FD. Neurotrophin signal transduction in the nervous system. *Curr Opin Neurobiol*. 2000;10:381–391.

Lee FS, Chao MV. Activation of Trk neurotrophin receptors in the absence of neurotrophins. *Proc Natl Acad Sci U S A*. 2001;98:3555–3560.

Lee FS, Kim AH, Khursigara G, Chao MV. The uniqueness of being a neurotrophin receptor. *Curr Opin Neurobiol*. 2001;11:281–286.

Lee R, Kermani P, Teng KK, Hempstead BL. Regulation of cell survival by secreted proneurotrophins. *Science*. 2001;294:1945–1948.

Levi-Montalcini R. The nerve growth factor: thirty-five years later. *Science*. 1987;237:1154–1162.

*Lyons WE, Mamounas LA, Ricaurte GA, et al. Brain-derived neurotrophic factor-deficient mice develop aggressiveness and hyperphagia in conjunction with brain serotonergic abnormalities. *Proc Natl Acad Sci U S A*. 1999;96:15239–15244.

Mariga A, Mitre M, Chao MV. Consequences of brain-derived neurotrophic factor withdrawal in CNS neurons and implications in disease. *Neurobiol Dis*. 2016.

Minichiello L, Calella AM, Medina DL, Bonhoeffer T, Klein R. Mechanism of TrkB-mediated hippocampal long-term potentiation. *Neuron*. 2002;36:121–137.

*Monteggia LM, Zarate C. Antidepressant actions of ketamine: from molecular mechanisms to clinical practice. *Curr Opin Neurobiol*. 2015;30:139–143.

Notaras M, Hill R, van den Buuse M. The BDNF gene Val66Met polymorphism as a modifier of psychiatric disorder susceptibility: progress and controversy. *Mol Psychiatry*. 2015;20:916–930.

Park H, Poo MM. Neurotrophin regulation of neural circuit development and function. *Nat Rev Neurosci*. 2013;14:7–23.

Petryshen TL, Sabeti PC, Aldinger KA, et al. Population genetic study of the brain-derived neurotrophic factor (BDNF) gene. *Mol Psychiatry*. 2010;15(8):810–815.

Riccio A, Ahn S, Davenport CM, Blendy JA, Ginty DD. Mediation by a CREB family transcription factor of NGF-dependent survival of sympathetic neurons. *Science*. 1999;286:2358–2361.

Rios M. BDNF and the central control of feeding: accidental bystander or essential player? *Trends Neurosci*. 2013;36(2):83–90.

Shirayama Y, Chen AC, Nakagawa S, Russell DS, Duman RS. Brain-derived neurotrophic factor produces antidepressant effects in behavioral models of depression. *J Neurosci*. 2002;22:3251–3261.

Snider WD. Functions of the neurotrophins during nervous system development: what the knockouts are teaching us. *Cell*. 1994;77:627–638.

Strand AD, Baquet ZC, Aragaki AK, et al. Expression of profiling of Huntington's disease models suggests BDNF depletion plays a major role in striatal degeneration. *J Neurosci*.

2007;27:11758–11768.

Thoenen H, Sendtner M. Neurotrophins: From enthusiastic expectations through sobering experiences to rational therapeutic approaches. *Nat Neurosci.* 2002;5:1046–1050.

Yamashita N, Kuruvilla R. Neurotrophin signaling endosomes: biogenesis, regulation, and functions. *Curr Opin Neurobiol.* 2016;39:139–145.

▲ 1.8 Novel Neurotransmitters

THOMAS W. SEDLAK, M.D., PH.D., AND ADAM I. KAPLIN, M.D., PH.D.

Neurotransmitters are chemicals that amplify or inhibit the depolarization signal from one neuron to that of an adjacent neuron. A neurotransmitter is typically released from a *presynaptic neuron* and travels across a small space, the *synaptic cleft* or *synapse,* to act upon the *postsynaptic neuron*. Neurotransmitters may be excitatory (e.g., glutamate) in that they promote depolarization of the target cell, or inhibitory by promoting hyperpolarization (e.g., most GABA). Neurotransmitter receptors that are themselves ion channels may elicit depolarization via the neurons they conduct, typically depolarization via sodium or calcium, and hyperpolarization via chloride. In addition to ion channel receptors are metabotropic receptors that produce second messengers that may have complex effects upon target cells, such as altering gene expression or modulating the propensity to depolarize.

An *action potential* is a wave of depolarization that travels down a neuronal axon to the *presynaptic terminal,* a specialized appendage where neurotransmitters are stored in specialized *vesicles*. The action potential opens voltage-sensitive calcium channels in the membrane, allowing for an increase in cellular calcium that results in vesicles releasing their contents into the synaptic cleft and acting upon receptors on the postsynaptic neuron membrane. When binding to their targets, neurotransmitters are typically full or partial agonists, and therapeutic drugs range from partial to full agonists, to antagonists or inverse agonists (Fig. 1.8–1).

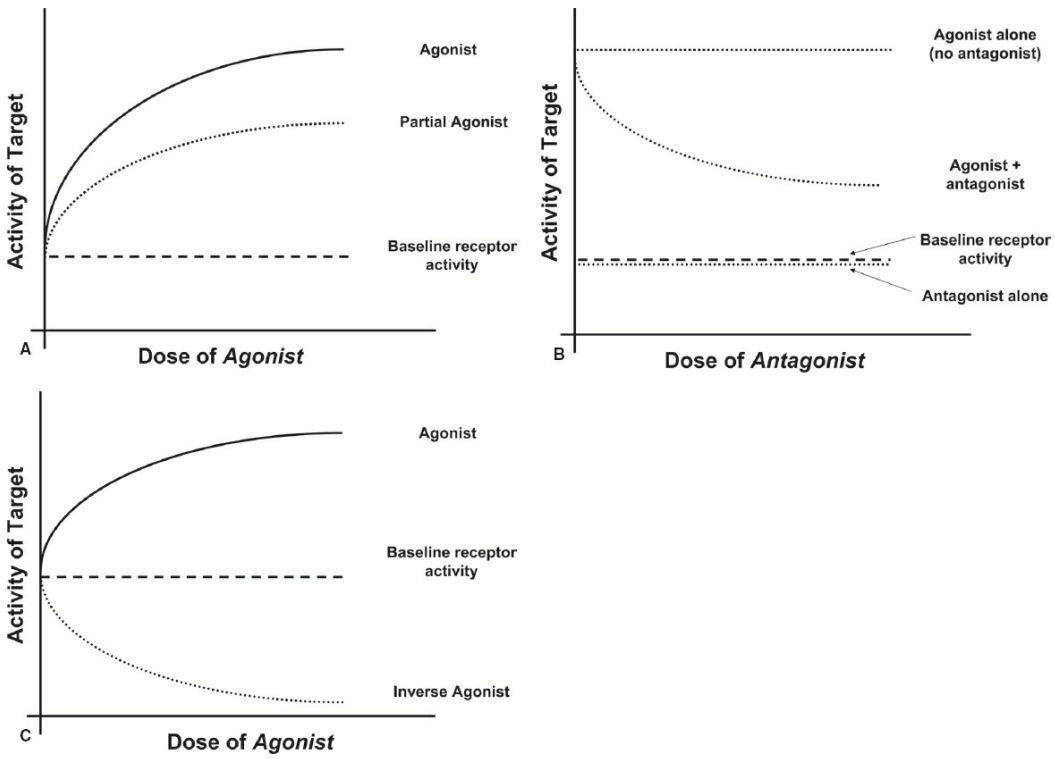

FIGURE 1.8–1. Agonists, antagonists, partial agonists, and inverse agonists. **A:** *Agonist* drugs bind to a target, such as a neurotransmitter receptor, and increase its activity until the receptors saturate. A *partial agonist* increases the activity of its target, but to a level below that of its maximum. In some cases, a partial agonist leads to diminished overall activity of a neurotransmitter receptor because it competes with the full agonist neurotransmitter. Occasionally a drug considered to be an agonist becomes reclassified as a partial agonist once a stronger agonist is found. THC was once considered a full agonist for the CB1 receptor, but was found to be a partial agonist after the discovery of the more potent synthetic cannabinoids, CP55,940 and WIN55,212–2. **B:** An *antagonist* has no intrinsic activity for activating or inhibiting a receptor. It acts to inhibit the activity of an agonist, in many cases by competing with an agonist for the binding site. **C:** An *inverse agonist* inhibits the activity of its target to a level below that of its baseline level with no drug present. Rimonabant is a CB1 receptor inverse agonist. It can block the baseline activity of the receptor even in the absence of cannabinoid agonists.

The definition of what a neurotransmitter is and is not has changed over the decades. The neurotransmitters initially discovered were small chemicals, first acetylcholine and later biogenic amines such as serotonin, dopamine, norepinephrine, epinephrine, and histamine. Later it was found that amino acids and peptides could act as neurotransmitters, such as the case of enkephalin being the transmitter acting upon the opiate receptor. By the 1990s, it became apparent that the neurotransmitter acting on cannabinoid receptors was derived from cellular lipids. Furthermore, even a gas, nitric oxide, could be a neurotransmitter, bypassing the requirement for postsynaptic receptors and acting directly within postsynaptic neurons.

GASES AS NEUROTRANSMITTERS

Nitric Oxide

The discovery that gases could function as neurotransmitters revealed that highly atypical modes of signaling existed between neurons. In the early 1990s, nitric oxide was the first gas to be ascribed a neurotransmitter function and proved to be an atypical neurotransmitter for several reasons. First, it was not stored in or released from synaptic vesicles; as it was a small gas, it could freely diffuse across cell membranes and into the target neuron. Second, its target was not a specific receptor on the surface of a target neuron, but intracellular proteins whose activity could directly be modulated. Nitric oxide also lacks a reuptake mechanism to remove it from the synapse. Although its enzymatic inactivation is postulated to exist, nitric oxide is also limited by a short half life of a few seconds.

Nitric oxide was initially discovered as a bactericidal compound released from macrophages and as an endothelial-cell-derived relaxation factor allowing for the dilation of blood vessels. A role for nitric oxide in the brain followed, revealing a role for the gas in neurotransmission, learning and memory processes, neurogenesis, and neurodegenerative diseases.

Synthesis of Nitric Oxide . Nitric oxide is chemically designated NO., with the dot representing that the molecule is a free radical, also imparting a highly reactive nature. Nitric oxide is occasionally confused with nitrous oxide (N_2O), the gaseous anesthetic, and nitrogen dioxide (NO_2), a pollutant found in exhaust fumes, although these are not synthesized endogenously in mammals. A specific enzyme exists to generate nitric oxide within cells, nitric oxide synthase (NOS) (Fig. 1.8–2). This enzyme generates nitric oxide by abstracting nitrogen from the amino acid arginine and reacting it with an oxygen atom. The enzyme utilizes nicotinamide adenine dinucleotide phosphate (NADPH) and generates citrulline as a byproduct.

Three distinct enzymatic forms of NOS exist, each with differing locations and activation patterns within the body. *Neuronal nitric oxide synthase* (nNOS) was the first form discovered, by Bredt and Snyder, and is the predominant form in the brain. Neuronal NOS is expressed only in neurons, especially those of the cortex, dentate gyrus of the hippocampus, corpus striatum, and cerebellum. Although nNOS-containing neurons are only 1 percent of cortical neurons, their neuronal processes are so extensively distributed that almost all neurons make contact with an nNOS-containing nerve terminus. nNOS enzyme activity is markedly augmented by calcium levels via the accessory protein calmodulin. Thus, nitric oxide may be synthesized following neuronal depolarization, in which calcium levels increase transiently.

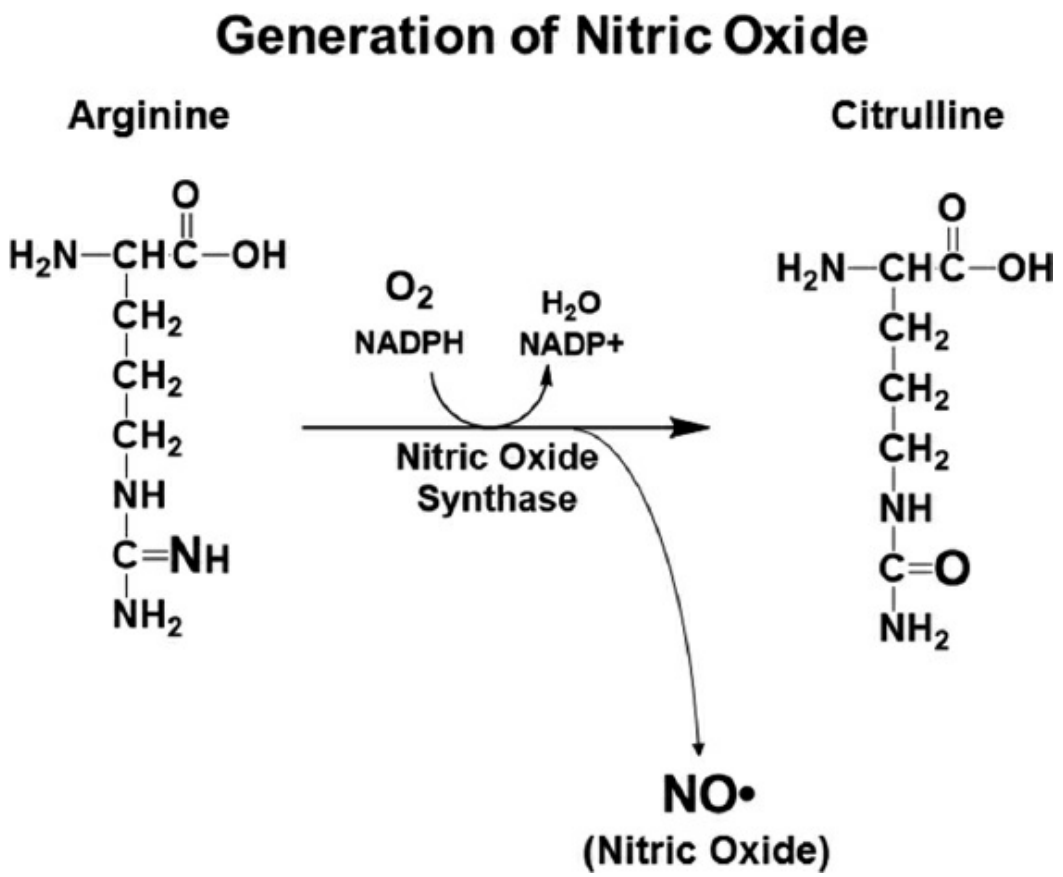

FIGURE 1.8–2. Enzymatic generation of nitric oxide. The gaseous neurotransmitter, nitric oxide, is generated by the enzyme nitric oxide synthase. The amino acid arginine is converted to citrulline and nitric oxide, employing oxygen and the reductant NADPH. Of note, nitric oxide is not preformed and stored in synaptic vesicles, but synthesized on demand.

Endothelial NOS (eNOS) is predominantly found in blood vessels, where it plays a crucial role in allowing for the relaxation and dilation of blood vessels. Nitroglycerine and sodium nitroprusside exert their vasodilatory effects via conversion to nitric oxide. eNOS activity is augmented by phosphorylation and increases in intracellular calcium. The third NOS, *inducible NOS* (iNOS), exists in many tissues in miniscule amounts. However, its levels are strongly increased by a great variety of cell stressors, especially inflammation. In the brain, it is largely induced in glial cells, but also neurons.

Mechanism of Action of Nitric Oxide: Cyclic GMP Pathway. Long-term changes in brain function, such as learning and memory, involve a great variety of cellular processes, including changes in the patterns of gene and protein expression, and physical remodeling of neuronal architecture, such as dendrites. *Signal transduction* is the process by which extracellular signals, such as neuronal depolarization and receptor activation, lead to modified cellular function. Cyclic guanosine monophosphate (cGMP) is a prototypic intracellular messenger whose synthesis by guanylyl cyclase is

stimulated after a membrane receptor is activated. cGMP then activates protein kinases that phosphorylate proteins as part of the cascade that elicits the biological effect. As is the case with other neurotransmitters, such as serotonin, nitric oxide also activates cGMP production in neurons (Fig. 1.8–3). While a typical neurotransmitter activates guanylyl cyclase via a G protein coupled to a membrane receptor, nitric oxide directly activates soluble guanylyl cyclase, the enzymatic form found in cytoplasm. The active site of guanylyl cyclase contains a heme group cofactor whose iron atom is bound by nitric oxide, leading to a protein conformation change and production of cGMP. Nitric oxide can also interact with the heme groups of other proteins including hemoglobin/myoglobin, ferritin, and cytochrome P450.

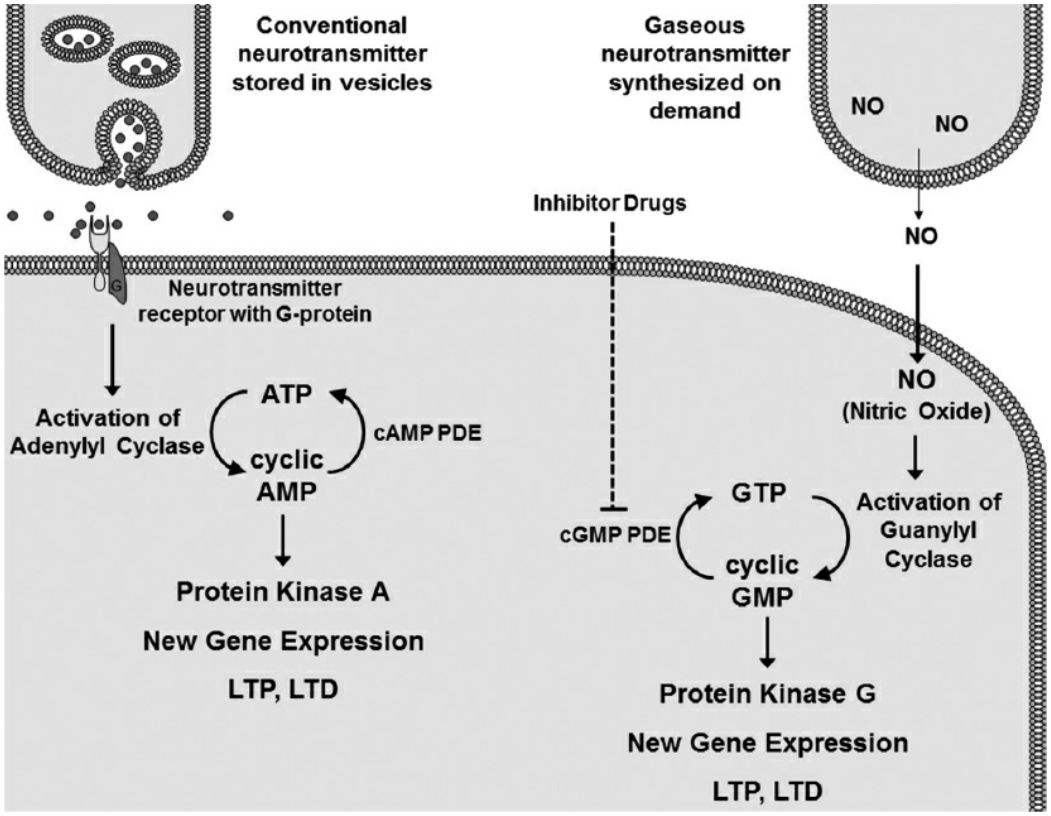

FIGURE 1.8–3. Neurotransmitter and signaling functions of nitric oxide (NO) via production of cyclic GMP. Gaseous nitric oxide is generated enzymatically and freely diffuses into an adjacent neuron (*upper right*). In comparison to traditional neurotransmitters (*upper left*), NO does not act via a specific neurotransmitter receptor on the surface membrane of a neuron. In contrast, NO freely diffuses across the neuronal membrane and activates the enzyme, guanylyl cyclase, which converts GTP into the second messenger, cyclic GMP. Nitric oxide effects are mediated, in part, by cyclic GMP activation of neuronal protein kinases, new gene expression and effects on neuronal long term potentiation (LTP) and long-term depression (LTD). ATP, adenosine triphosphate; PDE, phosphodiesterase.

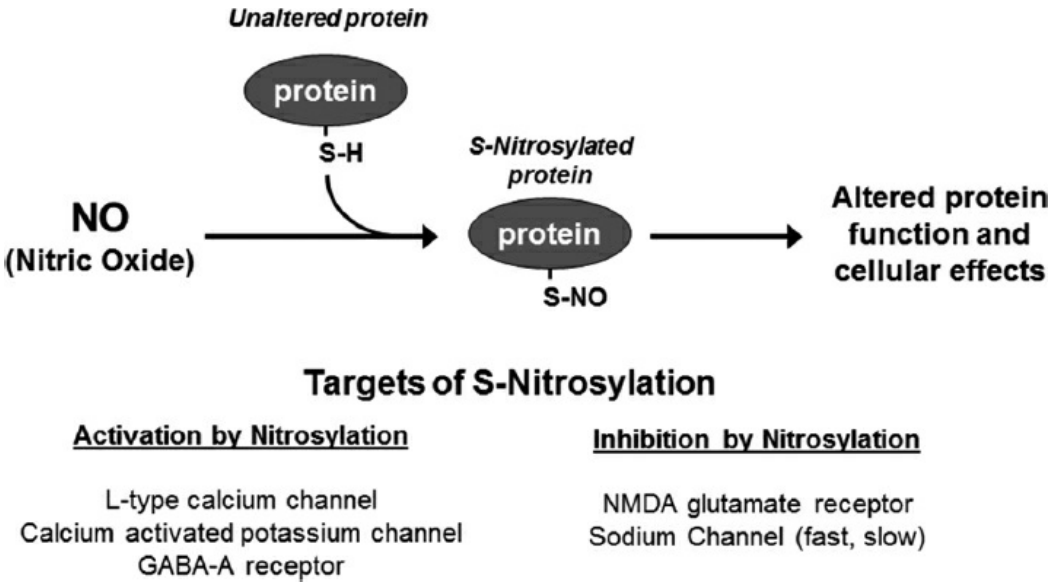

FIGURE 1.8–4. Nitric oxide (NO) signaling via *S*-nitrosylation. In addition to NO activation of guanylyl cyclase, NO may also directly alter protein function via the process of *S*-nitrosylation. In this process, which does not require enzymatic catalysis, NO reacts with -SH groups of protein cysteine residues, resulting in an –SNO modification and altered protein function. Some proteins demonstrate robust activation following *S*-nitrosylation, whereas others are inhibited by the process.

Mechanism of Action of Nitric Oxide: S-Nitrosylation Pathway. A second method by which nitric oxide exerts it effects upon cells is by the process of *S*-nitrosylation. In this signaling mechanism, nitric oxide directly modifies the sulfur atom of a protein cysteine residue, forming an *S*-nitrosothiol group (Fig. 1.8–4) that modifies the activity of select proteins. This generation of this modification requires no enzyme, though there is evidence that some proteins may carry the modification around to different areas of the cell and transfer it to a destination protein that elicits the target effect. Proteins that have been nitrosylated vary in their response to this modification; some are activated and others inactivated.

The number of protein targets of *S*-nitrosylation is rapidly expanding and includes molecules involved in signal transduction, programmed cell death, transcription factors, cytoskeletal proteins, ion pumps, and ion channels. In many cases, modification of a single cysteine residue in a target protein is sufficient for nitric oxide to regulate its activity. Specific targets that are *activated* by *S*-nitrosylation include L-type calcium channels, calcium-activated potassium channels, and GABA-A receptors. Proteins *inhibited* by nitrosylation include several types of sodium channels, the *N*-methyl-D-aspartate (NMDA) subtype of the glutamate neurotransmitter receptor, and several metabolic enzymes. *S*-nitrosylation as a means of signal transduction is somewhat analogous to protein phosphorylation, as both are reversible covalent modifications that regulate protein function to change cell activity. *S*-nitrosylation might participate in memory, learning, and behavior as many brain proteins are nitrosylated

through the activity of neuronal nitric oxide generation.

Nitric Oxide and Neurotransmission. LTP is the process by which repetitive stimulation of a presynaptic neuron leads to stronger firing of a postsynaptic neuron, a process that underlies changes in learning and behavior. Induction of LTP depends on activation of postsynaptic NMDA receptors, while LTP maintenance relies on presynaptic mechanisms. Neurotransmission through the NMDA receptor facilitates LTP, in part, through the activity of nitric oxide. Activation of the NMDA receptor leads to a cellular calcium increase, promoting nitric oxide synthesis and cGMP formation in the postsynaptic cell (Fig. 1.8–5).

Pharmacological inhibitors of NOS have revealed deficits in LTP in rodent, bird, and honeybee models, and nitric oxide has been implicated in both short-term and long-term memory acquisition. Genetically modified mice lacking either nNOS or eNOS demonstrate no changes in LTP in the hippocampus; however, mice deficient in *both* nNOS and eNOS show decreased LTP in the CA1 region of the hippocampus. One form of NOS may compensate for the absence of the other, or the two may function cooperatively in LTP. Studies of the enteric nervous system have also revealed roles for nitric oxide in relaxation of the pyloric sphincter, and mice deficient in nNOS reveal a marked hypertrophy of the pylorus. Nitric oxide may also regulate monoaminergic neurotransmission. Inhibition of NOS in rats enhances the effects of cocaine and amphetamine, while the reverse is observed by increasing nitric oxide.

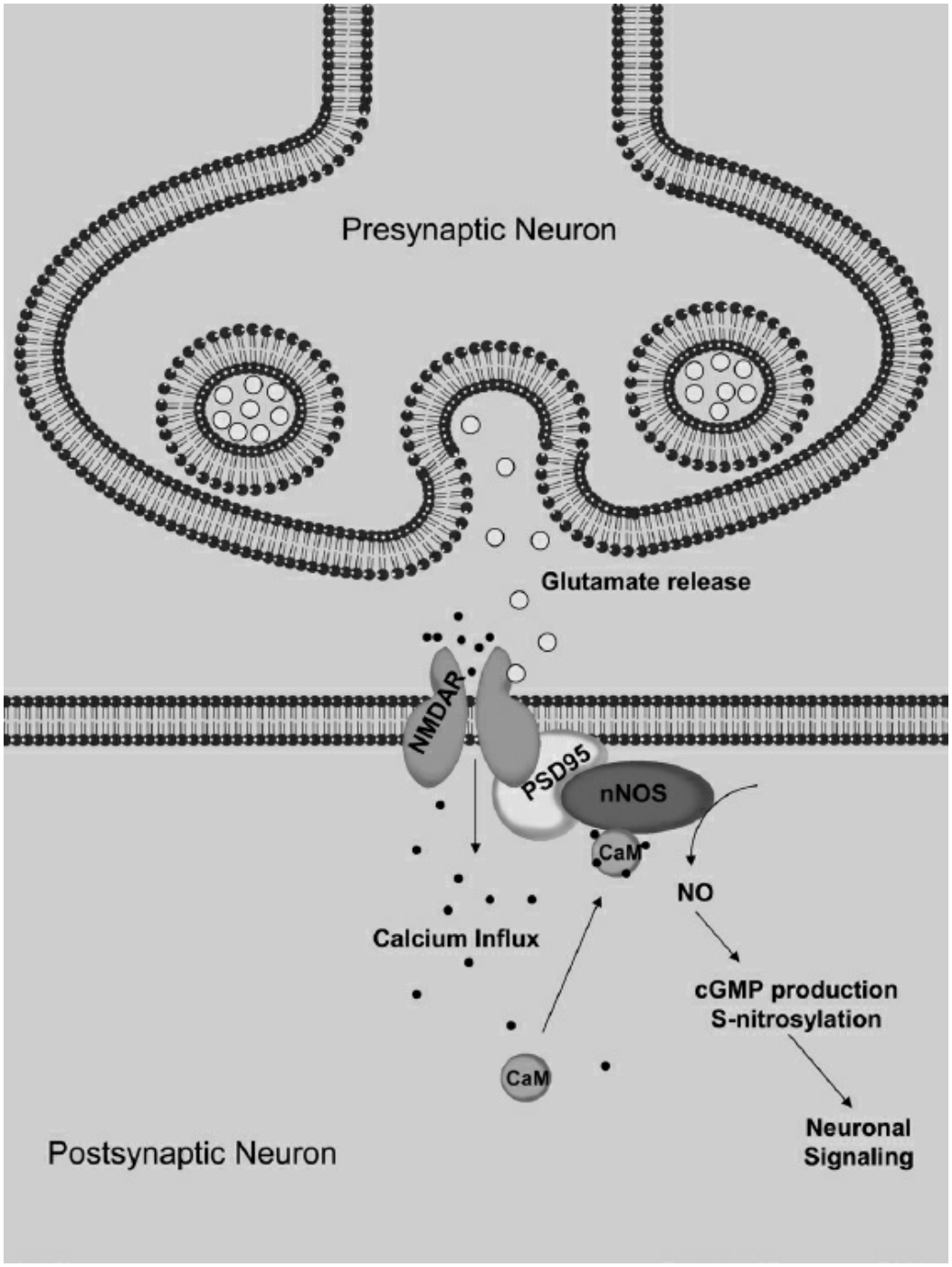

FIGURE 1.8–5. Nitric oxide (NO) generation following NMDA receptor activation. The presynaptic neuron (top) releases glutamate (not shown), activating the NMDA glutamate receptor and allowing for calcium entry into the postsynaptic neuron. Calcium binds to the protein calmodulin (CaM), which in turn activates neuronal nitric oxide synthase (nNOS) to synthesize NO. A freely diffusible gas, NO exerts effects upon the target neuron via formation of cyclic GMP and *S*-nitrosylation (Figs. 1.8–3 and 1.8–4).

Nitric Oxide and Behavior. Nitric oxide neurotransmission may play a role in behavior as nNOS-deficient male mice display exaggerated aggressive tendencies and increased sexual activity. In female mice, the contrary is true, as they have reduced aggression. As manic bipolar patients may show both hypersexuality and aggression, the nitric oxide pathway

may participate in the psychopathology of affective states.

In the periphery, nNOS localizes to neurons that innervate blood vessels of the penis, including the corpus cavernosa. Stimulation of these nerves releases nitric oxide, leading to cGMP formation, blood vessel wall relaxation and vasodilatation, penile engorgement, and initial erection. The sustained phase of erection also depends on nitric oxide; turbulent blood flow leads to phosphorylation of eNOS and sustained nitric oxide production. Drugs used in treatment of erectile dysfunction, sildenafil (Viagra), tadalafil (Cialis), and vardenafil (Levitra), act to inhibit type 5 phosphodiesterase, an enzyme that degrades cGMP in the penis (Fig. 1.8–3), thereby potentiating nitric oxide neurotransmission and penile erection.

Numerous lines of evidence have suggested a role for nitric oxide in the regulation of sleep-wake cycles. nNOS-expressing neurons occur in several areas that initiate rapid eye movement (REM) sleep, including the pons, dorsal raphe nucleus, laterodorsal tegmentum, and pedunculopontine tegmentum. In animal models, microinjection of compounds that release nitric oxide result in decreased wakefulness and increased slow-wave sleep (SWS). Consistent with this, NOS inhibitors show a trend toward decreasing slow wave and REM sleep. Studies of NOS-deficient mice suggest that nitric oxide may serve a more complex role than merely promoting sleep. nNOS-deficient animals also show reduced REM sleep; however, iNOS-deficient mice demonstrate the reverse, suggesting a complex interplay between NOS enzymatic isoforms.

Nitric Oxide and Mood Disorders. NOS-expressing neurons are well-represented in brain areas implicated in depression, including the dorsal raphe nucleus and prefrontal cortex. A role for nitric oxide has been suggested in antidepressant response as SSRI antidepressants can directly inhibit NOS activity. Moreover, in animal studies such as the forced swim test, NOS, and soluble guanylyl cyclase inhibitors can achieve antidepressant-like effects. Plasma nitric oxide levels were elevated in patients with bipolar disorder compared to healthy control subjects. However, in depressed subjects, studies have found decreased nitric oxide levels and increased plasma nitrite, a byproduct of nitric oxide. Reduced NOS has also been described in the PVN of patients with schizophrenia and depression compared to controls.

Neurogenesis, the process by which new neurons are generated in the adult brain, is increasingly appreciated to participate in both mood disorder pathophysiology and antidepressant response. Increased hippocampal neurogenesis is associated with antidepressant response, and smaller hippocampal volume may be a risk factor for mood and anxiety disorders. Serotonin itself appears to promote neurogenesis in the hippocampus, while nitric oxide has been found to inhibit neurogenesis. Pharmacologic inhibitors of NOS result in increased serotonin and

neurogenesis in the dentate gyrus of the hippocampus, a paramount site of this process. These NOS inhibitors also lead to an increase in serotonin in the dentate gyrus. Unsurprisingly, nNOS-deficient animals also manifest increased proliferation in the dentate gyrus. As steroids appear to induce NOS expression, nitric oxide may contribute to effects upon mood and anxiety often observed in those treated with these agents.

Nitric oxide has been interrogated for its ability to regulate neurotransmission at serotonin, norepinephrine, and dopamine nerve termini. Nitric oxide appears to possess the capability of either increasing or decreasing activity at these neurons depending upon the specific brain region and timing of its activation.

Nitric Oxide and Schizophrenia. Nitric oxide has been investigated as a candidate molecule contributing to symptoms of schizophrenia. Two genetic studies have identified schizophrenia-associated SNPs in CAPON, a protein that associates with nNOS. SNPs in nNOS itself have been associated with schizophrenia, although others have not been able to reproduce such findings. Changes in NOS levels have been reported in postmortem brain samples of individuals with schizophrenia. Abnormalities have been noted in the cortex, cerebellum, hypothalamus, and brainstem, although no specific trend can be discerned. Elevated NOS activity has been noted in platelets from drug-naïve and drug-treated individuals with schizophrenia. Some investigators find increased nitric oxide activity and others the reverse. In autopsy samples, patients with schizophrenia were found to have abnormally localized NOS-expressing neurons in the prefrontal cortex, hippocampus, and lateral temporal lobe, consistent with abnormal migration of these neuronal types during development. In a rat model, prenatal stress led to reduced NOS-expressing neurons in the fascia dentate and hippocampus.

Neuropathologic Roles of Nitric Oxide. Abundant evidence exists that nitric oxide is a direct participant in a variety of neuropathic events. Superoxide, a byproduct of cellular metabolism, can react with nitric oxide to form peroxynitrite (chemical formula, $ONOO^-$). This labile and toxic compound forms chemical adducts with protein tyrosine residues, a process termed *protein nitration,* and DNA, leading to cellular dysfunction. Cell loss resulting from ischemic stroke is mediated in part by overstimulation of the glutamate NMDA receptor, a process termed *excitotoxicity*. Nitric oxide produced by NMDA activation appears to mediate a significant portion of this excitotoxic neuronal death, and stroke damage is reduced in mice with a genetic deletion of nNOS.

S-Nitrosylation has also been implicated in pathologic processes in the brain. Mutations in the Parkin protein are associated with early-onset Parkinson disease. Parkin is an E3 ubiquitin ligase, adding ubiquitin

molecules to proteins and targeting them for destruction in the cell proteasome. In sporadic Parkinson disease, that is, without the early-onset mutation, nitric oxide can nitrosylate the Parkin protein and inhibit its protective E3 ubiquitin ligase function. An overabundance of nitric oxide signaling may thus predispose to the dysfunction and cell death of dopaminergic neurons in Parkinson disease by interfering with proteins essential for cell functioning. In Alzheimer disease, excess oxidation of brain protein, lipids, and carbohydrates has long been appreciated, but nitrosative stress from excess nitric oxide also appears to participate in the disease. Protein disulfide isomerase (PDI) is a cellular protective protein that may help combat the accumulation of misfolded proteins such as the amyloid fibrils occurring in the disease. In both Alzheimer and Parkinson disease brains, PDI appears to be *S*-nitrosylated in a harmful way that impedes its cellular protective function.

The discovery that nitric oxide participates in neurodegenerative processes raises the possibility for improved diagnostic processes, such as detecting damage to cellular components produced by nitric oxide prior to the onset of full-blown symptoms. In addition, drugs may be designed to attenuate the damage to crucial neuronal proteins that protect against disease onset. However, completely and nonspecifically inhibiting or stimulating nitric oxide synthesis is likely to produce significant side effects because of its wide-ranging activities throughout the body.

Nitric oxide is one of the most intensively studied compounds in the body, and possesses pleiotropic activities in different organs. While nitric oxide has physiologic roles in the brain, vasculature, and immune system, it is a complex and incompletely understood molecule that also participates in disease. However, oxygen, which is a highly reactive molecule like nitric oxide, also possesses the capacity to contribute to disease pathogenesis, such as in oxidative damage, while still being essential for life.

Carbon Monoxide

Although carbon monoxide (chemical formula, CO) is most well known as an air pollutant derived from combustion reactions, it is produced physiologically in a great variety of organisms ranging from human to bacterium. While once thought to be a toxic byproduct of metabolic reactions, carbon monoxide is increasingly recognized to play an important role in regulating a variety of physiological processes in the brain and other organs. These varied effects include regulation of olfactory neurotransmission, blood vessel relaxation, smooth muscle cell proliferation, and platelet aggregation.

Carbon monoxide is far more well known for its toxic effects than its activities at physiologic concentrations. It binds tightly to heme molecules within hemoglobin, forming carboxyhemoglobin, which can no longer transport oxygen to tissues. One to two packs per day smokers typically

have 3 to 8 percent of their hemoglobin as carboxyhemoglobin, while nonsmokers have less than 2 percent. Following acute carbon monoxide poisoning, 5 to 10 percent carboxyhemoglobin is associated with impaired alertness and cognition, and 30 to 50 percent carboxyhemoglobin leads to significant drops in oxygen transport to tissues.

Enzymatic Generation of Carbon Monoxide. Carbon monoxide is produced during the metabolism of heme by the action of *heme oxygenase* (HO). This enzyme utilizes oxygen, the reducing equivalents of NADPH, and the electron donor cytochrome P450 reductase to break open the carbon ring of heme and release a one-carbon fragment as carbon monoxide (Fig. 1.8–6). The reaction also produces the green pigment biliverdin and free iron. Biliverdin is converted to the yellow pigment bilirubin, which like carbon monoxide is no longer solely considered a toxic byproduct. At physiologic concentrations, bilirubin is an enormously potent antioxidant, which can be converted back to its precursor, biliverdin.

Three forms of HO exist. HO1 is similar to iNOS in that it typically exists at very low levels, but its expression may be potently induced by a great variety of stimuli, such as oxidative stress, inflammation, dopamine, steroids, and growth factors. Indeed, HO1 is one of the most easily induced proteins known. HO_2 is not an inducible protein and is predominantly expressed in the brain and testis. HO_2 is expressed in discrete neuronal populations throughout the brain, including cortical and hippocampal pyramidal cells, dentate gyrus granule cells, the olfactory bulb, thalamus, hypothalamus, brainstem, and cerebellum. HO3 is an isoform whose significance is poorly understood.

Molecular Actions of Carbon Monoxide. Similar to nitric oxide, gaseous carbon monoxide can freely diffuse across membranes and directly activate soluble guanylyl cyclase, though it is approximately 30-fold less potent than nitric oxide in doing so. With the ensuing increase in cGMP, protein kinases are activated, leading to some of the manifold effects of carbon monoxide upon cells. The expression of HO_2 closely mirrors the expression pattern of guanylyl cyclase, implicating the two as part of a common pathway of neuronal signaling, and inhibitors of HO_2 block the generation of cGMP. Similar to the case of nNOS, HO_2 can be activated by calcium/calmodulin and phosphorylation, fulfilling the important criterion that a neurotransmitter needs to be rapidly released in response to neuronal depolarization. Carbon monoxide can also activate p38 MAP kinase, though a poorly understood mechanism that does not require cGMP. This important kinase promotes a variety of cellular effects including inhibition of inflammation, cell proliferation, and programmed

cell death (apoptosis).

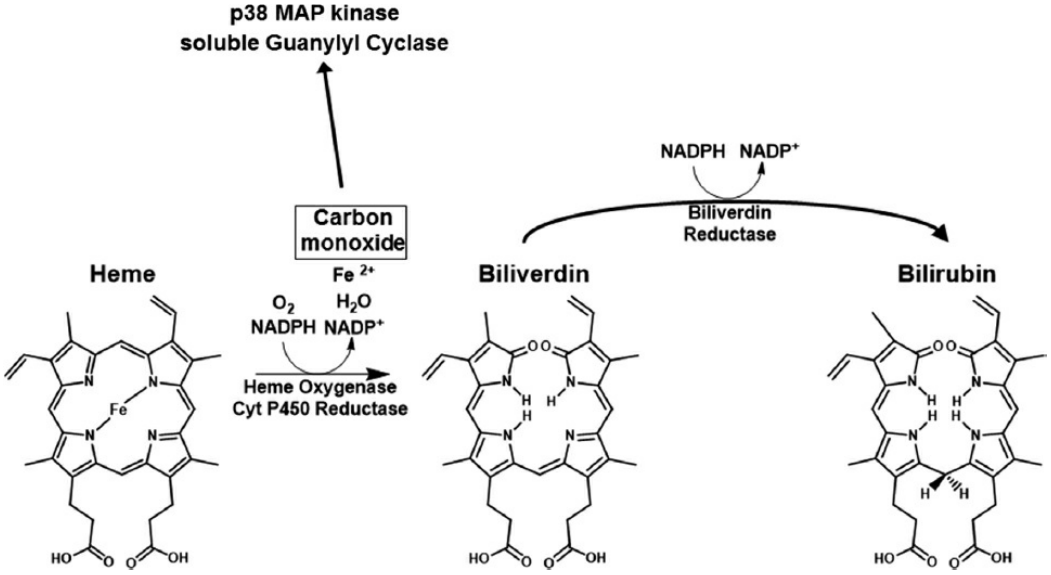

FIGURE 1.8–6. Synthesis of carbon monoxide, an atypical neurotransmitter. Gaseous carbon monoxide (CO) is enzymatically synthesized in neurons via the enzyme, heme oxygenase, also converting heme into the molecule biliverdin, and liberating free iron (Fe). Similar to nitric oxide, CO is not stored in neuronal vesicles and can freely diffuse across neuronal membranes. CO also similarly activates soluble guanylyl cyclase, and leads to activation of multiple intracellular signaling molecules such as p38 MAP kinase. CO exerts its neurotransmitter and signaling functions at concentrations far below that at which classical CO toxicity occurs. The significance of this pathway in neurons is underlined by the existence of two distinct heme oxygenase enzymes, one of which is predominantly expressed in the brain. Biliverdin is converted to bilirubin via the enzyme biliverdin reductase. Similar to CO, bilirubin is no longer relegated to the status of toxic byproduct and may be an important antioxidant.

Carbon Monoxide and Neurotransmission. Carbon monoxide appears to participate in the neurotransmission of odorant perception. Odorants lead to carbon monoxide production and subsequent cGMP synthesis that promotes long-term adaptation to odor stimuli. Carbon monoxide has the potential to regulate a variety of perceptual and cognitive processes that are yet untested. Similarly, in the rat retina, long periods of light exposure lead to increased HO1 expression, carbon monoxide production, and cGMP signaling. Carbon monoxide may also participate in adaptation to chronic pain. HO_2-deficient animals manifest reduced hyperalgesia and allodynia after exposure to chronic pain stimuli. Carbon monoxide may thus set the threshold for pain perception, although it is unclear whether the effect occurs in the CNS and the PNS. Aside from its role in promoting cGMP production, carbon monoxide may also directly bind to and open the BK_{Ca} channel, leading to as yet uncharacterized effects on neurotransmission.

In the gastrointestinal nervous system, carbon monoxide serves as a neurotransmitter to relax the internal anal sphincter in response to nonadrenergic noncholinergic (NANC) nerve stimulation and VIP. Mice

rendered genetically deficient of HO_2 demonstrate a 50 percent reduction in NANC neurotransmission, as do nNOS-deficient animals. Mice that are bred to have loss of both HO_2 and nNOS have their NANC neurotransmission abolished, establishing a physiologic process mediated entirely by gaseous neurotransmitters.

Carbon monoxide has been implicated in the development of hippocampal LTP, although lines of evidence are contradictory. Carbon monoxide and tetanic stimulation of nerves lead to increased excitatory postsynaptic potentials (EPSPs). HO inhibitors that block carbon monoxide production lead to impaired induction of LTP and reduced calcium-dependent release of glutamate neurotransmitter. However, HO_2-deficient animals fail to manifest any differences in LTP. These disparate findings may be explained by a role for HO1 in LTP, or an ability of HO inhibitors to nonspecifically block some other processes important to LTP induction.

Animals with loss of HO_2 demonstrate a reduced fear of falling from a suspended wire and greater exploratory behavior in open field testing. These animal-model correlates are consistent with a role for carbon monoxide signaling in anxiety states. Subtle abnormalities on memory have been noted in these HO_2 knockout mice, although findings have been inconsistent. As HO_2 is abundantly expressed in male testis and penile autonomic ganglia, it is unsurprising that carbon monoxide and nitric oxide both appear to regulate the male reproductive ANS. HO_2-deficient mice manifest abnormal neurotransmission of the myenteric plexus and diminished bulbospongiosus muscle reflex, leading to ejaculatory abnormalities.

Other Signaling Roles of Carbon Monoxide. The expression of HO_2 in the hypothalamus led investigators to test whether carbon monoxide could regulate the release of peptide hormones. In animal models, carbon monoxide was found to block the secretion of both oxytocin and vasopressin from the hypothalamus. Cell culture systems have also suggested that carbon monoxide inhibits corticotrophin releasing factor (CRF) release from hypothalamic cells, but near-toxic levels do the reverse, stimulating CRF release.

Heme metabolism and carbon monoxide production appear to be regulated in a circadian fashion, consistent with a role in the regulation of sleep/wake cycles. The mammalian transcription factor NPAS2 binds BMAL1 to form a complex (NPAS2/BMAL1) that, along with Clock/BMAL1, activates transcription of period and cryptochrome proteins. Period and cryptochrome have dual functions, inhibiting the circadian rhythm machinery, but also blocking NPAS2/BMAL1 and Clock/BMAL1, forming a periodic circuit. NPAS2 contains two heme moieties that can bind carbon monoxide, leading to the inhibition of the NPAS2/BMAL1 complex being

able to bind to DNA and regulate transcription.

At toxic levels, carbon monoxide is well known to impair oxygen transport by binding to hemoglobin with higher affinity than oxygen. Amazingly, carbon monoxide itself plays a physiologic role in the mechanism by which the carotid body senses oxygen. HO, expressed in glomus cells of the carotid body, uses oxygen as a substrate in the production of carbon monoxide (Fig. 1.8–6). When oxygen levels drop, so does carbon monoxide production, leading to a resetting of the threshold in which the carotid body senses oxygen. The molecular mechanism may occur via carbon monoxide regulation of the carotid body BK ion channel.

Neuroprotective Roles of the Heme Oxygenase Pathway. Mice rendered genetically deficient in HO_2 manifest increased susceptibility to traumatic brain injury and stroke damage, consistent with a role for the pathway in protecting the brain against neurotoxic insults. The neuroprotective function of HO may be impaired in Alzheimer disease as HO is found in amyloid plaques. The amyloid precursor protein (APP), a source for toxic amyloid beta fragments, can bind to and inhibit HO neuroprotective function, and APP mutants associated with early-onset Alzheimer disease are the most potent at blocking HO function.

Hydrogen Sulfide: The Newest Gaseous Messenger Molecule

As carbon monoxide and bilirubin had reputations of toxicity prior to the appreciation of their physiologic functions, a similar tale is unfolding for the gas hydrogen sulfide (chemical formula, H_2S). While still abundantly recognized as a foul-smelling and toxic emission of bacteria and sewage treatment plants, hydrogen sulfide may yet prove to be a significant neuromodulator and neurotransmitter.

At least two different enzymes can generate hydrogen sulfide, cystathionine beta-synthase (CBS) and cystathionine gamma-lyase (CSE). Each catalyzes the same reaction, converting cysteine and water to hydrogen sulfide, pyruvate, and ammonia. Intriguingly, each enzyme also catalyzes an independent reaction. CBS converts homocysteine and serine to cystathionine, which CSE uses to produce cysteine, ammonia, and 2-oxobutyric acid.

In the brain, hydrogen sulfide exists at concentrations as high as 160 micromolar, consistent with a role in regulating brain function. CBS is abundantly expressed in brain, while CSE is undetectable. Similar to the enzymes that generate the other gaseous neurotransmitters, CBS is activated by calcium/calmodulin. CBS-deficient mice have altered hippocampal LTP, and hydrogen sulfide potentiates NMDA receptor currents. While much remains to be discovered, hydrogen sulfide can activate ATP-sensitive potassium channels, dilate arterioles, and increase

the activity of the signaling kinase, ERK.

ENDOCANNABINOIDS: FROM MARIJUANA TO NEUROTRANSMISSION

Known as cannabis, hemp, hashish, ma-fen, or a variety of slang terms, marijuana has been cultivated and utilized by human populations for thousands of years. Despite long debates as to whether its risks and benefits are evenly matched, it has only been in recent decades that some of the mystery has been revealed by which marijuana exerts its effects on the brain. The "high" users experience, euphoria and tranquility, relates to cannabis acting upon a neural pathway involving cannabinoids endogenous to the human brain, or endocannabinoids.

The first described medicinal use of cannabis dates to approximately 2700 B.C. in the pharmacopeia of Chinese Emperor Shen Nung, who recommended its use for a variety of ailments. At this time, adverse properties were also apparent, and a large amount of the fruits of hemp could lead to "seeing devils" or a user might "communicate with spirits and lightens one's body." For centuries, cannabis was employed in India as an appetite stimulant, and habitual marijuana users remain well acquainted with "the munchies."

For many years, the mechanisms by which the active components of marijuana, *cannabinoids,* exerted their psychoactive effects remained a mystery. Chemists sought to isolate the psychoactive components of cannabis from the many components of the plant oil (Table 1.8–1). First elucidated in 1940 was cannabinol, a compound now appreciated to be an oxidation product of other cannabinoids and not a psychoactive compound (Fig. 1.8–7). However, with a structure in hand, chemists were now able to synthesize synthetic cannabinoids that did possess psychoactive properties. Within a few years, it was apparent that tetrahydrocannabinols might be the active components of cannabis. Following advances in chemical separation techniques, Mechoulam and Gaoni in 1964 identified delta-9-tetrahydrocannabinol (THC), a compound that accounts for nearly all of the psychoactive effects of cannabis. THC acid is the predominant form of plant THC, and this is readily converted to THC upon heating, such as when cannabis is smoked.

Table 1.8–1.
Selected Discoveries in Cannabinoid Research

1899	► Cannabinol isolated from cannabis resin
1940	► Identification of cannabinol structure
1964	► Discovery of the structure of THC, the most psychoactive component of cannabis
1988	► Specific THC-binding sites identified in brain
1990s	► Identification of a brain cannabinoid receptor, CB1
	► Discovery of the first endogenous brain endocannabinoid, anandamide
	► Identification of a second cannabinoid receptor, CB2

	► Rimonabant, a CB1 receptor blocker, is developed ► Report of a second endocannabinoid, 2-AG ► FAAH, an endocannabinoid degrading enzyme, is discovered
2000s	► FAAH inhibitors reduce anxiety-like behaviors in animal studies ► Identification of enzymes that synthesize endocannabinoids ► MAGL, a second endocannabinoid degrading enzyme, is found ► Abuse of synthetic cannabinoids is increasingly recognized ► Rimonabant approved for use in Europe for weight-loss ► Rimonabant meta-analysis finds increased anxiety and depressive symptoms in humans without a history of psychiatric illness

Discovery of the Brain Endocannabinoid System

Estimates suggest that 20 to 80 μg of THC reach the brain after one smokes a marijuana cigarette (i.e., "joint"). This is comparable to the 100 to 200 μg of norepinephrine neurotransmitter present in the entire human brain. Thus the effects of THC might be explained by effects upon neurotransmitter systems. In the 1960s, there were at least two schools of thought on how THC exerted its psychoactive effects. One held that THC worked in a manner similar to the inhaled volatile anesthetics, that is, no specific receptor existed and it might have a generalized effect on neuronal membranes or widespread actions upon neurotransmitter receptors. A competing school of thought speculated that specific receptors for cannabinoids existed in the brain, but they were difficult to identify due to the lipophilic nature of these chemicals. Novel cannabinoids were synthesized that were more water soluble, and in the late 1980s, this allowed for the discovery of a specific cannabinoid receptor, named CB1.

Now that a cannabinoid receptor was known to exist in the brain, the research community felt that it was unlikely that such receptors would have evolved solely to allow for action of plant cannabinoids. Indeed, after Pert and Snyder discovered opiate receptors in the 1970s, it was soon apparent that these did not evolve for the purpose of morphine drugs, but as the targets of the endogenous enkephalin neurotransmitters, discovered soon after. A hunt for an endogenous brain ligand for the CB1 receptor was underway, and the existence of such a substance was hypothesized to be an endogenous cannabinoid. This proved true in 1992 when Mechoulam and colleagues reported the discovery of anandamide, a lipid produced endogenously in the brain that could activate cannabinoid receptors and function as a neurotransmitter (Fig. 1.8–8). The name of this substance was derived from the Sanskrit word, *ananda,* which translates as bliss. Anandamide could also mimic THC in a variety of animal behavioral tests that generally predict whether a substance will have psychoactive properties in humans, including inhibition of spontaneous movement, promotion of freezing spells, reducing pain sensitivity, and decreasing body temperature.

Several additional endocannabinoids were soon discovered: 2-

arachidonylglycerol (2-AG), *N*-arachidonyldopamine (NADA), 2-arachidonoylglycerol ether (noladin ether), and virodhamine (Fig. 1.8–8). The reason for having several different endocannabinoids may lie with their differing affinities for the cannabinoid receptors CB1 and CB2. Anandamide appears to have the greatest selectivity for the CB1 receptor, followed by NADA and noladin ether. In contrast, virodhamine prefers CB2 receptors and has only partial agonist activity at CB1. 2-AG appears not to discriminate between CB1 and CB2.

Biosynthesis of Endocannabinoids

Arachidonic acid is utilized as a building block for biosynthesis of endocannabinoids, prostaglandins, and leukotrienes, and is found within cellular phospholipids of the plasma membrane and other intracellular membranes. Synthesis of anandamide requires the sequential action of two enzymes (Fig. 1.8–9). In the first reaction, the enzyme NAT transfers an arachidonic acid side chain from a phospholipid to phosphatidylethanolamine (PE), generating NAPE. In the second reaction, the enzyme NAPD-PLD converts NAPE to anandamide. As NAPE is already a natural component of mammalian membranes, it is the second step which generates anandamide that is most crucial to neurotransmission.

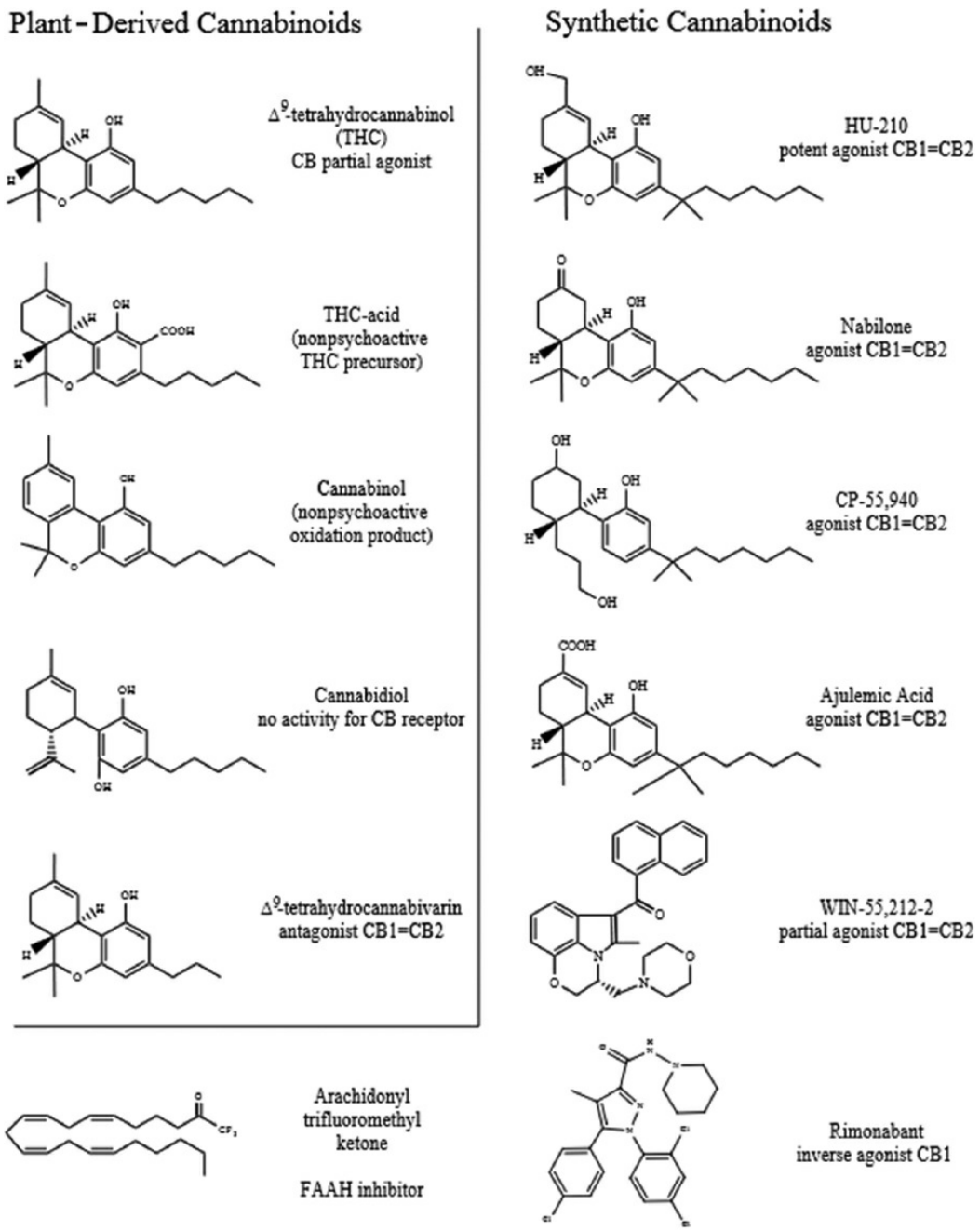

FIGURE 1.8–7. Selected plant-derived and synthetic cannabinoids. THC is the main psychoactive component of cannabis. The drug rimonabant (Acomplia) is a potent inverse agonist for cannabinoid CB1 receptors. FAAH, fatty acid amide hydrolase.

Biosynthesis of 2-AG also requires two enzymes. In the first step, a phospholipid containing arachidonic acid at the middle position is converted to sn-1-Acyl-2-arachidonyl glycerol (DAG) via the action of the enzyme PLC. The second reaction generates 2-AG via either of two specific diacylglycerol lipases (DAGLs). The enzymes for biosynthesis of the other endocannabinoids are undefined.

Endocannabinoids are not stored in synaptic vesicles for later use, but are synthesized on demand as is done for the gaseous neurotransmitters. An important criterion for a signaling molecule to be considered a

neurotransmitter is that neuronal depolarization should lead to its release. Depolarization leads to increases in cellular calcium, which in turn promotes synthesis of the endocannabinoids and their release. The mechanism is explained in part by calcium activation of NAPE-PLD and DAGL, leading to augmented biosynthesis of anandamide and 2-AG, respectively.

Endocannabinoids generated in a neuron must cross the synaptic cleft to act upon cannabinoid receptors. Similar to THC, endocannabinoids are highly lipophilic, and thus poorly soluble in cerebrospinal fluid. It is hypothesized that a specific endocannabinoid transporter exists to allow endocannabinoids to cross the synaptic cleft and allow entry into the target neuron.

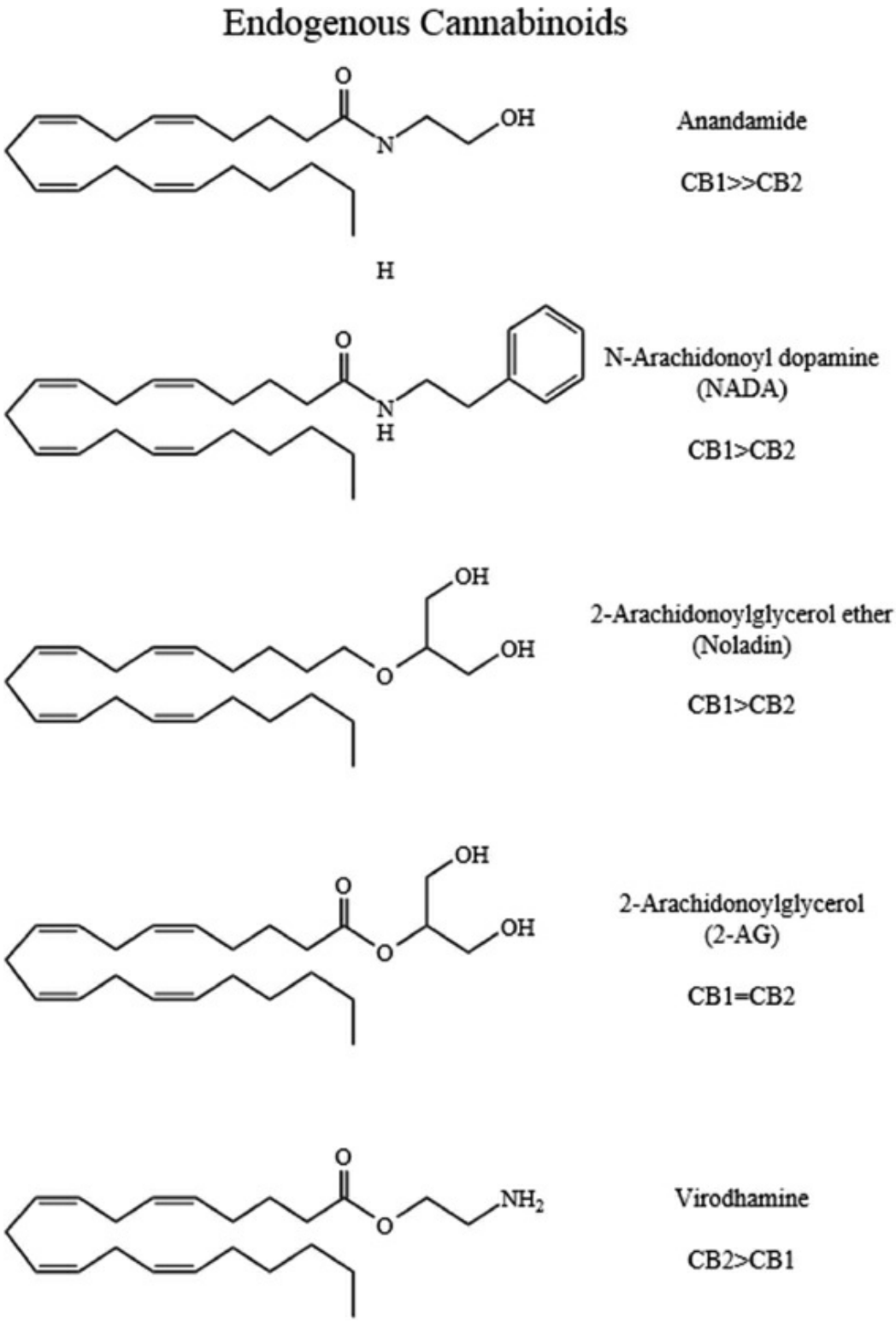

FIGURE 1.8–8 Endogenous cannabinoids. At least five endocannabinoids exist in the mammalian brain, each differing in affinity for CB1 and CB2 cannabinoid receptors. All are derived from the essential omega-6 fatty acid, arachidonic acid, which is also a substrate in the formation of prostaglandins and leukotrienes.

Inactivation of Endocannabinoids

Neurotransmitters are typically inactivated either by reuptake from the

neurons that release them or degradation by highly specific enzymes, such as the example of acetylcholine being hydrolyzed by AChE. At least two enzymes exist to target the destruction of endocannabinoids and attenuate their neurotransmission. Fatty acid amide hydrolase (FAAH) converts anandamide to arachidonic acid and ethanolamine (Fig. 1.8–9). FAAH is found in regions of the brain where CB1 receptors are predominant, and localizes to postsynaptic neurons where anandamide is made. Rapid degradation of anandamide in part explains its relatively low potency compared to THC. Confirming a role of FAAH in anandamide inactivation, knockout mice without FAAH exhibit a 15-fold increase of anandamide, but not 2-AG. These mice have greater behavioral responses to exogenous anandamide, owing to its decreased degradation. The endocannabinoid 2-AG is inactivated by FAAH, but also by a MAGL located in presynaptic neurons.

Pharmacologic inhibitors of FAAH have analgesic effects and reduce anxiety in animal models, but do not have the undesirable effects of THC such as immobility, lowered body temperature or greater appetite. Such a pharmacologic strategy would be analogous to MAOI and catechol o-methyltransferase inhibitors (COMTIs). MAOI, used in depression, inhibit the breakdown of serotonin and other monoamines, thereby increasing serotonin, while COMTI serve an analogous role in blocking the destruction of dopamine and other catecholamines.

Cannabinoid Receptors

Underscoring their importance in neural functions, CB1 receptors are possibly the most abundant GPCRs in the brain. They occur at highest density in the basal ganglia, cerebellum, hippocampus, hypothalamus, ACC, and cerebral cortex, particularly the frontal cortex. Humans or animals that receive large doses of THC develop catalepsy, a reduction of spontaneous movement and freezing into bizarre and unnatural postures. The action of cannabinoids in the basal ganglia and cerebellum may be associated these behaviors, which may prove relevant in understanding catatonic symptoms in schizophrenia.

CB1 receptors are predominantly found on axons and nerve termini, with little present on neuronal dendrites and the cell body. CB1 receptors tend to be localized to the presynaptic rather than postsynaptic side of the neuronal cleft, suggesting a role in regulation of neurotransmission. A second cannabinoid receptor, CB2, is predominantly expressed on the surface of white blood cells of the immune system, but small amounts appear to be present in the brainstem.

Effects on Neurotransmission. The cannabinoid CB1 receptor is associated with G proteins that mediate its intracellular signaling, in part, through inhibition of adenylyl cyclase. This leads to a decrease in levels of

the important second messenger, cAMP. Activation of the CB1 receptor also leads to activation of potassium channels and inhibition of N-type calcium channels. As calcium is integral to neurotransmitter release, cannabinoids can block neurotransmission through this mechanism. Cannabinoid receptors also activate mitogen-activated protein kinases (MAPK).

Via the use of cell culture models and slices of brain, cannabinoids have been shown to block the release of a variety of neurotransmitters, including GABA, norepinephrine, and acetylcholine. Norepinephrine and acetylcholine tend to be excitatory neurotransmitters, and cannabinoid inhibition of their release would be expected to have an overall inhibitory effect. However, GABA is an inhibitory neurotransmitter, and cannabinoid inhibition of it would lead to overall excitatory effects, demonstrating that cannabinoids can have complex effects on neurotransmission depending upon the specific context. Cannabinoids also appear to increase the release of brain endorphin neurotransmitters and increase dopamine release in the nucleus accumbens, a "reward center" relevant to addiction and learning. The endocannabinoids have been implicated in a variety of forms of synaptic plasticity, including LTP and LTD.

Retrograde Transmission Regulated by Endocannabinoids

It has long been apparent to neuroscientists that a postsynaptic neuron could regulate the activity of a presynaptic neuron, for instance, inhibiting further release of neurotransmitters by the presynaptic neuron. Endocannabinoids may be the best candidate to date as the *retrograde messenger* that diffuses from a postsynaptic neuron to act upon a presynaptic neuron. During development, the enzymes responsible for cannabinoid synthesis are expressed in both pre- and postsynaptic neurons. However, in adult brains, the synthesis of endocannabinoids appears to be predominantly in postsynaptic neurons. This suggested that they might work *backwards* and regulate neurotransmission of a presynaptic neuron.

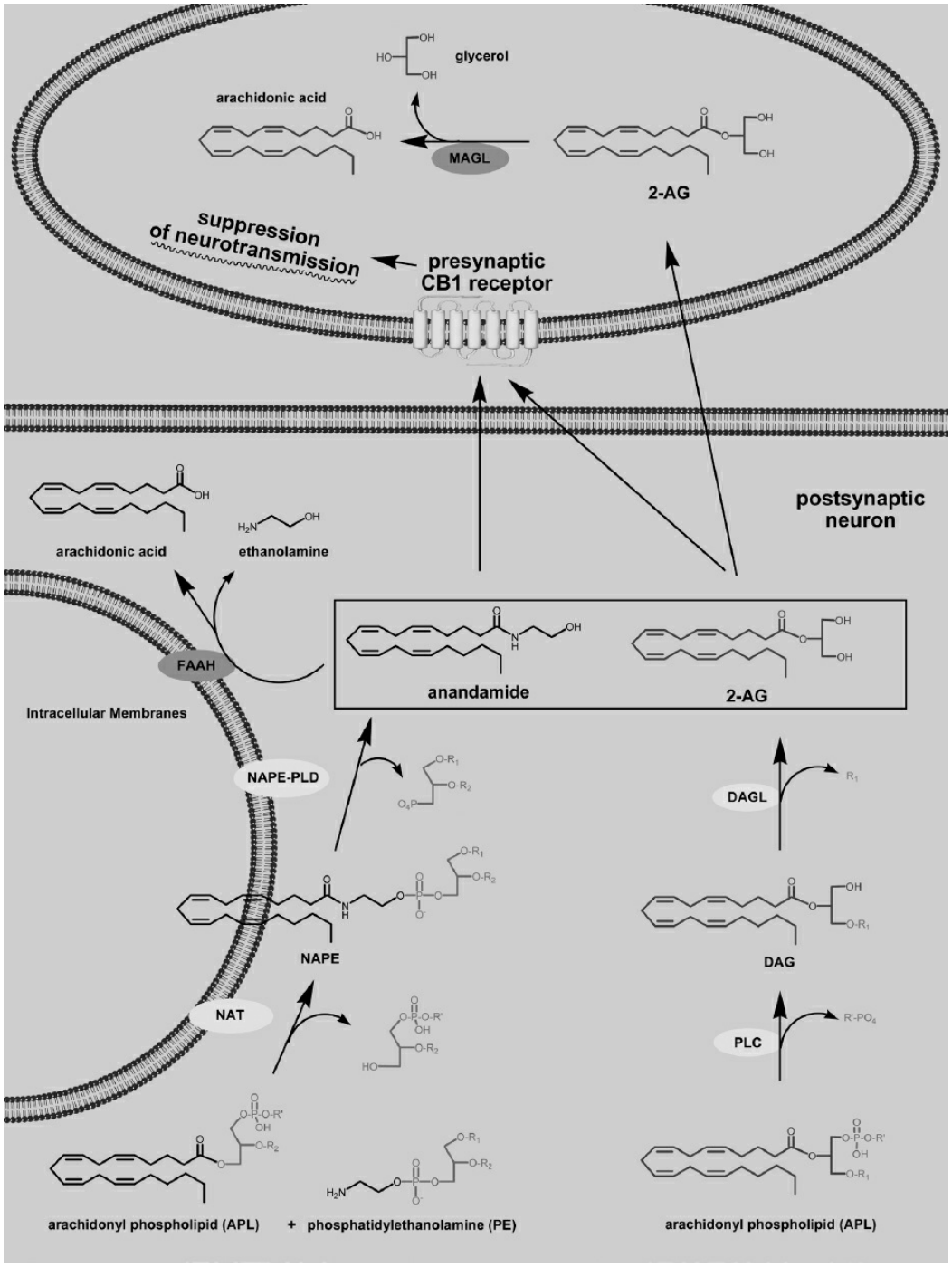

FIGURE 1.8–9. Retrograde neurotransmission of the endocannabinoids, anandamide and 2-arachidonylglycerol (2-AG). Anandamide (*black*) is synthesized on demand for neurotransmission via a two-step process. The enzyme NAT transfers the arachidonic acid chain from a phospholipid (APL) to phosphatidyl ethanolamine (PE), producing NAPE. A second enzyme, NAPE-PLD, generates anandamide. 2-AG (*black*) is similarly synthesized in two steps by the enzymes PLC and DAGL. The endocannabinoids made in a postsynaptic neuron cross the synapse and activate presynaptic CB1 receptors, suppressing neurotransmission of the presynaptic neuron (although activation of the presynaptic neuron occurs in some cases). Enzymes involved in endocannabinoid synthesis are *light gray*, those that break them down in *dark gray*. 2-AG is predominantly inactivated in the presynaptic neuron by MAGL, whereas anandamide is destroyed in the postsynaptic neuron by FAAH. PE, phosphatidylethanolamine; APL, arachidonyl phospholipids; NAT, N-acyltransferase; NAPE, N-arachidonyl-phosphatidylethanolamine; NAPE-PLD, N-arachidonyl-phosphatidylethanolamine phospholipase D; FAAH, Fatty acid amide hydrolase;

MAGL, monoacylglycerol lipase; PLC, phospholipase C; PDE, phosphodiesterase; DAG, diacylglycerol; DAGL, diacylglycerol lipase; R_1-R_3, various acyl or alkyl side chains of phospholipids; R', side chain of phospholipid head group.

A typical presynaptic neuron containing dopamine or glutamate releases its neurotransmitter, leading to depolarization of the postsynaptic neuron. This second neuron can release endocannabinoid that diffuses across the synaptic cleft and inhibits further neurotransmitter release from the presynaptic neuron. Such a mechanism has been identified in the rat hippocampus, in which CB1 receptor antagonists block suppression of the presynaptic neuron. Similarly, mice rendered genetically deficient in CB1 receptors lose hippocampal suppression of the presynaptic GABA neurons (a process also known as depolarization-induced suppression of inhibition, or DSI). The reverse has also been demonstrated, as inhibitors of endocannabinoid destruction enhance retrograde neurotransmission in the hippocampus. Endocannabinoids have also been implicated in regulation of other neurotransmission processes, such as LTD and depolarization-induced suppression of excitation (DSE). The ability of endocannabinoids to inhibit neurotransmitter release may be an important general mechanism by which neurotransmission is regulated.

Endocannabinoids in Anxiety and Mood

Endocannabinoid neurotransmission may be an important regulator of anxiety, and cannabis users regularly describe a tranquilizing effect of THC. Loss of signaling by the endocannabinoid system appears to promote anxiety-like states in animal studies. CB1 receptor-deficient animals exhibit more pronounced anxiety behavior when exposed to stress or new environs.

Marsicano and colleagues suggested a role for cannabinoid signaling in forgetting painful memories. A mouse model was employed in which a tone was paired with an electric shock. Typically, when exposed to the tone, mice "freeze" in anticipation of the shock. Once the shock is no longer paired with the tone, animals demonstrate "extinction," that is, they no longer freeze in response to just the tone by itself. Remarkably, animals deficient in CB1 receptors failed to demonstrate this normal extinction. Thus, endocannabinoid neurotransmission may mediate the ability to "forget" the anxiety associated with a painful memory. CB1 antagonist drugs, given just before the tone, revealed a similar effect. The amygdala participates in many anxiety responses, and endocannabinoids may act upon this brain region to attenuate anxiety. In support of this, levels of anandamide and 2-AG were found to increase in the amygdala immediately following exposure of mice to the tone.

FAAH knockout mice lack the enzyme that degrades endocannabinoids, and exhibit both increased anandamide levels and reduced anxiety in

behavioral tests. Enhancing levels of endocannabinoids may represent a therapeutic target for anxiety. Whereas an agonist might over-activate cannabinoid receptors where little neurotransmission normally occurs, blocking the breakdown of endocannabinoids would be expected to facilitate activity in areas already utilizing these molecules, and thereby having fewer side effects. Novel FAAH inhibitors reduce the breakdown of anandamide and reduce anxiety-like behaviors in animals. Though the "forced swim test" and "tail suspension test" are far from perfect models of depression in the mouse, FAAH inhibitors improved the ability of the animals to cope with these stresses, a benefit also observed by treatment with antidepressant drugs.

The endocannabinoid pathway may represent an attractive target in understanding posttraumatic stress responses and phobias. Although one cannot yet routinely measure brain endocannabinoid levels in human subjects, this model is supported by clinical trials of the cannabinoid receptor blocker, rimonabant, which may offer promise as a strategy for weight loss (see below). A frequent adverse reaction to the drug is increased anxiety and depression. In a 2007 meta-analysis, Christensen et al. reported that those receiving rimonabant had a 2.5 times greater risk of stopping treatment because of depression and a threefold greater risk of stopping due to anxiety. These psychiatric side effects occurred despite the studies having excluded those who had a history of recurrent depressive or anxiety disorders, underlining an important role for this system in the regulation of anxiety and mood.

Endocannabinoids may play a role in mood disorders, as cannabis use is associated with a tranquilizing effect upon mood, while some users experience paradoxical anxiety. A postmortem study of depressed suicides found increased CB1 receptors in the prefrontal cortex, also observed in a study of alcoholic suicides. Genetic association studies have sought links between the CB1 receptor and psychiatric illness. While results have been mixed, some have suggested links to depression in Parkinson disease and an association with attention deficit disorder in alcoholic patients.

Addiction. The endocannabinoid system may be an attractive target for the understanding of addiction. Mice deficient in CB1 receptors are unsurprisingly resistant to the behavioral effects of cannabinoids; however, they also appear to have reduced addiction to and withdrawal from opiates. Further interaction has also been found between the opioid and cannabinoid systems as cannabinoids appear to increase the release of dopamine in the nucleus accumbens, a key reward area of the brain implicated in addiction. This dopamine release appears to require mu-opioid receptors, as pharmacologically inhibiting these receptors blocks the ability of cannabinoids to increase dopamine release. Rats with a preference for alcohol have decreased FAAH activity, suggestive of greater

cannabinoid signaling. CB1 receptor antagonists dampen their alcohol consumption, while inhibiting FAAH increases their alcohol consumption. Furthermore, CB1-deficient animals also appear to have reduced alcohol intake. A single amino acid mutation in human FAAH has been found to be associated with drug abuse, and this abnormal enzyme appears to be less stable than its wild-type counterpart.

Endocannabinoids in Psychosis

Heavy use of cannabis can produce psychotic symptoms in individuals with no prior history of psychiatric disorder, though it is unclear whether this is solely due to the drug or an underlying vulnerability to psychosis in such persons. Cannabis use often worsens psychosis in schizophrenia and heavy use has been associated with developing schizophrenia, although some suggest that this association is an accelerated development of symptoms in those who would eventually manifest schizophrenia. Nonetheless, the endocannabinoid system has implications for the pathophysiology of schizophrenia, as cannabinoid signaling appears to increase the release of dopamine. Medications that act as antagonists of dopamine D2 receptors will likely remain a component of schizophrenia treatment for some time.

Elevated levels of anandamide have been noted in the CSF from individuals with schizophrenia, including in a follow-up study of medication-naïve patients. Elevated anandamide levels have also been noted in the blood of those with schizophrenia, and such elevations normalized with clinical improvement.

Elevated CB1 receptor levels have been reported in postmortem brain samples from those with schizophrenia, particularly the DLPFC and cingulate cortex. While some have identified polymorphisms in the CB1 receptors associated with schizophrenia, others found no association. Furthermore, the implications of such polymorphisms on CB1 function are unclear.

Feeding. Following drug ingestion, THC users develop an increased appetite ("the munchies"), and cannabis has been utilized as an appetite stimulant for centuries. This effect may depend on the CB1 receptors present in the hypothalamus. Endocannabinoid levels increase in the hypothalamus and limbic system when animals are deprived of food. Mice genetically deficient in CB1 receptors become resistant to developing obesity after they are given a high-fat diet. Similarly, the CB1 receptor antagonist rimonabant appears to facilitate weight loss by blocking cannabinoid signaling. In a clinical trial of over 3,000 obese patients, those treated with 20 mg/day of rimonabant lost 6.3 kg at 1 year, compared to 1.6 kg in the placebo group. Nausea was the most common side effect reported. A 2007 meta-analysis of clinical trials reported an overall 4.7 kg weight loss with rimonabant treatment, besting the weight-loss drugs orlistat (2.9 kg)

and sibutramine (4.2 kg).

Effects on Brain Injury and Pain

In mouse models of traumatic brain injury, 2-AG appears neuroprotective, reducing brain edema, infarct size, and cell death while improving functional outcomes. Anandamide also protected against brain injury in a model of multiple sclerosis, and human patients with the disease have increased production of anandamide. A study of cannabinoid agonist HU-211 led to more rapid clinical improvement following head trauma. FAAH inhibitors improved motor symptoms in a mouse model of Parkinson disease, likely via cannabinoids increasing dopamine neurotransmission.

Neurotransmission via the endocannabinoid pathway is increasingly appreciated to regulate pain perception. THC and cannabinoid agonists have proven effective in animal models of acute and chronic pain, ranging from burn injury to nerve damage and inflammation. The CB1 receptor plays an important role in these effects as the analgesic effects of cannabinoid drugs are lost when CB1 antagonist rimonabant is given. Similarly, the analgesic effect of THC is lost in mice genetically deficient in the CB1 receptor.

Stress has long been associated with diminished pain perception, such as in cases of injured military personnel who demonstrate remarkable pain tolerance, a phenomenon known as stress-induced analgesia. The endocannabinoid system may mediate these effects. Animal models reveal anandamide and 2-AG production after stress, and stress-induced analgesia is blocked by CB1 blocker rimonabant in these animals. In a placebo-controlled, randomized study of over 600 individuals with multiple sclerosis, Zajicek and colleagues found that oral THC administration led to improvement in mobility and pain. However, THC offered little benefit in postoperative pain following hysterectomy.

Endocannabinoid regulation of pain perception appears to be distinct from that of the endogenous opiate system, but the two pathways may share overlapping neural pathways. Evidence for this has been provided using CB1 blocker rimonabant and naloxone, which blocks opiate receptors. Rimonabant attenuates analgesia provided by THC and cannabinoids, but only partly blocks the response to morphine. However, the opposite is true for opiates: naloxone blocks morphine-induced analgesia but also partially blocks the analgesia of THC and cannabinoid drugs. Combinations of cannabinoid and opiate drugs evince synergistic analgesic effects in animal models.

Although it was initially assumed that cannabinoids exert their analgesic effects via the CNS, in animal models it has been shown that localized administration of cannabinoids may also be effective, including drugs selective for the CB2 receptor, whose expression is minimal in the CNS.

Endocannabinoids may also influence pain sensitivity by mechanisms that do not involve the CB1 and CB2 receptors. Both anandamide and NADA can also activate a calcium channel known as the vanilloid receptor (also known as TRPV-1) found on sensory nerves. This same receptor is notable for being activated by capsaicin, which causes the hot sensation of chili peppers. Thus, endocannabinoids can exert opposing functions: promoting analgesia through the CB1 and CB2 receptors, but potentially increasing pain via TRP channels. While CB2 receptors are largely expressed in the periphery, postmortem analyses reveal an upregulation in brain from those with Alzheimer disease.

The rapid development of novel cannabinoid drugs may allow for targeting of specific symptoms, rather than elicit all of the typical effects of THC. For instance, ajulemic acid demonstrates analgesic and anti-inflammatory properties, but may offer a benefit of limited psychoactive side effects. In a randomized clinical trial of this compound, ajulemic acid proved efficacious in reducing chronic neuropathic pain.

Effects in the Periphery

Cannabinoids lead to direct relaxation of vascular smooth muscle by local CB1 receptors. This vasodilatation extends to the conjunctiva of the eyes, leading to a “bloodshot” appearance in some cannabis users. Relaxation of ocular arteries by cannabinoids may offer utility as a treatment for glaucoma, a condition of high intraocular pressure, and activation of CB1 receptors in the kidney can improve renal blood flow. A role in generalized blood pressure regulation is unproven, and blood pressure is unaltered in persons treated with rimonabant or animals deficient in CB1 receptors. Cannabinoid signaling may also be relevant to ectopic pregnancy, as CB1-deficient mice retain many embryos in the oviduct.

Nonpsychoactive Cannabinoids

Although THC is the principal psychoactive component of cannabis, the many nonpsychoactive cannabinoids also have intriguing properties and may regulate neurotransmission. Cannabidiol may offer potential therapeutic effects, and appears to stimulate TRP-V1 receptors and influence endocannabinoid degradation. Cannabidiol also demonstrated a protective effect in a mouse model of inflammatory arthritis and has been purported to have anticonvulsant properties. Although results have been mixed, purified cannabidiol may also exert antipsychotic activity, although the net effect of plant cannabis use typically exacerbates schizophrenia symptoms owing to the THC component. Tetrahydrocannabivarin is a plant cannabinoid that antagonizes CB1 receptors. It is a candidate marker to distinguish whether a patient has been using plant-derived cannabis or prescription THC, which contains no tetrahydrocannabivarin.

EICOSANOIDS

Overview

Clinical findings suggest that the dietary supplements omega-3 fatty acids-eicosapentaenoic acid (EPA), its ester ethyl-eicosapentaenoic acid (E-EPA) and docosahexaenoic acid (DHA), help relieve symptoms of depression, bipolar illness, schizophrenia, and cognitive impairment. DHA and EPA may help reduce behavioral outburst and improve attention in children.

Chemistry

Essential fatty acids are a group of polyunsaturated fats that contain a carbon–carbon double bond in the third position from the methyl end group in the fatty acid chain. They are essential because unlike monosaturated and saturated fatty acids, polyunsaturated fatty acids cannot be synthesized de novo and can only be acquired through diet from natural fats and oils. Linoleic acid (LA) is the parent compound of omega-6 fatty acids, and α-linolenic acid (ALA) is the parent compound of omega-3 fatty acids. Both omega-3 and omega-6 groups use the same enzymes for desaturation and chain elongation. Omega-3 fatty acids are synthesized by algae and plankton. Fish such as herring, salmon, mackerel, and anchovy feed on these aquatic species and become a rich, dietary source of omega-3. EPA and DHA are highly unsaturated omega-3 fatty acids that contain 6 and 5 double bonds on their long structural chain, respectively. They are positioned in the cell membrane by phospholipids and play a crucial role in cell membrane signaling.

Effects on Specific Organs and Systems

The strongest scientific evidence for treatment with fatty acid supplements comes from cardiovascular literature. Several human trials have demonstrated that omega-3 fatty acids lower blood pressure, reduce the rate of recurrent myocardial infarction, and lower triglyceride levels. In the nervous system, fatty acids are essential components of neurons, immune cells, and glial phospholipid membrane structures. They increase CBF, decrease platelet aggregation, and delay progression of atherosclerosis in the cardiovascular system. Omega-6 fatty acids appear to reduce inflammation and neuronal apoptosis, and decrease phosphatidylinositol second-messenger activity. Omega-3 fatty acids have been suggested to alter gene expression.

In the CNS, fatty acids are selectively concentrated in neuronal membranes and involved in cell membrane structure. Omega-6 arachidonic acid has been shown to enhance glutamate neurotransmission, stimulate stress hormone secretion, and trigger glial cell activation in the setting of oxidative toxicity and neurodegeneration. The omega-3 fatty acids DHA

and EPA appear to protect neurons from inflammatory and oxidative toxicities. Increases in serotonin, enhancement of dopamine, and regulation of corticotrophin-releasing factor have been demonstrated in cell culture models.

In rodent models of depression, chronic EPA treatment normalized behavior in open-field tests. Serotonin and norepinephrine were also increased in the limbic regions. Mice fed Omega-3-poor diets had reduced memory, altered learning patterns, and more behavioral problems.

Therapeutic Indications

Clinical research with the use of fish oil for mood disorders has been based on epidemiology studies where there appears to be negative correlation between fish consumption and depressive symptoms. Countries with lower per capita fish consumption had up to 60 times increased rates of major depression, bipolar disorder, and postpartum depression. Observational studies concluded that the lower incidence of seasonal affective disorder (SAD) in Iceland and Japan-than latitude predicted is related to the amount of fatty acid these populations consume in their diet. A study in Norway showed that use of cod-liver oil decreased depressive symptoms. Depression after a myocardial infarction shows higher arachidonic acid to EPA ratio. Postmortem studies in brains of patients diagnosed with MDD show reduced DHA in the orbitofrontal cortex.

The first randomized, controlled pilot study of omega-3 fatty acids focused on adjunctive treatment in both bipolar and unipolar patients with depression, in addition to their standard lithium or VPA treatment. The omega-3 fatty acid group had significant improvement on the Hamilton Depression Rating Scale and a longer period of remission than the placebo group. A subsequent larger study supported a benefit from treatment with E-EPA for bipolar illness. However, a study group of patients with either bipolar disorder or rapid cycling treated with E-EPA showed no significant difference on any outcome measure between the EPA and placebo groups. Bleeding time was also increased in the treatment group. There is no current data on monotherapy in bipolar illness or depression.

In 2002, two landmark double-blind, placebo-controlled studies showed that supplementation with E-EPA or DHA, in addition to standard treatment for unipolar depression, led to a significant reduction in Hamilton Depression Rating scores. A later study with longer duration and a larger number of subjects failed to see an effect of EPA and DHA supplementation on depressed mood or on cognitive function, but also found no adverse effects from the supplements. Although meta-analysis shows significant antidepressant efficacy of omega-3 fatty acids, the authors recognize publication bias and heterogeneity. More large-scale, well-controlled trials are needed to find out the favorable target subjects, therapeutic dose of EPA, and the composition of omega-3 P in treating

depression.

Limited consumption of seafood during pregnancy is recommended in American guidelines, but Japanese studies suggest that low DHA but not EPA increased risk for post-partum depression. A randomized trial of omega-3 fatty acids as monotherapy for MDD during pregnancy showed a reduction in Hamilton Depression Rating scores and compliance with treatment compared to the control group.

The most convincing evidence comes from early brain development and learning studies. Pregnant mothers who consumed foods rich in DHA gave birth to infants that had improved problem-solving skill, but not necessarily improved memory. Visual acuity and eye development are also associated with DHA supplementation during pregnancy.

The Oxford-Durham study of dietary supplementation with omega-3 fatty acids in children with developmental coordination disorder suggests a potentially controversial role in learning disabilities, ADHD, and autism. The authors saw significant reductions in inattention, hyperactivity, and impulsivity. The crossover placebo group improved after switching to fish oil supplements and a multivitamin showed no additional benefits for ADHD symptoms. This led to a plan by education officials in the Durham County Council in England to spend £1 million on omega-3 fish oils to help 5,000 children as they approach their school placement exams in order help to improve their performance.

In behavioral studies, prisoners in England that consumed higher amounts of seafood containing omega-3 fatty acids saw a decrease in the assault rates. A Finnish study of violent criminals identified lower levels of omega-3 fatty acids in their system compared to the nonviolent offenders.

The negative and psychotic symptoms of schizophrenia may be improved with supplementation with omega-3 fatty acids. Antipsychotic medications such as haloperidol appear to have fewer extrapyramidal side effects when combined with antioxidants and omega-3 fatty acids.

EPA and DHA have been associated with decreased dementia incidence. After reviewing the Rotterdam study of a longitudinal cohort of over 5,300 patients, fish consumption appeared to be inversely related to development of new cases of dementia. A later analysis of the study after 6 years demonstrated that low intake of omega-3 fatty acids was not associated with increased risk of dementia. In contrast, the Zutphen study, also in the Netherlands, concluded that fish consumption was inversely related to cognitive decline at both 3 years and after 5 years. Well-designed clinical trials are needed before omega-3 fatty acids can be recommended for prevention of cognitive impairment.

Precautions and Adverse Reactions

The most adverse complication is increased risk of bleeding. Dietary sources can contain heavy metals, and there is no standard preparation for

capsule formulations. Treatment studies have yielded a variety of different doses, but evidence for the therapeutic dose and clinical guidelines are almost nonexistent. The length of treatment still needs to be determined.

NEUROSTEROIDS

Background

While steroids are critical for the maintenance of body homeostasis, neurosteroids are synthesized from cholesterol in the brain and independent of peripheral formation in the adrenals and gonads. Neurosteroids are produced by a sequence of enzymatic processes governed by P450 and non-P450 enzymes either within or outside the mitochondria of several types of CNS and PNS cells.

Recent work has shown that neurosteroids can operate through a nongenomic pathway to regulate neuronal excitability via their effects on neurotransmitter-gated ion-channels. Receptors are generally located in the nucleus, membrane, or microtubules of the CNS and PNS. Although steroids and neurosteroids may act on the same nuclear receptors, neurosteroids differ from steroids in their topological distribution and regional synthesis. The most well-known effect of neurosteroids is on the GABA receptor, particularly the GABAA receptor. Neurosteroids acting primarily at this site include allopregnanolone (3α5α tetrahydroprogesterone), pregnenolone (PREG), and tetrahydrodeoxycorticosterone (THDOC). Dehydroepiandrosterone sulfate (DHEA-S), the most prevalent neurosteroid, acts as a noncompetitive modulator of GABA, and its precursor dehydroepiandrosterone (DHEA) has also been shown to exert inhibitory effects at the GABA receptor. Some neurosteroids may also act at the NMDA, alpha-amino-3-hydroxy-5-methyl-4-isoxazole-propanoic acid (AMPA), kainate, glycine, serotonin, sigma type-1, and nicotinic acetylcholine receptors. Progesterone is also considered a neurosteroid and has the ability to regulate gene expression at progesterone receptors.

Neurosteroids in Neurodevelopment and Neuroprotection

In general, neurosteroids stimulate axonal growth and promote synaptic transmission. Specific neuroprotective effects are unique to each neurosteroid. DHEA acts to regulate brain serotonin and dopamine levels, suppress cortisol, increase hippocampal primed burst potentiation and cholinergic function, decrease amyloid β protein, inhibit the production of proinflammatory cytokines, and prevent free radical scavenging. DHEA and DHEA-S have both been shown to have a role in glial development and neuronal growth, and to promote their survival in animals; their injection into the brains of mice promoted long-term memory while reversing amnesia. Progesterone is linked to myelinating processes such as aiding in

the repair of damaged neural myelination. Allopregnanolone contributes to the reduction of contacts during axonal regression.

Role of Neurosteroids in Mental Illness

Neurosteroids have distinct implications for the maintenance of normal neurologic function and may also contribute to neuropathology. Neurosteroids are differentially regulated in males and females, and may impact the manifestation of psychological disorders in these two populations. Specifically, they play a distinct role in depression and anxiety disorders, and may be targeted by psychiatric medications in the near future.

Depression. When compared with nondepressed controls, studies show that depressed patients have lower plasma and CSF concentrations of allopregnanolone; additionally, this research has elucidated an inverse relationship between allopregnanolone concentrations and severity of depressive illness. However, there are no allopregnanolone-based therapies available for humans, so their direct efficacy is unsubstantiated. Antidepressant drugs, specifically fluoxetine, have been shown in multiple studies to increase the levels of certain neurosteroids. Nonetheless, there is debate over the therapeutic properties of neurosteroids, prompting the investigation of neurosteroid concentrations in patients undergoing nonpharmacological therapies. Preliminary results indicate that the lack of modifications in neurosteroid levels during nonpharmacological treatments supports the validity of the pharmacological properties of antidepressants, not their therapeutic action, in the elevation of neurosteroid levels in medicated populations. In a 2006 clinical study with mirtazapine, allopregnanolone concentrations increased in patients with MDD regardless of the therapeutic benefit.

Anxiety Disorders. In patients with anxiety disorders, the major mechanism of action is on the GABA receptor. Homeostasis characterized by normal GABAergic activity is restored after panic attacks as neurosteroids are released in response to stress. Allopregnanolone stimulates GABAergic activity with 20 times the strength of benzodiazepines and 200 times the potency of barbiturates. The positive and negative regulation of the GABAA receptor is correlated with anxiolytic and anxiogenic action, respectively.

Psychotic Disorders. In addition to their primary relevance to the pharmacological treatment of mood and anxiety disorders, neurosteroids contribute to psychotic, childhood, substance abuse, eating, and postpartum disorders. The effect of neurosteroids on psychotic disorders such as schizophrenia is mediated by DHEA and DHEA-S. DHEA has been

dispensed to decrease anxiety in schizophrenics, as DHEA and DHEA-S suppress GABA inhibition and heighten the neuronal response at the NMDA and sigma receptors. DHEA and DHEA-S levels are typically elevated in schizophrenics' initial episode, indicating neurosteroids are upregulated by the onset of psychosis. Because neurosteroids levels are studied across various illness stages, some questions still exist regarding the role of neurosteroids in psychosis.

Attention Deficit Disorder. In children, the clinical symptoms of ADHD are inversely correlated with DHEA and pregnenolone levels.

Substance Abuse. Alcohol is theorized to regulate the GABA receptor and induce de novo steroid synthesis in the brain; specifically pregnenolone, allopregnanolone, and allotetrahydrodeoxycorticosterone levels are increased in the brain and periphery in response to increases in peripheral alcohol levels. It is hypothesized that sharp increases in ethanol concentration may mimic the acute stress response and elevate neurosteroid concentrations by the HPA axis. To prevent ethanol dependence, researchers are investigating fluctuations in neurosteroids levels and in vivo neurosteroid responsiveness. Neurosteroids (increased allopregnanolone levels in particular) are associated with drug abuse. However, DHEA-S may actually check the acquisition of morphine tolerance. Past research has shown DHEA-S levels were also increased in patients who abstained from cocaine use as per a treatment program, and as patients relapsed, DHEA-S concentrations decreased accordingly.

Eating Disorders. With regard to eating disorders, DHEA has been shown to diminish food intake, temper obesity, moderate insulin resistance, and lower lipids in rats with a model of youth-onset, hyperphagic, genetic obesity. By regulating the serotonergic system, DHEA is hypothesized to promote a reduced caloric load. Although hypothetical, low levels of DHEA and DHEA-S are recorded in young women with anorexia nervosa, and 3 months of oral DHEA supplementation increased bone density and tempered the emotional problems associated with the disorder.

Postpartum and Gynecological Disorders. As estrogen and progesterone levels fluctuate during the course of pregnancy and drop markedly after delivery, neurosteroids are thought to contribute to postpartum disorders. Low postpartum DHEA concentrations have been linked to mood instability. In addition, allopregnanolone levels correlated with mood disorders during pregnancy and in pre-menstrual syndrome (PMS). It has been noted that women with pre-menstrual dysphoric disorder have higher allopregnanolone/progesterone ratios than normal controls; women

treated for this disorder reported improvement as allopregnanolone levels deceased.

Neurosteroids, Memory Disorders, and Aging. Neurosteroid levels may be irregular in neurodegenerative disorders and aging conditions such as Alzheimer and Parkinson diseases. DHEA levels at age 70 are only about 20 percent of their maximum value recorded in the late twenties, and some researchers believe DHEA supplementation can prevent or slow the cognitive declines associated with the aging process. However, conflicting studies have indicated that DHEA administration does not improve cognitive measures in patients. Additionally, in patients with Alzheimer disease, DHEA concentrations have been found to be markedly decreased.

Conclusion

Neurosteroids may play a role in a variety of psychiatric pathologies. Breakthroughs with animal models do not always correlate with advances in understanding the role of neurosteroids in humans, complicating research. Moreover, further research is exploring neurosteroid levels over the treatment course of a diverse array of psychiatric illnesses to get a more complete picture of disease management.

NOVEL NEUROTRANSMITTERS: BEYOND THE CLASSICAL DEFINITION OF NEUROTRANSMITTER

The classical criteria for a chemical to be consider a neurotransmitter were (a) synthesis in a presynaptic neuron, (b) storage and release from a presynaptic neuron, (c) binding to a receptor on a postsynaptic membrane, and (d) removal from the synaptic cleft by reuptake or degradation. Within the past few decades the discovery of novel neurotransmitters has led to a reformulation of these strict criteria. Messengers such as the gases, cannabinoids, and eicosanoids are not stored in vesicles in presynaptic neurons, but appear to be generated and released "on demand." The endocannabinoids appear to have an important role as *retrograde messengers,* being released from a conventional postsynaptic neuron and acting upon a presynaptic neuron (Fig. 1.8–9). Finally, the gases do not act upon a receptor on the extracellular membrane of a postsynaptic neuron, but diffuse into the cell and act directly upon multiple cellular proteins, bypassing membrane receptors entirely. Additional, yet undiscovered, chemical messengers may further refine the concept of what a neurotransmitter is.

REFERENCES

Appleton KM, Sallis HM, Perry R, Ness AR, Churchill R. Omega-3 fatty acids for depression in adults. *Cochrane Database Syst Rev*. 2015;(11):CD004692.

Cutajar MC, Edwards TM. Evidence for the role of endogenous carbon monoxide in memory processing. *J Cogn Neurosci*. 2007;19:557–562.

Deng J, Lei C, Chen Y, et al. Neuroprotective gases–fantasy or reality for clinical use? *Prog Neurobiol*. 2014;115:210–245.

Di Marzo V, Stella N, Zimmer A. Endocannabinoid signalling and the deteriorating brain. *Nat Rev Neurosci*. 2015;16:30–42.

Grosso G, Galvano F, Marventano S, et al. Omega-3 fatty acids and depression: scientific evidence and biological mechanisms. *Oxid Med Cell Longev*. 2014;2014:313570.

*Iseger TA, Bossong MG. A systematic review of the antipsychotic properties of cannabidiol in humans. *Schizophrenia Research*. 2015;162:153–161.

Iversen LL. *The Science of Marijuana*. Oxford University Press; 2008.

Joy CB, Mumby-Croft R, Joy LA. Polyunsaturated fatty acid supplementation for schizophrenia. *Cochrane Database Syst Rev*. 2006:CD001257.

Keck PE Jr, Mintz J, McElroy SL, et al. Double-blind, randomized, placebo-controlled trials of ethyl-eicosapentanoate in the treatment of bipolar depression and rapid cycling bipolar disorder. *Biol Psychiatry*. 2006;60:1020–1022.

Kidd PM. Omega-3 DHA and EPA for cognition, behavior, and mood: clinical findings and structural-functional synergies with cell membrane phospholipids. *Altern Med Rev*. 2007;12:207–227.

Kreitzer AC, Malenka RC. Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. *Nature*. 2007;445:643–647.

Longone P, Rupprecht R, Manieri GA, Bernardi G, Romeo E, Pasini A. The complex roles of neurosteroids in depression and anxiety disorders. *Neurochem Int*. 2008;52:596–601.

Lynch AM, Loane DJ, Minogue AM, et al. Eicosapentaenoic acid confers neuroprotection in the amyloid-beta challenged aged hippocampus. *Neurobiol Aging*. 2007;28:845–855.

*Maccarrone M, Guzman M, Mackie K, Doherty P, Harkany T. Programming of neural cells by (endo)cannabinoids: from physiological rules to emerging therapies. *Nat Rev Neurosci*. 2014;15:786–801.

Mazier W, Saucisse N, Gatta-Cherifi B, Cota D. The endocannabinoid system: Pivotal orchestrator of obesity and metabolic disease. *Trends Endocrinol Metab*. 2015;26:524–537.

*McEwen BS, Bowles NP, Gray JD, et al. Mechanisms of stress in the brain. *Nat Neurosci*. 2015;18:1353–1363.

Mifflin K, Benson C, Kerr B, et al. Involvement of neuroactive steroids in pain, depression and anxiety. *Mod Trends Pharmacopsychiatri*. 2015;30:94–102.

*Mischoulon D, Freeman MP. Omega-3 fatty acids in psychiatry. *Psychiatr Clin North Am*. 2013;36:15–23.

Moncada S, Bolanos JP. Nitric oxide, cell bioenergetics and neurodegeneration. *J Neurochem*. 2006;97:1676–1689.

Nemets H, Nemets B, Apter A, Bracha Z, Belmaker RH. Omega-3 treatment of childhood depression: a controlled, double-blind pilot study. *Am J Psychiatry*. 2006;163:1098–1100.

Paul BD, Snyder SH. H2 S: A Novel Gasotransmitter that Signals by Sulfhydration. *Trends Biochem Sci*. 2015;40:687–700.

Pi-Sunyer FX, Aronne LJ, Heshmati HM, Devin J, Rosenstock J. Effect of rimonabant, a

cannabinoid-1 receptor blocker, on weight and cardiometabolic risk factors in overweight or obese patients: RIO-North America: a randomized controlled trial. *JAMA*. 2006;295:761–775.

Porter J, Van Vrancken M, Corll C, Thompson H, Svec F. The influence of dehydroepiandrosterone and 8-OH-DPAT on the caloric intake and hypothalamic neurotransmitters of lean and obese Zucker rats. *Am J Physiol Regul Integr Comp Physiol*. 2005;288:R928–R935.

Renard J, Krebs MO, Le-Pen G, Jay TM. Long-term consequences of adolescent cannabinoid exposure in adult psychopathology. *Frontiers Neuroscience*. 2014;8:361.

Richardson AJ, Montgomery P. The Oxford-Durham study: a randomized, controlled trial of dietary supplementation with fatty acids in children with developmental coordination disorder. *Pediatrics*. 2005;115:1360–1366.

Rogers PJ, Appleton KM, Kessler D, et al. No effect of n-3 long-chain polyunsaturated fatty acid (EPA and DHA) supplementation on depressed mood and cognitive function: a randomised controlled trial. *Br J Nutr*. 2008;99:421–431.

Rubino T, Parolaro D. The impact of exposure to cannabinoids in adolescence: Insights from animal models. *Biol Psychiatry*. 2016;79(7):578–585.

Schule C, Romeo E, Uzunov DP, et al. Influence of mirtazapine on plasma concentrations of neuroactive steroids in major depression and on 3alpha-hydroxysteroid dehydrogenase activity. *Mol Psychiatry*. 2006;11:261–272.

Sedlak TW, Snyder SH. Messenger molecules and cell death: Therapeutic implications. *JAMA*. 2006;295:81–89.

Song C, Shieh CH, Wu YS, Kalueff A, Gaikwad S, Su KP. The role of omega-3 polyunsaturated fatty acids eicosapentaenoic and docosahexaenoic acids in the treatment of major depression and Alzheimer's disease: acting separately or synergistically? *Prog Lipid Res*. 2016;62:41–54.

Strous RD, Maayan R, Weizman A. The relevance of neurosteroids to clinical psychiatry: from the laboratory to the bedside. *Eur Neuropsychopharmacol*. 2006;16:155–169.

Su KP, Huang SY, Chiu TH, et al. Omega-3 fatty acids for major depressive disorder during pregnancy: results from a randomized, double-blind, placebo-controlled trial. *J Clin Psychiatry*. 2008;69:644–651.

Sulaieva O, Wallace JL. Gaseous mediator-based anti-inflammatory drugs. *Curr Opin Pharmacol*. 2015;25:1–6.

Vincent SR. Nitric oxide neurons and neurotransmission. *Prog Neurobiol*. 2010;90:246–255.

*Volk DW, Lewis DA. The role of endocannabinoid signaling in cortical inhibitory neuron dysfunction in schizophrenia. *Biol Psychiatry*. 2016;79(7):595–603.

▲ 1.9 Intraneuronal Signaling

JOHN A. GRAY, M.D., PH.D., AND BRYAN L. ROTH, M.D., PH.D.

OVERVIEW OF NEURONAL SIGNAL TRANSDUCTION

Signal transduction is the process by which a cell converts extracellular signals to intracellular signals and the subsequent cascade of events that

leads to alterations in cellular function. The initial step in signal transduction usually involves the binding of an extracellular signal, such as a neurotransmitter, to a designated plasma membrane receptor. Binding of this molecule, or ligand, to its receptor stabilizes a unique conformation of the receptor, resulting in the transmission of the signal across the plasma membrane. In its simplest form, the binding of a neurotransmitter to an ion channel stabilizes a conformation of the channel protein that allows the channel to open and ions to flow into or out of the cell, setting off a cascade of intracellular events. In more complex examples of signal transduction, the stabilized conformation of the receptor allows for the binding of other proteins to the intracellular portions of the receptor. These intracellular proteins subsequently become "activated" and go on to initiate various downstream events.

Before the details of each signaling pathway are discussed, it is useful to understand the common themes in the flow of information from receptor binding to the final alterations in neuronal function. In general, these systems are organized into several layers (Fig. 1.9–1). First, extracellular signals are detected by receptors and transmitted across the plasma membrane to adaptor proteins. These adaptor proteins then activate one or more intracellular signaling pathways, which, in turn, alter the function of effector proteins, either directly or via intermediates, such as protein kinases. A protein kinase is an enzyme that adds a phosphate group to a protein. Protein phosphorylation is a primary mechanism in signal transduction as phosphorylation changes a protein's conformation, which can alter its enzymatic activity or its ability to bind with other proteins. Typically, protein phosphorylation leads to the activation of a protein. The eventual outcome of these signaling pathways is the alteration of neuronal activity and changes in the expression of various genes.

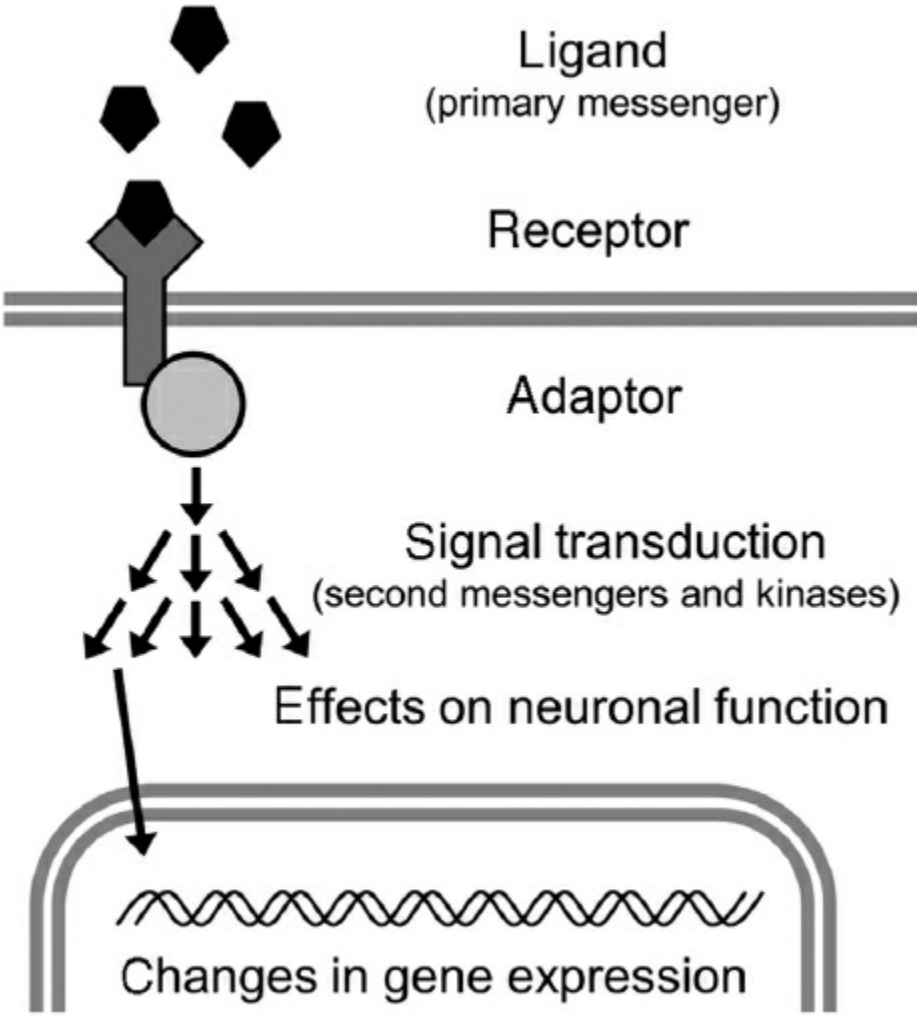

FIGURE 1.9–1. General outline of signal transduction. Transmission of an extracellular signal to an intracellular signal usually involves the binding of a ligand, such as a neurotransmitter, to a receptor. Conformational changes in the receptor protein transmits the signal across the plasma membrane to adaptor proteins which then activate one or more intracellular signaling pathways. These signaling pathways often involve the generation of small-molecule second messengers or the activation of kinase cascades leading to changes in neuronal function and gene expression.

So why do neurons and other cells have these complex intracellular signaling pathways? First, in addition to transmitting an extracellular signal across the plasma membrane, these signal transduction pathways amplify the signal exponentially, allowing for cells to have large responses to very minute quantities of extracellular stimuli. Furthermore, the multiplicity of intracellular signaling pathways allows signals to be directed in a specific manner, thus enabling cells to maintain separate channels of information that can be integrated only when appropriate. For example, these separate channels of information allow neurons to detect when different stimuli are presented concurrently, thus enabling them to alter their response accordingly. Additionally, each signaling pathway has distinctive spatial and temporal characteristics that allow for the optimal handling of different types of information. For example, in some instances

it may be advantageous for a neuron to have extremely high sensitivity to an uncommon stimulus but to ignore repetitive inputs. Thus, these complex signaling pathways determine a neuron's sensitivity and responsiveness to its environment in the context of its current circumstances and its past experiences.

Overall, a detailed understanding of the complex biochemical processes operating inside neurons is critical to appreciate how the brain not only responds to individual stimuli but also how it can continuously adapt to endless environmental changes. In addition, advances in the understanding of the molecular processes occurring within neurons will lead to improved insight into the basis of behavior and psychotropic drug action and will likely guide the development of improved psychiatric treatments and diagnostic tools.

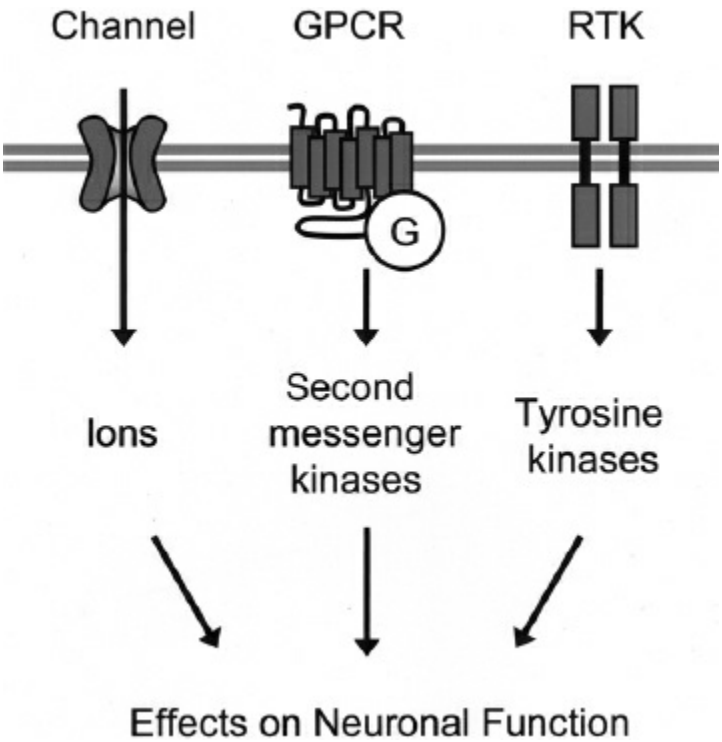

FIGURE 1.9–2. Outline of the three major receptor types mediating intraneuronal signal transduction in neurons: ligand-gated ion channels, G-protein–coupled receptors (GPCRs), and receptor tyrosine kinases (RTKs). Each receptor type, when activated by its extracellular signal, induces intracellular signaling pathways that result in alterations of neuronal function. G, heterotrimeric G protein.

Major Neuronal Signaling Pathways

There are three main schemes of signal transduction in neurons (Fig. 1.9–2). The first, which will be discussed in more detail in Section 1.10, involves ligand-gated ion channels. Ligand-gated ion channels are the primary mechanism of signal transduction for amino acid neurotransmitters such as glutamate and GABA. In addition, a number of other neurotransmitters, including acetylcholine and serotonin, have a subset of receptors that are

ligand-gated ion channels. Upon binding of neurotransmitter, these ion channels are stabilized in a conformation that alters the conductance of the channel to particular ions, usually Na^+, K^+, Ca^{2+}, or Cl^-. At synapses, these receptors can very quickly convert an extracellular signal into a postsynaptic electrical signal constituting so-called "fast synaptic transmission." A prototypical example of a ligand-gated ion channel is the nicotinic acetylcholine receptor, which consists of five protein subunits with two binding sites for acetylcholine enclosing a central aqueous pore. When both acetylcholine binding sites are filled, the internal pore of the channel opens, allowing Na^+ ions to flow into the cell, down their electrochemical gradient. Channels that allow Ca^{2+} to enter the cell can have effects on the electrical properties of the cell and can stimulate additional calcium-mediated intracellular signaling cascades as discussed below.

A second primary scheme of neuronal signal transduction involves the binding of a neurotransmitter to seven-transmembrane-domain receptors. These receptors are also known as GPCRs because they activate heterotrimeric guanine nucleotide-binding proteins (G proteins). GPCRs represent the single largest family of receptors (with more than 700 members) in the genome and are the primary form of receptor for many of the neurotransmitters, including serotonin and dopamine. GPCRs primarily signal by activating G proteins that subsequently activate effector enzymes that generate small molecules termed "second messengers" (the "first messenger" being the extracellular signal itself). Second messengers in turn mediate many of the downstream intracellular signaling cascades, largely involving protein kinases. Because of the additional step of creating small-molecule second messengers, signaling through GPCRs generally requires more time to develop than the opening of a ligand-gated ion channel, and thus this scheme accounts for the majority of "slow synaptic transmission" in neurons.

The third common scheme of signal transduction in the brain involves the activation of a distinct class of protein kinases that phosphorylate proteins on tyrosine residues. Activation of these protein tyrosine kinases is the primary signaling pathway for most neurotrophic factors, such as NGF and BDNF, as well as various cytokines and chemokines. The binding of a neurotrophic factor or chemokine to its respective plasma membrane receptor leads to the dimerization of the receptor with another copy of the receptor transmitting the signal across the plasma membrane. This results in the activation of a cascade of protein kinases. In some cases, the intracellular portion of the receptor itself contains the first tyrosine kinase in the cascade, and in other cases the dimerized receptors recruit cytoplasmic tyrosine kinases that then become activated. While the cascades of kinases can be complicated, the ultimate purpose is to amplify the initial signal and affect numerous changes in neuronal function.

G-PROTEIN–COUPLED RECEPTOR SIGNALING

Accounting for at least 2 percent of the genes in the genome, GPCRs comprise a very large family of proteins that represent targets for a wide array of molecules ranging from hormones and neurotransmitters to odorants and even light. The binding of an agonist to its cognate GPCR stabilizes an active conformation of the receptor, inducing the dissociation and activation of a receptor-specific heterotrimeric G protein into its α- and βγ-subunits (Fig. 1.9–3). G proteins were discovered by Alfred Gilman, Martin Rodbell, and colleagues and are a family of proteins named because they use the exchange of guanosine diphosphate (GDP) and guanosine triphosphate (GTP) as a molecular "switch" to regulate cell processes. In their baseline state, G proteins are heterotrimeric proteins consisting of α-, β-, and γ-subunits, with a GDP molecule bound to the α-subunit. When a GPCR is activated by an agonist, the G-protein α-subunit becomes associated with the receptor, causing conformational changes that release the GDP molecule. This allows a GTP molecule to bind, thus activating the α-subunit. GTP binding to the α-subunit also causes the dissociation of the α-subunit from the βγ-subunits and from the receptor. These dissociated subunits are now biologically active and activate or inhibit a number of downstream effectors, such as nucleotide cyclases, phospholipases, and kinases, resulting in a variety of downstream cellular effects. The system is returned to its basal state when the α-subunit, which has intrinsic GTPase activity, hydrolyzes the GTP back to GDP. This hydrolysis of GTP to GDP within the α-subunit also leads to the reassociation of the α-subunit with the βγ-subunits and thus restoration of the inactive heterotrimer. If the receptor remains bound to its agonist, then the GDP can again dissociate from the α-subunit and another G-protein cycle begins; however, if the receptor becomes inactive, the hydrolysis of GTP to GDP halts intracellular signaling.

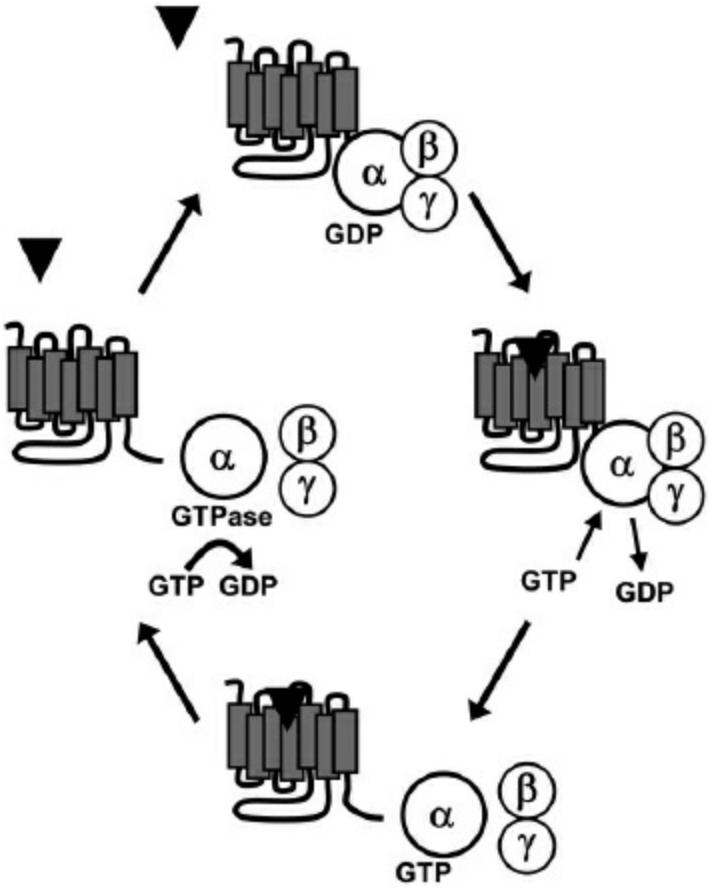

FIGURE 1.9–3. Outline of G-protein function. Under basal conditions, G proteins exist as heterotrimers consisting of single α-, β-, and γ-subunits, with the inactive α-subunits bound to GDP. After the G-protein–coupled receptor is activated by its ligand, it causes the associated G protein to release GDP, allowing GTP to bind. GTP binding to the α-subunit causes the dissociation of the α-subunit from the βγ-subunits and from the receptor. The free G-protein subunits are functionally active and can activate and regulate a variety of downstream effector proteins. The α-subunit has intrinsic GTPase activity that hydrolyzes the GTP back to GDP, causing the reassociation of the α- and βγ-subunit, restoring the system to its basal state.

The primary effector systems mediated by G proteins share a common overall form, whereby the activated G protein activates an effector enzyme that generates small-molecule second messengers that then initiate specific protein kinase cascades. The particular class of G proteins associated with each individual receptor dictates which specific second messenger system is activated. The four major classes of G proteins that are involved in neurotransmitter-mediated signaling are called G_s, G_i, G_q, and $G_{12/13}$ and are discussed below with the details of each second messenger system. Briefly, both the G_s and G_i classes of G proteins regulate the cAMP system, while the G_q class regulates the phosphoinositide signaling system. The $G_{12/13}$ class modulates Rho/Rac signaling cascades but will not be discussed. Importantly, a single receptor can activate multiple G proteins, which can then activate multiple effector enzymes that synthesize many second messenger molecules, leading to an exponential amplification of the

initial signal.

Cyclic Adenosine Monophosphate Pathway

The discovery of cAMP in the 1950s by Earl Sutherland and Theodore Rall established the concept that small intracellular molecules can act as second messengers that convey information from cell-surface receptors to their targets within the cell. Since that seminal discovery, decades of intensive research on the cAMP system has elucidated its operating principles, making it the prototypical second messenger system.

GPCRs that activate the cAMP signaling pathway are coupled to the G_s class of G proteins. When G_s becomes activated by a receptor, it dissociates from the receptor and stimulates a membrane-bound effector enzyme called adenylate cyclase that converts ATP to cAMP (Fig. 1.9–4). Conversely, other receptors are coupled to the G_i class of G proteins that inhibit adenylate cyclase and thus decrease cAMP production. The net level of cAMP production by a given neurotransmitter is thus determined by the specific G_s- and G_i-coupled receptor subtypes expressed on a given neuron or synapse. For example, norepinephrine stimulates adenylate cyclase via its interaction with β-adrenergic receptors and inhibits adenylate cyclase via stimulation of α_2-adrenergic receptors.

The major target of action of cAMP in most cells is the cAMP-dependent protein kinase, also known as protein kinase A (PKA), which mediates many of the actions cAMP has on neuronal function. PKA is a multisubunit serine/threonine kinase consisting of two regulatory subunits and two catalytic subunits. In the absence of cAMP, the regulatory subunits are bound to the catalytic subunits, thus keeping the kinase inactive. However, when cAMP is present, it binds to the regulatory subunits, thus causing a conformation change that dissociates the regulatory subunits from the catalytic subunits. The release of the regulatory subunits activates the catalytic subunits, which are then free to phosphorylate various cellular proteins on serine and threonine residues. This kinase has a broad range of substrate proteins involved in regulating virtually every aspect of neuronal function. In particular, several important neuronal targets for PKA have been identified, including various ion channels, synaptic vesicle machinery, neurotransmitter synthetic enzymes, and proteins involved in regulating gene transcription. Thus, alterations in cAMP levels are able to affect neuronal function over a broad range of time scales. For example, rapid effects are achieved by targeting ion channel gating and neurotransmitter release machinery, while slower effects occur with the targeting of neurotransmitter synthesis and cellular energy metabolism. Furthermore, cAMP elicits longer-lasting changes in neuronal function by controlling the expression of specific target genes.

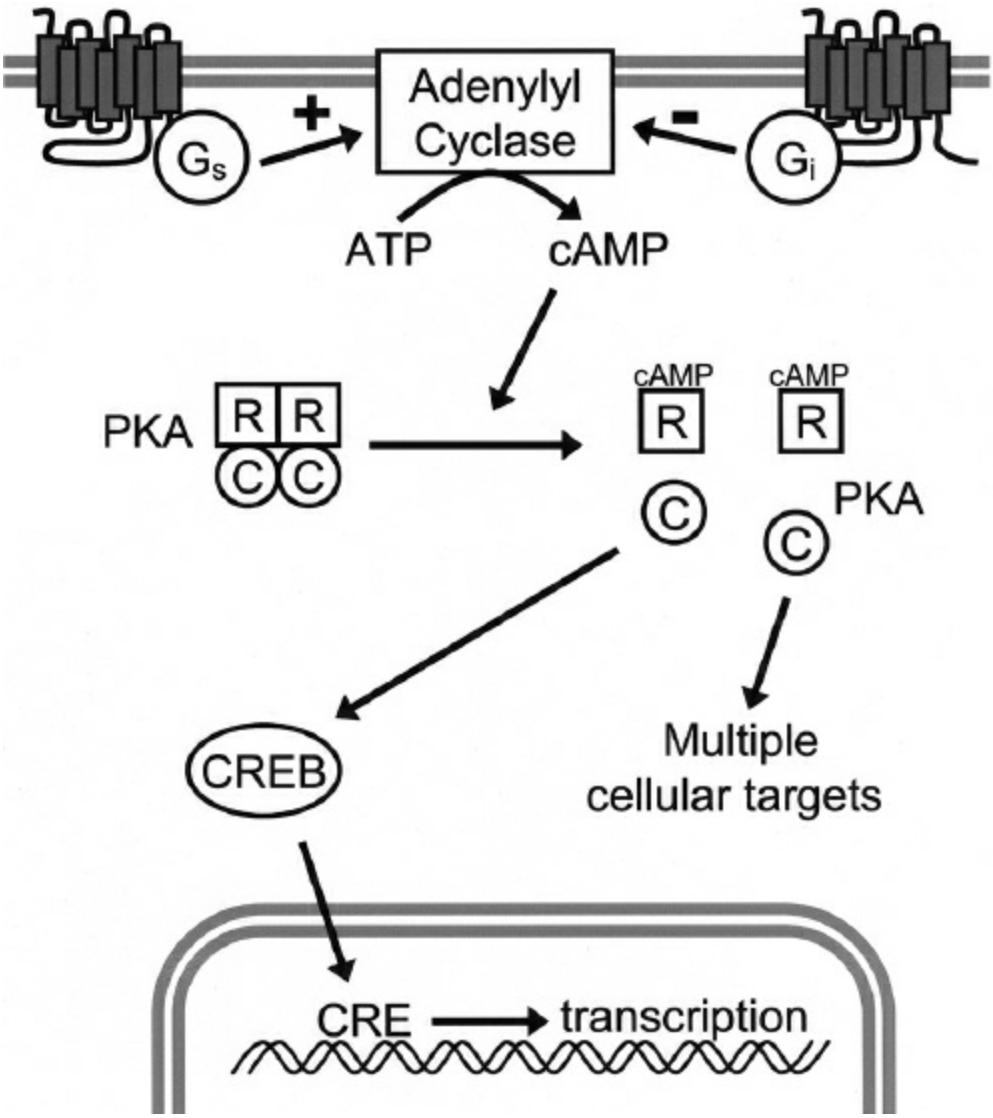

FIGURE 1.9–4. Basic organization of the cyclic adenosine monophosphate (cAMP) signaling pathway. Production of cAMP from ATP by adenylate cyclase can either be stimulated or inhibited by G-protein–coupled receptors. Receptors coupled to the G-protein G_s stimulate cAMP synthesis whereas those coupled to G_i inhibit adenylate cyclase. Many of the cellular actions mediated by cAMP occur through protein kinase A (PKA), which exists at baseline as a tetramer of two regulatory subunits (R) that tonically inhibit two catalytic subunits (C). When the R subunits become bound by cAMP, they dissociate from the C subunits, which become activated and can phosphorylate multiple cellular proteins. One such target is the transcription factor cAMP response element binding (CREB) protein, which, when activated by PKA, can bind to DNA sequences called cAMP response elements (CREs) and promote gene transcription.

One important substrate of PKA is a transcription factor that enables elevations in cAMP to regulate gene expression. This transcription factor, called CREB protein, regulates the expression of various genes by binding to short deoxyribonucleic acid (DNA) sequences called cAMP response elements (CREs). Phosphorylation of CREB by PKA activates it, thus allowing it to bind to CRE sequences in the regulatory regions of target genes where it increases or decreases the transcription of certain genes.

Proteins whose expression is regulated by cAMP through CREB are thought to be involved in various neuronal processes, including neuronal development and survival and the formation of long-term memories.

Termination of cAMP signaling is mediated by the actions of phosphodiesterases (PDEs), enzymes that cleave cAMP to AMP. There are multiple isoforms of PDEs expressed throughout the brain that are differentially regulated, adding a level of complexity to the precise cellular control of cAMP signaling. At high concentrations, caffeine nonselectively inhibits phosphodiesterases, possibly contributing to some of its pharmacologic effects. In addition, considerable efforts have been made to develop PDE inhibitors that are selective for individual isoforms. For example, rolipram, an inhibitor selective for type IV PDEs, initially showed promise as an antidepressant but was limited due to side effects. In addition, PDE10A is a recently identified isoform expressed at high levels in the brain, and PDE10A inhibitors have been shown to antagonize the effects of both amphetamine and phencyclidine in rodents, thus suggesting antipsychotic potential.

Phosphatidylinositol Pathway

After the discovery of the cAMP system, it became apparent that there were many neurotransmitter receptors that did not act via cAMP, suggesting the possible existence of other second messenger systems. Beginning in the 1950s, there were hints that phosphoinositides (PIs) may be involved in various cellular pathways, though a coherent view of this second messenger system did not emerge until the early 1980s. The phosphatidylinositol signaling pathway parallels many basic aspects of the cAMP system (Fig. 1.9–5), though it also includes several unique features. The phosphatidylinositol signaling pathway is initiated by receptors that are coupled to the G_q class of G proteins, which activate the effector enzyme PLC. This enzyme cleaves phosphatidylinositol bisphosphate (PIP_2), an inositol-containing phospholipid located in the cytoplasmic leaf of the plasma membrane, into two second messengers, DAG and inositol trisphosphate (IP_3). These two second messengers can then go on to affect distinct cellular pathways.

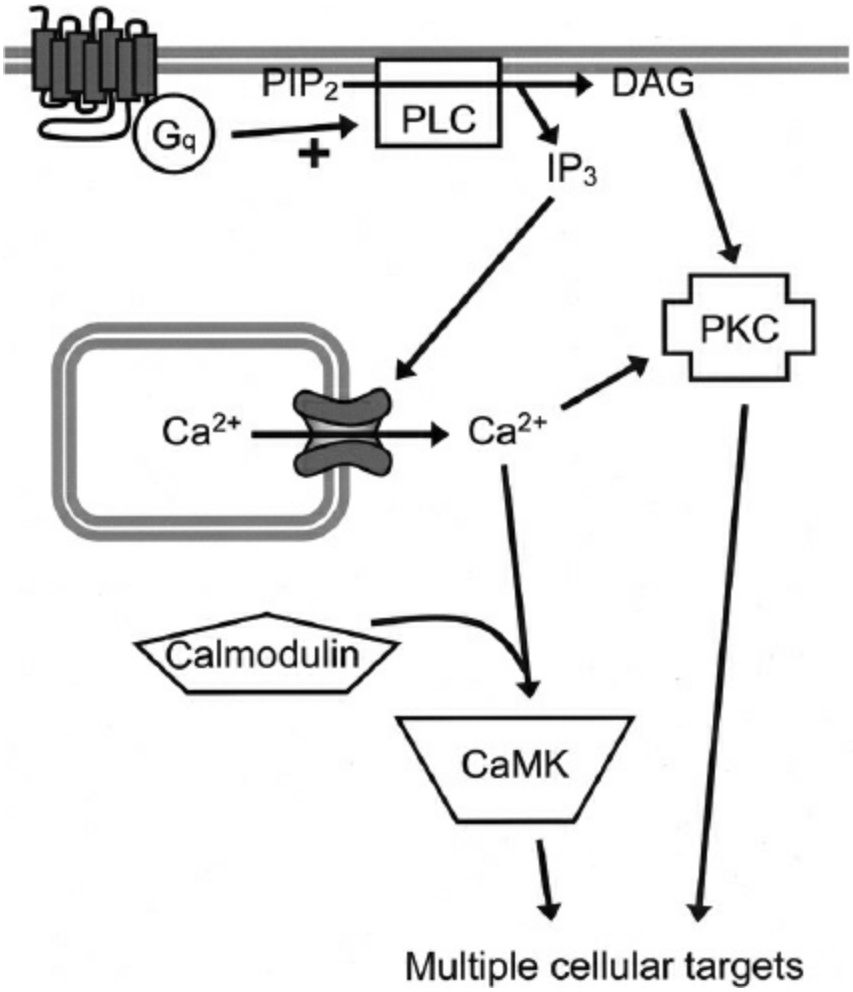

FIGURE 1.9–5. Basic organization of the phosphatidylinositol signaling pathway. The G_q class of G proteins activates phospholipase C (PLC), which cleaves the membrane phospholipid phosphatidylinositol bisphosphate (PIP_2) into two second messengers, diacylglycerol (DAG) and inositol trisphosphate (IP_3). IP_3 diffuses to the endoplasmic reticulum where it binds to the inositol trisphosphate receptor, rapidly releasing large stores of Ca^{2+}, which also functions as a second messenger. DAG activates protein kinase C (PKC), though some isoforms also require Ca^{2+} to become activated. The released Ca^{2+} mediates many of its functions by binding to calmodulin, which can then activate multiple cellular targets, including calcium/calmodulin-dependent kinase (CaMK). PKC and CaMK can phosphorylate multiple cellular targets, including various transcription factors.

DAG, which is hydrophobic, remains in the plasma membrane where it activates various isoforms of PKC, a serine/threonine protein kinase. In its inactive state, PKC is found in the cytoplasm, but when DAG is generated, PKC relocates to the plasma membrane, becomes activated, and phosphorylates multiple cellular substrates. The water-soluble IP_3 second messenger is released from the plasma membrane and diffuses to the endoplasmic reticulum where it binds to the inositol trisphosphate receptor. This receptor is a ligand-gated ion channel that, when bound by IP_3, rapidly releases the large stores of Ca^{2+} from the endoplasmic reticulum. This released Ca^{2+} also functions as a second messenger,

regulating various cellular functions. In addition to DAG, some isoforms of PKC also require Ca^{2+} to become activated.

The rapid rise in intracellular Ca^{2+}, mediated by IP_3-induced release of cellular stores, has both immediate and more delayed effects on neuronal functioning. Immediate effects are triggered by direct binding of Ca^{2+} itself to various effector proteins and include the release of synaptic vesicles and the opening of calcium-activated ion channels in the plasma membrane. Delayed effects of Ca^{2+} signaling are similar to those of cAMP, such as effects on cellular energy metabolism and gene expression. Many of these slower effects of calcium signaling are mediated by the association of calcium with calmodulin, a small, ubiquitous, calcium-binding protein. Calmodulin is activated when the intracellular concentration of Ca^{2+} is high enough for four ions to bind to the calmodulin protein. Activated calmodulin has multiple cellular targets, including activating various kinases, such as the calcium/calmodulin-dependent (CaM) kinases. It is important to note that calcium-based signaling pathways can also be activated by an influx of Ca^{2+} from the cell surface by various voltage- and ligand-gated ion channels, independent of G-protein signaling.

Termination of the phosphatidylinositol pathway involves multiple steps. DAG is degraded by lipases into glycerol and fatty acids or recycled into membrane phospholipids. Ca^{2+} is rapidly cleared from the cytoplasm by Ca^{2+}–ATPase pumps on the plasma membrane and endoplasmic reticulum, the action of which is enhanced by Ca^{2+} itself through the interaction of activated calmodulin with the transport pump. IP_3 is sequentially dephosphorylated by inositol phosphatases to inositol, which can then be reintegrated into membrane phospholipids. Interestingly, lithium is an inhibitor of these inositol phosphatases and leads to an accumulation of IP_3 and other inositol phosphates within cells. This leads to a depletion of the free cellular inositol needed to replenish membrane PIP_2 for further signaling, prompting the hypothesis that the rundown of the phosphatidylinositol cycle may underlie lithium's therapeutic action, though this remains controversial. Indeed, lithium is also known to inhibit several adenylate cyclases and protein kinases.

Other Second Messenger Systems

In addition to cAMP, another cyclic nucleotide, cGMP, is a second messenger that is regulated by neurotransmitter receptor stimulation. However, there are significant differences between the two systems. Guanylate cyclases are primarily cytoplasmic enzymes that are not directly activated by G proteins but are activated by the gas nitric oxide. Nitric oxide is synthesized in cells by NOS, which is activated by calmodulin and is thus mediated by increases in intracellular Ca^{2+} levels (Fig. 1.9–6). This

demonstration that a gas can act as a second messenger blurs the distinction between extracellular and intracellular messengers, as nitric oxide is capable of diffusing across cell membranes and at synapses may act as a retrograde signal to the presynaptic neuron. Synthesis of cGMP also leads to various downstream effects, many through its activation of protein kinase G. Like cAMP, cGMP is degraded by various PDEs. Indeed, drugs for erectile dysfunction, such as sildenafil (Viagra), act by selectively inhibiting a PDE isoform that is highly enriched in the vascular smooth muscle of the penis, highlighting the utility of developing drugs that target intracellular signaling pathways.

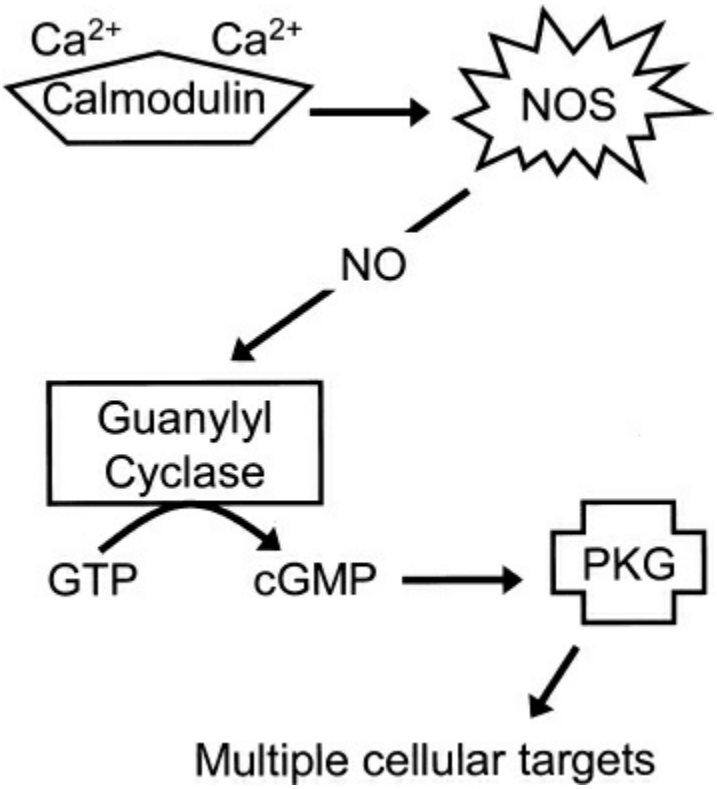

FIGURE 1.9–6. The cyclic guanosine monophosphate (cGMP) signaling pathway. In contrast to the cyclic adenosine monophosphate (cAMP) system, cGMP synthesis via guanylate cyclase is not regulated by G proteins. Instead, guanylate cyclase is activated by nitric oxide, which is synthesized by nitric oxide synthase (NOS) after it is activated by a calcium/calmodulin complex. Like cAMP, cGMP affects neuronal function by stimulating its cognate kinase, protein kinase G (PKG).

Another intracellular signaling system that appears to play an important role in neuronal function involves metabolites of the fatty acid arachidonic acid. Various receptors activate an enzyme called phospholipase A_2, possibly through an unidentified G protein or elevations in cytoplasmic calcium levels. Phospholipase A_2 cleaves membrane phospholipids, typically PIP_2, releasing free arachidonic acid, which is rapidly converted to a number of active metabolites (Fig. 1.9–7). For example, arachidonic acid may be cleaved by cyclooxygenase to yield, after

multiple enzymatic steps, several types of prostaglandins and thromboxanes. Alternatively, arachidonic acid may be cleaved by lipoxygenases to yield the leukotrienes. These active metabolites then regulate many intracellular functions, including ion channels and protein kinases, and are important for modulating signaling through other pathways by regulating adenylate and guanylate cyclases. Additionally, as these compounds are lipophilic, they can diffuse out of the neuron and act as ligands for their own GPCRs on other neurons. Interestingly, cyclooxygenase inhibitors have been hypothesized to improve cognitive performance in schizophrenia, possibly by reducing inflammatory processes in the brain.

Direct Regulation of Ion Channels by G Proteins

As discussed, the primary function of G proteins is to initiate second messenger signaling cascades that go on to affect a multitude of cellular functions including modulating ion channel gating through phosphorylation by kinases such as PKA, PKC, and CaM kinases. In addition, second messengers such as cAMP and cGMP can regulate ion channels directly. However, it is now clear that the G proteins themselves are also able to directly bind to and regulate ion channels independent of second messenger cascades. In particular, this process is best established for receptors that couple to the G_i family of G proteins, such as muscarinic acetylcholinergic, α_2-adrenergic, D_2-dopaminergic, and 5-HT_{1A}-serotonergic receptors. As before, the activation of these G_i-coupled receptors causes the dissociation of the G-protein α- and βγ-subunits. While the α-subunit goes on to inhibit adenylate cyclase, the βγ-subunits bind directly to the cytoplasmic regions of two different ion channels, depending on the cell type. In some cells, βγ-subunits bind to and directly open specific K^+ channels known as G-protein–regulated inwardly rectifying K^+ (GIRK) channels. These channels are called inwardly rectifying because, if under no electrochemical gradient, they more readily pass current inward; however, under normal physiological circumstances, K^+ flow through GIRKs is primarily outward. In other cells, βγ-subunits directly inhibit voltage-gated Ca^{2+} channels, limiting the opening of these channels in response to membrane polarization. In addition to G_i, there is some evidence suggesting that certain members of the G_s family can increase the opening of certain voltage-gated Ca^{2+} channels, though it remains unclear if this is mediated by α- or βγ-subunits. Furthermore, recent evidence suggests that GIRKs may interact directly with GPCRs likely to promote near-instantaneous opening of the ion channel following G-protein activation. Overall, cellular signaling has evolved such that each of the major components of the signaling pathway can act on effector

targets themselves or recruit downstream components of the signaling cascade.

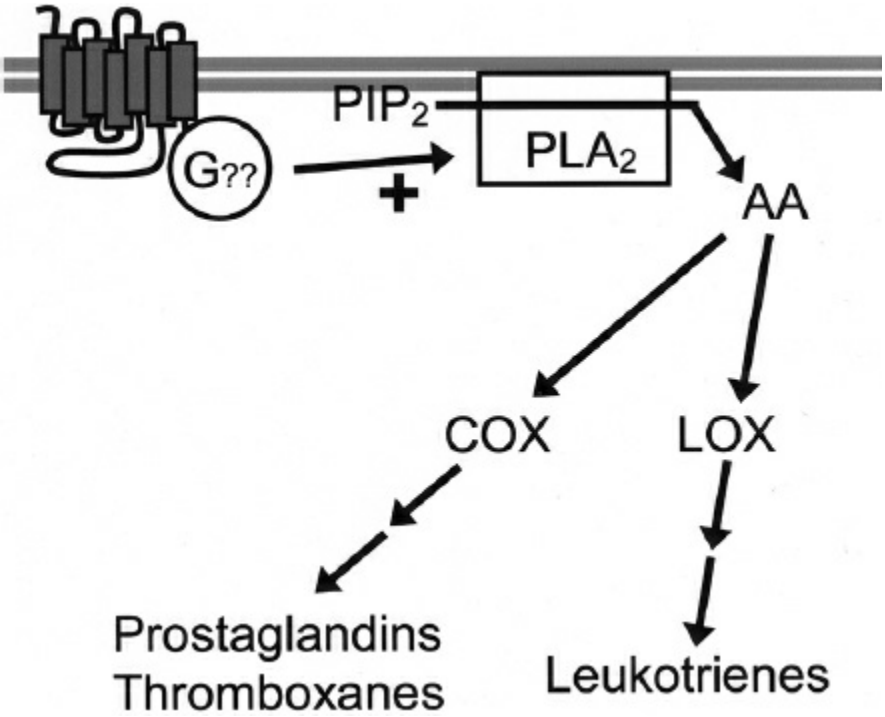

FIGURE 1.9–7. Organization of the arachidonic acid signaling pathway. Various G-protein–coupled receptors activate an enzyme called phospholipase A_2 (PLA_2), possibly through an unidentified G protein ($G_{??}$). PLA_2 primarily cleaves the membrane phospholipid phosphatidylinositol bisphosphate (PIP_2), releasing free arachidonic acid (AA), which is rapidly converted to a number of active metabolites. For example, arachidonic acid may be cleaved by cyclooxygenase (COX) to yield prostaglandins and thromboxanes and by lipoxygenases (LOXs) to yield leukotrienes. These AA metabolites can regulate many intracellular functions and can diffuse out of the neuron and act as ligands for their own G-protein–coupled receptors on other neurons.

Regulation of GPCR Signaling

Because GPCRs play such a key role in cellular signaling, it is not surprising that their activity is tightly regulated. Indeed, regulation occurs at almost every point along the signaling pathways. Above, the termination of second messenger signaling was discussed, such as with the degradation of cAMP and cGMP by PDEs. In this section, regulation of the G proteins and GPCRs will be briefly discussed.

The termination of G-protein signaling is mediated by the intrinsic GTPase activity of the α-subunit, which hydrolyzes GTP to GDP and thus inactivates the G protein. A separate class of proteins, called regulators of G-protein signaling (RGS) proteins, can accelerate the GTPase activity of the α-subunits thereby shortening the duration of G-protein signaling. There are more than 20 subtypes of RGS proteins that are differentially expressed throughout the brain and are involved in the regulation of all G-protein α-subunits, except for the G_s family. Specific subtypes of RGS

proteins have been shown to regulate important neuronal functions, including behavior. For example, mice with a targeted deletion of the gene encoding RGS2, which regulates G_q, show increased anxiety and decreased male aggression, highlighting the important role of RGS proteins in modulating behavior. Interestingly, expression levels of RGS9, which inhibits G_i-mediated dopamine signaling in the striatum, were found to be reduced in the postmortem brains of patients with schizophrenia, consistent with the hypothesis of excessive dopamine signaling resulting in psychosis. Overall, the diversity and heterogeneous distributions of the RGS proteins make them an attractive target for drug development, as drugs affecting individual isoforms may exert highly selective effects.

In addition to initiating intracellular signal transduction cascades, agonist activation of GPCRs also triggers cellular and molecular mechanisms that lead to the attenuation (desensitization) of receptor signaling. Desensitization is an adaptive mechanism to attenuate receptor responsiveness to repetitive environmental stimuli. At the level of the whole organism, the mechanisms underlying receptor desensitization are likely responsible for the development of tolerance to psychopharmacological agents such as opiates as well as the delayed therapeutic response to antidepressants and antipsychotics.

Receptor desensitization is typically mediated by feedback phosphorylation of the receptor by a class of kinases called G protein–coupled receptor kinases (GRKs) (Fig. 1.9–8). The GRKs phosphorylate the intracellular domains of receptors only when the receptors are bound by an agonist. Receptor phosphorylation by GRKs enables a protein called arrestin to bind to the receptor, preventing the G protein from recoupling, thus rendering the receptor inactive. Arrestin binding to a GPCR also causes the receptor to be internalized into endocytic vesicles. This process is mediated by an interaction of the arrestin molecule with proteins in clathrin-coated pits, membrane invaginations that are pinched off during endocytosis. After internalization, receptors may be recycled back to the cell surface or degraded. Interestingly, mice with a targeted deletion of a particular arrestin isoform do not develop tolerance to the analgesic effects of morphine, suggesting that tolerance may be mediated by GRK and arrestin interactions with the opioid receptor. However, these “arrestin knockout” mice still develop morphine dependence, providing an elegant dissociation between these features of chronic morphine administration. While the precise role of receptor endocytosis is not entirely clear, targeting receptors for degradation may be involved in the downregulation of brain receptor level following chronic drug administration. Indeed, accumulating evidence suggests a high degree of specificity and plasticity in the regulation of GPCRs by endocytic membrane trafficking that may provide for the development of novel pharmacologic agents.

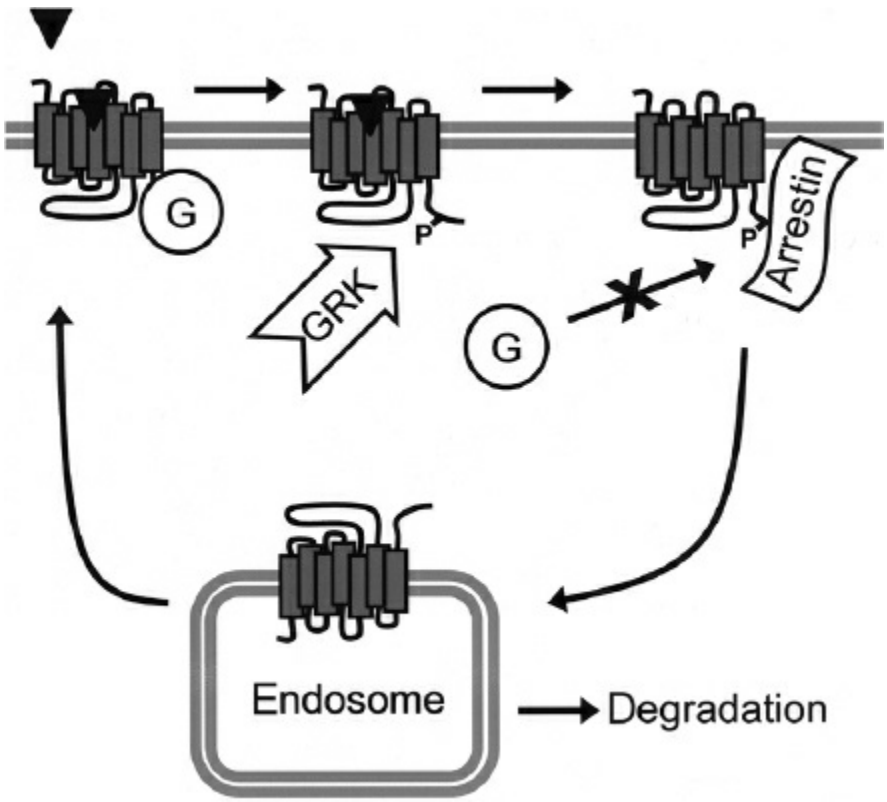

FIGURE 1.9–8. Regulation of G-protein–coupled receptors (GPCRs) by desensitization and internalization. Desensitization is an adaptive mechanism that attenuates receptor responsiveness to repetitive stimuli. GPCR desensitization is typically mediated by feedback phosphorylation (P) by a specific G-protein–coupled-receptor kinase (GRK). Receptor phosphorylation by a GRK causes a protein called arrestin to bind to the receptor, effectively preventing G proteins (G) from recoupling to the receptor. Arrestin binding also causes the receptor to be internalized into endocytic vesicles, which may then recycle the receptors back to the cell surface or target them for degradation.

Role of Phosphatases

Because phosphorylation plays such a central role in intracellular signaling pathways, it is not surprising that protein phosphatases, which reverse the effect of protein kinases, also have a major impact on these signaling pathways. There are four major protein phosphatases that are differentially distributed in the brain that dephosphorylate targets of the second messenger kinases, named protein phosphatases 1, 2A, 2B, and 2C. For example, protein phosphatase 2B, also called calcineurin, is activated by the binding Ca^{2+}/calmodulin. Thus, neurotransmitters coupled to G_q proteins as well as Ca^{2+} channels can activate calcineurin and influence the phosphorylation of various cellular proteins. Indeed, phosphatases can be targets for pharmacological agents, as demonstrated by tacrolimus, an immunosuppressant agent used to prevent organ transplant rejection, which is a selective inhibitor of calcineurin that interferes with T lymphocyte signaling.

Another key mechanism for regulating protein phosphatases involves a separate class of proteins called protein phosphatase inhibitors. These

proteins, such as phosphatase inhibitors 1 and 2, are highly potent inhibitors of protein phosphatase 1, a major neuronal phosphatase, and their inhibitory activity is greatly enhanced when they are phosphorylated by PKA and other second messenger kinases. Thus, neurotransmitters that signal through cAMP can influence the phosphorylation of target proteins, both by PKA activation and through PKA-induced indirect inhibition of protein phosphatase 1.

Another protein phosphatase inhibitor, called dopamine- and cAMP-regulated phosphoprotein of 32 kDa (DARPP-32), is of particular interest because it is highly concentrated in regions of the brain that receive dopaminergic input. Similar to other kinase events, phosphorylation of DARPP-32 by PKA greatly enhances its ability to inhibit protein phosphatase 1. Interestingly, DARPP-32 is dephosphorylated by calcineurin resulting in an increase in protein phosphatase 1 activity. Because calcineurin is activated by increases in intracellular Ca^{2+} concentrations, DARPP-32 may be involved in the integration of current signals from different pathways. Indeed, DARPP-32 is a key mediator of dopaminergic signaling and appears to play an important role in the effects of drugs of abuse. A share of the 2000 Nobel Prize for Physiology or Medicine was awarded to Paul Greengard for elucidating this role of DARPP-32.

TYROSINE KINASE PATHWAYS

While the vast majority of the protein phosphorylation that occurs in cells is on serine and threonine residues, the phosphorylation of tyrosine residues has an extremely important role in a distinct set of intracellular signaling pathways. In particular, tyrosine-phosphorylation–based signaling is mediated by receptors for neurotrophic factors such as NGF and BNDF. Functionally, neurotrophic factors modulate a wide variety of cellular events, such as cell growth and differentiation, metabolism, and cell survival, and thus have been classically studied for their role in neurodevelopment. However, these factors have been shown to be expressed throughout the entire life span, and exciting new research is delineating roles that neurotrophic factors play in regulating behavior and responses to stress.

Protein tyrosine kinases represent a diverse superfamily of proteins. These include the receptor tyrosine kinases, which are transmembrane receptors with a tyrosine kinase built into the intracellular domains, and nonreceptor tyrosine kinases, which are soluble cytoplasmic enzymes that are often recruited to membrane receptors to become activated. Neurotrophins such as NGF and BDNF bind to the Trk family of receptor tyrosine kinases that will be the primary focus of this section. Neurotrophins bind to two individual Trk receptors, resulting in the dimerization of the two receptors activating the protein tyrosine kinases

that reside in the cytoplasmic domain of each receptor (Fig. 1.9–9). The activated Trk receptors subsequently phosphorylate the opposite dimer on tyrosine residues, a process called autophosphorylation. These phosphorylation events produce new binding sites for various other intracellular signaling proteins. For example, an adaptor protein called growth-factor-receptor-bound protein 2 (Grb2) contains an Src homology 2 (SH2) domain that binds to specific phosphorylated tyrosine residues on Trk and leads off a complex signaling cascade.

As with neurotransmitter receptors, G proteins play a major role in the signal transduction from activated receptor tyrosine kinases. In this case, however, the G proteins are members of the Ras, Rho, and Ral families, collectively referred to as small G proteins. Like the classic G proteins described previously, small G proteins are bound to GDP in their inactive state and become active when GTP is bound. However, unlike the classic G proteins, the small G proteins are not directly activated by the receptor but by distinct proteins called guanine nucleotide exchange factors (GEFs). For Trk, binding of Grb2 to the phosphorylated tyrosine residues recruits a GEF protein called SOS (for Son of Sevenless), which then activates the small G-protein Ras by exchanging GTP for GDP. Once recruited to the receptor complex, SOS can activate many molecules of Ras, thus amplifying the initial signal. In its active GTP-bound form, Ras is able to activate the multiple downstream effector pathways described below. Ras is inactivated by hydrolysis of GTP to GDP, which can be accelerated by a specific GTPase-activating protein (GAP) called RasGAP, which is analogous to the action of RGS proteins on classic G proteins.

Mitogen-Activated Protein Kinase Cascades

In contrast to the classic second messenger signaling pathway, the activation of small G proteins such as Ras by receptor tyrosine kinases does not lead to the production of small-molecule intermediates. Instead, the small G proteins stimulate a signaling pathway that is organized as a kinase cascade in which a series of three or more kinases sequentially phosphorylate another kinase. Several parallel kinase cascades can be activated by various receptors and are collectively referred to as MAPK pathways. MAPKs are serine/threonine kinases making up three main classes characterized in mammals: The extracellular signal-regulated kinases (ERKs), the c-Jun N-terminal kinases (JNKs), and isoforms of p38. The ERK pathway is the classical MAPK pathway that is preferentially activated by the neurotrophins and other growth factors while the JNK and p38 pathways are activated by various forms of cellular stress.

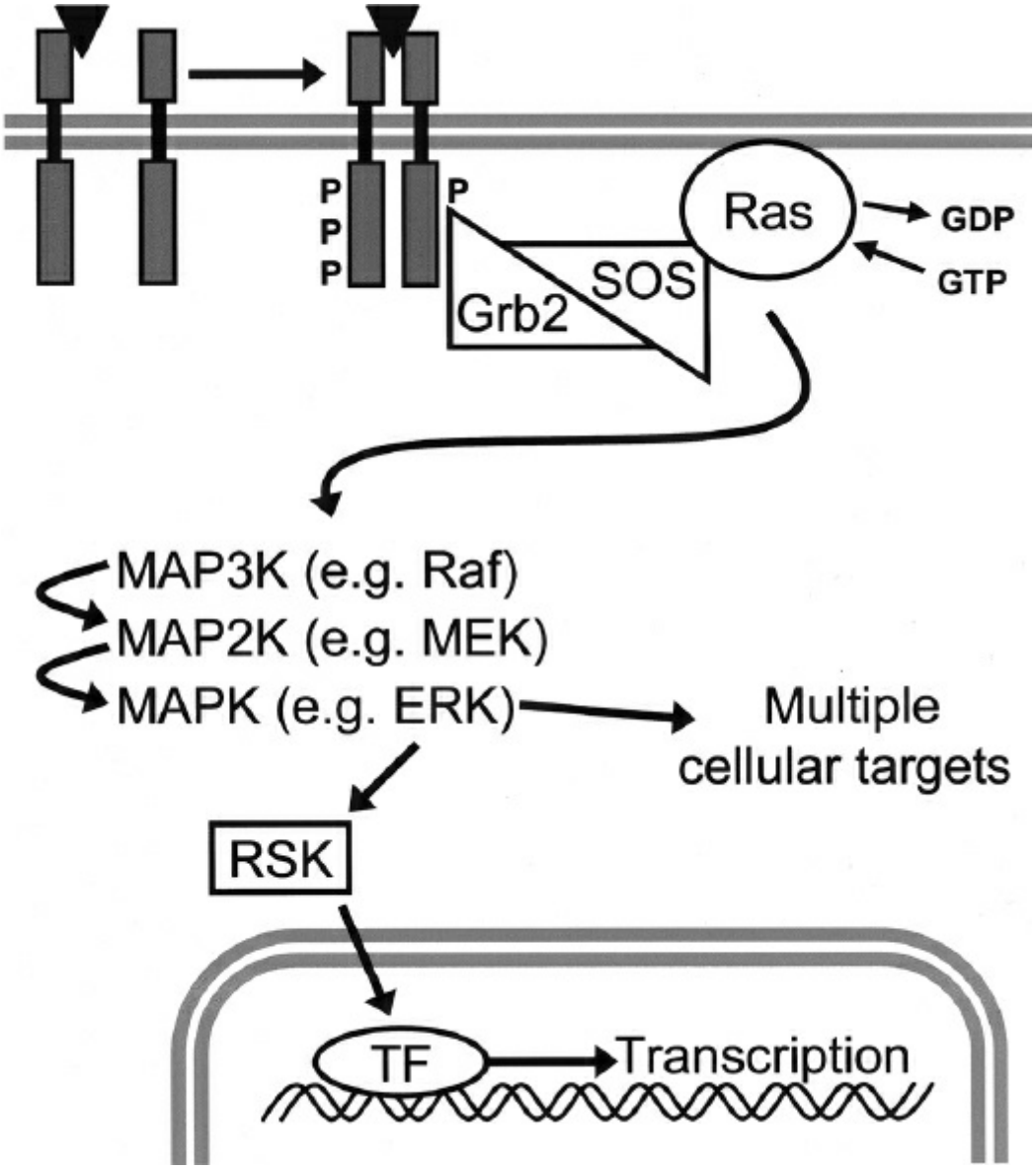

FIGURE 1.9–9. General organization of neurotrophic factor signaling through receptor tyrosine kinases. Receptor tyrosine kinases are transmembrane receptors with a tyrosine kinase built into the intracellular domains. Neurotrophic factor binding induces the dimerization of two receptors and the activation and autophosphorylation of their intrinsic tyrosine kinase domains. These phosphorylated (P) tyrosines become the binding sites for adaptor proteins such as growth-factor-receptor–bound protein 2 (Grb2), which can then attract a protein called Son of Sevenless (SOS) that activates the small G-protein Ras by enhancing the exchange of GTP for GDP. In its active GTP-bound form, Ras activates multiple downstream effector pathways, including mitogen-activated protein kinase (MAPK) cascades. A MAPK such as extracellular signal-regulated kinase (ERK) is activated by a MAPK kinase (MAP2K) such as the MAPK/ERK kinase (MEK), which is activated by a MAPK kinase kinase (MAP3K) such as Raf. After the cascade, ERK can activate various cellular targets including ribosomal S6 kinase (RSK), which translocates to the nucleus and activates various transcription factors (TFs) and regulates gene expression.

The kinase cascades that lead to MAPK activation follow an organization that is evolutionally well conserved from yeast to mammals: A MAPK kinase kinase kinase (MAP4 K) phosphorylates a MAPK kinase kinase (MAP3 K) that phosphorylates a MAPK kinase (MAP2 K) that then

phosphorylates the MAPK (Fig. 1.9–9). The cascade of events leading to the activation of ERKs by neurotrophins through the Trk receptor begins with the activation of the small G-protein Ras. When Ras is active, it recruits a MAP3 K called Raf to the cell surface where it is phosphorylated by a MAP4K that is not yet well described. Raf then phosphorylates and activates a MAP2K called MEK (for MAP kinase/ERK kinase) that then phosphorylates and activates ERK.

The ERK pathway is the subject of a considerable amount of current biomedical research, as it is involved in regulating a wide variety of cytoplasmic proteins as well as multiple transcription factors. For example, ERK phosphorylates and activates protein kinases such as ribosomal S6 kinase (RSK), which in turn phosphorylates an array of transcription factors including c-myc and CREB. Interestingly, stimulation of the ERK signaling pathway has also been linked to neurotransmitter receptors through PKC. Thus, ERK activation plays a key role in modulating long-term neuronal function and may represent a critical node for the interplay between other signaling pathways.

Phosphoinositide 3-Kinase Pathway

The elucidation of the signaling pathways downstream of Ras has identified another major kinase cascade that mediates many of the powerful effects of neurotrophins on neuronal differentiation and survival. This cascade involves the phosphoinositide 3-kinase (PI3K) pathway (Fig. 1.9–10). In this pathway, PIP_2, the same membrane phospholipid that is cleaved to DAG and IP_3 by PLC, is phosphorylated by PI3K, a lipid kinase, to yield PIP_3, which then acts to recruit various proteins to the membrane. One of the proteins that PIP_3 recruits to the membrane is Akt, a serine/threonine kinase that, upon translocation, becomes activated, dissociates from the membrane, and phosphorylates several substrate proteins important for controlling cell survival. For example, Akt activates the "rapid-acting" transcription factor nuclear factor-κB (NF-κB), resulting in the transcription of prosurvival genes. NF-κB is present in cells in an inactive state bound to a protein called inhibitor of κB (IκB). Akt activates a kinase called IκB kinase that phosphorylates IκB, thus tagging it for degradation. This degradation releases NF-κB, which can then migrate to the nucleus and regulate gene expression. Akt may also inhibit glycogen synthase kinase 3 (GSK-3), a metabolic regulatory protein that may be a cellular target for lithium (see below). Indeed, a major challenge of current research involves determining which of the many effects of neurotrophins are mediated by these various signaling cascades. Interestingly, some G_i-coupled neurotransmitter receptors can also trigger the activation of PI3K and Akt, suggesting that agonists of these receptors may represent novel strategies for enhancing neuronal survival.

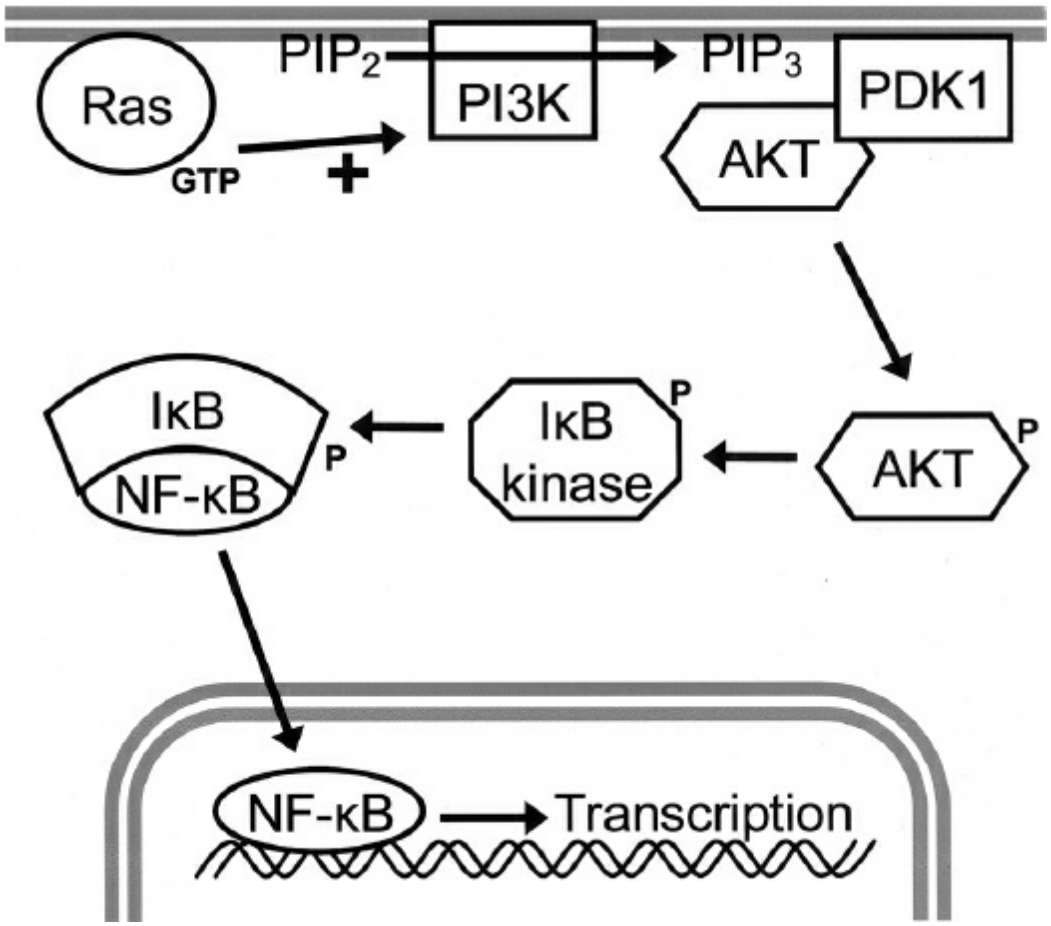

FIGURE 1.9–10. The phosphoinositide 3-kinase (PI3K) pathway. In addition to the mitogen-activated protein kinases (MAPKs), receptor tyrosine kinase activation of Ras can activate the PI3K pathway. PI3K adds another phosphate to the membrane phospholipid phosphatidylinositol bisphosphate (PIP_2) to yield phosphatidylinositol trisphosphate (PIP_3). PIP_3 recruits various proteins to the membrane including 3-phosphoinositide-dependent protein kinase 1 (PDK1) and a kinase called Akt. PDK1 phosphorylates Akt, which then dissociates from the membrane and can phosphorylate multiple cellular proteins important for controlling cell survival. For example, Akt can lead to the activation of a transcription factor called nuclear factor-κB (NF-κB), which is normally found in an inactive state bound to "inhibitor of κB" (IκB). Akt activates a kinase called IκB kinase, which phosphorylates IκB, tagging it for degradation, which releases NF-κB that can then translocate to the nucleus and regulate gene expression.

Mammalian Target of Rapamycin

A downstream target of the PI3K, and a central node of integration of multiple extracellular signals is an atypical kinase called mammalian target of rapamycin (mTOR). mTOR is a central controller of local protein synthesis that is required for the formation of new synaptic connections. mTOR function is influenced by the activity of each of the major classes of receptors including ligand-gated ion channels, GPCRs, and neurotrophin receptors suggesting that mTOR may act as a signaling "node" at the convergence of several signaling pathways. Indeed, mTOR activity is regulated by multiple signaling pathways, notably the PI3K-Akt pathways, though the details are not fully characterized.

Protein synthesis is a highly regulated process that can be separated into three general phases: initiation, elongation, and termination. mTOR is a master controller of protein synthesis by regulating both the initiation and elongation steps of protein synthesis (Fig. 1.9–11). One of the first mTOR targets to be discovered was 4E-BP (eukaryotic initiation factor 4E-binding protein) that, when phosphorylated by mTOR, releases eIF4E (eukaryotic initiation factor 4E) which can then initiate local protein translation by binding to the 7-methyl-guanosine cap of mRNAs. In addition, mTOR activates the ribosomal S6 kinases S6 K which then activates eEF2 K (eukaryotic elongation factor 2 kinase) facilitates protein synthesis.

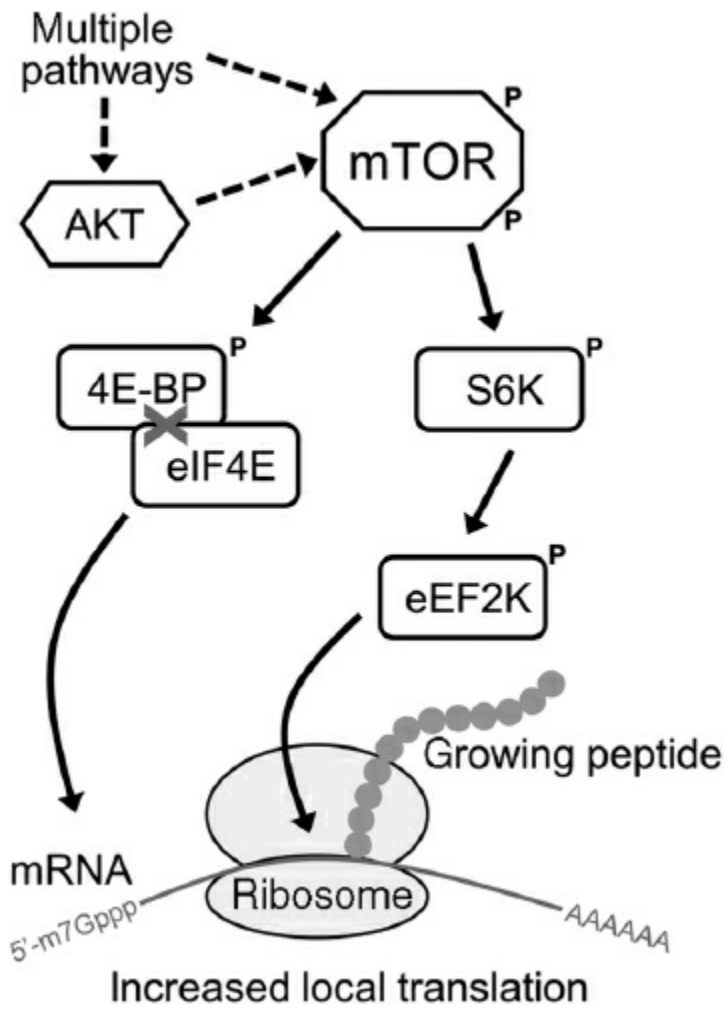

FIGURE 1.9–11. Regulation of protein translation by the mammalian target of rapamycin (mTOR). Research on mTOR regulation has suggested that mTOR is a key point of convergence of multiple signaling pathways, notably via Akt. mTOR is a central controller of local protein synthesis by regulating proteins involved in protein translation initiation and elongation at ribosomes. The activated mTOR complex phosphorylates 4E-BP (eukaryotic initiation factor 4E-binding protein) which dissociates it with eIF4E (eukaryotic initiation factor 4E). eIF4E can then facilitate translation initiation by binding mRNA at the 7-methyl-guanosine cap (5'-m7Gppp). mTOR also facilitates the elongation of nascent polypeptides by phosphorylating ribosomal S6 kinase (S6K) which can then activate the eukaryotic elongation factor 2 kinase (eEF2K).

Recently, there has been accumulating evidence linking mTOR signaling to long-lasting synaptic plasticity and various neuropsychiatric disorders. For example, the rapid antidepressant activity of NMDA receptor antagonists such as ketamine, may be mediated by the activation of the mTOR leading to increased translation of synaptic proteins and generation of new spine, particularly in the prefrontal cortex. In addition, genetic mutations that lead to pathologically sustained mTOR activity may underlie several neuropsychiatric disorders, including Fragile X syndrome, autism and tuberous sclerosis.

WNT SIGNALING

Another signaling pathway gaining interest in psychiatry and neurobiology is the Wnt signaling pathway. Wnts are a family of secreted glycoproteins known to play a critical role in embryogenesis. However, components of the Wnt signaling pathway are expressed in the adult brain, and Wnt signaling is important in adult behavior and possibly the pathophysiology of psychiatric and neurological disorders. The primary, or canonical, Wnt signaling pathway begins with binding of secreted Wnt proteins to cell-surface receptors of the Frizzled family. Frizzled receptors are seven-transmembrane-domain receptors similar to the GPCRs, though it remains unclear whether Frizzled interacts with a heterotrimeric G protein. It is clear, however, that Frizzled receptors activate a cytoplasmic protein called Dishevelled that ultimately leads to the regulation of gene expression through an increase in a transcriptional coactivator called β-catenin.

In the absence of Wnt signaling, a complex of proteins that includes axin, GSK-3, and the protein adenomatous polyposis coli (APC) regulates the intracellular levels of β-catenin (Fig. 1.9–12). Through phosphorylation by GSK-3, this protein complex promotes the proteolytic degradation of β-catenin. However, when Wnt signaling is initiated, the activation of Dishevelled causes this protein complex to dissociate and other proteins to associate and inhibit GSK-3 activity, preventing the degradation of β-catenin. Thus, the cytoplasmic levels of β-catenin increase, and β-catenin translocates to the nucleus, interacting with transcription factors and promoting specific gene expression changes. In addition to the canonical signaling pathway, Wnt signaling has also been shown to follow other pathways, including increasing intracellular calcium and activating the MAPK JNK. Thus, while the details of the Wnt signaling pathway are not fully delineated, there is likely some intersection and cross-regulation with other signaling pathways, and this is an area of active research.

Glycogen Synthase Kinase 3

Though initially discovered as a kinase involved in the regulation of glucose metabolism, GSK-3 is emerging as a promising target for the development

of psychiatric and neurological medications. In 1996, it was discovered that lithium inhibited GSK-3, raising the possibility that GSK-3 inhibition might play a role in the treatment of bipolar disorder. Recently, there has been an emergence of research supporting the hypothesis that the inhibition of GSK-3 represents a therapeutically relevant target for mood stabilization.

GSK-3 is a ubiquitous kinase, found in both neurons and glia, and has two isoforms that are highly homologous but may have slightly different biological effects. It is generally considered to be constitutively active, meaning that it phosphorylates target proteins until a signal regulates it to stop. For example, as described above, constitutive phosphorylation of β-catenin by GSK-3 leads to its proteolytic degradation, but signaling through the Wnt pathway turns off GSK-3, releasing β-catenin to affect gene expression. GSK-3 was initially characterized in 1980 as an enzyme that phosphorylated and deactivated glycogen synthase, leading to studies of its role in insulin signaling and diabetes mellitus. Indeed, insulin binding to the insulin receptor, a receptor tyrosine kinase, leads to the activation of Akt, which also phosphorylates and deactivates GSK-3. Because GSK-3 normally phosphorylates and inactivates glycogen synthase, insulin's ability to turn off GSK-3 allows cells to utilize the elevated plasma glucose levels to make glycogen.

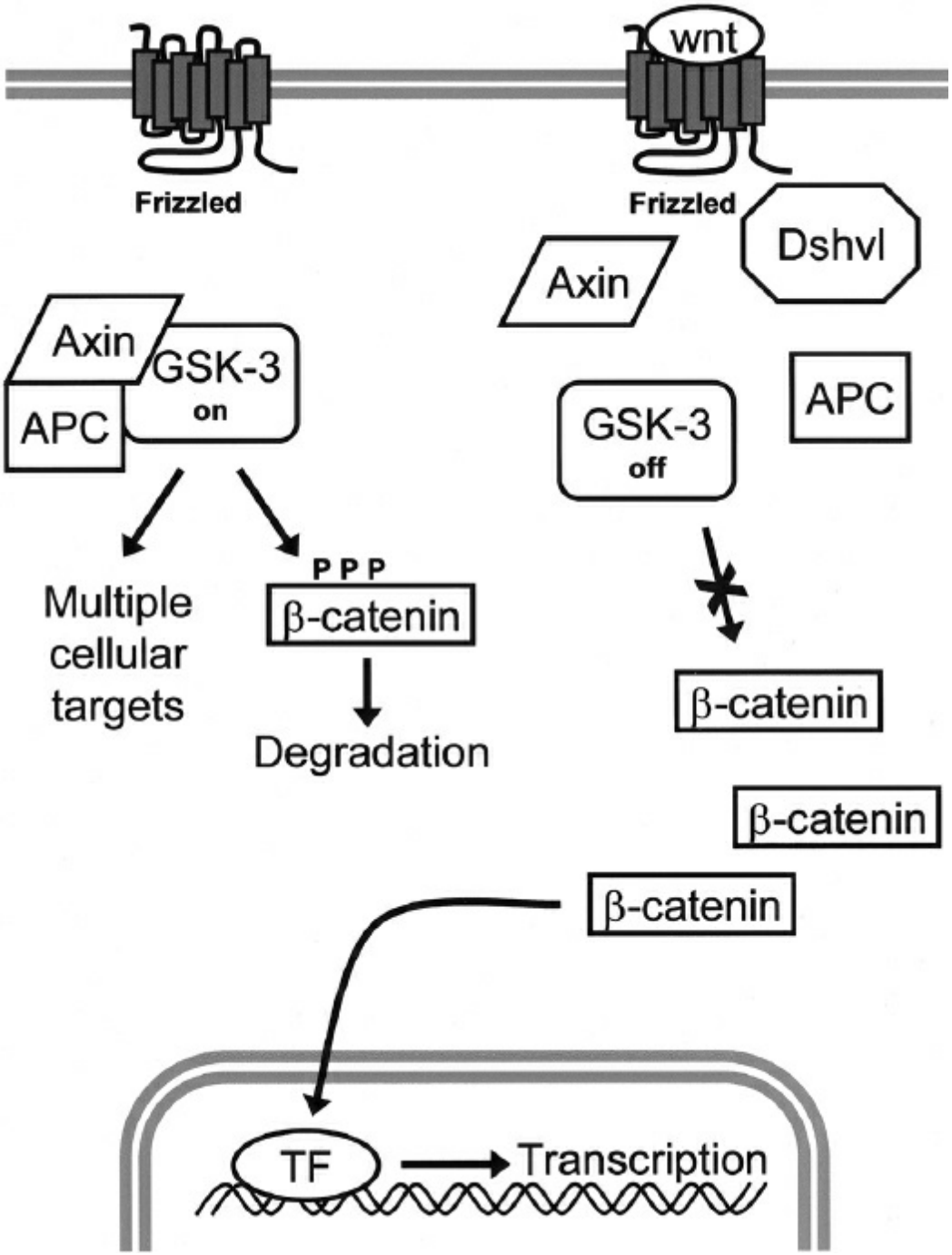

FIGURE 1.9–12. Basic organization of the canonical Wnt signaling pathway. The primary, or canonical, Wnt signaling pathway begins with the binding of secreted Wnt proteins to cell-surface receptors of the Frizzled family. In the absence of Wnt signaling, a complex of proteins that includes axin, adenomatosis polyposis coli (APC) protein, and glycogen synthase kinase 3 (GSK-3) maintains an active form of GSK-3 (on), which phosphorylates multiple cellular proteins, including the transcriptional coactivator β-catenin. When β-catenin is phosphorylated, it is targeted for degradation. When Wnt signaling is initiated, Frizzled activates a protein called Dishevelled (Dshvl), which causes the dissociation of the axin/APC/GSK-3 complex, resulting in an inhibition of GSK-3 activity. Decreased GSK-3 activity reduces the degradation of β-catenin, which then translocates to the nucleus, interacts with transcription factors (TFs), and promotes specific gene expression changes.

As discussed above, Akt is also activated by neurotrophic factors such as BDNF, low levels of which are implicated in depression and other neuropsychiatric disorders. Other kinases that regulate GSK-3 include PKA, PKC, and RSK, demonstrating that the mechanisms of regulation and biological targets of GSK-3 are quite diverse. This convergence of diverse signaling pathways onto GSK-3 is a characteristic that has led to the labeling of GSK-3 as a crucial signaling "node." The precise mechanisms that regulate the cross-talk among these distinct pathways are not well established, and this is an area of active research. However, it is likely that the compartmentalization of GSK-3 to distinct regions of the cell minimizes much of the potential cross-talk among pathways. Interestingly, there is growing evidence that GSK-3 is involved in synaptic plasticity (see below), possibly by "funneling" diverse inputs into the processes that regulate synaptic strength.

SYNAPTIC PLASTICITY

Changes in the strength and efficiency of synaptic signaling, termed synaptic plasticity, underlie one of the most important neurochemical foundations of learning and memory. Because these processes play a prominent role in a variety of psychiatric disorders and psychotherapies, there has been intense interest in defining the cellular and molecular events mediating these processes. Because synaptic plasticity is activity-dependent, various intraneuronal signaling pathways are important for coordinating these changes. Indeed, there are several mechanisms that cooperatively affect synaptic plasticity, including changes in the release of presynaptic neurotransmitters and changes in how effectively the postsynaptic neuron responds to those neurotransmitters. A postsynaptic mechanism that is widely considered to be a major mediator for enhancing synaptic efficacy is called LTP.

LTP is roughly defined as an increase in the strength of a synapse that lasts from hours to several days and is widely considered one of the major mechanisms by which memories are formed and stored in the brain. Given the diversity of neuronal cell types in the brain, there are many variations in the processes involved in LTP; however, the prototypical model is the CA1 region of the hippocampus, which has glutamatergic synapses. At these synapses, there are both early and late stages of LTP that are initiated by the actions of two glutamate-gated ion channels, α-amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA) and NMDA receptors. Early LTP is mediated by modulating proteins already at the synapse while late LTP requires new protein synthesis.

The initiation of early LTP begins with glutamate binding to AMPA receptors, which allows Na^+ to enter the synapse, depolarizing the membrane (see Section 1.10 for details) (Fig. 1.9–13). When the postsynaptic membrane is sufficiently depolarized, NMDA receptors open,

leading to a rapid increase in intracellular Ca^{2+} concentrations. The rise in Ca^{2+} levels leads to the activation of CaM kinase II and PKC, which rapidly phosphorylate the AMPA receptors, increasing the efficiency of synaptic transmission, and regulate the insertion of additional AMPA receptors into the postsynaptic membrane. By increasing the number of AMPA receptors at the synapse, future signaling stimuli are able to generate larger postsynaptic responses. This process may additionally be regulated by the concomitant activation of other signaling pathways through various GPCRs.

Although these phosphorylation events underlie the rapid changes in synaptic efficacy during early LTP, enduring changes characteristic of "late LTP" depend on the targeting of newly synthesized proteins to the synapses. These newly targeted proteins, which can include additional receptors and scaffolding proteins, induce a remodeling of the synapse and can profoundly strengthen postsynaptic responses to stimuli. At least some of this new protein synthesis appears to occur locally in the dendrites allowing direct targeting of the new proteins to the activated synapse, a process largely regulated by mTOR activity.

Overall, LTP and the other cellular mechanisms involved in synaptic plasticity, including a process called LTD, are highly active areas of neuroscience research. The current understanding of these processes highlights the complexity and importance of tight temporal and spatial regulation of synaptic signaling. Continued elucidation of these biochemical mechanisms underlying synaptic plasticity will enhance the understanding of learning and behavior, provide insights into psychiatric diseases, and may allow the development of pharmacological agents that can improve cognition.

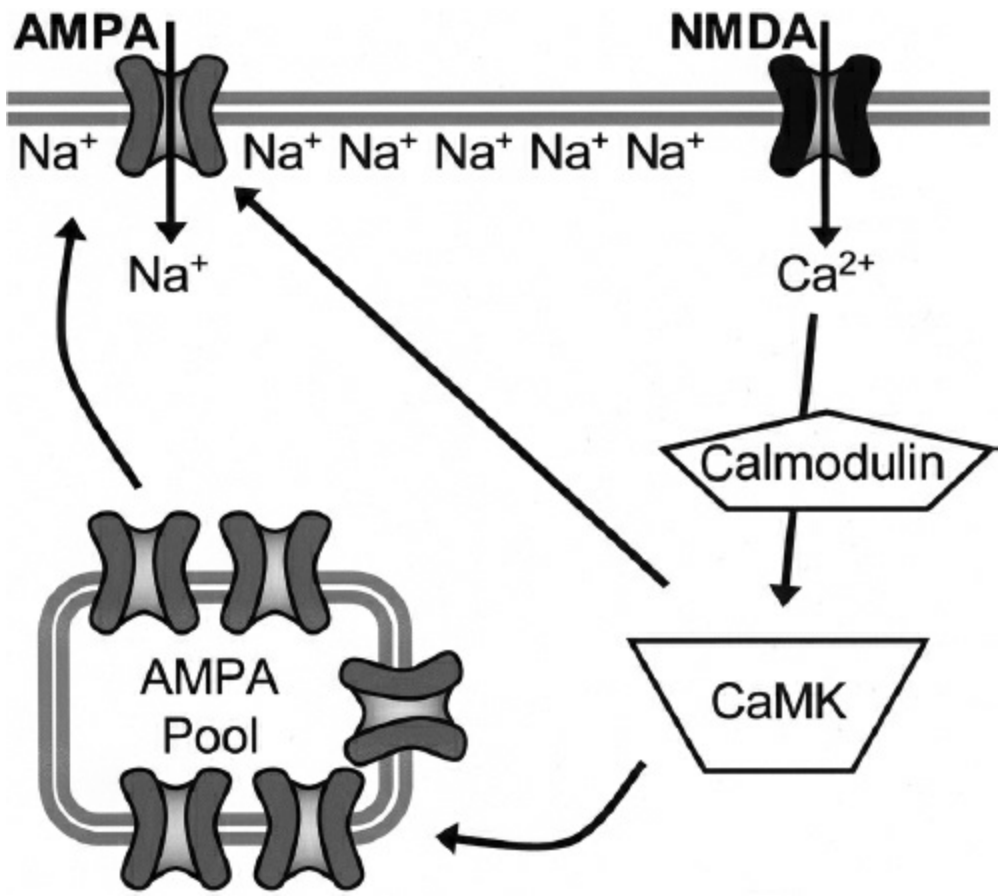

FIGURE 1.9–13. An example schematic of a postsynaptic mechanism involved in long-term potentiation (LTP) and synaptic plasticity. Initiation of LTP begins with glutamate binding to α-amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA) receptors gating the influx of Na^+ ions into the synapse, depolarizing the membrane. When the postsynaptic membrane is sufficiently depolarized, *N*-methyl-D-aspartic acid (NMDA) receptors open, leading to a rapid increase in intracellular Ca^{2+} levels, which, through binding to calmodulin, activate Ca^{2+}/calmodulin-dependent kinases (CaMKs) and protein kinase C (PKC). These activated protein kinases can then both phosphorylate AMPA receptors and regulate the insertion of additional AMPA receptors into the postsynaptic membrane intracellular pool. Together, these processes increase the efficiency of future synaptic transmission.

SIGNALING COMPLEXES AND SCAFFOLDS

Several additional types of proteins are central to the organization of signaling pathways. These include scaffolding and anchoring proteins, which provide mechanisms to ensure that information being signaled in cells is transferred to the appropriate targets in a timely and efficient manner. They do this by mediating the localized assembly of multiprotein complexes that contain, for example, receptors, second-messenger–generating enzymes, kinases, phosphatases, and substrates. By keeping many of the components of a signaling cascade in close proximity, these complexes minimize the need for activated proteins to diffuse through a dense cytoplasm to find their targets, thus greatly enhancing signaling efficiency. Additionally, these systems both maintain the separation, or compartmentalization, of distinct signals when simultaneous signaling

events are occurring and are crucially involved in moderating and integrating this information. The characterization of the multitude of ways that signaling pathways interact via these complexes is an active area of research.

Many of these scaffolding and adaptor proteins use specific protein–protein interactions to mediate the transport and localization of signaling proteins and form these specialized multiprotein complexes. Protein interactions are often formed by distinct domains within adaptor proteins that are responsible for recognizing and binding to specific regions of other proteins. An example mentioned above is the Src homology (SH) domain involved heavily in neurotrophin signaling. SH2 domains, found in adaptor proteins such as Grb2, are roughly 100 amino acids long and specifically bind to short amino acid sequences that contain a phosphorylated tyrosine residue. Multiple proteins can bind to these phosphotyrosine sequences via their SH2 domains, some of which are subsequently phosphorylated and activated, and others act as adaptors that recruit other substrates to the kinase. Indeed, activated receptor tyrosine kinases, such as Trk, serve as scaffolds for an array of activated signaling molecules. Another important example is the localization of PKA to distinct sites within cells by a family of scaffolding proteins called A-kinase-anchoring proteins (AKAPs). These AKAPs bind to the regulatory subunits of PKA, holding it near particular substrates while it is awaiting activation by cAMP. Some AKAPs, for example, are thought to localize PKA next to synaptic ion channels, greatly enhancing the rate of substrate phosphorylation by eliminating delays due to protein diffusion.

Scaffolding also organizes signaling complexes involving the classic neurotransmitter receptors and ion channels at synapses. A common protein domain involved in protein scaffolding in these systems is the PDZ domain. PDZ domains are found in more than 400 proteins in humans and bind tightly to the extreme C-terminal segment of proteins in which the last three amino acids are S/TXV (i.e., serine [S] or threonine [T], followed by any amino acid [X], followed by valine [V] or another hydrophobic amino acid). Given the large number of proteins containing PDZ domains and PDZ recognition elements, scaffolds containing these proteins can assemble into very large molecular complexes. The best-known example of these large scaffolds is the PSD of excitatory neurons, which organizes glutamate receptors and their associated signaling proteins at the postsynaptic membrane and helps to determine the size and strength of synapses. Indeed, scores of proteins have been identified in the PSD: Ion-gated and GPCRs, kinases, and phosphatases and cytoskeletal proteins, all targeted and maintained in the PSD by various adaptor and scaffolding proteins. Thus, scaffolding proteins are major players in the organization of the postsynaptic signaling machinery.

Functional Selectivity

Even though most cell-surface receptors have classically been described as activating a single primary intracellular signaling cascade, as discussed above, most receptors also activate one or more additional pathways. Some GPCRs, for example, have been shown to signal through the phosphatidylinositol pathway, the arachidonic acid pathway, and the MAPK/ERK pathway, among others. In addition, activation of these receptors stimulates the biochemical mechanisms involved in their desensitization and internalization. A classical concept of receptor pharmacology is that a receptor ligand can be either classified as a full agonist, partial agonist, or antagonist at that receptor and that this classification will be consistent for all of the signaling pathways for that receptor. In other words, if a receptor ligand fully activates the phosphatidylinositol pathway, then it is expected to fully activate all of the other signaling and regulatory pathways linked to that receptor. However, there is now evidence that some ligands may inherently be able to produce different levels of signaling among the various pathways (Fig. 1.9–14). This phenomenon is most often referred to as "functional selectivity" (also called biased agonism).

An example of functional selectivity is seen with 5-HT_{2C} serotonin receptors, which activate the phosphatidylinositol signaling pathway as well as arachidonic acid release. Pharmacological studies looking at a panel of different 5-HT_{2C} receptor agonists showed that full agonism for increasing IP_3 and Ca^{2+} was not correlated with the efficacy of the ligand to increase arachidonic acid. In addition, it was demonstrated that the ability of agonists to activate 5-HT_{2C} receptor signaling pathways did not predict their ability to desensitize the receptor to that pathway. For example, the 5-HT_{2C} receptor ligand *meta*-chlorophenylpiperazine (mCPP) is a partial agonist for IP_3 signaling with 80 to 90 percent of the efficacy of the endogenous full agonist 5-HT and causes a similar level of receptor desensitization. However, mCPP is a full agonist for the arachidonic acid pathway, but induces little or no 5-HT_{2C} receptor desensitization. At the extreme end, functionally selective ligands may act both as agonists and antagonists at different receptor-mediated functions. As an interesting example, 5-HT_{2A} serotonin receptor antagonists, while unable to induce the stimulation of any classical signaling pathways, have been shown to induce receptor internalization and downregulation. Similar antagonist-induced internalization has also been demonstrated with CCK and other peptide receptors.

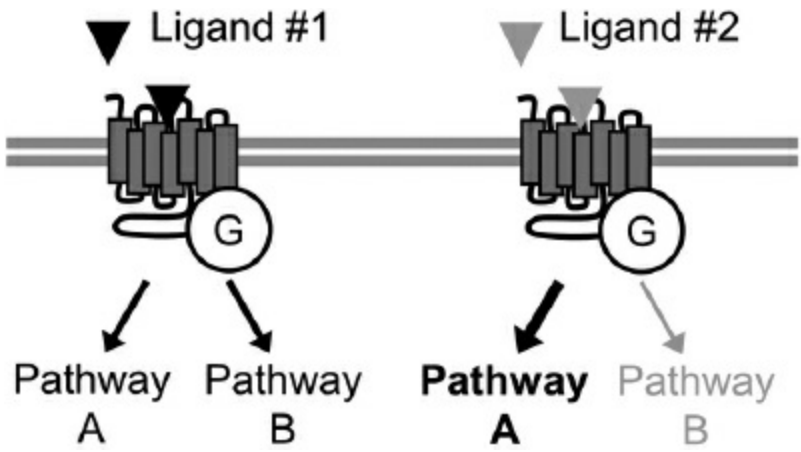

FIGURE 1.9–14. Functional selectivity. In addition to the canonical G-protein–mediated signaling pathways, most GPCRs can activate other signaling pathways. While classical receptor pharmacology theory posits that a full agonist at a receptor will activate all signaling pathways equally, it is now clear that some ligands may inherently be able to produce different levels of signaling from the various pathways. This phenomenon is referred to as "functional selectivity."

While the examples above focus on serotonin receptors, functional selectivity has been demonstrated in many GPCRs and may be mediated by a variety of mechanisms. First, different ligands may sample and stabilize unique conformational changes in the receptor protein, resulting in a differential activation of the various signaling pathways. Second, functional selectivity may be affected by the diversity of G proteins and other signaling and scaffolding proteins or may be related to the observed ability of GPCRs to dimerize and oligomerize, the function of which remains poorly understood. Thus, not only is the concept of functional selectivity an interesting one, but it is likely to have an important impact on future psychiatric drug development.

REMOTE CONTROL OF INTRANEURONAL SIGNALING

In recent years, emerging synthetic biology technologies have allowed for the targeted activation of intracellular signaling pathways with spatial and temporal precision. These techniques have facilitated the study of complex circuits and behaviors and are revolutionizing the understanding of brain function. In one approach, chimeric receptors were engineered that replaced the intracellular loops of the light-activated rhodopsin with those of specific adrenergic or serotonergic receptors. These artificial receptors, collectively called "opto-XRs," allow the use of light to activate the signaling pathways normally recruited by the receptor used for the intracellular loops. Because of the speed of optical manipulation, this approach allows for rapid stimulation of intracellular signaling and examination of the effects of tonic versus phasic activation, though it requires the invasive insertion of optical fibers into the brain. Another approach uses specifically engineered GPCRs that are only activated by synthetic ligands that do not

activate endogenous receptors. Termed DREADDs (designer receptors exclusively activated by designer drugs) or RASSLs (receptors activated solely by synthetic ligands), these technologies similarly allow for precise spatiotemporal control of intracellular signaling in vivo. Various versions of DREADDs and RASSLs activate the major GPCR signaling pathways (G_s, G_i, and G_q). In addition, since they are activated by drugs, DREADDs and RASSLs allow the noninvasive modulation of intracellular signaling and thus have potential to be used for transformative therapeutic approaches for a host of neuropsychiatric disorders.

FUTURE DIRECTIONS

Translating the advances in molecular neurobiology into improved diagnostic and therapeutic capabilities represents the greatest opportunity and challenge facing modern psychiatry. The current armamentarium of medications used in treating psychiatric diseases has facilitated decades of progress in understanding intercellular signaling mediated by cell-surface receptors. However, dramatic and ongoing advances in the understanding of the intraneuronal signaling pathways activated by these receptors will likely lead to novel, innovative, and improved pharmacological agents for psychiatric diseases, as has been achieved in other branches of medicine. Thus, future efforts in drug discovery should move beyond the current strategies of solely targeting synaptic neurotransmission at the receptor level to the development of agents acting on components of intracellular signaling pathways. By nature, signaling pathways have significant redundancy and interactions. Thus, identifying and targeting critical points within these networks may lead to improved molecular diagnostic tests and treatments.

REFERENCES

Airan RD, Thompson KR, Fenno LE, Bernstein H, Deisseroth K. Temporally precise in vivo control of intracellular signalling. *Nature*. 2009;458:1025–1029.

*Alberts B, Johnson A, Lewis J, Raff M, Roberts K. *Molecular Biology of the Cell*. New York: Garland; 2002.

Angelucci F, Brene S, Mathe AA. BDNF in schizophrenia, depression and corresponding animal models. *Mol Psychiatry*. 2005;10:345–352.

*Blitzer RD, Iyengar R, Landau EM. Postsynaptic signaling networks: Cellular cogwheels underlying long-term plasticity. *Biol Psychiatry*. 2005;57:113–119.

Boeckers TM. The postsynaptic density. *Cell Tissue Res*. 2006;326:409–422.

Bohn LM, Gainetdinov RR, Lin FT, Lefkowitz RJ, Caron MG. Mu-opioid receptor desensitization by β-arrestin-2 determines morphine tolerance but not dependence. *Nature*. 2000;408:720–723.

Cheyette BN, Moon RT. Wnt protein family. In: Henry HL, Norman AW, eds. *Encyclopedia of Hormones*. San Diego: Academic Press; 2003.

Doupnik CA. GPCR-Kir channel signaling complexes: defining rules of engagement. *J Recept Signal Transduct Res*. 2008;28:83–91.

Duman RS, Voleti B. Signaling pathways underlying the pathophysiology and treatment of depression: novel mechanisms for rapid-acting agents. *Trends in Neurosci*. 2012;35:47–56.

Gainetdinov RR, Premont RT, Bohn LM, Lefkowitz RJ, Caron MG. Desensitization of G protein-coupled receptors and neuronal functions. *Annu Rev Neurosci*. 2004;27:107–144.

Hanyaloglu AC, von Zastrow M. Regulation of GPCRs by endocytic membrane trafficking and its potential implications. *Annu Rev Pharmacol Toxicol*. 2008;48:537–568.

Kobayashi T, Ikeda K. G protein-activated inwardly rectifying potassium channels as potential therapeutic targets. *Curr Pharm Des*. 2006;12:4513–4523.

Kroeze WK, Sheffler DJ, Roth BL. G-protein-coupled receptors at a glance. *J Cell Sci*. 2003;116:4867–4869.

Lee HK, Takamiya K, Han JS, et al. Phosphorylation of the AMPA receptor GluR1 subunit is required for synaptic plasticity and retention of spatial memory. *Cell*. 2003;112:631–643.

Le Novère N, Li L, Girault JA. DARPP-32: molecular integration of phosphorylation potential. *Cell Mol Life Sci*. 2008;65:2125–2127.

Malinow R, Malenka RC. AMPA receptor trafficking and synaptic plasticity. *Annu Rev Neurosci*. 2002;25:103–126.

Maxwell CR, Kanes SJ, Abel T, Siegel SJ. Phosphodiesterase inhibitors: A novel mechanism for receptor-independent antipsychotic medications. *Neuroscience*. 2004;29:101–107.

Michel JJ, Scott JD. AKAP mediated signal transduction. *Annu Rev Pharmacol Toxicol*. 2002;42:235–257.

Miyakawa T, Leiter LM, Gerber DJ, Gainetdinov RR, Sotnikova TD, Zeng H. Conditional calcineurin knockout mice exhibit multiple abnormal behaviors related to schizophrenia. *Proc Natl Acad Sci U S A*. 2003;100:8987–8992.

Müller N, Schwarz MJ. COX-2 inhibition in schizophrenia and major depression. *Curr Pharm Des*. 2008;14:1452–1465.

*Nestler EJ, Hyman SE, Malenka RC. *Molecular Neuropharmacology*. New York: McGraw-Hill; 2001.

Pastalkova E, Serrano P, Pinkhasova D, Wallace E, Fenton AA, Sacktor TC. Storage of spatial information by the maintenance mechanism of LTP. *Science*. 2006;313:1141–1144.

*Patapoutian A, Reichardt LF. Trk receptors: Mediators of neurotrophin action. *Curr Opin Neurobiol*. 2001;11:272–280.

Peineau S, Bradley C, Taghibiglou C, et al. The role of GSK-3 in synaptic plasticity. *Br J Pharmacol*. 2008;153(Suppl 1):S428–S437.

*Rogan SC, Roth BL. Remote control of neuronal signaling. *Pharmacol Rev*. 2011;63:291–315.

*Urban JD, Clarke WP, von Zastrow M, et al. Functional selectivity and classical concepts of quantitative pharmacology. *J Pharmacol Exp Ther*. 2007;320:1–13.

Willars GB. Mammalian RGS proteins: Multifunctional regulators of cellular signalling. *Semin Cell Dev Biol*. 2006;17:363–376.

Yaffe M, Cantley L. Grabbing phosphoproteins (what happens to proteins after phosphorylation). *Nature*. 1999;402:30–31.

▲ 1.10 Cellular and Synaptic Basis of Neural Signaling

CHARLES F. ZORUMSKI, M.D., AND STEVEN MENNERICK, PH.D.

INTRODUCTION

Neuropsychiatric disorders involve defects in intercellular communication. Although these defects often include changes in synaptic interactions between neurons, changes in the excitability of neurons also contribute to pathogenesis. Furthermore, pharmacological and brain stimulation treatments aimed at altering neuronal excitability are becoming standard for neurological and psychiatric disorders. The clearest example is epilepsy where abnormal neuronal excitability is a hallmark of the disorder. Altered excitability also contributes to primary psychiatric disorders, and many anticonvulsants used to treat mood disorders affect neuronal excitability and secondarily influence synaptic function. Similarly, brain stimulation techniques directly exploit electrical properties of neurons for therapeutic purposes.

Although the brain is not an "electrical organ," neurons use electrical signals to send and receive information. These electrical signals determine the brain's local and network properties and result from the flow of ions across cell membranes through macromolecular pores called **ion channels**. Neurons express two major classes of ion channels, gated and nongated. Nongated (or leakage) channels are open at rest and contribute to the cellular **resting membrane potential**. The opening and closing of most ion channels is regulated (**gated**) by changes in transmembrane voltage, extracellular chemicals, or intracellular messengers. Certain **voltage-gated sodium channels** open and close rapidly and provide the basis for communication within and between neurons. These rapid signals (**action potentials**) are generated near the neuronal cell body and are transmitted with little decrement in amplitude along the neuron's axon to nerve terminals. This high-fidelity signal propagation involves the regenerative nature of action potentials, imparted by the presence of voltage-gated channels along the length of the axon. In myelinated axons, action potential propagation is enhanced by **saltatory conduction**, which refers to the ability of electrical signals to "jump" between axonal nodes of Ranvier.

At nerve terminals, the wave of action potential-induced depolarization opens **voltage-gated calcium channels**, and the influx of calcium promotes release of a chemical neurotransmitter into the extracellular space where the transmitter then influences a receptive cell. **Neurotransmitters** bind two broad classes of **receptors** and alter neuronal excitability via actions on ion channels. **Ligand-gated ion**

channels are directly opened by transmitter binding whereas **GPCRs** influence ion channels indirectly via guanine nucleotide binding proteins (G proteins) or intracellular chemical messengers. This chapter outlines the basic principles of cellular electrophysiology and discusses the importance of electrical signals to brain function and neuropsychiatry.

PRINCIPLES OF CELLULAR ELECTROPHYSIOLOGY

Resting Membrane Potential

The concentration of potassium ion [K^+] is much higher inside the cell than outside, reflecting the selective permeability of most cell membranes at rest to K^+. The basis for this selective permeability is the presence of nongated (leakage) K^+ channels in the cell membrane. Potassium channels are a class of transmembrane proteins with a hydrophilic pore region that selectively conducts K^+. Positively charged K^+ ions are initially attracted into the cell by large, impermeant intracellular anions (acids and proteins). As K^+ accumulates in the cell, the membrane potential of the cell becomes more depolarized (less negative), and therefore, K^+ entry is driven less and less by the electrical gradient. Intracellular concentrations of K^+ achieve levels of ~100 mM while extracellular [K^+] is between 2 and 6 mM, setting up a chemical gradient, which in isolation would result in net K^+ efflux from the cell. Thus, two gradients act on K^+, the intracellular electronegativity resulting in K^+ influx, and the chemical gradient resulting in K^+ efflux. At a specific membrane potential (around −90 mV), the electrical and chemical gradients for K^+ are equal and opposite. This membrane potential is known as the **equilibrium potential** (also called the **Nernst potential**) for K^+. The equilibrium potential is the transmembrane potential at which the electrical and chemical gradients are balanced and there is no net flux of a specific ion. In a cell whose membrane is exclusively permeable to K^+, the resting potential of the cell would be equal to the equilibrium potential for K^+.

The situation in most neurons is not this simple because other ions, with different electrochemical gradients, are slightly permeant through ion channels that are open in the resting cell membrane. Each of these ions has its own characteristic equilibrium potential depending on the ion concentrations inside and outside the cell. The cations Na^+ and Ca^{2+} are present at higher concentrations outside than inside the cell. Therefore, at negative membrane potentials, both the electrical and chemical gradients for these cations are inwardly directed, and the equilibrium potentials are positive to 0 mV. Chloride (Cl^-) concentrations are usually higher outside the cell, but because of this ion's negative charge, the equilibrium potential for chloride is near the resting potential (Fig. 1.10–1). The actual resting potential of a given cell is determined by the average of the equilibrium

potentials of all the permeant ions, weighted by the relative permeability of each ion. At rest, K^+ and Cl^- are much more permeant than the other ions so the resting potential is closest to the equilibrium potentials for these ions. Na^+ and Ca^{2+} are less permeant and thus contribute less to the resting potential, but the small permeability of nongated channels to these cations renders the actual resting potential more positive than the equilibrium potentials of K^+ or Cl^-. Typical values for neuronal resting potentials are between −55 and −70 mV. Astrocytes, by contrast, have a membrane more purely permeable to K^+ and therefore a more deeply negative resting membrane potential (~−90 mV), very close to the K^+ equilibrium potential.

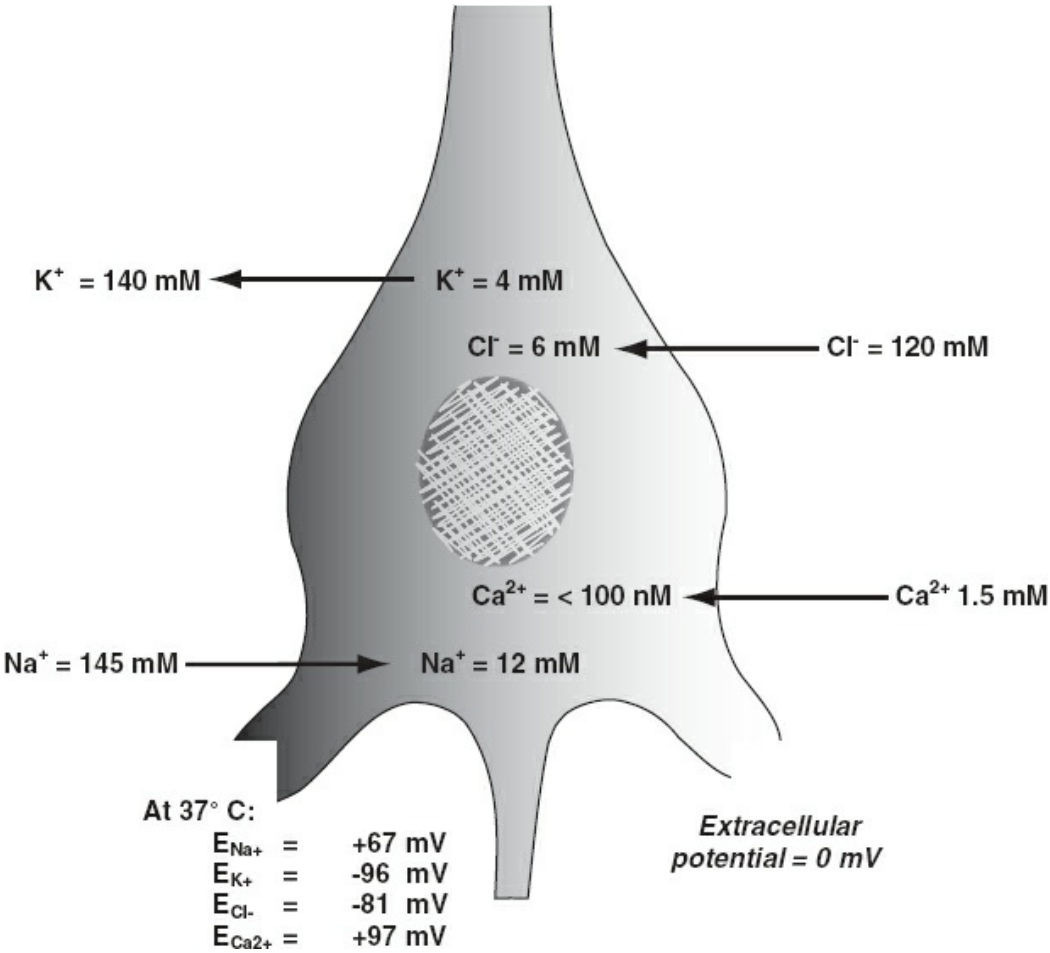

FIGURE 1.10–1. The figure shows the distribution of Na^+, K^+, Ca^{2+}, and Cl^- across the membrane of a neuron; the arrows show the direction of current flow down the chemical gradient. Using the indicated ion concentrations, the equilibrium (Nernst) potentials for these ions at 37°C are shown at the lower left.

The concepts of equilibrium potential and membrane potential outlined above can be described quantitatively by the Nernst equation. The equilibrium potential for any ionic species can be calculated based on the ion concentrations on either side of the membrane. For K^+, the equilibrium potential (designated E_K) is expressed as: $\boldsymbol{E_K = (RT/zF)\ \ddot{e}\ \ln([K]_o/[K]_i)}$,

where ***R*** is the ideal gas constant (8.31 joules/degree/mole), ***T*** is the temperature in degrees Kelvin, ***z*** is the valence of the ion, ***F*** is Faraday's constant (96,500 coulombs/mole, the charge on a mole of monovalent ions), and $[\mathbf{K}]_o$ and $[\boldsymbol{K}]_i$ are the concentrations of K^+ outside and inside the cell. At 37°C, the equilibrium potential for K^+ is −96 mV, E_{Na} is +67 mV, E_{Cl} is −81 mV, and E_{Ca} is greater than +97 mV. These equilibrium potentials are important in determining what happens to the membrane potential when an ion channel that is permeable to a specific ion opens or closes because the opening of a specific ion channel drives the membrane potential toward the equilibrium potential for that ion. For example, when K^+-selective ion channels open the neuronal membrane potential moves toward −96 mV. This makes the inside of the cell more negative, an effect called **hyperpolarization**. Na^+ and Ca^{2+} channel opening has the opposite effect, making the inside of the cell less negative (**depolarization**).

Because resting cell membranes are permeable to more than one ion, the true membrane potential is never exactly equal to the equilibrium potential for any one ion. The Goldman–Hodgkin–Katz (GHK) equation quantitatively describes the actual resting potential as the average of the various equilibrium potentials, weighted by the relative permeability of each ion. The equation is of the form:

$$E_m = \frac{RT}{F} \ln \frac{P_k[K]_o + P_{Na}[Na]_o + P_{Cl}[Cl]_i}{P_k[K]_i + P_{Na}[Na]_i + P_{Cl}[Cl]_o}$$

Most of the variables are familiar from the Nernst equation. E_m is the membrane potential and P_{ion} is the permeability of the membrane to the ion. The resting membrane potential is thus a reversal potential (potential at which no net inward or outward current flows) for the various conductances open at rest.

The bulk solutions on either side of the membrane are electrically neutral, with most of the intracellular negative charge being contributed by large intracellular organic anions (acids and proteins). The differential distribution of ions across neuronal membranes is maintained by membrane pumps that use energy from ATP hydrolysis to move ions against a concentration gradient into or out of the cell. The best-characterized pump is the Na^+-K^+ ATPase (**sodium pump**) that transports 3 Na^+ out of and 2 K^+ into the cell during each cycle. Because an unequal amount of charge is moved during each cycle, the pump is **electrogenic** and contributes to intracellular negativity of the membrane potential. Na^+-K^+ ATPase activity is a major contributor to brain energy utilization, with as much as 40 percent of brain oxygen consumption

resulting from the pump activity required to reestablish ionic homeostasis following action potential firing and synaptic transmission. The cardiac glycosides, digoxin and ouabain, inhibit Na^+-K^+ ATPase in the heart and improve myocardial contractility by depolarizing cardiac myocytes and increasing intracellular Ca^{2+}.

The resting potential is relatively static and represents the potential energy available for neuronal signaling. Negative resting potentials are not unique to excitable cells, but neurons and other excitable cells make unique use of the energy stored in the resting potential to generate transient membrane potential *changes,* the true currency of neuronal information exchange. A membrane potential change is typically initiated by a change in current flow across the membrane, usually resulting from the opening or closing of specific ion channels discussed below. The number of ions needed to change the membrane potential is very small relative to concentrations in the bulk intracellular and extracellular solutions. For example, a potential change of 100 mV across a 1 cm^2 area of membrane requires the movement of only about 10^{-12} moles of a monovalent (singly charged) ion. By comparison, Na^+ and K^+ are present at about 10^{-1} M in the extracellular and intracellular fluids, respectively.

Passive Membrane Properties

To understand how ion concentration gradients, electrical gradients, ion channels, and the distribution of charges across the membrane are related, it is helpful to think of the cell membrane as an electrical circuit consisting of resistors (conductors), batteries, and capacitors. Because ions flow through ion channels and do not directly penetrate the cell membrane, ion channels can be viewed as variable resistors. Physiologists describe ion channels in terms of their **selectivity** (which ions flow through the channel) and their **conductance** (relative ease of passing ions). Conductance (g) is the inverse of **resistance** (R) in an electrical circuit ($g = 1/R$). The presence of a voltage across the membrane provides an electrical driving force for the flow of ions through ion channels resulting in a transmembrane **current**. The relationship among voltage (V), ionic current (I), and resistance (conductance) is given by the physiologist's version of Ohm's law: $\boldsymbol{I}_{\text{ionic}} = \mathbf{g}(\boldsymbol{V}_{\text{m}} - \boldsymbol{E}_{\text{rev}})$, where $\boldsymbol{V}_{\text{m}}$ is the membrane potential, $\boldsymbol{E}_{\text{rev}}$ is the equilibrium (Nernst) potential for the ion(s) flowing through the channel, and $(\boldsymbol{V}_{\text{m}} - \boldsymbol{E}_{\text{rev}})$ is the **driving force** for ion flow.

Capacitance is another important passive membrane electrical property. A capacitor consists of two conductors separated by an insulating material that is capable of storing charges of opposite sign on the two conductors. In the case of neurons, the conductors are the extracellular and intracellular fluids while the lipid membrane is the insulator. Whenever current flows through the membrane, some current must be used to charge

the membrane capacitance (C_m). The expression describing this capacitive current is: $\boldsymbol{I}_{cap} = \boldsymbol{C}_m(\boldsymbol{dV}/\boldsymbol{dt})$. Note that capacitive current flows only when the membrane potential is changing (i.e., there is some change in voltage [$\boldsymbol{dV}$] over time [$\boldsymbol{dt}$]). The total current flowing across a membrane at any given time is a sum of $\boldsymbol{I}_{cap}$ and $\boldsymbol{I}_{ionic}$. The membrane capacitance along with the leak conductance of the membrane at rest helps to set the low pass filtering property of a neuron. A major tool used by physiologists to study ionic currents is a **voltage clamp**, often implemented using the **patch clamp** method. These techniques employ specialized amplifiers to keep the membrane potential constant and eliminate capacitive currents during physiological studies by preventing changes in voltage, making it possible to measure ionic currents directly.

One way to view the operation of an ion channel is as a battery (voltage source) in series with a conductor (resistor). The different types of ion channels are in parallel with each other and with the membrane capacitance. Thus, the neuronal membrane can be represented by an equivalent electrical circuit (Fig. 1.10–2), which describes membrane current flow in response to stimulation.

Active Membrane Properties: Action Potentials

Changes in membrane potential have important effects on excitability because some ion channels are regulated (gated) by voltage changes. When neurons depolarize, specific Na^+ channels open rapidly and drive the membrane potential toward the Na^+ equilibrium potential (~+66 mV). Because of leakage channels that are open at rest, there is initially a balance between leakage currents and the currents flowing through Na^+ channels opened by depolarization. At a specific depolarized membrane potential, the current flowing through Na^+ channels exceeds the current through the leak channels. The membrane potential at which Na^+ currents exceed leakage currents is called the **threshold potential**. This potential is typically between −45 mV and −30 mV in neurons. At membrane potentials depolarized to threshold, the entry of more Na^+ into the neuron produces further depolarization, which in turn opens more Na^+ channels in a positive-feedback cycle. During this process, the neuronal membrane potential becomes more positive than 0 mV but never reaches the Na^+ equilibrium potential for three reasons. First, the leak channels continue to play a role in determining the membrane potential during the action potential. Because of the relative K^+ selectivity of these channels, the membrane potential never reaches the Na^+ equilibrium potential. Second, during depolarization, Na^+ channels not only activate but also rapidly inactivate. **Inactivation** is a process by which voltage-gated channels enter a nonconducting state despite the continued presence of the

activating stimulus (in this case, depolarization). Third, the depolarization produced by Na^+ entry also opens voltage-gated K^+ channels that drive the membrane potential toward the K^+ equilibrium potential (~−90 mV). The net effect of activation and inactivation of Na^+ channels and the delayed opening of voltage-gated K^+ channels is that the neuronal membrane potential rapidly changes to values more positive than 0 mV and then returns rapidly to the resting membrane potential. This rapid sequence occurs over several milliseconds and is referred to as an **action potential** (or sodium **spike**) (Fig. 1.10–3). The fact that the membrane potential transiently exceeds 0 mV is called an **overshoot**. To a first approximation, action potentials represent all-or-none increases in electrical excitability and are important contributors to information transfer within and between neurons. Because they propagate (see below), action potentials allow the neuronal cell body to communicate rapidly with its axon terminals. In axon terminals, the spike provides the depolarization that promotes Ca^{2+} channel opening and Ca^{2+}-dependent release of neurotransmitters.

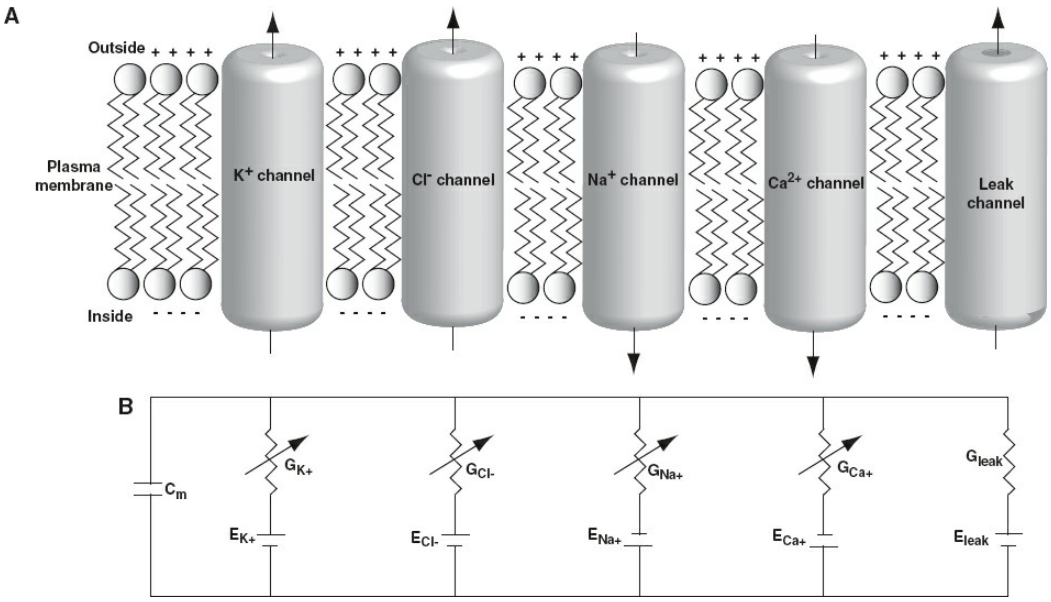

FIGURE 1.10–2. **A:** Ion channels are protein pores that span the lipid bilayer of cell membranes (upper panel). Because of membrane pumps, the extracellular surface of the membrane has a net positive charge with respect to the intracellular surface. The figure shows a membrane schematic that includes major leak and/or voltage-activated ion channels and the predominant direction of ion flux under physiological conditions. **B:** As a result of the transmembrane potential and the presence of ion channels, the neuronal membrane can be described as an equivalent electrical circuit in which each ion channel is a variable resistor (conductor, G_x) in series with a battery (E_x) (lower panel). Different ion channels are in parallel with each other and in parallel with the membrane capacitance (C_m).

In most neurons, the K^+ equilibrium potential is negative with respect to the resting membrane potential. Thus, the action potential is often followed by a transient **undershoot** (or afterhyperpolarization [AHP])

that decays back to the resting potential as the voltage-sensitive K^+ channels responsible for action potential repolarization close (Fig. 1.10–3). Following an action potential, there is a time during which depolarization either cannot elicit an action potential or during which it takes a very strong depolarization to evoke an action potential. These are called the **absolute** and **relative refractory periods**. The absolute refractory period results from nearly complete inactivation of Na^+ channels immediately following an action potential. The relative refractory period reflects the time it takes for voltage-gated K^+ channels to close following membrane repolarization.

Action Potential Conduction in Axons

Action potentials are usually generated in the initial segment of the axon (called the **axon hillock**) within about 50 μm of the neuronal cell body. Because action potentials are initiated at a distance from nerve terminals where neurotransmitters are released, an important question concerns how action potentials propagate to synaptic terminals. In a strictly passive nerve fiber, leakage of current across the membrane results in **decremental conduction** with the signal (voltage change) fading over a distance that is determined by the longitudinal (axial) resistance of the fiber, the membrane capacitance, and the transmembrane resistance. Passive decremental conduction is more typical of how voltage changes spread along dendrites back to the neuronal cell body, although dendrites also express voltage-gated ion channels that support back-propagating action potentials and that play important roles in modifying synaptically generated voltage changes in dendrites.

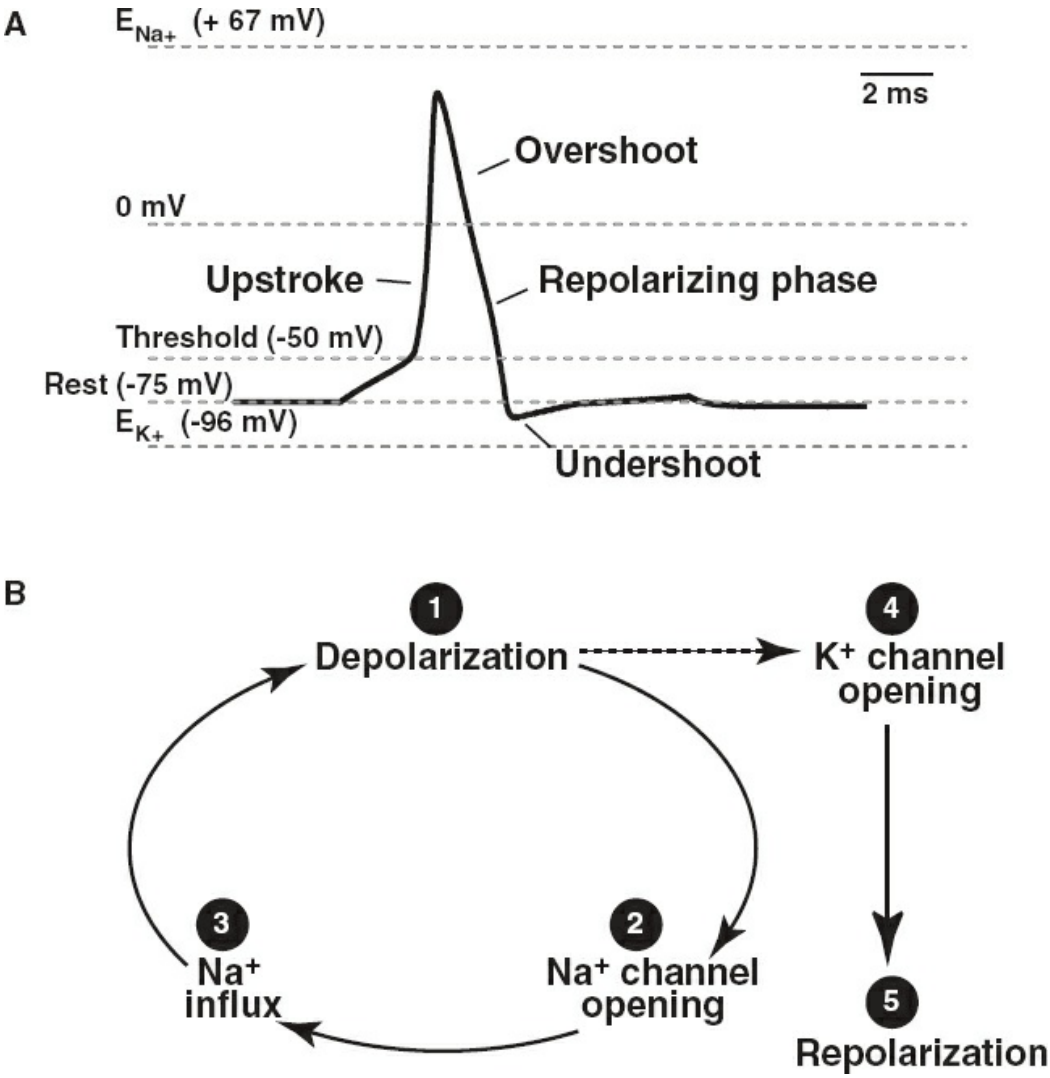

FIGURE 1.10–3. **A:** The trace shows a neuronal action potential recorded by an intracellular microelectrode. The portions of the action potential are described in the text. **B:** The lower portion of the figure shows the sequence of events underlying an action potential.

Many, but not all, axons are encased in myelin sheaths that allow them to transmit action potentials more efficiently. As a result of myelination, axons are electrically insulated except at **nodes of Ranvier** where there are dense collections of the voltage-gated Na^+ channels involved in action potential initiation (Fig. 1.10–4). The myelin sheath greatly increases transmembrane resistance and decreases membrane capacitance. This diminishes leakage of current from the axon, making it easier for current to flow down the length of the axon, thus changing the membrane potential distally. Once generated, action potentials propagate rapidly, and the wave of depolarization jumps from node to node in a form that transmits the signal faithfully to nerve terminals. This process of action potential spread through myelinated axons is called **saltatory conduction** (from the Latin word *saltare* meaning "to jump") and is important because of the speed and fidelity with which electrical information is passed from a neuronal cell

body to its terminals. Note that while depolarization-gated currents initiate comparatively sluggishly because of the time-dependent changes in channel conformational state required, the passive current flow between nodes is essentially instantaneous. Thus, passive spread of current longitudinally down the axon is important in determining the conduction speed. A typical value for conduction velocity in large myelinated axons is ~100 m/s, while propagation in small unmyelinated fibers is ~0.3 m/s. The importance of saltatory conduction is readily appreciated when considering the distances over which impulses must travel from motor cortex to cause movement in the toes. In several human illnesses, including multiple sclerosis and Guillain–Barré syndrome, demyelination of axons produces changes in axon conduction and specific neurological defects.

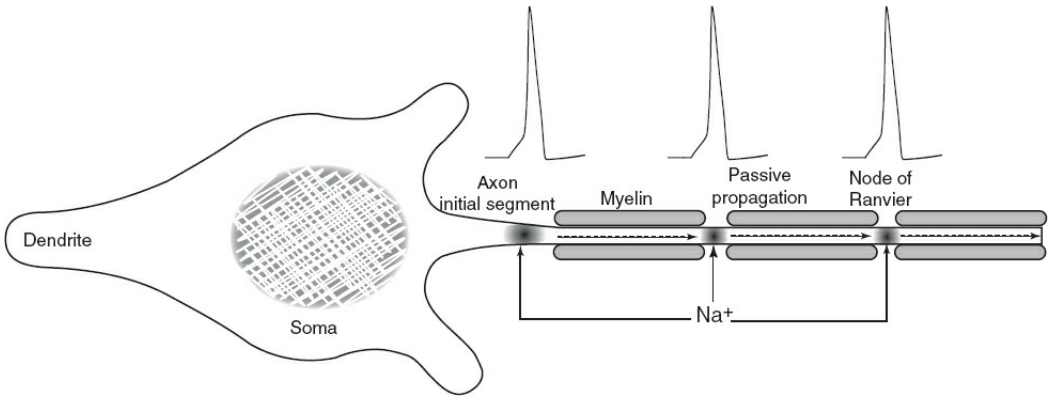

FIGURE 1.10–4. The diagram depicts saltatory conduction of an action potential in a neuron with a myelinated axon. The action potential is generated in the axon initial segment. As the waveform moves along the axon, current tends to leak from the cell diminishing the amplitude. However, myelin insulates the axon and markedly diminishes current leakage, thus enhancing flow to the first node of Ranvier. At the node of Ranvier, Na^+ channels open in response to the wave of depolarization and reproduce the all-or-none action potential. The sequence is repeated at subsequent nodes of Ranvier until the action potential reaches the nerve terminal.

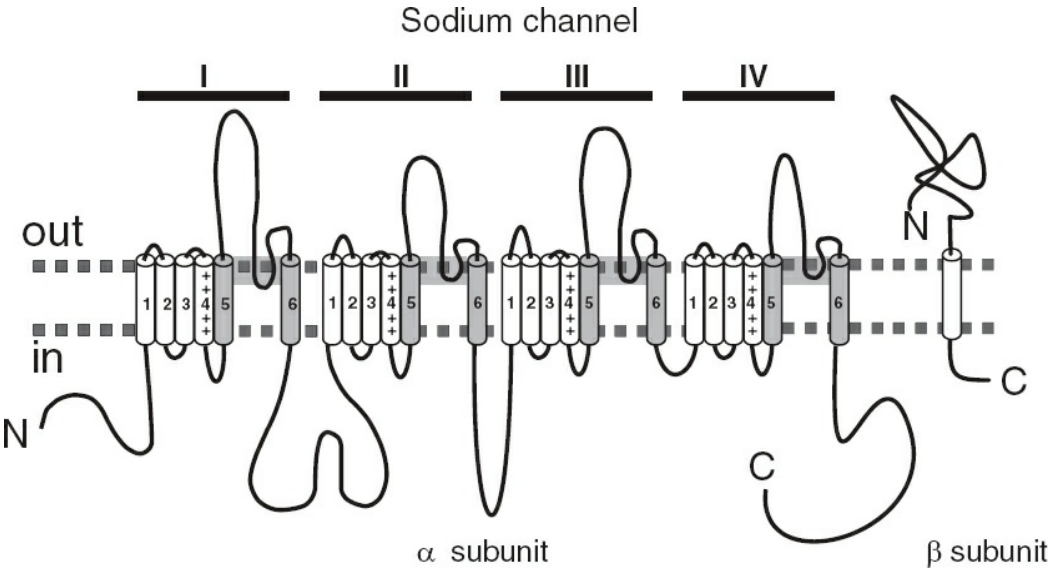

FIGURE 1.10–5. The diagram shows the proposed secondary structure of voltage-gated Na^+ channels based on α-subunit amino acid sequences. Na^+ channels consist of four homologous domains (I, II, III, and IV), each of which has six membrane-spanning regions (numbered 1 to 6). Both the amino (N) and carboxy (C) terminals are located intracellularly. A stretch of amino acids between S5 and S6, called the P-loop, forms two antiparallel β-sheets that line the ion channel pore. Positive charges in the fourth membrane spanning (S4) region comprise the voltage sensor. An accessory (β) subunit is shown to the right.

ION CHANNELS

Structure and Function of Voltage-Gated Ion Channels

Voltage-gated ion channels allow the flow of ions in response to changes in transmembrane voltage and are key elements in regulating neuronal excitation. Although ion channels can usually pass more than a single ionic species, voltage-gated channels are named according to their predominant permeant ion, and neurons express ion channels selective for Na^+, K^+, Ca^{2+}, or Cl^-. Certain ion channels, including those opened by chemical neurotransmitters such as glutamate and acetylcholine, are selective for positively charged ions (Na^+, K^+, and Ca^{2+}) but exclude Cl^-, and are called **nonselective cation channels**. To give some idea about the complexity and diversity of the "voltage-gated ion channel superfamily," it is estimated that this group of proteins has more than 140 members, and data from the human genome project (HGP) predict that there may be more than 300 ion channels. Only the families of GPCRs and protein kinases have more members. The International Union of Basic and Clinical Pharmacology (IUPHAR) maintains a website with updated nomenclature on ion channels and receptors.

Sodium (Na^+) Channels

Na^+ channels are responsible for the fast upstroke of neuronal action potentials. In some neurons, Na^+ channels also contribute to smaller depolarizations and pacemaker firing. Pacemaker activity refers to the ability of some neurons to depolarize spontaneously and to coordinate activity in a network of connected cells. Na^+ channels **activate** (open) rapidly in response to depolarization and most also **inactivate** rapidly and nearly completely in response to prolonged depolarization.

Na^+ channels cloned from rat brain have three protein subunits—a main (or α) subunit with a molecular weight of 240 to 280 kiloDaltons (kDa) and two minor (β) subunits with molecular weights of 30 to 40 kDa expressed in a 1:1:1 ratio. The α-subunit is a glycoprotein consisting of four structurally similar (homologous) domains that have six membrane-spanning (transmembrane) domains, referred to as S1 to S6 (Figs. 1.10–5 and 1.10–6). The α-subunit alone can form a functional channel. Unlike voltage-gated K^+ channels, which are tetramers of distinct subunits with

each subunit containing six transmembrane domains, functional sodium channels arise from a single α-subunit gene product. There are at least 10 mammalian genes that encode sodium channel α-subunits. These channels are named Na_v 1.1 to 1.9 and Na_x. Na_v 1.1 to 1.3 and Na_v 1.6 to 1.9 are neuronal channels while Na_v 1.4 is expressed in muscle and Na_v 1.5 is expressed in heart. Some neuronal channels are expressed primarily in the CNS, others in the PNS. The properties of voltage dependence, ion permeation, pharmacological sensitivity, activation, and inactivation are conferred by specific regions of these proteins.

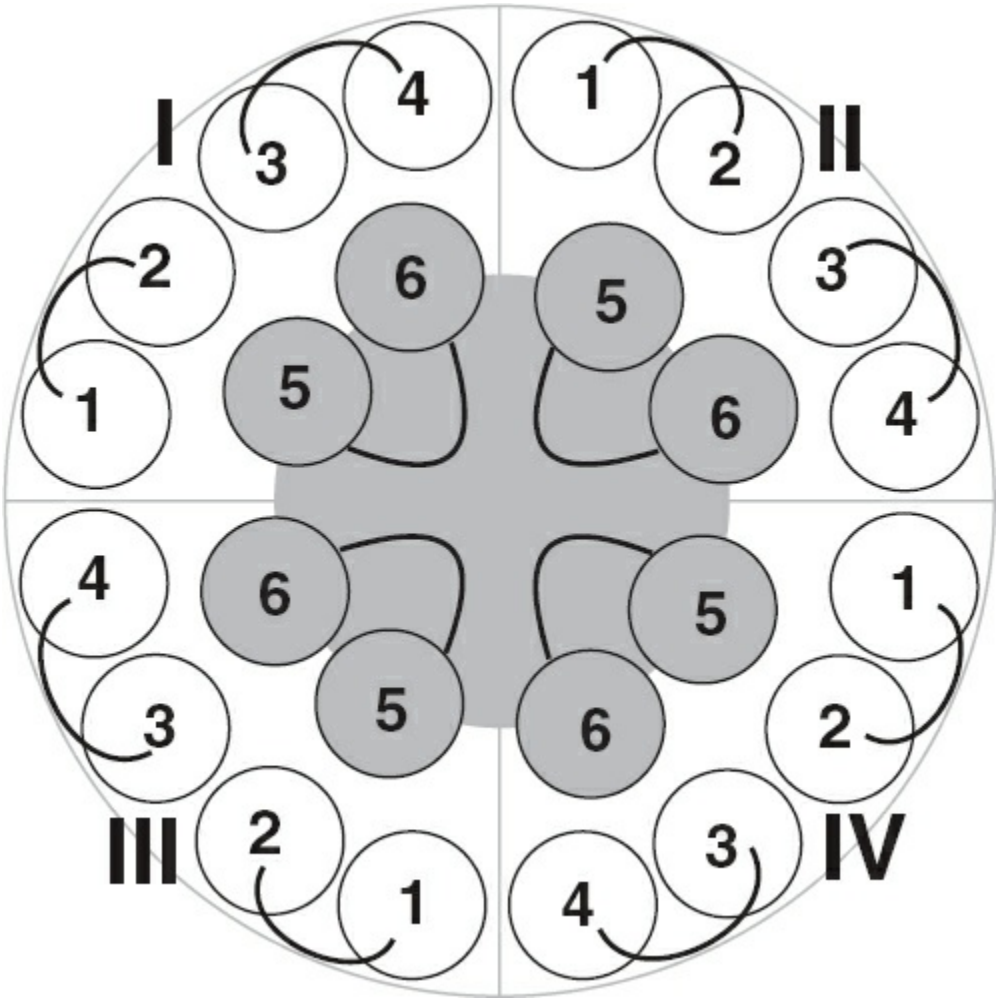

FIGURE 1.10–6. The diagram depicts the arrangement of the four homologous repeats (I to IV) in the cell membrane. The view is looking at the channel en face from the outside of the cell. Transmembrane helices are labeled 1 to 6. Extracellular linker regions are depicted by connecting lines. The P-loop is depicted as the linker between S5 and S6. Note the P-domains from each of the four homologous repeats contribute to lining the ion channel located at the center of the diagram.

Relationships between primary protein structure and ion channel function in Na^+ channels have been examined using mutations of specific amino acid residues. Both the amino- and carboxy-terminals of the α-subunits are located intracellularly. The fourth membrane-spanning region

(S4), which contains several electrically charged amino acid residues, plays a key role in sensing transmembrane voltage changes that allow channel opening. Between the S5 and S6 membrane-spanning regions, there is a segment of hydrophobic amino acids that does not completely cross the lipid membrane bilayer. This group of amino acids (called a "P-loop") forms the lining of the ion channel PORE (Figs. 1.10–5 and 1.10–6). The P-loop is a feature shared by other voltage-gated ion channels and some nonselective cation channels (Figs. 1.10–7 and 1.10–9). A short intracellular loop between the third and fourth homologous domains plays a role in channel inactivation and physically blocks the ion pore during longer periods of depolarization. A high-resolution crystal structure of a bacterial Na^+ channel (NavAb) provides insights into structure–function relationships. This channel has a large external vestibule, a narrow ion selectivity filter, and a large, water-filled central cavity lined by S6 segments. The ion selectivity filter is formed by negatively charged side chains of glutamate residues that remove most, but not all, waters of Na^+ hydration. An intracellular activation gate involves crossing parts of the S6 segments.

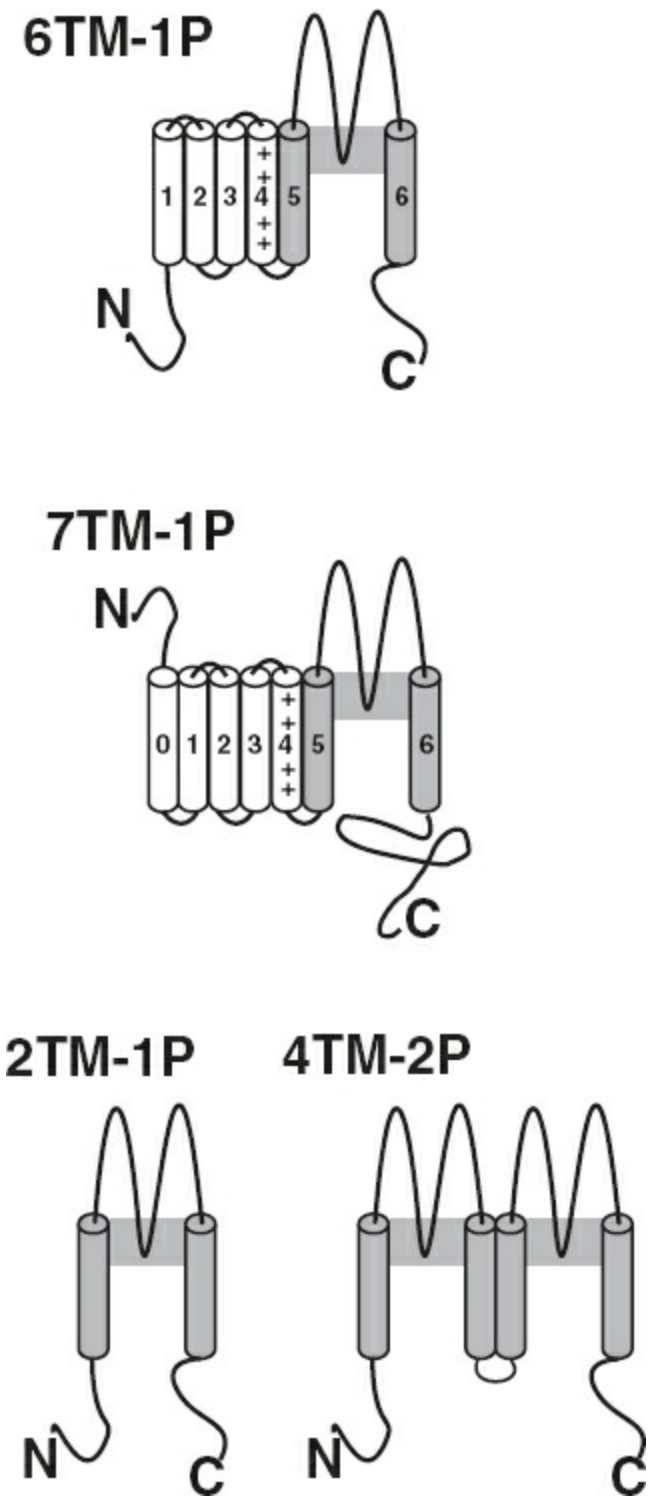

FIGURE 1.10–7. Potassium channels have diverse molecular structures. Four families of K^+ channels defined by the number of transmembrane domains and number of P-domains are shown. Of the channels described in the text, classical depolarization-gated K^+ channels and KCNQ subunits are members of the six transmembrane (TM) domain, 1 P-domain family. Some Ca^{2+}-dependent K^+ channels are also members of this family, but the BK Ca^{2+} and voltage-dependent channels are members of a separate family because of an extra transmembrane domain and an extracellular N-terminus. Inward rectifiers, including K_{ATP} and astrocyte leak channels, are members of the 2 transmembrane domain, 1 P-domain family. The tandem pore (or TWIK) family of K^+ channels has four membrane-spanning regions and two P-domains.

Sodium channel β-subunits (termed β_1 to β_4) are glycoproteins with a large extracellular amino terminus, a single transmembrane domain, and a short intracellular carboxy terminus. Two β-subunits associate with a single α-subunit and regulate functional expression, channel kinetics and gating. The large extracellular N-terminus is involved in cell adhesion via an immunoglobulin-like fold. A mutation in a cysteine residue in this extracellular fold is linked to a form of familial epilepsy.

Na^+ channels contain at least seven sites at which neurotoxins and clinical drugs act. Most, but not all, Na^+ channels contain an extracellular site at which tetrodotoxin (TTX) and saxitoxin (STX) block ion flow. TTX is a neurotoxin isolated from puffer fish that is used experimentally to block Na^+ channel function. At a site on Na^+ channels distinct from the TTX site, α-scorpion and sea anemone toxins act to modulate gating. α-scorpion toxins slow channel inactivation while β-scorpion toxins, acting at a distinct site, shift the voltage of activation and allow channels to open at voltages closer to the resting potential. The net effect of the scorpion toxins is to enhance excitation, contributing to the increased firing in pain fibers and paralysis associated with a scorpion sting. Mutations in α-subunits of skeletal muscle Na^+ channels cause the human disorder hyperkalemic periodic paralysis. Like anemone and α-scorpion toxins, these mutations slow channel inactivation.

Other toxins from the buttercup family (aconitine), the lily family (veratridine) and frogs used for arrow poisons in South America (batrachotoxin) promote direct opening of Na^+ channels and prolong the duration that channels stay open. The net effect is similar to the scorpion toxins. Finally, certain local anesthetic drugs, including lidocaine and procaine, block Na^+ channels by binding to sites within the hydrophobic regions of the ion channel. The blockade of Na^+ channels contributes to local anesthetic effects and to antiarrhythmic effects in the heart. Certain clinically important anticonvulsants (carbamazepine, lamotrigine, phenytoin, riluzole) bind a site similar to that bound by procaine, and have become important psychotherapeutic agents. All blockers mentioned, with the exceptions of TTX and STX, block Na^+ channels in a **use-dependent** manner. That is, the drugs become more effective as cells become more depolarized. Mechanistically, this may be achieved by drugs stabilizing an inactivated state (conformation) of the channel. This may lead to a clinically beneficial situation where normal CNS activity is relatively spared by the drugs, but abnormal hyperexcitation is blocked. Further clinical benefit from these drugs may result from their ability to dampen excitatory synaptic transmission selectively while sparing inhibitory transmission by mechanisms not fully understood. It is important to emphasize that not all Na^+ channels in neurons are sensitive to all of the above agents. For example, TTX-insensitive Na^+ channels exist in some excitable cells and

play a prominent role in certain nociceptive neurons.

Potassium (K^+) Channels

K^+ channels are the most diverse family of ion channels and determine the resting and firing properties of neurons. At least 75 different K^+ channels are expressed in various cells. Among these are 40 mammalian genes for voltage-activated K^+ channels grouped in 12 families according to their major (α) subunit (termed K_v1 to K_v12). The various families have distinct properties and structural motifs (Fig. 1.10–7). These include six-transmembrane domain/one P-domain channel subunits (classical depolarization-gated channels), inwardly rectifying channel subunits with two transmembrane domains, two-pore channel subunits with 4 transmembrane domains, and two-pore channel subunits with eight transmembrane domains (found in invertebrates). A fifth class of K^+ channels is a group of calcium and depolarization-gated channels known as BK (for their big single-channel conductance); these channels are similar to six-transmembrane domain channels but also contain a seventh membrane-spanning domain (S0 region) and an extracellular amino terminus. Adding to the diversity, there is evidence for both homo- and heteromeric K^+ channels. There are also several auxiliary subunits (β_1 to β_3) that associate with the α-subunits in a α4β4 stoichiometry. Based on crystallographic studies, K^+ channels have served as a model for relating protein structure to the functional properties of ion flow and channel gating. The diversity of potassium channels is daunting, and molecular and structural differences among these channels have broad functional implications. Perhaps more than any other class of ion channels, K^+ channels shape the pattern of membrane potential changes in response to input signals.

K^+ channels with six transmembrane domains are particularly important in neurons, with depolarization-gated channels representing a major subgroup within the family. The voltage-gated K^+ channel group is divided into subfamilies, K_V1 (sometimes called *Shaker* for the *Drosophila* gene), K_V2 (*Shab*) K_V3 (*Shal*), and K_V4 (*Shaw*), each with constituent subunits (e.g., $K_V1.1$ to 1.4). A functional channel contains four subunits (a tetramer) from within the same subfamily. Specific domains within the amino terminal regions are responsible for forming tetramers. Likewise, domains within the individual subunits modulate gating, inactivation, and interactions with accessory proteins.

A common structural motif among voltage-gated K^+ channels is the presence of six transmembrane domains (called S1 to S6) with a reentrant loop (P-domain) between S4 and S5. The reentrant loops of the four subunits coordinate to line the ion channel pore. Carbonyl groups in the P-

domain directly interact with K^+ and determine ion selectivity. These interactions mimic the hydration shell for K^+, and no water molecules intervene between K^+ and the carbonyl groups, differing from ion selectivity in Na^+ channels. The specificity of the interactions within the channel imparts selectivity for K^+ over other ions, and high efficacy channel permeation involves electrostatic ion–ion repulsion between adjacent K^+. Not surprisingly, the reentrant loop motif is common to all other families of K^+ channels, including the two P-domain families and inwardly rectifying K^+ channels.

Voltage-gated K^+ channels help to define the electrophysiological "signature," the characteristic spike shape and firing pattern, of neurons. The fast repolarization mediated by certain K^+ channels allows an increased rate of action potential firing that can then be used in frequency-dependent information coding (Fig. 1.10–8). Most neurons express multiple types of K^+ channels that differ in activation and inactivation kinetics, voltage dependence, and pharmacology. Because the equilibrium potential for K^+ is about −90 mV, the opening of K^+ channels results in K^+ efflux, membrane hyperpolarization, and a decrease in excitability.

The first K^+ channel identified was called a **delayed rectifier**. These channels were predicted in the experiments of Hodgkin and Huxley on squid giant axons and were so named because their currents activate more slowly than the Na^+ channels that produce the upstroke of the action potential (i.e., channel opening is "delayed" relative to Na^+ channel opening). A "rectifier" (diode) is an electrical device that passes current better at certain voltages than at others; in other words, conductance varies with voltage. The K^+ current is described as a rectifier because the channel is more effective in allowing K^+ ions to exit than to enter the cell, by virtue of depolarization opening the channels. Delayed rectifier channels show little inactivation during prolonged depolarization. These channels help to determine the frequency with which neurons fire action potentials. For instance, "fast-spiking" interneurons of the hippocampus and cortex possess a rapidly activating and deactivating delayed rectifier channel encoded by K_v3 family members. These channels are largely responsible for the narrow spike and brief refractory period of this interneuron class. Structurally, delayed rectifiers are members of the six-transmembrane K^+ channel subfamily.

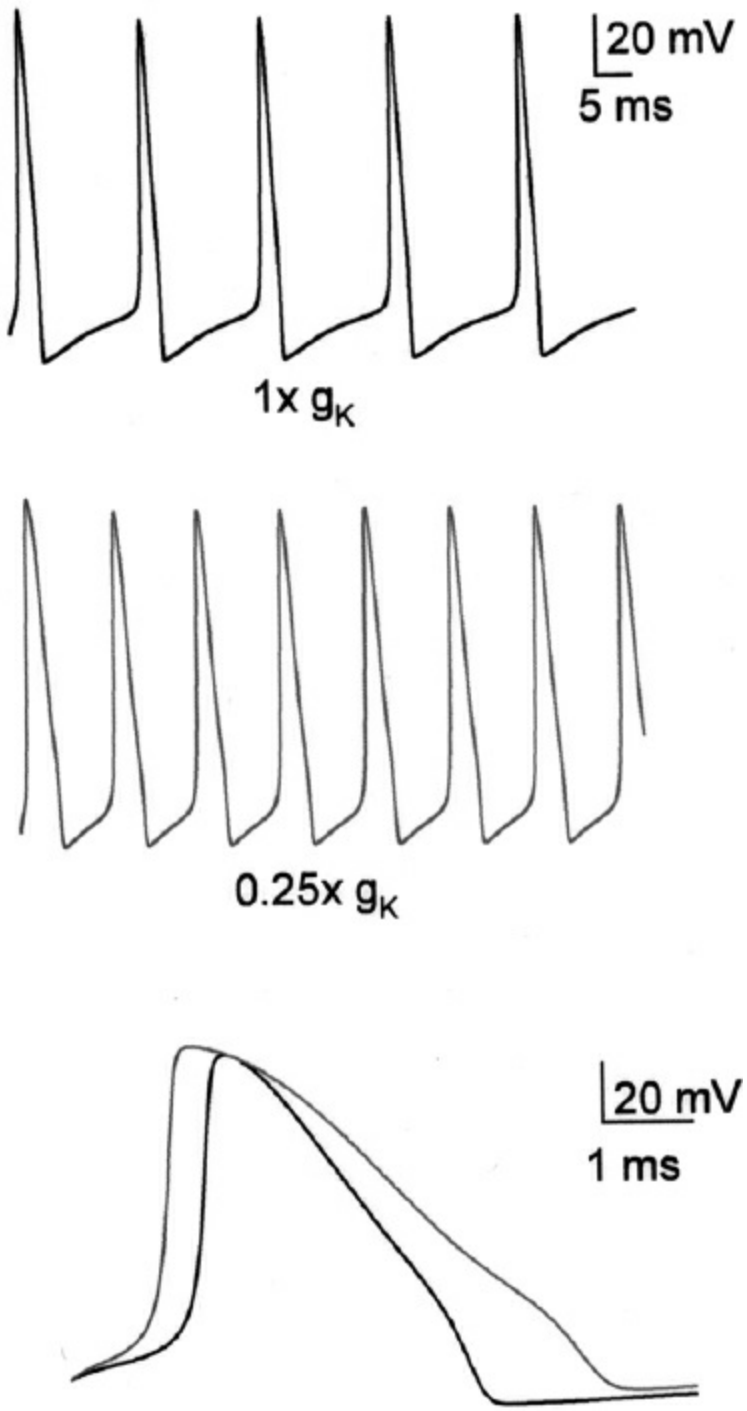

FIGURE 1.10–8. The traces show the effect of inhibiting K^+ channels involved in action potential repolarization. The traces were constructed using a simulation program that includes nongated leak channels and voltage-gated sodium and potassium channels similar to those described by Hodgkin and Huxley. The simulation includes a sustained depolarizing current injection of fixed amplitude to elicit spiking. The gray traces show the effect of reducing the delayed rectifier potassium conductance (g_K) to 25 percent of the initial baseline level (black traces). Note that with fewer voltage-gated potassium channels, the cell is hyperexcitable, exhibiting more action potentials to the same depolarizing input. The bottom panel shows the first 7 milliseconds of the simulations superimposed. Following K^+ channel block, individual action potentials are broadened and show a diminished undershoot.

Early experiments in squid axons indicated that delayed rectifier currents were the primary K^+ currents involved in action potential repolarization. In neurons, the situation is more complex, with rapidly activating K^+ channels contributing significantly. These include two classes of **calcium-activated K^+ channels** that are opened by increases in intracellular Ca^{2+} (some are also opened by depolarization). These channels are important in action potential repolarization and in generating an AHP characteristic of some neurons. In cells that possess it, this AHP is responsible for the **accommodation** (adaptation) that diminishes repetitive action potential firing during prolonged depolarization (Fig. 1.10–8). The AHP has several temporal components that are mediated by BK (big conductance) and SK (small conductance) calcium-activated K^+ channels. BK channels mediate the fast component of the AHP and have a large single channel conductance. BK channels belong to the SLO family of channels named for the *slowpoke* gene in *Drosophila* and include Ca^{2+}-, Na^+-, and H^+-sensitive K^+ channels. The large conductance of these channels results from a ring of negative charges in the inner and outer pore regions and a large inner pore. Some evidence suggests an association between a defect in the α-subunit of BK channels and autism with mental retardation. SK channels (SK1 to 3 in the CNS) are gated by intracellular Ca^{2+} but are voltage insensitive, and contribute to the medium and slow components of the AHP. Some evidence has associated SK channel variants with schizophrenia and other psychotic disorders.

A-Type K^+ **channels** rapidly activate with depolarization to potentials above −60 mV and rapidly inactivate at depolarized potentials. A-channels are involved in setting the interspike frequency with which neurons can fire and contribute to action potential repolarization. They are important for frequency coding, the encoding of input strength in the firing frequency of a neuron. The *Shaker* A-channel from *Drosophila* was the first K^+ channel cloned.

M-**Channels** are K^+ channels that are activated in a time- and voltage-dependent fashion and are inhibited by the neurotransmitter, acetylcholine, acting at muscarinic receptors. These channels are slow to activate and contribute little to action potential repolarization, but their activation at negative membrane potentials helps to slow repetitive firing. In pyramidal neurons, M-channels are expressed in perisomatic regions where they regulate synaptic integration and adaptation to repetitive firing. M-currents arise from heteromeric association of two members of the six-transmembrane domain group, KCNQ2 and KCNQ3 (members of the K_v7 subfamily) and possibly other KCNQ family members. Mutations in KCNQ subunits result in benign familial neonatal convulsions (BFNC), deafness, and a form of cardiac long QT syndrome. The name of the subunits in this class of channels (KCNQ) derives from their role in long QT syndrome.

Retigabine, an anticonvulsant, potently opens M-channels and slows their closing.

In the sea snail *Aplysia californica,* certain K^+ channels that contribute to action potential repolarization are inhibited by the neurotransmitter serotonin and are called S-**channels**. The activity of S-channels is diminished during acute behavioral sensitization of the gill-withdrawal reflex in *Aplysia,* and studies of the role of these channels in synaptic function provided important insights into the cellular basis of certain forms of learning and memory.

Some K^+ channels (K_{ir} family) are opened by hyperpolarization instead of depolarization. These **inward rectifiers** (anomalous rectifiers) allow K^+ to more easily enter than exit the cell, and strongly rectify near the equilibrium potential for K^+. The primary mechanism underlying rectification is channel block by positively charged intracellular magnesium or polyamines. This channel block is relieved by K^+ influx through the channel, accounting for rectification that follows the K^+ equilibrium potential. Despite their inward rectification, the physiological importance of these channels for neurons likely lies in passing small outward (hyperpolarizing) currents because neurons are rarely hyperpolarized beyond the K^+ equilibrium potential. Inward rectifiers are members of the two-transmembrane domain group of K^+ channels. This class of K^+ channels contains a single pore region and lacks a voltage sensor (Fig. 1.10–7).

Brain astrocytes and Muller (glial) cells of the retina express inwardly rectifying K^+ channels of the $K_{ir}4$ and possibly $K_{ir}2$ subfamily that may be responsible for buffering extracellular increases in K^+ during neuronal activity. Expression of the inward rectifiers in these cells is positioned to siphon K^+ away from axons and axon terminals where extracellular accumulation may occur. The unique gating properties of these channels favor influx of K^+.

An additional feature of K^+ channels is that certain neurotransmitters can alter their function via GPCRs. For example, acetylcholine, acting at muscarinic receptors, blocks several K^+ currents leading to enhanced excitability. At different muscarinic receptors, acetylcholine activates inwardly rectifying K^+ channels in a variety of tissues including heart and brain. In the hippocampus and other brain regions, the neurotransmitters GABA, serotonin, and adenosine indirectly open a class of inwardly rectifying K^+ channels. These channels allow divergent synaptic inputs to exert regulatory influences over neuronal firing through a single class of ion channels.

In peripheral tissues and in some neurons, a class of K^+ channels ($K_{ir}6$ subfamily) is regulated by intracellular ATP. These channels are members

of the two transmembrane domain inward rectifier class (Fig. 1.10–7). In the pancreas, K_{ATP} channels regulate insulin release and are a site of action of the hypoglycemic sulfonylurea drugs, tolbutamide and glibenclamide, used to treat diabetes mellitus. The hypoglycemic drugs promote insulin release by blocking ATP-sensitive K^+ channels. This leads to membrane depolarization, calcium influx, and hormone release. Diazoxide, an antihypertensive drug that increases blood glucose levels, has the opposite effect on pancreatic ATP-sensitive K^+ channels, opening the channels and diminishing the release of insulin. The sulfonylurea drugs do not interact directly with K_{ir} subunits of K_{ATP} channels, but rather bind to high-affinity sulfonylurea receptors (SURs) that combine with members of the $K_{ir}6$ family in heteromeric complexes of four K_{ir} subunits and four SURs. K_{ATP} channels are also expressed in the CNS and regulate the release of certain neurotransmitters and perhaps determine responses to changes in intracellular energy levels. These channels are also expressed intracellularly in mitochondrial membranes and may play a role in regulating apoptotic cell death.

HCN (hyperpolarization and cyclic-nucleotide-gated) or **H-channels** are a class of nonselective cationic channels structurally related to K^+ channels. These channels are strongly expressed in certain neurons, including dopaminergic neurons of the substantia nigra and VTA, cells implicated in motor behavior, arousal, attention, reward, motivation, and addiction. H-currents contribute to resting membrane potential and pacemaker firing in certain neurons. H-currents are also expressed in dendrites of some pyramidal neurons and, along with other classes of voltage-gated channels, modulate synaptic signals in dendrites. HCN channels are similar to K_{ir} channels in that they activate at hyperpolarized voltages and close with depolarization. HCN channels differ in being permeable to both K^+ and Na^+, providing a depolarizing current at membrane potentials near rest. H-channels help stabilize neuronal membrane potentials, resisting changes that depolarize or hyperpolarize the cell. The anticonvulsant and mood stabilizer, lamotrigine, activates dendritic H-channels in pyramidal neurons as one of its mechanisms of action, and the anesthetic, psychotomimetic, and rapidly acting antidepressant drug, ketamine, blocks these channels at concentrations relevant to its clinical effects. HCN channels have a structure that differs from K_{ir} channels but is similar to other voltage-activated channels with six membrane-spanning regions, an S4 voltage sensor, and a reentrant P-loop. H-channels have an intracellular cyclic nucleotide binding domain near the carboxy terminus, and the binding of cyclic AMP shifts the voltage range for channel gating.

The K^+ channels described above have either two- or six-

transmembrane regions and a single P-loop. Another class of K^+ channels has four-membrane spanning regions and two P-domains (Fig. 1.10–7). These **tandem pore** (K_{2P}) channels are widely expressed in the CNS and periphery, and serve, at least in part, as leak conductances that help to establish the resting membrane potential. Eighteen members of this family have been identified and it is believed that these proteins form functional dimers. The various K_{2P} channels vary in electrical properties and in sensitivity to activators (e.g., neurotransmitters, arachidonic acid, protons, heat, stretch) and modulators (second messengers). These channels go by a variety of names based on the TWIK (tandem pore weak inward rectifying K^+) channels that were the first cloned. For example, TWIK-related arachidonic acid-sensitive K^+ channels are called TRAAKs, while acid-sensitive channels are called TASKs. These channels are activated by certain volatile anesthetics and anticonvulsants (riluzole), contributing to CNS-depressant effects of these agents.

Compared to voltage- and transmitter-sensitive channels, members of the TWIK family are activated by a variety of novel stimuli including acidic pH, heat, and mechanical activity. Another class of channels, called ASICs (acid-sensing ion channels), is also responsive to these stimuli. ASICs have two transmembrane domains with an apparent reentrant loop between. At least nine mammalian family members falling into five subfamilies have been identified. Unlike TWIKs, ASICs are voltage-insensitive cation channels that are more permeable to Na^+ than to K^+ and least permeable to Ca^{2+}. While information about the function of these channels is limited, ASICs appear to form functional trimers and participate in peripheral sensory processing including touch, heat, taste, and pain, and also contribute to certain forms of long-term synaptic plasticity in the hippocampus and amygdala, where they are involved in fear-associated learning.

Calcium (Ca^{2+}) Channels

Ca^{2+} serves as both an important messenger regulating intracellular chemistry (enzyme activity, metabolism, transmitter release) and as an electrical signal for membrane depolarization. In some neurons, Ca^{2+} influx can underlie action potentials ("calcium spikes"). Additionally, excessive and prolonged increases in intracellular Ca^{2+} contribute to neurodegeneration. Under some conditions, prolonged deficiency of Ca^{2+} influx may also lead to neuronal death, particularly in developing neurons. These features make the regulation of intracellular Ca^{2+} levels vital to cellular function and survival.

Voltage-activated Ca^{2+} channels provide a major source of the Ca^{2+} signals that regulate cellular function, and, in conjunction with Ca^{2+}-

permeable ligand-gated ion channels (e.g., NMDA-type glutamate receptors and neuronal nicotinic acetylcholine receptors), represent major conduits for Ca^{2+} entry from the extracellular environment. Neurons possess multiple classes of voltage-gated Ca^{2+} channels that are classified by biophysical and pharmacological properties. Some Ca^{2+} channels are activated by relatively small depolarizations over the range of −80 to −50 mV and are called **low-voltage–activated** (LVA, Ca_v3) Ca^{2+} **channels**. These LVA channels inactivate rapidly and are relatively insensitive to dihydropyridine Ca^{2+} channel blockers, such as nifedipine and nimodipine. LVA channels are also called T-**type** Ca^{2+} **channels** because of their "transient" (inactivating) currents. Because LVA Ca^{2+} channels are activated at membrane potentials near rest, these channels contribute to burst firing and oscillatory neuronal activity. Oscillatory neuronal firing is important in driving coordinated movements and in maintaining complex behavioral states such as wakefulness. LVA channels are also expressed in neuronal dendrites and contribute to synaptic integration and spike-timing–dependent synaptic plasticity. With a few exceptions, LVA channels do not participate in neurotransmitter release. LVA channels may be a target for the actions of some antipsychotic drugs, and the diphenylbutylpiperidines, pimozide and penfluridol, inhibit LVA channels at concentrations similar to those affecting D2 dopamine receptors. Other antipsychotics also inhibit T-type channels but do so at concentrations above those required at dopamine receptors.

A second class of Ca^{2+} channels, called **high-voltage–activated** Ca^{2+} **channels** (HVA, Ca_v1, and Ca_v2 family members), is activated by stronger membrane depolarizations to potentials positive to −50 mV. In many neurons, even when Na^+ channels involved in the upstroke of action potentials are blocked, HVA Ca^{2+} channels can produce regenerative spikes. These calcium spikes are slower in onset and longer in duration than Na^+ spikes, reflecting the kinetics of HVA channels.

HVA Ca^{2+} channels are heterogeneous with several major subtypes. L-**type** Ca^{2+} **channels** (Ca_v1 family; named for their "long-lasting" responses) show slow inactivation during sustained depolarizations and are sensitive to block by dihydropyridines. L-type channels provide sufficient Ca^{2+} influx during action potentials to activate Ca^{2+}-dependent second messengers and gene expression. N-**type** Ca^{2+} **channels** ($Ca_v2.2$ named because they were "neither" L- nor T-type) are HVA channels involved in regulating neurotransmitter release from some presynaptic structures, particularly those of sensory neurons in the retina and inner ear. N-type channels are blocked irreversibly by ω-conotoxin GVIA, a poison from the snail *conus geographus*. Another conotoxin, ω-conotoxin MVIIA (Prialt), is a reversible N-channel blocker that is used to treat chronic pain. P/Q-**type**

Ca^{2+} **channels** ($Ca_v2.1$) are a third class of HVA channels and are named because of their presence in Purkinje cells of the cerebellum and pyramidal neurons of the hippocampus and cortex. P/Q-channels are insensitive to dihydropyridines and ω-conotoxin GVIA but are blocked by ω-Aga-IVA, a toxin from the funnel web spider *Agelenopsis aperta*. P/Q channels, like N channels, regulate the release of neurotransmitters in the CNS. N- and P/Q-type channels have an intracellular loop between channel domains II and III that binds certain presynaptic proteins including syntaxin 1, Rim, and synaptotagmin 1. This site is referred to as a "synprint" region and plays a role in allowing synaptic proteins to modulate channel function.

R-Type Ca^{2+} **channels** ($Ca_v2.3$) also contribute to CNS function, but their actions are less well understood. R-type channels are resistant to the Ca^{2+} channel antagonists described above, but are inhibited by SNX-482, a toxin derived from the African tarantula *Hysterocrates gigas*. Although their presynaptic role is not as prominent as N and P/Q channels, R-channels participate in transmitter release at fast excitatory synaptic synapses in the CNS.

The cloning of specific Ca^{2+} channel subunits has provided insights into the structural mechanisms of these channels and highlighted even further complexity than outlined above. Skeletal muscle HVA Ca^{2+} channels were the first cloned and serve as a model for understanding the structure of other voltage-gated Ca^{2+} channels. These channels are involved in excitation–contraction coupling at neuromuscular junctions and consist of five distinct subunits termed α1 (165 to 195 kDa), α2 (~150 kDa), δ (17 to 25 kDa), β (50 to 60 kDa), and γ (25 to 35 kDa) arranged in a 1:1:1:1:1 stoichiometry. The δ-subunit is produced by cleavage of α2/δ peptides while the other subunits are encoded by separate genes. α1-subunits have about 30 percent sequence homology to voltage-gated Na^+ channels and form the ion channel pore. A recurring theme in α1-subunits is the existence of four homologous internal repeats that each contains six putative membrane-spanning regions and a pore-forming P-loop (Figs. 1.10–5 and 1.10–6). The HVA α1 Ca^{2+} channel subunit from skeletal muscle contains the dihydropyridine binding site. Point mutations in the α1-subunit of skeletal muscle T-tubule Ca^{2+} channels ($Ca_v1.1$) and the skeletal muscle Na^+ channel ($Na_v1.4$) cause the human disorder, hypokalemic periodic paralysis. The critical mutations occur in the S4 region of the channel involved in voltage sensing and result in a gain of function current that is open at rest and results in membrane depolarization and action potential failure. Similar gating pore mutations may occur in other channelopathies. The functions of the β (β_{1-4}) and γ ($\gamma_{1\ to\ 8}$) subunits are less certain, but appear to regulate membrane expression and trafficking of α1-subunits. Loss of the γ_2-subunit (*stargazin*) in the stargazer mutant

mouse markedly diminishes cell surface expression of certain glutamate receptors. The importance of auxiliary subunits in the function of Ca^{2+} channels is highlighted by recent evidence that analgesic effects of the anticonvulsants gabapentin and pregabalin involve binding to specific residues in α_2–δ_1 subunits.

α1-subunits differ structurally among the different Ca^{2+} channel subtypes and 10 different α1 genes have been cloned. These include four different α1 subtypes contributing to L-type channels (Ca_v1.1, 1.2, 1.3, 1.4), three α1 genes contributing to P/Q-, N-, and R-type channels (Ca_v2.1, 2.2, 2.3), and three variants of T-type Ca^{2+} channels (Ca_v3.1, 3.2, 3.3). Adding to this complexity, human genes typically go by other names. For example, the skeletal muscle L-type channel Ca_v 1.1 is referred to as α1S and CACNA1S; other human calcium channel genes are named accordingly. The importance of calcium channels to neuropsychiatric disorders is highlighted by the finding that mutations in the human CACNA1A gene encoding the α1 subunit of P/Q calcium channels are associated with familial hemiplegic migraine.

An important question in ion channel biology concerns how channels establish selectivity for one ion over another. At an initial level, ionic charge and charged amino acid residues present in ion channel proteins help to select for cations over anions. It is a more difficult and complex problem, however, for channels to select among different cations. In the case of Ca^{2+} channels, this is particularly vexing because hydrated Ca^{2+} ions are significantly larger than Na^+ or K^+ ions. Thus, size of the ion channel pore cannot determine selectivity for Ca^{2+}. Rather, selectivity in Ca^{2+} channels involves high-affinity binding of divalent cations within the ion channel. When Ca^{2+} is present, its binding within the pore excludes monovalent cations, rendering the channels selective for Ca^{2+}. As might be expected, when Ca^{2+} is not present, these channels will pass monovalent cations. This principle of ions binding to specific sites within a channel to regulate permeability and gating is a recurring theme in ion channel biology that can sometimes be exploited for the development of drugs that modulate channel function.

In some regions of the CNS, particularly retinal photoreceptors and olfactory epithelial cells, intracellular cyclic nucleotides (e.g., cyclic AMP and cyclic GMP) gate specific classes of ion channels. These cyclic nucleotide–gated (CNG) channels have structural features similar to voltage-gated channels, including the presence of six membrane-spanning regions and a P-loop that lines the ion channel. CNG channels also have an S4-like voltage sensing region, although the channels are not regulated by voltage. Three α and three β CNG subunits have been cloned. CNG channels are nonselectively permeable to cations, but like voltage-gated

calcium channels bind divalent cations in the extracellular pore region. The binding of divalent cations restricts the flow of monovalent cations through CNG channels much like voltage-activated Ca^{2+} channels rendering them relatively selective for Ca^{2+} over Na^{+}. In some respects, CNG channels are similar to HCN channels, but are described in this section because of their higher calcium permeability. Mutations in retinal photoreceptor CNG channels contribute to color blindness in humans.

The TRP **superfamily** represents another class of cationic channels with six membrane-spanning regions and high calcium permeability. This family contains more than 25 members in at least seven subfamilies (TRPC, TRPV, TRPM, TRPML, TRPP, TRPA, and TRPN) that participate in multiple processes in the nervous system and in nonexcitable cells, ranging from sensory processing to vascular and cell cycle control. TRP channels are named for the first member to be identified, the *trp* (transient receptor potential) gene in *Drosophila,* and have been linked to several human disorders including polycystic kidney disease and mucolipidosis, a neurodegenerative illness. TRP channels are regulated by intracellular and extracellular signals including changes in pH, temperature, capsaicin (the active ingredient in hot peppers) and anandamide (an endogenous ligand for cannabinoid receptors). Endogenous lipid mediators like the endocannabinoid, anandamide, are important regulators of the TRPC/VM subfamilies. TRP family members (TRPC1 and TRPC4) may also contribute to "store-operated" channels that mediate extracellular calcium influx following calcium release from intracellular stores. However, other proteins, Orai and Stim1, are more essential components of this source of Ca^{2+} influx.

Chloride (Cl^{-}) Channels

In most neurons, Cl^{-} is present at higher concentrations outside than inside cells, and the equilibrium potential for Cl^{-} is near the resting membrane potential. Thus, the opening of Cl^{-} channels tends to keep the membrane potential near rest, and in conjunction with K^{+} channels, serves as a mechanism to dampen neuronal excitability. Some Cl^{-} channels also conduct bicarbonate, another physiological anion. Cl^{-} channels contribute to the resting membrane potential in certain neurons and muscle cells. These channels are spontaneously open at resting membrane potentials and exhibit weak voltage dependence. In muscle fibers, the background Cl^{-} conductance is the largest resting conductance and the distribution of Cl^{-} is near equilibrium. Cl^{-} channels fall into three main gene families, voltage-gated Cl^{-} channels (ClCs), the cystic fibrosis transmembrane conductance regulator (CFTR), and ligand-gated Cl^{-} channels. The ClC family includes nine different channels that have a structure unlike any cation channel.

ClCs have 10 or 12 membrane-spanning regions and, while several of the transmembrane regions participate in ion channel pore formation, there is no defined S4 voltage sensing region as in voltage-gated cation channels. Crystallographic studies indicate that ClC channels are double-barreled homodimers with each subunit containing its own pore. In the human illness, myotonia congenita, an abnormality of a muscle Cl^- channel (ClC-1) results in abnormally low Cl– conductance leading to increased muscular excitability and fatigue with exercise.

There are at least five activating stimuli for Cl^- channels. These include changes in membrane voltage (depolarization or hyperpolarization), increases in intracellular Ca^{2+}, ligand binding (usually GABA or glycine), cellular swelling, and phosphorylation by cyclic-AMP–dependent protein kinase (PKA). Ca^{2+}-activated Cl^- channels may help to determine the interspike frequency of neuronal firing, while swelling-activated channels help to protect cells from damage during osmotic stress. These channels are pentamers with a single ion pore and a hydrophobic neck that forms the channel gate controlling ion flow. Ca^{2+} binds to a large cytosolic region of the protein to regulate channel opening. Cl^- channels also serve important functions in secretory cells, providing a major source of Cl^- in tears, sweat, and digestive juices. A defect in CFTR secretory Cl^- channels that renders the channels insensitive to activating stimuli is important in cystic fibrosis. CFTR has a structure that differs from the ClC family and belongs to a larger family of ATP-binding cassette (ABC) proteins that require phosphorylation by PKA and hydrolysis of ATP for activation. CFTR is the only member of the ABC family that also serves as a Cl^- channel. Structurally, CFTR has two repeats with six membrane-spanning regions (12 transmembrane regions in total), a nucleotide binding domain and a regulatory domain.

Mitochondrial membranes express unique voltage-dependent anion channels (VDACs) that pass negatively charged ions and have unusual gating properties, being open at potentials near 0 mV and closing with voltage changes in either direction. VDACs (or **porins**) participate in releasing metabolites from mitochondria and are components of the mitochondrial permeability transition pore (PTP) that regulates apoptotic cell death. PTP is a multiprotein complex that includes adenine nucleotide translocase, hexokinase, cyclophilin D, and a VDAC. PTP activity is regulated by peripheral-type benzodiazepine receptors present on mitochondria. VDACs may also be expressed in plasma membranes, particularly in postsynaptic densities, and these channels may complex with some neurotransmitter receptors, although the function of plasma membranes VDACs remains uncertain. Three VDACs (VDAC 1 to 3) have been cloned and are β-sheet proteins that differ from the α-helical configuration of most ion channels.

NEUROTRANSMITTERS AND ION CHANNELS

Classes of Neurotransmitters

Most information transfer between neurons occurs via chemical synapses. These connections use a wide array of chemical messengers (**neurotransmitters**) that are released in a Ca^{2+}-dependent fashion from presynaptic terminals and act on specific membrane proteins (**receptors**) to produce biochemical and excitability changes in the receiving cell. There are two primary groups of neurotransmitters—low–molecular-weight amines and neuroactive peptides. These agents act on two classes of receptors, ligand-gated ion channels, at which binding of the transmitter directly opens an ion channel, and GPCRs, where an activated G-protein modulates ion channels or alters biochemical **second messenger systems**. Physiologists classify synaptic transmission according to the speed of transmission (fast or slow) and according to the excitability change (excitatory or inhibitory). **Fast** transmission occurs on a time scale of up to several 100 milliseconds and is mediated primarily by amine neurotransmitters acting at ligand-gated ion channels. **Slow** synaptic communication occurs on a scale of seconds to minutes or longer, usually through the actions of amine or peptide transmitters acting on GPCRs. The specific ion channels involved and the electrochemical gradients of the permeant ions determine whether effects are **excitatory** (depolarizing) or **inhibitory** (hyperpolarizing). Moreover, an excitatory synaptic input can exert a net inhibitory influence on regional excitability and vice versa. For example, release of an excitatory neurotransmitter onto an inhibitory neuron can stimulate the inhibitory neuron and dampen activity of a local population of cells. Conversely, inhibition of inhibitory neurons can enhance regional excitability. These features provide great flexibility in regulating circuit activity.

There are at least nine low-molecular-weight amines that serve as neurotransmitters. These include **glutamate**, the major fast excitatory transmitter in the mammalian brain, **acetylcholine**, the excitatory transmitter at the vertebrate neuromuscular junction, GABA, and **glycine**, the major fast inhibitory transmitters in the brain and spinal cord, respectively, and the biogenic amines **dopamine, norepinephrine, epinephrine, serotonin**, and **histamine**. Purines, ATP, can also act as transmitters. A large number of **neuroactive peptides** modulate neuronal excitability. It is uncertain, however, whether all of these substances function as neurotransmitters. Many peptides, including vasopressin and CCK, were first identified as hormones in the vasculature and gut. ATP and certain neuroactive peptides coexist with amine neurotransmitters in some nerve terminals, and can be coreleased. Thus, interactions between classes of neurotransmitters may determine the ultimate effects of a presynaptic neuron on its postsynaptic target.

Conductance Mechanisms Underlying Neurotransmitter Actions

Neurotransmitter actions can also be described by effects on membrane conductances. Transmitters that act at ligand-gated ion channels increase the conductance of cell membranes to specific ions. Excitatory transmitters, such as acetylcholine and glutamate, directly activate nonselective cation channels, increasing the flow of Na^+, K^+, and in some cases, Ca^{2+}. Because of the mixed permeability of these channels, the reversal potential of the currents generated by the channels is a weighted average of the equilibrium potentials of the individual permeant ions, typically near 0 mV. Thus, at membrane potentials negative to 0 mV, channel opening depolarizes (excites) the cell by allowing inward flow of cations.

In contrast, the fast inhibitory transmitters, GABA and glycine, open ligand-gated channels permeable to Cl^-. Because Cl^- usually has an equilibrium potential near the resting membrane potential, opening of these channels "clamps" the membrane potential negative to spike threshold, decreasing (inhibiting) the likelihood of spike initiation and shunting excitatory influences. It can be appreciated, however, that the excitatory or inhibitory nature of transmitter actions is dictated by the electrochemical gradients of the permeant ions. Early in development, many neurons possess a sufficiently high intracellular Cl^- concentration that the Cl^- equilibrium potential is positive relative to spike threshold, resulting in excitatory effects of GABA and glycine. With development and changes in the expression of specific Cl^- transporters, intracellular Cl^- levels diminish, and GABA and glycine become inhibitory transmitters. There are pathological conditions, including behavioral stress, that alter the expression of Cl^- transporters. Under these conditions, GABA (and Cl^- channel opening in general) becomes excitatory. This is thought to contribute to the epileptogenesis associated with human brain tumors.

A second group of transmitters indirectly increases membrane conductance through G proteins. For example, GABA, serotonin, and adenosine, by binding GPCRs, promote G-protein–mediated opening of inwardly rectifying K^+ channels (GIRKs). A third set of transmitter actions involves indirect effects on voltage-gated or leakage ion channels. These transmitters typically decrease membrane conductance by activating chemical second messenger systems via GPCRs. Certain voltage-gated K^+ and Ca^{2+} channels are specific targets of this inhibition, resulting in cellular excitation or inhibition, respectively. Most transmitters that act on GPCRs exert at least some of their effects by these decreased conductance mechanisms. The electrical principles underlying synaptic excitation or inhibition are identical to those described for other ion channels and are based on the relative permeabilities of the ion channels and the equilibrium

potentials of the ions involved.

Several transmitters (e.g., GABA, glutamate, acetylcholine, and serotonin) act at both ligand-gated ion channels and GPCRs. Thus, receptors for these neurotransmitters and their effects are heterogeneous, with the nature of the effect depending on the specific receptor to which the transmitter binds. Cloning studies have demonstrated that receptors for most neurotransmitters are structurally complex with multiple receptor subtypes being the rule rather than the exception. At the receptor level there is tremendous flexibility in determining the effects of a neurotransmitter on a single neuron or on a set of neurons.

Structure of Neurotransmitter Receptors

Most transmitter-gated ion channels are multimeric proteins consisting of several (usually four or five) subunits that have multiple (two to five) membrane-spanning regions (Fig. 1.10–9). Functional receptors typically have large amino-terminal regions that extend into the aqueous extracellular environment and are sites at which neurotransmitters bind and at which sugar molecules attach to the receptor (glycosylation sites). Receptor glycosylation likely participates in determining optimal conformations for channel gating. The intracellular regions of the receptor often contain sites at which phosphate groups can be attached. Phosphorylation is a mechanism by which second messenger systems modulate receptor and ion channel function. Many receptors are multiprotein complexes in which the receptor subunits that comprise the ion channel pore are in physical proximity with intracellular proteins (in some cases 100 or more intracellular proteins). The intracellular proteins regulate receptor trafficking and expression as well as ion channel function and participation in intracellular processes.

The neuromuscular nicotinic acetylcholine receptor was the first transmitter-gated channel to be purified and cloned, and five neuromuscular nicotinic receptor subunits have been identified. Each of these subunits has four membrane-spanning regions and a pair of cysteine residues located 13 amino acids apart in the extracellular region of the protein (Fig. 1.10–9). These cysteine residues form a disulfide bridge that contributes to transmitter binding. The nicotinic ion channel is a nonselective cation channel permeable to Na^+, K^+, and Ca^{2+}. Membrane-spanning regions form the ion channel with the second transmembrane region lining the channel pore. Muscle nicotinic subunits assemble to form a pentamer with a stoichiometry of α_2, β, δ, and γ or ε depending on the age of the animal. Nicotinic receptors are part of a superfamily that includes $GABA_A$, glycine and serotonin-type 3 (5HT-3) receptors. $GABA_A$ and glycine receptors are anion-selective, passing primarily Cl^- in physiological solutions, whereas nicotinic and 5HT-3 receptors are cation selective.

Because all members of the family contain the characteristic extracellular cysteine residues, the family is referred to as the **cys-loop** family. Differences in charges on amino acids at the entrance to the channel pore determine whether the channel passes cations or anions. A recent crystal structure of a human homomeric β3 $GABA_A$ receptor shows the GABA channel to have a cylindrical structure about 110 Å high and 60 to 80 Å wide, extending 65 Å into the extracellular space. A positively charged ring of amino acids halfway down the channel allows anion binding and determines ion selectivity. A closed gate at the base of the pore provides a mechanism for desensitization. Crystal structures are extremely important because they offer the hope of more rational and selective drug design.

Ligand-gated ion channels gated by extracellular ATP (P2X receptors) are exceptions to the scheme outlined above and have structures more similar to inwardly rectifying K^+ channels (Figs. 1.10–7 and 1.10–9). ATP receptors have two membrane-spanning regions and a pore-forming region (P-loop) that are connected by a large loop of extracellular amino acids. A major difference between P2X receptors and inwardly rectifying K^+ channels is that P2X receptors are mostly extracellular, whereas the bulk of the K^+ channel is intracellular. P2X channels are cation selective and have high Ca^{2+} permeability. These receptors participate in fast excitatory synaptic transmission in certain regions of the CNS including the thalamus. Native ATP receptors may consist of combinations of P2X subunits.

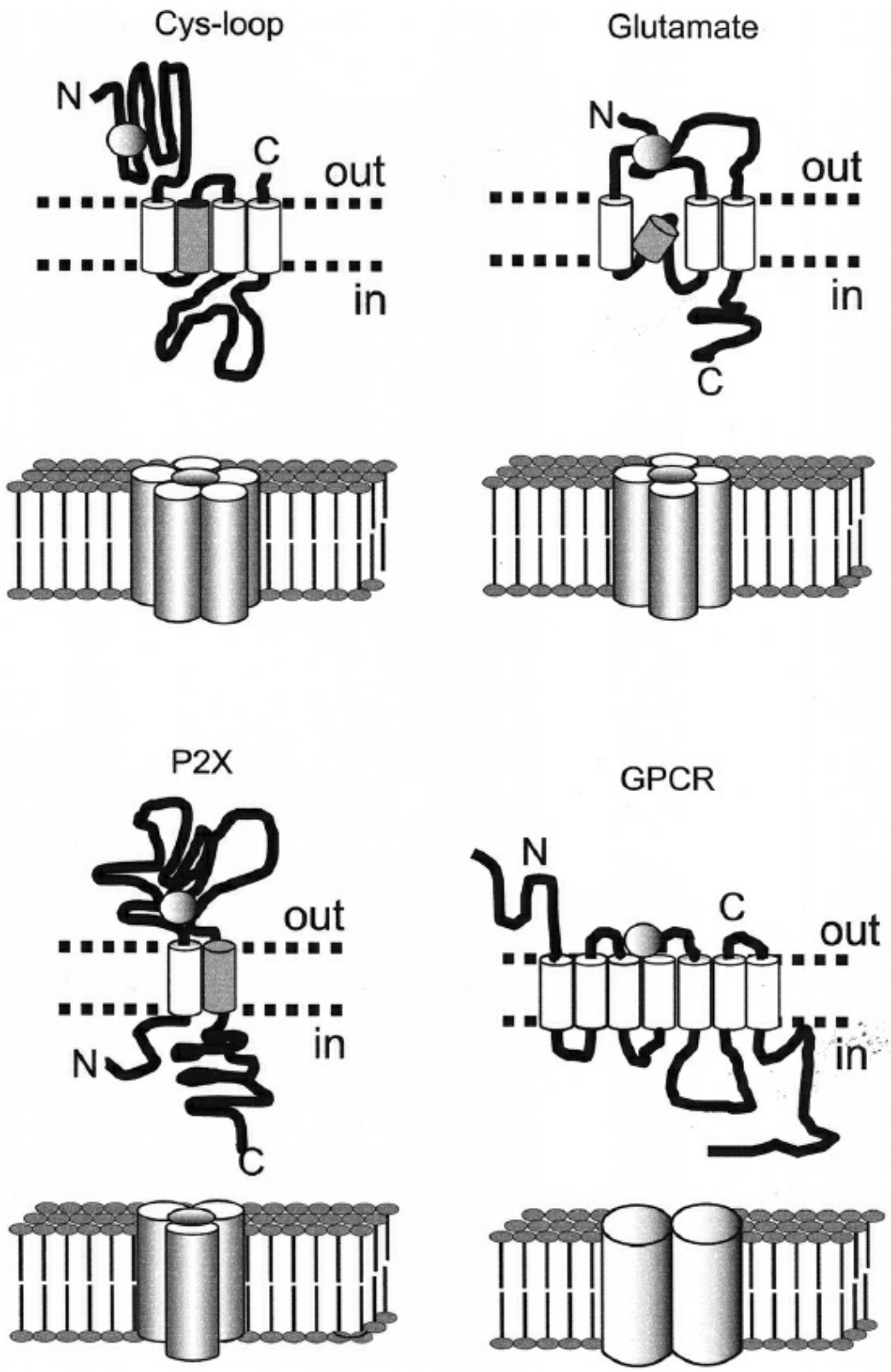

FIGURE 1.10–9. The diagrams show the proposed secondary structure of receptors for several neurotransmitters, including a $GABA_A$ receptor (member of the Cys-loop family), an ion-channel–linked glutamate receptor, a channel gated by extracellular ATP (P2X) and a G-protein–coupled receptor (GPCR). The ligand binding domains of these receptors are depicted by the circles in the extracellular regions.

Ionotropic glutamate receptors are also exceptions to the structural

scheme proposed for $GABA_A$ and nicotinic receptors. Glutamate receptor subunits have three membrane-spanning regions and a reentrant P-loop between the first and second transmembrane regions that does not completely cross the membrane (Fig. 1.10–9). The P-loops in glutamate-gated channels like those in voltage-gated ion channels line the ion channel. A difference from voltage-gated channels is that the glutamate receptor P-loops enter the membrane from the cytoplasmic side. Crystallographic data on homomeric GluA2 AMPA-type glutamate receptors indicate that these channels contain four subunits, each of which binds a glutamate molecule. Subunits form a dimer of dimers, and the receptor resembles a letter "Y" when viewed perpendicular to the membrane. The glutamate binding region has a hinged cleft resembling a venus flytrap. When agonists bind, the cleft closes to varying degrees depending upon the ligand. This imparts structural changes necessary for ion channel opening. Competitive antagonists that block glutamate binding stabilize the open cleft configuration of the binding pocket, providing an explanation for how an agent can bind to the agonist recognition site but not produce the conformational changes that cause channel opening. A recent crystal structure for heterotetrameric GluN1-GluN2B NMDA receptors has also been generated and reveals differences from AMPA receptors. The crystallized NMDA receptor has a much more tightly packed amino terminal domain and ligand-binding domain, possibly explaining why the amino terminus more effectively regulates ion channel activity in these receptors. In contrast to the Y-like shape of AMPA receptors, NMDA receptors better resemble a hot air balloon with an extracellular "balloon" and a transmembrane "basket."

GPCRs have structures that differ completely from ligand-gated ion channels. These receptors have seven membrane-spanning regions (Fig. 1.10–9), and many GPCRs exist as homo- or heterodimers. Transmitter binding occurs in a pocket formed by the intramembranous domains of the receptor, except for the glutamate family of GPCRs, where binding is in a large extracellular amino terminal region. The coupling of GPCRs to their G protein occurs at intracellular loops of the receptor. GPCRs also have sites for glycosylation and phosphorylation.

CLINICAL ASPECTS OF ION CHANNELS

CNS information processing depends critically upon ion channels. Most rapid processing involves action potential firing and fast neurotransmission. While it is beyond the scope of this chapter to detail all clinical arenas in which ion channels are important, this chapter highlights some areas where the function of ion channels is particularly important to psychiatry.

Oscillatory Neuronal Firing and Complex Behavioral States

Complex behavioral states, including wakefulness, attention, mood, and sleep require coherent activity within and between neuronal circuits. Neuronal networks show rhythmic activity (oscillation) over a wide range of frequencies from less than 1 to several 100 cycles/s (Hz). These oscillations provide energy efficient mechanisms that determine optimal input and output frequencies and provide coherence in functional networks. Activity in oscillating circuits involves interplay of intrinsic electrical properties of neurons, local balance of excitatory and inhibitory inputs (E/I balance), and sustained effects of more diffusely acting neuromodulator systems including muscarinic and monoaminergic systems. With regard to intrinsic electrical properties, certain neurons have voltage-gated conductances that allow them to fire rhythmically and spontaneously, thus having properties of a **pacemaker** or **oscillator**. For example, inferior olivary neurons fire action potentials spontaneously and sustain this firing for relatively long periods in the absence of external inputs. These neurons fire conventional fast Na^+ spikes that provide the depolarization needed to open high voltage-activated (HVA) Ca^{2+} channels. In turn, Ca^{2+} influx activates a Ca^{2+}-dependent K^+ conductance that rapidly hyperpolarizes the membrane. When the membrane hyperpolarizes, LVA Ca^{2+} channels open and depolarize the membrane potential to threshold for firing Na^+ spikes, which then activates another cycle. In the case of inferior olivary neurons, it is the properties of the LVA Ca^{2+} channels that foster oscillatory firing. LVA channels are inactivated at the neuronal resting membrane potential but become activatable when the membrane is hyperpolarized with respect to rest. In effect, hyperpolarization becomes a priming stimulus that allows LVA channels to open. Pacemaker firing of inferior olivary neurons drives Purkinje neurons in the cerebellum to **resonate** at the inferior olivary neuron's preferred firing frequency. This resonating circuit contributes to the physiological resting tremor that oscillates at about 10 cycles per second.

Pacemaker activity is also found in thalamic neurons where similar, though not identical, mechanisms are used. In the thalamocortical system, network activity is mediated by both intrinsic neuronal conductances and synaptic connections. This activity drives EEG changes during different stages of sleep and vigilance. Thalamocortical neurons exhibit two distinct activity states. During sleep, neurons show synchronized rhythms that resemble delta (1 to 3 Hz activity) and other slow waves on the EEG. During wakefulness and REM sleep, these neurons exhibit tonic activity. LVA calcium channels are important participants in thalamocortical network activity. The transition from sleep to wakefulness is mediated by depolarization of thalamic reticular neurons and inactivation of LVA calcium channels. Specific abnormalities in thalamocortical neurons may also be critical in the generation of 3 Hz spike and wave activity observed in

childhood absence epilepsy. In spike-wave discharges, interplay between LVA calcium currents and H-currents appears to be critical in generating abnormal firing. Drugs targeted at these channels (e.g., ethosuximide, an inhibitor of LVA channels and lamotrigine, an activator of H-channels) are useful clinically. Mutations in $Ca_v3.2$ LVA channels have been associated with absence of epilepsy.

In some brain regions, the outputs of pacemaker cells are mediated by fast excitatory or inhibitory transmitters. However, some neurons are capable of firing bursts of action potentials. **Bursts** are periods of frequent spike firing followed by quiescent periods. This type of firing can be used to drive activity in a local or distributed neural network. Additionally, burst-like firing can provide sufficient intracellular Ca^{2+} to stimulate release of peptide transmitters. In turn, the slow synaptic actions of the peptides in combination with or independent of other GPCR systems can alter the frequency of oscillatory firing and bursting. An example of this is the repeated firing that occurs when spike frequency adaptation is inhibited by blocking Ca^{2+}-activated K^+ conductances. In this case, both the intrinsic electrical properties of neurons and the effect of modulatory transmitters conspire to determine a background level of activity (or tone) in specific neuronal systems.

Ion Channels and Neuropsychiatric Disorders

There is increasing evidence that defects in ion channels (called **channelopathies**) contribute to the pathophysiology of neuropsychiatric illnesses. Channelopathies typically involve altered function of specific ion channels that result from genetic mutations, transcriptional abnormalities, or immune processes. Although many of these illnesses are not purely "psychiatric" disorders, the involvement of specific ion channels in illnesses is instructive for understanding the importance of ion channels in physiology and illnesses.

Given the importance of ion channels in determining the excitability of neurons, the observation that epilepsy, a group of disorders characterized by recurrent bouts of abnormal paroxysmal electrical activity, is associated with mutations in specific ion channels is not surprising. For example, BFNC are associated with mutations in K^+ channel genes, KCNQ2 and KCNQ3. These proteins ($K_v7.2$, 7.3) form M-channels that produce slowly activating and slowly inactivating K^+ currents that blunt neuronal firing. The expression of a single mutant allele of either gene—BFNC is a genetically dominant disorder—decreases channel number sufficiently to result in neuronal hyperexcitability. Generalized epilepsy with febrile seizures plus (GEFS+) is an autosomal dominant syndrome that results from mutations in Na^+ channel subunits. The role that ion channels play in these rare genetic syndromes suggests that idiopathic epilepsy might also

be a channelopathy resulting from an interaction between genetic defects in ion channels and adverse environmental effects. A recent exome sequencing study examined more than 230 channel genes in sporadic idiopathic epilepsy, the most common neuronal excitability disorder. This study reported rare missense variants in both epilepsy patients and unaffected individuals, including deleterious mutations in both groups. These results strongly suggest that the ion channel genes confer uncertain illness risk and that illness expression depends on other variants, even though the ion channel defects likely contribute. Identification of defective channel genes, nonetheless, offers an opportunity to improve clinical outcomes based on individualized treatment.

The calcium channel CACNA1A gene, which encodes the α_1-subunit of HVA P/Q-type Ca^{2+} channels ($Ca_V2.1$), is associated with several rare genetic diseases. Familial hemiplegic migraine is an autosomal dominant form of migraine with childhood onset and an aura that includes transient hemiparesis or hemiplegia lasting hours to days. Otherwise, the headache is indistinguishable from other migraine syndromes associated with aura. In some families, the disorder is associated with progressive ataxia. Familial hemiplegic migraine is genetically heterogeneous with about 50 percent of cases involving mutations in CACNA1A. The heterogeneity extends to the molecular level with more than a dozen mutations associated with illness risk. The effect of the mutations was anticipated to result in a gain-of-function, given the pathological impact of the single mutant allele. However, expression of mutant genes in heterologous systems reveals a more complex picture with the most frequently identified mutation reducing Ca^{2+} currents, presumably a loss-of-function, and another mutation associated with increased Ca^{2+} flux, or gain-of-function. Episodic ataxia type 2 (EAT2) is another dominant disorder associated with mutations in CACNA1A. Patients with EAT2 experience episodes of nystagmus and ataxia lasting hours to days; for some, the disorder is progressive and associated with cerebellar atrophy. About 50 percent of patients have migraine. Although most mutations associated with EAT2 grossly disrupt protein expression, several are point mutations that change a single amino acid. Expression of the gene with one of the point mutations in a heterologous system results in complete loss of function without a change in protein expression. Finally, a type of autosomal dominant spinocerebellar ataxia (SCA6) is linked to a mutation leading to an expanded series of glutamine residues in the carboxyl terminus of the $Ca_V2.1$ protein. It is perhaps easier to conceptualize the mutant protein producing a chronic condition such as ataxia rather than an episodic disorder such as migraine or epilepsy because the P/Q channel regulates transmitter release in cerebellar circuits responsible for gait. It is conceivable that the P/Q mutations associated with migraine are important

for the function of a neural circuit activated by adverse environmental exposure and/or dependent on a transmitter such as serotonin (thought to be involved in pathogenesis of migraine). The various syndromes associated with altered CACNA1A gene highlight the heterogeneity that is expected from exploration of genes for psychiatric disorders.

Genetic studies indicate that certain pain syndromes are also channelopathies involving altered Na^+ channel function and expression. Mutations in the SCN9A gene on chromosome 2 that encodes $Na_v1.7$ are implicated in both paroxysmal extreme pain disorder (PEPD) and congenital inability to experience pain. In the case of PEPD, the causative mutations result in fast inactivation of the channel with a persisting tonic Na^+ current. In the pain insensitivity disorder, the mutations result in a loss of function. These syndromes are instructive not only from the perspective of ion channels and disease but also in guiding future work aimed at developing more specific pain therapies.

Studies examining ion channels in psychiatric disorders remain in their infancy. There is evidence that polymorphisms in SK type calcium-activated K^+ channels are associated with certain forms of psychosis. A polymorphism in the KCNH2 gene that encodes a K^+ channel is also associated with schizophrenia and predicts lower intelligence and altered memory-related changes in functional neuroimaging. Other work suggests the involvement of BK channels in a form of autism and mental retardation, and mutations in T-type (LVA) calcium channels have been linked to autism. Similarly, a polymorphism in the α5 neuronal nicotinic acetylcholine receptor is linked to nicotine dependence. This latter polymorphism is in a coding region of the protein in a position that influences ion channel kinetics. A recent large-scale study revealed genetic loci in schizophrenia that likely code several voltage-gated ion channels including α1 C- and β2-subunits of L-type (HVA) Ca^{2+} channels, $K_v2.1$ delayed rectifier K^+ channels, ClC-3 voltage-activated Cl^- channels, and HCN1 channels, among others. These observations suggest that understanding how genetic and environmental factors influence membrane excitability will play an increasing role in psychiatry. Changes in ion channel function are likely to contribute to diagnosis, pathophysiology, and response to psychotropic medications, including predispositions to serious side effects including cardiac arrhythmias.

"E/I Balance" and Psychiatric Illnesses

A principle emerging from modern genetics is that certain psychiatric illnesses, including schizophrenia, bipolar disorder, developmental cognitive disorders, and autism share genetic risk factors and that genes associated with these complex disorders involve pathways that regulate neuronal excitability and synaptic function. Anatomical studies indicate

that some of these psychiatric illnesses also involve abnormalities in inhibitory interneurons that use GABA as a neurotransmitter and that regulate regional excitability. These observations have led to hypotheses about neurocircuitry defects that underlie the disorders. One hypothesis posits that the various disorders involve shifts in the relative degrees of regional excitation and inhibition (referred to as E/I **balance**). According to this hypothesis, both synaptic and intrinsic excitability changes contribute to neurocircuitry abnormalities. In schizophrenia and autism, there appears to be augmented excitation relative to inhibition in certain brain regions, and this increased E/I ratio may account for abnormal hyperactivity in frontal and hippocampal circuits that underlie cognitive and social defects. This E/I balance hypothesis is consistent with functional neuroimaging findings and increases in high-frequency cortical oscillations recorded by electroencephalography (EEG) and magnetoencephalograhy (MEG). Changes in beta and gamma oscillations (activity above 20 Hz) are thought to underlie defects in sensory processing and learning, and inefficient information transfer within and between neural networks. Specific classes of inhibitory GABAergic interneurons play a critical role in regulating these high-frequency rhythms by influencing the firing frequency of excitatory glutamatergic (pyramidal) neurons.

There are few direct tests of the E/I balance hypothesis in humans, but studies using optogenetic methods in mouse models have provided tentative support (see below). Studies using bistable step function opsins to depolarize specific neuronal types for prolonged periods have reproduced increases in high frequency (30 to 80 Hz) cortical gamma oscillations and some altered social behaviors akin to those observed in human illnesses. These electrical and behavioral changes are reversed by increasing the activity of specific GABAergic inhibitory interneurons that express the calcium-binding protein parvalbumin (PV). PV^+ interneurons are fast spiking cells and a model for understanding how expression of specific classes of ion channels drives unique function, allowing interneurons to regulate rhythmic activity within and across brain regions. Simulation studies provide further support for the E/I hypothesis and indicate that local feedback inhibition enhances information capacity and discrimination accuracy of more global and distributed networks. These studies suggest that local E/I balance within brain regions helps to regulate regional spontaneous activity and information transfer across brain regions at large-scale levels.

While the E/I balance hypothesis is still not proven and may represent an overly simplistic view, this conceptualization is having a significant influence on clinical neuroscience. Efforts to manipulate E/I balance and the plasticity of excitatory (glutamate) synapses are leading to novel ideas about the treatment of developmental neurocognitive disorders that have heretofore proven therapeutically intractable, including fragile X

syndrome, Rett syndrome, and ASDs, among others.

Neuronal Activity and the Fetal Alcohol Syndrome

Ion channels play major roles in brain development. During development, more neurons are produced than are needed for mature functioning. Depending on the region involved, 50 percent or more of neurons normally do not survive, and die by a process of programmed (apoptotic) cell death that is regulated in part by electrical activity. This activity involves intrinsic action potential firing mediated by voltage-activated ion channels, and synaptic transmission, particularly excitatory synaptic input. When intrinsic neuronal activity is inhibited by blocking voltage-activated sodium or calcium channels, neurons undergo apoptotic neuronal degeneration. Studies in rodents indicate that agents that diminish glutamate-mediated synaptic transmission, particularly the component mediated by NMDA receptors, cause massive abnormal loss of neurons throughout the forebrain during the period of synaptogenesis. Treatments that enhance GABA-mediated inhibition also promote neuronal apoptosis during the same developmental period. Multiple clinically used and abused drugs modulate voltage-gated ion channels, NMDA receptors, and GABA receptors, and these drugs promote developmental neuronal apoptosis. For example, ketamine-like drugs that inhibit NMDA ion channels are potently neurotoxic in rodents when administered on a single day during synaptogenesis. Similar neurotoxicity is observed with benzodiazepines and barbiturates that act via $GABA_A$ receptors. General anesthetic and anticonvulsants dampen brain electrical activity and have also been linked to developmental neuroapoptosis.

In humans, synaptogenesis extends from the third trimester of pregnancy through the first several years of life, and in some regions, like prefrontal cortex, may persist even longer. Exposure of the developing human nervous system to ethanol produces a syndrome known as FAS or in its milder form, fetal alcohol effects (FAEs). FAS is characterized by microcephaly, short stature, facial abnormalities, and learning defects. Rodents exposed to intoxicating levels of ethanol for several hours on a single day during synaptogenesis develop widespread neurodegeneration, which in some regions results in loss of more than half the neurons. Ethanol is a drug with complex effects, including the ability to inhibit NMDA receptors and enhance $GABA_A$ receptors.

Major psychiatric disorders result from complex interactions of genes with environmental variables, and some studies indicate that individuals with fetal alcohol exposure exhibit significant psychopathology as they mature to adulthood, including increased prevalence of major depression and psychotic disorders. It is reasonable to be concerned that early exposure to drugs that alter neuronal activity during development may

have major impact on the development of psychiatric syndromes in adulthood, and some epidemiological data support this notion beyond FAS.

Optogenetics, Neuronal Excitability, and Psychiatric Neuroscience

Psychiatric disorders involve dysfunction within and across neural networks underlying cognition, emotion, and motivation. The activity of neurons in these circuits underlies information processing and is at the base level of pathophysiology. While much progress is being made in understanding molecular mechanisms of neuronal excitability and synaptic function, considerable gaps exist in relating neuronal activity to the function of neural circuits. Over the past decade, driven in large part by the work of psychiatrist Karl Deisseroth and others, the field of **optogenetics** has evolved and is filling this gap by revolutionizing how psychiatrists think about and study the role of neural circuits in mental processing, behavior, and psychiatric disorders. This section discusses the principles of optogenetic methods as they pertain to neuronal function. These principles directly involve the points raised earlier in this chapter.

Optogenetics combines optics and genetics, and takes advantage of the fact that certain microorganisms make proteins that use visible light to regulate the flow of ions across their membranes. These proteins link opsins (light-activated proteins) to the function of ion channels and membrane pumps. A leading example of this is channelrhodopsin (ChR), a nonselective cation channel that is activated by blue light. As is true of nonselective cation channels described elsewhere in this chapter, ChR provides a depolarizing stimulus to neurons when stimulated by the appropriate wavelength of light (~480 nm). By expressing ChR in specific brain regions or certain cell types within a brain region, it is possible to use light to precisely drive action potential firing of certain neurons but not others, and to study the influence of these neurons on circuit function and animal behavior. The selective expression of ChR in specific neurons is controlled by cell-specific genetic promoters. Thus, compared to electrical methods that stimulate all cells in a local brain region, optogenetic approaches provide a way to more precisely activate specific neuronal populations.

Other techniques exploit the idea of introducing a foreign receptor into selected neurons to investigate the influence of particular cell types on behavior. These methods are sometimes classified as **chemical genetics**. In these techniques, a ligand-gated receptor or a GPCR, activated by an exogenous small molecule, is expressed in specific neurons. When the animal is administered the ligand for these designer receptors, the receptor is activated and behavioral effects are measured. Compared with optogenetics, which requires an implanted brain light source in addition to a genetically introduced ion channel, chemical genetics is somewhat less invasive.

Advances in optogenetics and chemical genetics have greatly expanded the toolbox available to scientists. For instance, optogenetic variants are available that inhibit neurons (e.g., the chloride and proton pumps halorhodopsin and Archaerhodopsin), and that even more precisely regulate the timing of neuronal changes (e.g., step function opsins that require pulses of light to turn on and off as described above). Use of different opsins also allows expression of channels/pumps that are regulated by different wavelengths of light across the visible spectrum. Longer wavelengths of light (red-shifted variants) allow stimulation with lower light energy levels, producing less tissue heating and deeper light penetration. Optogenetics has promoted exploration of psychiatric models in mice, including studies of fear and anxiety, addiction and reward, depression and anhedonia, schizophrenia, and autism. How these findings translate to human illnesses remains to be determined and suffer the limitation that mouse models of psychiatric disorders are abstractions of clinical conditions. Importantly, optogenetic approaches can be expanded beyond mice to study brain function and behavior in higher-order species. Whether optogenetics or chemical genetics can be developed for the treatment of human psychiatric illnesses remains to be determined.

Neural Excitability and Brain Stimulation Methods in Neuropsychiatry

There is increasing interest in the use of brain stimulation (neuromodulation) methods to treat psychiatric and neurological disorders. These methods include ECT, vagus nerve stimulation (VNS), rTMS, DBS, and transcortical direct current stimulation (tDCS). The development of optimal stimulation parameters for these treatments requires knowledge about the effects of electrical stimulation on neuronal function. A major advance in ECT has been the recognition that electrical stimuli play a key role in determining therapeutic and adverse effects. Because the goal of an ECT session is to cause a generalized brain seizure while minimizing cognitive side effects, understanding the factors that determine seizure threshold and optimizing stimulus parameters becomes extremely important. To stimulate nerve cells, brief square-wave pulses of electrical current (0.1 to 2.0 milliseconds) are much more effective than more prolonged pulses or sine-wave stimuli. The rate of delivery of the current pulses is also important with lower frequencies (30 to 40 Hz) being more efficient than very high frequency trains (>100 Hz). These features reflect the fact that entrainment of neurons in a seizure is more likely using parameters that mimic neuronal firing patterns. Long current pulses (particularly sine waves) are inefficient because of the relatively slow onset and offset of the stimulus, and because much of the stimulus is delivered during the absolute and relative refractory periods when neurons are less excitable. Similarly, very high frequencies of stimulation also result in pulses being delivered during refractory periods. For reasons that are

unclear, ultrabrief electrical pulses (~0.2-millisecond duration) are effective for unilateral but not bilateral ECT. The degree to which electrical doses exceed seizure threshold is also important. For bilateral ECT, electrical doses ~1.5 times threshold result in highly effective treatments that minimize cognitive impairment. For nondominant hemisphere (unilateral) ECT, these same electrical doses are ineffective and stimuli that are five to six times threshold are required for clinical improvement. The implication of these observations is that electrical dosing and stimulation parameters have a major impact on ECT outcomes, and both bilateral and unilateral ECT can be made effective or ineffective depending on the stimulus.

More recent additions to the brain stimulation methods, particularly DBS, also require an understanding of the effects of electrical stimulation on neuronal activity, and raise issues about the potential role of homeostatic plasticity as a therapeutic mechanism. The parameters used for DBS typically consist of brief (0.06 milliseconds) pulses administered continuously at high frequency (>100 Hz). Modeling studies suggest that this stimulation has complex effects on network function with suppression of intrinsic firing at the neuronal cell bodies in the stimulated region but enhanced axonal responses and increased efferent output to downstream targets. Whether similar considerations are important for other brain stimulation methods such as rTMS and VNS is uncertain. A potential vista in the treatment of neuropsychiatric disorders may involve the ability to regulate neural activity in focal regions of the brain using optogenetics or small molecules that are inactive until exposed to light of an appropriate wavelength. These methods offer the hope of targeting specific neurocircuits in specific disorders.

SUMMARY

The diversity of voltage-gated and ligand-gated ion channels is becoming better understood at structural, biophysical, and genetic levels. The properties of ion channels are being harnessed to yield unprecedented understanding of the roles of particular neural circuits in normal and abnormal behavior in animals. Electrical properties of neurons are also fundamental to modern brain stimulation methodologies. Although the electrical events underlying neuronal excitability are relatively stereotyped, the numerous ion channels contributing to neuronal excitability offer a great deal of flexibility in the control of cellular and network activity. Furthermore, the diversity of ion channels involved in electrical signaling provides complex and powerful mechanisms by which excitability can be modulated by neurotransmitters and drugs. Determining how alterations in ion channel function contribute to development, behavior, cognitive processing, and clinical syndromes continues to be a major goal in this field.

REFERENCES

Cannon SC. Pathomechanisms in channelopathies of skeletal muscle and brain. *Annu Rev Neurosci*. 2006;29:387–415.

*Catterall WA. Sodium channels, inherited epilepsy, and antiepileptic drugs. *Annu Rev Pharmacol Toxicol*. 2014;4:317–338.

Cattarall WA, Swanson TM. Structural basis for pharmacology of voltage-gated sodium and calcium channels. *Mol Pharmacol*. 2015;88:141–150.

Crisp SJ, Kullmann DM, Vincent A. Autoimmune synaptopathies. *Nature Rev Neurosci*. 2016;17:103–116.

Deco B, Ponce-Alvarez A, Hagmann P, Romani GL, Mantini D, Corbetta M. How local excitation-inhibition ratio impacts the whole brain dynamics. *J Neuroscience*. 2014;34:7886–7898.

Delorme R, Ey E, Toro R, Leboyer M, Gillberg C, Bourgeron T. Progress toward treatments for synaptic defects in autism. *Nat Med*. 2013;19:685–694.

Doyle DA. Structural changes during ion channel gating. *Trends Neurosci*. 2004;27:298–302.

Herguedas B, Garcia-Nafria J, Cais O, Fernandez-Leiro R, Krieger J, Ho H, Greger IH. Structure and organization of heteromeric AMPA-type glutamate receptors. *Science*. 2016;352:549.

*Hille B. *Ion Channels of Excitable Membranes*. Sunderland, MA: Sinauer Associates Inc.; 2001.

*Hu H, Gan J, Jonas P. Fast-spiking, parvalbumin+ GABAergic interneurons: from cellular design to microcircuit function. *Science*. 2014;345:1255263.

*Isaacson JS, Scanziani M. How inhibition shapes cortical activity. *Neuron*. 2011;72:231–243.

Jan LY, Jan YN. Voltage-gated potassium channels and the diversity of electrical signaling. *J Physiol*. 2012;590:2591–2599.

Karakas E, Furukawa H. Crystal structure of a heterotrimeric NMDA receptor ion channel. *Science*. 2014;344:992–997.

Khosravani H, Zamponi GW. Voltage-gated calcium channels and idiopathic generalized epilepsies. *Physiol Rev*. 2006;86:941–966.

Klassen T, Davis C, Goldman A, et al. Exome sequencing of ion channel genes reveals complex profiles confounding personal risk assessment in epilepsy. *Cell*. 2011;145:1036–1048.

Kopfer DA, Song C, Gruene T, Sheldrick GM, Zachariae U, de Groot BL. Ion permeation in K+ channels occurs by direct Coulomb knock-on. *Science*. 2014;346:352–355.

Lee CH, Lu W, Michel JC, et al. NMDA receptor structures reveal subunit arrangement and pore architecture. *Nature*. 2014;511:191–197.

McCarroll SA, Hyman SE. Progress in the genetics of polygenic brain disorders: significant new challenges for neurobiology. *Neuron*. 2013;80:578–587.

Miller PS, Aricescu AR. Crystal structure of a human $GABA_A$ receptor. *Nature*. 2014;512;270–275.

Ranada SS, Syeda R, Patapoutian A. Mechanically activated ion channels. *Neuron*. 2015;87:1162–1179.

Shah MM. Cortical HCN channels: function, trafficking and plasticity. *J Physiol*. 2014;592;2711–2719.

*Schizophrenia Working Group of the Psychiatric Genetics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511:421–427.

*Tye KM, Deisseroth K. Optogenetic investigation of neural circuits underlying brain disease in animal models. *Nature Rev Neurosci*. 2012;13:251–266.

Waxman SG. Channel, neuronal and clinical function in sodium channelopathies: From genotype to phenotype. *Nature Neurosci*. 2007;10:405–409.

Womelsdorf T, Valiante TK, Sahin NT, Miller KJ, Tiesinga P. Dynamic circuit motifs underlying rhythmic gain control, gating and integration. *Nature Neurosci*. 2014;17:1031–1039.

Yizhar O, Fenno LE, Prigge M, et al. Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature*. 2011;477:171–178.

Zhu S, Stein RA, Yoshioka C, Lee C-H, Goehring A, Mchaourab HS, Gouaux E. Mechanism of NMDA receptor inhibition and activation. *Cell*. 2016;165:704–714.

Zorumski CF, Rubin EH. *Psychiatry and Clinical Neuroscience: A Primer*. New York: Oxford University Press; 2011.

▲ 1.11 Genome, Transcriptome, and Proteome: The Molecular Genetics and Biochemistry Underlying the Neurobiology of Mental Disorders

MARJAN BOZINOSKI, M.S., AND CHRISTOPHER E. MASON, PH.D.

This is an era in which the global study of the DNA, ribonucleic acid (RNA), and protein building blocks of cells has become feasible and increasingly routine. Genome-wide association studies (GWAS) of a million or more common genetic markers from large populations are now relatively inexpensive and conducted routinely. The same can be said for studies that sequence the exome, that is, the DNA of expressed genes, as well as the entire genome. The growing understanding of genome variation provided by the International HapMap Consortium and continued major advances in genotyping technology together have made it possible to conduct high-throughput, cost-effective GWAS in large numbers of individuals with detailed information on phenotypic traits and environmental exposures. The resulting data will be used to identify genetic variants potentially related to mental disorders, to assess the prevalence of these variants in large and diverse samples, and to examine possible modifiers of gene-disease relationships. Soon, the 1,000 genome project followed, organizing a comprehensive database of SNPs, copy number variations (CNVs), insertions and deletions (indels), as well as large structural rearrangements. ENCODE project (described later in this chapter) expanded these studies further into functional genomics. Functional genomics is already becoming routine in brain research and is being applied to the study of postmortem human brains from individuals with mental illness and to animal models of relevance to clinical neuroscience.

These exciting molecular genetic approaches permit the study of biological information from a global perspective. This information is

contained in the **human genome**, the three billion base sequence comprising the full complement of DNA in human cells. The genome of each individual is contained within every cell in the body that carries a nucleus and represents the raw material necessary to allow for normal development. This genetic material also plays a causal or contributory role in much of human disease, including mental illness. **Genomics** is the study of the full complement of genetic material of an organism. The ability to consider humans in this light is a fairly recent development, one that has been energized by the sequencing of the human genome and the subsequent large-scale efforts to identify and characterize human genetic variation and all functional and regulatory elements. The **transcriptome** is a term that can be applied to the sum total of RNAs transcribed from genes in an organism, a process often referred to as gene "expression." Studies of gene expression vary depending on the experimental design, and may include only mature, polyadenylated mRNA or total RNA (mRNA plus noncoding RNA) in the cell. Studies of gene expression usually focus on (**mRNA**), the intermediate between genes and proteins. Global studies of gene expression compose the field of **functional genomics** and may use such tools as New Generation Sequencing (NGS), a methodology that relies on the power of massive parallel sequencing of short reads derived from genomic DNA or RNA. A simplified diagram of NGS workflow is shown in Figure 1.11–1.

The **proteome** denotes the full panoply of proteins expressed in an organism and studied globally by proteomics. As the general understanding of these biological domains in humans and many other organisms has grown, so too has the appreciation of its complexity. The deeper we delve into its structure and function, the more we are pressed to reconsider basic concepts. For instance, the amount of genetic variation in apparently normal healthy people is underestimated. While it has long been appreciated that the sequence of the DNA varies slightly between individuals, it has recently learned that as much as 20 percent of the human genome can possess large-scale structural variation as well. Hundreds of thousands to millions of nucleotides of the genetic code may be missing, duplicated, or have multiple copies within a single individual without obvious adverse effects. These variations, which are referred to as **copy number polymorphisms** (CNPs) or **copy number variants** (CNVs), are just beginning to be catalogued and studied in depth. In addition, there are extremely large numbers of common differences between individuals on a single nucleotide level, known as SNPs. Understanding the importance of these many forms of genetic variation promises to revolutionize the current thinking about the role of genetic structural variation in normal development, shed light on the heritability and biology of traits associated with psychiatric disorders, and perhaps ultimately to reshape psychiatric diagnostic schema.

The notion of "gene" is evolving as well. It was long thought that a single gene led to the production of single protein. However, by 1977, it was discovered learned that one gene may code for multiple versions of a protein that differed functionally. Recent data challenge the conventional definition of the boundaries of an individual gene, findings which will likely lead to a major reappraisal of genomic function and regulation. For example, most genes have multiple 3′ and 5′ termini and various splicing sites, and harbor various regulatory elements within their gene bodies. Also, elements called enhancers can control gene expression from long-range, and TFBSs can regulate the ability of genes to be activated, attenuated, or silenced.

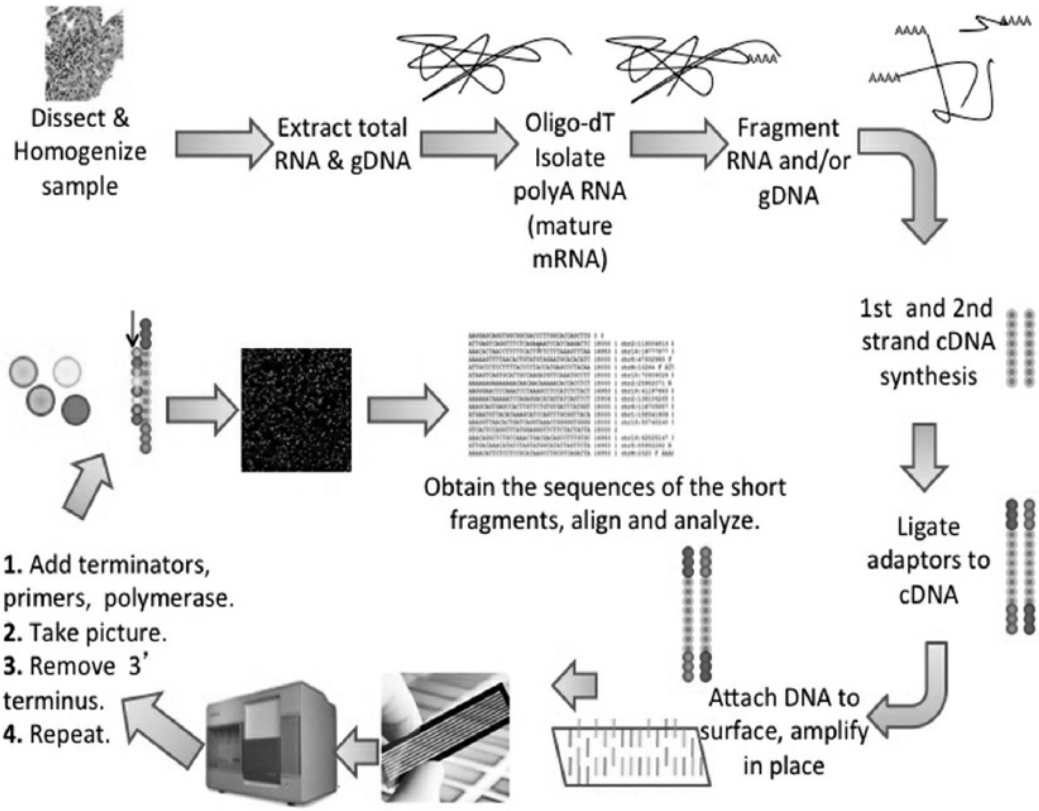

FIGURE 1.11–1. Overview of the NGS workflow. Tissue is being dissected and homogenized, and genomic DNA and/or RNA are extracted. In case of studying mature coding mRNAs, oligo-dT purification is performed. Next, DNA and/or RNA are fragmented into short pieces (~200 nt) and cDNA library is derived. Next, adaptors are ligated, which help anchor the cDNA to a hard surface, and sequencing by synthesis is performed. The final results are the sequences of the short fragments of DNA or RNA.

In addition to forcing a reconsideration of fundamental genetic concepts, the sequencing of the entire human genome also set the stage for a far deeper understanding of how genetic variation alters the regulation of gene expression and how variation either in functional or regulatory domains of genes affects brain structure and function, and how this variation contributes to the risk for mental disorders.

This chapter will review fundamental concepts, including the nature of

DNA, chromosomes, RNA, proteins, and the concept of coding and noncoding genes. This chapter will then review what is currently known about the sequence characteristics of the human genome on a global scale and consider how this sequence varies between individuals and among populations. Subsequently, This chapter will consider how the genome is regulated and the manner in which the study of proteomics and systems biology promises to leverage these recent advances in functional genomics to enhance the general understanding of normal and pathological brain development and functioning.

THE ORGANIZATION AND STRUCTURE OF THE HUMAN GENOME

A **genome** is defined as the total complement of DNA replicated in a living organism. The sequencing of genomes of free-living organisms began in 1995 with bacteria and progressed to larger and more complicated organisms, such as yeast (1996), worms (1998), and fruit flies (2000). A critical milestone was reached in 2001 with the completion of the first draft of the human genome. Currently, there are thousands of genomes sequenced. In 2011, the Genome in a Bottle Consortium was established, with the goal of providing an entire technical infrastructure that will enable data from genome sequencing to be used in translational medicine.

As these data became available, four curious patterns emerged: (1) The number of genes is lower than expected. Estimates prior to the completion of the draft human sequence ran as high as 160,000. As of the last issue of GENCODE annotation, the official annotation of the ENCODE project, there are 60,155 genes, with 19,881 of them being protein-coding genes. It is now evident that there are in total about 20,000 protein-coding genes in *Homo sapiens;* (2) Gene number is not a predictor of organismal complexity. For example, amoebas are simple, single-celled organisms, but are predicted to have many more genes than humans (Table 1.11–1); mice and humans have been found to have approximately the same number of protein coding regions; (3) Genome size is not a predictor of organismal complexity. An amoeba (*Amoeba dubia)* has the largest known genome (670 gigabases); humans and chimpanzee genomes are essentially of the same size, but they hold obvious and important differences with respect to the development of the cerebral cortex; and (4) A small percentage of eukaryotic genomes codes for proteins. Only 2 percent of the human genome encodes proteins. The vast majority is comprised of intronic and repetitive sequences, some of which are clearly functional, and much of which remains a mystery. Apart from protein-coding genes, there is a substantial number of actively transcribed genes that don't code any proteins (noncoding RNAs, to be discussed below), sometimes called the "dark matter" of the genome. Some of this "dark matter" is known to be functional, although the functional significance of much of noncoding RNAs is yet to be resolved.

Cumulatively, these observations pose a central question for psychiatry: if the genome of humans is not very distinctive from an evolutionary perspective, how has it produced the unique qualities of the human CNS? Recent developments have begun to shed light on this critical question.

Table 1.11–1.
Number of Genes versus Genome Size across Different Organisms

Organism	Gene Count	Genome Size (Base Pairs)	Number of Chromosomes
Amoeba	50,000	670,000,000,000	13
Plant (Fern)	37,500	100,000,000,000	90
Human	25,000	3,000,000,000	46
Chimpanzee	25,000	3,000,000,000	48
Mouse	25,000	2,500,000,000	40
Honey bee	15,000	300,000,000	32
Fruit fly	14,000	130,000,000	10
Worm	19,000	97,000,000	12
Fungus	6,000	13,000,000	32
Bacterium	3,000	5,000,000	1
Mycoplasma genitalium	500	580,000	1
DNA virus	450	50,000	0
RNA virus	20	10,000	0
Viroid	1	500	0

DNA and Chromosomes

DNA is made of four **nucleic acids**, also known as **nucleotides** or **bases**, Adenosine (A), Cytosine (C), Guanine (G), and Thymine (T). In total, human genomic DNA is comprised of approximately 3 billion nucleotides, and this full complement is found in every cell in the body that contains a nucleus. Within the nucleus, the genome is organized into 46 strands of DNA that complex with multiple proteins to form **chromosomes** (23 inherited from mother and 23 inherited from father). In the nuclei of cells, the strands of DNA are combined with **histone proteins** that enable DNA to be wound tightly into histone-DNA complexes called **nucleosomes**, and then organized into a super structure called **chromatin**.

From the first observations of chromosomes under the light microscope, chromatin has been divided into two types corresponding to the familiar light and dark banding patterns: **euchromatin** (lighter, less dense material) and **heterochromatin** (darker, more dense material). In one sense, chromatin acts simply as a spool around which DNA is wound, ensuring that the entire genome fits within the nucleus. This property enabled cells to remain small, a property that influenced the evolution of multicellular organisms. However, it also plays a key role in coordinating the function of the genome. The histone protein cores around which the DNA is organized may be altered by chemical reactions including

acetylation, phosphorylation, and **methylation**. The addition or subtraction of these and other chemical modifiers are able to dictate the conformation of regions of DNA within the nucleus, a process which helps coordinate the regulation of gene expression. A field of study, epigenetics, has emerged to study the way that gene expression is altered by modifying chromatin or DNA methylation but not altering the DNA sequence (see section on Epigenetics of this chapter). Finally, the organizational structure of chromatin creates scaffolding that helps guide the movement of DNA necessary for chromosomal replication and during cellular division (**mitosis**).

The darker staining of heterochromatin is due to the very tight packing of nucleosomes, making the DNA largely inaccessible to cellular machinery, whereas the light color of the euchromatin reflects a more open, unfolded, and usually active state. Euchromatic regions were found to be distinct in other ways as well: they hold many more genes and contain less repetitive sequences. However, heterochromatic regions of the genome can become active, and some heterochromatic regions are modified back and forth, becoming active only when needed. These processes will be discussed in more detail subsequently when epigenetic mechanisms are considered at the conclusion of this chapter.

Genes, RNAs, and Proteins

Gregory Mendel first discovered genes in 1866, when he identified hereditary units he called "factors" and elaborated the principles of their inheritance through his experiments with pea plants. However, he had no physical or mechanistic understanding of what constituted a gene. The first medical application of genes came in 1902, when Archibald Garrod found that alkaptonuria, a rare disorder characterized in part by the urine turning black when exposed to air, followed Mendel laws of inheritance. He identified that the transmission of this condition followed an autosomal recessive pattern. However, while it was clear that genes were passed from one generation to the next and carried disease liability, it was still unclear precisely how or in what substance.

In 1910, Thomas Morgan demonstrated that genes were discrete units on chromosomes though his work with fruit flies, and he first proposed the idea of genetic linkage due to chromosomes exchanging material in a process called **crossing over** (Fig. 1.11–2).

In 1941, experiments in bread molds (*Neurospora crassa*) by Edward Tatum and George Beadle showed further that certain enzymes were functionless if the genes were mutated. This led to the widely accepted notion that specific genes make specific proteins, commonly referred to as the "**one gene, one enzyme**" hypothesis.

Despite these remarkable discoveries, the physical substrate for the transmission of genetic information was still a matter of debate. Early in

the 20th century, proteins were favored as they were known to be comprised of 20 different amino acids. As DNA was the product of only of four different nucleotides, it was thought that the former was most likely to allow for the diversity required to make the complex instructions necessary for sustaining life. However, in 1944, genes were shown to be composed of DNA by Oswald Avery, through his demonstration that DNA could create a heritable transformation in bacteria, but proteins could not.

Finally, in 1953, Watson and Crick published the chemical structure of DNA, showing that it was a **double helix** formed by two sugar-phosphate backbones supporting the four nucleotides. These, in turn, were recognized to form specific pairs between them, such that an "A" on one strand paired with "T" on the other and "C" on one paired with "G" on the other.

Coding Genes

The last several decades of progress has now led to an appreciation that there are, in a broad sense, two classes of genes: those that lead directly to the production of proteins (**coding genes**) and those that do not (**noncoding genes**). Coding genes are stretches of DNA that are **transcribed** by an enzyme (RNA polymerase) into a temporary mRNA, which is then **translated** into a protein (made of peptides), following what is known as the central dogma of molecular biology: DNA→RNA→Protein (Fig. 1.11–3).

In the 1960s, it was discovered that coding DNA is "read" in three-letter segments called **codons** (Fig. 1.11–3), each of which becomes an amino acid in a peptide or protein. However, unlike simpler prokaryotic organisms, human coding genes are transcribed initially as sequences that require editing. Coding genes have segments that are read as codons (called **exons**), and also noncoding sequences that are transcribed but not translated into a protein. Noncoding sequences can be present in three areas of a gene: in front (**upstream** with respect to the sequence of transcription) of the gene, inside the gene, and behind (**downstream**) the gene. Those sequences upstream of the gene are called the **5′ UTR**, while noncoding sequences flanked by exons are called **introns.** The noncoding region downstream of the coding interval is the **3′ UTR**. Mature mRNA contains a polyadenylation signal (i.e., AAUAA) that directs transcriptional machinery to insert a long string of adenosines, known as a poly-A tail, at the 3′ terminus of the mRNA. A single gene may contain several polyadenylation signals in series in its 3′ UTR, but the mechanism of determining which one is going to be used has yet to be established. In turn, multiple polyadenylation signals allow for transcripts with variable 3′ UTR length to be produced from a single gene. Transcripts with longer 3′ UTRs are relatively prevalent in brain, giving rise to higher numbers of miRNA binding sites that enable fine tuning of RNA regulation. Additionally, longer 3′ UTRs are known to harbor protein-binding motifs

that shuttle proteins to a specific subcellular location. The BDNF gene is known to possess a form with long 3′ UTR with a binding motif for BDNF that enables it to be transported to dendrites. Thus, the nascent mRNA transcript produced from a coding gene must be first processed to lead to the coherent set of instructions necessary to produce a protein product at the appropriate subcellular location. The removal of introns, through a process called **gene splicing**, is one critical step in determining a gene's function, and it will be addressed in more detail when considering the function of each part of the genome.

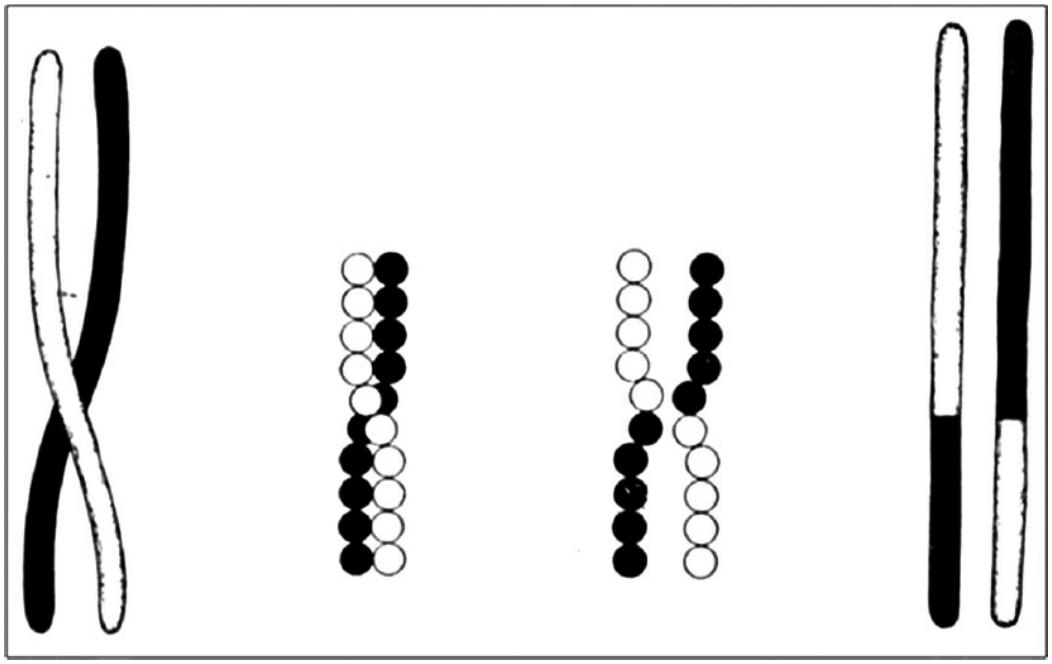

FIGURE 1.11–2. Crossing over. A schematic of genes positioned along chromosomes (*circles*) and exchanging information (crossing over) with a different series of genes (*black circles suddenly with white circles*) during meiosis. The farther apart genes are on a chromosome, the more often they cross over—a phenomenon known as recombination.

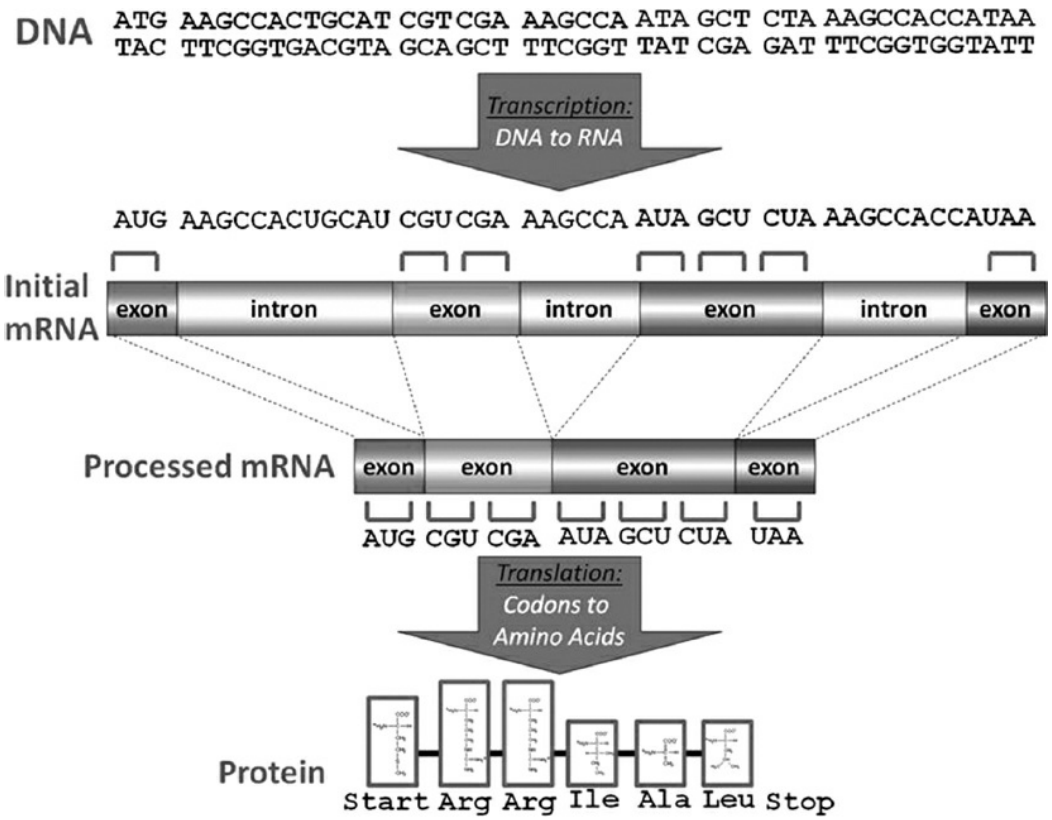

FIGURE 1.11–3. Transcription and translation of a gene. During transcription, the thymine (T) in DNA is represented with a different nucleic acid (U, for Uracil), so that the codon can be understood by the cell's machinery. Introns are removed from the final messenger RNA (mRNA), leaving a product that is ready to become a protein through translation of the codons.

Second Letter of Codon

First Letter of Codon	A	G	C	U	Third Letter of Codon
A	AAA, AAG – Lys AAC, AAU – Asn	AGA, AGG – Ser AGC, AGU – Arg	ACA, ACG, ACC, ACU – Thr	AUA – Ile AUG Start AUC, AUU – Ile	A G C U
G	GAA, GAG – Asp GAC, GAU – Glu	GGA, GGG, GGC, GGU – Gly	GCA, GCG, GCC, GCU – Ala	GUA, GUG, GUC, GUU – Val	A G C U
C	CAA, CAG – Gln CAC, CAU – His	CGA, CGG, CGC, CGU – Arg	CCA, CCG, CCC, CCU – Pro	CUA, CUG, CUC, CUU – Leu	A G C U
U	UAA Stop UAG Stop UAC, UAU – Tyr	UGA Stop UGG – Trp UGC, UGU – Cys	UCA, UCG, UCC, UCU – Ser	UUA, UUG – Leu UUC, UUU – Phe	A G C U

FIGURE 1.11–4. The genetic code. Genes are read in three-letter nucleotide segments (codons) and are then read by the cellular machinery to correspond to particular amino acids.

Since gene transcripts are read in three-base codons, and DNA is made up of four nucleotides, there are 64 possible codons (4 × 4 × 4). In most organisms, mRNA that are translated into proteins possess a three-base sequence, called a **start codon** (AUG), delineating where the reading or translation of the codons should begin. Three **stop codons** (UGA, UAA, UAG) signal the last amino acid in the growing polypeptide chain. Start and stop codons are the punctuation needed for reading a gene, while the other 60 codons determine which amino acids are included in the peptide or protein. Since the cell uses only 20 different amino acids to make proteins, there exists some redundancy in the genetic code, such that different codons can encode for the same amino acid (Fig. 1.11–4). This redundancy in codon usage may have a specific function. There is a so-called codon usage bias, where translation favors some codons over others. Differences in codon translation efficiency are related to the concentration of tRNAs that correspond to those codons, thereby implying a that there is a complex interplay in between them. Such complex evolutionary changes may fine tune the translation process, especially in tissues with extremely broad repertoire of expressed genes, such as neurons.

The processed RNAs are shuttled through the cell to organelles called **ribosomes** to be turned into a protein chain which then folds into increasingly complex levels of organization (Table 1.11–2).

The total of all the proteins made by the genome is called the proteome. Many proteins serve as subunits for larger protein complexes (including the large enzyme that performs transcription, RNA polymerase II), and the study of many proteins at once is known as proteomics, a topic that will also be addressed in more detail later in this chapter.

Noncoding Genes (or Noncoding RNAs)

General statistics of the GENCODE annotation for the past 4 years shows a very interesting trend: the total number of annotated genes is rising; however, the number of protein-coding genes decreases, at the expense of rising number of noncoding RNAs, especially pseudogenes and long intergenic noncoding RNAs (LINCs), as shown in Table 1.11–3.

Table 1.11–2.
Increasing Levels of Protein Folding Complexity

Name	Definition
Primary	A chain of amino acid (peptides), making a polypeptide
Secondary	The amino acids are bound within a chain by hydrogen bonds
Tertiary	The polypeptide forms entire sheets and helices for a 3D structure

Table 1.11–3.
Distribution of Annotated Genes by Their Coding Capacity and Biotype for GENCODE 2010–2014

General Stats	Gencode Version 21 (June 2014)	Gencode Version 10 (July 2011)	Gencode Version 7 (December 2010)
Total no. of genes	60,155	52,376	51,082
Protein-coding genes	19,881	20,007	20,687
Long noncoding RNA genes	15,877	10,840	9,640
Small noncoding RNA genes	9,534	8,801	8,801
Pseudogenes	14,467	12,358	11,580

While coding genes conform most closely to early notions of how hereditary information is stored and processed, over time it has become clear that these represent only about 2 percent of the genome. Moreover, an alternative set of instructions contained within the DNA governs transcription that is not followed by the production of a protein. Over 30,000 noncoding genes have so far been catalogued in the human genome. These are comprised of RNA (and consequently are also called **noncoding RNAs**, or **ncRNAs**). Table 1.11–4 lists several types of RNAs. Some are involved in regulating normal cellular processes of gene transcription and protein translation, whereas others have only recently been discovered and their functions are not as well understood. Moreover, some ncRNAs can function all by themselves, catalyzing reactions and acting autonomously in the cell.

Some of the best characterized ncRNAs are called siRNAs. These are transcripts coded for in the DNA that are complementary (or antisense) to another transcribed sequence from the genome. Once a siRNA binds to the complementary mRNA sequence, the cell degrades the transcript, thus functionally silencing the gene that produced the mRNA without altering the DNA or mRNA sequence. Similarly, a different subset of noncoding genes, called micro-RNAs (miRNAs), are short (18–22 bp) sequences that bind most often to the 3′ UTR of a transcript and most commonly lead to a decrease in the level of the resulting protein. Both siRNAs and miRNAs operate posttranscriptionally to regulate gene expression. Thus, there are multiple extremely important ways along the process from DNA to protein in which gene expression may be altered is of major importance to scientists seeking to understand the function of the genome. This topic will be addressed in the second section of this chapter. The recognition that ncRNAs are widespread and play a key role in the regulation of gene

expression strongly suggests that they may also confer risk for mental disorders and other complex diseases. Moreover, it was quickly realized that siRNAs could be useful to assess the consequences of loss of expression of a particular gene at the level of the cell or the model organism; as a result, siRNAs have quickly become an indispensable research tool.

Two other important classes of ncRNAs are pseudogenes and LINCs. Pseudogenes are ubiquitous in the brain transcriptome, and encompass several types of transcripts, transcriptionally active genes that lost their coding potential—unitary pseudogenes (GULO-vitamin C–producing enzyme), duplicated genes that lost their coding potential, and reverse transcribed genes (processed pseudogenes). The reason their number keeps growing in each subsequent annotation is due to the fact that they share a substantial portion of their sequence with the functional genes, which makes them extremely hard to identify. Bioinformaticians have developed highly sophisticated algorithms to be able to distinguish pseudogenes from the functional genes. The "sponge" hypothesis suggests that pseudogenes may compete with functional genes for a limited pool of miRNAs. If so, then pseudogenes may influence overall gene regulation without possessing intrinsic function.

LINCs also appear to influence neural function. Most LINCs alter transcription (H19, XIST, AIR, HOTAIR) in various ways (i.e., direct effect on transcription, changing the epigenetic signature). Others exert their function on the translational level (GAS5).

While a significant amount of ncDNA is transcribed to functional ncRNAs (Table 1.11–4), it has not been clear how much of the genome was actively processed nor how much of this transcription might contribute directly to biological function. To answer this question, the National Human Genome Research Institute (NHGRI) launched a project in 2003 to identify all the functional elements in the genome, called ENCODE (the ENCyclopedia Of DNA Elements). By the second phase of the ENCODE project, it appeared that 75 percent of the human genome was transcriptionally active. However, 70 percent of the genome showed extremely low expression levels, raising questions regarding functional significance. Also, 11 percent of the genome was associated with motifs in transcription factor-bound regions or high-resolution DNase footprints in one or more cell types. Promoter-associated histone modifications or enhancers marked ~20 percent of the genome, while as much as one-third of the genome was marked by modifications associated with transcriptional elongation. Taken together, it appears that the vast majority of the genome is either transcribed or harbors regulatory elements (Fig. 1.11–5).

Table 1.11–4.

Types of RNA

Name	Abbreviation	Year	Function
Messenger	(mRNA)	1956	Carries the transcribed message from the DNA
Ribosomal	(rRNA)	1958	Components of the protein-making machinery, the ribosome
Transfer	(tRNA)	1962	Function in translation to bring amino acids to ribosome
Small nuclear	(snRNA)	1977	Help the splicing of immature mRNAs to reach their final form
Small nucleolar	(snoRNA)	1986	Direct chemical modifications of rRNAs and other RNAs
Micro-	(miRNA)	1993	Single-stranded RNA that regulates gene expression
Small interfering	(siRNA)	1999	Small, double-stranded RNAs that interfere with transcription
Piwi-acting	(piRNA)	2006	Germ-line acting RNA that stops parasitic genetic elements

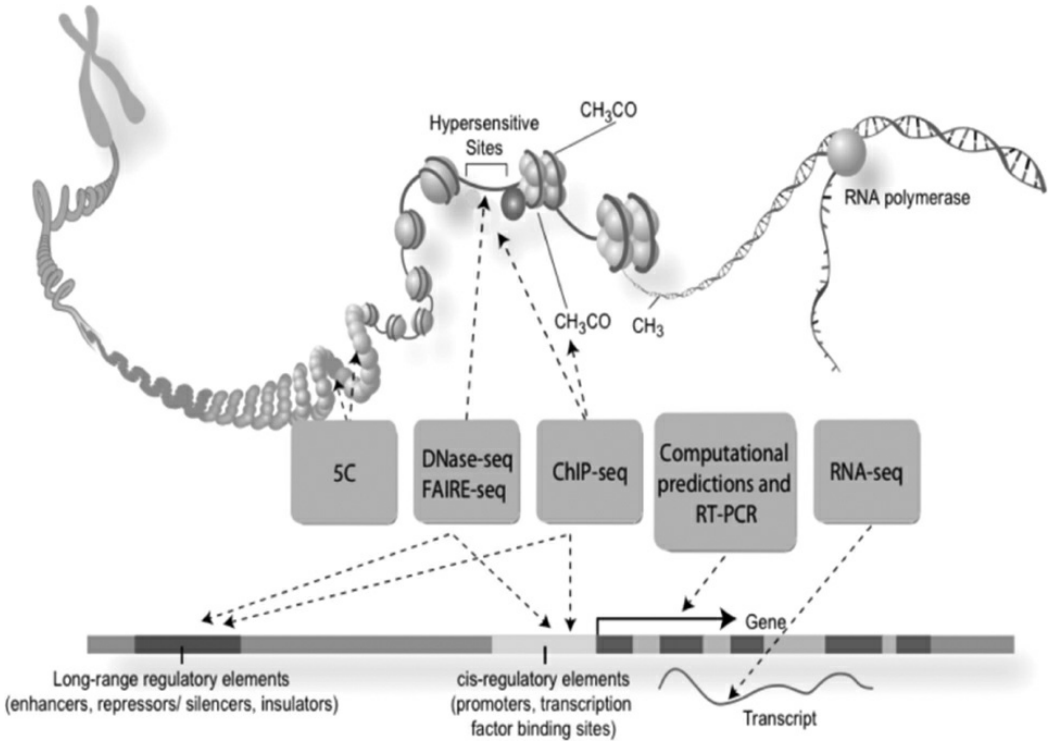

FIGURE 1.11–5. Summary of the ENCODE project: The discovery and annotation of TARs was performed by sequencing RNA from a diverse range of cell lines, followed by comparative genomics, integrative bioinformatic methods, and human curation. Investigation of the regulatory elements involved DNA hypersensitivity assays, assays of DNA methylation, and chromatin immunoprecipitation (ChIP) of proteins that interact with DNA, including modified histones and transcription factors, followed by sequencing (ChIP-Seq).

These observations strained the ability to classify genetic material. While there is a clear distinction between coding and noncoding genes, a large amount of transcriptionally active DNA produces RNAs that do not fit into current functional categories. For clarity, the term ncRNA will be used to describe noncoding genes such as siRNAs and miRNAs whose general structure is well characterized. The additional transcribed material is referred to as **transcripts of unknown function (TUFs)** or **transcriptionally active regions (TARs)**.

TUFs and TARs may represent: (1) new protein-coding genes that were

missed by previous experiments and gene-finding algorithms; (2) new noncoding genes; (3) novel antisense transcriptional units (siRNAs and miRNAs) that may regulate other genes; (4) alternative isoforms of known genes that include intronic regions in the final mRNA product; (5) mis-annotated genes that should be longer or shorter, divided into two genes, or merged into one larger gene; (6) biological artifacts representing aberrant transcription or accidental transcriptional read-through; or (7) experimental artifacts due to small amounts of genomic DNA contamination.

Future experiments will be needed to clarify the role of these TARs, but some conclusions are already possible. For instance, the majority (93 percent) of the euchromatic genome is functional in some capacity (either transcribed or regulating another sequence). Also, gene regulation is symmetrical along the genome, with no bias for upstream or downstream placement of regulatory structures. This contrasts with previous notions that regions 5′ or upstream to the starting codon of a gene were most likely to modulate regulation of that transcript. Finally, surprisingly, many of these regulatory or active regions of the genome are not highly conserved across species. Taken together, these data show that any search for genes and genetic mechanisms contributing to brain development and mental disorders must extend far beyond the horizon of coding sequences.

Repetitive DNA

With the sequencing of the human genome and the recognition that only a small proportion is present in the form of coding DNA or noncoding RNAs, interest in the remainder of the genome has increased, and much of this material is present in the form of repetitive elements.

The structural organization of the human genome is shown in Figure 1.11–6. Nearly one-third of the genome consists of repeated elements. Some repeated sections of the human genome serve structural purposes at the end or middle of chromosomes. However, many types of repeats are of unknown function including simple sequence repeats (SSRs) and segmental duplications (SegDups). SSRs are small sequences (2 to 6 bp) that are tandemly repeated in the genome. A short trinucleotide repeat (CAG) in the coding part of the HGG gene responsible for coding of the huntingtin protein is known to be the cause of Huntington disease, a progressive neurodegenerative disorder that is ultimately fatal. While healthy individuals have 10 to 35 repeats in the HGG gene, affected individuals can have as much as 120. The onset and severity of the disease correlates very strongly with the number of repeats. This is one of the few neuropsychiatric diseases with a clearly defined genetic etiology that can be diagnosed prior to the onset of symptoms with a simple test. Children of a parent with Huntington disease have a 50 percent risk of developing this disorder. Thus, the capacity to diagnose an illness in presymptomatic

individuals raises important ethical issues related to the role of genetics in neuropsychiatric clinical practice. SegDups are sequences of 1,000 bp or greater that appear at least twice throughout the genome. Interestingly, some of these SegDups contain extra copies of entire genes.

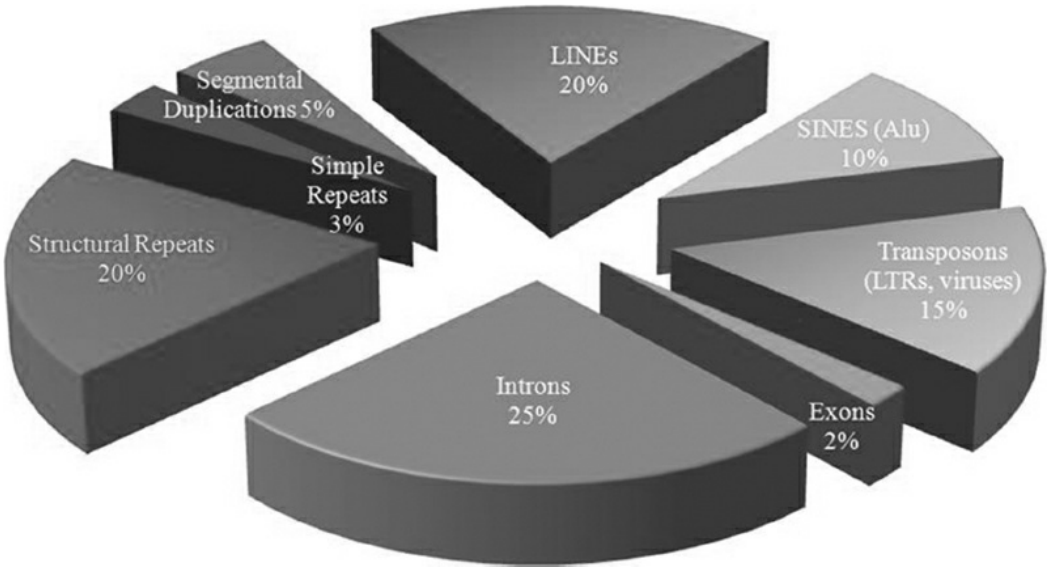

FIGURE 1.11–6. Structural organization of the human genome. Most of the human genome is repetitive DNA sequences and transposable elements (LINES, SINEs, LTRs, and viruses). Very little of the genome is coding sequence (exons) but there is great room for gene flexibility and change with many of the gene's long intron sequences (introns).

The most ubiquitous repeated genetic elements in the genome are transposable elements (TEs or **transposons**), which can jump from one place to another in the genome. This process may lead to new forms of genes that may be useful to humans, but it also poses a danger of disrupting essential genes. TEs are divided into several subgroups: **Long interspersed elements** (LINEs), **short interspersed elements** (SINEs), and the small 300 bp *Alu* element (considered a SINE). The *Alu* element is present in 75 percent of introns and accounts for 10 percent of the entire genome, whereas the 6,000 bp (6 kb) LINEs account for 20 percent of the human genome, and can be found in thousands of genes (in 5′ UTRs, exons, 3′ UTRs, and introns). Almost all genes in the human genome have at least one TE.

Alu elements are both transposons and retrotransposons, which alter DNA through reverse transcription in the manner of retroviruses, such as the 7SL RNA-derived primate-specific repeat family. Their distribution throughout the evolution of primate lineage has resulted in as much as 1.1 million copies in the human genome. Insertions of the Alu elements occurred in a few major waves, deriving old AluJ subfamily (including AluJo and AluJb; about 81 million years ago), the middle-aged AluS subfamily (including AluS, AluSx, AluSg, AluSq, AluSp, AluSc, and AluSb; about 48 to 19 million years ago), and the young and still active AluY wave

(including AluY, AluYa5, AluYa8, and AluYb8; starting about 6 million years ago and continuing to the present) (M7). Albeit ubiquitous in the genome, Alu repeats (as any other repeat or low-complexity sequence) are extremely hard to map accurately using the current short-read sequencing. This fact, along with the limited RNA sequence data derived from primates, explains the lack of proper mapping of these elements. The body of the *Alu* element is approximately 300 bp bases in length, formed from two diverged dimers, separated by a short A-rich region. Such structure makes enables them to fold perfectly into double-stranded structures, thereby being substrates for a process called RNA editing (discussed below).

Novel findings show that Alu elements are actively transposing in neurons in culture as well as in neurons in vivo. These changes in the neuronal genomes might generate neurons with unique genetic, transcriptomic, and proteomic makeup, increasing the complexity of the brain.

Since the human's divergence from the common ancestor with the chimpanzee, approximately 98,000 viruses have invaded the human genome. These viruses, called **endogenous retroviruses**, total 8 percent of the human genome and are made up of **long terminal repeats** (LTRs) that reverse transcribe themselves from RNA to DNA, DNA transposons, and some viruses that lay dormant and can no longer replicate (about 4 percent of the human genome). Some of these sequences are essentially dead genomes that have incorporated into human DNA and are now replicated along with the genome, whereas others are still actively infecting humans by retrotransposition.

The large volume of repetitive noncoding DNA present in the human genome total 98 percent of the human genome. These regions were thought to be nonfunctional, earning the misnomer, "junk" DNA. However, in light of the many functions ascribed to noncoding DNA, it is generally thought that the functions of this "dark matter" in the genome remain to be discovered. Even if a function for a specific sequence has not been identified, it may still be critical to the functioning of other genes, either nearby or surprisingly remote. Even if a sequence is a repetitive TE that is present in thousands of places in the genome, the location of a single TE may be needed for correct activity of that specific gene that now possesses the element. In a complete reversal from only a decade ago, every sequence of the genome is now considered putatively functional, opening the door for a better understanding of the genomic biology of psychiatry and revealing a much larger genome that requires critical examination.

The Mitochondrial Genome

All eukaryotes, including humans, possess two fully-functional living genomes, that is, the DNA particular to that species and the genomes of mitochondria (in animals) or chloroplasts (in plants). The theory of

endosymbiosis suggests that in the early days of evolving life on Earth (1.5 to 2 billion years ago), multicellular life forms merged with respiratory bacteria, which were able to produce cellular energy in a more accessible form (ATP). This relationship proved to be mutually beneficial for both the protoeukaryotes and the respiratory bacteria, since the bacteria could evade prey and the protoeukaryotes had a new source of energy. This theory is supported by the fact that mitochondria have their own genome and separate replication cycles, and that the closest living relatives of mitochondria (rickettsia bacteria) are parasites of eukaryotic cells.

Table 1.11–5.
Human Diseases Arising from Mitochondrial-Encoded Genes

Disorder	Affected Gene(s)
NARP (neuropathy, ataxia, retinitis, pigmentosa)	ATPase6
MELAS (mitochondrial encephalomyopathy, lactic acidosis, stroke)	tRNA (Leu), Cytochrome C
MERRF (myoclonic epilepsy; ragged red fibers)	tRNA (Lys), tRNA (Ser)
LHON (Leber's; hereditary; optic; neuropathy)	MTND4
FRDA (Friedreich ataxia)	FRDA
Leigh's syndrome	ATPase6, MTND5

The human genome and the mitochondrial genome now have a completely intertwined coexistence. Nearly 300 genes from mitochondria are present on the 23 human chromosomes. These are referred to as **nuclear DNA of mitochondrial origin** (NUMTs), and they are produced by the nuclear genome to support the function of mitochondria. Further, 27 of these NUMTs do not appear in the chimpanzee or other genomes, and have therefore been incorporated into the human genome within the last 4 to 6 million years, the period since the divergence with the last common ancestor with chimpanzees. Most of them (23/27) are present within known or predicted human genes, indicating that the symbiosis between these two genomes is significantly gene-centered, perhaps ensuring stronger protection against transposable elements **and/or protect mitochondrial DNA (mtDNA) from damage inflicted by reactive oxygen species (ROS)**. Finally, some neuromuscular diseases and metabolic defects arise from mutations or errors specific to mtDNA or NUMTs (Table 1.11–5), underscoring the importance of the mitochondrial genome in influencing risk to mental disorders.

HUMAN GENETIC VARIATION

The whole-genome sequencing of a reference human in 2001 set the stage for intensive studies of genetic variation between individuals and among populations. Generally, between any two unrelated persons, differences are present at approximately 1 in every 1,000 base pairs, that is, 0.1 percent of their genomes. It also appears that not everyone has two copies of every

part of the genome. These larger regions of variability (>1 kb) occur in heterochromatic and euchromatic regions and are estimated to cover 5 to 20 percent of the total human genomic sequence. The prevalence of these variants has dramatically altered thinking about the total burden of variation within the human species. The search for disease susceptibility genes in psychiatry may be more precisely thought of as the hunt for that small proportion of genetic variation that is either causal or contributing to psychopathology.

Sequence Variation

DNA sequences of individuals vary in terms of the genetic code at the nucleotide level. Such changes may involve the substitution of one nucleotide for another, or in some cases, the insertion or deletion of one or a small number of bases, or **indels.** In addition, short repetitive sequences in the DNA may differ in terms of the number of nucleotides present in these regions, variations that are known as **short tandem repeats** (or STRs) and **small sequence repeats** (SSRs).

There are varying thresholds for classifying a variant as common. Typically, if a sequence variation is found in less than 1 percent of the population, it is considered rare, though some authors use a threshold of 5 percent. Specific genetic variants found above this frequency in a given population are typically referred to as **polymorphisms**, but this term may also be applied to any change in the genome regardless of its frequency and regardless of whether or not it is deleterious to the function of the RNA or protein that it encodes or regulates. Similarly, any permanent change in the DNA nucleotide sequence can be called a **mutation,** although this term is most commonly applied to disease-related genetic variation.

Common single base substitutions with a population frequency greater than 1 percent are often called **SNPs**. These have been intensively studied and proven critical to gene discovery efforts. SNPs are thought to have occurred spontaneously in human history, to have only happened at one point in that history, and to have distributed subsequently throughout a population over time. SNPs are generally stable from generation to generation and a very large number have accumulated in the genome as the human species has evolved. Currently, 30 million known SNPs have been identified. These individual SNPs, or the arrangement of SNPs on a single chromosome (called **haplotypes**), allow scientists to trace inheritance within families or to identify variation that may carry risks for disease within a population. In addition, the distribution of SNPs and haplotypes among various populations has provided critical insight into human evolution. For instance, this type of data, along with other genetic and archeological evidence, has provided strong evidence for the theory that humans originated from African hominids.

Structural Variation

While large microscopic variations including balanced and unbalanced translocations, large deletions or inversions, and extra chromosomes have been studied for decades, only recently has it become apparent that there is also a tremendous amount of submicroscopic structural variation in the genome.

In the late 1990s, several labs developed tools to identify chromosomal aberrations that were below the resolution of the light microscope. Given the known relationship between chromosomal abnormalities and various human disorders, including cancer and mental retardation, it was anticipated that the ability to identify formerly unseen chromosomal anomalies would lead to the identification of additional disease susceptibility genes. The surprising result of these efforts was the identification of differences in the number of copies of chromosomal segments found among the control samples used to validate the technology. Given the clear evidence that these changes were widespread among individuals without any obvious phenotype, these changes became known as **CNPs** or **CNVs**.

Subsequent studies have widely replicated these initial results, and the development of higher resolution technologies has demonstrated that deviations from the expected two copies of each genetic locus are common throughout the genome. These variants are not replacements or changes in sequence (like mutations or SNPs), but are more comparable to indels, where regions of the genome are missing a copy (**heterozygous deletions**), missing both copies (**homozygous deletions**), or have extra copies (**amplifications**). Healthy human beings possess from several dozen to several hundreds CNVs, and may carry as many as a thousand total CNVs. When the sum total of all the identified variants are considered (SNPs, CNVs, and larger DNA deletions and/or amplifications), as much of 20 percent of the genome has been shown to vary between individuals (Fig. 1.11–7).

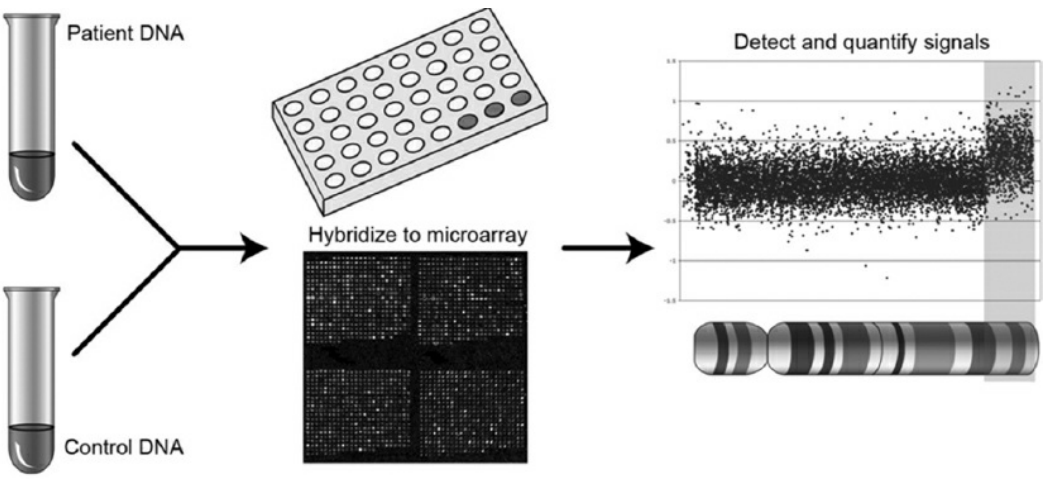

FIGURE 1.11–7. Detecting structural variation in the human genome. Using microarrays, which hold thousands of spots of DNA from the human genome, a comparison between amount of patient DNA and control DNA along the genome can be made. If a region of a chromosome is altered (upward shift in intensity, in *red* along the chromosome), it is then defined as a CNV.

Gene (and large genomic segments) duplications have resulted in many evolutionary changes during the evolution of hominid lineage. Trichromatic vision was developed in the primate lineage around 40 million years ago as a result of gene duplication. Such rearrangement was facilitated by the Alu repeats, which flank the opsin genes, allowing the duplication to occur. The genes coding for cytochrome P450 enzymes frequently contain CNVs in humans. They are key enzymes in drug metabolism, and as such, extremely important in the novel field of pharmacogenomics (discussed below).

The discovery of CNVs has had an enormous impact. Approximately 10 percent of the genes in the human genome, including genes influencing development, overlap common CNVs, suggesting that the human genome is able to tolerate haploinsufficiency (the loss of one of two haploid copies) to a far greater extent than previously anticipated. Several mechanisms lead to CNV formation. Some CNVs appear to have developed once in human history with subsequent distribution among populations. Other CNVs appear to be more dynamic. In this case, a region of the genome may be primed to lead to CNVs, but this process may occur essentially randomly among different individuals and lead to multiple different CNVs in a single chromosomal region.

The relationship between CNVs and human disease is a topic of intense study. It is now clear that the relationship between genotype and phenotype with regard to structural variation is quite varied. Previously, if a deletion was identified in a patient being evaluated for mental retardation, it was presumed that the deletion was the likely cause of the clinical problem. Presently, the identification of a loss of a gene in such a patient, even one known to be involved in brain development, is no longer *prima face* evidence of causality. At present, it is generally presumed that: (1) some CNVs have no phenotypic consequences; (2) as with other forms of variation, some common CNVs will be found to contribute incrementally to complex, multigenic disease; and (3) some copy number changes, particularly those that are rare and or *de novo* (newly created in a single person), may carry large risks for some disorders. Given that the characterization of these types of variation in the human genome is in the early stages, distinguishing among these alternatives in a given patient or population remains a formidable challenge.

Common versus Rare Genetic Variation

Irrespective of whether a genetic variant is sequence-based or structural, once it is introduced into the human genome, one expects that it will be subject to natural selection, that is, changes that do not alter reproductive fitness may be readily passed from generation to generation and, over time, have the potential to become common. Alternatively, changes that result in reduced fitness are likely to be subject to purifying selection, leading the frequency of that variant to decline over time in the population.

The impact on fitness is only one of several forces that influence the frequencies of genetic variants. For instance, a new mutation is, by definition, rare until it has had time to disseminate into a population and be subject to selective pressures. Moreover, the history of ethnic populations and migration patterns can dramatically influence the dynamics of genetic variants over time.

One would presume that rare variants contribute to disorders that are lethal early in development and those that reduce fertility or otherwise lead to decreased reproductive fitness. One would conversely presume that common genetic variants contribute to disorders that appear later in life or for which there is not a negative impact on fitness. Common variants are also presumed to be implicated in conditions (e.g., sickle cell trait) where a positive effect (malaria resistance in this case) of a given genetic variation might counterbalance negative consequences. The competing impacts of positive and negative effects of a gene variant is called **balancing selection**.

The field of human genetics has a record of tremendous accomplishment in those disorders in which a single rare genetic change causes or dramatically increases the risk for a disease or syndrome. However, the task of clarifying disorders in which multiple gene variants contribute to the manifestation of a disease has remained a daunting challenge. Two alternative, but not mutually exclusive, paradigms have emerged to account for the genetics of common complex disease: a ***common variant: common disease*** model and a ***rare variant: common disease*** alternative.

The former has been largely favored in mental disorder; for example, it is widely held that schizophrenia, major depression, bipolar disorder, and autism result from the combined effect of multiple common genetic variants, each with modest effect and interacting with environmental factors to exceed a biological threshold. For these and other complex disorders, both modeling and early experimental evidence have essentially ruled out the contribution of a single gene of major effect, assuming that there is a single phenotype (with relatively little variation) for a particular illness. Extended family members can be found to show signs of subtle, subclinical phenotypes (so-called **endophenotypes**) that appear to be near but not beyond a critical liability threshold. As most of the variation within a population is carried in common variants, it is thought that

common disorders will likely reflect this underlying genetic architecture.

An alternative model posits that many individually rare variants with relatively large effects contribute alone or in combination to the risk for common disorders. This hypothesis seems intuitive for disorders with early onset and those that alter reproductive fitness. In the absence of balancing selection, one would expect that alleles with large effects, even if these are contributory and not causal, would be likely to be rare. In addition, one would expect a significant burden of rare mutations for disorders in which so-called sporadic or *de novo* variation played an important role. Indeed, recent studies show that *de novo* mutations (mostly *de novo* missense and *de novo* likely gene-disrupting mutations) in children affected by autistic spectrum disorder (ASD) may account for more than 10 percent of the diagnosed cases.

These two possibilities are not mutually exclusive. As more research is completed, it is likely that both common and rare variants will be found to contribute to the variable manifestation for many common mental disorders. However, the distinction is of tremendous importance with respect to current genetic studies because the methodologies employed to show a causal or contributory role for genes differ tremendously in terms of their ability to detect the contribution of rare versus common genetic variation. In addition, diseases with complex and extremely variable phenotypes, such as schizophrenia or ASDs, may be caused by various combinations of rare and common variations, which would, at least in part, explain the large spectra of pathologies seen in aforementioned conditions. This issue will be addressed in the Population Genetics and Genetic Epidemiology chapter.

THE FUNCTION OF THE GENOME

Now that the sequence of the human genome has been elaborated, research now focuses on how its 3 billion nucleotides work in a coordinated fashion. Given that the human genome appears largely unremarkable compared to other species in terms of size, gene number, or percentage of the genome devoted to protein coding, one must presume that the capabilities of the human brain must be powerfully influenced by variation in a relatively limited number of genes or influenced by factors other than raw DNA sequence. The characterization of brain-specific functional components of the genome, and the consequences of genetic variation in these elements, will undoubtedly become a prime concern for psychiatric genetics in the coming era. More broadly, such an understanding will identify those unique molecular changes that define *Homo sapiens* as a species.

As recently as the early 1970s, the importance of gene regulation for shaping the development of the human brain remained in question. Several hypotheses were entertained regarding the origin of the differences between humans and other closely related species. It was thought that

speciation might be the result of small-scale sequence changes (DNA or protein sequence chances), regulatory changes (from controller genes or other similar elements), or large-scale sequence change (gene and genome duplication). In 1975, Mary-Claire King and Allan Wilson published a paper in which they sequenced and compared 44 human and chimpanzee genes. To the surprise of many, the vast majority of these sequences were found to be nearly identical among chimp and human and virtually all of the small-scale changes identified were synonymous. Moreover, the gene and genomic organization of the transcripts was essentially unchanged between the two species. These data demonstrated for the first time that neither coding sequence nor genomic organization was likely to account for the differences between the two species.

King and Wilson's work supported the alternative conclusion, that is, the sequence surrounding genes that regulated gene expression accounted for key evolutionary changes that include dramatic differences in cortical architecture. This work also implied that sequencing of the human genome and that of closely-related species would not be sufficient for understanding key differences in their biology. These conclusions largely have been borne out: it is clear that in order to understand brain function, the question of whether a gene sequence is present is often not as important as determining *when* and *how* the transcript is active in that organism. In cases of highly divergent neurodevelopmental disorders, such as autism, it has been shown that there are distinctive spatiotemporal expression patterns of the ASD-associated genes in early developmental stages of the brain. Aberrant spatiotemporal expression of these genes could easily contribute toward developmental impairments, resulting in autism.

Yet another confirmation of King and Wilson's work came up recently with obtaining the genomes of the Old World Apes. Namely, several genomic regions that underwent rapid evolution in humans were identified —human accelerated regions (HARs). While HAR1 is a gene expressed during development of the neocortex, and HAR2 (also known as HACNS1) is a human-specific developmental enhancer, another 721 HARs were identified comparing conservation scores of the sequences. Surprisingly, 92 percent of them are noncoding genes, emphasizing the likely importance of regulatory sequences in recent human genome evolution.

REGULATION OF TRANSCRIPTION AND THE TRANSCRIPTOME

Perhaps the largest contribution to the complexity of eukaryotes comes from **RNA splicing**, the process by which the exons of a gene can be combined in multiple ways (Fig. 1.11–8).

The discovery that genes were transcribed as both exons and introns came from studying the adenovirus gene *hexon* in 1977. The implications of this discovery for human evolution were quickly recognized: a single region

of DNA could lead to a variety of transcripts and multiple versions of a protein. These in turn could become targets of evolutionary selection that might generate new function without having to sacrifice the utility of the original transcript. Thus, the "universe" of the transcriptome suddenly became much larger.

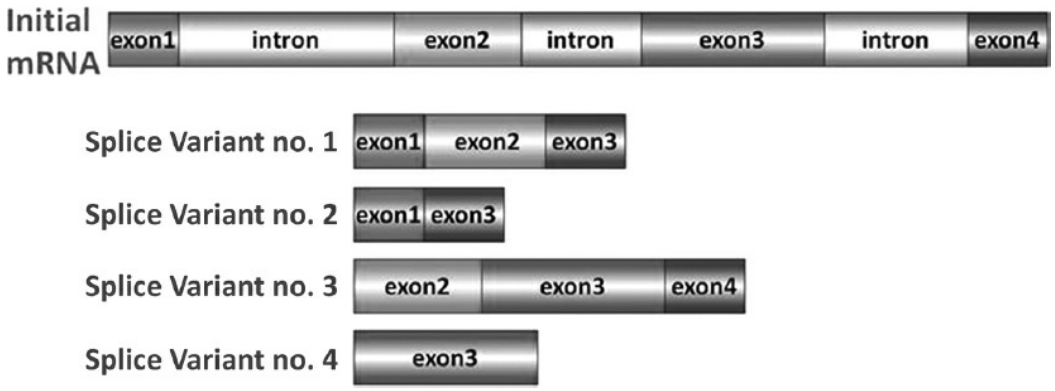

FIGURE 1.11–8. Alternative splicing in eukaryotic genes. Different variations of certain multiexon genes can be generated by the cell, depending on environmental conditions or developmental stage. These different splice variants will then create a slightly modified version of the gene's protein.

During RNA splicing, exons can either be retained in the mature message or targeted for removal in different combinations to create a diverse array of mRNAs from a single pre-mRNA, a process referred to as **alternative RNA** splicing. By 1994, estimates of the number of human genes exhibiting alternative splicing ranged from 1 to 5 percent. However, these estimates were known to be imprecise due to the sparse availability of sequence data. Interest in resolving this issue led to a portion of the HGP being devoted to studying the relationship between gene structure and gene expression, and two methods were utilized: **complementary DNAs** (cDNAs) and **expressed sequence tags** (ESTs). Both methods depend on the mRNA being processed first, which removes all the introns and then adds a long polyadenosine (poly-A) tail to the initial transcript. The presence of these poly-A tails provides a means for transcripts to be bound and isolated in the laboratory by running the total extracted RNA through columns packed with complementary poly-T sequences. Once isolated, the single-stranded mRNAs serve as a template in a reaction to create the complementary sequence, using an enzyme called reverse transcriptase. This creates the cDNAs, which can then be sequenced. ESTs are small fragments of cDNAs, and they are slightly different because they use cloning techniques to digest the cDNA and merge the sequence with a bacterial fragment in order to discern the sequence. Once created and analyzed, these EST and cDNA libraries confirmed a massive amount of alternative splicing in the genome and showed that within genes some exons are always used (constitutively spliced), while others are very rare

and specific to a certain time or place in the body (alternatively spliced).

As these sequence libraries were being created, other technologies simultaneously emerged to allow researchers a clearer view of gene expression and splicing. In 1995, Pat Brown and colleagues first measured expression for 45 genes in parallel in the plant *Arabidopsis* by creating a microarray. They placed complementary sequences to the 45 genes in small spots on a glass microscope slide and then isolated mRNAs from two different *Arabidopsis* samples. The relative expression of each of the 45 genes could be determined between the two samples based on the fluorescence intensity determined at each spot on the slide.

Since this time, a series of technical and methodological developments have led to an astonishingly rapid increase in the density and specificity of the probes (spots) that may be placed on microarrays. Current technologies allow for millions of sequences to be arrayed on a single slide, providing the basis for comprehensive and global studies of gene expression within and among organisms. With hundreds of thousands of cDNAs and ESTs isolated, the current estimate is that at least 80 percent of human genes undergo alternative splicing.

While the size and gene number found in an organism do not generally correlate with biological sophistication of an organism, such is not the case for alternative splicing. At the most fundamental level, this distinction is made clear by the contrast between prokaryotes (which have no alternative splicing) and eukaryotes (which do). There is also some evidence that within higher organisms, increases in splicing rates correspond to greater complexity. For instance, in yeast, most genes (~96 percent) have no introns, and very few splicing events have been discovered. Moreover, recent evidence shows that invertebrates have fewer genes undergoing alternative splicing (50 to 60 percent in the fruit fly) than do vertebrates (80 percent in human).

While the total amount of alternative splicing in the human genome remains a question, it is already clear that brain shows the largest amount of any tissue in the human body. This observation suggests that alternative splicing and its regulation may be quite important for understanding what makes the human brain so different from other closely related species, and in understanding the neurobiology of mental disorders.

Indeed, the process of alternative splicing bestows an opportunity for genes to undergo accelerated evolutionary change. Most genes are under strong negative selection pressure, meaning that it is difficult for them to change, especially if they are critical to early development of an organism. Alternative splicing leading to the formation of a gene's less common products can be an evolutionary shortcut that relaxes this negative selection pressure. Since enough of the major (or more common) form is still present, the new, minor splice form would not be subject to the same degree of negative selection, and may even create a "neutral space" for the

novel gene product to acquire a new function for the organism.

Ample evidence of this process can be found within the DNA sequence. The exonic constituents of the major versus minor forms of a given gene are largely conserved between the mouse and human genomes. The major-use exons typically show higher degrees of sequence conservation than minor-use exons, suggesting stronger negative selection on the former. Cross-species comparisons demonstrate that minor-use exons are a more recent addition in evolutionary time, supporting the hypothesis that the basic biological functions of a gene can be augmented through the addition of new splice variants. There is strong evidence that when a minor-use exon subserves a new important function, its level of conservation quickly approaches that of more ancient exons. These phenomena take place even if the new splice variant is highly restricted temporally or spatially, and this is amply demonstrated in human brain.

It has often been the case that researchers have focused on changes that affect the most highly conserved regions of a gene or the most highly conserved amino acids in a protein sequence, based on the notion that these underpin the most vital functions of the transcript. Recent cross-species comparisons and the enhanced knowledge of splice variation suggests that it is also possible that changes in less-conserved sequence or those that alter only minor splice variants that are spatially or temporally restricted could be quite relevant.

An appreciation of the importance of splice variations points to the critical nature of its regulation. As with many other aspects of the genome, the mechanisms of the process are incompletely understood. It has long been appreciated that there are canonical splice sites that define the boundaries of introns: the 5′ splice site typically has the sequence GT and 3′ site most often has the sequence AG. However, these are not absolute, as they are absent in approximately 2 percent of known introns. It has been appreciated that regulation of gene splicing is also influenced by sequences nearby or within the gene. These include **exonic splicing enhancers** (ESEs) and **exonic splicing silencers** (ESSs), which are specific sequences contained within exons that regulate splicing proliferation or attenuation. Changes in these sequences can have a profound impact on gene function without altering the amino acid composition of a protein. Synonymous sequence mutations, that is, those that alter the nucleotide sequence but do not change the amino acid sequence, involving these motifs have been shown in some cases to lead to serious developmental disorders.

In addition to the accumulating knowledge on gene splicing, several other aspects of regulating the transcriptome have been well characterized. Individual genes are regulated by transcription factors, which determine the timing and degree of a gene's activity. Transcription factors bind to specific sequences upstream of genes called TFBSs or **response elements**

(REs) (Fig. 1.11–9). Most transcription factors consist of a DNA-binding domain which confers some target specificity and an activation domain which confers its regulatory function.

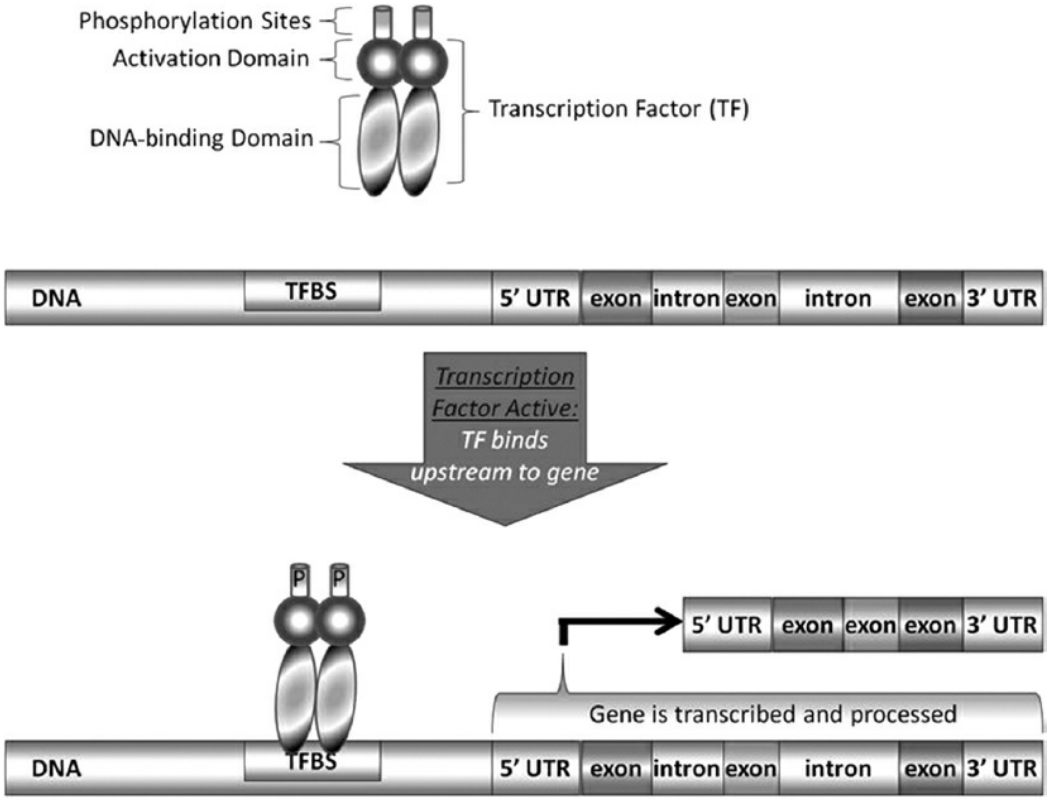

FIGURE 1.11–9. Gene regulation by transcription factors. Once a transcription factor is activated in its activation domain, usually by phosphorylation (the P sites on the TF), the DNA-binding domain is then activated and finds the binding site in front of the appropriate gene (TFBS). Once bound, the gene is then activated.

RNA Editing

Another way of regulating and diversifying the transcriptome is RNA editing. RNA editing is one of the most unusual posttranscriptional RNA modifications, since it can modify the information content of an organism without changing the underlying DNA in the genome. It is mediated by the ADAR type of enzymes, deaminating adenosine (A), and turning it into inosine (I), which is then recognized as guanosine (G) by both the spliceosome and the ribosome (Fig. 1.11–10). This type of transcriptome diversification contributed to evolutionary changes in many cases instead of gene duplication/mutation or alternative splicing, possibly due to its efficiency and low energy consumption. RNA editing may produce changes in the coding sequence, miRNA binding sites, and may create novel splicing sites. Until recently, only a small number of edited human genes were discovered, and most of them have important roles in the CNS. Brain shows the most versatile repertoire of RNA editing events, including AMPA, HTR2C receptors, sodium channels, and calcium channels. Humans have

the highest amount of A to I editing, as compared to rodents, rhesus macaque, and chimp, which may contribute to human traits.

RNA editing occurs almost exclusively on RNAs that form double-stranded secondary structures (dsRNA). Alu elements, due to their specific sequence, are perfect target for RNA editing. Due to the fact that so many genes involved in brain functions undergo editing (calcium, sodium, potassium channels; GABA, serotonin, glutamate receptors), it is thought that RNA editing is a "fine tuner" of the brain transcriptome.

One of the best known eukaryotic regulatory elements is the TATA box, so named because the sequence TATA is present upstream of a gene. This motif binds a protein complex that allows for the initiation of transcription. The core promoter of individual genes need not contain all elements. Many promoters lack a TATA box and use instead the functionally analogous **initiator element** (INR). These sequences are related to the number of protein-coding genes in the genome. Determining the number of coding genes in an organism is simply a matter of counting the number of TATA boxes and INRs in promoters that are nearby coding regions. Sequences like the TATA box, which are immediately upstream of a gene, are called **cis-regulatory** (meaning "on the same side"), due to their presence nearby the gene and on the same strand of DNA. Regulatory elements, TFBSs, or gene products that act on a gene from the other strand of DNA or from a distant region of the genome are called **trans-regulatory elements**.

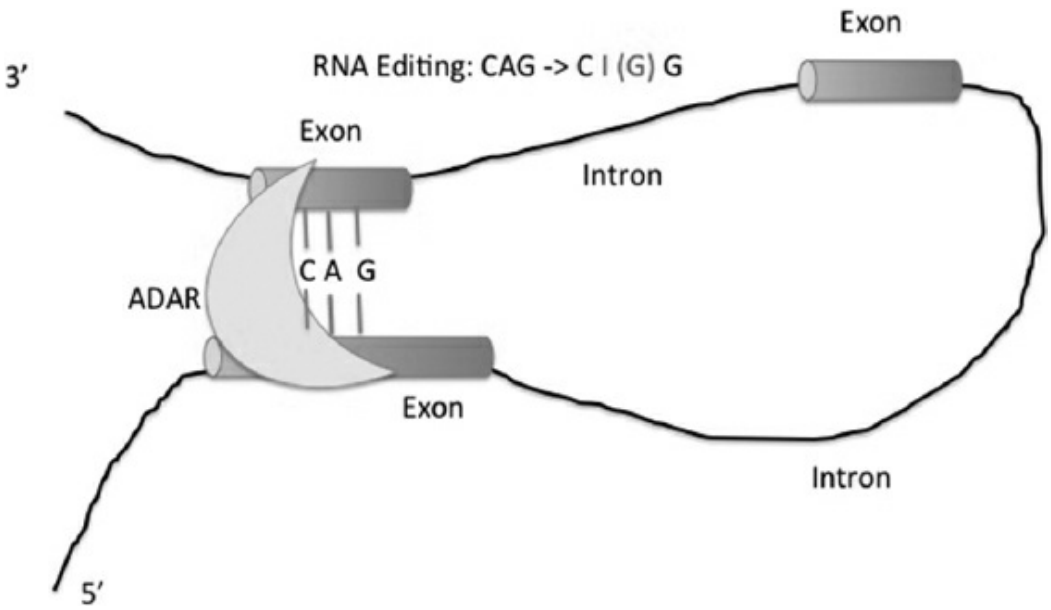

FIGURE 1.11–10. Representation of RNA editing in a coding region of a gene. ADAR-mediated RNA editing. Double-stranded secondary structure of the RNA (*black line*) enables ADAR binding (*light gray*) and deamination of the adenosine in the exon. Deamination derives inosine, which then acts like guanosine in each interaction that requires Watson-Crick base pairing. Nascent CAG codon showed on the scheme, has been changed to CI(G)G.

After the human genome was sequenced, this simple dichotomy of regulation (cis- and trans) has become more complicated. The most well-understood promoter elements (TATA box and INR) are only present in the promoter regions of 54 percent of human genes, where promoters are defined within 100 bp in either direction of the **transcriptional start site** (TSS). This suggests that unknown promoter sequences exist to regulate the expression and function of these genes, or much of gene regulation occurs in trans-regulatory elements. A great deal of bioinformatics research is now focused on this very question.

Over 10 other core promoter elements are now predicted, but there is no longer an expectation that these elements will be universally present in all or even most genes. Rather, genes are activated in large simultaneous groups, or come in multiple waves of activation, and these coexpressed genes can have a specific set of regulatory elements and modules that restrict expression to a certain time or place. Once a large set of coexpressed genes is found, an examination of the upstream regions of those genes can be performed to ascertain the sequences of the regulatory elements.

For a set of coexpressed genes, it is not typical to identify precisely the same regulatory sequence upstream of all of these transcripts. Due to the flexibility in the TF's DNA binding domain, there is often some sequence variability that is tolerated in gene regulation. Studies of many TFBSs have demonstrated that certain nucleotides in the binding site are critical, while others may vary. Aligning these sites across many genes results in the identification of a consensus sequence or consensus motif for a TFBS (Fig. 1.11–11).

Hundreds of transcription factors have now been studied with a technique called ChIP (Chromatin Immuno-Precipitation). The dynamic process of TF binding to regulatory elements is frozen in a moment in time by fixing bound TFs to their cognate regulatory site via a chemical process known as **crosslinking**. Genomic DNA, some of which is now bound by TFs, is then fragmented in a manner that preserves this crosslinking. Using antibodies specific to the known TFs, the bound genomic sequences can be identified, separated from the rest of the genome (the unbound fragments), and then characterized. These bound fragments can sequenced directly to identify where precisely a specific TF is functioning, or may be labeled with different dyes and studied on microarray. This latter method has become known as **ChIP-chip** (chromatin immunoprecipitation on a chip).

These experiments identified a large amount of regulatory complexity in the human genome. The position of the TSS, where the beginning of transcription occurs for a gene, can sometimes be found hundreds of thousands of bases (>100 kb) away from the first exon. Also, many of these regulatory elements, TSSs, and TFBSs for genes are not found beside their respective gene but within introns or within different genes altogether.

Sometimes these regulatory elements allow for genes to behave as discrete elements; however, at times, these regulatory motifs allow for a group of genes to merge together. There have been observations of trans-splicing, whereby exons from one gene can combine with exons from a completely different gene, creating chimeric mRNAs and chimera proteins that contain some or all of the elements from the two independent proteins.

Important points in understanding gene regulation include the following: any regulatory element in the genome should not be presumed to be regulating the gene to which it is adjacent; multiple genes may be activated in concert by a single, distant regulatory element; and a single gene is not always a single gene, and can (if needed) merge with some or all of the parts of another gene, even at relative great distances.

Epigenetics

One additional realm of genetic regulation rests outside of genomic sequence in the **epigenome**. Epigenetic regulation is any heritable change in the genome that does not change the nucleotide sequence. These changes can occur by either a chemical modification to the DNA known as methylation or by chemical changes to the protein complexes of histones. Histone modifications are far more numerous, with phosphorylation, acetylation, methylation, and ubiquitination modulating the activity of a gene or group of genes by altering DNA-protein complex conformation. This in turn has an impact on DNA availability to transcription factors and the recruitment of other molecules that inhibit or enhance gene expression.

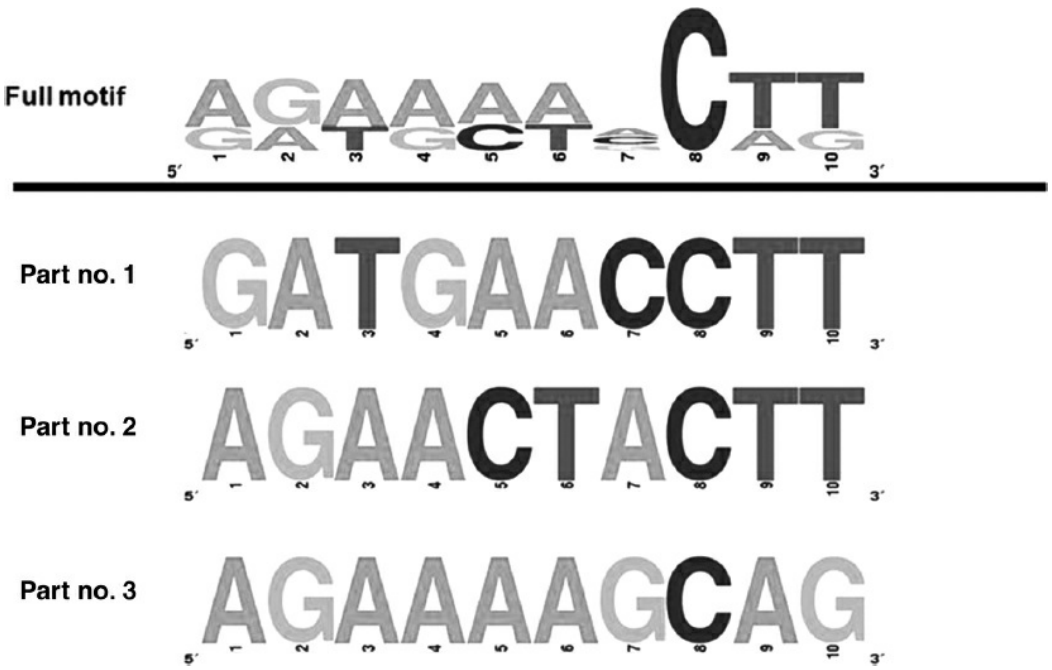

FIGURE 1.11–11. Building a consensus sequence for a transcription factor binding site (TFBS). For any TFBS upstream of many genes, there are often several different functional motifs (Part no. 1, no. 2, and no. 3) that allow for binding of that specific TF. When merged,

these three 10 bp sequences create a consensus motif that degenerates across all nucleotides except for position 8, which is always a cytosine (**C**).

Table 1.11–6.
Effects on Gene Expression of Mono-, Di-, Trimethylation, as well as Acetylation of Lysine for the Most Prominent Histone Markers

	Histone Modification Type			
Histone Type	**Monomethylation**	**Dimethylation**	**Trimethylation**	**Acetylation**
H3K4	Activation		Activation	
H3K9	Activation	Repression	Repression	Activation
H3K14				Activation
H3K27	Activation	Repression	Repression	Activation
H3K79	Activation	Activation	Activation	
H3K36			Activation	
H4K20	Activation			
H2BK5	Activation		Repression	

Epigenetic marks on DNA consist of methylation of cytosine, by specific methyltransferases, thereby deriving 5-methyl Cytosine (5 mC). 5 mC presence is predominant at regions of the genome that contain neighboring CG tandems—CpG sites. Cytosine methylation of CpG sites located at the gene promoters is a hallmark of transcriptionally silenced genes. DNA methylation is a reversible process, and it is thought that the oxidative product of 5 mC, 5-hydroxymethylcytosine (5 hmC) is an intermediate in conversion of 5 mC back into cytosine. Presence of 5hmC in germ cells is in favor of this model, as demethylation (resetting of the epigenetic code) is needed for production of gametes. Additionally, 5hmC shows high concentrations in CNS, with novel data suggesting roles of 5hmC in regulating splicing and gene expression in the brain.

Important histone modifications and their effect on the gene expression are presented in Table 1.11–6.

Probably the most reliable signs of gene repression is the presence of H3K27Me3 at gene promoters as well as H3K9Me2/3, which is a strong mark for the presence of heterochromatin at the specific location. In contrast, H3K4Me3 presence at the promoters is a sign of open, actively transcribed chromatin regions, while H3K36Me3 marks during the length of the gene body is also an indication of an actively transcribed region.

Epigenetic programming allows for tissue-specific regulation of gene function and structural and functional specialization throughout the body. Epigenetics also clearly helps explain differences between identical twins who share the same DNA sequence. This nonsequence-based, heritable change can also lead to parent-specific effects during fertilization, known as **imprinting**. Each person, when making eggs or sperm, imprints the epigenetic characteristics of his or her own sex, since they must reprogram

the epigenetic signature that was given by their own parents. Prader–Willi and Angelman syndromes are examples of diseases attributable to abnormal epigenetic imprinting.

Epigenetic programming not only regulates single genes but also can alter the behavior of entire chromosomes. X-chromosome inactivation is a normal process that occurs in every human female, in which one X chromosome is genetically silenced in every cell. This allows for dosage compensation, in which the amount of expression of any gene mapped to the X chromosome is essentially the same for males and females (even though males carry only a single X). The choice of which chromosome is active (X_a) and inactive (X_i) within a given cell is typically a random process. However, in some cases a particular chromosome is more likely to be inactivated, contributing to aspects of Rett syndrome and other X-linked disorders.

Parent of origin effects, defined as transmission of a risk for disease based on whether an autosome has been transmitted by the mother or father, has been hypothesized to play a role in several common mental disorders. The strongest evidence that this may be the case is the observation that duplications of chromosome 15q inherited from the mother, but not the father, accounts for a small but significant proportion of cases of autism.

Epitranscriptome

Another layer of complexity arises from modifications to the bases in RNA —the epitranscriptome. The current database list 112 forms of RNA modification. One type, methyl-6-Adenosine (m^6A), is found in all of the kingdoms of organisms. m^6A is enriched within 3′ UTRs, and long exons in both mice and humans. Sequences enriched with m^6A are evolutionary conserved. Preliminary analysis showed that m^6A may hinder RNA editing. Also, as m^6A can be demethylated back into adenosine, it may constitute a mechanism for regulating RNA editing. m^6A is implicated in the modulation of the circadian rhythm, perhaps by influencing the trafficking of circadian cycle regulatory genes through the nuclear envelope. Additionally, it may play a role in regulation of translational dynamics for transcripts enriched with m^6A. These functional relationships suggest that posttranscriptional modifications of RNA may modify transcript's properties and function, generating the so-called epitranscriptome.

PROTEIN REGULATION AND THE PROTEOME

Just as gene expression is regulated at multiple steps, so too can the protein products specified by coding genes (Table 1.11–7). Over 20 known modifications may occur to proteins. These modifications include chemical modification of specific amino acids, exchanging one amino acid for

another, attaching other biochemical functional groups, attaching larger molecules (e.g., lipids or carbohydrates), or making structural changes to a protein. Even after a protein is fully formed and biochemically active, it can still be selectively targeted for degradation.

These types of protein modifications play an important role in neurotransmission, since most neurotransmitters are shuttled across synaptic membranes and clefts by proteins that have been specially created to accomplish this task. For example, the protein that transports dopamine across synaptic membranes, the DAT, undergoes a range of posttranslational changes that are specific to its needed function in each cell.

Table 1.11–7.
Regulatory Mechanisms across the Genome, Transcriptome, and Proteome

Name	
1. Epigenetic	Allelic exclusion (imprinting one allele of a gene, so only the other is expressed) X-chromosome inactivation Chromatin remodeling and chemical alteration (methyl, acetyl, phosphate) Short-range cell–cell signaling
2. Pretranscriptional	Binding of tissue-specific transcription factors to activation sites in single genes Binding of a competing, inhibitory factor on the binding site Hormones binding to response elements in inducible genes Use of different transcription start sites, stop sites, and promoters in a gene
3. Posttranscriptional	Alternative splicing within the same gene Tissue-specific RNA editing Translational control mechanisms Interference and degradation by small RNAs (siRNA) Trans-splicing between different genes
4. Posttranslational	Proteolytic cleavage (removal of the start codon methionine) Protein merging into a larger protein complex Glycosylation Targeted for degradation (ubiquitination) Chemical alteration of the protein (methyl, acyl, phosphate, or sulfate groups) Addition of heavy metals

To measure and understand these changes for each protein, a new discipline, **proteomics**, has emerged that examines each individual protein's various chemical modifications and how those proteins interact with each other to create larger cellular complexes and machinery. Components of the proteome have been the focus for intensive study, for example, all the active and interacting proteins in a synapse are dubbed the **synapse proteome** and those in the brain are termed the **brain**

proteome. All of these protein characterizations and interactions require extensive experimental support, large datasets, and overlapping experimental methods.

Mass spectrometry is useful for rapid cataloging of all proteins present within a sample, such as the ~1,000 proteins in the synapse proteome (~1,000 proteins). High-pressure liquid chromatography (HPLC) coupled with mass spectrometry has the capabilities of resolving the full proteome of the brain. Two-dimensional (2D) electrophoresis is used to study protein interactions. In this approach, proteins are pushed across each other's path in a gel, with a change in the migration pattern suggesting a direct interaction. Yeast two-hybrid experiments are another approach to characterizing protein binding partners. In this case, two proteins of interest are cloned into a yeast cell. Each protein is a hybrid, one containing the first protein and the DNA-binding domain of a specific transcription factor, while the other contains the second protein and the activation domain of the same transcription factor. If the two proteins interact, the transcription factor will have both necessary components to activate its target reporter gene, which is typically designed to be easy to assay. More recently, **protein chips**, which are similar to expression microarrays, have been created. These chips place thousands of proteins on a microscope slide, and then they are washed over by a single sample, thus detecting every interaction between a protein and the rest of the queried proteome.

All these methods have enabled discovery and characterization of the proteins made by the ~25,000 genes in the human genome. Current estimates suggest that from this relatively small number of coding genes, as many as 2 million distinct proteins may result. Several groups have emerged to tackle the task of accurately cataloging all of these proteins including the Human Proteome Initiative (HPI) and the Human Proteome Organization (HUPO).

In summary, the complex machinery that regulates the genome can be thought of as encompassing four levels: pretranscriptional regulation (exemplified by the functioning of transcription factors), posttranscriptional regulation (exemplified by miRNAs and siRNAs), posttranslational regulation (exemplified by chemical modification and processing of proteins), and epigenetic regulation.

If the notion of understanding how these processes contribute to the complex functioning of the human brain and how they might contribute to mental illness was not sufficiently daunting, one need only consider the astronomical potential for diversity inherent in the transcriptome and proteome: Though the human genome is only 3 billion letters long, for every gene with more than one exon, the number of possible transcripts increases at the rate of 2^n-1. Consequently, a two-exon gene can be spliced into product containing each one of the two exons, or both combined, totaling three possible transcripts. Similarly, a three-exon gene can be

spliced into one of the three exons, or four other combinations (1–2, 2–3, 1–3, 1–2–3), totaling seven, and so on. When this formula is applied to the gene with the most exons in the human genome (*titin,* with 234 exons), one can see that this one coding unit could theoretically combine into 2.76×10^{70} different transcripts. Additionally, given about 2 million proteins in the human proteome, there exist 4 trillion possible protein–protein interactions. In total, this becomes about 7.0×10^{82} possible gene products and subsequent interactions to examine. As a comparison, the number of atoms in the universe is estimated to be around 4.0×10^{80}. Though these numbers are theoretical approximations, they illustrate the vast, potential complexity of the human genome and the need for sophisticated methods to understand it.

PHARMACOGENOMICS IN NEUROPSYCHIATRIC DISORDERS

Pharmacogenomics is the effort to characterize gene variants that influence the response to medications with the long-term goal of optimizing the treatment of individual patients. Pharmacogenomic tests are already established in oncology clinical practice.

In psychiatry, pharmacogenomics has had its greatest impact through characterizing gene variants coding for P450 microsomal enzymes that influence the rate of metabolism of psychotropic medications. These tests have been applied clinically as an alternative to measuring drug blood levels in order to identify whether lack of clinical response could be attributed to rapid drug clearance, necessitating treatment with higher medication doses. They might identify slow metabolizers, who might be more suspectable to medication side effects. To date, pharmacogenomics studies related to the pharmacodynamics actions of psychotropic medications do not yet have sufficient validity or explanatory power for routine clinical use.

THE BRAIN AS A SYSTEM

The field of **systems biology** focuses on the emergent properties of biological systems. There is increasing interest in developing sophisticated computational approaches that enable scientists to evaluate the interplay of genome, transcriptome, proteome, and other emergent properties of the brain, such as behavior.

Systems biology is a particularly powerful approach to extraordinarily large datasets generated by large-scale international collaborations aiming for a complete functional annotation of the human genome, transcriptome, and proteome. These efforts have been made possible by technological advances that have allowed for the accumulation of very large datasets that encompass the all of the types discussed in this chapter: regulatory sequences, the transcription factors that bind them, the expressed genes (coding and noncoding), their subsequent modifications by other enzymes,

the final protein product(s), and the interaction of that protein with any other protein (**protein–protein interactions** [PPIs]). These data are derived from three sources: (1) computational predictions of transcriptional activity, splicing potential, or protein interactions; (2) experimentally-based observations extracted from published scientific reports and databases; and (3) the use of algorithms that scan all of the world's published scientific literature and report any interaction with a specific gene or protein of interest (**text mining**). Additionally, large databases of protein sequences have been compiled and curated, such as UniProt.

Over the past several years, major advances in bioinformatics and computational biology have allowed entire biochemical or neurological pathway to be modeled and a change in one segment of the pathway (from a mutated exon to a drop in synaptic plasticity) to be traced to every other possibly interacting biological entity. This technique—known as **pathway analysis**—creates predictions for expected changes to other genes or proteins in the same pathway. Already, several such tools are publicly available online, including Ingenuity, DAVID, KEGG, and Interactome, an effort at Cold Spring Harbor Laboratory to build a working model for every interaction (gene–gene, gene–protein, protein–protein) that exists in the human body.

An extraordinary public resource has recently been created that catalogues region-specific expression analyses in the mouse brain, known as Allen Brain Atlas. This database currently contains a three-dimensional rendering of almost every gene's expression pattern. The mouse model serves as a good proxy for understanding region-specific brain expression in humans, and, as a result of this comprehensive resource, essentially any gene of interest can be analyzed with all the others for instance to investigate spatial and temporal correlation among groups of genes or across the entire genome. These and other resources are publically available on the World Wide Web (Table 1.11–8).

FUTURE PROSPECTS

The brain is the most complex organ in the human body and perhaps the most complex structure known to humanity. Scientists have never before had so many tools and technologies to bring to bear to study the brain. These approaches create an unparalleled ability to gather and merge information about a biological system, but the exact methods by which clinicians and researchers can merge patient data and abstract predictions is a work in progress.

Though microarrays and cloning have been indispensible techniques, they are already being eclipsed by new technologies. All of the methods currently used for analyzing and understanding an individual patient's genome are restricted to specific regions, but this is only because of

limitations of the current technology. Once each patient's genome can be sequenced cheaply and quickly, another revolution in the postgenomics era will begin. Instead of searching for mutations or structural abnormalities only in genes or places where a change is suspected, every small- and large-scale genetic variation will be detected for every location in the genome.

Table 1.11–8.
Scientific Resources on the World Wide Web

Electronic Address	Description
http://www.ncbi.nlm.nih.gov	National Center for Biotechnology Information
http://www.gdb.org	The Genome Database
http://www.wellcome.ac.uk/	Wellcome Trust genome research center
http://www.hapmap.org	The HapMap Consortium
http://www.genome.gov	Nation Human Genome Research Institute
http://www.rcsb.org/pdb	Protein Data Bank
http://www.harvester.embl.de	Collection of human proteins and all predictions
http://www.genome.gov/10005107	The ENCODE project
http://genome.ucsc.edu/	UCSC Genome Bioinformatics Center
http://arep.med.harvard.edu/PGP/	Personal Genome Project
http://www.brainatlas.org/aba/	Allen Brain Atlas from Mouse Brains
http://www.geneontology.org/	Functional Annotation of Genes (Ontology)
http://bluebrain.epfl.ch/	Reverse Engineering of the Mammalian Brain
http://www.e-cell.org/	Model and Reconstruct Biology *in silico*
http://www.proteinlounge.com/	Database and Research Tools for Proteins
http://openwetware.org/	Group to share tools for making organisms
http://interactome.org/	Modeling all interactions in Biology

Clearly, the availability of a complete picture of each patient's DNA will vastly empower efforts to identify that proportion of genetic variation that contributes to mental disorders, influences therapeutic response, and mediates the results of environmental stressors. There is no doubt that recent discoveries have already heralded the first steps toward the development of personalized genomic medicine. The so-called $1,000 genome will be a major step toward realizing this potential. The tremendous payoff from molecular genetics will greatly enhance the understanding of DNA sequence and structural changes; these important first steps will be critical in dissecting the molecular neurobiology of mental disorders and in developing new therapeutics.

References

An JJ, Gharami K, Liao GY, et al. Distinct role of long 3′ UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell*. 2008;134(1):175–187.

Bass BL. RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem*. 2002;71:817–846.

Capra JA, Erwin GD, McKinsey G, Rubenstein JL, Pollard KS. Many human accelerated regions

are developmental enhancers. *Philos Trans R Soc Lond B Biol Sci*. 2013;368(1632):20130025.

Drozda K, Müller DJ, Bishop JR Pharmacogenomic testing for neuropsychiatric drugs: Current status of drug labeling, guidelines for using genetic information, and test options. *Pharmacotherapy*. 2014;34(2):166–184.

Dulai KS, von Dornum M, Mollon JD, Hunt DM. The evolution of trichromatic color vision by opsin gene duplication in New World and Old World primates. *Genome Res*. 1999;9(7):629–638.

ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*. 2007;447(7146):799–716.

Frankish A, Harrow J. *GENCODE Pseudogenes*. *Pseudogenes*. New York: Springer; 2014:129–155.

Gerstein MB, Bruce C, Rozowsky JS, et al. What is a gene, post-ENCODE? History and updated definition. *Genome Res*. 2007;17(6):669–681.

Jens M, Nikolaus R. Competition between target sites of regulators shapes post-transcriptional gene regulation. *Nat Rev Genet*. 2015;16(2):113–126.

Kapranov P, Willingham AT, Gingeras TR. Genome-wide transcription and the implications for genomic organization. *Nat Rev Genet*. 2007;8(6):413–423.

Kellis M, Wold B, Snyder MP, et al. Defining functional DNA elements in the human genome. *Proc Natl Acad Sci USA*. 2014;111(17):6131–6138.

Khaja R, Zhang J, MacDonald JR, et al. Genome assembly comparison identifies structural variants in the human genome. *Nat Genet*. 2006;38(12):1413–1418.

King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. *Science*. 1975;188:107–116.

Lander ES, et al, International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860–921.

Levy S, Sutton G, Ng PC, et al. The diploid genome sequence of an individual human. *PLoS Biol*. 2007;5(10):e254.

Liu GE, Alkan C, Jiang L, et al. Comparative analysis of *Alu* repeats in primate genomes. *Genome Res*. 2009;19:876–885.

Mendel JG. Versuche über pflanzenhybriden. In: *Verhandlungen des naturforschenden Vereines in Brünn 4*. Abhandlungen; 1866:3–47.

Meyer KD, Saletore Y, Zumbo P, et al. Comprehensive analysis of mRNA Methylation reveals enrichment in 3' UTRs and near stop Codons. *Cell*. 2012;149(7):1635–1646.

Oldham MC, Horvath S, Geschwind DH. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *PNAS*. 2006;103(47):17973–17978.

Paz-Yaacov N, Levanon EY, Nevo E, et al. Adenosine-to-inosine RNA editing shapes transcriptome diversity in primates. *Proc Natl Acad Sci USA*. 2010;107(27):12174–12179.

Pipes L, Li S, Bozinoski M, et al. The non-human primate reference transcriptome resource (NHPRTR) for comparative functional genomics. *Nucleic Acids Res*. 2013;41(D1):D906–D914.

Plotkin JB, Grzegorz K. Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genetics*. 2011;12(1):32–42.

Pollard KS, Salama SR, King B, et al. Forces shaping the fastest evolving regions in the human genome. *PLoS Genet*. 2006;2(10):e168.

Pollard KS, Salama SR, Lambert N, et al. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature*. 2006;443(7108):167–172.

Stolc V, Gauhar Z, Mason C, et al. A gene expression map for the euchromatic genome of Drosophila melanogaster. *Science*. 2004;306(5696):655–660.

Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science*. 2001;291(5507):1304–1351.

Wang HY, Chien HC, Osada N, et al. Rate of evolution in brain-expressed genes in humans and other primates. *PLoS Biol*. 2007;5(2):e13.

Wilson FH, Hariri A, Farhi A, et al. A cluster of metabolic defects caused by mutation in a mitochondrial tRNA. *Science*. 2004;306(5699):1190–1194.

Zhang ZD, Paccanaro A, Fu Y, et al. Statistical analysis of the genomic distribution and correlation of regulatory elements in the ENCODE regions. *Genome Res*. 2007;17(6):787–797.

▲ 1.12 Psychoneuroendocrinology

DEBRA S. HARRIS, M.D., OWEN M. WOLKOWITZ, M.D., AND VICTOR I. REUS, M.D.

Endocrine disorders are frequently associated with secondary psychiatric symptoms, such as depressed mood and disturbances in thought. In addition, certain psychiatric syndromes are associated with distinct patterns of endocrine dysfunction. The term *psychoneuroendocrinology* encompasses the inextricable structural and functional relationships between hormonal systems and the CNS and behaviors that modulate and are derived from both. Classically, *hormones* have been defined as the products of endocrine glands transported by the blood to exert their action at sites distant from their release. Advances in neuroscience have shown, however, that in the CNS the brain not only serves as a target site for regulatory control of hormonal release but also has secretory functions of its own and serves as an end-organ for some hormonal actions. These complex interrelationships make classic distinctions between the origin, structure, and function of neurons and those of endocrine cells dependent on physiological context. Multiple interactions between the endocrine and immune systems have pointed to a parallel regulatory complexity.

HORMONE EVOLUTION

Over the course of evolution, as organisms have increased in complexity, hormones that first appeared in unicellular organisms have been recruited to serve a multiplicity of functions, a quality that is referred to as *pleiotropy*. A single hormone may act at multiple sites, including binding to receptors on the membrane, cytoplasm, or nucleus, each with different effects, and subtle differences in molecular structure or metabolic processing can have profound physiological consequences. Hormones are thus ideally suited to regulate complex behavioral activities and to play a role in the *plasticity* of the organism, allowing it to adapt to the changing demands of the environment, as, for example, in the alteration of sexual phenotype in certain amphibians and reptiles in response to changing environmental conditions.

HORMONE CLASSIFICATION

Hormones are divided into two general classes by structure—(1) proteins, polypeptides, and glycoproteins; and (2) steroids and steroid-like compounds—and into three classes by location of function (Tables 1.12–1 and 1.12–2). In addition to classical action on a target tissue, hormones may also act as *neuromodulators,* regulating the effects of neurotransmitters and, in some cases, meeting criteria for

neurotransmitter function independently.

HORMONE SECRETION

Hormone secretion is stimulated by the action of a neuronal secretory product of neuroendocrine transducer cells of the hypothalamus. Examples of hormone regulators (Table 1.12–3) include corticotropin-releasing hormone (CRH), which stimulates adrenocorticotropin (ACTH); TRH, which stimulates release of TSH; gonadotropin-releasing hormone (GnRH), which stimulates release of luteinizing hormone (LH) and follicle-stimulating hormone (FSH); and somatostatin (somatotropin release-inhibiting factor [SRIF]) and growth hormone–releasing hormone (GHRH), which influence growth hormone (GH) release. Chemical signals cause the release of these neurohormones from the median eminence of the hypothalamus into the portal hypophyseal bloodstream and subsequent transport to the pituitary to regulate the release of target hormones. Pituitary hormones in turn act directly on target cells (e.g., ACTH on the adrenal gland) or stimulate the release of other hormones from peripheral endocrine organs. In addition, these hormones have feedback actions that regulate secretion and exert neuromodulatory effects in the CNS.

HORMONE SYNTHESIS AND STRUCTURE

Peptide hormones represent subsections of larger amino acid chains or polypeptides called *prohormones*. Production of a peptide hormone occurs by the cleavage of its prohormone chain at a given site on the chain by the appropriate enzyme. Proopiomelanocortin (POMC) is an example of a prohormone that contains the sequences for ACTH, β-endorphin, β-lipotropin, and melanocyte-stimulating hormone (MSH). Some hormones, called *dimers,* contain two or more peptide chains (e.g., FSH, LH, and TSH). Further cleavage of these hormone peptide chains in the course of metabolism may create additional biologically active peptides that have different effects from those of the parent peptide, and even minor modifications of structure can drastically change binding properties and metabolic processing.

Table 1.12–1.
Classification of Hormones by Structure

Structure	Examples	Storage	Lipid Soluble
Proteins, polypeptides, and glycoproteins	Adrenocorticotropic hormone, β-endorphin, thyrotropin-releasing hormone, luteinizing hormone, follicle-stimulating hormone	Vesicles	No
Steroids and steroid-like	Cortisol, estrogens, testosterone, progesterone, dehydroepiandrosterone	Diffusion after	Yes

Tropic hormones, such as ACTH and gonadotropins, in turn, induce steroidogenesis in two distinct ways. Acute regulation occurs through activation and rapid synthesis (over minutes) of steroidogenic acute regulatory (StAR) protein, which regulates the rate-limiting step of steroid hormone synthesis, the transport of cholesterol from the outer to the inner mitochondrial membrane. In contrast, chronic stimulation induces transcription, increases P450scc protein, and steroidogenesis over hours to days.

CELLULAR MODE OF ACTION

Genomic

The first known mode of action of steroid hormones (glucocorticoids, estrogen, and testosterone) and thyroid hormones (triiodothyronine [T_3] and thyroxine [T_4]) is by binding to intracellular receptors in the cytoplasm. The hormone–receptor complex in turn binds to common response elements on chromosomal DNA and alters transcription through a conformational change that unmasks the binding site. The hormone complex can also interact with transcription factors such as those produced by c-*fos,* c-*jun,* or activator protein-1 (AP-1) to amplify or to inhibit gene expression and, by these mechanisms, regulate the induction of such gene products as enzymes and other cell proteins that affect metabolic change.

Nongenomic

Alternatively, as in the case of estrogen-stimulated prolactin release, certain hormones may exert a physiological effect within seconds to minutes, a time course precluding a genomic mechanism. Nongenomic action is hypothesized to involve membrane hormone receptors. Some nongenomic effects appear to be mediated through distinct, nonclassical membrane receptors in that they are not blocked by classical receptor inhibitors and do not require gene transcription, protein synthesis, or a coagonist. Hormones also may act through ion-gated neurotransmitter receptors as coagonists or antagonists, as in the modulation of GABA type A ($GABA_A$) receptors by neurosteroids, or by altering the fluidity and microenvironment of membrane receptors through the intercalation of the steroid in the phospholipid bilayers. Genomic and nongenomic mechanisms may be active simultaneously, with cross-talk a likely occurrence.

Table 1.12–2.
Classification of Hormones by Location of Function

Hormone Classification	Function
Autocrine	Self-regulatory effects
Paracrine	Local or adjacent cellular action
Endocrine	Distant target site

Combined Action

An example of a hormone-induced behavioral response with both genomic and nongenomic mechanisms is corticosteroid stimulation of aggressive behavior in rats. A rapid surge of glucocorticoids precedes the aggressive behavior, but early and later stimulation of aggressive behavior is mediated differently. The initial phase is promoted by a nongenomic mechanism needed for rapid response, which can play a decisive role in the outcome of the encounter. (In the case of rats, an early aggressive response may scare off the intruder and avoid a Fight.) The effects of this mechanism rapidly subside, and aggressive behavior in later phases is stimulated by genomic mechanisms. This later, genomic corticosteroid-stimulated behavior is blocked by a protein synthesis inhibitor, but earlier nongenomic stimulation of aggression is not. Besides meeting the immediate need for rapid action, the nongenomic mechanism may serve the purpose of preparing for the genomic response by activating changes needed for aggressive response, such as changes in energy metabolism.

Tissue Specificity

Many hormones, such as estrogen, act through a multistep process involving a hormone–receptor complex. This complex binds to specific DNA sites called hormone response elements. Formation of a cluster with coactivators, corepressors, and transcriptional factors stimulates transcription and protein synthesis, with the response specific to the type of cell. Tissue specificity may arise from several mechanisms. Selective estrogen receptor modulators (SERMs) are synthetic hormones developed to target certain actions while avoiding other unwanted responses through coactivators or repressors or other elements of the cluster that do not recognize the SERM in certain tissues. Alternatively, many hormones, such as estrogen or DHEA, may act directly or indirectly as a prohormone to be converted into other hormones in specific tissues. For example, tibolone is a synthetic hormone structurally related to 19-nortestosterone derivatives that has weak estrogen, progestogen, and androgen effects in specific tissues, in part because it is differentially metabolized. Accordingly, it has been used to reduce hot flashes and sweating, improve mood and libido and sexual functioning, decrease bone loss, stimulate semantic memory, and lessen vaginal atrophy without stimulating the growth of endometrium or breast tissue as do estrogens.

Table 1.12–3.
Examples of Regulating Hormones

Regulating Hormone	Hormone Stimulated (or Inhibited)
Corticotropin-releasing hormone	Adrenocorticotropic hormone
Thyrotropin-releasing hormone	Thyroid-stimulated hormone
Luteinizing-hormone-releasing hormone	Luteinizing hormone
Gonadotropin-releasing hormone	Follicle-stimulating hormone
Somatostatin	Growth hormone (inhibited)
Growth hormone–releasing hormone	Growth hormone
Progesterone, oxytocin	Prolactin
Arginine vasopressin	Adrenocorticotropic hormone

CHARACTERISTICS OF ENDOCRINE ACTIVITY

In general, hormonal compounds often exert their effect in a tonic rather than phasic fashion, being diffused in a less precise manner than a neurotransmitter and over a longer time period. Theoretically, such a characterization would allow hormones to be more closely linked to integrated behavioral responses. Release of many hormones is pulsatile, and the pattern of these pulses (i.e., duration, interpulse interval, slope of increase or decrease in rate, and amplitude) is crucial to their effects. Other factors that can influence the regulation and effect of a hormone in a given individual include genetic differences, a history of exposure to the hormone during critical developmental encoding periods, the frequency and chronicity of past exposure, the time since last exposure, and the status of other influences on the target system. A decrease in the amplitude of response after repeated exposure is referred to as *habituation,* and an enhancement is termed *sensitization*. Facilitation of a previously habituated response after exposure to a novel stimulus or more severe stressor is called *dishabituation* (i.e., this allows enhancement of the previously habituated response to cope more effectively with the stimulus). In the case of the HPA system, the release of cortisol by the adrenal gland is dependent on an integration of three separate control systems. These include an underlying circadian rhythm regulated by the SCN; a stress-responsive circuit involving inputs to the hypothalamus from the brainstem, limbic system, and cerebral cortex; and a feedback control system exerted through two classes of corticosteroid receptor.

PHARMACOGENETICS AND PHARMACOGENOMICS

The study of pharmacogenetics and pharmacogenomics is leading to a better understanding of the mechanisms of psychopathology and its response to treatment. Pharmacogenetics refers to the effects of single genes. Pharmacogenomics is the genome-wide approach to discovering individual differences in the outcome of drug therapy. Recent studies of treatment response to antidepressants suggest that polymorphisms in some

genes, such as for glucocorticoid receptors (GRs), CRH receptor 1, and a GR-regulating cochaperone, can predict treatment outcome or rapidity of response, but to date studies have been of insufficient size, and the findings are difficult to replicate. Larger scale GWAS are needed. Polymorphisms in genes that regulate hormonal response or in genes for other chemical messengers that modulate hormonal activity may also explain treatment responsiveness through more indirect mechanisms. Examples include genes for neurotrophic factors; excitatory cytokines, such as TNF; α-adrenergic receptors; GABA receptors; and metabolizing enzymes, such as MAO. Treatment interventions, including nonpharmacological ones such as exercise, can produce changes in the expression of certain genes that may be more beneficial to those with particular genotypes. Use of genetically engineered mice has allowed the targeting of specific genes for the study of the dysfunction of hormonal systems in relation to anxiety and depression, including peptides modulating hormone release, receptors, binding proteins, and proteins controlling the access of hormones into the CNS. This research is still in its early stages, but the clinical implications of discovering genetic differences as biomarkers for subtyping of the disorder and selection of treatment agent are exciting.

DEVELOPMENTAL PSYCHONEUROENDOCRINOLOGY AND EPIGENETIC TRANSMISSION

Although a review of the effect of hormones on brain development is beyond the scope of this chapter, it is important to note that hormones can have organizational as well as activational effects. Exposure to gonadal hormones during critical stages of neural development directs changes in brain morphology and function (e.g., sex-specific behavior in adulthood) and differentiation of dopaminergic neurons. Similarly, thyroid hormones are essential for the normal development of the CNS, and thyroid deficiency during critical stages of postnatal life severely impairs growth and development of the brain, resulting in behavioral disturbances that may be permanent if replacement therapy is not instituted. Prenatal exposure in animals to endogenous or exogenous glucocorticoids or to stressful circumstances reduces offspring birth weight and may result in long-lasting changes in immune response, hypertension, hyperglycemia, hyperinsulinemia, cardiovascular function, and neuroendocrine responses and behavior, including attentional deficits, increased anxiety, and disturbed social behavior. Maternal deprivation in strains of rats with increased glucocorticoid response to stress has similarly been shown to lead to increased startle responses, anxiety-like behavior, increased alcohol preference, and difficulties with spatial learning in adulthood.

Epigenetic transmission occurs through changes in chromatin and DNA structure that do not involve changes in the sequence of DNA but which alter gene expression and phenotype. Maternal behavior can cause

epigenetic alterations to steroid receptor genes and produce long-term changes influencing postpartum behavior. Cross-fostered rats show maternal care behaviors similar to those of their "adoptive" mother, rather than to those of their biological parent. This change in hormone-influenced behavior is thought to involve estrogen and oxytocin receptor changes.

PSYCHONEUROENDOCRINOLOGY METHODOLOGY IN HUMANS

Human studies are often limited to examining the relationship between hormone concentrations or changes in concentration and psychiatric disease states, symptoms, neurotransmitter function, or response to treatment. Concentrations may be measured in plasma, urine, saliva, CSF, or postmortem tissue and are sometimes used as indicators of regulatory neurotransmitter function in response to a given stimulus. For example, cortisol or prolactin response to d-fenfluramine has been used to assess serotonin activity and GH response to clonidine (Catapres) to assess dopaminergic function.

One example of a provocative psychoendocrine test used to assess the HPA system is the combined dexamethasone–CRH test, which assesses response to two hormonal stimuli, one inhibiting (dexamethasone [Decadron]) and the other stimulating (CRH), on cortisol secretion. Typically, 1.5 mg of dexamethasone is given in the evening, and plasma cortisol concentration is measured 16 hours later on the following day. An infusion of 100 μg of CRH is then given, and cortisol level and ACTH are measured again several times over the next 75 minutes. Abnormalities in this test are found in a variety of psychiatric illnesses, including bipolar disorder, major depression, schizophrenia, and PTSD, although the sensitivity or specificity for these disorders does not make the test a reliable diagnostic tool. Several studies or case reports have suggested an association between changes in suppression and outcome for these disorders.

Another method of studying the relationship of hormones to psychiatric disorders is the administration of hormones or other secretagogues (substances that cause other substances to be secreted) to experimentally correct an abnormal hormone concentration and examine the effects. Some secretagogues can be given orally to replace a parenterally administered hormone and be given multiple times during a day to mimic circadian rhythms or other variations in concentration when studying circadian effects.

Functional brain imaging studies can help localize the areas of changed activity produced by hormonal action affecting a variety of cognitive and behavioral activities.

HYPOTHALAMIC–PITUITARY–ADRENAL AXIS

Since the earliest conceptions of the stress response by Hans Selye and

others, investigation of HPA function has occupied a central position in psychoendocrine research. CRH, ACTH, and cortisol are all elevated in response to a variety of physical and psychological stresses and serve as prime factors in the maintenance of homeostasis and the development of adaptive responses to novel or challenging stimuli. Glucocorticoids regulate glucose metabolism, blood pressure, immune response, lipid metabolism, glycogen deposition, and energy homeostasis for Fight or flight response and are essential for embryonic development and neonatal survival. A normal glucocorticoid stress response helps to recover after the challenge and helps to store the experience for coping with future encounters, although sustained levels may impair some forms of memory. The hormonal response is dependent not only on the characteristics of the stressor itself but also on how the individual assesses and is able to cope with it. In primates, social status can influence adrenocortical profiles and, in turn, can be affected by exogenously induced changes in hormone concentration. Aside from generalized effects on arousal, distinct effects on sensory processing, stimulus habituation and sensitization, pain, sleep, and memory storage and retrieval also have been documented with CRH, ACTH, and cortisol (or corticosterone).

Exposure to chronic stress produces increased concentrations of CRH and AVP in the PVN of the hypothalamus and, over time, leads to a reduction in CRH receptor number in the anterior pituitary. Release of CRH results in a simultaneous activation of the locus coeruleus noradrenergic circuit, which functionally increases arousal and selective attention and decreases vegetative functions, such as appetite and sex drive. ACTH concentrations are increased in acute stress but diminish over time in chronic stress.

GRs are ubiquitously distributed throughout the body. At least two intracellular receptor subtypes bind corticosteroids: The mineralocorticoid receptor (MR) (or type I receptor) and the GR (or type II receptor). The human GR has an α and a β form, the α form showing high affinity for dexamethasone, modest affinity for cortisol, and low affinity for aldosterone, deoxycortisol, and the sex steroids; and the β form acting as a negative regulator. MRs have high affinity but low capacity, whereas GRs have low affinity but high capacity. Owing to the difference in affinity, low corticosteroid levels generally result in a predominant MR occupation, with higher steroid levels shifting the balance in favor of the GR. “Permissive actions” of MR activation before the onset of stress are tonically involved in the mediation of the initial stress response. When more GRs become occupied, local excitability may decrease or, in some cases, may increase on a short-term basis to inhibit or enhance the initial effects of stress-responsive hormones. Over time, continued stress produces increasing “allostatic load,” the sustained effects of hypercortisolemia giving rise to hyperglycemia, increased visceral fat, elevated blood pressure, decreased

bone density, hyperlipidemia, and changes in electrolytes and immune response. Interactions between MR and GR in the hippocampus may be relevant to understanding the regulation of stress response in depression and the efficacy of antidepressants. Studies of GR function have pointed to a relevant decrement in response to agonist in depressed patients, but MR function is generally preserved.

Three types of inhibitory feedback of glucocorticoids on CRH and ACTH have been characterized. Fast, rate-sensitive feedback occurs, while plasma concentrations of the glucocorticoid are rising and regulate release rather than synthesis of CRH and ACTH. Intermediate, delayed feedback occurs from 1 to 2 hours after steroid administration, is dose-sensitive and duration-sensitive rather than rate-sensitive, and inhibits the release of CRH and ACTH as well as the synthesis of CRH. Slow feedback is similar to intermediate feedback but occurs over a longer period of time (hours) and is distinguished by decreased synthesis of CRH and ACTH (and other POMC derivatives).

Glucocorticoid release is amplified, at least acutely, by serotonergic and cholinergic input and is inhibited by GABA and opioids. Catecholamines play a role in response to stress and interact with the limbic–hypothalamic–pituitary–adrenal axis. Acute addition of glucocorticoids can increase dopaminergic activity in certain areas of the brain, but chronic hypercortisolemia may decrease dopamine activity (depending on the region involved). Glucocorticoids may exert their influence on dopamine activities by functioning as a transcriptional regulator, by acting on promoter regions of dopamine receptors, and by modulating catecholamine biosynthesis.

Pathological alterations in HPA function have been associated with a number of psychiatric disorders, including mood disorders, PTSD, dementias, and substance use disorders (SUDs). Disturbances of mood are found in more than 50 percent of patients with Cushing syndrome (characterized by elevated cortisol concentrations), with psychosis or suicidal thought apparent in more than 10 percent of cases studied. Cognitive impairments similar to those seen in MDD are common and relate to the degree of hypercortisolemia present and a possible reduction in hippocampal size. In general, therapeutically induced reductions in cortisol levels result in a normalization of mood and mental status. Mifepristone (RU486) has been reported to ameliorate psychosis and depression in Cushing's patients, and several studies have reported that it also alleviated psychosis or depression in psychotic depression not associated with Cushing syndrome. In Addison disease (characterized by adrenal insufficiency and diminished glucocorticoid output), apathy, social withdrawal, impaired sleep, and decreased concentration frequently accompany prominent fatigue. Replacement of glucocorticoid results in resolution of behavioral symptomatology, although correction of the

associated electrolyte disturbances by itself does not.

Exogenous administration of synthetic corticosteroids is commonly associated with mild activation, but a higher dose and more sustained treatment may produce depression, mood lability, memory and attentional impairment, and sometimes psychosis. Alterations in HPA function associated with depression include elevated cortisol concentrations, failure to suppress cortisol in response to dexamethasone, increased adrenal size and sensitivity to ACTH, a blunted ACTH response to CRH, and elevated CRH concentrations in brain. In addition to altered slow feedback, several groups have demonstrated decreased sensitivity to glucocorticoid fast feedback as well, and increased cortisol near sleep onset and on first awakening have been shown to be predictive of increased risk of future depression. The pattern of these abnormalities has not, thus far, led to a definitive theory of mechanism, and other elements, such as AVP, are important to understanding the change in homeostasis. The finding that corticosteroids have multiple regulatory effects on serotonergic function, particularly on the serotonin (5-hydroxytryptamine type 1A [5-HT_{1A}]) receptor, may also be relevant, as may be the state-dependent stimulant-like effects that glucocorticoids can exert on mesencephalic dopamine transmission.

Excessive glucocorticoid activity may contribute to the symptoms of psychotic mood disorders. Some mood stabilizers, such as lithium, carbamazepine (Tegretol) and valproate, inhibit the transcription activity of GRs and may exert some of their therapeutic effects in this way, and the development of CRH-1 antagonists as potential pharmacotherapies for depression and anxiety disorders, alone or in combination with current antidepressant medications, is currently underway. Cortisol response to CRH or ACTH is abnormal in some psychiatric illnesses, although studies in this area have been highly variable and confusing. For example, in one study, patients with PTSD exhibited low cortisol levels despite high CRH activity, while in another cortisol response to ACTH was increased.

A number of studies have linked abnormalities in HPA activity, hippocampal volume, and certain memory functions, but it is not clear which is causal. One imaging study of identical twins discordant for PTSD showed similarly diminished hippocampal volume, but treatment of the symptoms of PTSD with SSRIs is followed by increase in hippocampal volume and improvement in memory. Genomic and nongenomic actions of cortisol in hippocampal cells promote enzymatic processes that can lead to cell death, but as the hippocampus mediates negative feedback of cortisol release, hippocampal damage may impair cortisol negative feedback. Increasing glucocorticoid levels by administration or other methods has decreased hippocampal volume in several (but not all) studies in NHPs and other species. PTSD symptoms and cortisol at baseline predicted hippocampal size reduction later in one study, and a decrease in cortisol

following the treatment of Cushing disease has been shown to reverse hippocampal atrophy.

In schizophrenia, difficulty suppressing cortisol after dexamethasone is associated with negative symptoms and cognitive impairment. Clozapine (Clozaril) improves cognitive functioning, possibly in part through preventing or reversing cortisol-induced hippocampal damage, and reverses stress-induced impairment of LTP, a measure of synaptic plasticity important for the storage of information and memory, in the frontal cortex. The GR antagonist mifepristone also blocks cortisol-induced impairment of LTP.

Stress is often reported as a reason for relapse in substance-dependent individuals, and animal studies suggest that glucocorticoid administration is involved in drug self-administration. Glucocorticoid acute enhancement of dopamine activity likely contributes to the motivational changes involved with drug use. Alcohol usage and withdrawal produce profound changes in HPA regulation, pseudocushingoid features are a phenotypic feature of chronic alcohol intake, and HPA adaptation to alcohol withdrawal varies by family history of alcoholism. There is also suggestive evidence that alteration in HPA response to acute alcohol challenge may represent an endophenotype of genetic risk of dependence. CRH is involved not only in the HPA axis but also in extrahypothalamic systems that play a role in the relapse to alcohol and other drug use after stress.

Another disorder associated with a disturbance in glucocorticoid negative feedback is polydipsia. Polydipsic schizophrenic patients, particularly those with hyponatremia, show marked impairment in cortisol suppression of ACTH.

METABOLIC SYNDROME

Metabolic syndrome is a cluster of multiple metabolic risk factors, including elevated insulin levels and resistance, hyperglycemia, visceral obesity, hyperlipidemia, and hypertension. Glucocorticoids interfere with glucose transport and utilization. The insulin resistance and increased insulin concentrations that develop decrease lipid deployment and increase lipid accumulation. GRs are found in high concentrations in intra-abdominal fat tissue, accounting for the truncal obesity resembling that in Cushing disease. Elevated intracellular glucocorticoid tone is thought to be an etiology of metabolic syndrome. Selective inhibitors of 11β-hydroxysteroid dehydrogenase 1, which decrease cortisol production, are thus being tested as a therapeutic intervention. Stress can produce these changes, along with increased vascular resistance through effects on the SNS. Many patients with schizophrenia who exhibit increased cortisol and epinephrine have more central obesity, higher plasma cortisol, and an increased risk of diabetes, even when medication free. Many of the atypical antipsychotics and some of the other psychotropic medications also can

cause hyperglycemia, hyperlipidemia, and visceral obesity, limiting their use. Reports of changes in HPA axis hormones with treatment by typical and atypical antipsychotics have been inconsistent. However, atypical antipsychotics may cause more glucose elevation after glucose challenge compared to typicals, putting patients at greater risk to develop diabetes.

PROOPIOMELANOCORTIN, MELANOCORTINS, MELANOCYTE-STIMULATING HORMONE, AND MELANONIN

POMC is a prohormone from which several hormones called melanocortins are derived. These include ACTH (described above), melanocortins, and MSH.

Studies of rodent sexual behavior suggest that the activation of melanocortin receptors may be an avenue to treat women with hypoactive sexual desire, and the development of melanocortin medications for sexual disorders is underway. Blockade of the melanocortin-4 receptor decreases the reinforcing effects of cocaine, and melanocortins may be neuroprotective for ischemic injury, reducing hippocampal cell death, an area important in learning. Inactivating the melanocortin-5 receptor in rats also reduces pheromone-induced aggression.

MSH, an anterior pituitary peptide, controls the secretion of melatonin and melanin. Melanin is a pigment not only in hair, skin, and eyes but also in neurons in the substantia nigra and the locus coeruleus. Melatonin (described below) is a hormone that plays a major role in regulating the circadian rhythm. Phenothiazines increase pituitary MSH secretion and pigmentation in some patients. In a double-blind, cross-over trial in humans, an infusion of α-MSH resulted in a significant improvement in verbal memory but little change in mood. A dose-related biphasic effect on mood has been reported for MSH-release-inhibiting factor. Recent data indicate that MSH interacts with leptin to counteract NPY, decrease food intake, and increase energy expenditure. It may also antagonize the antidepressant effects and anxiolytic effects of NPY, while also antagonizing the anxiogenic effects of the proinflammatory cytokine interleukin-1β. MSH inhibitors have been found to decrease dyskinetic movements in animals.

Melatonin is a pineal hormone that is derived from the serotonin molecule and controls photoperiodically mediated endocrine events (particularly those of the hypothalamic–pituitary–gonadal [HPG] axis). It also modulates immune function, mood, and reproductive performance; is a potent antioxidant and free-radical scavenger; and may have oncostatic effects. Melatonin has a depressive effect on CNS excitability and exerts neuroprotective effects against excitotoxicity. Melatonin has analgesic effects through its actions on opiate receptors and has regulatory effects on serotonin metabolism. Altered secretory patterns and levels of melatonin have been found in various psychiatric disorders, such as unipolar and

bipolar depression, SAD, bulimia, anorexia, schizophrenia, panic disorder, and OCD. Although suppression of melatonin is not necessary for the efficacy of light therapy in SAD, melatonin can be a useful therapeutic agent in the treatment of circadian phase disorders, such as jet lag, and intake of melatonin increases the speed of falling asleep as well as its duration and quality. A number of synthetic melatonin-like drugs have recently been or are being developed as hypnotic agents.

ENDOGENOUS OPIOIDS

Since the discovery of endogenous opioid receptors and their endogenous ligands in the early 1970s, research into the possible behavioral roles of such compounds has grown at a rapid pace. At least three different receptor systems for these ligands have been identified (μ, δ, and κ), and each of these has subtypes. Opioid receptors μ, δ, and κ are activated by the endogenous ligands β-endorphin, enkephalins, and dynorphin, respectively, among others. σ-Receptors, originally included because some opioids that suppressed coughing acted on them, are no longer considered opioid receptors. β-Endorphin is the principal opioid peptide prototype and, like ACTH, MSH, and β-lipotropin, is derived from POMC. Methionine enkephalin (met-enkephalin) and leucine enkephalin (leu-enkephalin) are two small pentapeptides that also possess direct opioid activity, met-enkephalin being contained in POMC and another precursor, proenkephalin, and leu-enkephalin being contained in the prohormones proenkephalin and prodynorphin. The best-documented function of the endogenous opiates is analgesia and alteration of pain perception, but effects on stress, appetite regulation, learning and memory, motor activity, and immune function also appear to be of physiological importance. CRF and endogenous opioids also interact to coregulate the locus coeruleus, a role important in early adaptation to stress.

Early enthusiasm for the idea that a dysfunctional opiate system was etiologically related to schizophrenia has waned in the face of contradictory findings. Increases in various endorphin compounds have been reported in plasma as well as in postmortem brain tissue of patients with schizophrenia, but studies of short-term and long-term treatment with opioid antagonists show no consistent or reproducible effects on psychopathology. However, hypersecretion of opioids in the CNS of patients with PTSD has been postulated to be an adaptive response to traumatic experience, and CSF β-endorphin concentration is inversely related to intrusive and avoidant symptoms of PTSD. Naltrexone (Revia), an opioid receptor antagonist, has decreased symptoms in autistic children and can improve functioning, with decreases in social withdrawal, stereotypy, and abnormal speech being directly related to decreases in β-endorphin levels.

In animal models, a number of stressors, including those that are purely

psychological, induce opiate-mediated effects such as analgesia and hypomotility that are reversed by the opiate antagonist naloxone (Narcan). Several studies have found that concentrations of plasma β-endorphin in humans are correlated with measures of stress elicited by surgery, exercise, parachuting, or pain. Short-term administration of opioid agonists also increases eating, whereas antagonists reduce food intake by as much as 30 percent, diminish intake of fats and highly palatable foods, and increase caloric expenditure. Thus far, however, their long-term use in obesity and eating disorders has not proven clinically useful. Some studies of opioid antagonist treatment have found certain binge parameters (e.g., duration) to be reduced in bulimia, but no studies have demonstrated weight loss in obese subjects. Naltrexone is helpful as an adjunct in the treatment of alcohol as well as opioid dependence, reducing drinking, craving, the high derived from drinking alcohol, and the likelihood that sampling alcohol would precipitate a relapse. In addition to the μ agonist methadone, buprenorphine, a partial μ agonist, has been helpful for opioid dependence probably both because of its alleviation of withdrawal and because of its blockade of opioid-induced euphoria.

It is well known that exogenous opioids (e.g., heroin and morphine) can induce a euphoric mood state and that exercise increases the release of endogenous opioids and is associated with mood enhancement; these observations, together with findings that exercise-induced mood enhancement is blocked by naloxone, suggest that endogenous opioids are also involved in the mediation of mood. Such conclusions must be moderated, however, by the recognition that additional specific and nonspecific effects on other neurochemical systems are possible contributors to exercise-related mood effects.

HYPOTHALAMIC–PITUITARY–GONADAL AXIS

GnRH is a decapeptide that was sequenced and synthesized by Andrew Schally and colleagues in 1971. GnRH administration results in the rapid release of LH and FSH from the pituitary in healthy subjects and, in some pathological states, such as acromegaly, an abnormal release of GH or prolactin. The cell bodies of GnRH are located principally over the optic chiasm in the arcuate area, with projections to the median eminence, and in the lamina terminalis. GnRH release is stimulated by norepinephrine and is inhibited through negative feedback of gonadal steroids. Administration of GnRH can result in a depressive-like state, characterized by hot flashes, anxiety, insomnia, decreased libido, and fatigue in euthymic subjects, but it is not known whether this is a direct effect of the agent or is caused by the hypoestrogenic state that is produced when GnRH is given continuously. A GnRH analog has been found to have some efficacy in the treatment of paraphilia by decreasing testosterone.

The gonadal hormones (progesterone, androstenedione, testosterone,

E_2, and others) are steroids that are secreted principally by the ovary and testis, but significant amounts of androgens arise from the adrenal cortex as well. The prostate gland and adipose tissue are also involved in the synthesis and storage of dihydrotestosterone and contribute to individual variance in sexual function and behavior.

The timing and presence of gonadal hormones play a critical role in the development of sexual dimorphisms in the brain. Developmentally, these hormones direct the organization of many sexually dimorphic CNS structures and functions, such as the size of the hypothalamic nuclei (INAH3) and corpus callosum, the neuronal density in the temporal cortex, the organization of language ability, and the responsiveness in Broca area. Women with congenital adrenal hyperplasia, a deficiency of the enzyme 21-hydroxylase leading to high exposure to adrenal androgens in prenatal and postnatal life, have been found to be more aggressive and less interested in "traditional female roles" than control female subjects. Sexual dimorphisms may also reflect acute and reversible actions of relative steroid concentrations (e.g., higher estrogen levels transiently increase CNS sensitivity to serotonin). The importance of timing in the exposure to sex steroids is highlighted by studies of administration in gender identity disordered adults. These individuals perform on cognitive tests according to their physical gender and not to their perceived gender. Treatment with sex steroids produces substantial cross-sex changes but no changes in cognitive performance.

Testosterone

Testosterone is the primary androgenic steroid, having androgenic (i.e., facilitating male gonadal development) and anabolic (i.e., facilitating linear body growth and somatic growth) functions. Testosterone is important for sexual desire in men and women. In males, muscle mass and strength, sexual activity, desire, thoughts, and intensity of sexual feelings are dependent on normal testosterone levels, but these functions are not clearly augmented by supplementing testosterone in those with normal androgen levels. The addition of small amounts of testosterone to normal hormonal replacement in postmenopausal women has, however, proven to be as beneficial as its use in hypogonadal men. Testosterone has both genomic and nongenomic actions, including acting directly at the cell membrane or modulating the activity of other membrane receptors or second messengers and regulating the actions of a wide range of neurotransmitters.

Testosterone administration has been shown to result in increased violence and aggression in animals, and testosterone level tends to be correlated with aggression in humans, but anecdotal reports of increased aggression with testosterone treatment have not been uniformly

substantiated in human scientific investigations. Reports may be confounded by factors such as past history and social factors, which are particularly important determinants of the effects of hormones in primates and in humans. For instance, in the cynomolgus monkey, administration of testosterone increases dominant behavior in dominant monkeys and submissive behavior in submissive monkeys; in hypogonadal men, it improves mood and decreases irritability. Psychological tests such as the Point Subtraction Aggression Paradigm have been developed to measure aggression in a human laboratory setting and support the clinically observed association between increased testosterone and aggression, although personality characteristics are confounding factors. Free testosterone in the CSF has been associated with measures of aggression, sensation seeking and monotony avoidance, suspiciousness, and reduced socialization, and testosterone-induced aggression in mice has been associated with a decrease in brain allopregnanolone. Abnormal testosterone levels have been inconsistently reported in a variety of disorders, including schizophrenia, PTSD, depression, and anorexia. Studies of testosterone treatment in depression have generally been inconclusive, although two randomized, placebo-controlled trials found intramuscular (IM) therapy effective in eugonadal men with late-life depression and testosterone gel supplementation efficacious in refractory depression. Testosterone may play a role in PMS, as concentrations are higher in those with the disorder. Testosterone has also been reported to improve mood and fatigue in HIV-positive men.

Anabolic steroids are synthetic derivatives of testosterone modified to enhance their anabolic actions, such as muscle growth, rather than their androgenic actions. Varying effects of anabolic-androgenic steroids on a wide variety of moods have been noted anecdotally. In one prospective placebo-controlled study of anabolic-androgenic steroid administration in normal subjects, positive mood symptoms, including euphoria, increased energy, and sexual arousal, were reported, in addition to increases in negative mood symptoms of irritability, mood swings, violent feelings, anger, and hostility. In addition to a variety of adverse physical effects, anabolic steroids can cause a variety of adverse psychiatric effects, including aggressive behavior, mood and psychotic disturbances, and psychological dependence, particularly when taken in high doses.

Dehydroepiandrosterone

DHEA and DHEA-S are adrenal androgens secreted in response to ACTH and represent the most abundant circulating steroids. DHEA is also a neurosteroid that is synthesized in situ in the brain. DHEA has many physiological effects, including reduction in neuronal damage from glucocorticoid excess and oxidative stress. Behavioral interest has centered on its possible involvement in memory, mood, and a number of psychiatric

disorders. *Adrenarche* is the prepubertal onset of adrenal production of DHEA-S and may play a role in human maturation through increasing the activity of the amygdala and hippocampus and promoting synaptogenesis in the cerebral cortex. DHEA has been shown to act as an excitatory neurosteroid and to enhance memory retention in mice, but studies of DHEA administration to humans have not consistently shown any improvement in cognition. Several trials of DHEA administration point to an improvement in well-being, mood, energy, libido, and functional status in depressed individuals. Administration of DHEA to women with adrenal insufficiency (e.g., Addison disease) has repeatedly been demonstrated to enhance mood, energy, and sexual function; effects in men remain to be assessed. Mood, fatigue, and libido improved in HIV-positive patients treated with DHEA in one study, and DHEA and DHEA-S have been found to be inversely correlated with severity in ADHD. Women diagnosed with fibromyalgia have significantly decreased DHEA-S levels, but supplementation does not improve outcome. Several cases of possible DHEA-induced mania have been reported, and DHEA has been reported to be inversely related to extrapyramidal symptoms (EPS) in schizophrenics treated with antipsychotics. DHEA administration in these cases improves EPS.

Double-blind treatment studies have shown antidepressant effects of DHEA in patients with major depression, midlife-onset dysthymia, and schizophrenia, although beneficial effects on memory have not been reliably demonstrated. A small, double-blind trial of DHEA treatment of Alzheimer disease failed to reveal significant benefit, although a near-significant improvement in cognitive function was seen after 3 months of treatment.

Animal studies suggest that DHEA may be involved in eating behavior, aggressiveness, and anxiety as well, with its effects resulting from its transformation into estrogen, testosterone, or androsterone from its antiglucocorticoid activity, or from direct effects on $GABA_A$, NMDA, and σ-receptors. Because of the putative antiglucocorticoid effects, the ratio of cortisol to DHEA levels may be particularly important in understanding adaptive responses to stress. Both cortisol and DHEA appear to be involved in fear conditioning, with the cortisol/DHEA ratio hypothesized to be an index of the degree to which an individual is buffered against the negative effects of stress. This ratio has been found to be related to some measures of psychopathology and response to treatment, predicting the persistence of the first episode major depression and being related to degree of depression, anxiety, and hostility in schizophrenics and response to antipsychotic treatment. PTSD patients have higher DHEA levels and lower cortisol/DHEA ratios related to symptom severity, suggesting a role in PTSD recovery. Fear-potentiated startle is larger in individuals with high as compared to low cortisol/DHEA-S ratios and is positively associated with

cortisol and negatively with DHEA-S. Greater DHEA response to ACTH is related to lower PTSD ratings, and the cortisol/DHEA ratio to negative mood symptoms. A genetic variation in an ACTH receptor promoter has been found to influence DHEA secretion in response to dexamethasone and may underlie some individual differences in stress response.

Estrogen and Progesterone

Sex differences in the prevalence or expression of psychopathology or response to treatment may be caused by differences in the concentrations of hormones or gender-related CNS differences in brain morphology and function. Mood and some other psychiatric disturbances are particularly likely to occur during times of sex hormone changes in women, such as the postpartum period, premenstrually, and perimenopause. The primary estrogens are estradiol (E_2), estrone (E_1), and estriol (E_3), with E_2 being the major secretory product of the ovaries. Two different estrogen receptors have been identified (α and β), each with different anatomical distributions and physiological effects. Estrogens can influence neural activity in the hypothalamus and limbic system directly through the modulation of neuronal excitability and have complex multiphasic effects on nigrostriatal dopamine receptor sensitivity. Estrogens also enhance dopamine synthesis and release, modify basal firing rates, and can lead to stereotypical behavior in rodents. Accordingly, there is evidence that the antipsychotic effects of psychiatric drugs may change over the menstrual cycle and that the risk of tardive dyskinesia is partly dependent on estrogen concentrations. However, studies of symptom changes in schizophrenic women show significant differences for anxiety–depression and withdrawal–retardation subscales but not for psychotic subscales across the menstrual cycle. Nevertheless, lower levels of estrogen are associated with episodes of acute psychosis in both women and men and with more severe negative symptomatology as well as poorer cognitive function. Estrogen pretreatment attenuates anticholinergic drug-induced problems with attention.

Several studies have suggested that gonadal steroids modulate spatial cognition and verbal memory and are involved in impeding age-related neuronal degeneration. There is also some evidence that estrogen administration may decrease the risk and delay the onset of dementia of the Alzheimer type in postmenopausal women, but acute treatment in dementia has been ineffective in reducing symptoms.

Estrogen has mood-enhancing properties and can also increase sensitivity to serotonin, possibly by inhibiting MAO; in animal studies, long-term estrogen treatment results in a decrease in $5\text{-}HT_1$ and an increase in $5\text{-}HT_2$ receptors. In oophorectomized women, significant reductions in tritiated imipramine binding sites (which modulate

presynaptic serotonin uptake) were restored with estrogen treatment. Severe postpartum depression has been successfully treated with sublingual 17-β-estradiol, as have depressive disorders in perimenopausal women in a large, randomized, double-blind trial. Prophylactic estrogen administration has been reported to prevent recurrence of postpartum depression.

Associations between response to antidepressant treatment and age group in women have been found. Low-dose estrogen augmentation of antidepressant medication for perimenopausal women is reported to improve mood, and poorer response to SSRI treatment in older women can be eliminated by hormone replacement therapy. Inadequate response to SSRIs may be due to a chronic hypoestrogenic state, as serotonin receptor binding appears to be estrogen-dependent. Tamoxifen (Nolvadex), a SERM used to treat breast cancer, showed beneficial effects on mania in one study. The greater sensitivity of women to certain kinds of stress may also be attributed in part to differences in tissue sensitivity. Brain norepinephrine system activation, for example, may be stronger in women because of the differential postsynaptic sensitivity of locus coeruleus neurons to CRH. Many of estrogen's psychiatric effects also could be mediated through its stimulation of BDNF or NPY through the estrogen–BDNF–NPY cascade (see Neuropeptide Y).

Progesterone, the primary progestin, is produced by the corpus luteum of the ovary. Although progesterone itself may be anxiogenic, metabolites of progesterone (allopregnanolone and pregnenolone) appear to have anxiolytic and hypnotic properties via $GABA_A$ agonistic activity. Progesterone is colocalized with serotonin in cells of the median raphe and causes increased serotonin uptake and turnover in the brain in several species. Progesterone, which has antiestrogen effects, such as downregulating estrogen receptors and increasing MAO activity, is often associated with dysphoric mood. The ratio of progestin to estrogen in oral contraceptives has been associated with negative mood change, although this effect varies with depression history.

The association of these hormones with serotonin is hypothetically relevant to mood change in premenstrual and postpartum mood disturbances. Women with a history of depression have higher FSH and LH and lower E_2 levels and are at risk to begin perimenopause at a younger age. Similarly, the 2:1 female-to-male ratio in depression prevalence may speculatively be related to the rapid changes in hormone levels in menarche. Premenstrual dysphoric disorder is a disorder in which a constellation of symptoms resembling MDD (in some respects) occurs in most menstrual cycles, appearing in the luteal phase and disappearing within a few days of the onset of menses. No definitive abnormalities in estrogen or progesterone levels have been demonstrated in women with premenstrual dysphoric disorder, but decreased serotonin uptake with

premenstrual decreases in steroid levels have been correlated with the severity of symptoms in some studies. Progesterone downregulates the estrogen receptor, and it has been suggested that, despite high circulating concentrations of estrogen, the luteal phase is a period of functional estrogen withdrawal, with concomitant effects on the serotonergic system. Recent evidence indicates that the abrupt decline of progesterone and allopregnanolone in the luteal phase results in increased production of the α_4 subunit of $GABA_A$ and changes in receptor sensitivity that could account for the typical behavioral symptoms noted. This effect is correlated with the insensitivity of the GABA receptor to modulation by the benzodiazepine class of tranquilizers (and hence is anxiogenic). SSRIs, particularly fluoxetine (Prozac), have demonstrated efficacy, and as many as 50 percent of women may respond to fluoxetine administered only in the second half of each cycle. Alprazolam (Xanax), a $GABA_A$ agonist, has been found to be more effective than placebo in several studies for the treatment of premenstrual dysphoric disorder. In women with severe symptoms that are not responsive to these treatments, the long-term use of a GnRH agonist to abolish menstrual cycling, with added estrogen–progestin, may be therapeutic.

Menstrual phase also has been associated with aspects of substance abuse. Although reports vary, craving for cigarettes and tobacco withdrawal appear to vary with menstrual phase (worse in the luteal phase). Women show greater heart rate and pleasurable drug effects after cocaine administration during the follicular phase but report that cocaine improves dysphoric mood during the luteal phase.

The bulk of the psychological symptoms associated with menopause are actually reported during perimenopause rather than after complete cessation of menses. Reported symptoms include worry, fatigue, crying spells, mood swings, diminished ability to cope, and diminished libido or intensity of orgasm. Estrogen replacement alone may be beneficial, but combination androgen–estrogen replacement may be superior for reinstating energy, a sense of well-being, and libido. In women with an intact uterus, the addition of progestin is necessary to protect against endometrial hyperplasia but can attenuate the beneficial effects of estrogen on mood.

The postpartum period appears to be a particularly risky time for the emergence or relapse of psychiatric illnesses. Women with bipolar disorder appear to be particularly sensitive to alterations in gonadal steroid level. A high risk of relapse during the postpartum period has been observed and appears to be maintained over subsequent pregnancies; evidence for familial preponderance of the puerperal trigger also exists, suggesting a genetic contribution. During pregnancy a number of mechanism play a role in dampening the response to stress in order to promote maternal care and survival of the offspring. Hormones such as oxytocin and prolactin produce

inhibitory effects on the HPA axis, decreased excitatory activity, and a more positive mood state. The reversal of many of these effects may contribute to the vulnerability of women to postpartum psychiatric illnesses. Lactation has the potential to dampen some postpartum changes and has been associated with decreased maternal postpartum depression.

PREGNENOLONE AND ALLOPREGNANOLONE

Neuroactive steroids, such as pregnenolone and allopregnanolone (allo), modulate activity at $GABA_A$, NMDA, σ-1, $5\text{-}HT_3$, nicotinic, kainate, oxytocin, and glycine receptors, among others. Pregnenolone is a neurosteroid synthesized from cholesterol in the brain and is partially metabolized to all subsequent steroids. It, and especially its sulfate, pregnenolone sulfate, appear to have memory-enhancing properties in animal studies. Pregnenolone increases the rate and extent of formation of microtubules, important in neuronal plasticity and function. Progesterone counteracts this effect. Pregnenolone sulfate is an excitatory neurosteroid and has GABA inhibitory effects.

Allo is a neurosteroid derived from progesterone with high concentrations in the CNS. It acts as a $GABA_A$ receptor agonist, decreases CRH concentrations in the hypothalamus, and reduces the anxiety evoked by CRH in rats. Testosterone-induced aggression in mice is thought to be mediated by a downregulation in allo, and normalization of allo with progesterone and estrogen prevents aggression. SSRIs increase brain allo levels in rodents. Social isolation also downregulates allo and increases aggression. Fluoxetine reduces both allo downregulation and aggression. CSF levels of allo in humans have been correlated with clinical improvement over the first 8 to 10 weeks of SSRI treatment, an effect that appears to be independent of the serotonin reuptake inhibition. Allo increases in response to ethanol administration and may play a role in ethanol withdrawal through modulation of $GABA_A$. Administration of allo also has been shown to reduce anxiety and hyperlocomotion associated with benzodiazepine withdrawal in animals. Allo is currently unavailable for clinical use, but studies have already commenced on clinical treatment with allo or with related compounds.

PROLACTIN

Since its identification in 1970, the anterior pituitary hormone prolactin has been examined as a potential index of dopamine activity, dopamine receptor sensitivity, and antipsychotic drug concentration in studies of CNS function in psychiatric patients and as a correlate of stress responsivity. The secretion of prolactin is under direct inhibitory regulation by dopamine neurons located in the tuberoinfundibular section of the hypothalamus and is, therefore, increased by classical antipsychotic medications. Prolactin

also inhibits its own secretion by means of a short-loop feedback circuit to the hypothalamus. In addition, a great number of prolactin-releasing or -modifying factors have been identified, including estrogen, serotonin (particularly through the 5-HT_2 and 5-HT_3 receptors), norepinephrine, opioids, TRH, T_4, histamine, glutamate, cortisol, CRH, and oxytocin, with interaction effects possible. For example, estrogen may promote the serotonin-stimulated release of prolactin.

Prolactin is primarily involved in reproductive functions. During maturation, prolactin secretion participates in gonadal development, whereas, in adults, prolactin contributes to the regulation of the behavioral aspects of reproduction and infant care, including estrogen-dependent sexual receptivity and breast-feeding. In female rats, prolactin secretion is strongly stimulated with exposure to pups. In women, basal prolactin levels are elevated in the postpartum period before weaning, and prolactin release is stimulated by suckling. Hyperprolactinemia is associated with low testosterone in men and reduced libido in men and women. In rodents, prolactin is increased along with corticosterone in response to such stressful stimuli as immobilization, hypoglycemia, surgery, and cold exposure and may be specifically associated with the use of passive coping in the face of a stressor. Prolactin promotes various stress-related behaviors in rats, depending on the condition, such as increasing object-directed exploration while decreasing other exploration.

Hyperprolactinemic patients often complain of depression, decreased libido, stress intolerance, anxiety, and increased irritability. These behavioral symptoms usually resolve in parallel with decrements in serum prolactin when surgical or pharmacological treatments are used. In psychotic patients, prolactin concentrations and prolactin-related sexual disturbances have been positively correlated with the severity of tardive dyskinesia. Prolactin levels are also positively correlated with negative symptoms.

HYPOTHALAMIC–PITUITARY–THYROID AXIS

Thyroid hormones are involved in the regulation of nearly every organ system, particularly those integral to the metabolism of food and the regulation of temperature, and are responsible for the optimal development and function of all body tissues. Moreover, rates of secretion and metabolism of all other major hormones (cortisol, gonadal hormones, and insulin) and catecholamines depend on thyroid status. The thyroid gland secretes two thyroid hormones: T_3 and T_4. T_3 is the more potent of the two, and most of the T_3 circulating in the blood is created by the peripheral metabolism of T_4. The brain relies on its own conversion of T_4 to T_3 rather than on circulating T_3. The hypothalamus secretes TRH into the capillaries of the pituitary portal venous system, and the pituitary responds with the

synthesis and secretion of TSH, which stimulates thyroid cells. Negative feedback regulation occurs when T_3 and T_4 act in the pituitary and hypothalamus to inhibit TSH and TRH, respectively. Finally, a corticotropin-release-inhibiting factor (CRIF) has been identified in the rat that inhibits the synthesis and secretion of ACTH. This peptide, prepro-TRH 178–199, is derived from the prohormone TRH and may play a role in integrating the regulation of the HPA and HPT axes. The expression of genes that encode for thyroid hormone receptors, such as TRα and TRβ, plays a major role in the regulation of neuronal differentiation and the action of immediate early genes. As in the case of steroids, thyroid hormones regulate the transcription of a variety of genes through binding to thyroid response elements (TREs) in regulatory sequence elements.

There is general agreement that central noradrenergic systems are primarily stimulatory to TSH secretion and that central dopamine neurons inhibit TSH release. Thyroid hormones, in turn, are important regulators of central adrenoceptor function, generally decreasing presynaptic noradrenaline release and increasing postsynaptic β-adrenergic receptor number. Hypothyroidism is conversely associated with decreased β-receptor number. Changes in serotonin function are also apparent, with T_3 increasing 5-HT in frontal cortex and inducing downregulation of $5\text{-}HT_{1A}$ autoreceptors. These changes in neurotransmitter release and receptors in response to thyroid hormones parallel the alteration in α- and β-receptor sensitivity associated with pharmacological and electroconvulsive antidepressant treatments and may explain the therapeutic efficacy of supplemental thyroid hormone in treatment-resistant depression. Alternatively, therapeutic benefit may be secondary to the alteration of gene expression and a remodeling of synaptic connectivity. In addition to its prime endocrine function, TRH has direct effects on neuronal excitability, behavior, and neurotransmitter regulation, particularly on central cholinergic systems located in the septohippocampal band and on mesolimbic and nigrostriatal dopamine systems. In lower animals, TRH possesses mild stimulant properties. Initial reports of its mood-elevating effects in healthy human subjects led to a number of projects investigating its short-term and long-term antidepressant effects in clinical populations. Despite some initial enthusiasm, the degree of mood alteration does not seem to be great nor is its occurrence reliable.

Given these observations, it is not surprising that alterations in behavioral function have been observed in patients with primary thyroid gland dysfunction, beginning with the earliest reports in the medical literature. It has been noted that thyroid disorders may induce virtually any psychiatric symptom or syndrome, although regular associations of specific syndromes and thyroid conditions are not consistently found. Hyperthyroidism is commonly associated with fatigue, irritability,

insomnia, anxiety, restlessness, weight loss, and emotional lability; marked impairment in concentration and memory may also be evident. Such states can progress into delirium or mania, or they can be episodic in nature. On occasion, a true psychosis develops, with paranoia being a particularly common presenting feature. In some cases, psychomotor retardation, apathy, and withdrawal rather than agitation and anxiety are the presenting features. Symptoms of mania also have been reported after rapid normalization of thyroid status in hypothyroid individuals and may covary with thyroid level in individuals with episodic endocrine dysfunction. In general, behavioral abnormalities resolve with a normalization of thyroid function and are responsive symptomatically to traditional psychopharmacological regimens. Long-term residual complaints of fatigue, cognitive impairment, and emotional distress have been reported in some individuals even after remission of the precipitating thyroid dysfunction. Caution should be exerted, however, regarding the use of MAOIs or TCA medications in hyperthyroid states because of possible synergistic cardiotoxicity. In several case reports, haloperidol (Haldol) has been linked to increasing thyrotoxicity, and hyperthyroidism has been associated with an enhancement of the neurotoxic effects of antipsychotic medications.

The psychiatric symptoms of chronic hypothyroidism are generally well recognized. Most classically, fatigue, decreased libido, memory impairment, and irritability are noted, but a true secondary psychotic disorder or dementia-like state also can develop. In milder, subclinical states of hypothyroidism, the absence of gross signs accompanying endocrine dysfunction may result in its being overlooked as a possible cause of a mental disorder. Accordingly, the evaluation of the basal TSH concentration or the TSH response to TRH infusion is necessary to arrive at the proper diagnosis. Figure 1.12–1 illustrates a characteristic physical sign of advanced hypothyroidism. Hypothyroidism impairs neurogenesis in the hippocampus, and this impairment is associated with rat models of depression. Thyroid hormone administration corrects the impairment and reverses the depression.

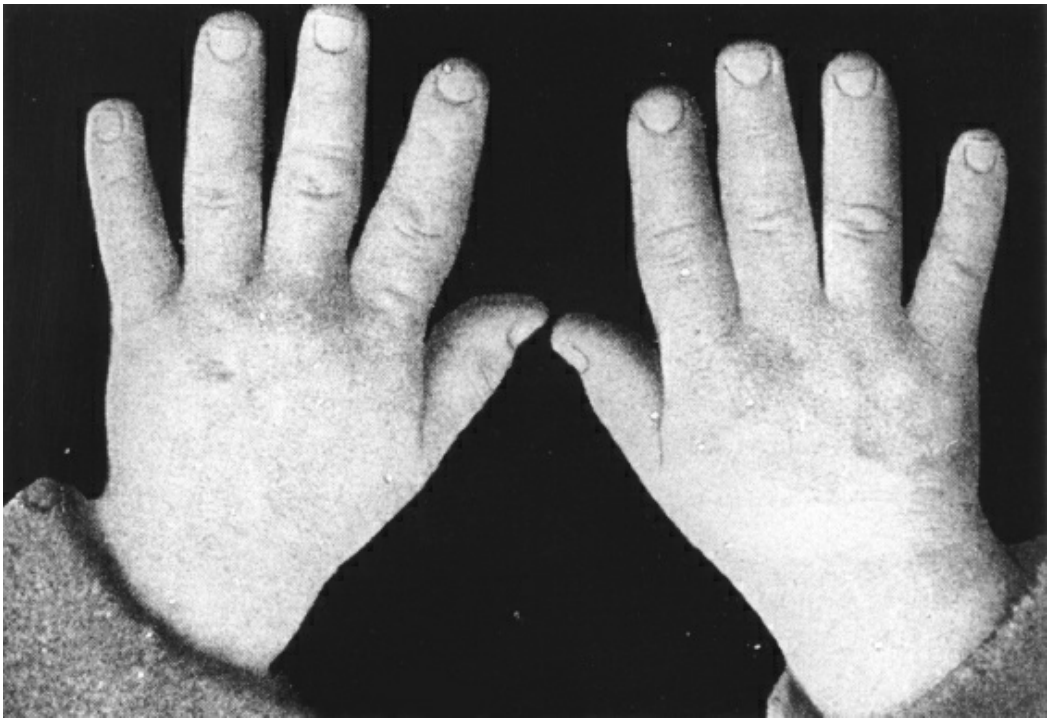

FIGURE 1.12–1. Hands of a patient who has hypothyroidism (myxedema), illustrating the swelling of the soft parts, the broadening of the fingers, and their consequent stumpy or pudgy appearance. (From Waterfield RL. Anæmia. In: Douthwaite AH, ed. *French's Index of Differential Diagnosis*. 7th ed. Baltimore, MA: Williams & Wilkins; 1954.)

A blunted response of TSH to TRH infusion has been found in a significant percentage of patients with a variety of disorders, including eating disorders, panic disorder, alcoholism, schizophrenia, and, most commonly, MDD, and probably reflects a transient hyperthyroxinemia. No evidence of TRH hypersecretion has been shown. Large-scale studies suggest that such subjects are in fact euthyroid, and predictive sensitivity of the test is low. Thyroid autoimmunity may play a role in some psychiatric illnesses. Antithyroid antibodies are found more frequently in women with depression than in control subjects and in patients with bipolar disorder and may contribute to relative treatment resistance as well as to postpartum behavioral disturbance. The frequency of thyroid antibodies may also be higher in certain nonpsychiatric diseases with prominent psychiatric symptoms, such as fibromyalgia and rheumatoid arthritis. Patients with major depression also have been found to have low levels of CSF transthyretin, a protein involved in thyroid transport. Basal T_3 level has been inversely related to time to episode recurrence, and basal TSH positively correlated with overall severity of depressed mood in an unselected inpatient population. Depressed patients who show an improvement in mood after one night of sleep deprivation also appear to have lower T_3 uptake at baseline and greater nocturnal TSH release. Thyroid receptor alterations can produce symptoms of ADHD in animals, and a genetic resistance to thyroid hormone has been associated with

ADHD symptoms in humans. Thyroid receptor coactivators also may play a role in increased vulnerability to a number of psychiatric disorders, including depression and psychosis.

Most antidepressant therapies have some influence on thyroid concentrations at baseline; T_4 and T_3 concentrations have been correlated with antidepressant response, as have antidepressant-induced changes in thyroid hormones as well as changes induced by ECT. Lithium increases antithyroid antibodies and inhibits iodine uptake into the thyroid, iodination of tyrosine, release of T_3 and T_4 from the thyroid, and peripheral breakdown of thyroid hormones. It also regulates *TR* gene expression, blocks the thyroid-stimulating effects of TSH through interference with adenylate cyclase, and may, in certain circumstances, precipitate a rebound thyrotoxicosis. Approximately 30 percent of patients receiving lithium have an elevated TSH level during treatment, and approximately one-sixth of these patients go on to develop frank hypothyroidism. Attention to subtle alteration in thyroid status induced by lithium treatment is important in the clinical evaluation of symptomatic complaints, such as fatigue, memory impairment, and anhedonia; a specific association between lower serum T_4 and mood instability during lithium maintenance suggests even subclinical changes may be clinically relevant. Carbamazepine, an anticonvulsant, shown to have antimanic properties akin to lithium also decreases peripheral thyroid hormone concentrations while increasing TSH. Administration of T_3 accelerates clinical response to TCAs and is sometimes helpful in patients with treatment-resistant depression, whereas adjunctive T_4 contributes to decreasing cycling in patients with rapid-cycling bipolar I disorder. Supraphysiological dosing is sometimes required, with 200 to 500 μg per day being given as an adjunct in treatment-resistant depression and intractable bipolar disorder. Administration of mirtazapine (Kemeron), a mixed action antidepressant, increases free T_3 levels and decreases free T_4. Higher T_3 concentrations predicted improvements in depression in one study.

PARATHYROID HORMONE

Parathyroid hormone (PTH) was originally isolated as an endocrine factor having effects on bone, gut, and kidney and contributing to calcium and phosphorus homeostasis. However, the frequent, and often profound, neuropsychiatric changes that can result from altered parathyroid gland function are consistent with other central actions of PTH that have been described in recent years. Hyperparathyroidism can cause lethargy, stupor, coma, depression, delirium, psychosis, primarily visual hallucinations, or anxiety. Hypoparathyroidism can cause cognitive impairment, psychosis, depression, or anxiety by alterations in calcium and magnesium levels. PTH administration can impair the active uptake and release of

norepinephrine and dopamine and result in adrenergic-like effects (not blocked by β-adrenergic antagonist), learning and memory problems, and a state of hyperalgesia.

Lithium treatment can raise the concentrations of serum calcium and may increase PTH over a period of months to years by a direct stimulation of PTH secretion and through a shift in the set point for inhibition of PTH secretion by calcium. When such effects are associated with somatic or behavioral changes, discontinuation of lithium should result in rapid symptomatic improvement. When this does not happen, a parathyroid adenoma is sometimes discovered fortuitously. Primary hyperparathyroidism most commonly occurs secondary to a single parathyroid adenoma, the removal of which almost invariably results in a lysis of behavioral symptoms, regardless of severity or chronicity. Animal studies suggest that long-term lithium administration can stimulate the development of extant parathyroid tumors but does not induce tumors in normal parathyroid tissue. Thus, reinstitution of lithium treatment after surgical removal of the tumor should be possible.

GROWTH HORMONE

Somatotropin or GH, a hormone required for normal growth, is synthesized and released by the anterior pituitary gland. Dopamine, serotonin acting at the 5-HT_{1D} receptor, and norepinephrine acting at the α_2-adrenergic receptor appear to play a role in its release. GH acutely stimulates lipolysis and ketogenesis, important in the adaptation to stress, and prevents hypoglycemia. Most psychiatric studies of the regulation of GH have used strategies similar to those described for prolactin. Accordingly, studies of GH response to various provocative stimuli, such as to GHRH or psychotherapeutic drugs, have been seen as a means to evaluate central neurotransmitter function. Augmentation of GH secretion in response to GHRH, LH-releasing hormone (LHRH), or TRH in patients with schizophrenia or dementia of the Alzheimer type has been interpreted as reflecting an alteration in catecholamine and, possibly, prostaglandin regulation, which facilitate the secretion of human GH. In general, however, there is a large variation in GH response to GHRH; a blunted response has been variably linked to length of illness, presence of negative symptoms, and platelet MAO activity, but the validity of the conclusions drawn from this test is controversial.

The stress responsiveness of somatotrophs is well established but species-dependent, with increases in circulating GH noted in humans and inhibition of secretion noted in rodents. In humans, the direction of the GH stress response may depend on the persistence of the stressor. GH appears to be relatively more responsive to exercise and hypoglycemic stress than to psychological stress. However, GH has been reported to increase in

response to psychological stress in anxious subjects, perhaps due to hyperactivity of the noradrenergic system. Case reports have documented reversible GH deficiencies and marked growth retardation and delay of puberty secondary to stressful experience. Administration of GH to individuals with GH deficiency has a beneficial effect on cognitive function in addition to its more obvious somatic effects. A significant proportion of adult-onset patients with GH deficiency are depressed, and GH therapy significantly improves their depression scores. Some prepubertal as well as adult patients with diagnoses of MDD show hyposecretion of GHRH during an insulin tolerance test, a deficit that has been interpreted as reflecting alterations in cholinergic and serotonergic mechanisms. Blunted response to 5-HT_{1D} agonists also has been found. Panic disorder patients may have a blunted GH response to clonidine (Catapres), an α_2-adrenergic agonist, that does not normalize with antidepressant treatment. A number of GH abnormalities also have been noted in patients with anorexia nervosa, but secondary factors, such as weight loss, may be responsible for such alterations in endocrine release in depression and eating disorders. At least one study has reported that GHRH stimulates food consumption in patients with anorexia nervosa and attenuates elevated food consumption in patients with bulimia. Administration of GH to elderly men results in an increase in lean body mass, but controlled trials have been unable to replicate anecdotal reports of improved mental clarity, muscle strength, or vigor. Recent evidence indicates that a novel GH secretagogue (GHS), ghrelin, may represent an important alternative regulatory influence over food intake and sleep pattern.

Many GHS can be administered orally, in contrast with GH that must be injected on a daily basis. In addition they may be used to fine-tune body concentrations in a way that daily injections cannot. GHB, a potent GHS, has been used to mimic the physiological secretory pattern of GH and as a way to increase SWS in patients with fibromyalgia and, secondarily, to a reduction in pain and fatigue. The use of the GHS GHB also has been used by bodybuilders as a way of increasing muscle mass, but respiratory depression and sedation and dependence with severe withdrawal effects may result.

SOMATOSTATIN

Somatostatin (SRIF) is a hypothalamic tetradecapeptide that is located principally in the nerve endings of the median eminence and in neurosecretory neurons located in the PVN. SRIF inhibits anterior pituitary secretion of ACTH, thyrotropin, GH, and prolactin, alters release of catecholamine neurotransmitters, and stimulates serotonin release. A number of receptor subtypes have been cloned, and receptor-specific ligands have been developed. SRIF was so named because of its action in

inhibiting the release of immunoreactive GH, a function that is subserved by SRIF-2 receptors.

In rats, SRIF delays the extinction of active avoidance behavior and antagonizes amnesia induced by electric shock. Alterations in the concentration of SRIF have been associated with a number of conditions in which cognitive dysfunction is present, including Huntington disease, Parkinson disease, multiple sclerosis, and Alzheimer disease. Decreases in SRIF are highly correlated with decreases in AChE, suggesting a close relationship between the cholinergic and somatostatinergic systems. Decreased concentrations of SRIF in the CSF are inconsistently found in patients with depression, and central injection of SRIF in rats causes decreased slow wave and REM sleep, altered appetite and locomotor activity, impaired cognition, and decreased sensitivity to pain. Early stressful experiences also have been related to sustained elevations of CRH and somatostatin in the CSF of adult primates. Altered somatostatin concentrations have been reported in a number of other illnesses and with treatment, but the physiological relevance of these changes is still unclear.

ARGININE VASOPRESSIN

AVP (or antidiuretic hormone [ADH]) is a posterior pituitary hormone that maintains plasma osmolarity through the regulation of renal water excretion and that stimulates hepatic glycogenolysis. AVP release is triggered by pain, emotional stress, dehydration, increased plasma osmolarity, or decreases in blood volume and acts synergistically with CRH to control ACTH release. AVP potentiates the stimulatory effect of CRF. An AVP receptor antagonist blocks ACTH release, norepinephrine release, and hyperthermic response to stress and attenuates some stress-related behaviors. AVP receptor blockade does not impair motor or cognitive processes or produce tolerance, as do many other anxiolytics.

Animal and normal human studies of AVP administration (or longer-acting synthetic analog compounds) have indicated that the hormone may enhance the consolidation and retrieval of memory, particularly that associated with aversive learning. AVP has been shown to prevent the loss of tolerance to the incoordinating, sedative-hypnotic, and hypothermic effects of alcohol after cessation of ingestion and to delay the loss of sexual behavior after castration.

Altered AVP function has been reported in depression and in eating disorders. Anorexic and bulimic patients show hypersecretion of centrally directed AVP, and patients with bulimia nervosa or depression may have an attenuated AVP response to hypertonic saline. Vasopressin delays the extinction of behaviors acquired during aversive conditioning and may be related to obsessional preoccupation with the aversive consequences of eating and weight gain. An inverse relationship between AVP concentration and motor activity in depression and an increased number of vasopressin

and oxytocin neurons also have been reported in the hypothalamus of depressed patients. Although dexamethasone suppression of ACTH and cortisol release is attenuated in depressed patients, suppression of ACTH and cortisol release in response to vasopressin is not.

Profound alterations in fluid ingestion and excretion have been observed in psychiatric patients. Polydipsia occurs in 10 to 15 percent of hospitalized psychiatric patients and is unrelated to diagnosis; in many cases, the syndrome is secondary to inappropriate secretion of AVP, which occurs as a feature of the altered behavioral state itself and resolves with treatment or, conversely, is precipitated by a variety of antidepressant or antipsychotic agents. In contrast, fluoxetine treatment of depression decreases the CSF vasopressin levels. Alprazolam, an inhibitor of CRH secretion, inhibits the vasopressin-stimulated release of ACTH and cortisol.

Several subtypes of AVP receptors and been discovered, including $V_{1a}R$, $V_{1b}R$, and V_2R. $V_{1a}R$ knock-out (KO) mice show impairments in spatial learning, and both $V_{1a}R$ and $V_{1b}R$ KO mice show impairments in prepulse inhibition and social behavior. AVP $V_{1a}R$ KO mice have reduced social recognition. With transgenic technology, this gene, when introduced into the nonmonogamous mouse, increased a behavior associated with monogamy. AVP antagonists prevent bonding in the monogamous prairie vole. Animal studies show that centrally released AVP produces anxiogenic and depression-like actions, including generating passive coping strategies in stressful situations. That AVP modulates the stress response is shown by the ability of $V_{1a/b}$ receptor antagonist to normalize the usually abnormal dexamethasone–CRH test in rats bred for high anxiety (HAB). Serotonin increases the release of AVP, but SSRIs reduce AVP secretion or normalize AVP overactivity in HAB rats, resulting in the normalization of the dexamethasone–CRH test and less depression-like behavior in response to stress. Animal studies of the role of AVP in social behavior suggest that receptor variations may be involved in such disorders as schizophrenia and autism. In fact, several studies have found an association between an AVP receptor gene and autism and social behavior in a nonclinical population.

OXYTOCIN

Oxytocin is a posterior pituitary hormone that is involved in osmoregulation, the milk ejection reflex, food intake, and female maternal and sexual behaviors and has many effects reciprocal to those of vasopressin. Convergent evidence, using a range of methodologies, indicates that oxytocin inhibits food and sodium intake. Oxytocin binding in the hypothalamus is increased by estrogen and glucocorticoids and in estrogen-primed women. Oxytocin also can act as a neuromodulator of limbic dopamine concentrations and thus may be involved in the adaptation to substances of abuse, and it can act as a mediator of the effect

of CRH on ACTH. Oxytocin has anxiolytic activity.

Many of oxytocin's behavioral effects are affiliative, and it promotes a variety of reproductive (grooming, arousal, lordosis, orgasm, nesting, and birthing) and maternal behaviors (breast-feeding and mother–infant bonding). Infusion of oxytocin in female subjects of monogamous species facilitates pair bonding in the absence of mating, and administration of an oxytocin antagonist prevents pair bonding. It has been called the *amnestic neuropeptide* owing to its ability to attenuate memory consolidation and retrieval. Patients with autism and anorexia have been reported to have reduced levels of oxytocin. Autistic children also do not show the expected increase in oxytocin with age. In adults with autism, oxytocin infusion reduces repetitive behaviors. In one animal study, chronic administration of phencyclidine (PCP), which lowers hypothalamic oxytocin, also decreased social interaction. An association between an oxytocin receptor gene and autism has recently been identified. Oxytocin interacts with the mesolimbic dopamine system and is believed to facilitate the acquisition of drug use disorders, particularly of some of the "party drugs," such as MDMA and GHB, which are often used to promote social behavior.

NEUROPEPTIDE Y

NPY is closely linked with stress response and with the action of a number of steroid hormones. NPY is widely distributed throughout the CNS and is one of the most conserved peptides in evolution, suggesting an important role in the regulation of basic physiological function, including learning and memory. NPY and NPY-related peptide bind to at least five receptors, which are widely distributed but relatively concentrated in the hypothalamus, the hippocampus, and several other limbic regions. NPY is synthesized in the arcuate nucleus of the hypothalamus. Immunoreactive NPY is found in the serotonin-containing raphe nucleus, and it has been implicated in the modulation of emotional processing. Marijuana use appears to elevate the expression of NPY-1 receptor mRNA levels, perhaps explaining some of the drug's effects. NPY has been found to increase feeding, particularly carbohydrate ingestion, and to counteract leptin effects in a variety of animal models. It has a mutually inhibitory relationship with insulin, and its release is stimulated by stress and corticosteroids and associated with norepinephrine release.

NPY has been studied for its potential anxiolytic, antinociceptive, antihypertensive, and memory-enhancing effects and for a possible role in seizure disorder, schizophrenia, and depression. Treatments for depression, such as some antidepressants, lithium (Eskalith), and ECT, increase NPY concentrations in a number of brain areas in rats, while significantly low levels of NPY have been found in the temporal cortices of patients with schizophrenia. Greater NPY release during stress is associated with less psychological distress in humans and the modulation

of the activity of GABAergic neurons. Neuroactive steroids in turn regulate NPY transmission. NPY appears to influence alcohol consumption and alcohol and morphine withdrawal effects. Reduced levels of NPY are found in discrete brain regions of alcohol-preferring rats, and polymorphisms of the NPY gene are associated with alcohol dependence and with alcohol withdrawal seizures in humans. Polymorphisms in the genes for the NPY peptide and for a promoter region also have been found to be associated with depression in some individuals.

GALANIN

Galanin is an inhibitory peptide that is stimulated in a coordinated fashion with gonadal steroid release. Its documented actions include increased release of GH and inhibition of insulin release, locus coeruleus noradrenergic firing, and acetylcholine release as well as impairment of memory and the mediation of some emotional responses. Studies are underway using galanin receptor agonists and antagonists and galanin receptor KO mice to explore its role in mediating anxiety- and depression-related behavior and in decreasing opiate withdrawal.

INSULIN

Insulin is a protein hormone secreted by the β cells of the pancreas in response to elevations of glucose and amino acids; insulin receptors occur in high density in the hippocampus and are believed to help neurons to metabolize glucose by controlling the transport across cell membranes. Some atypical antipsychotics impair response to insulin and raise blood glucose, increasing the risk of developing diabetes (see Metabolic Syndrome). Psychotic stress itself may impair insulin sensitivity. Those antidepressants that predominantly increase catecholamine activity, such as many of the TCAs, also reduce sensitivity to insulin. Those that increase serotonergic function, such as the SSRIs, increase sensitivity to insulin and are preferred in the treatment of depression comorbid with diabetes. However, insulin may play a more active role in psychiatric symptoms as increasing evidence indicates that insulin may be integrally involved in learning, memory, and mood. Depression is frequent in patients with diabetes, as are indices of impaired hormonal response to stress, but it is not known whether these findings represent direct effects of the disease or are secondary, nonspecific effects. Higher fasting insulin levels also have been associated with better psychopathology scores in schizophrenic patients.

LEPTIN

Leptin is a protein hormone synthesized and secreted in a pulsatile fashion by adipose tissue and involved in the regulation of food intake. Obesity is

associated with leptin resistance, principally its metabolic actions, because sympathetic effects are preserved. Leptin also affects the HPG axis, inhibits insulin-induced steroidogenesis and human chorionic gonadotropin-induced testosterone secretion, and may play a role in menstruation, pregnancy, lactation, puberty, and amenorrhea due to weight loss in anorexia nervosa. Leptin stimulates hematopoiesis, T-cell activation, phagocytosis, and cytokine production and decreases susceptibility to infection. Mediators of leptin action include orexigenic neuropeptides, such as NPY, galanin and galanin-like peptide, and melanin-concentrating hormone (MCH), and anorexigenic neuropeptides, such as CRH and α-MSH hormone. Weight gain produced by some atypical antipsychotics may be mediated in part through increases in leptin. Patients with major depression have lower leptin levels, and these levels are inversely correlated with depression severity. However, higher leptin levels have been found in subjects exposed to trauma with hyperarousal, and these levels are related to hypervigilance. Because these relationships were correlations, the causal relationship is not known. Increased leptin may be an attempt at adaptation, as postulated for changes in opioids and neuroactive steroids in PTSD.

CHOLECYSTOKININ

CCK is a peptide neurotransmitter originally isolated from the gut. In addition to its presence in pancreas and the gastrointestinal (GI) tract, CCK is found in high concentrations in the cerebral cortex, limbic system, and hypothalamus. CCK is involved in the regulation of such behavioral functions as inhibition of intake of solid and liquid food, production of satiety, pain relief (probably from modulation of the endogenous opioid system), cardiovascular and respiratory function, neurotoxicity and seizures, sexual and reproductive behaviors, and memory. Of the two identified receptor subtypes, CCK type A (CCK-A) is found primarily in the periphery and in some discrete brain areas, whereas CCK type B (CCK-B) is plentiful in the brain. The primary form of CCK, CCK-8S (a sulfated octapeptide), coexists with dopamine in the VTA and substantia nigra, and its interactions with dopamine are context- and location-specific. CCK stimulates the synthesis of NGF and plays a neuroprotective role. CCK also modulates HPA activity.

Of particular interest to psychiatry is the colocalization of CCK with dopamine in mesolimbic and mesocortical, but not in nigrostriatal, systems. CCK contributes to the modulation of dopamine-mediated behavior and might be dysregulated in psychiatric syndromes thought to involve altered dopamine transmission. CCK-A receptor antagonists have been proposed for the treatment of schizophrenia, and initial evidence suggests that CCK agonists may be useful for decreasing the severity of parkinsonian symptoms. Cerulein, a mixed CCK-A and CCK-B agonist, has

weak, neuroleptic-like effects on prepulse inhibition in schizophrenic patients. Evidence is stronger for a role of CCK-B receptor antagonists in the treatment of anxiety. CCK has anxiogenic effects in some animal models, and several small-scale human studies have demonstrated that the administration of CCK-4 or pentagastrin can induce panic attacks and can increase neurosteroid release in a significant percentage of healthy volunteers as well as in anxiety disorder patients, even in the absence of arousal or environmental stress. Panic attacks in subjects with preexisting panic disorder can be elicited at doses of CCK-4 that do not reliably induce panic in healthy subjects, indicating enhanced sensitivity. Not only can selective CCK-B antagonists completely abolish the anxiogenic effects of CCK-4, but in an animal model of anxiety used to evaluate the efficacy of benzodiazepines, CCK-B antagonists also demonstrated independent anxiolytic properties. However, CCK overactivity may be involved in submissive behavior in animal models, and CCK receptor expression is higher in suicide victims.

Modulation of mesolimbic dopamine-related behavior, including exploratory and rewarded behaviors may underlie CCK's effects on substance use. Central CCK activity has been linked with preference for drugs of abuse, such as cocaine or alcohol, and polymorphisms of the CCK gene may be one of the risk factors for smoking. This suggests a potential role for CCK receptor antagonists in the treatment of drug dependence.

GASTRIN AND GASTRIN-RELEASING PEPTIDE

Gastrin is a peptide hormone closely related to CCK that stimulates the secretion of gastric acid by the stomach. Pentagastrin is a synthetic polypeptide with effects similar to those of gastrin. It is a CCK agonist and produces anxiety and panic in patients with anxiety disorders and to a lesser extent in those without anxiety disorders. It increases ACTH and cortisol release. Gastrin-releasing peptide (GRP), as its name implies, stimulates gastrin release but has a number of other actions, including interacting with GABA, dopamine, and GRs. GRP appears to enhance memory storage, and a GRP receptor antagonist impairs emotionally motivated memory tasks in rats. GRP-receptor-deficient mice show increased locomotor activity and changes in social behavior.

NEUROTENSIN

Neurotensin is a tridecapeptide that appears to play a role in neuroendocrine regulation and coordination as a signaling molecule. Gonadal and adrenal steroids and thyroid hormones alter neurotensin levels in the hypothalamus, preoptic area, and arcuate nucleus. Neurotensin has a close neuroanatomical relation with serotonin and dopaminergic pathways and is involved in the control of anterior pituitary activity, stimulating the release of prolactin and TSH, as well as in the

regulation of a subpopulation of serotonergic neurons in the dorsal raphe and frontal cortex and GABAergic and glutamatergic neurons. Stimulation of serotonin neurons may be responsible for its analgesic effects and reduction of stress response, whereas the effects on dopamine suggest a possible antipsychotic role. Subgroups of drug-free schizophrenic patients have low neurotensin CSF concentrations and altered neurotensin receptor binding in the entorhinal cortex. Psychotogenic drugs (e.g., methamphetamine) inhibit the release of striatal neurotensin via an inhibitory effect of the dopamine type 1 (D_1) receptor. Most antipsychotic drugs increase neurotensin concentrations in the nucleus accumbens and caudate nucleus; schizophrenic patients with decreased CSF neurotensin show an increase compared to baseline values after antipsychotic drug treatment and clinical improvement. Because of neurotensin's association with the nigrostriatal dopamine and the serotonin systems, it is suspected of playing a role in movement disorders caused by antipsychotic drugs. Central administration of neurotensin in rats produces motor effects seen in animal models of parkinsonian and dystonic reactions (catalepsy) and tardive dyskinesia. Neurotensin may exert an antipsychotic action through intramembrane receptor interactions that reduce affinity of the dopamine type 2 (D_2) agonist binding site. Overexpression of neurotensin 1 receptors in rats results in decreased activation of the mesolimbic dopamine system, similar to that produced by atypical antipsychotics but without changing prepulse inhibition or locomotor behavior, as atypical antipsychotic drugs do. This suggests that neurotensin receptor agonists may be candidates for the treatment of psychosis and attenuate dopamine-induced motor behaviors.

An involvement in the development of drug dependence also has been hypothesized. Blocking the effects of neurotensin with antiserum or a receptor antagonist enhances dopamine release in the nucleus accumbens, and neurotensin itself blocks stimulant-induced motor activity. However, doses that block hyperlocomotion do not attenuate the self-administration of cocaine and even enhance conditioned place preference, an animal model of rewarding effects. Neurotensin's modulating effects on dopamine activity may depend on stimulus intensity, enhancing the rewarding properties of a subthreshold stimulus from intracranial self-stimulation (ICSS), as do psychostimulants, but decreasing maximal stimulation rate, as do antipsychotic drugs. Acute administration of stimulants increases neurotensin in the nucleus accumbens, but with chronic administration, levels return to normal. Ibogaine, a hallucinogen used in indigenous religious ceremonies, has been shown to interrupt cocaine and methamphetamine abuse in patients and may act similarly to ICSS, increasing neurotensin concentrations in the nucleus accumbens when given alone but attenuating cocaine-induced increases in neurotensin. Neurotensin also mimics many of the effects of alcohol, chronic alcohol

downregulates neurotensin receptors, and lower concentrations of neurotensin have been found in the frontal cortices of alcohol-preferring rats.

FUTURE DIRECTIONS: ENDOCRINE VARIABLES IN THE DIAGNOSIS AND TREATMENT OF PSYCHIATRIC DISORDERS

Although it is clear that alterations in endocrine regulation are involved in the pathophysiology and treatment responses of many psychiatric disorders, incorporating these findings into clinical diagnostic assessment and decision making remains problematic. Questions about state/trait differences and the role of confounding variables plague nearly all observations, and most of the findings to date are based on small numbers of subjects studied under experimental conditions. Large-scale longitudinal or cost/effectiveness studies are rare, despite indications that baseline alterations in glucocorticoid regulation and thyroid status (two of the best studied abnormalities) may actually be useful in subtyping psychiatric disorders and in prediction of outcome. Although the dexamethasone-suppression test (DST) was perhaps prematurely put forward as a diagnostic aid, the fact remains that alterations in HPA/stress regulation underlie a number of psychiatric diagnoses and may serve as complementary independent variables in defining treatment response and course of illness to the classical behavioral categories that have thus far defined current practice. Studying genetic polymorphisms in factors regulating hormonal response may help us better understand the influence of hormonal variability on the illness and also possible underlying differences in the nature of the illness reflected in these genetic subtypes. Incorporation of endocrine variables into the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) is premature, but routine assessment of endocrine status in clinical trials and epidemiologic surveys is long overdue. Incorporation of biological variables into psychiatric diagnosis and treatment decision-making requires prospective proof of concept studies that unequivocally demonstrate the clinical superiority of such assessments over that of behavioral variables alone. Unfortunately no such data exist at the present time.

REFERENCES

Amin Z, Mason GF, Cavus I, Krystal JH, Rothman DL. The interaction of neuroactive steroids and GABA in the development of neuropsychiatric disorders in women. *Pharmacol Biochem Behav*. 2006;84:635.

Baribeau DA, Anagnostou E. Oxytocin and vasopressin: linking pituitary neuropeptides and their receptors to social neurocircuits. *Front Neurosci*. 2015;9:335. Review.

Boules M, Shaw A, Fredrickson P, Richelson E. Neurotensin agonists: potential in the treatment of schizophrenia. *CNS Drugs*. 2007;21:13.

Caceda R, Kinkead B, Nemeroff CB. Neurotensin: role in psychiatric and neurological diseases. *Peptides*. 2006;27:2385.

Celec P, Ostatníková D, Hodosy J. On the effects of testosterone on brain behavioral functions. *Front Neurosci*. 2015;9:12.

Champagne FA. Epigenetic mechanisms and the transgenerational effects of maternal care. *Front Neuroendocrinol*. 2008;29(3):386–397.

*Chrousos GP, Kino T. Glucocorticoid action networks and complex psychiatric and/or somatic disorders. *Stress*. 2007;10:213.

Dubrovsky B. Neurosteroids, neuroactive steroids, and symptoms of affective disorders. *Pharmacol Biochem Behav*. 2006;84:644.

Duval F, Mokrani MC, Ortiz JA, Schulz P, Champeval C. Neuroendocrine predictors of the evolution of depression. *Dialogues Clin Neurosci*. 2005;7:273.

Faron-Górecka A, Kuśmider M, Solich J, et al. Involvement of prolactin and somatostatin depression and the mechanism of action of antidepressant drugs. *Pharmacol Rep*. 2013;65(6):1640–1646. Review.

Fliers E, Alkemade A, Wiersinga WM, Swaab DF. Hypothalamic thyroid hormone feedback in health and disease. *Prog Brain Res*. 2006;153:189.

Frye CA. Progestins influence motivation, reward, conditioning, stress, and/or response to drugs of abuse. *Pharmacol Biochem Behav*. 2007;86:209.

Genazzani AR, Pluchino N, Luisi S, Luisi M. Estrogen, cognition and female aging. *Hum Reprod Update*. 2007;13:175.

Heinrichs M, Gaab J. Neuroendocrine mechanisms of stress and social interaction: implications for mental disorders. *Curr Opin Psychiatry*. 2007;20:158.

Hervieu GJ. Further insights into the neurobiology of melanin-concentrating hormone in energy and mood balances. *Expert Opin Ther Targets*. 2006;10:211.

Hofmann HA. Gonadotropin-releasing hormone signaling in behavioral plasticity. *Curr Opin Neurobiol*. 2006;16:343.

Holsboer F, Ising M. Central CRH system in depression and anxiety—evidence from clinical studies with CRH1 receptor antagonists. *Eur J Pharmacol*. 2008;583(2–3):350–357.

Karl T, Herzog H. Behavioral profiling of NPY in aggression and neuropsychiatric diseases. *Peptides*. 2007;28:326.

Kehne JH. The CRF1 receptor, a novel target for the treatment of depression, anxiety, and stress-related disorders. *CNS Neurol Disord Drug Targets*. 2007;6:163.

Korol DL, Pisani SL. Estrogens and cognition: friends or foes? An evaluation of the opposing effects of estrogens on learning and memory. *Horm Behav*. 2015;74:105–115. Review.

Lifschytz T, Segman R, Shalom G, Lerer B, Gur E. Basic mechanisms of augmentation of antidepressant effects with thyroid hormone. *Curr Drug Targets*. 2006;7:203.

McEwen BS, Bowles NP, Gray JD, et al. Mechanisms of stress in the brain. *Nat Neurosci*. 2015;18(10):1353–1363. Review.

Miller GE, Chen E, Zhou ES. If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans. *Psychol Bull*. 2007;133:25.

Nishino S. The hypothalamic peptidergic system, hypocretin/orexin and vigilance control. *Neuropeptides*. 2007;41:117.

Olza-Fernández I, Marín Gabriel MA, Gil-Sanchez A, Garcia-Segura LM, Arevalo MA. Neuroendocrinology of childbirth and mother-child attachment: the basis of an etiopathogenic model of perinatal neurobiological disorders. *Front Neuroendocrinol*. 2014;35(4):459–472. Review.

Overli O, Sorensen C, Pulman KG, Pottinger TG, Korzan W. Evolutionary background for stress-coping styles: relationships between physiological, behavioral, and cognitive traits in non-mammalian vertebrates. *Neurosci Biobehav Rev*. 2007;31:396.

*Phillips DI. Programming of the stress response: a fundamental mechanism underlying the long-term effects of the fetal environment? *J Intern Med*. 2007;261:453.

Rosell DR, Siever LJ. The neurobiology of aggression and violence. *CNS Spectr*. 2015;20(3):254–279. Review.

Schatzberg AF, Lindley S. Glucocorticoid antagonists in neuropsychotic disorders. *Eur J Pharmacol*. 2008;583(2–3):358–364.

Schneider JE. Metabolic and hormonal control of the desire for food and sex: implications for obesity and eating disorders. *Horm Behav*. 2006;50:562.

Shamlian NT, Cole MG. Androgen treatment of depressive symptoms in older men: a systematic review of feasibility and effectiveness. *Can J Psychiatry*. 2006;51:295.

Slattery DA, Neumann ID. No stress please! Mechanisms of stress hyporesponsiveness of the maternal brain. *J Physiol*. 2008;586(2):377–385.

Sodersten P, Bergh C, Zandian M. Psychoneuroendocrinology of anorexia nervosa. *Psychoneuroendocrinology*. 2006;31:1149.

*Strous RD, Maayan R, Weizman A. The relevance of Neurosteroids to clinical psychiatry: from the laboratory to the bedside. *Eur Neuropsychopharmacol*. 2006;16:155.

Wrenn CC, Holmes A. The role of galanin in modulating stress-related neural pathways. *Drug News Perspect*. 2006;19:461.

▲ 1.13 Immune System and Central Nervous System Interactions

CHARLES L. RAISON, M.D., REBECCA G. REED, PH.D., GRAHAM A.W. ROOK, B.CHIR., M.D., AND ANDREW H. MILLER, M.D.

Interactions between the immune system and the CNS play a critical role in the maintenance of bodily homeostasis and the development of diseases, including psychiatric disease. Alterations in CNS function brought about by a variety of stressors have been shown to influence both the immune system and diseases that involve the immune system. Moreover, many of the relevant hormonal and neurotransmitter pathways that mediate these effects have been elucidated. With this has come an increased understanding of the central role that cytokines, which derive from immune cells and microglia, play in brain–immune interactions. Indeed, interest in the role of cytokines and their signaling pathways in the various psychiatric diseases has exploded in recent years and with this has come an expansion of research into the role of infectious and autoimmune diseases

in the pathophysiology of psychiatric disorders. Most recently, a previously unrecognized lymphatic system that drains the brain—and directly connects the CNS with peripheral immunity—highlights the fact that the central CNS and immunity are really best understood as different facets of a larger system evolved to enhance survival and reproduction. Taken together, these findings highlight the importance of interdisciplinary efforts involving the neurosciences and immunology for gaining new insights into the etiology of psychiatric syndromes.

OVERVIEW OF IMMUNE SYSTEM

The immune system has the capacity to protect the body from the invasion of foreign pathogens, such as viruses, bacteria, fungi, and parasites. In addition, the immune system can detect and eliminate cells that have become neoplastically transformed. Finally, it has emerged recently that the immune system surveys, controls, and maintains the various symbiotic and commensal microbiota, notably in the gut. These functions are accomplished through highly specific receptors on immune cells for molecules derived from invading organisms and a rich intercellular communication network that involves direct cell-to-cell interactions and signaling between cells of the immune system by soluble factors called *cytokines*. The body's absolute dependence on the efficient functioning of the immune system is illustrated by the less than 1-year survival rate of untreated infants born with severe combined immunodeficiency disease and the devastating opportunistic infections and cancers that arise during untreated AIDS.

Cells and Tissues

The immune system must be able to survey all tissues of the body for the presence of infectious agents or neoplastic cells and to mobilize its effector components to specific sites in the body where infectious agents may invade. Therefore, an important requirement of the immune system is that it be systemic and mobile. Cells of hematopoietic origin largely accomplish this function. Like all other blood cells, immune cells are derived from hematopoietic precursor stem cells, which in the adult originate in the bone marrow. The stem cells are pluripotent and capable of differentiating into any one of the various mature hematopoietic cells. There are two major paths of immune cell differentiation that are regulated in part by cytokines and other factors (Fig. 1.13–1). The lymphoid pathway leads to the formation of the mature lymphocytes, B cells, T cells, and natural killer (NK) cells, and the myeloid path of differentiation leads to other cells that participate in the immune response, including monocytes and granulocytes, which include neutrophils, eosinophils, and basophils. Monocytes and basophils may further differentiate into macrophages and

mast cells, respectively, which take up residence in tissues throughout the body.

Lymphocyte maturation occurs in *primary immune tissues*. In humans, the bone marrow serves as the primary site for B-cell maturation, and the thymus is the primary site for T-cell maturation. An important part of the maturation process is the screening out of cells that are reactive to the body's own constituents (*self-reactive*). After maturation, lymphocytes exit the primary immune tissues and circulate through the bloodstream and the lymphatic system into and out of the secondary immune tissues, including the spleen and widely distributed lymph nodes. *Secondary immune tissues* provide a venue for interactions between different immune cells and circulating pathogens.

Innate and Adaptive Immunity

The immune system is often divided on a functional basis into two separate categories: innate or natural immunity and adaptive or acquired immunity (Table 1.13–1). The components of innate immunity act rapidly and in a relatively nonspecific manner against pathogens or infected cells and are evolutionarily more primitive than the specialized T and B lymphocytes that mediate acquired immunity. Operationally, however, the two modes of immunity interact and cooperate.

Innate Immunity

The cells mediating innate immunity are maintained at a constant background level of activation, partly by exposure to microbial components taken in from the gut, so they do not require prior activation and/or specific recognition of invading pathogens to be functional. They provide an important first line of defense against infectious agents during the early stages of an immune response. Mononuclear phagocytic cells and NK cells are examples of immune cells that mediate innate immunity. Mononuclear phagocytic cells, such as macrophages, microglia (the macrophage equivalent in the brain), dendritic cells, reticular cells, and certain endothelial cells of lymphoid organs are all part of the *reticuloendothelial system,* which surveys circulating antigen and mobilizes an immune response upon its discovery. These cells recognize extracellular pathogens (e.g., bacteria and parasites) through relatively crude pattern recognition molecules (e.g., toll-like receptors [TLRs]) and in some cases destroy these pathogens by engulfing and degrading them (Fig. 1.13–2). In addition to recognizing pathogen-associated molecular patterns (PAMPs), TLRs and other innate immune receptors can be activated by noninfectious molecules, including heat-shock proteins, high-mobility group box 1, ATP and DNA, that collectively have come to be known as damage-associated (or danger-associated) molecular pattern molecules (DAMPs). Recent

studies suggest that DAMPs may be especially relevant for the ability of psychological stress to activate inflammation in the absence of an infectious stimulus.

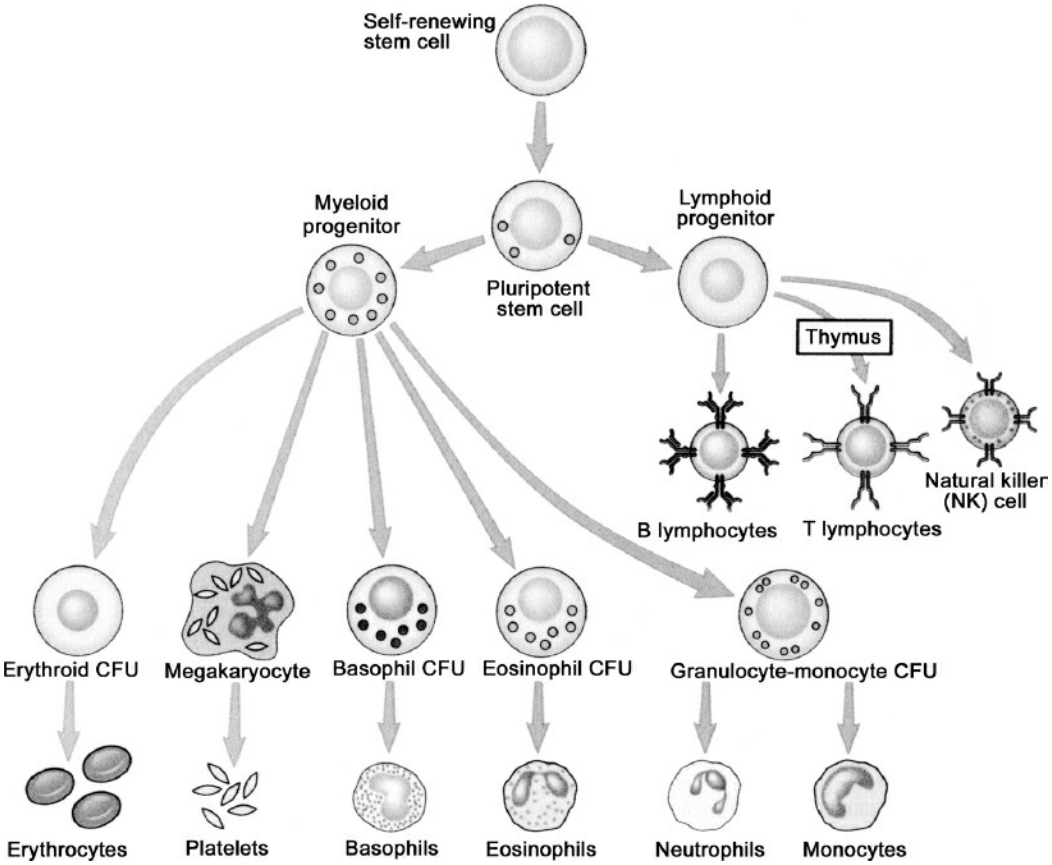

FIGURE 1.13–1. Hematopoietic tree. The development of different lineages of blood cells is depicted in this hematopoietic tree. CFU, colony forming unit. (From Abbas AK, Lichtman AH, Pober JS. *Cellular and Molecular Immunology*. Philadelphia, PA: WB Saunders; 2000, with permission.)

Table 1.13–1.
Divisions of the Immune System: Innate versus Acquired

	Innate	Acquired
Physicochemical barriers	Skin, mucous membranes	Cutaneous and mucosal immune systems
Cells	Phagocytes (macrophages, neutrophils, and natural killer cells)	Lymphocytes (B and T cells)
Soluble mediators that affect other cells	Macrophage-derived cytokines, i.e., IL-1, IL-6, TNF-α, IFN-α	Lymphocyte-derived cytokines, i.e., IL-2, IL-4, IL-5, IL-6, IL-10, IFN-γ
Memory	None	Yes
Circulating molecules	Complement, acute-phase reactants	Antibodies

IL, interleukin; IFN, interferon; TNF, tumor necrosis factor.

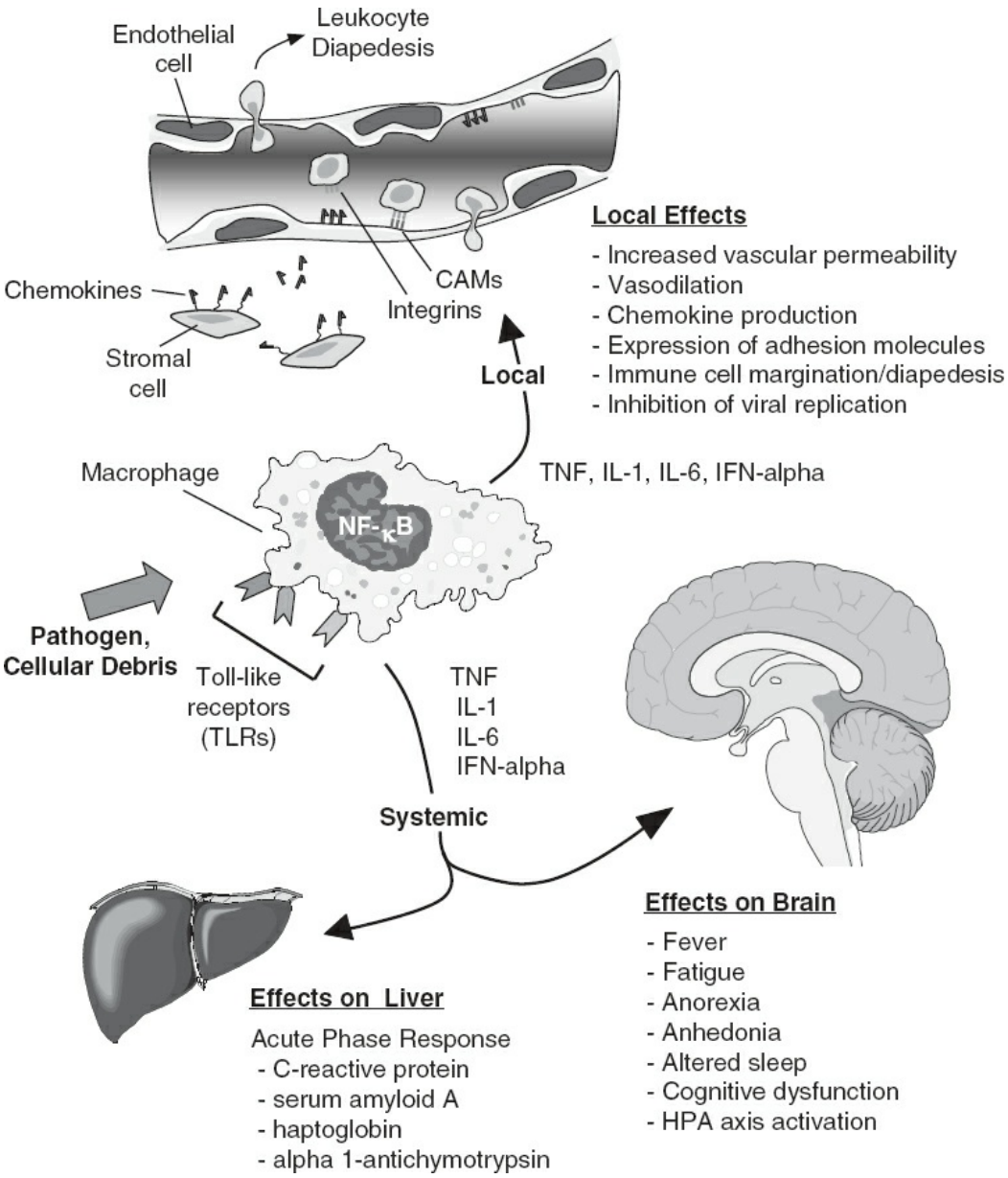

FIGURE 1.13–2. Innate immunity. Local and systemic responses to cytokine release secondary to tissue injury/infection. Locally cytokines act on endothelial and tissue stromal cells. Tissue stromal cells produce chemotactic factors, recruiting other immune cells to the site of injury. The endothelial cells produce adhesion molecules, enhancing immune cell margination and diapedesis. In the brain, proinflammatory cytokines—including interleukin (IL)-1, IL-6, and tumor necrosis factor α (TNF-α)—activate the hypothalamic–pituitary–adrenal (HPA) axis and induce behavioral changes that subserve the metabolic demands of fever and inflammation. The proinflammatory cytokines also induce the liver to produce acute phase proteins. CAM, cellular adhesion molecule; IFN, interferon; NF-κB, nuclear factor κB; PGE, prostaglandin E; TLR, toll-like receptor. (From Cowles MK, Miller AH. Stress, cytokines and depressive illness. In: Squire LR, ed. *The New Encyclopedia of Neuroscience*. Oxford: Academic Press; 2008.)

Table 1.13–2. Cytokine Mediators of Innate Immunity

Cytokine	Source	Principal Cell Targets and Biological Effects
Tumor	Macrophages, T cells	Endothelial cells: activation (inflammation, coagulation)

necrosis factor (TNF)		
		Neutrophils: activation
		Hypothalamus: fever
		Liver: synthesis of acute-phase proteins
		Muscle, fat: catabolism (cachexia)
		Many cell types: apoptosis
Interleukin-1 (IL-1)	Macrophages, endothelial cells, some epithelial cells	Endothelial cells: activation (inflammation, coagulation)
		Hypothalamus: fever
		Liver: synthesis of acute-phase proteins
Chemokines	Macrophages, endothelial cells, T cells, fibroblasts, platelets	Leukocytes: chemotaxis, activation; migration into tissues
Interleukin-12 (IL-12)	Macrophages, dendritic cells	T cells: Th1 differentiation
		NK cells and T cells: IFN-γ synthesis, increased cytolytic activity
Type 1 IFNs (IFN-α, IFN-β)	IFN-α: macrophages	All cells: antiviral state, increased class I major histocompatibility complex (MHC) expression
	IFN-β: fibroblasts	NK cells: activation
Interleukin-10 (IL-10)	Macrophages, T cells (mainly Th2)	Macrophages, dendritic cells: inhibition of Th1 cell production of IFN-γ and IL-2 and expression of costimulators and class II MHC molecules
		NF-κB and TNF-α: inhibition
Interleukin-6 (IL-6)	Macrophages, endothelial cells, T cells	Liver: synthesis of acute-phase proteins
		B cells: proliferation of antibody-producing cells
Interleukin-15 (IL-15)	Macrophages, others	NK cells: proliferation
		T cells: proliferation (memory $CD8^+$ cells)
Interleukin-18 (IL-18)	Macrophages	NK cells and T cells: IFN-γ synthesis

From Abbas AK, Lichtman AH, Pober JS. *Cellular and Molecular Immunology*. Philadelphia, PA: WB Saunders; 2007, with permission.

Important intracellular signaling molecules that are triggered by TLRs (as well as a number of cytokines) include NF-κB and the MAPKs including p38 MAPK. NF-κB and MAPK play a key role in initiating the innate immune inflammatory response. Activated mononuclear phagocytes release type I interferons (e.g., IFN-α), which have direct antiviral properties, and proinflammatory cytokines, including TNF, interleukin (IL)-1, and IL-6. TNF is the principal mediator of the response to gram-negative bacteria and is one of the earliest cytokines released in the proinflammatory cascade that includes IL-1 followed by IL-6. TNF and IL-1 are endogenous pyrogens capable of inducing fever by increasing the

synthesis of prostaglandins by cytokine-stimulated hypothalamic cells. TNF also leads to cachexia, characterized by wasting of muscle cells, in part secondary to appetite suppression.

The combination of liver-derived plasma proteins induced by TNF and IL-1 with those induced by IL-6 constitutes the acute-phase response (Fig. 1.13–2). The acute-phase response is designed to limit tissue damage, isolate, and destroy invading pathogens, and set repair functions in motion. These objectives are achieved by rapid changes in plasma protein composition characterized by increases in *acute-phase reactants,* including C-reactive protein (CRP) (which coats bacteria to facilitate phagocytosis-opsonization), macroglobulin, and other antiproteases (which neutralize tissue destructive proteases), and the clotting protein fibrinogen. Albumin and transferrin (the iron transport protein) decline during the acute-phase response and are therefore called *negative acute-phase reactants*. During the acute-phase response, inflammatory cytokines also coordinate the systemic response to infection, having potent effects on the neuroendocrine system (especially the HPA axis) and the CNS where they mediate many symptoms of illness, including fever, loss of appetite, social withdrawal, and sleep changes. Innate immune responses can also be triggered in response to cellular constituents that are released as a consequence of cell death or destruction (as may occur during trauma or ischemia).

Complement factor proteins, which are produced by the liver, represent another important humoral component of the innate immune response. These functionally linked proteins interact with one another in a highly regulated manner and subserve many of the effector functions of the immune system, including cell lysis, opsonization, activation of inflammation by attracting inflammatory cells (chemotaxis), stimulation of immune cells to release chemical mediators of inflammation, and neutralization of antigen–antibody complexes that can damage tissues.

NK cells are also an important component of innate immunity. These cells can destroy virally infected cells by binding to them and releasing cytolytic factors, including perforin. NK cells also have the ability to recognize and destroy neoplastically transformed host cells, especially those of hematopoietic origin, thus providing protection against some cancers.

Adaptive Immunity

T and B lymphocytes are the crowning achievement of the evolution of immune cell specialization. These cells account for the diversity, specificity, and adaptive functionality within the immune system. An effective acquired immune response includes a recognition phase, an activation phase, and an effector phase of antigen elimination.

During the *recognition* phase, the presence of an infectious agent, antigen, or neoplastically transformed cell is detected. This is achieved

through specialized receptors for antigens on the surfaces of B and T lymphocytes. The receptors for an antigen on a particular B or T cell are identical and unique to that cell and its descendants (*clones*). A family of lymphocytes with identical antigen receptor specificity is called a *clonal line*. Diversity in antigen recognition is derived from the vast number of different B- and T-cell clonal lines present in each individual. When an antigen is detected, the corresponding clone is selected and activated.

The *activation* phase includes the proliferation and mobilization of the immune cells relevant to the eradication of the infectious agent. The binding of foreign antigens by B and T cells is usually not sufficient to produce cell activation; an accessory signal must also be provided. Important accessory signals are generated by a group of cytokines called *interleukins* that are secreted by T helper (Th) cells and antigen-presenting cells (APCs), such as macrophages. Th cells and APCs cooperate; APCs secrete IL-1 and other cytokines that stimulate Th cells to secrete a host of cytokines including interferon γ (IFN-γ), which then increases the phagocytic ability of APCs, ultimately improving their antigen-presenting capacity. Cytokines involved in the acquired immune response often simultaneously serve multiple functions. For example, in addition to the effects noted above, IFN-γ also has direct antiviral properties, and IL-1 stimulates the expression of IL-2, which in turn activates multiple lymphocyte functions (Fig. 1.13–3).

Cytokines can be broadly divided into categories based on their role in the initiation, regulation, and maintenance of the immune response. Hence, there are cytokines that mediate innate immunity and inflammation (Table 1.13–2), cytokines that regulate acquired immunity (Table 1.13–3), and cytokines that control proliferation and differentiation of immature immune cells (Table 1.13–4). Although structurally distinct, these cytokines overlap in function and act together to govern the dynamic events of immunity. Within the acquired immune response, significant interest has been paid to the concept that there are two varieties of Th cells (designated as clusters of differentiation [CD] 4^+) known as Th1 and Th2. Th1 cells produce cytokines such as IL-2 and IFN-γ that promote T cell and inflammatory responses that are especially relevant for protection against intracellular pathogens. Th2 cells produce cytokines such as IL-4 that promote antibody production and provide protection against parasites. When excessive, Th2 responses also have been closely associated with allergic and hypersensitivity reactions as well as asthma. A third subset of Th cells that are developmentally distinct from Th1 or Th2 cells and that produce IL-17 (Th17 cells) has been identified as an important source of inflammation central to the development of many autoimmune and allergic conditions that were once thought to be more closely involved with Th1 or Th2 dysregulation. More recently additional subsets of effector cells have been found, notably Th9 cells, which secrete IL-9 and seem to be

predominantly proinflammatory, and Th22 cells, which secrete IL-22 and can have either proinflammatory or anti-inflammatory effects. Finally, CD4 cells can differentiate into T regulatory cells (Tregs) that dampen immune activation both via direct cell-to-cell effects and via production of anti-inflammatory cytokines, especially IL-10. This crucial function is discussed later.

After binding antigen in the presence of stimulatory cytokines, T and B lymphocytes with the appropriate binding sites are activated, leading to cell growth, division, and proliferation. Activation also results in the clonal expansion of immune cells with the identical high-affinity specificity for the foreign antigen. Some of the progeny during clonal expansion undergo further differentiation into mature effector cells, such as antibody-secreting plasma B cells and cytotoxic $CD8^+$ T-lymphocytes (CTLs). In contrast, some descendants of activated B or T cells become memory cells that are primed for activation on future stimulation by the same antigen. Reexposure to that antigen results in a secondary immune response (thus the name *acquired* immunity), which is typically more rapid and robust than the first or primary immune response to that antigen. Memory cells may live for many years providing long-lasting acquired immunity, as demonstrated by individuals who have received a vaccine or had their first contact with a specific infectious agent during infancy.

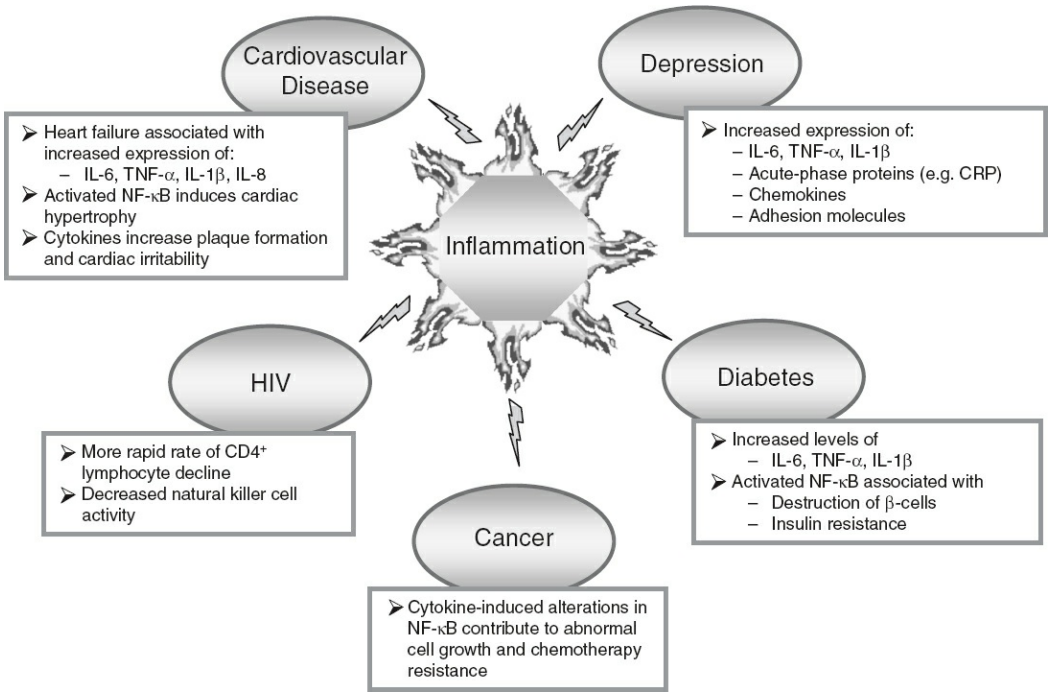

FIGURE 1.13–3. Inflammation and disease. IL, interleukin; TNF, tumor necrosis factor; NF-κ*B,* nuclear factor kappa beta; CRP, C-reactive protein. (From Cowles MK, Miller AH. Stress, cytokines and depressive illness. In: Squire LR, ed. *The New Encyclopedia of Neuroscience*. Oxford: Academic Press; 2008.)

Table 1.13–3.

Cytokine Mediators of Acquired Immunity

Cytokines	Source	Target	Primary Effect
Interleukin-2 (IL-2)	T cells	T cell	Growth, cytokine production
		NK cell	Growth, activation
		B cell	Growth, antibody production
Interleukin-4 (IL-4)	$CD4^+$ T cells	T cell	Growth, differentiation
		B cell	Isotype switching to IgE
Interleukin-5 (IL-5)	T cell	Eosinophil	Activation
		B cell	Growth, IgA production
Transforming growth factor-β (TGF-β)	T cells, others	T cell	Inhibit growth and activation
		B cell	Inhibit growth
		Macrophage	Inhibit activation

Adapted from Abbas AK, Lichtman AH, Pober JS. *Cellular and Molecular Immunology*. Philadelphia, PA: WB Saunders; 2000, with permission.

During the *effector* phase, the pathogen is neutralized and eliminated. The principal effector mechanisms of acquired immunity are mediated by antibodies (*humoral immunity*) secreted from B cells, and by CTLs that kill target cells directly and macrophage-activating cytokines such as IFN-γ that enhance antimicrobial functions of macrophages including microglia (*cellular immunity*). Humoral immunity is especially effective in combating extracellular pathogens, such as bacteria and parasites, while cellular immunity is effective in protecting against viral and intracellular bacterial infection and, as with NK cells, may provide some protection against tumor cells. Regulation of the acquired immune response is another important component of the effector phase and includes $CD4^+CD25^+$ regulatory T cells, which are immunosuppressive in nature and serve to restrict immune responses to self-antigens. In addition, the inhibitory costimulatory molecule called programmed death-1 (PD-1) and its ligands have been shown to play an important role in regulating T-cell activation. Indeed, loss of PD-1 has been associated with an autoimmune diathesis in laboratory animals.

The effector components of innate immunity are also recruited, enhanced, and directed toward specific pathogens as a result of the actions of B and T cells. For example, circulating antibodies can neutralize pathogens by binding to and coating the pathogens (*opsonization*). Pathogens that are opsonized are made susceptible to lysis by complement factors and phagocytosis. NK cells and *phagocytic cells,* such as neutrophils and macrophages, have receptors for the Fc fragment (fragment crystallizable region) of antibodies. Furthermore, complement proteins bind to and are activated by the Fc fragments of some types of antibodies. Thus, antibodies can link effector cells and cytolytic proteins of innate immunity with pathogens, lending a level of specificity that is not

inherent in the effector innate immune processes themselves.

Regulation of the Immune Response

An effective immune response requires the cooperation of many components of the immune system, often resulting in the augmentation of each component's contribution to the overall immune response. However, the simultaneous indiscriminate amplification of all aspects of the immune system would not be efficient and could even be disastrous. An overactive immune system may contribute to autoimmunity. Furthermore, the inflammatory component of immune responses can be damaging if not controlled, as is seen in immune complex diseases and septic shock. Therefore, regulation of the immune response is necessary to make sure that the response is energy efficient, focused on the infectious agent, counterbalanced in a fashion that does not cause self-damage, and reversible once the pathogen has been eliminated. The immune system is potentially dangerous to the host, and needs to be controlled by efficient negative feedback. A pathological failure to shut off inappropriate inflammation is now implicated in several psychiatric disorders. Interestingly, several organisms found within the gut microbiota are known to have major impacts on Treg development, so the immune system both regulates, and is regulated by the microbiota.

Probably the most important form of *intrinsic regulation* of the immune system is mediated by the various cytokines. Several examples of the facilitatory effects of cytokines have been cited. Conversely, cytokines such as transforming growth factor beta (TGF-β) and IL-10 potently inhibit lymphocyte activation and proliferation and antagonize the activity of proinflammatory cytokines (Tables 1.13–2 and 1.13–3). In addition, as noted above, regulatory T cells play a pivotal role in suppressing the function of other immune cell types.

Table 1.13–4.
Cytokine Mediators of Immune Cell Growth and Differentiation

Cytokines	Source	Target	Primary Effect
Interleukin-3 (IL-3)	T cell	Immature progenitor	Growth and differentiation to many cell lines
Granulocyte–monocyte colony-stimulating factor (G–M CSF)	T cell, monocyte, macrophage, others	Immature and committed progenitors, macrophages	Growth and differentiation to all cell lines
Macrophage CSF	Macrophage, others	Committed progenitor	Differentiation to monocyte, macrophage
Granulocyte CSF	Monocyte, macrophage, others	Committed progenitor	Differentiation to neutrophil, eosinophil, basophil
Interleukin-7	Fibroblast,	Immature progenitor	Growth of T- and B-cell lines

	bone marrow stromal cells		
Leukemia inhibitory factor (LIF)	Fibroblast, bone marrow stromal cells	Immature progenitors, others	Governs growth and differentiation of hematopoietic and monocyte cell lines

Adapted from Abbas AK, Lichtman AH, Pober JS. *Cellular and Molecular Immunology*. Philadelphia, PA: WB Saunders; 2000, with permission.

Table 1.13–5.
Foundations of Nervous, Endocrine, and Immune System Interactions

1. Expression of receptors for neurotransmitters, hormones, and neuropeptides on immune cells
2. Autonomic nervous system innervation of lymphoid tissues
3. Conditioning of the immune response
4. Stress effects on immune function
5. Expression of cytokines and their receptors in the CNS
6. Influence of the immune system on neurotransmitter turnover, neuroendocrine function synaptic plasticity, regional brain activity, and behavior

The relative significance of *extrinsic regulation* of the immune response remains to be fully established. However, increasing evidence of neural-immune interactions indicates that extrinsic factors of CNS origin play an important role in the modulation of the immune system. These data provide the foundation for nervous, endocrine, and immune system interactions, many of which are relevant to psychiatry (Table 1.13–5).

EVIDENCE OF NERVOUS SYSTEM AND IMMUNE SYSTEM INTERACTIONS

Immune Cell Receptors

As outlined in Table 1.13–6, cells from the immune system express receptors for a wide variety of molecules that are, in part, regulated by or derived from the nervous system. One of the first receptors to be characterized in lymphocytes was the β-adrenergic receptor, which is the predominant adrenergic receptor subtype expressed on T and B cells. Subsequently, receptors for the other small molecule neurotransmitters have been described. As in the nervous system, receptors for neurotransmitters on immune cells are located in the cell membrane and in most cases are coupled to G proteins and their associated second-messenger pathways.

Several important concepts from research on receptors in immune cells and tissues are central to understanding the effects of neutrally derived molecules on immune function. First, the expression of receptors is heterogeneous. For example, of the two types of receptors for adrenal steroids, mineralocorticoid receptors and glucocorticoid receptors (GRs), only GRs are expressed in the thymus, whereas both glucocorticoid and mineralocorticoid receptors are expressed in the spleen. Related to heterogeneity in receptor expression in immune cells and tissues is heterogeneity in receptor density. Heterogeneity of receptor expression/density is relevant for determining the net effect of circulating transmitters on immune function. For example, the β2 adrenergic receptor is expressed on resting and activated B cells, naive $CD4^+$ T cells, newly generated Th1 cells, and Th1-cell clones. However, it is not expressed on newly generated Th2 cells or Th2 cell clones. Consistent with these findings, norepinephrine (NE) has been found to enhance IL-12–induced differentiation of naive $CD4^+$ T cells into Th1 cells and to promote production of IFN-γ by these cells. No effect was found on IL-4–induced Th2 cell differentiation. The effect of NE on Th1-type responses is also manifested by the ability of NE to help Th1 cells support B-cell antibody production.

Table 1.13–6.
Receptors for Neurotransmitters, Hormones, and Peptides on Immune Cells

Neurotransmitters	Hormones	Peptides
Acetylcholine	Corticosteroids–glucocorticoids, mineralocorticoids	ACTH
Dopamine	Gonadal steroids–estrogen, progesterone, testosterone	α-MSH
Histamine	Growth hormone	AVP
Norepinephrine	Prolactin	Calcitonin
Serotonin	Opioids (endorphins, enkephalins)	CGRP

Thyroid hormone	CRH GHRH GnRH IGF-1 Melatonin NPY PTH Somatostatin Substance P TRH TSH VIP

ACTH, adrenocorticotropin; α-MSH, α-melanocyte-stimulating hormone; AVP, arginine vasopressin; CGRP, calcitonin gene–related peptide; CRH, corticotrophin-releasing hormone; GHRH, growth hormone–releasing hormone; GnRH, gonadotropin-releasing hormone; IGF-1, insulin-like growth factor-1; NPY, neuropeptide Y; PTH, parathyroid hormone; TRH, thyrotropin-releasing hormone; TSH, thyroid-stimulating hormone; VIP, vasoactive intestinal peptide.

Another important concept is that a change in the circulating concentration of a hormone or transmitter is not necessarily reflected equally in all immune compartments. For example, stress-related increases in glucocorticoids are more effective in activating GRs in the peripheral blood and the thymus than in the spleen. Thus, the microenvironment of any given tissue is critical in determining hormonal or neurotransmitter influences on immune function. Taken together with the heterogeneity in receptor expression and density, the data demonstrate that the influence of any given molecule on the immune system is a function of (1) the type of cell that exhibits the relevant receptor, (2) the density of the receptors on that cell, and (3) whether that cell is located in an immune compartment that allows access of the relevant molecule to the receptor under the conditions being studied.

Cross talk between receptor-associated second-messenger pathways is another important mechanism by which neutrally derived or regulated molecules can influence the immune response and vice versa. For example, activation of cytokine signaling pathways including p38 MAPK by IL-2 and IL-4 (as well as IL-1, TNF, IL-6 and IL-8) has been shown to downregulate GR expression or function and may account for the glucocorticoid resistance seen in some inflammatory disorders (such as asthma) as well as major depression (see below).

In Vitro and In Vivo Effects

Numerous chemical messengers derived from or regulated by the nervous system are capable of altering immune cell function and distribution. Table 1.13–7 provides a necessarily simplified, representative listing of selected neurotransmitters and neuropeptides and their immune effects. Table 1.13–8 lists some of the immune activities of hormone messengers. The

immunological effects of these agents depend on a number of factors aside from those involving the relevant expression, density, and activation of receptors on target immune cells. For example, the effect of any given molecule on the immune system depends on the phase of the immune response (recognition, activation, or effector) that is involved. NE, for example, has been found to promote immune function during the recognition phase, both potentiate and inhibit immune function during the activation and proliferation phase, and inhibit the effector phase. The potentiation of the activation phase occurs at low concentrations of NE, but inhibition occurs at high NE concentrations. These findings indicate that both the timing of exposure and the dose are important.

Table 1.13–7.
Immunological Effects of Representative Neurotransmitters and Neuropeptides

Chemical Messengers	Immunological Activity	
	In Vitro	**In Vivo**
Neurotransmitters		
Norepinephrine	Stimulation of T-cell proliferation at low concentrations, high concentrations are inhibitory; enhancement of IL-12–induced differentiation of naïve CD4 cells into Th1 cells; inhibition of Th1-type cytokines (IFN-γ) and stimulation of Th2- type cytokines (IL-10) in PBMC; activation of NF-κB and proinflammatory cytokines	Influences immune cell trafficking; redistribution of NK cells from spleen to blood; inhibition of NK cell activity and cytolytic T-cell activity; inhibits generation of antigen-specific T cells in draining lymph nodes but increases inflammation in joints during autoimmune arthritis
Serotonin	At suprapharmacological concentrations: suppression of lymphocyte reactivity to mitogens and antigens At physiogical concentrations: inhibition of monocyte-induced suppression of NK cell activity; promotes capacity of macrophages to enhance T-cell activation; enhancement of macrophage superoxide production and IFN-γ–induced phagocytosis; stimulation of chemotactic factors; contributes to DTH	Suppression of humoral and cellular immune responses; enhancement of immune activity when serotonin availability is decreased
Neuropeptides		
Opioids	Enhancement of T-cell proliferation, NK activity, cytokine production, and generation of cytotoxic T cells	Mediation of immunosuppressive effects of stress on NK activity; inhibition of mitogen-induced lymphocyte proliferation and phagocytic cell function; inhibition of antibody production; diminished DTH;

		promotion of splenic immune cell apoptosis
Substance P	Enhancement of lymphocyte proliferation, lymphocyte and monocyte chemotaxis, and monocyte production of IL-1, IL-6, and TNF; augmentation of IgA synthesis; induction of mast cell degranulation; promotion of superoxide anion release from neutrophils and eosinophils; increases infectivity of HIV	Increased severity of adjuvant-induced arthritis; associated with hypersensitivity reactions and chronic inflammatory disorders
ACTH	Suppression of antibody production and disruption of macrophage-mediated tumoricidal activity	Activation of immunoregulatory glucocorticoids
CRH	Stimulation of T- and B-cell proliferation; enhances IL-1 and IL-6 secretion but inhibits IFN-γ secretion	Exerts mixed proinflammatory (e.g., in the periphery) and immunosuppressive (central) actions; promotes IL-1 production in CNS and increases IL-2; activates the HPA axis; suppresses NK cell activity in the spleen; inhibits antibody production and decreases T-cell numbers; inhibits mitogen-induced T-cell proliferation
VIP	Enhancement of monocyte chemotaxis; inhibition of Ig and IL-2 production; inhibition of NK activity and one-way MLR	Inhibition of egress of lymphocytes from sheep lymph nodes
α-MSH	Inhibits proinflammatory cytokine release from PBMCs	Antipyretic and anti-inflammatory

ACTH, adrenocorticotropic hormone; CRH, corticotrophin-releasing hormone; DTH, delayed-type hypersensitivity; HPA, hypothalamic-pituitary-adrenal; IFN, interferon; Ig, immunoglobulin; IL, interleukin; MLR, mixed lymphocyte reaction; MSH, melanocyte-stimulating hormone; NF-κB, nuclear factor kappa B; NK, natural killer; PBMC, peripheral blood mononuclear cell; TNF, tumor necrosis factor; VIP, vasoactive intestinal peptide.

Issues of timing are also relevant in terms of development and aging. In aged rats, for example, a progressive loss of noradrenergic innervation of the spleen is accompanied by a progressive increase in the density of β-receptors on splenic lymphocytes. However, there is also an age-related dysfunction that involves impaired coupling between the β-receptor and adenylate cyclase, indicating that noradrenergic agents may have variable, unpredictable effects on immune function in old animals.

Related to the phase of the immune response and developmental stage of the animal is the type of immune response as it relates to pathophysiology. Substances that are primarily inhibitory to immune function may promote tumor development in animals with cancer but may attenuate the development of autoimmune disease. For example,

glucocorticoids accelerate the growth of tumors in mice, whereas they inhibit the development of several types of autoimmune disorders, including experimental allergic encephalitis (a model of multiple sclerosis) and streptococcal cell-wall–induced polyarthritis (a model of rheumatoid arthritis). Relevant to the effects of gonadal steroids on immune function, estrogens tend to promote Th1-type responses, while progesterone tends to promote Th2-type responses. Accordingly, during pregnancy, Th2-type immune responses prevail (possibly secondary to the increased influence of progesterone), and autoimmune disorders related to excessive Th1-like activity (multiple sclerosis and rheumatoid arthritis) may improve. In contrast, diseases related to Th2-like activity (e.g., systemic lupus erythematosus) may be exacerbated during pregnancy. Normal pregnancy is also associated with an increase in Treg activity and a decrease in Th17 activity. This shift toward an anti-inflammatory Treg/Th17 balance can fail to occur in some cases of unexplained recurrent spontaneous abortion.

Table 1.13–8.
Representative Listing of Hormone Messengers and Their Immunological Effects

Chemical Messengers	Immunological Activity	
	In Vitro	**In Vivo**
Hormones		
Adrenal steroids	Inhibition of IL-1, IL-2, and interferon; augmentation of IL-4 production; inhibition of NK activity, mitogen proliferation, and antigen presentation; promotion of T-cell differentiation toward Th2 profile and away from Th1 profile	Thymic involution; lymphopenia; monocytopenia; neutrophilia; suppression of inflammation and cell-mediated immunity during chronic exposure; enhancement of immune responsiveness during acute exposure, modulation of apoptosis, immune cell trafficking
Estrogen	Inhibition of T suppressor cell and NK cell activity; increased macrophage phagocytosis and lysosomal activity; promotion of T-cell differentiation toward Th1 profile; increased IL-6 and IL-10; inhibition of B-cell apoptosis	Lymphopenia; decreased mitogen responsiveness, NK activity and macrophage phagocytosis; increased plasma cells in spleen; promotion of autoantibodies
Progesterone	Decreased mitogen responsiveness at high concentrations; inhibition of T-cell activation and cytotoxicity; inhibition of NK cell activity; inhibition of prostaglandin synthesis; promotion of T-cell differentiation toward Th2 profile	Increased skin graft survival; increased survival of xenographic tumor cells; increases risk of viral infection; decreased $CD4^+$ cell numbers
Growth hormone	Enhancement of mitogen responsiveness and cytotoxic T-cell activity; priming of macrophages and neutrophils for superoxide anion release; augmentation of neutrophil differentiation;	Increases thymus and spleen size and cellularity; augmentation of antibody synthesis, T- and B-cell proliferation, IL-2 production, mitogen responsiveness, and NK activity; promotes survival

	enhancement of neutrophil and macrophage phagocytosis; increases synthesis of IFN-γ	during bacterial infection
Insulin-like growth factor	Prevention of promyeloid cell apoptosis; promotes priming of macrophages and neutrophils for superoxide anion release; enhancement of neutrophil and macrophage phagocytosis; inhibition of nuclear translocation of NF-κB following TNF-α exposure	Promotes hematopoiesis and lymphopoiesis; increases thymus and spleen size and cellularity; enhances overall immune responsiveness in aged animals; blocks TNF-α induction during septic shock; blocks TNF-α–induced sickness behavior
Prolactin	Removal of PRL from culture media inhibits DNA synthesis and cell proliferation; comitogenic with IL-2; inhibits T-cell apoptosis; promotion of Th2 cytokine production from T cells	Increases thymus and spleen size and cellularity; counteracts glucocorticoid-mediated immunosuppression; PRL removal reduces NK activity and T-cell proliferation and increases lethality of Listeria challenge; stimulates autoimmunity

IL, interleukin; NF-κB, nuclear factor kappa B; NK, natural killer; TNF, tumor necrosis factor; PRL, prolactin.

Another important factor in determining the immunological effect of a particular molecule is its indirect effects, as well as its direct effects, on the immune system. In vitro studies provide important information on the direct effects of the various chemical messengers, but the influence of those agents in vivo may be completely different. For example, a number of in vitro studies have shown that opioid peptides are capable of enhancing NK cell activity (NKCA). However, in vivo, opioid peptides play an important role in mediating the inhibitory effects of shock stress on NKCA, most likely through effects in the brain. In vivo, neutrally derived molecules act against a complicated background of multiple hormones that may have synergistic or antagonist effects or both. Furthermore, many of the hormones and transmitters influence other bodily systems, including the cardiovascular system, which may influence the traffic of immune cells to various organs, immunological and otherwise. Changes in immunocyte distribution may ultimately have effects on cellular function.

Neural Innervation of Lymphoid Tissue

Identification of nerve fibers derived from the SNS in immune tissues was one of the first indications that communication between the CNS and the immune system was possible. Sympathetic nerve fibers have been identified in organs that are responsible for the development, education (e.g., self–nonself discrimination), and function of lymphocytes. Specifically, nerve fibers are found in the bone marrow, thymus, spleen, and lymph nodes. The nerves that innervate the thymus gland are derived from the vagus, phrenic, and recurrent laryngeal nerves and from the stellate and other small ganglia of the thoracic sympathetic chain. The

nonmyelinated nerves that innervate the bone marrow arise from the level of the spinal cord associated with the location of the bone. The spleen obtains its sympathetic nerves from the celiac ganglion. SNS innervation of lymph nodes is not as dense or as uniquely distributed as that of the spleen and thymus.

In general, sympathetic nerve fibers enter lymphoid tissues in association with the vascular supply. Because these nerves play an important role in vascular tone, their presence in association with the smooth muscle cells of the blood vessels is not unexpected. However, the nerve fibers also travel with small blood vessels devoid of smooth muscle cells and are present in the parenchyma of the lymphoid tissue (Fig. 1.13–4). Electron microscopy has shown that sympathetic nerve terminals exist in close approximation with lymphocytes and macrophages. Thus, the sympathetic branch of the ANS is in a position to influence the immune system either by changing the vascular tone and blood flow into lymphoid organs or by directly influencing immune cell function via locally released neurotransmitters, especially NE as well as neuropeptides such as NPY, substance P, VIP, calcitonin gene–related peptide (CGRP), and corticotropin-releasing hormone (CRH), which, in turn, interact with specific receptors on nearby immune cells.

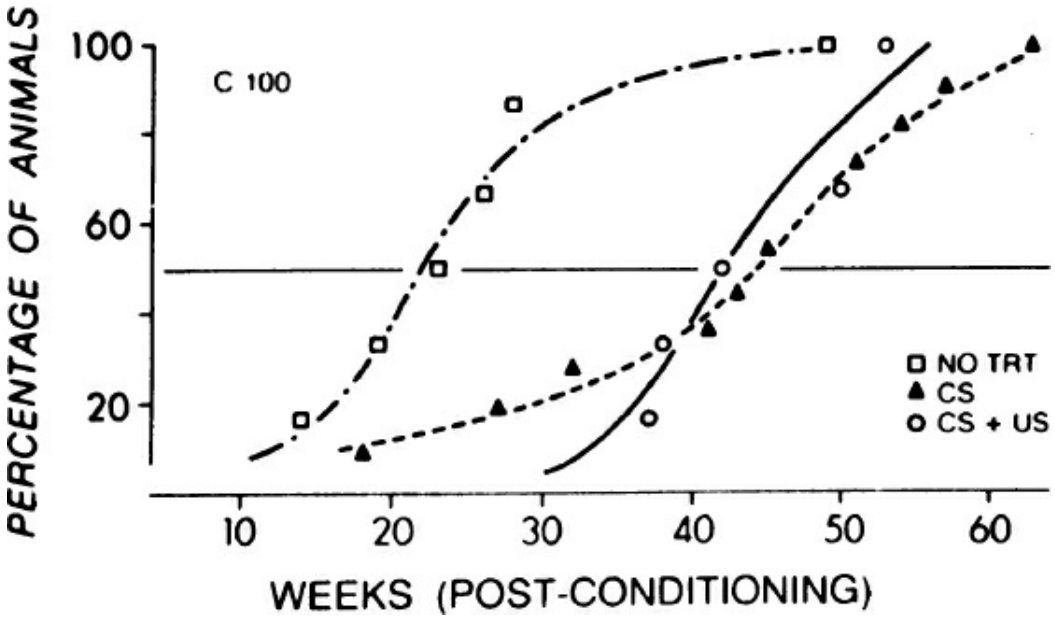

FIGURE 1.13–4. Conditioned immunosuppression. Mortality rate in first filial generation female mice (New Zealand black × New Zealand white) treated with saccharin and cyclophosphamide (CY) weekly and then continued on a regimen of saccharin and CY (group conditioned stimulus [CS] + unconditioned stimulus [US], *N* = 6), continued on saccharin alone (group CS, *N* = 11), or deprived of both saccharin and CY (no treatment [TRT], *N* = 6). (From Ader R. Behaviorally conditioned modulation of immunity. In: Guillemin R, Cohen M, Melnechuk T, eds. *Neural Modulation of Immunity*. New York: Raven Press; 1985, with permission.)

Regarding the functional implications of sympathetic nervous

innervation of immune tissues, chemical sympathectomy in laboratory animals has variable effects on immune function, depending in part on the phase of the immune response studied. The reported effects of sympathectomy include suppressed antibody responses to sheep red blood cells, suppressed cytolytic T-cell activity, and enhanced NKCA. Splenic sympathectomy also leads to an upregulation of β-adrenergic receptors on lymphocytes and a decrease in suppressor lymphocyte (T-cell) function. In multiple model systems, sympathectomy reduces peripheral inflammatory tone, an effects associated with suppression of Th1 and Th17 activity. Aside from the phase of the immune response, other factors that influence the effects of sympathetic nervous innervation on immune function include the animal's age, sex, and strain. Of note, there is also evidence that local immune responses within the microenvironment of immune tissues may be able to interact directly with sympathetic nerve fibers through the effects of cytokines on neurotransmitter release.

While SNS effects on immune function have been well-established, more recently studies have shown that the parasympathetic branch of the ANS also contributes to immune regulation. Via an efferent neural signaling pathway referred to as the *cholinergic anti-inflammatory reflex,* studies have found that stimulation of the vagus nerve attenuates immune system activation and the physiological signs of septic shock in response to lipopolysaccharide (LPS). These effects are mediated by vagal release of acetylcholine that interacts with the α7 subunit of the nicotinic AChR (α7 nAChR) on relevant immune cells, and suppresses the production of a host of cytokines, including TNF-α, via inhibition of NF-κB, as well as other inflammatory signaling molecules (Fig. 1.13–5). This influence of efferent vagal pathways on the immune response has been demonstrated in the context of a variety of inflammatory processes including myocardial ischemia, hemorrhagic shock, ischemia/reperfusion, and pancreatitis.

Finally, in addition to the immunological influences of the PNS and SNS *efferent* nerve fibers, increasing attention is now being directed to the study of sensory *afferent* fibers and their relevance to immune system communication with the brain. For example, sensory afferent fibers have been shown to relay immune signals to the brain through cytokine receptors on paraganglia of the vagus nerve. Vagal nerve fibers transmit cytokine signals to the nucleus of the solitary tract, where nervous system pathways project to other brain regions, including the hypothalamus, hippocampus, and amygdala, which are relevant to the CNS response to immune system activation (see below). Although much of the research on these "sensory" immune functions has focused on the vagus nerve, it should be noted that sensory fibers distributed throughout the body, such as skin, muscle, and all mucosal surfaces, can respond to immunological stimuli and transmit this information to the CNS.

SOCIAL ADVERSITY, ENVIRONMENTAL CHALLENGES, AND THE IMMUNE RESPONSE

As shown in Table 1.13–9, most factors that increase the risk of depression in the modern world also increase levels of peripheral inflammatory biomarkers, and many factors that enhance emotional well-being have been shown to lower inflammation. Interest in the effects of the larger environment on the immune system grew out of a series of animal and human studies suggesting that stressful stimuli can influence the development of immune-related disorders, including infectious diseases, cancer, and autoimmune disorders. While stress was initially thought to be associated with suppression of immune function, ever-increasing data indicate that such a conclusion oversimplifies the complexities of the mammalian immune response to environmental perturbation and that stress may also activate certain aspects of the immune system, particularly the innate proinflammatory immune response (Fig. 1.13–6).

Stress and Illness

Experiments conducted on laboratory animals in the late 1950s and the early 1960s indicated that a wide variety of stressors—including isolation, rotation, crowding, exposure to a predator, and electric shock—increased morbidity and mortality in response to several types of tumors and infectious diseases caused by viruses and parasites. However, as research progressed it became increasingly clear that "stress" is too variegated a concept to have singular effects on immunity and that, in fact, the effects of stress on immunity can be opposite depending on a number of factors. Chief among these factors is whether a stressor is acute or chronic. Other critical variables include stressor severity and type, as well as the timing of stressor application and the type of tumor or infectious agent investigated. For example, mice subjected to electric grid shock 1 to 3 days before the infection of Maloney murine sarcoma virus–induced tumor cells exhibited a decreased tumor size and incidence. In contrast, mice exposed to grid shock 2 days after tumor cell injection exhibited an increase in tumor size and number.

The relevance of the effects of stress on immune-related health outcomes in humans has been demonstrated in studies that have shown an association between chronic stress and increased susceptibility to the common cold, reduced antibody responses to vaccination, and delayed wound healing. In addition, stress, as well as depression, through their effects on inflammation has been linked to increased morbidity and mortality in infectious diseases, such as HIV infection, autoimmune disorders, neoplastic diseases, as well as diabetes and cardiovascular disorders, which are increasingly being recognized as diseases in which the immune system, and inflammation in particular, plays a pivotal role (Fig.

1.13–7).

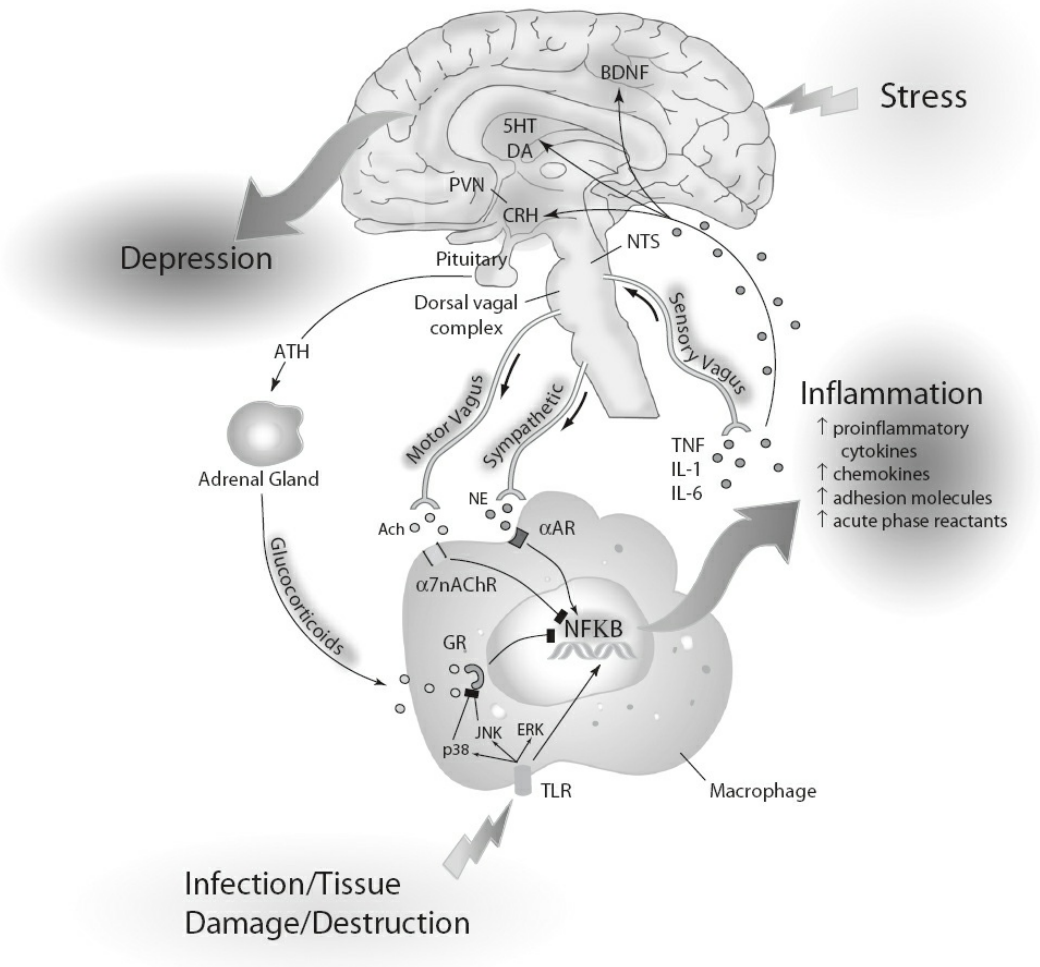

FIGURE 1.13–5. Stress and bidirectional neural–immune interactions. (**A**) Activation of NF-κB through toll-like receptors (TLRs) during immune challenge leads to an inflammatory response including (**B**) the release of proinflammatory cytokines, TNF-α, IL-1, and IL-6. **C:** These cytokines, in turn, access the brain via leaky regions in the blood–brain barrier, active transport molecules, and afferent nerve fibers (e.g., sensory vagus), which relay information through the nucleus tractus solitarius (NTS). **D:** Once in the brain, cytokine signals participate in pathways known to be involved in the development of depression, including: (i) altered metabolism of relevant neurotransmitters such as serotonin (5HT) and dopamine (DA), (ii) activation of CRH in the paraventricular nucleus (PVN) and the subsequent production and/or release of ACTH and glucocorticoids (cortisol), and (iii) disruption of synaptic plasticity through alterations in relevant growth factors (e.g., brain-derived neurotrophic factor [BDNF]). **(E)** Exposure to environmental stressors promotes activation of inflammatory signaling (NF-κB) through increased outflow of proinflammatory sympathetic nervous system responses (release of NE, which binds to the α [αAR] and β [βAR] adrenoceptors). **(F)** Stressors also induce withdrawal of inhibitory motor vagal input (release of acetylcholine [Ach], which binds to the α7 subunit of the nicotinic acetylcholine receptor [α7nACRh]). **(G)** Activation of the mitogen-activated protein kinase pathways, including p38 and Jun N-terminal kinase (JNK), inhibit the function of glucocorticoid receptors (GRs), thereby releasing NF-κB from negative regulation by glucocorticoids released as a result of the hypothalamic–pituitary–adrenal (HPA) axis in response to stress. (See color atlas.) (From Raison CL, Capuron L, Miller AH. Trends in Immunology, 27, Cytokines sing the blues: inflammation and pathogenesis of depression, 24–31, 2006, with permission from Elsevier.)

Acute/Mild Stress

Although a number of studies have demonstrated decreases in a variety of functional immune parameters (primarily related to acquired immune responses following acute stress), more recent data have suggested that the effects of acute stress are quite complex and involve considerations of the nature and location of the immune stimulus. Indeed, data from laboratory animals suggest that brief and/or mild stressors may actually enhance adaptive immunity. For example, in a series of elegant experiments in rats, it has been shown that a brief stressor (e.g., 2 hours of physical restraint) applied prior to antigen challenge significantly enhances delayed-type hypersensitivity (DTH), an antigen-specific reaction mediated by $CD4^+$ T lymphocytes, in the skin. Similarly, rodents exposed to an acute and/or mild stressor prior to antigen presentation have been reported to demonstrate enhanced humoral immunity by producing more antibodies than animals exposed to the same antigen in a control condition. Interestingly, the enhanced DTH and humoral immunity observed with mild/acute stressors are opposite to the effects seen if the stress intensity or duration is increased, even if the stressor type remains the same. In the case of DTH, the enhancing effect of acute stress appears to depend on stress-related increases of the glucocorticoid hormone, corticosterone, which can significantly increase the trafficking of immune cells to the site of antigen challenge.

Table 1.13–9.
Factors in the Modern World That Impact Inflammation and Depression

Factors that Increase Inflammation And Increase Risk of Depression	Factors that Decrease Inflammation And Decrease Risk of Depression
Early life adversity	Mediterranean diet
Obesity	Exercise
Sedentary lifestyle	Positive social connectivity
Processed foods	Early contact with pets and farm animals
Psychosocial stress	Experiencing reverence and awe
Social isolation	Antidepressants
Poor sleep/insomnia	Electroconvulsive therapy
Smoking	Psychotherapy
Drug abuse/dependence	Meditation
Low socioeconomic status	
Female sex	
Reduced childhood exposure to co-evolved micro-organism	

Initial studies of acute stress in humans focused on changes in enumerative and in vitro immune system functional assessments. On the basis of a number of studies and three meta-analyses, a clear pattern of

immune changes emerges from a host of acute stressors ranging from psychosocial laboratory stressors to first-time parachute jumping (Table 1.13–10). Enumerative changes include increased numbers of white blood cells, CTLs, and NK cells and decreased numbers of total B lymphocytes. Functional changes associated with acute stressors in humans include a decrease in lymphocyte responses to several nonspecific mitogens and an increase in NKCA, although it should be noted that this increased activity may largely reflect an increase in NK cell number. More recently, acute stressors have also been shown to suppress circulating number of Treg cells.

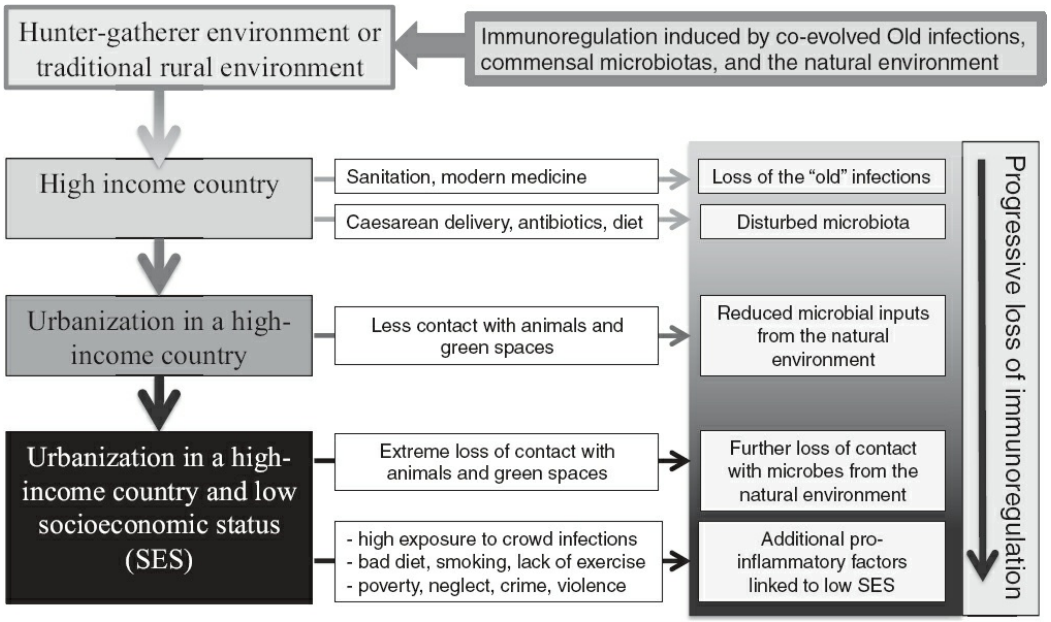

FIGURE 1.13–6. Progressive loss, in high income settings, of contact with the immunoregulatory organisms with which humans coevolved. In communities with low socioeconomic status, there are numerous additional factors that also tend to increase background inflammation, including psychosocial stressors, bad diet, obesity and continued exposure to crowd infections. (Modified from Rook GA, Raison CL, Lowry CA. Microbial "old friends," immunoregulation and socioeconomic status. *Clin Exp Immunol*. 2014;177(1):1–12.)

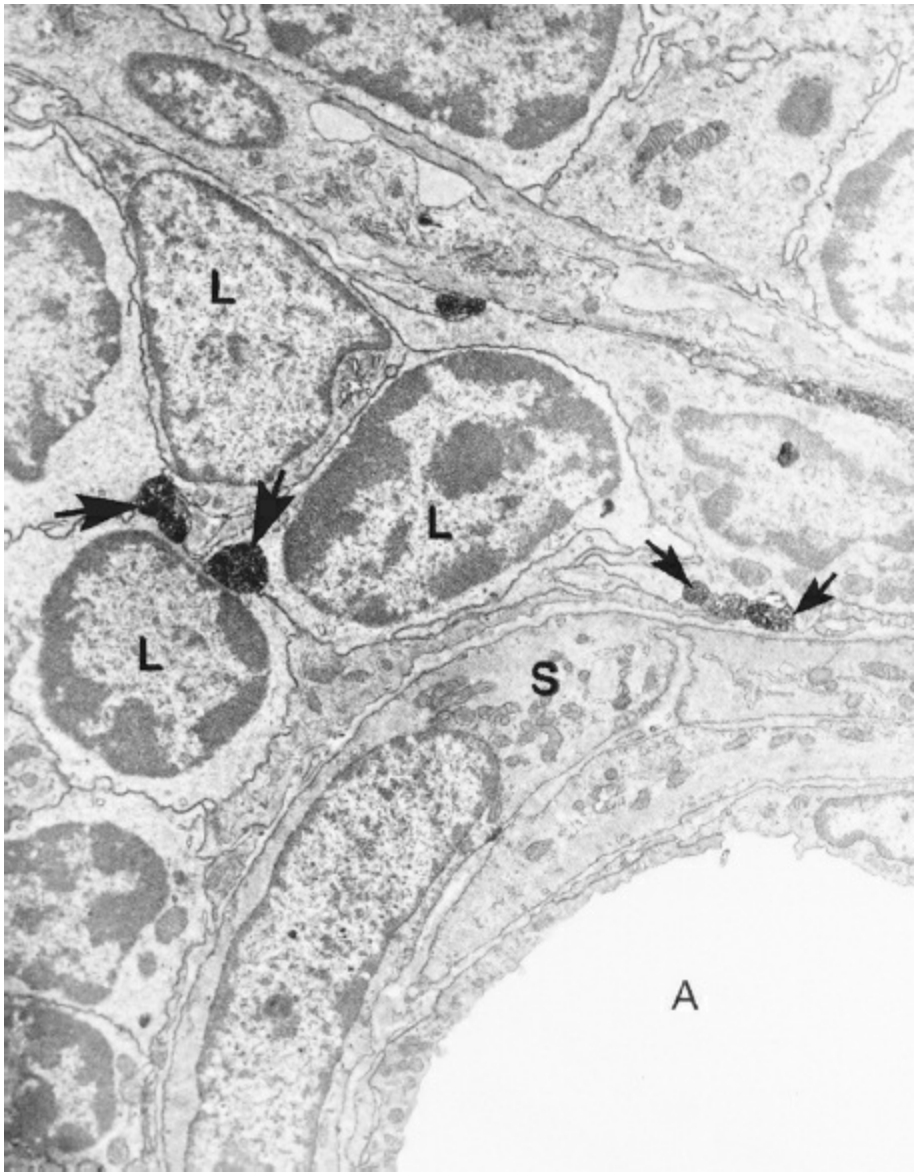

FIGURE 1.13–7. Sympathetic nervous system innervation of lymphoid tissue. Tyrosine hydroxylase-immunoreactive nerve processes (*small arrowheads*) in contact with the smooth muscle (*S*) of the central arteriole (*A*) and nerve processes (*large arrowheads*) in direct contact with lymphocytes (*L*) in the periarteriolar lymphatic sheath of the rat spleen. Transmission electron micrograph, ×6,732. (Courtesy of Denise L. Bellinger, Center for Neuroimmunology, Loma Linda University, Loma Linda, CA, and Suzanne Y. Stevens, Department of Neurobiology and Anatomy, University of Rochester School of Medicine and Dentistry, Rochester, NY.)

Table 1.13–10. Effects of Acute Stress, Chronic Stress, and Depression on Immune Parameters in Humans

Immune Variable	Acute Stress	Chronic Stress	Major Depression
Leukocytes	↑/↑	↑/↑	↑/↑
Lymphocytes	↑/↑	↔	↓/↓
Monocytes	0	↔	0
T lymphocytes	0	↓/↓	↔

B lymphocytes	↓/↓	↓/↓	↔
$CD4^+$ cells	0	↓/↓	↔
$CD8^+$ cells	↑/↑	↓/↓	↔
NK cells	↑/↑	↓/↓	0
$CD4^+/CD8^+$	↓/↓	↓/↓	↔
% B cells	↓/↓	↔	↔
% T cells	0	↓/↓	↔
% $CD4^+$	↓/↓	↔	↑/↑
% $CD8^+$	0	↓/↓	↔
% $CD4^+/CD8^+$	↓/↓	↔	↔
% NK cells	↑/↑	↔	↔
NKCA (total)	↑/↑	↓/↓	↓/↓
NKCA (per cell)	↔	↓	↓
Mitogen-induced proliferation	↓/↓	↓/↓	↓/↓
Response to vaccine	↑	↓/↓	↓
Antibody titers to EBV	↑/↑	↓/↓	?
Cytolytic T-cell response to antigen	?	↓	?
Wound healing	?	↓	?
Th1/Th2 balance	↑	↓	?
DTH	↑	↔	↓
TNF-α	0	?	↑
IFN-γ	↑/↑	?	?
IL-1	↑/↑	?	↑
IL-2	?	↓/↓	↑
IL-4	↓/↓	?	?
IL-6	↑/↑	↑	↑/↑
NF-κB	↑	?	?
CRP	↑/↑	↑	↑

↑/↑, positive effect confirmed by meta-analysis; ↑, majority of studies suggest positive effect; ↓/↓, negative effect confirmed by meta-analysis; ↓, majority of studies suggest negative effect; 0, meta-analysis suggests no effect; ↔, conflicting findings;?, not enough data to suggest positive or negative relationship; NK cells, natural killer cells; CD4/CD8, ratio of $CD4^+$ T lymphocytes to $CD8^+$ T lymphocytes; CRP, C-reactive protein; NKCA, natural killer cell activity; EBV, Epstein–Barr virus; Th1/Th2, the ratio of T helper cell type 1 cytokines to T helper cell type 2 cytokines; DTH, delayed type hypersensitivity; NF-κB, nuclear factor κB.

From Raison CL, Gumnick JF, Miller AH. Neuroendocrine–immune interactions: Implications for health and behavior. In: *Hormones, Brain and Behavior*. Vol 5. San Diego, CA: Academic Press; 2002, with permission.

Regarding innate immune inflammatory responses, acute stress exposure in humans and laboratory animals has been shown to increase the expression of innate immune cytokines, activate microglia, and sensitize subsequent immune responses to inflammatory immune challenge. For example, peripheral blood concentrations of proinflammatory cytokines and/or their soluble receptors, especially IL-6, have been reported to be increased in the context of several acutely stressful situations, including

public speaking, mental arithmetic, exercise, and academic examinations. While the mechanism by which stress induces cytokine production has yet to be fully elucidated, it has been shown that catecholamines may play an important role (see below). The effects of stress on innate immune responses appear to be mediated in part by activation of inflammatory signaling pathways including NF-κB, which is a lynchpin in the initiation of the inflammatory response following the stimulation of TLRs as well as relevant cytokine receptors. Recent studies have highlighted the health relevance of even mild and/or acute stress in humans. Two studies have now shown, for example, that men and women with histories of early life adversity (ELA) show markedly increased inflammatory responses to an acute laboratory stressor (the Trier Social Stress Test [TSST]), as assessed by increases in plasma concentrations of IL-6 and by increased production of the key intracellular inflammatory signaling molecule NF-κB (Fig. 1.13–8). And circulating levels of Il-1-beta following the TSST were found to mediate the relationship between increased negative emotionality in response to the stressor and higher levels of depressive symptoms a year later.

Chronic/Severe Stress

In keeping with the well-documented health risks associated with chronic or severe stress, many studies confirm that more pernicious types of stressors are associated with immune alterations that may predispose toward disease development and are, in some cases, opposite to changes seen in the context of acute and/or mild stress (Table 1.13–10). Nevertheless, like acute stress, chronic stress also appears to be associated with enhanced proinflammatory activity.

Many individual studies and three large meta-analyses of the chronic stress literature have found that, in humans, naturalistic stressors lasting from days to years evince a consistent pattern of functional immune changes that include a decrease in NKCA, a decrease in mitogen-induced lymphocyte proliferation, and a functional resistance to glucocorticoids in monocytes. Enumerative changes such as increases in circulating white blood cells and decreases in NK and T cells as well as alterations in the number and ratio of $CD4^+$ and $CD8^+$ T cells also have been reported. However, the consistency of these findings across studies is in part dependent upon the definition of chronic stress.

Chronic and/or severe stress has been found to affect both humoral (i.e., antibody) and cellular immunity in ways directly relevant to "real world" immune functioning. Stressors such as exam taking or caring for a demented spouse have been repeatedly shown to impair the body's ability to suppress the activity of latent viruses (especially Epstein–Barr virus [EBV]) as measured by an increase in latent viral antibodies and to

interfere with antibody development following vaccination. Consistent with these findings, both examination and caregiving stress are associated with decrements in memory T-cell responses to latent virus antigens and to vaccines, both in terms of antigen-induced T-cell proliferation and T-cell–mediated killing of virally transformed B lymphocytes.

In addition to effects on enumerative, functional and naturalistic measures of immune functioning, it has become increasingly recognized that both acute and chronic stress alter the mix of cytokines produced by T lymphocytes and that this alteration may contribute to the effects of stress on immunity and disease. While acute/mild stress in animals and humans may favor the in vitro production of the Th1 cytokine IFN-γ without increasing the Th2 cytokine IL-4, chronic/severe stressors in animals and humans tend to suppress lymphocyte production of the Th1 cytokines IFN-γ and IL-2, while not effecting or actually increasing the Th2 cytokines IL-4 and IL-10. This shift from Th1 to Th2 may increase the susceptibility to allergic or hypersensitivity reactions, where Th2 responses play an important role. Interestingly, in a recent study in mice, chronic restraint stress was found to increase the susceptibility to skin cancer in association with decreases in IFN-γ production and increased expression of $CD4^+$ $CD25^+$ regulatory (suppressive) T cells.

As noted above, chronic stress has also been associated with increased innate immune responses. For example, increased chronic stress prior to experimental inoculation with influenza A virus was found to correlate with higher IL-6 concentrations in nasal lavage and with increased behavioral symptom scores, suggesting that stress-related inflammatory activation may impair the body's ability to control the immune response to viral infection. Moreover, recent data indicate that early life stress including parental loss or physical and/or sexual abuse during childhood is associated with increased markers of innate immune system activation, including CRP, in both nondepressed and depressed individuals at rest and in response to both acute and chronic stress. As noted in Figure 1.13–3, inflammation is believed to be involved in a number of medical illnesses, and therefore the relationship between early life stress and inflammation may represent an important pathophysiological mechanism that explains the increased medical morbidity in individuals exposed to early life stress.

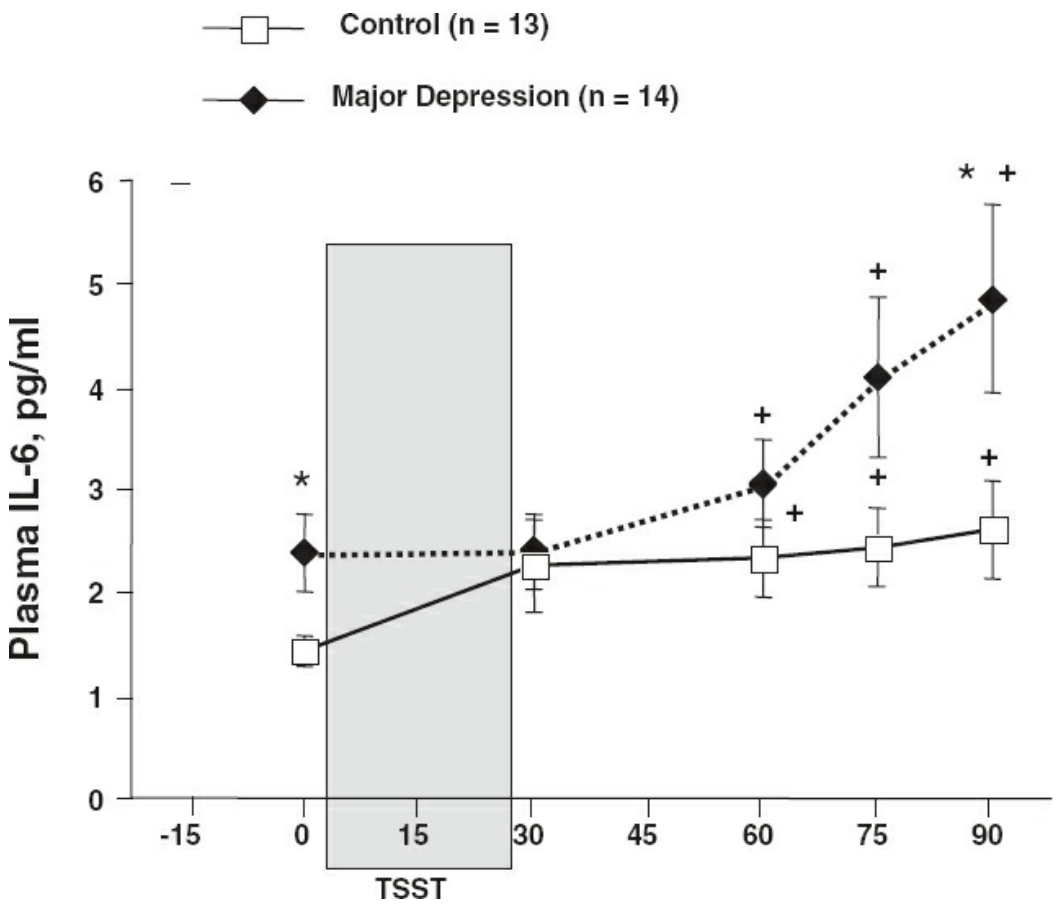

FIGURE 1.13–8. Early life stress, depression, and inflammation: Plasma interleukin (IL)-6 in patients with major depression and nondepressed comparison subjects before and after psychosocial stressor challenge. +Significant difference from baseline ($p < 0.05$). *Significant difference between groups ($p < 0.05$). (Modified from Pace TWW, Mletzko TC, Alagbe O, et al. Increased stress-induced inflammatory responses in male patients with major depression and increased early life stress. *Am J Psychiatry.* 2006;163:1630–1633.)

Stress-Relevant Social Processes and Immune Functioning

Physical and mental health are often thought of as intrapersonal phenomena, and stress is often conceptualized as a process that happens within the individual. Importantly, however, individuals live in social environments and continually interact with, and are influenced by, significant social others, such as spouses, partners, parents, and children. Recent evidence suggests that these social factors can influence immune regulation and functioning in ways that transcend simpler associations between stress and immunity.

Distressing relationships, such as those that involve stressful encounters (e.g., martial strain) or negative interaction behaviors (e.g., showing anger or hostility, and interrupting or criticizing) can dysregulate immune function. Observational studies of relationship interactions, which use behavioral coding systems that assess actual behaviors as opposed to self-reported marital quality, have demonstrated that greater hostility displayed during relationship interactions is associated with several

indicators of poorer immune function; these findings include larger increases in circulating markers of inflammation post-discussion, slower wound healing from suction blisters, smaller antibody responses to an influenza virus vaccine, and higher EBV antibody titers—the latter of which indicates poorer ability to control this latent herpes virus that infects most adults. A recent longitudinal study suggests a similar pattern of results: spouses who were in more distressing marriages at baseline had larger declines in cellular immune function (proliferative responses to two mitogens, concanavalin A and phytohemagglutinin) 2 years later when compared to spouses in less distressing marriages.

Overtly negative social interactions or behaviors may not be required to adversely affect immune function. Interestingly, recent evidence suggests that being ambivalent about relationship partners—i.e., feeling simultaneously negative and positive toward a close relationship partner—may also influence immunity. Perceiving ambivalence toward one's spouse in a support context was linked to greater inflammation (higher IL-6 and fibrinogen, and marginally higher CRP levels) even when considering health behaviors, relationship-specific romantic attachment style, spouse negativity/positivity ratings, and overall marital satisfaction. Furthermore, coronary-artery calcification scores were highest for individuals who both viewed and were viewed by their spouse in an ambivalent manner, suggesting that ambivalence may affect clinically relevant outcomes (e.g., cardiovascular disease, stroke), and potentially through immune mechanisms.

However, supportive relationships, including those that involve positive interaction behaviors, supportive communication patterns, and warm touch, may confer immunoprotective effects. The majority of perceived social support research suggests that people who feel more supported have better immune function than those who feel less supported. Benefits for people who feel more supported include lower systemic inflammation, larger antibody responses to pneumococcal pneumonia, hepatitis B, and influenza virus vaccines, and lower EBV titers, than those who feel less social support. Additionally, couples who displayed more positive behaviors (e.g., acceptance, self-disclosure, humor) and who used more cognitive processing words during interaction tasks exhibited faster wound repair and had lower systemic inflammatory responses even 24 hours after the interaction task than couples who displayed fewer positive behaviors and used fewer cognitive processing words. Findings also suggest that warm physical contact (hugging and kissing) may be immune enhancing; circulating levels of IFN-γ decreased significantly in couples after an hour-long experimental induction of warm physical contact (hugging and kissing), whereas levels did not change in the control condition (couples who read books in separate rooms).

Relationship processes that involve social rejection also have

immunological consequences. Social rejection has often been studied in the context of depression due to the sustained inflammatory process that may elicit sickness behaviors and precipitate depression for vulnerable individuals. In a longitudinal study of 147 adolescent girls at elevated risk for depression, participants had significantly higher levels of mRNA for both proinflammatory transcription factor NF-κB and I-κB, which regulates the effects of NF-κB, at visits when they had experienced a recent targeted rejection life event compared to visits when no such event had occurred. A growing body of research suggests that stressors involving social rejection and exclusion activate neural regions involved in processing negative affect, including the dorsal anterior cingulate cortex (dACC) and anterior insula. These neural regions activate multiple biological systems, including, in particular, the HPA axis and sympathetic–adrenal–medullary axis, which produce cortisol and catecholamines that can bind to receptors on immune cells, which then modulate the release of proinflammatory cytokines. Thus, social stress-related implications on the neurocognitive pathway involving the dACC and anterior insula may be one mechanism linking social threat and rejection with elevated inflammation and risk for depression. Indeed individuals with elevated dACC responses to social rejection during a neuroimaging paradigm also show increased inflammatory responses to a laboratory psychosocial stressor.

In a series of groundbreaking papers, Steve Cole and colleagues have shown that many types of psychosocial adversity (i.e., loneliness and conflict) share in common a pattern of gene expression characterized by increased NF-κB–mediated inflammation and reduced type I IFN signaling. Interestingly this pattern appears to subserve an adaptive evolutionary logic. Across evolutionary time, individuals in close interpersonal groupings would have benefited from enhanced antiviral immunity conferred by type I IFN activity, whereas isolated individuals would be more likely to face infectious threat from dangers in the environment, especially bacteria delivered by wounding (because ostracized humans were at significantly increased risk of predation and/or physical injury from other social groups).

Disruptions in Relations with the Microbial World as a Source of Inflammation in the Modern World

Increasing evidence suggests that modern environmental practices have disrupted human relationships with a range of micro-organisms in ways that may help to account for a wide range of immune-based pathologies that are far more common in industrialized societies than in environments more consistent with those in which humans evolved. Initial theorizing regarding the role of human–microbial interactions in the explosion of immune disorders in the twentieth century focused on the possibility that reduced exposure to childhood infectious diseases in the modern world

might be the culprit. As originally articulated in the "hygiene hypothesis," it was theorized that factors unique to industrialized societies, including improved sanitation, modern medicine, and smaller family size, reduce the prevalence and change the timing of childhood infections in a way that promotes the development of allergy and asthma. Because infections typically mobilize Th1-type inflammatory responses, a necessary correlate of this idea was that the loss of infection-induced Th1 activation early in life released Th2 processes from appropriate regulatory control, with resultant increases in allergy and asthma. However, rather than decreasing (as would be predicted from the Th1–Th2 balance idea), an array of Th1-mediated autoimmune conditions was also exploding in prevalence in exactly the same countries in which allergy was on the rise. A potential resolution to this inconsistency, implicating other types of microbial and parasitic agents with powerful effects on regulatory dendritic cells and Tregs, began to emerge in the late 1990s. As eventually articulated in the "old friends" hypothesis, the link between modernity and increased inflammatory, atopic, and autoimmune disease might be better explained by disruptions of ancient associations with mostly nonpathogenic micro-organisms once ubiquitous in both the external and internal (i.e., microbiota) human environment, which have been largely (or completely) banished from developed countries (Fig. 1.13–3). These organisms induce and maintain an adaptive level of immune suppression by stimulating T cells to differentiate along regulatory, rather than Th1 or Th2 lines, with a resultant increased production of anti-inflammatory, immunoregulatory cytokines, especially IL-10 and TGF-β.

Significant in vitro and in vivo preclinical data indicate that all three classes of "Old Friend" organisms (microbiota, pseudocommensals, and parasites) have the potential to prevent or ameliorate pathology in animal models for autoimmune, allergic, and inflammatory bowel diseases, as well as neoplasms and certain infections, and do so via reductions in inflammatory activity. While data implicating alterations in human–microbial interactions in the development of psychiatric disease are limited, many animal studies indicate that these interactions have immune and behavioral effects of direct relevance to psychiatric functioning. Perhaps the most compelling animal data demonstrating the ability of the microbiota to influence brain functioning and stress-relevant behavior come from studies showing that transplanting the microbiota from one strain of mouse to another causes the recipient mice to take on the behavioral repertoire of the donors. In addition to direct effects on immune functioning, there is evidence to suggest that the gut microbiome profoundly affects other whole body physiological processes, including peripheral pain sensitivity, sleep, and metabolism, that are affected by activity in inflammatory signaling pathways and that are abnormal in a variety of psychiatric conditions. Moreover, an increasing number of

animal studies suggest that host–microbial interactions early in development have profound effects on stress physiology and behavior by affecting the development and functioning of neuroendocrine and CNS processes relevant to the development of resilience or disease in response to stress.

As with many processes relevant to psychiatry, the relationship between the microbiota and stress observed in animal models is bidirectional. Just as the composition of the microbiota appears to affect stress responses, stress changes the composition of the microbiota. For example, in mice the stress of maternal separation for 3 hours per day from postnatal days 2 to 12 has long-term effects on the subsequent 16S ribosomal RNA diversity of the microbiota that persists into adulthood. In addition to the ability of early life stressors to program baseline microbiota composition across the lifetime, animal studies demonstrate that acute stressors in adulthood can powerfully alter the composition of the microbiota. These alterations may then contribute to the behavioral changes observed in response to stress. For example, chronic exposure of mice to a grid flooring stressor produced significant changes in microbiota composition, which correlated with several circulating immune markers and with anxiety as assessed by closed arm entries and total time spent in the elevated plus-maze, time spent in a light/dark box, and time spent in the inner zone and total time spent in the open field test.

The data linking the microbiota and other environmental "old friends" organisms with psychiatric illness in humans lags far behind the more conclusive evidence derived from animal models. Nonetheless, several small studies suggest that probiotics may hold impact brain function and hold promise for the treatment of behavioral disorders. Similarly, a large study in patients with advanced lung cancer found that coadministration of *Mycobacterium vaccae* (an environmental saprophyte) with chemotherapy improved mood, energy and related symptom domains. Finally, the rapidly expanding understanding of the relevance of human–microbial and human–parasitic interactions to disease pathogenesis suggests that in addition to symptomatic treatment, organism-based interventions may show promise, if delivered early in life, as preventive or ameliorative strategies not just for depression and anxiety but perhaps for schizophrenia and autism as well.

Immunological Mechanisms

Investigators have examined the immunological mechanisms through which stress may affect the immune system. In general, a stressor can alter immune function in two major ways. First, the stressor can lead to changes in the distribution of immune cells in any given part of the body. Second, stress can alter the function of the immune cells themselves. Because the immune response depends on the interplay of various immune cell

subtypes, a redistribution of relevant cell types into or out of a particular immune compartment can directly influence the local immune response. As mentioned previously, significant and selective changes in immune cell distribution have been described in rats undergoing a mild acute stress (2 hours of restraint), including decreased numbers and percentages of cells in the peripheral blood and a concomitant increase in immune cells in the skin. It appears that cells leaving the blood during acute stress may migrate to the skin where they might be more likely to encounter pathogens, thus facilitating the immune response to pathogen exposure in wounded tissues (a distinct evolutionary advantage).

Interestingly, data suggest that stress-induced activation of proinflammatory pathways may contribute to many of the classic immunosuppressive findings associated with stress. For example, production of NO by macrophages (an early step in inflammatory activation) has been shown to be involved in the biochemical mechanisms underlying the ability of stress to reduce lymphocyte proliferation. Both depletion of macrophages and inhibition of NO synthesis have been shown to attenuate stress-induced changes in acquired immune responses. Consistent with this finding, it is known that chronic proinflammatory cytokine production inhibits several indices of T-cell–mediated immunity. Similarly, the ability of a brief laboratory stressor to attenuate antibody responses to the antigen keyhole limpet hemocyanin is largely reversed when an antagonist to IL-1 is coadministered with the stressor. Other data demonstrate that TNF can disrupt T-cell signaling and thus impair cell-mediated immune mechanisms. In humans, depressive symptoms have been associated with increased inflammatory responses (increased plasma IL-6) to an influenza vaccine, while at the same time portending reduced antibody responses to the vaccine. Taken together, these data suggest that chronic stress may lead to a state in which innate immune mechanisms are hyperactive at the expense of functionality in acquired immune pathways.

Neuroendocrine Mechanisms

A number of studies have focused on neuroendocrine mechanisms by which stress may influence the immune response. The two systems that have received the most attention are the HPA axis and the ANS.

Sympathetic Nervous System

As discussed previously, catecholamines released by local SNS nerve fibers within immune tissues can have complex suppressive and enhancing effects on immune responses. Regarding the capacity of stress to activate innate immune responses, catecholamines have been shown to stimulate the production of proinflammatory cytokines, especially IL-6, in both the peripheral blood and CNS, and activate inflammatory signaling cascades

including NF-κB in peripheral blood mononuclear cells (Fig. 1.13–5). In rodents, β-adrenergic receptors are critical for central production of inflammatory cytokines, whereas both α- and β-adrenergic receptors contribute to the induction of cytokines in the peripheral blood following stress. Of note, activation of catecholamines in the context of immune stimulation with LPS is associated with decreased cytokine (e.g., TNF-α) production in the spleen, an effect that can be eliminated by sympathectomy. Catecholamines via the β2 adrenergic receptor (β2AR) also reliably inhibit LPS-induced cytokine release in vitro, while having a stimulatory effect on cytokine production when administered alone. The ability of catecholamines to induce inflammation may be related in part to a "switch" that occurs whereby PKA activation leads to β-adrenergic receptor phosphorylation, which in turn switches β-receptor signaling from G_s to G_i. G_i has been associated with activation of the ras–raf signaling cascade. In addition, activation of the α1 adrenergic receptor (α1AR) may be involved in innate immune system activation following NE exposure. Regarding the effects of stress on acquired immune responses, catecholamines via activation of β2ARs have been shown to inhibit Th1 cytokines (IFN-γ), an effect that is enhanced by the presence of glucocorticoids, which can increase β2AR expression. The health-relevance of adrenergic—immune interactions has recently been highlighted by studies suggesting that beta-blockers may hold promise in the prevention of cancer or its recurrence, likely via immunomodulatory effects.

HPA Axis

In concert with the ANS, the HPA axis serves as a central component of the mammalian stress response system and ultimately functions to maintain bodily homeostasis via mediation of immunosuppression and immune system activation. This is exemplified by the varied effects of both CRH (an early product of HPA axis activity) and glucocorticoids (the final product of HPA axis activation).

CRH applied within the CNS suppresses several measures of immunity, including splenic NKCA, mitogen-stimulated lymphocyte proliferation, and in vivo and in vitro antibody formation, as well as T-cell responses to T-cell receptor antibody. CRH-overproducing mice demonstrate a profound decrease in the number of B cells and severely diminished primary and memory antibody responses. These immunosuppressive effects appear to be mediated by stress response outflow pathways activated by CRH, given that blockade of the SNS abolishes CRH effects on NKCA, and adrenalectomy obviates CRH effects on lymphocyte proliferation. In addition, the B-cell decreases in CRH-overproducing mice are consistent with the marked reduction in rodent B cells observed after chronic glucocorticoid exposure.

In contrast to its immunosuppressive properties, CRH has also been shown to enhance proinflammatory cytokine production in animals and humans when administered peripherally or within the CNS. Chronic ICV administration of CRH in rats leads to induction of IL-1β mRNA in splenocytes, and acute intravenous administration in humans has been reported to cause a fourfold increase in the induction of IL-1α. Similarly, the addition of CRH to in vitro mononuclear cell preparations induces the release of IL-1 and IL-6. Both chronic and acute CRH infusion have also been reported to increase production of IL-2 in humans and animals. In addition to potential proinflammatory activities of CRH within the CNS, peripheral production of CRH has been demonstrated in inflammatory diseases, such as arthritis and ulcerative colitis, in which CRH appears to act as a local proinflammatory agent. Indeed, CRH acting via CRH receptor 2 (CRHR2) appears to play a direct role in mediating intestinal inflammatory responses to enterotoxins.

Of the effects of all the neurotransmitters or hormones known to modulate immune functioning, the actions of glucocorticoids, albeit complicated, are probably best understood. Glucocorticoids modulate the following: immune cell trafficking throughout the body, cell death pathways (i.e., apoptosis), and Th1/Th2 cellular immune response patterns in a manner that inhibits Th1 (cell-mediated) responses and promotes Th2 (antibody) responses. At the same time, glucocorticoids inhibit the following: arachidonic acid pathway products (e.g., prostaglandins) that mediate inflammation and sickness symptoms (e.g., fever), T-cell– and NK-cell–mediated cytotoxicity, and cytokine production and function through interaction of GRs with transcription factors (NF-κB, in particular), which in turn regulate cytokine gene expression and/or the expression of cytokine-inducible genes.

Although glucocorticoids may actually enhance certain aspects of naturalistic immune functioning (e.g., when produced for brief periods, at low to moderate doses, or in the context of acute and/or mild stress as exemplified by DTH via effects on immune cell trafficking), they in general play a primary role in restraining excessive or prolonged inflammatory activation. This property has long been exploited for the treatment of autoimmune and other chronic inflammatory conditions, with the result that glucocorticoids remain a cornerstone of the clinical anti-inflammatory armamentarium. Consistent with their pharmacological uses, endogenous glucocorticoids have been shown to be essential for inflammatory regulation in response to immune system activation. For example, neutralization of endogenous glucocorticoid function results in enhanced pathology and mortality in animals exposed to LPS as well as other inflammatory stimulators, such as streptococcal cell wall antigen or myelin basic protein. Similarly, rodents that have been rendered glucocorticoid deficient by adrenalectomy have markedly increased death rates following

infection with murine cytomegalovirus, an effect that arises from unrestrained activity of TNF-α. Finally, blockade of GRs within the CNS promotes profound neurodegeneration in response to LPS exposure, an effect not seen when glucocorticoid signaling is intact.

IMMUNE SYSTEM EFFECTS ON THE CNS

Tremendous interest has been generated in the neurosciences by the discovery that the immune system—largely via cytokine activity—is capable of exerting profound effects on CNS function (Fig. 1.13–5). Cytokines are involved in the regulation of sleep, temperature, feeding behavior, motor activity, cognition, and hedonic tone, not only in the context of infection but also in response to stress-induced and circadian changes in cytokine production. In addition, cytokines play a role in the regulation of neurotransmitter metabolism, neuroendocrine function, synaptic plasticity, and regional brain activity.

Cytokine Effects on Behavior

When challenged with a medical illness or chronic psychological stressor, complex interactions between the immune and nervous systems promote a constellation of immune-induced behavioral changes, alternatively referred to as "sickness syndrome" or "sickness behavior." These behavioral changes include dysphoria, anhedonia, fatigue, social withdrawal, hyperalgesia, anorexia, altered sleep-wake patterns, and cognitive dysfunction. Although seen in response to infection, the full syndrome can be reproduced in humans and laboratory animals by administration of innate immune cytokines, such as IFN-α, IL-1, TNF-α, IL-6, and IL-2. Blocking cytokine activity with IL-1 receptor antagonist (IL-1ra), α-melanocyte-stimulating hormone (αMSH), IGF-1 or IL-10 diminishes or prevents the development of sickness behavior in laboratory animals, even when such behavior develops as a result of psychological stress. Evidence that cytokine-induced behavioral toxicity is related to major depression comes in part from studies showing that in humans and laboratory animals, antidepressants are able to abolish or attenuate the development of sickness behavior in response to cytokine administration (see below).

Cytokine Effects on the Brain

Significant evidence indicates that the CNS is a primary site for the mediation of cytokine effects on behavior. Consistent with this notion, proinflammatory cytokines released in the periphery are capable of rapidly affecting CNS function. However, because cytokines are relatively large molecules that do not readily cross the BBB, considerable attention has been paid to the mechanisms by which peripherally released cytokines communicate with the brain (Fig. 1.13–9). Five major pathways have been

identified including (1) active transport of cytokines across the BBB; (2) access of cytokines to brain areas in which the BBB is relatively porous or "leaky," such as the organum vasculosum of the lamina terminalis; (3) conversion of cytokine signals into secondary signals, such as prostaglandin or NO, by endothelial cells that line the blood vessels of the brain; (4) transmission of cytokine signals via sensory afferents of the vagus nerve; and (5) a cellular pathway, which involves the trafficking of activated immune cells, typically monocytes, to the brain vasculature and parenchyma. There are data to support each of these mechanisms, and it appears that the pathway(s) by which cytokines signal the CNS depend in part on the concentration of cytokine in the peripheral blood or local tissue compartment. Indeed, evidence suggests that high concentrations of circulating cytokines enter the brain through leaky regions of the BBB, whereas lower concentrations of cytokines signal the brain through afferent nerve fibers. The cellular pathway is the one most recently identified, perhaps because of a long-standing belief that the brain is an immune privileged organ. It is increasingly apparent that this is only partially true, and this realization is one of the most exciting recent developments in the understanding of how peripheral inflammatory activity can influence brain function. Studies have now shown that the release of TNF from inflamed peripheral tissues stimulates microglial cell production of chemokine (C-C motif) ligand 2 (CCL2 also known as monocyte chemoattractant protein-1) that then attracts monocytes to the brain. In response to social defeat stress in rodents, monocytes infiltrate multiple brain regions associated with the detection of threat (e.g., amygdala); an effect that is dependent on CCL2 and requires mobilization of monocytes from the bone marrow in dependence upon stress-induced release of catecholamines. Recently the relevance of the cellular pathway of brain-immune communication has received support from postmortem analyses of brain tissue from patients with depression who committed suicide. These brains showed increased numbers of perivascular macrophages that was associated with increased gene expression of CCL2 and allograft inflammatory factor 1 (AIF1, also known as IBA1), both of which are associated with macrophage activation and cellular trafficking.

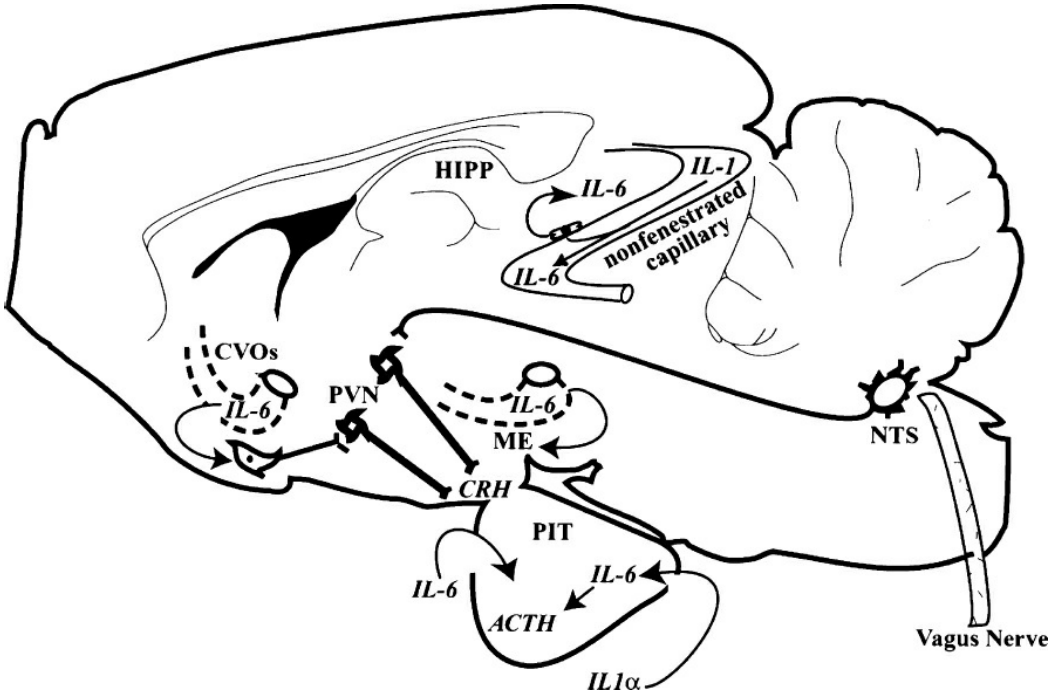

FIGURE 1.13–9. Schematic of possible pathways for translation of IL-1 and IL-6 into neuroendocrine signals during viral infection. Both IL-1 and IL-6 are thought to play a role in activating the hypothalamic–pituitary–adrenal (HPA) axis during viral infections, and there is evidence for various pathways by which this occurs. During infection, IL-6 produced in the liver (or other abdominal viscera) may stimulate vagal afferents, thereby activating central neurons in the nucleus of the solitary tract (NTS), which in turn send catecholaminergic projections to the paraventricular nucleus (PVN) of the hypothalamus. Alternatively, circulating IL-1 or IL-6 (which are confined to the vascular lumen of nonfenestrated capillaries) may act indirectly by recruiting the brain's cytokine network. In such a scenario, IL-1 could induce production of IL-6 from capillary endothelium, microvascular pericytes, or perivascular glia. Once in the brain parenchyma, IL-6 might act directly on HPA axis regulatory neurons or, more likely, modulate HPA axis activity through intermediates such as prostaglandins. In addition, IL-6 could enter the brain parenchyma passively by diffusing through the fenestrated capillaries of circumventricular organs (CVOs). Capillaries of the median eminence (ME) are also fenestrated, allowing IL-6 to travel from the vascular lumen to nerve terminals of the PVN, consequently placing this cytokine in position to mediate CRH release. The pituitary gland is intimately associated with the brain and represents another site where virus-induced cytokines could mediate ACTH release. Because the anterior pituitary is outside the blood–brain barrier, IL-1 and IL-6 presumably have direct access to pituitary corticotrophs. The foregoing mechanisms are not mutually exclusive. For example, CRH could play a permissive role for the action of IL-6 at the pituitary. (From Pearce BD, Biron CA, Miller AH. In: Buchmeier MJ, Campbell IL, eds. *Advances in Virus Research*. Vol. 56. New York: Academic Press; 2001, with permission.)

Cytokine Network in the Brain

Once peripheral signals reach the brain, it appears that in many instances these signals are translated back into cytokine signals (i.e., peripheral cytokines beget central cytokines) in part through activation of resident cytokine-producing cells such as microglia. For example, peripheral administration of either IL-1 itself, or substances such as LPS that induce IL-1, is associated with a rapid increase in IL-1 immunoreactivity and bioactivity in several brain regions, including the hippocampus and

hypothalamus. NF-κB appears to play a pivotal role in this process, in that inhibition of NF-κB can block the elaboration of neuronal markers of activation (*c*-fos) as well as behavioral changes following peripheral IL-1 administration. Cytokines mediate their effects in the CNS through their receptors that are expressed throughout the brain with especially high concentrations in brain regions involved in the regulation of affect, stress responsiveness, and learning and memory including the hippocampus, amygdala, striatum, prefrontal cortex, and hypothalamus. Soluble receptors exist for many cytokines, as do decoy (inactive) receptors (such as the IL-1 receptor type II), both of which serve to limit cytokine action and are expressed in the brain. The endogenous soluble IL-1 receptor antagonist (sIL-1ra) has also been described in CNS tissues.

Cytokines, Neurotrophic Factors, Neurogenesis, and Neurodegeneration

Stress and other aversive manipulations have been repeatedly shown to suppress the production and release of neurotrophic factors, such as BDNF, that play an important role in synaptic plasticity in brain areas such as the hippocampus. Concomitant with this reduced trophic support, physical and psychological stressors suppress neurogenesis in the hippocampus, promote apoptotic cell death, and reduce density of synaptic connectivity between nerve cells. Increasing evidence suggests that proinflammatory signaling within the CNS may play an important role in these detrimental stress-induced processes. For example, social isolation stress in rodents has been shown to impair memory consolidation, suppress hippocampal neurogenesis, and reduce hippocampal BDNF levels. In separate experiments, administration of IL-1 receptor antagonist within the brain prior to the social isolation stressor has been shown to reverse each of these effects. Proinflammatory cytokines including TNF-α, IL-1, and IL-6 within the CNS have also been shown to play key roles in neurodegenerative changes found in illnesses such as Alzheimer disease, multiple sclerosis, and HIV dementia as well as following a variety of insults such as ischemia, trauma, and radiation exposure.

Cytokine Effects on Neurotransmitters

In laboratory animals, the acute administration of a host of cytokines has been shown to produce rapid and significant alterations in the metabolism of multiple monoamines, including serotonin, NE, and dopamine (DA), in numerous brain regions. Relevant to serotonin (5HT) metabolism, activation of the enzyme indoleamine 2,3 dioxygenase (IDO) is believed to be involved in cytokine-induced neurotransmitter alterations, in part by diverting the metabolism of tryptophan (the primary amino acid precursor of serotonin) into kynurenine, a compound that can be further converted into the neurotoxic metabolite quinolinic acid by activated microglia and

infiltrating monocytes and macrophages in the brain. Quinolinic acid directly activates receptors for glutamate (NMDA) while also stimulating glutamate release and blocking glutamate reuptake by astrocytes. Quinolinic acid—like the cytokines that stimulate its production—impacts glutamate metabolism by decreasing the expression of astrocyte glutamate reuptake pumps and stimulating astrocytic glutamate release, ultimately contributing to excessive glutamate both within and outside the synapse. Binding of glutamate to extrasynaptic NMDA receptors decreases production of BDNF and increased excitotoxicity.

Increasing evidence implicates kynurenine and its metabolites in depression in general and in inflammation-induced depression in particular. Indeed, administration of kynurenine in rodents induces depressive-like behavior in a dose-dependent manner. Moreover, increased levels of quinolinic acid have been found in microglia in the ACC of suicide victims who suffered from depression. Tying these observations to inflammation are studies using IFN-α, an innate immune cytokine used to treat infectious diseases and cancer, that is notorious for causing profound behavioral disturbances, including major depression in up to 50 percent of patients depending on the dose and duration of treatment. IFN-α activates IDO leading to significant reductions in plasma concentrations of tryptophan. Several studies have shown an association between the degree of peripheral tryptophan decline and the development of depressive symptoms. However, the only study to examine associations between IFN-α–induced changes in CSF metabolites and the development of depression in humans found that the cytokine had no impact on central levels of tryptophan while significantly increasing CSF concentrations of both kynurenine and quinolinic acid. Importantly, the development of depressive symptoms correlated closely with increased CSF concentrations of both kynurenine and quinolinic acid, while showing no association with CSF tryptophan.

Innate immune cytokines, including TNF-α and IL-1, have also been shown to increase the expression and function of synaptic reuptake pumps for serotonin (and NE) via stimulation of p38 MAPK-linked pathways. The relevance of this mechanism to depression is highlighted by a recent human study showing that (1) increased circulating levels of TNF were associated with increased brainstem serotonin reuptake availability, (2) that a cytokine antagonist lowered brainstem serotonin transporter availability, and (3) that levels of both the transporter and peripheral TNF were associated with depressive symptoms.

Via the generation of reactive oxygen and nitrogen species, inflammatory cytokines have also been found to decrease the availability of tetrahydrobiopterin (BH4), an enzyme cofactor essential for the synthesis of all monoamines, likely as a result of BH4 being highly sensitive to oxidation. For example, CSF concentrations of BH4 correlate negatively

with CSF levels of IL-6 in patients treated with IFN-α, and the plasma phenylalanine to tyrosine ratio, an indirect measure of BH4 activity, correlates with CSF concentrations of DA as well as symptoms of depression in IFN-α–treated patients. Finally, treatment with the serotonin reuptake inhibitor, paroxetine, has been shown to attenuate the behavioral consequences of IFN-α administration, further highlighting the importance of effects on serotonin metabolism in cytokine-induced mood disturbances.

In addition to the effects of chronic cytokine exposure on serotonergic transmission, it has been reported that chronic administration of IL-2 or IFN-α significantly alters DA metabolism. For example, IFN-α–treated rhesus monkeys who displayed depressive-like behavior exhibited significantly greater decreases in CSF concentrations of the DA metabolite, homovanillic acid, than monkeys who did not exhibit such behavior. Moreover, IFN-α has been shown to lead to altered metabolic activity in brain regions high in dopaminergic neurocircuits including the basal ganglia. Relevant to cytokine-induced activation of IDO, data indicate that intrastriatal administration of kynurenic acid, a breakdown product of kynurenine, dramatically reduces extracellular DA in the rat striatum. Of note, cytokine induction of NO also has been shown to inhibit the activity of tyrosine hydroxylase (TH) (the rate-limiting enzyme in the synthesis of DA) through effects on the TH coenzyme tetrahydrobiopterin.

Although initial interest in the impact of inflammation on neurotransmitters focused on the monoamines (i.e., serotonin, NE, DA), more recent studies implicate changes in glutamate neurotransmission as being of central importance to the behavioral effects of cytokines. For example, increased levels of glutamate in the basal ganglia and dACC have been observed in patients receiving IFN-α, and higher levels of glutamate correlated with an increase in depressive symptoms. Interestingly, blocking glutamate receptors with ketamine or inhibiting IDO activity protects mice from LPS- or stress-induced depressive-like behavior, but leaves the inflammatory response intact. Taken together, these results indicate that increased activation of glutamate receptors by glutamate and/or quinolinic acid may be a common pathway through which inflammation causes depressive-like behavior.

Regional Brain Activity

Results from studies utilizing PET and functional magnetic resonance imaging (fMRI) provide further evidence that peripheral cytokine activity can induce centrally mediated behavioral changes. These and other imaging modalities provide a means by which various behavioral alterations can be associated with changes in the functioning of specific brain regions. For example, during an fMRI task of visuospatial attention, in comparison to controls, patients administered IFN-α exhibited significantly greater activation of the dACC. In patients, but not controls,

the degree of dACC activation correlated highly with the number of errors made during the task. Interestingly, increased dACC activity during cognitive tasks has also been demonstrated in patients vulnerable to mood disorders, such as those with high trait anxiety, neuroticism, or OCD. IFN-α has also been shown to lead to changes in basal ganglia metabolic activity (as measured by PET) that resemble those seen in Parkinson disease. These changes also correlate with IFN-α–induced fatigue-related symptoms and may be a function of IFN-α effects on DA metabolism as indicated above.

CYTOKINE EFFECTS ON THE NEUROENDOCRINE SYSTEM

Effects on the HPA Axis

Inflammatory cytokines have well-described effects on the HPA axis, including increased production of CRH and cortisol and decreased tissue sensitivity to glucocorticoid hormones. Although cytokines have been shown to be capable of activating the HPA axis at multiple levels, with a resultant increase in glucocorticoid release, evidence suggests that a major final common pathway for cytokine activation involves stimulation of CRH production in the PVN of the hypothalamus (Figs. 1.13–5 and 1.13–9). Several lines of evidence suggest that this increase in CRH activity may contribute to cytokine-induced depression/sickness behavior. CRH has behavioral effects in animals that are similar to those seen in patients with depression and/or sickness syndrome, including alterations in appetite, activity, and sleep. Patients with major depression frequently demonstrate increased CRH production, as assessed by increased CRH in CSF, increased mRNA in the PVN, downregulated frontal CRH receptors, and a blunted ACTH response to CRH challenge (likely reflecting downregulation of pituitary CRH receptors). Although never receiving approval for use, preliminary data suggest that agents that block the CRH type I receptor exhibit antidepressant and anxiolytic effects in humans. In animals, blocking CRH reverses some of the behavioral sequelae of proinflammatory cytokine administration. Indirect evidence for a role of CRH in cytokine-induced depression in humans comes from a study in which individuals that developed depression during IFN-α administration exhibited significantly higher ACTH and cortisol responses to the first injection of IFN-α compared to those of controls. These findings suggest that sensitized CRH pathways may serve as a vulnerability factor for cytokine-induced behavioral changes.

In addition to a direct stimulatory effect on CRH within the CNS, in vivo and in vitro studies suggest that inflammation may induce resistance to circulating glucocorticoids in nervous, endocrine, and immune system tissues. This is of great potential relevance, given the high rates of relative glucocorticoid resistance in HPA axis tissues (as assessed in vivo by the dexamethasone suppression test [DST] or the dexamethasone–CRH

stimulation test) and the immune system (as measured in vitro) found in patients with major depression and in animals and humans exposed to chronic and/or severe stressors. Supporting a role for cytokines in the induction of glucocorticoid resistance is the observation that many chronic inflammatory conditions, including steroid resistant asthma, rheumatoid arthritis, multiple sclerosis, and HIV infection, are characterized by a decrease in sensitivity to glucocorticoids. Of note, in HIV infection, glucocorticoid resistance has been shown to correlate with increased IFN-α plasma levels. Consistent with the ability of inflammatory cytokines to induce glucocorticoid resistance in animal models, IFN-α treatment in patients with hepatitis C virus infection has been shown to increase cortisol concentrations during the DST (consistent with glucocorticoid resistance). Increased DST cortisol levels, in turn, correlated with a more flattened diurnal cortisol slope ($r = 0.39$, $p = 0.038$) and with increased depression scores ($r = 0.57$, $p = 0.002$). Furthermore, the change in post-DEX cortisol was also associated with induction of the soluble tumor necrosis factor receptor 2 (sTNFR2; $r = 0.51$, $p = 0.006$) in response to IFN-α treatment.

Glucocorticoid Resistance

There are several mechanisms by which proinflammatory cytokines can disrupt GR function and contribute to glucocorticoid resistance. In addition to downregulating the expression of GR protein, proinflammatory cytokines have been found to increase the expression of the inert, β-isoform of the GR. Exposure of cells that constitutively express both GR-α (the active isoform) and GR-β to either TNF-α or IL-1β in vitro results in a marked increase in GR-β production, which is associated with the development of glucocorticoid resistance as demonstrated by a significant reduction in dexamethasone-stimulated activity of a GR-sensitive reporter gene in cytokine-treated cells. That overproduction of the negative GR-β isoform has a clinically relevant effect on glucocorticoid sensitivity is suggested by several studies documenting that patients with a variety of inflammatory and immune system disorders, including asthma, ulcerative colitis, and chronic lymphocytic leukemia, whose conditions are resistant to steroid treatment, demonstrate a significantly increased GR-β-to-GR-α ratio.

Another mechanism by which inflammatory cytokines may attenuate GR signal transduction, and hence cause glucocorticoid resistance, is through the induction of inflammatory signaling pathways that directly influence GR function. For example, adding IL-1α to an in vitro preparation of mouse fibroblast cells has been shown to suppress the ability of dexamethasone to induce translocation of the GR from the cytoplasm to its site of action in the nucleus. This IL-1α–mediated blockade of GR translocation from the cytoplasm to nucleus inhibits GR activity, as indicated by a decrease in the ability of dexamethasone to activate a

glucocorticoid-sensitive reporter gene construct. The signaling pathways involved in this effect include p38 MAPK, which has been shown to phosphorylate the GR. Other inflammatory signaling pathways have also been shown to alter GR function, including NF-κB, JNK, and signal transducers and activators of transcription 5.

RELEVANCE OF IMMUNE–CNS INTERACTIONS TO PSYCHIATRIC DISORDERS

Major Depression

For many years, major depression was seen as a quintessential example of how stress-related disorders may decrease immunocompetence. More recently, however, it has become evident that stress also activates inflammatory pathways, even while suppressing measures of acquired immunity. Not surprisingly, studies now indicate that, in addition to immunosuppression, major depression is also frequently associated with inflammatory activation. Recent research showing that proinflammatory cytokines are capable of suppressing many of the immune measures examined in major depression may provide a mechanism to account for how chronic stress–induced inflammatory activity may give rise to depression-related suppression of in vitro functional assays, such as lymphocyte proliferation and NKCA. Building upon a range of research findings, recent theorizing has turned old conceptions of MDD as an immunocompromised state on their head by suggesting that genes promoting depression may have evolved and persisted in the human genome at least in part because in ancestral environments many physiological and behavioral aspects of depression may have actually enhanced host defense against pathogens. From this perspective, the immunosuppressive changes that have been associated with depression reflect not a pathology per se (despite potential negative impacts on health in the modern world), but rather an adaptive compromise to best utilize limited metabolic energy for immune responses most likely to have promoted survival across the wide range of environmental challenges faced by hominids during evolution.

Effects on Acquired Immune (Lymphocyte) Responses

Despite heterogeneity across individual studies, many studies from the period 1980 to 2010 indicate that major depression is associated with a number of immunosuppressive changes also seen in individuals without depression but who are undergoing chronic and/or severe stress (Table 1.13–10). This is hardly surprising, given the many indices of stress system hyperactivity/dysregulation that are apparent in patients with major depression, including increased CRH and cortisol production and augmented SNS/reduced parasympathetic activity (as manifested in part by increased peripheral blood catecholamines and reductions in overall and

spectral measures of heart rate variability). Enumerative immune changes shared by major depression and chronic/severe stress include a decrease in lymphocytes, B cells, and T cells and a decrease in the ratio of $CD4^+$ to $CD8^+$ T-cell subsets. Shared functional changes include a decrease in NKCA and lymphocyte proliferation in response to nonspecific mitogens. Less is known about the effects of major depression in vivo (i.e., naturalistic immune functioning); however, available evidence suggests that depression may impair T-cell function in ways relevant to disease vulnerability. For example, one study reports that patients with major depression have a marked decrement in their ability to generate lymphocytes that respond to the herpes zoster virus. Also consistent with impaired T-cell function is the observation that depressed patients, especially those with melancholia, demonstrate impaired DTH.

Because major depression is a heterogenous condition, immune changes are not uniform across all patients. In general, immunosuppression as described above tends to be most robust in patients who are older, hospitalized, or have more severe and/or melancholic types of depression. There is also some indication that certain depressive symptoms might disproportionately account for immune alterations. For example, it should be noted that patients with primary sleep disorders who are not depressed exhibit immune alterations similar to those seen in major depression. Nevertheless, data from a meta-analysis of relevant literature indicates that age, hospitalization status, depression severity, or specific symptoms of depression cannot completely account for the association between major depression and alterations in enumerative and functional immune measures.

Effects on Innate Immune Inflammatory Responses

Taken as a whole, currently available data suggest that patients with MDD manifest all of the cardinal features of an inflammatory response including increased proinflammatory cytokines and their receptors and increased acute phase reactants, chemokines, and soluble adhesion molecules in peripheral blood and CSF. While results are not entirely consistent between studies and associations between many inflammatory markers and depression rely on findings from a limited number of studies, meta-analyses confirm reliable associations between depression and increased plasma concentrations of IL-1B, IL-6, TNF, and CRP. Because these meta-analyses only report cross-sectional associations, they do not provide insight into causal relationships between inflammation and depression. Said simply, they do not answer the question of whether inflammation causes depression, depression causes inflammation, or both cause each other. However, a recent meta-analysis shows that increased inflammation (measured as either CRP or to a lesser degree IL-6) predicts the future

development of depression, thereby increasing the probability inflammation does represent one causal pathway for depression, even in individuals who are otherwise medically healthy.

In addition to meta-analysis-confirmed increases in circulating inflammatory cytokines, recent studies show that peripheral blood gene expression profiles consistent with an M1 proinflammatory phenotype and an overrepresentation of IL-6, IL-8 and type I IFN-related signaling pathways have also been described. In addition, increased expression of a variety of innate immune genes and proteins including IL-1β, IL-6, TNF, and TLR 3 and 4 has been found in postmortem brain samples of depressed suicide victims. Moreover, genes implicated in depression based on GWAS have been linked to the immune response in particular and to antipathogen mechanisms more generally.

Interestingly, given the role of IL-6 and CRP in predicting disease outcome in both cardiovascular disorders and diabetes, as well as data indicating that inflammation may play a role in cancer, the relationship between depression and activation of the innate immune inflammatory response may provide a mechanism that explains the negative impact of depression on a number of illnesses, as well as the fact that depression predicts the later development of a range of medical illnesses including cardiovascular disease, stroke, diabetes, and dementia. Moreover, immune activation in major depression may be involved in several of the pathophysiological changes that are common in the context of depression, including bone loss, insulin resistance, cachexia, increased body temperature, and hippocampal volume loss. Interestingly, activation of innate immune responses following stress and depression may also contribute to stress- and depression-induced decreases in acquired immune (lymphocyte) responses. For example, as already discussed, administration of IL-1ra prior to stressor exposure has been found to reduce the inhibitory impact of stress on antibody production.

There are a number of potential factors that may contribute to increased innate immune responses in depressed patients. Several lines of evidence suggest that ELA may be a primary pathway by which inflammation and depression become associated. Long-term longitudinal studies have shown that individuals subjected to ELA demonstrate increased plasma concentrations of CRP in adulthood. Interestingly, individuals with ELA but no current depression were found to have higher levels of CRP than individuals with current depression but no history of ELA. Not surprisingly individuals with both ELA and current depression were found to have the highest levels of CRP. These findings suggest that ELA may program humans (as it does other mammals) to maintain high levels of inflammation in adulthood, likely as an adaptive response to the fact that dangerous childhoods typically predicted dangerous adulthoods rife with the risk of tissue damage and ensuing death from infection. Further

confirmation of the primacy of ELA in driving the increased inflammation seen in depression and other stress-related conditions comes from at least three other sources: (1) individuals with ELA but no current depression show the same elevated inflammatory cytokine (IL-6) responses to a laboratory psychological stress test (TSST) as do individuals with depression and ELA; (2) increased plasma IL-6 concentrations have been associated with diminished hippocampal volume in healthy, nondepressed adults and only depressed patients with a history of ELA show reduced hippocampal volume; and (3) IL-6 concentrations have been shown to predict depression 6 months later in adolescent females with family loading for depression and ELA, but not to predict depression in adolescent females with a family loading for depression but no ELA (in these subjects increased IL-6 actually nonsignificantly predicted *reduced* depression 6 months later).

A second factor that has received significant attention is body mass index (BMI). BMI has been reliably correlated with peripheral markers of inflammation including IL-6, likely related in part to the capacity of adipocytes to produce inflammatory cytokines and in part to the proclivity of abdominal adipocytes to recruit cytokine-producing macrophages. Relevant in this regard, an analysis of data from the Third National Health and Nutrition Examination Survey revealed that, after adjustment for a multitude of variables including BMI, there was a strong association between major depression and elevated levels of CRP in men but not women. It should be noted that obesity and ELA are not entirely separate phenomena, given that ELA strongly predicts the development of high BMI in adulthood. Nonetheless, how these two interrelated mechanisms interact in promoting the pathophysiology of depression remains to be fully elucidated.

Despite years' of evidence that depression is not a single disease state, the hope that a unifying mechanism for the disorder will be found dies hard. Once the link between inflammation and depression was clearly established, initial theorizing proposed the monolithic notion that depression was an inflammatory disease. However, convergent data from multiple sources now confirms that increased inflammation only occurs in a subgroup of depressed patients (frequently those with ELA and/or obesity). As discussed below, the finding that cytokine antagonism performs worse than placebo in depressed patients with low levels of inflammation strongly argues against the possibility that hyperactivity in inflammatory pathways is relevant to all—or perhaps even most—patients with major depression.

Bipolar Disorder

Although less extensively studied than unipolar major depression, bipolar disorder is increasingly being examined for the evidence of altered immune

system functioning. The majority of studies conducted in recent years have focused on a variety of measures of inflammation, in keeping with the general drift of the field away from assessments of acquired immune responses toward examination of innate immune system activation. While not entirely consistent, these studies—taken as a whole—suggest that patients with bipolar disorder evince many of the immune alterations frequently observed in the context of unipolar depression. Several studies have observed that bipolar patients, especially when manic, demonstrate increased plasma concentrations of the soluble receptor for the Th1-promoting inflammatory cytokine IL2 (sIL-2R). Bipolar manic patients have also been reported to have increased plasma concentrations of IL-8 and TNF-α, as well as the acute phase reactant CRP. One study reports that stimulated peripheral blood mononuclear cells (monocytes and lymphocytes) from acutely manic patients demonstrate increased production of IL-6 and TNF-α and reduced production of IL-4. Several studies indicate that treatments for mania, such as lithium, lower plasma concentrations of a number of cytokines. Interestingly, the available literature seems to suggest that patients in the manic phase of the disorder may be more likely than depressed patients to demonstrate increased inflammatory markers. It should not be surprising that mania—which seems the phenomenological opposite of depression—should be associated with increased inflammation, given that mania and depression have also been reported to show identical neuroendocrine and autonomic abnormalities, such as dexamethasone nonsuppression and increased sympathetic activity, both of which would be expected to promote inflammatory activity. Finally, examination of gene expression profiles in postmortem brain samples from patients with bipolar disorder revealed that 20 transcripts related to the immune response were upregulated, including members of the complement cascade as well as the TNF receptor family.

Schizophrenia

Inflammation produces many core symptoms of mood and anxiety disorders, but generally does not promote the types of psychotic symptoms that define schizophrenia, a disorder strongly associated in cross-sectional studies with increased inflammation. At first glance this might suggest that inflammation is less likely to cause schizophrenia than depression, but before accepting this conclusion, one consider what is meant by “cause.” While inflammatory activation in adults does not routinely produce psychosis, substantial animal data, as well as powerful circumstantial evidence in humans, suggest that increased inflammation very early in life, especially in the prenatal and perinatal periods, promotes changes in CNS development and function that significantly increase the risk of manifesting schizophrenia in adulthood. Moreover, increasing evidence suggests that

reducing inflammation in youth at high risk for schizophrenia may prevent the transition to full-blown psychosis.

Consistent with a role for in utero or early life inflammation in the pathogenesis of schizophrenia, multiple epidemiological studies have found associations between in utero exposure to a range of infectious organisms (e.g., influenza virus, *Toxoplasma gondii,* rubella virus, genital–reproductive infections) and the development of psychosis in general, and schizophrenia in particular in adult life. Although not the only potentially causative pathway, many other prenatal factors linked to an increased risk for the development of schizophrenia have also been associated with increased levels of inflammatory biomarkers during pregnancy, including obstetric complications, gestational diabetes, severe physical (e.g., famine) or psychosocial stressors, and maternal depression. Directly supporting a role for increased inflammation in the pathogenesis of schizophrenia are several longitudinal studies showing that increased maternal inflammatory biomarkers predict the development of schizophrenia in adult offspring. Although not entirely consistent in terms of the specific cytokines involved, studies have implicated elevated TNF-α and IL-8 in this association. Increased maternal levels of IL-8 in the second and third trimesters have also been found to be associated greater ventricular CSF volume in adult schizophrenia spectrum cases, as well as reduced volume of the left entorhinal cortex and right posterior cingulate.

These human findings are strongly reinforced by animal studies demonstrating that a variety of immune stimuli (i.e., stimuli that activate both antiviral and antibacterial immunity) administered to pregnant females affect CNS development and adult functioning in ways directly relevant to schizophrenia. Various prenatal immune challenge models in rodents produce an array of adult abnormalities relevant to schizophrenia, including impaired or deficient social interaction, repetitive and stereotyped behaviors, deficient working memory, impaired executive functioning and impaired sensorimotor gating as assessed by prepulse inhibition. Offspring exposed to certain highly pathogenic viruses in utero have reduced numbers of dopaminergic neurons and behavioral abnormalities in adulthood. These viral strains show strong tropism for dopaminergic neurons in vitro, with resultant activation of the lynchpin intracellular inflammatory molecule NF-κB and cellular apoptosis. Consistent with these data from rodents, exposure of pregnant macaques to low-dose LPS was found to produce an array of behavioral abnormalities in the offspring, including deficient prepulse inhibition.

While human studies linking in utero or early postnatal inflammation to the later development of schizophrenia remain correlative, animal studies strongly suggest that inflammatory mediators cause the long-term effects of immune stimulation early in development on CNS morphology and function. For example, the long-term disease-relevant effects of

administration of either IL-1β or IL-6 are abrogated if these cytokines are delivered in concert with strategies that block their biological effects (e.g., antibody against IL-6 or IL-1-receptor antagonist). Similarly, mice genetically modified to overexpress the anti-inflammatory cytokine IL-10 demonstrated significant reductions in behavioral abnormalities induced by in utero exposure to poly I:C (an immune stimuli that mimics many aspects of viral infection). Finally, rats exposed to LPS in the neonatal period had disrupted patterns of spontaneous synchronized cortical activity known to be essential for the formation of neural networks and in the regulation of neuronal survival and developmentally appropriate apoptosis (programmed cell death). These effects, which are highly relevant to both schizophrenia and autism, were mediated by TNF-α derived from activated microglia cells, consistent with increasing evidence that immune molecules in general, and microglial activity in particular, is essential for many aspects of normal neural development, including neurogenesis, neuronal migration, axon guidance, synapse formation, and activity-dependent refinement of circuits. Importantly, blockade of TNF-α activity reversed the effects of neonatal LPS on neural firing patterns and apoptosis.

Both animal and human data suggest that the long-term impact of inflammatory activation on the developing CNS does not stop at birth but persists across early life. For example, large, population-based epidemiological studies have found that individuals with a history of an autoimmune condition early in life have a 29 to 45 percent increased risk of subsequently developing schizophrenia. A history of severe early life infection (as indexed by hospitalization) increased the risk of subsequent schizophrenia by 60 percent, and the effects of autoimmunity and infection were synergistic when combined in the same individuals, increasing the risk of subsequent schizophrenia by 125 percent, with these associations remaining significant after adjustment for other potentially confounding risk factors. Animal studies are generally consistent with these human findings in suggesting that inflammatory stimulation very early in postnatal life has similar long-term effects to stimulation in utero.

The data linking various surrogates for immune activation (including measures of inflammatory biomarkers) remain circumstantial, but animal studies have been far more rigorous in demonstrating that inflammatory mediators drive the long-term effects of immune stimulation early in development on CNS morphology and function. For example, the long-term disease-relevant effects of administration of either IL-1β or IL-6 are abrogated if these cytokines are delivered in concert with strategies that block their biological effects (e.g., antibody against IL-6 or IL-1-receptor antagonist). Similarly, mice genetically modified to overexpress the anti-inflammatory cytokine IL-10 demonstrated significant reductions in behavioral abnormalities induced by in utero exposure to poly I:C. Finally, rats exposed to LPS in the neonatal period had disrupted patterns of

spontaneous synchronized cortical activity known to be essential for the formation of neural networks and in the regulation of neuronal survival and developmentally appropriate apoptosis (programmed cell death). These effects, which are highly relevant to both schizophrenia and autism, were mediated by TNF-α derived from activated microglia cells, consistent with increasing evidence that immune molecules in general, and microglial activity in particular, is essential for many aspects of normal neural development, including neurogenesis, neuronal migration, axon guidance, synapse formation, and activity-dependent refinement of circuits. Importantly, blockade of TNF-α activity reversed the effects of neonatal LPS on neural firing patterns and apoptosis.

Although inflammatory activation only rarely produces psychotic symptoms in humans (i.e., less than 1 percent of patients receiving IFN-α develop psychosis), patients with schizophrenia have been shown in numerous studies to demonstrate increases in a number of immune biomarkers. Most commonly reported (and confirmed by meta-analysis) have been elevations in IL-6, IL-2, and sIL-2R. Importantly, these alterations have been found in first-episode patients, strongly suggesting that immune activation is associated with the disease itself, rather than merely with its treatment or with other factors (such as obesity) that ensue as the disease progresses. In addition to these cytokines, a meta-analysis of over 85,000 individuals found modest, but reliable increases in plasma CRP in patients with schizophrenia compared to healthy controls, with CRP levels correlating with positive, but not negative symptoms, and with antipsychotic treatment having no effect on CRP levels.

Schizophrenia has also been associated with alterations in T-cell function and number. Summarizing results from 16 studies, a recent meta-analysis found that schizophrenia was characterized by increased absolute levels of total lymphocytes, as well as CD3 and CD4 T cells. In drug-naïve first-episode psychosis the CD4/CD8 ratio was significantly increased and the percentage of CD3 cells was significantly decreased. A shift in the Th1/Th2 cytokine profile has also been reported in schizophrenia, although results are complex. For example, a recent meta-analysis found a reduced Th1/Th2 balance in serum but an increased Th1/Th2 balance based on ex vivo stimulation testing in patients with schizophrenia. Further evidence of T-cell involvement in at least some cases of schizophrenia comes from mounting evidence that schizophrenic patients have increased autoantibodies when compared to healthy controls. For example, a meta-analysis of 81 studies found an increased prevalence of positive titers for 20 different autoantibodies in patients with schizophrenia compared to controls; however, there was marked heterogeneity between studies and correlations between autoantibodies and symptom profiles were inconsistent.

Perhaps the strongest evidence linking the immune system to the

pathophysiology of schizophrenia comes from the multiple large GWAS that have consistently identified immune genes as contributing to disease risk.

Autism

Autism has been associated with a number of immune abnormalities including unbalanced Th1/Th2 cytokine production, reduced NK and T-cell activation, and increased TNF-α and IL-12 especially in patients with comorbid gastrointestinal symptoms. In addition, patients with autism exhibit increased autoimmune-based genes including HLA-DRB1*04 and the complement C4B null allele. More recent meta-analytic studies in ASD report increased circulating concentrations of IL-1-β, IL-6, IL8, IFN-γ and monocyte chemotactic protein (MCP)-1 and reduced concentrations of TGF-β.

Still, while a convincing case can be made for a significant immune component to autism, the relationship of immune abnormalities to the neurobehavioral symptoms of the disease remains controversial. The claim that autism is triggered by childhood vaccines has not been substantiated by recent epidemiological studies, and immune-based therapies for autism have not been reliably effective. However, in support of a causative role for the immune system at least some cases of autism, a large epidemiological study from Denmark found that in utero exposure to maternal influenza infection associated with a twofold increased risk and prolonged episodes of maternal fever caused a threefold increased risk of infantile autism. Similarly, maternal autoimmune disease has been shown via meta-analysis to increase the risk of autism in children by 34 percent.

Animal studies also support the possibility that inflammatory activation either in utero or early in life may increase the risk for autism. For example, exposure of pregnant macaques to low-dose LPS resulted in significantly increased brain volume in the affected offspring, a finding repeatedly observed in human children with autism. It is increasingly recognized that immune processes play central roles in brain development. Perhaps the most striking example of the power of immune processes to both guide and disrupt neural development in ways relevant to autism comes from a study using a mouse model of maternal immune activation (MIA) that produces a variety of behaviors in offspring relevant to autism in humans. Prenatal immune stimulation reduced the number of circulating $CD4^+$ $TCR\beta^+$ Forkhead Box 3 $(Foxp3)^+$ $CD25^+$ Tregs, while increasing IL-6 and IL-17 production by $CD4^+$ T cells. Transplantation of immunologically normal bone marrow into irradiated MIA offspring reversed the autism-relevant behavioral abnormalities observed in these animals, even though the heads of these animals were protected from any direct effects of the radiation. These findings are consistent with work demonstrating that peripheral T

and monocytic cell functioning—specifically the homing of these cells to the meninges of the brain in response to psychosocial stress—is required for learning in rodents in a process that drives hippocampal BDNF production via an IL-4-dependent mechanism.

Other Disorders

Far less attention has been given to the role of immunity in general, and to inflammation in particular, in the pathogenesis of neuropsychiatric conditions other than major depression, bipolar disorder, schizophrenia and autism. Nonetheless, increasing evidence suggests that increased inflammation is a hallmark of nearly all types of psychiatric disturbance. Indeed, it is intriguing that no currently defined psychiatric condition has been reliably associated with *reduced* levels of inflammatory biomarkers.

Consistent with the ability of stress to activate inflammatory pathways (as discussed above), many studies now report that PTSD is associated with increased peripheral inflammatory biomarkers, regardless of whether the PTSD results from civilian or military trauma. Moreover, veterans with PTSD show elevated inflammatory biomarkers compared to veterans without PTSD who were exposed to similar degree of combat. PTSD has also been associated with increased ex vivo stimulated cytokine release from peripheral blood mononuclear cells and with increased intracellular NF-κ*B*. Importantly, cytokine levels (i.e., IL-6) immediately following a stressor have been found to predict the later development of PTSD symptoms, and predeployment plasma levels of CRP predict the presence/severity of PTSD symptoms following combat exposure.

There appears to be a spectrum of pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections (PANDAS), presumably as a result of antineuronal antibodies that develop following infection with group A β-hemolytic streptococcal infections, as exemplified by Sydenham chorea. Considering that children with Sydenham chorea frequently exhibit obsessive–compulsive symptoms, emotional lability, and hyperactivity, In particular, sudden onset of OCD, tics, ADHD, and other psychiatric syndromes has been characterized in children following infection with group A β-hemolytic streptococcus. Moreover, children with repeated streptococcal infections exhibit higher rates of distal choreiform movements and behaviors consistent with ADHD. Interestingly, recent data suggest that OCD more generally may also be associated with increased levels of a number of cytokines, including IL-2, IL-4, IL-6, IL-10 and TNF.

THERAPEUTIC IMPLICATIONS

The bidirectional nature of CNS–immune system interactions implies the therapeutic possibility that agents known to positively alter stress system

activity might benefit immune functioning and, conversely, that agents that modulate immune functioning may be of potential benefit in the treatment of neuropsychiatric disturbance, especially in the context of medical illness. Increasing evidence supports both hypotheses.

Antidepressants and the Immune System

Emerging data indicate that in animals and humans, antidepressants attenuate or abolish behavioral symptoms induced by inflammatory cytokine exposure. For example, pretreatment of rats with either imipramine or fluoxetine (a TCA and SSRI, respectively) for 5 weeks prior to endotoxin administration significantly attenuated endotoxin-induced decrements in saccharine preference (commonly accepted as a measure for anhedonia), as well as weight loss, anorexia, and reduced exploratory, locomotor, and social behavior. Similarly, several studies in humans suggest that antidepressants are able to ameliorate mood disturbances in the context of chronic cytokine therapies, especially if given prophylactically prior to cytokine exposure. For example, the SSRI paroxetine significantly decreased the development of major depression in patients receiving high doses of IFN-α for malignant melanoma. Following 3 months of IFN-α, 45 percent of patients receiving placebo developed major depression, compared to only 11 percent in patients receiving paroxetine. Moreover, patients on placebo had significantly higher rates of IFN-α treatment discontinuation secondary to depression when compared to patients receiving paroxetine, indicating that by blocking cytokine-induced mood disturbance, antidepressants significantly contribute to improved treatment adherence. Interestingly, in a dimensional analysis, it became apparent that paroxetine prevented major depression by blocking mood, anxiety, and cognitive symptoms. In contrast, the antidepressant was no more effective than placebo in preventing neurovegetative symptoms such as fatigue and anorexia. These findings raise the possibility that major depression in the context of chronic cytokine exposure may represent a phenomenon comprised of at least two separate syndromes mediated by two distinct pathways: an affective syndrome that is antidepressant responsive and a neurovegetative syndrome that is antidepressant resistant. The importance of both these syndromes is highlighted by the somewhat contradictory findings that, on the one hand, affective, but not somatic symptoms, are associated with depression-associated increases in mortality in hospitalized, medically ill patients and with increased viral load in patients with HIV, while on the other hand that neurovegetative, but not affective symptoms, predict the progression of atherosclerosis in patients with coronary artery disease.

Many studies indicate that, in general, antidepressants decrease immune system responsivity in ways that are likely to be of benefit in the context of immune activation. A number of antidepressants have been

shown to diminish proinflammatory cytokine production, not just from peripheral immune cells but also within the CNS, where, for example, desipramine has been reported to attenuate TNF-α release within the locus coeruleus. Rolipram, a phosphodiesterase type 4 inhibitor, has antidepressant properties and has been shown to suppress NF-κB activity, which is a primary pathway for the induction of genes that encode proinflammatory cytokine production. Concomitant with attenuating proinflammatory cytokine production, antidepressants enhance production of the anti-inflammatory cytokine IL-10. Antidepressants are also known to enhance activity in intracellular second messenger pathways (such as the cAMP cascade) known to suppress production of proinflammatory cytokines. Anti-inflammatory effects have also been observed for other modalities with antidepressant activity including ECT and exercise.

Finally, antidepressants also reverse the relative glucocorticoid resistance found in many patients with major depression, suggesting that antidepressants may diminish activity in CRH and immune pathways by restoring normal glucocorticoid-mediated feedback inhibition.

Importantly, conventional antidepressant medications act by increasing synaptic availability of monoamines and increasing neurogenesis through induction of BDNF[79]. Therefore, cytokines such as IL-β and TNF serve to undermine these activities as they decrease the synaptic availability of monoamines while also decreasing BDNF and increasing extracellular glutamate, which is not a target of conventional antidepressant therapy. These cytokine-driven effects may explain the association between inflammation and treatment nonresponse in patients with depression.

Immune Modulators and Psychiatric Illness

A logical correlate to the idea that antidepressants modulate inflammation is that agents with the capacity to directly diminish cytokine production may be of value in stress-related conditions, such as depression. Because both schizophrenia and autism appear to be reliably associated with increased inflammation, anti-inflammatory agents are of interest in these disorders, too.

In animal models, endogenous inhibitors of proinflammatory cytokines, such as the sIL-1ra, have been reported to attenuate or abolish sickness symptoms following endotoxin or cytokine administration. In addition to direct anti-inflammatory activities, sIL-1ra also blocks many of the sequelae of psychological stress in rodents. For example, direct injection of sIL-1ra blunts HPA axis responses to psychological stressors, such as restraint, and prevents stress from causing learned helplessness (a frequent animal model for depression). Of note, mice whose TNF-α receptor genes have been knocked out exhibit an antidepressant phenotype and are resistant to anxiety conditioning paradigms and virus-induced anxiety behaviors. As discussed above, inflammatory mediators likely induce

depression, at least in part, by impacting glutamatergic signaling in the CNS. Given this, it is intriguing that ketamine administration prevents more specific depressive-like symptoms following LPS, while having no effect on sickness behavior per se. However, it is not currently known whether ketamine is more effective in depressed patients with increased inflammatory biomarkers than in patients without evidence of increased immune activation.

Despite high initial hopes, recent meta-analytic explorations of the potential antidepressant effects of anti-inflammatory agents provide a mixed picture in terms of therapeutic potential. Studies of nonsteroidal anti-inflammatory agents (NSAIDs) as augmenting agents for SSRIs in major depression are either methodologically weak or have failed to be replicated and epidemiological data suggest that while some NSAIDs may be of benefit others may actually increase the prevalence of psychiatric symptoms. No evidence suggests that NSAIDs have appreciable efficacy as monotherapy for depression. In addition to these weak empirical findings, it should be noted that NSAIDs have a number of "off-target" anti- and prodepressant effects making them less than ideal as probes for testing the idea that blocking inflammation holds antidepressant promise.

Because cytokine antagonists have no appreciable "off-target" effects they are superior to NSAIDs and other less specific agents for examining the potential of anti-inflammatory strategies for the treatment of major depression. But as with NSAIDs, results of cytokine antagonist trials are complex. With one exception, their antidepressant potential has only been examined in patients with underlying disease states associated with increased inflammation (e.g., psoriasis, Crohn disease). Two meta-analyses find that cytokine antagonists (thus far only TNF antagonists have been examined) reduce depressive symptoms when compared to placebo in these inflamed patients, with the caveat that effect sizes are generally small, likely because baseline rates of depressive symptoms were low in these medically ill populations. While intriguing, these findings only point to the logical conclusion that lowering inflammation in inflamed patients improves their mood—either because inflammation was driving depressive symptoms or because the cytokine antagonists treated the underlying immune conditions caused by inflammation and the improvement in depression was secondary to reductions in primary disease burden. Left unanswered by these studies is the question of whether cytokine antagonists might work as "all-purpose" antidepressants in medically healthy patients with major depression, as would be predicted by the idea that depression is an inflammatory condition.

To answer this question, a randomized, double-blind, placebo-controlled study was conducted in medically healthy patients with treatment resistant depression using the TNF antagonist infliximab. Sixty participants were randomized to receive either three infusions of infliximab

or a salt water placebo, with infusions delivered at baseline, study week 2 and study week 6 and clinician- and self-report–based assessments of depressive symptoms and related constructs obtained at baseline (i.e., pretreatment) and at study weeks 1, 2, 3, 4, 6, 8, 10, and 12. Enrolled subjects were either off antidepressants or on a stable antidepressant regimen for at least 4 weeks prior to study entry without appreciable clinical response.

With the caveat that placebo response rates were high (50 percent), the results from the study were unequivocal. The groups were as close to each other in depressive symptom outcome as could be expected by chance (i.e., $p = 0.92$), and—in fact—placebo outperformed infliximab on a numeric basis. Thus, results of this study do not suggest that anti-inflammatory agents have general antidepressant effects. However, consistent with their efficacy for depressive symptoms in autoimmune conditions, a striking correlation was observed between baseline levels of peripheral inflammatory biomarkers (CRP and TNF) and clinical outcomes: a linear relationship was observed between increasing plasma concentrations of high-sensitivity CRP and antidepressant response to infliximab. Participants with CRP levels greater than 5 mg/L did better with infliximab than placebo, with a medium effect size of 0.41, which is similar to the efficacy of antidepressants against placebo in most studies. However, participants with hs-CRP below 5 mg/L did better on placebo than infliximab (effect size 0.82). Importantly, in participants with hs-CRP levels about 5 mg/L the response to infliximab was not the result of only impacting "sickness symptoms" such as fatigue, but resulted from a reduction in the core MDD symptoms of depressed mood and anhedonia, and from other symptoms often considered "emotional" as opposed to "somatic," including suicidal ideation and psychic anxiety.

Behavioral Interventions and Immunity

It has been known for years that psychosocial factors can mitigate or worsen the effects of stress, not only on immune functioning but also on the long-term outcomes of medical conditions in which the immune system is known to play a role. Therefore, behavioral interventions aimed at maximizing protective psychosocial factors might be predicted to have a beneficial effect, not only in terms of mitigating the effect of stress on immune functioning but perhaps also on diminishing emotional disturbances that arise in the context of immune system dysregulation (Table 1.13–10).

Two factors that have been repeatedly identified as protective against stress-induced immune alterations are social support and the ability to see stressors as being to some degree under the individual's control. In this regard, a recent study that conducted a genome-wide scan to assess gene expression activity in socially isolated versus nonisolated individuals found

that social isolation was associated with increased activation of a number of proinflammatory, cytokine-related pathways and reduced activity in anti-inflammatory cytokine pathways, as well as in the GR, which plays an important role in neuroendocrine control of inflammatory processes. Interestingly, the two types of psychotherapy most often examined in illnesses associated with immune dysregulation are group therapy, which provides social support, and CBT, which provides cognitive reframing techniques aimed at enhancing one's sense of agency (and hence control). In addition to treating depression, both cognitive behavioral and group therapy have also been shown to diminish IFN-γ production in patients with multiple sclerosis over a 12-week treatment course. Finally, although mixed results have been obtained across multiple trials, two well-designed studies have indicated that group therapy can prolong survival in cancer patients, an effect that was correlated in one study with enhancement of NKCA in patients with malignant melanoma.

Such interventions may limit the impact of stress on the immune response and may have direct effects on neuroendocrine–immune interactions. Indeed, psychological interventions such as cognitive-behavioral stress management and mindfulness-based stress reduction have been shown to alleviate psychological distress in breast cancer patients, while increasing lymphocyte proliferative responses and normalizing diurnal cortisol secretion. There is also evidence that aerobic exercise can lead to reductions in inflammatory markers in cancer survivors. The possibility that behavioral interventions such as exercise (and possibly other behavioral therapies) may attenuate cancer-related comorbidities is an important avenue for future research.

REFERENCES

Abbas AK, Lichtman AH, Pillai S. *Cellular and Molecular Immunology*. 7th ed. Philadelphia, PA: WB Saunders; 2012.

Bierhaus A, Wolf J, Andrassy M, Rohleder N, Humpert PM. A mechanism converting psychosocial stress into mononuclear cell activation. *Proc Natl Acad Sci U S A*. 2003;100:1920–1925.

Cryan JF, Dinan TG. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat Rev Neurosci*. 2012;13:701–712.

Danese A, Moffitt TE, Pariante CM, Ambler A, Poulton R. Elevated inflammation levels in depressed adults with a history of childhood maltreatment. *Arch Gen Psychiatry*. 2008;65:409–415.

Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat Rev Neurosci*. 2008;9:46–56.

Dantzer R, Walker AK. Is there a role for glutamate-mediated excitotoxicity in inflammation-induced depression? *J Neural Transm (Vienna)*. 2014;121:925–932.

Ford DE, Erlinger TP. Depression and C-reactive protein in US adults: data from the Third National Health and Nutrition Examination Survey. *Arch Intern Med*. 2004;164:1010–1014.

Frank MG, Baratta MV, Sprunger DB, Watkins LR, Maier SF. Microglia serve as a neuroimmune substrate for stress-induced potentiation of CNS proinflammatory cytokine responses. *Brain Behav Immun*. 2007;21:47–59.

Glaser R, Kiecolt-Glaser JK. Stress-induced dysfunction: implications for health. *Nat Rev Immunol*. 2005;5:243–251.

Glezer I, Simard AR, Rivest S. Neuroprotective role of the innate immune system by microglia. *Neuroscience*. 2007;147:867–883.

Irwin MR, Olmstead R, Carroll JE. Sleep disturbance, sleep duration, and inflammation: a systematic review and meta-analysis of cohort studies and experimental sleep deprivation. *Biol Psychiatry*. 2015; Epub.

*Köhler O, Benros ME, Nordentoft M, Farkouh ME, Iyengar RL, Mors O, et al. Effect of anti-inflammatory treatment on depression, depressive symptoms, and adverse effects: a systematic review and meta-analysis of randomized clinical trials. *JAMA Psychiatry*. 2014;71:1381–1391.

Koo JW, Duman RS. IL-1beta is an essential mediator of the antineurogenic and anhedonic effects of stress. *Proc Natl Acad Sci U S A*. 2008;105.751–756.

*Louveau A, Harris TH, Kipnis J. Revisiting the mechanisms of CNS immune privilege. *Trends Immunol*. 2015;36:569–577.

*Miller AH, Raison CL. The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nat Rev Immun*. 2016;16(1):22–34.

Miller GE, Chen E, Sze J, Marin T, Arevalo JM. A functional genomic fingerprint of chronic stress in humans: blunted glucocorticoid and increased NF-kappaB signaling. *Biol Psychiatry*. 2008;64:266–272.

Müller N, Weidinger E, Leitner B, Schwarz MJ. The role of inflammation in schizophrenia. *Front Neurosci*. 2015;9:372.

Musselman DL, Lawson DH, Gumnick JF, Manatunga AK, Penna S. Paroxetine for the prevention of depression induced by high-dose interferon alfa. *N Engl J Med*. 2001;344:961–966.

Nance DM, Sanders VM. Autonomic innervation and regulation of the immune system (1987–2007). *Brain Behav Immun*. 2007;21:736–745.

Pace TW, Hu F, Miller AH. Cytokine-effects on glucocorticoid receptor function: relevance to glucocorticoid resistance and the pathophysiology and treatment of major depression. *Brain Behav Immun*. 2007;21:9–19.

Pace TW, Mletzko TC, Alagbe O, Musselman DL, Nemeroff CB. Increased stress-induced inflammatory responses in male patients with major depression and increased early life stress. *Am J Psychiatry*. 2006;163:1630–1633.

Popovich PG, Longbrake EE. Can the immune system be harnessed to repair the CNS? *Nat Rev Neurosci*. 2008;9:481–493.

Raison CL, Miller AH. The evolutionary significance of depression in Pathogen Host Defense (PATHOS-D). *Mol Psychiatry*. 2013;18:15–37.

Raison CL, Rutherford RE, Woolwine BJ, Shuo C, Schettler P, Drake DF, et al. A randomized controlled trial of the tumor necrosis factor antagonist infliximab for treatment-resistant

depression: the role of baseline inflammatory biomarkers. *JAMA Psychiatry*. 2013;70:31–41.

Rook GA, Raison CL, Lowry CA. Microbiota, immunoregulatory old friends and psychiatric disorders. *Adv Exp Med Biol*. 2014;817:319–356.

Rosenblat JD, Brietzke E, Mansur RB, Maruschak NA, Lee Y, McIntyre RS. Inflammation as a neurobiological substrate of cognitive impairment in bipolar disorder: evidence, pathophysiology and treatment implications. *J Affect Disord*. 2015;188:149–159.

Savitz J, Drevets WC, Wurfel BE, Ford BN, Bellgowan PS, Victor TA, et al. Reduction of kynurenic acid to quinolinic acid ratio in both the depressed and remitted phases of major depressive disorder. *Brain Behav Immun*. 2015;46:55–59.

*Slavich GM, Cole SW. The emerging field of human social genomics. *Clin Psychol Sci*. 2013;1:331–348.

Steptoe A, Hamer M, Chida Y. The effects of acute psychological stress on circulating inflammatory factors in humans: a review and meta-analysis. *Brain Behav Immun*. 2007;21:901

*Tracey KJ. Physiology and immunology of the cholinergic anti-inflammatory pathway. *J Clin Invest*. 2007;117:289–296.

Valkanova V, Ebmeier KP, Allan CL. CRP, IL-6 and depression: a systematic review and meta-analysis of longitudinal studies. *J Affect Disor*. 2013;150(3):736–744.

Zorilla E, Luborsky L, McKay J, Roesnthal R. The relationship of depression and stressors to immunological assays: a meta-analytic review. *Brain Behav Immun*. 2001;15:199–226.

▲ 1.14 Chronobiology

MEGAN A. MILLER, M.S., KATHRYN A. ROECKLEIN, PH.D., AND IGNACIO PROVENCIO, PH.D.

Chronobiology is the study of biological time. The rotation of the Earth about its axis imposes a 24-hour cyclicity upon the biosphere. Although it is widely accepted that organisms have evolved to occupy geographical niches that can be defined by the three spatial dimensions, it is less appreciated that organisms have also evolved to occupy temporal niches that are defined by the fourth dimension of time. Much like light represents a small portion of the electromagnetic spectrum, the 24-hour periodicity represents a small time domain within the spectrum of temporal biology. A broad range of frequencies exist throughout biology, ranging from millisecond oscillations in ocular field potentials to the 17-year cycle of emergence seen in the periodic cicada (*Magicicada* spp.). While these differing periodicities all fall within the realm of chronobiology, *circadian* (Latin: *circa,* about; *dies,* day) rhythms that have a period of about 1 day are among the most extensively studied and best understood biological rhythms.

A defining feature of circadian rhythms is that they persist in the absence of time cues and are not simply driven by the 24-hour environmental cycle. Experimental animals housed for several months under constant darkness, temperature, and humidity continue to exhibit robust circadian rhythms. Maintenance of rhythmicity in a "timeless" environment points to the existence of an internal biological timing system that is responsible for generating these endogenous rhythms.

The site of the primary circadian oscillator in mammals, including humans, is the SCN, located in the anterior hypothalamus (Fig. 1.14–1). The mean circadian period generated by the human SCN is approximately 24.18 hours. Like a watch that ticks 10 minutes and 48 seconds too slowly per day, an individual with such a period gradually comes out of synchrony with the astronomical day. In slightly more than 3 months, a normally diurnal human would be in antiphase to the day–night cycle and thus would become transiently nocturnal. Therefore, a circadian clock must be reset on a regular basis to be effective at maintaining the proper phase relationships of behavioral and physiological processes within the context of the 24-hour day.

Although factors such as temperature and humidity exhibit daily fluctuations, the environmental parameter that most reliably corresponds to the period of Earth's rotation around its axis is the change in illuminance associated with the day–night cycle. Accordingly, organisms have evolved to use this daily change in light levels as a time cue or *zeitgeber* (German:

zeit, time; *geber,* giver) to reset the endogenous circadian clock. Regulation of the circadian pacemaker through the detection of changes in illuminance requires a photoreceptive apparatus that communicates with the central oscillator. In humans, as in other mammals, this apparatus is known to reside in the eyes, because surgical removal of the eyes renders an animal incapable of resetting its clock in response to light.

The circadian clock drives many rhythms, including rhythms in behavior, core body temperature, sleep, feeding, drinking, and hormonal levels. One such circadian-regulated hormone is the indoleamine, melatonin. Melatonin synthesis is controlled through a multisynaptic pathway from the SCN to the pineal gland. Serum levels of melatonin become elevated at night and return to baseline during the day. The nocturnal rise in melatonin is a convenient marker of circadian phase. Exposure to light elicits two distinct effects on the daily melatonin profile. First, light acutely suppresses elevated melatonin levels, immediately decreasing them to baseline levels. Second, light shifts the phase of the circadian rhythm of melatonin synthesis. Because melatonin can be assayed easily, and reflects any perturbation of the clock, it can be used as an output measure to study the regulation of the central circadian pacemaker.

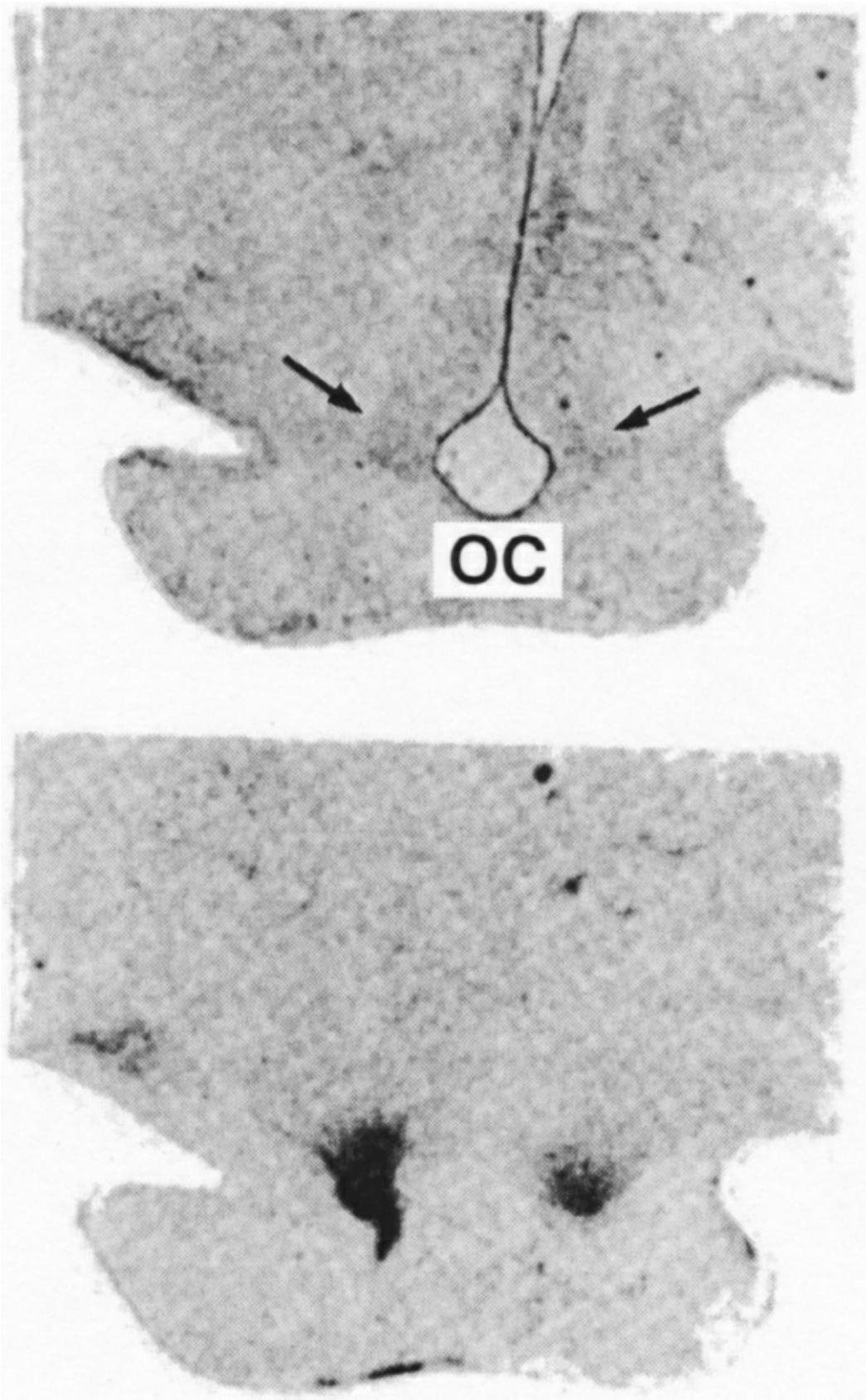

FIGURE 1.14–1. The human suprachiasmatic nucleus. **Top:** Nissl stain of section through the human hypothalamus. The suprachiasmatic nuclei are indicated by arrows and are located dorsal to the optic chiasm (OC). **Bottom:** Autoradiograph of same section. Specific binding of a radioiodinated analog of melatonin is indicated by the darkening of the

suprachiasmatic nuclei. (From Reppert SM, Weaver DR, Rivkees SA, Stopa EG. Putative melatonin receptors in a human biological clock. *Science.* 1988;242:78, with permission.)

Taken together, the circadian axis of mammals can be divided into three distinct functional components: (1) a master pacemaker situated in the SCN, (2) a photoreceptive input to the SCN that originates in the eye, and (3) the myriad of rhythmic outputs that provide insight into the clockwork of the circadian pacemaker.

CIRCADIAN PACEMAKERS

Anatomy

The mammalian circadian system is organized as a hierarchy of pacemakers. The SCN is the master oscillator that orchestrates a multitude of oscillators outside of the SCN. These peripheral oscillators are found in a wide range of peripheral tissues including the kidney, liver, lung, and other sites in the brain. Most of the current understanding of the SCN and its peripheral oscillators is derived from rodent studies.

The SCN are small, paired, hypothalamic structures situated immediately dorsal to the optic chiasm. They were recognized as the site of the primary circadian pacemaker because lesions in the ventral hypothalamus that encompassed the SCN rendered rodents behaviorally arrhythmic. Transplantation of SCN tissue from mutant hamsters that expressed abnormally short circadian periods into the brains of SCN-lesioned host hamsters with normal prelesion circadian periods results in a transfer of the abnormally short period. This transfer of a distinguishing parameter of the circadian clock indicates that the SCN is a true biological pacemaker and not simply a neural relay for a rhythm generator located elsewhere in the brain.

Metabolically, the SCN show peak activity during the subjective day. This increased level of metabolism is paralleled by the increased electrophysiological activity evident from brain slice recordings. SCN neurons that are isolated and maintained in culture for several days also continue to show approximately 24-hour rhythms in action potential frequency (Fig. 1.14–2).

This observation indicates that the rhythmicity of the SCN is not an emergent property of the system but rather an inherent feature of individual SCN neurons. Molecular studies have confirmed that the oscillatory machinery of the SCN is indeed contained within the individual neurons. It is likely, however, that the general output of the SCN is a result of coupling between individual cellular oscillators, resulting in a coordinated rhythmic signal. The prevailing view of SCN oscillator organization is that the individually oscillating neurons are largely synchronized and the composite output of the SCN reflects the mean phase of these oscillators. Recent studies, however, have shown that discrete

phase groupings of oscillators exist. The relative contributions of these "phase ensembles" to the overall rhythmic output of the SCN are likely to be modulated by entraining agents such as light. In addition to the variable contributions of the ensembles to the global output of the SCN, the amplitude and relative phase of the ensembles may also be modified by entraining agents. The potential to manipulate so many parameters of these sets of oscillatory neurons affords a tremendous degree of flexibility with respect to clock-resetting mechanisms. Furthermore, the relative phasing of heterogeneous oscillator groups provides a strategy by which seasonality can be transduced into a biological signal.

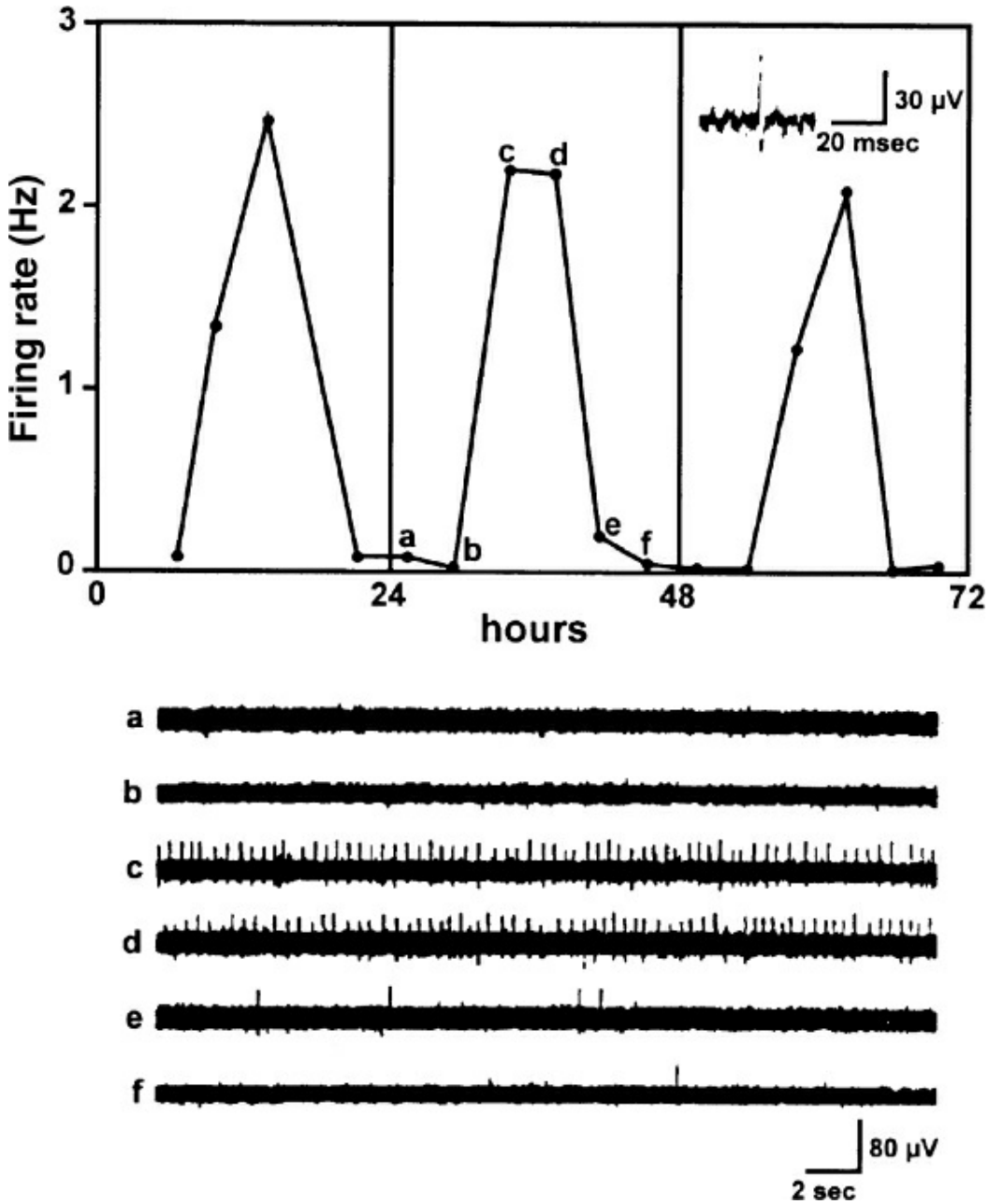

FIGURE 1.14–2. Circadian rhythm in the firing rate of an individual suprachiasmatic nucleus (SCN) neuron. **Top:** Firing rate of an individual action potential. **Bottom:** Extracellular recording traces corresponding to the labeled times indicated during day 2 in

the top panel. (From Welsh DK, Logothetis DE, Meister M, Reppert SM. Individual neurons dissociated from rat suprachiasmatic nucleus express independently phased circadian firing rhythms. *Neuron.* 1995;14:697, with permission.)

The neurons of the SCN are among the smallest neurons in the entire brain. They possess short dendrites that are not extensively branched. A consequence of these cellular dimensions is the high packing density of the nucleus. Virtually every neuron in the SCN is immunopositive for the inhibitory neurotransmitter GABA. Subdivisions of the SCN have also been defined according to immunohistochemical and neural tract tracing criteria. Perhaps the most obvious anatomical subdivision is the *core* of the SCN, defined by the presence of calbindin-positive neurons. The remainder of the SCN that surrounds the core is considered the *shell*. Discrete functions have not been firmly assigned to subdivisions of the SCN, but the afferent and efferent projections to and from these subdivisions are beginning to provide insights into their putative functions.

Afferent Projections

The retinohypothalamic tract is the primary afferent input to the SCN. It originates in the retina from a small subset of retinal ganglion cells and innervates the entire volume of the SCN, although the specific subregions of the SCN that are most heavily innervated vary among different species of mammals.

The excitatory neurotransmitter glutamate is the primary neurotransmitter of the retinohypothalamic tract, with PACAP modulating glutamate's effect in the SCN. Glutamate receptor antagonists can block the effects of light on the circadian axis, illustrating the importance of this neurotransmitter in conveying photic information from the retina to the SCN.

The SCN also receives afferents from the ipsilateral intergeniculate leaflet (IGL), a subnucleus of the lateral geniculate complex. The IGL, in turn, receives input directly from the retina, thus providing a secondary indirect pathway from the retina to the SCN. NPY is the predominant transmitter of the IGL-to-SCN projection. Although the function of the IGL is not well established, it is purported to be involved in encoding environmental luminance. Other, less understood projections to the SCN also are known to exist.

Most prominent among these is a distinct serotonergic projection from the midbrain raphe. Serotonin (5-hydroxytryptamine [5-HT]) is known to modulate light's effect on SCN function. Systemic administration of a 5-HT_{1B} receptor agonist before the application of a light pulse attenuates light-induced phase shifts of circadian locomotor activity in hamsters and mice. Light-induced expression of *Fos* (an immediate early gene) within the SCN is also attenuated with this agonist. 5-HT_7, a novel serotonin receptor

subclass, and has been implicated in mediating serotonin's modulation of the glutamatergic input to the SCN. Electron microscopy has been used to localize 5-HT_{1B} and 5-HT_7 receptors in pre- and postsynaptic membranes within the SCN. Behavioral data, in conjunction with the pharmacological and anatomical findings, highlight the importance of serotonin in regulating the photic information reaching the SCN via the retinohypothalamic tract. It has been hypothesized that serotonin may adjust the gain of the circadian system's response to light.

Efferent Projections

Most SCN efferent projections remain within the limits of the hypothalamus. The best-studied projection originating in the SCN is the multisynaptic projection to the pineal gland. Axons of inhibitory GABAergic SCN neurons traverse the hypothalamus dorsally to the autonomic division of the PVN. The tonic activity of the PVN is suppressed during the day, when the SCN firing rate is high, and is uninhibited during the night, when the SCN is quiescent. The PVN sends a descending glutamatergic projection through the medial forebrain bundle to the intermediolateral cell column of the spinal cord at the upper thoracic levels (T1 and T2). Cholinergic preganglionic sympathetic neurons propagate this signal by synapsing on adrenergic postganglionic sympathetics within the superior cervical ganglion. These postganglionic fibers finally innervate pinealocytes to stimulate melatonin synthesis. Norepinephrine release from the terminals of these fibers increases intracellular cAMP levels and, consequently, the activity of melatonin synthetic pathway. Melatonin is not stored or released via a secretory pathway. Its lipophilic nature permits it to passively diffuse through membranes. Thus, the release of melatonin is directly related to its rate of synthesis.

The action of norepinephrine is mediated through β- and α_1-adrenergic receptors. β-Adrenergic receptors stimulate the production of cAMP, whereas the α_1-adrenergic receptors potentiate the action of β-adrenergic receptors. The ultimate outcome of this efferent pathway is that increased melatonin synthesis occurs when SCN activity is low. This antiphasic relationship is established by the GABAergic sign-changing synapse at the PVN. Melatonin receptors localized to the SCN are likely to provide a feedback mechanism by which the antiphasic relationship between the SCN and the pineal gland is maintained and possibly reinforced.

A second, less understood efferent pathway from the SCN plays an important role in the control of cortisol. Systemic cortisol levels increase in response to stress. However, these levels also have a strong circadian component, being highest in the early morning in humans. Peak cortisol levels occur at approximately 6 AM to 8 AM, just as melatonin levels approach baseline. This stress-independent component was identified

through SCN-lesion studies in rodents that abolished circadian rhythms of cortisol but left the acute response to stress intact. An inhibitory GABAergic projection from the SCN to the hypothalamic PVN is the first element in the neural pathway regulating rhythmic cortisol levels. Axons of this projection synapse on the parvicellular neurons of the PVN that contain CRH. The terminals of these CRH-containing neurons reside in the median eminence, where they release CRH into the pituitary portal system and stimulate the release of adrenocorticotrophic hormone (ACTH) from the adenohypophysis. ACTH, in turn, acts on the zona fasciculata of the adrenal gland to release cortisol.

A secondary inhibitory pathway to the PVN via a vasopressinergic projection through the dorsomedial hypothalamus has been implicated in cortisol regulation. Analogous to the circuit controlling pineal melatonin, the circuit regulating cortisol contains a single inhibitory synapse. Accordingly, one would expect a circadian cortisol secretion profile similar in phase and shape to that of pineal melatonin. This is not the case, suggesting the presence of other factors that may be involved in shaping the circadian profile of cortisol. First among these is the mechanism of information transfer. Although the SCN–pineal circuit is exclusively neural, the SCN–adrenal circuit involves the stimulated release of hormones that are subject to the vagaries of diffusion and transport through the circulation. Second, the sensitivity of the adrenal cortex to ACTH also exhibits a circadian variation. Finally, it has been proposed that cortisol itself may provide a feedback on the brain to inhibit CRH and ACTH production.

Molecular Clockwork

As mentioned previously, isolated SCN neurons can generate circadian rhythms. For decades, however, the lack of knowledge regarding the inner clockwork of the circadian pacemaker forced investigators to treat the SCN as a "black box." Since the mid-1980s, progress in understanding rhythmic biochemical processes has led to advances in the identification and characterization of the molecular gears of circadian pacemakers. A cohort of principal clock genes has been identified in mammals since 1997. Many of these molecular components were initially discovered in the fruit fly, *Drosophila melanogaster,* leading to the discovery of mammalian orthologs. The basic architecture of the circadian clockwork is generally conserved among species of the animal kingdom; however, some of the clock genes have been duplicated in the mammalian genome. These duplications have added another level of complexity and the possibility of partially redundant functions.

The molecular clockwork of the master clock in the SCN is virtually identical to that of the peripheral oscillators. The SCN are so small that their size precludes most biochemical analyses. However, investigators

have been able to characterize clock proteins from abundant clock-containing peripheral tissues, such as liver, to understand the workings of the central clock in the SCN. From these and other studies, the following concepts are now established.

The mammalian circadian clockwork consists of interacting positive and negative transcriptional and translational feedback loops (Fig. 1.14–3). The expression of multiple homologs of the *Drosophila* period gene and cryptochrome genes are positively regulated by the binding of Clock-Bmal1 (Bmal1 is also known as Mop3) heterodimers to E-box enhancers in the promoters of these genes. The products of the *Per* and *Cry* genes translocate back into the nucleus and repress their own transcription. This series of events constitutes the negative feedback limb of the core oscillation. Interestingly, light directly influences the molecular clock via its impact on the expression of Per2, which has either phase advancing or delaying effects on the feedback loop depending on timing within the cycle described above.

In addition to activating the transcription of the *Per* and *Cry* genes, the Clock-Bmal1 complex also activates expression of the orphan nuclear receptor gene *Rev-Erb*a. The gene product of *Rev-Erb*a, in turn, translocates into the nucleus and represses transcription of the *Bmal1* gene through *Rev-Erb/Ror* response elements in the *Bmal1* gene promoter. Bmal1 subsequently heterodimerizes with Clock and again activates expression of the *Per, Cry,* and *Rev-Erb*a genes. This derepression (or activation) of the *Bmal1* gene, subsequent heterodimerization with Clock, and activation of the *Per, Cry,* and *Rev-Erb*a genes constitutes the positive feedback limb of the core oscillator.

The interactions of clock proteins and the translocations of these proteins between cellular compartments are tightly regulated by posttranslational modifications. For example, phosphorylation of some of the Per proteins by casein kinase Iε or δ is important for their translocation into the nucleus or their degradation in the cytoplasm. Other kinases, and presumably phosphatases, are emerging as critical regulators of the circadian molecular clockwork. More global regulatory mechanisms, such as histone acetylation or phosphorylation, are also likely to control the rhythmic expression and function of clock genes.

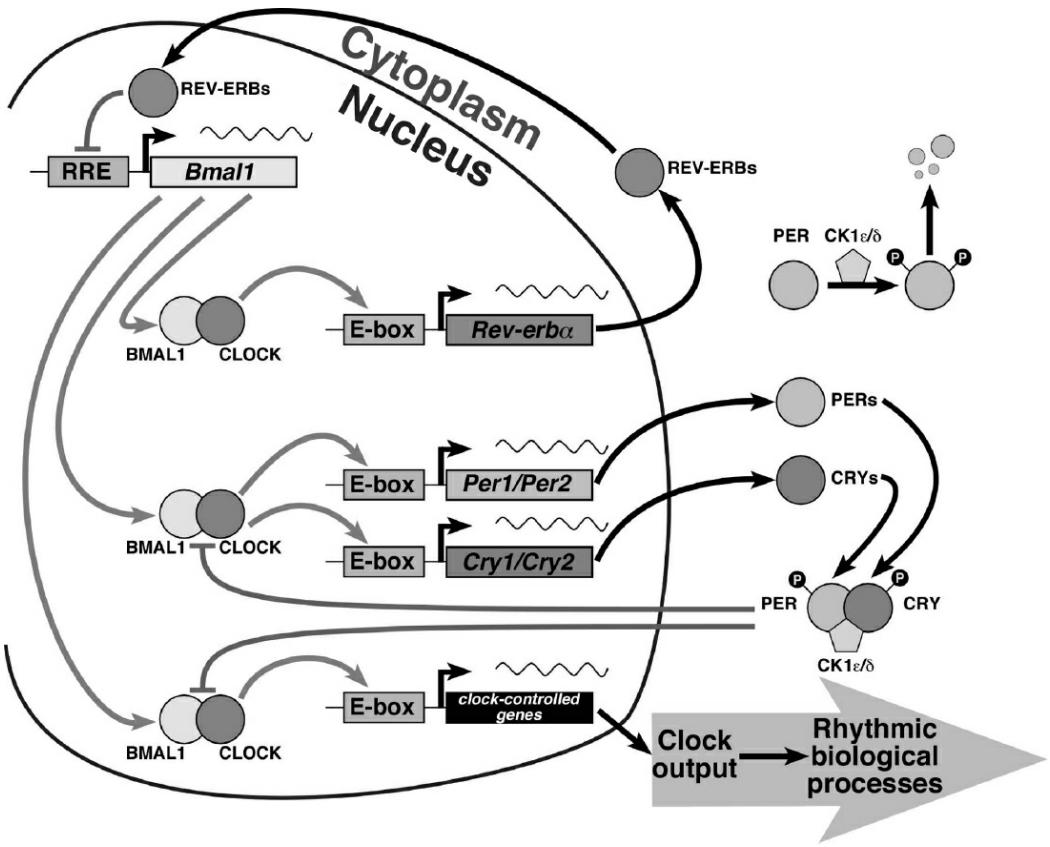

FIGURE 1.14–3. Transcriptional–translational feedback loops that constitute the molecular clockworks of the mammalian circadian clock. (Adapted from Ko CH, Takahashi JS. Molecular components of the mammalian circadian clock. *Human Mol Genet*. 2006;15:R271.)

RESETTING THE CIRCADIAN CLOCK

Sensory Parameters

In humans and other mammals, light perceived through the eyes is the most effective agent for entraining (synchronizing) the circadian system to the 24-hour day. Bilateral removal of the eyes renders an individual incapable of resetting the circadian clock, indicating that the photosensitive apparatus necessary for resetting must be ocular. However, several lines of evidence suggest that this apparatus is distinct from the rod and cone photoreceptors that are required for vision.

First, the photic sensitivities of the visual and circadian systems are quite different. The visual system can be activated by intensities of light ranging from dim starlight to bright daylight. This dynamic range represents approximately 14 log units of light intensity measured in photons per second per square centimeter. The dynamic range of the circadian system is only 3 log units, and the activation threshold is much higher than that of the visual system. Additionally, the circadian system requires light stimuli of much longer duration to activate clock resetting than that required by the visual system to construct images.

These differences in activation parameters are consistent with the

differences in the photoreceptive tasks performed by these functionally distinct systems. The principal task of the visual system is to construct a spatiotemporal representation of the environment. In this sense, the eye functions like a movie camera, acquiring a series of still shots that the brain can interpret as a dynamic visual scene. Thus, information about the relative positions of different stimuli within the visual field must be maintained throughout processing. The ability to detect motion within the visual field also requires a relatively fast integration time analogous to a fast International Standards Organization (ISO) rating for a roll of 35-mm film. These spatial and temporal requirements can be satisfied by a fine two-dimensional array of narrow-capture, highly sensitive, photoreceptive elements, such as the photoreceptor layer of the retina that contains the rod and cone photoreceptors.

By contrast, the task of the photoreceptive input to the circadian system is the measurement of ambient illumination. In essence, the circadian photoreceptive system must function as a light meter rather than a camera. The requirements of a light meter are different than those of a camera. Spatial information is not important and may, in fact, confound the system. Luminous point sources of light, such as the moon, could prove confusing for narrow-capture photoreceptive elements. If, however, the system used relatively insensitive, broad-capture photoreceptive elements capable of integrating large sectors of visual space, then the relative contribution of the moon's irradiance to the total ambient irradiance would be minimized and thus would not be mistaken for a daytime light level. Theoretically, a fine, two-dimensional array of narrow-capture photoreceptive elements could serve in this capacity if the output of the array were averaged. Alternatively, an anatomically distinct, coarse array of relatively few broad-capture photoreceptive elements would be optimal for wide spatial integration at a reduced absolute sensitivity. Lightning could also potentially baffle the circadian system. Ambient levels of illumination achieved by lightning can equal those levels of daylight. However, the circadian system's insensitivity to stimuli of short duration essentially filters out this source of photic noise. Taken together, the spatial and temporal stimulus parameters required to activate the circadian photic input system ensure that only the relevant stimulus, that is, the change in ambient irradiance associated with the day:night cycle, is conferred to the central circadian pacemaker.

The different sensory demands of the circadian and visual systems raise the possibility that a novel photoreceptive apparatus subserves light-mediated entrainment of the circadian system to the day–night cycle. Visually blind rodents that have a genetically induced loss of rods and cones remain capable of light-mediated circadian clock resetting. Paradoxically, the loss of rods and cones has no effect on the sensitivity of the circadian system to light, despite the fact that these animals are

incapable of forming images. A similar situation has been observed in blind humans. A subset of blind individuals retains the ability to photically regulate the rhythmic synthesis of melatonin.

Some blind individuals report no cognitive visual perception, show no electrophysiological evidence for ocular light detection as determined by electroretinogram analysis, and exhibit no pupillary light response. However, some of these individuals continue to show an acute suppression of nocturnal melatonin and a phase shift in the circadian rhythm of melatonin production. All too frequently, blind individuals are bilaterally enucleated and fitted with prosthetic eyes for reasons ranging from susceptibility to ocular infections to aesthetics. Perhaps some of these decisions should be reconsidered in light of the fact that the eyes of these individuals may retain a function in clock resetting despite being useless for vision.

NOVEL CLASS OF RETINAL PHOTORECEPTOR

Intrinsically Photosensitive Retinal Ganglion Cells

The findings from retinally degenerate animal models and blind humans indicate that photoreceptors other than rods and cones are likely to be involved in the photoregulation of the circadian axis. Blue wavelengths of light most efficiently suppress melatonin in humans; however, the spectral profile of melatonin suppression does not match that of any of the photopigments found in human rods or cones. A small subset of rodent retinal ganglion cells has been shown to be intrinsically photosensitive. The spectral sensitivity of these cells matches the spectral sensitivity of the circadian system. Most compelling, however, is that the intrinsically photosensitive retinal ganglion cells project directly to the SCN. They also project to the IGL and the olivary pretectal nuclei, other brain structures involved in the interpretation of illuminance information. The intrinsically photosensitive cells contain melanopsin, a photopigment initially discovered in the pigmented skin cells (melanophores) of tadpoles and subsequently identified in human and mouse retinas. The anatomy and physiology of melanopsin-containing retinal ganglion cells are consistent with the previously described characteristics expected of cells involved in illuminance detection. Namely, these cells are few in number. They represent 1 to 2 percent of all of the retinal ganglion cells in the rodent retina. These cells are also distributed over the entire retinal expanse. The dendritic arbors of melanopsin-containing retinal ganglion cells are vast, with arbors in the mouse retina ranging from 400 to 500 μm in diameter.

Melanopsin itself is localized to the plasma membrane of the cell body, axons, and dendrites. The size of the receptive fields of these cells matches the size of the dendritic arbors, indicating that the entire arbor has the ability to initiate phototransduction and therefore is capable of spatially

integrating large sectors of the visual field. The average mouse eye is approximately 3 mm in diameter. Therefore, a photoreceptor with a receptive field diameter of 400 to 500 μm is able to spatially integrate 15 to 20 degrees of visual space. By comparison, the diameter of the full moon at its highest point in the sky is approximately equivalent to 1 degree of the human visual field. Melanopsin-containing retinal ganglion cells are clearly broad-capture photoreceptors. Furthermore, because the dendritic arbors overlap, these cells form a photoreceptive net in the inner mammalian retina. This complex of cells represents the coarse array of photoreceptive elements expected of an irradiance detector. In addition, the activation parameters of these cells parallel the parameters observed for the circadian system as a whole. The melanopsin-containing ganglion cells are significantly less sensitive to light compared to the rod and cone photoreceptors of the visual system, and they require light stimuli of relatively long duration to be activated. Finally, these cells are maximally sensitive to wavelengths of light similar to those required to acutely suppress nocturnal melatonin levels in humans.

Function

Although the anatomy and physiology of melanopsin retinal ganglion cells are highly suggestive that these cells function as circadian photoreceptors, recent studies provide the most compelling evidence. Mice with a targeted disruption of both copies of the melanopsin gene show profound deficits in their ability to phase-shift circadian locomotor rhythms in response to pulses of light. These deficits were observed at all irradiances tested (Fig. 1.14–4). Thus, the photopigment melanopsin and, presumably, the retinal ganglion cells containing melanopsin are required for normal photic regulation of circadian rhythms. Perhaps the most surprising aspect of these "knock-out" studies is that disrupting both copies of the melanopsin gene does not completely abolish light-induced circadian phase shifting; some capacity for phase shifting remains. The photoreceptors mediating this residual sensitivity are likely to be the rods or the cones; however, one cannot exclude the possibility that an unrecognized class of ocular photoreceptor fulfills this role.

To test the contribution of rod and cone photoreceptors to photoentrainment, melanopsin-null mice have been crossed with mice lacking rods and cones. The progeny of this cross that are rodless, coneless, and melanopsin-null are incapable of photoentrainment, even at high irradiances of ambient light (Fig. 1.14–5). Other nonvisual photophysiology is also abolished in these mice, such as the photoregulation of the melatonin biosynthetic pathway, the pupillary light response, and the acute light-induced inhibition of activity. From these studies, it can be concluded that at least partial functional redundancy exists for nonvisual photoreception between rods, cones, or both, and the melanopsin-

containing retinal ganglion cells. It should be the relative contribution of the visual photoreceptors versus that of the melanopsin retinal ganglion cells appears to be different among the various nonvisual responses. For example, melanopsin plays a rather significant role in the phase shifting of circadian locomotor activity; however, the pupillary light response is relatively insensitive to the loss of melanopsin. Importantly, the complete loss of photic responses in melanopsin-null mice lacking rods and cones demonstrates that no additional photopigments, such as cryptochromes, are required for nonvisual photic signaling. Thus, it appears that multiple photoreceptor systems subserve nonvisual photoreception, a phenomenon observed across phylogeny.

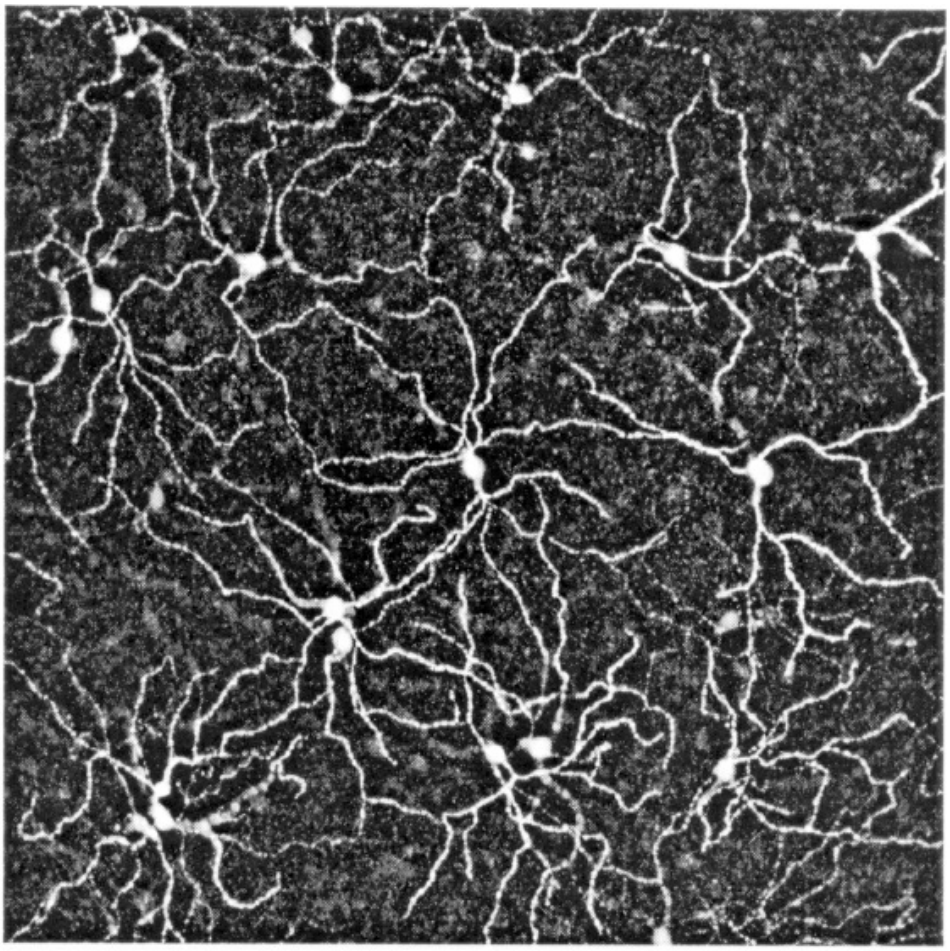

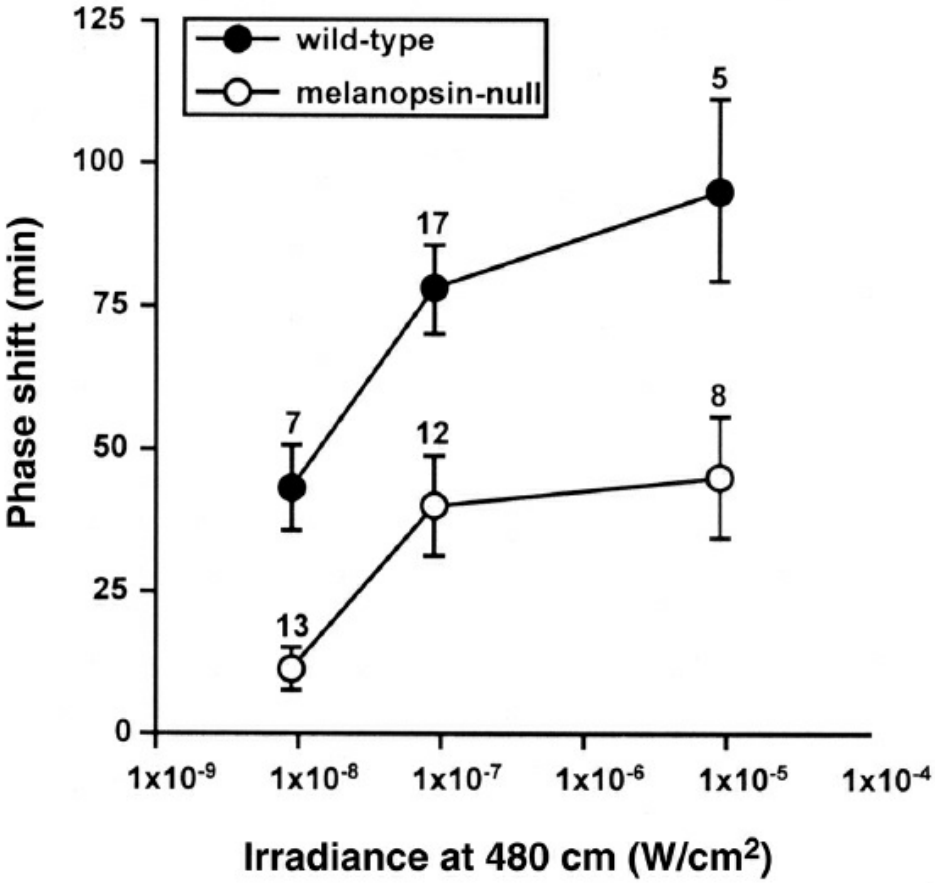
wild-type
melanopsin-null
Phase shift (min)
125
100
75
50
25
0
5
17
7
12
8
13
1x10-9
1x10-8
1x10-7
1x10-6
1x10-5
1x10-4
Irradiance at 480 cm (W/cm2)

FIGURE 1.14–4. Melanopsin in the mouse. **Top:** Retinal ganglion cells in a flat-mounted mouse retina, labeled by indirect immunofluorescence with an antibody against melanopsin. Note the *photoreceptive net* formed by the overlapping dendritic arbors. (Courtesy of Dr. Ana Castrucci.) **Bottom:** Fluence–response relationship in wild-type and melanopsin-null mice in response to a 15-minute pulse of blue (480-nm wavelength) light at circadian time 15 hours. Melanopsin-null mice exhibited attenuated phase shifting of circadian locomotor rhythms relative to wild-type siblings at all irradiances tested. (Adapted from Panda S, Sato TK, Castrucci AM, et al. Melanopsin [*Opn4*] requirement for normal light-induced circadian phase shifting. *Science.* 2002;298:2213.)

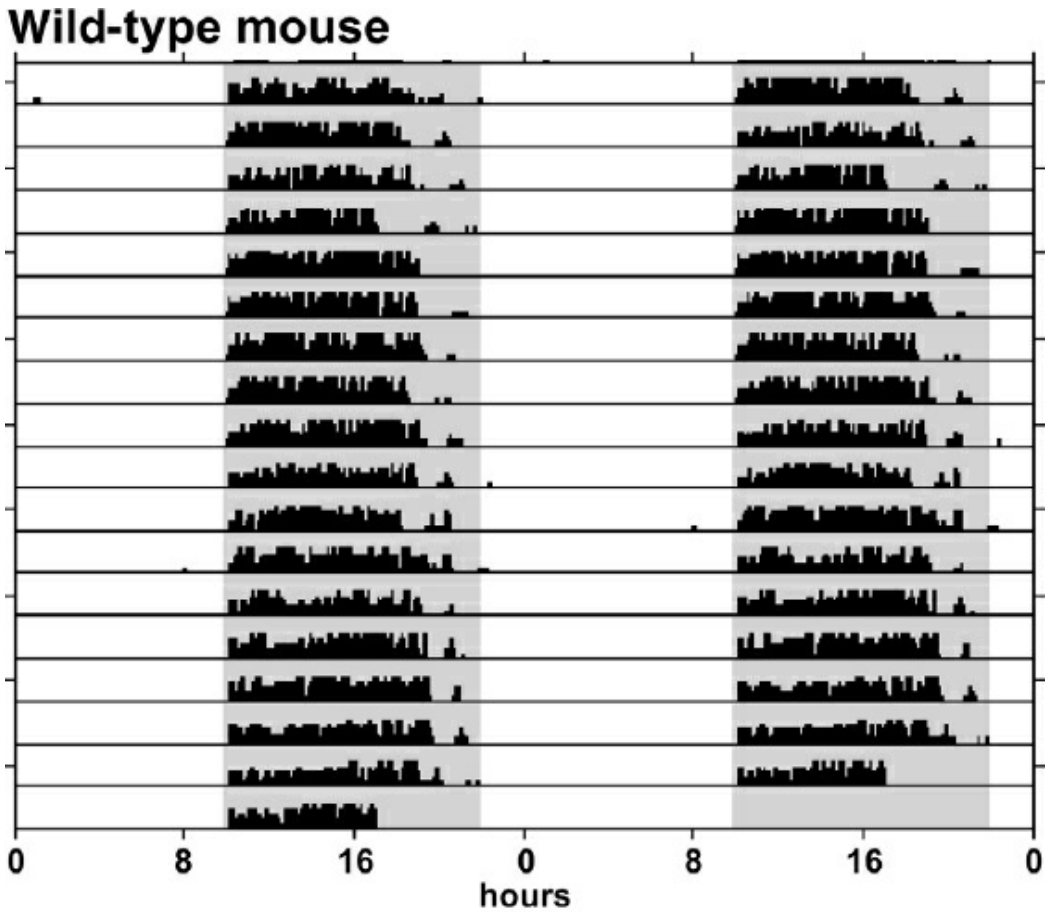
Wild-type mouse
0
8
16
0
8
16
0
hours

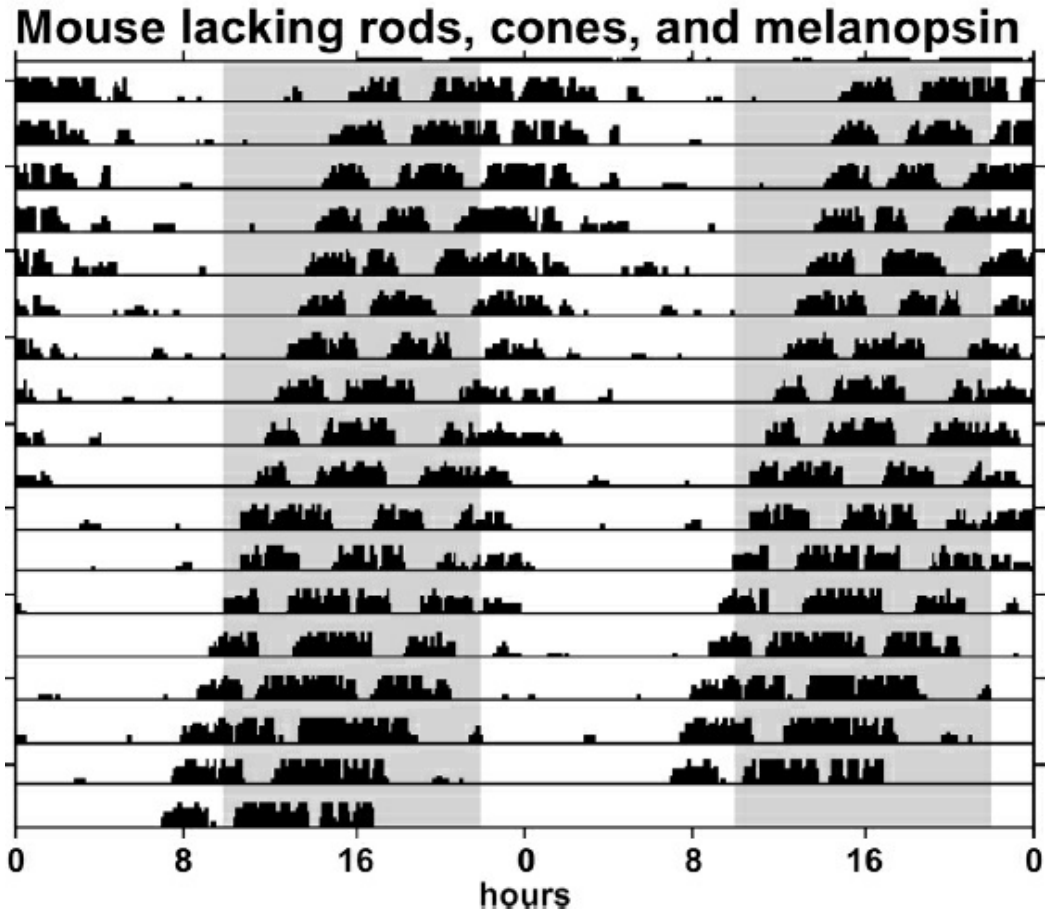
Mouse lacking rods, cones, and melanopsin
0
8
16
0
8
16
0
hours

FIGURE 1.14–5. Wheel running activity records of a wild-type mouse and a mouse lacking rods, cones, and melanopsin. **Top:** Representative double-plotted activity record of a wild-type mouse under entraining conditions of 12 hours of white light (800 lux) and 12 hours of darkness (*gray box*). **Bottom:** Representative double-plotted activity records of a mouse lacking rods, cones, and melanopsin under identical entraining conditions. Whereas wild-type mice consolidate their activity to the dark phase and the time of activity onset is coincident with the light to dark transition, the mice lacking rods, cones, and melanopsin continue to free run with an intrinsic period length of less than 24 hours. (Adapted from supplementary data from Panda S, Provencio I, Tu DC, Pires SS, Rollage MD. Melanopsin is required for non-image-forming photic responses in blind mice. *Science.* 2003;301:525.)

The dawn of air travel has introduced society to the phenomenon of jet lag, a dramatic example of circadian desynchrony. Simply stated, jet lag is the condition of one's circadian clock being desynchronized from the local time. Shift work can also cause circadian desynchrony. The invention of artificial lighting has permitted the manufacturing and service industries to work around the clock. As a result, shift workers are constantly experiencing the effects of circadian desynchrony as they try to entrain to an ever-changing light–dark cycle. Some deleterious effects of shift work include elevated stress, deficits in alertness, decreased cognitive function, and gastric distress. Additionally, individuals may experience a phenomenon termed *social jet lag,* in which those who naturally exhibit slightly delayed sleep–wake patterns are able to sleep on their normal patterns on the weekends but are forced to wake earlier during work days, replicating the desynchronization of one's circadian clock with the social environment. Similar to shift work and jet lag, social jet lag has been associated with several psychological and biological outcomes including increased risk of depression and obesity. Although no therapy currently exists for jet lag or shift work, an efficacious treatment ultimately must involve the appropriate resetting of the clock. Such a treatment may include timed administration of light stimuli of spectrally optimal wavelengths. A more complete understanding of how melanopsin-containing retinal ganglion cells convert such stimuli into neural signals may present investigators with pharmacological entry points against which chronobiotic drugs can be designed.

Extra SCN Zeitgebers

In addition to photic input to the SCN, several other zeitgebers can synchronize rhythms in the periphery. Restricted access to food has been shown to restore circadian rhythms in peripheral tissues in animals with lesioned SCN, suggesting that food has the ability to entrain circadian rhythms in the absence of the master clock. Similarly, administration of methamphetamine in SCN-lesioned mice entrained peripheral clock gene expression and, when administered to intact animals, methamphetamine created phase desynchrony in the periphery relative to the master clock.

Together, this indicates that methamphetamine acts as a synchronizing stimulus in the periphery and may compete with other stimuli leading to desynchrony if administered in intact SCN individual. Lastly, time-specific restricted wheel access has been shown to shift heart rate and clock gene expression rhythms in mice, implicating exercise as another zeitgeber. Given the entrainment value of these zeitgebers, especially food and exercise, they may provide novel treatment targets in efforts to stabilize or shift circadian rhythms.

SLEEP AND CIRCADIAN RHYTHMS

Sleep Regulation

Restful consolidated sleep is most appreciated when sleep disturbances are experienced. Sleep is the integrated product of two oscillatory processes. The first process, frequently referred to as the *sleep homeostat,* arises from the accumulation and dissipation of sleep debt. The biological substrates encoding sleep debt are not known, although adenosine is emerging as a primary candidate neuromodulator of the sleep homeostat. The second oscillatory process is governed by the circadian clock and controls a daily rhythm in sleep propensity or, conversely, arousal. The fact that these two rhythms can be forced to desynchronize establishes that they are two separate processes.

The circadian cycle in arousal (wakefulness) steadily increases throughout the day, reaching a maximum immediately before the circadian increase in plasma melatonin (Fig. 1.14–6). Arousal subsequently decreases to coincide with the circadian trough in core body temperature. Experiments imposing forced sleep schedules throughout the circadian day have shown that an uninterrupted 8-hour bout of sleep can only be obtained if sleep is initiated approximately 6 hours before the temperature nadir. This nadir typically occurs at approximately 5:00 AM to 6:00 AM. In healthy individuals, initiating sleep between 11:00 PM and 12:00 AM affords the highest probability of getting 8 solid hours of sleep.

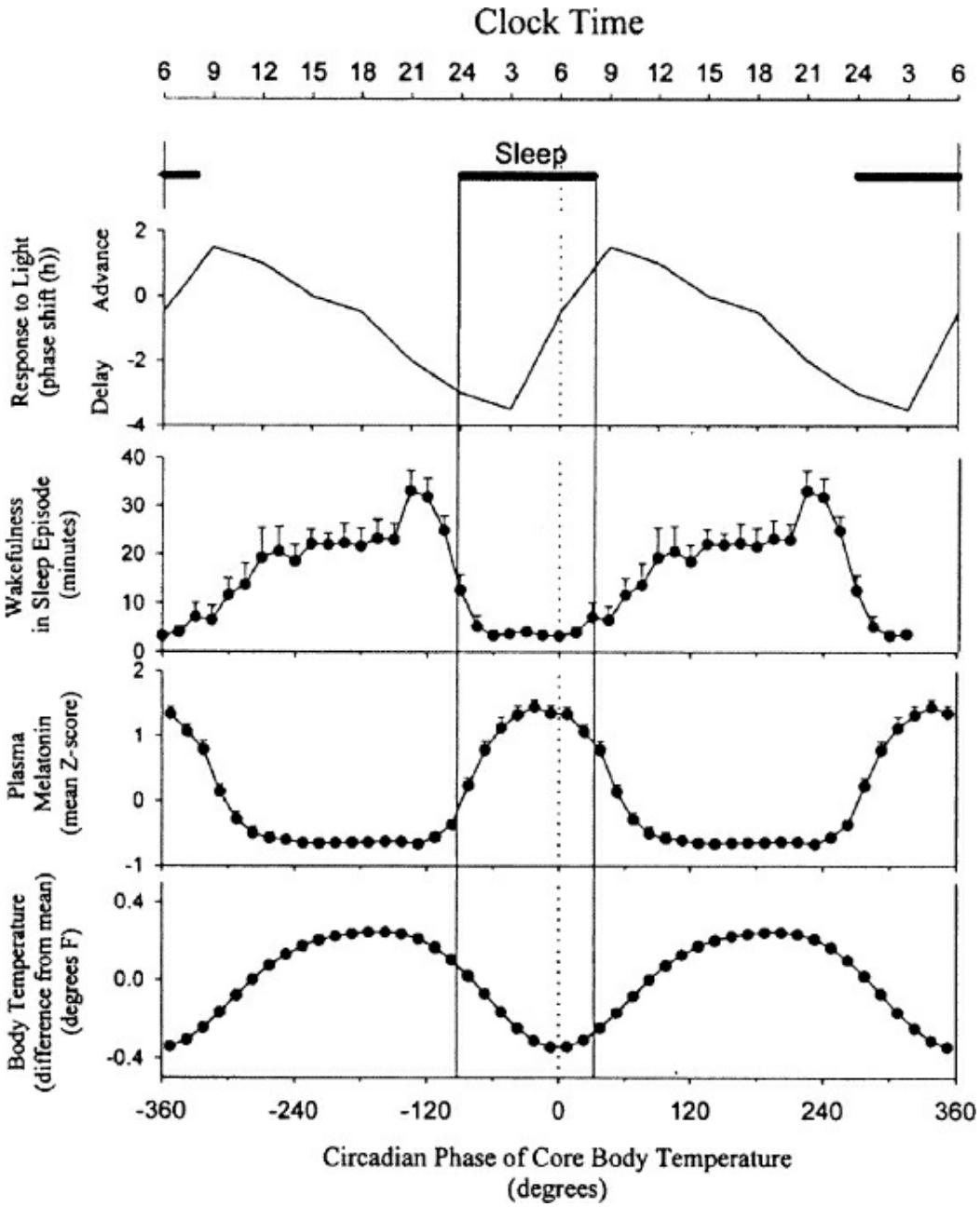

FIGURE 1.14–6. Relative phase relationship of sleep in young adults to other circadian phase markers. (From Dijk D-J, Lockley SW. Invited review: Integration of human sleep-wake regulation and circadian rhythmicity. *J Appl Physiol*. 2002;92:852, with permission.)

It should be stressed that diurnal preference varies among individuals as a function of age, endogenous circadian periods, and other factors. This variability is paralleled by physiology. Clinically, diurnal preference, or chronotype, can be quantified by one of several available morningness–eveningness questionnaires. In qualitative terms, individuals who are labeled as morning types or *morning larks* tend to awaken earlier and experience the core body temperature minimum at an earlier clock time relative to those who are labeled as evening types or *night owls*. Sleep deprivation studies have shown that the homeostatic component of sleep is remarkably similar among individuals of similar age. (It should be noted

that there is a well-established age-dependent decline in sleep need.) Therefore, chronotype is dictated almost exclusively by the circadian process of sleep regulation. Morning and evening chronotype loosely reflect advanced and delayed circadian rhythms relative to the environmental light cycle.

The advent of the light bulb has extended the human day into the natural night. This encroachment on the night, although increasing productivity, has affected human sleep patterns (Fig. 1.14–7). Typical use of artificial lights results in a single, consolidated bout of sleep lasting approximately 8 hours. This pattern of sleep is uncommon among most other mammals, which typically experience more fractured sleep. Human sleep under more natural photoperiods, where the duration of the night is longer, becomes decompressed. Specifically, a bimodal distribution of sleep is observed; bouts of sleep occur in early and late night. Periods of *quiet wakefulness* are interspersed between the two primary bouts of sleep. This bimodal sleep pattern is more similar to the sleep patterns of other mammals.

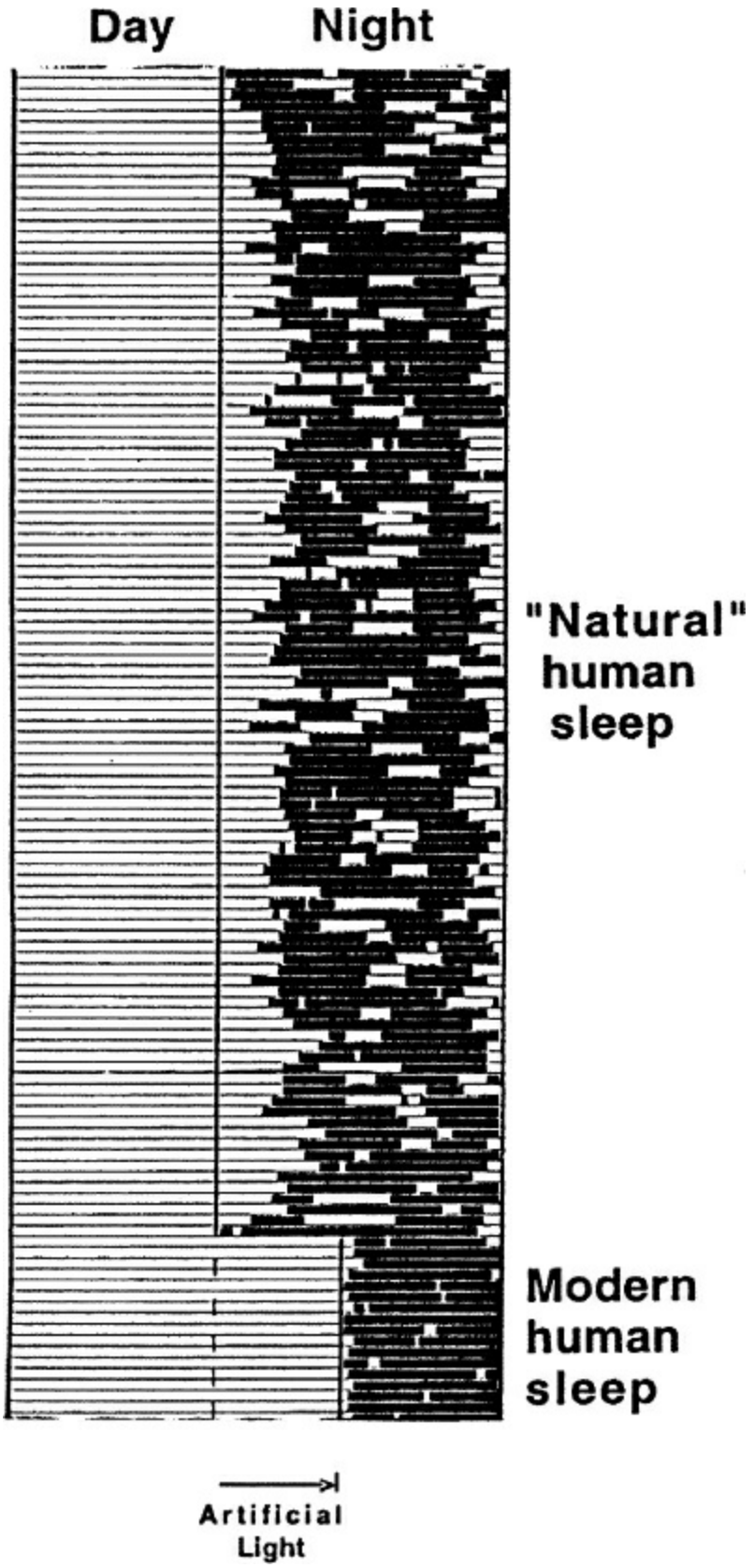

FIGURE 1.14–7. Change of sleep structure in response to artificial lighting. Total sleep time is reduced, and periods of quiet wakefulness are abolished by extending daytime into nighttime through artificial lighting. (From Wehr TA, Moul DE, Barbato G, et al. Conservation of photoperiod-responsive mechanisms in humans. *Am J Physiol*. 1993;265:R846, with permission.)

SEASONALITY

The 24-hour period of the Earth's rotation around its axis is unchanging. However, the Earth's axis is tilted 23.45 degrees from the plane of its own orbit around the sun (the ecliptic). As a result, the relative proportion of daytime to nighttime within the 24-hour astronomical day varies as the Earth proceeds through its orbit of the sun. Many organisms are capable of synchronizing physiology to the seasonal cycle to maximize survival. For example, precise seasonal cycles in reproduction are seen throughout the plant and animal kingdoms. Large mammals that typically have long gestation periods, such as sheep, conceive in the fall when the nights are long and the days are short, so birth occurs during the relatively mild season of spring. These animals are referred to as *short-day breeders*. Conversely, mammals with gestation periods of only a few weeks, such as hamsters, conceive and give birth during spring and summer, when the days are long and the nights are short. Hence, these animals are referred to as *long-day breeders*. Like circadian rhythms, many of these yearly (circannual) rhythms tend to persist in the absence of seasonal cues with endogenous periods of approximately 1 year.

Melatonin and Seasonality

The most reliable environmental parameter providing a faithful representation of the solar day is the day–night cycle. Similarly, the most reliable environmental parameter reflecting the progression through the seasons is the change in day length, the fraction of the 24-hour day between sunrise and sunset. In seasonally breeding animals, day length is physiologically encoded through the melatonin profile. As described previously, melatonin levels are elevated during the night. A long night, such as that experienced during the short day lengths of winter, results in an elevated melatonin profile of a relatively long duration. A short summer night, by contrast, results in a short duration of elevated melatonin. This seasonal signal is interpreted by the reproductive axis, resulting in an appropriate reproductive response. Melatonin's role in transducing day length was elucidated by pinealectomizing seasonally breeding animals, thereby removing the primary endogenous source of melatonin. Melatonin was then infused in profiles mimicking long days or short days. The duration of elevated melatonin was the primary determinant of seasonal reproductive status, even when the infused profile was administered under a conflicting day length. Variations in other parameters, such as the amplitude of the melatonin profile, the amount of total melatonin synthesized, or the phase relationship of the profile to the light–dark cycle, are of limited importance in producing a humoral signal that transduces day length.

Reproductive responses to changing day length can be dramatic. A male

Siberian hamster (*Phodopus sungorus*) maintained in long days is reproductively competent and typically has a testicular weight of approximately 250 mg per testis. Under short days, however, the testes regress to approximately 15 mg per testis, representing a 94 percent decrease in testicular mass. The same degree of regression is observed in response to melatonin infusions that mimic short days. Communication of the hormonally transduced day length to the reproductive axis is likely to be mediated, at least partially, through melatonin receptors in the *pars tuberalis* near the pituitary stalk. The exact mechanism remains unknown, but activation of these receptors is hypothesized to indirectly regulate an unidentified factor putatively named *tuberalin*. Tuberalin, in turn, controls gene expression and prolactin release from lactotrophs in the adenohypophysis of the pituitary.

Seasonality in Humans

Seasonal variations in human behaviors include seasonal peaks in birth rate, sexual activity, conception, pain thresholds, suicide, and crimes including rape. Seasonal variations are seen in biological processes including growth rate, cortisol, and testosterone levels. The general population experiences a winter worsening of mood, heightened appetite, lengthened sleep duration, lower energy, and less social behavior. While such seasonal variations may have once been adaptive as a mechanism to conserve energy in winter and increase energy in summer to capitalize on seasonally available resources, the industrial revolution has diminished the need for such seasonal variations, and has created an environment in which seasonal variations in mood and behavior are potentially maladaptive.

The decompressed bimodal structure of human sleep during long nights indicates that the length of natural sleep is related to the length of the night. Potentially, a two-oscillator system could function to maintain proper sleep patterns during changing photoperiods. Such a proposed system would consist of an evening oscillator that tracks the transition from day to night (dusk) and a morning oscillator that tracks the transition from night to day (dawn). The relative phase differences between these oscillators may encode the changing day lengths associated with the passing of the seasons. Biological evidence for a two-oscillator system exists in rodents and humans.

The melatonin profile of many vertebrates, including some humans, is bimodal, with evening and morning peaks. In rodents, metabolic and electrophysiological studies of the SCN typically have been done in brain slices cut in the coronal plane. Results of electrophysiological studies conducted in brain slices cut in the horizontal plane have provided new insights. The action potential frequency in SCN neurons from horizontally cut preparations is bimodal, with peaks in the early and late subjective day. Furthermore, the interpeak interval varies as a function of the photoperiod

in which the animal was housed. These studies lend credence to long-standing suspicions that the SCN of seasonally breeding mammals and, perhaps, nonseasonal mammals harbor a morning and evening oscillator that interact to convey day-length information.

In seasonally reproductive mammals, the duration of the nightly increase in melatonin effectively encodes day length (or, more accurately, night length). By contrast, in the vast majority of humans, the duration of elevated melatonin is invariant throughout the year. Recent studies have shown that healthy men living in their usual home environment had winter and summer melatonin profiles that were indistinguishable. However, healthy men enrolled in a carefully controlled photoperiod experiment gave surprisingly different results. In this cohort, long nights elicited an extended period of melatonin elevation. Conversely, short nights produced a compressed period of elevated melatonin. In essence, humans retain the capacity to encode day length, although this capacity is masked by the self-imposed artificial lighting regimens of modern society. It should be noted that a small percentage of individuals residing in their usual environment exhibits melatonin profiles that track day length, much like seasonally breeding mammals. Of particular interest are male patients experiencing SAD, some of whom exhibit this apparent seasonality.

CIRCADIAN RHYTHMS IN PSYCHIATRIC DISORDERS

Seasonal Affective Disorder

SAD is the most overt manifestation of seasonality in humans. It is characterized by recurrent major depressive episodes followed by periods of remission, or change to mania or hypomania, that recur on a seasonal basis. SAD is diagnosed as MDD, bipolar I or bipolar II disorder, with the specifier "with seasonal pattern" in the Fifth Edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). The seasonal pattern specifier criteria include:

A. Regular temporal relationship between the onset of major depressive episodes and a particular time of the year (unrelated to obvious season-related psychosocial stressors).
B. Full remissions (or a change to mania or hypomania) also occur at a characteristic time of the year.
C. Two major depressive episodes meeting criteria A and B have occurred in the last 2 years, and no nonseasonal episodes have occurred in the same period.
D. Seasonal major depressive episodes substantially outnumber the nonseasonal episodes over the individual's lifetime.

The most prevalent form of SAD has an onset in the late fall and early winter and remits in the late spring and early summer. This condition is frequently referred to as *winter SAD,* or *winter depression,* but is a

clinically significant level of depression unlike the "winter blues." Symptoms atypical of Major Depressive Disorder can present with winter SAD. These can include a significant increase in weight, increased appetite, an increase in sleep, a heightened sensitivity to interpersonal rejection, and a leaden feeling in the extremities or psychomotor retardation. Most distinct, however, is a craving for carbohydrates.

Surveys indicate the prevalence of winter SAD among the general population to be between 4 and 9 percent. Women are up to four times as likely as men to be affected, and as much as 20 percent of the population may have subsyndromal symptoms. Rates of SAD are higher among relatives of those with a confirmed diagnosis of SAD and twin studies identified a heritability of 29 to 43 percent for seasonality and/or SAD.

The gold standard treatment for winter SAD is light therapy. About half of individuals with SAD experience remission in response to light therapy, representing one of the highest treatment response rates in the field. However, the mechanism behind light therapy is not yet clear. A typical prescription for light therapy involves 45 to 90 minutes daily exposure to a broad spectrum, ultraviolet-filtered, white light source of relatively high irradiance (5,000 to 10,000 lux). Recent studies have suggested that CBT may be equally efficacious compared to light therapy for acute treatment, and may offer prophylactic benefits against recurrences in subsequent years. Multiple antidepressant medications have also been used successfully to treat winter SAD.

The antidepressant effect of phototherapy in winter SAD patients has given rise to several hypotheses regarding the etiology of the disorder. One hypothesis proposes that SAD patients experience the consequences of a chronically phase delayed circadian clock, suggesting that the aberrant phase angle between the clock and the environment is causative of winter SAD. Although most individuals with SAD report being evening chronotypes, nearly half do not have significantly delayed rhythms. The morning offset of the nightly release of melatonin appears delayed among men with SAD relative to that of healthy controls. However, the onset of melatonin increase is not phase-shifted relative to that of controls. In essence, these patients have a longer duration of elevated melatonin that increases at the same clock time as that of controls but stays elevated longer, thus impinging on the morning hours. Although these data are not consistent with the phase delay hypothesis, they may explain the effectiveness of morning bright light therapy that would acutely suppress the extended melatonin profile of winter SAD patients. It should be noted that the sculpting of the melatonin profile by morning light exposure cannot entirely explain the proven success of phototherapy. In some winter SAD patients, bright light administered during the evening is also antidepressant. In fact, some phototherapy treatment paradigms prescribe morning and evening light exposures, although data suggest that morning

light therapy is most effective for the average patient. Individually tailoring light therapy timing based on individual chronotype may confer some advantage, although the data for such tailored prescriptions is limited.

The success of light therapy administered at various clock times suggests that winter SAD may not have a circadian-based etiology. An alternate hypothesis proposes that patients experiencing winter SAD are generally less sensitive to light than healthy counterparts. Such photic insensitivity would become apparent during the decreased light levels of late fall and early winter, depriving these individuals of the threshold of light required to stave off depression. Accordingly, daily supplementation of light through bright phototherapy would be expected to exceed this theoretical threshold. Multiple studies have identified a retinal subsensitivity in SAD, including a subsensitivity that appears specific to the melanopsin containing retinal cells responsible for conveying light information to the SCN. However, the link between sensitivity of these input cells and treatment response has not been established yet.

A basic tenet in the development of pharmacological treatments is that a dose dependence must exist to implicate the effectiveness of the drug in question. Similarly, a dose or fluence dependence should be observed with respect to treatment of winter SAD with bright phototherapy. Several studies attempting to document a fluence–response relationship in the treatment of winter SAD have provided the field with conflicting data. Moreover, light therapy, like other photobiological responses, should show a wavelength dependence that reflects the spectral sensitivity of the photopigments mediating that response. Indeed, light therapy focused in the range of blue wavelengths appears to be as effective as brighter white or full spectrum light therapy, although the safety of long-term use of blue-enriched light therapy devices has not been established.

While there is evidence for retinal subsensitivity in SAD, another possibility is that SAD patients do not experience an inherent insensitivity to light but rather fail to respond appropriately to light. Such responsiveness is hypothetically mediated by central mechanisms downstream of retinal photoreception. Several physiological responses to light have been tested among SAD patients, and no striking deficits in photoresponsiveness were observed relative to healthy controls. Light exposure, however, elicits a myriad of biological responses, some of which are subtle. Studies comparing the photoresponsiveness of SAD and control subjects are far from comprehensive.

In general, it cannot be denied that phototherapy has proven to be an effective treatment for winter SAD. The mechanisms by which bright light ameliorates the symptoms of this disease remain unknown.

NONSEASONAL UNIPOLAR DEPRESSION AND CIRCADIAN RHYTHMS

Abnormal circadian rhythms have been documented in those with

unipolar, nonseasonal depression. For instance, individuals with depression are more likely to be evening chronotypes and often show delayed rhythms and reduced peak amplitude of melatonin and core body temperature relative to healthy controls. Daily patterns of worsening mood symptoms in morning or afternoon are evident in nonseasonal depression, and coincide with brain glucose metabolism in affect regulation areas of the brain. Several variants in clock genes have been found to be associated with depression diagnosis, including *CLOCK* and *CRY,* although reports are inconsistent.

In addition, sleep disturbances are characteristic of nonseasonal depression, specifically insomnia. Additionally, REM latency or the time from sleep onset to the first REM episode is shortened, while the duration of REM is lengthened in depression. As improvement in sleep is an important marker of lower recurrence rates, circadian control of sleep is a potential avenue to improve long-term outcomes in depression. A curious phenomenon related to depression and sleep is that total sleep deprivation can provide a transient antidepressant effect in a majority (approximately 60 percent) of depressed patients. This "wake therapy" involves sleep restriction for one night, or for the latter half of one night, but cannot be used as a monotherapy because depressive symptoms will return after a full night of sleep, and the treatment effect must be maintained with another type of treatment. Other chronotherapies, beyond this type of sleep deprivation, include advancing sleep phase, bright light therapy, and melatonin administration. Bright light therapy for nonseasonal depression demonstrates moderate improvement for nonseasonal depression based on a meta-analysis of 20 studies, and is administered to nonseasonally depressed patients similarly as to those with SAD. As is also the case in SAD, evidence for misaligned or shifted circadian phase in nonseasonal depression is inconsistent. This discrepancy suggests that light could have noncircadian antidepressant effects in addition to circadian effects. However, light therapy for nonseasonal, unipolar depression represents an option for adjunctive treatment in addition to antidepressant medication, or another treatment option for cases in which medications are undesirable as in some cases of pregnancy.

Given that wake therapy can result in fast antidepressant effects, the practice of combining wake therapy with antidepressant medications and/or light therapy shows promise for achieving a rapid response followed by sufficient maintenance of the antidepressant response. In addition, it appears that effective antidepressant medications have therapeutic effects on sleep processes such as inhibiting excessive REM sleep or lengthening REM latency, possibly reflecting underlying circadian resynchronization. Melatonin may be directly linked with monoaminergic activity due to a recent finding that agomelatine (an antidepressant medication that acts as a melatonin receptor agonist) increases monoaminergic activity.

Administration of either melatonin or agomelatine 5 hours prior to bedtime can advance rhythms and improve mood.

When negative life events occur in depression, this may lead to a disruption of social rhythms (e.g., timing of first daily social encounter after the death of a spouse). Limited evidence suggests that such a disruption or irregularity of social rhythms leads to similar disruptions in biological rhythms, and that treatments to restore the regularity of social rhythms may be effective (see IPSRT below). Circadian regulation of hormonal, temperature, and mood rhythms differ, may be differentially influenced by social zeitgebers as well as genetic factors, and are likely to interact, underlining the complexity of circadian influences on mood disorders.

Bipolar Disorder

As in nonseasonal depression, circadian abnormalities have been suggested as a potential important etiological mechanism in the development of Bipolar Disorder. Bipolar Disorder has been linked to variants in the core clock genes in humans and in animal models. For example, mice with a mutation in the *Clock* gene (*Clock*Δ19) show manic-like behaviors, including hyperactivity and decreased sleep. These *Clock*Δ19 mice have offered a helpful model to better understand the contribution of circadian disruption in Bipolar Disorder, specifically in recognizing the potential importance of the VTA. *Clock*Δ19 mice have increased dopamine synthesis in the VTA, suggesting the importance of the *Clock* gene in dopaminergic activity in the VTA. Further, expression of CLOCK protein in the VTA rescues most of the *Clock* mutant mice manic-like behavior, including hyperactivity and anxiety. Together, *Clock*Δ19 mice models suggest the importance of the circadian modulation of dopaminergic activity within the VTA in Bipolar Disorder.

Human studies in Bipolar Disorder also point to abnormal circadian regulation as a potential etiological mechanism. Circadian abnormalities seem to vary based on which state an individual is experiencing. For example, depression is associated with a phase delay, whereas mania is associated with a phase advance. Therefore, stabilization of these rhythms may help in preventing or reducing the occurrence of manic or depressive episodes in bipolar individuals. Indeed, Interpersonal Social Rhythm Therapy (IPSRT), a type of interpersonal therapy, was developed due to the observation that irregular daily routines resulting in circadian dysregulation precipitate mood episodes. Importantly light therapy has been shown to unmask manic tendencies in those with a predisposition for developing Bipolar Disorder, perhaps due to the relatively abrupt circadian changes, and evidence for the efficacy of light therapy in Bipolar Disorder is currently limited.

Circadian Sleep Disorders

In the DSM-5, Circadian Rhythm Sleep–Wake disorders include multiple disorders with circadian symptoms and presumed circadian etiology. Among these, advanced sleep phase syndrome (ASPS) is a pathological extreme of the morning lark phenotype. An autosomal-dominant familial form of ASPS (FASPS) recently has been genetically characterized. Afflicted family members exhibit a striking 4-hour advance of the daily sleep–wake rhythm. They typically fall asleep at approximately 7:30 PM and spontaneously awaken at approximately 4:30 AM. Affected individuals have a SNP in the *PER2* gene. This adenine-to-guanine nucleotide polymorphism results in serine-to-glycine amino acid substitution that causes the mutant protein to be inefficiently phosphorylated by casein kinase Iε, an established component of the circadian molecular clockwork. Similarly, delayed sleep phase syndrome (DSPS) is likely influenced by genetics. A length polymorphism in a repeat region of the *PER3* gene appears to be associated with morning chronotype in DSPS patients, the shorter allele being associated with evening chronotype.

Although sleep disorders are discussed in Chapter 21, the following Sleep–Wake and Circadian Rhythm Disorders show improvement in response to bright light therapy: DSPS, ASPS, Hypersomnia, Sleep Maintenance Insomnia in the Elderly, Jet Lag and Shift Work Disorders, and Non-24-Hour and Irregular Sleep–Wake Type Circadian Rhythm Disorders. Light therapy is also helpful in normalizing sleep during withdrawal from alcohol, and could therefore be considered as an adjunctive option for many disorders in which sleep and mood are impaired. Chronotherapeutics, treatments administered either according to biological rhythms or in order to normalize biological rhythms, include bright light therapy, melatonin and its agonists, and wake and dark therapy. Such treatments are often the first line of treatment in Circadian Rhythm Sleep–Wake disorders.

Memory, Aging, and Neurocognitive Disorders

Converging circadian disruption paradigms in animals, including both photoperiod shift and clock gene knock-out (KO) models, suggest an important role for circadian organization in proper memory performance. In several studies exposing rats to photoperiod shifts, the shifts lead to poor performance on both acquisition and retention phases of the Morris water maze task relative to rats housed under a naturalistic light dark cycle. Additionally, permanently arrhythmic hamsters are unable to differentiate between novel and familiar objects 20 and 60 minutes after training, representing a significant deviation from normal memory of previously presented objects. *Clock* gene KO mice also demonstrate abnormal memory performance, indicating the importance of appropriate circadian alignment in normative memory performance.

In general, as humans age and normative age-related memory problems begin, individuals experience a shorting of circadian period as well as phase advances resulting in earlier waking times and bedtimes and decreased the amplitudes of most circadian rhythms. Further, aging individuals may be less tolerant of dramatic phase shifts such as those caused by jet lag. Again, a mouse model has provided interesting insight into the interaction of the aging process and the circadian clock. The effect of chronic jet lag on aged mice has dramatic consequences on mortality. About half of aged mice forced to phase advance 6 hours once per week survive this treatment compared with an 83 percent survival rate in unshifted age-matched mice. Aged mice subjected to weekly 6-hour phase delays show an intermediate survival of 68 percent. These profound age effects on mortality in response to phase shifts are not observed in younger mice. The pathogenesis of chronic jet lag remains to be determined. Interestingly, these mice did not suffer an increased rate of tumorigenesis. It is likely that, in humans, as in mice the internal desynchrony of oscillators that results from a shifting light schedule may have dire consequences that are exacerbated by aging.

Age-related circadian changes, like decreases in amplitude, are exacerbated in age-related neurocognitive disorders including Alzheimer disease (AD). For example, those with moderate dementia exhibited fragmented rhythms and decreased amplitude in activity rhythms relative to mild dementia and healthy control counterparts. The amount of disruption in activity rhythms has also been associated with the severity of neuropathology found in postmortem brains, such that those with the most aberrant activity rhythms around their time of death also exhibited the most neuropathology. Dysregulated circadian rhythms have been found to predate the occurrence of dementia and/or clinical diagnosis of AD and significantly predict the occurrence of AD or mild cognitive impairment (MCI). Together, these pieces of evidence would suggest that circadian disruption may play a central role in the development of the neuropathology and neurocognitive deficits associated with AD. Indeed, light therapy aimed to normalize an individual's circadian rhythms has been shown to improve AD patient's sleep and cognitive scores.

Other Psychiatric Disorders

Understanding the mechanism of action of light therapy may aid in our understanding of the etiology and treatment of multiple other psychiatric disorders that show improvement in response to light therapy. In addition to SAD and nonseasonal mood disorders, other disorders that improve in response to light therapy include Ante- and Post-Partum Depression, Pre-Menstrual Dysphoric Disorder, Eating Disorders Including Bulimia Nervosa, Attention Deficit Hyperactive Disorder and disordered sleep due to traumatic brain injury.

OTHER CIRCADIAN-CLOCK-ASSOCIATED PATHOLOGIES

There have been many reports of circadian-clock-associated pathologies. It has been suggested that desynchrony between and among the SCN and the various oscillators in peripheral tissues lies at the heart of these maladies. Examples include cancer, heart disease, and metabolic disorders. Cardiovascular disease risk factors such as obesity, low high-density lipoprotein (HDL) cholesterol levels, and high triglycerides are more prevalent among shift workers than day workers, as an example. Furthermore, many of these associations with health problems increase in aged shift workers. Epidemiological studies have shown that women working night shifts have a significantly elevated risk of breast cancer. The advent of rodent genetic models with compromised circadian systems has provided new insight into the pathogenesis of such conditions.

CIRCADIAN RHYTHMS AND PHARMACOTHERAPY

Circadian rhythmicity can be affected by drugs, and conversely, the circadian clock can modulate the efficacy of drugs throughout the course of the day. A better understanding of these interactions will lead to more effective pharmacotherapies. Some of the best-studied interactions between medications and the circadian clock have included the circadian effects of antidepressants. Elevated nocturnal body temperature is a common feature among depressed patients. This effect may be due to a reduced amplitude of the master circadian oscillator in the hypothalamus that drives body temperature. TCAs and SSRIs reduce elevated nocturnal body temperature while simultaneously enhancing circadian amplitude. Similarly, many depressed patients exhibit a dampened amplitude in daily activity rhythms. Like body temperature, the amplitude in daily activity cycles of depressed individuals may be augmented by TCA or SSRI treatment.

The use of lithium to treat Bipolar Disorder has been long established. However, lithium also impacts the circadian system, resulting in a lengthening of circadian period. The molecular mechanism by which this occurs remains unknown. Glycogen synthase kinase 3β (GSK3β) participates within the molecular clock mechanism. Interestingly, GSK3β is inhibited by lithium. In cell culture, GSK3β has been shown to stabilize the negative clockwork regulator REV-ERBα via phosphorylation. REV-ERBα typically represses the transcription of the *BMAL1* gene. In the presence of lithium, however, GSK3β is inhibited, thereby preventing the phosphorylation and stabilization of REV-ERBα, which as a consequence is targeted for proteasomal degradation. The degradation of REV-ERBα results in the derepression of *BMAL1* transcription. Whether lithium's influence on circadian behavior is attributable to its inhibitory effect on GSK3β-mediated stabilization of REV-ERBα remains to be determined.

Short-acting benzodiazepines (e.g., triazolam [Halcion] and brotizolam [Lendormin]) also exert chronobiological effects. In hamsters, triazolam or brotizolam administered during the middle of the subjective day induces circadian phase advances in activity. Brotizolam has been shown to reduce the light-induced expression of clock genes *Per1* and *Per2* in the SCN. While benzodiazepines are allosteric modulators of γ-aminobutyric acid A receptors ($GABA_A$), several lines of evidence indicate that the circadian effects of short-acting benzodiazepines require an intact serotonergic system. When the $5\text{-}HT_{1A/7}$ receptor agonist 8-hydroxy-2-(di-*n*-propylamino)-tetralin (8-OH-DPAT) is injected into hamsters at subjective midday, phase advances in locomotor behavior and SCN neuronal activity are observed in addition to a reduction in *Per1* and *Per2* gene expression in the SCN. Recreational drugs of abuse also impact the circadian system. 3,4-Methylenedioxymethamphetamine (MDMA) or "ecstasy" can act as a serotonin neurotoxin. Hamsters treated with MDMA showed reduced triazolam-induced phase shifts in circadian locomotor activity and a diminished ability to reentrain rhythms posttreatment. MDMA-treated animals exhibited a reduction in serotonergic axonal terminals in the SCN, again emphasizing the importance of an intact serotonergic system in the regulation of the circadian axis. Recreational use of methamphetamine has increased dramatically within the past decade. Chronic administration of methamphetamine disorganizes rodent activity rhythms. However, administration of methamphetamine to rodents rendered arrhythmic through ablation of the SCN results in a reemergence of rhythmicity. The mechanism involved in the rescue of rhythmicity or site of action remains unknown.

The efficacy and toxicity of many pharmacotherapeutics vary as a function of circadian phase. Daily variations in fixed-dose lethal toxicity have been appreciated in rodents for years. Many anticancer drugs, ranging in mechanism from antimetabolites to DNA intercalators to mitotic inhibitors, have been shown to have 2- to 10-fold changes in tolerability in rodents over the course of the day. Much of this difference is attributed to circadian variations in the body's ability to absorb, distribute, metabolize, and eliminate toxic compounds. These four processes are affected by underlying circadian rhythms in physiological processes such as daily variations in gastric pH, gastrointestinal mobility, glomerular filtration rate, and membrane viscosity. The rhythmic intake of food during traditionally timed meals also influences how therapeutic drugs are handled by the body. It is becoming clear that to maximize efficacy and minimize toxicity of drugs circadian phase of administration must be considered. Appropriate circadian timing of the administration of multiple drugs can be a daunting challenge to infirmed individuals or their caretakers. The development of small implanted programmable pumps that can be directed to administer anticancer drugs or other therapeutics at

particular times of day may provide a limited solution to this challenge. The emergence of the field of chronotherapy is a reflection of the increased understanding of the impact of the circadian system on the effectiveness of pharmacological treatments.

FUTURE CONSIDERATIONS

Evidence is emerging that the influence of the circadian system is much broader than previously believed. Desynchrony between the SCN and peripheral oscillators is likely to be involved in the pathogenesis of multiple maladies not limited to psychiatric disorders. Additionally, treatments designed to manipulate the light input, central clock, or clock output mechanisms show promise in treating multiple psychiatric disorders. Bright light therapy is the first line of treatment for MDD with seasonal pattern, and is considered a useful adjunct in nonseasonal depression as well. An important issue not yet resolved is whether the antidepressant effect of bright light therapy is circadian or also has noncircadian antidepressant effects. Resolving this question of mechanism may lead to better improvement rates in response to light therapy.

REFERENCES

Berger M, Vollmann J, Hohagen F, Konig A, Lohner H. Sleep deprivation combined with consecutive sleep phase advance as a fast-acting therapy in depression: an open pilot trial in medicated and unmedicated patients. *Am J Psychiatry*. 1997;154:870.

Bohle P, Di Milla L, Fletcher A, Rajaratnam S. Introduction: aging and the multifaceted influence on adaptation to working time. *Chronobiol Int*. 2008;25:155–164.

Carter DS, Goldman BD. Antigonadal effects of timed melatonin infusion in pinealectomized male Djungarian hamsters (*Phodopus sungorus sungorus*): duration is the critical parameter. *Endocrinology*. 1983;113:1261.

Chen-Goodspeed M, Lee CC. Tumor suppression and circadian function. *J Biol Rhythms*. 2007;22:291.

Czeisler CA, Duffy JF, Shanahan TL, Brown EN, Mitchell JF. Stability, precision, and near-24-hour period of the human circadian pacemaker. *Science*. 1999;284:2177.

*Davidson AJ, Sellix MT, Daniel J, Yamazaki S, Menaker M. Chronic jet-lag increases mortality in aged mice. *Curr Biol*. 2006;16:R914.

Dijk DJ, Lockley SW. Integration of human sleep-wake regulation and circadian rhythmicity. *J Appl Physiol*. 2002;92:852.

Germain A, Kupfer DJ. Circadian rhythm disturbances in depression. *Human Psychopharmacology*. 2008;23:571.

Harvey AG. Sleep and circadian rythms in bipolar disorder: seeking synchrony, harmony, and regulation. *Am J Psychiatry*. 2008;165:820–29.

Herzog ED, Schwartz WJ. A neural clockwork for encoding circadian time. *J Appl Physiol*. 2002;92:401.

*Klein DC, Moore RY, Reppert SM, eds. *Suprachiasmatic Nucleus: The Mind's Clock*. New York:

Oxford University Press; 1991.

Klerman EB, Shanahan TL, Brotman DJ, Rimmer DW, Emens JS. Photic resetting of the human circadian pacemaker in the absence of conscious vision. *J Biol Rhythms*. 2002;17:548.

*Ko CH, Takahashi JS. Molecular components of the mammalian circadian clock. *Hum Mol Genet*. 2006;15:R271.

Kondratova AA, Kondratov RV. The circadian clock and pathology of the ageing brain. *Nat Rev Neurosci*. 2012;13(5):325–35.

*Levi F, Schibler U. Circadian rhythms: mechanisms and therapeutic implications. *Annu Rev Pharmacol Toxicol*. 2007;47:593.

McClung CA. How might circadian rhythms control mood? Let me count the ways... *Biol Psy*. 2013;74(4):242–49.

Miller JD, Morin LP, Schwartz WJ, Moore RY: New insights into the mammalian circadian clock. *Sleep*. 1996;19:641.

Moore RY. Circadian rhythms: basic neurobiology and clinical applications. *Annu Rev Med*. 1997;48:253.

Morin LP. The circadian visual system. *Brain Res Rev*. 1994;19:102.

Nelson DE, Takahashi JS. Sensitivity and integration in a visual pathway for circadian entrainment in the hamster (*Mesocricetus auratus*). *J Physiol*. 1991;439:115.

Quintero JE, Kuhlman SJ, McMahon DG. The biological clock nucleus: a multiphasic oscillator network regulated by light. *J Neurosci*. 2003;23:8070.

Ralph MR, Foster RG, Davis FC, Menaker M. Transplanted suprachiasmatic nucleus determines circadian period. *Science*. 1990;247:975.

Roecklein KA, Wong PM, Miller MA, Donofry SD, Kamarck ML, Brainard GC. Melanopsin, photosensitive ganglion cells, and seasonal affective disorder. *Neurosci Biobehav Rev*. 2013;37:229.

*Roybal K, Theobold D, Graham A, et al. Mania-like behavior induced by disruption of CLOCK. *P Natl Acad Sci USA*. 2007;104(15):6406–11.

Ruby NF, Brennan TJ, Xie X, Cao V, Franken P. Role of melanopsin in circadian responses to light. *Science*. 2002;298:2211.

Smith BN, Sollars PJ, Dudek FE, Pickard GE. Serotonergic modulation of retinal input to the mouse suprachiasmatic nucleus mediated by 5-HT_{1B} and 5-HT_7 receptors. *J Biol Rhythms*. 2001;16:25.

*Takahashi JS, Turek FW, Moore RY, Takahashi JS, Turek FW, eds. *Handbook of Behavioral Neurobiology: Circadian Clocks*. New York: Kluwer Academic Publishers; 2001.

Toh KL, Jones CR, He Y, Eide EJ, Hinz WA. An hPer2 phosphorylation site mutation in familial advanced sleep phase syndrome. *Science*. 2001;291:1040.

Weekes NY, Lewis RS, Goto G, Garrison-Jakel J. The effect of an environmental stressor on gender differences on the awakening cortisol response. *Psychoneuroendocrinology*. 2008;33:766–772.

Yamazaki S, Goto M, Menaker M. No evidence for extraocular photoreceptors in the circadian system of the Syrian hamster. *J Biol Rhythms*. 1999;14:197.

▲ 1.15 Applied Electrophysiology

JASON K. JOHANNESEN, PH.D., AND PATRICK D. SKOSNIK, PH.D.

INTRODUCTION

Over the past few decades, electrophysiology has contributed substantially to the understanding of normal brain functions as well as brain function deviations in psychopathological conditions. Monitoring brain processes in real time requires millisecond-level resolution to follow the typical timing and rapid unfolding of neural events. Electroencephalographic (EEG) techniques enable the study of brain activity with a high temporal resolution, making it possible to measure differences in the time course of brain electrical activation during both elementary sensory processes (e.g., sensory registration and discrimination) as well as complex cognitive processes (e.g., mental calculations and working memory). Despite the continued development of high-density digital EEG recoding and associated methods of interpolating sources of underlying neural activations, the strength of EEG remains to be nested in ascertainment of events as they occur in time rather than from which brain structures they evolve. This distinction is commonly discussed in terms of the tradeoff between "temporal" and "spatial" resolution, which applies to all methods of brain imaging and, thereby, should guide the selection of methods most appropriate to investigate questions concerning neural processes versus brain structures associated with a particular clinical phenomenon.

A renascence of EEG methods has been promoted by the accessibility of these instruments to cognitive and clinical scientists who seek cost-efficient and noninvasive means of applying measures of brain function to new experimental, diagnostic, and computational questions. EEG methodology is exquisitely applied in perusal of current theories on the pathogenesis of psychiatric illness, no longer conceptualized as the result of a dysfunction of one or more brain regions, but as a consequence of the failure to integrate neural activity through networks engaged across multiple brain regions. Advances in electrophysiological measurement of brain function has been aided by technological advances at levels of bioamplification, electrode sensors, and computational algorithms for signal processing and quantification. As will be discussed further in this chapter, EEG remains early in its development as an *applied* area of psychiatry, yet to be widely accepted as a diagnostic tool beyond detection of organic conditions that may mimic psychiatry symptomatology. Nevertheless, EEG methodology and its interface with experimental neuroscience continue to hold an important position at the forefront of psychiatry research, contributing both to areas of illness characterization and treatment development. In

time, and with necessary large-scale replication, the ultimate goal will be to translate research findings to clinical protocols that can be used to inform early prediction of illness, differential diagnosis, and treatment response of common psychiatric disorders.

Applied electrophysiology remains a growing discipline within the field of medicine broadly, encompassing a range of both diagnostic neuroevaluative techniques and therapeutic brain stimulation procedures. EEG is one example, used as a diagnostic tool for seizure and epilepsy, sleep disorders, head trauma, tumors, infections, encephalopathies, anoxia, and cases where acute toxicity may be suspected. There are a number of examples from clinical research, such as the case of traumatic brain injury secondary to blast pressure waves, where EEG has demonstrated superior sensitivity to a pathophysiological process than structural brain scans. In the practice of clinical psychiatry, EEG is used primarily as screening tool for ruling out organic etiology, with differential diagnosis among psychiatric disorders informed by history and clinical interview. Although this chapter will focus primarily on EEG, additional diagnostic techniques worth mention include evoked potentials (EP), entailing the visual evoked response (VER), auditory brainstem evoked potentials (ABEPs), and somatosensory evoked response (SER), reflecting the earliest stages of CNS registration of sensory stimulation. EP diagnostic procedures are commonplace to neurologic evaluation, however, paradigms derived from this work, such as the steady-state evoked potential (SSEP), have also been adopted in psychiatric research. While not discussed in this chapter, a range of *therapeutic* brain stimulation procedures, including transcranial magnetic stimulation (TMS), vagal nerve stimulation (VNS), and Deep Brain Stimulation (DBS) may hold new promise as treatment options in psychiatry. TMS, in particular, has received U.S. Food and Drug Administration (FDA) approval for patients suffering from depression who have not achieved satisfactory improvement from conventional therapies. Given these trends, it is conceivable that in the not too distant future psychiatrists could be offered specialized training programs to perform these diagnostic and therapeutic procedures. A generation of clinical psychiatric electrophysiologists (CPEs) thus could begin to provide a much needed service within the psychiatric community, as well as lead and galvanize clinically driven psychiatric electrophysiology research. Growth of this field nonetheless is expected to come mainly from various techniques used to extract information from EEG records, such as quantitative EEG (Q-EEG), event-related potentials (ERP), time–frequency analysis, coherence measurement, and source analysis, which will be discussed in subsequent sections.

HISTORY

Experimental evidence of brain-derived electrical potentials was first

established in 1874 with Richard Caton's (1842 to 1926) discovery of spontaneous electrical activity recorded from the exposed cortices of cats, rabbits, and monkeys. However, the lengthy transition from laboratory experiment to eventual acceptance of EEG was plagued by intense controversy. Despite the continuing accumulation of evidence, the notion of electrical potentials emanating from the brain was rejected for nearly 50 years by leading authorities. Caton's work was replicated by Vasili Danilevsky's 1877 report that electrical oscillations recorded from the animal brain could be altered by strong sensory stimuli. Three further historical highlights of note include (1) the 1891 demonstration by Adolph Beck that the dog visual cortex produced large electrical potentials when the eyes were rhythmically illuminated (thus laying the experimental foundation for EEG photic driving), (2) Beck and Napoleon Cybulski's 1892 report that local injury to the cortex could alter the characteristics of recorded spontaneous electrical activity, and (3) Cybulski's 1914 report that brain wave seizure discharges could be induced in the cortex by applying electrical stimulation to the cortex (thus presaging the use of EEG in epilepsy). Despite this stream of successful experimental work, the EEG phenomenon continued to be scrutinized by the scientific community.

Caton's seminal work was finally brought to light when cited in 1929 by Hans Berger (1873 to 1941), reporting the discovery of the alpha waves at the emergence of the field of *electroencephalography*. The focused perseverance of Hans Berger, a biologically oriented Professor of Psychiatry and Director of the Psychiatric Clinic in Jena, Germany, finally brought EEG to a position of acceptance and clinical utility. After years of unsuccessful attempts to record brain waves from humans (he was able to obtain recordings from animals), he finally succeeded in recording the human EEG in 1924, and in 1929 published the first in his classic series of 23 papers describing many aspects of the human EEG. Among his vast achievements, Berger demonstrated that brain electrical activity came from neurons and not blood vessels or connective tissue, that recordings from patients with brain tumors contained high-voltage slow waves (his recording technique did not permit localization), that waking *alpha* waves were blocked by eye opening, and that the characteristics of EEG activity change with age, sensory stimulation, state of consciousness, and physiochemical state of the body. However, acceptance was still temporarily delayed when Lord Adrian, a Nobel laureate neurophysiologist, claimed that Berger's findings were "impossible." Later, in 1934, Lord Adrian publicly confirmed Berger's work, and the field of electroencephalography was born.

BASIC PRINCIPLES OF ELECTROENCEPHALOGRAPHY

The term *electroencephalogram,* coined by Hans Berger, refers to the graphic record of brain electrical activity, or brain waves. The related term,

electroencephalograph, refers to specialized instrumentation used to record electrical potentials at the scalp surface. The abbreviated term, *EEG,* may be used interchangeably to refer to either the record or the recording instrument. As rendered by EEG, the brain wave reflects the transient difference in electrical potential (greatly amplified) between any two points on the scalp or between an electrode placed on the scalp and one located at a neutral off-scalp location (e.g., ear lobe, mastoid, nose), used as a reference. In a simplistic sense, the EEG is an extremely sensitive voltmeter, with the unit of measurement being the microvolt (μV), or millionth of a volt. Typical EEG signals range from approximately 30 to 80 μV, but can be as low as 10 μV or as high as 150 or 200 μV in some high-voltage "spike" discharges. The difference in electrical potential measured between any two EEG electrodes fluctuates or oscillates rapidly, usually many times per second. It is this oscillation that produces the characteristic "squiggly line" that even many lay persons now recognize as the appearance of "brain waves."

The earliest EEG recordings involved only one pair of electrodes, or one *channel* of recording, and although this could detect certain normal and abnormal features, effective clinical application remained for the future. Soon, the ability to record two channels of EEG was accomplished, making it possible to record activity simultaneously from homologous locations in each hemisphere. Before long, the rapid advances in recording technology allowed four- and eight-channel recordings to be made, and EEG became a viable clinical tool. Eventually, 10-, 12-, and 16-channel recording machines became the standard workhorses of clinical and research EEG laboratories around the world. EEG equipment capable of simultaneous recording from 32 to 64 channels is now commonplace in most electrophysiology laboratories, with 128 and 256 channel systems also used in research where "high-density" recording may lend further accuracy to estimation of current sources. The ability to simultaneously record brain waves from many scalp locations is important, because it allows direct comparisons between homologous cortical regions, permits recording arrays to locate focal or regional abnormal features more clearly, and increases the ability to detect various *artifacts,* sources of signal of noncerebral origin that contaminate the EEG.

Brain wave activity captured on the scalp by EEG is a result of the summed extracellular electric potential fields generated by large numbers of neurons. Thus, EEG cannot detect the electrical activity generated by a single neuron or even by several neurons close to the scalp. Rather, the scalp-recorded EEG signals are the result of summated field potentials generated by excitatory and inhibitory postsynaptic potentials (i.e., EPSPs and IPSPs) in vertically oriented pyramidal cells of the cortex. An EPSP in a dendrite produces electrical negativity in the immediately surrounding area, but the electrical field becomes positive with increasing distance from

the source. The reverse occurs with an IPSP, generating an electrical positivity nearby and a negative field at a distance. The summation of EPSPs and IPSPs is enhanced, because the neurons are tightly packed together and oriented vertically in parallel. In order to be detectable by surface EEG electrodes, neural activity has to be strong enough to cross other tissue and fluid (i.e., CSF, dura, skull, scalp). It is the transduction of current across tissue and fluid of different densities, with different electrical resistance, that contributes to the wide diffusion of current as it is recorded over the scalp. This is one important aspect of EEG recording that must be considered when making assumptions of underlying source and the direction of current projection. The transduction of electrical activity from cortex to scalp relies on the ability of the brain to synchronize large populations of cells, which can occur either spontaneously or in response to a stimulating event. Neuronal activity that is unsynchronized cancels out due to destructive interference, thus does not contribute to the electric field potentials recorded as scalp EEG. Because of the manner in which the dominant intrinsic brain waves are generated, EEG is maximally sensitive to cortical neuronal activity and relatively insensitive to electrical potentials generated from subcortical regions. However, there are exceptions, as subcortical neuronal events can influence cortical neuronal firing via afferent transmissions along subcortical–cortical tracts.

Probably the first observation about brain waves, going back to the time of Caton, was that many recorded potentials oscillate and repeat in a rhythmic fashion. Indeed, the term *intrinsic rhythms* is often used for normal activity, and the term *dysrhythmic* is used for activity that might be abnormal. Within reasonable limits, the repetitive rhythmic nature of the EEG is stable across individuals and within individuals over time, barring the introduction of pathophysiological events. Indeed, the quantified EEG signal has been shown to have high test–retest reliability in a number of recording situations, with coefficients of $r = 0.90$ reported in healthy adults and in individuals with schizophrenia, while work by Thomas Lund and collaborators actually found higher test–retest reliability in schizophrenia ($r = 0.94$) than in healthy comparison subjects ($r = 0.70$). Examination of test–retest reliability in task-related EEG, recorded under conditions of mental challenge, tends to produce higher reliability coefficients than obtained from conventional resting EEG. The stability of EEG spectral characteristics may also have a genetic basis, with an average biometric heritability estimate of 0.78 reported for a large-scale ($N = 4026$) community sample of adolescent twins and their parents. Although the EEG does reflect state-dependent changes, sensitive to states of wakefulness and relaxation for instance, its reliability within individuals over time has been interpreted by some authors as evidence of an EEG set point, a natural spontaneous rhythm that is individually characteristic and stable over time. The concept of a set point suggests that repeated testing

with averaging across test sessions would help to eliminate measurement error, thus maximizing the chances of detecting illness-related changes. Given further evidence that EEG characteristics are to some extent heritable, EEG measures may be particularly useful for the early detection of pathophysiological features in individuals regarded to be at elevated risk for developing a psychiatric disorder due to family history.

Despite the more than 130 years that have passed since the original recording of EEG by Richard Caton, over which a variety of abnormal findings have been reported in many psychiatric conditions, electrophysiological methods have yet to find a prominent role in the training and clinical practice of psychiatry. This is in contrast to EEG's well-established position in clinical neurology as an approved subspecialty, sanctioned by the American Board of Medical Specialties (ABMS) and the American Board of Psychiatry and Neurology (ABPN). De-emphasis of EEG training in psychiatry is most likely secondary to many factors. Prominent among these factors are the complex and heterogeneous nature of psychiatric disorders in general, for which there are currently no formally accepted diagnostic techniques beyond the clinical interview. The lack of familiarity of psychiatrists with electrophysiological techniques, and heightened standards for evidence-based practices, are additional factors to contend with when considering the feasibility of dissemination of electrophysiological techniques in clinical psychiatry practice.

Basics of EEG Interpretation

The interpretation of the EEG record is essentially a problem of learning to recognize all of the myriad intrinsic waveforms, their expected distributions over cortical regions, and their range of variation in amplitude and degree of symmetry. Although there is little doubt that 64, 32, or even 16 channels of oscillating waves appear quite confusing to the beginner, there is an inherent regularity to the brain's electrical output that an experienced electroencephalographer comes to recognize. Because of this inherent regularity, the intrinsic EEG rhythms and their range of variation that characterizes the wake, drowsy, and various sleep states become easily identified through practice and experience. Various electrophysiological phenomenon of nonencephalic origin, which need to be distinguished from EEG signal, also produce localized or regional waveforms of stereotypical appearance that one can learn to identify. The primary job of the interpreter is to identify those EEG waveforms that appear to fall outside of what one might call the *normal range of variation* of the intrinsic background activity, and therefore are (1) frankly abnormal with known pathophysiological correlates or (2) controversial abnormal waveforms of potential clinical relevance, pending the results of further continuing clinical investigation.

Interpreter Reliability

Because the EEG record is obtained by precisely measuring the microvolt fluctuations of electrical potentials over the scalp, a misconception often emerges that EEG interpretation is purely objective in the sense that measurements of temperature, weight, length, and volume are objective. In truth, there is a large, subjective element of judgment in EEG interpretation, and the achievement of skill in this area only follows a period of thorough training and the guiding hand of clinical experience. Accepted EEG abnormalities do not always appear in the EEG tracing in clear-cut textbook form, and there are always ambiguities in areas where EEG activity that is clearly normal becomes blurred and shades off into waveforms that are unequivocally abnormal.

Although the important area of reliability of clinical EEG interpretation has not received extensive study, it has generally been assumed that high levels of interpretation reliability can be obtained through training and clinical experience. Two formal methods of assessing reliability involve comparing the independent readings of two or more EEG interpreters (*interrater reliability*) and asking one electroencephalographer to blindly reinterpret a sample of EEGs after an elapsed time (*intrarater reliability*). When these methods were employed in a recent study of six board-certified clinical neurophysiologists, asked to classify 300 EEGs into one or more of seven diagnostic categories, the findings suggested that clinical interpretation of EEG should receive much greater attention as a matter of training and practice. It was reported that five of the six readers provided confidence ratings of their interpretation with a median score of ≥99 percent. However, intrarater reliability coefficients ranged from 0.33 to 0.73 with an aggregate value of 0.59, while interrater coefficients ranged from 0.29 to 0.62, with an aggregate value of 0.44. The take-home point of this study is that while experienced epileptologists have very high confidence in their EEG interpretations, this confidence may not be founded in actual performance, and that further training to improve EEG accuracy and reliability is needed. One suggestion for accomplishing this goal is through continuing medical education modules that measure and report EEG reliability scores to individual users.

From earlier meta-analytic work on this topic, it should be noted that interpretive differences usually involve disagreements over ambiguous areas in which alterations in the frequency, symmetry, or amplitude of the intrinsic EEG approach what, by consensus, would be considered outside of normal limits. In such transition areas the dividing line is not sharply precise, and statements such as *borderline slowing* are sometimes used to denote the understandable uncertainty that is involved. When disagreements in this borderline region are removed from consideration, interpretive agreements tend to improve well into high ranges of reliability.

As a final point on reliability, the reader is referred to a detailed review and analysis by Robert Thatcher, published in 2010 in the Journal of Neurotherapy. Dr. Thatcher contrasts reliability results obtained through modern, digitally analyzed, Q-EEG techniques with clinical interpretation by traditional "eye ball" examination, which has remained the standard for clinical EEG examination since its inception. This review highlights advances in EEG analysis that have been primarily driven through scientific research, and clearly exemplify the advantages that could be achieved with greater acceptance of these methods in clinical practice.

EEG RECORDING AND RELATED METHODOLOGY

Much has been written about the complexities of EEG recording, interpretation, and the corresponding level of skill needed to obtain an adequate EEG. What may be insufficiently recognized is the fact that there are also clinical situations in which a greatly simplified EEG obtained by a properly trained registered nurse or resident can have substantial diagnostic usefulness. There are some important EEG findings of particular relevance to emergency room settings and, possibly, even to some acute psychiatric admission or triage units that can be obtained in only 10 minutes by clinicians with a minimum level of technical skill. Clinical cases presenting moderate to marked confusion and agitation, delirium, or possible nonconvulsive status epilepticus (NCSE) may have diffuse EEG abnormalities that are more or less continuous in the tracing, detectable almost as soon as the recording is turned on. Such findings do not require sophisticated localization studies, and their presence, as well as their absence, is diagnostically relevant. Prompt access to an EEG laboratory may not always be possible, especially on weekends or evenings, thus on-site screening may be most helpful.

Other than the circumscribed screening EEGs described above, recording the EEG does, in fact, require a considerable amount of skill and experience. It is not merely a technical act performed by a technician. The unfolding clinical EEG record is a constantly moving and shifting display of complex waveforms recorded simultaneously from numerous scalp locations, and the EEG patterns differ dramatically during wakefulness, drowsiness, and various sleep levels. The appearance of the EEG also changes from one recording montage to another, along with differences in the reference selected, while a host of normal and abnormal EEG waveforms and contaminating artifacts must be identified in their obvious and subtle forms. Identifying the source of artifacts and minimizing their intrusion on the EEG is most critical during the recording procedure, as application of offline removal and correction procedures require an advanced skill set and appreciation of intricacies of how such transforms may inadvertently change the appearance of features retained in the EEG. In addition to having the skills to correctly apply electrodes and operate

amplifier and recording equipment, the better technologists are also capable of sophisticated EEG interpretation. It may not be obvious how important this is. EEG abnormalities do not always emerge in clear-cut, textbook form but instead may be distorted and, hence, ambiguous. Interpretative ability is necessary to recognize probable abnormalities and then to arrange recording montages and states of patient alertness (wake or sleep) in ways that can enhance the patterns and allow a more definite interpretation by the electroencephalographer. A minimum of 1 year of full-time training, including didactic instruction and supervised, hands-on recording and interpretation experience, is necessary for an EEG technologist to achieve competence. Formal training schools for EEG technologists exist in many places, and the graduates can become registered EEG technologists by taking and passing a two-part written and practical examination.

Working in a psychiatry environment has its own challenges that technicians must be comfortable handling. First and foremost is being able to deal with a disturbed individual. It should be appreciated that clinical EEG is typically administered under conditions proximal to an acute change in mental state; therefore, psychiatric symptoms may be exacerbated. An EEG technologist who practices in a psychiatric setting must know when to terminate a procedure, call for help, and try to de-escalate a situation. Even the most cooperative psychiatric patients can also present challenges for EEG recording due to increased eye and involuntary body movements associated with tardive dystonia and akathisia. Muscle activity, particularly of the eyes, mouth, and jaw, produces changes in electrical activity on a much larger scale than brain activity and can contaminate an EEG record to the point where no reliable interpretation can be made. Given that good awake recording is necessary, sedation is not an acceptable procedure. On the other hand, the need for sleep in order to more completely assess *paroxysmal* activity, such as spike-wave discharges associated with seizure and other forms of neuropathology, mandates that all efforts be taken to get the patient to fall asleep. This may require decreasing the lighting and allowing the patient to relax. Room temperature in EEG recording environments is ideally kept between 70° and 75°F (21° to 24°C), a comfortable temperature at which an individual would not be expected to perspire. It is important to minimize changes in body temperature that could contribute to perspiration at the scalp, as sweat interacts with electrolytic gel applied to maintain conductance and reduce electrode impedance, produces slow drift and other artifact affecting EEG signal, and can contribute to “salt bridging” of adjacent electrodes (essentially connecting electrodes together so no differentiation can be made between signal recorded at the bridged sites). Perspiration and associated changes in skin conductance are also associated with increased sympathetic arousal, which should be recognized in agitated or highly

anxious patients. For these reasons, patients should be encouraged to remove heavy clothing or excessive layers if tested in street attire and, as stated previously, every effort should be made to keep the patient comfortable and relaxed. Explaining the purpose of the procedure, the EEG apparatus, and the signals obtained by EEG recording (e.g., an analogy to an EKG of cardiac activity but with electrodes place on the head to record brain activity) during electrode application can go a long way in helping the patient feel informed and at ease with the process. Careful procedures will prolong the recording time in psychiatric patients. However, in laboratories where psychiatric patients are not routinely evaluated, technicians may be inclined to hurry the procedure up instead of slowing it down.

EEG Apparatus

The basic components of the EEG recording apparatus are the (1) electrodes, (2) amplifier, and (3) recording device including means of visualizing, acquiring, and storing the EEG record. Each will be discussed in following sections.

Electrodes. As EEG emerged into the clinical arena, electrodes simply were placed on the scalp symmetrically by eye, using salient landmarks on the head as reference points, and not all laboratories used the same placement system. Eventually, in 1947, it was decided at an International EEG Congress held in London that some effort should be made to standardize the system of electrode placement, so that clinical and research findings would be more directly comparable across different laboratories. The challenge was taken up by Herbert Jasper, who developed the *10–20 International System* of electrode placement, which has become standard worldwide since 1958. The system is based on two anatomical landmarks for positioning; the *nasion,* identified as the depressed area between the eyes above the bridge of the nose and the *inion,* which is the highest point of the external occipital protuberance, located by a prominent bump at back of head near the lowest point of the skull. Briefly, measurement is taken from the total nasion-to-inion distance, and the midpoint is marked at a distance 50 percent between these points (and 50 percent of the transverse distance between the preauricular points on the ears). The 10–20 nomenclature refers to the distance between electrodes, positioned at either 10 percent or 20 percent of the actual anterior–posterior or transverse distance of the skull (Fig. 1.15–1A,B). Each electrode location is associated with an underlying area of cerebral cortex, indicated by the letters F, T, C, P and O, reflecting frontal, temporal, central, parietal, and occipital lobes, respectively, followed by either the lowercase letter "z" for midline electrodes or a number indicating homologous left- (odd numbers) or right- (even numbers) sided locations. Thus, electrodes O1 and O2, for example, are placed over the left and right sides of the occipital region,

respectively. Electrode Cz, also referred to as the vertex, is the most centrally located electrode and is used as the midpoint for alignment to the intersection of nasion-to-inion and transverse distance measures. The standard 10–20 system is based on a total of 21 electrodes, with 19 scalp sites including two electrodes at frontal polar sites (Fp1, Fp2), and two electrodes affixed at earlobes (A1, A2) which have traditionally served as a common reference for scalp electrodes.

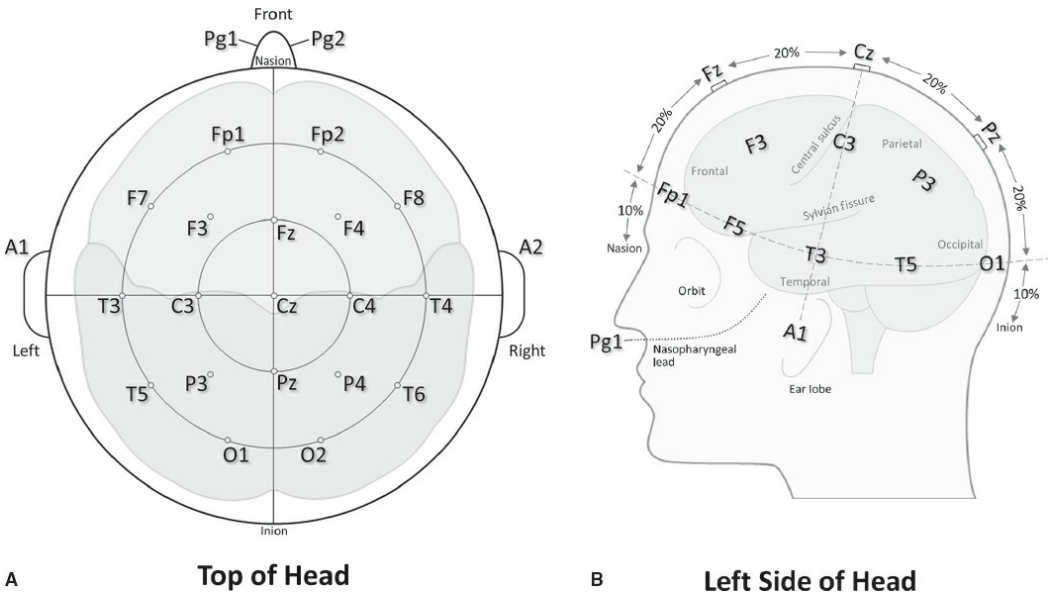

FIGURE 1.15–1. **A:** International 10–20 Electrode Placement System, as viewed from the top of the head. **B:** A left-lateral diagram of the head showing the locations of the routine 10–20 electrodes. Note the relationships between electrodes placed on the scalp and underlying brain structures.

Although precise localization with respect to underlying functional anatomy should not be assumed, some basic generalizations have been made concerning electrode location and the nature of brain activity detected in their region. For instance, the frontal–central electrode Fz is located in proximity to frontal lobe structures important to intention and motivational processes, while central sites near the vertex (C3, Cz, C4) are sensitive to sensory (particularly auditory registration) and motor processes, and occipital locations O1 and O2 to early stage visual processes. In addition, temporal sites (T3, T4, T5, T6) have shown sensitivity to memory and emotional processes, with temporoparietal electrode sites (P7, P8) typically selected for experimental studies on neural activity involved in structural facial encoding. Again, report of these findings is not intended to insinuate that EEG is an effective tool for localization of function to a known source, but rather to provide examples that illustrate the importance of multichannel recording and standardized electrode placement for characterizing unique and spatially distributed brain signals.

Although most laboratories use 21 scalp electrodes for standard

recordings, the 10–20 system enables expansion of electrodes to provide greater coverage, with additional electrodes typically placed at 10 percent of the distance between the original 10–20 locations. The American EEG Society (currently the American Clinical Neurophysiology Society [ACNS]) has developed a nomenclature for the designation of as many as 75 defined electrode locations. However, it must be stressed that extremely large numbers of scalp electrodes are unnecessary for currently established clinical EEG applications. For currently accepted clinical indications, optimally useful EEG recordings can be achieved with as few as 16 (minimum recommended standard) or, at the most, 32 scalp electrodes. While preparation and setup time must always be considered as a constraint on the number of electrodes that can feasibly be applied in a given recording situation, arrays of 128 or even 256 scalp leads are not uncommon in specialized research applications where aims of data collection include current source analyses and three-dimensional modeling (benefitting from additional electrode placements below scalp). An example of an extended 64 channel electrode array can be seen in Figure 1.15–2.

The 10–20 placement system has several advantages. Because the placement system is based on rigorous measurements, electrode placement error, particularly asymmetrical placement of electrodes for homologous electrode pairs, is greatly minimized. The system also renders recordings entirely comparable between laboratories, as well as across serial recordings obtained from a single subject. Because percentages of distances between landmarks on the head are used for placement locations, scalp electrodes overlie the same cortical regions despite differences in head size. Furthermore, the relationships between electrodes placed on the scalp and underlying brain structures have been well established (Fig. 1.15–1B) by using placements on cadavers (with holes drilled at the electrode sites to later identify the cortical area under the electrode), as well as studies using CT scanning.

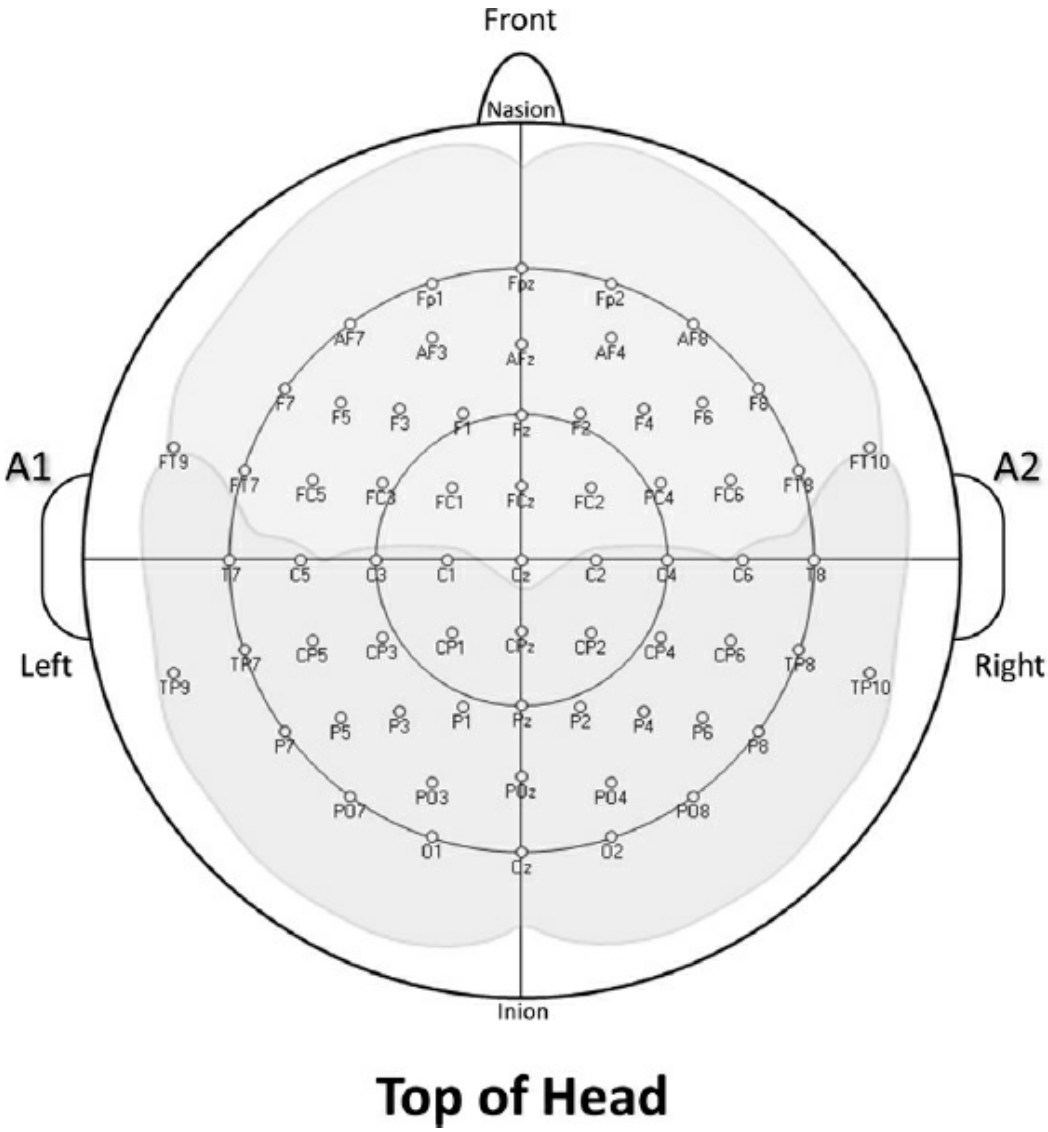

FIGURE 1.15–2. An expanded 64-electrode array. Such extensive placement systems are primarily used for special research studies and are rarely used for clinical recordings.

It has been suggested that in cases of suspected temporal lobe abnormality that is unconfirmed by conventional EEG recording, closer examination of the temporal area should be attempted. This is due to the fact that the anterior temporal lobe is not well covered by the standard 10–20 placement system. The F7 and F8 electrodes are over the posterior–inferior–frontal lobe and, hence, forward of the temporal pole, whereas the T3 and T4 electrodes are behind the anterior temporal region. Some laboratories now add new electrodes (T1 and T2) or simply relocate the F7 and F8 electrodes to this new position. The placement of the T1 and T2 anterior temporal electrodes is based on the distance from the lateral canthus of the eye to the external auditory canal, with electrodes placed at one-third of this total distance anterior to the auditory canal and 1 cm up from a line connecting these two landmarks. However, the F7 and F8 electrodes may detect potentials spreading from the anterior temporal

cortex, particularly if the voltages of the discharges are high.

Scalp electrodes must be applied carefully. The skin under the electrode must be clean and completely free of oil or grease. A common practice is to rub the area with a slightly abrasive cleansing electrolyte material that also removes some of the superficial epidermis. When this is done, a metal disc electrode can be applied to the scalp by using a conducting electrode gel. Electrode impedance should be maintained at equal to or less than 5,000 ohms (5 K Ohms) and closely matched across electrodes, ideally within 1 K Ohm of each other. The electrode application procedure should not be uncomfortable for the subject, however, scalp abrasion has typically been used to perforate the most superficial layer of skin as a method to reduce scalp impedance. Modern EEG systems designed for research enable high-impedance recording through the use of *active electrodes* (of note, active systems have not yet been approved for clinical use). Active electrode systems do not require skin abrasion, rather, acceptable signal-to-noise quality is achieved through preamplification at the electrode stage. Placing the first stage of amplification nearer to the scalp reduces the propagation of electrical interference through the shielded electrode cable to the amplifier, at which point line noise is amplified at the same scale as EEG signal. Active electrodes do not overcome the necessity to assure good conductance at the scalp site through careful setup and application of electrolyte gel, but may reduce overall preparation time by foregoing the need for scalp abrasion common to passive electrode systems.

SPECIAL ELECTRODES. Nasopharyngeal (NP) electrodes are inserted into the nostrils to achieve EEG recording with closer proximity to basal temporal lobe that can be obtained by scalp electrodes (Fig. 1.15–1B); these leads are designated Pg1 and Pg2 in the 10–20 placement system). No actual penetration of tissue occurs. The NP lead is a long (as long as 15 cm for adults), curved S- or Z-shaped insulated wire with a silver ball (the electrode) on the tip, which is rotated laterally after insertion in order to place the ball in contact with the roof of the nasopharynx. With a cooperative patient and a skilled technologist, the procedure can be well tolerated. Although this lead is presumed to be better positioned to detect activity from the orbitofrontal cortex, temporal pole, and hippocampus, it has numerous disadvantages. Chief among these are a high propensity to produce pulse and respiration artifacts and the fact that NP leads cannot be used when a deviated septum or nasal inflammatory process is present. They also are contraindicated with many psychiatric patients displaying behaviors, such as confusion, agitation, or belligerence who could pull the leads out, possibly lacerating the nasal passage. Their use also may interfere with obtaining a sleep-activated EEG and, not infrequently, otherwise cooperative patients simply refuse the procedure.

Sphenoidal electrodes use a hollow needle through which a fine

electrode that is insulated, except at the tip, is inserted between the zygoma and the sigmoid notch in the mandible, until it is in contact with the base of the skull lateral to the foramen ovale, the intended location. This is an invasive procedure that must be done by a physician and requires a signed consent form. While the goal of the sphenoidal electrode is to improve characterization of epileptiform activity with mesial–basal–temporal origins, the yield of positive results from these specialized electrodes over and above conventional scalp recordings has been somewhat controversial. Traditionally, sphenoidal electrode placement was conducted blindly, based only on anatomic measures and patient's report of mandibular pain, but radiographic studies suggested that this approach did not reliably produce accurate localization with respect to distance from the foramen ovale. Fluoroscopically guided placement was subsequently introduced to better ensure accuracy. A substantial body of empirical evidence suggests that in general, while not routinely used by all epilepsy centers, sphenoidal leads improve detection of seizure activity with anterior temporal origin over temporal lobe scalp recordings. In some studies, treatment decisions based on sphenoidal recording were improved by as much as 90 percent following inconclusive results from noninvasive scalp recording.

MONTAGE SELECTION. A common misconception is that the EEG records the voltage detected at each electrode site. Instead, each "squiggly line" on the EEG chart represents the shifting or oscillating difference in electrical potential *between* two electrodes. Thus, in a multichannel recording, the activity from each channel represents the shifting difference in voltage between two selected electrodes. When 16, 21, or more electrodes are placed on the head, the number of possible electrode pairs becomes large, and how these pairs are arranged among the recording channels can become complex. In EEG parlance, the way electrode pairs are arranged for a recording is called a *montage,* and although many montages are possible, only a limited number have become popular and useful. Prior to the advance of digital EEG equipment, selection of the montage was done prior to recording and would determine how the EEG was visualized in the paper record. When digital EEG equipment is used, montages are preprogrammed. Recording is performed using one montage, but the recorded (digitized) signal can be viewed offline in any configuration desired.

Montages are designed to facilitate the detection of EEG abnormalities in different brain regions and to facilitate comparisons between left and right hemisphere activity. There are general guidelines for how montages are to be set up. The most important rule is simplicity for the sake of interpretation. Additional rules include the stipulation that odd numbers refer to the left side, whereas even numbers refer to right-side electrodes (Fig. 1.15–1A). Furthermore, left-side electrodes are routinely displayed at

the top of or before right-side electrodes on the record. Similarly, anterior electrodes are displayed on top of or before more posteriorly placed electrodes.

There are two main types of montages: referential and bipolar. With referential montages, all electrodes are referenced to a single common reference point that traditionally consisted of linked ears (the mastoid prominence can be used in place of the ear lobe), with variations being left or right ear reference alone, ipsilateral ear reference in which all electrodes in one hemisphere are referenced to the ear on that side, or a contralateral ear reference in which all electrodes in one hemisphere are referenced to the opposite-side ear. Given the risk of unintended distortion of EEG (causing artifactual asymmetry between hemispheres) that can occur when using references based on bilateral sites, particularly if impedance is not carefully balanced between the two electrodes, single, midline, sites including the nose or vertex (Cz) have been recommended in some cases. An in-depth discussion of arguments in favor and against the various reference options is beyond the scope of this chapter, but suffice it to say that no one option will be ideal (or collectively accepted) for all EEG recording applications. A host of "reference-free" recording approaches have also been employed in digital EEG recordings to circumvent a choice of reference altogether. Reference-free montages are based on common average or other mathematical derivations of signal from all electrodes such as *current source density* (CSD), defined by the second spatial derivative of the local field potential (LFP) (for detailed discussion of field potentials in EEG the reader is referred to the text of Nunez and Srinivasan, 2006). As some suggested guidelines, selection of a reference electrode should entail a location that is electrically neutral, uncontaminated by obvious physiological (e.g., mastoid reference can be sensitive to cardiac signal if not properly located) or motion (e.g., nose reference can be sensitive to head movement if not properly affixed) artifact, and relatively distal to electrodes in the region of interest (e.g., EEG activity most proximal to reference is most extensively subtracted). A referential montage would most likely be chosen for applications in which it is important to ascertain the magnitude of an abnormality, in terms of how large a sharp wave or slow wave is in microvolts, in one region compared to another.

Bipolar montages, on the other hand, are useful (and, indeed, are much more widely used in clinical practice) for pinpointing the area of maximal abnormality or the exact source of abnormal activity. In bipolar montages, electrodes are referenced from one scalp location to an adjacent scalp location in chains of electrodes going across the head from front to back (Fig. 1.15–3) or from left to right (Fig. 1.15–4). The majority of abnormal cerebral activities tend to appear at the surface as negative potentials.

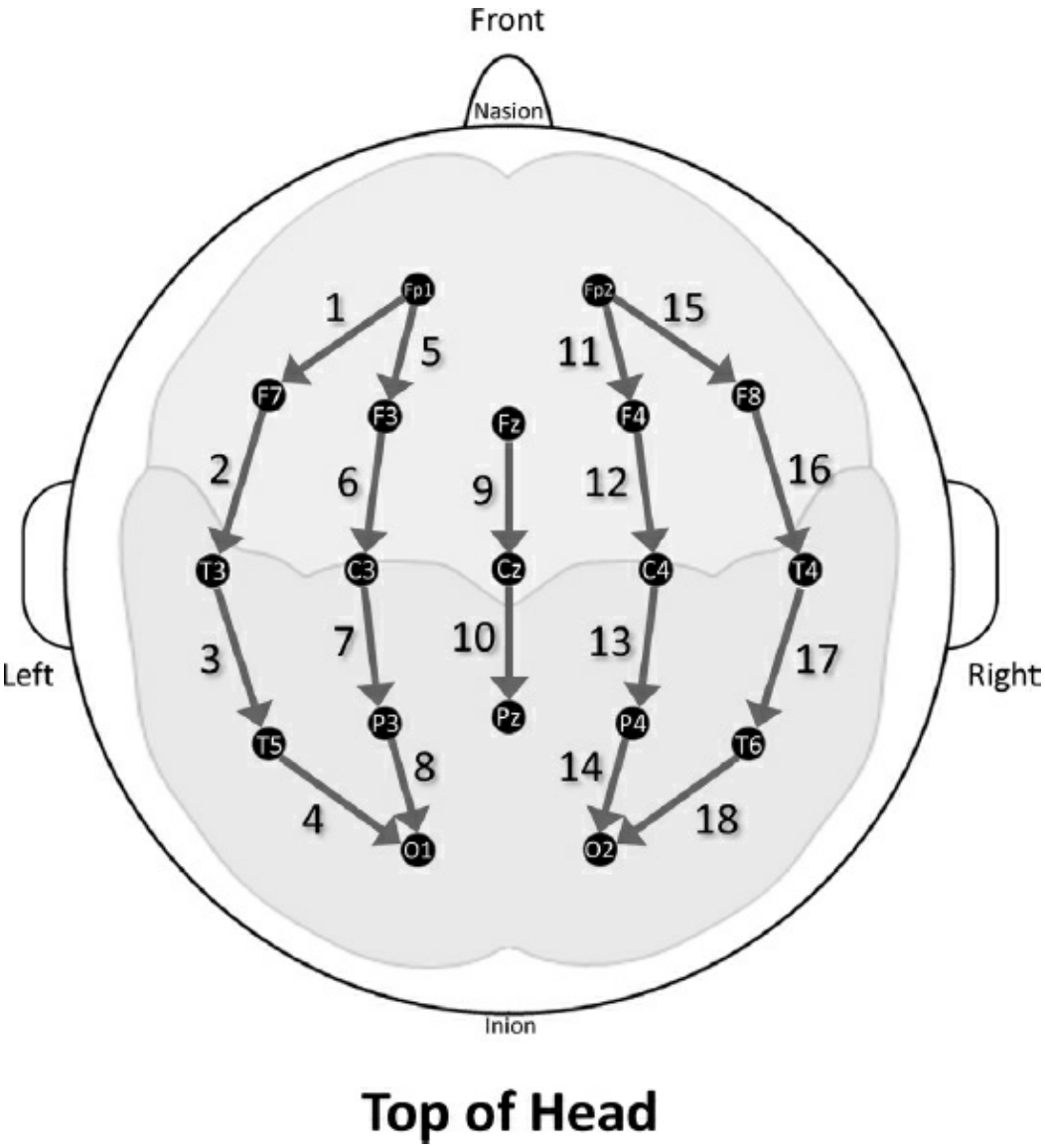

FIGURE 1.15–3. Example of an 18-channel bipolar montage with anterior to posterior linkages. The numbers between electrode locations designate recording channels. Thus the number 6 means channel 6, which measures the difference in electrical potential between F3 and C3 electrodes.

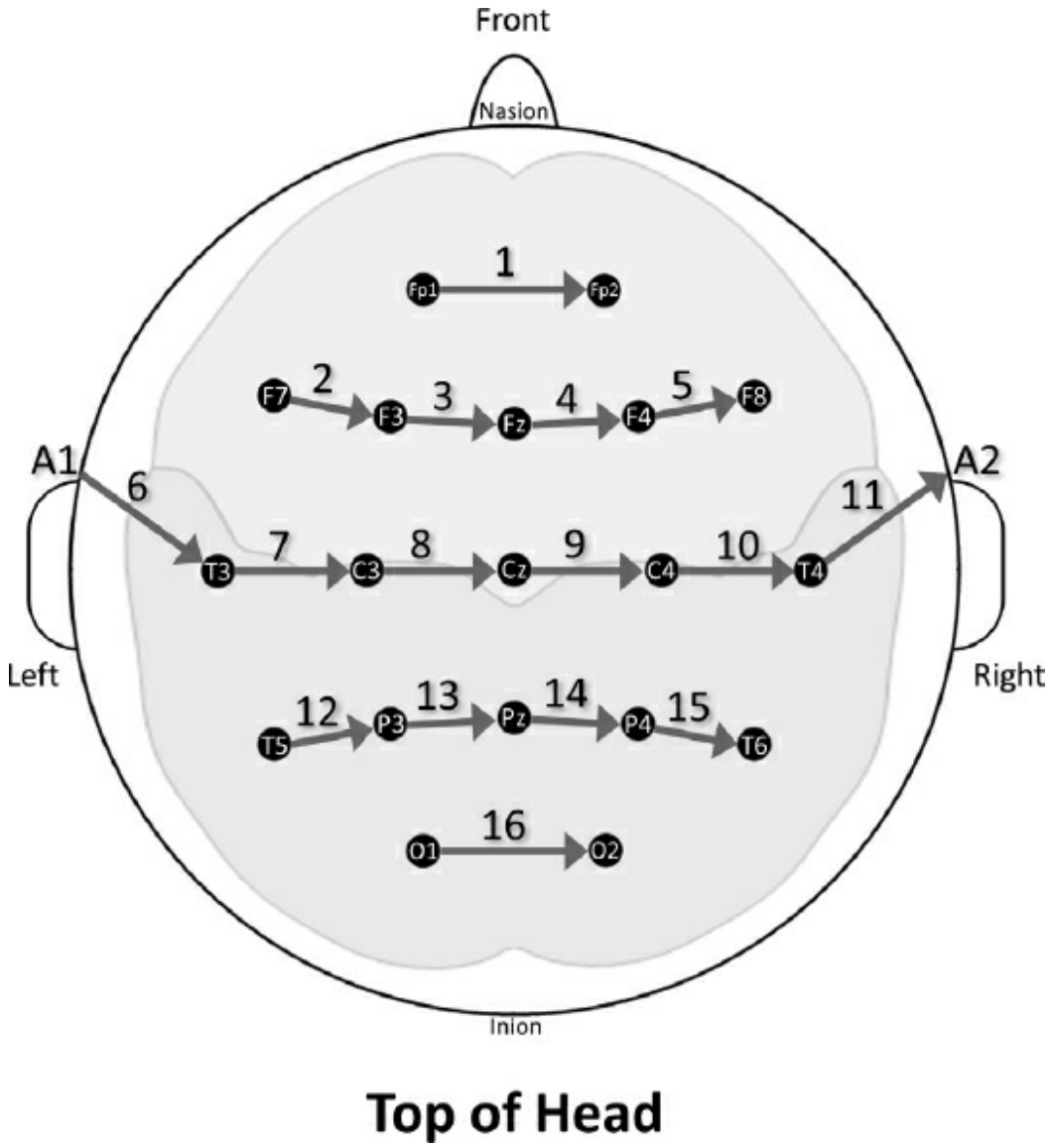

FIGURE 1.15–4. Example of a 16-channel transverse bipolar montage.

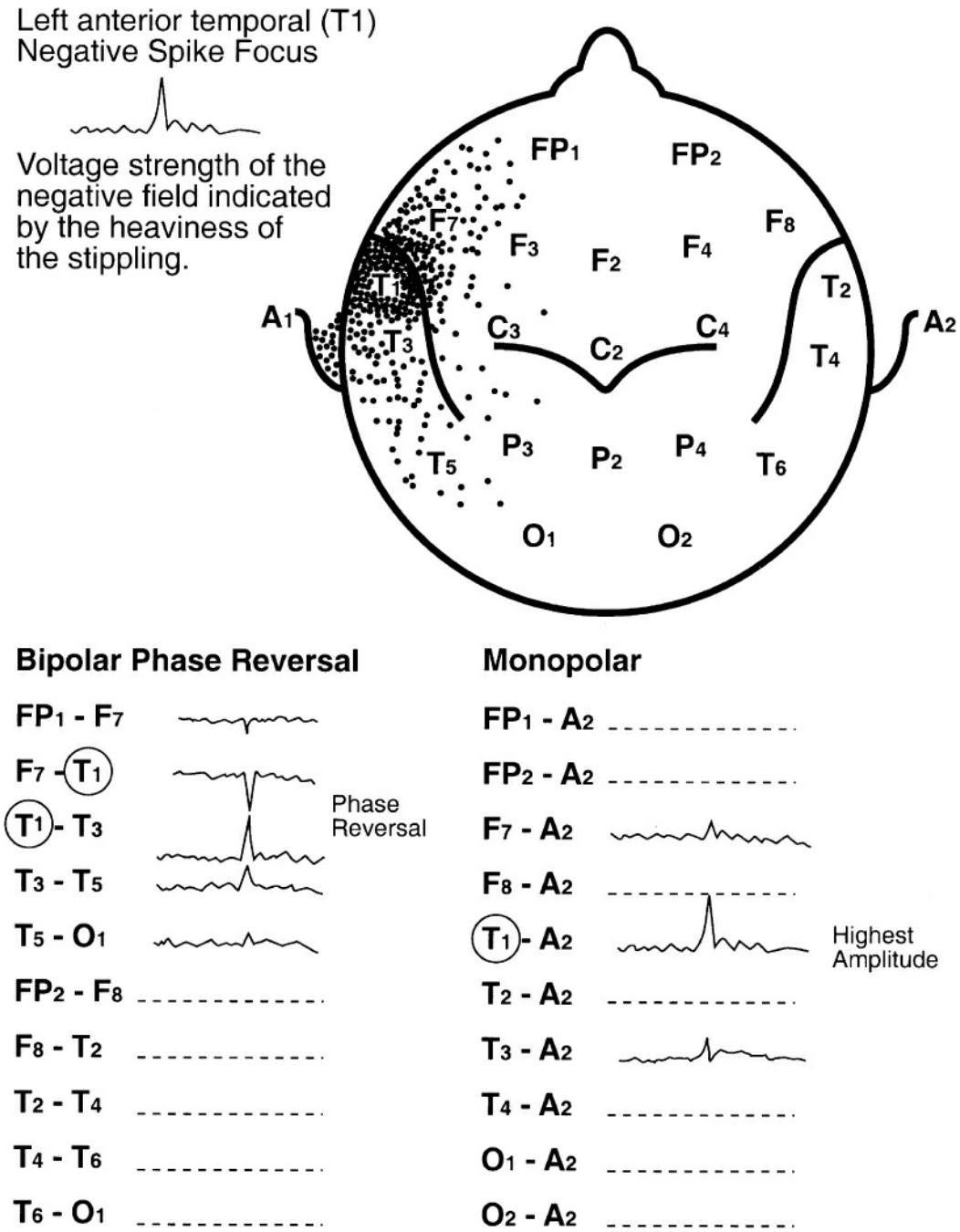

FIGURE 1.15–5. Illustration of bipolar (phase-reversal) and monopolar (highest amplitude) localization of a focal negative spike discharge at the left anterior temporal (T1) electrode. See text for explanation.

One can think of a given channel of EEG activity as being derived from two inputs.

By convention, the first electrode of a pair constitutes input 1, whereas the second electrode provides input 2. Thus, in the electrode pair C3–P3, for example, the first electrode C3 constitutes input 1. The direction of deflection in the EEG waveform is based on whether input 1 is, relatively speaking, "more negative" or "less negative" (i.e., relatively more "positive") than input 2. If the electrode designated as input 1 is closer to the source of

a negative field and, hence, more "negative" than input 2, an upward deflection (i.e., negative to positive/up) will be observed as an indication of input 1's closer proximity to the source of an event. Conversely, if input 1 is more distant from the source of the field than input 2 and, hence, less negative, then the deflection of the bipolar pair is downward (i.e., positive to negative/down). As bipolar pairs of electrodes move in a longitudinal or transverse direction from one side of a strong and highly localized negative field to the other side, the deflection changes direction across bipolar pairs as input 1 shifts from being relatively more negative to relatively more positive than input 2. This change of deflection is called a *phase reversal* and is a powerful method for localization of sharply focal abnormalities (Fig. 1.15–5).

One additional montage configuration deserves special mention, because it may be particularly useful in psychiatric EEG. This is a montage that combines referential and bipolar electrode arrangements. Following four bipolar connections from the frontal regions through the temporal areas and ending in the occipital region, a referential placement connects each posterior temporal region (T5 and T6) to the opposite ear. This arrangement allows activity of low amplitude to be highlighted by the referential electrodes for further examination via the bipolar electrode pairs. This montage is commonly referred to as the *Queen Square montage* (Fig. 1.15–6).

The appearance of EEG activity varies greatly from one recording montage to another. Large interelectrode distances often yields higher voltages, whereas a close spacing between electrodes in bipolar pairings tends to reduce voltage since both electrodes will overlie approximately the same portion of an electrical field and the signal common to both cancels out. Furthermore, specific EEG patterns visible in one montage may be distorted or even completely canceled out in another montage. Although some montages may permit a differentiation of activity between two or more brain regions, other montage choices may not do so. For example, EEG sleep patterns are well visualized and differentiated in central and occipital regions when a common (monopolar) reference recording is made. However, differentiation between central and occipital sleep activity is no longer possible when bipolar anterior–posterior linkages are used (C3–O1 and C4–O2), and, with transverse bipolar links between homologous electrodes, the sleep patterns may not be visible at all (Fig. 1.15–7). The issue is not merely academic. Discharges of interest to the electroencephalographer that are detectable in some recording montages may be completely or nearly undetectable, despite their neurophysiological presence, when a different montage is being used (Fig. 1.15–8). As previously mentioned, while traditional analog EEG required any change in montage to be done before recording, modern digital EEG is much more flexible in allowing changes of reference (i.e., re-referencing) or montage to

be done as an *offline* procedure (i.e., after data is recorded). In the case of digital EEG a change in reference does not alter the EEG record, only the appearance of the data for purpose of visualization.

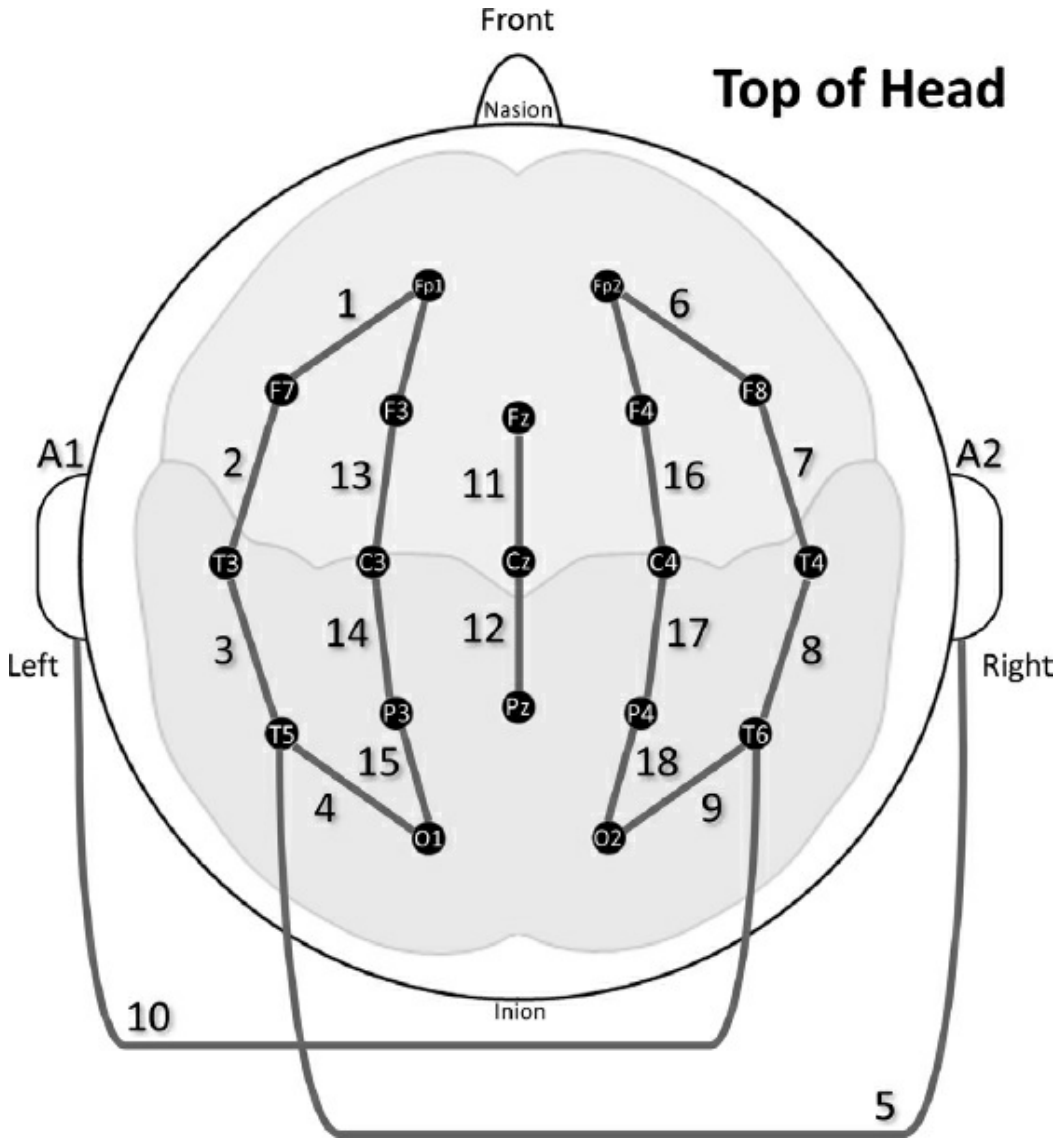

FIGURE 1.15–6. Diagram of the Queen's Square montage. This is an 18-channel montage modified to include two referential leads to highlight temporal lobe activity.

EEG Amplifiers. Biopotentials recorded at the scalp are very small and require amplification (typically with voltage gain of 1,000 to 100,000 times or 60 to 100 dB) in order to be visualized and, subsequently, analyzed. Indeed, scalp potentials are typically smaller in scale than many sources of interference, including nonencephalographic biopotentials (e.g., myographic, ocular, cardiac), electric line noise, and shifts in voltage that occur with poor electrode conductance at recording surface. Therefore, a critical requirement for an EEG amplifier is the ability to boost brain signal while minimizing amplification of these contaminating signals. This is achieved through use of a *differential amplifier* using two input sources,

one from the electrode and the other from a common system reference. The term "differential" refers to the manner in which signal is amplified, wherein the difference between the two input voltages is passed through to the amplification stage while any common voltage shared between inputs (regarded as noise) is rejected. The proportion of common voltage rejection relative to the amplified difference signal can be quantified as the *common mode rejection ratio* (CMRR), expressed in decibels (dB) as a unit of relative gain. As previously mentioned, the quality of EEG recording depends highly on the ability to achieve and maintain low impedance at scalp and reference sites. With regard to CMRR, it is especially important to ensure that impedance is balanced across electrodes, in that unequal impedance reduces the differential amplifier's ability to detect, and reject, common mode signals. Early analog EEG amplifiers required adjustment of *sensitivity* (i.e., ratio of input voltage to pen deflection) to find the range of amplification that would boost low-voltage signals high enough to be visualized, but prevent high-voltage signals from pushing recording pens above their deflection limits, where "squaring off" of waveform peak would occur. The necessity to adjust sensitivity, or gain, in this manner is eliminated in modern EEG amplifiers that employ analogue-to-digital (A/D) conversion, replacing pen and paper with software for control of EEG acquisition, visualization, and data storage. Most software allows input gain can be adjusted as needed for visualization at any point during EEG acquisition, with no consequence to the actual recording of the signal.

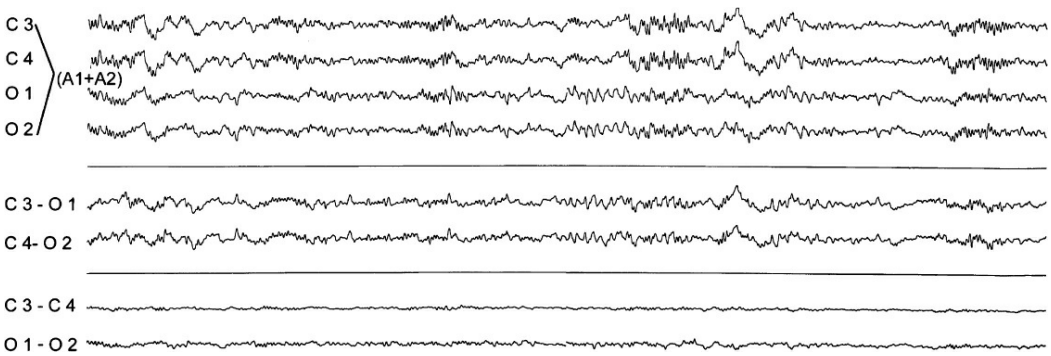

FIGURE 1.15–7. Alteration of appearance of brain waves (sleep patterns) with changes of recording montages. Note that the monopolar montage (*top four channels*) yields higher amplitudes and greater differentiation between central and occipital activity. Similar input to members of an electrode pair (C3–O1 and C4–O2) can reduce voltage in the bipolar derivation. Note the absence of differentiation between central and occipital activity in bipolar derivation. Note extreme cancellation of activity in the last two bipolar derivations.

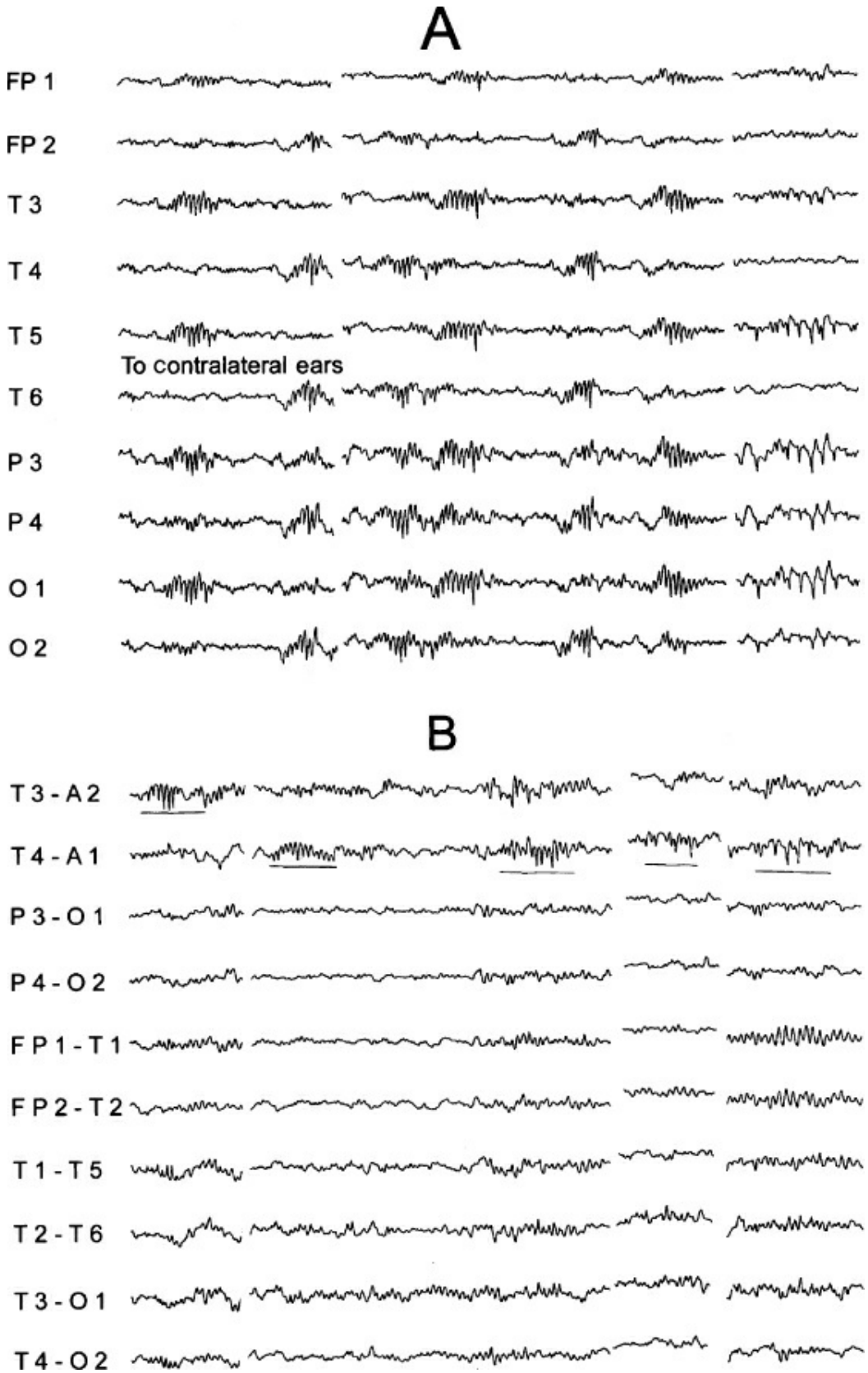

FIGURE 1.15–8. **A:** Fourteen-per-second and six-per-second positive spike discharges, independent left and right temporal–parietal–occipital area (monopolar montage). **B:** The top two channels show these discharges with the same monopolar montage as channels 3 and 4 in **A,** whereas lower channels show bipolar cancellation of the discharges even though all electrodes in montage **A** are present. The female patient was 32 years of age

with a closed head injury.

A feature found both in traditional analogue and modern digital EEG amplifiers is an online filter, placed before the sampling stage, where it functions to limit bandwidth of incoming signal in accord with sampling theorem. This filter is set to the amplifier's *Nyquist frequency,* defining the highest frequency of a discrete signal that can be accurately characterized at a specific sampling rate. High signal frequency relative to sampling rate can produce *aliasing* in the recording, a distorted rendering of a sinusoid that results from undersampling. *Nyquist theorem* dictates that the sampling frequency (F_S) must be at least twice the frequency of the highest frequency contained in the signal (F_C). Accordingly, Nyquist frequency is constrained by amplifier sampling rate, as expressed mathematically in the equation $F_{\text{Nyquist}} = F_C < F_S/2$, with both signal frequency and sampling rate measured in cycles/samples per second using the common unit Hertz (Hz). Thus, the first stage of filtering may be regarded as an, "*antialiasing filter,*" used to attenuate EEG frequencies at and above the limit that can be accurately sampled. Aliasing has become less of a concern over time as amplifiers used in clinical applications sample minimally at rates of 256 to 512 Hz, which is more than sufficient to characterize EEG within the range of most human brain functions (e.g., ~1 to 100 Hz). Indeed, modern digital amplifiers are capable of sampling well beyond 1,000 Hz, and usually up to 20 kHz, though doing so comes at the cost of much larger data files, with higher computer processing speed and storage space requirements, and little useful information gained. One notable exception to this generalization is testing of the auditory brainstem response (ABR), which can include events sampled at rates of 7 kHz to 50 kHz using specialized amplifiers.

Digital Recording. Over the last 30 years, the analog EEG traces of the old EEG machines have been largely replaced by computer systems providing an analog to digital conversion of the recorded EEG signals, as well as amplification, digital filtering, storing, and quantitative analysis of multichannel EEG. Digital systems are now commonplace in laboratory and clinical settings, are relatively inexpensive, and offer several practical advantages. They can help reduce the space problem of storing paper records over many years and also allow review of an EEG record using multiple montages, filters, and vertical (gain or sensitivity) and horizontal scaling (e.g., paper speed) selected after the original recording. Moreover, digital recording systems allow many new and exciting possibilities for EEG signal processing and quantification that rely not only on time-series analysis but also spectral-frequency, time–frequency, related methods of coherence and phase synchrony, and spatial–temporal source localization and modeling (to be discussed in subsequent sections).

FREQUENCY FILTER SETTINGS. Nearly all of the EEG activity that is analyzed for clinical or research purposes falls within the frequency range of 0.5 to 50.0 Hz (and occasionally up to 100 Hz). Filters are used essentially to limit the range of frequencies to a bandwidth containing features of interest. It has become conventional in clinical EEG to use a *low-pass filter* setting of 70 Hz, which means that EEG frequencies above this filter setting become progressively attenuated while those below the filter setting as allowed to "pass" through and are retained for further analysis. At the other end of the spectrum, most laboratories also employ a *high-pass filter* between 0.5 and 1.0 Hz to reduce the registration of frequencies below this level. Because the EEG is represented as a two-dimensional image of amplitude over time, with the visualized waveform representing a summation of all frequencies, a *band-pass filter* (combines a high- and low-pass filter) may be used to isolate discrete frequency bands (e.g., alpha band, 8 to 12 Hz) of interest from the broadband signal.

Unfortunately, scalp electrodes pick up a variety of electrical potentials of noncerebral origin, and many of these have frequencies within or close to the EEG frequency spectrum. Frequency filters may, to some degree, mitigate against the distorting effects of frequencies generated by artifacts. However, filters must be used judiciously and with caution, because they also can filter out real brain waves that one wishes to observe. For example, slow waves generated by changes in skin conductance during sweating (galvanic skin response) can be of sufficiently high magnitude to completely obliterate genuine EEG activity in the affected recording channels (bilateral frontal–anterior temporal areas around the forehead are particularly prone). Raising the high-pass filter setting to 5 Hz effectively eliminates this source of contamination in the recording (Fig. 1.15–9) but does so at the expense of attenuating any real generalized or focal slow activity that also may be present. It is also common practice to adjust the low-pass filter downward from 70 to 35 or even 15 Hz to eliminate or reduce unwanted electromyographic (EMG) activity from the recording (Fig. 1.15–10). Again, the choice to do this involves a compromise, because attenuating high-frequency activity may obscure the accurate detection of certain fast spike discharges. Of note, adjusting filters in this way is permissible for clinical applications, with the goal of identifying significant features at the individual level. Naturally, the selection of filter settings in research applications must be kept standard across data samples, as not to systematically introduce differences that would invalidate the measure, and would be based primarily on resolution of the frequencies important to the experiment or neurophysiological processes being evaluated.

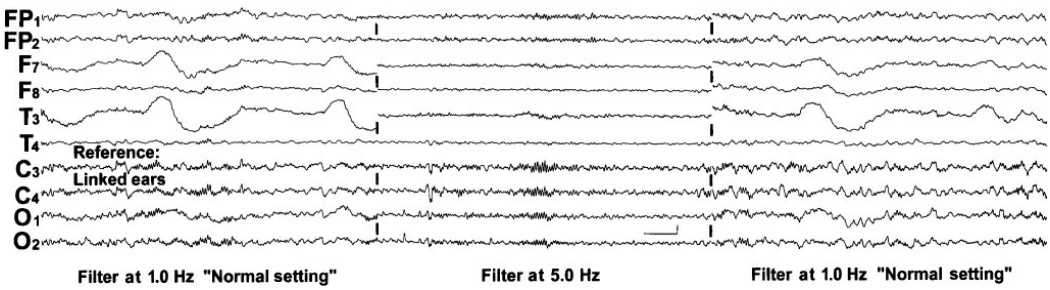

FIGURE 1.15–9. Effect of low-frequency filter setting on perspiration artifact (F7 and T3 electrodes) during sleep recordings. Adjusting low-frequency filter upward to 5 Hz completely eliminates the slow rhythm artifact and also eliminates the normal slow wave components of sleep but does not alter the faster 14-Hz sleep spindles.

EEG CHARACTERISTIC FEATURES

While the appearance of the raw EEG may vary somewhat between individuals, the range of frequencies, voltages, and waveforms that characterize the normal EEG during wake and sleep has been well established. The normal EEG tracing is composed of a complex mixture of many different frequencies, and based upon modern spectral analysis, consistent recognizable patterns can be observed in most healthy individuals. Some frequency bands are expressed more strongly over some cortical regions than others (e.g., alpha over occipital cortex). Moreover, the frequency profile varies considerably as the subject being recorded transitions from alert wakefulness into sleep. Following the lead of Berger, discrete frequency bands within the broad EEG frequency spectrum are designated with Greek letters. However, it should be noted that the sequence of Greek letters is not entirely logical, and in fact, derives from historical precedents. For example, the term alpha rhythm discovered by Berger was chosen because it was the first rhythm detected and described. By contrast, the designation theta derives from the fact that activity in this frequency range was initially thought to originate in thalamus.

Normal Intrinsic Frequencies

A brief discussion of the most established and recognizable *intrinsic rhythms* follows. Examples of each of these rhythms can be found in Figure 1.15–11. Intrinsic rhythms, which provide descriptions, examples, and spectral analysis (time × frequency) for each of the rhythms described below. For a thorough treatment of this topic, please refer to Chang, Schomer, and Neidermeyer (2011).

Delta. Delta activity (≤3 Hz) is not present in the normal waking EEG, but is a prominent feature of deeper stages of sleep, particularly stage 3. A slower delta rhythm (<1 Hz) is also observed when an individual is under anesthesia. The delta rhythm has been shown to be primarily generated by thalamocortical circuits. The presence of significant generalized or focal

delta in the waking EEG is strongly indicative of a pathophysiological process.

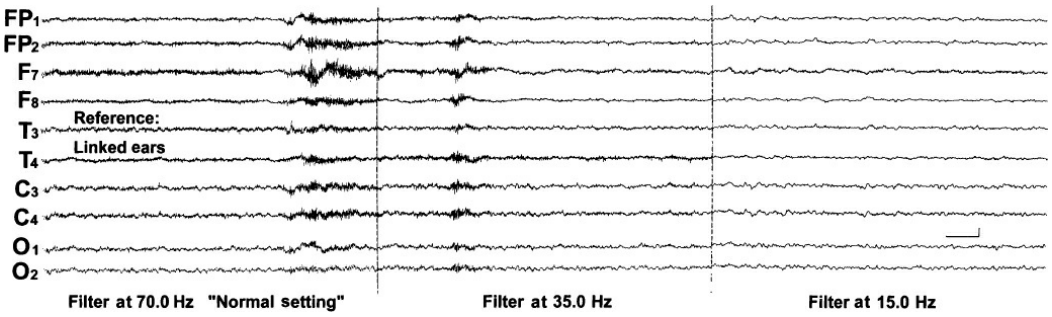

FIGURE 1.15–10. Effect of adjusting high-frequency filter setting on muscle potential artifact (generated by having the patient grind his teeth repeatedly). Muscle potential seen at the "normal" filter setting of 70 Hz is attenuated when the filter setting is lowered to 35 Hz and completely removed when it is set at 15 Hz. Lowering the high-frequency filter introduces the risk of attenuating or removing (i.e., filtering out) abnormal spike discharges from the tracing.

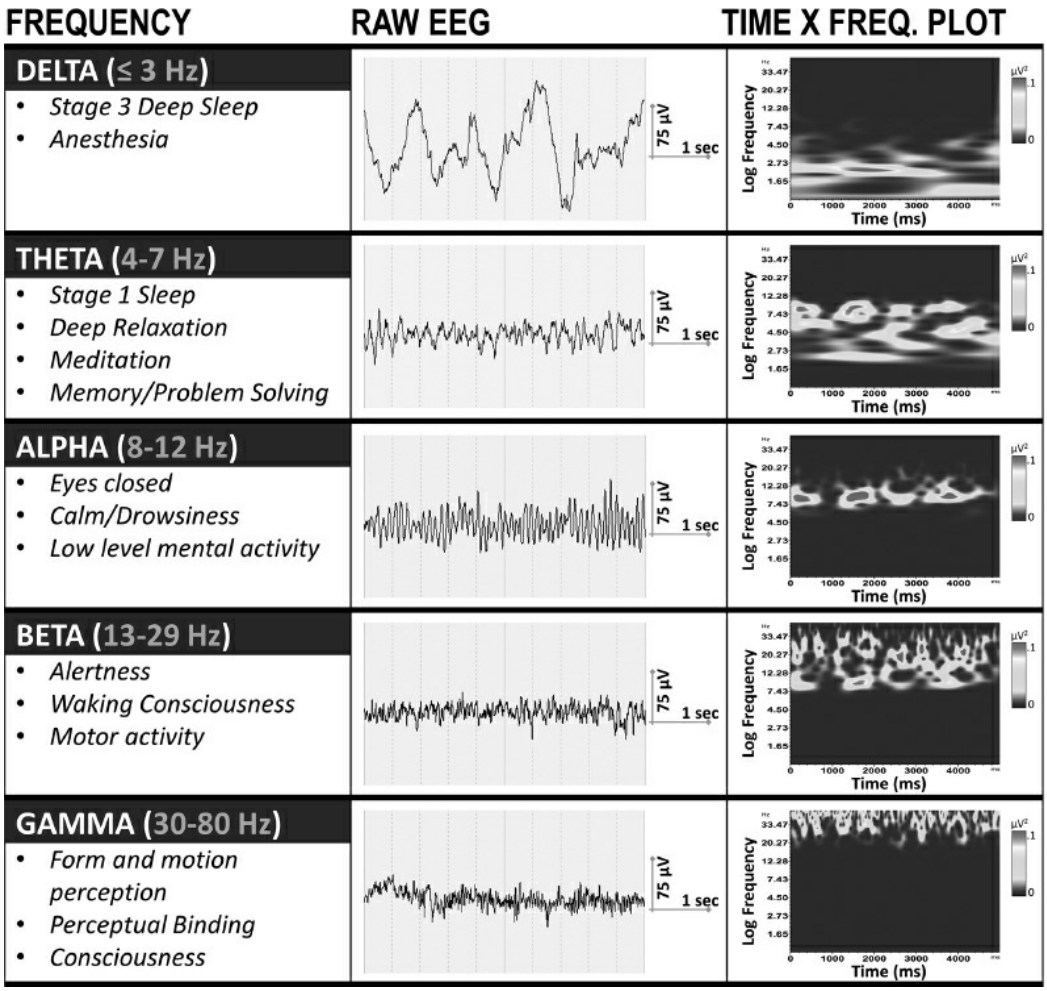

FIGURE 1.15–11. Example of the typical intrinsic brain rhythms observed in the normal EEG. The middle panel illustrates a prototypal example of each rhythm in the raw EEG. The far right panel demonstrates how each rhythm would appear after a time × frequency transform.

Theta. EEG with a frequency of 4 to 7 Hz is referred to as theta activity. A small amount of sporadic, arrhythmic, and isolated theta activity can be seen in many normal waking EEGs, particularly in frontal–temporal regions. While ongoing theta activity is somewhat limited in the waking EEG, it can be a feature of drowsiness during sleep tracing, and is highly prominent in stage 1 sleep. Brief periods of theta activity have also been shown to be associated with mental activities such as problem solving and processes related to memory. The fact that theta appears to be generated in hippocampus supports the role of theta in memory. Theta rhythms are also generated in the neocortex, particularly during working memory tasks. Excessive theta during waking states, generalized or focal in nature, suggests a focal pathological process such as CBF reduction or metabolic encephalopathies.

Alpha. Highly rhythmic EEG with a frequency range from 8 to 12 Hz is termed alpha activity, and was the first brain rhythm to be described by Berger in 1929 (i.e., the Berger rhythm). Alpha constitutes the dominant brain wave frequency of the normal closed-eye waking EEG, and is also common in states of relaxation or during periods of minimal mental activity. Hence, alpha has routinely been interpreted as constituting an EEG correlate of an "idling brain." In middle-aged adults, the vast majority of individuals exhibit an alpha frequency at, or close to, 10 Hz, whereas, with normal geriatric populations, a slower alpha frequency of 8 to 9 Hz is not uncommon. Alpha activity is most prominent over the posterior cortex, particularly the parietal, posterior temporal, and occipital cortex, with the occipital region being best suited to detect this activity. In fact, the neural generators of alpha activity appear to be visual cortex, with some influence from visual thalamus. The detection of alpha activity diminishes as one records from more anterior locations, and this frequency is rare at prefrontal electrode sites. Alpha activity is abolished, or at least severely attenuated, by eye opening. Further, alpha activity also disappears as one enters sleep (indeed, one of the criteria for stage 1 sleep is <50 percent alpha activity per epoch). Alpha activity can be highly responsive to cognitive activity, such as focused attention or concentration. For example, under closed-eye recording conditions, alpha can be blocked or attenuated by engaging in visual imagery, numeric calculation, or almost anything requiring significant concentration (Fig. 1.15–12). Alpha frequency can be increased or decreased by a wide variety of pharmacological, metabolic, or endocrine variables.

Beta. Frequencies in the 13 to 29 Hz range are termed beta waves, and they are one of the dominant frequencies in normal adult waking EEGs, particularly over frontal–central regions. Beta rhythms are typically associated with alertness, and it is not uncommon for beta to appear as

runs of rhythmic activity as opposed to sporadic isolated waves. Mixed frequency EEG, including beta, is also observed in REM sleep. In addition, research has shown that beta frequencies are correlated with peripheral EMG activity, suggesting that rhythmic beta in the motor cortex plays a key role in the generation of motor commands. The voltage of beta activity is also almost always of a smaller amplitude than that of the lower frequencies described above. Hence, along with gamma (see below), beta rhythms are not as easily visualized in the raw EEG compared to delta, theta, and alpha (Fig. 1.15–11).

Gamma. One of the most studied EEG frequencies during the past several years is the gamma rhythm, which is typically considered to be activity in the 30 to 80 Hz range (note that some investigators extend gamma up to 100 Hz). Gamma rhythms were largely ignored throughout most of the history of EEG research, which was due to in part to (1) recording methodologies, and (2) the notion that such higher frequencies were nonneural in origin (i.e., noise). However, seminal work in the 1990s by researchers such as Rodolfo Llinás, Andreas Engel, Wolf Singer, and others, demonstrated that transient gamma rhythms play a key role the integration or "binding" together of neural networks in order to generate coherent cortical representations. Since that time, the role of gamma rhythms in human sensation, perception, and cognition has gained wide support. Abnormalities in gamma rhythms are now thought to be a core feature of several psychiatric disorders, most notably schizophrenia.

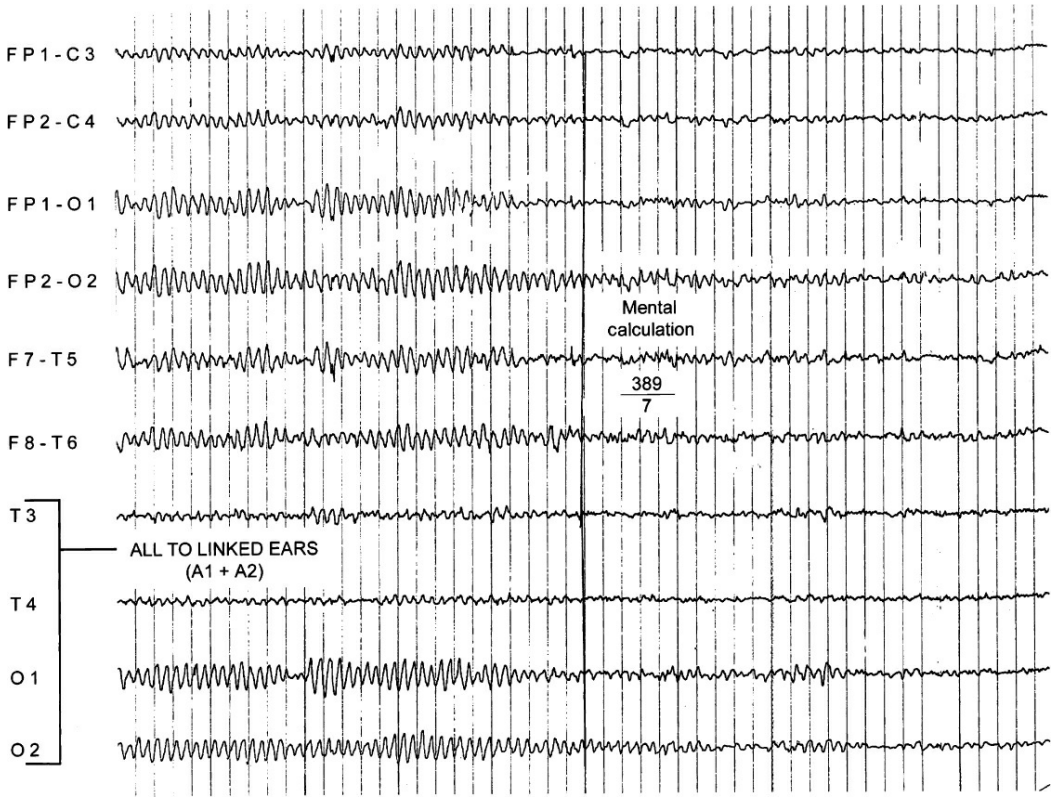

FIGURE 1.15–12. Resting, eyes-closed, awake electroencephalography recording. Effect of mental concentration on alpha activity. While instructed to keep the eyes closed, at the heavy vertical line, the subject is asked to divide 389 by 7. Note immediate blocking of alpha activity.

EEG Artifacts

EEG artifacts, as previously alluded to, refer to any source of recorded activity that is not of cerebral origin and, therefore, contaminates accurate characterization of the desired signal. The various sources of artifact are unfortunately abundant, but can generally be reduced to two sources: *physiological* and *extraphysiological*. Physiological artifacts are real biosignals recorded from the patient, but generated by sources other than the brain. Extraphysiological artifacts are those caused by external sources, including EEG or other medical equipment, and typically electronic devices located on the person or in the recording environment. Familiarity with the appearance of each artifact in the EEG, and its associated source, is critical for the EEG technician as most of these contaminating signals can be mitigated during recording but tend to be much more difficult to remove from the recorded data.

Physiological Artifacts. Important physiologic features to be aware of during EEG recording and interpretation are ocular artifact, muscle artifact, perspiration, and cardiac artifact, though others (e.g., lingual motion) also exist. Eye blinks and vertical shifts in eye movement can be prominently observed in the EEG record, identified by rapid changes of voltage over brief intervals over frontal electrode sites, particularly the frontal poles (Fig. 1.15–13). Blinks and *vertical electrooculogram* (VEOG) responses can be further captured in EEG recoding by affixing electrodes above and below the orbit of one eye as a bipolar channel. Similarly, horizontal eye movement, or "saccades," produces a characteristic artifact that appears as a change in voltage that tends to flatten over a brief period of time with phase reversal observed between left- and right-sided frontal electrode channels (the side of positive deflection indicates the direction of eye movement). Like VEOG, *horizontal electrooculogram* (HEOG) responses are detected in scalp EEG channels (particularly F7 and F8) but optimal characterization benefits from a bipolar channel of electrodes placed near the outer canthi of the eyes. Fortunately, several "blink correction" algorithms have been developed and can be employed in digital EEG data to remove VEOG/HEOG artifact from the EEG record (Fig. 1.15–14).

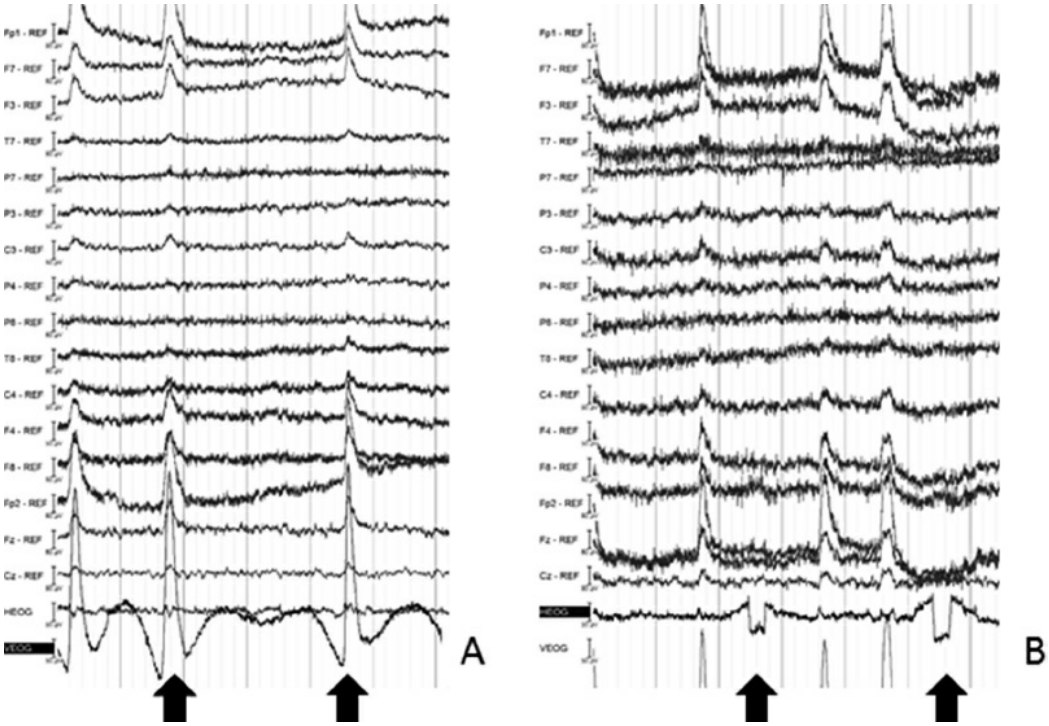

FIGURE 1.15–13. Electrooculographic (EOG) artifact. **Panel A:** Vertical EOG (VEOG) activity caused by eye blinks. VEOG recorded from electrodes affixed above and below orbit of right eye, also evident in frontal poles (Fp1, Fp2) and adjacent frontal sites with attenuated signal distributed posteriorly. **Panel B:** Horizontal EOG (HEOG) causes by lateral shifts in gaze, recorded at electrodes affixed bilaterally at outer canthi of eyes.

EMG activity of the temporalis muscle, indicating jaw clenching or chewing, and of the frontalis muscle, indicating tension and contraction of the forehead, appear as high-frequency artifact that can spread across a large portion scalp electrodes (Fig. 1.15–15). It is more difficult to correct EEG for this type of EMG artifact offline, and could result in rejection of large segments of data if prominent. As mentioned previously, a low-pass filter may be used to attenuate EMG artifact but use of this option may equally affect frequencies of interest. Therefore, it is best to reduce this form of artifact during EEG recording by training the patient to relax facial muscles beforehand and stopping recording, as necessary, to remind them. Obviously, anything that contributes to use of these muscles, such as squinting to read in a dark room, chewing gum, or talking, should be avoided during EEG recording.

Perspiration artifacts (galvanic skin response) can be detected by a slow, undulating, drift in signal over long durations (e.g., 2 seconds), thus are easy to differentiate from cerebrally generated activity, however, also resembles slow roving eye motion. Perspiration artifact is caused by sweat (sodium chloride) interacting with electrolytic gel and electrode metals and tend to be most prominent over frontal electrode sites (Fig. 1.15–16). Properly temperature-regulated recording environments and instructions to patients to dress in layers that can be removed as needed are the best

solutions for avoiding perspiration, which can be difficult to reduce once it begins. Less frequently seen are artifacts from an electrocardiogram (ECG), though certain physical characteristics of the patient, such as a short neck or heavy, "barrel-chested," build, tend to increase their presence. ECG artifacts reflect both electrical and mechanical actions of the heart, and can be detected by a consistent, rhythmic pattern with a period that aligns closely to pulse rate (thus, near 1 Hz), and are observed most commonly in temporal electrode sites and references placed at ears or mastoid (when electrodes placed too low and nearer to carotid artery). Although ECG artifact is not likely to directly affect many scalp EEG channels, the intrusion of ECG into reference channels could give this appearance during data acquisition using a reference montage or after data has been re-referenced offline. These problems are then best avoided by moving reference channels, when possible, to a location where ECG signal is not picked up.

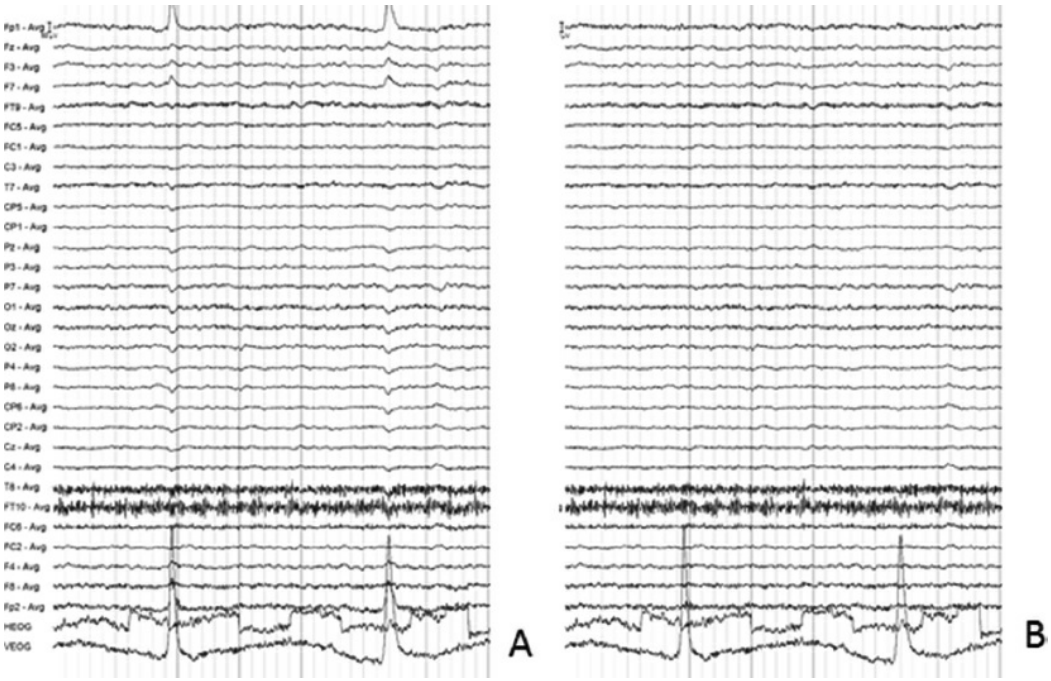

FIGURE 1.15–14. Ocular correction. **Panel A:** EEG shows ocular motion artifact on VEOG and HEOG channels with propagation to scalp electrodes. **Panel B** shows the same data following offline ocular correction by Gratton & Coles method. Note that correction is applied only to scalp electrodes in this case, leaving the original artifact present in VEOG and HEOG.

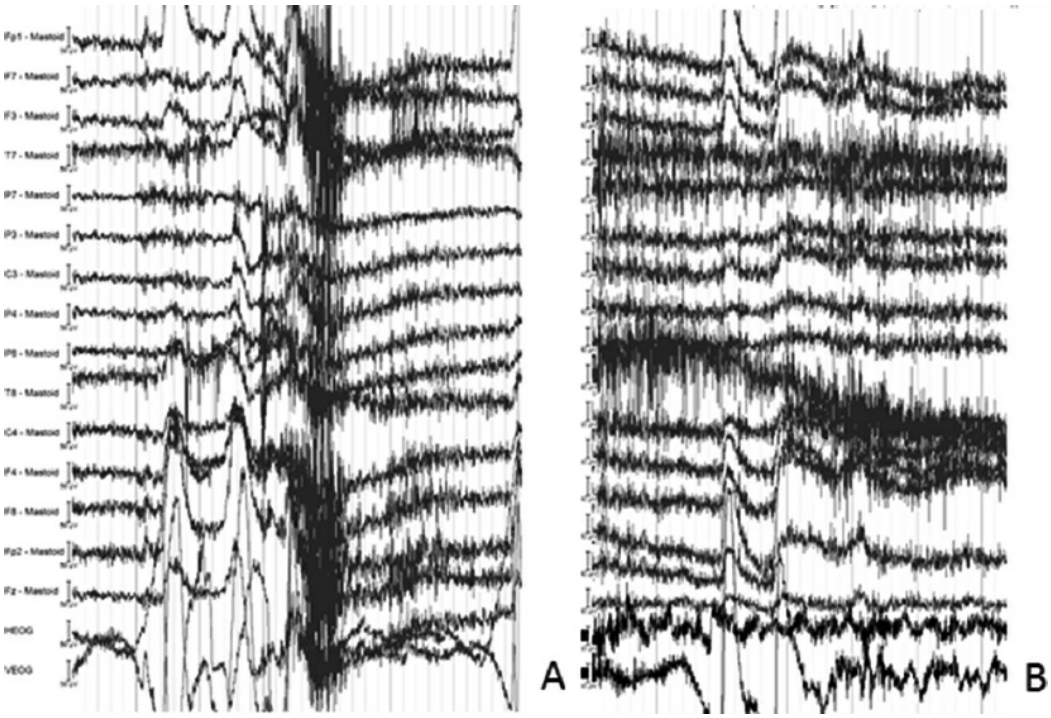

FIGURE 1.15–15. Electromyographic artifact. **Panel A:** Jaw (temporalis) muscle tension. **Panel B:** Forehead (frontalis) muscle tension. These artifacts can typically be ameliorated by asking subject to relax facial muscles when observed.

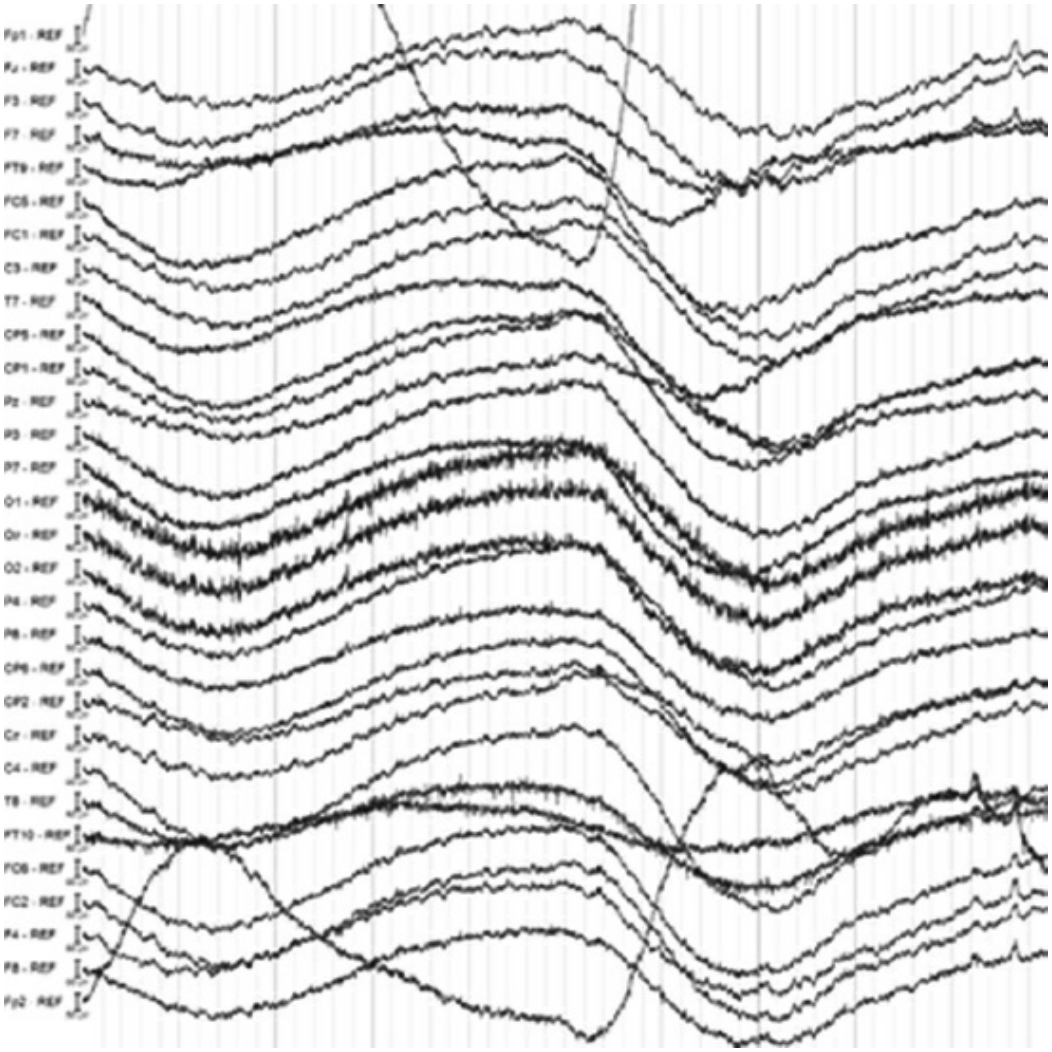

FIGURE 1.15–16. Sweat drift artifact in EEG record referenced to mastoid. Drift at frontal poles (electrode Fp1, Fp2) suggests sweat on forehead. Drift common across remaining scalp electrodes suggests sweat at mastoid reference electrodes.

Extraphysiological Artifacts. The main sources of extraphysiological artifacts have to do with equipment and environmental issues that can be minimized through careful attention to proper electrode setup, cleaning and maintenance of equipment between uses, and mitigation of ambient electromagnetic noise sources in the recording room. Electrical devices such as computer monitors, lights, and poorly shielded power cables emit 60 Hz signal, the line frequency of AC delivery in the United States (50 Hz in most other parts of the world). Electrodes with high or unstable impedance, due to poor contact with gel/scalp or break in lead will be particularly susceptible to picking up ambient noise in the recording environment, which is identified by persistent 60 Hz activity that appears as a fast sine wave imposed over background EEG (Fig. 1.15–17). Some EEG acquisition software packages enable viewing EEG in spectral frequency mode, which can quickly identify the characteristic peak at 60 Hz if there is any doubt.

Maintaining proper electrode contact over long EEG recordings may require reapplication of gel, which can always be a first troubleshooting step for addressing with noisy channels. Detection of a connection problem, which may also be caused by electrode movement, is not difficult at as the resulting effect on recorded signal in no way resembles true EEG (Fig. 1.15–18). In these cases, there tend to be very large changes of voltage, which may "rail" out of range at gain settings used for regular EEG monitoring, with no regularity in the signal.

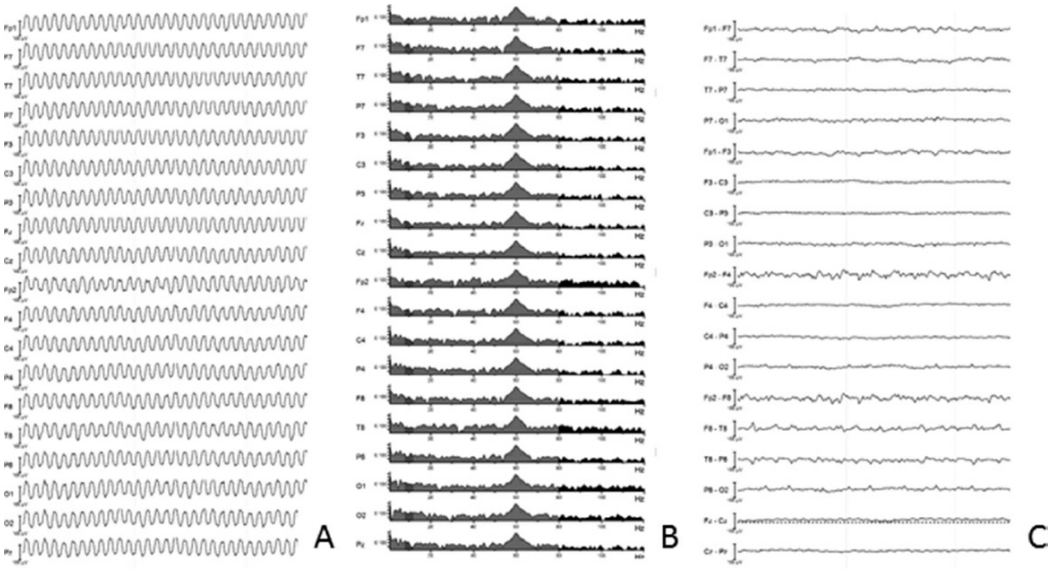

FIGURE 1.15–17. Panel A: 60-Hz ambient noise in unreferenced EEG viewed at range of 100 µV. **Panel B:** Power spectral analysis by Fast Fourier Transform (FFT) depicting prominent 60 Hz activity in Panel A segment. **Panel C:** Panel A EEG segment viewed in bipolar reference montage at 100 µV. Note that 60-Hz activity remains present in EEG but cancels out as common signal when referenced.

One common artifact that is more difficult to prevent is the electrode is a "pop," appearing as a steep and rapid deflection in voltage, usually isolated to a single electrode (unless a reference or a channel included in a bipolar montage) and occurrence (or may reoccur a few times sporadically), that is caused by the discharge of electrical potential between the scalp and electrode where electrical charges can store (as if in a capacitor). The pop is the only extraphysiological artifact mentioned here (though this list is not exhaustive) that may resemble abnormal EEG activity, with some characteristics similar to epileptiform activity. Importantly, pops can be distinguished from true neurophysiological paroxysmal activity by the fact that they tend to occur in only one electrode at a time, thus do not give the appearance of voltage change from an electrical field source, and appear with an asymmetric morphology characterized by relatively steep rise and slow recovery. These characteristics should differentiate pops from interictal spikes to the trained eye.

EEG Quantification and Feature Detection

EEG recording, combined with stimulus-evoked paradigms and experimental measures of mental task performance, provides powerful methodology for studying neural dynamics of human brain function and pathology. EEG data is dimensional and complex, based on a time series of events sampled with high temporal resolution (i.e., millisecond level) and distributed spatially across multiple scalp locations (e.g., montages of 32 to 256 channels). If one considers a sampling rate of 256 samples per second, across 32 scalp channels, for 10 minutes of recording, each parameter representing a lower bound of what is possible for digital EEG recording and typical for a recording session (which may last hours), the resulting data matrix could conceivably include 4,915,200 discrete data points. While examples of clinical EEG discussed thus far would not entail a detailed analysis of each sample point by electrode (this would never be done under any circumstance), with objectives that are routinely achieved by visual inspection of data over time scales of several seconds, such abundant data no doubt contains important features that cannot be ascertained (or even seen) through visual inspection and clinical judgment alone. The advent of digital EEG recording technology, among other advantages over analog paper EEG, makes highly detailed EEG analysis possible through computerized signal processing. Basic processes such as *segmentation,* dividing the time series record into shorter periods or epochs based on experimental events, and *averaging* over multiple segments (consistencies in signal average together while noise averages out), provide efficient means of isolating important features in data, upon which a number of additional signal extraction methods can be applied. Digital EEG analysis is broadly referred to as "*Quantitative EEG*," abbreviated to *Q-EEG or qEEG,* emphasizing measurement of the magnitude and differentiation of features contained within the EEG rather than qualitative decisions over whether or not an observed feature indicates a pathophysiological process. However, even while the heavy computational burden is now passed on to the machine, Q-EEG still requires the same human judgments concerning data quality and interpretation. When the issue of artifact contamination of the Q-EEG is not properly addressed, one has a situation described a generation ago by computer programmers, namely, "garbage in–garbage out," leading to erroneous conclusions of the processed data. Software used for Q-EEG analysis includes automatic or semiautomatic functions for removal of artifact based on thresholds for acceptable amplitude or change in amplitude over time, which can be used to mark and remove some sources of EEG contamination. However, these *artifact rejection* procedures cannot typically mitigate all problematic conditions, such as low-frequency activity of a drowsy patient, and visual data inspection should always be

undertaken as a first step of analysis.

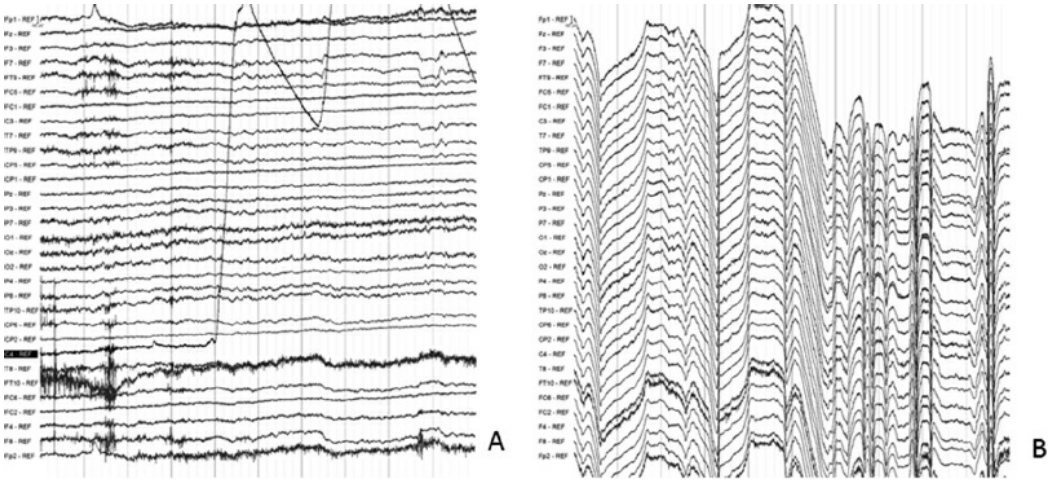

FIGURE 1.15–18. Physical electrode movement in referential montage. **Panel A** shows effect at single electrode (C4) suggesting electrode is being physically shifted in position and unseated from point of scalp contact. **Panel B** shows consistent pattern of shift across all electrodes, indicating movement at the reference electrode site.

Beyond the initial data cleaning, the technician or experimenter is faced with many more technical decisions regarding the selection of time points and signal extraction methods used to best characterize the neurophysiological phenomena under examination. It is sensible, if not required, for some processing steps to be taken before others, but each involve a number of decisions over parameters for which there is not common agreement or standardization. Furthermore, there is usually more than one way to implement a given function, which may differ between software packages, and the user (typically lacking extensive training in computer science and engineering) may be unaware of the intricacies and assumptions involved in each. There are also a number of free, open-source, toolkits available for EEG analysis that enable the inclined user to customize various analysis routines. Therefore, while digital recording and analysis greatly increases the capacity to explore and extract features from EEG, the application of such methods remains to be more of an exquisite art than a standardized practice.

The most commonly used method for EEG quantification is the spectral analysis, principally based on the *Fourier transform*. The Fourier transform decomposes a signal, represented as a function of time, into its constituent frequency components. Given the analogy of a chord played in music, playing the chord as an arpeggio (e.g., the A major chord played as individual notes A-C#-E) enables the listener to appreciate the frequency and loudness of each note as a component of chord's overall harmonic quality. In a similar way, the Fourier transform represents signal as a measure of *spectral power* (square of signal amplitude) according to *total power* of the broadband (e.g., chord) and *band power* of individual

frequencies (e.g., notes). In EEG signal processing, the Fourier transform is most commonly implemented as the *Fast Fourier Transform* (FFT), which applies an efficient algorithm for computing the Fourier transform achieved by reducing the total number of samples submitted to computation by their base 2 logarithm. Applying FFT therefore assumes that the data represents a fixed length waveform with a total number of samples that can be expressed as a factor of 2-to-the-nth power (e.g., 64, 128, 256, 512). Frequency resolution of the FFT is also dependent on the length of the data sampled, with longer segments of data needed to resolve lower frequencies. For example, the lowest frequency that could be characterized accurately using a segment length of 1 second is 1 Hz, while resolving 0.1 Hz signal would require segment length of 10 seconds. When using FFT, it is also important to keep in mind that the analysis treats all samples within the segment equally without regard to time, therefore, is insensitive to time-varying functions and most appropriate for signals assumed to be stationary. Quantification of changes in spectral power that occur over a period of time is made possible through *time–frequency decomposition*. A variation on the FFT, called the *short-time Fourier transform* (STFT), has been developed for this purpose, but EEG analyses typically employ more sophisticated algorithms based on the *wavelet* transform. The theory and mathematical computation involved in the wavelet transform is beyond the scope of this chapter, but an eloquent description written in context of EEG analysis can be found in the methods handbook of Todd Handy. The basic point to keep in mind in differentiating FFT, or time to frequency transformation, and wavelet, or time–frequency transformation, is that frequency extraction by FFT applies to the entire time window of analysis (signal is not localized in time) while wavelet analysis provides localization of frequency extraction to each sample of the time window. The latter is advantageous when examining change in frequency component structure as signals evolve over time, and so tends to be the more appropriate analysis for perceptual-cognitive experiments. That said, there are still cases where FFT may be preferred given that time resolution does come at a cost to frequency resolution.

In both the clinical and research literature, several types of time–frequency outcome measures are often discussed, which vary depending the paradigms used to elicit them and the signal processing techniques utilized. Most commonly, oscillatory activity measured by time–frequency decomposition is differentiated into three categories: transient evoked, steady-state evoked, and induced (Fig. 1.15–19).

Q-EEG using frequency-based analyses has contributed largely to the appreciation of EEG as an ensemble of neurophysiology generated oscillatory components with unique functional properties. For instance, it has been suggested that the delta rhythm, traditionally regarded as a pathological finding in the awake EEG, has been related to internal

concentration (i.e., conditions in which subjects have to disregard external input). Further examples include the association of frontally generated theta activity with encoding and working memory functions, the slow alpha band (7.5 to 9.5 Hz) with attentional processes, fast alpha band (9.5 to 13.5 Hz) with memory retrieval, and beta activity with the integration of sensory information. Furthermore, functions associated with frequencies above 30 Hz, the so-called gamma band, have been more recently discovered and are currently a major focus in psychiatric research. The demonstration of a stimulus-induced gamma oscillatory activity in the mammalian brain was the first empirical evidence that synchronization of neural groups (individually coding for different aspects of the sensory input) serve a basic "binding" process in the CNS, enabling information communicated over sensory afferents to combine in forming a unified percept. While high-frequency gamma synchronization is thought to support communication of short-range local networks, slower oscillatory rhythms have been implicated in cognitive processes requiring integration over large cortical distances (e.g., among areas of different sensory modalities) that extend over longer periods of time (e.g., memory encoding).

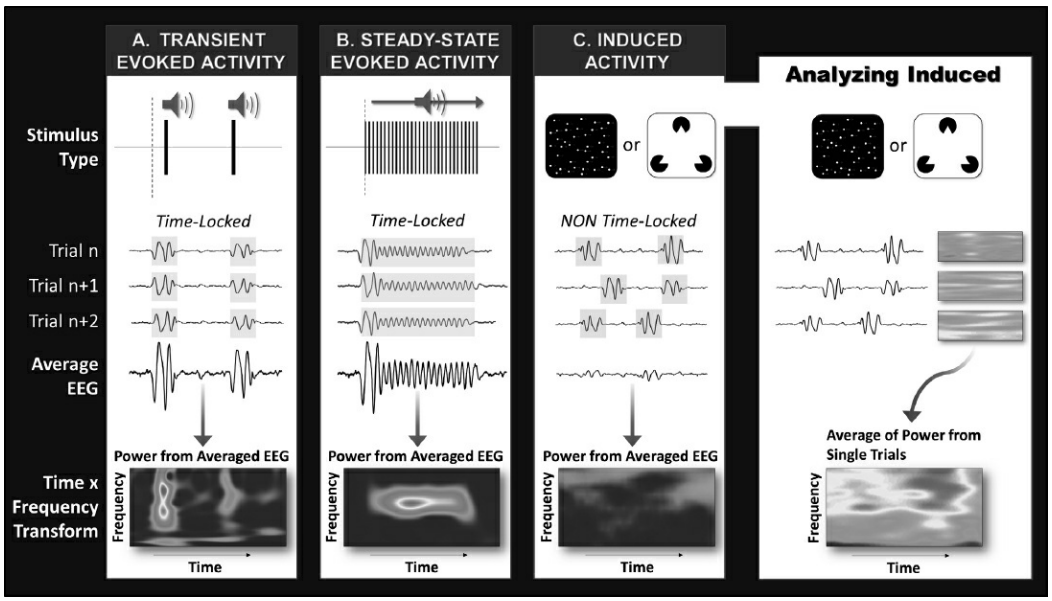

FIGURE 1.15–19. Transient and steady-stated evoked activity are phase- and time-locked oscillations (**A** and **B**) elicited by discrete or steady-state sensory stimuli. For both of these types of activity, averaging across trials has the effect of cancelling out non–phase-locked oscillations (due to destructive interference), leaving the evoked response intact and magnified. Hence, time–frequency analysis can be performed on the averaged EEG (i.e., after averaging all the individual trials). By contrast, induced activity is typically elicited during higher perceptual processing, for example, during the perception of coherent motion or of illusory shapes (see **Panel C**). These responses are often non time-locked (jittered across time from trial to trial). Hence, averaging the trials cancels out the response of interest due to destructive interference, and is not visible in the time–frequency transform (bottom of **Panel C**, left). However, these responses can be detected if time–frequency analysis is performed on the single trials, and the time–frequency transforms of individual trials are then averaged (**Panel C**, right). (See color atlas.) (Adapted from Skosnik PD, Cortes-Briones JA. Electroencephalography (EEG) and cannabis: from event-related

potentials (ERPs) to oscillations. In: Preedy VR, ed. *The Neuropathology of Drug Addictions and Substance Misuse*. Amsterdam, Netherlands: Elsevier Science; 2016.)

Time- and frequency-based analyses mentioned thus far have paved the wave for more precise characterization of neural activity and function, however, may fall short in providing a fully accurate representation of neural communication in that the information extracted by these methods remains spatially localized (e.g., analyzed based on 1 electrode or field of electrodes). Furthermore, while these methods effectively quantify the magnitude of neural activity as an average representation of time, the degree of consistency in activation across time segments is lost. Relationships between frequency ensembles located at different spatial locations or occurring at different time points, determined as a measure of synchronization, can be ascertained using *coherence* measures. Coherence analysis follows frequency (e.g., FFT) and time–frequency (e.g., wavelet) decomposition of signal, and is applied to a complex form of these data as a measure of strength of association in the linear dependence of two signals. Coherence analysis can be applied to stationary and time-series data as a measure of relationship of spectral power between two regions (i.e., *spatial coherence*), dependence of one frequency on another (i.e., *cross-spectrum coherence*), and of consistency in the phasic cycle of oscillatory activity over repeated trials (i.e., *intertrial coherence*). Taken together, the combination of spectral power and coherence measures greatly enhances the ascertainment of neural dynamics comprising the human EEG, and increase capability to understand the functional interactions of brain areas on a large spatial scale.

Q-EEG Topographic Analysis and Source Analysis

Computerized EEG topography (CET) or Q-EEG mapping was developed for multichannel Q-EEG data. The technique enables the construction of a bi- or three-dimensional matrix for a topographic representation of Q-EEG parameters, such as instant amplitude or band power (Fig. 1.15–20).

Matrix points that do not correspond to recording electrodes are calculated by interpolating the values of the nearest three or four recording electrodes. As for other imaging techniques, statistical probability maps can also be implemented for comparisons between different subject populations or experimental conditions. Color-coded maps make the multichannel information more immediately intelligible; however, the topographic information provided by the technique is critically dependent on the reference electrode used for EEG recording (Fig. 1.15–21). To address this issue, several transformations have been proposed that provide reference-independent data, including average-reference (i.e., subtraction of the mean of all electrodes from each electrode value) and surface Laplacian or CSD approaches previously mentioned. However, it is

imperative to realize that the topography at the scalp, even when resulting from reference-independent calculations, does NOT allow physiological inferences on the underlying brain generators, due to the so-called inverse solution problem. Essentially, this problem appreciates the imprecision of localizing sources of field potentials as recorded at the level of the scalp, since numerous neural activation patterns can generate the same scalp topographies (Figs. 1.15–21 and 1.15–22).

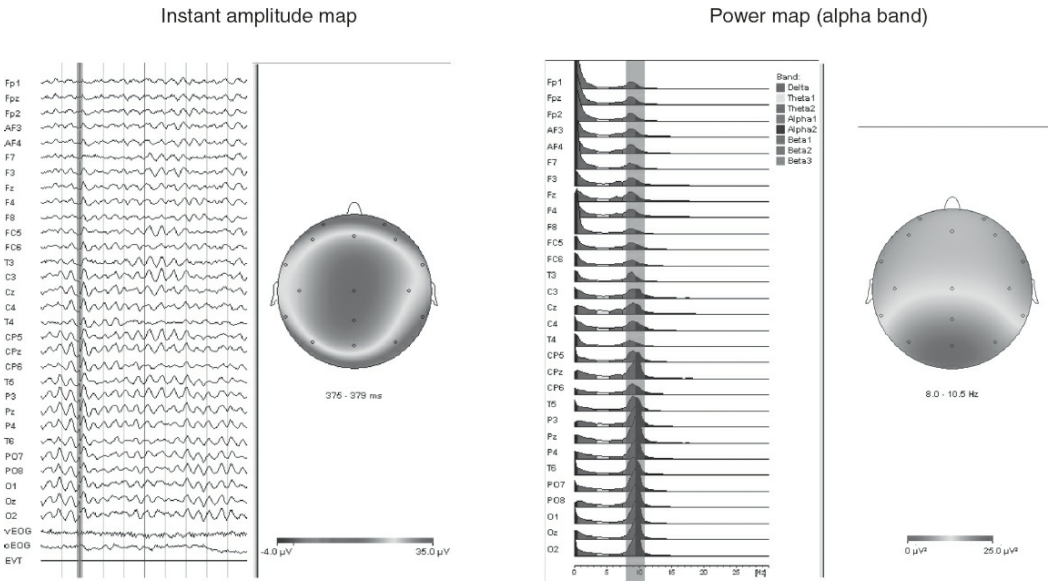

FIGURE 1.15–20. Quantitative electroencephalography maps. Instant amplitude map on the left and alpha absolute power map on the right. Units are µV and µV^2, respectively.

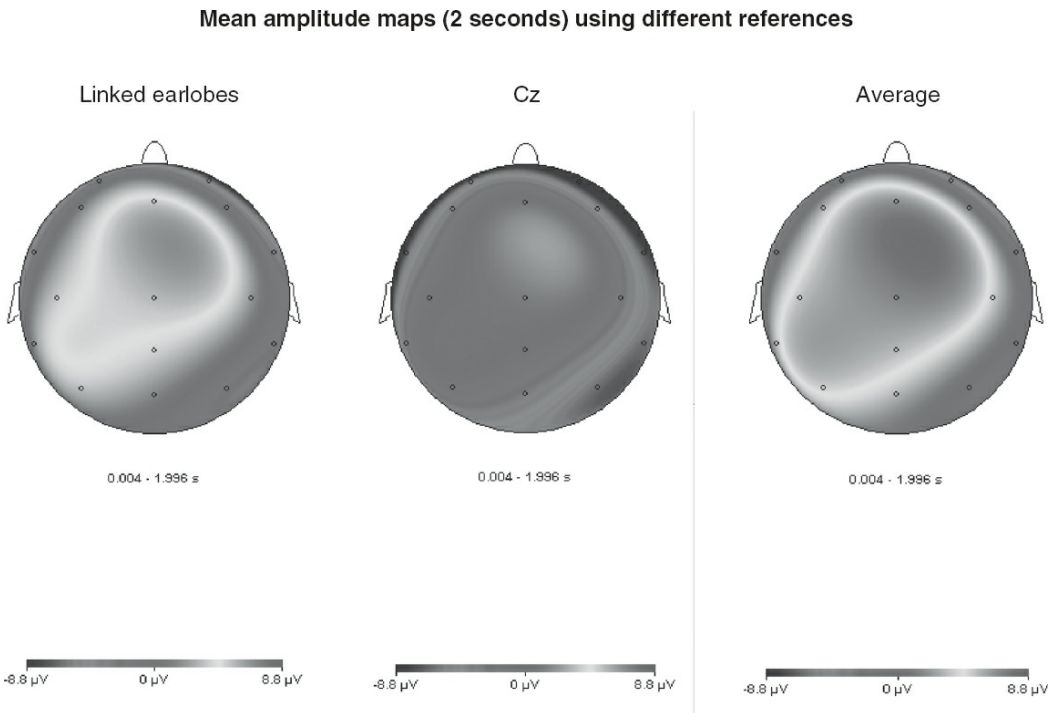

FIGURE 1.15–21. Quantitative electroencephalography (EEG) maps. Mean amplitude maps obtained from the same segment of EEG data (2 seconds) using different references. Left-side map, linked earlobes reference; middle map, Cz reference; right-side map,

average-reference. Units are µV.

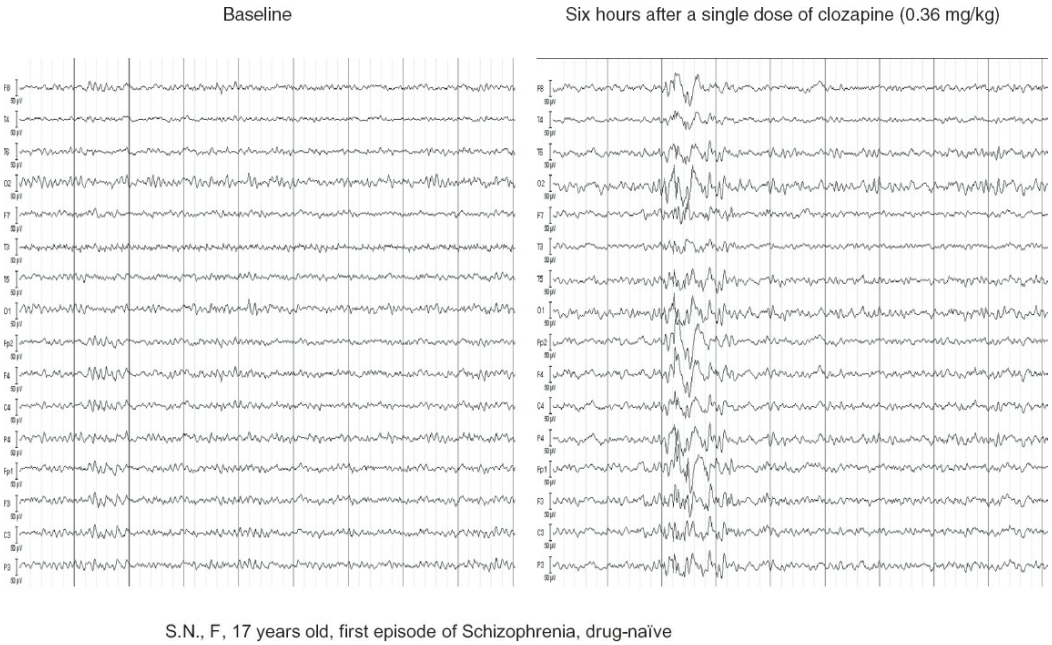

FIGURE 1.15–22. Clozapine-induced electroencephalography (EEG) alterations. On the left a sample of the baseline EEG of a 17-year-old, first episode, drug-naive patient with schizophrenia is shown. On the right a sample of the EEG recorded in the same subject, six hours after a single dose of clozapine (0.36 mg/kg) is shown. A high-amplitude, slow wave complex, with maximal amplitude over the right hemisphere, is observed. In both EEG samples, the interval between two vertical lines is 1 second.

Thus, in 1989 Frank Duffy, Dietrich Lehman, Fernando Lopes da Silva, and other eminent scientists formed the International Society for Brain Electromagnetic Topography (ISBET), established to promote the application and further development of techniques investigating brain electromagnetic activity, topography, and source analysis. In recent years, different algorithms have been proposed to solve the inverse problem. They can be divided into *equivalent current dipole models* and *current distributed source models*. According to recent comparative studies, the dipole models are suitable only when a single source is expected. Among the distributed source methods, *low-resolution brain electromagnetic tomography* (LORETA), or more recently, *standardized low-resolution brain electromagnetic tomography* (sLORETA) have proven greater effectiveness in terms of localization error, particularly when multiple sources and noise are present. Briefly, the sLORETA method is a properly standardized discrete, three-dimensional distributed, linear, minimum norm inverse solution. The particular form of standardization used in sLORETA endows the tomography with the property of exact localization to test point sources, yielding images of standardized current density with exact localization, albeit with low spatial resolution (i.e., neighboring neuronal sources will be highly correlated). sLORETA is currently a method of choice since it has no localization bias even in the presence of

measurement and biological noise. In the current implementation of sLORETA, computations are made in a realistic head model using the MNI152 template with the three-dimensional solution space restricted to cortical gray matter, as determined by the probabilistic Talairach atlas. Thus, sLORETA images represent the standardized electric activity at each voxel in neuroanatomic Montreal Neurological Institute (MNI) space as the exact magnitude of the estimated current density (Fig. 1.15–23). Several studies have reported consistency between sLORETA and other neuroimaging methods such as fMRI. The LORETA/sLORETA software is based on the pioneering work of Roberto Pascual-Marqui and his group, and relevant references can be freely downloaded (http://www.uzh.ch/keyinst/loreta.htm).

Status of Q-EEG in Applied Clinical Practice

Despite advances in EEG signal extraction enabled by digital technology, there are few specific conditions for which Q-EEG has gained acceptance as a clinical diagnostic tool, and in these cases it is used only to enhance interpretation of conventional clinical EEG. According to the AETNA Health Insurance list of approved indications (last reviewed April 28, 2015), Q-EEG is medically necessary only as an adjunct to traditional EEG for the following conditions (from http://www.aetna.com/cpb/medical/data/200_299/0221.html, accessed September 15, 2015):

1. For ambulatory recording of EEG to facilitate subsequent expert visual EEG interpretation; *or*
2. For continuous EEG monitoring by frequency trending to detect early, acute intracranial complications in the operating room or ICU; *or*
3. For evaluation of certain members with symptoms of cerebrovascular disease whose neuroimaging and routine EEG studies are not conclusive; *or*
4. For evaluation of dementia and encephalopathy when the diagnosis remains unresolved after initial clinical evaluation; *or*
5. For screening for possible epileptic seizures in high-risk ICU members; *or*
6. For screening for possible epileptic spikes or seizures in long-term EEG monitoring; *or*
7. For topographic voltage and dipole analysis in presurgical evaluations for intractable epilepsy.

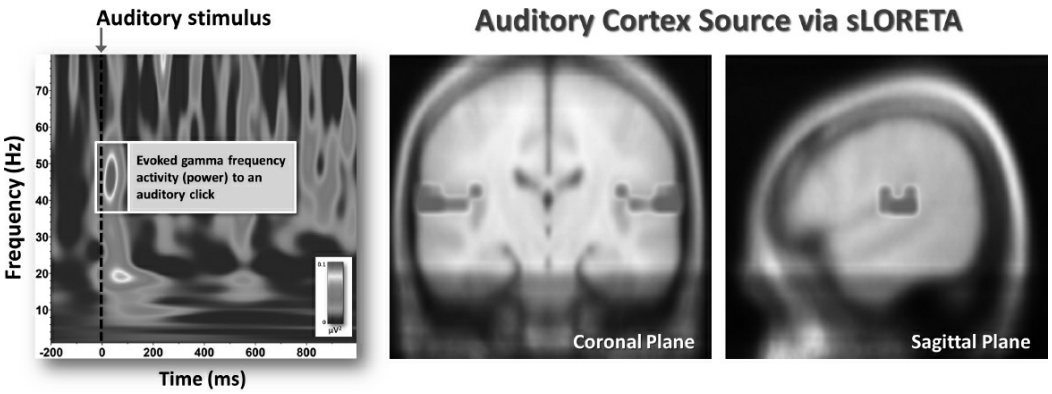

FIGURE 1.15–23. Example of EEG source analysis using sLORETA. The time–frequency transform on the left indicates a burst of transient gamma activity to an auditory click stimulus, which was localized to primary auditory cortex (right). (See color atlas.)

The list of indications described is no more inclusive than those recommended nearly 20 years prior by the ACNS and American Academy of Neurology (AAN) Executive Board's publication, appearing in the journal *Neurology* in 1997. Also consistent with conclusions of the ACNS/AAN report, AETNA considers the use of Q-EEG "experimental and investigational" for all other indications due to lack of adequate scientific evidence to support its clinical utility. Psychiatric conditions listed explicitly among unsupported indications include alcohol and SUDs, Asperger syndrome and other ASDs, attention deficit disorders, major depression and bipolar disorder, panic disorder, and schizophrenia. Furthermore, despite favorable reports in the research literature, Q-EEG has yet to receive approval for use in predicting efficacy of a psychotropic medication.

EEG IN APPLIED CLINICAL PRACTICE

Despite the fact that EEG originated in psychiatry, the strongest initial impetus for its use came from neurology, particularly the study of epilepsy. The years from 1934 to 1940 saw a marked proliferation of EEG studies focused on structural brain lesions and a variety of seizure disorders. In 1934, the team led by Fred Gibbs discovered the classic three-per-second spike-and-wave complex, which proved to be specific to petit mal absence attacks. Before the decade ended, they had described EEG patterns associated with grand mal and myoclonic seizures and a diffuse spike-and-wave pattern that was slower in frequency than petit mal (and given the confusing name of petit mal variant) and that was associated with grand mal seizures and a high incidence of mental retardation. They also introduced the term *psychomotor seizures* (now *complex partial seizures* [CPSs]) and described the EEG manifestations characterizing an ictal psychomotor attack. Later, in 1947 and 1948, they described the anterior temporal spike focus that became the interictal EEG correlate of this disorder. On the other side of the neurological coin, the study of structural

brain lesions was advanced through landmark EEG discoveries of the 1930s to 1950s. In 1935, Otfried Foerster and Helmut Altenberger reported from that focal slow waves in the EEG recording often appeared near brain tumors and, 1 year later, Grey Walter made a major advance by demonstrating a technique for EEG localization of brain tumors. Denis Williams at Oxford subsequently began using EEG recordings to study and localize traumatic intracranial injuries received during World War II. Finally, through the collaboration of neurosurgeon Wilder Penfield and psychologist (turned neurophysiology research) Herbert Jasper, who brought EEG to the Montreal Neurologic Institute in 1937, a technique for direct intracranial EEG recording termed *electrocorticography* (ECoG) was developed. ECoG was applied to the detection of eptileptogenic zones for surgical resection as part of the groundbreaking *Montreal procedure* in the 1950s.

Normal versus Abnormal: General Considerations

One of the factors that may have limited the perception of EEG as a useful assessment tool in psychiatry is the simplistic attempt to conceptualize EEG findings as a normal-versus-abnormal dichotomy. Psychiatric practitioners commonly ask if their patient's EEG is abnormal and seem content with an affirmative or negative answer. The expectation was that if the EEG was abnormal then this meant that the patient belonged to "neurology" and if normal that would confirm the psychiatric status of the patient. Furthermore, much of the published research literature reports the percentage of abnormal EEG findings in a given study population without stating the exact nature of the abnormalities. In actuality, the simple normal–abnormal dichotomy is inconsistent with the dynamic nature of EEG, particularly when interpreted by conventional visual inspection without use of more objective means of quantification that could equate EEG to more common, and reliably discernable, measures of physiological function used in medicine (e.g., blood pressure). Rather than being considered as a dichotomy, the range of EEG findings exists on a broad continuum anchored on one end by unequivocally normal EEGs and extending in the other direction to features of possible and, eventually, more definite clinical significance. Some EEG patterns that are correctly classified as abnormal are on the low end of the continuum of clinical expressivity and may not always contribute strongly to diagnostic decisions. The psychiatrist receiving a report of a minor finding of limited clinical relevance may understandably begin to question the value of EEG when no current or past history of organic signs are detected. The bottom line is that one should not ask simply if an EEG is abnormal, as if it were one side of a dichotomy, but instead should ask what specific features were present and what, if any, known clinical correlates may be associated with that feature. Even when questions remain about the clinical significance of

an abnormal feature, this finding should not simply be disregarded and, at least, could indicate a need for follow-up EEG, additional brain imaging, neuropsychological assessment, review of medications or illicit drugs used, or other laboratory measures of cardiac and metabolic function.

In addition to the confusion caused by the unwise reliance on the normal–abnormal dichotomy, literature reports stating that 10 to 20 percent of the normal population have abnormal EEG findings could lend uncertainty regarding the value of the EEG in psychiatric practices. Because such statements are almost never properly clarified, they tend to be terribly misleading. Clearly, there can be an understandable reluctance to follow up an abnormal EEG report if one has read in a presumably authoritative source that 20 percent of the normal population have abnormal EEG findings, decreasing the impression of their diagnostic specificity and, logically, clinical utility. More importantly, there also may be an equally understandable reluctance to refer a patient for EEG study for the same reason. To place the issue in proper perspective, the reports of substantial percentages of abnormal EEGs within the normal population are heavily skewed by the inclusion of minor findings that often have low levels of clinical or diagnostic relevance. A comprehensive review of the literature on EEG findings in normal control populations by Nash Boutros and colleagues highlights the fact that most control studies have been seriously compromised by failure to screen subjects for a variety of medical and psychiatric variables that can impact EEG results. This group had previously reanalyzed data from numerous published EEG control population studies. When borderline EEG findings were removed from the data, the incidence of EEG abnormality dropped to 3.2 percent of 6,182 healthy adults and 3.5 percent of 1,450 healthy children, with the prevalence of accepted paroxysmal EEG abnormalities reduced to only 1.14 percent of a sample of 11,560 mixed-age healthy subjects.

Variants of Normal EEG

Sleep Patterns. The EEG patterns that characterize drowsy and sleep states are different from the patterns seen during wake. A detailed accounting of the nuances of sleep EEG is beyond the scope of this chapter. Briefly, the rhythmic posterior alpha activity of the closed-eye waking state subsides during progressive drowsiness and is replaced by irregular low-voltage theta activity and sporadic vertex sharp waves, which indicate entry into stage 1 sleep. The progression into stage 2 sleep is marked by the appearance of 12- to 14-Hz sleep spindles (also called sigma waves) and k complexes. These patterns eventually give way to high-voltage (>75 μV) delta waves, which indicates stage 3 deep sleep. Lastly, REM sleep is characterized by low-voltage, mixed frequencies that resembles waking EEG. Some individuals may also exhibit low-amplitude sawtooth waves in REM sleep.

Changes with Age. The appearance of the EEG changes dramatically from birth to advanced age. From a preponderance of irregular medium- to high-voltage delta activity in the tracing of the infant, EEG activity gradually increases in frequency and becomes more rhythmic with increasing age. Rhythmic activity in the upper theta–lower alpha range (7 to 8 Hz) can be seen in posterior areas by early childhood, and, by the time mid-adolescence is reached, the EEG essentially has the appearance of an adult tracing. The interpretation of EEGs obtained from children demands a solid grounding in the age-related changes during this period and is best performed by one specializing in pediatric EEG. For sake of the current chapter, it is sufficient to raise awareness that developmental period (among other factors to be discussed) must be considered in making absolute determinations regarding whether or not the presence of a given feature (e.g., presence of waking delta wave activity) is clinically significant. This topic of age-related change in EEG and trajectories of pathology receive detailed treatment in the text *Developmental Psychophysiology,* edited by Louis Schmidt and Sidney Segalowitz.

Controversial Sharp Waves or Spike Patterns

Early in the emergence of clinical EEG, a considerable amount of heated controversy developed around the clinical significance of certain sharp wave or spike patterns that were easily identified in the EEG tracing. These so-called controversial patterns deserve a specific mention here, because they have been observed consistently in numerous studies to be more prevalent in psychiatric populations than healthy or general medical patient populations. Despite the increased prevalence in psychiatric patients, defining the clinical and neurobiological correlates of these patterns has proved to be an elusive goal. The task is made difficult because none of these patterns is specific for a diagnostic category, and none predictably correlates with seizure disorders. Because the physiological or pathological correlates of these patterns have not been well studied using currently available methodologies and assessment techniques, the assertion that they are irrelevant to neuropsychiatric conditions is not fully supportable. At a minimum, the appearance of such patterns may give support to the possible organic nature of presenting symptoms.

The number of accepted and still controversial EEG abnormalities is quite large and well beyond the scope of this chapter. For the purpose of illustration only, four classic abnormal EEG patterns are described in this section (Figs. 1.15–24 to 1.15–27).

The reader should note that each of these four patterns has two different labels. One label usually is used by those who ascribe clinical significance to the pattern (i.e., *small sharp spikes* [SSSs]), and the other is used by those who consider these patterns to be normal variants (i.e.,

benign epileptiform transients of sleep [BETS]). Labels given in parentheses in the following section are the ones used by those who reject the potential clinical significance of the patterns.

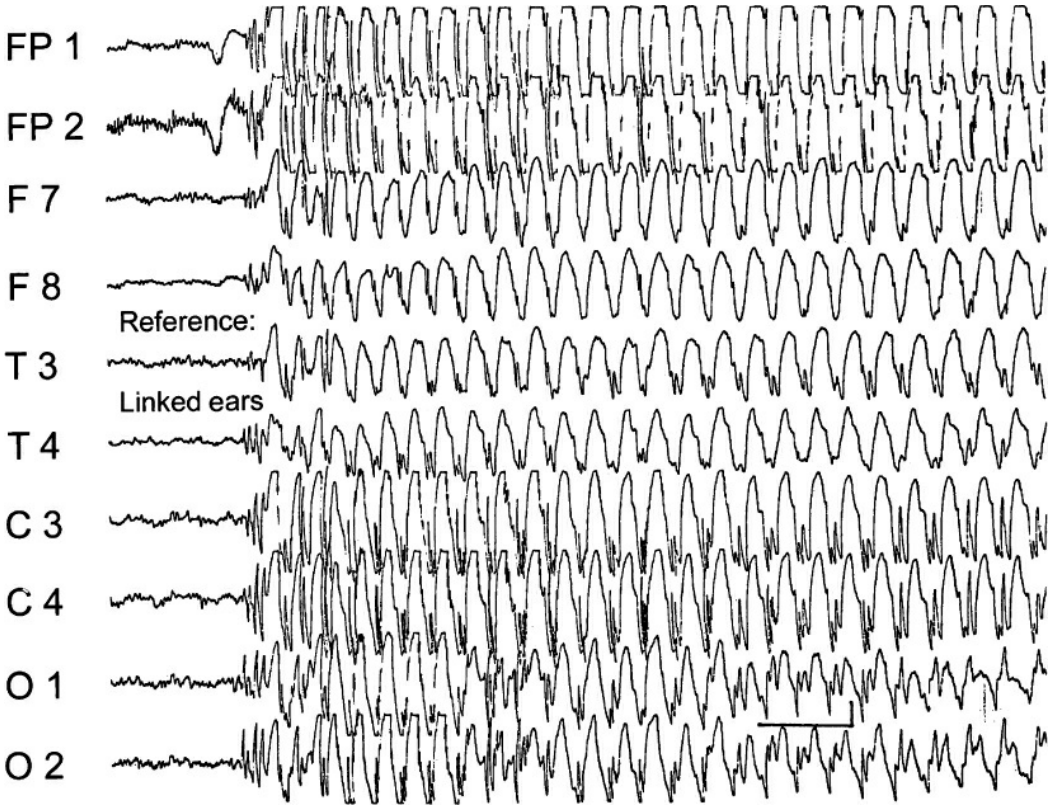

FIGURE 1.15–24. Diffuse three-per-second, spike-and-wave discharges of the petit mal type with a multiple spike component. The patient is a female 16 years of age with a previous history of petit mal seizures and rare grand mal attacks. At a previous psychiatric facility, the diagnosis was changed to "hysterical seizures" following a psychological evaluation.

Fourteen- and Six-per-Second Positive Spikes (Ctenoids). Since its description by Fred and Erna Gibbs 50 years ago, the fourteen- and six-per-second positive spike pattern (hereafter called *PS*) has become the most frequently encountered of the controversial waveforms and the one with the most extensive literature (Fig. 1.15–28).

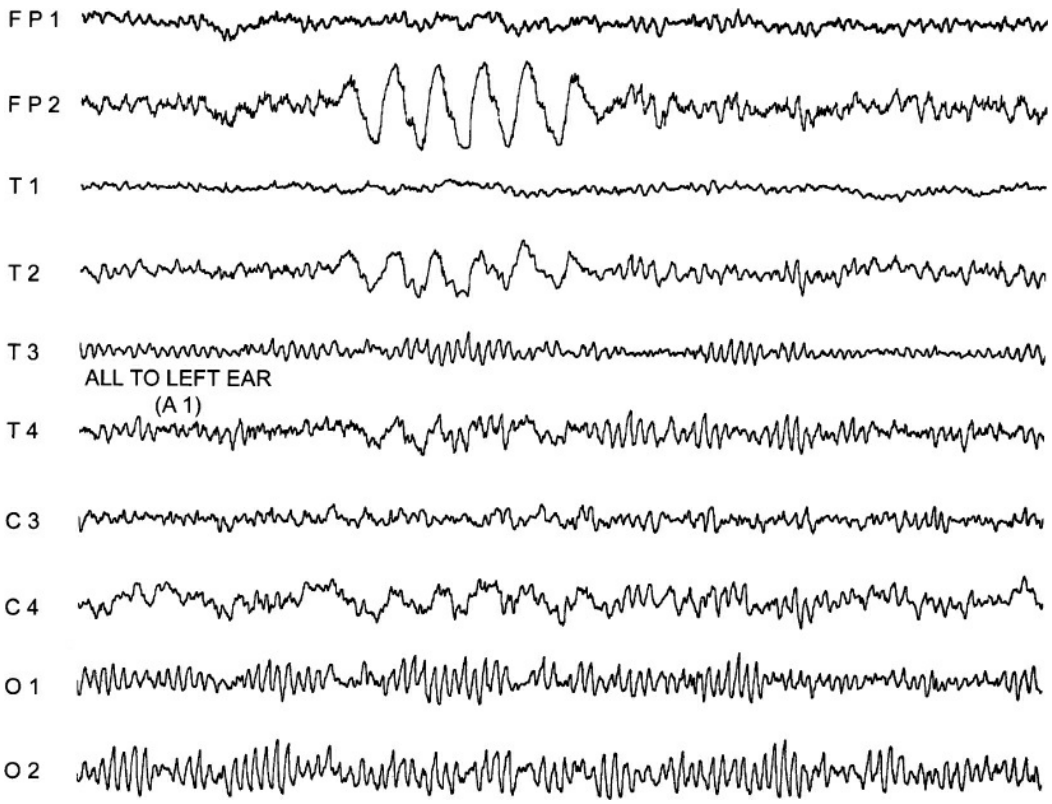

FIGURE 1.15–25. Focal slow (delta), right prefrontal spreading with reduced voltage to right anterior and mid-temporal and right central regions. During the sleep recording (not shown), sleep spindles were absent in the right hemisphere. The patient is a female 49 years of age with no prior psychiatric history, recent emergence (over 6 months) of depression, mild episodic confusion, symptoms of hyperthyroidism, and pathological crying. Glioblastoma was found at surgery.

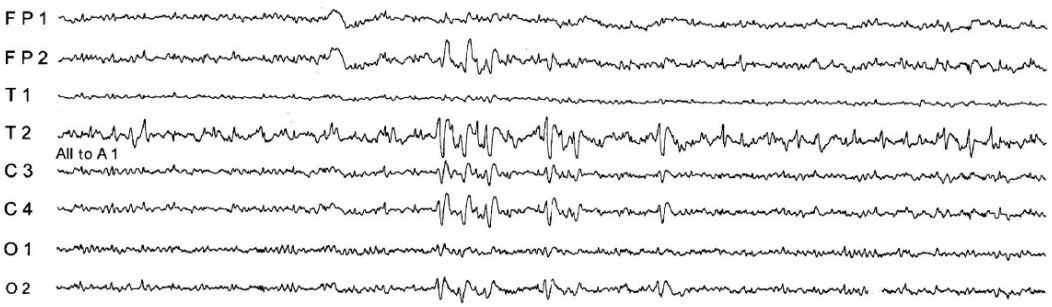

FIGURE 1.15–26. Focal negative spike and spike–wave discharges in the right anterior temporal cortex and often spreading with reduced voltage throughout the right hemisphere. The patient is an adult male with complex partial seizures (psychomotor) and a psychiatric outpatient with infrequent periods of sudden unresponsiveness during which time he would make grunting and guttural sounds and speak with incoherent words and syllables with amnesia for these events. (His girlfriend, who was involved in New Age phenomena, was convinced that he was communicating with spirits during these unresponsive spells and did not want him to be medicated.)

They are typically observed during sleep, and are described as bursts of arch- or comb-shaped waves at 14 and/or 6 Hz over posterior temporal areas when a contralateral ear reference is utilized. Indeed, the term "ctenoid" was coined by Lombroso and others to describe their comb-like

shape. They are most often seen in states of drowsiness or in light non-REM sleep. The pattern is highly age-dependent, which starts after around the age of 4, and then declines as one enters adulthood. While originally thought to represent an indication of an underlying pathology (e.g., epilepsy, brain lesions), PSs have been shown to be quite common in healthy individuals, and are now only regarded as abnormal if they are pronounced or occur at an extremely high rate.

Rhythmic Mid-Temporal Discharges (Psychomotor Variant). Rhythmic mid-temporal discharges (RMTDs) are long runs of activity in the 5 to 6.5 Hz range with a typical maximum over mid-temporal regions (Fig. 1.15–29). They most commonly during drowsiness or light sleep. This pattern is rare, and is observed in less than 0.1 to 0.12 percent of the population. Earlier studies reported an increased incidence of seizures in patients with RMTD, but the association of this pattern with ictal phenomena remains controversial. Although there have been several efforts to link this EEG dysrhythmia with behavioral symptoms, such as temper dysregulation, personality disorder, or autonomic phenomena, other studies have not been able to demonstrate such associations. There have been rare documented cases of RMTD status epilepticus associated with prolonged periods of confusion.

Small Sharp Spikes (Benign Epileptiform Transients of Sleep). The SSS pattern consists of rather low-voltage sharp negative or biphasic spikes that are best seen over the frontal–temporal cortex, sometimes alternating between the left and the right hemispheres (Fig. 1.15–30). It is very inconspicuous, and often overlooked. It is rarely seen in children, and tends to peak in those in the 30 to 60 age range (Chang et al., 2011). It is most often seen in drowsiness or in light non-REM sleep. They also have been associated with the presence of neurovegetative symptoms (i.e., headache, dizziness, and episodic abdominal pain). Furthermore, the likelihood of a patient experiencing epileptic seizures is higher in association with this pattern as compared to those of the other three controversial patterns. For example, the incidence of seizures in patients with SSS has ranged as high as 53 percent and even 67.4 percent, thus stressing the moderate degree of epileptogenicity of this pattern.

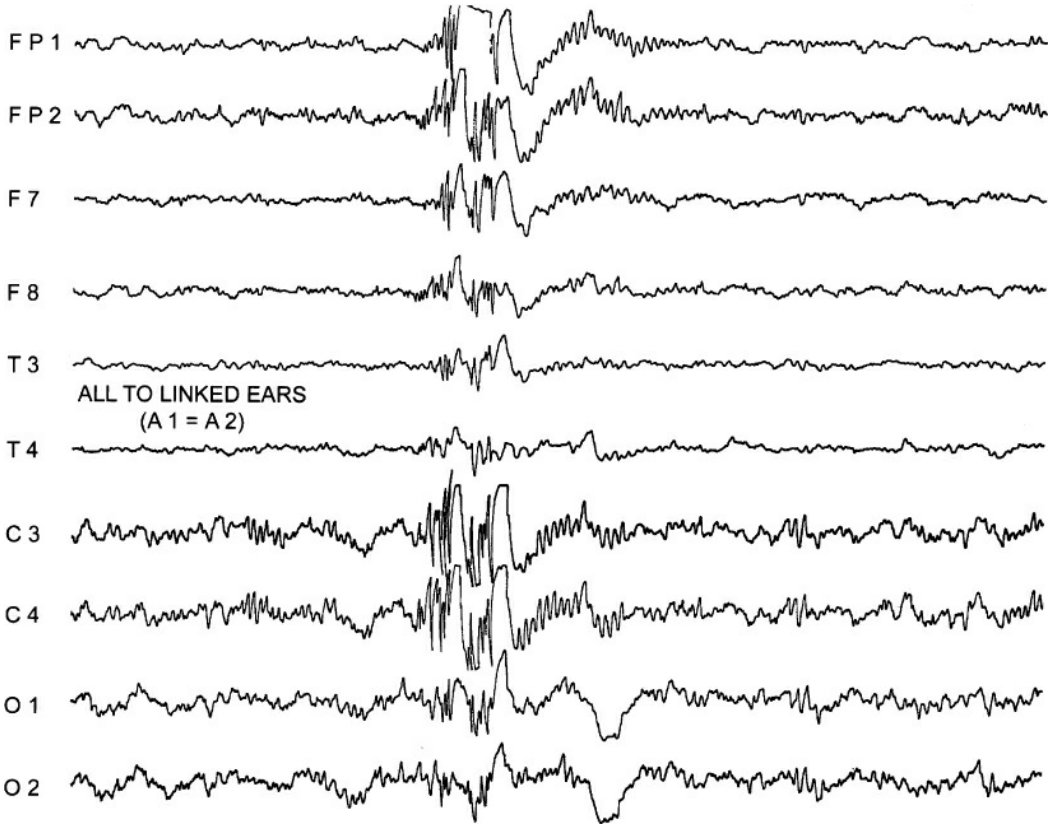

FIGURE 1.15–27. Diffuse spike-and-wave discharges during the sleep recording. The waking electroencephalography was within normal limits. The patient is a female 15 years of age who was admitted to an inpatient adolescent unit with complaints of serious temper dyscontrol and episodic dizzy spells.

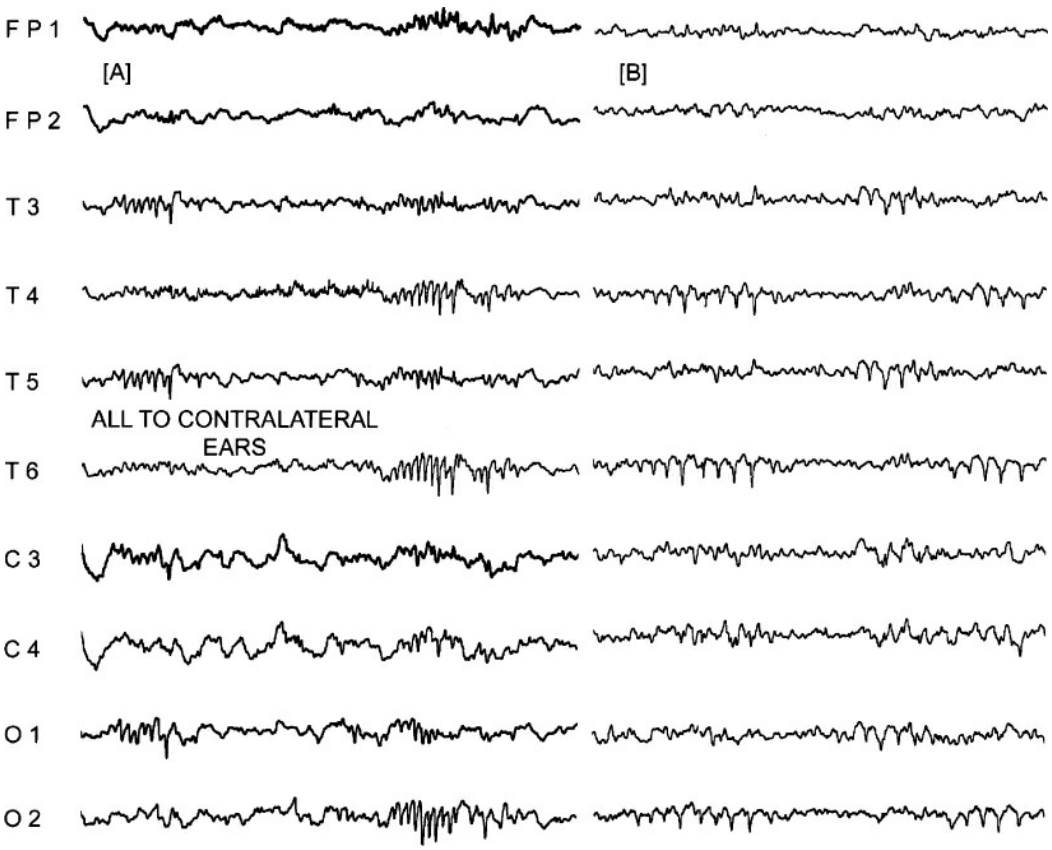

FIGURE 1.15–28. Fourteen- and six-per-second positive spikes, independent left and right mid-temporal posterior–temporal–occipital areas. **A:** Fourteen-per-second variety; female patient, 13 years of age. **B:** Six-per-second variety; male patient, 24 years of age. An electroencephalograph may contain either variety alone or both combined.

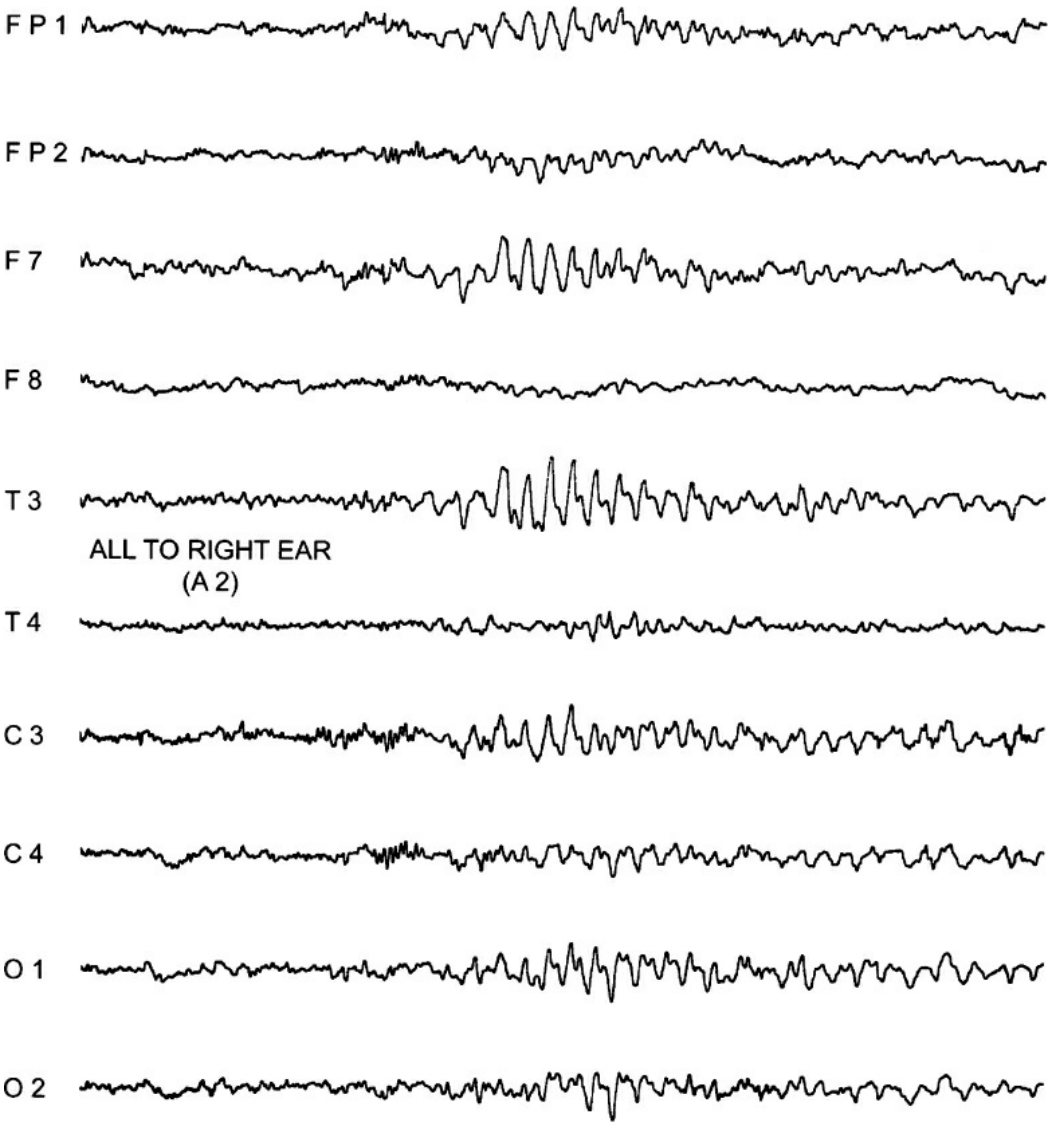

FIGURE 1.15–29. Rhythmic mid-temporal discharges (psychomotor variant) during drowsy recording. Left mid-temporal focus. The patient is a male 20 years of age with episodic temper dyscontrol, frequent intense homicidal urges, paresthesias, headaches, abdominal pain, and suicide plans (as a way of avoiding acting on strong homicidal impulses). Patient responded well to phenytoin (Dilantin) therapy.

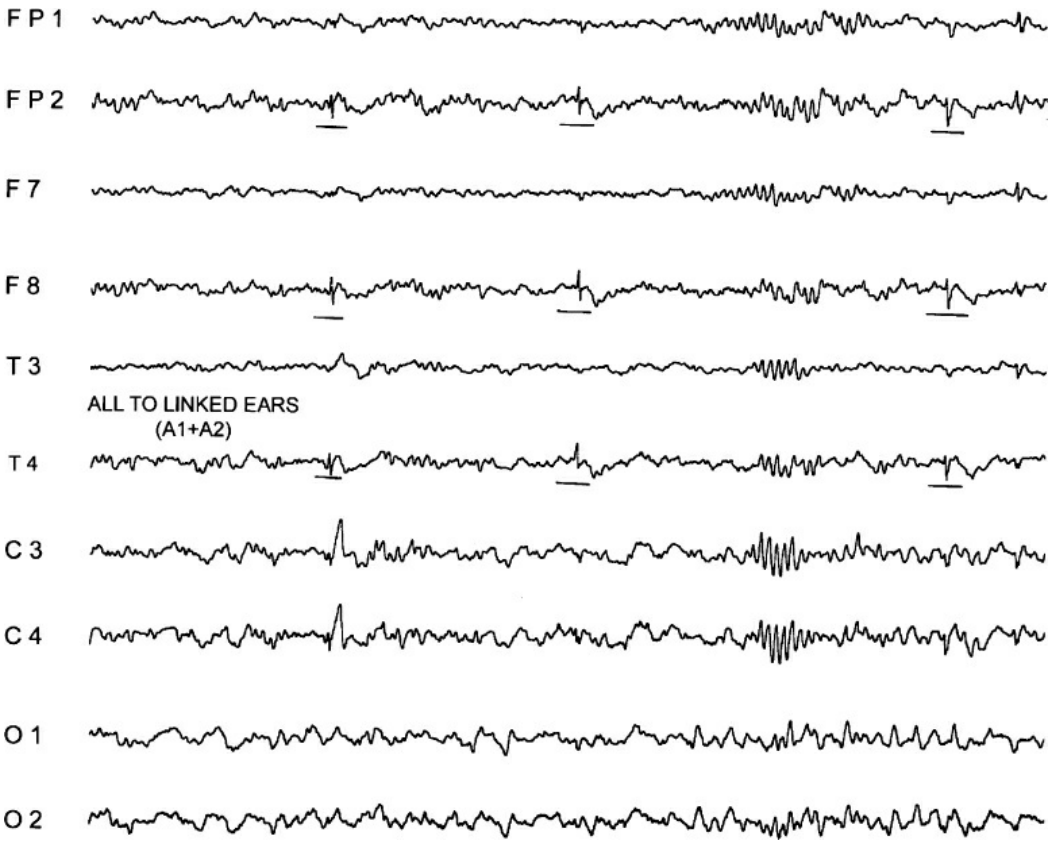

FIGURE 1.15–30. Small sharp spikes (underlined), right frontotemporal during stage II sleep. The patient is a female 24 years of age with a history negative for head trauma, encephalitis, or seizures. The psychiatric diagnosis was major depression. Patient also complained of headaches and infrequent parethesias.

Six-per-Second Spike and Wave (Phantom Waves). Finally, the six-per-second spike-and-wave pattern (also called *spike–wave phantom*) is often difficult to recognize, because the discharges tend to be low in amplitude (Fig. 1.15–31). It is usually observed in adulthood, but can also be seen in children or adolescents. It occurs while awake, during drowsiness, or during light non-REM sleep. This pattern is now considered a normal variant, with uncertain clinical relevance.

Paradigms for Eliciting Pathophysiological Features

Over the years, electroencephalographers have recognized that certain *activating procedures* tend to increase the probability that abnormal discharges, particularly spike or spike–wave seizure discharges, will occur. Some activating techniques remain standard in many laboratories, others are used only rarely for specific purposes, and still others introduced in the past largely have been abandoned because they were not easy to use or involved risk. It has long been recognized that seizure manifestations or aberrant behaviors with a pathophysiological basis involving seizures can be triggered by specific stimuli. In this regard, cases of audiogenic, musicogenic, photogenic, and reading epilepsy readily come to mind, even

though such cases are rare, and most practitioners have never encountered them. Seizure phenomena related to other sensory system input (e.g., somatosensory and gustatory) are even more rare. Sometimes, it may be possible for laboratory personnel to duplicate or approximate sensory triggers in various modalities to determine if they activate EEG seizure discharges combined with overt symptoms. Some of the more conventional activating techniques are described below.

Photic Stimulation. In the earliest days of EEG, it was known that the frequency of normal EEG activity recorded from posterior scalp regions could be made (within narrow limits) to follow the frequency of a flickering light that was flashed slightly faster or slower than the intrinsic brain wave frequencies, a phenomenon that came to be referred to as *photic-driving*. An illustration of this phenomenon is presented in Figure 1.15–32, depicting the instantiation of a steady-state response of approximately 9 Hz over visual cortex following stimulation at a 9 Hz flicker rate. When it also became known that photic-driving would sometimes cause paroxysmal discharges to occur in the EEG, *intermittent photic stimulation* (IPS) emerged as a technique for eliciting EEG abnormalities referred to, under conditions of stimulation, as *paroxysmal photosensitive responses* (PRRs). Although there is some variation between laboratories, IPS generally involves placing an intense strobe light (at least 0.70 J) approximately 30-cm in from the nasion and flashing at frequencies that can range from 1 to 50 Hz, depending on how the procedure is carried out. It is recommended that a round stimulator be used to best ensure uniform stimulation of the entire retina. In addition to flashing light, flickering grid stimuli may be used to further assess pattern-specific photosensitive responses, which may be indicated in patients who tend to have seizures induced by television watching or video gaming. Some laboratories sample independent flash frequencies separately and randomly, without necessarily sampling all frequencies between 1 and 50 Hz, while others use a *zoom technique* in which the flashes start at a low frequency, such as 1 Hz, and are then gradually and continuously increased to higher frequencies. In some individuals, IPS produces facial and eye muscle jerks, called a *photomyoclonic response* (PMR). More than 40 years ago, Henri Gastaut observed that PMR occurs in 0.3 percent of normal subjects, 3 percent of epileptic patients, and 17 percent of patients with psychiatric disorders. The PMR also is enhanced in early stages of alcohol withdrawal in chronic alcoholics and after a sudden withdrawal from barbiturates and other sedatives. The clinical relevance of PMR in psychiatric patients is yet to be fully explored in systematic research.

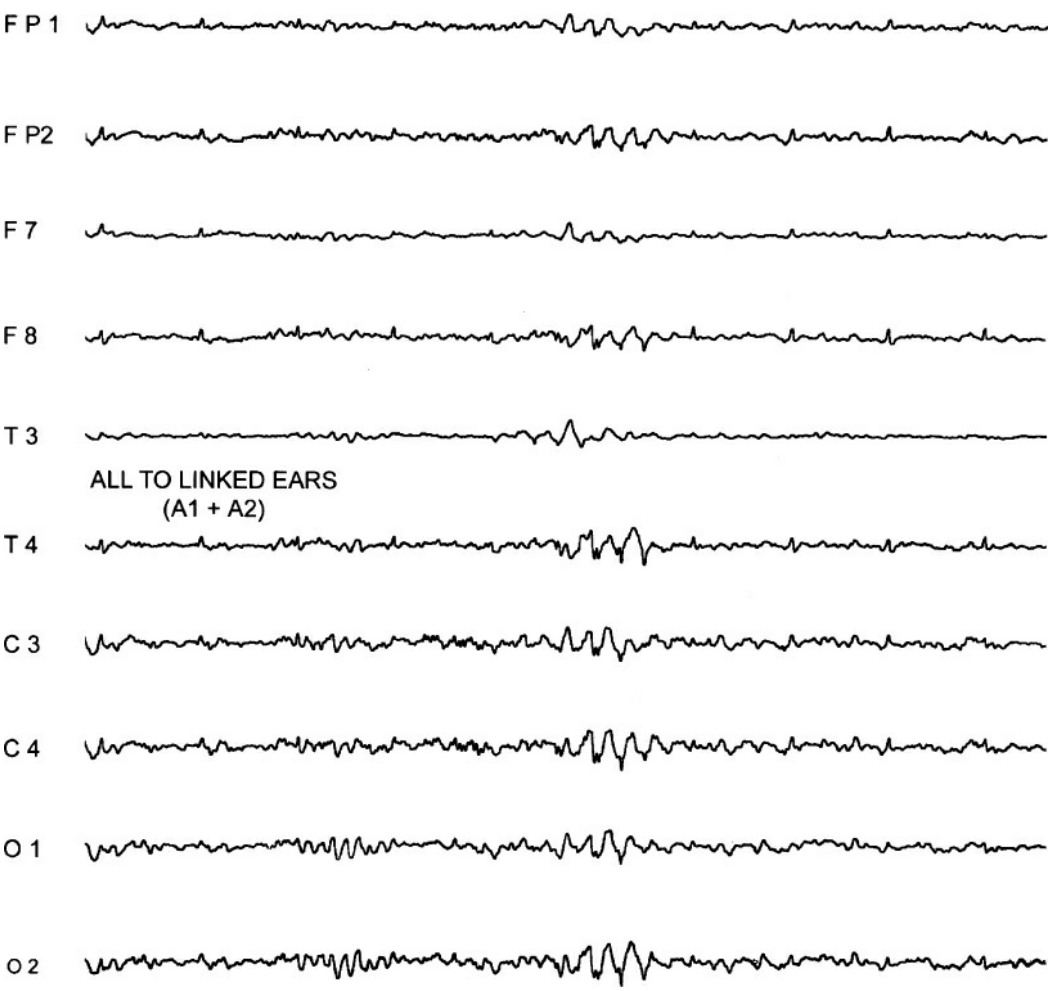

FIGURE 1.15–31. Six-per-second spike and wave during drowsy recording. Small electrocardiogram artifact present. The patient is a male 24 years of age diagnosed with schizophrenia.

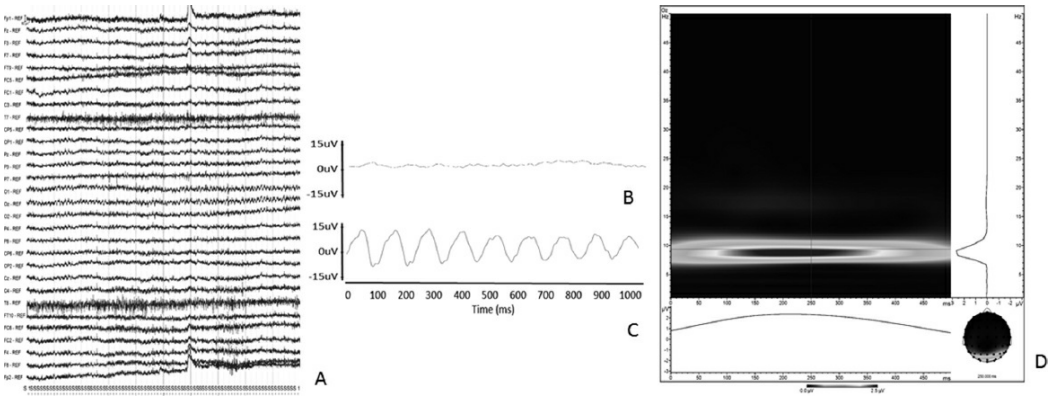

FIGURE 1.15–32. Photic driving using checkerboard pattern presented at 9-Hz flicker rate. **Panel A:** 10-second display of EEG at beginning of photic stimulation. Onset of visual checkerboard display indicated by "s" at bottom of display. **Panel B:** 1-second segment based on average EEG for 10-second period prior to stimulation recorded at electrode Oz. **Panel C:** Visual evoked potential (VEP) based on average of 10 seconds recorded during stimulation at electrode Oz. The VEP represents an oscillatory pattern of 9 cycles over 1 second segment. **Panel D:** VEP represented as time–frequency decomposition by wavelet transform of frequencies of 1 to 50 Hz at electrode Oz, with peak activity at approximately 9 Hz. Scalp plot (bottom right) indicates activity is generated primarily over occipital electrode sites.

In terms of activating EEG abnormalities, one looks for a PRR that consists of bilaterally synchronous, usually diffuse, spike-and-wave discharges of various frequencies or diffuse, multiple spike–wave complexes. When the resting EEG is normal and a seizure disorder or behavior that is suspected to be a manifestation of a paroxysmal EEG dysrhythmia is suspected, IPS can be a valuable activation to use. Its primary limitations, especially as a routine procedure, are the "false positive" incidence of PRR in individuals with no history of seizure disorder or current symptoms suggestive of such, and similarly low sensitivity in cases of verified seizure disorder. The persistence of the spike–wave discharges after the cessation of the IPS may also be an indicative of a PRR. Moreover, a pathological process might be indicated by the spread of the spike–wave activity to leads outside the occipital region.

Polysomnography—Sleep Study. Largely because of the pioneering studies and perseverance of Fred and Erna Gibbs, EEG recording during sleep, natural or sedated, now is widely accepted as an essential technique for eliciting a variety of paroxysmal discharges, when the wake tracing is normal, or for increasing the number of abnormal discharges to permit a more definitive interpretation to be made. A variety of focal and diffuse spike and spike–wave discharges, as well as several minor or controversial paroxysmal patterns, occur much more often during drowsiness and light sleep than during the wake recording, and some of them are seen almost exclusively during the sleep recording. Paroxysmal patterns differ substantially among themselves in the degree to which their appearance in the tracing is sleep-activation dependent (Fig. 1.15–33). Although most clinical EEGs should contain drowsy- and light-sleep tracings to be complete, stage 3 deep sleep with generalized high-voltage delta slowing have almost no activating property and is not clinically useful.

Sleep Deprivation. It has been shown that the stress produced by 24 hours of sleep deprivation alone can lead to the activation of paroxysmal EEG discharges in some cases. This effect is presumably independent of the known activating properties of natural or sedated sleep itself. Sleep deprivation is without risk for the healthy patient but may be contraindicated for patients medically or physically compromised. The primary disadvantage is the tendency for sleep-deprived individuals to enter into deep sleep (stage 3) immediately at the start of the recording, thus reducing the chances of detecting spike activity. The optimal method is to ensure that the subject stays awake until the recording begins and remains so for the initial recording period, after which a gradual transition into drowsy sleep can be observed.

Drug Activation and Suppression Tests. Psychiatrists evaluating a

patient with an atypical behavioral reaction to a drug (e.g., an unusually severe response to a small amount of alcohol or other abused drug or some other highly idiosyncratic response) may consider a drug-activated EEG to determine if ingestion of the drug activates a seizure or other paroxysmal abnormality of clinical significance. In the past, drug activation using hypoglycemic agents (e.g., tolbutamide) and convulsants (e.g., pentylene tetrazol) was used to activate epileptiform discharges. However, this type of activation has been essentially eliminated from clinical work in recent years due to the high incidence of discharges produced by these drugs in healthy subjects and low value of positive findings in individuals with epilepsy. Today, drug modulation of EEG remains to be used as an investigative procedure for ascertaining drug effects on presumed neurophysiological mechanisms, both with relevance to enhancing functions in the case of drug development, and in modeling pathological conditions in healthy subjects using the pharmacological challenge design. While drug activation is no longer used in routine clinical practice, drug suppression paradigms are still regarded as useful for the identification of focal epileptiform activity. Seizure foci can be difficult to detect on EEG when in the case of generalized discharges and accompanying spread of activation across electrodes. In a drug suppression test, generalized discharges can be reduced by intravenous injection of diazepam or barbiturate methohexital (e.g., methohexital suppression test; MHXT). The rationale behind this approach is that by using drugs that block excitatory neurotransmission and/or enhance inhibitory transmission, much of the background EEG is suppressed while autonomous epileptogenic lesions, which are resistant to synaptic blockade, continue to fire. This approach can therefore improve localization of epileptiform lesions against background EEG and generalized activity that is dependent on synaptic input.

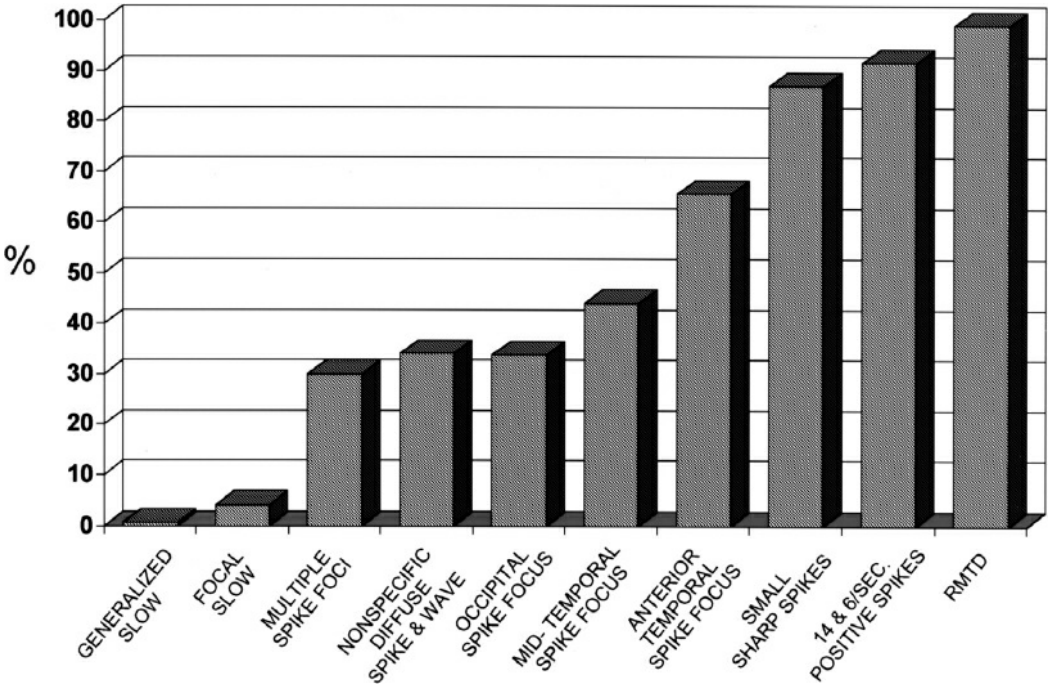

FIGURE 1.15–33. Percent of various electroencephalography patterns detected only during drowsiness or sleep, or both. To be read as follows: Of all cases of multiple spike foci, 30 percent required a drowsy or sleep, or both, recording for their detection, and 70 percent were detected during a wake recording. RMTD, rhythmic mid-temporal discharge.

Hyperventilation. Hyperventilation is one of the oldest, and still one of the most frequently used, activation procedures in clinical laboratories. While remaining reclined with the eyes closed, the patient is asked to overbreathe through their open mouth for 1 to 5 minutes (not useful beyond 5 minutes), emphasizing deep breathing rather than rapid breathing. The purpose is to empty the lungs to the extent of producing hypocarbia, and to reduce oxygen and glucose supply to the brain through cerebral vasoconstriction. The normal EEG response to hyperventilation (referred to in EEG parlance as a *buildup*) consists of an increase in generalized medium- to high-voltage synchronous slow waves in the delta range, which then quickly subside when overbreathing stops (within 1 to 2 minutes normally). Not everyone has a build-up response to hyperventilation, and children are far more likely to respond with diffuse EEG slowing than are adults. Indeed, this procedure is not effective in the elderly, who show minimal reactivity of the cerebral vasculature to hypocarbia, and is contraindicated those aged 65 and older. In terms of activating EEG abnormality, hyperventilation is especially effective in eliciting the classic diffuse three-per-second spike-and-wave complex of petit mal seizures when the pattern does not first appear in the standard waking EEG. Hyperventilation may, to a lesser degree, activate other synchronous diffuse spike–wave patterns and activation of focal seizure activity has been infrequently reported. An interesting observation is that significantly low blood glucose levels have been associated with large,

synchronous delta wave hyperventilation buildups that may persist beyond 2 minutes after overbreathing is stopped. Accordingly, a large delta wave buildup in an adult may signal the existence of covert, unsuspected pathological hypoglycemia. If this suspicion should present itself during a recording, a good idea would be to give a sugar drink to the patient and then to repeat hyperventilation later. If the glucose ingestion reduces or abolishes the large hyperventilation buildup, then the suspicion of hypoglycemia is likely confirmed. In general, hyperventilation is one of the safest EEG activating procedures, and, for the majority of the population, will present no physical risk. However, it may pose a risk for patients with cardiopulmonary disease or risk factors for cerebral vascular pathophysiology, thus would also be contraindicated in these cases.

Brief Synopsis of EEG Findings in Organic Pathophysiology

The EEG has a high degree of sensitivity to a wide variety of medical conditions and events that impinge on CNS function. For this reason, EEGs sometimes have contributed to the detection of unsuspected organic pathophysiologies influencing a psychiatric presentation. The range of EEG findings in medical disease is enormous and far beyond the scope of this chapter, but this topic is covered in depth in the large edited volume by Ernst Niedermeyer and Lopes da Silva. Findings associated with some of the major organic conditions are discussed below.

Seizure Disorder. The hallmark EEG finding for a seizure disorder is the generalized, hemispheric, or focal spike or spike–wave discharge, or both (Figs. 1.15–24, 1.15–26, and 1.15–27). However, this statement constitutes a large oversimplification, because many types of EEG abnormalities have been associated with seizures at one time or another, and some patients with a bona fide seizure disorder have been known to have normal interictal EEG tracings. Indeed, while epileptiform activity detected on EEG is a reasonably specific indicator of epilepsy, sensitivity can be quite low, ranging from 25 to 56 percent, with only about 50 percent of patients diagnosed with epilepsy evidencing *interictal epileptiform discharge* (IED) of their first EEG examination (of note, this rate can increase to 80 percent in sleep EEG). Furthermore, seizure disorders have many possible etiologies, whether genetically based or a consequence of head trauma, cerebrovascular pathophysiology, metabolic disorders, structural brain lesions, infectious or toxic encephalopathies, certain drug-abuse withdrawal states, or iatrogenic causes to name only a few. The diagnosis of seizure disorder based on EEG findings should consider family and personal medical history, while also keeping in mind that seizures can be idiopathic.

The classification of seizure types is also broad, and the majority of specific seizure manifestations may be of little concern to the practicing

psychiatrist. One of the exceptions may be *absence* (i.e., petit mal) seizures, during which the patient presents with grossly impaired consciousness marked by pronounced confusion, greatly slowed mental processes, and stuporous or somnolent behavior. Although these seizures tend to be brief, normally lasting only 15 to 30 seconds, absence episodes can reoccur as many as 100 times per day. The psychiatrist is most likely to encounter this clinical presentation in the emergency room, where it is often confused with other functional or organic syndromes or intoxicated states. If suspected, then the status can be confirmed quickly by an EEG demonstration of continuous, diffuse spike–wave activity. Typical absence seizures also may be of relevance to the child psychiatrist as this type of epilepsy has a peak incidence from childhood to early adolescence and, in some cases, may be mistaken for inattention or other functional behavioral reactions. Spike foci in the anterior temporal region, associated with CPSs (also known as psychomotor or temporal lobe seizures) should also be of interest to the psychiatrist due to the wide range of possible seizure manifestations, consisting of almost any combination of "automatic" motor movements, sensory disturbances, psychiatric symptoms, or autonomic signs. The symptom cluster almost always remains the same from one attack to another, and the "automatic" behavior during the ictal event may be quite unusual, with the patient undressing, picking at clothing and lip smacking, walking in circles and yelling, trying to climb on a table while making guttural noises, experiencing hallucinatory symptoms, or almost any other action. When continuous or closely spaced CPS attacks (complex partial status) lasting hours or longer than a day occur, they may mimic hysterical dissociative states.

Spike activity of various kinds and locations is sometimes found in the EEGs of psychiatric patients with no obvious seizure manifestations. In such cases, one might consider possible relationships to episodic aberrant behavior (particularly explosive aggressiveness), if present, as well as a trial of an anticonvulsant medication. Sometimes, such spike discharges in the nonseizure psychiatric patient may indicate an elevated risk for iatrogenic seizures after treatment with neuroleptic medication (Fig. 1.15–34).

Structural Lesions. Brain tumors are the leading cause of new-onset seizures in adults. EEG is rarely used for the diagnosis of structural and space-occupying lesions, which are typically assessed with CT or MRI. Nonetheless, it should be noted that 90 percent of cortical brain tumors can be detected and localized on the scalp with routine EEG. For a thorough treatment of the effect of tumors or other space-occupying lesions on the EEG, please see Hartman and Lesser (2011).

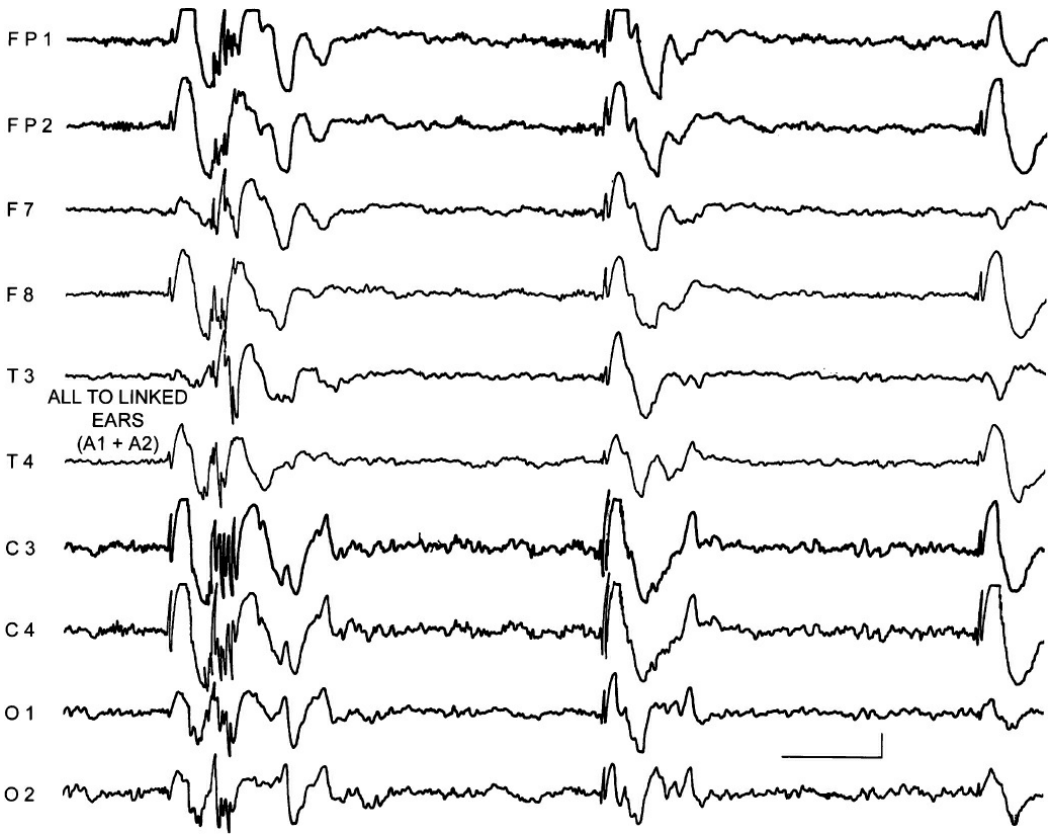

FIGURE 1.15–34. Routine admission electroencephalography (EEG). Diffuse spike-and-wave with multiple spike components during drowsiness recording (the wake EEG was normal). The patient is a male, 16 years of age with no present or past history of seizures and no history of head trauma, encephalitis, or other plausible causes for the spiking. Patient later developed iatrogenic grand mal seizures during treatment with neuroleptics.

Closed Head Injury. Generalized or focal slowing is the main characteristic of EEG following head trauma, appearing over the site of the trauma or over a *contrecoup* (e.g., the side opposite the side of impact) location. Attenuation of posterior alpha is another hallmark feature associated with head injury. The basic premise guiding the interpretation of these observations is that the frequency and power of scalp-recorded EEG is dependent on the numbers of activated neurons and the integrity of their connections. Therefore, in the case of injury, reductions in neurons and disruption of normal connectivity result in attenuation of signal power and a shift to slower frequency. The EEG change may be relatively transient with resolution over days or weeks, or it may persist for extended periods of time. Focal spiking may be taken as an early indication of a posttraumatic seizure disorder. Subdural hematomas after head injury may present with focal delta slowing or more widespread slowing, and, on occasion, amplitude asymmetry of waking EEG with the voltage decreased over the site of the hematoma. That said, conventional EEG is generally not regarded as clinically useful in cases of mild traumatic brain injury (mTBI), particularly beyond the first several hours following injury and when loss of consciousness (LOC) with injury is 2 minutes or less. EEG findings may

appear normal early in the postinjury period following mTBI despite the presence of postconcussive symptoms, such as headache, dizziness, fatigue, irritability, impaired memory and concentration, insomnia, and sound and light sensitivity. Clinically significant EEG changes in association with mTBI, if at all present, may not be recognized as deviant until compared with follow-up EEG taken later in the postinjury period. Of note, literature reviews suggest that as many as 20 percent of cases with noncomplicated mTBI recover from clinical symptoms within the first 24 hours, and as many as 85 percent within 1 week following injury. Similarly, abnormalities detectable by conventional clinical EEG tend to resolve within months of injury.

While routine clinical EEG may not be regarded as a useful tool for routine screening of mTBI, at least not beyond detection of posttraumatic seizure activity, there is evidence that Q-EEG may hold some promise. Numerous studies by Robert Thatcher and colleagues, as well as others, have reported Q-EEG findings associated with the postconcussion syndromes following mild to moderate closed head trauma. Reported Q-EEG changes involve a combination of increased theta power, decreased alpha power, increased theta–alpha ratios, and increased coherence between frontal and temporal areas in the resting EEG. A number of studies have applied discriminative analysis to Q-EEG collected in mTBI versus control sample, yielding high classification accuracy, and have demonstrated the utility of Q-EEG in predicting functional disability. Q-EEG and related topographic analysis are likely to outperform routine clinical EEG in these cases because the quantification process aims to identify sharply focal or regional alterations in activity according to statistically significant deviations from normative parameters, which could remain below the threshold for accurate visual detection (Fig. 1.15–35). The use of Q-EEG in assessment of mTBI nevertheless remains a heavily debated topic, with supporting evidence criticized for the composition of study samples and other methodological choices. David Arciniegas has published a detailed review on this issue including many suggestions for further reading.

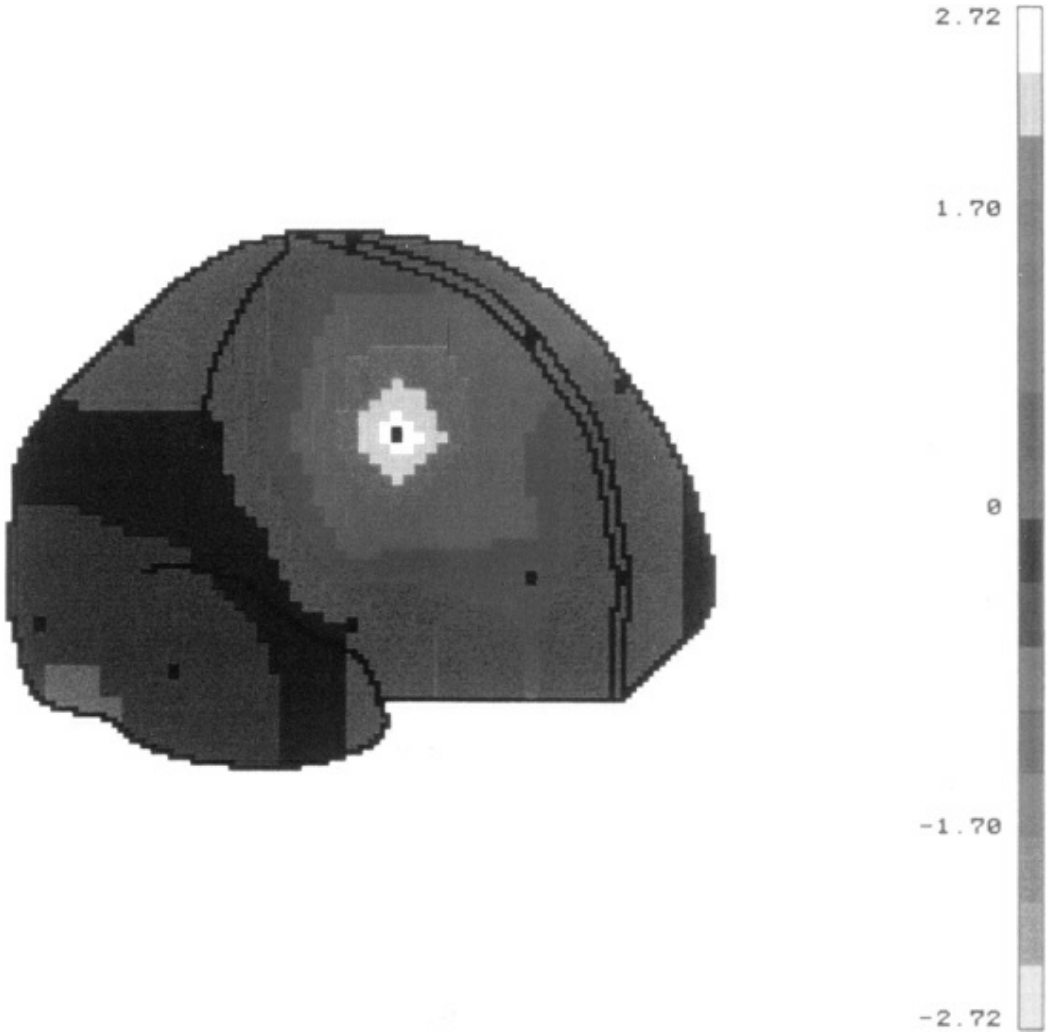

FIGURE 1.15–35. Topographic quantitative electroencephalography map of theta absolute power (*z* score departures from normative database mean). The patient is a male, 24 years of age with a closed head injury. The focus of increased theta voltage is at the locus of an earlier head injury that occurred about approximately 2 years before the recording. This was an unexpected finding. After recording and quantification, the color bar scale was adjusted to maximize the bull's-eye localization effect. Theta voltage was also elevated to a lesser extent over a wide, right frontal area and even spread somewhat across the midline. Theta relative power (not shown) was also elevated over the right frontal region, but mapping did not produce a sharp relative power focus at the locus of injury. Important note: A very sharply defined bull's eye can also be produced by artifact from a faulty electrode, and it is imperative to monitor electrical impedance and check the integrity of the electrode in the event that deviant activity appears confined to only one lead with no spread to adjoining electrodes.

Metabolic and Endocrine Disorders. EEG is not a tool that one would use in assessing endocrine or metabolic disorders. Yet the majority of these conditions alter the EEG, and unexpected abnormal findings during routine EEG screening have occasionally been the first positive laboratory results leading to the subsequent identification of mild endocrine or metabolic disturbance. Because metabolic and endocrine disorders, as well as certain vitamin deficiencies and toxic exposures, affect the brain in a global fashion, they tend to produce diffuse generalized slowing of wake

frequencies, with the degree of abnormality correlated with the severity of the disorder. As would be expected, focal slowing is almost never seen, and paroxysmal spike activity is quite rare, although spike activity has been observed in cases of hypocalcemia, acute intermittent porphyria, and exposures to some toxic substances (e.g., lead, carbon monoxide). Early hepatic disease is commonly associated with generalized slowing in the theta range, which may increase in severity with elevated blood ammonia levels. If the disorder progresses to hepatic encephalopathy, a distinctive and diagnostically relevant EEG pattern called *triphasic waves* (or, more informally, *liver waves*) can emerge, characterized by frontally dominant or diffuse 1.5- to 3.0-per-second high-voltage slow waves initiated by blunt or rounded spike-like transients. For a thorough review of the effect of various metabolic disorders on EEG, the reader is directed to referenced chapter by Pang and colleagues.

Psychiatry Overview

Starting in approximately 1938, a flurry of continuing EEG investigations began to reveal an increase in the overall prevalence of minor abnormalities in almost all psychiatric populations as compared to healthy or nonpsychiatric individuals, which continue to be investigated to this day. On the other hand, two major factors led to the rapid disillusionment of the field of psychiatry with EEG. The first was the lack of specificity of EEG abnormalities to known psychiatric syndromes. The second factor, alluded to previously, was the continuing discovery of EEG abnormalities correlating with epilepsy, tumors, encephalopathies, stroke syndromes, and coma. With discoveries of significant EEG changes accompanying neurological problems on the rise and EEG abnormalities associated with psychiatric symptomatology remaining to be relatively obscure and noncontributory to the diagnostic process, the field of *clinical EEG* (and later *clinical neurophysiology*) became more firmly established as a subspecialty of neurology, with de-emphasis of EEG among psychiatrists.

The recent surge of interest in the neurobiology of psychiatric disorders, the emergence of the clinical field of *neuropsychiatry,* and the unprecedented advances in computerized analyses of EEG and other neurophysiological signals have resulted in a strong rekindling of interest in electrophysiology among psychiatrists. In 1995, John Hughes undertook the mammoth task of compiling a comprehensive outline of the broad area of EEG and psychiatry with 181 significant references selected for citation from before 1950 until 1994. When such compilations are inspected, the findings reveal that more than one-half of the EEG–psychiatry references appear after 1980, and one-third were written within the 5 years preceding Hughes' 1995 report. Continued inspection of the literature indicates that this trend has not abated. A *PubMed* search using the terms "EEG + Psychiatry" identified an additional 6,144 articles published in this field

since 1995, including both clinical and basic science research. Furthermore, two scientific organizations have been established for the primary purpose of accelerating the translation of electrophysiological research findings to clinically utilizable laboratory tests. In 1991, the American Psychiatric Electrophysiology Association (APEA) was founded. In 1999, the APEA merged with the American Medical EEG Association (AMEEGA), and the EEG and Clinical Neuroscience Society (ECNS) was formed (http://www.ecnsweb.com).

In spite of the enormous research potential recognized in psychiatry, electrophysiological methods are still of limited impact in clinical practice, where their primary applications are to aid in differential diagnostic purposes (i.e., the exclusion of "organic" brain pathology), and to the investigation of sleep disorders. There is no accepted indication of EEG methods for the diagnosis DSM-5 psychiatric disorders or for drug treatment choice and monitoring. This may seem surprising when considering that (1) robust results of quantitative Q-EEG abnormalities have been reported and independently confirmed for major psychiatric disorders, especially for schizophrenia; (2) evidence has been provided that abnormalities of event-related potentials and of eye movements are related to genetic risk factors, symptom dimensions, prognosis, and diagnostic subtypes of schizophrenia and depression; and (3) EEG and related methodologies already provide translational paradigms and have demonstrated promise in predicting the likelihood that novel therapies and compounds will exhibit clinical efficacy in preclinical development as well as in clinical trials. However, it should also be acknowledged that despite acceptance of psychiatric disorders as "brain disorders" there are currently no tests from any discipline of brain imaging or behavioral assessment that have been formally integrated into clinical diagnostic formulation. Perhaps this disconnect reflects the fact that it is not brain function, *per se,* but disordered behavior (e.g., reduction in normal routine behavior, elevation in abnormal behavior causing distress to self/others) that signifies the need for clinical attention and, thereby, serves as the basis for diagnosis and treatment decisions. In this respect it is understandable that even the strongest evidence linking an electrophysiological feature to a symptom or disorder may not necessarily indicate use of that feature as a diagnostic measure when clinical symptoms are acutely present.

Where then could electrophysiology find a place in applied clinical psychiatry? One promising area is in the early detection of pathophysiological processes during preclinical periods that could eventuate to psychiatric illness. For example, the prodromal period of psychosis is now well recognized as a stage of disease progression during which at-risk individuals (e.g., presenting genetic liability, attenuated symptoms) begin to express functional impairment, which may correspond to a combination of factors affecting critical periods of brain maturation.

Early detection is currently a topic of great interest to psychiatric research as well as to the *National Institute of Mental Health* (NIMH), who, under directorship Thomas Insel, coalesced 8 independent research laboratories in forming the *North American Prodromal Longitudinal Study* (NAPLS) consortium. Early *p*sychosis *p*rediction and *p*revention (i.e., EP3) was subsequently identified as an NIMH priority issue in 2013, with plans to focus funding efforts on this area for the following 5 years. These developments are consistent with aims to reduce disability of severe psychiatric illness through early detection and intervention, reducing long periods of untreated illness that may precede a first clinical encounter. EEG methodology has figured prominently in the conceptualization and data collection of NAPLS investigations, with paradigms demonstrated to be sensitive to the pathophysiology of schizophrenia used as predictors of conversion to formal psychotic disorder during the prodromal stage. EEG has provided a highly valuable assessment tool for this work, having advantages of being low-risk, noninvasive, and well-tolerated by this study population. Importantly, EEG has established sensitivity to normal neurodevelopmental processes that are hypothesized to be altered in the schizophrenia disease process. Finally, it should not be overlooked that EEG systems can be purchased at relatively low cost (~$60k), enabling equipment to be standardized across centers participating in multisite research. Innovations of the NAPLS study have included traveling subjects and additional calibration procedures to assess reliability of instrumentation and measures across sites, and has firmly established the feasibility of large-scale data collection efforts in psychiatry.

Nonspecificity of Results

Two particular problems plague and hinder research in the field of EEG abnormalities in psychiatric populations. First is the belief that EEG abnormalities can be found in up to 20 percent of otherwise healthy individuals. The second major problem is another pervasive belief that interictal epileptiform activities, in the absence of clinically diagnosable seizures, are incidental (or nonspecific) findings without clinical relevance. Well-designed studies addressing these two important issues are needed to allow further research in this field to be more meaningful.

Moreover, there are only a limited number of ways in which brain electrical activity can respond to normal or pathological influences. Brain waves change simply by becoming faster or slower in frequency or lower or higher in voltage, or perhaps some combination of these two responses. Thus, very similar patterns of EEG abnormality can emerge from different etiological causes. For example, a neoplasm, subdural hematoma, brain abscess, cerebral vascular accident, or closed head injury, may result in similar, although not always identical, focal EEG slowing. Generalized slowing is a common abnormal finding for which the etiological causes

include cortical atrophy, drug-induced encephalopathy, ECT, encephalitis, certain endocrine disorders (hypothyroidism and hypopituitarism), porphyria, head trauma, lead exposure, hypocalcemia, and Wernicke encephalopathy, to name but a few. The nonspecificity of results or the failure to specifically denote etiology is a genuine limitation, but one that is not as bleak as one may think.

Concerns raised regarding the sensitivity and specificity of EEG abnormalities does pertain primarily to conventional clinical EEG, which realistically only permits examination of the few dimensions of data that can be ascertained by visual inspection. The same concerns, though still appropriate, may in time be alleviated to an acceptable level given advances in the field of Q-EEG. Q-EEG profiles, as well as deviations on evoked responses, have produced high sensitivity and specificity in classification analysis comparing subjects with psychiatric disorders to healthy comparison and other diagnostic groups. Although, by and large, Q-EEG remains a research tool, its promise for aiding the differential diagnostic process and predicting treatment response to pharmacological agents is becoming clearly evident.

Applied EEG in Psychiatry

As stated previously, the role of EEG in applied clinical psychiatry remains fairly limited to differential diagnosis of organic etiologies. In this regard, the EEG can be most useful when the presenting symptoms, age of first symptom onset, or response to treatment is atypical, suggesting that a more structural cerebral pathology may be contributing to the syndrome being evaluated. In such cases, the presence of paroxysmal EEG activity or focal or diffuse slowing should be ruled out (see the example in Fig. 1.15–36). If paroxysmal activity is identified, then it is important to be clear that such an EEG finding is not sufficient grounds to diagnose epilepsy, unless, retrospectively, the symptoms can be seen as being of an epileptic nature. The presence of a paroxysmal EEG abnormality in a patient with an atypical presentation, particularly in the absence of a family history of the diagnosis that is being considered, should lead the clinician to be wary of and less confident in the diagnosis. It may be a good practice at the current state of knowledge to take advantage of the *not otherwise specified* category. Whether anticonvulsants are indicated in these situations remains a question open to clinical judgement. Early studies in this area suggested that anticonvulsants can sometimes block EEG epileptiform discharges and also can lead to dramatic clinical improvement in individuals exhibiting aggressive behavior. Furthermore, carbamazepine (Tegretol) was thought to be clinically useful in individuals with schizophrenia who exhibited temporal lobe EEG findings and aggressive tendencies without history of seizure disorder. Based on current practice, it would seem as though prescribers are comfortable using adjunctive

anticonvulsant therapy for many indications, irrespective of EEG findings, as would be suggested by common prescription of this medication class in anxiety disorders, dementia, schizophrenia, depression and bipolar disorders, and in cases of drug and alcohol withdrawal syndromes. Prescription in these cases, perhaps with exception of drug withdrawal, is not based on question of seizures but, rather, on the efficacy of anticonvulsants for treating psychiatric symptoms including aggression, mood swings, mania, and panic symptoms.

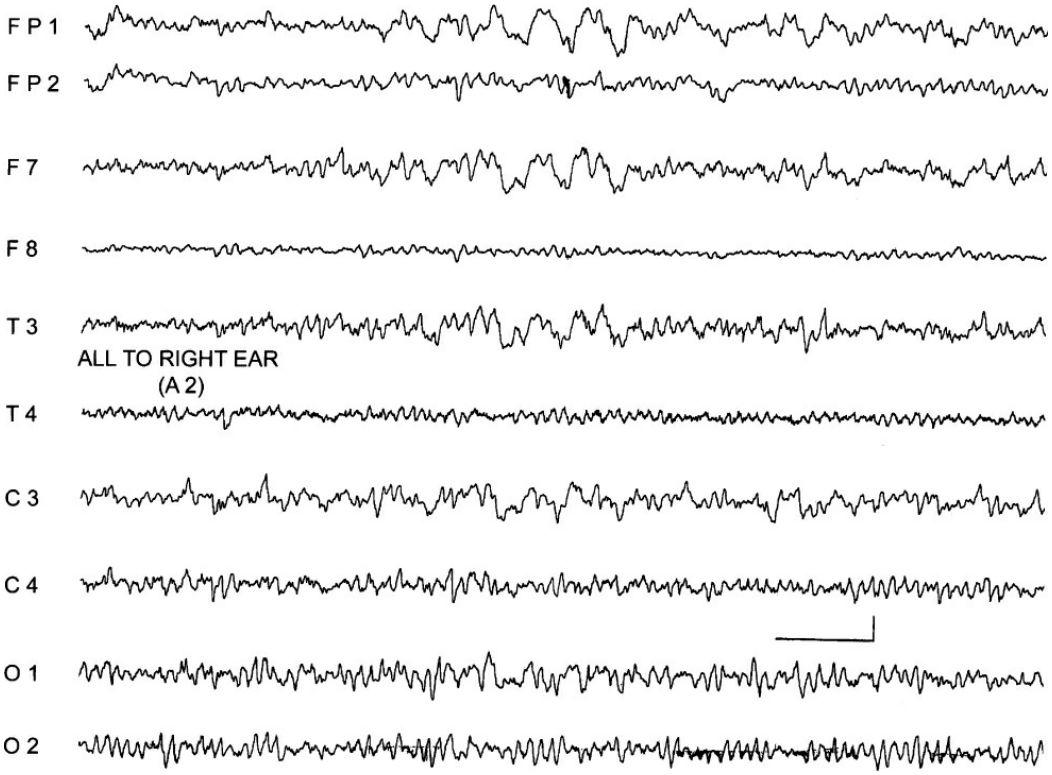

FIGURE 1.15–36. The patient is a single female, 40 years of age admitted to psychiatry for "functional" anorexia nervosa. This was the first episode with no prior psychiatric history. Patient would frequently vomit after meals and insisted that it was not self-induced. Medical examination before admission was negative. Routine admission electroencephalography: Focal slowing (delta), left frontal-temporal with reduced voltage spread through left hemisphere. Left frontal–temporal deep neoplasm found. Neurosurgical intervention was unsuccessful.

Similar comments can be made regarding the presence of a focal slow wave abnormality in the EEG. A minimal focal slow wave abnormality is unlikely to reflect a serious space-occupying lesion in a patient with a normal neurological examination. However, if no logical explanation for the minor focal slowing is apparent (e.g., old head injury) and the finding did not appear on earlier EEGs, then a repeat EEG may be valuable to determine whether the finding is static or progressive. Clinicians should use their clinical judgment regarding when imaging (and what imaging) should be performed. If a space-occupying lesion is suspected, then a CT scan is usually sufficient. An MRI is more likely to yield abnormalities if the

status of the lesion is less acute (e.g., old head injury). The presence of a focal slow abnormality, similar to the presence of paroxysmal activity, should cause the clinician to be wary of the diagnosis. Consideration also should be given to obtaining a neuropsychological evaluation to assess whether the focal cerebral abnormality identified by the EEG has any clinically correlating deficits.

The situation is different when an EEG is obtained in the course of managing an unstable or nonresponsive patient. The most important data to be gained from an EEG performed on an acutely agitated patient are the presence of diffuse slowing (indicating a delirious encephalopathic state) or ongoing epileptic activity NCSE. In both conditions, the EEG abnormality is fairly obvious and can be detected in a short period of time by temporarily retraining the patient, if needed, and using an electrode cap to minimize setup time. The EEG can be extremely useful in ruling out delirium secondary to medication toxicity, and, in this respect, it should be noted that diffuse EEG slowing (suggestive of a diffuse encephalopathic process) may be seen despite serum plasma levels within the therapeutic range.

When dealing with geriatric populations, sensitivity to organic pathology should be particularly heightened. In addition to the evaluation for dementia and delirium, EEG could be invaluable when there is a suspicion of seizures, and, in this respect, it is noted that elderly patients may account for as many as one-fourth of the cases of new onset epilepsy. One also must keep in mind that a normal EEG in a patient diagnosed with dementias, once past a mild stage, should raise the suspicion that a depressive disorder, or other reversible cause, may be contributing to the clinical picture.

EEG Findings in Psychiatric Disorders

In addition to the general considerations regarding the interface between EEG and psychiatry, a number of psychiatric disorders deserve specific mention. Findings deemed promising for clinical applications will be briefly reviewed.

Schizophrenia. EEG abnormalities have been extensively examined in schizophrenia patients. A literature review conducted by Boutros and colleagues in 2008 examined whether findings on EEG spectral abnormalities warranted additional effort toward developing them into a clinical diagnostic test for schizophrenia. Fifty-three papers were classified based on a four-step approach for evaluating the clinical usefulness of a diagnostic test (Table 1.15–1), and 15 of the papers were included in a meta-analysis. The review and meta-analysis revealed that most of the abnormalities were replicable across studies, with the most consistent results related to the prominence of resting low-frequency schizophrenia

patients regardless of medication status.

Table 1.15–1.
Proposed Four-Step Approach

Step	Design	Purposes	Desired Outcomes
1	Target group vs. healthy controls.	1. Demonstration of significant deviance in the target group. 2. Demonstration of test–retest reliability of finding.	Provide evidence of a consistent biological abnormality in the target group.
2	Target group vs. healthy and appropriate patient control groups.	Demonstration of significant differential prevalence of abnormality between illnesses that frequently need to be differentiated from one another.	Demonstration of potential clinical utility.
3	Target vs. proper control groups (may include within target group subpopulations).	Definition of test-performance characteristics.	Defining clinical utility.
4	Same as in Step 3 but across centers, ideally in a multicenter design.	Demonstration and standardization of clinical application.	Setting up standards for clinical application.

Suggested guidelines for evaluating the clinical usefulness of a diagnostic test, as published by Boutros and colleagues (2007) and applied by Boutros and colleagues (2008).

A recent study published by the *Bipolar and Schizophrenia Network on Intermediate Phenotypes* (BSNIP) multicenter consortium examined the polygenic architecture of resting state EEG frequency activity in 105 schizophrenia, 145 psychotic bipolar disorder, and 56 healthy individuals using 1 million single-nucleotide polymorphisms (SNPs) derived from DNA. Multivariate data clustering and component extraction methods identified five frequency and nine SNP components, with theta and delta activity found to correlate with gene components involved in brain development, neurogenesis and synaptogenesis. Increased anterior delta and posterior theta abnormality was present in both schizophrenia and bipolar samples, while theta differed between the two disorders. Theta abnormalities were also mediated by unique gene clusters involved in glutamic acid pathway and synaptic cell adhesion processes. These results replicated prior findings pertaining to excess resting slow wave activity in schizophrenia, and further differentiated the component structure of this abnormality from psychotic bipolar disorder patients, thus elevating specificity of this feature to a diagnostic differential that is difficult to make using conventional clinical assessment. These findings further point to plausible genetic networks, consisting of genes previously implicated in schizophrenia as well as novel mechanisms, that could underlie abnormal features observed in resting EEG.

Catatonia. Although the routine EEGs of catatonic patients with known functional etiology tend to be normal, catatonia could be a symptom of a number of serious medical conditions, including epilepsy and encephalopathy secondary to a variety of general medical conditions. Neuroleptic malignant syndrome has presented, at times, as catatonia, and, in such cases, the EEG is likely to show a picture of diffuse slowing indicative of an encephalopathic process. Although a well-known patient whose usual presentation includes catatonic symptoms may not need an EEG, it is advisable to obtain an EEG on every new patient who presents with catatonic symptoms.

Panic Disorder. Panic attack symptoms carry a significant resemblance to symptoms induced by temporolimbic epileptic activity, particularly those originating from the sylvian fissure. Fear, derealization, tachycardia, diaphoresis, and abdominal discomfort are characteristic symptoms of simple partial seizures with psychiatric and autonomic symptomatology. Studies comparing symptomatology of patients with panic disorder to patients with CPSs have reported much similarity, suggesting that there may be a common neurophysiological substrate linking the disorders. Notable conclusions of this literature are that individuals suffering from panic disorder show a higher rate of unspecific EEG abnormalities that include a unique increase in beta power, thought to align with the common clinical observation of hyperarousal in these patients. Although some evidence may suggest that IEDs occur at higher than normative baseline rates in panic disorder, relationships between panic symptoms and seizure activity remains far from clear, with EEG findings ranging from paroxysmal epileptiform discharges to asymmetrical increases in slow wave activity and focal slow wave abnormalities (reportedly observed in up to 25 percent of patients with panic disorder). Along these lines, it has been suggested that use of anticonvulsant drugs, including VPA (Depakene), be considered in cases of panic disorder if conventional anxiolytic treatment is inefficacious.

On the other side of the coin, panic attacks also have been demonstrated to occur more frequently in epileptic patients. In some cases, these attacks could lead to overmedication for seizures if their nature is not precisely defined. Again, EEG monitoring during multiple attacks is indispensable in rendering such an evaluation. Other reports have concluded that the most common psychiatric disorder that must be differentiated from temporal lobe epilepsy is panic disorder. Epileptic seizures are commonly briefer and more stereotyped than panic attacks. Additionally, aphasia and dysmnesia often accompany seizure activity.

Obsessive-Compulsive Disorder. EEG studies in patients with OCD reported a widespread increase of slow waves, with additional report of epileptiform activities over the left temporal lobe. The frequency of these

abnormalities varies among studies.

Q-EEG studies have more often involved the anterior regions of the scalp, supporting the hypothesis of a frontal dysfunction in the pathogenesis of OCD. The cingulate gyrus, in particular, has been implicated as a locus of dysfunction by Q-EEG topographic brain mapping. Data concerning the involvement of a specific frequency band have been heterogeneous, probably due to differences in methodologies, small study sample sizes, and a broadly heterogeneous patient population in terms of the presentation of OCD symptoms. Although not given the coverage that might be deserving of the topic in the current chapter, EEG biofeedback or, more precisely, *neurofeedback,* has gained interest as a nonpharmacologic treatment option for individuals suffering OCD. Basically, neurofeedback involves a graphic representation of ongoing EEG to the patient, who may learn to implicitly modulate activity upward or downward toward what would be regarded as normalization from a pathophysiological baseline. Again, although varied in terms of the specific EEG frequencies and locations involved in OCD therapy, neurofeedback has been approved by the FDA as a treatment for stress management and, like stress, may benefit from improved self-regulation that can be achieved as a common therapeutic mechanism of neurofeedback.

Mood Disorders. Depression may result from a variety of brain and other somatic diseases (e.g., tumors, dementia, and vascular accidents) that EEG can help to identify. With this exception, diagnosis of mood disorders is typically not informed by EEG. Nevertheless, research findings of EEG abnormality in depression, starting with the first publication by Fredrick Lemere in 1936, have provided many intuitive interpretations relating the clinical features of depression to brain mechanisms. Of note, Lemere's first paper, in which he proposed *"…The ability to produce 'good' alpha waves seems to be a neurophysiological characteristic which is related in some way to the affective capacity of the individual…,"* followed Hans Berger's original report on the human EEG by only a few years and, indeed, referred to 10 Hz alpha waves as "Berger rhythms." Observations involving abnormality in the alpha frequency have continued to be the major story pertaining to EEG research and conceptualization of depression up through current times. Since the early 1980s, work credited to Richard Davidson and colleagues has highlighted frontal alpha asymmetry and the associated *Approach-Withdrawal* model (aka the dispositional model) as an explanatory synthesis of the neural basis of depression. Briefly, an asymmetry of frontal alpha involving hyperactivation of the right hemisphere relative to the left has been interpreted as the interplay of *approach* (facilitating appetitive behaviors in association with positive affect) and *withdrawal* (facilitating avoidance of aversive stimuli in association with negative affect), with reduction in left-sided alpha thought

to reflect deficient approach system activation in depression. The model further stipulates that relatively higher right-sided withdrawal system activation contributes to anxiety and other forms of psychopathology. EEG research has continued to support this seminal finding of frontal alpha asymmetry and, to some extent, this formulation of depression, with studies by other groups suggesting that alpha asymmetry more likely reflected a state rather than trait feature in major depression. New models based on alpha asymmetry have since been proposed, including the capability model, which postulates that brain activity elicited under emotional challenge will be a stronger indicator of psychopathology than when measured at rest, and therefore entails testing with emotionally provocative stimuli.

Few large scale EEG studies have been carried out in acute, drug-free subjects during a manic episode, probably due in part to the difficulty of obtaining such data. However, Joyce Small and colleagues published several papers in the late 1990s on a sample of approximately 200 patients admitted to the Larue D. Carter Memorial Hospital of Indianapolis in acute manic states, many of whom received follow-up EEGs when medicated, after medication-washout, and when rehospitalized during subsequent episodes. Q-EEG was also performed in a subsample of about 40 patients who were cooperative enough to test. Remarkably, the majority of EEG data collected was interpreted to be normal. A slight increase in abnormal features (16 percent of sample) was associated with positive family history of mood disorder and spike–wave activity was observed in 17 percent of the sample. Q-EEG differentiated patients on the basis of prescribed pharmacotherapy, with higher theta activity at baseline predicative of poorer response to psychotropic drugs. Abnormal hemispheric laterality was detected in a smaller portion of patients (9 percent), which implicated left hemisphere involvement.

Recent evidence suggests the usefulness of baseline Q-EEG measurements in the prediction of positive as well as negative outcomes of treatments with antidepressant drugs. Andrew Leuchter and his team introduced cordance, a Q-EEG measure combining relative and absolute power, to investigate correlates of response to fluoxetine in unipolar depression, with a double-blind, placebo-controlled paradigm. Significantly more depressed subjects with high cordance in the theta band responded to fluoxetine in comparison to those with low cordance. The results were interpreted as suggesting that only depressed subjects with low baseline dysregulation respond to treatment. More recent studies conducted by Leucther's group have also found change in prefrontal cordance measures to be predictive of a positive response (e.g., reductions in Hamilton Depression Scale scores) to placebo conditions, however, the direction of change in concordance indicating improvement has not been entirely consistent across studies.

Attention-Deficit/Hyperactivity Disorder. ADHD is a prevalent disorder among children and adolescents, as well as adults. The prevalence of abnormalities in the conventional EEG is higher in ADHD than that in age- and gender-matched healthy control subjects, with approximately one-third of ADHD children having abnormal routine EEGs and a significant proportion exhibiting spike or spike–wave discharges. Some studies even report abnormal EEG incidence to be as high as 60 percent for this disorder. In addition, an association between child and adolescent ADHD and the controversial 14- and 6-per-second positive spike pattern also has been reported. Writing in a 2001 issue of *Pediatric Neurology,* Sarah Hemmer and colleagues reported that 15.4 percent of 234 ADHD children had frank epileptiform EEG abnormalities (controversial waveforms were not included) and went on to offer the extremely important observation that epileptiform EEG findings might increase the risk for seizures with methylphenidate (Ritalin) therapy. In their data, there was a seizure incidence of 0.6 percent in patients with pre-methylphenidate normal EEGs, as contrasted with a 6 percent seizure incidence in patients with pre-methylphenidate epileptiform EEGs (an example of premedication seizure discharges in a patient developing iatrogenic grand mal seizures after neuroleptic treatment was given earlier in Figure 1.15–34.

Q-EEG has identified increased theta activity, particularly in the frontal lobe regions, as a consistent finding in ADHD. Even when learning disability is removed as a confounding variable, increased theta remains as one of the most prevalent EEG findings in this disorder, replicated by a number of groups from different parts of the world. In addition to being a consistent finding, the theta increase seen in ADHD is resistant to age effects, whereas abnormalities of beta EEG frequency tend to decrease with age. It also should be noted that early work suggests that increased theta abnormality frontally may be a strong predictor of response to methylphenidate and other psychostimulants and that favorable clinical responses may be associated with a normalization of the EEG abnormality. Because decreased beta activity has sometimes been reported in ADHD, some investigators have tried to examine the theta-to-beta ratio as a stronger measure of ADHD-related EEG abnormality with a higher detection power. There are, however, additional Q-EEG reports of an increase (instead of decrease) in beta activity in similar patients, raising questions concerning whether or not different subtypes of ADHD differentially affect beta activity. Finally, discriminant function analyses of Q-EEG data have been shown to be capable of significantly discriminating between healthy, attention deficit disorder and ADHD, and learning disability subpopulations.

Dementia and Delirium. Because patients with advanced dementia rarely have normal EEG, a normal test result can play an important role in

diagnosing cases of pseudodementia (symptoms of dementia secondary to depression or psychosis). When dementia and depression coexist, it becomes important to have some idea about the relative contribution of each disorder to the overall clinical presentation. In a comparison of the EEGs of patients with depression, dementia, pseudodementia, and dementia plus depression with the EEGs of normal age-matched controls, the degree of EEG abnormality shows a significant inverse association with clinical response to antidepressants.

The routine EEG is also useful in following the progression of Alzheimer disease, and may be useful in differentiating the Alzheimer etiology from that of frontotemporal dementia (FTD), which has been associated with either normal routine EEGs or mild frontal slowing despite progressive behavioral deterioration. Relatively normal EEG in FTD should therefore not be taken as an indication that behavioral features in these patients result from primary psychiatric causes.

A number of recent publications strongly support the utility of Q-EEG as diagnostic and prognostic tests for early Alzheimer disease. Collaborative work between centers in Stockholm and New York City, under the leadership of Christoph Lehmann, provided diagnostic efficiency statistics for EEG-based classification of mild Alzheimer disease from healthy controls of 85 percent sensitivity and 78 percent specificity. On the basis of available literature, EEG measures are likely to rapidly emerge as a cornerstone of diagnostic workup for suspected Alzheimer disease.

The differential diagnosis of acutely disturbed and disorganized patients often includes delirium. In acutely agitated delirious patients, the EEG is often helpful in indicating whether the alteration in consciousness is due to (1) a diffuse encephalopathic process, (2) a focal brain lesion, or (3) continued epileptic activity without motor manifestations. Most often, patients with delirium have a toxic metabolic encephalopathy. In general, with the progression of the encephalopathy, there is diffuse slowing of the background rhythms from alpha (8 to 12 Hz) to theta (4.0 to 7.5 Hz) activity. Delta activity (<3.5 Hz) usually does not become prominent until the patient approaches nonresponsiveness. The major exception to this rule is seen during withdrawal from alcohol and during delirium tremens. Instead of diffuse slowing, as described previously, excessive fast activity dominates the EEG in patients with alcohol withdrawal delirium. Patients in alcohol withdrawal who are not delirious could have a normal EEG. Low-voltage fast activity also can be induced by benzodiazepines, and this should be borne in mind when interpreting the EEG. The EEG thus can be helpful in differentiating between alcohol withdrawal delirium tremens and encephalopathies from other causes, including iatrogenic encephalopathy. The seminal work of George L. Engel and John Romano described the clinical features and the role of the EEG in detecting delirium. The typical changes are slowing or even loss of the posterior dominant rhythm, slow-

wave increase, reduced reactivity to eye opening, and intermittent, especially frontally located, rhythmic delta activity. These neurophysiological changes relate to the dopaminergic–cholinergic dysbalance as the final common pathway of delirium with a lack of cholinergic stimulation. In contrast to delirium cases revealing a cholinergic deficit as a common pathophysiological feature, alcohol or benzodiazepine abstinence delirium results in decreased EEG voltage and an increase of fast activities, while nonconvulsive seizure state may be indicated by rhythmic activity and typical epileptiform discharges. Serial EEG assessments are considered the gold standard of delirium diagnostics.

Q-EEG is capable of detecting subtle changes and is especially useful for the evaluation of frail elderly with a high prevalence of dementia, accumulated organic brain disease, and age-related EEG changes. Using relative power measures alpha decrease and delta increase, as well as reduction in peak or mean frequency and diminished fast-to-slow-wave ratios, have been found to indicate delirium. In the presence of dementia, Q-EEG greatly improves delirium detection and monitoring, especially when supplemented with a simple activation procedure as demonstrated by Charles Thomas and colleagues.

Autistic Spectrum Disorders. Several types of EEG abnormalities have been found in ASDs including epileptiform activity, background slowing, and abnormal EEG hemispheric asymmetry. Altered mu rhythms (transient, nonoccipital alpha rhythms that occur over motor cortex) have also been reported, which some have postulated to indicate altered "mirror neuron" function, referring to neurons that fire both when an individual performs an action and when observing the same action performed by another. Of the various EEG abnormalities reported, epileptiform activity has been the most consistent finding in ASDs, detected at rates of 18.9 percent up to 60.7 percent of individuals with autism. Interestingly, approximately one-third of children with ASDs go on to develop epilepsy. That being said, several reviews have concluded that there is no justification for the use of anticonvulsant medications or surgery in children with pervasive developmental disorders without seizures; that is, there is no evidence that treatments to eliminate EEG spikes will have a therapeutic benefit on the behavioral abnormalities in autism.

EEG ALTERATIONS FROM MEDICATIONS AND DRUGS

A great many medications, as well as substances consumed for therapeutic or recreational purposes, can produce some degree of alteration in the EEG. This section attempts to highlight those compounds most relevant to clinical and research psychiatry.

Pharmaco-EEG

During the 1970s, several laboratories consistently obtained a classification of psychotropic drugs on the basis of EEG changes induced by a single dose of each of these drugs in healthy subjects. This history was reviewed by Silvana Galderisi and Walter G. Sannita. The availability of Q-EEG analysis led to the development of a new research field, *pharmaco-EEG*. Pharmaco-EEG methods were included in preclinical studies to identify, at early stages of drug development, the therapeutic indications of new drugs, test drug bioavailability at the CNS level, determining onset, peak effect, and duration of its CNS effects, predict therapeutically useful dosages, and compare the bioavailability of different forms of psychotropic drugs (e.g., oral vs. parenteral). Pharmaco-EEG studies identified the antidepressant activities of mianserin and doxepin, which had been classified as antiallergic and anxiolytic, respectively, by preclinical studies, as well as the sedative activity of fenfluramine, classified as a psychostimulant by animal studies. The discovery of the antidepressant activity of mianserin was instrumental in the development of new animal tests which, in turn, promoted the discovery of new antidepressants, such as fluvoxamine (Luvox) and fluoxetine (Prozac). For fluvoxamine and sertraline (Zoloft), pharmaco-EEG studies enabled the identification of therapeutic doses in early stages of drug development.

Few attempts have been made to translate pharmaco-EEG methods to psychiatric clinical settings, where this approach could inform the expectation for therapeutic benefit in response to a given psychotropic drug or dose level. Pharmaco-EEG methods for response prediction, where applied in research, mainly involve examination of Q-EEG changes following either a single administration of a drug (the so-called test-dose procedure) or at a point early in the course of treatment. Such methods, if proven capable of providing reliable prediction of the therapeutic response to treatment, could have an enormous clinical impact when considering that failure to respond to initial treatment tends to increase attrition from care, severity of prognosis, and ultimately the costs associated with psychiatric illness.

Based on a review conducted by Armida Mucci and colleagues, early evidence favored the use of pharmaco-EEG monitoring of first-generation antipsychotic- and lithium-induced EEG abnormalities, however, these procedures were never adopted as a clinical routine. Regarding early prediction of treatment response, several QEEG studies of first-generation antipsychotics (FGA) in schizophrenia produced consistent findings, but with ultimately no clinical impact. For instance, administration of high-potency FGA medication was found to produce an increase of alpha activity, particularly in the slow alpha range (7.5 to 9.5 Hz), that several independent groups found predictive of a favorable clinical response. Using the test-dose procedure, Silvana Galderisi and colleagues demonstrated that Q-EEG changes in the slow alpha band classified FGA responders

versus nonresponders with near 90 percent overall accuracy.

The test-dose procedure seldom has been used in pharmaco-EEG studies of response prediction in depressed patients. In a study by Martin Bares and colleagues, Q-EEG cordance (an index combining relative and absolute power) was measured pretreatment and following 1 and 4 weeks of an antidepressant therapy in a sample of treatment-resistant inpatients. Responders showed a decrease, while nonresponders showed an increase, in prefrontal theta cordance after the first week of treatment. These findings, in line with other brain imaging results, suggest that early changes of prefrontal activity mediate the clinical response to antidepressants.

In conclusion, Q-EEG indices could be a valuable complement to a patient's clinical assessment, providing further guidance to inform clinician's choice of the appropriate drug treatment. Toward this end, new guidelines have been published by the International Pharmaco-EEG Society (IPEG), with the goal of standardizing the acquisition and processing of data obtained in pharmaco-EEG studies. The introduction of these methods in clinical routine will require the replication of findings in large patient populations as well as the investigation of new drugs recently introduced in the clinical practice.

Psychotropic Medications—General Considerations

It is well known that psychotropic agents can affect the EEG. For the routine EEG, with the exception of the benzodiazepines and some compounds with a propensity to induce paroxysmal EEG discharges, there is little, if any, clinically relevant effect when the medication is not causing any toxicity. For example, benzodiazepines, even in small doses, generate a significant amount of beta activity that is seen diffusely (see below), and can last substantially longer (up to 2 weeks) than might be expected given the rapid response and peak occupancy (typically 1 to 4 hours) of these drugs. This response is so universal that it has been suggested that, if a particular brain region fails to exhibit the expected benzodiazepine-induced beta activity, that the area may be in some way damaged.

Psychotropic Medications—Toxic Effects. It has long been accepted that the EEG is sensitive to the neurotoxic effects of psychotropic medications, and clinical vignettes illustrating the value of EEG in detecting a neurotoxic reaction to medications commensurate with clinical deterioration have been reported. It is often stressed that such a scenario could occur at any time during the course of treatment, because many factors impact the patient, and the symptoms could be subtle. An EEG investigation may be useful when patients undergoing long-term therapy present with clinical deterioration, particularly if the patient is known to be taking the medication, and the serum plasma levels are within the therapeutic range.

Such clinical situations also may highlight the need for having baseline EEGs available for comparison when a patient presents with clinical exacerbation. Normative EEG databases are available, but are based on cross-sectional evaluations and do not take into account the dynamic nature of psychiatric disorders or changes due to medication status. Thus, for the purpose of detecting toxicity, routine clinical EEG remains of immediate relevance to the everyday management of patients prescribed psychotropic medication. The appearance of significant diffuse EEG slowing in a patient who is receiving psychotropic medications and whose clinical condition is not stable (particularly the elderly) should prompt the clinician to consider medication toxicity, in addition to other causes of encephalopathy (e.g., electrolyte imbalance and thyroid problems, to name only two).

Psychotropic Drug-Induced EEG Abnormality (Nonparoxysmal and Paroxysmal). Almost from the time psychotropic medications were introduced, it was known that some of these compounds could precipitate EEG abnormalities, including paroxysmal EEG discharges (spike and spike–wave activity) in some individuals. Usually, medication-induced paroxysmal EEG activity remains behaviorally silent and is not accompanied by iatrogenic overt seizure manifestations.

GENERAL ANESTHETICS. General anesthetics primarily induce their effects by binding to affecting neuronal ion channels. In terms of EEG signatures, anesthetics initially evoke a frontal fast activity followed by a disruption of alpha rhythms. As the anesthesia deepens, EEG activity becomes slower and of low voltage (not unlike deep sleep), and burst-suppression patterns can be observed. At the deepest levels, EEG activity ceases, and can even resemble patterns seen in comatose states.

BARBITURATES. Barbiturates are now minimally used for medical or recreational purposes. The mechanism of action of this class of drugs involves binding to a specific site on the $GABA_A$ receptor, thus potentiating the effects of the inhibitory neurotransmitter GABA. When barbiturates are taken for medical or recreational purposes, beta activity is introduced into the EEG, particularly over frontal regions. Higher doses are associated with lower frequencies and dissolution of the alpha rhythm (as drowsiness and sleep ensue). Abrupt withdrawal from barbiturates after long-term exposure can produce EEGs containing generalized paroxysmal activity and spike discharges, which often are not associated with overt motor seizure manifestations. These effects may be seen for days up to 3 to 4 weeks after the last drug exposure.

TRADITIONAL NEUROLEPTICS. The primary mechanism of action of the typical antipsychotics/neuroleptics (e.g., chlorpromazine, haloperidol) is

antagonism of the D2 dopamine receptor. EEG changes with the neuroleptics include an increase in alpha activity, a shift to lower frequencies, and a decrease in beta activity. At high doses, seizures and spikes may occur.

ATYPICAL ANTIPSYCHOTICS. The newer, atypical antipsychotic agents such as clozapine and risperidone have different mechanisms of action compared to traditional neuroleptics, and have affinity for 5HT2a, alpha-adrenergic, and D4 receptors, among others. EEG frequency changes during the administration of clozapine include increases in delta, theta, and beta, which are prominent over frontal, central, and parietal areas. Some paroxysmal slowing and spikes can occur, along with rare instances of seizure. Little is known regarding the EEG correlates of other atypical antipsychotic drugs.

TRICYCLIC ANTIDEPRESSANTS. Older generation TCAs (e.g., imipramine, dexepin) typically act as both serotonin and norepinephrine reuptake inhibitors. They can increase the amount of slow and fast rhythms, decrease the frequency of alpha, and can induce paroxysmal slow waves and spikes. At higher doses, they can induce seizures, which can be particularly concerning given that tricyclics have a somewhat narrow therapeutic range.

SELECTIVE SEROTONIN REUPTAKE INHIBITORS. By definition, SSRIs increase synaptic levels of serotonin by binding to the serotonin transporter. These agents tend to decrease alpha activity while increasing theta frequencies. Further, the rates of seizure are less in SSRIs compared to the tricyclics.

LITHIUM. Lithium is still commonly used in the treatment of bipolar disorder, and can induce marked changes in the EEG. Lithium salts, particularly at intoxicating levels, are associated with diffuse EEG slowing, and paroxysmal spikes and triphasic waves, which can remain after serum levels have normalized.

BENZODIAZEPINES. Like the barbiturates, benzodiazepines potentiate the action of GABA, although the mechanism of action entails binding to a separate benzodiazepine site on the $GABA_A$ receptor. As such, benzodiazepines are commonly used for their anxiolytic and anticonvulsant properties. Benzodiazepines at therapeutic doses are potent enhancers of beta rhythms, and in addition, slightly increase theta power. Decreases in alpha frequencies are also noted. Benzodiazepines are well known to reduce epileptiform activity, and indeed, are the first drugs of choice for status epilepticus.

OTHER THERAPEUTIC AGENTS. While beyond the scope of the current chapter, a host of other nonpsychiatric therapeutic agents have been shown to

induce robust effects on the EEG. Most notably, EEG changes in the context of antiepileptic drug treatment are well documented. Further, antibiotics, immunosuppressants, and antineoplastic agents, among others, have all been shown to affect EEG signatures. For a thorough treatment of these medication types, see referenced chapter by Bauer and Bauer.

Drugs of Abuse

The recreational use of addictive drugs remains a significant public health issue. In fact, in terms of cost of illness, it has been estimated that drug addiction ranks second only to cardiovascular disease. Hence, drug dependence is becoming of increasing concern for those involved in the assessment and treatment of psychiatric disorders. Nearly all drugs of abuse are capable of altering the frequency spectrum of the EEG, and the degree of alteration varies with recreational versus heavy use and whether the EEG was obtained during acute intoxication, during intervening periods of abstinence, or during clinical withdrawal in the addicted individual.

With only infrequent exceptions, the use of an abuse drug does not introduce frank clinical abnormalities into the visually analyzed EEG tracing. This is especially true for recreational drug use, and is even the case for dependent and addictive use as well. For this reason, drug abuse alone is not a sufficient reason for EEG referral. Although the alterations of EEG frequency and voltage in the visually analyzed EEG produced by many abuse drugs are often unimpressive, topographic Q-EEG analyses can reveal marked alterations of the EEG spectrum that constitute significant deviations from population norms (even though the clinical implications may not be clear). Nonetheless, it is important to appreciate the various effects of different drug classes on the EEG.

Alcohol. Ethyl alcohol, or ethanol, exhibits many of the same pharmacodynamics as the benzodiazepines (i.e., acting as a positive allosteric modulator of the $GABA_A$ receptor). Hence, many of its behavioral and EEG effects are similar. There is considerable consensus that an increase in the amount of alpha activity and a slight slowing of alpha frequency typically accompany alcohol consumption, and that higher blood alcohol levels increase the amount of theta. Some reports have indicated that chronic alcohol consumption may be associated with a lower voltage and slightly faster resting EEG, although the clinical relevance of this remains obscure.

Alcohol withdrawal can lead to epileptic seizures as a result of perturbations of GABA and glutamate neurotransmitter/receptor function. Higher-frequency beta activity may be substantially increased in the addicted alcoholic undergoing withdrawal, and, if delirium tremens complicates the clinical picture, then excessive fast activity may dominate

the EEG. EEG signatures during acute alcohol withdrawal can be desynchronized with low voltage, and generalized spikes may be observed. Alcohol-related seizures need to be differentiated from true epileptiform activity with comorbid alcohol dependence. Routine EEG can be extremely helpful in this regard. A normal EEG during periods of sobriety, particularly if associated with an abnormal EEG during early withdrawal, strongly suggests that the seizures are withdrawal-induced. It should be noted that the EEG tends to normalize faster (with the exception of generalized decrease in amplitude) during withdrawal from alcohol as compared to that from barbiturates.

Opiates. The effects of opiates such as morphine and heroin on the EEG are similar to those of alcohol and involve slight reductions in alpha frequency. They also may increase the voltage of the EEG, particularly the power of theta and delta activity. Opioid anesthesia has also been shown to induce spike discharges resembling benign epileptiform transients of sleep. However, when an opiate overdose produces a comatose clinical presentation, the EEG usually consists of clinically abnormal diffuse slowing.

Cannabis. Marijuana, or cannabis, remains one of the most prevalent drugs of abuse. Moreover, rates of its consumption are on the rise due to the trend of decriminalization and the acceptance of the use of "medical marijuana." Cannabis, or its primary psychoactive constituent delta-9-tetrahydrocannabinol (THC), affects the brain via central cannabinoid receptors (CB1Rs). CB1Rs act primarily by affecting the release of the neurotransmitters GABA and glutamate. As reviewed by Skosnik and Cortes, the effects of cannabis/THC on EEG can be broken down into effects on resting EEG (intrinsic rhythms), event-related potentials (ERPs), and evoked/induced oscillations.

Acute cannabis smoking typically produces minimal visible effects on traditional EEG. In terms of contemporary studies utilizing Q-EEG, chronic cannabis users typically show reductions in power in the alpha-, theta-, and beta-frequency bands while in the nonintoxicated state. Acute THC has been shown to reduce resting theta power, which has been correlated with altered memory performance. Both chronic and acute cannabis/THC have been shown to disrupt the P300 ERP, which is congruent with cannabis' effect on attention and memory. Finally, both chronic and acute cannabis/THC have been shown to disrupt sensory evoked and induced gamma rhythms, which may provide a mechanism for the perceptual-altering effects of cannabinoids.

Finally, it should also be noted that one of the other, nonpsychoactive constituents in cannabis (cannabidiol) has been shown to have antiepileptic properties. Hence, this may represent a potential future direction in the

"medicinal" use of cannabinoids.

Psychostimulants. Psychostimulants such as cocaine, amphetamine, and methylphenidate act primarily increasing central dopamine function. They typically increase the amounts of beta and alpha activity, and generally reduce EEG voltage. Overdose with psychostimulants produce typical coma patterns in the EEG. The effect of psychostimulants on ERPs have yielded mixed results, although they may exhibit some effects similar to those observed in disorders of dopamine function such as schizophrenia.

Inhalants. Inhalation abuse of volatile substances (e.g., airplane glue, cleaning fluid, paint thinner, and gasoline) can produce a nearly instantaneous sensation of euphoria, and, in the early period of use, there may be no obvious residuals after the acute response subsides. However, with continued inhalant abuse, serious neurological and neurocognitive deficits can emerge that may not be completely reversible with abstinence. The immediate effects of inhalation of volatiles on the human EEG appear not to have been well studied. Where persistent neurological or neurocognitive sequelae follow chronic inhalant abuse, clinically abnormal diffuse EEG slowing in the lower theta to upper delta range may be seen.

Hallucinogens. Drugs such as lysergic acid diethylamide (LSD) and mescaline appear to have only minor effects on the visualized EEG and do not produce clinically relevant changes. However, recent studies of the drug ecstasy (MDMA) have shown abnormalities in several ERPs, which may be related to the purported memory-impairing effects of long-term MDMA use.

Tobacco. Tobacco (nicotine) does not appear to produce dramatic alterations in the analog EEG. However, topographic Q-EEG analyses reveal striking EEG changes with acute exposure to, as well as withdrawal from, tobacco. The immediate effects of smoking include a decrease in slower frequencies (especially theta), increased power of frequencies in the upper one-half of the alpha-frequency band, and beta activity. Twenty-four hours of tobacco deprivation produce a marked decrease in alpha frequency, with a corresponding marked increase in the relative power (amount) of theta activity. The effects of acute smoking and abstinence are essentially opposite to one another. Further, nicotine has been associated with alterations in "sensory gating" as indexed by studies of P50 ERP suppression. However, these effects can vary depending on whether they are examined during acute nicotine exposure use or during withdrawal from nicotine.

Caffeine. The use of coffee and products containing caffeine are quite

prevalent. The use of caffeine is of little concern to the clinician interpreting the visually analyzed EEG. Withdrawal from caffeine in the caffeine-dependent individual, however, produces a markedly significant increase in the amplitude, or voltage, of theta activity—an effect that is reversible within 15 minutes of consuming one cup of coffee.

Other Drugs of Abuse. In recent years, several novel synthetic drugs of abuse have emerged, most notably "K2/spice," which are synthetic cannabinoids, and "bath salts," which are synthetic psychostimulants/hallucinogens typically in the cathinone family. However, little research has been performed on these substances, particularly in the context of EEG.

EEG AND EXPERIMENTAL RESEARCH PARADIGMS

The majority of EEG conducted in the field of psychiatry may indeed be for research purposes, including studies both on the intrinsic rhythms ascertained under normal resting conditions as well as a number of experimental paradigms associated with different sensory-perceptual and cognitive tasks. The high temporal resolution of EEG essentially provides "real-time" measurement of changes in brain activity in response to experimental conditions. EEG allows the experimenter to study changes in brain activity related to overt processes, such as differences in the auditory response to frequently versus infrequently presented stimuli, as well as more covert processes ranging from anticipation of stimulus delivery (e.g., contingent negative variation; CNV) to presupplementary motor area activation (e.g., lateralized "readiness" potential; LRP) in preparation of a button press, or consolidation-retrieval processes associated with working memory following presentation of a stimulus set. As illustrated by these examples, EEG is an ideal tool for assessing differences or change in neural activity generated as a sequence of mental operations over a time course. Regardless of whether or not such functions can be localized to discrete cortical or subcortical regions, which is never the first aim of EEG research, the ability to precisely record change in neural activity in response to known events (e.g., under controlled test conditions) enables the researcher to ask, and potentially answer, a limitless set of questions concerning human information processing and abilities affected by psychiatric illness.

Evoked and Event-Related Potentials

The genesis of experimental EEG research no doubt has roots in *evoked potential* (EP) methods, which continue to see routine clinical use in neurology practice as a test of CNS function. Briefly, EPs are electrophysiological responses to external stimuli, usually administered over several repetitions with changes in stimulus properties (e.g.,

frequency, intensity) across presentations. EPs can be generated in any sensory modality but in most clinical applications involve *visual evoked potentials* (VEPs) and *brainstem auditory evoked potentials* (BAEPs) measured by EEG, or *somatosensory evoked potentials* (SSEPs) measured electromyographically (i.e., EMG). An important point to keep in mind regarding the EP is that the response is time-locked to delivery of a stimulus and, therefore, can be strictly regarded as "evoked." This point distinguishes the EP from the conceptually, and semantically, similar *event-related potential* (ERP), which also represents an electrophysiological response to stimulation. However, the subtle difference that distinguishes these measures can be appreciated in the examples of so-called overt versus covert brain processes provided in the preceding paragraph; whereas the overt auditory-evoked response can be regarded as an EP, examples of ERPs including the CNV and LRP occur prior to stimulus delivery and, therefore, reflect brain activity engaged in process of performing mental operations that is not necessarily time-locked, or "evoked" in response, to a stimulus. The term "event-related" has therefore replaced "evoked" over time in the research arena, where EEG recording is used to study both "bottom-up" sensory-evoked responses as well as changes in neural activity associated with higher-order cognitive processes that may be represented by induced brain signals.

Another basic principle regarding the EP or ERP is that these potentials are usually of small magnitude (as small as a few microvolts) compared to the ongoing EEG activity (reaching 50 μV or more). This is due to the fact that EPs/ERPs represent only the components of the EEG that are consistently present across repeated trials of stimulus delivery, which may consist of hundreds of trials. The EP/ERP is extracted from the digital recording by averaging, resulting in a signal that reflects change in neural activity in response to stimulus events with ongoing background EEG activity (which is not time locked to the stimuli) canceled out by the averaging process. The resulting average is then graphed as a change in voltage over time plots and can be quantified by measures of peak amplitude and latency.

By convention, brainstem auditory EPs are labeled with Roman number I, II, III, IV, and V, reflecting the order of the component, with components I to III generated by the auditory branch of cranial nerve VII and components IV and V by the upper brainstem. ERPs (as well as SSEPs and VEPs) are labeled according to both their polarity and latency. The letters "P" or "N" indicate positive or negative direction of deflection relative to prestimulus baseline, though whether polarity is depicted graphically in up or down direction seems to be matter of personal choice and, in many cases, polarity does reverse depending on electrode location due to the dipolar distribution of associated electrical fields. Labeling nomenclature representing ERP peak latency is done more simply, with exceptions. While

the ERP associated with the label "P300" is a positive deflection peaking at around 300 ms following stimulus delivery, the short-hand "P3" can also be used, signifying the third positive deflection of the ERP, which follows P1 or P100 at around 100 ms and P2 or P200 at around 200 ms. Time-courses and conventional labeling of ERPs most commonly studied in psychiatry are presented in Figure 1.15–37.

When a stimulus is presented to the brain, it travels through the receiving sensory organ and the brainstem (except for visual and olfactory stimuli) on its way to the primary cortical regions representing the particular sensory input. EPs recorded over primary sensory cortices indicate how the stimulus was received and registered by the brain, thereby providing a test of the integrity of the underlying neural architecture. For instance, the BAEP is measured in EEG waves arising within the first 10 ms following delivery of an auditory stimulus, thought to reflect the intactness of the auditory pathway, and most are useful for the diagnosis of hearing loss. It cannot be stated with certainty at this time whether basic EPs are broadly or specifically affected in psychiatric disorders, as ERPs have been the focus of most contemporary neuropsychiatric research. However, disturbances in early sensory processes are thought to be fundamental to schizophrenia pathophysiology, for example, and may conceivably involve impairment detected at the earliest stages of sensory processing.

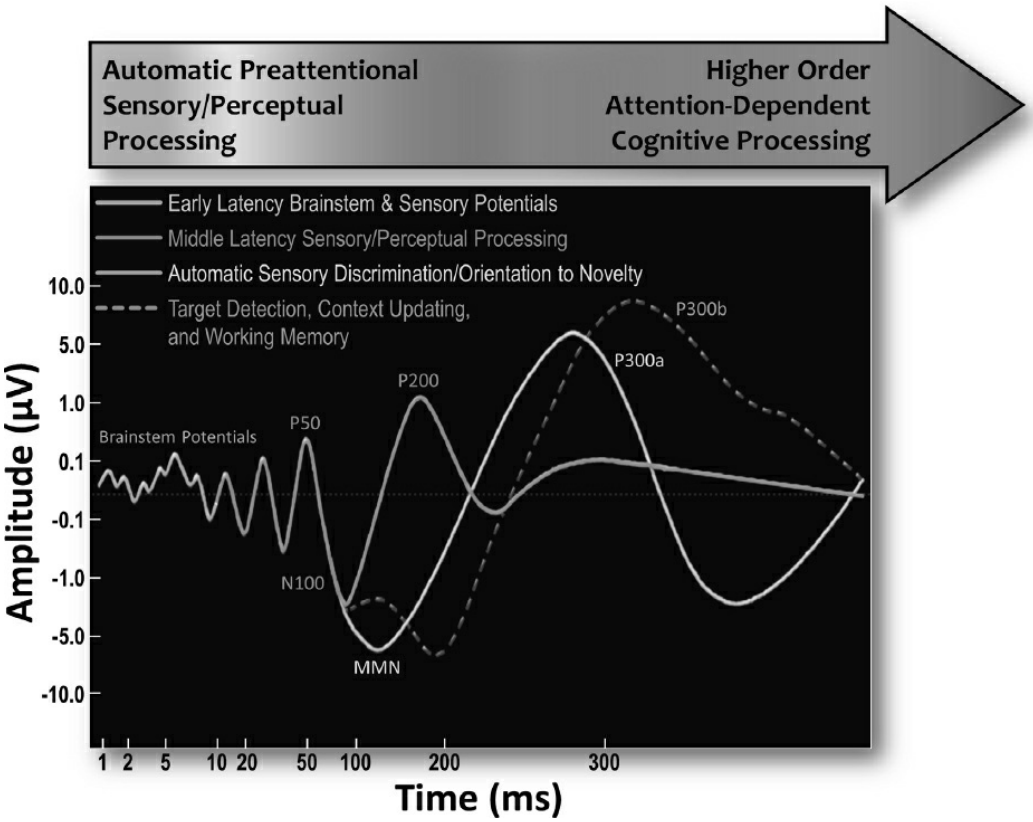

FIGURE 1.15–37. Schematic illustrating prototypical ERPs. Time is shown using a logarithmic scale. Stages of information processing from automatic preattentional to higher order attentional processes are shown from left to right, respectively. (Adapted from

Skosnik PD, Cortes-Briones JA. Electroencephalography (EEG) and cannabis: from event-related potentials (ERPs) to oscillations. In: Preedy VR, ed. *The Neuropathology of Drug Addictions and Substance Misuse*. Amsterdam, Netherlands: Elsevier Science; 2016.)

Regarding stage of information processing, a basic delineation has been made between *early-latency potentials,* which would include the BAEP and other components occurring before 10 ms poststimulus, *mid-latency potentials* occurring up to 200 ms, and *long latency potentials,* occurring after 200 ms. As a general rule of thumb, early-latency potentials are thought to mainly reflect processing dependent on *exogenous* factors, determined outside of the CNS, such as stimulus intensity or rate. Exogenous potentials can be measured regardless of a subject's level of consciousness, are thought to be fundamentally unaffected by cognitive processes, and represent basic sensory registration. Once received by the brain, a stimulus undergoes psychological processing and evaluation, and these operations influence later ERP waves that are broadly distributed outside of primary sensory cortices. While ERPs also contain exogenous components (though typically recorded at sampling rates too slow to accurately measure early latency potentials), mid-latency and long latency potentials tend to be of most interest to experimental research. In later stages of information processing, ERPs increasingly reflect activity relating to cognitive processes, such as anticipation, encoding, and recognition, thus can be influenced by level of alertness and attentional resources devoted to the experimental task. Accordingly, mid-latency and late ERPs mainly reflect *endogenous potentials,* with measurements of amplitude or other characteristics reflecting the interaction of stimulus properties and experimental task demands with individual differences in information processing. Because ERPs are sensitive to experimental manipulations, various paradigms have been developed to study brain function at the interface of sensory and cognitive processing. These paradigms have demonstrated many abnormalities during mid-latency and later processes across the psychiatry research literature, and thus are of considerable interest for their potential as clinical tools.

Middle Latency Components (P50 and N100)

Three of the most examined ERP components in the auditory modality are designated P50 (a positive component occurring between 35 and 80 ms after stimulus onset), N100 (a negative component occurring between 80 and 150 ms after stimulus onset), and P200 (a positive component occurring between 150 and 250 ms after stimulus onset). Figure 1.15–37 shows a typical example of these mid-latency ERPs. These components share the characteristic that their amplitude decreases with repetition (habituation or sensory gating). The P50 and N100 components have been studied extensively in psychiatric patients.

The P50 is a positive voltage, middle latency ERP component peaking at around 50 ms poststimulus, best elicited by discrete auditory stimuli (e.g., brief white noise clicks). In the standard P50 gating or "dual-click" paradigm, the amplitude of the P50 to a second paired click (S2) is attenuated relative to the P50 amplitude to the first click (S1). A 500-ms interstimulus interval is typically utilized, and is optimum to elicit gating. While the P50 to S1 is thought to be related to the capacity of the brain to register salient stimuli, the reduction of the P50 amplitude to S2 (compared to S1) reflects the automatic suppression or inhibitory gating of redundant and irrelevant stimuli. It is thought to be mediated by the hippocampus, temporoparietal region, auditory cortex, and prefrontal cortex. Deficient P50 suppression is commonly associated with schizophrenia, and has been proposed as a candidate endophenotype for psychosis risk (representing a physiologically detectable feature linked to underlying genetic cause of illness), supported in part by evidence for linkage to several psychosis risk genes and by similar reduction in P50 suppression observed in first-degree family members of schizophrenia probands. Evidence for an inhibitory abnormality (i.e., sensory gating deficit) in schizophrenia patients has been accumulating over the past 30 years. In an early meta-analysis of proposed endophenotypes of schizophrenia, sensory gating of the P50 response was found to represent a magnitude of deficit equal to or greater than the most robust findings reported across neuroimaging and behavioral studies.

The N100 component is a large exogenous ERP that exists regardless of task demands (although it can be modulated by attention). The N100 is generally associated with early perceptual processing with origins in the primary sensory cortices and is elicited in response to various forms of auditory and visual stimuli, regardless of experimental context. Of note, while the visual N100 is regionally localized over occipital and posterior parietal electrode sites, the N100 (like all auditory ERPs) tends to be maximal over the central or fronto-central sites, representing the confluence of potentials generated in bilateral auditory cortices. The N100 component has also been shown to exhibit suppression similar to the P50 ERP when assessed using the dual-click paradigm, and is similarly affected in schizophrenia patients. However, work by Jason Johannesen and colleagues has demonstrated that reductions in N100 amplitude in schizophrenia, whether elicited by paired clicks or by pure tones spaced approximately 1,200 ms apart, have equally large effect sizes when contrasted against healthy comparison subjects, and exceed effect sizes obtained by measures of P50 suppression and P300 amplitude and latency measures in the same sample. Furthermore, while P50 and P300 measures were not found to differentiate schizophrenia versus bipolar disorder above chance level, N100 provided unique variance in classification analysis differentiating these psychiatric groups with 72 percent overall accuracy,

only slightly less than achieved in an optimal model classifying schizophrenia against matched healthy participants (79 percent accuracy). Therefore, auditory processes indexed by the N100 appear to represent a stable (e.g., obtained in different experimental contexts), robust, and perhaps diagnostically specific, neurophysiologic deficit among individuals with schizophrenia.

P300

The P300 is a late positive, postattentional ERP component thought to be related to directed attention, contextual updating of working memory, and the attribution of salience to deviant or novel stimuli. This response is typically elicited utilizing standard "oddball" paradigms in which low-probability target stimuli (~10 to 20 percent) are embedded within a repeating sequence of high-probability standard stimuli (~80 to 90 percent) differing in some physical dimension (e.g., frequency, duration, amplitude). It this thought that the P300, particularly the P300b component (elicited in response to designated targets), reflects activity from a distributed neural network including areas, such as, thalamus, hippocampus, inferior parietal lobe, superior temporal gyrus, and frontal cortex. Cognitive and affective neuroscientists have been conducting basic research on P300 since it was discovered over 40 years ago, providing a rich knowledge base that has yet to be fully exploited by clinical scientists.

Mismatch Negativity

The mismatch negativity (MMN) is an automatic, preattentive, negative voltage ERP component that is generated primarily in the superior temporal and prefrontal cortex. The MMN is hypothesized to index the automatic, preattentive processing of the "mismatch" between standard and periodic deviant stimuli, indexed by the amplitude of a negative component over latencies of 100 to 200 ms measured in waveforms created by the subtraction of ERPs produced by the deviant and standard stimuli. The MMN experiment is similar to the P300 oddball paradigm where serial presentation of a standard stimulus (e.g., 10-ms presentation of 1,000 Hz tone) is occasionally interrupted by deviants that differ in subtle ways (e.g., 10-ms presentation of 1,500 Hz tone, 50-ms presentation of 1,000 Hz tone). However, unlike the P300, attention to the stimuli is not required, and is normally redirected from the primary auditory discrimination task by a concurrent, noninterfering, task (e.g., reading, visual target detection). Also unlike the P300, which is sensitive to both auditory and visual stimuli, MMN has been studied mostly as it relates to acoustic information processing. Because it is virtually sensitive to any feature of an acoustic stimulus, including frequency, duration, and learned features such as rhythmic patterns and timing of stimulus delivery, the MMN has been

especially useful in the evaluation of problems associated with specific aspects of auditory processing. Deviations in MMN are reported across several psychiatry disorders, including depression, alcoholism, and dementia, but have received the most attention in schizophrenia.

In a longitudinal investigation of first-hospitalized schizophrenia patients, Dean F. Salisbury and colleagues have shown that reduction in the MMN over a 1.5-year period following first hospitalization was associated with progressive reductions in Heschl gyrus gray matter volumes in schizophrenia but not in psychotic bipolar or healthy comparison subjects, who evidenced no significant change in or correlation between MMN and gray matter measures. This important finding links an experimental electrophysiological process to brain structure, and highlights the sensitivity of MMN to a neurodegenerative processes that may be uniquely involved in the etiology of schizophrenia.

Error-Related Negativity

The error-related negativity (ERN), typically studied in a choice reaction time task, occurs approximately 60 to 150 ms after an incorrect choice has been made. The ERN has been posited to reflect sensitivity to error detection, the remedial actions that are taken to compensate for the error, and the conflict generated by the process. Time–frequency analysis indicates that the ERN is mainly comprised of anterior theta activity that appears to originate from the ACC, part of a neural system involved in behavioral regulation and self-monitoring. Because it appears to involve both affective and cognitive information processing and evaluation, it is not surprising that the ERN is also sensitive to clinical features of various psychiatric disorders, including OCD and other anxiety disorders, major depression, and externalizing psychopathology related to behavioral disinhibition. The findings for these studies suggest that anxious individuals, because they generate larger ERNs, are more sensitive to making errors, while depressed individuals and those with disorders associated with behavioral disinhibition are less so. These results are generally consistent with expectations given the clinical characterization of these disorders. For instance, Jason Hall and colleagues note that a common externalizing behavioral pattern involves a failure to learn from experience, possibly suggesting a deficit in the ability to self-monitor ongoing behavior for errors, with this impairment reflected in the diminished ERN.

EEG in the Systematic Study of Psychopathology. It is now firmly established that ERP anomalies are associated with a wide variety of psychiatric conditions, including schizophrenia, mood disorder, anxiety, ADHD, and substance abuse disorders, which a voluminous research literature associated with each. For sake of illustrating the depth of science

pertaining to early risk indication, genetics, neurochemical involvement, and psychiatric comorbidities, major findings from the field of alcohol research are briefly summarized.

It has been shown repeatedly that the young offspring of alcoholics demonstrate P300 amplitude reduction despite their lack of exposure to alcohol, thus suggesting that this anomaly may represent an endophenotype tapping the underlying genetic risk for alcoholism. Further work supporting diminished P300 amplitude as an endophenotype for alcoholism includes studies showing that reductions in P300 amplitude precede the subsequent development of alcoholism and that unaffected twins from pairs discordant for alcoholism have, like their affected co-twin, reduced P300.

Research with this putative endophenotype also is helping to explain patterns of psychiatric comorbidity often observed with alcoholism. These include childhood disruptive disorders that often precede it, such as ADHD and conduct disorder, and other disorders associated with behavioral disinhibition, such as antisocial personality disorder and other SUDs. Figure 1.15–38 illustrates how reduced P300 amplitude may be a common feature among these disorders. To the extent that P300 amplitude reduction indexes underlying genetic risk, the observed pattern is consistent with the hypothesis that the covariation among these disorders reflects shared genetic influences. There is ample evidence to support this notion, including twin and family linkage studies, showing that shared genes account for most of variance in P300 amplitude among these disorders.

Because ERPs are derived from the EEG, it is possible to use time–frequency analysis to identify the constituent EEG waves, which, because they are time-locked to the stimulus presentation, compose the ERP waveform. This type of analysis in effect renders a possible determination of the amount of each EEG frequency present at each point in time elapsing from the onset of a stimulus to the resolution of the associated ERP. These transient-evoked oscillations (Fig. 1.15–19) can be partitioned into the same frequency bands identified for spontaneous EEG, but represent stimulus-elicited rhythms that are likely to serve roles that are functionally different than those associated with their appearance in resting EEG. For example, while higher resting low frequency (e.g., delta, theta) is associated with pathology in a number of clinical conditions, higher activation of low-frequency activity under task-related conditions (e.g., working memory) is often interpreted to indicate better brain function. With respect to the P300 ERP, higher amplitude (better function) is mainly dependent on activation of delta and theta frequencies, accompanied by a reduction in alpha frequency.

Transient-evoked oscillations also are beginning to provide insights regarding brain dynamics and psychopathology, especially regarding the

development of alcohol dependence. Reduced delta- and theta-evoked oscillations have been associated with risk for the development of alcoholism over and above the risk associated with reduced P300 amplitude, suggesting that the decomposition of the P300 ERP into constituent oscillatory components may improve sensitivity of this measure over the broadband signal (typically filtered at 0.5 to 12 Hz or as high as 24 Hz) characterized by the ERP. These evoked oscillations have also been linked to gene variants involved in cholinergic and GABAergic neurotransmitter systems, which have been implicated in the neurochemical pathways involved in many psychiatric conditions. This line of ERP or evoked oscillation research conducted in alcoholism exemplifies an evolving scientific effort to use experimental electrophysiological assays of human information processing as fundamentally biologic, and perhaps genetically related, indicators of risk for the development of psychiatric disorders.

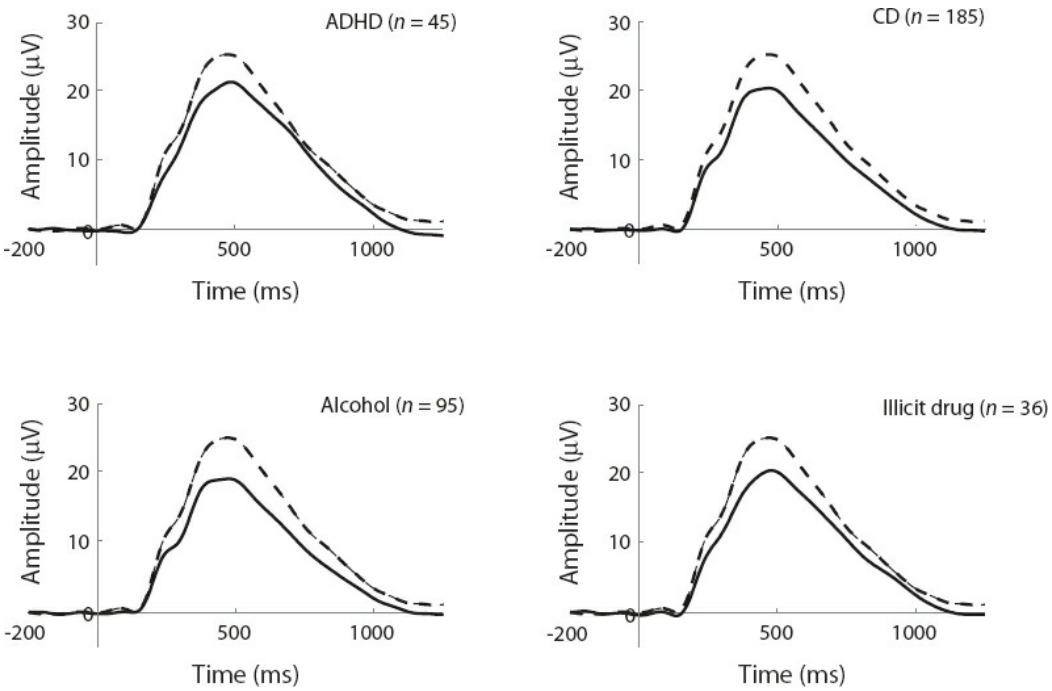

FIGURE 1.15–38. The grand average event-related potential (ERP) waveforms for male, 17-year-old participants in the Minnesota Twin Family Study with a diagnosis of attention-deficit/hyperactivity disorder (ADHD), conduct disorder (CD), alcohol use disorder, and illicit drug use disorder. The diagnostic group (*solid line*) is plotted against that of a comparison group (*dashed line, n* = 71) selected to be free of these disorders. In each case, P300 amplitude, here with an onset latency of about 300 ms and a peak latency of 500 ms, is significantly attenuated in the affected teenagers.

TRANSLATION OF RESEARCH FINDINGS TO CLINICAL APPLICATIONS

In the absence of a guiding system or a systematized and accepted approach, biological findings that carry significant promise as diagnostic tests are likely to continue to be interesting research findings without translating into clinically meaningful lab tests or be disseminated to wide

use prematurely, hindering a steady evolution of the field of applied electrophysiology. To address this translational gap, a four-step approach for developing diagnostic tests in psychiatry has been proposed by Nash Boutros and Cynthia Arfken (Table 1.15–1). These final thoughts are offered as future directions and, perhaps, a roadmap from eventually moving reliable and clinically informative electrophysiological findings from the laboratory to the clinic.

Step 1, a biological variable, is observed to be deviant from healthy controls in a particular patient population. The demonstration of test–retest reliability of the finding using blinding procedures is an essential component of this early step. Step 2 is the demonstration of potential clinical usefulness of the specific finding. The two most important objectives at this step are demonstration of difference between the target patient population and appropriate control groups (these should be groups of patients with diagnoses that commonly appear on the differential diagnostic lists of the target disorder). Estimation of the effect size of the finding could be a reasonable guide to which findings should be considered good candidates for Step 3 studies. During step 3, the performance characteristics of the test should be established. Specifically, the sensitivity, specificity, and positive and negative predictive values of the biological marker should be examined. Step 4 defines the clinical application of the test and helps standardize the technique used in large and multicenter clinical trials. Multicenter trials should guide methodology for standardization of laboratory procedures used to conduct the test as well as obtaining data regarding cost effectiveness and impact on both short-term and long-term clinical outcomes. It is strongly recommended that studies describing efforts toward developing diagnostic testing should adhere to the guidelines proposed by the Standards for Reporting of Diagnostic Accuracy (STARD) initiative.

References

Amzica F, Lopes da Silva FH. Cellular substrates of brain rhythms. In: Schomer DL, Lopes da Silva FH, eds. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA: Lippincott Williams & Wilkins, a Wolters Kluwer business; 2011.

*Arciniegas DB. Clinical electrophysiologic assessments and mild traumatic brain injury: state-of-the-science and implications for clinical practice. *Int J Psychophysiol*. 2011;82:41.

Bauer G, Bauer R. EEG, drug effects, and central nervous system poisoning. In: Schomer DL, Lopes da Silva FH, eds. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA: Lippincott Williams & Wilkins, a Wolters Kluwer business; 2011.

Boutros NN, Mirolo HA, Struve F. Normative data for the unquantified EEG: Examination of adequacy for neuropsychiatric research. *J Neuropsychiatry Clin Neurosci*. 2005;17:84.

Boutros NN, Arfken CL. A four-step approach to developing diagnostic testing in psychiatry. *Clin EEG Neurosci*. 2007;38:62.

Boutros NN, Arfken C, Galderisi S, Warrick J, Pratt G. The status of EEG abnormality as a diagnostic test for schizophrenia. *Schizophr Res*. 2008;99:225.

Chang BS, Schomer DL, Niedermeyer E. Normal EEG and sleep: adults and elderly. In: Schomer DL, Lopes da Silva FH, eds. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA: Lippincott Williams & Wilkins, a Wolters Kluwer business; 2011.

Chang BS, Schomer DL, Niedermeyer E. Epilepsy in adults and the elderly. In: Schomer DL, Lopes da Silva FH, eds. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA: Lippincott Williams & Wilkins, a Wolters Kluwer business; 2011.

Galderisi S, Sannita WG. Pharmaco-EEG: a history of progress and a missed opportunity. *Clin EEG Neurosci*. 2006;37:61.

*Grant AC, Abdel-Baki SG, Weedon J, et al. EEG interpretation reliability and interpreter confidence: a large single-center study. *Epilepsy Behav*. 2014;32:102.

Handy TC. *Event-Related Potentials: A Methods Handbook*. MIT Press; 2005.

Hartman AL, Lesser RP. Brain tumors and other space-occupying lesions. In: Schomer DL, Lopes da Silva FH, eds. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA. Lippincott Williams & Wilkins, a Wolters Kluwer business; 2011.

Herrmann CS, Fründ I, Lenz D. Human gamma-band activity: a review on cognitive and behavioral correlates and network models. *Neurosci Biobehav Rev*. 2010;34:981.

*Jobert M, Wilson FJ, Ruigt GS, Brunovsky M, Prichep LS, Drinkenburg WH. IPEG Pharmaco-EEG guidelines committee: guidelines for the recording and evaluation of pharmaco-EEG data in man: the International Pharmaco-EEG Society (IPEG). *Neuropsychobiology*. 2012;66:201.

Johannesen JK, O'Donnell BF, Shekhar A, McGrew JH, Hetrick WP. Diagnostic specificity of neurophysiological endophenotypes in schizophrenia and bipolar disorder. *Schizophr Bull*. 2013;39:1219.

Kockaya G, Wertheimer A. What are the top most costly diseases for USA? The alignment of burden of illness with prevention and screening expenditures. *Health*. 2010;2:1174.

Krystal AD. Psychiatric disorders and EEG. In: Schomer DL, Lopes da Silva FH, eds. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA: Lippincott Williams & Wilkins, a Wolters Kluwer business; 2011.

Lehmann C, Koenig T, Jelic V, Prichep L, John RE. Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). *J Neurosci Methods*. 2007;61:342.

*Leiser SC, Dunlop J, Bowlby MR, Devilbiss DM. Aligning strategies for using EEG as a surrogate biomarker: a review of preclinical and clinical research. *Biochem Pharmacol*. 2011;81:1408.

*Malone SM, Burwell SJ, Vaidyanathan U, Miller MB, McGue M, Iacono WG. Heritability and molecular-genetic basis of resting EEG activity: a genome-wide association study. *Psychophysiology*. 2014;51:1225.

Mucci A, Volpe U, Merlotti E, Bucci P, Galderisi S. Pharmaco-EEG in psychiatry. *Clin EEG*

Neurosci. 2006;37:81.

Narayanan B, Soh P, Calhoun VD, et al. Multivariate genetic determinants of EEG oscillations in schizophrenia and psychotic bipolar disorder from the BSNIP study. *Transl Psychiatry*. 2015;5:e588.

Niedermeyer E, Lopes da Silva FH, eds. *Neidermeyer's Electroencephalography: Basic Principles, Clinical Applications, And Related Fields*. Philadelphia, PA: Lippincott Williams & Wilkins, a Wolters Kluwer business; 2011.

Nunez PL, Srinivasan R. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press; 2006.

Pang T, Selvitelli M, Schomer DI, Niedermeyer E. Metabolic disorders in EEG. In: Schomer DL, Lopes da Silva FH, eds. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA: Lippincott Williams & Wilkins, a Wolters Kluwer business; 2011.

Salisbury DF, Kuroki N, Kasai K, Shenton ME, McCarley RW. Progressive and interrelated functional and structural evidence of post-onset brain reduction in schizophrenia. *Arch Gen Psychiatry*. 2007;64:521.

Schmidt LA, Segalowitz SJ. *Developmental Psychophysiology: Theory, Systems, and Methods*. Cambridge University Press; 2007.

Skosnik PD, Cortes-Briones JA. Electroencephalography (EEG) and cannabis: from event-related potentials (ERPs) to oscillations. In: Preedy VR, ed. *The Neuropathology of Drug Addictions and Substance Misuse*. Amsterdam, Netherlands: Elsevier Science; 2016.

Smith SJ. EEG in the diagnosis, classification, and management of patients with epilepsy. *J Neurol Neurosurg Psychiatry*. 2005;76:ii2.

Thatcher RW. Validity and reliability of quantitative electroencephalography. *J Neurother*. 2010;14:122.

▲ 1.16 Nuclear Magnetic Resonance Imaging and Spectroscopy: Basic Principles and Recent Findings in Neuropsychiatric Disorders

GRAEME F. MASON, PH.D., JOHN H. KRYSTAL, M.D., AND GERARD SANACORA, M.D., PH.D.

Historically the role of MRI in psychiatry has been limited to the use of structural MRI in ruling out known organic causes of emotional, cognitive, and behavioral symptoms. At present there is a growing interest in the use of MR technology in psychiatry largely owing to its noninvasive nature and its increasing ability to provide information on the structure, function, and physiology of the brain. While the potential usefulness of MR in the diagnosis of psychiatric disorders is immense, the current clinical applications remain extremely limited. However, the last decade has seen an explosion of MR research applications related to the study of the pathophysiology of psychiatric disorders and the mechanisms of action related to various forms of treatment. Imaging measures are also now included explicitly in the National Institute of Health's Research Domain

Criteria (RDoC). In the context of the RDoC, the imaging results are specifically for research purposes, attempting to characterize pathophysiological processes across a spectrum of disorders with various forms of measurements, rather than limiting imaging to its standard diagnostic purposes.

HISTORY

MRI and other MR techniques were originally called nuclear magnetic resonance (NMR). The "nuclear" part of the name derives from the basis of the technique in energy transfer into and out of the nuclei of atoms. The "magnetic" in the name refers to the use of magnetic fields in the experimental environment. The "resonance" component of the name comes from the use of particular frequencies of energy used for the measurements. Since its introduction, NMR has been used for basic investigations of chemistry and physics, but with its introduction to medical applications, the word "nuclear" has generally been dropped because patients may associate it erroneously with ionizing radiation, nuclear power, and nuclear weapons. A history and survey of many MR techniques has been reviewed recently (Ai et al., 2012).

In 1946, Edward Purcell (1912–1997) and Felix Bloch (1905–1983) independently demonstrated NMR at Harvard and Stanford Universities, and for this discovery they shared the 1952 Nobel Prize in Physics. Early NMR measurements were performed with a *swept field,* in which a narrow-banded radio frequency application was performed as the sample was exposed to a time-varying magnetic field. To enhance the efficiency of the measurement, pulse-and-Fourier-transform NMR was developed and successfully demonstrated at Varian Instruments in Palo Alto, California, in 1964 by Richard Ernst (born 1933), who proceeded to extend the concept to MRI and received the Nobel Prize in Chemistry in 1991. Raymond Damadian, a New York physician, observed in 1969 that tumors had different MR-related characteristics that provide the potential to differentiate cancerous tissue from healthy tissue, thereby demonstrating a potential clinical application of NMR. Dr. Damadian applied for the first patent for an MRI machine in 1972, in which he described the purpose of the device. In 1973, Zenuemon Abe and colleagues published a method that produced images by a slow process of acquiring signals sequentially as the position of the object was shifted through a series of positions, and in 1976 Dr. Damadian published a similar method applied to a mouse. A dramatic breakthrough for MRI occurred when Paul Lauterbur (1929–2007) of SUNY Stony Brook and Sir Peter Mansfield (born 1933) of Nottingham, England, independently developed the acquisition of MRI by means of magnetic field gradients, which will be discussed in this chapter, and that achievement led them to share the 2003 Nobel Prize in Physiology or Medicine.

BASIC PHYSICS OF NMR, MRS, AND MRI

For human use, the techniques are typically called MRI and MRS. Although MR methods are based on a quantum mechanical property called *spin*, they make use of phenomena that for most purposes are most appropriately described using classical mechanics.

Spin, the Larmor Frequency, and Net Magnetization

The primary points of this section are (1) only certain species of nuclei can be detected, (2) the frequency of detection depends on the species of nucleus observed and the strength of the magnetic field of the scanner, and (3) only a very small fraction of the nuclei can actually be detected.

Table 1.16–1.
MR-Detectible Nuclei and Some of Their Applications to Brain Research

Nucleus	γ (rad/Gauss s)	Sensitivity (% of 1H)	Application
^{1}H	26,752	100.0	Amino acid neurotransmitters and neuronal and glial compounds; imaging of water and lipids
^{31}P	10,839	.0666	High-energy phosphate levels, intracellular pH, phospholipid head groups
^{13}C	6,728	.0174	Metabolism of amino acid neurotransmitters, glucose, lipids, acetate
^{15}N	−2,713	-3.83×10^{-4}	Metabolism of ammonia, amino acids
^{17}O	−3,628	1.89×10^{-7}	Cerebral oxygen consumption, production of H_2O
^{19}F	25,182	83.5	Cerebral pharmacokinetics of fluorinated drugs like fluoxetine and fluvoxamine
^{23}Na	7,081	1.60×10^{-3}	Effects of hypoxia, pharmacologic challenges to sodium pumps
^{2}H	4,107	1.89×10^{-8}	Metabolism of fats, glucose
^{7}Li	10,398	4.62×10^{-3}	Cerebral pharmacokinetics of lithium

The fundamental basis of MRI and MRS lies in quantum mechanics, because the atomic nuclei that produce the signals are quantum particles. The nuclei carry a positive electrical charge, and observable nuclei have a quantum mechanical characteristic called *spin*. Spin of ½ is one of the basic properties of protons and neutrons. That is, a proton or a neutron can have spin of either +½ or −½, and when neutrons and protons combine to form a nucleus, their spin values combine to yield a nucleus with the net of their spins. The most abundant form of hydrogen has a nucleus that is simply a proton, so the spin of the nucleus of ^{1}H is ±½. The most common form of carbon has a nucleus that is composed of six protons and six neutrons, half with spin +½ and half with spin −½, such that in combination they yield a nucleus for ^{12}C that has no spin; thus, ^{12}C is not observable with MR techniques. However, 1.1 percent of naturally

occurring carbon is ^{13}C, and the additional neutron that resides within its nucleus confers upon ^{13}C a net spin of ½; therefore, ^{13}C can be observed with MR techniques. Table 1.16–1 lists several MR-observable nuclei that have been applied to study the human brain.

Nuclei that have nonzero spin can be detected with NMR by applying energy and observing their responses, and so investigators who develop MR technology often discuss the behavior of spins in a magnetic field and the manipulation of spins during a measurement. When a charged particle like the atomic nucleus has spin, it produces a *magnetic dipole,* which can be imagined as a small bar magnet with North and South ends.

When the nuclei with spin are placed in a strong magnetic field of magnitude B_0, they *precess* around the axis of that field (Fig. 1.16–1) with a frequency called the *Larmor frequency* or the *resonance frequency*.

$$v = \gamma B_0/2\pi \tag{1}$$

The Larmor frequency is the frequency that must be transmitted to affect the sample, and it is also the frequency that is received by the MR scanner. Equation 1 shows that the Larmor frequency depends on the strength of the static magnetic field, as well as on γ. The frequency of detection of a nucleus in a clinical magnet operating at a magnetic field strength B_0 of 3 Tesla (abbreviated 3 T) is twice the frequency of detection of the same nucleus at 1.5 T. The Larmor frequency of the most abundant isotope of hydrogen, ^{1}H, is 63 MHz at 1.5 T and 126 MHz at 3 T. These frequencies lie in the FM band of the radio, and accordingly MR scanners are usually in rooms that are electrically sealed to prevent ambient signals from contributing what looks like noise in the data. The other variable that affects the frequency is γ, the *gyromagnetic ratio,* which differs for each species of nucleus. So, because the value of γ for the common hydrogen nucleus is approximately four times as large as that for ^{13}C, the Larmor frequency of hydrogen is about four times as large as that of ^{13}C.

The reason for the large static magnetic field is to create a condition called *net magnetization*. In the absence of an external magnetic field, the spins of the nuclei in the magnetic field are oriented randomly in space. To create the potential for a signal, some nonrandomness must be introduced. It is this small fraction of the total nuclei that MR scanners can detect, and the greater the difference in the number of particles in the two energy states, the larger will be the signal that can be detected from the sample. The net magnetization expressed as a fraction the total spins is given by the *Boltzmann distribution*

$$N_{\text{high}}/N_{\text{low}} = 1 - vhB_0/2\pi kT \tag{2}$$

where k is Boltzmann constant and T is the absolute temperature. Most

people have little occasion to use Eq. 2, but it illustrates several important features of NMR measurements. First, a larger net magnetization can be obtained if a nucleus with a larger gyromagnetic ratio is used. Hydrogen has the largest ratio and has the sensitivity to be used for routine anatomic and other imaging. Sodium has a low ratio and is less sensitive and more difficult to use for imaging. Second, larger B_0 yields more net magnetization, so a stronger magnet can yield more signal. This is why Earth's field MR has not been used for studies of the brain: a 3-T magnet is 67,000 times stronger than the Earth's average magnetic field, and so the signal available at Earth's field is thousands of times lower. Third, lower temperature adds signal, but for measurements in humans, T is usually fixed at about 37°C. In practical terms, measurements on 3 T machines have twice the signal available from a 1.5 T machine. Finally, if one plugs in typical temperatures and magnetic field strengths used for human clinical scans, the equation shows that for hydrogen the net magnetization consists of only 0.0004 to 0.0008 percent of the nuclei in the sample.

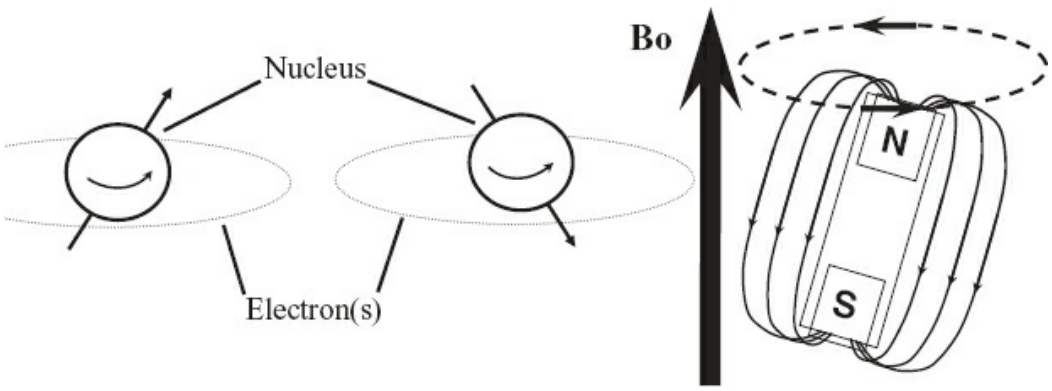

FIGURE 1.16–1. Individually, nuclei are particles that have the quantum mechanical property *spin.* By virtue of their charge and the spin, they behave as small magnets that can be aligned either with or against the large magnetic field of the MR scanner (B_0), rotating at the *Larmor frequency.*

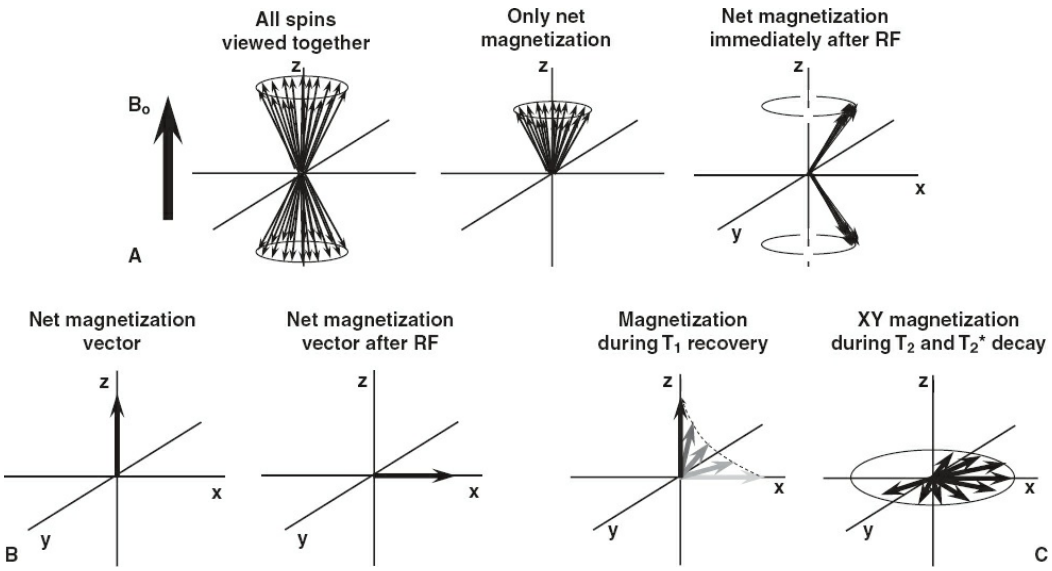

FIGURE 1.16–2. Magnetization in an MR experiment shown with perspectives of quantum and classical mechanics. **A:** Quantum mechanical: In the presence of an external magnetic field, the individual spins are aligned with or against the field, and a slight majority, on the order of just 0.0004 percent, is aligned with the field. The very small size of this ratio is the reason that MR is considered to be insensitive compared to other forms of imaging. When a radiofrequency burst is applied at the Larmor frequency of the spins, to apply what is called a 90-degree pulse, the net distribution is randomized so that the spins are evenly distributed with and against the field. At the same time that the spins are randomized, they acquire *phase coherence,* coming together in one orientation with respect to the *XY* plane. **B:** Classic mechanical: In the presence of the external magnetic field, there exists a magnetization vector whose size is proportional to the sum of the small majority of spins that have net alignment with the field. When a 90-degree pulse is applied, the magnetization vector rotates to the *XY* plane. **C:** After the excitation, the magnetization vector gradually returns to the *Z* axis according to T_1 relaxation, as shown by the progressively darkening arrows that represent the continual movement of the vector toward *Z.* At the same time, phase coherence is lost due to interactions among spins, according to T_2 relaxation, and inhomogeneities of magnetic field in the sample cause additional losses of phase coherence, according to T_2* decay.

Detection of the Net Magnetization

The techniques of MRS and MRI in vivo detect the quantities of net magnetization in various locations and elucidate properties of the cellular environment. For MRS, there is the particular purpose of discovering in what chemicals the magnetization resides, and for MRI, the specific purpose is to measure the spatial distribution of the magnetization. Having the nuclei precess at the Larmor frequency and follow a net distribution oriented along the scanner's magnetic field set the stage for MR detection, but one more component is necessary to obtain a signal.

Within seconds of placing a subject in the scanner, the nuclei in the brain are in a *fully relaxed* state. In the jargon of MR, a fully relaxed state is one in which the nuclei have their slight orientation with the magnetic field, according to their Boltzmann distribution. In this condition, there is no energy emitted from the sample, the phases of the nuclei in their precession are random, and the net magnetization cannot be detected. At this point, no signal is detected in the sample because of the lack of emission of energy, and even if energy were emitted by individual nuclei, their phases of precession are distributed randomly around the *Z* axis, so any energy emitted by one nucleus would be canceled by energy emitted by one that is opposite in phase. To detect the net magnetization, a magnetic field oscillating at the Larmor frequency (called radio frequency, or *RF*) is first applied, generally lasting a few hundred microseconds to a few milliseconds, and the nuclei absorb energy from the transmission, which causes the net magnetization to rotate out of alignment with the large, static magnetic field of the scanner (Fig. 1.16–2A). At the same time, the nuclei acquire *phase coherence*. That is, they precess together instead of having a random distribution about the *Z* axis. As the net magnetization returns to its alignment with the magnetic field, the nuclei emit energy, also at the Larmor frequency. The energy is detected with a *probe,* sometimes also called a *coil,* which is a placed around or against the part of the body that is being studied. In many cases, the same coil is used for transmission and detection.

As the net magnetization is rotated away from the magnetic field orientation, any magnetization that is in the *XY* plane is called *transverse magnetization* (Fig. 1.16–2B) In the terminology of classical mechanics, MR can only detect the component of the magnetization vector that lies in the *XY* plane. If less energy is used, then a lower flip angle is achieved, and less signal can be detected. It is possible to rotate the magnetization through any angle, including 180 degrees, which is a special rotation called an *inversion,* and it is used frequently for MRI and MRS.

Quantum and Classical Mechanics

Many explanations of MRI and MRS are based principally on quantum mechanics. The explanations are often daunting and are generally unnecessary. A phenomenon is called "quantum mechanical" if classical mechanics are unable to explain the observed behavior. For example, a rotating bicycle wheel can be described with great effort using quantum mechanical calculations, but the wheel's behavior is explained with great precision and considerably greater ease with classical mechanics. Therefore, a rotating bicycle wheel is considered a classical phenomenon, not a quantum mechanical phenomenon. In the case of MR, observations of a single spin can be explained only with quantum mechanics, but when many spins are observed, as is the case for an MRI study of the brain, the nuclei behave as a set of spins that have a continuum of orientations, though slightly biased along the magnetic field. The behavior is explained completely in the realm of classical mechanics. This point is described in detail by Hanson (1998).

T_1, T_2, and T_2* Relaxation

Once the magnetization vector has been rotated to the *XY* plane, phenomena called *relaxation* eliminate the signal and return the vector to the *Z* axis through exponential processes called T_1 and T_2 relaxation, or *longitudinal* and *transverse* relaxation, respectively. Longitudinal relaxation returns the magnetization to the *Z* axis at a rate governed by the exponential time constant T_1 (Fig. 1.16–2B). As will be explained later, this relationship can be used to create contrast among materials with different values of T_1, such as gray matter and white matter.

Repeating scans too rapidly with too large an excitation angle leads to a reduction in signal amplitude through a process called *saturation*. Transverse magnetization can only be generated from *Z* magnetization, which is created during the T_1 relaxation process. Therefore, if too little time is allowed for recovery of *Z* magnetization, then the signal will be reduced on repeated scans. For full relaxation after a 90 degrees pulse, typically a delay of several seconds is required. Less saturation occurs when a lower angle of excitation is used, but that also results in the generation of less transverse magnetization. In practice, the most efficient acquisitions incorporate a compromise, in which an excitation angle of something less than 90 degrees is used and an interscan delay of less than three T_1's is allowed between scans.

In vivo, T_1 relaxation depends on the static magnetic field strength, the rate of tumbling of the molecules in which the spins reside, and the distance between the relaxing nucleus and other charged spins in the sample. As a consequence, as magnetic field strength rises, the value of T_1 increases. As will be shown later, the dependence of T_1 on field strength can

be exploited to make more effective experiments but can also place some limitations on measurements at different magnetic field strengths.

For measurements in vivo, long before the magnetization vector has returned to the Z axis, the signal decays to nothing because of T_2 relaxation (Fig. 1.16–2B). T_2 relaxation destroys the phase coherence of the component of the magnetization in the XY plane, such that the MR signal decays in an exponential process. The signal gradually dies away, yielding a *free induction decay,* or FID. T_2 relaxation is primarily due to interactions of the nucleus with other spins in the sample, and in the brain in vivo its value is largely independent of magnetic field strength. As phase coherence is lost, the signal from each spin in the sample becomes increasingly canceled by another spin in the sample, so that gradually the average signal coming from the sample approaches zero. The value of T_2 depends largely on the mobility of the molecule in which the nucleus resides: The greater the restrictions on motion, the shorter will be the T_2. For example, white matter has abundant, rigid structure that prevents water from diffusing rapidly and rotating freely, and white matter has a shorter T_1 than CSF, in which the water is free to rotate. The differences between white matter and gray matter provide an effective basis for high contrast for diseases in which fluid must be emphasized in MRI. The T_2 of bone is very short and is therefore not typically detected in MRI or MRS.

A parameter related to T_2 is T_2^*, which results from inhomogeneity of the magnetic field. Such inhomogeneity can derive from macroscopic causes, such as an interface between tissue and bone or air. The sinuses, ear canals, and temporal bones cause large inhomogeneities that are often troublesome for MRI and MRS studies. Inhomogeneities also result from microscopic influences, such as the presence of deoxyhemoglobin. T_2^* is the time constant for an exponential loss of phase coherence, but the process is illustrated as a dephasing of magnetization in different parts of the sample (Fig. 1.16–2B). The sample volume could be an image pixel or region selected for an MRS measurement. If one corner of an image pixel experiences a slightly higher magnetic field strength, then the Larmor frequency in that corner will be a little higher. Perhaps the center of the image pixel has a lower frequency, so the Larmor frequency in the middle of the pixel will be lower. As time progresses, the magnetization in different parts of the sample will eventually randomize their phase, such that any signal coming from one part of the sample will be canceled by signals with the opposite phase from another part of the sample. In certain kinds of MRI and MRS studies, such effects can dramatically reduce a signal, causing regional darkening of images in MRI and blurring and loss of signal in MRS.

Influences on Frequency and Identification of Molecules in MRS Measurements

The main points of this section are that the major factors that influence a resonance frequency are (1) magnetic field strength and (2) gyromagnetic ratio, which affect primarily what transmitter frequency will be used. Smaller influences are the electron cloud conformation and the interactions with other nuclei in the molecule, which together imbue MRS with its ability to distinguish neurochemicals from one another.

The primary actor in the differentiation of chemicals is *shielding,* a phenomenon in which the electron cloud in a molecule partially blocks the static magnetic field that penetrates to the nucleus. Consider the cases of glutamate and glutamine (Fig. 1.16–3). The fourth carbon in the chain (C4) of glutamate resides next to a carboxyl group whose oxygen atoms are highly electrophilic. The nearby oxygen atoms of the carboxyl group draw the electrons at that end of the molecule to themselves, so that the C4 carbon nucleus of glutamate is more exposed to the large magnetic field and therefore has a larger Larmor frequency. Stated differently, the carboxyl group reduces the shielding of the carbon nucleus, slightly modifying the value of B_0, and thereby causing the carbon to resonate at a slightly higher frequency. Now consider glutamine, which has only one oxygen near the C4 carbon position, in addition to the nonelectrophilic amide group, and so the electrons more fully shield the glutamine C4 carbon nucleus. The C4 of glutamine resonates at a lower frequency than the C4 carbon of glutamate.

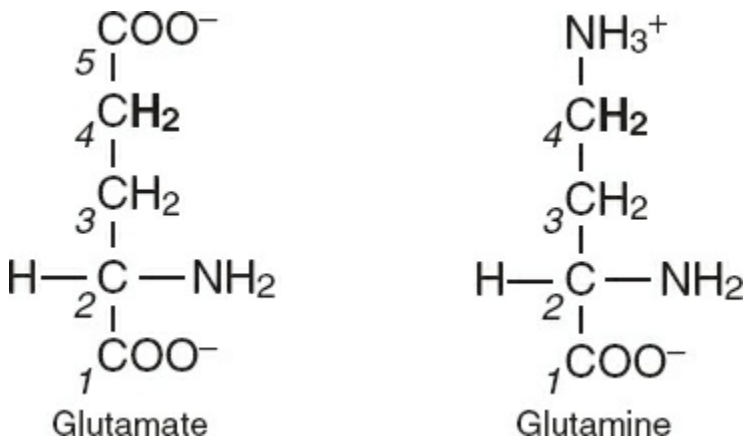

FIGURE 1.16–3. Neurochemicals differ slightly in the Larmor frequencies of their atoms, according to chemical identity and location within the molecule. The frequencies differ according to their magnetic environment, and an electron cloud provides shielding for the nuclei. In the case of the $-CH_2-$ of glutamate whose carbon is numbered 4, adjacent to the carboxyl group, the electrophilic oxygen atoms on the carboxyl group draw the electron cloud toward themselves, leaving the hydrogen nuclei exposed to the large static magnetic field. The nuclei of the corresponding hydrogen atoms in glutamine do not have the adjacent electrophilic oxygen atoms to unshield them and therefore experience a slightly lower magnetic field strength. Therefore, the resonance frequency of glutamate at 1.5 T is about 7 Hz higher than that of glutamine.

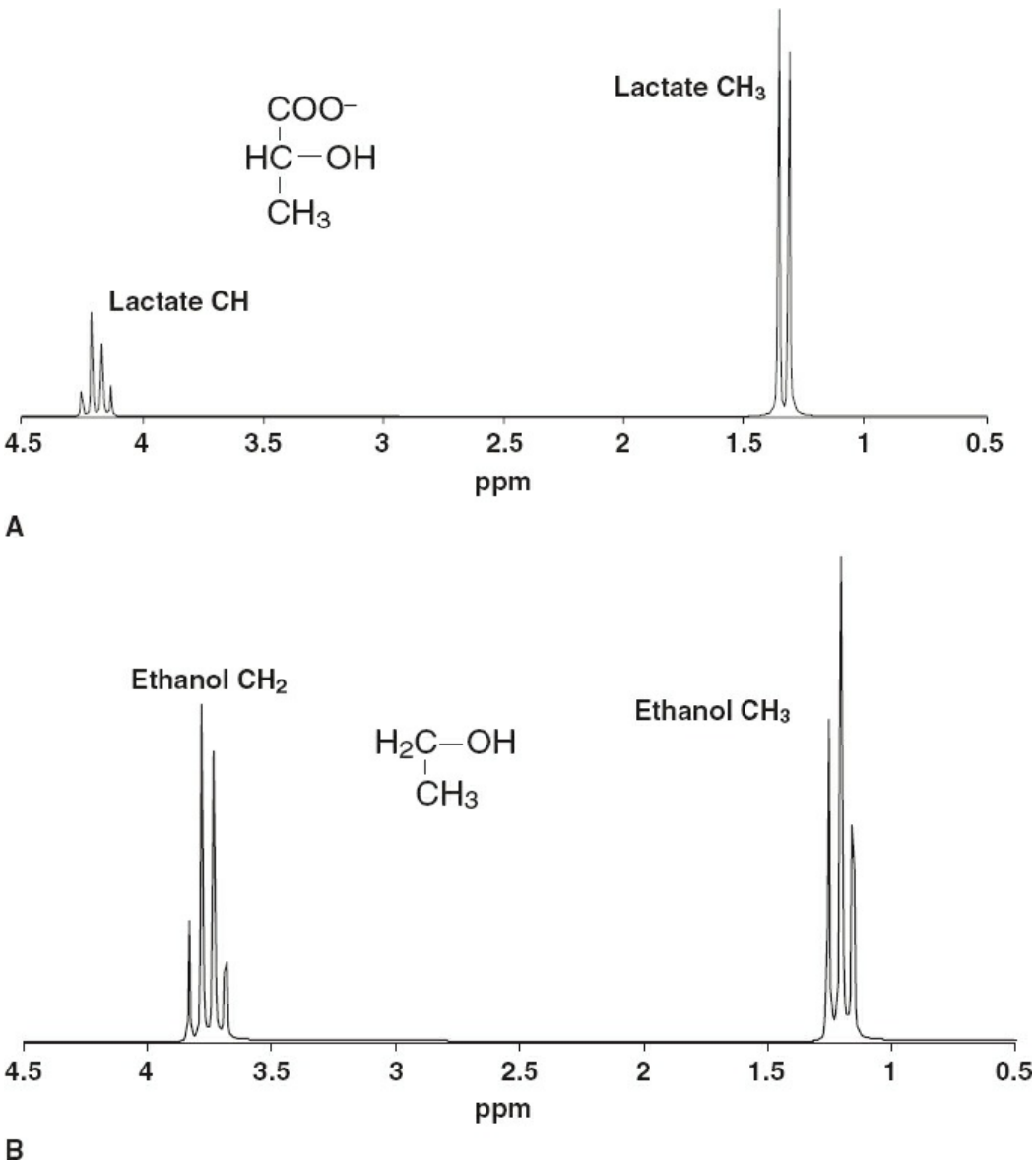

FIGURE 1.16–4. Effects of chemical shift and *J*-coupling. Chemicals in MRS studies are identified primarily on the basis of their chemical shifts, but significant information also arises from the *J*-coupling interaction of spins. **A:** Lactate has resonances at 4.15 and 1.33 ppm from TSP. The methyl protons are *equivalent* and are all centered at 1.33 ppm. However, they interact with the adjacent single proton at 4.15 ppm that splits their resonance into a doublet centered at 1.33 ppm. The single proton at 4.15 ppm interacts with all three methyl protons and is split into a quartet. **B:** Ethanol has a methyl group whose resonance is centered at 1.2 ppm and interacts with an ethylene group whose resonance is centered at 3.7 ppm. The two protons split the methyl resonance into a triplet, while the triple divides the ethylene resonance into a quartet.

Glutamate and glutamine are also separated in the ^{1}H spectrum. At a magnetic field strength of 1.5 T, the resonances of the two hydrogen nuclei attached to the C4 of glutamate and glutamine are centered 7.7 Hz apart for glutamate relative to glutamine. At 3 T, the same hydrogen nuclei resonate 15.4 Hz apart. At 7 T, they resonate 36 Hz apart, which separates the signals more effectively at the higher magnetic field strength. Because MRS studies are performed across a wide range of magnetic field strengths,

some common set of units is used to ease identification of chemicals based on their resonance frequency. The frequency of resonance for nuclei at a particular position in a given molecule is called the *chemical shift,* and the units are parts per million, or *ppm*. The notation of ppm is obtained by dividing the frequency difference in hertz by the Larmor frequency in megahertz (i.e., 10^6 Hz). The result has units of hertz per million hertz, or ppm. In these units, those protons of glutamate and glutamine are said to be centered .11 ppm apart, no matter what the static magnetic field strength.

Chemical shifts must be seen as relative differences in frequency location. By convention, for ^{1}H and ^{13}C MRS, the field has agreed that the signal from trimethylsilylpropionate (TSP) be designated as the reference compound with a shift of 0 ppm. All other compounds by convention are described by their frequencies relative to that of TSP. For example, lactate has protons that resonate 1.33 and 4.15 ppm away from TSP, respectively (Fig. 1.16–4). For reasons based on the original techniques used to obtain MRS measurements, called *swept field NMR,* lower chemical shifts are termed *upfield,* and the higher chemical shifts are called *downfield.*

The final influence on the frequency of a nucleus is *J-coupling,* which is the influence of neighboring nuclei with spin upon one another. The state of the spin of one nucleus exerts a very slight effect on the *transition energy* of the neighboring spin. To describe J-coupling, a quantum mechanical explanation is necessary. The transition energy is the amount of energy that a spin can absorb and change its energy state, and it is a defined, fixed amount and must be delivered at precisely the Larmor frequency. If one spin is coupled to another, usually through a covalent chemical bond, one spin will change the other's transition energy to a slightly higher or a slightly lower value. Because the transition energy is ultimately what defines the resonant frequency of a nucleus, half the spins have a slightly higher frequency and half a lower frequency. A nucleus that is affected by one nearby spin, such as the $-CH_3$ group at 1.33 ppm in lactic acid, has a *doublet* or two resonances instead of one, each representing half of the total resonance for $-CH_3$, as is the case for the methyl group of lactate, in which three identical hydrogen nuclei are affected identically by a single adjacent hydrogen (Fig. 1.16–4A). If nuclei are adjacent to two spins, then their resonance is divided into a *triplet,* as demonstrated by the spectrum of ethanol (Fig. 1.16–4B). Nuclei adjacent to three spins, such as seen in the hydrogen adjacent to the methyl group of lactic acid, yield a quartet that arises from the effect of the three methyl protons of lactic acid on the single ^{1}H at 4.11 ppm (Fig. 1.16–4A) or the $-CH_2$ protons of ethanol (Fig. 1.16–4B). The spectral distance between each resonance in the multipeak structure is J Hz and is typically reported in hertz, not ppm, because the splitting due to J-coupling is independent of magnetic field

strength. For the $^1H–^1H$ splitting of Figure 1.16–4, the value of J usually lies between 3 and 15 Hz.

SYSTEM HARDWARE

Before moving from the basic physics of MR to methods of MRI and MRS, it is helpful to examine the MR machine itself. MR machines are composed of three or four main components: A magnet, a *transmit path,* and a *receive path,* and almost all MR scanners also have *gradient coils.* The magnet is the source of B_0, which is what creates the net magnetization. The transmit path generates the transverse magnetization. The receive path detects the signal from the sample and carries it back to the computer for storage.

The magnet (Fig. 1.16–5) consists of hundreds of miles of superconducting wire wound in precise formations. When a magnet is installed, a strong vacuum is applied to remove all ambient air and water from around the wire. Then the evacuated container is filled with liquid helium to maintain its superconductivity at a temperature of 4 K, or –453°F. The helium bath is insulated by a vacuum chamber to reduce boil-off. Liquid helium is expensive, and a recent worldwide shortage is expected to continue for several years, so in many magnets, the helium structure lies within a bath of liquid nitrogen or a refrigeration unit. The helium in the magnet must be maintained at high purity, because humidity or air that enters the helium chamber (called a *dewar*) will solidify and may create a blockage that can result in an explosion. Another sort of failure is a *quench,* which is a sudden loss of charge in the electrical windings of the magnet, simultaneously with the almost instantaneous expulsion of helium from the magnet. Such events can be caused by a loss of vacuum around the helium container, a failure to maintain the helium at an adequate level, or the introduction of a ferromagnetic object in the vicinity of the magnet.

The magnet is constructed with additional circuits called *shims,* which are used to fine-tune the homogeneity of the magnetic field. The quality of MRI and MRS improves dramatically with the homogeneity of B_0. The magnets are designed and built to achieve a homogeneous magnetic field within a sphere in the center of the magnet, but when a magnet is placed in the site where it will operate, the homogeneity is disrupted. Initial adjustments to B_0 homogeneity are made using small pieces of metal, called *shims,* placed strategically at different places in the bore of the magnet in a painstaking process, and then engineers adjust the magnetic field homogeneity further with the shim circuits. Likewise, when a person is placed in the scanner, the homogeneity is altered, and it is necessary to adjust the shim circuits to compensate. For many cases, it is only necessary to apply three linear adjustments to the magnetic field, but for some applications, including MRS, second- or third-order changes must often be

applied.

The transmit path begins with a frequency source whose output is modified to obtain a desired amplitude, phase, and shape. The modulated signal is sent to a radio frequency amplifier to boost the transmission, often into the range of kilovolts. The transmission is directed to the coil that is around the head, where it generates an oscillating magnetic field, which by convention is called B_1.

Part of the transmit path is the radio frequency coil. A typical coil for brain imaging and spectroscopy is shown in Figure 1.16–6. The coil surrounds the brain so that relatively uniform coverage is achieved. Another type of detector, a *surface coil,* is used for cases in which localized coverage with high sensitivity is required. All coils must be tuned to the Larmor frequency, and to do that, adjustable capacitors form part of the circuit. The coil not only transmits to the brain but also receives the signal from the brain and therefore forms the beginning of the receive path. The signal in the receive coil has an amplitude of microvolts at the Larmor frequency. If it traveled unaltered through many feet of wires to the computers for processing and storage, then the noise that accumulates along its path would overwhelm the signal. Therefore, as close as possible to the coil, magnet-compatible *preamplifiers* are used to boost the received signal to a level of millivolts. The signal is modified to shift the frequency into the auditory range, covering kilohertz, and stored on a computer. Typical causes of coil failure include a preamplifier damaged by leakage of radio frequency transmission into the receive path or damage to capacitors from power-induced cracking or sparking. Such failures become more likely in a head coil by transmitting high power levels into coils, especially when a sample whose size and conductivity differ significantly from a human head.

FIGURE 1.16–5. A whole-body MR scanner. Magnets for MRI are usually cylindrical, with a hole through the middle. The hole accommodates a set of cylindrical gradient coils that are used for shimming and experiments that require spatial information, such as imaging. The gradient coil cylinder has a bore large enough to accommodate a human torso or, in some new models, just a person's head. A patient bed slides in and out of the scanner. On top of the magnet are vents to allow for boil off of helium and, in many cases, nitrogen gas. There is also an entry point to allow charging of the superconducting windings when the magnet is installed and to allow controlled shut down if the magnet needs to be de-energized after some years of operation.

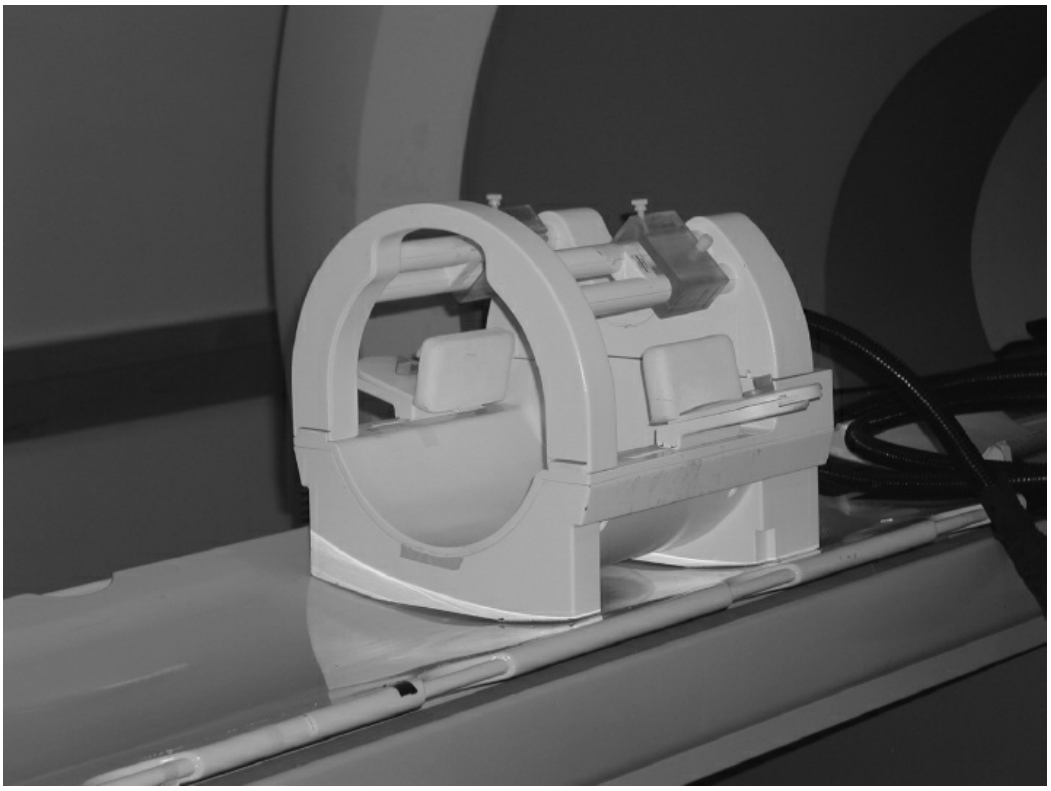

FIGURE 1.16–6. A typical coil used for MRI and MRS of the brain. Built for whole-brain

coverage, this coil surrounds the head. The ideal would be a long, empty conducting tube, but to accommodate a head, it must be short, and to allow for patient comfort and construction constraints, it must have discrete conducting elements. Four of the elements are housed within the bars that can be seen along the upper part of the coil. The other elements are housed in the bottom section of the coil, beneath where the head rests during a scan.

Gradient coils form an almost universal fourth component of hardware in an MR scanner. In the absence of gradient coils, MR signals are obtained with little or no information about what part of the sample that they represent. For example, if a standard brain imaging protocol is run without gradient coils, then the result would be an MR spectrum of a large water signal and small fat signal. The *X, Y,* and *Z* gradient coils provide the positional information. When current is directed through the *X* gradient coil, a spatially linear ramp is added to B_0, such that $B_{0\text{effective}} = B_0 + G_x x$, where G_x is the gradient strength in units of millitesla/centimeter, and x is the distance from the magnet's center. Similar relationships exist for the *Y* and *Z* gradients and the y and z directions in the magnet. When power is applied to the gradients, the Larmor frequency depends linearly on the position within the magnet, and by this encoding images can be obtained. Patients commonly ask the source of the loud noises of an MR scanner, and the answer is the gradient coils. When the current is switched on and off rapidly, it generates a magnetic field that opposes the primary field of the magnet and generates large forces that are manifest as loud noises. Patients may also sense vibrations inside the scanner, from the same cause.

FOURIER TRANSFORM

Having examined the basic physics and the MR machine, it is helpful to turn to one of the most basic analytical tools that are used for MR studies. The purpose of this section is to explain and illustrate the relationships between the observed signals and the processed images and spectra of MRI and MRS. The signals are observed over periods of milliseconds, and those time-dependent signals must be transformed into spatial pictures in the case of MRI and spectra in the case of MRS.

MRS and MRI both make use of a mathematical operation called a *Fourier transform,* which is used in a many fields to switch between time and temporal frequency (e.g., seconds and hertz) or between space and *k-space* (centimeters and reciprocal centimeters). Every time-dependent signal can be represented as a set of frequencies, measured in hertz, and every object can be expressed as a set of signals in *k*-space. The idea behind Fourier transforms is that any shape can be created by adding the right frequencies and amplitudes of sinusoidal signals together. Figure 1.16–7 shows the decomposition and recreation of a profile of the Taj Mahal in India. The Fourier transformation at the upper right is plotted as a histogram of weightings associated with the range of frequencies that make

up the profile, with low frequencies in the center of the graph and higher frequencies to either side. The panels at the center left show that the contributions of the lowest frequencies to the profile have to do with large features. From the low frequencies, one can see that the profile has a large object in the middle and something large on either side, but there is little detail. The next panels to the right show the effects of adding progressively higher frequencies, which gradually produces the finer features. That is because the higher-frequency components of the Fourier transform contain the details of a signal. When all of the frequencies are included, the original image is obtained. The Taj Mahal is built from sine waves.

Aside from its appearance in dreams, the Taj Mahal may have little to do with psychiatric brain imaging, but it illustrates fundamentals of the Fourier transform. Signals in MRS are acquired as a function of time, just as one can record sounds as time-dependent vibrations (Fig. 1.16–8A). The Fourier transform of the sounds reveals the frequencies and volumes of sound that comprise a note, just as the Fourier transform of an MRS signal reveals the frequencies and amplitudes of signals that uniquely identify neurochemicals that make up the FID (Fig. 1.16–8B). The MR signals in brain imaging are acquired in the *k*-space domain and undergo Fourier transformation to generate the standard anatomic pictures. Low-frequency data are those that reside near the center of *k*-space, and just as for the picture of the Taj Mahal, the data that reside near the center of *k*-space lack details but indicate primarily how large the head is and that the shape is roughly circular. The information about sulci and other fine details are found farther from the center of *k*-space (Fig. 1.16–9). Schemes exist to acquire only the center or the outer edges of *k*-space to save time in cases that do not require all of the information about the image, but the great majority of brain imaging is done with standard coverage of *k*-space.

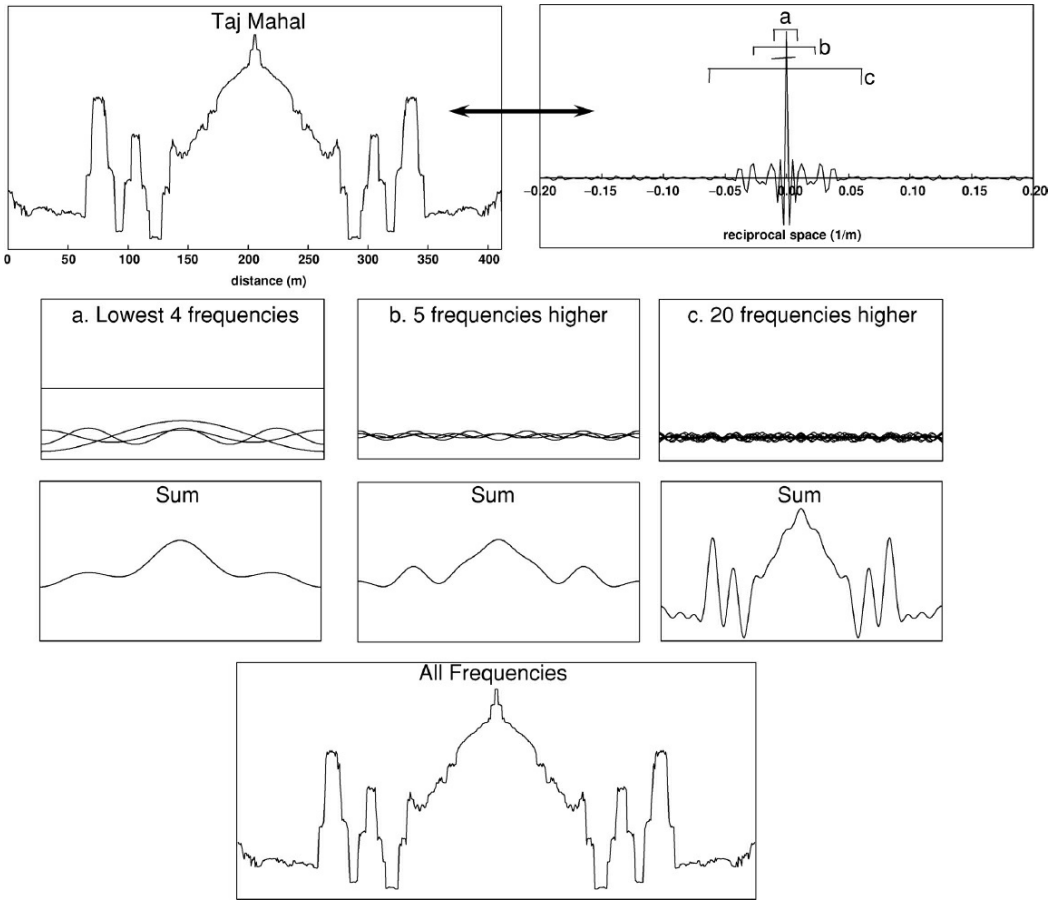

FIGURE 1.16–7. The Fourier transform is based on the premise that the shape of any object can be created with the right combination of sinusoidal waves. In this example, a profile of the Taj Mahal is shown at the upper left, and its Fourier transform is shown at the right. If the center group of frequencies, labeled a, are plotted and summed (**two center left panels**), then a general shape of the profile is achieved. If the next group of frequencies, labeled b, are plotted and added (**two middle panels**), then the shape of the center dome begins to take shape. The next band of frequencies, labeled c, can be included (**right center panels**) to show the presence of the pairs of towers on each side of the central dome. When all frequencies are included, the original picture is seen (**bottom**). Two principles are shown by this example: An object can be expressed as a series of sinusoids, and the details of an object lie primarily with the frequencies that lie farthest from zero.

MRI METHODOLOGY

MRI methodology is about seeing boundaries and differences in tissue composition. That requires the generation of *image contrast*. The word "contrast" is often used clinically to mean an injectable *contrast agent,* but the general meaning of contrast in MRI is anything that differentiates the appearance of one material from another. To that end, contrast in brain imaging is obtained by exploitation of differences in T_1, T_2, blood oxygenation, diffusion coefficient, and concentration of water or fat. The most basic neuropsychiatric application of MRI is the mapping of brain anatomy, and this is generally done on the basis of the natural contrast from T_1, T_2, and T_2* of brain water. On the side of clinical applications, the purpose is often to check for particular causes of symptoms, such as tumors

or multiple sclerosis. All image sequences consist of one or more region-specific radio frequency excitations followed by applications of gradients to scan *k*-space. The resulting *k*-space distributions undergo Fourier transformation to create images. By convention, MRI results are displayed such that the bottom of sagittal or coronal images points toward the feet and the right side of axial or coronal images points toward the patient's left.

The most basic MRI sequences are called *gradient-echo* and *spin-echo*. The gradient-echo sequence is simply an excitation pulse in conjunction with gradients to localize the sources of the signals (Fig. 1.16–10). For each slice, the measurement consists of a slice-selective excitation pulse applied, while a slice-selection gradient is applied. The longer the pulse and the stronger the gradient, the thinner the slice will be; scanners equipped with stronger gradients are generally able to provide thinner slices. A *phase-encoding gradient* is then applied perpendicular to the slice selection, in one of the in-plane dimensions, to shift the acquisition to a fixed distance in *k*-space; larger phase-encoding gradients shift the acquisition farther in *k*-space. After the phase-encoding gradient, a preparatory gradient is applied in the second in-plane dimension to shift the acquisition in the second dimension of *k*-space. Finally, a *read gradient* is applied in that second dimension of *k*-space. The process is repeated, using a slightly different phase-encoding gradient strength each time to provide appropriate coverage of two dimensions of *k*-space, allowing the generation of two-dimensional image.

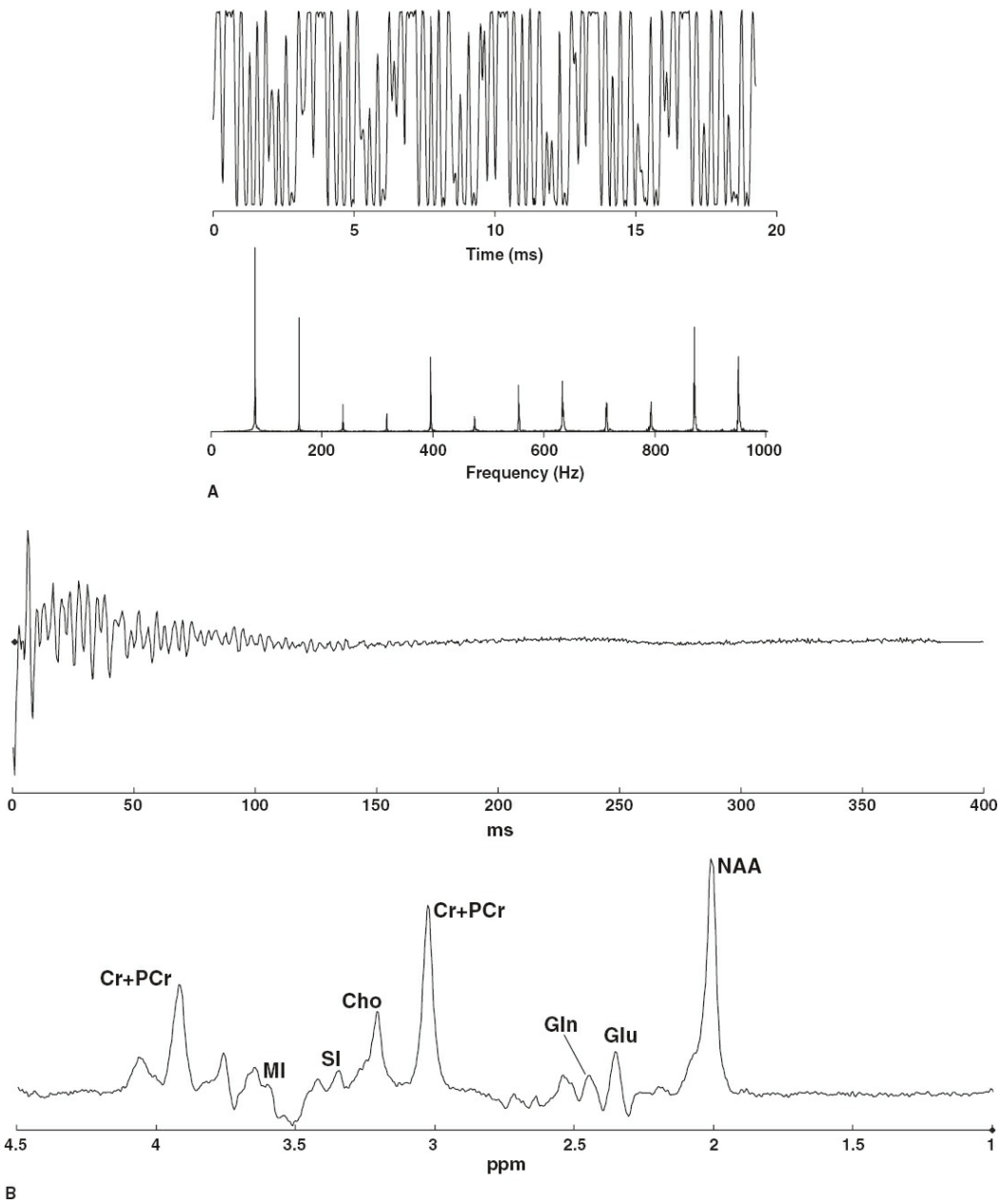

FIGURE 1.16–8. The Fourier transform allows time-dependent signals to be expressed in terms of frequency, and in this respect MRI is indistinguishable from sounds. **A:** The upper graph shows the sound of a note played on a saxophone and cymbal by the author's daughter and son. Fourier transformation reveals some of the frequencies that are generated from the instruments. The richness of the sound of a saxophone derives largely from the combination of frequencies and various amplitudes of tones that it emits, and the cymbal's characteristic sound derives from the multitude of frequencies in addition to its primary pitch at 540 Hz. **B:** The MRS signal, viewed relative to the Larmor frequency, is indistinguishable from sound; the MRS signal also appears as a time-dependent oscillation. Some early MR scanners were connected to audio speakers to allow the operators to make quick evaluations of signal quality. In this case, the Fourier transform converts a complicated time-domain measurement into a spectrum with easily identifiable features that are typical for a sample of healthy brain tissue. The volume contains *N*-acetylaspartate (NAA), creatine + phosphocreatine (Cr + PCr), myo-inositol (MI), choline, scyllo-inositol (SI), glutamate (Glu), and glutamine (Gln).

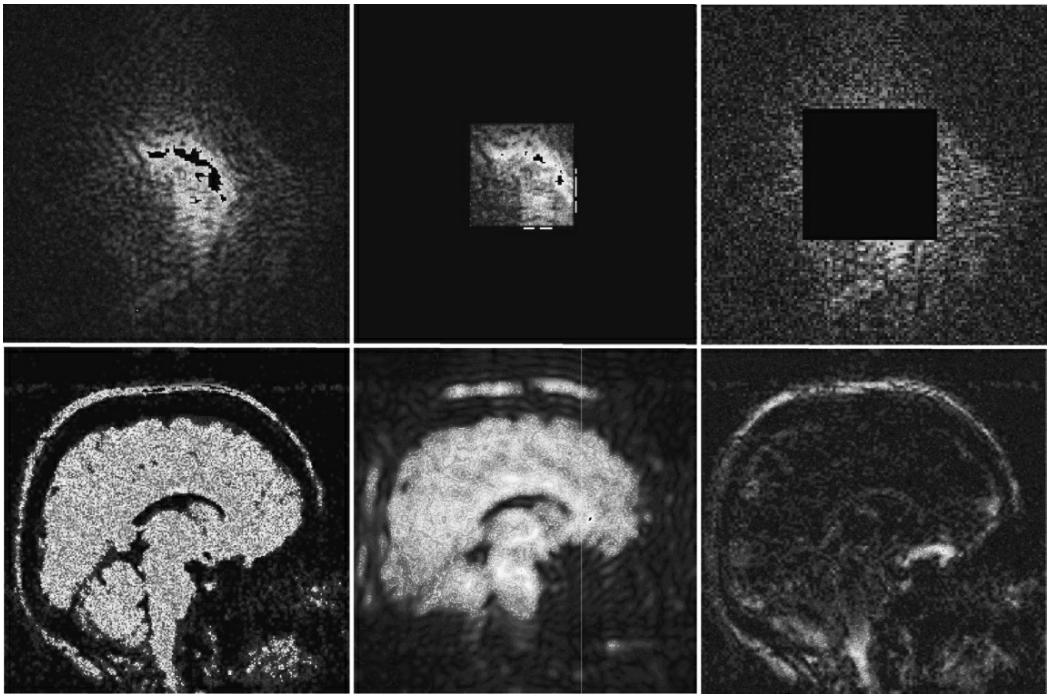

FIGURE 1.16–9. *K*-space data (**top**) paired with associated images (**bottom**). Typical *k*-space coverage is uniform and yields a standard brain image (**left pair**). If only the center of *k*-space is acquired (**center pair**), then the image loses much of its detail, which is analogous to the example of the reconstruction of the profile of the Taj Mahal in the earlier example. If only the outer regions of *k*-space are acquired (**right pair**), then only the fine details remain.

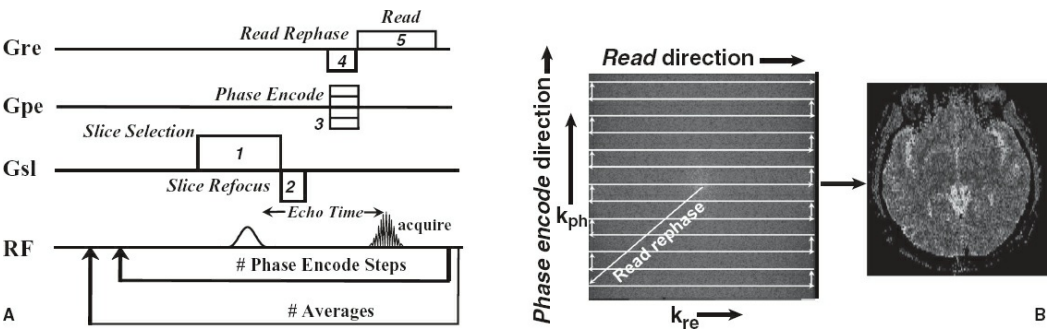

FIGURE 1.16–10. Coverage of *k*-space for brain imaging. **A:** The gradient-echo sequence is a pulse-acquire measurement that includes gradients to localize the signal. (1) A slice-selection gradient is applied with a frequency-selective pulse to excite the magnetization in a slice. (2) The slice selection gradient disperses the magnetization, so a slice refocus gradient is applied to refocus it before the acquisition. (3) Phase encode gradients move the magnetization in the phase encode direction. (4) A read rephase gradient opposite to the read gradient is applied to move to one edge of *k*-space. (5) A read gradient moves the signal through *k*-space in the read direction. **B:** The *k*-space data are shown at left, with indications of the path that was covered by the read and phase-encoding gradients. Transformation of the *k*-space data results in the image shown at the right.

T_1 Weighting, Injected Contrast Agents, and T_2 Weighting

T_1-based contrast exploits the different rates of realignment of spins with the static magnetic field. T_1 weighting is exploited in different ways,

including injections of exogenous agents, but a common noninvasive approach is *inversion recovery* (IR), with timing set to null, or blacken, regions of the image with the long T_1 of CSF. IR imaging works on the principle that the detectible *XY* magnetization can only be created from *Z* magnetization, which means that elimination of *Z* magnetization in the CSF will eliminate the source of the signal in areas that contain CSF. When a 180-degree pulse is applied to the magnetization, the net magnetization is inverted, and it immediately begins to return to the alignment with the magnetic field. The T_1 of white matter is the smallest, so the magnetization in white matter recovers first. The T_1 of CSF is longest, so its magnetization recovers most slowly, and if the images are acquired near the point at which its *Z* magnetization reaches 0, the CSF will appear black in the images. The gray matter has an intermediate T_1, so its value will be recovered enough to detect but not as much as that of white matter: Gray matter will appear significantly darker than white matter. Figure 1.16–9 shows such an image. T_1-weighted images are commonly used for structural analysis, including volumetric research.

Another way to obtain T_1-based contrast is with the use of an injectable agent, most of which contain gadolinium. The contrast agents pass through the blood but do not cross the healthy BBB. The gadolinium that is trapped in the molecule decreases the T_1 of the hydrogen nuclei in the water around it and brightens the image intensity wherever it appears. In cases in which the BBB has been disrupted, as is the case with some tumors, the gadolinium agent enters the abnormal tissue easily, and the region appears bright in the images (Fig. 1.16–11).

Losses of signal due to the T_2 decay of the water signal can be used to highlight CSF relative to white matter and gray matter, which lose signal significantly more quickly during echo delays. A T_2-weighted image typically uses a spin-echo type of image, which combines an excitation pulse to create *XY* magnetization and then one or more 180-degree pulses to track the T_2 decay that progresses with time after the excitation pulse. In the case of multiple sclerosis (Fig. 1.16–12), the fluid in pockets within the ventricles, sulci, and lesions has longer T_2 values than the other tissue. If a long echo delay is used, then the signals from gray and white matter decay to 50 percent or less of their maximum values, while the signal from the fluid remains strong.

One disadvantage of T_2-weighted images is the poor contrast between gray and white matter. However, an advantage is that the types of image sequences used for T_2-weighting typically eliminate inhomogeneity-related signal loss, called T_2*-loss, that typically darkens images near the ear canals and sinuses. The sequences employed for T_2 weighting avoid this behavior.

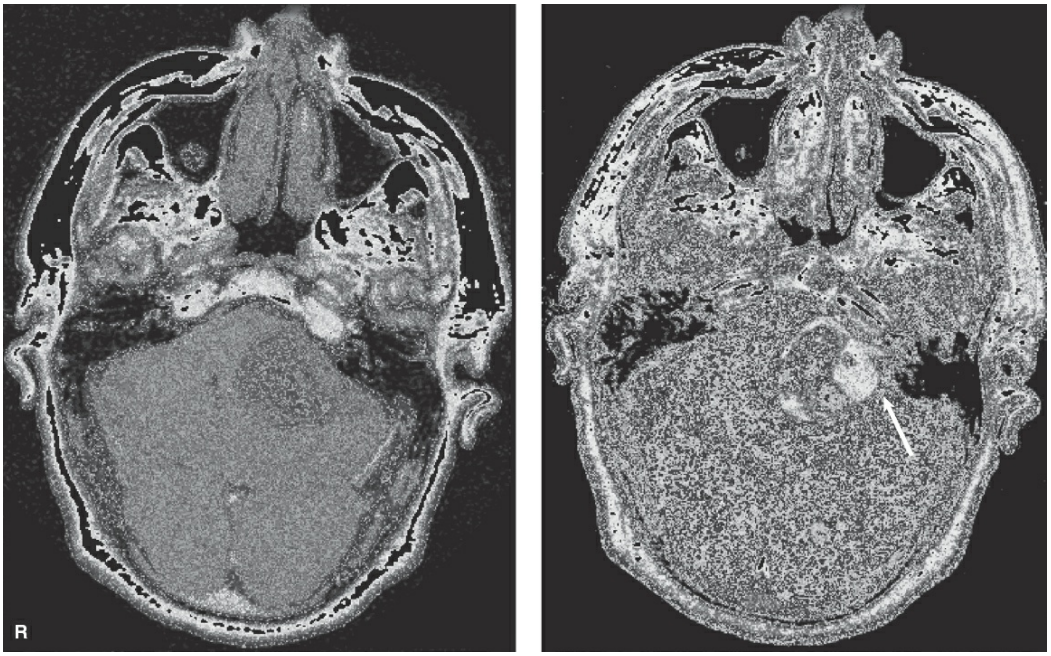

FIGURE 1.16–11. Acoustic schwannoma in the left cerebellar–pontine angle. The T_1-weighted image before the injection of the contrast agent (**left**) demonstrates a mass in the left cerebellar–pontine angle. After administration of the gadolinium contrast agent (**right**), blood vessels are highlighted, and because the tumor is not protected by the blood–brain barrier, the mass is more clearly characterized as being partially cystic. There is enhancement in the internal auditory canal (*arrow*), which is consistent with an acoustic schwannoma. (Courtesy of Robert Fulbright, M.D., Department of Diagnostic Radiology, Yale University.)

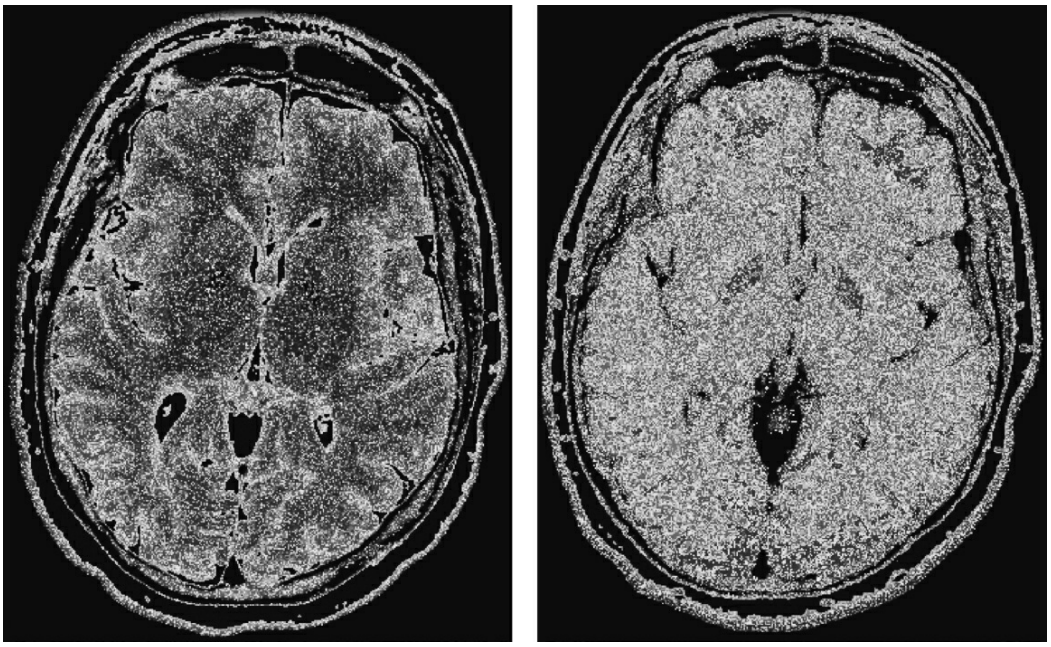

FIGURE 1.16–12. Multiple sclerosis (MS) in periventricular regions. **Left:** A fast spin-echo image enhances materials with large values of T_2, including cerebral spinal fluid (CSF) in the sulci and the ventricles. The MS lesions are enhanced, but it is difficult to differentiate them from the normal CSF. The periventricular MS lesions have T_1 values that are significantly lower than the value for CSF, so T_1 weighting can be used to differentiate the lesions from fluid. **Right:** The lesions are more conspicuous with the fast spin-echo, fluid-attenuated inversion recovery image (FLAIR) image. The FLAIR sequence incorporates

an inversion-recovery pulse that uses T_1 relaxation to suppress signal from the CSF, making it easier to detect lesions that are adjacent to CSF spaces. (Courtesy of Robert Fulbright, M.D., Department of Diagnostic Radiology, Yale University.)

Most MR images are acquired to evaluate structure. A treating physician may be interested in checking for tumors or multiple sclerosis, as already demonstrated in Figures 1.16–11 and 1.16–12. Other common clinical purposes are the detection of atrophy or cerebral malformations. An area of structural MRI that may move beyond its current status as a research tool is volumetric analysis (Fig. 1.16–13). A variety of approaches exist, both manual and automated, but their purpose is to measure the volumes of brain regions, often with separate quantification of white matter and cortical tissue. The specific goal of the research is usually to characterize particular populations of patients with respect to the volumes of various brain structures, commonly the frontal lobe, the hippocampus, or the corpus callosum.

Diffusion MRI

The diffusivity of water can provide a mechanism of contrast for MRI. The most common form, *diffusion-weighted imaging* (DWI), reduces the signal in each image pixel based on the *diffusivity* of water. The process by which the signal is reduced is similar to that of T_2* decay: Gradients are used as an artificial source of T_2* with a strong spatial dependence. A spin-echo-type sequence is applied with the diffusion gradients on each side of the refocusing pulse. Nuclei that diffuse very little between the gradients experience the same magnetic field strength and are refocused by the spin-echo, but nuclei that diffuse appreciably *along the direction of the gradient* experience two different magnetic fields and are incompletely refocused, yielding less signal than is seen for the stationary nuclei. DWI can reveal abnormalities in otherwise normal-appearing tissue, as in regions of edema that surround a cerebral infarct. Regions with edema show greater signal reductions, because water can diffuse more readily when the tissue structure is disrupted and filled with more fluid.

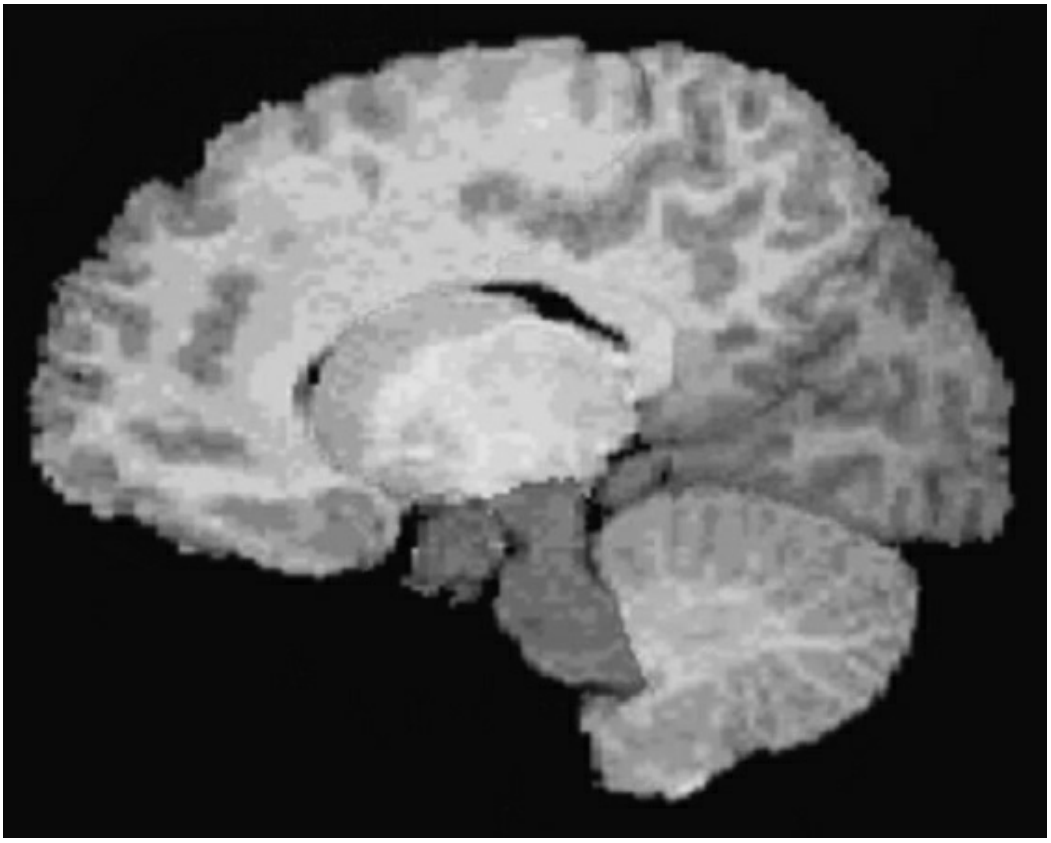

FIGURE 1.16–13. Structural T_1-weighted image that has been *segmented* by brain region for volumetric analysis in a group of patients with posttraumatic stress disorder. In this study, sets of contiguous slices were acquired to cover the entire brain, and then anatomic landmarks were used to demarcate brain regions. (Courtesy of Andrea Jackowski, Ph.D., Universidade Federal de São Paulo, Brazil, and Joan Kaufman, Ph.D., Yale University.)

Water does not diffuse equally in all directions in the brain, because the microenvironment restricts the mobility of water. For example, in white matter, water diffuses much more easily along the length of the fiber tracts than across them. That means that fiber tracts that run in the same direction as the diffusion gradient will show a greater signal reduction than fibers that run perpendicular to the orientation of the diffusion gradient. Application of diffusion gradients in six or more different orientations allows the directionality of the diffusion of the water, the *diffusion tensor,* to be measured for every image pixel, generating *diffusion tensor images* (DTIs). Measurement of the orientation of water diffusion permits a measurement of *fractional anisotropy,* which provides a measure of the randomness of tissue structure. For example, if tissue is highly ordered, then the diffusion of water will be restricted and more oriented than is the case for tissue that is disordered. A more calculation-intensive use of DTIs comes from following the diffusion tensors through chains of pixels to track fibers. Fiber tracking allows detailed analyses of anatomic connectivity (Fig. 1.16–14), which holds great potential clinically but for now remains a research tool.

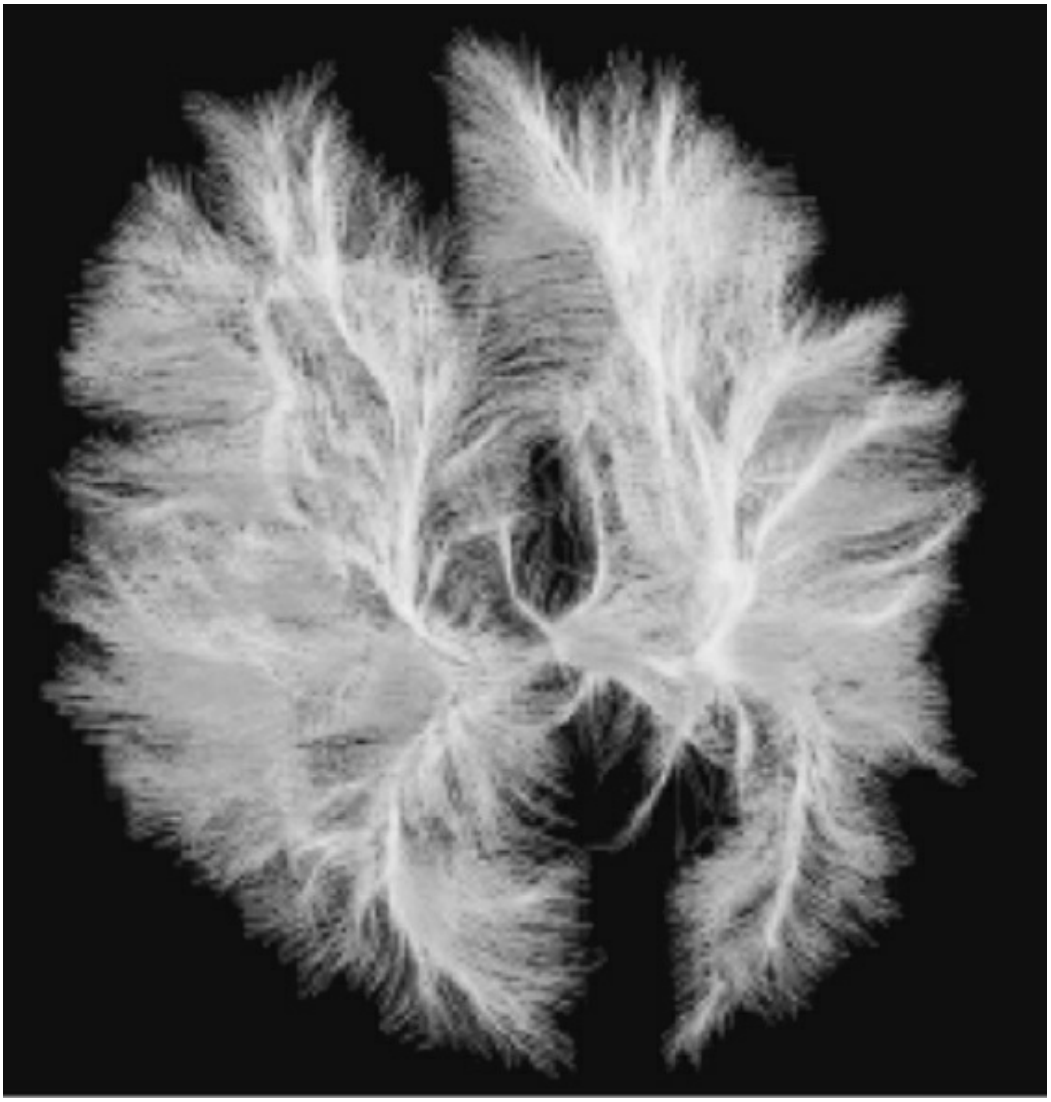

FIGURE 1.16–14. Fiber tracks determined from DTI, presented in an axial perspective. The diffusion tensor was mapped by the acquisition of images with diffusion weighting in multiple directions. A numerical integration technique followed the principal direction of diffusivity from pixel to pixel and slice to slice to track fibers through the brain. (Courtesy of Marcel Jackowski, Ph.D., Universidade de São Paulo, Brazil, and Todd Constable, Ph.D., Yale University.)

Imaging Blood Flow

MRI is used clinically to image blood flow in the head. A physician may need to differentiate between a degenerative disorder and a compromised blood supply for a region of the brain. For that purpose, *MR angiography* (MRA) and perfusion MRI can provide helpful information. The example of Figure 1.16–15 shows *time-of-flight* angiography, which uses a gradient-echo measurement to make images in which inflowing water has more magnetization than stationary spins and thereby appears brighter than the surrounding tissue. Another approach to image perfusion is to use an injection of an intravascular contrast agent that enhances the T_1 relaxation and increases the signal on a T_1-weighted MRI. A T_1-weighted MRI is

obtained without the agent and then with the agent, and the image pixels that are associated with flowing blood will be brighter with the contrast agent.

Another approach is *arterial spin labeling* (ASL), which is based on the manipulation of nuclei in water outside an image slice and the detection of changes in image intensity as those nuclei arrive in the slice of interest. One approach is to acquire two images, one a baseline image and the other identical but for the application of an inversion pulse (*spin tagging*) to generate negative magnetization in the water in a slab of the neck, including any water that happens to be flowing up the arteries when the inversion pulse is applied. If images are acquired through slices of the brain, then any of the water that flows with the blood, up from the neck, will have a negative signal. Upon subtraction of the pair of images from one another, pixels that represent static brain water will approach zero, but pixels that lie in the path of blood flow will appear up to twice as bright as those in the baseline image. In this way, blood flow can be mapped without injection of a contrast agent.

APPLICATIONS OF STRUCTURAL IMAGING TO PSYCHIATRY

To date, the clinical utility of structural imaging in psychiatry has been limited to ruling out organic brain diseases. However, the methodology has provided information that has helped to shape many of the current theories related to the pathophysiological mechanisms associated with several psychiatric disorders, suggesting that both neurodevelopmental and neurodegenerative abnormalities are associated with the various disorders. The use of imaging to contribute to the understanding of disease forms a key part of the RDoC criteria of NIH. Recent advances in the methodology now provide promise that the technique may soon find clinical utility in the diagnosis and prognosis of some psychiatric disorders.

MRI Applications to Dementia

The diagnosis of dementia comprises several illnesses that are characterized by memory loss and other cognitive disturbances. The most common illness associated with dementia is Alzheimer disease, which is characterized by the presence of amyloid plaques and neurofibrillary tangles in several cortical and subcortical limbic regions. The disease is associated with neuronal and glial pathology, as well as global and regional brain atrophy and loss of gray matter. Structural imaging is now beginning to emerge as a clinically relevant component of Alzheimer disease diagnosis. Although the current diagnostic criteria for Alzheimer disease used by the National Institute of Neurological Disorders and Stroke—Alzheimer Disease and Related Disorders and *Diagnostic and Statistical Manual of Mental Disorders,* 4th edition, text revision (DSM IV-TR) do

not consider MRI data in the diagnostic process, a recent position paper published on the basis of an international conference held to reevaluate the National Institute of Neurological Disorders and Stroke—Alzheimer Disease and Related Disorders diagnostic criteria suggested that MRI be used diagnostically as a supportive feature in the diagnosis of Alzheimer disease.

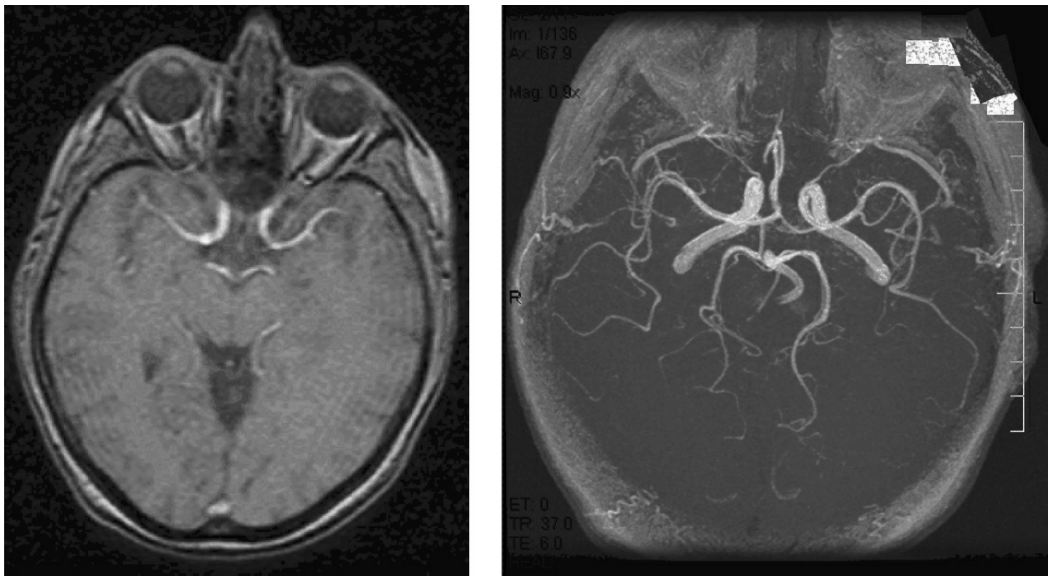

FIGURE 1.16–15. Imaging vasculature by *time-of-flight* (TOF) angiography. **Left:** Images are acquired with rapid repetitions that saturate the signal from nonvascular nuclei, thereby reducing the signals from tissues other than flowing blood. **Right:** After being processed to select image regions that show enhanced signal intensity, flowing blood is identified.

Several MRI changes including increased numbers of subcortical hyperintensities, generalized atrophy, and ventricular enlargement are associated with normal aging. However, it is well established that some changes seem more specific to the diagnosis of Alzheimer disease and may be clinically useful in formulating the diagnosis and prognosis of the disorder. MRI evidence of MTL atrophy appears to be most closely associated with the disorder. MTL atrophy is extremely common in individuals with Alzheimer disease with rates reported to be between 71 and 96 percent depending on disease severity. MTL atrophy is often observed in the early stages of MCI, making it potentially useful as a prognostic tool. However, the sensitivity and specificity associated with the current MRI methodology limit its clinical usefulness in prognostic imaging. One approach that may help to improve the clinical utility of MRI in the diagnosis and prognosis of Alzheimer disease and other forms of dementia is to follow the rate of change in brain structures over time. Longitudinal follow-up studies have shown the rates of volume loss to be significantly greater in subjects with prodromal Alzheimer disease (up to 5 percent brain volume/year) compared to those experiencing normal age-

related reductions (0.1 percent brain volume/year). Other methodological advances are also in development that may help to improve the clinical utility of MRI in the diagnosis of Alzheimer disease. Methods providing an improved ability to rapidly and reliably assess the volumes of specific subregions of the MTL such as the hippocampus and the entorhinal cortex provide some promise. Several recent studies suggest hippocampal volume loss determined by MRI volumetric analyses in subjects with MCI is capable of predicting conversion to Alzheimer disease in the subsequent years. Further advances in the ability to assess the specific subfields of these regions could lead to dramatic improvements in the specificity and sensitivity of the studies. Examples of these methods include voxel-based morphometry, an image analysis technique that can simultaneously identify multiple areas of cortical and subcortical degeneration, and surface mapping techniques that use specific cortical features as constraints to guide the matching of anatomy across subjects to provide information on cortical gray matter thickness for specific brain regions. All of these techniques are aided in part by the development of what are today considered high-field (3 to 4 T) and ultrahigh-field magnets (7 T and above). The possibility of increasing accuracy by the addition of other measures like FDG-PET is showing promise. Although these techniques were originally developed for neuroscience research, their potential clinical relevance is becoming appreciated in providing increased structural resolution and even the potential application of imaging microscopy.

DTI studies have identified significant frontal, temporal, and parietal white matter diffusion tensor changes in MCI and Alzheimer disease patients that appear to correlate with cognitive functioning. Other recent studies have also found relevant microstructural changes in the normal-appearing white matter of patients with Alzheimer disease and MCI, although fractional anisotropy may not be particularly sensitive to the disorder. The changes are consistent with a hypothesized loss of axonal processes in the frontal and temporal regions in Alzheimer disease.

MRI Applications to Schizophrenia

Structural neuroimaging studies have provided some of the most important evidence of abnormal cortical development in patients with schizophrenia. Initial studies reported consistent cerebroventriculomegaly, reductions in total cortical volumes, and regional volumetric deficits, particularly in frontal and temporal cortical regions. One consistent exception to this rule was the striatum where treatment with typical antipsychotics produces increased volume in patients compared to healthy subjects or never-medicated patients. With the application of segmentation techniques to separate cortical gray and white matter within MR images, it became clear that both tissue compartments showed regional deficiencies in patients. The application of more recent techniques have enabled investigators to

describe abnormal patterns of cortical gyrification, curvature, and reduced cortical thickness in patients, although to date, abnormalities are insufficient to discriminate patients from healthy comparison subjects on an individual basis.

Structural MRI also has been useful for characterizing features of the heritable risk for schizophrenia. Schizophrenia is associated with a heritability estimate of 70 to 90 percent. Consistent with this view, family members of individuals with schizophrenia show a pattern of reductions in cortical gray and white matter volume that resembles, but is milder than, that associated with schizophrenia. Further, among healthy individuals and patients diagnosed with schizophrenia, risk genes (COMT, PRODH, NRG1, DISC1, and GAD1) appear to influence MRI findings. One special case, velocardiofacial syndrome (DiGeorge syndrome, 22q11 deletion), is associated with structural MRI, functional MRI (fMRI), and clinical features of schizophrenia.

Structural MRI also has provided some insights into the progressive course of this disorder and the impact of antipsychotic treatment. The most striking findings have been in childhood onset schizophrenia, where a series of neuroimaging studies using a variety of MRI approaches including specified regions of interest, automated voxel-based morphological analysis, and deformation analysis (an extremely sensitive approach for measuring within-subject change) have clearly shown that the development and evolution of schizophrenia are associated with a progressive deviation of regional cortical volumes from those of healthy comparison populations. A surprising finding from these longitudinal studies has been the prominent involvement of posterior cortical structures, including primary and secondary sensory cortices, relatively early in the course of schizophrenia associated with a wave-like progressive emergence of deficits in additional regions that advances anteriorly as the disease advances. This pattern of progression mirrors the development of the cortex and, as a result, would seem to support a disturbance in the development of the cortex. In adolescents and young adults who manifest symptoms of the schizophrenia prodrome, “ultra high risk” state, or who are in their first episode of schizophrenia, many of the changes associated with chronic schizophrenia are already present. Despite earlier data to the contrary from computed tomography (CT) imaging, there are now a growing number of studies in adults with chronic schizophrenia that indicate that even at this stage of the disorder there is subtle but progressive cortical gray and white matter volume loss in a number of cortical regions, relative to time-dependent volumetric reductions in age-matched healthy subjects. In an era where nearly all patients with schizophrenia receive some degree of pharmacotherapy, studies face the challenge of distinguishing the long-term trajectory of cortical volume changes associated with the illness from the long-term effects of antipsychotics that might emerge over decades of

treatment. In this regard, it is interesting to note that some data suggest the possibility that treatment with some "second generation" or "atypical" antipsychotic medications may reduce the short-term, that is, 1 to 2 years, progression of cortical volume reductions or the typical antipsychotic treatment-related increase in basal ganglia volume in patients with schizophrenia. There is hope that these types of studies might reveal neuroprotective effects of future pharmacotherapies for schizophrenia.

DTI has been employed in a growing number of studies that have attempted to characterize disturbances in the organization and integrity of many major white matter pathways associated with schizophrenia. Because DTI measures the movement of water molecules, it is particularly sensitive to the constraining effects of neural projections and their associated glial sheaths upon the normally random diffusion of water as well as the impact of disturbances in neural and glial integrity. DTI findings in schizophrenia include (1) abnormalities in the degree of anisotropy (lack of organization in the direction that water molecules move within a given voxel), presumably reflecting altered integrity or degree of branching of cortical white matter pathways, (2) reduced total diffusivity (total water movement within a given voxel), a dependent measure related to anisotropy and related to integrity of white matter pathways, and (3) reduced size, integrity, or branchedness of specific white matter pathways identified using DTI tractography. Using standard techniques one cannot differentiate the contributions of local dendritic fields, local axonal pathways, and converging inputs from distant neurons.

MR-based perfusion studies have begun to complement other techniques, principally PET and single photon emission computed tomography (SPECT) in describing cortical perfusion deficits associated with schizophrenia. An early generation of PET studies of perfusion and metabolism deficits associated with schizophrenia yielded the impression that schizophrenia was associated with resting "hypofrontality," that is, a relative deficit in the metabolism or perfusion of the frontal cortical structures. However, careful review of this first generation of studies did not yield a firm conclusion of the nature and or validity of resting hypofrontality associated with schizophrenia. More recently, PET, SPECT, and MR-based perfusion studies have been used to characterize task-related alterations in regional cortical perfusion associated with schizophrenia. The novel opportunity provided by MR-based perfusion studies, largely unrealized, is to take advantage of the potential for absolute perfusion quantification in order to reduce interpretation errors when comparing across subjects or across test days where there might be baseline differences that cannot be appreciated using semiquantitative imaging techniques, such as Echo-Planar Imaging (EPI) with Blood-Oxygen Level-Dependent (BOLD) contrast. Studies of cortical perfusion in schizophrenia have been useful to help to evaluate the potential direct

vascular effects of substances that might also be of interest with respect to their modulation of neural and glial function, as in the case of nicotine.

MRI Applications to Mood Disorders

Older MRI studies evaluating larger brain regions have yielded few replicable findings related to mood disorders beyond vague reports of lateral ventriculomegaly. However, more recent findings examining specific subregions of the brain have yielded interesting and replicable findings of smaller volumes in several brain regions including of the basal ganglia, thalamus, hippocampus, frontal lobe, orbitofrontal cortex, and gyrus rectus. Perhaps MRI's greatest contribution to the field of mood disorders research comes from the studies demonstrating reduced hippocampal volume in individuals suffering from MDD. Hippocampal volumetrics have been studied in several neuropsychiatric disorders. Reduced hippocampal volumes have been reported in epilepsy, Parkinson disease, Huntington disease, Cushing disease, and Alzheimer disease, traumatic brain disorder as well as alcoholism, obsessive-compulsive disorder (OCD), and schizophrenia. Despite the confounds introduced by study heterogeneity (age, gender, duration of illness, and responsiveness to treatment) that characterizes the studies in MDD, a 2004 meta-analysis of 12 MRI studies revealed a significant reduction in hippocampal volume in both hemispheres of individuals with MDD. Average reductions of 8 percent on the left and 10 percent on the right were reported. Other evidence suggests that the total number of episodes, duration of illness, and exposure to childhood abuse may further contribute to the degree of volume reduction associated with MDD. This data has been taken as strong support for a model of stress-induced hippocampal toxicity involving glucocorticoid receptor activation, hypoxia, and elevated levels of glutamate neurotransmission. However, recent studies also suggest that pediatric patients with familial depression also have reduced hippocampal volumes, suggesting that reduced hippocampal volume may be suggestive of a risk factor for developing MDD.

The anterior cingulate gyrus is another region that has gained much attention due to findings of significant abnormalities associated with mood disorders. Structural MRI has in one study demonstrated reductions of 39 and 48 percent in the mean gray matter volume in the prefrontal cortex ventral to the genu of the corpus callosum in both familial bipolar depressives and familial unipolar depressives. Although this finding has not been replicated consistently in later structural imaging studies, functional imaging studies have consistently found abnormalities in this brain region, and histological studies have further identified significant glial cell reductions in this area.

Patterns of structural abnormalities in neural regions, including the hippocampus and anterior cingulated gyrus, supporting emotion

processing were recently found to differentiate MDD and BPD suggesting a multivariate pattern classification techniques may hold promise as future diagnostic aids. Similar to the case in dementia, future studies using higher field strengths and other methods to allow for improved measures of hippocampal subfields are likely to provide marked advantages over the currently used methods.

Findings remain inconclusive with regard to the amygdala for which several studies have reported no relationship between amygdala volume and mood disorders, while other studies have found either an increase or a decrease in amygdala volumes in mood disorder subjects compared with those of healthy control subjects. Other studies have attempted to define morphometric abnormalities in the basal ganglia related to mood disorders. While there is some evidence suggesting a possible decrease in basal ganglia volume associated with MDD and a potential increase associated with BPD, in general the data remain inconclusive. Recent advances with high-field magnets such as 7 T may clarify the disparate results, possibly even providing a way to differentiate bipolar psychosis from schizophrenia.

In the area of late-life depression, structural MRI findings have had a large impact on guiding the direction of research. White matter hyperintensities (WMHs) are areas of increased signal intensity that are particularly apparent on T_2-weighted images. As mentioned above, they are commonly associated with aging and are thought to reflect diverse neuropathological underpinnings related to ischemia, hypertension, myelin pallor, inflammatory disease, amyloidosis, and other causes. WMHs are considered important markers in the study of the consequences of vascular risk factors on the brain and are commonly interpreted as a hallmark of chronic cerebral ischemia. The consistent findings of significantly increased numbers of WMHs in individuals with late-life depression, especially late onset depression, suggest an etiologically different subtype of MDD. These findings have fueled the cerebrovascular hypothesis of late-life depression that posits that white matter lesions resulting from cerebrovascular disease disrupt fiber tracks within the frontostriatal circuitry.

Recent studies suggest that white matter changes predate the development of late onset depression and that the severity of the changes predicted subsequent depressive symptoms in the following year even after controlling for age, baseline cognitive function, and depressive symptoms. This adds support to the claim of a causative relationship, but it remains to be determined how important the location and lateralization of the actual WMHs are to the pathogenesis of late-life depression or whether they are merely a marker of a progressing of vascular disease. Future standardized and automated grading systems of WMH, currently in development, should greatly facilitate the use of this modality in clinical practice.

MRI Applications to Alcohol Dependence

MRI studies have been the principal tool to describe in vivo the many sources of neurotoxicity associated with alcoholism including (1) the direct neurotoxic and gliotoxic effects of ethanol, (2) the neurotoxic effects of poor nutrition that often accompany the abuse of alcohol, (3) the excitotoxicity associated with the ethanol withdrawal state, and (4) the possible disruption in adult-neurogenesis–associated ethanol intoxication and withdrawal. These studies documented a striking age dependence of the overall neurotoxicity associated with alcoholism. Alcohol dependence in adolescents has been associated with pronounced and lasting reductions in cortical volume, suggesting a disruption of a neurodevelopmental process associated with the familial risk for alcoholism or exposure to alcohol. In healthy young adults, careful studies suggest that there are relatively subtle neurotoxic effects of ethanol that may not be present in most heavy drinkers. However, with the onset of "middle age" and progressively thereafter, coincident with increasing cumulative exposure to alcohol and bouts of alcohol withdrawal, there appears to be a progressive sensitivity to alcohol-related neurotoxicity. This toxicity is reflected as loss of both gray and white matter, and these losses appear to be related to cognitive impairments. Gray matter microstructure has been reported to change first, and the damage is more severe in those who drink early, possibly because early heavy drinkers often proceed to heavier drinking than people who begin drinking at a later age. DTI studies provide evidence of disturbance in the integrity of white matter tracks associated with alcoholism. The volumetric changes are particularly severe in patients with Wernicke–Korsakoff syndrome, arising from nutritional deficiency. Generally, there is gradual and partial recovery of volume loss with extended sobriety, that is, a more rapid phase of volume restoration over several months with some continued recovery over several years. White matter shows evidence of earlier recovery than gray matter, although relapse leads to accelerated white matter damage. Structural MRI research has disproven an early hypothesis that the brain volumetric changes associated with alcohol dependence and the initiation of sobriety reflect the dehydration and rehydration of the brain, respectively.

FUNCTIONAL MRI METHODOLOGY

Brain function consumes energy, which is supplied by blood flow to carry oxygen and nutrients to the tissue. Changes in brain function lead to alterations in hemodynamics and blood oxygenation, and such changes affect the magnetization that is detected by MRI. The effects are mediated through the magnetic susceptibilities of oxygenated and deoxygenated hemoglobin, working through local T_2^* effects. Deoxygenated hemoglobin perturbs the magnetic field homogeneity in its vicinity, while oxygenated

hemoglobin does not. When the level of deoxyhemoglobin increases, the signal in the vicinity of the blood decreases and lower deoxyhemoglobin leads to larger signals around blood vessels. This leads to the phenomenon of BOLD contrast. Sets of images are acquired under two or more conditions, using different functional or pharmacologic challenges, in which case the technique is sometimes called *ph*MRI, and changes in brain function are detected as increased or decreased image intensity (Fig. 1.16–16).

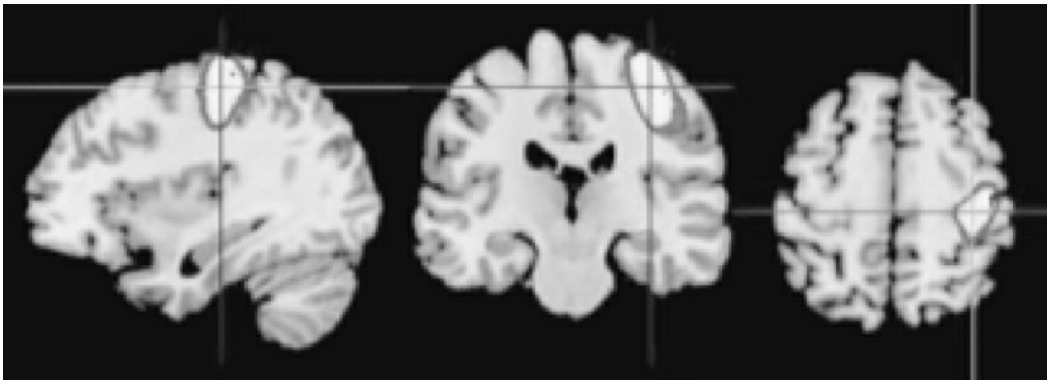

FIGURE 1.16–16. Functional MRI at 3 T during right-hand movement, with sagittal (**left**), coronal (**center**), and axial (**right**) views. The T_1-weighted image forms the base of the display for anatomic identification. A series of images were acquired with the hand alternating between stationary and moving conditions, and then the groups of images were compared to identify regions that showed statistically significant changes in signal during hand motion. The left motor cortex showed activation, with regions of most intense change highlighted.

Because brain function is linked to hemodynamics and the extraction of oxygen from the blood, the level of oxygenated hemoglobin depends on brain function. Paradoxically, when brain activity increases, the fraction of deoxygenated hemoglobin *falls* because blood flow increases so dramatically that it overcomes the extraction, and the MR signal in the vicinity of the blood vessel rises. Therefore, if MRI data are acquired in one condition, and then while the patient is performing some cognitive task that increases the oxygen extraction in Area 9, for example, the signal in Area 9 will be higher, with an increase typically on the order of 1 or less. If a cognitive task is selected that instead decreases the utilization of Area 9, then the second image will be darker by a small percentage in Area 9.

Although fMRI is often said simply to reflect blood oxygenation and/or blood flow, the contrast in fMRI depends on a variety of sources that vary in their degrees of specificity. The mechanisms depend on the diameter of the blood vessels, the blood volume, the rate of blood flow, the amount of deoxygenated hemoglobin in the blood, the strength of the magnet used for the scans, and the method of MRI acquisition. For blood vessels that are larger than 20 μm, it is the *static dephasing* that dominates the signal reduction from deoxygenated hemoglobin. The static dephasing results

from the proximity of deoxyhemoglobin in water in the tissue around a blood vessel, and it can be eliminated with a refocusing pulse in a spin-echo-type MRI sequence. For capillary beds and other small blood vessels, the dominant mechanism is *intravascular dynamic dephasing* and *extravascular dynamic dephasing,* which occur when water molecules wander randomly through the area of distorted magnetic field that surrounds the deoxyhemoglobin. Although intravascular dynamic dephasing provides information about activation, it yields signals that are not always located at the site of activation. The intravascular dynamic dephasing dominates the BOLD response at 1.5 T but exerts progressively smaller effects as field strengths rise. At high magnetic fields such as 7 T, there is a negligible contribution from intravascular dynamic dephasing. The degree of signal change from these mechanisms is also influenced by the orientation of each blood vessel with respect to the static magnetic field B_0. The random orientations in capillary beds yield random shifts in B_0 that, once averaged, reduce the signal in a particular image pixel. Larger vessels exert a net change in phase that will also reduce the signal but will do so in an orientation-dependent manner. A change in oxygenation in a large vessel that flows straight across the thickness of a slice will yield an fMRI change that differs from that of a vessel of the same size that flows along the plane of the slice: A different slice orientation can yield a different degree of fMRI signal change even with the same functional task.

Another mechanism of contrast in fMRI is inflow. During an fMRI study, the nuclei in each image slice are repeatedly excited, and the repeated excitations lead to *saturation*. That is, the observed signal from a series of rapid scans is smaller because too little time is allowed for full T_1 relaxation to the maximum Z magnetization. The saturation only occurs for nuclei that have been excited, which are those that are in the slice and remain in the slice. Nuclei that are carried into the slice by blood flow between the scans have experienced no saturation and therefore have greater Z magnetization. They contribute more Z magnetization per unit volume than their resident counterparts. The more saturated the magnetization within the slice, the greater the impact of inflow contrast will be. Inflow contrast is greater when faster scan repetitions and larger slice excitation angles are used. If the blood flow increases, then the spins that arrive will be less influenced by the slice excitation, so increased blood flow contributes greater signal to the fMRI scan. Any influence that increases blood flow will increase the signal intensity from blood vessels, which may or may not lie in a region of activated tissue, so the inflow contrast contributes nonspecific changes to the fMRI maps.

Once the fMRI signal changes have been evaluated, an additional factor must be considered: The meaning of the baseline image. A healthy adult might show an increase of 1 percent of signal amplitude in the DLPFC while performing a memory task, while a patient shows a 2 percent rise in the

same region for the same task. In both cases, the baseline activity of the neurons is significant with respect to a variety of pertinent parameters, including electroencephalography (EEG), perfusion, and metabolic measurements (Fig. 1.16–17). One interpretation is that the patient's neural circuitry functions less efficiently and requires more energy to do the same work. An equally viable explanation is that the patient had less baseline activity than the healthy individual, but that when the task began, the same region in the patient's DLPFC behaved normally and used a similar amount of energy. In both cases, an abnormality has been identified, but the understanding of the underlying pathology differs with the interpretation. When the challenge is pharmacologic, an additional layer of complexity should be considered: The potential influences of the drug, like acute or chronic tobacco use or heavy drinking on the cerebral vasculature, are independent of the functional demands of the tissue.

Taken together, the numerous mechanisms of contrast show that fMRI maps must be considered with respect to their method of acquisition, slice orientation, magnetic field strength, and pharmacologic influence on the cerebral vascular system. The various sources of contrast must be considered when evaluating the meaning and quantity of change that is observed.

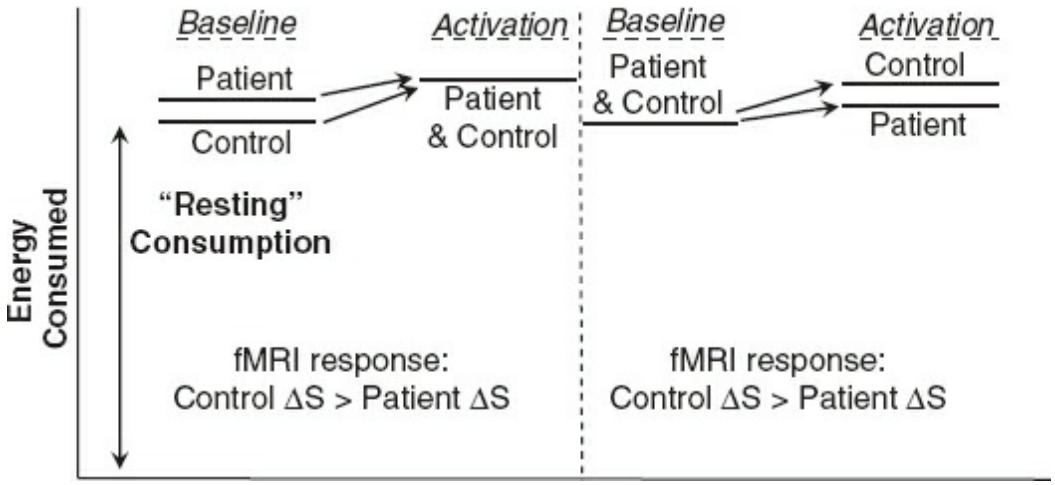

FIGURE 1.16–17. Relative changes of brain energy utilization when changing functional state and their impact on fMRI. The baseline state in the awake human corresponds to a high level of neuronal activity, blood flow, and metabolism. Changes in brain function are typically represented by small changes in brain energy utilization and, as a result, small changes in fMRI image amplitudes. Because fMRI reflects conditions relative to a large baseline, the interpretation of fMRI responses should consider baseline conditions. **Left:** The patient has a higher baseline activity than the control subject, and when the functional task shifts both to the same energetic condition, the patient shows a smaller fMRI response than the control subject. **Right:** The patient and the control subject have the same baseline energy utilization, but the control subject increases energy consumption more than the patient, such that in this case also, the patient has a smaller fMRI response than the control subject. Other factors that may lead to ambiguities are differences in cerebral vascular coupling to energy demand, due to disease or the pharmaceutical interventions.

APPLICATIONS OF fMRI IN PSYCHIATRY

fMRI is a noninvasive neuroimaging technique that can be used to study the neural correlates of complex cognitive, emotional, and perceptive processes. fMRI research holds tremendous promise for advancing the current understanding of the pathophysiology associated with psychiatric disorders where gross brain structure is preserved but disruptions of brain function exist. By revealing the neural correlates of symptoms, cognitive biases, and emotional responses, fMRI studies have already been used to elucidate the pathophysiology of dementia, mood, anxiety, psychotic, and addictive disorders and the mechanisms of psychopharmacological treatments for these conditions. To date, the clinical capabilities of fMRI for informing diagnostic or treatment decisions have not been established. The abnormalities identified by fMRI have thus far lacked sufficient sensitivity and specificity to discriminate individual patients from healthy subjects or from subjects with other illnesses. However, with future advances in the field, such as quantitative fMRI, it is hoped that the technique may ultimately lead the way to a pathophysiology-based classification of psychiatric phenotypes that can be used to improve both research and clinical practice. Examples of such research are provided below.

fMRI of Dementia

fMRI methods provide information that can potentially be used in the study, diagnosis, and prognosis of Alzheimer disease and other forms of dementia as well as providing insights into normal age-related changes in cognitive processing. Evidence that aging is associated with weaker and more diffuse activations as well as decreased hemispheric lateralization suggests either a compensation for lost regional intensity or a dedifferentiation of processing. The weaker activations, especially prefrontally, suggest potential encoding-stage dysfunctions associated with aging. The alterations in these processes that occur with neurodegenerative disease appear to be superimposed on the course of normal aging. fMRI studies have consistently demonstrated that patients with Alzheimer disease have decreased fMRI activation in the hippocampus and related structures within the MTL during the encoding of new memories compared to cognitively intact older subjects, including alterations in what is called the default mode network. More recently, combined PET/structural and FMRI revealed that decreased connectivity was associated with more white matter lesions and greater amyloid burden. Further studies are needed to validate the use of fMRI studies in these populations, particularly longitudinal studies to investigate the pattern of alterations in functional activity over the course of prodromal Alzheimer disease and the relationship to Alzheimer disease pathology. A major factor to consider is

how age- and disease-related changes in vasculature affect the representation of neuronal function by fMRI.

fMRI of Schizophrenia

fMRI has emerged as the primary approach for probing disturbances in the activity of particular brain regions and specified circuits associated with the risk for developing schizophrenia, the symptoms and cognitive impairments associated with schizophrenia, and the impact of antipsychotic treatments. Given the large number of published studies (over 5,000 related papers since 1987 according to PubMed), the range of MR-based approaches included within the rubric of fMRI, and the enormous range of cognitive tests that have been used to stimulate cortical activity, a comprehensive review of fMRI studies of schizophrenia is beyond the scope of this review.

Studies of Low-Frequency Cortical Connectivity

Studies of functional connectivity of brain regions as assessed by fMRI are complementing the structural evidence emerging from DTI studies that the organization and integrity of white matter tracks is compromised in patients with schizophrenia. In these studies, functional connectivity is evaluated by quantifying the covariance of cortical activity across brain regions at rest over extended periods of time.

fMRI studies of perception and cognition alterations associated with schizophrenia support the central hypothesis emerging from MRI, DTI, and functional connectivity studies, that is, that there are widespread alterations in regional brain functional activation that disturb the normal pattern of circuit activation associated with perception and higher cognitive functions including attention, learning and memory, and judgment. Here on the neuroimaging issues related to this research are discussed. A common assumption of this research is that these alterations in regional brain activation constitute the basis of the impairments in perceptual and cognitive processes associated with schizophrenia. While this assumption probably applies in a general way to these studies, EPI-BOLD imaging is actually measuring blood flow and volume changes rather than directly measuring neural activation and, as reviewed earlier current clinical research imaging, typically at 3.0 T, has modest temporal and spatial resolution with respect to the processes that it attempts to measure. Therefore, it is more accurate to assume that these studies provide an indirect measurement, a biomarker, of neural and glial processes that predominantly occur with spatial and temporal characteristics that exceed the resolution of fMRI. An additional complexity is that the relationship between the demands of a task and its ability to stimulate circuit activation is often nonlinear, that is, follows an inverted-U pattern. This issue

emerged in the study of working memory deficits in schizophrenia where it was first shown that patients exhibited reduced prefrontal cortex activation when performing tasks that put demands on working memory. Later, using tasks of graded difficulty, it was shown that patients actually showed prefrontal hyperactivation relative to healthy subjects in response to easy tasks but began to show functional activation deficits in response to moderately difficult tasks when healthy individuals were showing further increases in prefrontal cortex activation.

There are a number of additional challenges associated with the interpretation of fMRI studies of schizophrenia. First, because of widespread disturbances in brain structure, it is difficult to ascribe the underlying abnormality to the structure that performs abnormally, as opposed to its inputs. This limitation has profound implications for cognitive functions that are organized in a hierarchical or "bottom up" fashion, that is, judgment depends on memory, which depends on attention and perception. But there are also executive or "top down" controls of these processes, that is, we are more effective in perceiving and remembering stimuli that we expect or seek. From this perspective, the effort to determine the "central" cognitive impairments of schizophrenia and to ascribe particular cognitive impairments to particular brain regions is a challenging task. Second, because of their circuitry dysfunctions, individuals with schizophrenia may perform more poorly on tasks, or they may successfully perform particular tasks differently than individuals without schizophrenia, analogous to the way that individuals with hearing deficits learn to read lips. With regard to performance failure, studies typically select tasks that can be performed successfully by individuals with and without schizophrenia. However, there are many questions related to the performance deficits themselves. In this case, studies often use tasks with graded levels of difficulty so that the relationships between task difficulty, cortical activation, and performance may be determined. With regard to the second issue, it can be very difficult to know whether divergence between people with and without schizophrenia arises from the primary circuitry dysfunction or from the different strategies that these groups tend to use to solve the same types of problems. This highlights a critical issue for functional neuroimaging, which is the need to define structures of what people do and how they do it. Since one cannot obtain perfect control of these issues, it is useful and important to evaluate performance and, if possible, strategy. The effort to map circuitry dysfunction associated with schizophrenia, already challenging for the reasons outlined above, is made more difficult because the engagement of brain circuits in cognitive processes is highly state-dependent and therefore varies with level of emotional arousal, the impact of other state-dependent processes (such as the symptoms of schizophrenia), the presence of other neurobiological modulators (such as the impact of therapeutic medications,

exposure to substances of abuse, and other psychiatric and medical comorbidities).

fMRI of Mood Disorders

Over 1,200 fMRI studies exploring various aspects of circuitry related to mood and mood disorders have been reported. Studies investigating the neural circuitry mediating adaptive emotion regulation and cognition have been instrumental in advancing the current understanding of the pathophysiology of mood and anxiety disorders. Functional imaging studies using PET, SPECT, and fMRI spectroscopy performed over the last decade have provided us with a much greater understanding of the structure and functional correlates of emotional and cognitive processing as well as mood regulation. Since individuals showing vulnerability to, or suffering from, episodes of depression, mania, and anxiety appear to exhibit impaired cognitive processing and a decreased ability to effectively regulate mood and affect, these findings are considered highly informative in identifying regions of specific interest to the neurobiology of the disorders. Functional brain imaging studies have identified critical neural circuits that modulate emotional behavior. Several frontolimbic networks including (1) the limbic–thalamic–cortical (LTC) circuit, comprising the amygdala, mediodorsal thalamus, and orbital and medial prefrontal cortices, and (2) the limbic–cortical–striatal–pallidal–thalamic (LCSPT) circuit, which includes components of the LTC circuit as well as areas of the striatum and pallidum, have emerged. Reciprocal connections between the various regions within the circuits are believed to modulate emotional responses. The use of fMRI to study the contributions of these networks to the pathophysiology of mood and anxiety disorders provides several unique advantages. The modality offers excellent spatial and temporal resolution relative to other methodologies, allowing the integration of cognitive and emotional tasks that serve as probes to activate the specific networks and further elucidates the interconnectiveness of the individual structures. However, it is important to understand that the measures are always made relative to other measures of brain activation and cannot be considered as absolute measures of brain activity.

Although, as would be expected based on the heterogeneity of the illnesses and the multitude of study designs employed in the exploration, the method has yielded many inconsistent findings related to the function of these networks in individuals with mood and anxiety disorders, there is a growing consensus that disruptions in the function of these circuits do contribute to the pathophysiology of these disorders. For example, several groups have found activation of the amygdala, a region involved in identification and signaling of emotionally significant stimuli, to be increased and abnormally sustained during emotional task performance in subjects diagnosed with mood and anxiety disorders. Interestingly,

abnormal amygdalar activation has also been observed in healthy individuals with the short allele of the functional promoter polymorphism for the serotonin transporter gene, suggesting a possible link of the increased stress-related susceptibility to depression associated with the genotype.

The cingulate cortex, another region commonly found to exhibit abnormal patterns of CBF and metabolism in mood and anxiety disorder, patients receives inputs from the thalamus, as well as several cortical regions, and projects to the amygdala and entorhinal cortex via the cingulum. It functions as an integral component of the limbic system, being specifically involved in emotional processing, learning, memory, and the suppression of inappropriate unconscious priming. Activation of the subgenual ACC has repeatedly been measured to be decreased in unipolar and bipolar depression relative to control samples. However, after volume reductions in this region identified by MRI studies have been accounted, it is possible that the remaining tissue in this area actually exhibits a relative hypermetabolism. Depressed patients have been found to have greater activation of this region during processing of sad stimuli.

Abnormal activation patterns of two other prefrontal cortical regions have also consistently been found to be associated with mood and anxiety disorders. Several studies have reported abnormally elevated levels of CBF and metabolism in the orbital cortex as well as in the VLPFC (a region believed to play a role in the learning of new affective stimulus associations and the reversal of previously learned associations) in individuals with mood and anxiety disorders. Dysfunction of these regions is postulated to predispose individuals to the perseverative cognitions and emotional responses characteristic of depression. fMRI studies have demonstrated an inverse relationship between amygdala and VMPFC activation when healthy individuals reappraise the affective meaning of negative pictures. Individuals showing increased VMPFC and decreased amygdala activation during reappraisal show steeper, more adaptive diurnal patterns of daily cortisol, suggesting that the capacity to successfully modulate limbic circuitry has implications for well-being in daily life. A recent study demonstrating a significantly different correlation between VMPFC and amygdala activation in control versus depressed individuals during an effortful affective reappraisal task further supports this notion.

The DLPFC, a region specifically activated by cognitive tasks related to working memory and attention, has been consistently shown to exhibit abnormal reductions in CBF and metabolism associated with depression. The reduced activity of this region in response to cognitive task performance has been related to the deficits in cognitive performance characteristically observed in conjunction with mood disorders. Several fMRI studies demonstrate decreased DLPFC activity in response to cognitive tasks. Similar to the VLPFC, recent studies also find that

depressed individuals display decreased relationships between amygdala and DLPFC activity, potentially signifying decreased functional relationships among these structures.

The thalamus and the basal ganglia have extensive connections with the amygdala as well as the orbital cortex, VLPFC, and cingulate cortex. In addition the ventral striatum receives key projections from the VTA comprising part of a critical reward pathway of the brain. Unipolar depressed and bipolar depressed subjects both manifest abnormal increases of metabolism and CBF in the left mediodorsal nucleus of the thalamus. Functional imaging studies of these regions in mood disorder subjects have shown them to have a decreased striatal response to happy stimuli during episodes of depression and increased striatal activity in a manic state.

In summary, functional imaging studies afford a new window into the neural circuitry of emotional and cognitive processing, thus providing novel information that can be used in the generation and testing of new hypotheses. In conjunction with work derived from other branches of cognitive neuroscience, these studies have led to the general hypothesis that mood and anxiety disorders result from impaired function of the LCSPT circuits responsible for regulating emotional and cognitive processing. Specifically a "bottom up" model has been presented in which overactivation of limbic structures such as the amygdala drives enhanced emotional responsiveness to stimuli and further decreases the ability to effectively cognitively process the incoming stimuli due to reciprocal inhibition of the DLPFC. A second "top down" model postulates that the ability of higher-order cortical structures to modulate limbic activation is impaired in individuals with mood and anxiety disorders. Several studies have used fMRI protocols to identify specific activation patterns reflecting the top down or bottom up models that seem to be associated with treatment response. Additional studies are now underway to further explore the potential clinical relevance of these models and to determine if the patterns of activation could be used to predict differential treatment response to the various treatment interventions available.

fMRI of Alcohol Dependence

fMRI studies have provided insights into the functional consequences of alcoholism-related neurotoxicity. Studies suggest that recovering alcohol-dependent patients show abnormal activation patterns in frontal cortex, thalamus, striatum, cerebellum, and hippocampus related to impairments in attention, learning and memory, motor coordination, and inhibitory control of behavior. Studies have begun to explore pharmacologic modulation of resting circuit activity to probe mechanisms underlying circuit dysfunction in alcoholism, illustrated by blunted responses to benzodiazepines.

Studies also have provided insights into the vulnerability to alcoholism. Adolescents and adults who abuse alcohol or who are at risk for abusing alcohol show circuit-related dysfunction associated with reward anticipation, response inhibition, control of attention, and other dimensions of cognition. In addition, these individuals show increased limbic and orbitofrontal cortex activation when exposed to alcohol-related cues that elicit alcohol craving, associated with structural abnormalities of the brain. Studies are now attempting to utilize these craving-related changes in fMRI for testing putative pharmacotherapies for alcoholism.

^{1}H MRS

In 1983 spectra of small metabolites, including glutamate and GABA, were observed in the living rat brain by Kevin Behar in the laboratory of Robert Shulman at Yale University. Since that time MRS studies of the brain have expanded enormously in their variety and in their applications to the study of the human brain. Today, MRS is applied through much of the world for diagnosis and basic research on the brain in health and disease.

MRS employs the differentiation of molecules, primarily on the basis of their frequencies. The most commonly used nuclei have been ^{1}H, ^{31}P, and ^{13}C. Figure 1.16–9B shows an example of a ^{1}H MR spectrum obtained in a human brain. The primary components of the MR spectrum are methyl groups of creatine and phosphocreatine, choline, and *N*-acetylaspartate (NAA). The two creatine compounds have resonances that overlap in vivo, so they appear as a single peak. Although creatine and phosphocreatine play key roles in cellular energetics, their combined level remains constant in the face of even severe acute challenges such as ischemia. Some chronic effects such as aging have been reported to increase, decrease, or have no effect on levels of tissue creatine, and this is one of the areas that is currently under active investigation in the MRS community. The choline resonance, comprised primarily of phosphocholine and glycerophosphocholine, is believed to reflect membrane degradation and synthesis, with a negligible contribution from acetylcholine. NAA is found primarily in glutamatergic neurons, including neuronal processes, and it is synthesized by *N*-acetyltransferase, a mitochondrial enzyme. Glia have transporters to import NAA and also contain the enzyme that degrades NAA, raising the possibility of a shuttle of NAA from neurons to glia. Indeed, there is evidence that the acetyl group of NAA is incorporated into myelin. Until the early 1990s, decreases in NAA were believed to signify neuronal death, but since that time studies have found conditions in which NAA decreases and then recovers. NAA is now believed to represent neuronal health, viability, or mitochondrial integrity.

Much of MRS research has been based on these three metabolites that dominate the ^{1}H MR spectrum. Clinical applications have been proposed, including the diagnosis and evaluation of AD, hepatic encephalopathy, and

other disorders. Other metabolites of neurochemical interest that can be detected using ^{1}H MRS are myo-inositol, glutamate, glutamine, GABA, and glutathione (GSH) and these have all been scrutinized from a research perspective in a variety of neuropsychiatric disorders. Glutamate, glutamine, and particularly GABA have required advanced MRS methods for detection in the human brain because they overlap with one another and are modulated into complex patterns under the influence of *J*-coupling. Methods classed as *spectral editing* are used to isolate the signals of GABA, glutamate, GSH, and other compounds from larger, overlapping resonances. Research is currently under way to assess if and how total tissue levels of glutamate and glutamine with respect to neurotransmitter activity and whether any of the MRS-detected chemicals can indicate neuroinflammation.

Magnetic resonance spectroscopic imaging (MRSI), also called *chemical shift imaging* (CSI), combines the power of chemical discrimination with the ability to map their spatial distribution, and it is most commonly implemented for ^{1}H MRS. Because the signals from most detected neurochemicals are typically 5 to 6 orders of magnitude lower than that of water, the signal-to-noise ratio is much lower for MRSI measurements, significantly longer acquisition times are required to obtain adequate sensitivity, and the resolution is lower than what is normally used for MRI. However, the ability to view large areas of the brain simultaneously has led to the beginnings of clinical applications of this approach to conditions in which abnormalities must be localized, as in epilepsy, and conditions of dispersed changes in the brain, such as multiple sclerosis and alcoholism.

^{31}P MRS

In 1978, Britton Chance and colleagues published MR spectra of the phosphorylated metabolites in the brain in a living mouse. ^{31}P MRS yields measurements of high-energy compounds, including phosphocreatine and ATP, and it can provide measurements of phosphomonoesters (PMEs) and phosphodiesters (PDEs) (Fig. 1.16–18), with applications to study schizophrenia, depression, and other disorders. The ^{31}P isotope nucleus is much less sensitive than that of ^{1}H on a per-atom basis, and further sensitivity is lost when one considers that many of the ^{1}H resonances represent methyl or ethylene groups that provide three or two nuclei per resonance, instead of just one. Favoring sensitivity is the fact that 100 percent of naturally occurring phosphorus is the isotope ^{31}P. The unique information that it provides on high-energy phosphates and other metabolites without destroying the tissue makes it a powerful tool for brain research. Some laboratories are applying MRSI of ^{31}P to map these neurochemicals throughout the brain.

The dominant resonance in the ^{31}P spectrum of the brain is that of

phosphocreatine, which by convention is assigned a chemical shift of 0 ppm. To the right are three resonances that represent the γ, α, and β resonances of nucleotide triphosphates (NTPs). To the left are the PME, PDE, and inorganic phosphate. The PME and PDE resonances are believed to reflect mobile components of lipids and membrane metabolites. Inorganic phosphate is free phosphate that has been released from phosphocreatine and ATP during energy consumption. In cases of extreme insults, such as ischemia, the phosphocreatine resonance falls, the NTP resonances decrease, and the inorganic phosphate peak rises as the high-energy phosphate pool is exhausted and the phosphate groups are released. In psychiatric disorders, the changes are generally more subtle, with some reports of altered PME or PDE, although the directionality and extent of the changes are uncertain at this time. Recent advances have moved kinetic studies of ^{31}P metabolism to application in the brain, from muscle, where they have been used for many years. A technique called *saturation transfer* reduces the signal of ^{31}P in one chemical, like phosphocreatine, and then tracks the signal reduction in ATP when the phosphorus is transferred from phosphocreatine to ATP. This approach has allowed the measurement of phosphocreatine and ATP synthesis in the brain.

^{13}C MRS

Less common is ^{13}C MRS, used to detect the nonradioactive isotope carbon-13, which naturally occurs as 1.1 percent of the world's carbon. Like most ^{1}H MRS and all ^{31}P MRS, ^{13}C MRS remains a research tool, and in that role it has proven useful for the measurement of kinetics of neuroenergetic processes and neurotransmission. Its primary application in the brain has been the measurement of carbon oxidation and glutamate–glutamine neurotransmitter cycling.

^{13}C carries the powerful advantage of a broad chemical shift dispersion that allows the resolution of many metabolites that cannot be resolved in the ^{1}H MR spectrum. For example, glutamate and glutamine C4 and C3 resonances are completely resolved even at a magnetic field strength of 1.5 T, with the resolution improving further at higher field strengths. The great disadvantage of ^{13}C is that its sensitivity is even less than that of ^{31}P on a per-atom basis, and the sensitivity is further reduced by its low natural abundance. However, that low natural abundance can be turned into an advantage if viewed from the perspective that it contributes very little background signal (Fig. 1.16–19). That low level of background means that if ^{13}C is introduced into metabolic pathways and is incorporated into neurochemicals, then almost everything that is detected is known to result from the added carbon. This was first done in cell suspensions and then rabbits and rats and is now applied by several laboratories to studies of metabolism in the human brain. The most commonly used method today is the administration of ^{3}C-glucose or ^{13}C-acetate, which leads to labeling of

glutamate, glutamine, and GABA through glycolysis, oxidation, and the release and trafficking of neurotransmitters among neurons and astrocytes. The incorporation of ^{13}C is tracked over time simultaneously in multiple metabolites, and the more rapid the metabolic processes in the relevant pathways, the faster will be the appearance of the ^{13}C in the MR spectra. The glucose generally provides more information about neurons, while acetate yields greater sensitivity to glial metabolism.

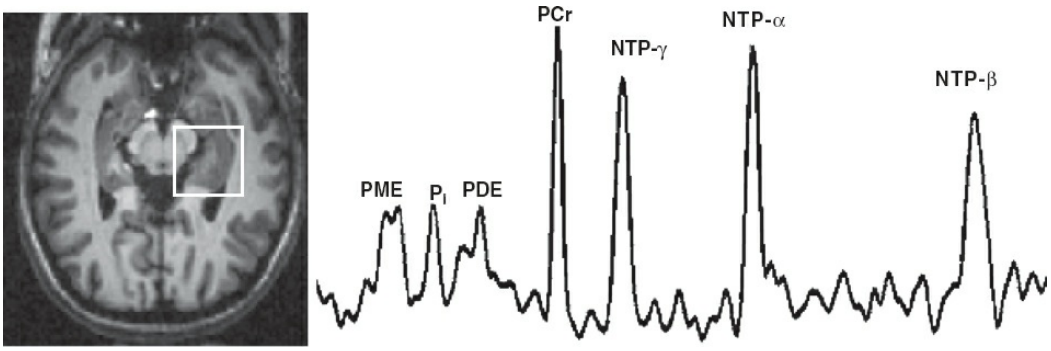

FIGURE 1.16–18. ^{31}P MRS in a healthy volunteer, using a 4-T magnet. **Left:** T_1-weighted image obtained with inversion recovery to null the CSF and create contrast between gray and white matter. **Right:** Spectrum that shows, from left to right, phosphomonoesters (PMEs), inorganic phosphate (P_i), phosphodiesters (PDEs), phosphocreatine (PCr), and the γ, α, and β phosphate residues of the nucleotide phosphates. (Courtesy of Jullie Pan, M.D., Ph.D., Yale University.)

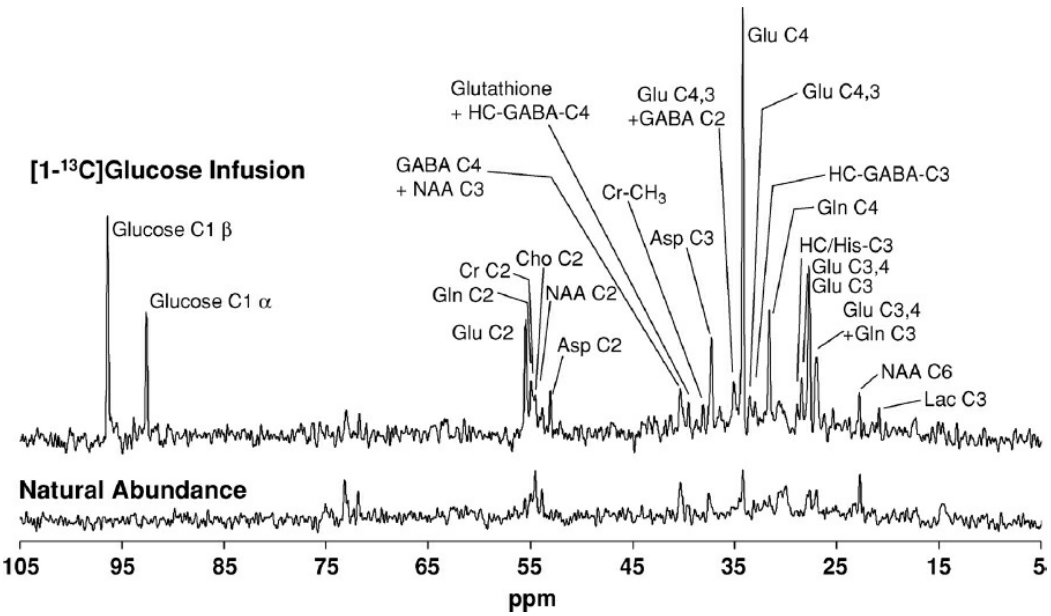

FIGURE 1.16–19. ^{13}C MR spectra from a human brain before (**bottom**) and during (**top**) a 2-hour infusion of [1–^{13}C]glucose. Before the infusion, only the 1.1 percent natural abundance signals can be seen in the spectrum. When exogenous ^{13}C is provided in glucose, through metabolism the isotope appears in numerous products in the brain and can be detected with ^{13}C MRS. At left can be seen the glucose itself. The largest single resonance is that of the C4 of glutamate (Glu C4). C3 and C2 of glutamate are also present, as are C4, C3, and C2 of glutamine (Gln). GABA is labeled at the C2, C3, and C4 positions, and aspartate (Asp) is labeled at C2. Lactate (Lac) is labeled at the C3 position. NAA and creatine (Cr) appear similar in the natural abundance and labeled spectra, because their

synthesis is too slow for significant labeling to occur in 2 hours.

Different methods have been developed to detect ^{13}C, with the appropriate choices based on the applications. The simplest approach is *direct detection,* which requires excitation and detection only of the ^{13}C nuclei. The sensitivity can be increased by up to threefold by preapplying radio frequency at the ^{1}H frequency, in what is called the nuclear Overhauser effect (NOE). *Polarization transfer,* a more complex approach, can increase the sensitivity by up to a factor of four by transferring some of the larger polarization of coupled ^{1}H nuclei to the ^{13}C nuclei through covalent bonds, and for the acquisition the ^{13}C nuclei are observed. *Proton-observed/carbon-edited* (POCE) spectroscopy provides a greater improvement in sensitivity. POCE detects ^{1}H that are bonded to ^{13}C, and because the ^{1}H nuclei are observed, the detection has the much greater sensitivity of ^{1}H MRS. The greater sensitivity means that volumes of a few milliliters can be studied, instead of 75 mL or more for the methods that acquire data in the ^{13}C domain. The primary disadvantage is that the detection in the ^{1}H domain eliminates the benefits of the high resolution available from the ^{13}C chemical shift dispersion. *Heteronuclear single quantum coherence spectroscopy* (HSQC) provides a way to achieve the resolution available in the ^{13}C spectrum with the sensitivity of the ^{1}H acquisition, by yielding data that simultaneously contain the information that would be seen by acquiring both nuclei, and this holds promise for achieving higher spatial resolution.

APPLICATIONS OF MRS IN PSYCHIATRY

As MRI provides information about brain structure and fMRI provides information about regional brain function, MRS provides information related to brain chemistry and metabolism. In many ways it is the youngest and least developed of the MR methodologies, and until more recently its application to psychiatry has also been limited by its relatively poor spatial and temporal resolution and need for technical expertise. However, information provided by MRS studies in vivo over the past decade has proven valuable to identify neurochemical and metabolic abnormalities associated with several psychiatric disorders that have helped to reshape the current thinking related to the pathophysiological models of these disorders. With the advent of improved MRS imaging techniques and dissemination of expertise, the modality is gaining greater application to psychiatric research that is allowing the field to examine novel hypotheses related to pathophysiology and pathogenesis of psychiatric disorders.

MRS in Dementia

^{1}H MRS presents the opportunity to obtain noninvasive measures of several neurochemicals related to neurotransmission, energy metabolism,

and cellular function. Studies using ^{1}H MRS have shown a trend for a general reduction in NAA measures with increasing age in medial temporal and frontal cortical brain regions, but with greater decreases in dementia and NAA levels related to plaque density. NAA is perhaps the most reliably measured metabolite in the brain due to it being present between 7 and 13 mmol/kg and the physical characteristics of the molecule (presence of identical three hydrogen atoms on a methyl group that provide triple the sensitivity of what would be provided by a single proton). As mentioned above, the significance of brain NAA content is not completely understood. It is commonly believed to serve as a marker of neuronal health, but it is also known to reflect aspects of mitochondrial energy metabolism and myelin maintenance. The studies in MCI and AD are in general agreement, reporting patients with these disorders to have decreased levels of NAA and increased levels of myo-inositol (a form of inositol normally found in the brain that contributes to osmotic regulation) compared to those of age-matched comparison subjects. One interpretation is that decreased NAA and increased myo-inositol are related to the loss of neurons and an increase in gliosis. However, the loss of NAA may reflect decreases in neural function. This hypothesis is supported by the fact that the changes in NAA can be reversed in some conditions. Several groups have also reported a correlation between the reductions in NAA content and clinical neuropsychological scores. Interestingly, this correlation has also recently been reported in healthy volunteer subjects.

Overall, studies of other metabolites have been less convincing in relationship to AD and other dementias, but a few studies suggest decreased levels of glutamate and elevated levels of glutamine can be found in subjects with MCI impairment and AD. Together these findings are believed to reflect a combination of ongoing metabolic dysfunction and increased gliosis. At present there is no clinical use of MRS in the diagnosis or treatment of dementia.

MRS in Schizophrenia

^{1}H MRS has been applied widely in studies of cortical chemistry in schizophrenia. These studies documented reductions in NAA levels in many cortical and limbic brain regions in schizophrenic individuals and smaller reductions in family members of people diagnosed with schizophrenia. NAA level reductions have been described in some brain regions in medication-naïve patients, but there also do appear to be progressive reductions with advancing illness and continued antipsychotic treatment. Because NAA is localized to neurons, reductions in NAA levels may reflect the postmortem findings of reduced neuronal size and reduced dendritic arborization as well in vivo evidence from DTI studies of disturbances in the integrity of long fiber pathways. However, NAA is synthesized by a mitochondrial enzyme. As a result, reduced NAA levels

may reflect metabolic disturbances that may be related to mitochondrial dysfunction suggested in postmortem studies or regional metabolic alterations associated with the pathophysiology of schizophrenia or the impact of treatment. This possibility is consistent with the emerging differences between the impact of typical and atypical antipsychotic treatment upon NAA levels in patients. Other metabolites have been measured in ^{1}H MRS studies of schizophrenic patients. The most interesting findings to date may be the description of normal or low levels of glutamate and increased levels of glutamine and GABA in medication-free patients with schizophrenia, with evidence accumulating that medication may be responsible for some of the abnormal metabolite levels reported in schizophrenia. Brain glutamine is synthesized primarily by glia from glutamate that emerges as a by-product of energy metabolism as well as from glutamate that is taken up by glia following its release by neurons and glia in the context of excitatory neurotransmission. Thus, it is possible that the alterations of glutamine and glutamate in individuals diagnosed with schizophrenia could reflect changes in glutamatergic neurotransmission. Consistent with this view, one preliminary study suggested that glutamine elevations were not present in medication-free patients who were receiving benzodiazepines, drugs that would be predicted to suppress excitatory neurotransmission.

A number of other metabolites have been examined using ^{1}H MRS and ^{31}P MRS in patients with schizophrenia including glutathione, choline, PME, and PDE. These metabolites are of potential interest because of studies of these metabolites that suggest that schizophrenia is associated with inflammation and abnormalities in membrane integrity and might be reflective of disturbances in the processing of myelin suggested by postmortem and clinical molecular genetic studies. However, the relationships of these metabolites to specific membrane disturbances associated with schizophrenia remain unclear. Also, in the ^{31}P MRS studies, inconsistencies in the findings made by various groups have called these findings into question, generally. However, they may be consistent with recent findings using magnetization transfer that oxidative phosphorylation is slower in schizophrenia, using a volume of tissue that included a large fraction of white matter, indirectly suggest that myelin may be compromised in patients with schizophrenia.

MRS in Mood Disorders

There has been a rapid increase in the use of MRS in the study of mood and anxiety disorders over the past decade. While the methodology does not yet enter into the clinical practice related to the diagnosis and treatment of mood disorders, it has in part led to an evolution in the concepts of the pathophysiology of the disorders. The majority of studies have used either ^{1}H MRS or ^{31}P MRS to provide information about the abnormalities in

neurochemistry and energy metabolism associated with the illnesses. A few other studies have used ^{7}Li and ^{19}F MRS to study the pharmacokinetics and brain concentrations of drugs used in the treatment of mood and anxiety disorders. Recent new work suggests that the use of ^{13}C MRS will be useful in further defining the pathophysiology of the disorders.

To date, ^{1}H MRS has been the most widely used MRS methodology in the study of mood disorders. While the technology continues to evolve, several of the measures are becoming more routine and reliable, thus allowing for investigators to more confidently draw conclusions from data obtained at different sites using different protocols and spectrometers.

There have been over 200 published studies examining NAA levels in relation to mood and anxiety disorders. However, due to the heterogeneity of the illnesses studied and variety of regions examined it remains difficult to draw firm conclusions regarding NAA's contributions to the neurobiology of mood disorders. In general, there have not been consistent findings showing NAA changes in association with major depression. Yet, there have been specific studies suggesting that NAA may be reduced in the hippocampus of depressed and anxious patients. There is also some evidence suggesting that NAA levels may be reduced in the frontal lobe of bipolar patients, potentially serving as a diagnostic marker. The creatine peak is believed to remain relatively constant in most disorders, leading to the use of creatine as an internal standard in many studies. Studies examining creatine in relation to mood disorders appear to be in general agreement with this view, but exceptions do exist. Choline is an essential precursor to membrane lipids and the neurotransmitter acetylcholine. A series of studies seem to demonstrate elevated choline levels in the basal ganglia of mood disorder subjects compared to those of healthy comparisons. These studies have been suggested as evidence of altered membrane turnover and impaired signal transduction mechanisms within the basal ganglia of patients.

As the major inhibitory and excitatory neurotransmitters in the brain, it is not surprising that the amino acid neurotransmitters GABA and glutamate have been linked to the pathophysiology of several neuropsychiatric disorders, thus making them extremely interesting targets for MRS investigation. Due to overlapping resonances, these compounds are more difficult to measure using standard ^{1}H MRS methods. This has delayed the use of MRS to study the amino acid neurotransmitters and led many to use the term Glx to refer to the combined measure of the GABA, glutamate, and glutamine. With the development of several new editing techniques, several groups are now capable of isolating the individual peaks. The relatively recent application of MRS to investigate amino acid neurotransmitter systems has proven highly productive in the area of mood disorders research, and there is mounting evidence that suggests markedly abnormal concentrations of GABA and glutamate in several brain regions

of mood and anxiety disorder patients. Furthermore, recent studies suggest that antidepressant treatments may be capable of reversing some of these abnormalities. These data helped to increase the awareness of a potential role of the amino acid neurotransmitters in the neurobiology of mood disorders and the mechanism of antidepressant action. Since it is now recognized that glial cell function is critical to amino acid neurotransmitter metabolism, these data in conjunction with several postmortem findings have also provided support for a pathophysiological model of glial cell impairment in relation to MDD. Recent observations of altered brain glutathione levels in mood disorders may also be related to alterations in glial/neuronal interactions.

^{31}P MRS provides in vivo measures of brain membrane phospholipid metabolism, high-energy phosphate metabolism, and intracellular pH. Initial ^{31}P MRS studies in bipolar disorder set out to test the hypothesis that the phosphoinositide pathway is enhanced in bipolar disorder and to explore the contributions of membrane phospholipids and membrane defects in bipolar disorder by measuring levels of PME and PDE. A meta-analysis of these studies found significant diagnosis- and mood state-dependent abnormalities in PME content in patients with bipolar disorder, which suggests dysregulation of brain-signal transduction systems and membrane metabolism may have relevance in bipolar disorder. Further, they suggest that lithium treatment increases PME levels, providing a possible clue into the mechanism of action.

Other studies have attempted to evaluate whether neuroenergetic defects are related to mood disorders by measuring the concentrations of nucleotide phosphates (the three phosphate residues from NTP) and phosphocreatine. In general, these studies have found lower levels of nucleoside phosphates in the basal ganglia of subjects in depressive episodes, suggesting that high-energy phosphate metabolism is altered in the basal ganglia of subjects with depression and drawing attention to the possible role that the mitochondria may play in mood disorders.

Reduced pH_i has been observed in pathological states arising from ischemic insult to the brain. Abnormalities in pH_i concentration have also been reported in bipolar subjects. Several studies have found significantly reduced pH_i in both the basal ganglia and whole brain of bipolar subjects in the euthymic state, while other studies have found increased pH_i levels to be associated with the bipolar and depressed states. While the specific pathophysiology underlying these effects remains unknown, the findings have helped to focus attention on the potential contributions of neuroenergetics to mood disorders.

As described earlier in this chapter, the unique physical properties of ^{13}C make it an excellent tool to measure rates of glucose metabolism as well as amino acid neurotransmitter cycling, and being a stable isotope, it does

not present the risks and challenges associated with radiochemical research. Studies using glucose labeled with ^{13}C have been used to measure the rates of glucose oxidation, glutamate/glutamine cycling, and GABA synthesis. Studies using ^{13}C-labeled acetate have recently yielded measures of glial cell metabolism, a potentially valuable tool considering the recent findings suggesting that glial cell pathology may play a prominent role in the neurobiology of mood disorders. Although the application of ^{13}C MRS in vivo is relatively new to psychiatry, recent data show significantly lower rates of oxidative metabolism in unmedicated patients with major depression. Future studies using this modality are likely to provide new insights into the association relationships between mood disorders, neuroenergetics, amino acid neurotransmitter functioning, and glial cell function.

MRS in Alcohol Dependence

^{1}H MRS studies evaluating NAA and choline have provided neurochemical evidence that complements the MRI findings related to the emergence and recovery from alcohol-related neurotoxicity. ^{1}H MRS studies of GABA have provided insights into alterations in cortical inhibitory neurotransmissions associated with the recovery from alcohol dependence. During acute withdrawal, cortical GABA levels appear to be normal, consistent with animal studies. With recovery from alcohol dependence, cortical GABA levels appear to decline and may be significantly below the level seen in healthy subjects with extended sobriety. These time-dependent changes in cortical GABA levels during sobriety have been hypothesized to reflect recovery from the adaptations in $GABA_A$ receptor populations produced by chronic ethanol exposure. ^{13}C measures may reveal factors that promote energetic factors, which promote heavy drinking and relapse, such as enhanced ability of the brain to oxidize alcohol and its metabolic product, acetate. To the extent that the brain becomes dependent on acetate as a fuel, withdrawal from alcohol may be compounded by the near-absence of circulating acetate once patients stop drinking. Further, when alcohol is oxidized, it creates acetaldehyde, which has been shown to have rewarding properties inside the brain, even though systemically acetaldehyde is aversive.

^{7}Li and ^{19}F MRS

A few studies have attempted to use ^{7}Li MRS to study the pharmacokinetics of lithium as it relates to the treatment of mood disorders. These studies have demonstrated that lithium relatively rapidly clears the BBB and has approximately a 28-hour half-life in the brain. A potentially powerful clinical application of the methodology is the ability to measure brain concentrations of Li relative to serum levels and to assess intracellular

versus extracellular lithium. While limited in number, these studies seem to suggest only a modest correlation between serum lithium levels that are commonly used to titrate treatment dosage and brain lithium concentration. Similarly, studies have employed ^{19}F MRS to determine the pharmacokinetics and the relationship between serum and brain levels of fluorine-containing drugs such as fluoxetine and fluvoxamine. While such studies are of limited clinical usefulness at present, they may be useful in determining dose–response relationships for other compounds in the future.

PRACTICAL CONSIDERATIONS

Although the technology invokes some complexities, there are practical considerations of safety and comfort that are of great importance when considering a possible MRI or MRS scan for a patient.

Safety

The primary consideration is the safety of the patient. MRI and MRS are considered to be among the safest ways to examine the human body. Because MRI and MRS use magnetism and radio waves, and not x-rays, they are not known to be mutagenic or associated with cancer risk like methods that use ionizing radiation. Radio waves can cause heating of tissue, including of the brain, so the United States FDA has set guidelines for magnet strength and exposure to radio waves, and the manufacturers and operators of MR equipment carefully observe those guidelines. For exposure to the radio frequencies, powers, and durations that are used in human MR scans, no harmful effects have been seen.

MRI and MRS pose some risks for certain people. People who have a pacemaker or certain metal objects inside their bodies are generally not candidates for MR scans, because the strong magnets in the MR scanner might cause the implants to malfunction or move. Aneurysm clips pose a particularly major risk. As current technologies increase magnetic field strength, one must take care with metallic products that are designated to be MR-safe. As far as can be achieved, companies attempt to maintain the necessary qualities of their products while removing as much of the ferromagnetism from the material as possible. However, trace levels of ferromagnetism may remain, posing a greater risk as magnetic field strengths increase. The claims of MR compatibility are generally true for the magnetic field strength of 1.5 T or lower, and in some cases, the devices are safe at 3 T. However, many devices have not been tested above 1.5 T, and fewer still above 3 T, so for now, many metallic implants pose an unknown and unacceptable risk above 3 T and should be checked individually for compatibility at a given magnetic field strength.

Another risk is that of a metallic object flying through the air toward the

magnet and hitting the patient. To reduce this risk, MR facilities usually require all personnel, including the patient, to remove all metal from their clothing and all metal objects from their pockets. In many cases, patients will be required to wear hospital garb. Nothing metal can be brought into a magnet room except for equipment or supplies that have been specially screened by the MR operators. The doors to the magnet room should be kept closed except for brief entry and exit by the operators and patient or by personnel directly approved by the operators. Failure to safeguard against this risk has resulted in catastrophes, such as the death of a child when a magnetic oxygen cylinder crushed his skull at speeds in excess of 50 miles per hour. A less dramatic but more common risk is that of injury to a patient's eye from a flying ballpoint pen or a paperclip.

A common question is about the safety of pregnancy. At magnetic field strengths of 1.5 T, in use since the early 1980s, and 3 T, in use since the early 1990s, no adverse effects of the magnetic field have been associated with pregnancy or reproductive health. Above magnet field strengths of 3 T, far fewer data exist, although no ill effects have been reported at this time. At 7 T and higher, some individuals experience nausea or dizziness when entering or leaving the magnetic field, due to magnetic field effects on the inner ear while moving through the region at which the magnetic field changes near the bore of the magnet.

The use of injectable contrast agents poses an additional but small source of risk. There are the risks associated with the placement of an intravenous line, including bruising and infection, although the lines should be placed with care under sterile conditions that minimize those risks. The FDA has approved contrast agents that contain the metal gadolinium for use in human subjects and does not recognize any major risks with its use. Less than 3 percent of patients may experience mild nausea, headache, or dizziness after the injection, and those side effects usually resolve themselves without treatment. Less than 1 percent of patients experience an anaphylactic reaction, including hives, itching, or difficulty breathing. Subjects should be asked if they have a history of allergic reactions to MR or other injected contrast agents or if they have a history of kidney disease, asthma, allergic respiratory disorders, or anemia or other diseases that affect red blood cells. Women who are pregnant are generally not given gadolinium contrast agents in the United States, although in Europe it has been judged to be safe and is permitted. Women who are breastfeeding are instructed to express milk for 48 hours after the scan before returning to breastfeeding.

A magnet quench was described earlier in this chapter as the sudden de-energization of the magnet, accompanied by a rapid boil off of liquid helium and possibly nitrogen. Nitrogen and helium are inert and nontoxic, but in confined spaces they displace oxygen and can cause suffocation. Therefore, MR scanners are equipped with ventilation systems to remove

those gases in the event of a quench. Furthermore, on the very rare occasion that a quench occurs, the operators are trained to remove a patient immediately from the magnet room. Patients are unlikely to be present at the times that a quench is most likely to occur, these being during magnet energization that occurs at installation or after major equipment upgrades, or when helium is being added to the magnet to replace material that has slowly been lost from the system. Quenches may also occur if a large metal object strikes the magnet. To date, no patients have been injured or killed by quenches.

What to Tell the Patient before a Scan

A patient should be told that the study requires lying on a bed, usually with a detector that has some rods that are placed around the head. The bed is rolled into the scanner, which typically looks like a tube, so that the individual's head is approximately centered along the length of the tube, and the feet, legs, and sometimes even the waist lie outside the tube. Patients should be told that the machine does not move, although they will hear knocking and other sounds and possibly feel some vibrations while the scanner is acquiring the data. It is crucial to stress the need to hold still. Movement can lead to images that are blurry or otherwise uninterpretable. If there is an injection of contrast agent, the individual may feel some discomfort associated with the intravenous line, just as any intravenous catheter may cause some bruising or other discomfort. Otherwise, MR has no painful parts.

Some people feel uncomfortable or anxious because of having to lie still in a tube that is generally less than three feet in diameter and several feet long. If a subject is feeling anxious, then a small dose of an anxiolytic drug such as a benzodiazepine is sometimes administered, as long as it does not adversely affect the measurements. A subject should be told that the operators will monitor them constantly by sight and sound, so a patient can ask the operator to be brought out of the scanner. It is important to warn a patient that some people feel dizzy, develop some stomach upset, experience a metallic taste, or find tingling sensations or muscle twitches. These sensations usually go away quickly, but it is important to tell the research staff if they occur.

The patient should be presented with an MR safety questionnaire to be answered very carefully, and it is important to verify the answers verbally with each patient. If parents accompany a child into the MRI room, the parents should be screened likewise. The questionnaire will ask about surgical history and the presence of a variety of potentially MR-incompatible items, including pacemakers, intrauterine contraceptive devices, aneurysm clips, and injuries from shrapnel and metal work, especially to the eye. Patients commonly ask questions about implanted metal including dental work and screws used to repair broken bones.

Fillings are safe. Most permanently installed dental work is safe. Partial plates that can be removed should be removed before the scan, partly for safety but primarily because ferromagnetic material in some of them can cause T_2* effects that compromise the quality of the MRI or MRS. Bone screws and other surgically implanted materials should be examined in conjunction with the MR facility's safety officer. An x-ray or review of a patient's surgical records may be necessary to evaluate the presence and MRI compatibility of internal materials. The patient should be encouraged more than once to read the safety form carefully and ask questions about any uncertainties, even if the patient thinks that it is probably safe. It is best for a patient to ask about something that poses no risk than to keep quiet out of embarrassment and die.

Jewelry is a major source of questions and sometimes of resistance. Although patients often refuse, they should not enter the magnet room without first removing piercing jewelry, such as tongue rings, nose rings, belly button rings, and earrings. They may claim that a favorite necklace is pure gold, which is nonmagnetic, only to have the MR machine demonstrate the opposite. In some cases, an inadequately anchored piece, such as the backing of an earring, can come loose and pose a flying hazard or become lodged in a joint somewhere in magnet cowell or gradient coil assembly. A more common problem is that the small quantities of ferromagnetic materials distort or, in some cases, generate huge holes to appear in the MRI, even when the patient clearly has an intact head. Safety is enhanced if a blanket ban on jewelry is followed.

A significant number of subjects in the general population are likely to show abnormalities. When these are unexpected, they are called *incidentalomas,* and the proper course of action upon finding an incidentaloma has been and continues to be a subject of debate. Observations that are of potential clinical significance should be referred to a radiologist for evaluation and decisions about further courses of action. For example, aneurysms pose cause for further evaluation. Some conditions, such as an asymptomatic meningioma in an 85-year-old woman, might be best monitored but untreated unless symptoms later emerge. Each case must be evaluated on its own circumstances and according the policies of each institution.

FUTURE DIRECTIONS

Today, established MR technologies provide a limited but valuable clinical service to psychiatry through their ability to rule out gross structural abnormalities. Presently, most of the methodologies described in this chapter are used to acquire data in academic research settings in efforts to help psychiatrists understand the pathophysiology, pathogenesis, and treatment of psychiatric disorders.

The primary limitations that face MR today are spatial resolution,

sensitivity, and cost. Regarding spatial resolution, MRI data in humans are rarely acquired with grid sizes finer than 0.5 to 0.75 μL and more typically 1.5 to 2 μL. The sensitivity of MRS is typically far less than that of MRI, because the concentrations of the metabolites are far less than those of the water and fat that are usually detected for MRI, with a general limitation of 0.5 to 1 mmol/kg as the detection limit in the human brain in vivo, with volumes in some cases as small as 0.25 mL but more typically 8 mL for ^{1}H MRS and larger volumes for ^{31}P and ^{13}C. Because of the need to detect lower concentrations, MRSI fares more poorly than MRI with regard to resolution and/or sensitivity, but the ability to map chemicals through the brain provides a unique power to investigate mental illnesses. Although whole-brain ^{1}H MRS can be implemented, much of the brain is poorly shimmed and yields no useful data. Methods that greatly improve problematic shimming are also under development and may resolve this problem in the coming years. fMRI receives great attention in the psychiatric research community today, and a major limitation that must be kept in mind when interpreting the imaging results is that the measurement is relative, and any function-related changes must be evaluated in the context of possible baseline conditions. Cost poses an impediment to increases of diagnostic approaches with new MR methodologies. With scans now costing $500 to $4,000, depending on the procedures ordered and the institution where the scan is done, additional methods will only gain widespread acceptance if they are clearly shown to add clinically significant information.

The future of MR can be viewed from several perspectives. The noninvasive, nonradioactive nature of MR provides a strong impetus for continued development in the diagnostic evaluation of patients and use in basic science research. By virtue of its safety profile, MR is likely to support research on childhood manifestations of psychiatric disorders. Magnetic field strengths for research scanners are increasing, having reached 9.4 T for human use; higher magnetic field strengths for most brain studies represent greater sensitivity. As manufacturers gain experience with higher-field research machines, the technology has in the past lowered the cost of medium-field magnets and led to the improvements in reliability and stability. A second potential advance lies with the elimination of helium as the coolant for the superconducting wire in the magnet. Helium currently poses a significant cost of installation and operation of MR scanners. The cost is rising with increasing limitations on worldwide helium supplies, and some industry sources predict exhaustion of the world supply by the year 2030. Researchers around the world are working to develop superconductors with the necessary physical flexibility and ability to carry large electrical currents at liquid nitrogen temperatures and also more powerful permanent magnets (not produced by electric currents); such breakthroughs should lower MR operating costs and eliminate the

need for the limited helium supply. Another area of intense investigation is hyperpolarization of ^{13}C, increasing its sensitivity by several orders of magnitude. Researchers are actively pursuing the use of hyperpolarized ^{13}C for use as nontoxic contrast agents for angiography and metabolism. A very different approach, *molecular MRI,* is increasing the ability to monitor processes such as the migration of nerve cells, gene expression, and targeting of particular cell types. Such molecular imaging abilities remain in the realm of research on animals, but as methods continue to develop, they appear likely to find their way to applications to human diseases.

In the future, MR is likely to find increasing utility as a research tool, because it provides a quantitative window on neuroanatomy, function, and neurochemistry. Developers are also working to turn the research methodologies into robust clinical techniques for the testing and use of biomarkers to monitor treatment efficacies. MR studies are becoming increasingly multimodal, with various types of MRI and MRS acquired in single scan sessions and on individual subjects, as well as in complementary measurements with methods like PET and EEG recordings. As outlined above, MR has already begun to change the current thinking with respect to the pathophysiology of psychiatric disorders, neuroenergetics, amino acid neurotransmission, and plasticity. From most perspectives, MR appears likely to continue to do so in the future.

REFERENCES

Abdallah CG, Jiang L, De Feyter HM, et al. Glutamate metabolism in major depressive disorder. *Am J Psychiatry*. 2014;171:1320–1327.

Ai T, Morelli JN, Hu X, et al. A historical overview of magnetic resonance imaging, focusing on technological innovations. *Invest Radiol*. 2012;47:725–741.

Attwell D, Laughlin SB. An energy budget for signaling in the grey matter of the brain. *J Cereb Blood Flow Metab*. 2001;21:1133.

Brown GG, Eyler LT. Methodological and conceptual issues in functional magnetic resonance imaging: applications to schizophrenia research. *Annu Rev Clin Psychol*. 2006;2:51.

Carlson PJ, Singh JB, Zarate CA Jr., Drevets WC, Manji HK. Neural circuitry and neuroplasticity in mood disorders: insights for novel therapeutic targets. *NeuroRx*. 2006;3:22.

Chang L, Friedman J, Ernst T, Zhong K, Tsopelas ND. Brain metabolite abnormalities in the white matter of elderly schizophrenic subjects: implication for glial dysfunction. *Biol Psychiatry*. 2007;62:1396.

Drevets WC, Price JL, Simpson JR, Todd RD, Reich T. Subgenual prefrontal cortex abnormalities in mood disorder. *Nature*. 1997;386:824.

Du F, Cooper AJ, Thida T, et al. In vivo evidence for cerebral bioenergetic abnormalities in schizophrenia measured using ^{31}P magnetization transfer spectroscopy. *JAMA Psychiatry*. 2014;71:19–27.

Dukart J, Mueller K, Barthel H, Villringer A, Sabri O, Schroeter ML, for the Alzheimer's Disease Neuroimaging Initiative. Meta-analysis based SVM classification enables accurate detection of

Alzheimer's disease across different clinical centers using FDG-PET and MRI. *Psychiatry Res Neuroimag*. 2013;212:230–236.

Feldstein Ewing SW, Sakhardande A, Blakemore SJ. The effect of alcohol consumption on the adolescent brain: a systematic review of MRI and fMRI studies of alcohol-using youth. *Neuroimage: Clinical*. 2014;5:420–437.

Goghari VM, Rehm K, Carter CS, MacDonald AW III. Regionally specific cortical thinning and gray matter abnormalities in the healthy relatives of schizophrenia patients. *Cereb Cortex*. 2007;17:415.

Hanson LG. Is Quantum Mechanics Necessary for Understanding Magnetic Resonance? Concepts in Magnetic Resonance Part A. 2008;32:329–340.

Jiang L, Gulanski BI, de Feyter Weinzimer S, et al. Increased brain uptake and oxidation of acetate in heavy drinkers. *J Clin Inv*. 2013;123:1605–1614.

Kegeles LS, Mao X, Stanford AD, et al. Elevated prefrontal cortex γ-aminobutyric acid and glutamate-glutamine levels in schizophrenia measured in vivo with proton magnetic resonance spectroscopy. *Arch Gen Psychiatry*. 2012;69:449–459.

Kempton MJ, Salvador Z, Munafò MR, et al. Structural neuroimaging studies in major depressive disorder. Meta-analysis and comparison with bipolar disorder. *Arch Gen Psychiatry*. 2011;68:675–690.

Kroenke CD, Rohlfing T, Park B, Sullivan EV, Pfefferbaum A, Kathleen Grant KA. Monkeys that voluntarily and chronically drink alcohol damage their brains: a longitudinal MRI study. *Neuropsychopharm*. 2014;39:823–830.

Kugaya A, Sanacora G. Beyond monoamines: glutamatergic function in mood disorders. *CNS Spectr*. 2005;10:808.

Lee MH, Smyser CD, Shimony JS. Resting-state fMRI: a review of methods and clinical applications. *Am J Neurorad*. 2013;34:1866–1872.

Mahon PB, Lee DS, Trinh H, et al. Morphometry of the amygdala in schizophrenia and psychotic bipolar disorder. *Schizophrenia Res*. 2015; in press.

Minati L, Grisoli M, Bruzzone MG. MR spectroscopy, functional MRI, and diffusion-tensor imaging in the aging brain: a conceptual review. *J Geriatr Psychiatry Neurol*. 2007;20:3.

Niciu MJ, Mason GF. Neuroimaging in alcohol and drug dependence. *Curr Behav Neurosci Reports*. 2014;1:45–54.

Ogawa S, Menon RS, Tank DW, Kim SG, Merkle H. Functional brain mapping by blood oxygenation level-dependent contrast magnetic-resonance imaging—A comparison of signal characteristics with a biophysical model. *Biophys J*. 1993;64:803.

Pfefferbaum A, Rosenbloom MJ, Chu W, et al. White matter microstructural recovery with abstinence and decline with relapse in alcohol dependence interacts with normal ageing: a controlled longitudinal DTI study. *Lancet Psychiatry*. 2014;1:202–212.

Redlich R, Almeida JJ, Grotegerd D, et al. Brain morphometric biomarkers distinguishing unipolar and bipolar depression. A voxel-based morphometry-pattern classification approach. *JAMA Psychiatry*. 2014;71:1222–1230.

Sheline YI. Neuroimaging studies of mood disorder effects on the brain. *Biol Psychiatry*. 2003;54:338.

Urenjak J, Williams SR, Gadian DG, Noble M. Specific expression of N-acetylaspartate in

neurons, oligodendrocytes in vitro. *J Neurochem*. 1992;59:55.

Wang J, Du H, Jiang L, Ma X, de Graaf R, Behar K, et al. Oxidation and effects of ethanol in the rat brain after chronic ethanol exposure. *Proc Natl Acad Sci USA*. 2013;110:14444–14449.

Wang L, Hermens DF, Hickie IB, Lagopoulos J. A systematic review of resting-state functional-MRI studies in major depression. *J Affect Disord*. 2012;142:6–12.

Zahr NM. Structural and microstructural imaging of the brain in alcohol use disorders. In: Sullivan EV, Pfefferbaum A, eds. *Handbook of Clinical Neurology: Alcohol and the Nervous System*. Vol 125:257–290.

Zhang B, Ferman TJ, Boeve BF, et al. MRS in mild cognitive impairment: early differentiation of dementia with Lewy bodies and Alzheimer's disease. *J Neuroimag*. 2015;25:269–274.

▲ 1.17 Radiotracer Imaging with Positron Emission Tomography and Single Photon Emission Computed Tomography: Fundamental Principles, Methodology, and Role in Neuropsychiatric Research

MARK SLIFSTEIN, PH.D., JOHN H. KRYSTAL, M.D., AND ANISSA ABI-DARGHAM, M.D.

OVERVIEW

Radiotracer imaging is a methodology with the unique ability to quantify spatially localized neurochemical information in vivo in the brains of humans and preclinical species. Radiotracers are chemical probes designed to interact with neurochemical systems, either by binding to target molecules such as neuroreceptors, reuptake transporters or intracellular enzymes, or by acting as substrates for the metabolic pathways of endogenous substances in parallel with them. "Tracer" refers to an experimental approach in which a chemical substance containing a quantifiable reporter probe is introduced into a biological system and its interaction with the system is "traced" by the probe. It is implicit that the tracer will be introduced in small enough concentration so that the normal function of the system is not noticeably perturbed by the presence of the tracer itself. Tracer methods were in use for decades before the advent of modern imaging methods and technology. George De Hevesy pioneered the use of radioactive isotopes incorporated in tracers in the mid-20th century to study a wide range of chemical phenomena, including biological systems. Other investigators developed pharmacokinetic models based on tracer methods to quantify processes such as the perfusion of blood in biological tissues, that form the foundation of the pharmacokinetic models used in modern radiotracer imaging, but acquisition of experimental data frequently involved invasive measurements in animals. Approaches that were noninvasive enough to use routinely in living humans did not become available until the late 1970s and 1980s when technological advances in imaging systems and computational power made such methods feasible. In radiotracers, the reporter probe is a radioactive atomic isotope incorporated in a molecule that is designed to interact with a (neuro)biological system. Radioactive isotopes have nuclear configurations that are unstable; individual radioactive atoms decay to a more stable state at unpredictably random times, but when a large number of radioactive atoms are present, the rate of decay at any time is a very well-characterized fraction of the initial rate. There are many different schemes by which radioactive isotopes decay, but all involve the release of energy. Radioactive isotopes used in imaging release, when they decay, high-energy photons (also called γ rays) that readily pass through biological tissue and are then detected in imaging systems external to the body. Photon energy is usually

described in units of electron volts (eV, 1 eV = 1.6×10^{-19} Joules, the energy change that an electron would experience when passed through a 1 volt potential difference. 1 keV = 10^3 eV, 1 MeV = 10^6 eV), and photons released by isotopes used in imaging range from ~150 keV for some isotopes to 511 keV for others; most materials other than very dense ones, such as lead or the inorganic scintillating crystals used as detectors in imaging systems, are transparent to photons in this energy range. Radioactive atoms are incorporated into a very small portion of the tracer molecules; these are termed labeled molecules, and are assumed to be well mixed within the tracer sample, so that at any given time the rate of decay measured in a location within the tissue will have a known proportionality to the number of unlabeled tracer molecules there. By counting the number of photons that originate in a local volume within the brain over an interval of time, the concentration of the radiotracer in that volume can be inferred. The ability to spatially localize the source of radioactivity requires the use of computational methods known as tomographic reconstruction. The mathematical theory behind tomography was already well established in the early part of the 20th century, but was only practically implemented in the final decades of the century with the advent of fast, powerful computers in combination with the development of electronics and detector materials that could provide the necessary spatial, temporal and energy resolution. There are two imaging modalities that utilize similar experimental methods, but differ in the way photons are counted and the mode of decay of the radioactive isotopes used. These are PET and SPECT. The term *emission* refers to the fact that the radioactive isotope is part of a molecule that is injected intravenously (or, less frequently, inhaled or ingested) so that the detected photons are emitted from within the body. This is in contrast to transmission imaging such as x-ray and computed tomography (CT) scans in which a beam of photons is directed at the body from an external source and the portion that pass through without interacting with the intervening tissues are counted on the opposite side. Psychiatry researchers have used PET and SPECT imaging to characterize many aspects of neurochemistry in sample cohorts of subjects with many different neuropsychiatric conditions. While the number of radioactive isotopes used in imaging is small, the number of radiotracers that these can be incorporated into, and biochemical systems that can be studied, are virtually limitless, and many different tracers have been used in research studies with human subjects. Currently, in psychiatry these imaging methods are used almost exclusively in research. They do not yet have the sensitivity or specificity to be used diagnostically. Most radiotracer studies in psychiatric populations have based their conclusions on statistical inference from moderately large sample cohorts in which there is considerable overlap in the measured quantities between patients and comparison subjects. But these studies provide important clues to other

investigators working in basic neuroscience or drug development about processes that may be pathological in psychiatric conditions and can be examined more closely using in vitro methods or in preclinical species, or that might suggest molecular targets for pharmacotherapy. In this sense, radiotracer imaging is an important translational tool for understanding brain neurochemistry in psychiatric patient populations. In the remainder of this chapter, the physics and technology of imaging, pharmacokinetic models, and experimental designs used in psychiatric imaging research will be described, followed by a survey of the radiotracers, molecular targets, psychiatric populations that have been studied, and some well-replicated results.

BASIC PHYSICAL PRINCIPLES

Radioactivity, Radionuclides, and Photon Emission

Half-Life. As stated above, the time at which any individual radioactive atom will undergo decay is random and unpredictable. However, radiotracer doses contain a very large number of labeled molecules. For example, a typical PET dose contains on the order of 10^{10} labeled molecules, each of which has the same probability of decay per unit time, independently of all the others. In this setting, the rate, or number of decays per unit time follows an exponential formula, $A(t) = A(0)e^{-\lambda t}$ where $A(t)$, the activity, is the number of decays per unit time at time t, $A(0)$ is the initial activity and λ is a very well-characterized rate constant that is different for each radioactive isotope. The Système International (SI) unit of radioactivity is the Bequerel (Bq), equal to 1 decay per second. An older, though still frequently used unit is the Curie, equal to 3.7×10^{10} decays per second. For the range of activity used in imaging, it is more convenient to use units of MBq (10^6 Bq) and mCi (10^{-3} Ci). It is also convenient to factor λ into $\ln(2)/T_{1/2}$, where $T_{1/2}$, the half-life, is the time it takes for A to decrease to half its initial value. The choice of time and the total activity when the initial measurement $A(0)$ is made does not affect the half-life; following any measurement, the activity will decrease by 50 percent after an additional $T_{1/2}$ has passed. Half-lives and other parameters for the isotopes most commonly used in PET and SPECT are given in Table 1.17–1. The isotopes used in radiotracer imaging are also referred to as radionuclides.

Decay by Positron Emission and PET Radionuclides. In decay by positron emission, a proton within the atomic nucleus is converted into a neutron and a positron (and an extremely low mass, uncharged particle called a neutrino). A positron is a particle with the same rest mass as an electron but opposite electric charge. When the positron is ejected from the nucleus, it will have an initial kinetic energy that will be reduced as the particle

interacts with the surrounding matter. The sequence of collisions with nearby matter will cause the positron to follow an unpredictable path with many changes of direction and it will usually come to rest within nanoseconds of its release and within 1 to 2 mm of its source atom for the most commonly used PET radionuclides (Table 1.17–1). At that point, the positron and a local electron will be attracted to each other by their opposite charges and undergo an annihilation reaction, in which their masses are converted to energy. Two photons will be produced, each having the energy of the rest mass, 511 keV, of an electron, as the energy from the rest masses of the positron and electron is conserved. The photons travel away from each other in directions that are very nearly opposite to each other (the total momentum is conserved: the momenta of the positron and the electron are nearly zero at the time of the annihilation, and the momenta of the photons are opposite and sum to zero, although the angle is usually slightly different than 180 degrees due to a small amount of residual momentum in the particles at the time of the annihilation event) (Fig. 1.17–1). PET technology exploits the fact that two nearly collinear photons are emitted simultaneously from the location of the annihilation event by utilizing the principle of coincidence detection (see below). The two radionuclides most commonly used in PET brain imaging are ^{11}Carbon (^{11}C, $T_{1/2} \cong 20$ minutes) and ^{18}Fluorine (^{18}F, $T_{1/2} \cong 110$ minutes). ^{15}Oxygen (^{15}O, $T_{1/2} \cong 2$ minutes), incorporated in ^{15}O-labeled water, can be used to assess CBF, but its use for this purpose has been mostly superseded by MRI techniques. Also ^{13}Nitrogen (^{13}N, $T_{1/2} \cong 10$ minutes), incorporated in ^{13}N-labeled ammonia, is used frequently in other areas such as cardiology. Other positron emitters such as ^{64}Copper (^{64}Cu, $T_{1/2} \cong 12.7$ hours) have been used as radionuclides in settings where longer half-lives are desirable, but to date ^{64}Cu has not been applied in neuroimaging. The radionuclides used in PET are not naturally occurring, and their short half-lives place strict constraints on how long before the study, and therefore how far away from the study site, they can be produced and incorporated into radiotracers. Production of PET radionuclides is performed in a particle accelerator called a cyclotron, an instrument that is produced by several medical device manufacturers, but is large, expensive, requires a high level of skill to operate and maintain, and is subject to regulatory scrutiny because it produces radioactivity. Rapid incorporation of the radionuclide in a radiotracer requires elaborate chemistry and quality control instruments and highly skilled chemists to perform these tasks. An approximate rule of thumb is that a radiotracer must be able to be synthesized and undergo tests for sterility and chemical purity in an hour or less to be suitable as a ^{11}C-labeled radiotracer. The use of ^{11}C is therefore restricted to sites that can produce it locally at the time of the PET scan, and PET studies using ^{11}C radiotracers are performed almost exclusively at

large academic medical centers that can support the required staffing and equipment. While the requirements for producing ^{11}C-labeled radiotracers are clearly demanding, there are some major advantages to ^{11}C that makes its use very desirable. First, the ubiquity of carbon in virtually all molecules of interest as tracers means that labeling can potentially be performed by substituting ^{11}C for ^{12}C in an organic molecule without altering its biochemical properties. Second, the 20-minute half-life allows for multiple scans on the same day so that scanning can be performed under different experimental conditions on a subject in a single study session. Academic medical centers also routinely produce ^{18}F-labeled radiotracers. There is clearly less time pressure in the production of ^{18}F-labeled tracers than with ^{11}C. Additionally, the 110-minute half-life of ^{18}F allows enough of a time window for ^{18}F tracers to be produced by specialized manufacturers with facilities located within several hours traveling time from the study site. Furthermore, ^{18}F labeling is preferable for cases where the pharmacokinetics of the radiotracer require scans of several hours duration, as in this case, ^{18}F will continue to provide a reliable signal for the entire scan duration, whereas ^{11}C activity would become very low leading to unreliable data late in the scan. However, multiple scans cannot be done on the same day in a single subject, due to the confound that would result from activity injected in the first scan being detected in later scans.

Table 1.17–1.
Characteristics of Commonly Used PET and SPECT Radionuclides

Nuclide	$T_{1/2}$ (minutes)	Imaging Photon Energy (keV)	Usual Imaging Modality	Mean Positron Range (mm)
^{11}C	20.4	511	PET	1.1
^{18}F	109.8	511	PET	0.6
^{15}O	2.03	511	PET	2.4
^{123}I	787.8	159	SPECT	
^{99m}Tc	360.4	140	SPECT	

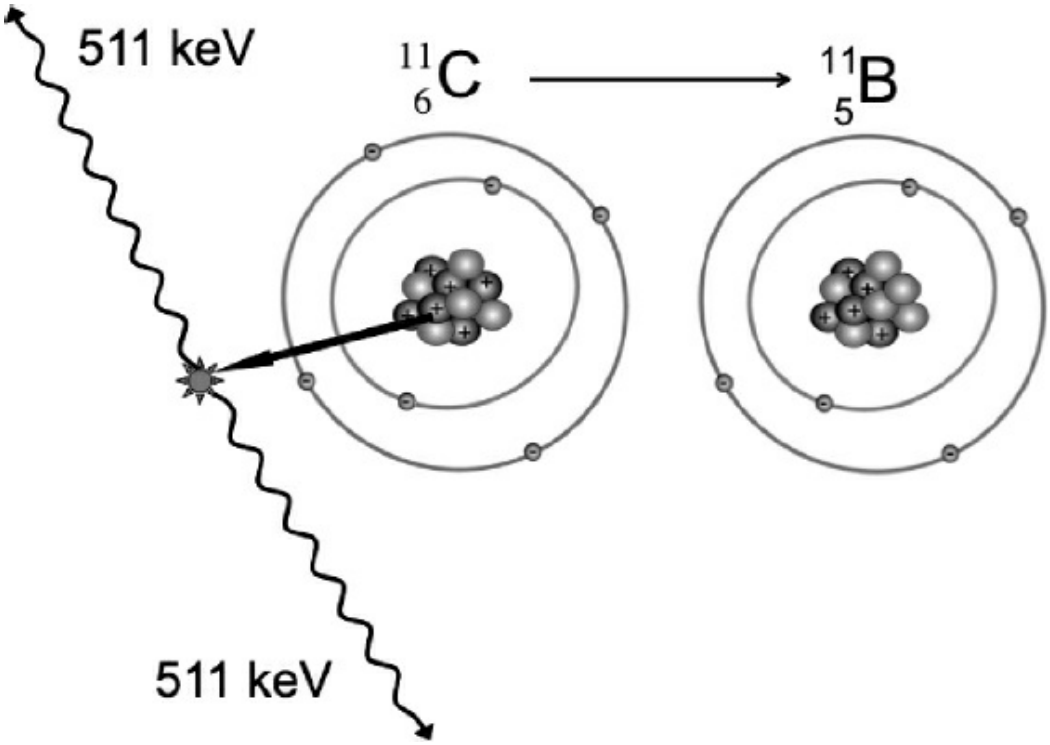

FIGURE 1.17–1. Schematic representation of the decay of unstable carbon 11 to stable boron. A positron is ejected from the carbon nucleus, eventually colliding with an electron. The pair annihilate and two 511 keV photons are released in opposite directions. Distances are not drawn to scale.

SPECT Radionuclides. As the name implies, SPECT scanners measure single emitted photons, unlike PET scanners that utilize coincidence detection. In theory, PET radionuclides could be used in SPECT systems, but other radionuclides that also emit photons in a usable energy range and decay with longer half-lives are more practical. A SPECT radionuclide that has proven to be particularly useful for brain imaging is ^{123}Iodine (^{123}I, $T_{1/2} \cong 13.1$ hours). ^{123}I decays by the process of electron capture, in which an orbital electron combines with a proton to form a neutron (plus a neutrino) and a characteristic x-ray photon is emitted when another, outer orbital electron, fills the orbital shell vacated by the transformed electron. There are multiple electron capture decay schemes that ^{123}I can undergo, leading to release of photons of different energies, but approximately 83 percent of decays result in the emission of 159 keV photons, and these are the ones detected by SPECT scanners. ^{123}I can be produced in a cyclotron, but the long half-life allows for ^{123}I or ^{123}I-labeled radiotracers to be distributed within a several thousand mile radius of the manufacturing site. Across all SPECT applications, the most commonly used radionuclide is ^{99m}Technicium (^{99m}Tc, $T_{1/2} \cong 6$ hours). ^{99m}Tc is a "metastable" product of the decay of ^{99}Molybdenum (^{99}Mo, $T_{1/2} \cong 66$ hours), meaning that its nucleus is in a moderately excited state that will decay to the more stable ^{99}Tc form without transmuting to a different element. ^{99m}Tc can decay to

stable ^{99}Tc by several processes, but approximately 88 percent of decays result in 140 keV photon release, and these photons are detected by SPECT scanners. ^{99m}Tc is most commonly distributed to imaging sites in ^{99}Mo generators, portable devices containing ^{99}Mo produced in large nuclear reactors. The ^{99}Mo decays to ^{99m}Tc in the generator and the ^{99m}Tc can be separated from ^{99}Mo by chromatographic methods. ^{99m}Tc, however, does not readily cross the BBB, whereas ^{123}I does so more easily due to the more lipophilic nature of iodine. As a result, the majority of SPECT radiotracers for brain imaging are ^{123}I-labeled.

Physical Principles of Emission Tomography

Gamma Ray Detection. The same tissue penetrating property that makes the 511 keV photons emitted in PET and the 140 to 159 keV photons emitted in SPECT useful for imaging leads to the necessity for special materials to detect and transduce their energy into a form more amenable to processing by the electronic systems that count them. Detectors in PET and SPECT are dense inorganic scintillating crystals that interact with, that is, stop, the high-energy photons resulting from radioactive decay and emit a portion of the absorbed energy in the form of lower energy photons in the visible to UV range. The most commonly used scintillator in SPECT is thallium-doped sodium iodide (NaI(Tl)). Many scintillating detectors have been used in PET; some of the most frequently used ones are bismuth germinate (BGO), lutetium oxyorthosilicate (LSO), and lutetium yttrium orthosilicate (LYSO). Each has different properties that are more optimal for some applications than others.

After the scintillator releases its lower energy photons, the next stage in signal transduction for both PET and SPECT is to convert that signal into an electrical current that represents the detected event in the electronic circuits that perform the remainder of the activity counting process. This component is called the photosensor, and the most commonly used photosensor is the photomultiplier tube (PMT). PMTs are evacuated glass tubes with a light-sensitive material on the end that interfaces with the scintillator (the cathode) and a sequence of metal plates (dynodes) along the length of the tube that are maintained at successively higher voltages. Light striking the cathode will induce electrons to leave the surface and strike the nearest dynode, inducing more electrons to break free from that plate and strike the next one. The cascade continues, with enough amplification of the electron stream by the last stage (the anode) to form a current in the range that is suitable for the next stage of the electronics (Fig. 1.17–2). Common configurations for detectors and PMTs in PET are rectangular blocks of detector material attached to a small number of PMTs. The blocks are assembled into polygonal rings that surround the imaging field of view (Fig. 1.17–3), and a number of rings are assembled concentrically to form a cylindrical imaging gantry. Similar designs are

feasible in SPECT, but more commonly, SPECT scanners contain several large slabs of detector materials attached to large arrays of PMT that rotate around the subject to obtain data from all angles (Fig. 1.17–3). Recently, there has been a great deal of interest in developing solid-state photosensors to replace PMTs, driven largely by the development of hybrid PET-MR scanners (see below). PMTs interact with, and do not operate properly in, the strong magnetic fields generated by MRI scanners, whereas solid-state photosensors can operate in that environment. Two designs that have been studied extensively are called avalanche photodiodes (APMs) and APMs operating in Geiger mode (referred to as silicon photomultipliers, SiPMs). APMs are used in the PET-MR hybrid scanners currently produced by commercial manufacturers. As the technology of APMs and SiPMs becomes more standardized and less expensive, they will have the potential to greatly impact the design and capabilities of future scanners.

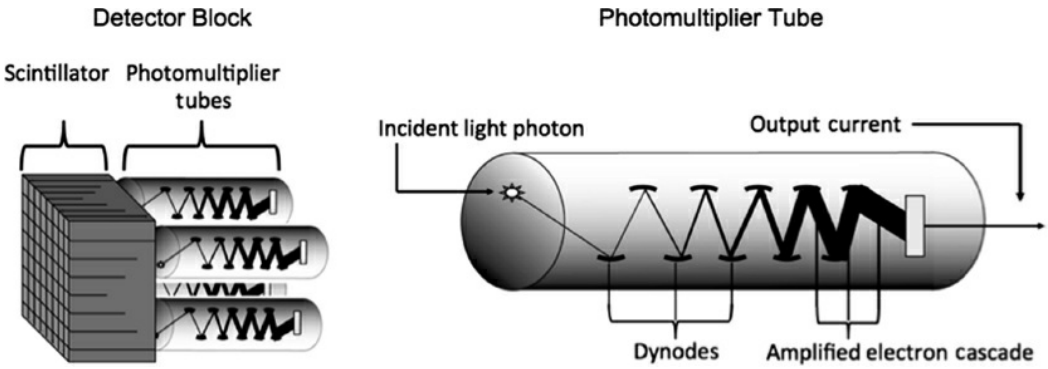

FIGURE 1.17–2. Representation of a detector block with scintillator crystal attached to an array of photomultiplier tubes (**left**) and a photomultiplier tube (**right**).

Coincidence Detection in PET. In PET, radioactive decay events are recorded when two 511 keV photons strike two detectors simultaneously. The path between the two detectors is referred to as a line of response (LOR). An electronic circuit detects if two photon arrival events occur within a very narrow time interval called a coincidence timing window. The coincidence timing windows on modern scanners are in the range of 3 to 6 nanoseconds. Each detected coincidence event causes a computer to increment the count total for that LOR by one unit (Fig. 1.17–3).

Photon Detection in SPECT by Collimation. In SPECT, decay events are localized to an LOR by use of lead collimators. These are usually formed from blocks of lead that are attached to the detectors and have long narrow channels that only allow photons from a narrow set of angles with respect to the detector to pass through. The simplest design utilizes parallel cylindrical channels, all of which are perpendicular to the flat face of the detector, but many different configurations are in use with different

scanner designs and different image reconstruction algorithms. The necessity for collimation in SPECT causes its scanners to utilize a smaller portion of the photons that are emitted from radioactive decay than PET scanners. The overall efficiency with which scanners convert radioactivity in the field of view to detected counts is referred to as sensitivity. Because of collimation, SPECT systems have lower sensitivity than PET systems.

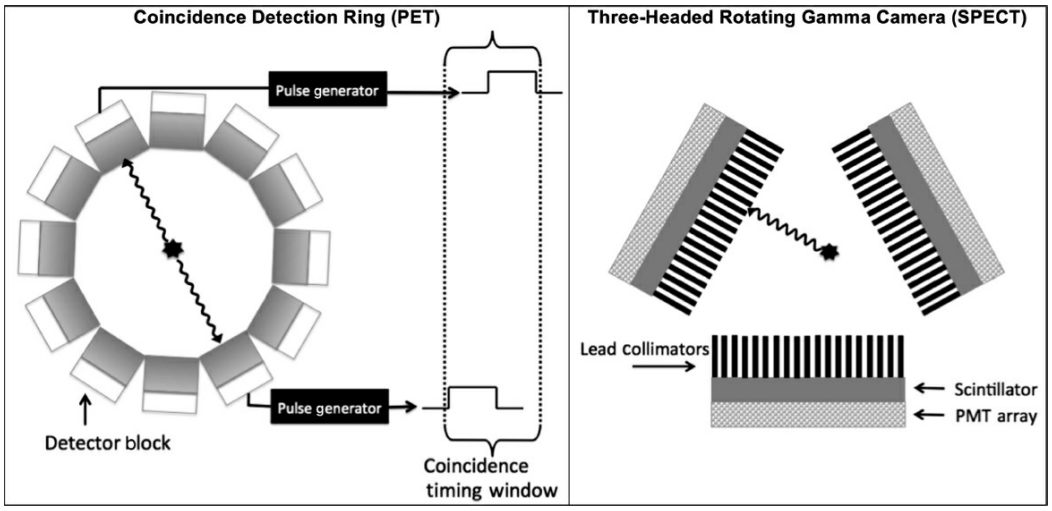

FIGURE 1.17–3. A PET ring detector (**left**) with a representation of a coincidence timing circuit and a 3-headed gamma camera for SPECT (**right**). In PET, an electronic pulse generator sends a signal to a coincidence detection circuit. If two pulses are detected within a coincidence timing window from two detectors, an event is counted on the LOR between them. In SPECT, photon detections are restricted to LORs defined by the collimators. The detector arrays rotate around the imaged object.

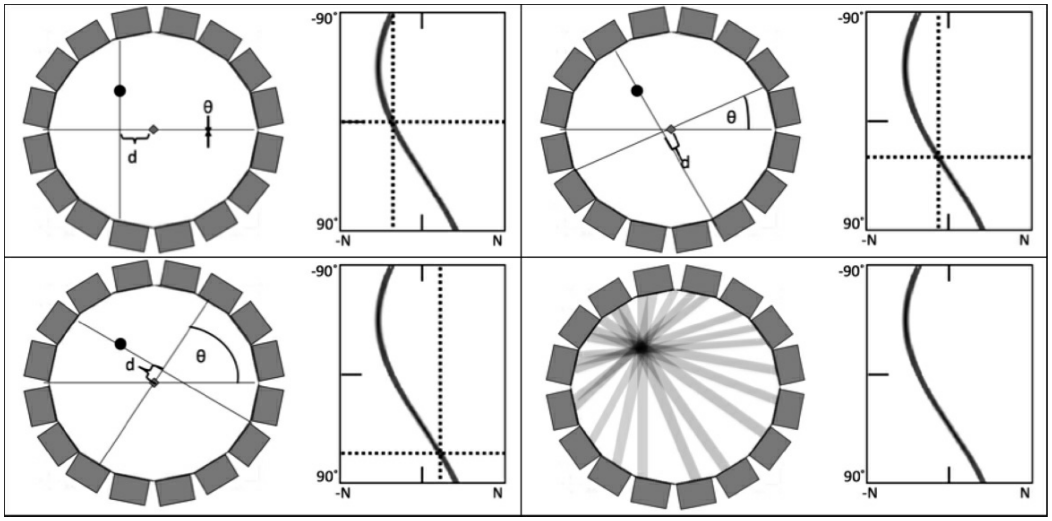

FIGURE 1.17–4. Depiction of backprojection. A point source of activity is placed in the upper left quadrant of the field of view. The top row and bottom left panel depict three different LORs and their locations in the sinogram of the point source. The bottom right panel shows the result of backprojecting all LORs; the only ones that are nonzero are those that pass through the location of the point source.

Image Reconstruction

The data obtained with both PET and SPECT are called projection data. For each LOR, a single number is recorded that contains counts from all loci along the LOR. Conversely, as the direction in which photons are released following a decay event is random, each location in the imaging field of view that contains activity will contribute counts to several, possibly many, LORs. The LOR data contain enough information to reconstruct a facsimile of the original distribution of decay events, but this requires the application of mathematical models of the projection data and numerical implementations of the models on computers. There are two general approaches to reconstruction: filtered backprojection (FBP) and iterative reconstruction. Both involve some degree of mathematical complexity, and there are many variants of each. Below, only the general concepts of both methods are described, not specific implementations. The projection data are stored as a numerical array called a sinogram, which is organized by angle and radial distance from the origin (the center of the field of view). For each angle θ and distance *d,* the value recorded in the sinogram at the point (θ, *d*) is the number of detected events (the counts) in the LOR that is perpendicular to the line through the origin at angle θ and passes through that line at distance *d* from the origin (Fig. 1.17–4). The name sinogram derives from the fact that an object with radioactivity only at a single spatial point would be recorded as a sine wave in the (θ, *d*) representation. The reconstructed image is usually represented with Cartesian coordinates in a rectangular grid of pixels (in two dimensions) or voxels (for "volume element" in three dimensions).

Filtered Backprojection. FBP, sometimes called analytic reconstruction, was the most common approach to image reconstruction from the inception of CT up until the early 2000s, partly because the iterative methods, which appeared in several published papers in the 1980s, were too computationally intensive, and therefore too slow for the computers of that time period. Iterative reconstruction programs, which lead to visually clearer reconstructed images, are now routinely provided as part of the software package with commercially produced scanners, but many researchers still prefer FBP reconstruction due to statistical issues described below. The basic equations of FBP are derived from mathematical formulas called the Radon transform and the inverse Radon transform, theoretical constructs first described in the early part of the 20th century explaining the relationship between a mathematical function defined in two dimensional space, that is, $f(x,y)$ which defines a single numerical value for each (x,y) pair of Cartesian coordinates, and its projections. FBP is a discretized version of the inverse Radon transform; that is, a version based on a finite number of LORs rather than a full continuum of angles and distances from the origin. The first step can be thought of as overlaying the projections from all the LORs and summing

(Fig. 1.17–4). This step is called backprojection. Because the loci that contain radioactivity contribute to all the LORs that pass through them, the backprojected data will have elevated intensities in these positions. However, pure backprojection would also attribute activity along each entire LOR containing counts, so an adjustment called spatial filtering is applied prior to backprojection to remove the starburst-like artifact that would otherwise result (Fig. 1.17–5). This is a conceptual description of FBP; real implementations use more sophisticated computer algorithms and mathematical procedures such as Fourier transforms. The main advantage of FBP is that it is fast. However, there are disadvantages. The filtering step and the fact that there are a finite number of LORs combine to reduce spatial resolution and introduce so-called streak artifacts (Fig. 1.17–6).

Iterative Reconstruction. A number of algorithms have been proposed that treat the relationship between projection data and voxel values probabilistically by framing reconstruction as a statistical estimation problem. These treat the voxel values as a sample drawn from a probability distribution and seek to estimate the most probable distribution of counts in voxels, given the measured projection data and the known physical properties of the imaging system such as the detector geometry. They are called iterative because they start with an arbitrary estimate of the voxel values and then successively improve the estimate by updating a formula that takes the current estimate as input and returns an improved estimate as output. Each cycle through the procedure is called an iteration. The iterative algorithm that is available in the software packages provided by most commercial scanner manufacturers is called ordered subset expectation maximization (OSEM), a modified version of the expectation maximization (EM) algorithm. EM and OSEM are based on the maximum likelihood (ML) principle. ML estimation procedures are statistical methods that establish formulas to estimate the parameters in a parametric probability distribution from observed sample data. For example, the conventional formula for estimating the mean μ of a Gaussian random variable as the sample average can be derived from the ML principle. The Likelihood is the formula for the probability of the observed sample data, expressed as a function of these data and the parameters of the distribution. As the name suggests, ML methods seek the value of the parameter(s) that maximizes the Likelihood of the observed data. Voxel data are usually modeled as Poisson random variables; the Poisson distribution can take on only nonnegative integer values and has a single parameter λ, the mean of the distribution. In the EM and OSEM approaches, counts emanating from each voxel are treated as statistically independent of every other voxel. The mean or expected value of the counts at voxel j is assumed to be λ_j, the parameter of the Poisson distribution for

that voxel. These, however, are hidden variables, because the observed data are the projection data on LORs. The EM algorithm is designed to obtain the ML parameter estimate for the distribution of a hidden variable when only incomplete data, such as the projection data, are available. For the particular case of projection data and the assumed Poisson distribution of voxel values, the procedure leads to a simple and elegant formula that is guaranteed to converge, over the course of successive iterations, to the set $\{\lambda_j\}$ of ML estimates of voxel means. However, the full Poisson EM algorithm requires a huge number of computations, given the large number of LORs and voxels in PET and SPECT image data, and was therefore not practical to apply on the computers available at the time the EM algorithm was proposed. The OSEM algorithm is a modified version introduced in the 1990s that made iterative reconstruction feasible. By cleverly ordering the full set of LOR data into a sequence of smaller subsets and updating the voxel estimates based on computations performed on each successive subset rather than only once after each complete iteration through all the LORs, fewer equations need to be solved and the necessary number of iterations, and therefore the computational load, is reduced significantly. Unlike full EM, OSEM is not guaranteed to converge to the ML estimate of the voxel values, but practical experience has shown that it leads to very satisfactory image reconstructions after a fixed, small number of iterations (Fig. 1.17–6). The main advantage of OSEM reconstruction is that it leads to clearer images without streak artifacts or the level of random noise that is seen in FBP reconstruction (Fig. 1.17–6). A disadvantage is that, following necessary preliminary corrections to the sinogram data prior to reconstruction for several noise sources that occur, especially in PET, the Poisson distribution assumptions are not completely satisfied, and this can lead to bias in which the true proportionality between brain regions with high and low activity is not preserved, although the bias is relatively small. For this reason, many neuroimaging researchers, who are more interested in quantification than visual presentation, prefer FBP to OSEM.

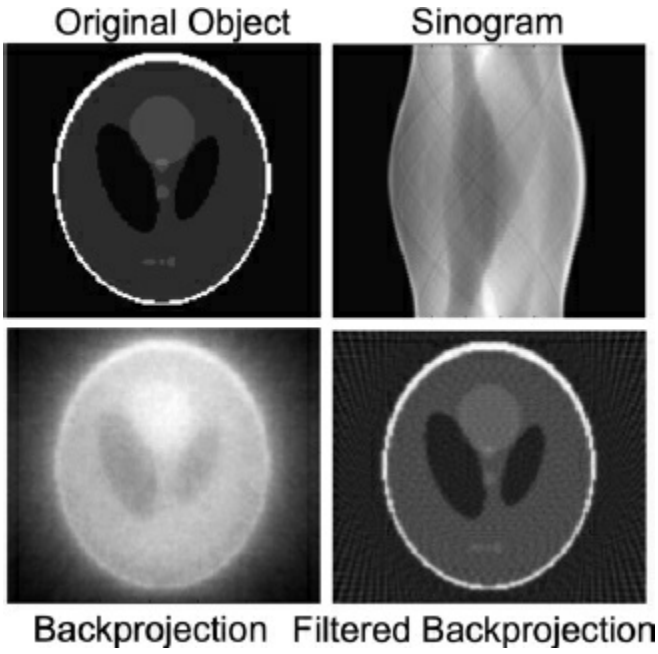

FIGURE 1.17–5. Filtered backprojection. The object is a Shepp-Logan phantom, a set of superimposed ellipses of different intensities used as a standard for comparing reconstruction algorithms. There are several different filters that can be used. The one used here is called a ramp filter.

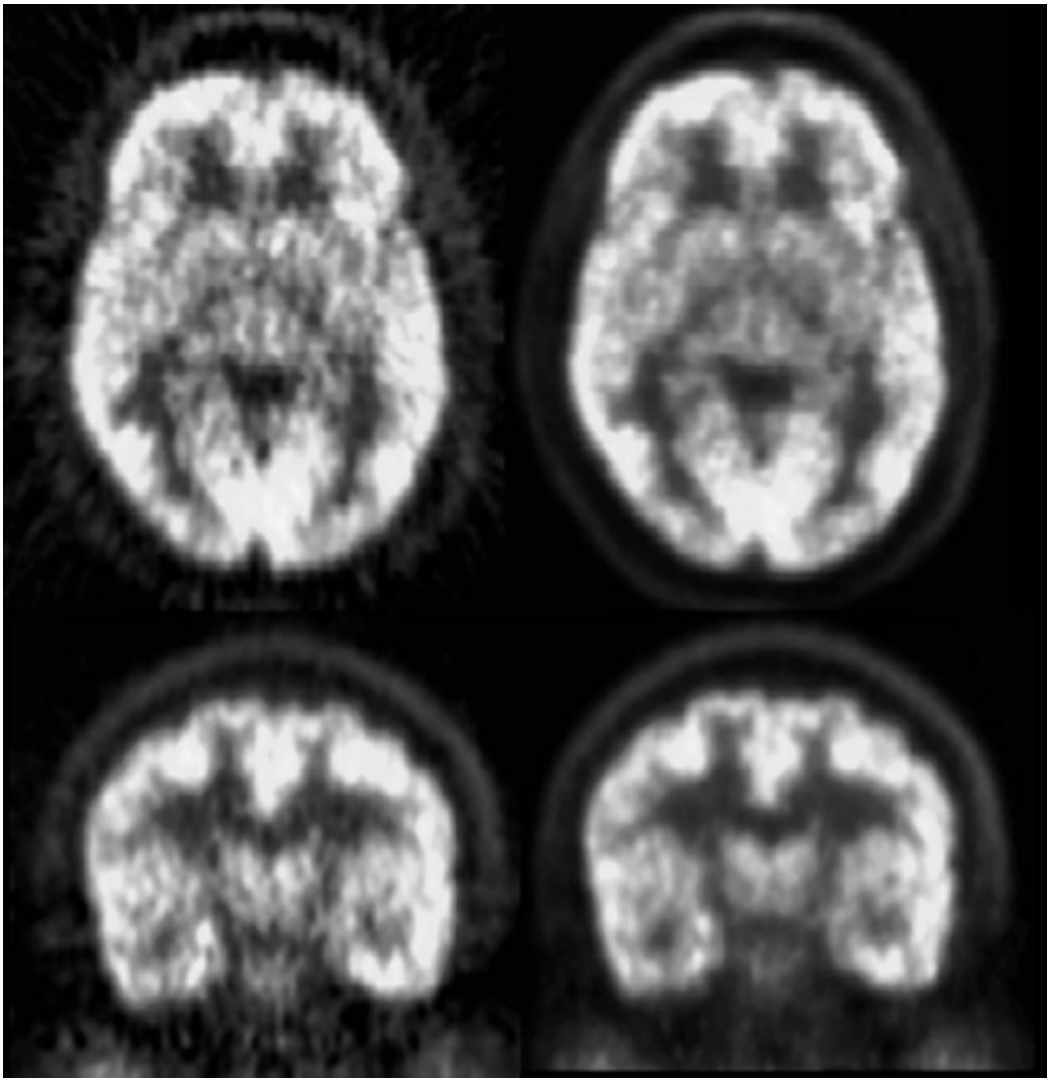

FIGURE 1.17–6. Comparison between FBP (**left**) and OSEM (**right**). The image is summed data from a 120 minute [^{11}C]MDL100907 scan, shown in transverse (**top**) and coronal (**bottom**) views. The radiotracer binds to the serotonin 5-HT_{2A} receptor. The summed image represents a combination of distribution of the tracer throughout the brain during the early part of the scan and specific binding to receptors in the later part. Streak artifacts are visible in the FBP, especially in the space between scalp and brain. The FBP reconstruction also has a more mottled appearance than the OSEM reconstruction.

Factors Affecting Image Quantification and Image Quality

The sections above describing the release and subsequent detection of photons are idealized in the sense that it is assumed that 100 percent of photons pass through biological tissue without any interaction and that 100 percent of photons or coincident pairs striking the detectors are perfectly converted into detected counts. In reality, there is some interaction between photons and tissue, LORs are not really lines but three-dimensional columns in which the probability of detection is greater at the center than toward the edges, and electronics have finite limits to their temporal and energy resolution. These factors and others degrade the overall image quality and spatial resolution; some severely enough to

require corrective procedures to be applied prior to the reconstruction phase.

Photon Attenuation. While many of the photons from radioactive decay do pass through tissue unimpeded, a fraction of them do not. For γ rays in the energy range of those released by PET and SPECT isotopes and matter like biological tissue, the majority of interactions are due to Compton scattering, in which the photon collides with an atom and is deflected from its original path, with some of its energy absorbed by an orbital electron from the atom. Less frequently, the photon is completely absorbed in the collision and the atom ejects an electron, a process called photoelectric effect. The reduction in detected photons due to these interactions is called attenuation. The probability of attenuation decreases with increasing γ ray energy, and increases with the density and thickness of the intervening matter between the photon source and the detector. For a given kind of matter such as soft tissue, the probability that a photon passes through without being attenuated is an exponential function of the depth of the tissue where the photon is released; the probability is proportional to $e^{-\mu d}$ where d is the depth in centimeters and μ is the linear attenuation coefficient of the material in units of cm^{-1}. The attenuation in air is nearly zero, so only the thickness of the tissue is a factor in attenuation, not the total distance to the detectors. The linear coefficient μ is a function of both the energy of the photon and the nature of the material through which the photon is passing; the larger μ is, the greater the attenuation. Because PET photons have higher energy than SPECT photons, they will undergo less attenuation. For example, μ in water for 511 keV photons is 0.097 cm^{-1}, whereas for 150 keV photons it is approximately 0.15 cm^{-1}, and μ in soft tissue is similar to μ in water. Without correction for attenuation, voxels located deep in tissue would have lower detected counts than voxels closer to the surface having the same concentration of radiotracer. Therefore, attenuation correction is a necessary step in quantitative imaging. The fact that PET utilizes coincidence detection allows the implementation of a very accurate form of attenuation correction. This is because the probability of either photon from an annihilation not being attenuated is independent of whether or not the other is attenuated, and therefore the probability of both photons not being attenuated is the product of their probabilities; it is proportional to $e^{-\mu d_1} e^{-\mu d_2} = e^{-\mu(d_1+d_2)}$ that depends only on the total length of tissue, $d_1 + d_2$, traversed by the LOR, and does not depend on the location within the LOR where the annihilation event occurs (Fig. 1.17–7). Attenuation correction in PET is usually applied to the sinogram prior to reconstruction by multiplying the value of each LOR by a correction factor measured in a separate scan called a transmission scan. In dedicated PET systems, this is

usually performed by rotating a point source or a rod containing multiple point sources of a positron emitting material of known activity around the measured object (the subject's head) and recording the fraction of activity detected on the opposite side. Different approaches are required for hybrid PET-CT and PET-MR scanners (see below). In single photon detection as in SPECT, other strategies must be employed. While some sophisticated algorithms for measuring attenuation using a transmission source have been devised for application in SPECT, uniform attenuation correction is most frequently applied. In this approach, an image of an elliptical object with voxel values equal to the correction factors for water is aligned to the reconstructed SPECT image and the values in the two images are multiplied on a voxel by voxel basis.

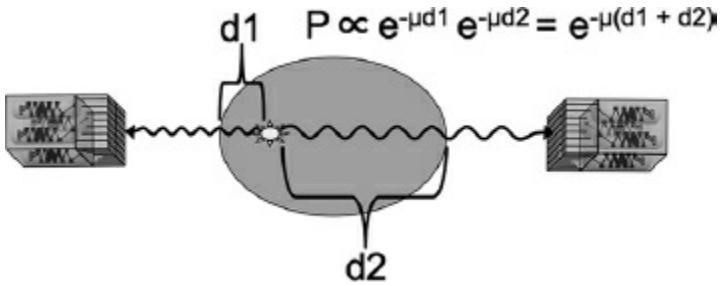

FIGURE 1.17–7. Attenuation. The ellipse in the center represents a hypothetical object (i.e., a head) in the field of view. In SPECT, the photon traveling to the left would have the probability proportional to $e^{-\mu d1}$ of not being attenuated and the one traveling to the right would have probability proportional to $e^{-\mu d2}$ of not being attenuated. In PET, the coincidence event along this LOR would have probability proportional to $e^{-\mu(d1+d2)}$ regardless of where along the LOR it originated.

Photon Scatter. Compton scattered photons are not absorbed in the tissue completely, but rather, change their trajectory and continue at reduced energy. The direction of change is random, and the reduction in energy increases as the angle of deflection increases. In PET imaging, the scattered photon and its partner will arrive at two detectors within a coincidence timing window and therefore the annihilation event will be misattributed to an LOR that does not pass through the true location. In SPECT imaging, the photon will be misread as a photon emanating along an LOR from the location of the scattering event in the scattered direction. In both modalities, detector circuitry can discriminate photon energy and reject low-energy photons to some degree, but there is enough overlap between the range of accepted energies and the energies of scattered photons such that many scattered photons are counted as true decay events. Scatter adds random noise unrelated to the true radiotracer distribution and would degrade the image quality considerably if not corrected. There have been many scatter correction methods proposed and implemented. Modern scanners incorporate sophisticated software algorithms that utilize information from the attenuation map generated

from the transmission image to make probabilistic models of the scatter distribution, which are subtracted from sinogram data prior to reconstruction.

Random Coincidences in PET. An additional source of noise specific to PET is accidental, or random, coincidences. These arise when two photons from distinct annihilation events arrive at two detectors within a coincidence timing window, and are counted as an annihilation event along the LOR connecting the two detectors. Random coincidences are uniformly distributed across the field of view, and cause an artifact of elevated background activity level, roughly in proportion to the square of the singles rate (the rate at which individual photons are detected). As with scatter, random coincidences would considerably degrade the image quality if not addressed. Most scanners have a delayed coincidence timing channel, an electronic circuit that allows estimation of random coincidences by purposely counting pairs of detected photons as coincidence events that arrive several nanoseconds further apart than a coincidence timing window and are therefore known not to have originated from the same annihilation event. The estimated random coincidences are subtracted from sinogram data prior to any other processing. In regions of the field of view where there is no, or very low, true activity, this step can lead to apparent negative counts.

Spatial Resolution. All imaging systems have characteristic spatial resolutions and indices of the smallest object dimensions that they can distinguish. Resolution in PET and SPECT systems is limited by physical characteristics of the scanner, such as the finite width of LORs (as opposed to infinitesimally thin lines) or the fact that projection data are measured at a finite number of angles and distances from the origin at each angle, rather than the full continuum of angle and length values. Resolution limiting factors specific to PET are the range of positrons and the fact that the annihilation events are only localized to within a sphere of 1 to 2 mm radius from the sites of the decay, and that the angle between released photons is not identically 180 degree. Resolution in SPECT is limited by collimator design and distance of the imaged object from the collimators. Resolution of imaging systems is measured empirically, by imaging radioactive sources of standard shapes and characterizing the accuracy of the image. If a small point source or a thin line source of activity is placed in the field of view, the reconstructed image spreads in a bell-shaped curve around the point or line called the point spread function (PSF) or the line spread function (LSF). Resolution is usually reported as the full width at half maximum (FWHM) of the PSF. This is the length, in millimeters, between the two points at which the value of the PSF is half of it maximum, central value (Fig. 1.17–8). The PSF and FWHM are usually not constant

across the field of view; they vary slightly with position, but are typically reported as one, average number. They are also dependent on the method of image reconstruction applied; performance evaluation reports are usually conducted using standardized methods to facilitate comparisons across scanners and reports. Most modern production PET scanners have FWHM in the range of 4 to 5 mm. Some specialized PET scanners have FWHM in the range of 2 to 3 mm. Resolution in SPECT is dependent on the collimator system in use. Clinical, rotating head SPECT scanners have FWHM in the range 10 to 12 mm. Specialized small animal SPECT scanners, that can exploit the close proximity of the animal to the collimator and use special collimator designs, can have submillimeter FWHM. Images are usually reconstructed with voxel dimensions that are smaller than the FWHM to avoid having a checkerboard-like or pixelated appearance. As a result, each voxel will contain overlapping information from the voxels that are adjacent to it.

Hybrid Imaging Systems

In the years since 2000, scanner manufacturers have begun producing multimodal scanners having a PET or SPECT component combined with an additional imaging system such as CT or magnetic resonance (MR). The adaptations necessary to use these systems and their potential to expand the information obtained in the context of neuroimaging are briefly described.

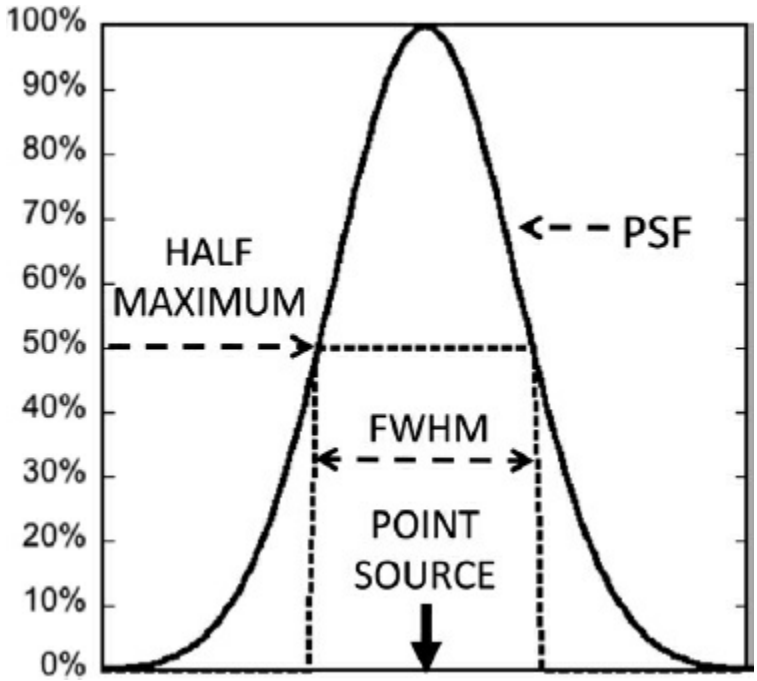

FIGURE 1.17–8. Schematic representation of a PSF from a point source of activity at the center of the horizontal axis.

PET-CT and SPECT-CT. A CT system is mounted along the same axis as the PET or SPECT system. The CT and radionuclide imaging are performed in sequence, not simultaneously, as the locations of the imaging detectors for the two systems are different. Very accurate coregistration of the images is possible because the distance between the patient positions for the two modalities is known precisely and can be used for image fusion in which the neurochemical information from the radiotracer image is superimposed on the more anatomically precise, high-resolution CT image. This has proven to be very valuable for imaging of the chest or abdomen, as organs are easily identifiable on CT images but not on PET or SPECT images. It is of less value for brain imaging, because CT images do not distinguish between soft tissue types. MRIs, however, do distinguish between gray matter, white matter and CSF, and MRI is therefore a preferable modality for anatomical coregistration and structure identification. Frequently in neuroimaging studies, MRIs will be obtained in a separate session and coregistration will be performed by software methods. PET-CT and SPECT-CT systems utilize the CT scan as transmission scan for attenuation correction. An advantage of this use is the high signal to noise ratio obtainable in the CT images. The photons from CT scanners are x-rays spread over a relatively broad energy spectrum from ~40 to 140 keV. These have different (higher) attenuation than the higher energy photons from PET or SPECT and require use of a lookup table method to transform their μ values into appropriate values for the other modalities.

PET-MR. PET-MR scanners are now being commercially produced. Unlike PET-CT, the PET and MRI components are concentric and images are acquired simultaneously. For whole body imaging, the main advantages are the ability to obtain anatomical images without the additional radiation burden of CT, and reduced overall imaging time. For neuroimaging, there is the potential, in addition to the anatomical information from the MR, to simultaneously utilize other MR techniques such as perfusion imaging or BOLD fMRI, greatly expanding the experimental information that can be obtained in a single imaging session. An ongoing technical challenge is attenuation correction. The concentric scanner ring design and the current size of the components from both modalities preclude the inclusion of rod or point sources of radioactivity used for conventional transmission scans. However, the MRIs, while providing very precise anatomical information about brain tissue, provide no direct information about attenuation of high-energy photons. Additionally, skull bone and other materials in the PET field of view such as the patient bed or the MRI radiofrequency coil (which is outside the MR field of view) that do attenuate photons are not detected with MRI (skull can be detected using special fast echo radiofrequency pulse sequences, but these are not yet widely available). Current approaches for brain imaging involve coregistering acquired MRIs to

standard atlases and using standardized μ values from lookup tables. Other, simpler methods, such as fat/water tissue segmentation algorithms in which each voxel is assigned to a tissue type with a standardized μ value, have proven adequate for whole body scanning. Attenuation correction in PET-MR is currently an active area of research; improvements and refinements can be expected as the technology matures and interest grows.

Safety

Radiation Exposure and Internal Dosimetry. Radioactive decay leads to deposition of energy in tissue when positrons, electrons, or photons interact with molecules in tissues. While the type and amounts of radiation that PET and SPECT research subjects are exposed to are very unlikely to lead to acute effects, the long-term impact and the possibility of genetic damage are less well understood. Risk/benefit considerations have led governments and international advisory groups to propose guidelines and regulations for exposure limits to research subjects. These are based on the linear nonthreshold model that assumes observations made at much higher exposures with qualitatively different radiation can be extrapolated to the levels that research subjects are exposed to in PET and SPECT studies. There is not currently a consensus among radiation biologists as to the validity of this assumption, but in the absence of definitive data a conservative approach has been adapted. In the United States, there are FDA regulations limiting maximum radiation exposure to research subjects. Other countries are guided by the recommendations of the International Commission on Radiological Protection (ICRP). The amount of energy deposited in any organ or part of the body is a complicated function of the tissue type, the form of the organ, the isotope, and the distribution of the radiotracer throughout the body. The estimated radiation exposure associated with a radiotracer is most often determined through a combination of mathematical models, computer simulations, and test scans known collectively as internal dosimetry studies. The distribution of the radiolabel throughout the body will differ between radiotracers according to how the tracer is metabolized and cleared, and this is determined empirically by acquiring whole body scans, either in a small cohort of human subjects or in animals with application of scaling factors to account for size differences between species. Organ radiation doses are calculated as weighted sums of the energy due to decays occurring in the organ itself and the energy deposited in it from decay events originating in other parts of the body. The unit of absorbed dose is the Gray (Gy, joules per kg tissue). An older unit is the radiation absorbed dose (rad, 1 rad = 0.01 Gy). Exposure is usually reported as absorbed dose per injected unit of radioactivity, such as mGy/MBq (1 mGy = 10^{-3} Gy) or rad/mCi. In the ICRP guidelines, absorbed doses are multiplied by an additional quality

factor that accounts for the type of radiation, leading to different units (the guidelines also apply to workers with occupational exposures to more damaging types of radioactivity than used in PET and SPECT). After multiplication by the quality factors, Gy are transformed to units called Sieverts (Sv), and rads become Roentgen-equivalent man (rem); however, the quality factors are all equal to 1 for the types of radiation from PET and SPECT isotopes so that Sv are numerically equivalent to Gy and rems are numerically equivalent to rads. The weighting coefficients used to calculate the radiation dose to each organ as a sum of contributions of activity originating from all organs are derived from tables generated through detailed computer simulations that account for organ shape, size, composition, and distance from other organs in mathematical models of humans of different sizes, ages, and sexes. The FDA regulations are based on the concept of the critical organ, in which the total allowable radiation per study is limited by the absorbed radiation dose to the single organ with the highest exposure. The ICRP guidelines are based on weighted averages of doses to a predetermined set of organs. There have been multiple revisions of these guidelines and the names and definitions of the weighted sums. The names have been somewhat confusingly similar (Dose Equivalent, Effective Dose, Equivalent Dose, Effective Dose Equivalent, etc.) but the standard as of ICRP Publication 103 (2007) is the Effective Dose (E, Sv) equal to a weighted sum of absorbed doses from 28 specified organs and tissues. While the precise dose limits differ from tracer to tracer, several scans per subject per study can be performed with most radiotracers while remaining within the allowable exposure range.

Pharmacological Toxicity. Compounds that are administered to human research subjects must undergo extensive toxicology screening prior to approval. The low mass dose usually administered to research subjects, in the range of several μg or less, generally keeps the dose many orders of magnitude lower than the level at which physiological effects are noticed. This is especially true of radiotracers that are neuroreceptor antagonists. At tracer doses, these maximally occupy several percent of the receptors, well below the amount necessary for pharmacological effects on neurotransmission. However, some agonist radiotracers have been known to elicit a pharmacological response even at very low mass doses. Two examples are the μ-opioid agonist [^{11}C]carfentanil, which can cause sedation and hypoventilation at μg doses, and the dopamine D2/D3 receptor agonist [^{11}C]-(+)-PHNO, which can cause transient nausea several minutes after a bolus injection.

PHARMACOKINETIC MODELS AND METHODS, EXPERIMENTAL DESIGNS

General Framework and Compartment Models

Above, some of the physical and image processing theory, technology, and algorithms used to create a PET or SPECT image were described. From the experimenter's perspective, all these components lead to one key characteristic of the imaging data that makes quantitative neuroimaging feasible: in each spatial location within the image, the measured amount of activity is proportional to the concentration of the tracer at that location. These measured amounts can be extracted as the average activity over a well-defined brain region of interest that can be as large as thousands of voxels or as small as a single voxel. Imaging data can then be statistically fitted to mathematical models of the observed biological processes, and parameter estimates can be derived to quantitatively characterize the processes. The imaging data can range from single, static snap shots of the tracer distribution obtained from 10 to 20 minutes of scanning, to sequential dynamic images that record the kinetic profile of the tracer over the course of hours. The mathematical models can range from simple measurements of the ratio of radiotracer concentrations in different tissues, to detailed differential equation models that relate the pharmacokinetic profile of the brain concentration of the radiotracer over hours to the concentration in arterial blood plasma, collected in parallel with the brain images over the same time period. The radiotracer enters the brain from the bloodstream, and the most comprehensive models combine an equation for transport across the BBB with equations for the interaction of the radiotracer with biological processes of interest in brain tissue; simpler models can sometimes be derived from these if appropriate conditions are met. The transport equation is usually based on Fick's principle, in which the net rate of influx or efflux across the BBB is proportional to the gradient in radiotracer concentrations between brain tissue and blood. Models of biological processes can be broadly categorized according to whether the radiotracer is a ligand that binds reversibly to a target molecule, or is a substrate for a metabolic pathway and becomes irreversibly trapped at some stage. A general, flexible framework that encompasses the reversible and irreversible cases is the compartment model (Fig. 1.17–9). A compartment model describes radiotracer kinetics with a simple set of rules, in which states of the radiotracer are represented as compartments, drawn as boxes, and changes of state are represented as flow between compartments, denoted by arrows showing allowable directions of flow between compartments, at rates that are proportional to the tracer concentration in the source compartment. Corresponding to the compartment diagram is a system of differential equations describing the rates. Outcome measures containing relevant biological or pharmacokinetic information are formed as functions of the rate constants. The rate constants are unknown parameters; they are estimated by using a computer-based data fitting procedure that provides the best fit of the model to the measured PET or SPECT data. The most common criterion for

best fit is least squares minimization, in which the program searches through the entire plausible range of rate constants and finds the ones that minimize the sum of the squared differences between the measured data, recorded as discrete "frames" at a sequence of time points over the course of the scan, and the model predictions at these time points, although many data fitting and parameter extraction procedures are in use in addition to the conventional least squares approach.

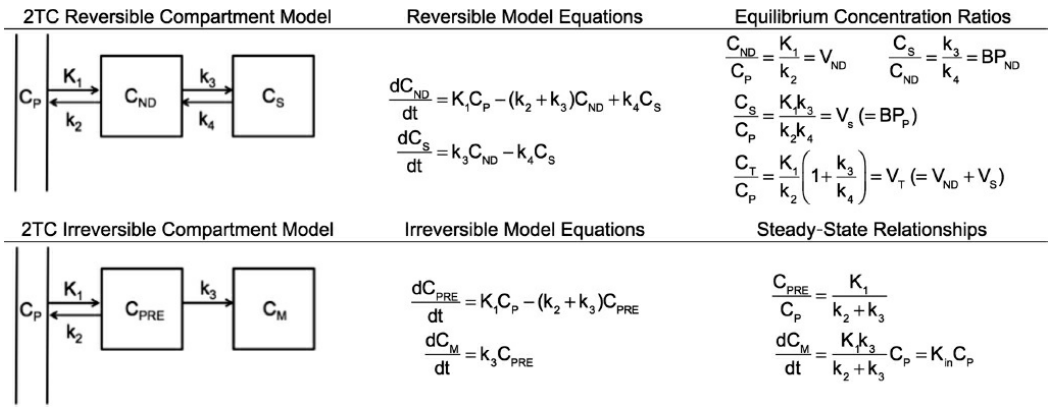

FIGURE 1.17–9. Some of the mathematics of compartment models. The box representations (left) for reversible and irreversible 2TC models. The center column shows the associated differential equations. These are derived by the simple rules that the overall rate of change of a compartment is a weighted sum of all source compartment concentrations (ones with *arrows* pointing into it) minus a weighted sum of the compartment concentration itself if the compartment is reversible (*arrows* pointing out of it). The weights are the rate constants, labeled K_1, k_2, k_3, and k_4. The right column shows outcome measures. For the reversible tracers, all outcome measures are equilibrium ratios of concentrations. Because they are equilibrium ratios, the rates of change of each compartment (the left side of the equations in the center column) are 0, and the concentration ratios can be derived as functions of the rate constants. For the irreversible model, the steady-state representation of C_{PRE} can be derived in a similar way, and then substitution into the differential equation for the rate of accumulation of C_M leads to a rate constant expression for K_{in}.

Reversible Models. Reversibly binding tracers are used most frequently to quantify receptor, transporter, enzyme, or other protein density in brain. These tracers are also referred to as radioligands. The modeled processes for reversible radioligand binding are bimolecular association between the radioligand and target protein and unimolecular dissociation of the target–ligand complex. The differential equation for this process is the mass action law. The canonic compartment model applied to reversible radioligands is the two-tissue compartment model (2TC, Fig. 1.17–9). The 2TC model requires that the concentration of radiotracer in arterial plasma be measured during the scan as an input to the model. This is invasive, as it requires arterial catheterization, and technically demanding, as it requires application of chromatography to estimate the fraction of radioactivity in blood that is due to unmetabolized radioligand rather than radiolabeled metabolites that form in the periphery and

accumulate in the blood over the course of the scan. For some radiotracers, there are methods for estimating relevant parameters that do not require an arterial plasma input function, but even these are often conceptually derived from the 2TC, highlighting its central role in models of reversible tracers. There are two main simplifying assumptions in the 2TC. When the radioligand crosses the BBB, it is initially in a free, unbound state. From the free state, it can associate specifically with the target protein it was designed to bind to, or, it can loosely bind to membranes nonspecifically. Specific binding is characterized by certain properties: it is saturable (it cannot exceed the concentration of the available target protein regardless of how high the ligand concentration is) and it is displaceable by other ligands that bind to the same site. Conversely, nonspecific binding is not saturable and not displaceable. Also, compared to the rate of the specific binding process, free and nonspecifically bound radiotracers equilibrate rapidly. The first simplifying assumption is that this equilibration occurs instantaneously, so that free and nonspecifically bound radiotracers are combined into a single compartment called the nondisplaceable compartment (C_{ND}). Only the free fraction of C_{ND}, denoted f_{ND}, is able to bind specifically to the target molecule. The bound state is the specific binding compartment (C_S). The second simplification is called the tracer dose approximation. The bimolecular association term of the mass action law contains the product of the unbound target protein concentration and the free radiotracer concentration. The tracer dose approximation is that the concentration of radioligand is so low that at most several percent of the target is bound, and therefore the target concentration can be treated as if it is constant. To satisfy the tracer dose approximation condition, very low mass, usually several micrograms or less, of radioligand are injected. This also limits the possibility that the radiotracer will pharmacologically perturb the system under study. In theory, models could be formed that do not include these simplifications—free and nonspecifically bound tracer could be treated as separate compartments and the equations of the full mass action law without the tracer dose assumption could be implemented, but experience has proven the 2TC model to be the most comprehensive model that is also robust and stable enough for reliable parameter estimation. The two compartments, C_{ND} and C_S represent states that are occurring simultaneously in the same voxels and cannot be separated in the image. Only their sum, $C_T = C_{ND} + C_S$ (T for Total concentration) can be observed; the relationship between C_{ND} and C_S must be inferred through the model. Concentration of the radiotracer in arterial plasma is denoted C_P. There are two types of related outcome measures derivable from the 2TC model. These are distribution volumes and binding potentials. These represent equilibrium ratios of radiotracer concentrations under tracer dose conditions (Fig. 1.17–9), but they are rarely estimated under

equilibrium conditions. Rather, they are derived from dynamic scan data, usually acquired following a bolus injection of the radiotracer, as functions of the estimated rate constants. Some of the mathematics leading to the derived relationships between distribution volumes and rate constants are provided in greater detail in the legend to Figure 1.17–9. Distribution volumes represent equilibrium ratios between tracer concentration in compartments and tracer concentration in arterial plasma. For the 2TC model, there are three relevant distribution volumes: $V_{ND} = C_{ND}/C_P$, $V_S = C_S/C_P$, and $V_T = V_{ND} + V_S$. Binding potentials are parameters that are proportional to the ratio of specifically bound to free radiotracer at equilibrium. There are several forms of binding potential with different constants of proportionality, resulting from different ways of approximating the free concentration, but the version that is frequently reported in PET and SPECT studies is BP_{ND}, the ratio C_S/C_{ND} at equilibrium. From the distribution volume definitions, BP_{ND} can be seen to equal V_S/V_{ND}. Physiologically, BP_{ND} is equivalent to $f_{ND}B_{MAX}/K_D$, where B_{MAX} (nM) is the concentration of target protein. Receptor studies report this as B_{AVAIL} to emphasize the fact that that some portion of the total concentration B_{MAX} may not be available to bind to the radiotracer due to phenomena such as baseline receptor occupancy by the endogenous neurotransmitter or trafficking of receptors between the cell surface and cytoplasm. K_D (nM) is the equilibrium dissociation constant (reciprocal of the affinity). From the equilibrium definition C_S/C_{ND}, it is apparent that BP_{ND} is a ratio of specific binding, the quantity of most interest, to nondisplaceable concentration, which incorporates nonspecific binding, essentially a nuisance variable. For this reason, BP_{ND} is sometimes thought of as a signal to noise ratio. In studies where C_P is measured, V_T can be estimated in all brain regions. BP_{ND} can be estimated through use of the 2TC model provided there is a reference tissue, a brain region with negligible target protein concentration (Fig. 1.17–10). In a reference tissue, $V_T = V_{ND}$. An additional assumption applied in all compartment models that is well supported by data to within experimental error is that V_{ND} is the same in all brain tissues, so that V_T in the reference tissue provides an estimate of V_{ND} throughout the brain (Fig. 1.17–10). A particularly robust way to estimate BP_{ND} in all brain regions is through distribution volumes as $[V_T(\text{region}) - V_T(\text{reference tissue})]/V_T(\text{reference tissue})$. Also, if there is a reference tissue, there are mathematical models that eliminate the arterial plasma concentration C_P from the equations and estimate BP_{ND} directly from the relationship between the target-rich brain regions and the reference tissue; these allow BP_{ND} estimation without the requirement for arterial blood samples.

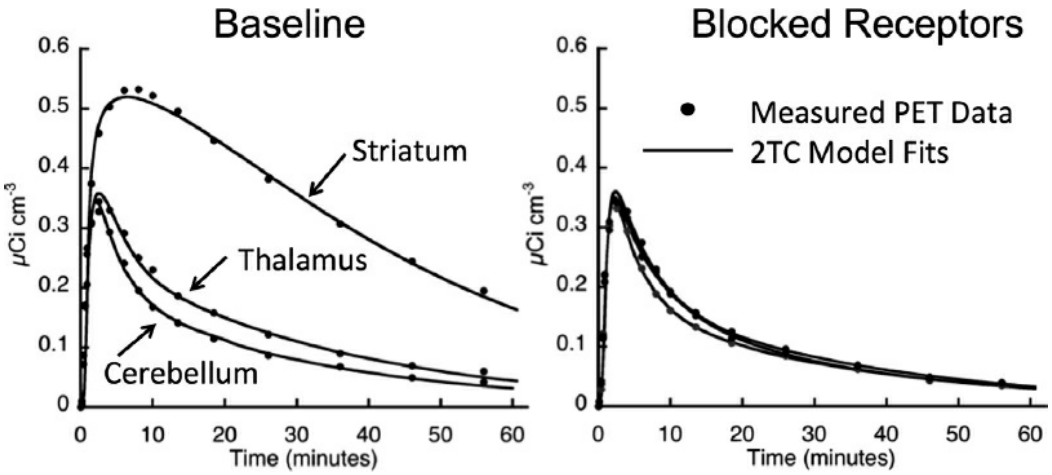

FIGURE 1.17–10. Time activity curves and 2TC fit at baseline and after receptor blockade. The data are from [^{11}C]raclopride scans, a dopamine D2-like receptor tracer, in a nonhuman primate at baseline (left) and following D2 receptor blockade with haloperidol (right). The figure illustrates several aspects of 2TC models. Distribution volumes (V_T) at baseline are 3.9 in striatum, 1.4 in thalamus, and 1.0 in cerebellum. Using cerebellum V_T as an estimate of V_{ND} leads to BP_{ND} = 2.9 in striatum and 0.4 in thalamus. Following receptor blockade, all regions appear similar to cerebellum and all V_T are close to 1, illustrating the concept that V_T in the reference tissue, cerebellum, provides a good estimate for V_{ND} in regions containing receptors.

Experimental Designs for Reversible Models. The most direct application of reversible models is the comparison of target protein density across groups or conditions. If the assumption is made that K_D and f_{ND} do not differ, on average, between scans being compared, then BP_{ND} can be taken to represent differences in B_{MAX} or B_{AVAIL}. Another flexible paradigm that can model many different phenomena is to compare the relative change in BP_{ND} across conditions within subjects, sometimes denoted as ΔBP_{ND}: $\Delta BP_{ND} = [BP_{ND}(\text{condition 2}) - BP_{ND}(\text{condition 1})]/BP_{ND}(\text{condition 1})$. The most common applications are those in which condition 1 is a baseline measurement and condition 2 is a scan following a pharmacological challenge that perturbs the system. Stimulant challenges induce an endogenous ligand for the binding site, such as a neurotransmitter in the case of receptor binding, to be released and compete with the radioligand for binding sites. Another example is a depletion paradigm, in which the challenge prevents or reduces synthesis of the endogenous substance that normally binds to the site under baseline conditions, leading to the unmasking of the underlying B_{MAX} or B_{AVAIL}. ΔBP_{ND} also provides an estimate of the target fraction, or percent occupancy, attained by an exogenously administered drug. In the simplest case when there is only a single competing ligand occupying some portion of the target molecules in condition 2 and no competitors in condition 1, BP_{ND} will be reduced by the factor $1/(1+[L]/K_{DL})$ where [L] is the

concentration of the competitor and K_{DL} is the equilibrium dissociation constant of L for the target. The factor $1/(1+[L]/K_{DL})$, equals 1 – occupancy, where occupancy is the fraction of the target concentration occupied at equilibrium by L at concentration [L]. Therefore, ΔBP_{ND} provides an estimate of the target occupancy attained by the competitor L, and PET imaging has become an important tool in the drug development and characterization of existing drugs. In vivo occupancy studies with PET have become nearly de rigueur for demonstrating target engagement by candidate drugs, and the need for such studies has motivated the development of tracers for new target molecules under investigation as sites for drug therapy. Occupancy study designs have included both simple measurement of receptor occupancy in stably-dosed patients and more comprehensive designs utilizing experimental ranges of doses to form models of the relationship between doses or drug plasma levels and target occupancy, with estimates of related parameters such as ED_{50} or EC_{50} (dose and blood plasma concentration at which 50 percent of the target is occupied). Studies where ΔBP_{ND} is used to measure changes other than target occupancy by an exogenously administered drug, especially ones that utilize pharmacological challenges to induce changes in neurochemical systems outside of the normal physiological range, likely involve underlying pharmacokinetics or biophysical processes that are more complicated than the simple competition model implies, but ΔBP_{ND} is still an informative parameter to quantify the magnitude of the challenge effect.

Irreversible Models. The most commonly used irreversible compartment model also contains two tissues, but has no route of return from the terminal compartment to the intermediate compartment as in the reversible model; the radiotracer or radiolabeled metabolic byproduct is effectively trapped in the final state. This model is most often applied in cases where the radiotracer is a substrate for several, but not all of the steps in the metabolic pathway of an endogenous substance, with the radiolabel retained on the metabolic byproduct of each of these steps. If the process by which clearance from the last shared step is slow enough such that virtually all of the accumulated radiolabeled metabolite remains in that state for the duration of the scan, it can be treated as a trapped or irreversible state. Unlike the reversible models, irreversible models do not have a broadly agreed upon nomenclature, but here the terminology C_{PRE} (for precursor, subsuming all reversible states in brain prior to the last, trapped state) and C_M (for metabolite, the trapped state) are adapted. While there is no equilibrium state in an irreversible model, a hypothetical steady state can be envisioned in which the efflux and influx of C_{PRE} to and from the arterial plasma are just balanced by the forward transformation of C_{PRE} to C_M and the concentration C_{PRE} remains constant. From the

compartment model equations, this steady state would satisfy the condition $C_{\mathrm{PRE}} = [K_1/(k_2 + k_3)]C_{\mathrm{P}}$. The overall rate of trapping, equal to $k_3 C_{\mathrm{PRE}}$, would then equal $[K_1\ k_3/(k_2 + k_3)]C_{\mathrm{P}}$. The expression $K_1\ k_3/(k_2 + k_3)$ is called K_{in}, the steady-state uptake rate constant of the tracer, and is the usual outcome measure estimated in irreversible models. In similar fashion to reversible models, if a reference tissue for the irreversible process exists, a modified version of K_{in} can be estimated without the need for collecting arterial plasma.

Examples of Irreversible Models. The assumptions in irreversible models are more specific to the individual applications than with reversible models. A general sense for the way they work is most readily described through the examples of [^{18}F] DOPA and ^{18}FDG, two of the most widely used metabolism tracers in brain PET.

[^{18}F] DOPA is a substrate for AADC, the enzyme that catalyzes L-dihydroxyphenylalanine (DOPA) into dopamine within the dopaminergic terminals. [^{18}F] DOPA readily passes the BBB and cell membranes, and is metabolized by AADC into 6-fluorodopamine (6-FDA), effectively ^{18}F-labeled dopamine, in dopaminergic terminals. 6-FDA, like endogenous dopamine, does not cross the BBB and so is "trapped" in the proximity of the terminals and the cycle of vesicular loading, release through exocytosis and reuptake into the terminal, in parallel with endogenous dopamine in the striatum, where reuptake through DATs is the dominant mode of clearance from the synapse. K_{in}, estimated with the irreversible trapping model, represents a marker of presynaptic dopaminergic condition. The trapping is only approximate, because 6-FDA is a substrate for MAO and for catechol-*O*-methyltransferase (COMT), and the radiolabeled metabolic by-products from 6-FDA metabolism can readily diffuse across the BBB. Several approaches have been proposed to incorporate the slow efflux from brain, sometimes called k_{loss}, into the modeling and design of [^{18}F] DOPA experiments, but many studies quantify K_{in} with the more basic irreversible model. [^{18}F] DOPA data sets can comprise anywhere between one and several hours of dynamic imaging. Several brain regions, including cerebellum, can function as a reference tissue for [^{18}F] DOPA imaging.

^{18}F-labeled 2-fluoro-2-deoxy-D-glucose (^{18}FDG) is by far the most widely used PET radiotracer. Originally developed for quantifying brain glucose metabolism, it was observed early on that ^{18}FDG is taken up more rapidly by some types of tumor cell than by the surrounding tissue, and ^{18}FDG imaging became a widely used diagnostic and staging tool in oncology and other medical fields. It is still frequently used in research for quantification of brain glucose metabolism as well. ^{18}FDG is a substrate for the same carrier protein that transports glucose into brain tissue. It is also a substrate for hexokinase, the enzyme that metabolizes glucose, and is

phosphorylated into ^{18}FDG-6-PO4. ^{18}FDG-6-PO4 does not follow further steps of glucose metabolism. It does dephosphorylate, but at a rate that is very slow compared to the forward process, and ^{18}FDG-6-PO4 is treated as a trapped state. Here, K_{in} represents the steady-state phosphorylation rate of ^{18}FDG. This is proportional, not identical, to the rate of these steps in brain glucose metabolism, owing to the fact that the transport and phosphorylation rates are different for the two compounds. When multiplied by a conversion factor accounting for this difference (1/LC, the "lumped constant," because it lumps two conversion factors together), K_{in}/LC times the plasma glucose concentration is taken as an estimate of brain glucose metabolism. As glucose is utilized throughout the brain, there are no reference tissues. The most comprehensive approach to estimating K_{in} with ^{18}FDG is to obtain arterial samples and fit the PET data to the 2TC irreversible model with arterial input, or a modified version that accounts for the slow dephosphorylation, but this is rarely done in clinical research studies, where simplifications of both the blood input and the PET data acquisition are frequently employed. Simplification of blood input data can entail utilizing venous blood that has been "arterialized" by warming the site of the venous catheter to produce vasodilation and more rapid equilibration between arterial and venous ^{18}FDG concentrations, or even simpler, collecting only a small number of samples and using these to scale a standardized population input. PET data acquisition can be reduced by approximating K_{in} as the ^{18}FDG concentration in brain at steady state divided by the area under the curve of the arterial plasma input; the image of the average ^{18}FDG concentration over a relatively short period of time, typically 30 minutes or less of acquisition starting between 45 and 60 minutes after a bolus injection, is used in this approach. Finally, the steady-state approach can be simplified even further by estimating a parameter called the standard uptake value (SUV), in which the area under the curve of the ^{18}FDG in blood is replaced by the injected radioactivity divided by the subject's body weight. SUV is a widely used outcome measure in clinical scans, but has come under criticism as a research tool because without any direct assessment of the concentration of ^{18}FDG in blood, the method cannot account for differences across subjects or conditions in the rate of peripheral clearance of ^{18}FDG, and therefore can only provide information about relative metabolism across the brain within a scan.

RADIOTRACERS

Table 1.17–2 lists PET and SPECT radiotracers that have been successfully used in human subjects in recent years. Tracers for new targets and improved tracers for established targets are constantly under development; the table is intended to convey a sense of the resources available in 2015.

Many more tracers have been proposed but have either failed to be adapted or have not yet progressed beyond testing in vitro or in preclinical species. The pharmacokinetic properties that lead to useful reversibly binding radiotracers are delicate and somewhat different than those that might make a compound suitable for in vitro binding experiments or as a therapeutic drug. Generally, balance is required between a number of different properties in order for quantification to be reliable. For example, a prerequisite is that a tracer can be adequately supplied to brain tissues in order to provide a detectable signal. Several factors can affect this. The tracer will bind nonspecifically to plasma proteins in the blood. The protein-bound fraction (the unbound fraction is called f_P; $1 - f_P$ is the bound fraction) can be very high; for some radiotracers $1 - f_P > 99$ percent. In some cases this may prevent an adequate concentration of radiotracer from entering the brain for reliable image quantification. BBB penetrability is an issue that must be addressed in the in vivo setting that does not arise in cell cultures or in vitro assays utilizing brain tissue. A property related to this is lipophilicity. High lipophilicity generally improves the potential of a compound for BBB penetration, but it also leads to high nonspecific binding. If the nonspecific binding is too high relative to specific binding, the specific binding signal will be difficult to separate from background, leading effectively, to low BP_{ND} and unreliable quantification. Thus, lipophilicity must have an intermediate value (log P in the range 1.5 to 4) to balance between the need for brain penetrance and the undesirable effect of high nonspecific binding. Also related to BBB penetrability is the fact that some radiotracer candidates are substrates or weak substrates for P-glycoprotein (P-gp) or other efflux mechanisms that actively prevent brain entry; this can be hard to predict prior to in vivo testing. Another property that needs to be considered is the relationship between affinity of the tracer for its target and the target density. As $BP_{ND} = f_{ND}B_{MAX}/K_D$, it can be seen that for a given target availability and f_{ND}, BP_{ND} varies directly with affinity. If BP_{ND} is too low, specific binding cannot be reliably measured relative to nonspecific binding (Fig. 1.17–11). Conversely, if target density is very high, a high affinity tracer may not reach a washout phase within a reasonable time frame and parameter estimates will not be stable—small perturbations of the data late in the scan will lead to large changes in outcome measures (Fig. 1.17–11). Targets that are distributed very heterogeneously in the brain may require multiple radiotracers, with moderate affinity tracers for high B_{MAX} regions and high-affinity tracers for low B_{MAX} regions. Some tracers that appear promising in development may turn out to be too far to one end or the other of this spectrum to be useful in humans: either BP_{ND} too low everywhere or wash out too slow everywhere. Another phenomenon that can limit the usefulness of a radiotracer is the discovery that in vivo, its selectivity for the primary target molecule

compared to a secondary target is less than predicted in vitro. Generally, tracers are screened in vitro for binding affinity to many brain proteins, but there have been cases where the in vitro affinity for a primary target is several hundred-fold greater than for another protein, but in vivo the affinities are found to differ by much less and binding to the secondary target presents a confound for quantification. Finally, a tracer may appear to have high potential according to all of the above criteria and then turn out to have a radiolabeled metabolite in the blood that crosses the BBB and makes quantification difficult or impossible.

Chemists working in radiotracer development can control some of these parameters, but much of the process is trial and error. Development might entail starting with an initial pharmacophore or a molecule known to have specificity for the target and then iterating through the sequence of devising a radiolabeling procedure, testing the result in vitro, testing in preclinical species, altering the molecule to improve properties based on test results, and repeating until a suitable candidate has strong enough potential for human testing. This process can be long and arduous, often taking years to arrive at a suitable radiotracer for a given target. Some recent work has focused on ways to streamline the process or on the development of a rational approach to predicting the quantitative potential of new tracers early in their development.

NEUROCHEMICAL IMAGING IN PSYCHIATRIC RESEARCH

As Table 1.17–2 demonstrates, radiotracers for imaging many different neurochemical systems have progressed as far as human imaging. Some of these are recently developed or have only been used in one or several studies, whereas dozens or even hundreds of studies of neuropsychiatric illnesses and other conditions related to brain function have been published using some of the more established tracers. Some of the most studied areas are schizophrenia and related conditions, major depression and other mood and affect disorders, substance and alcohol abuse and addiction, and more recently, Alzheimer disease and other dementias, but many more have been investigated. Some of these extensively studied areas and some well-replicated results are described here. This section is not intended to be a comprehensive review of all neuroimaging results in psychiatry, but rather, a selective sample to convey the nature of the research, some of the more well-established or informative results, and cases illustrating the capabilities as well as limitations of radiotracer imaging.

Dopamine

The dopaminergic system is one of the most extensively studied neurotransmission systems with both PET and SPECT imaging. Studies

have examined receptor and transporter availability as well as transmitter synthesis, storage and release capacity in cohorts from a broad range of patient populations. Most recently, several tracers have been developed for intracellular enzymes involved in the signaling pathways of dopamine receptors such and phosphodiesterase 10A (PDE10A). Because all antipsychotic drugs block dopamine D2 receptor transmission to some degree, and because of the role of dopamine transmission in functions such as attribution of salience to environmental stimuli or fine tuning of working memory, dopaminergic dysfunction is considerably implicated in schizophrenia and there have been a particularly large number of imaging studies characterizing different aspects of the dopaminergic system in patients with schizophrenia. Another area that has produced many imaging studies of the dopaminergic system is substance dependence. Many drugs of abuse induce synaptic release of dopamine. This is especially true of drugs such as cocaine and amphetamines that act directly on dopamine terminals, but is also true for other substances such as alcohol or nicotine that act indirectly. Because of this, and because of dopamine's roles in motivation, identification of rewarding stimuli, and reinforcement of rewarding behaviors, dopaminergic dysfunction has been implicated in drug abuse and addiction, and this has been an area of focus for imaging.

Table 1.17–2.
PET and SPECT Radiotracers in Published Studies with Human Subjects

Target System	Specific Target	PET Radioligand	SPECT Radioligand
Dopamine	D2/D3 receptor	[^{11}C]raclopride	[^{123}I]IBZM
		[^{18}F]fallypride	[^{123}I]epidepride
		[^{11}C]fallypride	
		[^{11}C]FLB457	
		[^{11}C]-(+)-PHNO	
		[^{11}C]NPA	
		[^{11}C]MNPA	
	D1/D5 receptor	[^{11}C]NNC112	
		[^{11}C]SCH23390	
	Reuptake Transporter	[^{11}C]PE2I	[^{123}I]PE2I
		[^{11}C]altropane	[^{99m}Tc]TRODAT-1
		[^{18}F]CFT	[^{123}I]β-CIT
		[^{11}C]CFT	[^{123}I]FP-CIT
	DA terminals	[^{18}F]DOPA	
		[^{18}F]FMT	
	VMAT2	[^{11}C]DTBZ	
		[^{18}F]AV133	
Serotonin	SERT	[^{11}C]McN5652	[^{123}I]β-CIT
		[^{11}C]DASB	[^{123}I]ADAM
		[^{11}C]AFM	

		[^{11}C]HOMADAM	
	5-HT$_{1A}$ receptor	[*carbonyl* -^{11}C]WAY100635	
		[^{18}F]MPPF	
		[^{18}F]FCWAY	
		[^{18}F]MEFWAY	
		[^{11}C]CUMI101	
	5-HT$_{2A}$ receptor	[^{11}C]MDL100907	[^{123}I]5-I-R91150
		[^{11}C]CIMBI-36	
		[^{18}F]Altanserin	
	5-HT$_{1B}$ receptor	[^{11}C]P943	
		[^{11}C]AZ10419369	
	5-HT$_{6}$ receptor	[^{11}C]GSK215083	
Norepinephrine	NET	[^{18}F]MeNER	
		[^{11}C]MRB	
Opioid	μ opioid receptor	[^{11}C]carfentanil	
	opioid receptor	[^{11}C]diprenorphine	
Acetylcholine	α4β2* nicotinic receptor	2-[^{18}F]-FA-85380	[^{123}I]5–1A-85380
		[^{18}F]AZAN	
	nicotinic receptors	[^{11}C]nicotine	
	muscarinic M2	[^{18}F]FP-TZTP	
Glycine	GlyT1	[^{11}C]GSK931145	
		[^{11}C]RO5013853	
GABA	GABA$_{A}$ receptor	[^{11}C]flumazenil	[^{123}I]iomazenil
Adenosine	Adenosine 2A receptor	[^{18}F]CPFPX	[^{123}I]MNI420
Glutamate	mGluR5	[^{11}C]ABP688	
		[^{18}F]FPEB	
Substance P	NK1 receptor	[^{18}F]SPA-RQ	
Enzymes	MAO B	[^{11}C]deprenyl	
	MAO A	[^{11}C]clorgyline	
		[^{11}C]harmine	
	PDE4	[^{11}C]rolipram	
	PDE10	[^{11}C]IMA107	
		[^{18}F]MNI659	
Neuroinflammation	TSPO	[^{11}C]PK11195	[^{123}I]PK11195
		[^{11}C]PBR28	[^{123}I]CLINDE
		[^{18}F]PBR06	
		[^{18}F]PBR111	
		[^{18}F]FEPPA	
Cannabinoid	CB1 receptor	[^{11}C]OMAR	[^{123}I]AM281
		[^{18}F]MK9470	
	FAAH	[^{11}C]CURB	
Misfolded Proteins	β-Amyloid	[^{11}C]PiB	
		[^{11}C]AZD2184	
		[^{11}C]AZD4694	
		[^{18}F]GE067	

	[^{11}C]SB13
	[^{18}F]BAY94–9172
	[^{18}F]AV45
	[^{18}F]FDDNP†
Tau	[^{18}F]T807
	[^{18}F]-THK105
	[^{18}F]-THK523
	[^{11}C]-PBB3

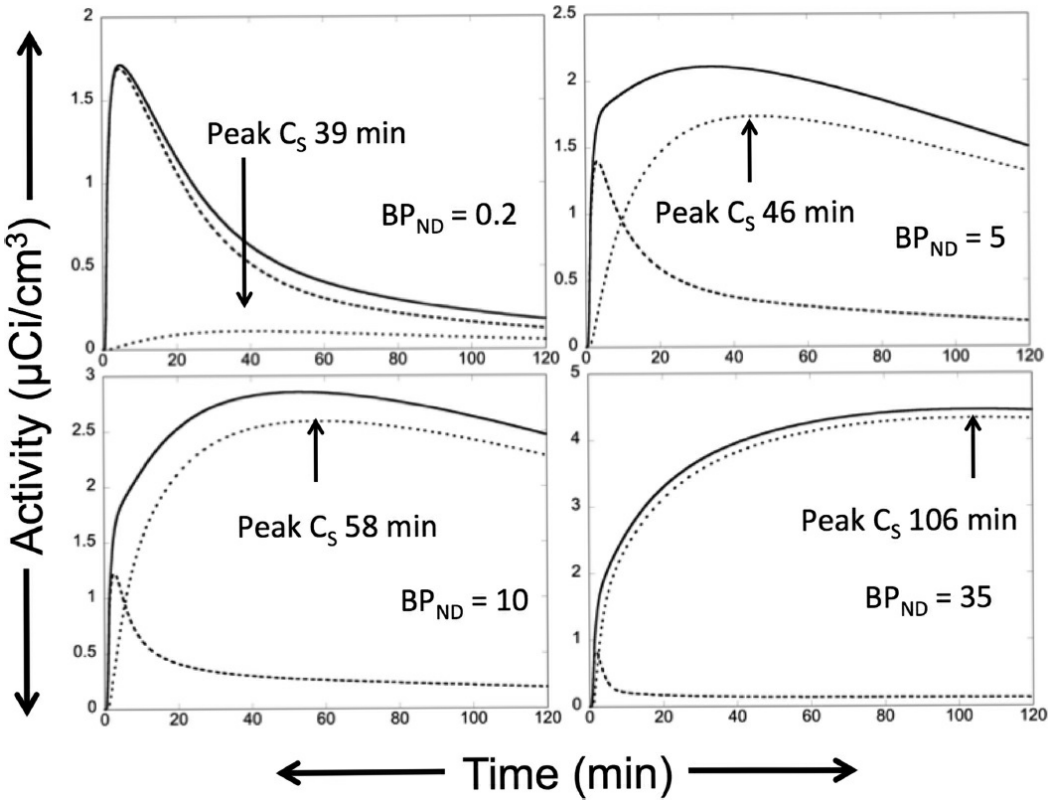

FIGURE 1.17–11. Time to peak-specific binding in simulated 2TC time activity curves. These curves were all created with the same input function and the same K_1, k_2 and k_4 values. They differ only in k_3, which is proportional to B_{AVAIL}. They illustrate the fact that the time of the peak-specific binding increases with increasing B_{AVAIL}. In the upper left panel (BP_{ND} = 0.2), the majority of the activity is attributable to the C_{ND} curve so that it will be difficult to distinguish C_T from C_{ND} and BP_{ND} will be difficult to estimate accurately; the signal to noise ratio is too low. In the lower right panel (BP_{ND} = 35), the peak-specific binding occurs so late in the scan that small perturbations of the last several data points could lead to large differences in estimated V_T or BP_{ND}. The other two panels (BP_{ND} = 5, 10) illustrate cases where parameter estimation would be robust, both because C_S comprises a large part of C_T, and peak-specific binding and a washout phase are reached during the course of the scan. Dashed curves are C_{ND}, dotted curves are C_S, and solid curves are C_T.

D1 Receptor Availability. Dopamine D1-like receptors are distributed throughout most of the brain, with high concentration in the striatum and low-to-moderate concentration in frontal cortex. D1 concentration in the cerebellum is below the detectable level with currently available tracers and cerebellum is treated as a reference tissue. Because of the involvement of

cortical dopaminergic function in cognition and working memory, there has been great interest in assessing cortical D1 in conditions where these are affected, especially schizophrenia. There are two PET radiotracers that have been most frequently used to assess D1 receptor availability. These are [^{11}C]NNC112 and [^{11}C]SCH23390. Neither tracer can distinguish between D1 and D5 receptors. The signal to noise ratio, embodied in BP_{ND} values, is approximately threefold higher for [^{11}C]NNC112 than for [^{11}C]SCH23390. Also, the rapid peripheral metabolism of [^{11}C]SCH23390 has made it difficult to measure arterial plasma concentration reliably, so that [^{11}C]SCH23390 is used exclusively with reference tissue methods, whereas [^{11}C]NNC112 can be used with both arterial input and reference tissue methods. Several studies have been performed to examine D1 availability in patients with schizophrenia. The studies utilizing [^{11}C]NNC112 have found increased cortical D1 in antipsychotic drug-free patients and especially patients without prior history of antipsychotic drug treatment compared to demographically matched control subjects, whereas the studies utilizing [^{11}C]SCH23390 have found decreased D1. Initially it was thought this might reflect differences in tracer properties, but subsequent studies from both sets of investigators, comparing subsamples of the original cohorts who had been administered both tracers, have yielded results in the same directions as originally reported, suggesting this may be more related to population heterogeneity than to the tracers. Over the past decade, several investigators have demonstrated that the selectivity of both [^{11}C]NNC112 and [^{11}C]SCH23390 for D1 compared to serotonin 5-HT_{2A} receptors is approximately 10-fold rather than the 100-fold difference measured in vitro, and that 25 to 35 percent of cortical BP_{ND} is attributable to 5-HT_{2A} binding for both tracers. Investigators presented some evidence that 5-HT_{2A} binding probably did not contribute to group differences in the schizophrenia studies, but unless more D1 selective radiotracers are developed, future studies will need to account for this factor when interpreting their results.

D2/D3 Receptor Availability. Dopamine D2-like receptors are heterogeneously distributed, with high concentration in the striatum, low concentration in thalamus, midbrain (substantia nigra and ventral tegmental area, SN/VTA) and medial temporal cortical regions (amygdala, hippocampus), and very low concentration in frontal cortex. These receptors are among the most studied with imaging and many tracers have been developed. Widely used tracers are [^{11}C]raclopride for PET and [^{123}I]IBZM for SPECT. Both are suitable for striatal imaging but do not have high enough BP_{ND} for imaging outside the striatum. Cerebellum is frequently used as a reference tissue for each, and neither tracer distinguishes between D2 and D3 receptors. Several higher affinity tracers

have been developed for extrastriatal imaging, including [^{18}F]fallypride and [^{11}C]FLB457 for PET and [^{123}I]epidepride for SPECT. [^{11}C]FLB457 can only be used quantitatively in cortex and limbic regions as striatal binding does not reach equilibrium within a suitable time frame for ^{11}C imaging. Also, unlike most other tracers for D2-like receptors, specific binding of [^{11}C]FLB457 is detectable in cerebellum. [^{18}F]fallypride, with its more long-lived radioisotope, can be used to simultaneously image the striatum and extrastriatal regions, though cortical BP_{ND} is still very low, and striatal imaging requires lengthy scanning of at least 3 to 4 hours. Specific binding in cerebellum has also been detected with [^{123}I]epidepride, which is chemically similar to [^{11}C]FLB457. The long half-life of ^{123}I allows simultaneous striatal and extrastriatal imaging, but again, with very long scan times in excess of 4 hours. Additionally, some investigators have detected radiolabeled metabolites of [^{123}I]epidepride in brain. None of these high-affinity tracers can be used to distinguish between D2 and D3 receptors. [^{11}C]-(+)-PHNO is an agonist radiotracer that is strongly D3-preferring. Estimates of its in vivo D3:D2 selectivity range from 20- to 50-fold. D3 receptors are much less widely distributed and in lower concentrations than D2 receptors, but the D3-preferrring nature of [^{11}C]-(+)-PHNO has allowed noninvasive in vivo characterization of D3 receptors for the first time. Binding is quantifiable in the striatum, the globus pallidus, the thalamus, the SN/VTA, and hypothalamus. Dorsal striatum BP_{ND} is attributable mostly to D2 binding, SN/VTA and hypothalamus BP_{ND} are attributable mostly to D3 binding, and globus pallidus, thalamus, and ventral striatum BP_{ND} represent a mixture of both receptor subtypes. Cerebellum is used as a reference tissue.

D2 RECEPTOR AVAILABILITY IN SCHIZOPHRENIA. Studies of striatal D2 receptor availability in schizophrenia have been mixed, with meta-analyses suggesting that striatal D2-like receptor availability may be slightly elevated if at all. Most of these studies were performed with [^{11}C]raclopride or [^{123}I]IBZM. One recent study also failed to detect a difference in striatal [^{11}C]-(+)-PHNO BP_{ND} between patients with schizophrenia and matched control subjects. Studies have also been published assessing extrastriatal D2 availability in schizophrenia using [^{18}F]fallypride, [^{11}C]FLB457 and [^{123}I]epidepride imaging. Individual studies have reported differences between patients and controls in several brain regions such as temporal cortex, anterior cingulate or thalamus, but these have not been seen consistently across studies or even go in opposite directions, suggesting that observed changes were either state dependent or due to population heterogeneity. In summary, the imaging literature to date suggests that dopaminergic dysfunction in schizophrenia is not likely mediated through receptor expression levels, but rather, through either pre- or postsynaptic

phenomena.

D2 RECEPTOR AVAILABILITY IN SUBSTANCE DEPENDENCE. In substance dependence, a very different picture has emerged in which studies have consistently shown decreased striatal D2 availability in subjects who are dependent on a variety of substances including cocaine, methamphetamine, alcohol, or heroin. An exception is cannabis use in which such differences have not been detected in several studies using [^{11}C]raclopride, but this may be related to the moderate use levels of the subjects enrolled in those studies and imaging studies focused on heavy users have yet to be reported.

Dopamine Release, Synthesis, and Storage Capacity. Many studies have used imaging methods to indirectly assess aspects of dopamine transmission. These include studies measuring ΔBP_{ND} of a receptor-binding tracer to infer levels, or changes in levels, of synaptic dopamine following pharmacological challenges, and metabolism studies using mainly [^{18}F]DOPA to assess presynaptic dopamine tone. In theory, any radiotracer that binds to receptors should be sensitive to changes in the synaptic concentration of the endogenous neurotransmitter, but for reasons that are not yet well understood, radiotracers for D2-like receptors have turned out to be the most sensitive. Thus, studies examining dopamine transmission have mainly used [^{11}C]raclopride or [^{123}I]IBZM. More recently, several studies have been published using [^{18}F]fallypride or [^{11}C]FLB457 to examine extrastriatal dopamine transmission, or [^{11}C]-(+)-PHNO, which provides another perspective due to its agonist and D3-preferring properties. Two of the most widely used pharmacological compounds for increasing synaptic dopamine are dextroamphetamine (d-AMPH) and methylphenidate (MPH). d-AMPH causes dopamine to be extruded from intracellular synaptic vesicles and to be released in synapses by reversing the direction of action of the dopamine reuptake transporter (DAT). ΔBP_{ND} in the striatum between a baseline scan and a post d-AMPH scan shows decreased binding ($\Delta BP_{ND} < 0$) consistent with competition at the receptors between the radiotracer and dopamine. The d-AMPH paradigm has been carefully validated; preclinical testing has demonstrated that measured ΔBP_{ND} is strongly correlated with released dopamine levels measured by microdialysis, that the measured levels are d-AMPH dose dependent, and that the d-AMPH effect is negated ($\Delta BP_{ND} \sim 0$) when dopamine is depleted prior to administration of d-AMPH, that is, the effect requires the presence of dopamine. There is some question as to whether, in addition to competition between the radiotracer and dopamine, the d-AMPH effect is mediated by other processes such as trafficking of receptors between the cell surface and cytoplasm, because the decreased BP_{ND} is sustained for at least several hours longer than the dopamine surge

following administration, but nonetheless ΔBP_{ND} following d-AMPH has been accepted as a reliable indicator of the capacity for dopamine synthesis, storage, and release. Microdialysis results from NHPs show increased extracellular concentration of dopamine following d-AMPH of 1,000 percent or more and suggest that in the dose range administered to human subjects dopamine release is high, but the effects on BP_{ND} are still small, with striatal ΔBP_{ND} in the range of 10 to 20 percent for [^{11}C]raclopride, [^{123}I]IBZM and [^{18}F]fallypride. Recently, a small number of studies have applied the d-AMPH paradigm with [^{11}C]-(+)-PHNO imaging and these suggest that this tracer may be more sensitive to dopamine than previous tracers, that is, ΔBP_{ND} is in the range of 20 to 30 percent; this may be related to the agonist properties of [^{11}C]-(+)-PHNO. MPH blocks reuptake through the transporter so that released dopamine accumulates in the synapse and extracellular space. The magnitude of accumulated dopamine concentration following MPH is less than the amount released following d-AMPH, but the effect is still reliably detected with imaging and MPH challenge has also been widely used as a measure of presynaptic dopamine tone. Because radiotracers for D2-like receptors are sensitive to binding competition from endogenous dopamine, and there is a sustained, low baseline level of dopamine binding to receptors, there is some question as to the extent to which apparent baseline BP_{ND} is reduced by baseline dopamine. Depletion studies, in which preadministration of a compound that pharmacologically prevents dopamine synthesis or vesicular storage, are designed to unmask the contribution of baseline dopamine to apparent BP_{ND}. The tyrosine hydroxylase inhibitor alpha-methyl-para-tyrosine (AMPT) prevents dopamine synthesis and causes depletion of dopamine stores when administered over a period of 48 hours or more; AMPT has been used in human subjects to unmask baseline occupancy of D2 receptors by dopamine.

DOPAMINE TRANSMISSION IN SCHIZOPHRENIA. Studies using the d-AMPH challenge have consistently shown greater dopamine release (larger magnitude of ΔBP_{ND}) in patients with schizophrenia compared to matched control subjects. These have been observed at the level of the whole striatum, using [^{123}I]IBZM or [^{11}C]raclopride imaging. Studies using the AMPT depletion paradigm have found elevated baseline dopamine at the level of the whole striatum using [^{123}I]IBZM, and, more recently, specifically in the associative striatum (caudate and precommissural putamen), using [^{11}C]raclopride imaging. These studies suggest that in humans, baseline occupancy of D2 receptors by dopamine is in the range of 10 to 20 percent. One study in which a small cohort of patients was imaged with both d-AMPH and AMPT challenges showed a high correlation between these, that is, greater AMPH-induced dopamine release was

associated with greater baseline dopamine binding as inferred from AMPT unmasking. Thus, imaging measures of dopamine transmission have consistently found elevated presynaptic tone in the striatum in schizophrenia. This is also in accord with [^{18}F] DOPA imaging in patients with schizophrenia and prodromal subjects at high risk of schizophrenia that have shown elevated K_{in} compared to matched controls in patients and intermediate elevation in prodromal subjects, with an association between the magnitude of [^{18}F] DOPA K_{in} and the likelihood of progression to psychosis in the prodrome.

DOPAMINE TRANSMISSION IN SUBSTANCE DEPENDENCE. In contrast to patients with schizophrenia, in substance-dependent subjects, reduced dopamine release relative to matched controls has been most frequently observed following either d-AMPH or MPH administration. This has been observed in cocaine, heroin, and alcohol dependent subjects. One study also used the AMPT depletion model to show that baseline dopamine levels were reduced in cocaine users. Combined with observations of reduced D2-like receptor availability, the imaging literature suggests an overall reduced function, or blunting, of the dopaminergic system in substance dependence. These studies cannot address the question of whether low dopamine function is a predisposing factor or consequence of drug use, but one recent study using an NHP model of cocaine dependence found that presynaptic vesicle availability in dopamine terminals, as measured with the VMAT2 radiotracer [^{11}C]DTBZ, is lowered following chronic cocaine administration, suggesting a direct role for cocaine itself in the observed reduction of dopamine tone in users.

Dopaminergic Target Engagement. As far back as the late 1980s and early 1990s, PET studies were published using ΔBP_{ND} to infer receptor occupancy by commonly used antipsychotic drugs. It was shown using [^{11}C]raclopride, that haloperidol (Haldol) and sulpiride (Dogmatil, Dolmatil, Eglonyl, Espiride, Modal, Sulpor) administered at therapeutic doses occupied between 75 and 90 percent of D2-like receptors and, using [^{11}C]SCH23390, that binding of both drugs to D1-like receptors was below detectable levels. The atypical antipsychotic clozapine (Clozaril) was shown to occupy ~50 percent of D2-like receptors and ~40 percent of D1-like receptors. These papers also were the first to demonstrate a direct link between EPS and the level of D2 receptor occupancy, showing a threshold for appearance of EPS at approximately 80 percent occupancy. Since that time, receptor occupancy has been characterized for many antipsychotic drugs as well as drugs from many other classes. Studies using [^{11}C]raclopride have repeatedly shown some widely used atypical antipsychotics such as risperidone (Risperdal) or olanzapine (Zyprexa) occupy 70 to 80 percent of striatal D2-like receptors in patients on stable

doses and have also replicated earlier results that D2 occupancy by therapeutic levels of clozapine is lower than for other antipsychotics. Additionally, risperidone, olanzapine, and clozapine have all been shown to occupy a greater percent of serotonin 5-HT_{2A} receptors than dopamine D2 receptors using radiotracers for 5-HT_{2A}. Studies measuring striatal and extrastriatal occupancy of D2-like receptors by the partial agonist antipsychotic aripiprazole (Abilify) using [^{18}F]fallypride have demonstrated very high occupancy, ~90 percent at stable doses, without EPS, consistent with the expectation that partial agonists should support higher levels of D2 receptor blockade without EPS than antagonists.

Serotonin

The serotonergic (5-HT) system has also been widely studied with neuroimaging. The involvement of 5-HT in major depression and other mood disorders has led to an extensive body of imaging studies measuring availability of receptors and reuptake transporters in those conditions, but serotonergic indices have been characterized across a wide spectrum of neuropsychiatric and related conditions.

5-HT_{1A} Receptor Availability. 5-HT_{1A} receptors are distributed throughout the brain with highest concentration in the insula and medial temporal cortex, high concentration in cingulate, and moderate concentration in prefrontal cortex and posterior cortex (parietal and occipital cortices). The receptors are also found in the dorsal raphe nuclei (DRN) where they act as inhibitory autoreceptors. DRN concentration is low, but still of particular interest because of the role of 5-HT_{1A} autoreceptors in regulation of serotonin transmission, as evidence from preclinical studies and other sources suggests that the mechanism of SSRIs may involve desensitization of 5-HT_{1A} autoreceptors in DRN over the first several weeks of treatment, leading eventually to increased serotonin transmission. [*carbonyl* -^{11}C]WAY100635, usually abbreviated as [^{11}C]WAY100635, a selective 5-HT_{1A} receptor antagonist, has been the most widely used radiotracer for this target. Cerebellum is frequently used as a reference tissue with [^{11}C]WAY100635, but there is detectable specific binding in the cerebellar vermis, so the reference tissue must be defined to exclude that part of cerebellum. [^{11}C]WAY100635 is difficult to produce, due to a long and complicated synthesis procedure, and quantification is also difficult due to rapid peripheral metabolism and very low signal in cerebellum, so there have been numerous attempts to developed modified versions and alternatives. Many of these use ^{18}F as the radiolabel and have run into problems with radiolabeled metabolites in brain or uptake of [^{18}F]fluoride ion in skull bone. Most recently, a new ^{11}C radiotracer, [^{11}C]CUMI101 has been introduced, but its properties are still being

characterized and it remains to be seen how widely it will be adapted. Thus, despite the difficulties associated with its use, most published studies to date in patient populations have been performed with [^{11}C]WAY100635.

5-HT1A RECEPTOR AVAILABILITY IN DEPRESSION. There have been a large number of studies examining 5-HT_{1A} availability with [^{11}C]WAY100635 in major depression disorder (MDD) and related mood conditions. The majority of these have reported decreased BP_{ND} in the DRN and many other brain regions in MDD patients compared to matched controls; a few have reported no differences between groups. A small number of investigators, however, report increased binding in DRN and other regions when using the outcome measure BP_F, derived as V_S/f_P. BP_F is a form of the binding potential that uses nonprotein bound tracer in arterial plasma as the reference value rather than the nondisplaceable concentration in brain used with BP_{ND}. There has been debate as to whether these exceptional findings are due to methodological differences, differences in patient samples, or a combination of these. Studies using BP_{ND} with cerebellum as reference tissue may be biased by group differences in cerebellum binding of the tracer or by group differences in f_{ND}. Conversely, arterial plasma and f_P are challenging to measure and may contribute variability to data leading to discrepant results. Some evidence suggests that [^{11}C]WAY100635 is a weak substrate for P-gp. Possible group differences in P-gp activity have not been investigated in MDD to date, but if they exist they might bias plasma-based outcome measures more than BP_{ND}. Thus, despite the importance of understanding the role of 5-HT_{1A} availability in MDD, there are currently controversies in the field that may only ultimately be resolvable with the development and use of more quantitatively reliable radiotracers for this target.

Serotonin Reuptake Transporter Availability. Serotonin reuptake transporter (SERT) is heterogeneously distributed in brain with highest concentration in midbrain, high concentration in striatum, amygdala and thalamus, moderate concentration in cingulate and hippocampus, and low concentration in frontal cortex. Cerebellum is treated as a reference tissue for SERT tracers. The first radiotracer used to image SERT was [^{123}I]β-CIT. This tracer has affinity for both SERT and the DAT and was thus not suitable for imaging SERT in striatum where DAT concentration is high. Effectively, [^{123}I]β-CIT could only be used for SERT measurement in midbrain. Subsequently, several PET radiotracers were developed for SERT imaging. The first of these, developed in the early 1990s, was [^{11}C]McN5652, which has a relatively low signal to noise ratio, but is suitable for imaging in midbrain and striatum. [^{11}C]DASB was developed in the late 1990s and has BP_{ND} values that are nearly double those of

[^{11}C]McN5652, but still very low in prefrontal cortex. [^{11}C]DASB has become the most widely used radiotracer for SERT imaging. More recently, higher affinity tracers such as [^{11}C]AFM and [^{11}C]HOMADAM have been developed that may provide reliable quantification in cortex, but studies using these have not yet been widely reported. SERT availability, measured with imaging, has been interpreted as a marker of 5-HT innervation and the integrity of the serotonergic system.

SERT AVAILABILITY IN MDMA USERS. 3,4-methylenedioxymethamphetamine (MDMA, ecstasy) releases 5-HT and increases its synaptic concentration by binding to and reversing the direction of action of SERT. MDMA has been used as a recreational drug, especially within the so-called "rave" subculture, due to its euphoria- and empathy-inducing qualities. Animal studies performed in the late 1980s indicated MDMA produced neural toxicity and reduced SERT at doses comparable to those of recreational users, raising the possibility that MDMA use could lead to reversible or even permanent neural damage. Initial studies, performed in the late 1990s, of SERT availability using [^{11}C]McN5652 found dramatically lowered BP_{ND} in MDMA users compared to matched control subjects. Several subsequent studies using [^{11}C]McN5652 did not replicate the dramatic result observed in the first study, but did show moderately decreased BP_{ND}, in the range of 20 to 30 percent, in current MDMA users that appeared to be reversible both in cross-sectional studies that included former users, as well as in correlations, with duration of abstinence or amount of use, in another study. Since that time, at least eight studies have been performed using [^{11}C]DASB in current and former MDMA users. These have consistently shown moderate decreases in cortical SERT availability in recently abstinent current MDMA users. One study examining former users found no difference compared to matched controls, suggesting the cautious interpretation of reversibility of MDMA-related neurotoxicity. An interesting aspect to this set of findings is that reported [^{11}C]DASB BP_{ND} values in prefrontal cortex range from approximately 0.05 to 0.15 in healthy controls; these are usually considered well below the limit of quantitative reliability, but nonetheless the result that there is reduced cortical SERT in recently abstinent MDMA users has been repeatedly replicated.

Serotonin Release, Synthesis, and Storage Capacity . Attempts to indirectly measure 5-HT release through competition with radiotracers following stimulant challenges have not been as successful as they have been for dopamine. The options for experimental design are more limited, because strong 5-HT releasing agents such as fenfluramine or MDMA are not suitable for human administration and studies using human subjects have had to rely on SSRI administration to increase 5-HT levels, which

likely leads to more modest effects than compounds that release 5-HT directly, but studies in rodents and NHPs using fenfluramine to release 5-HT have failed to consistently demonstrate sensitivity of most 5-HT_{1A} and 5-HT_{2A} tracers to 5-HT, despite microdialysis evidence that extracellular 5-HT levels rise dramatically in response to the pharmacological challenge. Recently, though, two radiotracers for the 5-HT_{1B} receptor, [^{11}C]AZ10419369 and [^{11}C]P943, have shown more promise. Both have proven sensitive to 5-HT release by fenfluramine and to blockade of SERT by SSRIs in anesthetized NHPs, with robust ΔBP_{ND} following fenfluramine administration that was similar for both tracers when comparable doses were administered, and moderate response to an SSRI. These studies are promising, but whereas both 5-HT_{1B} tracers have begun to be used in human subjects to measure 5-HT_{1B} availability, to date no studies have been published measuring 5-HT transmission in humans, and so it remains to be seen if effects are detected in awake humans with the available pharmacological challenges.

The only currently available approach for reducing 5-HT levels to unmask possible baseline binding to receptors is rapid tryptophan depletion (RTD). Tryptophan is the precursor of 5-HT. The only source of tryptophan is dietary and by having subjects ingest preparations that have high concentrations of other amino acids but not tryptophan, plasma levels decrease rapidly, leading, in turn, to reduced 5-HT in brain. Behavioral studies have shown that RTD induces symptoms of depression in remitted depression patients. Glucose metabolism as measured with PET and ^{18}FDG, CBF measured with PET, and ^{15}O water and neural activity measured with fMRI all suggest a detectable effect of RTD on brain function. Neurochemical imaging with PET has been more equivocal. Studies using [^{11}C]WAY100635 or [^{11}C]DASB failed to detect effects at 5-HT_{1A} receptors or SERT. Two studies using the 5-HT_{2A} tracers [^{11}C]MDL100907 or [^{18}F]Setoperone detected paradoxical effects, in which tracer binding was slightly reduced, rather than increased, in several brain regions following RTD. Thus, radiotracer studies to date have failed to definitively detect baseline levels of 5-HT in brain using depletion methods; it remains to be seen if the apparent increased sensitivity of 5-HT_{1B} tracers to the presence of 5-HT will prove useful in this context as well.

Serotonergic Target Engagement. PET imaging has been used to establish SERT occupancy necessary for SSRI efficacy. These studies have been remarkably consistent across patient groups, drugs, and radiotracers in demonstrating that therapeutically effective SSRI doses attain 80 percent or greater occupancy of SERT. Studies have included patients with MDD, panic disorder, OCD, and other conditions and have used both [^{11}C]McN5652 and [^{11}C]DASB as radiotracers. Tested drugs have included

citalopram (Celexa), fluoxetine (Prozac), paroxetine (Paxil), sertraline (Zoloft), venlafaxine (Effexor), and others. These findings provide a quantitative dimension to the understanding of SERT blockade as an antidepressive mechanism, and a benchmark for the drug development and clinical trial design. An interesting related set of studies provide an instructive example of the application of imaging studies in the drug development. The long latency to therapeutic effects of SSRIs is believed to be related to a period of reduced 5-HT release following an initial transient increase, due to inhibition of serotonergic cell firing in the DRN caused by stimulation of 5-HT_{1A} autoreceptors. The inhibition was thought to subside over the course of several weeks due to desensitization of 5-HT_{1A} autoreceptors. Thus, a strategy for reducing the latency would be to block 5-HT_{1A} receptors in the DRN. But while this might reduce autoinhibition in the DRN, systemic administration of a 5-HT_{1A} antagonist would also block postsynaptic 5-HT_{1A} receptors in the terminal fields, possibly interfering with antidepressant activity. Pindolol (Visken) is a β-adrenergic receptor blocker that is also a weak partial agonist at 5-HT_{1A} receptors, and in the 1990s, some preclinical and postmortem data indicated that it might bind preferentially to DRN autoreceptors, suggesting that augmentation of SSRI treatment with pindolol might decrease the latency and increase the efficacy of SSRIs. Double blind, placebo controlled clinical trials showed mixed results, with some reporting beneficial effects and others not. In the early 2000s, several PET studies were performed to measure occupancy of pre- and postsynaptic 5-HT_{1A} receptors by pindolol using [^{11}C]WAY100635 imaging. These studies showed that at the doses of pindolol administered as augmentation therapy in the clinical trials, probably less than 30 percent of DRN 5-HT_{1A} receptors were occupied, calling into question whether or not the tested doses provided enough blockade of autoreceptors to be effective. In the particular case of pindolol, higher doses could not be tested due to adrenergic effects, but these studies demonstrated the key role that imaging can play in quantifying target engagement by therapeutic agents. Among the several PET studies measuring pindolol occupancy, the most comprehensive one did observe higher occupancy in DRN than in cortex. Theoretical models have been suggested, but the mechanism for this difference has never been carefully characterized.

FUTURE DIRECTIONS

The field of radiotracer imaging with PET and SPECT continues to grow and expand, with the number of peer-reviewed publications using these methods increasing annually. Progress in certain areas, both in the technology and the biological science, may lead to expanded utility and adaptation of imaging. Manufacture of scanners or cyclotrons that are less expensive or more portable would facilitate even wider use. Some

manufacturers have already considered developing lines of lower cost PET instruments, for example, by utilizing BGO detector crystals, which have some less optimal properties but are also less expensive than LSO and other state of the art scintillators. Another direction being pursued is to make scanners more sensitive, therefore requiring less radioactivity, by increasing the length of the stack of detector rings. Sensitivity would be increased because the solid angle, the total fraction of a spherical surface surrounding the field of view accessible to the scanner, would be increased. The most immediate way to expand the scientific questions that can be addressed with PET or SPECT is the development of tracers for more, currently inaccessible target molecules or processes. Despite numerous attempts, progress has been slow in developing tracers for the glutamatergic and GABAergic transmitter systems, but clearly, direct observation of the major excitatory and inhibitory neurochemical systems would benefit all brain research. Finally, it is possible that eventually, neurochemical information currently obtained by PET and SPECT radiotracer imaging will be replaced by less invasive methods, for example, by the development of neurochemical probes that can be detected with MRI. But research in this area is in its early stages and for the foreseeable future, PET and SPECT neuroimaging methods remain essential translational tools for understanding the brain in health and disease, and to inform the drug development process.

REFERENCES

Abi-Dargham A, Gil R, Krystal J, et al. Increased striatal dopamine transmission in schizophrenia: confirmation in a second cohort. *Am J Psychiatry*. 1998;155:761–767.

Abi-Dargham A, Mawlawi O, Lombardo I, et al. Prefrontal dopamine D1 receptors and working memory in schizophrenia. *J Neurosci*. 2002;22:3708–3719.

Abi-Dargham A, Xu X, Thompson JL, et al. Increased prefrontal cortical D1 receptors in drug naive patients with schizophrenia: a PET study with [^{11}C]NNC112. *J Psychopharmacol*. 2012;26:794–805.

Bailey DL, Townsend DW, Malsey MN (eds): *Positron Emission Tomography: Basic science*. London: Springer-Verlag London Limited; 2005.

Brooks RA, Dichiro G. Principles of computer-assisted tomography (CAT) in radiographic and radioisotopic imaging. *Phys Med Biol*. 1976;21:689–732.

Buchert R, Thiele F, Thomasius R, et al. Ecstasy-induced reduction of the availability of the brain serotonin transporter as revealed by [^{11}C](+)McN5652-PET and the multi-linear reference tissue model: loss of transporters or artifact of tracer kinetic modelling? *J Psychopharmacol*. 2007;21:628–634.

Bushberg J, Siebert J, Leidholdt E, Boone J. *The Essential Physics of Nuclear Medicine*. Philadelphia: Lippincott Williams & Wilkins; 2012

Cherry S, Sorenson J, Phelps M. *Physics in Nuclear Medicine*. Philadelphia: Elsevier Saunders; 2003.

Farde L, Wiesel FA, Halldin C, Sedvall G. Central D2-dopamine receptor occupancy in schizophrenic patients treated with antipsychotic drugs. *Arch Gen Psychiatry*. 1988;45:71–76.

Farde L, Nordstrom AL, Wiesel FA, Pauli S, Halldin C, Sedvall G. Positron emission tomographic analysis of central D1-Dopamine and D2-Dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine - relation to extrapyramidal side-effects. *Arch Gen Psychiatry*. 1992;49:538–544.

Howes OD, Kambeitz J, Kim E, et al. The nature of dopamine dysfunction in schizophrenia and what this means for treatment. *Arch Gen Psychiatry*. 2012;69:776–786.

Hudson HM, Larkin RS. Accelerated image-reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging*. 1994;13:601–609.

*Innis RB, Cunningham VJ, Delforge J, et al. Consensus nomenclature for in vivo imaging of reversibly binding radioligands. *J Cereb Blood Flow Metab*. 2007;27:1533–1539.

Kegeles L, Abi-Dargham A, Frankle WG, et al. Increased synaptic dopamine in associative regions of the striatum in schizophrenia. *Arch Gen Psychiatry*. 2010;67:231–239.

Kish SJ, Lerch J, Furukawa Y, et al. Decreased cerebral cortical serotonin transporter binding in ecstasy users: a positron emission tomography/[(11)C]DASB and structural brain imaging study. *Brain*. 2010;133:1779–1797.

Lammertsma AA, Hume SP. Simplified reference tissue model for PET receptor studies. *Neuroimage*. 1996;4:153–158.

Laruelle M. Imaging dopamine transmission in schizophrenia. A review and meta-analysis. *Q J Nucl Med*. 1998;42:211–221.

Laruelle M. Imaging synaptic neurotransmission with in vivo binding competition techniques: a critical review. *J Cereb Blood Flow Metab*. 2000;20:423–451.

Laruelle M, Slifstein M, Huang Y. Relationships between radiotracer properties and image quality in molecular imaging of the brain with positron emission tomography. *Mol Imaging Biol*. 2003;5:363–375.

Martinez D, Narendran R, Foltin RW, et al. Amphetamine-induced dopamine release: Markedly blunted in cocaine dependence and predictive of the choice to self-administer cocaine. *Am J Psychiatry*. 2007;164:622–629.

Martinez D, Slifstein M, Broft A, et al. Imaging human mesolimbic dopamine transmission with positron emission tomography. Part II: amphetamine-induced dopamine release in the functional subdivisions of the striatum. *J Cereb Blood Flow Metab*. 2003;23:285–300.

Okubo Y, Suhara T, Suzuki K, et al. Decreased prefrontal dopamine D1 receptors in schizophrenia revealed by PET. *Nature*. 1997;385:634–636.

Parsey RV, Ogden RT, Miller JM, et al. Higher serotonin 1 A binding in a second major depression cohort: modeling and reference region considerations. *Biol Psychiatry*. 2010;68:170–178.

Savitz JB., Drevets WC. Neuroreceptor imaging in depression. *Neurobiol Dis*. 2013;52:49–65.

Shepp LA, Vardi Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging*. 1982;1:113–122.

Shrestha S, Hirvonen J, Hines CS, et al. Serotonin-1A receptors in major depression quantified using PET: controversies, confounds, and recommendations. *Neuroimage*. 2012;59:3243–3251.

Slifstein M, Laruelle M. Models and methods for derivation of *in vivo* neuroreceptor parameters with PET and SPECT reversible radiotracers. *Nucl Med Bio*. 2001;28:595–608.

Urban NB, Girgis RR, Talbot PS, et al. Sustained recreational use of ecstasy is associated with altered pre and postsynaptic markers of serotonin transmission in neocortical areas: a PET study with [(1)(1)C]DASB and [(1)(1)C]MDL 100907. *Neuropsychopharmacology*. 2012;37:1465–1473.

Volkow ND, Fowler JS, Wang GJ, et al. Decreased dopamine D2 receptors availability is associated with reduced frontal metabolism in cocaine abusers. *Synapse*. 1993;14:169–177.

Volkow ND, Wang GJ, Fowler JS, et al. Decreased striatal dopaminergic responsiveness in detoxified cocaine-dependent subjects. *Nature*. 1997;386:830–833.

▲ 1.18 Gene Mapping Investigations of Psychiatric Disorders

SCOTT C. FEARS, M.D., PH.D.

INTRODUCTION: PROGRESS AND PITFALLS IN PSYCHIATRIC GENETICS

The field of genetics essentially started from the rediscovery of Gregor Mendel's basic concepts at the turn of the 20th century, and has matured into an essential cornerstone not only of the biological sciences but also of all of medicine. The discovery of the basic structure and properties of DNA in the middle of the century led to an exponential acceleration in the current understanding of all aspects of the life sciences, including deciphering the complete sequence of the human genome, and those of myriad other species. Massive databases of such sequences now provide 21st century biologists with the task of decoding the functional significance of all this information. In particular, attention has turned to determining how sequence variations contribute to the phenotypic variation between species and between individuals within a species; in humans, it is hoped that discoveries about the relationship between genotypes and phenotypes will revolutionize the understanding of why and how some individuals but not others develop common diseases. This hope is particularly strong for psychiatry, as current knowledge of the pathogenic mechanisms of psychiatric disease remains sparse.

Since the publication of the previous edition of this chapter, the fruits of many decades of frustrating working in the field of psychiatric genetics have finally begun to be realized. For example, no confirmed genetic variants influencing risk for schizophrenia had been discovered at the time of writing the previous edition, but since then, over 100 risk variants have been identified and replicated in rigorously designed gene mapping investigations. The growing list of genes associated with psychiatric disorders promises to redefine the current understanding of the biological basis of mental disorders. The emerging catalog of gene mapping findings does not support the popular theories of the last 50 years that have focused

on alterations of monoaminergic systems as fundamental to the development of mood and psychotic disorders. Rather, genetic findings indicate that the etiologies of psychiatric disorders are linked to a complex network of partially overlapping cellular processes including neurodevelopment, synapse functioning, ion regulation, immunity/inflammation, and intracellular signaling. Further, genetic analyses are confirming what has long been suspected; that there is substantial heterogeneity within categorically defined diagnostic groups and significant genetic overlap between different categories of disorders. In contrast to traditional genetic approaches that aimed to identify one or a few rare genetic variants with a large impact on the development of disease (i.e., Mendelian genes), much of the recent success has been gained by focusing on the identification of common genetic variants, each of which contributes only a small increase in disease risk. There are still major issues to be resolved, for example, deciphering the relative contributions of common versus rare genetic mutations; however, as a whole, the field is finally able to begin building a empirically rigorous foundation upon which many of the deep unknowns regarding the pathophysiology of mental disorders can begin to take shape.

Genetic mapping studies aim to identify the genes implicated in heritable diseases, based on their chromosomal location. These studies are carried out by investigating affected individuals and their families through two approaches: linkage and association (Figs. 1.18–1 and 1.18–2). It is now straightforward to genetically map Mendelian traits (traits for which a specific genotype at one particular locus is both necessary and sufficient to cause the trait). Psychiatric diseases, however, do not follow simple Mendelian inheritance patterns but rather are examples of etiologically complex traits. Etiological complexity may be due to many factors, including incomplete penetrance (expression of the phenotype in only some of the individuals carrying the disease-related genotype), the presence of phenocopies (forms of the disease that are not caused by genetic factors), locus heterogeneity (different genes associated with the same disease in different families or populations), or polygenic inheritance (risk for disease increases only if susceptibility variants at multiple genes act in concert). Mapping a complex disorder involves several component steps, including definition of the phenotype to be studied, epidemiological studies to determine the evidence for genetic transmission of that phenotype, choice of an informative study population, and determination of the appropriate experimental and statistical approaches. This chapter will start with an overview of the epidemiological methods used to identify genetically influenced traits and disorders and then describe the various approaches that are used to map the specific genetic variants associated with complex traits. The final part of this chapter will review the progress to date in mapping genes for several key psychiatric disorders.

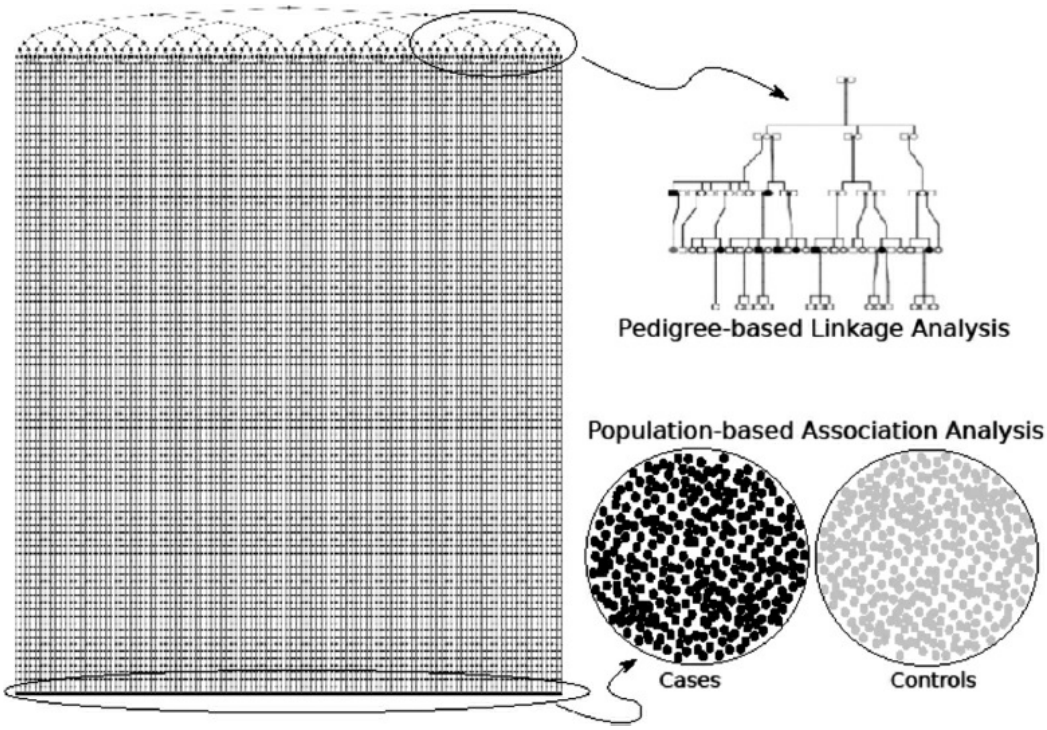

FIGURE 1.18–1. There are two basic approaches to mapping genes: linkage (top right) and associations (bottom right). All gene-mapping approaches are based on identifying genetic mutations that are inherited identical by descent (IBD). The left side of the figure shows a schematic representation of a founder at the top center of the figure and all of his/her descendants inheriting the risk allele IBD. Linkage analysis (top right) samples multiple generations of individuals within a family and aims to identify cosegregation of genetic markers of known location with the disease trait in affected individuals. Linkage analysis is optimal for detecting genes of large effect in family members relatively close to the origination of the genetic mutation (before natural selection results in the removal of the mutation). In contrast, association analysis (bottom right) samples generations very distant from the founder and attempts to identify enrichment of specific genetic variants within affected individuals compared to controls. Genetic mutations that have survived for many generations tend to confer a relatively small increase in disease risk (which is generally expected to lower fitness), which is reflected in the small effect sizes of variants identified by association studies.

Gene Mapping Strategies

	Linkage Analysis		Genome Wide Association	
	Pedigree Analysis	Affected Sib Pair Analysis	Case-Control	Family-Trios
		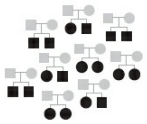	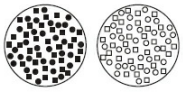	
Study Subjects	Multigenerational families with multiple affected individuals	Two or more affected siblings	Affected individuals and matched unaffected controls sampled from population	Affected individual and parents
Basic Idea	Identify genetic markers that cosegregate with disease phenotype	Identify chromosomal regions shared by siblings concordant for disease.	Tests for statistical association of alleles and disease in cases versus controls.	Tests for association using nontransmitted parental chromosome as control.
Strengths	1) Can detect rare variants of large effect. 2) Gains power by incorporating information about familial relationships into the model.	1) Robust to differences in genetic composition of study population. 2) Easier to collect clinical samples compared to special pedigrees. 3) Allows incorporation of enviromental data.	1) Can detect common variants of small effect. 2) Does not require collection of family data.	1) Can detect common variants of small effect. 2) Robust to problems of population stratification.
Limitations	1) Limited power to identify common variants of small effect. 2) Cost intensive.	1) Limited power to identify common variants of small effect.	1) Increased false-positive rate in the presence of population stratification. 2) Requires large sample sizes.	1) About two-thirds as powerful as case-control designs. 2) Difficult to collect samples for late onset diseases.

FIGURE 1.18–2. Genetic mapping approaches can be divided into those that rely on linkage analysis and those that rely on association analysis. Linkage studies can be further categorized as either focused on investigation of pedigrees or focused on investigation of sib pairs. Association studies can be categorized as either case-control or family-based. Some of the key features as well as advantages and disadvantages of these different approaches are shown in the figure.

EPIDEMIOLOGY

Genetic Epidemiological Approaches

Genetic epidemiological investigations provide quantitative evidence regarding the degree to which a given trait aggregates in families and, furthermore, can suggest to what degree such aggregation reflects a genetic contribution to the etiology of the trait. Family studies compare the aggregation of disease among the relatives of affected individuals compared to control samples. Because these studies do not differentiate between genetic and environmental contributions to such familial aggregation, they provide only indirect evidence regarding the heritability of a trait. Often these studies measure the relative risk (λ), defined as the rate of occurrence of a disease among specified categories of relatives of an affected individual divided by the rate of occurrence of the disease for the general population. A relative risk of >1 suggests a genetic etiology, and the magnitude of the measure gives an estimate of the genetic contribution to the disease. Relative risks can be calculated for sibling pairs, parent–offspring pairs, and various other types of family relationships. Likely modes of transmission can be assessed by comparing the degree of relative risk for each type of relationship. Multiple family studies have been carried out for

many of the major psychiatric disorders, including major depression, bipolar disorder, schizophrenia, and OCD. While these studies have consistently reported familial aggregation for all of these disorders, the degree of such aggregation has varied substantially across studies, largely reflecting differences in phenotype definition and how study samples were ascertained and assessed.

Twin studies examine the concordance rates of a particular disorder (the percentage of twin pairs where both twins have the disorder) in monozygotic (MZ) and dizygotic (DZ) twins. For a disorder that is strictly determined by genetic factors, the concordance rate should be 100 percent in MZ twin pairs (who share 100 percent of their genetic material) and 25 or 50 percent in DZ twin pairs (who are no more closely related than any siblings), depending on whether the disease is recessive or dominant, respectively. For a disorder where genetic factors play a role in disease causation but are not the exclusive cause of disease, the concordance rates should be greater for MZ twins than those for DZ twins. The higher the degree of concordance of MZ twins, the higher the trait heritability or the evidence for a genetic contribution to disease risk. When genetic factors do not play a role, the concordance rates should not differ between the twin pairs, under the simplifying assumption that the environment for MZ twin pairs is no more similar than that for DZ twin pairs. The twin studies that have been conducted for traits such as autism, bipolar disorder, and schizophrenia have consistently suggested high heritability and have therefore spurred efforts to genetically map loci for each of these conditions. Different twin studies may, however, generate varying estimates for the heritability of any given disorder. When evaluating the results of twin studies, it is therefore important to scrutinize how the phenotype was ascertained because, as with family studies, the different heritability estimates are likely due to differences in the mode of assessing and defining phenotypes. For example, early twin studies of psychiatric disorders often relied for their phenotypes on unstructured interviews by a single clinician. In contrast, modern studies generally utilize standardized assessments and review of diagnostic material by a panel of expert clinicians. Similarly, part of the apparent variation in heritability between different twin studies can be attributed to the fact that some studies employ narrow definitions of affectedness for a given phenotype, while other studies employ broader phenotype definitions (e.g., considering a twin with MDD to be phenotypically concordant with a cotwin diagnosed with bipolar disorder). Because of such differences in approach across studies, it is usually prudent to view such investigations as providing a rough estimate of the genetic contribution to trait variability. Nevertheless, even such estimates are useful in deciding which traits are likely to be mappable.

BASIC CONCEPTS OF GENE MAPPING

Introduction

Once genetic epidemiological studies of particular phenotypes have suggested that these phenotypes are heritable, genetic mapping studies are conducted to identify the specific genetic variants that contribute to the risk of the disorder. The goal of gene mapping is to identify the order of genetic loci along a chromosome and determine the distance between them. Prior to the discovery of DNA as the molecular basis of genes, mapping studies typically sought to identify patterns of transmission in families between disease phenotypes and enzymatic or physiological phenotypes like the ABO blood group to establish linkage. Current mapping studies start with an extensive set of DNA polymorphisms of known location and methodically traverse through the genome to identify the location of an unknown disease locus. There are two basic approaches to mapping genes: linkage and association (Figs. 1.18–1 and 1.18–2). Linkage analysis uses information from family members to identify loci that are transmitted together across generations more often than would be expected by chance. In contrast, association designs do not make explicit use of family information and aim to identify nonrandom co-occurrence of genetic variants of know location with disease traits. Linkage analysis is ideally suited to identify rare genetic variants that have a relatively large effect on the disease trait, whereas population-based association methods are designed to identify common genetic variants that have modest effects on the disease trait. The following sections will provide an overview of the basic concepts of gene mapping starting with the phenomena of recombination and linkage upon which all mapping methods are based.

Recombination and Linkage

All genetic mapping methods aim to identify disease-associated variants based on their chromosomal position and the principle of genetic linkage. All cells contain two copies of each chromosome (called homologs): one inherited from the mother and one inherited from the father. During meiosis, the parental homologs cross over, or recombine, creating unique new chromosomes that are then passed on to the progeny. Genes that are physically close to one another on a chromosome are genetically linked, and those that are farther apart or are on different chromosomes are genetically unlinked. Genes that are unlinked will recombine at random (i.e., there is a 50 percent chance of recombination with each meiosis). Genetic loci that are linked will recombine less frequently than expected by random segregation, with the degree of recombination proportional to the physical distance between them. The principle of linkage underlies the use of genetic markers, segments of DNA of known chromosomal location that contain variations or polymorphisms (described in more detail below). Strategies to map disease genes are based on identifying genetic marker

alleles that are shared—to a greater extent than expected by chance—by affected individuals. It is presumed that such sharing reflects linkage between a disease locus and a marker locus, that is, the alleles at both loci are inherited "identical by descent" (IBD), from a common ancestor, and, furthermore, that this linkage pinpoints the chromosomal site of the disease locus.

The evidence for linkage between two loci depends on the recombination frequency between them. Recombination frequency is measured by the recombination fraction (Θ) and is equal to the genetic distance between the two loci [1 percent recombination equals 1 centimorgan (cM) in genetic distance and, on average, covers a physical distance of about 1 megabase (mB) of DNA]. A recombination fraction of 0.5 or 50 percent indicates that two loci are not linked but rather that they are segregating independently. A LOD (logarithm of the odds ratio) score is calculated to determine the likelihood that two loci are linked at any particular genetic distance. The LOD score is calculated by dividing the likelihood of acquiring the data if the loci are linked at a given recombination fraction by the likelihood of acquiring the data if the loci are unlinked ($\Theta = 0.5$). This step gives an odds ratio, and the log (base 10) of this odds ratio is the LOD score. A LOD score can be obtained for various values of the recombination fraction, from $\Theta = 0$ (completely linked) to $\Theta = 0.5$ (unlinked). The value of Θ that gives the largest LOD score is considered to be the best estimate of the recombination fraction between the disease locus and the marker locus. This recombination fraction can then be converted into a genetic map distance between the two loci.

Linkage Disequilibrium

Linkage disequilibrium (LD) is a phenomenon that is used to evaluate the genetic distance between loci in populations rather than in families. When alleles at two loci occur together in the population more often than would be expected given the allele frequencies at the two loci, those alleles are said to be in LD. When strong LD is observed between two loci, it usually indicates that the two loci are sited in very close physical proximity to one another on a given chromosome and is useful in mapping disease susceptibility loci because one locus can be used to predict the presence of another locus. This predictability is important because current gene-mapping strategies are able to sample only a subset of the estimated 10 million common human polymorphisms. Because of the existence of LD, one can use data from a subset of genotyped polymorphisms to infer genotypes at nearby loci. Clusters of alleles that are in LD and inherited as a single unit are termed haplotypes. Thus, LD mapping "consolidates" genomic information by identifying haplotypes in populations that can then be used to infer IBD sharing among unrelated individuals.

There are several methods to measure the extent of LD. One of the most

commonly used measures of LD is r^2, a measure of the difference between observed and expected haplotype probabilities. Unlike D', another widely used measure of LD, r^2 values do not depend on the allele frequencies of the loci being assessed. A large r^2 value indicates that the observed frequency of association between two alleles is greater than that expected by chance; that is, the alleles are in LD. LD studies have traditionally been used to complement traditional pedigree analyses, for example, to hone in on a locus that has been mapped by linkage analysis. However, LD-based association analysis has become the method of choice for whole genome screens, particularly for diseases where traditional linkage studies have been unsuccessful. These studies have one great advantage over a traditional family analysis; because affected individuals are chosen from an entire population rather than from one or a few pedigrees, the number of potential subjects is limited only by the size of the population and the frequency of the disease. Maximizing the potential number of affected individuals that can be included in the analysis is extremely important for disorders where genetic heterogeneity or incomplete penetrance is likely to be factors.

Genetic Markers

Mapping studies, regardless of their type, depend on the availability of genetic markers. The most widely used markers are microsatellite markers (also called simple tandem repeats [STRs], or simple sequence length polymorphisms [SSLPs]) and single nucleotide polymorphisms (SNPs). SSLPs are stretches of variable numbers of repeated nucleotides two to four base pairs in length. These markers are highly polymorphic, as the number of repeat units at any given STR locus varies substantially between individuals. SNPs, as the name implies, are single base pair changes at a specific nucleotide; they are the most common form of sequence variation in the genome. SNPs are widely used for genetic mapping studies because they are distributed so widely across the genome and because they can be assessed in a high-throughput, automated fashion. Other forms of genetic variation that have been investigated for use as genetic markers include small insertion or deletion polymorphisms, termed indels, that generally range between 1 and 30 base pairs and CNVs, which can refer to either deletions or duplications. Recent genome-wide surveys have revealed that CNVs are common and can range in length from several base pairs to several million base pairs. CNVs may contribute to chromosomal recombination and rearrangements, thereby playing an important role in generating genetic diversity, and also, as many of these variants are sizable, it is hypothesized that they may significantly influence the expression of genes that encompass or are adjacent to the variant.

MAPPING STRATEGIES

The genetic variants that contribute to disease susceptibility can be roughly categorized into those that are highly penetrant and those that are of low penetrance. High-penetrance variants by definition have a large effect on phenotype, and therefore identifying these variants usually provides fundamental insights into pathobiology. Because individuals carrying high-penetrance variants have a high probability of expressing a disease phenotype, such variants tend to be rare and to segregate in families and are generally most powerfully mapped using pedigree-based approaches (Figs. 1.18–1 and 1.18–2). In contrast, low-penetrance variants have a relatively weak effect on phenotype, and therefore identifying individual low-penetrance variants may, at least initially, provide relatively little new biological knowledge. However, because of their small effects, such variants are typically common in the population, and therefore identifying them may add to the understanding of disease risk in the population as a whole. It is not expected that these variants will segregate strongly with the disease phenotype in pedigrees, efforts to identify them focus on population samples. In the following sections, two common family-based linkage strategies (pedigree and sib-pair analyses) are described followed by an overview of population and family-based association analyses.

Pedigree Analysis

A pedigree analysis, which is conducted in multigenerational families, consists of scanning the genome or a portion of the genome with a series of markers in one or more affected pedigrees, calculating a LOD score at each marker position, and identifying the chromosomal regions that show a significant deviation from what would be expected under independent assortment. The primary goal of pedigree analysis is to determine if two or more genetic loci (i.e., a genetic marker of known location and the unknown disease loci) are cosegregating within a pedigree.

Following the successful application of pedigree analysis to map Mendelian disorders such as Huntington disease, many investigators adopted this strategy for mapping psychiatric disease genes with, at best, mixed success. In the late 1980s and mid-1990s, several pedigree-based studies reported the mapping of susceptibility loci for Alzheimer disease, bipolar disorder, and schizophrenia. Although the linkage findings for three Alzheimer disease loci were relatively quickly replicated, the findings reported for bipolar disorder and schizophrenia were ultimately determined to have been false positives. While a number of different explanations have been proposed for the failure of pedigree-based approaches to map psychiatric loci, most investigators now recognize that these studies were generally drastically underpowered considering the apparent etiological complexity of psychiatric disorders.

Pedigree analysis in psychiatry has increasingly turned toward an application for which it is more appropriately powered, namely, the

mapping of quantitative trait loci (QTLs). QTLs are defined as genetic loci that contribute to the variation in continuously varying traits (as opposed to categorical traits such as disease diagnoses). QTLs are typically loci of small effect that only contribute to a portion of the observed variance of a trait in the population. It is now generally accepted that, using analytical methods developed in the late 1990s, it may be possible to use pedigree studies to map a wide range of quantitative traits that are relevant for understanding psychiatric disorders including brain and behavioral phenotypes. Several such studies are now being undertaken, typically with multiple phenotypes being assessed in each individual in the pedigree. A more detailed discussion of quantitative phenotypes is provided after the overview of population-based mapping strategies.

Sib Pair Analysis

Affected sib pair (ASP) analysis, first proposed in 1935, became widely used during the 1990s for the genetic mapping of complex traits, including many psychiatric disorders. Sib pair analysis examines the frequency with which sibling pairs concordant for a trait share a particular region of the genome compared with the frequency that is expected under random segregation. Sib pair analysis is based on the fact that siblings share approximately 50 percent of their genomes IBD. Therefore, if a set of unrelated sib pairs affected with a given trait shares a particular area of the genome at a frequency significantly greater than 50 percent (the proportion of sharing expected under conditions of random segregation), then that area of the genome is likely to be linked to the trait in question. In this method, siblings are genotyped, and population frequencies and parental genotypes are used to estimate the proportion of genes shared IBD at each site for each sib pair. The linkage analysis then compares those pairs concordant and discordant for each locus.

Like pedigree studies, ASP studies have more power to locate genes of large effect than genes of small effect. This limitation can be partially addressed by a two-tiered design that incorporates additional markers or family members after an initial linkage study in affected siblings or by increased sample size. It generally requires less effort to identify and assess even large sets of affected sibs than to identify and assess all members of extended pedigrees, particularly when investigators can take advantage of data repositories that include samples and phenotype data from sib pairs ascertained from multiple sites. For example, the US NIMH maintains such repositories for sizable collections of sib pairs affected with schizophrenia, bipolar disorder, autism, and Alzheimer disease. An additional benefit of the ASP design is that it allows for the incorporation of epidemiological information, permitting the simultaneous examination of environmental and gene–environment interactions.

Association Studies

In the past few years, there has been increasing acceptance of the notion that association studies are more powerful than linkage approaches for mapping the loci of relatively small effect that are thought to underlie much of the risk for complex disorders. While linkage studies attempt to find cosegregation of a genetic marker and a disease locus within a family or families, association studies examine whether a particular allele occurs more frequently than expected in affected individuals within a population. As noted previously in this chapter, mapping genes using association studies is based on the idea that certain alleles at markers closely surrounding a disease gene will be in LD with the gene; that is, these alleles will be carried in affected individuals more often than expected by random segregation, because they are inherited IBD.

There are two common approaches to association studies (Fig. 1.18–2), case-control designs and family-based designs, which typically investigate trios (mother, father, and an affected offspring). In a case-control study, allele frequencies are compared between a group of unrelated affected individuals and a matched control sample. This design is generally more powerful than a family-based design, as large samples of cases and controls are easier to collect than trios and are less expensive as they require the genotyping of fewer individuals. Case-control samples may be the only practical design for traits with a late age of onset (such as Alzheimer disease) for which parents of affected individuals are typically unavailable. The main drawback of the case-control approach is the potential problem of population stratification; if the cases and controls are not carefully matched demographically, then they may display substantial differences in allele frequency that reflect population differences rather than associations to the disease.

Family-based association studies are designed to ameliorate the problem of population stratification. In this design, the nontransmitted chromosomes (the copy of each chromosome that is not passed from parent to child) are used as control chromosomes, and differences between allele frequencies in the transmitted and nontransmitted chromosomes are examined, eliminating the problem of stratification, as the comparison group is by definition genetically similar to the case group. Although more robust to population stratification than a case-control study, family-based studies are only about two-thirds as powerful using the same number of affected individuals, as noted previously.

Until recently, it was not practical to conduct association studies on a genome-wide basis, as relatively few SNPs were available. Therefore, association studies focused on testing one or a few markers in candidate genes chosen on the basis of their hypothesized function in relation to a given disease. Recently, however, as a result of international efforts that

have identified millions of SNPs distributed relatively evenly across the genome and the development of technology for genotyping them relatively inexpensively, genome-wide association (GWA) studies are now a reality. Such studies hold much promise for the identification of common variants contributing to common diseases. Early GWA studies in psychiatric disorders failed to identify any risk variants; however, recent studies have made use of very large samples (in some cases up to ~40,000 cases and controls), providing further support for the hypothesis that underpowered study designs bear much of the responsibility for the disappointing results of previous psychiatric genetic investigations.

Statistical Considerations

Scientists in other biomedical research fields are often surprised by the apparently high level of statistical evidence that geneticists require to consider a linkage or association result to be significant. Most simply, this requirement can be thought of in terms of the very low expectation that any two loci selected from the genome are either linked or associated with one another. The likelihood that any two given loci are linked (i.e., the prior probability of linkage) is expected to be approximately 1:50, based on the genetic length of the genome. To compensate for this low prior probability of linkage and bring the posterior (or overall) probability of linkage to about 1:20, which corresponds to the commonly accepted significance level of $P = .05$, a conditional probability of 1,000:1 odds in favor of linkage is required, corresponding to the traditionally accepted LOD score threshold of 3. This generally provides an acceptable false-positive rate (Fig. 1.18–3), but some false-positive findings have exceeded even this threshold.

Geneticists generally assume that the expectation that any two loci in the genome are associated with one another is even lower than that of their being in linkage, and typically a P value of less than about 10^{-7} is considered to indicate "genome-wide significance." This standard essentially discounts the prior probability that some investigators assign to variants in candidate genes chosen on the basis of their hypothesized functional relevance to a given disorder or trait. GWA studies are now replicating associations with very low P values for a wide range of complex traits, while the vast majority of candidate gene associations (which usually report as significant much higher P values) remain unreplicated. It is therefore increasingly apparent that genome-wide levels of significance are appropriately applied to all initial association studies for a given trait.

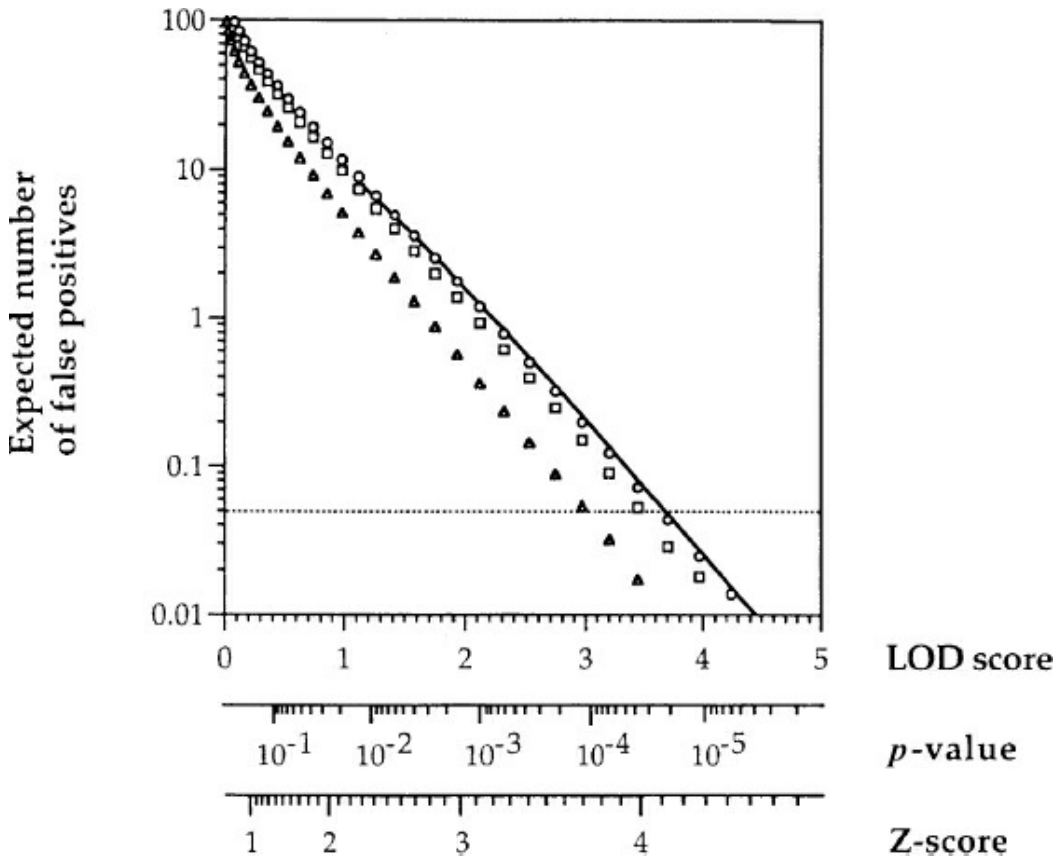

FIGURE 1.18–3. Number of false positives expected in a whole genome scan for a given threshold of logarithm of odds (LOD) score. Solid line represents the expectation for a perfect genetic map. Symbols represent the results for 100 sib pairs using genetic maps with markers spaced every .1 cM (*circles*), every 1 cM (*squares*), and every 10 cM (*triangles*). The dotted line indicates the 5 percent genome-wide significance level. Lander, E. and Kruglyak, L. Genetic dissection of complex traits: guidelines for interpreting linkage results, *Nature Genetics.* 1995;11(3):241–247.

THE PHENOTYPE

The generally disappointing results of psychiatric genetic mapping studies have focused increasing attention on the problem of defining and assessing phenotypes for such studies. Most psychiatric mapping studies to date have relied on categorical disease diagnoses, as exemplified by the *Diagnostic and Statistical Manual* (DSM) classification scheme. Criticisms of this approach rest on two arguments. First, diagnosis of psychiatric disease depends on subjective clinical evaluation, a fact that underscores the difficulty in ascertaining individuals who can be considered definitely affected with a given disease. Second, even when a psychiatric diagnosis can be established unambiguously, the menu-based system used for psychiatric classification provides the possibility that any two individuals affected with a given disorder may display largely nonoverlapping sets of symptoms, likely reflecting distinct etiologies. Concern that the diagnosis-based approach to phenotyping may represent one of the chief obstacles to the genetic mapping of psychiatric phenotypes has generated considerable interest in mapping heritable traits known to demonstrate continuous

variation in the population. Continuous measures that are hypothesized to be related to psychiatric disorders include biochemical measures (e.g., serum or CSF levels of neurotransmitter metabolites or hormones), cognitive measures, personality assessments, structural or functional brain images, biophysical markers such as responses to evoked potentials, or molecular assays such as gene expression profiles. Key features of categorical and continuous phenotyping strategies are shown in Figures 1.18–4 and 1.18–5, and each is discussed in more detail below.

Categorical Phenotypes

The most commonly used categorical phenotypes in psychiatry are DSM diagnoses. Some studies focus on a single DSM diagnosis, while other studies include individuals with a range of different diagnoses. The latter approach is typically used for disorders that are hypothesized to represent a single disease spectrum, such as mood disorders. Using the categorical approach, it is important to be able to classify subjects as unambiguously as possible. Several strategies are used to accomplish this goal. The first strategy involves deciding on the appropriate diagnostic criteria for the study in question and deciding how these criteria will be applied to individuals in the study. One way of standardizing the procedures used to identify and assess potential study subjects is to use only experienced clinicians in the diagnostic process and to train them in the administration of the instruments and the diagnostic criteria to be employed. Additionally, a "best estimate" procedure and/or a consensus diagnosis are frequently used. The best estimate process involves making use of every piece of available information, including medical records, interviews, and videotapes, to arrive at a diagnosis. For a consensus diagnosis, two or more diagnosticians independently review the material and make a diagnosis for each individual. The diagnoses are then compared, and individuals for whom an agreement in diagnosis cannot be reached are not entered as "affected" into the study.

A well-designed study makes use of all available information about the genetic epidemiology of the disorder to choose a sample of affected individuals to study. It is often the case that a subset of families carries the disorder in what appears to be a simple Mendelian pattern, while the inheritance pattern is less clear for other families or groups. In a disorder where there are likely to be multiple genes contributing to the phenotype, it makes sense to begin with a study sample where there may be major loci. Redefining the disease phenotype can often simplify the mapping process by identifying such groups or families. For example, in the search for a genetic defect for Alzheimer disease, the process was advanced enormously by limiting the study population to those individuals who had early age of onset (before age 65); the early onset trait segregated in an autosomal dominant fashion. Other ways of redefining the phenotype include focusing

on factors such as ethnic background, treatment response, symptom severity, or the presence of comorbid disorders.

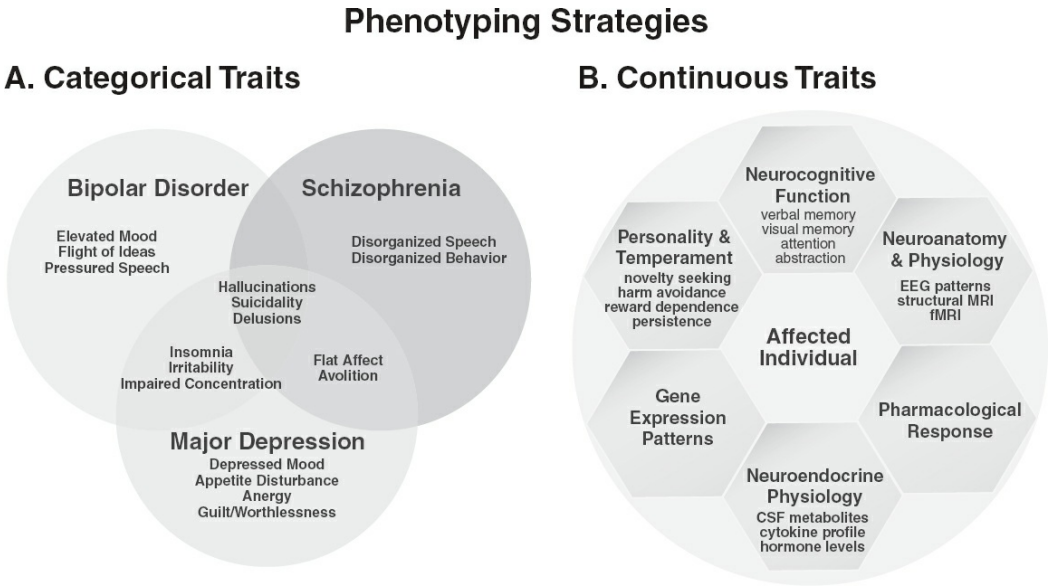

FIGURE 1.18–4. Two alternate schemes for conceptualizing psychiatric phenotypes. **A:** Categorical Traits as conceptualized by the *Diagnostic and Statistical Manual* (DSM) represent a "menu-based" approach to psychiatric disorders. Individuals are assessed for a checklist of signs and symptoms that are then used to categorize the individual as "affected" according to a specific diagnosis. Not all symptoms are present in samples of individuals who carry a particular DSM diagnosis, and many of these symptoms occur across diagnostic boundaries, as illustrated in this Venn diagram. DSM phenotypes therefore probably represent etiologically heterogeneous categories. **B:** Alternatively, in the Continuous Traits model, "affectedness" can be conceptualized in terms of an expectation that an individual will demonstrate extreme values on a set of continuous measures that correlate with psychopathology and thus are hypothesized to underlie the disorder (as illustrated by examples of six different types of measures shown in the hexagon). Such measures may also be associated with particular components of categorical phenotypes, such as those depicted in the Venn diagram. The justification for using continuous measures as the phenotypes for genetic mapping studies is that they are considered etiologically simpler and more reliably assessed compared to categorical phenotypes. In addition, mapping such traits combines information from all members of the study population (affected and unaffected individuals alike), which adds considerably to power.

Narrowing the phenotype using the approaches discussed above may increase the chances of finding a genetic defect in complex diseases, but it can also greatly reduce the power of the study by limiting the number of available affected individuals. For this reason, it has been argued that for some disorders broadening the phenotype is an appropriate strategy. The suggestion is that for some complex diseases the phenotype of interest may represent the extreme end of a spectrum and that to have enough power to map genes other phenotypes within the spectrum must also be included. For example, mapping studies of bipolar disorder might include as affected individuals with MDD as well as those individuals diagnosed with bipolar disorder.

Although the two approaches of narrowing the disease phenotype and broadening the disease phenotype may seem to be mutually exclusive, many groups studying complex disorders have incorporated both approaches into their study designs. One way to do this is to create stratified diagnostic categories, ranging from a narrow diagnostic category to a broad diagnostic category, and test for genetic linkage under each of these schemas. Some investigators argue that for complex diseases that are part of a spectrum, this strategy decreases the rate of false negatives, that is, of missing an existing linkage because of misspecification. Others argue that using several models and picking the one that gives the highest scores greatly increase the rates of false positives, that is, of identifying an area of linkage where none exists. One problem that clearly exists with the use of multiple diagnostic categories is that as more models are used (and therefore more statistical tests are performed); increasingly stringent levels of evidence are required to consider a result significant.

While categorical phenotypes remain the mainstay of psychiatric genetic studies, the limitations of DSM nosology as the basis of phenotyping for genetic studies are becoming clear. Genetic investigations are focusing increasingly on traits that may be components of one or more DSM diagnostic categories. For example, there is growing evidence that genetic susceptibility to psychosis, broadly defined, contributes to both severe bipolar disorder and schizophrenia, and a number of investigative approaches are being employed to attempt to identify genes that underlie such susceptibility and even to explore possible etiological relationships between psychiatric and nonpsychiatric disorders. For example, bioinformatics models have been employed to investigate medical records databases and have uncovered extensive pairwise correlations among a diverse list of psychiatric disorders, neurological disorders, autoimmune disorders, and infectious diseases. Eventually, the results of such model-fitting experiments may provide a framework to design more powerful linkage and association studies that can search for alleles that contribute to susceptibility to multiple disorders.

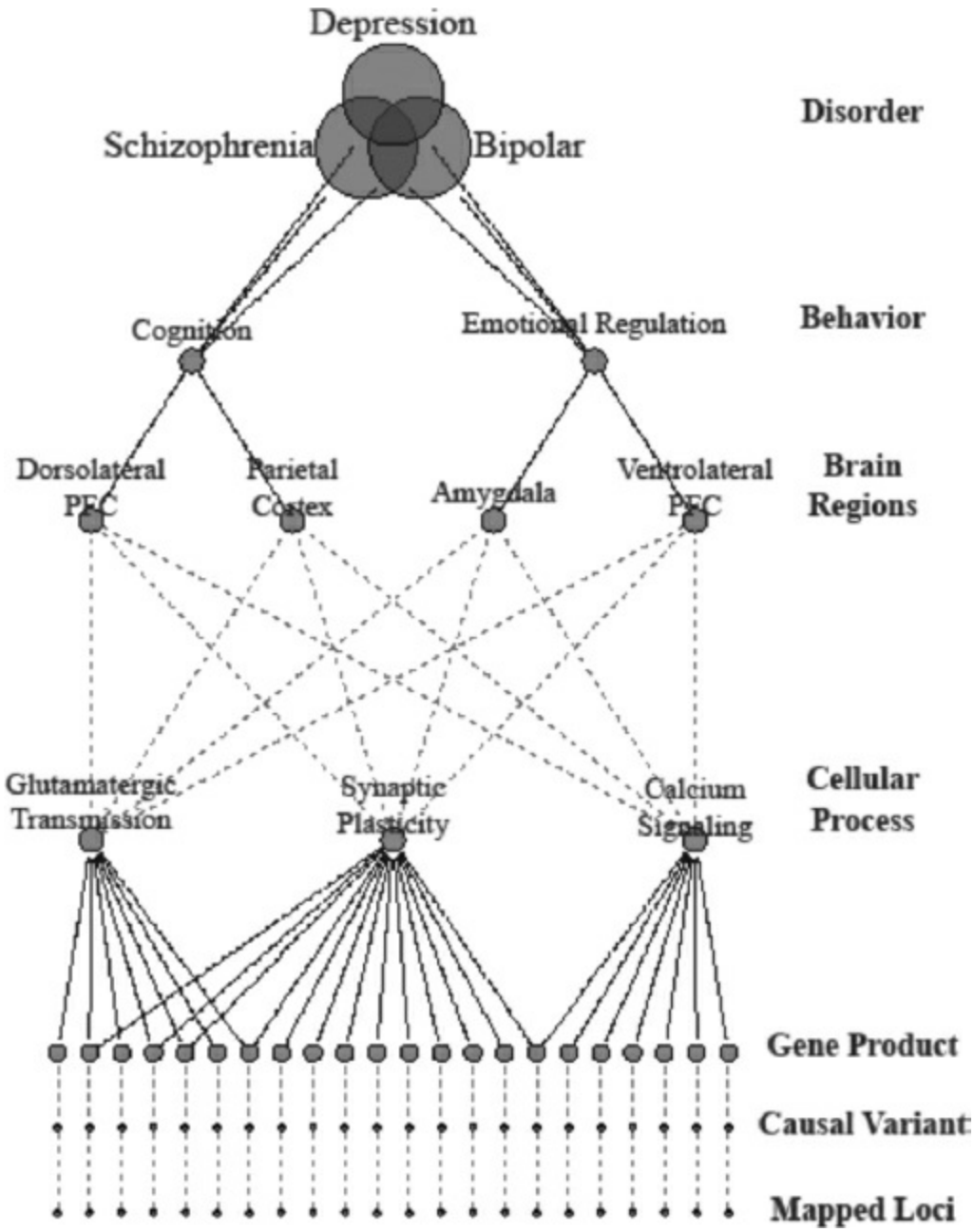

FIGURE 1.18–5. The endophenotype (also known as intermediate phenotype) strategy aims to map genes associated with quantitative measures acquired from multiple levels of physiology in order to characterize the biological pathways that cause psychiatric disorders. In contrast to conventional studies that map genes associated with dichotomous disease categories, the endophenotype method can be implemented across diagnostic categories to identify the differences/commonalities in biological systems that contribute to psychiatric disorders. Because intermediate phenotypes are "closer to genes" compared to diagnostic categories, they are expected to have more power in gene mapping studies.

Continuous Phenotypes

Because of the difficulties experienced in genetic mapping of categorical diagnoses, neurobehavioral geneticists are increasingly focused on investigating quantitative traits that are hypothesized to underlie a particular psychiatric diagnosis and that may be simpler to genetically map.

The rationale for efforts to map such alternative phenotypes, termed endophenotypes or intermediate phenotypes, is that the genes identified through such efforts may provide clues regarding the biological pathways that are relevant to understanding a particular disorder or multiple disorders that may have overlapping etiological pathways (Fig. 1.18–5). Several features characterize useful endophenotypes. First, they should be state independent; that is, they should not fluctuate as a function of the disease course or medication treatment and should show adequate test–retest stability. Second, they should be heritable; that is, there should be evidence that genetic factors are responsible for a substantial proportion of the variability of the trait within the population. Third, the endophenotype should be correlated with the disease under investigation; that is, different values of the trait measure are observed in patients compared to unrelated control subjects.

Measures of brain structure and function are now under investigation as endophenotypes for psychiatric disorders. For example, several features of brain morphometry (as assessed by MRI) are highly heritable (in the range of 60 to 95 percent) including total brain volume, cerebellar volume, gray and white matter density, amygdala and hippocampal volume, and regional cortical volume. Several studies show that brain structural features that are correlated in clinical samples with disorders such as schizophrenia or bipolar disorder are also abnormal in relatives of affected individuals. Physiological measures of brain activity that have been employed as candidate endophenotypes for psychiatric disorders include electroencephalography (EEG) patterns. In addition to brain measures, behavioral traits relevant to psychiatric disorders are an active field of gene mapping studies including neurocognition, impulsivity, emotional regulation, and temperament (i.e., neuroticism).

Animal Models

In contrast to categorical phenotypes, endophenotypes can be more straightforwardly related to phenotypes that can be assessed in animal models. Studies of genetic variations that affect circadian rhythms provide a good example. Variations in circadian rhythms have long been recognized as important features of mood disorders, and quantitative assessments of activity patterns have been proposed as endophenotypes for such disorders. Numerous studies in animal models have demonstrated that genetically controlled biological clocks determine circadian activity and that variations in clock genes are associated with variations in such activity from bacteria to humans. Genetic mapping efforts in fruit flies starting in the early 1970s resulted in the identification of at least seven “clock genes,” beginning with *period*. Subsequent studies showed that the homologs of several of these genes play essential roles in regulating mammalian circadian rhythms. Genetic mapping studies in mice also have identified previously unknown

circadian rhythm genes, beginning with the discovery and characterization in the early 1990s of *clock*. These genetic discoveries have not only explicated the cellular networks and neurophysiological circuits responsible for the control of mammalian circadian rhythms but have also generated animal models that may shed light on the pathobiology of psychiatric syndromes such as bipolar disorder. For example, mice carrying a targeted mutation in *clock* demonstrate abnormal activity patterns, such as hyperactivity and decreased sleep, which are apparently modified by administration of lithium.

PROGRESS IN THE GENETIC DISSECTION OF SPECIFIC DISORDERS

Taken as a whole, the progress in identifying susceptibility genes for psychiatric disorders has been slower compared to that observed for nonpsychiatric disorders. The final sections of this chapter will review the progress that has been made in identifying the genetic underpinnings of several specific psychiatric disorders. Alzheimer disease represents the most successful application of gene-mapping strategies to complex neurobehavioral disorders, and the section on this disease provides an example of how genetic linkage studies add to understanding the pathogenesis of a complex trait. Gene mapping studies of autism provide examples of genetic investigations of syndromes that have relatively simple inheritance patterns and demonstrate how these studies have provided starting points for investigations of more complex autism spectrum disorders (ASDs). Finally, recent progress identifying risk variants for schizophrenia, bipolar disorder, and major depression is used to illustrate the exciting directions that are beginning to revolutionize the current understanding of the etiology of these complex disorders.

ALZHEIMER DISEASE

Alzheimer disease provides an excellent example of the power of genetics to elucidate the complex biology of a neuropsychiatric disorder. Alzheimer disease is a well-defined form of dementia characterized by progressive impairment of memory and intellectual functioning. The clinical signs and symptoms, although characteristic, are not limited to Alzheimer disease but are also found in several other types of dementia. For this reason, the diagnosis of Alzheimer disease can only be confirmed histopathologically at autopsy. The presence of senile plaques (made up of a core of β-amyloid fibrils surrounded by dystrophic neurites), tau-rich neurofibrillary tangles, and congophilic angiopathy in the brain parenchyma and associated blood vessels are pathognomonic for Alzheimer disease.

A variable age of onset has been noted for Alzheimer disease, ranging from as early as age 35 to as late as age 95. Family history is the second greatest risk factor after age for the disease. The concordance rate for Alzheimer disease in MZ twin pairs is about 50 percent, indicating a

moderately strong genetic contribution to disease risk. It is now evident from a wide range of genetic studies that Alzheimer disease can be divided into two broad categories: familial forms, which account for about 5 percent of Alzheimer disease cases and are characterized by early onset and autosomal dominant inheritance with high penetrance; and sporadic forms, in which the genetic contribution is hypothesized to be similar to that characterizing other common neuropsychiatric diseases.

The search for the genetic basis of familial Alzheimer disease began with traditional linkage studies. First, an investigation of a candidate locus on chromosome 21 in humans identified mutations in the *APP* gene in a small number of families in which significant linkage had previously been observed to markers from this region. Transgenic mice with different *APP* mutations were created and have been shown to produce β-amyloid deposits and senile plaques as well as to show synapse loss, astrocytosis, and microgliosis, all part of the pathology of Alzheimer disease. Mutations in the genes that encode β-APP all lead to an increase in the extracellular concentration of longer fragments of β-amyloid (Aβ42). Most of the strains of transgenic mice with mutations in APP exhibit increased rates of behavioral changes and impairment in several memory tasks, indicating dysfunction in object-recognition memory and working memory among others. These findings represent striking evidence that mutations in the β-amyloid gene are indeed responsible for at least some of the histopathological elements of Alzheimer disease.

Even as the above findings were being reported, it was clear that mutations in the β-amyloid gene could not completely explain the etiology and pathology of Alzheimer disease, not least because it was shown that linkage to chromosome 21 was excluded in most early onset Alzheimer disease families. Additionally, no neurofibrillary tangles are observed in most of the different β-amyloid transgenic mice. The subsequent search for the genetic underpinnings of Alzheimer disease using genome-wide linkage analysis of early onset Alzheimer disease families resulted in the identification of two additional Alzheimer disease susceptibility genes: *presenilin-1* (*PS-1*) on chromosome 14q24.3 and *presenilin-2* (*PS-2*) on chromosome 1q. PS-1 and PS-2 are integral transmembrane proteins with at least seven transmembrane domains. The *presenilins* are aspartyl proteases that carry out γ-secretase cleavage of *APP* to generate β-amyloid. Similar to mutations in *APP,* genetic variation in *presenilins* causes an increase in the expression of longer fragments of β-amyloid.

These findings emphasize one of the strengths of using family-based linkage analysis. Pedigree-based studies are especially suited to identify highly penetrant disease genes that serve critical roles in important biological processes. Mutations in *APP* and *presenilin* are rare, but different mutations in these genes have occurred in different founders; to date 185 mutations have been identified in *PS-1,* 13 mutations for *PS-2,* and

24 mutations for *APP*. Virtually all of these mutations are transmitted in an autosomal dominant and fully penetrant manner. The convergence of different mutations within the same genes clearly confirms the involvement of these genes in the disorder. Although the mutations in *APP* and *presenilin* are rare, research into the biology of the expressed proteins has provided key insights into the pathophysiology of dementia. Because these highly penetrant mutations elucidate important biological functions, they also provide a firm ground to design therapeutic interventions. For example, β-amyloid "vaccines" designed to induce an immunogenic response to pathogenic amyloid are now in advanced clinical trials. Additionally, γ-secretase modulators have been developed to target dysfunction of the *presenilin* gene products. Unlike the current psychopharmacological treatments for Alzheimer disease that nonspecifically target cholinergic and glutamatergic systems, the β-amyloid vaccines and γ-secretase modulators specifically treat the causes of Alzheimer disease by generating an immune response or altering the processing of *APP* and may actually prevent or reverse the deposition of senile plaques.

Sporadic and Late Onset Alzheimer Disease

Mutations in *APP, PS-1,* or *PS-2* are present in a majority of familial cases of early onset Alzheimer disease but do not account for sporadic or familial late onset Alzheimer disease. For this reason, investigators turned to other approaches to search for the evidence of linkage in a large number of small families with late onset Alzheimer disease. In 1991, the results of a nonparametric linkage study using 36 markers in late onset Alzheimer disease families provided evidence for a susceptibility gene on the long arm of chromosome 19. In 1993, association studies revealed that the e4 allele of the *apolipoprotein E* gene was strongly associated with late onset Alzheimer disease and that this association almost certainly was responsible for the previously observed linkage signal on chromosome 19. There are three known alleles of this gene—e2, e3, and e4. These three alleles are distinguished by combinations of amino acids at residues 112 and 158; e2: Cys_{112}/Cys_{158}, e3: Cys_{112}/Arg_{158}, e4: Arg_{112}/Arg_{158}. In most populations, the e3 allele is the most common. However, in familial late onset Alzheimer disease, the incidence of e4 is approximately 50 percent, and in sporadic late onset Alzheimer disease it is 40 percent, compared with about 16 percent in normal controls. Inheritance of a single copy of e4 increases the risk for the disease by 4-fold and inheritance of two copies of the allele increases the risk to 10-fold. Epidemiological studies suggest that between 30 and 60 percent of late onset Alzheimer disease cases have at least one *apoE-e4* allele. The e4 genotype appears to be a more important risk factor for Alzheimer disease in populations of European and Asian

origin when compared with populations of African origin. The *apoE* gene has a well-characterized role in the metabolism and transport of lipids and is believed to play a role in the clearance of β-amyloid from the brain. Overall, the association of *apoE-e4* with Alzheimer disease remains probably the strongest association yet identified for a common human disease.

The establishment of *apoE-e4* as a susceptibility allele for late onset Alzheimer disease has led to the search for additional alleles that might interact with *apoE-e4* to modify disease risk. In 2007, investigators used GWA strategies (in histologically confirmed cases and controls) to identify *GAB2* (GRB-associated binding protein 2) as an additional risk allele in *apoE-e4* carriers (but not in Alzheimer disease patients who were not e4 carriers). Initial studies suggest that carriers of both *apoE-e4* and *GAB2* risk alleles have an almost 25-fold greater risk for Alzheimer disease than individuals who do not carry either risk allele. *GAB2* appears to modulate β-amyloid production through its interaction with GRB2 (growth factor receptor binding protein 2), which physically interacts with APP and the presenilins.

Since the initial GWA studies that identified *GAB2,* there have been more than a dozen additional GWA studies and together have identified more than 10 confirmed Alzheimer risk genes. Unlike *APP* and the *presenilins,* these genes have small effect sizes (odds ratio < 1.2), and are relatively common in the general population. Their role in the pathophysiology is less clear, though current animal models and cellular studies have expanded the understanding of the basic biology of the risk genes. Some of the risk genes identified in GWA studies have been linked to pathways involving β-amyloid processing. For example, the *ATXN1* gene product modulates γ-secretase levels and impacts cleavage of APP and murine knockout model of *ATXN1* shows cognitive impairment by age 6 months. Other genes identified by GWA studies appear to influence the risk of Alzheimer disease by impacting immune and inflammation pathways including factors that regular cellular interactions required for immune reactions (e.g., CD33) and the factors involved in the inflammatory process (e.g., *CR1,* the receptor for the complement molecule C3b). These immune/inflammatory processes have been postulated to mediate the neurotoxic response to β-amyloid.

Summary

Progress in the field of Alzheimer research has achieved significant momentum, and there are now more than a dozen genes implicated in the pathogenesis of this disorder. Linkage studies of rare familial forms of Alzheimer disease led to the discovery of high-penetrance variants that have had a profound impact in the understanding of Alzheimer disease pathogenesis and on the basic understanding of a wide range of cellular

processes within the CNS. Association studies have unequivocally identified lower-penetrance variants that together explain much of the genetic contribution to disease risk at the population level. Taken together, the risk genes converge on a few basic biological pathways: processing of β-amyloid and immunity and inflammation. These findings have provided the promise of two types of medical breakthroughs. New therapies are in development that target the molecular pathways identified through these studies. In addition, the emerging picture of genetic risk for common forms of Alzheimer disease suggests that it may soon be possible to focus prevention and early intervention strategies on individuals who are at high risk.

AUTISM

Autism is a severe neurodevelopmental disorder that was originally described as "enclosure in one's self" by Leo Kanner in 1943, and is characterized by three primary features: impaired language and communication, abnormal or impaired social interaction, and restricted, repetitive, and stereotyped patterns of behavior. Understanding of the etiology of autism has proceeded slowly, but there is now convincing evidence that alterations in specific cellular and molecular neurodevelopmental pathways are important in its etiology. There is particularly strong evidence for a genetic contribution to the risk of autism and ASDs. The sibling recurrence risk for autism and/or ASD is between 2 and 6 percent. Given a population prevalence of about 1 in 2,000 (.04 percent), this means that the siblings of autistic individuals are approximately 50 to 100 times more likely to develop autism than a person in the general population. Twin studies of autism not only show an extraordinarily high heritability (as demonstrated by MZ twin concordance of 80 to 92 percent) but also demonstrate the genetic complexity of these disorders, with the DZ twin concordance rate of 1 to 10 percent suggesting a highly multigenic mode of inheritance.

Increasing interest is focused on the possibility that large-scale chromosomal aberrations (CNVs) contribute an important role in autism. Early studies found an increased frequency of CNVs in autistic individuals at a rate of 5 to 10 percent compared to 1 to 3 percent in controls. Subsequent studies failed to replicate the initially identified high rate of CNVs compared to controls; however, recent work has suggested that the specific location and functional consequence of CNVs are more critical than the relative number of abnormalities. Two large studies identified a heterogeneous collection of CNVs that were not randomly distributed within the genome, but rather concentrated within a subset of loci. Functional analysis of these loci showed that they are involved in synapse development, neuronal migration, and axon targeting. These functional findings converge nicely with investigations of well-characterized non–

CNV-related abnormalities (discussed in detail below) and provide confidence that sporadic or inherited CNVs represent an important etiological factor for some cases of the disorder. Collectively, the large population studies estimated that CNVs account for 5 to 8 percent of syndromic cases of autism.

Even prior to the demonstration of high rates of de novo mutations in autism, epidemiological studies had strongly suggested that the genetic basis of this disorder is likely complex. For example, although the risk of autism in first-degree relatives of autistic probands is high, there is a substantial falloff for second- and third-degree relatives of such probands, suggesting that multiple genetic variants must interact to increase susceptibility to this syndrome. Segregation analyses of autism also support the hypothesis that it is a heterogeneous disorder that reflects the actions of multiple genetic variants of small effect. Genetic studies of autism have included whole genome screens, candidate gene studies, chromosome rearrangement studies, mutation analyses, and comparative genomic hybridization studies. Taken together, these studies have contributed to an emerging picture of autism susceptibility that includes genes involved in three major systems: those involving synapse formation and maintenance, morphogenesis, and calcium regulation and signaling.

Synapse Formation and Maintenance

Perhaps the biggest breakthroughs in identifying susceptibility genes for autism have come from studies of disorders that display clinical features associated with autism or ASDs but with simpler inheritance patterns, including fragile X syndrome, tuberous sclerosis, and Rett syndrome. In general, the genetic defects associated with these disorders affect synapse formation and maintenance. Fragile X, which accounts for 3 to 4 percent of autism cases, is caused by an unstable trinucleotide repeat in the 5′ region of the *FMR1* gene at Xq27.3. This repeat expands as it is transmitted to succeeding generations, resulting in abnormal methylation and inhibition of expression of *FMR1*. *FMR1* produces an RNA-binding protein that acts as a chaperone for the transport of RNA from the nucleus to the cytoplasm and is involved in mRNA translation at the synapse. Abnormalities in dendritic spine density (increased over normal) and morphology (longer and thinner than normal) have been reported in individuals with fragile X as well as in mouse models of this disorder. Tuberous sclerosis, which accounts for perhaps 2 to 10 percent of autism cases (the rate of tuberous sclerosis is higher among autistic individuals with seizure disorders), results from mutations in one of two tumor suppressor genes, *TSC1* on 9q34, and *TSC2* on 16p13, both of which are involved in guanosine triphosphatase (GTPase) inactivation. Loss of a single copy of *TSC1* in mice has been shown to disrupt cytoskeletal dynamics and dendritic spine structure.

The *neuroligin* (*NLGN*), *SHANK,* and *neurexin* gene families are additional risk loci that appear to play a role in synapse formation and may be affected by mutations observed in some individuals affected with autism. The *NLGN* genes, situated on the X chromosome, produce cell adhesion molecules that are located on postsynaptic glutamatergic neurons. When mutated in rodents, these genes show defective trafficking and synapse induction. In nonmutated form, their expression induces the formation of normal, presynaptic terminals in axons. In mutants, there is defective transport of vesicles to the synaptic membrane and reduced synaptic induction. Mice carrying a human mutation of *NLGN3* show an interesting phenotype characterized by impaired social novelty preference and reduced vocalizations, two of the central features of the disorder in humans. The *SHANK* gene family encodes scaffolding proteins that are located in the postsynaptic membrane and serve as master organizers of the PSD. *SHANK3* is one of three members of the *SHANK* gene family and is located in the minimal region of the 22q13 deletion syndrome, known as Phelan-McDermid syndrome that includes prominent autistic features. Larger studies have found that *SHANK3* mutations are present in about 0.85 percent of all autistic subjects. A third family of synapse-related genes associated with autism are the neurexins. Neurexins interact with neuroligins and are critical to the formation of presynaptic boutons. Neuroexins also link calcium channels to synaptic vessels and mediate neurotransmitter release through exocytosis. Like neuroligins and SHANK3, mutations in neurexins (*NRXN1*) occur rarely, but have emerged multiple times in independent familial lines of individuals with autism.

Morphogenesis

Many individuals with syndromic forms of autism display facial dysmorphisms and/or abnormal body growth rates. About 20 percent of autistic children show macrocephaly with increased trajectories of head growth during the first few years of life, although a smaller proportion of patients show microsomia. Although the etiology of the abnormal growth patterns is not yet clear for most cases, several syndromic forms of autism have been associated with rare mutations of genes involved in embryogenesis and organogenesis. Mutations in *HOXA1,* a gene necessary for normal development of the brain stem, cerebellum, cranial nerves, and portions of the calvarium, have been identified in several families with heterogeneous phenotypes that include autistic features. Missense mutations have been identified in the tumor suppressor gene *PTEN* on chromosome 10q23. *PTEN* mutations have been associated with a broad cluster of syndromes including Cowden syndrome, Bannayan-Riley-Ruvalcaba syndrome, Proteus syndrome, and Lhermitte–Duclos disease. *PTEN* has been shown to interact in the ERK/P13K/mTOR pathway that is critical for incorporating information regarding nutrient availability and

growth factor signals to regulate cell proliferation and body growth. *PTEN* mice knockout models show macrocephaly, CNS overgrowth, and excessive dendritic and axonal growth. Syndromes involving *PTEN* mutations have heterogeneous features, including nonpsychiatric phenotypes like cancer, highlighting that the final expression of specific genetic mutations depends on a complex interplay of genetic background (other mutations) and environmental factors.

Another syndromic form of autism is Rett syndrome, an X-linked pervasive developmental disorder (the first with a known genetic etiology) that occurs only in girls and is associated with normal early development followed by loss of skills, particularly social engagement and purposeful hand skills by age 4. The syndrome also points to abnormalities in synapse formation and maintenance in ASD and ASD-like disorders. Rett syndrome is caused by mutations in *MeCP2,* which makes a methylated-DNA-binding protein that regulates gene expression and chromatin structure and is required for normal brain development. Loss of *MeCP2* is associated with delayed neuronal maturation and abnormal synaptogenesis. Loss-of-function mutations in *MeCP2* are lethal in males and cause severe symptoms in about 70 percent of females; however, some mutations result in relatively mild phenotypes with some intellectual disability and/or verbal deficits. The heterogeneous expression of these mutations on the syndromic phenotype again highlights the key role that genetic background plays in determining the final manifestation of symptoms.

Calcium Regulation and Signaling

Another class of genes that has been linked to at least some cases of autism includes ion channels, receptors, and Ca^{2+}-regulated signaling molecules. Gain-of-function mutations in several L-type voltage-gated Ca^{2+} channels that express constitutively active channels and lead to excessive Ca^{2+} influx have been associated with autism. Timothy syndrome is caused by gain-of-function mutations in *CACNA1C,* an L-type voltage-gated Ca^{2+} channel, and is characterized by intellectual impairment and autism. Interestingly, *CACNA1C* has been associated with schizophrenia and bipolar disorder, providing an example of a genetic risk factor that can cross diagnostic categories. Gain-of-function mutations in *CACNA1F,* another L-type voltage-gated Ca^{2+} channel, are the cause of congenital night blindness that can be accompanied by autistic features. Several other genetic mutations have been identified in other genes that are linked to calcium regulation. In one autism case, a mutation was identified in *SCN2A,* which expresses a sodium channel that is regulated by calmodulin. The mutation disrupts the binding affinity between calmodulin and the channel, leading to sustained channel activity during depolarization. In another case, a chromosomal translocation was identified that disrupted *KCNMA1* that codes for a

potassium channel that is activated by depolarization or intracellular Ca^{2+} and leads to constitutively depolarized resting membrane potential and anomalous neuronal excitability.

The Role of Common Variants in Autism

The genetic mutations described in the preceding sections are rare and collectively explain less than 10 percent of autism cases. The specific genetic causes for the remaining cases have not been conclusively identified. In large samples of genotyped individuals, it has been estimated that common variants explain between 40 and 60 percent of the liability for the disorder. These estimates have motivated the use of GWA methods to identify common variants of small effect. However, to date, GWA studies of autism have yet to rigorously identify and replicate any findings. This failure is almost certainly due to the relatively small sample sizes that have been investigated; the largest sample size to date is less than 2,500 individuals, which is much smaller than successful studies for other complex psychiatric disorders like schizophrenia, where sample sizes of almost 40,000 cases has been achieved (see below for more details).

Summary

Genetic investigation of autism has progressed considerably in the last several years. Successes in autism genetics may be attributed largely to three factors: (1) the high heritability of this disorder, (2) international collaborations that have made large samples of autism families readily available to the scientific community for a wide range of investigations, and (3) the relatively consistent findings obtained using complementary approaches, including linkage studies, investigation of known chromosomal abnormalities, comparative genomic hybridization, mutation analyses, and investigation of animal models that display phenotypes relevant to autism. Compared to other psychiatric disorders like schizophrenia, bipolar and major depression, there are now relatively elaborated models of the molecular biology and cellular pathways that are involved in the etiology of at least some of the cases. This progress is attributable to the investigation of syndromic cases of autism, which are caused by rare mutations of large effect. Unanswered at this point is whether and to what extent the majority of autism cases (nonsyndromic) share overlapping pathophysiology with syndromic cases; a question that is likely answerable by increasing sample sizes for GWA studies.

BIPOLAR DISORDER, SCHIZOPHRENIA, AND MAJOR DEPRESSION

Bipolar Disorder

The search for the genetic basis of bipolar affective disorder has been

fraught with missteps and partial answers. The history of genetic mapping attempts for bipolar disorder illustrates not only the extreme complexity of psychiatric disorders but also the evolution of genetic approaches to such diseases. Bipolar disorder is an episodic illness characterized by recurrent periods of both mania and depression. Psychotic symptoms are often a part of the clinical picture, particularly in more severely affected individuals.

Numerous genetic epidemiological investigations conducted over several decades have strongly supported a genetic contribution to risk for bipolar disorder. As with other psychiatric disorders, however, the definition of the bipolar disorder phenotype in these studies has varied substantially, and this in turn has resulted in a wide range in estimates of its heritability. For example, many early studies into the genetic basis of mood disorders did not distinguish between unipolar and bipolar mood disorders. Furthermore, the diagnostic methodology used in such early studies differs substantially from that employed in current-day genetic studies. For example, a Danish twin study that suggested a very high heritability for bipolar disorder and thereby had a heavy influence on the design of initial genetic mapping studies of mood disorders employed only unstructured diagnostic interviews by a single clinician rather than the structured assessments used in current studies, which have suggested somewhat lower heritabilities.

Current estimates of concordance for bipolar disorder range between 65 and 100 percent in MZ twins and between 10 and 30 percent in DZ twins, indicating that the disorder is highly heritable (between about 60 and 80 percent). Several studies have shown that bipolar disorder is substantially more heritable than unipolar major depression, which has an estimated heritability between 30 and 40 percent. Early family studies suggested that bipolar disorder segregation patterns were compatible with single gene inheritance of a locus of major effect. Although it is possible that some bipolar disorder pedigrees segregate such a locus, mounting evidence indicates that if such pedigrees exist they must be quite rare. Furthermore, the fact that genetic linkage studies have failed to uncover such a locus with unequivocal evidence in any pedigrees argues against this possibility. The observed rapid decrease in recurrence risk for bipolar disorder from MZ cotwins to first-degree relatives is also not consistent with single gene inheritance models but rather suggests models of multiple interacting genes.

Tremendous excitement followed the first reports of linkage to bipolar disorder on chromosomes X and 11 in 1987. Investigators noted that in several families, bipolar disorder and other affective disorders appeared to be inherited in an X-linked fashion. Likewise, these disorders appeared to cosegregate in several Israeli families with color blindness and G6PD deficiency, which map to the X chromosome. Linkage studies in these pedigrees, using color blindness or G6PD deficiency as marker loci, gave

LOD scores between 4 and 9. Early studies of chromosome 11 were similar to those for chromosome X in that they reported significant linkage after testing only a few markers in a single region, in this case in an extended Old Order Amish pedigree heavily loaded for bipolar disorder.

Not surprisingly, these findings generated a great deal of interest. Both studies showed high LOD scores and seemed to provide clear evidence for linkage. However, replication studies in other populations failed to produce positive results for either the X chromosome or chromosome 11, and evidence for linkage essentially disappeared in both chromosomal regions in the samples in which linkage was originally reported when the pedigrees were extended to include additional affected individuals and when additional markers were typed in the putative linkage regions. The most likely explanation in each case is that the original linkage results were false-positive findings and may have reflected overoptimistic interpretation of evidence that, in retrospect, was relatively scanty.

The early linkage studies of bipolar disorder evaluated only a few markers because they were all that were available. With the construction of genetic linkage maps of the genome in the 1990s, linkage studies of most complex traits, including bipolar disorder, began to search genome-wide. The advantage of genome-wide mapping studies is that they do not require a priori knowledge of the biological underpinnings of a particular phenotype. Complete genome screens provide an opportunity to evaluate the evidence of linkage at all points in the genome without bias. While genome-wide studies clearly had greater power to detect true linkage than studies focused on only a few markers in arbitrary locations or around a few candidate genes, these investigations have also generally had disappointing results. The challenge of achieving replicated significant linkage results for bipolar disorder and other complex traits is apparent when one reviews the many gene-mapping studies that have suggested—but not demonstrated unequivocally—bipolar disorder susceptibility loci on chromosome 18.

The first report of linkage came from a partial genome screen that examined 11 markers on chromosome 18 and identified suggestive linkage near the centromere. Because the inheritance patterns for bipolar disorder are unknown, the results were analyzed using both recessive and dominant models. Some of the markers were positive under a recessive model in some families, some were positive under a dominant model in other families, and some markers gave positive LOD scores in a subset of families under both models. Attempts to replicate this finding in other populations have been mixed. So far at least two groups have found no evidence for linkage to the pericentromeric region of chromosome 18 in their samples, although one other group has found evidence to support linkage to this region. Other studies have found suggestive evidence for linkage on chromosome 18, including a complete genome screen in two large Costa

Rican pedigrees that gave evidence for linkage on chromosome 18q22–23 as well as in an area on 18p. The combined evidence of these several studies, although somewhat contradictory and confusing, points to at least two different susceptibility loci on chromosome 18: one on 18p and another on 18q.

The complexities outlined above for chromosome 18 highlight the potential difficulties encountered in gene mapping studies. Replications of significant findings in small studies often fail in samples of different families. One approach to overcome the low power of small linkage studies is to use meta-analyses to combine numerical results of several individual studies to estimate the average effect of interest. A meta-analysis combining the results of 18 genome-wide linkage investigations of bipolar disorder failed to identify any loci at a genome-wide level of significance. The most promising finding from this meta-analysis was for a location on chromosome 9p21–22. In a separate meta-analysis that combined the genotyping data from 11 scans, linkage signals were detected on chromosomes 6q21–25 and 8q24 that were interpreted by some geneticists as meeting genome-wide levels of significance. The largest meta-analysis to date included almost 1,000 families, but even for this large sample, it is still debatable whether any unequivocally linkage signals were identified. Various explanations have been put forward to interpret the lack of findings for these meta-analyses; at the top of the list is the possibility that they have been statistically underpowered. Another consideration is that multiple genes of small effect are interacting to cause the disorder, in which case these genes will be difficult to detect with linkage approaches. This problem will be especially difficult in the presence of genetic heterogeneity. Another possibility is that each family is segregating a unique set of genetic variants and linkage analysis has little power to identify risk loci in these situations. Ultimately, it may turn out that the genetic architecture of bipolar disorder includes a mix of common variants segregating in the general population and rare alleles that are specific to recent ancestors and only segregating within a relatively small set of families. Much of the recent gene mapping effort in psychiatric genetics has focused on identifying common risk loci of small effect using GWA methods. Compared to several decades of frustrating linkage studies, recent GWA investigations are appreciating considerable success (described below). However, individual common variants identified in GWA studies account for only a small proportion of disease risk (~1 to 2 percent). Therefore, family-based linkage methods have a potentially important role for the full elaboration of the genetic architecture of psychiatric disorder by identifying rare risk variants that occur on the background of common risk variants and determine which individuals develop the disorder.

In contrast to the disappointing results from linkage analysis described above, GWA studies of bipolar have begun to yield statistically rigorous

significant findings. In 2007, the Wellcome Trust Case Control Consortium (WTCCC) conducted the first GWA study of bipolar disorder in 2,000 cases and 3,000 controls genotyped at ~5,00,000 SNPs, but did not identify any genome-wide significant risk variants. The negative finding was notable given that six other common disorders were included in the WTCCC, all of which generated some significant findings. In 2008, a similarly sized investigation was implemented in samples from the Systematic Treatment Enhancement Program for Bipolar Disorder and University College of London (STEP and UCL), and also failed to identify any genome-wide significant risk loci. The WTCCC, STEP, and UCL then combined their data, to obtain a sample size of 4,387 cases and 6,209 controls, and finally reported the first genome-wide significant association for the disorder for a SNP residing in the *ANK3* (ankyrin G) gene. The *ANK3* finding has subsequently been replicated multiple times in additional studies, which have also extended the list of risk loci such that by the end of 2013 about a dozen loci had been established as confirmed risk factors for bipolar disorder. As expected for population-based GWA findings, the effect sizes of the risk alleles are low (odds ratios less than 1.2), indicating that individually these common variants only contribute to a small increase in disease risk. Although the growing list of bona fide variants is not yet large enough to make definitive statements regarding the underlying etiological mechanisms involved in the disorder, functional patterns are beginning to emerge. Additionally, innovative variations in the GWA method, commonly termed pathway analysis, have been developed to assess the significance of gene clusters with similar biological function instead of individual genetic variants. Pathway analysis leverages the growing body of evidence in gene ontology databases to group genes into specific biological pathways and thereby improves statistical power by reducing the overall number of tests. In combination with traditional GWA approaches, such approaches have provided evidence for the involvement of a range of cellular processes in bipolar disorder including ion channelopathies, endocrine regulation, and posttranscriptional modification. Current effort for gene mapping in bipolar disorder is focused on amassing larger samples sizes. As will be described below for similar studies in schizophrenia, greater sample sizes have yielded a greater number of significant findings. The current understanding of the genetic architecture of bipolar disorder, which was murky just 5 years ago, has started to come into focus, and in combination with family-based studies and intermediate phenotype approaches, it promises to translate into clear mechanistic disease models.

Schizophrenia

Perhaps more than any other psychiatric disorders, genetic studies of schizophrenia have appreciated the greatest acceleration of new findings in the past several years. The success has come through the establishment of

several large consortia that have amassed extensive samples with sufficient power to detect risk variants of small effect. Prior to successful gene mapping, family studies had clearly established that schizophrenia was strongly influenced by genetic factors with heritability estimates between 60 and 80 percent. Despite the high heritability, no meaningful genome-wide population findings emerged until 2009, when the International Schizophrenia Consortium, Molecular Genetics of Schizophrenia, and the Schizophrenia Working Group of the Psychiatric Genomics Consortium combined samples and identified a SNP in the MHC on chromosome 6p22 as a risk locus for the disorder. The mapping of a schizophrenia risk locus to the MHC region was an exciting discovery, although the full significance of the finding is difficult to interpret. It has long been known that exposure to maternal infection is a risk factor for schizophrenia. Therefore, specific MHC configurations may lead to anomalous immune responses that impact brain development and/or processes that predispose to the disorder. Additionally, neurons express MHC molecules where they are involved in several aspects of brain development including synaptogenesis, neurite outgrowth, and long-term plasticity. Therefore, genetic variations in the MHC may directly impact brain development outside immunomodulatory mechanisms.

Shortly after the discovery of the MHC as a risk locus, GWA studies identified more than a dozen additional risk loci, and as sample sizes increased, the number of novel and replicated risk loci accumulated. In 2014, the Schizophrenia Working Group of the Psychiatric Genomics Consortium, which by then had reached a sample size of 37,000 cases and 113,000 controls, identified 108 distinct loci contributing to risk for schizophrenia. It is important to note that the identification of a genetic association at a specific SNP does not imply that the genotyped SNP is the causal allele. In fact, it is unlikely that the genotyped SNP on most current genotyping platforms has functional significance. The genotyped SNPs are more likely in LD in the immediate vicinity (~100 kb) of the causal variant. Additionally, given the extensive sequence variation even within relatively small regions, establishing a specific variant or set of variants as causal factors in the development of a complex disorder is not trivial. Complicating the definitive identification of variants is the expectation that many of the risk loci identified by GWA studies will harbor causal variants that impact more subtle aspects of gene functioning compared to Mendelian genes of large effect. While Mendelian mutations often lead to major disruptions in the coding sequencing of genes, common risk variants of small effect have been hypothesized to impact gene functioning at the level of expression or regulation. In contrast to mutations that impact protein-coding regions, it is more difficult to identify mutations that lie outside the intron/exon structure of genes that may be affecting regulatory regions that control expression or chromatin structure. Ongoing initiatives

like the ENCODE project aim to characterize biochemically active regions of the genome including regulatory protein binding sites and open chromatin regions, thereby providing a valuable resource for guiding the interpretation of mutations in noncoding regions.

Although the elaboration of the specific biological pathways involved in the pathogenesis of schizophrenia will require definitive identification of causal mutations and their impact on specific genes, the recent accumulation of risk loci provides suggestive directions. The risk loci do not appear to be randomly distributed among genes involved in all cellular processes, but rather appear to be enriched near genes involved in a smaller subset of processes important for brain functioning including neurodevelopment (FXR1 on 3q26.33, SATB2 on 2q33.1, PODXL on 7q32-q33, BCL11B on 14q32.2, TLE1 on 9q21.32, TLE3 on 15q22, and BRINP2 on 1q24), glutamatergic neurotransmission (GRIN2 A on 16p13.2, GRIA1 on 5q33.2, SRR on 17p13.3, CLCN3 on 4q33, and GRM3 on 7q21.12), neuronal calcium signaling (CACNA1C on 12p13.3, CACNB2 on 10p12, CAMKK2 on 12q24.2, NRGN on 11q24, ATP2A2 on 12q24.11, CACNA1I on 22q13.1, and RIMS1 on 6q12–13), and synaptic function and plasticity (KCTD13 on 16p11.2, NLGN4X on Xp21.33–32, IGSF9B on 11q25, CNTN4 on 3p26.3, MEF2C on 5q14.3, PTN on 7q33, CNKSR2 on Xp22.12, PAK6 on 15q14, and SNAP91 on 6q14.2).

Although GWA studies have appreciated substantial success over the past several years, it has become clear that common variants segregating in the general population are not able to account for all of the estimated genetic contribution to psychiatric disorders; this fact has been labeled as the “missing heritability” problem. For example, based on twin and family studies, the heritability for schizophrenia is around 70 percent, but the total amount of variance explained by common genetic polymorphisms is estimated to be about 30 percent, such that there is a “gap” of unexplained heritability. Various explanations have been put forward to account for the missing heritability problem. One factor that is likely to explain at least some of the missing heritability is the possibility that rare genes of medium-to-large effect that are not detectable with GWA methods are present in affected individuals. Prior to the recent GWA results outlined above, a number of family-based linkage studies had started making progress in the identification of loci segregating within families genetically loaded for schizophrenia. Chromosome 6p24–22 was among the first regions to be implicated by a complete genome screen for schizophrenia, in this case from a study of Irish families heavily loaded for schizophrenia. The linkage results were strongest under a broad diagnostic definition that included schizophrenia spectrum disorders, such as schizotypal personality disorder. Six additional linkage studies have shown positive results over approximately the same region, but at least three studies have found no linkage to the region. Fine-scale mapping of this region using association

analysis in the original Irish kindreds led to the proposal of *Dysbindin* (*DTNB1*) as a candidate gene for schizophrenia. Additional association studies of *Dysbindin* have been equivocal. Although multiple association studies in a variety of populations have shown positive results, interpretation of the results has been difficult. Different association studies have not used the same SNP marker sets. Meta-analysis of five "positive" association studies using a high-resolution haplotype map designed to compare the five studies showed significant inconsistencies with regard to the identified disease-associated *Dysbindin* allele. Although it is possible that several different variants in the same gene could each contribute to disease susceptibility in different families or populations, this possibility does not explain the inconsistencies between the several *Dysbindin* association studies.

Separate linkage studies subsequently pointed to a region on chromosome 1 containing the candidate genes *DISC1* and *DISC2* (*disrupted in schizophrenia 1* and *2*) located on chromosome 1q21–22 and 1q32–42. These genes were initially identified in a large Scottish pedigree in the early 1990s. A balanced translocation between chromosomes 1 and 11 was segregated in this pedigree and was possibly associated with serious mental illness. *DISC1* and *2* were identified in the original Scottish family because of their location near the chromosomal translocation breakpoint. As with *Dysbindin,* follow-up studies of *DISC1* and *2* have been equivocal.

Genome screens, including a screen focused on extended Icelandic kindreds, have identified a schizophrenia candidate region on chromosome 8p21–22. Fine mapping of the region narrowed the search and eventually led to the proposal of *Neuregulin 1* (*NRG1*) as a schizophrenia candidate gene. Association studies again provided equivocal and difficult-to-interpret results. Meta-analysis of 14 separate studies using the SNP marker that demonstrated an association in the original study showed significant heterogeneity between the follow-up studies. It also showed that there is no consistent association between the specific risk allele "tagged" by the marker SNP and schizophrenia in different populations. However, after taking account of the statistical power of each association study, the meta-analysis showed a positive association between *NRG1* at the level of the gene (as opposed to the SNP or haplotype level).

Despite the equivocal genetic studies, significant resources have been channeled into molecular and neurophysiological investigations of the functional products of *Dysbindin, DISC1*, and *Neuregulin*. Mutant mice for each of the three genes are now available and have been used to demonstrate interesting biological findings. For example, *Dysbindin* is expressed in the hippocampus and dorsolateral prefrontal cortex. The dysbindin protein binds to β-dystrobrevin and has been implicated in synaptic structure and signaling. *DISC1* has been shown to influence neurite formation in cellular studies, and mutant mice for *DISC1* show

impairments in a wide variety of tests including learning, memory, and sociability. Neuregulin belongs to a family of growth factors that mediate numerous functions including synapse formation, neuronal migration, and neurotransmission. Targeted disruption of *erbB4,* the postsynaptic target of neuregulin, leads to synaptic glutamatergic hypofunction. Despite the interesting biology uncovered, it remains unclear whether and to what extent any of these genes contribute to the etiology of schizophrenia in humans, and many geneticists have been cautious in their endorsement of the legitimacy of the mutant mice generated from the current list of candidate genes as models of psychiatric disorders.

Major Depression

Major depression is the most common psychiatric disorder and is one of the leading causes of morbidity worldwide. Lifetime prevalence varies across countries but generally ranges between 8 and 12 percent. Evidence for genetic contribution to major depression comes from family studies that show an increased risk of the disorder in first-degree relatives of probands that is 2 to 3 times greater than the general population. The heritability of major depression is generally lower than other mood and psychotic disorders and is estimated to be between 30 and 40 percent. Large population studies have consistently shown that there is a higher prevalence of the disorder in women and twin studies have shown that the disorder is in fact more genetically influenced in women with an estimated heritability of ~40 percent compared to ~30 percent for men. These same studies have estimated that although a substantial proportion of genetic causes of major depression are shared between men and women, a significant proportion of genetic factors are sex specific.

In contrast to the recent success of investigations in schizophrenia and bipolar disorder populations, GWA studies of major depression have not yet identified any replicated findings. A major factor in this failure is that fact the sample sizes for gene mapping studies of major depression have not attained the numbers compared to schizophrenia and bipolar disorder, with the largest studies to date having about 10,000 patients. The lack of statistically significant findings with sample sizes of this magnitude is informative; on modern genotyping array, studies with as sample size of 10,000 cases have ~85 percent power to detect a variants with odds ratio of >1.2; the absence of findings to date provides support for the growing consensus that common genetic variants influencing major depression individual contribute only a small increase in risk (odds ratio <1.2). Additionally, the lack of significant findings in genome-wide mapping studies of major depression provides a context in which to interpret the vast and confusing literature of candidate gene association studies of major depression. During the past two decades, over 1,000 studies focusing on over 200 specific genes have tested specific genetic variants for association

with major depression and have produced a complex mix of conflicting results. The recent genome-wide interrogations described have adequate power to identify significant associations reported in candidate gene studies, but have consistently failed to do so indicating that the vast majority of "significant" candidate gene findings reported to date are false positives.

GENETIC OVERLAP AMONG PSYCHIATRIC DISORDERS

The preceding sections have emphasized that categorically defined psychiatric disorders display substantial genetic heterogeneity, indicating that the symptom clusters that specify diagnostic categories can emerge from mutations in different sets of genes. The converse is also true; mutations in specific genes can cross diagnostic boundaries. Early family studies showed that relatives of an individual diagnosed with a specific disorder (i.e., schizophrenia) are at higher risk to develop not only schizophrenia but also major depression or bipolar disorder. Gene mapping studies have confirmed that genetic risk loci are shared by different disorders. For example, mutations in the calcium channel gene, *CACNA1C,* have been linked to autism, bipolar disorder, and schizophrenia. A recent study investigated the overlap of genome-wide common variants in a large sample (~30,000 cases and ~30,000 controls) that included patients from five categorically defined psychiatric disorders: autism, attention deficit hyperactivity disorder (ADHD), bipolar disorder, major depression, and schizophrenia. The investigation used the dense SNP genotype information to estimate the extent to which genetic risk factors were shared between the five disorders and found a remarkable degree of overlap between schizophrenia and bipolar disorder; about 68 percent of the polygenic risk attributable to common variants was shared between these two disorders. Major depression also showed a substantial overlap with other disorders, where about 45 percent of the genetic risk attributable to common variants was shared with both schizophrenia and bipolar disorder, and about 30 percent was shared with ADHD. Schizophrenia showed a low (~15 percent) but significant overlap with ASDs. Autism did not show significant genetic overlap with bipolar disorder, major depression, or ADHD. Similarly, aside from the genetic overlap with major depression, ADHD did not show significant overlap with any of the other disorders. These findings seem to parallel clinical scenarios, where, for example, schizophrenia and bipolar disorder can sometimes be difficult to disentangle, but bipolar disorder and autism are usually easy to distinguish. The findings also provide hints into the emerging pattern of cellular pathways underlying psychiatric disorders; alterations in some molecular pathways (e.g., calcium signaling) can perturb neural systems nonspecifically such that the final psychiatric presentation is determined by other genetic and/or environmental factors. This scenario may be particularly true for schizophrenia, major depression,

and bipolar disorder, which share a substantial proportion of common genetic risk factors. In contrast, disorders like attention deficit-hyperactivity do not share as many common genetic risk factors with other disorders and may emerge through alterations of cellular pathways distinct from those leading to the psychotic and mood spectrum disorders.

FUTURE DIRECTIONS

The field of psychiatric genetics is in a state of transition. New technologies have made the genetic exploration of complex behaviors a real possibility. While 25 years ago human gene-mapping studies involved the use of a few markers in a small number of individuals, it is now routine to genotype millions of markers in samples of tens of thousands of individuals. The development of exciting new technologies for phenotyping the human nervous system—such as high-resolution neuroimaging—and the implementation of many such assays in genetic mapping studies have created extraordinary opportunities for the large-scale association of phenotypes and genotypes in psychiatric research. While the analysis of such vast and complex datasets remains a substantial challenge, the development of methodologies for this purpose is now a central focus of research in statistical genetics and bioinformatics.

As the field of psychiatric genetics progresses, several major questions that have emerged may soon be answered, at least partially, through the analysis of large-scale datasets. One major question is whether genetic risk for psychiatric disorders derives mainly from a few rare variants of large effect or many common variants of small effect (or some combination of rare and common variants). While well-powered GWA studies should identify common variants, they have very little power to identify rare variants. The failure thus far of linkage studies to identify rare causative variants for psychiatric disorders does not preclude the possibility that rare variants of somewhat lesser effect could play an important role in susceptibility to these diseases. Whole genome sequencing of large samples is now feasible and is being implemented with the goal of identifying rare variants segregating in families with heavy genetic loading for psychiatric disorders.

The second question concerns the degree to which psychopathology will be dividable into discrete disorders, as in the current DSM nosology, or if the heterogeneity and overlap among syndromes is so great as to require a thorough overhaul of the current classification systems. Estimates from population-based investigations of common genetic variation show that the mood and psychotic disorders share a substantial overlap of genetic risk factors, whereas disorders like autism and attention deficit-hyperactivity may be more genetically distinct from other psychiatric disorders. To elaborate this issue more completely, it will be necessary to begin examining how genetic risk factors interact with other risk genes and the

environment to predict the symptomatology in specific individuals. Additionally, the increasing focus of the field on genetic investigation of brain, behavioral, and other endophenotypes for major psychiatric disorders may be a step in the direction of such a revolution.

The final question concerns the impact of anticipated psychiatric genetic discoveries on health outcomes—in terms of either prevention or improved treatments. For Alzheimer disease, concrete advances have been made and answers to this question may be obtained soon. For the other disorders discussed in this chapter, rapid advances have been made in the last several years, but the emerging picture is still fluid, and it may still be several years before enough is understood to frame this question in answerable terms.

References

Balding DJ. A tutorial on statistical methods for population association studies. Nature reviews. *Genetics*. 2006;7(10):781–791.

Bienvenu OJ, Davydow DS, Kendler KS. Psychiatric 'diseases' versus behavioral disorders and degree of genetic influence. *Psychol Med*. 2011;41(1):33–40.

Cardno AG, Rijsdijk FV, West RM, et al. A twin study of schizoaffective-mania, schizoaffective-depression, and other psychotic syndromes. *Am J Med Gen B Neuropsychiatr Genet*. 2012;159B(2):172–182.

Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype-phenotype associations. *Nature*. 2007;447(7145):655–660.

*Civelek M, Lusis AJ. Systems genetics approaches to understand complex traits; Nature reviews. *Genetics*. 2014;15(1):34–48.

Craddock N, Sklar P. Genetics of bipolar disorder. *Lancet*. 2013;381(9878):1654–1662.

*Flint J, Kendler KS. The genetics of major depression. *Neuron*. 2014;81(3):484–503.

Flint J, Munafò M. Schizophrenia: genesis of a complex disease. *Nature*. 2014;511(7510):412–413.

Glahn DC, Knowles EE, McKay DR, et al. Arguments for the sake of endophenotypes: examining common misconceptions about the use of endophenotypes in psychiatric genetics. *Am J Med Genet B Neuropsychiatr Genet*. 2014;165B(2):122–130.

Kellis M, Wold B, Snyder MP, et al. Defining functional DNA elements in the human genome. *Proc Natl Acad Sci U S A*. 2014;111(17):6131–6138.

*Lee SH, Ripke S, Neale BM, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet*. 2013;45(9):984–994.

Merikangas AK, Segurado R, Heron EA, et al. The phenotypic manifestations of rare genic CNVs in autism spectrum disorder. *Mol Psychiatry*. 2015;20(11):1366–1372.

Munafò MR, Thiselton DL, Clark TG, Flint J. Association of the NRG1 gene and schizophrenia: a meta-analysis. *Mol Psychiatry*. 2006;11(6):539–546.

Mutsuddi M, Morris DW, Waggoner SG, Daly MJ, Scolnick EM, Sklar P. Analysis of high-resolution HapMap of DTNBP1 (Dysbindin) suggests no consistency between reported common variant associations and schizophrenia. *Am J Hum Genet*. 2006;79(5):903–909.

Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. *Nat Reviews Genet*. 2015;16:275–284.

O'Tuathaigh CM, Babovic D, O'Meara G, Clifford JJ, Croke DT, Waddington JL. Susceptibility genes for schizophrenia: characterisation of mutant mouse models at the level of phenotypic behaviour. *Neurosci Biobehav Rev*. 2007;31(1):60–78.

Persico AM, Napolioni V. Autism genetics. *Behav Brain Res*. 2013;251:95–112.

Pinto D, Delaby E, Merico D, et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Hum Genet*. 2014;94(5):677–694.

Reiman EM, Webster JA, Myers AJ, et al. GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. *Neuron*. 2007;54(5):713–720.

Roybal K, Theobold D, Graham A, et al. Mania-like behavior induced by disruption of CLOCK. *Proc Natl Acad Sci U S A*. 2007;104(15):6406–6411.

Rzhetsky A, Wajngurt D, Park N, Zheng T. Probing genetic overlap among complex human phenotypes. *Proc Natl Acad Sci U S A*. 2007;104(28):11694–11699.

*Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511(7510):421–427.

Shih RA, Belmonte PL, Zandi PP. A review of the evidence from family, twin and adoption studies for a genetic contribution to adult psychiatric disorders. *Int Rev Psychiatry*. (Abingdon, England). 2004;16(4):260–283.

Sklar P, Pato MT, Kirby A, et al. Genome-wide scan in Portuguese Island families identifies 5q31–5q35 as a susceptibility locus for schizophrenia and psychosis. *Mol Psychiatry*. 2004;9(2):213–218.

Stanewsky R. Genetic analysis of the circadian system in Drosophila melanogaster and mammals. *J Neurobiol*. 2003;54(1):111–147.

*Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet*. 2012;13(8):537–551.

Tanzi RE. The genetics of Alzheimer disease. *Cold Spring Harb Perspect Med*. 2012;2(10):1–10.

Thomas D. *Statistical Methods in Genetic Epidemiology*. New York: Oxford University Press; 2004.

Visscher PM, Hill WG, Wray NR. Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet*. 2008;9(4):255–266.

Witte JS, Visscher PM, Wray NR. The contribution of genetic variants to disease depends on the ruler. *Nat Rev Genet*. 2014;15(11):765–776.

▲ 1.19 Animal Models in Psychiatry

SIETSE JONKMAN, PH.D., AND PAUL J. KENNY, PH.D.

THE USE OF ANIMAL MODELS IN PSYCHIATRIC RESEARCH

Animal research allows experimenters to manipulate and measure behavioral and biological processes that are relevant to psychiatry. Animal research in psychiatry falls into two broad categories. First, basic research

aims to improve the current understanding of psychological and neural processes in their species-specific context using established and novel technical approaches. Second, in applied research, animal experiments are used to model human psychiatry, either to help identify novel targets for therapeutic drug development or to provide a preclinical test of the effectiveness of a compound on a behavior considered relevant to a psychiatric disorder. The usefulness of the latter type of research depends entirely on the degree to which findings can be translated between species.

There are clearly substantial differences between humans and NHPs and rodents, which are the species most often used in animal research. Abstract thought in humans allows for self-regulation, exemplified by the fact that humans, unlike any other animal, will diet and willingly forego the opportunity to consume tasty food in order to lose weight. Conversely, self-reflection can lead to rumination, which may play a role in human susceptibility to depression. Language allows humans to engage in uniquely complex communication, and insights gained during therapy may be instrumental in coping with psychiatric disorders. Biological differences probably underlie these differences, including a more developed cortex and prefrontal cortex compared to NHPs and particularly rodents. More broadly, although the vast majority of human genes have homologous versions in NHPs and rodents, there are substantial differences at the level of gene regulation that likely explain differences in organismal complexity.

Table 1.19–1.
Implication of Class of Symptoms in Major Psychiatric Disorders

Domain	Autism	ADHD	Schizophrenia	Bipolar Disorder	Depression	Anxiety	PTSD	OCD	Eating Disorder	Impulse-Control
Negative valence	++	+	++	+++	+++	+++	+++	++	++	+
Positive valence	++	+++	++	+++	+	+	++	+++	+++	+++
Cognitive	+++	+++	+++	+	+	+	+	+++	+	+++
Social	+++	+	++	++	++	++	+	+	+	++
Sleep and arousal	++	++	+	+++	++	++	+++	+	++	+

+++, strongly impacted; ++, moderately implicated; +, slightly implicated.

These differences notwithstanding, there are many similarities between humans and experimental animals in both functional and biological terms. Thus, basic psychological processes such as stress, panic, anxiety, motivation, attention, and reward are present in rodents and NHPs, albeit in species-specific expressions. In addition, subcortical structures are quite well conserved, with a common brain design and substantial homologies for most neural structures.

Based on these considerations, and on the limited success in using animal models to develop new therapeutics for psychiatric disorders, it has been suggested that animals should be used to model particular symptoms, instead of entire human psychiatric disorders. In particular, productive animal research in psychiatry should focus on well-defined phenotypes that have a known biological basis and that can be measured to facilitate experimentation, and preferentially focus on brain substrates that have reasonable homologies between humans and the animal model of choice. In addition, when genetic models are used in animals, it is useful to focus on highly penetrant genetic variants to ensure relevance to human disease. In a general sense, the usefulness of animal models in modeling aspects of human psychiatric disorders can be assessed by three criteria.

- *Face validity* refers to the extent to which a behavior that is studied in animals resembles the behavior in humans it is intended to model.
- *Predictive validity* refers to the extent to which drug effects in animal models predict clinical effectiveness in psychiatric patients, and an ideal model would have low false positive and false negative rates.
- *Construct validity* refers to the extent to which an animal model reproduces the etiology of the human psychiatric disorder.

Following recent guidelines from the NIMH, this chapter is organized by a description of available animal models in five broad classes of symptoms, which may be observed across psychiatric disorders (see Table 1.19-1).

NEGATIVE VALENCE

Defensive Behaviors

Ethological studies of defensive behavior in rodents have shown that threatening situations evoke defensive behaviors that depend on the proximity of the threat. When a threat is very near and acute, panic responses consisting of flight or defensive attack are elicited, while fear responses consisting of concealment and freezing occur when a threat is present but more distal. Finally, potential threats elicit lower intensity anxiety-like behavior, with risk assessment and reduction in exploratory behaviors for extended periods of time. In general, animal models of defensive behavior are focused on acute normal defensive responses, while only few models involve chronic stress manipulations.

Panic-Like Behavior. Panic responses can be studied in rats and mice in its ethological context in defensive tests that use a predator (humans or rats, respectively) closing in on subjects in a circular runway. In this test, called the defense test battery, the predator approaches the subject in a long oval runway, and flight responses can be quantified. If the runway is blocked at one point, imminent approach of the rat or mice elicits defensive threat and attack responses. In addition to good face validity, pharmacologic agents that are clinically effective at reducing panic attacks have been shown to reduce flight in this model, while anxiolytic agents that are less effective in reducing panic attacks do not have a strong effect on flight, suggesting good predictive validity. However, the main drawback of this model is its relative complexity, and it has not been widely adopted to date. Panic responses can also be evoked artificially in rats by electrical stimulation of the dorsal periaqueductal grey matter (dPAG). As the dPAG harbors a hypoxia sensitive alarm system, this model is most relevant to the putative respiratory subtype of panic disorder.

Fear-Like Behavior. Unconditioned fear responses to predators can also be measured in the defense battery tests for rats or mice. In both cases, freezing behavior is observed in the blocked runway when the predator approaches up to a certain point. Further approach is followed by active defensive attack.

Conditioned fear responses have traditionally been studied to elucidate the neural basis of learning and memory, but are increasingly used to study normal and pathological defensive behaviors. In the basic fear conditioning setup, experimenters consistently present a conditioned stimulus (a tone or a light) before an aversive unconditioned stimulus (typically a mild foot shock). After repeated pairings, the subject will be able to predict presentation of the aversive stimulus when the conditioned stimulus is

presented. Conditioned fear can be measured in the fear potentiated startle paradigm, where presentation of the conditioned stimulus potentiates the locomotor response to a loud startle tone. This procedure can be used in both rodents and humans with minor adaptations. Alternatively, conditioned fear is often measured in rodents in enclosed environments as a freezing response. Although both tests can be confounded by manipulations that affect locomotion in opposite ways, when a pharmacological manipulation reduces fear potentiated startle and increases locomotion by reducing freezing, the results can be interpreted as fear-reducing with confidence.

Generalization of conditioned fear is a learning phenomenon that refers to the transfer of conditioned fear responses from the proper conditioned stimulus that predicts an aversive outcome to stimuli that are related, but distinct. Theoretically, putative abnormal generalization of conditioned fear to related stimuli is an attractive explanation for the pervasiveness of defensive responses in panic disorder and PTSD. Operationally, rodents, NHPs, or humans can be tested for the generalization of conditioned fear by presenting stimuli that are more or less similar to the original conditioned stimulus, and measuring fear responses.

Extinction of conditioned fear is another learning phenomenon that has clear relevance to clinical psychiatry. In this procedure, after fear conditioning in rodents, NHPs, or humans, the conditioned stimulus is presented repeatedly in the absence of the unconditioned stimulus, gradually reducing conditioned fear responses. This procedure has a clear parallel in exposure therapy that is used to successfully treat generalized anxiety disorder, specific phobias, and PTSD. Pharmacological treatments that increase extinction such as D-cycloserine have been shown to enhance the beneficial effect of exposure therapy.

Anxiety-Like Behavior. The most common way to measure anxiety-like behavior in rodents is based on approach–avoidance conflicts. Rats and mice are naturally cautious in open, elevated, and brightly lit environments because of the risk of predation, and the avoidance of such places can be a measure of anxiety-like behavior. Typically, animals are allowed to explore an apparatus with an enclosed or dark compartment, and a more open or brightly lit or elevated compartment. The elevated plus-maze, light dark exploration, and open field tests are the most popular tests in this category. There is also a human version of the open field test, called the human behavioral pattern monitor, which combines free exploration of a novel environment with presentation of novel objects, thus allowing for translational studies of anxiety-like behavior. In addition, the elevated plus-maze has been adapted to measure stretch-attend postures, reflecting risk assessment. Tests of approach–avoidance behavior have been criticized as a test of anxiety-like behavior, as drugs may affect the drive to

explore independently of anxiety, which could be misinterpreted as an anxiety-like phenotype. In addition, these tests have been developed to test the anxiolytic effect of benzodiazepine drugs, and might therefore not be very effective at detecting novel anxiolytics that work through alternative mechanisms. Indeed, clinically effective serotonin-selective reuptake inhibitors (SSRIs) generally do not have strong anxiolytic effects in approach–avoidance tests.

A second approach to study anxiety-like behavior is through conflict tests, in which hungry or thirsty animals are punished with a mild electric shock for lever pressing for food (Geller–Seifter conflict test) or licking a spout for water respectively (Vogel conflict test). The resulting suppression of appetitive behavior is taken as anxiety-like behavior and is sensitive to anxiolytic drugs.

In the shock probe burying test for rats, animals are confronted with an electrified probe that delivers shocks on contact. Defensive burying of the probe with bedding material serves as the measure of anxiety-like effects. Defensive burying is also measured as anxiety-like behavior in the marble burying test, where animals are presented with novel marbles in a cage with bedding.

Importantly, there are tests for anxiety-like responses that do not measure anxiety-like responses as behavioral inhibition, as general sedative effects might be mistaken for anxiolytic effects in the tests described above. In the stress-induced hyperthermia test, animals are stressed with rectal placement of a thermometer, and the gradual increase in body temperature is measured as index of the effect of stress. In addition, light-enhanced startle is a test that uses a continuous bright light in the startle chamber, which gradually potentiates the startle response.

Social Anxiety-Like Behavior. Several tests of social anxiety have been developed in rodents and NHPs. In the social interaction test for rats and mice, two animals are either allowed to interact freely or through a wire bar enclosure. Time spent interacting with a con-specific is quantified, and treatments that increase interaction time are considered anxiolytic. In addition, ultrasonic distress vocalizations are social signals that serve to indicate distress, and such signals in pups separated from mothers, or in mature animals in distressing conditions, can be taken as a measure of anxiety.

Although the most common animal models of anxiety are based on rat or mouse studies, the human intruder test for primates is well validated and offers particularly good face and constructs validity as a model of human social anxiety. In this test, confined macaque or marmoset primates are confronted with the presence of a human who does not look at them directly. The mere presence of the human introducer causes behavioral inhibition, aggression, and displacement stress responses, such as

scratching and autogrooming. The behavior of animals in this test has been shown to be sensitive to clinically effective anxiolytic drugs.

Depressive Behaviors

Reduced Effort to Escape Aversive Stimuli. Hopelessness is critical symptom of depression, and the rodent forced swim test is probably the most widely used acute behavioral screen for antidepressant activity. In this test, rodents are placed in a cylinder filled with water, and while the animals will initially try to escape the cylinder with vigorous escape movements, eventually the animals will show immobility. Antidepressant effects are inferred by an increase in the time spent in escape movements. A related task for mice is the tail-suspension test, in which mice are hung by the end of their tail and a similar transition from active escape movements to immobility occurs. Various classes of clinically effective antidepressants produce acute reductions in immobility in these tests, suggesting good predictive validity. In addition, lower doses of antidepressants show an effect after chronic but not acute treatment, in line with the delayed effectiveness of antidepressants for humans. It has been observed that the forced swim test has been developed to detect the effects of antidepressants, and so it is unclear whether this test will measure putative clinically effective antidepressant treatments that operate through a different mechanism of action. In the learned helplessness paradigm, uncontrollable stress has been shown the lead to reductions in escape attempts in subsequent situations where the stressor can actually be evaded. Such escape deficits are also reversible by antidepressant drugs.

Reduced Effort to Obtain Rewards. Anhedonia, or the loss of interest in otherwise rewarding activities, is also a key symptom of depression, which can be modeled in rodents. Preference for palatable sucrose over water, or the maximum number of lever presses an animal is willing to expand to obtain food in the progressive ratio test, can both provide measures of anhedonia. In addition, chronic implantation of ICSS electrodes in reward centers of the brain can allow for repeated measures of the threshold current at which animals are willing to self-stimulate, thus allowing for a more robust measure of reward function.

Chronic Stress Models. Finally, antidepressant activity can also be measured in rodent tests that use chronic stress manipulations. In the chronic mild stress paradigm, experimenters administer various mild unpredictable stressors for prolonged periods, and the effect of chronic stress can be measured as reduced sucrose preference and reduced grooming behavior in mice. Although the chronic mild stress model has good face and predictive validity, attempts to establish the model in different laboratories have not always been successful.

In the chronic social defeat model, subject mice are introduced in the home environment of a bigger mouse each day for 10 days. The sequential defeat experience leads to a relative avoidance of a novel mouse in the social interaction test, as well as reduced sucrose preference. The study of susceptible and resilient animals in the chronic social defeat stress model has yielded interesting insights into the molecular biology underlying this behavior.

POSITIVE VALENCE

Mania

Sleep disturbance often triggers manic episodes in bipolar patients, and sleep deprivation paradigms in rodents can produce short-term effects that resemble some aspects of mania, such as hyperactivity, stereotypy, insomnia, aggressive behavior, and hypersexuality. Typically, rats or mice are placed in a small platform surrounded by water that does not allow for sleep without falling into the water, and 72 hours of sleep deprivation is sufficient to produce a brief manic-like state. Interestingly, these behaviors are sensitive to chronic treatment of mood stabilizing drugs such as lithium. In addition, acute amphetamine injections in rats have been used to induce hyperactivity.

Impulsivity

Impulsivity can broadly be divided into actions that are initiated prematurely or without appropriate cognitive control, and actions that favor options with short-term payoff, without taking longer term consequences into consideration.

Impulsive Action. Behavioral inhibition can be measured in rodents, NHPs, and humans in the stop signal reaction time task. In the rat version, animals are trained to respond to a Go stimulus, unless a No-go stimulus is concurrently presented. The No-go stimulus is presented at various delays after the Go stimulus, enabling a quantification of the maximal delay at which a prepotent response can still be stopped. Clinically effective anti-impulsive drugs such as methylphenidate and atomoxetine decrease reaction time in both rodents and humans. Another test for impulsive action is embedded in the 5-choice serial reaction time task for rats. In this visuospatial attention task, after trial initiation, animals are faced with a standard intertrial interval before a visual stimulus prompts the correct response. When this intertrial interval is lengthened, the rats tend to show premature responses, indicative of motor impulsivity. Clinically effective anti-impulsive medications generally reduce premature responses in this task. Thus, both measures of motor impulsivity have good face and predictive validity, but their operation requires substantial training.

Impulsive Choice. Impulsive choices can be measured in delayed discounting paradigms in rodents, NHPs, and humans. Typically, subjects can choose between one option that results in an immediate small reward and another option that results in a delayed larger reward. The relative discounting of future rewards can be deduced from choice behavior and can serve as an index of impulsivity. In animals, delayed reinforcement can only be delivered through experience. In contrast, human studies can use either experience of actual delayed rewards inside the timeframe of the experiment or description of future rewards (outside of the timeframe of the lab visit). Clearly, animal experiments can best be compared with human studies using experience of delayed outcomes. Clinically effective anti-impulsivity drugs such as atomoxetine and methylphenidate have been shown to reduce choice of smaller immediate rewards.

Compulsive Behaviors

Repetitive invariant behavior with no apparent function or goal can easily be measured in its natural context. For example, barbering in mice, circling behaviors in dogs and rocking, and scratching or head banging in monkeys can occur naturally in stressful situations. Experimentally induced forms of compulsive behaviors include burying of unfamiliar objects such as marbles, and the exaggerated drinking behavior that is observed in rats that respond for food under a lean fixed time schedule of reinforcement (schedule induced polydipsia). Both types of compulsive behaviors respond well to clinically effective SSRIs treatment for patients suffering from OCD.

More complex cognitive measure of perseverative responding can be obtained in operant tasks for rodents and primates. Continued responding to a previously reinforced lever in reversal learning is one such measure of perseverative responding. An analogous measure in the 5-choice serial reaction time task is based on repeated responding to a previously rewarded magazine. Signal attenuation is a similar, but more complex, task that involves lever pressing by rats that are presented with a sensory signal to mark reward delivery. When that cue is then presented without reward delivery (cue-extinction), subsequent lever pressing under extinction conditions is not always accompanied by magazine entries to check for food rewards. The resulting lever pressing is considered to be perseverative. Of these three cognitive tasks, signal attenuation is most responsive to clinically effective SSRIs.

Abnormal Eating

The combination of under-eating, over activity and severe weight loss that is often seen in anorexia nervosa patients can be modeled in rodents. In the activity-based anorexia model, rats or mice are allowed to exercise in a running wheel, while food is only available for a limited amount of time per

day. Paradoxically, access to the running wheel decreases food intake, which combines to produce severe weight loss that can lead to self-starvation. To date, no pharmacotherapies have been approved for the treatment of anorexia nervosa, so it is not clear what the predictive validity of this model is.

Excessive calorie intake can be measured by feeding rats or mice high fat and carbohydrate food in the so-called cafeteria diet, which induces gradual weight gain and various metabolic changes. Binge eating models restrict access to highly palatable food to short periods of time, and rats and mice show considerable increases in the intake of the highly palatable food over time, while intake of normal chow food decreases.

Substance Abuse

Drug Reward. Drug reward can be investigated in simple form by administering drugs of abuse to rodents and measuring the behavioral consequences. In the conditioned place preference test, drugs are repeatedly injected when the animal is confined to one of two compartments, while control injections are paired with the alternative compartment. Subsequently, in a free exploration test, animals will generally spend more time in the drug-paired environment, which is taken as an index of conditioned drug reward. Although this test is relatively simple to administer, it has limited face validity because the critical pathology of substance abuse is that patients actively seek out and control drug consumption. Another way to quantify the rewarding effects of drugs of abuse is to use the ICSS procedure, where a rewarding effect of drug is measured as a decrease in ICSS reward threshold.

Drug Self-Administration

Rodents and NHPs can be trained to orally or intravenously, via indwelling catheters, self-administer drugs of abuse. Animals will self-administer almost all drugs with abuse potential, except for hallucinogens like D-lysergic acid die thylamide (LSD). In the simplest form, stable drug self-administration can be established by requiring the animal to perform a simple response, and the effect of environmental, pharmacological, or genetic manipulations on self-administration can be studied. There are numerous behavioral manipulations, including altering the response requirements, which can be implemented to better understand variables that influence the willingness of the animal to seek or consume the drug. For example, a dose–response function can be established to assess changes in sensitivity to drug dosing (leftward or rightward shift), or the reinforcing effect of the drug (upward or downward shift). The progressive ratio is a related test that measures the maximal number of lever presses animals are willing to emit to obtain a fixed dose of the drug. Although

intake of drugs of abuse tends to be well controlled and stable in short daily sessions, extended access to many drugs of abuse leads to a progressive escalation of drug intake over days. Finally, after prolonged drug self-administration, some animals will continue to seek drugs when their drug-seeking responses result in contingent punishment such as foot shock, suggesting compulsive drug seeking.

Substance abusers typically attempt to quit drug consumption, but often fail to do so, and many models of addiction involve an abstinence phase after self-administration. Immediately after cessation of self-administration, animals may undergo psychological or physical withdrawal symptoms. After a period of abstinence or extinction of the drug-seeking response, reinstatement of drug seeking can be precipitated by the drug-seeking context, response-contingent presentation of the drug-paired cue, stressors like foot shock, or administration of the drug by the experimenter. Incubation of craving refers to the finding that the magnitude of reinstatement tends to increase with the duration of abstinence. Finally, although most drug self-administration tests are based on free operant procedures, discrete trial procedures allow for discrete choice between drugs of abuse and natural rewards.

There are few effective and approved medications for the various substance abuse disorders, so it is mostly difficult to evaluate the predictive validity of substance abuse models. However, since substance abuse in humans is the result of prolonged substance consumption, these models have excellent face and construct validity.

COGNITION

Preattentive Processing

Prior to conscious processing, sensory information is selected and processed in the brain to facilitate further analysis. Preattentive processing can be measured in rodents, NHPs, and humans in the prepulse inhibition test. In this test, a loud startle tone is preceded by a less loud pretone with a specific latency. The presentation of the less loud pretone leads to an inhibition of the startle response to the loud tone, thus producing prepulse inhibition. Prepulse inhibition is not only disrupted in psychotic schizophrenia patients but also in many other psychiatric patients. Antipsychotic activity of therapeutic agents can be measured in combination with administration of dopaminergic agents that impair prepulse inhibition, such as apomorphine. The P50 response is a similar measure of preattentive processing in rodents and humans. In this test, two identical auditory stimuli are presented in quick succession, and the second stimulus normally produces a less strong P50 event related potential.

Attention

Training subjects to report the presentation of discrete signals over prolonged periods of time, typically tests attention, or the process of selectively focusing on specific information. There are several sustained attention tasks for rodents, based on analogous continuous performance tasks for humans. These tasks differ somewhat in that humans typically do not undergo training before testing. In the aforementioned 5-choice serial reaction time task, rats are trained to report the presentation of a stimulus light in one of five locations. In the behavioral vigilance task, rats are required to report the presence or absence of a visual stimulus by different responses, and background noise is used to increase the difficulty of the task.

Working Memory

The capacity to transiently hold on to and process new and already stored information can be tested in rodents, NHPs, and humans. Behavioral models of working memory depend on a delay between information that informs correct responding and the opportunity to execute the response.

In delayed alternation tasks, rats are trained to select alternating arms in a T-maze, either based on sequential free choice or with alternating forced choices and free choice trials. In the radial arm maze, eight arms are baited with a food reward, and the frequency of reentry in already visited arms serves as an index of working memory function. Operant delayed nonmatching to place is another rodent task that has a direct primate analogue. In this task, rats or mice are presented with a stimulus lever (right or left) that is subsequently retracted, and after a variable delay the animals have to press the alternative lever in order to earn a reward.

Rodent models can provide valuable information on the neural basis of working memory, and can serve as a first screen for potential new drugs. However, evolutionary differences in cognitive capacity and prefrontal cortex architecture suggest that NHP models of working memory provide a closer approximation of human working memory. Primate tests of working memory focus on spatial or object working memory, which depend on dissociable parts of the prefrontal cortex. In the spatial task, a reward is presented in one of several locations, after which a screen is lowered for various delays, and when the screen is lifted up, the animal can find a reward if it looks in the same location. In the delayed match to sample object task, primates are presented with multiple objects, and they have to select the matching object after a delay in order to be rewarded. Critically, the objects have to be reused frequently for the same animals, so that they cannot solve the task based on recognition memory. The CANTAB cognitive test battery for humans contains an analogous task.

Declarative Memory

Animal models of declarative memories have been developed for rodents and NHPs. The Morris water maze for rats tests spatial learning and memory. Rats are introduced at varying start locations in a large swimming pool with opaque water with a hidden submerged platform. Over multiple trials, rats will learn to swim to the platform regardless of the fact that their starting location is constantly varied. Subsequently, when the platform is removed, the rats will spend more time swimming in the location where the platform was previously located, indicating good spatial memory. The spontaneous object recognition task for rats is based on a sampling phase with two identical and novel objects at two ends of a y-maze. After a variable retention delay, one of the objects is replaced by a new object, and rats with intact memory will naturally explore the novel object more. The most popular NHP model of object recognition is the delayed nonmatch to sample task. The basic setup of the task is similar to the object working memory task described above, except that the objects are new to the animal at the time of testing, so that the task can be solved based on recognition memory.

SOCIAL FUNCTIONING

There are many rodent tests of social functioning. In addition to social anxiety-like behavior described above, there are specific models of social attachment, dominance, and aggression.

Social Attachment

Mother–infant attachment is often studied in rodents in the homing test, where pups at postnatal days 8 to 10 are separated from their mothers, and allowed to choose between bedding from their own nest and bedding from another age-matched litter. The pups will prefer bedding from their own nest, indicating normal social function. Similarly, isolated pups will emit specific ultrasonic vocalizations upon separation from the mother. Further in development, the social functioning of mice can be assessed in the social preference test, in which the time spent near a confined social stimulus versus a nonsocial stimulus (social preference) or familiar social stimulus versus novel social stimulus (novelty preference) can be measured. In addition, social interactions can also be studied in freely interacting con-specifics. Social play in young animals can be quantified in rats by placing two unfamiliar young rats together and observing the frequency of species-specific playful behavior. In addition, the rewarding qualities of social play can be quantified by a place preference procedure. Finally, the formation of monogamous pair-bonds between partners is often studied in another rodent species: the prairie vole. If a male prairie vole stays with the female for 1 day after mating, they will form a lifelong pair-bond, even to the extent that males will not seek a new mate after the female dies.

Social Hierarchy and Aggression

Both mice and rats tend to form stable hierarchies in the natural setting and engage in aggressive behaviors under particular circumstances. Male aggression tends to occur between males in order to establish dominance hierarchies and toward intruders in order to protect reproductive success. Female aggression in rodents tends to occur specifically postpartum to protect offspring. Dominance and aggression can be observed in natural and seminatural settings such as the visible burrow system. Dominance hierarchies can also be studied in mice in the tube test, in which pairs of mice are placed opposite of each other at the ends of a tube, and the mouse that forces the other mouse to retreat is considered to be the dominant mouse. Further, aggression is often provoked by the resident intruder test, sometimes after social isolation, which enhances resident aggression.

SLEEP AND AROUSAL

Arousal

Arousal and hyperactivity can be measured in rats and mice, typically through automated analysis of locomotor behavior in the home cage or in testing equipment like the open field test. Such tests can quantify overall movement and the relative occurrence of rest and activity.

Sleep

Natural Sleep. Natural sleeping behaviors are most often studied in rats and mice. Rodent models of sleep have been informative for human sleep regulation; however, unlike humans, rodents tend to sleep during the dark phase and show highly fragmented sleep episodes. Thus, results from rodent experiments should be interpreted with some caution. Sleep can be quantified on the basis of behavioral observation or electroencephalographic (EEG) measures. More detailed analyses focus on the amount of sleep per day and the temporal distribution of sleep.

Insomnia. Inability to sleep has been modeled in rats and mice through the effects of stress manipulations on sleep architecture. Stressors may include immobilization, social stress, or fear conditioning, and these manipulations generally prolong the latency to fall asleep, reduce sleep duration, and increase fragmentation of sleep. Most of these changes are reversible by the sleep medications.

Narcolepsy. Drowsiness during normal waking hours and sudden attacks of sleep have been modeled successfully in dogs and rodents. There are several strains of dogs that have naturally occurring narcolepsy, and the investigation of the genetic determinants of narcolepsy in dogs led to the identification of the orexin 2 receptor and its natural peptide ligand orexin

as critical for normal arousal and wakefulness. Sudden arrest of movement, or cataplexy, is a hallmark of narcolepsy and can be measured in dogs in the food-elicited cataplexy test, in which bits of food are placed along a semicircle and sudden arrests in movement are taken as partial or complete cataplexy.

Analogous mouse models of cataplexy have been developed to facilitate the investigation of the effect of genetic manipulations on arousal and sleep systems. In susceptible mice, cataplexy can be elicited by various manipulations including anticipation of food, wheel running, and social interactions. Rodent cataplexy can be measured as brief episodes of behavioral immobility, in particular atonia of muscles in the back of the neck and a sleep-like EEG signature.

Sexual Function

Normal Sexual Behavior. Normal and dysfunctional sexual behavior have most often been studied in rodents. Although sexual interaction in rodents has important species-specific components, such as a strong role for olfaction and a series of approaches and withdrawals that does not typically occur in humans, many aspects of sexual behavior are preserved between rodents and humans. Normal sexual function can be characterized in rodents by observation studies, and male ultrasonic vocalizations toward females can be measured, in addition to approach, female lordosis, mounting, intromission, and ejaculation.

Sexual Dysfunction. Although not often studied, stress can have detrimental effects on male sexual behaviors in rats, leading to longer latencies to mounting and intromission, but shorter latencies to ejaculation. In addition, there are many models of male erectile dysfunction that are based on physiological manipulations, such as cavernous nerve damage.

FUTURE DIRECTIONS

The value of animal models rests on their ability to reveal insights into psychological and neural processes that contribute to psychiatric disorders in humans, largely reflected by their success in facilitating the development of novel and efficacious therapeutics. The complement of behaviors displayed by rodents, and alterations in those behaviors resulting from perturbations in brain processes, is often very different from those observed in humans. This highlights the difficulties and challenges in modeling human psychiatric illness based primarily on behavior in rodents and other nonhuman species. The advent of optogenetic and chemogenetic technologies has provided powerful new mechanisms to modify discrete brain circuits in laboratory animals. Similarly, methods to manipulate

brain circuits in humans continue to improve, such as transcranial magnetic stimulation (TMS), DBS, and emerging new technologies. It seems likely that progress in understanding biological processes relevant to psychiatry will be seen when there is a better understanding of the impact of manipulating activity of brain circuits in humans and nonhuman species relevant to the five broad classes of symptoms identified by NIMH and outlined above. Those behaviors that show consistent cross-species alterations in response to manipulating symptom-relevant brain circuits are likely to be those that yield better insights to psychiatric illness and support new medications development. Also important for future progress in animal modeling of psychiatric will be the ability to introduce alleles into the genomes of laboratory animals that in humans are associated with the risk of psychiatric illness. Emerging genome editing technologies that permit the introduction of one of more risk alleles into specific populations of neurons in laboratory animals will likely drive better understanding of precisely how psychiatric disorders develop in humans, and the behaviors in nonhuman laboratory rodents that may reflect the impact of such risk alleles on the functioning of impacted brain circuits.

References

Agmo A. Animal models of female sexual dysfunction: basic considerations on drugs, arousal, motivation and behavior. *Pharmacol Biochem Behav*. 2014;121:3–15.

Belin D, Jonkman S, Dickinson A, Robbins TW, Everitt BJ. Parallel and interactive processes within the basal ganglia: relevance for the understanding of addiction. *Behav Brain Res*. 2009;199:89–102.

Blanchard RJ, Flannelly KJ, Blanchard CD. Life-span studies of dominance and aggression in established colonies of laboratory rats. *Physiol Behav*. 1988;43:1–7.

Blanchard CD, Griebel G, Blanchard RJ. The mouse defense test battery: pharmacological and behavioral assays for anxiety and panic. *European J Pharm*. 2003;463:97–116.

Burkett JP, Young LJ. The behavioral, anatomical and pharmacological parallels between social attachment, love and addiction. *Psychopharmacology (Berl)*. 2012;224:1–26.

Camilla D'Angelo LC, Eagle DM, Grant JE, Fineberg NA, Robbins TW, Chamberlain SR. Animal models of obsessive-compulsive spectrum disorders. *CNS Spectr*. 2014;19:28–49.

Castner SA, Goldman-Rakic PS, Williams GV. Animal models of working memory: insights for targeting cognitive dysfunction in schizophrenia. *Psychopharmacology (Berl)*. 2004;174:111–125.

Chung E, De Young L, Brock GB. Investigative models in erectile dysfunction: a state-of-the-art review of current animal models. *J Sex Med*. 2011;8:3291–3305.

Coleman K, Pierre PJ. Assessing anxiety in nonhuman primates. *ILAR J*. 2014;55:333–346.

*Cryan JF, Holmes A. The ascent of mouse: advances in modelling human depression and anxiety. *Nat Rev Drug Discov*. 2005;4:775–790.

Cryan JF, Markou A, Lucki I. Assessing antidepressant activity in rodents: recent developments and future needs. *Trends Pharmacol Sci*. 2002;23:238–245.

*Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med*. 2013;11:126.

*Eichenbaum H. Exploring declarative memory using animal models. In: *The Cognitive Neuroscience of Memory: An Introduction*. New York: Oxford University Press; 2011.

*Evenden JL. Varieties of impulsivity. *Psychopharmacology (Berl)*. 1999;146:348–361.

Gestel van MA, Kostrzewa E, Adan RAH, Janhunen SK. Pharmacological manipulations in animal models of anorexia and binge eating in relation to humans. *Br J Pharmacol*. 2014;171:4767–4784.

Griebel G, Holmes A. 50 years of hurdles and hope in anxiolytic drug discovery. *Nat Rev Drug Discov*. 2013;12:667–687.

Hyman SE. Revitalizing psychiatric therapeutics. *Neuropsychopharmacology*. 2014;39:220–229.

Hyman SE, Malenka RC, Nestler EJ. Neural mechanisms of addiction: the role for reward-related learning and memory. *Annu Rev Neurosci*. 2006;29:565–598.

Markou A, Chiamulera C, Geyer MA, Tricklebank M, Steckler T. Removing obstacles in neuroscience drug discovery: the future path for animal models. *Neuropsychopharmacology*. 2009;34:74–89.

Miczek KA, Maxson SC, Fish EW, Faccidomo S. Aggressive behavioral phenotypes in mice. *Behav Brain Res*. 2001;125:167–181.

Oddi D, Crusio WE, D'Amato FR, Pietropaolo S. Monogenic mouse models of social dysfunction: implications for autism. *Behav Brain Res*. 2013;251:75–84.

Peleg-Raibstein D, Feldom J, Meyer U. Behavioral models of antipsychotic drug actions. In: *Current Antipsychotics, Handbook of Experimental Pharmacology*. Heidelberg: Springer-Verlag; 2012.

*Sarter M, Givens B, Bruno JP. The cognitive neuroscience of sustained attention: where top-down meets bottom-up. *Brain Res Brain Res Rev*. 2001;35:146–160.

Schenberg LC, Schimitel FG, de Souza Armini R, et al. Translational approach to studying panic disorder in rats: hits and misses. *Neurosci Biobehavioral Rev*. 2014;46:472–496.

Shaham Y, Shalev U, Lu L, de Wit H, Stewart J. The reinstatement model of drug relapse: history, methodology and major findings. *Psychopharmacology (Berl)*. 2003;168:3–20.

Toth LA, Bhargava P. Animal models of sleep disorders. *Comp Med*. 2013;63:91–104.

Wang F, Zhu J, Zhu H, Zhang Q, Lin Z, Hu H. Bidirectional control of social hierarchy by synaptic efficacy in medial prefrontal cortex. *Science*. 2011;334:693–697.

▲ 1.20 Pain Systems: Interface with Affective and Motivational Mechanisms

JON-KAR ZUBIETA, M.D., PH.D.

INTRODUCTION

Pain is a cardinal symptom of many somatic diseases. Acutely, it serves to warn of danger and shapes behavior to avoid the environment eliciting the stimulus. It is also a subjective experience that incorporates not only sensory (e.g., intensity, location, quality), but also affective and cognitive dimensions. Its chronification is associated with adaptive and maladaptive neurobiological changes that serve to perpetuate the experience, and leads to an increase in its aversive and motivational significance. In individuals for whom acute pain transitions into a persistent condition, its experience no longer serves a protective role and is frequently associated with the development of depression and anxiety comorbid conditions in approximately 50 percent of cases, with bidirectional exacerbation.

While the modulation of acute pain can be achieved pharmacologically by the introduction of agents that reduce nociceptive input (e.g., opioid compounds, pharmacological conduction blockade and modulation, monoamine reuptake inhibitors), the relief of persistent pain is not adequately achieved in many conditions, particularly those without an obvious or identifiable pathophysiology (e.g., fibromyalgia, irritable bowel syndrome, nonneuropathic back pain). The negative effects of chronic opioid use have received much attention because of the development of tolerance to these compounds, withdrawal symptoms and their potential for misuse, albeit it remains a commonly used strategy. The potentiation of monoaminergic neurotransmission through inhibition of the reuptake or similar mechanisms, together with agents that reduce inflammatory responses and neuronal excitability remain the main nonopioid strategies to reduce pain severity and comorbidity, albeit they rarely completely eliminate the experience of pain once it becomes persistent.

PAIN SYSTEMS

Pain perception is typically initiated by the activation of peripheral nociceptors, sensory end organs in the skin, muscle, joints, and viscera that selectively respond to noxious or potentially damaging stimuli. An important property of nociceptors is that they sensitize as a consequence of tissue injury and inflammation, reducing their threshold of response and increasing the magnitude of their responses to noxious stimulation. Nociceptor endings in tissue are unencapsulated ("free") and are commonly associated with medium diameter myelinated Aδ axons which transmit

acute, well-localized "first pain" signals, or with small diameter, unmyelinated slowly conducting C fibers that convey poorly localized, "second pain," "burning," or slow pain. Aδ and C nociceptors are further subdivided into multimodal or more specialized types, responding to mechanical, chemical, or heat stimuli with varying degrees of sensitivity. Larger, rapidly conducting myelinated Aβ fibers respond to innocuous mechanical stimulation such as touch.

The cell bodies of the nociceptors are located in the dorsal root ganglia, or in the trigeminal ganglion in the case of the facial area, and are organized into anatomically distinct laminae. Axonal projections from these neurons transmit information through excitatory glutamatergic and peptidergic neurotransmission to second-order neurons of the dorsal horn of the spinal cord. Second-order neuron axons then cross contralaterally and form ascending pathways that synapse in the brainstem (spinoreticulothalamic tract) and thalamic nuclei (spinothalamic tract), thought to respectively subserve less localized and more sensory-discriminative (e.g., intensity and localization) elements of pain.

Spinoreticulothalamic projections to the rostroventral medulla (RVM) and periaqueductal gray (PAG) engage descending pain inhibitory pathways in feedback loops that regulate dorsal horn neuronal excitability, while thalamic nuclei project to cortical structures. Other spinoreticulothalamic projections reach the parabrachial nucleus of the pons, which is directly connected to the amygdala, a region implicated in the representation and modulation of aversive, affectively negative elements of the pain experience, and which provides further input to the insular cortex.

The regulatory functions of the PAG and RVM have received much attention since they are engaged as a modulatory mechanism of the nociceptive signal which is heavily influenced by neurotransmitters such as the enkephalinergic endogenous opioid, serotonin, and norepinephrine, among others. These are the targets of commonly used pain control agents, such as the monoamine reuptake inhibitor antidepressants, and the opioid medications. Pain signal to the PAG activate excitatory and peptidergic transmission which projects to the RVM. In the RVM, "on" cells become activated in response to noxious painful stimuli and are involved in the formation of protective "nocifensive" behaviors, such as withdrawal reflexes. "Off" cells, on the other hand, respond acutely to endogenous opioids, GABA, and serotonin, to induce analgesic responses through descending connections to the dorsal horn of the spinal cord. Norepinephrine, through alpha 1 and alpha 2 receptors, respectively, can have both pro- and antinociceptive effects, by modulation of "on" and "off" cells of the RVM (Fig. 1.20–1).

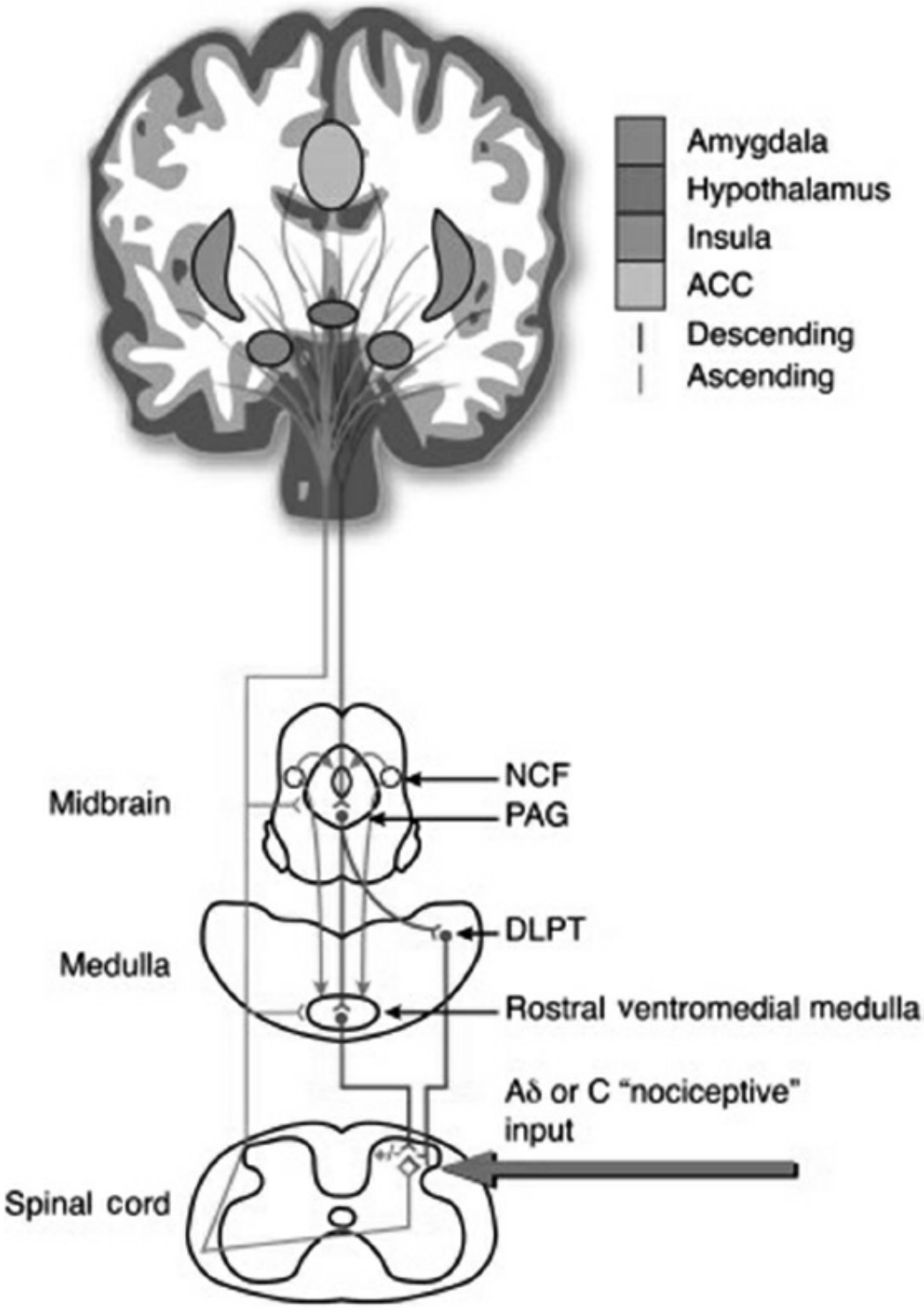

FIGURE 1.20–1. Descending pain modulatory system. NCF, nucleus cuneiformis; PAG, periaqueductal gray; DLPT, dorsolateral pontine tegmentum; ACC, anterior cingulate cortex; +/− indicates both pro- and antinociceptive influences, respectively. (From Tracey I, Mantyh PW. The cerebral signature for pain perception and its modulation. *Neuron*. 2007;55(3):377–391.)

Thalamic input reaching the somatosensory cortex is implicated in sensory-discriminatory functions, while projections to regions such as the anterior cingulate and insular cortex are thought to integrate sensory and

emotional aspects of the experience. Often referred as “lateral” (sensory-discriminatory) and medial (affective–cognitive) projections, they form a large distributed brain network, “pain matrix” that reflects and processes the complex, multifaceted experience of pain. During acute pain, primary and secondary somatosensory cortices, insular cortex, anterior cingulate, and prefrontal cortices become consistently activated. Basal ganglia, amygdala, and hippocampus are also involved in these processes.

It should noted that many of these “medial projection” integrative areas overlap with those implicated in the representation of other, nonpainful, salient or stressful stimuli, and are thought to be dysregulated when pain becomes persistent, chronic, reflecting alterations in reward and motivational mechanisms observed in persistent pain conditions. Notably among these, anterior insular, anterior cingulate, and prefrontal cortices, as well as the nucleus accumbens and amygdala are thought to be implicated in the representation and regulation of emotional, motivational, and evaluative aspects of the pain experience and will be discussed in further detail in the following paragraphs. Conversely, ventrolateral thalamus, somatosensory cortex, dorsal posterior insula, and some, but not all, cellular groups of the anterior cingulate and anterior insular cortex are thought to be more pain-specific, responding to variations in pain intensity and being differentially activated by pain stimuli as opposed to non–pain-specific sensory stimuli or other components of the experience of pain. In addition to the modulatory role of the PAG and RVM on nociceptive suppression, telencephalic regions that include the DLPFC, anterior cingulate and insular cortices, thalamus, nucleus accumbens, amygdala are implicated in regulating the pain experience across sensory, unpleasantness, and affective components through the local activation of endogenous opioid and other neurotransmitter systems such as the noradrenergic, GABA, endocannabinoid, and glycinergic, as well as through descending projections to the PAG and RVM (Fig. 1.20–2).

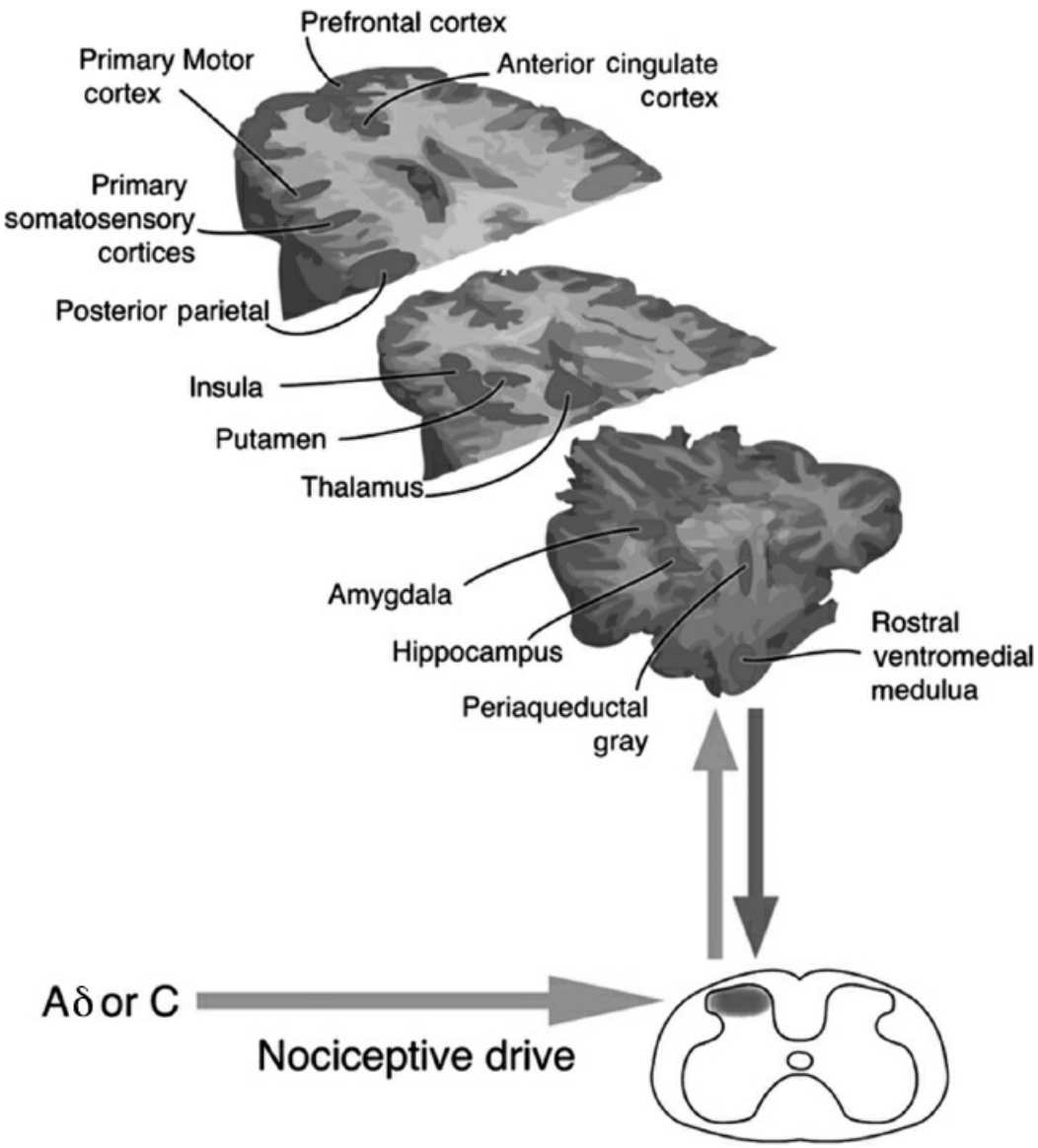

FIGURE 1.20–2. Neuroanatomy of pain processing. Brain regions that activate during a painful experience, with bilateral representation (contralateral to the noxious stimulus). (From Tracey I, Mantyh PW. The cerebral signature for pain perception and its modulation. *Neuron*. 2007;55(3):377–391.)

PAIN AND MOTIVATIONAL MECHANISMS

From an evolutionary perspective, pain is a strong motivational signal with negative affective connotations, while pain relief can be conceptualized as a form of relief and reward, and therefore represent potent factors directing behavior. The basal ganglia and associated limbic cortex are thought to integrate internal and external signals to encode the value of the sensory signals received. They are engaged during anticipation of pain or expectations of pain relief, and together with prefrontal cortical structures, are implicated in the integration of sensory information but also the modulation of pain through endogenous opioid modulatory systems.

During acute pain, brain functional activity of the saliency/reward/motivation network, which includes the orbitofrontal

cortex, nucleus accumbens, and VTA, is increased and associated with the capacity to express endogenous opioid analgesia in humans. Interindividual variations in the capacity of the nucleus accumbens to respond to reward value have been further linked to the capacity to develop placebo-induced analgesic responses, suggesting that reward and saliency evaluation mechanisms are also engaged to induce appropriate analgesic responses when required to maintain homeostasis.

The nucleus accumbens receives pain information through regions involved in pain and emotional processing, such as the prefrontal cortex, anterior cingulate, and the amygdala, and direct projections from spinal cord neurons to the nucleus accumbens have been described. During experimental pain in humans, dopamine neurotransmission is activated in the basal ganglia. In dorsal regions, receiving input from somatosensory cortex, dopamine is released in a manner proportional to the intensity and sensory elements of the pain signal, while in the nucleus accumbens it was related to negative affective responses in humans. This effect appears to be localized in the nucleus accumbens core, based on animal model data, while in the shell area of this region, it is released after pain stimulus termination. This is consistent with the concept that pain relief represents a form of positive reinforcement upon its termination, as shown using neuroimaging data in humans diagnosed with persistent pain. The transition from injury and acute pain to persistent low back pain in humans has been further associated with alterations in the functional connectivity of the basal ganglia to prefrontal cortical structures, suggesting that persistent pain-induced alterations in motivational systems from a critical part in the perpetuation of the experience of pain in humans.

In patients diagnosed with persistent pain conditions, such as fibromyalgia and chronic back pain, dopamine receptors and dopaminergic neurotransmission in the nucleus accumbens are downregulated compared to those of healthy individuals. These reductions appear adaptive in nature, hampering the salience of the chronic noxious stimulus but also enhancing the functional capacity of downstream antinociceptive endogenous opioid mechanisms. However, dysregulation of nucleus accumbens function in the context of chronic pain have also been linked to reductions in motivation, such as the capacity to perform effortful tasks to gain rewards in animal models. Therefore, nucleus accumbens function and dopaminergic neurotransmission in this area are involved in the signaling of the negative value and salience of noxious input, and has been implicated in the transition from acute to persistent pain. Dopaminergic mechanisms also seem to become adaptively downregulated in the context of persistent pain, however with potentially negative behavioral consequences, such as the capacity to engage motivational mechanisms and the effortful pursuit of rewards, core features of anhedonia and depression.

PAIN AND AFFECTIVE REGULATION AND DYSREGULATION

Aside from its adaptive value as a warning signal, pain, when persistent, becomes a form of physical and emotional stressor that induces negative mood states, part of maladaptive responses to chronic noxious stimuli. Clinically significant symptoms of depression and anxiety appear in approximately half of the patients diagnosed with persistent pain conditions, and form part of the comorbid pathology associated with chronic pain.

Brain regions such as the anterior cingulate, which receives nociceptive signals through the spinothalamic pathway, are implicated in the encoding of pain unpleasantness, which can be modulated by cognitive processes such as hypnotic suggestions and where endogenous opioid mechanisms are engaged in the context of suggestions of analgesia during placebo administration. Opioid receptors mediating endogenous opioid analgesia are found downregulated in this and other regions, such as the thalamus, nucleus accumbens, and amygdala in chronic pain states such as inflammatory, central, and neuropathic pain as well as in fibromyalgia, with accompanying enhancements in negative affective and pain unpleasantness ratings. Similarly, and in chronic back pain patients, deficits in endogenous opioid neurotransmission have been encountered in the amygdala, a region implicated in both antinociception but also responses and regulation of affectively valenced stimuli, and have been associated with both clinical pain ratings and greater negative affect. These findings parallel those encountered in patients diagnosed with major depression, and are an example of alterations in biological processes that overlap between pathologies.

The anterior insular cortex, one of the regions that mediates responses to emotional stimuli and is implicated in the pathophysiology of depression has also been shown to have an important role in the encoding of interoceptive information in the context of pain. It has been linked to the formation of attentional biases toward physical physiological cues, including pain, and implicated in fear learning and anxiety responses to painful stimuli. Clinical interventions such as meditation-based therapies, geared toward the reduction of interoceptive sensitivity and anxiety in patients diagnosed with persistent pain have been shown to regulate anterior cingulate and anterior insular cortex in a manner proportional to clinical benefit.

As noted above, substantial overlap exists among brain regions and neurobiological mechanisms implicated in the processing of emotionally salient stimuli and those found to represent and regulate pain signaling, consistent with the concept of persistent pain as an allostatic process that promotes negative affective responses and enhances the individual propensity to the development of depression and anxiety symptomatology.

COGNITIVE PROCESSING AND PAIN

The perception of pain is sensitive to mental processes such as attentional biases toward the noxious stimulus. Heightened attention and expectations of threat toward painful stimuli have been shown to induce catastrophic thinking and higher pain sensitivity. Conversely, distraction from pain and reframing of its meaning have been associated with reductions in pain-associated responses of somatosensory cortex, thalamus, insula, and ACC, with accompanying changes in sensory, cognitive, and affective appraisals of the painful stimulus. Therefore, both enhancements and reductions in pain report can be achieved by the influence of higher-order cognitive processes, effects that at least in part, take place through the top-down engagement of the PAG and descending pain modulatory circuits impacting spinal nociceptive transmission. Executive cortical regions such as the dorsolateral and VLPFC are thought to provide input to anterior cingulate, insular and thalamic circuitry, which in turn project to PAG, a region that potently influences the intensity of ascending signals from the spinal cord.

Another example of how cognitive processes can affect pain processing is the so-called placebo analgesic effect. Besides the perspective that placebo effects confound controlled trials of potentially therapeutic agents, the information so far acquired points to neurobiological systems that when activated by positive expectations and maintained through reward learning are capable of inducing physiological changes that lead to the experience of analgesia and improvements in emotional state. This line of research has described neural networks and neurotransmitter systems implicated in the formation of placebo analgesic responses. These include the DLPFC, implicated in the processing of environmental cues and expectations, as well as VLPFC, anterior cingulate, insular cortex, nucleus accumbens, amygdala, and PAG. Salience assessments mediated by dopamine neurotransmission in the nucleus accumbens appear to engage pain regulatory networks across these regions, inducing the adaptive activation of pain and affect regulatory neurotransmitters and neural networks (or maladaptive in the case of nocebo responses, adverse effects in response to therapeutic cues). Endogenous opioids are centrally involved in these processes, but also the endocannabinoid and CCK systems. The formation of biological placebo effects therefore appear to implicate the engagement of resiliency mechanisms, with interindividual variations partially determined by genetic factors, and uncovered by the cognitive emotional integration of the expectations created by the therapeutic environment. Of note, similar mechanisms appear to take play in the formation of placebo antidepressant effects, with partial regional overlap with those engaged during the development of placebo-induced analgesia.

FUTURE DIRECTIONS

The majority of the work carried out in the study of pain perception and regulation has been concerned with the transmission of nociceptive input and its regulation at the level of peripheral, spinal, and brainstem structures, providing exquisite detail on these processes. By comparison, much less is known regarding the contribution of telencephalic mechanisms that are engaged during acute, sustained, and persistent pain, particularly in humans, and which undoubtedly contribute to chronicity and comorbidity with depression, anxiety, and alterations in reward and motivation. However, an emerging literature demonstrates substantial overlaps between neural systems engaged during pain and those that respond to other stressors, the dysregulation of which would explain the complex presentation of patients diagnosed with persistent pain conditions. Future advancements in the area will require particular attention to those interfaces to exert actual change in pain chronicity and associated neuropsychiatric symptomatology and to provide new insights into novel treatment targets for these conditions.

References

Apkarian AV, Bushnell MC, Treede RD, Zubieta JK. Human brain mechanisms of pain perception and regulation in health and disease. *Eur J Pain*. 2005;9:463–484.

Baliki MN, Geha PY, Fields HL, Apkarian AV. Predicting value of pain and analgesia: nucleus accumbens response to noxious stimuli changes in the presence of chronic pain. *Neuron*. 2010;66:149–160.

Baliki MN, Petre B, Torbey S, Herrmann KM, Huang L, Schnitzer TJ, et al. Corticostriatal functional connectivity predicts transition to chronic back pain. *Nat Neuroscience*. 2012;15:1117–1119.

Basbaum AI, Bautista DM, Scherrer G, Julius D. Cellular and molecular mechanisms of pain. *Cell*. 2009;139:267–284.

Bushnell MC, Ceko M, Low LA. Cognitive and emotional control of pain and its disruption in chronic pain. *Nat Rev Neurosci*. 2013;14:502–511.

Fields H. State-dependent opioid control of pain. *Nat Rev Neurosci*. 2004;5:565–575.

Gold MS, Gebhart GF. Nociceptor sensitization in pain pathogenesis. *Nat Med*. 2010;16:1248–1257.

Harris RE, Clauw DJ, Scott DJ, McLean SA, Gracely RH, Zubieta JK. Decreased central mu-opioid receptor availability in fibromyalgia.*J Neurosci*. 2007;27:10000–10006.

Hsu DT, Sanford BJ, Meyers KK, et al. It still hurts: altered endogenous opioid activity in the brain during social rejection and acceptance in major depressive disorder. *Mol Psychiatry*. 2015;20:193–200.

Lutz A, McFarlin DR, Perlman DM, Salomons TV, Davidson RJ. Altered anterior insula activation during anticipation and experience of painful stimuli in expert meditators. *Neuroimage*. 2013;64:538–546.

Martikainen IK, Nuechterlein EB, Pecina M, et al. Chronic back pain is associated with alterations in dopamine neurotransmission in the ventral striatum. *J Neurosci*. 2015;35:9957–

9965.

Martikainen IK, Pecina M, Love TM, et al. Alterations in endogenous opioid functional measures in chronic back pain. *J Neurosci.* 2013;33:14729–14737.

Navratilova E, Porreca F. Reward and motivation in pain and pain relief. *Nat Neurosci.* 2014;17:1304–1312.

Pecina M, Bohnert AS, Sikora M, et al. Association between placebo-activated neural systems and antidepressant responses: neurochemistry of placebo effects in major depression. *JAMA Psychiatry*. 2015;72:1087–1094.

Pecina M, Zubieta JK. Molecular mechanisms of placebo responses in humans. *Mol Psychiatry*. 2015;20:416–423.

Pertovaara A. The noradrenergic pain regulation system: a potential target for pain therapy. *Eur J Pharmacol.* 2013;716:2–7.

Phan KL, Wager T, Taylor SF, Liberzon I. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*. 2002;16:331–348.

Rainville P, Carrier B, Hofbauer RK, Bushnell MC, Duncan GH. Dissociation of sensory and affective dimensions of pain using hypnotic modulation. *Pain*. 1999;82:159–171.

Schwartz N, Temkin P, Jurado S, et al. Chronic pain. Decreased motivation during chronic pain requires long-term depression in the nucleus accumbens. *Science*. 2012;345:535–542.

Scott DJ, Heitzeg MM, Koeppe RA, Stohler CS, Zubieta JK. Variations in the human pain stress experience mediated by ventral and dorsal basal ganglia dopamine activity. *J Neurosci.* 2006;26:10789–10795.

Scott DJ, Stohler CS, Egnatuk CM, Wang H, Koeppe RA, Zubieta JK. Individual differences in reward responding explain placebo-induced expectations and effects. *Neuron*. 2007;55:325–336.

Scott DJ, Stohler CS, Egnatuk CM, Wang H, Koeppe RA, Zubieta JK. Placebo and nocebo effects are defined by opposite opioid and dopaminergic responses. *Arch Gen Psychiatry*. 2008;65:220–231.

Segerdahl AR, Mezue M, Okell TW, Farrar JT, Tracey I. The dorsal posterior insula subserves a fundamental role in human pain. *Nat Neurosci.* 2015;18:499–500.

Simons LE, Elman I, Borsook D. Psychological processing in chronic pain: a neural systems approach. *Neurosci Biobehav Rev*. 2014;39:61–78.

Smith TH, Sim-Selley LJ, Selley DE. Cannabinoid CB1 receptor-interacting proteins: novel targets for central nervous system drug discovery? *Br J Pharmacol.* 2010;160:454–466.

Tracey I, Mantyh PW. The cerebral signature for pain perception and its modulation. *Neuron*. 2007;55:377–391.

Vandenberg RJ, Ryan RM, Carland JE, Imlach WL, Christie MJ. Glycine transport inhibitors for the treatment of pain. *Trends Pharmacol Sci.* 2014;35:423–430.

Wager TD, Atlas LY, Lindquist MA, Roy M, Woo CW, Kross E. An fMRI-based neurologic signature of physical pain. *N Engl J Med*. 2013;368:1388–1397.

Wanigasekera V, Lee MC, Rogers R, et al. Baseline reward circuitry activity and trait reward responsiveness predict expression of opioid analgesia in healthy subjects. *Proc Natl Acad Sci U S A*. 2012;109:17705–17710.

Wiech K, Ploner M, Tracey I. Neurocognitive aspects of pain perception. *Trends Cogn Sci.*

2008;12:306–313.
Wood PB, Schweinhardt P, Jaeger E, et al. Fibromyalgia patients show an abnormal dopamine response to pain. *Eur J Neurosci*. 2007;25:3576–3582.
Zubieta JK, Smith YR, Bueller JA, et al. mu-opioid receptor-mediated antinociceptive responses differ in men and women. *J Neurosci*. 2002;22:5100–5107.
Zubieta JK, Smith YR, Bueller JM, et al. Regional mu opioid receptor regulation of sensory and affective dimensions of pain. *Science*. 2001;293:311–315.

▲ 1.21 Basic Science of Self

DEBRA A. GUSNARD, M.D.

This chapter is intended to provide an overview of considerations for what may be regarded as a basic science of self and its relevance to psychiatry. Terminological issues are discussed first to provide some perspective on historical and other phenomena that initially complicated and then, more recently, encouraged thinking that such a "science" might be possible. It should be noted at the outset that a single paradigm for research on "self" does not exist, as separate disciplines and even different investigators within disciplines have come to the study of the individual (qua self) and its unique perspective differently. This will be apparent in the descriptions of representative scientific work that follows. Despite these differences, however, some conceptual convergence in this scientific enterprise has begun to emerge along with signs of where a more advanced program for a science of self might be heading.

INTRODUCTION

Stated simply, self-representation is central to human behavior in health and disease. That is, an individual's abilities to experience and recognize his physical being as distinctly his own and to act on various kinds of self-knowledges are critical for his being able to organize and regulate his own behavior and engage well, or not so well, in social relationships. Until recently, the importance of self-representation and the value of understanding mechanisms by which it may be achieved have been generally under-appreciated in scientific circles. But this has begun to change as clinical disciplines concerned with various disorders bearing on aspects of self-representation have started to collaborate and communicate more directly with experimentalists and neuroscientists and the latter's methods and technologies have begun to be more widely applied.

Historically, the relative dearth of scientific studies targeting self-representation has been largely a consequence of the fact that, unlike most other domains in science where commonly agreed upon and well-elaborated conceptual frameworks guide investigation and constrain the

interpretation of data, the scientific investigation of self-representation has not been grounded by such a framework and, as a consequence, has been a relatively fragmented enterprise. Even acknowledging the meaningfulness of such work requires an endorsement of both a naturalistic perspective and the view that self-representation includes having a functionality for individuals that, in its typical operations, serves to enhance or, in its dysfunction, may hinder their survival. In this view, forms of self-representation inhere operations that support the distinguishing and recognition of an individual's perspective *as* his/her own spatiotemporal perspective as well as an individual's sense of ownership of his/her own mental and bodily states. Clearly, these forms of self-representation and their constitutive operations are necessarily grounded in material, specifically neural processes. General appreciation of such a view is becoming more widespread so that, like the study of consciousness to which it is related, the study of various means by which individuals represent aspects of "self" to themselves is acquiring increasing acceptance.

Certain scientific discoveries facilitated the burgeoning acceptance of such scientific inquiry. For example, the split-brain studies carried out by Roger Sperry and colleagues in the 1960s on individuals who had undergone separation of their two cerebral hemispheres by corpus callosotomy for intractable epilepsy revealed significant disconnection of the perceptions and decisions of one hemisphere from the other. These findings among others contributed to deflating the notion of a necessary unity and coherence in self and the recognition that the normal experience of such unity masks underlying neural organizational processes that may be subject to disruption and scientific manipulation.

Recent scientific collaborations between clinical and nonclinical disciplines have also provided impetus to this burgeoning acceptance. Such collaborations have stimulated new thinking on the subject due to the required communication and resulting awareness of their inherently different, though still functionalist, perspectives on considerations of self. Philosophers of mind are sometimes invited to participate in such cross-disciplinary collaborations to assist in refining and clarifying concepts to help with the design and interpretation of experimental studies. As a result, scientific inquiry into the nature of self-representation is now broadening both across levels of description (from psychological and behavioral levels of description to neurophysiological levels of description) and in terms of relevant capacities that change across human development and may exist in other animal species as well.

HISTORICAL CONSIDERATIONS

To appreciate the scope of the conceptual challenges faced by current investigators in refining a scientific agenda, however, an understanding of some relevant historical forces shaping considerations of self is useful. First

of all, the human drive to reflect upon what constitutes self has been recorded by philosophers, poets, artists, and others for millennia. But societal and cultural forces have played an important role in determining the expression of socially accepted intellectual interest in the self, so that such expression has clearly varied across time. For example, in the west, before the decline of the middle ages and the attendant decline of religion as the dominant organizational culture, emphasis was placed on the concept of community rather than on self in all aspects of social life. The significance of the community tended to overshadow the significance of individuals and their personal interests, which were viewed by the dominant culture as arising from a human nature that was inherently base and "selfish." Ideally, self-interest was to be suppressed in the interests of the common good and spiritual attainments. Major shifts in social thought began to occur with the dawning of the renaissance in the 14th century with its greater focus on the personal and temporal that was then followed by the enlightenment period in the 18th century with its emphasis on rationality, systematized thinking, and science. During these periods, the individual, personal identity, and personal expression became increasingly valued. Technological developments also played a role. Interestingly, it has been suggested that one of the more significant technological developments in this regard was the rise of the glass industry in Western Europe and the establishment of mirrors as a common commodity. The widespread availability of mirrors provided individuals for the first time in human history with a stable and accurate representation of their own faces, which allowed them to regard themselves in a different way and encouraged greater self-inspection, including more frequent introspection (for further discussion of this point, see *How We Got To Now: Six Innovations That Made the Modern World* (2014) by Steven Johnson).

Since then, political and economic changes and the rise of capitalist societies with their encouragement of personal initiative and positive regard for individual striving and success have promoted a veritable culture of individualism, particularly in North America and Western Europe. Such forces have all contributed to a modern intellectual climate that has made the concept of the individual and the nature of the self and subjectivity available for consideration and conceptual elaboration as never before.

Conceptualizations of self have been far from homogeneous as discussions in the modern era have unfolded, however. One of the fundamental questions targeted in the intellectual debate has, in fact, been whether there is such a thing as the self. David Hume in his *Treatise on Human Nature* in 1739 notoriously stated that, when he looked inside himself, he could find many perceptions but nothing linking them together. Since Hume's denial of the self, other influential philosophers, including Friedrich Nietzsche in the German tradition and Ludwig Wittgenstein in the analytical tradition, have also argued that the self is most reasonably

understood to be a fiction that may have some utility for those using the term but that it has no actual substance or basis in reality, that "the self [is simply] a center of narrative gravity," as Daniel Dennett, the American philosopher has famously put it. It is important to note that these careful arguments have largely tended to be in the context of being directed *against* particular notions of the self, such as notions that self is fundamentally a kind of disembodied and/or unitary essence, for example. Against such notions, the self may now generally be agreed to be a "fiction." But such arguments have not been undertaken from a perspective of attempting to describe and understand many of the details of functions that self-representation might actually serve, let alone the details of their neural instantiation.

Nonetheless, such apparent dismissals of "a self" have had significant influence and have found kinship with the scientific position that has appropriately regarded as untenable notion of an "entity" (i.e., homunculus) controlling brain processes and behavior. For some scientists, the very term "self" has been associated with this problematic notion of a homunculus within the brain and, as a consequence, has been rejected as a subject worthy of investigation.

Other scientists, however, have not been so dismissive and have approached the study of operations supporting an individual organism's subjectivity and sense of personal-state ownership from a conceptual vantage point that includes the consideration of biological necessity and evolutionary pressures. Neuroscientists Rodolfo Llinás and Walter Freeman, for example, have both used the term "self" to refer to that locus of properties that permit a mobile organism to sustain its viability *as* an individual organism. Though the details of their formulations differ, both consider the self from a comparative perspective and regard the primary *need* for a "self" to be related to mobile organisms' capacities for goal-directed behavior in an environment. That is, in contrast to organisms incapable of active movement, mobile organisms need to be able to predict the effects of their actions on their environment. Otherwise the outcomes of their actions would be purposeless and potentially dangerous. They need to be able to "own" their behavior (though this need not be conscious) and distinguish their behavior's effects from effects attributable to other forces in their environment.

There are also other scientists, notably social and personality psychologists, who have routinely employed self terminology in their work. However, their use of this term has tended to be relatively specific to the consideration of humans and their functioning as social beings. Usually the "self" under consideration in this research tradition has referred to some socially constructed aspect of a person, such as social role or other aspect of social identity.

Such differences in terminology regarding "self" have led to some

confusion as individuals from diverse disciplines have begun to use the same technologies, such as neuroimaging, in their experimental work. Because investigators have not always articulated their assumptions or defined their terms when referring to the self, and they have not always made clear what functions of the self they are interested in studying, it has not been uncommon for there to be a "clash of cultures" as investigators attempt to communicate with others about issues in this scientific area. However, it is also in such settings that a large part of the impetus for more careful articulation of concepts and specific questions to be addressed has been growing.

MULTIDIMENSIONALITY OF SELF

In those areas of scientific discourse where progress has been made, it has become increasingly clear that a single discrete referent for the concept "self" does not exist. There is also no property of the concept "self" that is present in all instances of "self" processing and absent in all instances of "nonself" or "other" processing. Properties or characteristics by which "self" is sometimes defined (e.g., being a particular kind of mental phenomenon or having a particular kind of function) do not apply to all instances in which the concept is used, and none is unique to it. Concepts like that of "self" have fuzzy boundaries. Such concepts are fuzzy because the content or boundaries of application vary according to context or conditions.

Work that has been undertaken from comparative evolutionary and human developmental perspectives is particularly useful in illustrating this with regard to notions of the self. Importantly, both of these perspectives are grounded in considerations of the biology and adaptive functioning of organisms. Across the animal kingdom and in human ontogeny, self-representational processes, which are necessarily operationalized in the nervous system, clearly vary in their degree of complexity and in the self-representational phenomena that they support. Yet it is also the case that, from both of these perspectives, "self" may be regarded as the set of mechanisms or means by which individuals (and, more specifically, their nervous systems) organize their perceptions, plans, and decisions, allowing them to act as a coherent whole rather than as a group of independently operating systems with competing interests. It is, in fact, when such coherence significantly breaks down that self-pathology is thought to exist, and affected individuals (if they are capable of sufficient insight) may actually report that aspects of their experience or behavior seem "foreign" to them or unmanageable by them.

Comparative Evolutionary Perspective

Scientists who study various animal species are among those that have been interested in the question of how organisms represent themselves to

themselves. Because such scientists consider the conservation or loss of certain bodily structures and capacities across species (in this case, conservation or loss of organisms' capacities to recognize aspects of themselves as being or belonging to themselves), they are often able to offer physiological (proximate) and evolutionary (ultimate) explanations for "why" such capacities exist.

In contrast to organisms of greater complexity, simple vertebrates are observed to be very limited in their self-representational capacities. In these simple animals, such capacities are confined to mechanisms that coordinate visceral and other internal signals with perceptions of the outside world that allow them to distinguish the boundaries of themselves or the effects of their own actions from other phenomena in the environment and generate responses that enable their survival (e.g., an appendage withdrawal when a noxious stimulus is perceived to have impinged upon it). At this simple level, coordination of responses is automatic and implicit. There is no conscious awareness of a "self" who is the subject of experience or the agent of the action.

The centralization of such coordinated, even if limited, processing is the neurobiological foundation upon which higher levels of self-representation rest. One widely recognized view on the interdependent relationships between this implicitly coordinated reflexive processing that subserves basic survival functions of the individual organism and the computationally more sophisticated reflective processing that is associated with self-awareness has been provided by Antonio Damasio. He has referred to the most basic level of inner coordination and regulation which in humans is grounded in the brainstem–hypothalamic axis as the "protoself." Conscious awareness that in humans includes conscious awareness of a uniquely situated subjectivity and sense of ownership, however, requires a nervous system sufficiently evolved and complex that the organism can hold in mind the image of a protoself's moving through and interacting with the world. These more complex aspects of self that also engender greater flexibility and capacity for decision-making (representation and choice among a range of action alternatives) require greater computational resources. For example, increased accuracy in planning and execution of movement in space–time is believed to have been achieved by the development of cortical (over and above subcortical) models of the body in relation to its environment, so that animals with many limbs are capable of moving in very specific ways and at very specific times to meet current demands.

Human Developmental Perspective

Researchers in child and adolescent development also operate on the basis of regarding human beings not as unitary but as integrated multidimensional beings whose capacities emerge and mature over time.

Those researchers explicitly concerned with aspects of self-representation view typical human ontogeny as being associated with the development of a complex set of increasingly differentiated and subsequently integrated self-representational capacities. Some developmental theorists have invoked William James for establishing a useful framework for considering relationships among these various capacities.

James made a distinction between two separable but intimately interrelated aspects of the self: self as *subject* (the I-self) and self as *object* (the Me-self). He also identified subcomponents within these. Components of his I-self included self-agency (the sense of ownership of one's thoughts and actions), self-awareness (an appreciation of one's internal states, needs, thoughts, and emotions), self-continuity (the sense that one remains the same person over time), and self-coherence (a sense that one is a single, coherent, bounded entity). Components of his Me-self included the "material me," the "social me," and the "spiritual me," which were conceptualized as aggregates of various kinds of knowledge about the self that one acquired by adopting an observer's perspective, the details of which also had the potential for coming into conflict with each other.

While details of the Jamesian framework have been challenged, his general distinction between the sense of the I-self and that of the Me-self has remained a viable heuristic for much subsequent research in this area. Until recently, more empirical attention had been devoted to the latter, the self as an object of one's knowledge and evaluation, because of the relative ease with which it and its derivatives (e.g., individuals' "self-concepts" and "self-esteem") are able to be queried and articulated by investigators and study participants. The more implicit and thereby elusive cognitive processes that define the individual as a subject have only recently gained increasing emphasis in theories and studies of self-development. This has happened largely because of advances that have been made in experimental paradigms that are better able to characterize individual differences in cognitive abilities as well as the development of new technologies, such as neuroimaging, which are able to tap some of the previously hidden workings of brain processes when individuals think and act.

There is evidence that there is a significant developmental interdependence between these two general domains of self-representation. Specifically, the structure and content of *what* individuals say they think about themselves at any developmental stage depends on the acquisition of specific cognitive abilities that influence *how* individuals are able to know things about themselves at these stages in development.

For example, young children between the ages of 2 and 4 are able to construct separate attributes of themselves that may be physical (e.g., "I have blonde hair"), social (e.g., "I have three brothers"), or psychological (e.g., "I am happy"). However, they are unable to coordinate two such constructs into a coherent self-portrait, in part because of working memory

limitations at that age that prevent them from holding several features in mind simultaneously. They are also unable to acknowledge the possibility that they (or others) can possess opposing attributes (e.g., good and bad) or different emotions of the same valence, such as mad and sad, or opposite valence, such as happy and sad, at the same time. It is not until middle childhood, between the ages of 7 and 9, that children acquire the ability to integrate positive and negative concepts about the self and become less likely to engage in the kind of all-or-none thinking about self that they display earlier.

Similarly, while even infants are capable of perceiving that other individuals (particularly their significant socializing agents) have reactions toward them, it is not until early to middle childhood that children acquire a cognitive appreciation for the perspective of others and recognize that others have a particular viewpoint and are actively evaluating them and their behavior. As this perspective-taking ability develops, it begins to function as a guide for children so that they are able to identify more with what others expect of them and are increasingly capable of regulating their own behavior. It is not until adolescence, however, that perspective-taking goes beyond recognition of others' expectations for the self and extends to sophisticated abilities for evaluating the self independently and abstractly. It is during this developmental phase that dramatic increases in introspection are observed, which in some individuals appear to be associated with enhanced vulnerabilities to depression or social anxiety.

Despite the general utility of the Jamesian framework, however, there is no clear consensus on a standard taxonomy of self-representational phenomena at this time. This is particularly apparent in research on adult populations, where debates about self and its conceptualizations are, at times, still fairly contentious. Significant coordination between the comparative evolutionary, human developmental, and traditional (adult human) investigative approaches with the elaboration of a broadly applicable theoretical framework for research on "self" has yet to take place, though work undertaken through the relatively new lenses of dynamical systems theory and theoretical-computational neuroscience may offer some possibilities for promoting such a framework.

CLINICAL PERSPECTIVE

Much of the earliest work encouraging scientific investigation of aspects of self-representation actually arose in the clinical context of treatment of adolescent and adult patients. Observation of patients with psychiatric disorders, neurological syndromes, and known structural brain lesions have all contributed to illustrating many of the phenomena that will need to be explained by any scientific formulation of neural substrates of self-representation.

Schizophrenia and Other Psychiatric Disorders

In psychiatry, schizophrenia may be viewed as a disorder of the self. This is presumably because symptoms observed in patients with schizophrenia, such as auditory hallucinations ("voices" that likely represent patients' own inner speech), delusions (that commonly involve patients' own identities and powers), thought withdrawal and insertion (delusions regarding ownership and control of patients' own thought processes), and negative symptoms (that include deficits in emotional responsiveness, spontaneous speech, and volition) typify dysfunction of some of the traditional subject and object conceptions of the self already noted.

In the very earliest days of psychiatry, Emil Kraepelin claimed that disunity in schizophrenic patients' consciousness ("an orchestra without a conductor") was a core feature of the illness and that this disunity was linked to "a peculiar destruction of the psychic personality's inner integrity, whereby emotion and volition in particular are impaired." Such a framing of a core dysfunction in schizophrenia is striking for its resonance with current thinking from dynamical systems' and theoretical neuroscience perspectives that consciousness, including aspects of self-representation, fundamentally depends on the brain's *capacity to "properly" integrate and coordinate* information within and across widely separated brain regions.

Several other psychiatric disorders exhibit features that current thinking (according to their DSM-5 diagnostic criteria) would suggest involve a critical dysfunction of some aspect or aspects of self-representation. Examples include anorexia nervosa (with its "disturbance in the way in which one's body weight or shape is experienced"), various dissociative disorders (with their "losses of continuity in subjective experience, such as fragmentation of identity, depersonalization, and derealization and/or inability to access information or control mental functions that normally are amenable to such access or control"), social anxiety disorder (with its "marked fear or anxiety of social situations in which the individual is exposed to possible scrutiny by others"), and various personality disorders, notably borderline personality disorder, with its pattern of instability of self-images, affects, and interpersonal relationships. This nosological "list" is likely to become more refined and expanded as research progresses and the current understanding of the neural bases of self-representational capacities increases.

Neurological Syndromes

Some patients who are known to have sustained brain damage, usually as a result of a cerebrovascular accident, display dissociations between aspects of their behavior and their reported experiences or stated beliefs about themselves. For example, there are patients with alien limb and alien hand syndromes who by definition exhibit movements of all or a portion of one

of their upper extremities that are dissociated from their stated intentions and are often in conflict with those of their opposite extremity. The patients react with surprise and concern at such movements. These involuntary movements are not spontaneous but are in response to stimuli in patients' peripersonal space. The qualities of these movements are different depending on the location of their brain lesions. Patients with damage to the medial surface of the frontal lobe, for example, typically display intermanual conflict or grasping, groping movements of the hand represented by the lesioned cortex and an inability to release grasped objects. Patients with parieto-occipital lesions by contrast display involuntary movements that involve avoidance of contacts so that the affected hand and limb often levitate away from support surfaces.

In addition to neurological disorders that exhibit dissociations between self-perceived will and purposive action, there are other disorders which involve an apparent lack of self-awareness with regard to some acquired impairment of the body that is clearly apparent to observers. In some cases, this lack of awareness is associated with unshakeable false beliefs about the alteration in their bodies or their impairments. The syndrome of unilateral spatial neglect, for example, is seen in a heterogeneous group of patients who have suffered damage usually to the right parietal lobe (right frontal and left parietal variants are uncommon). Though specific symptoms and their severity vary, all patients exhibit some failure to detect or respond to stimuli on the side opposite to the brain damage. Along with this, patients may have an isolated delusion where they deny ownership of the limb or entire side of the body represented by the lesioned cortex. At the same time, they continue to retain awareness of subjective sensations such as whether they are tired or hungry and continue to have normal autobiographical memory.

Patients who have sustained brain injuries in various locations (again, typically in the right hemisphere) may also remain unaware of and deny their handicaps (anosognosia). This is seen acutely in 20 to 30 percent of cases of hemiplegia/hemiparesis after stroke but can occur with virtually any neurological impairment, including bilateral cortical blindness, prosopagnosia (face blindness), amnesia, aphasia, or apraxia, as well as in cases of multiple impairments where lack of awareness may be for all or limited to just one of the impairments. Such denials of one's own impairments do not appear to be directly related to sensory loss but rather to damage to higher level cognitive processes, including attentional mechanisms, which likely disrupt some capacity for self-monitoring. Nonlocalizability (except to the right hemisphere) also suggests a disturbance of a brain "dynamical" mechanism (e.g., an integrative mechanism, such as one that permits the coordination of sensory information with other processes that support spatial or bodily representations) as its basis. Such attentional and "dynamical" mechanisms

need not be mutually exclusive.

There are also patients who exhibit dissociations between their behavior and reported sense of themselves of another sort. These patients have often been characterized as having disorders of self-regulation. Adult patients with acquired lesions involving the prefrontal cortex, particularly the ventromedial region, often exhibit changes in their ability to control their behavior. While in testing situations they can verbalize norms for acceptable social behavior, in actual practice they behave in ways that conflict with their reportable knowledge. It is as if despite knowing "the rules," they are unable to implement them. Such patients often appear impulsive, socially insensitive, and as if their decision-making, particularly in social contexts, is impaired. They are commonly described by those familiar with them as having undergone a "personality" change. This was seen in the case of Phineas Gage, who sustained a traumatic injury to his frontal lobes (Fig. 1.21–1). When the prefrontal damage is of childhood onset, affected individuals exhibit defective social and moral reasoning as well, suggesting that capacities for acquiring knowledge of complex social conventions and making moral judgments has also been impaired. The resulting syndrome in these early-onset cases then resembles psychopathy.

These various clinical syndromes demonstrate some of the ways in which impairments in capacities for self-representation manifest themselves and result in disunity of functioning that is perceptible to the subject in some cases and only to outside observers in others. Patients' verbal reports and clinical observations of patients' behavior are clearly limited in what they can contribute to the development of a taxonomy of self-representation functions and underlying mechanisms, however. As more experimental work is undertaken and differences in nervous system functioning identified between such individuals and those not afflicted, more will be learned about how to characterize the instantiation of such impaired aspects of self-representation in nervous systems and may permit an even wider variety of pathology involving the disordering of aspects of human subjectivity and behavior to be better understood.

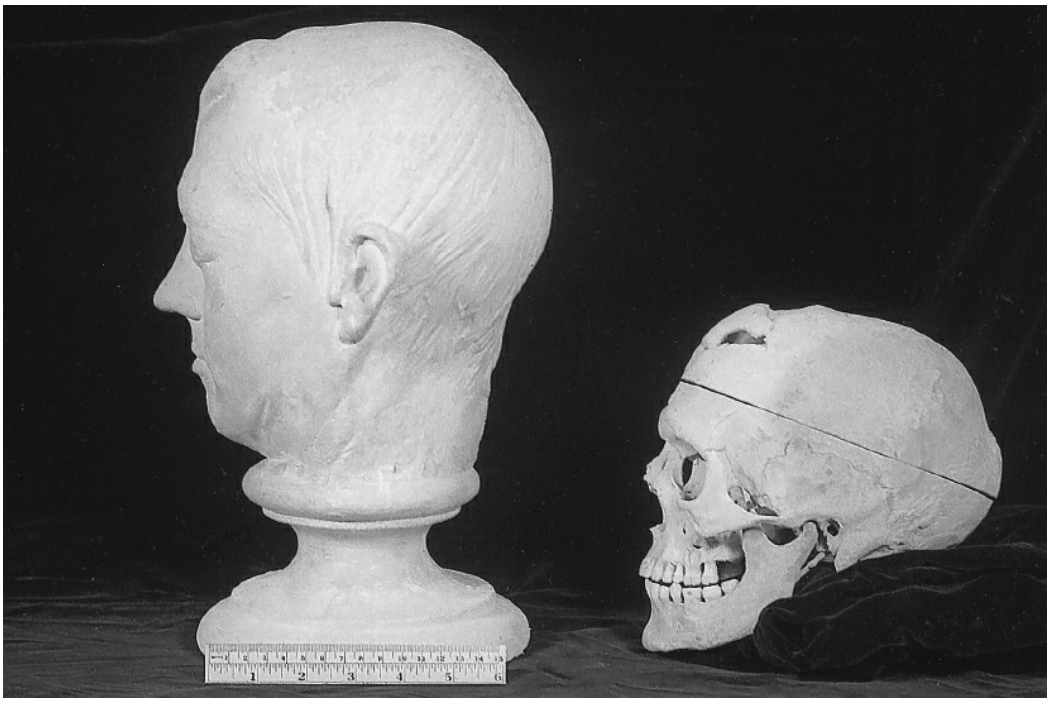

FIGURE 1.21–1. The life mask and skull of Phineas Gage. Note damage to the frontal region. "A famous case illustrating the result of frontal lobe damage involves Phineas Gage, a 25-year-old railroad worker. While he was working with explosives, an accident drove an iron rod through Gage's head. He survived, but both frontal lobes were severely damaged. After the accident, his behavior changed dramatically. The case was written up by J. M. Harlow, M.D., in 1868, as follows: [Gage] is fitful, irreverent, indulging at times in the grossest profanity (which was not previously his custom), manifesting but little deference for his fellows, impatient of restraint or advice when it conflicts his desires… His mind was radically changed, so decidedly that his friends and acquaintances said he was 'no longer Gage.'" (Courtesy of Anthony A. Walsh, Ph.D.)

SCIENTIFIC INVESTIGATIONS

In the past 50 years, experimental work has been done that has supported the increasing shift from examining questions of self-representation using only the approaches employed in philosophy, literary, social, and cultural studies, and clinical settings to those employed in the cognitive and neurosciences as well. Much of this became possible because of developments in scientific methodologies and the emergence of new technologies, such as fMRI and transcranial magnetic stimulation (TMS).

Split-Brain Studies

Some of the most significant data providing conceptual traction on the issue of self-representation in the human nervous system arose from the aforementioned ground-breaking studies of patients who had undergone complete transection of the corpus callosum for the treatment of medically intractable epilepsy that were first undertaken in the 1960s. Functional specializations of the two cerebral hemispheres (e.g., the dominant role of the left hemisphere in supporting most language functions) were the most widely reported findings of the initial studies that were conducted by Roger

Sperry and for which he won a Nobel Prize in 1981. Sperry used several ingenious tasks in order to investigate lateralization of brain function. The tasks were carried out in highly standardized laboratory conditions using specialized equipment. The experiments all involved setting tasks separately to each of the two cerebral hemispheres in each experimental subject. For example, after blindfolding one of the subject's eyes, subjects would be asked to fixate with the seeing eye on a point in the middle of a screen to the left or right of which a stimulus would subsequently be presented for less than one-tenth of a second. This very short duration ensured that there would not be time for subjects to move their eyes and permit the visual information to "spread" across both left and right visual fields and therefore across both sides of the brain. Analogous procedures were used for presenting unseen tactile stimuli to one or the other of subjects' hands or auditory and olfactory stimuli to only one side of their brains. The subjects would then be asked to name the various stimuli. Because language is largely processed in the left hemisphere, it was only when stimuli were presented so that processing occurred in the left hemisphere that subjects were able to name the stimuli (i.e., because of fiber crossing this would occur when stimuli were presented to subjects' right visual fields or right hands). It was also observed that when subjects were presented with a stimulus first in one of the two visual hemifields and then the other, they responded in the second case as if they had never seen the stimulus before. But when the stimulus was represented to the original hemifield, subjects were able to recognize it as the one they had seen before. Other examples of the separability of processing in the two hemispheres were seen in related experiments where different objects would be placed in each of the participants' hands at the same time and then placed for retrieval in a pile of test items. Each hand was observed to search out its own object, rejecting the other item for which the other hand was searching.

Besides revealing specialization of the two hemispheres, these studies also showed that a manipulation (in this case, a disconnection of the two hemispheres) was able to result in a real-time separability of aspects of an individual person's consciousness. Long-term observations of such patients showed that each disconnected hemisphere possesses not only a separate sensorimotor interface with the environment, with its own perceptual, mnestic, and linguistic repertoires, but also its own characteristic likes and dislikes and styles of decision-making. Researchers involved in conducting experiments with such patients reported that the experience of "interacting" with each of the separated hemispheres in individual patients felt akin to interacting with a distinct personality.

It is not the case, however, that patients who have undergone a complete callosotomy appear to have a "split personality." One of the most significant observations since those initial experiments is that there is an

apparent difference in the patients' ability to transfer information between the two hemispheres in testing situations compared to routine social (free-field) situations. Aside from short-term memory deficits and attentional limitations, such individuals may appear quite ordinary, "unified," and purposeful in their everyday behavior. This appears to be made possible by several facts.

In free-field settings, patients are able to use strategies that involve bihemispheric explorations of space, including bimanual explorations and conjugate eye movements, so that both hemispheres receive salient information for working together in particular contexts. There are also pathways at levels below that of the corpus callosum that indirectly connect the two hemispheres via the cerebellum, brainstem, and the thalamus. These may facilitate interhemispheric cooperation for compatible stimuli in the two visual fields and the typically observed unity in motor responses. And it is also the case that much socially observable behavior, particularly involving speech and language, appears to be under the control of the left hemisphere, so its responses are in effect dominant.

Recognizing that the disconnected right hemisphere demonstrates awareness in the absence of an ability to express such awareness verbally has provided additional evidence that language is not necessary for human consciousness. Also, while the normal experience of consciousness is often thought to be inherently unified, it now appears that normal consciousness may be fundamentally "dual" (if not "multiple"). That is, it is grounded in parallel processes in the two hemispheres that tend to interact and accommodate the resources of each other in most settings but whose partially separable workings may sometimes be in evidence, as in the common human experience of subjective conflict.

And, despite the fact that the two cerebral hemispheres in a single split-brain individual clearly exhibit signs of distinctiveness in processing style that occur in parallel, the aforementioned free-field observations suggest that mechanisms exist for "driving" coherence in the nervous system in natural settings. Findings such as these that have led increasing numbers of theoretical neuroscientists and others to begin viewing the brain and its operations as a kind of "complex adaptive system," subject to self-organizing principles and dynamics similar to those observed in other domains.

Distinguishing Self from Nonself

Brains, and the nervous systems of which they are a part, are thus being increasingly viewed as inherently dynamical. Their ongoing activity, the patterns of communication between brain areas residing in such activity, and adaptive "problems" they "solve," likely subserve various forms of self-representation. For example, the fundamental problem of distinguishing self from nonself exists for all mobile organisms. It has been studied in

lower animals as well as humans. One well-studied animal model is that of the electric fish. Studies of this animal have been useful for identifying strategies that a relatively simple sensorimotor system may use for solving the inverse problem of determining which of all the patterns of neural activity in its sensory system represent those arising from environmental events (nonself). All such strategies are made possible by the animal's movement and exploration of its environment, which result in modification of the kind of sensory information that it acquires and help constrain the solutions to this problem.

Some of the modifications that help obtain such a solution arise from nervous systems' determination of different temporal relationships between sensory signals. In addition, prediction signals arising from the motor system and their comparison with sensory signals are thought to play a major role in neural processes that permit separation of self-generated movements from movements caused by another individual (passive movement) or from other activity that is going on in the environment. That is, well-established theoretical frameworks (called feed-forward models) have been developed for computational considerations of processes describing such action execution and motor control by nervous systems. In these frameworks, at least one kind of neural signal that is thought likely to be a component of this preparatory activity for movement execution is one *predicting* the sensory consequences of the movement.

An interesting example of how such processing may operate in healthy humans is illustrated by the phenomenon of tickling. Individuals can usually experience tickling by others but not when trying to tickle themselves. Now consider how the nervous system could distinguish such similar events and yet produce these very different experiences. Sensations that are associated with self-generated movements can be *correctly* predicted on the basis of the timing of the motor command. They are also associated with little or no temporal discrepancy resulting from the comparison between predicted and actual sensory feedback. As the discrepancy from this comparison gets larger, however, so does the probability that this sensation is externally produced. In such a system, it would be possible to "cancel out" sensory effects (e.g., tickling sensations) that are induced by self-motion and thus distinguish them from the sensory feedback caused by effects from the environment.

Sarah-Jayne Blakemore, Daniel Wolpert, and Chris Frith conducted a now classic experiment in which normal subjects underwent self-produced tactile stimulation and robotic stimulation using the same stimulus (a sinusoidal movement having a particular amplitude and frequency of a piece of soft foam across the right palm). In a related experiment, subjects underwent the same kind of stimulation, but it was achieved via a robotic interface that the subjects manipulated with their own left hands and into which different time delays were introduced (between when the subjects

moved their left hand and produced the stimulation on their right palm). Under these controlled conditions, subjects rated the self-produced tactile stimulation as significantly less tickling and pleasant than that produced by the robot. They also reported a progressive increase in the tickling and pleasant sensation as the temporal delay increased from no delay to a 200-ms delay. Such data are consistent with the feed-forward model's hypothesis of perception of self-produced tactile stimulation being able to be attenuated due to precise sensory predictions. A follow-up neuroimaging study of the self versus robotic stimulation comparison supported this by showing that there was relatively attenuated activity in the parietal operculum (secondary somatosensory cortex) and anterior cingulate in the self condition.

In similar experiments as well as those using other designs, individuals with delusions of control have been shown to differ in their responses from individuals without such delusions. Other designs include the "rubber-hand illusion," an established method for manipulating the sense of body(part) ownership, where subjects see feedback of their own hand movement or that of an experimenter's hand making similar movements. In such experiments, individuals with delusions of control have found it more difficult than control subjects to distinguish sensations arising from self-generated actions from those arising from actions that were not self-generated. They also required greater temporal discrepancies between such actions to distinguish between actions arising from themselves and those not arising from themselves. These and other results have suggested that some component of the feed-forward modeling process for making such discriminations is impaired in these individuals.

The Self as Agent

With modern scientific discussions of self moving away from a relatively "static" concept to more dynamical, process-based accounts, there are new considerations of self as agent promoting investigations that target *experiences* of agency (e.g., experiences of mental causation or of being the "source" of intentions and motor commands) or the actual *structure* of agency (the organizational relationships among perception, an intention or goal, planning or selecting a movement consciously or unconsciously, and a motor command that may or may not require subsequent monitoring for reaching its goal). Disorders of the neural mechanisms underlying one or both of these are thought likely to account for disorders of volition—that is, the psychiatric or neurological conditions where individuals' own actions are either denied having been self-generated or are reported as feeling as if they have not been self-generated.

Experiences of agency have representational content and represent individual agents and their actions as being in a certain way. For example, normal individuals do not simply find themselves walking toward a

stairway and then, on the basis of this, acquire a belief that they must be intending to go down the stairs. Rather, the experience is of walking toward the stairway in order to go down stairs. Individuals usually experience a sense of purposiveness with their actions, which may be related to intentions that are conscious or not. When there is no awareness of a prior intention, it is still the case that normal individuals' actions involve experiences of having acted in the service of a goal and thus having implemented an intention that they experience as their own. In normal individuals the sense of intention and of being the "owner" and source of their actions is typically very strong. Normal individuals tend not to doubt that their actions are self-generated. There is now a variety of data on how normal individuals experience their actions as self-generated and the form that this self-generation takes.

Action Awareness

Experiments by Marc Jeannerod and others have addressed fundamental questions regarding the aspects of action that individuals are aware of. In these experiments, subjects were typically asked to perform a task, such as drawing a straight line on a computer screen, while they were unable to see their arm or hand. They would be given false feedback about the trajectories of their movements, so that they would have to make significant deviations from straight movements in order to reach their goal. However, when asked about their experiences, subjects' verbal reports indicated that they had been unaware that they were making deviant movements. Other experiments involved discrepancies in the timing between expected and actual movements with similar results. Investigators also determined thresholds for the degrees of discrepancy that could exist before individuals reported awareness of a discrepancy when they were asked. When they were not asked, subjects appeared not to notice even when the degrees of discrepancy were higher than these thresholds. Thus it appears that there are aspects of their own movements about which even normal individuals are unaware. In addition, individuals may be more aware of movements they *intend* to make than those they actually make in particular settings.

Sense of Agency and Voluntary ("Willful") Action

There are deep normal human intuitions as well as folk-psychological beliefs that intentions and thoughts *cause* human action. As already noted, however, scientific accounts of intention and its relationship to action reject dualistic accounts, where the "mind" is thought to cause changes in the body, resulting in action. In the scientific view, both conscious experience ("thought") and action are *results* of brain activity. There must, therefore, be some organized relationships in neural processes that

instantiate this process. That is, there must be a structure of agency. Despite the inherent challenges, some investigators have specifically targeted this structure, including considerations of how components of this structure may interface with mechanisms for the human experience of agency.

A classic study by Benjamin Libet and colleagues has been commonly cited as evidence that consciousness (specifically, individuals' experience of the intention to move) is not the initial cause of behavior and that behavior occurs instead as the result of a chain of events initiated by unconscious brain events. Libet monitored subjects' electroencephalogram (EEG) and muscle activity while subjects performed simple finger movements and asked them when they were aware of the "urge" to move. Though subjects did consistently anticipate the starting time of the movement by over 50 ms, there were signs of preparatory brain activity for this movement that preceded these "awareness judgments" by several hundred milliseconds. Many authors have had concerns about some of the technical details of this experiment, however, and have consequently also disagreed that the results of this study support the idea that "free will" does not exist.

Other investigators have undertaken experiments based on the Libet paradigm to pursue this issue further. These experiments are notable in two respects. First, the basic result that the generation of action (from the beginning of preparatory brain activity to the onset of muscular activity) extends over time and has a phase that is unavailable to awareness has been replicated. However, the details of the relationships among brain activity, conscious intention, and action have also been determined to be more complex than initially surmised. What is now recognized is that action generation is not a "simple" process. That is, it is not simply an underlying neural process of preparation to move and its execution that may (or may not) be experienced. Rather it now appears that the brain processes of preparation, execution of movement, and conscious experience of these phenomena each have several components.

The actual structure of agency itself is thought to depend upon particular organizational relationships among several variables, including intention or goal, action selection, a motor command, and its outcome. Models have been proposed for this structure with a particularly notable one being a "comparator model" formulated by Wolpert and colleagues in the 1990s. This model and its derivatives were originally articulated as models of motor control. An important question, however, is whether or not the sense of agency associated with or emerging from this structure depends solely upon comparing the degree of matching between *predicted* and actually experienced consequences of movement (so that sense of agency is strong when there is a close match between predicted and actual sensory consequences of an action and is reduced or absent when the match is poor). Such an account would necessarily imply that sense of

agency occurs *after* an action has been performed and information about its sensory consequences becomes available. While the "comparator model" and its retrospective account of agency have received significant behavioral and neuroimaging support (recall the "tickling" and "rubber-hand illusion" experiments, as examples), other experimental data challenge the notion that sense of agency is *fully* explained by such an account.

Patrick Haggard and his colleagues have undertaken several experiments investigating these issues and their data have yielded results suggesting that sense of agency may also be generated *prospectively* via the process by which intentions (here defined as prospective signals involved in action preparation) are transformed into specific actions to reach desired effects. In a variant of the Libet paradigm, for example, they stimulated the motor cortex using TMS and observed that, relative to the nonstimulated condition, the perceived time of movement was only slightly delayed (~75 ms) while signs of initiation of the movement (muscular activity) were more significantly delayed (~200 ms). Such data support the idea that awareness of initiating a movement is not derived from any sensory signals arising from the moving limb (because such signals are not available until after the limb has started moving). Rather awareness appears to be linked, at least in part, to some signal that precedes the movement. Additional behavioral and neuroimaging experiments by this group have shown that sense of agency is also able to be modulated by subliminal priming maneuvers that affect the fluency of action selection processes (i.e., the sense that something may or may not be "wrong" when selecting or following a course of action). Thus, at least under certain circumstances, individuals' sense of agency may be more an effect of such prospective signals than an effect of the consequences of the movements themselves.

As previously noted, prediction signals are crucial in the nervous systems of organisms that are capable of movement. For human beings, who are capable of both voluntary and involuntary movement, being able to compare signals that predict the likely effects of commands arising from the motor system with sensory information acquired during movement is particularly important for adjusting movements easily and checking whether movements are being completed as planned. That is, such a system includes both feed-forward (prediction) and feedback (signal comparison) components. The work of Haggard and colleagues suggests, however, that even earlier prospective signals involving action *selection* (and action programming) may participate along with these prediction signals and signal comparisons in acting as cues that have a role in the experience of agency. Additional research will be required to further clarify the details of these processes and the circumstances under which each may have a particularly important role.

Self versus Other: Perspective-Taking and Mental State Attribution

In human beings, there is another set of functions by which an individual's experience of self may be uniquely characterized. These involve the manipulation of perspective-taking and mental state attribution.

First- versus third-person perspectives are quite different from each other. First-person perspective refers to the centralization of the subjective multimodal experiential space around one's own body, while third-person perspective refers to a variety of processes that all include ascribing mental states to others.

Individuals may adopt first- versus third-person perspectives with regard to movement through or location in space. Spatially, an individual is situated in what is called an egocentric reference frame, where locations of objects are represented relative to that individual's body. However, when individuals try to understand what someone else sees, for example, they need to be aware that others' percepts are actually different from their own *and* they need to be able to perform the mental computation that permits reorienting the egocentric reference frame and centers it on that other person's body. Electrophysiological studies in primates have yielded much data on the role of the posterior parietal cortex in multisensory integration and spatial coordinate transformations for this purpose. (These processes are also required for converting sensory input centered on an organism's body into motor output that is directed at objects located in what is referred to as an allocentric reference frame—that reference frame of object—object relations whose internal coordinates are independent of an organism's body.) In human brain imaging studies, the posterior parietal cortex usually on the right side has routinely been activated when people have been asked to execute tasks that involve adopting a third-person perspective.

Imaging studies have also revealed sites in the prefrontal cortex that are thought to be involved in the other major component of the perspective-taking process—the recognition of others' mental states. Many functional brain imaging studies have shown that the medial prefrontal cortex (MPFC) is significantly engaged when people are asked to perform tasks that require representation of mental states (Fig. 1.21–2). The mental states may be those of the subjects themselves or of others. This imaging finding affords additional evidence that a property by which "self" is often defined is not unique to it. That is, at least with regard to the brain regions involved, there is overlap in generating considerations of one's own and others' mental states. For example, the MPFC has shown activation when subjects have been asked to perform various acts of self-evaluation. Experimental tasks have included asking subjects to make trait judgments about themselves ("does this word [e.g., friendly, unhappy, industrious] apply to you"), and make preference judgments ("which do you prefer—item 1 or item 2").

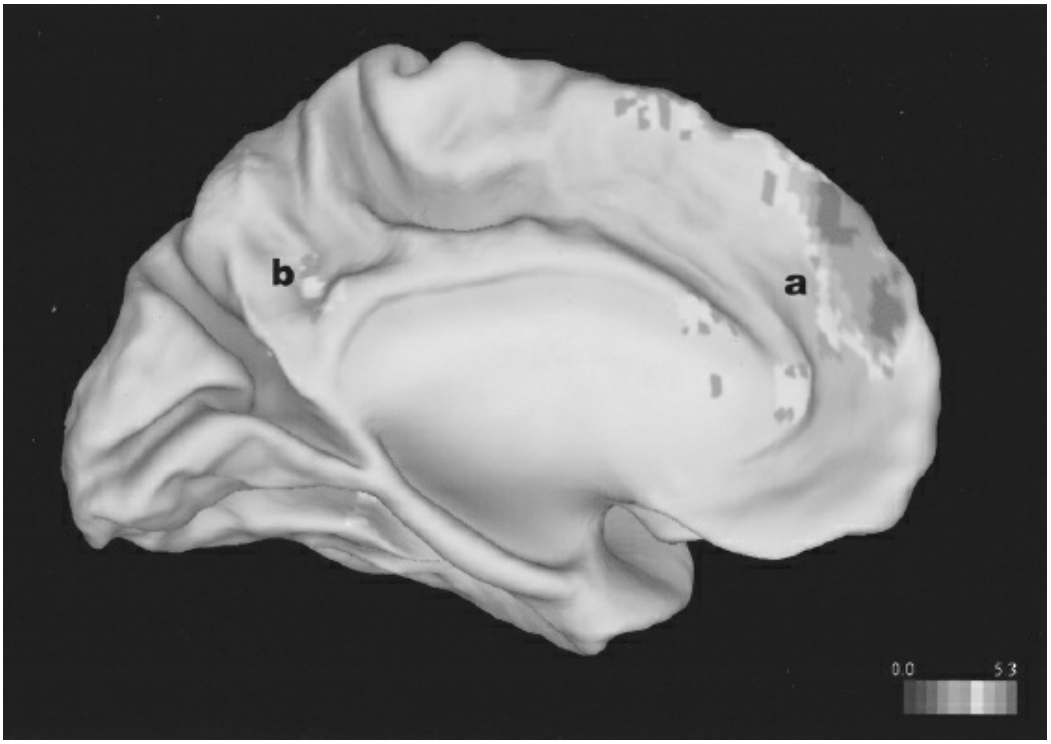

FIGURE 1.21–2. Statistical activation map of the contrast between participants' judgments of their own feelings about variably evocative scenes and their judgments of whether these same scenes appear to be indoors or outdoors (*n* = 24 normal participants). Relatively greater activation is commonly seen in the dorsal medial prefrontal/paracingulate (**a**) and posterior cingulate/retrosplenial (**b**) regions in such experimental tasks, where attention to subjective states is of primary interest. Other experimental data suggest that the medial prefrontal region may be more significant for the instrumental aspects of self-reflection, while the posterior medial cortex may be more significant for experiential (including memory-related) aspects of self-reflection. (See also Gusnard DA, Akbudak E, Shulman GL, Raichle ME. Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proc Natl Acad Sci U S A*. 2001;98:4259.)

The dorsal MPFC has also shown activation when subjects have been asked to assess what others' think or feel, as in so-called theory-of-mind tasks. Theory-of-mind refers to the cognitive ability to attribute mental states, such as beliefs, desires, and intentions, to oneself and others and to understand that others have mental states that are different from one's own. The most commonly used task to assess theory-of-mind abilities is the false-belief task. In one version of this task, subjects (usually children or patients, such as those with autism) are shown a story involving two characters. The two characters might be two girls, for example, playing with a doll. The girls put the doll away in a box and then one girl leaves. The other girl removes the doll from the box, plays with it again, and then puts it in a different box. The girl who had left returns and subjects are asked where they think she will look for the doll. Subjects pass the test if they realize that the girl will look in the first box where the doll was placed before she left the room. They fail if they believe she will look in the second box, where they know the doll is hidden. This is because the girl character

could not know this, since she had not seen it placed there.

The capacity to be aware that others have mental states that are different from one's own is critical for individuals' ability to participate successfully in human social interaction. Much of human cooperative behavior depends on the ability and motivation to exchange information, the recognition that others lack important knowledge, have beliefs that are false, or have intentions that complement or conflict with one's own.

Human social interaction is also supported by individuals' abilities to understand the emotional and affective states of selves and others. Abilities to reflect on one's own feelings as well as to be empathic, feel sympathy, and have compassion for others contribute to moral decision-making and help promote social cohesiveness. Cognitive studies and functional imaging experiments targeting empathy have also been done. Experimental work on this issue has shown that the same network of brain regions are often engaged when subjects experience something (e.g., pain) *and* when they observe someone else appearing to be in the same experiential state (e.g., when shown pictures of others in pain). Such data again support the view that certain brain networks *share* in representing self and others' states and, in this case, account for humans being able to achieve empathic understanding. Brain networks that have been identified in such contexts include but are not limited to portions of prefrontal cortex, medial parietal cortex, portions of the ACC, and insula.

This shared-representation explanation is also consistent with, though not completely explained by, the tenets of "simulation theory," which has been a proposed framework for describing how individuals come to be able to understand others' behavior and subjective states. Identical circuitry for "sharing" or "simulating" others' experiences is unlikely to be the sole mechanism underlying abilities to appreciate and recognize others' subjective states as such, however. Being empathic also requires maintaining a separateness of oneself from another, so that one can act. In settings such as psychotherapy, for example, both must occur for progress to be made. Thus mechanisms underlying personal agency likely need to be integrated with those that underlie this kind of equivalence between self and other to accomplish this in practice.

General Neuroscientific Theories of Self-Representation and Self-Processing

It is critical for the integrity and survival of all mobile organisms that they be able to distinguish "self" from "nonself" and act on the basis of such a distinction. Evolution of nervous systems from the simplest vertebrates to humans has resulted in marked increases in complexity and diversity of capacities for achieving various forms of such "self-recognition." Damasio has emphasized this point, observing that the limitations of the "protoself" grounded in the brainstem–hypothalamic axis of humans have been surmounted by the development of a cortex with its latest addition, the

neocortex, affording not only a greater range and flexibility in action selection and execution, but also a capacity for consciousness and reflective processing supporting various aspects of self-awareness. Current evidence suggests that the relatively primitive capacity to recognize features of one's physical self, such as features of one's face or body as assessed by the behavioral mirror self-recognition test, for example, is limited to a small set of primate species, including humans, while more cognitively abstract forms of self-awareness, such as introspection and mental state attribution, are unique to humans.

Given this exceedingly wide range of difference in degree of complexity (as illustrated here for "self-recognition," for example) and the diversity of functions that self-representational capacities serve across organisms and species, it is not surprising that attempts at any unifying scientific theories targeting self-representation have been extremely limited to date. Two examples of preliminary efforts in this direction, however, are mentioned here.

One of these, by Legrand and Ruby (2009), has attempted to gain some traction on this problem of general self-representation by carrying out a meta-analysis of a large number of self-related neuroimaging studies performed (on adults) over approximately the past decade. Though their focus remained limited to humans, this kind of effort is still an important one, given the lack of conceptual clarity that still remains at the level of the human species itself (let alone across species). The primary aim of their analysis was to determine whether any cerebral correlates specific to the "self" condition could be identified across the neuroimaging studies, which involved a variety of tasks and stimuli that employed diverse self versus nonself contrasts. They observed that, while the results of individual studies and assessments of small sets of studies were often argued to identify a self-specific brain region or network of brain regions, the results of their more inclusive analysis did not support such a conclusion, but rather pointed to a lack of brain regions specifically dedicated to self-condition processing (which is not a completely novel point and is, in fact, one that some others have maintained, though typically on the basis of little data or to offering any other constructive suggestion addressing the main question of how it is that sense of self is engendered in brains). On the other hand, several brain regions, including the medial prefrontal, precuneus, and temporoparietal regions, appeared regularly across such studies, though they were identified not only in the self-condition but in conditions such as interpreting others' mental states as well. They also observed that these same brain regions were often seen to be activated in studies where subjects were asked to perform memory recall and reasoning tasks. Legrand and Ruby then proceeded to propose a framework for describing self-processing as a product of more general mechanisms that, taken separately, are not self-processing specific, but in their aggregate are

self-defining. More specifically, they proposed that these mechanisms together operationalize a subjective perspective and permit relating represented objects to a representing subject. Their proposal aligns in several respects with other investigators' emphases on aspects of agency, underscoring the role of functional processes that effectively integrate an individual's motor activity with its predicted sensory consequences in the course of acting in the world.

Another attempt at a broad theoretical framework has used a different tack with potential applicability beyond the human species, approaching the problem of self-recognition, specifically. Apps and Tsakiris have offered an account that focuses on the relatively low-level problem of how the brain processes incoming visual, auditory, somatosensory, or interoceptive input as "self" ("me") and distinguishes it from "nonself" ("not-me"), allowing recognition of an individual's physical self (face, body and movement, and voice) as his or her own. They describe how the recently developed "free-energy principle" (itself an attempt at a unifying theory of the brain viewed as a self-organizing biological system developed by Karl Friston and colleagues) can be used to explain many of the current findings in self-recognition research and also relate to higher cognitive processes of self-awareness. Though the details of their theoretical account (particularly the Bayesian principles and mathematical descriptions upon which it is based) lie outside the purview of the present chapter, the basic notion is that there are mechanisms by which the brain optimizes predictions about the sensory consequences of events occurring in the environment (including the results of subjects' own actions) in order to minimize the amount of "free-energy" (i.e., "surprise" or deviation from expectation) and be in states of high predictability (recall the previously described tickling experiment by Blakemore and colleagues). Apps and Tsakiris also argued that several of the brain regions identified by others for so-called self-processing could be interpreted as participating in the mechanisms of minimizing prediction errors and integrating multisensory information for recognizing the "self" that are critical to their account. Like Legrand and Ruby, they propose that there are general mechanisms underlying the defining of a particular sense of self that are not self-processing specific but are grounded in the operations of brain regions whose dynamic and coordinated activity are what are, in fact, responsible for such self-definition.

Though the proposals of both groups afford some explanation for a range of extant self-related empirical data, additional work will be required to see how much each account will be able to be validated by future research that is guided by its specific predictions. It is also interesting to note, however, that as investigators like these focus on the brain and its activity in discussing "self," discussion trends away from notions of self-representation per se to more dynamic notions of self-*processing* (i.e., *how*

such representations are constructed).

FUTURE DIRECTIONS

With the increasing emphasis on considerations of self-processing, the question arises whether current technologies are capable of providing the data that may ultimately prove necessary for describing the operations involved in such processing. Progress in theory is often stimulated by technological progress as well as by the use of available technology in new and different ways. Diffusion MRI and fMRI methods for assessing anatomical and functional connectivity now exist but they are continuing to undergo development, particularly in terms of the speed and accuracy with which the imaging data may be acquired and processed. Various "connectome projects" are currently underway, including the "Human Connectome Project," sponsored by the United States' National Institutes of Health. This several year project is using current MRI technology to map human brain connectivity in a large sample of healthy adults, offering the possibility of not only greater understanding of typical human brain networks' structures, but a sense of the range of individual variability in their structures as well. Researchers and computational neuroscientists are also beginning to use this kind of data to assess patterns of connectivity disturbance and network breakdown in various developmental, degenerative, and mental disorders, which is providing preliminary markers of what constitutes robust versus disordered patterns of brain networks along with their implications for (information) processing.

These brain network and other brain dynamical approaches are highly sophisticated and currently relatively labor-intensive techniques, whose ramifications for furthering the understanding of self-processing are in their infancy. So, while understanding "the self" in all the various and nuanced ways in which that term is used will clearly never be achieved, continued pursuit of the scientific agenda on self-processing within the limits and with some of the caveats outlined in this chapter holds promise for continued progress in enriching the understanding of mechanisms underlying many psychiatric disorders.

REFERENCES

*Apps MA, Tsakiris M. The free-energy self: A predictive coding account of self-recognition. *Neurosci Biobehav Rev*. 2014;41:85–97.

Blakemore SJ, Frith C. Self-awareness and action. *Curr Opin Neurobiol*. 2003;13:219–224.

Blanke O, Arzy S, Landis T. Illusory perceptions of the human body and self. In: Goldenberg G, Miller B. *Handbook of Neurology. Neuropsychology and Behavioral Neurology*. Vol. 88. Amsterdam: Elsevier; 2008.

Brass M, Haggard P. The what, when, whether model of intentional action. *Neuroscientist*. 2008;14:319–325.

*Chambon V, Sidarus N, Haggard P. From action intentions to action effects: How does the sense of agency come about? *Front Hum Neurosci*. 2014;8:320.

*Churchland PS. Self-representation in nervous systems. *Science*. 2002;296:308–310.

Damasio AR. *The Feeling of What Happens*. New York: Harcourt; 1999.

Decety J, Jackson PL. A social-neuroscience perspective on empathy. *Curr Dir Psychol Sci*. 2006;15:54.

Denton DA. *The Primordial Emotions: The Dawning of Consciousness*. New York: Oxford University Press; 2006.

Freeman WJ. *How Brains Make Up Their Minds*. New York: Columbia University Press; 2000.

Gallagher S. *The Oxford Handbook of the Self*. New York: Oxford University Press; 2011.

*Harter S, Bukowski WM. *The Construction of the Self: Developmental and Sociocultural Foundations*. 2nd ed. New York: Guilford Press; 2012.

Jeannerod M. The sense of agency and its disturbances in schizophrenia: a reappraisal. *Exp Brain Res*. 2009;192:527–532.

Leary MR. *The Curse of the Self: Self-Awareness, Egotism, and the Quality of Human Life*. New York: Oxford University Press; 2004.

LeDoux J, Debiec J, Moss H. *The Self: From Soul to Brain. Annals of the New York Academy of Sciences*. Vol. 1001. New York: New York Academy of Sciences; 2003.

*Legrand D, Ruby P. What is self-specific? Theoretical investigation and critical review of neuroimaging results. *Psychol Rev*. 2009;116:252–282.

Llinás RR. *I of the Vortex: From Neurons to Self*. Cambridge, MA: MIT Press; 2001.

Montague PR. Free will. *Curr Biol*. 2008;18:R584–R585.

Rabinovich MI, Friston KJ, Varona P. *Principles of Brain Dynamics: Global State Interactions*. Cambridge, MA: MIT Press; 2012.

Sorabji R. *Self Ancient and Modern Insights about Individuality, Life, and Death*. Chicago, IL: University of Chicago Press; 2006.

Sperry RW. Hemisphere deconnection and unity in conscious awareness. *Am Psychol*. 1968;23:723–733.

Sporns O. *Networks of the Brain*. Cambridge, MA: MIT Press; 2011.

Vogeley K, May M, Ritzl A, Falkai P, Zilles K. Neural correlates of first-person perspective as one constituent of human self-consciousness. *J Cogn Neurosci*. 2004;16:817–827.

Wegner D. Who is the controller of controlled processes? In: Hassin RR, Uleman JS, Bargh JA, eds. *The New Unconscious*. New York: Oxford University Press; 2005.

▲ 1.22 Basic Science of Sleep

RUTH M. BENCA, M.D., PH.D., CHIARA CIRELLI, M.D., PH.D., AND GIULIO TONONI, M.D., PH.D.

Sleep is a fundamental behavior of all animal species, although its specific functions are not yet fully understood. Sleep occupies approximately one-third of the human lifespan, and loss of sleep can lead to cognitive, emotional, and physical impairment. Systems involved in regulation of

sleep and wakefulness appear to overlap or interact with systems involved in the regulation of emotion and other behaviors. It is therefore not surprising that sleep abnormalities are commonly found in patients with psychiatric disorders, part of the diagnostic criteria for several disorders, and have predictive value for the future development of psychiatric disorders. Furthermore, a majority of commonly used psychiatric medications have effects on sleep.

NORMAL HUMAN SLEEP

Definition of Sleep

From a behavioral standpoint, sleep is a state of decreased awareness of environmental stimuli that is distinguished from states such as coma or hibernation by its relatively rapid reversibility. Sleeping individuals move little and tend to adopt stereotypical postures. Although sleep is characterized by a relative unconsciousness of the external world and a general lack of memory of the state, unlike in comatose states, people generally recognize when they feel sleepy and are aware that they have been asleep at the termination of an episode.

For clinical and research purposes, sleep is generally characterized by combining behavioral observation with electrophysiological recording. Humans, like most other mammals, express two types of sleep: rapid eye movement (REM) and nonrapid eye movement (NREM) sleep. These states have distinctive neurophysiological and psychophysiological characteristics. REM sleep derives its name from the frequent bursts of eye movement activity that occur. It is also referred to as *paradoxical sleep* because the electroencephalogram (EEG) during REM sleep is similar to that of waking. In infants, the equivalent of REM sleep is called *active sleep* because of prominent phasic muscle twitches. NREM sleep, or *orthodox sleep,* is characterized by decreased activation of the EEG; in infants it is called *quiet sleep* because of the relative lack of motor activity.

Stages of Sleep

Within REM and NREM sleep, there are further classifications called *stages* (Table 1.22–1 and Fig. 1.22–1). For clinical and research applications, sleep is typically scored in epochs of 30 seconds with stages of sleep defined by the visual scoring of three parameters: EEG, electrooculogram (EOG), and electromyogram (EMG) recorded beneath the chin. Most of the criteria defined by Rechtschaffen and Kales in 1968 are still accepted in clinical practice and for research around the world, although new rules that modify the older criteria were adopted in 2007 by the American Academy of Sleep Medicine (AASM) in *The AASM Manual for the Scoring of Sleep and Associated Events.* During wakefulness, the EEG shows a low-voltage fast activity or activated pattern. The EMG has a

high tonic activity with additional transient muscle activity related to voluntary movements. Voluntary eye movements and eye blinks can also be observed during wakefulness. When the eyes are closed in preparation for sleep, alpha activity (8 to 13 Hz) becomes prominent, particularly in the occipital regions.

NREM sleep, which usually precedes REM sleep, is subdivided into three (N1 to N3) stages (Table 1.22–1 and Fig. 1.22–1). Sleep usually begins with a transitional state, stage N1 (formerly stage 1 sleep), characterized by the loss of alpha activity and the appearance of a low-voltage, mixed-frequency EEG pattern with prominent theta activity (4 to 7 Hz), and occasional vertex sharp waves (V waves) over the central regions. Eye movements become slow and rolling, and skeletal muscle tone relaxes. Subjectively, stage N1 may not be perceived as sleep although there is a decreased awareness of sensory stimuli, particularly visual, and mental activity becomes more dreamlike. Motor activity may persist for a number of seconds during stage N1. Occasionally individuals experience sudden muscle contractions, sometimes accompanied by a sense of falling and/or dreamlike imagery; these hypnic jerks or sleep starts are generally benign and may be exacerbated by sleep deprivation. Typically, sleep-deprived individuals will enter periods of "microsleep" that consist of brief (5 to 10 seconds) bouts of stage N1 sleep; these episodes are unavoidable in sleepy individuals and can have serious consequences in situations that demand constant attention, such as driving a motor vehicle.

Table 1.22–1.
Stages of Sleep—Electrophysiological Criteria

	EEG	EOG	EMG
Wakefulness			
Stage W	Low-voltage, mixed frequency Alpha (8–13 Hz) with eyes closed, vertex sharp waves	Eye movements and eye blinks	High tonic activity and voluntary movements
Nonrapid eye movement			
Stage N1	Low-voltage, mixed frequency Theta (4–7 Hz) and vertex sharp waves may be present	Slow eye movements	Tonic activity slightly decreased from wakefulness
Stage N2	Low-voltage, mixed frequency background with sleep spindles (12–14 Hz bursts) and K-complexes (negative sharp wave followed by positive slow wave)	None	Low tonic activity
Stage N3	High-amplitude (≥75 μV) slow waves (≤2 Hz) occupying at least 20% of epoch	None	Low tonic activity
Rapid eye	Low-voltage, mixed frequency	Rapid eye	Tonic atonia

movement	Saw-tooth waves, theta activity, slow alpha activity may be present	movements	with phasic twitches

Criteria from Iber C, Ancoli-Israel S, Chesson AL, Quan SF. *The AASM Scoring Manual for the Scoring of Sleep and Associated Events.* Westchester, IL: American Academy of Sleep; 2007.

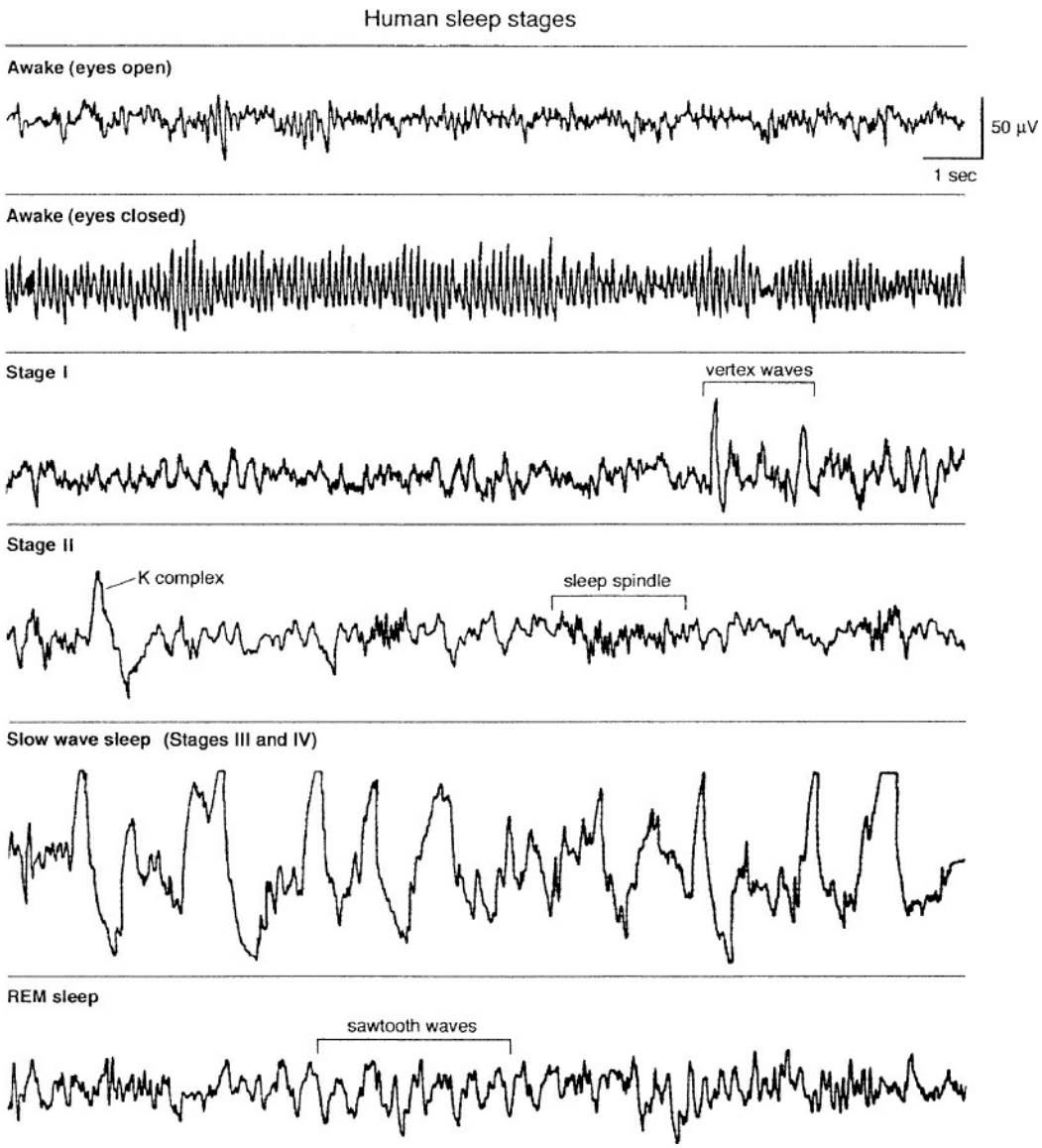

FIGURE 1.22–1. Electroencephalogram patterns for stages of sleep and wakefulness. REM, rapid eye movement. (From Butkov N. Atlas of Clinical Polysomnography. Medford, Oregon: Synapse Media; 2010, with permission.)

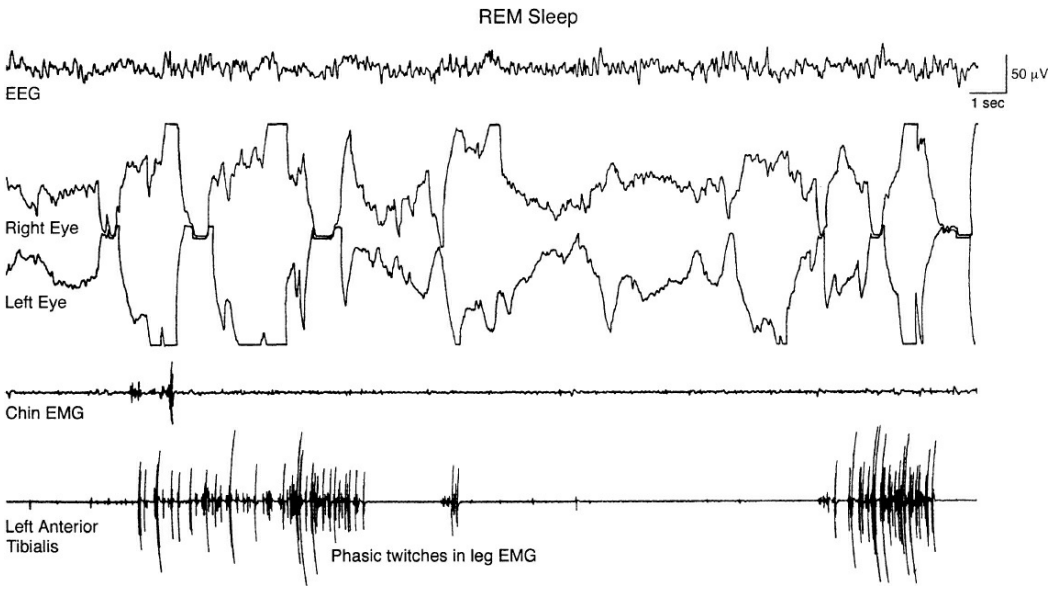

FIGURE 1.22–2. Example of human rapid eye movement (REM) sleep. EEG, electroencephalogram; EMG, electromyogram. (From Butkov N. Atlas of Clinical Polysomnography. Medford, Oregon: Synapse Media; 2010, with permission.)

After a few minutes of stage N1, sleep usually progresses to stage N2 (formerly stage 2), which is heralded by the appearance of sleep spindles (11 to 16 Hz, lasting ≥0.5 seconds) and K-complexes (high-amplitude, negative sharp waves followed by positive slow waves) in the EEG. Stage N2 and subsequent stages of NREM and REM sleep are all subjectively perceived as sleep. Particularly at the beginning of the night, stage N2 is generally followed by N3 (formerly stages 3 and 4), a period when 20 percent or more of each sleep epoch consists of slow waves, that is, waves of 0.5- to 2-Hz frequencies with peak-to-peak amplitudes of >75 μV over frontal regions. N3 is also defined as *slow wave sleep (SWS)*, *delta sleep,* or *deep sleep,* because the arousal threshold increases incrementally from stages N1 to N3. Previously, SWS was subdivided according to the proportion of slow waves in the epoch (stage 3, 20 to 50 percent; stage 4, >50 percent) and older studies use this nomenclature, but the validity and biological significance of this subdivision was called into question. Eye movements typically cease during stages N2 and N3, and EMG activity decreases further.

REM sleep, or stage R, is not subdivided into stages but is rather described in terms of tonic (persistent) and phasic (episodic) components. Tonic aspects of REM sleep include the activated EEG similar to that of stage N1, which may exhibit increased activity in the theta band and a generalized decrease of the tone of skeletal muscles except for the extraocular muscles and the diaphragm. Sawtooth waves, trains of triangular, serrated 2 to 6 Hz waves may be present as well. Phasic features of REM include irregular bursts of REMs and muscle twitches. An example of stage R showing tonic and phasic components is shown in Figure 1.22–2.

Organization of Sleep

The amount of sleep obtained during the night varies among individuals and with age; the National Sleep Foundation recently convened an expert panel to review the scientific literature to make recommendations for sleep duration across the lifespan. They concluded that most adults need about 7 to 9 hours of sleep per night to function optimally, although those over 65 might only require 7 to 8 hours. Although these are population-based recommendations, there also exist short sleepers who appear to function adequately with less than 6 hours per night as well as long sleepers who may need 12 or more hours per night. In addition to age and genetic factors that influence daily sleep needs, medical or psychiatric disorders also strongly influence sleep patterns. Regardless of the number of hours needed, the proportion of time spent in each stage and the pattern of stages across the night is fairly consistent in normal adults (Fig. 1.22–3). A healthy young adult will typically spend about 5 percent of the sleep period in stage N1 sleep, about 50 percent in stage N2, and 20 to 25 percent in each of stages N3 and R. Sleep occurs in cycles of NREM–REM sleep, each lasting approximately 90 to 110 minutes. SWS (stage N3) is most prominent early in the night, especially during the first NREM period, and diminishes as the night progresses. As SWS wanes, periods of REM sleep lengthen, while showing greater phasic activity and generally more intense dreaming later in the night. Older adults have decreased amounts of SWS.

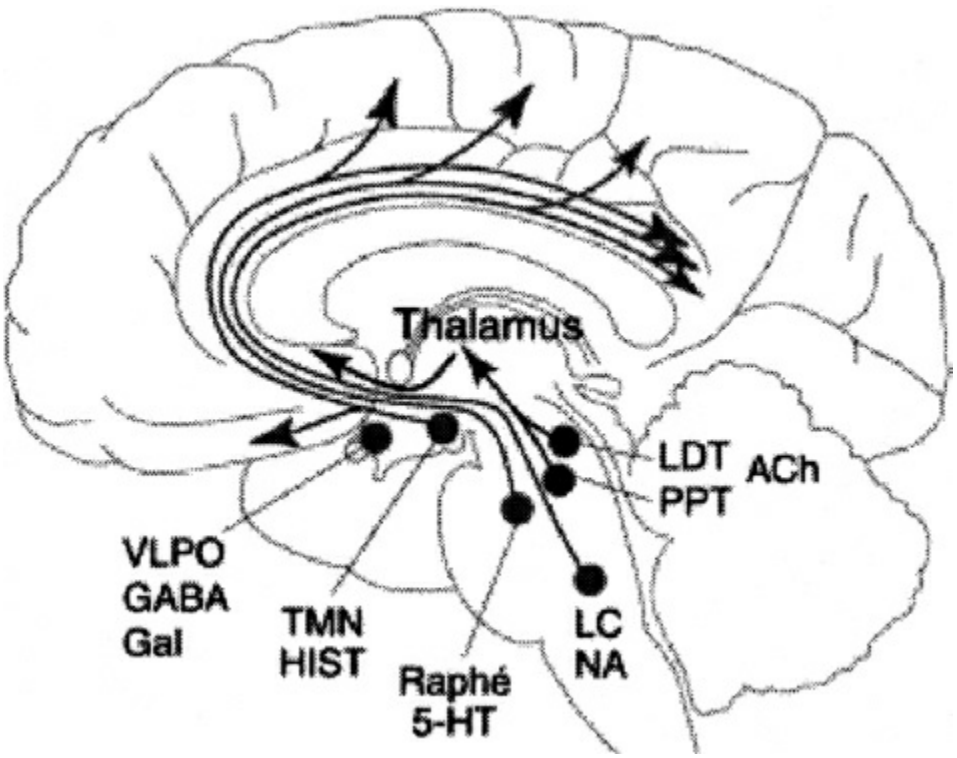

FIGURE 1.22–3. The ascending arousal system sends projections from the brainstem and posterior hypothalamus throughout the forebrain. Neurons of the laterodorsal tegmental (LDT) nuclei and pedunculopontine tegmental (PPT) nuclei send cholinergic

fibers (Ach) to many forebrain targets, including the thalamus, which then regulate cortical activity. Aminergic nuclei diffusely project throughout much of the forebrain, regulating the activity of cortical hypothalamic targets directly. Neurons of the tuberomammillary nucleus (TMN) contain histamine (HIST), neurons of the raphe nuclei contain serotonin (5-HT), and neurons of the locus coeruleus (LC) contain noradrenaline (NA). Sleep-promoting neurons of the ventrolateral preoptic (VLPO) nucleus contain γ-aminobutyric acid (GABA) and galanin (Gal). (From Saper CB, Chou TC, Scammell TE. The sleep switch: hypothalamic control sleep and wakefulness. *Trends Neurosci*. 2001;24:726.)

The ascending arousal system sends projections from the brainstem and posterior hypothalamus throughout the forebrain. Neurons of the LDT nuclei and pedunculopontine tegmental (PPT) nuclei send cholinergic fibers (Ach) to many forebrain targets, including the thalamus, which then regulate cortical activity. Aminergic nuclei diffusely project throughout much of the forebrain, regulating the activity of cortical and hypothalamic targets directly. Neurons of the tuberomammillary nucleus (TMN) contain histamine (HIST), neurons of the raphe nuclei contain serotonin (5-HT), and neurons of the locus coeruleus (LC) contain noradrenaline (NA). Sleep-promoting neurons of the ventrolateral preoptic (VLPO) nucleus contain GABA and galanin (Gal).

EFFECTS OF AGE ON SLEEP

Sleep patterns change markedly across the lifespan, with the most rapid changes occurring during the first years of life. Development of EEG patterns of sleep and wakefulness begins at about 24 weeks of gestational age, and differentiation into active (REM) and quiet (NREM) sleep occurs during the last trimester. Newborn infants spend about 14 to 17 hours per day sleeping, and premature infants may sleep even more. Infants do not show evidence of strong diurnal sleep patterns for the first several months of life; they exhibit short sleep–wake cycles of about 3 to 4 hours as well as a reduced length of active–quiet sleep cycles (about 50 minutes). At birth, active sleep occupies about half of their sleep time, and they tend to enter sleep through active rather than quiet sleep. At approximately 3 to 4 months of age, several important developmental changes occur: Babies shift to the adultlike pattern of initiating sleep with NREM, sleep starts to become consolidated during the night, and the sleep EEG shows more mature waveforms characteristic of NREM and REM sleep. Sleep duration also drops slightly, to 12 to 15 hours per day.

During early childhood, total sleep time continues to decrease, such that recommended sleep amounts for preadolescent school-aged children range from 9 to 11 hours per night, and REM sleep proportion drops to adult levels (20 to 25 percent). Napping normally continues during the preschool years but is often abandoned once children begin school full-time. Young children show the highest percentages of SWS, with particularly high arousal thresholds; this accounts for both for the difficulty in arousing them at the beginning of the sleep period as well as the high incidence of

bedwetting and SWS-related parasomnias such as sleepwalking and night terrors.

SWS amounts diminish significantly during adolescence, possibly related to cortical synaptic pruning. Irwin Feinberg has suggested that abnormalities in synaptic elimination may account for the seemingly coincidental timing of the maturation of sleep patterns and the increasing incidence of schizophrenia in late adolescence and early adulthood. In addition to showing decreases in SWS amounts, adolescents often decrease their total sleep time significantly, although this is probably due to behavioral changes rather than representing a true decrease in sleep need; their recommended sleep amount is estimated at 8 to 10 hours per night. They also show a tendency to become "night owls," preferring to stay up late rather than wake up early. This shift to "eveningness" may be related to an increase in the intrinsic period of the circadian clock.

In addition the changes in amounts of SWS, there are dramatic changes in the topography of scalp slow-wave activity that mirror the pattern of cortical brain maturation. Occipital regions, that develop first, show the most slow wave activity in very young children. During development, as the cortex matures in a posteroanterior direction, maximal slow-wave activity also shifts forward, showing the adult pattern of dominance over frontal cortex by late adolescence/early adulthood.

Older adults show an increased incidence of primary sleep disorders (e.g., sleep apnea and periodic limb movements), medical illnesses, and psychiatric disorders that all may interfere with sleep; it is therefore difficult to determine which sleep changes represent normal versus pathological aging. SWS declines across adulthood and may disappear entirely by age 60; slow-wave activity declines globally over the scalp. Sleep also becomes more fragmented, with prolonged latency to sleep onset, increased numbers of arousals, more time spent awake during the sleep period, and increased daytime napping. Older adults thus may spend more time in bed but obtain less sleep; 5 to 6 hours may be a sufficient amount of sleep for some individuals. The elderly also show a trend to wake up earlier and feel more alert in the morning in comparison to the evening. The percentage of REM sleep only shows a small decline with age in normal individuals. Patients with Alzheimer disease and other degenerative disorders of the CNS, however, show a loss of REM sleep and deterioration in diurnal patterns of sleep–wakefulness; this deterioration may become so severe that nursing home patients may not spend a single hour of the day where they are consistently either asleep or awake.

MONITORING HUMAN SLEEP

Sleep–wakefulness-related alterations in the human EEG were first reported by Hans Berger in 1929. Since this discovery, the EEG and other electrophysiological parameters have been used to investigate normal

human sleep. Starting in the 1970s, researchers and clinicians began to use similar monitoring techniques to characterize and diagnose sleep disorders. In addition to the EEG, the EOG and EMG are recorded to measure eye movements and muscle activity, respectively, both requisites to distinguish wakefulness and the stages of sleep. The EEG is recorded from electrodes affixed to the scalp overlying specific regions of the brain according to the International 10–20 system of electrode placement. Because the EEG features that define wakefulness and the stages of sleep are most readily recorded from different regions of the brain, a minimum of three strategically chosen regions (frontal, central, and occipital) are required. K-complexes and slow waves are optimally recorded with the frontal electrode. Sleep spindles (Fig. 1.22–1) are instead best recorded from electrodes placed over the central region. The third electrode is placed over the occipital lobe to optimize the detection of alpha activity, correlated to a relaxed waking state with closed eyes. Under certain research or clinical circumstances (e.g., diagnosis of sleep-related seizure disorders) additional electrodes are applied to obtain a higher spatial resolution of EEG activity. The EOG recording is used to detect REMs associated with wakefulness and REM sleep as well as the slow rolling eye movements that occur during stage N1 sleep. Because the retina is electrically negative compared to the cornea, eye movements generate small electrical fields that can be detected from electrodes attached to the skin near the eyes. The EMG recording is used to detect tonic and phasic changes in muscle activity that correlate with changes in behavioral state. In particular, during REM sleep skeletal muscle tone reaches the lowest tonic levels, reflecting the general paralysis associated with this stage of sleep. Typically, the EMG is recorded from electrodes attached to the chin. In clinical settings, additional EMG electrodes are placed over the anterior tibialis and intercostal muscles to detect leg movements and respiratory effort, respectively. Depending on the presenting symptoms, clinical monitoring may also include additional monitoring such as respiratory effort and flow sensors, ECG, oxyhemoglobin saturation, and limb EMG recording.

Standard Sleep Stage Scoring

Standardized visual scoring rules were first instituted by Rechtschaffen and Kales in 1968 and are still used, with more recent modifications, to quantify time spent in wakefulness and each sleep stage as well as the temporal distribution, or architecture, of sleep across the recording period. In addition, mathematical techniques, such as power spectral analysis based on the *Fast Fourier Transform*, are frequently employed to quantify the relative contributions of various brain wave frequencies to the overall EEG recording. For example, power spectral analysis has shown that slow-wave activity is greatest early in the sleep period and progressively declines across successive periods of SWS, a finding that provided the basis for the

two-process model of sleep regulation.

Sleep stage scoring yields several measures of sleep quantity and quality, including clinically relevant markers of sleep and psychiatric disorders. For each sleep measure, specific clinical or research circumstances may call for slightly different definitions. *Sleep latency* is the time elapsed from the start of the recording to the onset of any stage of sleep. *REM sleep latency* is defined as the time elapsed from sleep onset to the first epoch of stage R sleep. *Total sleep time* is the cumulative time spent in all sleep stages, and *sleep efficiency* is how much of the total recording time was spent in any stage of sleep. Other measures of interest include the proportion of sleep spent in each sleep stage.

Measuring Daytime Sleepiness

In addition to characterizing sleep patterns across the major sleep period, researchers and clinicians are sometimes interested in quantifying daytime sleepiness. Daytime sleepiness manifests as an increased propensity to fall asleep and/or a decreased ability to maintain wakefulness. Sleepiness can be measured using subjective questionnaires or objective electrophysiological monitoring. Sleep questionnaires such as the Stanford Sleepiness Scale (SSS) or Epworth Sleepiness Scale (ESS) are easy to use in the laboratory or physician's office. The SSS asks the individual to rate their current level of sleepiness, whereas the ESS asks the individual to rate their probability of falling asleep under various circumstances. Both questionnaires may be influenced by the individual's ability to assess their own level of alertness or propensity to fall asleep as well as their motivation to obtain treatment.

Objective measures of sleepiness generally require time-consuming, laboratory-based testing but yield a more reliable estimate of sleepiness. Two standardized objective tests are typically employed to measure daytime sleepiness; the multiple sleep latency test (MSLT) measures the propensity to fall asleep, whereas the maintenance of wakefulness test (MWT) measures the ability to maintain wakefulness. Both tests employ the electrophysiological sleep-monitoring techniques described above. The MSLT consists of four or five sleep latency tests spaced 2 hours apart. At the start of each test individuals are asked to lie quietly in a darkened room and allow themselves to fall asleep. The latency to sleep onset is recorded, and if sleep does not occur within 20 minutes, then the test is terminated. In the clinical setting, patients that fall asleep are usually allowed to sleep for 15 minutes to determine whether they exhibit a tendency to enter REM sleep in an abnormally short period of time. The mean sleep latency is calculated as well as the number of tests in which REM sleep was detected. A mean sleep latency <7 minutes reflects excessive sleepiness and may be indicative of a CNS hypersomnia such as narcolepsy or idiopathic hypersomnia, whereas a mean sleep latency ≥15 minutes is considered

normal. In conjunction with other symptoms, two or more tests with REM sleep are suggestive of narcolepsy. During the MWT, individuals are placed under similar conditions to that described for the MSLT, but rather than being asked to fall asleep, they are asked to stay awake. Although 20 to 40 minute tests can be used, the 40-minute protocol has the advantage of minimizing ceiling effects. A mean sleep latency >35 minutes on the 40-minute MWT is considered normal. Although both the MSLT and the MWT are objective tests of sleepiness, the results can be influenced by various factors such as an individual's prior sleep history. In the clinical setting, in particular, the goal is to assess whether a patient exhibits a pathological degree of sleepiness despite obtaining normal amounts of sleep at night; short sleep latencies on the MSLT are of limited clinical utility if the patient is clearly sleep-deprived. Consequently, it is critical that the patient's prior sleep history be evaluated by monitoring in the sleep laboratory during the preceding night as well as questionnaires logging sleep habits for the previous week. Under certain circumstances, patients may wear small devices on their wrist that detect and log motion, a correlate of wakefulness, for several days or weeks before undergoing laboratory testing. Such activity monitoring (or actigraphy) is also used to obtain objective measures of sleep at home in patients complaining of insomnia. Other objective measures of sleepiness based on power spectral analysis of the EEG, pupillary activity, and performance on cognitive tests are currently under development.

PHYSIOLOGY IN SLEEP

An appreciation of the physiological changes that occur during sleep is helpful for understanding the effects of normal and disturbed sleep on medical disorders. Various parameters of neuroendocrine and ANS physiology show changes related to circadian rhythms and/or sleep itself (Table 1.22–1).

Autonomic Nervous System

During NREM sleep and tonic REM sleep, there is a relative increase in parasympathetic activity relative to sympathetic activity. The ANS reaches its most stable state during SWS in comparison to wakefulness; for example, blood pressure, heart rate, and respiratory rate are at their lowest mean values and least variable during SWS. During phasic REM sleep, however, there are brief surges in both sympathetic and parasympathetic activity, resulting in a high degree of autonomic instability.

Cardiovascular System

Blood pressure, heart rate, and cardiac output decrease during NREM sleep, reaching their lowest average values and least variability in SWS.

Although, on average, these parameters remain somewhat reduced during REM sleep in comparison to waking, they attain their peak values during REM sleep. Arrhythmias are also more prevalent during REM sleep, which may contribute to the increased rate of cardiovascular mortality in the early morning, the time of greatest REM sleep propensity. It is also possible that increased rates of mortality related to cardiovascular causes seen in major depression might be related to the tendency of depressed patients to have increased amounts of REM sleep with greater phasic activity.

Pulmonary System

Temporary breathing instability and/or periodic breathing may occur at the onset of sleep related to the loss of waking-related respiratory drive as well as a decreased sensitivity of central chemoreceptors to pCO_2; sensitivity to pCO_2 declines further during REM sleep, along with a decrease in the ventilatory response to reduced pO_2. Respiratory rate and minute ventilation decrease during sleep, and upper airway resistance increases as a result of muscle relaxation, most significantly during REM sleep. These changes contribute to exacerbations of underlying pulmonary disease as well as sleep-related breathing disorders such as sleep apnea.

Thermoregulation. In addition to the circadian-related nocturnal decrease in body temperature, sleep has direct effects on thermoregulation. Brain and body temperature are downregulated during NREM sleep, particularly SWS, as a result of both a decreased hypothalamic temperature set point as well as active heat loss. People commonly experience this phenomenon when they go to sleep feeling somewhat cold and wake up several hours later to throw off their extra covers because they feel too warm and may have been sweating. During REM sleep, there is a decreased ability to regulate body temperature through sweating and shivering.

Neuroendocrine Changes. Most hormones that are regulated by the circadian system also show significant interactions with sleep–wakefulness patterns. Growth hormone (GH) is released primarily during the early part of the night, and its secretion is enhanced by SWS. Sleep also stimulates prolactin secretion, although prolactin peaks after GH, usually during the middle portion of the night. Pulses of GH and prolactin can occur after the onset of sleep, regardless of its timing, however. Both GH and prolactin may have feedback effects on sleep as well; GH seems to enhance SWS, whereas prolactin may increase REM sleep. In contrast, *TSH* reaches its peak level in the evening just prior to sleep onset; its secretion is inhibited by sleep and stimulated by sleep deprivation. The *HPA axis* is usually at its most inactive state at nocturnal sleep onset. Sleep onset inhibits cortisol release, whereas *ACTH* and *cortisol* levels rise at the end of the sleep

period, shortly before awakening, and likely contribute to morning arousal. Severe sleep disruption or sleep deprivation may have significant clinical effects on the endocrine system; for example, patients with obstructive sleep apnea show decreased levels of GH and prolactin, and sleep deprivation produces evidence of HPA axis activation in the evening of the day following deprivation.

Melatonin secretion is mediated by a combination of circadian control and effects of the light–dark cycle. It can only be released at night if it is dark; darkness during the day does not stimulate melatonin secretion. Thus melatonin can transduce the duration of the photoperiod; it has a significant role in timing reproductive function in some mammals, although its function in humans is less clear. Melatonin can feedback on the circadian clock and may serve to maintain entrainment, which is why is it sometimes recommended for the treatment of jet lag or sleep schedule disorders. It may also have a modest hypnotic effect in humans but can produce arousal in nocturnal animals, suggesting that it acts as a modulator of nocturnal behaviors.

Sexual Function

One of the characteristics of REM sleep in men is the occurrence of penile erections, beginning in infancy and persisting into old age. Nocturnal penile tumescence studies are therefore helpful in determining whether cases of impotence are related to organic or psychogenic etiologies. In women, REM sleep produces increased vaginal blood flow and clitoral erection. These changes are not necessarily linked to sexual content in associated dreams.

NEUROBIOLOGY OF SLEEP AND WAKEFULNESS

Sleep and wakefulness are governed by separate yet interacting systems. Although the specific mechanisms are not fully understood, it is clear that multiple structures and systems in the medulla, brainstem, hypothalamus, and basal forebrain are involved in the orchestration of wakefulness, NREM sleep, and REM sleep. No single brain lesion has been able to produce persistent insomnia or sleep. Given the importance of these behaviors to survival, it is not surprising that there is a certain amount of apparent redundancy in their mechanisms.

Wakefulness

As mentioned above, the waking EEG is characterized by an activated pattern with low-voltage fast activity. Correspondingly, PET studies show that during resting wakefulness blood flow and metabolic activity are higher than those in NREM sleep. The most active brain areas, as indicated by increased regional cerebral blood flow (rCBF), include the prefrontal

cortex, anterior cingulate parietal cortex, and precuneus, which are areas known to be involved in attention, cognition, and memory.

Maintenance of wakefulness is dependent on the *ascending reticular activating system* (ARAS), which is comprised of inputs from the oral pontine and midbrain tegmentum as well as posterior hypothalamus (Fig. 1.22–4). In animals, electrical stimulation of these regions produces an activated EEG pattern and behavioral arousal. Lesions may produce a comalike state, although recovery of cortical activity can sometimes occur given sufficient time, suggesting that structures outside the brainstem reticular formation are also involved in maintaining wakefulness. Clinically, lesions in the midbrain, diencephalon, or posterior hypothalamus can produce somnolence, stupor, or coma.

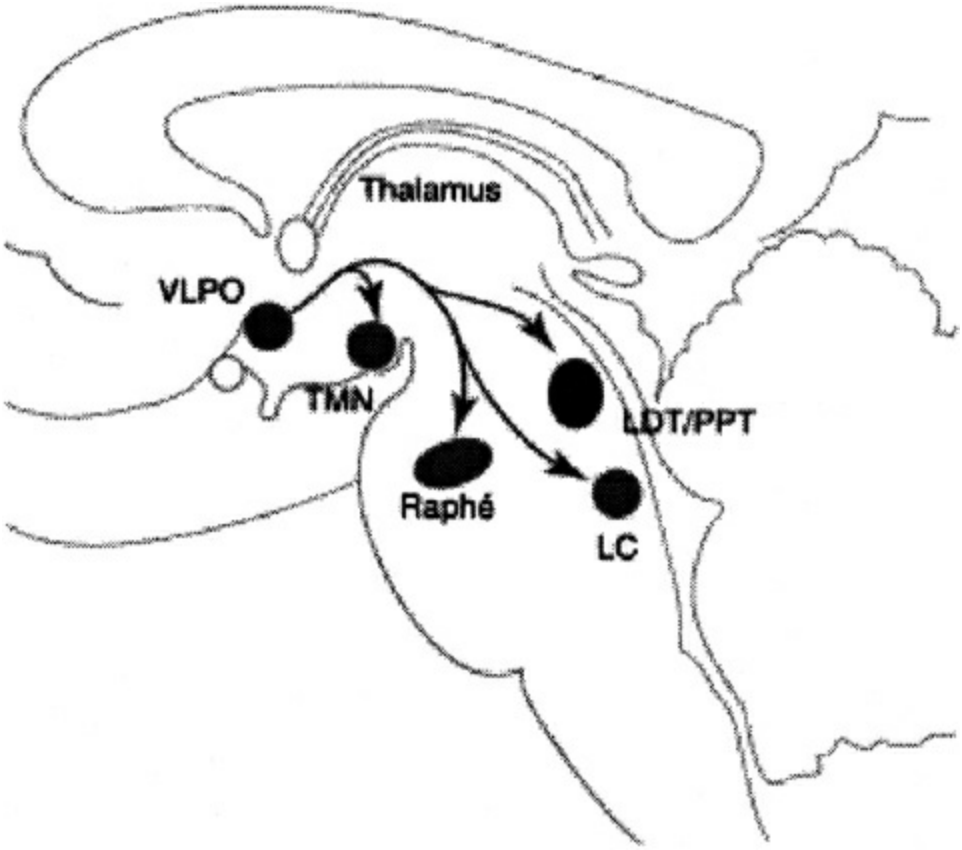

FIGURE 1.22–4. The projections from the ventrolateral preoptic (VLPO) nucleus to the main components of the ascending arousal system. LC, locus coeruleus; LDT, laterodorsal tegmental nuclei; PPT, pedunculopontine tegmental nuclei; TMN, tuberomammillary nucleus. (From Saper CB, Chou TC, Scammell TE: The sleep switch: hypothalamic control of sleep and wakefulness. *Trends Neurosci*. 2001;24:726, with permission.)

It is now known that several distinct structures and neurochemical systems with diffuse projections are involved in wakefulness, including noradrenergic cells in the *locus coeruleus* (LC), cholinergic cells in the *PPT* and *LDT* nuclei, histaminergic cells in the *TMN* of the posterior hypothalamus, orexin/hypocretin cells in the lateral hypothalamus, and

glutamatergic and GABAergic neurons in various structures in the CNS. The ARAS produces cortical activation via input to the thalamus as well as through an extrathalamic pathway with projections to the hypothalamus and basal forebrain.

Noradrenergic cells from the LC project directly throughout the forebrain and cerebral cortex and show their highest discharge rates during wakefulness. These cells decrease their firing during NREM sleep and cease firing altogether during REM sleep. Converging evidence suggests that wakefulness and sleep differ not only in terms of behavior, metabolism, and neuronal activity but also in terms of gene expression. LC cells are responsible for at least some of the changes in gene expression that occur in the brain between wakefulness and sleep, and specifically for the upregulation of plasticity-related genes involved in synaptic potentiation during wake. The LC has long been considered a homogeneous cell group with broad and uniform action on the cortex, but recent experiments found little overlap among the LC projections to three subregions of the rat prefrontal cortex, including the medial prefrontal cortex (mPFC), or between the LC afferents to each of these prefrontal areas and the primary motor cortex, M1. Moreover, relative to the neurons targeting M1, LC cells projecting to mPFC have spontaneous firing rate three times as high, and are more excitable, suggesting that decision-making prefrontal areas are under a stronger modulatory influence by the LC than movement-generating cortical areas.

Cholinergic cells from the oral pontine region fire at high rates when the EEG is activated, that is, during wakefulness and REM sleep, but reduce firing during NREM sleep. They promote cortical activation through inputs to the thalamus, hypothalamus, and basal forebrain. In addition, cholinergic cell bodies in the basal forebrain, including the nucleus basalis, substantia innominata, diagonal band of Broca, and septum receive input from the ARAS and in turn provide excitatory input to the entire cortex. Stimulation of the basal forebrain results in cortical release of acetylcholine, EEG activation, and depolarization of cortical neurons. Moreover, the cholinergic cells of the basal forebrain also fire maximally during wakefulness and REM sleep and minimally during NREM sleep. Optogenetic experiments in mice suggest that cholinergic cells in the pons promote REM sleep more than wake, while cholinergic cells in the basal forebrain may promote both. Loss of cholinergic cells in Alzheimer patients is associated with slowing of the cortical EEG. Drugs with anticholinergic activity, including TCAs and atropine, can cause sedation, suppress REM sleep, and increase slow-wave activity. However, cholinergic agonists (e.g., nicotine) or anticholinesterase inhibitors (e.g., neostigmine) enhance arousal.

Histaminergic neurons in the TMN of the posterior hypothalamus also appear to have an important wakefulness-promoting function, in part

inferred from the fact that antihistaminergic drugs typically produce sedation. The significance of this region for waking was first identified by Constantin von Economo in the early part of the 20th century following an outbreak of viral encephalitis; encephalitis lethargica, as it was called, produced lesions of the posterior hypothalamus and profound somnolence. Histaminergic TMN neurons project throughout the cortex and, like noradrenergic cells, fire at the highest rates during wakefulness and are inhibited during sleep. However, cholinergic and noradrenergic cells in the pons increase firing hundreds of milliseconds or even a few seconds before the EEG switches from the synchronized pattern of NREM sleep to the activated pattern of wake, suggesting a causal role in arousal, while histaminergic neurons increase firing only at wake onset. Histamine infusion into the CNS causes arousal, whereas experimental lesions of the TMN decrease waking and increase SWS and REM sleep. The wakefulness-promoting effect of histamine is mediated by H_1 receptors. In the thalamus, cortex, basal forebrain, and pontine tegmentum, histamine promotes wakefulness by enhancing glutamatergic and cholinergic transmission.

The dopaminergic system also appears to modulate arousal. Dopamine-containing neurons in the substantia nigra and VTA innervate the frontal cortex, basal forebrain, and limbic structures. The mean firing rate of these cells does not change across behavioral states. However, their bursting activity, which is known to induce synaptic dopamine release, increases during the consumption of palatable food and during REM sleep relative to NREM sleep. Lesions of areas containing dopaminergic cell bodies in the ventral midbrain or their ascending pathways can lead to the loss of behavioral arousal while maintaining cortical activation. Psychostimulants such as amphetamines and cocaine that block reuptake of monoamines including norepinephrine, dopamine, and serotonin promote prolonged wakefulness and increase both cortical activation and behavioral arousal.

The importance of the peptide hypocretin (orexin) in the maintenance of wakefulness was defined through discovering its role in the disorder narcolepsy. Hypocretin is produced by cells in the lateral hypothalamus that provide excitatory input to all components of the ARAS, including the LC, PPT and LDT, VTA, basal forebrain, and TMN. These cells, too, are most active during waking, especially in relation to motor activity and exploratory behavior, almost completely stop firing during both NREM and REM sleep, and start firing again before the EEG activates upon arousal. Optogenetic stimulation of hypocretin cells promotes arousal from sleep, although not as quickly as the stimulation of cholinergic and noradrenergic cells. Narcolepsy in animal models is related to deficits in the hypocretin system; canine narcolepsy is caused by a mutation in the hypocretin type 2 receptor gene, whereas narcoleptic symptoms (sleep attacks and sleep-onset REM periods) occur in hypocretin knock-out mice. Mouse experiments also show that the slow and progressive elimination of

hypocretin cells mimics many aspects of human narcolepsy: sleep fragmentation is the first symptom, and cataplexy develops only after ~95 percent of all hypocretin cells are lost. Human narcoleptics show the loss of hypocretin cells and/or protein in the CSF. Hypocretin cell loss is also present in Parkinson disease, which is often associated with sleep disturbances similar to narcolepsy.

Serotonergic cells from the *dorsal raphe* nucleus also project widely throughout the cortex. Serotonergic neurons, like noradrenergic neurons, fire at higher levels during waking, lower levels in NREM sleep, and fall silent during REM sleep. However, in sharp contrast to noradrenergic neurons, serotoninergic neurons are inactivated during orientation to salient stimuli and are activated instead during repetitive motor activity such as locomotion, grooming, or feeding. SSRIs tend to decrease sleep time and increase arousal during sleep. The role of serotonin in sleep is not straightforward, however, since there is also evidence that serotonin may be involved in sleep induction (see below).

A number of other neurotransmitters and neuromodulators appear to have wakefulness-promoting effects. These include substance P, neurotensin, epinephrine, and hypothalamic peptides such as corticotrophin-releasing factor, VIP, and thyrotropin-releasing factor, all of which can increase arousal levels. Cortisol also promotes wakefulness. It is thus possible that sleep disturbance in depression including early morning awakening could be related in part to the associated hyperactivity of the HPA axis.

NREM Sleep

The EEG of NREM sleep is dramatically different from that of waking and is characterized by oscillatory waveforms such as sleep spindles, K-complexes, slow waves (0.5 to 2 Hz), and slow oscillations (mainly 0.7 to 1 Hz). Brain activation generally decreases in NREM sleep, particularly SWS, which is characterized by an overall decrease in CBF. PET imaging studies show the deactivation of many structures, including the brainstem, thalamus, anterior hypothalamus, basal forebrain, basal ganglia, cerebellum, and frontal, parietal, and mesiotemporal cortical areas.

The control of NREM sleep, like wakefulness, involves multiple structures ranging from the lower brainstem through the thalamus, hypothalamus, and forebrain. Electrophysiological studies in animals have clarified the production of various thalamocortical oscillations that produce the characteristic waveforms in NREM sleep. The generation of sleep oscillations requires the interplay between intrinsic cellular properties and synaptic activity mediated by cortico-cortical, cortico-thalamo-cortical, and thalamoreticular loops. Work in animals has shown that, shortly before the transition from waking to sleep, changes in the activity of cholinergic, noradrenergic, histaminergic, hypocretinergic, and glutamatergic

neuromodulatory systems with diffuse projections belonging to the ARAS bring about a change in the firing mode of thalamic and cortical neurons. Thalamocortical cells are hyperpolarized, whereas reticulothalamic cells are facilitated and further inhibit thalamocortical cells, with the consequence that sensory stimuli are gated at the thalamic level and often fail to reach the cortex. Rebound firing due to the activation of intrinsic currents in thalamocortical cells leads to the emergence of oscillations in the spindle frequency range within local thalamoreticular circuits. Local thalamic spindle sequences are globally synchronized and grouped with other rhythmic activities by the slow oscillations that originate in the cortex. Intracellular recordings have shown that the slow oscillation is the result of a brief hyperpolarization of cortical neurons (lasting a few hundred milliseconds), which is seen in the surface EEG as a high-amplitude negative wave. The hyperpolarization phase, also known as the down state, is followed by a slightly longer depolarization phase, known as the up state, during which the firing of cortical neurons entrains and synchronizes spindle sequences in thalamic neurons, resulting in EEG-detectable spindles. K-complexes are made up of the cortical depolarization phase followed by its triggered spindle. The slow oscillation also organizes delta waves, which can be generated both within the thalamus and in the cortex. The hyperpolarization phase of the slow oscillation is associated with the virtual absence of synaptic activity within cortical networks. By contrast, during the depolarized phase, cortical cells fire at rates that can be higher than those seen in waking. Moreover, during the depolarized phase, thalamocortical neurons often fire rhythmically in the gamma range (around 40 Hz), which was previously thought to be exclusively associated with activated states such as wakefulness and REM sleep. The slow oscillation is found in virtually every cortical neuron and is synchronized across the cortical mantle by cortico-cortical connections. Intrinsic currents are also involved in initiating and terminating the oscillation.

The importance of hypothalamic structures for sleep induction was recognized in early studies in which electrical stimulation of the anterior hypothalamus resulted in increased slow-wave activity in the cortex. Conversely, some cases of encephalitis lethargica, in which lesions occurred in the anterior rather than posterior hypothalamus, were characterized by severe insomnia. A few years ago attention became focused on a small portion of the anterior hypothalamus, the VLPO area, as a possible sleep switch. It is now clear, however, that many other neurons scattered through the anterior hypothalamus (for instance, in the median preoptic nucleus, MnPO) and the basal forebrain, also play a major role in initiating and maintaining sleep. These neurons tend to fire during sleep and stop firing during wakefulness. MnPO neurons can also track sleep pressure, because their firing is higher during the first sleep cycle, when sleep pressure is high, and declines in later cycles. When they are active, many of them

release GABA and the peptide galanin and inhibit most wakefulness-promoting areas, including cholinergic, noradrenergic, histaminergic, hypocretinergic, and serotonergic cells. In turn, the latter groups of cells inhibit several sleep-promoting neuronal groups. This reciprocal inhibition provides state stability, in that each state reinforces itself as well as inhibits the opposing state. A subset of cortical GABAergic cells is also active during sleep and may track sleep pressure. Finally sleep-promoting GABAergic cells were recently identified in the medulla, in the parafacial zone. Stimulation of this cell group promotes NREM sleep by inhibiting the wake-promoting neurons of the parabrachial nucleus, located at the junction between midbrain and pons, which send glutamatergic projections to the basal forebrain.

In terms of NREM sleep neurochemistry, many substances modulate sleep, but no unique sleep factor has been identified. GABA, the major inhibitory neurotransmitter in the CNS, appears to be involved in thalamocortical oscillations and in the inhibition of waking centers by sleep-active cells. Most hypnotics, including barbiturates, benzodiazepines, and several of the newer nonbenzodiazepine hypnotics act by enhancing GABA transmission. However, many GABAergic neurons in the brainstem and basal forebrain are active during wake.

Adenosine has been increasingly recognized as having a role in sleep. Caffeine probably exerts its stimulant effects by blocking adenosine receptors. Adenosine, a degradation product of adenosine triphosphate (ATP), is a byproduct of brain metabolism whose levels increase as a result of increased glutamatergic activity. It accumulates in the basal forebrain and cerebral cortex during prolonged wakefulness and decreases during sleep, suggesting that it may serve to transmit the homeostatic signal for sleep. Adenosine inhibits cholinergic neurons in the pons and basal forebrain, as well as hypocretin neurons and cortical glutamatergic neurons. By contrast, adenosine excites VLPO neurons, and its infusion promotes NREM sleep.

Early studies raised the possibility that serotonin might also be involved in SWS, because lesions of serotonergic nerve cells in the dorsal raphe led to insomnia. However, more recent evidence suggests that some of the early results may have been confounded by lesion-induced loss of thermoregulation, which caused increased locomotion (and thus decreased sleep) as a compensatory response to maintain body heat. Serotonergic neurons decrease firing rates in NREM sleep and are completely inhibited during REM. However, they inhibit cholinergic neurons and produce behavioral inhibition, raising the possibility that they help facilitate sleep onset although they may not be involved in directly inducing or maintaining sleep. It remains controversial as to how much serotonin contributes to sleep versus arousal.

A large number of other substances, including peptides and

neuromodulators, have been attributed with sleep-promoting properties. These include a variety of hormones, such as melatonin, αMSH, growth-hormone-releasing factor, insulin, CCK, and bombesin; cytokines such as interleukin-1, interleukin-6, and TNF; muramyl peptides produced from gut bacteria and dozens of other substances. Most of these factors have mild hypnotic and/or circadian effects consistent with the usual timing of their release (e.g., melatonin or growth–hormone-releasing factor, which is normally released at night) or the physiological state of the organism (e.g., cytokines produced during infectious illness promote sleep).

REM Sleep

REM sleep is characterized by an activated EEG and increased neuronal activity and CBF. Recent studies using functional imaging techniques have shown that during REM sleep there are some brain regions showing increased activation as well as others with decreased activation in comparison to wakefulness. Areas involved in REM sleep generation in the mesopontine tegmentum, thalamus, posterior cortical areas, and limbic areas, particularly the amygdala, are highly activated during REM sleep. In contrast, frontal and parietal cortices are relatively deactivated.

REM sleep is somewhat unique among the three states because a specific brain region—the *pons* and *caudal midbrain*—are both necessary and sufficient to generate the features of REM sleep and represent the final common pathway for the induction of REM sleep. A series of transection studies has demonstrated that REM sleep is preserved only in the portion of the CNS containing these structures. Bilateral lesions within the pons and caudal midbrain can completely eliminate REM sleep. Although these findings have created a tendency to focus on brainstem mechanisms of REM sleep, more rostral brain regions, including the preoptic area, are also important in the homeostatic regulation of REM sleep and in organizing REM episodes through the night.

As in wakefulness, cholinergic neurons produce EEG activation and a hippocampal theta rhythm during REM sleep. LDT/PPT neurons provide input to the thalamus and cholinergic basal forebrain neurons that in turn activate the limbic system and cortex. Hobson and McCarley proposed the reciprocal interaction hypothesis to explain NREM–REM cycles based on interactions between cholinergic and aminergic neurons in the mesopontine junction. Cholinoceptive and/or cholinergic REM-on cells in the PPT and LDT regions become activated during REM sleep, whereas noradrenergic or serotonergic REM-off cells are inhibitory of the REM-on cells. The aminergic cell groups are most active during waking; they decrease activity somewhat during NREM sleep, and meanwhile cholinergic activity increases to turn on REM sleep. REM sleep episodes are terminated because REM-on cells are self-inhibitory and provide excitatory input to the REM-off cells. Recent experiments show that

GABAergic and glutamatergic neurons in the mesopontine tegmentum are also very important in the control of REM sleep.

The role of cholinergic–monoaminergic interactions in regulating REM sleep is supported by a variety of experimental data. Local infusion of cholinergic agonists such as carbachol in the region of the LDT/PPT of cats or systemic administration of physostigmine, arecoline, pilocarpine, or other cholinergic agonists in humans produces prolonged REM sleep episodes and reduced latency to REM sleep. Cholinergic induction of REM sleep appears to be related primarily to activation of M_2 muscarinic receptors in the pontine reticular formation. Facilitation of REM sleep also results from the depletion of brainstem monoaminergic activity, for example, in human subjects with an acute depletion of serotonin after being fed a tryptophan-deficient diet. Most antidepressants cause significant reductions in REM sleep, particularly those that increase synaptic availability of norepinephrine and/or serotonin; anticholinergic effects seen in TCAs and MAOIs may also contribute to REM sleep suppression.

Muscle atonia in REM sleep can be eliminated by small lesions in the pontine reticular formation lateral to the LC or lesions in the medial medulla that eliminate inhibitory input from this area to the spinal motoneurons. Spinal motoneurons are inhibited by both GABA and glycine, and excited by noradrenaline, serotonin, and glutamate. The brainstem circuit responsible for REM sleep atonia includes two major pathways: (1) a glutamatergic projection from the subcoeruleus area to the medullary reticular formation, which then sends GABAergic and glycinergic inputs to motoneurons; (2) a glutamatergic projection from the subcoeruleus area to GABAergic interneurons, which then inhibit motoneurons. Disfacilitation of spinal motoneurons resulting from decreased noradrenergic and serotonergic activity also contributes to suppression of muscle tone during REM sleep. Several clinical conditions illustrate how atonia may be separated from the state of REM sleep. In narcolepsy, muscle atonia can occur during wakefulness, either as cataplexy, a sudden loss of muscle tone usually brought on by emotional stimuli, or sleep paralysis, in which atonia persist briefly after waking out of REM sleep. In contrast, patients with REM sleep behavior disorder do not develop atonia during REM sleep and act out their dreams, sometimes with such violence that they may injure themselves or their bed partners. REM-sleep–suppressing antidepressants such as TCAs, MAOIs, and SSRIs may induce REM sleep behavior disorder in some individuals.

In animals, ponto-geniculo-occipital (PGO) waves that occur sequentially in the pons, lateral geniculate nucleus of the thalamus and occipital cortex appear just prior to the initiation of REM periods and in conjunction with phasic activity in REM sleep, including eye movements and muscle twitches. They originate from cholinergic burst cells in the

peribrachial region and are suppressed by serotonergic cells in the raphe nuclei. Eye movements during REM sleep are tightly linked to PGO waves and are mediated by inputs to vestibular neurons that in turn activate oculomotor cells. PGO waves are thought to be the internal representation of orienting responses, because they can also be observed during a startle response during waking. In humans PGO-like potentials have been recorded in the pons before and during REM sleep, and in the subthalamic nucleus.

In addition to imaging data showing dramatically increased activity in forebrain structures during REM sleep, previous anatomical work has also suggested forebrain involvement in REM sleep regulation. Transection studies that separate the forebrain from pons disrupt NREM–REM cycling caudal to the transection. The amygdala has reciprocal connections with REM-generating brainstem regions, and electrical stimulation or infusion of carbachol into the amygdala can increase REM sleep. Abnormal activation of limbic structures occurs in depression and may contribute to associated changes in REM sleep.

CIRCADIAN AND HOMEOSTATIC REGULATION OF SLEEP

The regulation of sleep—both NREM and REM—involves at least two key components—a circadian one and a homeostatic one. The circadian component is responsible for the change in sleep propensity that is tied to the time of day, with obvious adaptive advantages. The homeostatic component refers to the fact that the longer one stays awake, the greater the propensity to sleep, and it represents the essential aspect of sleep whose function remains mysterious.

The Two-Process Model

Several models of sleep regulation positing a circadian and a homeostatic process have been proposed and validated on the basis of a large amount of data. One of the most influential is the *two-process model* developed by Alexander A. Borbély and colleagues, which predicts sleep propensity based on the interaction between the homeostatic *process S* and the circadian *process C*. Process S builds up across the day in response to the increase in sleep pressure caused by wakefulness and decreases during sleep. The circadian process C for sleep propensity, however, reaches its peak during the latter half of the night. Thus nocturnal sleep onset is primarily driven by process S, whereas process C maintains sleep through the latter part of the night. It is common for humans to have a brief period of arousal in the middle of the night, possibly related to a reduction in overall sleep drive from the fall in process S before process C has reached its maximal values. Similarly, the tendency for afternoon napping may be caused by the increase in process S across the day before process C has reached its lowest

values in the late afternoon/early evening. Although napping during the day has become relatively uncommon in industrialized societies, most people experience a period of increased sleepiness in the afternoon. This afternoon "dip" can produce significant daytime sleepiness in individuals who are already somewhat sleep deprived, are taking sedating medications, and/or have sleep disorders causing excessive daytime sleepiness such as sleep apnea.

Circadian Rhythms

In humans and other mammals, the primary pacemaker for generating circadian rhythms lies in the *SCN* of the hypothalamus. The SCN regulates a number of neuroendocrine and behavioral parameters, including sleep propensity as measured by process C, to coordinate the state of the organism with the 24-hour light–dark cycle. Circadian sleep regulation is strongly linked to the endogenous temperature rhythm; subjective sleepiness, sleep propensity, as well as REM sleep propensity are all maximal at the minimum (nadir) of core body temperature, usually in the very early morning, several hours prior to waking up. Sleep tendency is greater on the falling phase of the temperature curve, during the night. When core body temperature begins its rising phase in the morning hours, people tend to wake up; arousal levels, performance, and cognitive function are maximal in association with the rise of body temperature across the day.

In animals with lesions of the SCN, sleep is no longer concentrated in one main episode but is dispersed across the entire 24-hour cycle. Still, in these animals, sleep propensity increases as a function of previous waking. Thus, although process C and process S normally work together and interact significantly, they can to a large extent be separated. Several approaches have been used to investigate specifically the circadian regulation of sleep in humans. The *constant routine* protocol was designed to minimize the influences of factors other than the circadian clock on behavioral state. In this protocol, subjects are kept awake, usually for more than 24 hours, while sitting in bed in a dimly lit room. This technique has been used successfully to demonstrate the persistence of various circadian outputs in the absence of external cues, but the need to enforce sleep deprivation limits the duration of the experiment. With *temporal isolation,* subjects are placed in an environment without time cues for periods of weeks to months. Although circadian rhythms of sleep persist, the "day length" as defined by the period between successive bedtimes is close to 25 hours. These results led to the conclusion that the endogenous period of the human circadian clock is about 25 hours and that the light–dark cycle serves to entrain it to the 24-hour day.

To accurately determine the endogenous period of the circadian pacemaker, it is necessary to disentangle the circadian output from the

effects of sleep or prolonged sleeplessness; however, neither of the two protocols discussed above can satisfactorily achieve this. The *forced desynchrony* protocol, originally developed by Nathaniel Kleitman in 1938, consists of having subjects live on a 28-hour "day" while living in temporal isolation. Since it is not possible to entrain the human circadian clock to a period of 28 hours, after a sufficient period of time, sleep periods will have occurred at all phases of the circadian cycle. Thus, circadian and homeostatic influences on human sleep organization can be teased apart. Data from forced desynchrony studies suggest that the endogenous human circadian period is, in fact, close to 24 hours (24.1 to 24.2 hours). Moreover, they show that SWS is primarily regulated by the homeostatic sleep drive, whereas REM sleep is primarily regulated by the circadian clock. However, REM sleep is also homeostatically regulated, as indicated by increased attempts to initiate REM sleep if that stage of sleep is prevented.

More recent studies have demonstrated that the circadian timekeeping is more complex; in addition to the central pacemaker in the SCN, there are peripheral clocks in many cells and organs. These peripheral clocks with their local circadian oscillations are likely important in the regulation of secretion of multiple hormones, and both misalignment of the central SCN oscillator with the light-dark cycle as well as internal desynchronization between the SCN and peripheral oscillators may be contributory to metabolic and endocrine disorders, as well as to a number of psychiatric disorders.

HOMEOSTATIC REGULATION OF SLEEP AND THE EFFECTS OF SLEEP DEPRIVATION

In humans, as well as in virtually all animal species in which sleep has been carefully studied, sleep deprivation produces sleepiness and increased sleep pressure that soon become overwhelming. Sleep deprivation is followed by a "sleep rebound," that is, a compensatory increase in the duration and/or the intensity of sleep. After sleep deprivation, sleep latency is decreased and sleep efficiency is increased; that is, sleep is less fragmented. The amount of NREM sleep (especially stage N3 in humans) increases, together with markers of NREM sleep intensity such as slow-wave activity, which measures the predominance of slow waves in the cortical EEG. REM sleep amount also increases, but it is unclear whether this is also true for REM sleep "intensity." Such exquisite homeostatic regulation is one of the most important indications that there must be a distinct physiological, biochemical, or molecular process that builds up beyond its usual level in the brain if sleep initiation is postponed.

The most prominent effect of total sleep deprivation in humans is cognitive impairment, with striking practical consequences. It is estimated that 35 to 40 percent of Americans suffer from chronic sleep disorders resulting in difficulty falling asleep or daytime sleepiness. Each year, the

costs attributed to sleep problems, including the direct costs of treatment as well as the indirect costs of absenteeism, decreased productivity, accidents, and increased morbidity and mortality range into the hundreds of billions of dollars in the United States alone. Moreover, the National Highway Traffic Safety Administration estimates conservatively that each year drowsy driving is responsible for at least 100,000 automobile crashes, 71,000 injuries, and 1,550 fatalities (National Sleep Foundation, 2002). A sleep-deprived person tends to take longer to respond to stimuli, particularly when tasks are monotonous and low in cognitive demands. However, sleep deprivation produces more than just decreased alertness. Tasks requiring higher cognitive functions, such as logical reasoning, encoding, decoding and parsing complex sentences; complex subtraction tasks, and tasks requiring divergent thinking, such as those involving the ability to focus on a large number of goals simultaneously, are all significantly affected even after one single night of sleep deprivation. Tasks requiring sustained attention, such as those including goal-directed activities, can also be impaired by even a few hours of sleep loss. Thus, sleep loss causes attention deficits, decreases in short-term memory, speech impairments, perseveration, and inflexible thinking. These deficits can explain why sleep-deprived subjects underestimate the severity of their cognitive impairment, often with tragic consequences. Another reason people may underestimate their impairment due to sleep loss is that the lack of sleep does not completely eliminate the capacity to perform but rather makes the performance inconsistent and unreliable. Thus, a sleepy driver will either respond normally to an emergency or not at all, due to rapid changes in vigilance state and the sudden intrusion of microsleeps during waking. Similarly, subjects may still be able to transiently perform at baseline levels in short tests even after 3 to 4 days of sleep deprivation. However, the same subjects will perform very poorly when engaged in tasks requiring sustained attention. New evidence suggests that not just a few hours of sleep but several days of normal sleep–wake patterns are required to normalize cognitive performance after sleep deprivation.

Cognitive impairment is unfortunately not only the consequence of total sleep deprivation. Cognitive performance is also affected by sleep restriction (6 hours per night or less) if it continues for several days and by chronic sleep discontinuity such as that occurring in patients with chronic pain, sleep apnea, or other sleep disorders. According to the 2002 “Sleep in America” poll conducted by the National Sleep Foundation, United States residents of at least 18 years of age slept on average 6.9 hours during the weekdays and 7.5 hours on weekends. Twenty-four percent of the respondents in this poll reported that during weekdays they sleep less than they needed to in order to avoid feeling sleepy the next day. Whether this chronic sleep restriction is sufficient to affect objective measures of cognitive performance is not known, but it is certainly concerning given the

data from sleep restriction studies that point to impairment and decreased performance. It is becoming apparent that almost all sleep measures, such as the duration of total sleep or of SWS and the amount of slow-wave activity in the cortical EEG during NREM sleep, show great variability across different subjects but high consistency within the same individual from one night to another. The extent of the cognitive impairment after sleep deprivation also varies significantly across subjects. Converging evidence shows that individual differences in the cognitive impairment caused by sleep deprivation are stable over time and trait-like, suggesting a genetic basis.

Brain and peripheral tissues respond differently to sleep loss. Like in sleep-deprived animals, the peripheral metabolic rate is increased in sleep-deprived human subjects and in normal sleepers on nights of poor sleep relative to baseline nights; this increased rate is also present generally in people with insomnia relative to normal sleepers. In both animals and humans, glucose metabolism is higher in many brain regions in waking than that in NREM sleep. After 1 day of sleep deprivation, selective brain areas can still be activated metabolically when the subject is engaged in specific tasks. However, the global cerebral metabolic rate does not increase and actually decreases relative to normal waking values in areas such as the thalamus and the midbrain. Thus, while peripheral metabolic rate is persistently increased during sleep deprivation, brain metabolic rate is not. This may be an indication that the brain cannot sustain high-energy metabolism for too long.

In addition to causing cognitive impairment, sleep deprivation in humans may also affect various physiological systems with impacts on overall health. It has been suggested that sleep loss can affect host defense systems; for example, sleep-deprived rats show increased rates of bacteremia. Sleep deprivation has also been shown to lead to decreased glucose tolerance, increased SNS activation, and elevated cortisol levels, suggesting that it may contribute to disorders such as diabetes, hypertension, and obesity. Patients with insomnia have increased rates of health problems, including cardiac disease, further suggesting a possible causal relationship between reduced sleep amounts and health outcomes. Some studies in both humans and animals have linked long-term disturbances in sleep with reduced longevity. Others have suggested that in fact short sleep and insomnia are associated with little risk distinct from comorbidities. Whether there is a U-shaped association between sleep duration and mortality risk (i.e., both short sleep and long sleep are associated with increased mortality), remains unclear and meta-analyses have come to conflicting conclusions. The largest analysis, published in 2003 and based on 42 prospective studies in 35 distinct study populations worldwide, concluded that the evidence for a robust U-shaped association is still lacking.

The most extreme case of prolonged sleep loss in humans is observed in patients affected by fatal familial insomnia (FFI), a prion disorder characterized by near-complete loss of sleep and associated with various neurological symptoms and spongiform degeneration in selected brain regions. The disease is invariably fatal after a course of a few months to 2 to 3 years. In FFI with a short clinical course (death in <1 year), insomnia is almost complete from the onset, and spongiform degeneration is mainly restricted to the thalamus and inferior olive but does not extend to the cerebral cortex. In FFI cases with a longer clinical course (death in 2 to 3 years), insomnia develops more gradually, and spongiform degeneration is found in most cortical regions. While several clinical aspects of FFI can be attributed to diffuse spongiform degeneration, it is likely that sleep loss per se plays an important role in determining the evolution of the disease, especially in short-course FFI. Indeed, the severity of the clinical course correlates with the severity of the insomnia rather than with the accumulation of prion protein.

In humans, sleep deprivation is never enforced for more than 3 to 4 days (the record is 11 days, but in only one subject), making it impossible to determine whether there are other more severe effects of very prolonged sleep loss. Prolonged sleep deprivation is possible in animals, where several techniques have been used, from forced locomotion to pharmacological stimulation with amphetamines. Irrespective of the method used, an important and consistent finding is that even drastic manipulations such as brain electrical stimulation are unable to enforce complete and uninterrupted wakefulness after the first 24 hours. Even prolonged sleep deprivation attempts cannot enforce complete and uninterrupted wakefulness for more than one day. Evidently, sleep pressure overcomes whatever method is used to maintain wakefulness, in animals as well as in humans. This pressure can manifest itself in different ways. The most obvious way is “microsleeps,” short lapses lasting a few seconds during which the waking EEG clearly switches to a sleep pattern. Microsleeps can cause obvious impairment in performance, for instance losing the control of the car while driving. Less obvious changes consist in an overall “slowing of the EEG,” with an increase in low range frequencies (delta/theta activity). This change is detectable by quantitative analysis of the EEG but not by simple visual observation, although it may still affect performance in cognitive tasks. Finally, an even more subtle change is the recently identified phenomenon of “local sleep” during wake. This phenomenon, first described in rodents using intracortical recordings, refers to the fact that during prolonged wakefulness rats can move around and show a normal wake activated EEG, while some cortical neurons may go “offline” and stop firing, showing down states of membrane hyperpolarization and neuronal silence similar to those normally present only during NREM sleep. Because these down states are local and short lasting, they can only

be detected using invasive intracortical recordings. However, in humans, a useful EEG correlate of this phenomenon is the occurrence of local delta/theta waves (2 to 6 Hz), which can be identified using high-density EEG (hd-EEG, 256 channels). When human subjects stay awake for 24 hours and engage in intensive learning of cognitive tasks, hd-EEG can detect a local increase in delta/theta waves mainly in the cortical areas mostly involved in these tasks. During recovery sleep after sleep deprivation, the same areas will show a local increase in SWA.

The most comprehensive series of sleep deprivation studies in animals has been performed using the *disk-over-water apparatus* (DOW), which can prevent sleep for days and even weeks. This method uses minimal stimulation to enforce chronic sleep deprivation in the sleep-deprived rat, while it simultaneously applies the same stimulation to the control rat but without severely limiting its sleep. Sleep deprivation in rats with the DOW produces a series of dramatic physiological changes that culminate invariably in death after 2 to 3 weeks of sleep loss. Within the first 1 to 2 days, rats develop a syndrome characterized by an increase in food intake, energy expenditure, and heart rate, followed by a decrease in body weight and a decline in body and brain temperature. The sleep deprivation syndrome and its lethal consequences have also been observed after selective REM sleep deprivation, although the pathology associated with the loss of sleep takes longer to appear, the survival time is longer (4 to 5 weeks rather than 2 to 3 weeks), and body and brain temperature are not significantly decreased. Despite a long series of studies, the DOW sleep deprivation syndrome has not been fully explained nor is it clear why the animals die of sleep deprivation. It is evident, however, that the syndrome produced by the DOW is not unique and other methods of chronic sleep deprivation used in different animal species have produced similar effects. Also in dogs, rabbits, and to a lesser extent cats, sleep deprivation for several days causes an increase in food intake and in heart rate, weight loss, and eventually death. Even fruit flies, if prevented from sleeping for several days, die of sleep deprivation. Again, these findings suggest that there must be at least one potentially vital function for sleep and that this function is conserved across different animal species.

DREAMING

For a long time, sleep has been regarded as the annihilation of consciousness, save for the occasional dream that is remembered when one wakes up. Just as the old notion that the brain is silent during sleep has been disproved, so has the myth of cognitive death during sleep. Most subjects, if awakened and asked to report whatever may be going through their minds, report some kind of mental activity most of the time. Often such mental activity constitutes a dream, which can be defined as a complex, temporally unfolding, hallucinatory episode that occurs during

sleep.

The Relationship between Dreaming and Normal Waking Consciousness

Dream hallucinations are typically more vivid than waking images. Images are predominantly visual, although all modalities can be represented, and impossible motor activities, such as flying, may occur. Hearing speech or conversation is extremely frequent. Dreaming is generally delusional—events and characters in the dream are taken for real—and confabulatory—a dream involves making up a story. There is often disorientation, that is, uncertainty about space (where one is in the dream), about time (when the dream is taking place in personal history), and confusion about the gender, age, and identity of dream characters. However, dreams appear to run in real time, as there are good correlations between the subjective duration of a dream, the length of the dream report, and the time it would take to reenact a dream. Emotions are prominent in many dreams, especially fear and anxiety. While the self is almost always at the center of the dream, there is some reduction of reflective consciousness. Dreams have been described as single-minded, that is, missing the alternation of primary and reflective consciousness that characterize wakefulness. Self-monitoring, directed thinking, and volition are reduced, leading to an inability to analyze situations, to question assumptions, and to make appropriate decisions. The ability to form new memories is drastically impaired (dream amnesia). Not surprisingly, psychiatrists have pointed out similarities between some aspects of dreaming cognition and certain psychotic symptoms as well as delirium.

Despite these psychopathological traits, perhaps the most important fact about dreaming consciousness is how remarkably similar it can be to waking consciousness. That is, the sleeping brain, disconnected from the "real" world, is capable of generating an imagined world, a "virtual reality," which is fairly similar to the real one and is indeed experienced as equally real. Considering dreams from the perspective of the waking state, people are often intrigued by the bizarre quality of some dreams, especially morning dreams. Bizarre dreams may be more memorable; however, if collected systematically throughout the night, dreams are much more mundane—they are a faithful replica of waking life. As a rule, one can only dream what one can imagine, although dreams can be much more vivid, probably because of the reduced competition from external signals.

Conditions that affect people's brains during wakefulness extend to their dreams. For instance, if blind people can still construct visual images, then they have visual dreams, otherwise not. If a stroke abolishes the ability to perceive color during waking, then visual images as well as dreams become achromatopsic. Like the ability to form mental images, dreaming seems to depend most strongly on the integrity of cortical areas higher than the primary visual cortex. Similarly, somatosensory/motor and

audioverbal/motor imagery are normal in the dreams of hemiplegic and aphasic patients. Another example of the close connection between waking and dreaming cognitive competence comes from the study of children's dreams. Dream reports are rare and extremely short until 4 to 5 years of age—in agreement with the limited ability of children below that age to imagine and narrate complex stories with high emotional content. Finally, there is a good correlation between waking and dreaming mood, imaginativeness, and personality. Studies of dream content in psychiatric populations, especially in depression and PTSD, have generally been unremarkable. However, some intriguing findings have been reported, for example, that changes in dream content may anticipate changes in the course of the disorder.

Dreaming and Sleep Stages

It was initially suggested that full-fledged dreams could be elicited almost exclusively during REM sleep. Later studies have shown conclusively, however, that dreamlike mental activity can be elicited also from NREM sleep, especially at sleep onset and during the last part of the night—the times when NREM sleep is less deep. Up to 15 percent of subjects, however, never recall any mental content when awakened from NREM sleep. Moreover, typical REM dream reports can easily be distinguished from NREM ones by being on average much longer (up to seven times). When dream reports are longer, this suggests longer duration of the actual dream, but they may also be due to an increased density or bizarreness of the experienced scenes. Whether REM dreams are not only longer but also qualitatively different—more bizarre, hallucinatory, delusional, narrative, and emotional than NREM dreaming—remain controversial. So is the suggestion that NREM dreams may be "covert" REM dreams, due to an intrusion of REM characteristics in NREM sleep. There is no doubt, however, that typical dreams, as well as nightmares, can be experienced in certain phases of NREM sleep. Dreamlike mental activity having a hallucinatory–delusional character can also occur during quiet wakefulness, especially under conditions of reduced sensory input (daydreaming). Conversely, some wakinglike mental activity—non hallucinatory and nondelusional—is occasionally reported at sleep onset. A good example of a mixed state is lucid dreaming, where dreamers are aware that they are dreaming and, to some extent, can control the course of their dreams.

The initial equating of the cognitive state of dreaming with the physiological state of REM sleep was encouraged by the remarkable similarity between the EEG of REM sleep with that of conscious waking and by their equivalent differences from the NREM sleep EEG. It seemed natural to infer that the activated (low voltage, high frequency) EEG of waking and REM sleep would support vivid conscious experience, while the

deactivated (high voltage, low frequency) EEG of NREM sleep would not. While frequent dream reports at sleep onset could still be reconciled with the mixed EEG of stage N1 sleep, the presence of dreamlike experiences, although shorter, during NREM stages characterized by EEG slow waves seemed paradoxical. This paradox may be resolved by considering the time course of neural excitability during the slow oscillation of NREM sleep. It is now known that during the depolarized phase of the slow oscillation, neural activity is as intense as in waking or REM sleep and neurons are highly excitable. However, during sleep the depolarized phase is interrupted by a hyperpolarized phase during which neural activity ceases throughout the cerebral cortex and neurons are much less excitable. The inevitable occurrence of hyperpolarization (down state) after a period of depolarization (up state) occurs exclusively during NREM sleep and reflects a peculiar bistability of cortical circuits. This bistability is especially strong during SWS early in the night—the depolarized phase may not last longer than a second or so before it is followed by the hyperpolarized phase—and becomes less marked during stage N2 sleep late in the night, when longer periods of depolarization are possible. Thus, if dreamlike experiences occur during the depolarized phase, then they will be short and poor during SWS early in the night and become progressively longer and richer later in the night. Altogether, it would seem that the likelihood of conscious experience during sleep is related to moment-to-moment cortical activation, where activation should be understood as the readiness of cortical neurons to respond in a rich and differentiated manner to incoming signals.

Neural Correlates of Dreaming Consciousness

Recently, lesion and imaging studies have provided new insights into brain correlates of the characteristic differences between dreaming and waking consciousness. The most obvious difference is that dreaming consciousness is only marginally influenced by external stimuli, very few of which are incorporated into the dream narrative. During NREM sleep, partial functional disconnection is mediated by thalamic inhibition. During REM sleep, external stimuli more easily pass the thalamic gate and reach primary cortical areas, but they do not seem to influence higher cortical areas, as if the brain were not paying attention to them. The hallucinatory character of dreams is obviously facilitated by such sensory disconnection. Visual hallucinosis of dreams is indeed associated with increased activity in higher visual areas, while the primary visual cortex is less active, as indicated by PET studies. Whether the intense visual experience and scene changes characteristic of dreams are triggered primarily by phasic signals from the brainstem, areas involved in visual imagery, or both is an unresolved issue.

Another relevant difference between dreaming and waking

consciousness concerns the ability to reflect on oneself and one's experience, especially the ability to judge the verisimilitude of dreaming experience. Imaging studies indicate that the DLPFC, a brain region implicated in volitional control and self-monitoring, is less active in sleep compared to waking. Reduced activation of the DLPFC may also contribute to the disorientation and reduction of directed thinking and working memory observed in dreams and contribute to dreaming amnesia. Recent episodic memories are conspicuously absent in dreams, and memory for dreams is strikingly labile, unless one wakes up and rehearses the dream.

Dreams are also characterized by a high degree of emotional involvement, especially fear and anxiety. Correspondingly, imaging studies have revealed a marked activation of limbic and paralimbic structures such as the amygdala and anterior cingulate, insular, and medial orbitofrontal cortices during REM sleep. Altogether, cognitive activity during sleep provides a powerful indication of the extent to which fluctuating levels of several neuromodulators, whose dysfunction is implicated in several psychiatric disorders, can affect mental function in healthy subjects. Indeed, J. Allan Hobson has suggested that most aspects of dreaming cognition can be explained by considering the level of three processes—brain activation, input source (external or internal), and neuromodulation—across the sleep–wakefulness continuum.

Recently, high density EEG (hd-EEG) has also been used to characterize the neural correlates of conscious experience during sleep. These experiments have used a no-task, within-state paradigm in which sleeping subjects recorded with hd-EEG reported retrospectively whether they had been conscious or not. It was found that the neural correlates of consciousness in both NREM and REM sleep map to two localized hotspots: (1) a parieto-occipital hotspot characterized by reduced low-frequency EEG activity and (2) a frontal hotspot characterized by increased high-frequency EEG activity. It was also found that the frontal hotspot is involved in highly thought-like experiences, while the parieto-occipital hotspot is implicated in highly perceptual experiences, suggesting that the neural correlate of consciousness during sleep may be restricted to a single hotspot and change in location depending on the nature of the subject's dream. These studies have also shown that specific contents of a subject's dreams—such as faces and spatial setting—are associated with increased high-frequency EEG activity in specific cortical areas that closely match those engaged during waking perception of the same contents. These results provide strong evidence that dream reports reflect conscious experience that occurs during sleep, and are not simply confabulations that are produced upon awakening.

Theories of Dreaming

In what he considered his most important work—*The Interpretation of*

Dreams—Freud suggested that dreams provide disguised wish fulfillment and, if properly interpreted, would provide essential clues to the most profound determinants of psychic life, "the royal road to the unconscious." The systematic investigation of dreams has not provided much support for this notion.

A radically different suggestion was made by Hobson and McCarley based on their neurophysiological studies of REM–sleep-generating mechanisms in the brainstem. According to their activation-synthesis hypothesis, dreams were the forebrain's attempt to make sense of the random activation of thalamocortical networks by the upper brainstem, like the music produced by unmusical fingers wandering over the keys of a piano. Another suggestion, based on a comprehensive evaluation of dream reports, was made by Foulkes and others. According to this view, dreams reveal not so much the psychodynamic unconscious but instead the cognitive development, competence, and style of the dreamer, just as in waking cognition. This view carefully eschews the one-to-one conflation of dreams with REM sleep.

Along these lines, Solms has recently reviewed neuropsychological evidence and shown that the ability to dream depends not on the upper brainstem but on forebrain regions. In more than 100 cases of cessation of dreaming, the responsible lesion was either the parieto-temporo-occipital junction (uni- or bilaterally) or white matter near the orbitomesial prefrontal cortex (bilaterally). Despite the cessation of dreaming, REM sleep was almost always preserved. The parieto-temporo-occipital junction is important for mental imagery, spatial cognition (on the right side), and symbolic cognition (on the left side), all central features of dreaming. The white matter underlying the ventromesial prefrontal cortex that is necessary for dreaming is the same brain area that is targeted by modified prefrontal leucotomy. Its lesion reduces the positive symptoms of schizophrenia but produces adynamia. Chemical activation of these circuits by stimulants such as amphetamines, as well as by 3,4-dihydroxy-L-phenylalanine (L-DOPA), can produce hallucinations and delusions, suggesting that dreams may be facilitated by the activation of the mesolimbic and mesocortical dopaminergic systems.

THE FUNCTIONS OF SLEEP

The function of sleep is still a mystery. Sleep may have evolved from the circadian rest–activity cycle and may thus represent a default state. However, it is likely that sleep serves some more fundamental function. Otherwise, why should animals engage in prolonged periods of quiescence with increased arousal thresholds during which they cannot monitor potential dangers in the environment? Sleep seems to be universal. All animal species studied so far sleep, from invertebrates such as fruit flies and bees to birds and mammals, although mammals and birds generally

display more elaborate sleep cycles that include an alternation of NREM and REM sleep. Ridding or even reducing the body's need for sleep does not seem to be easy. Some marine mammals who may need continuous vigilance while flying or swimming, such as certain dolphins and porpoises, have developed alternating unihemispheric sleep rather than eliminating sleep altogether. Sleep deprivation in all species studied leads to an increase in sleep pressure, manifested as sleepiness, which rapidly becomes irresistible. Sleep pressure may lead to microsleeps or piecemeal sleep that result in cognitive impairment. Sleep deprivation is followed by longer and more intense sleep, suggesting a regulated need for sleep. If sleep is prevented for several weeks, then the consequences are fatal. Rats deprived of sleep for more than 2 to 3 weeks invariably die. Humans affected by a rare prion disorder called FFI also die, although it is not certain whether death is due to sleep deprivation per se.

Many hypotheses have been formulated about the functions of sleep. A good hypothesis should account for the following facts: (1) Sleep involves partial disconnection from the environment, which is potentially dangerous; sleep must therefore provide something not provided by quiet waking. (2) Sleep is accompanied by intense neural activity in most brain regions; an explanation is needed for the brain's activity in the absence of overt behavior. (3) Sleep seems to constitute a universal requirement, but its amount varies a great deal across different species; conservation of sleep across species must be accounted for, while accommodating key differences in sleep patterns. (4) In most species, sleep is prevalent early in life; sleep should perform important functions both during development and in the adult. (5) Sleep is often made up of both NREM and REM stages; it must be explained whether the sequence of the two stages is important and whether both stages serve a similar function. These various functions and characteristics of sleep have proved difficult to accommodate within one hypothesis; no single hypothesis has been proposed that can account for all these facts.

Sleep and Brain Restitution

Most hypotheses currently under investigation are concerned with a role of sleep in restoring some metabolic function or in serving neural plasticity. It is likely that sleep may preserve energy by enforcing body rest in animals with high metabolic rates. Indeed, animals and humans eat more during sleep deprivation. However, in humans the metabolic efficiency of sleep is only marginally better than that of quiet waking. Most bodily organs can obtain rest through quiet wakefulness, except for the brain; thus sleep may be especially important for the brain. Some molecular pathway or chemical in the brain may be depleted during waking and restored during sleep. For example, it has been suggested that sleep may favor the replenishment of glycogen in glial stores, although recent evidence shows that this may only

be true in a few brain regions. Alternatively, sleep could counteract synaptic fatigue by favoring the replenishment of calcium in presynaptic stores, the replenishment of glutamate vesicles, the resting of mitochondria, recycling of membranes, or transfer of proteins along axons and dendrites. While recent studies have revealed that molecular changes do occur between sleep and waking and after sleep deprivation, the significance of such changes is still unknown. If sleep restores something lost during waking, it is not known what it is.

Sleep, Learning, and Memory

A connection between sleep and memory was noted a long time ago. For example, after struggling to learn a new piece of music for much of the day, one will often play it better after a night of sleep. Recently, the importance of sleep and even naps occurring after certain types of declarative and nondeclarative learning has been documented in well-controlled experiments. Moreover, recent evidence suggests that sleep deprivation impairs not only the ability to consolidate previously learned tasks but also the capacity to acquire new information. Sleep could indeed offer a favorable context for certain aspects of learning and memory. The sensory disconnection associated with sleep reduces interference between ongoing activities and the consolidation of previously acquired memories. Moreover, sleep has been suggested to permit the repeated reactivation, in an off-line mode, of the neural circuits originally activated during a memorable experience. Studies using multielectrode recordings in animals and PET in humans have shown that brain areas or cells activated during waking are preferentially reactivated during subsequent sleep, although this reactivation also occurs during periods of quiet wakefulness following the learning experience. A further advantage of sleep could be that the relevant neural circuits can be reactivated in a spaced and interleaved fashion. This would favor the integration of new with old memories and would avoid catastrophic interference. The intense, high-frequency bursts of spontaneous neural activity that occur during sleep may be particularly important for both triggering molecular mechanisms of synaptic consolidation and enlarging the network of associations.

Many unknowns remain, however. Whether sleep may favor the consolidation of newly established memories or the maintenance of older ones is not clear. The molecular correlates of such processes are still not known. Molecular markers of memory acquisition are turned off during sleep, which may be advantageous given that the intense neural activity of sleep occurs while the animal is disconnected from the environment. Much of the early literature connecting sleep and memory was concerned with REM sleep. However, prolonged inhibition of REM sleep in humans through MAOIs does not seem to disrupt memory, nor does the complete disappearance of REM sleep after certain brainstem lesions. Perhaps the

most convincing evidence concerns the role of sleep in developmental plasticity. Recent experiments have shown that REM sleep deprivation in newborn rats and NREM sleep deprivation in kittens influence the activity-dependent development of visual system circuits. If sleep affects synaptic maturation and plasticity, then one can expect that sleep disturbances in early life may affect psychopathological development.

SLEEP AND PSYCHIATRY

Historically, psychiatrists have been interested in sleep and dreaming and their relationship with mental illness. From Sigmund Freud and Carl Jung in the latter part of the 19th century, the interpretation of dreams became an important tool in psychoanalytical psychotherapy. After the discovery of REM sleep in 1953, psychiatrists began to explore whether specific sleep abnormalities could be correlated with psychiatric disorders. One of the first questions addressed was whether schizophrenia might represent a disorder of REM sleep, given the similarities between dreaming and psychosis. Although psychosis could not be explained as an REM sleep disorder, both REM sleep and schizophrenia were later found to be associated with decreased activity in the DLPFC. Narcolepsy, however, turned out to be caused by the abnormal intrusion of REM sleep phenomena into wakefulness.

Research has focused more on sleep in depression than any other psychiatric aspects of sleep by far. Even prior to the discovery of REM sleep, people with depression were known to have disrupted sleep. Sleep EEG recordings in the 1950s and 1960s also showed that they had a relative loss of SWS in comparison to age-matched control subjects as well as specific changes in REM sleep, including reduced latency to REM sleep, greater proportion of REM sleep during the first third of the night, increased frequency of REMs during REM sleep (i.e., increased REM density), and increased percentage of sleep time spent in REM sleep. Although not every patient with depression shows changes in REM sleep and not every patient with short REM latency has depression, REM sleep abnormalities are one of the more robust biological markers for depression. Both reduced REM sleep latency and loss of SWS appear to be trait markers for depression in that they persist even during clinical remission and are found at higher rates in first-degree family members of people with depression. Although the mechanisms for sleep changes in depression are not fully understood, there is a convergence of data suggesting that sleep and mood are regulated by common systems. For example, the cholinergic–monoaminergic imbalance hypothesis of depression is consistent with the observed increase in REM sleep and reduction in SWS that would be caused by increased cholinergic activity. Individuals with depression show a heightened sensitivity to REM sleep induction by cholinergic drugs in comparison to nondepressed control subjects as well.

Sleep deprivation studies are even more suggestive of a functional relationship between sleep and mood. Total deprivation of a single night of sleep or even partial deprivation of sleep in the latter half of the night can have an immediate antidepressant response in many moderately to severely depressed individuals. Sleep loss can induce or perpetuate mania in bipolar patients, who may go for periods of several days with little or no sleep. In contrast, even a short bout of sleep can reverse the antidepressant effect of sleep deprivation, and prolonged sleep can induce depression in some individuals. Functional imaging studies have shown that sleep deprivation, like antidepressant drug therapy, normalizes the increased metabolic activity seen in the anterior cingulate gyrus in individuals with depression. Selective REM sleep deprivation has also been shown to have antidepressant effects, and it has been suggested that REM-sleep–suppressing antidepressants may act in part through their effects on sleep. REM sleep suppression, however, is not a requirement for antidepressant efficacy, because some agents such as bupropion and nefazodone appear to cause no significant reduction of REM sleep. More recent evidence suggests that both sleep deprivation and antidepressant medications may act by similar mechanisms, namely, by selectively upregulating genes involved in neural plasticity and synaptic potentiation.

Although therapeutic sleep deprivation may be beneficial for some cases of depression, the preponderance of evidence suggests that disturbed sleep is a risk factor for the development of psychiatric illness. People with insomnia as well as hypersomnia have higher rates of psychiatric disorders, particularly mood and anxiety disorders; in primary care settings, insomnia is more strongly associated with depression than with any other medical disorder. Individuals who have insomnia or even difficulty sleeping during times of stress are significantly more likely to develop depression in the future. A causal link between sleep and depression has not yet been established, but the fact that both sleep deprivation and depression are becoming more prevalent in society suggests that the relationship between the two must be clarified. In patients with psychiatric, insomnia is associated with a significantly increased risk factor for suicidal behaviors.

From an epidemiological perspective, although sleep disturbance is most strongly associated with psychiatric disorders, there are a number of other significant health-related correlates. People with insomnia have higher rates of other medical illnesses, use more health care services, have higher rates of absenteeism, accidents, and disability, and have poorer outcomes with some medical disorders, including cardiac disease. Whether treatment of sleep problems can prevent the development of any of these comorbidities remains to be seen.

Sleep abnormalities are also seen in virtually all other psychiatric disorders, particularly disturbances in sleep continuity, including prolonged latency to sleep onset, diminished efficiency of sleep, and

decreased amounts of total sleep. REM sleep abnormalities similar to depression have been described in some studies of patients with schizophrenia, alcoholism, eating disorders, and borderline personality disorder. Patients with panic disorder may have panic attacks arising from sleep, usually at the transition into SWS, which further emphasizes the biological etiology of this disorder. Recent studies using hd-EEG have demonstrated a loss of spindle activity and decrements in slow-wave activity in schizophrenic subjects, which may be related to a dysfunction in the thalamocortical network and associated with clinical features of sleep disruption and cognitive impairments.

An appreciation of sleep neurobiology is essential for clinicians, because psychiatric patients often have sleep problems associated with their illnesses and most psychiatric drugs have significant effects on sleep; these range from sedation caused by many antipsychotics, benzodiazepines, TCAs, and antiparkinsonian agents to sleep disturbance caused by MAOIs, SSRIs, and psychostimulants. Antidepressants may precipitate some sleep disorders, such as periodic movements in sleep and REM sleep behavior disorder, and abrupt withdrawal of REM-suppressing antidepressants including tricyclics, MAOIs, and SSRIs can lead to REM sleep rebound, characterized by increased duration and intensity of REM sleep and sleep disruption. Sleep apnea may be exacerbated by drugs that produce muscle relaxation (e.g., benzodiazepines or barbiturates) or weight gain (e.g., antipsychotics, antidepressants, and mood stabilizers).

Sleep is a revealing, though not yet transparent, window into the functional state of the human brain. The study of sleep has shed light on many aspects of consciousness and the workings of the human brain. Clarifying the functions and mechanisms of sleep will undoubtedly provide us with a greater understanding of psychiatric disorders and their treatments.

References

Abel T, Havekes R, Saletin JM, Walker MP. Sleep, plasticity and memory from molecules to whole-brain networks. *Curr Biol*. 2013;23(17):R774–R788.

Benca RM, Obermeyer WH, Thisted RA, Gillin JC. Sleep and psychiatric disorders: A meta-analysis. *Arch Gen Psychiatry*. 1992;49:651.

Borbely AA, Achermann P. Sleep homeostasis and models of sleep regulation. *J Biol Rhythms*. 1999;14:557.

Brown RE, Basheer R, McKenna JT, Strecker RE, McCarley RW. Control of sleep and wakefulness. *Physiol Rev*. 2012;92:1087.

Cipolli C, Ferrara M, De Gennaro L, Plazzi G. Beyond the neuropsychology of dreaming: Insights into the neural basis of dreaming with new techniques of sleep recording and analysis. *Sleep Medicine Reviews*. 2016, http://dx.doi.org/10.1016/j.smrv2016.07.005

Cirelli C. Cellular consequences of sleep deprivation in the brain. *Sleep Med Rev*. 2006;10:307.

Cirelli C, Tononi G. Is sleep essential? *PLoS Biol*. 2008;6(8):e216.

Cirelli C, Tononi G. Cortical development, EEG rhythms, and the sleep/wake cycle. *Biol Psychiatry*. 2015;77(12):1071–1078.

*Colten HR, Altevogt BM, eds. *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*. Washington, DC: The National Academies Press; 2006.

Consensus Conference Panel: Watson NF, Badr MS, Belenky G, Bliwise DL, Buxton OM, Buysse D, Dinges DF, Gangwisch J, Grandner MA, Kushida C, Malhotra RK, Martin JL, Patel SR, Quan SF, Tasali E. Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society on the Recommended Amount of Sleep for a Healthy Adult: Methodology and Discussion. *J Clin Sleep Med*. 2015;11(8):931–952.

Dang-Vu TT, Schabus M, Desseilles M, Sterpenich V, Bonjean M, Maquet P. Functional neuroimaging insights into the physiology of human sleep. *Sleep*. 2010:33(12):1589–1603.

*Diekelmann S, Born J. The memory function of sleep. *Nat Rev Neurosci*. 2010;11(2):114–126.

Ferrarelli F. Endophenotypes and biological markers of schizophrenia: from biological signs of illness to novel treatment targets. *Curr Pharm Des*. 2013;19(36):6462–6479.

Foster RG, Peirson SN, Wulff K, Winnebock E, Vetter C, Roenneberg T. Sleep and circadian rhythm disruption in social jetlag and mental illness. *Prog Mol Biol Transl Sci*. 2013;119:325–346.

Fuller PM, Gooley JJ, Saper CB. Neurobiology of the sleep-wake cycle: Sleep architecture, circadian regulation, and regulatory feedback. *J Biol Rhythms*. 2006;21:482.

Gamble KL, Berry R, Frank SJ, Young ME. Circadian clock control of endocrine factors. *Nat Rev Endocrinol*. 2014;10:466.

Goel N, Basner M, Rao H, Dinges DF. Circadian rhythms, sleep deprivation and human performance. *Prog Mol Biol Transl Sci*. 2013;119:155–190.

Hirshkowitz M, Whiton K, Albert SM, et al. National Sleep Foundation's sleep time duration recommendations: methodology and results summary, Sleep Health. 2015: http://dx.doi.org/10.1016/j.sleh.2014.12.010

Iber C, Ancoli-Israel S, Chesson AL, Quan SF. *The AASM Scoring Manual for the Scoring of Sleep and Associated Events*. Westchester, IL: American Academy of Sleep Medicine; 2007.

*Kryger MH, Roth T, Dement WC. *Principles and Practices of Sleep Medicine*. Philadelphia, PA: WB Saunders; 2017.

Kurina LM, McClintock MK, Chen JH, et al. Sleep duration and all-cause mortality: a critical review of measurement and associations. *Ann Epidemiol*. 2013;23:361.

Kurth S, Ringli M, Geiger A, LeBourgeois M, Jenni OG, Huber R. Mapping of cortical activity in the first two decades of life: A high-density sleep electroencephalogram study. *J Neurosci*. 2010;30:13211.

*Llinás RR, Steriade M. Bursting of neurons and states of vigilance. *J Neurophysiol*. 2006;95:3297.

Mignot E. Why we sleep: the temporal organization of recovery. *PLoS Biol*. 2008;29:6.

Neske GT. The slow oscillation in cortical and thalamic networks: Mechanisms and functions. *Front Neural Circuits*. 2016;9:88. Doi: 10.3389/fncir.2015.00088.

Palagini L, Baglioni C, Ciapparelli A, Gemignani A, Riemann D. REM sleep dysregulation in depression: state of the art. *Sleep Med Rev*. 2013;17(5):377–390.

Saper CB, Fuller PM, Pedersen NP, Lu J, Scammell TE. Sleep state switching. *Neuron*. 2010:68(6):1023–1042.

*Silber MH, Ancoli-Israel S, Bonnet MH. The visual scoring of sleep in adults. *J Clin Sleep Med*. 2007;3:5.

Tononi G, Cirelli C. Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. *Neuron*. 2014;81:12.

Tononi G, Koch C. The neural correlates of consciousness: an update. *Ann N Y Acad Sci*. 2008;1124:239–261.

▲ 1.23 Basic Science of Appetite

NORI GEARY, PH.D., AND TIMOTHY H. MORAN, PH.D.

INTRODUCTION

Appetite comprises the complexly determined and highly individuated phenomena of eating and the subjective phenomena surrounding it. Basic research in appetite spans experimental psychology, gastrointestinal (GI) physiology, metabolism, neuroendocrinology, and genetics. This chapter emphasizes the core physiology of appetite: eating behavior, the sensory signals that control it, the brain mechanisms that integrate these signals, and some of the system's important modulators. Appetite is an important research problem. Despite intense research efforts, both normal and disordered eating remain poorly understood. As a result, only minimally effective treatment strategies are presently available for disordered eating, whether presenting in psychiatric eating disorders or simple obesity.

Appetite Is a Network Function

According to current understanding, eating is controlled by a widely distributed neural network in the brain that integrates oropharyngeal food stimuli, GI signals, metabolic signals (including signals related to body adiposity), numerous other internal stimuli (e.g., circadian signals, stress responses, reproductive-axis function), and external stimuli with learned meanings. Furthermore, the system displays structural and functional plasticity throughout the life span. This perspective emphasizes the multiple factors affecting eating, their varied effects on information processing in the brain, and their complex synergistic and antagonistic interactions. It contrasts with the traditional view that only one or a few signals, for example, blood glucose levels, acting on a few circumscribed hypothalamic sites produce a unitary control of appetite.

The view that appetite is a neural-network function has several important implications. (1) There may be no brain locus where all neural signals controlling appetite converge. (2) It is unlikely that a neural basis for a unitary motivational process for appetite will be found. (3) Controls of eating that are elaborated in different parts of the network may not always operate in a coordinated manner. For example, eating might simultaneously be inhibited by some controls (e.g., homeostatic signals related to metabolic fuel availability or GI nutrient sensing) and stimulated by others (e.g., oral sensory hedonics). For this reason, patterns of within-meal eating rates or patterns of spontaneous meal sizes and frequencies may change in the absence of changes in the total amount eaten. (4) The existence of such partially autonomous controls may contribute to the

modest effects of pharmacological therapies for disordered eating that are based on manipulation of single signaling molecules.

The Meal as the Unit of Analysis

The meal is the appropriate unit of analysis for both functional and physiological analyses of appetite. The meal is the functional unit of eating. (1) The timing, size, and content of meals provide a complete description of the organism's response to the basic challenges of nutrition, namely, what, when, and how much to eat. (2) Meal onset, the rate of eating during meals, and meal termination provide concrete metrics for the study of appetite. For example, accounts of hunger cannot succeed if they do not explain the timing of meal onset. (3) Disordered meals are the behavioral phenomenology of disordered eating. Abnormally large meals are the defining behavioral change in bulimia nervosa and in binge-eating disorder, and abnormally small meals are the defining behavioral change in anorexia nervosa.

The meal is also the biological unit of eating. Humans and most other vertebrates organize eating into discrete bouts, or meals, separated by intervals of noneating, or intermeal intervals, with species-specific variations related to circadian, seasonal, and other influences. Meal-ending satiation is an active process: meals are followed by a period of sedation or sleep, which is clearest in animals and infants. These characteristics are undoubtedly products of both natural selection and cultural evolution, shaped by such influences as the digestive consequences of cooked rather than raw food and the logistics of food acquisition and preparation.

Behavioral Neuroscience of Eating

The behavioral neuroscience of eating follows three general approaches. (1) *Input approaches* are aimed at understanding the environmental and experiential contingencies and the exteroceptive stimuli and interoceptive sensory signals that influence eating. Exteroceptive stimuli include stimuli whose meaning is learned as well as a few innately meaningful stimuli, such as the hedonic significance of sweet taste. Important interoceptive stimuli include circulating levels of hormones and metabolites. (2) *Network approaches* focus on the central neural networks that integrate inputs and produce the motivational processes that drive eating. These approaches are most meaningful when meals are studied, rather than summary measures, for example, calories ingested per day, and so forth. (3) *Motor approaches* include efforts to understand the neural control of the movements of eating, often called the microstructure of eating. Food is licked, sucked, bitten, or masticated prior to being moved by lingual and palatal movements to the oropharynx and swallowed. Analysis of these movements has the potential to track their neural control through the lower motor

neurons of cranial nerves V, VII, IX, X, and XII into (i) the local circuits in the motor nuclei of these nerves, (ii) the central pattern generators in the hindbrain that project onto these nuclei, (iii) upstream integratory interneuronal networks, and (iv) ultimately to sensory and brain-autonomous inputs, such as memories, circadian rhythms, and so forth.

Subjective Experience of Eating

The subjective experiences associated with meals are scientifically accessible in humans and provide important insights into the expression of normal and disordered eating. Figure 1.23–1 shows examples. Optimal methods to measure subjective aspects of eating continue to evolve. Because a convincing taxonomy of the natural categories of appetite is still lacking, subjective measures of appetite are also most reliable when based on operational definitions that can be linked to meals.

SIGNALS FOR MEAL INITIATION

Hunger Processes and Learning

The patterns of neural activity leading to food seeking and meal initiation are conceptualized as hunger motivation. Three distinct motivational processes are thought to be involved: one negative-reinforcement process and two positive-reinforcement processes, one related to satiation and the other to orosensory hedonics (Fig. 1.23–2). Each of these suffices to stimulate eating, and eating leads to negative feedbacks that reduce the hunger. However, only negative-reinforcement hunger involves an aversive state that one wishes to escape. In contrast, the expectation of satiation or of a positive hedonic state is usually pleasant. Most people rarely experience the aversive state of strong deprivation-induced hunger (among those who do are the poor, strict dieters, and people with anorexia nervosas). Although the GI tract is relatively empty after overnight fasts, people generally do not report intense hunger in the morning, and breakfast is usually the smallest meal of the day. Thus, normal adult hunger probably represents a lifetime of positive-reinforcement learning, with innate "homeostatic" mechanisms not usually playing important roles.

Physiological investigations have suggested the utility of subdividing the brain process of hedonics into a motivational aspect, known as incentive motivation or wanting, and a subjective aspect, known as liking. This is best considered a provisional categorization, as it makes little sense to conceptualize a hedonic experience that lacks motivational power.

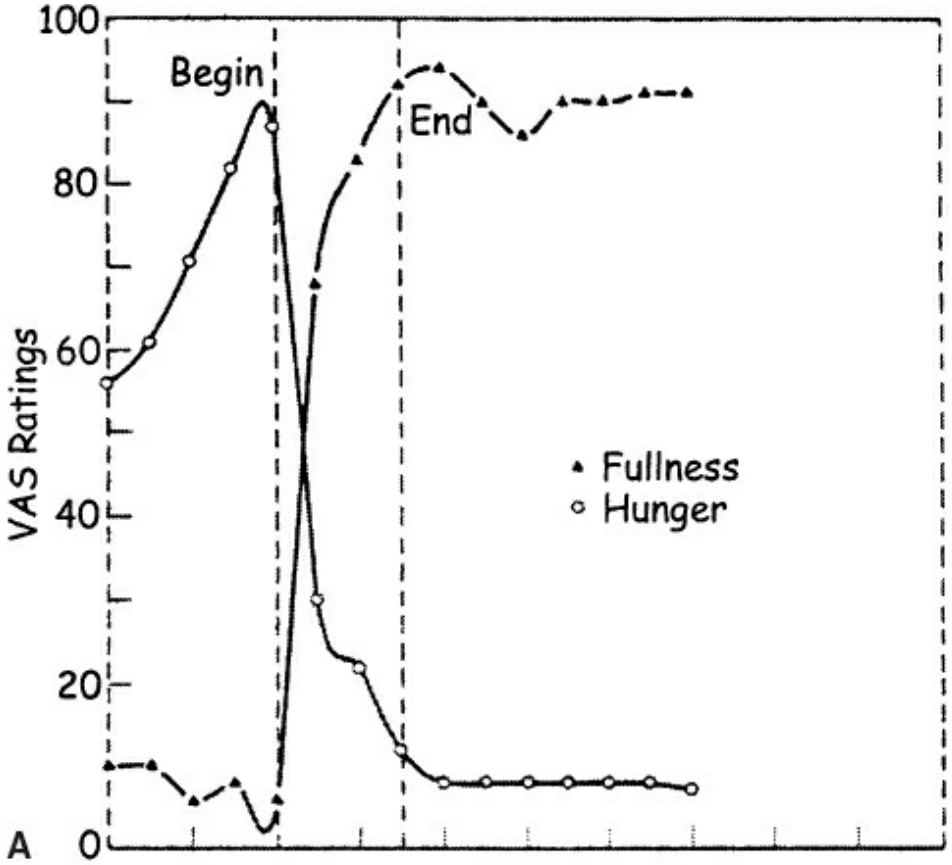
100
80
60
40
20
0
VAS Ratings
Begin
End
Fullness
Hunger
A

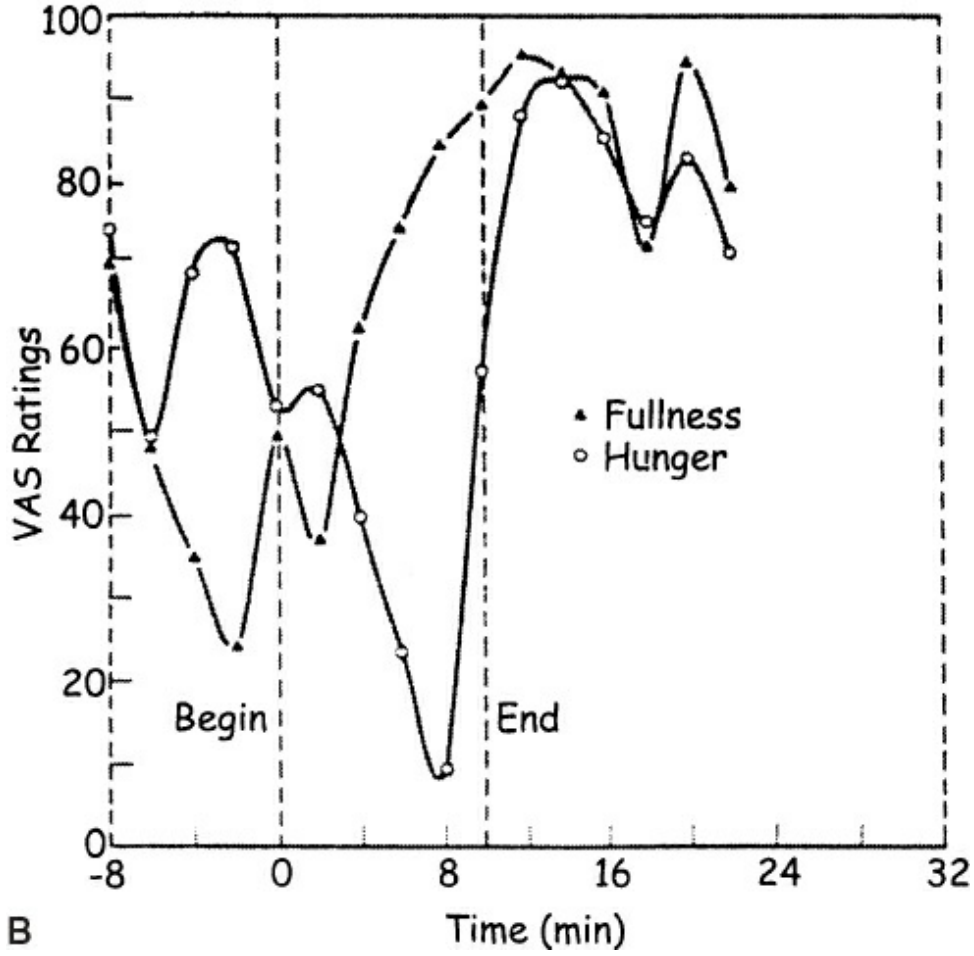
100
80
60
40
20
0
VAS Ratings
Fullness
Hunger
Begin
End
-8
0
8
16
24
32
Time (min)
B

FIGURE 1.23–1. Example of the use of the visual analog scale (VAS) (essentially a 100-mm line that the subject marks with a pencil to indicate the momentary intensity of a percept; the line is anchored with descriptors such as "most possible" or "not at all") to measure meal-related changes in hunger and fullness in (**A**) a healthy woman and (**B**) a woman with anorexia nervosa. In the healthy woman, hunger is high before the meal, fullness is low before the meal, and the two percepts change in a reciprocal fashion before, during, and after eating; this is typical of most normal subjects. In contrast, in this patient with anorexia nervosa, hunger and fullness change irregularly before, during, and after the meal and are often not reciprocally related; other patients with anorexia nervosa showed a variety of responses, some more and some less normal than these. Note that the figure does not show many experiences associated with meals, such as pleasure and tranquilization. The challenge for this research is to disentangle the various psychological and physiological influences producing these percepts. (From Owen WP, Halmi KA, Gibbs J, Smith GP. Satiety responses in eating disorders. *J Psychiatr Res*. 1985;19:279, with permission.)

Circadian Rhythms

Circadian rhythms importantly affect hunger, and may explain low morning hunger (Fig. 1.23–3). Such circadian rhythms presumably originate in the circadian pacemaker in the SCN of the hypothalamus. This area has extensive neural connections with arcuate nucleus of the hypothalamus (ARC), dorsomedial nucleus of the hypothalamus (DMN), LHA, and other areas through which it could modulate the neural and endocrine signals controlling eating, as reviewed below.

Metabolic Signals

Stimuli resulting from decreased glucose or fatty-acid utilization can elicit hunger and meal initiation. These effects are clear in extreme conditions, such as during extended nutrient depletion, abrupt shifts in the macronutrient composition of the diet, and biochemical hypoglycemia. It seems unlikely that they operate under usual physiological conditions.

Small transient declines in plasma glucose sometimes occur prior to spontaneous meals in rats and humans and may be a normally operating hunger stimulus. These are too small (~10 percent) to influence cellular glucose availability. Rather, the temporal dynamics of the decline may signal hunger.

Ghrelin

Ghrelin is a peptide hormone synthesized and released primarily by enteroendocrine cells in the stomach and proximal small intestine. Ghrelin is secreted in an inactive form and metabolized to the biologically active, acylated form, octanoyl- or acyl-ghrelin, by the enzyme ghrelin O-acyltransferase. Plasma levels are high in the morning, decrease after meals, and increase during intermeal intervals. The mechanism that stimulates ghrelin secretion is unknown. The gradual reduction in GI contents in the hours after meals apparently releases a tonic stimulatory

control of ghrelin secretion from inhibition. Inhibition is reinstated when nutrients reach the small intestine during and after meals. Carbohydrates and proteins inhibit ghrelin secretion more than isoenergetic lipid loads.

Ghrelin became a candidate endocrine hunger signal when it was discovered that ghrelin infusion decreased the latency to eat and increased meal size in rats and mice; indeed, repeated ghrelin administration induced obesity. Animal experiments suggest that ghrelin may activate both negative- and positive-reinforcement hunger: (i) systemic ghrelin administration stimulated rats' ingestion of palatable foods more than normal chow and led to the formation of conditioned place preferences, suggesting that ghrelin stimulates positive-reinforcement hunger, and (ii) ghrelin injected into the third cerebral ventricle failed to affect rats' initial rate of eating during the meal, which is an index of hedonic hunger, but did increase the rate of responding in a progressive-ratio test, an index of motivation, which together suggest that ghrelin elicited negative-reinforcement hunger. These results also suggest that ghrelin acts in different brain loci to stimulate eating. Among the sites of ghrelin receptors suggested to mediate hunger are the ARC, the nucleus of the solitary tract (NTS), the VTA, and nucleus accumbens (NAc); these areas are discussed further below.

Evaluating Physiological Signaling. Endocrinology and neuroscience have evolved criteria for identifying normal physiological functions of chemical signals. Table 1.23–1 presents a current set. Criteria 1 and 2 address the plausibility of the candidate hormonal signal to control a particular function; criteria 3 to 5 concern its sufficiency; and criterion 6, its necessity. For candidate endocrine signals such as ghrelin, plasma measurements and intravenous infusions suffice to test these criteria. Some of the GI hormones that are discussed also act locally, in paracrine, autocrine, and so forth, modes. For these, measurements and administration in the local environment, that is, the intestinal lamina propria, are required. Molecules acting as neurotransmitters or neuromodulators require measurements and manipulations in the synaptic cleft.

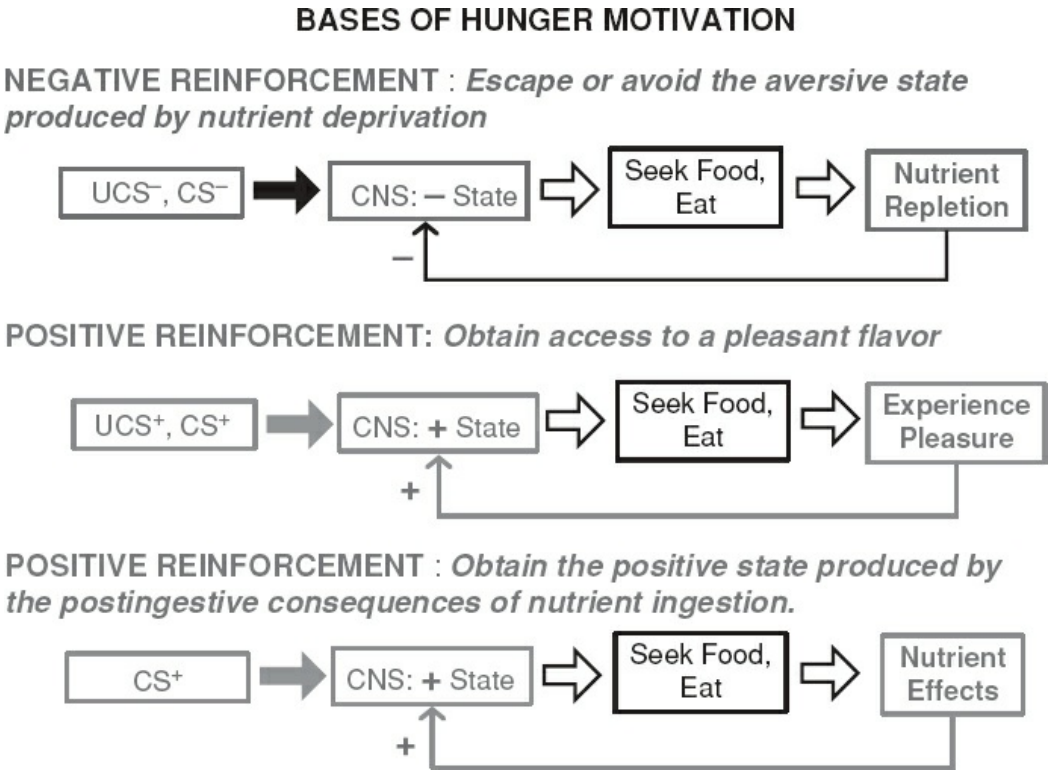

FIGURE 1.23–2. Hunger-motivation constructs developed to explain food seeking, meal initiation, and eating. **Top:** Negative-reinforcement hunger is the classical construct of hunger generated by nutrient depletion, or homeostatic need. Depletion of metabolic energy supplies or of certain micronutrients such as sodium lead to aversive CNS states (UCS^-), that are escaped by ingestion of the appropriate commodity. Such aversive states act as unconditioned negative reinforcers, such that with experience, initially arbitrary stimuli can elicit a similar state; these CS^- permit avoidance as well as escape. **Middle:** Positive-reinforcement hunger generated by flavor hedonics. Some flavors, such as sweet flavor, lead to positive CNS states (UCS^+) that elicit approach and ingestion. Other flavor stimuli (CS^+) come to elicit similar states via a variety of learning mechanisms, often not directly associated with UCS; even initially aversive flavors such as chili pepper can become such CS^+. Such CS^+ are likely to be the most important hunger motivation mechanism in affluent humans. **Bottom:** Positive-reinforcement hunger generated by stimuli associated with postingestive GI or metabolic effects produced by ingested food. These effects generate positive CNS states that act as unconditioned positive reinforcers, which, with experience can lead to the development of conditioned satiating stimuli (CS^+) that can elicit approach and ingestion. This process has been referred to as *appetition* because it is not clear that the GI and metabolic stimuli that generate it are identical to those that generate satiation. Note (i) that negative-reinforcement hunger is a negative feedback process (*solid black arrows*), whereas the positive-reinforcement hungers are positive-feedback processes (*gray arrows*); and (ii) that the conditioned stimuli in all three types of hunger motivation can come to motivate eating without deprivation, which is presumably a cause of overeating.

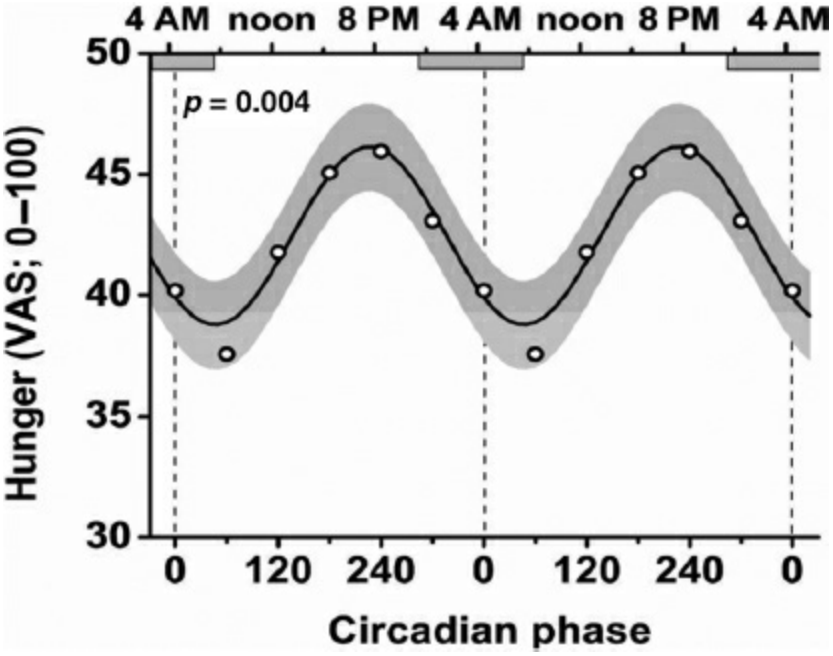

FIGURE 1.23–3. The circadian rhythm of hunger. Hunger was rated using visual analog scales (VAS; 0 = not at all hungry, 100 = extremely hungry) by healthy, normal weight men who underwent a 13-day "forced desynchronization" protocol that balanced all behaviors, including meals and sleep periods, evenly across the endogenous circadian cycle under constant dim light. Data are double plotted; points are mean ratings in 4-hour bins; the black line is a cosine function ± 95 percent confidence interval (*gray area*) fit to the rating data; and horizontal gray rectangles indicate average habitual scheduled sleep episodes. Note that hunger minima occurred in the biological morning and hunger maxima in the evening and that there was a 17 percent maximum–minimum difference. (From Shear FA, Morris CJ, Shea SA. The internal circadian clock increases hunger and appetite in the evening independent of food intake and other behaviors. *Obesity (Silver Spring)*. 2013;21:421, with permission.)

Table 1.23–1.
Empirical Criteria to Assess Physiological Status of Chemical Signals[a] in Meal-Related Functions

1. Concentrations of the molecule change at the site of action in a pattern consistent with the effect.[b]
2. Cognate receptors for the molecule are expressed at its site(s) of action.
3. Exogenous administration of the molecule in amounts duplicating the meal-related changes in endogenous patterns at the site of action produces the effect.
4. Administration of secretagogues for the molecule produces effects similar to the effect of the hormone.
5. The molecule's effect occurs in the absence of abnormal behavioral, physiological, or subjective effects.
6. Administration of selective agonists and antagonists of the molecule's receptors produce effects that are consistent with their receptor pharmacologies.[c,d]

[a]Signaling molecules may be hormones, cytokines, neurotransmitters, and so forth.

[b]At a minimum, the change in concentrations of the proposed signal should precede the effect.

[c]For example, administration of specific and potent receptor antagonists should delay or reduce eating in the case of a hunger signal or increase eating in the case of a satiation signal.

[d]We do not include phenotypic evaluation of global transgenic or spontaneous genetic loss-of-function

models in this criterion. These are valuable research tools, but complications due to developmental compensatory effects, pleiotropic actions, and species differences preclude their use as a "necessity" criterion for physiological function. Rapidly inducible, tissue-specific reductions in gene function, however, may complement the use of receptor antagonists in establishing necessity.

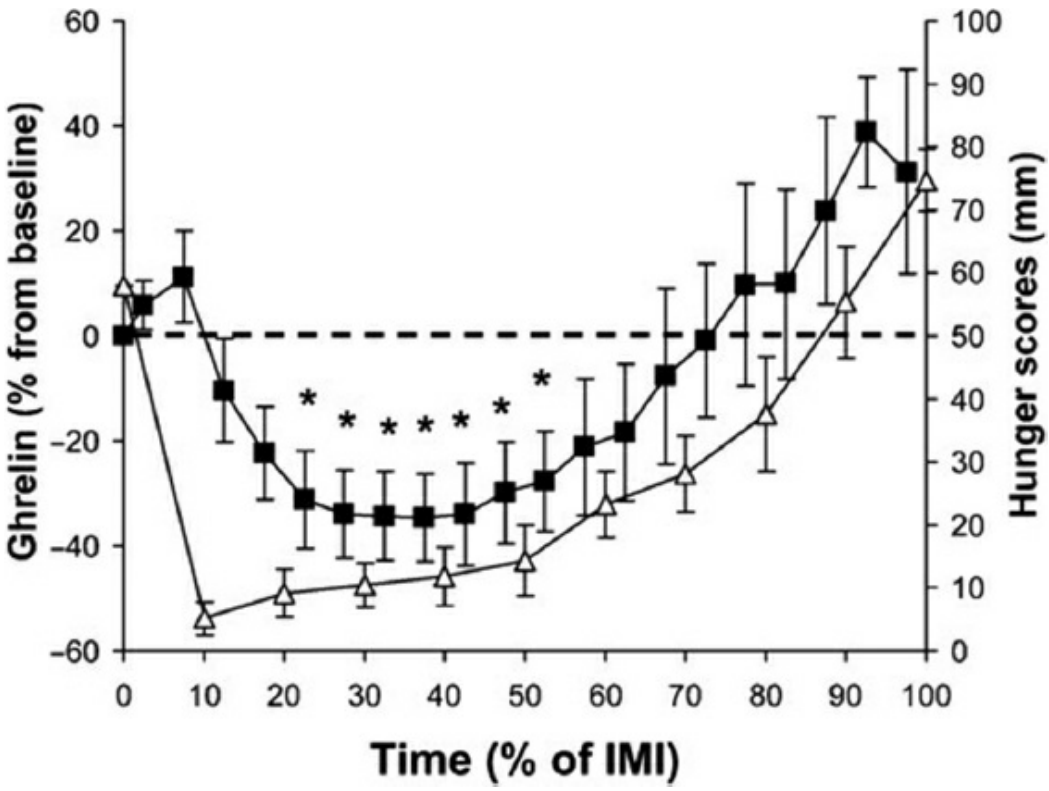

FIGURE 1.23–4. Close association of hunger scores (*open triangles,* rated on a 100-mm visual analog scale) and plasma ghrelin concentration (*filled squares,* total ghrelin-like immunoreactivity) during the interval between lunch (begun at time 0) and a freely requested dinner (time 100 percent). Subjects were six healthy, normal-weight males. Mean lunch size was 800 kcal, mean lunch duration was 20 minutes, and mean intermeal interval was 359 minutes. Data are mean ± standard error of the mean. Baselines are values at lunch onset. *Significantly different from time 0. (From Cummings DE, Fayo RS, Marmonier C, Aubert R, Chapelot D. Plasma ghrelin levels and hunger scores in humans initiating meals voluntarily without time- and food-related cues. *Am J Physiol*. 2004;287:E297, with permission.)

At present, ghrelin fulfills some, but not all of these criteria. It is a plausible hunger signal (criteria 1 and 2) because plasma ghrelin levels increase before meals, increase during food deprivation, and decrease after meals in animals and humans (Fig. 1.23–4) and because infusion of ghrelin to brain areas that contain ghrelin receptors simulates eating in animals. Such effects can be obtained without side effects (criterion 5). Two cardinal criteria, however, have not yet been met in humans: it is not known whether (i) ghrelin infusions that mimic the changes in ghrelin produced by premeal increases in endogenous ghrelin are sufficient to stimulate eating (criterion 3), and (ii) antagonism of endogenous ghrelin signaling before meals is sufficient to delay meal onset (criterion 6), although this

has been shown in animals.

MEAL MAINTENANCE

Flavor Hedonics

The maintenance of eating during the meal is attributed to the ongoing influence of the hunger processes reviewed above together with the eating-stimulatory effect of positive flavor (or orosensory) hedonics. Flavor stimuli elicited during eating arise from olfactory (cranial nerve I), gustatory (VII, IX, X), and somatosensory (V) receptors in the oronasopharynx. Flavor stimuli contribute to (i) detection and selection of foods; (ii) stimulation or inhibition of eating, mainly through positive or negative hedonic processes, presumably via mechanisms overlapping with those producing hedonic hunger discussed above; and (iii) associative learning processes, both as conditioned stimuli (CS) and, in the case of sweet and fat flavors, unconditioned or reinforcing stimuli. A number of well-validated methods can be used to assess flavor hedonics.

Psychophysics. Subjective aspects of flavor hedonics are typically measured with visual analog scales (VAS). One common scheme is to instruct subjects to indicate how pleasant a flavor is by marking a 100-mm line anchored at the ends as "not at all" and "extremely." Most subjects produce reliable and coherent VAS ratings, but patients with anorexia nervosa often do not. In addition, it is now clear that VAS ratings are not appropriate for between-subject comparisons because different categories of subjects, such as men versus women or obese versus nonobese subjects, use VAS differently. The generalized labeled-magnitude scale (gLMS), shown in Figure 1.23–5, avoids this.

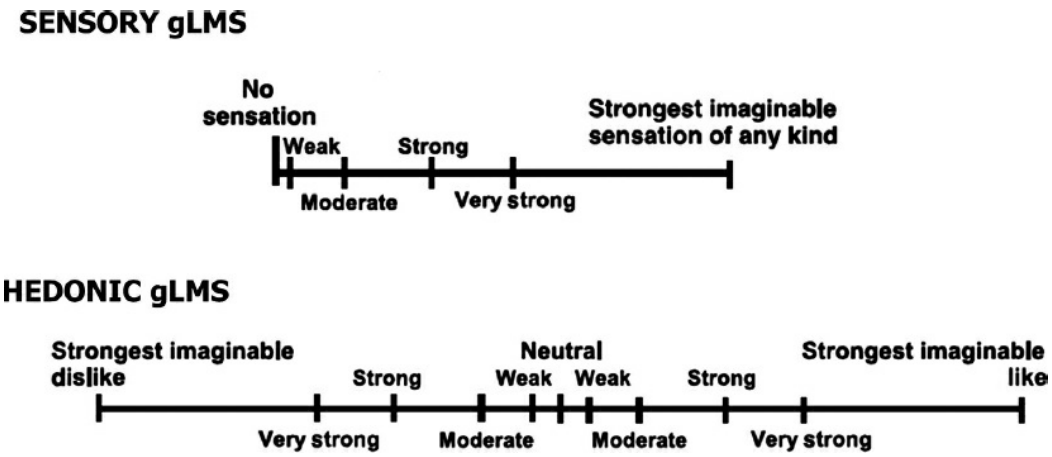

FIGURE 1.23–5. Generalized labeled-magnitude scales (gLMS) for estimations of the sensory intensity (**top**) and hedonic intensity (**bottom**) of subjective experience. These scales are anchored by maxima and minima of sensory experiences of all kinds; for example, a subject might consider the brightness of the sun to be the maximum possible sensory intensity and rate the sweetness of a food against that standard. Similarly, a subject might consider the pleasure of orgasm to be the maximum possible positive

hedonic experience and rate the pleasantness of a food against that standard. Because subjects rate their experiences relative to a standard (e.g., brightness of the sun) that is independent of the comparison of interest (e.g., flavor intensity), valid comparison can be made across groups, for example obese and nonobese subjects, who may experience flavors differently. (From Kalva JJ, Sims CA, Puentes LA, Snyder DJ, Bartoshuk LM. Comparison of the hedonic general labeled magnitude scale with the hedonic 9-point scale. *J Food Sci*. 2014;79:S238.)

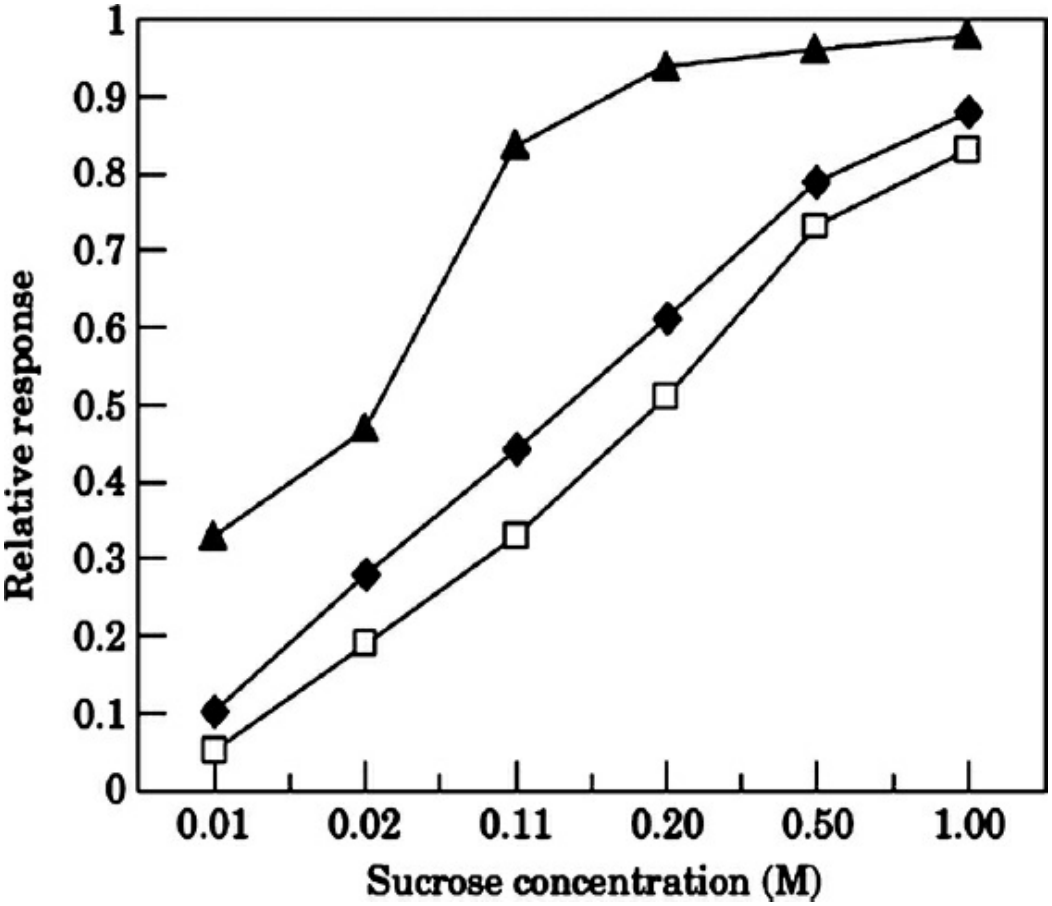

FIGURE 1.23–6. Flavor hedonics stimulates ingestion in rats. Rats' responses (expressed relative to maximal) to varying sucrose concentration in 30-minute sham-feeding tests (*open squares*), in 30-second brief-access licking tests (*filled triangles*), and from electrophysiological recordings from the greater superficial petrosal nerve (a branch of cranial nerve VII, *filled diamonds*). (From Smith JC. The history of the "Davis Rig." *Appetite*. 2001;36:93, with permission.)

Sham and Brief-Access Eating. Sham-eating and brief-access eating tests are designed to isolate the orosensory controls of eating in the absence of postingestive consequences. An elegant method for brief-access tests in rats involves measurement of licks of different fluids during repeated 10-second presentation periods, with presentation order randomized so that although the fluids enter the stomach normally, their mixture prevents the association of particular flavors with postingestive effects. Sham eating eliminates postingestive food stimuli altogether. It is done in rats with a surgically implanted gastric cannula, which when opened causes liquid food to drain out rather than accumulating in the stomach or enter the

intestines in appreciable amounts. Rats' brief-access and sham intakes of sweet solutions or of oil emulsions increase linearly over a wide range of concentrations, indicating the potent effect of flavor hedonics on eating (Fig. 1.23–6). Sham eating is done in humans simply by having subjects take food into the mouth and expel it without swallowing. This method has been used successfully with women with anorexia nervosa or bulimia nervosa. Interestingly, it revealed a greater propensity to sham eat in women with bulimia nervosa than in women without a history of eating disorders.

Eating in the Absence of Hunger. Eating in the absence of hunger (EAH) denotes eating driven by hedonics after satiation is achieved with less preferred foods (Fig. 1.23–7). Pediatric EAH, assessed by offering snack foods after a standard meal, is strongly associated with overweight and obesity. Note that EAH refers only to the presumed absence of negative-reinforcement hunger; thus, "hedonic hunger" would be less ambiguous.

Innate and Learned Hedonics

Although some preferences (sweet) and aversions (bitter or sour) for basic gustatory stimuli appear to be innate, preferences and aversions for the vast majority of flavors are predominately learned. Foods' GI and metabolic consequences can reinforce such learning. This occurs in conditioned satiations, conditioned aversions (including the marked aversions for flavors associated with acute upper GI illness), and "specific hungers" (preferences for flavors associated with foods containing some, but by no means all, vitamins or minerals that can be learned during states of nutritional deficiency). Note that in these situations, with the exception of sodium appetite during sodium depletion, it is the discriminative, that is, nonhedonic, aspects of the flavors that are important, although increased or decreased hedonic responses are part of what is learned. The majority of human flavor preferences seem to arise from the emotional, cognitive, and cultural associations attached to various foods, independent of their nutritional or physiological properties. Most individuals endorse especially strong preferences, or *cravings,* for particular foods. Because they dramatically affect an individual's tonic level of food intake and their adherence to therapeutic dietary regimens, the origins and plasticity of human food preferences are important research challenges.

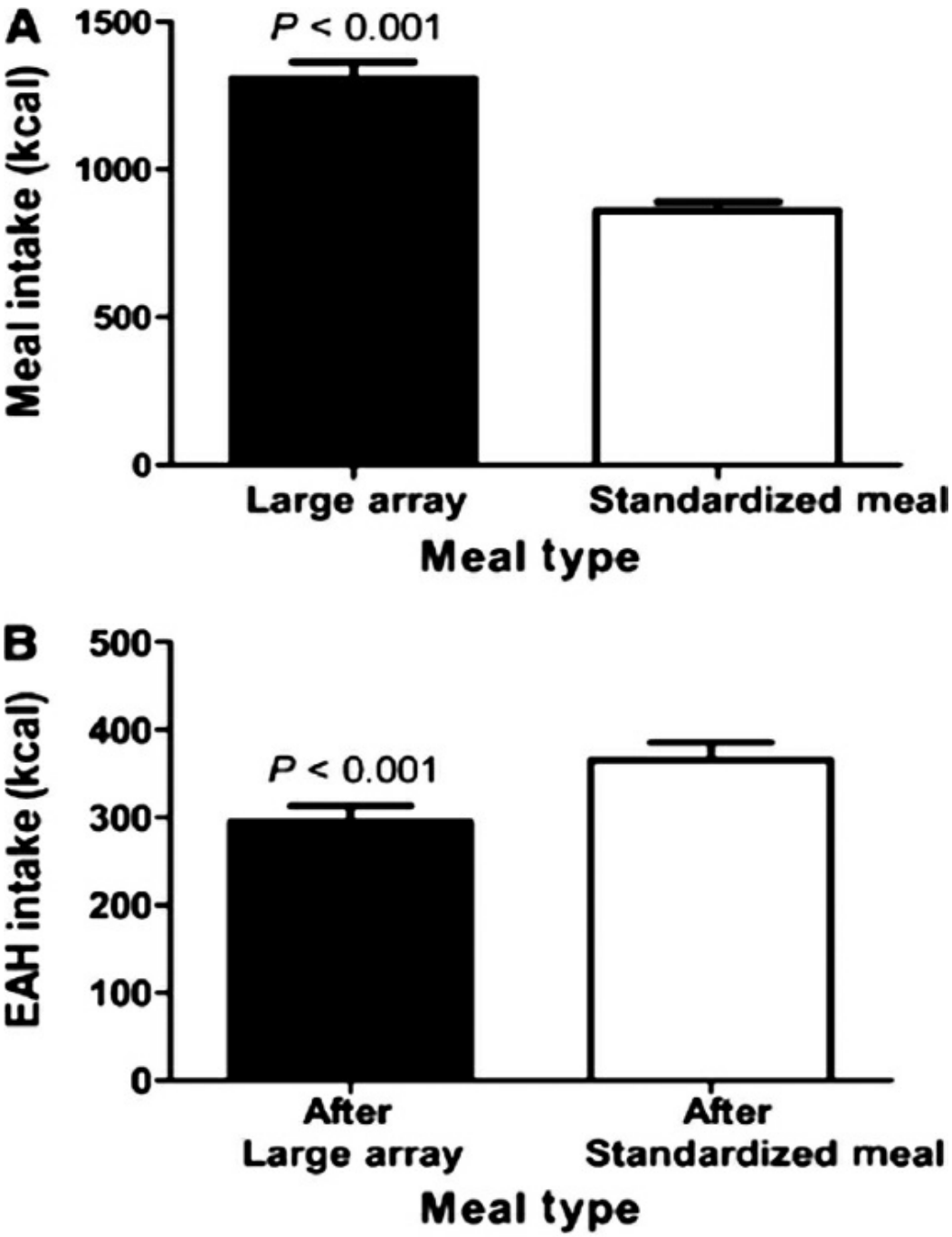

FIGURE 1.23–7. Two demonstrations of "eating in the absence of hunger." **A:** Adolescent subjects were first offered a midday meal consisting of either a large array of lunch foods in a buffet setting or a single standardized meal consisting of a subset of foods from the large-array meal (chicken nuggets, grapes, baby carrots, tortilla chips, cookies, 2 percent fat milk, etc.) served in excess. Subjects were instructed to eat until they were no longer hungry and to open the room when they were done. Note the effect of variety: subjects ate significantly larger meals in the large-array condition. **B:** Thirty minutes after their meals, subjects were offered unexpectedly an array of novel, highly palatable snack food items in generous portions. Note that they ate 300 to 400 kcal more regardless of the type of first course served, although there was a significant difference in the conditions. (From Shomaker LB, Tanofsky-Kraff M, Zocca JM, et al. Eating in the absence of hunger in adolescents: intake after a large-array meal compared with that after a standardized meal. *Am J Clin Nutr*. 2010;92:697, with permission.)

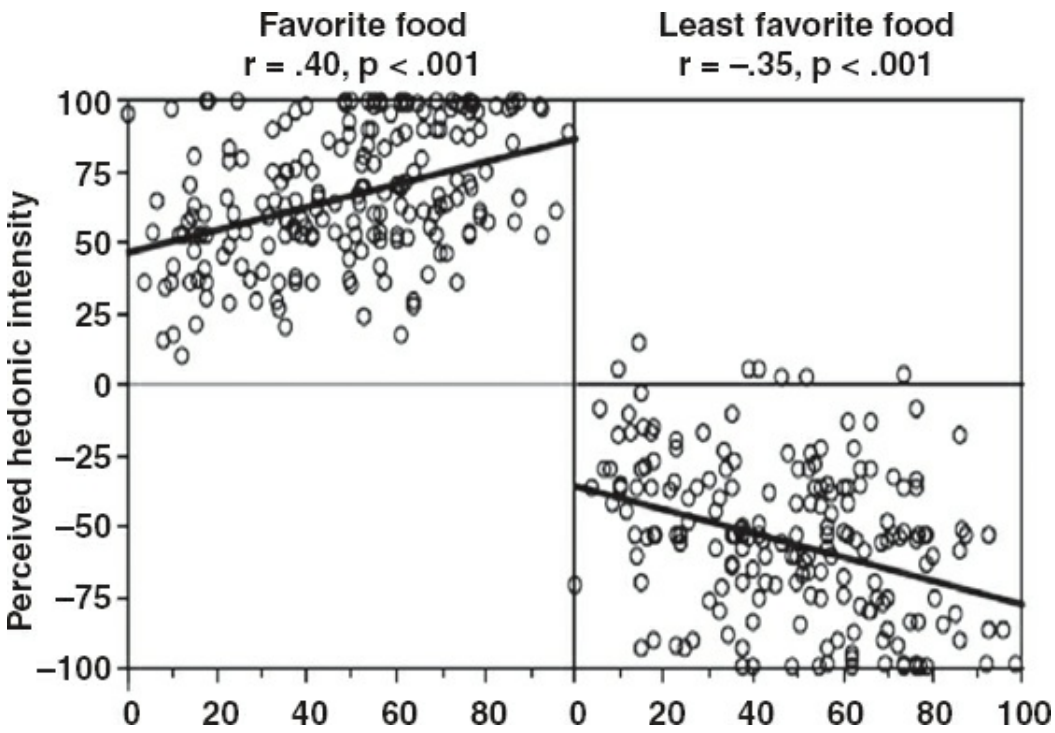

FIGURE 1.23–8. Hedonic intensity rated with the gLMS scale for favorite and least favorite foods plotted against subjects' taste capacity, assessed by the average sensory intensity rating of four standard taste solutions (1 M NaCl, 1 M sucrose, 0.032 M citric acid, and 0.001 M quinine hydrochloride). Note the significant correlations between taste capacity and both positive (favorite food) and negative (least favorite food) flavor hedonics. (From Kalva JJ, Sims CA, Puentes LA, Snyder DJ, Bartoshuk LM. Comparison of the hedonic general labeled magnitude scale with the hedonic 9-point scale. *J Food Sci*. 2014;79:S238.)

Variety. Variety is a strong hedonic control of eating. Offering a variety of nutritionally similar foods with different, preferred flavors leads to larger meals than does offering only one of the alternatives, even the single most preferred one. Figure 1.23–7 shows an example. The effect of variety on intake is referred to as sensory-specific satiety, reflecting the decline in the preference for a consumed food in relation to preferences for foods that have not been recently consumed. Flavor variety can tonically increase rats' intake and increase body weight.

Familiarity. Familiarity is sufficient to condition flavor preferences. This phenomenon likely explains much of the marked cultural variety of preferred foods, the social contexts or times of day when they are eaten, and so forth. For example, children repeatedly faced with foods that they find unpalatable and with the sight of adults enjoying them learn to enjoy and eat them as well. This learning has to overcome the innate tendency to avoid novel flavors (neophobia), which is especially strong in young children.

Individual Differences. Heterogeneity in human gustatory capacity is a source of individual differences in hedonics. For example, the density of

fungiform papillae on the anterior tongue is usually greater in women than in men, and this is associated with both individual and sex differences in bitter and sweet perception and hedonics. Furthermore, individual differences in sweet or bitter taste capacity are associated with the hedonic valence of a wide variety of flavors (Fig. 1.23–8). Such hedonic differences, in turn, are associated with *BMI* (weight in kg/height in m^2), indicating that they tonically increase meal size and caloric intake.

SATIATION AND POSTPRANDIAL SATIETY

Eating-inhibitory processes that occur during meals and contribute to meal termination are known as satiation processes. These are differentiated from longer-lasting processes that inhibit eating during the subsequent intermeal interval, known as postprandial or across-meal satiety processes. Because some signals appear to contribute to both, these processes are discussed together.

Gastric Mechanoreception

Although experimental dissection of gastric and postgastric signals in the control of eating is challenging, a variety of evidence indicates that gastric volume contributes directly to satiation via mechanoreceptors that are tuned to both tension and stretch or length. For example, (i) imaging studies have revealed significant correlations between fullness and gastric volume, (ii) increasing a meal's volume by mixing in air decreased meal size, and (iii) incorporating water into foods maintained gastric volume at higher levels, increased fullness, and decreased subsequent intake more than ingestion of the same amounts of food and water served separately. Gastric mechanoreceptors are linked to the brain mainly by vagal afferents.

Intestinal Nutrient Sensing

Food stimuli in the small intestine activate mechano-, osmo-, and chemoreceptors that initiate a number of potent satiation signals. Most important are the chemoreceptors tuned to detection of glucose, fatty acids, amino acids, and other digestive products. Negative-feedback signals generated by these nutrient sensors are relayed to the brain via endocrine and neural pathways, some of which are described below.

Oral preloads can be used to investigate the contributions of a food's metabolic energy content, nutrient composition, colligative and osmotic properties, and so forth to its satiating potency. When mixed nutrient preloads are used, decreases in eating are often proportional to the sizes of the preloads; that is, to their metabolizable energy contents. When the macronutrient compositions of the preloads vary, however, protein is typically more satiating on a calorie-for-calorie basis than carbohydrate or fat. Comparisons of intraduodenal and intravenous infusions indicate that

intestinal signals inhibit eating more potently than postabsorptive signals. For example, in one comparison of intraduodenal and intravenous glucose infusions that produced identical increases in systemic glucose levels, intraduodenal glucose but not intravenous glucose decreased premeal hunger ratings.

The next sections describe just three of the many intestinal hormones thought to contribute to the control of eating. The probable existence of numerous signaling mechanisms for satiation raises the question of their interactions and their relative contributions in various normal and abnormal contexts. As yet, little is known about these questions. Finally, the peripheral origin and, in most cases, peripheral sites of action on eating of GI hormones make them attractive targets for the development of pharmacological therapies.

Cholecystokinin. CCK is secreted by small-intestinal enteroendocrine I-cells. CCK secretion is stimulated by glucose and other sugars, amino acids, and long-chain fatty acids. CCK secretion begins soon after meal onset and usually returns to basal levels during the intermeal interval.

CCK's role as an endocrine satiation signal is uniquely well established. Intravenous CCK infusions mimicking postprandial plasma levels reduced meal size without adverse physical or subjective effects in men and women (Table 1.23–1, criterion 3), intraduodenal lipid infusions stimulated satiation, which was reversed by a CCKA receptor antagonist (criteria 4 and 6), and infusions of the antagonist alone increased meal size (criterion 6) (Fig. 1.23–9). No comparable antagonist effects have been reported in humans for any other hypothesized endocrine control of eating. Finally, occasional individuals bearing CCKA receptor null mutations are extremely obese, and more common CCKA receptor polymorphisms are associated with increased meal size, increased food intake, and obesity. CCK also fulfills the criteria for a normal physiological satiation signal in animals. Indeed, CCK research in rats and mice has been paradigmatic for the study of the control of eating in animal models. CCK elicits highly specific, dose-related decreases in meal size. The specificity is indicated by: (i) CCK fails to inhibit water intake in water-deprived rats under conditions where it inhibits eating liquid food; (ii) CCK elicits behavioral signs of satiation, including postprandial grooming and sleep, in rats that are sham feeding with open gastric cannulas, which otherwise sham feed essentially indefinitely; and (iii) CCK can decrease food intake without signs of illness, such as pica or formation of conditioned taste aversions (of course, large doses can induce illness). In summary, a wide variety of data indicate CCK is an important part of the natural process of satiation. As well, they indicate that the physiological system controlling satiation, however complicated, is not completely redundant.

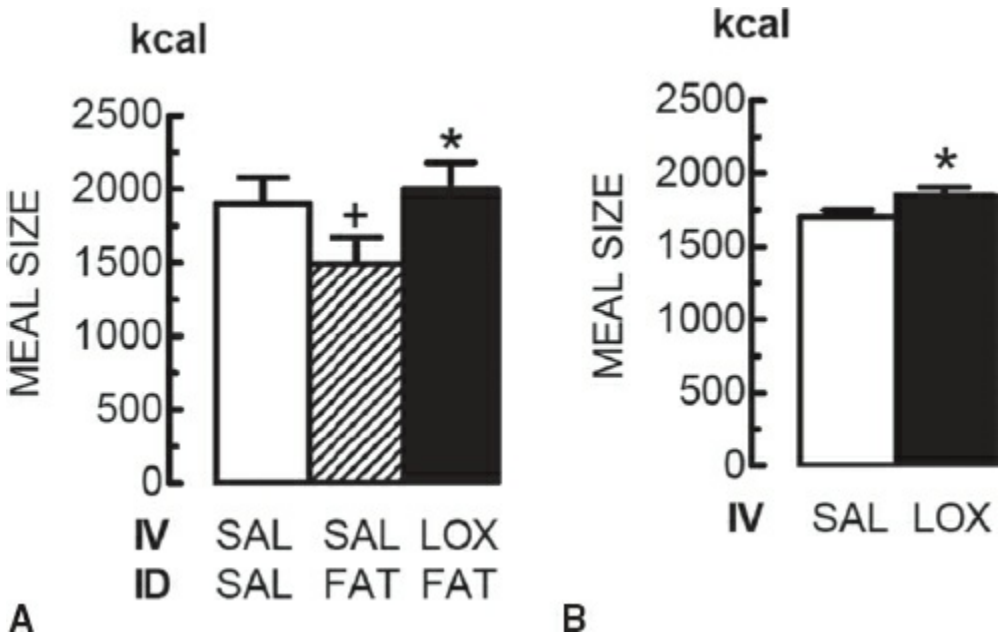

FIGURE 1.23–9. Evidence from CCKA receptor-antagonist studies that endogenous CCK is a physiological satiation signal. **A:** The CCKA receptor antagonist loxiglumide (LOX) antagonized the satiating action of endogenous CCK stimulated by intraduodenal (ID) infusion of a fat emulsion. Healthy-weight adult males began a midday lunch buffet 4 hours after a standard breakfast, 90 minutes after onset of an IV infusion of 10 $\mu mol \cdot kg^{-1} \cdot h^{-1}$ LOX or SAL, 60 minutes after an ID infusion of 0.4 $mL \cdot min^{-1}$ corn oil (FAT) or SAL, and 20 minutes after an oral preload of 400 mL of a low-fat banana milkshake (bananas were used because they do not elicit CCK secretion). Infusions were continued throughout the meal. ID fat infusion significantly reduced the size of the lunch meal (+), and that this was reversed by LOX (*); no physical or subjective side effects occurred in any condition. (From Matzinger D, Gutzwiller J-P, Drewe J, et al. Inhibition of food intake in response to intestinal lipid is mediated by cholecystokinin in humans. *Am J Physiol*. 1999;277:R1718, with permission.) **B:** Antagonism of CCK signaling with the CCKA receptor antagonist LOX stimulated eating. Healthy-weight adult males began a midday lunch buffet 4 hours after a standard breakfast and 60 minutes after beginning an IV infusion of 22 $\mu mol \cdot kg^{-1} \cdot h^{-1}$ LOX or SAL. Infusions were continued throughout the meal. LOX significantly increased meal size (*) without physical or subjective side effects. (From Beglinger C, Degen L, Matzinger D, D'Amato M, Drewe J. Loxiglumide, a CCK-A receptor antagonist, stimulates calorie intake and hunger feedings in humans. *Am J Physiol*. 2001;280(4):R1149, with permission.)

CCK appears to act on CCKA (or CCK1) receptors in the pyloric area, liver, and brain to elicit satiation. At least in rats, CCK also acts locally to elicit satiation, that is, in the intestinal lamina propria before absorption. For example, intravenous infusion of a small-molecule CCKA receptor antagonist, which presumably diffused from the capillaries into the lamina propria, increased food intake, but infusions of a CCK antibody, which would not escape the vasculature, did not. Activation of abdominal CCKA receptors stimulates vagal afferent neurons, which relay the feedback information to the NTS in the caudal brainstem.

Glucagon-like Peptide-1. Glucagon-like peptide-1 (GLP-1) is produced by enteroendocrine L-cells, which are found in both small and large intestine, and by a small population of neurons in the NTS. All three

macronutrients lead to intestinal GLP-1 secretion. Secreted GLP-1 is rapidly inactivated by dipeptidyl peptidase-IV (DPP-IV), a proline/alanine specific peptidase found on the luminal surface of capillary endothelial cells, in the liver, and in the blood. Despite this, and in contrast to CCK, GLP-1 secretion and plasma levels are elevated for hours after meals and return to basal only overnight. Thus, GLP-1 may function as a postprandial satiety or across-meal satiety signal. GLP-1 receptors on abdominal vagal afferents and possibly in the brain mediate GLP-1 satiation.

Whether GLP-1 signals satiation in humans has not been fully established. Consistent with a satiating effect, intravenous infusion of physiological doses of GLP-1 reduced meal size in the absence of adverse effects in normal-weight men (it has not been tested in women). In two tests, however, GLP-1 receptor antagonism failed to increase meal size, although the same antagonist had clear effects on insulin secretion and glycemia in other tests. GLP-1's status as a normal satiation signal deserves further study for several reasons: (i) GLP-1 antagonism stimulated secretion of glucagon and PYY(3–36) (see below), which may have interfered with its desatiating effect; (ii) the same antagonist increased eating in several tests in rats and mice; (iii) the antagonist may not have reached GLP-1 receptors in the lamina propria that may be important for satiation; (iv) the possibility that GLP-1 may function as an across-meal control of eating has not been explicitly tested; and (v) long-lasting GLP-1 agonists such as liraglutide (Victoza or Saxenda, Novo Nordisk) can reduce eating and body weight in obese patients.

Peptide Tyrosine Tyrosine. Intestinal L-cells are also the major source of peptide tyrosine tyrosine (PYY) (although most proximal L cells apparently secrete only GLP-1). All three macronutrients stimulate PYY secretion, with long-chain fatty acids the most potent. DPP-IV acts on secreted PYY (i.e., PYY(1–36)) to produce the active form, PYY(3–36). Interestingly, in rats, peripheral administration of PYY(3–36) selectively stimulates Y2 receptors and inhibits eating, whereas central administration of PYY(1–36) stimulates other Y-receptor types and stimulates eating.

PYY(3–36)'s role in postprandial satiety has been tested using intermeal infusions that began after standard meals and ended before test meals. Although these tests led to reductions in eating without signals of illness, it was unclear whether the doses were physiological. Infusions of supraphysiological doses of PYY(3–36) during meals also reduced intake, but in this case the subjects usually became ill. Thus, whether PYY(3–36) has a normal role in the control of eating is uncertain. It will be important for future research on this issue to focus on the possibility that PYY signals locally. Although all GI hormones may signal in part locally, this possibility is especially likely in the case of PYY because of the recent discovery that PYY-expressing L-cells have axon-like protrusions that form synapse-like

appositions with glial cells of the enteric nervous system.

Vagal Afferent Signaling. A variety of peripheral eating-inhibitory feedbacks are signaled to the brain via vagal afferents. Vagal afferent nerve endings densely innervate the lamina propria in most of the GI tract and, less densely, the hepatic paraganglia, the portal vein, and the bile duct. Spinal visceral afferents of the splanchnic nerves have a similar innervation pattern, but their role in eating is less well understood. Abdominal vagal afferents are stimulated by glucose and other digestive products, by CCK, GLP-1 and other hormones, and by serotonin (5-HT) and other signaling molecules that are released by food stimuli. Their importance in satiation is reflected in the increase in meal size that follows selective surgical section of sensory abdominal vagal fibers in rats (Fig. 1.23–10). Total vagotomy may not result in increased meal size because of the impairment of GI motor function caused by loss of vagal efferents.

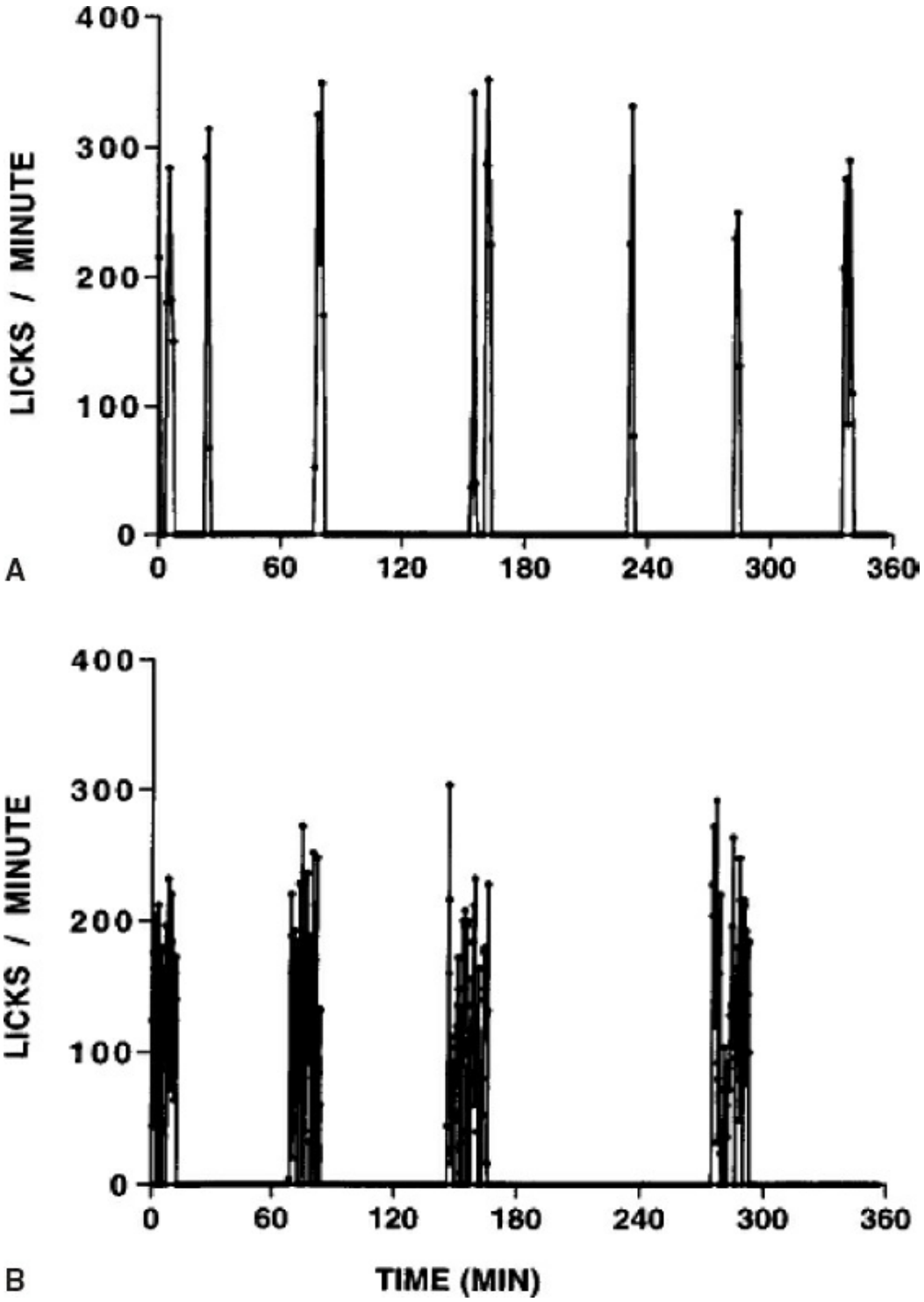

FIGURE 1.23–10. Disconnection of subdiaphragmatic vagal afferents leads to chronic increases in spontaneous meal size in rats. Subdiaphragmatic vagal deafferentation (SDA) combines unilateral dorsal vagal rhizotomy at the brainstem and section of the ipsilateral abdominal vagal trunk, which arises from the contralateral side of the brainstem. Data shown are representative patterns of licking (licks/minute) of liquid diet (Ensure, Ross Laboratories) before (**A**) and after (**B**) SDA. The increase in meal size and reduction in meal number evident in this rat was similar to the statistically significant mean increase in meal size of about one-third and mean decrease in meal frequency of about one-quarter observed in a group of SDA rats compared to sham-operated control rats. (From Schwartz GJ, Salorio CF, Skoglund C, Moran TH. Gut vagal afferent lesions increase meal size but do not block gastric preload-induced feeding suppression. *Am J Physiol*. 1999;276:R1623.)

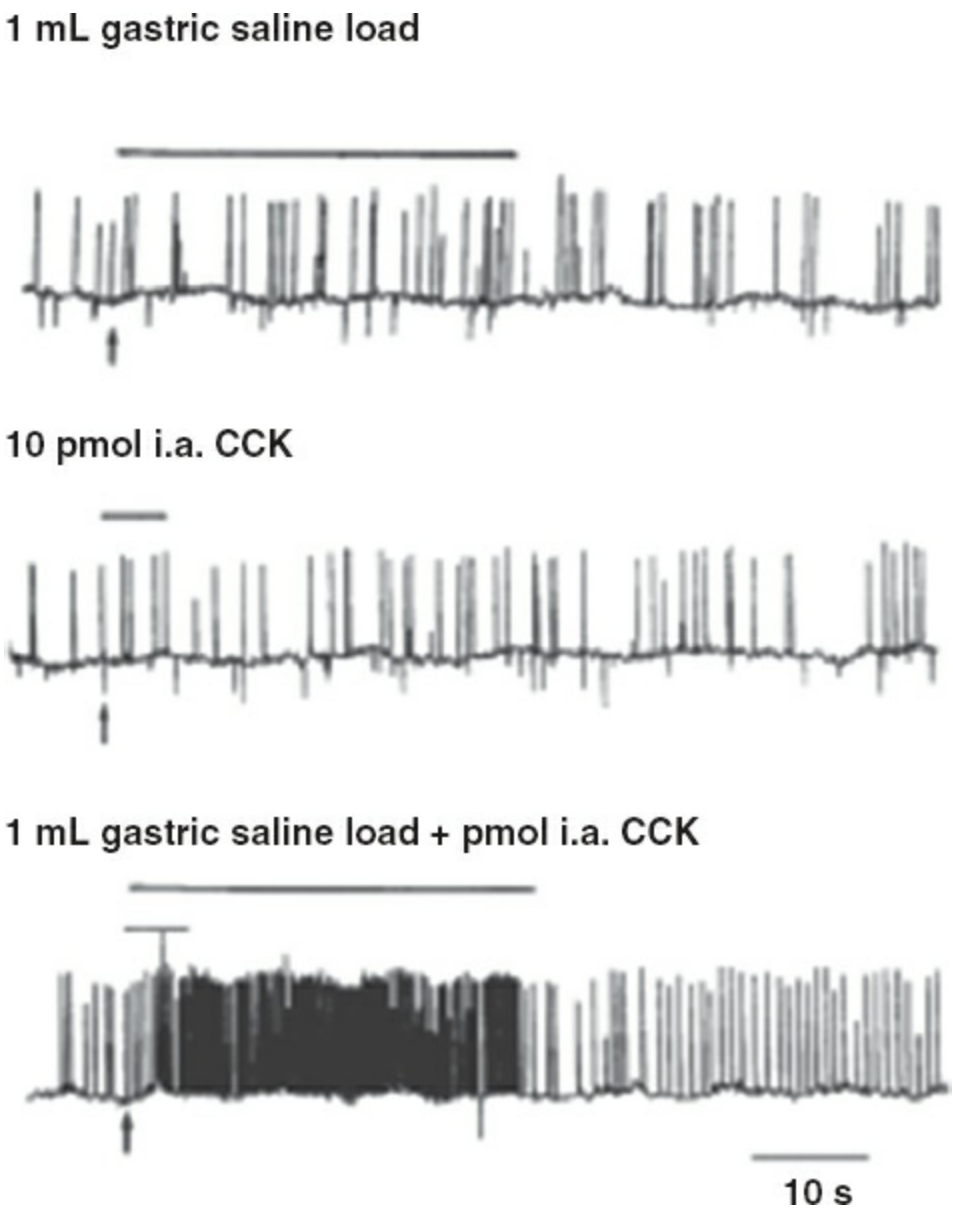

FIGURE 1.23–11. Synergistic effect of gastric mechanoreception and CCK on vagal afferent signaling in anesthetized rats. Data are extracellular recordings of action potentials in single vagal afferents in response to 1 mL intragastric saline loads and intraarterial (i.a.) infusions of 10 pMol CCK. (From Schwartz GJ, McHugh PR, Moran TH. Gastric loads and cholecystokinin synergistically stimulate rat gastric vagal afferents. *Am J Physiol*. 1993;265:R872.)

Vagal afferents integrate neural signals as well as relaying them to the brain. For example, in rats (i) some individual vagal afferent fibers are stimulated by both gastric fill and CCK infusion, and combinations of the two greatly amplify the vagal response (Fig. 1.23–11), and (ii) other fibers are sensitive to both ghrelin and CCK, and the two functionally antagonist stimuli sum to determine the net vagal electrophysiological response. Vagal afferent cell bodies in the nodose ganglion also demonstrate integrative properties in that patterns of mRNA expression of orexigenic and anorexigenic peptides and receptors vary depending on nutritional state.

Vagal (and spinal visceral) afferents project to the NTS, described in the CNS section.

Adiposity Sensing

Adiposity and Energy Homeostasis. The body contains both on-line and stored substrates for energy metabolism. On-line energy substrates include glucose and other metabolites used for energy that are circulating in the blood. Blood glucose levels, of course, are actively regulated. Stored energy substrates include glycogen in the muscle and liver, which amounts to a few thousand calories, and triacylglyceride stored in the adipocytes, which amounts to more than 70,000 calories in people at healthy weights. The relative constancy of body weight displayed by most people through adulthood, during which one eats about ten times as many calories per year as are stored in the adipose tissue, strongly suggests that adipose-tissue mass is physiologically regulated. At the same time, the epidemic rates of type-2 diabetes mellitus and obesity in many societies indicate that lifestyles in these societies stress the capacities for glycemic and adipose-tissue mass regulation beyond their capacities to handle surfeits.

Although much has been learned about the physiological mechanisms of adipose-tissue mass regulation, the feedback signals used by the brain to detect the mass of adipose tissue are not well understood. At present no strong candidate for a signal informing the brain about overweight has been identified.

Leptin. Leptin is a hormone secreted by adipocytes in proportion to their volume. It was discovered in 1994 in the search for the causes of the obesity–diabetes phenotype of mutant *ob/ob* and *db/db* mice, which had been shown by clever parabiosis experiments to lack a hormone and its receptor, respectively. Leptin's role in energy homeostasis, however, is to defend against underweight, not to defend against overweight. That is, decreases in leptin levels below the usual level are sufficient to increase eating and decrease energy expenditure, but increases in leptin above the usual level have little effect. Note that "usual level" means the level to which the organism is adapted. Thus, long-term obese persons will increase eating and decrease energy expenditure when leptin is reduced to levels that are still much higher than those in never-obese persons. How long a stable weight level must be maintained for such adaptation is unclear. Leptin physiology makes leptin monotherapy ineffective against obesity. Leptin has, however, shown promise in preclinical studies when administered in combination with other hormone-based therapies, such as long-acting forms of amylin, GLP-1, or CCK. Leptin is actively transported through the BBB and acts in a variety of brain loci to affect, among others, eating, energy expenditure, and HPD axis function.

BRAIN MECHANISMS OF APPETITE

Eating is mediated by a complex, anatomically diffuse neural network that extends from the caudal brainstem to the frontal lobes of the cerebral cortex. Figure 1.23–12 schematizes its general organization. In addition, a large number of interneuronal signaling molecules involved in this network have been identified. Specific localization of particular aspects of appetite to particular patterns of activity within this network remains difficult for several reasons: (i) the network's nodes are directly and indirectly interconnected by numerous converging and diverging neural projections, including many reciprocal connections; (ii) each node contains many small intrinsic neurons organized into a variety of local circuits, whose complex functions are poorly understood; and (iii) most of these areas also are involved in functions other than eating. Nevertheless, the application of increasingly powerful molecular-genetic and sophisticated behavioral methods has led to rapid progress. It is no criticism of these demanding experiments to say that they approach, but do not quite attain, the truly physiological. That is, what normally leads to the excitation and inhibition of these neurons and under what condition the signaling molecules involved meet the criteria of Table 1.23–1 for endogenous neurocrine function requires experiments that are not yet doable.

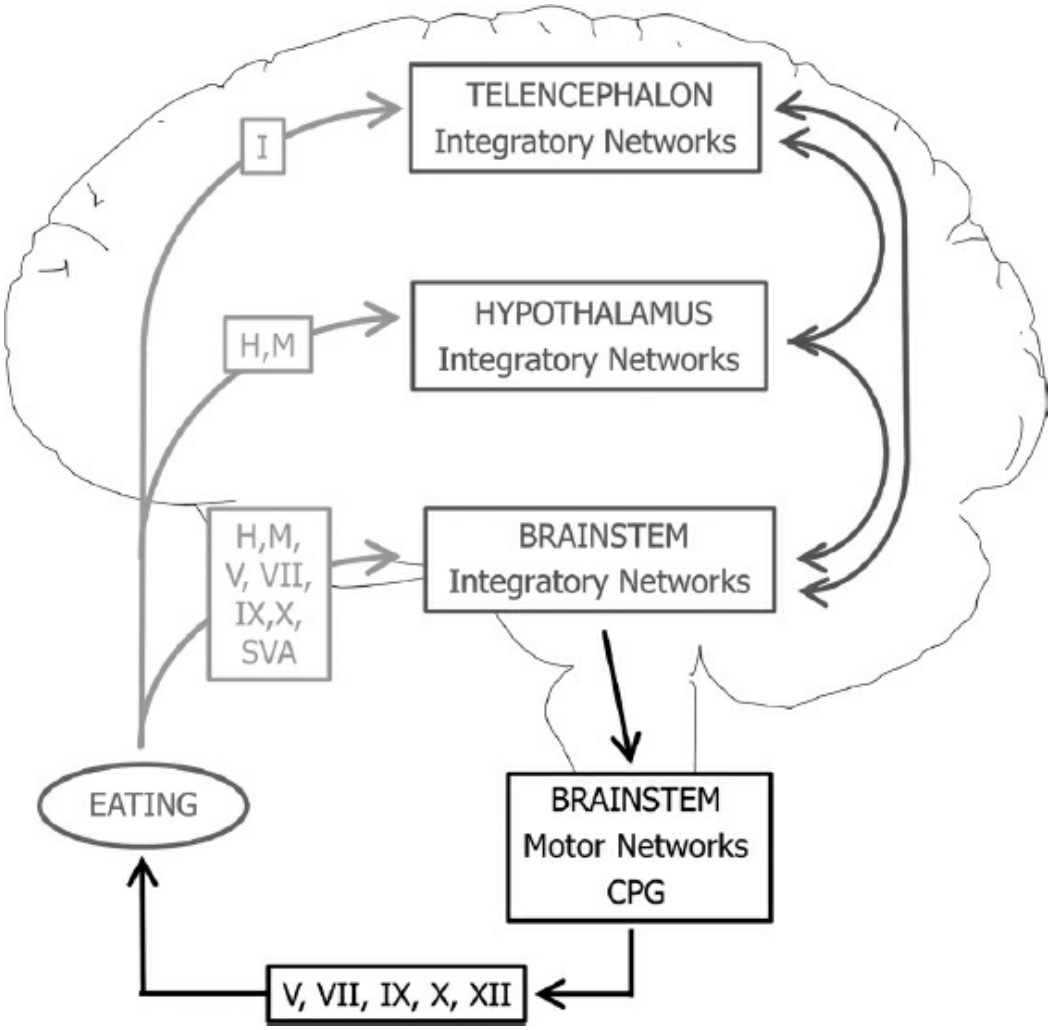

FIGURE 1.23–12. Schematic of the major brain areas involved in the control of eating as well as their principal afferent connections (*light gray arrows*), interconnections (*dark gray arrows*), and motor outputs (*black arrows*). Roman numerals indicate cranial nerves; H, hormones; M, metabolites; SVA, spinal visceral afferents; CPG, central pattern generator.

Brainstem

The brainstem has vital sensory, motor, and integrative functions related to eating. As mentioned above, the NTS receives vagal and spinal visceral afferents. It also is the target of primary gustatory afferents. In addition, it expresses a variety of metabolite receptors and hormone receptors (e.g., for ghrelin and leptin) that directly sense eating-relevant information. The NTS is also the target of neural efferents from the area postrema, which, like other circumventricular structures, lacks a normal BBB and has receptors for a variety of circulating molecules, including: (i) receptors for the hormone amylin, which is a candidate satiation signal; (ii) receptors for several toxins, which mediate conditioned taste aversion learning and elicit the emesis reflex; and (iii) osmoreceptors. The brainstem also contains the lower motor neurons that produce eating and swallowing movements, the upper motor neurons that drive the respective cranial nerve nuclei, and the central pattern generators that coordinate these movements. Accordingly,

difficulty eating and swallowing are signs of motor neuron diseases, such as bulbar palsy, that affect these motor networks. Finally, the NTS integrates a wide variety of information contributing to the control of eating. In addition to the sensory inputs described above, it is the target of projections from a variety of brainstem and hypothalamic areas, some of which are described below.

Tests of decerebrate rats, that is, rats bearing transections of the brain at the level of the midbrain, reveal the impressive integratory capacity of the caudal brainstem. When liquid food is infused into their mouths via oral catheters, they (i) ingest food in a meal-like pattern (which they end by food drain from the mouth) and after meals groom and become quiescent, similar to intact rats, (ii) eat progressively more of sweeter food (Fig. 1.23–13), (iii) display the orofacial movements associated with preferred and nonpreferred foods in intact rats, (iv) eat less after CCK injection, and (v) eat more after pharmacological blockade of glucose metabolism. The caudal brainstem does not, however, have the capacity to initiate eating spontaneously or to increase eating in response to food deprivation; thus, these functions require the forebrain.

Ascending Projections. Several ascending neural projections originate in the brainstem. Each contributes to a range of arousal, affective, and motivational functions related to eating.

A2 NORADRENERGIC NEURONS . A2 noradrenergic neurons have cell bodies in a region of the caudal medulla containing the NTS. They receive multiple inputs and project directly to the hypothalamus, cerebral cortex, and many other sites throughout the brain. They contribute to functions including vagovagal reflexes, satiation (including CCK satiation), and several aspects of affect and learning, especially related to stress or illness. Although noradrenergic neurons originating in the locus coeruleus (A4 and A6 cell groups) receive more attention in the study of stress, stress-related changes in eating are more likely to involve A2 noradrenergic neurons because these provide greater inputs to the PVN of the hypothalamus, central nucleus of the amygdala (CeA) and NAc, and other sites mediating eating.

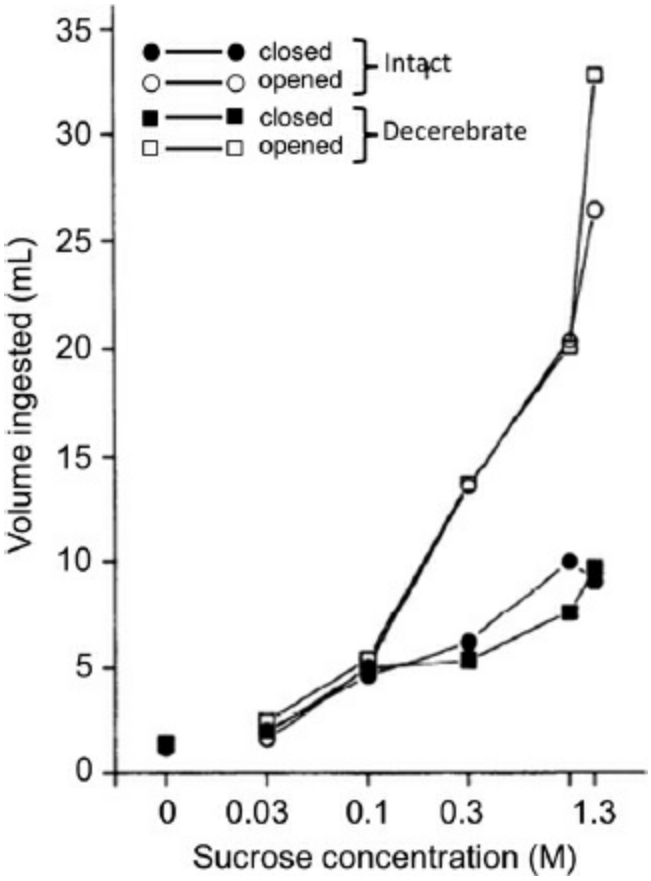

FIGURE 1.23–13. Eating behavior in intact and chronic decerebrate rats. Various concentrations of sucrose were delivered by intraoral catheters so that rats could either ingest it or passively reject it by allowing it to drip from the mouth. Intake was an increasing function of sucrose concentration in decerebrate rats that fed normally (closed condition). The gain of this function was dramatically increased in sham-feeding rats, in which postingestive controls of meal size were minimized by opening a gastric cannula (opened condition). Note the close correspondence, in intact and decerebrate rats, of the stimulatory effect on eating of increasing sucrose concentration and the interaction of this stimulatory effect with the inhibitory effect of normal postingestive food stimuli. (From Grill HJ, Kaplan JM. Sham feeding in intact and chronic decerebrate rats. *Am J Physiol*. 1992;262:R1070, with permission.)

DOPAMINE. The VTA contains dopaminergic neurons that project via the mesocorticolimbic pathways to numerous hypothalamic and telencephalic targets as well as to the NTS and other caudal sites. Each of these areas is also reciprocally connected to the VTA. This network is important in hedonic controls of eating, as discussed further below, as well as having numerous other affective and cognitive functions. Importantly, dopamine in different brain areas can have very different influences on eating. For example, dopamine neurotransmission in the NAc stimulates eating, whereas dopamine in the perifornical hypothalamus inhibits eating (a phenomenon potentially explaining the clinical finding that neuroleptics that antagonize dopamine increase body weight, apparently by increasing food intake).

SEROTONIN. Serotonin (5-hydroxytryptamine, 5HT) neurons in the midbrain raphe nuclei give rise to another diffuse ascending projection system that is involved in affect and motivation, including the control of eating. For example, 5HT signaling via the 2C receptor has been implicated in the eating-inhibitory effects of CCK, GLP-1, leptin, and the bacterial toxin LPS. Some of these effects may involve melanocortin signaling in the ARC (discussed below).

Hypothalamus

The importance of the hypothalamus for the neural control of eating has been recognized since the classic descriptions by John Brobeck and his colleagues of the syndromes of hyperphagia and aphagia that result from lesions of the mediobasal (VMH and ARC) and lateral hypothalamus (LHA), respectively. Originally understood as functional centers, these and a number of other hypothalamic areas are now conceptualized as nodes in a distributed neural network, with each node containing a number of signaling molecules with different, context-specific roles in eating. This is illustrated by the network element shown in Figure 1.23–14. This circuit is based on two chemically distinct neuronal populations in the ARC. One population expresses αMSH. The second population expresses NPY, agouti-related peptide (AgRP), and GABA. Both sets of neurons express receptors to ghrelin and leptin (as well as receiving a variety of additional inputs), with the hormones' electrophysiological effects opposite in sign within and between the two sets of ARC neurons.

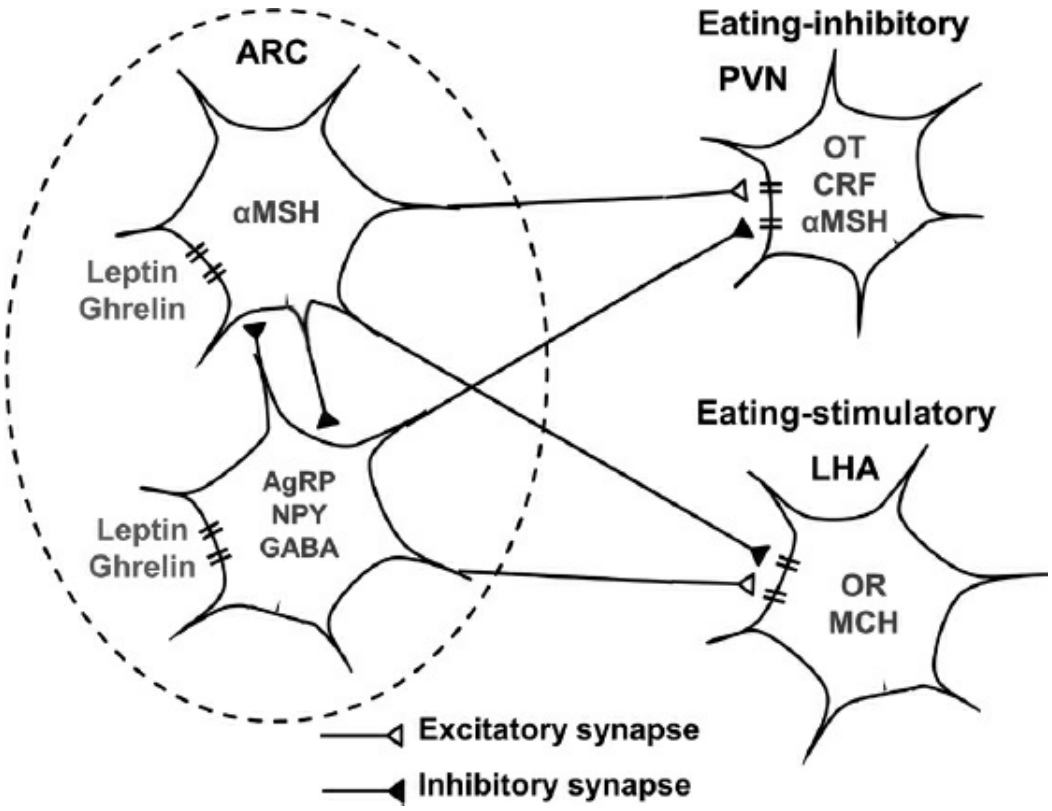

FIGURE 1.23–14. Schematic diagram of part of the hypothalamic component of the neural network controlling eating. Two populations of neurons in the hypothalamic arcuate nucleus (ARC) receive inputs from the hormones leptin and ghrelin, as well as other humoral signals and neural afferents (not shown). One of these populations expresses α-melanocyte–stimulating hormone (αMSH) as a neurotransmitter; these neurons are stimulated by leptin and inhibited by ghrelin. The other population expresses agouti-related peptide (AgRP), neuropeptide Y (NPY), and γ-amino butyric acid (GABA) as neurotransmitters; these neurons are inhibited by leptin and stimulated by ghrelin. The two populations of ARC neurons have mutually inhibitory connections. Among their principal projection targets are neurons in the paraventricular nucleus of the hypothalamus (PVN) that express oxytocin (OT), corticotropin-releasing factor (CRF), and αMSH as neurotransmitters and neurons in the lateral hypothalamic area (LHA) that express orexin (OR) and melanin-concentrating hormone (MCH) as neurotransmitters. αMSH neurons stimulate PVN neurons, and AgRP/NPY/GABA neurons inhibit them; in contrast, αMSH neurons inhibit LHA neurons, and AgRP/NPY/GABA neurons stimulate them. If the net balance of activity in this network is to stimulate the LHA neurons, eating is stimulated (thus, AgRP, NPY, OR, and MCH are often called orexigenic peptides; GABA is too widely expressed to have such a simplistic designation). If the net balance of activity in this network is to stimulate the PVN neurons, eating is inhibited (thus, αMSH, OT, and CRF are called anorexigenic peptides). Both these LHA and PVN neurons project to many sites in the brain to control eating.

Both sets of ARC neurons project to both the paraventricular hypothalamic nucleus (PVN) and the LHA, again with electrophysiologically opposite effects. Although the ARC, PVN, and LHA are also interconnected with other hypothalamic areas, the PVN and LHA appear to be the source of the most important extrahypothalamic projections related to eating. Thus, if the net result of neural activity in this circuit is to stimulate PVN neurons and inhibit LHA neurons, then eating is

inhibited; conversely, if the net effect is to stimulate LHA neurons and inhibit PVN neurons, then eating is stimulated. The effects of αMSH and AgRP are primarily mediated through their opposite effects on the melanocortin-4 receptor (MC4R; αMSH stimulates the receptor and AgRP is an inverse agonist; i.e., a functional antagonist). Mutations in *POMC,* the gene expressing αMSH, or *MC4R* comprise the commonest form of monogenic human obesity, occurring in 1 to 2 percent of obese patients.

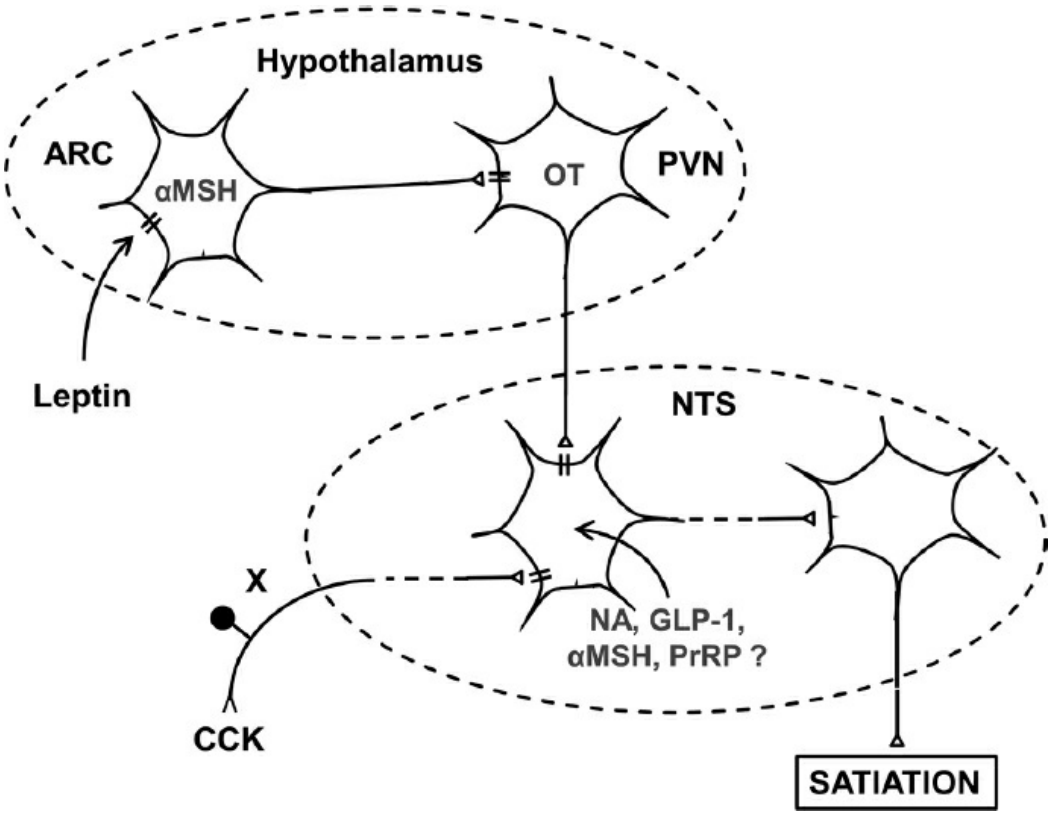

FIGURE 1.23–15. Schematic depiction of the neural network through which plasma leptin levels modulate the satiating potency of cholecystokinin (CCK). Increases in plasma leptin levels stimulate neurons in the hypothalamic arcuate nucleus (ARC) that express α-melanocyte–stimulating hormone (αMSH), and decreases in plasma leptin levels inhibit them (this appears to be the more physiologically relevant effect). ARC αMSH neurons project to neurons in the paraventricular nucleus of the hypothalamus (PVN) that express oxytocin (OT). These in turn project to neurons in the nucleus of the solitary tract (NTS) in the brainstem that also receive direct or indirect inputs from vagal afferent neurons (X) that are stimulated by CCK. Increases or decreases in the activity of this network lead to increases or decreases in satiation and meal size. The NTS network mediating CCK satiation appears to involve noradrenergic (NA), glucagon-like peptide-1 (GLP-1), αMSH, and prolactin-releasing peptide (PrRP) neurotransmission. Note: *broken lines* on the schematic axons indicate polysynaptic connections.

Latest-generation molecular genetic methods have led to fascinating insights concerning the roles of AgRP in eating. Optogenetic stimulation of AgRP neurons is sufficient to elicit eating in mice, with a rapid eating response mediated by release of NPY and GABA, and a slower onset, more sustained eating response elicited by release of AgRP. More selective

stimulation of specific axonal targets of AgRP neurons identified three projections, each of which mimicked the effect of global AgRp stimulation, the LHA, PVN, and anterior bed nucleus of the stria terminalis (aBNST). Others discovered that the AgRP projection the lateral parabrachial nucleus (lPBN) inhibits CGRP neurons, whose activation mediates the anorexia produced by visceral malaise.

It is important to note that leptin and ghrelin receptors outside the ARC also contribute to the control of eating, including in the NTS, VTA, and in abdominal vagal afferents. The multiple points of entry of these hormones into the brain, similar to the multiplicity of AgRP neural projections, presumably reflect both network redundancy and as yet unappreciated selectivity of function.

Descending Projections. Forebrain eating-control signals are relayed to the integratory networks in the caudal brainstem rather than directly to the motor networks for eating. Examples of this include: (i) leptin administration in intact rats increased the satiating potency of both gastric loads and CCK; (ii) leptin increased the neuronal activation produced by gastric loads and CCK in the NTS; (iii) rats bearing mutations of the leptin receptor had reduced sensitivity to exogenous CCK; and (iv) transgenic restoration of the leptin receptor function to these rats' ARC restored CCK's potency to inhibit eating and activate the NTS. Some of the neural pathways mediating these effects have been identified. As shown in Figure 1.23–15, ARC leptin-sensitive αMSH neurons project to oxytocin neurons the PVN, which then project to NTS neurons that also mediate CCK satiation.

Telencephalon

The best understood telencephalic contribution to appetite is the *telencephalic reward network,* which comprises a number of subcortical areas (e.g., NAc, ventral pallidum [VP]) and cortical areas (e.g., limbic, orbitofrontal, anterior cingulate, opercular, and insular cortex). Most of these structures receive dopaminergic, noradrenergic, and serotonergic inputs. In addition, opioid, cannabinoid, acetylcholine, orexin, benzodiazepine, glutamate, and GABA neurotransmission has been implicated in flavor reward. In addition to its roles in hedonics and reinforcement related to eating, this network is involved in other natural rewards, such as sex, and in drug reward.

Subcortical Reward Nodes. The NAc and VP have been extensively analyzed in rodents. Each has connections with the other and has numerous connections with more rostrally and caudally located structures. Dopamine is released in the NAc in a dose-dependent fashion as rats sham feed sucrose or liquid fat (Fig. 1.23–16), and local administration of

dopamine-receptor antagonists into the NAc reduces sham feeding of sucrose solutions, indicate that NAc dopamine neurotransmission plays an important role in flavor reward. This is supported by anatomical studies indicating that all the sensory modalities involved in flavor, that is, gustatory, olfactory, and somatosensory, converge in the NAc. Furthermore, these projections are relatively direct, that is, involve a minimal number of synaptic connections.

A dense projection from the NAc targets the VP, a subdivision of the substantia innominata. Opioid neurotransmission via μ-opioid receptors appears to mediate reward in this area. In a fascinating rat study, lesion of the caudal VP caused rats to display aversive facial reactions to sucrose rather than positive reactions, suggesting that neural processing in this area is necessary for generation of positive flavor hedonics.

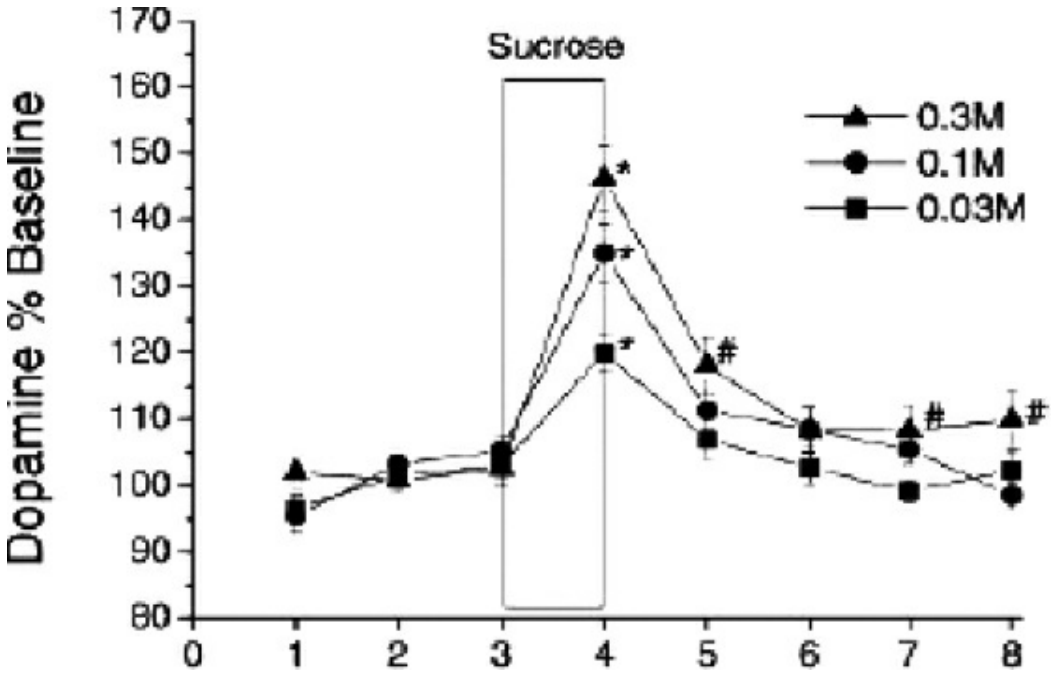

FIGURE 1.23–16. Evidence for the role of dopaminergic neurotransmission in the nucleus accumbens (NAc) in the orosensory reward of sweet taste. Different molar sucrose concentrations were presented for 20 minutes per day to *ad libitum*-fed rats, which sham fed with open gastric cannulas to minimize postingestive negative feedbacks. Rats ingested increasing amounts of more concentrated sucrose (means of 18, 31, and 43 mL of 0.03, 0.1, and 0.1 M sucrose, respectively), and simultaneous extracellular microdialysis revealed a graded increase in the release of dopamine in the NAc during ingestion. Control tests (not shown) revealed that similar graded dopamine release occurred when the volume of sucrose ingested was held constant. Data are shown for 20-minute periods as mean ± standard error of the mean of baseline. $*P < 0.01$; $\#P < 0.05$ versus baseline. (From Hajnal A, Smith GP, Norgren R. Oral sucrose stimulation increases accumbens dopamine in the rat. *Am J Physiol*. 2004;286:R31, with permission.)

Cortical Reward Nodes. Several areas of the frontal and temporal lobes of the cortex have been implicated in flavor reward. These include both evolutionarily older areas (e.g., the hippocampus and cingulate cortex) and

evolutionarily newer or neocortical areas (e.g., orbitofrontal cortex and prefrontal cortex). Functional imaging studies in human subjects as well as single-neuron recording studies in monkeys indicate that flavor hedonics is best mirrored by the activity of the mid-anterior orbitofrontal cortex. For example, orbitofrontal cortical activity is much higher if palatable foods are sampled when hungry than when satiated. In contrast, the medial orbitofrontal, ventromedial prefrontal, and middle anterior insula cortices may be less involved in generating the experience of flavor hedonics than in monitoring and predicting flavor rewards.

PHYSIOLOGICAL MODULATORS OF APPETITE

Learning

Meal initiation, food selection, and meal size are all readily conditionable in animals and humans, as indicated in Figure 1.23–2 above. Both classical (or Pavlovian) and instrumental (or operant) learning are involved. Environmental context and flavor usually provide the CS. For example, when a sound/light CS was presented to rats before each of six scheduled meals for several days and then tested during "extinction," that is, when the rats had free access to the same diet, the CS elicited initiation of a very large meal on each daily presentation for 3 weeks. Thus, cues that predict food availability during food deprivation can provoke the initiation of a large meal in the absence of deprivation. The unconditioned stimulus (UCS) for this learning has not been identified.

Satiation is also conditionable. The UCS for conditioned satiation are the postingestive consequences of eating. For example, rats that drink flavored water while proteins, carbohydrates, or fats are intragastrically infused subsequently eat less food with that flavor. This is satiation, not aversion. When preference is tested, the rats choose the flavor associated with the nutrient infusion. The importance of conditioned controls can be demonstrated in the absence of explicit learning contingencies, for example, by using the sham-feeding procedure to extinguish conditioned controls of meal size (Fig. 1.23–17).

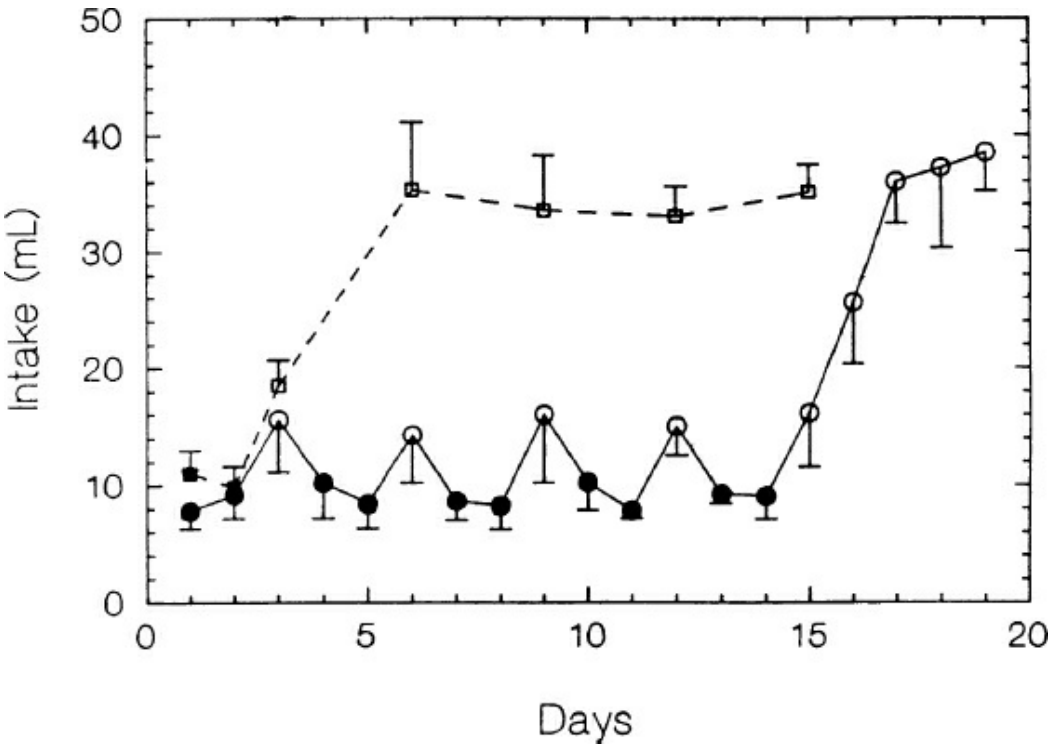

FIGURE 1.23–17. The role of conditioned satiation in eating in the rat. Data are 30-minute intakes of 0.8 M sucrose in two groups of rats in successive tests. Rats were prepared with steel gastric cannulas that were either closed for real-feeding tests (closed symbols) or open for sham-feeding tests (open symbols), during which sucrose did not accumulate in the stomach or enter the intestines. One group (*squares*) real-fed sucrose once and then sham fed. The other group (*circles*) alternated real- and sham-feeding tests. The progressive increase in sucrose intake during consecutive sham-feeding tests was prevented by interspersing two normal feeding tests between each sham-feeding test. When this was done, sham intake was significantly larger than real intake but significantly smaller than the asymptotic sham intake. These data indicate that normal feeding of sucrose produces a learned association between the taste of the sucrose and its postingestive effects, that this association limits intake, and that it extinguishes during consecutive sham-feeding tests. (From Davis JD, Smith GP. Learning to sham feed: Behavioral adjustments to loss of physiological postingestional stimuli. *Am J Physiol*. 1990;259:R1228, with permission.)

Higher-order conditioning is likely to play important roles in human eating. For example, cultural socialization may result in a preference for the flavor of capsaicin (chili), which all infants avoid. Indeed, with the exception of a few unconditioned gustatory preferences and aversions (such as for sweet and bitter taste), *all* food selection appears to be learned. Finally some forms of learned controls of eating are special forms of conditioning. Flavor aversions conditioned by upper GI food poisoning, for example, can be learned after only one CS–UCS pairing despite extraordinarily long CS–UCS delays (hours) and are extremely resistant to extinction.

The analysis of the influence of learning on human eating should be a high priority. The number of unconditioned hindbrain mechanisms, through which all learned influences of meal size presumably operate, is

small enough to imagine achieving control of them in humans. Furthermore, behavior therapy programs, because they are relatively highly structured and can present specific food stimuli, should provide excellent opportunities for the clinical use of conditioned physiological controls of eating.

Exercise

Although exercise contributes primarily to the expenditure side of energy balance, some recent work has suggested inhibitory effects of exercise on food intake, especially in obesity models. One example of such an action occurs in Otsuka Long Evans Tokushima Fatty (OTELF) rats, which lack CCK1 receptors and are hyperphagic and obese. When given access to running wheels, they are equally active to the control strain, but they reduce their food intake within a day down to levels of the control rats and maintain this lower intake for as long as running wheel access is maintained. This results in normalization of body weight. When they no longer have access to running wheels, food intake increases. Importantly, the magnitude and extent of the increase depends upon the developmental stage during which exercise is experienced. In mature rats, intake and body weight return to pre-exercise levels. However, when access to exercise is provided to young OLETF rats, the effects appear to be permanent. Food intake temporarily increases but not to levels in OLETF rats that did not have running wheel access, and body weights are maintained at levels significantly below those of rats that did not have access to running wheels. Similar effects of exercise have been demonstrated for rats bred to be susceptible to dietary obesity and in mice lacking melanocortin-4 receptors. The mechanisms underlying such actions have yet to be determined.

Recent work in rats has also demonstrated effects of exercise on diet preference. Access to a running wheel results in the reduction of an existing preference for a high fat diet and can produce an avoidance of a novel normally preferred palatable diet.

Sex Differences

Ovarian cycling and other reproductive states affect eating. Adult women and female animals of many mammalian species decrease eating during the periovulatory phase of the cycle. In women, this decrease, which is not consciously appreciated, amounts to a few hundred kilocalories per day, which over the long term can affect body weight (Fig. 1.23–18). This is presumably caused by the increase in plasma estradiol levels during the follicular phase. Studies of oophorectomy and estrogen treatment in rats and mice indicate that estrogens acting on estrogen receptor-1 (or α) in the medial NTS mediate these effects.

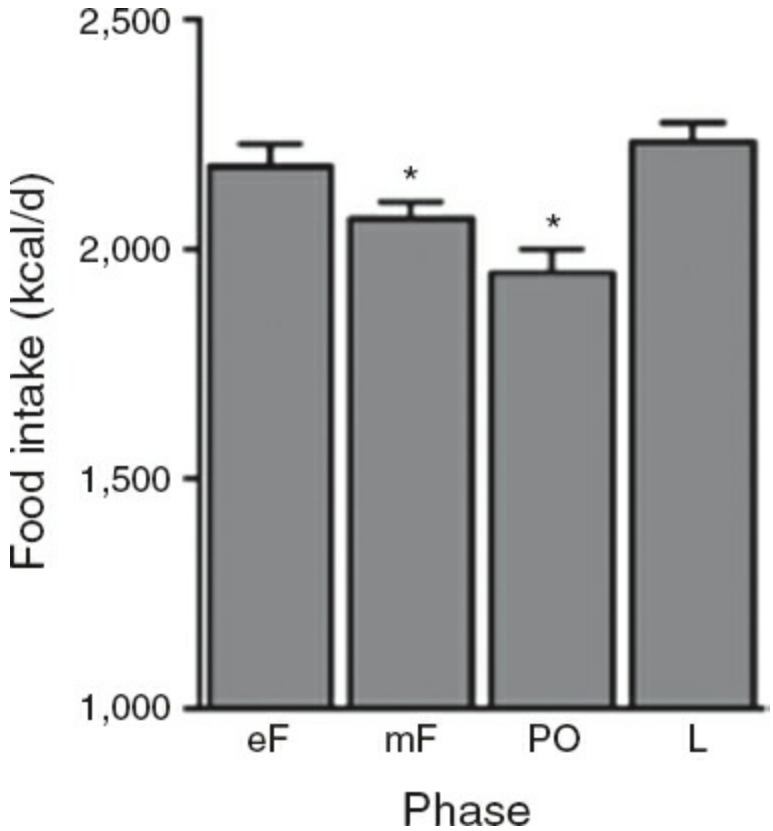

FIGURE 1.23–18. Daily food intake during the ovarian cycle in women, averaged across the early-follicular (eF; 4 days), mid-follicular (mF; about 9 days), periovulatory (PO; 4 days), and luteal (L; about 11 days) phases. Note the progressive decrease in food intake during the follicular phase. Old-world monkeys, whose cycle is apparently identical to the human cycle, show a similar follicular decrease in daily food intake and also a high, constant level of food intake during most of the luteal phase, which has not been verified in women. *Significantly different from luteal phase. (From Asarian L, Geary N. Sex differences in the physiology of eating. *Am J Physiol*. 2013;305:R1215.)

In rats and mice the periovulatory decrease and the postoophorectomy increase in food intake are caused by selective changes in meal size; meal frequency does not contribute. Indeed, when the hyperphagia of oophorectomized rats abates after they have increased body weight about 25 percent, it is because meal frequency decreases; the oophorectomy-induced increase in meal size is permanent. This is a good example of the importance of meal pattern analysis in understanding the controls of eating. Estrogens decrease meal size during the periovulatory phase in rats at least in part by selectively increasing the satiating potency of CCK.

Both human and animal research indicate that there is a sex difference in the avidity for sweets, with females preferring them more than men. One reflection of this is that women more frequently endorse sweets as their most preferred or craved foods, whereas men more frequently endorse savory foods. The mechanisms underlying this sex difference are unknown.

Illness Anorexia

Anorexia of varying intensity and duration is a common element of both

innate and acquired immune responses to infection, trauma, neoplasm, and other challenges. The transient anorexia of the acute phase response of the innate immune system, like fever, is an unpleasant but adaptive response that can facilitate recovery. More chronic illness anorexia, however, is a maladaptive response that can increase illness severity and interfere with therapy.

A number of cytokines, including interleukin-1 and TNF-α, as well as prostaglandin-E_2 and other immune signaling molecules are involved in illness anorexia. Peripheral immune signals apparently affect the brain both directly and indirectly, via endocrine and peripheral neural responses, and converge on the same neural networks that mediate normal eating. For example, serotonergic neurons in the DRN, which project into the hypothalamus and telencephalon, have been implicated in both the normal control of eating and illness anorexia. In animals, illness can lead to reduced meal size, reduced meal number, or both. Illness anorexia in humans remains very poorly understood but is presumably equally complex.

PSYCHIATRIC EATING DISORDERS

Behavioral Studies

The translation of behavioral neuroscience approaches into clinical research has contributed importantly to understanding the phenomenology of eating disorders. This is exemplified by two studies of eating in patients with bulimia nervosa. First, it was discovered that under laboratory conditions patients with bulimia nervosa eat significantly larger meals than normal volunteers. This difference can be obtained in individual test meals of a single test diet or in residential laboratory settings in which subjects have free access to a variety of foods for one or more days. A study of the latter design by Walter Kaye and colleagues showed that patients with bulimia nervosa and control subjects took similar numbers of meals and that most of the meals taken by patients with bulimia were within the range of meal sizes displayed by the controls but that about a quarter of the meals taken by the patients were larger than any meal taken by control subjects, on average by a factor of 10. Second, it is now clear that experimentally manipulated cognitive stimuli are sufficient to induce binges in the laboratory. That is, when subjects are instructed to eat large meals, patients with bulimia nervosa eat much larger meals than do controls and report that these meals have the subjective character of binges.

Satiation in Bulimia Nervosa

Postingestive negative-feedback satiation signals are less potent in patients with bulimia. (1) Equivalent preloads of food decrease intake less in patients with bulimia nervosa than in controls, particularly when the

patients are eating large meals. (2) Patients with bulimia must eat larger amounts of food to produce equivalent self-reports of fullness during a meal. (3) Volume distention of the stomach produces a decreased perceptual and mechanical response in patients with bulimia, presumably because their stomachs are larger than normal as the result of accommodation to the frequent ingestion of large meals. (4) Food-stimulated CCK release is less in bulimic patients than that in controls (Fig. 1.23–19). These abnormalities appear to resolve as binging decreases, suggesting that they are not the initial cause of bulimia. Nevertheless, it is possible that they facilitate the development of the disorder once it has begun and impede recovery from it.

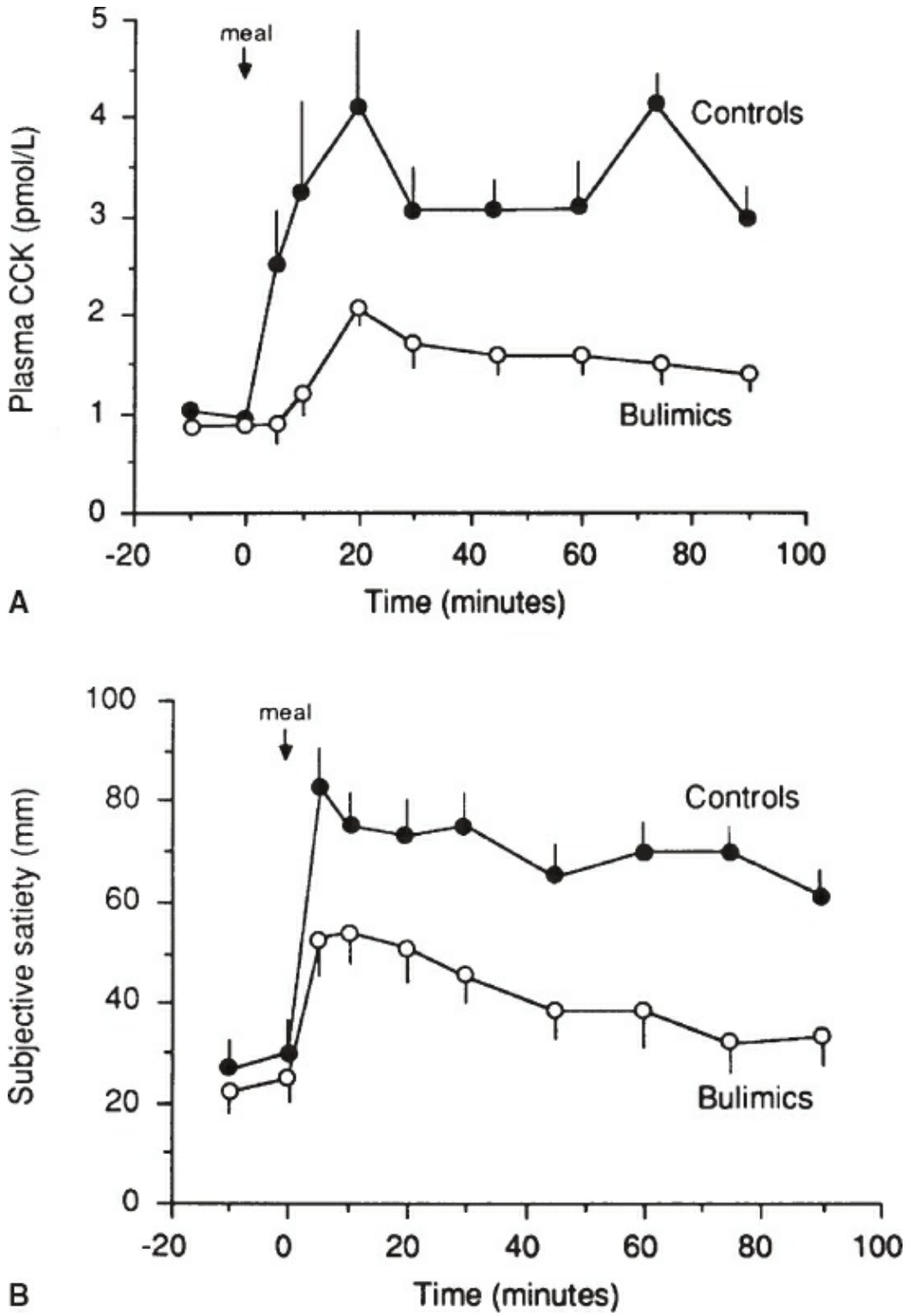

FIGURE 1.23–19. Prandial concentrations of cholecystokinin (CCK) (**A**) and the subjective experience of satiety (**B**) are reduced in patients with bulimia nervosa. Fourteen patients and ten control women matched for age and weight were offered a 400-mL liquid meal after an overnight fast (*arrows*) and ate it in 1 to 2 minutes. Plasma CCK was measured with a selective bioassay. Satiety was measured by a 100-mm visual analog scale (0 = empty; 100 = full). Both the peak CCK concentration and the integrated CCK response (area under the curve) were significantly reduced in patients with bulimia nervosa, and this correlated with reports of significantly less satiety beginning 5 minutes after meal onset. (From Geracioti TD, Jr., Liddle RA. Impaired cholecystokinin secretion in bulimia nervosa. *N Engl J Med*. 1988;319:683, with permission.)

Serotonin and Eating Disorders

Functional brain imaging, often involving radiolabeling methods to identify neurotransmitter function, is a major research focus in eating disorders. The most progress to date involves the study of serotonergic mechanisms. For example, women with bulimia nervosa still displayed reductions in $5HT_{2A}$ receptor binding in the orbital frontal cortex following recovery, suggesting that this might reflect a trait difference underlying vulnerability for this disorder. Similarly, women recovered from anorexia nervosa also displayed reduced $5HT_{2A}$ receptor binding, albeit in different brain areas. A key challenge in research of this type is to identify the functional nature of the vulnerability. For example, serotonergic changes might be related to anxiety disorders, which occur in the majority of patients with anorexia nervosa and bulimia nervosa, rather than specifically to the control of eating. Unfortunately, relatively few brain imaging studies of eating disorders in patients occur in the context of eating. More work of this type might reveal state or trait differences related to the physiology of appetite between eating disorders, and suggest new targets for improved treatment.

The acute tryptophan depletion (ATD) test has been used to investigate the role of serotonin in eating disorder behaviors. The test involves ingestion of a tryptophan-free protein load, which diverts plasma tryptophan into peripheral protein synthesis, leading to a decrease in central serotonin synthesis. A recent study by Kenneth Bruce and colleagues indicated that ATP (1) decreased mood in both patients with bulimia nervosa and healthy controls and (2) increased the urge to binge in patients who were medicated with serotonin-reuptake inhibitors at the time of testing, but not in patients who were not so medicated and not in controls. These data are consistent with a specific role of serotonin in acute binge urges. Further research is warranted to determine, first, why ATD affected binging only in the medicated group (perhaps they had a more severe disorder or perhaps it depended on the pharmacologically reduced urge to binge), second, where in the brain the critical effect on serotoninergic neurotransmission occurred, and, third, what functional system it influenced.

Sex Differences in Eating Disorders

Whether physiological sex differences contribute to the increased vulnerability of women to eating disorders is unknown. There are many reasons to think that they do. For example: (1) the role of estrogens in the increased vulnerability of women to eating disorders has not been established. This possibility is suggested by their potent influence on the satiating action of peripheral CCK in animals together with the changes in CCK satiation associated with bulimia nervosa, both described above. In addition, because the estrogenic inhibition of eating in animals first

appears at puberty, it seems possible that estradiol's inhibitory influence on eating may be part of the reason that anorexia nervosa most frequently develops shortly after menarche. (Note, however, that in anorexia nervosa ovarian secretion is suppressed as body weight drops, so any defect in the estrogenic control of feeding would have to be a precipitating factor not required for the continued course of the disorder.) (2) Family gene-association studies have linked polymorphisms in ESR1, which encodes the estrogen receptor gene most associated with the physiology of eating, to increased risk for anorexia nervosa, although this has not yet been verified in population-wide studies. (3) Female rats are markedly more vulnerable to display binge-like eating in animal models of binge eating than are male rats (Fig. 1.23–20).

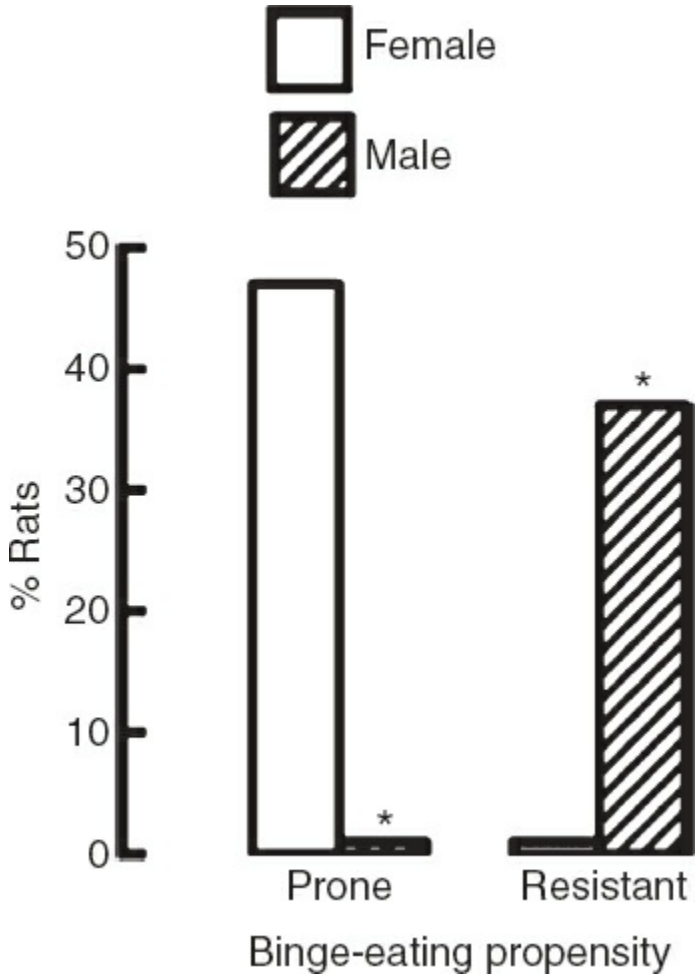

FIGURE 1.23–20. Male–female sex difference in rats' propensities to binge eat. Male and female rats were offered standard chow *ad libitum* and a palatable commercial cake frosting 3 days a week for 2 weeks, a procedure that leads to increased palatable food intake in comparison to *ad libitum* access to the same food. Four-hour palatable food intakes in each of the six tests were ranked across all rats, and individual differences in binge-eating propensity were scored. Data are percentages of male and female rats that were binge-eating prone (i.e., scored in the highest tertile of palatable food intake in three or more of the six tests and never scored in the lowest tertile) or binge-eating resistant (i.e., scored in the lowest tertile of palatable food intake in three or more of the six tests and never scored in the lowest tertile). *Significant sex difference. (From Klump KL, Racine S,

Hildebrandt B, Sisk CL. Sex differences in binge eating patterns in male and female adult rats. *Int J Eat Dis*. 2013;46:729, with permission.)

REFERENCES

Asarian L, Geary N. Sex differences in the physiology of eating. *Am J Physiol*. 2013;305:R1215.

Asarian L, Gloy V, Geary N. Homeostasis. In: Ramachandran VS, ed. *The Encyclopedia of Human Behavior,* vol. 2. San Diego, CA: Academic Press; 2012.

Asarian L, Langhans W. A new look on brain mechanisms of acute illness anorexia. *Physiol Behav*. 2010;100:464.

Atasoy D, Betley JN, Su HH, Sternson SM. Deconstruction of a neural circuit for hunger. *Nature*. 2012;488:172.

Bailer UF, Kaye WH. Serotonin: Imaging findings in eating disorders. In: Adan RA, Kaye WH, eds. *Behavioral Neurobiology of Eating Disorders*. Berlin, Germany: Springer Verlag; 2011.

Berridge KC, Kringelbach ML. Neuroscience of affect: brain mechanisms of pleasure and displeasure. *Curr Opin Neurobiol*. 2013;23:294.

Blevins JE, Ho JM. Role of oxytocin signaling in the regulation of body weight. *Rev Endocrinol Metab Disorders*. 2014;14:311.

Blundell J, de Graaf C, Hulshof T, et al. Appetite control: methodological aspects of the evaluation of foods. *Obes Rev*. 2010;11:251.

Bohórquez DV, Shahid RA, Erdmann A, et al. Neuroepithelial circuit formed by innervation of sensory enteroendocrine cells. *J Clin Invest*. 2015;125:782.

Brobeck JR. Remembrance of experiments almost forgotten. *Appetite*. 1993;21:225.

Bruce KR, Steiger H, Young SN, Ng Ying Kin NM, Israël M, Lévesque M. Impact of acute tryptophan depletion on mood and eating-related urges in bulimic and nonbulimic women. *J Psychiat Neurosci*. 2009;34:376.

Carter ME, Soden ME, Zweifel LS, Palmiter RD. Genetic identification of a neural circuit that suppresses appetite. *Nature*. 2013;503:111.

Castro DC, Berridge KC. Advances in the neurobiological bases for food 'liking' versus 'wanting.' *Physiol Behav*. 2014;136:22.

French SA, Epstein, LH, Jeffrey, RW, Blundell JE, Wardle J. Eating behavior dimensions. Associations with energy intake and obesity. A review. *Appetite*. 2012;59:541.

Grill HJ, Hayes MR. Hindbrain neurons as an essential hub in the neuroanatomically distributed control of energy balance. *Cell Metab*. 2012;5:296.

Higgs S, Robinson E, Lee M. Learning and memory processes and their role in eating: Implications for limiting food intake in overeaters. *Curr Obes Rep*. 2012;1:91.

Kaye WH, Weltzin TE, McKee M, McConaha C, Hansen D, Hsu LK. Laboratory assessment of feeding behavior in bulimia nervosa and healthy women: methods for developing a human-feeding laboratory. *Am J Clin Nutr*. 1992;55:372.

Lam DD, Garfield AS, Marston OJ, Shaw J, Heisler LK. Brain serotonin system in the coordination of food intake and body weight. *Pharmacol Biochem Behav*. 2010;97:84.

Leibel RL. Molecular physiology of weight regulation in mice and humans. *Int J Obes (London)*. 2008;32(Suppl 7):S98.

Levin BE, Magnan C, Dunn-Meynell A, Le Foll C. Metabolic sensing and the brain: who, what,

where and how? *Endocrinol*. 2011;152:2552.

Meye FJ, Adan RA. Feelings about food: the ventral tegmental area in food reward and emotional eating. *Trends Pharmacol Sci*. 2014;35:31.

Moody L, Liang J, Choi PP, Moran TH, Liang NC. Wheel running decreases palatable diet preference in Sprague-Dawley rats. *Physiol Behav*. 2015;150:53–63.

Moran TH, Dailey MJ. Intestinal feedback signaling and satiety. *Physiol Behav*. 2011;105:77.

Perello M, Dickson SL. Ghrelin signaling on food reward: a salient link between the gut and the mesolimbic system. *J Neuroendocrinol*. 2015;27(6):424–434.

Potenza MN, Grilo CM. How relevant is food craving to obesity and itt Treatment? *Front Psychiat*. 2014;5:164.

Ravussin Y, Leibel RL, Ferrante AW Jr. A missing link in body weight homeostasis: the catabolic signal of the overfed state. *Cell Metab*. 2014;20:565.

Rinaman L. Hindbrain noradrenergic A2 neurons: diverse roles in autonomic, endocrine, cognitive, and behavioral functions. *Am J Physiol*. 2011;300:R222.

Rolls BJ. The relationship between dietary energy density and energy intake. *Physiol Behav*. 2009;97:609.

Rolls ET. Taste, olfactory and food texture reward processing in the brain and obesity. *Int J Obes*. 2011;35:550.

Rosenbaum M, Leibel RL. 20 years of leptin: role of leptin in energy homeostasis in humans. *J Endocrinol*. 2014;223:T83.

Schwartz GJ, Moran TH. Sub-diaphragmatic vagal afferent integration of meal-related gastrointestinal signals. *Neurosci Biobehav Rev*. 1996;20:47.

Schwartz GJ. Central leucine signaling in the control of energy homeostasis. *Endocrinol Metab Clinic N A*. 2013;42:81.

Sclafani A. Gut-brain nutrient signaling. Appetition vs. satiation. *Appetite*. 2013;71:454.

Smith GP, Gibbs J. The development and proof of the cholecystokinin hypothesis of satiety. In: Dourish CT, Cooper SJ, Iversen SD, Iversen LL. *Multiple Cholecystokinin Receptors in the CNS*. Oxford, England: Oxford University Press; 1992.

Steinert R, Feinle-Bisset C, Asarian L, Horowitz M, Beglinger C, Geary N. Ghrelin, CCK, GLP-1 and PYY(3-36): Secretory controls and physiological roles in eating and glycemia in health, obesity and after RYGB. Physiological Reviews, in press.

Thompson RH, Swanson LW. Hypothesis-driven structural connectivity analysis supports network over hierarchical model of brain architecture. *Proc Natl Acad Sci USA*. 2010;107:15235.

Treesukosol Y, Moran TH. Analyses of meal patterns across dietary shifts. *Appetite*. 2014;75:21.

Walsh BT, Kissileff HR, Cassidy SM, Dantzic S. Eating behavior of women with bulimia. *Arch Gen Psychiat*. 1989;46:54.

Weingarten HP. Conditioned cues elicit feeding in sated rats: a role for learning in meal initiation. *Science*. 1983;220:431.

▲ 1.24 Neuroscience of Substance Use Disorders

INTRODUCTION

"Drug addiction" will likely remain in current vernacular, but it fails as a medical term because it is laden with semantics and negative stigma. In particular, it is impossible to ascribe a precise point at which drug abuse can be classified as drug addiction. Likewise, "drug dependence" fails as a medical term for a disorder because dependence can be a normal physiological response to a substance. SUD, as defined in DSM-5, solves these issues by combining the DSM-IV categories of substance abuse and substance dependence into a single disorder measured on a continuum from mild to severe. In DSM-5, each specific substance (hereafter referred to as "drug") class (e.g., alcohol, stimulants) is addressed separately and diagnostic criteria are provided for three distinct drug-related phenomena: SUD, substance intoxication, and substance withdrawal. Although disorders with the majority of drugs are diagnosed based on the same overall criteria (Table 1.24–1), it is increasingly clear that chronic exposure to different drugs produces different patterns of neuroadaptations.

Table 1.24–1.
Generic Diagnostic Criteria for Substance Use Disorder, DSM-5

1. Drug is taken in larger amounts or over a longer period than intended.
2. There is a persistent desire or unsuccessful effort to reduce drug use.
3. A great deal of time is spent in activities necessary to obtain, use, or recover from the drug.
4. Craving for the drug.
5. Recurrent drug use resulting in a failure to fulfill major obligations at work, home, or school.
6. Continued drug use despite persistent of recurrent social or interpersonal problems caused by drug use.
7. Important activities are given up or reduced because of drug use.
8. Recurrent drug use in situations where it is physically hazardous.
9. Continued drug use despite knowledge of having a persistent or recurring physical or psychological problem that is likely to have been caused by or exacerbated by the drug.
10. Tolerance to the drug.
11. Withdrawal: either experiencing a drug withdrawal syndrome or taking drug to relieve or avoid the withdrawal syndrome.

From a neurobiological perspective, severe SUD is a relapsing CNS disorder characterized by craving and compulsive drug seeking and taking along with negative emotional states (e.g., dysphoria, anxiety, irritability) when access to the drug is prevented. Neural adaptations persist beyond detoxification and are responsible for craving and relapse when an individual has drug access. Research over the last four decades has significantly advanced the understanding of the neurobiology underlying SUD, which has led to improved treatments. Through the use of increasingly sophisticated experimental methods in laboratory animals and

humans, detailed information from molecular to behavioral levels exists on the neural pathways that underlie each phase of the addiction process. Although animal models cannot fully reconstitute psychiatric disorders in humans, studies in laboratory animals permit investigation of specific elements of SUD. As neuroscience and psychiatry have evolved, these elements have become more concrete and tractable.

As described in the DSM-5, SUD comprise a collection of cognitive, behavioral, and physiological symptoms suggesting that an individual continues to abuse drugs in a pathological pattern despite adverse consequences. SUD diagnostic criteria (Table 1.24–1) can be roughly grouped into four categories:

1. Impaired control (Criteria 1 to 4)
2. Social impairment (Criteria 5 to 7)
3. Risky use (Criteria 8 and 9)
4. Pharmacological (Criteria 10 and 11)

PHASES OF THE ADDICTION PROCESS

The distinction between drug use, abuse, and SUD is critical to understand. Drug use includes legal use of substances such as caffeine, nicotine, and alcohol as well as medically prescribed use of drugs such as benzodiazepines, marijuana, and opioid painkillers. Drug abuse includes illicit drug use or when a psychoactive drug is taken for a nonmedical use. It is estimated that 15 to 16 percent of the US adult population engage in drug abuse at some point in their lives. Of those, approximately 3 percent go on to develop SUD. Implicit in this is the idea that drug use (drug exposure) is required for the expression of SUD, wherein increasing levels of drug consumption leads to compulsive drug use despite adverse consequences.

The transition from casual drug use to SUD has been neurobiologically conceptualized as a shift from impulsive to compulsive use and hedonic dysregulation driven by a transition from positive to negative reinforcement mechanisms. Reinforcement is a motivational process by which a response-contingent delivery or removal of a stimulus increases the probability of a subsequent response. There are two forms: positive reinforcement, which is associated with euphoria (e.g., initial drug “high”) and negative reinforcement, which is associated with cessation of an aversive state (e.g., drug withdrawal). Casual drug users who go on to develop SUD display characteristics of impulse control disorders, in which there is an increasing sense of tension or arousal before the impulsive act followed by a sense of relief or pleasure when the act is committed. Impulse control disorders are thought to involve positive reinforcement mechanisms. In contrast, individuals who meet criteria for SUD display characteristics of compulsive disorders, in which there are feelings of

anxiety and stress before engaging in a compulsive and repetitive behavior followed by relief from stress when the compulsive behavior is enacted. Compulsive disorders are thought to involve negative reinforcement mechanisms and habitual, automatic behaviors. When viewed within this psychiatric-motivational framework, the development of SUD can be described as a three-stage cycle: binge/intoxication, withdrawal/negative affect, and preoccupation/anticipation (Fig. 1.24–1). Impulsivity is most evident in the early stages whereas compulsivity dominates at later stages. The transition from substance use to SUD involves pathological neuroplasticity at each stage within neural circuits that regulate affective state, motor patterns, and cognition.

Binge/Intoxication

Initial drug use may begin in vulnerable individuals. However, it is difficult to study this early stage of substance use because what constitutes vulnerability and reasons for initiating drug use can vary from person to person and culture to culture. Furthermore, the fact that only a small subset of people who use drugs casually will go on to develop SUD suggests that there is no underlying pathology in the majority of people who experiment with drugs use. Indeed, the DSM-5 provides diagnostic criteria for drug intoxication that are separate from those comprising SUD. Whether or not a person tries a drug and what the hedonic valence of initial intoxication will be is influenced by a confluence of innate traits such as impulsivity, sensitivity, anxiety, risk-taking behavior, and biological and environmental influences such as socioeconomic status, education, exposure to trauma and abuse, peer pressure, drug availability, family drug use, and the social construct of gender. Despite these complexities, recent studies have begun to identify some of the critical behavioral and biological markers that may predict a propensity to SUD. Animal models have been developed that reflect the innate or environmental factors thought to contribute to acquisition of drug taking. These models continue to be tested in behavioral paradigms that reflect the binge/intoxication stage of the SUD cycle (Table 1.24–2).

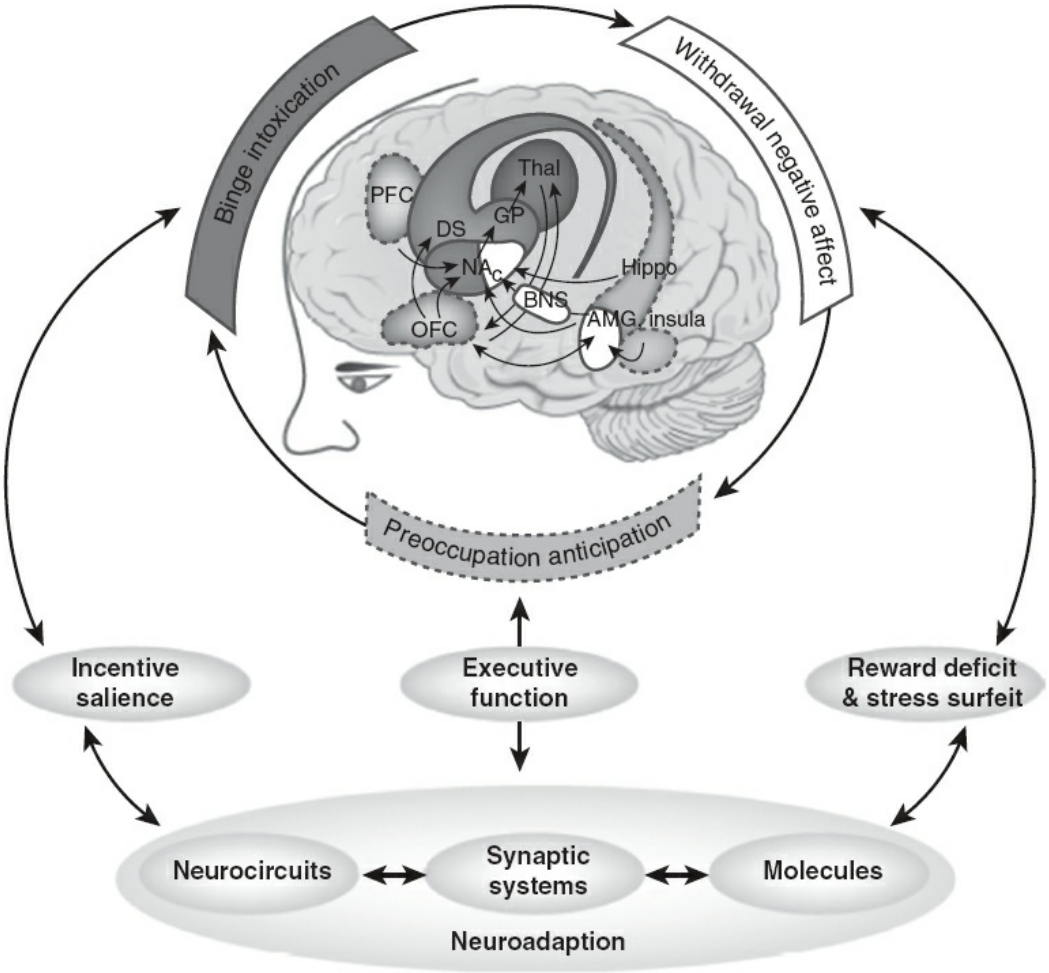

FIGURE 1.24–1. Diagram showing the neurocircuitry of the addiction cycle divided into three stages: binge/intoxication (*dark gray*); withdrawal/negative affect (*white*); and preoccupation/anticipation (*light gray*). The neurocircuits involved are also color-coded, with the basal ganglia, including the nucleus accumbens (NAc), dorsal striatum (DS), globus pallidum (GP), and thalamus (Thal) as key elements of the binge/intoxication stage; the extended amygdala, including the central nucleus of the amygdala (AMG), bed nucleus of the stria terminals (BNST), and a transition area in the shell of the nucleus accumbens (NAc) as key elements of the withdrawal/negative affect stage; and the frontal cortex and allocortex, including the prefrontal cortex (PFC), orbitofrontal cortex (OFC), hippocampus (Hippo), and insula (Insula) as key elements of the preoccupation/anticipation stage. Molecular, synaptic, and neurocircuitry neuroadaptations combine to render the four key elements of the transition to addiction: increased incentive salience, decreased reward, increased stress, and decreased executive function. (From Wise RA, Koob GF. The development and maintenance of drug addiction. *Neuropsychopharmacology*. 2014;39:254.)

Withdrawal/Negative Affect

As substance use progresses toward SUD, periods of chronic drug use are typically followed by periods of abstinence and increasingly severe somatic and/or psychological withdrawal signs. People with SUD have described withdrawal states as a "hunger" or a primary need that is only satisfied by taking more drug (i.e., negative reinforcement). As drug use escalates and becomes compulsive in nature, neuroadaptations start to develop that

result in decreased cognitive control and increased motivation for drugs that outweigh the motivation for natural rewards. Motivation is defined as the biological, emotional, social, and cognitive forces that activate behavior, and it can be modulated by drug-induced changes in hedonic and affective states that occur with the transition from substance use to SUD. This has been conceptualized in the opponent process theory of motivation (Fig. 1.24–2). According to Solomon and Corbit (1974), "hedonic, affective, or emotional states are automatically opposed by CNS mechanisms which reduce the intensity of hedonic feelings, both pleasant and aversive." Positive hedonic responses to drug use (i.e., euphoria, reward) occur shortly after taking the drug, correlate closely with the intensity, quality, and duration of the drug, and show tolerance. Upon drug withdrawal, negative affective states are unmasked. These are slow to develop, slow to decay, and grow in magnitude with repeated exposure to drug. In the context of SUD, opponent processes begin early in drug taking and reflect compensatory changes in the neural circuitry that controls reward and stress responses.

Preoccupation/Anticipation

Periods of drug withdrawal are followed by instances of craving and relapse, whereby various triggers such as stress, exposure to the drug itself or to previously drug-paired environmental cues reinstate drug seeking and drug taking. The subjective experience of craving frequently precedes relapse. Craving can be defined as an acquired desire characterized by wanting/anticipating drug use. Its importance is exemplified by its inclusion in the DSM-5 criteria for diagnosis of SUD (Table 1.24–1). In studies of people with psychostimulant use disorder, increasing levels of craving have been associated with increasing levels of anxiety, irritability, restlessness, increases in heart rate, nausea, and other symptoms that overlap with stress-related arousal and that are transiently relieved with drug use.

Table 1.24–2.
Using Animal Models to Study the Stages of Substance Use Disorders

Stage	Source of Reinforcement	Animal Models
Binge/intoxication	Positive reinforcement	Conditioned place preference, drug self-administration, choice procedure, increased brain stimulation reward
Withdrawal/negative affect	Negative reinforcement	Increased anxiety-like responses, decreased brain stimulation reward, conditioned place aversion, increased motivation for drug self-administration during dependence
Preoccupation/anticipation	Conditioned	Drug-, cue-, stress-induced reinstatement of drug

positive reinforcement	seeking
Conditioned negative reinforcement	Protracted abstinence

Adapted from Koob GF and Volkow ND. *Neuropsychopharmacology*. 2009;35:217.

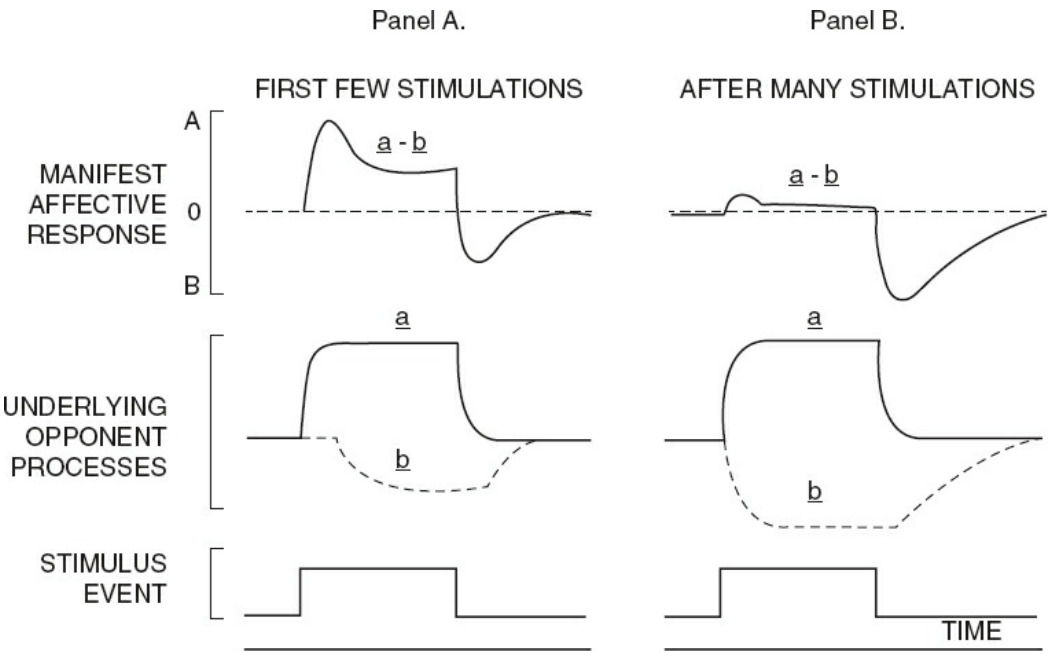

FIGURE 1.24–2. Opponent process theory. **Panel A:** The operation of the summing device for the first few stimulations. **Panel B:** The operation of the summing device after repeated stimulations. The summation of the underlying opponent processes (schematics in middle row), "a" and "b," yields the manifest affective response (schematics on top row). In the context of drugs of abuse, manifest affective response "A" represents acute drug reward and "B" represents drug-opposite effects (aversion) that emerge on stimulus off-set (schematics on lower row). (From Solomon RL, Corbit JD. An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychol Rev*. 1974;81:119.)

The objective measurement of craving in the laboratory has historically presented a major challenge for researchers; however, animal models provide a variety of objective and quantifiable indices of drug-taking and drug-seeking behavior from which neurobiological substrates of craving may be deduced. The best empirical model to study mechanisms of drug-seeking and drug-taking behavior is self-administration, an established operant procedure for measuring the reinforcing effects of drugs. Drug self-administration in animals meets face, construct, and predictive validity criteria. Animals contingently administer drugs of abuse in a manner akin to human drug users. Furthermore, the common use of intravenous injections, which allows for rapid drug delivery to the brain, models human methods of drug administration. Drug self-administration studies in animals are also widely used for preclinical evaluation of putative pharmacotherapies for the treatment of SUD. Preclinical studies suggest that the underlying mechanisms and neural circuits mediating the

reinforcing effects of drugs in animals overlap substantially with those mediating the abuse-related effects of drugs in humans. These studies have also increased the understanding of the normal function of the neural substrates of reward related behavior.

The study of the neurobiological substrates of relapse has received particularly intense focus, because this point in the cycle of SUD underlies the chronicity of the disorder, and preventing relapse is critical for successful treatment approaches. Drug self-administration paradigms have been adapted to model relapse by focusing on the reinstatement of operant behavior (i.e., lever pressing or nose poking) previously associated with drug delivery. "Reinstatement"—a common abbreviation for "reinstatement of extinguished drug self-administration behaviors"—is an animal model of drug craving and relapse. In this procedure, laboratory animals are trained to self-administer a drug of abuse and then put through extinction training in which operant responding is no longer paired with drug delivery. When a predetermined criteria of operant responding under extinction conditions is achieved, the animals are tested for reinstatement of drug-seeking behavior by presentation of a drug prime, discrete cues, contextual cues, or a stressor. Although reinstatement models in animals can be further refined to more closely approximate relapse in humans, they possess good face and construct validity for modeling the activation of drug craving and relapse produced by various stimuli in drug-dependent humans.

THEORIES OF ADDICTION

A number of theories have been proposed to explain the development and maintenance of SUD. Given the heterogeneity of the disorder (e.g., contributions of genes, environment, and the drugs themselves), it is likely that aspects of each theory are applicable under given conditions. The major theories are summarized below and framed in the context of the SUD cycle described above and in Figure 1.24–1. It is important to note that these theories are not mutually exclusive but rather tend to emphasize particular aspects of the SUD cycle. There is a general consensus that positive reinforcement kicks off the cycle and leads to the initial repetition of drug taking that eventually becomes habitual and compulsive.

Psychomotor stimulant theory: Addictive drugs share common features: they cause psychomotor activation, are reinforcing, and elicit approach behaviors. Seminal work done in the 1960s to 1980s demonstrated that the mesocorticolimbic dopamine system is a major neural substrate of the natural approach system. In this theory, people initiate and maintain drug-taking behavior because drug-induced dopamine release has reinforcing effects and activates motor systems that result in repetitive drug "approaches." Proponents of this theory argue that drug withdrawal symptoms do not play a major part in compulsive drug

use.

Incentive sensitization theory: This theory builds upon the psychomotor stimulant theory of addiction in that drugs are hypothesized to hijack neural systems (e.g., mesocorticolimbic dopamine system) that motivate animals to approach stimuli with salience (importance) for preservation of the species. Contexts, cues, and other stimuli—including drugs of abuse—that activate these neural systems become tagged with incentive salience (motivational importance, "wanting"). This theory posits that repeated exposure to drugs of abuse results in persistent neural adaptations in these systems such that the animal becomes hypersensitive to drugs and to any cues that have become associated with the drug. Thus, as SUD develops, a person's motivational focus narrows to drug seeking at the expense of natural reward seeking. A key aspect of this theory is that sensitization of neural systems that control incentive salience (drug "wanting") is independent of changes in neural systems that control reward (drug "liking"). Finally, proponents of this theory argue that the systems that mediate incentive salience are distinct from those controlling drug withdrawal-induced negative hedonic states.

Aberrant learning theory: The increase in the incentive salience of drugs and drug cues at the expense of natural rewards is explained in the context of aberrant learning. As with Incentive Sensitization theory, this theory holds that repeated exposure to drugs heightens responsiveness to drug-associated cues by modulating neural systems that control normal responses to nondrug associated cues. This aberrant learning process is thought to represent changes in associative brain structures such that drug seeking and taking becomes automatic or habitual. Evidence supports the idea that this is mediated by a gradual shift from ventral to dorsal striatal mechanisms. An important aspect of this theory is that the sensitized responsiveness to drugs and drug cues is immune to punishment, resulting in a core symptom of SUDs: continued drug use despite adverse consequences.

Frontostriatal-dysfunction theory: This theory incorporates mechanisms to explain the clinical observation that people with SUD tend to have poor impulse control by proposing that repeated drug exposure leads to alterations in cortical and limbic circuits such that an individual displays pathological responses to drugs and drug cues while simultaneously displaying impaired inhibitory, "top-down" executive control over drug-taking behavior. This leads to loss of impulse control, poor decision-making, heightened responsiveness to drug-associated cues, and compulsive drug use.

Hedonic-allostasis theory: This theory is rooted in the opponent-process theory of motivation, and postulates that initial drug abuse is motivated by the positive reinforcement obtained from the drug's rewarding effects, whereas chronic drug abuse is motivated by the negative

reinforcement resulting from the recruitment of stress-related systems and emergence of negative hedonic states. This chronic drug-induced negative emotional state, termed the "hedonic allostatic" state, reflects a persistent change in the animal's normal reward set point. Hedonic allostasis contributes to compulsive drug taking through negative reinforcement mechanisms that involve circuits connecting cortical, striatal, and thalamic regions.

BRAIN PATHWAYS IN ADDICTION

There is considerable evidence from animal models and human studies that all drugs of abuse converge on a common circuitry in the brain's limbic system. Synthesizing (and simplifying) decades of work, there are three primary circuits upon which chronic exposure to all drugs of abuse act. First, chronic exposure to addictive drugs causes a homeostatic impairment of the dopamine system, in which natural rewards (e.g., food, water, sex, social companionship) become less effective at increasing dopamine release in a typical fashion. These homeostatic changes are thought to contribute to negative affective states observed between periods of drug taking or during periods of drug withdrawal and abstinence. It is also observed that chronic, intermittent drug exposure can sensitize the dopamine system, such that exposure to the drug or drug-associated stimuli elicits greater increases in dopaminergic transmission. This sensitization of dopamine responses can persist long after drug taking ceases and may relate to drug craving and relapse. A second neural circuit common to all drugs of abuse involves the extrahypothalamic actions of CRF. Acute drug withdrawal activates CRF-containing neurons in the amygdala, which project to numerous forebrain and brainstem regions discussed below. The stress-like effects of CRF are thought to contribute to the negative affective and somatic symptoms that occur upon drug withdrawal. A final adaptation common to chronic exposure of most classes of abused drugs is cortical "hypofrontality," in which there is reduced basal activity of cortical neurons and a decreased ability of natural rewards to activate certain cortical regions such as the prefrontal cortex, ACC, and orbitofrontal cortex. These cortical regions control executive function, working memory, attention, and behavioral inhibition through glutamatergic projections to the nucleus accumbens and VTA. Chronic drug-induced neural adaptations in these frontal cortical regions and their glutamatergic outputs are thought to contribute to both impulsive and compulsive aspects of SUD.

Table 1.24–3.
Generic Diagnostic Criteria for Substance Intoxication, DSM-5

A. Recent ingestion of substance.
B. Clinically significant problematic behavioral or psychological changes that develop shortly during or

shortly after substance ingestion*.

C. A minimum number of physiological/psychological symptoms that are substance-specific*.
D. The signs or symptoms are not attributable to another medical condition and are not better explained by another mental disorder, including intoxication with another substance.

*DSM-5 includes euphoria for cannabis, inhalants, opioids, and stimulants.

Neural Pathways Underlying Acute Drug Reward

The DSM-5 separates substance intoxication, withdrawal, and SUD into distinct diagnostic categories. Intoxication comprises four criteria (Table 1.24–3), two of which can include feelings of euphoria that occur during or shortly after substance ingestion. Although effects other than euphoria can certainly be reinforcing, this section will focus on brain pathways known to play a role in the acute rewarding effects of drugs of abuse. It is this feature of intoxication that is thought to reinforce drug taking and lead to SUD in vulnerable individuals.

The discovery in the 1950s that rats would learn to press a lever to self-deliver electrical stimulation into specific brain regions (an operant behavior known as ICSS) allowed for the discovery of neural circuits involved in reward function. The efficacy of ICSS to produce brain stimulation reward is quantified by measures such as the minimum frequency of electrical stimulation that supports operant responding. Although brain stimulation reward involves multiple brain regions, the most sensitive site—defined by the lowest stimulation thresholds—is within the medial forebrain bundle at the level of the lateral hypothalamus. The medial forebrain bundle connects the VTA to the basal forebrain and carries ascending norepinephrine and dopamine monoaminergic fibers (Fig. 1.24–3). Acute administration of most drugs of abuse decreases brain stimulation reward thresholds (i.e., increase reward) whereas chronic administration increases reward thresholds during withdrawal (i.e., decrease reward). In summary, these early studies provided strong support for monoaminergic neurotransmission as a critical substrate of the reinforcing effects of drugs of abuse.

Emerging data is accumulating that mesolimbic dopamine pathway activity is required for the primary reinforcing effects of psychostimulants but has a more enabling function for other drug classes. These differences are due to the varied pharmacological mechanisms of action of each drug class. Psychostimulants such as cocaine and amphetamine block monoamine reuptake transporters at nerve terminals to directly increase levels of dopamine and other monoamines, whereas other abused drugs indirectly activate mesocorticolimbic pathways (Fig. 1.24–4). Opioids, such as oxycodone and heroin, act on μ-opiate receptors in the VTA to decrease the activity of inhibitory GABA interneurons, which results in a greater release of dopamine in forebrain regions. Other drugs, including nicotine and cannabinoids, activate their respective receptors, which cause

disinhibition or excitation of dopamine neurons. Drugs with more complex pharmacological profiles, including alcohol, also lead to increased dopamine release, albeit through actions at a variety of receptor subtypes, including serotonergic, glutamatergic, and GABAergic receptors. While somewhat controversial, there are also theories that the rewarding effects of some of these substances (e.g., opiates, phencyclidine, alcohol, benzodiazepines) may be regulated by actions directly upon the medium spiny neurons of the nucleus accumbens, independent of dopamine.

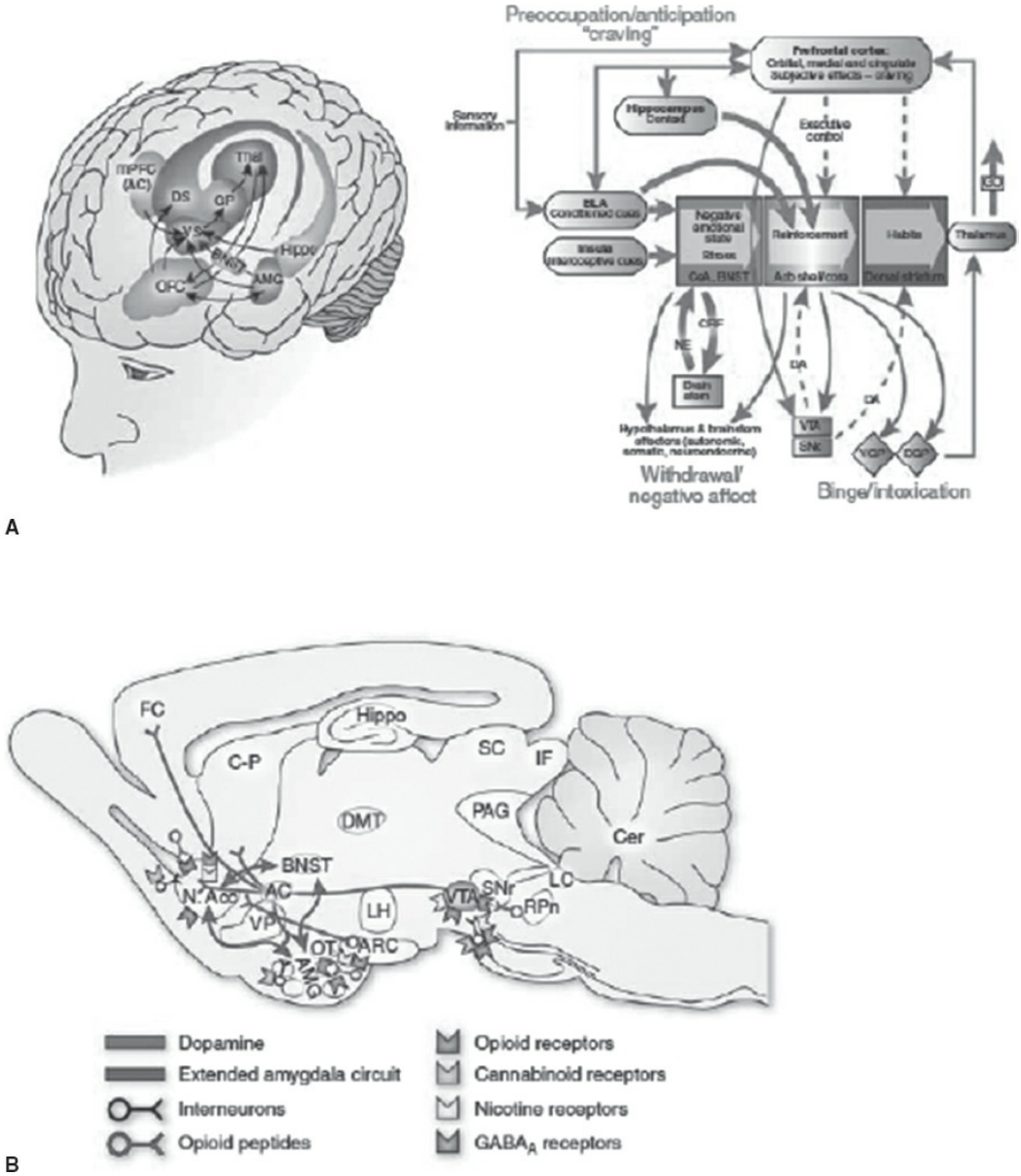

FIGURE 1.24–3. Diagram of neural circuits implicated in substance abuse and substance use disorders. **A:** Schematic illustrating the combination of neuroadaptations that are thought to occur within brain circuitry required for the three stages of the addiction cycle. The frontal cortex system is compromised, producing deficits in executive function and contributing to the incentive salience of drugs compared to natural reinforcers. Dopamine systems are compromised, and brain stress systems such as CRF are activated to reset further the salience of drugs and drug-related stimuli in the context of an aversive dysphoric state. **B:** Sagittal section through a representative rodent brain

illustrating the pathways and receptor systems implicated in the acute reinforcing actions of drugs of abuse. The *blue arrows* represent interactions within the extended amygdala system hypothesized to have a key function in drug reinforcement. AC, anterior commissure; AMG, amygdala; ARC, arcuate nucleus; BNST, bed nucleus of the stria terminalis; Cer, cerebellum; CP, caudate–putamen; DMT, dorsomedial thalamus; FC, frontal cortex; Hippo, hippocampus; IF, inferior colliculus; LC, locus coeruleus; LH, lateral hypothalamus; N Acc., nucleus accumbens; OT, olfactory tract; PAG, periaqueductal gray; RPn, reticular pontine nucleus; SC, superior colliculus; SNr, substantia nigra pars reticulata; VP, ventral pallidum; VTA, ventral tegmental area. (See color atlas.) (From Koob GF, Volkow ND. Neurocircuitry of addiction. *Neuropsychopharmacology*. 2009;35:217.)

The mesocorticolimbic dopamine pathway is consistently identified as a common neural substrate of all drugs of abuse. Indeed, studies in rodents have shown that all drugs with addictive potential increase dopamine release in forebrain structures such as the nucleus accumbens (Fig. 1.24–3). The majority of dopamine cells resides in the ventral mesencephalon, particularly the VTA, and projects to forebrain regions including the nucleus accumbens, prefrontal cortex, and amygdala. As an example of the importance of this pathway in acute drug reward, it has been shown that lesioning the VTA or nucleus accumbens prevents self-administration of cocaine in rodents. Also, opioids and alcohol are self-administered directly into the VTA, while blockade of μ-opioid receptors in the VTA suppresses heroin self-administration. In general, repeated exposure to nondrug rewards (e.g., food) activates mesocorticolimbic dopamine pathway in a consistent manner that does not result in abnormal (too much or too little) dopamine release. In contrast, drugs of abuse “hijack” this pathway such that acute drug ingestion produces dramatic and supranormative increases in dopamine release, with subsequently profound and long-lasting adaptive changes following repeated exposure to the drug and intervening withdrawal periods. The last several decades of research have identified multiple roles for dopamine: it stamps in salience to stimuli in the environment and it promotes behavioral activation and performance of goal-directed behavior. The time course of dopamine signaling is also a key factor. Rapid, phasic dopamine release tends to code the experience of reward and in determining the value of predicted outcomes of behavior, whereas tonic activation of dopamine release tends to facilitate and enable reward-related neural systems.

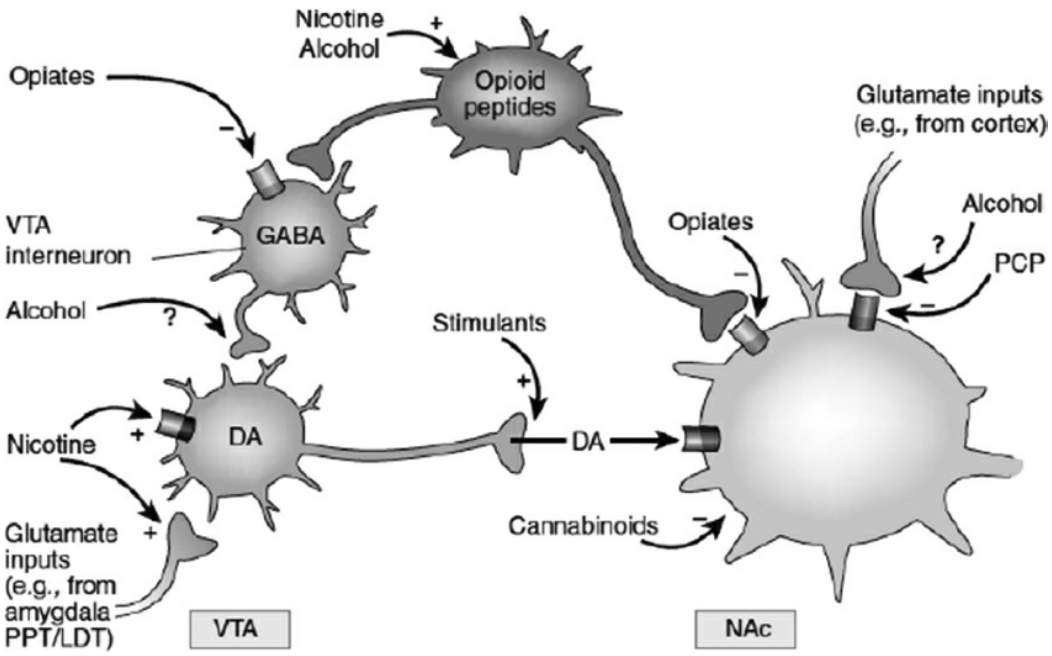

FIGURE 1.24–4. Simplified schematic of converging acute actions of drugs of abuse on the VTA–NAc pathway. Stimulants directly increase dopaminergic transmission in the nucleus accumbens (NAc). Opiates indirectly increase dopaminergic transmission in the NAc by inhibiting GABAergic interneurons in the ventral tegmental area (VTA), which disinhibits VTA dopamine (DA) neurons. Opiates also act directly on μ-opioid receptors on NAc neurons. μ-Opioid receptors, like D2 dopamine receptors, signal via the inhibitory G-protein Gαi; hence, the two mechanisms converge within some NAc neurons. The actions of other drugs of abuse remain more conjectural. Nicotine seems to activate VTA dopamine neurons directly via stimulation of nicotinic cholinergic receptors on those neurons and indirectly via stimulation of its receptors on glutamatergic nerve terminals that innervate the dopamine cells. Alcohol promotes $GABA_A$ receptor function, which may inhibit GABAergic terminals in the VTA and hence disinhibit VTA dopamine neurons. It may similarly inhibit glutamatergic terminals that innervate NAc neurons. Many additional mechanisms (not shown) are proposed for alcohol. Cannabinoid mechanisms are complex and involve activation of CB1 receptors (which, like D2 and μ-opioid receptors, are Gαι linked) on glutamatergic and GABAergic nerve terminals in the NAc, and on NAc neurons themselves. Phencyclidine (PCP) may act by inhibiting postsynaptic NMDA glutamate receptors in the NAc. Finally, there is some evidence that nicotine and alcohol may activate endogenous opioid pathways and that these and other drugs of abuse (such as opiates) may activate endogenous cannabinoid pathways (not shown). PPT/LDT, peduncular pontine tegmentum/lateral dorsal tegmentum. (From Nestler EJ. Epigenetic mechanisms of drug addiction. *Neuropharmacology*. 2014;76 Pt B:259–268.)

The nucleus accumbens is a small subcomponent of the ventral striatum that is a substrate for both dopamine-dependent and dopamine-independent drug reinforcement. The nucleus accumbens receives information related to hedonic states from glutamatergic afferents originating in the amygdala, frontal cortex, and hippocampus and converts it to motivational action via connections with the extrapyramidal motor system. Glutamate neurotransmission is tightly regulated: too much or too little can have devastating effects (e.g., cytotoxicity), whereas stimulus-dependent changes in glutamatergic transmission are necessary for learning. Indeed, a long-standing concept is that the nucleus accumbens is

a "motivation to movement interface." The nucleus accumbens itself can be subdivided into core and shell regions whose functional connectivity and neuronal phenotypes reflect this functional integration of motivation to movement. The shell has an important role in acute drug reward, motivated behavior, behavioral sensitization, and changes in affective state. The core is a central portion of the ventral striatum and is a functional continuation of the neighboring, motor-related dorsal striatum.

Neural Pathways Underlying Tolerance and Withdrawal

When drugs are used repeatedly over time, tolerance may develop. Tolerance is defined as a reduced sensitivity to the effects of the drug (e.g., euphoria) such that a higher dose of the drug is required to achieve the same level of response achieved initially. Importantly, the development of tolerance is not in itself addiction. In the context of addiction, tolerance engages mechanisms that serve to oppose reward and, paradoxically, reinforce drug-seeking behavior. Opponent process theory has been used as a framework to understand tolerance and the transition from occasional drug use to addiction: immediate rewarding effects of a drug of abuse are followed by delayed aversive effects that are due to drug-dependent recruitment of opponent processes (Fig. 1.24–2). These aversive effects emerge in force during drug withdrawal and include acute and protracted aversive states including anxiety, anhedonia, irritability, and depressed mood that resemble MDD. The mechanisms underlying tolerance to drugs and opponent processes occur at the molecular, cellular, and circuit levels and can be classified as within- or between-system adaptations.

A primary example of a "within-system" neural adaptation to chronic drug administration is hypofunction of the mesocorticolimbic dopamine system during withdrawal, which leads to decreased motivation for nondrug-related stimuli and increased sensitivity to, and motivation for, the abused drug (Fig. 1.24–3A). Consistent with hypofunction of mesocorticolimbic dopamine systems, psychostimulant withdrawal in humans is associated with fatigue, depressed mood, and sluggishness, and in animals is associated with decreased motivation to work for natural rewards and decreased locomotor activity. Similarly, withdrawal from chronic opioids essentially unmasks all the neural adaptations the brain produced in its attempts to equilibrate in the presence of drug. Consequently, neural circuits regulating everything from gastrointestinal function to affective states are instantly unbalanced, and an opioid withdrawal syndrome emerges. A primary cause of affective withdrawal signs such as anxiety, dysphoria, depression, and irritability, is thought to be the dramatic reduction in dopamine neuron firing and dopamine release in efferent targets such as the nucleus accumbens.

Despite its long tenure as the "reward center" of the brain, increasing evidence supports the idea that the nucleus accumbens is a bivalent

structure that processes positive and negative emotional stimuli into either approach or avoidance behavior. Electrophysiological data collected over the last several years has shown that the nucleus accumbens may encode aversive states via increases in firing rates of intrinsic medium spiny neurons, and that individual nucleus accumbens neurons can track the hedonic valence of a stimulus by varying their firing-rate response to it. This has important ramifications for understanding addiction, since drugs of abuse provide hyperbolic positive (drug "high") and negative (drug withdrawal, "crash") emotional signals to the nucleus accumbens. Neural circuits other than the mesocorticolimbic system also play critical roles in somatic and psychological withdrawal syndromes. These include norepinephrine, corticotropin releasing factor (CRF), orexin, dynorphin, substance P, vasopressin, NPY, endocannabinoids, orphanin, and many others.

Between-system neuroadaptations also contribute to tolerance and withdrawal syndromes. These involve neurochemical systems different from those directly targeted by drugs of abuse. The extended amygdala, which is composed of the central nucleus of the amygdala, BNST, and a medial portion of the nucleus accumbens shell, represents a common anatomical substrate that integrates stress and reward systems to produce drug withdrawal-induced negative affective states and promote negative reinforcement processes associated with the development of SUD (Fig. 1.24–3A). Engagement of the extended amygdala is one mechanism by which the brain attempts to restore normal function despite the chronic presence of drug. The stress peptide CRF mediates stress responses of both the HPA axis and the extended amygdala, and CRF is released during withdrawal from all major drugs of abuse. In summary, drug withdrawal is associated with within-system decreases in mesocorticolimbic dopamine activity and between-system recruitment of stress-like neurotransmitter systems.

Chronic exposure to drugs of abuse increases the activity of CRF and dynorphin systems within mesocorticolimbic and extended amygdala circuits. In general, CRF and dynorphin produce states that have characteristics of both withdrawal and stress. Dynorphin binds to and activates kappa-opioid receptors (KORs) and CRF binds to and activates CRF type 1 and 2 receptors (CRF_1-R, CRF_2-R) in the brain. KOR activation in mesolimbic structures including the VTA and nucleus accumbens decreases dopamine transmission, whereas the effects of CRF are more complex. On balance, evidence supports a serial CRF to dynorphin directional relationship between the two stress peptide systems such that stress triggers CRF release, which then functions to increase dynorphin release and subsequent activation of the KOR system within neural stress circuits. Withdrawal from drugs of abuse can be conceptualized as a stressor that causes a dysphoric state encoded by CRF and dynorphin

release. This is supported by studies showing that CRF infusions into the VTA can trigger reinstatement of drug seeking—perhaps through activation of dopamine neurons—and CRF antagonists reduce stress- but not drug-induced reinstatement of drug seeking. Similarly, KOR antagonism blocks the development and expression of cocaine withdrawal-induced anhedonia and attenuates the expression of nicotine and morphine withdrawal signs in rats. KOR-induced dysphoria is thought to enhance the rewarding valence of drugs of abuse—although empirical evidence centers primarily on psychostimulants. This role of KORs in addictive behavior has been expanded to include a role in stress-induced reinstatement of drug seeking. In rodents, KOR antagonism and genetic deletion of the dynorphin gene block stress-induced, but not cocaine prime-induced, reinstatement of drug seeking.

Neural Pathways Underlying Substance Use Disorder

Impaired Control (Impulsive ã Compulsive Use). It has been hypothesized that the transition from voluntary (goal-oriented) drug abuse to habitual (not goal oriented) and compulsive drug abuse is encoded through drug-dependent transitions from prefrontal cortical to striatal and from ventral to dorsal striatal control over drug-seeking and drug-taking behavior (Fig. 1.24–5). The basic organization of cortical inputs to the striatum comprises a functional gradient from limbic to associative to motor domains. A similar gradient is found in midbrain to nucleus accumbens connections. There is topographical organization such that a medial (i.e., shell) to lateral (i.e., core/dorsal striatum) series of projection loops allows emotion-based information from limbic-associated structures to transfer to motor-related areas of the basal ganglia. It has been conceptualized that the nucleus accumbens (ventral striatum) is an especially critical substrate for the rewarding drug effects that contribute to the development of SUDs, whereas the dorsal striatum is an especially critical substrate for the motor habits that maintain SUDs. A provocative corollary of this conceptualization is that once motor habits are established, whether or not a drug continues to be perceived as rewarding becomes irrelevant. This phenomenon of a transition from voluntary to habitual drug use as SUD develops is broadly consistent with functional neuroimaging studies in humans showing that the ventral striatum is implicated in associative conditioning and the dorsal striatum in habit learning. In summary, the connections among prefrontal cortex, the striatum, and midbrain dopamine neurons are organized in an ascending spiral that creates a feed-forward organization from reward-related regions of the striatum to cognitive and motor areas. Importantly, the neural networks underlying the transition from voluntary to habitual drug abuse are drug-dependent. For example, it has been shown in heroin self-administration studies that information in ventral medial cortical to ventral

and dorsal striatal circuits may be processed in parallel rather than in a spiraling, feed-forward series. As such, it is possible that heroin intake remains more goal-directed, whereas cocaine intake becomes more habitual. This is consistent with the finding that animals given unlimited access to cocaine rapidly escalate their intake and can overdose, whereas animals given unlimited access to heroin maintain a relatively stable pattern of intake.

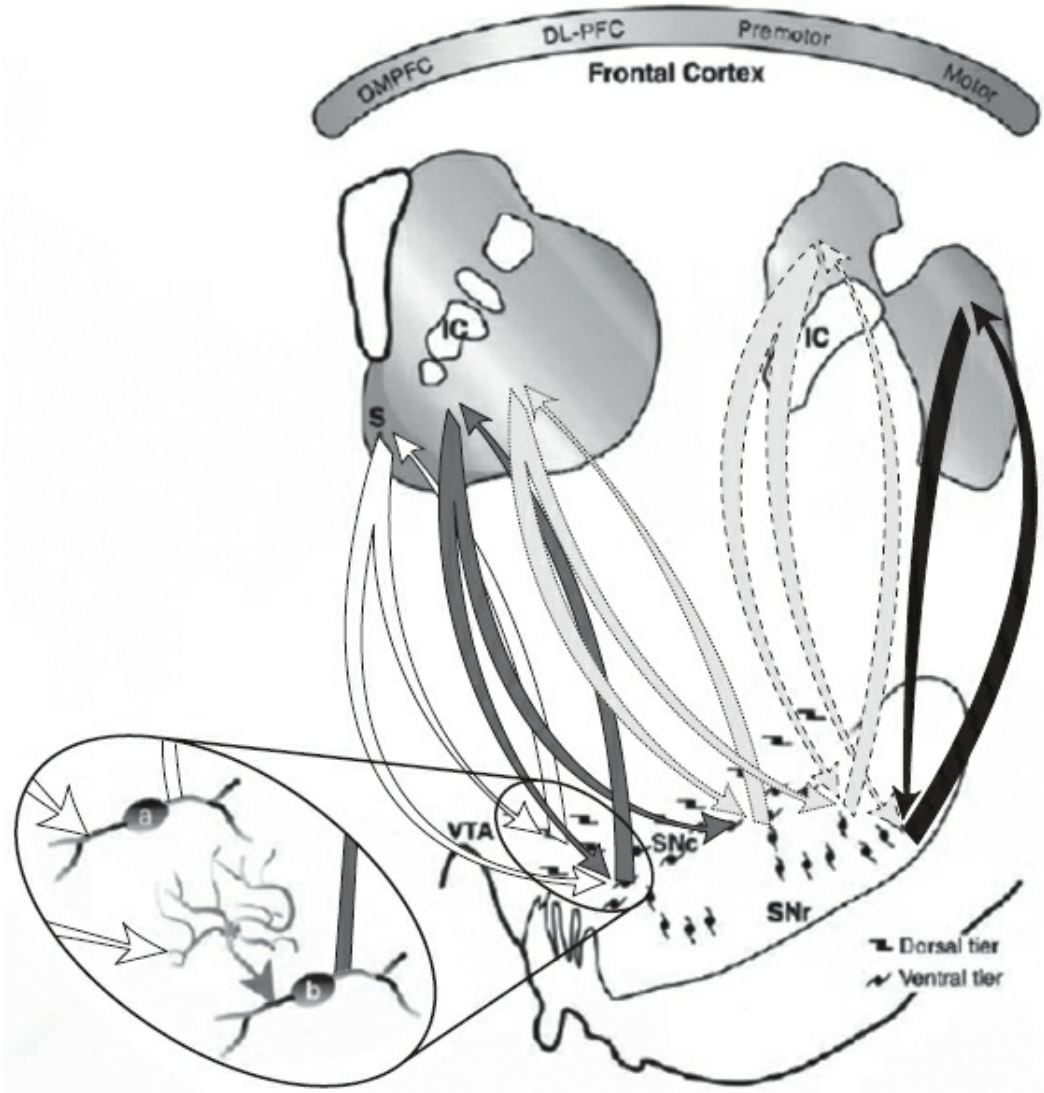

FIGURE 1.24–5. Diagram of the organization of striatonigralstriatal (SNS) projections that comprise spiraling limbic–motor loops. The colored gradient in rostral and caudal schematics of the striatum illustrates the organization of functional corticostriatal inputs (*white* = limbic, *light gray* = associative, *black* = motor). The shell receives forebrain input primarily from the amygdala, hippocampus, and cortical areas 25 and Ia. The core receives input from the entire orbital and medial prefrontal cortex (OMPFC). The dorsolateral prefrontal cortex (DL-PFC) projects to the central striatum and premotor and motor cortex project to the dorsolateral striatum. Midbrain projections from the shell target both the ventral tegmental area (VTA) and ventromedial substantia nigra pars compacta (SNc) (*white arrows*). Midbrain projections from the VTA to the nucleus accumbens shell form a "closed," reciprocal SNS loop (*white arrow*). Projections from the medial SN feed-forward to the nucleus accumbens core forming the first part of a spiral (*dark gray arrow*). The

spiral continues through the SNS projections (*light gray arrows* and *dotted arrows*) with pathways originating in the core and projecting more dorsally (*black arrows*). In this way ventral striatal regions influence more dorsal striatal regions via spiraling SNS projections. Magnified oval region shows a hypothetical model of the synaptic interactions of SNS projections in reciprocal versus feed-forward loops. The reciprocal component (*white arrows*) of each limb of the SNS projection terminates directly (a) on a dopamine cell, resulting in inhibition. The nonreciprocal, or feed-forward, component (*dark gray arrow*) terminates indirectly (b) on a dopamine cell via a GABAergic interneuron (*light gray cell*), resulting in disinhibition and facilitation of dopaminergic cell burst firing. IC, internal capsule; S, nucleus accumbens shell; SNr, substantia nigra, pars reticulata. (From Haber SN, Fudge JL, McFarland NR. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsal striatum. *J Neurosci*. 2000;20(6):2369–2382.)

Craving. Drug craving is an important construct in addiction, and it is now included in the DSM-5 as a clinical symptom of SUD. Craving can be defined as an acquired desire characterized by wanting to use drug in that moment. Drug craving is evident after chronic drug use along with the associated increased salience of drug and drug cues. Clinical studies show that acute increases in anxiety or distress that occur during periods of drug abstinence or withdrawal do not in themselves predict drug relapse as proposed in models of negative reinforcement. Rather, concomitant increases in aversive states and drug craving predict subsequent drug taking and relapse.

Craving is difficult to study in animal models of SUD as it is a highly personal and subjective, emotion-based state. Recent advances in neuroimaging techniques and ecological momentary assessment approaches to studying craving and drug use episodes in real time in individuals with SUD has allowed identification of neural circuits involved in craving. It has been shown that drug-cue and stress-induced craving for cocaine, nicotine, methamphetamine, and opiates activates regions of the prefrontal cortex, amygdala, hippocampus, insula, and the VTA. In human studies, cocaine-dependent females show greater reactivity in these brain regions to stress and drug cues during drug craving whereas cocaine-dependent males show greater responses to drug cues. In contrast, alcohol-dependent individuals show a hyperactivity in the ventral striatum and the VMPFC during neutral, relaxed states that is blunted during stress- and drug cue-induced craving. Indeed, it has been shown that hypoactivation of the VMPFC and the insula during stress is predictive of time until alcohol relapse as well as the severity of relapse.

Relapse. A major problem in the treatment of SUD is relapse to drug use after periods of detoxification and abstinence. Unlike craving, which is difficult to model in laboratory settings, the phenomenon of relapse is well modeled in laboratory animals using the reinstatement procedure (described above). Evidence from animal studies suggests that neural circuit changes associated with drug- and cue-induced reinstatement after extinction of operant responding for drug are linked to glutamatergic

projections from the prefrontal cortex to the nucleus accumbens core, dopamine projections from the VTA to the medial prefrontal cortex, and GABA projections from the nucleus accumbens to the ventral pallidum. In contrast, stress-induced reinstatement of drug seeking depends more on activation of CRF and norepinephrine systems within the extended amygdala.

Drug-Primed Relapse. Small doses of an abused drug (often called "priming") can initiate subjective states of drug desire that prompt renewed drug consumption in humans. Early studies indicated that dopamine neurons in the VTA and glutamatergic neurons in the prefrontal cortex (prelimbic, anterior cingulate, media) that project to the nucleus accumbens shell and core, respectively, play critical roles in cocaine- and heroin-induced reinstatement of drug seeking. Indeed, a key feature of the circuitry underlying cocaine and heroin seeking is enhanced glutamate release into the nucleus accumbens core by afferents from the prelimbic cortex. Inhibiting AMPA-type glutamate receptors or activating presynaptic GluR2 autoreceptors in the nucleus accumbens core blocks both cocaine- and cue-induced cocaine seeking. Dopamine transmission within mesocorticolimbic circuitry is also required for drug-primed relapse as blockade of dopamine D1 receptors in the medial prefrontal cortex decreases reinstatement. Additional brain areas have been implicated in drug-primed reinstatement through reversible, pharmacological inactivation studies. These include the ventral subiculum, ventral pallidum, and central and basolateral amygdala.

Converging evidence from all domains of neuroscience research on addiction highlights the point that the actions and consequences of different drug classes can be distinct from one another. As a case in point, the neural circuitry that underlies heroin-primed reinstatement of heroin seeking has both distinct and overlapping components. Compared to cocaine-primed reinstatement, a more diffuse circuit appears to be engaged during heroin-primed reinstatement because inactivation of multiple cortical and limbic structures abolishes reinstatement of heroin seeking. Some of these heroin-specific brain regions include the substantia nigra, BNST, and mediodorsal thalamus. Furthermore, the role of dopamine has distinct roles. It has been shown that blockade of either D1 or D2 dopamine receptors in the nucleus accumbens prevented heroin-, but not cocaine-primed reinstatement of drug seeking.

Conditioned Cue-Induced Relapse. In the animal model of conditioned-cue-induced reinstatement of drug seeking, several lines of research have implicated cortico-striato-limbic pathways in the development and maintenance of drug–cue associations that drive drug-seeking behavior after periods of withdrawal. Of particular interest has been the amygdala,

which is well established as a critical structure in the learning of affectively relevant stimuli for both appetitive and negative reinforcers. Studies have found that permanent excitotoxic lesions or temporary pharmacological inhibition of the basolateral amygdala does not effect cocaine-taking during daily cocaine self-administration, but abolishes the reinstatement of cocaine-seeking produced by cocaine-paired cues long after the cessation of cocaine self-administration. Additional studies have demonstrated that amygdala modulation of conditioned-cued reinstatement is dopamine-dependent. Blockade of dopamine D1 receptors in the amygdala blocks cue-induced reinstatement, whereas increasing dopamine signaling in the amygdala during cue presentation potentiates cocaine seeking.

Discrete subregions of the prefrontal cortex (anterior cingulate and prelimbic cortex) and striatum are also involved in conditioned cue-induced reinstatement. Pharmacological inactivation of the dorsal medial prefrontal cortex, the lateral orbitofrontal cortex, or the nucleus accumbens core attenuates cue-induced cocaine seeking. This amygdala, prefrontal cortex, and nucleus accumbens core circuitry in the animal model of cue-induced relapse shows striking homology with data from humans. Brain imaging studies in cocaine-dependent human subjects have shown that cocaine-paired cues can increase metabolic activation of the amygdala, the anterior cingulate region of the cortex, the nucleus accumbens, and the orbitofrontal cortex.

While limited in scope, imaging studies on the neural circuitry of cue-induced relapse in humans have generally found overlapping patterns of brain activation of cortical and limbic structures across various drugs of abuse including methamphetamine, alcohol, opiates, and nicotine. In rodent models, the amygdala is necessary for both heroin- and methamphetamine-paired cue-induced relapse in animal models. However, a more diffuse circuitry is engaged during heroin-paired cue reinstatement, suggesting mechanistic differences that can be best parsed out in animal models.

In humans, craving and relapse can occur after prolonged periods of abstinence. For unknown reasons, drug craving waxes and wanes over time. In an attempt to account for temporal oscillations of craving and the persistent capacity for relapse, it was hypothesized that cue-induced drug craving progressively increases over the first weeks of abstinence and can remain high over extended periods. This effect was first described in the context of alcoholism, and called the “alcohol deprivation effect.” More recently, this effect has been more broadly conceptualized as “incubation of drug craving” and can be modeled in rodents, as shown by time-dependent increases in cue-induced cocaine or heroin seeking that occur after withdrawal from the drug. Many of the same brain regions implicated in cue-induced reinstatement of drug seeking (discussed above) also contribute to incubation of drug craving. Increased neuronal activity has

been demonstrated in the ventral medial prefrontal cortex, nucleus accumbens core, and central nucleus of the amygdala when cocaine craving —as assessed in reinstatement tests after 1 or 30 days of drug withdrawal (extinction)—is measured.

Stress-Induced Relapse. Stress plays a role in all facets of drug abuse and SUD. On balance, people with an SUD—even when currently abstinent —have an increased behavioral sensitivity to stress, which supports the clinical relevance of negative emotionality for protracted abstinence and relapse. By definition, stress is the nonspecific response of an organism to any demand (stressor) placed upon it that requires adaptation. Stressors typically engage the HPA axis to produce physiological and emotional responses. The biological effects of stress are relatively hardwired and can be measured objectively, whereas emotional effects can depend on an individual's perception of his/her ability to cope with the stressor and are therefore more subjective. A popular conceptualization of stress is that it is a constellation of negative, depressive-like emotions (anhedonia, anxiety, irritability, feeling overwhelmed) that should be avoided, although stimuli that would seem to qualify as stressors can be experienced as positive or even thrilling (e.g., new job, skydiving). Controlled laboratory studies in human drug addicts have shown that drug craving can be elicited with stressors and that this stress-induced response predicts relapse. The fact that the same stressor may trigger pleasure in one individual but aversion in another has made it very difficult to develop blanket "anti-stress" pharmacotherapeutics as a panacea for addiction treatment.

Some of the same circuitry required for drug-primed or conditioned cue-induced reinstatement of drug seeking is also necessary for stress-induced reinstatement, including the prelimbic cortex and nucleus accumbens as well as structures within the extended amygdala. Inactivation of the central, but not basolateral, amygdala and the BNST will attenuate stress-induced reinstatement. For intermittent footshock-induced reinstatement, studies indicated a role of CRF and glutamate in the VTA and a role of CRF projections from the central amygdala to the BNST. Central infusions of CRF produce reinstatement, whereas CRF receptor and ionotropic glutamate receptor antagonists block stress-induced reinstatement of cocaine and heroin seeking, as do noradrenergic α_2-receptor agonists.

Numerous other neurotransmitter systems known to play a role in stress responses likely contribute to stress-induced relapse. For example, enhanced vulnerability to stress during withdrawal and protracted withdrawal may be due in part to hypodopaminergia in the mesocorticolimbic system. Drug dependence and early drug withdrawal has been associated with decreased dopamine levels and dopamine receptors in limbic areas such as the nucleus accumbens. Reduced tonic

dopamine levels during withdrawal may provide a better contrast, or "signal-to-noise ratio," to a phasic dopamine signal in response to a stressor. As discussed earlier, the neuropeptide dynorphin is increased with chronic drug administration and acts at KORs on dopaminergic nerve terminals to decrease dopamine release. This raises the possibility that chronically elevated dynorphin expression and release could contribute to drug withdrawal-induced hypodopaminergia and increased stress sensitivity. Consistent with this, blockade of KORs in the VTA with the selective antagonist nor-BNI decreases swim stress-induced reinstatement of cocaine seeking in rats. A summary schematic of the shared and distinct neurocircuitry for the reinstatement of drug-seeking behavior produced by conditioned cues, drug priming, and stress is illustrated in Figure 1.24–6.

CELLULAR AND MOLECULAR SUBSTRATES OF ADDICTION

Defining SUD as a relapsing CNS disorder accompanied by adaptations in the brain that can persist beyond detoxification and are responsible for maintenance of drug taking as well as craving and relapse occurring during periods of abstinence requires identifying and understanding such neural adaptations. This has proven to be difficult, as chronic exposure to drugs of abuse elicits a plethora of cellular and molecular adaptations that vary depending on the type and route of administration of the drug, the brain region studied, developmental history, and biological sex to name a few. A current struggle is to understand how different drug classes (e.g., psychostimulants and opioids) can produce similar behaviors but seemingly different molecular and cellular adaptations. Following are a few representative examples of drug-induced neuroadaptations and current theories on their functional significance. This is a rapidly changing field of study, with extraordinarily high stakes: development of effective SUD treatments.

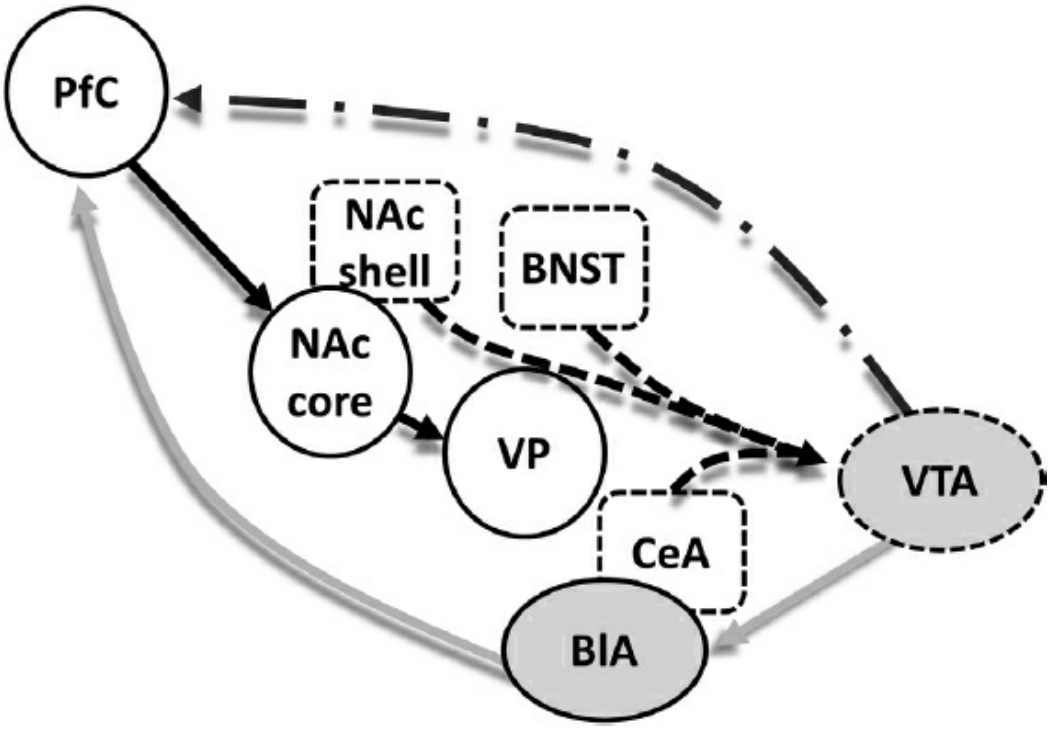

FIGURE 1.24–6. Neural circuitry mediating drug seeking. Serial projections from the prefrontal cortex to the nucleus accumbens core to the ventral pallidum (*white circles*) is a final common pathway for stress-induced drug seeking initiated within extended amygdala circuits (*white, dashed line* boxes), by drug-associated cues (*gray ovals*), or by the drug itself (*dash-dot line* from ventral tegmental area to prefrontal cortex indicating increased dopamine release). PfC, prefrontal cortex; NAc, nucleus accumbens; VP, ventral pallidum; BNST, bed nucleus of the stria terminalis; CeA, central amygdala; BlA, basolateral amygdala; VTA, ventral tegmental area.

Gene Expression

The majority of studies examining drug-induced molecular changes have focused on the mesocorticolimbic system (VTA, nucleus accumbens, prefrontal cortex) and the consequences of altered dopamine signaling. As discussed above, a common initial action of all drugs of abuse is an increase in extracellular dopamine levels in the nucleus accumbens. Dopamine binds to and activates D1- or D2-like G-protein coupled dopamine receptors, which leads to either stimulation or inhibition, respectively, of adenylate cyclase and cAMP production. As a molecular second messenger, cAMP has multiple intracellular effects, including activation of cAMP-dependent PKA and subsequent phosphorylation of an activating site on response element binding protein (CREB), a transcription factor (Fig. 1.24–7). As such, a reliable acute effect of most drugs of abuse is modest D1 receptor-mediated CREB phosphorylation in the nucleus accumbens. More importantly for the development of SUD, chronic administration of drugs of abuse produces compensatory adaptations in cAMP-PKA signaling pathways. Repeated exposure to cocaine, amphetamine, or opiates significantly induces CREB activity in the nucleus accumbens, whereas alcohol and nicotine reduce CREB phosphorylation in this region. Drug-

induced modulation of CREB activity is not limited to the nucleus accumbens. For example, acute morphine binds to inhibitory μ-opioid receptors and inhibits CREB phosphorylation in the locus coeruleus, a brain region important for opiate physical dependence and withdrawal. This inhibition undergoes tolerance with chronic morphine exposure due to compensatory upregulation of adenylate cyclase–mediated cAMP production and activation of PKA. Upon morphine withdrawal, this upregulated cAMP/PKA pathway results in a spike in CREB activation, which is thought to contribute to withdrawal signs. Finally, there is evidence that chronic administration of several drugs of abuse activates CREB in the VTA, an effect that directly impacts dopamine synthesis and excitability of VTA dopamine neurons. CREB induces expression of tyrosine hydroxylase (TH), the rate-limiting enzyme in dopamine biosynthesis, in the VTA. Despite increased TH gene expression in the VTA, decreased levels of TH protein and/or activity have been reported in VTA nerve terminals in the nucleus accumbens under certain experimental conditions. Decreases in TH protein levels could mediate chronic drug-induced hypodopaminergia and emergence of negative affective states.

Although CREB phosphorylation is a transient effect on the order of minutes, repeated brief activations of CREB triggers numerous changes in the neuronal transcriptome that in turn cause significant alterations in the physiology of individual cells as well as the circuits in which they are embedded. Some examples of CREB-regulated genes that likely play important roles in SUD include other transcription factors such as c-Fos, intracellular messengers (e.g., adenylate cyclase VI. VIII), neurotransmitter biosynthetic enzymes (e.g., TH), peptide neuromodulators (e.g., CRF and somatostatin), growth factors (e.g., BDNF), opioid peptides (e.g., dynorphin and enkephalin), and neurotransmitter receptor subunits (e.g., GluR1) (see Fig. 1.24–7). The behavioral effects of chronic drug-induced activation of CREB in the nucleus accumbens tend to be similar across drugs. Increased CREB activity in the nucleus accumbens decreases behavioral responses to cocaine, opiates, and alcohol, whereas decreased CREB activity increases such responses. This finding, which has been observed in many experimental systems with a variety of methods, has led to the hypothesis that nucleus accumbens CREB activation is a molecular opponent process engaged to counteract the continued presence of drug. In support of this hypothesis, CREB-induced dynorphin activates KORs on VTA dopamine neurons and their nerve terminals resulting in decreased dopaminergic tone and opposition of drug reward in the presence of drug or production of negative affective states in the absence of drug (i.e., withdrawal).

In addition to CREB activation, drugs of abuse can activate other transcription factors, including members of the Fos protein family (c-fos, FosB, Fra-1, and Fra-2) and NAC-1 in the nucleus accumbens. Direct

involvement of these gene products in acute drug reward is doubtful, as this depends upon more immediate signaling events associated with dopamine and opioid receptor stimulation. ΔFosB, a highly stable form of FosB, has been shown to accumulate over time with repeated administration of virtually all classes of drugs of abuse. Animals in which ΔFosB is activated have been shown to have exaggerated sensitivity to the rewarding effects of morphine and cocaine. This has led to the hypothesis that ΔFosB may be a sustained molecular "switch" that helps to initiate and maintain a state of addiction. As with CREB, ΔFosB itself does not contribute to cellular functioning, but facilitates transcription of numerous genes that encode proteins involved in neural plasticity. Taken together, the current understanding of how drugs of abuse modulate gene expression comes primarily from studies that focus on one drug class (e.g., psychostimulants or opioids). A postmortem study in people with cocaine and heroin use disorder showed that out of approximately 39,000 gene transcripts that were investigated in the nucleus accumbens, only 25 genes had altered expression in both cocaine and heroin users. For about half of these genes, the drugs had opposite effects on expression.

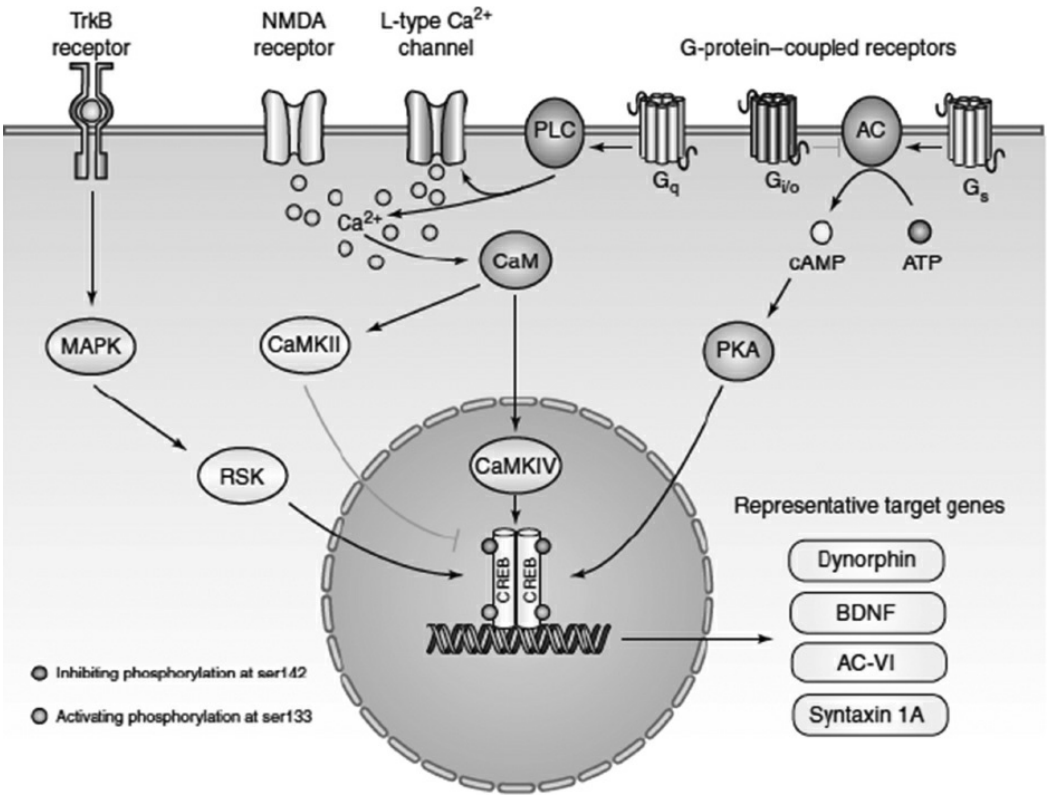

FIGURE 1.24–7. Simplified depiction of how extracellular signals influence gene transcription via CREB. Phosphorylation of CREB at ser133 increases its transcriptional activity and target gene expression. CREB activity can be attenuated by phosphorylation of CREB at ser142. Within the nucleus accumbens (NAc), neurons expressing Gαs-coupled receptors (e.g., dopamine D1) activate adenylate cyclase (AC), whereas neurons expressing Gαi/o-coupled receptors (e.g., dopamine D2) inhibit AC. Exposure to drug-induced increases in dopamine in the NAc can also trigger activation of MAPK- and calcium- ($Ca2^{+}$)

sensitive signal transduction pathways. For clarity, cross talk between second messengers is not included. *Black lines* signify activation; *gray lines* signify inhibition. (From Muschamp JW, Carlezon WA, Jr. Roles of nucleus accumbens CREB and dynorphin in dysregulation of motivation. *Cold Spring Harb Perspect Med*. 2013;3(2):a012005.)

A final example of how drugs of abuse regulate gene expression is the finding that addictive drugs upregulate BDNF in the VTA, nucleus accumbens shell, and amygdala, where it is thought to act as a "proaddictive" modulator. BDNF promotes presynaptic axonal branching and organization as well as postsynaptic dendritic arborization, number, size, and motility of dendritic spines, and enhances the synthesis of synaptic proteins. BDNF is increased by acute drug administration, but BDNF protein levels continue to increase in the nucleus accumbens during withdrawal from cocaine self-administration, with maximum levels observed after approximately 30 days of withdrawal. Elevated BDNF can remain elevated for months. The temporal pattern of BDNF accumulation maps onto the phenomenon of incubation of drug craving (as discussed above in Conditioned-cue Induced Relapse), and intra-VTA or nucleus accumbens shell BDNF injections cause delayed and long-lasting increases in cocaine seeking after withdrawal. However, it has been shown that BDNF projections from the prefrontal cortex to the nucleus accumbens reduce vulnerability to cocaine seeking. This provides another important example of how the behavioral effects of proteins and other modulators depend on the neural circuits within which they act.

Neuronal Morphology and Synaptic Plasticity

In addition to dopamine- and opioid receptor–mediated changes in gene expression within mesocorticolimbic pathways after repeated drug exposure, the transition from drug use to addiction involves persistent restructuring of cortical circuitry. Drug-induced structural plasticity includes changes in the size of cell bodies and changes in dendritic arborizations or spine morphology. The most consistent and recognizable cellular adaptations observed are morphological changes in dendritic spine density. These changes are most likely due to drug-induced regulation of genes that encode cytoskeleton regulatory proteins, including decreased levels of Homer 1 and postsynaptic density 95 (PSD95), both of which are scaffolding proteins associated with the postsynaptic actin cytoskeleton. In addition, morphine and cocaine have been shown to decrease RhoA, Rac1, and cell division cycle 42 (Cdc42), which are small GTPases that regulate the actin cytoskeleton.

A current paradox in addiction research involves the generalization that opiates decrease, whereas psychostimulants increase, the number and complexity of dendritic spines on neurons in the nucleus accumbens, medial prefrontal cortex, and on hippocampus pyramidal neurons (Fig.

1.24–8). Likewise, opiates decrease, whereas psychostimulants increase, the soma size of VTA dopaminergic neurons. One hypothesis toward reconciliation of these different cellular effects of opiates and psychostimulants involves the finding that both repeated morphine and cocaine administration increase actin cycling. This can theoretically lead to increases or decreases in spine density, but more importantly it is associated with greater plasticity in responses to drug-induced changes in glutamate release.

Despite the apparent opposite effects of different drug classes on neuronal morphology and synaptic strength, an implicit assumption has been that neuroadaptations that occur independently of drug class are particularly important for core elements of SUD that exist across drug classes. Relapse is a core element of SUD, and it has been shown that, regardless of drug class, cue-induced reinstatement is accompanied by a rapid, transient synaptic potentiation of glutamatergic synapses in the core of the nucleus accumbens. In addition, enlargement of dendritic spine heads and increases in AMPA-mediated glutamatergic transmission are observed. A relatively recent discovery is that the extracellular matrix constitutes an environment replete with neurotransmission modulators—many of which are likely involved in addiction. As one example, matrix metalloproteinases (MMPs) are a family of extracellular proteases that strongly modulate synaptic plasticity. Drugs of abuse and reinstatement to drug seeking rapidly induce MMP activity and subsequent enlargement of dendritic spine heads and increased synaptic strength in the nucleus accumbens core. Likewise, inhibiting MMP activity prevents reinstatement to drug seeking. Finally, a relatively consistent finding is that drug seeking is associated with increased glutamate release in the nucleus accumbens. The mechanisms underlying enhanced glutamate release include downregulation of cystine–glutamate exchange in glia, decreased signaling through presynaptic mGluRs, and reduced elimination of glutamate through glutamate transporters.

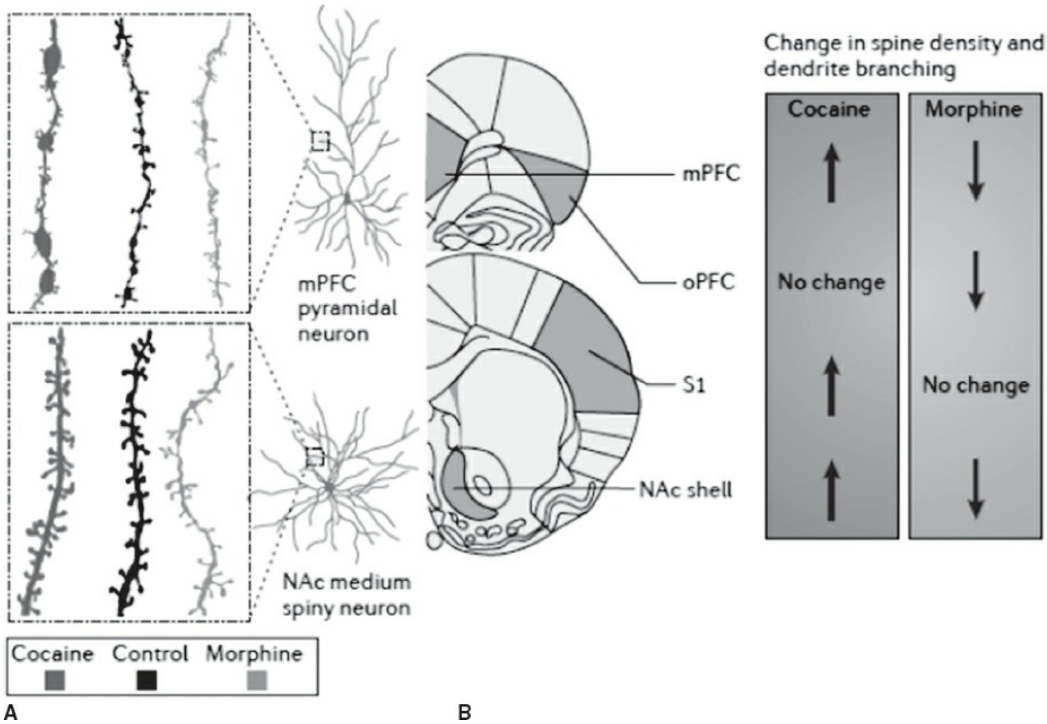

FIGURE 1.24–8. Morphine and cocaine have opposite effects on structural neuroplasticity in the nucleus accumbens (NAc) and medial prefrontal cortex (mPFC). **A:** Groups of rats were trained to self-administer morphine or cocaine intravenously for several weeks. The control groups were given daily intravenous infusions of vehicle for the same period of time. After 1 month of withdrawal from the drugs, the rats' brains were processed using the Golgi staining procedure. Rats that were exposed to cocaine showed increased dendritic branching and increased spine density in both NAc medium spiny neurons and mPFC pyramidal neurons. By contrast, rats that were exposed to morphine had both reduced dendritic branching and reduced spine density in these brain regions. **B:** A summary of changes in spine density and dendritic branching that occur after exposure to cocaine or morphine relative to controls. A dissociation between the effects of cocaine and morphine was also observed in the orbital prefrontal cortex (oPFC) and in the primary somatosensory cortex (S1). (From Badiani A, Belin D, Epstein D, Calu D, Shaham Y. Opiate versus psychostimulant addiction: the differences do matter. *Nat Rev Neurosci*. 2011;12:685.)

Epigenetics

Epigenetics can be broadly defined as a series of biochemical processes through which changes in gene expression are achieved throughout the life of an animal without a change in DNA sequence. Epigenetics is a means through which environment and genes can interact to produce enduring structural and functional changes underlying experience-dependent neuronal development and plasticity, as well as addiction-related behaviors. DNA is organized into chromatin, and the fundamental unit of chromatin is the nucleosome, which consists of DNA wrapped around a core group of histones. Epigenetic mechanisms control inter-nucleosome spacing and the degree to which they are condensed, which is a determinant of gene activity. Heterochromatin refers to the condensed, inactivated state, which does not allow gene transcription, and

euchromatin refers to the activated, open state, which allows individual genes to be transcribed. There are three major types of epigenetic regulation: histone modifications (e.g., acetylation and methylation), DNA methylation, and regulatory noncoding RNAs.

Epigenetics likely plays a major role in addiction. Repeated exposure to drugs of abuse or other environmental insults can engage epigenetic processes in early life (or adulthood). These epigenetic processes act at specific genes or sets of genes resulting in stable changes in gene expression or changes in the inducibility of other genes that can lead to addiction in vulnerable individuals (Fig. 1.24–9A). It is also possible that drugs or other environmental insults produce epigenetic changes in gametes, which can be passed on to offspring and increase their vulnerability to SUDs. There is increasing evidence supporting this idea of transgenerational epigenetic inheritance of vulnerability to addiction, but it remains controversial. Epigenetics is a nascent field, with much still to learn. As such, two salient examples of epigenetic changes are described below.

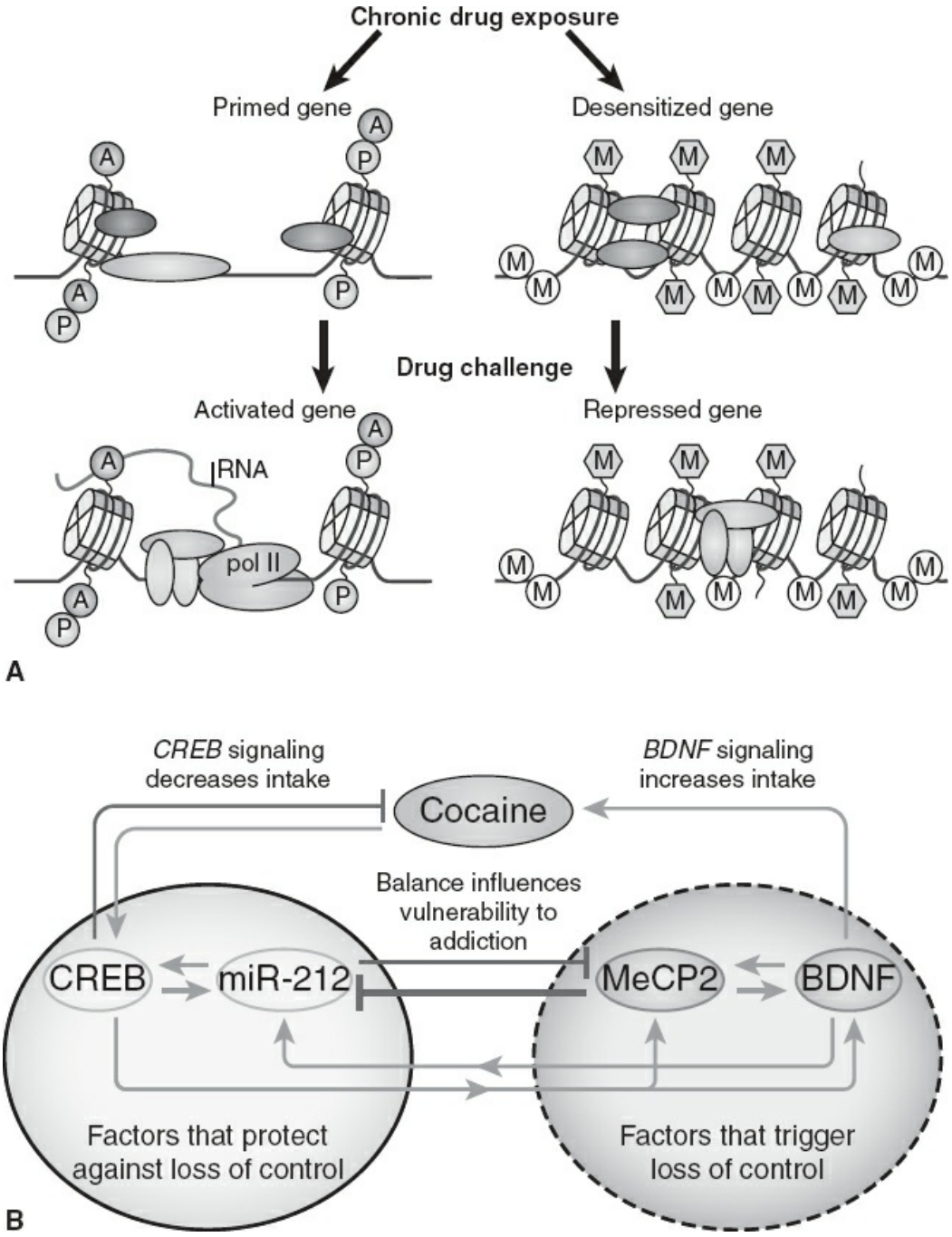

FIGURE 1.24–9. A: Epigenetic mechanisms implicated in acute and chronic drug effects on gene priming and desensitization. Many epigenetic changes are latent, meaning that they are not reflected by stable changes in steady-state mRNA levels. Instead, such changes alter chromatin structure, so that a later drug challenge induces a given gene to a greater (primed) or lesser (desensitized) extent based on the epigenetic modifications that are induced by previous chronic drug exposure. A, acetylation; M, methylation; P, phosphorylation; pol II, RNA polymerase II. **B:** The complex interactions between micro RNA 212 (miR-212), cAMP response element-binding protein (CREB), methyl CpG-binding protein 2 (MeCP2), and brain-derived neurotrophic factor (BDNF) and their putative effects on cocaine intake. The *black outlined circle* implies that CREB signaling protects against the development of escalating cocaine intake, whereas the *dotted circle* implies that MeCP2-BDNF signaling promotes escalation of intake. *Light gray arrows* indicate a stimulatory relationship, whereas *dark gray lines* indicate an inhibitory relationship. Cocaine is shown to activate both CREB and MeCP2-BDNF signaling, and the balance between these two pathways, coordinated by miR-212, likely regulates escalation of cocaine intake in extended

access rats, and perhaps vulnerability to cocaine addiction. (**A,** from Robison AJ, Nestler EJ. Transcriptional and epigenetic mechanisms of addiction. *Nat Rev Neurosci*. 2011;12:632. **B,** from Jonkman S, Kenny PJ. Molecular, cellular, and structural mechanisms of cocaine addiction: A key role for microRNAs. *Neuropsychopharmacology*. 2013;38:198.)

Histone Acetylation. Histone acetyltransferases (HATs) catalyze acetylation and HDACs catalyze deacetylation of histones. Acetylation promotes decondensation of chromatin and—in general—increases gene activity. The bulk of addiction research on histone acetylation has been with psychostimulants, which increase total cellular levels of acetylated histones in the nucleus accumbens. The acetylation state of individual genes in response to drug exposure is starting to be examined. Acetylation of the c-Fos gene promoter is acetylated briefly after drug treatment, with no observable chronic changes, which is consistent with the rapid on–off expression pattern of c-Fos mRNA. In contrast, BDNF promoters become acetylated only after chronic cocaine, which is consistent with the accumulation of BDNF protein only after chronic cocaine and during withdrawal. Unfortunately, most genes that show altered histone acetylation in response to cocaine do not exhibit changes in mRNA expression expected with hyper- or hypoacetylation. Furthermore, the behavioral effects of histone acetylation are difficult to parse out because each gene promoter has multiple sites at which epigenetic regulation can occur.

Micro RNAs. Micro RNAs (miRNAs) are small, noncoding RNA transcripts (~22 nucleotides) that bind to target mRNAs and either inhibit their translation or induce their degradation. To date, hundreds of miRNAs have been identified, and each miRNA can potentially target hundreds to thousands of mRNA transcripts. In a series of studies focused on the miRNA miR-212 (Fig. 1.24–9B), it was shown that extended access to cocaine self-administration in rats increases levels of miR-212 in the striatum. Experimenter-controlled overexpression of miR-212 in the striatum reduces the motivation to self-administer cocaine specifically in rats given extended access to the drug, but had no effect on self-administration of cocaine in rats with restricted drug access. CREB increases the expression of miR-212, which can itself dramatically increase CREB activity in the striatum. Although BDNF is a CREB target gene (see above), elevations in miR-212 actually repress BDNF expression in the striatum. Taken together, the epigenetic regulator, miR-212, is increased with chronic cocaine self-administration, activates CREB, and suppresses BDNF. These findings suggest that miR-212 may protect against the development of cocaine addiction.

SEX DIFFERENCES IN SUBSTANCE USE DISORDERS

In humans there are numerous sex differences in almost every facet of SUD

(see Table 1.24–4). For most drugs of abuse, men consistently have higher rates of use, abuse, and dependence compared to women, with an overall odds ratio of 2:1 (men:women). In contrast, data regarding prescription opioids are strikingly equivocal: based on data from several large epidemiological surveys, women are equally or more likely than men to use prescription drugs nonmedically. Furthermore, accumulating evidence suggests that women progress from substance use to SUD faster than men, a phenomenon termed "telescoping." Sex differences have also been found in drug sensitivity and self-administration behaviors in laboratory animal studies. In general, females are more vulnerable than males to the reinforcing effects of psychostimulants, opiates, and nicotine during each phase of the addiction cycle: acquisition, maintenance, escalation, and relapse.

Table 1.24–4.
Summary of Human Studies on Sex and Ovarian Hormone Influences on Drug Abuse. ↑ Small, ↑↑ Moderate, ↑↑↑ Large Effect

Dependent Measure/Phase	Drug	Finding*
Subjective effects	Cocaine	Women ↑↑ nervousness
		Women ↓↓ euphoria & dysphoria
		↑↑ high during follicular
	Amphetamine	↑↑↑ euphoria & high during follicular
		↑ pleasant stimulation with estrogen
Levels/patterns of use	Cocaine	No sex difference in total levels
		Women ↓ abstinence
	Opiates	No sex difference in total levels
		No sex difference in patterns of use
Relapse	Cocaine	Women ↑ cue reactivity
		Women ↑↑↑ craving during abstinence

*Adapted from Lynch WJ, Roth ME, Carroll ME. *Psychopharm*. 2002;164:121. References for Findings found therein.

The biological mechanisms that underlie sex differences in substance use and SUD are still being elucidated. While the basic neural systems involved in positive and negative reinforcement processes that contribute to SUD development are similar in males and females, sex differences have been described in the organization and connectivity of these systems. When thinking about how sex can influence SUD, it is important to keep in mind that sex differences can be occur at behavioral and/or neural levels. In other words, sex differences in behavior imply sex differences in underlying mechanism, but it is also true that sex differences in neural mechanisms may underlie similar behavioral outputs. Given that rigorous examination of the role of sex in SUD is relatively recent, there are still many gaps in the current knowledge. It has become clear that factors such as drug type,

experimental paradigm, animal species, and estrous cycle contribute to behavioral and neurobiological responses to drugs of abuse. As such, much of the data currently available are widely variable. Following are two relatively robust examples of putative mechanisms underlying sex differences in substance use and addiction.

Activational Effects of Gonadal Hormones. Evidence from both preclinical and clinical studies demonstrates that fluctuations in gonadal sex hormones, particularly estrogen and progesterone, over the course of the female reproductive cycle (Fig. 1.24–10) play important roles in sex differences in the response to (at least) psychostimulants. For example, several of the reinforcing effects of amphetamine, such as euphoria and increased energy are potentiated in women during the follicular phase compared to the luteal phase and are positively correlated with estrogen levels. Consistent with this, administration of exogenous estrogen during the follicular phase enhances the subjective effects of psychostimulants in women. In contrast to the potentiating effects of estrogen, progesterone levels are negatively correlated with the subjective effects of psychostimulants. Furthermore, administration of exogenous progesterone during the follicular phase decreases many of the positive subjective effects of cocaine. The use of animal models allows a more precise examination of the effects of gonadal hormones on behavior because gonadectomy can eliminate the source of hormones, which can then be reintroduced with experimental control of dosing and timing. Using these methods, it has been shown that ovariectomy reduces cocaine self-administration, and exogenous estrogen treatment is sufficient to facilitate acquisition of cocaine, morphine, and heroin self-administration in ovariectomized rats. Interestingly, there is no effect of castration on acquisition of cocaine self-administration in male rats.

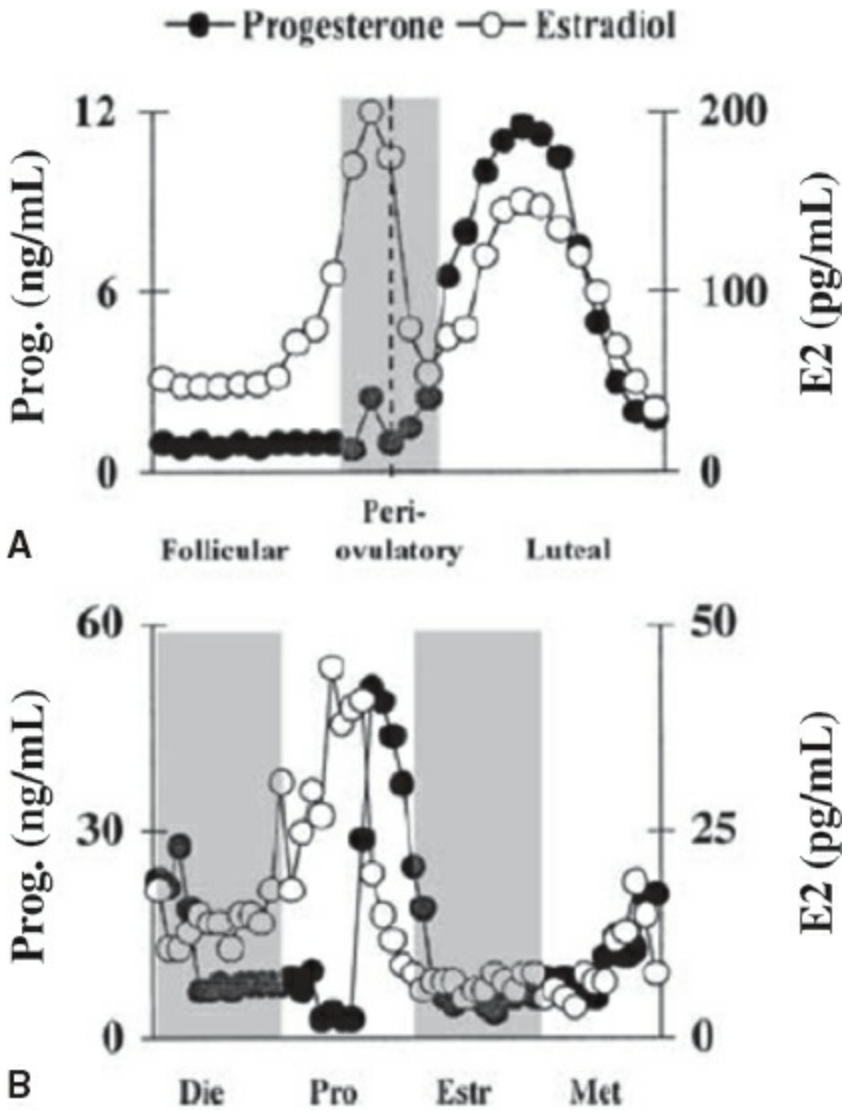

FIGURE 1.24–10. Gonadal hormone fluctuations throughout the female reproductive cycle. **A:** Changes in levels of estrogen (E2) and progesterone (Prog) as a function of days from the preovulatory luteinizing hormone (LH) surge for the human menstrual cycle. The shaded bar separates the periovulatory phase from the follicular and luteal phases. **B:** Changes in levels of estrogen and progesterone throughout the phases of the rat estrous cycle. The *shaded bars* separate the successive estrous cycle phases diestrus (DIE), proestrus (PRO), estrus (ESTR), and metestrus (MET) to identify the start and end of each phase. (From Lynch WJ, Roth ME, Carroll ME. Biological basis of sex differences in drug abuse: preclinical and clinical studies. *Psychopharmacology*. 2002;64:121.)

Sex Differences in Dopamine Signaling. In a second example of a possible mechanism underlying sex differences in addiction, the number and function of midbrain dopamine neurons has been reported to be sexually dimorphic. For example, women have been reported to have a greater capacity for dopamine synthesis and reuptake through the DAT relative to men. Consistent with this, female rats have more dopamine neurons in the VTA than males. Gonadal hormones can regulate the density of dopamine terminals in the nucleus accumbens and dorsal striatum. The density of dopamine terminals impacts basal dopamine tone,

which has been shown to fluctuate in females over the course of the reproductive cycle but is relatively stable in males. There is also evidence that—regardless of estrous cycle—females show greater basal and cocaine-stimulated dopamine release and reuptake compared to males.

CURRENT TREATMENTS

Treatment of SUD occurs in three phases: detoxification, initial recovery, and relapse prevention. After detoxification, the goal of the initial recovery phase is to learn to tolerate craving induced by external or internal cues and to learn new patterns of behavior that help the individual avoid relapse. The goal of relapse prevention treatments is for the individual to develop long-term strategies that allow them to replace addictive behaviors with new, healthy behaviors. Treatments can be pharmacological, behavioral, social, and/or cognitive. There is a general consensus among clinical and preclinical investigators that the most effective treatments will be based on scientific evidence that comes from the understanding of basic neural mechanisms underlying SUD.

Evidence-based clinical interventions may be understood as effective interventions according to the neurobiological substrates targeted during the engagement process and early recovery treatment of those with SUDs (Table 1.24–5). For example, medication interventions are designed to ameliorate allostatic states associated with chronic substance exposure and the development of physiological dependence and withdrawal syndromes. As such, stabilization of these systems includes the improvement of withdrawal symptoms, substance craving, and hyperarousal associated with early abstinence. This translates clinically to improved individual capacity for executive functioning and impulse regulation, learning and attention, lowered stress- and cue-reactivity, and social normalization.

Psychosocial interventions may be similarly conceptualized based on the components of their structure (Table 1.24–5). Psychosocial interventions include both community-based interventions such as education about SUDs and their treatments, access to exercise programs as a component of health and wellness, and mutual help peer support programs aimed at supporting abstinence (e.g., the 12-step model of recovery founded as Alcoholics Anonymous), as well as professional interventions delivered by trained clinicians such as motivational interviewing, cognitive-behavioral relapse prevention, and positive reinforcer contingency management. Typical interventions will provide for some, but not all, of the neurobiological targets thought to be important in an effective early recovery treatment. For instance, education is a prevention and treatment modality that instructs individuals about risk factors for SUDs and also provides a basic understanding of SUD as a brain-based illness that may be arrested with abstaining from substance use. As such, it will primarily assist individuals in making healthy choices,

promoting healthy social networks, and avoiding the use of substances as a means of coping with stressors. It will not have much impact on altering reward-seeking behavior and related incentive salience, nor will it help to stabilize withdrawal syndromes. An engagement intervention, known as motivational interviewing, refers to a trained process of dialoguing with an individual to help that person better define his/her goals, values, and current incentivized behavioral patterns, with a goal of assisting that person in focusing on incentivized behaviors that are consistent with personal goals and values, and diminishing incentivized behaviors (such as substance misuse) resulting in consequences that are inconsistent with personal goals and values. While motivational interviewing may successfully engage an individual in planning for change, it is unlikely on its own to significantly alter subjective reward-driven behavior and/or stress reduction, and thus motivational interviewing best serves as a platform to support engagement in other therapies, such as cognitive-behavioral relapse prevention or contingency management, which are more specifically aimed to alter these aspects during the abstinence recovery phase of treatment. Interestingly, mutual help peer support may provide the broadest neural substrate "coverage" in recovery by virtue of its being a socially facilitated system of training the individual to understand addiction cues ("people, places, and things"), avoid near occasions of substance use, develop healthier alternatives for stress reduction and natural reward, and to integrate meaningful social connection into planned recovery and meaningful life pursuits.

Following are several clinical vignettes that cover different SUD severities and drug classes. Included are the clinically recommended treatments currently available.

Table 1.24–5.
Evidence-based clinical interventions for substance use disorder and their hypothesized functional neurobiological targets. Treatment gaps exist for effective therapeutic interventions targeting the transition of substance-related reward to healthier, natural rewards as well as to the extinction of substance cue-associated incentive salience.

Treatment Modality	Executive Function (Fronto-Temporal Cortical Circuitry)	Reward Circuitry (Mesolimbic Dopaminergic Circuitry)	Stress Circuitry (Basolateral Striatum, Extended Amygdala)	Incentive Salience (Amygdala and Hippocampus)	Social Isolation (Amygdala, Fusiform Face Area, Superior Temporal Sulcus, Ventrolateral Prefontal Cortex, Perietal Cortex, Subthalamic Nucleus)
Medication	×	×	×	×	Connection with prescriber
Education	×		×		×
Cognitive-	×		×	×	×

behavioral relapse prevention					
Motivational interviewing	×			×	×
Mutual help	×	×	×	×	×
Contingency management	×	×		×	×
Exercise		×	×		×

Woman with Opioid Use Disorder

A 23-year-old single woman, no children, working part-time in a retail department store, and weekend work dancing, presents for first treatment of injection heroin use disorder. She has a significant history of childhood and adult sexual trauma. She began using substances at age 14, with cigarette smoking, drinking with friends episodically, and regular marijuana use. At age 18 her boyfriend introduced her to using illicit prescription opioids first orally and then intranasally. At age 19, he left the relationship and she experienced opioid withdrawal for the first time. This led her to increasingly risky drug-seeking behaviors, including dancing at a strip club, and at age 20 she was introduced to IV heroin use. She has maintained a daily habit of 1 to 2 g IV for years and she has not had medical or mental health care. She presents for treatment in the context of severe weight loss and having been assaulted two nights ago. She is in early opioid withdrawal and has been smoking marijuana for relief. She is a one pack-per-day cigarette smoker.

Treatment approach: In this case, this young woman has multiple life-threatening problems related to her SUDs: lack of stable housing, injection drug use, malnutrition, and sexual assault. The first step in treatment is securing environmental safety in the form of residentially contained housing to support engagement in health care and to prevent access to substances and dangerous peers. The severity of early opioid withdrawal symptoms will not be adequately diminished by cannabis use and consideration of agonist stabilization as a medication-assisted treatment (i.e., buprenorphine or methadone maintenance) would be recommended to engage her in care and to reduce opioid craving that would drive her to leave the residential setting. Effective care must also be comprehensive, and would include assessment of depression and PTSD as well as medical assessment of nutritional status, infection (HCV and HIV) and sexually transmitted diseases, and smoking cessation. Women's peer support would provide important social structure for breaking dangerous behavioral patterns and optimizing reproductive health. Vocational counseling is necessary to support independent recovery in a safe manner. In the setting of injection opioid use disorder, best practices would include a component

of contingency management to incentivize recovery-oriented behaviors, and training and access to intranasal naloxone rescue in the event of future relapse to opioid use.

Man with Cocaine Use Disorder

A 55-year-old divorced man, father of three adult children, who is referred by his primary care provider for stabilization of cocaine use disorder following medical treatment for recurrent sinus and upper respiratory infections that ultimately detected nasal perforation associated with intranasal cocaine use. He owns his own plumbing business, which has been successful and expanded over the past 20 years. He presents with significant grief, depressed mood, and passive thoughts of suicide in the context of recent loss of his homosexual partner of 5 years, with whom he has been using cocaine 3 days per week on average. His family is not aware of this and he has lack of personal supports. Recently he has been using online sites for rapid "hook ups" with other men who use cocaine, and he has begun to experience financial strain in supporting some of these evenings. He presents abstinent 1 week with severe cravings to use. He does not use other drugs, has never injected, and infrequently drinks alcohol.

Treatment approach: In this case, in addition to counseling immediate disengagement with potentially dangerous online partner-seeking sites, and assessing his environmental safety, the most imminent risk of harm in his bereaved, early abstinent state is that related to suicidal ideation, which could progress at any time to planning and suicidal behavior. Careful assessment of the need to hospitalize or otherwise contain him in a residential treatment setting is appropriate, as SUDs are associated with significantly increased suicidality. While there are currently no specific medication-assisted treatments for cocaine or other stimulant use disorders, aggressive treatment of dysphoria/depression, a hallmark of early stimulant abstinence, is warranted; consideration of bupropion would be optimal due to its weak dopaminergic activity which may contribute to amelioration of cocaine craving. Culturally sensitive therapy addressing "coming out" and recent complicated bereavement is essential for recovery, and mutual help peer support designed for homosexual men and/or for cocaine use disorder would support facilitation of healthier social networks for healing. Although family education is typically an integral component in care, some aspects concerning this case would suggest caution and consideration that the individual be further along in his recovery process prior to integrating family directly into the treatment planning of his health care.

References

Badiani A, Belin D, Epstein D, Calu D, Shaham Y. Opiate versus psychostimulant addiction: the differences do matter. *Nat Rev Neuroscie*. 2011;12:685–700.

Becker JB, Perry AN, Westenbroek C. Sex differences in the neural mechanisms mediating addiction: a new synthesis and hypothesis. *Biol Sex Differ*. 2012;3:14.

Berridge KC, Robinson TE, Aldridge JW. Dissecting components of reward: 'liking', 'wanting', and learning. *Curr Opin Pharmacol*. 2009;9:65–73.

Bossert JM, Marchant NJ, Calu DJ, Shaham Y. The reinstatement model of drug relapse: recent neurobiological findings, emerging research topics, and translational research. *Psychopharmacology (Berl)*. 2013;229:453–476.

Bruchas MR, Land BB, Chavkin C. The dynorphin/kappa opioid system as a modulator of stress-induced and pro-addictive behaviors. *Brain Res*. 2010;1314:44–55.

Carlezon WA Jr, Thomas MJ. Biological substrates of reward and aversion: a nucleus accumbens activity hypothesis. *Neuropharmacology*. 2009;56(Suppl 1):122–132.

Chartoff EH, Carlezon WA Jr. Drug withdrawal conceptualized as a stressor. *Behav Pharmacol*. 2014;25:473–492.

Dong Y, Nestler EJ. The neural rejuvenation hypothesis of cocaine addiction. *Trends Pharmacol Sci*. 2014;35:374–383.

Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci*. 2005;8:1481–1489.

Gipson CD, Kupchik YM, Kalivas PW. Rapid, transient synaptic plasticity in addiction. *Neuropharmacology*. 2014;76(Pt B):276–286.

Haber SN, Knutson B. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*. 2010;35:4–26.

Heyer MP, Kenny PJ. Corticostriatal microRNAs in addiction. *Brain Res*. 2015;1628(PtA):2–16.

*Jonkman S, Kenny PJ. Molecular, cellular, and structural mechanisms of cocaine addiction: a key role for microRNAs. *Neuropsychopharmacology*. 2013;38:198–211.

Kalivas PW, Lalumiere RT, Knackstedt L, Shen H. Glutamate transmission in addiction. *Neuropharmacology*. 2009;56(Suppl 1):169–173.

*Kalivas PW, Volkow ND. The neural basis of addiction: a pathology of motivation and choice. *Am J Psychiatry*. 2005;162:1403–1413.

Kauer JA, Malenka RC. Synaptic plasticity and addiction. *Nat Rev Neurosci*. 2007;8:844–858.

Koob GF. A role for brain stress systems in addiction. *Neuron*. 2008;59:11–34.

*Koob GF, Volkow ND. Neurocircuitry of addiction. *Neuropsychopharmacology*. 2010;35:217–238.

Lynch WJ, Roth ME, Carroll ME. Biological basis of sex differences in drug abuse: preclinical and clinical studies. *Psychopharmacology (Berl)*. 2002;164:121–137.

Muschamp JW, Carlezon WA Jr. Roles of nucleus accumbens CREB and dynorphin in dysregulation of motivation. *Cold Spring Harb Perspect Med*. 2013;3:a012005.

*Nestler EJ. Is there a common molecular pathway for addiction? *Nat Neurosci*. 2005;8:1445–1449.

Nestler EJ. Epigenetic mechanisms of drug addiction. *Neuropharmacology*. 2014;76(PtB):259–268.

Pickens CL, Airavaara M, Theberge F, Fanous S, Hope BT, Shaham Y. Neurobiology of the incubation of drug craving. *Trends Neurosci.* 2011;34:411–420.

*Potenza MN, Sofuoglu M, Carroll KM, Rounsaville BJ. Neuroscience of behavioral and pharmacological treatments for addictions. *Neuron.* 2011;69:695–712.

Reid AG, Lingford-Hughes AR, Cancela LM, Kalivas PW. Substance abuse disorders. *Handb Clin Neurol.* 2012;106:419–431.

Robison AJ, Nestler EJ. Transcriptional and epigenetic mechanisms of addiction. *Nat Rev Neurosci.* 2011;12:623–637.

Russo SJ, Dietz DM, Dumitriu D, Morrison JH, Malenka RC, Nestler EJ. The addicted synapse: mechanisms of synaptic and structural plasticity in nucleus accumbens. *Trends Neurosci.* 2010;33:267–276.

Sinha R. The clinical neurobiology of drug craving. *Curr Opin Neurobiol.* 2013;23:649–654.

Smith AC, Scofield MD, Kalivas PW. The tetrapartite synapse: Extracellular matrix remodeling contributes to corticoaccumbens plasticity underlying drug addiction. *Brain Res.* 2015;1628(Pt A):29–39.

Wise RA, Koob GF. The development and maintenance of drug addiction. *Neuropsychopharmacology.* 2014;39:254–262.

▲ 1.25 Epigenetics in Psychiatry: The Promise for New Biomarkers and Treatments

JENNIFER B. DWYER, M.D., PH.D., AND ARIE KAFFMAN, M.D., PH.D.

INTRODUCTION

In 1942 Conrad Waddington coined the term "*epigenetics*" by fusing the terms "*genetics*" and "*epigenesis*"—an old term used to describe the developmental program by which the zygote gives rise to different cell types in a multicellular organism. Epigenetics, as defined by Waddington, represented the cooperation between genetic traits and environmental conditions that together established a stable cellular phenotype. The notion that nongenetic factors play an important role in development was further formalized by defining epigenetics as "a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence." More recently, the term epigenetics has been further broadened to emphasize the contribution that environmental factors make to the formation of a stable cellular phenotype that may or may not be heritable. In this review, the latter definition was adopted, and the term epigenetics is used to describe "changes in gene expression that lead to predictable, stable, and possibly heritable alterations in cell function that are not due to changes in DNA sequence." This definition is particularly relevant in psychiatry where environmental factors, such as stress or substance abuse, alter the expression of genes in postmitotic neurons that are not heritable, but contribute to psychiatric pathologies such as major depression, anxiety, and addiction.

Table 1.25–1 summarizes some of the key similarities and differences between the "*genetic*" and the "*epigenetic*" codes. Structurally, the genetic code is made of a one-dimensional sequence of nucleotides that are joined together to form a long DNA strand. The epigenetic code, on the other hand, is composed of a host of proteins and noncoding RNAs that coat the DNA and transform it into a three-dimensional structure that can be compressed, decompressed, and twisted into loops. The genetic code is identical in all the cells in the body, while the epigenetic code is highly variable between different cell types, and may also change during different phases of the cell cycle within the same cell. Epigenetics thus comprises the mechanisms by which the same genetic blueprint can be used to generate different cell types. This concept is illustrated in Figure 1.25–1 where the same genetic code, represented as a linear sequence of letters, is selectively masked by two different epigenetic programs to read STAND (top) or SIT (bottom), depending on what letters within the underlying code are made visible. The genetic code is static whereas the epigenetic code is highly

dynamic and can change rapidly in response to environmental cues. This dynamic property allows postmitotic neurons to induce the expression of genes that modify synaptic plasticity in brain circuits that regulate behavior. Finally, the genetic and the epigenetic codes share some similarities. These include the ability to modify gene expression in a manner that can be transmitted across generations.

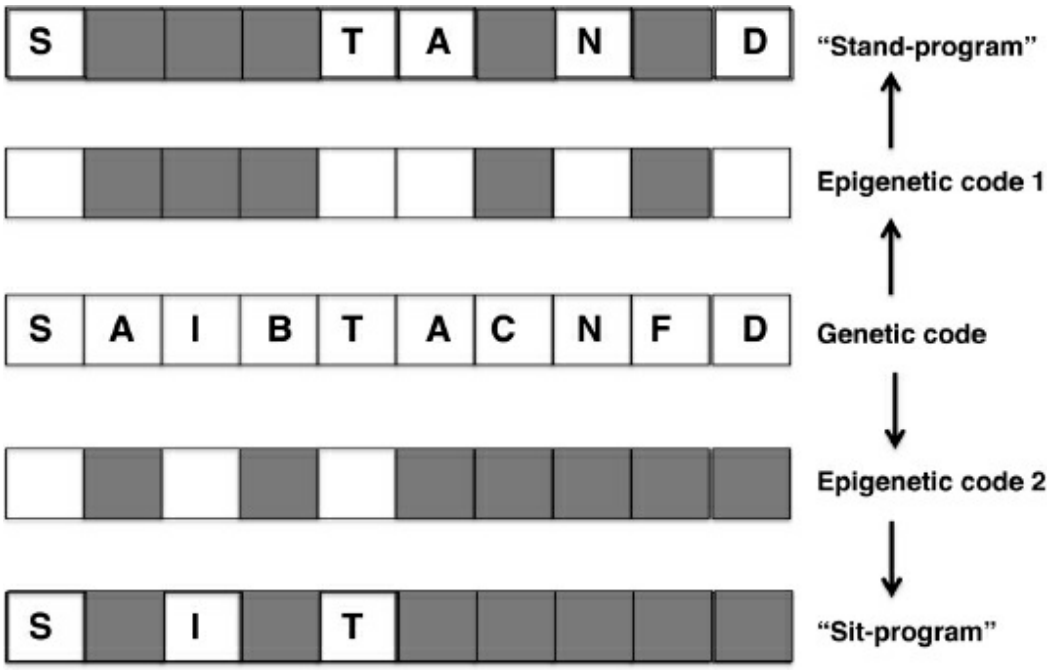

FIGURE 1.25–1. The epigenetic code modifies the genetic instructions in response to environmental demands.

Epigenetics research is one of the most exciting and rapidly growing frontiers in psychiatric research. This is due to its potential to advance the field in six specific areas (summarized in Table 1.25–2). First, epigenetics provides a molecular mechanism to explain how environmental factors shape vulnerability to psychiatric conditions. In some cases, such as substance abuse or PTSD, the contribution of the environment is quite obvious. In other examples, such as schizophrenia or bipolar disorder, a strong environmental influence can be surmised form the observation that the concordance between monozygotic twins is slightly less than 50 percent. One of the goals of this chapter is to provide specific preclinical and clinical examples in which epigenetic changes explain how specific environmental conditions modify the risk for psychopathology. It is important to recognize that environmental manipulations can also induce epigenetic changes that alleviate psychiatric symptoms. In fact, many of the known interventions in psychiatry such as antidepressant medications, electric convulsive shock (ECS), diet, and exercise have been shown to cause epigenetic changes that modify vulnerability to stress, anhedonia, and drug-seeking behaviors in animals. A better understanding of how

epigenetics mediate environmental risk and protective factors is likely to generate new and perhaps more effective strategies to treat mental illness.

Table 1.25–1.
Differences and Similarities between the Genetic and the Epigenetic Codes

Variable	Genetic Code	Epigenetic Code
Components	Long strand of DNA composed of four nucleotides (A, G, C, T)	Histones, histone-modifying enzymes, structural proteins, DNA-modifying enzymes, ATP-dependent chromatin-remodeling machinery
Structure	One-dimensional linear strand	Dynamic three-dimensional structure
Cellular content	Identical in all cells	Unique for each cell type. Dynamic, may change in response to environmental cues and cell cycle progression
Environmental impact	No change in response to environmental changes	Rapid changes in response to environmental changes, could also be long lasting
Can modify gene expression	Yes	Yes
Heritable	Yes	Sometimes

Table 1.25–2.
Epigenetics Promise in Psychiatry

Contribution	Practical Implications
1. Provides molecular mechanisms to explain how environmental change modify behavior and psychopathology	Novel therapies
2. Clarifies how abnormal neurodevelopment leads to psychopathology	More effective early interventions
3. Allows for the generation of inducible pluripotent cells	Gaining access to “brain tissue” from live patients with mental illness
4. Provides additional ways to explain how psychopathology is inherited	Novel diagnosis and interventions
5. Elucidates the mechanisms by which gene and environmental factors interact to modify vulnerability to mental illness	More specific biomarkers and interventions
6. Provides reliable biomarkers to diagnose psychiatric vulnerabilities	The availability of objective and quantifiable markers to diagnose and assess treatment response

Second, epigenetics play a central role in guiding neurodevelopment and therefore provides a possible mechanism to explain how abnormal brain development contributes to the vulnerability to many psychiatric conditions. This observation is evidenced by the large number of neurodevelopmental disorders that are caused by epigenetic dysregulation (Table 1.25–3). A better understanding of this process is likely to lead to the

development of new treatments (see The Epigenetic Etiology of Fragile X Syndrome). Moreover, a growing body of work, some of which is summarized in this chapter, demonstrates that epigenetics play a critical role in modifying brain development and the vulnerability to several psychopathologies in children exposed to abuse or neglect (discussed in detail below).

Third, epigenetics research is responsible for recent advances in generating inducible pluripotent cells. This technology allows researchers to harvest somatic cells (e.g., skin cells) from individuals with specific psychiatric illnesses (e.g., schizophrenia) and reprogram their epigenetic code in a way that transforms them into neurons or glial cells. These cells can then be grown and studied in culture. This remarkable process solves one of the key hurdles in psychiatric research, which is how to gain access to functional brain tissue.

Fourth, despite significant progress over the past decade to map genetic causes of highly heritable psychiatric conditions such as schizophrenia (81 percent) and bipolar disorder (85 percent), genetic mutations account for only a small fraction of the cases. There is a growing consensus that epigenetics may also contribute to the heritability of these disorders, because heritability is calculated as the relative rates of concordance between monozygotic and dizygotic twins, and monozygotic twins share many more epigenetic similarities than dizygotic twins. In addition, the ability of some epigenetic markers to be inherited may explain how exposure to a specific stressor in the parental generation may modify the risk for psychopathology in the offspring. An example of this form of epigenetic heritability is described in the preclinical section in which the response to olfactory fear conditioning is transmitted across generations in rodents.

Fifth, the dynamic interactions between the underlying DNA sequence and its packaging chromatin provide a molecular mechanism to explain how genes modify the ability of certain environmental conditions to induce psychopathology. Such gene-by-environment interactions have been elegantly demonstrated between different alleles of the serotonin transporter and the ability of childhood maltreatment to modify risk for adult depression. This is illustrated in the clinical section of this chapter by describing the molecular details in which the genetic and epigenetic codes coordinate the expression of FKBP5 and thus influence the risk for the development of PTSD.

Table 1.25–3.
Chromatin Mediated Neurodevelopmental Disorders

Disease	Protein/Gene	Chromatin Modifications	Clinical Features

α-Thalassemia/mental retardation syndrome, X-linked (ATR-X)	ATRX/XH2, member of SWI/SNF family of proteins (chromatin-remodeling enzyme)	Defective chromatin remodeling that is implicated in downregulating the α-globin locus	X-linked inheritance; mental retardation, hemolytic anemia, splenomegaly, facial, skeletal and genital anomalies
Angelman syndrome	Loss of UBE3A expression from the maternal chromosome	Abnormal DNA methylation and silencing (imprinting) of maternal chromosomal region 15q11–13	Cortical atrophy, cerebellar abnormal myelination, seizures, mental retardation
Coffin–Lowry syndrome	RSK2 (Kinase)	Abnormal histone acetylation and phosphorylation	X-linked inheritance; psychomotor retardation, skeletal and craniofacial deformities
Fragile X syndrome	FMR1 and FMR2	DNA hypermethylation of expanded trinucleotide repeats at the FMR1 and FMR2 promoters	Most common inherited form of intellectual disability, signs of autistic behavior, long and narrow face with large ears and macroorchidism
Immunodeficiency–centromeric instability–facial anomalies syndrome (ICF)	Dnmt3B (De-novo DNA methyltransferase)	Hypomethylation at centromeric regions of chromosomes 1, 9, and 16	Mild mental retardation, immunodeficiency, facial anomalies
Rubinstein–Taybi syndrome	CBP (HAT)	Reduced histone acetylation, autosomal dominant inheritance	Mental retardation, abnormal facial features, blunted growth
Rett syndrome	MeCP2 (MBP)	Abnormal recruitment of MeCP2 to methylated DNA	X-linked, pervasive developmental disorder, predominantly affecting girls with arrested brain development, cognitive decline, and autistic-like behavior
The 9q34 subtelomeric deletion syndrome	GLP (HMT)	Abnormal H3K9 methylation	Severe mental retardation, hypotonia, facial anomalies

Sixth, many epigenetic markers, such as DNA methylation, are stable biomarkers that can be accurately quantified using a very small number of cells. These properties make epigenetic changes a potentially valuable biomarker for diagnosing and assessing treatment response. Biomarkers are desperately needed in psychiatry where no objective disease markers are currently available to diagnose or assess treatment response. A recent finding suggesting that the methylation rate of a stress-related gene, SKA2,

can be used to predict suicidal behavior is presented in the last section of this chapter. This kind of work exemplifies the utility of epigenetics to identify useful biomarkers of psychiatric conditions.

This chapter is divided into three sections. The first section provides a comprehensive review of the cellular machinery involved in epigenetic regulation. The second section describes several preclinical examples in which epigenetic changes have been shown to modify psychiatrically relevant phenotypes in animals. The third and final section provides several clinical examples to illustrate how epigenetics may modify clinical understanding and practice in psychiatry.

THE EPIGENETICS MACHINERY

Evolutionary studies in archaebacteria suggest that chromatin was originally evolved as a mechanism to rapidly regulate gene expression in response to environmental demands. This more complex form of transcriptional regulation was later adapted to support cellular differentiation, in which the same genetic code is used to generate different cell types. With the expansion of the genome, chromatin was essential to compact the lengthy DNA so it could fit into a small nucleus. Finally, chromatin serves many other functions that are essential for maintaining the integrity of the genetic code. These include DNA repair and the formation of telomeres and centromeres.

The basic folding unit of chromatin is the nucleosome. The nucleosome consists of a histone octamer core wrapped by a DNA loop (Fig. 1.25–2A). The octamer core is assembled from four highly conserved, positively charged, histone proteins (H2A, H2B, H3, and H4). H3 and H4 histones form a tetramer that is sandwiched between two H2A/H2B dimers forming the octamer core. The octamer core is wrapped by a 146bp DNA loop, a structure that is stabilized by large numbers of ionic contacts made between the negatively charged DNA and positively charged amino acids found in histones. Each histone protein has a long unstructured N-terminal tail that protrudes and probes the surrounding space. This tail plays a critical role in the ability of the chromatin to condense or open up.

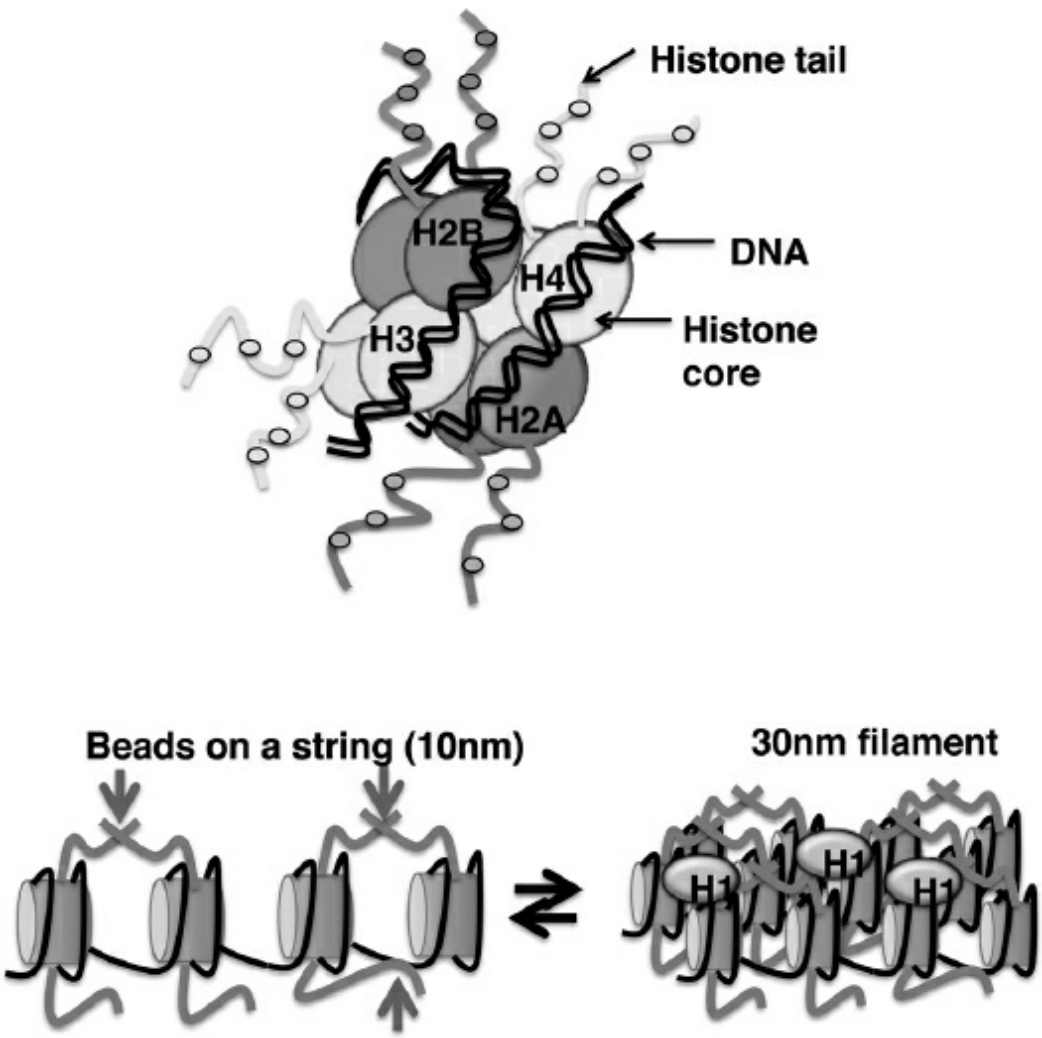

FIGURE 1.25–2. **Top:** The nucleosome as a basic folding unit of chromatin. **Bottom:** Chromatin can transition from a "beads on a string" 10-nm filament to a more compact 30-nm filament structure that is stabilized by the presence of histone H1 (*yellow circles*). Histone tails (*gray*) make contacts between neighboring nucleosomes (internucleosomal contacts) and with the underlying DNA (*red arrows*). (See color atlas.)

In the open chromatin conformation, also known as "beads on a string," neighboring nucleosomes are separated by a short (10 to 80 bp) linker DNA (Fig. 1.25–2B). The "beads on a string" chromatin structure alters the properties of the underlying DNA in two important ways. First, it compacts the DNA by about five- to tenfold. Second, significant portions of the wrapped DNA are not accessible for the binding of transcriptional activators. As a result, even in its most accessible form, chromatin represents a barrier for basal transcription. Moreover, the long N-terminal tails on histones make contacts between neighboring nucleosomes that tend to compact the chromatin. These higher-order filaments can be stabilized by the binding of a fifth histone, histone H1, forming higher-order 30-nm filaments that further compact the DNA, making it even less accessible (Fig. 1.25–2B). This natural tendency of chromatin to coil into less accessible, more compact structures is regulated by several

mechanisms that include posttranslational modifications of the histone N-terminal tails, recruitment of chromatin-remodeling machinery, and the use of invariant histones. These processes are described in some detail below.

Histones Modifications

The N-terminal tails of histones are decorated by a large number of posttranslational modifications such as acetylation, methylation, phosphorylation, ubiquitylation, and sumoylation as some of the main examples. These modifications are reversible and are carried out by large multisubunit protein complexes that coordinate their activities in a way that can either condense or open the chromatin in response to environmental cues. Histone modifications affect chromatin structure via three distinct mechanisms. First, some modifications disrupt ionic contacts between positively charged residues on histone tails and the negatively charged DNA backbone, loosening DNA–histone contacts. Second, nucleosomes contact one another via their N-terminal tails. These internucleosomal contacts can be strengthened or disrupted by posttranslational modifications. Third, alterations of the histone tails provide binding sites that recruit chromatin-remodeling machinery. These proteins can use ATP hydrolysis to slide or remove nucleosomes from chromatin (Fig. 1.25–3).

Histone Acetylation. Acetylation neutralizes lysine residues, rendering the N-terminal tails on histones less positively charged. This decrease in positive charge is thought to disrupt contacts with the negatively charged DNA and with neighboring nucleosomes. Histone acetylation is associated with transcriptional activation and other cellular processes that require access to the underlying DNA, such as DNA replication and repair (Table 1.25–4). Histone acetyltransferases (HATs) acetylate lysine residues on the tails of histones H3 and H4, while Histone deacetylases (HDACs) remove these acetylations (Fig. 1.25–4). Acetylation modifies lysines (K) that are systematically spaced on histone 3 (e.g., 9, 14, 18, and 23) and histone 4 (e.g., 5, 8, 12, and 16). Different HATs and HDACs are responsible for acetylating and deacetylating specific residues. On histone H3, for example, K14 and K18 are acetylated by CREB protein (CBP/P300), while K9, K14, and K18 are acetylated by a different HAT, PCAF/GCN5. This enzyme specificity provides the opportunity to develop more selective and targeted pharmacotherapies to act on this system. Recent work using reconstituted purified histones found that a single acetylation at lysine 14 in histone 4 (H4K14-Ac) was sufficient to maintain chromatin in the “beads on a string” state by preventing interactions between neighboring nucleosomes. Interestingly, the H4K14-Ac modification inhibited the ability of the chromatin-remodeling enzyme, ACF, to slide nucleosomes, suggesting that

this modification "locks" chromatin into an open state.

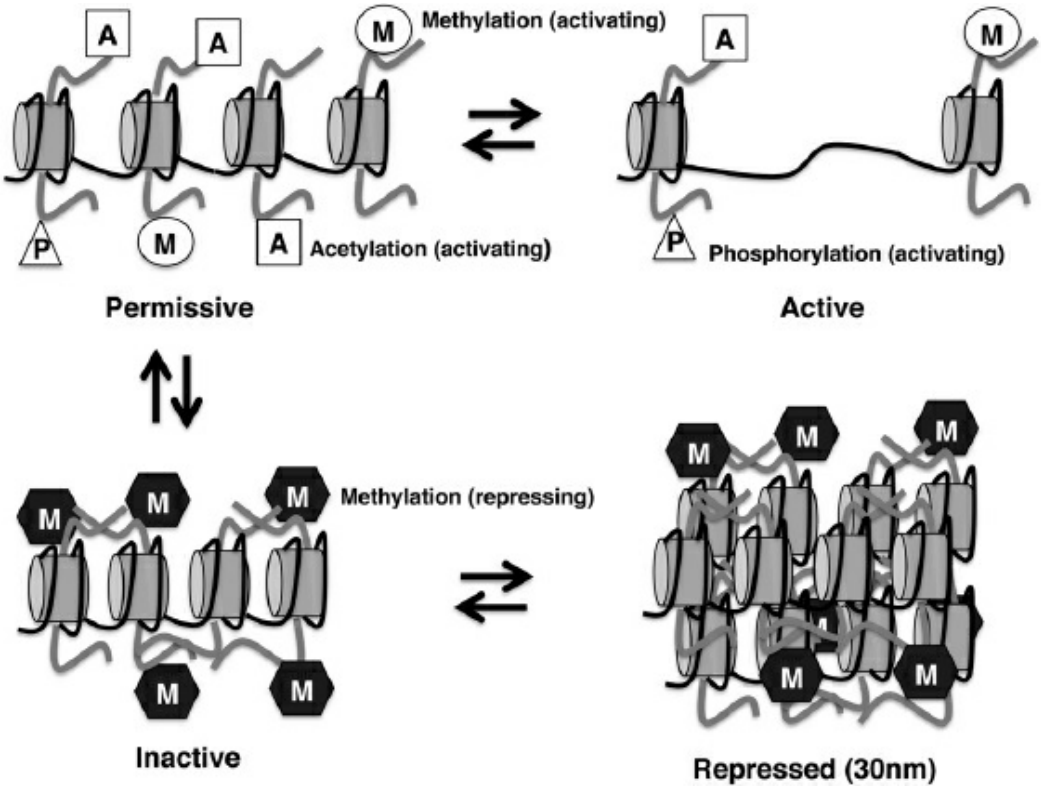

FIGURE 1.25–3. "Permissive chromatin" (**top left**) is acetylated (*open squares*), methylated (*open ovals*), and phosphorylated (*open triangle*). These histone modifications open up the chromatin by disrupting internucleosomal contacts and DNA–nucleosome contacts and recruit ATP-dependent chromatin-remodeling enzymes (not shown) that further open up the chromatin to form "active chromatin" (**top right**). Permissive chromatin can be modified to form "inactive chromatin" (**bottom left**) by removing activating histone modifications and the addition of inhibitory histone methylations (*dark hexagons*). These repressive modifications compact chromatin by promoting internucleosome and DNA–nucleosome contacts. Inactive chromatin can be further condensed to form "repressed chromatin" (**bottom right**) by recruiting ATP-dependent chromatin-remodeling enzymes and other structural proteins (not shown).

Consistent with histone acetylation opening chromatin, deacetylation of histones leads to chromatin compaction and transcriptional repression (Fig. 1.25–4). HDACs fall into three classes: Class I, Class II, and the NAD-dependent Class III of the SIR family. These classes are based on structural similarities and are therefore inhibited by different pharmacological agents. In general, HDACs seem to have less specificity than HATs with regard to the lysine resides they deacetylate. Different HDACs are targeted and bind to different compartments within the nucleus, allowing them to conduct somewhat distinct functions. For example, HDAC1 seems to play a general role in chromatin condensation and repression, while HDAC3 is tethered to the nuclear periphery where it forms highly repressive null chromatin (see Chromatin Domains).

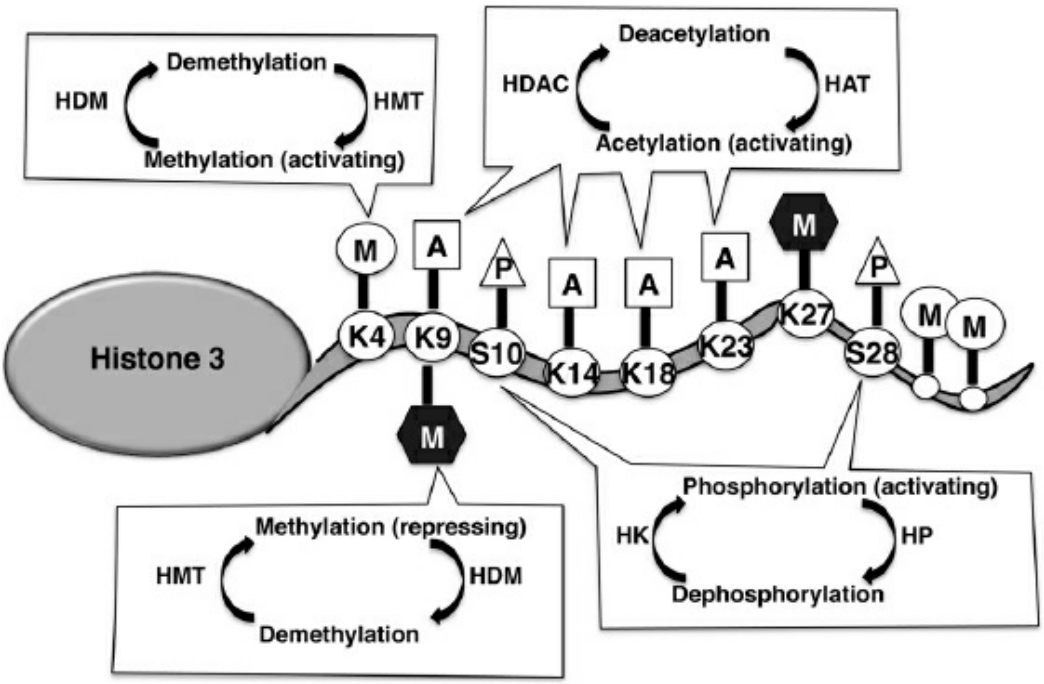

FIGURE 1.25–4. Histone tails are reversibly modified at specific residues. HAT, histone acetyltransferases; HDAC, histone deacetylase; HMT, histone methyltransferase; HDMT, histone demethyltransferase; HK, histone kinase; HP, histone phosphatase.

Proteins that have a bromodomain-folding motif are able to bind acetylated histones in order to maintain open chromatin. Mutations in the HAT domain of CBP cause Rubinstein–Taybi syndrome that is associated with mental retardation and cardiac abnormalities (Table 1.25–3). Interestingly, a mouse model of this condition showed that the cognitive deficits associated with low levels of CBP could be reversed by administrating HDAC2 inhibitors to adult mice. These findings suggest that the relative activities of HATs and HDACs in postmitotic neurons play an important role in neuroplasticity with broad implications for neurodevelopmental disorders, aging, and drug addiction (see Epigenetics in Preclinical Work). Indeed, increased HDAC2 has been shown to silence genes that are necessary for learning and memory and HDAC2 activity was found to be elevated in Alzheimer disease. Clinical work is currently underway to test whether specific HDAC inhibitors can be used to improve cognition in conditions such as early Alzheimer disease, schizophrenia, and traumatic brain injury.

Table 1.25–4.
Enzymatic Activities That Modify Chromatin

Modification	Residues Modified	Impact on Chromatin	Modifying Enzyme	Demodifying Enzyme
Histone acetylation	H3K9, H3K14, H3K18, H3K23,	Opens chromatin/transcriptional activation	HAT (GCN5, SRC-1, CBP)	HDAC (HDAC1, HDAC2, HDAC5)

	H4K5, H4K8, H4K12, H4K16			
Histone methylation (activating)	H3K4, H3K36, H3K79	Opens chromatin/ transcriptional activation	HMT (Set7/9, MLL, SUV39H/H2)	HDMT (LSD1, JMJD2A)
Histone methylation (repressing)	H3K9, H3K27, H4K20	Condensed chromatin/transcriptional repression	HMT (GLP, G9a, EZH2,	HDMT (KDM3A/ JMJD1A, JMJD2A, JMJD3)
Histone phosphorylation	H3S10, H3S28	Opens chromatin/transcriptional activation (immediate early genes)	HK (AURORA/IP1, MSK1/MSK2)	HP (DUSP1)
Chromatin-remodeling enzymes	NA	Open or condense chromatin	NA	NA
DNA methylation	Cytidine within the CpG dinucleotide	Condensed chromatin/transcriptional repression (mainly)	DNMT1, DMT3a, DNTM3b	Tet family, Gadd45 family, AID, SMUG1/TDG

Histone Methylation. Histone methylation is more complex than acetylation because both lysines (K) and arginines (R) can be methylated. Moreover, each lysine can be methylated once, twice, or three times, providing another layer of complexity. Lysine methyltransferases are more selective compared to acetylating enzymes; targeting specific lysine residues within the histone tail. Methylation on three sites is associated with transcriptional activation (H3K4, H3K36, H3K79), whereas methylation on three other lysine resides is associated with repression (H3K9, H3K27, H4K20) (see Table 1.25–3 for a summary). In addition to transcriptional regulation, histone methylation also plays a role in DNA repair.

As with other chromatin modifications, methylated lysines can be demethylated by histone demethylases, allowing for dynamic and reversible chromatin changes (Fig. 1.25–4). Proteins that contain chromo-like domains (chromo, tudor, MBT), or the structurally unrelated PHD motif are able to bind methylated histones. Such proteins form large complexes that contain chromatin-remodeling enzymes and other histone modifiers. These complexes are recruited to methylated histones and depending on the exact composition of the complex are able to efficiently open or close the chromatin. For example, appropriate activation of the HOXC8 gene is essential for normal axis development in all vertebrates. Activation of the HOXC8 gene requires histone methylation during a specific period in development. This methylation step is required to recruit the nucleosome-remodeling factor (NURF) via the bromodomain PHD finger transcription factor (BPTF). The NURF complex contains a

chromatin-remodeling enzyme (SNF2L) that removes nucleosomes from the HOXC8 promoter allowing for transcriptional activation. Mutations that interfere with the ability of BPTF to bind the methylated promoter cause abnormal expression of HOXC8, and abnormal axis formation in tadpoles, demonstrating that this histone modification plays an important role in normal axis development (Fig. 1.25–5).

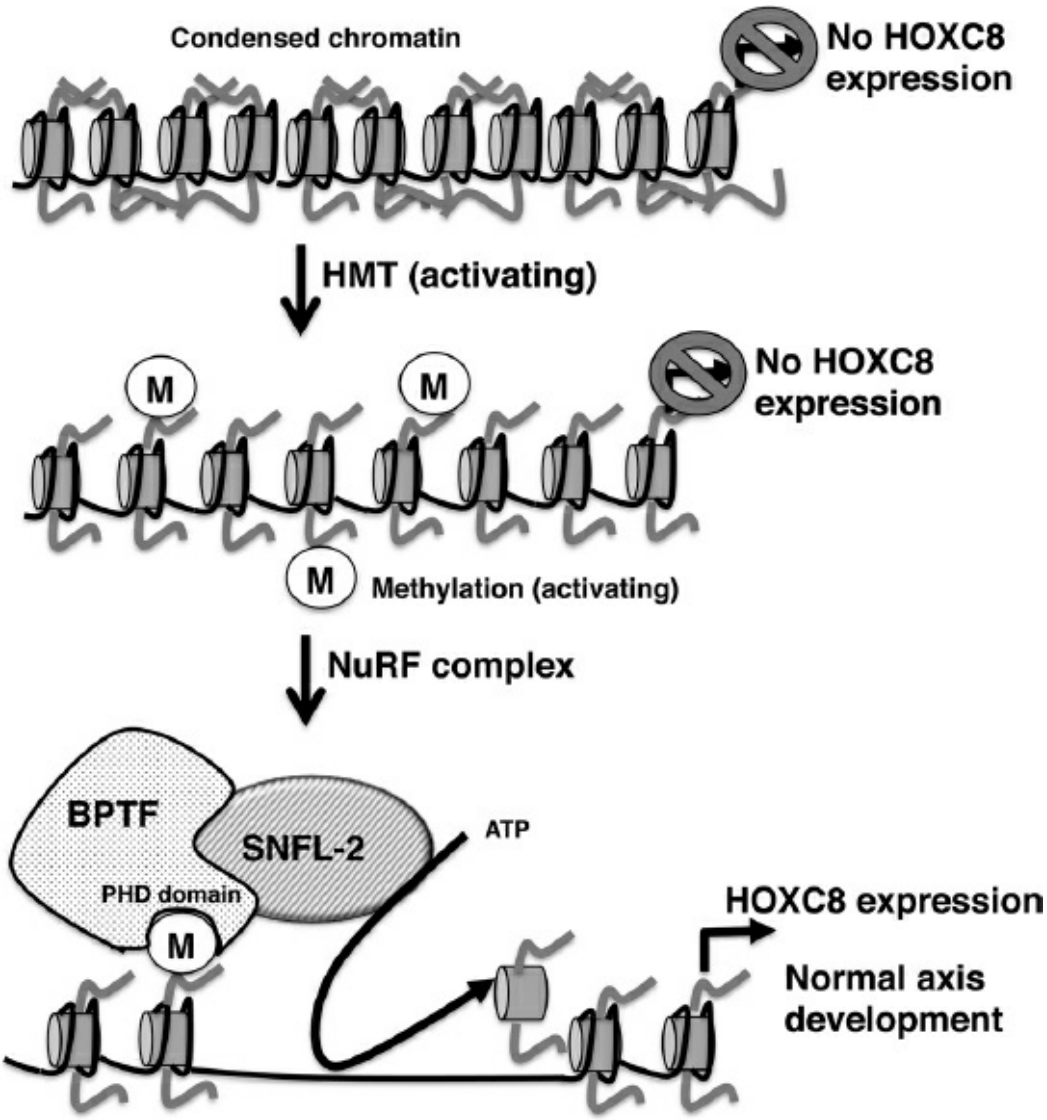

FIGURE 1.25–5. Histone methylation recruits BPTF via its PHD domain brining with it the chromatin-remodeling enzyme SNFL-2, as part of the NuRF complex. Binding of the NuRF complex to the HOXC8 promoter induces HOXC8 expression, which is necessary for normal axis development in vertebrates.

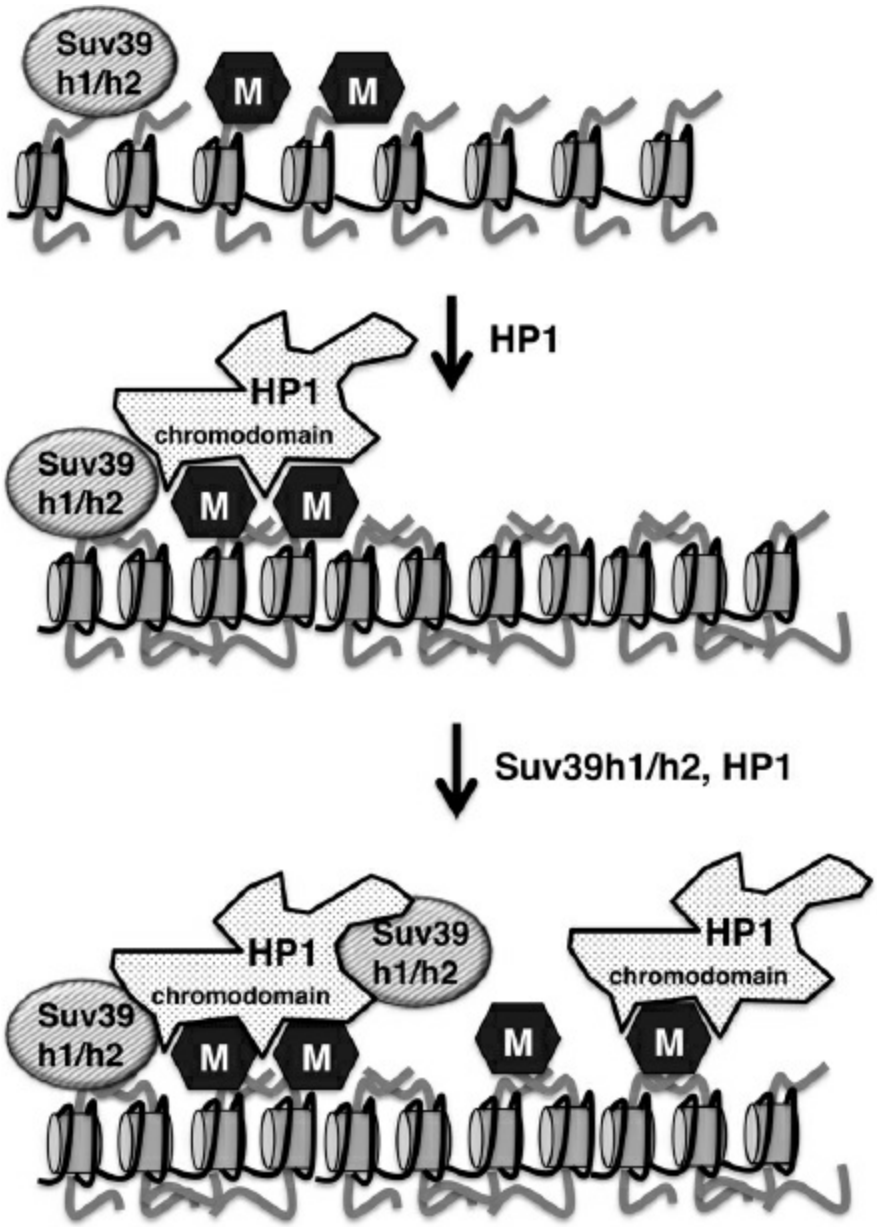

FIGURE 1.25–6. The HMT Suv39h1/h2 methylates H3K9 (**top**) recruiting HP1 via its chromodomain (**middle**). HP1 binds Suv39h1/h2 allowing for the propagation of this repressive pericentric/constitutive heterochromatin (**bottom**).

Methylation at lysine 9 in histone 3 (H3K9-me2/H3K9-me3) plays a critical role in maintaining a specialized from of chromatin known as pericentric/constitutive heterochromatin. One of the main functions of the pericentric chromatin is to support the formation of centromeres that are used to align and segregate chromosomes during cell division (Table 1.25–4). H3K9 is methylated by a specific histone methyltransferase (Suv39h1/h2), which in turn triggers the binding of the repressor protein HP1 via its chromodomain. HP1 then associates with several other proteins to form a complex that extends this specialized repressed chromatin structure through large sections of the underlying DNA. This is accomplished by the ability of HP1 to recruit additional Suv39h1/h2,

allowing further histone methylation and HP1 recruitment to propagate down the DNA (Fig. 1.25–6). HP1 is also associated with DNA methyltransferases (DNMTs), DNMT1 and DNMT3a, allowing for DNA methylation and further chromatin compaction (see DNA Methylation). Finally, during replication, HP1 is associated with chromatin factor 1 (CAF1) protein, which deposits histones at the replication fork. The presence of a complex containing HP1, Suv39h1/h2, DNMT1, and CAF1 at the replication fork provides a possible mechanism by which pericentric chromatin is inherited in daughter cells.

Histone methylation is also critical for the creation of repressive polycomb chromatin structures, which are instrumental in maintaining the pluripotency of embryonic stem cells, regulating cell cycle progression, cell differentiation, and chromosome-X inactivation. In this process, histone methylation at H3K27 is catalyzed by the Polycomb Repressive Complex (PRC2), a highly evolutionarily conserved protein complex. PRC2 is recruited to specific regions of DNA via interactions with other DNA-binding proteins and noncoding RNA. Once recruited, PRC2 methylates H3K27 to form the di and the tri methylated states (H3K27-me3), which then serve as tethering sites for the entire complex. H3K27-me3 also recruits the chromatin-remodeling machinery (PRC1) that further condenses chromatin into the highly repressive polycomb conformation (Fig. 1.25–7 and Table 1.25–4). Interestingly, the PRC2 complex maintains its binding to DNA during DNA synthesis, providing a mechanism by which H3K27-me3 and polycomb chromatin repression is maintained in daughter cells. Mutations in the histone methylase activity of PRC2 prevent the stable propagation of polycomb repression into daughter cells, suggesting an essential role for methylation in this epigenetic form of inheritance, with critical implications for many developmental programs including X-inactivation and the transition from cellular pluripotency to differentiation.

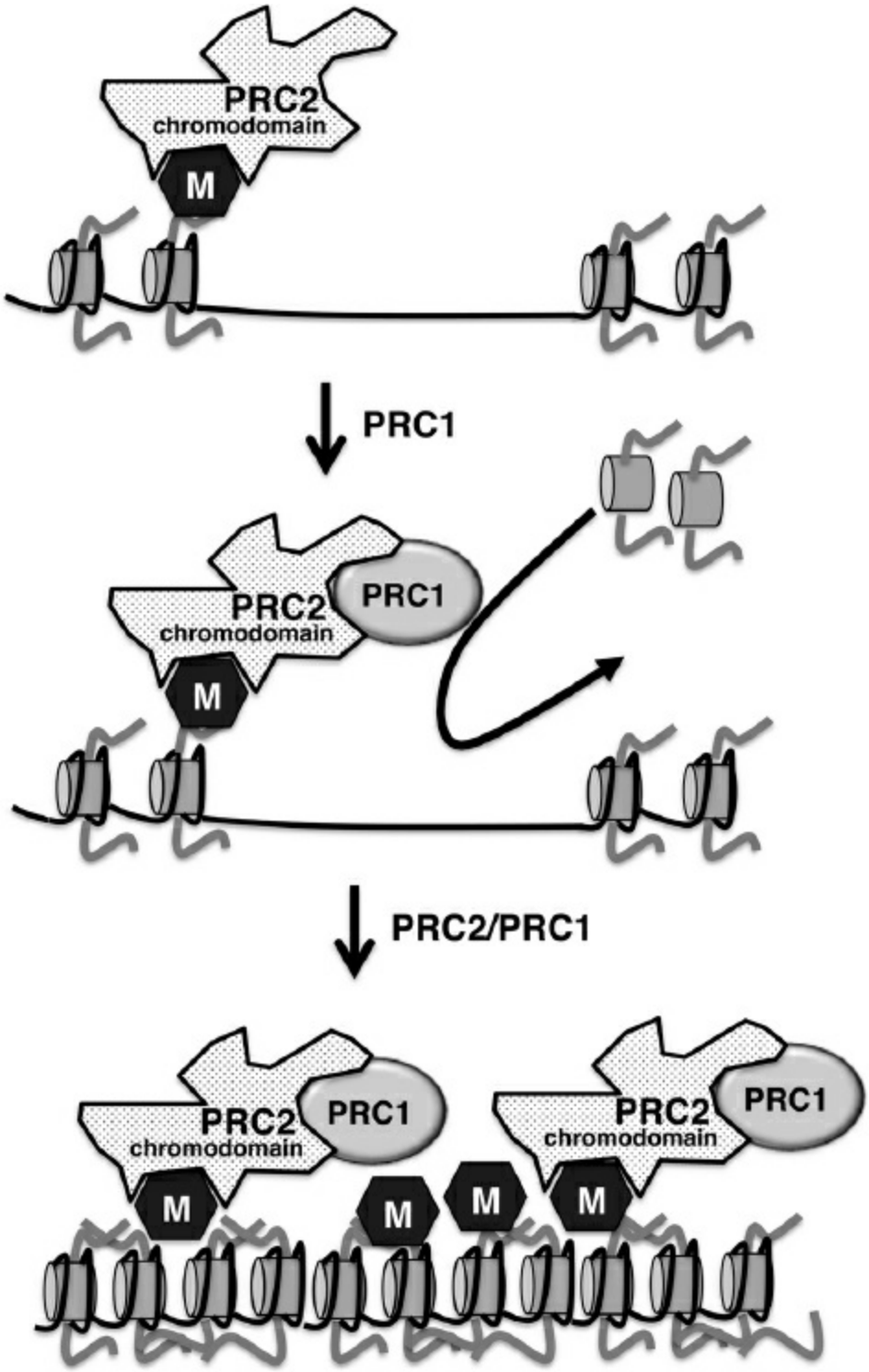

FIGURE 1.25–7. The PRC2 complex has a HMT activity that generates H3K27-me3, which further stabilizes PRC2 binding to chromatin (**top**). PRC2 recruits the PRC1 complex that further compact chromatin by increasing nucleosome density (**middle**). These newly recruited nucleosomes are methylated by PRC2 allowing for the propagation of this polycomb repressive chromatin (**bottom**).

Histone Phosphorylation. Phosphorylation of histones plays a role in chromosome segregation during mitosis, induction of early response genes, and DNA repair. Histone kinases such as Aurora/Ip1 and the mitogen and

stress activated protein kinases 1 and 2 (MSK1/MSK2) phosphorylate serines (S) 10 and 28 on histone 3. The proximity of these serine residues to lysines 9 and 27 influence the decision as to whether these lysines will be acetylated or methylated. This influence is critical because acetylation at lysines 9 and 27 tends to activate transcription whereas methylation at these same sites represses transcription (Fig. 1.25–4). Several environmental stimuli, such as stress, cocaine, changes in neurotransmission, and UV light induce H3S10 phosphorylation by activating MSK1 and MSK2. Activated MSK1 is recruited to the promoters of immediate early genes, such as c-fos and Jun, in a complex containing 14–3–3 proteins, chromatin-remodeling machinery, and a HAT. This complex efficiently activates expression of immediate early genes that provide appropriate cellular responses depending on the type of cells and the environmental context. In the dentate gyrus, this response facilitates helplessness behavior, while in the striatum this response modifies drug-seeking behavior (see Epigenetics in Preclinical Work).

ATP-Dependent Chromatin Remodeling. Nucleosomes are distributed in a stereotypical pattern along actively transcribed housekeeping genes. This pattern is highly conserved from yeasts to mammals and is characterized by a 250 bp nucleosome free region (NFR) that is flanked by two positioned nucleosomes in the promoter of housekeeping genes (Fig. 1.25–8). Similar NFRs are also found at the 3′ end of the gene and in enhancer regions. This pattern allows the transcriptional machinery to bind and constitutively express housekeeping genes. Nucleosome positioning is accomplished in vivo by two mechanisms. First, specific DNA sequences within the promoter region can guide the nucleosome's position, providing an important example of how the genetic code influences the epigenetic code. However, this arrangement is static and does not allow for rapid reorganization of chromatin in response to environmental demands. The second and perhaps more important mechanism by which nucleosomes are arranged along the DNA is mediated by chromatin-remodeling enzymes. These enzymes are equipped with chromatin-binding motifs and use ATP to remove, add, or slide nucleosomes along the DNA strand. In some cases, chromatin-remodeling enzymes maintain NFR by removing nucleosomes from specific promoter sequences. This in turn allows transcription factors to bind DNA and activate transcription (Fig. 1.25–8). Alternatively, nucleosomes can be pushed by chromatin-remodeling enzymes into NFRs where they block transcription and form higher-order chromatin structures (Fig. 1.25–8). Recruitment of chromatin-remodeling enzymes has been implicated in the development of addictive-like behaviors after chronic cocaine exposure in mice via modification of the promoters of late-onset addiction-related genes, such as cyclin-dependent kinase 5 (CDK5) (see Epigenetic Mechanisms of Cocaine Addiction).

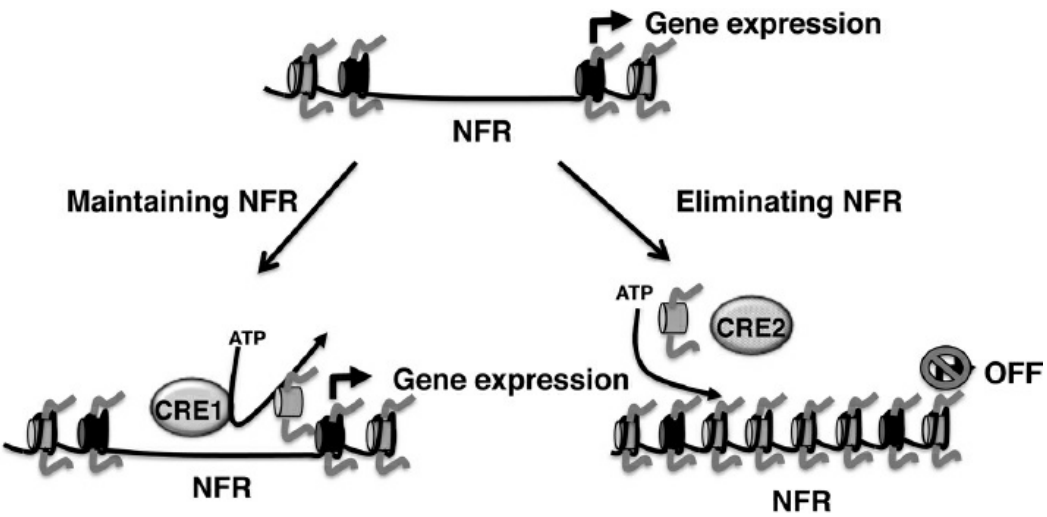

FIGURE 1.25–8. **Top:** Nucleosome free region (NFR) can be passively maintained by DNA sequences that position nucleosomes (*shown in black*) on both sides of the NFR. This in turn allows transcription factors to bind and activate gene expression. In many cases NFR is actively maintained by ATP-dependent removal of nucleosomes (depicted as CRE1on the lower left panel). ATP-remodeling activities (CRE2, on the right lower panel) use ATP to deposit nucleosomes making the NFR inaccessible and therefore turning transcription off.

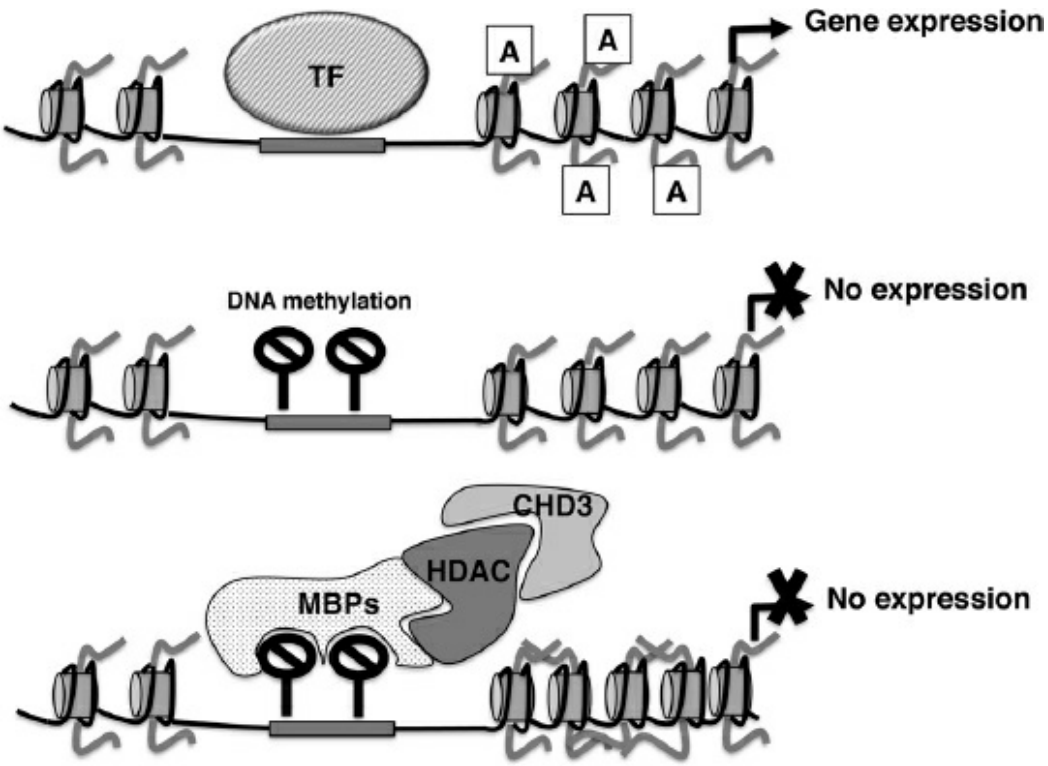

FIGURE 1.25–9. **Top:** In the absence of DNA methylation a generic transcription factor (TF) binds a specific DNA sequence within the promoter and activates gene expression. **Middle:** DNA methylation (black no entry signs) prevents the TF from binding its binding site and therefore preventing gene expression. **Bottom:** The NuRD complex is recruited to methylated DNA via methyl-binding proteins (MBPs). These MBPs are complexed with enzymatic activities such as HDACs and the chromatin-remodeling protein CHD3 to compact DNA and terminate gene expression.

DNA Methylation. Chromatin accessibility can also be modified by

DNA methylation. DNA methylation is accomplished by transferring a methyl group from a donor molecule to the 5' position on the cytosine ring to generate 5-methylcytidine (5mc). In vertebrates, DNMT enzymes methylate cytosines in a dinucleotide (CpG) palindromic sequence. Cytosine methylation does not affect its ability to form correct Watson–Crick pairing with a guanosine on the antiparallel DNA strand, and thus does not affect the underlying genetic code. Within the mammalian genome, CpG dinucleotides are either clustered in a high-density region known as CpG islands (CGIs) or found sporadically across the genome. CGIs usually mark promoters of active housekeeping genes and are thus generally unmethylated. In rare cases, when CGIs are methylated they form a highly repressive chromatin structure that inhibits gene expression and in many cases is inherited. Some well-known examples include X-inactivation and imprinted genes. The observation that abnormal regulation of imprinted genes leads to several forms of intellectual disability (Table 1.25–3) demonstrates the importance of correct methylation of these CGIs for normal neurodevelopment.

Methylation of CpG in sporadic regions usually downregulates gene expression. This downregulation is achieved in one of two ways. First, DNA methylation can prevent the binding of a transcription factor to its binding site (Fig. 1.25–9). A second mechanism involves the direct binding of a methyl-binding protein (MBP) to methylated cytosines. Docking of MBP to methylated DNA recruits enzymes that further compress chromatin and inhibit gene expression. A good example of this repressive cascade is the nucleosome-remodeling deacetylases (NuRD) repressive complex. NuRD is a highly conserved multisubunit complex that synergistically represses chromatin. The NuRD complex includes two HDACs (HDAC1 and HDAC2), a chromatin-remodeling subunit (CHD3 or CHD4), and several methyl-binding proteins (MBD2 and MBD3). The MBPs recruit the NuRD complex to methylated cytosines where CHD3/CHD4-mediated ATP hydrolysis compacts nucleosomes and stimulates deacetylation by HDACs. This synergism between ATP hydrolysis and histone deacetylation is thought to be due to the ability of chromatin-remodeling enzymes to make histone tails more accessible for deacetylation. The result is a highly repressive chromatin structure around methylated sites (Fig. 1.25–9). DNA methylation can therefore be thought of as a "molecular bridge" between the genetic code (i.e., CpG within specific binding sites) and the epigenetic machinery (e.g., DNMT, MBP, histone modifiers, and ATP-dependent chromatin-remodeling enzymes). The clinical section of this chapter provides an example of this concept, showing how the genetic code and changes in DNA methylation induced by early life stress interact to modify expression of the FKBP5 gene and subsequent vulnerability to PTSD.

Two de-novo DNMTs enzymes (DNMT3a and DNT3b) are able to methylate CpG sequences on naked or unmethylated DNA, whereas the

maintenance DNMT (DNMT1) methylates only hemimethylated sequence (Fig. 1.25–10). De-novo DNMTs are responsible for remethylating chromatin after two global demethylation events that occur during normal development. The first global demethylation occurs in the embryonic gametes in order to reprogram these cells to become a sperm or an oocyte. The second global demethylation occurs immediately after fertilization and is necessary to establish a pluripotent state in the zygote. Deletion of either DNMT3a or DNMT3b is embryonic lethal underscoring the importance of de-novo methylation for normal development. De-novo DNMTs are also expressed in postmitotic neurons where they methylate promoter elements that regulate the expression of genes implicated in synaptic plasticity, learning, and the response to cocaine.

In contrast to the two de-novo DNMTs (e.g., DNMT3a, DNMT3b), DNMT1 does not efficiently methylate naked DNA (i.e., unmethylated DNA). Instead, DNMT1 efficiently methylates DNA in which only one strand is methylated (Fig. 1.25–10). This hemimethylated DNA is commonly formed after DNA replication of a fully methylated CpG site. The presence of DNMT1 at the replication fork allows it to faithfully copy the DNA methylation pattern present on the maternal strand onto the newly formed daughter strand. This mechanism is responsible for the heritability of DNA methylation across mitosis. Deletion of DNMT1 in neural stem cells (NSCs) causes premature induction of gliogenesis (i.e., the production of glia cells) during embryogenesis and multiple defects in neuronal differentiation and synaptogenesis. In addition, exposure of rodents to early life stress increases expression of DNMT1 in NSCs, which is associated with increased methylation in promoter regions of the retinoic acid receptor. This promoter methylation results in decreased expression of the retinoic acid receptor, which leads to defects in the ability of NSCs to differentiate into neurons. In summary, appropriate DNMT1 activity is required for normal differentiation of NSCs.

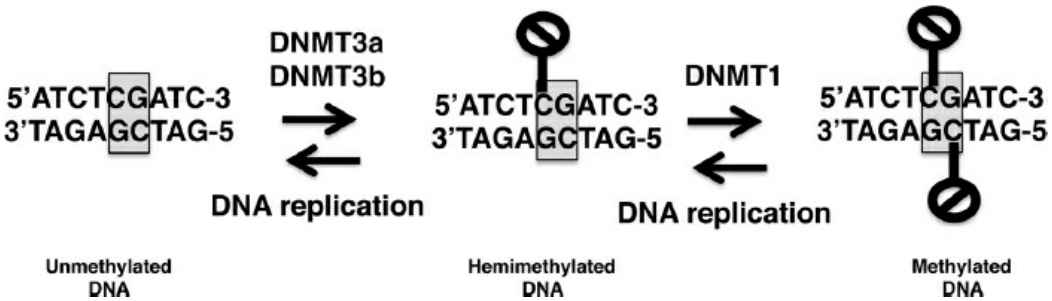

FIGURE 1.25–10. Unmethylated DNA is methylated by the de-novo DNMTs, DNMT3a, and DNMT3b, to form hemimethylated DNA. Hemimethylated DNA is methylated further by the maintenance DNMT, DNMT1. DNA methylation is lost after each round of DNA replication. The CpG dinucleotide is shown in *gray* and the DNA methylation is shown as *black "no entry" signs.*

The prevailing dogma for many years has been that DNA methylation is

a stable epigenetic marker that can be removed only by DNA replication coupled with inactivation of DNMT1 (Fig. 1.25–10), a process known as passive demethylation. Therefore, the presence of DNA methylation in postmitotic neurons was thought to be a stable marker that caused long-term suppression of gene expression. This assertion was based on the observations that the carbon–carbon covalent bond formed between the methyl group and the cytosine ring is highly stable, and no enzymatic activities have been conclusively shown to cleave this bond in vivo. However, a growing body of work has consistently shown that demethylation of some CpG sites can occur rapidly in postmitotic neurons. For example, from a total of 219,000 CpG sites examined in postmitotic neurons in the dentate gyrus of the hippocampus, roughly 3,000 sites showed a change in DNA methylation 4 hours after the administration of a single ECS. About two-thirds of these sites showed an increase in DNA methylation while the remaining third showed a decrease in DNA methylation, indicating rapid and active demethylation. DNA methylation and demethylation modified expression of many genes involved in synaptogenesis providing a possible mechanism to explain how ECS modifies neuroplasticity in the hippocampus. Interestingly, similar changes in DNA methylation were observed after 3 days of voluntary running, suggesting that ECS and exercise share similar epigenetic mechanisms for modifying synaptogenesis in the hippocampus.

The observation that DNA can be rapidly demethylated in neurons has led to an intense search for the responsible mechanisms. Different models have been proposed and additional work is needed to further clarify this issue, especially given that several pathways are likely to be involved. Nevertheless, compelling evidence has emerged to support the role of the ten-eleven translocation (TET) family of proteins, especially Tet1, in this process. This family of Fe(II)/2-oxoglutarate-dependent dioxygenases catalyzes the hydroxylation of methylated cytosine (5mC) to 5-hydroxymethyl-cytosine (5hmc). 5mhc is poorly recognized by DNMT1 providing a mechanism by which passive DNA demethylation can occur in mitotic cells (see also above). In neurons, 5mhc is recognized by deaminases, which catalyze the formation of 5-hydroxyuricil (5mhU). This alteration is recognized as a U:G mismatch that is repaired by the base excision repair (BER) machinery (Fig. 1.25–11). The BER complex consists of a DNA glycosylase that hydrolyzes the N-glycosidic bond creating an apyrimidinic/apurinic (AP) site that is removed by an AP endonuclease, and the resulting gap is filled by DNA polymerases. The growth arrest and DNA damage 45 protein family (Gadd45a, Gadd45b, Gadd45c) facilitate this process by recruiting deaminases or glycosylating enzymes to 5mhc sites (Fig. 1.25–11). Interestingly, the C to T transition in a CpG dinucleotide is the most common mutation in human disease suggesting that this convoluted DNA demethylation/repair mechanism is associated

with a significant cost for the organism. This observation further highlights the intimate relationship between the epigenetic and the genetic code.

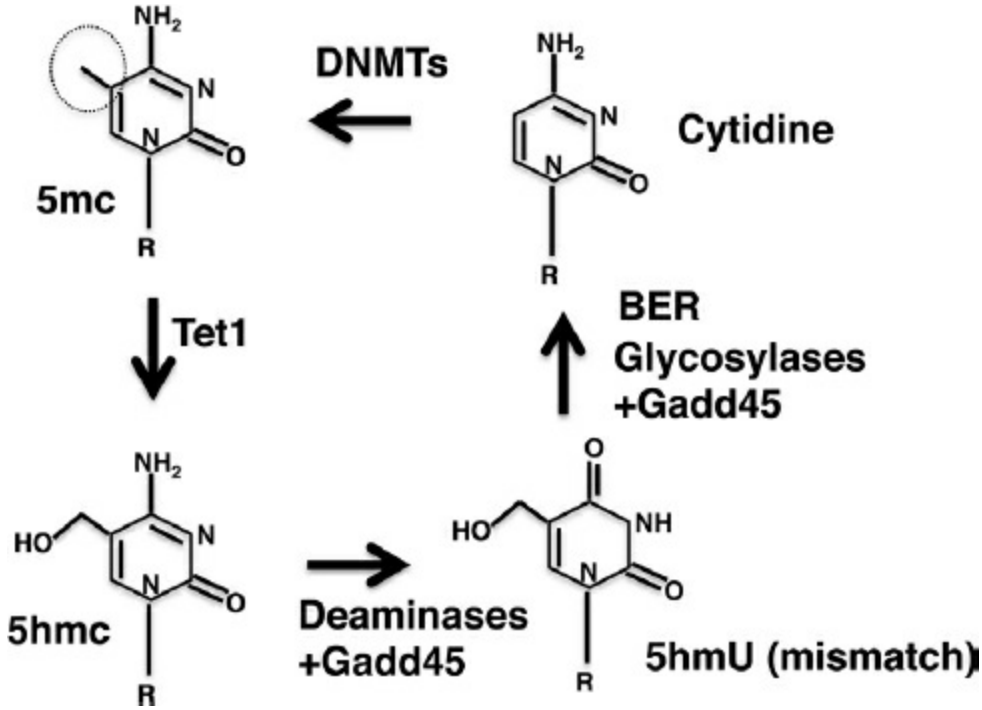

FIGURE 1.25–11. Pathways for DNA demethylation (methylation site in 5mc is encircled with a *dotted line*).

Chromatin Domains. Chromatin has been traditionally divided into euchromatin (active chromatin) and heterochromatin (repressed chromatin). However, recent advances in chromatin precipitation and high throughput sequencing allow for more precise chromatin mapping. This work has shown that in the mammalian cell chromatin is organized in topologically associating domains (TAD) that divide each chromosome into roughly 1Mb segments. TADs can be grouped into four major categories of chromatin: active chromatin (equivalent to euchromatin), and three distinct subtypes of heterochromatin known as polycomb-repressed chromatin, null chromatin, and constitutive heterochromatin (Table 1.25–5). The boundaries between different TADs are fairly well delineated by CGIs of actively transcribed housekeeping genes and binding sites for insulator proteins. However, some flexibility and shift within these boundaries can occur in response to environmental cues. The positions and boundaries of TADs are programed by the genetic code and are therefore similar in all cells in the body. However, the exact chromatin content within each TAD is different in different cell types (Fig. 1.25–12). These changes in TAD content are induced by histone modifications, chromatin-remodeling enzymes, and DNA methylation allowing cells to alter gene expression in response to different environmental conditions, and to use the same genetic code to form different cell types (Fig. 1.25–1). This

segmented organization further demonstrates how the genetic and the epigenetic codes interact with each other.

1. **Active chromatin.** Active chromatin is characterized by the presence of specific histone acetylations (H3K9-ac) and methylations (H3meK4) that are associated with open chromatin and active transcription (Table 1.25–4). Genes within an active TAD are usually coexpressed. Active TADs are mainly found in the interior of the nucleus and in some cases are aggregated together in specific sub-regions of the nucleus to form efficient "assembly lines" (Fig. 1.25–12). Perhaps the most well-characterized example is the nucleolus, where ribosomal RNAs (rRNAs) are synthesized and assembled with ribosomal proteins to form ribozymes. The rRNA transcriptional unit is arranged in a tandem series of about 100 such units clustered on different chromosomes in regions known as nucleolus organizer regions (NORs). Not all NORs are active, and the activation process is at least partly dependent on the ability of the UBF transcription factor to recruit NORs into the nucleolus. This recruitment process helps maintain the rRNA chromatin in an active state, creating a more efficient assembly of ribosomes. This ability of chromatin to loop enhancer elements, some of which are hundreds of kilobase pairs away from each other, in and out of contact with the each other and the transcriptional machinery, represents an important mechanism by which chromatin regulates gene expression. The effect of such looping on the risk for developing PTSD is discussed in the clinical section.
2. **Polycomb-repressed chromatin.** Polycomb TADs are characterized by polycomb-group response elements (PREs). These sequences are bound by specific transcription factors that recruit the PRC1 and PRC2 complexes to these TADs (Fig. 1.25–7). Polycomb TADs are therefore characterized by the presence of H3K27-me3, polycomb proteins, and repressed chromatin. Polycomb TADs regulate many important cellular processes such as X-inactivation, imprinting, NSC differentiation, and cortical development.
3. **Null chromatin.** In mammals, null chromatin is characterized by H3K9me2 histone methylation. This modification is required to tether TAD null chromatin to the nuclear envelope via laminin proteins. Null TADs can be significantly larger than other TADs extending up to 10 Mb in humans. They consist of highly repressive, late replicating chromatin that is characterized by high levels of H1 histones and laminin. HDAC1 is required for general silencing of null chromatin, whereas HDAC3 seems to play an important role in maintaining null chromatin tethered to the nuclear envelope. Mutations in laminin cause the movement of null chromatin from the nuclear envelope to the central part of the nucleus and are associated with chromatin decompaction, gene activation, and

severe developmental defects such as Hutchinson–Gilford progeria syndrome.

4. **Constitutive heterochromatin.** Constitutive heterochromatin is compacted during the entire cell cycle and is usually located around centromeres, telomeres, and regions that are enriched with tandem repeats. About 30 percent of the human chromatin is constitutive chromatin and is characterized by the presence of methylated histones, such as H4K20me3 and H3K9me2/me3, specialized histone variants, and high density of the HP1 protein (Fig. 1.25–6). Constitutive heterochromatin provides a more repressive chromatin than active chromatin, but is less repressive compared to polycomb or null chromatins.

Table 1.25–5.
Chromatin Can Be Categorized into Four Different Topologically Associated Domains (TADs)

TAD Chromatin	Histone Modifications and Associated Proteins	Nuclear/Chromosomal Position	Chromatin Structure and Transcription Activation
Active	H3K9-ac, H3meK4	Interior of the nucleus	Open chromatin, active transcription
Polycomb	H3K27-me3, PRC1, PRC2	Peripheral in differentiated cells	Compact chromatin, transcriptional repression
Null	H3K9me2, nuclear laminin, H1 histones	Nuclear envelope	Compact chromatin, transcriptional repression
Constitutive/Pericentric	H4K20me3, H3K9me2/me3, HP1, H1 histones	Chromocenters, centromeres, telomeres, tandem repeats	Compacted chromatin with intermediate transcriptional repression

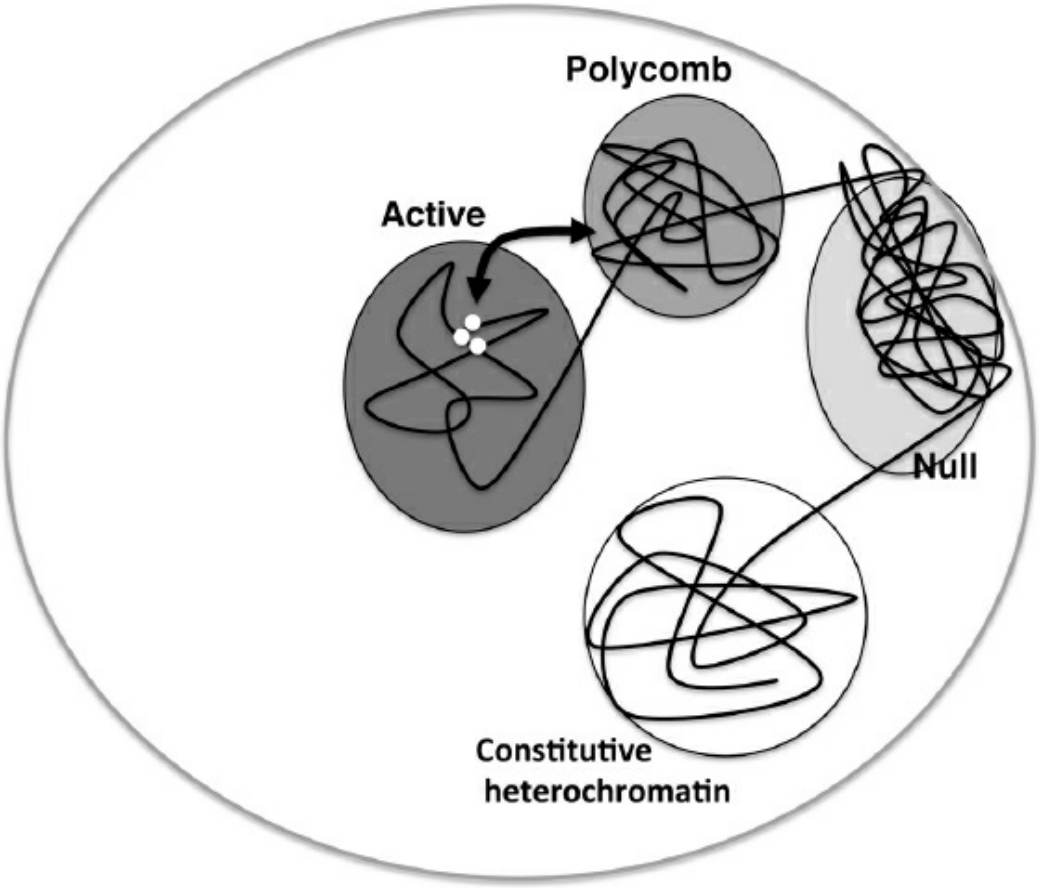

FIGURE 1.25–12. Chromatin in the nucleus is organized into TADs that are separated by housekeeping genes or insulator proteins (not shown). TADs can be categorized as active, polycomb, null, or constitutive/heterochromatin. The chromatin identity of each TAD may be different between different cell types or across the cell cycle within the same cell. Long distance loops of chromatin bring DNA regions that are far apart in the genetic code into close proximity in the epigenetic code, allowing the formation of assembly lines (*white dots*). These chromatin loops are dynamic and can change in response to environmental cues (*black arrow*).

EPIGENETICS IN PRECLINICAL WORK

Several lines of animal research elegantly demonstrate how epigenetic changes cause stable alteration of neuronal circuitry and psychiatrically relevant phenotypes such as stress reactivity, drug seeking behaviors, and helplessness behavior. Four examples are discussed to illustrate how epigenetic changes modify complex behaviors in animals in a manner that, in some cases, can be transmitted to the offspring. The potential therapeutic implications of these findings for psychiatric care are also discussed.

Epigenetic Transmission of Stress Reactivity and Maternal Behavior

The importance of early life experience for the development of the organism has long been appreciated from both psychodynamic and biologic perspectives. Cases of early life abuse and neglect often fit into a transgenerational pattern in which childhood maltreatment is subsequently

associated with maladaptive parenting styles later in life. Exposure to early life stress is linked with enhanced endocrine and autonomic responses to stress that persist into adulthood. This enhanced stress reactivity is closely related to increased risk for developing several psychopathologies including affective disorders, PTSD, and substance abuse (see Epigenetics Mechanisms in Psychiatric Illness). Animal models taking advantage of the normal heterogeneity in "parenting styles" in rats has allowed investigation of the biological mechanisms by which parental neglect increases stress reactivity in a manner that persists into adulthood and can be transmitted across generations.

Maternal behavior in rats is normally distributed in the population, as some mothers lick and groom (LG) their pups almost three times more frequently compared to others. Despite the fact that differences in the frequency of LG are seen only during the first postpartum week, they have profound effects on many behavioral phenotypes in the offspring (summarized in Table 1.25–6). For example, offspring raised by high LG mothers perform better in several memory tests, are less fearful, and show decreased corticosterone secretion in response to stress compared to "neglected" offspring raised by low LG mothers. Several observations suggest that these behavioral changes are due to the frequency of LG during the first week of life and not to genetic differences between the high and low LG mothers. First, cross-fostering studies show that offspring born to low LG mothers but raised by high LG dams are behaviorally indistinguishable from pups born and raised by high LG mothers. Similarly, offspring born to high LG dams but raised by low LG mothers resemble the biological offspring of low LG mothers. Second, mimicking maternal behavior by providing rat pups with different amounts of paintbrush strokes during the first weeks of life can induce similar behavioral changes. These distinct phenotypes persist into adulthood, indicating that exposure to different levels of tactile stimulation during the first week of life in some way programs adult behavior and physiology. Importantly, female offspring raised by high LG dams become high LG mothers themselves, while those born to low LG mothers become low LG mothers (Table 1.25–6). This pattern provides a behavioral mechanism by which the low and high LG phenotypes can be propagated transgenerationally.

Table 1.25–6.
Differences between Offspring Exposed to High and Low LG Mothers

Behavioral Phenotype in Adulthood	High LG Offspring	Low LG Offspring
Hypothalamic pituitary axis (HPA) reactivity (GR resistance)	Low	High
Anxiety	High	Low
Hippocampal dependent learning	Intact	Impaired

Maternal behavior (LG)	High	Low

These observational studies beg the question of how 7 days of early life experience determine stress reactivity and maternal behavior in adulthood. The emerging answer is that the frequency of LG during the first week of life causes stable changes in DNA methylation in promoter elements that regulate the expression of genes in neuronal circuits that control stress-reactivity and maternal behavior. For example, the promoter of the glucocorticoid receptor (GR) is heavily methylated at birth, leading to low GR levels in the hippocampus of newborn pups. Exposure to high LG during the first week of life induces several transient changes in the hippocampus that ultimately lead to extensive demethylation of the GR promoter. Demethylation of the GR promoter allows the transcription activator NGF inducible protein A (NGFI-A) to bind the GR promoter. Binding of NGFI-A recruits the HAT, CBP, to acetylate nearby histones, which further opens chromatin and enhances GR expression in the hippocampus of high LG offspring (Fig. 1.25–13). High levels of GR persist in the hippocampus of adult high LG offspring and exert an inhibitory influence on neurons in paraventricular nucleus (PVN) of the hypothalamus that secrete CRH in response to stress. The promoter of the GR in offspring of low LG mothers is kept hypermethylated. This hypermethylation prevents NGFI-A from binding to the GR promoter, leading to reduced GR levels in the hippocampus of low LG offspring (Fig. 1.25–13). In sum, exposure to high LG during the first week of life decreases stress reactivity in adulthood by demethylating the GR promoter during a critical period of development.

Exposure to high LG behavior also decreases DNA methylation at the estrogen receptor alpha (ERα) promoter in the medial preoptic area (MPOA) of the hypothalamus. Demethylation of this promoter recruits the transcription factor Stat5 to the ERα promoter allowing for increased ERα levels in the MPOA of high LG female offspring. Offspring raised by low LG mothers maintain low levels of ERα in the MPOA due to the persistence of DNA methylation at the ERα promoter. Differences in ERα levels in the MPOA persist into adulthood, regulate oxytocin expression, and are thought to program subsequent LG behavior in adult female offspring.

An important clinical question regarding the capacity of LG to cause stable alterations in DNA methylation, gene expression, and behavioral phenotype is whether these changes can be reversed pharmacologically in adulthood. The answer to this question seems to be yes. For example, intraventricular infusion of an HDAC inhibitor decreases DNA methylation at the GR promoter, increases GR levels in the hippocampus, and decreases stress reactivity in adult low LG offspring. While HDAC inhibitors do not change DNA methylation directly, they increase chromatin accessibility, a process that may facilitate demethylation of DNA. Conversely,

intraventricular infusion of methionine (a donor for methylation reactions) increases DNA methylation at the GR promoter, reduces hippocampal GR expression, and increases stress reactivity in adult high LG offspring. These findings suggest that the epigenetic landscape at the GR promoter can be modified pharmacologically in adulthood by tilting the balance between activities that promote open chromatin (histone acetylation, DNA demethylation) and activities that compact chromatin (methylating agents). It is currently unclear whether the developmental phenotypes induced by LG at the GR promoter will be reestablished after prolonged washout of the HDAC inhibitors or methionine. Addressing this question will have important implications regarding the true reversibility of these developmentally sensitive epigenetic changes.

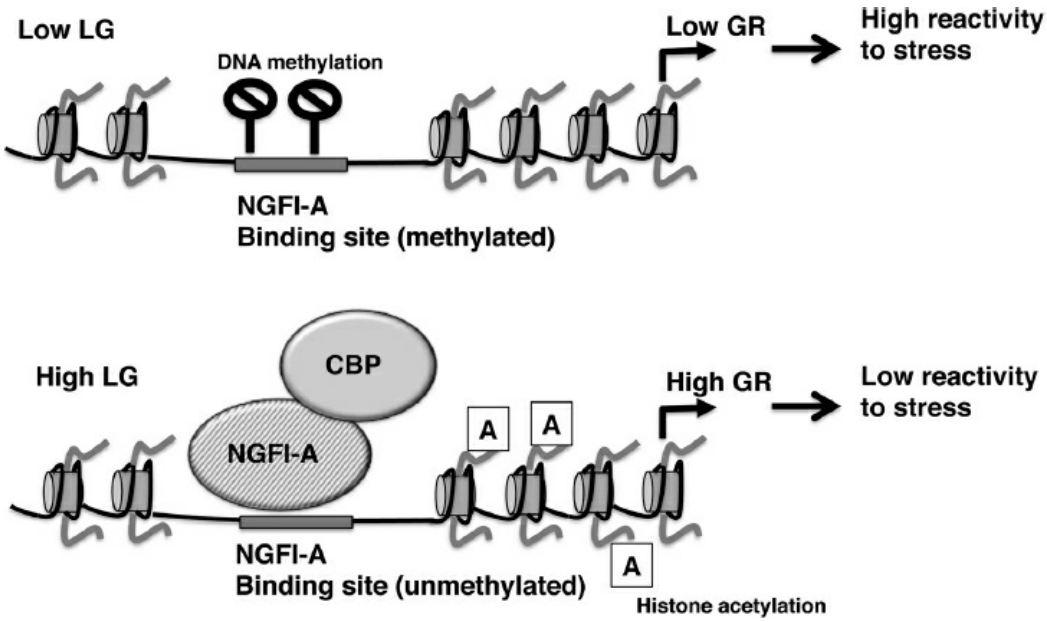

FIGURE 1.25–13. **Top:** The promoter of the GR is maintained in a methylated state in the offspring of low LG mothers, leading to low levels of GR expression in the hippocampus and high stress reactivity. **Bottom:** High LG during the first week of life leads to demethylation of the GR promoter, allowing the NGFI-A transcription factor to bind and activate GR expression by recruiting the HAT, CBP. Elevated levels of GR in the hippocampus decrease stress reactivity in the offspring of high LG mothers.

The relevance of these findings to human psychopathology has been examined in postmortem studies by assessing the brains of suicide victims with documented histories of childhood abuse or neglect compared to suicide victims with no childhood trauma. Similar to the preclinical work in rats, childhood maltreatment was associated with increased DNA methylation at the GR promoter and reduced GR levels in the hippocampus. While the challenges of assessing epigenetic signatures in living patients remain (see Epigenetics Mechanisms of Psychiatric Illness), these studies suggest that preclinical work in rodents can provide important insights into the mechanisms by which epigenetics contribute to

psychiatric conditions in humans.

Epigenetic Inheritance of Olfactory Fear Memory

A critical survival function in both animals and humans is the ability to recognize, remember, and avoid dangerous environmental signals. While humans rely strongly on visual cues, rodents scout their environment with their highly sensitive sense of smell. In mice, the repeated pairing of a specific odor, referred to here as the "*offending odor,*" with a noxious stimulus such as a foot shock, results in a fear memory that is manifested as an exaggerated startle when the rodent encounters the offending odor alone (no shock). Olfactory-mediated memory is therefore considered an adaptive from of sensitization to environmental threats that helps the mouse predict and avoid potential dangers. The ability to transmit this sensitization strategy to one's offspring would theoretically provide a considerable advantage to young animals learning to navigate a dangerous world. Recent animal data suggest that parental experience with olfactory-fear learning can be transmitted to subsequent generations through epigenetic mechanisms.

To test the inheritance of sensitivity to danger-associated olfactory stimuli, mice were fear-conditioned to a specific offending odor. This procedure was done in mice in which the neurons expressing the olfactory receptor coding for that specific offending odor were labeled with a LacZ reporter. The advantage of this experimental system is that it allows the researchers to assess whether the sensitivity to the offending odor is accompanied by an increase in the expression of its receptor (i.e., detection system) in the olfactory epithelium. Male mice exposed to olfactory fear conditioning were then mated to naïve female mice that were not exposed to the offending odor or to shock. The offspring from this mating (i.e., the F1 generation) showed an exaggerated startle response when exposed to the offending smell despite the fact that these mice had no prior exposure to the offending odor. Importantly, the F1 offspring showed no startle response when exposed to another odor, indicating that the fear response was specific for the offending odor. Similar findings were seen in the F2 generation, when F1 male mice were mated with naïve females, now two generations removed from the initial fear learning. Neuroanatomical examination of F1 and F2 mice showed that the expression of the olfactory receptor responsive to the conditioned odor was selectively upregulated in the olfactory epithelium. These findings suggest that exposure to a specific olfactory threat in the parental generation (F0) somehow amplifies the detection apparatus for this specific threat in the olfactory epithelium in a manner that can be transmitted to subsequent generations, which may better prepare them to avoid specific environmental threats.

To test whether this form of transmission was mediated by epigenetic modifications in the parental sperm, sperm were collected from fear-

conditioned males and were used for in vitro fertilization (IVF) experiments with fear-naïve female mice. The offspring of these IVF experiments also showed transgenerational fear memory and the upregulation of olfactory receptors. The DNA isolated from sperm of the F0 and F1 male mice showed hypomethylation in the vicinity of the olfactory receptor responsible for the specific odor detection. No changes in DNA methylation were seen around other olfactory receptors, indicating precise epigenetic marking of the appropriate detection system. Curiously, however, this hypomethylation was seen only in the sperm DNA, and was not present in DNA extracted from the olfactory epithelium, raising questions about the mechanisms by which hypomethylation in the sperm affects receptor expression in the olfactory epithelium. Moreover, additional work is needed to elucidate the "imprinting mechanism" by which hypomethylation in the F0 sperm escapes two global demethylation events in the F1 mice (see The Epigenetic Machinery, DNA Methylation). While more studies are needed to clarify the mechanisms by which olfactory memory is propagated across generations, these studies demonstrate how exposure to environmental stress in the parental generation can be transmitted via stable epigenetic alterations in the germline. This form of epigenetic heritability, that of germline transfer, differs from the one described above for maternal care where the transgenerational transmission is mediated by tactile stimulation provided by the mother during the first week of life.

The Epigenetic Mechanisms of Depressive Behavior and Antidepressant Action

Rodent models of depression and PTSD typically employ some form of chronic, unpredictable, or severe stress that leads to despair, avoidance, or anhedonic-like behaviors. The observations that these behavioral consequences could be reversed with chronic treatment with antidepressants suggest that they provide reasonable models of human depression and PTSD. One such paradigm involves chronic social defeat in which a mouse is briefly exposed to an aggressive intruder over the course of 10 days. Defeated mice show a decrease in sucrose consumption, suggesting an anhedonic-like state, and increased helplessness in the forced-swim and the tail-suspension tests. Like human depression, many of these behavioral changes are reversed by chronic but not acute antidepressant treatment.

Several observations suggest that the behavioral consequences of social defeat and the rescue of depressive phenotypes by antidepressant medications are due to epigenetic changes in the expression of genes implicated in neuroplasticity. This property is not unique to stress or antidepressants but seems to be a common pathway seen with many other psychiatrically relevant interventions (e.g., ECS, exercise, learning) and risk factors (e.g., drug abuse, inflammation). One of the best examples of

epigenetic regulation of neuroplasticity genes is the opposing effects of stress and antidepressants on the expression of BDNF in limbic brain areas. BDNF expression in the hippocampus is downregulated in defeated mice in a manner that persists for many weeks after terminating the stress, an effect that can be reversed with chronic treatment with the TCA, imipramine. The ability of social defeat to decrease hippocampal BDNF levels appears to be mediated by increasing H3K27 methylation at the BDNF promoter. These histone modifications lead to a repressive chromatin state that inhibits BDNF expression (Fig. 1.25–14). These findings are consistent with human postmortem studies showing that depressed individuals have increased H3K27 dimethylation (H3K27-me2) at the BDNF promoter and reduced BDNF levels. Treatment with imipramine rescued the depressive phenotype and increased BDNF levels in the hippocampus of defeated mice, but did not alter H3K27-me2 levels. Instead, imipramine treatment increased acetylation at H3K9 residues at the BDNF promoter, a modification that is associated with open chromatin and transcriptional activation (Fig. 1.25–14). These results underscore the combinatorial nature of the histone code, in which acetylation can override the repressive nature of H3K27-me2. Treatment with imipramine decreased HDAC5 expression in defeated, but not control mice, raising the possibility that reduced HDAC5 in defeated mice might promote H3K9 acetylation and therefore mediate the therapeutic effects of imipramine (Fig. 1.25–14). This assertion is supported by data showing that overexpression of HDAC5 in the hippocampus blocks imipramine's antidepressant effects in defeated animals.

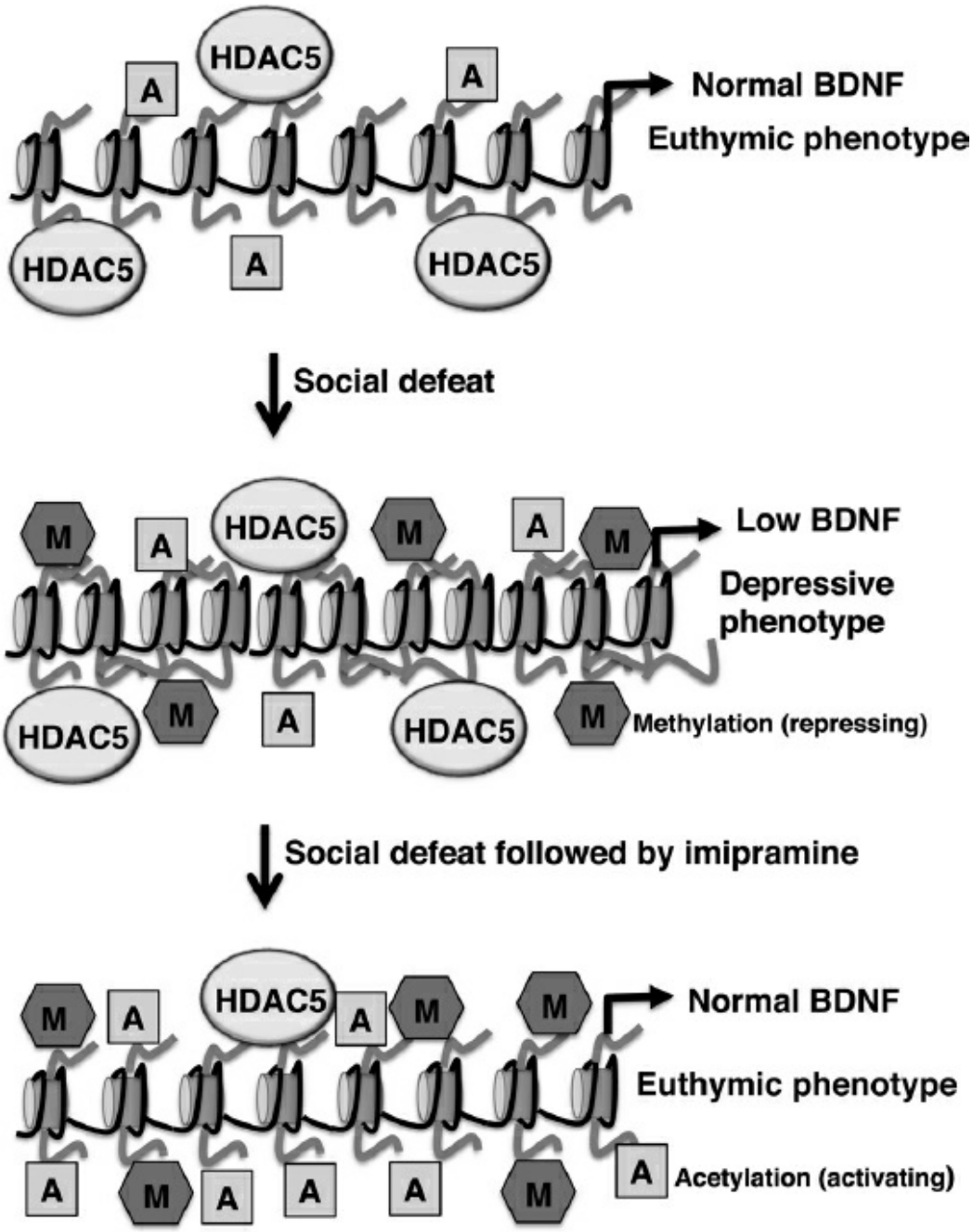

FIGURE 1.25–14. **Top:** Control mice show normal BDNF levels and euthymic-like phenotype. **Middle:** Exposure to repeated social defeat leads to an increase in repressive HK27-me2 methylation (*red hexagons*), leading to condensed chromatin, decreased BDNF levels, and a depressive-like phenotype. **Bottom:** Treatment of socially defeated mice with imipramine does not alter HK27-me2 levels, but instead decreases HDAC5 levels (*yellow circles*). Reduced HDAC5 levels contribute to the increase in histone acetylation (*green squares*), chromatin decondensation, increased BDNF levels, and restoration of euthymic phenotype. (See color atlas.)

Epigenetic Mechanisms of Cocaine Addiction

While the prior sections have addressed how negative life experiences (fear conditioning, chronic social defeat, or maternal neglect) cause long-lasting changes in chromatin structure, intensely positive experiences can also be

coded epigenetically. A classic example is drug addiction, in which the natural reward system is hijacked by highly rewarding exogenous substances, resulting in maladaptive compulsive drug-seeking behavior.

The transition from the initial experimentation phase of drug consumption, driven by the rewarding experience, to the addictive phase, characterized by compulsive maladaptive behavior, is a gradual process that is seen across diverse species from humans to rodents. In rodents, the transition to addictive behavior is mediated by epigenetic changes in the expression of genes that modify synaptic plasticity in neurons located within the reward system. This section provides a brief summary of some of the main lessons learned about this process by focusing on the role that histone acetylation in the nucleus accumbens (NAc) plays in the development of drug-seeking behaviors.

The immediate rewarding effects of cocaine are closely associated with an increase in dopamine levels in several brain circuits, with the NAc being one of the most important and well-studied targets. The initial rewarding effects of cocaine are experienced within minutes and lead to the transient expression of immediate early genes such as c-fos. Chronic, but not acute, administration of cocaine is necessary to increase expression of an additional group of late onset genes such as FosB/ΔFosB, CDK5, and BDNF. Expression of many of these late onset genes can persist for many weeks even in the absence of cocaine and is necessary to elicit robust drug-seeking behaviors in rodents. This effect has been extensively studied for the highly stable transcription factor ΔFosB. Repeated activation of c-fos leads to the accumulation of ΔFosB in a subset of NAc neurons. High levels of ΔFosB are necessary to express synaptic proteins that enhance synaptic plasticity in NAc neurons, which results in the stable changes that are necessary for drug seeking behavior. Accumulation of ΔFosB is also seen in other compulsively consumptive behaviors such as exercise and feeding, consistent with the notion that cocaine is able to hijack endogenous reward pathways.

Histone acetylation plays a critical role in coordinating the transition from inducing immediate early genes to the stable expression of late onset genes such as ΔFosB. Acute cocaine administration induces c-fos expression by increasing H4 acetylation and H3 phosphorylation. With repeated cocaine use, these chromatin changes are no longer present at the c-fos promoter leading to a decline in c-fos levels. Instead, repeated cocaine administration leads to an increase in H3 acetylation at promoters of late onset genes (e.g., FosB, ΔFosB, Cdk5 and BDNF). Manipulations that inhibit histone acetylation such as viral knockdown of the HAT CBP, or the over expression of different HDACs in the NAc block the ability of chronic cocaine to induce late onset genes and reduce compulsive drug-seeking behaviors. These findings suggest a model in which the transition to drug addiction is mediated by an epigenetic switch from an initial rapid H4

acetylation to a slower wave of H3 acetylation that is necessary to induce late onset genes.

It is important to note however that this is a highly simplified model and that the epigenetic effects of chronic cocaine are far more complex. For example, prolonged inhibition of HDAC1 in the NAc reduces cocaine locomotor sensitization, which is inconsistent with the "expected" increase in histone acetylation and the promotion of drug-seeking behavior. This unexpected result is likely due to the ability of chronic HDAC1 inhibition to maintain the repressive H3K9-me2 marker at the promoters of GABA receptor genes even after chronic cocaine exposure. Thus, the retention of the H3K9-me2 repressive marker prevents expression of these genes even in the presence of high levels of acetylation. This example underscores the complexity by which different histone modifications interact with one another to modify chromatin structure and gene expression (Fig. 1.25–14).

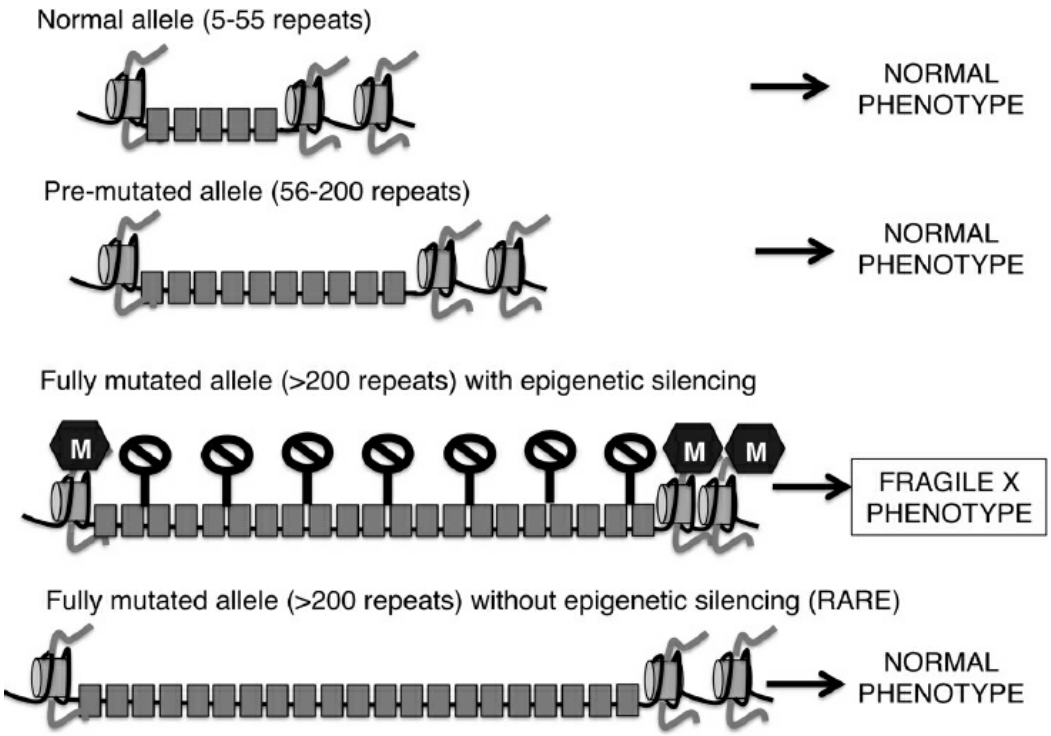

FIGURE 1.25–15. In normal individuals, the FRM1 gene contains 5 to 55 CGG repeats (each gray box represents 10 CGG repeats). Individuals with the premutation have 56 to 200 repeats and are phenotypically normal. However, they are at increased risk for repeat expansion during meiosis resulting in the fully mutated phenotype in their offspring. The FRM1 promoter of fully mutated individuals (>200 CGG repeats) becomes highly methylated (*black no entry signs*) leading to abnormally reduced FMR1 expression and the Fragile-X phenotype. Rarely, individuals have been identified who possess the full mutation, yet lack epigenetic silencing. These children are phenotypically normal, suggesting that DNA methylation is necessary to induce the clinical syndrome.

Chronic cocaine also downregulates expression of DNMT3a in the NAc, and manipulations that decrease DNMT3a enhance addictive phenotypes. These findings suggest that some of the behavioral consequences of cocaine

addiction are mediated by demethylation of DNA in the NAc. Finally, some of the epigenetic changes seen after cocaine administration appear to be cell-type specific. For example, knockout of the HMT (G9a) in D1 neurons decreases the behavioral consequences of cocaine, but the opposite phenotype is seen when the same enzyme is knocked out in D2 neurons. This cell-type specificity is not surprising given that different cell types possess very different chromatin structures and functions (see Introduction and Fig. 1.25–1). Nevertheless, this cell-type specificity represents a challenge for pharmacological interventions that target specific chromatin activities (e.g., the HMT G9a) and for interpreting chromatin changes seen in heterogeneous brain regions such as the NAc.

EPIGENETIC MECHANISMS IN PSYCHIATRIC ILLNESS

While epigenetics in psychiatric research is still in its infancy, studies of chromatin modifications in clinical populations hold significant promise for understanding several fundamental unanswered questions in psychiatric research (see Table 1.25–2 in the introduction section). Furthermore, this kind of research will be critical for confirming the translational potential of preclinical work (see Epigenetics in Preclinical Work) and for identifying novel drug targets and biomarkers. This last section provides three clinical examples that illustrate the promise that epigenetic research holds for revolutionizing psychiatric care.

The Epigenetic Etiology of Fragile X Syndrome

Fragile X syndrome is the leading cause of inherited intellectual disability in the world, with an estimated prevalence of 1:2,500 males. It is an X-linked monogenetic neurodevelopmental psychiatric disorder characterized by moderate intellectual disability, social deficits, a high incidence of ADHD, and multiple physical abnormalities (Table 1.25–3). Fragile X syndrome is caused by the loss of function of the fragile X mental retardation gene 1 (FRM1). The FMR1 gene codes for an RNA-binding protein that plays an important role in dendritic spine formation, explaining its impact on cognitive function and behavior. While the term "monogenetic" implies mutated genetic code, the actual disease-causing mechanism involves epigenetic changes that silence the promoter of this gene.

The number of CGG repeats in Fragile-X syndrome can be categorized into three groups: normal, premutated, and fully mutated alleles. The normal allele consists of 5 to 55 CGG repeats allowing for normal expression of the gene. The premutated allele has 56 to 200 repeats and maintains normal expression of the FRM1 protein. However, the premutated allele is associated with increased risk for further expansion, pushing it into the fully mutated range, defined as greater than 200

repeats. Since each CGG repeat contains a CpG dinucleotide, this expanded repeat is recognized as an unstable genomic region that becomes heavily methylated. As in the case of other heavily methylated CGIs (see The Epigenetic Machinery, DNA Methylation), this methylated region recruits additional remodeling enzymes that form a highly repressive chromatin structure that prevents expression of the FRM1 gene (Fig. 1.25–15).

In a subset of rare cases, individuals with fully mutated alleles (i.e., >200 repeats) do not develop clinical symptoms of Fragile X syndrome. In these rare cases, the expanded repeats somehow escape DNA methylation providing compelling evidence that DNA methylation is necessary to induce the clinical syndrome in individuals with the full mutation. Therefore, therapies that prevent DNA methylation could potentially rescue the neurodevelopmental sequelae of individuals with the fully mutated allele. Indeed, preliminary experiments using a lymphoblastoid cell line have been able to demethylate the fully mutated promoter and induce protein expression using a combination of DNMT and HDAC inhibitors. The compounds used in these cell culture experiments lack specificity and carry a certain level of toxicity, limiting their translation into clinical practice thus far. However, studies are currently underway to develop novel, nontoxic, epigenetic interventions that could potentially cure this syndrome in the future.

Gene by Environment Interaction Modulates FKBP5 Expression and the Risk for PTSD

In many cases the effects of stress on complex behaviors are mediated by genetic vulnerability. Several examples of this type of gene by environment interaction have been described in the psychiatric literature; however, the molecular details by which these interactions take place have only just started to emerge. This was recently illustrated in a series of studies showing how childhood trauma preferentially interacts with a specific genetic predisposition to increase expression of the FK506-binding protein 5 (FKBP5) and the vulnerability to PTSD and depression.

FKBP5 is a stress-related gene that exerts its effects via interaction with the GR. Under basal conditions, the GR is maintained in the cytoplasm by the heat shock 90 (Hsp90) and FKBP5 proteins, which bind GR and mask its nuclear localization signal. High levels of cortisol induced by stress dissociate Hsp90 and FKBP5 from GR, allowing it to enter the nucleus and induce two transcriptional programs. The first transcriptional program is essential for terminating further release of cortisol from the adrenal gland (see Epigenetic Transmission of Stress Reactivity and Maternal Behavior). The second task involves the induction of FKBP5, a negative feedback loop that prevents further entry of GR into the nucleus. Retaining GR in the cytoplasm is thought to reset the stress response so it can be reactivated in response to additional threats (Fig. 1.25–16).

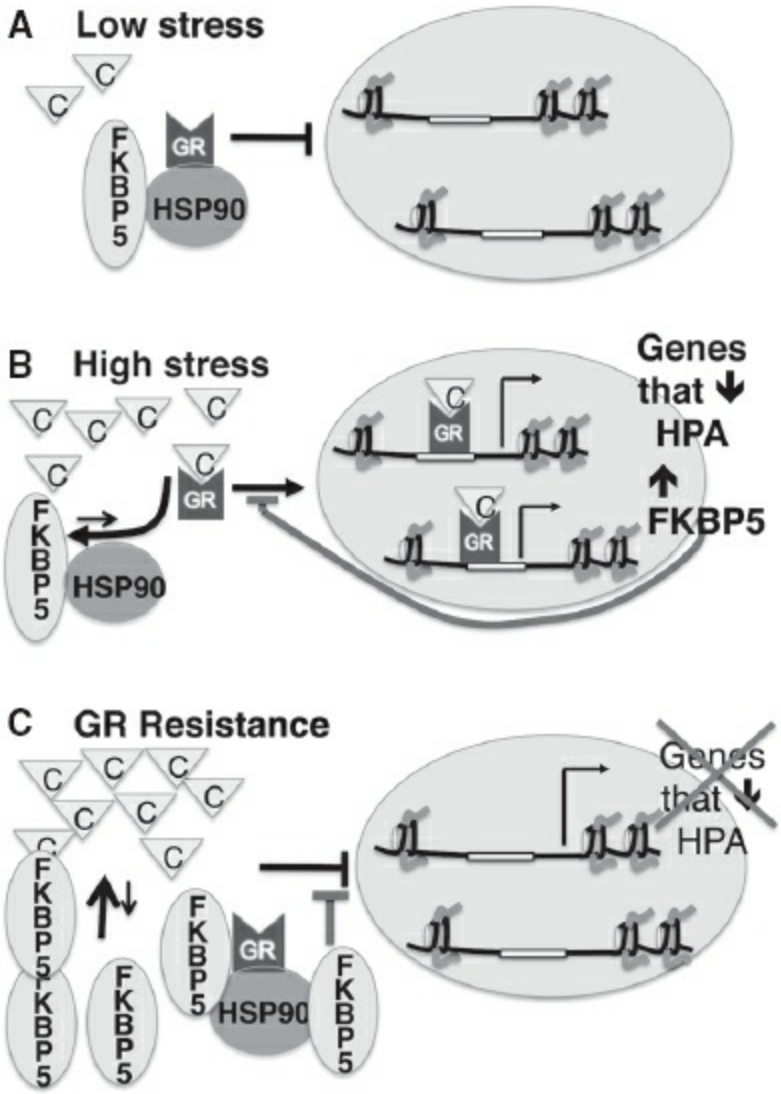

FIGURE 1.25–16. A: In the absence of stress, when cortisol levels (*green "C" triangles*) are low, the GR (*blue*) is bound to the chaperone proteins FKBP5 (*yellow circle*), and HSP90 (*brown circle*). In this bound state, the GR is prevented from entering the nucleus and activating transcription. **B:** Exposure to stress increases cortisol levels allowing GR to enter the nucleus by displacing the FKBP5/HSP90 complex. Inside the nucleus, GR activates transcription of genes that terminate further cortisol release and activates transcription of FKBP5. Increased FKBP5 levels prevent further GR entry into the nucleus, resetting the system so it can respond to subsequent waves of stress. **C:** Overexpression of FKBP5 leads to GR resistance by preventing high levels of cortisol from driving GR into the nucleus. As a result, additional levels of cortisol are being released leading to high levels of circulating cortisol. (See color atlas.)

Conditions in which FKBP5 is overexpressed prevent normal activation of the GR, leading to inefficient termination of the stress response (see also Epigenetic Transmission of Stress Reactivity and Maternal Behavior). This inability to terminate HPA activity leads to prolonged exposure to high levels of cortisol and is referred to as "*GR resistance*." GR resistance is associated with increased risk for depression, PTSD, and several other medical complications (Fig. 1.25–16). This disease model is similar to that described for insulin resistance, in which ineffective insulin signaling

causes end-organ toxicity by maintaining high levels of glucose in the circulation.

The rs1360780 haplotype of the FKBP5 gene has an A/T and a C/G allele. At baseline, the A/T allele is associated with higher levels of GR-induced FKBP5 expression, leading to more robust HPA activation compared to the protective C/G allele (Fig. 1.25–17). While the A/T allele alone is not associated with increased risk for stress-related psychiatric disorders, the risk allele interacts with early life trauma in a dose-dependent manner to increase the risk for PTSD. In other words, individuals with the A/T allele that were exposed to more severe childhood maltreatment are more likely to develop PTSD and depression. This is not the case for individuals with the protective C/G allele in which exposure to childhood trauma is not associated with increased risk for PTSD or depression.

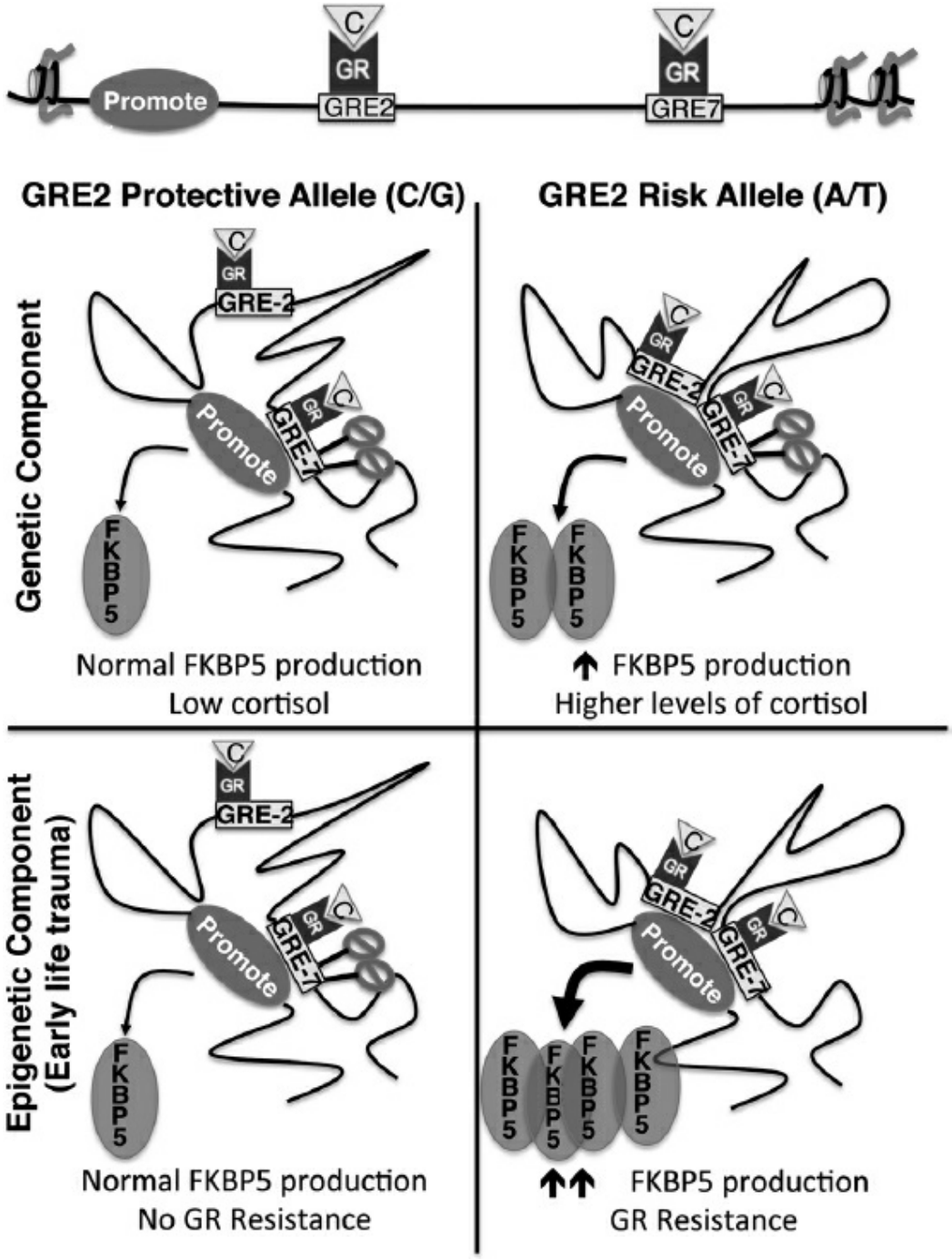

FIGURE 1.25–17. The FKBP5 gene has two GREs located in introns 2 (GRE-2) and 7 (GRE-7). The **top** two panels demonstrate the impact of the genetic risk on FKPB5 expression and GR resistance. On the **left,** GRE-2 containing the protective C/G allele is not recruited to the promoter/GRE-7 complex producing relatively low levels of FKBP5, and low cortisol levels in response to mild stress. The presence of the GRE-2 at the promoter/GRE-7 complex in the A/T allele (**top right**) produces higher levels of FKBP5, and hence higher levels of cortisol in response to mild stress. The bottom panels demonstrate the epigenetic contribution of early life stress to FKPB5 expression. Early life stress results in relatively low levels of cortisol and normal methylation at the GRE-7 in individuals with the C/G protective allele (**lower left panel**). In contrast, early life stress leads to higher cortisol levels and a decrease in the DNA methylation at the GRE-7 of individuals with the A/T risk allele. Demethylation at GRE-7 leads to more efficient GR binding to this site and therefore higher FKBP5 levels and GR resistance (**lower right panel**). (See color atlas.)

The mechanistic explanation for these clinical observations points to a synergy between genetic and epigenetic regulation of FKBP5 levels and GR resistance. At the genetic level, A/T susceptible individuals show a more efficient GR-mediated FKBP5, producing higher levels of FKBP5 than C/G individuals. The FKBP5 gene has two glucocorticoid response elements (GREs) that are occupied by GR in response to stress. One GRE is found in intron 2 (GRE-2), is roughly 50 kb from the TSS, while a second GRE is found in intron 7 (GRE-7), about 100 kb from the TSS of FKBP5. Despite these long distances, both GREs are able to contact the TSS via long-distance loops. These chromatin-mediated loops recruit these GREs to the vicinity of the TSS, driving FKBP5 expression. When GR is activated, the GRE-7 site is recruited to the TSS regardless of the genotype. In contrast, GRE-2 is recruited to the TSS only in AT individuals. The presence of GRE-2 at the vicinity of the TSS, results in a more robust expression of FKBP5 in response to stress, and hence greater HPA activation in individuals with the A/T allele (Fig. 1.25–17). It is currently unclear how these genetic changes contribute to these differential looping configurations, but the presence of these loops is an important function of chromatin (see Active Chromatin).

As a result of these genetic differences, A/T individuals exposed to childhood abuse are predicted to secrete higher levels of cortisol compared to C/G individuals. This increased cortisol response from early life trauma is thought to decrease DNA methylation at the GRE-7 binding site in individuals with the risk allele. Decreased DNA methylation at GRE-7 leads to more efficient recruitment of GR, further augmenting FKBP5 expression and GR resistance in A/T individuals that were exposed to childhood trauma (Fig. 1.25–17). GRE-7 continues to be methylated in G/C individuals that were abused or neglected, demonstrating specific gene by environment interactions for this epigenetic marker. Ex vivo experiments with patient-derived monocytes show a strong negative correlation between GRE-7 methylation and glucocorticoid potency, supporting a functional role for this epigenetic biomarker in mediating GR resistance.

The ability of trauma to induce these epigenetic changes appears to be confined to a sensitive period, as exposure to trauma in adulthood does not lead to a decrease in the GRE-7 methylation rate. These clinical findings are supported by cell culture experiments in which dexamethasone decreases GRE-7 methylation in proliferating NSCs, but not in fully differentiated postmitotic neurons. When exposed during proliferation, this decrease in methylation is maintained after the NSCs differentiate into neurons, even once the dexamethasone stimulus is removed, suggesting stable propagation of demethylated GRE-7 into differentiated neurons. These observations suggest that exposure to stress during childhood, when high levels of neurogenesis are present, is responsible for this gene by early life stress interaction. Therefore, interventions that are aimed at decreasing stress early in life are likely be more effective in reducing vulnerability to

PTSD compared to interventions that eliminate stress in adulthood.

Epigenetic Suicide Signatures: Advance Toward a Biomarker

Suicide is the eleventh most common cause of death in the United States. The importance of accurate risk assessment is highlighted by the observation that roughly half of the individuals who commit suicide are seen by a mental health worker prior to their death. In the absence of reliable biomarkers to predict suicide, psychiatric evaluation relies on a long list of risk factors to assess the likelihood of lethal self-harm. These population risk factors have poor predictive value, as most people who have these risks do not commit suicide, and some individuals who ultimately take their own life may not share these risks. Therefore, the recent findings that certain peripheral epigenetic markers can predict suicide with more than 80 percent accuracy provide a possible breakthrough for preventing these tragedies.

To search for epigenetic markers of suicide, DNA methylation patterns from the prefrontal cortex of depressed individuals who committed suicide were compared to depressed individuals who died by other causes. DNA methylation is amenable to study in postmortem tissue because of its relative stability and its ability to predict gene expression (see DNA Methylation). Only one gene, coding for the spindle and kinetochore associated complex subunit 2 (SKA2), was found to be differentially methylated in neurons of suicide completers compared to the control group. SKA2 is a scaffolding protein that chaperones the GR into the nucleus, and therefore helps terminate HPA activation in response to stress. Increased DNA methylation at the 3′ UTR of SKA2 was replicated in a separate postmortem cohort, suggesting that it represents a reliable retrospective biomarker of suicide. Methylation of SKA2 was associated with a decrease in SKA2 protein levels, suggesting that this epigenetic modification regulates SKA2 activity in suicide victims. Interestingly, the CpG methylation site at the 3′ UTR of SKA2 contains an alternative SNP allele in which the cytidine is replaced with thymidine. This T allele cannot be methylated and is therefore considered to be the protective allele.

While the postmortem replication is an important first step, the more important question is whether the presence of this epigenetic biomarker in peripheral tissue can predict future suicidal behavior in living patients. This was assessed by first comparing gene expression in prefrontal cortex neurons and glial cells obtained from the postmortem analysis, and peripheral immune cells obtained from a different cohort of individuals with depression. This analysis showed high correlation between gene expression in neurons, but not glia, and peripheral immune cells, suggesting that the two cell populations share similar epigenetic signatures, which is a critical feature in establishing a peripheral biomarker. Next, the relationship between SKA2 methylation state in peripheral immune cells,

genetic makeup (T vs. C alleles), and suicidal ideation were assessed prospectively in three separate cohorts. The remarkable findings were a significant interaction between genetic disposition (T vs. C alleles) and epigenetic status (low vs. high methylation) in predicting suicidal ideation. In addition, levels of SKA2 methylation in immune cells were highly correlated with morning (waking) cortisol levels suggesting a possible functional link between SKA2 methylation and stress levels. In some cases high levels of methylation at SKA2 were detected months before the reported onset of suicidal ideation raising the possibility that this biomarker was present in these individuals before they developed suicidal thoughts.

The inclusion of additional variables such as cortisol levels and anxiety further improved the predictive value of the model. This revised model was tested in another prospective cohort using genotype (T vs. C alleles), methylation rate of SKA2, and anxiety scores. This model was able to predict suicidal behavior with 71 percent accuracy that was increased to 82 percent when levels of salivary cortisol were added. Further refinement of the stringency of categorizing suicidal ideation improved the predicative value to 97 percent, successfully predicting all four eventual suicide attempts in this particular cohort.

These provocative findings are yet to be replicated by an independent group. However, they demonstrate why epigenetics is such an exciting frontier in psychiatry, providing long sought biomarkers for predicting complex and devastating human behaviors such as suicide. This study highlights the challenges of relating peripheral biomarkers to human psychopathology and behavior. This was elegantly addressed in this study by showing that the transcriptional and the epigenetic programs in the peripheral cells resemble those seen in postmortem neuronal tissue, and by relating methylation rate of SKA2 in the periphery with morning cortisol levels and suicidal ideation. It will be interesting to see if the peripheral methylation rate of SKA2 could be modified by treatment with ketamine, which has recently emerged as a rapid and promising treatment for suicidal ideation.

References

Babenko O, Kovalchuk I, Metz GA. Stress-induced perinatal and transgenerational epigenetic programming of brain development and mental health. *Neurosci Biobehav Rev*. 2015;48:70–91.

Ciabrelli F, Cavalli G. Chromatin-driven behavior of topologically associating domains. *J Mol Biol*. 2015;427:608–625.

Dias BG, Ressler KJ. Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat Neurosci*. 2014;17:89–96.

Guintivano J, Brown T, Newcomer A, et al. Identification and replication of a combined epigenetic and genetic biomarker predicting suicide and suicidal behaviors. *Am J Psychiatry*.

2014;171:1287–1296.

Guo JU, Ma DK, Mo H, et al. Neuronal activity modifies the DNA methylation landscape in the adult brain. *Nat Neurosci*. 2011;14:1345–1351.

Guo JU, Su Y, Zhong C, Ming GL, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*. 2011;145:423–434.

Hansen KH, Bracken AP, Pasini D, et al. A model for transmission of the H3K27me3 epigenetic mark. *Nat Cell Biol*. 2008;10:1291–1300.

Heller EA, Hamilton PJ, Burek DD, et al. Targeted Epigenetic Remodeling of the Cdk5 Gene in Nucleus Accumbens Regulates Cocaine- and Stress-Evoked Behavior. *J Neurosci*. 2016;36:4690–4697.

Kaffman A, Meaney MJ. Neurodevelopmental sequelae of postnatal maternal care in rodents: clinical and research implications of molecular insights. *J Child Psychol Psychiatry*. 2007;48:224–244.

Kaminsky ZA, Tang T, Wang SC, et al. DNA methylation profiles in monozygotic and dizygotic twins. *Nat Genet*. 2009;41:240–245.

Klengel T, Mehta D, Anacker C, et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat Neurosci*. 2013;16:33–41.

Kouzarides T. Chromatin modifications and their function. *Cell*. 2007;128:693–705.

Millan MJ. An epigenetic framework for neurodevelopmental disorders: from pathogenesis to potential therapy. *Neuropharmacology*. 2013;68:2–82.

Nestler EJ. Epigenetic mechanisms of depression. *JAMA Psychiatry*. 2014;71:454–456.

Niculescu AB, Levey DF, Phalen PL, et al. Understanding and predicting suicidality using a combined genomic and clinical risk assessment approach. *Mol Psychiatry*. 2015;20:1266–1285.

Pena CJ, Bagot RC, Labonte B, Nestler EJ. Epigenetic signaling in psychiatric disorders. *J Mol Biol*. 2014;426:3389–3412.

Peter CJ, Akbarian S. Balancing histone methylation activities in psychiatric disorders. *Trends Mol Med*. 2011;17:372–379.

Piatti P, Zeilner A, Lusser A. ATP-dependent chromatin remodeling factors and their roles in affecting nucleosome fiber composition. *Int J Mol Sci*. 2011;12:6544–6565.

Pirozzi F, Tabolacci E, Neri G. The FRAXopathies: definition, overview, and update. *Am J Med Genet Part A*. 2011;155A:1803–1816.

Reul JM. Making memories of stressful events: a journey along epigenetic, gene transcription, and signaling pathways. *Frontiers in Psychiatry*. 2014;5:5.

Rogge GA, Wood MA. The role of histone acetylation in cocaine-induced neural plasticity and behavior. *Neuropsychopharmacology*. 2013;38:94–110.

Schafer A. Gadd45 proteins: key players of repair-mediated DNA demethylation. *Adv Exp Med Biol*. 2013;793:35–50.

Swygert SG, Peterson CL. Chromatin dynamics: interplay between remodeling enzymes and histone modifications. *Biochim Biophys Acta*. 2014;1839:728–736.

Szyf M. Epigenetics, a key for unlocking complex CNS disorders? Therapeutic implications. *Eur Neuropsychopharmacol*. 2014;25(5):682–702.

Vizan P, Beringer M, Ballare C, Di Croce L. Role of PRC2-associated factors in stem cells and

disease. *FEBS J*. 2014;282(9):1723–1735.

▲ 1.26 Basic Systems Neuroscience

ANDRÉ ANTONIO FENTON, PH.D., AND DINO DVORAK, Ph.D.

INTRODUCTION

Systems neuroscientists study how neural circuits represent, analyze, and compute information by generating internal, neural representations of variables in the external world, to create knowledge, make decisions, and execute actions. Systems neuroscience aims to understand the causal relationships between the molecular and cellular levels of understanding brain structure and function, including the high-level mental functions that systems of neural circuits implement. For the basic systems neuroscientist, schizophrenia is at once a fascinating and frustrating subject to contemplate. Schizophrenia is fascinating because the core symptoms are of such substantial consequence. As such, schizophrenic symptomology are profound reminders that thoughts and neural representations are neurobiologically mechanistic constructions that should be within one's ability to understand and decode, despite the apparently immense complexity. Schizophrenia is frustrating because the disorder currently defies mechanistic definitions and therefore compelling mechanistic theories are lacking, which threatens to make futile the challenge and hope of mimicking the disorder and core symptoms with simpler, more tractable organisms, which is where basic science derives much of its strength. The pendulum is decidedly on an optimistic swing as we enjoy the present era of advanced research tools that embrace the fact that the problem of schizophrenia, like that of understanding any mental phenomena, is fundamentally a systems problem that engages and demands integration across all the levels of biological understanding from genes and molecules to cells and circuits and brain-regional specializations to individual and social behaviors (Fig. 1.26–1).

This chapter will review fundamentals of the basic systems neuroscience of cognitive circuit function in rodent models as may be relevant to concepts in schizophrenia. This chapter will focus on the hippocampus, not merely because it is specifically abnormal in schizophrenia, but because the hippocampus circuit is a particularly well-characterized model system within which the systems neuroscience perspective on neural computation has been developed. This point invites the opportunity to clarify what can be productively expected of an animal model of schizophrenia in particular, and any neuropsychiatric illness in general, at least from the systems perspective.

The Importance and Utility of Animal Models

The importance of animal research for translating basic science to understanding mental illness like schizophrenia is highlighted by the facts that knowledge of basic neural and pharmacological mechanisms has grown enormously while the expansion of treatment options has been much more modest. A central difficulty is the wide separation of biological levels between the behavioral and mental spheres of the clinical manifestations and the biochemical domain of therapies and potential interventions. This separation is known as the "problem of the missing middle"; whereas research tools are optimized for studies at the microscopic levels of biological function (genes and proteins and cells), as well as the large ensemble level, like the fMRI and EEG methodologies, the technologies and concepts for investigating the middle levels of biological function are far less developed or possible to implement in people. These missing levels include the temporally organized action potential discharge within ensembles, that is, subpopulations of single neurons and the temporally activated subpopulations of activated synapses and the consequences of the patterns of intracellular and extracellular current flow. Without this middle level, it has proven rather difficult to connect the nitty-gritty of receptors and transmitters to the clinical observables that drive treatment decisions. Though crucial, this middle level of investigation is only practical in animal research, and is especially well developed for rodent preparations where it is feasible to record the extracellular electrical discharge from several tens to hundreds of single neurons, concurrent with the population synaptic potentials that manifest as oscillations in the LFP.

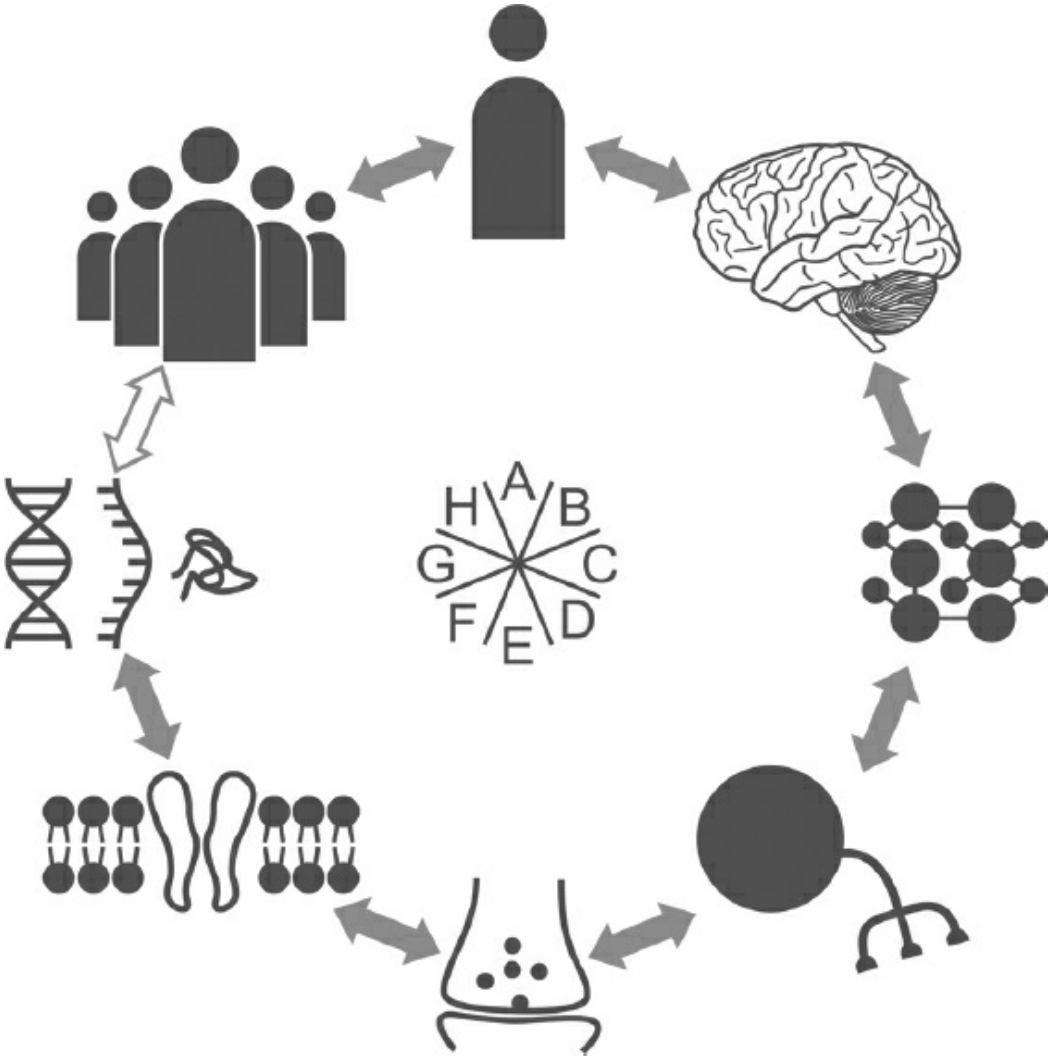

FIGURE 1.26–1. Circle of Being. This illustration shows distinct levels of biological organization, each of which has a direct mechanistic influence on the adjacent levels of biological function, and thus an indirect influence on any other level. If we start arbitrarily at the level of an individual's behavior (A) the schematic illustrates that individual behavior both influences and is influenced by social behavior (H) as well as brain activity (B). Brain activity influences and is influenced by neural circuit operations (C) and that activity influences and is influenced by the activity and properties of single cells (D). Single cell activity influences and is influenced by synaptic function (E). Synaptic properties and function are determined by membrane and channel properties (F) and those in turn influence and are influenced by genetic function, depicted here as DNA, RNA, and peptides (G). While genetic function is not known to directly influence social behavior, the effects are indirect, via the other levels of biological function depicted here. Nonetheless, much of what is known about the genetics of mental illness derives from studies that seek to correlate gene function with social and individual behavior, and this extends to attempts to use this genetic knowledge for therapeutic interventions.

Since much of this chapter will be concerned with animal research that is conducted in the effort to understand schizophrenia-related issues, it is necessary to clarify what is and is not meant by the term "animal model." Although the term is commonly used to indicate the animal is being used to mimic some aspect of the clinical disorder, this is unlikely to be a valid or productive conceptualization. The mental phenomena that are the

foundation of a schizophrenia diagnosis are hardly applicable to animals. There is great uncertainty about whether animals have the corresponding mental capacities and if they do, it is highly unlikely that they will manifest as they do in people. There has been considerable discussion over the validity of the current DSM-based diagnostic criteria, and whether there should be greater emphasis on objective biological criteria as proposed by the NIMH's RDoC project. A consequence, from the vantage of animal research, is that uncertainty in defining the disease translates into making it impossible to mimic something that is poorly defined. But even basing diagnosis on objective criteria like measurements of functions like attention, working memory, and long-term memory one has to be aware and careful that such functions can manifest in species-specific ways. Consider for example the Stroop test as a tool to assess cognitive control, surely rodents also use cognitive control, but effectively assessing cognitive control in the rodent certainly does not resemble a Stroop test. There is substantial overlap in the genetic abnormalities that predispose to the diverse conditions that are classified as, ADHD, bipolar disorder, MDD, and schizophrenia. Which disorder would be modeled by the genetic abnormality in a mouse? Even though it is clearly practical and advantageous to mimic genetic alterations associated with schizophrenia in mutant mice, one must also be wary that pleiotropy has species-specific outcomes. A gene expressed in one species may confer considerably different phenotypes in different species making it challenging to even interpret how a human genetic alteration might be mimicked in a model organism. Lastly, the penetrance of genetic variants is typically a few percent, which from the perspective of mimicry, should mean that there is a low probability of detecting that a specific genetic modification leads to an abnormal phenotype. There are a multitude of reasons not to seriously consider that animal models of neuropsychiatric disorders mimic the disorder. Yet, beyond the problem of the missing middle, animal models are among the most powerful tools to test hypotheses about schizophrenia and other neuropsychiatric disorders. This chapter does not consider any of the rodent models reviewed here to be models of schizophrenia itself. The animals are themselves tools, somewhat akin to a reagent, which, can be used to effectively evaluate the key hypotheses that drive understanding of schizophrenia.

Neural Electrical Activity and Neural Codes

A major general challenge in neuroscience that may in particular provide an opportunity for progress toward a systems-level understanding of schizophrenia is to determine how information is represented in the electrical activity of neurons. The standard, textbook treatment of this fundamental question is the "dedicated coding" hypothesis in which the action potential discharge of single neurons is interpreted to signal the

presence (or absence) of a complex stimulus like a object or the occurrence of a complex event like a movement or being in a particular location of space, as is the case for hippocampus place cells (Fig. 1.26–2). There is ample cardinal cell evidence for the dedicated coding hypothesis. These are neurons dedicated to signaling high-order stimuli and concepts, like "face" or "grandmother." Especially dramatic, well-studied examples are the place cells, head-direction, border cells, and grid cells in the rodent hippocampus and related structures that signal location, direction, environmental boundaries, and distance in external space. Another dramatic example, recorded from the same brain regions of human epilepsy-surgery candidates, are the cells with firing that is tuned to well-known individuals like the actresses Jennifer Aniston and Halle Berry. A problem with dedicated coding is that the capacity of a dedicated code is limited, to the number of cells, but a major attraction of the hypothesis is it is conceptually straightforward to conceive of how it is possible for neurons to read and transform the information in the firing of such cardinal cells, by for example, higher-order representations like of a place emerging from the integration of converging connections from neurons representing lower-order concepts like the direction and distance to borders.

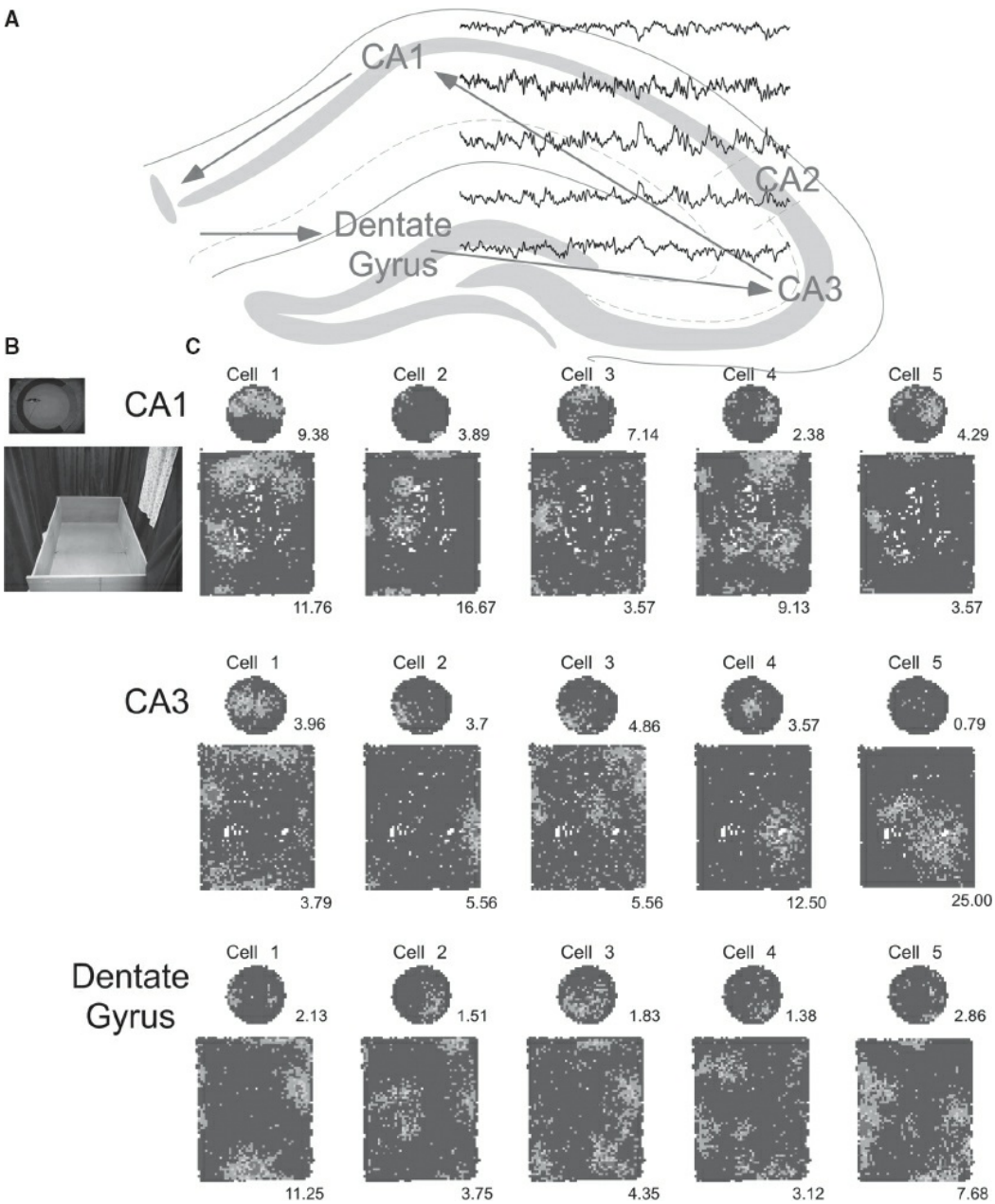

FIGURE 1.26–2. Hippocampus place cells generate an ensemble place code. **A:** Schematic of the dorsal hippocampus with the major subdivisions labeled and arrows indicating the feed-forward flow of information through the trisynaptic circuit. A 2-s example of the location-specific LFP has been superimposed such that the start of the LFP corresponds to the location where it was recorded. **B:** Photos of the two familiar environments in which place cells were recorded, a 76-cm diameter cylinder and a 1.8 × 1.4 m box. **C:** Blue-to-red color-coded average firing rate maps document the spatial tuning of place cells recorded as a rat foraged for scattered food in the cylinder and box. Each set of five cells was recorded simultaneously from the same anatomical subdivision of the hippocampus circuit. Notice that in the cylinder most place cells only discharge robustly in a single location, called the cell's place field, which suggests that the discharge provides a dedicated neural code for location. However, the fundamental nature of the place code is actually an ensemble code because in the larger environment of the box, the same cells typically discharge in multiple place fields. If each cell's discharge in the box is interpreted as a dedicated code, the activity of each cell provides an ambiguous indication of location that cannot be reliably disambiguated within the dedicated coding framework. If however, the discharge is interpreted as an ensemble code, the pattern of activity across the set of cells provides an accurate indication of location because the across-cell pattern of activity is

unique at each location. The minimum firing rate (measured as spikes/s) in the *red color* (highest firing rate) category is given below each firing rate map. The *blue color* indicates a firing rate of 0 spikes/s. Recordings by Dr. EunHye Park. (See color atlas.)

Much of the contemporary research into the neural activity correlates of information representation in animal models that are relevant to schizophrenia has been conducted and interpreted within the framework of the "ensemble-coding" hypothesis. The hypothesis asserts that information is encoded by the pattern of activity across many neurons. An ensemble code is analogous to a stadium jumbotron display that uses many light bulbs to represent a message. There is ample evidence of ensemble coding, even from the same spatially tuned cells in the hippocampal and related neocortical areas that is interpreted as evidence of dedicated-coding (Figs. 1.26–2 and 1.26–3). Analogous to the jumbotron bulbs, the activity of a given neuron contributes to multiple representations, yet no particular neuron is essential to a representation; many more representations can be encoded than there are neurons, and crucially, the same neuron can only represent one message at a single moment. Like the jumbotron, these properties require that the activity within the system of neurons (or light bulbs) is temporally coordinated so that the proper subsets of cells are coactive at the proper time intervals to represent and transmit information, perhaps at multiple time scales. Hebb's idea of the cell assembly, a set of cells that fire together to represent information by the pattern of active and inactive cells is an example of an ensemble code. While there is substantial evidence of ensemble codes operating in the mammalian and invertebrate brains, there is a major issue with the hypothesis. Unlike a dedicated code, it is unclear how neurons might detect, read, and transform information from an ensemble code. The difficulty arises from the facts that (a) the information in an ensemble code is distributed across time and space and (b) single neurons tend not to possess the necessary connectivity to collect an overview of all the cellular activity that defines the representational patterns carrying information. How might people fare, either individually or collectively, to read messages on a jumbotron display if each person could only see a subset of the light bulbs?

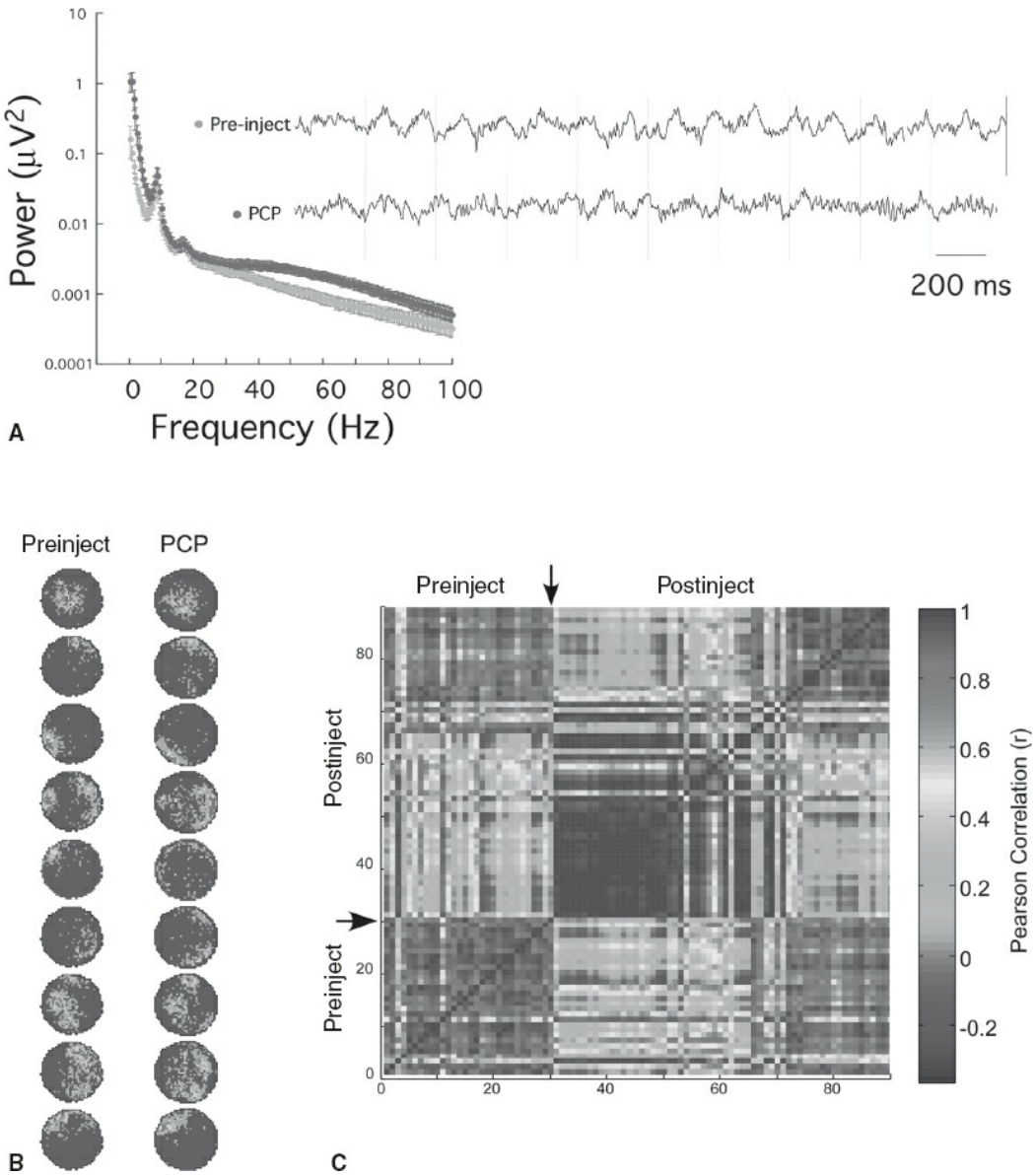

FIGURE 1.26–3. The psychotomimetic phencyclidine (PCP) discoordinates neural discharge with minimal effects on the spatial tuning of individual place cells. **A:** The blue-to-red color-coded firing rate maps of an ensemble of nine CA1 place cells that were recorded as a rat foraged for scattered food in a familiar cylinder before and after a systemic injection of 5 mg/kg of PCP. The PCP had no discernable effect on the spatial firing of the individual place cells. **B:** In contrast, when the same data are investigated as a time-varying across-cell pattern of discharge, it is obvious that the PCP radically changed the ensemble firing of the cells (PCP was administered at 30 minutes). To generate the correlation matrix that is shown, every minute the number of spikes that each cell discharged was counted and organized as an activity vector that described the temporal pattern of discharge across the population. Each activity vector was correlated with every other activity vector and the correlation is represented using the color code. The correlation at any point in the matrix represents the activity at time x correlated with the activity at time y. The correlations along the x = y line are exactly 1, because it is the activity vector correlated with itself. The activity vectors under PCP immediately changed relative to the preinjection vectors (independent or negative correlations) but the PCP activity formed a stable pattern indicated by the correlations of approximately 1.30 to 40 minutes after PCP the place cell ensemble activity switched back and forth between the preinjection and the PCP patterns until it settled back to the preinjection pattern. Notice that at three and eleven minutes during the preinjection period, the PCP pattern was expressed (in the absence of PCP). Not

only do these data demonstrate discoordination of neural activity without affecting the information tuning of individual neurons, but also they indicate that ensemble activity is multistable. The system prefers to exist in one effectively stable activity pattern and PCP radically changes the probability of observing the distinct patterns. Recordings by Dr. Hsin-Yi Kao. (See color atlas.)

The ensemble coding hypothesis offers a natural systems-level account for the cognitive and perceptual abnormalities that are observed in schizophrenia. Like the jumbotron, a neural network operating to represent information in an ensemble code must avoid the total loss of information resulting from simultaneously activating multiple representations with the same cells. A jumbotron can display two messages with the same lights, but only one message after the other as a temporally coordinated series of patterns, not simultaneously. Similarly, two cell assemblies that share neurons and are each defined as the subset of coactive neurons can represent multiple items so long as the two-coactivation patterns are separated in time. But there is catastrophic loss of information if the two cell assembly patterns coactivate, because the coactivity merges all the active cells into a distinct, coactivity-defined cell assembly, that represents something other than what is present or intended. Like any system that uses an ensemble code, a neural system can only operate effectively if mechanisms are in place to coordinate that cells within a single assembly coactivate, and that multiple coactivity defined assemblies do not coactivate. We will presently examine neural oscillations as one class of these coordinating mechanisms, but let us first consider what kinds of conditions present a challenge to the ability to coordinate multiple neuronal ensemble representations of information.

A hypothesis for the information processing, perceptual, and cognitive impairments in schizophrenia falls out of the foregoing discussion of ensemble coding. Loss or aberrant coordination of neural activity will generate information processing failures in a network of neurons and such failures are more likely when multiple sources of conflicting information compete for the same neural resources. Such ideas have been put forth in a number of ways; in essence, the notion asserts that discoordinated patterns of electrical neural activity are the core deficit in schizophrenia from which other deficits derive. This “discoordination” hypothesis is a systems-level attempt to account for the cognitive deficits in schizophrenia. It asserts that by and large, cognitive and perceptual deficits of the disease are caused by the aberrant coordination of the many individually “normal” neurons that comprise a neural network or circuit. This discoordination is asserted to occur both within and between neural networks. Note that the coordination and discoordination of neural activity can only really be investigated and assessed directly by measuring the timing relationships of neural discharge among two or more neurons or other measures of neural activity, such as the EEG that reflects spatiotemporal patterns of synaptic currents.

An instructive example of the neural discoordination that is hypothesized to underlie the cognitive deficits in schizophrenia is demonstrated by the effect of the psychotomimetic phencyclidine (PCP) on rat hippocampus place cells. The drug selectively increases the power of gamma oscillations in the hippocampal LFP (Fig. 1.26–3A) but it does not cause a detectible change in the spatial tuning of hippocampal place cells (Fig. 1.26–3B). Interpreting the effect on place cells from the dedicated coding perspective, one must conclude that PCP has a minimal effect on the discharge of these cognition-related cells. Remarkably, the conclusion is rather different if the same data are interpreted from the ensemble-coding perspective, and evaluated as an across-cell pattern of discharge (Fig. 1.26–3C). The correlation matrix shows that PCP has a large and immediate effect on the temporally organized pattern of discharge across the sample of hippocampal place cells. PCP causes the familiar, stable pattern of discharge to change to a new, different stable pattern of discharge that does not correspond to the familiar representation of this familiar environment. As the PCP effect begins to wear off, about 30 minutes after administration, the neural activity pattern alternates between the preinjection and the PCP patterns before it settles into the familiar preinjection pattern. It is notable that the PCP discharge pattern had been transiently expressed at minutes 3 and 11 of the recording, *before* PCP was administered. These data indicate that hippocampal neural activity is multistable and reveals that PCP changed the probabilities of the system being in the familiar preinjection and the aberrant PCP states, where the network state is defined by the coordinated temporally defined ensemble pattern of discharge, and not on the basis of a measure of the discharge properties of the individual cells.

The discoordination hypothesis conceptualizes the coordinating processes as distinct or at least dissociable from the more unitary processes that characterize neural activity like the firing rate and the tuning curves of individual neurons, and the frequency of oscillations in the field potentials that can be recorded as EEG from the scalp or as an LFP when it can be localized to a specific location of a neural circuit. However, since both the unitary physiological properties of cells and their coordinating processes result from the deployment and operation of the cellular machinery, including protein channels, at some sufficiently detailed level, alterations of the coordinating activity also influences these unitary properties. As currently conceived, the discoordination hypothesis predicts that basic properties and neural network functions like the tuning of neural responses to stimuli, the behavioral responses to simple, unitary stimuli and the expression and use of memory will be relatively normal when conflicting information is minimal. Whereas functional and discoordination abnormalities will be prominent with increasing likelihood under complex conditions that require ignoring and selectively using the accessible information to manage and meet competing demands. A canonical example

of such a situation has been formalized as the Stroop test. The basic form of the test presents stimuli to the subject that provide two distinct forms of information, such as the words of colors, for example “blue” and “red” written in either a congruent (“blue” written in blue) or incongruent color (“blue” written in red). Because the subject is instructed to report the color, and thus ignore the meaning of the word, to be effective the subject’s relevant brain circuits must selectively represent the color and not the meaning. Using tasks that present a cognitive control challenge like this, one can observe the discharge across an ensemble of neurons in rodents coalesce into coactivity-defined groups or assemblies of cells that represent one class of information but the activity toggles between distinct patterns of coactive cells each representing one or the other class of information. At any moment the network of cells prefers to represent the information with the greater behavioral relevance.

Origins of Local Field Potentials and Neural Oscillations

Whether it is an action potential or a synaptic potential, neural electrical activity derives from changes in the transmembrane potentials of excitable cells like neurons. The extracellular currents emanate through the conductive extracellular medium and are recorded as field potentials, the origins of which can be localized to the vicinity of an intracranial electrode and are hence called an LFP, or in the case of the EEG recorded with a scalp electrode. The spatial localization of the origins of the EEG potentials is much reduced.

Extracellular fields are generated by the flow of positively or negatively charged ions through the semi-permeable membranes of neurons. When positively charged ions (e.g., Na^+ or Ca^{2+}) leave the extracellular space into the cell, they cause local electrical sinks, which appears as a negative deflection in the LFP, measured at the proximity of the active membrane. Passive current in an opposite direction balances the current influx and is termed the return current. Excitatory neurons, such as those that release the neurotransmitter glutamate, cause a current influx in postsynaptic cells and therefore a depolarization of the postsynaptic membrane from its resting potential and an increased probability of a postsynaptic action potential. Inhibitory cells typically release GABA, which in the postsynaptic cell causes hyperpolarizing and/or shunting inhibition where negative chloride ions can enter or leave the cell through chloride ion channels associated with the GABA receptors. This hyperpolarizes or stabilizes at rest the membrane potential of the postsynaptic cell, both of which makes postsynaptic action potentials less likely. Currents through the membrane are extremely small and therefore any measurable changes in the LFP are a result of many synchronized events. The largest contribution to the LFP comes from slow and synchronized currents activated by inhibitory

neurons. While fast Na^+ and K^+ conductance-based action potentials generate the largest extracellular currents (an extracellular action potential or spike) they only have a limited contribution to the LFP spectrum below 100 Hz, because of the short duration of the action potential events (<1 milliseconds).

Neuronal oscillations are periodic fluctuations in neural activity in the CNS. Oscillations can be observed across multiple spatial scales. At the level of a single neuron they either have a form of rhythmic, subthreshold changes of the membrane potential or rhythmic changes in action potential activity. The synchronized activity of large numbers of neurons gives rise to macroscopic oscillatory electric fields that can be observed in the LFP or electroencephalogram (EEG). Oscillations differ in several aspects such as the typical frequency, anatomical localization, generating mechanism, relative power, and relationship to the ongoing brain state. Here, the importance of theta (5 to 12 Hz) and gamma (30 to 100 Hz) oscillations is stressed, because their biophysical origins are well known and their interactions provide a functional coordinating infrastructure within which information-carrying action potentials are organized and controlled.

Theta oscillations can be detected in brain structures such as hippocampus, entorhinal cortex, amygdala, sensory cortex, and prefrontal cortex. The theta rhythm has been associated with encoding and retrieval of memory, synaptic plasticity, long-range synchronization, and top-down control. The hippocampal theta rhythm is the most prominent oscillation associated with the online state of the hippocampus; it is most notable during intake of sensory information such as active locomotion, whisking, and sniffing although it is also consistently present in the REM stage of sleep. Theta rhythmicity is generated by pacemaker cells located in the medial septum-diagonal band of Broca, which provides GABAergic and cholinergic drive to hippocampal CA1 interneurons. On the other hand, excitatory CA3 hippocampal subfield and entorhinal cortical inputs to the dendrites of the CA1 hippocampal subfield may also produce stronger and higher frequency theta oscillations. The presence of multiple, weakly coupled theta oscillators might allow the necessary flexibility in the selection of appropriate cell assemblies for various cognitive tasks. One of the most important features of the theta rhythm is its ability to modulate the discharge of interneurons that generate local gamma oscillations by the postsynaptic inhibitory effects of their release of GABA. This mechanism was proposed as a general mechanism for chunking information throughout the brain. The result of such bias or so-called “nesting of oscillations” reflects the observation that the slower theta rhythm modulates the faster gamma oscillations at a specific phase of its rhythm (Fig. 1.26–4).

Gamma oscillations are the most ubiquitous rhythm; they have been localized in all brain structures where fast GABA-mediated inhibition is

provided by soma-targeting interneurons. It is not surprising that something so wide spread has been associated with perception, attention, memory, synaptic plasticity, and consciousness. In contrast to the theta rhythm, gamma oscillations are more locally generated, faster oscillations that cover a wide range of frequencies from 25 Hz to approximately 150 Hz. The upper limit of the gamma band has been a subject of intensive discussions during the past decade. While the gamma rhythm was originally defined as the 35- to 85-Hz band, in more recent work, authors have subdivided the gamma band into slow (<50 Hz) and "high" or "fast" (>50 Hz) components. Furthermore, fast gamma was recently extended up to 140 Hz. However, caution is necessary for oscillatory activity above 100 Hz because of their possible contamination by the wideband spectral content of action potentials and associated after potentials. Gamma oscillations are generated by the interplay of excitatory and inhibitory neurons. GABAergic interneurons, especially interneurons that express the Ca^{2+}-binding protein parvalbumin, have been shown to be a particularly important part of the pacemaker system that generates high-frequency oscillations. These interneurons receive glutamate receptor-mediated excitatory inputs. Transient excitation of the network of parvalbumin-expressing interneurons, which are mutually interconnected through electrical gap junctions and GABAergic synapses, generates rhythmic IPSPs, which can synchronize the firing of large populations of pyramidal neurons.

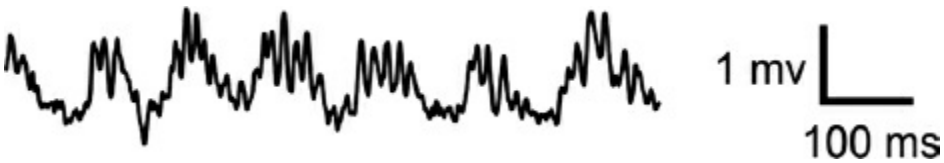

FIGURE 1.26–4. Phase–amplitude coupling of distinct neural oscillations. This LFP example illustrates "nesting" of fast (60 to 100 Hz) gamma oscillations within theta (~8 Hz) oscillations. There are approximately 6.5 theta cycles depicted here. Notice that the largest gamma oscillations tend to occur preferentially following the peak of the concurrent theta oscillation.

Canonical pattern generating networks are shown in Figure 1.26–6. Feed-forward inhibition (Fig. 1.26–6 left) provides a mechanism for attenuation of the afferent excitation, where a distant excitatory cell activates a local interneuron or a group of local interneurons connected via gap junctions, which in turn inhibits multiple principal cells in the local network. This mechanism increases the temporal precision of firing by narrowing temporal windows of discharge probability of the principal neurons. Because the periodic repetition of the "window of opportunity" of the principal neuron discharge is defined by the time constant of the inhibitory neuron, this pairing of excitatory and inhibitory neurons is proposed as one of the generating mechanisms of gamma and other types

of oscillations. In feedback inhibition (Fig. 1.26–6 middle) the interneuron receives excitatory glutamatergic input from the local principal cell and in turn provides inhibition to a specific compartment of the principal cells in the same network. Feedback inhibition provides a mechanism for reducing the excitability of principal cells and a mechanism for inhibition of a specific dendritic segment of the principal cell. Lateral inhibition (Fig. 1.26–6 right) provides an effective "winner-take-all" mechanism for selective activation of distinct ensemble patterns by increased activation of selected principal neurons (left population), which in turn inhibits other networks of neurons, therefore reducing their chance to fire (right population).

Gamma activity is not actively propagated; instead its role in neural coordination is to synchronize the spiking of neurons. From the perspective of the dedicated-coding hypothesis it may not seem intuitive that inhibition would be synchronizing; however, from the perspective of the ensemble-coding hypothesis, inhibition synchronizes by defining narrow windows of opportunity for principal cells to discharge. The coupling between spiking activity and gamma oscillations could arise because local inhibitory neurons (crucial for generating gamma) fire preferably at the trough of the local gamma cycle. This mechanism makes the spiking of excitatory projection neurons more likely at the offset phase, when inhibition is weaker. This feature of coupling between gamma and spiking has been hypothesized as a central aspect of cortical processing. Several theories have also been proposed for the role of the gamma rhythm. According to one hypothesis, the phase of firing relative to gamma encodes the stimulus strength, because stronger afferent excitation will permit firing to occur earlier in the window of opportunity that is set by ongoing gamma oscillations. Another idea is that gamma influences the communication between neuronal populations when there is gamma phase-coherent spiking in the two populations.

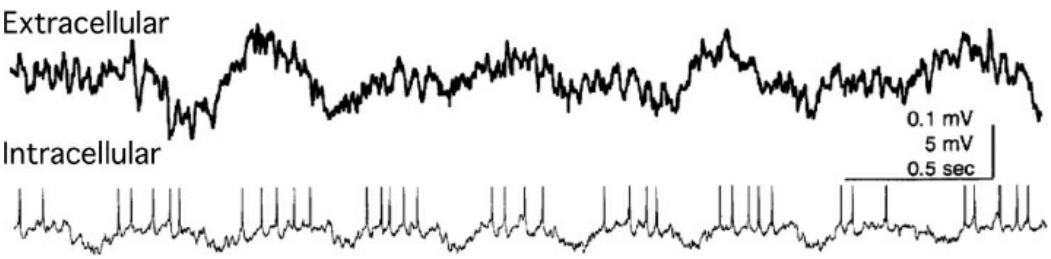

FIGURE 1.26–5. Theta-paced gamma frequency discharge of a parvalbumin-expressing basket cell type interneuron. A potential source of theta and gamma oscillations in the field potential is illustrated in this pair of simultaneous records from an intracellular recording of a parvalbumin-expressing basket cell and an extracellular recording of the LFP. (Adapted from Ylinen A, Soltesz I, Bragin A, et al. Intracellular correlates of hippocampal theta rhythm in identified pyramidal cells, granule cells, and basket cells. *Hippocampus.* 1995;5:78–90.)

Oscillations

Various forms of neural synchrony, the coordinated, synchronized activation of same-function cells, and the active desynchronization of different-function cells have been proposed as a fundamental mechanism of neural computation. Neural synchrony appears across multiple scales. On the cellular (microscopic) scale, neural synchrony mediated by action potential and synaptic activity, manifests as a transient coupling between functionally distinct neuronal subpopulations. This coupling of neural subpopulations creates a form of synchrony that manifests as oscillations in the LFP and the synchrony that is observed on a mesoscopic scale of neural networks. The synchronized activity of a large number of neural networks in turn gives rise to the signals that are observed at the macroscopic level such as the scalp EEG.

An important aspect of neural synchrony is its cross-scale nature. For example, neural oscillations such as gamma synchronize action potentials within and between networks of neurons. Furthermore, slow oscillations, such as theta, synchronize transient bursts of fast oscillations such as gamma (Figs. 1.26–5 and 1.26–6). Coordinated neural interactions therefore appear across scales and across frequency bands. Although it is an extremely important and well-studied area, because of the difficulty of recording spiking data in clinical settings, spike-field synchrony will not be discussed further, instead synchrony within and amongst the oscillations in field potentials becomes the main focus.

Oscillations can be observed at multiple temporal and spatial scales. Cortical oscillations can be recorded noninvasively by conventional EEG. Circular cup electrodes of approximately 10 mm in diameter are placed on the scalp with a small amount of conductive gel between the scalp and the electrode. In order to fix the electrode to a specific location of the head, electrodes are sometimes embedded in a fabric cap, which stabilizes their location during the recording. The location of electrodes has been standardized, as in the international 10 to 20 systems. Signals from electrodes are usually recorded using differential amplifiers where the common reference can be one of the electrodes (typically *CZ* electrode at the center of the 10 to 20 system) or linked mastoid electrodes. Scalp EEG has several important limitations. The tissue between the cortex and the electrode is formed by the CSF, pia mater, arachnoid, dura mater, skull, periosteum, and scalp. The tissue is highly heterogeneous, most importantly in terms of its conductance. The presence of the low conductance bone causes signal smearing that decreases spatial resolution and source localization. Another problem is related to the distance between the electrode and the recorded neural tissue. Because the intensity of the electric field falls with the square of the distance, the measurable voltage change on the electrode (several tens of microvolts) corresponds to the

synchronized activity of approximately 6 cm² of cortical gyri tissue corresponding to 60 million neurons forming the electrical dipole layer. The last important limitation of the scalp EEG is in its frequency range. Because the spectral magnitudes of brain oscillations are inversely related to their frequency, according to a power law, it is progressively more difficult to record higher frequencies on the scalp because of the progressively worse signal/noise ratio. This fact limits the upper bandwidth of the EEG to approximately 70 Hz although several techniques have been used to circumvent this limitation such as trial-based averaging (event-related potentials, ERPs) or high-resolution EEG.

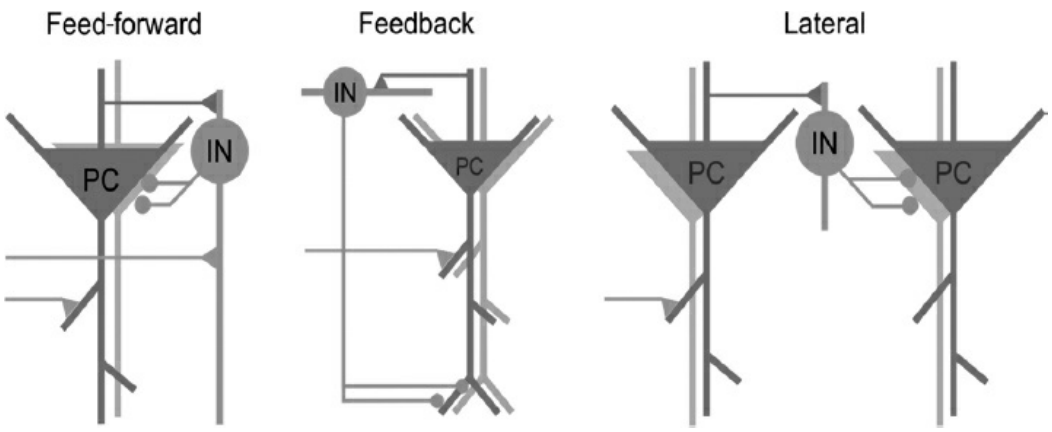

FIGURE 1.26–6. Basic mechanisms of network excitation–inhibition coordination. In these schemes, each cell type that is depicted, principal cell (PC) and inhibitory interneuron (IN), represents a population of the cell type, not an individual neuron. Feed-forward inhibition (**left**) provides a mechanism for attenuation of the afferent excitation. Feedback inhibition (**middle**) provides a mechanism for reducing excitability of principal cells and a mechanism for inhibition of a specific dendritic segment of the principal cell. Lateral inhibition (**right**) provides an effective "winner-take-all" mechanism for selective activation of distinct ensemble patterns.

Higher spatial and frequency resolution than the EEG can be achieved by ECoG by using subdural electrodes which come into direct contact with brain tissue. In humans, this technique is limited to presurgery monitoring of epilepsy patients where epileptogenic source localization provides the justification for this highly invasive technique. Compared to standard EEG electrodes, subdural electrodes are usually implemented in the form of grids or strips with typical electrode diameters of 2 to 3 mm and spacing of 10 mm between the electrodes. The close contact between the electrode and the neural tissue allows recording of LFPs up to 500 Hz.

The highest spatial and frequency resolution can be obtained by using penetrating high-impedance microelectrodes placed directly into the neuronal dipole sheet to record the electric signal of a small groups of neurons. Microelectrodes come in several different forms such as high-density microelectrode arrays used primarily in NHP research or micromachined multicontact silicon probes used primarily with rodents.

When a microelectrode is placed in the CA1 region of the hippocampus, it can record the summed activity of up to 1,000 neurons within a radius of 140 μm.

The remaining parts of the recording system are principally the same no matter if scalp EEG or LFPs are acquired. The signal is first routed into a signal follower, an operational amplifier with unity gain which reduces the signal impedance in order to minimize additional electric noise caused by the signal's passing through the cable that connects the subject to the main amplifier and digitization stages. After this stage, the signal is amplified using one or multiple operational amplifiers so the signal's voltage is increased from tens-to-hundreds of microvolts to several volts. At this point, the signal is high-pass filtered to avoid saturation of the analog-to-digital converter (ADC) and then low-pass filtered to avoid aliasing effects in the ADC stage and finally routed to the ADC for digital conversion. Output of the ADC is a binary number corresponding to the voltage between the active electrode and the reference electrode. The binary number is stored in computer memory for further processing.

Analysis of Oscillations

The most basic approach for the analysis of oscillations is based on Fourier analysis, where the time representation of the voltage signal is transformed into the frequency domain using the FFT. The result of this transformation is a power spectra, which defines the original voltage time series as a function of oscillations of different frequencies. Specifically, the power spectrum is the amount (power) of each frequency-defined sine function from which the original voltage time series signal is composed. Importantly, computing the FFT requires a sufficiently long time series of data, and it is assumed that the frequency content of the signal did not vary in any meaningful way in time. Thus, appropriate analysis using the power spectra assumes stationarity of the signal and as such, cannot provide any information about the dynamics of oscillations, should the stationarity assumption be invalid. Given that neural signals are time varying, especially those that might contribute to cognitive processes, it is important to recognize that several methods that have been developed for digital signal processing engineering applications are now available and are growing in their use and utility for systems neuroscience applications. One such development is the called the Short Time Fourier Transform (STFT), an extension of the Fourier Transform. The STFT adds the time dimension to the power spectra, which allows analysis of the dynamical changes of the power at individual frequencies over time. Limitations of the time–frequency resolution of the STFT had been partially overcome by Wavelet analysis, which enables time–frequency analysis at various levels of detail or so called scales. When a complex wavelet is used, one can also estimate the instantaneous phase and amplitude information from the signal. This

concept is similar to the application of a pass-band filter followed by the Hilbert Transform, which will also yield a time series of the phase and amplitude of a frequency-defined signal. The ability to measure the time varying spectral content of EEG and LFP signals has made it possible to study the interactions of distinct oscillations, like that of the theta and gamma oscillations that can be see in the LFP fragment in Figures 1.26–3 and 1.26–4. Interactions of oscillations can be studied by methods of bivariate data analysis. In general, these methods analyze the relationships between various aspects of signals such as their amplitudes in time (cross-correlation), their phase relationship (coherence, phase locking), or the relationship between amplitude and phase (cross-frequency coupling, CFC). The most advanced methods such as phase locking and CFC are based on a similar scheme. First, both signals under analysis are band-pass filtered, after that the Hilbert Transform separates the phase and amplitude information of both signals. In the next step, the phase and/or the amplitude information is compared using indicators such as circular variance or mutual entropy. The last important step is testing of the statistical significance of any phase and/or amplitude relationships, for example by using surrogate data techniques.

There are several types of synchrony that can be readily measured. Amplitude synchrony (Fig. 1.26–7 left) is maximal when the amplitudes of two signals covary and coincide in time. This form of synchrony can be detected by linear cross-correlation. Phase synchrony (Fig. 1.26–7 middle) is maximal when the difference between the instantaneous phases of two signals is constant. Phase synchrony can be detected by measures such as the phase locking value or coherence, although amplitude synchrony also contributes to coherence. Phase–amplitude synchrony (Fig. 1.26–7 right) is maximal when the amplitude of a fast signal like gamma is modulated by the phase of a slow signal, like theta. That coordination is called phase–amplitude coupling (PAC), but because such interactions also happen between different frequency bands, it is also known as CFC. The term "nesting" has also been used to describe PAC and it emphasizes the fact that the phase of the slow oscillation is modulating the amplitude of the fast oscillation within the cycle of the slow oscillation. While all forms of neural synchrony have a role in neural circuit and brain function, the present goal is to be illustrative rather than comprehensive so the discussion of PAC will be expanded to illustrate key principles of how neural oscillations participate in neural circuit computation.

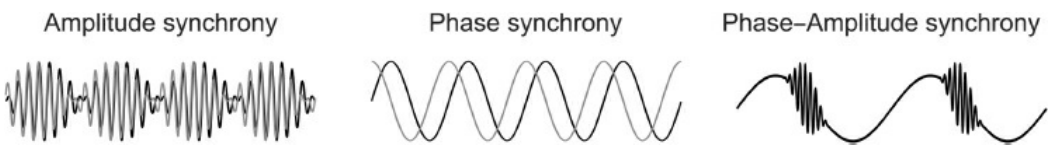

FIGURE 1.26–7. Types of synchrony. In amplitude synchrony (**left**), the amplitudes of two signals (depicted as *gray* and *black*) follow similar temporal patterns. In phase

synchrony (**middle**), the phases of two signals maintain a constant phase offset over time. In phase–amplitude synchrony (**right**), the amplitude of the fast oscillation appears at a specific phase of the slower oscillation.

Phase–Amplitude Coupling

PAC has been observed across a variety of brain regions as well as across different animal species. Although how PAC is computed is not considered here, the reader referred to a review of the methods and their limitations. The power of fast gamma oscillations (80 to 150 Hz) is phase-locked to ~8 Hz theta oscillations in human neocortex and different behavioral tasks evoke distinct patterns of theta–gamma coupling across cortex. The properties of theta-slow gamma (25 to 50 Hz) coupling in the human neocortex reflect the number of items stored in working memory. Increasing memory load is associated with a reduction in the frequency of the slow oscillation, consistent with the idea that longer oscillation cycles are required to accommodate an increased number of information representations, each of which is paced by a gamma cycle.

PAC is thought to coordinate and route information flow into computationally different compartments of a neural circuit. In the rat CA1 region of the hippocampus, the theta phase of local oscillations is demonstrated to modulate the amplitude of two somewhat overlapping gamma frequency bands, slow and fast gamma. A role of the slow delta (1 to 4 Hz) rhythm in the sensory selection of the task-relevant input stream has also been demonstrated. Although the physiological mechanisms of PAC are not well understood, it is hypothesized that perisomatic inhibitory basket cells in the hippocampus play a role in the spatiotemporal organization of cell assemblies by firing theta rhythmic trains of action potentials at gamma frequency and therefore this activity contributes to both rhythms and their phase–amplitude and phase–phase relationships (Fig. 1.26–5).

Several conceptual frameworks were also proposed for the computational role of PAC. The concept of a cross-scale organization of neural activity offers a possible neural mechanism for integrating information between several functionally distinct networks, to accomplish perceptual binding, selective attention, cognitive control, and the recruitment of specific computational cell assemblies. Neural activity in macroscopic (slow oscillations), mesoscopic (high frequency oscillations), and microscopic (single neuron activity) scales are braided together such that a progressively faster activity occurs within a short time window of a slower activity. A particularly clear example is seen in the hippocampus LFP of the freely behaving rat, the phase of the hippocampal theta rhythm modulates the amplitude of hippocampal gamma oscillations (see Fig. 1.26–4). In turn, gamma oscillations provide millisecond-scale “windows of opportunity” when the discharge probability of hippocampal pyramidal

cells is the greatest. It has been suggested that because of conduction delays, slow oscillations such as theta are particularly suitable for synchronization of functional networks over longer distances. On the contrary, fast oscillations such as gamma might transiently synchronize neuronal action potential discharge into same-function assemblies over relatively short spatial scales that can be as long as a few hundred microns. The interplay between the two oscillations might then provide a mechanism for spatial integration from a local to a global scale. In this scheme, the faster gamma oscillations that emerge at the specific phase of the slower theta oscillations might recruit subsets of cells that signal the same information into a cell assembly of co-active cells for neuronal information processing. Theta might then serve as an integration mechanism between several functional networks, each associated with a particular gamma episode, across longer periods of time. Another conceptual framework for the functional significance of PAC is based on the idea of phase coding, where PAC is a mechanism for encoding different information at different phases of the theta cycle. In this scenario, faster gamma oscillations appearing at specific phases of slower theta oscillations are hypothesized to activate distinct neural ensembles, which encode distinct representations by the across-cell pattern of discharge as illustrated in Figure 1.26–4.

Discoordination of Neural Discharge and Oscillations

Disturbances of neural synchrony are proposed to underlie the cognitive deficits that associate with diverse brain disorders that include schizophrenia, autism, Alzheimer disease, Parkinson disease, and epilepsy. It is not a simple matter that dysfunction emerges from too little synchrony, rather the idea is that neural synchrony is optimized for cognitive function and either too much or too little manifests as abnormal cognition. Gestalt perception of Mooney faces (black and white contour pictures of human faces) is impaired in schizophrenia patients and the impairment is associated with reduced phase synchrony in the gamma band. It has also been reported there is a selective alteration of theta frequency oscillations, suggesting deficits in the large-scale synchronization while local synchrony in the gamma band is largely intact. Several studies have found that the positive symptoms, specifically visual and auditory hallucinations and delusions are correlated with enhanced amplitude and phase synchronization of evoked and induced beta- and gamma-band activity. On the other hand, negative symptoms have been associated with both enhanced and reduced high frequency oscillations. In fact, there is substantial evidence that resting, baseline (i.e., unprovoked) gamma power is elevated in schizophrenia. This might in part explain why stimulus-evoked gamma elevations appear to be attenuated in schizophrenia patients. While exact mechanisms are not known to explain these diverse

findings of neural discoordination in schizophrenia, it is hypothesized that reduced parvalbumin expression in the cortical GABAergic interneuron population indicates GABAergic dysfunction and this results in impaired gamma band synchronization, which is crucial for the local synchronized firing of large populations of pyramidal neurons and may lead to deficits in the temporal coordination of distributed neural activity.

Neural Discoordination in Neurodevelopmental Models of Mental Disorders

An important development in understanding mental illness in general and schizophrenia in particular, is the recognition that these are neurodevelopmental disorders. The effect of disturbing neurodevelopment has been investigated using animal models, which provide some of the most compelling support for the assertion that disorders like schizophrenia have a neurodevelopmental origin. In the MAM-E17 rat model the mitotoxin methylazoxymethanol-acetate (MAM) is administered to pregnant dams on embryonic day 17 to preferentially disrupt ongoing embryogenesis of the hippocampus and prefrontal cortex. Adult MAM rats exhibit deficits in cognitive flexibility and hyperresponsiveness to dopaminergic agonists, hippocampal glutamatergic dysfunction, and reduced parvalbumin-positive interneuron expression. These abnormalities accompany reduced cortical up-state and down-state transitions that underlie state-selective neural responses and slow oscillations in sleep. The coupling of hippocampal (150 to 250 Hz) ripple oscillations and sleep thalamocortical spindles in prefrontal cortex is also disrupted in adult MAM rats. Thus the MAM neurodevelopmental insult causes a wide variety of abnormalities with general relevance to schizophrenia and clear evidence of discoordination of neural network function.

The neonatal ventral hippocampus lesion (NVHL) rat is one of the most studied neurodevelopmental models in schizophrenia research. It was developed on the hypothesis that an early life brain insult causes aberrant development of neural circuits. The excitotoxic lesion of ventral hippocampus occurs on postnatal day 7, corresponding to the third trimester of pregnancy in human brain development. NVHL studies provide some of the strongest evidence that early brain insults can disturb neural circuit and cognitive functions relevant to schizophrenia. Adult NVHL rats demonstrate abnormalities that are relevant to mental illness, especially schizophrenia. These include abnormal cortical dendritic morphology reduced GABAergic metabolism, excessive dopaminergic signaling, hyper responsive behavior to dopaminergic agonists, and various behavioral abnormalities that emerge in adulthood. There is increased prevalence of seizure in schizophrenia and the incidence of epileptiform activity is increased in adult NVHL rats. Making the lesion in adulthood does not produce the same NVHL abnormalities, consistent with a developmental component to the consequences.

Single neuron responses in the prefrontal cortex and nucleus accumbens are aberrant in adult NVHL rats. Prefrontal firing normally decreases in response to dopamine from stimulation of the VTA but prefrontal activity increases in NVHL adults. Such disruptions of excitation–inhibition coordination occur, in part, because of impaired interneuron activation by dopamine. Disinhibition of the adult NVHL neocortex is associated with cognitive impairments, such as in decision tasks. In control rats, the adolescent prefrontal network is also disinhibited and less coordinated than in adults, suggesting that NVHL information processing networks are functionally immature.

NVHL cortical network activity patterns are discoordinated within and between the hippocampus and prefrontal cortices during cognitively challenging behaviors. During an odor discrimination task, rats generate behavioral sampling event-related increases in low frequency (1 to 30 Hz) prefrontal oscillations. These increases are weak in NVHL rats. Consistent with reduced oscillations being an indicator of reduced excitation–inhibition organization, prefrontal neurons in NVHL rats are hyperresponsive when the animals sample odors and this was exacerbated on large reward trials when dopamine levels are expected to increase. These abnormalities were not evident during baseline behavior, when differences in oscillatory power are hard to detect, indicating that network activity patterns are sensitive to what the subject is doing and presumably also thinking. It follows that group and treatment differences tend to emerge during cognitive effort, but it is also the case that physiological differences may be irrelevant to cognitive performance and this is difficult to know unless the behavior and physiology are assessed concurrently.

The interdependence of cognitive behavior and network activity is evident in studies of NVHL adults performing an active place avoidance task on a slowly rotating arena. The task requires intact hippocampus function and maintained hippocampal long-term potentiation (LTP) as well as cognitive control to use stationary spatial cues and ignore irrelevant rotating olfactory cues. NVHL adults are impaired at place avoidance but with extended training they can approach the performance of controls. NVHL rats are not impaired in adolescence, nor are they impaired if the cognitive control demand is reduced, indicating that an increased incidence of epileptiform activity is not sufficient to cause basic performance, learning and memory deficits. Hippocampal and prefrontal LFPs are reduced across a wide band of frequencies during walking but this is insufficient to explain the cognitive deficits, which depend on experience and the availability of cues. Compared to during the task, interhippocampal and interprefrontal phase and amplitude synchrony of oscillations is reduced during home cage behaviors such as quiet rest, grooming, and eating, except in the gamma range, which is similar in NVHL and controls in all conditions. Interhemispheric phase synchrony increases during place

avoidance compared to home cage behaviors, but prefrontal synchrony is not different between NVHL and control groups during place avoidance. Theta phase synchrony between the hippocampi is weaker in NVHL adults, even when the NVHL rats are performing as well as controls. These differences in adulthood were prevented by cognitive training in adolescence, presumably because of persistent changes of neural network function. The neural discoordination and cognitive abnormalities are also partly attenuated by reducing excitability with the antiepileptic ethosuximide, consistent with the notion that excitation–inhibition disbalance causes discoordination in NVHL adults, and the possibility that this discoordination may be sufficient to explain the cognitive abnormalities.

Summary

The systems neuroscience perspective embraces the diverse levels of biological organization and when this is practiced in an integrative framework, the goal is to make measurements at the different biological levels from the same subject to look for cross-scale relationships and principles. This systems approach can be especially powerful when it is applied to testing hypotheses using animal models that are associated with mental illness, such as models that are relevant to schizophrenia, because it is extremely difficult, if not impossible to perform the requisite multiscale systems-level analyses in a clinical population. As a closing example, let us consider the loss of GABAergic tone that is associated with schizophrenia and manifests as a reduction of parvalbumin-immunoreactivity that is associated with GABAergic interneurons like the one depicted in Figure 1.26–5. Such compromised GABAergic function is consistent with the systems-level discoordination hypothesis that asserts that cognitive dysfunction is a result of multi-scale disturbances that impact the different levels of biological organization (Fig. 1.26–1). According to this viewpoint, genetic abnormalities and altered developmental programs cause, at the level of synapses and cells, impaired molecular interactions due to inefficient calcium binding, which is the normal function of parvalbumin. This can have diverse consequences including abnormal synaptic signaling, and the dysregulation of synthesis of synaptic proteins, all of which would contribute to disbalancing excitation and inhibition at the level of neural circuits. Because these circuits are dynamic and nonlinear, at least in part because of the network mechanisms depicted in Figure 1.26–6, the consequences of reduced parvalbumin expression in GABAergic neurons is a time varying excitation-inhibition discoordination. The excitation-inhibition discoordination results in discoordinated neural activity that manifests as aberrant temporal patterns of discharge in neural ensembles as depicted in Figure 1.26–3, disruptions of the kind of spike-field synchrony that is illustrated in Figure 1.26–4, and the discoordination of

the various forms of synchrony between oscillations in the field potentials that reflect population synaptic activity. According to the hypothesis, these diverse forms of discoordination cause cognitive impairments exemplified by failures of cognitive control and the disorganization of purposeful goal-directed behavior.

From the systems perspective, it is not obvious that standard pharmacological interventions could be effective for improving cognitive deficits, at least because the systems view is often that interventions are needed at specific brain locations at specific times, depending on the particular ongoing cognitive effort. More optimistically, as exemplified by the research on the NVHL rat neurodevelopmental model, the systems neuroscience perspective suggests there is an opportunity to harness experience itself to aid in the location- and time-specific correction and adjustment of dysfunctional neural circuit function. The hope and promise is that as knowledge of the detailed recruitment and input-specific activation and suppression of neural circuit operations develops and becomes more elaborate, it will be possible design and implement specific experience-based therapies, like CBT and computer-based cognitive training programs that harness the mechanisms of synaptic plasticity to tune and adjust dysfunctional neural circuits and in so doing, compensate and perhaps even correct impaired cognitive behavior. Although it is still a hypothesis, there is substantial optimism that if pharmaceutical, or brain stimulation-based interventions to regulate neuroplasticity can be coupled with the cognition-driven exercise of specific neural circuits, then this combination can be a powerful strategy to use the self-organizing neural consequences of experience to improve or normalize abnormal brain operations. There is a much work to do but also reason for considerable optimism, because basic systems neuroscience has never had such a powerful and vast toolbox to deploy.

References

Buzsaki G. Neural syntax: cell assemblies, synapsembles, and readers. *Neuron*. 2010;68:362–385.

Buzsaki G, Anastassiou CA, Koch C. The origin of extracellular fields and currents–EEG, ECoG, LFP and spikes. *Nat Rev Neurosci*. 2012;13:407–420.

Cross-Disorder Group of the Psychiatric Genomics Consortium; Lee SH, Ripke S, Neale BM, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet*. 2013;45:984–994.

Cuthbert BN, Insel TR. Toward new approaches to psychotic disorders: the NIMH Research Domain Criteria project. *Schizophr Bull*. 2010;36:1061–1062.

Dvorak D, Fenton AA. Toward a proper estimation of phase-amplitude coupling in neural oscillations. *J Neurosci Methods*. 2014;225:42–56.

Gandal MJ, Edgar JC, Klook K, Siegel SJ. Gamma synchrony: Towards a translational biomarker

for the treatment-resistant symptoms of schizophrenia. *Neuropharmacology*. 2012;62(3):1504–1518.

Insel TR. Rethinking schizophrenia. *Nature*. 2010;468:187–193.

Kelemen E, Fenton AA. Dynamic grouping of hippocampal neural activity during cognitive control of two spatial frames. *PLoS Biol*. 2010;8:e1000403.

Laughlin RB, Pines D, Schmalian J, Stojkovic BP, Wolynes P. The middle way. *Proc Natl Acad Sci U S A*. 2000;97:32–37.

Le Van Quyen M, Bragin A. Analysis of dynamic brain oscillations: methodological advances. *Trends Neurosci*. 2007;30:365–373.

Lee H, Dvorak D, Kao HY, et al. Early cognitive experience prevents adult deficits in a neurodevelopmental schizophrenia model. *Neuron*. 2012;75:714–724.

Lieberman JA, Stroup TS, McEvoy JP, et al. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med*. 2005;353:1209–1223.

Mitchell KJ, O'Donnell P, Durstewitz D, et al. A framework for the use of models in Schizophrenia. In: Silverstein SM, Moghaddam B, Wykes T, eds. *Schizophrenia: Evolution and Synthesis*. Cambridge, MA: MIT Press; 212–226.

Nunez PL, Srinivasan R. *Electric Fields of the Brain: the Neurophysics of EEG*. New York: Oxford University Press; 2006.

Phillips WA, Silverstein SM. Convergence of biological and psychological perspectives on cognitive coordination in schizophrenia. *Behav Brain Sci*. 2003;26:65–82; discussion 82–137.

Subramaniam K, Luks TL, Fisher M, et al. Computerized cognitive training restores neural activity within the reality monitoring network in schizophrenia. *Neuron*. 2012;73:842–853.

Weinberger DR. On the plausibility of "the neurodevelopmental hypothesis" of schizophrenia. *Neuropsychopharmacology*. 1996;14:1S–11S.

Ylinen A, Soltesz I, Bragin A, et al. Intracellular correlates of hippocampal theta rhythm in identified pyramidal cells, granule cells, and basket cells. *Hippocampus*. 1995;5:78–90.

▲ 1.27 Learning Theory and Psychosis, Anxiety, and Addiction

PHILIP R. CORLETT, PH.D.

Associative learning theorists concern themselves with how humans and other animals learn and use information about the relationships between events. Ivan Pavlov is perhaps the most famous learning theorist; his empirical studies of anticipatory learning about food rewards in dogs gave provided a framework to examine how relationships between environmental stimuli (lights, tones) and salient events (food, electric shocks) are acquired. In a curious mistranslation from his original Russian, this process is called conditioning and the cues conditioned are called stimuli (they ought to have been conditional, as in the salient events were conditional upon those cues).

Conditioning has obvious relevance to psychiatry (which is broadly considered as the study and treatment of mental illness, emotional disturbance, and abnormal behavior); given that conditioning guides behavior and emotions, abnormal conditioning may portend emotional disturbance, abnormal behavior, and thus mental illness. Anxiety and addiction have been studied most extensively from a learning theory perspective. More recently, psychosis has been likewise investigated in learning theory terms. Given that learning theory is inherently translational, studying mental illnesses from this perspective may portend a translational or reductionist understanding of mental illness. This chapter will review the impact of learning theory on the current understanding of psychosis, addiction, and anxiety. Furthermore, this chapter concludes by highlighting how computational neuroscience and, thereby, computational psychiatry are grounded in learning theory.

Pavlov himself appreciated the role of conditioning in mental illness. Toward the end of his career, he sought to explain hysteria and neuroses in conditioning terms. He attributed emotional perturbations to dysfunction the low level first signaling system that dealt with unconscious, preverbal associations. The verbal, second signaling system, Pavlov considered to be unique to humans and he asserted that dysfunction in the second signaling system led to psychosis. This distinction between simple associations and complex propositions continues to be a focus of contemporary learning theory.

A BRIEF HISTORY OF LEARNING THEORIES

Learning theory is sometimes referred to as associative learning theory, its currency being associations between mental concepts, although this is the subject of ongoing debate. Proponents of learning theory are sometimes

called associationists.

Associationism began with Plato, but Aristotle developed the concept and outlined the first associative learning theory by expounding the laws of association: **The law of contiguity**; things that occur close to each other in space or time are associated in mind. **The law of frequency**; the more often two things occur together, the more strongly they will be associated. If you think of Oliver Hardy, it is rather difficult not to think of Stan Laurel. **The law of similarity;** thoughts about one thing will trigger thoughts of similar things. **The law of contrast**; however, recollecting something will also trigger its complete opposite. When thinking of the tallest person you know, you may suddenly recall the shortest one as well.

Although John Locke is considered the first to discuss the role of improper association of ideas in mental illness, David Hartley believed that the principles of association were the fundamental operations of cognition and a means through which to relate psychology and physiology. In a neural doctrine inspired by Newton, Hartley argued that neural vibrations passing from the representation of one idea to another formed the physiological basis for their association (pre-empting the contemporary Hebbian notion of neurons firing together and wiring together). Hume added cause and effect (contiguity in time) as a law of association. British empiricist philosophers rejected Descartes notion of innate faculties. At birth, they thought of the mind as a tabula rasa—a blank slate—that accumulated ideas via sensory experience and built more complex ideas from simpler ones, through the laws of association.

Pavlov was perhaps the first experimental associationist. He explored the physiological and psychological mechanisms of association. He developed paradigms for presenting associations to experimental animals and tracking their learning, as well as theories for explaining how the cortex and subcortex underpinned that learning. Pavlov's work provided a framework for broad experimental exploration, some of which indicated that mere contiguity was not sufficient for learning, for example, Leon Kamin's blocking effect. When a novel cue (e.g., a tone) is paired with a stimulus (e.g., a light) that already predicts an outcome (electric shock), the pretrained cue—the light—"blocks" new learning about the novel cue—the tone; clearly a problem for contiguity-based theories of association.

Blocking requires that associative learning is sensitive to surprise—the degree of predictability of one associate to another modulates the formation of new associations. Hebb and Hartley's physiological explanations of learning are driven by contiguity and hence cannot explain blocking. It seems error correction—learning from mistakes—is crucial to the formation and updating of associations. For example, an algorithm developed to solve engineering problems using a simple connectionist neural network learned via error correction; the networks represented their inputs and outputs across distributed nodes. Associations or connections

were formed and strengthened between nodes. The weights of connections between particular nodes were updated using the mismatch between the desired output from a given input and the output that actually occurred. By reducing this error signal, the network's responses to particular inputs improved the network learned associations.

A similar algorithm was proposed to explain conditioning in rodents by Bob Rescorla and Allan Wagner. Here, environmental stimuli engender a set of expectations about potential subsequent states of the world by exciting representations of those states (e.g., the availability of food or the presence of a predator). Any mismatch, or prediction error, between those expectancies and actual experience is used as a teaching signal to update future expectancies by changing the associative strength. They explain Kamin blocking in terms of a lack of prediction error. The outcome is already completely predicted by the pretrained cue. When the novel cue is added, there is no prediction error, and hence no new learning.

Attention is also critical for associative learning. Associationists theorize that attention changes associability—the readiness with which an event or its representation (cue, outcome) can enter into associative relationships. Some theorists argue that cues that are predictably associated with important outcomes are more associable. However, stimuli with an uncertain predictive history also garner attention; hence, cues whose outcomes were surprising in the past are allocated more attention and learned about more readily. These theories are incompatible, but they are both supported by behavioral evidence. Attention must be important to association formation in different ways under different circumstances. Theories have reconciled these observations by positing a role for both confirmation and violation of expectancy in associative learning; the occurrence of an event that matches an active expectancy amplifies its representation in short-term memory, increasing the likelihood that it will be consolidated within long-term memory. By contrast, when an unexpected event violates the active expectancy, an orienting system is activated that resets short-term memory (dropping active expectancies) and engages an orienting response, permitting the acquisition of new explanatory associations. Using both of these mechanisms, organisms learn associations between stimuli, events, thoughts, and perceptions to build an internal model of their environment from which to make predictions.

The authors suggest that the principles of contiguity, similarity, and prediction error drive association formation and, like the British empiricists, more complex cognitive structures emerge by combining these simple representational units. There is some evidence for such an approach in human causal learning; subjects learn new causal models (propositional knowledge structures that they use to learn about causes) by minimizing prediction errors within and across possible causal models, that is, associative links can be used to form more complex causal representations

and the principles of associative learning (namely minimizing prediction error) may aid decisions about which structures to impose on an inference task. The authors note too that simple connectionist architectures in which Widrow–Hoff/Rescorla–Wagner like error correcting mechanisms are embedded often demonstrate rather complex emergent properties akin to reasoning and inference. In some of the earliest computational models of cognition, McCulloch and Pitts elaborated upon Gottfried Leibniz, the 17th-century philosopher who attempted to create an alphabet of human thought, each letter of which represented a concept and could be combined and manipulated according to a set of logical rules. Their building block was the proposition—the simplest possible statement, either true or false. From there, they employed the fundamental operations of logic, like the conjunction ("and"), disjunction ("or"), and negation ("not"), to link propositions into increasingly complicated networks. McCulloch–Pitts networks have also been combined with Widrow–Hoff error correction in a foreshadowing of propositional causal model theory.

Causal model theories also highlight an important new direction in cognitive science, the rise of Bayesianism. Thomas Bayes was a British mathematician whose doctrine of probabilities (published posthumously in 1873) has had a striking impact upon science in recent years. The theorem embodies a formal approach to reasoning about data using hypotheses and captures the probabilistic nature of many of the tasks faced by organisms; The authors believe a Bayesian approach to the mind and brain might bridge David Marr's levels of analysis and achieve a consilient understanding of cognition and its implementation in the brain. To wit, the problem faced by an organism is to predict its environment and respond appropriately; the organism's brain does this by minimizing its uncertainty about subsequent inputs by forming and maintaining a set of prior expectancies (predictive associations) about incoming stimulation.

Such prediction error minimization may even manifest in the signaling properties of single neurons that shuttle ion channels and receptors into and out of their membranes to modulate the cell's receptivity to incident stimuli. It is also embedded in the hierarchical neuroanatomy of the brain; in which stimuli enter through sense organs and their neural representations are communicated through the brain. Top-down expectations are communicated from areas with more abstract representations downward through the hierarchy. Any mismatch between prior expectation and current input (prediction error) needs to be dealt with. It is either cancelled by the top-down expectancy or propagated and used to update subsequent expectations with new learning. The neuroanatomy and neurochemistry of backward and forward connections across cortical layers are exquisitely suited to this approach.

The task of visual perception, conceived of as unconscious inference about current inputs based on a world model, embodies similar Bayesian

principles; indeed, the various visual illusions to which we are susceptible profoundly illustrate the overriding role of prior visual experience and expectancies on current perception. Pavlov believed that Helmholtz unconscious inferences were aligned with his conditioned responses; an assertion realized at the algorithmic level by the use of similar Bayesian formalisms to explain Kamin blocking, perceptual learning, and visual perception. Indeed, color-after effects in the visual system appear to be influenced to some degree by selective learning; they are subject to Kamin blocking, underlining the crucial role of learning across cognitive, perceptual, and neural systems.

Taken together, these theories conceive of perception, cognition, learning, action, and belief as processes of inference about latent causes that unite environmental stimuli and neural responses, cues and outcomes, choices, and rewards. These latent causes form the basis of the associations on which associative learning theories, by extension minds, are predicated. Disruptions to these processes of inference can lead to perturbed world models, percepts, beliefs, and actions. In the remainder of this chapter, learning theory is used to explain the symptoms of psychosis, addiction, and anxiety.

Psychosis

Psychotic symptoms involve a profound departure from consensual reality and manifest as perceptual experiences (hallucinations) and beliefs (delusions) that others in ones' culture do not share. One of the cardinal delusions, which occurs both after neurologic injury and as a symptom of psychiatric illness, is addressed. The Capgras delusion, originally termed the illusion of doubles, is the belief that familiar individuals (usually family members) have been replaced by imposters. While the individual confronting the sufferer looks very much like the familiar person, they are described as subtly different in appearance and crucially the sufferer reports lacking the feeling of emotional recognition that characterizes our experience of those familiar to us.

First, delusions are explained by using learning theory as a theoretical framework that bridges mind and brain, perception and belief. The unit through which this bridging takes place is the association. The following processes are central to this general explanation of delusions; first, the notion of expectation and that, through experience, one learns to expect a certain set of circumstances. Second, the consequences when these expectancies are violated that are either to ignore or discard the dissonant experience or to embrace and incorporate it into the world model, engendering an updated set of expectancies. That is, when confronted with someone whose features very much resemble a loved-one, one expects to feel a sense of familiarity. The absence of such a feeling induces a prediction error that demands explanation.

Absent but expected events are crucial in the phenomenology of psychosis. They contribute greatly to the feeling of delusional mood or atmosphere that attends the formative stages of delusions. In learning theory too, absent but expected events garner new learning, indeed, events that involve updating a causal belief about an absent but expected event engender a prediction error brain signal in the same circuit engaged when we learn beliefs about events that were actually experienced.

This error signal guides the Capgras sufferer toward a particular conclusion; that a replica has replaced their loved-one. The learning model challenges the strict distinction between perception and belief, and therefore the necessity for two factors to explain delusions.

The prediction error model provides little insight into the exact form that the new learning will take; the particular beliefs it will reinforce. The simplest associative learning models argue that the belief formed will track the evidence experienced; that the weight given to a particular belief will be proportional to the prediction error that it minimizes. Nearly all delusions are socially relevant: they are ideas about oneself and one's relationships with others. Their content is crucially related to the individual's specific fears, needs, or security.

Historically, associative learning research has also failed to consider mental representations of causality; that is, going beyond how correlations are represented to try to understand how people make causal attributions. For example, how do people learn about and represent foods causing allergies or drugs causing side effects? Such causal relationships can be complex, for example, in the case of foods causing allergic reactions, do two foods that cause allergies interact to cause a stronger allergy? As described above, individuals bring a rich set of expectancies regarding causal interactions to bear on the causal relationships with which they are confronted and, in the presence of competing candidate causes, they make all kinds of inferences about those candidates and how they might interact. There may be minimization of prediction error within a causal model (when learning about particular cues) as well as across causal models (when forming beliefs about causal mechanisms) as beliefs form.

One remarkable feature of delusional beliefs is their elasticity: they expand and morph to include new contradictory data. The person with Capgras, claiming their spouse is an imposter, might respond to other family members who greet the spouse warmly by saying "of course they hug her [the impostor spouse]—they're in on it!" This can be particularly difficult to understand as the patients can often learn about other new things (they don't have an all-encompassing deficit in learning). How can an associative learning, and in particular prediction error, account of delusion formation and maintenance account for elasticity? Given their explanatory utility, delusions are rehearsed extensively. When delusions are questioned, bringing them to mind may actually serve to reinforce

rather than to disrupt the memory.

In summary then, delusions form in response to aberrant prediction errors that imbue stimuli, thoughts, and percepts with inappropriate salience. Delusions formed to explain these experiences relieve the uncertainty associated with the prodromal "delusional mood." These beliefs are maintained in memory more strongly through cycles of recall, reactivation, and reconsolidation based strengthening. Confronted with contradictory evidence, people are more likely to discount the evidence than relinquish the belief (whether or not it is delusion).

Hallucinations

Hallucinations (percepts without external stimulus, often auditory and verbal in patients with schizophrenia—auditory verbal hallucinations, AVH) may also be explicable in terms of aberrant associative learning. Horace Barlow pointed out the similarities between Pavlovian conditioning and the unconscious inference processes that underpin perception.

Statistical regularities abound in the world around us, and many of them are actually, or potentially, important for human survival. Barlow suggested that perception too adapted to statistical regularities of the environment and that indeed all learning can be regarded as the internalization of environmental regularities, for it is driven by the statistically regular occurrence of reinforcement following particular sensory stimuli or self-initiated actions. Helmholtz argued that perception results from the interaction of apperception—the immediate impact of sensory messages—with remembered ideas resulting from past experience. He wrote of perceptions that "…by their peculiar nature they may be classed as conclusions, inductive conclusions unconsciously formed." Thus, he held the view that experience of the environment was internalized or remembered and provided the basis, together with current sensory messages, for the statistical conclusions—mostly valid that constitute the current perception of the world. These conclusions constitute symbolic working models based on the associative structure of objects and events in the environment, and are therefore expressions of environmental regularities. Can those models be misappropriated to give rise to perception without stimulus (Hallucination)?

This chapter highlights the "conditioned hallucination" research, published most extensively by Douglas Ellson in the 1940s. Here, subjects believed they were undergoing a test of auditory acuity: every time they saw a light bulb illuminate they would be presented with an auditory stimulus. After a series of pairings between tone and light, the light was presented. Subjects reported continuing to perceive the tone. Patients who hallucinate are more susceptible to this effect, suggesting a common mechanism between pathological and conditioned hallucinations. At first, it seems conditioning theories might have some difficulty explaining the

phenomenology of AVHs—for example, why are they voices rather than other sounds? Theorists like Chris Frith have proposed self-monitoring theories, whereby AVH arise as a result of misattribution of inner-speech to an external source. Can prediction error and conditioning be reconciled with this notion of source monitoring?

In perhaps the paradigmatic example of computational neuroscience, Wolfram Schultz has demonstrated in a series of studies that the prediction error signals that drive conditioning are generated in midbrain dopamine neurons. The signals are communicated to the striatum and prefrontal cortex via glutamate co-release where they have a role in ascribing salience to coincident stimuli, thoughts and percepts and determining agency for important events. Peter Redgrave argued that the dopamine signals in VTA do respond to expectancy violations; however, their role in the behavior is one of agency attribution—did I cause that surprising event to happen with my actions and is it something I want to happen again? This is achieved by discerning whether the VTA prediction error was coincident with forward-models of action, reverberating representations of intended actions in the striatum. When predictions and prediction errors are disrupted, agency for actions can be misattributed externally, leading to delusions of control (feeling that ones' actions are under the control of an external agent). Since speech is a motor act, it too may be generated and attributed via prediction error mechanisms and through similar mechanisms, inner speech may be externalized.

In terms of the Bayesian brain theory, slower neuromodulators like acetylcholine and dopamine gate the impact of a particular prediction error on new learning and future percepts. This precision weighting has much in common with associability in earlier learning theories, or attention in more common parlance. Hallucinations may arise top-down due to higher precision weighting of predictions such that they shape expectations, in a sort of listening attitude. Exactly how that expectation more commonly gives rise to voices than other sounds remains to be established. There are some intriguing findings from functional neuroimaging studies of patients who hear voices. First, in a predictive learning task, aberrant resting activity in auditory cortex and thus weakened responses to unexpected speech were observed in patients with auditory hallucinations. Similarly, the connectivity between Wernicke's area and Broca areas is disrupted during inner speech perception in patients who hear voices. Specifically, there is reduced input from temporal to frontal language areas in schizophrenia patients with AVH, suggesting that Broca activity may be less constrained by perceptual information received from the temporal cortex that could in turn prime increased top-down efforts from Broca area. Indeed, it has been suggested that reduced perceptual input triggers top-down influences in perception. With regard to speech processing, an fMRI study showed that only Broca area (BA44) was activated by unintelligible

speech presented at low signal-to-noise ratios, whereas an extended frontotemporal network (Wernicke's areas) was active during intelligible speech at high signal-to-noise ratios. The agency mechanism hypothesized by Redgrave and Gurney has yet to be examined in the context of speech studies in patients with schizophrenia; however, there is evidence for auditory cortex prediction error dysfunction and evidence (in healthy controls) that striatal prediction errors can train perceptual expectations in auditory cortex. Given the growing role that linguistic representations, like propositions, are having in associative learning, the links between voice hearing and conditioning bear further investigation.

Negative Symptoms

Under a conditioning account, the negative symptoms of schizophrenia (like anhedonia, amotivation, alogia) might occur when predictions are unreliable and sensory data are noisy. This would lead to a state in which decisions are difficult and ultimately actions are fruitless. One can easily imagine how such a situation would lead, for example, to the apathy and amotivation that are debilitating, strongly associated with poor functional outcome and particularly poorly treated with current pharmacotherapies. Furthermore, some current pharmacotherapies interact with the specification of reward predictions and prediction errors to exacerbate negative symptoms iatrogenically, which may explain some of the unpleasant side effects of these drugs like weight gain and sexual dysfunction.

Learning theory may further shed light upon the relationships between psychotic symptoms. Philosophers and clinicians have challenged the notion that delusions are even real beliefs, equating them with "empty speech acts" without manifest conviction. That is, patients with delusions espouse their beliefs verbally but do not act upon them in a manner consistent with such fantastical or terrifying claims; for example, a patient may claim that their food is being poisoned, but they will nevertheless eat it. This is termed double book-keeping. Of course, some patients do act on their delusions with disastrous consequences, both suicidal and homicidal. The cognitive neuropsychiatric account of delusions is committed to conceiving of delusions as a broken from of belief, but belief nevertheless. One answer is that even those without delusions fail to act consistently on their beliefs. Perhaps a more satisfactory explanation is that motivation plays a crucial mediating role, dictating whether one acts upon his or her beliefs or not. Given that patients who have delusions in the context of schizophrenia often have cooccurring negative symptoms including motivational deficits, double book-keeping may represent a confluence of delusions and negative symptoms.

ADDICTION

Laboratory animals will self-administer drugs and learn the cues and actions that predict drug access. Those cues drive subsequent drug seeking, even after periods of abstinence. Thus, associative models of addiction in animals have face and construct validity for human drug abuse. However, clinical research determining how human subjects learn about drug predicting stimuli is lacking. First, there are ethical considerations when administering drugs to human subjects. Furthermore, it is difficult to know exactly what might serve as a drug-predicting cue to a human addict, for example, the places in which drugs were procured or consumed, the sight and smell of drugs and drug paraphernalia, or even the friends with whom drugs have been taken. In one study smokers recorded the place, situation, and feeling experienced during smoking. Different contexts were associated with differing degrees of craving and enjoyment.

But how does this learning take place? Thus far, with a few notable exceptions clinical research into cue processing in addiction has focused on exposing drug users to generic cues associated with drug use such as stimuli from the International Affective Pictures System or dramatized video representations of cocaine consumption. However, conditioning histories are unique to each subject and hence, these standard pictures will vary in their relevance to subjects (e.g., a cocaine user may not snort their cocaine from a mirror). Nevertheless, stimuli related to the abused substance evoke neural activation in the striatum of smokers, alcoholics, and cocaine users, exactly as one would predict from the preclinical literature and from an aberrant learning and memory model of addiction (note however that one should be careful of making reverse inferences with functional imaging data; merely observing striatal activation in a subject does not mean that the subject was necessarily experiencing reward, motivation or engaging in learning).

PET scans with radiolabelled tracers sensitive to dopamine binding reveal dopamine release in the dorsal striatum in cocaine users in response to cocaine associated cues. Alcoholics too release dopamine in the striatum in response to alcohol cues, particularly in situations when alcohol was expected but not delivered. Such expectancy violations or prediction errors may provide a strong impetus to drug seeking and may drive new learning of Pavlovian cue-drug associations and the instrumental actions involved in drug procurement.

Learning to predict important environmental events and therefore behave adaptively toward natural rewards, aversive experiences, and other salient shifts in environmental contingencies are highly conserved across phylogeny. Drugs and drug-predicting stimuli appear to usurp the systems responsible for such adaptive anticipation, driving behavior toward anticipation and consumption of drug rewards and biasing memory consolidation toward associations between cues, drugs and the actions that procure them, perhaps by spuriously elevating reward prediction error

(RPE) signals in the midbrain, striatum, and frontal cortex. These aberrant errors may drive addiction related plasticity toward a particular behavioral repertoire, but once strongly consolidated this learning becomes insensitive to updating and is maintained despite large discrepancies between expected and experienced reinforcement.

During the earliest phases of drug use, individuals appear to show sensitized responses to all salient events but with repeated drug use, and hence conditioning, behavioral repertoires become more and more focused on the procurement and consumption of drug rewards. This conditioned attentional orienting response (CAOR) is elicited by initial sensory processing of the cues and engenders a shift in the focus of attention toward the spatial location of drug-predicting cues. It is speculated further that the CAOR gates other drug-associated conditioned responses and Pavlovian-to-instrumental-transfer (PIT)—the energizing effect of drug predicting cues on drug-seeking behavior, which supports and constitutes drug-seeking and drug-taking.

In a series of studies with smokers, Hogarth, Dickinson, and Duka have established the utility of eye tracking as an outcome measure in studies of predictive learning in addiction. They trained a set of novel arbitrary visual cues as discriminative stimuli, that is, subjects were trained that instrumental keypress responses in the presence of these stimuli were always rewarded with smoke delivery. In a subsequent test phase, they found that subjects demonstrated preferential attention to those stimuli relative to cues that did not predict smoke delivery. There are competing theories of the role of attentional allocation in conditioning and suggestions from Hogarth and colleagues that attentional biases may differ depending on the stage of drug use and abuse.

Folk psychology conceives of drug use as a bad habit, something in which addicts indulge despite an expressed desire not to. Preclinical behavioral neuroscience supports this conception of addiction. Ascertaining whether a behavior is goal directed or habitual involves a simple yet elegant procedure—if behavior is governed by the incentive value of a reward then, devaluing that reward (by, e.g., pre-feeding to satiation or associating reward consumption with gastric illness) will decrease the vigor with which an organism engages in reward-seeking behaviors. However, if behavior is habitual, it is no longer governed by a representation of reward value. Instead it involves reflexive responses to environmental cues previously associated with drug use, hence devaluing the drug reward has no effect on behavioral choice or vigor and the organism continues to engage in reward seeking. Although much of this work has been conducted in rodents, recent reports have confirmed similar behavioral and neural processes are involved in human habit learning. That is, with overtraining, devaluing a particular food will not prevent human subjects from engaging in the instrumental action that procures that food,

and such behavior is associated with activation in the dorsal caudate and putamen, the regions aligned with habitual responding in laboratory animals.

To date, no studies have examined the neural mechanisms of associative learning where drugs are the reinforcers. However, studies using monetary outcomes associated with previously neutral cues have revealed neural differences between addicted individuals and nonsubstance abusing controls. These studies indicate a more global dysfunction in learning and memory suggestive of rapid habit learning and an insensitivity of actions to their consequences. For example, in a financial investment game, smoker's brains compute a fictive error signal (indicating what they would have won had they chosen an alternative option). However, they do not use that signal to guide their choices in the same way nonsmokers do. Likewise, alcoholics learning action choices generate striatal prediction errors, but those errors appear to not be communicated to the frontal cortex effectively such that they are not used to update action choices. Interestingly, in this study, the degree of frontostriatal dysconnectivity predicted current alcohol craving in the problem-drinking group. Of course, these effects with monetary rewards may indicate a relative lack of salience of monetary reinforcers when compared with drugs, but, nevertheless, they implicate associative learning dysfunction in the pathology of addiction. Furthermore, these studies suggest an insensitivity to the consequences of actions in addicts that resonates with the theoretical conception of habitual responding.

Some learning-based treatments for addiction have emerged. Psychotherapies based upon cue reactivity involve repeatedly exposing a client to the sights, smells, or paraphernalia associated with a substance in the absence of actual consumption, thus attenuating the power of those cues to elicit an expectation of drug effects. In other words, they no longer readily predict a 'high' is coming so the cue–drug associations are extinguished. Such therapies are often effective in the therapeutic setting, but they can fail to transfer to the real world outside of the clinic, that is, they are context dependent.

ANXIETY

Watson and Rayner (1920) originally argued that phobias are associative fears that develop when a neutral stimulus is paired with a traumatic event. They conditioned a phobia in their subject, Little Albert (a 9-month-old boy) by pairing a loud and frightening gong with the presence of a rat. It is said that the boy became afraid of white furry animals subsequently and that fear even generalized to Santa Claus—whose beard was white and fluffy.

One criticism of conditioning theories though is the observation that many people with phobias do not appear to recall a specific conditioning

episode. It is possible that simply observing others experiencing a trauma or behaving fearfully could be sufficient for some phobias to develop. Such vicarious conditioning of a fear of snakes fear has been observed in primates. Laboratory-reared monkeys without initial fear of snakes watched unrelated monkeys reacting fearfully to snakes. The observers rapidly acquired an intense fear of snakes. This vicarious conditioning also occurred via watching videotapes of model monkeys behaving fearfully. This has interesting implications for human susceptibility to vicarious conditioning from watching movies and TV.

However, many individuals who actually have traumatic experiences do not develop phobias. There may be a genetic vulnerability rendering some individuals more susceptible than others. There are also relevant differences in life experiences that may occur before, during, or following a fear-conditioning experience, and they can affect how much fear is experienced, acquired, or maintained over time.

Latent inhibition is a conditioning phenomenon in which prior exposure to a conditioned stimulus before the it is ever paired with an unconditioned stimulus retards subsequent conditioning. Children who have had more previous nontraumatic encounters with a dentist are less likely to develop dental anxiety if subsequently traumatized at the dentist's office.

Instrumental learning (learning to make responses, rather than passively experience contingencies) is also relevant to fear conditioning and anxiety. As discussed so far, Pavlovian conditioning elicits fear in response to situational cues. That fear response then mediates avoidance responses. The avoidance response is maintained, according to two-factor theory, by escape from the conditioned stimulus and by an immediate reduction in its fear-eliciting properties. These ideas aided the development of the behavior therapy, particularly exposure and response prevention approaches.

In addressing pathophysiological and potentially therapeutic processes, what happens after fear learning should also be considered. Extinction refers to the well-demonstrated effect that if a conditioned predictor that engenders a response is presented alone (without the salient outcome), then response strength will decline over successive trials until the stimulus no longer elicits a response. Exposure to the cue without the traumatic outcome should allow the anxious response to extinguish, which is the basis for exposure-based treatments for anxiety.

Despite this success, conditioning as an explanation of how phobias develop has come under fire: basic conditioning theory predicts that fear should decrease over successive nonreinforced presentations of the cue; however, the opposite is often true. When someone with an established spider phobia subsequently comes into contact with spiders without traumatic consequences, fear should extinguish. Instead spider fear increases, it incubates. Incubation may proceed through memory

reconsolidation. Reconsolidation occurs when memories are recalled into a labile state, integrated with new information, and consolidated once more. This process depends on prediction error and evokes competing processes: However, there is extinction initiated by a negative prediction error signal and engendering competing learning that overrides the original memory. However, the positive prediction error response to predictive cues leads to the reminder-based strengthening of the belief (much like delusions—see above). The precise balance between these processes is driven by the magnitude, sign, and timing of the accompanying prediction error signal. Incubation may occur because of this reminder-based reconsolidation, that is, fear memories strengthen because they are reactivated, ruminated upon, reconsolidated, and laid down more strongly.

Although exposure and other interventions based on conditioning have already proved successful, it is clear that therapeutic interventions are relatively simplistic and do not incorporate, for example, the possibility for incubation. Of course, therapists already try to discern cues (e.g., in exposure therapy), but realizing that such cues might also trigger incubation, and understanding the role of reconsolidation in fear memory maintenance might add to and improve the current therapeutic arsenal. For example, although reconsolidation might contribute to fear memory maintenance, it can be blocked pharmacologically and psychologically and might therefore be employed therapeutically.

SYNTHESIS AND FUTURE DIRECTIONS

Associative learning lends itself to the burgeoning field of computational psychiatry in which patients' behavioral and brain responses can be compared to and combined with algorithmic predictions of optimal performance in order to characterize which processes are related to (or are perhaps responsible for) their psychopathology. If the underlying psychological and neural mechanism of those processes can be characterized, and we can ameliorate those deficits through behavioral pharmacological means can be ameliorated, it stands to reason that computational modeling (of behavioral and brain responses) in individual psychiatric patients could perhaps be used to divert patients toward particular treatment approaches that address their specific deficits. Thus far, goal-directed and habitual behavior has largely been modeled in terms of model-based and model-free reinforcement learning, respectively. These computational approaches capture the initial flexibility and outcome sensitivity (model-based/goal-directed) and its transition to inflexible stimulus driven responding (model-free/habitual). However, many do not appreciate that, in order for these models to work, they necessitate an assumption that agents have complete knowledge of the task and environmental contingencies. This is not realistic. Rather than assuming that subjects will learn associations (e.g., between stimuli-actions and

outcomes), they instead learn beliefs or causal hypotheses—that the same causal factors lead to the stimuli and outcomes, and that actions influence those factors. Hence, PIT occurs because of the agent's assumptions about common causes between the factors generating the relationships between Pavlovian cues and outcomes and the factors influencing whether (and which) actions lead to outcomes. Similar latent cause models have been applied to Pavlovian conditioning and attentional allocation. Furthermore, these latent cause models can discern goal-directed and habitual behavior. These models can be applied to the behavioral and neural data generated in current studies to discern the representational structures and processes that are guiding learning in clinical disorders.

REFERENCES

Aristotle, Ross WD, ed. *On Memory and Reminiscence. The Works of Aristotle*. Vol. 3. Oxford: Clarendon Press; 350 B.C./1930.

Barlow H. Conditions for versatile learning, Helmholtz's unconscious inference, and the task of perception. *Vision Res*. 1990;30(11):1561–1571.

Bortolotti L, Broome M. Affective dimensions of the phenomenon of double Bookkeeping in delusions. *Emot Rev*. 2012;4(2):187–191.

Corlett PR, Aitken MR, Dickinson A, et al. Prediction error during retrospective revaluation of causal associations in humans: fMRI evidence in favor of an associative model of learning. *Neuron*. 2004;44(5):877–888.

Corlett PR, Fletcher PC. Computational psychiatry: a Rosetta Stone linking the brain to mental illness. *Lancet Psychiatry*. 2014;1(5):399–402.

Corlett PR, Taylor JR, Wang XJ, Fletcher PC, Krystal JH. Toward a neurobiology of delusions. *Prog Neurobiol*. 2010;92(3):345–369.

Ellson DG. Hallucinations produced by sensory conditioning. *J Exp Psychol*. American Psychological Association; 1941:1.

Everitt BJ, Parkinson JA, Olmstead MC, Arroyo M, Robledo P, Robbins TW. Associative processes in addiction and reward. The role of amygdala-ventral striatal subsystems. *Ann N Y Acad Sci*. 1999;877:412–438.

Frith C. The neural basis of hallucinations and delusions. *C R Biol*. 2005;328(2):169–175.

Gray JA, Feldon J, Rawlins JN, Hemsley D, Smith AD. The neuropsychology of schizophrenia. *Behav Brain Sci*. 1991;14:1–84.

Grossberg S. Processing of expected and unexpected events during conditioning and attention: a psychophysiological theory. *Psychol Rev*. 1982;89(5):529–572.

Hartley D. *Observations on Man, His Frame, His Duty, and His Expectations*. New York: Delmar; 1749/1976.

Haselgrove M, Hogarth L. *Clinical Applications of Learning Theory*. 1st ed. East Sussex: Psychology Press; 2011:252.

Helmholtz H von. The facts of perception. In: Kahl R, ed. *Selected Writings of Herman von Helmholtz*. Wesleyan University Press; 1878/1971.

Hogarth L, Dickinson A, Duka T. The associative basis of cue-elicited drug taking in humans.

Psychopharmacology (Berl). 2010;208(3):337–351.

Hui S, Zak SH. The Widrow-Hoff algorithm for McCulloch-Pitts type neurons. *IEEE Trans Neural Netw*. 1994;5(6):924–929.

Hume D. *A Treatise of Human Nature*. Oxford: Oxford University Press; 1739/2007.

Kamin L. Predictability, surprise, attention, and conditioning. In: Campbell BA, Church RM, eds. *Punishment and Aversive Behavior*. New York: Appleton-Century-Crofts; 1969.

Mackintosh NJ. A theory of attention: variations in the associability of stimuli with reinforcement. *Psychol Rev*. 1975;82:276–298.

Mineka S, Zinbarg R. A contemporary learning theory perspective on the etiology of anxiety disorders: it's not what you thought it was. *Am Psychol*. 2006;61(1):10–26.

Mitchell CJ, De Houwer J, Lovibond PF. The propositional nature of human associative learning. *Behav Brain Sci*. 2009;32(2):183–198; discussion 198–246.

Pavlov IP. *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. 1927.

Pearce JM, Hall G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev*. 1980;87(6):532–552.

Redgrave P, Gurney K. The short-latency dopamine signal: a role in discovering novel actions? *Nat Rev Neurosci*. 2006;7(12):967–975.

Reed GF. *The Psychology of Anomalous Experience: A Cognitive Approach*. London: Hutchinson; 1972.

Rescorla RA, Wagner AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In: Black AH, Prokasy WF, eds. *Classical Conditioning II: Current Research and Theory*. New York: Appleton-Century-Crofts; 1972:64–99.

Sass LA. *Paradoxes of Delusion: Wittgenstein, Schreber, and the Schizophrenic Mind*. Ithaca: Cornell University Press; 1994

Shanks DR. The associative nature of human associative learning. *Behav Brain Sci*. 2009;32:226–227.

Waldmann MR, Martignon L. A Bayesian network model of causal learning. In: *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Earlbaum; 1998.

Widrow B, Hoff ME Jr. Adaptive switching circuits. In: *IRE WESCON Convention Rec*. 1960;96–104.

▲ 1.28 Functional Brain Connectivity and Psychopathology

DEANNA M. BARCH, PH.D.

Since the advent of modern "Biological Psychiatry" in the 1950s, much work on understanding the etiology of mental illness has focused on identifying the neural mechanisms that contribute to the development and maintenance of varying forms of psychopathology. This work has been aided by advances in novel neuroimaging technologies, such as the use of

SPECT, PET, and MRI for the assessment of either or both structural and functional characteristics of the human brain. The focus of this chapter will be on the use of measures of human brain functional connectivity and the ways in which understanding brain connectivity can inform the general understanding of the mechanisms that lead to the development and maintenance of various forms of psychopathology. The chapter begins with a brief history of the development of methods to measure and understand human functional brain connectivity. Next, the chapters provides a brief overview of the currently available methods for studying functional brain connectivity, with some discussion of their relative pros and cons. An overview of the current knowledge about core "intrinsic" functional brain connectivity networks in humans is also provided. Lastly, the chapter selectively describes some of the work using functional brain connectivity to understand the neurobiology of mental illness, with a specific focus on psychosis and depression, as these are two facets of psychopathology with perhaps the largest existing literatures on functional connectivity impairments.

In terms of brain function, historically, much of the functional neuroimaging work in psychopathology has focused on the use of either SPECT/PET or fMRI. These approaches provide indirect estimates of neural activity through the measurement of either glucose or oxygen metabolism (PET), blood flow (SPECT), or the measurement of blood oxygen level dependent (BOLD) activity. The initial studies addressing questions about the neural mechanisms of psychopathology primarily used task-based PET or fMRI to assess differential activation of brain regions while participants performed various cognitive, effective, or social tasks. These studies have been essential for the development and testing of hypotheses about which brain regions may be affected in different manifestations of mental illness, under the inference that abnormal BOLD activity during performance of cognitive or effective tasks indicates abnormal underlying neuronal processes. As some examples, this research has led to the development of theories about the role of the DLPFC in cognitive dysfunction in psychosis, the role of the amygdala in depression, the role of the amygdala and insula in at least some anxiety disorders, and the role of striatal mechanisms in addictive disorders.

Both the basic and translational research on task-related brain activation has been hugely informative in terms of the current understanding of the neural substrates of particular cognitive and effective processes and how these may go awry in psychopathology. However, over time, it has become clear that rarely does any particular task activate only a single brain region, and rarely is any particular domain of psychopathology associated with abnormal activation in only a single brain region. Further, basic neuroscience research has long made it clear that activity in any individual brain region (or any individual neuron!) is the result of inputs

from and outputs to different areas of the brain. Such realizations led to a shift in focus to questions regarding neural circuits rather than a focus only on specific brain regions. More specifically, this shift has been to questions about the relationship *between* and *among* different brain regions in producing successful cognitive and effective function in health, and the ways in which abnormalities at the level of circuits contribute to the development and maintenance of particular dimensions of psychopathology. Questions regarding the role of the DLPFC in cognitive dysfunction in schizophrenia provide an excellent case study in this regard. Much of the early work on DLPFC dysfunction associated with cognitive impairment in psychosis focused almost exclusively on the DLPFC itself. However, what if, for example, the abnormal activity observed in the DLPFC during a cognitive control task was actually the result of faulty inputs from another brain region(s)? In other words, what if the DLPFC itself functioned normally in schizophrenia, but the information it received was leading to the observed abnormalities? Alternatively, what if the problem was not in the activity within any specific brain region, but rather in the connections among, or the ability to coordinate activity among, brain regions? The growing work on the role of brain oscillations in coordinating activity among and between neutral networks is consistent with such hypotheses that localize impairment in psychopathology at the circuit level of function rather than within specific individual brain regions. Further, in many ways, it is easier to connect circuit level hypotheses with the larger literature on impairments at other levels of analysis in psychopathology (e.g., neurotransmitter availability, receptors, E-I balance). Such impairments are frequently not regionally specific and involve processes and mechanisms that would reasonably be expected to alter inter-regional communication, either with cortex or between cortical and subcortical regions.

THE ADVENT OF STUDYING FUNCTIONAL CONNECTIVITY

The ability to examine these types of questions has been propelled by the development of functional connectivity methods for use in humans. Originally, the concept of functional connectivity was applied to simultaneous recordings of neuronal spike trains, which are thought to contribute to the functional connectivity observed in human using noninvasive neuroimaging methods. A main inference of functional connectivity is that, if two regions have highly correlated neuronal activity (i.e., have high functional connectivity), then they are more likely to engaged in a common set of processing mechanisms. If so, then functional connectivity provides a tool for understanding which brain regions may be communicating during the completion of cognitive or effective demands, and therefore what brain circuits support performance and ability in different domains of cognition, emotion, and/or social processing.

RESTING STATE FUNCTIONAL CONNECTIVITY

A major breakthrough in the development and application of functional connectivity methods for humans came in 1995, when Biswal and colleagues reported that spontaneous activity from regions in the right and left motor cortices were highly correlated even while an individual was resting, suggesting functional connectivity between these brain regions even when the individual was not performing a task. Interestingly, this correlated activity was observed in the spontaneous low-frequency fluctuations of the BOLD signal (0.01 to 0.10 Hz), a frequency band that has often been discarded as noise in task-based studies, though such correlations can be seen at other frequencies as well. This work spurred a major field of exploration of what is now referred to as resting state functional connectivity. Importantly, this resting brain state activity is thought to consume a major portion of the body's energy (~20 percent), despite the brain only being 2 percent of the body's total mass. Further, changes in metabolism due to engagement in a specific task are typically less than 5 percent, which suggests that ongoing resting state activity may provide a richer a source, if not richer, of disease-related variability than changes observed due to task performance. In addition, there is also robust work demonstrating that a large portion of the trial-to-trial variability in task-related activity reflects coherent and organized spontaneous fluctuations in brain activity, providing another piece of evidence that this is meaningful source of variation in human brain function.

RESTING STATE FUNCTIONAL CONNECTIVITY NETWORKS

One of the key aspects of resting state functional connectivity that has spurred interest in this aspect of brain function is the realization that *resting state functional connectivity* reveals intrinsically organized networks of brain regions that are consistently functionally connected, even in the absence of task-induced perturbations in ongoing brain activity This has been supported by numerous "network" mapping studies that have identified consistent, robust, and reproducible networks of brain regions that show coordinated activity at rest. These include the default mode network, the frontal–parietal network, the cingulo–opercular network, and the dorsal attention network, to name a few. One recent example of this is shown in Figure 1.28–1, where 356 regions ("parcels") identified as having highly homogenous patterns of resting state functional connectivity were grouped into 12 resting state networks that map nicely to networks identified in previous resting state studies. Importantly, many of these networks map closely or at least partially to networks that have been identified in task-based studies, providing validation to the functional meaningfulness of this coordinated activity.

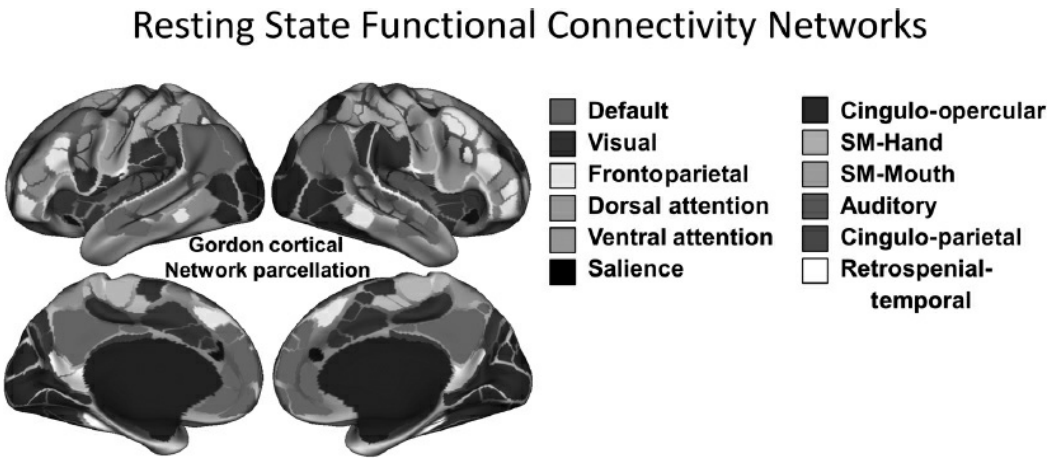

FIGURE 1.28–1. Functional connectivity networks derived from resting state data. The top row is a lateral view of the brain and the bottom row is a medial view of the brain. The colored "parcels" were derived from a boundary identification approach using a "watershed by flooding" algorithm to identify boundaries in the gradients of change in whole brain maps of functional connectivity. The community structure of the parcels derived from this boundary mapping was generated using an InfoMap algorithm. (See color atlas.) (From Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb Cortex*. 2016;26(1):288–303, with permission.)

Such functional connectivity networks have their basis in part in structural connectivity, but are not isomorphic with regions that show direct structural connectivity. As such, the mapping of resting state networks to known task networks has led to the hypothesis that resting state networks reflect in part organized interactions that arise from a history of coactivation over development. Such networks are now frequently referred to as "intrinsic" connectivity networks, given that they are not dependent on performance of a particular task for identification, are present even at rest, and are relatively consistent across task and environmental states. A recent major project to study the human functional and structural connectome across multiple scales and modalities—the Human Connectome Project—has corroborated the presence of these core networks in the human brain. Moreover, the results of this project have confirmed and extended current knowledge about the critical relationship between individual variation in function brain connectivity and individual variation in a range of behaviors relevant to mental health and both negative and positive life outcomes.

Although in principle any human brain network is relevant to the understanding of psychopathology, several of the most commonly identified networks are particularly relevant to the functions and processes often found to be impaired in psychopathology. For example, both the *frontal–parietal* and the *cingulo–opercular* networks are ones that have been consistently associated with a variety of cognitive control functions. The frontal–parietal network includes dorsal regions of both the lateral prefrontal cortex and parietal cortex, while the cingulo–opercular network

includes the dACC, bilateral dorsal anterior insula, and sometimes both thalamic regions and anterior prefrontal regions. As will be discussed in more detail below, these networks have been found to be abnormal in populations with known impairments in cognitive control, such as schizophrenia, or in emotion regulation, such as major depression. The *dorsal* and *ventral attention* networks have also been associated with cognitive function, though more so top-down and/or stimulus driven attention functions. The dorsal attention network is closely aligned with the frontal–parietal network in that it not only includes both dorsal frontal and parietal regions but also more dorsal supplementary motor and eye field areas. The ventral attention system includes the temporal–parietal junction and the VLPFC, and has been associated with attention to salient events in the environment, often activated when events in the word that disrupt ongoing processing occur. Altered connectivity has been associated with anxiety disorders, under the hypotheses that it could contribute to altered attention to threat relevant stimuli in anxiety disorders.

The *default mode network* has been associated with a variety of functions, but in particular may be associated with attention to internal emotional states and the ability to distinguish or shift between internal and external modes of attention. Further, there is large body of literature showing that the default mode network decreases activity during engaged task states, and some hints that a failure to "shut down" the default mode network may impair cognitive function. The default mode network (see Fig. 1.28–1) includes a wide variety of regions, but most notably medial prefrontal, medial posterior cingulate, and precuneus regions. Increased connectivity among default mode network regions has been associated with rumination in major depression. The *salience* network is one that has been identified somewhat more recently than some of the other networks, and includes more rostral regions of the anterior cingulate and insula than typically associated with the cingulo–opercular network (see Fig. 1.28–1), though both have strong connectivity with limbic and subcortical regions. Connectivity of the salience network has been associated with anxiety and arousal and is hypothesized to be a network of regions that serve to process and coordinate reactions to salient events in the environment.

MODALITIES FOR ASSESSING RESTING STATE FUNCTIONAL CONNECTIVITY

Resting state functional connectivity in humans can be assessed using a variety of methods. By far the most common is to use fMRI with a focus on spontaneous fluctuations in BOLD-related activity. This approach is noninvasive and provides good-to-excellent spatial resolution, depending on the pulse sequences used, but it is inherently limited by the fact that BOLD-related signals are indirect measures of neural activity (i.e., reflecting a compensatory blood flow response in reaction to neural activity that alters local magnetic homogeneity) and by the fact that BOLD signals

are slow and temporally extended over time. Resting state functional connectivity can also be measured using EEG or MEG, which can provide much finer grained temporal resolution and the ability to resolve connectivity at different frequencies, though these methods can be limited in their ability to localize signals (i.e., EEG) or to address deep sources of activity in regions such as the striatum or thalamus (i.e., MEG). Functional connectivity can also be assessed using more invasive techniques, such as electrocorticography (ECoG), which involves placement of grids of electrodes subdurally. Such an approach has the obvious advantage of direct mapping to electrical activity generated by neuronal activity along with good spatial localization, but is limited to use in populations in need of intracranial recordings for medical reasons.

Further, although much of the focus on functional connectivity has been on data acquired while participants are resting quietly with their eyes closed or open, numerous studies have now demonstrated that very similar results can be derived from data acquired during the performance of specific tasks, though with some caveats. One such approach is to use the "fixation" blocks that occur in between blocks of task performance in task-based activation studies that use what is referred to as a "blocked" design. If the study design involves a sufficient number of such fixation blocks, they can be concatenated and used to assess intrinsic connectivity networks, with relatively similar results to more traditional resting state studies. An alternative approach is to use data acquired during the performance of a task, but to try to remove the deterministic influence of the task structure on BOLD activity through regression methods, and then to base the functional connectivity analyses on the residuals from the regression. Results from such studies can be very similar to results from traditional resting state studies. However, the degree to which this is true may depend on the number of different task states over which the data is integrated. Doing this for only a single task state is essentially functionally equivalent to the data generating using psychophysiological interaction analysis approaches (PPI), which focus on comparing connectivity differences in one task state versus another task state. Further, there is some evidence in the literature that prior "states" or performance of specific types of tasks can influence subsequent assessments of resting state connectivity. As such, integrating over several tasks states maybe more effective at eliminating deterministic or residual effects of any one task state, and thus increase generalizability and similar to resting state connectivity.

METHODS FOR ANALYZING RESTING STATE FUNCTIONAL CONNECTIVITY

As is true for most neuroimaging methods, the use of resting state functional connectivity data introduces many choice-points for the researcher, primarily in terms of how functional connectivity should be

calculated, how to identify different brain regions, and how noise should be handled. Each of these choices will influence the structure and quality of the data, but there is no clear consensus about a "best" method for all resting state studies, only trends of commonly used approaches in the literature. A review of all possible variations or choices relevant to these issues is beyond the scope of this chapter and has been reviewed elsewhere. The focus of this chapter now turns to some of the choices most relevant to understand the current literature on psychopathology. For most studies of psychopathology, the main goal of resting state functional connectivity analyses is to assess variations in the relationships between different brain regions, either due to an interest in a specific, a priori region and its connectivity, or to better understand disruptions in functional networks. As such, one dimension (see Fig. 1.28–2) along which to characteristics analysis approaches is from relatively more a priori and hypothesis driven to relatively more data driven. Along this dimension, the three most common analysis techniques are region of interest (ROI) or a priori network-based analyses, seed/ROI to whole brain connectivity analysis, and independent component analysis (ICA). All three of these methods have pros and cons and it is up to the researcher to decide which analysis will allow them to most accurately test their research question at hand.

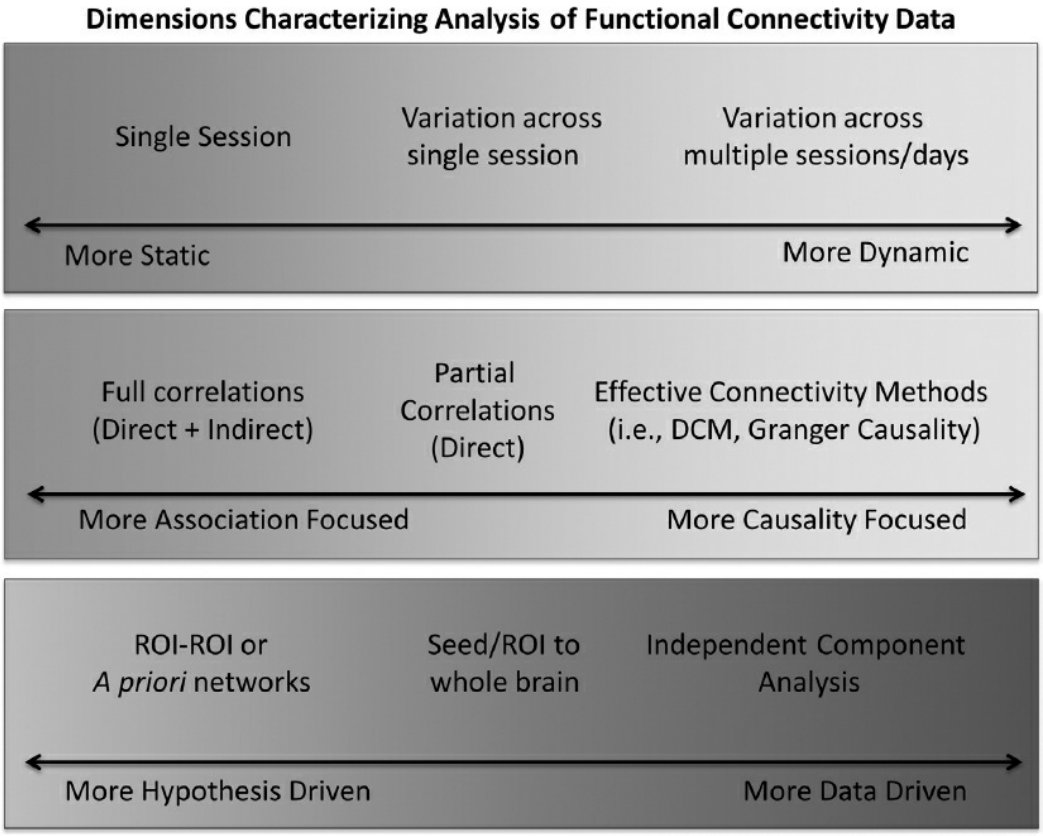

FIGURE 1.28–2. Dimensions characterizing functional connectivity analysis. The bottom row illustrates the dimension from more hypotheses driven to more data-driven approaches to the analysis of functional connectivity data. The middle row illustrates the dimension of analysis approaches more focused on characterizing association versus those

attempting to determine causality or directional relationships. The top row illustrates the dimension of analysis approaches focused on static estimates of connectivity in a single sessions, versus approaches more focused on the dynamics of change within a session or across sessions.

ROI-ROI or Network-Based Connectivity

ROI to ROI or network-based analyses typically take as their starting point a specific set of regions or a specific network of regions already identified and characterized in the literature, and focus on extracting the time series from these regions and analyzing the connections among those specific ROIs or among the regions in that specific network. As an example, one might wish to test the hypothesis that a particular dimension of psychopathology is associated with abnormal connectivity between two specific regions (e.g., medial prefrontal and precuneus regions from the default mode network) or with abnormal connectivity among all or some of the regions associated with the default mode network. For example, prior work has examined whether connectivity among or between regions in the frontal–parietal network, the cingulo–opercular network, or the default-mode network is altered in schizophrenia, and whether it relates to variation in cognitive function. As another example, Bishop and colleagues examined whether connectivity between the amygdala and the specific parts of the anterior cingulate related to anxious effect.

ROI/Seed-Based Connectivity

Although there are examples of specific ROI–ROI or network-focused analysis in the literature, perhaps the more common approach is to start with a particular ROI or "seed" and examine connectivity between the seed and the rest of the brain, or at least a subset of the brain. Seed-based connectivity represents the most common method employed to test hypotheses regarding functional connectivity abnormalities in schizophrenia and their relationship with cognition. Seed-based connectivity is a flexible method that can be performed, if so desired, but is more commonly used to understand the functional connectivity map for specific regions region of interest. There are numerous examples of this in the literature, such as work looking at connectivity of the amygdala, or subregions of the amygdala, in relationship to depression. Seed regions can also be used to identify functional networks, assuming that the seed region is a robust member of a specific network. However, just because many regions are strongly connected with, say, the DLPFC does not necessarily indicate that those regions form a unitary network. In the case of seed-derived functional networks, it is important to verify that most other nodes are functionally connected to one another as well as to the seed. Taken together, one of the biggest advantages of ROI-ROI, network, or seed-based connectivity methods is that they allow for a direct tests of a specific

research questions, making interpretation of findings in such studies relatively straightforward, modulo all of the potential interpretive caveats that arise in any functional connectivity study.

Importantly, however, this advantage to ROI or seed-focused studies may also contribute to some of the drawbacks of such studies. ROI/seed-based studies allow researchers to make specific predictions about relationships between certain regions or networks of the brain and to test those questions directly. However, by preselecting an anatomical region as the seed for all other correlations, researchers can introduce a bias in their data. There is no perfect way to provide the most accurate and reliable time series for the functional area of interest in all subjects, even when using anatomically based approaches (e.g., individual subject segmentations of specific areas). Consider the possibility that functional areas from two different networks could sit side by side in the brain, such that the inclusion of several voxels from the wrong region could result in connectivity maps that reflect two different functional networks. Add on top of that the consideration that at least some clinical populations may also have anatomical abnormalities that could make identification of regions challenging. Further, seed or ROI-focused studies may also suffer from the "streetlamp" problem, such that a researcher is biased to find significant results in a priori regions simply because those are the ones examined, and the researcher does not know whether even stronger or more consistent results might have been generated through the use of a different set of regions.

Another limitation of seed-based connectivity approaches is that they provide a univariate analysis of what is in reality a multivariate system. In such ROI or seed-based connectivity analyses, the time series of one or more regions is correlated with the time series of one or more other regions or voxels as an effort to understand pairwise relationships. What are being ignored in this process are the complex interactions that lead to the observed pairwise relationships, such as modulations of the connectivity of that pair of regions or voxels by other connections. Therefore, seed-based connectivity analysis not only restricts the anatomical definition of functional networks, as described above, but interpretations of these functional networks do not take into account the full complexity of the system in which the network is embedded.

Independent Components Analysis

The most common alternative to ROI or seed-based connectivity analysis is ICA, which is a multivariate data-driven approach to identifying functional networks of the brain. ICA is an algorithm that takes the entire set of BOLD data and decomposes it into spatially segregated, maximally statistically independent components. In doing so, ICA addresses several of the limitations associated with ROI/seed-based connectivity that were

discussed above. Namely, ICA does not require the definition of a priori regions, resolving the potential bias of selecting nonoptimal voxel clusters or of looking only at certain connections. ICA also takes into account the complexity of multivariate data, since the components identified by ICA are determined using the entire BOLD data set, which includes information about the time series from all voxels.

Not surprisingly, ICA also introduces several limitations that, similar to seed-based connectivity, have the potential to bias the data. Most notably, ICA identifies components of the data, but it is up to the researcher to determine what each component represents. For example, researcher must decide which components are noise and which components represent functional networks, though strategies exist for helping investigators to accomplish this, such as automated identification of previously validated noise components. Finally, interpretation of ICA data is not always as straightforward as interpretation of seed-based results. Specifically, the metric of ICA is the magnitude of the component, not the magnitude of specific pairwise relationship. The magnitude of a component can be reduced for a number of different reasons, including variations in spatial distribution, reduced variability in signal or reduced correlations among time series across regions, and it is not always clear which of these is driving the results.

Association versus Directionality

Another dimension (see Fig. 1.28–2) on which the analysis of resting state data can vary is the degree to which the methods focus on simply characterizing association versus attempting to identify causal or directional associations. By far the most common approach to the analysis of resting state data is to simply use full correlations or ICA, which do not presume any directionality to the associations and which capture covariance due both to direct connections between regions and to share variance with other regions. At the opposite end of the spectrum are approaches such as dynamic causal modeling and granger causality, which attempt to establish temporal or causative directionality from the patterns of covariance, though there are challenges to the use of this method with resting state fMRI data. Somewhat in between these two extremes is the use of partial correlations, which attempt to isolate the direct associations of connectivity between regions or voxels, partialling out the covariance due to shared connections with other regions.

Static versus Dynamic Assessments

Yet another dimension along which analysis of resting state data can vary is a focus on a static estimate of a pattern of function connectivity in an individual at one point in time versus dynamic variations in connectivity

"states" over time. A number of studies have now begun to examine how patterns of connectivity among brain regions may vary over time and be classified into a limited set of connectivity "states" or patterns of connectivity. These state patterns may also vary across clinical populations, either in terms of the number of states achieved, the amount of time spent in different states, or the speed with which transitions between states occur.

Network Approaches to Functional Connectivity

As a complement to seed/ROI-based and data-driven approaches, researchers are increasingly applying tools from network science understand brain organization, primarily through the use of graph theory. When adopting a graph theory approach, investigators have examined variations in several graph properties including, global and local efficiency, small-worldness, and participation coefficients. Global efficiency measures a brain network's functional integration, by indexing how easy it is for information to be communicated across brain regions in a network (e.g., across regions in the default mode network). Local efficiency measures how well a brain region's neighbors continue to communicate with each other following removal of that region ("node" in graph theory terminology) and therefore represents something like the fault tolerance of a network. Participation coefficients indicate the number of different networks with which a particular brain region communicates, providing information about how important that region is in overall brain communication. Regions with high participant coefficients are often referred to as "hubs" and are thought to be particularly important for information transfer. The hubs that have been identified in the human brain include the DLPFC, the precuneus, and the insula among others, and those that are particularly highly connected to each other are referred to as the "rich club." Graph theory approaches have been used to examine brain organization in a number of forms of psychopathology, but perhaps most frequently in schizophrenia, as will be discussed in more detail below.

BENEFITS AND CHALLENGES IN APPLYING RESTING STATE FUNCTIONAL CONNECTIVITY ANALYSES TO THE STUDY OF PSYCHOPATHOLOGY

Studies of resting state functional connectivity have not only shaped the way that human brain activity and organizations is understood but also have had important implications for research in clinical populations. One potential advantage of examining resting state data over task-based data in individuals with psychiatric impairments is that may not be as complicated by task performance as task-based imaging data, which has proved to be a difficult confound in studies comparing healthy individuals and various patient populations. Primarily, a concern of task-based studies is often

whether the differences in observed BOLD activity between controls and patients might be due to differences in ability to complete the task, rather than due to inherent problems using those brain regions, since modulation of task difficulty often influences BOLD activity within subjects. This has led many researchers into a challenging situation of designing tasks that are difficult enough for healthy controls to yield variability in BOLD signal and performance, but easy enough for patients to learn and perform reliably. In addition, this task performance confound has shaped the recruitment strategies of researchers who tend to focus on higher functioning patients who can tolerate a cognitive or effective task within the scanner. Because of this, the majority of task-based fMRI studies in schizophrenia may be biased by the use of populations of higher functioning patients and therefore may not always generalize to more symptomatic individuals. Therefore, resting state functional connectivity potentially represents an opportunity to assess neurobiological abnormalities in a wider range of individuals with schizophrenia.

However, it is also important to keep in mind that resting state studies come with their own particular set of challenges when using such data to understand the neural mechanisms of psychopathology. The first and foremost challenge is the pernicious confound introduced by participant movement, which is often greater among clinical populations. Movement can introduce all sorts of artifacts and confounds into resting state data, and if not properly addressed during processing, can lead to spurious conclusions about the nature and source of connectivity changes in clinical populations. For example, one of the effects of movement is to increase the magnitude of connections among regions that are close together, but to decrease the strength of connections among more distant regions, a confound referred to as distance-dependent movement artifact. The existence of such artifacts challenges the interpretation of studies claiming to find global reductions in long range connections coupled with increased local connectivity in clinical populations, when movement confounds have not been sufficiently addressed in the data.

Insights from Resting State Functional Connectivity Studies in Psychopathology

To date, there are now 100s and 100s of resting state functional connectivity studies that have been conducted in various populations with different aspects of psychopathology. As such, it is well beyond the scope of this chapter to review all of these studies. However, to illustrate some of the insights gained from using resting state functional connectivity to study psychopathology, this chapter will selectively review studies focused from the two domains with arguably the largest resting state literatures—schizophrenia and major depression. The specific focus will be to identify areas where there is starting to be a convergence of results that have replicated across studies and which point to clear pathways or directions

for future research.

Functional Connectivity Abnormalities in Schizophrenia Spectrum Disorders

To date there are over 200+ empirical studies examining various aspects of functional connectivity in schizophrenia spectrum disorders, with the vast major conducted in individuals with *Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition* (DSM-IV) diagnosed schizophrenia. These studies have used a range of methods and approaches, and thus have generated quite a diversity of findings. However, there are areas with increasingly consistent results that point to clear and compelling aspects of altered functional connectivity in individuals with schizophrenia.

Altered Thalamic Connectivity. A strong focus on functional connectivity in schizophrenia is consistent with long-standing hypotheses about schizophrenia as a "dysconnection" syndrome, where impairments in cognition and behavior occur because of a failure of coordinated action across multiple brain regions. As many researchers have noted versions of this hypothesis were put forth as early as the work of Wernicke and Bleuler. Further, theories about abnormalities in connections among brain regions have also played a central role in more modern theories of the pathophysiology of schizophrenia. One such prominent theory, put forth by Andreasen and colleagues, suggested that schizophrenia involves a disruption in the integration of cortical-subcortical-cerebellar circuits a hypothesis termed "cognitive dysmetria." Alterations in the structure and function of the thalamus are prominent in the schizophrenia literature.

As shown in anatomical studies of primates, the thalamus is topographically organized into parallel pathways connecting specific thalamic nuclei to different regions of cortex brain regions, helping to form parallel loops for various types of information processing between subcortical and cortical regions. In particular, the medial dorsal and anterior nuclei of the thalamus project to the DLPFC. A large and consistent body of evidence has now shown connectivity from bilateral thalamic regions, medial dorsal and anterior nuclei in particular, to bilateral DLPFC, dACC, parts of the striatum, and bilateral cerebellum. In a number of studies, these patterns of decreased connectivity have also been coupled with increased connectivity between the thalamus. These patterns of altered connectivity are illustrated in Figure 1.28–3, which shows regions showing reduced thalamic connectivity in schizophrenia in blue, and those showing increased connectivity in red. Like the published literature, individuals with schizophrenia in this sample show reduced thalamic connectivity with bilateral prefrontal and parietal cortices, the ACC, and bilateral striatum and cerebellum. In contrast, they show increased connectivity bilaterally with motor, somatosensory and visual cortices.

Similar patterns of altered connectivity are seen in individual at both clinical and genetic high risk for schizophrenia, and even to some extent in other disorders characterized by psychosis, such as bipolar disorder. Further, among individuals at clinical high risk, the patterns of altered thalamic connectivity are most pronounced among those who subsequently converted to a full blown psychosis diagnosis. In addition, studies of structural connectivity are also starting to support altered thalamic (especially medial dorsal) to prefrontal connectivity including reduced fractional anisotropy of the anterior limb of the internal capsule. Importantly, reduced connectivity between the thalamus and the DLPFC predicts cognitive impairment. These robust and consistent findings across studies of altered thalamic connectivity are fully consistent with the idea that coordination among these brain regions is disrupted in individuals with schizophrenia spectrum disorders. Further, the relationship of this impaired thalamic connectivity with cognitive function is in line with the idea that thalamocortical networks are critical for organizing brain oscillations important for effective cognitive function, and that disruptions in such circuits due to impairments at other level of analysis (e.g., E-I balance) may significantly impair large-scale brain function.

Pattern of Altered Thalamic Connectivity in Schizophrenia

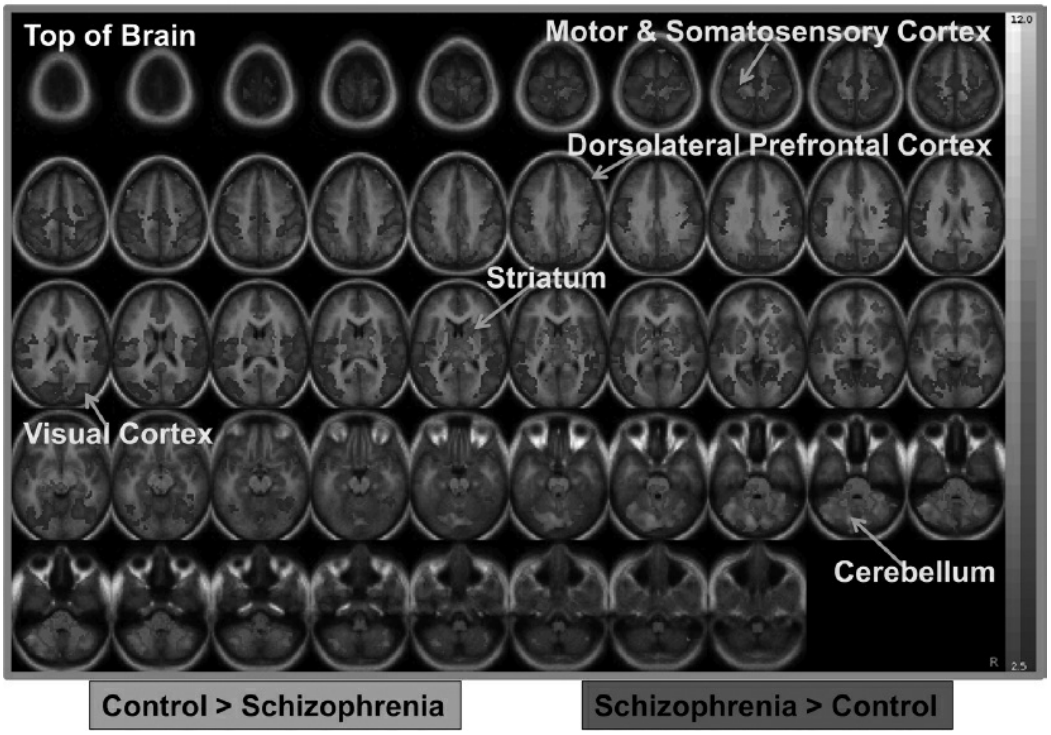

FIGURE 1.28–3. Patterns of altered thalamic connectivity in schizophrenia. Unpublished data showing regions (in *red*) where healthy individuals show greater thalamic connectivity than individuals with schizophrenia, and regions (in *blue*) where individuals with schizophrenia show greater thalamic connectivity than healthy individuals. The regions showing decreased thalamic connectivity in schizophrenia includes cerebellum, striatum, dorsolateral prefrontal cortex, and the dorsal anterior cingulate. The regions showing increased thalamic connectivity in individuals with schizophrenia include motor and

somatosensory cortices, visual cortex, and anterior temporal poles. (See color atlas.)

A variety of aspects to the results regarding thalamic connectivity suggest that it needs to be given major weight as a potential biomarker for the risk of psychosis. The first is the fact that such results have been found even among individuals who have never taken medications. The second is the fact that it is present across the course of the disorder and even in those who share risk of psychosis via either clinical symptom manifestation or genetic risk. As such, a next important step will be to relate thalamic connectivity to other early indicators or risk markers for the development of psychosis. Such markers would include the evidence for increased presynaptic dopamine availability in the striatum, as well as altered glutamate levels in the thalamus.

Altered Default Mode Connectivity. Many researchers have been interested in the idea that individuals with schizophrenia might have alterations in connectivity of the default mode network. As described above, one of the major functions ascribed to the default mode network is attention to internal emotional states. A number of researchers have hypothesized that this attention to internal states must be dampened when individuals need to engage in specific tasks or attend to the external world, and that this leads to the reduced activity of the default mode network typically seen when individuals engage in a focused cognitive task. Thus, the interest in the default mode network from the perspective of schizophrenia is the idea that some of the altered perceptual experiences that individuals with schizophrenia might have, or even some of the delusional material, might reflect *enhanced* attention to internal thoughts, feelings, and experiences, due to altered connectivity as well as activity of the default mode network. Further, it is has been hypothesized that a failure to suppress default mode network activity during cognitive task performance may be one of the mechanisms, leading to disrupted cognitive function in schizophrenia.

Although such hypotheses are intriguing, the empirical literature on abnormalities in default mode connectivity in schizophrenia is actually quite mixed, as recently reviewed by Narr and Leaver. An early small-scale study found evidence for hyperconnectivity and hyperactivity of the default mode network in schizophrenia, which spurred interest in this line of functional connectivity work. However, the results of subsequent studies have been quite varied, with some studies also finding hyperconnectivity, some finding hypoconnectivity and some finding no differences in default mode network connectivity. One of the challenges in this literature may be that many of the studies have simply focused on comparing individuals with schizophrenia, regardless of symptom profile, to healthy controls. Given the fact that symptom profiles and severity can vary widely across

individuals with the same diagnosis and across time in the same individuals, it may be critical to also examine default mode connectivity in relationship to variations in the severity of specific symptoms, such as hallucinations, delusions, and cognitive impairment. For example, some work links auditory hallucinations to altered connectivity of the temporal–parietal junction and cortical midline structures that are part of the default mode network (anterior and posterior cingulate, precuneus). However, the direction of the relationship is such that the presence of auditory verbal hallucinations tends to be associated with reduced connectivity among these default mode structures. This is perhaps not the direction one would have predicted under the hypotheses that auditory hallucinations reflect abnormal attention to internal states due to hyperconnectivity or hyperactivity of default mode regions.

Graph Theory Analyses of Functional Brain Connectivity in Schizophrenia. Earlier this chapter described the application of graph theory approaches to understanding functional brain connectivity, as a means to summarize major properties of brain organization. Such approaches have been increasingly applied to study functional (and structural) brain connectivity in schizophrenia. Although the findings have been somewhat mixed and complex, an emerging theme is that individuals with schizophrenia show reductions in both global and local efficiency, particularly when examining specific networks. The networks most frequently associated with reduced global and/or local efficiency tend to involve the anterior cingulate and/or insula (i.e., the cingulo–opercular network) or frontal regions (i.e., the frontal–parietal network). Such findings are consistent with a large body of other work using ROI-based or ICA-based methods that have also identified reduced functional connectivity in or between frontal–parietal and cingulo–opercular networks in schizophrenia. Interestingly, reductions in both the local and global connectivity of the frontal–parietal and cingulo–opercular networks have also been associated with impaired cognition in schizophrenia, suggesting functional relevance of these network alterations.

There is also intriguing evidence that individuals with schizophrenia show impaired function and connectivity of “hub” regions in the brain, particular those that are thought to be part of the “rich club” of highly interconnected hubs. More specifically, individuals with schizophrenia show significant reductions in the connections among regions characterized as hubs, including the insula, the superior frontal cortex, the precuneus, and the superior parietal lobule. However, they do not show reductions in the connections between these hub brain regions and nonhub regions (referred to as “feeder” connections), or in the connections among nonhub regions. Importantly, for some regions, such as the anterior insula, individual differences in the degree of hubness, as indexed by participation

coefficients, also predict the level of cognitive impairment in schizophrenia. As a piece of converging evidence, numerous studies have also shown reduced gray matter volume in at least some of these hub regions, including the anterior insula.

Summary of Schizophrenia Functional Connectivity Literature. As noted above, there is now a very large literature on functional connectivity in schizophrenia, and some of the findings are mixed, including the results regarding default mode connectivity. However, several findings stand out as consistently replicable. By far the most and robust set of findings is of reduced thalamic connectivity with frontal, cingulate, striatal, and cerebellar regions, a result that is strongly consistent with long-standing theories about schizophrenia as a disconnection syndrome. In addition, there is a growing literature suggesting reduced global and local efficiency in schizophrenia, particularly in frontal–parietal and cingulo–opercular networks, as well as evidence for reduced connectivity of hub regions in this illness, both of which predict cognitive function. It is highly likely that these two sets of findings are related, as the large-scale thalamic disconnectivity seen in schizophrenia may contribute to reductions in the efficiency with which networks can process and transfer information and may also interfere with the ability of hub regions to communicate information across networks. In many cases, these impairments are present at the start of illness, or even before the start, and are more prominent among those individuals who go on to develop full blown psychosis. Thus, these findings converge to suggest that abnormalities in functional connectivity may be a key component of the pathophysiology of psychosis, and a promising avenue for future biomarker and intervention work.

Functional Connectivity Abnormalities in Major Depression

Like the literature on schizophrenia, there are now a burgeoning number of empirical studies examining various aspects of functional connectivity in major depression, using a variety of different analytic approaches. Again, like the schizophrenia literature, there is diversity in the findings, which may relate to methodology, stage of illness, and medication factors. However, there are several emerging trends that provide important clues about the ways in which altered functional brain connectivity maybe contributing to the development and/or maintenance of mood disorders.

Amygdala Connectivity and Emotion Regulation. The amygdala is a subcortical structure that has received massive research attention because of its role in fear conditioning, threat sensitivity, and general responding to salient events in the environment. In addition, numerous studies have shown that poverty, stress, and trauma influence the development,

function, and structure of the amygdala. Further, a number of investigators have suggested that such alterations in amygdala structure/function are a mediating mechanism, linking early adversity to later negative mental health outcomes. Importantly, these findings in humans are consistent with a large animal literature also showing detrimental effects of stress and maternal care on amygdala function and structure, as well as behavior.

The amygdala shows a normative pattern of both positive and negative connections with other regions of the brain. Figure 1.28–4 illustrates the typical pattern of amygdala connectivity found in both children and adults, using a sample of school-age children and adolescents. As can be seen, the amygdala shows positive resting state functional connectivity with other regions in the medial and anterior temporal lobes, including the hippocampus, as well with as the ventral medial prefrontal cortex (at age 10+ for this region). These regions showing positive connectivity with the amygdala are also important for emotion and stress responsivity. In some studies, the amygdala has been found to couple with other regions in the salience network, including ventral regions of the anterior cingulate (ventral medial prefrontal cortex, as in Fig. 1.28–4) and ventral regions of the insula and anterior temporal poles. In contrast, as shown in Figure 1.28–4, the amygdala also shows negative functional connectivity—anticorrelations—with bilateral dorsal prefrontal and parietal regions, as well as the dorsal anterior cingulate. These regions showing negative correlations with the hippocampus and amygdala are the same regions that are activated in studies of emotion regulation and overlap in large part with regions that are part of the frontal–parietal and cingulo–opercular networks. Importantly, these "anti-correlations" are thought to indicate top-down regulation of emotion and stress responsivity occurring in amygdala regions, suggesting a critical role for the integrity of these connections between amygdala and frontal–parietal and cingulo–opercular networks in mood and effective function.

Importantly, several studies have shown that early life stress and/or maternal deprivation are associated with altered connectivity between amygdala and ventral medial prefrontal cortex and pregenual anterior cingulate, as well as other regions, connections that are thought to be critical for effective emotion regulation. Further, patterns of altered amygdala connectivity with components of both the salience network and the frontal–parietal and cingulo–opercular networks have also been found in individuals with MDD. The results have been somewhat variable across studies, but one of the more consistent findings has been reduced connectivity between the amygdala and various regions of the prefrontal cortex, including dorsolateral prefrontal, rostral prefrontal, and ventral prefrontal regions, as well as part of the anterior cingulate. As noted above, one hypothesis is that such reduced amygdala to prefrontal/cingulate connectivity reflects impaired top-down regulation of emotion processing

and that such reduced connectivity contributes to emotion regulation difficulties in this form of psychopathology. Interestingly, these patterns of altered connectivity have been found across the life span, including among individuals with depression who are children, adolescents, young and middle aged adults, and older adults, suggesting that altered amygdala to prefrontal/cingulate connectivity may be a life span characteristic of depressive disorders. Further, such connectivity alterations are present even in those who are at risk of the development of depression by virtue of having a mother with depression, suggesting that such connectivity disruptions precede the onset of manifest illness, and might even be part of the pathophysiological pathway.

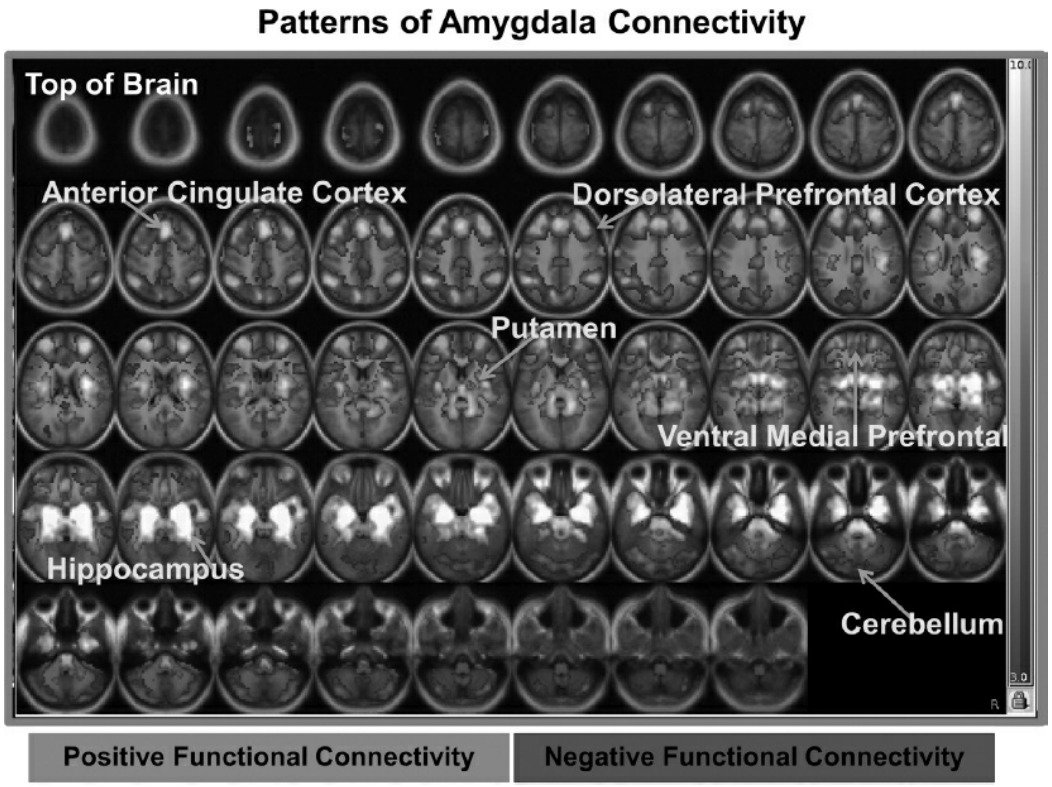

FIGURE 1.28–4. Patterns of amygdala connectivity. Unpublished data showing regions (in *light gray*) demonstrating positive functional connectivity with the amygdala, and regions (in *dark gray*) showing negative functional connectivity or "anti-correlations" with the amygdala. The regions showing positive functional connectivity with the amygdala include the hippocampus, the ventral medial prefrontal cortex, and the anterior temporal poles. The regions showing negative functional connectivity with the amygdala include the dorsolateral prefrontal cortex.

Default Mode Network Connectivity and Rumination. The second area of consistent findings of altered connectivity associated with major depression is in the domain of default mode network connectivity. As described above, one of the major functions ascribed to the default mode network is to modulate attention to internal emotional states and to govern self-referential processing. Phenomenologically, one of the major

characteristics of depression is a preoccupation with negative mood, negative self-evaluation, and frequently a ruminative style of coping with negative events and thoughts. It is not surprising, then, that research points to an *increase* in connectivity among default mode regions in major depression, especially those forming the more anterior part of the default mode network, including the dorsomedial prefrontal cortex. Interestingly, work has also shown an increase in connectivity between the dorsomedial prefrontal cortex and the rostral anterior insula, part of the salience network, among individuals with major depression. This increased connectivity between parts of the default mode network and the salience network has been interpreted as potentially reflecting an overly strong integration or dominance of default mode network activity over activity related to attention to salient events, potentially contributing to a negative evaluation bias in regard to salient information in the surrounding environment.

A number of studies have also found evidence for increased connectivity of the subgenual prefrontal cortex with the default mode network (see Fig. 1.28–5). The subgenual prefrontal cortex is a brain region that has received consistent attention in the literature on mood disorders, as studies have demonstrated that it shows glial cell reduction, altered metabolism, and reduced volume and/or thinning among individuals with MDD, again across the life span. It has been hypothesized that the subgenual prefrontal cortex is important for mechanisms of behavioral withdrawal and negative effective appraisal, with activation of this region associated with conserving resources (i.e., withdrawal), promoting safety oriented behaviors, and experiencing guilt. Importantly, research has shown that the degree of hyperconnectivity between the subgenual prefrontal cortex and the default mode network predicts the level of rumination among individuals with depression. This work has led to the hypothesis that this overly strong integration serves to link the self-referential processing supported by the default mode network with the negative affectivity and withdrawal supported by the subgenual prefrontal cortex, leading to a cycle of rumination and impaired emotion regulation among individuals with MDD.

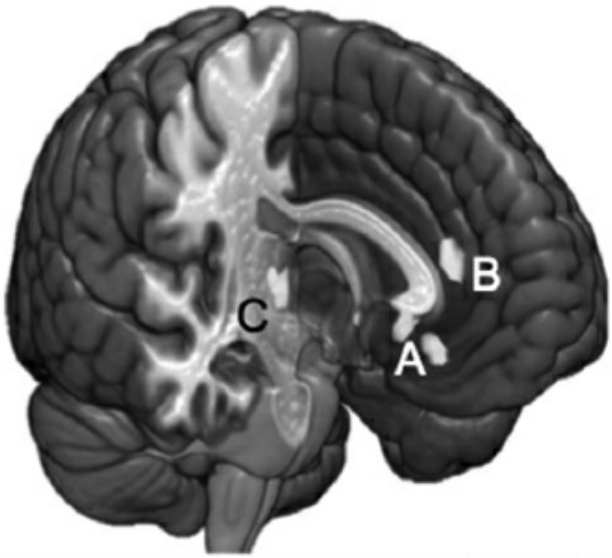

Region	Voxels	X	Y	Z	Figure
Subgenual prefrontal cortex (bimodal cluster)	1281	0	26	-10	A
Left dorsal anterior cingulate	976	-6	37	12	B
Right medial dorsal thalamus	891	9	-18	5	C
Right posterior lateral parietal cortex	734	42	-67	29	

FIGURE 1.28–5. Default mode network connectivity in depression. The figure is an image depicting regions showing increased connectivity with the default mode network in individuals with major depression. This includes regions of the subgenual prefrontal cortex that is not always found to be a part of the default mode network in studies with nondepressed individuals. (From Hamilton JP, Farmer M, Fogelman P, Gotlib IH. Depressive rumination, the default-mode network, and the dark matter of clinical neuroscience. *Biol Psychiatry*. 2015;78(4):224–230, with permission.)

Summary of Functional Connectivity Literature in Major Depression. As with the literature on functional connectivity in schizophrenia, the literature on connectivity in depression is large and somewhat mixed. However, several findings again stand out as consistently replicable. The first is evidence for altered connectivity between the amygdala and prefrontal/cingulate regions that are part of the frontal–parietal and cingulo–opercular networks. These connections are thought to be crucial for top-down emotion regulation and control over regions such as the amygdala that are important for responding to stress and salient environmental events. These connectivity alterations are present in individual experiencing various forms of early stress or trauma, are present across the life span in individuals with major depression, and are present even among those at risk of depression. As such, that may represent an important clue as to the etiological pathway linking stress and trauma to subsequent development of emotion dysregulation and mood pathology.

The second area of relatively consistent results is evidence for increased connectivity of the default mode network, including links between the subgenual prefrontal cortex and other regions of this network. This increased connectivity has been specifically linked to rumination among individuals with major depression, and may reflect an overly tight integration between brain regions supporting self-referential processing and those influenced in negative effect processing and withdrawal.

CONCLUSION

The goal of this chapter was to provide a brief overview of the current state of the field in terms of methods for assessing the human functional connectome, the existing knowledge about the core intrinsic functional connectivity networks present in the human brain, and the ways in which connectivity abnormalities in these networks are relevant for understanding the pathophysiology of major mental illnesses. Although in many ways this field is still in its infancy, the remarkable consistency and robustness of the identified functional brain networks across imaging modalities, functional connectivity methods, and populations attest to the importance and relevance of understanding their impact on human behavior. Work from the Human Connectome Project has confirmed that individual variation in intrinsic network connectivity relates to a wide range of both positive and negative traits and characteristics in humans. Further, the work selectively reviewed in this chapter illustrates that there are robust, replicable, and differentiable alterations in functional connectivity associated with differing forms of psychopathology. This includes robust evidence for altered thalamic connectivity associated with psychosis, impairments in global and local efficiency of frontal–parietal and cingulo–opercular networks in relation to cognitive function in psychosis, and alterations in the hub structure of regions such as the anterior insula in schizophrenia. In major depression, there is consistent evidence for altered amygdala to prefrontal–cingulate connectivity relevant for understanding emotion regulation, as well as evidence for increased default mode connectivity in relation to rumination. The next key steps will be to examine the role that these connectivity alterations play in the development of psychosis and depression, or specific symptom dimensions, as well as the degree to which connectivity can be modified by treatment and whether it serves as a mediator of symptom improvement. Further, it will be crucial to link alterations at this neural circuit level of analysis with alterations at other levels of analysis, such as genes, gene expression, neurotransmitter signaling, and/or cellular architecture.

REFERENCES

Alderson-Day B, McCarthy-Jones S, Fernyhough C. Hearing voices in the resting brain: a review of intrinsic functional connectivity research on auditory verbal hallucinations. *Neurosci*

Biobehav Rev. 2015;55:78–87.

Andreasen NC, Paradiso S, O'Leary DS. Cognitive dysmetria as an integrative theory of schizophrenia: a dysfunction in cortical-subcortical-cerebellar circuitry? *Schizophr Bull*. 1998;24(2):203–218.

Anticevic A, Cole MW, Repovs G, et al. Characterizing thalamo-cortical disturbances in schizophrenia and bipolar illness. *Cerebral Cortex*. 2014;24(12):3116–3130.

*Anticevic A, Haut K, Murray JD, et al. Association of thalamic dysconnectivity and conversion to psychosis in youth and young adults at elevated clinical risk. *JAMA Psychiatry*. 2015;72(9):882–891.

Bijsterbosch J, Smith S, Forster S, John OP, Bishop SJ. Resting state correlates of subdimensions of anxious affect. *J Cogn Neurosci*. 2014;26(4):914–926.

Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med*. 1995;34(4):537–541.

Canu E, Agosta F, Filippi M. A selective review of structural connectivity abnormalities of schizophrenic patients at different stages of the disease. *Schizophr Res*. 2015;161(1):19–28.

Cole MW, Bassett DS, Power JD, Braver TS, Petersen SE. Intrinsic and task-evoked network architectures of the human brain. *Neuron*. 2014;83(1):238–251.

Dauvermann MR, Whalley HC, Romaniuk L, et al. The application of nonlinear Dynamic Causal Modelling for fMRI in subjects at high genetic risk of schizophrenia. *Neuroimage*. 2013;73:16–29.

Dosenbach NU, Fair DA, Miezin FM, et al. Distinct brain networks for adaptive and stable task control in humans. *Proc Nat Acad Sci*. 2007;104(26):11073–11078.

Fair DA, Schlaggar BL, Cohen AL, et al. A method for using blocked and event-related fMRI data to study "resting state" functional connectivity. *Neuroimage*. (2007);35(1):396–405.

Fornito A, Zalesky A, Pantelis C, Bullmore ET. Schizophrenia, neuroimaging and connectomics. *Neuroimage*. 2012;62(4):2296–2314.

Fox MD, Raichle ME. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*. 2007;8(9):700–711.

Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ. Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*. 1997;6(3):218–229.

Gonzalez-Burgos G, Lewis DA. NMDA receptor hypofunction, parvalbumin-positive neurons, and cortical gamma oscillations in schizophrenia. *Schizophr Bull*. 2012;38(5):950–957.

Gordon EM, Breeden AL, Bean SE, Vaidya CJ. Working memory-related changes in functional connectivity persist beyond task disengagement. *Hum Brain Mapp*. 2014;35(3):1004–1017.

Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb Cortex*. 2016;26(1):288–303.

Haijma SV, Van Haren N, Cahn W, Koolschijn PC, Hulshoff Pol HE, Kahn RS. Brain volumes in schizophrenia: a meta-analysis in over 18 000 subjects. *Schizophr Bull*. 2013;39(5):1129–1138.

Hamilton JP, Chen MC, Gotlib IH. Neural systems approaches to understanding major depressive disorder: an intrinsic functional organization perspective. *Neurobiol Dis*. 2013;52:4–11.

*Hamilton JP, Farmer M, Fogelman P, Gotlib IH. Depressive Rumination, the Default-Mode

Network, and the Dark Matter of Clinical Neuroscience. *Biol Psychiatry*. 2015;78(4):224–230.

Howes O, McCutcheon R, Stone J. Glutamate and dopamine in schizophrenia: an update for the 21st century. *J Psychopharmacol*. 2015;29(2):97–115.

Ma S, Calhoun VD, Phlypo R, Adali T. Dynamic changes of spatial functional network connectivity in healthy individuals and schizophrenia patients using independent vector analysis. *Neuroimage*. 2014;90:196–206.

Menon V, Uddin LQ. Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct*. 2010;214(5-6):655–667.

Minzenberg MJ, Laird AR, Thelen S, Carter CS, Glahn DC. Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Arch Gen Psychiatry*. 2009;66(8):811–822.

*Mulders PC, van Eijndhoven PF, Schene AH, Beckmann CF, Tendolkar I. Resting-state functional connectivity in major depressive disorder: a review. *Neurosci Biobehav Rev*. 2015;56:330–344.

*Narr KL, Leaver AM. Connectome and schizophrenia. *Curr Opin Psychiatry*. 2015;28(3):229–235.

Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*. 2014;84:320–341.

Repovs G, Barch DM. Working memory related brain network connectivity in individuals with schizophrenia and their siblings. *Front Hum Neurosci*. 2012;6:137.

Repovs G, Csernansky JG, Barch DM. Brain network connectivity in individuals with schizophrenia and their siblings. *Biol Psychiatry*. 2011;69(10):967–973.

Sheffield JM, Repovs G, Harms MP, et al. Fronto-parietal and cingulo-opercular network integrity and cognition in health and schizophrenia. *Neuropsychologia*. 2015;73:82–93.

Smith SM, Miller KL, Salimi-Khorshidi G, et al. Network modelling methods for FMRI. *Neuroimage*. 2011;54(2):875–891.

*Smith SM, Nichols TE, Vidaurre D, et al. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat Neurosci*. 2015;18(11):1565–1567.

Stephan KE, Friston KJ, Frith CD. Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. *Schizophr Bull*. 2009;35(3):509–527.

Sylvester CM, Corbetta M, Raichle ME, et al. Functional network dysfunction in anxiety and anxiety disorders. *Trends Neurosci*. 2012;35(9):527–535.

Tang Y, Kong L, Wu F, et al. Decreased functional connectivity between the amygdala and the left ventral prefrontal cortex in treatment-naive patients with major depressive disorder: a resting-state functional magnetic resonance imaging study. *Psychol Med*. 2013;43(9):1921–1927.

Tottenham N. The importance of early experiences for neuro-affective development. *Curr Top Behav Neurosci*. 2014;16:109–129.

Uhlhaas PJ, Singer W. Oscillations and neuronal dynamics in schizophrenia: the search for basic symptoms and translational opportunities. *Biol Psychiatry*. 2015;77(12):1001–1009.

van den Heuvel M, Sporns PO, Collin G, et al. Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA Psychiatry*. 2013;70(8):783–792.

Woodward ND, Heckers S. Mapping thalamocortical functional connectivity in chronic and early

stages of psychotic disorders. *Biol Psychiatry*. 2016;79(12):1016–1025.
Yeo BT, Krienen FM, Sepulcre J, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*. 2011;106(3):1125–1165.

▲ 1.29 Computational Modeling Approaches to Psychiatry

JOHN D. MURRAY, PH.D., AND ALAN ANTICEVIC, PH.D.

INTRODUCTION

A primary goal of psychiatry is to understand the neural bases and mechanisms underlying mental disorders. Mechanistic understanding is integral to the development of new diagnostic and therapeutic approaches. A great challenge faced by the field of psychiatry is bridging the vast explanatory gaps across levels of analysis: from neurons, to circuits, to complex human behavior. At present, there is a lack of unifying understanding across this computational hierarchy for any psychiatric symptom. This profoundly limits the field's capacity to develop novel and rationally guided therapies that are grounded in well-established neural mechanisms.

It is no longer controversial to conceptualize severe mental illness as brain-based disease. The explosion of research in neuroscience has established that the brain is essentially a massively complex, nonlinear computational device that governs expression of an organism's behavior. This perspective is also captured by the ongoing paradigm shift within the field of psychiatry, from categorical, behavioral psychiatric diagnoses toward understanding of dimensional disruptions in complex behaviors that map onto well-defined neural substrates. This perspective is not intended to argue for a reductionistic model of complex mental illness. Critically, this view does not posit that psychiatric disorders cannot emerge as a result of external causal factors or perturbations (e.g., long-term stress or extreme trauma) that affect the brain. Nevertheless, such perturbations ultimately impact a set of neural processes and computations that govern a behavioral response (e.g., fear expression). Put simply, there is an organ that psychiatrists are attempting to treat—namely the CNS.

Given the complexity of the brain and the vast number of possible disruptions, across distinct levels of analysis, the field of psychiatry should attempt to harness all available tools to mechanistically understand abnormal behavior, regardless of whether the upstream cause lies in genes or external stressors. One promising way to link these levels is to combine empirical findings from clinical neuroscience with approaches from computational and theoretical neuroscience, in an emerging field often called "Computational Psychiatry." The key logic here is to use mathematical principles and formalism to generate consistent, rigorous,

and testable hypotheses that can lead to better understanding of mechanism across levels of analyses.

In that sense, computational psychiatry offers the opportunity to interface with experimental and treatment studies at the level of neurons, neural systems, and ultimately abnormal behavior (Fig. 1.29–1). Put differently, disrupted mechanisms that contribute to a particular psychiatric disease can occur at the level of neurons and synapses, whereas psychiatric symptoms express at the level of behavior, which involves large-scale brain networks. Linking these disparate levels is vital for gaining mechanistic insight into mental illness, and for rational development of pharmacological treatments, which act at the synaptic level. The primary objective of computational psychiatry is to help bridge this existing gap and to mathematically formalize computations across noted levels of analysis.

To date, mathematical models of neural systems have contributed greatly to our understanding of how neural dynamics and behavior arise from underlying mechanisms. Building on this mechanistic understanding, disruptions in the underlying elements of the model can then be linked to neural and behavioral dysfunction. In particular, researchers can evaluate which models or model parameters better describe empirical findings in a disease state compared to controls. By examining these disruptions within the models, we can describe key differences in underlying processes in a mathematically precise way. This approach can also test hypotheses about lower-level phenomena (e.g., specific synaptic disruptions) with experimentally accessible noninvasive measures at higher levels in humans (e.g., fMRI or psychophysical behavior).

Indeed, in the past several decades the field of psychiatry has profited tremendously from the ongoing development of noninvasive neuroimaging technology, more sophisticated human cognitive neuroscience, as well as advances in basic neuroscience. The combination of these approaches has directly impacted the broad evidence base in psychiatry, providing support for neural alterations across psychiatric conditions. Such evidence for neural alterations in psychiatric disease has emerged from a wide range of experimental modalities. Analysis of postmortem brain tissue reveals differences in neuronal and synaptic structure across psychiatric disorders. Genetic analyses implicate specific neuronal and synaptic alterations. Neurochemical alterations can be assessed in vivo using noninvasive methods such as PET and MRS. The spatiotemporal properties of brain activity during various tasks and at rest can be measured with noninvasive neuroimaging such as fMRI, EEG, and magnetoencephalography (MEG). Neuroimaging can also reveal structural differences in gray-matter distribution and long-range white-matter pathways. As noted, these neural measurements are complemented by rapid and parallel advances in cognitive neuroscience that allow us to quantitatively characterize, at the behavioral level, cognitive, affective, and perceptual impairments in

patients.

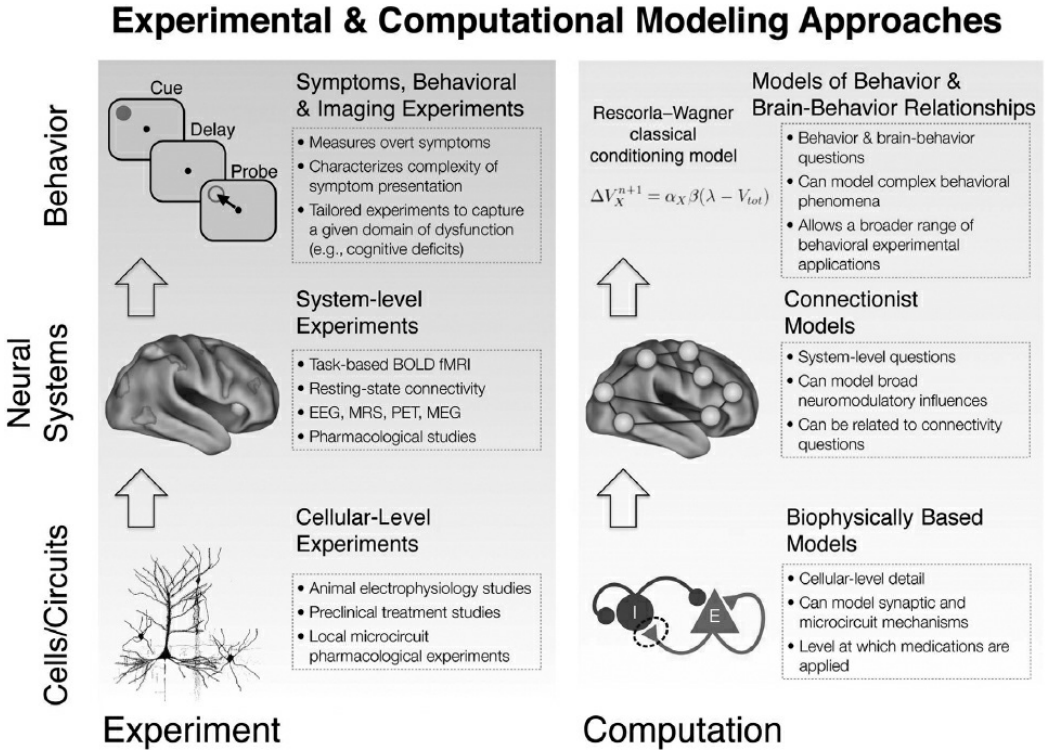

FIGURE 1.29–1. Multilevel analysis of hypothesized disruption informed by computational psychiatry. While current noninvasive neuroimaging methods offer major insights into neural system dysfunction in neuropsychiatric conditions, they cannot reveal synaptic or neuronal-level mechanisms. This leaves important explanatory gaps between neuroimaging systems-level observations and cellular-level hypotheses, such as those proposing cellular alterations on specific receptor subtypes. Computational psychiatry aims to help close these gaps by generating testable and translatable hypotheses across distinct levels of inquiry.

Despite this explosive technological and experimental progress, the field of psychiatry has yet to produce systematic and mechanistic disease models that are amenable to rational development of therapies. As noted, to achieve such a multilevel understanding of psychiatric diseases, the field of psychiatry must ultimately link these levels of analysis: from synapses, to cells, to neural circuits, to large-scale systems, and to behavioral disturbances. The emerging field of computational psychiatry is in a unique position to help provide the theoretical foundation needed to close these gaps. Computational psychiatry benefits from the diverse range of complementary modeling approaches within the broader field of computational neuroscience. Different types of models are best suited to address different questions and different types of experimental data. In that sense, the level of computational modeling is matched to the appropriate levels of experimental inquiry, enabling the interplay between experiment and theory. Multilevel understanding of complex abnormalities in behavior (i.e., psychiatric symptoms) directly benefits from these complementary modeling approaches (Fig. 1.29–2).

This chapter describes three broadly defined levels of computational modeling. Biophysically based models are constrained by realistic physiological properties of synapses, neurons, and circuits. These models are well suited for testing synapse-level hypotheses and to allow comparison to physiological and pharmacological studies in psychiatry. This type of modeling is also well suited to study neurophysiological biomarkers of disease states. Connectionist models use more abstract neural elements but incorporate a systems-level neural architecture. These models can be applied to more complex behaviors and are also well suited to interface with task-based neuroimaging and behavior—namely cognitive neuroscience approaches. Mathematical models of behavior formalize the algorithms underlying cognitive computations, which can in turn formalize behavioral symptoms and cognitive deficits. These models are well suited to quantitatively fit behavior and symptoms. Collectively, this chapter provides a conceptual overview of the current state-of-the-art in computational psychiatry, spanning from local neuronal circuits to complex behavior such as reinforcement learning (RL). Finally, ongoing challenges and future directions in computational psychiatry are discussed.

BIOPHYSICALLY BASED CIRCUIT MODELS

Biophysically based circuit modeling refers to a computational framework for modeling neural systems that is grounded in neurobiological mechanisms. These models can simulate neural circuit activity and computation at the levels of neurons and synapses, constrained by detailed physiological characterization of neuronal and synaptic dynamics. These models can be further constrained with systems-level experimental measurements of circuit architectures, neural activity, and behavior in tasks. In these models, dynamic activity is typically simulated through systems of differential equations governing the biophysical properties of neurons, synapses, and ion channels. Patterns of activity across neurons are shaped by connectivity within the circuit along with external inputs, which are also grounded in basic anatomy principles. This allows such models to be tested with empirical measures of neural activity. In certain models, behavioral responses can be read out from the patterns of activity, allowing comparison with behavioral data from various task paradigms.

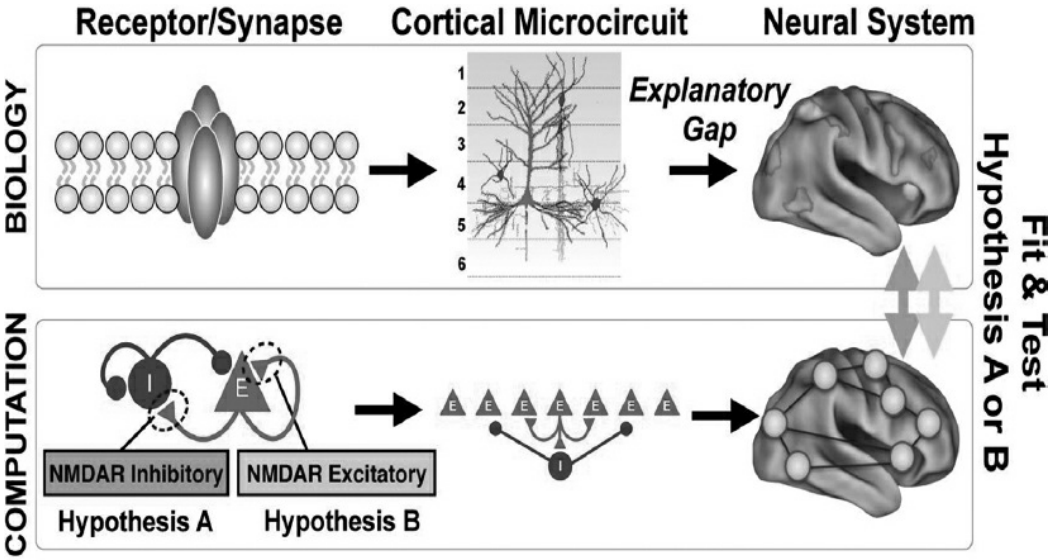

FIGURE 1.29–2. Conceptual illustration of the computational modeling and experimental interplay across levels of analysis. The utility of computational modeling, particularly in the study of schizophrenia, is its ability to inform a given level of experimental study. Because we study abnormalities in schizophrenia from the cellular level at the neural system level, and ultimately at the level of behavior, we have to utilize our modeling approaches to best fit the experimental framework. For instance, cellular-level experiments use techniques and produce measurements that are best captured using models that contain the necessary level of biophysical realism (bottom panels). Such models can, for instance, inform synaptic processes that may govern the microcircuit phenomena under study such as neural oscillations. In turn, a number of neuroimaging studies have focused on characterizing system-level disturbances in schizophrenia both using task-based paradigms and resting-state functional connectivity approaches. To best inform such system-level cognitive neuroscience experiments, models should capture the relevant detail and complexity of larger-scale neural systems (middle panel). Such models can perhaps better inform the role of systemic pharmacological manipulations on blood-oxygen-level-dependent (BOLD) functional magnetic resonance imaging (fMRI) or can be used to predict results of functional connectivity studies in schizophrenia. Finally, schizophrenia produces complex and devastating behavioral symptoms, which can be measured via increasingly sophisticated behavioral paradigms. Here the use of models that formalize complex behavior can provide a powerful tool to quantitatively examine a given behavioral process in patients (e.g., reinforcement learning).

In computational psychiatry, the biophysical basis of circuit models provides opportunities to probe how synapse-level disruptions produce aberrant neural activity and deficits in cognition and behavior. For instance, one can perturb specific synaptic parameters in the model that are physiologically interpretable and implicated in hypothesized disease mechanisms. This allows direct implementation of hypothesized perturbations related to disease mechanisms and pharmacological manipulations. One can then characterize the impact of these perturbations on emergent neural activity and behavior. Therefore, the model results can be related to human neuroimaging and psychophysics in healthy and clinical populations, as well as in animal models of disease states. Biophysically based circuit models have the specific potential to inform our understanding and design of pharmacological therapies, because they

incorporate details of synaptic and neuronal properties that are altered by psychiatric medications.

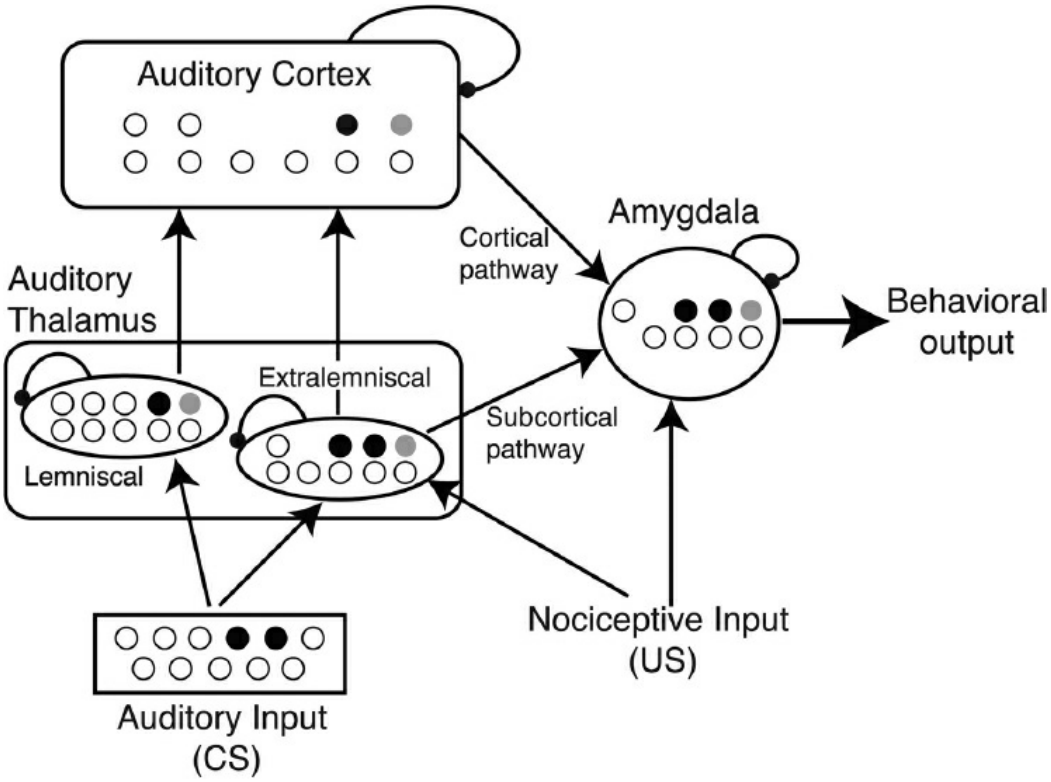

FIGURE 1.29–3. Illustrative example of a connectionist PDP-style model in the context of fear conditioning. In the case of fear learning, PDP models have successfully merged connectionist-level architecture across distinct "modules" that represent specific neuroanatomical structures while remaining committed to a certain level of biophysical realism (i.e., by modeling the amygdala based on animal physiology experiments). In this exemplar case, each larger module is comprised of individual nonlinear units (black units = maximum activation; white units = zero activation; gray units = intermediate level of activation), as is typically done in PDP models. Large arrows denote connections across modules, which are modeled as feedforward and excitatory, corresponding to pathways in the actual rat brain (although they are simplified here). Connection strengths across the modules are modulated as a function of learning by incorporating an extended Hebbian rule. (Adapted from Armony JL, Servan-Schreiber D, Romanski LM, Cohen JD, LeDoux JE. Stimulus generalization effects of auditory cortex lesions in a computational model and in rats. *Cereb Cortex.* 1997;7(2):157–165.)

Despite these inherent advantages, due to the level of detail in biophysically based modeling, it remains a challenge to relate circuit models to specific complex cognitive functions. For some core cognitive computations, we have knowledge of the neural circuit basis underlying these processes, which typically involve contributions from preclinical animal studies. For these cases, detailed circuit models can be developed rigorously to provide the link from synaptic disruptions to behavior (e.g., cognitive deficits discussed below). However, due to the large number of physiological parameters, these models are more difficult to fit quantitatively to behavioral performance. In other cases, psychiatric

symptoms relate to complex functions for which we lack understanding of the underlying neural circuit processes. At present, these circuit models are limited and cannot be applied to complex behavioral tasks, for which we lack understanding of neural circuit correlates.

Biophysically based circuit models can also provide insight into the mechanisms underlying neurophysiological biomarkers of disease states (see section below). These biomarkers can include electrophysiological measurements of aberrant neural oscillatory dynamics and neuroimaging characterization of large-scale dysconnectivity in resting-state activity. Put differently, some neural measures are directly translatable from cells, to neural systems and from animals to humans (e.g., electrophysiological signatures measured via single electrodes, electrode arrays or via EEG in humans). In such cases biophysically based models can provide testable predictions that can be readily translated from animal to human experiments.

Linking Synapses to Cognition: Working Memory

An important application of biophysically based neural circuit models to cognitive processing has been in the topic of working memory. Working memory refers to the ability to encode and transiently maintain or manipulate information internally, in the absence of sensory stimulation (e.g., remembering a phone number for a few seconds). This neural computation is fundamental for normal expression of behavior and relies on the delicate balance of local microcircuit neurons. Impairment in working memory is a prominent cognitive deficit in a number of psychiatric and neurological disorders, including schizophrenia and ADHD.

Working memory function is also a promising candidate in clinical neuroscience as an endophenotype, a quantitatively measurable core trait that is intermediate between genetic risk factors and a psychiatric disorder. Such transdiagnostic endophenotypes may be more tractable for research into mechanistic underpinnings of symptom dimensions. In that sense, working memory is a useful test bed for applying biophysically based modeling to cognitive function, because there are experimental paradigms that provide multilevel data that can constrain and test models.

Neural Circuit Basis of Working Memory. Neuroscientific studies in animals and humans have revealed important neural circuits and systems involved in working memory. Working memory engages a distributed network of cortical and subcortical areas, with prefrontal cortex (PFC) as a key node in this network. Single-neuron recording experiments in monkeys have found that a neural correlate of working memory in PFC is stimulus-selective persistent activity, in which neurons show elevated firing rates during the mnemonic delay. For instance, in one well-studied experimental paradigm, the oculomotor delayed response task, the subject must

maintain in working memory the spatial location of a visual cue across a delay period to guide a saccadic eye movement toward that location. During the mnemonic delay, a subset of prefrontal neurons exhibit tuned persistent activity patterns, with single neurons firing at elevated rates for a preferred spatial location. This paradigm has been successfully extended to human studies and extensively tested via functional neuroimaging, confirming the neural correlates of working memory in humans. In that sense, this complex cognitive behavior is an ideal example where the synaptic, neuronal, system-level, and behavioral mechanisms are well linked in both animal and human studies. We focus here on biophysically based models of working memory that can offer specific insights into the circuit mechanisms causing disruptions in this complex behavior.

Attractor Network Models. A class of circuit models called *attractor networks* provides a leading framework for understanding the neural circuit basis of persistent activity supporting working memory. The term *attractor* comes from the mathematical formalism of dynamical systems: an attractor state is an activity pattern that is stable in time, so that after a small perturbation to the state of the system, the system converges back to the attractor state. In attractor network models, these dynamics are typically formalized by a set of differential equations that capture the neuronal dynamics over time. An attractor network model of working memory typically possesses multiple attractor states: a low-activity baseline state and multiple stimulus-selective memory states in which a subset of neurons are persistently active. Because the memory state is a stable attractor state, it is self-reinforcing and resistant to noise or perturbation, allowing the memory to be maintained over time. This critical property of such models has been shown to rely on the slow currents mediated via the NMDARs on excitatory pyramidal cells. It is important to highlight here that it is precisely this level of detail in biophysically based models (i.e., NMDAR conductance) that allows for perturbing and testing of specific synaptic contributions to the neural circuit or ultimately behavior (i.e., working memory performance).

In a typical attractor network, subpopulations of neurons are selective to different stimuli. Neurons of similar selectivity show a Hebbian pattern of strong recurrent excitatory connections between them. When the strength of recurrent excitatory connections is strong enough, they can support a persistent-activity memory state. Strong, structured recurrent excitation sustains the persistent activity. Inhibitory GABAergic interneurons mediate strong feedback inhibition that serves two roles. First, it stabilizes the low-activity baseline state. Second, during persistent activity they provide lateral inhibition that enforces selectivity of the working memory representation, preventing the spread of excitation to the entire population of neurons in the network. The stability of attractor

dynamics, which depends on excitation and inhibition, provides the network with a natural resistance against noise and distractors that intervene during the delay (i.e., a stable attractor state). This robustness of working memory against distraction is a key cognitive property that supports cognitive processing.

Detailed biophysically based models have revealed important constraints on circuit function and made specific predictions that are experimentally testable. Strongly recurrent networks are prone to oscillations, which can destabilize persistent activity patterns and destroy working memory function. As noted, a decade of computational modeling work has predicted that persistent activity can be stabilized if recurrent excitation is mediated primarily by slow NMDAR, rather than by fast α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors, to provide temporal stability of the delay activity. A number of model predictions have been tested through studies that record the activity of single neurons while focally applying pharmacological agents localized to the cortical tissue surrounding that neuron. In line with model predictions, it was experimentally established that locally blocking NMDA receptors can abolish delay-period persistent activity, and that locally blocking GABA receptors leads to decreased stimulus selectivity during the delay through elevated firing. These experimental tests support the critical roles of balanced excitation and inhibition in PFC in supporting working memory function, a neural property that is hypothesized to be affected by severe psychiatric conditions (e.g., schizophrenia).

Dysfunction in Attractor Models. A number of computational modeling studies have analyzed how attractor network function is impacted by alterations of synaptic and neuronal properties. Optimal cortical function depends on the proper balance of excitation and inhibition (E/I balance). Disruption of E/I balance in cortex is implicated in the neuropathology of a number of psychiatric disorders, including schizophrenia, autism, and ADHD. As described above, attractor models require strong recurrent interactions between pyramidal neurons and interneurons, and are therefore sensitive to perturbations of E/I balance. Because these models also implement the core cognitive process of working memory, which is known to be disrupted in psychiatric disorders, attractor models provide a useful test bed to link disrupted E/I balance to cognitive deficits and generate critical experimental predictions across levels of analysis (see Fig. 1.29–2).

These models have been most often applied to circuit disruptions implicated in schizophrenia, for which a number of cortical alterations are proposed. Although neuronal microcircuit alterations in schizophrenia are complex, many studies have converged on interneuron dysfunction as a key component of its pathophysiology. Impaired interneuron function can

result in weakened recruitment of feedback inhibition, causing disinhibition of pyramidal neurons and an elevated E/I balance in the circuit. Computational models have explored the neural and cognitive effects of disinhibition in working memory circuits. A mild disinhibition can induce a broadening of working memory representations, which can be understood from the role of inhibition in maintaining the selectivity of persistent activity patterns. A strong disinhibition can destabilize the baseline state, causing working memory patterns to spontaneously appear (e.g., false memory representations). Importantly, these distinct regimes make dissociable predictions for performance in working memory tasks, which allows testing of these hypotheses with behavior. In addition to inhibitory deficits, there is also strong evidence that recurrent excitation is compromised in PFC in schizophrenia. In the attractor models, reductions in recurrent excitation can abolish persistent activity, preventing the circuit from implementing working memory. This regime again makes specific testable predictions for behavior and patterns of neural activity in the context of working memory tasks.

Attractor models have also been applied to study the effects of neuromodulation on working memory. Neuromodulators are known to alter the effective conductances of a number of synaptic receptor and neuronal channels. Dopamine (DA) has diverse effects of prefrontal cortical circuits, modulating multiple cellular and synaptic sites on both pyramidal cells and interneurons. DA also modulates working memory function, and DA signaling is found to be dysregulated in a number of psychiatric disorders. Numerous lines of evidence support the hypothesis that in schizophrenia there is impairment of DA signaling, via the D1 receptor, in PFC. One circuit modeling study found that D1 stimulation increases stability of both baseline and memory states in the network, which would be functionally beneficial during working memory maintenance. This in turn predicts that a hypodopaminergic state in PFC impairs working memory by weakening persistent activity and thereby rendering it vulnerable to disruption by noise or distraction. In contrast, an aberrantly hyperdopaminergic state could cause attractor states to be too strong, rendering the network inflexible to signals that should normally induce transitions between internal states.

In turn, biophysically based models can also be scaled to the level of neural systems. Extending biophysically based models in this way can generate experimental predictions for neuroimaging experiments that can be tested via causal pharmacological manipulations in humans or during cognitive tasks in psychiatric populations. More specifically, one of the most established findings coming out of human neuroimaging experiments implicates an “anti-correlated” neural architecture. There are brain regions specifically involved in orienting to an externally demanding cognitive task (i.e., task-positive areas) or areas that are more active when engaged in

internal mental states but suppressed during cognitive demand (i.e., task-negative areas or sometimes referred to as the default mode network, DMN). Studies have found that during demanding cognitive tasks patients with schizophrenia typically cannot suppress task-negative areas and concomitantly cannot activate task-positive regions—a process also shown to occur following NMDAR antagonist administration in healthy individuals during working memory. Biophysically based microcircuit computational models have been expanded to capture the reciprocal putative antagonism between task-positive and task-negative neural systems in a biologically plausible fashion by incorporating mutual synaptic inhibition. In this way, such models can generate system-level predictions for task-based neuroimaging experiments in humans while still capturing the relevant synaptic detail. Nevertheless, as noted previously, these models are still limited in the repertoire of complex behaviors they are able to capture due to the inherent constraint placed by biological realism in such models. Therefore, this branch of computational psychiatry is not yet positioned to address the range of complex cognitive and affective computations that can be represented in humans. This is at present much better accomplished by another type of modeling—namely connections models—which we turn to next.

COGNITIVE PSYCHIATRIC NEUROSCIENCE AND CONNECTIONIST MODELS

The sections above introduce the interface between biophysically based computational modeling and clinical neuroscience. As noted, this branch of modeling is distinguished by building on the physical properties of real neurons and synapses, producing realistic cellular-level dynamics. This level of detail is ultimately vital to generate mechanistic predictions at the level of treatments or to elucidate synaptic mechanisms governing complex behaviors in psychiatric patients. Nevertheless, a parallel line of computational psychiatry research has emerged in lockstep with the cognitive neuroscience revolution, building on the rapid and productive advances in noninvasive neuroimaging technology. Specifically, this branch of modeling can be broadly referred to as *connectionist models* that are designed to capture the function of large-scale neural networks and generate behavioral predictions in conjunction with human cognitive neuroscience experiments.

Emergence of Cognitive Neuroscience in Psychiatry

Before detailing the contributions connectionist modeling has made to psychiatry, it is important to set the stage for the type of research questions these models are best suited for—namely cognitive neuropsychiatry. The past two decades have witnessed an explosive growth of knowledge regarding the neural correlates of various cognitive and affective processes

in healthy individuals. This burgeoning field of cognitive neuroscience has generated an increasingly robust platform for interpreting clinical neuropsychiatric phenomena. This is primarily accomplished by garnering an increased understanding of neural systems known to be involved in various cognitive operations in healthy individuals. This understanding in turn constrains our search space for what aspects of brain circuitry may be abnormal in clinical populations exhibiting deficits in these same cognitive operations. Therefore, using a cognitive neuroscience framework with the specific application toward understanding computational mechanisms offers a promising tool for elucidating and ultimately treating psychiatric symptoms by delineating abnormalities in neural circuits whose functions are increasingly understood in healthy populations.

This cognitive neuroscience framework has emerged in parallel with the development of computational modeling in psychiatry. Historically, the types of computational psychiatry models that first appeared were most directly aligned with the neural systems and behavioral levels of analyses. Therefore, there exists a productive and ongoing interplay between cognitive neuroscience in psychiatry and connectionist modeling approaches.

Connectionist and Parallel Distributed Processing Models

There is a long history of using models to understand CNS function. Such a review is beyond the scope of this chapter and we refer the reader to one of the references at the end of the chapter. However, briefly it is worth considering the long-standing analogy of the brain as a "computer." Growing sophistication of algorithms and principles in basic computer science has cross-fertilized the emerging application of such approaches in neuroscience. One particular framework that has emerged out of this effort is modeling of neural networks. More formally, models of this class are called parallel distributed processing (PDP) or connectionist models. This broad class of models has emerged in an effort to elucidate the neural correlates of complex psychological functions, including cognitive control, language, and fear processing.

Importantly, these models are distinct in their flavor, architecture, and algorithmic implementation from the biophysically based (or behavioral) models reviewed in this chapter. The key distinction is represented in their ability to incorporate some level of abstraction beyond strict biological realism. A major objective of PDP models is to develop a deeper understanding of the mechanisms driving cognition from a psychological perspective, but using principles thought to capture aspects of computations as they might be carried out in the brain. Put differently, such models do not necessarily build in explicit neurobiological constraints or assumptions.

A typical PDP model architecture is conceptualized as a series of "units"

that are interconnected in some fashion and that represent some computational process. Each unit implements a simple transfer function that summates its inputs from other model units and transforms that into an output activity level. The state of the network is described by the graded activity patterns over populations of these simple units, and information processing occurs through the transformation of activity patterns from one set, or layer, of units to another. To achieve the desired information processing, the connection strengths between units are modified through a learning algorithm, such as *back-propagation*. This training of connection weights is meant to capture some elements of the type of Hebbian learning known to govern changes in connection strength between neurons.

In the back-propagation learning algorithm, the network is presented with an input pattern, and the output layer is allowed to generate its response. At each layer of the network, an error signal is computed as the difference between the response generated by the network and the desired target response. These error signals are propagated back through the network toward the input layer, used to modify the connection weights between units. The implementation of back-propagation learning provides an example of how connectionist models, while capturing complex neural phenomena, may not map onto specific neurobiological mechanisms (in contrast to biophysically based models described above). Nevertheless, this approach captures the idea that learning occurs through an incremental minimization of error by small changes in the connections between units (a phenomena referred to as *gradient descent*).

In typical PDP models, the units are abstracted, without a direct correspondence to individual neurons or explicit neurobiological assumptions. Units instead represent processing accomplished by neuronal assemblies and capture the general principles of neural information processing. However, some modeling studies have incorporated neuroanatomical constraints and specific neurotransmitter modifications, at varying levels of resolution.

Connectionist models help to build bridges between understanding of low-level properties of neural systems and their participation in higher-level (systems) behavior. They are able to capture a wide range of complex behaviors and neural systems interactions, which cannot at present be modeled as effectively by biophysically based models. Therefore, PDP models continue to be successfully applied in computational psychiatry to address questions and scenarios in which the field lacks firm empirical evidence to inform strong biological assumptions.

Applications of Connectionist Models in Psychiatry

This section presents three example areas of research where PDP models have provided novel insights and established a vital interplay between basic neuroscience and psychiatric questions. First, the focus is placed on how

PDP models have been used to understand the formations of hallucinations and delusions in schizophrenia. In turn, we review the role of PDP models in understanding cognitive control deficits, which cut across many psychiatric conditions. Third, we discuss the use of connectionist models in the context of fear conditioning, specifically from a systems neuroscience perspective with possible implications for understanding PTSD.

Connectionist Modeling of Hallucinations in Schizophrenia. Some of the seminal models using the PDP framework were developed to examine the impact of abnormal synaptic pruning in cortex and its possible consequences on symptoms observed in schizophrenia, specifically formation of hallucinations, which are abnormal internal sensory experiences in the absence of external sensory input. This mechanism is based on the neurobiological conceptualization of schizophrenia as a neurodevelopmental illness associated with abnormalities of synaptic pruning over the course of development, which in turn impacts the functioning of large-scale neural systems.

PDP models implemented this mechanism by eliminating the connections between units thought to represent synaptic connections. This line of research showed that if a sufficient number of connections were removed, then the model would repeatedly return to an abnormal state referred to as "parasitic foci." That is, researchers observed patterns of activity across units in the model that did not correspond to any learned memory, but to the "state" to which the model kept returning—perhaps capturing an abnormal internally generated state that could occur during hallucinations. Researchers in this area hypothesized that these parasitic foci may behaviorally map onto delusions and hallucinations. A concrete example of how PDP models may help explain hallucinations came from applications to understanding language processing. Using the PDP-style architecture, research found that language could be represented within an "Elman" network—namely a sequential recurrent network that consists of an input and an output layer connected by a "hidden layer" that computed internal processing. This architecture included an additional layer connected to the hidden layer that could temporarily store information, thought to support working memory functions (temporary storage of information) during speech perception. In this architecture, it was shown that elimination of connections between units (i.e., "synapses") in the working memory subnetwork of the model initially improved speech perception. However, as the severity of the elimination increased, the model started to fail to detect words correctly and started to generate "hallucinations." In this PDP model, this was computationally captured as patterns of activity across the output layer that corresponded to words even when there was no structured input into the model. In contrast to synapse elimination, a small of amount of "cell" removal (e.g., elimination of units)

also initially improved speech perception, but greater cell loss did not generate hallucinations. Critically, these early PDP models provided an important "proof of concept" for how various neurobiological changes might affect connectionist networks.

In turn, such PDP models can also incorporate some aspects of broad neuromodulation. For instance, researchers that extended the PDP models of hallucinations in schizophrenia also captured the potential influence of increased DA in the model. This is concretely accomplished as a gain function within the model by abstracting a neurobiologically plausible broad "neuro-modulatory" influence on the entire PDP network. In schizophrenia, the role of DA is hypothesized to play a key role in cortical "tuning." Based on this hypothesis, PDP models incorporated increased DA by changing the bias onto the units in the working memory module, effectively requiring them to accumulate greater information to maintain information. This manipulation only relates to one aspect of the influence of DA on neuronal information processing and by itself it did not fully capture clinical observations. However, a combination of increased DA and synaptic pruning more successfully captured the performance of those patients who experienced hallucinations.

These types of models incorporate aspects of brain-style computational principles, but did not make strong commitments to key aspects of neurobiology during model development. Nonetheless, they provided critical steps and a mathematical formalism to understand how neurobiological or neural network deficits might contribute to the formation of core symptoms of psychosis, such as hallucinations. Although the models themselves did not produce behaviors with the same format and content as shown by human patients, the "in principle" examples provided conceptual leverage to bridge the gap between biology and behavior, and even extend to treatment recommendations (e.g., work in the area of transcranial magnetic stimulation that explicitly leveraged insights from these models). Therefore, despite not explicitly committing to biological realism, PDP models can provide useful predictions for treatment development.

Connectionist Modeling of Cognitive Control Defi cits. A parallel line of research leveraged PDP models to understanding cognitive control deficits that emerge across psychiatric conditions but are particularly prominent in schizophrenia. These models developed a more neurobiologically plausible architecture by postulating that "cognitive control" in the CNS arises from interactions between the DA neurotransmitter system and the PFC. These models suggested that goal-related information, or context information, was maintained in PFC and used to bias stimulus response mappings represented in posterior cortex, serving as a source of top-down support for controlling behavior. In these models, the "context" module was associated

with the functions of the DLPFC. Active maintenance in the absence of external inputs was assumed to occur via recurrent excitation, as suggested by neurobiological data that was reviewed in the sections above. This branch of models assumed that feedback projections from the context/DLPFC module biased processing in posterior systems via direct excitation, but that competition between representations within a processing stream occurred via lateral inhibition. Further, the effects of DA were assumed to be modulatory rather than purely inhibitory, such that changes in DA activity could either increase or decrease activity depending on the nature of inputs that DA was modulating. Put differently, increases in DA were assumed to increase the signal-to-noise ratio of a unit's activation value in relation to its input.

Concretely, these models have been successfully used to generate predictions about a specific pattern of complex behavioral errors that are observed during cognitive paradigms evaluating context processing deficits in schizophrenia. Specifically, the models were used to test the hypothesis that reductions in DA input into lateral PFC lead to impairments in the ability to represent and maintain context information in lateral PFC, and that this impairment in the representation of context leads to deficits in a range of cognitive tasks among individuals with schizophrenia. Behaviorally, the model predicted that individuals with cognitive deficits should exhibit a specific pattern of behavioral errors when context information needed to be maintained. This pattern of errors was largely confirmed across a series of studies in schizophrenia patients. In turn, this type of PDP model was also successfully leveraged to generate predictions about patterns of lateral PFC signals found during task-based cognitive activation neuroimaging paradigms. The specific prediction from these models suggested that individuals with schizophrenia should show reduced lateral PFC activity in conditions with a strong demand for context maintenance. Again, a number of empirical reports confirmed this prediction, showing that patients with schizophrenia exhibited deficits in recruitment of lateral PFC specifically in line with model predictions. Subsequent work in this area suggested that deficits in DA signaling may contribute to abnormal context maintenance and information gating in lateral PFC, which in part contributes to cognitive processing deficits seen in this patient population.

Collectively, this line of research illustrates how PDP models can be used to make both behavioral and neural predictions of a given behavioral symptom (i.e., cognitive deficits) that can in turn be explicitly tested via cognitive neuroscience paradigms in patients. This line of research has also been successfully leveraged to understand the neural underpinnings of specific affective deficits in psychiatric populations—namely fear conditioning.

Connectionist Modeling of Fear Conditioning. It is well established that severe deficits in emotion processing cut across different diagnostic categories, including schizophrenia, bipolar illness, and anxiety disorders to name a few. Here we specifically focus on computational modeling studies in a specific area of affect processing—namely fear expression. Abnormal fear expression provides a powerful translational target that can be understood computationally across mammalian species given the well-preserved neurobiological mechanisms underlying basic fear expression. At the basic preclinical level, the role of the amygdala in fear generation is perhaps one of the best-studied neural computational mechanisms, having been translated across levels of analysis, including human studies. Clinically, abnormal fear expression can be found across a host of anxiety disorders, including generalized anxiety, social anxiety, specific phobias, and PTSD. Therefore, computationally formalizing abnormal fear circuits represents a particularly broad and important challenge for computational psychiatry.

Computational studies in this area have used the amygdala as a central starting point. The amygdala is a small structure comprised of several nuclei located anterior to the hippocampal formation, and is critical for the normal expression of fear and learning of fear-related associations (i.e., classical fear conditioning). In that sense, PDP models in this area build on both well-understood biophysical principles of this neural structure and its well-established large-scale anatomical connectivity. This is a good example of how PDP-style models can at times incorporate more explicit commitments to biologically plausible circuits and instantiate biophysical realism while maintaining a system-level architecture with useful behavioral predictions (e.g., fear learning). Specifically, such anatomically constrained PDP models of fear circuitry have been used to generate predictions for how multiple interacting neural structures could work jointly to accomplish a pairing between the unconditioned stimulus (US, nociceptive stimulus) and the conditioned stimulus (CS, auditory input) (see Fig. 1.29–2). Using this model architecture, it was found that fear learning can be established via the joint pattern of activation across specific unit pairs that map the CS and UC. Furthermore, the model provided a powerful way to generate predictions for how focal lesions may affect fear learning—a specific advantage of PDP models that are developed based and constrained via neuroanatomically plausible principles. This line of work in particular can help generate useful predictions for abnormalities in fear expressions that may result from neuropathology to one or more key components of the fear circuit.

Collectively, this work demonstrates how connectionist models can be used creatively and flexibly to generate behavioral and neuroimaging predictions across a wide range of neural systems and behaviors. As noted, this flexible repertoire of PDP-style models is possible because their

architecture is not always explicitly developed or constrained based on neurophysiology or neuroanatomy. That said, these models still make a number of strong commitments to represent neural computations. For this reason, PDP models are not always capable of formalizing even more complex behaviors that may occur in psychiatric patients (e.g., formation of complex and abnormal belief patterns such as delusions, or deficits in motivation and reward learning). In such cases where neither PDP models nor biophysically based models can provide the final links to behavior, we turn to mathematical models of behavior that are explicitly developed to formalize the patterns of complex psychiatric symptoms.

MATHEMATICAL MODELS OF BEHAVIOR

The computational psychiatry approaches discussed above, biophysically based circuit modeling and connectionist modeling, aim to develop models with elemental components that instantiate how neural systems transform their inputs into outputs. These approaches are often limited in the range of cognitive behaviors to which they can be applied, due to our limited knowledge of the relevant neurophysiological mechanisms. An alternative approach to modeling cognition is to develop mathematical models that provide an account of the psychological computations underlying a given cognitive function. These models are often grounded in so-called normative accounts that aim to define the algorithms that optimize performance at a task and to characterize constraints that limit such task performance.

Although developed to explain behavior, these models can potentially make important links to neural circuits. For example, experiments can reveal that activity in specific brain areas represents the internal variables of the model, providing support for this computational account and providing insight into the division of labor among brain systems. In turn, such neuroscientific findings can inform the algorithms and modules that compose the computational architecture of the model.

Impairments in decision making are a common cognitive symptom across psychiatric and neurological disorders. Mathematical models of behavior can formalize these symptoms in rigorous and quantitative terms, potentially isolating specific components of decision-making computations that are altered in a given psychiatric condition. These models typically have fewer parameters than the neural systems models described above, which enables them to be more easily fit to behavioral data or to explore the full range of a given parameter. In that sense, behavioral models can provide a deeper intuition on the full range of a possible behavioral repertoire and help rigorously define the behavioral states that may be abnormal or maladaptive.

Reinforcement Learning

One form of dysfunction in decision making is disruption in the process through which one learns the values of available actions. Mathematical modeling and experiments have made great progress in studying this process in the context of RL. In RL, an agent learns the values of stimuli and actions through successive trials of feedback (reward and punishment), and selects actions based on these values to maximize one's overall outcomes.

A class of *temporal-difference RL* (TDRL) models has been widely applied to explain the process of learning the mapping from stimuli and states to actions. In these models, actions are selected on the basis of the total expected value for that action. The value of each action is learned through sequential updating based on a *RPE*. In its simplest form, the RPE is the difference between the reward received after that action and the expected value of the action. Thus, if the reward is greater than expected, a positive RPE signals that the value should be increased. If the reward is equal to that expected, the RPE should be zero. The formulation of the RPE can be extended to incorporate additional factors, such as temporal discounting of future outcomes. TDRL models are well supported by behavioral studies, as well as neuroscientific studies that have found brain areas that signal action values and RPEs. In particular, DA neurons in the subcortical VTA quantitatively reflect internally computed RPEs through bursts of action potentials.

Many key aspects of addiction can be modeled as dysfunction in a TDRL model. It is thought that drugs of abuse access the natural RL systems in the brain, leading an agent to over-select drug-seeking actions. Specifically, drugs of abuse can intervene in the pathways that activate the DA neurons that are involved in signaling RPEs. For instance, cocaine can induce phasic firing of DA neurons in the VTA. In TDRL, this can interfere with RPE signaling, strongly increasing the learned values for drug-seeking actions so that they come to dominate behavior. Modeling studies have found that implementing this drug-induced component to RPE signaling in a TDRL model can parsimoniously explain a number of important aspects of addiction-related behavior.

The TDRL model described above, which learns the values of actions for a given state, applies to habit-related learning. More elaborate RL models have also been developed to explore the roles of multiple learning systems in the brain. In addition to this habit-related form of learning (also called model-free RL), there is another type called model-based RL. Model-based RL calculates the values of actions from the values of future states, using internal representation of the causal structure of the environment (i.e., knowing which states are reached through the actions). Model-based RL is more computationally demanding for the agent (but also more flexible) than model-free RL, and is associated with cognitive processing involving in the PFC.

By comparing behavior in RL tasks to model performance, researchers can probe the relative reliance on model-free versus model-based RL. Cognitive deficits may impair function of model-based RL, leading to an increased reliance on model-free RL. Researchers have designed behavioral paradigms to measure the relative contributions of model-free versus model-based RL in decision making. Studies have found, both in addiction and in OCDs, that there is a favoring of the habitual model-free RL compared to healthy controls. These findings support the view of these disorders exhibiting habitual, repetitive behaviors.

Decision Making via Sequential Sampling

Another topic of mathematical modeling of decision making is the process of categorically selecting among available options. This has been most extensively studied in the context of two-alternative forced choice (2AFC) tasks. In 2AFC tasks, the decision-making process can be formulated as the sequential accumulation over time of evidence in favor of one option or the other. The drift diffusion model (DDM) is a phenomenological but mechanistic model that is widely applied to 2AFC tasks, and likely the optimal solution in certain contexts.

In the DDM, a decision variable evolves in time, accumulating noisy evidence. Because evidence sampling is inherently noisy, the decision variable undergoes a biased random walk, with a mean drift rate in one direction determined by the net evidence in favor of one option over the other. Computationally, the two choices correspond to a lower and an upper bound; when the decision variable crosses a bound, the corresponding choice is made. Because the noise is different on each decision-making trial, the choice and reaction time vary across trials. One strength of the DDM is that it generates not just a proportion of choices but a full distribution of reaction times that can be used to fit empirical choice data. At the neural level, the DDM has received support from electrophysiological recordings from a number of cortical areas that show slow ramping of neural activity during decision making.

The DDM can also explain aspects of impaired decision making, by describing behavior in terms of the values of DDM parameters. This has the potential to more precisely isolate and quantify differences in decision-making behavior. Consider, for instance, a clinical group that shows a propensity toward selecting one option over the other. In the DDM framework, this could be due to one decision boundary being closer than the other. Alternatively, the mean drift rate may include a component of the drift rate in favor of that option. Critically, these two potential mechanisms have dissociable effects on the patterns of errors and reaction times. Therefore, fitting subjects' behavioral data can isolate which alteration better describes behavior in a group of subjects.

Bayesian Approaches to Inference

Another fruitful area of research is the mathematical formalism of belief formation. This is most clinically relevant in cases where strong but abnormal beliefs can result in internal states that are clinically characterized as delusions. Here the biggest impact has been achieved by a class of mathematical models of *Bayesian inference*. Bayesian inference describes how the estimated probability of event *A* should be updated by observation of event *B,* according to Bayes' theorem:

$$P(A/B) = P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

where $P(A \mid B)$ is the probability of *A* given that *B* is true, $P(B \mid A)$ is the probability of *B* given that *A* is true, $P(A)$ is the probability of *A,* and $P(B)$ is the probability of *B*. Therefore, the *posterior* $P(A \mid B)$ is proportional to the *likelihood* $P(B \mid A)$ times the *prior* $P(A)$. The prior probability, the expectation for event *A* to occur, plays a key role in the inference of the posterior estimate from observation of event *B*.

In the context of delusion formation, Bayesian models hypothesize that psychosis is a neural and psychological state that is marked by abnormal priors about the world. Abnormal priors may result from a state of aberrant salience, that is, events in the environment are hypothesized to be more salient than usual, perhaps due to elevated DA signaling in the striatum. Bayesian models can in turn be applied to a range of behavioral experimental contexts where expectations are formed and then confirmed or violated, with the reported posterior estimates providing a window into the abnormal priors. Studies in this area support the view that abnormal priors (or posterior probability) lead to formation of abnormal inferences, which could lead to a stable internal belief that is not supported by current evidence (i.e., a delusion). Put simply, Bayesian models of behavior when applied to psychiatry have postulated that delusions may be an extreme state of abnormal inference. This framework has been subsequently linked to abnormal prediction error coding and extended to a neuroimaging context, providing a putative neural marker of abnormal belief formation.

MODELING NEUROPHYSIOLOGICAL BIOMARKERS

A major goal of clinical neuroscience is the discovery of robust neurophysiological biomarkers for psychiatric disorders. Understanding the mechanistic basis underlying these biomarkers could aid in the development of treatments, by mapping specific circuit perturbations to neurophysiological measurements. Biophysically based circuit modeling is well suited for linking these levels, as the models link synapse-level parameters to network-level dynamics. Models can make predictions for

how physiological biomarkers are affected by putative synaptic disruptions. In turn, model predictions provide the basis for inference in the other direction, to test hypotheses regarding site of a synaptic perturbation with measured changes in physiological biomarkers. Therefore, biophysically based circuit modeling provides a means to interpret specific biomarkers and generate hypotheses that can be tested further with synapse-level and cellular-level measurements. This understanding allows for interpretation of aberrant neural dynamics measured in patients in terms of circuit perturbations, which generates specific and testable hypotheses for experimental studies in animal models and clinical populations. Two topics studied in this way are cortical oscillations and large-scale resting-state brain networks.

Oscillations

Cortical circuits generate a wide range of oscillatory dynamics that arise from strong recurrent network interactions. Abnormal oscillations have been observed in a number of psychiatric and neurological disorders, typically measured through EEG or MEG. Such oscillatory deficits are observed both in local dynamics and in long-range coupling between brain areas. In schizophrenia, oscillatory abnormalities in the gamma range (30 to 80 Hz) are of particular interest. Cortical circuit models, comprised of pyramidal cells and interneurons, have been used to explore the disruption of gamma oscillations in response to perturbations implicated in psychiatric disorders such as schizophrenia.

Computational and experimental studies support a model in which neocortical gamma oscillations arise from a feedback loop of excitatory pyramidal cells reciprocally connected with perisomatic-targeting, fast-spiking (and parvalbumin-expressing) inhibitory interneurons. In these circuit models, gamma oscillations can arise through a patterned cycle of activation of pyramidal cells and interneurons. Excitatory input causes pyramidal cells to fire. This activity then recruits inhibitory interneurons to fire short delay. This feedback inhibition suppresses firing in the circuit. After this synaptic inhibition decays, the cycle begins again. Models show that such oscillatory behavior arises naturally in systems with strong recurrent excitation and inhibition if the feedback inhibition is slow or delayed with respect to the excitation. As in the case of working memory function described above, gamma oscillations depend on strong recurrent interactions between excitatory and inhibitory cells, and oscillatory dynamics are sensitive to E/I disruptions.

A number of computational studies have developed models of cortical microcircuits generating gamma oscillations to study how gamma-range synchronization is affected by synaptic perturbations implicated in disease states, with a specific focus on schizophrenia. These studies have characterized how the schizophrenia-associated disruptions of

parvalbumin-expressing interneurons can impair generation of gamma oscillations. Modeling studies have also characterized how gamma oscillations may be affected by dopaminergic modulation and antagonism of NMDA receptors on specific cell types. In each of these cases, the models provide testable hypotheses for the physiological effects of synaptic manipulations.

These models have proven useful in understanding the circuit basis of aberrant cortical dynamics in disease states. However, the roles of aberrant oscillations in cognitive deficits and psychiatric symptoms remain unclear. There is still controversy regarding whether oscillatory synchronization is fundamentally necessary for the computations performed by neocortical circuits during cognition. Theoretical models have shown that oscillatory synchronization can subserve specific computational functions, such as efficient information transmission between neural populations. Further computational studies are needed that posit specific roles for oscillations in cognitive processes and make specific predictions for how behavior could be impaired in disease states.

Large-Scale Brain Dynamics

An emerging topic in computational neuroscience is the dynamics of large-scale brain networks. Progress in this area has the potential to provide insight into the growing body of experimental literature characterizing psychiatric disorders with whole-brain functional neuroimaging via fMRI. These data may provide important biomarkers of disease states. One common approach in fMRI is to analyze low-frequency fluctuations present in the BOLD signal during rest. In particular, resting-state fMRI can provide a pattern of functional connectivity, that is, the variability of the BOLD signal across brain parcels. These patterns of functional connectivity can be related to the underlying structural connectivity between brain areas, which can be measured noninvasively in humans through diffusion tractography.

Recent computational models have been developed to capture how the global pattern of resting-state functional connectivity emerges through long-range interactions among cortical areas. These large-scale network models simulate the dynamics of multiple nodes, each representing a cortical parcel. The nodes interact through an anatomical coupling matrix reflecting long-range connections, typically derived from tractography studies. Ongoing temporal dynamics in the nodes produces patterns of correlated fluctuations in the activity across the network. The functional connectivity of the model can then be calculated and compared to functional connectivity observed in experiments.

In one class of these resting-state models, each node is simulated with reciprocally connected pools of excitatory pyramidal cells and inhibitory interneurons, with diffusion tractography in humans used to define the

structural connectivity pattern of excitatory long-range projections between nodes. Each node receives fluctuating background input to drive dynamics. As in the microcircuit models, the biophysical basis of these large-scale models makes them directly applicable to address the dynamical consequences of anatomical and physiological changes induced by disease processes or pharmacological manipulation. Changes in the strengths of local and long-range connections may induce differential effects in the patterns of functional connectivity. The models reveal that the resting-state functional connectivity is sensitive to both the structure of underlying anatomical connectivity and the physiological parameters that scale local and long-range connections. This can be used to test synapse-level hypotheses for a disease state by their predictions for resting-state alterations observed via fMRI. For instance, the hypotheses of elevated versus reduced E/I balance in local cortical circuits make dissociable predictions for the patterns of variability and covariability in the cortical BOLD signals that can be tested in patients with schizophrenia.

FUTURE CHALLENGES

Integration across Levels of Analysis

This chapter describes how computational psychiatry has offered a method to bridge levels of analysis, from local neuronal circuits, to large-scale neural systems and ultimately psychiatric symptoms. The challenge that still faces this rapidly growing field is how to provide a formal bridge between these levels of inquiry for most psychiatric symptoms. We highlighted example cases where such progress has occurred—namely in the context of working memory and fear conditioning. In such cases, computational psychiatry has begun to establish a formalism for how to link local circuits to large-scale neuronal dynamics and expression of behavior. However, the current models in this area of computational psychiatry are still simplistic in their repertoire of behavior and are designed to capture very limited range of experimental contexts.

To accurately capture more complex cognitive and affective behaviors, models must incorporate multiple interacting brain areas possessing various distributed and modular computations. In this area, biophysically based modeling will be informed by connectionist-style models that are more abstract and removed from biophysical detail, but that can be readily applied to modeling psychological processes and interactions among distributed neural systems. Similarly, a highly related challenge facing computational psychiatry is to provide a formal extension of biophysically based modeling to large-scale distributed brain networks, beyond the microcircuit level, as this level of analysis offers the potential for the development of disease-related biomarkers. For instance, a key extension of such models would be to capture, in a biophysically realistic way, the

large-scale thalamocortical functional loops that may play a critical role across a range of psychiatric disorders. The next generation of computational psychiatry studies will be vital to formalize these bridges across levels of analysis.

Computational Phenotyping

It is appreciated by both researchers and clinicians that individuals' psychiatric symptoms present with substantial heterogeneity, which cuts both within and across diagnostic categories. This conceptual problem is captured in the case of existing tension between category-based DSM-style diagnostic rubrics and the dimensional RDoC framework. Instead of making an a priori decision regarding which framework may be better for a given psychiatric symptom or syndrome, we posit that perhaps there should be a parallel movement toward "precision psychiatry." Simply put, it may be possible to harness advances in computational psychiatry to generate individual-level inferences regarding the type of underlying computational mechanisms that are governing the observed patterns of neuroimaging effects or behavioral symptoms.

The current state-of-the-art in this area of research focuses on group effects. That is, the majority of work that has emerged in computational psychiatry has focused on explaining group-level effects across a given diagnosis. We argue that one highly promising future application of computational psychiatry may be to provide individual-level computational phenotyping based on mechanistic models. Such level of precision may possibly provide a method to manage the massive clinical heterogeneity that exists both within and across our current clinical rubrics in psychiatry. Ultimately, mapping such clinical heterogeneity onto underlying computational mechanisms will be vital to develop better mechanistically informed treatments.

In computational phenotyping, the parameters of a model are fit to best capture the empirical data for a single subject. Computational phenotyping could take advantage of a range of empirical data and modeling approaches. For instance, an RL model could be fit to a subject's decision-making behavioral data. Similarly, a biophysically based network model could be fit to a subject's resting-state neuroimaging data. The promise of this approach is that these model parameters may have diagnostic value, potentially leading to more predictive biomarkers. These model parameters may allow effective classification according to diagnostic criteria, and may reveal clinically relevant clusters or dimensions within a diagnostic category, or be predictive of a patient's response to treatment. Longitudinal studies will be needed to validate the diagnostic potential of computational phenotyping.

Modeling Treatments

Finally, of utmost relevance to psychiatry, computational modeling has the potential to provide a method for simulating possible effects of treatments (or compensations) that can occur across the described levels of analysis. For instance, some modeling efforts have already begun to generate predictions for how compensations onto NMDAR or GABAR synapses in local circuits can impact cortical disinhibition that adversely affects working memory computations. However, this area of research is at present underdeveloped and has future potential for guiding intuition in cases where complex neural dynamics are involved.

Here further development and refinement of biophysically based models have particular advantages. Given the relevant cellular detail, such models can actually generate both biomarker and behavioral predictions of disruptions on specific synaptic sites. Such level of detail may be out of reach in cases that are not informed by animal models. Nevertheless, this is precisely the point of cutting-edge interplay between future research in computational psychiatry and clinical neuroscience. We posit that future studies in this area hold the exciting promise of rationally guiding treatment development in psychiatry, grounded in basic neuroscience.

References

*Abbott LF, Dayan P. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Boston, MA: MIT Press; 2001.

Adams RA, Brown HR, Friston KJ. Bayesian inference, predictive coding and delusions. *AVANT*. 2015;3:51.

Anticevic A, Cole MW, Repovs G, et al. Connectivity, pharmacology, and computation: toward a mechanistic understanding of neural system dysfunction in schizophrenia. *Front Psychiatry*. 2013;4:169.

*Anticevic A, Murray JD, Barch DM. Bridging levels of understanding in schizophrenia through computational modeling. *Clin Psychol Sci*. 2015;3(3):433–459.

Armony JL, Servan-Schreiber D, Cohen JD, Ledoux JE. Computational modeling of emotion: explorations through the anatomy and physiology of fear conditioning. *Trends Cogn Sci*. 1997;1(1):28–34.

Cohen JD, Braver TS, O'Reilly RC. A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. *Philos Trans R Soc Lond B Biol Sci*. 1996;351(1346):1515–1527.

Corlett PR, Taylor JR, Wang XJ, Fletcher PC, Krystal JH. Toward a neurobiology of delusions. *Prog Neurobiol*. 2010;92(3):345–369.

Deco G, Kringelbach ML. Great expectations: using whole-brain computational connectomics for understanding neuropsychiatric disorders. *Neuron*. 2014;84(5):892–905.

Huys QJ, Guitart-Masip M, Dolan RJ, Dayan P. Decision-theoretic psychiatry. *Clin Psycholo Sci*. 2015;3(3):400–421.

Insel TR. The NIMH research domain criteria (RDoC) project: precision medicine for psychiatry. *Am J Psychiatry*. 2014;171(4):395–397.

Kishida KT, King-Casas B, Montague PR. Neuroeconomic approaches to mental disorders. *Neuron*. 2010;67(4):543–554.

Lee D. Decision making: from neuroscience to psychiatry. *Neuron*. 2013;78(2):233–248.

Lisman J. Excitation, inhibition, local oscillations, or large-scale loops: what causes the symptoms of schizophrenia? *Curr Opin Neurobiol*. 2012;22(3):537–544.

*Maia TV, Frank MJ. From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci*. 2011;14(2):154–162.

*Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. *Trends Cogn Sci*. 2012;16(1):72–80.

Redish AD, Jensen S, Johnson A. A unified framework for addiction: vulnerabilities in the decision process. *Behav Brain Sci*. 2008;31(4):415–437; discussion 437.

Stein DJ, Ludik J. *Neural Networks and Psychopathology: Connectionist Models in Practice and Research*. Cambridge: Cambridge University Press; 2007.

Stephan KE, Mathys C. Computational approaches to psychiatry. *Curr Opin Neurobiol*. 2014;25(0):85–92.

Thomas MS, McClelland JL. Connectionist models of cognition. In: Sun R, ed. *The Cambridge Handbook of Computational Psychology*. Cambridge: Cambridge University Press; 2008.

Uhlhaas PJ, Singer W. Neuronal dynamics and neuropsychiatric disorders: toward a translational paradigm for dysfunctional large-scale networks. *Neuron*. 2012;75(6):963–980.

Wang XJ. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol Rev*. 2010;90(3):1195–1268.

*Wang XJ, Krystal JH. Computational psychiatry. *Neuron*. 2014;84(3):638–654.

▲ 1.30 Functional Genomics of Human Brain Development

SIRISHA POCHAREDDY, PH.D., JOHN C. SILBEREIS, PH.D., AND NENAD SESTAN, M.D., PH.D.

The human central nervous system (CNS) is possibly the most complex biological tissue comprising approximately 86 billion neurons connected by hundreds of thousands of kilometers of myelinated axons and several hundred trillion to well over a quadrillion synapses, along with a roughly equal number of glial cells. Reflecting this complexity, building the human CNS requires the precise regulation of molecular and cellular processes. Through this regulation, intrinsic processes encoded in the human genome and extrinsic modulators from the environment converge in development to shape the circuitry, anatomical features, synaptic interactions, and cellular processes that carry out neural function. The greater size and complexity of the human brain are of course foundational to the expanded intellectual, cognitive, and emotional capacities of humans. As a result of its size and complexity, the human brain as a whole, but most especially the association areas of the cerebral neocortex, takes relatively more time than the brains of other mammals to develop, and humans have a particularly long gestational time, as well as childhood and adolescence (Fig. 1.30–1).

This prolonged developmental course and period of dependency allows, more so than in other species, environmental factors to shape the development of cognitive, emotional, and social capabilities and capacities. Indicative of the biological challenge of precisely regulating diverse molecular and cellular processes over a protracted period of time and across myriad cell types and regions, the CNS exhibits regionally and temporally distinct patterns of vulnerability to various diseases and insults (Fig. 1.30–1). Thus, the many neurodevelopmental processes that enable human brain complexity seem to have required an evolutionary trade-off that rendered it especially susceptible to certain disorders.

Recent technological advances and increased focus on human neurodevelopment have enabled a more comprehensive characterization of the human CNS and its development in both health and disease. In particular, the genomic revolution has considerably advanced knowledge about the genetic variations that confer risk for various neurological and psychiatric disorders, confirming that psychiatric disease is by and large genetically complex and etiologically diverse.

In addition, "-omic"-level characterizations of the transcriptional (e.g., gene expression properties), gene regulatory, and epigenetic processes of the developing human CNS across developmental periods and brain regions have generated considerable new insights into the molecular landscapes that define human CNS development and their dysfunction in disease.

This chapter details the current understanding of the transcriptional, and regulatory architecture of the developing human neocortex and other regions of the CNS. It further provides examples of how broad dysfunction of these processes likely underlies neurodevelopmental and psychiatric disease. Emphasis will be given to studies of the neocortex, because of its importance in higher-order cognition, emotional regulation, complex behaviors, and the pathobiology of many neurodevelopmental and psychiatric disorders. Another chapter of this book titled "Neurodevelopment and Neurogenesis (Chapter 1.3)" describes the basic principles of the organization and development of the CNS, and it is recommended as companion reading.

THE TRANSCRIPTIONAL LANDSCAPE OF THE DEVELOPING HUMAN CNS

At the root of the cell and molecular processes governing human brain development is the immensely complex genome (see section "Organization of the Human Nuclear Genome"). Transcription refers to the conversion of genetic information encoded in DNA to mRNA; mRNA is then translated into the proteins that ultimately carry out the common and unique functions of specific cell types. Therefore, the selection of genes for transcription and the determination of gene expression levels are largely responsible for establishing the molecular and physiological properties of

distinct cell types and brain regions. Transcription is therefore tightly governed by a diverse array of regulatory mechanisms. The development of high-throughput technologies (see section "High-Throughput Techniques to Study Molecular Landscapes") has greatly advanced the ability to explore the transcriptome in the developing human CNS in order to understand the diverse cell and molecular phenotypes that underlie CNS function and disease processes.

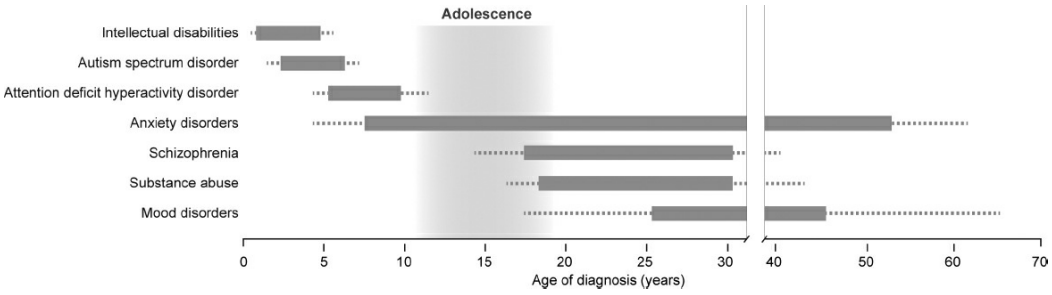

FIGURE 1.30–1. Psychiatric and neurodevelopmental disorders have discrete ages of onset. An illustration detailing the ages of onset of major psychiatric and neurodevelopmental disorders. The bars indicate the age range that each disorder commonly affects, with less frequent ages of diagnosis denoted as dotted lines. Note that the age of diagnosis is highly variable between brain disorders. Many psychiatric disorders emerge in adolescence and early adulthood, while neurodevelopmental disorders emerge earlier. This variability is indicative of dysregulation of tightly regulated developmental processes. (Based in part on data from Kessler RC, Amminger GP, Aguilar-Gaxiola S, Alonso J, Lee S, Ustun TB. Age of onset of mental disorders: a review of recent literature. *Curr Opin Psychiatry*. 2007;20:359–364; Silbereis John C, Pochareddy S, Zhu Y, Li M, Sestan N. The cellular and molecular landscapes of the developing human central nervous system. *Neuron*. 2016;89:248–268.)

Organization of the Human Nuclear Genome

The human nuclear genome comprises approximately 6.2 (in males) or 6.4 (in females) billion base pairs distributed across 46 chromosomes and provides the blueprint for building the body and its function. Current GENCODE estimates (/www.gencodegenes.org/stats/current.html; version 23) predict 60,498 genes. Of these, 19,881 are protein-coding genes, 25,813 are noncoding RNA (long and short) genes and the remaining are pseudogenes. A significant part of the remaining genome is believed to be *cis*-regulatory elements (CREs), the DNA sequences that regulate transcription. According to some estimates, the human nuclear genome harbors approximately 400,000 putative enhancers and 70,000 promoters. Access to CREs by transcription factors and other components of the transcriptional machinery is regulated by epigenetic modification of both DNA and chromatin. Most notably, methylation of DNA generally represses gene expression by blocking DNA binding, while acetylation of histones promotes transcription by opening up the chromatin to make DNA more accessible. This highlights both the complexity of the human genome and the challenges in elucidating the molecular mechanisms

underlying complex processes like neurodevelopment.

High-Throughput Techniques to Study Molecular Landscapes

Several major techniques have been in the brain applied over the past decade to study the human genome and its functions at a global scale. Many of these techniques rely on either hybridization or deep sequencing to identify and quantify the molecular species (RNA, DNA) tested. Some of the popular high-throughput techniques applied to investigate gene expression and its regulation at a genome-wide scale are discussed below:

Exon arrays – A microarray platform to quantify RNA expression at the exon level through hybridization to sequences of exons for known genes.

RNA-seq – Next-generation sequencing of RNA to identify and quantify the various RNA species, including coding and noncoding RNAs in a given sample. Allows unbiased detection of transcripts thus leading to identification of novel transcripts.

ChIP-seq – Chromatin immunoprecipitation (ChIP) followed by next-generation sequencing identifies regions of genomic DNA at which a particular protein of interest is interacting. Typically, DNA interacting with the protein of interest (e.g., transcription factor or histone mark) is precipitated along with the protein by using an antibody specific to the protein of interest. The precipitated DNA is then sequenced to identify the genomic sequences. This technique has been applied to identify genomic regions enriched for certain histone modifications or bound by specific transcription factors.

Methylation arrays – Microarrays that quantify DNA methylation levels at an array of regulatory sites by selective hybridization of methylated DNA that utilizes a reaction with bisulfate that marks methylated but not unmethylated DNA by converting methylated cytosine to uracil.

WGBS – Whole genome bisulfite sequencing is a sequencing based approach to determine DNA methylation levels on a genome-wide scale.

Spatial–Temporal Transcriptional Dynamics of Human Brain Development

A number of recent genome-wide profiling studies of the developmental transcriptome (see Silbereis et al., for comprehensive review) have produced comprehensive datasets describing coding and noncoding RNAs across multiple time points and regions of the human brain. These analyses have shed new light on the transcriptional architecture of brain regions and neocortical areas, neurodevelopmental processes, and the underlying biology of complex neurodevelopmental disorders. This work indicates that most protein coding genes (at least 86 percent according to Kang et al., 2011) and a growing number of noncoding genes (see section “Organization of the Human Nuclear Genome”) are involved in building the human CNS

at some point in development. Of the protein coding genes expressed in the brain, about 90 percent were either differentially expressed or spliced across brain regions and/or time. Such differences in the transcriptional architecture of brain regions and neocortical areas reflect their distinct ontologies, cellular compositions, biological processes, and developmental timing (Fig. 1.30–2).

Another common finding of these studies is that transcriptomes differ more prominently across time and space than they do between sexes, ethnicities, or individuals, despite the underlying genetic differences. In particular, remarkable dynamicity of gene expression is observed during prenatal and early postnatal development (4 postconceptual weeks through approximately 6 months of age), accounting for approximately two-thirds of the variance in global expression. In contrast, expression changes over several decades of adulthood are relatively minute (Fig. 1.30–2A). When compared with dorsal pallial structures (amygdala, hippocampus and neocortex), the cerebellum is the most transcriptionally distinctive region in the developing and adult brain, followed by thalamus and striatum (Fig. 1.30–2B). The transcriptional differences in primordial areas of the developing neocortex are less robust than those between neocortex and the other analyzed brain regions. A majority of these spatially enriched genes are also temporally dynamic; notably, some are transiently enriched during a specific time window, including protracted epochs of postnatal development such as late childhood and adolescence. These findings reveal that regional transcriptomes are developmentally regulated and may reflect cellular and functional differences.

Transcriptional studies have also demonstrated correlations between gene expression dynamics and the morphological and functional development of brain regions, and consequently shed light on the timing of developmental processes and the onset of specific biological functions (Fig. 1.30–2C). For example, gene coexpression network analyses have revealed that the developmental brain transcriptome is segregated into distinct modules or clusters of genes with highly correlated expression. The expression trajectories of such modules often exhibit dynamic patterns across development, indicating that they play roles in distinct and dynamic developmental processes. For example, in a study by Kang et al., 2011, the two largest modules were composed of genes that are coregulated across brain regions and time, indicating that many transcriptional processes are shared and coordinated across different regions as the brain develops. Many modules could be defined by functional ontologies, as they were highly enriched for genes important for specific biological processes.

Looking broadly at the expression levels of genes implicated in major developmental processes by functional ontology analysis sheds light on the molecular dynamics of human brain maturation and the timing of key neurodevelopmental events across areas. For instance, in the prefrontal

cortex, genes promoting cell proliferation are most highly expressed embryonically and expression declines markedly during the fetal period; genes associated with neuronal differentiation peak in the fetal period and are downregulated in infancy; and genes important for myelination rise markedly during perinatal development, plateauing around 1 year of age (Fig. 1.30–2C). Taken together, this recent transcriptomic analysis of human brain development across the spatial, temporal, and cellular dimensions has provided novel insights into normal neurodevelopmental processes and the interpretation of disease-associated mutations and underlying neuropathology, as discussed later in this chapter.

Transcriptional Organization of the Fetal Neocortex

These gene expression differences play essential roles in establishing the complex anatomical, circuit, and functional organization of the CNS, as can be exemplified in the development of the cerebral neocortex. At least two types of spatial information must be encoded in nascent projection neurons of the cortical plate (CP). First, their position in the radial direction, corresponding to their laminar position, with distinct cell types, synaptic connectivities, and functions within each lamina or layer; and second, their position in the tangential plane, corresponding to their particular cortical region/area identity. The human cerebral neocortex exhibits the basic organizing principle common to all mammals; however, there are human-enriched and defining features that are largely responsible for differences in cognition, emotion, behavior, and manifestations of disease. Spatiotemporal gene expression patterns in distinct zones of the fetal neocortical wall and different brain regions and neocortical areas have been characterized in several recent studies.

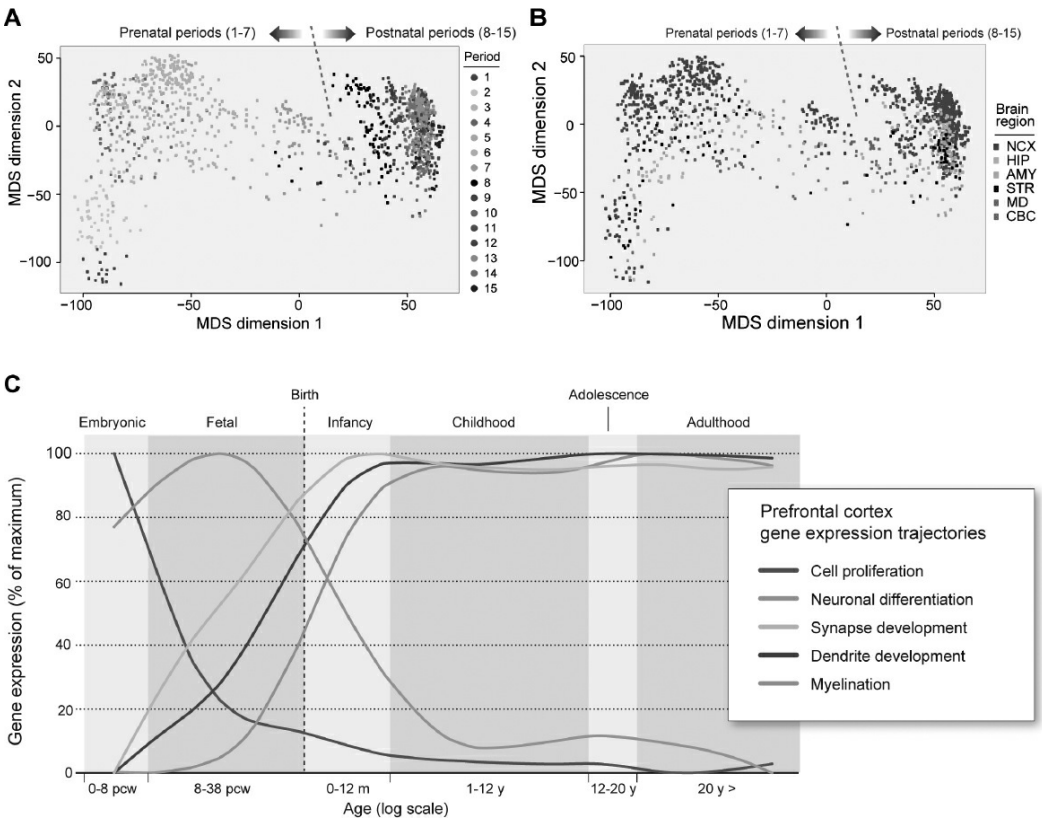

FIGURE 1.30–2. Global spatiotemporal dynamics of the human brain transcriptome. **A:** This panel shows a multidimensional scaling (MDS) plot that describes gene expression differences between developmental periods. MDS is statistical/graphical technique that reduces the complexity of data by plotting individual data points apart from one another at distances corresponding to the degree to which they are different. In the plot, each dot represents a sample and is colored according to period as defined by Kang et al., 2011; the dotted line indicates birth. Note that the most pronounced differences (approximately two-thirds) occur during prenatal development (periods 1 to 7). By contrast, over four decades of adulthood (periods 13–15), less than 1 percent of genes are differentially expressed. **B:** Here the same MDS plot as in (**A**) is shown except that the new data points are colored by brain region. Note that the most prominent transcriptional differences between regions are between cerebellar cortex (CBC) and forebrain regions. NCX, neocortex; HIP, hippocampus; AMY, amygdala; STR, striatum; MD, mediodorsal nucleus of thalamus. **C:** A line graph illustrating the developmental trajectory of expression of genes important for major neurodevelopmental processes demonstrating how gene expression profiles are linked to the cellular processes that take place at different periods of development. The y-axis represents the percentage of gene expression relative to maximum and the x-axis corresponds to age. (See color atlas.)

Most of the neurons of the cortex are generated from the late embryonic period to the late fetal period, as is detailed in depth in Chapter 1.3. During this period, the neocortical wall becomes segregated into laminar zones, each composed of characteristic cell compositions carrying out discrete neurodevelopmental processes. These lamina exhibits highly distinct gene expression patterns that define their biological roles. For instance, in early and midfetal development genes associated with proliferation and stem cell identity are enriched in the ventricular and subventricular zones (VZ and SVZ), which are germinal lamina overlying the ventricles. In contrast, the

CP, which is the primordium for the cerebral cortex, is characterized by high expression of genes promoting neuronal differentiation and synaptogenesis.

The CP will ultimately form the mature neocortex, which is a six-layered structure with the upper layers (e.g., through 4) sending intracerebral projections, and the lower layers 5 and 6 sending projections to the subcerebrum, brain stem, spinal cord, and other cortical areas, as well as, in the case of layer 6, reciprocal connections to the thalamus. Each of these layers is also characterized by distinct gene expression patterns, which are important for conveying the unique projection identities, synaptic receptivities, and functional properties within them. Defining these layers requires the establishment of spatially restricted gene expression patterns through transcriptional regulation, which shall be described later in this chapter.

Transcriptional Patterns Define Neocortical Regions and Areas in Development

While the neocortex has a highly specialized laminar organization, its organization into functionally discrete regions and areas along the tangential dimension is arguably far more complex, especially in humans. Work in mouse shows this is governed by graded expression of transcription factors during early embryonic development, followed by extrinsic signaling from thalamic axonal inputs. However, the molecular processes underlying the regional/areal compartmentalization of the neocortical CP remain poorly understood, particularly in humans. Recent transcriptome studies have revealed substantial regional/areal gene expression differences in the developing human neocortex (Fig. 1.30–3). Indeed, each region/area exhibits discrete temporally specified transcriptional properties that are likely reflective of underlying biological processes involved in the patterning and specialization of neural circuits encompassing particular regions and areas (Fig. 1.30–3A). Interestingly, these regional/areal transcriptional profiles were largely symmetric between hemispheres at the population level from fetal development through adulthood. This is surprising given the functionally distinct roles and different developmental trajectories of the left and right hemispheres, which arise in the late midfetal period and become more prominent during early postnatal periods. These findings therefore suggest that transcriptional neocortical left-right asymmetry is either too subtle for detection by tissue-level transcriptomic methods or that nontranscriptional mechanisms, such as neural activity, play more critical roles.

Interareal transcriptional differences were found to be most robust during early and midfetal development (Fig. 1.30–3B, C. During these periods, there were prominent and specific transcriptional signatures associated with prospective prefrontal and perisylvian areas, which are involved in some of the most distinctly human aspects of cognition and

behavior. In addition, prefrontal/frontal-enriched gradients of expression along the anteroposterior axis of the CP, as well as gradients with enrichment in temporal, occipital, occipitotemporal, perisylvian, and ventromedial areas, were observed. These strong early and midfetal interareal transcriptional differences and gradients diminish perinatally and increase again after adolescence (Fig. 1.30–3BC). The authors hypothesize that the greater interareal differences observed during early and midfetal development may in part reflect distinct transcriptional programs that are necessary to establish area-specific patterns of connectivity. By contrast, in late fetal periods and early postnatal development, perhaps more general molecular programs governing glial development, synaptogenesis, synaptic plasticity, and dendritic maturation predominate. The progressive increase in regional differences that emerge in adolescence could be reflective of variations in the maturational trajectories between certain neocortical areas (e.g., associative vs. primary areas) and may be important for late cognitive, emotional, and sociobehavioral development.

Taken together, these findings indicate that, similar to other species, the generation and differentiation of neocortical cell types and their assembly into functional neural circuits in humans are achieved through precise regulation of spatiotemporal gene expression. Psychiatric disorders and neurodevelopmental disorders are associated with both selective dysfunction of specific brain regions and neocortical areas and abnormal functional and anatomical connectivity between them. It is therefore intriguing to speculate that the distinct etiologies of psychiatric disorders are defined by abnormal gene expression at specific times and places in the developing brain.

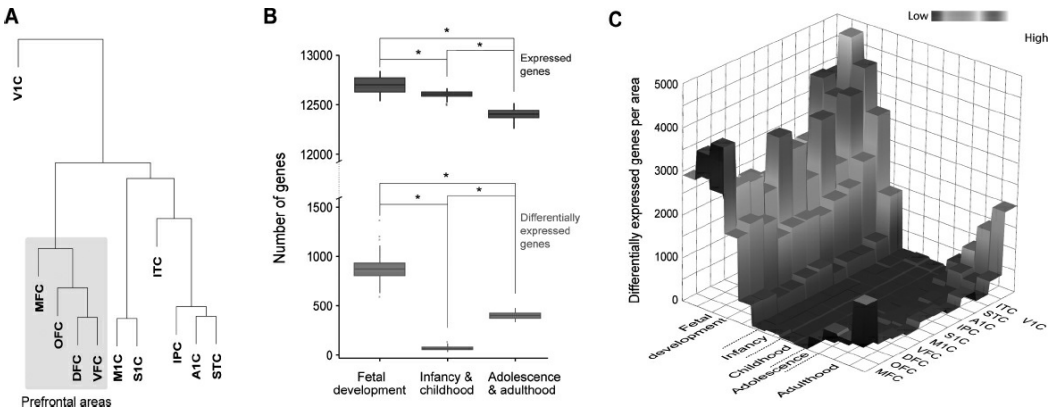

FIGURE 1.30–3. Transcriptional differences among human neocortical areas are temporally regulated. **A:** Illustration of the relative transcriptional differences between brain areas in an unsupervised hierarchical clustering chart that arranges the 11 neocortical regions/areas profiled by Pletikos et al., 2014, based on the similarity of the transcriptome of each area from early fetal development throughout adulthood. OFC, orbital prefrontal cortex; DFC, dorsal prefrontal cortex; VFC, ventral prefrontal cortex; MFC, medial prefrontal cortex; M1C, primary motor cortex; S1C, primary somatosensory cortex; IPC, posterior inferior parietal cortex; A1C, primary auditory cortex; STC, superior temporal cortex; ITC,

inferior temporal cortex; V1C, primary visual cortex. **B:** Boxplots showing the number of expressed (*dark gray*, top panel) and differentially expressed (*light gray*, bottom panel) genes among neocortical areas across fetal development (periods 3 to 7), infancy (periods 8 and 9), childhood (periods 10 and 11), adolescence (period 12), and adulthood (periods 13 to 15). Note that the total number of genes expressed decreases over development. However, the number of differentially expressed genes observed over development exhibits a temporal hourglass pattern with the most substantial differential expression occurring in fetal development, a marked decline in infancy and childhood, and an increase in adolescence through adulthood. **C:** A three-dimensional Manhattan plot showing the number (post hoc Tukey test) of genes with differential expression between any two neocortical areas, demonstrating that the hourglass pattern of interareal differential expression persists in all neocortical areas, but is most prominent in MFC and V1C. Differential expression corresponds to the y-axis, the x-axis delineates developmental periods, and the z-axis delineates brain region.

THE REGULATORY LANDSCAPE OF THE DEVELOPING HUMAN CNS

The spatiotemporal patterns of gene expression discussed above are established by elaborate regulatory mechanisms that govern each step of transcription. The current understanding of the general properties of transcriptional regulation of the developing human CNS is discussed below.

Regulation of Gene Transcription

A gene is transcribed when core transcriptional machinery including RNA polymerase and other proteins is recruited to the gene's proximal core promoter, a minimal region of contiguous DNA sequence that is sufficient to accurately initiate a basal level of gene expression. But the tissue specificity, timing, and level of a gene's expression are greatly controlled by distal DNA sequences like enhancers (sequences that increase expression), silencers (sequences that repress expression), and insulators (sequences that act as barriers), which along with core promoters are referred to as "cis-regulatory elements" (CREs) for their ability to mediate transcriptional activity in cis (i.e., proximal regions of the genome). Enhancers are believed to activate transcription by recruiting important accessory proteins to the promoter. They have binding sites for multiple transacting regulatory proteins, primarily transcription factors, and can function on their target promoters from locations as remote as another chromosome. Bound transcription factors recruit other transcription factors or chromatin modifying enzymes and interact with mediator complexes. One model of activation proposes that enhancer bound transcription factors along with mediator complex proteins loop the intervening DNA and bring enhancers and core promoters in close proximity, leading to activation of gene expression. Like enhancers, silencers or repressors recruit proteins but these proteins inhibit transcription. Silencer proteins act by inhibiting transcription initiation or elongation, or by sequestering activator proteins or altering chromatin state. Insulators also bind proteins, and these DNA–protein complexes act as barriers in the genome, blocking the action of an enhancer/silencer on its noncognate promoter or preventing the spread of

heterochromatin.

cis-Regulation of Transcription in Neocortical Development

Comparative genomic studies, which assess sequence conservation across different organisms, have shown that many of the regulatory networks involved in neurodevelopment are highly conserved across mammals. For instance, the regulation of *FEZF2*, a gene that is necessary for corticospinal (CS) tract development and that has remarkably similar neocortical expression patterns between mouse and human, is highly similar in these species. Furthermore, studies of *Fezf2* regulation in mouse neocortical development exemplify key concepts that describe developmental regulatory networks.

The *Fezf2* gene encodes a transcription factor and the formation of the CS tract is disrupted when *Fezf2* is knocked out in mutant mice. *Fezf2* expression is highly enriched in the layer projection neurons of the cerebral cortex that give rise to the CS tract. This pattern suggests that the expression has to be activated in L5 and inhibited in L6 and other layers of the cortex. The gene regulatory network that accomplishes this selective expression has recently been defined (Fig. 1.30–4) and provides a model of spatial control of gene expression. In Shim et al., four different putative enhancers (E1 to E4) of *Fezf2* were identified (Fig. 1.30–4A) and their individual roles in *Fezf2* regulation were tested (Fig. 1.30–4B,C). Of these CREs, E4, which lies downstream of the Fezf2 transcription start site, was identified to be an enhancer (see Figure 1.30–4) essential for *Fezf2* expression in the developing neocortical plate and for specification of the CS tract. Two related transcription factors of the SOX family, SOX4 and SOX11, were shown to be essential for FEZF2 expression by binding and activating the E4 enhancer in the cortex. SOX5, another member of the SOX family of transcription factors, also binds E4 but represses FEZF2 expression. It was observed that SOX4 and SOX11 functionally compete with SOX5 to determine if *Fezf2* transcription is activated or repressed. The spatiotemporal dynamics of neocortical FEZF2 expression are further controlled by TBR1, another transcription factor expressed in L6 neurons. TBR1 binds to a different enhancer in a highly conserved region downstream of the *Fezf2* transcription start site and represses the expression of FEZF2 in L6. This process is indicative of a common principle of gene regulatory networks in which competitive binding of transcriptional activators and repressors in key regulatory elements form gradients that determine the areas and levels of gene expression.

Transcriptional Regulation of Human Neurodevelopment

Although studies of individual CREs in model organisms have provided great advances in the current understanding of neurodevelopment, it is also

highly evident that comprehensive understanding of gene regulation in human neurodevelopment requires a catalogue of all CREs that are active in humans. Based on conservation of genomic sequence during evolution, many putative CREs have been identified, a large number of which are proposed to regulate expression of genes enriched in human brain. Conversely, DNA sequences with dramatically increased substitution rates in the human lineage compared to other organisms have also been used to predict CREs in humans.

Similar to studying the transcriptome, the development of high-throughput technologies (see section "High-Throughput Techniques to Study Molecular Landscapes") has greatly improved the ability to detect CREs in the human brain. Traditionally, chromatin accessibility and nucleosome depletion and, more recently, enrichment for specific posttranslational histone modifications have been considered to be among the defining characteristics of CREs. High-throughput DNAase hypersensitivity and ChIP assays followed by deep sequencing (ChIP-seq) have been applied in postmortem human brain tissue. Visel et al. determined the genome-wide occupancy of the enhancer-associated proteins p300/CBP in human fetal neocortex to identify active enhancers. By mapping enrichment of two enhancer-enriched posttranslational modifications of histones, histone 3 lysine 4 monomethylation (H3K4me1) and histone 3 lysine 27 acetylation (H3K27ac), enhancers active in different brain regions and at different time points were deciphered. These studies have identified region-specific enhancers and enhancers that have been gained in humans. In addition, coregulated enhancers that form cell-type and context-specific networks have been identified. Many of these enhancer networks are correlated with modules of coexpressed genes that function in neuronal proliferation, migration, and cortical-map organization. Thus, the regulatory networks that govern spatial–temporal gene expression in the developing human CNS are beginning to be elucidated; however, this information has yet to be fully placed into a functional context.

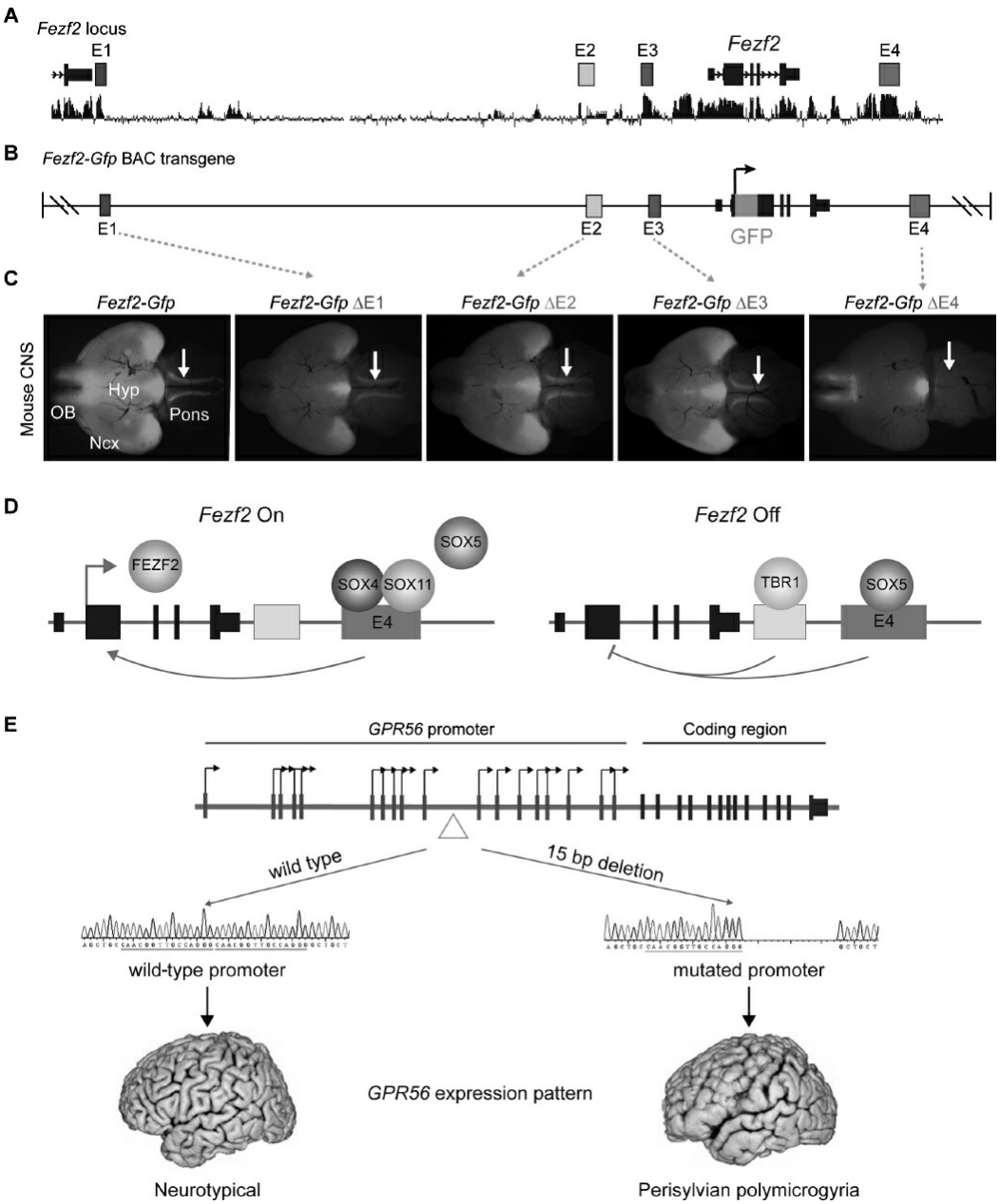

FIGURE 1.30–4. cis-Regulatory control of gene expression in brain development and disease. **A:** Schematic of the *FezF2* gene locus, including its regulatory elements, demonstrating interspecies conservation. The top panel schematizes the FezF2 locus in mouse and the position of its regulatory enhancer elements. The bottom panel plots the level of conservation between species; note that enhancers and other regulatory elements are typically far more conserved than other noncoding regions. **B:** Schematic showing the construction of a bacterial artificial chromosome (BAC) that was used to drive green fluorescent protein expression (FezF2-GFP) from specific FezF2 enhancers in transgenic mice, as shown in (**C**). **C:** Micrographs of GFP expression in transgenic lines in which individual enhancers were deleted. The E4 enhancer was essential for the formation of the cortical spinal tract (see *arrows*). This illustrates how multiple regulatory elements interact to determine spatial gene expression patterns (micrographs from Shim et al. 2014). **D:** Schematic describing the cis-regulatory mechanism that drives FezF2 expression. When SOX4 and SOX11 bind the E4 enhancer FezF2, it is highly expressed (left panel). When SOX5 binds the E4 enhancers and TBR1 binds an adjacent regulatory element, FezF2 expression is repressed (right panel). **E:** Perisylvian microgyria as an example of a how mutation of an enhancer element can lead to a developmental disorder. In this disease, deletion of one of 17 transcriptional start sites in the human promoter for GPR56, a gene important for

driving proliferation of cerebral cortical progenitor cells, leads to reduced expression of GPR56 in the primordium of the perisylvian neocortex, resulting in polymicrogyria specifically in this region (adapted from Bae et al. 2015).

EPIGENETIC REGULATION OF HUMAN BRAIN DEVELOPMENT

Epigenetics refers to the chemical modification of DNA, chromatin structure, and RNA that can modulate gene expression without changing the DNA sequence. Epigenetic states are dynamic over neurodevelopment, heavily influenced by the environment, and can be heritable. Therefore, epigenetic regulation is a mechanism by which genes and the environment interact to convey susceptibility and heritability to psychiatric disorders. The two most commonly studied epigenetic mechanisms are DNA methylation, which generally represses gene expression by inhibiting transcription factors from binding DNA, and posttranslational histone modifications, which alter chromatic structure to make it either more or less accessible to transcriptional machinery.

DNA methylation of cytosine, primarily at CpG nucleotides, which are enriched in gene promoters, plays a key role in neural development and function and can be profiled at hundreds of thousands to millions of CpGs in the human genome in an unbiased manner. A widespread methylome reconfiguration during fetal to young adult development has been observed by multiple studies. Global DNA methylation in the human prefrontal cortex was shown to be greatly reconfigured during fetal-to-postnatal development. These changes were more rapid during fetal development and slowed down after birth and with aging. Many of these methylation changes correlate significantly with the changing cellular composition of the human cortex during development, highlighting the important role of DNA methylation in human neurodevelopment.

In addition to methylation CpGs, accumulation of non-CpG methylation (mCH) during the early postnatal developmental period that coincides with the primary phase of synaptogenesis has recently been discovered in the human brain. This accumulation exhibits cell-type specificity, with abundant mCH in neurons and insignificant levels in glial cells. For both CpG and mCH methylation, a strong correlation with the expression of adjacent genes was observed. Another key epigenetic mechanism for establishing cellular diversity and regulating the timing of neurodevelopment is the posttranslational modification of histones. Methylation (mono-, di-, and tri-) and acetylation are the most extensively studied modifications, although many more have been observed. Studies have generated chromatin state maps in a limited number of postmortem human brains by applying ChIP coupled with next-generation sequencing (ChIP-seq). Cheung et al. profiled histone 3 lysine 4 trimethylation (H3K4me3) across development in neuronal and nonneuronal cells of the human prefrontal cortex of 11 individuals ranging in age from 0.5 to 69

postnatal years. A key finding of their study was significant remodeling during postnatal development and aging of prefrontal cerebral cortical neurons. Taken together, these data confirm that epigenetic mechanisms play a key role in establishing the regional and temporal landscape of gene expression and contributed to human brain evolution. It is intriguing to speculate that increased DNA methylation in infancy and childhood may lead to the reduction in interareal differences in gene expression and the increase in the expression of genes associated with synapse formation/regulation observed during these periods.

TRANSCRIPTIONAL AND EPIGENETIC DYSREGULATION IN NEURODEVELOPMENTAL AND PSYCHIATRIC DISORDERS

It is becoming increasingly apparent that psychiatric disease is rooted in genetic and environmental perturbations of neurodevelopment. Over the last few years, genetics research has made significant strides in identifying genetic risk factors for psychiatric disorders. Most psychiatric disorders are highly complex, with multiple genes contributing to disease risk. There is very strong evidence that the genetic factors involved in psychiatric disease comprise both common and rare variants (SNPs) and submicroscopic variations in chromosomal structure (i.e., CNVs). For example, in schizophrenia and bipolar disorder, gene discovery has centered on linkage between these disorders with common SNPs, while in ASD, studies of rare and de novo single nucleotide variants have been the most productive. Even when a disease is manifest by a single causative gene, as in many neural developmental disorders (e.g., fragile X syndrome [FXS]), there is broad dysregulation of neurodevelopmental processes that ultimately lead to disease phenotypes independent of the mutated gene's direct biological roles. The polygenicity of these disorders makes identifying the molecular mechanisms underlying the pathobiology of psychiatric diseases difficult. Further complicating the ability to understand genetic bases of disease is that, in addition to the above-mentioned germline mutations, somatic mutations—de novo mutations occurring postfertilization and often in specific cell and tissue types—can contribute to psychiatric and neurodevelopmental disorders.

Genetics and complex polygenic risk architecture do not completely explain susceptibility to many psychiatric diseases. Indeed, for many disorders, the possibility of two identical twins developing the disease is lower than the genetic risk, indicating that environmental factors influence disease susceptibility. For example, environmental factors such as obstetric complications have been reported to increase the risk of developing schizophrenia. Aversive stimuli from the environment have broad effects on the cellular processes in the developing brain, ranging from altered gene expression levels to dysregulated neural activity, which then interact in neural development to convey increased susceptibility to psychiatric

problems. Moreover, adverse events in development can lead to epigenetic changes that might have permanent effects on gene expression.

Despite this etiologic complexity, the neurodevelopmental disruptions underlying psychiatric diseases must logically converge on distinct and therefore definable cell and molecular processes, brain circuits, and epochs of development. Resolving these processes conveys a better understanding of the underlying pathology, and may lead to more targeted therapeutic strategies. As described in several examples below, increasingly comprehensive studies of the cell and molecular biology of neurotypical human neurodevelopment can be integrated with studies of psychiatric disease, genetics, and pathology to interrogate disease mechanisms and identify the time and place where disrupted molecular processes are likely to act.

Fragile X Syndrome: Single Gene Mutations Can Have Broad Downstream Consequences

FXS is the leading inherited cause of intellectual disability and is often accompanied by autistic-like features, aggression, attention deficits, and delays in speech and language development. FXS is caused by loss of function of the FMR1 gene, which encodes an RNA-binding protein (FMRP) involved in mRNA localization, stability, and translation. Therefore, loss of FMRP function has broad downstream consequences on gene expression and translation. For example, in a recent study, FMRP was shown to interact with the human neuronal *nitric oxide synthase 1* (*NOS1*) *gene* through both conserved and species-dependent processes. *In the cerebral cortex,* NOS1 protein expression is restricted to layer 5 (L5) subcerebrally projecting pyramidal neurons and exhibits expression in alternating minicolumns in the frontoparietal operculum. In humans, FMRP interacts with sequences in the *NOS1*-coding region and leads to increased NOS1 protein expression. Concordantly, neocortical NOS1 protein levels are dramatically reduced in the developing FXS human neocortex, where NOS1 and FMRP are transiently coexpressed during synaptogenesis in subpopulations of pyramidal neurons in regions involved in speech, language, and complex social behaviors.

Down Syndrome and ASD: Polygenic Disruptions Converging on Specific Developmental Processes

Down syndrome (DS) provides an interesting neurodevelopmental disorder for studying the biological consequences of polygenic, complex disorders, as the source of polygenic disruption (trisomy 21) is well established. Studies have indicated that trisomy 21 causes genome-wide disruption of gene expression. Among these, a characterization of the DS developmental brain transcriptome found that approximately 5 percent of genes on all

chromosomes were dysregulated at some point in development, with a roughly equal number of up- and downregulated genes, with the exception of 18 percent of Chr21 genes, all of which were upregulated (Fig. 1.30–5). Gene coexpression analysis revealed gene networks that were implicated in processes previously associated with DS (e.g., synaptic function and neuronal specification) and aberrations in genes for particular cellular and molecular processes that had not been well established in DS, such as altered expression of genes associated with myelination and oligodendrocyte development. Subsequent analysis of postmortem human DS brain specimens revealed prominent hypomyelination, disorganization of myelinated fibers, and slower conductance through myelinated fiber tracts as features of DS neuropathology.

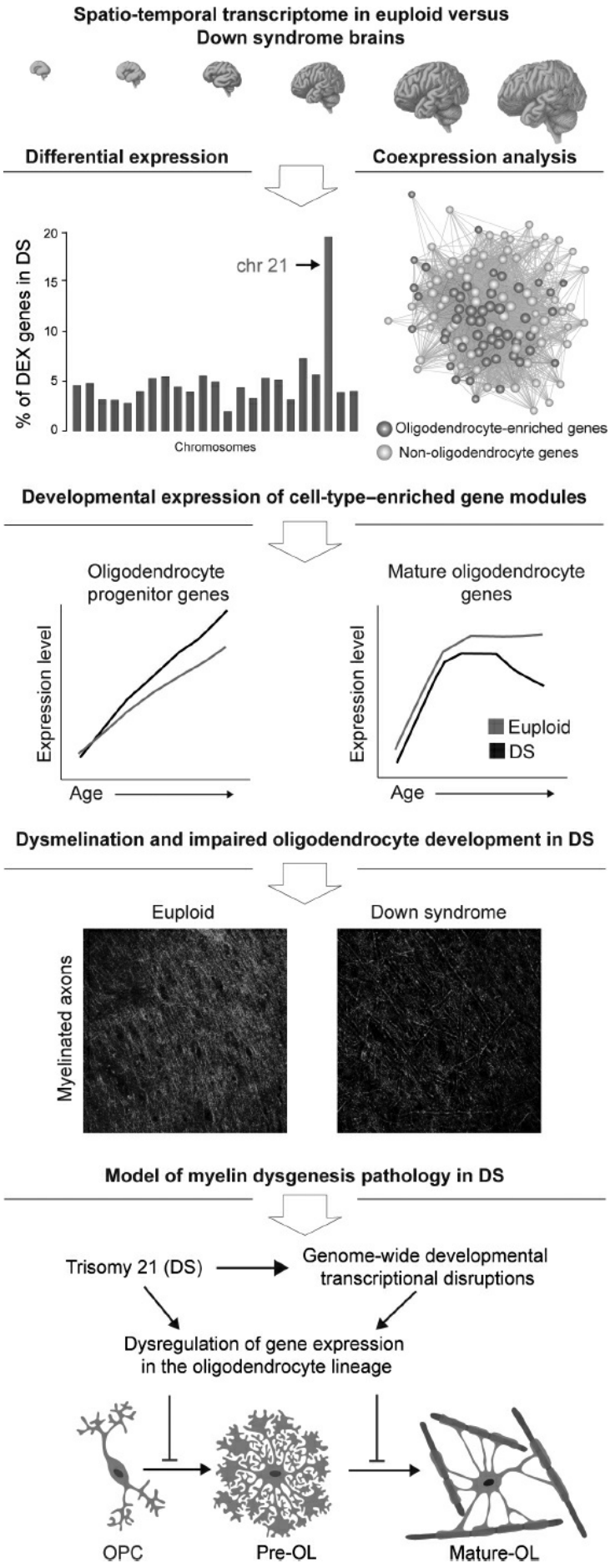
Spatio-temporal transcriptome in euploid versus Down syndrome brains
Differential expression
Coexpression analysis
% of DEX genes in DS
20
15
10
5
0
chr 21
Chromosomes
Oligodendrocyte-enriched genes
Non-oligodendrocyte genes
Developmental expression of cell-type–enriched gene modules
Oligodendrocyte progenitor genes
Mature oligodendrocyte genes
Expression level
Age
Euploid
DS
Dysmelination and impaired oligodendrocyte development in DS
Euploid
Down syndrome
Myelinated axons
Model of myelin dysgenesis pathology in DS
Trisomy 21 (DS)
Genome-wide developmental transcriptional disruptions
Dysregulation of gene expression in the oligodendrocyte lineage
OPC
Pre-OL
Mature-OL

FIGURE 1.30–5. Genome-wide dysregulation in Down syndrome (DS) leads to impaired oligodendrocyte development. Schematic of a study by Olmos-Serrano et al. showing how developmental transcriptome characterizations of DS (Trisomy 21, DS) have provided insight into how polygenic gene disruptions lead to far broader impacts on gene expression and ultimately cell and molecular function. In this study, global gene expression was characterized from midfetal development to adulthood in several brain regions of individuals with DS (first panel). It was determined that approximately 20 percent of genes of chromosome 21 genes were misexpressed at some point in development, while 3 to 8 percent of the genes on all other chromosomes were misexpressed. Thus, in total, far more nonchromosome 21 genes are misexpressed in DS (second panel). Network analysis of coexpressed genes revealed a module that was highly enriched in myelination genes. Subsequent analysis of gene networks enriched in genes highly expressed in either oligodendrocyte precursors or mature oligodendrocytes revealed substantial dysregulation of developmental gene expression dynamics in the oligodendrocyte lineage: progenitor genes were more highly expressed and mature genes were expressed at lower levels (third panel). Analysis of myelinated fiber tracts in the prefrontal cortex revealed hypomyelination and disorganization of myelinated axons in DS (fourth panel). This study suggests that trisomy of chromosome 21 leads to genome-wide disruptions in developmental gene regulation, which were particularly prominent in oligodendrocytes and resulted in impaired oligodendrocyte maturation and hypomyelination (fifth panel).

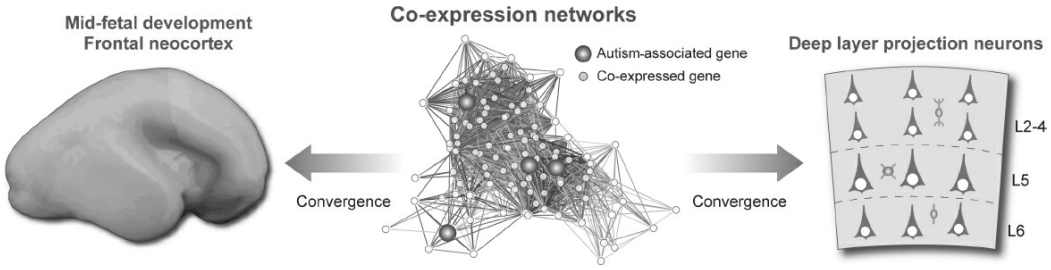

FIGURE 1.30–6. Dysregulation of autism spectrum disorder (ASD)-implicated genes converge on gene networks enriched in midfetal cortical projection neurons. This figure illustrates how a coexpression network analysis of ASD-associated genes with a spatiotemporal human transcriptome database inferred a convergence of autism-associated genes in frontal neocortex and deep layer projection neurons, suggesting that abnormalities in this cell type may play a role in ASD pathogenesis. ASD-associated genes identified in genome-wide sequencing studies were used as seed genes to build coexpression networks (middle panel). The networks of genes highly interconnected with ASD-associated genes were found to be enriched for genes highly expressed in the frontal cerebral cortex (left panel) and deep layer projection neurons (right panel). (Adapted from Tebbenkamp AT, Willsey AJ, State MW, Sestan N. The developmental transcriptome of the human brain: implications for neurodevelopmental disorders. *Curr Opin Neurol*. 2014;27(2):149; Willsey AJ, Sanders SJ, Li M, et al. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell*. 2013;155(5):997–1007.)

Similarly, to interrogate ASD pathogenesis, recent studies have used computational approaches to integrate data on risk variants from GWAS with spatiotemporal brain transcriptome data. For instance, by building coexpression networks of genes previously associated with ASD, researchers have identified modules of highly related genes and determined the time in development and cell types in the brain where these gene networks are enriched. Intriguingly, it was found that ASD-risk genes were highly enriched in cortical projection neurons during midfetal development

(Fig. 1.30–6). These data indicate that, in ASD, development may be particularly disrupted in cortical projection neurons when they are in the process of terminal migration and beginning to elaborate dendrites and early axonal projections. Moreover, functional and anatomical imaging studies have revealed that brain circuits composed of these neurons are abnormal in both ASD and schizophrenia.

Noncoding Mutations in Neurodevelopmental Disorders

The majority of disease-associated sequence variants occur in noncoding regions of the genome. While the consequences of these mutations are difficult to study, many likely have developmental effects and represent risk factors for psychiatric disease. This is particularly the case for variants that lie within regulatory regions such as enhancers and promoters. An illustrative example is a recently described study of perisylvian polymicrogyria that results from a deletion immediately upstream of one of 17 transcriptional start sites (TSSs) of *GPR56*, a G-coupled protein receptor encoding gene that is important for stem cell proliferation and cerebral cortical development. In five affected individuals, a 15 base pair deletion was detected in a region of the GPR56 core promoter that is essential for driving expression in perisylvian regions of the brain during prenatal development. This deletion likely impaired expression of GPR56 specifically in perisylvian cortex during neurogenesis and led to reduced progenitor cell proliferation and microgryrification specifically in the perisylvian cortex (Fig. 1.30–4). While this example is extreme, it demonstrates how differences in noncoding sequences, particularly in regulatory elements, can convey different disease susceptibilities and neurodevelopmental phenotypes.

Somatic Mutations: Regionally Discrete Gene Mutations and Patterns of Gene Misexpression

Although somatic mutations in the brain cannot be readily detected using standard genomic screening techniques, recent advances in deep next-generation sequencing and single cell sequencing have enabled the detection of somatic mutations in specific cells and tissues and advanced the current understanding of their contribution to disease etiology. This work has broad implications, from uncovering new disease mechanisms to providing new insights into neurotypical development. For example, these studies have demonstrated that certain neurodevelopmental diseases may be caused by *de novo* somatic mutations in neural progenitors that then expand clonally to produce aberrations in specific neural cell types and brain regions that can lead to brain malformations and more subtle phenotypes. Recent studies suggest that ASD, and by extension risk for psychiatric diseases, results from a combination of germ line and somatic

mutations that converge on specific brain regions, cell types, and periods of development.

Epigenetic Dysregulation in Schizophrenia

Dysregulation of epigenetic mechanisms can lead to the transcriptome changes observed in disease states. Understanding the role of epigenetic modifications in psychiatric diseases is in its nascent phase, although potential roles for DNA methylation have been observed in some psychiatric diseases including depression, schizophrenia, and addiction. To date, the majority of such studies have analyzed changes in promoters of individual candidate genes. In schizophrenia, several lines of evidence suggest that reelin gene (*RELN*) expression, which encodes an extracellular matrix protein that plays critical roles in the developing and adult brain, is downregulated due to hypermethylation of its promoter. A similar observation linked the promoter of the glutamic acid decarboxylase 67 (GAD1) gene, a gene encoding an enzyme that is required for synthesis of GABA, with schizophrenia. In addition, decreased enrichment of H3K4me3 and an excess of H3K27me3, a repressive mark, has also been observed in the GAD1 promoter. Highlighting the potential importance of GAD1 dysregulation in schizophrenia are the well-established findings that GAD1 and GABA are not only expressed at lower levels in schizophrenia but also that mutations in GAD1 have been implicated in childhood-onset schizophrenia and that GAD autoantibody syndrome typically presents as schizophrenia.

Thus far, very few studies have explored possible roles of epigenetic mechanisms of psychiatric disease at a genome-wide scale. Upon integrating the DNA methylation differences observed in the normal brain development and risk loci associated with schizophrenia in a genome-wide association study, Jaffe et al. observed that global CpG methylation is different between neurotypical prenatal and postnatal samples. CpG sites that were hypermethylated prenatally were enriched for schizophrenia risk loci. However, a significant enrichment for risk was not observed at the age of onset of illness (around 25 years). A direct comparison of the DNA methylation profiles between control and schizophrenia adult prefrontal cortex identified approximately 2,100 CpG loci that were differentially methylated. Although these loci were not enriched in schizophrenia risk alleles, they again exhibited pre-and postnatal changes in neurotypical cortex. Together these findings suggest that epigenetic changes during development play a critical role in disease risk.

FUTURE DIRECTIONS

Neuropsychiatric disorders owe in part to broad dysregulation of the genome over the course of development. Transcriptome data suggest that,

in complex psychiatric and neurodevelopmental disorders, disruptions in many genes can converge on particular cell types and biological process to produce disease phenotypes. It is therefore intriguing to speculate that selective dysfunction of spatially and temporally regulated gene expression may in part explain differences in the age of onset and affected neural circuits in neurological and psychiatric disorders (Fig. 1.30–1). However, a great deal of work must be undertaken to further elucidate mechanisms of diseases, determine how genetic and somatic variations cause diseases, and develop more rational therapeutic strategies. Multidimensional genomic data, including more comprehensive transcriptomic, epigenomic, and regulatory maps of the developing human brain, will greatly advance the current understanding of the pathophysiology of complex diseases such as psychiatric disorders. These genomic data may one day reveal genetic and molecular phenotypes of diseases and disease subtypes that might be useful for diagnosis and more targeted treatment paradigms. In addition to generating these genomic data, integrative analyses with pathological, neuroimaging, behavioral, and clinical studies will be critical for a full understanding of psychiatric diseases.

References

Bae BI, Jayaraman D, Walsh CA. Genetic changes shaping the human brain. *Dev Cell*. 2015;32:423–434.

Bassell GJ, Warren ST. Fragile X syndrome: Loss of local mRNA regulation alters synaptic development and function. *Neuron*. 2008;60:201–214.

Dehaene-Lambertz G, Spelke ES. The infancy of the human brain. *Neuron*. 2015;88:93–109.

Geschwind DH, Rakic P. Cortical evolution: judge the brain by its cover. *Neuron*. 2013;80:633–647.

Han W, Kwan KY, Shim S, et al. TBR1 directly represses Fezf2 to control the laminar origin and development of the corticospinal tract. *Proc Natl Acad Sci U S A*. 2011;108:3041–3046.

Jaffe AE, Gao Y, Deep-Soboslay A, et al. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. *Nat Neurosci*. 2016;19:40–47.

Johnson MB, Kawasawa YI, Mason CE, et al. Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron*. 2009;62:494–509.

*Kang HJ, Kawasawa YI, Cheng F, et al. Spatio-temporal transcriptome of the human brain. *Nature*. 2011;478:483–489.

Kasprian G, Langs G, Brugger PC, et al. The prenatal origin of hemispheric asymmetry: an in utero Neuroimaging study. *Cereb Cortex*. 2011;21:1076–1083.

Kessler RC, Amminger GP, Aguilar-Gaxiola S, Alonso J, Lee S, Ustun TB. Age of onset of mental disorders: a review of recent literature. *Curr Opin Psychiatry*. 2007;20:359–364.

Kwan KY, Lam MM, Johnson MB, et al. Species-dependent posttranscriptional regulation of NOS1 by FMRP in the developing cerebral cortex. *Cell*. 2012;149:899–911.

Kwan KY, Lam MM, Krsnik Z, Kawasawa YI, Lefebvre V, Sestan N. SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer

neocortical neurons. *Proc Natl Acad Sci U S A*. 2008;105:16021–16026.

Kwan KY, Sestan N, Anton ES. Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. *Development*. 2012;139:1535–1546.

Lister R, Mukamel EA, Nery JR, et al. Global epigenomic reconfiguration during mammalian brain development. *Science*. 2013;341:1237905.

Miller JA, Ding SL, Sunkin SM, et al. Transcriptional landscape of the prenatal human brain. *Nature*. 2014;508:199–206.

O'Leary DD, Chou SJ, Sahara S. Area patterning of the mammalian cortex. *Neuron*. 2007;56:252–269.

Olmos-Serrano JL, Kang HJ, Tyler WA, et al. Down syndrome developmental brain transcriptome reveals defective oligodendrocyte differentiation and myelination. *Neuron*. 2016;89(6):1208–1222.

Parikshak NN, Luo R, Zhang A, et al. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*. 2013;155:1008–1021.

Pletikos M, Sousa AM, Sedmak G, et al. Temporal specification and bilaterality of human neocortical topographic gene expression. *Neuron*. 2014;81:321–332.

*Shibata M, Gulden FO, Sestan N. From trans to cis: transcriptional regulatory networks in neocortical development. *Trends Genet*. 2015;31:77–87.

Shim S, Kwan KY, Li M, Lefebvre V, Sestan N. Cis-regulatory control of corticospinal system development and evolution. *Nature*. 2012;486:74–79.

Shorter KR, Miller BH. Epigenetic mechanisms in schizophrenia. *Prog Biophy Mol Biol*. 2015;118:1–7.

*Silbereis John C, Pochareddy S, Zhu Y, Li M, Sestan N. The cellular and molecular landscapes of the developing human central nervous system. *Neuron*. 2016;89:248–268.

Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Gen*. 2012;13:537–551.

*Tebbenkamp AT, Willsey AJ, State MW, Sestan N. The developmental transcriptome of the human brain: implications for neurodevelopmental disorders. *Curr Opin Neurol*. 2014;27:149–156.

Visel A, Taher L, Girgis H, et al. A high-resolution enhancer atlas of the developing telencephalon. 2013;152:895–908.

Watson JD, Baker TA, Bell SP, Gann A, Levine M, Losick R. *Molecular Biology of the Gene*. San Francisco, CA: Benjamin Cummings; 2013.

Willsey AJ, Sanders SJ, Li M, et al. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell*. 2013;155:997–1007.

2

Neuropsychiatry and Behavioral Neurology

▲ 2.1 The Neuropsychiatric Approach to the Patient

FRED OVSIEW, M.D.

Psychiatry has eliminated the term "organic" from the official nomenclature, but the word persists in vernacular use because the care of patients with identifiable, acquired brain disease—such as those with epilepsy, movement disorders, and traumatic brain injury—requires the physician to have a knowledge base and a familiarity with assessment and treatment methods not usually required for patients with primary psychiatric disorders. Patients with organic mental syndromes are common in clinical practice and often difficult to manage for the general psychiatrist, even with consultation from other specialists who may themselves not be expert in the mental and emotional phenomena accompanying brain disease.

Neuropsychiatry is the psychiatric subspecialty that deals with the psychological and behavioral manifestations of brain disease. Neuropsychiatry is closely allied with cognitive and behavioral neurology, the neurological subspecialty that interests itself in psychological phenomena in patients with brain disease. In addition to expert management of patients with organic mental disorders, from its clinical vantage point neuropsychiatry can offer a distinctive perspective on idiopathic psychiatric disorders, although later in this section the limitations on the usefulness of this perspective will be noted.

A few preliminary words about the history of neuropsychiatry and its terminology will help to sketch the neuropsychiatric perspective. The seemingly obvious view that neuropsychiatry is the offspring of psychiatry and neurology is historically mistaken. Psychiatry differentiated itself as a medical specialty in the early part of the 19th century; neurology did so somewhat later. Evidence is ample to show that early asylum physicians, the precursors of psychiatrists, considered their patients to be suffering from brain diseases, and moreover that a large proportion of their patients evinced organic disease, even with the limited tools of that time. General

paresis of the insane (neurosyphilis, as it was later discovered to be), epilepsy, intellectual disablity, and the complications of alcohol abuse were all common in the 19th-century asylum. Based on this evidence, one might say that general psychiatry, as it was understood for the larger part of the 20th century, was derived from an earlier neuropsychiatry.

Early neurology, on the other hand, took little part in the care of patients with major psychiatric disorders, at least those requiring hospitalization; but what would later become outpatient psychiatry—the care of patients with milder mood and anxiety disorders not requiring asylum management, for example—fell into the province of the early neurologists. The theories by virtue of which they understood their patients have fortunately been consigned to the dustbin of history. The mainstream of Anglo-American neurology was ill-equipped to give rise to a scientific neuropsychiatry, and it was not until (as a convenient and meaningful landmark) Norman Geschwind in 1965 awakened interest in the continental tradition of a behavioral neurology *avant la lettre* that the contributions of John Hughlings-Jackson, Ludwig Lichtheim, Hugo Liepmann, Karl Wernicke, and others could provide impetus to the development of a clinical specialty devoted to scientific understanding of the cerebral basis of mental and behavioral disorder.

To say that neuropsychiatry is devoted to the care of patients with brain disease is not to depreciate the role of psychological and social factors in the understanding of the genesis of symptoms or in the formulation of interventions to assist patients. To the contrary, patients with brain disease are often inordinately reactive to or dependent on influences from the outside world, notably the social world. Neuropsychiatric case formulation takes into account both the vulnerability and the setting. To the extent that patients suffer from brain-based impairments in processing information from their environment, their need for assistance in dealing with instrumental and interpersonal tasks increases. Much of the brain, after all, is devoted to processing social information and devising ways of meeting internal needs in a social context. The neuropsychiatrist requires a detailed assessment of the patient's functional deficits and the contexts in which they arise.

To say that the neuropsychiatrist regards deficits, or for that matter intact behavior, as the manifestations of brain-based processes is not to imply that idiopathic psychiatric disorders occur in people with normal brains, nor that general psychiatrists are unaware of the cerebral origin of these disorders. To the contrary, the evidence for abnormal brain structure and function in the major psychiatric disorders is unmistakable, and general psychiatrists often assert the neurobiological nature of these illnesses. However, evidence for such assertions is often not demonstrable in the individual case, with all available laboratory investigations characteristically falling within the broad range of normal. Moreover, the

neurobiological abnormalities in question are believed to be, at least in large part and at least in most illnesses, genetic in nature and developmental in pathogenesis. The recognition and understanding of the mental consequences of acquired diseases of the brain, which form the bulk of the neuropsychiatrist's concern, are likely to require different tools from those required by the general psychiatrist treating idiopathic disorders. Although a bright line between the two situations is not possible, and although many neuropsychiatrists maintain a lively interest in such disorders as schizophrenia and autism, the distinction supports the continued use of the term "organic" to refer to these acquired disorders, with pathology identifiable at the bedside and by the clinical laboratory, as much as some would like to see the term interred.

NEUROPSYCHIATRIC NEUROANATOMY

The neuropsychiatric brain is more complex than the general psychiatric brain. The latter is a soup of neurotransmitters, perhaps in "chemical imbalance" (as patients are wont to say), with considerable pharmacological but little anatomic specificity. Although the benefits of psychopharmacological intervention are indisputable, the locus of these effects is rarely of concern to clinicians. A neuropsychiatric approach relies on greater differentiation among brain circuits and systems.

Lateralization

The two hemispheres differentially subserve many cerebral functions, although in many instances both hemispheres participate in naturally occurring behavior, albeit contributing differently to the complex outcome. Brain asymmetries arise early in vertebrate evolution, and the two hemispheres display regional lateral asymmetries in size and differentially innervate viscera and peripheral endocrine tissues. For example, the pars opercularis of the third frontal gyrus (Broca area) and the planum temporale (infolded cortex in the posterior portion of the sylvian fissure) are typically larger on the left, with greater dendritic branching of the neurons therein. (For simplicity sake, "left" and "right" here refer to the situation in the average dextral patient.) These cortical regions are parts of the substrate of language processing. Insular cortex of the right hemisphere regulates cardiac sympathetic drive, of the left hemisphere parasympathetic drive. In consequence, left hemisphere stroke involving insula produces more cardiac destabilization and morbidity than right, and lateralization of seizure discharges may have implications for autonomic function and unexplained sudden death in patients with epilepsy. Lateral differences in limbic (hypothalamic and amygdalar) regulation of sexual function also have clinical implications; for example, polycystic ovary syndrome in woman may be more commonly associated with left-sided

limbic epilepsy. Hemispheric side of lesion also affects the immunological consequences of brain injury.

Whether a single tag can accurately contrast the processing "styles" of the hemispheres across multiple functions—local versus global or linear versus context-dependent, for example—is doubtful. Although left lateralization of language and right lateralization of visuospatial function are widely recognized, lateral specialization in the prefrontal regions is less obvious but of clinical significance. Frontal lobe degeneration involving the right more than the left frontal lobe is particularly associated with disinhibition. Traumatic injury to the right hemisphere is more associated with depression and anxiety, to the left with anger and hostility. Women and sinistrals tend to show less lateralization of language (and perhaps of other functions), so that left hemisphere lesions are less likely to produce severe impairment.

Of considerable importance for neuropsychiatric practice is the question of lateralization of emotional processing. An array of evidence supports the notion of differential emotional valences in the two hemispheres. On this account, the left hemisphere is specialized for positive emotions, the right for negative emotions. Thus, it has been claimed that left hemisphere destructive lesions are associated with pathological crying, right hemisphere ones with pathological laughing; contrariwise, left hemisphere discharging lesions produce gelastic (laughing) epilepsy, right hemisphere ones dacrystic (crying) epilepsy. However, these associations have not been confirmed by further studies.

To the contrary, much evidence favors assigning a prepotent role in emotional processing in general to the right hemisphere. Patients with right hemisphere damage appear to be more impaired at perceiving emotion, regardless of the valence or input medium. Lesions of the right hemisphere are associated with impairments in processing emotion in speech, a defect known as aprosodia. Patients may lack the capacity to modulate prosody, so as to encode emotional information into speech, or the capacity to recognize emotional intonations produced by others. Subtler clinically may be deficits in recognizing emotion in faces or visual scenes. Such deficits may be part of the basis for a finding that may seem counterintuitive, namely that patients with right hemisphere injury have a poorer rehabilitation outcome than their left hemisphere counterparts.

Frontosubcortical Circuits

The projection of prefrontal cortex to subcortical structures in multiple closed loops is a crucial feature of behavioral neuroanatomy. The key concept is that, in each loop, a distinct region of prefrontal cortex projects to a distinct portion of the striatum, then to an output nucleus of the basal ganglia, then in turn to a specific region of the thalamus, which itself projects to the given area of cortex. Thus a set of parallel closed loops of

frontosubcortical connections process information in separate domains. In the motor system, premotor cortex and supplementary motor area project primarily to putamen, the output of which projects via ventrolateral globus pallidus and caudolateral substantia nigra pars reticulata (SNr) to ventrolateral/ventroanterior and centromedianum nuclei of thalamus and then back to the originating cortical structures. Of particular interest to neuropsychiatrists are the loops involving dorsolateral prefrontal, medial and lateral orbitofrontal, and anterior cingulate cortex:

- Dorsolateral prefrontal cortex projects to dorsolateral caudate; projections from caudate go to dorsolateral globus pallidus and SNr. The output from basal ganglia flows primarily to ventrolateral and ventral anterior nuclei of thalamus (but also to dorsomedial nucleus of thalamus), where it projects to areas 9 and 46 of dorsolateral prefrontal cortex.
- Lateral orbitofrontal cortex projects to ventromedial caudate, thence to the caudomedial aspect of SNr. The thalamic level of this loop is represented in ventral anterior and dorsomedial nuclei, whence projections arise back to the lateral aspect of area 12 in orbitofrontal cortex.
- The medial orbitofrontal cortex loop features projections from gyrus rectus and medial orbital gyrus to ventromedial caudate; output from the basal ganglia arises in SNr and flows to dorsomedial nucleus of thalamus as well as ventrolateral and ventral anterior nuclei, then back to medial orbitofrontal cortex. Anterior cingulate cortex, in the dorsomedial aspect of the hemisphere, projects to ventral striatum, including nucleus accumbens and olfactory tubercle (termini of the mesolimbic dopamine system), with output from SNr flowing through ventral anterior thalamus on its way back to anterior cingulate cortex.

Disruption of each of these loops produces a distinctive clinical syndrome. As is implied by the concept of a circuit, deficits similar to those produced by cortical damage can also occur with damage to the subcortical connections of the cortical region. Before sketching each of these syndromes, it should be noted that most naturally occurring lesions do not respect the anatomic boundaries, so that clinical presentations are commonly mixed. Nonetheless, for analytic purposes, the anatomic specificity is of interest and importance.

Interference with the loop involving dorsolateral prefrontal cortex prominently produces executive cognitive impairment, with decrements in working memory, problem solving, and related capacities. Damage to this loop commonly arises from traumatic brain injury, stroke, and basal ganglion degenerative diseases, such as Parkinson disease. Involvement of the white matter of the frontal lobes by small vessel disease commonly leads to interruption of corticosubcortical connections in this circuit,

resulting in the picture of subcortical dementia.

Damage to orbitofrontal cortex and its connections produces impulsivity, disinhibition, dampening of the experience of emotion, irritability and lability of affect, poor judgment and decision making (especially in regard to social behavior), and insightlessness about these impairments. These impairments are generally seen with bilateral damage, although unilateral right-sided injury may also produce them. As a neighborhood sign, damage often involves the olfactory nerve (which runs along the orbital surface of the brain) with consequent anosmia, at times the only neurological sign. Cognitive function as tested by the usual bedside or neuropsychological probes may be unaffected, even in the presence of devastating personality change. Trauma is a common etiology.

Damage to dorsomedial prefrontal structures may arise from tumor or stroke. Abulia and apathy, disorders of initiation of action and the experience of motivation, are the result. Abnormalities of initiation of movement, with akinetic mutism as the most extreme state, may occur. Cingulate cortex is a structure of particular interest. Evidence from animal and imaging studies demonstrates its importance in orienting attention under conflicting stimulus demands, modulating focused problem-solving and monitoring performance to optimize reward. A cell type seen only in cingulate cortex, the spindle cell, appears in evolution only with the great apes and in ontogeny only at age 4 months, concomitant with the infant's increasing capacity to focus attention. Interference with the output of cingulate gyrus, namely by interrupting the cingulum—the procedure of cingulotomy—appears to be beneficial in a disorder of excessive attention, namely obsessive-compulsive disorder (OCD).

Limbic System

Le grand lobe limbique was delineated in the mid-19th century (by Broca of aphasia fame) as a ring of cortical and subcortical structures on the medial aspect of the hemispheres. Papez drew attention to the circuit formed by projections from hippocampus via fornix to mamillary bodies of hypothalamus, then to anterior nucleus of thalamus, then via the anterior limb of internal capsule to cingulate gyrus, then back to hippocampus via presubiculum, entorhinal cortex, and the perforant pathway. In addition to this "Papez circuit," amygdala and its reciprocally connected orbitofrontal cortex are taken to form part of a limbic system, a term first used by MacLean. Although some anatomists bristle at its inclusiveness, the concept is nearly universally used, probably because it focuses attention on the "emotional brain."

The core limbic structures are characterized by rich reciprocal monosynaptic connections with the hypothalamus. These are the (1) hippocampus, (2) amygdala, (3) piriform cortex, anterior to amygdala on the medial surface of the temporal lobe, (4) septal nuclei, in the medial wall

of the hemispheres, immediately rostral to lamina terminalis, and (5) substantia innominata in the basal forebrain. Paralimbic cortices reside in temporopolar, insular, and orbitofrontal regions, which have primary affiliations with amygdala, and in parahippocampal, retrosplenial/posterior cingulate, and subcallosal regions, with primary affiliations with hippocampus.

In the limbic system, broad and direct input from sensory cortices into amygdala and hippocampus is extensively processed on its way to effector neurons in hypothalamus that regulate autonomic and endocrine activity. In addition to this mediation of the regulation of the internal milieu, the limbic system gates the activity of the motor systems in the basal ganglia, regulating action in the external milieu. This occurs by prefrontal cortical integration of information in the limbic frontosubcortical circuit, which reaches the cortex via projections from ventral pallidum to dorsomedial nucleus of thalamus.

One reason for the central importance of the limbic system in neuropsychiatry is that the threshold for production of epileptic discharges is lowest in amygdala and hippocampus. Thus most epilepsy in adults is limbic epilepsy. One consequence is the "voluminous mental state" first identified by Hughlings-Jackson. This refers to the range of experiential phenomena encountered as auras in limbic epilepsy: déjà vu, depersonalization/derealization, micropsia and macropsia, and so on. Such symptoms are seen not only in epilepsy but also in mood disorders and as putative pointers to limbic involvement in paroxysmal disorders not of clear epileptic nature, including those associated with borderline personality disorder and with childhood abuse. Their presence, therefore, does not unequivocally mark an organic diagnosis.

Another reason for the centrality of the limbic system is that hippocampus in particular has a crucial role in explicit memory, further discussed below. Persisting substantial amnestic deficits in multiple modalities require limbic system damage.

Cerebellum

Against the prevailing notion that the cerebellum is a motor structure, anatomic evidence shows that cerebellar inputs access areas of prefrontal cortex, with a relay in thalamus. These areas of cortex project reciprocally to cerebellum, creating, as with the prefrontal-basal ganglia circuits previously discussed, a set of parallel (relatively) closed loops, or channels. These crossed connections from the cerebellar hemispheres and the further crossing of descending cerebrofugal long tracts mean that motor deficits are manifest ipsilateral to lateralized cerebellar injuries. Additional reciprocal connections link cerebellum with hypothalamus monosynaptically and with other areas of the limbic system via a relay in the basis pontis. The phylogenetically older vermis and fastigial nucleus

can be differentiated from the neocerebellum of the cerebellar hemispheres and considered a "limbic cerebellum."

Growing evidence of cerebellar contributions to cognition and affect comes from clinical data and neuropsychological studies. The data are fraught with uncertainty, however, because many cerebellar patients have disorders that may not be limited to cerebellum; for example, cerebellar degenerations may include cortical degeneration, and tumors (and their treatment with radiation and chemotherapy) may have remote effects. Moreover, the phenomenon of crossed cerebellar diaschisis—the reduction in blood flow to connected neocortical areas after cerebellar damage—means that interpretation of deficits as due to abnormal cerebellar processing, as opposed to shut-down of cerebral cortical processing, is treacherous. Nonetheless, patients with stroke lesions clinically and by neuroimaging limited to cerebellum may have deficits in executive cognitive function, memory, language, and visuospatial function. The data suggest that lateralized cerebellar damage is associated with the predicted lateralized cognitive phenomena (right cerebellar damage with language impairment, left with visuospatial impairment). Reports of an affective syndrome after cerebellar injury are less systematic. Defects in affect regulation, with irritability and lability, are proposed to be associated with damage to the limbic cerebellum, notably the vermis.

White Matter and Cerebral Connectivity

Although the volume of neocortex has increased over the course of phylogenetic history, the volume of white matter has increased disproportionately. In human beings, the white matter tracts occupy some 42 percent of the volume of the hemispheres. The great majority of these fibers serve corticocortical connectivity rather than projections between cortical regions and subcortical sites; for example, thalamic input is estimated to represent only 5 percent of the total input into primary sensory cortex, the remainder being from other cortical areas.

The fibers in white matter are of several types. First are the longer intrahemispheric fiber tracts:

- arcuate fasciculus, which connects superior and middle frontal gyri to the temporal lobe and (via a superior portion of the fasciculus called superior longitudinal fasciculus) the parietal and occipital lobes;
- uncinate fasciculus, which connects orbitofrontal cortex to temporal cortex and (via an inferior portion of the fasciculus called inferior occipitofrontal fasciculus) the occipital lobe;
- cingulum, which lies medially beneath cingulate cortex in cingulate gyrus and connects frontal and parietal lobes with parahippocampal gyrus and adjacent structures.

Second are the long projection systems linking cortex, subcortical

nuclei, and lower portions of the neuraxis. Medial forebrain bundle is the primary connection between limbic structures and the brainstem and carries projections from the monoaminergic cells in the midbrain and pons. Others are the thalamic peduncle, with reciprocal fibers between thalamus and parietal lobe, and the corticopontine and corticospinal tracts, descending through the corona radiata and internal capsule. Fibers from prefrontal cortex descend into the anterior limb of internal capsule, so that lesions there may have predominant behavioral and a paucity of elementary sensorimotor effects.

Lacunes and degeneration of the white matter due to hypertensive small vessel disease (Binswanger disease) interrupt these corticocortical fibers and corticosubcortical projections. The result of progressive loss of communication among cortical regions and between cortex and subcortical gray matter is the clinical state of subcortical dementia, which is prominently characterized by slowing of mental processing and failure of executive control processes. The latter may be explained in part by the preferential occurrence of lacunes in frontal locations but also by the impairment of connectivity.

Third, U-fibers are the short, juxtacortical fibers connecting adjacent cortical regions. These fibers are characteristically spared in Binswanger disease, cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), and certain other disease processes.

Fourth are the many specific projection systems linking delimited regions, such as mamillothalamic tract, which connects mamillary bodies with anterior nucleus of thalamus, and fornix, which connects mamillary bodies with hippocampus. Interesting neurobehavioral syndromes have been described related to rare cases of focal interruption of such pathways. For example, interruption of the mamillothalamic tract or of the fornix is implicated in amnesia.

Fifth, several pathways connect the two hemispheres, notably corpus callosum but also anterior and posterior commissures and massa intermedia of thalamus. Syndromes due to interruption of the smaller commissures have not yet been described, although absence of massa intermedia is reported to be associated with schizophrenia in women, and anterior commissure and massa intermedia are larger in women than in men. Corpus callosum is congenitally absent in numerous neurodevelopmental syndromes, and its absence has been associated with schizophrenia. Congenital absence is not, however, associated with the interesting disconnection symptoms seen in lesional interruption of the callosum, such as by anterior or posterior cerebral artery stroke or by surgical callosotomy for control of epilepsy.

Two callosal disconnection syndromes are worthy of specific mention. After anterior cerebral artery occlusion with anterior callosal infarction, the right hemisphere is deprived of verbal information; a left-hand apraxia is

seen, and the patient cannot name unseen objects placed in the left hand. Reciprocally, the right hand shows constructional apraxia. This is termed the anterior disconnection syndrome. After occlusion of the left posterior cerebral artery with infarction of the left occipital lobe and the splenium (posterior portion) of corpus callosum, the language cortices of the left hemisphere lose access to visual information: The left visual cortex is damaged, as are the projections from the right visual cortex, which cross in the splenium. Thus, reading becomes impossible, although other language functions are unaffected—the syndrome of alexia without agraphia.

Cerebral Cortex

The cerebral cortex develops through complex but increasingly well understood processes of cell proliferation and migration, axonal projection, and dendritic proliferation and pruning. Abnormalities in these processes lead to cortical dysplasia, with clinical consequences including intellectual disability and epilepsy. Some 10 percent of intractable epilepsy may be due to such disorders, and increasingly migration abnormalities are recognizable by imaging prior to neuropathological examination. Failure of normal pruning of synapses by elimination of dendrites is now known to be crucial in the pathogenesis of the fragile X syndrome and has been speculatively linked to schizophrenia. Rarely cortical dysplasia may be present without epilepsy or intellectual disability; the neurobehavioral consequences of this abnormality are just coming under investigation.

The organization of sensory cortices follows a regular plan. Each primary sensory cortical area projects to unimodal association cortices specialized for the extraction of features in that particular modality; the unimodal association cortices are densely and reciprocally interconnected. For example, visual association cortex has specialized regions for color, motion, and shape. Unimodal association cortices project in turn to heteromodal cortices, which receive inputs from more than a single sensory modality. Heteromodal cortices are located in prefrontal, posterior parietal, lateral temporal, and parahippocampal regions. Unimodal cortices do not project to unimodal cortices in other modalities, only to the higher-level heteromodal cortices. Further, widespread hippocampal projections to cortex arrive only at association cortices, not primary sensory cortices. These structural features amount to the isolation of sensory processing from top-down influences over the first several synaptic stages and presumably increase its fidelity to external phenomena.

Lesions of cortical association areas produce an array of behavioral and cognitive disorders of intriguing specificity. The specificity can be demonstrated by the occurrence of double dissociations: a lesion in area A produces a deficit in function X but not Y; a lesion in area B produces a deficit in function Y but not X. This pattern of findings provides crucial confirmation that the deficits arise not from task difficulty (if Y were simply

more difficult than X, then Y would always be disturbed when X was disturbed), but from separable processing components. For example, some patients show a greater impairment for naming living things than for naming artifacts after a brain injury. However, occasionally patients show the opposite pattern, greater impairment in naming living things: a double dissociation. The explanation of the discrepancy thus cannot depend on insufficient processing resources but must reveal a property of the organization of the semantic system.

Cognitive disorders of the visual system can serve as a paradigm of the syndromes seen with damage to association cortex. Lesions of primary visual cortex (V1, or Brodmann area [BA] 17) produce cortical blindness, in a quadrant, hemifield, or the entire visual field. Despite the genuine blindness, accuracy above chance in localizing visual stimuli can be achieved without awareness of vision, the phenomenon of blindsight, which testifies to subcortical visual processing inaccessible to consciousness. V1 projects to adjacent cortical regions (BA18 and BA19), which contain neurons that respond to specific features of visual stimuli, such as color, movement, or shape. Lesions in these cortices produce deficits in identification of these features. Thus arise syndromes such as central achromatopsia, demonstrated by inability to sort (as well as to name) colors. The information transfer divides into dorsal and ventral streams, the former specialized for localization of visual stimuli ("where") and the latter for identification of the stimuli ("what"). Dorsal lesions involving superior parietal lobule can produce impaired reaching under visual guidance (optic ataxia), a part of the Balint syndrome; the deficit testifies to the integration of visual information with motor output in association cortex. Ventral lesions, involving inferotemporal cortex, produce defects in recognition (agnosia). Agnosic patients are not only unable to name elements within the domain of agnosia but also unable to demonstrate their use or show recognition of the items in other nonverbal ways.

Central auditory disorders include cortical (or central) deafness; pure word deafness, the inability to recognize words presented in the auditory modality despite preserved visual–verbal function; and auditory agnosia, the inability to recognize words or complex sounds (e.g., the meaning of the ringing of a telephone). Central deafness requires bilateral lesions involving primary auditory cortex in superior temporal gyrus or auditory radiations in white matter. Patients with pure word deafness generally have bilateral lesions of association cortex more anteriorly in superior temporal gyrus, although unilateral left lesions, presumably disconnecting left from right cortices by subcortical damage, also are reported. Primary auditory cortex is partially spared in these cases. In agnosia for nonverbal environmental sounds, right-hemisphere damage is sufficient to produce the deficit. Amusia, the incapacity to recognize musical sounds, is associated with cortical damage, but the issue of laterality is complex, dependent in part on

the preinjury level of musical skills.

Full evaluation of these disorders requires techniques that go well beyond bedside examination or routine paraclinical tools. At issue in the agnosic disorders is the extent to which a deficit is apperceptive (i.e., due to impairment in analysis of subtle perceptual elements presumably dependent on more upstream cortical regions), and to what extent associative (i.e., occurring in the absence of definable perceptual abnormalities and presumably due to dysfunction of more downstream cortical analyzers). This distinction requires detailed neuropsychological and often psychophysical assessment.

Modulators of Brain States

This account of cognitive processing in cortex will seem to many psychiatrists to leave out of consideration the matters with which they are most concerned, namely pervasive states of altered mood, drive, and behavior. That such states are behaviorally pervasive does not entail that they are anatomically global. Limbic structures discussed above provide in part the anatomic substrate for emotional states. Further, several systems with diffuse cortical projections have the capacity to modulate processing in widespread brain regions. These originate in:

- Intralaminar thalamic nuclei, which project to cortex (especially prefrontal and cingulate cortex) and to striatum;
- Histaminergic cells in posterior hypothalamus;
- Serotonergic cells in pontine raphe nuclei;
- Noradrenergic cells in locus ceruleus;
- Dopaminergic cells in the midbrain ventral tegmental area, giving rise to the mesocortical and mesolimbic systems;
- Cholinergic cells in basal forebrain nuclei, such as nucleus basalis of Meynert.

The last of these is of relevance to cholinesterase inhibitor treatment of dementia, the preceding three of relevance to treatment of mood, anxiety, and psychotic disorders. The hypothalamic histaminergic projections are involved in arousal. “Nonspecific” thalamic projections may have an important role in executive dysfunction seen after thalamic lesions. For fear of complacency in understanding of such pathways, it should be recalled that only within the past few years have a previously unknown neurotransmitter and its pathways been recognized, and the discovery of orexin/hypocretin and its hypothalamic anatomy exposed the secrets of narcolepsy. Neuropsychiatric anatomy is not a closed book.

MODULARITY AND NEUROPSYCHIATRY

These focal behavioral syndromes, and many others, compel attention to local processing in the brain and almost irresistibly suggest a particular

model of brain organization. One imagines a box-and-arrow diagram, in which each box—representing an elementary cognitive function—maps on to a specialized region of cortex. Each area of cortex has its job to do, and a lesion of any area produces a distinctive, delimited, and predictable deficit.

This model raises the issue of modular organization of the brain. The general topic of modularity in cognitive processing deserves further consideration because it is crucial to the theoretic perspective of neuropsychiatry, and in particular because it bears on the value of neuropsychiatric data for the understanding of idiopathic psychiatric disorders. Modularity in cognitive neuroscience refers to a brain organization characterized by multiple computational devices, each of which operates on characteristically encapsulated input with prewired (perhaps innate) rules, thus being rapid, efficient, and reliable. For example, elementary visual processing can be considered modular, inasmuch as it utilizes restricted input with hardwired feature extraction (e.g., motion, color, shape). In another domain, consider that it is easier to teach an animal to associate a taste than a visual stimulus with the aversive effects of an ingested toxin. This finding implies domain-specific, innate learning constraints. The classic cognitive example of domain-specific prewiring is language, for example, Chomsky's observation that children generate language errors that they have never heard: "He bringed me here," the small child might say, although he or she has never heard an adult say "bringed." The implication is that a language-processing module possesses innate grammatical rules that have generated a grammatical form without experiential foundation.

Evolutionary psychologists have forcefully argued the case for modular processing, as opposed to domain-general problem-solving devices. The core of the evolutionary argument is that cerebral organization is the result of natural selection operating on the adaptational fitness of mankind's Pleistocene hunter-gatherer ancestors. Domain-specific processing has advantages of speed and efficiency that necessarily lead to an advantage in fitness. The availability of pre-experiential information about the content of domain-specific processing carries a large advantage over the "combinatorial explosion" of informational possibilities requiring evaluation by a domain-general processor. For example, detection of cheating in social exchanges is an essential element of adaptation in a population group featuring cooperative behavior. Is it a function of a domain-general logical problem-solving device, or is there a cheater-detection module? Cross-cultural evidence shows that people are far better at detecting violations of social exchange rules than at solving problems of equivalent logical complexity when posed in other terms, and focal lesions can differentially affect cheater detection. The implication is that prewired mechanisms, presumably located in a particular brain area, are "tuned" to recognize and reason about this adaptationally crucial behavior, just as

innate mechanisms subserve language learning and toxin recognition. One of the strengths of the evolutionary approach is to direct attention to those processing domains the modularity of which is plausible on adaptational grounds.

However, many of the "modules" that have attracted clinical interest are not plausibly directly the objects of natural selection. Reading and writing are clear examples. These have arisen too recently in evolutionary time to have been the product of natural selection and thus must depend on the workings of processors that are, at least to this extent, domain general.

Moreover, much of the literature on modularity is written from a cognitive psychological or philosophical perspective, with less attention to the "wetware" (i.e., actual brain substance) implementation of the processing devices. A foundation in cognitive neuroscience and evolutionary biology can enrich clinical theories, but it creates the potential for misunderstanding by clinicians interested in the functioning of patients with brain lesions. The modules of the evolutionary biologists and philosophically inclined cognitivists do not map directly on to brain areas. One of the striking results of functional neuroimaging experiments is that, however the function under study is delimited, multiple areas of brain activation are found. Even functions that seem psychologically fundamental are not implemented in a simple way, and local processing components may be recruited into multiple networks subserving a variety of tasks. This seems to be the case in respect to the limited number of frontal sites involved in a wide range of executive tasks. The specialization of regions is dependent on input from other regions; specialization is not entirely dependent on intrinsic properties but partially on top-down influences.

The issue is not whether different cerebral regions carry out different modes of information processing. This is unquestionably so, and neither unreconstructed holists who believe in the equipotentiality of cortex nor strict localizationists who believe only in fully autonomous processing devices figure on the current neuroscientific scene. The question is how regions are linked in carrying out tasks. Functions are implemented by networks, most or all of the nodes of which participate in multiple functional networks. This pattern of cerebral organization has been termed selectively distributed processing or sparsely distributed networks. Although cortical regions have specialized capacities for information processing, functions cannot be localized to regions (as Hughlings-Jackson explicitly warned a century and a half ago). It would be erroneous, for example, to believe that an area crucial for face recognition contained all of the neurons, and only neurons, that respond to faces.

Moreover, normal individuals may differ in how they recruit regions into networks. The methods used in studying groups of subjects in functional imaging experiments may obscure such individual differences.

For example, robust individual differences in patterns of activation emerged in a memory task, differences putatively reflecting different strategies in performing the task. The differences were stable within individuals over time, yet analysis of group data revealed activations in regions activated in only some of the subjects and failed to disclose activations in regions consistently activated in others. Individual differences in organization of language cortex are evident clinically in the unusual, but not negligible, occurrence of crossed aphasia (aphasia due to right hemisphere injury in a dextral), crossed nonaphasia (lack of aphasia with a left hemisphere injury that should cause aphasia in a dextral), and aphasic deficits anomalous in respect to the predicted effects of lesions in both dextrals and sinistrals.

Another critique of the modularity hypothesis crucial for neuropsychiatry derives from developmental psychology. Trenchant arguments contradict the assumption that a mapping of deficits to specific brain structures could be static over developmental time. To the contrary, how the brain performs cognitive tasks changes with development. Development entails changing patterns of interaction among brain components, and localization may alter as neurons and regions become "tuned" in responsiveness, based on their initial characteristic processing biases and their patterns of inputs and connectivity. This reorganization of cortical function could mean that the same behavior has different substrates at different developmental epochs. For example, in adult subjects with Williams syndrome, poor function at number processing and good language skills are characteristic; however, in infancy the opposite pattern is seen. Whatever the fundamental processing disorder of genetic origin may be, it cannot be seen as having knocked out a module. A large expanse of nonlinear brain development lies between the gene and the clinical phenomena, an expanse that can be understood only with a better theory than neophrenology. The very development of modularity can be anomalous. Indeed, in the Williams syndrome cases, functional magnetic resonance imaging (fMRI) data disclose an anomalous, diffuse pattern of activation for music perception, an area of preserved or enhanced ability in these patients. Focal syndromes in adults provide an appropriate place to start formulating hypotheses, but a deficit seen in an idiopathic disorder cannot be assumed to have its basis in dysfunction in the same simple locus as a phenomenologically similar deficit seen after a focal brain lesion occurring in an adult.

Nothing in this line of argument diminishes the interest of focal neurobehavioral syndromes, which are clinical facts that have a substantial heuristic value for the cognitive neurosciences. Neuropsychiatry, along with other brain specialties, has the task of importing into clinical theory the understanding of the mind and brain that is developing in cognitive neuroscience. The search for psychopathological understanding based on

identification of deficits in cognitive modules that are relatively well understood in normal subjects has been termed cognitive neuropsychiatry. This pursuit inevitably results in deconstruction of the psychiatric diagnoses of the Diagnostic and Statistical Manual of Mental Disorders (DSM) or the International Statistical Classification of Diseases and Related Health Problems (ICD) into symptoms or syndromes, because the standard diagnostic categories are generally based on folk-psychological notions (such as the division between "thought" and "affective" disorders).

Much of this section is devoted to the anatomical mode of thought practiced by neuropsychiatrists. However, some clinicians hope that neuropsychiatry will provide a localizing taxonomy of behavioral syndromes, so that particular psychiatric disorders will carry the same localizing power as, say, the Babinski sign for the corticospinal tract: the nuclear syndrome of schizophrenia to the left temporal lobe, for example. From the contemporary cognitive neuroscience perspective just reviewed, this seems likely to be false hope. The Babinski sign is a limiting, not a paradigmatic, case of brain–behavior relationships. For the fullest understanding of complex mental syndromes, notably those traditionally in the realm of psychiatry, a more adequate theory of brain function is needed than can be offered by the localizationist tradition.

CLINICAL EVALUATION

The neuropsychiatric perspective places great reliance on information that can be gathered at the bedside. No practical inquiry and examination can include all possible items; rather, the clinician selects from a toolbox of probes of the history and of the patient's functioning in the examination room in order to confirm or refute hypotheses generated by the emerging clinical picture. Screening items should have high sensitivity but not necessarily high specificity. Beyond screening, elements of the history and examination that might potentially elucidate the nature of the disease process under consideration form the entire corpus of medical assessment. For example, the neuropsychiatrist considering liver disease as the explanation of delirium will want to estimate the liver span during the physical examination.

The Neuropsychiatric History

The initial steps in screening for the presence of organic disease in patients with mental symptoms are easily taken. The physician should obtain a general medical history, including a history of diseases possibly relevant to the neuropsychiatric symptoms under consideration, and a review of systems in potentially relevant areas. With a cognitively impaired or psychotic patient, such history taking may be unreliable. Collateral history from a family member or other informant and review of medical records

are almost always essential.

With virtually every patient, the clinician should inquire as to a history of (1) heart, lung, liver, kidney, skin, joint, eye disease; (2) hypertension; (3) diabetes; (4) traumatic brain injury; (5) seizures, including febrile convulsions in childhood; (6) unexplained medical symptoms; (7) substance misuse; (8) current medication; and (9) family history of neuropsychiatric disorder. The inquiry about these disorders in some settings can be quite general. For example, the question "Have you ever had heart problems?" along with a few questions in the review of systems may suffice to screen for heart disease in a young apparently healthy patient. In other settings, more detailed information must be gathered.

The review of systems as well should vary according to the setting. Positive responses should of course lead to further inquiry. The clinician should be practiced in inquiring about:

- Constitutional symptoms: Fever, malaise, weight loss, pain complaints;
- Neurological symptoms: Headache, blurred or double vision, impairment of balance, impairments of visual or auditory acuity, swallowing disturbance, focal or transient weakness or sensory loss, clumsiness, gait disturbance, alteration of urinary or defecatory function, altered sexual function;
- Paroxysmal limbic phenomena: Micropsia, macropsia, metamorphopsia, déjà vu, jamais vu, forced thoughts or emotions, depersonalization/derealization, autoscopy, paranormal experiences such as clairvoyance or telepathy;
- Thyroid symptoms: Heat or cold sensitivity, constipation or diarrhea, rapid heart rate, alopecia or change in texture of hair;
- Rheumatic disease symptoms: Joint pain or swelling, mouth ulcers, dry mouth or eyes, rash, past spontaneous abortions.

Birth History and Early Development

Because brain development starts before birth, so too does the neuropsychiatric history. The clinician should note:

- Maternal substance misuse, bleeding, and infections during the pregnancy;
- The course of labor;
- Fetal distress at birth, including Apgar scores if available;
- Perinatal infection or jaundice;
- Motor and cognitive milestones, such as the age the child crawled, walked, spoke words, spoke sentences;
- The infant's temperament;
- The child's school performance (including special education and anomalous profiles of intellectual strengths and weaknesses), usually the best guide (absent psychometric data) to premorbid intellectual

function.

The role of perinatal injury in cerebral palsy and mental retardation has commonly been overestimated; in many instances developmental disorder was present in gestation prior to the perinatal misadventure, which may in fact have arisen from the pre-existing abnormality. However, perinatal injury, in particular hypoxic injury, is probably associated with later schizophrenia.

Head Injury and Its Sequelae

Head injury is commonly a potential factor in later mood and psychotic disorders as well as cognitive impairment, epilepsy, and posttraumatic stress disorder (PTSD). The clinician should inquire about a history of head injury in virtually every patient. The nature of the injury should be clarified by eliciting the circumstances, including risk-taking behaviors that may have predisposed to injury and others who were injured in the same incident, often an emotionally powerful aspect of the event. The loss of consciousness is not a prerequisite to important sequelae; even a period of being stunned, "seeing stars," can presage later neuropsychiatric symptoms. The period of loss of consciousness, or coma, should be established, ideally with the assistance of contemporaneous medical records. The period of retrograde amnesia—from last memory before the injury to the injury itself—and of anterograde amnesia—from injury to recovery of the capacity for consecutive memory—should be noted.

Paroxysmal Disorders

Paroxysmal disorders of neuropsychiatric interest include epilepsy, migraine, panic attacks, and episodic dyscontrol of aggression. Taking a history of an episode has common features irrespective of the nature of the disorder. The clinician should track through the chronology of the attack. This starts with the possible presence of a prodrome, a warning of an impending episode in the hours or days prior to one. The episode itself may be presaged by an aura, lasting seconds to minutes. In the case of an epileptic seizure, this represents the core of the seizure itself and may carry important localizing information about the hemisphere side and site of the focus. The pace of buildup, from onset to peak of the ictus, is of differential diagnostic importance. For example, epileptic seizures begin abruptly; panic attacks may have a more gradual development to peak intensity. The mental and behavioral features of the ictus itself should be elicited in detail, if possible from collateral informants as well as from the patient. The duration of the spell and the mode of its termination should be elicited. Inquiring whether the patient has just one sort of spell or more than one is an essential prelude to establishing the frequency of episodes, both at present and at maximum and minimum in the past.

By interviewing the patient and collateral informants, information necessary to make a differential diagnosis can usually be elicited. The differential diagnosis between epilepsy and pseudoseizures can be difficult; but at times, if asked properly, the patient will make the diagnosis for the clinician by reporting "two kinds of seizures," one of which is clearly epileptic and the other of which is clearly dependent on emotional states.

Cognitive Symptoms

Recognizing cognitive symptoms in patients without established dementia is a crucial element of neuropsychiatric history taking. Such symptoms may be outweighed by more dramatic behavior or mood change, but identification of cognitive impairment can reorient the diagnostic evaluation of a late-life depression, for example. No doubt the commonest complaint along cognitive lines is of memory problems. In the setting of depression, the more intense the complaint of memory impairment, the less likely it is to have an organic basis and the more likely to testify to depressive ideation and attentional failure. The clinician should establish whether forgotten material (say, an acquaintance's name or a task meant to be performed) comes to the patient later, as a matter of absentmindedness rather than memory failure. Certain other complaints are highly characteristic of organic disease. These include a loss of the capacity for divided attention or for the automatic performance of familiar tasks. A patient might report, for example, no longer being able to read and listen to the radio at the same time. Getting lost or beginning to use aids for recall, such as a notebook, is suggestive of organic cognitive failure.

Appetitive Symptoms and Personality Change

Alterations of sleep, appetite, and energy are common in idiopathic psychiatric disorders as well as transiently in the healthy population and cannot be interpreted as implying brain disease. Certain patterns of altered sleeping and eating behavior and personality, however, are pointers to organic disease. Excessive daytime sleepiness or sleep attacks raise the question of sleep apnea or narcolepsy, or in a different temporal pattern, Kleine–Levin syndrome. Abnormal behavior during sleep raises the question of a parasomnia. Of particular interest is rapid eye movement (REM) behavior disorder, which may rarely be due to a pontine lesion, but when a focal lesion is absent strongly points to ingravescent Lewy body disease. Much more rarely nocturnal dream-enacting behavior represents a prion disease, notably fatal familial insomnia. Loss of dreaming occurs with parietal or bifrontal damage; loss of visual imagery in dreams occurs with ventral occipitotemporal damage, part of the Charcot–Wilbrand syndrome (loss of visual imagery with brain damage). In medial hypothalamic disease, eating behavior is marked by lack of satiety and resultant obesity.

In the Klüver–Bucy syndrome of bilateral anterior temporal damage (involving amygdala), patients mouth nonfood items. With frontal damage, patients may stuff food into the mouth, a form of utilization behavior, sometimes with alarming or even fatal consequences. A "gourmand" syndrome of excessive concern with fine eating has been associated with right anterior injury.

Changes in sexual behavior are common consequences of brain disease. Hyposexuality is common in epilepsy, possibly as a consequence of limbic discharges. A change in habitual sexual interests, quantitative or qualitative, developing in midlife suggests organic disease. It is possible, although understudied, that relevant organic disease, such as the sequelae of traumatic brain injury, is common in sexual offenders. Other changes in personality, such as the development of shallowness of affect, irritability, loss of sense of humor, or a coarsening of sensibilities may indicate ingravescent organic disease, for example frontotemporal dementia.

Handedness

About 90 percent of people designate themselves as dextral, almost all the rest as sinistral, and a very few as ambidextrous. The true state of affairs is somewhat more complicated, in that handedness may be considered dimensionally (i.e., as a matter of degree rather than categories). A patient may call himself or herself right handed but use the left hand preferentially for certain tasks. Inquiring about a few specific tasks—writing, throwing, drawing, using a scissors or toothbrush—yields helpful information. A family history of sinistrality may also be relevant.

THE NEUROPSYCHIATRIC PHYSICAL EXAMINATION

To the neuropsychiatrist, the physical examination is a central feature of clinical evaluation. In principle, any aspect of the general physical or neurological examination may be relevant to neuropsychiatric diagnosis, if only in revealing an incidental clinical problem in a neuropsychiatric patient. Here, the focus is on elements of the physical examination with specific relevance to detection and identification of organic disease in patients with mental presentations. The mental examination, including the cognitive examination, is discussed below, in association with syndromes of behavioral disorder and insofar as it can elicit or elucidate these syndromes in the consultation room.

The General Physical Examination

General Appearance. Dysmorphic features include so-called minor physical anomalies, some of which are captured in the widely used Waldrop scale. These are associated with developmental disorders, including schizophrenia. These features center on the head, hands, and feet. No

single minor anomaly is diagnostic of pathological development, but the coincidence of multiple anomalies argues that development has gone awry. Many specific developmental disabilities syndromes can be diagnosed by the constellation of dysmorphic features presented. Cleft lip or palate may be associated with brain malformations and frontal cognitive impairment. Asymmetry of the extremities, often best seen in the thumbnails, or of the cranial vault, points to a developmental abnormality. Occasionally, a patient even reports wearing shoes of different sizes on the two feet. The larger extremity and the smaller side of the head are ipsilateral to the abnormal cerebral hemisphere. Short stature is an important feature of many developmental syndromes, both common, such as fetal alcohol syndrome and Down syndrome, and uncommon, such as mitochondrial cytopathies. Abnormal habitus, such as the marfanoid habitus of homocystinuria, may be a clue to diagnosis. Weight loss is an important clue to systemic disease, such as neoplasia; it should not be dismissed without further ado even in a patient with depression, which may—but may not—account for the weight loss. Weight gain equally may point to limbic or systemic disease, especially an endocrinopathy, or may reflect toxicity of psychotropic drugs.

Vital Signs. Elevated temperature or heart or respiratory rate should never be ignored, even in a patient whose agitation or anxiety might seem to explain the abnormality. Doing so risks missing infection, neuroleptic malignant syndrome, connective tissue disease, or other important causes of morbidity. Abnormal respiratory patterns occur in hyperkinetic movement disorders (including tardive dyskinesia). Yawning is a feature of opiate withdrawal and serotonergic toxicity.

Skin. Alopecia or rash may point to systemic connective tissue disease. Alopecia is also a feature of drug toxicity and hypothyroidism (where thinning of the lateral part of the eyebrow is characteristic). The malar rash of systemic lupus erythematosus is typically slightly raised and tender and extends to both cheeks in a “butterfly” pattern, while sparing the nasolabial folds. Discoid rashes in lupus are characterized by hyperkeratosis, atrophy, and loss of pigment; the strong tendency to scarring means that the presence of a discoid rash does not necessarily indicate active disease. A pink periungual rash is also characteristic of lupus. A vasculitic rash is classically palpable purpura and may be seen in lupus or other connective tissue diseases. Livedo reticularis, a net-like violaceous pattern on the trunk and lower extremities, is not specific but raises the question of Sneddon syndrome when stroke or dementia is a clinical accompaniment. The neurocutaneous syndromes have typical skin manifestations: adenoma sebaceum (facial angiofibromas), ash-leaf macules, depigmented nevi, and shagreen patches (thickened, yellowish skin over the lumbosacral area) in

tuberous sclerosis; a port-wine stain (typically involving both upper and lower eyelids) in Sturge–Weber syndrome; neurofibromas, café au lait spots, and axillary freckling in neurofibromatosis.

Head. Head circumference should be measured in patients with a question of developmental disorder. Most reference works give normal ranges for head circumference in developing children but not for adults, and extrapolation would be inaccurate. Fortunately, adequate data to establish normal ranges do exist. Although height and weight need to be taken into account along with gender, roughly the normal range for adult males is 54 to 60 cm (21.25 to 23.5 in); for females, 52 to 58 cm (20.5 to 22.75 in). Old skull fracture or intracranial surgery usually leaves palpable evidence.

Eyes. Exophthalmos usually indicates Graves disease; especially if unilateral it may reveal a space-occupying lesion. Dry eyes, along with dry mouth, raise the question of Sjögren syndrome, although drug toxicity or the aging process are common confounds. Inflammation in the anterior portion of the eye, uveitis, can be appreciated at the bedside by the presence of pain, redness, and a constricted pupil; this is commonly associated with connective tissue disease. The Kayser–Fleischer ring is a brownish-green discoloration at the limbus of the cornea; it sensitively and specifically indicates Wilson disease. The pupils, optic discs, visual fields, and eye movements are discussed below.

Mouth. Oral ulcers can be seen in lupus, Behçet disease, and other connective tissue diseases. Dry mouth is a part of the sicca syndrome, along with dry eyes, discussed above. Vitamin B_{12} deficiency produces atrophic glossitis, a smooth, painful, red tongue.

Heart and Vessels. A carotid bruit indicates turbulent flow in the vessel but is a poor predictor of the degree or potential risk of the vascular lesion. A thickened, tender temporal artery points to giant cell arteritis; here the physical examination is an excellent guide to clinical significance. Cardiac valvular disease, marked by cardiac murmurs, is important in assessing the cause of stroke, and congestive heart failure may be relevant in delirium. In a schizophrenic patient, a murmur may imply the velo-cardio-facial syndrome. Patients with developmental disabilities may have multiple anomalies, including structural heart disease.

Extremities. Joint inflammation as a pointer to systemic rheumatic disease is distinguished from noninflammatory degenerative joint disease (osteoarthritis) by the presence of swelling, warmth, and erythema and is characteristically seen in wrists, ankles, and metacarpophalangeal joints, as

opposed to the involvement of the base of the thumb, distal interphalangeal joints, and spine in degenerative joint disease. Raynaud phenomenon and sclerodactyly are signs of connective tissue disease.

The Neurological Examination

Olfaction. Hyposmia is common in neurological disease, but even more common in local disease of the nasal mucosa, which must be excluded before a defect is taken to be of neuropsychiatric significance. Assessment of olfaction is often ignored ("cranial nerves II through XII normal"), but it is easily performed and gives clues to the integrity of regions otherwise hard to assess, notably orbitofrontal cortex. The olfactory nerve lies underneath orbitofrontal cortex; projections go to olfactory tubercle, entorhinal and piriform cortex in the temporal lobes, amygdala, and orbitofrontal cortex. Testing of olfaction is best performed using a floral odorant, such as scented lip balms, which are inexpensive and simple to carry. Although a distinction can be made between the threshold for odor detection and that for identification of the stimulus, with differing anatomies, at the bedside without special equipment the best one can achieve is recognition of a decrement in sensitivity (i.e., whether the patient smells anything, even without being able to identify it).

Eyes. Pupillary dilation may indicate anticholinergic toxicity; pupillary constriction is a characteristic feature of opiate toxicity. Argyll Robertson pupils are bilateral, small, irregular, and reactive to accommodation but not to light; the finding is characteristic of paretic neurosyphilis but also present in other conditions. Papilledema indicates elevated intracranial pressure; the earliest and most sensitive feature is loss of venous pulsations at the optic disc. A homonymous upper quadrantic field defect is present when temporal lobe disease affects Meyer loop, the portion of the optic radiation that dips into the temporal lobe. A field defect in a delirious patient may point to an etiology in focal vascular disease (as discussed below). The normal spontaneous blink rate is 16 ± 8 per minute. Hypodopaminergia is associated with a reduction in blink rate. Impairment of voluntary eye-opening is seen in association with extrapyramidal signs, making the common appellation "apraxia" of eye opening a misnomer. Impairment of voluntary eye closure is seen after frontal or basal ganglia damage.

Both saccadic and pursuit eye movements should be examined. The former are assessed by asking the patient to look to the left and the right, up and down, and at the examiner's finger on the left, right, up, and down. Pursuit eye movements are examined by asking the patient to follow a moving stimulus in both the horizontal and vertical planes. These maneuvers test supranuclear control of eye movements; the oculocephalic maneuver (doll's head eyes), that is, moving the patient's head, tests the

brainstem pathways and may be added to the assessment if saccades or pursuit is abnormal. Limitation of voluntary upgaze is common in the normal elderly. A limitation of voluntary downgaze, however, in a patient with extrapyramidal signs or frontal cognitive impairment may suggest progressive supranuclear palsy. Slowed saccades are characteristic of Huntington disease. Impairment of initiation of voluntary saccades, requiring a head thrust or head turning, amounts to apraxia of gaze and is seen in developmental disorders as well as Huntington disease and parietal damage. Contrariwise, impairment of inhibition of saccades represents a visual grasp, with forced gaze at environmental stimuli. This can be usefully tested by placing stimuli (a finger and a fist) in the left and right visual fields of the patient and asking the patient to look at the fist when the finger moves, and vice versa. The patient's inability to perform horizontal pursuit or saccadic movements without turning the head may represent the same impairment of inhibition.

Facial Movement. Both spontaneous movements of emotional expression and movement to command should be tested. In pyramidal disorders, spontaneous movements may be relatively spared when the face is hemiparetic for voluntary movements. Contrariwise, in nonpyramidal motor disorders, voluntary movement may be possible despite a hemiparesis of spontaneous movement. The latter situation is seen *inter alia* in temporal lobe disease, including temporal lobe epilepsy, for which it has lateralizing value. Vertical furrows between the eyebrows are known as Veraguth folds and are associated with depression.

Speech. A variety of speech abnormalities are listed in Table 2.1–1. A systematic examination of speech may include asking the patient to produce a sustained vowel ("ahhh"), the performance being assessed for voice quality, steadiness, and loudness; then strings of consonants ("puh-puh-puh") and alternating consonants ("puh-tuh-kuh-puh-tuh-kuh"), the performance being assessed for rate, rhythm, and clarity.

Table 2.1–1.
Speech Syndromes

Syndrome	Output	Characteristic Lesion Location or Associations
Aphemia	Initial mutism, recovery without agrammatism	Broca area (BA44), foot of left third frontal gyrus
Apraxia of speech	Inconsistent and slowed articulation, flattened volume, abnormal prosody	Left insula
Ataxic dysarthria	Slowed, equalization of or erratic stress (scanning), imprecise articulation	Cerebellum, especially superior-anterior vermis, left hemisphere to right

Pyramidal dysarthria	Slowed, strained, slurred	Anterior hemispheres, usually bilateral; may be accompanied by pseudobulbar palsy (dysphagia, drooling, pathological laughing and crying)
Extrapyramidal dysarthria	Hypophonia, monotony of intonation, tailing off with longer phrases	Basal ganglia
Bulbar dysarthria	Nasality, breathiness, slurred articulation	Brain stem
Expressive aprosodia	Loss of emotional "melody of speech"	Right hemisphere
Foreign accent syndrome	Phonetic and prosodic alterations like those of dysarthria after cortical damage but giving listener feeling of foreign accent	Motor or premotor cortex or subjacent white matter of left hemisphere
Developmental stuttering	Repetition, prolongation, arrest of sounds; if overcome in childhood, may re-emerge after stroke, onset of Parkinson disease	Various hemisphere sites
Acquired stuttering	No dystonic facial movements as are seen in developmental stuttering	Various hemisphere sites
Cessation of stuttering	Not an abnormality but the reversal of an abnormality	Various hemisphere sites
Echolalia	Automatic repetition of interlocutor's speech or words heard in environment, sometimes with reversal of pronouns, correction of grammar, completion of well-known phrases	Various anatomies, but seen in frontotemporal dementia, transcortical aphasias, other settings
Palilalia	Automatic repetition of own final word or phrase, with increasing rapidity and decreasing volume	Usually extrapyramidal system
"Blurting," "echoing approval"	Automatic utterance of stereotyped or simple responses (e.g., "yes, yes")	Frontal system

The mute patient poses a special problem in neuropsychiatric assessment. Mutism may occur at the onset of aphemia or transcortical aphasia due to vascular lesions, and it commonly develops late in the course of patients with frontotemporal dementia or primary progressive aphasia. The examiner should assess nonspeech movements of the relevant musculature, for example tongue movements, swallowing, and coughing. Other means of communication should be attempted, such as gesture, writing, or pointing on a letter board or word board.

Abnormalities of Movement. The neuropsychiatric examiner should pay attention to weakness, abnormality of muscle tone, abnormal gait, and involuntary movements. Weakness due to muscle, peripheral nerve, or lower motor neuron disease is associated with atrophy, fasciculations, characteristic distributions, loss of reflexes, and tenderness in the case of muscle disease. Of greater relevance to cerebral mechanisms, pyramidal weakness, greatest in the distal musculature, is accompanied by increased

muscle tone in a spastic pattern (flexors in the upper extremity, extensors in the lower extremity, with the sudden loss of increased tone during passive movement, the "clasp-knife" phenomenon), loss of control of fine movements, brisk tendon jerks, and the presence of abnormal reflexes such as the Babinski sign (discussed below). Less well recognized is the nonpyramidal motor syndrome, such as is seen in caudate or premotor cortical lesions: clumsiness, decreased spontaneous use of affected limbs, apparent weakness but production of full strength with coaxing. Mild degrees of impairment can be elicited with the pronator test by seeking pronation of the outstretched supinated arms; the forearm rolling test, by asking the patient to roll the forearms around each other first in one direction then in the other, looking for one side that moves less thus appearing to be an axis with the other circling around; or fine finger movements, with the hands resting facing up on the thighs, the patient touching each finger to the thumb in turn and repeatedly.

Muscle tone can be increased not only in the pyramidal fashion just described but also as a manifestation of extrapyramidal or diffuse brain disease. In the latter case, paratonic rigidity, or Gegenhalten, is manifested by an erratic, "pseudoactive" increase in resistance to passive movement. The fluctuating quality of the resistance reflects the presence of both oppositional and facilitatory aspects of the patient's response to passive movement. The facilitatory aspect can be evoked by repeatedly flexing and extending the patient's arm at the elbow, then abruptly ceasing and letting go when the arm is extended; the abnormal response, facilitory paratonia, is for the patient to continue the sequence by flexion. In the case of extrapyramidal disease, tone is increased in both extensors and flexors and throughout the range of movement, so-called lead-pipe rigidity. The "cogwheel" or ratchety feel to the rigidity is imparted by a coexisting tremor and is not intrinsic to the hypertonus; when paratonic rigidity co-occurs with a metabolic tremor a delirious patient may mistakenly be thought to have Parkinson disease. Extrapyramidal rigidity can be amplified by activation techniques, such as movement of the limb contralateral to the one being examined (the Froment maneuver). However, this can occur bilaterally in healthy subjects so is most informative if unilateral.

Gait should always be tested, if only by focused attention to the patient's entering or leaving the room. Attention should be paid to the patient's station, postural reflexes, stride length and base, and turning. Postural reflexes can be assessed by asking the patient to stand in a comfortable fashion, then pushing gently on the chest or back, with care taken to avoid a fall. Gait should be stressed by asking the patient to walk in tandem fashion and on the outer aspects of the feet. This may reveal not only mild ataxia (representing cerebellar vermis dysfunction) but also asymmetric posturing of the upper extremity (in nonpyramidal motor dysfunction).

Akinesia is manifested by delay in initiation, slowness of execution, and

difficulty with complex or simultaneous movements. Mild akinesia may be observed in the patient's lack of spontaneous movements of the body while sitting, or of the face, or elicited by asking the patient to make repeated large amplitude taps of the forefinger on the thumb (looking for decay of the amplitude). Akinesia is characteristically accompanied by rigidity. These plus rest tremor and postural instability represent the core features of the parkinsonian syndrome, seen not only in idiopathic Parkinson disease (IPD) but in several other degenerative, "Parkinson-plus" disorders such as progressive supranuclear palsy and multiple system atrophy as well as in vascular white-matter disease. Rest tremor is less common in these other disorders than in IPD.

Dystonia is sustained muscle contraction with consequent twisting movements or abnormal postures. Typically dystonia in the upper extremity is manifested as hyperpronation, in the lower extremity as inversion of the foot with plantar flexion. Dystonia may occur only with certain actions, such as writer's cramp; focally, such as blepharospasm or oculogyric crisis; or in a generalized pattern, such as torsion dystonia associated with mutations in the DYT1 gene. The symptoms and signs often do not comport with a naïve idea of how things should be in organic disease; only expert knowledge will suffice for recognition. For example, a patient with early torsion dystonia may be able to run but not walk, because the latter action elicits leg dystonia. Or a patient with intense neck muscle contraction may be able to bring the head to the midline by a light touch on the chin, a *geste antagoniste* diagnostic for dystonia.

Tremor is a regular oscillating movement around a joint. In rest tremor, the movement occurs in a relaxed, supported extremity and is reduced by action. Often an upper extremity rest tremor is exaggerated by ambulation. The frequency is usually 4 to 8 Hz. This is the distinctive tremor of Parkinson disease. In postural tremor, sustained posture elicits tremor. It may be made more visible by placing a sheet of paper over the outstretched hand. Hereditary essential tremor presents as postural tremor, predominantly in upper extremities but also at times involving head, jaw, and voice. A coarse, irregular, rapid postural tremor is often seen in metabolic encephalopathy. In intention tremor, the active limb oscillates more prominently when approaching its target, such as touching with the index finger the examiner's finger. Maximizing the range of the movement increases the sensitivity of the test. Intention tremor is one form of kinetic tremor, which is tremor elicited by movement; another sort of kinetic tremor is that elicited by a specific action, such as writing tremor or orthostatic tremor upon standing upright. The examiner can characterize tremor by observing the patient with arms supported and fully at rest, then with arms outstretched and pronated, then with arms abducted to 90 degrees at the shoulders and bent at the elbows while the hands are held palms down with the fingers pointing at each other in front of the chest.

The patient should also be observed during ambulation. Anxiety exaggerates tremor; this normal phenomenon, for example when the patient is conscious of being observed, should not be mistaken for psychogenesis. A good test for psychogenic tremor relies on the fact that although organic tremor may vary in amplitude, it varies little in frequency. A patient can be asked to tap a hand at a frequency different from the tremor frequency; if another tremulous body part entrains to the tapped frequency, psychogenic tremor is likely.

Choreic movements are random and arrhythmic movements of small amplitude that dance over the patient's body. They may be more evident when the patient is engaged in an activity such as ambulation. When the movements are of large amplitude and forceful, the disorder is called ballism. Ballistic movements are usually unilateral.

Myoclonus is a sudden, jerky, shock-like movement. It is more discontinuous than chorea or tremor. The negative of myoclonus is asterixis, a sudden lapse of muscle contraction in the context of attempted maintenance of posture. Both phenomena, but more sensitively asterixis, are common in toxic metabolic encephalopathy (not just hepatic encephalopathy). Asterixis should be carefully sought by observation of the patient's attempt to maintain extension of the hands with the arms outstretched, because it is pathognomonic for organic disease and is never seen in acute idiopathic psychosis or other nonorganic disorders. Myoclonus is additionally an important feature of nonconvulsive generalized status epilepticus, Hashimoto encephalopathy, and Creutzfeldt–Jakob disease. Unilateral asterixis may be rarely seen in parietal, frontal, or (most often) thalamic structural disease.

Tics are sudden, jerky movements as well, but they may be more complex than myoclonic jerks and are subjectively characterized by an impulse to perform the act and a sense of relief for having done so (or mounting tension if restrained from doing so). Compulsions are not easy to differentiate from complex tics; the tiqueur may, like the patient with compulsions, report deliberately performing the act. Repetitive behavior superficially like compulsions may occur in organic disease but represent environment-driven behavior rather than having the same subjective structure as compulsive behavior. For example, a patient with frontal disease may repeatedly touch an alluring object without an elicitable subjective impulse and without anxiety if separated from the object. Organic obsessions and compulsions occur as well and have been associated with globus pallidus lesions among others.

Akathisia is defined by both its subjective and its objective features. The patient exhibits motor restlessness, for example, by shifting weight from foot to foot while standing, and expresses an urge to move. At times psychotic or cognitively impaired patients cannot convey the subjective experience clearly, and the examiner must be alert for the objective signs in

order to differentiate akathisia from agitation due to anxiety or psychosis. The complaints and the signs in akathisia are referable to the lower, not the upper, extremities; the anxious patient may wring his or her hands; the akathisic patient shuffles his or her feet. Myoclonic jerks of the legs may be evident in the recumbent patient. The phenomenon occurs in IPD and with drug-induced dopamine blockade, but also rarely with extensive frontal or temporal structural lesions.

Ataxia is a disorder of coordinating the rate, range, and force of movement and is characteristic of disease of cerebellum and its connections. In the limbs, dysmetria represents disordered determination of the distance to be moved, so that the patient overshoots or undershoots the target; if the reaching limb oscillates in the process, the clinician observes intention tremor. Asking the patient to touch the examiner's finger, then his or her own nose, tests this system. Accurately touching one's own nose with eyes closed requires both cerebellar and proprioceptive function. Eye movements also may be hypermetric or hypometric. The patient's difficulty in performing rapid alternating movements, such as supination/pronation of the hand or tapping of the foot, is called dysdiadochokinesia. The failure of coordination of movement is also demonstrated by loss of check, which should not be elicited by arranging for the patient to hit himself or herself when the examiner's hand is removed. In the normal situation, if the outstretched arms are tapped, only a slight waver is produced; the ataxic patient fails to damp the movement. Gait may be affected by midline cerebellar (vermis) disease in the absence of limb ataxia, which is related to cerebellar hemisphere disease. Gait is unsteady, with irregular stride length and a widened base. (In the normal subject, the feet nearly touch at their nearest point; even a few inches of separation represents widening of the base.) Gait and limb ataxia may be complemented by cerebellar dysarthria (described in Table 2.1–1) and by eye movement disorders, including nystagmus (usually gaze-paretic), slowed saccades, saccadic pursuit, and gaze apraxia.

The catatonic syndrome has been variously defined. The core of the syndrome is a mute motionless state; variably added are abnormal movements including grimacing, stereotypy, echopraxia, and catalepsy. The latter, known also as *flexibilitas cerea* (waxy flexibility), refers to posturing of a limb in the position in which it is placed by the examiner, or in some other unnatural position. It is not seen in all or even most cases of catatonia, and it can be seen apart from the catatonic syndrome in patients with contralateral parietal lesions (as described below as the avoidance sign of parietal disease). Catatonic excitement refers to the sudden eruption into over-activity of a catatonic patient; this probably usually represents psychotic mania. The catatonic syndrome occurs in the course of schizophrenia or mood disorder, or without other psychopathology as idiopathic catatonia, or in the setting of acute cerebral metabolic or

structural derangements. In the latter case it is best thought of as a nonspecific reaction pattern, such as is delirium, requiring a comprehensive clinical and laboratory evaluation to seek the cause of the behavioral disturbance. An important instance is catatonia as part of the neuroleptic malignant syndrome, the diagnosis of which requires exclusion of other metabolic encephalopathy, notably systemic infection. Catatonia is thus a medical emergency, requiring prompt attention to diagnostic evaluation as well as supportive care (fluids, nutrition, measures to avoid complications of immobility including venous thrombosis).

Motor sequencing tests assay function of premotor cortical areas and striatum and are related to deficits in executive cognitive function seen with dysfunction of the dorsolateral prefrontal loop. The ring/fist test involves asking the patient to alternate between making a ring with his thumb and first finger and making a fist with the same hand ("ring, fist, ring, fist"). The abnormal response is perseveration of one or the other posture or disorganization of the sequence. At times, patients will be unable to perform the correct series even when repeating the verbal cues aloud. A more complex alternation is between striking the table gently with the fist, then the edge of the hand, then the palm: ("fist, edge, palm, fist, edge, palm"). A different approach is to ask the patient to extend the arms, make a fist with one hand while keeping the other hand flat, then switching hands. The abnormal response has the patient ending up with two fists, or two palms, outstretched.

Review of the material in this section on the motor system will reveal how much can be accomplished in the neurological examination by asking the patient to stretch out his or her arms. With a few additional maneuvers (tapping the outstretched pronated hands, supinating them and asking the patient to close his or her eyes, then asking the patient to touch his or her nose with the eyes still closed, then asking the patient to perform the alternating fists test) all of the following can be assessed in a matter of a minute or so: postural and intention tremor, loss of check, asterixis and myoclonus, a pronator sign, dysmetria, and motor sequencing. Doing this, plus testing muscle tone, plus observing the patient's natural and stressed gait, plus checking tendon jerks and abnormal reflexes, takes just a few minutes and does not elucidate disorders of muscle, nerve, and spinal cord but represents a rather extensive assessment of the central organization of the motor system.

Abnormalities of Sensation. Disorders of sensation are sometimes difficult to assess reliably in patients with cognitive and behavior disorders. Nonetheless, several points should be familiar to the neuropsychiatrist. Distal loss of sensation, often accompanied by loss of ankle jerks, is characteristic of peripheral neuropathy. Often all modalities of sensation are disturbed. If proprioception is sufficiently severely reduced, a Romberg

sign will be present. The Romberg sign means that closing the eyes produces substantial impairment of balance; it is elicited by asking the patient to stand, allowing the patient to seek a comfortably balanced position, then asking the patient to close the eyes (ensuring against a fall).

Loss of sensation from sensory cortex injury is classically limited to complex discriminations, such as graphesthesia (recognizing numbers written on the palm), stereognosis (identifying unseen objects in the hand), and two-point discrimination (telling whether the examiner is touching with one or two points, as these come closer together in space). However, patients with parietal stroke may have a pseudothalamic sensory syndrome (with impairments in elementary sensory modalities and subsequent dysesthesia) or other anomalous patterns of sensory loss. At times these patients will present with pseudomotor deficits: ataxia, fluctuating muscle tone, and strength (dependent in part on visual cueing), "levitation," and awkward positioning of the arm contralateral (or at times ipsilateral) to the lesion. In the acute phase, the combination of deficits can amount to motor helplessness. These deficits result from the loss of sensory input to regions in which motor programs arise. The lessons here are that "cortical" sensory deficits should be sought if there is a question of cortical involvement, and that more dramatic or unusual sensory abnormalities may also occur with cortical lesions.

Soft Signs. An extensive literature about the "soft signs" of neurological dysfunction is difficult to comprehend because of the varied definitions and batteries used in the various studies. Most of the signs sought in these batteries are discussed in this section under their more specific headings, such as graphesthesia under abnormalities of sensation and the alternating fist (Oseretsky) test under motor sequencing. From the corpus of test batteries, a few simple maneuvers can be extracted that can contribute to the neurological examination of the patient with a mental presentation.

While the patient is touching each finger to the thumb, as described above in the section on weakness, the examiner can watch the opposite hand for mirror movements. Obligatory bimanual synkinesia is seen specifically in disorders of the pyramidal pathways, such as the Klippel–Feil syndrome, and in agenesis of the corpus callosum, but also in putative neurodevelopmental disorders such as schizophrenia. Asking the patient, with eyes closed, to report whether the examiner is touching one or the other hand (with the patient's hands on the patient's lap), or one or the other sides of the face, or a combination, makes up the face–hand test. The examiner touches the left hand and right face simultaneously. If the patient reports only the touch on the face (i.e., extinguishes the peripheral stimulus), then the examiner can prompt (once), "Anywhere else?" Then the examiner touches the right hand and left cheek, left hand and left

cheek, right hand and right cheek, both hands, and both cheeks. Extinction of the peripheral stimulus is the pathological response and has been associated with schizophrenia and dementia.

Abnormal Reflexes. The Babinski sign is the shibboleth of the neurological examination. It should be elicited by stroking the lateral aspect of the foot from back to front, with the leg extended at the knee, using a pointed object such as an orange stick or a key. The response of extension of the great toe with or without fanning of the other toes indicates corticospinal tract disease. Two confounding factors in assessment of the Babinski sign are the striatal toe and the plantar grasp. The striatal toe is extension of the hallux without fanning of the other toes or a flexion synergy in the other muscles of flexion of the leg. It may occur spontaneously or upon elicitation in patients with Parkinson disease in the absence of evidence of pyramidal dysfunction. The plantar grasp, the equivalent of the familiar palmar grasp, may mask an extensor response to lateral foot stimulation when stimulation in the midfoot brings about flexion of the toes.

Other important reflexes for the neuropsychiatrist are:

- Myerson sign, a failure to habituate to regular, one per second, taps on the glabella (with the tapping hand outside of the patient's visual field), present in parkinsonism and diffuse brain disease;
- Hoffmann sign, flexion of the thumb with snapping of the distal phalanx of the patient's middle finger, an upper extremity sign of pyramidal dysfunction though sometimes present bilaterally in normal subjects;
- Grasp, flexion of the fingers with stroking of the patient's palm toward the fingers during distraction, despite instructions to relax, associated with disease of the contralateral supplementary motor area;
- Avoidance, extension of the wrist and fingers to the same stimulus as the grasp, a less well-known sign that points to contralateral parietal cortex abnormality.

Several other "primitive reflexes" are less specific, in that they are commonly present in the normal subject and are thus less useful for diagnostic purposes. These include the suck, snout, and palmomental reflexes.

FOCAL NEUROBEHAVIORAL SYNDROMES

The idea that the brain is regionally specialized had a difficult gestation in the 19th century, and the key to its acceptance lies with recognition of the effects of focal brain lesions. At the end of the 18th and early in the 19th centuries, phrenology drew adherents to the claim that personality traits could be inferred by inspection of the cranium. This claim was faulty, but

phrenology had an underlying theory that was an important step forward for the brain sciences; in particular, the beliefs that the brain was the organ of the mind and that mind could be fractionated into functions gave impetus to the development of neuroscience in a modern form. In the middle of that century, the gradual realization that aphasia occurred with damage to specific areas of the left hemisphere was another crucial step. The subsequent identification of numerous syndromes of localized damage —syndromes such as apraxia, agnosia, visuospatial impairment in its various forms, and so on—is a fascinating story of astute and painstaking clinicopathological, and later clinicoradiological, correlation.

Patients' introspective access to their deficits may be limited. Much of cognitive processing is unconscious, not in the sense of being excluded from awareness by motivated defense, but in the sense that it is not even in principle open to introspection. Jonathan Miller, in his television show *The Body in Question*, displayed this point by asking passersby, in a man-in-the-street interview, "Sir, where is your spleen?" No one can say from introspection where the spleen is. The same is true of much of cognitive processing. Explanations provided by patients may be confabulations that fill in such introspective gaps, in a situation where the brain is functioning abnormally in a way not foreseen in its design. In neuropsychiatry, subjective experience and behavior are separate explicanda. The patient's appraisal of his or her own situation is always relevant to collaboration with treatment and its outcome and should be explored in every clinical encounter.

Dementia

Although characteristically considered a syndrome of "global" cognitive impairment, implying global or diffuse brain dysfunction, in fact each dementing disorder produces a distinct pattern of brain pathology and corresponding pattern of cognitive dysfunction. For this reason, and against tradition, dementia will be discussed under the rubric of focal neurobehavioral syndromes.

In Alzheimer disease, the earliest neuropathologic abnormality is characteristically medial temporal accumulation of plaques and tangles, initially in entorhinal cortex and subiculum (the input and output zones of hippocampus). The disease progressively involves association cortices in temporoparietal and prefrontal regions. This burden of pathology determines the characteristic early memory impairment with ensuing anomia, failure of grasp, and coarsening of personality. On occasion, Alzheimer disease pathology is predominantly posterior, with concomitant predominance of visuospatial impairment in the clinical course, or predominantly frontal, with executive dysfunction and characteristic behavioral alterations. By contrast, in frontotemporal dementia, the earliest disease manifestations are pathologically in the frontal or temporal

cortex, clinically presenting as primary progressive aphasia, semantic dementia, or a frontal apathy or disinhibition syndrome. In none of these situations is a view of dementia as a "global" impairment of brain function justified by the clinical or pathological facts; rather, selective disruption of anatomic networks corresponds to symptomatic features.

Extensive subcortical white and gray matter damage due to small vessel disease is a common cause of dementia, and in this situation the clinical picture is dominated by slowed mental processing, forgetfulness with relative preservation of recognition memory (as opposed to free recall), and executive cognitive dysfunction. "Strategically" located single infarctions can also produce dementia. These strokes can involve left angular gyrus, genu of internal capsule, and (perhaps most commonly) medial thalamus. The thalamic and internal capsule strokes may produce cognitive impairment by interfering with frontal networks.

Delirium

Classically a syndrome of "global" brain dysfunction due to toxic metabolic infectious encephalopathy, delirium may also point to focal brain disease. Delirium may be caused by infarction in the right posterior superior temporal gyrus due to occlusion of the inferior division of the right middle cerebral artery, or by infarction in the inferior temporo-occipital cortex, on the left or bilaterally, due to posterior cerebral artery occlusion. In both instances focal neurological signs may be limited to a visual field cut or absent entirely. Finding bilateral asterixis or multifocal myoclonus strongly indicates a toxic-metabolic brain derangement, and the history and physical examination should provide pointers for the essential laboratory confirmation of the abnormality. Features of the mental state are of little use in determining the cause of the syndrome, except that agitation is far more common in certain disorders, such as substance withdrawal, hypoxia, and the syndromes of left posterior cerebral artery stroke or of right middle cerebral artery territory stroke with involvement of the temporal lobe.

Aphasia

Acquired impairment of lexical or syntactic performance is termed aphasia. Lexicon and syntax do not exhaust the domain of language, and attention is devoted below to prosody and discourse pragmatics. At the bedside, the clinician should be able to distinguish language impairment from other sources of abnormal discourse (such as psychosis), delineate the features of abnormality in the patient's linguistic function, and tentatively identify the locus of brain injury.

A simple distinction between "expressive" and "receptive" defects has some power to distinguish between anterior and posterior lesion sites, but it is not in current use in aphasiology because most aphasiogenic lesions

produce some impairment in both production and comprehension of language, and these impairments are of multiple sorts.

A widely accepted approach to examination and classification in aphasia identifies six domains for elucidation: spontaneous speech, naming, comprehension, repetition, reading, and writing. Attention to spontaneous speech reveals dysfluency and word-finding difficulties. The dysfluent speaker produces shorter phrases and utterances without a natural flow. Substantives (nouns and verbs) may be preserved at the expense of function words (such as prepositions) and grammatical morphemes (such as tense endings), leading to agrammatical utterances that are nonetheless relatively information-rich. Lesions disrupting fluency are characteristically anterior in the left hemisphere or involve putamen. Naming performance requires the adequate functioning of a network including posterior temporal, temporoparietal, and inferior frontal sites. This is ordinarily tested by confrontation ("What do you call this?"), which can be conveniently done using body parts or common items at the bedside. Naming from description ("What do you call the vehicle that travels underwater?") is an alternative mode of testing, particularly useful for visually impaired or agnosic patients. Comprehension is tested best using probes with minimal demand on output, so "yes/no" questions ("Does a stone sink in water?") are better than motor commands, which may be impaired by concurrent apraxia. Impairment of comprehension results from posterior temporal lesions. Disordered repetition is disclosed by asking the patient to produce progressively longer utterances reiterating the examiner: "Airplane, he and she are here, the spy fled to Greece, ..." Repetition may be surprisingly spared (the so-called transcortical aphasias) or disproportionately impaired (conduction aphasia). The latter depends on lesions of insula or external capsule. Reading comprehension (not reading aloud, a different skill) tests comprehension with a different input modality from aural comprehension, and some patients will show significant dissociations. Similarly writing tests output in a different modality from speech. Writing is a particularly sensitive probe for the anomia seen in early Alzheimer disease and the disorganization seen in delirium.

A classification of the aphasias using the data from an examination, as just outlined, is shown in Table 2.1–2. Clinicians recognize, however, that many patients will not fit well into the categories created by this scheme.

Ideomotor apraxia commonly occurs together with aphasia. This is a disorder of performance of skilled movements to command in the absence of explanatory elementary sensory and motor disturbances. Oral apraxia is revealed by the patient's incapacity to show, for example, how to blow out a match or lick an envelope. Limb apraxia is shown by the patient's incapacity to show, for example, how to wave good-bye or use a hammer or screwdriver. Patients with these deficits may nonetheless be able to follow

whole body commands: "Show me how a boxer stands," for example. Patients rarely complain of apraxic deficits, in part because they are artifacts of the examination, in the sense that they may not be present in utilization of real-world items.

Attention

Several related phenomena cluster under the clinical description of attentional disorders. At the most fundamental level, alertness represents a continuum ranging from coma to normal wakefulness. The clinician faced with a patient who is less than fully alert should quantify the disorder by assessing the patient's response to a graded series of probes: does the patient orient to the examiner's presence in the room, what does the patient do when his or her name is called or when touched or when shaken or when a (harmless) painful stimulus is applied, and so on. These responses should be recorded in detail, rather than summarized by an ambiguous term such as "lethargic."

Alert patients may show deficits in sustained attention to external stimuli (vigilance) or internal stimuli (concentration). Attentional deficits of these sorts are characteristic of delirium. Vigilance can be assessed with a bedside adaptation of a continuous performance task, for example, by asking the patient to lift a hand each time the examiner says the letter "A," or to raise the difficulty, each time the examiner says the letter "A" after the letter "D." The examiner then produces a series of random letters at a deliberate and steady rate over an extended period of time. Any error of omission or commission represents a failure. By asking the patient to count from 20 to 1 or give the days of the week or the months of the year in reverse, the examiner can appraise concentration. Digit span—asking the patient to repeat a list of numbers spoken at a slow, steady rate without separation into chunks—is a classic test of attention; the lower limit of normal for digit span is five.

A "higher" level of attentional function is the capacity to manipulate information kept in consciousness over a short period of time—a test of working memory. An excellent example is alphanumeric sequencing. The patient is asked to alternate between numbers and letters; the examiner provides "1-A-2-B-3-C" as a model then allows 30 seconds for the patient to start at 1 and give as many alternations as possible. If only the number of correct alternations is counted (ignoring errors), the lower limit of normal is 15. A comparable, simple task is alphabetizing the letters of the word "world" (or any similar word). Working memory is known to require intact processing in dorsolateral prefrontal cortex.

Hemineglect is a focal disorder of attention, almost always of left hemispace in a patient with acute right hemisphere disease. Most characteristically the lesion is parietal, but distinguishable patterns of hemineglect occur with frontal and cingulate lesions. Gross neglect can be

recognized in the patient's ignoring, or even denying the ownership of, the left limbs or failing to attend to objects and people in left hemispace. Subtler degrees of neglect can be elicited by presenting an array in which the patient must search into both hemifields to point to items, for example, all the yellow dots in a stimulus card with dots of varied colors to both left and right.

Table 2.1–2.
Aphasia Syndromes

Aphasic Syndrome	Spontaneous Speech	Repetition	Naming	Aural Comprehension	Reading for Comprehension	Writin
Global	Impaired	Impaired	Impaired	Impaired	Impaired	Impair
Broca	Dysfluent, effortful, agrammatic	Impaired	Impaired	Spared	Spared	Impair
Wernicke	Fluent, paraphasic, absence of substantive words	Impaired	Impaired	Impaired	Impaired (but not always to same degree as aural comprehension)	Impair
Conduction	Fluent, paraphasic with phonemic errors	Impaired	Impaired	Spared	Spared	Impair
Transcortical motor	Dysfluent	Spared	Impaired	Spared	Spared	Impair
Transcortical sensory	Fluent, paraphasic	Spared	Impaired	Impaired	Impaired	Impair
Mixed transcortical	Dysfluent	Relatively spared	Impaired	Impaired	Impaired	Impair
Anomic	Fluent,	Spared	Impaired	Spared	Spared	Spared

In the phenomenon of hypermetamorphosis, included as part of the Klüver–Bucy syndrome of bilateral anterior temporal damage, animals or patients exhibit an increased level of attention to individual items in the environment.

Amnesia

The term *memory* is used in several ways by clinicians and psychologists. The amnestic syndrome features impairment of learning of new material (anterograde amnesia) and a variable period of impaired recall prior to the onset of the syndrome (retrograde amnesia). It is due to damage to the hippocampus or to the anterior thalamus (including mamillothalamic tract). Memory proper is distinguished from retention in consciousness of material over the course of a few seconds, which may be called "working" or "iconic" or "short-term" memory. This function is spared in the amnestic syndrome because it depends on frontoparietal mechanisms distinct from the hippocampal and thalamic pathways damaged in amnesia. Deficits due to hippocampal and thalamic lesions are dependent on the lateralization of the damage, left-sided damage producing verbal and right-sided damage figural memory impairments.

A distinction between free recall and recognition memory is of neuropsychological significance and generally can be made adequately if imperfectly at the bedside. Hippocampal and thalamic patients show accelerated forgetting so that cues (such as providing the semantic category) are relatively ineffective in aiding recall. Recognition memory is always better than free recall on an absolute scale; exaggeration of the disparity (i.e., sparing of recognition memory) is characteristic of memory impairment due to dysfunction of frontal mechanisms of effortful search.

In addition, frontal patients show impairments of memory for the temporal context or source of information. This deficit is probably relevant to the occurrence of confabulation. Spontaneous confabulation occurs in only a minority of amnestic patients and depends on ventromedial frontal damage.

Memory is so commonly impaired in brain disorders that it should be tested in all patients undergoing neuropsychiatric evaluation. Recall of a test phrase (for example, a name and address) over a several minute distraction is a valid and simple screening test. However, more detailed analysis of memory is necessary in patients with disorders likely to affect memory mechanisms, including (among many others) head injury, epilepsy, and dementing disorders. Testing should include both verbal and nonverbal material. For example, testing recall over several minutes' delay of three words and three shapes, or three words and three pointed directions, or three object names and the locations in the room in which

those objects were hidden is easily performed. Further, the examiner should be prepared with cues, including appropriate (incorrect) foils, to assess sparing of the capacity to make use of cues in frontal memory impairment. For example, if one of the words provided is "piano," the examiner could cue, "One was a musical instrument" and further provide "guitar, piano, violin" as multiple-choice options. Only rough inferences can be drawn from this bedside assessment, as compared with formal neuropsychological evaluation.

Apart from patients with persisting amnestic syndromes, the neuropsychiatrist may be presented with patients who suffered an amnestic state transiently, or rarely may see one during a transient amnestic state. The syndromes of transient global amnesia and transient epileptic amnesia, and their differentiation, have been fully described and require thorough history taking, neuropsychological evaluation, and electroencephalogram (EEG) recordings. The neuropsychiatrist should also know that amnesia for criminal offenses is common; certainly it is not confined to those who committed a crime while in a delirious, ictal, or postictal state, as is sometimes claimed for legal reasons.

Visuospatial Dysfunction

The requirement to test visual as well as verbal memory has just been mentioned. Drawing and copying tasks can further the assessment. Copying intersecting pentagons (as in the Mini-Mental State Examination) or the shapes used in the memory task begins the assessment. With more complex figures, failures with a slavish element-by-element strategy are characteristic of patients with right hemisphere damage, as is neglect of the left side of the stimulus.

The variety of disorders of higher visual function has already been mentioned in describing the complex structure of visual association cortices. Prosopagnosia is a defect in recognition of faces. Such a defect may be obvious from the history or may be a more subtle abnormality; it can be spotted at the bedside, albeit insensitively, with the use of a few pictures of famous people. Defects of topographic skill, although rarely presenting in an isolated form, also occur with right hemisphere dysfunction. The patient can be asked to describe a route between familiar places or a geographical question thought to be within premorbid capacities ("If you're going from New York to Los Angeles, is the Atlantic Ocean in front of you, behind you, to your left, to your right?").

The incapacity to grasp in attention multiple visual objects at once is known as simultanagnosia. The patient may fail in describing a complex visual scene by reporting only a single, perhaps peripheral, element. Together with psychic paralysis of gaze (inability to direct gaze voluntarily, or ocular apraxia) and optic ataxia (misreaching under visual guidance), it makes up Balint syndrome, the archetypal disorder of the dorsal visual

pathway. The patient with impairment of reaching under visual guidance should be examined reaching without visual guidance (e.g., pointing to parts of his or her own body with eyes closed) to confirm the defect.

Executive Cognitive Dysfunction

This term refers to initiation of cognition and action, their maintenance in the face of distraction, organized but flexible pursuit of goals, and self-monitoring with error correction. Executive processes are crucial in adaptive function, and performance in this realm is better correlated with real-world outcomes of brain-injured patients than are many other domains traditionally analyzed in neuropsychology or many psychosocial variables. Bedside exploration of executive function is of central importance in the neuropsychiatric examination. Analysis of behavioral disturbance and neuropsychological deficits in patients with cerebral injury suggests that multiple dissociable processes compose executive cognitive function, and certainly these processes are instantiated by anatomically distributed systems. Curiously, however, many different tasks recruit a similar or identical set of regions in the mid-dorsolateral prefrontal cortex, the mid-ventrolateral prefrontal cortex, and anterior cingulate. Nonetheless, the clinical examiner must know that no single probe can screen for all dysfunctions.

Many aspects of executive function are illuminated by attention to the patient's performance of elements of the history taking and examination. Disinhibition may be noted in abnormalities of comportment during social interaction. Motor impersistence, the failure to sustain actions that can be initiated properly, may be noted in the patient's peeking when asked to keep the eyes closed, repeatedly looking back at the examiner's face when lateral gaze (especially to the left) is attempted, or failing to keep the arms extended or the tongue protruded when instructed to do so. Perseveration is the continuation of elements of past actions into present activity. Perseverative responses may be noted when testing naming or attention. Echopraxia, for example the patient's crossing the arms when the examiner (spontaneously) does so, even when some other behavior has been requested, can be observed during the interview and examination. Utilization behavior is an automatic tendency to make use of objects in the environment, for example, picking up a pen and starting to write, despite this behavior's being inappropriate to the setting.

More focused efforts to assess executive function are almost always indicated in the neuropsychiatric examination. Perseveration may be specifically sought in the motor sequencing tasks described above or with a sample of spontaneous writing. A tapping task with conflicting instructions may illuminate the inflexibility of goal-directed behavior that gives rise to perseverative responding. The patient is instructed to tap once if the examiner taps twice, and twice if the examiner taps once. The examiner

then taps on the table in a random series of one tap or two taps. This can be directly followed by a go/no-go tapping task, in which the patient is instructed to tap once if the examiner taps once, not at all if the examiner taps twice. Intrusions from the previous task's instructions represent perseverative responding; echopraxic responses (tapping just like the examiner) represent failures of inhibition. Inhibition of reflexive gaze can be tested during the examination of eye movements, as described above. Looking at the moving stimulus rather than in the opposite direction as instructed amounts to a visual grasp response and represents failure of inhibition. Spontaneous word-list generation ("Tell me all the animals you can think of," or "Tell me all the words that start with 's'") depends on the capacity for effortful search of semantic stores. A greater decrement in fluency to semantic cues ("animals") than to phonemic cues ("words with 's'") is seen in Alzheimer disease because of the degradation of semantic stores due to temporoparietal damage. Working memory can be assessed with the alphanumeric sequencing task described above.

Anatomic inferences from dissociations in performance on these tasks are limited. Go/no-go tasks depend on the integrity of orbitofrontal cortex, and other tasks on the dorsolateral prefrontal cortex and its circuit. Impersistence is associated with right hemisphere dysfunction. A further dissociation is between executive cognitive impairments and personality change in frontal injury. Especially with orbitofrontal lesions, executive function can be spared even in the face of grave alterations in emotional and comportment; the two domains cannot simply be considered two sides of the same coin. Nonetheless, it bears repeating that neuropsychiatric examiners always should consider executive cognitive function in their formulation of cases.

Disordered Mood and Emotion

Several syndromes of disordered emotion in organic disease can be delineated. Disturbances of recognition and expression of emotion with right hemisphere lesions have already been mentioned. Patients and their families are rarely aware of these deficits and do not complain of them; rather the examiner must recognize the deficit and figure it into a formulation of the patient's social and functional decline. Impairment of prosodic expression should not be mistaken for depressed affect. Testing of affective prosody can be undertaken at the bedside without special equipment. The examiner should ask the patient to say emotionally loaded sentences in an emotional manner, expressing surprise, fear, pleasure, and anger. People vary considerably in their native acting talents, and the range of normal performance is wide. The examiner also can utter neutral sentences in various emotional tones: "I am going to the store," stated with surprise, and so on, with the examiner's face turned away from the patient to avoid providing a second input channel. Patients should be able to

recognize the affect. Separately, if testing materials are available, the examiner can assess the patient's capacity to identify emotions in visual scenes and facial expressions. Lesions that involve both limbic and heteromodal cortices in the right hemisphere especially impair performance in recognizing emotional facial expressions.

Pathological laughing and crying are stereotyped displays of affect incongruent with inner experience and elicited by inappropriate, nonemotional, or inadequate stimuli. The examiner may, in the extreme, be able to elicit full displays of affect by waving a hand in front of the patient's face. The patient is often embarrassed by the pathological expression of affect. Features of pseudobulbar palsy are often present in these patients. In these cases, damage to a volitional control system has disinhibited a network subserving emotional display. This network includes amygdala, hypothalamus, periaqueductal grey, and dorsal tegmentum. A broader form of affective dysregulation, which may be called emotionalism or emotional lability, is commonly seen, usually in the direction of tearfulness. Patients report that they are more emotional than previously and that the tears are sudden, unexpected, and uncontrollable. However, they are generally congruent with the patient's subjective state. Such patients are often cognitively impaired. The rare phenomenon of *fou rire prodromique* (mad prodromal laughter) presages acute vascular lesions of the brainstem or thalamus.

Apathy is an emotional disturbance marked by reduction of affect and motivation. Goal-directed behavior is reduced, and emotional responses are lacking. The distinction from depression is crucial: patients do not report negative emotional states or ideational content. Although they may meet criteria for depression because of the loss of interest in activities, they are mentally empty rather than full of distress. Recognizing apathy rather than mistaking it for depression may imply treatment with different pharmacological agents, for example, use of dopamine agonists. Euphoria refers to a persistent and unrealistic sense of well-being, without the increased mental or motor rate of mania. Although often mentioned in connection with multiple sclerosis (MS), it is unusual and almost always associated with extensive disease and substantial cognitive impairment.

The Klüver–Bucy syndrome, as described in the captive monkey, includes reduction in aggression (tameness), excessive and indiscriminate sexual behavior, hypermetamorphosis (forced attention to environmental stimuli), and hyperorality (mouthing nonfood items). This mixture of emotional, perceptual, and motivational changes is dependent on bilateral damage to amygdala. In human patients, pathologies including trauma, herpes simplex encephalitis, and frontotemporal dementia can produce the syndrome, usually in partial form.

Depression is common in patients with brain diseases including stroke, MS, traumatic brain injury, and Parkinson disease. Certainly this is in part

a reaction to altered circumstances and distressing disability. Nonetheless, the syndromal nature of the depressed state and its imperfect correlation with measures of disability have prompted extensive efforts to seek anatomic correlations. Converging evidence leads to a model of alterations in a distributed network involving neocortical and limbic elements. In particular, a dorsal compartment involving dorsolateral prefrontal cortex, inferior parietal cortex, and the dorsal and posterior portions of cingulate gyrus show underactivity in the depressed state; these regions are thought to mediate the cognitive alterations and impairments of depression. Inversely, a ventral compartment containing anterior insula, subgenual cingulate, hippocampus, and hypothalamus are overactive; these elements are thought to mediate somatic ("vegetative") features of the depressed state. Interactions between the two compartments are mediated through the thalamus, basal ganglia, and especially rostral cingulate.

Mania is substantially less common than depression after brain injury. Mania is associated with right-sided lesions involving paralimbic cortices in orbitofrontal or basotemporal regions or subcortical sites in caudate or thalamus. Some evidence suggests that subcortical lesions are more likely to produce a bipolar picture, and cortical lesions unipolar mania. As with depression, the abnormal mood state does not necessarily appear in immediate temporal association with the injury, so determining whether the mood disorder is organic or idiopathic is not always straightforward. The absence of a personal history of mood disorder is an obvious criterion, but the presence of a family history of mood disorder may mark a vulnerability factor not operative in the absence of the brain lesion. Age of onset is relevant, especially for mania: the onset of idiopathic mania after age 40 is rare.

A common issue in neuropsychiatric assessment is the patient with late-onset depression in whom evaluation reveals executive cognitive dysfunction and subcortical white-matter disease. This state of vascular depression is marked by the presence of vascular risk factors, notably hypertension, a tendency to psychomotor retardation and anhedonia and not psychosis or guilty ideation, and poor outcome with usual treatments. Some but not all of these patients have apathy rather than depression.

Abnormalities in Agency

Ordinarily, the person performing an action has the sense of being the one performing it. The prototype abnormality of this normal subjective sense is the "alien hand" phenomenon. Patients with parietal lesions may report a sense of strangeness of the hand, and the limb may exhibit levitation or avoidance reactions. More dramatically, with medial frontal or callosal lesions, the hand may engage in unwilled behavior (representing unilateral utilization behavior), or intermanual conflict may occur.

Abnormal Social Behavior

The multitude of behaviors exhibited in social interaction has, of course, multiple underpinnings. Several behavioral complexes, the neurobiology of which has come under scrutiny, can be observed in their abnormal form in patients and, at times, understood from an anatomic and physiological point of view.

Failures of empathic understanding are common in patients with frontal injury. These impairments result both from cognitive inflexibility in assessing complex social situations, especially in patients with dorsolateral prefrontal lesions, and from emotional impoverishment, especially in patients with orbitofrontal lesions. The intensity of social interaction manifested by patients with temporal lobe epilepsy may be due to deficits in social cognition or to a limbic lesion that reinforces social cohesiveness.

The capacity of human beings to understand the mental states of others —and thus to recognize not just another's goals or intentions but also the other's deceptions or pretense—has been termed mentalization or "theory of mind." Imaging and lesion data suggest that this capacity depends critically on the prefrontal cortex (particularly right medial prefrontal cortex adjacent to anterior cingulate gyrus), right temporoparietal cortices, and amygdala. Patients with isolated lesions of amygdala are rare, but deficits in theory of mind are seen in patients with frontal disease and may contribute to their social failure. Patients with right hemisphere lesions have a range of deficits in social interaction that may be characterized as a disorder of pragmatics. Although they may grasp the propositional content of language correctly, they mistake aspects of communication that require appraisal of the interlocutor's intent, for example, whether an utterance was intended as a joke. Pragmatic disorders due to frontal and right hemisphere damage may impair narrative coherence through verbosity, vagueness, and disregard for the listener's informational needs. Thus, for example, pronoun use may be syntactically correct, but the referents of pronouns are obscure to the listener. Although language itself is normal, the way language is embedded in social interaction is not.

Abnormal Beliefs and Experiences

Hallucinations are a common feature of diseases of the brain. Visual hallucinations in the absence of auditory hallucinations are suggestive of organic disease. Visual hallucinations may occur in a hemifield blind from cerebral disease, so-called release hallucinations. Visual hallucinations in the setting of visual impairment due to ocular disease, usually in the elderly, are known as the Charles Bonnet syndrome. The hallucinations are characteristically vivid images of living figures, and the patient is aware of their unreality. Other psychopathology is absent, but treatment aimed at the hallucinations is usually ineffective. Elaborate formed visual

hallucinations may occur with lesions of thalamus or upper brainstem, so-called peduncular hallucinosis. The symptoms are worse in the evening (crepuscular), and again the patient is aware of the unreality of the visual experiences. Prominent, early visual hallucinations in the context of progressive dementia suggest dementia with Lewy bodies.

Auditory hallucinations occur rarely with pontine lesions. More common are musical hallucinations in the setting of hearing impairment, akin to the Bonnet syndrome. Unilateral hallucinations are characteristically ipsilateral to the deaf ear. Olfactory hallucinations occur as a limbic aura in partial epilepsy, but they also occur in idiopathic psychiatric illness.

Palinopsia and palinacousis refer to persisting or recurrent perceptual experiences after the object is gone, in the visual and auditory domains, respectively. Lesions in association cortex—parieto-occipital and temporal, respectively—are responsible, although (for the visual sphere more than the auditory) drug toxicity is often the explanation.

The content of delusions may yield clues to causative organic disease and its nature. Most notably, misidentification delusions have been associated with dysfunction of face processing and clearly linked—in some but not all cases—to right hemisphere dysfunction. Misidentification of place is regularly associated with visuospatial and executive cognitive dysfunction. Misidentification delusions have been of special interest in cognitive neuropsychiatry. Perceptual recognition without a sense of familiarity (as in the Capgras syndrome and perhaps the nihilistic delusions of the Cotard syndrome) may reflect a disruption of visual-limbic connections. In a sense it is the reverse of déjà vu, which amounts to familiarity without perceptual recognition. However, many patients with misidentification delusions have no evidence of organic disease. Although such patients may have dysfunction of underlying mechanisms similar to patients with ascertainable organic disease, the similarity of clinical phenomena cannot be taken to prove an identity of mechanism.

Particular delusional themes may mark delirious thinking, such as a focus on danger to or harm to others, as opposed to the more self-centered constructions in idiopathic psychotic disorders. However, most delusions in patients with brain disease are of more banal nature, often with persecutory elements that bespeak cognitive failure (the theft of one's purse, for example, deriving from a failure of memory as to its location). Complexity or elaborateness of delusional ideation is associated with preservation of intellect, and delusions tend to become less complex with progression of dementia.

Paraclinical Investigations

Specialized paraclinical investigations form a major part of the neuropsychiatrist's toolbox. Sometimes patients are referred for

neuropsychiatric consultation when a routine investigation—such as a screening MRI or EEG—gives an unexpected abnormal result; the neuropsychiatrist is called on to assess the meaning of the finding in the psychiatric context. Late onset, atypical features or course, absence of a relevant psychiatric family history or presence of a relevant general medical history, presence of cognitive impairment, and presence of abnormal physical signs should prompt increased scrutiny for organic disease in a patient presenting psychiatrically. Catatonia should also lead to exploration for organic disease: Data from several centers suggest that one catatonic patient in five has explanatory organic disease.

Neuroimaging

Structural neuroimaging with computed tomography (CT) and later with MRI revolutionized practice in the clinical neurosciences. No longer was the organ of interest invisible within the carapace of the skull. CT relies on the differential absorption of x-rays by brain tissues and on the power of computerized methods to integrate data from multiple perspectives. The strengths of CT are its speed and its sensitivity to blood and bone. Thus for neuropsychiatric purposes, situations in which a patient cannot tolerate a prolonged imaging procedure may mandate CT. This problem often arises with an agitated demented or psychotic patient. Bony abnormalities, parenchymal deposition of calcium, and intracranial hemorrhage are particularly well assessed by CT. Such questions arise in the acute aftermath of trauma in particular.

The advent of MRI was an advance over CT in several respects. The anatomic resolution is substantially better, and the discrimination of white matter abnormalities exceptionally so. The capacity to display data from a single acquisition in multiple views—sagittal, axial, and coronal—allows improved anatomic understanding. Gradient-echo and susceptibility-weighted images sensitively reveal the sequelae of hemorrhage and may be useful in assessing the damage from trauma. Infusion of gadolinium for contrast enhancement is not necessary for delineation of nonvascular anatomic structures, such as is the goal in the case of atrophy or old stroke or trauma, but can identify areas of breakdown of the blood–brain barrier, such as in the meninges in meningitis or in parenchymal lesions of active multiple sclerosis, tumor, or acute stroke. Special imaging sequences should be used for the identification of cortical dysplasia or mesial temporal sclerosis. Volumetric MRI allows diagnosis by quantitative assessment of delineated brain structures, such as hippocampus, in the case of temporal epilepsy and potentially Alzheimer disease. One imagines the day in the near future when the scans will come (as electrocardiograms now do) with quantitative information routinely accompanying the analog image. MR angiography allows the delineation of medium and large vessels without the administration of contrast material, as is required for

conventional angiography. Stenosis of these vessels, such as the vessels of the neck, or the presence of vascular malformations or aneurysms is reliably ascertained. However, resolution is not sufficient to allow assessment of small vessels; thus some forms of vasculitis cannot be excluded with MR angiography and require contrast angiography.

Additional MRI sequences include diffusion-weighted imaging, which captures acute vascular injury; diffusion-tensor imaging, which discloses patterns of connectivity in white matter; and magnetization-transfer imaging, which promises even greater sensitivity to brain lesions than FLAIR imaging. Diffusion-weighted imaging is essential in evaluation of stroke, and high intensities in the cortex ("ribboning") are characteristic of Creutzfeld-Jakob disease (CJD). Magnetic resonance spectroscopy (MRS) is a method for analyzing the regional chemical composition of brain. The benefits of its ability to identify neuronal loss and glial proliferation are still under investigation, although in certain circumstances—such as distinguishing radiation necrosis from recurrent brain tumor—it is of proven utility. Dopamine-transporter imaging (DAT-scan) sensitively and specifically shows loss of nigrostriatal neurons and aids in the diagnosis of dementia with Lewy bodies and drug-induced parkinsonism.

Functional Neuroimaging

Four methods of functional neuroimaging are available: single-photon emission computed tomography (SPECT), positron emission tomography (PET), functional MRI, and brain mapping by quantitative electroencephalography. All are exciting research avenues, but the established clinical role for functional imaging is limited. All the techniques have a place in the presurgical evaluation of epileptic patients. SPECT and PET in the patient with frontotemporal dementia typically disclose the lobar nature of the dysfunction, although their value diagnostically over and above neuropsychological demonstration of the same phenomenon is uncertain. Similarly, under exactly which circumstances demonstrating bilateral temporoparietal hypoperfusion advances the diagnosis of Alzheimer disease is not yet clear. The demonstration of occipital hypoperfusion strongly supports a diagnosis of dementia with Lewy bodies. The evidence for other clinical uses of functional imaging is at present limited or anecdotal.

Electroencephalography

The expectation of the originators of EEG was that it would allow tracking of mental processes. This hope has not been realized. EEG does have the advantage over other clinically available brain imaging tools in that it reflects function at high temporal resolution, resolution corresponding to the time course of mental processing. Thus, at least from a research

perspective, measurement of brain potentials in relation to stimuli—the technique of evoked potentials—has the capacity to identify anomalous modes of cerebral processing. Recordings from electrode placements in subdural or cortical sites provide irreplaceable information about the origin and spread of epileptic discharges, but this invasive technique is justified only under exceptional circumstances. From today's practical point of view, scalp EEG can distinguish epileptic from nonepileptic spells and clarify the type of epilepsy. It can also distinguish acute encephalopathy from nonorganic psychosis.

Only 30 to 50 percent of patients with epilepsy show an epileptic abnormality on a single interictal waking EEG. With sleep deprivation, sleep during the recording, and repeated recordings, sensitivity improves to 70 to 80 percent. Anterior temporal electrodes add to the sensitivity and localizing power of the EEG, but nasopharyngeal electrodes, which are quite uncomfortable for the patient, do not provide additional sensitivity and are not recommended. A reasonable protocol would start with a routine EEG including anterior temporal leads; if this is negative but suspicion remains high, a second EEG with sleep deprivation can be undertaken. A third and fourth EEG may be useful, but the rate of discovery of abnormalities declines after that. Even then, some epileptic patients will not have been shown to have interictal abnormalities. At times ambulatory EEG is of use to ascertain the epileptic nature of undiagnosed events, but the restricted montage of the ambulatory equipment limits its utility. Hospitalization for video EEG recording may be essential for clarifying the nature of puzzling spells.

Delirium is characterized by slowing of the EEG, a finding never seen in acute idiopathic psychosis. This differential point can be decisive in a confusing clinical setting. However, EEG is not indicated as routine in the screening of psychotic patients. Among the dementing disorders, frontotemporal dementia is distinctive in having a normal EEG even as the clinical state becomes moderately severe. In CDJ, the EEG is always slow and may ultimately (not necessarily immediately) show the diagnostic feature of pseudoperiodic complexes.

Evoked potentials can identify abnormalities in neural transmission along myelinated pathways, such as the visual pathway or the sensory pathways of the spinal cord and brainstem. This can help in the diagnosis of disorders such as MS or B_{12} deficiency.

Laboratory Investigations

In general, empirical evidence for the utility of laboratory studies supports only a limited role for "routine" or screening investigations; for the most part, laboratory tests should be performed as guided by the history and examination. A full discussion of laboratory strategies for all

neuropsychiatric situations is beyond the scope of this section. In regard to dementia, a complete blood count (CBC), chemistry panel, B_{12} assay, and thyrotropin (TSH) assay are indicated as screening tests, in addition to a test for syphilis (the fluorescent treponemal antibody test [FTA]) in those areas of the United States in which the prevalence of syphilis justifies the testing. (The region of high incidence is a broad belt across the South in addition to some urban areas in the North; 30 US counties contribute more than half the national total of cases.) The reason the FTA is the test of choice is that reagin tests (the venereal disease research laboratory test [VDRL] or rapid plasma reagin [RPR]) may revert to normal after intercurrent antibiotic treatment or with the passage of time and thus are insufficiently sensitive to serve as screening tests for neurosyphilis.

Table 2.1–3.
Metabolic Disorders with Potential Psychiatric Presentations

Disease	Psychiatric Features	Other Features (All May Have Cognitive Impairment)	Initial Diagnostic Test
Metachromatic leukodystrophy	Psychosis	Peripheral neuropathy, cerebellar and pyramidal signs, anterior white matter abnormality on MRI	Arylsulfatase A
GM2 gangliosidosis	Psychosis, mood disorder	Pyramidal, cerebellar, lower motor neuron signs	β-Hexosaminidase A, genetic testing
Niemann–Pick type C	Mood disorder, psychosis	Hepatomegaly, splenomegaly, cerebellar signs, dystonia, vertical oculomotor apraxia	Fibroblast culture, genetic testing
Adrenoleukodystrophy	Mood disorder, psychosis	Pyramidal signs, peripheral neuropathy, visual loss, adrenal insufficiency, posterior white matter abnormality on MRI	Very long chain fatty acids, genetic testing
Neuronal ceroid lipofuscinosis (Kufs disease in adolescents)	Mood disorder, psychosis	Extrapyramidal, cerebellar signs	Electron microscopy of skin/rectal biopsy
Cerebrotendinous xanthomatosis	Psychosis	Cerebellar, pyramidal signs, peripheral neuropathy, cataracts, xanthomas	Cholestanol
Urea cycle disorders	Acute confusion/psychosis	Protein intolerance, headache, nausea, pyramidal signs, seizures	Ammonia
Porphyria (acute intermittent or variegate)	Acute confusion/psychosis	Peripheral neuropathy, abdominal pain	Urine porphobilinogen
Wilson disease	Personality change, rarely psychosis	Extrapyramidal signs, Kayser–Fleischer ring, liver disease, basal ganglia hyperintensity	Ceruloplasmin

Appropriate screening tests for mental presentations other than dementia, for example, first-episode psychosis, are less well established. Unfortunately, no cohort studies applying a consistent laboratory diagnostic approach are available to provide guidance as to the sensitivity and specificity of testing or even as to the prevalence of organic disease in this situation. The first step should be a neuropsychiatric history and examination. A reasonable laboratory screen might include CBC, chemistry panel, TSH, urinalysis, and urine toxicology. (A fasting lipid panel and in women a pregnancy test are appropriate as well.) If it is considered justified to screen for rheumatic disease, an antinuclear antibody test (ANA) is adequate for this purpose, being abnormal in almost all cases of lupus, although not sufficient to confirm that diagnosis. (False positives from psychotropic drug-induced ANA assays will be an important confound.)

Excessive laboratory testing is to be deplored; on the other hand, limiting laboratory testing to generally familiar diseases is inexpert. Many rare metabolic diseases can present in adolescence or young-adulthood with purely psychiatric manifestations, or with psychiatric manifestations on the background of developmental intellectual impairment, or with prominent psychiatric symptoms mixed with motor or systemic abnormalities. The family history, which may be of severe intellectual deficit or early mortality in a sibling, may be a pointer to the need for more thorough investigation. Some of these disorders are listed in Table 2.1–3.

Genetic testing is relevant in certain neuropsychiatric situations, some of which are mentioned in Table 2.1–3. Genetic testing in the degenerative dementias is still a research tool for the most part, though under certain circumstances (e.g., identifying a mutation in the prion protein gene) it may play a practical role in diagnosis. “Premutation”-range repeat expansions in the Fragile X gene may at times provide an explanation for an anxiety disorder.

Examination of the Cerebrospinal Fluid

Examination of CSF obtained through lumbar puncture is sometimes a crucial element of the diagnostic process, in particular to diagnose infection or inflammation, more rarely in neuropsychiatric practice to seek evidence of neoplasia (such as meningeal carcinomatosis). Specific assays are available for the diagnosis of neuropsychiatrically relevant infectious agents, such as polymerase chain reaction (PCR) for the herpes simplex virus (HSV) genome to diagnose herpes encephalitis or cryptococcal antigen assay to diagnose this fungal meningitis. In rheumatic diseases involving the brain, the white cell count may not be elevated, but elevated protein and evidence of intrathecal elaboration of antibodies may give evidence of inflammatory activity. The latter is sought by the ratio of

immunoglobulin-G (IgG) to albumin or, better, by the IgG index, which requires measurement of serum IgG by immunoelectrophoresis. CSF antineuronal antibodies are uncommon but specific for cerebral lupus. In the future, assay of CSF cytokines may provide assistance in the difficult diagnosis of these inflammatory diseases.

Measurement of the neuron-derived 14–3–3 protein in CSF has been used to assist in the diagnosis of CJD, but the ratio of phosphorylated to total tau protein in spinal fluid has proven to be more sensitive and highly specific. A new test, RT-QUIC, directly identifies prions in the CSF and promises even greater sensitivity and specificity for the diagnosis of CJD. Increased CSF total and phosphorylated tau and reduced amyloid protein, in combination, show a high sensitivity and high specificity for Alzheimer disease.

Removal of CSF by lumbar puncture or external drainage also plays an important role in the evaluation of patients suspected of shunt-reversible normal-pressure hydrocephalus.

Neuropsychological Assessment

Neuropsychological evaluation has an important role to play in neuropsychiatric care, both for diagnosis and for management. Sound use of the clinical neuropsychologist as a consultant requires as a first step formulation of a cogent consultative question. The more specific the consultant's question, the more able the neuropsychologist is to integrate the psychometric data with the rest of the clinical picture. Much of the early literature on neuropsychological assessment focused on identifying and localizing organic brain disease. With the advent of neuroimaging, neuropsychological testing is seldom the most powerful means of addressing this issue, although it certainly continues to play such a role, for example, in lateralizing cognitive deficits as a preoperative tool in epileptic patients. Nor is the role of the neuropsychologist to make a disease diagnosis, although at times the psychometric picture is strongly suggestive of a particular diagnosis.

In several areas of assessment, the neuropsychiatrist has particular reason to turn to the neuropsychologist. If substantial confounds make bedside diagnosis difficult, neuropsychological data may be of considerable assistance. For example, identifying supervening cognitive impairment in an intellectually disabled or poorly educated patient or subtle impairment in a highly intelligent patient may be impossible for the clinician to do with confidence, while quantitative assessment may allow these diagnoses. Another example of utilizing neuropsychological assessment as a probe of brain function is disclosing a pattern of cognitive strengths and weaknesses amounting to right hemisphere learning disabilities in a patient with a clinical picture suggestive of pervasive developmental disorder or a cluster A personality disorder. Obtaining neuropsychological data about a

dementing patient often allows more precise targeting of behavioral interventions, more specific education of families, and more confident assessment of decline or of benefit from pharmacological treatments.

One common use of neuropsychological assessment requires a word of caution: identifying cognitive impairment in an older patient presenting with mood disorder or psychosis. No neuropsychological findings should deter the clinician from aggressive treatment of the psychiatric symptoms, and nonspecific state-dependent attentional and motivational factors may confound the neuropsychological results. Rather than devoting resources of time and energy to pinpointing a moving target in the acute phase, deferring the assessment until symptoms are reduced is often the wiser course.

Another caution about neuropsychological assessment falls under the rubric of ecological validity. This term refers to the extrapolation of results obtained in the neuropsychological laboratory by artificial paper-and-pencil methods to real-world performance. The concern arises in particular with orbitofrontal lesions, which may produce a paucity of cognitive findings but devastating personality change. Deriving from the developing realm of affective neuroscience clinical measures suitable to characterize such patients is a current challenge to neuropsychology.

Brain Biopsy

Biopsy of the brain has a limited role in neuropsychiatric evaluation. The morbidity and mortality of the procedure, as performed by an experienced neurosurgeon, are low, but the sensitivity of the procedure is lower than one might expect. For example, the sensitivity of biopsy for primary angiitis of the central nervous system (CNS) may be only 75 percent. In some circumstances, biopsy of a peripheral tissue can substitute for brain biopsy in a patient with primarily cerebral symptoms at lower risk. For example, lung or muscle biopsy may make a diagnosis of sarcoid, skin biopsy a diagnosis of vasculitis if a rash is present or of CADASIL, and temporal artery biopsy of giant cell arteritis. In neuropsychiatric situations, the major indication for biopsy of the brain is consideration of inflammatory disease, when the nature and aggressiveness of treatment depend on a tissue-proven diagnosis.

Although it cannot be considered a clinical diagnostic test, the neuropsychiatrist should not neglect the autopsy as a learning tool.

COMMON NEUROPSYCHIATRIC CONDITIONS

This section provides a survey of some issues commonly brought to the attention of neuropsychiatrists. The emphasis is on the priorities for clinical and laboratory assessment for a variety of presentations. The organization is by disease and syndrome, as a complement to the anatomic

and symptom-oriented discussion provided so far. This perspective is distinctive for neuropsychiatry within psychiatry; the disease processes underlying symptoms in the idiopathic disorders are unknown. For neuropsychiatric patients one can hope and work to uncover the disease causing the symptoms and on fortunate occasions to provide disease-specific treatment.

Dementia

Many diseases produce the clinical state of dementia. A shotgun laboratory approach to "ruling out treatable disease" is unwise, if only because finding reversibility is unusual. Moreover, clinical clues to reversible disease are available in the history and examination: use of psychotoxic medicines, rapid course, mildness of cognitive impairment (even short of fully meeting criteria for dementia), subcortical features of the cognitive disorder, presence of motor signs. Table 2.1–4 provides specific guidance to be gained from clinical clues. The diagnostic approach to dementia also includes differential diagnosis among the degenerative disorders. Progressive dementia supervening on developmental intellectual disability raises a particular differential diagnosis, including many of the diseases listed in Table 2.1–3.

Table 2.1–4.
Clues to the Differential Diagnosis of Dementia in the Neurological Examination

Abnormal eye findings	Celiac disease Gaucher disease type 3 Mitochondrial cytopathy Multiple sclerosis Niemann–Pick disease type C Progressive supranuclear palsy	Syphilis Vascular dementia Wernicke–Korsakoff syndrome Whipple disease Wilson disease
Ataxia	Celiac disease Cerebellar degenerations GM2 gangliosidosis Hypothyroidism Multiple sclerosis Niemann–Pick disease type C Prion disease	Progressive multifocal leukoencephalopathy Toxic-metabolic encephalopathy Wernicke–Korsakoff syndrome Wilson disease
Dysarthria	Cerebellar degenerations Dementia pugilistica Dialysis dementia Motor neuron disease Multiple sclerosis Niemann–Pick disease type C Neuroacanthocytosis	Pantothenate kinase–associated neurodegeneration Progressive multifocal leukoencephalopathy Progressive supranuclear palsy Wilson disease
Extrapyramidal signs	Alzheimer disease Cerebellar degenerations Dementia pugilistica	Parkinson disease Progressive supranuclear palsy Postencephalitic parkinsonism

	Dementia with Lewy bodies Fahr syndrome GM1 gangliosidosis type 3 Huntington disease Multiple system atrophy Neuroacanthocytosis Niemann–Pick disease type C Normal-pressure hydrocephalus Pantothenate kinase–associated neurodegeneration	Subacute sclerosing panencephalitis Toxic-metabolic encephalopathy Vascular dementia Wilson disease
Gait disorder	Adrenomyeloneuropathy Cerebellar degenerations Dementia pugilistica HIV encephalopathy Multiple sclerosis Normal-pressure hydrocephalus	Parkinson disease Progressive supranuclear palsy Syphilis Vascular dementia Wernicke–Korsakoff syndrome
Myoclonus	Alzheimer disease Celiac disease Dialysis dementia Kufs disease Lafora body disease	Mitochondrial cytopathy Prion disease Subacute sclerosing panencephalitis
Peripheral neuropathy	Adrenomyeloneuropathy B_{12} deficiency HIV encephalopathy Metachromatic leukodystrophy	Porphyria Toxic-metabolic encephalopathy
Pyramidal signs	Adrenomyeloneuropathy B_{12} deficiency Cerebellar degenerations GM2 gangliosidosis HIV encephalopathy Kufs disease Metachromatic leukodystrophy Motor neuron disease Multiple sclerosis	Pantothenate kinase–associated neurodegeneration Polyglucosan body disease Progressive multifocal leukoencephalopathy Syphilis Vascular dementia

HIV, human immunodeficiency virus.

Modified from Sandson TA, Price BH. Diagnostic testing and dementia. *Neurol Clin.* 1996;14:45, with permission.

The clinician needs to gather data relevant to management issues other than purported reversibility, such as safety of living arrangements, driving ability, preparation of a will and advance directives, and so on. Dementing patients often develop psychiatric symptoms, which respond to pharmacological and behavioral treatment. All these considerations should prompt the clinician to cast a wide net in data gathering regarding the demented patient.

A 67-year-old woman presented with at least 1 year of progressive memory impairment, confusion, then irritability and suspiciousness. The mental state was typical of Alzheimer disease, and the physical examination disclosed only brisk tendon jerks. An EEG, done earlier because of a spell of uncertain nature, had shown left temporal spikes.

Neuropsychological assessment had shown a pattern typical for Alzheimer disease, with memory impairment characterized by rapid forgetting, semantic more than phonemic verbal fluency deficits, and anomia. MRI, however, demonstrated extensive white-matter disease, with bilateral confluent hyperintensities, which extended into the gyri and involved U-fibers. CSF examination was entirely normal. Skin biopsy for CADASIL and screening genetic assay for CADASIL were negative. Repeat EEG showed bilateral temporal spikes, and carbamazepine (Tegretol) was begun. The clinical diagnosis was leukoencephalopathy due to cerebral amyloid angiopathy, possibly with Alzheimer disease. The patient's mother had died at age 86, having suffered from "the same thing" as the patient. Four of the mother's five siblings demented in the eighth or ninth decade of life, none earlier, in most cases with a clinical diagnosis of Alzheimer disease. The patient herself was an only child. The patient's two daughters, both young adults, were very concerned that they might inherit the same disease as their mother, and they insisted the patient undergo the brain biopsy that a geriatrician had recommended. This disclosed pronounced congophilic angiopathy. Immunostaining for A-β confirmed the vessel abnormality and showed neuropil plaques; immunostaining for tau did not reveal neuritic plaques. Nonetheless, Alzheimer disease could not be excluded. No inflammation was seen. Unfortunately, several days after the biopsy she developed status epilepticus.

Among the dementias, behavioral-variant frontotemporal dementia (bvFTD) is of particular neuropsychiatric interest. In contrast to Alzheimer disease (AD), the incidence of which increases with age, in FTD the peak incidence is in the fifth and sixth decade of life. In contrast to AD, which characteristically has a cognitive presentation with memory failure most prominent, bvFTD presents with altered personality and behavior; typically executive cognitive dysfunction can be detected. Other neuropsychiatric features may be present early in the course, including psychotic symptoms, especially in bvFTD due to repeat expansions in the C9ORF72 gene.

Epilepsy

Major concerns in patients with epilepsy include differential diagnosis, psychosis, personality change, depression, violence and other episodic behaviors, and nonepileptic seizures. The last will be dealt with below, along with other conversion disorders.

Patients with paroxysmal disorders can be misdiagnosed to have epilepsy when they do not have it or as having a different disorder when epilepsy is the correct diagnosis. Paroxysmal symptoms from panics, cardiovascular disease with syncope or near-syncope, endocrine disorders (pheochromocytoma, carcinoid, systemic mastocytosis), or conversion disorder can be mistakenly labeled epileptic; contrariwise, epilepsy can be missed when a diagnosis of panic disorder in particular is accepted.

A 59-year-old man was evaluated for 7 years of memory problems and spells refractory to treatment on a diagnosis of panic disorder. These spells characteristically lasted 5 to 10 seconds and recurred as often as hourly; he was sometimes amnestic for the spells afterward. During an attack, he had gooseflesh and his speech became garbled; once at church he was thought to be speaking in tongues. Extensive treatment trials with benzodiazepines and serotonergic drugs had given no consistent benefit. Apart from

hyperlipidemia, he had no significant medical history. EEG had been negative on three occasions, MRI on two occasions, Holter monitoring and SPECT on one occasion each. The neurological and mental state examinations were normal. An episode was witnessed during the examination: He showed 10 seconds of facial flushing and stereotyped hand movements. The attacks were subsequently abolished by a trial of an antiepileptic drug. The case illustrates that epilepsy is primarily a clinical, not an EEG, diagnosis.

A 52-year-old woman was referred for the evaluation of spells. In her 30s she had been hospitalized for depression and was subsequently treated intermittently as an outpatient. The family history included several members with depression or bipolar disorder. Two years before evaluation she presented with headache and proved to have an unruptured aneurysm, which was clipped through a craniotomy. Several months later she had a generalized convulsion. She went on to have spells at a rate of up to six a day. They were stereotyped and abrupt in onset and termination; she could not identify provocative factors or social contexts. During a spell she would feel cold and have gooseflesh for about 3 minutes. Then she would become rigid and unable to speak or interact, although able to hear others' speech. This would last several minutes. Then she would begin to cry. The whole sequence would last some 6 to 10 minutes. She was on phenytoin (Dilantin) with a therapeutic serum level. Previous trials of divalproex (Depakote) and topiramate (Topamax) were not tolerated. EEGs had shown only right frontal slowing with no epileptic features on several tracings. The neurological and cognitive examinations were normal, and she was not depressed at the time of evaluation. The clinical picture was inconclusive. In favor of epilepsy were the abrupt onset and termination, stereotyped nature of the spells, and background of craniotomy; against epilepsy were the weeping, length of the ictus (if all the phenomena were taken to be ictal), failure of response to treatment, relatively inactive EEG, and background of depression. Video EEG done after medication withdrawal recorded three complex partial seizures with right anterior-inferior temporal onset with her typical semiology.

In exploring the psychiatric concomitants of epilepsy, the clinician needs to be aware of the nature of the epilepsy. Most adult epilepsy is focal (localization-related epilepsy), with the ictal onset in the temporal lobe. However, other forms of epilepsy, including frontal epilepsy and primary generalized epilepsy, are common. These distinctions, and the laterality of the focus, can often be inferred from the semiology of the seizures as reported or as observed clinically. A history of febrile convulsions in childhood and age of onset of epilepsy are relevant to the likelihood of mesial temporal sclerosis as the underlying pathology. Body asymmetry and dissociated facial paresis should be sought as indicators of laterality. The MRI and EEG provide crucial information on pathology and seizure type. Almost all the findings relating psychiatric disorder to epilepsy are concerned with partial epilepsy of temporal onset; linking psychiatric symptoms to epileptic syndromes other than temporal lobe, or limbic, epilepsy would generally go beyond the evidence. Further, to what extent psychopathology is associated with the epilepsy per se and to what extent with the underlying brain disease remains controversial. Without question, cognitive impairments are related to the lateralization of the temporal focus.

Psychotic states in epileptic patients are usually divided into those

occurring during the epileptic ictus, often called epileptic twilight states; those occurring for a delimited period in the aftermath of a seizure or, more commonly, a flurry of seizures, called postictal psychosis; and those that are chronic, called interictal psychosis. Usually this chronology can be ascertained by inquiry, but at times EEG monitoring is necessary to identify the occurrence of seizures in relation to psychopathological phenomena, especially because patients can be amnestic for complex partial seizures. A further issue to be elucidated from the history is of a relationship between seizure treatment and control and the level of psychopathology, especially psychosis. An inverse relationship is sometimes noted, better seizure control being associated with occurrence of psychosis, a phenomenon known as forced normalization. On the other hand, frequent seizures certainly can cause an increase in confusion and related failure in functional capacity.

A 35-year-old woman with lupus and intractable epilepsy was admitted several times with persecutory and nihilistic delusions ("I'm dead") and depressive symptoms. Investigations to identify active cerebral lupus were unrevealing, even when she had evidence of peripheral activity of the disease. In fact, MRI disclosed the findings of hippocampal sclerosis, suggesting that the epilepsy was idiopathic and not due to cerebral lupus. Without specific treatment, the psychotic symptoms diminished over the course of several days; this also was thought to make a diagnosis of cerebral lupus unlikely. Between episodes she showed no psychotic phenomena.

The interictal personality syndrome of temporal lobe epilepsy (the Gastaut–Geschwind syndrome) is characterized by hypergraphia; religiosity or deepened metaphysical interest; intensified emotionality with a tendency to holding grudges and aggression; hyposexuality; and an alteration of social behavior with intensity of interaction, an inability to end interactions, and circumstantiality of discourse (phenomena confusingly denominated "viscosity"). The syndrome remains controversial; what is of importance for neuropsychiatric assessment is that inquiry be directed to phenomena such as hypergraphia that are not included in the usual review of symptoms of idiopathic psychiatric disorder.

Episodic aggression is often suspected of being ictal but rarely is. Aggression occurring during seizures is almost always disorganized, not deliberately directed. A high threshold is justified in attributing a violent act to epilepsy in the absence of typical epileptic features. Amnesia for serious violence is common and not a strong pointer to an epileptic origin.

A special issue in neuropsychiatric assessment of the epileptic patient is the presurgical evaluation. Surgical treatment, especially of temporal lobe epilepsy due to mesial temporal sclerosis, is underutilized; ideally more and more patients with medically refractory epilepsy will be evaluated for their suitability for surgery. Along with intensive electroencephalographic evaluation, volumetric MRI, and neuropsychological assessment, the

patient's psychiatric state should be systematically evaluated. The patient's ability to consent and issues such as the patient's capacity to cope with the stress of monitoring and surgery as well as the expectations held for surgery should be addressed. Neither depression nor psychosis is an absolute contraindication to surgery, although a chronic psychosis probably will not be alleviated by surgery. Indeed few if any psychiatric findings will contraindicate surgery, but psychiatric evaluation may well reveal deficits that need to be taken into account in developing a treatment plan.

Traumatic Brain Injury

Traumatic brain injury is epidemic in current society, with advances in emergency medical care leading to growth in the prevalence of survivors of severe injury. Issues commonly facing the neuropsychiatrist include aggression, depression and anxiety, and the delineation of deficits (sometimes for legal purposes) in patients with mild traumatic injury. The features of the head injury should be ascertained, ideally with confirmation from medical records.

The altered behavior and personality common after traumatic brain injury are more burdensome for families than are the physical disabilities. Disinhibition and aggression are particularly uncomfortable and often hard to treat. A complicating factor is that preinjury impulsivity and substance abuse are common, as they predispose to head injury.

A 24-year-old woman was seen 19 months after a car crash in which she was an unrestrained passenger. She had been comatose for 3 months and underwent surgical evacuation of a left-sided intracranial hematoma. She had a few weeks of rehabilitation after regaining consciousness and returned home after spending most of a year in a nursing home. The family was at wits' end over episodes of aggression, which appeared to be directed angry behavior elicited by frustration. She did not have depressive symptoms. On examination, she showed severe bilateral spasticity, including spastic dysarthria, drooling, and a brisk jaw jerk. She was able to recall dates and other details of her illness accurately, but she disclaimed behavioral or emotional alterations. Her language comprehension was adequate, but output was telegraphic. Affect was labile. Behavior during the consultation was initially appropriate, with an obvious effort to cooperate with the evaluation, although she had greeted the examiner with, "I love you." At the end of the examination, however, she urgently requested the examiner's business card and rammed him with her wheelchair while cutting off his access to the door. Chronic phase CT showed atrophy and left temporal encephalomalacia.

A 59-year-old man was referred by a court for assessment of his ability to take part in proceedings related to his divorce. Three months prior to evaluation, he was struck several times by an unknown assailant during an altercation regarding who was to get the use of a taxi. He suffered contusions of the left periorbital area but no other overt injury. He was able to recount the events in some detail, but he explained that this was because over time, by comparing notes with others, he had "put it all back together"; of his own recollection he could remember the first punch that struck him but not the second or subsequent events of

the altercation. Although he could not be certain of the duration of the gap in his recollection, it was clearly a matter of some seconds, possibly a minute or two, and at no time was he unconscious. In the aftermath, he was "confused" and had a headache. He found that he could not come up with names, dates, or numbers, although this information would generally come back to him later or with considerable unaccustomed effort. He also noticed that he "could not visualize" geographical scenes, so that in planning to go to a familiar place he was unable to picture it in his mind. Although he did not get lost, he found that he would turn the wrong way or miss a turn because of inattention and have to correct himself. He noted that his memory, previously highly trustworthy, could not be counted on: "I had to write everything down." He found that he had to "take time to think," "strain my brain to focus." He distanced himself from business decisions and relied on trusted subordinates to counsel him. He acknowledged sensitivity to light, noting that he had begun to wear sunglasses even when the weather was cloudy and to turn off the room lights when he was watching television. To a lesser extent he was bothered by noise. He noted that he was more readily irritated than was characteristic of him. He did not have depressive symptoms, intrusive recall of the altercation, or nightmares. The symptoms had gotten gradually less severe. The history included several head injuries in adolescence, with two of which he had loss of consciousness of a few hours without recognized sequelae. The noncognitive mental state and neurological examinations were normal. He scored 21 on the Mental Alternations Test, a normal performance on a task of mental speed and working memory. He scored 16 of 18 on the Frontal Assessment Battery, a collection of tasks assessing executive cognitive function. The two lost points were on the go/no-go task, on which he made perseverative errors. He was mildly disorganized on performing the ring/fist test of motor sequencing. MRI and EEG had been normal. The picture was thought to be consistent with organic sequelae of traumatic brain injury. The case underlines the importance of prior traumatic brain injury in determining the effects of seemingly mild trauma and that loss of consciousness is not a prerequisite for significant sequelae.

Movement Disorders

Cognitive impairment due to involvement of subcortical structures is a common neuropsychiatric feature of the movement disorders. This applies to cerebellar as well as basal ganglia diseases, for the anatomic reasons described above. The anatomy of the close relation between emotion and movement was also described above. Clinically, mood disorders are common in IPD and other movement disorders. Anxiety disorders, although less emphasized in the literature than depression, are also common. The evaluator should take into account that mood and anxiety can fluctuate according to the timing of doses of dopaminergic drugs. A mood disorder can occasionally present in advance of overt movement abnormalities, so IPD must be considered in the differential of late-onset mood disorders.

A 43-year-old woman with no personal or family history of psychiatric illness developed a psychotic depression. She had a severe extrapyramidal reaction to risperidone (Risperdal). Two years later, when euthymic and unmedicated, she developed progressive shuffling gait, upper extremity tremor, and micrographia. She then suffered another episode of depression. Three years later she had severe anxiety, no cognitive impairment, and the motor features of IPD.

Psychotic reactions to dopaminergic drugs are an important feature of movement disorders. Sometimes this is the result of overuse of prescribed dopaminergic agents, in an effort to increase time in the "on" state.

A 63-year-old man with long-standing IPD developed delusions while being treated with high-dose levodopa and carbidopa (Sinemet) on a five-times-a-day schedule, pramipexole (Mirapex), tolcapone (Tasmar), and amantadine (Symmetrel). Under inpatient observation for several days on the prescribed doses, he remained psychotic. He responded well to quetiapine (Seroquel).

Developmental Disabilities

Adult patients with developmental disabilities are enormously underserved by the medical and social service communities and are frequently referred for neuropsychiatric attention. Few of these patients will have had adequate diagnostic evaluation for the cause of the disability. Beyond clinical assessment, with particular attention to dysmorphology, because features of the mental state and neurological examination are generally nonspecific, the most useful diagnostic tests are MRI and karyotyping. Specific genetic probes can confirm tentative clinical recognition of syndromes of mental retardation. Patients with developmental disabilities are vulnerable—indeed especially vulnerable—to the mood, anxiety, and psychotic disorders that can afflict anyone and can be treated effectively for these; diagnostic overshadowing (attributing all psychological and behavioral disturbance to developmental disability *tout court*) is to be avoided. These syndromes may present atypically in the developmentally disabled population, and the clinician must be alert to indirect indicators of mood disturbance or psychotic experience. For example, while the patient may not report depressed mood verbally, the caregivers may report the loss of interest in favorite activities and the other features of a depressive syndrome.

Of particular neuropsychiatric interest is the question of behavioral phenotypes, specific psychological correlates of developmental syndromes. Syndromes recognized to have behavioral phenotypes (and their correlates) include:

- Lesch–Nyhan syndrome (self-injury);
- Prader–Willi syndrome (excessive eating);
- Williams syndrome (anomalous cognitive profile, elevated sociability);
- Velo-cardio-facial syndrome (schizophrenia).

Infectious and Inflammatory Diseases

Infectious and inflammatory diseases of the brain always need to be considered in acutely or subacutely evolving mental disorders. Among the infectious diseases, HSV encephalitis has a particular claim on attention,

because delay in diagnosis, even by hours, can lead to substantially increased morbidity and mortality. Definitive diagnosis is possible without biopsy by assaying for HSV in the CSF with the PCR, but treatment may be indicated if suspicion is high in advance of firm diagnosis. Chronic meningitis, for example, from infection with fungi, is a rare consideration in subacutely evolving dementia; the definitive diagnostic tests are CSF assays or serological tests (e.g., for toxoplasmosis). In the acquired immunodeficiency syndrome (AIDS) era, infection with opportunistic agents and with the human immunodeficiency virus itself needs to be kept in mind.

Noninfectious inflammatory diseases include the rheumatic diseases, of which the prototype is systemic lupus. A rheumatic disease review of systems is always of importance in exploring the differential diagnosis of a puzzling case, especially in a young woman. Although psychosis is often considered the psychiatric hallmark of lupus, in fact psychotic states (other than delirium) are unusual, and a variety of other psychiatric pictures need to be included in the clinician's consideration. Few clinical features of lupus are risk factors for cerebral disease, not even disease activity, which may be misleading in either a positive or a negative direction. One feature that is a risk factor for neuropsychiatric symptoms, including cognitive impairment, is the presence of antiphospholipid (aPL) antibodies; the primary aPL syndrome similarly carries mental risk.

A 40-year-old woman presented with the typical features of psychotic depression. There was a family history of depression, and she had suffered two episodes of depression earlier in her adult life, both of which were brief, nonpsychotic, and responsive to treatment. For the previous year, however, her depression had been poorly responsive to pharmacological treatment and electroconvulsive therapy (ECT). Examination disclosed no definite cognitive abnormality and brisker reflexes on the left. Review of MRI obtained at her previous treatment venue, performed as a routine prior to the administration of ECT, evinced striking areas of white matter abnormality in the right hemisphere. Extensive laboratory investigation, short of angiography and biopsy, revealed only high-titer IgA anti-β2-glycoprotein-1 antibodies. Neuropsychological assessment performed after partial remission of the depression showed deficits in attention and mental processing speed. The working diagnostic formulation was that an otherwise ordinary idiopathic depressive disorder had been rendered treatment resistant and gravely severe by a wave of cerebral injury due to the antiphospholipid syndrome. One lesson of the case is always look at the imaging yourself.

Other rheumatic diseases, such as Sjögren syndrome and the vasculitides, are also of neuropsychiatric importance. Hashimoto encephalopathy—subacutely developing cognitive impairment and myoclonus or seizures with elevated antithyroid antibodies—is important in the differential diagnosis of CJD and of subacute confusional states. Prominent among nonrheumatic inflammatory diseases is paraneoplastic limbic encephalitis, an autoimmune complication of several tumors, notably small cell carcinoma of the lung. In recent years, a number of new

antibodies have been discovered as the cause of autoimmune limbic encephalitis, commonly without a systemic neoplasm. These include antibodies to the n-methyl d-aspartate receptor and those to voltage-gated potassium channels. Limbic encephalitis, which can also be infectious (notably due to HSV), typically features memory impairment and altered mood with a subacute course. The role of such autoantibodies in cases of apparently idiopathic acute psychosis remains to be elucidated.

A 60-year-old woman was admitted for confusion. She had been drinking more heavily than usual after a forced retirement several months earlier. The family noted that she had been forgetful and behaviorally erratic for 1 to 2 months. She smoked cigarettes and had hypertension. On examination, she had mild gait instability but no other physical signs. Thought was disorganized, but no psychotic ideas were present. Psychomotor rate and affect were normal. She showed verbal memory impairment and disinhibition, with many errors of commission on a go/no-go task and failure to inhibit reflexive gaze. CSF examination revealed a mild lymphocytic pleocytosis with no other abnormalities. EEG showed intermittent frontal slowing. The following were negative or normal: serological studies, thyroid function and antibody tests, anti-Hu, MRI, magnetic resonance angiography (MRA), chest and abdominal CT (except a benign adrenal tumor), and cerebral angiogram. The patient's family refused brain biopsy. On a differential diagnosis of primary angiitis and paraneoplastic encephalitis the patient received a pulse of intravenous (IV) methylprednisolone, without benefit, then a course of oral cyclophosphamide and prednisone, again without benefit. Some months after discharge, she died suddenly. Autopsy revealed pulmonary embolus to be the cause of death. Perivascular T-cell infiltrates and activated microglia were seen in the medial temporal lobes, and to a lesser degree widespread in the cortex. No tumor was found in the lungs or elsewhere.

A 56-year-old man without prior psychiatric history presented because of "a few months of memory problems." Beginning 5 months prior to the neuropsychiatric evaluation, he seemed confused and forgetful to his family. For example, he repeatedly called his daughter at college and could not recall what she was studying. He declined over ensuing months, with anxiety, depression, reduced energy, and difficulty in maintaining sleep but without anhedonia or diurnal variation of mood. A psychiatrist treated him with quetiapine and lithium on a diagnosis of a mixed affective state but without benefit. The neurological examination was normal. On cognitive examination, he showed mild memory impairment but no disturbance of naming or executive function. MRI was unremarkable. Neuropsychological assessment showed impaired memory with both rapid forgetting and reduced encoding. A panel of autoimmune serologies showed antibodies to the voltage-gated potassium channel, and he responded to immunomodulatory treatment.

Conversion Disorder

Neuropsychiatrists often see patients whose symptoms appear to arise from brain disease but do not. These patients' condition has been described under various names: hysteria, functional disorder, psychogenic disorder, conversion disorder, or medically unexplained symptoms. None of the designations is entirely satisfactory. Whatever the designation, such patients are not uncommon. Complicating matters is the common coexistence of organic disease and conversion symptoms. For example, a

sizable minority of patients with nonepileptic seizures has epilepsy as well. Brain disease may, in some patients with conversion disorder, have produced organic personality change with a reduction in the maturity of defenses and the too-easy resort to somatization.

Various techniques have been advocated for identification of nonorganic disease from the physical examination. These have several shortcomings. First, they easily lend themselves to a counter-therapeutic alliance in which the examiner is trying to trick the patient—not a good start for the treatment whatever the findings. Second, they fail to distinguish deliberate falsification on the patient's part (i.e., malingering), from conversion disorder. Third, most such findings are commonly present in patients with organic disease who are trying to help the examiner make the diagnosis. That is, they may show a patient to be histrionic or suggestible but fail to rule out organic disease. Thus, for example, reporting a difference in vibratory sensation between the two sides of the sternum is by no means confined to patients with conversion disorder. Exceptions to this caution occur in cases where the nonphysiological finding is precisely the phenomenon of the complaint. Even then, however, the phenomena of brain disease are sufficiently odd that the examiner should maintain an attitude of humility about achieving diagnostic certainty by recognizing the nonphysiological at a glance. Of the described "signs of hysteria," perhaps the best is the Hoover sign. The examiner places a hand underneath the heel of the affected leg of a supine patient who complains of leg weakness. Asked to press down with the heel, the patient fails to generate power with the leg. Asked to raise the opposing leg, however, the patient produces an automatic synergistic downward movement of the affected leg.

Recent systematic findings of progressively greater methodological sophistication confirm the belief that experiences of abuse in childhood are common in the background of patients with conversion disorder. This may indirectly account for another progressively more solidly substantiated finding, namely that the prognosis of conversion disorder is poor. Although a given symptom may wax and wane or disappear, patients commonly have a chronic course of disability, interpersonal difficulties, psychiatric symptoms, and fruitless seeking after medical help. Although hysterical symptoms have often been taken to represent symbolically a psychological conflict, the fundamental difficulty is that patients who make prominent use of somatization have a disorder of the symbolic function itself. The goal of the examiner should be not to expose the patient, but to establish an alliance that allows exploration of areas of the patient's life outside the presenting symptoms and construction of a plan to reduce distress and manage emotions (including focused treatment of commonly coexisting depressive disorder) and to develop alternative ways of seeking attention and assistance for distress.

THE NEUROPSYCHIATRIC PERSPECTIVE

This section has surveyed a neuropsychiatric approach to the patient. The neuropsychiatrist thinks anatomically about mental state disorders, even as cognitive neuroscientists attempt to construct a sufficiently sophisticated model of large-scale brain function to do justice to the complex mental states of neuropsychiatric interest. The neuropsychiatrist relies on rich data gathered at the bedside and on laboratory methods of investigating brain structure and function and of diagnosing disease. The effort is to identify not just behavioral syndromes as found in DSM or ICD but the pathological processes underlying them, in two senses. First, neuropsychiatry seeks medical diagnoses of systemic or brain diseases that account for the patient's illness. Second, neuropsychiatry seeks to understand clinical phenomena in terms of the disruption of elementary mental processes, the nature of which is beginning to be elucidated by the cognitive neurosciences. The result is a highly differentiated diagnostic enterprise. With continual refreshment from a multidisciplinary base—ranging from cognitive neuroscience to general medicine—the neuropsychiatric approach to the patient is certain to remain exciting.

REFERENCES

Erro R, Brigo F, Trinka E, Turri G, Edwards M, Tinazzi M. Psychogenic nonepileptic seizures and movement disorders: A comparative review. *Neurology Clin Pract*. 2016;6:138–149.

Ferro J, Caeiro L, Figueira M. Neuropsychiatric sequelae of stroke. *Nat Rev Neurol*. 2016;12:269–280.

Frith C. In praise of cognitive neuropsychiatry. *Cogn Neuropsychiatry*. 2008;13:1–7.

Geschwind N. Disconnexion syndromes in animals and man. I. *Brain*. 1965;88:237–294.

Geschwind N. Disconnexion syndromes in animals and man. II. *Brain*. 1965;88:585–644.

Habib M. Athymhormia and disorders of motivation in basal ganglia disease. *J Neuropsychiatry Clin Neurosci*. 2004;16:509–524.

*Heimer L, Van Hoesen GW, Trimble M, Zahm DS. *Anatomy of Neuropsychiatry: The New Anatomy of the Basal Forebrain and its Implications for Neuropsychiatric Illness*. Boston: Academic Press/Elsevier; 2008.

Hillis A. Inability to empathize: brain lesions that disrupt sharing and understanding another's emotions. *Brain*. 2014;137:981–997.

*Hodges JR. *Cognitive Assessment for Clinicians*. New York: Oxford University Press; 2007.

Kopelman MD. Disorders of memory. *Brain*. 2002;125:2152–2190.

Kopelman M. Opinion What's Wrong and What's Right in Neuropsychology and Neuropsychiatry. *Cognitive Neuropsychiatry*. 1996;1:275–288.

Krishnan V, Leung LY, Caplan LR. A neurologist's approach to delirium: diagnosis and management of toxic metabolic encephalopathies. *Eur J Int Med*. 2014;25:112–116.

Lamont PJ. Cognitive decline in a young adult with pre-existent developmental delay–what the adult neurologist needs to know. *Practical Neurology*. 2004;4:70–87.

Lanata S, Miller B. The behavioural variant frontotemporal dementia (bvFTD) syndrome in

psychiatry. *J Neurology Neurosurg Psychiatry*. 2016;87:501–511.

Lauterbach E, Cummings J, Kuppuswamy P. Toward a more precise, clinically–informed pathophysiology of pathological laughing and crying. *Neurosci Biobehav Rev*. 2013;37:1893–1916.

Leypoldt F, Armangue T, Dalmau J. Autoimmune encephalopathies. *Ann Ny Acad Sci*. 2015;1338:94–114.

Mesulam MM. Behavioral neuroanatomy: large-scale networks, association cortex, frontal syndromes, the limbic system, and hemispheric specializations. In: Mesulam MM, ed. *Principles of Behavioral and Cognitive Neurology*. London: Oxford University Press; 2000:1–120.

Ovsiew F. Seeking reversibility and treatability in dementia. *Semin Clin Neuropsychiatry*. 2003;8:3–11.

Ovsiew F. Introduction to the psychiatry of conversion disorders. In: Hallett M, Lang AE, Jankovic J, et al., eds. *Psychogenic Movement Disorders and Other Conversion Disorders*. Cambridge: Cambridge University Press; 2011:1–5.

*Ovsiew F, Price B, Murray E. Neuropsychiatric evaluation of the psychiatric inpatient. In: Ovsiew F, Munich R, eds. *Principles of Inpatient Psychiatry*. Philadelphia, PA: Lippincott Williams & Wilkins, 2008:97–124.

Ovsiew F, Silver JM. Unexplained neuropsychiatric symptoms. In: Coffey CE, McAllister TW, Silver J, eds. *Guide to Neuropsychiatric Therapeutics*. Philadelphia, PA: Lippincott Williams & Wilkins, 2007:355–357.

Pievani M, de Haan W, Wu T, Seeley WW, Frisoni GB. Functional network disruption in the degenerative dementias. *Lancet Neurol*. 2011;10:829–843.

Schmahmann JD, Pandya DN. *Fiber Pathways of the Brain*. New York: Oxford University Press; 2006.

Schmahmann JD, Sherman JC. The cerebellar cognitive affective syndrome. *Brain*. 1998;121(Pt 4):561–579.

Sedel F, Baumann N, Turpin J-C, Lyon-Caen O, Saudubray J-M, Cohen D. Psychiatric manifestations revealing inborn errors of metabolism in adolescents and adults. *J Inherit Metab Dis*. 2007;30:631–641.

Seeley WW, Crawford RK, Zhou J, Miller BL, Greicius MD. Neurodegenerative diseases target large-scale human brain networks. *Neuron*. 2009;62:42–52.

Somers K, Lennon V, Rundell J, Pittock S, Drubach D, Trenerry M, et al. Psychiatric manifestations of voltage-gated potassium-channel complex autoimmunity. *J Neuropsychiatry Clin Neurosci*. 2011;23:425–433.

*Stuss DT, Alexander MP. Is there a dysexecutive syndrome? *Philos Trans R Soc Lond B Biol Sci*. 2007;362:901–915.

Walterfang M, Bonnot O, Mocellin R, Velakoulis D. The neuropsychiatry of inborn errors of metabolism. *J Inherit Metab Dis*. 2013;36:687–670.

Wortzel HS, Arciniegas DB. A forensic neuropsychiatric approach to traumatic brain injury, aggression, and suicide. *J Am Acad Psychiatry Law*. 2013;41:274–286.

▲ 2.2 Neuropsychiatric Aspects of Cerebrovascular Disorders

SARAH T. MCCUTCHEON, M.D., AND ROBERT G. ROBINSON, M.D.

INTRODUCTION

Definition

Cerebrovascular disorders consist of a large number of disorders that disrupt the vascular supply to the brain. These disorders are divided into ischemic (i.e., absence of blood flow) and hemorrhagic (i.e., bleeding in or around the coverings of the brain) etiologies. They may also be classified according to the size of the affected blood vessel. Large artery disease produces large lesions of the brain, while small artery (arteriolar) disease results in small lacunar or hyperintensity lesions on MRI. Vascular lesions may occur in any region of the brain.

Clinicians have recognized the association between neuropsychiatric disorders and cerebrovascular disease for over a century. However, this subject has only been the subject of systematic scientific scrutiny for the last 30 years. As the vast majority of studies have focused on patients with stroke, this chapter will focus on the neuropsychiatric correlates of stroke.

A stroke is defined as the sudden loss of blood supply to the brain leading to permanent tissue damage. Nearly 90 percent of strokes are due to ischemia while the remaining 10 percent are caused by hemorrhagic events. While less common, hemorrhagic strokes tend to be more severe and are associated with greater mortality than ischemic strokes.

As the most common serious neurological disorder in the world, stroke is associated with an enormous global public health burden. The World Health Organization estimates that 15 million people are diagnosed with a stroke each year worldwide. The age-specific incidence of stroke varies dramatically over the lifespan. According to the 2015 Update of Heart Disease and Stroke Statistics published by the American Heart Association, the overall annual incidence of stroke in children in the United States is 4.6 to 6.4 per 100,000 children. In adults, stroke incidence increases considerably with age, and the incidence of first stroke is greatest in those 85 years or older. In adults ages 55 to 64, the annual incidence of first-ever stroke is 2.2 to 5.3 per 1,000 persons, while the incidence is 8.9 to 13.8 per 1,000 persons in those 75 to 84 years old. In the United States alone, 795,000 strokes and approximately 130,000 stroke-related deaths occur annually. Stroke is the fourth-leading cause of death in the United States and the second-leading cause of death globally.

History

Early reports of depression after brain damage (usually caused by cerebrovascular disease) were made by neurologists and psychiatrists in case descriptions. Adolf Meyer warned that new discoveries of cerebral localization, such as language function, in the early 1900s led to an overly hasty identification of centers and functions of the brain. He identified that several clinical syndromes such as delirium, dementia, and aphasia were the direct result of brain injury. However, in keeping with his view of biopsychosocial causes of most mental "reactions," he saw manic-depressive illness and paranoiac conditions as arising from a combination of brain injury (specifically citing left frontal lobe and cortical convexities), family history of psychiatric disorder, and premorbid personal psychiatric disorders. Eugen Bleuler described "melancholic moods lasting for months and sometimes longer appear frequently" in patients after strokes. Emil Kraepelin recognized an association between manic-depressive insanity and cerebrovascular disease. He stated, "the diagnosis of states of depression may offer difficulties, especially when arteriosclerosis is involved." Kraepelin concluded that cerebrovascular disorder may be an accompanying phenomenon of manic-depressive disease or may itself produce depressive disorder.

Pathological laughter and crying (PLAC) is another emotional disorder that has only been associated with brain injuries, including cerebral infarction, and represents one of the differential diagnoses for depression in stroke patients. In 1956, Redvers Ironside described the clinical manifestations of this condition, noting that patients' emotional displays were characteristically unrelated to their reported inner emotional state. Laughing and/or crying may occur spontaneously or after some seemingly minor provocation. This phenomenon has been given various names, including emotional incontinence, emotional lability, pseudobulbar affect, pathological emotionalism, and involuntary emotional expression disorder.

Another emotional abnormality, also thought to be characteristic of brain injury, is the indifference reaction described by Derek Denny-Brown in 1952. Associated with right-hemispheric lesions, this reaction was described as indifference toward failures; lack of interest in family and friends; enjoyment of foolish jokes; and minimization of physical difficulties. In the late 19th century, Leonore Welt first described euphoria and loquaciousness associated with orbital frontal injury. Hermann Oppenheim used the term "witzelsucht" to refer to the inappropriate humor in these patients, while Karl Kleist stated that the orbital frontal cortex was the center of emotional life and the dorsal lateral frontal cortex was the source of psychomotor and intellectual activity.

Kurt Goldstein described the catastrophic reaction, another neuropsychiatric disorder historically associated with brain injury such as stroke. He characterized this reaction as an emotional outburst involving various degrees of anger, frustration, depression, tearfulness, refusal,

shouting, swearing, and sometimes aggressive behavior. Goldstein ascribed this reaction to the inability of the organism to cope when faced with a serious defect in its physical or cognitive functions.

Anosognosia is a term introduced by Joseph Jules François Félix Babinski to indicate a patient's lack of awareness of his or her hemiplegia. In Babinski's words, he described this condition as "a mental disorder which I had the opportunity to observe in cerebral hemiplegia, consisting in the fact that the patients ignore or seem to ignore the paralysis which affects them." This term has been extended to unawareness of other deficits after stroke, such as cortical blindness, hemianopia, and amnesia.

Comparative Nosology. The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) defines poststroke psychotic, mood, and anxiety disorders as disorders due to another medical condition, stroke. Relevant specifier(s) are applied to the diagnosis to provide additional details about the patient's clinical presentation. For example, as per DSM-5, a poststroke depressive syndrome meeting full criteria for a major depressive episode would be called depressive disorder due to stroke, with major depressive-like episode. Poststroke subsyndromal depression, also known as minor depression, provides another instructive case example. Common following stroke, subsyndromal depressive symptoms are characterized by more than 2 but less than 5 of the required symptoms for a major depressive episode and persist for at least 2 weeks. According to DSM-5 nomenclature, such a clinical scenario is classified as depressive disorder due to stroke, with depressive features.

The diagnosis of poststroke anxiety based on DSM-5 criteria is termed "anxiety disorder due to stroke" regardless of the phenomenological features of the anxiety. For example, one patient may have anxiety symptoms suggestive of panic disorder, and another patient may have symptoms more consistent with generalized anxiety disorder (GAD). Both of these patients would receive diagnoses of "anxiety disorder due to stroke" according to DSM-5.

The only DSM-5 disorders that are specific for cerebrovascular disease are the major or minor vascular neurocognitive disorders. DSM-5 identifies two forms of cognitive impairment due to vascular disease: major and minor vascular neurocognitive disorders. The term major neurocognitive disorder due to vascular disease replaces the term vascular dementia which was used in the fourth edition of the DSM (DSM-IV). Minor vascular neurocognitive disorder is now used in lieu of mild cognitive impairment. The other aforementioned neuropsychiatric disorders which are specific to brain injuries, such as PLAC or catastrophic reactions, do not have defined diagnostic criteria in DSM-5.

Investigators of depression that co-occurs with physical illness have also debated the most appropriate method for the diagnosis of these disorders

when some symptoms (e.g., sleep or appetite disturbances) could result directly from the physical illness. The following four approaches have been used to assess depression in this complicated scenario:

1. *The inclusive approach:* depressive diagnostic symptoms are counted regardless of whether they may be related to physical illness.
2. *The etiological approach:* a depressive symptom is counted only if the diagnostician feels that it is not caused by the physical illness. Of note, this is the method to be used when making DSM-5 diagnoses.
3. *The substitutive approach:* other psychological symptoms of depression replace the neurovegetative symptoms.
4. *The exclusive approach:* symptoms are removed from the diagnostic criteria if they are not found to be more frequent in depressed than nondepressed patients.

In 1997, Sergio Paradiso and colleagues examined the utility of these methods in the diagnosis of depression during the first 2 years following stroke. Among 205 patients with acute stroke, 142 patients were followed up for examination at 3, 6, 12, or 24 months following stroke. Of 142 patients with follow-up, 60 (42 percent) reported the presence of a depressed mood (depressed group) while they were in hospital, and the remaining 82 patients were nondepressed. Depressive symptoms were divided into vegetative (e.g., autonomic anxiety or subjective anergia) and psychological (e.g., worrying or hopelessness) subtypes. The only vegetative symptoms that were not significantly more frequent in the depressed compared to the nondepressed patients at the different follow-up intervals were as follows: at the initial evaluation, weight loss and early morning awakening; at 6 months, weight loss and early morning awakening; at 1 year, weight loss, early morning awakening, anxious foreboding, and loss of libido; and at 2 years, weight loss and loss of libido. Among the psychological symptoms, the depressed patients had a higher frequency of most of the psychological symptoms throughout the 2-year follow-up. The only psychological symptoms that were not significantly more frequent in the depressed than in the nondepressed group were as follows: at 3 months, suicidal plans, simple ideas of reference, and pathological guilt; at 6 months, pathological guilt; at 1 year, pathological guilt, suicidal plans, guilty ideas of reference, and irritability; and at 2 years, pathological guilt and self-depreciation.

These findings are supported by a more recent study by Cumming and colleagues from 2010. This study included 149 stroke patients at 18 months poststroke and 745 age- and gender-matched controls from the general population. A major depressive episode was diagnosed in 25 of 149 (17 percent) stroke patients and 61 of 745 (8 percent) of controls. When symptom profiles were compared across these two depressed groups, the controls reported greater “inability to feel” in comparison to the stroke

group. Otherwise, there were no statistically significant differences between symptom profiles in the groups. Moreover, a factor analysis of 10 psychological and somatic symptoms of depression showed no difference between the stroke and control groups, regardless of depression diagnosis.

In sum, these two studies suggest that the phenomenology of poststroke depression (PSD) is similar to depression in the general population. In other words, the use of the same diagnostic criteria for depression in these two populations appears to be valid. Thus, DSM-5 defines depressive disorder due to stroke according to the same symptoms as major depression occurring in patients without stroke.

EPIDEMIOLOGY

Major Vascular Neurocognitive Disorder

In 2000, Antonio Lobo and colleagues performed a review of the prevalence of dementia in Europe based on population-based studies. Based on a total of 2,346 cases of mild to severe dementia, the age-standardized prevalence was 6.4 percent for all-cause dementia, 4.4 percent for dementia due to Alzheimer disease (dAD), and 1.6 percent for vascular dementia. The authors found a consistent increase in the lifetime prevalence of vascular dementia with advancing age. The pooled prevalences for those ages 65 to 69 years was 0.5 percent in men and 0.1 percent in women, while the rates were 3.6 percent in men and 5.8 percent in women 90 years and older. In 2007, the Aging, Demographics, and Memory Study estimated the prevalence of dementia in the United States. Composed of approximately 850 elders aged 71 and older, the prevalence of dementia overall was 13.9 percent, equating to 3.4 million people. The prevalence of vascular dementia was 2.43 percent. In Asia, vascular dementia was more common than dAD until the last two decades. In a review of the epidemiology of dementias in developing countries, Raj Kalaria and coauthors concluded that the prevalence of vascular dementia appears to be low except for some of Latin America and Asian countries. Overall, they estimate the prevalence of vascular dementia in developing countries to be 0.6 to 2.1 percent.

Depression

Several reviews have examined the prevalence of depression following stroke. In the most recent, Luis Ayerbe and colleagues evaluated data from 43 studies including 20,293 patients. They found the pooled prevalence of depression observed at any time point was 29 percent. The prevalence rates did not differ significantly over time or in studies of different settings. The cumulative incidence of PSD was 39 to 52 percent within 5 years of stroke. In another large review of 51 cohorts and 25,207 patients, Maree L. Hackett's group estimated that 33 percent of stroke patients experience

depression. Both reviews emphasize that the estimation of the prevalence of PSD is confounded by heterogeneity in important clinical variables, such as the type of depression (e.g., syndromic vs. subsyndromic depression) or population type (e.g., hospital-based vs. outpatient sample). To address these limitations, Bob Robinson and coinvestigators conducted a systematic review of these factors in 2006. Among the 2,108 patients in community-based settings, the prevalence rate of major depression was 14.1 percent, while the rate of minor depression was 9.1 percent. A total of 2,769 patients were treated in acute or rehabilitation hospitals. Of these patients, 21.6 percent met criteria for major depression, and an additional 20.0 percent had minor depression. Among studies of outpatients, the prevalence rates of major and minor depression were 24.0 and 23.9 percent, respectively.

Mania

Compared to PSD, mania after stroke is a rare neuropsychiatric occurrence, and the literature consists primarily of case reports or small case series. Catarina Santosa and colleagues performed a systematic review of poststroke mania in 2011 to better understand its demographic and clinical features. Their literature search covered approximately 50 years and yielded a total of 74 cases of mania in adult stroke patients. The authors observed the typical patient was male without a personal or family history of psychiatric illness; with at least one vascular risk factor; without subcortical atrophy; and with a right cerebral infarct. Overall, this study demonstrates the rarity of mania after stroke.

Anxiety

A recent meta-analysis from 2013 of 44 reports and 5,760 stroke patients found the pooled estimate of poststroke anxiety disorders, as determined by clinical interview, was 18 percent. When anxiety symptoms were assessed by a rating scale, the prevalence of anxiety was 25 percent. Three of the included studies which examined the relative frequency of the different anxiety disorders and found that phobic disorders and GAD were the most common. In addition, the prevalence of anxiety tended to increase with time. The prevalence was 20 percent in the month following stroke; 23 percent between 1 and 5 months; and 24 percent had anxiety symptoms at 6 months or more. These results suggest that anxiety symptoms are common after stroke and tend to persist. The review also found similar prevalence estimates regardless of study setting (population-based vs. hospital vs. rehabilitation vs. community-based).

A significant number of patients with a poststroke anxiety disorder also have depression. A review of 1,445 patients within 3 months following stroke found 49 percent met criteria for GAD without comorbid depression,

while 14 percent had comorbid GAD and major depression. Among community samples, the rate of GAD alone was 2 percent and GAD with depression was 8 percent. Hospital and outpatient samples found that GAD alone occurred in 5.5 percent of cases, while GAD with depression occurred in 15.2 percent. Thus, anxiety disorders following stroke are frequently comorbid with a depressive disorder, although a significant number of patients have anxiety alone.

While numerous case reports exist, a paucity of data is available from systematic studies of the frequency of different types of anxiety disorders (e.g., phobia or panic disorder) after stroke. However, a growing literature now pertains to the occurrence of PTSD after stroke. For instance, Judith Z. Goldfinger and coinvestigators evaluated 535 patients who had a stroke or transient ischemic attack (TIA) and found that 18 percent likely had PTSD. In general, those with likely PTSD had worse physical and mental health. Predictors of poststroke PTSD included low income, a history of recurrent stroke or TIA, more disability, and greater medical comorbidity. Older age, presence of a romantic partner, and having emotional support were protective against developing PTSD. While this literature is still in its infancy, such studies are a valuable contribution to studies of PTSD after a variety of acute medical events.

Psychosis

Although rare, case reports and empirical studies have documented that psychosis, primarily identified by the presence of delusions or hallucinations, may occur after stroke. The incidence or prevalence of psychosis following stroke has yet to be evaluated in a systematic epidemiological study. Two hospital registry studies evaluated the epidemiological characteristics of poststroke psychosis. The reported frequency of psychotic disorders was 0.4 percent and 3.1 percent of stroke patients with an incidence of 1.1 per 1,000 person-years and a median time to onset of 6.1 months following stroke.

Apathy

Although DSM-5 does not include diagnostic criteria for this condition, apathy is typified by the absence or lack of feeling, emotion, interest, concern, or motivation and has been reported frequently among patients with brain injury, including stroke. Over the last 5 years, apathy has become one of the most intensely studied stroke-related disorders. Table 2.2–1, a summary of several recent studies, shows the reported prevalence of apathy in poststroke patients.

A recent meta-analysis of 2,706 patients from 24 cohorts found a mean prevalence of apathy of 34.6 percent at a median of 120 days following stroke. Depression occurred in 40.1 percent of patients with apathy, which

suggests that apathy often occurs distinctly from depression. The estimated relative risk of depression among apathetic patients was 1.8. A very recent longitudinal, 1-year study found that 41.1 percent of patients met diagnostic criteria for apathy during the first year after stroke. Further, the mean months to onset was 3.8 ± 3.3 months, and the mean duration of meeting diagnostic criteria for apathy was 5.6 ± 2.3 months, as seen in Figure 2.2–1. In a linear mixed model, the authors controlled for age, initial severity of impairment, and major depression and found that patients in the apathy group had significantly less recovery in cognition and activities of daily living (ADLs) during the first year after stroke compared with nonapathic patients.

Table 2.2–1.
Recent Studies of the Prevalence of Poststroke Apathy

Study	N	Apathy Frequency	Associated Variables based on Multivariate Models
Brodaty et al. (2013)	Index: 135 Year 5: 70	Index: 32% Year 1: 33% Year 3: 37% Year 5: 37%	Dementia New stroke ADLs impairment
Caeiro et al. (2012)	94	38%	Cerebral hemorrhage Low education
Caeiro et al. (2013)	Meta-analysis: 2,221	36%	Depression Cognitive impairment
Mayo et al. (2009)	408	50% low score 3% high score 7% improve	High score related to poor recovery
Tang et al. (2013)	75	33%	Depression
Tang et al. (2014)	391	9%	Poor mental quality of life
van Dalen et al. (2013)	Meta-analysis: 2,706	35%	Female gender Older age Cognitive impairment Greater disability Depression

ADLs, activities of daily living.

Brodaty H, Liu Z, Withall A, Sachdev PS. The longitudinal course of post-stroke apathy over five years. *J Neuropsychiatry Clinical Neurosci*. 2013;25(4):283.

Caeiro L, Ferro JM, Costa J. Apathy secondary to stroke: a systematic review and meta-analysis. *Cerebrovasc Dis*. 2013;35(1):23.

Caeiro L, Ferro JM, Figueira ML. Apathy in acute stroke patients. *Eur J Neurol*. 2012;19(2):291.

Mayo NE, Fellows LK, Scott SC, Cameron J, Wood-Dauphinee S. A longitudinal view of apathy and its impact after stroke. *Stroke*. 2009;40(10):3299.

Tang WK, Chen YK, Liang HJ, Chu WC, Mok VC, Ungvari GS, Wong KS. Location of infarcts and apathy in ischemic stroke. *Cerebrovasc Dis*. 2013;35(6):566.

Tang WK, Lau CG, Mok V, Ungvari GS, Wong KS. Apathy and health-related quality of life in stroke. *Arch*

Phys Med Rehabil. 2014;95(5):857.
van Dalen JW, Moll van Charante EP, Nederkoorn PJ van Gool WA, Richard E. Poststroke apathy. *Stroke.* 2013;44:851.

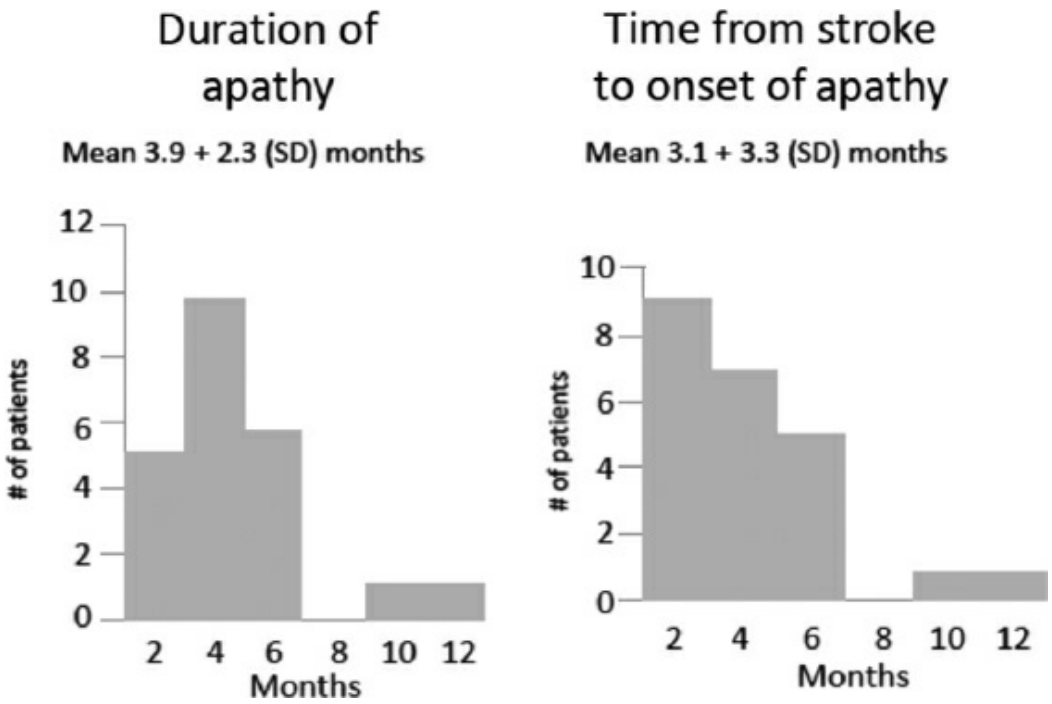

FIGURE 2.2–1. Incident apathy after stroke. Left: The number of patients who met diagnostic criteria as proposed by Robert et al. for apathy for the duration indicated by the right edge of the column (e.g., 5 patients had 3 months' duration, whereas in the next column 10 patients had 4.5 months' duration). Right: The approximate time from initial evaluation to onset of apathy. The column on the left shows nine patients had onset at the initial evaluation, whereas the next column shows seven patients had onset at 1.5 months (i.e., no apathy at initial evaluation but present at 3 months). Thus, except in a few cases, apathy occurred within the first 5 months after stroke and lasted almost 6 months. (Reprinted with permission from Mikami KI, Jorge RE, Moser DJ, Jang M, Robinson RG. Incident apathy during the first year after stroke and its effect on physical and cognitive recovery. *Am J Geriatr Psychiatry*. 2013;21(9):848.)

Catastrophic Reaction

First described by Goldstein, catastrophic reactions are characterized by anxiety, tears, aggressive behavior, swearing, displacement, refusal, renouncement, and/or, sometimes, compensatory boasting. In an early study, the Catastrophic Reaction Scale (CRS), which was developed to assess the existence and severity of catastrophic reactions, was introduced. Based on a survey of 62 consecutive patients, 12 (19 percent) with acute stroke lesions were found to have catastrophic reactions. Of these 12, 9 (75 percent) met criteria for major depression, and an additional 2 had minor depression. Further, the 12 patients with catastrophic reactions showed elevated scores on a measure of anxiety, the Hamilton Anxiety Scale. In a larger, prospective study, 202 consecutive patients received the CRS within 4 days of stroke. Anger was seen in 71 (35 percent) patients, and 26 of these

were severely angry. In a third study, the 10-item Spielberger Trait Anger Scale was used, and 47 of 145 (32 percent) of outpatients had evidence of the inability to control anger or aggression 3 to 12 months poststroke.

Pathological Laughter and Crying

In a recent cohort study of 508 consecutive patients with ischemic stroke, 9.4 percent of patients had evidence of pathological emotions during their admission for care of acute stroke. Three months later, the frequency increased to 11.7 percent. Yet, in a separate study of 516 patients at 3 months after stroke, the estimate of PLAC was 7.6 percent. In a third study, 127 stroke patients who initially received care at a university-affiliated regional hospital were evaluated for PLAC 3 months poststroke. When caregivers were involved in the assessment, the prevalence of PLAC was 17.9 percent. In contrast, the estimate fell to 6.3 percent when only patients were evaluated. Such studies illustrate how relevant clinical factors (e.g., timing of study; involvement of informant) can significantly influence prevalence rates.

Anosognosia

Estimates of the prevalence of anosognosia vary considerably in the literature. Several factors contribute to this heterogeneity. For instance, the inclusion of patients with aphasia or dementia tends to conflate estimates due to significant language and cognitive limitations, respectively. Further, anosognosia is uncommon beyond 3 months poststroke. Therefore, the timing of study assessment is another important factor in this literature. In the only meta-analysis to date, Lorenzo Pia and colleagues identified 52 studies from 1938 to 2001 for study inclusion. They found the frequency of anosognosia for hemiplegia to range from 20 to 44 percent depending on when the study was conducted relative to stroke occurrence. Further, of patients with complete contralateral motor impairment, the prevalence of anosognosia was 32.3 percent.

ETIOLOGY

Depression

A number of hypotheses about the etiology of PSD have been postulated. Numerous studies have found abnormalities in biomarkers linked to mood dysregulation such as the serotonin transporter gene and certain inflammatory cytokines, particularly interleukin 6 and 18. Elevated levels of glucocorticoids, which affect glutamate transmission, and stroke location, which results in the interruption of biogenic amine transmission, are additional findings. Of these factors, lesion location has been a topic of considerable interest.

In studies performed 3 to 12 months after stroke, the prevalence of PSD

is similar in those with left- and right-hemispheric lesions. However, in the first few months after stroke, lesion location does appear to increase the risk of PSD. This relationship has been the subject of recent meta-analyses. The study conducted by Robinson and colleagues evaluated the frequency of depression within 2 months following stroke. Of the 128 patients with left anterior and left posterior strokes, the presence of a left anterior stroke increased the odds for PSD by 2.3 in comparison to the posterior lesion. In a comparison of right anterior and left anterior strokes, having a left anterior lesion increased the odds of PSD by 2.18 (Fig. 2.2–2). In a separate meta-analysis of 8 studies conducted within 6 months of stroke, proximity of the anterior border of the stroke lesion to the left or right frontal pole and severity of PSD was evaluated. The authors found a significant inverse relationship between severity of depression and distance of the stroke from the left frontal pole, but not right frontal pole. The third and most recent meta-analysis paints a different picture. Performed by Na Wei and colleagues, the meta-analysis consisted of 43 studies and approximately 5,500 patients. The overall analysis failed to show a relationship between lesion location and PSD. However, the impact of time since stroke was evaluated in subsequent subgroup analyses, and a modest, albeit statistically significant relationship (odds ratio of 0.79), was found between right-hemispheric stroke and PSD risk within the subacute (i.e., 1 to 6 months poststroke) group.

In sum, while patients appear to be at elevated risk for PSD within the first few months after stroke, meta-analyses have reported conflicting results regarding the association between PSD and lesion location. Such conflicting results exemplify a fundamental concept seen across neuropsychiatry: specific clinical symptoms are unlikely to be the product of an isolated lesion, but instead, due to dysfunction of key neurocircuits. Further, a major problem with these meta-analyses, as indicated earlier, is that combining findings across studies frequently overlooks potential modifying variables. For example, specific types (i.e., major vs. minor depression), specific symptoms of depression (i.e., neurovegetative vs. psychological), and type of brain injury (e.g., ischemic vs. hemorrhagic) may be other crucial variables that are overlooked by meta-analytic techniques.

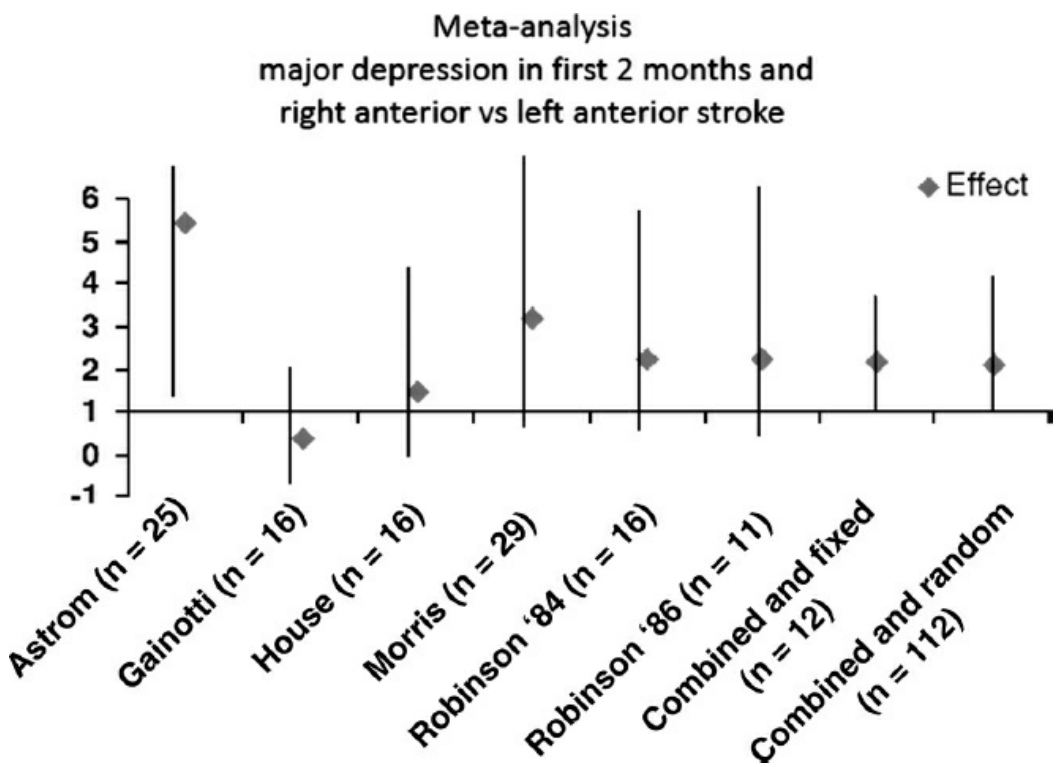

FIGURE 2.2–2. A meta-analysis involving 112 patients was conducted comparing the relative risk of major depression following left anterior versus right posterior stroke. Meta-analysis based on the fixed model effect was 2.18 (CI 1.40 to 3.30, p = 0.000) and on the random model was 2.16 (CI 1.30 to 3.60, p = 0.004). The random model is a more conservative statistic that assumes that there are random variations in the interstudy variance that the fixed model does not. These findings mean that, during the acute stroke period, patients with left frontal or left basal ganglia lesions are more than two times more likely to have a major depression than patients with comparable lesions of the right hemisphere. (Reprinted with permission from Narushima K, Kosier JT, Robinson RG. A reappraisal of poststroke depression, intra- and interhemispheric lesion location using meta-analysis. *J Neuropsychiatry Clin Neurosci.* 2003;15:422.)

Several studies have evaluated the clinical correlates of PSD. The most consistent finding has been increased severity of impairment in ADLs in those with PSD. In 2005, Hackett and colleagues performed a systematic review of observational studies. In 9 of 11 studies evaluating this factor, disability after stroke showed a positive and significant association with PSD. In a separate review by Robinson's group, 15 out of 18 studies (83 percent) found a statistically significant relationship between the severity or existence of PSD and severity of impairment in ADLs. However, the temporal relationship between these factors (i.e., whether the depression preceded or followed the severe functional impairment) was not examined in most of these studies.

In addition, numerous studies have demonstrated that a personal history or a family history of psychiatric disorders is an important risk factor for PSD. For example, an Australian study of 99 patients in a poststroke rehabilitation hospital found that 11 of 16 patients (69 percent) with major depression following right- or left-hemispheric stroke had a

family history of mood or anxiety disorders compared with 5 of 18 with minor depression (28 percent) and 20 of 54 (37 percent) who were not depressed. There were similar findings for major depression and prior personal history of psychiatric disorder (i.e., 8 of 16 with major depression vs. 14 of 54 nondepressed, $p = 0.04$).

This combination of biological, clinical, and psychological correlations has led many investigators to hypothesize a biopsychosocial etiology for PSD. This may be a helpful clinical construct for the explanation for PSD as a whole, but this remains to be tested systematically.

Mania

Similar to other neuropsychiatric disorders, the etiology of mania remains unknown. The relatively rare occurrence of mania following stroke suggests that a combination of variables may be needed to produce this disorder. For example, a study of 17 patients with stroke and mania (i.e., DSM-IV diagnosis of mood disorder due to stroke, with manic features) found that 12 had unilateral right-hemispheric lesions. The frequency of right-hemispheric lesions was significantly higher compared to 28 patients with major depression, who tended to have left frontal or basal ganglia lesions, or patients with no mood disorder following stroke. Lesions associated with mania were either cortical (basotemporal or orbitofrontal cortices) or subcortical (frontal white matter, basal ganglia, or thalamus). A PET study using [^{18}F]fluorodeoxyglucose (FDG) showed focal hypometabolic deficiency in the right basotemporal cortex in three manic patients with right subcortical lesions, but this hypometabolism was not seen in seven, age-comparable, normal controls. A recent systematic review of single or small series of cases of mania found, among 49 patients, associations with male gender, vascular risk factors, and right-hemispheric infarcts.

Thus, mania appears to be provoked by injury to specific right-hemispheric structures that have connections to the limbic system. The right basotemporal cortex may be particularly important because direct lesions as well as distant hypometabolic effects of this cortical region, which is suggestive of diaschisis, were associated with secondary mania. The basotemporal cortex has strong efferent connections to the orbitofrontal cortex suggesting that the lateral orbitofrontal circuit in the right hemisphere may play a role in the etiology of mania. A combination of biogenic amine system dysfunction and release of tonic inhibitory input to the orbitofrontal–thalamic circuit may lead to the production of mania.

Another study compared patients with secondary mania to patients with no mood disturbances. Participants were matched for size, location, and etiology of brain lesion. Patients with secondary mania had a significantly greater degree of subcortical atrophy, as measured by the ratio of bifrontal and third ventricular volume to brain volume. Moreover, of the patients who developed secondary mania, those with a positive family history of

psychiatric disorders had significantly less atrophy than those without such a family history. This implies that genetic predisposition to affective disorders and brain atrophy may be independent risk factors for poststroke mania. Although the aforementioned review study of 49 cases did not report associations with subcortical atrophy or family history of psychiatric disorder, the included studies did not tend to examine for measures of these two factors.

Psychosis

Information about the mechanism of poststroke psychosis is derived from anecdotal reports or small case series. One study of five patients with psychosis following stroke found that all patients had right-hemispheric lesions, primarily involving frontoparietal regions. When compared with five patients matched for age, education, and lesion size and location, but no psychosis, patients with secondary psychosis had significantly greater subcortical atrophy. This atrophy was manifested by larger areas of both the frontal horn of the lateral ventricle and the body of the lateral ventricle on the side contralateral to the brain lesion. Several investigators have also reported a high frequency of seizures among patients with secondary psychosis. These seizures usually started after the brain lesion but before the onset of psychosis. The aforementioned study of patients with poststroke psychosis compared with lesion-matched controls found a seizure disorder among three of five patients with poststroke psychosis, as compared to zero of five stroke patients without psychosis. Overall, a right-hemispheric lesion, seizures, and subcortical brain atrophy appear to be important factors in the pathogenesis of poststroke psychosis.

Apathy

The mechanism of poststroke apathy has been the focus of many recent studies. For example, in a sample of 55 chronic stroke patients, patients who showed little behavioral response to positive reinforcement in the laboratory reported higher levels of apathy. These patients were also found to have lesions of the ventral putamen and globus pallidus, as well as dorsal thalamus and prefrontal cortex. Another study of 106 acute stroke patients found that apathy at the time of acute stroke predicted apathy at 1-year follow-up. A meta-analysis reported that apathy among 2,221 patients was associated with PSD and greater cognitive impairment. Imaging studies have found that apathy is significantly associated with lesions of the pons, frontal cortex, basal ganglia, dorsal thalamus, posterior limb of the internal capsule, and temporal cortex (Table 2.2–2).

Overall, lesions along the anterior cingulate subcortical circuit (including cingulate gyrus, frontal cortex, ventral striatum, ventral pallidum, and magnocellular dorsomedial thalamus) have been frequently

associated with the occurrence of apathy and may constitute an anatomical substrate for poststroke apathy.

Table 2.2–2.
Recent Imaging Findings in Poststroke Apathy

Author	N	Caseness	Main Findings
Hama et al. (2007)	243	126 with apathy or depression	Apathy associated with bilateral BG damage; PSD associated with left frontal lobe damage
Murakami et al. (2013)	149	149 total cases	Apathy associated with lesions of brainstem and bilateral striatum
Onoda et al. (2011)	102	37 with apathy, 65 controls	CBF reduced in basal ganglia in apathetic cases; hypoperfusion in bilateral-basal ganglia associated with presence of apathy
Tang et al. (2013)	185	20 with apathy, 165 controls	Pontine infarcts associated with apathy; volume of pontine infarct significantly predicted occurrence of poststroke apathy

BG, basal ganglia; CBF cerebral blood flow; PSD, poststroke depression.

Hama S, Yamashita H, Shigenobu M, et al. Post-stroke affective or apathetic depression and lesion location: left frontal lobe and bilateral basal ganglia. *Eur Arch Psychiatry Clin Neurosci*. 2007;257(3):149.

Murakami T, Hama S, Yamashita H, et al. Neuroanatomic pathways associated with poststroke affective and apathetic depression. *Am J Geriatr Psychiatry*. 2013;21(9):840.

Onoda K, Kuroda Y, Yamamoto Y, et al. Post-stroke apathy and hypoperfusion in basal ganglia: SPECT study. *Cerebrovascular Dis*. 2011;31(1):6.

Tang WK, Chen YK, Liang HJ, et al. Location of infarcts and apathy in ischemic stroke. *Cerebrovascular Dis*. 2013;35(6):566.

Catastrophic Reaction

In a study of 62 patients with acute stroke, those demonstrating catastrophic reactions had a significantly higher frequency of lesions involving the basal ganglia compared to acute stroke controls. When 10 depressed patients with a catastrophic reaction were compared to 10 depressed patients without a catastrophic reaction, the catastrophic reaction group had significantly more anterior lesions, which were located primarily in subcortical regions (i.e., eight of nine depressed patients with catastrophic reaction had subcortical lesions; three of nine depressed patients without catastrophic reaction had subcortical lesions).

On the basis of these findings, the catastrophic reaction may result from neurophysiological dysfunction rather than realization of intellectual impairment. Catastrophic reactions occurred predominantly in patients with major depression associated with anterior subcortical lesions. Subcortical damage has also been hypothesized to underlie the “release” of emotional display by removing inhibitory input to limbic areas of the

cortex.

Pathological Laughter and Crying

Pathological emotions have traditionally been explained as secondary to the bilateral interruption of descending neocortical upper motor neuron innervation of bulbar motor nuclei. Some patients with pathological emotions have such bilateral lesions, but this is not a uniform finding. For instance, one study found that patients with frontal or temporal lesions in either hemisphere had a significantly increased frequency of pathological emotions. Examination of lesion size and location in 12 patients with pathological crying found that patients with the most frequent crying episodes had relatively large bilateral pontine lesions. The intermediate group had large bilateral lesions. The least affected patients had relatively large unilateral subcortical lesions. Such evidence has led to the hypothesis that pathological emotions may arise from partial destruction of raphe serotonergic neurons or their projections. More recently, investigators at the University of Iowa suggested that the critical lesions eliciting PLAC are located along fronto-ponto-cerebellar pathways.

DIAGNOSIS AND CLINICAL FEATURES

Mild and Major Vascular Neurocognitive Disorders

As per DSM-5, the most recent classification of vascular cognitive impairment includes both patients with dementia (i.e., major vascular neurocognitive disorder) and its precursor (i.e., mild neurocognitive disorder). In comparison to prior versions of the DSM, the category of vascular neurocognitive disorders has expanded considerably in DSM-5. The following three conditions must be met to reach diagnostic criteria for major vascular neurocognitive disorder:

1. Criteria are met for major neurocognitive disorder (i.e., the presence of acquired impairment in one or more cognitive domains and significant impairment in activities of daily living not better explained by another medical condition);
2. The clinical features are consistent with a vascular etiology as suggested by either (a) cognitive deficits that are temporally related to one or more cerebrovascular events or (b) presence of prominent executive dysfunction (e.g., inattentiveness or slow processing speed);
3. The patient's history, physical examination, and/or neuroimaging findings are consistent with cerebrovascular disease of sufficient severity to account for patient's cognitive impairment.

Patients with mild vascular neurocognitive disorder have impairment in one or more cognitive domains, but daily functioning remains relatively intact. However, affected individuals may take longer than usual or engage

in compensatory techniques to overcome their cognitive disturbances.

Cognitive impairment due to vascular disease is expected to occur in the presence of multifocal motor and/or sensory deficits on the physical examination. Single defects in cognition, such as amnestic states, aphasia, or apraxia, do not fulfill the criteria for either major or mild neurocognitive disorder. On the other hand, a single vascular lesion may result in a clinical picture consistent with major vascular neurocognitive disorder.

Depression

As indicated in the section on Comparative Nosology, the assessment for depression in patients with stroke or other physical illness has been a topic of considerable debate. According to DSM-5, two types of poststroke depressive disorders are recognized: depressive disorder due to stroke with major depression-like episode and depressive disorder due to stroke with depressive symptoms (i.e., sub-syndromal depression). The diagnostic criteria, however, require the clinician to determine whether he or she believes that the mood disorder is the direct physiological consequence of the stroke. If this judgment is made, then the patient is diagnosed with "depression due to stroke with major depressive-like episode."

The problem with this diagnostic schema is that even the most experienced clinician cannot be certain if the depression is due to the direct physiological consequences of the stroke. Even if the patient has a history of depression prior to the stroke, the physiological response to brain injury may or may not provoke a new depressive episode. Furthermore, the clinical manifestations of depression appear to be the same regardless of whether or not the depression followed a stroke. That is, empiric evidence has shown that all of the symptoms of major depression with poststroke onset have been shown to be significantly more frequent in depressed compared with non-depressed stroke patients. Thus, the symptoms of depressed mood, anhedonia, weight loss, insomnia, psychomotor agitation or retardation, loss of energy, worthlessness, poor concentration, and suicidal thoughts characterize patients with PSD just as well as depressed patients without structural brain injury. Thus, a better diagnostic system may be to make stroke a specifier, akin to *postpartum onset*. Namely, the diagnosis of "major" or "minor depression would be qualified with *poststroke onset*. Empirical data appears to support such nomenclature more so than the current diagnostic classification.

To promote the early detection of PSD in patients hospitalized for acute stroke, Janneke de Man-van Ginkel and colleagues developed the Post-Stroke Depression Prediction Scale. The authors validated this scale in a study of 410 stroke patients. Within the first week after stroke, participants were assessed for various PSD predictive factors and then reassessed 6 to 8 weeks after stroke. Based on a multivariate model, a medical history of depression or other psychiatric disorder was the most important predictor

(odds ratio 7.2) of PSD. Other predictive factors included medical history of hypertension, angina pectoris, and impaired ability to dress oneself based on the dressing item of the Barthel Index of ADLs. Further, the lowest risk score at 1-week poststroke was associated with a 2 percent risk of depression at 6 to 8 weeks follow-up while 82 percent of the highest risk group developed depression at 6 to 8 weeks.

Mania

The symptoms of mania were examined in a series of 25 consecutive patients who met DSM-IV criteria for a mood disorder due to brain injury with manic features. These patients, who developed mania after a stroke, traumatic brain injury, or tumors, were compared to 25 patients with primary mania (i.e., no known neuropathology). Both groups of patients showed similar frequencies of elation, pressured speech, flight of ideas, grandiose thoughts, insomnia, hallucinations, and paranoid delusions. Thus, the symptoms of mania that occurred after stroke (now diagnosed by DSM-5 as bipolar disorder due to stroke with manic-like episode if full criteria for mania are met, or bipolar disorder due to stroke with manic features if full criteria for mania are not met) appear to be the same as those found in idiopathic cases of mania (i.e., bipolar disorder). As with depression, a better diagnostic classification might be mania with *poststroke onset*.

Anxiety

A 2013 study including 374 patients found that 23 percent had anxiety defined by a score of 8 or more on the Anxiety Section of the Hospital Anxiety Depression Scale. Patients with poststroke anxiety had poor stroke-related quality of life in general and in several particular areas including energy, mood, personality, thinking, and work. Another 2013 study looked at long-term outcomes of poststroke anxiety in 220 patients followed longitudinally. At 5 years poststroke, 64 (29 percent) of patients were deemed anxious and 73 (33 percent) were depressed, with anxiety and depression comorbidity in 43 (20 percent). Higher anxiety scores at 6 months following stroke were associated with anxiety at 5 years poststroke but not ADLs scores at 6 months. Finally, an imaging study of 693 patients with acute stroke reported that the anxious patients, as defined by a score of 8 or more on the Hospital Anxiety Depression Scale, had a significantly increased frequency of right frontal infarcts even after controlling for other relevant variables.

In sum, anxiety is common after stroke, and its prevalence remains elevated even several years out from the stroke. Further, depression and anxiety after stroke are strongly associated. Given that both depression and anxiety are associated with poor recovery, regular assessments for PSD and

anxiety are warranted.

The following case illustrates a typical clinical course of poststroke anxiety and its frequent comorbidity with depression:

> The patient was a 71-year-old farmer who suffered from thrombosis of the basilar artery. He developed blurring of his vision, gait impairment, and paresthesias of his face. Within 2 months of the stroke, the patient developed panic attacks and GAD. His GAD was characterized by almost constant anxiety about minor issues, insomnia, restlessness, and poor concentration. The panic attacks were typified by the rapid onset of anxiety with tachycardia, shortness of breath, sweating, and fearfulness that he would pass out or die from another stroke or a heart attack. The panic attacks occurred first when he was away from home, and later, while he was at home. The panic attacks were controlled by alprazolam, but in spite of taking this medication four times per day, he continued to have significant anxiety marked by worrying, restlessness, muscle tension, poor concentration, and insomnia. About 2 months later, the patient also developed depression with symptoms of low mood, loss of interest, poor concentration, self-blame, hopelessness, and psychomotor slowing. He responded to electroconvulsive therapy but relapsed quickly. He was then treated with nortriptyline, which led to remission of both his depressive and anxiety disorders.

Table 2.2–3.
Questions from the Apathy Scale Used to Determine the Existence of Lack of Emotion, Slowed Activity, and Impaired Cognitive Activity Seen in Apathy

Lack of Emotion	Slowed Motor Activity	Impaired Cognitive Activity
Are you concerned about your condition?	Does someone have to tell you what to do each day?	Are you interested in learning new things?
Are you indifferent to things?	Do you put much effort into things?	Do you have anything that interests you?

The proposed criteria for the diagnosis of apathy is lack of motivation (e.g., "Do you have motivation?") plus two of the three symptoms above (i.e., lack of emotion, slowed motor activity, and/or impaired cognitive activity) present most of the time for at least 4 weeks based. (Data from Mikami K, Jorge RE, Moser DJ, et al. Prevention of poststroke apathy using escitalopram or problem-solving therapy. *Am J Geriatr Psychiatry*. 2013;21(9):855.)

Apathy

The clinical manifestations of apathy were utilized by Marin and Starkstein in the 1990s to develop an apathy rating scale and to propose diagnostic criteria. The essential symptom of apathy is the loss of motivation or drive which is the required symptom for the diagnosis of apathy. Lack of motivation may be manifested emotionally, cognitively, or behaviorally as reflected in the recently proposed diagnostic criteria for apathy by Roberts and colleagues, as reflected in Table 2.2–3.

A 2013 meta-analysis of 2,221 stroke patients found that apathy was associated with older age, depression, and cognitive impairment. The

frequency of apathy did not differ when the following factors were considered: gender, stroke type (i.e., ischemic vs. hemorrhagic), and lesion location (i.e., left- vs. right-hemispheric lesions). Interestingly, the presence of apathy did not appear to affect clinical outcome when measures of ADLs were examined.

In a study of 80 acute stroke patients, 18 (22.5 percent) had apathy compared to 62 stroke patients without apathy. Nine (11.3 percent) had comorbid apathy and depression. The apathetic patients, with or without depression, were significantly older than the nonapathetic patients. Also, the apathetic patients showed significantly more severe deficits in ADLs. The nine cases who met criteria for both apathy and depression were significantly more impaired in their ADLs than patients with apathy or depression alone. Seiji Hama and colleagues further examined the differential impact of depression and apathy on functional recovery. Based on sample size of 237 patients with hemorrhagic or ischemic stroke, the authors found that apathy was a stronger predictor of poor functional recovery than depression.

Catastrophic Reactions

Currently, no diagnostic criteria are available for catastrophic reactions. As described earlier in the chapter, patients typically present with sudden outbursts of refusal, swearing, tears, and sometimes aggressive behavior. This intense emotional reaction is usually provoked by stressful situations such as cognitive testing and typically subsides within a few minutes.

Catastrophic reactions occurred in 12 of 62 patients (19 percent) admitted to the hospital with acute stroke. Patients with catastrophic reactions were found to have a significantly higher frequency of familial and personal history of psychiatric disorders, mostly depression, than patients without catastrophic reactions. Catastrophic reactions, however, were not significantly more frequent among aphasic compared with nonaphasic patients. This finding did not support the contention that catastrophic reactions represent an understandable psychological response of “frustrated” aphasic patients. Furthermore, 9 of 12 (75 percent) patients with catastrophic reactions also had PSD, 2 (17 percent) had minor depression, and only 1 was not depressed. Thus, catastrophic reaction was closely linked with PSD.

Pathological Laughter and Crying

There are no established diagnostic criteria for PLAC. This condition is characterized by the sudden onset of crying, or more rarely, laughing, which is out of proportion to the conversation or situation in which the emotional reaction occurred. The emotion may last from a few seconds to a few minutes with no residual feelings of sadness or happiness.

At least five studies have examined the clinical features of pathological emotions. The hallmark features of PLAC are as follows: an inability to control crying or laughter; an increased frequency of such emotional display; and recognition by the patient that the emotional display is inconsistent or excessive to his or her underlying emotional feelings. Several rating scales are available for PLAC, but only the Pathological Laughter and Crying Scale (PLACS) was specifically developed for poststroke pathological emotions. As seen in Figure 2.2–3, this scale assesses the existence and severity of pathological emotion among patients with stroke.

PATHOLOGY AND LABORATORY EXAMINATION

Mild and Major Vascular Neurocognitive Disorders

The clinical identification of vascular neurocognitive disorders requires a medical history, neurological examination, psychiatric interview, and neuropsychological assessment. Diagnostic testing also plays an important role and is informed by blood chemistries (including B_{12}, folate, and thyroid function), head CT and/or brain MRI, and, in certain cases, cerebral angiography. These laboratory and imaging data are used to evaluate for potentially treatable forms of dementias caused by tumor, vascular malformation, cerebral hematoma, normal pressure hydrocephalus, infections, endocrine abnormalities, or vitamin deficiencies. Structural imaging studies (i.e., head CT or brain MRI) should document the existence of one or, more likely, several cerebrovascular lesions. Cerebral angiography may identify vascular lesions (e.g., aneurysms or cerebral artery stenosis) that may be reversible through interventional radiology or neurosurgery procedures. Additional tests that may be helpful include CSF analysis, an electroencephalogram (EEG), and/or an EEG with evoked responses. In diagnostically challenging patients, CSF analysis may be helpful to rule out less common causes of dementia (e.g., paraneoplastic syndromes), and the EEG can help discriminate delirium from dementia.

Poststroke Depression

Currently, no laboratory or neuroimaging studies are routinely used to inform the clinical diagnosis of PSD. This topic, nevertheless, is an area of considerable interest for the scientific community. Studies to date have largely focused on endocrine markers linked to major depression. For instance, the dexamethasone suppression test (DST) has been investigated as a possible biological marker. A total of nine studies involving 327 patients demonstrated a statistical association between PSD and a failure to suppress serum cortisol in response to the administration of dexamethasone. The mean pooled data specificity of the test was 87 percent

while the sensitivity was 47 percent. These values are insufficient to be diagnostically useful. The inadequacy of the DST as a diagnostic tool was also seen in a study of 65 patients whose acute strokes had occurred within the preceding year. Of these cases, 67 percent with major depression failed to suppress serum cortisol compared to 25 percent of patients with minor depression and 32 percent of nondepressed patients. The sensitivity of the DST for major depression was 67 percent with a positive predictive value of 50 percent. The specificity was only 70 percent with a negative predictive value of 80 percent. False-positive tests, found in 33 percent of patients, were related to large lesion volumes.

Pathological Laughing and Crying Scale

Part 1: Patient Interview

1. Have you recently experienced sudden episodes of laughter?
2. Have you recently experienced sudden episodes of crying?

If you have experienced sudden episodes of laughter, please answer the following (questions 3-10), otherwise skip to question 11.

3. Have these episodes occurred without any cause in your surroundings?
4. Have these episodes lasted for a long period of time?
5. Have these episodes been uncontrollable by you?
6. Have these episodes occurred as a result of feelings of happiness?
7. Have these episodes occurred in excess of feelings of happiness?
8. Have these episodes of laughter occurred with feelings of sadness?
9. Have these episodes occurred with any emotions other than happiness, or sadness, such as nervousness, anger, fear , etc.?
10. Have these episodes caused you any distress or social embarrassment?

If you have experienced sudden episodes of crying, please answer the following (questions 11-18).

11. Have these episodes occurred without any cause in your surroundings?
12. Have these episodes lasted for a long period of time?
13. Have these episodes been uncontrollable by you?
14. Have these episodes occurred as a result of feelings of sadness?
15. Have these episodes occurred in excess of feelings of sadness?
16. Have these episodes of crying occurred with feelings of happiness?
17. Have these episodes occurred with any emotions other than happiness, or sadness, such as nervousness, anger, fear , etc.?
18. Have these episodes caused you any distress or social embarrassment?

FIGURE 2.2–3. Pathological Laughing and Crying Scale. (Reprinted with permission from Robinson RG. *The Clinical Neuropsychiatry of Stroke*. Cambridge, UK: Cambridge University Press; 2006: 155.)

In a different line of work, a study of growth-hormone response to desipramine found that growth-hormone responses were significantly blunted in patients with PSD. These results suggest that diminished α_2-adrenergic receptor function may be an important marker for PSD. The sensitivity of the test was 100 percent, and the specificity was 75 percent. Future studies of endocrine responses as markers of PSD may prove to have diagnostic validity. Alternatively, such markers may help define more homogeneous subtypes of PSD for research purposes.

Other Disorders

The utility of laboratory examinations in the diagnosis or prognosis of poststroke mania, anxiety disorders, psychosis, apathy, catastrophic reactions, and anosognosia have not been established except as discussed under etiology.

COURSE AND PROGNOSIS

Mild and Major Vascular Neurocognitive Disorders

The course of major vascular neurocognitive disorder is characterized by a close temporal relationship between stroke occurrence and deterioration of cognitive function. The probability of recurrent stroke is approximately 7 percent per year. The course and prognosis of poststroke cognitive impairment, however, can be influenced by preventative measures.

A recent 2014 report from the Framingham Heart Study was the first to prospectively examine the effect of new onset stroke on cognition. In this study, 132 new-onset, dementia-free stroke patients were prospectively followed and received serial cognitive testing. At 6 months after stroke, significantly greater impairment in the following cognitive domains was seen: immediate recall of logical and visual memory, verbal learning, naming, executive function, visuospatial, and motor skills. These deficits were independent of prestroke cognitive function and vascular risk factors.

Hypertension has been shown to be the most common modifiable risk factor for stroke worldwide, and it may potentially have a mediating effect on the strong association between vascular-related cognitive impairment and Alzheimer disease. A longitudinal study of 173 patients with vascular dementia examined the frequency of risk factors for stroke and cerebral atherosclerosis. Although hypertension was the single most potent risk factor for cerebral atherosclerosis and stroke, hypotension, present in 66 percent of cases, was, by far, the most common risk factor for vascular dementia in this sample. Heart disease of the atherosclerotic type, with or without cardiac arrhythmia, was also present in the majority of cases of vascular dementia. Cardiac disease may provide a source of cerebral emboli leading to vascular dementia. Cigarette smoking of one or more packs per day was a risk factor in 21 percent. Hyperlipidemia of the type 4 form, that is, hypertriglyceridemia, was present in 29 percent of cases. Diabetes mellitus of sufficient clinical severity to require medical treatment was found in 20 percent of the cases, while symptomatic peripheral vascular disease with ischemic symptoms referable to the lower extremities was present in 6 percent. Vascular dementia was also associated with limited education, suggesting that prevention efforts that are related to education may be effective in slowing or preventing the disease. Alternatively, the association with limited education may suggest some social or neurobiological benefits of learning that may inhibit the development of this disease.

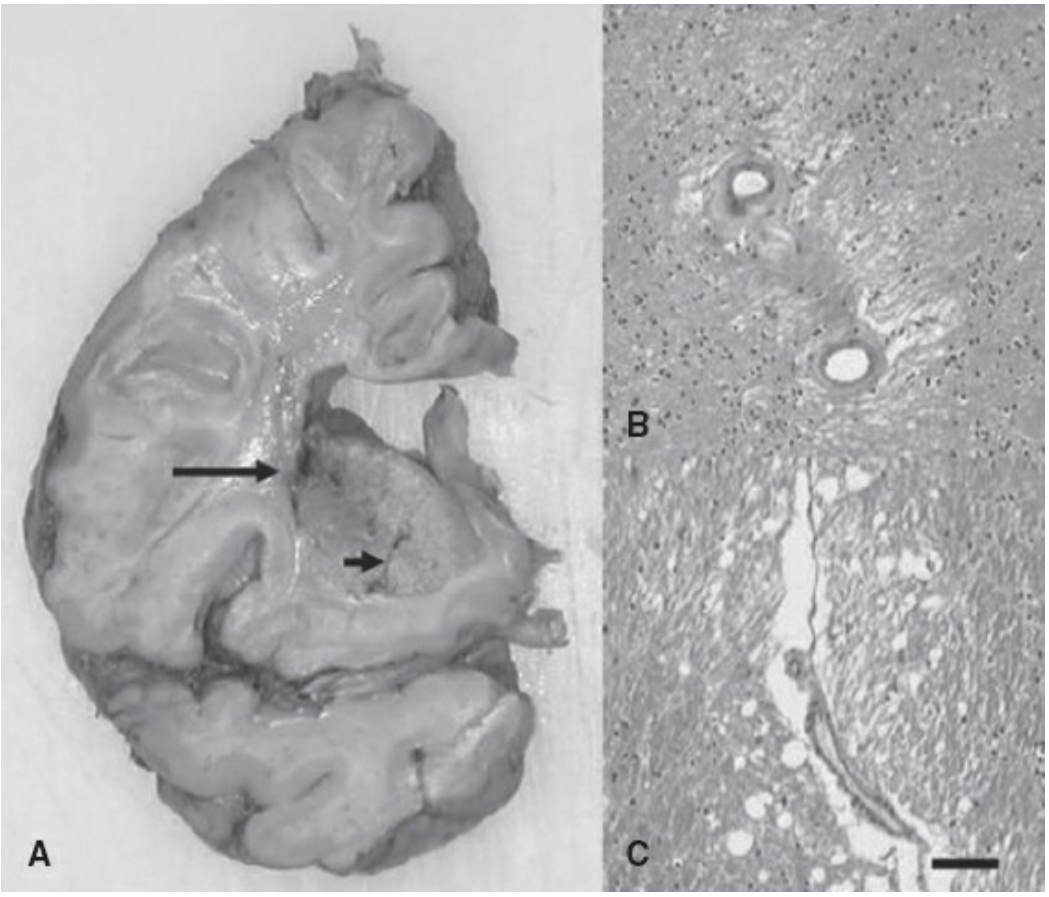

FIGURE 2.2–4. Pathological lesions associated with small vessel disease. **A:** Small infarct (*arrow*) and lacunes (*arrowhead*) in the basal ganglia of a 78-year-old man with cognitive impairment. **B** and **C:** Hyalinsed vessels with perivascular rarefaction and microinfarct in the white matter of a 78-year-old man. Moderate gliosis in the surrounding region is also evident in both cases. Magnification Bar: **A** = 2 cm; **B, C** = 50 µm. (Reprinted with permission from Karalia RN, Erkinjuntii T. Small vessel disease and subcortical vascular dementia. *J Clin Neurol.* 2006;2(1):1–11.)

A 2014 update on subcortical vascular neurocognitive disorder noted that subcortical vascular dementia is driven by severe stenosis of the small vasculature of white matter and small vessel disease (Fig. 2.2–4). In turn, these processes lead to white matter ischemia, slow progression of symptoms, and cholinergic neuronal damage that makes it difficult to differentiate cognitive impairment due to subcortical vascular disease from dAD.

Depression

The longitudinal course of PSD has been examined in a number of studies. At the time of the initial acute or rehabilitation hospital evaluation (e.g., 1 week poststroke), 21.6 percent patients had a symptom cluster consistent with major depression, and 20 percent had a symptom cluster of minor depression (i.e., subsyndromal depression). Although both major and minor depressive disorders appear to continue for months, the observed course has varied from one study to the next. Philip Morris and colleagues

calculated a mean duration of major depression of 39.0 ± 31.8 weeks and a mean duration of minor depression of 12.2 ± 18.2 weeks. Another report by Robinson and colleagues of 142 patients over 2 years following acute stroke found a mean duration of 31.2 weeks for major depression but 11.9 months for minor depression. A mean of 26 percent patients continued to have major depression at 1 year after initial diagnosis. A recent study of 1,233 stroke patients followed over 15 years found that 50 percent of patients with PSD at 3 months had recovered by 1 year. A 2013 meta-analysis of 4,022 patients from 50 studies found, among patients depressed a few months following stroke, recovery at 1 year ranged from 15 to 57 percent. In sum, these findings indicate that a significant number of patients who develop either major or minor depression following stroke may experience persistent depression for 2 or more years after stroke.

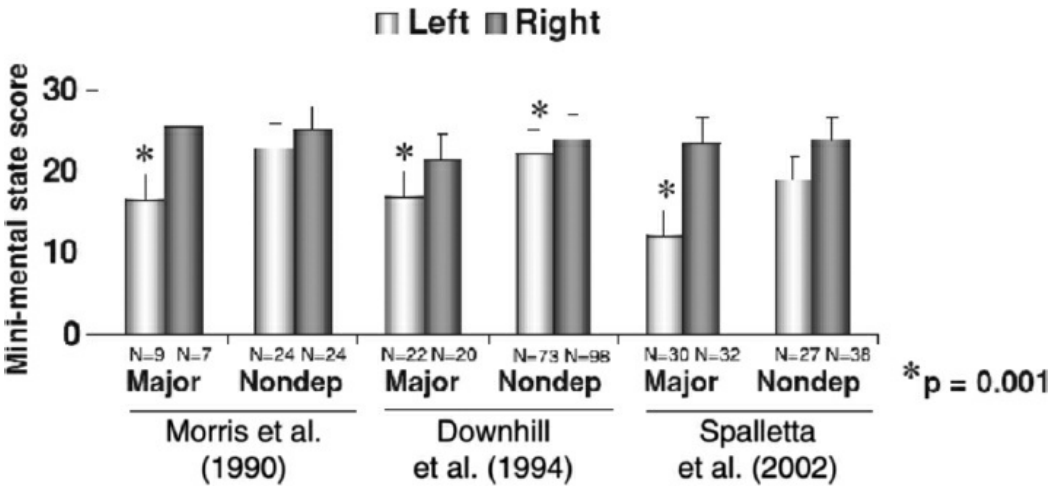

FIGURE 2.2–5. Cognitive impairment related to depression and lesion site. Mini-mental state examination scores following acute stroke in three studies among patients with major or no mood disturbance grouped according to the hemisphere of ischemia. In all three studies, there was a significant difference between patients with major depression following left-hemispheric stroke and nondepressed patients with similar lesions (p = 0.001). Major depression following right-hemispheric lesions did not lead to the same phenomenon. (Reprinted with permission from Robinson RG. *The Clinical Neuropsychiatry of Stroke*. Cambridge, UK: Cambridge University Press; 2006:155.)

Numerous studies have examined the relationship between depression and physical/functional recovery from stroke as measured by ADLs. Virtually all studies have found that the most impaired patients in ADLs have the most severe depressions. Furthermore, a subset of these studies has demonstrated that depression severity after acute stroke is an independent predictor of severity of ADLs impairment at 1 or more years after stroke. Thus, the prognosis for recovery in ADLs is significantly worse if a patient has depression following their acute stroke. Interestingly, Cevdet Bilge and colleagues found that patients with PSD who responded to treatment with citalopram showed significantly better improvement in ADLs than patients who were never depressed.

Similarly, the prognosis for cognitive functioning is also dependent, in part, upon the existence of PSD. Three separate studies have shown that major depression following acute stroke is associated with more severe cognitive impairment if the stroke occurred in the left hemisphere (Fig. 2.2–5). This effect of right versus left hemispheric stroke is not seen among nondepressed patients with similar lesions.

Longitudinal studies have also shown that patients with major depression following left-hemispheric stroke have greater severity of cognitive impairment in the 12 months following stroke in comparison to patients with comparable stroke but no major depression. Between 12 and 24 months following stroke, however, cognitive function among these patients tends to improve. By 24 months follow-up, no difference exists in the severity of cognitive impairment among patients regardless of lesion location (i.e., right- vs. left-hemispheric strokes) or depression status (i.e., depressed or nondepressed patients) following the acute stroke.

Importantly, PSD has a dramatic association with increased mortality. The first study to report this phenomenon using standardized interviews and diagnosis was performed in 1993 on a sample of 103 acute stroke patients. At 10 years poststroke, patients with major or minor depression had a 3.4-fold increase in the odds of mortality in comparison to those with no depression. A more recent study of 976 patients with stroke found that patients with depression, assessed at 3 weeks poststroke using the Wakefield Self-Assessment Depression Inventory, had 50 percent higher mortality at 1 year compared to nondepressed patients. A 2013 meta-analysis of 50 studies found that patients with depression within 3 months of stroke had poor clinical outcomes defined by lower quality of life, increased disability, and increased mortality at follow-up compared with nondepressed stroke patients.

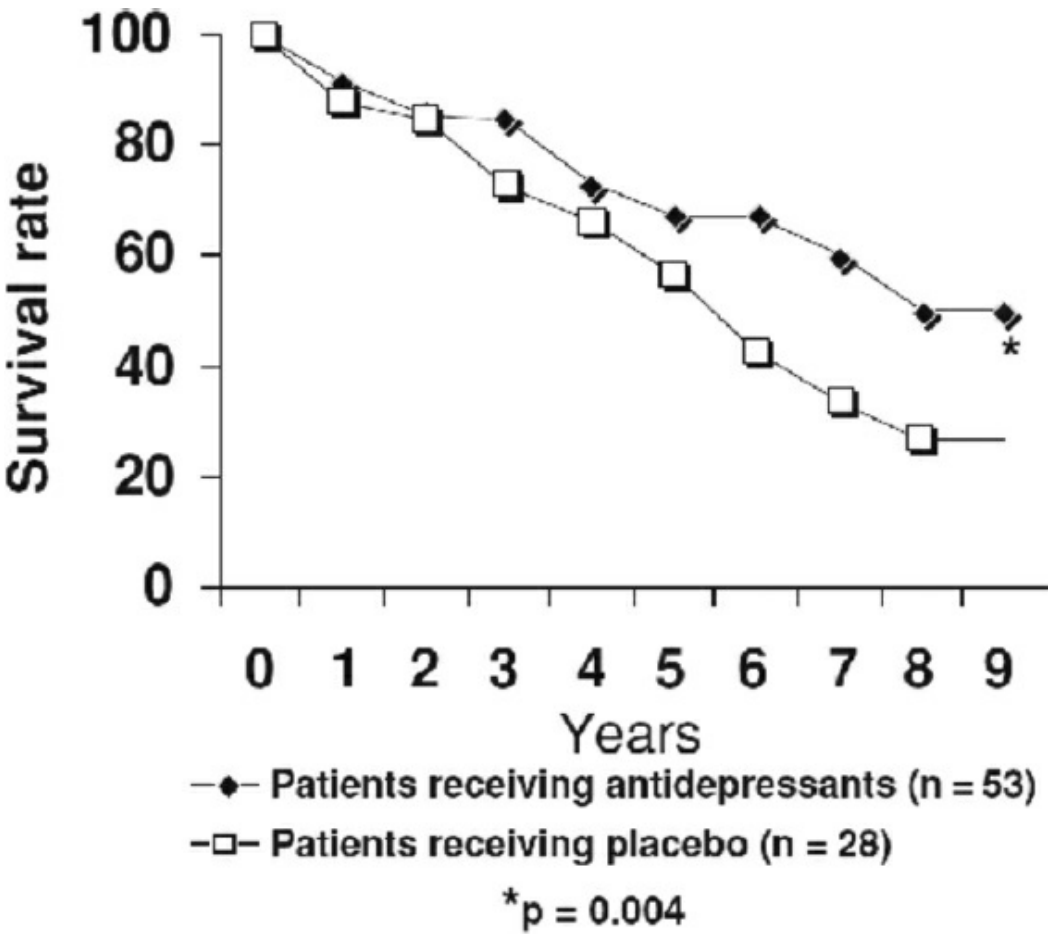

FIGURE 2.2–6. Poststroke mortality after 3 months of treatment. Survival rates over 9-year follow-up for stroke patients who received a 12-week course of antidepressants or placebo during the first 6 months following stroke. Probability of survival was significantly greater in the patients receiving antidepressants ($\chi^2 = 8.2$, df = 1, $p = 0.004$, Kaplan–Meier survival analysis, log-rank test). (Reprinted with permission from Robinson RG. *The Clinical Neuropsychiatry of Stroke*. Cambridge, UK: Cambridge University Press; 2006:155.)

Perhaps the most provocative finding, however, is the relationship of mortality following PSD to treatment with antidepressant therapy. A 9-year follow-up of patients who had been treated for PSD found that active treatment with nortriptyline or fluoxetine ($n = 53$) versus placebo ($n = 28$) over 12 weeks resulted in increased probability of survival (i.e., 59.2 percent for treated vs. 36.4 percent for placebo patients) at 6 years follow-up (Fig. 2.2–6). A logistic regression that examined the effects of age, diabetes, relapsing depression, and antidepressant use found that antidepressant use and the existence of diabetes mellitus independently and significantly predicted survival.

A series of recent findings, which have been important additions to the literature, found that the use of fluoxetine (20 mg/day) or nortriptyline (100 mg/day) were associated with improved recovery from stroke at 3 or 12 months, respectively, after 3 months of treatment. Interestingly, this effect was independent of PSD.

The course of poststroke mood disorders is exemplified by the following

case history:

Ms. A. was a 35-year-old woman who had been the regional director of marketing for a national company. Deadlines, frequent travel, and sales quotas were all part of her high-pressure work. She developed hypertension during her first pregnancy, but in spite of this, she kept up her hectic work schedule. While on a business trip during her subsequent pregnancy, however, she suffered a stroke that caused mild weakness of her right side as well as aphasia characterized by difficulty producing speech but not comprehension (i.e., a nonfluent, expressive aphasia). These motor and language impairments were relatively mild and cleared up within several months of the stroke.

When she was approximately 6 months poststroke, she was convinced that there was still something wrong with her as a result of the stroke. Several physicians told her that there was nothing physically wrong with her, and all she needed to do was to get back to work. She had never experienced prolonged depressive symptoms prior to the stroke, and her affect did not appear depressed at the time of the assessment. She was talkative, and her thoughts and speech were not slowed, as often occurs in depression. She was not tearful or suicidal. She did, however, feel depressed and had loss of interest, concentration, and motivation. She had returned to work for a couple of hours a day but was unable to concentrate well enough to accomplish even simple tasks. She had lost interest and pleasure in virtually all of her work and social activities. She no longer had the ambition to climb the corporate hierarchy. She also had sleep disturbance with early morning awakening, loss of appetite and weight, decreased sexual interest, and decreased energy. A diagnosis of PSD was made, and antidepressant therapy with nortriptyline was started. Her response to antidepressant treatment was dramatic. Between 4 and 6 weeks after beginning nortriptyline, her mood greatly improved, and she returned to work. She found that she was able to concentrate and experience interest and pleasure in her work. Over a period of 2 to 3 months, she reverted from somebody who was virtually immobilized by depression vocationally and socially to an effective, energetic woman. She also had a return of some of her previous ambition, though she still lacked her prior drive to reach the top of the corporate hierarchy.

After 9 months of taking nortriptyline, she wanted to discontinue this medication. Her rationale was that she felt she had fully recovered and did not want to continue taking medication that produced a dry mouth and constipation. The medication was tapered over a period of about 6 weeks and then stopped. She remained well for approximately 1 year but then had a recurrence of the same symptoms that were initially observed. Namely, she was uninterested in work, had no feeling of pleasure in any of her usual activities, and was unable to attend to the demands of work or home. She also had difficulty sleeping, lost her appetite, and felt depressed. These symptoms subsided after restarting her antidepressant medication, which she continued to take for another year. After that year, she again insisted on stopping her antidepressant medication. Over the next 2 years of follow-up, she remained free of depressive symptoms, but she was counseled that her previous history placed her at increased risk for another recurrence of depression. She was advised to remain vigilant for any signs or symptoms of her depression and to report them immediately so resumption of treatment could be discussed in a timely fashion.

Mania

The course of mania following stroke has not been systematically examined. Anecdotal cases have been reported indicating that recurrent episodes of mania or depression may occur in these patients (Fig. 2.2–7).

Anxiety

The prevalence of GAD and major depression based on standardized interviews and diagnostic criteria was documented in two separate studies

(Fig. 2.2–8). Note that the rates of GAD with and without comorbid depression are stable at about 20 percent over 3 years poststroke. A recent 5-year longitudinal study of 220 patients found that at 5 years poststroke, the prevalence of anxiety was 29 percent, and those with anxiety at 6 months were significantly more likely than nonanxious patients to have anxiety at 5 years. Another recent 10-year study reported that anxiety ranged from 17 to 24 percent with a cumulative incidence of 57 percent over 5 years. Further, comorbid anxiety and depression at each year ranged from 57 to 73 percent.

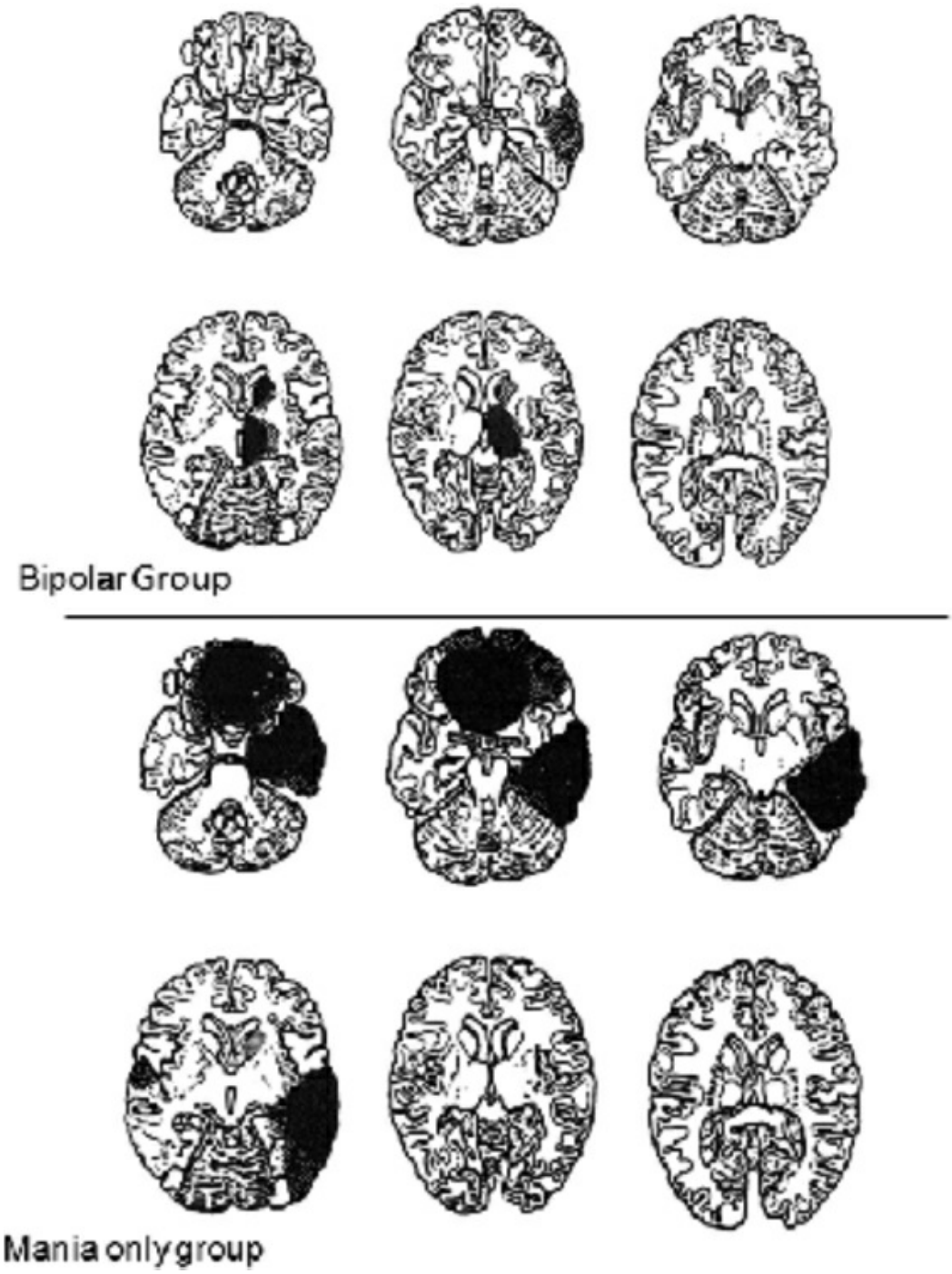

FIGURE 2.2–7. Brain imaging of stroke lesions for patients with bipolar disorder or mania only. Bipolar Group: schematic templates of computed tomographic scan slices showing largest cross-sectional area of lesion for manic-depressive patients. Lesions mainly involved the right head of the caudate and right thalamus. Mania only group: schematic templates of computed tomography scan slices showing largest cross-sectional area of lesion for mania only patients. Lesions mainly involved the orbitofrontal cortex and the right basitemporal cortex. (Reprinted with permission from Robinson RG. *The Clinical Neuropsychiatry of Stroke*. Cambridge, UK: Cambridge University Press; 2006:155.)

Another study of 142 patients over 2 years found that early onset (i.e., within 3 months after stroke) anxiety disorder without depression had a mean duration of 1.5 months while delayed onset (i.e., more than 6 months after stroke) GAD without depression had a mean duration of 3 months. In addition, the existence of anxiety also influenced the duration of depression. Patients with GAD and major depression had a mean duration of depression that was significantly longer than the duration of depression without anxiety disorder.

The recent longitudinal study of poststroke anxiety by Luis Ayerbe and colleagues in 2014 reported that anxiety at 3 months poststroke was associated with lower quality of life at 10 years follow-up. In 2014, Maya Lambiase and coauthors found among 6,019 patients with stroke that higher scores of baseline anxiety were associated with increased risk of incident stroke within the 16-year follow-up period (Fig. 2.2–9).

Other Disorders

The course and prognosis of patients with psychosis, apathy, catastrophic reaction, PLAC, and anosognosia have not been systematically studied.

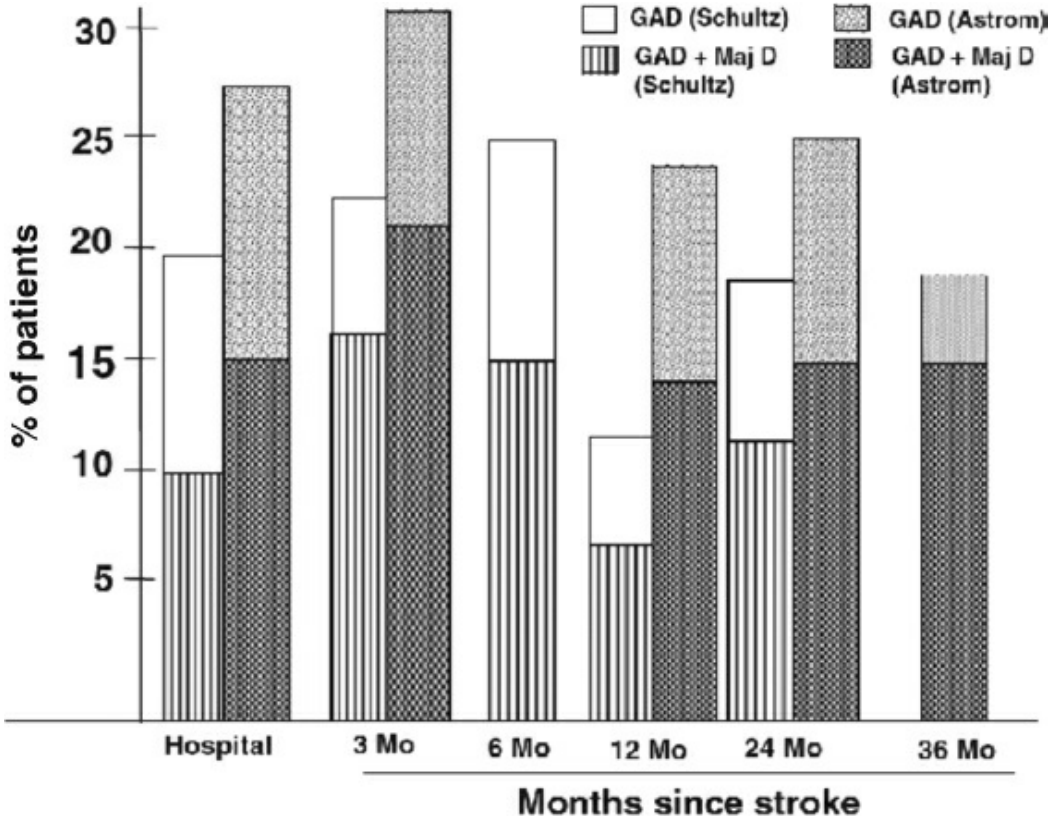

FIGURE 2.2–8. Longitudinal association of major depression with anxiety disorder after stroke. The frequency of generalized anxiety disorder (GAD) with and without major depression over the 3 years following acute stroke. Results obtained from Schultz et al. (1997) using the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) criteria were compared to the results of Astrom (1996) using the third revised edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III-R) criteria. Results indicate a slightly lower frequency of GAD using the DSM-IV criteria and emphasize the prominence of major depression in this population of patients with poststroke anxiety disorder. (Reprinted with permission from Robinson RG. *The Clinical Neuropsychiatry of Stroke*. Cambridge, UK: Cambridge University Press; 2006:155.)

TREATMENT

Mild and Major Vascular Neurocognitive Disorders

Some of the risk factors for stroke can be effectively treated, thus giving hope that cognitive impairment due to vascular disease may someday be a preventable condition. In the past decade, treatment with antiplatelet aggregate drugs has reduced the number of repeated ischemic vascular episodes in patients with TIAs. Acetylsalicylic acid (ASA) and other antiplatelet drugs have been shown to be effective in the secondary prevention of stroke. The United Kingdom-TIA Aspirin Trial, with 2,435 patients using two different dosages of ASA, found that there were 21.7 and 25.1 percent reductions, based on dosage, in the risk of nonfatal strokes, myocardial infarction, or death compared with placebo treatment.

Given that hypertension is the biggest risk factor for stroke, its impact on the incidence of vascular dementia has been an area of great interest. A 2011 meta-analysis by Sally Sharp and co-scientists of 768 cases of vascular

dementia and nearly 10,000 control cases found a significant association between hypertension and reduced risk of vascular dementia with an odds ratio of 1.6 and 4.8, respectively, based on longitudinal studies and cross-sectional studies. A separate meta-analysis from 2011 based on 14 studies demonstrated that individuals prescribed antihypertensive medications had a reduced risk of vascular dementia (pooled relative risk of 0.67). To date, it is unclear if this benefit is attributable to a certain class(es) of antihypertensives and/or from the treatment of the hypertension alone.

For patients who are at particular risk for stroke-related cognitive impairment, but are yet to show signs of dementia (i.e., those with a history of TIAs, stroke, or silent cerebral infarctions), carotid endarterectomy, when carotid stenosis ranges from 70 to 99 percent, may be indicated to minimize the risk of future strokes. After cardioembolic stroke, anticoagulation is an effective treatment to reduce the risk of stroke recurrence. Stroke of cardioembolic origin is responsible for about 15 percent of all ischemic strokes, and this percentage is even higher among younger patients.

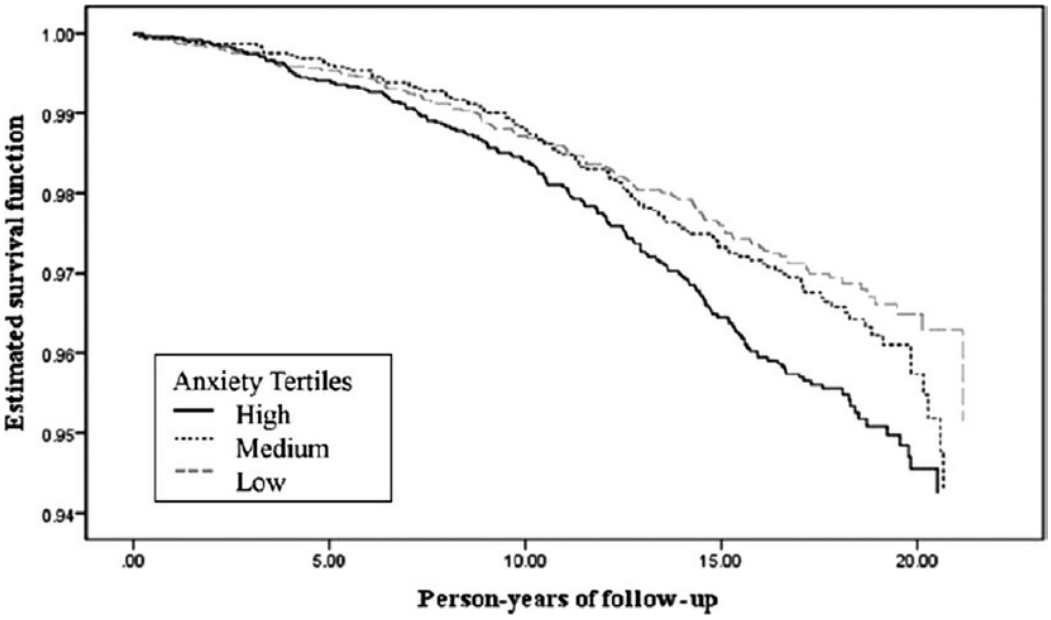

FIGURE 2.2–9. Estimated survival function based on the relationship between anxiety symptoms and incident stroke after adjustment for age and gender. Higher anxiety symptoms levels were associated with higher risk of stroke. (Reprinted with permission from Lambiase MJ, Kubzansky LD, Thurston RC. Prospective study of anxiety and incident stroke. *Stroke*. 2014;45(2):438.)

In patients with vascular dementia, ongoing treatment of vascular risk factors is generally recommended, unless the risks clearly outweigh the benefits. For example, generally antihypertensive treatment is recommended to minimize the risk of further stroke and thus further stroke-related cognitive decline. However, if antihypertensive treatment is leading to orthostatic hypotension, the risk of falls may outweigh the

benefit of such treatment. Therapeutic measures aimed at reducing vascular risk factors include antihypertensives (e.g., beta-blockers or calcium channel blockers), lipid lowering agents such as statins, smoking cessation, and prevention or careful management of diabetes mellitus. Finally, treatment may also include antidepressant agents, cholinergic agonists (e.g., donepezil), antiplatelet aggregation agents, statins, and neurotrophic factors.

Depression

Reported in 1984, the first randomized, blinded, placebo-controlled study in PSD found that 14 patients treated with nortriptyline (50 to 100 mg/day) had significantly greater reduction in Hamilton Depression Scores over 6 weeks than 20 patients given placebo. The first double-blind, placebo-controlled trial to examine the efficacy of selective serotonin inhibitors (SSRIs) was reported by Grethe Anderson and colleagues in 1993. Among 33 poststroke patients given citalopram (20 mg/day under age 65; 10 mg/day over age 65), a significantly greater reduction in Hamilton Depression Scores over 6 weeks was observed in comparison to 33 similar patients given placebo. Consisting of 16 trials and 1,655 subjects, a Cochrane Review from 2008 by Hackett and colleagues found an odds ratio 0.47, favoring pharmacological treatment over control for the treatment of PSD. Thus, the efficacy of antidepressant treatment for PSD, particularly severe depression, has been established with the exception that three studies failed to find fluoxetine any better than placebo for the treatment of PSD.

Perhaps the most promising recent advance in the treatment of PSD has been the study of preventative treatments (Fig. 2.2–10). To date, at least four randomized studies have been designed to evaluate the prevention of PSD using SSRI antidepressants. Results of these studies are shown in Figure 2.2–8. Although three of the four studies failed to show a significant effect on prevention, the similarities in rate of prevention suggest that lack of power may have been the cause of the lack of an effect. The fourth study, conducted by Robinson and colleagues, involved 58 nondepressed acute stroke patients who were treated with escitalopram (5 mg/day over age 65; 10 mg/day under age 65) over 1 year. The incidence of PSD was 8.5 percent compared with 11.9 percent among 57 patients receiving problem solving therapy (PST) and 22.4 percent among 58 patients receiving placebo (escitalopram vs. placebo, $p \leq 0.001$; PST vs. placebo, $p \leq 0.001$).

Mania

Data on individual patients with single or recurrent episodes of mania suggest that they respond to lithium, although some fail to respond to either lithium or carbamazepine. No randomized controlled studies of the

treatment of poststroke mania have been conducted, but one longitudinal study of 23 patients showed that most patients' mania remit within a few months.

Anxiety Disorders

Data from three randomized, double-blind treatment studies were merged to evaluate nortriptyline (n = 13) versus placebo (n = 14) in the treatment of patients with comorbid GAD and depression following stroke. Severity of anxiety was measured using the Hamilton Rating Scale for Anxiety, and severity of depression was measured using the Hamilton Rating Scale for Depression. Although there were no significant differences between the nortriptyline and placebo groups in demographic characteristics, stroke type, and neurological findings, patients receiving nortriptyline treatment showed significantly more rapid improvement on the Hamilton anxiety scores compared with the Hamilton depression scores, suggesting that anxiety may respond more quickly to treatment than depression. Furthermore, the anxiety symptoms showed greater improvement in response to nortriptyline treatment compared with placebo. Examining 175 patients in two trials of paroxetine or buspirone, a 2011 Cochrane Review of treatment of poststroke anxiety without placebo control concluded that there is insufficient evidence to guide the treatment of anxiety after stroke.

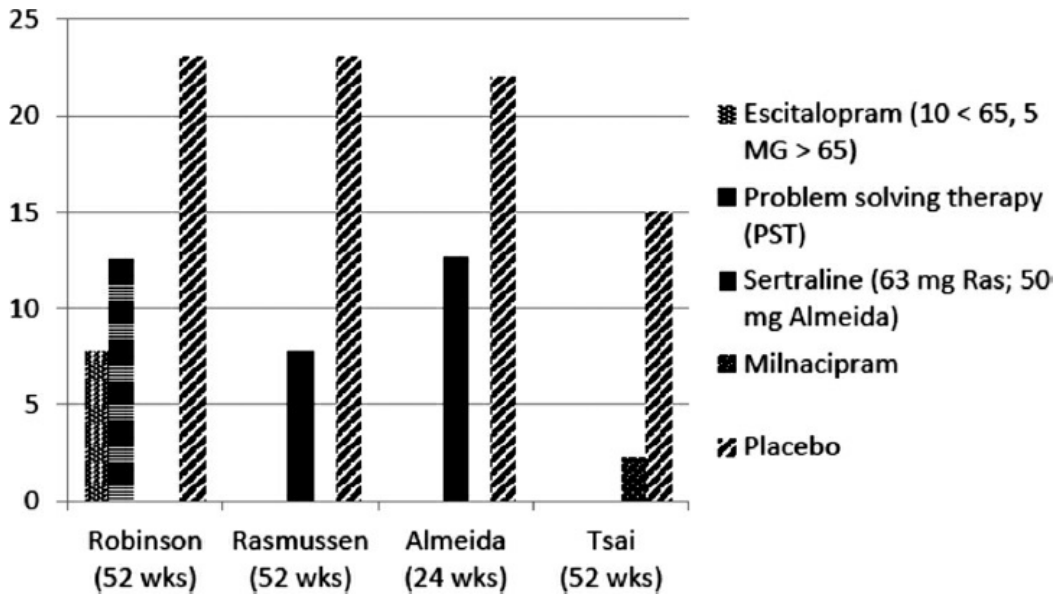

FIGURE 2.2–10. A summary of four randomized trials for the prevention of poststroke depression. For each study, the percent of patients who developed depression during the duration of the prevention trial is shown. The Robinson and Rasmussen studies showed remarkably similar findings. The difference between antidepressant and placebo were statistically significant for the Robinson and the Tsai studies. (The data was extracted from Robinson RG, Jorge RE, Moser DJ, et al. Escitalopram and problem-solving therapy for prevention of poststroke depression: a randomized controlled trial. *JAMA: The Journal of the American Medical Association.* 2008;299(20):2391; Rasmussen A, Lunde M, Poulsen DL, Sorensen K, Qvitzau S, Bech P. A double-blind, placebo-controlled study of sertraline in the prevention of depression in stroke patients. *Psychosomatics.* 2003;44(3):216; Almeida OP,

Waterreus A, Hankey GJ. Preventing depression after stroke: Results from a randomized placebo-controlled trial. *J Clin Psychiatry.* 2006;67(7):1104; and Tsai CS, Wu CL, Chou SY, et al. Prevention of poststroke depression with milnacipran in patients with acute ischemic stroke: a double-blind randomized placebo-controlled trial. *International Clinical Psychopharmacology.* 2011;26(5):263.)

Katsunaka Mikami and colleagues performed a secondary analysis of patients in the previously described prevention trial of PSD conducted by Robinson's group. Of 52 patients without poststroke GAD treated for 1 year with escitalopram, only three cases developed GAD compared with 3 of 53 treated with PST and 9 of 49 patients given placebo ($p = 0.0005$).

Psychosis

There are no controlled treatment trials among patients with delusions or hallucinations following stroke. Anecdotal reports have suggested two basic approaches to treatment: utilization of anticonvulsant therapy or antipsychotic medication. The use of anticonvulsants has its rationale in the frequent coexistence of seizures with psychotic disorders following stroke.

Apathy

Apathy following stroke has been treated with nortriptyline, bromocriptine, methylphenidate, amantadine, selegiline, and tacrine with some success. In a double-blind trial of nefiracetam (900 mg/day) for the treatment of comorbid apathy and depression by Robinson and colleagues, nefiracetam was significantly better in reducing scores on the Apathy Rating Scale compared with placebo. Of note, nefiracetam has not been approved by the U.S. Food and Drug Administration (FDA) for any indications. Since apathy is a relatively common consequence of stroke and can have devastating effects on poststroke recovery, treatment trials are urgently needed.

Mikami and colleagues performed a secondary analysis of the previously mentioned depression trial to evaluate the prevention of apathy after stroke. Among 51 nonapathetic patients treated for a year with escitalopram (10 mg/day for those less than age 65; 5 mg/day for those greater than age 65), 2 developed apathy compared with 3 of 56 treated with PST and 7 of 47 given placebo ($p = 0.0031$). Thus, treatment trials with a primary outcome of apathy prevention are needed to determine if the administration of antidepressant medication such as escitalopram may prove helpful for the prevention of poststroke apathy.

Pathological Laughter and Crying

A Cochran Review of treatment trials of pathological emotions in 2010 included seven trials totaling 739 people who were 6 days to 13 years after

stroke. Antidepressant treatment, including both tricyclic and SSRI/selective serotonin–norepinephrine reuptake inhibitors (SNRIs), significantly reduce the frequency and severity of PLAC. For example, a double-blind, placebo controlled cross-over study of 16 patients reported a greater than 50 percent reduction in the number of crying episodes while receiving citalopram. Eight patients responded to treatment within 24 hours of taking citalopram (20 mg/day); three patients responded within 3 days; and four patients took more than a week to respond. Further, an FDA-approved combination drug consisting of dextromethorphan and quinidine (Nuedexta) is now available with a specific indication for the treatment of PLAC. Thus, treatments that have shown to be effective in treating PLAC are available for clinical use.

The clinical manifestations of PLAC can be appreciated in this case history:

A 64-year-old right-handed, retired, married woman with no prior history of stroke suffered a thrombotic right-hemispheric stroke with hemiparesis but no sensory deficit. Beginning within a few days of the stroke, the patient had uncontrollable crying episodes that occurred 5 to 10 times per day and each lasted about 1 to 2 minutes. In addition to the crying episodes, the patient had major depression with a Hamilton Depression Scale score of 19. She stated that she felt sad but the crying was greatly in excess of her sadness at the time of her episodes, and she had no sense of being able to control the crying. The pathological emotions were more troublesome to her than the depression. Although she and her husband previously had an active social life, she stopped seeing friends and leaving the house for fear of being embarrassed socially by these crying episodes. Her PLACS score was 24, which was severe. She improved greatly after a course of nortriptyline.

Other Disorders

Effective treatments have not been established for catastrophic reactions or anosognosia.

References

Andersen G, Vestergaard K, Lauritzen L. Effective treatment of poststroke depression with the selective serotonin reuptake inhibitor citalopram. *Stroke*. 1994;25:1099–1104.

Ayerbe L, Ayis S, Crichton SL, Rudd AG, Wolfe CD. Explanatory factors for the increased mortality of stroke patients with depression. *Neurology*. 2014;83(22):2007–2012.

Ayerbe L, Ayis S, Wolfe CD, Rudd AG. Natural history, predictors and outcomes of depression after stroke: systematic review and meta-analysis. *Br J Psychiatry*. 2013;202(1):14–21.

Blige C, Koçer E, Koçer A, Türk Börü U. Depression and functional outcome after stroke: the effect of antidepressant therapy on functional recovery. *Eur J Phys Rehabil Medicine*. 2008;44(1):13–18.

Campbell Burton CA, Holmes J, Murray J, et al. Interventions for treating anxiety after stroke. *Cochrane Database Syst Rev*. 2011;CD008860.

Denny-Brown D, Meyer JS, Horenstein S. The significance of perceptual rivalry resulting from

parietal lesions. *Brain*. 1952;75:433–471.

Goldstein K. *The Organism: A Holistic Approach to Biology Derived from Pathological Data in Man*. New York: American Books; 1939.

Hackett ML, Anderson CS. Predictors of depression after stroke: a systematic review of observational studies. *Stroke*. 2005;36(10):2296–2301.

Hackett ML, Kohler S, O'Brien JT, Mead GE. Neuropsychiatric outcomes of stroke. *The Lancet Neurology*. 2014;13(5):525–534.

*House A, Knapp P, Bamford J, Vail A. Mortality at 12 and 24 months after stroke may be associated with depressive symptoms at 1 month. *Stroke*. 2001;32:696–701.

*Jorge RE, Robinson RG, Arndt S, Starkstein S. Mortality and poststroke depression: a placebo-controlled trial of antidepressants. *Am J Psychiatry*. 2003;160:1823–1829.

Kohler S, Hackett ML, O'Brien JT, Mead GE. Neuropsychiatric outcomes after stroke – Authors' reply. *Lancet Neurol*. 2014;13(12):1168–1169.

Kraepelin E. *Manic Depressive Insanity and Paranoia*. Edinburgh, UK: E & S Livingstone; 1921.

Li W, Ling S, Yang Y, Hu Z, Davies H, Fang M. Systematic hypothesis for post-stroke depression caused inflammation and neurotransmission and resultant on possible treatments. *Neuro Endocrinol Lett*. 2014;35(2):104–109.

Lincoln NB, Brinkmann N, Cunningham S, et al. Anxiety and depression after stroke: a 5 year follow-up. *Disabil Rehabil*. 2013;35(2):140–145.

Mattioli F, Gialanella B, Stampatori C, Scarpazza C. General intellectual impairment in chronic right hemisphere damaged patients with anosognosia: a group study. *Neuropsychol Rehabil*. 2012;22(4):501–515.

Mead GE, Hsieh CF, Lee R, et al. Selective serotonin reuptake inhibitors (SSRIs) for stroke recovery. The Cochrane Library. *Cochrane Database Syst Rev*. 2012;11:CD009286.

Meyer A. The anatomical facts and clinical varieties of traumatic insanity. *Am J Insanity*. 1904;60:373.

Morris PLP, Robinson RG, Raphael B. Lesion location and depression in hospitalized stroke patients: evidence supporting a specific relationship in the left hemisphere. *Neuropsychiatry Neuropsychol Behav Neurol*. 1992;3:75.

*Narushima K, Kosier JT, Robinson RG. A reappraisal of poststroke depression, intra and inter-hemispheric lesion location using meta-analysis. *J Neuropsychiatry Clin Neurosci*. 2003;15:442–449.

Robert P, Onyike CU, Leentjens AF, et al. Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders. *Eur Psychiatry*. 2009;24(2):98–104.

*Robinson RG. *The Clinical Neuropsychiatry of Stroke*. Cambridge, UK: Cambridge University Press; 2006.

Robinson RG, Jorge RE. Post-Stroke Depression: A Review. *Am J Psychiatry*. 2016;173(3):221–231.

Robinson RG, Jorge RE, Long J. Prevention of Post-Stroke mortality using problem solving therapy or Escitalopram. *Am J Geriat Psychiatry*. 2016; in press.

Rochat L, Van der Linden M, Renaud O, et al. Poor reward sensitivity and apathy after stroke: implication of basal ganglia. *Neurology*. 2013;81(19):1674–1680.

Santos CO, Caeiro L, Ferro JM, Figueira ML. Mania and stroke: a systematic review. *Cerebrovasc*

Dis. 2011;32(1):11–21.

*Spalletta G, Bossu P, Ciaramella A, Bria P, Caltagirone C. The etiology of poststroke depression: a review of the literature and a new hypothesis involving inflammatory cytokines. *Mol Psychiatry*. 2006;11:984–991.

Starkstein SE, Fedoroff JP, Price TR, Leiguarda R, Robinson RG. Catastrophic reaction after cerebrovascular lesions: frequency, correlates, and validation of a scale. *J Neurol Neurosurg Psychiatry*. 1993;5:189.

Tang WK, Chen Y, Lu J, et al. Frontal infarcts and anxiety in stroke. *Stroke*. 2012; 43(5):1426.

Tang WK, Lau CG, Mok V, Ungvari GS, Wong KS. Impact of anxiety on health-related quality of life after stroke: a cross-sectional study. *Arch Phys Med Rehabil*. 2013.94(12):2535.

Wiesmann M, Kiliaan AJ, Claassen, JA. Vascular aspects of cognitive impairment and dementia. *J Cereb Blood Flow Metab*. 2013;33(11):1696.

▲ 2.3 The Neuropsychiatry of Brain Tumors

ALASDAIR G. ROONEY, M.D.

NOSOLOGY AND CONTEXT

It can be challenging to consistently apply the DSM system when diagnosing psychiatric disorders in brain tumor patients. Although some psychiatric disorders pose no major difficulty (e.g., a new-onset psychosis in a patient who has commenced high-dose corticosteroids), for others, a rigidly operationalized approach to diagnosis presents intellectual challenges. This is not a new problem; nor is it surprising: the original development of the DSM system in the 1970s was not conducted primarily with brain tumor patients in mind. Indeed, studies of the nature, course, and covariation of putative psychiatric symptoms *specifically in brain tumor patients* were probably absent from the synthesis of evidence which led to the first set of diagnostic criteria and from which the DSM-5 is ultimately descended. The same thing could of course be said about most other medical diseases. The reason the nosology of mental disorder can cause particular difficulties in brain tumor patients is that in some respects they differ in important ways from other clinical groups. These differences could, in certain cases, call the validity and reliability of DSM syndromes into question. Two arguments will be made briefly here in support of this proposition and by way of contextualizing the rest of this section. The first is with respect to the internal structure of the DSM philosophy itself, and the second a function of how that philosophy is applied.

A central assumption in psychiatry is that symptoms arise, in whole or in part, through some form of dysfunction in the brain. Although nominally atheoretical, the DSM system is rooted in the same assumption: its core symptoms are invariably focused on the internal psychic reality experienced by patients. At this point in history no one can convincingly

prove that this perception of internal reality is centered anywhere but the brain. With this assumption, the DSM philosophy is to standardize diagnoses between patients in the name of diagnostic reliability. Yet, the presence of a tumor invading, distorting, and destroying brain tissue with often multiple toxic brain treatments is unique context. It can be unclear whether the fundamental nature of brain-derived psychiatric symptoms are, compared to people without a brain tumor, altered to such a degree as to affect diagnostic validity and reliability. In other words, a "brain manual" which is designed without a specific focus on validity and reliability in patients with alien abnormalities of brain biology could, at times, be neither valid nor reliable. These patients have—confoundingly—a tumor growing in the seat of the psyche.

The second argument stems from how the DSM-operationalized criteria are generally applied. In order to boost reproducibility and minimize theoretical influence, much of the context of symptomatology is stripped away. The jettisoning in DSM-5 of the "bereavement exclusion" criterion for major depressive disorder is emblematic of this strategy. The natural trend induced is, at least, to minimize contextual influence in certain ways. The problem—as many psychiatrists who spend time with brain tumor patients will affirm—is that it is difficult to think of a patient group in whom the contextual palette is starker and potentially matters more. A catastrophic existential threat occurring at any age, without warning, with multiple sudden, persisting, and progressive losses in multiple domains of personhood and life, an onslaught of treatments with side effects identical to key psychiatric symptoms, cognitive impairment, epilepsy, and often eventually death: This is indeed a challenging test for the DSM philosophy of minimizing or abandoning context. Few other patient groups for whom the DSM is intended will face quite the same severe mix of stressors. Does this at all matter? The only way to know the answer to this question is to study it. As things stand, however, brain tumor specific studies of DSM validity and reliability are painfully rare.

The solution is to conduct prospective studies of the reliability and validity of psychiatric diagnoses in large cohorts of brain tumor patients of all ages and stages of disease. In the meantime and in general, the day-to-day psychiatric management of brain tumor patients can be summed up in three words: pragmatism, not dogmatism. There is relatively little randomized controlled trial (RCT) evidence to inform much of what psychiatrists will be asked to decide on for these patients. But this statement of affairs serves to highlight the argument that clinical practice ultimately should be based on evidence. The ambition is surely this: to fully realize the potential of psychiatry in neuro-oncology. To this end, psychiatrists will have to systematically study, comprehensively understand, and rationally and effectively treat the abnormal emotional, behavioral, and neurocognitive consequences of brain tumors.

BRAIN TUMORS

Epidemiology

In the United States and in many other countries, primary brain tumors are reportable diseases. The word "primary" in this context refers to a tumor which arises de novo in the brain and is not the result of metastatic spread from a site out-with the CNS. The Central Brain Tumor Registry of the United States (CBTRUS) is a nonprofit research organization which collaborates with centralized state registries to provide an annual statistical analysis of their data. As such it is an authoritative resource for descriptive epidemiological statistics on new primary brain tumor diagnoses, albeit limited in scope to the United States. CBTRUS data show that primary brain and CNS tumors are the single most common cancer in children and adolescents (age 0 to 19 years) with an average annual incidence of 5.4 per 100,000. In adults, the incidence of primary brain or CNS tumors lags behind that of prostate and breast cancer (the two most common adult cancers in the United States), but at 27.9 per 100,000, it is similar to that for bladder cancer, malignant melanoma, or non-Hodgkin lymphoma. Overall, primary brain tumors are more common in females (58% vs. 42%), but women are more likely to have tumors which are benign: Malignant tumors are slightly more common in men. Particularly given public anxiety over the possible impact of increasing mobile phone use, there is currently no evidence that the incidence of primary brain tumors—benign or malignant—is increasing.

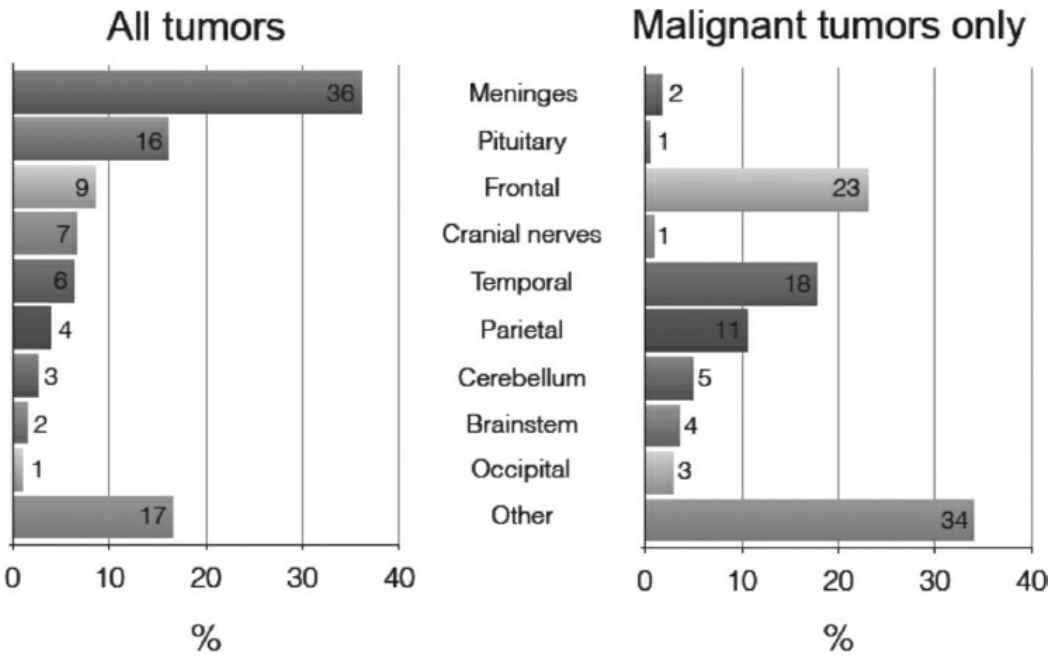

FIGURE 2.3–1. Tumor type and location.

The epidemiology of primary brain tumors is difficult to summarize

concisely. They are a heterogeneous group of diseases with a wide variety of etiologies, ages of onset, anatomical locations, clinical courses, responses to treatment, and prognoses. In very general terms, they can be categorized most simply either as benign or malignant. Most benign tumors arise in the meninges, pituitary gland, or cranial nerve sheaths. Benign tumors are slightly more common than malignant, so psychiatrists in routine general adult practice may be more likely to encounter them incidentally, or very occasionally, as occult causes for psychiatric symptoms. Incidence rates for certain tumors also vary according to ethnicity. Oversimplifying, some benign tumors occur slightly more frequently in people of black or Hispanic origin.

Most malignant primary tumors, meanwhile, are gliomas and are more common in whites. Glioma is cancer arising from neuroepithelial tissue in the brain. It is particularly relevant to psychiatrists working in neuro-oncology because—unlike, for example, meningiomas which are usually anatomically circumscribed and curable by surgery—glioma is generally invasive and incurable. The gliomas are often accompanied by markedly impaired neurological functioning and long-term neurocognitive damage. In adults, most malignant gliomas are supratentorial and situated in the frontal, temporal, parietal, and occasionally occipital lobes (Fig. 2.3–1). In children, by contrast, glioma is usually infratentorial. Treatments in children are therefore often aimed at the posterior fossa and brainstem with considerable functional consequences for survivors.

Table 2.3–1.
The Main Histological Cell of Origin of Primary Brain Tumors Differs at Different Ages

Age	Most Common Histology	Second Most Common Histology	Third Most Common Histology
0–4	Embryonal tumor	Pilocytic astrocytoma	Malignant glioma NOS
5–9	Pilocytic astrocytoma	Malignant glioma NOS	Embryonal tumor
10–14	Pilocytic astrocytoma	Malignant glioma NOS	Neuronal and mixed glial tumor
15–19	Pituitary tumor	Pilocytic astrocytoma	Neuronal and mixed glial tumor
15–39	Pituitary tumor	Meningioma	Nerve sheath tumor
20–34	Pituitary tumor	Meningioma	Nerve sheath tumor
35–44	Meningioma	Pituitary tumor	Nerve sheath tumor
45–54	Meningioma	Pituitary tumor	Glioblastoma
55–64	Meningioma	Glioblastoma	Pituitary tumor
65–74	Meningioma	Glioblastoma	Pituitary tumor
75–84	Meningioma	Glioblastoma	Pituitary tumor
85+	Meningioma	Glioblastoma	Pituitary tumor
Total	*Meningioma*	*Pituitary tumor*	*Glioblastoma*

Adapted from Quinn T, Ostrom QT, Gittleman H, et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2007–2011. *Neuro Oncol*. 2014 Oct;16 Suppl 4:iv1–63.

The gliomas are categorized by the World Health Organization (WHO) according to the predominant tissue histological profile (astrocytic, oligodendrocytic, etc.) and by their malignancy. Malignancy is judged according to histological features known to confer a poorer prognosis. The WHO grades glioma on a level from I to IV where IV is the most malignant type, also known as glioblastoma multiforme (GBM). By convention, WHO grades I and II are called "low-grade" while WHO grades III and IV are "high-grade." In reality, though, nearly all low-grade tumors eventually progress and transform into a higher WHO grade, before killing the patient.

In general, the incidence of most major classes of primary brain tumor (e.g., meningioma, pituitary tumor, and GBM) (Table 2.3–1) rises continuously with advancing age. The main exceptions are pilocytic astrocytoma, germ cell tumors, and embryonal tumors such as medulloblastoma, each of which is more common in younger children and rare in adults. The median age at diagnosis for all primary brain and CNS tumors is 59 years.

In contrast to primary brain tumors, the epidemiology of metastatic brain tumors is less well understood. There is no requirement on National Cancer Registries to document their incidence. There is general awareness that metastatic or "secondary" tumors far outnumber primary tumors. Metastatic brain tumors occur in a sizeable minority of all cancer patients, are usually multiple, and especially more likely to occur when the primary site is lung, breast, skin, kidney, or colon. The fact that their presentation and management is scattered across multiple clinical pathways, and their particularly poor prognosis, combine to make studies of their psychiatric effects difficult. This section will therefore focus only on primary brain tumors and mostly on glioma, since this is the best-studied tumor type to date.

Management

In many countries, the main treatment decisions for someone with a brain tumor are made by a multidisciplinary neuro-oncology team. In the United Kingdom, the minimum core personnel requirement for this team is stated in NHS clinical guidance and includes a team co-ordinator, neurosurgeon, neuroradiologist, neuropathologist, neurologist, oncologist, clinical nurse specialist, palliative care specialist, and neuropsychologist. Psychiatry and neuropsychiatry are listed as core members "if required"; in other words, they are in some sense optional. One consequence of this, at least in the United Kingdom, is that psychiatric involvement is often sporadic and restricted mainly to centers with particularly well-developed links to liaison or neuropsychiatry, or a particular local interest. Yet, the same neuro-oncology teams affirm that more input from psychiatry would be desirable.

A survey of 29 UK neuro-oncology teams showed a desire for greater psychiatric involvement. Data collected by Cliodhna Carroll (PhD student, University of Cambridge, UK) and the author, with supervision from Dr. Howard Ring (Consultant neuropsychiatrist, University of Cambridge) and Dr. Robin Grant (Consultant neurologist, NHS Lothian, UK).

Patients with primary brain tumors are at high risk of psychiatric illness. However psychiatric symptoms are believed to be under-treated and current neuropsychiatric services may be inadequate. A postal questionnaire survey of all UK neuro-oncology multidisciplinary teams (MDTs) was conducted to determine: the reported proportion of new patients referred for psychiatric assessment in the previous year; the person most often responsible for identifying patients requiring specialist assessment; whether neuropsychiatry services were available locally; and whether the team perceived an unmet need for neuropsychiatry.

Of 39 neuro-oncology MDTs, twenty-nine (74%) responded, treating an estimated 3500 new patients/year. A median of 1% (mean 4%, range 0–25%) of these new patients were reported as having been referred to psychiatry in the previous year. Eight MDTs stated that they had made no psychiatric referrals. Most often, significant emotional difficulties were either first identified by the clinical nurse specialist (CNS) or brought to clinical attention by the patient's relatives. Psychiatric referrals were usually made by the CNS. A minority of MDTs had access to neuropsychiatry (12/29; 41%) and 24/29 (82%) perceived an unmet need for neuropsychiatry.

In summary, the neuro-oncology CNS was central to referral procedures in most UK centres. Most neuro-oncology MDTs had no regular access to neuropsychiatry assessment. There was a clear national and clinical demand for more neuropsychiatry input to neuro-oncology services.

The risk of psychiatric patients developing a brain tumor may be up to 10 fold higher than that of the general population. Prevalence estimates from retrospective psychiatric cohorts range from 0.1% to 3% depending on whether autopsy information was included. However large studies in this area are rare and mostly conducted before the widespread availability of CT scanning.

Allowing for some differences in international health care models, the main treatment decisions are shown diagrammatically in Figure 2.3–2. Newly diagnosed patients can be treated with varying combinations of surgery, radiotherapy, chemotherapy, watchful waiting, or supportive (palliative) care. The most common active primary treatment for the most common glioma (GBM) is a combination of maximal surgical debulking, radiotherapy, and chemotherapy. High-dose corticosteroids are often started with the intention of reducing cerebral edema as soon as the radiological diagnosis is suspected. Antiepileptic drugs are also very often prescribed at this time. A large research effort is going on to find better curative treatments and the range of "experimental" additional therapies is wide. Childhood brain tumors are often managed with treatments focused on the posterior fossa or brainstem, again including surgery, radiotherapy, and chemotherapy.

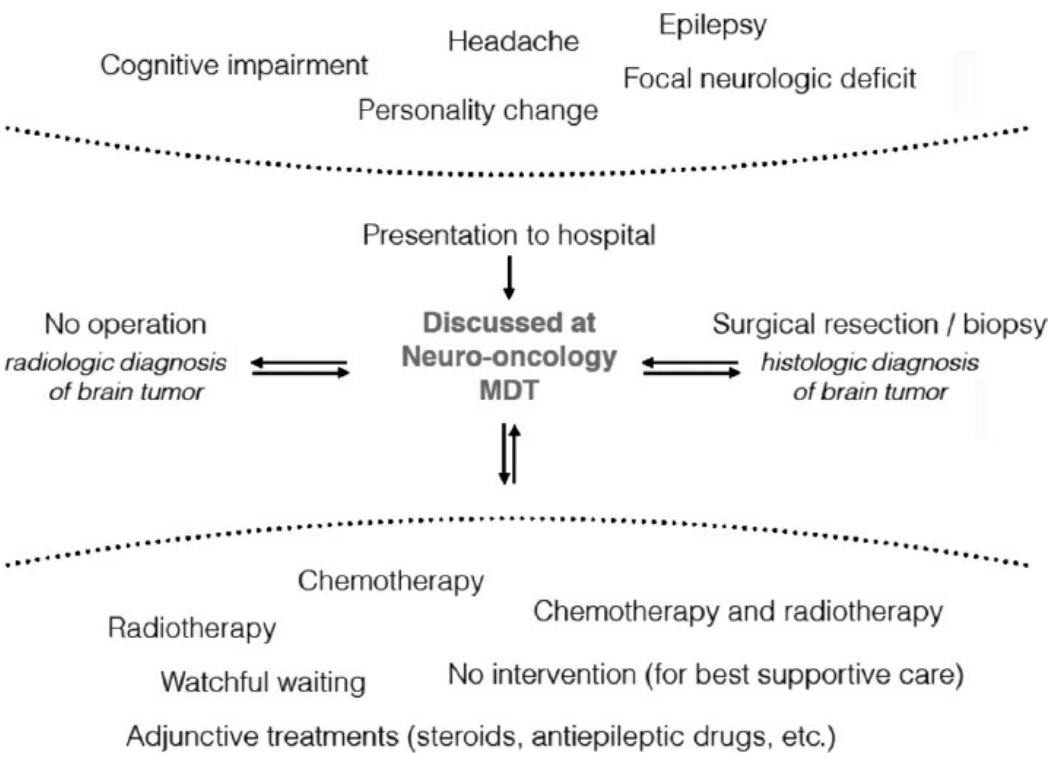

FIGURE 2.3–2. MDT pathway.

Following primary treatment, patients enter follow-up during which they are regularly reviewed by the treating neuro-oncology team. Many patients report that this is an unexpectedly difficult period. The frantic energy of the primary treatment phase is over and they face having to readjust to a life that may now be very different. Physical disability, epilepsy, cognitive impairment, emotional difficulties, and behavioral changes are common. Most patients will be forbidden to drive by national authorities and unable to work, placing further strain on previously stable family routines. Although some patients—especially younger patients with low-grade tumors and well-controlled epilepsy—cope very well and can regain much of their independence, the reality for many is a partially dependent existence and an uncertain future.

Prognosis

The prognosis of a primary brain tumor depends on the type. Meningiomas and pituitary tumors can often be cured surgically, as can some kinds of childhood brain tumor, with relatively little in the way of long-term deficit. Glioma presents an insurmountable challenge to surgical cure because by the time of clinical presentation, tumor cells are dispersed throughout much of the brain (not just in the area seen on an MRI or CT scan). The certain consequence of somehow cutting them all out would be the death of the patient. Low-grade glioma typically follows an indolent course increasing in size only gradually before invariably transforming into a higher grade after a period of years. Higher grade tumors are more rapidly

fatal. The median survival of patients with GBM in the field-leading RCT was 14 months: Around one-quarter of these patients survived to 2 years post diagnosis and only 10 percent were alive after 5 years. Within each of these broad categories, it is becoming increasingly clear that molecular subtypes play an additional role, with some mutations conferring a better and some a poorer prognosis. The likely presence of as-yet-undiscovered molecular prognostic indicators means it is impossible to confidently predict a duration of survival for any given individual patient at the point of diagnosis.

Thus, the tempo of the clinical course of a primary brain tumor is characterized by sudden presentation, a flurry of multidisciplinary management amid a period of intensive assessment, and primary treatment lasting for several months, which is followed by a variable period of follow-up in which patients must attempt to rediscover a normality of life. The profile of disability in the patient is initially moderate and often relieved by primary treatment, before a progressive but unpredictable cumulation of impairments that typically end in the death of the patient after months or years. It is in this context that this chapter will consider the neuropsychiatry of brain tumors.

BRAIN TUMORS IN PSYCHIATRIC PATIENTS

How can psychiatrists tell if their patient has a brain tumor? The idea that psychiatric symptoms are in some way indicative of an underlying brain tumor is reinforced by popular culture. In the movie "Phenomenon," the main character experiences unusual psychic changes which are ultimately attributed to a developing astrocytoma. A character in the United States TV series "Grey's Anatomy" experiences visual hallucinations secondary to brain metastasis. The clearest real-life cultural example, meanwhile, is probably George Gershwin the musical composer and pianist, whose olfactory hallucinations and increasingly erratic behavior were initially attributed to hysteria before he collapsed and died from a brain tumor in 1938 at the age of 37. Similar associations between psychopathology and an occult brain tumor are drawn by hundreds of single case studies in the academic literature. Even collectively, these case studies prove little, but informative and high-quality studies do exist. Large population-based epidemiological datasets suggest that among all occult developing cancers, the risk for psychiatric hospitalization in the period before cancer diagnosis is highest for brain tumors. Thus, there does seem to be a genuine risk that some psychiatric patients will experience symptoms because they actually have a brain tumor. Many such patients (and their doctors) may initially attribute psychiatric symptoms to various environmental or situational stresses, such as economic difficulties, losses of loved ones, or other major life stresses. The problem for the general psychiatrist is that "secret tumor" patients constitute only a vanishing minority: in all likelihood, under 1

percent of their entire case load. How then can psychiatric patients whose symptoms are due to an occult brain tumor be identified? Can investigations be targeted or is indiscriminate screening the only answer?

The importance of targeted investigation is illustrated by autopsy data from chronic psychiatric patients dying in mental hospitals from other causes. These show unsuspected and undiagnosed brain tumors in roughly 3 percent of patients examined. More recently, several radiological studies converge on a similar conclusion: The frequency of clinically unsuspected space-occupying lesions on brain CT or MRI scanning of inpatient psychiatric populations is in the order of 3 to 4 percent. Likewise, the frequency of incidental abnormalities observed in the general nonpsychiatric population is also approximately 3 percent. The conclusion is that many such abnormalities encountered in routine scanning of psychiatric patients will be incidental (i.e., functionally silent). Because at least some of these tumors will never progress to cause any kind of clinical problem, indiscriminate brain scanning may increase the risk of an unnecessary operation.

Head CT for suspected brain tumor in psychiatric inpatients usually shows no evidence of intracranial abnormality. Data collected by Anita Vas-Falcao, (year 4 medical student, University of Edinburgh) with supervision from the author and Dr. Robin Grant. We identified adult psychiatric inpatients referred consecutively over a four-year period for a head CT on suspicion of organic pathology. Patients were excluded if they were diagnosed with brain tumor before the psychiatric admission. The primary aim was to determine the prevalence of brain tumor. A secondary aim was to describe presenting psychiatric symptoms and neurological signs. Consultant radiologist CT reports were obtained from computerised records and reviewed. Psychiatric notes were traced for patients noted to have suspected or possible primary brain tumor on CT report. To simplify analysis, presenting psychiatric symptoms were grouped heuristically into the following broad categories: low mood (anxiety, depression or psychomotor retardation); cognitive dysfunction (confusion or cognitive impairment); personality/behavior change; and psychosis (delusions, mania, or paranoia).

A total of 328 psychiatric inpatients (mean age 65; range 37–86; 43.3% male) were identified as referred for a head CT. Space occupying lesions were identified from the CT report in 19 cases, of which 12/328 (3.7±2.0%) were newly discovered tumors. The most common radiological diagnosis was meningioma (5/12). Eight tumors involved the cerebral hemispheres (5 right-sided and 3 left-sided). The rest were located in the midline or involved the ventricular system.

Most patients had symptoms from more than one psychiatric category. The most common were *low mood* and *cognitive dysfunction,* each seen in 9/12 patients. *Personality/behaviour change* was reported in 6/12 and *psychosis* in 5/12. In four patients psychiatric symptoms were the only presenting features. In five patients meanwhile, no neurological signs were documented in the psychiatric notes. In the remainder, brisk reflexes (five patients) and speech deficits (four patients) were the most common signs. Papilledema was seen in one patient and was associated with a ventricular tumor.

In summary, few (in this study, 1 in 27) head CTs for suspected brain pathology in adult psychiatric inpatients reveal a new case of brain tumor.

"Red flag" Symptoms and Signs

Although newer and better brain imaging techniques have led to earlier diagnosis of many brain tumors, they can still present initially with vague and nonspecific psychiatric and behavioral changes. Psychiatrists must

have a low threshold for considering the possibility of a brain tumor in the differential diagnosis of patients with new-onset psychiatric symptoms. This is especially important if the patient is older but has a negative past personal and family history for psychiatric illnesses, if the psychiatric symptoms have atypical features, if they are associated with additional unexplained personality change, or if there is neurological or neurocognitive dysfunction. In these cases, the psychiatrist should take a careful history from the patient and their close carer(s), looking for symptoms that are commonly associated with brain tumors. Psychiatric patients to be particularly vigilant of include those with a history of recent mental state change in association with *unexplained:*

1. Headache. The classical headache of raised intracranial pressure (ICP) is worse on wakening, presumed owing to relative hypercapnia and cerebral vasodilatation during sleep. It gradually resolves during the morning and may transiently worsen with coughing, straining, or sneezing. Not all brain tumor-associated headaches fit this pattern though, and not all tumors cause headache. Most headaches in psychiatric patients are not caused by brain tumors at all: other explanations may include opiate medication overuse, somatization, tension headache, and selective serotonin reuptake inhibitor (SSRI) discontinuation syndrome.
2. Nausea or vomiting. Presumed secondary to raised ICP compressing vagal centers in the medulla, nausea or vomiting is another classic "red flag" which is nonetheless absent in many brain tumor patients. Projectile vomiting without nausea should raise particular suspicion.
3. Cognitive impairment. As discussed below, cognitive impairment is almost universal in brain tumor patients and ranges in severity from subtle to profound. The more subtle deficits can often be picked up only by detailed neuropsychological testing, or in some respects by a careful history. One of the most frequently reported general cognitive deficits is a reduction in the speed of information processing. Patients may appear slowed up and rather vague. Since impaired concentration is a feature of many psychiatric conditions, cognitively impaired patients, in particular, need a full neurological history and examination.
4. Epilepsy. Most brain tumor patients will have a seizure at some point, but epilepsy is more commonly a presenting symptom for patients with a low-grade glioma. Alternative explanations in psychiatric patients may include alcohol or benzodiazepine withdrawal, lowering of the seizure threshold secondary to antipsychotic use, or pseudoseizures.
5. Weakness or clumsiness. This can be subtle and, if present, weakness is classically pyramidal. If due to a brain tumor it tends to be both subacute (duration of a few weeks rather than many months) and gradually progressive. A high index of suspicion is helpful while listening to the

history. Apraxia may be heralded, for example, by a passing mention of mild difficulties in dressing appropriately, managing the TV remote control, or similar complex movements. An unexplained fall in any psychiatric patient should trigger suspicion.

6. Visual or sensory disturbance. Clues to these sometimes subtle symptoms may again be uncovered by an attentive history and examination. The patient may not be fully aware of any deficit. Hemisensory or visual neglect may, for example, manifest in the history as the patient bumping into doorways on the same side or walking out dangerously into traffic, or in the examination by the observation of asymmetry of personal grooming.
7. Unusual ("noncanonical") psychiatric symptoms. Although general statements should be only cautiously extrapolated from a body of case reports, the literature is quite striking in how often odd symptoms which "don't quite fit" with the classical psychiatric syndromes herald an occult brain tumor. Oft-reported descriptive psychopathology includes visual or musical hallucinations, palinacousis, palinopsia, or similar phenomena, with a scattering of rare and exotic symptoms (e.g., polyopic heautoscopy). The unusual development of pervasive new personality abnormalities in adulthood despite normal prior personality development is also a well-recognized and noncanonical symptom.
8. Objective neurological deficit. A comprehensive neurological examination can be difficult to perform in patients who are psychiatrically unwell but is a core part of the full assessment of someone with a possible brain tumor. Abnormalities may be picked up with relatively little effort by: testing pupils and eye movements (the III and IV cranial nerves have the longest intracranial course and may be particularly affected by raised ICP); looking for pronator drift (a sensitive test for pyramidal weakness in the upper limbs); testing muscle tone and spinal reflexes; assessing gait including a heel-to-toe test; and screening for dysmetria with the finger-to-nose test. It must be stressed that these examination components are very much just a basic and relatively pragmatic skim: a full examination should be done immediately or as soon as possible. In isolation, neurological "soft signs" such as sensory abnormalities and the subtle persistence of developmental reflexes lack specificity in psychiatric patients. If there is doubt, a neurologist or neuropsychiatrist can give a second opinion.
9. Papilledema. Maybe the most difficult "high-yield" sign of a brain tumor to elicit accurately in psychiatrically ill patients, papilledema is suggestive of raised ICP. There are many other causes but the presence of papilledema should always trigger further assessment (for a helpful website demonstrating fundoscopy and the different stages of papilledema, see Table 2.3–2 later in this chapter).
10. Endocrinological abnormality. Pituitary tumors can cause a variety of

endocrinological disturbances. Systemic physical or blood biochemical signs of acromegaly, Cushing disease, abnormal lactation, or altered menstruation should raise clinical suspicion. The classical neurological deficit of pituitary tumor is a bitemporal hemianopia secondary to midline compression of the optic chasm. These patients may, also, walk out into traffic.

Table 2.3–2.
A Selection of Potentially Useful Websites

URL	Comments
www.soc-neuro-onc.org	The Society for Neuro Oncology (North America, multidisciplinary)
https://www.cancerinstitute.org.au/how-we-help/reports-and-publications?ct=29&f=	Factsheets for patients and carers about managing the cognitive, emotional, and behavioral consequences of brain tumor
http://www.assbi.com.au/making%20sense.html	Purchase link for the "Making Sense of Brain Tumour" multimodal psychological intervention manual. Aimed at professionals.
https://issuu.com/brainstrust/docs/150309_what_the_community_needs_fin	Discussion document on the need for neurorehabilitation from the perspective of Braintrust, a national UK brain tumor charity
https://www.youtube.com/watch?v=EWFFAtNenRQ	A video tutorial about papilloedema.

All accessed January 8, 2015. Readers are also directed to local and national brain tumor charity websites for further information on general support.

Even a careful neurological assessment may not initially (or for a considerable period of time) elicit neurological symptoms or signs. This is especially the case with slow-growing tumors involving a relatively neurologically silent brain region such as the ventricles, corpus callosum, nondominant temporal and parietal lobes, and the frontal pole.

Investigations

Definitive brain imaging is certainly indicated in psychiatric patients with psychiatric symptoms accompanied by focal neurological findings and/or one or more of the symptoms listed above. The thing to do is of course obtain an urgent brain CT or MRI scan. In countries where these modalities are not available, other first-line options may include plain skull x-ray, electroencephalography (EEG), and lumbar puncture with CSF examination. The risk of causing fatal coning by lumbar puncture in the presence of raised ICP must be fully considered. In the absence of radiological evidence of a fully patent IV ventricle it may be safer to avoid lumbar puncture unless all signs of raised ICP have been sought and documented as excluded. Lumbar puncture may however be helpful when cytology studies are required in the assessment of tumors such as leukemias, lymphomas, and meningeal metastasis. In all cases of suspected brain tumor, a chest x-ray will help inform a basic assessment of the likelihood and source of any metastasis.

CT scanning can diagnose about 90 percent of cerebral tumors and is enhanced by the use of intravenous (IV) contrast. CT scans may be especially useful in the evaluation of tumors which are calcified, eroding bony intracranial structures, or causing shifts in middle cerebral structures. They can effectively evaluate abnormalities involving the ventricular system such as obstructive hydrocephalus, and gauge the extent of cerebral edema. Certain types of tumors remain difficult to identify with CT scans, though. These include: lesions smaller than 0.5 cm in diameter; tumors occurring next to bony structures (such as acoustic neuromas, pituitary tumors, some meningiomas, and anything in the lower posterior fossa); low-grade astrocytomas; and meningeal carcinomatosis. CAT scanning also involves a degree of radiation exposure.

MRI outperforms CT scanning in these respects. Owing to greater image resolution, MRI is also better at identifying the architecture of brain tumors (whether they are solid or cystic, or both), and in more precisely defining the relationship of a tumor to nearby vascular structures. However, MRI scanning is claustrophobic and takes longer: Some psychiatric patients may better tolerate CT, and CT is cheaper.

Newer quantitative diagnostic techniques can further improve the diagnosis and treatment of brain tumors. These techniques include single-photon emission computed tomography (SPECT), PET, magnetoencephalography (MEG), and diffusion tensor imaging (DTI) tractography. Although not yet in routine clinical use, these techniques may be useful in special situations. SPECT and PET may enhance the ability to differentiate tumor recurrence from radiation necrosis in patients who have received prior radiotherapy and in whom structural imaging suggests progression. They may also help discriminate between CNS lymphoma and opportunistic infections (such as toxoplasmic encephalitis in patients with AIDS). MEG and DTI tractography may be helpful in more precisely

characterizing the phenomenon of diaschisis and the various disconnection syndromes which frequently occur in brain tumor patients. MEG—along with functional MRI (fMRI)—is also used in the preoperative localization of specialized cortical function, such as motor, speech, and vision. These modalities aid neurosurgeons in planning stereotactic surgery that removes as much tumor as possible while minimizing the risk of critical cortical damage. Increasingly such location-critical operations are performed as "awake craniotomies" under sedation but with maintained consciousness. With the patient awake, surgeons can stimulate an area of brain before its removal and check whether it is functionally vital.

If radiological investigations reveal a likely brain tumor, the psychiatric patient should immediately be referred to the local neuro-oncology MDT for further discussion and appropriate work-up.

PSYCHIATRIC SYMPTOMS IN BRAIN TUMOR PATIENTS

Switching focus to the neuro-oncology team, it is time to consider the range of neuropsychiatric symptoms that may arise in the 18,000 or so people who are (for example) diagnosed with a primary brain tumor each year in the United States.

General Comments

Predisposing Factors

TUMOR LOCATION. The notion that brain tumors in particular lobes of the brain cause specific psychiatric symptoms is simplistic. Although there is a relationship of sorts, anatomical tumor location is only one of many factors to consider when assessing the nature and severity of psychopathology. For example, a "frontal lobe" dysexecutive syndrome can arise in the context of a brainstem tumor. Indeed, neuropsychiatric symptoms arising from structures far removed from the tumor is a well-recognized puzzle. It may partly result from the fascinating phenomenon of diaschisis, and partly from various white matter disconnection syndromes secondary to tumor-associated damage or disruption of interconnecting neural pathways, including the corpus callosum.

Similarly, many readers will have encountered the belief that tumors in the left hemisphere cause dysphoria and depression whereas right-hemisphere tumors cause euphoria and symptom denial. Indeed, right-sided lesions of the temporal skull base were associated with a disinhibition syndrome in one small but intriguing prospective cohort study conducted after CT scanning became widespread. However, the association between laterality and psychopathology is inconsistent. Most large studies of this issue predate the use of more recent psychiatric diagnostic classification schema, many the advent of CT scanning, and a few the introduction of modern research standards of methodology and reporting. Meanwhile,

most recent studies are case studies. Many of the clinical inferences from this literature are difficult to interpret.

Unpicking the etiological relationship between tumor location and psychopathology will need the use of more sophisticated neuroimaging techniques than CT or MRI. Few studies have taken advantage of the likes of DTI tractography or SPECT to address this issue. Those that have are limited to cognitive assessment of childhood survivors of posterior fossa tumors and are discussed briefly later.

TUMOR MALIGNANCY. Does tumor malignancy influence the development of neuropsychiatric symptoms? Like tumor location, various authors have proposed a causal relationship between WHO grade and psychopathology. The hypothesis is that rapidly growing high-grade tumors are associated with a greater severity and faster progression of psychopathology. Patients with indolent low-grade tumors, meanwhile, have more limited and stable symptoms. This high/low-grade difference makes intuitive sense and feels clinically right. It could be explained mechanistically by several factors. These include: differential dynamics of shifts in ICP; a differential metabolic profile in the cellular microenvironment between high- and low-grade tumors; and impairment of neuronal functioning by contact, or destruction, from invading cancer cells in more invasive tumors.

Once more, the picture is not as simple in reality. The question above is too vague: The extent and direction of any possible relationship may depend on several factors (e.g., prediagnosis or postsurgery; psychiatric or neurocognitive symptoms). There is also a lack of robust supporting evidence from prospective studies that directly compare high- and low-grade populations. The few studies that do exist lack power because low-grade glioma is rare. If psychopathology is with respect to more traditional psychiatric syndromes, then evidence supporting a role for tumor grade is at best inconsistent. For example, there is about as much supporting evidence for greater generalized distress in *low-grade* tumors, an association that could be partly explained by the younger median age of low-grade brain tumor patients, and the greater prevalence of epilepsy.

This rather bleak appraisal of the roles of tumor location and malignancy illustrates four things about general statements regarding the influence of tumor biology on neuropsychiatric symptoms. Such statements run the risk of: (1) minimizing the potential role of sociodemographic confounds; (2) eliding cognitive with more traditional psychiatric psychopathology when the two are probably mechanistically different; (3) forgetting that different time-points in the disease process might show a different relationship; and (4) sounding so incredibly right that the first three things are forgotten.

Properly powered comparative studies are needed to identify and allow support to be directed towards high-risk groups. Recent but as-yet-

unpublished conference data *does* suggest an intriguing relationship between a certain molecular subtype of faster growing tumors and greater cognitive impairment. The notion deserves further study, not least because it may offer useful insights into the mechanisms of how neuropsychiatric symptoms develop in other clinical populations.

PREMORBID AND PSYCHOSOCIAL FACTORS. In contrast to most variables relating to tumor biology, the patient's premorbid status and psychiatric history clearly influence their risk of postoperative psychiatric complications. Even after adjusting to the diagnosis, the ongoing stress may be harrowing: a constant threat of clinical progression, neurocognitive and physical disability, as well as the risk of morbidity from surgery, radiotherapy, or chemotherapy. The patient's premorbid cognitive capacity, coping skills, and adaptive or maladaptive behavioral styles, in conjunction with the adequacy and availability of psychosocial support systems, play important roles in determining the impact and degree of neuropsychiatric symptoms.

The contributing factors that determine the type and severity of psychiatric and behavioral symptoms that co-occur with brain tumors are therefore multiple and complex. To varying degrees they may include tumor type, the rate and extent of tumor growth, the anatomical location, the presence or absence of increased ICP, the types of treatment used and the type and severity of complication associated with them, premorbid patient characteristics, psychiatric history, the adequacy of coping skills, and the quality of psychosocial and family support systems. In adults, on balance and with a slight risk of over-generalizing, there is probably a higher likelihood of neuropsychiatric and/or neurocognitive dysfunction in patients with: frontotemporolimbic and bifrontal or deep midline tumors; large (>5 cm) or multifocal tumors; aggressively growing malignant tumors; increased ICP; obstructive hydrocephalus; significant tumor-related physical functional impairment; a history of psychiatric illness including drug or alcohol addiction; prior brain injury; a family history of psychiatric illness; lower premorbid intellectual capabilities; poorer adaptive coping skills; or lower psychosocial support.

Treatment

GENERAL APPROACH. When a psychiatric disturbance is directly caused by a cerebral tumor, its surgical removal may lead to complete remission of the behavioral or neurocognitive symptoms. Some readers may imagine that this covers the majority of brain tumor-associated neuropsychiatry. For isolated cases of (e.g.) meningioma or a functioning pituitary adenoma, surgery can indeed cure psychiatric symptoms, as the case report literature attests. In general though, most psychiatric morbidity in brain tumor patients occurs after surgery, during primary treatment and follow-up. The psychiatrist is most likely to be consulted when the patient's behavior or

neurocognitive symptoms persist, become more severe, or arise de novo during follow-up. Appropriate treatment may reduce symptomatic distress, improve functional ability, and enhance overall well-being and quality of life. Treatment typically involves nonpharmacological support, pharmacotherapy when necessary, psychoeducation and support of the family, and clear bidirectional communication about treatment with the neuro-oncology team. It can be very helpful to ameliorate nonspecific agitation, irritability, dysphoria, and anxiety, as well as any specific psychiatric symptomatology that may be present.

As noted above, only a small proportion of brain tumor patients have psychiatric disturbances exclusively due to direct tumor effects. The lifetime prevalence of psychiatric disorders in the general population is high: the same applies to people who will eventually develop brain tumors. In many cases, patients relapse in response to the fear and ongoing stress of being diagnosed with a brain tumor. Neurocognitive impairment is the biggest exception to this general rule: this very common problem can often be attributed directly to the tumor, or from side effects or complications related to its treatment. It is also itself one of the most challenging problems to treat.

In planning a therapeutic approach, the psychiatrist may find it useful to attempt to characterize the patient's psychiatric and behavioral symptoms as (1) a relapse, or (2) arising de novo in response to either the psychological or biological stress of having a brain tumor. This broad framework can be clinically helpful in communicating with patients and their families, and in planning optimal drug and nondrug interventions.

PRE-EXISTING PSYCHIATRIC DISORDERS. The management of pre-existing psychiatric disorders in neuro-oncology is grounded in common sense and best practice. The balance of risks and benefits of treatment should be fully discussed with the patient. Drug treatment follows the same general principles as in the general psychiatric population. The prescriber should adjust doses (and occasionally drugs) as necessary to accommodate for the increased risk of CNS side effects. Common side effects include delirium and the increased potential for an epileptic seizure. Prescribers should also anticipate cytochrome P450 (CYP450) system pharmacodynamic interactions between psychiatric drugs and the array of medication these patients may be on.

Many older psychopharmacological agents are potentially deliriogenic. Examples include the tricyclic antidepressants (TCAs), typical antipsychotics, anticholinergic antiparkinsonian drugs, the benzodiazepines as a group, and lithium carbonate. As a rule of thumb, these older agents are first-line only when they have previously been effective and well tolerated by the patient. The drug should be reintroduced at a low dose and gradually titrated to efficacy. In most patients, newer

drugs may be better as first-line treatment. These include the SSRIs, other relatively nonanticholinergic antidepressants, atypical antipsychotics, nonbenzodiazepine anxiolytics, and anticonvulsant mood stabilizers. Again start low and titrate slowly. Caution is especially needed in the elderly and those with multiple medical conditions who may be on many other medications.

Epilepsy is another concern. Many psychotropic drugs have the potential to variably lower seizure threshold. The risk of inducing seizures with different psychotropic medications in the different populations of brain tumor patients is however unknown. The newer atypical antipsychotics may have less epileptogenic potential at comparable doses than older typicals. If high-potency older antipsychotics are required, haloperidol (Haldol) may be less likely to lower seizure threshold than others in its class. For antidepressants, licensing data from RCTs in patients without a brain tumor suggest that bupropion (e.g., Wellbutrin) and maprotiline have a relatively high risk of inducing seizures. These particular drugs may therefore be best avoided, where possible, in brain tumor patients. A reasonable first choice would be an SSRI with few effects on the CYP450 system.

Pharmacodynamic interactions may occur through the CYP450 system; for example, some antidepressants are CYP450 system inhibitors. If an AED is metabolized by the same enzyme, administration of a directly competing SSRI could predispose to an interaction. In this particular example, the clinical risk would be of AED toxicity (and perhaps not, as might be assumed, increased AED metabolism and inefficacy). In the reverse direction, coprescription of an enzyme-inducing AED might necessitate higher therapeutic levels of an antidepressant. This example is extremely general: prescribers should consider interactions on a case-by-case basis, consulting the pharmacist, if necessary. If interactions cannot be avoided, clinicians may wish to monitor anticonvulsant drug levels. Others may simply elect to closely monitor the patient clinically.

In brain tumor patients with a history of schizophrenia or other psychotic disorders, atypical antipsychotics should be used where possible. In some circumstances, it may still be necessary to use high-potency agents with a long history of successful use in medical disorders (such as haloperidol), or other individual drugs known to be well-suited to a particular patient. With appropriate dosage adjustments these may well be tolerated.

Patients with longstanding type I bipolar disorder are at the same risk of developing a brain tumor as everyone else. As is also the case with schizophrenia, the combination of these two devastating illnesses is rare. A complicating factor in this population is that lithium carbonate can cause seizures and delirium and should be avoided if possible. There may be cases where lithium has been used for many years and where the relative

risks and benefits can only be weighed on an individual basis. Mostly though, in acute type I bipolar mania, a mood-stabilizing agent with documented efficacy in acute mania such as valproic acid (e.g., Depakene) would be a reasonable first choice. An anticonvulsant mood stabilizer may have obvious additional advantages in patients with epilepsy. An atypical antipsychotic may be added if required. In type I bipolar depression there is a well-known risk of precipitating secondary mania or rapid cycling when traditional antidepressants are used. The anticonvulsant lamotrigine (Lamictal) may be effective here and appears to have the advantage of not precipitating secondary mania or causing rapid cycling, although it also has the potential for causing rare but serious and potentially life-threatening dermatological side effects.

Common Neuropsychiatric Syndromes

Neuropsychiatric problems often arise after the diagnosis of a primary brain tumor. Here, the most frequent ones are discussed (for delirium which can also be a common problem, see Section 10.2). Childhood brain tumor survivors are a distinct group and are discussed generally.

Depressive Disorder due to Brain Tumor

NOSOLOGY. The psychiatrist diagnosing DSM-5 major depressive disorder (MDD) in brain tumor patients faces a series of difficult judgments that broadly fall into two groups. One is whether the somatic symptoms (weight, sleep, and psychomotor changes, and fatigue) should be counted when there is considerable risk of confounding from biological effects of the tumor and its treatment. The other broad group of judgments is whether psychological symptoms (depressed mood, anhedonia, guilt, poor concentration, and suicidal ideation) are abnormally severe, when considered in the context of multiple personal losses and a profound existential threat. Because these judgments are difficult they can cause variability in clinical practice. This might threaten the intended reliability of the DSM system: indeed, this is what happens.

Inter-rater reliability of DSM criteria for MDD in a typical brain tumor patient is questionable. Data collected by the author with supervision from Dr. Robin Grant and Dr. Alan Carson (Consultant neuropsychiatrist, NHS Lothian). When assessing patients with a brain tumor for depression doctors have to decide whether to attribute symptoms to a likely mood disorder, or else to the direct consequences of the tumor and its treatment. Different doctors may interpret the same diagnostic information in different ways, causing unreliability but still influencing their treatment decisions. To test this, a clinical case study was constructed for a hypothetical brain tumor patient meeting criteria for Major Depressive Episode (DSM-IV-TR 296.2). The vignette was sent to NHS Consultants in five neuro-oncology related specialities in Scotland. The aims were to determine how many doctors diagnosed depression; whether their diagnosis influence their preferred treatment; and among doctors who diagnosed depression, specifically whether concern over a history of epilepsy influenced their readiness to recommend an antidepressant.

A written vignette was developed for a typical patient with malignant glioma and a history of epilepsy. Contained within the vignette were all nine of the symptoms from the diagnostic criteria for DSM-IV-TR MDD. However, by design, each symptom also had a realistic alternative explanation directly related to the tumor or its treatment. Respondents were asked to state the diagnosis ("major depression" or "an understandable reaction"); specify treatment; and state to what extent the history of epilepsy had influenced their treatment decision ("not at all," "a little," "moderately," or "markedly"). All NHS consultant neurologists, clinical oncologists with neuro-oncology interest, neurosurgeons, and liaison psychiatrists working in Scotland, as well as the general practitioners (GPs) of all local glioma patients were invited to participate. Statistical analyses were conducted, where appropriate, with Fisher's exact test or Chi square. The vignette read as follows:

"You are seeing Mrs. Smith, a previously fit and well 54-year-old language teacher, in your clinic. Four months ago she presented with complex partial seizures, followed by a prolonged generalized tonic–clonic seizure which required admission to ITU. CT had revealed a left temporal lobe tumor which was debulked and found to be a high-grade glioma. She completed a course of radical radiotherapy last month and continues on chemotherapy. There is no other past medical history and no past psychiatric history.

About a month ago Mrs. Smith found a website on the internet saying chemotherapy gives her a 20% chance of being alive in 2 years. She has thus been thinking a lot about her death and how it would affect her family, and has been tearful and feeling sad since then. Her hair fell out toward the end of the radiotherapy, which she found distressing. She used to enjoy sport, and meeting friends, but says that both the self-consciousness and a marked increase in tiredness have left her totally disinterested in her old activities for some weeks. She feels bad about this, worries that she is letting her friends down, and struggles with thoughts that she has become a "weak" person instead of being her usual strong self. She has put on 5 kg in weight since she stopped going out.

Mrs. Smith is currently unable to return to her teaching job due to persisting word-finding difficulties. She has begun to lie awake at night worrying about her finances. Since the radiotherapy finished she has also been much more forgetful, and finds it hard to read the newspaper. She smokes 10 cigarettes per day, does not drink, and has a supportive relationship with her husband and two adult daughters. One has just become pregnant and Mrs. Smith finds the thought of not seeing her first grandchild particularly painful. She would consider going to Switzerland for euthanasia if she deteriorated and was no longer able to walk, but otherwise hopes to live to see the child.

Steroids were stopped soon after the operation. She was initially started on phenytoin but switched to a low dose of lamotrigine after the operation. It was increased slightly to control partial seizures a few weeks later. She remains on the same dose and has had no seizures since. She is on no other medicines and has no allergies.

Today's brain MRI shows postoperative scarring with a small amount of residual tumor. At interview, Mrs. Smith sits with little spontaneous movement, and cries when talking about her family. She is not psychotic. Her speech is quiet, and halting at times. Her concentration occasionally wanders but she is orientated to time and place. Mrs. Smith asks you if this is a normal reaction to her situation and treatment. She has no strong feelings about how you manage her."

Of 228 questionnaires distributed, 105 were returned complete (46%

response rate). Sixty-four doctors (61%) diagnosed MDD and 41 (39%) diagnosed an "understandable reaction." There were no significant interspeciality differences.

Choice of diagnosis strongly influenced treatment. Antidepressants were more likely to be prescribed for MDD (38/64 vs. 2/41; $p < 0.00001$). For an "understandable reaction," counseling (23/41 vs. 15/64; $p < 0.001$) or watchful waiting (13/41 vs. 0/64; $p < 0.00001$) were more likely to be suggested as the first-line management. Eleven distinct types of counseling were identified with no consensus on the best sort.

Just over half of the doctors diagnosing MDD stated that the history of epilepsy had "moderately" or "markedly" influenced their choice of treatment (34/64). There was a trend for concern about epilepsy to be associated with lower antidepressant prescription. Only 17/34 moderately or markedly influenced doctors prescribed an antidepressant versus 21/29 little- or not-at-all–influenced ($p = 0.079$). GPs and psychiatrists appeared to be more concerned about epilepsy than neurologists ($p = 0.007$ and $p = 0.041$, respectively).

In summary, a sample of over 100 doctors presented with the same clinical information showed marked disagreement in whether to diagnose MDD in a brain tumor patient. Importantly, preferred diagnosis strongly influenced preferred management. In considering treatment, psychiatrists were significantly more concerned about a history of epilepsy than neurologists.

Attempting to mitigate these problems by setting diagnostic "rules" for individual symptoms often raises further questions. For example, a doctor may decide to exclude all reports of insomnia if the patient is also on high-dose corticosteroids. But insomnia is a common symptom of depression and the relationship between steroids and sleep is not straightforward. In a local series of over 150 glioma patients, most of those taking dexamethasone shortly after primary surgery reported either no sleep changes at all or hypersomnia at clinical interview (author's nonpeer–reviewed data). Thus, it may often be impossible to confidently attribute causality to individual depressive symptoms.

What can be done? In the "general comments" section above, readers were urged to consider whether the patient's depressive disorder arose (1) as a recurrence of prior depressive disorder, or (2) de novo in response to either the psychological or biological stress of having a brain tumor. This provides a clinically useful framework by which to communicate the meaning of a diagnosis to patients. It may also suggest particular therapeutic strategies. However, as can be seen above, the DSM entities of MDE (Major Depressive Episode) or MDD do not make it easy to secure the diagnosis in the first place. A better academic strategy may be needed as well.

DSM-5 explicitly instructs users to discount depressive symptoms if the

symptoms are clearly due to another medical condition. Yet as shown in the vignette above, every single symptom of MDD/E can often quite plausibly be attributed to the brain tumor or its treatment. Given this bind, psychiatrists may therefore choose instead to *assess all potentially depressed patients with a brain tumor as per the entity "mood disorder due to brain tumor"* (hereafter "Depressive Disorder," DSM-5 293.83). This can be done regardless of any (theorized) cause such as tumor effect or a psychological reaction. For the same reason, the presence or absence of a prior history of MDD/E can also be seen as secondary. The criteria for depressive disorder are discussed below: crucially they are different from MDD and do not rest on theoretical judgments. Only additional specifiers discriminate between various kinds of depressive syndrome. In short, when assessing depression the diagnosis sought by psychiatrists could be automatically assumed, in all brain tumor patients, to be due *in some way* —presumably pathophysiological at some level—to the tumor.

Objections to a blanket approach might include: (1) it risks implying that some recurrences of prior depressive disorder after brain tumor diagnosis are caused by a new mechanism rather than being a simple relapse; (2) the diagnostic threshold is potentially reduced because the number of required symptoms is lower, which could lead to overdiagnosis; (3) psychiatrists still have to grapple with the relevance of each individual symptom when assigning specifiers; and (4) it is simplistic. But equally, the strategy: (1) is testable: for example, studies could assess its validity by grouping brain tumor patients with depressive disorder into those with, and without, a prior history of depression, and seek differences in symptom profile, clinical course, and/or response to treatment; (2) minimizes interclinician variability in practice, helping to standardize diagnosis; (3) cuts much of the Gordian knot of whether to count somatic and psychological symptoms toward the main diagnosis because they are not initially relevant; and (4) benefits from simplicity in terms of communicability and tractability. It may also pragmatically solve a real-life problem whilst remaining true to the DSM directive for MDD. Perhaps helpfully, it would both allow this patient group to be more easily identified in future epidemiological studies, and accommodate the ubiquitous and pressing context: there is a brain tumor in the person's head that has comprehensively changed their life. It would therefore be a relatively simple, pragmatic, reliable, testable, and academically useful solution to the problems associated with diagnosing MDD in a brain tumor patient.

Apart from the upcoming section on "Diagnosis and clinical features" which for consistency also focuses on the suggested pragmatic diagnosis of depressive disorder, the rest of this section is based on the historic literature. The authors of this literature, including the present author, used either traditional DSM diagnoses of MDD or questionnaires which were validated against MDD (or its temporal equivalent).

EPIDEMIOLOGY. MDD is considerably more common in glioma patients than in the general population. The biggest individual study to use a "clinical interview" diagnosis in these patients found that the 6-month prevalence was 20%; reviews estimate a corroborative median frequency of 15%. If depressive symptoms are measured by rating scale, ballpark estimates vary depending on the scale. For example, the Beck Depression Inventory (BDI) returns a median frequency of 39%, whereas the Hospital Anxiety and Depression Scale (HADS) suggests 16%. Original data and reviews show that MDD is most consistently associated with impaired physical and cognitive functioning, a prior history of depression, and reduced quality of life. Unlike in the general population, female sex is not a clear risk factor. Depressive symptoms may predict increased mortality but further studies are needed on this issue. Longitudinal studies of depression are rare: most likely a clinical diagnosis of depression persists in about half of patients, for at least 3 months. The inflating contribution of (nonmood disordered) biological symptoms of tumor or treatment to these estimates is not known.

ETIOLOGY. The etiology of MDD in brain tumor patients is also unknown. At a population level, there is a relationship in many observational studies between a past history of depression and an increased risk of depression posttumor diagnosis. This suggests that in some cases, the cause is a relapse of a constitutive depressive disorder. Genetic influences can be inferred when there is a family history of depression. Posttumor diagnosis, a new case of depression, could be explained by biological theories including: tumor- or edema-associated disruption of critical tracts, or a functional diaschisis; critical metabolic alterations of the neuronal microenvironment; a putative depressogenic effect of long-term corticosteroids; and possible long-term effects of radiotherapy or chemotherapy on hippocampal neurogenesis or white matter tract biology. An equally strong case can often be made for emphasizing a psychological reaction to the devastating array of losses that accompany the onset, treatment, and natural course of many brain tumors. Various neuroendocrine abnormalities may be etiologically implicated in pituitary tumors including hypercortisolemia and hypothyroidism.

A subset of brain tumor patients meet criteria for a depressive disorder by having marked anhedonia but no depressed mood. The clinical picture can be striking, is recognized in both pituitary tumor and glioma, and conveys the strong sense of a disorder distinct from the classical depressive picture of pervasively sad mood. The biological basis of any such difference is unknown.

DIAGNOSIS AND CLINICAL FEATURES. Mood disorder due to brain tumor requires clinical interview evidence of a prominent, persistent, and socially

disruptive change of mood. In the current context the change is depressive, apathetic, or both. Symptoms occur in the context of a radiologically or histologically diagnosed brain tumor. Adjustment disorder with depressed mood should be excluded along with delirium. It can be further specified with respect to the full criteria for MDD/E (Section 10.5): w*ith major depressive-like episode* if they are met, and *with depressive features* if not.

The traditionally required minimum duration of symptoms is 2 weeks. While still observing this requirement, some psychiatrists allow at least 1 month to pass from the point of brain tumor diagnosis before *first* considering whether a mood disorder is present, unless extremely severe symptoms manifest before then. Waves of grief are common post diagnosis and regardless of the timing of assessment, the persistence criterion is only met if symptoms are present for most of the day, on most days of the week. Social disruption can be difficult to assess. Many brain tumor patients are unable to drive. They are often more socially isolated than previously; they are also more likely to be unable to work. Questions about social interaction can focus on what the patient *is* able to do. The psychiatrist can assess their daily interaction with carers, family, and friends, and whether this has changed along with mood. If anhedonia is the main clinical feature, a similar problem arises: questions about "enjoying things they normally enjoy" may be met with the reply that they can't do those things now they have a brain tumor. One useful question may be whether they still feel the same pleasure from spending time with grandchildren or other close family. All symptoms should be new and not attributable to the patient's long-standing personality. Excluding adjustment disorder is a matter of assessing severity and persistence of the cardinal depressive symptoms. Excluding hypoactive delirium may be difficult if the patient is already cognitively impaired, but in the history, the absence of fluctuating agitation or disorientation in time should be recorded. Proxy report (e.g., from the patient's main carer) is valuable and adds to diagnostic reliability.

PATHOLOGY AND LABORATORY EXAMINATION. Investigations are as those required to exclude delirium (Section 10.2) and a physical cause for MDD (Section 10). An endocrine screen is particularly indicated in patients with a remote history of radiotherapy to the brain, if the radiation field included the pituitary gland.

DIFFERENTIAL DIAGNOSIS. The main competing psychiatric diagnoses are delirium, neurocognitive disorder due to a brain tumor, and (probably most commonly) adjustment disorder with depressed mood. The medical differential is as broad as that for MDD in patients without a brain tumor (Section 10). In brain tumor patients specifically, potential medical causes include: postepileptic personality change; neuroendocrine abnormalities (including systemic diseases such as secondary hypothyroidism); and

steroid-induced depressive symptoms.

COURSE AND PROGNOSIS. Longitudinal studies of MDD in brain tumor patients are rare, partly because MDD is most likely in functionally impaired patients who generally have a poor prognosis. Available evidence suggests that of those patients who can be followed up, the disorder persists for at least 3 months in around one half. This is in keeping with median duration in the nonbrain tumor population and lends some support to the argument that depressive symptoms cause significant clinical burden in these patients despite concerns about nosology.

MDD and moderate–severe depressive symptoms have been clearly shown to correlate with shorter survival times. They also correlate with functional impairment, a strong negative prognostic factor for brain tumor patients. An open question is therefore whether depression is independently predictive of mortality after controlling for functional impairment. Some studies suggest that it is and others do not. Median survival times for the most often studied tumor (high-grade glioma) are naturally so short that large sample sizes are likely needed to show even a modest effect size. An individual patient data meta-analysis may be one way to clarify this issue.

MDD is associated with an increased risk of epileptic seizures in the general population. It is possible that left untreated it predisposes these vulnerable patients yet further to epilepsy. The frequency of DVT/PE is also high in brain tumor (approaching one-quarter in some series). The apathy and psychomotor retardation of untreated severe depression may exacerbate this problem too. Suicidal ideation, though rare in some cohorts in the clinically busy first few months after brain tumor diagnosis, is also a complication of uncontrolled MDD. Brain tumor patients have one of the highest risk ratios for completed suicide in the year following diagnosis of tumor.

TREATMENT. No phase III RCT has examined the effectiveness of antidepressants for MDD in brain tumor patients. Although the same can be said for nearly every psychiatric condition in this patient group, the absence is felt more acutely here. Depression is a common condition which many psychiatrists will feel relatively comfortable recognizing and prescribing antidepressants for, yet in brain tumor patients there is clinical equipoise between their effectiveness and risk. The equipoise is on the one hand whether these drugs are more effective than placebo in the grossly abnormal cerebral environment of a primary tumor (glioma especially). On the other, the likelihoods that they cause AED or chemotherapeutic toxicities by CYP450 interaction, worsen fatigue as per the main recognized side-effect profiles of all antidepressants, contribute via anticholinergic mechanisms to cognitive impairment, worsen seizure control in this highly

susceptible population, or even by contrast improve seizure control by treating depression, are in each case also unknown. These factors (professional confidence, and for want of a better term academic ignorance) might combine to create a rather risky "blind optimism" in favor of automatic prescribing for MDD. Or doctors might fear unquantified consequences like epilepsy, and hesitate to prescribe (as suggested by the study "Inter-rater reliability") when treatment may actually be effective. The best answer is to conduct RCTs in depressed brain tumor patients who meet treatment thresholds at enrolment. In the meantime, treatment is based on commonly accepted principles and empirical best practice. Patients should be fully informed about the possible risks and benefits, treated cautiously, and reviewed regularly.

In some patients, antidepressants certainly appear clinically to be helpful in treating new depressive symptoms. Given that depression plays a major role in decreasing the overall quality of life that these patients experience, its early recognition and rapid treatment is potentially of great benefit. An SSRI with a relatively clean CYP450 side-effect profile such as sertraline (Zoloft) may be a reasonable first-line treatment. The SSRIs are increasingly clinically familiar, seem to be generally well-tolerated, and may be less likely to cause delirium than older antidepressants. A particular clinical drawback with these agents can be troublesome sexual side effects: brain tumor patients are already at high risk of sexual difficulties. The prescribing psychiatrist should check that patients who have been prescribed corticosteroids to reduce brain edema also have gastroprotective cover to minimize the risk of a gastrointestinal bleed. An ECG should be checked before prescribing to confirm that the QTc interval has not been prolonged by chemotherapy.

TCAs have a side effect profile which includes sedation, anticholinergic effects, and orthostatic hypotension. For this reason they are often considered second-line unless previously effective for the patient. If TCAs are considered for refractory depression, some clinicians report that nortriptyline (Pamelor) may be generally well tolerated. In some circumstances (such as severe refractory depression or a long history of previous effective treatment), monoamine oxidase inhibitors (MAOIs) may also be appropriate. MAOIs should not pose undue risks provided a tyramine-free diet is maintained, and interactions are checked as usual. Before using these agents the psychiatrist should assess the patient's cognitive capacity with respect to successfully observing the dietary restrictions.

Rarely, and in severely depressed brain tumor patients refractory to drug treatment, electroconvulsive therapy (ECT) may be considered. ECT was once thought to be absolutely contraindicated in the presence of brain tumor. Case studies, however, report that unilateral brief pulse ECT can be safe, effective, and well tolerated. Important precautions include close

attention to the patient's capacity to consent to treatment, careful dose titration, and the need to monitor the patient carefully for post-ECT delirium.

Psychostimulants can be effective in treating cancer-related fatigue and have been proposed as a treatment for depression in patients with brain tumor. Systematic trial evidence is currently quite limited. At present there is no good evidence that modafinil (Provigil) is effective for depression. Methylphenidate (Concerta) may reduce subthreshold depressive symptoms, but studies to date measured these only as secondary outcomes and conclusions are also inconsistent. Whether these drugs are more effective than placebo in treating clinical MDD is not known at all. Although anecdotally some clinicians find them helpful in some patients, studies do suggest a considerable placebo effect. If a trial of treatment is considered, methylphenidate has the advantages of a negligible effect on seizure threshold and no major sedating or deliriant properties. Adverse effects are still three times more likely than with placebo, however. In a recent meta-analysis, vertigo, anxiety, anorexia, and nausea were significantly more frequent in the treatment arm.

For all patients with MDD, psychological treatments should be considered and offered, wherever possible, alongside antidepressants. Once more it is unclear whether these treatments work (and if so which ones, and in what circumstances) in patients with progressive cognitive and communication deficits of varying severity. Decisions about their suitability for individual patients should be made empirically. Patients with a lower severity of subthreshold depressive symptoms and distress, meanwhile, may benefit from supportive or group therapy or referral to external services such as a brain tumor charity.

Anxiety

EPIDEMIOLOGY. Depending on the diagnostic method, screens for anxiety can capture elements of "normal" distress as well as the more severe and specific symptoms that present to adult psychiatry. In the brain tumor literature, anxiety has generally been measured using self-report questionnaires of the more general type. Frequency estimates usually range between 30% and 50%: how much of this morbidity is clinically diagnosable is not well characterized. These figures probably reflect, in part, reaction to the natural fear, agitation, uncertainty, and stress of having a brain tumor. Anxious reaction is especially likely in low-grade glioma, in younger adult patients, in women, and in those with lower premorbid IQ.

ETIOLOGY. Rarely, anxiety can be caused directly by the brain tumor. One example would be a coarsening of pre-existing anxious personality traits as the tumor disrupts higher inhibitory circuits. Another direct tumor cause

would be stereotyped panic-like attacks occurring in the context of complex partial epileptic seizure. Isolated case reports further record the association of brain tumor and new-onset obsessive or compulsive symptoms. However, as noted, most anxiety and "distress," even when severe, can realistically be framed in cognitive behavioral terms as a psychological reaction to the multimodal stress of having a brain tumor. Anxiety presenting during outpatient follow-up, meanwhile, is sometimes related to occupational stress. For example, a hairdresser with a low-grade tumor who has returned to work may have subtle cognitive deficits that cause difficulty in such a high-pressure environment.

DIAGNOSIS AND CLINICAL FEATURES. Generalized anxiety disorder (GAD, 300.02) has a minimum symptom duration requirement of at least 6 months. Additionally and as with MDD, several of the necessary symptoms (e.g., being easily fatigued, or difficulty concentrating) can nearly always be plausibly alternatively attributed to physiological effects of the tumor or its treatment. This combination of features means GAD is unlikely to be often the preferred diagnosis. The most likely clinically diagnosable anxiety disorder in brain tumor patients is adjustment disorder (309.xx).

A diagnosis of adjustment disorder requires emotional or behavioral symptoms occurring in response to the stress of having a brain tumor. The specific stressor may be the diagnosis itself or one of the many possible physical or social complications that can arise during treatment and follow-up. Symptoms should arise within 3 months of the given stressor. The criterion of "disproportionate" distress may be difficult to assess in the circumstances; the same is true for occupational impact. The significance of impairment in social functioning can be probed, though, by asking about the impact of symptoms on home life and contacts with friends. The exclusion criterion of "normal bereavement" should be taken at face value. The difference between bereavement and the many personal losses that can accompany brain tumor is unclear but perhaps of secondary import in the immediate context. Additional specifiers code for prevailing mood and behavioral state.

Differentiating adjustment disorder "with depressed mood" from the competing syndrome of depressive disorder may be difficult. One pragmatic approach might be to focus initially on the cardinal symptoms of depressive disorder (low mood and anhedonia). If either would be severe enough to contribute toward a diagnosis of MDD in the general population, the preferred diagnosis would be depressive disorder. If not, then screening for adjustment disorder would be appropriate (Fig. 2.3–3).

Diagnostic features of panic disorder (300.xx) are fully discussed in Section 10. In the general population, classical panic attacks have a characteristically sudden onset, stereotyped (repeating) pattern of features, and relatively short duration. Unfortunately, in brain tumor patients, much

the same things can be said about epileptic seizures. Thus in the brain tumor patient, it may be difficult, purely on the basis of a history, to distinguish between these two causes of apparent panic. One symptom which may point toward complex partial epilepsy is the presence of an unusual prodromal aura such as déjà vu, rising gastric sensations, a strange smell, or visual disturbance. Another is the presence of orofacial movements during the seizure. For this reason, a corroborative eyewitness account is recommended when assessing panic in a brain tumor patient. With the patient's consent, carers may be asked to video-record an attack on their mobile phone for the psychiatrist's review.

DIFFERENTIAL DIAGNOSIS. Psychosis and depressive disorder should be ruled out. Alongside the medical differential for anxiety disorders, it may be helpful to consider in particular complex partial epileptic seizures, pituitary-related endocrinological abnormalities (hypopituitarism being a late consequence of certain radiotherapy regimes), and PE/DVT presenting as difficulty breathing or hyperventilation.

PATHOLOGY AND LABORATORY EXAMINATION. Most anxiety will be—for want of a better term—psychological in origin. It is still valuable to exclude potential reversible physical causes. Brain tumor-specific investigations may include EEG, endocrine blood screen, and a chest x-ray as per the medical differential.

COURSE AND PROGNOSIS. The general cancer literature shows that anxiety symptoms characteristically peak at the times of diagnosis, end of primary treatment, relapse, and the terminal phase. Distress occurring at these points can be observed supportively unless it is very severe. Apart from the terminal phase when palliative services should be consulted, patients and their families can usually be reassured that it is normal and will pass. For example, much of the generalized distress at diagnosis resolves (as one might expect) within the first few months. In many instances, brain tumor patients face the added and recurring stressor of routine interval scanning: repeat neuroimaging at intervals usually ranging from every 3 months to every year and tracking progress of the tumor. As representatives from the charity sector confirm, these times can be highly distressing for those affected. They are temporary and supportive care is again indicated.

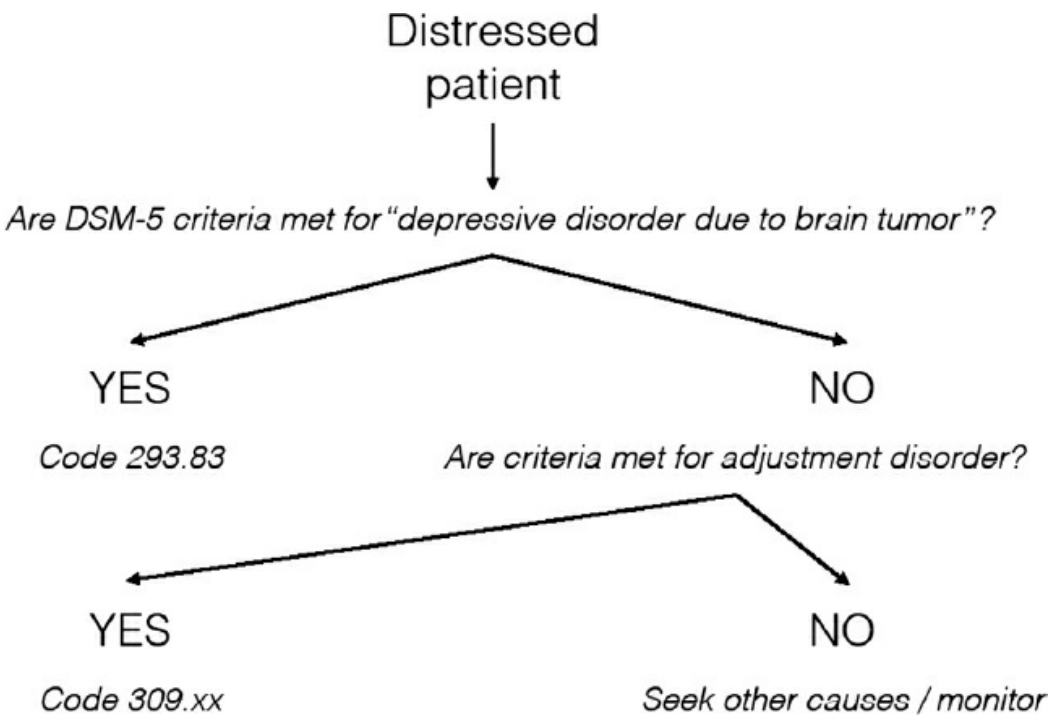

FIGURE 2.3–3. Suggested initial decision tree for assessing a distressed brain tumor patient.

Symptoms may persist preferentially in the minority of patients with other risk factors for psychiatric disorders, a premorbid anxious predisposition, or a direct tumor-related cause of symptoms. In these patients, or in those with severe levels of distress or anxiety at the points noted above, more systematic treatment may be indicated.

TREATMENT. Currently, no phase III RCTs have studied interventions for anxiety in brain tumor patients. General management principles from adult psychiatry can usually be applied, however. In a telling demonstration of the power of these principles, anxiety management techniques can even be used intraoperatively to facilitate awake craniotomy without sedation and only minimal analgesia. A cognitive behavioral formulation of the genesis and maintenance of anxious feelings may also be a useful starting point for treatment decisions.

In treating low-level anxiety or distress occurring in reaction to the psychological stress of having a brain tumor, full and detailed explanations of all diagnostic procedures and proposed treatments may help. There is observational evidence that giving information lowers anxiety, for example, in the perioperative phase. The patient and family should have the opportunity to ask questions and to have them as fully answered as is possible. Patients can also be referred to a trusted local brain tumor charity. Often, no further treatment will be needed. In patients whose anxiety persists or is moderately severe, supportive or structured psychotherapy is indicated. Not all patients may be suitable for these

therapies. Severe cognitive, functional, or practical concerns may preclude their effective implementation. To save distress and health care resources, when considering a psychological therapy the psychiatrist should consider the likelihood of patient compliance. Where possible, adjustments should be made to allow the patient to participate. Questions to consider include: (1) Does the patient have cognitive dysfunction that is too severe for them to engage with treatment and internalize therapeutic strategies? (2) Do they have a functional physical disability that may limit their participation, such as hemiparesis or hemianopia making traditional written "homework" impossible? (3) If the patient is unable to drive, how will they get to regular sessions? and (4) If there *is* a cognitive, physical, or practical difficulty, is it possible to adapt therapy to compensate? These questions and any other foreseeable problems should ideally be considered by the psychiatrist and treating psychologist together in close liaison with the neuro-oncology team.

With respect to drug treatment, most clinicians feel that antipsychotics should be avoided unless specific psychotic symptoms are associated with the patient's anxiety and agitation. The mainstays of anxiolytic drug treatment in brain tumor patients are the SSRIs, buspirone, and low-dose, long-acting benzodiazepines, in conjunction with supportive or structured psychotherapy. Alternative medications such as pregabalin (Lyrica) may be helpful, as in some circumstances may a low-dose TCA.

The general principle for patients with a pre-existing, longstanding anxiety disorder who subsequently develop a brain tumor is to continue previously effective treatment where possible. In line with the general comments above, attention should be paid to the possibility of interactions with other prescribed medicines. One or more of the SSRIs, buspirone or clonazepam (Klonopin), alone or in combination, may be effective and well tolerated in treating acute symptomatic exacerbations. Patients with fear or panic disorder symptoms occurring as a part of a temporal lobe tumor-induced complex partial seizure may respond to treatment with AEDs. If such symptoms occur during the interictal period, anxiolytic agents discussed previously may be helpful. Close collaboration with the neuro-oncology team is essential and particularly with the neurologist in respect of monitoring and maintaining epileptic control.

Steroid-Induced Psychotic Disorder

EPIDEMIOLOGY. Corticosteroids are prescribed to brain tumor patients to reduce brain edema, lower ICP, and thereby improve physical and cognitive functioning. The typically high doses (in the United Kingdom the standard initial dose is 8 mg dexamethasone orally, twice per day) and often urgent necessity mean that newly diagnosed brain tumor patients are at risk of developing a steroid-induced psychotic disorder (hereafter "steroid-

induced psychosis"). This sounds worrying but most brain tumor patients do not become psychotic from any cause, let alone steroids. The frequency of steroid-induced psychosis is estimated at under 5% from review of case and clinical study literature. Much of this literature is focused on rheumatology patients receiving prednisolone. It suggests that women may be more at risk, but the most consistently identified risk factor is higher steroid dose.

ETIOLOGY. The relationship between dose and emergence of psychotic symptoms is incompletely understood: whilst probably not linear, psychotic symptoms are clearly more likely in patients on very high doses. The molecular pathological mechanism is unknown and may be due to a primary effect of the corticosteroid on brain function or a secondary effect (e.g., severe sleep disruption), coupled with particular—and currently unknown—biological predisposition(s).

DIAGNOSIS AND CLINICAL FEATURES. The diagnosis of steroid-induced psychosis rests upon psychotic symptoms (delusions and/or hallucinations, together with other symptoms consistent with psychosis) occurring in association with the use of corticosteroids. The risk is highest soon after starting treatment. High dose (>40 mg prednisolone equivalent/day) makes the syndrome more likely but is not essential: the reaction appears to be idiosyncratic. The psychiatrist should exclude a prior history of functional psychosis and, if possible, delirium. The latter may not be straightforward: the clinical picture may fluctuate from psychosis and white cell count is likely to be raised secondary to the corticosteroid. To help exclude delirium, clues to infection should be sought in the physical examination. These might include fever, tachycardia, and a raised respiratory rate, as well as various focal signs of infection. Investigations can help include metabolic and infective causes (electrolyte or liver function abnormalities, urinalysis, chest x-ray if possible). The drug prescription sheet should be reviewed for any potentially deliriogenic medicines.

Steroid-induced psychosis may often have a prominent affective rather than schizophreniform flavor. There is however no single characteristic phenotype other than the presence of delusions or hallucinations. Anecdotally, oncologists treating brain tumor patients have commented that their patients (none of whom being psychotic) are often subtly elated for the first few weeks of corticosteroid treatment.

DIFFERENTIAL DIAGNOSIS. The differential diagnosis of steroid-induced psychosis in brain tumor patients includes delirium, psychosis secondary to complex partial epilepsy, MDD with psychotic features, and a stress-induced relapse of a pre-existing functional/affective psychosis. There is no clear evidence that either paranoid schizophrenia or type I bipolar disorder

are associated with brain tumor whether in cause or effect. In the fast-moving environment after presentation to hospital however, on occasion the clerking medical history may be incomplete, especially for historical psychiatric episodes.

PATHOLOGY AND LABORATORY INVESTIGATION. In cases of immediate risk, the priority is to treat the patient as per any severe psychotic episode. Physical investigations to consider as the clinical situation allows are those relevant to delirium and the major functional psychotic syndromes. Occasionally, an EEG may be clinically useful. Epileptogenic foci in deep frontal lobe structures can occasionally present with a bizarre behavioral syndrome that looks quite unlike either a classical generalized or complex partial epileptic seizure.

COURSE AND PROGNOSIS. Psychotic symptoms should resolve with the combination of antipsychotic and a lowering or reduction of steroid dose. If there is no evidence of improvement, perhaps particularly in patients with schizophreniform features, clinical attention should be given to the possibility of an emergent or previously undiagnosed functional psychotic illness such as schizophrenia or bipolar I disorder. In this situation a detailed family history may be of additional value. Coemergence of the catastrophic events of brain tumor and a functional psychosis is rare but does occasionally happen in younger patients. There is currently no robust evidence that it is anything other than a coincidence of timing.

The literature is sparse on prognostic guidance. What evidence there is suggests that prior steroid-induced psychosis may not accurately predict the likelihood of future episodes. Unlike—say—puerperal psychosis where even a single previous episode greatly increases the future risk, florid psychic reactions to corticosteroids appear to be largely idiosyncratic both between individuals, and within the same individual over time.

TREATMENT. If possible, the offending corticosteroid should be reduced or stopped. This may have consequences such as the re-emergence of cognitive or neurological symptoms; the decision rests on a balance of risks specific to each individual case. The psychiatrist should urgently consult with the treating medical or surgical team. Infrequently, the risk of altering steroid dose may be deemed too high: one example might be in the case of a brainstem tumor threatening to compress breathing centers.

Standard nonpharmacological methods to reduce patient agitation, anxiety, and disorientation can be applied. These methods are identical to those used widely for delirium and dementia and include nursing by familiar faces, adequate ambient lighting, regular reassurance and re-orientation, and de-escalation and gentle distraction when necessary. Antipsychotic medication is also indicated. Brain tumor patients are likely to be antipsychotic naive; they are also at very high risk of epilepsy. Neither

efficacy nor the risks of causing a dystonic reaction or epileptic seizure by prescribing any antipsychotic to brain tumor patients are conclusively known. Case studies in varying clinical populations have reported successful treatment with haloperidol, quetiapine (Seroquel), risperidone (Risperdal), olanzapine (Zyprexa), and aripiprazole (Abilify). If the patient is manic, a mood-stabilizing agent may also be required (e.g., sodium valproate, or lithium). If the patient is depressed, TCAs are best avoided: There is some evidence from older studies that these drugs may worsen symptoms. The successful use of ECT for steroid-induced psychotic depression has been reported.

The acutely disturbed brain tumor patient who presents an immediate danger to themselves or others should be managed according to local rapid tranquillisation protocols. Such extreme situations are rare. If an intramuscular injection is necessary to rapidly sedate the patient, close nursing observation is essential afterwards.

The patient should be reviewed regularly by the psychiatrist, since the neuro-oncology team may wish to restart steroids. It is probably sensible to seek to minimize the future doses of both corticosteroid and any ongoing antipsychotic as a starting point. Clinicians should also check that the patient's carers know how to contact services in the event of any future emergency.

Personality and Behavioral Change

INTRODUCTION. Until relatively recently, personality and behavioral change had some claim to be the "Cinderella" of brain tumor neuropsychiatry. It is one of the most distressing symptoms for families and carers. At a recent Society for Neuro-Oncology conference platform session, this topic was identified by carers as the most pressing area for future carer-focused research. Yet, evidence specific to brain tumor patients on how best to manage these problems has been lacking. This historical academic inertia may partly be because the concept itself is slippery: defining "personality change" is difficult. Studies make variable use of questionnaires loading on attitudinal change, behavioral change, and cognitive impairment. It is challenging to untangle in particular the influence of cognition—which is thoroughly affected by a brain tumor—on personality. Probably the most reproducible and accessible component, in terms of family understanding, clinical intervention, and systematic study, is a clear focus on the identification and formulation of problem behaviors.

EPIDEMIOLOGY. Personality changes are surprisingly frequent: troublesome behavioral changes are reported across tumor subtypes, ages, and stages of disease. Studies conducted in various primary brain tumor populations report a point prevalence ranging from 16% in patients with low-grade oligodendroglioma during follow-up to over 70% in patients with

primary CNS lymphoma (in whom, interestingly, behavioral change may be the most common presenting symptom). Although only a minority of patients present with clear behavioral symptoms from the outset, it probably becomes more common as the disease progresses. Behavioral symptoms are highly prevalent in the palliative stage of glioma. There is no evidence in adult patients that men and women are affected differently.

ETIOLOGY. Lesion location (especially when the tumor is in the frontal lobe) is often assumed to be the "cause" of personality or behavioral change. As discussed above, there is plainly a relationship of sorts, but it may be tempting to overextrapolate from neurology where destruction of a critical circuit or nucleus causes predictable neurological signs. When predicting wider psychiatric syndromes, the relationship is not so clear: the recent literature, which has the benefit of using sophisticated neuroimaging and more up-to-date nosologies to link tumor location with behavioral phenotype, unfortunately consists mainly of case reports. Accordingly, the etiology of personality and behavioral change in brain tumor patients remains obscure in any sort of detailed respect: in all likelihood, it is multifactorial, encompassing biological, psychological, social, and iatrogenic causes. The balance of factors will be different in different individuals.

DIAGNOSIS AND CLINICAL FEATURES. There is general consensus between clinicians and patient/carer representative groups that assessment of behavioral change should be person-centered and holistic, encompassing behavior, cognitive functioning, and emotional state. Assessment by both neuropsychiatry and neuropsychology is recommended. Reversible causes should be excluded or addressed. These include depression, anxiety, delirium, epilepsy, hypoxia, hypoglycemia, pain, confusion, and frustration at physical or communication deficits. Brain tumor medication may also contribute to unwanted behavioral symptoms, including a disinhibition syndrome secondary to corticosteroids or AEDs. The AED levetiracetam (Keppra) seems especially prone to an association with increased irritability. However, these may also be essential medicines for control of brain edema and epilepsy. As always, full discussion with the treating team and a collaborative management approach is vital.

Clinical diagnosis centers around identifying and characterizing the main behavioral syndrome. It is essential to take and give considerable weight to a collateral history from a reliable carer who knows the patient well. The problems most likely to be reported are explosive anger, apathy, emotional lability, and/or disinhibited inappropriate behavior. The assessing psychiatrist should establish the patient's premorbid personality and behavioral characteristics and confirm that the onset of the troublesome behavior was, at the earliest, in the months leading up to

initial presentation of the brain tumor. The severity of impact (on the patient or his immediate environment including carers) should at least be moderate. Day-to-day persistence of symptoms may be variable given the intermittent and often highly contextual nature of behavioral dysregulation. In the long term, however, there should be a clear pattern of new, harmful behavior. An assessment of exacerbating and alleviating factors is helpful therapeutically but probably not critical to the diagnosis, unless it identifies an effective reversible cause.

DSM-5 does not easily accommodate coding diagnoses to discriminate between clinically different syndromes of tumor-associated personality or behavioral change. The closest "catch-all" approach may be to first specify the etiological medical code for the type of brain tumor in question and then code for Major Neurocognitive Disorder Due to Another Medical Condition. Here the fifth digit specifies the presence of behavioral disturbance, which in the immediate context is assumed. This diagnosis is not ideal if cognitive function is relatively preserved. Additionally, different kinds of behavioral disturbances (apathy, disinhibition, etc.), which presumably have different biological mechanisms, will be grouped together by this approach.

DIFFERENTIAL DIAGNOSIS. The psychiatric differential includes MDD and generalized anxiety disorder. The psychiatrist should also consider reversible causes noted above, with particular respect to the possibility of delirium, and the possible role of any medication changes contemporaneous with onset of the troublesome behavior.

PATHOLOGY AND LABORATORY INVESTIGATION. Physical investigations should be targeted at reversible medical causes. They may include as appropriate: standard hematological and biochemical screens for causes of delirium; pulse oximeter reading to rule out frank hypoxia (DVT/PE is very common in these patients); blood glucose measurement to exclude steroid-associated diabetes and consequent hypoglycemic episodes; chest x-ray to exclude chest infection or DVT/PE; and a repeat head CT or MRI to exclude tumor progression or intratumoral bleed.

COURSE AND PROGNOSIS. There is limited observational evidence of the longitudinal evolution of personality changes in adult brain tumor patients. Anecdotally, once established these symptoms tend to persist and are difficult to treat, perhaps unsurprisingly in a progressive disease. The patient's family can learn effective behavioral management strategies to mitigate the causes and consequences of the patient's symptoms. Psychiatry and psychology services may helpfully contribute by teaching and reinforcing these strategies, prescribing medication if necessary to manage any unacceptable risks to the patient or carers, and signposting carers to trusted support services.

TREATMENT. Despite the considerable impact behavioral change can have on families and carers of brain tumor patients, few RCTs have enrolled them to study interventions for behavioral syndromes. A home-based psychosocial intervention (the "Making Sense of Brain Tumor" program [see Table 2.3–2 for a URL to the manual]) combines neuropsychological assessment, psychological therapy, and a person-centered approach to treatment. A small single-center randomized control study has recently shown some benefit from this program. In general though, there is a lack of specific psychological therapy resources tailored to problematic behaviors in brain tumor patients. Clinical management is therefore mostly based on common sense, discussion with the patient and family, and a pragmatic behavioral approach. Carer education and support is essential. Many situations can be managed acceptably by treating reversible causes, educating the patient and carer, and using behavioral strategies.

In terms of medication, there may be some benefit from treatment trials for psychomotor retardation, abulia, and apathy which (case studies suggest) may respond to dopamine agonists such as bromocriptine (Parlodel), or stimulants such as methylphenidate or modafinil. In the more overtly risky settings of severe uncontrolled disinhibition and explosive anger, options include treatment of disinhibition with lithium, carbamazepine (e.g., Tegretol), oxcarbazepine (e.g., Trileptal), or valproic acid. Explosive anger might respond to phenytoin (e.g., Dilantin), atypical antipsychotics, β-blockers, or short-acting benzodiazepines. Risperidone has been reported as of some effect for these problems in a case study set in palliative care. It is advisable to weigh the risks and benefits of treatment, and then conduct systematic empirical therapeutic trials until medications have been optimized for the individual in question. Behavioral agitation associated with the terminal stages of disease can be complex to manage and palliative care / hospice services should be involved.

A small number of RCTs have examined cognitive rehabilitation strategies for tumor- and treatment-associated cognitive dysfunction. Though these trials were not aimed specifically at the problem of personality change, cognitive changes may often affect behavior. Cognitive rehabilitation techniques might therefore be of use in treating personality or behavioral change in these patients. They can be delivered successfully both in the early postoperative and the later outpatient phases of tumor treatment. In general, studies show that these treatments lead to significant improvements, but usually in narrow psychometric parameters (mainly attention and visual memory). It is unclear whether cognitive retraining is effective for more complex social behaviors underpinned by executive function and impulse control.

Similarly, a few RCTs have tested the efficacy of various medications for either the treatment, or prevention, of cognitive changes in brain tumor patients. Methylphenidate, modafinil and donepezil (Aricept) have shown

some efficacy in small phase II RCTs. These studies recruited general samples of brain tumor patients rather than targeting those with frank behavioral disturbances. They showed modest improvements in relatively narrow neuropsychological parameters, again mostly attention and some aspects of memory. The clinical utility of these treatments has not been studied in large phase III studies with "real-world" behavioral outcomes. Psychiatrists may consider therapeutic trials as appropriate in individual patients.

Childhood Brain Tumor Survivors

INTRODUCTION. As can be seen from the general epidemiology of brain tumors, childhood tumors are different diseases from the common adult tumors. Myriad biological differences translate in some cases to differences in prognosis and in many children, even if the tumor is not actually cured by surgical resection, survival is often measured in decades. Increasingly therefore and as treatment strategies for childhood tumors continue to evolve, patients will present to adult psychiatric services as survivors of childhood brain tumors. They can display a uniquely challenging constellation of symptoms and disabilities, at the crux of which is the long-term cognitive effects of radiotherapy and/or chemotherapy on the vulnerable, developing brain.

EPIDEMIOLOGY. Most childhood brain tumors are infratentorial but the clinical problem in adult survivors is often dysexecutive. This can be traced back to the time of operation: compared to controls and in the year after diagnosis of a childhood BT, patients are consistently shown to have significantly lower performance IQ, processing speed, memory, and attention. These neurocognitive difficulties are especially likely in patients treated at a younger age, for malignant tumors, with a higher radiotherapy dose, or with chemotherapy. Radiation dose to the hippocampus may be particularly important. Hydrocephalus and the need for a CSF shunt are also associated with poorer functioning. Adult patients with these factors in their history may therefore be at higher risk.

Neuropsychiatrists working in specialized pediatric neuro-oncology centers may also occasionally encounter pediatric patients with an acute postsurgical posterior fossa syndrome (PFS). This is discussed as relevant below.

ETIOLOGY. Several possible mechanisms may explain persisting neurocognitive impairment in childhood survivors. For over a decade, it has been known that radiotherapy to the brain adversely affects neural stem cell proliferation and neurogenesis in the rodent hippocampal subgranular zone. The precise functional contribution of postnatal neurogenesis to human cognition is not yet clear. However, the effect of

radiotherapy is probably functionally relevant in humans too, given the well-recognized role of the hippocampus in modulating learning, memory, and emotion. A second population of particularly vulnerable cells may be oligodendrocyte precursor cells (OPCs). These cells generate oligodendrocytes which ensheath neurons in myelin. Accordingly, they support the biological integration and normal functioning of white matter tracts, which may be critical to cognition. Animal and human studies suggest that OPCs are vulnerable to the effects of certain types of chemotherapy: intriguingly there may be a differential and more pronounced adverse effect on OPCs in white relative to gray matter.

Genetic studies provide some preliminary evidence that children with a history of cranial radiotherapy and COMT gene heterozygosity (Val/Met) may have better working memory than patients with COMT gene homozygosity (Met/Met or Val/Val). Structural candidates under study include supratentorial white matter tract changes: loss of volume and atrophy of cortical and subcortical structures correlating to poorer functioning. Meanwhile, functional MEG data suggest that loss of high-frequency resting-state and task-oriented gamma oscillations may be associated with greater cognitive impairment and slower processing speeds in children undergoing radiotherapy. The clinical utility of these and other recent discoveries remains to be shown.

In acute pediatric neuro-oncology meanwhile, the PFS has been associated with abnormal functional neuroimaging in mute postoperative patients. These abnormalities affect multiple supratentorial sites during the acute episode. SPECT abnormalities appear to resolve in tandem with mutism. This observation has led to speculation that the cause of PFS may be a cerebellar–cerebral diaschisis. Quite how cerebellar damage may cause vascular or neuronal functional changes in the higher centers is a fascinating question for future study.

DIAGNOSIS AND CLINICAL FEATURES. Directions for diagnosing common individual psychiatric syndromes can be found in the appropriate sections. In brain tumor survivors, psychiatrists should also consider the possible contribution of subtle impairments which can persist many years after treatment of the tumor. Neuropsychologically speaking, these are often dysexecutive: working memory, attention, and processing speed are commonly lower than normal levels for the patient's age and developmental stage. These deficits often become more apparent with age. The working model in the field is that the *rate* of development of cognitive function, though constant after treatment, is reduced. The "slope" of attainment is altered. The older the patient, the more likely they are to fall below normal levels relative to peers (Fig. 2.3–4).

The PFS meanwhile consists of variable cognitive, emotional, and behavioral disturbances characteristically presenting in the days following

cerebellar surgery. Mutism can be present to varying degrees (milder cases may be dysarthric) but is often complete. Many patients may have shown similar, if less marked, symptomatology preoperatively. Whether such preoperative abnormalities might indicate a higher risk of developing the postsurgical syndrome is currently unknown. The PFS typically moves from complete mutism, through a period of cerebellar dysarthria, to resolution. The few case studies or small series charting its recovery suggest that this period can often measured in weeks, rather than days or months. Even after 1 year however, many children who have recovered from PFS may remain cognitively impaired in multiple domains. Whether the PFS is a risk factor for even greater cognitive impairment than normal for adult survivors is currently unclear.

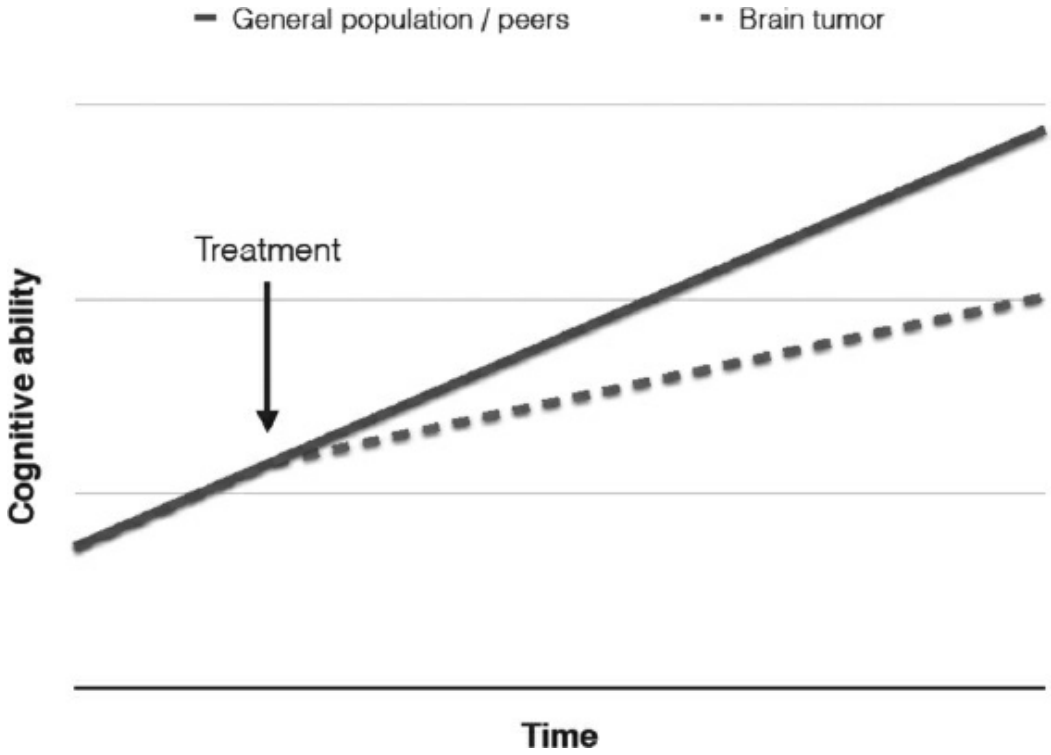

FIGURE 2.3–4. Reduced rate of acquisition of cognitive skills.

DIFFERENTIAL DIAGNOSIS. In childhood brain tumor survivors, the psychiatric differential will depend on the particular syndrome presented by the individual patient. Considering a general syndrome of subtle cognitive dysfunction in a young adult, treatable conditions to be variously excluded include MDD, anxiety, and alcohol or drug use, or dependence. It may help to pinpoint and quantify specific cognitive deficits through formal neuropsychological assessment. Medical causes to exclude will again depend on the specific presentation, but the general health of these patients is often poor. In particular, hypertension, diabetes, and endocrine abnormalities secondary to radiation-induced pituitary dysfunction should be excluded. Some forms of chemotherapy can cause deafness which may

contribute to psychosocial or emotional difficulties.

Children presenting acutely with an apparent PFS should receive an MRI brain scan to rule out postoperative bleeding or cerebral edema that might otherwise explain the clinical picture.

COURSE AND PROGNOSIS. The longer survivorship of childhood brain tumor patients lends itself to longitudinal study. The literature on course and prognosis is richer than for adults, both in the number of studies and a broader emphasis on psychosocial outcomes. There is high agreement that childhood brain tumor survivors tend to have considerable impairments of cognitive and emotional functioning. These often manifest as social, educational, and vocational disability.

In childhood, executive impairments rooted in the primary treatment of the tumor often persist. Attention, working memory, and processing speed may all remain affected, often subsyndromally. Educational achievement is often compromised. Reading, spelling, and arithmetic may all be affected. Impaired emotional functioning with increased anxiety is seen in some children, with some evidence of defective cognitive processing of emotions. Behavioral problems may manifest at school, including aggression, and possibly a higher level of somatization. Peer relationships suffer accordingly and adaptive and developmentally important behaviors such as communication and socialization are often adversely affected. Many patients may report cognitive and behavioral difficulties that do not meet diagnostic criteria suggesting a large hidden morbidity that current psychiatric classification systems may not reflect.

In late adolescence and adulthood, the general picture is similar. Medulloblastoma survivors are highly likely to need some form of accommodation for acquired impairments in learning skills. The physical health of this group can be strikingly poor and even barely out of their teens, hearing impairment, secondary cancers, diabetes, hypertension, and endocrinopathies may be common. Most of these patients are likely to remain dependent on parental income into their fourth decade. Fully one-third of adult survivors of posterior fossa tumors display clinically significant levels of apathy. Females may be at greater risk, and interestingly, apathy again seems to be a disorder distinct from clinical depression. Themes of impaired identity formation, social isolation, difficulty securing financial independence, and limited career options are common in young adult survivors generally. Completion of higher education is relatively rare. Unemployment is often the norm, although by no means universal and if there is a glimmer of silver lining it is that despite an often-formidable array of psychosocial disadvantages, most adult patients will be functionally independent in daily life.

TREATMENT. Treating these chronic, biologically mediated, multimodal

impairments is naturally difficult. A person-centered, problem-focused, holistic, and realistic approach is recommended. In adult survivors, neuropsychology referral should be made where possible and treatment plans ideally developed in a collaboration of psychology, psychiatry, the patient, and if relevant, their carers. Other professions including occupational therapy, physiotherapy, and dietetics may be involved as appropriate. If the neuro-oncology team is still involved in following patients up they should be consulted: they will know the patient better than nearly any other professional service. Recommendations and adaptations should be tractable and feasible to implement.

Cognitive rehabilitation—structured exercises to boost cognitive skills—has some evidence for efficacy in childhood survivors. Improved executive and memory skills are seen particularly in the domains of attention, processing speed, cognitive flexibility, and declarative memory. The "real-life" impact of these interventions is largely unknown because the literature is mainly of small unblinded pilot studies that do not show an impact on educational and social skills. Properly powered RCTs will be valuable. Whether exercise programs have a beneficial effect, and if so in what domains, is another interesting nonpharmacological question that is being addressed by ongoing studies.

Medication can also be considered as a treatment trial for troublesome neurocognitive impairment in survivors. There is some evidence that the psychostimulant methylphenidate improves attention and may improve behavior in the short term in childhood survivors of brain tumor. Boys, those of older age, and a higher baseline IQ may respond better. It has not been consistently shown to improve educational performance. Donepezil has limited evidence of efficacy for executive and memory impairments. RCTs are going on for modafinil. It is important to stress that evidence underpinning these and other drug treatments is, again, limited in this situation. As illustrated by Figure 2.3–4, to actively return adult patients closer to population norms, drug and nondrug treatments alike must cause a sustained and relative gain in the *rate* of acquisition of cognitive skills – a challenging ask. In fully informed patients, systematic therapeutic trials may certainly form part of a holistic treatment package though. Given their empirical nature, treatment trials should be time-limited, reviewed regularly, and discontinued if unsuccessful. Functional psychiatric syndromes arising in adult survivors, meanwhile, should be treated as outlined above.

As well as treating established dysfunction researchers are exploring strategies that may help minimize or prevent neurological damage at the time of primary treatment. In particular, hippocampal-sparing radiotherapy schedules have recently been associated with better preservation of cognitive functioning, in both pediatric and adult brain tumor samples. A clear understanding of the clinical scope and limitations

of hippocampal-sparing techniques will require further study, but they hold promise for the future.

Treatment of the PFS, meanwhile, is mostly supportive. In cases that do not recover as expected, a trial of the hypnotic Zolpidem (e.g., Ambien) may be of value. Families should be informed of the often gradual process of recovery. Where appropriate they should be supported in making arrangements for the eventual return of their child to school. Given the considerable residual impact that the condition can have on various cognitive skills at a critical period for educational and social development, detailed neuropsychological testing is again central to this process.

FUTURE DIRECTIONS

Clinical

Awareness is increasing of the potential for benefit of multidisciplinary neurorehabilitation for adult and pediatric brain tumor patients. Among the wide range of patients with differing types of primary brain tumor, those who could especially benefit from neurorehabilitation may be patients with persisting disability but a medium- to long-term expectancy of survival. At least in theory this would indicate patients with a low-grade or anaplastic (high-grade but not the most malignant) glioma, or those surviving a childhood brain tumor of nearly any sort.

The evidence that multidisciplinary interventions actually work is mostly observational and therefore at a high risk of bias. Available studies tentatively suggest that multidisciplinary rehabilitation for brain tumors can be effective, particularly for improving some functional outcomes, vocational skills, and quality of life. Evidence for their effectiveness in preventing or alleviating neuropsychiatric complications is lacking. What this situation mainly illustrates is the difficulty of designing trials that provide integrated care to such a complex population.

Here too therefore, a strong case can be made that psychiatrists have a critical role to play. Psychiatrists have sophisticated transferable professional skills that are directly relevant to both neuro-oncology and rehabilitation. For example, experience working in a multidisciplinary environment; the mindset, technical abilities, and time required to conduct comprehensive and holistic person-centered assessments; and knowledge of how to manage chronic and complex illnesses affecting vulnerable individuals. They also have direct expertise in the assessment and management of the many cognitive and psychiatric problems which, as outlined, very commonly affect these patients.

Despite a strong case for inclusion of psychiatrists in the rehabilitation of brain tumor patients, the current clinical reality is that in most centers, neuro-oncology, neurorehabilitation, and neuropsychiatric services remain separate entities. There are exceptions but mostly these systemic divisions

are a natural consequence of separate service evolution. They create practical barriers to clinical integration. Resistance to the idea that brain cancer patients can be meaningfully rehabilitated is common, if almost certainly wrong. Having a psychiatrist in the multidisciplinary neuro-oncology team is not mandated anywhere in the world.

One consequence is that psychiatric diagnoses are missed in surviving patients who are at very high risk of psychiatric morbidity. They are missed for many reasons. Skilled and caring clinicians nonetheless may think something like "I'd be depressed if I had a brain tumor"; patients themselves may be reluctant to "bother the doctor" with emotional symptoms; and the sheer pressure of time in a busy neuro-oncology clinic often mitigates against a thorough exploration of the psychic reality of a devastating diagnosis. One solution to these problems would be for psychiatrists themselves to recognize and make the compelling case for their skills in managing brain tumor patients. There is a clear need for psychiatry, both in established acute neuro-oncology, and in emerging rehabilitation neuro-oncology teams.

Research

Doctors diagnose, prognose, and treat. As noted at the beginning, the very presence of a brain tumor distorting and destroying critical cerebral pathways poses questions for traditional conceptualizations of some kinds of mental disorder. Part of the challenge ahead therefore lies in testing and, where necessary, refining diagnostic criteria. Large, longitudinal observational studies will be needed and will inform better prognostication. The relative rarity of brain tumors will promote multicenter collaborations. Research questions may be shaped by advances in the current understanding of the nature of mental disorder that arise from other fields and studies in other model systems. In particular, mechanistic insights will be vital.

Serendipitous discoveries apart, it is a general principle of science that thorough understanding of biological mechanisms also favors the development of novel and rational treatments for medical illness. Such understanding is needed of the mechanisms that contribute to the risk, onset, maintenance, and resolution of brain tumor-associated neuropsychiatric disorders. Because of the unique clinical context, these mechanisms should be studied in brain tumor patients, where possible, in order to tailor treatments for maximal effect in this population. Thus in rationalizing conceptions of illness, prognostication, and in developing new treatments, it will be critical to understand the biological mechanisms that underpin the neuropsychiatry of brain tumors.

References

Alper K, Schwartz KA, Kolts RL, Khan A. Seizure incidence in psychopharmacological clinical trials: an analysis of Food and Drug Administration (FDA) Summary Basis of Approval reports. *Biol Psychiatry*. 2007;62:345–354.

Arnold SD, Forman LM, Brigidi BD, et al. Evaluation and characterisation of generalised anxiety and depression in patients with primary brain tumors. *Neuro Oncol*. 2008;10:171–181.

Benros ME, Laursen TM, Dalton SO, Mortensen PB. Psychiatric disorder as the first manifestation of cancer: a 10-year population-based study. *Int J Cancer*. 2009;124:2917–2922.

Carrera E, Tononi G. Diaschisis: past, present, future in patients with malignant glioma. *Brain*. 2014;137:2408–2422.

Carroll C, Watson P, Spoudeas HA, et al. Prevalence, associations, and predictors of apathy in adult survivors of infantile (<5 years of age) posterior fossa brain tumors. *Neuro Oncol*. 2013;15:497–505.

Catsman-Berrevoets CE, Aarsen FK. The spectrum on neurobehavioural deficits in the Posterior Fossa Syndrome in children after cerebellar tumour surgery. *Cortex*. 2010;46:933–946.

Caudill JS, Brown PD, Cerhan JH, Rumens TA. Selective serotonin reuptake inhibitors, glioblastoma multiforme, and impact on toxicities and overall survival: the Mayo clinic experience. *Am J Clin Oncol*. 2011;34:385–387.

de Smet HJ, Baillieux H, Wackenier P, et al. Long-term cognitive deficits following posterior fossa tumor resection: a neuropsychological and functional neuroimaging follow-up study. *Neuropsychology*. 2009;23:694–704.

Dockstader C, Wang F, Bouffet E, Mabbot DJ. Gamma deficits as a neural signature of cognitive impairment in children treated for brain tumors. *J Neurosci*. 2014;34:8813–8824.

Fang F, Fall K, Mittleman MA, et al. Suicide and cardiovascular death after a cancer diagnosis. *New Engl J Med*. 2012;366:1310–1318.

Ford E, Catt S, Chalmers C, Fallowfield L. Systematic review of supportive care needs in patients with primary malignant brain tumors. *Neuro Oncol*. 2012;14:392–404.

Gehring K, Sitskoorn MM, Gundy CM, et al. Cognitive rehabilitation in patients with gliomas: a randomized, controlled trial. *J Clin Oncol*. 2009;27:3712.

Gong S, Sheng P, Jin H, et al. Effect of methylphenidate in patients with cancer-related fatigue: a systematic review and meta-analysis. *PLOS One*. 2014;9:e84391.

Howarth RA, Adamson AM, Ashford JM, et al. Investigating the relationship between COMT polymorphisms and working memory performance among childhood brain tumor survivors. *Pediatr Blood Cancer*. 2014;61:40–45.

Khan F, Amatya B, Ng L, Drummond K, Olver J. Multidisciplinary rehabilitation after primary brain tumour treatment. *Cochrane Database Syst Rev*. 2013, Issue 1. Art. No.: CD009509.

Lewis DA, Smith RE. Steroid-induced psychiatric syndromes. A report of 14 cases and a review of the literature. *J Affect Disord*. 1983;5:319–332.

Lucas MR. What brain tumor patients and their families have taught me. *J Neurosci Nurs*. 2013;45:171–175.

*Monje M, Mizumatsu S, Fike JR, Palmer TD. Irradiation induces neural precursor-cell dysfunction. *Nat Med*. 2002;8:955–962.

Nazemi KJ, Butler RW. Neuropsychological rehabilitation for survivors of childhood and adolescent brain tumors: a view of the past and a vision for a promising future. *J Pediatr*

Rehabil Med. 2011;4:37–46.

Omuro A, DeAngelis LM. Glioblastoma and other malignant gliomas: a clinical review. *JAMA*. 2013;310:1842–1850.

*Ostrom QT, Gittleman H, Liao P, et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2007–2011. *Neuro Oncol*. 2014;16:iv1.

*Ownsworth T, Chambers S, Damborg E, Casey L, Walker DG, Shum DH. Evaluation of the making sense of brain tumor program: a randomized controlled trial of a home-based psychosocial intervention. *Psychooncology*. 2014; (currently epub only) doi: 10.1002/pon.3687.

*Padovani L, André N, Constine LS, Muracciole X. Neurocognitive function after radiotherapy for paediatric brain tumours. *Nat Rev Neurol*. 2012;8:578–588.

Palmer SL, Armstrong C, Onar-Thomas A, et al. Processing speed, attention, and working memory after treatment for medulloblastoma: an international, prospective, and longitudinal study. *J Clin Oncol*. 2013;31:3494–3500.

Palmer SL, Hassall T, Evankovich K, et al. Neurocognitive outcome 12 months following cerebellar mutism syndrome in pediatric patients with medulloblastoma. *Neuro Oncol*. 2010;12:1311–1377.

Pitsika M, Tsitouras V. Cerebellar mutism. *J Neurosurg Pediatr*. 2013;12:601.

*Rooney AG, Brown PD, Reijneveld JC, Grant R. Depression in glioma: a primer for clinicians and researchers. *J Neurol Neurosurg Psychiatry*. 2014;85:230–235.

Rooney AG, McNamara S, Mackinnon M, et al. Frequency, clinical associations, and longitudinal course of Major Depressive Disorder in adults with cerebral glioma. *J Clin Oncol*. 2011;29:4307–4312.

Smithson EF, Phillips R, Harvey DW, Morrall MC. The use of stimulant medication to improve neurocognitive and learning outcomes in children diagnosed with brain tumours: a systematic review. *Eur J Cancer*. 2013;49:3029–3040.

Starkstein SE, Boston JD, Robinson RG. Mania after brain injury: a controlled study of causative factors. *Arch Neurol*. 1987;44:1069–1073.

▲ 2.4 Neuropsychiatric Aspects of Epilepsy

PONGSATORN PAHOLPAK, M.D., JOANNA JUAN GAN, M.D., AND MARIO F. MENDEZ, M.D., PH.D.

Clinicians have recognized the association of epilepsy with psychiatric disorders since antiquity. In modern times, this relationship has often been poorly recognized and inadequately investigated. Yet, the development of new antiepileptic and psychiatric therapies and novel neuroimaging techniques makes understanding the association of epileptic seizures and psychopathology increasingly important. Most recently, many psychiatrists and neurologists have taken up the investigation of neuropsychiatric aspects of epilepsy, as exemplified by an increase in research and publications in this field.

Epidemiological data support an increased risk for psychiatric comorbidity among epilepsy patients as compared to nonepileptic patients. The most established association is between epilepsy and depression or dysthymia, but a range of psychopathology occurs in one-fourth or more of patients. For example, epilepsy is associated with anxiety disorders, somatic symptom disorders, personality changes, hyposexuality, and, perhaps most dramatically, several forms of psychosis. These behaviors and others have different potential relationships to the ictus or seizure itself. Psychiatric manifestations may result directly from ictal discharges (e.g., psychic auras), as peri-ictal phenomena (e.g., postictal confusion), interictally or between seizures (e.g., interictal psychosis), or with a variable and less-established relationship to the seizure discharges (e.g., most mood disorders).

Psychiatric comorbidity has a serious impact on the quality of life and well-being of patients with epilepsy. Psychiatrists and neurologists need to maximize the mood-stabilizing and other psychotropic effects of antiepileptic drugs, consider the seizure threshold lowering effects of some psychotropic medications, and monitor the potential interaction of antiepileptic and psychotropic drugs. Before committing patients to antiepileptic treatment, psychiatrists and neurologists must also be able to distinguish epileptic seizures from other spells, particularly somatic symptoms such as nonepileptic seizures.

DEFINITION

Epileptic *seizures* are sudden, involuntary behavioral events associated with excessive or hypersynchronous electrical discharges in the brain. The seizure itself is known as the *ictus*. The *interictal period* refers to the period between the postictal abnormalities and the next ictus, and the *peri-ictal period* refers to the period just before or after the ictus and is applied when

there is insufficient information to know when the ictus ends or begins. Epileptic seizure can have a genetic basis or can be caused by an acquired insult such as trauma, stroke, inflammation, or neoplasm. *Epilepsy* is the recurrent tendency to seize, and *status epilepticus* is prolonged or repetitive seizures without intervening recovery.

In epilepsy, abnormal electrical discharges are due to hyperexcitable neurons with sustained postsynaptic depolarization. Proposed mechanisms for this sustained depolarization include changes in ionic conductance, decreased γ-aminobutyric acid (GABA) inhibition of cortical excitability, and increased glutamate-mediated cortical excitation. In animals, alumina-induced membrane changes alter the ratio of intracellular to extracellular ionic concentrations and result in abnormal neuronal firing. Antiepileptic drugs, such as phenytoin (Dilantin), carbamazepine, and valproate, reduce this repetitive firing through effects on sodium channels. Ethosuximide (Zarontin) works through blockage of calcium currents. Penicillin-induced cortical injury causes seizures through decreased GABA inhibition. Barbiturates and benzodiazepines may reduce seizures by enhancing GABA receptor current and valproate through blockage of GABA catabolism. Kainic acid, a glutamate agonist, induces seizures through increased synaptic action at its *N*-methyl-D-aspartate (NMDA) receptors. Much work is underway on potential antiepileptic drugs that may work through inhibition of this excitatory receptor mechanism.

The electroencephalogram (EEG) is a surface recording of brain wave activity used in the evaluation of seizures. Basic waves include normal waking alpha waves (8 to 13 Hz), which are most prominent over the occipital region, high-frequency beta waves (greater than 13 Hz), and theta waves (4.0 to 7.5 Hz) and delta slowing (3.5 Hz or less). Seizures are manifest as multiple spikes or spike and wave discharges on the EEG (Fig. 2.4–1). A spike is a sharp transient with a duration of 20 to 70 milliseconds. Interictally, single spikes and other markers of abnormal electrical activity may be seen, often emanating from a temporal lobe.

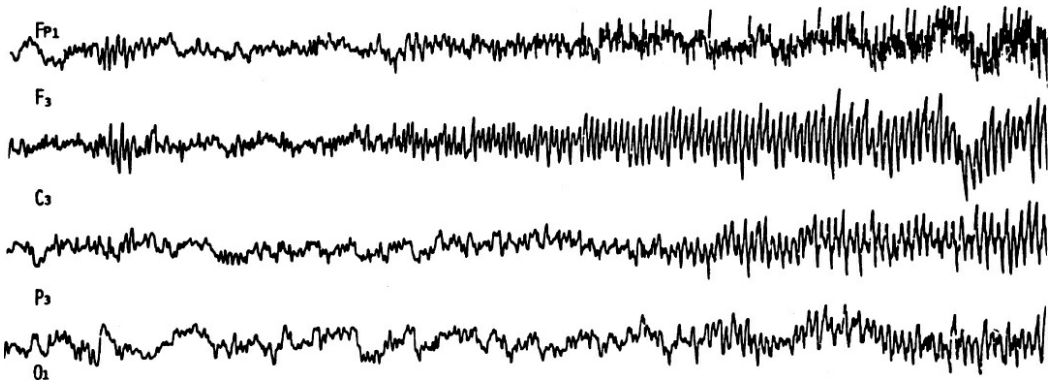

FIGURE 2.4–1. Electroencephalogram demonstrating the focal onset of seizure discharges from the left frontotemporal region that are consistent with the onset of focal dyscognitive seizures.

HISTORY

In his book on epilepsy, *On the Sacred Disease,* Hippocrates (460 to 377 BC) attacked the prevailing belief that those afflicted with epilepsy were possessed by gods or goddesses. He proposed that epilepsy was a brain disease caused by the blockage by phlegm of air-carrying vessels to the brain. Despite this initial view, throughout most of human history, epilepsy constituted demonic possession or the accumulation of bad humors, and attempts at exorcism involved trephination, cautery of the back of the skull, diuretics, emetics, bloodletting, purging, sweating, and even intercourse to release sperm. In the 18th century, the first so-called scientific treatise on epilepsy since ancient times attributed seizures to masturbation. By happenstance, bromides, which were introduced to diminish libido and masturbation, proved to be the first successful antiepileptic medication. With the development of effective antiepileptic drugs and the introduction of EEG, physicians have come full circle to Hippocrates' belief that epilepsy is rooted in organic brain disease.

The purported association of epilepsy with behavioral disorders also dates to antiquity. The brain was the seat of the falling sickness and madness, and both were related to phlegm. With demonic possession as a form of punishment, unusual or abnormal behaviors became associated with seizure patients, even during their seizure-free periods. At the turn of the 19th century, the psychiatric writings of Emil Kraepelin emphasized that epileptic patients possessed personality changes and a predisposition to psychosis. With the greater understanding of the physical basis of epilepsy, many clinicians sought to protect epileptic patients from the demonic stigma of their disease; in their view, psychiatric problems resulted from the psychosocial difficulties associated with having seizures rather than any unique relationship of epilepsy with psychiatric illness. The current age was initiated by the definition of temporal lobe epilepsy and the concept of a physiological disturbance in the limbic or emotional brain.

NOSOLOGY

The classification and terminology of seizure was revised by the International League Against Epilepsy in 2010 to reflect the advances in the understanding of epilepsy. Seizures are classified based on mode of etiology and mode of onset. Table 2.4–1 reflects changes in classification terminology. Etiology can be grouped into genetic, structural–metabolic, and unknown. Mode of onset can be *generalized* or *focal.* Generalized onset seizures are those with an initial widespread bihemispheric involvement, and focal onset seizures are those that emanate from a focus limited to part of one hemisphere. In adults, most generalized seizures are *tonic–clonic* seizures (grand mal seizures or convulsions) and are characterized by an abrupt loss of consciousness with tonic rigidity

followed by a synchronous, clonic release. Focal seizures (previously partial or localized seizures) can occur with or without impairment of consciousness or awareness (Table 2.4–1). Simple partial seizure is another term for focal onset without impairment of consciousness, whereas focal dyscognitive seizure is synonymous with focal onset with impairment of consciousness. Focal seizures produce isolated motor, sensory, autonomic, psychic, or mixed symptoms in a clear sensorium. Focal seizures that involve subjective sensory or psychic phenomena without observed motor or autonomic changes are considered *auras*. Focal dyscognitive seizures are usually characterized by motionless staring combined with simple automatisms, or automatic motor activity, and last approximately for 1 minute. Focal dyscognitive seizures can evolve to become a tonic, clonic, or tonic–clonic seizure. Finally, there is a form of generalized onset seizures, *absence* (petit mal) seizures, which occur less commonly in adults and are characterized by brief lapses of consciousness. Absence seizures can be differentiated from focal dyscognitive seizures in that they are generalized in mode of onset. They also differ in that they are short (10 seconds in length) and repetitive; lacking auras, postictal confusion, or complex automatisms; and in having characteristic generalized two to four counts per second spike and wave discharges on EEG.

Table 2.4–1.
2010 ILAE Classification of Epileptic Seizures

Focal onset seizures
Without impairment of consciousness or awareness[a]
Involving motor, autonomic, sensory, or psychic phenomena (latter two correspond with auras)
With impairment of consciousness or awareness[b]
Evolving to a bilateral, convulsive seizure (tonic, clonic, or tonic–clonic)[c]
Generalized onset seizures (convulsive or nonconvulsive)
Tonic–clonic
Myoclonic (myoclonic, myoclonic–atonic, myoclonic–tonic)
Absence (typical, atypical, with special features)
Clonic
Tonic
Atonic
Unclear mode of onset
Epileptic spasms

[a]Previously, simple partial seizure.

[b]Previously, corresponds to “complex partial seizure” or “dyscognitive” seizure.

[c]Previously, partial seizure with secondary generalization.

EPIDEMIOLOGY

Seizure disorders are common and usually have an early onset. Epilepsy affects 20 to 40 million people worldwide and has a prevalence of at least

0.63 percent and an annual incidence of approximately 0.05 percent. The overall incidence is high in the first year, drops to a minimum in the third and fourth decades of life, then peaks after age 75 years. More than 75 percent of patients have their first seizure before 18 years of age, and 12 to 20 percent have a familial incidence of seizures. Among adults, the most common seizures are focal dyscognitive and generalized tonic–clonic seizures. The most common cause of new-onset seizures in adults is cerebrovascular disease.

Psychopathology

Epidemiological studies from communities, psychiatric hospitals, and epilepsy clinics report a 20 to 60 percent prevalence of psychiatric problems among epilepsy patients. Epilepsy patients are prone to psychosis; mood disorder, especially depression; anxiety disorder; somatic symptom disorders; personality disorders; hyposexuality; and dissociative symptoms. These problems are approximately equally divided between those that occur ictally or peri-ictally and those that occur interictally or are variably related to the ictus. The percentage of epilepsy patients in psychiatric hospitals was also higher than the general prevalence of epilepsy and ranged from 4.7 percent of all inpatients in a British psychiatric hospital to 9.7 percent in a US Veterans Affairs psychiatric facility. Among patients attending epilepsy clinics, approximately 30 percent had a prior psychiatric hospitalization, and 10 to 20 percent were on at least one psychotropic drug. Furthermore, epidemiological studies indicate an increased interictal psychopathology among head-injured patients with epilepsy compared to head-injured patients without epilepsy. There exist a bidirectional relationship between psychiatric disorders and epilepsy, such that patients with epilepsy are more likely to develop depression than the general population, and patients with depression, anxiety, and psychosis are more likely to develop epilepsy. Despite criticisms of selection bias, these studies constitute a broad spectrum of sources that indicate greater overall psychopathology in epilepsy patients.

Do epilepsy patients have greater psychopathology than other similarly impaired patients? If this were so, then it would suggest that the psychopathology is of biological origin rather than a less specific reaction to chronic disease. Although disputed by some investigators, several studies report more psychopathology among epileptic patients than among patients with chronic diseases that do not directly affect the brain. Furthermore, the pattern of behavioral changes in seizure patients appear specific to epilepsy. For example, on the Minnesota Multiphasic Personality Inventory (MMPI) 2, despite a lack of difference in overall psychopathology, patients with epilepsy have higher schizophrenia scale and paranoia scale scores than patients with other neurological disabilities.

Many studies found a special relationship to psychopathology in

patients whose seizures emanated from mediobasal temporal lesions. Psychiatric disturbances, primarily psychosis and personality disorders, are two to three times more common in patients with focal dyscognitive seizures, most of whom have a temporal focus, compared to those with generalized tonic–clonic seizures; other studies have failed to find a difference. Major depression, dissociative symptoms, obsession, and compulsion were also seen more frequently in temporal lobe epilepsy compared to generalized epilepsy. Nevertheless, 60 to 76 percent of adults with epilepsy, regardless of seizure type, have a temporal lobe focus, and many generalized tonic–clonic seizures are secondarily generalized from a temporal lobe focus without a preceding focal dyscognitive seizure. Moreover, psychic auras from the temporal lobe, particularly if associated with negative feelings (e.g., *jamais vu* and fear), predispose to psychosis or personality disorders.

Psychosis

Psychosis is the specific psychiatric disorder most clearly associated with epilepsy. The lifelong prevalence of all psychotic disorders among epileptic patients ranges from 7 to 12 percent. In a follow-up of 100 children with focal dyscognitive seizures for as long as 30 years, of the 87 patients who survived to adulthood and who did not have mental retardation, 9 (10 percent) experienced a psychotic illness. Moreover, in temporal lobectomy studies, in which there is surgical removal of an epileptic focus, psychosis occurred in 7 to 8 percent of patients, even long after the seizures were arrested. The data from a recent meta-analysis also indicated that a prevalence of psychosis in temporal lobe epilepsy was 7 percent. That percentage is higher than in the other types of epilepsy (5.6 percent) and represents approximately a five- to eight-fold greater risk of psychosis for epileptic patients than for the general population.

Autoantibodies may play a role in psychotic symptoms in epilepsy. Data from many studies found that specific autoantibodies, for example, anti-NMDA, anti–voltage-gated potassium channel (VGKC), and anti-GABA-B receptor ($GABA_B$) can be found in up to 10 percent of these patients. Moreover, those patients who have autoimmune encephalitides usually develop convulsive seizure during the course of illness.

Regarding lateralization and neuropathology, results from many studies suggest an association of a left-sided focus and hippocampal sclerosis with psychosis. Although conclusions derived from surface EEG recording are open to criticism, depth recordings of presurgical patients show that twice as many patients with left temporal lesions have psychosis. Data from study using voxel-based morphometry analysis of MRI also show that patients with temporal lobe epilepsy and psychosis have more severe grey matter loss in an entire temporal lobe as well as in the left hippocampus, insulae,

anterior cingulum, and left inferior parietal lobule. The results from white matter data further demonstrate white matter reduction in widespread brain regions. Finally, PET scans and single photon emission computed tomography (SPECT) scans may show predominant left temporal hypometabolism among epilepsy patients with psychosis.

Depression

The prevalence of depression in different studies varies from 7.5 to 34 percent of patients with epilepsy. This discrepancy is attributed to various diagnostic measurements and different definition of depressive disorders across studies. Major depressive episode can be found in up to 20 percent of the patients. Those with focal dyscognitive seizures originated from mesial temporal lobe and those with poor seizure control are more likely to have mood disorders. Psychological studies also suggest a greater incidence of self-criticism, and depression among epilepsy patients with a left hemisphere focus. Patients with focal dyscognitive seizures of temporal limbic origin also have a higher incidence of depression than patients with other types of seizure disorders.

Table 2.4–2.
Psychic Auras

Type	Symptoms	Probable Source
Dysphasic[a]	Nonfluent	Left perisylvian language areas
	Impaired comprehension	
Dysmnesic	*Déjà vu, déjà vécu, déjà pensé, déjà entendu, jamais vu,* etc., prescience, illusion of memory	Mesobasal temporal,[b] especially on right
Cognitive	Dreamy state, altered time sense, derealization, depersonalization	Mesobasal temporal and temporal neocortex
	Forced thinking, forced actions, and altered or obscure thoughts	Frontal association cortex
Affective	Fear, anxiety, apprehension, depression, pleasure, displeasure	Mesobasal temporal and temporal neocortex
Illusions[c]	Macropsia, micropsia, teleopsia, movement, metamorphopsia, increased color intensity, increased stereopsis intensity	Lateral superior temporal neocortex, especially on right for visual illusions
Hallucinations[c]	Structured, hallucinatory remembrances, autoscopy	Mesobasal temporal and temporal neocortex

[a]Does not include speech arrest or simple vocalizations.

[b]Includes hippocampus, amygdala, and the parahippocampal gyrus.

[c]Includes interpretive (size, motion, shape, and stereopsis) or experiential (elements of past experience or involvement).

Other Behaviors

The prevalence of other specific behavioral disorders among patients with

epilepsy is less well established. However, recent evidences consistently demonstrated that personality disorders, suicidal behavior, anxiety disorder, and hyposexuality are more prevalent among epilepsy patients than among those without seizure disorders.

ETIOLOGY

There is increasing movement among the epileptology community toward identification of etiology, as diagnosing the specific cause can have important implications for management. Previous instances of idiopathic epilepsy are increasingly found to have evidence of specific genetic defects. Other frequent causes include trauma in the third and fourth decades of life, neoplasms in the fifth and sixth decades of life, and cerebrovascular disease in the elderly. Although some focal dyscognitive seizures originate from the frontal or temporal neocortex and other areas, at least two-thirds of focal dyscognitive seizures and generalized tonic–clonic seizures originate from the mediobasal temporal limbic structures (hippocampus, amygdala, and parahippocampal gyrus).

Psychopathology

The relationship of seizures, psychiatric syndromes, and the mediobasal temporal lobes implies that many behavioral changes are more than psychological reactions to the psychosocial stressors of epilepsy. Stimulation and ablation studies in humans and animals link temporal limbic structures to emotional behavior.

For example, temporal limbic stimulation in a person evokes psychic auras (Table 2.4–2) and automatisms, and amygdala stimulation and ablation in animals result in aggression or placidity. Moreover, psychotic behavior in cats occurs when their limbic structures undergo kindling (that is, the repeated application of epileptic agents to induce lasting behavioral changes).

There are several potential organic causes of psychiatric disturbances in epilepsy (Table 2.4–3). First, the pathology itself could be the source of seizures and behavioral changes. Left hemisphere and temporal lobe lesions may be associated with interictal psychosis, and psychosis in epilepsy may be particularly frequent if there is specific underlying pathology or ventricular enlargement. Psychotic disorders may be more common with temporal dysplasia or neurodevelopmental abnormalities and depression with mesial temporal sclerosis. Second, ictal or subictal epileptiform activity may promote behavioral changes by facilitating distributed neuronal connections, increasing limbic–sensory associations, or changing the overall balance between excitation and inhibition. This may occur not only with temporal lobe seizures but also with those that originate in the frontal lobes. Third, the absence or hyperactivity of

function, such as the interictal hypometabolism observed on PET scans (Fig. 2.4–2), may lead to depression or other interictal behavioral change, as well as may have a reduction in cerebral blood flow in the left medial temporal region from SPECT may be associated with interictal psychosis. In contrast, auditory hallucination during a focal onset seizure without impairment of consciousness may demonstrate a cerebral hyperperfusion in primary auditory cortex. Fourth, seizures may result in neuroendocrine or neurotransmitter changes, such as increased dopaminergic or inhibitory transmitters, decreased prolactin, increased testosterone, or increased endogenous opioids, all of which can affect behavior. Furthermore, neurobiological factors may be potentiated by psychodynamic factors, such as feelings of helplessness, learned helplessness, dependency, low self-esteem, and the disruption of reality testing. In summary, the psychiatric manifestations of epilepsy are heterogeneous disorders with potentially different causes.

Table 2.4–3.
Proposed Relationships of Psychiatric Disturbances to Epilepsy

Common neuropathology, genetics, or developmental disturbance
Ictal or subictal discharges potentiate abnormal behavior
Kindling or facilitation of a distributed neuronal matrix
Changes in spike frequency or inhibitory–excitatory balance
Altered receptor sensitivity, e.g., dopamine receptors
Secondary epileptogenesis
Absence of function at the seizure focus
Inhibition and hypometabolism surrounding the focus
Release or abnormal activity of remaining neurons
Dysfunction or downregulation of associated areas
Neurochemical
Dopamine and other neurotransmitters
Endorphins
Gonadotropins and other endocrine hormones
Psychodynamic and psychosocial effects of living with epilepsy
Dependence, learned helplessness, low self-esteem, weak defense mechanisms
Disruption of reality testing
Neurobiological and psychodynamic factors potentiate each other
Sleep disturbance
Antiepileptic drug related

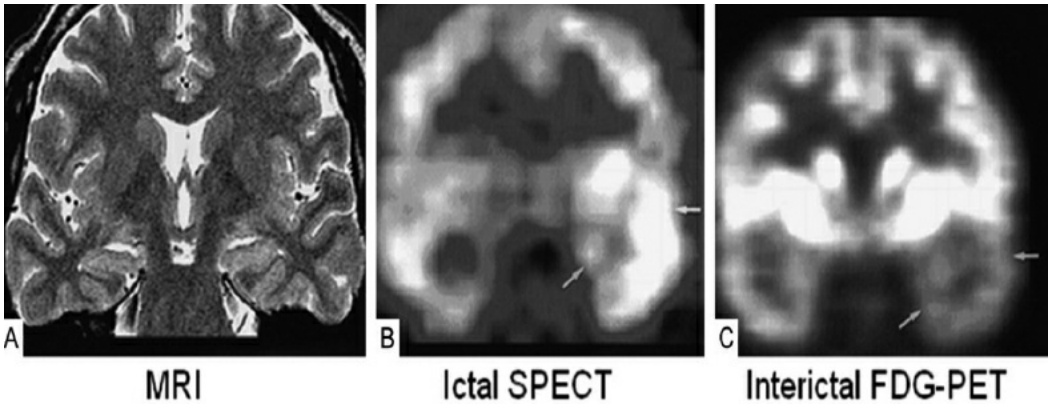

FIGURE 2.4–2. Neuroimaging from a 30-year-old man with left temporal origin dyscognitive seizures and ictal activity on invasive EEG. **A:** Coronal T2-weighted MR does not disclose abnormalities. **B:** Ictal single photon emission tomography image shows left temporal hyperperfusion. **C:** Fluorodeoxyglucose positron emission tomography confirms left temporal hypometabolism. The patient had a left temporal lobectomy and hippocampal sclerosis. (From Won HJ, Chang KH, Cheon JE, et al. Comparison of MR imaging with PET and ictal SPECT in 118 patients with intractable epilepsy. *AJNR Am J Neuroradiol*. 1999;20:593–599.)

DIAGNOSIS

Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) does not have an exclusive diagnostic entity for epilepsy-related mental disorders. Each psychiatric syndrome must be judged by a physician whether it is a direct consequence from epilepsy or not. If the syndrome was clearly associated with epilepsy, it will fall under "mental disorder due to another medical condition" diagnosis; if not, the diagnosis should be listed as a primary psychiatric disorder. Nevertheless, since the DSM-5 removed multiaxial system, all of the psychiatric diagnoses have to be recorded alongside a diagnosis of epilepsy. The DSM-5 also encourages clinicians to perform a dimensional assessment, measuring severity of specific psychiatric symptoms by certain scale, and record disability of the patients.

The other way to categorize psychiatric behaviors associated with epilepsy is their relationship to the ictus or seizure discharges. These behaviors may occur as part of the ictus, peri-ictally, or during the interictal period (Table 2.4–4). Moreover, a range of other behaviors appear to have some relation to the ictus but do not clearly fall into one of the former three categories.

Table 2.4–4.
Behavioral Disorders in Epilepsy

Ictal
Ictal psychic symptoms (aura)
Nonconvulsive status: simple partial seizures, focal dyscognitive seizures, and periodic lateralizing epileptiform discharges
Catatonia

- Peri-ictal (includes prodromal, postictal, and mixed ictal)
  - Prodromal symptoms: irritability, depression, headache, etc.
  - Postictal confusion
  - Peri-ictal psychoses
    - Concomitant with increased seizure frequency
    - Concomitant with decreased seizure frequency
- Postictal psychoses
- Interictal psychosis and personality disturbances
  - Interictal psychosis (schizophreniform psychosis)
  - Personality disorders
  - Gastaut–Geschwind syndrome
- Behavioral disturbances variably related to ictus
  - Mood disorders (depression and mania)
  - Anxiety disorders including panic and posttraumatic stress disorder
  - Aggression and violence
  - Hyposexuality
  - Suicide
- Other behaviors

Ictal Features

Seizure discharges can produce semipurposeful automatisms and psychic auras, such as mood changes, derealization and depersonalization, and forced thinking. Ictal fear, which ranges from a vague apprehension to abject fright, has occurred without any other seizure manifestation, and ictal depression has extended days or longer after the seizure has passed. Some patients have pleasurable auras. Fyodor Dostoyevsky had "ecstatic auras" in which he felt in perfect harmony with the entire universe and "would give 10 years of this life, perhaps all of it, for a few seconds of such bliss." The experience of epileptic derealization or depersonalization could impair reality testing. Another psychic aura is "forced thinking," characterized by recurrent intrusive thoughts, ideas, or crowding of thoughts. Forced thinking must be distinguished from obsessional thoughts and compulsive urges. Some patients may have vivid auditory and visual hallucination during having an epileptic discharge. Epileptic patients with forced thinking experience their thoughts as stereotypical, out-of-context, brief, and irrational, but not necessarily as ego dystonic.

A 36-year-old right-handed man presented with frontal headaches and 5 years of focal dyscognitive seizures. His seizures began with 15 seconds of a sense of "impending doom," speech arrest, and orobuccal movements followed by 30 seconds of altered consciousness. At seizure onset, the patient felt forced to think the phrase "tell me yes." The phrase repeated several times without his being able to control it. Concomitantly, his mouth would open, and he would attempt to say the phrase but could utter only unintelligible sounds. The patient interpreted this phrase as a call for help. On examination, he had a mild memory deficit, normal language testing, a right facial droop, and brisk right-sided reflexes with a right Babinski sign. Neuroimaging revealed a left frontal mass lesion. EEGs showed amplitude attenuation and polymorphic delta in the left frontal area, and intraoperative

electrocorticography disclosed polyspike and spike-wave discharges, associated with impaired language, from just below the lesion. The patient underwent subtotal resection of a 4.3 × 3 × 3 cm oligodendroglioma. Postoperatively, his forced thinking and seizures resolved, but he had a nonfluent aphasia.

Cognitive disorders follow status epilepticus with focal, focal dyscognitive, or absence seizures. Recurrent or prolonged focal seizures do not result in alteration of consciousness or invariable abnormalities on EEG, and, if manifested by psychic auras, focal seizures may be difficult to distinguish from primary psychiatric disturbances. Status epilepticus from focal dyscognitive seizures and absence seizures results in prolonged alterations of responsiveness. With the addition of various ictal auras, focal status epilepticus (epilepsia partialis continua) can appear psychotic. Nonconvulsive status epilepticus can also present with immobility, waxy flexibility, and behavioral negativism resembling catatonia. Occasionally, EEGs and a therapeutic trial of antiepileptic medications may be the only way to distinguish behavioral disturbances due to nonconvulsive status epilepticus. Finally, recurrent interictal EEG complexes, known as *periodic lateralizing epileptiform discharges,* may also be associated with prolonged confusional behavior and focal cognitive changes.

A 68-year-old man had a left temporal–parietal hemorrhagic stroke. An initial fluent aphasia and right hemiparesis completely resolved, but he developed poststroke epilepsy. His seizures began with speech arrest and were followed by secondary generalization to tonic–clonic seizures. The postictal periods lasted days due to continued left-hemisphere periodic lateralizing epileptiform discharges. During these prolonged postictal periods, he was confused and placid and had a return of his aphasia. One year later, after achieving seizure control, the patient developed mania for the first time in his life. His mania was in a clear sensorium without a change in his neurological examination or epileptiform activity on EEG. He did not sleep, had flight of ideas, and had grandiose ideation, including beliefs that he was a three-star general, had killed Adolf Hitler, and was now a millionaire. He exposed himself to everyone, including his daughter, and inserted pencils up his penis, because he believed that he needed catheterization. His psychosis lasted for 3 months until he had two generalized tonic–clonic seizures. Postictally, for 10 days he remained placid, confused, and aphasic, with a right beating nystagmus and periodic lateralizing epileptiform discharges maximal in the left temporal region (Fig. 2.4–3). With a new antiepileptic medication, he returned to normal with total resolution of his mania.

Peri-Ictal Features

Psychiatric disturbances can occur before seizures (prodromal), after seizures (postictal), or during intermittent seizure activity. Some patients experience prodromal symptoms that begin at least 30 minutes before seizure onset, last 10 minutes to 3 days, and are continuous with irritability, depression, headache, confusion, and other symptoms. The postictal period is characterized by a confusional state lasting minutes to hours or, occasionally, days. Prolonged, postictal confusion may

particularly follow right temporal focal dyscognitive seizures. Some "twilight states" result from a protracted period of intermixed ictal and postictal changes.

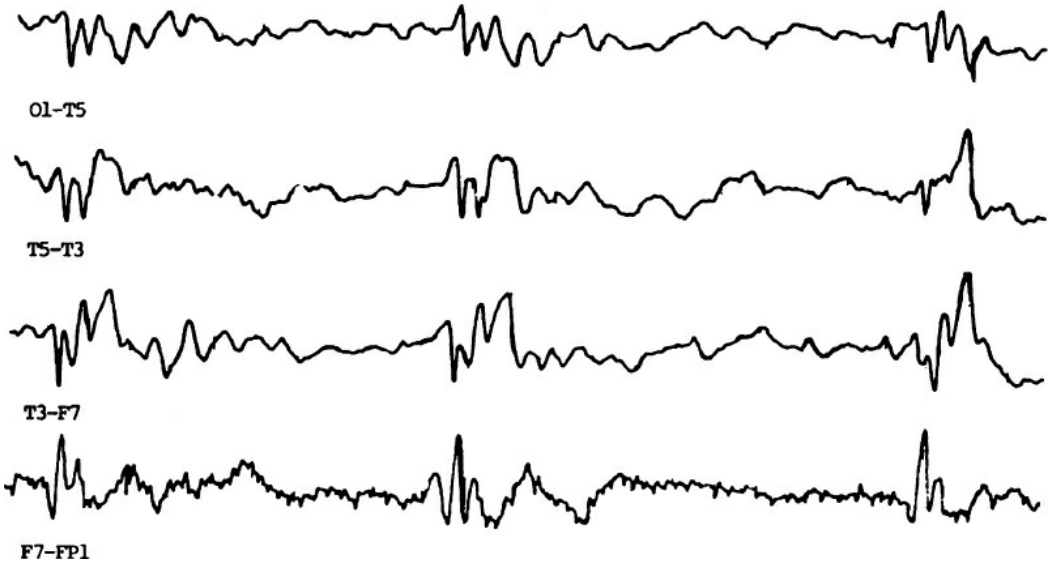

FIGURE 2.4–3. Periodic lateralizing epileptiform discharges. Left-sided electrodes per the International 10/20 System.

Peri-ictal psychotic symptoms often worsen with increasing seizure activity. Rarely, psychotic symptoms alternate with seizure activity. In this *alternating psychosis,* when patients are having seizures, they are free of psychotic symptoms, but when they are seizure free and their EEG has *forced* or *paradoxical normalization,* they manifest psychotic symptoms. This alternating pattern is much less common than the increased emergence of psychotic behavior with increasing seizure activity.

An important peri-ictal psychiatric disorder consists of brief psychotic episodes that follow clusters of generalized tonic–clonic seizures (i.e., postictal psychosis). These psychotic episodes occur in patients who have focal dyscognitive seizures, frequent secondary generalization to tonic-clonic seizures, bilateral interictal discharges, and frequent discharges involving the left amygdala. The postictal psychosis of epilepsy emerges after a lucid interval of 2 to 72 hours (with a mean of 1 day), during which the immediate postictal confusion resolves, and the patient appears to return to normal. The postictal psychotic episodes last 16 to 432 hours (with a mean of 3.5 days) and often include grandiose or religious delusions, elevated moods or sudden mood swings, agitation, paranoia, and impulsive behaviors, but no perceptual delusions or voices are heard. The postictal psychoses remit spontaneously or with the use of low-dose psychotropic medication.

A 33-year-old man with a 15-year history of generalized tonic–clonic seizures and a 4-year history of peri-ictal psychotic episodes had several hospitalizations for recurrent postictal

psychosis. The initial flurry of generalized tonic–clonic seizures was followed by a 24- to 48-hour latency period, and, subsequently, 2 to 7 days of delusions, hallucinations, and disordered thought processes. He believed that people could transmit messages and could read his thoughts, and voices commanded him to love his neighbor. The patient claimed to read the future and to communicate with a dead grandfather who voiced dissatisfaction with things on Earth. During these episodes, the patient had loose associations, euphoria, agitation, and occasional spike and waves on EEGs. Between psychotic episodes, he was psychiatrically and neurologically normal, and his EEGs showed left temporal interictal spikes. After the postictal psychosis, the patient returned to baseline without residual changes in behavior.

Interictal Features

Interictal Psychosis (Schizophreniform/Schizophrenia Like Psychosis). Most psychotic episodes that occur in epilepsy patients are interictal psychosis which usually had no known direct relationship to seizure events or ictal discharge. Duration of psychotic episodes vary greatly from days to years, but usually last longer than 1 month in the vast majority of cases. However, many of these patients develop worsening psychotic symptoms that are concomitant with an increase in seizure frequency or with antiepileptic drug withdrawal, and a few others have worsening psychotic symptoms on control of the seizures (alternating psychosis). The terms *alternating psychosis* and *forced* or *paradoxical normalization* refer to this demonstrable antagonism between the psychosis and the seizures or EEG discharges. Although a spontaneous resolution of interictal psychosis is not common, it may occur in patients with shorter duration of epilepsy history and later onset of epilepsy. The patients with chronic interictal psychosis often have an early age of onset of seizures and a decade or more of poorly controlled partial complex seizures with secondary generalized tonic–clonic seizures. The interictal psychosis may evolve from prior recurrent postictal psychotic episodes. Seizure control with antiepileptic drugs or removal of the seizure focus does not prevent the development of the interictal psychosis, which occasionally emerges for the first time after successful seizure treatment. This disorder sometimes resembles schizophreniform psychosis, and schizoaffective psychosis with intermixed affective symptoms. In addition, there are prominent paranoid delusions, relative preserved affect, normal premorbid personality, and no family history of schizophrenia. Other reported differences with idiopathic schizophrenia are outlined in Table 2.4–5.

A 23-year-old man developed paranoid delusions after his daily focal dyscognitive seizures were controlled for the first time. His seizures dated from 8 years of age and consisted of a rising epigastric sensation and facial flushing followed by a motionless stare and automatisms, often culminating in secondary generalized tonic–clonic seizures. Before initiating antiepileptic therapy, there was no history of paranoid or psychotic behavior. Afterward, he believed that people were sending energy to him through small concealed batteries. He believed that he was able to work this energy off with his fluorescent watch dial

and a one-armed plastic crucifix in his boot. The patient also felt that people were observing him, trying to manipulate him, and were threatening him through telephone lines and telephone poles. His examination was remarkable for the degree of emotion when relating his bizarre ideas. He had a lesion in the left anterior temporal area, probably consistent with an old calcified cyst, and left temporal spikes on EEG. His paranoid delusions subsequently abated with antipsychotic therapy.

Table 2.4–5.
Proposed Predisposing Factors for the Interictal Psychosis of Epilepsy

Epilepsy characteristics
Focal dyscognitive seizures with secondary generalized tonic-clonic seizures
More auras and automatisms than nonpsychotic epilepsy patients
Epilepsy present for 11 to 15 years before psychosis
Long interval of poorly controlled seizures
Recently diminished seizure frequency, especially generalized tonic–clonic seizures
Left temporal focus
Mediobasal temporal lesions, especially tumors
Psychosis characteristics
Atypical paranoid psychosis–paranoia with sudden onset
Psychosis alternating with seizures
Preserved affective warmth
Failure of personality deterioration
Less social withdrawal than schizophrenia
Less systematized delusions than schizophrenia
More hallucinations and affective symptoms than schizophrenia
More religiosity than schizophrenia
More positive, as opposed to negative, symptoms
Few schneiderian first-rank symptoms

Personality Disorders. Among epileptic patients, there is a high prevalence of personality disorders, including borderline, atypical or mixed, histrionic, and dependent disorders. Patients with personality disorders tend to show dependent and avoidant personality traits. The most common personality disorder in epilepsy is a borderline personality. Not surprisingly, epileptic patients frequently lack a stable character structure and can be immature and impulsive. This personality constellation partially explains the increased incidence of irritability, suicide attempts, and intermittent explosive disorder. Those with epilepsy are stigmatized, feared, and subject to difficulties in obtaining a job, driving an automobile, and maintaining a marriage. These psychosocial difficulties, along with any associated mental retardation, contribute to the dependency, low self-esteem, and overall borderline personality traits present in many such patients. In addition, the experience of epileptic auras may contribute to the development of personality disorders.

Another way to understand personality changes in patients with epilepsy is from the evidences of social dysfunction in this population.

Recent studies consistently showed that the patients with temporal lobe epilepsy usually have an impairment in ability to understand others' feeling, thought, and perspective, also known as "theory of mind (ToM)." It is also common for patients with temporal lobe epilepsy to have deficits in other social cognitions including affective and cognitive empathy, processing of implied meaning, and semantic social concept. These components may eventually lead to decreased personal warmth and inappropriate social behaviors.

Gastaut–Geschwind Syndrome. Although there is no general epileptic personality, a group of traits termed the *Gastaut–Geschwind syndrome* occurs in a subset of patients with focal dyscognitive seizures. Some epilepsy patients with a temporal limbic focus develop a sense of the heightened significance of things. These patients are serious, humorless, and overinclusive, and have an intense interest in philosophical, moral, or religious issues. Occasionally, epilepsy patients experience multiple religious conversions or experiences. In interpersonal encounters, they demonstrate *viscosity,* the tendency to talk repetitively and circumstantially about a restricted range of topics. They can spend a long time getting to the point, give detailed background information with multiple quotations, or write copiously about their thoughts and feelings (hypergraphia). Viscosity may particularly occur in patients with left-sided or bilateral temporal foci.

A 39-year-old man developed seizures after a contusion of the left temporal region. His seizures began with stereotypical voices and generalized to tonic–clonic seizures. He was extremely circumstantial and tangential, stressed every detail, and had difficulty getting to the point. Ironic and minor philosophical insights were fascinating to him. He wrote 30-page rambling letters to his physician, and his writings were full of metaphors and quotes. An example of his writing was as follows: "I became overwhelmed by the sentiment of a letter composed in my head before reaching paper. The sentiment of this letter continued to expand in all dimensions until it seemed no longer connected to any specific ideal, but more to an all pervasive color, yellow, and a smell, like burning leaves. I felt deliriously happy, but I felt in danger as well. Afterwards I got an acute attack of aphasia and could do nothing but shrug. My prior prophet voices which went away with the Dilantin were saying something profound that I needed to get down on paper. It seems as though I am a prophet and I will never have another problem for the rest of my life."

Conclusive proof that epileptic patients with a temporal lobe foci are disproportionately prone to the Gastaut–Geschwind syndrome has remained elusive. Most of the early studies used the MMPI, a test that proved insensitive to most of the specific traits attributed to epilepsy. Studies with the Bear–Fedio Inventory, an MMPI-like instrument developed to assess these "epileptic traits," found that epileptic patients with temporal lobe foci were sober and humorless, dependent, and circumstantial and had strong philosophical interests. In addition, those

with a left-sided focus had a more reflective ideational style and maximized their problems, whereas those with a right-sided focus had emotional tendencies and minimized their problems. Further investigations with the Bear–Fedio Inventory described these seizure patients as having viscosity in interactions, prominent religious interests, a pronounced sense of personal destiny, and deepened affect. However, other applications of this inventory found the same characteristics in nonepileptic patients with psychiatric disorders or with comparable physical disabilities. Although these personality characteristics do occur in some epileptic patients, they may not be specific for patients with seizure disorders.

Behavioral Disturbances Variably Related to Ictus

Mood Disorders. Depressive disorder is the most prevalent neuropsychiatric disorder in epilepsy; it is also the main diagnosis among epileptic patients in mental hospitals. Depression is twice as common in epilepsy patients as in comparably disabled populations, suggesting that much of the depression in epilepsy patients is more than just a psychological reaction to a disability.

Depression can manifest in several forms; MDD, dysthymic disorder, adjustment disorder, and depressive episode of bipolar disorder. Most patients with mood disorders and epilepsy have dysthymia; some investigators refer to this condition as a chronic interictal depression (*interictal dysphoric disorder of epilepsy*), and emphasize its association with paroxysmal irritability and agitation along with a good therapeutic response to antidepressant medications. These patients may have accompanying paranoia and hallucinations, emphasizing the continuum with psychotic disorders. Patients with interictal dysphoria tend to have frequent focal dyscognitive seizures, possibly with greater left-sided temporal foci, although this lateralization is not well established. The experience of certain psychic auras, especially those with cognitive content, may predispose to interictal depression. Several investigators also report increased seizure control or a decrease in seizures before the onset of interictal depressive symptoms. Patients with this “alternating depression” experience relief with a seizure or ECT.

There are other associations of depression with epilepsy. The rare occurrence of ictal depression may not only outlast the actual ictus but also may lead to suicide. Depression also occurs peri-ictally. Episodic mood disturbances, often with agitation, suicidal behavior, and psychotic symptoms, may occur with increasing seizure activity. Finally, postictal depression is common, and a prolonged depressive state occasionally follows focal dyscognitive seizures, even when ictal experiences do not include depression.

Mood disorder due to epilepsy with manic features or with mixed features is much rarer than mood disorder due to epilepsy with depressive

features or with major depressive-like features. Rarely, manic symptoms may emerge with an increase in seizure frequency or after seizure control. Although a right temporal focus is a possible source of mania in epilepsy, this laterality is not well established.

Anxiety Disorders. Anxiety disorders, including generalized anxiety disorder (GAD), panic disorder, obsessive compulsive disorder (OCD), PTSD, phobia and social anxiety disorder, are the second most common psychiatric comorbidities in patients with epilepsy after depressive disorders. The prevalence of anxiety disorders in epileptic patients may be as high as 30 percent. Anxiety disorders which occur during interictal phase must be distinguished from simple partial seizures manifesting as anxiety or panic attack. Anxiety may be present with depression or other psychopathology, as part of Cluster C personality disorders, or independently as a GAD. Some patients with epilepsy clearly have PTSD from the psychological trauma of their recurrent seizures. This may contribute to the prevalence of nonepileptic seizure epilepsy among patients with true epilepsy.

NONEPILEPTIC BEHAVIORAL DISORDERS, OR PSYCHOGENIC NONEPILEPTIC SEIZURES

Nonepileptic behavioral disorders, or psychogenic nonepileptic seizures (PNES) are involuntary, psychogenically induced spells that mimic many epileptic behaviors. Though the literature implies that it is a unique disorder, it is but one type of the functional neurological symptoms disorder in DSM-5, what was previously called conversion disorder in DSM-IV. In ICD-10, nonepileptic seizures are categorized as "dissociative convulsions." Up to 30 percent of patients with epilepsy are affected by nonepileptic seizures. A previously used term, *pseudoseizure* is discouraged as it suggests that the patient is faking or that the condition is not real, either can be invalidating for the patient. Although the common term *nonepileptic seizure* is used here, some experts believe that it is best to avoid the term "seizure." Since the behaviors are not seizures, psychogenic nonepileptic events or episodes may create less confusion among patients and families.

Differentiating epileptic seizures from nonepileptic seizures can be difficult based on history alone. Video EEG monitoring is the gold standard and can diagnose nonepileptic seizures with near certainty; however, a diagnosis of nonepileptic seizures does not rule out epileptic seizures. Approximately 10 to 15 percent of nonepileptic seizure patients have a true seizure disorder as well, and nonepileptic seizures may result from the elaborating or "highlighting" of their epileptic seizures. Patients with nonepileptic seizures are most commonly women between the ages of 26 and 32 years with psychological stressors and poor coping skills.

Nonepileptic seizures are most commonly characterized by unresponsiveness with motor activity that does not fit a typical focal dyscognitive or generalized tonic–clonic seizure (Table 2.4–6). In children, nonepileptic seizures are usually characterized by unresponsiveness, with violent and uncoordinated movements of the whole body. There is no female gender preference before adolescence. However, every epileptic behavior can occasionally occur, including tongue biting and incontinence, and nonepileptic seizures are especially difficult to differentiate from the atypical motor behavior of frontal lobe epilepsy. The most helpful differentiation feature may be an ictal duration of 2 minutes or more. In addition, nonepileptic seizures usually occur in the presence of a witness; can often be induced with injections, hypnosis, or suggestions; and are poorly responsive to antiepileptic medications. Ultimately, the differentiation may require continuous video EEG monitoring to demonstrate (1) the clinical event is incompatible for epileptic seizure, and (2) it does not have correlating seizures on EEG. Once a certain type of event has been captured and characterized, the diagnosis can be made regarding that particular event. It is important to note that the absence of EEG seizures during monitoring does not rule out a diagnosis of epileptic seizures, as there may be events that have yet to be captured and characterized.

Table 2.4–6.
Nonepileptic Seizures versus Epileptic Seizures

Nonepileptic Seizures	Epileptic Seizures
Preceding ictus	
Absence of explanatory disease or signs	Frequent evidence of neurological disease
Anxiety auras: palpitations, choking, etc.	Wide range of epileptic auras
Seizures may be induced or provoked	Rarely induced except for reactive seizures
During ictus	
Inconsistencies in clinical presentation	Fit specific seizure types
Seizures may differ from attack to attack	Stereotypical seizure pattern
Only occur when others are present	Often occur without witnesses or at night
Gradual onset, prolonged duration (>2 min)	Abrupt onset, short duration (<2 min)
Asymmetrical, out-of-phase movements, pelvic thrusts, and hyperarching	Decrescendo, symmetrical clonic activity in GTC seizure
Rare whole body rigidity	Tonic rigidity at onset of GTC seizure
Rare incontinence, tongue biting, self-injury	Incontinence, tongue biting if generalized
Normal autonomic reactivity, corneal reflex, and pupillary responses	Disturbed autonomic reactivity, corneal reflexes, and pupillary responses
Avoids noxious stimuli or eye opening	Cannot avoid noxious stimuli
Vocalizations may occur throughout ictus	Single vocalization, if present, at onset
Normal ictal EEG	Abnormal ictal EEG
After ictus	
No postictal delirium, pseudosleep	Typical postictal delirium
No increase in prolactin	Prolactin >1000 IU/L, 10 to 20 min postictally
Normal postictal EEG	Postictal slowing on EEG

Subsequent recall of events during ictus	No or fragmentary recall of ictal events
No relationship of ictal frequency to antiepileptic medications	Diminished seizure frequency with antiepileptic medications

EEG, electroencephalogram; GTC, generalized tonic–clonic.

Psychiatric Comorbidities

Patients with nonepileptic seizures have a high incidence of prior psychological trauma such as sexual or physical abuse. These psychosocial stressors may be ongoing and may play role in precipitating nonepileptic events. The patients with nonepileptic seizures also have a high rate of other psychiatric comorbidities, especially depressive disorders, dissociative disorders, anxiety disorders, PTSD, or personality disorders. Many evidences supported that both depressive disorder and anxiety disorder are very common in patients with epileptic seizure; but dissociative disorders, PTSD, and personality disorders, especially cluster B personality disorder, are more associated with PNES. Other psychiatric symptoms related to PNES include emotional dysregulation, impulsivity, and mental retardation. Psychosis may also be found in PNES; however, it is not common.

Treatment and Outcomes. Outcomes of functional neurological symptom disorder vary greatly. In example, patients who suffered from subtype with pure motor weakness usually recover in weeks to months, but the episode also frequently reoccurs. The patients with predominant somatic symptoms such as pain and paresthesia may have a stable course of illness but it tends to be more chronic. The outcome of nonepileptic seizures probably lies between those two. Overall outcome is mediocre; the course of illness can be chronic and fluctuate. Although most of the patients will have less frequent nonepileptic seizure episodes after the treatment, over half of patients do not achieve remission and remain disabled after 10 years of the symptom onset. The outcome may be impacted by delayed diagnosis and lack of access to treatment options. Patients with higher education, younger age of onset and diagnosis, events with less dramatic features, fewer psychiatric comorbidities, fewer additional psychogenic symptoms, lower dissociation scores, and lower scores on the personality dimensions in inhibitedness, emotional dysregulation, and compulsivity may have better outcome.

Treatment begins at diagnosis by the delivery of the diagnosis to patients and families. Because the patients' understanding and acceptance of the diagnosis greatly determine the outcome, the physician should be firm and confident, avoiding vague statements, while remaining compassionate. Psychiatric comorbidities should be optimally treated and psychological stressors have to be addressed. Pharmacological treatment alone may not be effective, despite SSRIs showing some benefits in some

patients. It is crucial to deliver a consistent and long-term care with psychological intervention such as cognitive behavioral therapy. Combination of these treatments helps the patients in regaining functional status, improved comorbid symptoms, and reducing the frequency of nonepileptic seizure event. There is a dearth of literature within the psychiatric community on somatic symptom disorders; this transcends into the limited access to consistent psychiatric care for patients with nonepileptic seizures.

Nonepileptic seizures must be differentiated from those specifically due to the malingering or feigning of epilepsy for secondary gain (Table 2.4–7). Epileptic seizures lend themselves to malingering because of their behavioral and episodic nature and the lack of consistent physical or diagnostic findings.

Table 2.4–7.
Malingered Seizures versus Nonmalingered Nonepileptic Seizures

Malingered Seizures	Nonmalingered Nonepileptic Seizures
Preceding ictus	
More common in men	Marked female predominance
Less likely to obtain prior abuse history	Prior history of physical or sexual abuse
Less likely to obtain prior psychiatry history	Prior psychiatric history
Evident secondary gain	No clear secondary gain
No clear emotional precipitants	Frequent emotional precipitants
Seizures are not suggestible	Seizures may be easily suggested
During ictus	
Seizures under volitional control	Seizures not under volitional control
Conscious awareness of seizures	Subconscious awareness of seizures only
Cannot maintain deficits over time	Able to maintain deficits over time
Errors in seizure behavior are likely to be major distortions	Errors in seizure behavior are likely to be omissions, perseverations, near misses
After ictus	
Angry, anxious on confrontation, with a lack of evidence for epileptic seizures	Indifferent, detached
Uncooperative, including circumstantial and evasive answers; may leave against medical advice	Cooperative with the workup, but answers may be devoid of content

A 33-year-old veteran presented with a complaint of epileptic spells, beginning 3 years after returning home from the war. The patient claimed that the stress of the war induced his seizure disorder, and he requested disability compensation. He described his seizures as the abrupt loss of consciousness associated with jerking movements of his extremities. His episodes occurred irregularly with a frequency of two to four per week. On admission to the hospital, medical staff observed several seizure-like spells in which the patient assumed a flexed posture of his upper and lower extremities and then shook them uncontrollably and in an asynchronous fashion. During this ictal period, the patient had normal pupillary and corneal reflexes. His seizures lasted nearly 5 minutes and then immediately resolved without postictal confusion. Postictally, he recalled comments and other environmental events that occurred during his seizure-like episodes. EEGs obtained immediately after an event did not

reveal postictal slowing, and prolactin levels obtained 15 minutes after a seizure episode were not significantly elevated over baseline levels. His seizures did not respond to antiepileptic medications, but they abated after he changed his strategy and began to explore compensation for other reasons.

Aggression. Lay people have accredited to epilepsy aggressive and violent acts and have even used this epilepsy defense in criminal proceedings. This belief peaked in the 19th century when the criminologist Cesare Lombroso promoted the association of epilepsy with aggressive, sociopathic tendencies. Investigators have bolstered this association with studies showing aggressive verbalizations with stimulation of the amygdala and interictal defensive rage in cats with epileptic hippocampal lesions. A minority of violent epilepsy patients have left-sided amygdala atrophy probably from a prior encephalitis. Among patients in a maximum-security mental hospital, the violent patients had focal temporal slowing or sharp waves on EEG and dilated temporal horns or small temporal lobes on CT scans. Moreover, on neuropsychological test, patients with left temporal lobe seizure foci have higher scores on hostile feelings than other patients with epilepsy. These results suggest that aggression is associated with abnormal temporal structure and electrical discharges.

A 37-year-old left-handed man with epilepsy presented with aggressive episodes. The seizures consisted of an olfactory aura followed by "spacing out" or alteration of consciousness for approximately 1 minute. In addition to these focal dyscognitive seizures, the patient had occasional secondary generalized tonic–clonic seizures with urinary incontinence and tongue biting. During the postictal period, as he began to recover consciousness, he experienced an overwhelming sense of threat or of having been harmed. These feelings became focused on any individual who was in his immediate environment. That person was believed to have beaten or otherwise hurt him and was going to harm him further. The patient felt compelled to attack these individuals, often inflicting significant physical injury. Although his postictal confusion would clear in approximately 1 hour, his sense of being harmed or threatened slowly diminished over approximately 24 hours after a seizure. After the resolution of these feelings, he felt great remorse over the harm that he had done. Nevertheless, on several occasions, he was charged with aggravated assault. Sleep-deprived EEGs confirmed the presence of left anterior temporal epileptiform activity. The patient's aggressive postictal episodes abated with control of his focal dyscognitive seizures with carbamazepine.

Although aggression can occur in relation to an ictus, as exemplified by this patient's subacute postictal aggression, most aggression among epilepsy patients is not related to epileptiform activity. Aggression in epilepsy is usually associated with psychosis or with intermittent explosive disorder and correlates with subnormal intelligence, lower socioeconomic status (SES), childhood behavior problems, prior head injuries, and possible orbital frontal damage. Moreover, although the prevalence of epilepsy among prison inmates has been two to four times that among the general population, studies from the United Kingdom and the United

States have not found more violent crimes among prisoners with epilepsy than among prisoners without epilepsy. Antiepileptic medication may also trigger aggression and mood lability in some patients, especially levetiracetam.

Can violence itself be a seizure? After the 1976 case of a New York City policeman who had never had seizures and successfully claimed the epilepsy defense, criteria for ictal violence were proposed that included video-EEG telemetry (Table 2.4–8). Since then, epilepsy has rarely, if ever, been proved to directly result in well-organized, premeditated, purposeful violence. Such acts require a series of coordinated steps that rarely occur as manifestations of seizures. Simple violent automatisms, such as spitting, verbal profanities, or flailing the arms, can occur at the onset of focal dyscognitive seizures, and secondary violent automatisms can occur as a response to an unpleasant or emotional aura or peri-ictal sensation (Table 2.4–9). More commonly, nondirected violent movements, aimless destructive behavior, or angry verbal outbursts occur during postictal delirium when patients misinterpret attempts to protect or restrain them (resistive violence) or as a manifestation of postictal psychosis and subacute postictal aggression.

Table 2.4–8.
Criteria for the Assessment of Ictal Violence in Epilepsy

The diagnosis of epilepsy is established by at least one specialist in epilepsy.
The presence of epileptic automatisms is documented by history and by closed circuit television EEG telemetry.
The presence of violence during epileptic automatisms is verified in a videotape-recorded seizure in which ictal epileptiform patterns are also recorded on the EEG.
The aggressive act is characteristic of the patient's habitual seizures, as elicited by history.
A clinical judgment is made by the epilepsy specialist attesting to the possibility that the aggressive act was part of a seizure.

EEG, electroencephalogram.

Sexuality. Patients with epilepsy tend to be hyposexual. Men and women experience disturbances of sexual arousal and a lower sexual drive. Some patients have a disinterest in "all the usual libidinous aspects of life," including loss of erotic fantasies or dreams, and may experience impotence or frigidity. Although comorbid depression and anxiety may affect sexual behavior, men with epilepsy have a two- to threefold increased risk of erectile dysfunction, suggesting a neurophysiological component, and studies of sex hormones suggest the possibility of a subclinical hypogonadotropic hypogonadism. Moreover, there can be undesirable effects on sexual performance of antiepileptic medication, including erectile dysfunction, decreased sexual desire, retrograde ejaculation, and altered semen quality. Hormone replacement therapy may normalize

sexual drive and sexual performance in some patients. Substantial improvement to the point of public hypersexuality can occur after seizures are brought under control. Moreover, before temporal lobectomy, most epileptic patients are hyposexual, but nearly one-third of them have an increase in libido after the operation, providing that their seizures are controlled.

Other sexuality changes are rare. Individual cases of homosexuality, transvestism, fetishism, and gender dysphoria are not frequent enough to exclude a coincident association. True ictal sexual manifestations are also unusual; however, libidinous feelings, erotic sensations, sexual remembrances, and even orgasm rarely occur, primarily in women and probably from seizure discharges in the amygdala. In addition, ictal masturbation has occurred with absence status. A woman with nymphomania proved to have incidental sexuality from sensory simple partial seizures caused by a tumor in the sensory cortex representing her genital region.

Table 2.4–9.
Mechanisms of Aggression among Epilepsy Patients

Period	Cause
Interictal	Impulse control disorder
	Mental retardation or cognitive impairments
	Personality disorders
	Schizophreniform/Schizophrenia like psychosis
	Medication related
Prodrome	Mounting tension, irritability
Ictal	Direct manifestation of the seizure
	Violent automatism
	Reaction to a negative aura
	Subtle seizure equivalents
Postictal	Resistive
	Postictal psychosis
	Subacute postictal aggression
	Poriomania and somnambulism

Suicide and Self-Harm. Patients with epilepsy are at risk of higher rates in self-harm, suicidal ideation, and complete suicide. The risk of completed suicide in epilepsy patients is four to five times greater than that among the nonepileptic population, and those with focal dyscognitive seizures of temporal lobe origin have a particularly high risk, as much as 25 times greater. Death by suicide occurs in 3 to 7 percent of epilepsy patients. A comparison of suicide attempts among patients with epilepsy and comparably handicapped nonepileptic controls has reported that 30 percent of those with epilepsy had attempted suicide as compared to 7

percent of the controls. This increased risk of suicide continues even long after temporal lobectomy and successful control of seizures. Although many studies showed that a severity of depressive symptoms and psychosocial status influence suicidal ideation in patients with epilepsy, most suicidal behavior among epileptic patients is not directly due to reactions to the psychosocial stressors of having a seizure disorder. Rather, these patients are likely to attempt suicide in conjunction with borderline personality behaviors and are likely to complete suicide during postictal psychosis. In addition, some evidences suggested that antiepileptic medications may also increase suicidality. This conclusion is still controversial since there is a lack of control groups. Contributors to successful suicides include paranoid hallucinations, agitated compunction to kill themselves, and occasional ictal command hallucinations to commit suicide. In sum, the cause of self-injurious behavior in patients with epilepsy is often multifactorial, and clinicians should consider any potential cause for an early intervention to prevent self-harm and suicide.

A 26-year-old woman had her initial seizure during her first pregnancy at 18 years of age. Her seizures included echoing sounds "like walking in a cave," a motionless stare with stereotypical automatisms, postictal confusion, and occasional secondary generalized tonic–clonic seizures. Because of a variable antiepileptic drug response, she underwent EEG and closed-circuit television video-EEG (CCTV-EEG) telemetry that documented focal dyscognitive seizures from a right temporal focus and nonepileptic seizures. The patient, who had six children by six different individuals, had prominent feelings of inadequacy and isolation and was considering cutting her wrists "just to see if anyone cared." Her multiple suicide attempts and threats resulted in five psychiatric hospitalizations. During one period of time, she complained of decreased menses, weight gain, stretch marks, increased appetite and sleep, and exhaustion. She insisted that she was pregnant despite six negative pregnancy tests and multiple evaluations.

Other Behavioral Changes. Other psychiatric disorders may be associated with epilepsy or epileptiform EEG activity. A specific association of epilepsy with dissociative identity disorder, depersonalization disorders, possession states, fugue states, and psychogenic amnesia is intriguing but unresolved. In one intracarotid amobarbital (Amytal) study of multiple personality disorder in two seizure patients, the different personalities were precipitated without seizure activity. Persistent alterations in the experience of self and feelings of being taken over by others may occur in patients with auras of derealization and depersonalization. In epilepsy, prolonged periods of compulsive wandering with amnesia have resulted from an admixture of ictal and postictal changes and have been termed *poriomania*. Finally, patients with epilepsy are subject to other behavioral difficulties stemming from their epilepsy, such as adjustment disorders, subtle cognitive effects of seizures, and the potential behavioral effects of antiepileptic medications.

PATHOLOGY AND LABORATORY EXAMINATION

Neurodiagnostic Tests

Seizure and epilepsy are clinical diagnoses. Additional tests can help identify provocative factors and diagnose seizure syndromes. In addition to the routine laboratory data and toxicology screens used to exclude reactive seizures, several neurodiagnostic tests are useful in the assessment of epilepsy. Scalp EEG is the most widely used confirmatory test for seizures; however, sensitivity ranges from 29 to 55 percent. This is because the source area of epileptic discharge has to be large enough in order to be detected by scalp EEG. Only 10 percent of cortical spikes having the source area less than 10 cm^2 produced scalp potentials; the source area, that is less than 6 cm^2 of gyral cortex, is rarely captured by scalp EEG. Single routine EEGs are frequently normal and must be repeated, particularly with provocative maneuvers, such as photic stimulation, hyperventilation, and sleep deprivation. Occasionally, video EEG telemetry for an extended period of time is necessary to characterize spells of unknown etiology, for localization of seizure focus, and for classification of seizure in a patient with known epilepsy. Inpatient epilepsy monitoring units are very helpful in distinguishing epilepsy from PNES. Neuroimaging procedures, such as CT scans and magnetic resonance imaging (MRI), can more precisely visualize a seizure focus or even mesial temporal sclerosis (Fig. 2.4–4). If the patient is a candidate for epilepsy surgery due to medication-resistant epilepsy, other tests that occasionally aid in localizing the seizure focus include invasive EEG, SPECT scans, and PET scans. PET scans may show interictal hypometabolism around the temporal seizure focus and are also useful in the presurgical assessment of medically intractable seizure patients. Neuropsychological examinations, particularly during a Wada test, further help in localizing and lateralizing memory and language before surgery.

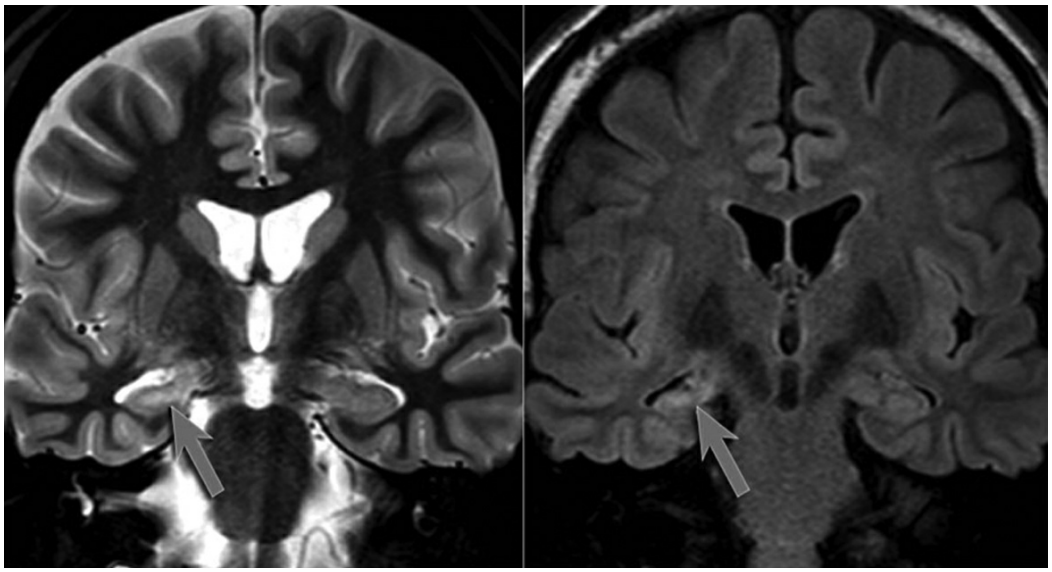

FIGURE 2.4–4. Coronal T2-weighted (**left**) and FLAIR (**right**) MR images showing prominent right-sided mesial temporal sclerosis. The high signal indicates gliosis. There is associated atrophy with enlargement of the right temporal horn of the lateral ventricle. (From Robin Smithius, MD. http://www.radiologyassistant.nl/)

Neuropathology

The common pathological findings in epilepsy are mediobasal temporal lobe lesions. Approximately two-thirds of epileptic adults have a temporal lobe focus, and two-thirds of these have mesial temporal sclerosis with pyramidal cell loss in the hippocampus. Theories about the cause of mesial temporal sclerosis include perinatal insults, dysgenesis, and kindling from reactive seizures. Another 20 to 25 percent of those with temporal lobe lesions have tumors, such as hamartomas and gangliogliomas. The rest have scars from trauma and other causes or lack a distinct histological lesion.

COURSE AND PROGNOSIS

Most epileptic patients have a good prognosis. The majority of seizures can be controlled sufficiently with antiepileptic medications so that the patient can live a productive life. Some seizures, such as absence seizures, tend to disappear by adulthood. For epileptic patients who are medically intractable, epilepsy surgery offers a good alternative (e.g., temporal lobectomy), provided that the focus can be localized. In addition, most epileptic patients do not have psychiatric disorders, and others have psychiatric difficulties only if they endure many years of poorly controlled seizures. For those with behavioral problems, antiepileptic drugs or epilepsy surgery may relieve some symptoms, such as aggression and impulsivity, but may not affect the emergence of others, such as psychosis and suicidal behavior.

TREATMENT

Antiepileptic Medications

In the treatment of psychiatrically disturbed epileptic patients, a first consideration is the behavioral effects of antiepileptic medications. Carbamazepine and valproate have significant antimanic and modest antidepressant properties. They have an efficacy in the long-term prophylaxis of manic episodes. Carbamazepine and valproate may also ameliorate some dyscontrolled, aggressive behavior in brain-injured patients. Lamotrigine is the only antiepileptic that has a well-established efficacy in preventing recurrence of depressive episode in bipolar disorder. Clonazepam, in addition to its anxiolytic properties, can serve as a supplement to other antimanic therapies. Gabapentin and pregabalin also decrease anxiety and improve general well-being in some epilepsy patients.

Carbamazepine and ethosuximide may have value for borderline personality disorder in reducing mood swing and behavioral disinhibition.

Encephalopathic changes occur at toxic levels of all antiepileptic drugs. Even at therapeutic levels, barbiturates may need discontinuation because of drug-induced depression, suicidal ideation, sedation, psychomotor slowing, and paradoxical hyperactivity in the very young and the very old. Levetiracetam may cause irritability and aggression. Gabapentin may induce aggressive behavior or hypomania, and vigabatrin (Sabril) may precipitate depression. In addition, clinicians need to be aware of the potential emergence of psychopathology on withdrawal of antiepileptic medications. Anxiety and depression are the most common emergent symptoms, but psychosis and other behaviors may also occur (Table 2.4–10).

Psychotropic Medications

Antipsychotics. Small number (up to 15 percent) of the interictal psychosis resolves without any treatment, but a majority of patients require antipsychotics. There are no distinct differences in efficacies between second generation antipsychotics and first generation antipsychotics. Therefore, a second consideration is the seizure threshold lowering effect of psychotropic medications (Table 2.4–11). This is usually not a problem but can occasionally reach clinical significance in poorly controlled epilepsy. Evidences from case reports suggest that haloperidol, risperidone, paliperidone, and aripiprazole may be the least seizure threshold lowering agents in contrast to clozapine, olanzapine, quetiapine, chlorpromazine, thioridazine, and perphenazine that may have a higher potency to induce seizure. Data on newer atypical antipsychotics such as lurasidone and asenapine are yet available. Nevertheless, these psychotropic drugs are most convulsive with rapid introduction of the drug and in high doses. Clozapine, for example, has induced seizures in 1 to 4.4 percent of patients, particularly when the dose was rapidly increased. In general, mild-to-moderate psychotic episode patients usually require 300 to 1,000 mg/day of chlorpromazine equivalent (CE) dose (Table 2.4–12). Although the patients with severe psychotic episode, for example, unshakable delusion with severe aggression and suicidal ideation, may require dosage higher than 1,000 mg/day of CE, it is still the best to initiate with low dose and titrate up slowly while monitoring antiepileptic levels and EEGs.

Table 2.4–10.
Behavioral Effects of Antiepileptic Drugs

Antiepileptics	Treatment–Emergent Behavioral Adverse Effects
Barbiturates	Depression, irritability, aggression, impaired cognition and attention, hyperactivity

Carbamazepine–Oxcarbazepine	Irritability, impaired attention
Ethosuximide	Behavioral abnormalities, psychosis
Felbamate	Depression, anxiety, irritability
Gabapentin	Behavioral problems in children
Lamotrigine	Insomnia, agitation, emotional lability
Levetiracetam	Irritability, emotional lability
Phenytoin	Encephalopathy, depression, impaired attention
Pregabalin	-[a]
Tiagabine	Depression (nonconvulsive status epilepticus), irritability
Topiramate	Depression, psychomotor slowing, psychosis, impaired cognition (word-finding and memory)
Valproate	Encephalopathy, depression
Vigabatrin	Depression, aggression, psychosis
Zonisamide	Agitation, depression, psychosis

[a]Less likely to cause behavioral problem.

Adapted from Mula M, Monaco F. Antiepileptic drugs and psychopathology of epilepsy: an update. *Epileptic Disord*. 2009;11(1):1–9.

Antidepressants. Prescribing antidepressants can alleviate both anxiety and mood symptoms at the same time. Given that the overall efficacies of antidepressants in treating depressive symptoms are the same, the principle of choosing antidepressants will be based on adverse effects and tolerability of the patients. The most concerned issue for antidepressants in epilepsy is the seizure threshold lowering property. Overdose and intoxication of tricyclic (TCA) and tetracyclic antidepressants (TeCA), SSRI, and serotonin norepinephrine reuptake inhibitor (SNRI) have been well documented as one of the potential causes of seizure. These medications in therapeutic dose usually do not adversely affect seizures. Some studies demonstrated that antidepressants may even help in reducing the frequency of seizure in some patients. Nevertheless, one has to bear in mind that these findings were mainly from studies using SSRI and SNRI, and not from randomized-controlled study. Therefore, epileptogenic property of both TCA and TeCA remain controversial.

Table 2.4–11.
Seizure Threshold Lowering Effect of Psychotropic Medications

Potential	Antipsychotic	Antidepressant	Other Psychotropic
High	Chlorpromazine (Thorazine)	Bupropion (Wellbutrin)	
	Clozapine (Clozaril)	Imipramine (Norfranil)	
	Thioridazine (Mellaril)	Maprotiline (Ludiomil)	
	Perphenazine (Trilafon)	Amitriptyline (Elavil)	
	Olanzapine (Zyprexa)	Amoxapine (Asendin)	
	Quetiapine (Seroquel)	Nortriptyline (Aventyl)	

Moderate	Most piperazines Thiothixene (Navane) Ziprasidone (Geodon)	Protriptyline (Vivactil) Clomipramine (Anafranil)	Lithium (Eskalith)
Low	Fluphenazine (Prolixin) Haloperidol (Haldol) Loxapine (Loxitane) Molindone (Moban) Pimozide (Orap) Risperidone (Risperdal) Paliperidone (Invega) Aripiprazole (Abilify)	Doxepin (Sinequan) Desipramine (Norpramin) Trazodone (Desyrel) Trimipramine (Surmontil) Selective serotonin reuptake inhibitors	Ethchlorvynol (Placidol) Glutethimide (Doriden) Hydroxyzine (Vistaril) Meprobamate (Equanil) Methaqualone (Quaalude)

Table 2.4–12.
Chlorpromazine Equivalent (CE) Dosages of Antipsychotics

Drug Classification	Name	CE (mg)
First generation antipsychotics	Chlorpromazine	1
	Thioridazine	1
	Fluphenazine	50
	Trifluoperazine	20
	Perphenazine	10
	Haloperidol	50
	Pimozide	50
	Sulpiride	0.5
	Zotepine	2
Second generation antipsychotics	Risperidone	50–100
	Olanzapine	30–50
	Quetiapine	1–2
	Ziprasidone	4–8
	Paliperidone	60–100
	Aripiprazole	30–50
	Clozapine	2

CE: 1 mg of each antipsychotic can be converted into a certain dose of chlorpromazine, for example, 1 mg of haloperidol is about 50 mg of chlorpromazine.

Adapted from Adachi N, Kanemoto K, de Toffol B, et al. Basic treatment principles for psychotic disorders in patients with epilepsy. *Epilepsia*. 2013;54(Suppl 1):19–33.

According to their better tolerability profile, SSRI and SNRI can be given as first-line antidepressants in patients with epilepsy (Table 2.4–13). SSRI may be more suitable in patients with depressive disorder associated with comorbid anxiety disorder, while SNRI may be more beneficial in patients with depressive disorder with psychomotor retardation, fatigue, anhedonia, and cognitive slowness. TCA and TeCA usually require higher dose for achieving their antidepressant property; therefore, they should be considered once the patients failed to respond with the first-line agents. In addition, these medications should not be discontinued abruptly because of withdrawal symptoms. Agomelatine, a newer antidepressant, demonstrated

an antiepileptic property in preclinical data, and seizure has not been reported as an adverse effect. It may be an alternative medication for depression in patients with epilepsy; however, its efficacy and safety profile are yet to be explored.

Table 2.4–13.
Suggested Antidepressants for Patients with Epilepsy

Suggested Level	Drug Name and Classification	Dose Range
First-line	SSRI	
	Fluoxetine	20–60
	Sertraline	50–150
	Escitalopram	10–20
	SNRI	
	Venlafaxine	75–300
	Duloxetine	30–90
Second-line	TCA	
	Amitriptyline	150–300
	Nortriptyline	50–150
	Imipramine	75–150
	Clomipramine	75–150
	TeCA	
	Maprotiline	75–150
	Mirtazapine	7.5–30

SSRI, selective serotonin reuptake inhibitor; SNRI, serotonin and norepinephrine reuptake inhibitor; TCA, Tricyclic antidepressants; TeCA, Tetracyclic antidepressants.

Antianxiety. Although anxiety disorders are common comorbidities with epilepsy, there is lack of evidences from good quality randomized-controlled trial epilepsy patients. Therefore, regimens of psychotropic prescription largely come from evidences in general psychiatric populations. Benzodiazepine may be particularly helpful in rapid relief of symptoms from anxiety, but a recent study demonstrated that chronic use of benzodiazepine in epilepsy patients with GAD did not decrease anxiety level. If the patients had comorbid panic disorder, GAD, OCD, or PTSD, SSRIs are probably the best choice of treatment. Pregabalin recently showed an antianxiety effect in small open-labeled trial in patients with epilepsy and GAD. Therefore, it can be an alternative choice if the patient did not respond to the first SSRI.

Drug Interactions

A third treatment consideration is the potential for interaction of antiepileptic and psychotropic medications. Most commonly, an antiepileptic drug increases the metabolism of a psychotropic drug with a consequent decrease in its therapeutic efficiency. Conversely, withdrawal of

antiepileptic drugs can precipitate rebound elevations in psychotropic levels. Moreover, the initiation of a psychotropic drug may result in competitive inhibition of antiepileptic drug metabolism with elevations of antiepileptic drug levels to toxicity. These interactions should be considered carefully especially in older antiepileptic medications; however, the newer antiepileptic medications have fewer potential interactions with psychotropic medications. Lacosamide (Vimpat), levetiracetam (Keppra), zonisamide (Zonegran), gabapentin (Neurontin), clobazam (Onfi), perampanel (Fycompa), ethosuximide (Zarontin), vigabatrin (Sabril), and tiagabine (Gabitril) are relatively free of enzyme-inducing or -inhibiting properties. Lamotrigine (Lamictal) and topiramate (Topamax) are the only two that have been reported for pharmacodynamics interactions as described in Table 2.4–14.

Surgery

Up to a third of epilepsy patients may have seizures that are not fully controlled with medications. For these patients, resection of epileptogenic brain may offer a far better chance of seizure control than medication alone. The most commonly performed operation is temporal lobectomy, involving resection of removal of 4 to 6 cm of the anterior temporal lobe. More than 80 percent of temporal lobectomy patients experience some reduction in their seizure frequency, and more than 60 percent of patients are entirely seizure-free. Less common epilepsy surgeries include resection of cortical dysplasia, tumors, vascular malformations, areas of old stroke, removal of the epileptogenic hemisphere, and ligation of the corpus callosum. Neocortical resections resulted in about 50 percent of seizure freedom.

Table 2.4–14.
Antiepileptic–Psychotropic Drug Effects on Blood Levels

Antiepileptic	Type of Seizure	Effects of Psychotropic Drug on Antiepileptic Drug[a]	Effects of Antiepileptic Drug on Psychotropic Drug[a]
Carbamazepine (Tegretol)	Focal, generalized	Potentially decreased	Decreased
Oxcarbazepine (Trileptal)	Focal, generalized	Mild increased, rarely toxic levels	Decreased
Phenytoin (Dilantin)	Focal, generalized	Potentially decreased or increased, rarely toxic levels	Decreased
Phenobarbital (Barbita) and primidone (Myidone)	Focal, generalized	Potentially decreased	Significantly decreased
Valproic acid (Depakene)	Focal, generalized,	Potentially increased, rarely toxic levels	Potentially decreased

	absence		
Ethosuximide (Zarontin)	Absence	None known	None known
Clonazepam (Klonopin)	Generalized	Potentially decreased	Potentially decreased
Gabapentin (Neurontin)	Focal, generalized	No significant interactions known	No significant interactions known
Lamotrigine (Lamictal)	Focal, generalized	No significant interactions known	Decrease quetiapine, increase olanzapine, Increase aripiprazole
Topiramate (Topamax)	Focal, generalized	No significant interactions known	Increase haloperidol, decrease risperidone
Vigabatrin (Sabril)	Focal, generalized	No significant interactions known	No significant interactions known
Tiagabine (Gabitril)	Focal, generalized	No significant interactions known	No significant interactions known
Felbamate (Felbatol)	Focal, generalized	No significant interactions known	No significant interactions known
Levetiracetam (Keppra)	Focal, generalized	No significant interactions known	No significant interactions known
Pregabalin (Lyrica)	Focal	No significant interactions known	No significant interactions known
Tiagabine (Gabitril)	Focal	No significant interactions known	No significant interactions known
Vigabatrin (Sabril)	Focal, generalized	No significant interactions known	No significant interactions known
Zonisamide (Zonegran)	Focal, generalized	No significant interactions known	No significant interactions known
Perampanel (Fycompa)	Focal, generalized	No significant interactions known	No significant interactions known

[a]Antipsychotic and antidepressant drugs; lithium and benzodiazepines have few drug interactions with antiepileptic drugs.

Focal, focal onset seizure; generalized, generalized onset seizure; absence, absence seizure.

Adapted from Patsalos PN. Drug interactions with the newer antiepileptic drugs (AEDs)—Part 2: pharmacokinetic and pharmacodynamic interactions between AEDs and drugs used to treat non-epilepsy disorders. *Clin Pharmacokinet*. 2013;52:1045–1061.

Neuropsychiatric manifestations perisurgery are beginning to be characterized and data is most robust in temporal lobectomies. Similar to medically managed patients, surgical patients suffer from depression and anxiety. Studies have shown that independent of postoperative seizure outcome, severity of depression and anxiety, independently, predict quality of life. These psychiatric symptoms gradually improve after surgery; only a minority of population will suffer from refractory depression and anxiety.

Removal of the amygdala and most of the hippocampus may have postoperative behavioral effects. Some patients have an anomia or a verbal memory deficit after resection of the dominant hemisphere, and patients occasionally develop a transient postoperative affective disorder. Others experience a reduction in postictal psychosis, depression, and hyposexuality, but epileptic patients may continue to develop interictal

psychosis, personality changes, and suicidal behavior even long after the temporal lobectomy. Moreover, patients with preoperative psychotic symptoms are at higher risk for a poor surgical outcome and postoperative psychosis. Corpus callosotomy, which aims to prevent the interhemispheric spread of seizures, results in a unique, transient disconnection syndrome of mutism, apathy, agnosia, apraxia of the nondominant limbs, difficulty naming, and writing with the nondominant hand.

Seizure Management

In treating the neuropsychiatric disorders of epilepsy, a final consideration is altering the seizure management itself. In addition to the occasional behavior alleviated by strict seizure control, allowing seizures under carefully controlled conditions, much like ECT, may relieve some cases of peri-ictal psychosis, depression, or other behaviors.

References

Austin JK, Caplan R. Behavioral and psychiatric comorbidities in pediatric epilepsy: Toward an integrated model. *Epilepsia*. 2007;48:1639–1651.

Baker GA. Depression and suicide in adolescents with epilepsy. *Neurology*. 2006;66:S5–S12.

Barry JJ, Ettinger AB, Friel P, Gilliam FG, Harden CL. Advisory group of the epilepsy foundation as part of its mood disorder: Consensus statement: the evaluation and treatment of people with epilepsy and affective disorders. *Epilepsy Behav*. 2008;13(Suppl 1):S1–S29.

*Blumer D, Montouris G, Davies K. The interictal dysphoric disorder: recognition, pathogenesis, and treatment of the major psychiatric disorder of epilepsy. *Epilepsy Behav*. 2004;5:826–840.

*Clancy MJ, Clarke MC, Connor DJ, Cannon M, Cotter DR. The prevalence of psychosis in epilepsy; a systematic review and meta-analysis. *BMC Psychiatry*. 2014;14:75.

Duncan R, Graham CD, Oto M. Neurologist assessment of reactions to the diagnosis of psychogenic nonepileptic seizures: Relationship to short- and long-term outcomes. *Epilepsy Behav*. 2014;41:79–82.

Ettinger AB. Psychotropic effects of antiepileptic drugs. *Neurology*. 2006;67:1916–1925.

*Ettinger AB, Kanner AM. *Psychiatric Issues in Epilepsy: A Practical Guide to Diagnosis and Treatment*. New York: Lippincott Williams & Wilkins; 2006.

Hamid H, Blackmon K, Cong X, et al. Mood, anxiety, and incomplete seizure control affect quality of life after epilepsy surgery. *Neurology*. 2014;82(10):887–894.

Hermann BP, Jones JE. Intractable epilepsy and patterns of psychiatric comorbidity. *Adv Neurol*. 2006;97:367–374.

*Hesdorffer DC, Ishihara L, Mynepalli L, Webb DJ, Weil J, Hauser WA. Epilepsy, suicidality, and psychiatric disorders: a bidirectional association. *Ann Neurol*. 2012;72(2):184–191.

Irwin LG, Fortune DG. Risk factors for psychosis secondary to temporal lobe epilepsy: a systematic review. *J Neuropsychiatry Clin Neurosci*. Winter 2014;26(1):5–23.

Jackson MJ, Turkington D. Depression and anxiety in epilepsy. *J Neurol Neurosurg Psychiatry*. 2005;76:i45–i47.

Jones JE, Hermann BP, Gilliam FG, Kanner AM, Meader KJ. Rates and risk factors for suicide,

suicidal ideation, and suicide attempts in chronic epilepsy. *Epilepsy Behav*. 2003;4:S31–S38.

LaFrance WC, Jr., Baird GL, Barry JJ, et al. Multicenter pilot treatment trial for psychogenic nonepileptic seizures: a randomized clinical trial. *JAMA Psychiatry*. 2014;71(9):997–1005.

*LaFrance WC, Jr., Baker GA, Duncan R, Goldstein LH, Reuber M. Minimum requirements for the diagnosis of psychogenic nonepileptic seizures: a staged approach: a report from the International League Against Epilepsy Nonepileptic Seizures Task Force. *Epilepsia*. 2013;54(11):2005–2018.

Lobello K, Morgenlander JC, Radtke RA, Bushnell CD. Video/EEG monitoring in the evaluation of paroxysmal behavioral events: duration, effectiveness, and limitations. *Epilepsy Behav*. 2006;8(1):261–266.

Manchanda R, Freeland A, Schaefer B, McLachlan RS, Blume WT. Auras, seizure focus, and psychiatric disorders. *Neuropsychiatry Neuropsychol Behav Neurol*. 2000;13(1):13–19.

Marsh L, Rao V. Psychiatric complications in patients with epilepsy: A review. *Epilepsy Res*. 2002;49:11–33.

Mula M. Treatment of anxiety disorders in epilepsy: an evidence-based approach. *Epilepsia*. 2013;54(Suppl 1):13–18.

Ott D, Siddarth P, Gurbani S, Koh S, Tournay A. Behavioral disorders in pediatric epilepsy. *Epilepsia*. 2003;44:591–597.

Oyebode F. The neurology of psychosis. *Med Princ Pract*. 2008;17:263–269.

Reuber M, Pukrop R, Bauer J, Helmstaedter C, Tessendorf N. Outcome in psychogenic nonepileptic seizures: 1 to 10-year follow-up in 164 patients. *Ann Neurol*. 2003;53:305–311.

Riggio S. Psychiatric manifestations of nonconvulsive status epilepticus. *Mt Sinai J Med*. 2006;73:960–966.

*Schachter SC, Holmes GL, Kasteleijn-Nolst Trenite DGA, eds. *Behavioral Aspects of Epilepsy: Principles and Practice*. New York: Demos Medical Publishing; 2008.

Schmitz B. Effects of antiepileptic drugs on mood and behavior. *Epilepsia*. 2006;47(Suppl 2):28–33.

Seethalakshmi R, Krishnamoorthy ES. Depression in epilepsy: Phenomenology, diagnosis and management. *Epileptic Disord*. 2007;9(1):1–10.

Swinkels WA, Kuyk J, van Dyck R, Spinhoven P. Psychiatric comorbidity in epilepsy. *Epilepsy Behav*. 2005;7:37–50.

Tao JX, Ray A, Hawes-Ebersole S, Ebersole JS. Intracranial EEG substrates of scalp EEG interictal spikes. *Epilepsia*. 2005;46(5):669–676.

▲ 2.5 Neuropsychiatric Consequences of Traumatic Brain Injury

JOANNE ALONSO BYARS, M.D., AND RICARDO E. JORGE, M.D.

INTRODUCTION

The neuropsychiatric consequences of TBI encompass disorders also seen in patients without brain injury and those unique to patients with brain damage. The disorders that also occur in patients without brain injury

cover the whole spectrum of psychiatric illness including substance abuse, mood, anxiety, psychotic, and personality disorders. The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) classifies many of these conditions as disorders due to another medical condition, in this case TBI (Table 2.5–1). Most of these disorders lack extensive study in the TBI population, and much research remains needed in this area. The disorders unique to brain injury also cover a wide range of conditions, including pseudobulbar affect (PBA), anosognosia, aprosody, and neglect. Most of these disorders have not undergone extensive examination in patients with TBI.

Table 2.5–1.
DSM-5 Classification of Some Behavioral Syndromes

Occurring after traumatic brain injury
Delirium due to traumatic brain injury
Major and minor neurocognitive disorder due to traumatic brain injury
Personality change due to traumatic brain injury
Labile, disinhibited, aggressive, apathetic, paranoid, combined, other, and unspecified types
Depressive disorder due to traumatic brain injury
With depressive features
With major-depressive-like episode
With mixed features
Bipolar and related disorder due to traumatic brain injury
With manic features
With manic- or hypomanic-like episode
With mixed features
Anxiety disorder due to traumatic brain injury
Post-traumatic stress disorder
Psychotic disorder due to traumatic brain injury

HISTORY

The earliest physical evidence of TBI due to assault occurred 1 million years ago. A damaged skull from an early hominid found in South Africa showed two posterior fractures that matched with the condylar surfaces of an antelope humerus discovered nearby. The earliest written evidence of brain injuries appeared 5,000 years ago on the Edwin Smith Papyrus, which contained the first 27 head injury records. The Hippocratic Corpus included a treatise on head injury with thoughtful comments on skull fractures, delirium, seizures, and coma.

The medical literature has reported associations between TBI and neuropsychiatric disorders for many years. Adolf Meyer, for example, identified a number of disorders that he referred to as the "traumatic insanities." Although he believed that these disorders were determined by a combination of psychological, social, historical, and biological factors, he suggested that there may exist some unique associations between these

disorders and specific lesion locations. Studies of war-related head injuries identified the high prevalence of psychiatric complications following TBI. Several of these studies emphasized the importance of frontal lesions in the pathogenesis of behavioral disturbances. The most famous case of frontal lobe injury, however, was Phineas Gage, who suffered a penetrating frontal brain injury after an explosion shot an iron bar through his skull. After the injury, he was described as childish, capricious, inconsiderate, profane, and having poor judgment.

Analysis of a large series of cases such as the Oxford Collection of Head Injury Records suggests that biological variables such as the extent of brain damage, lesion location, and the presence of post-traumatic epilepsy represent important etiological factors in determining the type and duration of psychiatric syndromes.

COMPARATIVE NOSOLOGY

The neurological and neurosurgical literature abounds with clinical descriptions of both early and delayed behavioral abnormalities that follow TBI. Acute syndromes include confusional states, agitation, restlessness, irritability, and posttraumatic amnesia (PTA). Delayed, often irreversible, consequences of TBI include cognitive disorders (e.g., amnesia or executive dysfunction), traumatic dementia, and organic personality change.

The spectrum of psychiatric disorders attributable to TBI spans almost the entire gamut of psychiatric disorders. The DSM-5 categorizes these disorders as due to TBI if evidence from the history, physical examination, or ancillary studies suggests that the disturbance is the direct physiological effect of brain trauma (Table 2.5–1).

EPIDEMIOLOGY

In the United States, in 2009, approximately 2.4 million individuals experienced a TBI significant enough to require medical attention or to result in death. Closed head injuries constitute the vast majority of TBIs, with penetrating head injury occurring much less frequently. Approximately 80 percent of patients with TBI have mild head injury, 10 percent have moderate head injury, and the remaining 10 percent severe.

Most TBIs occur among adolescents and young adults, with a second peak in elderly individuals. There is also a significant gender difference. Males are 40 percent more likely to suffer brain injury than females. African Americans and individuals of low socioeconomic status also experience higher rates of TBI than other groups. In the United States, falls represent the most common cause of TBI, both overall and in the specific populations of children and older adults. Among younger adults, motor vehicle accidents and assaults produce the most TBIs.

Patients with TBI have a greater risk of recurrent TBI than noninjured controls; among individuals with TBI, about 10 percent will sustain a

second TBI within the year following the initial event. The single greatest risk factor for TBI, however, is alcohol/drug abuse. One epidemiological study reported that close to one-third of brain injury patients had an identifiable alcohol problem before trauma, and more than 50 percent were intoxicated at the time of injury.

During the past years, TBI has been described as the "signature wound" of Operation Iraqi Freedom (OIF), Operation Enduring Freedom (OEF), and Operation New Dawn (OND), the conflicts involving US troops in Iraq and Afghanistan. Approximately 10 percent of veterans of these conflicts seeking care at VA medical facilities experienced a TBI while deployed. Blast exposure represents the most common cause of OIF/OEF-related TBI (88 percent). For instance, improvised explosive devices (IEDs) or mines produced the vast majority of injuries (97 percent) observed among a Marine unit in Iraq. In addition, an estimated 59 percent of soldiers with blast injuries sustain a TBI.

PATHOLOGY

Types of Primary and Secondary Brain Injury

The first level of categorization of head injuries divides them into closed or penetrating injuries, depending upon the integrity of meningeal coverings. Missile wounds are the most frequent cause of penetrating brain injuries. However, falls are the most frequent cause of closed trauma, which represents the majority of TBIs. Other causes of closed trauma include motor vehicle collisions, assault, and sport-related injuries. In addition, blast injuries due to IEDs, the signature wound of military operations in Iraq and Afghanistan, became the focus of intense research. Blast injuries preferentially involve the lungs, tympanic membranes, and eyes, and they produce concussion/contusion of the CNS. About 50 percent of soldiers wounded in Iraq have had TBIs of different severities.

Primary brain damage results from contact and inertial forces that occur at the time of injury. Contact forces may result in laceration to the scalp, skull fractures, intracranial hemorrhages, contusions, and intracerebral hemorrhages. Inertial loading consists of acceleration/deceleration and rotational forces that result in diffuse axonal injury and eventually acute subdural hematoma from the tearing of subdural bridging veins. Secondary brain damage is produced by pathological processes that are initiated at the moment of injury but span a variable period following the traumatic episode. These include brain damage secondary to ischemia (e.g., that resulting from associated hypotension and/or hypoxia), brain swelling, raised intracranial pressure (with consequent reduction of cerebral perfusion pressure), and infection.

Focal lesions consist of contusions and lacerations that usually occur at the surface of the brain. They are more prominent at the crest of cerebral

gyri and have a predilection for the frontal and temporal poles. Lacerations are usually accompanied by extracerebral hemorrhages (burst lobe). Intracranial–extracerebral hemorrhages occur in the epidural space (i.e., between the skull and the dura), the subdural (i.e., between the dura and the arachnoid), or the subarachnoid space. Intracerebral hemorrhages are often multiple, involving the frontal and temporal lobes and basal ganglia, and may have a delayed onset (i.e., hours or days after trauma). Finally, ischemic damage may produce focal lesions from traumatic vascular lesions (e.g., arterial dissection, vascular distortion and compression, or venous thrombosis) or from arterial vasospasm.

Diffuse lesions include traumatic axonal injury (TAI), which occurs preferentially within the corpus callosum, thalamus, and dorsolateral quadrants of the upper brain stem. Pathological processes associated with TAI include fragmentation of the axolemma, axonal transport disruption, axonal bulb formation, astrogliosis, and microglial activation. Although usually associated with immediate and persistent coma, lesser degrees of axonal disruption may also be seen in patients with a lucid interval or even in patients who suffered mild brain injuries (Fig. 2.5–1). Diffuse injury also includes diffuse ischemic damage (DID). DID is highly prevalent among patients with severe head injuries. In a series of 151 pathologically examined cases, 27 percent showed severe ischemic damage, 43 percent showed moderately severe ischemic damage, and 30 percent showed mild damage.

Blast injuries are complex in nature and result mainly from the biphasic high pressure wave produced by the explosion, which can produce mechanical tissue damage as well as injury due to the release of heat, acoustic, and electromagnetic energy; on top of the injury from the initial pressure wave, further damage can occur from reflection of the wave when it hits the skull. In addition, the person may be hit by an object set in motion by the blast or may be blown off against a stationary object as a result of the explosion. Finally, cerebral infarction may result from air embolism due to blast lung injury. The structural brain changes produced by blast injuries continue to undergo thorough investigation, though the evidence so far indicates that blast injury can produce diffuse axonal injury and other white matter abnormalities.

How these different patterns of injury influence psychiatric disorders following TBI has not received extensive investigation. Previous studies have stressed the importance of diffuse brain injury as a predictor of long-term disability of patients with TBI. Another study, however, did not find significant differences in neuropsychological outcome between two groups of patients with focal or diffuse patterns of injury. Furthermore, both focal and diffuse lesions usually coexist in patients with TBI.

Neurobiological Changes

TBI generates widespread neuropathological changes in diverse cortical areas, subcortical nuclei, and white matter tracts. Experimental models of TBI demonstrate cell death in the hippocampus and other medial temporal lobe structures that occur through both necrotic and apoptotic cascades. Furthermore, evidence from these models suggests that diffuse neuronal damage and cell loss may progress over weeks to months after the initial insult in selectively vulnerable regions of the prefrontal cortex, hippocampus, thalamus, striatum, amygdala, and forebrain nuclei. In addition, chronic axonal changes occur following experimental TBI that are not limited to Wallerian degeneration. Other complex delayed post-TBI processes include microglial activation and the release of inflammatory cytokines as well as mechanisms of repair and regeneration such as reactive synaptogenesis and axonal sprouting.

The complex interplay between neuronal degeneration and repair processes that takes place after TBI produces long-lasting changes in neuronal activity and neuronal circuitry. Moreover, the functional changes observed in neuronal circuitry may constitute the neurological substrate of the cognitive and behavioral deficits frequently seen following TBI. The chronology of these changes suggests that there is a relatively broad time window within which intervention may be possible to ameliorate the short- and long-term behavioral consequences of brain trauma. However, no clear-cut neuroprotective therapies for TBI currently exist. For instance, although early studies found a benefit to progesterone in acute TBI, a large phase III double-blind RCT did not confirm these results, demonstrating that progesterone does not improve outcomes in TBI.

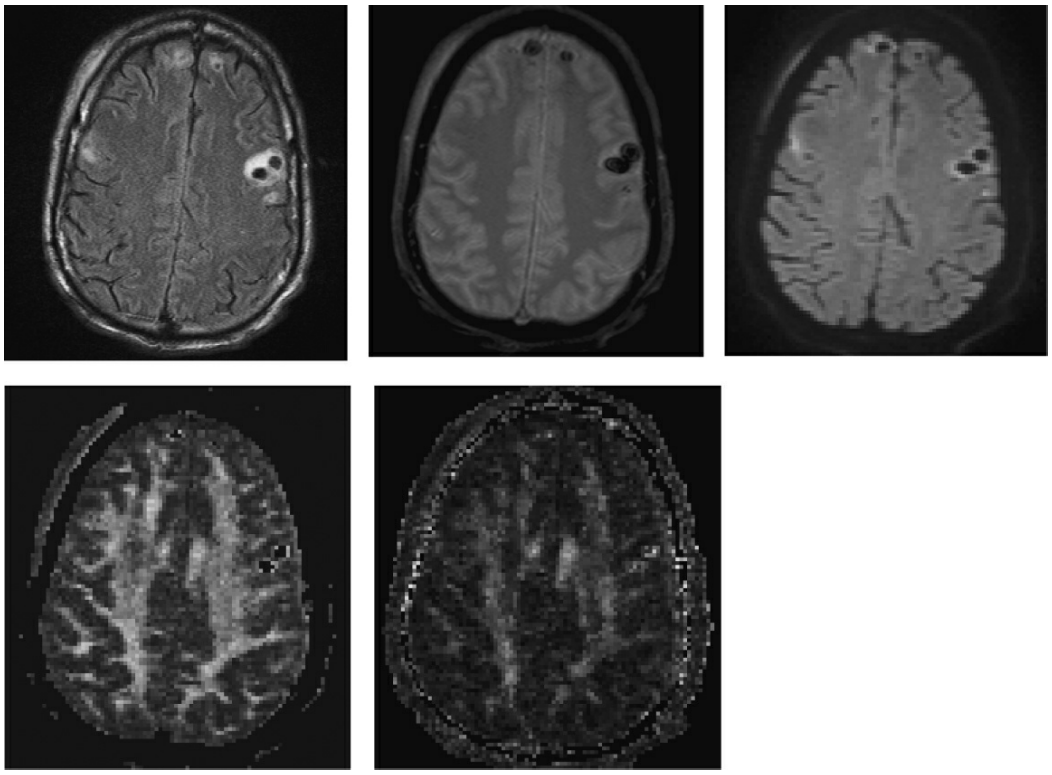

FIGURE 2.5–1. Diffuse axonal injury. Magnetic resonance images consistent with the presence of diffuse axonal injury at the gray–white matter junction.

Numerous studies have examined post-traumatic changes in the major neurotransmitter systems in the brain. Glutamate has been widely studied because of its role in excitotoxic neuronal injury. Excitotoxic injury has a sequential mechanism, beginning with sodium and chloride influx with resultant cytotoxic edema, and followed by calcium influx leading to increased expression of early transcription factors, acute phase proteins, and caspases, the proteolytic enzymes that mediate neuronal apoptosis. Clinical studies have reported that glutamate concentrations are significantly elevated for several days in the CSF of patients with TBI. Thus, timely pharmacologic intervention could conceivably prevent excitotoxic injury. Glutamate antagonists such as ketamine have shown a beneficial effect in experimental models of TBI. In addition, inhibitors of glutamate release, such as riluzole, may be an alternative to postsynaptic glutaminergic blockade, which is associated with severe psychiatric side effects. However, data on the use of these agents remains very preliminary.

Cholinergic neuronal activity appears to increase immediately after TBI. Blockade of massive acetylcholine release resulting from pathological excitation of basal forebrain nuclei at the time of injury may prevent neuronal cell loss and associated behavioral deficits. There is also evidence of a hypofunctional cholinergic state occurring later in the post-TBI course. A reduction in cholinergic transmission in hippocampal and neocortical areas has been observed following cortical contusion brain injury. In addition, experimental models in rats demonstrate dysfunction of the septo-hippocampal cholinergic pathway, which might play a significant role in the development of post-traumatic cognitive and behavioral deficits. Although cholinergic systems have not received systematic study in clinical populations of patients with TBI, cholinergic deficits observed in patients with Alzheimer disease have been associated with behavioral changes including lack of motivation, anhedonia, as well as agitation and disinhibited behavior. Preliminary clinical data—including some small RCTs—suggests that cholinesterase inhibitors improve cognition in TBI.

Ascending biogenic amine pathways may also influence the pathophysiological processes underlying the clinical presentation and even the long-term outcome of patients with TBI. Circulating levels of catecholamines correlate with TBI severity as measured by the Glasgow Coma Scale (GCS) score. One study found that patients with TBI showed an increase in both serotonergic and noradrenergic metabolites in the CSF. The investigators hypothesized that prolonged increase in the synaptic concentration of these neurotransmitters results in subacute or chronic downregulation of aminergic transmission and eventually depressive symptoms.

Dysregulation of mesolimbic and mesocortical dopaminergic pathways —perhaps as a result of prefrontal damage—might also underlie the mechanisms underlying manic and hypomanic syndromes. In addition, aminergic neurotransmitters may affect the restorative processes occurring in the chronic phase of TBI, possibly via neurotrophic factors.

LABORATORY TESTS

During the past few years, the search for diagnostic and predictive serum markers for TBI has inspired much research. These markers include both neuronal (e.g., neuron-specific enolase, cleaved Tau protein, neurofilament, ubiquitin C-terminal hydroxylase-L1) and glial (e.g., S-100B, glial fibrillary acidic protein) proteins. Serum S-100B is reported to have 90 percent sensitivity and 99 percent sensitivity for detecting intracranial pathology on head CT scans. In addition, serum S-100B levels greater than 2.0 μg/L predicted unfavorable outcome among patients with moderate-to-severe brain injury even more accurately than traditional clinical measures such as the GCS. In mild head injury, S-100B produces much less consistent results, though one study found that early detection of serum S-100B levels greater than 0.2 μg/L predicted long-term neuropsychological dysfunction, even in cases with negative MRI.

Ideally, serum markers could objectively diagnose TBI in symptomatic patients with normal neuroimaging findings; however, the most robust findings for biomarkers have occurred in patients with severe TBI, in whom the diagnosis is not likely in doubt. Sustained or secondary release of serum markers may prove useful in detecting disease progression or monitoring treatment response. Early identification of patients at high risk of long-term neurobehavioral dysfunction could greatly assist with the development of effective therapeutic interventions for TBI.

Neuroendocrine dysfunction may accompany head injury, particularly pituitary hormone deficiencies, as the pituitary's location makes it vulnerable to injury in TBI. Diabetes insipidus due to antidiuretic hormone (ADH) deficiency manifests clearly in the acute stage, but other pituitary hormone deficits may present insidiously, escaping detection for months or years. The emergence of the syndrome of growth hormone deficiency (GHD) in adults—characterized by decreases in strength, energy, and sense of well-being—highlights the potential importance of such deficiencies. Estimated rates of GHD following TBI range from 10 to 30 percent, making it likely the most common endocrinopathy; rates of luteinizing hormone (LH), follicle-stimulating hormone (FSH), adrenocorticotropic hormone (ACTH), thyroid-stimulating hormone (TSH), prolactin, and ADH abnormalities range from 2 to 17 percent. About 3 to 5 percent of patients with TBI may show multiple neuroendocrine abnormalities.

Electrophysiological Studies

Conventional EEG recordings are currently used in trauma intensive care units for monitoring procedures and for brain death diagnosis. EEG is also invaluable in the diagnosis of status epilepticus and is the primary diagnostic procedure to localize a post-traumatic epileptic focus. The use of nasopharyngeal, anterior temporal, or sphenoidal leads increases the chance of finding an abnormality on routine EEG. Video-EEG monitoring and 24-hour ambulatory recordings aid in the differential diagnosis of patients presenting with unclear paroxysmal behavioral disturbances. Quantitative EEG (Q-EEG) is currently used as an adjunctive diagnostic technique in the evaluation of slow-wave abnormalities associated with brain injuries and in the diagnosis of post-traumatic temporal lobe epilepsy. Polysomnography permits the diagnosis of atypical sleep disturbances that may occur in patients with TBI. These include atypical night terrors, sleep apnea, nocturnal myoclonus, and restless leg syndrome. Visual, auditory, and somatosensory evoked potentials help localize the level of injury to the CNS. Auditory evoked potentials are particularly effective in detecting brain stem pathology. Recently, interest in more complex electrophysiological responses has increased. One study using P50 physiology found evidence of deficits in sensory gating mechanisms in patients with TBI. The P50 ratio might serve as a marker of hippocampal dysfunction in this population.

Neuroimaging

Recent advances in neuroimaging techniques have improved the current understanding of the relationship between structural abnormalities and behavioral disturbances following TBI. CT and MRI are routinely used for the evaluation of patients with TBI. CT remains the most efficient means of detecting surgically treatable hematomas and is the study of choice for evaluating patients with rapid changes in their neurological status. MRI, however, shows more sensitivity in detecting the more prevalent post-traumatic nonhemorrhagic lesions (e.g., cortical contusions and deep white matter lesions) and in identifying small subdural collections. A delayed MRI scan (i.e., 2 weeks after injury) would be indicated in those patients whose initial instability favored an emergency CT scan but who received medical treatment or who continued to have persistent neurological deficits after a neurosurgical procedure.

CT and routine MRI lack the sensitivity to identify the whole extent of the pathological changes associated with TBI. Structural volumetric MRI studies, however, provide quantitative data on volumetric changes observed in gray matter, white matter, and CSF. For instance, Francesco Tomaiuolo and colleagues reported that patients with TBI who sustained severe closed head injury had significant volume reduction in the hippocampus, fornix, and corpus callosum compared with control subjects matched in terms of gender and age. Immediate and delayed recall of word

lists and visuospatial designs correlated with the volume of the fornix. Other investigators used voxel-based morphometry techniques to demonstrate reduced gray matter volumes in the basal forebrain, the hippocampal formation, and regions of the neocortex from 31 patients with moderate-to-severe TBI compared with healthy control of similar demographic characteristics. These structural changes were associated with the presence of memory impairment and, according to the authors, were consistent with cholinergic dysfunction. Elisabeth A. Wilde and colleagues used quantitative MRI techniques to analyze the effect of prior alcohol abuse on brain volumetric measures obtained after TBI. The authors observed increased general brain atrophy among patients with a history of moderate-to-heavy pre-injury alcohol use compared to patients with TBI who were abstinent or had a history of minimal alcohol use. This study gives some objective evidence of the synergistic effects of trauma and alcohol neurotoxicity on structural brain damage following TBI. Use of diffusion weighted imaging (DWI), apparent diffusion coefficient (ADC), gradient echo (GRE), and susceptibility weighted imaging (SWI) sequences also improves detection of axonal injury.

DTIis a sensitive method to study white matter changes after TBI. DTI provides an objective measure of directional water diffusion in space (e.g., fractional anisotropy [FA]), which in white matter regions reflects the integrity of myelin sheaths. White matter regions with reduced anisotropy appear in the first 24 hours after TBI. DTI reveals changes in the white matter that correlate with the severity of TBI and with functional outcome at the time of hospital discharge. One longitudinal study found that FA increased as clinical improvement occurred following TBI. Overall, these findings suggest that DTI may be a powerful technique for in vivo detection of TAI. Diffusion tensor tractography, which allows visualization and analysis of specific white matter pathways, offers promise for extending the potential of DTI in TBI; studies thus far have shown changes in corpus callosum tracts in mild, moderate, and severe TBI.

Functional imaging studies, such as fMRI, PET, single photon emission computed tomography (SPECT), and MRS can help detect brain changes not seen on structural imaging, making this a useful modality for exploring TBI.

Much recent imaging research in TBI has focused on fMRI, both task-based and default mode network (DMN), also known as resting state. Patients with mild, moderate, and severe TBI show abnormalities in the DMN; these abnormalities improve as clinical recovery occurs, and correlate with severity of cognitive impairment. In task-based fMRI studies of working memory and attention, individuals with TBI show different patterns of regional brain activation than healthy controls, even when performance on these cognitive tasks does not substantially differ from normal. This suggests that injury-related changes in the allocation of

working memory and attentional resources might underlie some of the cognitive complaints after mild head injury. Consistent with these findings, another study using an fMRI paradigm to explore functional changes in working memory circuits found that patients with TBI showed reduced activation—associated with increased processing load—in prefrontal and parietal cortical areas. Finally, investigators in Canada reported structural and functional changes in the anterior cingulate cortex in a group of depressed athletes with a history of mild TBI compared with a group of nondepressed athletes with TBI of similar characteristics.

One ^{18}F-FDG-PET investigation demonstrated significant differences in glucose metabolism in the thalamus, brain stem, and cerebellum between comatose and noncomatose patients acutely after severe TBI. It has also been shown that the ratio of gray matter to white matter cortical metabolic glucose rates correlates with initial GCS scores of patients with TBI. Furthermore, patients with higher ratios showed good recovery 12 months after TBI. However, earlier PET studies found only modest correlation between the longitudinal changes in cortical glucose metabolic rates and the degree of global functional recovery following TBI.

One study, using PET and SPECT to examine blood flow and metabolic changes among patients with mild TBI and persistent post concussive symptoms, found medial abnormalities. A recent meta-analysis found that SPECT improved the identification of mild TBI, with an almost 100 percent negative predictive value; frontal and temporal lobe abnormalities represented the most common finding.

MRS techniques acquire spectra simultaneously over a wide brain region, enabling the generation of maps of metabolite distributions (Fig. 2.5–2). Multislice short echo time (TE) H_1-MRS allows for the in vivo evaluation of the metabolic profiles of large brain regions including the neocortex, deep white matter, and basal ganglia. Specific metabolites studied with H_1-MRS include *N*-acetyl aspartate (NAA), a marker of neuron density and metabolism; choline-containing compounds (Cho), which reflect phospholipid metabolism; creatine-containing compounds (Cre), which reflects the cellular energetic status; myo-inositol (mI), a putative glial cell marker; the sum of glutamate and glutamine (Glx); lactic acid (Lac); and other peaks. M.R. Garnett and colleagues reported that, when compared to controls, patients with TBI showed significantly reduced NAA levels at the posterior part of an otherwise normal-appearing frontal lobe at the acute (~12 days) and chronic (~6 months) stages of TBI. In addition, the NAA levels dropped further between the acute and chronic stages in 10 patients imaged at both time points. Another recent study reported decreased NAA/Cho ratios in the basal ganglia and medial temporal region in patients with moderate-to-severe head injuries. The basal ganglia ratio correlated to measures of speed, motor scanning, and attention. Spectroscopic techniques have been also used to study mild traumatic

injury. For instance, a recent study showed evidence of widespread metabolic changes following mild TBI in regions that appear normal on diagnostic MRI scans. Finally, initial spectroscopic variables may help predict long-term cognitive outcome after pediatric TBI. However, one recent study found that SPECT outperformed MRS for predicting outcomes in mild-to-moderate TBI.

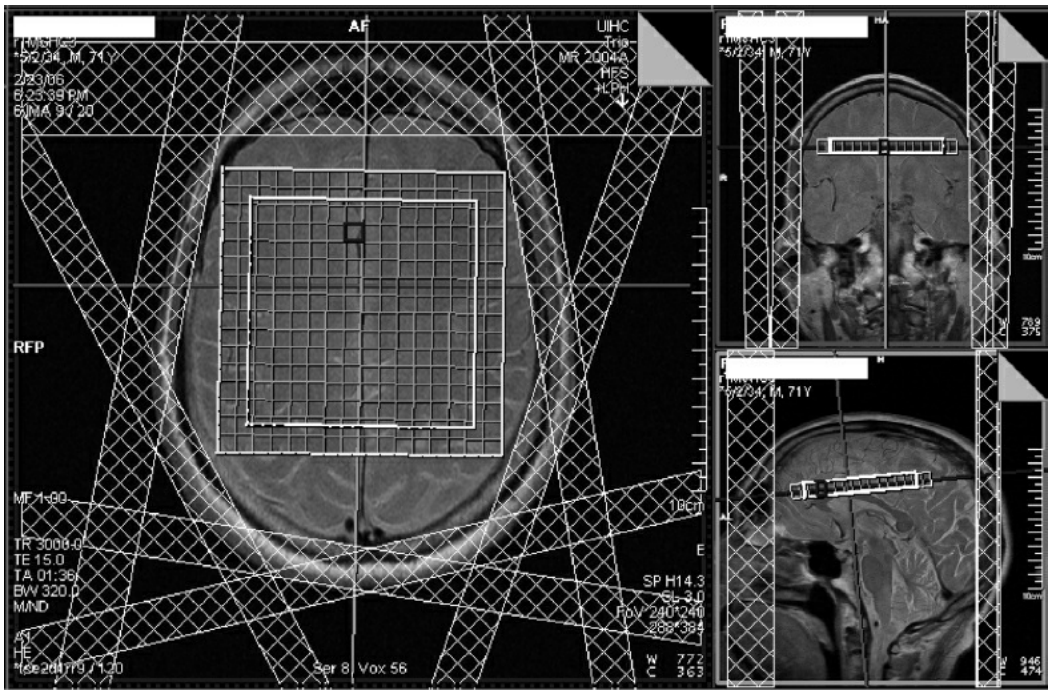

FIGURE 2.5–2. Proton spectroscopy dorsal–rostral cingulate gyrus PRESS CSI.

Functional imaging studies also have the potential to identify those brain areas involved in processing cognitive and emotional information as well as the areas activated during the implementation and monitoring of behavioral responses. PET and fMRI may further delineate the neural circuits involved in mood and cognitive regulation, and assess the effect of brain injury upon these neural circuits. Refined functional imaging studies could potentially guide the design of therapeutic interventions and even predict patients' response to treatment.

Consequences of TBI

Approximately 7 percent of patients hospitalized for TBI will die from their injuries, and over 5 million Americans experience ongoing disability due to TBI (MMWR 2013, MMWR 2007). Individuals who survive a TBI face an elevated risk of subsequent stroke and increased all-cause mortality. Direct and indirect costs of TBI exceed $76 billion per year (expressed in 2010 U.S. dollars).

Neuropsychiatric disorders are probably the most frequent complication of TBIs. One study compared the frequency of psychiatric

disorders between 939 patients with TBI and 2,817 controls enrolled in an adult health maintenance organization. The prevalence of any psychiatric illness in the first year was 49 percent following moderate-to-severe TBI, 34 percent following mild TBI, and 18 percent in the control group. Among patients without a history of psychiatric illness, the adjusted relative risk of any psychiatric illness in the 6 months following moderate-to-severe TBI was 4.0 and following mild TBI was 2.8 compared with those without TBI. Among subjects with previous psychiatric disorders, the adjusted relative risk of any psychiatric illness in the 6 months following moderate-to-severe TBI was 2.1 and following mild TBI was 1.6. Prior psychiatric illness was a significant predictor of psychiatric morbidity following TBI. Furthermore, the prevalence of psychiatric disorders remains significantly higher in patients with TBI than those in control groups many years after the traumatic injury.

The possibility that TBI might affect the onset, phenomenology clinical course, and even treatment response of major psychiatric illnesses is an exciting area of future research.

General Principles for Treatment of Neuropsychiatric Complications of TBI

Treatment of neuropsychiatric disorders occurring after TBI involves various strategies. Presumably, treatment of the neurobehavioral consequences of TBI should begin early in the acute phase following injury. If it is possible to modify the processes associated with acute and delayed neuronal damage, then the intervention should start as early as possible, in order to produce the greatest amount of recovery in cognition, motivation, activity levels, and emotional disorder. Although progress in basic research allows one to envision a promising future for therapeutic intervention following TBI, there is a lack of adequately controlled clinical studies providing a solid scientific basis for psychiatric treatment. Currently, anecdotal cases, small uncontrolled trials, and clinical experience are the only things that support many daily treatment decisions.

Patients with brain injury experience more sensitivity to medication side effects, especially with psychotropic drugs. Jonathan M. Silver and Stuart C. Yudofsky proposed several general guidelines for their use in this population. Doses of psychotropic drugs must be prudently increased, minimizing side effects (i.e., start low, go slow). However, the patient must receive an adequate therapeutic trial with regard to dosage and duration of treatment. Clinicians should frequently reassess brain-injured patients in order to determine changes in treatment schedules, and take special care in monitoring drug interactions. Finally, a partial response to a specific medication may warrant augmentation therapy, depending upon the augmenting drug's mechanism of action and potential side effects.

Behavioral deficits in self-care habits (e.g., feeding or personal hygiene), interpersonal skills (e.g., disinhibited behavior), problem solving, or

response to environmental stress may respond to behavioral intervention. Behavioral rehabilitation programs shape behavior based on the principles of operant conditioning (e.g., contingency contracts and token economies). Their goal is to increase the patient's repertoire of social and independent living skills, generalizing their use from the rehabilitation environment to the more demanding conditions of community life.

Psychotherapy forms an essential part of the treatment strategies implemented to improve the long-term outcome of patients with TBI. Preliminary data suggest that peer support and family therapies enhance coping for both patients with TBI and their family members. Behavioral interventions such as the Differential Reinforcement of Other Behavior (DRO) may successfully reduce the frequency of problematic behavior. Psychotherapy groups implemented in post-acute rehabilitation settings may focus on the treatment of substance abuse and anger management through education, social support, and development of interpersonal skills. Additionally, one study demonstrated the feasibility of using the telephone as a means of providing education and psychotherapeutic support during the first year after moderate-to-severe TBI.

CLINICAL FEATURES

Acute Behavioral Consequences of Traumatic Brain Injury

Head injuries span a wide range of severity, from patients who die at the moment of trauma to those who do not require medical evaluation or assistance. Most of the patients admitted to the hospital with a head injury diagnosis have a mild injury. A minority of these mildly affected patients will develop acute complications (e.g., brain swelling, delayed hematoma, or intracranial infection) or prolonged post-concussion symptoms. Neuroimaging studies (CT and MRI) have demonstrated the presence of structural brain lesions in some mild head injury patients who have not experienced clinical complications.

The most common consequence of head injury is impairment of consciousness, ranging from transient confusion to protracted coma. The GCS is commonly used to grade the severity of TBI. The scale gives a quantitative estimate of level of consciousness and neurological status based on patterns of eye opening, as well as best verbal and motor responses. GCS scores between 13 and 15 define mild brain injury, scores between 9 and 12 define moderate injury, and scores between 3 and 8 define severe injury.

The early phase of recovery from TBI is characterized by disorientation, confusion, and impaired memory function. Apathetic withdrawal, agitation, or severe delirium may also be observed in these patients.

PTA occurs during the period when the patient (who is usually emerging from coma) is disoriented, confused, and has disrupted memory

functioning. Deficits are observed in declarative memory (i.e., memory of recent events and times), affecting both anterograde and retrograde processes. Procedural memory, in contrast, appears to be relatively spared. Duration of PTA has been widely used as a measure of TBI severity. It may be assessed using the Galveston Orientation and Amnesia Test (GOAT), which evaluates orientation to person, place, and time, as well as awareness of the accident and its consequence. Alternatively, retrospective structured questionnaires have shown an excellent correlation with this prospective determination. Duration of PTA has proved to be a good predictor of the degree of disability, vocational outcome, and severity of personality change following TBI.

Clinical features of the early phase of recovery from TBI involve more than just memory impairment. Patients frequently have a decreased level of consciousness and meet DSM-IV-TR criteria for delirium. In addition, patients may present with perceptual disturbances (i.e., illusions or hallucinations), delusional thoughts, psychomotor agitation or retardation, affective lability, and vegetative symptoms (e.g., tachycardia, hypertension, diaphoresis, and sleep–wake cycle disruption). Symptoms usually have an acute onset and a fluctuating course. Delirium is most frequently observed in severe TBI cases.

A 19-year-old man was admitted to a trauma center after a motorcycle accident. He presented with a right epidural hematoma and bilateral contusions in the anterior temporal lobes. He also had a scalp laceration and right maxillary and zygomatic fractures. The post-resuscitation GCS score was 8. Following surgical evacuation of the hematoma, the patient transferred to the intensive care unit. Two days later, the patient became restless and agitated. He removed his intravenous (IV) lines and monitoring devices and tried to get out of bed. He was disoriented, incoherent, and aggressive. His behavior suggested that he was experiencing frightening visual hallucinations. He appeared to grab at nonexistent objects in the environment, and picked at the bedclothes. He responded well to a short course of high-potency antipsychotic agents.

Multiple conditions may contribute to the development of delirium in patients with TBI. These include structural brain damage, cerebral edema, brain hypoxia, seizures, electrolyte imbalance, infections, medications (e.g., barbiturates, opiates, or steroids), and drug or alcohol withdrawal. Old age, coexistent severe medical disease, polypharmacy, basal ganglia, and right hemisphere lesions have also been shown to be significant risk factors.

Agitation poses a frequent and significant problem in acute rehabilitation settings. Prospective studies among patients with TBI admitted to rehabilitation centers have found that one-third to one-half experience agitation, though it most commonly resolves within 10 days. Agitated patients experience longer hospital stays and are less likely to return home on discharge.

Another study found that 59 out of 85 patients with TBI (69.4 percent)

admitted to a rehabilitation unit met DSM-IV-TR criteria for delirium (Nakase-Thompson 2004). In approximately one-third of these patients, delirium had a protracted course. The authors emphasized that clinical measures that focus on orientation and anterograde memory functions do not address the complex phenomenology of acute confusional states observed among patients with TBI. Furthermore, agitated patients make less progress in rehabilitation not only because of greater injury severity but also because agitation disrupts engagement in rehabilitation therapies.

Although evidence from animal studies suggests that both typical and atypical neuroleptics might have a negative impact upon cognitive recovery, the relationship among the duration, clinical features (e.g., the presence of psychotic symptoms or seizure activity), and treatment of acute agitation syndromes (e.g., anticonvulsants or atypical antipsychotics) with the outcome of TBI has not been adequately studied. Systematic investigations of the use of antipsychotic agents to treat post-TBI agitation and delirium remain needed.

While prophylactic administration of antiepileptic drugs (AEDs) for the first 7 days following severe TBI effectively reduces the incidence of seizures during this timeframe, longer administration of AEDs or administration in mild or moderate TBI has no effect on subsequent development of post-traumatic epilepsy. As AEDs can have significant side effects, and may potentially inhibit the neuroplasticity needed for optimal recovery, clinicians should avoid prophylactic use of AEDs beyond 7 days postinjury in patients with severe TBI, and avoid prophylactic use completely in patients with mild or moderate TBI.

Chronic Behavioral Consequences

The chronic neurobehavioral consequences of TBI can be studied from a dimensional perspective using neuropsychological tests and behavioral scales that have been extensively validated in acute care settings and in rehabilitation services. However, cognitive and behavioral morbidity can be also assessed from a categorical, disease-based perspective, which assumes that psychiatric disorders, although diagnosed through a recognized constellation of symptoms, have an identifiable biological substrate, a distinct clinical prognosis, and a predictable treatment response.

Cognitive Disorders

Cognitive disorders following TBI can be divided into two categories—those which begin at the time of injury, and those which develop years to decades later. The first category is considered major or mild neurocognitive disorder due to TBI in the DSM-5; the second category is not well-represented in the DSM-5, but includes chronic traumatic encephalopathy (CTE) as well as TBI-related increased risk of other specific

neurodegenerative dementias of likely multifactorial etiology, such as Alzheimer disease.

Major or Mild Neurocognitive Disorder due to TBI. The DSM-5 defines major neurocognitive disorder due to TBI by the following: significant decline in at least one cognitive domain, interfering with everyday activities, and not occurring solely in the context of delirium; a history of head trauma accompanied by loss of consciousness, PTA, disorientation and confusion, and/or neurological changes; and symptoms beginning at the time of the injury and continuing after the acute post-injury phase. Mild neurocognitive disorder due to TBI follows the same criteria, except that the decline in cognitive domain(s) is modest and does not interfere with everyday life. Neurocognitive disorder due to head trauma is characterized by prominent memory and executive dysfunction with relatively preserved visuospatial, praxic, and primary linguistic functions. In addition, these patients may be severely apathetic and withdrawn and demonstrate markedly slow information processing. Physical examination may reveal the presence of extrapyramidal signs. Of note, a chronic subdural hematoma in the elderly may present as a progressive dementia.

A 40-year-old white man sustained severe injuries in a motor vehicle accident and experienced protracted coma of 1-month duration. An MRI performed at 2 weeks from injury revealed the presence of widespread diffuse axonal injury. At 6-month follow-up, the patient presented with a mild left hemiparesis, right hemidystonic symptoms, and a left peripheral facial palsy. Neuropsychological testing disclosed substantial memory deficits, frontal lobe dysfunction, and significantly impaired problem-solving ability. Visuospatial and linguistic skills ranked within the lower average range. His hygiene and self-care were poor, and he hoarded garbage in his pockets and under his bed. He had frequent bursts of severely aggressive behavior, but, overall, he remained abulic and withdrawn. Treatment with amantadine effectively controlled his aggressive behavior.

Cognitive disturbances are one of the most important long-term effects of severe TBI. A seminal study reported on the cognitive outcome of 127 severe brain-injured patients who could complete serial neuropsychological assessments during a 1-year follow-up period (i.e., excluding those patients with a persistent vegetative state or with very severe intellectual impairment). At 1-year follow-up, the brain-injured patients showed slower information processing and greater impairment in memory function compared with a neurologically intact control group. In contrast, their linguistic and visuospatial abilities were within the normal range.

While essentially all researchers and clinicians agree that moderate and severe TBI can produce significant, permanent cognitive changes, controversy exists as to whether mild TBI can as well. These patients complain of concentration and memory deficits during the first weeks following TBI, and some studies have found objective cognitive deficits

persisting up to 3 to 6 months post-injury. However, the majority of these patients spontaneously recover. It is possible, though, that the subtle frontal-executive deficits patients some patients report experiencing in real-world situations over the long term may not be well captured by standard cognitive batteries.

Repetitive mild head injury such as those observed among certain athletes (e.g., soccer and football players) requires special attention. A study of college football players showed that a history of multiple concussions was associated with reduced cognitive performance. In addition, amateur soccer players with repeated concussions show deficits in memory and executive functions.

Attention deficits are among the most frequent neuropsychological symptoms observed in patients with TBI following resolution of PTA. Attention consists of multiple processes based in the activity of interrelated neural networks. Patients with TBI may present with restricted verbal or visuospatial attentional span, altered vigilance patterns (i.e., sustained attention deficits), or slowed information processing. The most consistent findings, however, are associated with performance in the most demanding tasks (e.g., in divided attention paradigms such as the Paced Auditory Serial Addition Task).

TBI also produces distinct impairment in memory functions. Memory deficits are the most frequent cognitive disturbances reported by patients and relatives in the chronic phase of TBI. Memory dysfunction encompasses both anterograde and retrograde deficits, faulty sequencing of events, and inefficient encoding and storage strategies. For instance, HarveyLevin and Felicia Goldstein demonstrated that, when compared with control subjects, patients with TBI failed to organize recall of words by clustering them in appropriate semantic categories.

TBI also frequently affects linguistic competence. Approximately one-third of severely brain-injured patients admitted to a rehabilitation facility showed fluent (51 percent), nonfluent (35 percent), or global (14 percent) aphasic syndromes. Aphasia tends to resolve in the majority of cases during the first year following trauma. Anomia, however, constitutes the most prevalent long-term linguistic deficit following trauma. Patients with TBI may also have high-order language alterations and present with a defective narrative discourse, a lack of semantic coherence, aprosody, and impaired pragmatics of communication. All of these result in impoverished and disorganized language and in reduced communication proficiency.

A prominent defect in control or executive functions has been consistently described in patients surviving severe head injury. Executive functions include goal formation, planning, selection of adequate response patterns, and monitoring of ongoing behavior. Several neuropsychological tasks were specifically designed to quantify these deficits. These include the Wisconsin Card Sorting Test, the Goldberg Executive Control Battery, the

Tower of London Test, and the Trail Making Test. Simple cognitive tests such as the verbal and figural fluency tasks can document the lack of spontaneity and perseverative tendencies frequently observed in patients with traumatic brain damage. The executive dysfunction observed in patients with TBI strongly correlates with dysfunction of frontosubcortical pathways. When confronted with a demanding environment, the adaptive functioning of patients with TBI is also often impaired.

In contrast to what happens with memory and control functions, visuospatial and praxic abilities are usually preserved during the chronic phase of TBI, probably due to the relative sparing of posterior association cortices in TBI. Finally, patients with TBI frequently exhibit unawareness (anosognosia) or denial of deficits, particularly those who have suffered extensive frontal lobe damage. Unawareness of deficits impedes realistic goal setting and interferes with the rehabilitation process.

The type and extent of brain damage determine cognitive impairment following TBI. Genetic factors may also play a significant role in cognitive outcome, perhaps because of their effects on repair processes. For instance, studies among patients with severe brain injury demonstrated a strong association between the APOE-epsilon4 allele and a poor clinical outcome, implying genetic susceptibility to the effect of brain injury. However, other studies in mild and moderate TBI have not found that APOE-epsilon4 allele affects the recovery of patients with less severe types of injury.

Optimal treatment for post-TBI cognitive impairments may involve both medications and cognitive training. Prodopaminergic agents may modestly improve arousal and speed of information processing, reduce distractibility, and improve some aspects of executive function. This group of drugs may also ameliorate depressive symptoms. Nine double blind, randomized placebo-controlled studies of stimulants (eight using methylphenidate and one using lisdexamfetamine), have shown that these drugs improve overall cognitive function as well as attention, executive function, and memory. Another RCT found a cognitive benefit for bromocriptine in TBI. Amantadine, a drug with complex pharmacological effects on dopaminergic, cholinergic, and NMDA receptors, might improve motivational deficits; of the two double-blind RCTs of amantadine in TBI, one showed a benefit and one did not. Three double-blind placebo-controlled randomized trials have found that cholinesterase inhibitors, such as donepezil and rivastigmine, can improve cognition in TBI. These agents may also potentially improve motivation and general well-being.

Cognitive rehabilitation strategies with a good evidence base for efficacy in TBI include direct attention training, memory strategy training and compensatory strategies (such as notebook use), and pragmatic conversation training. Metacognitive training involving self-monitoring/regulating strategies has been shown to help remediate executive function, memory, and attention. Additionally, a comprehensive

rehabilitation approach involving emotional regulation and training in emotional, social, and functional skills in addition to cognitive training appears to produce the best outcomes.

TBI and Later-Life Neurodegenerative Dementia. Overall, the majority of the evidence—including retrospective, prospective, and meta-analytic studies—indicates that TBI increase the risk of developing dementia later in life. Among studies examining the severity of TBI, moderate-to-severe TBI have generally shown an association with dementia, while the role of mild TBI remains less clear.

Most prospective studies show a significant, albeit weak, association between dementia and a history of TBI. For instance, the MIRAGE study assessed the association between probable or definite Alzheimer disease and a history of TBI among 2,233 probands ascertained at 13 centers in the United States, Canada, and Germany. Comparison of Alzheimer disease patients with their unaffected spouse produced an odds ratio (OR) of 9.9 (95 percent confidence interval [CI] = 6.5–15.1) for TBI with loss of consciousness and an OR of 3.1 (95 percent CI = 2.3–4.0) for TBI without loss of consciousness. Other investigators carried out a population-based prospective cohort study that included World War II veterans who were hospitalized during their military service with either a TBI or another unrelated condition. TBI in early adult life was associated with an increased risk of Alzheimer disease or other dementias 40 to 50 years after the TBI (relative risk [RR] = 2.32, 95 percent CI = 1.04–5.17), and this risk increased with the severity of the injury (patients with mild TBI showed no significant risk of dementia). However, results from a prospective population-based study that included 6,645 inhabitants of Rotterdam aged 55 years or older and no dementia at baseline concluded that individuals with a history of TBI did not face an increased risk of dementia (RR = 1.0, 95 percent CI = 0.5–2.0). Some studies showed that a history of TBI may result in a relatively early onset of dementia, but others have not replicated this.

The EURODEM meta-analysis showed a significant association between TBI and Alzheimer disease (OR = 1.82, 95 percent CI 1.26–2.67). This association was strongest for cases without a family history of dementia and for men. However, a history of TBI did not increase the risk of dementia for women and did not influence the age at onset of dementia. A more recent systematic review of 15 case–control studies, seven of which postdated the EURODEM meta-analysis, reported an OR of 1.58 (95 percent CI = 1.21–2.06). Once again, the relative risk was increased for men but not for women.

While several studies have shown that moderate and severe TBI increase the risk of subsequent dementia, until recently, most studies have not found an association between mild TBI and increased risk of dementia

in later life. However, one recent large retrospective study using the Korean national health insurance database found that even a single mild TBI increases the risk for later developing dementia (hazard ratio [HR] 3.26, 95 percent 2.69–3.94), independent of other risk factors. Another recent large retrospective study using the state of California's emergency department and inpatient hospitalization database found that mild TBI increases the risk of subsequent dementia in patients age 65 or older, but not younger individuals.

TBI may accelerate the onset of Alzheimer disease in predisposed individuals by diminishing the cognitive reserve. Thus, dementia may become clinically evident after neuronal loss falls below a certain threshold. Several studies suggested that TBI may initiate a cascade of molecular events usually associated with Alzheimer disease. For example, amyloid precursor protein (APP) and amyloid-β-42 (A-β-42) peptide expression could increase secondary to the axonal damage secondary to TBI. One study reported a significant reduction on A-β-42 in the CSF after severe TBI and a significant association between low A-β-42 and poor outcome. Tau levels were significantly elevated immediately after the TBI and returned to normal about 6 weeks later. Other studies reported a significant increase in tau proteins in the CSF after TBI, as well as a significant association between clinical improvement and decreased CSF tau levels. Some authors suggest that the increased availability of A-β peptides may result in amyloid deposition, whereas others believe that a low level of A-β peptides in the CSF of TBI patients may reflect diversion of A-β peptides from the CSF to cerebral deposits.

Some of the pathology of Alzheimer disease (primarily amyloid deposits, but not neurofibrillary tangles) may be overrepresented in the brains of TBI individuals. Most of these neuropathologic analyses were carried out in cases with severe TBI who died within few days of trauma. Thus, it is possible that the final extent of Alzheimer disease-like pathological changes may be greater in individuals with longer survivals. However, some of these changes (e.g., diffuse amyloid deposits) are not specific for Alzheimer disease, whereas more specific Alzheimer disease changes (e.g., mature dense-cored plaques) are rare. Future studies will have to determine the significance of these changes.

One large epidemiologic study found that TBI increases the risk of subsequent ischemic stroke; as cerebrovascular disease is a known risk factor for dementia, both vascular and neurodegenerative, this could potentially represent another pathway for TBI to predispose to dementia.

In addition to facing higher risk of AD and other neurodegenerative disorders, individuals with repeated mild TBI may develop a unique cognitive disorder known as CTE, a dementia syndrome not explainable by an alternative pathophysiologic process. However, well-validated diagnostic criteria do not yet exist and different groups use different

classification schemes. Modern descriptions of CTE may likely represent a different condition from the classically described CTE. Classic CTE, also known as dementia pugilistica and initially described in professional boxers, primarily manifests with parkinsonism and nonprogressive cognitive deficits. Modern CTE presents with prominent psychiatric symptoms, and the cognitive deficit tends to progress, though parkinsonism occurs in both conditions. Both classic and modern CTE share some common neuropathological features (such as neurofibrillary tangles and amyloid plaques), though other pathologic features may differ.

TBI may influence cognition in aging even without frank dementia. A group of investigators from Finland examined the 30-year course of cognitive symptoms among 61 patients who had a TBI between 1966 and 1972. They concluded that most of the patients showed mild cognitive decline during the follow-up period, particularly among men who were older at the time of injury. However, they observed a relative preservation of semantic memory (Hinanen 2006).

Personality Changes. Patients with TBI may experience significant personality changes. These patients have been described as irritable, childish, inconsiderate, capricious, anxious, or aggressive. They lack foresight and misjudge the consequences of their actions. Disinhibition is a frequent and striking clinical feature that may lead to antisocial behavior. However, they may become apathetic, abulic, and withdrawn.

Some investigators group these changes into two distinct syndromes: first, a pseudodepressed personality syndrome that is characterized by apathy and blunted affect and, second, a pseudopsychopathic personality syndrome displaying disinhibition, egocentricity, and sexual inappropriateness as its outstanding features.

The DSM-5 defines personality change due to TBI as a persistent personality disturbance that represents a change from the individual's previous personality profile (or a deviation of normal development in children) and is attributable to the pathophysiological changes triggered by brain trauma (Table 2.5-1). The disturbance must not occur exclusively during the course of delirium and must not be better accounted for by another mental disorder (e.g., mood disorder or substance abuse).

A 42-year-old construction worker fell from the second floor of a new building. He spent 3 days in a coma and remained amnestic and disoriented for approximately 3 weeks. A CT scan showed bilateral orbitofrontal and anterior temporal hemorrhagic contusions. Six months after the injury, he exhibited a significant personality change. He spent most of his day watching television and refused to initiate his usual activities. He ate and gained excessive weight. His wife complained of his frequent and often inappropriate sexual demands and stressed his lack of intimacy. He also became easily upset, shouting and making threats when he felt provoked. He was less sensitive to other people's feelings. A trial of carbamazepine (Tegretol) with therapeutic blood concentrations and the maintenance of a numerical record of outbursts resulted in significantly reduced irritability and outbursts.

The DSM-5 further categorizes this condition into the following subtypes: labile (if the predominant symptom is affective lability), disinhibited, aggressive, apathetic, paranoid, combined, and unspecified (other) type. Disinhibition, poorly modulated emotional reactions, disturbances in decision-making and goal-directed behavior, social inappropriateness, hypersexuality, and lack of empathy and insight have all been linked to the occurrence of ventromedial frontal lesions. In addition, aggression and poor impulse control have been associated with lesions in the anterior temporal lobe.

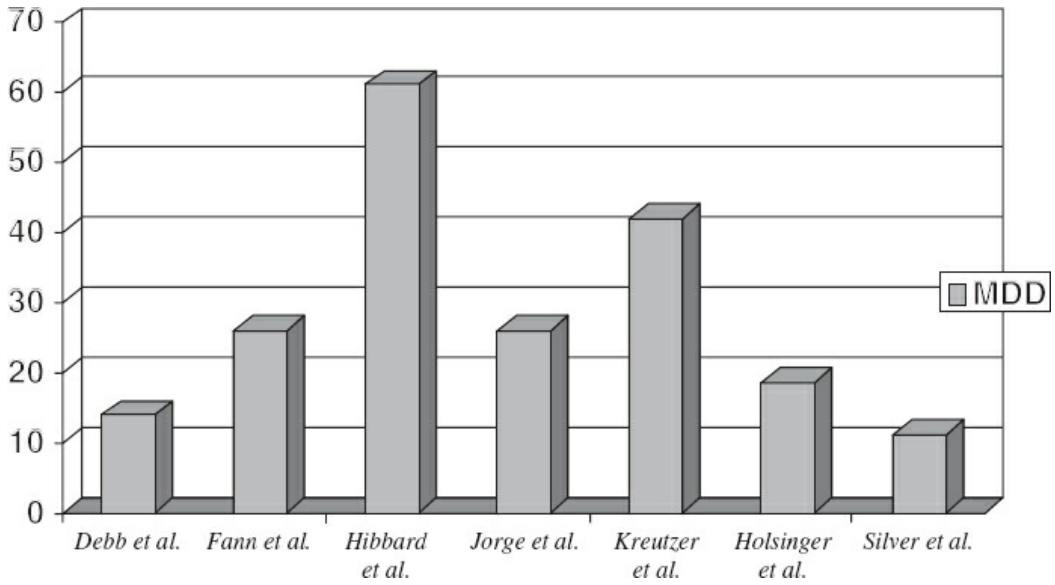

FIGURE 2.5–3. Frequency of mood disorder due to traumatic brain injury with major depressive features.

Depressive Disorders

Depressive disorders are a common psychiatric complication of TBITBI. Using the DSM-5 diagnostic criteria, depressive disorders associated with TBI are categorized as "Depressive Disorder due to Traumatic Brain Injury" with the following subtypes: (1) "With depressive features" (prominent depressed mood but full criteria for a major depressive episode are not met), (2) "With major depressive-like episode" (if the full criteria for a major depressive episode are met), or (3) "With mixed features" (manic or hypomanic symptoms present but not predominant).

The reported frequency of depressive disorders following TBI has varied from 6 to 77 percent (Fig. 2.5–3). This variability in the reported frequency of depressive disorders, particularly major depression, may reflect the lack of uniformity in the psychiatric diagnosis. Most early studies relied on rating scales or relatives' reports rather than on structured interviews and established diagnostic criteria (e.g., DSM-IV-TR). The consensus appears

to be that during the first year following TBI, depression prevalence ranges from 25 to 50 percent, with lifetime post-TBI prevalence ranging from 26 to 64 percent. While individuals in the first post-TBI year may face the greatest risk of depression, about 25 percent of those without depression in the first year will go on to develop it in the second year, and the risk may remain elevated for decades. Post-TBI depression tends to be persistent, with one study finding that ⅔ of patients experience symptoms for 1 year or longer. Another community study also suggested an association between a history of TBI and an increased lifetime prevalence of major depression: men who had suffered a TBI during World War II had an 18.5 percent lifetime prevalence of major depression, versus 13.4 percent for a comparable group without TBI. Overall, these findings suggest that Patients with TBI have recurrent depressive disorder throughout their lifetime at a significantly higher frequency than comparable patients without TBI.

Depressive disorders appear to be the most frequent psychiatric complication among patients with TBI. One observational study used a structured interview and the DSM-IV-TR criteria to identify Axis I psychopathology in 100 adults with TBI, evaluated on average 8 years after trauma. The prevalence of major depression in this series was 61 percent.

The standard DSM-5 criteria have a high sensitivity and specificity for identifying depressed patients when compared with alternative diagnostic criteria. This finding suggests that the widely used psychiatric nomenclature provides valid constructs to characterize affective disorders occurring after TBI, analyze their biological correlates, and follow their clinical course. One study found that the three symptoms of hopelessness, feeling worthless, and anhedonia best differentiated depressed from nondepressed patients. Risk factors for post-TBI depression identified by various studies include a prior history of mood or anxiety disorder, premorbid impairment of social functioning, current or past alcohol use, unemployment, and poverty.

One study that closely examined post-TBI depression consisted of 158 patients with closed head injury who came from two independent samples: the first one recruited at the University of Maryland R. Adams Cowley Shock Trauma Center, Baltimore, Maryland (n = 66), between 1989 and 1991, and the second at the University of Iowa Hospitals and Clinics (UIHC), Iowa City, Iowa (n = 61) or the specialized rehabilitation unit at Iowa Methodist Medical Center, Des Moines, Iowa (n = 31), between 1997 and 2001. In addition, 27 patients with multiple traumas but without clinical or radiological evidence of CNS involvement constituted the comparison group. Patients with penetrating head injuries, clinical or imaging findings suggesting spinal cord injury, or severe comprehension deficits that precluded a thorough neuropsychiatric evaluation were excluded from these studies. The comparison group did not significantly

differ from the TBI group in age, gender, and ethnic compositions, physical impairment, or psychosocial adjustment. The groups were also comparable in terms of their history of alcohol or other drug abuse and depressive and anxiety disorders. Moreover, the results obtained by stratifying by site were comparable to those without considering the stratification. Thus both sites were considered together in the following analyses.

According to their initial GCS and initial CT data, 98 patients (62.0 percent) had moderate-to-severe TBI, and 60 patients (38.0 percent) had a mild TBI. According to the Traumatic Coma Data Bank classification, 92 patients (58.2 percent) had diffuse CT patterns of injury, and 66 (41.8 percent) had focal patterns of injury. Most of the patients (74.6 percent) were injured in a motor vehicle crash; the remaining patients were injured as a result of falls (16.5 percent), assault (4.5 percent), and other traumas (4.4 percent).

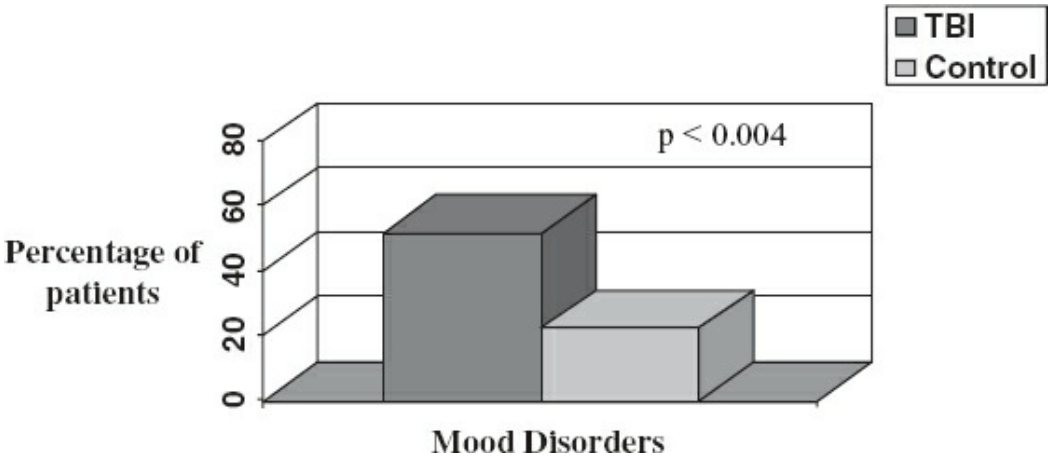

FIGURE 2.5–4. Frequency of mood disorders among patients with traumatic brain injury (TBI) compared to controls with orthopedic injuries. The frequency of mood disorders was significantly greater in TBI patients than in a comparison group of patients with orthopedic trauma.

The TBI and the comparison groups were evaluated at 3, 6, and 12 months of follow-up. The patients with TBI exhibited a significantly greater frequency of mood disorders than the comparison group of patients with orthopedic trauma (Fig. 2.5–4). In the TBI group, 54 percent of patients developed a mood disorder at some time during the first year after injury, 22 percent of patients with multiple traumatic injuries but without CNS involvement. In addition, the frequency of major depressive disorder (MDD) was also significantly greater in the group of patients with TBI when compared to the control group. Thus, mood disorders were significantly more frequent in patients who suffered TBIs than in patients with similar background characteristics who underwent similar levels of stress (e.g., motor vehicle accidents) but who did not sustain brain injury. This suggests that pathological processes associated with TBI constitute an important contributing factor to the development of these mood disorders.

Patients with post-TBI major depression did not differ from the nondepressed patients with respect to type or severity of brain injury, family history of psychiatric disorder, or the degree of physical impairment, but did have a significantly greater frequency of previous personal history of mood disorders. Patients with major depression had significantly poorer premorbid social functioning than the nondepressed group, and poor social functioning was the strongest and most consistent clinical correlate of major depression during follow-up (Fig. 2.5–5).

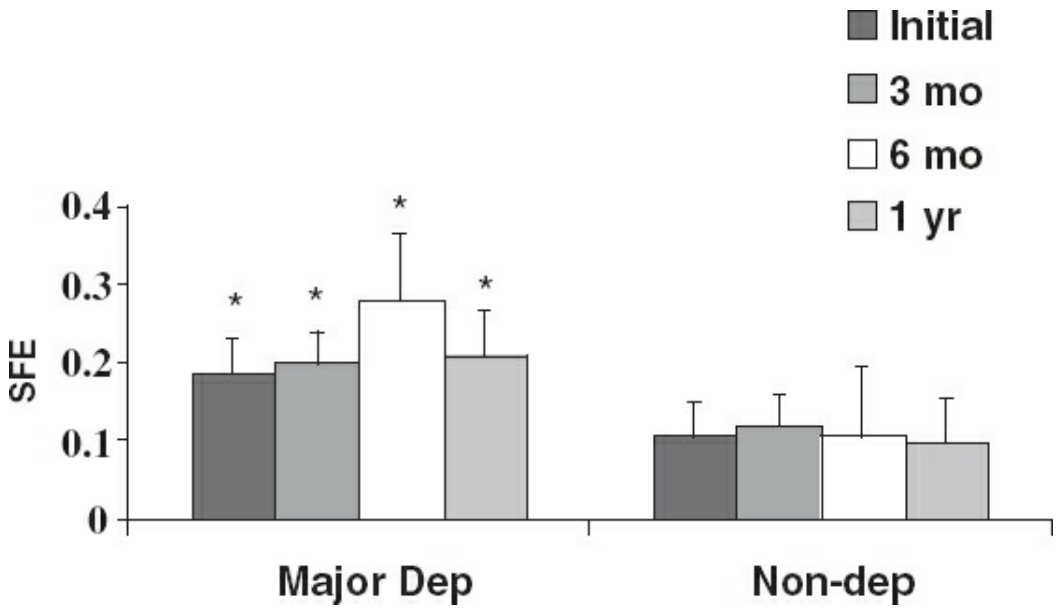

FIGURE 2.5–5. Comparison of Social Functioning Examination scores of major depressed and non-depressed patients during the first year following traumatic brain injury. Higher scores are associated with greater impairment. Poor social functioning was the strongest and most consistent clinical correlates of major depression during follow-up.

Considering the Iowa group only, MDD following TBI was significantly associated with the presence of anxiety disorders (Fig. 2.5–6). Out of 30 patients with MDD, 23 (77 percent) met diagnostic criteria for a comorbid anxiety disorder compared with 9 out of 44 patients (20 percent) who did not develop a mood disorder but met criteria for an anxiety disorder during the first year following TBI ($p < 0.001$). Of the 23 patients with comorbid anxiety and depression, 17 patients (74 percent) presented generalized anxiety features, and 6 patients (26 percent) had PTSD. However, of the 9 patients with anxiety but without comorbid depression, 6 patients (67 percent) had generalized anxiety features, and 3 patients (33 percent) had PTSD. Overall, the frequency of anxiety disorders with generalized features was 26.1 percent, and the frequency of PTSD was 14.3 percent during the first year following TBI. Major depression was also associated with the occurrence of aggressive behavior that was categorized using the Overt Aggression Scale (OAS). Of the 30 patients with major depression, 17 patients (57 percent) demonstrated significant aggressive behavior

compared with 10 of 44 patients who showed the same level of aggression without mood disorder during the first year after TBI ($p = 0.004$) (Fig. 2.5–6).

A 42-year-old engineer had a motor vehicle accident when returning from a convention. He had multiple injuries, including a diaphragmatic rupture and a left frontotemporoparietal subdural hematoma. When admitted to the hospital, the patient was hypotensive and hypoxic. With the urgent intervention of two surgical teams, he underwent repair of his diaphragm and evacuation of the subdural hematoma. The patient remained in a coma during the following 72 hours. PTA lasted for almost 3 weeks. At this point, the neurological examination disclosed a right hemiparesis, a left lateral rectus palsy, mild hypophonia, and dysarthria. Forty days after the accident, he transferred to a specialized rehabilitation unit. Once his PTA had cleared, neuropsychological tests were within normal limits. The patient conveyed a profoundly depressed mood and feelings of hopelessness. He stated that he would never be able to recover, that his career was ruined, and that it would have been better if he had died in the accident. He had no appetite and refused to participate in physical rehabilitation. He also had significant sleep problems. Treatment of depression was initiated with sertraline (Zoloft) at a dosage of 50 mg per day. After 3 weeks, the patient's mood significantly improved, and he became involved in the rehabilitation program. At 6-month follow-up, he was no longer depressed and had returned to work.

The differential diagnosis of post-TBI major depression includes adjustment disorder with depressed mood, apathetic syndromes, and PBA. Patients with adjustment disorders develop short-lived and relatively mild emotional disturbances within 3 months of a stressful life event, and symptoms do not persist longer than 6 months after the stressful event or its consequences cease.

As limited data regarding treatment of depression in the TBI population exist, clinical care generally follows principles derived from the treatment of idiopathic MDD. A comprehensive biopsychosocial approach may represent the most helpful treatment strategy for these complex patients, but no studies have yet established this.

Selection among competing antidepressants is usually guided by their side-effect profiles. Low anticholinergic activity, minimal lowering of seizure threshold, and low sedative effects are the most important factors to be considered in the choice of an antidepressant drug in this population. Based on these factors and the available evidence in the TBI population, SSRIs represent a good first-line choice for pharmacotherapy. Two small RCTs suggest sertraline improves depression in TBI; open-label studies of sertraline, citalopram, and fluoxetine have also found improvement in post-TBI depression. These studies also suggest that SSRIs are safe and well-tolerated in individuals with TBI. Data on the effectiveness of TCAs in TBI are mixed, with one small RCT finding improvement in depressive symptoms, but two open-label trials finding no effect. Some data, including one small RCT, has found that methylphenidate is effective for treating depression in TBI, and may also improve fatigue and cognitive functioning.

Bupropion, while not absolutely contraindicated in TBI, can lower the seizure threshold; using the lowest effective dose, the sustained-release formulation, and careful monitoring may help reduce the risk.

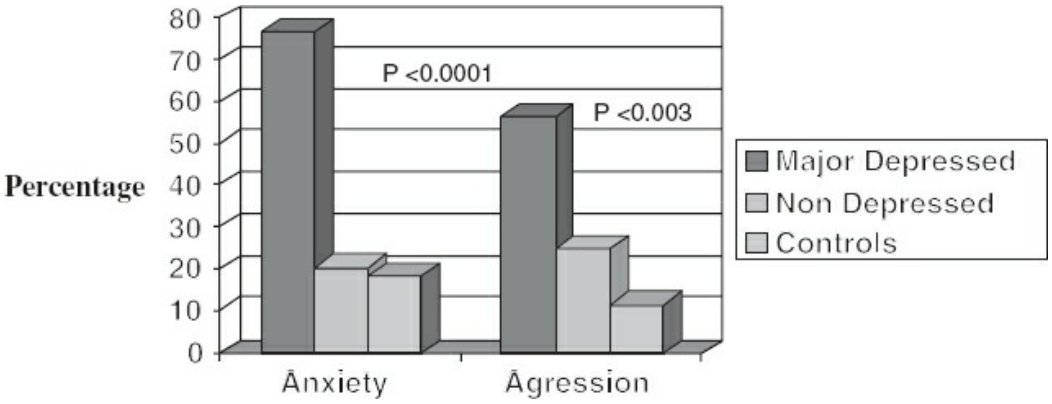

FIGURE 2.5–6. Coexistent anxiety and aggression among depressed traumatic brain injury (TBI) patients. Frequency of anxiety and aggressive disorders among depressed TBI patients, nondepressed TBI patients, and patients with orthopedic injuries. Major depressive disorder following TBI was significantly associated with the presence of anxiety disorders and aggressive behavior.

ECT is not contraindicated in patients with TBI and may be considered if other methods of treatment prove to be unsuccessful, though there are no systematic studies of its effects. ECT should be administered with the lowest possible effective energy, using brief-pulse, nondominant unilateral currents, with an interval of 2 to 5 days between treatments and four to six treatments for a complete course. Neurostimulation therapies such as transcranial direct current stimulation (tDCS) and low-frequency repetitive transcranial magnetic stimulation (rTMS) may also represent options, though data on their use in patients with TBI is limited. rTMS should be used with caution, as it may increase the risk of seizure.

Psychotherapy may help with post-TBI depression, though most evidence derives from small and often uncontrolled studies. Evidence from two small RCTs suggests that cognitive behavioral therapy (CBT) may improve mood in patients with TBI, either alone or as part of a neurorehabilitation program; another RCT found that both CBT and supportive psychotherapy equally improved depression following TBI. Uncontrolled studies indicate that mindfulness-based CBT and an internet-based self-guided CBT program also reduce symptoms of depression in TBI; these variations on standard CBT help patients who do not respond to standard CBT or who cannot drive to a therapist's office. CBT may also improve hopelessness following TBI. Another RCT found that a cognitive training intervention reduced depression symptoms, in addition to improving memory, postconcussive symptoms, PTSD symptoms, and employment. One uncontrolled study of music therapy in patients with severe TBI also found beneficial effects on mood.

TBI appears to increase the risk of subsequent suicide. A large epidemiological study of all Swedish adults treated for TBI over a 40-year period found a threefold increase in death by suicide compared to the general population (Fazel 2014). All clinicians caring for patients with a history of TBI should make sure to evaluate for suicide risk and take appropriate steps to mitigate harm.

Pseudobulbar Affect

PBA manifests as sudden and uncontrollable affective outbursts (e.g., crying or laughing), which may be congruent or incongruent with the patient's mood, and can occur spontaneously or may be triggered by minor stimuli. Patients recognize the emotional displays as excessive to the underlying mood. PBA lacks the pervasive alteration of mood, as well as the specific vegetative symptoms associated with a major depressive episode. One study examined the prevalence and clinical correlates of PBA assessed using the Pathological Laughter and Crying Scale in a group of 92 consecutive patients with TBI. Eleven percent of the patients developed PBA during the first year after TBI. PBA was associated with the presence of anxiety disorders and frontal lobe lesions involving the lateral and ventral aspects of the prefrontal region (Fig. 2.5–7). In addition, patients with PBA had significantly more frequent aggressive outbursts and poorer social functioning.

Treatment options for PBA in TBI remain understudied; however, based on their efficacy for this condition in other neurological disorders and their generally benign side-effect profile, SSRIs represent a reasonable first-line strategy. However, treatment response varies greatly among brain-injured patients, with some showing a rapid response at relatively low doses, while others require more time and higher doses. Combination therapy with dextromethorphan and quinidine can help treat PBA, but has not been specifically studied in TBI.

Apathy

Patients with TBI may present with apathetic syndromes that interfere with the rehabilitation process. Apathy consists of impairment of motivation, initiation, interest, activity, emotional responsiveness, and/or spontaneity. Apathy differs from depression in that it is not associated with sadness or low mood. Apathy, of whatever underlying etiology, frequently occurs comorbidly with depression, but the two conditions constitute distinct entities with different implications for prognosis and treatment; adherence to appropriate diagnostic criteria enables differentiation of the two conditions. While patients with apathy often do not appear distressed by their symptoms, apathy can markedly impair patients' ability to participate in rehabilitation therapies, function in the community, and maintain

interpersonal relationships.

Apathy occurs frequently with TBI, with estimated prevalence ranging from $^1/_5$ to almost ¾ of patients with TBI. Different studies have found contradictory results regarding the association between apathy and severity of TBI—some studies have found no relationship, but others have found that more severely injured patients show more apathy. A recent imaging study in patients with TBI found that patients with apathy showed damage to the left hemisphere, specifically in the frontal lobe, anterior cingulate, insula, corpus callosum, and corona radiate, consistent with these regions' involvement in emotion and goal-directed behavior.

As yet, no clear evidence-based treatments for apathy exist; however, case reports and case series suggest that dopamine agonists and stimulants may help. Some studies in the general population suggest that SSRIs can worsen apathy, but others have found that SSRIs improve apathy; no studies have yet examined SSRIs and apathy in TBI. Apathy frequently bothers the family more than the patient, providing psychoeducation on the nature of apathy and why it does not represent laziness or withdrawal of love may help. As patients with apathy may participate in and enjoy activities initiated by others, even if they do not initiate them on their own behalf, family members and friends should provide a structured environment and encourage the person with apathy to engage in activities.

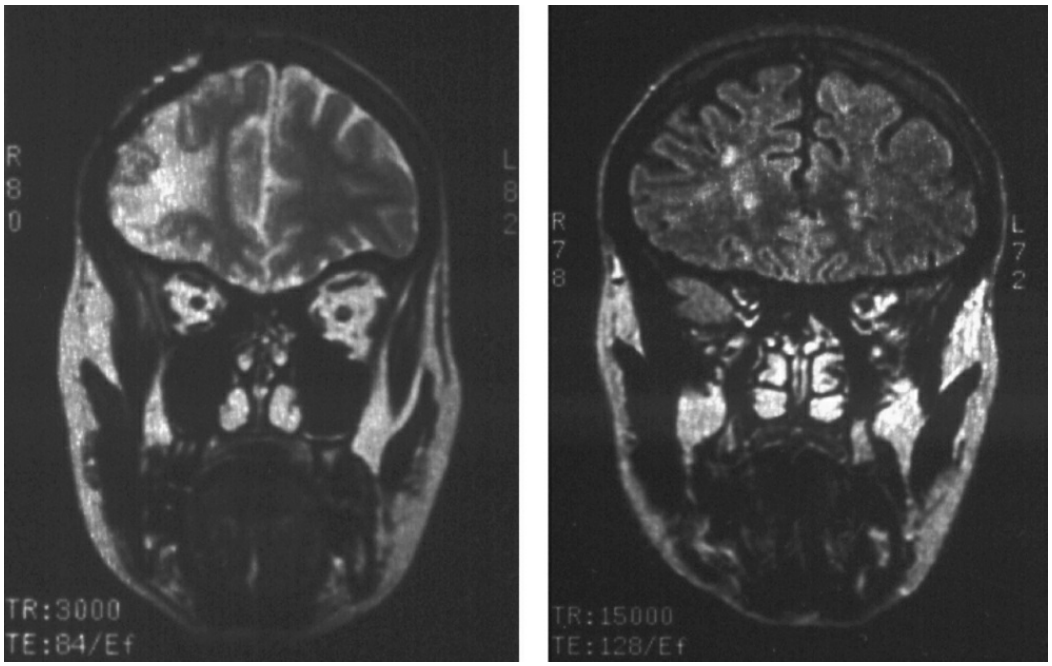

FIGURE 2.5–7. Emotional deregulation following TBI. Frontal lobe lesions involving the lateral and ventral aspects of the prefrontal region in a patient with recurrent episodes of pathological crying consistent with a diagnosis of pseudobulbar affect.

Bipolar and Related Disorders

The diagnosis of mania following TBI is based on the DSM-5 criteria for

bipolar and related disorder due to TBI, with three subtypes: (1) "With manic features" (full criteria not met for a manic or hypomanic episode), (2) "With manic- or hypomanic-like episode" (full criteria met), and (3) "With mixed features" (depressive symptoms present but not predominant) (Table 2.5–1). As in the case of depressive disorders, the mania must appear to stem directly from the physiological consequences of the brain injury. The diagnosis only requires that the predominant mood be elevated, euphoric, or irritable, and accompanied by abnormally elevated energy and/or activity. This obviously does not require the same specificity of syndrome as a manic episode or a major depressive episode.

A 1987 study reported on 20 patients who developed manic syndromes after closed head trauma. There was a significant association between mania and the presence of post-traumatic seizures, predominantly of the partial complex type (temporal lobe epilepsy). There was no association, however, with family history of bipolar disorder among first-degree relatives.

One study found that patients with TBI showed a higher rate of secondary mania during the year after traumatic injury than patients with orthopedic injuries. Although manic episodes lasted only approximately 2 months, the presence of an elevated, expansive, or irritable mood had a mean duration of 5.7 months. Manic syndromes were significantly associated with hyperactivity, disinhibition, aggressive behavior, and hypersexuality. Secondary mania was not related to the type or severity of brain injury, degree of physical or intellectual impairment, level of social functioning, or the presence of family or personal history of psychiatric disorder (Fig. 2.5–8). The development of abnormal electrical activation patterns in limbic networks, functional changes in aminergic inhibitory systems, and the presence of aberrant regeneration patterns may all play a role in the genesis of post-TBI mania.

A 31-year-old married man with two children was struck in the head when a welding tank exploded, sending large and small pieces of metal in all directions. He experienced intracerebral bleeding in the right basal ganglia and white matter surrounding the anterior horn of the right lateral ventricle. Within a few days, he became manic, staying up all night, selling things he owned for ridiculously low prices, and buying a new car he could not afford. He believed that he would become a famous songwriter and exhibited hypersexual behavior. He was admitted to a hospital, where he was treated successfully with lithium and low-dose haloperidol (Haldol). He remained euthymic for approximately 6 months, and he discontinued his medications. Although he continued to experience significant behavioral changes, manic symptoms did not recur during a 2-year follow-up period.

The differential diagnosis of mania following TBI should include substance/medication-induced bipolar and related disorder, psychotic disorder due to TBI, and personality change due to TBI. Substance-induced mood disorder occurs following intoxication or withdrawal from drugs and

is another important differential diagnosis among patients with TBI, who frequently have a prior history of substance abuse.

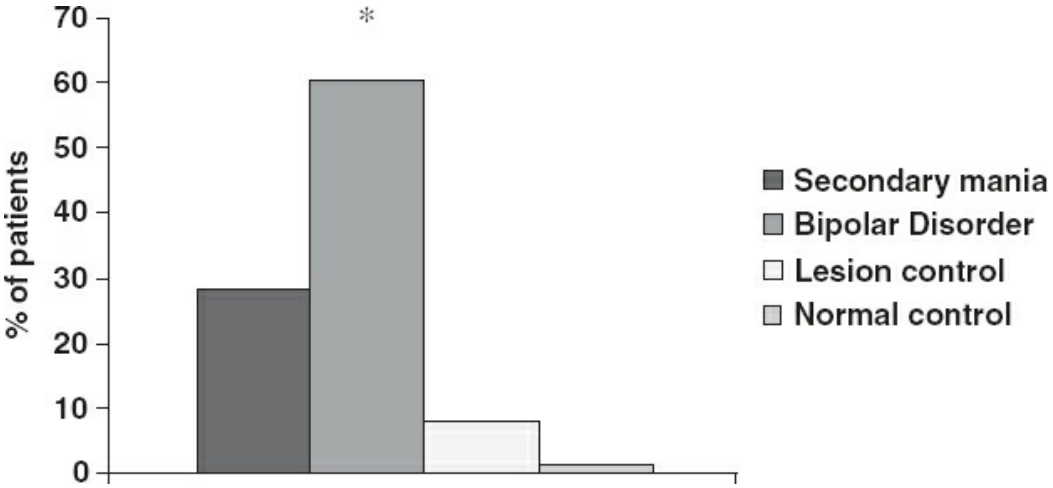

FIGURE 2.5–8. Family history of mood disorder among patients with manic syndromes due to traumatic brain injury (TBI). The frequency of family history of mood disorders is significantly higher among patients with bipolar disorder compared to patients with secondary mania, TBI patients without mood disorders, and healthy controls. However, family history of mood disorders was greater among patients with secondary mania compared to TBI patients without mood disorders and healthy controls.

There have been no systematic studies for the treatment of secondary mania. Lithium, carbamazepine, and valproate therapies have effectively treated in individual cases. Some researchers have theorized that lithium and valproate have neuroprotective effects, which would certainly constitute an important therapeutic advantage among brain-injured populations. However, data from the only controlled trial of valproate in TBI failed to identify a beneficial effect on cognitive and functional outcomes. The roles of the anticonvulsant lamotrigine for mood stabilization in TBI have not been tested.

Relationship between Mood Disorders and Functioning in TBI. It is well known that age and severity of brain injury are major determinants of recovery from TBI. However, other factors also influence long-term disability and, particularly, patients' return to a productive and meaningful life. Mood disorders have consistently shown an independent deleterious effect on the recovery of patients with TBI. For instance, one study of 100 patients with moderate-to-severe TBI at 6 months after trauma showed a significant association between depression and recovery status as measured by the Glasgow Outcome Scale (GOS). Another study of 170 patients with mild TBI found that those who developed major depression had objective evidence of poorer behavioral outcome. Taken together, these studies emphasize the need of recognizing and treating mood and anxiety disorders during the rehabilitation process.

Mood disorders and executive deficits interact to produce a detrimental effect on psychosocial outcomes. One study found that patients with TBI with major depression showed poorer psychosocial outcome; furthermore, half of the patients with major depression and initial ADL impairment had poor outcomes, while none of the nondepressed patients had a poor ADL outcome. Although patients with major depression did not differ from nondepressed TBI patients with regard to global measures of cognitive function such as the MMSE, they showed more impairment on tests of executive function.

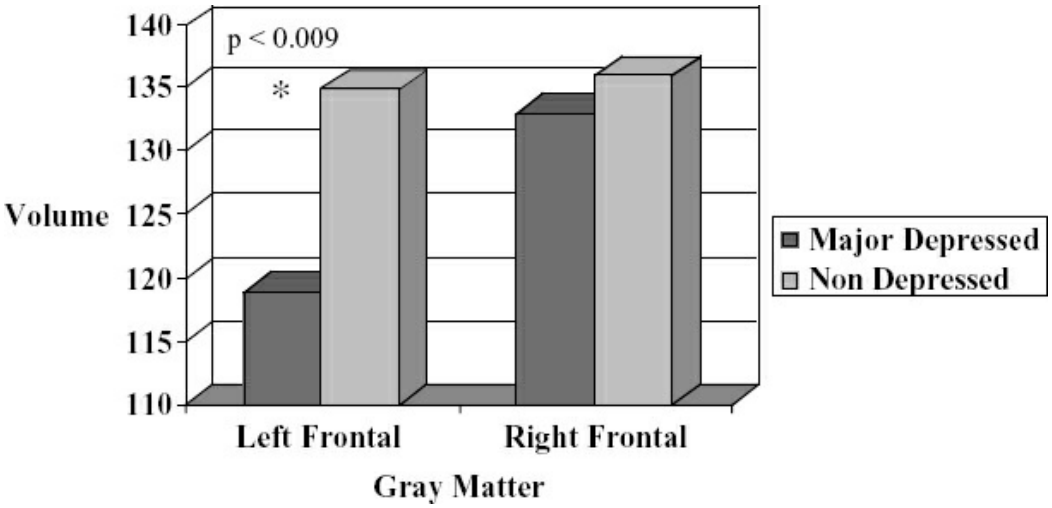

FIGURE 2.5–9. Frontal gray matter volumes in traumatic brain injury (TBI) patients with major depression. Magnetic resonance imaging volumetric studies showing that, after controlling for age and severity of trauma, left frontal gray matter volumes are significantly decreased in the group of depressed TBI patients compared with the nondepressed group.

Structural Brain Damage as an Etiological Factor for Mood Disorders Following TBI. Analysis of the Oxford Collection of Head Injury Records suggested that depressive symptoms were more common among patients with right hemisphere lesions. Depressive symptoms were also more frequent among patients with frontal and parietal lesions than among patients with other lesion locations. Furthermore, a study of Vietnam War veterans conducted several years following head injury reported that penetrating injuries involving the right orbitofrontal cortex showed the strongest association with depressive and anxiety symptoms. However, other studies have found a relationship between left hemisphere lesions and depressive symptoms as well. For instance, one study found an association between post-TBI major depression and the presence of left dorsolateral frontal and/or left basal ganglia lesions and, to a lesser extent, with right hemisphere and parieto-occipital lesions. In one series, post-TBI major depression was associated with reduced gray matter volume in the lateral aspects of the left prefrontal cortex (Fig. 2.5–9), a finding consistent with other studies of secondary depressive disorders that found decreased

metabolic rates in inferior frontal regions in patients with Parkinson disease, Huntington disease, and caudate stroke.

It is unclear if the reduced prefrontal volumes observed in patients with MDD result from the pathophysiological mechanisms initiated by TBI, or if they constitute a pre-existent trait associated with an increased risk to develop mood disorders. For instance, patients with substance abuse, antisocial personality, and chronic mood or anxiety disorders might show prefrontal volumetric changes unrelated to the traumatic insult. However, when controlling for these factors, evidence to support the idea that asymmetrical differences in frontal lobe volume predated the brain injury was not found, suggesting that the decreased left frontal lobe volume results from resolving lesions approximately 3 months postinjury. In this sense, major depression could result from the deactivation of more lateral and dorsal frontal cortex and increased activation in ventral limbic and paralimbic structures including the prelimbic cortex and the amygdala. The cognitive abnormalities observed in major depressed patients with TBI are consistent with left lateral prefrontal dysfunction. Interestingly, high levels of amygdala activation may be associated with an increased prevalence of anxiety symptoms and negative affect, a pattern of symptoms that closely resembles what was observed in another group of TBI patients. Moreover, faulty prefrontal modulation of medial limbic structures could explain the impulsive and aggressive behavior observed in these patients.

Patients who develop mood disorders after TBI have significantly lower hippocampal volumes than patients without mood disturbance; in addition, reduced hippocampal volumes have been associated with poor vocational outcomes at 1-year follow-up. These findings suggest a double-hit mechanism in which neural and glial elements already affected by trauma suffer further compromise by the functional changes associated with mood disorders (e.g., the neurotoxic effects of increased levels of cortisol or excitotoxic damage resulting from overactivation of glutaminergic pathways). Theoretically, early administration of antidepressants might prevent the occurrence of progressive structural and functional and, consequently, the psychiatric and functional morbidities associated with their disruption; however, the limited data available on prophylactic antidepressant treatment does not support this hypothesis (Fig. 2.5–10).

Trauma- and Stressor-Related Disorders

PTSD, the prototypical trauma-related disorder, is characterized by recurrent intrusive recollections, distressing dreams, and flashbacks of the traumatic event. Patients with PTSD show features of increased arousal such as difficulty falling or staying asleep, exaggerated startle response, and irritability. Patients lose interest in social activities and avoid thoughts, feelings, or circumstances associated with the event.

Individuals who sustain a TBI are more likely to develop PTSD than

those who suffer other types of traumatic injuries. Veterans who experienced a TBI show a threefold greater rate of PTSD than those without a TBI; civilians with TBI have almost twice the likelihood of developing PTSD as those with other types of traumatic injuries.

Estimates of how many individuals with TBI have comorbid PTSD vary widely, from 12 to 30 percent in civilians and from 12 to 89 percent in veterans. In addition to variance introduced by different methods of assessment, part of the uncertainty lies in the overlap between PTSD and TBI symptoms, such as impaired concentration/attention and executive function, and poor sleep.

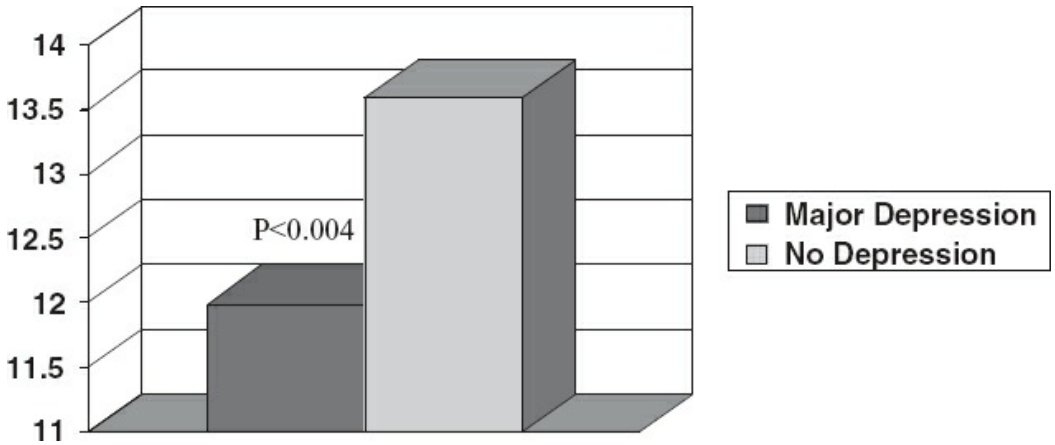

FIGURE 2.5–10. Hippocampal volumes in patients with moderate-to-severe traumatic brain injury (TBI) with and without mood disorders. Hippocampal volumetric findings among patients with moderate-to-severe traumatic brain injury. Volumes of the hippocampus have been adjusted for age and total intracranial volume. There was a significant interaction between mood disorder diagnosis and severity of TBI, by which patients with moderate-to-severe TBI who developed mood disorders had significantly smaller hippocampal volumes than patients with equivalent severe TBI who did not develop mood disturbance.

Although initially some researchers speculated that TBI was incompatible with the development of PTSD, theorizing that amnesia for the traumatic event could potentially protect against PTSD, more recent studies have not born this out. Even individuals with moderate-to-severe TBI—with its greater impairment of consciousness—can still develop PTSD. While some studies have found that patients with TBI with more PTA are less likely to develop PTSD, others have found that greater degrees of PTA actually increase the risk of PTSD. Patients with TBI can encode and retrieve trauma memories at an implicit level and these memories can influence ongoing emotions and behaviors. Limbic structures mediate fear conditioning to traumatic experiences independent of higher cortical processes, and this mechanism might elicit PTSD symptoms among patients with TBI with a loss of consciousness or severe PTA. One study found that individuals with severe TBI who had higher heart rates (a measure of sympathetic nervous system activity) at 1 week following TBI

were more likely to have PTSD 6 months after the injury, suggesting that fear conditioning can occur independently of the level of awareness and contribute to PTSD onset. Lack of declarative memory of the traumatic event could also interfere with attempts to place the trauma in the context of an overall life story and mobilize coping resources to address it.

Why do so many patients with TBI experience comorbid PTSD? Damage to specific brain structures a TBI may impair the brain's ability to modulate fear responses, thus predisposing to PTSD: for instance, studies have found that patients with TBI who experience damage to the left middle temporal gyrus or disruption of white matter show higher rates of TBI than those with other types of lesions. Cognitive deficits from TBI could impair the ability to effectively use coping skills to deal with PTSD; conversely, ongoing psychological stressors from PTSD could impair cognitive recovery from the TBI. Alternately, the types of incidents that result in TBI could produce more feelings of fear and helplessness than those responsible for other types of injuries.

The combination of TBI and PTSD appears to worsen outcomes compared to TBI alone or PTSD alone. Patients with PTSD at 6 months following severe TBI had significantly more impaired community reintegration than TBI patients without PTSD. Individuals with TBI complicated by PTSD self-assess their cognitive and health status as worse than those with TBI alone, and some data suggests that the combination of PTSD and TBI impairs objective cognitive performance more than TBI alone. Patients with both disorders report more symptoms of PTSD than those with PTSD alone, and experience greater PTSD symptom severity.

Some investigators have described a triad of TBI, PTSD, and chronic pain in trauma-exposed veterans of OIF/OEF/OND. An analysis of diagnostic records of all VA patients who served in OIF/OEF/OND found that almost two-thirds of veterans with TBI experienced this polytrauma triad; less than 10 percent of veterans with TBI were free of a comorbid pain disorder or PTSD (Cifu 2014). The long-term outcomes and optimal treatment of patients with this triad remain unclear.

While clear evidence-based strategies for treating PTSD in the context of TBI remain in need of development, clinicians can use the principles of PTSD treatment in the general population in the interim. Given their safety, tolerability, and good track record in individuals without TBI, SSRIs represent a reasonable first-line treatment strategy in PTSD associated with TBI. An open-label study of prazosin—effective for PTSD-related nightmares in the general population—in patients with TBI found improvements in cognitive performance, sleep, and headaches, suggesting a role for prazosin as well. Given the possibility of adverse cognitive side effects, anticholinergic medications and benzodiazepines should be avoided. CBT, cognitive processing therapy, and prolonged exposure therapy effectively treat PTSD in the general population; preliminary

studies indicate they may benefit PTSD associated with TBI. One study examined the efficacy of a cognitive behavioral strategy in 24 patients with acute stress disorder following mild head injuries; patients who received cognitive behavioral treatment had less re-experiencing and avoidance symptoms at 6-month evaluation than patients receiving supportive counseling.

A 33-year-old certified nurse was assaulted while walking to her job in a small hospital. She received several blows to her head and was admitted unconscious to the emergency room. A CT scan showed a depressed skull fracture and right temporal and parietal contusions. She regained consciousness during the first 12 hours of admission and was confused and disoriented for the following 3 days. After this, she had an uneventful course and a good recovery. One month after trauma, she received a complete neuropsychiatric evaluation. She complained of intrusive recollections of the moments previous to the assault, which were associated with great psychological distress, sweating, and palpitations. She was also disturbed by nightmares in which she suddenly found herself in an intensive care unit connected to a ventilator, incapable of talking or asking for help. She avoided the hospital's surroundings and was unable to walk alone. She also complained of decreased concentration and memory impairments. The results of her neuropsychological tests, however, fell within the normal range.

Anxiety Disorders

Few studies have explored the prevalence and clinical correlates of anxiety disorders in the TBI population. The available data, however, suggest that anxiety disorders frequently complicate TBI. Anxiety disorders following TBI may manifest themselves in a variety of ways. The DSM-5 classifies symptom patterns as anxiety disorders due to another medical condition (i.e., TBI) if anxiety or panic attacks predominate and occur from the direct physiological effects of the medical condition (Table 2.5–1).

The estimated prevalence of generalized anxiety disorder (GAD) after TBI varies from 2 to 32 percent, and may represent the most common post-TBI anxiety disorder. Rates of panic disorder following TBI range from 6 to 14 percent; one study found that TBI conferred double the risk of developing panic disorder compared to other traumatic injuries, and also doubled the risk of developing agoraphobia, with 7 percent of patients with TBI experiencing this condition. The limited data on post-TBI-specific phobias show rates of 1 to 10 percent, not necessarily elevated compared to the general population rate of 8 to 12 percent. One study found that among patients who experienced TBI due to a motor vehicle accident, the specific phobias that subsequently developed involved transportation; another group found that a substantial proportion of individuals with TBI experience a phobic-like fear of falling, which at times may interfere with activities. One study found that 5 percent of patients with TBI developed social anxiety disorder, twice the rate of those who experienced other types of injuries. Overall, the presence of anxiety disorder has not been

associated with any background characteristic including severity of head injury or severity of social, physical, or cognitive impairment resulting from TBI.

In the TBI population, anxiety and depression overlap substantially; pathological worrying, anxious foreboding, and autonomic symptoms are commonly associated with anxious depressions. One prospective study of 66 patients with TBI found that 42 percent of the individuals with an acute major depressive episode also had comorbid GAD; these patients with anxious depression had a more prolonged course of depression than those with depression alone (on average, 7.5 months vs. 1.5 months). Another study of patients with TBI found that 53 percent of those with a major depressive episode developed an anxiety disorder as well, versus only 20 percent of nondepressed patients. Another study found that patients with TBI with a major depressive episode had a 24-times higher rate of panic disorder than those without depression.

Previous experimental and clinical studies on the neurobiology of anxiety disorders have emphasized the etiological role of disrupted limbic and paralimbic circuits involving medial and orbital prefrontal cortices, the amygdala, and the hippocampus. At the present time, the effect of brain injury on the integrity and activation patterns of these neural circuits has not been extensively studied. A study of Vietnam War veterans with combat-associated TBI found an association between higher levels of anxiety and predominantly left-sided lesions, specifically in the temporal lobe, limbic areas, occipital lobe, and cerebellum. By contrast, one study of patients with TBI found that individuals with anxious depression, as opposed to depression without anxiety, were more likely to have a right hemisphere lesion. A study of 95 children and adolescents with TBI found that orbitofrontal lesions were actually associated with less risk of anxiety disorders; in this population, lesions in other brain areas did not correlate with anxiety symptoms.

The evidence base for treatment of anxiety disorders in TBI remains very preliminary; in general, the treatments used for idiopathic anxiety disorders likely represent the most practical options at this time. As with trauma-related disorders in TBI, SSRIs are likely the best first-line pharmacotherapy, and medications that could worsen cognition in this vulnerable population, such as benzodiazepines and drugs with anticholinergic effects, should be avoided. A few studies specifically examining psychotherapy for anxiety in TBI suggest CBT is effective.

Substance-Related and Addictive Disorders

Substance-related and addictive disorders are one of the major risk factors for the occurrence of TBI and also a major determinant of the clinical and psychosocial outcome of patients with TBI.

Substance misuse and TBI appear to share a bidirectional relationship:

substance misuse increases the risk of TBI and TBI increases the risk of substance misuse. Large studies of veterans have found that combat-related TBI increases the risk of binge drinking and the likelihood of being diagnosed with an addiction-related disorder; as military regulations prohibit alcohol and other substance use in war zones, this relationship is unlikely to reflect a greater propensity for intoxicated individuals to sustain a TBI. Mary R. Hibbard and colleagues found that 40 percent of patients with TBI met the DSM-IV-TR criteria for alcohol abuse or dependence before injury, and 28 percent of the sample received such diagnoses at the time of the study. A prospective study of 197 patients with TBI showed that 42 percent had blood alcohol levels consistent with legal intoxication at the time of admission to the emergency department and that blood alcohol levels suggested a history of problem drinking among these patients. The authors also observed that the frequency of alcohol abuse and associated disorders diminished immediately after TBI but increased again at 1-year follow-up, although not to pre-injury levels. Although alcohol use significantly decreased during the first year after injury, about 25 percent of the patients reported alcohol-related problems. As expected, pre-injury problematic alcohol use predicted alcohol-related problems after TBI. However, one large Australian community prospective cohort study found that individuals who misused alcohol or marijuana at baseline did not show an increased risk of TBI, and individuals with TBI did not show an increased risk of subsequently developing a substance use disorder.

In one study, 60 percent of patients with alcohol use disorders prior to TBI had resumed problematic alcohol use within 1 year of their injury. Patients who relapsed had significantly less education, showed significantly more severe injury as demonstrated by lower GCS scores and a greater frequency of moderate and severe injuries, had significantly higher frequencies of focal lesions on their initial CT or MRI scans, and showed impaired cognitive performance on memory and executive tasks.

Conceivably, the combined effects of traumatic injury and the toxicity of chronic substance abuse (particularly alcohol) might produce greater disruption of the neural circuits involved in mood regulation, motivation, and reward processing. In turn, these changes could result in more severe psychopathology, uncontrolled addictive behavior, and poor psychosocial outcome. One study found that inmates with head injuries had higher levels of alcohol and marijuana use, significantly higher frequency of depression, anxiety, and suicidal thinking, as well as more difficulties in concentrating and controlling violent behavior. In addition, another study reported that inmates who sustained a TBI during the year prior to neuropsychiatric assessment had significantly worse anger and aggression scores and a higher prevalence of psychiatric disorders than the group without TBI in the prior year.

Substance use disorders are associated with worse post-TBI outcomes.

Individuals with TBI who abuse substances show worse cognitive performance, higher rates of incident depression, decreased employment, and poorer rehabilitation participation than patients with TBI alone. One large study found that, among individuals with TBI, those who had consumed alcohol around the time of their injury were 20 percent more likely to die of their injuries, and were less likely to be able to return home at hospital discharge.

The optimal treatment strategy for comorbid substance use disorders and TBI remains to be determined. One study of Systematic Motivational Counseling (SMC) to treat substance abuse comorbidity among patients with TBI found that the SMC group showed significant improvements in motivational structure and a significant reduction in negative affect and the use of addictive substances. However, two RCTs of a brief motivational interviewing and educational intervention targeting alcohol use disorders among patients with TBI did not find effects on alcohol consumption, but one of the studies did indicate that the intervention increased awareness of the negative consequences of alcohol use. Physical disabilities, such as those which may accompany TBI, can pose barriers to substance abuse treatment, as many treatment programs do not accommodate them. Expanding substance abuse treatment access for these patients remains critical.

Psychotic Disorders

Following DSM-5 criteria, psychotic disorder due to TBI consists of prominent hallucination or delusions occurring as a direct physiological consequence of the TBI. Symptoms occurring exclusively in the context of delirium are excluded.

The reported incidence of psychosis following TBI varies widely. Retrospective reviews of World War II records found a rate of 7 percent, while one prospective study in an inpatient rehab setting found that 64 percent of individuals with moderate or severe TBI experienced hallucinations and/or delusions at some point in their postinjury course. Psychotic symptoms were associated with poorer sleep and worse cognitive impairment, and tended to occur earlier in the post-TBI recovery period.

Much of the literature on psychosis following TBI takes the form of case studies. Nineteen published reports describe patients with delusions only, with delusional content including Capgras syndrome, reduplicative paramnesia, delusional jealousy, Cotard syndrome, somatic delusions, and Fregoli syndrome. Seventy-one published reports describe patients with psychotic symptoms thought to resemble those of schizophrenia—i.e., hallucinations and delusions; almost all of these patients experienced auditory hallucinations, one in five also experienced visual hallucinations, 85 percent experienced delusions (most commonly persecutory), and 43 percent experienced negative symptoms. On average, these 90 patients

developed psychotic symptoms 3.3 years after the TBI.

In line with the current hypothesis on the physiopathology of psychotic disorders, post-traumatic psychosis has been associated with the presence of cognitive disturbance and imaging and EEG abnormalities in the frontal and temporal lobes. In addition, post-traumatic psychosis has been associated with seizure disorder in approximately one-third of cases.

One study examined neuropsychological functioning in individuals with psychotic disorder following TBI and compared it to that of patients with TBI without psychosis, patients with schizophrenia, and normal controls. Patients with TBI and psychosis and patients with schizophrenia, both showed impairments in measures of verbal learning and executive functioning compared with healthy controls. It is important to determine if psychotic symptoms occur in the context of more global and severe cognitive impairment, which would suggest a more appropriate diagnosis of post-traumatic dementia.

A related clinical problem is whether TBI constitutes a risk factor for later schizophrenia. A recent meta-analysis of nine studies with a total of 180,859 persons found an elevated risk of subsequent development of schizophrenia in individuals who had experienced a TBI (RR 1.65, 95 percent CI 1.17–2.32). Individuals with a family history of schizophrenia appeared to face a particularly heightened risk after TBI; conversely, a family history of schizophrenia also represented a risk factor for experiencing a TBI, suggesting the possibility of a complex interaction between the two disorders.

To date, no clinical trials have examined the pharmacotherapy of psychosis due to TBI. By extrapolation from the treatment of idiopathic psychosis, antipsychotic medications will likely represent the mainstay of treatment; however, given the possible association between post-traumatic psychosis and epilepsy, AEDs may also represent an option. Psychosocial treatments useful in patients with schizophrenia, such as psychoeducation of patients and family, vocational rehabilitation, psychotherapy, and assertive community treatment, may also help patients with psychosis due to TBI, though no studies have yet explored this.

COURSE AND PROGNOSIS

Assessing the course and prognosis of patients with TBI involves the longitudinal analysis of neurological, neuropsychological, psychiatric, and psychosocial variables. Some of these variables are readily operationally defined, while others are more elusive and difficult to quantify. The GOS has been widely used as a measure of the long-term outcome of patients with TBI. It consists of five levels of outcome: (1) death, (2) persistent vegetative state, (3) severe disability (conscious but dependent in activities of daily living), (4) moderate disability (disabled but living independently), and (5) good recovery (mild neuropsychiatric effects but able to resume an

otherwise normal life). Although crude, the scale has both validity and high reproducibility.

TBI, especially moderate and severe TBI, is associated with significant morbidity, mortality, and disability. One cohort study of 508 patients with moderate or severe TBI, 54 percent of patients with severe TBI, and 78 percent with moderate TBI survived to 1-year follow-up; among those who died, 85 percent of the deaths occurred in the emergency department or during hospitalization, and three-fourth were attributed to the head injury. Among individuals who survived to 1 year, 46 percent of patients with severe TBI and 55 percent of patients with moderate TBI showed good recovery. Another matched cohort study of 1,290 individuals hospitalized with TBI found no excess mortality among the 798 patients with mild TBI, but even these patients with mild TBI had increased hospitalizations and health services use during the 10 years following injury. A large epidemiological study of 18,988 patients hospitalized for TBI who survived to discharge found that these individuals showed 2.5 times higher risk of death at all times during follow-up, though the greatest excess mortality occurred in the month after discharge.

The long-term outcome of patients with TBI primarily relates to the severity of brain injury, type and location of intracranial lesion, age, and efficacy of acute medical and surgical treatment. Socioeconomic status, educational level, previous psychiatric disorders (e.g., history of alcohol and/or drug abuse, personality disorders), premorbid social functioning levels, the quality and extent of rehabilitation services, and the availability of social and vocational support also play a significant role in TBI outcome.

Cognitive disturbances are among the most important factors in long-term disability following severe TBI. A previous study reported on a group of 50 mild and moderate head injury patients and correlated different neuropsychological measures obtained in-hospital and at 1 and 3 months follow-up with serial MRI findings. The TBI group had significantly greater neuropsychological impairment than a control group. These cognitive deficits showed an impressive recovery during follow-up lacked a consistent relationship with specific lesion locations. MRI abnormalities were present at baseline in 80 percent of these patients and gradually resolved by 3 months.

Other studies have reported that by 6-month follow-up, 18 percent of severe head injury patients returned to gainful employment and 62 percent of former students returned to school. Among those not back to work or school at 6 months, 31 percent of the former workers and 66 percent of the former students returned by 12 months. The three most significant predictors for returning to work or school were age, intact verbal abilities, and speed of information processing. A series of moderate head injury cases revealed that 69 percent of patients who had been gainfully employed before the injury were unemployed at 3 months follow-up. Another study

found that among patients with moderate-to-severe TBI, 44 percent remained employed during the first 3 years following injury, with other patients employed for some of that time.

Behavioral changes associated with TBI may disrupt interpersonal relationships and pose a great burden to family members. During the first year after the injury, more than two-thirds of relatives experienced moderate-to-severe degrees of burden as a consequence of the behavioral changes in their family member. Thus, psychosocial adjustment and community re-entry have become the targets of rehabilitation efforts following TBI. Patient motivation and a history of alcohol/drug abuse as well as awareness of cognitive and physical impairments also exert a significant effect on rehabilitation outcome.

REFERENCES

Arciniegas DB, Harris SN, Brousseau KM. Psychosis following traumatic brain injury. *Int Rev Psychiatry*. 2003;15:328–340.

Bahraini NH, Breshears RE, Hernández TD, Schneider AL, Forster JE, Brenner LA. Traumatic brain injury and posttraumatic stress disorder. *Psychiatr Clin North Am*. 2014;37:55–75

Bigler ED. Neuropsychology and clinical neuroscience of persistent post-concussive syndrome. *J Int Neuropsychol Soc*. 2008;14:1–22.

Büki A, Povlishock JT. All roads lead to disconnection?—Traumatic axonal injury revisited. *Acta Neurochir (Wien)*. 2006;148:181–193; discussion 193.

Carroll LJ, Cassidy JD, Cancelliere C, Côté P, Hincapié CA, Kristman VL, et al. Systematic review of the prognosis after mild traumatic brain injury in adults: cognitive, psychiatric, and mortality outcomes: results of the International Collaboration on Mild Traumatic Brain Injury Prognosis. *Arch Phys Med Rehabil*. 2014;95:S152–S173.

Centers for Disease Control and Prevention (CDC). CDC grand rounds: reducing severe traumatic brain injury in the United States. *MMWR Morb Mortal Wkly Rep*. 2013;62(27):549–552.

Cicerone KD, Langenbahn DM, Braden C, Malec JF, Kalmar K, Fraas M, et al. Evidence-based cognitive rehabilitation: updated review of the literature from 2003 through 2008. *Arch Phys Med Rehabil*. 2011;92:519–530.

Cifu DX, Taylor BC, Carne WF, Bidelspach D, Sayer NA, Scholten J, et al. Traumatic brain injury, posttraumatic stress disorder, and pain diagnoses in OIF/OEF/OND Veterans. *J Rehabil Res Dev*. 2013;50(9):1169–1176.

Crowe SF. *The Behavioral and Emotional Complications of Traumatic Injury*. Philadelphia, PA: Taylor Francis; 2008.

Dikmen SS, Machamer JE, Powell JM, Temkin NR. Outcome 3 to 5 years after moderate to severe traumatic brain injury. *Arch Phys Med Rehabil*. 2003;84:1449–1457.

Edlow BL, Wu O. Advanced neuroimaging in traumatic brain injury. *Semin Neurol*. 2012;32:374–400.

Fann JR, Burington B, Leonetti A, Jaffe K, Katon WJ. Psychiatric illness following traumatic brain injury in an adult health maintenance organization population. *Arch Gen Psychiatry*. 2004;61:53–61.

Fann JR, Hart T, Schomer KG. Treatment for depression after traumatic brain injury: a systematic review. *J Neurotrauma*. 2009;26:2383–2402.

Fujii DE, Ahmed I. Psychotic disorder caused by traumatic brain injury. *Psychiatr Clin North Am*. 2014;37(1):113–124.

Gardner A, Iverson GL, McCrory P. Chronic traumatic encephalopathy in sport: a systematic review. *Br J Sports Med*. 2014;48(2):84–90.

Johnston MV, Shawaryn MA, Malec J, Kreutzer J, Hammond FM. The structure of functional and community outcomes following traumatic brain injury. *Brain Inj*. 2006;20:391–407.

Jorge RE, Acion L, Starkstein SE, Magnotta V. Hippocampal volume and mood disorders after traumatic brain injury. *Biol Psychiatry*. 2007;62:332–338.

Jorge RE, Robinson RG, Moser D, Tateno A, Crespo-Facorro B. Major depression following traumatic brain injury. *Arch Gen Psychiatry*. 2004;61:42–50.

Kim E, Lauterbach EC, Reeve A, Arciniegas DB, Coburn KL. Neuropsychiatric complications of traumatic brain injury: a critical review of the literature (a report by the ANPA Committee on Research). *J Neuropsychiatry Clin Neurosci*. 2007;19:106–127.

Lee YK, Hou SW, Lee CC, Hsu CY, Huang YS, Su YC. Increased risk of dementia in patients with mild traumatic brain injury: a nationwide cohort study. *PLoS One*. 2013;8:e62422.

Mallya S, Sutherland J, Pongracic S, Mainland B, Ornstein TJ. The manifestation of anxiety disorders after traumatic brain injury: a review. *J Neurotrauma*. 2014 Sep 16. [Epub ahead of print]

Molloy C, Conroy RM, Cotter DR, Cannon M. Is traumatic brain injury a risk factor for schizophrenia? A meta-analysis of case-controlled population-based studies. *Schizophr Bull*. 2011;37:1104–1110.

Seel RT, Kreutzer JS, Rosenthal M, Hammond FM, Corrigan JD. Depression after traumatic brain injury: A National Institute on Disability and Rehabilitation Research Model Systems multicenter investigation. *Arch Phys Med Rehabil*. 2003;84:177.

Singh R, Venkateshwara G, Nair KP, Khan M, Saad R. Agitation after traumatic brain injury and predictors of outcome. *Brain Inj*. 2014;28:336–340.

Starkstein SE, Jorge R. Dementia after traumatic brain injury. *Int Psychogeriatr*. 2005;17:S93–S107.

West SL. Substance use among persons with traumatic brain injury: a review. *NeuroRehabilitation*. 2011;29:1–8.

West TA, Sharp S. Neuroendocrine dysfunction following mild TBI: when to screen for it. *J Fam Pract*. 2014;63:11–16.

Wilde EA, Bigler ED, Pedroza C, Ryser DK. Post-traumatic amnesia predicts long-term cerebral atrophy in traumatic brain injury. *Brain Inj*. 2006;20:695.

Writer BW, Schillerstrom JE. Psychopharmacological treatment for cognitive impairment in survivors of traumatic brain injury: a critical review. *J Neuropsychiatry Clin Neurosci*. 2009;21:362–370.

Yokobori S, Hosein K, Burks S, Sharma I, Gajavelli S, Bullock R. Biomarkers for the clinical differential diagnosis in traumatic brain injury—a systematic review. *CNS Neurosci Ther*. 2013;19:556–565.

▲ 2.6 Neuropsychiatric Aspects of Movement Disorders

LAURA MARSH, M.D., AND RUSSELL L. MARGOLIS, M.D.

INTRODUCTION

Most often, and consistent with the focus of this chapter, the term *movement disorders* refers to a group of diseases of the CNS that primarily involve neurodegeneration of the basal ganglia, cerebellum, or both. From the clinical perspective, however, movement disorders mark the crossroads of the fields of psychiatry and neurology and underscore the importance of interdisciplinary care and perspectives. This is because psychiatric and cognitive abnormalities figure prominently in many patients, resulting in excess morbidity and caregiver burden. Diagnosis and treatment of neuropsychiatric conditions in patients with primary movement disorders are also challenged by the interplay of motor, cognitive, and psychiatric features over the course of the disease and their degree of overlap at the time of an examination. Examples are the masked face of parkinsonism and the flat affect in depressive disorders. Furthermore, in patients with movement disorders, management of psychiatric disorders is potentially complex because treatments of motor symptoms can exacerbate or cause cognitive or psychiatric symptoms, and, conversely, psychiatric treatments may result in adverse motor and cognitive side effects. Appreciation of the distinctive motor, cognitive, mood, and behavioral aspects of the various movement disorders and their underlying pathophysiology helps to target therapies that are effective and limited in their risks.

Comparative Nosology

Classification of Movement Disorders. Several approaches are used to classify movement disorders (Table 2.6–1). From the phenomenological standpoint, movement disorders are generally grouped into three main categories. Parkinsonian disorders are characterized by reduced amplitude of movements (hypokinesia), slowness of movements (bradykinesia), and difficulty initiating movements (akinesia). Dyskinetic disorders are characterized by increased motor activity and excessive (hyperkinetic) involuntary movements such as tremor, chorea, ballism, tics, dystonia, or myoclonus. Ataxias involve gait disturbances, dysmetria, and dysdiadokinesis. Nosologically, the phenomenological scheme is problematic because clinical features overlap across diseases. In addition, syndromes frequently involve several types of movement abnormalities, such as akinesia, dystonia, and tremor in Wilson disease (WD), while others, such as focal dystonia, consist of a single movement abnormality. Classification schemes based on neuropathologic findings generally

distinguish extrapyramidal diseases, in which the primary pathology is in the basal ganglia system, from diseases that primarily affect the cerebellum, cerebral cortex, or motor neurons. Again, common clinical features and neuronal system involvement across the different neuropathologic groups limit this approach. An etiologic classification has the advantage of linking a syndrome directly to its cause and directing attention at specific modes of pathogenesis, treatment, and prevention. However, the cause of many diseases is unknown (e.g., idiopathic Parkinson disease [PD]), and an etiologic nosology can lead to confusion since entities with clinical similarities are separated into different categories.

Classification of Psychiatric Disturbances Associated with Movement Disorders. In accordance with the *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition (DSM-5), psychiatric disturbances associated with movement disorders can generally be classified as "due to another medical condition," such as depression due to Huntington disease (HD), psychotic disorder secondary to dementia with Lewy bodies (DLBs), or anxiety disorder secondary to PD. Additional qualifiers designate specific features of the presentation, such as *with hallucinations or with major depressive-like episode*. The prominence of psychiatric and cognitive manifestations in movement disorders, however, points out the arbitrariness of this scheme; it could be equally valid for many of the movement disorders to be classified as neurocognitive disorders with associated movement and psychiatric phenomena, for example, probable major neurocognitive disorder with Lewy bodies, with behavioral disturbance. For that reason, many research studies establish diagnoses using the DSM phenomenological criteria for idiopathic psychiatric conditions, that is, from the revised fourth edition of DSM (DSM-IV-TR), and now the DSM-5, while ignoring the criterion stating that the disturbance is "not attributable to the physiological effects of a substance or to another medical condition." Another shortcoming of the DSM nosological system is that its terminology does not capture the nuances of many psychiatric disturbances. Examples are the disruptive perseverative behaviors seen with HD or pathological laughter or crying syndromes that may occur with progressive supranuclear palsy (PSP).

Bedside Assessment of Movement Disorders

Emotion, cognition, and movement are controlled by a series of inter-related and interdependent circuits involving the basal ganglia and frontal–subcortical circuits. It is therefore not surprising that patients with motor abnormalities often have emotional or cognitive disorders and patients with disorders of emotion or cognition frequently have motor abnormalities.

While often relegated to the specialty of neuropsychiatry, the careful appraisal and appreciation of movement abnormalities and movement disorders in psychiatric patients should be considered an important aspect of general psychiatric practice. First, side effects of many psychiatric medications include abnormal movements, some of which may be long-lasting or even permanent. Second, some patients have evidence of abnormal movements before exposure to psychiatric drugs, underscoring the need for baseline motor screening in all patients. Abnormal movements in such patients may indicate the presence of gross congenital disorders (such as cerebral malformations), developmental disorders without gross brain malformation (such as the "soft signs" seen in patients with schizophrenia or autism), neurodegenerative disorders (such as HD or PD), other cerebral insults (including strokes and tumors), or even normal aging. Third, powerful emotions can lead to abnormal movements in otherwise healthy individuals as well as those with CNS disturbances (e.g., tremulousness with anxiety). Fourth, primary psychiatric disorders may present with movement abnormalities, as illustrated by motor slowing seen with depression or the motor hyperactivity with mania. When the motor examination is accompanied by an adequate history and suspicion, overlapping motor and psychiatric signs can be recognized and distinguished according to the relative influences of a primary motor disorder, a psychiatric disturbance, a drug effect, or some interaction among them. For example, an elderly patient with melancholic depression may walk slowly with a hunched posture, but also display unilateral upper extremity rigidity, bradykinesia, and a rest tremor, suggesting a diagnosis of PD and the need for additional treatment beyond antidepressant agents.

Table 2.6–1.
Classification Schemes for Movement Disorders, with Examples

Major Clinical Feature	Location of Primary Neuropathology	Etiologic
Hypokinetic movements	Basal ganglia	Genetic
Idiopathic Parkinson disease	Idiopathic Parkinson disease	Huntington disease
Parkinsonian syndromes	Parkinsonian syndromes	Parkinson disease secondary to α-synuclein or Parkin mutations
Progressive supranuclear palsy	Progressive supranuclear palsy	Friedreich ataxia
Corticobasal degeneration	Huntington disease	Some spinocerebellar ataxias
Frontotemporal dementia	Wilson disease	Frontotemporal dementia with parkinsonism linked to chromosome 17
Dementia with Lewy bodies	Dystonias	Familial amyotrophic lateral sclerosis
Multiple system atrophy—parkinsonian type	Cerebellum (see also Table 2.6–5)	Wilsons disease
Hyperkinetic movements (See also Table 2.6–4)	Friedreich ataxia	Familial essential tremor
Huntington disease	Some spinocerebellar ataxias	Some dystonias
Wilson disease		Medication-induced
Essential tremor		Neuroleptic-induced parkinsonism or dystonia
		Toxin-induced

Dystonias	Cerebral cortex	MPTP, manganese, and carbon monoxide-induced parkinsonism
Ataxia	Frontotemporal dementia	Infectious
Spinocerebellar ataxia	Motor neuron	Post-encephalitic parkinsonism
Multiple system atrophy—cerebellar type	Amyotrophic lateral sclerosis	Traumatic
	Multiple affected regions	Post-traumatic parkinsonism, tremor, dystonia, or frontotemporal dementia
	Multiple system atrophy—both types	Vascular
	Corticobasal degeneration	Vascular parkinsonism or ataxia
	Dementia with Lewy bodies	Idiopathic
	Some spinocerebellar ataxias	Most Parkinson disease
		Progressive supranuclear palsy
		Corticobasal degeneration
		Some cases of frontotemporal dementia
		Dementia with Lewy bodies
		Multiple system atrophy
		Essential tremor
		Sporadic spinocerebellar ataxia

In patients with known movement disorders, the initial history and physical examination should confirm the presence of the diagnosed movement disorder. The causes, course, and treatment of the different movement disorders are diverse, but most share similar features. Given this, there is frequent misdiagnosis of movement disorders; only 75 to 80 percent of patients with parkinsonism are correctly diagnosed with idiopathic PD during life by movement disorder specialists and the rate of misdiagnosis is probably higher among nonspecialists. Examples of disorders that mimic PD, especially early in their course, are PSP and DLB, as well as neuroleptic drug effects in an elderly individual.

A second goal of the motor exam is to evaluate the extent of deficits relative to the reported impairment. Complaints of motor dysfunction can be exaggerated in patients with depression, whereas dementia or lack of insight, as is often the case with HD, can contribute to denial of illness or impulsive behaviors that increase the risk of falls. Lastly, as mentioned above, serial motor exams provide a barometer of motor side effects caused by agents used to treat psychiatric conditions.

The main elements of a motor function screen (Table 2.6–2) include assessment of gait, involuntary movements, coordination, muscle tone, and strength, as discussed further in Chapter 7.4 on Neuropsychiatric Assessment. Gait provides clues to the presence of upper versus lower motor neuron disease (MND), sensory loss, and cerebellar versus extrapyramidal pathology and should first be observed when a patient walks to the examining room. Ability to rise independently from a chair and without pushing off the armrests is a common but nonspecific screen for gait difficulties and increased fall risk in the geriatric population. Balance should be examined separately. Hyperkinetic and hypokinetic movement abnormalities can be observed during the interview while the

patient rests comfortably in a chair. The five main types of abnormal hyperkinetic involuntary movements are tremor, chorea, myoclonus, tics, and dystonia. Having the patient to hold their arms outstretched will reveal forms of nonrest tremor, other adventitious movements, and motor weakness. Finger-to-nose testing permits examination of large proximal arm movements, kinetic tremor, and dysmetria. Hypokinesia, characteristic of parkinsonian syndromes, is manifest by bradykinesia and akinesia. There is a relative absence of spontaneous movements or gestures, slowed movements, decreased arm swing, and diminished facial expressiveness (masked face) and spontaneous blinking. Gait disturbances are related to difficulties initiating and sustaining movements and turning. Coordination can be assessed with finger-thumb tapping. Motor tone is tested by passive movement of an extremity or the neck while the patient is at rest. When assessing tone, the examiner should note whether there is increased resistance to the passive movements, and whether that resistance is active, spastic, or "lead pipe" in character. Cogwheeling, a feature of a tremor, is often mistaken for a measure of tone.

Table 2.6–2.
Screening Exam for Motor System Dysfunction

Gait-observe walking, including on tiptoes and heels, and turning
Station-observe balance, pull test
Abnormal movements—With patient sitting at rest and with arms outstretched, observe limbs, trunk, face, and head for abnormal and excessive movements or hypokinesia
Coordination-rapid finger tapping, finger-to-nose test
Tone-passive movement of limbs and neck
Strength

PARKINSON DISEASE

Introduction and Definition

PD is a progressive neurodegenerative disorder that is associated with loss of dopaminergic neurons in the substantia nigra pars compacta (SNpc). The hallmarks of the disease are its triad of motor features—resting tremor, rigidity, and akinesia/bradykinesia (inability to initiate movement and slowness of movement, respectively). Gait and postural disturbances also characterize the disease. Parkinsonism, or parkinsonian syndromes, refers to the clinical appearance of PD without implying causation. Some patients, for instance, have parkinsonism secondary to antidopaminergic drugs, without pathology of the substantia nigra.

History

Descriptions of parkinsonian syndromes are found in the earliest medical writings. Idiopathic PD is named after James Parkinson, a general

practitioner in London who wrote about *paralysis agitans* as a distinct disease separate from other causes of tremor in his classic 1817 article, "Essay on the Shaking Palsy." Parkinson description of the disease, based on six cases, made a point of distinguishing the resting tremor and festinating gait as major symptoms of the disease and highlighted the cardinal motor signs that remain characteristic of the disease today. Parkinson case descriptions included neuropsychiatric as well as other physical changes, such as constipation. However, his conclusion of the "senses and intellect being uninjured" led to relatively less emphasis on these aspects of PD. Subsequent descriptions recognized memory loss and other neuropsychiatric features, but it is only recently that these have received greater attention. It is worth noting that profound akinesia and speech difficulties may have limited evaluation of the mental state in patients with PD until effective treatment with levodopa became available in the 1960s, thus leaving psychiatric disorders often unrecognized. Although anticholinergic belladonna alkaloids were used to treat PD in the 19th century, the discovery that dopaminergic precursors reversed reserpine-induced parkinsonism led to the treatment of PD with levodopa, still the gold standard of antiparkinsonian medications.

Epidemiology

Parkinsonian syndromes are frequently present in the elderly and their prevalence increases with advancing age, affecting up to 15 percent of individuals up to age 74 and over half of those over age 85. The most common parkinsonian syndrome is PD. It is the second most common neurodegenerative disease after Alzheimer disease, and it affects over 1 percent of the population over age 55 and nearly 3 percent of the population over age 70. Worldwide, an estimated 4 to 6 million people have PD. The mean age of onset for PD is 60 years of age and 80 percent of individuals develop the disorder between the ages of 40 to 70 years. About 5 percent of patients have symptom onset before age 40 and, in such cases, the disease is classified as young onset PD. Juvenile onset occurs in less than 1 percent of cases. The wide range in the age of onset of PD contributes to misdiagnosis, especially in younger patients.

Sex and ethnic differences in the frequency of PD can be an important source of variation because of the association of PD with genetic and environmental factors. Taken together, prevalence studies indicate some age-range associated differences in PD prevalence by geographic location and sex. For example, prevalence among individuals in their eighth decade is higher in European, North American, and Australian populations relative to Asians and men have a higher prevalence of PD as compared to women only for individuals in their sixth decade. The incidence of PD has been reported as lower among African-Americans than Caucasians; discrepant rates may be a function of socioeconomic factors, such as access to health

care and perception of the disease.

Etiology

The neuropathological hallmarks of PD include progressive loss of dopaminergic neurons in the ventral tier of the SNpc and intraneuronal aggregation of the α-synuclein protein in the form of Lewy bodies and Lewy neurites (Fig. 2.6–1A). When approximately 60 to 80 percent of the SNpc neurons degenerate and dopamine production is inadequate, the motor symptoms of PD emerge. What leads to selective and progressive cell death is unknown. Transformation of α-synuclein, a presynaptic protein, to fibrillary and oligomeric forms and their aggregation appears central to neurotoxicity. As such, PD belongs to the family of neurodegenerative disorders referred to as synucleinopathies, which also includes DLB and multiple system atrophy (MSA).

A prevailing theory of the pathophysiology of PD, referred to as the Braak hypothesis, suggests that initial symptoms of PD originate from pathology in the enteric nervous system, medulla, and olfactory bulb, followed by eventual progression to the SNpc over many years. The Braak hypothesis is supported by evidence that nonmotor symptoms, such as anosmia, hyposmia, sleep disturbances, constipation, and anxiety and depressive disturbances, often precede onset of motor symptoms by several years.

Research over the last 15 years has demonstrated that the contribution of genetic factors to the pathophysiology of PD is much greater than was previously recognized. While most cases (>85 percent) are sporadic (not inherited), it is increasingly evident that PD results from effects of multiple predisposing genetic factors, possibly in combination with select environmental insults. Proposed mechanisms of neuronal degeneration include mitochondrial dysfunction, effects of oxidative stress, excitotoxicity, perturbations of neuronal cytoskeleton and axonal transport, and programmed cell death. Though several risk factors have been proposed, the only unequivocal risk factor for PD is increasing age. Dopaminergic cell loss also occurs with normal aging, but the loss is less severe and involves different subpopulations of SNpc neurons than PD. Toxins associated with parkinsonism include viral infection of the CNS and industrial exposures, including manganese, carbon monoxide, organophosphates, and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP). Higher prevalence in industrialized regions and associations with well water consumption, farming, and rural living may reflect exposure to neurotoxins contained in pesticides and other agricultural agents that contain homologues of MPTP. By contrast, other environmental factors and behaviors, such as caffeine use and cigarette smoking, are associated with a reduced risk to develop PD.

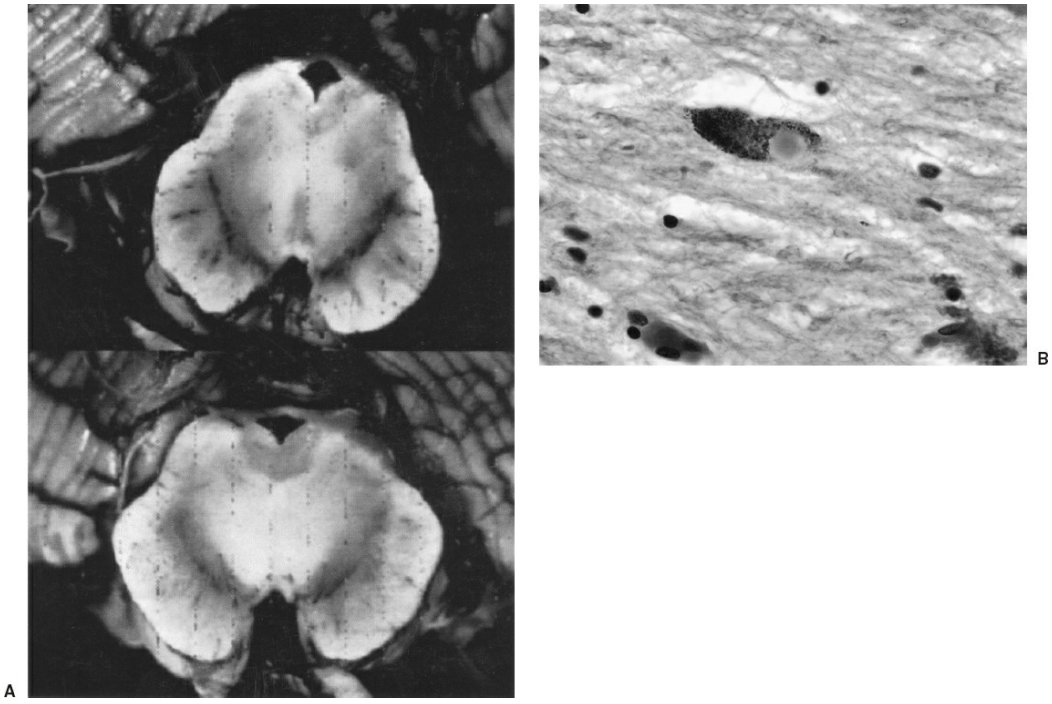

FIGURE 2.6–1. Parkinson disease neuropathology. **A:** Gross pathology in cross section of the midbrain. Note pallor and depigmentation of the substantia nigra in PD (bottom) compare to control (top) brain. **B:** Lewy body in a pigmented neuron of the substantia nigra. Lewy bodies are cytoplasmic inclusions with a dense eosinophilic core surrounded by a clear halo. (Photographs courtesy of The Morris K. Udall Parkinson's Disease Research Center and the Alzheimer's Disease Research Center at Johns Hopkins and Drs. Olga Pletnikova and Juan Troncoso.)

Familial forms of PD are present in about 15 percent of individuals with the disease. Studies of large family and twin samples with heritable parkinsonian conditions have led to identification of 15 causative genes, to date, as well as over 25 genetic variants that increase the risk for sporadic PD. One of the most prevalent genes associated with familial and sporadic PD is leucine-rich repeat kinase 2 (LRRK2), identified in 2004 and inherited in an autosomal dominant pattern. The locus for LRRK2, called PARK8, is located on chromosome 12q12. Its most common pathogenic mutation, G2019S, is found in up to 5 percent of Caucasian patients with familial PD. In Ashkenazi Jews and African Arabs, G2019S mutations are present in up to 30 percent of familial PD cases and 13 to 40 percent of sporadic PD cases. The G2385R LRRK2 variant is the most common genetic risk factor identified among Asian peoples, accounting for 23 percent of familial cases as well as 4 to 9 percent of early and typical late-onset cases. The physiological function of the LRRK2 protein remains unknown.

The α-synuclein gene (SNCA) mutation, located on chromosome 4q21-q23, was the first mutation identified in association with autosomal dominant PD. Although SNCA mutations are a rare cause of PD, the study of α-synuclein led to the discovery that it is a major component of Lewy bodies and demonstrated the role of protein aggregation in the

pathophysiology of PD. Since patients with sporadic PD, who have normal α-synuclein genes, have Lewy bodies that stain positive for α-synuclein and ubiquitin, it is thought that neuronal cell death in PD is related to abnormal accumulations of these proteins. Mutations in a gene on chromosome 6 that codes for Parkin (PARK2), a protein involved in the ubiquitination pathway, are the most common inherited defect and are particularly associated with a juvenile autosomal recessive form of early onset PD. Alterations in gene loci on chromosome 4 that affect ubiquitin processing (UCH-L1) and on chromosome 1 that affect β-glucocerebrosidase and lysosomal function (GBA) are also linked to PD. Other identified genes associated with parkinsonian syndromes include the oncogene DJ-1 (PARK7), the MAPT (tau) gene, progranulin (PGRN), FBXO7 (PARK-15), and PINK1. Mutations in some of these genes may be relevant to cases that appear to be sporadic.

Current models of basal ganglia/cortical physiology suggest the presence of five adjacent but anatomically distinct frontocorticostriatothalamic loops with motor, oculomotor, limbic, anterior cingulate, and prefrontal targets and that dysfunction of these neural loops is linked to the motor and nonmotor features in PD and other basal ganglia disorders. A common feature of these cortical–subcortical loops is their projection from higher level cortical areas to the basal ganglia via striatal areas with outflow through the internal globus pallidum (Gpi) or its histologic analogue, the substantia nigra pars reticulata (SNpr) (Fig. 2.6–2). Gpi/SNpr outputs then project to specific areas of the thalamus that complete the loops through projections to the original cortical area. Within each loop, the striatonigral component includes a “direct” pathway that projects directly from striatum to Gpi/SNpr and an “indirect” pathway that flows from the striatum through the subthalamic nucleus (STN) to reach Gpi/SNpr.

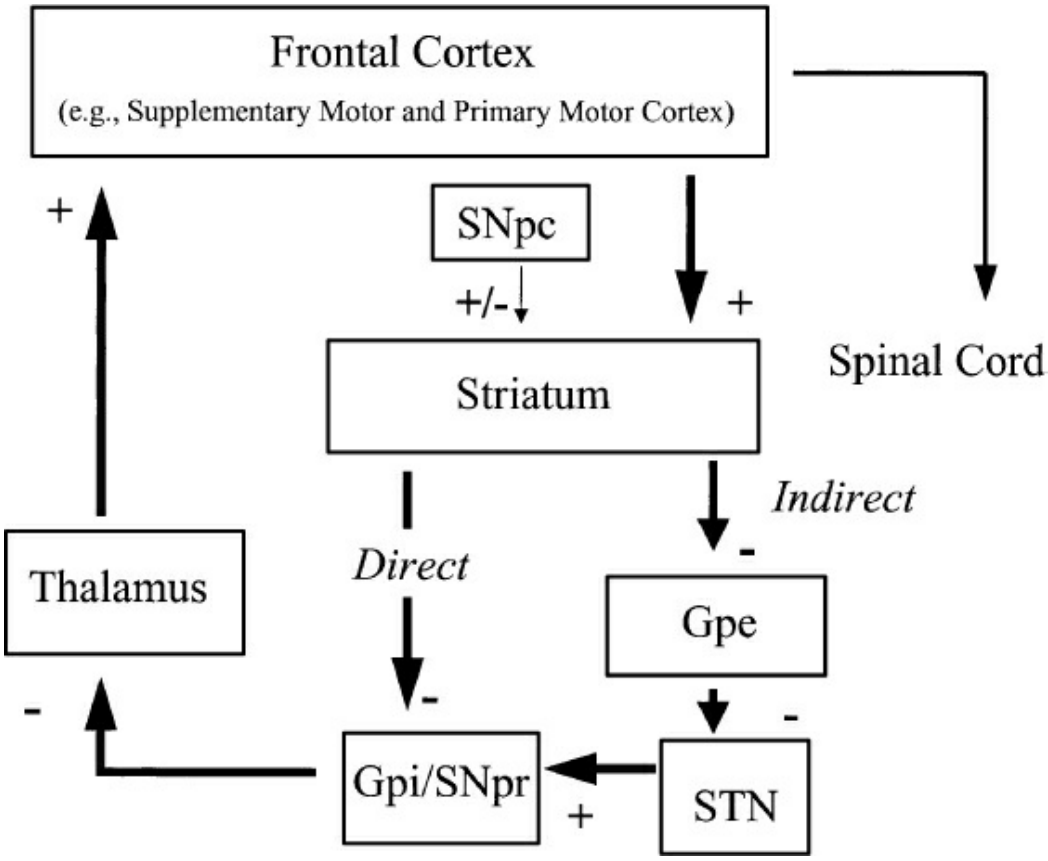

FIGURE 2.6–2. Corticostriatalthalamic circuits. In the motor circuit, the supplementary motor area and primary motor cortices in frontal regions project mainly to putamen and then "directly" to Gpi or "indirectly" via the external globus pallidus (Gpe) and then the STN to reach Gpi. Excitatory (glutamatergic) cortical projections generally decrease Gpi activity through the direct pathway and increase Gpi activity through the indirect pathway. Inhibitory (GABA-ergic) Gpi projections to thalamus and excitatory thalamocortical projections result in net excitation (increased motor activity) or net inhibition (decreased movement) through the direct and indirect pathways, respectively. Dopamine has differential inhibitory and excitatory effects on the nigrostriatal pathway.

Motor Manifestations. Human and animal models of PD demonstrate that loss of SNpc dopaminergic neurons affects frontocorticostriatothalamic circuits sufficiently to result in the motor manifestations of PD. The best studied of these neuronal circuits is the "motor circuit" as illustrated in Figure 2.6–2. In PD, dopaminergic degeneration appears to have differential effects on the nigrostriatal pathway. Projections from SNpc to striatal targets are inhibitory in the dopamine type 1 (D1) receptor-mediated direct path and excitatory in the D2 receptor-mediated indirect path. Normal movement results from coordinated regulation of thalamocortical excitation by the direct and indirect pathways, with the presence of nigrostriatal dopaminergic projections tending toward increased movement by facilitation of the direct pathway and inhibition of the indirect pathway. Loss of dopamine results in hypokinetic symptoms secondary to overactivity of the STN and Gpi via the indirect path, which increases inhibitory input to the thalamus and then

reduces the excitatory thalamocortical activity that ordinarily facilitates motor movements. Hyperkinetic movements, such as levodopa-induced dyskinesias, occur with overactivation of the direct pathway.

Cognitive Manifestations. Dementia in PD is associated with increasing age, a family history of dementia, depression, and more severe motor dysfunction. Neuropathological and imaging studies in patients with PD with dementia are associated with neuronal loss in cholinergic cells of the basal forebrain, suggesting a role for cholinergic deficits in PD-related cognitive impairment. Cortical Lewy bodies may also be associated with dementia. The histopathological changes of Alzheimer disease—neurofibrillary tangles and senile plaques—are seen in some patients with PD and dementia. Other cognitive impairments, in which higher-order associative functions are preserved, may involve dysfunction of nonmotor neural circuits, including mesocortical and mesolimbic dopaminergic projections. Bradyphrenia may be associated with noradrenergic cell loss.

Psychiatric Manifestations. Other than PD itself, there are no known specific risk factors for psychiatric disturbances in PD. Though psychological reactions influence occurrence of depression, most evidence suggests that depressive disorders are related to the underlying disease processes of PD. In particular, development of depressive and anxiety disorders before motor symptoms suggests that these mood symptoms are either neurological signs of the disease or a risk or causative factor for development of PD. Cross-sectional studies do not show consistent relationships between depression onset and severity to age onset or duration of PD, motor severity, disease subtype (tremor dominant versus akinetic-rigid), or disease stage. Several studies link depression with excessive serotonergic loss from the midbrain raphe. Depression in PD is also associated with decreased blood flow in frontal regions, a finding also seen in idiopathic depression. In apathetic syndromes, a role for noradrenergic dysfunction is suggested by a relationship between bradyphrenia and neuronal loss in the locus coeruleus. Several studies implicate noradrenergic dysfunction in anxiety disorders, including an imbalance in noradrenergic/dopaminergic tone. Psychosis in PD is most often associated with dopaminergic therapy, which may lead to hypersensitivity of mesocortical and mesolimbic dopaminergic receptors. Other factors related to the incidence of psychosis are cognitive impairment, disease severity, mood disorders, impaired visual acuity, other psychoactive medications, and delirium. The association of psychosis with cognitive impairment and mood disorders implicates more widespread brain pathology involving other neurotransmitter systems, including cholinergic deficits or a serotonergic/dopaminergic imbalance.

Diagnosis and Clinical Features

Motor Manifestations. Motor symptoms initially affect one side of the body and gradually progress to the contralateral side, though remaining asymmetric. The slow (3 to 6 Hz) classic resting tremor is often the presenting feature of PD, but other types of tremor may be present and 20 to 30 percent of cases have no tremor at all. Bradykinesia and akinesia are the most disabling aspects of PD and contribute to problems with dexterity and gait. Patients typically walk in a flexed forward posture without swinging their arms. The gait is slow and shuffling with small steps and a tendency to fall forward and hasten the rate of walking (festination). Difficulty initiating movements leads to start hesitation and "freezing." For example, a patient may suddenly be unable to move their legs while walking or their hands when reaching into a purse. Deteriorated handwriting, weakness, and clumsiness in one limb are also common initial symptoms of bradykinesia. Rigidity is appreciated as increased muscle tone on exam. Patients may note muscle cramps, weakness, or stiffness as the first indication of rigidity.

The diagnosis of PD is clinically based since there is no biological marker for the disease. Most experts require the presence of two of the three cardinal motor signs to diagnose PD along with a robust response to an initial trial of levodopa. Early in its course, PD can be difficult to diagnose since the signs of the disease may be subtle and not even perceived by the patient. Parkinsonian signs can also be a feature of normal aging or other parkinsonian syndromes. The initial symptoms in up to 20 percent of patients are nonmotor, including fatigue, musculoskeletal complaints, and depression. Other nonmotor somatic symptoms of PD include sialorrhea, dysarthria, visual and genitourinary dysfunction, sleep disturbances, sweating, seborrhea, edema, constipation, paresthesias, and hyposmia. In many patients, a prodromal phase, evident about 4 to 8 years before the onset of obvious motor signs, is often characterized by depression, anxiety, and musculoskeletal discomfort. Heightened vigilance for clinical signs of PD during this phase might result in earlier disease detection, especially in older adult patients who present with depression or anxiety.

Cognitive Manifestations. Some degree of cognitive impairment, ranging from mild executive dysfunction to dementia, affects nearly all patients with PD; about 25 to 40 percent of patients develop dementia. Commonly, cognitive decline is not obvious clinically and scores on tests that screen for dementia, such as the Mini Mental State Exam (MMSE), remain in the "normal" or nondemented range for some patients who nonetheless have dementia. The concept of "subcortical dementia" has been used to describe the cognitive impairment that is frequently

associated with the various movement disorders characterized primarily by subcortical neuropathology. Executive dysfunction is especially common, resulting from deficits that affect the ability to process new information and anticipate, plan, initiate, maintain, and change behaviors. Memory, visuospatial skills, information processing speed, attention, explicit recall, spatial planning, perseveration, verbal fluency, and poverty of thought are also affected. Compared to patients with Alzheimer disease, recognition memory, aphasia, agnosia, apraxia, and higher language functions are relatively spared. There are patients with PD, however, who demonstrate wider involvement of cortical functions, including aphasia, apraxia, and memory deficits, and, in some cases, a cognitive presentation indistinguishable from Alzheimer disease.

Psychiatric Manifestations. Psychiatric disturbances affect up to 90 percent of patients at some point during the course of PD and more than one psychiatric disturbance is often present. Early in the disease process, adult-onset anxiety and depressive disorders precede the obvious onset of motor symptoms in up to 30 percent of patients.

DEPRESSION. Normal psychological reactions to the illness can result in low mood, grief, demoralization, frustration, and embarrassment. Depressive disorders in PD involve more pervasive mood changes and generally resemble idiopathic forms of depression.

Reported prevalence rates range from 20 to 90 percent, with an average reported frequency of 40 to 50 percent. Up to half of those with depression, in most studies, have major depression. The majority of PD patients with depression have nonmajor depressive syndromes that are variably referred to as minor depression, persistent depressive disorder (dysthymia), or subsyndromal depression. Nonmajor forms of depression often progress to major depression if untreated. Anxiety symptoms and comorbid anxiety syndromes such as panic disorder are especially prominent, but self-blame, guilt, delusions, a negative self-attitude, self-destructive thoughts, a sense of failure and anhedonia may be less evident. Suicide can occur, but at a similar rate as for the general population. Other psychological, vegetative, and autonomic features of depression are also common, but early morning awakening, anergia, and retardation occur at comparable rates in nondepressed patients with PD. Depression is not consistently associated with a family history of mood disorders.

APATHY. Apathy occurs in at least 25 percent of patients and often coexists with depression. It manifests as indifference and a lack of motivation, initiative, perseverance, interest in new things, or concern for one's health. In most cases, apathy is not distressing to the patient, but the absence of spontaneous effort and interest can be frustrating for caregivers, who may perceive the patient as depressed. Inactivity, deconditioning, and

fatigue contribute to further loss of function and increased disability and suffering. Other neurological signs and symptoms such as akinesia, hypomimia, hypophonia, cognitive dysfunction, and bradyphrenia can confound recognition of apathy. SSRIs, the most frequently prescribed antidepressants for patients with PD, can also cause apathy, though this has not been demonstrated in PD specifically.

ANXIETY. Difficulties initiating movements can frequently lead to anxiety-provoking circumstances in the course of living with PD, but pathological anxiety is still more common in PD compared to other parkinsonian syndromes. Particularly common are generalized anxiety disorder, social phobia, and panic disorder, which have a prevalence of 25 percent in some series. In general, anxiety syndromes in PD resemble those in idiopathic conditions and frequently cooccur with depression. There are also anxiety disturbances unique to PD, particularly anxiety associated with fluctuations in levodopa levels and motor function, although there is not always a strict temporal relationship between anxiety and motor "off" states. Some patients describe episodic panic attack-like states that develop at the same time every day and can be associated with motor deficits. Sleep difficulties are very common and seem to aggravate daytime anxiety.

PSYCHOSIS. Psychosis affects about 25 to 40 percent of patients, with increased prevalence as the disease progresses. The temporal pattern and clinical profile of psychotic symptoms in PD are different from the pattern evident in other psychotic disorders such as substance-induced psychosis or schizophrenia. Lack of standardized diagnostic criteria for PD-associated psychosis contributed to variability in reported prevalence. As studies were often restricted to visual hallucinations, the most common psychotic symptom in PD, new criteria were proposed for in 2007 by a National Institutes of Health (NIH)-sponsored work group.

In PD, psychotic symptoms include hallucinations and delusions. Atypical symptoms, referred to as minor psychotic phenomena, affect over two-thirds of patients and include sense of presence, passage hallucinations, and illusions. It is a misnomer to refer to minor symptoms as "benign," as they, like other psychotic symptoms in PD, are associated with more depressive symptoms and worse quality of life. While visual hallucinations are most common, and are often related to dopaminergic medications, other sensory modalities can be involved. Typically, in earlier stages of PD, visual hallucinations include well-delineated animals or people that occur in a clear sensorium. They are not necessarily distressing, and insight may be retained. Later in the disease course, some patients experience persistent hallucinations and/or delusions in a clear sensorium but with diminished insight. Aggression, delusional jealousy and other paranoid accusations, delusional fears about eating or medications,

inappropriate sexual behaviors, and confusion are frequent manifestations that can be difficult to manage and may require hospitalization. Some patients develop chronic paranoid syndromes that resemble schizophrenia. Hallucinations and delusions also occur with major depression, mania, or anxiety disorders, and may be overlooked when patients are especially agitated. Fragmented sleep, vivid dreams, and nightmares may herald the onset of chronic psychosis. Hallucinations and delusions associated with advanced PD are a major reason for nursing home placement. A third category includes hallucinations or delusions in the context of delirium related to antiparkinsonian medications and/or other causes.

IMPULSE CONTROL DISORDERS. Impulse control disorders (ICDs), most commonly pathological gambling, hypersexuality, uncontrolled spending, and binge eating, are observed in up to 14 percent of patients with PD. About one-third of affected patients have more than one ICD. Sporadic cases of hypersexuality were reported in the medical literature in the 1970s, after the introduction of levodopa. However, development of ICDs is most strongly associated with use of dopamine agonists for PD (e.g., pramipexole, ropinirole, and rotigotine) as well as for other indications, for example, restless legs syndrome, warranting patient education and close monitoring for ICDs when dopamine agonists are prescribed. Since not all individuals taking dopamine agonists develop ICDs, additional monitoring is prudent in the context of other associated risk factors for development of ICDs, including younger age of PD onset, unmarried status, family history of problem gambling, cigarette smoking, caffeine use, and a novelty-seeking temperament. In most cases, ICDs resolve after the offending medications are discontinued.

ICDs are associated with devastating social consequences, including serious financial loss, bankruptcy, divorce, other legal problems, and loss of employment. Unfortunately, ICDs often elude detection or, when referred for psychiatric care, the link to a potential side effect of PD medications is not appreciated. In part, this is because ICD behaviors can begin as early as a few months after starting dopamine agonists, but may also develop only after years of treatment, such as when the dose is increased, or when levodopa and dopamine agonists are used together. Patients are unlikely to divulge uncontrollable urges to engage in sexual activity, gamble, or spend excessively, and the behaviors may be unusual or idiosyncratic. For example, hypersexuality, seen in two to six percent of patients with PD, usually presents as an increase from baseline sexual activity, but paraphilias including transvestitism, pedophilia, and flagrant promiscuity may develop. Spending may involve purchases of large quantities of unnecessary items, for example, lamps, or excessive and reckless generosity. Such behaviors may be misinterpreted as a way of “coping” with PD.

Other significantly disruptive behaviors, often referred to as compulsive behaviors, are associated with antiparkinsonian treatment, namely levodopa. Punding refers to intense time-consuming preoccupation and fascination with seemingly meaningless rituals or stereotyped activities, for example, arranging and rearranging objects. Hobbyism involves intense time-consuming preoccupation with specific hobbies or activities, for example, endlessly working on projects or hobbies, or computer addiction. Walkabout refers to excessive aimless walking or driving. The dopamine dysregulation syndrome is a challenging condition in that the patient has an addiction-like overuse of dopaminergic medications that is way beyond what is needed to control motor symptoms and despite harmful side effects. When agonists are eliminated, a dopamine-agonist withdrawal syndrome with addiction-like withdrawal symptoms (anxiety, panicky feelings, diaphoresis, irritability, dysphoria, drug craving, pain, and fatigue) can complicate clinical management. Mania or euphoric syndromes are generally related to dopaminergic medications and are often associated with mood swings and on-off phenomena, but can reflect a premorbid disorder. Emotionalism refers to heightened and excessive sentimentality, often with pathological tearfulness (emotional incontinence) that can be precipitated by a variety of emotions; it can be a feature of a depressive disorder.

PREMORBID PERSONALITY. There is controversy over whether certain premorbid personality characteristics prevail in patients with PD. Many observers have noted premorbid personality features that include industriousness, cautiousness, perfectionism, and punctuality in individuals who ultimately develop PD, with persistence of these traits over the course of the illness. One argument is that these traits represent a consequence of premorbid reductions in dopaminergic tone.

Pathology and Laboratory Examination. Unlike other common neurodegenerative disorders, there are no accepted neuropathological criteria for PD. Most neuropathologists agree that dopaminergic neuronal loss in the SNpc is a requisite finding. However, the pathological process precedes the clinical presentation since classic motor signs of PD are not evident until there is 70 to 80 percent loss of nigral neurons. In addition to dopaminergic cell loss, Lewy bodies, which are intraneuronal aggregates of the protein α-synuclein, are present in most forms of PD (Fig. 2.6–1B) and are seen in the substantia nigra, locus coeruleus, dorsal vagal nucleus, and nucleus basalis of Meynert. Lewy bodies are not specific to PD, and are seen in a variety of conditions, including Alzheimer disease. PD is also accompanied by degeneration of noradrenergic neurons in the locus coeruleus, serotonergic neurons in the dorsal raphe, and cholinergic neurons in the nucleus basalis and their attendant projection systems.

Alzheimer type neuropathological changes (i.e., senile plaques and neurofibrillary tangles) are present in some patients with PD and dementia. Differential degrees of pathology among affected neuronal systems are thought to underlie the heterogeneous motor, cognitive, and psychiatric features of PD.

It is important to rule out other causes of parkinsonism. Laboratory tests should be conducted to rule out Wilson's Disease and hypothyroidism. Neuropsychological testing can help to clarify the extent of cognitive deficits, evaluate cognitive status in potential candidates for neurosurgical treatments, or establish alternative diagnoses. Dementia in particular is an exclusionary criterion for surgical therapies since it is associated with a poor motor outcome but also limits compliance. An atypical cognitive profile for PD also predicts a poor outcome. Structural neuroimaging studies may help exclude other diagnoses, such as vascular disease, normal pressure hydrocephalus, and other Parkinsonian conditions, but are generally not needed to diagnose classic PD. With functional imaging, decreased dopaminergic uptake in the putamen is characteristic of PD and subclinical dopaminergic changes are evident in some at-risk relatives of kindreds with familial PD. Single-photon emission tomography (SPECT) and PET can be helpful for diagnosing PD in unclear cases with atypical signs and a poor response to levodopa and for assessing presynaptic dopaminergic integrity as a marker of disease progression.

Differential Diagnosis. Several clinical features help distinguish PD from other parkinsonian disorders. PD progresses more slowly, postural instability is a later finding, and there is an excellent and sustained response to levodopa. Alternative diagnoses, such as PSP or MSA, are suggested by rapid progression with development of postural instability and early falls (within 3 years), a poor or transient response to levodopa, orthostatic hypotension and other signs of autonomic dysfunction, early dysphagia, and supranuclear gaze palsy. Early hallucinations, unrelated to treatment, and early dementia suggest a diagnosis of DLB or Alzheimer disease with associated PD. Myoclonic jerks, pyramidal signs (e.g., Babinski sign), and cerebellar signs (e.g., wide-based gait), are not features of PD, but are seen in other parkinsonian disorders. Other non-PD associated findings include alien limb syndrome or ideomotor apraxia.

It is always important to consider alternative diagnoses to depression. Apathetic syndromes, dementia, or delirium can all cause changes in mood that are usually more apparent to families than the patient. The differential diagnosis of depression in PD also includes bipolar disorder, adjustment disorders, anxiety syndromes, psychosis, pathological tearfulness (emotional incontinence), substance abuse or dependence, mood lability associated with parkinsonian motor fluctuations, and increased medical debility. Independent of a depressive disorder, sleep disturbances and

daytime fatigue in PD can impair function and lead patients to feel discouraged. PD-related depression is also frequently associated with anxiety disorders (e.g., panic disorder) and dementia. The somatic and autonomic symptoms of anxiety can overlap with features of other medical conditions. Psychosis in the early stages of PD suggests alternative diagnoses, such as DLB or Alzheimer disease with extrapyramidal features. In some cases, pre-existing psychopathology emerges with dopaminergic treatment.

Course and Prognosis. PD is clinically heterogeneous. In each case, there is a different clinical presentation, age onset, natural history, and response to levodopa. Many cases of PD can be subdivided into either a tremor predominant form, or one characterized by prominent gait disorder and postural instability, the former having a more benign prognosis. Heterogeneity in clinical onset, progression, and levodopa response also influence the natural history of PD. This includes the onset of motor fluctuations and other treatment-related complications, longevity, and functional impairment. The rate and course of cognitive deterioration is also highly variable, ranging from mild-to-severe. Successful treatment of depression improves motor status but medical treatment of motor symptoms does not usually improve depressive syndromes. Even in the absence of depression, antidepressant medications and ECT have been shown to improve motor function. As in the general population, anxiety disorders often coexist with depressive episodes, but anxiety disorders tend to be chronic. The need for nursing home placement, often triggered by psychosis, is associated with a higher mortality rate.

Treatment

Motor Symptoms

PHARMACOLOGICAL TREATMENTS. Current treatment options for PD provide only symptomatic relief of the motor deficits; there are no known neuroprotective or regenerative strategies, though this is an active area of research. Consequently, the gold standard of treatment for PD remains levodopa, which increases the life expectancy among patients with PD and significantly improves function and quality of life. During early stages, anticholinergic medications such as trihexyphenidyl or benztropine, amantadine, and selegiline provide mild-to-moderate reductions in motor symptoms. Dopamine agonists such as pramipexole, ropinirole, and pergolide, and/or levodopa may then be added to address progressive disability. Dopamine agonists are especially effective early treatments that can delay initiation of levodopa therapy. Postural instability does not usually respond to antiparkinsonian medications.

There is debate over when to begin levodopa therapy since it is

associated with dose-limiting motor side effects (in addition to psychosis). Initially, there is tremendous improvement in motor function, but higher doses of levodopa are gradually required to achieve similar benefit and the duration of the motor response wanes. Fluctuating motor effects then replace sustained motor benefits, with improved motor function in response to medications (the "on" state) and reduced mobility when medication effects subside (the "off" state). Patients also develop abnormal involuntary hyperkinetic choreiform or dystonic movements during the "on" state, referred to as drug-induced dyskinesias. Agents that inhibit dopamine metabolism via catecho-o-methyltransferase inhibition, such as tolcapone and entacapone, can reduce motor fluctuations.

Psychiatric Disturbances Related to Fluctuating Effects of Dopaminergic Therapy. Emotional influences can trigger or intensify motor fluctuations and drug-induced dyskinesias, themselves, can be distressing or embarrassing. In the extreme, "on-off" fluctuations range from states of complete immobility or "freezing" to severe writhing. Fluctuating motor states may be accompanied by nonmotor phenomena, including akathisia, anxiety, dysphoria, autonomic symptoms, irritability, apathy, hallucinosis, psychosis, screaming, and cognitive slowing during "off" periods; hypersexuality, hypomania, and thought racing may occur during "on" periods.

SURGICAL TREATMENTS. Before availability of levodopa in the 1960s, neurosurgical procedures were standard treatments for tremor and rigidity in PD. More refined stereotactic neurosurgical therapies currently used for PD involve high-frequency deep brain stimulation (DBS) of the STN and globus pallidus pars interna (Gpi) or ablative procedures that target the GPi (pallidotomy) or thalamus (thalamotomy). Both improve motor function because they reduce the hyperactivity of the STN and its efferent targets, thus partially restoring thalamic outflow. Thalamotomy and thalamic DBS are indicated for disabling tremor. STN DBS and pallidal interventions also improve tremor, but are mainly indicated for severe dyskinesias, bradykinesia/akinesia, and rigidity. The procedures are aimed at alleviation of contralateral symptoms, but bilateral benefits are seen for dyskinesias and bradykinesia. Embryonic cell transplantation of dopamine neurons, first performed in the 1980s, remains an experimental procedure.

Neuropsychiatric Effects of Neurosurgical Therapy for PD. Initial studies showed limited if any adverse cognitive effects after ablative or DBS procedures when surgical candidates underwent careful screening for dementia. Some pallidotomy studies found changes in executive function, memory, verbal fluency, and attentional set shifting, but lesion size may have been influential. Stimulation parameters, laterality and site-specific

effects, electrode placement, and stimulation sites may also affect cognitive outcomes.

While many patients report improved quality of life after surgery, acute stimulation with DBS can be associated with profound depressive or manic symptoms, including in patients with no prior psychiatric history. In some patients, pre-existing disorders, such as depression, anxiety, or drug dependence, and interpersonal conflicts worsen and overshadow any motoric benefits. DBS treatment-emergent sexually deviant behaviors, gambling, mood lability, apathy, and aggression are also described. Reductions in anxiety and depression, unrelated to improved motor status, are reported after unilateral pallidotomy, but bilateral pallidotomy can result in depression and apathy. Studies of pallidotomy show no consistent effects on depression or anxiety, but a history of a prior or active affective disturbance at the time of surgery may be associated with development of a mood disorder.

COGNITIVE MANIFESTATIONS. Several large randomized controlled trials on treatments of cognitive impairment or dementia in PD or DLB have been conducted. Presence of cholinergic deficits in PD and their relationship to dementia suggests a role for cholinesterase inhibitors. Intervention trials for PD-related dementia or DLBs show that cognition and psychiatric symptoms, including psychosis, improved with donepezil and rivastigmine, though the data are more limited for galantamine. Cholinesterase inhibitors are not always well-tolerated; the most common side effects are nausea, vomiting, anorexia, dizziness, drowsiness, and insomnia and may be more common with rivastigmine. Memantine, while well-tolerated, has not demonstrated consistent benefit, though perhaps more so in DLB. A recent systematic review indicates that both cholinesterase inhibitors and memantine result in slight improvement in global impression, but cognitive performance improved only with cholinesterase inhibitors.

A few studies have investigated treatment of mild cognitive impairment in PD (MCI-PD), though inconsistent definitions of MCI-PD confound interpretation of the results. Cholinesterase inhibitors appear to have some potential benefits, but targeted treatment of noradrenergic deficits may also be beneficial. Cognitive rehabilitation programs are increasingly recognized as being beneficial to cognitive performance and functional abilities in patients with PD, including as adjunctive therapy to medications. Additionally, patients and families benefit from education about cognitive deficits in PD, especially their consequences on executive functioning and strategies for minimizing distress related to selective cognitive impairments, and, later in the course, dementia.

PSYCHIATRIC MANIFESTATIONS. A first step in the treatment of psychiatric disorders is to evaluate the contribution of the antiparkinsonian

medications, fluctuating medication effects, other medical conditions, cognitive deficits, and caregiver and other psychosocial variables to the presenting problem. The common occurrence of more than one psychiatric disturbance also influences treatment. Sleep is commonly disrupted in PD, independent of psychiatric problems, and addressing it alone can reduce daytime mood lability, psychosis, and motor impairments. Initial treatment involves consideration of the antiparkinsonian regimen, sleep hygiene practices, and use of a low dose sedating antidepressant or hypnotic such as trazodone (e.g., 25 to 100 mg). Benzodiazepines can be used judiciously, but are not considered first-line therapy for anxiety, agitation, or sleep disturbances because of effects on cognition, falls, and interdose withdrawal symptoms. Patient and caregiver education, especially about the psychiatric condition and its overlap with parkinsonian features, and supportive therapy are cornerstones of effective treatment. Parkinson support groups can be invaluable sources of education and community support. Ancillary treatment by speech, physical, and occupational therapists can help patients to maximize function, even early in the disease and periodically as the disease progresses and new impairments emerge. Regular exercise helps motor symptoms and overall well-being.

DEPRESSION. For many years, despite the prevalence and seriousness of PD-related depression, published controlled trials on antidepressant treatment in PD were of poor quality. Over the last decade, increased research on various modalities to treat depression in PD provides some guidance, though there remains a need for larger, well-designed trials including behavioral and pharmacological treatments. The most recent systematic review and meta-analysis indicates that pharmacological treatments, especially using serotonin-reuptake inhibitors (citalopram, sertraline, fluoxetine, and paroxetine), and CBT, result in significant improvements in depressive symptoms in both major and nonmajor depressive syndromes. Longer-term benefits of these treatments have not been assessed. Tricyclic antidepressants (nortriptyline, desipramine, and doxepin) showed a large pooled effect that was not statistically significant. Trials of nonantidepressant medications (pramipexole, pergolide, atomoxetine, and memantine) were not efficacious. rTMS studies showed improvement in depressive symptoms, but were not sufficiently powered to yield firm conclusions.

As with primary depression, antidepressant selection is usually based on consideration of side effect profiles. Serotonin reuptake inhibitors can aggravate motor symptoms, but this is not especially common. Use of selegiline with antidepressants carries the risk of inducing a serotonin syndrome or hypertensive reactions, but most patients tolerate this combination. In clinical practice, ECT is safe and effective for depression in PD, especially in the context of psychotic symptoms.

ANXIETY. In spite of the high prevalence of anxiety disturbances in PD and their impact, there is limited treatment data for this population. The most extensive data shows improvement of anxiety symptoms with CBT, though treatment targeted anxiety symptoms versus specific anxiety disorders. The one randomized controlled medication trial showed that bromazepam, a long-acting benzodiazepine, improved psychological and somatic anxiety symptoms. While benzodiazepines can be effective in reducing anxiety, unfavorable effects on alertness, cognition, and gait, and an increased risk of falls warrant their judicious use, with careful evaluation of known risks and putative benefits before prescribing even for short-term use. SSRI antidepressants have been shown to improve anxiety when measured as a secondary outcome in depression treatment trials in patients with PD. Buspirone, a nonbenzodiazepine anxiolytic with partial 5-HT agonism, has been used in PD but its specific effects on anxiety are not obvious. Gabapentin is used for treatment of anxiety disorders, but its use in PD has not been examined. The utility of antipsychotic medications as adjunctive therapy is also not known.

APATHY. Studies on treatment of apathy in PD are limited, in part because of practical and conceptual barriers, such as lack of established diagnostic criteria or a gold-standard assessment scale, and confounding conditions. Results of trials on the effects of rasagiline, an irreversible inhibitor of monoamine oxidase, and rotigotine, a dopamine agonist on apathy are pending publication. Treatment of comorbid depression is an obvious initial approach, but there may be residual or SSRI-induced apathy. Antidepressants that inhibit dopamine or norepinephrine reuptake are suggested by neurochemical theories of apathy. If depression is not an issue, dopaminergic agents, including levodopa, dopamine receptor agonists, selegiline, amphetamines, and amantadine, have been most consistently effective; the absence of cognitive dysfunction predicts a better response. In Alzheimer and PD, cholinesterase inhibitors may reduce apathy. Subthalamic DBS can induce or aggravate apathy in some patients

PSYCHOSIS. Treatment of psychosis in PD is challenged by the delicate balancing act between the use of medications that relieve psychosis without aggravating motor or cognitive dysfunction and treat motor symptoms without exacerbating the mental status. Caregiver safety is also an issue because of immobility, directed aggression, and combativeness during care. Reducing or simplifying the antiparkinsonian regimen usually involves elimination of anticholinergics, amantadine, and selegiline; in addition, levodopa is generally preferred to dopamine agonists and controlled release levodopa when there is psychosis. When higher doses of levodopa are required or disruptive symptoms persist, treatment usually requires addition of an atypical antipsychotic.

There is a solid evidence-base for treatment of psychosis in PD. The gold standard is clozapine; it is effective even at low doses (<25 mg), does not block D2 receptors, and has been evaluated in several controlled trials. Because of the requisite blood monitoring with clozapine, quetiapine is usually the initial drug of choice, but some patients eventually experience worsened motor function and patients with dementia may be especially vulnerable. Controlled trials of quetiapine, however, have not been shown to be efficacious. Orthostatic hypotension, headache, and nausea also contribute to its intolerance. Other "atypical" neuroleptic agents (olanzapine and risperidone) reduce or eliminate psychosis at low doses, but their effects on striatal D2 receptors are frequently associated with significant motor deterioration. Pimavanserin, a selective serotonin 5-HT2A inverse agonist without adverse motor effects, is a promising drug for PD-related psychosis that is undergoing regulatory review. ECT can be used to treat psychotic depression as well as dopamine-induced psychosis. Cholinergic agents can also have antipsychotic effects.

Other Parkinsonian Conditions

Progressive Supranuclear Palsy

EPIDEMIOLOGY AND ETIOLOGY. PSP, an important cause of parkinsonism, was first described in 1964. The prevalence is 5 to 6 cases per 100,000, and PSP accounts for 4 percent of all cases with parkinsonism. Onset of symptoms is usually after age 60 in all races and both genders, but men tend to be affected more often than women. There are rare familial cases of PSP, but its development is usually sporadic. PSP is classified as a tauopathy, a term referring to neurodegenerative disorders characterized by intracellular aggregations of abnormally phosphorylated tau protein. Tau normally stabilizes microtubules and its aggregation leads to formation of abnormal cytoskeletal fibrils. The cause of the tauopathy in PSP, and its role in neurodegeneration, remains uncertain.

CLINICAL FEATURES. The defining motor features of PSP are progressive parkinsonism with symmetric bradykinesia, axial rigidity, and paralysis of downward vertical gaze. The problems with ocular motility may not be evident for several years, but cause problems with walking downstairs. Dysphagia develops over the course of the illness. Unlike PD, postural instability leading to falls and dysarthria are early and often presenting features. Also, there is a poor response to levodopa, tremor is uncommon and stride length is erratic. Posture tends to be upright or sometimes opisthotonic because of axial rigidity. Instead of a lack of facial expression (hypomimia), patients appear "worried" or "astonished," resulting from dystonic contractions of facial muscles between the eyebrows, reduced blinking, lid retraction, and gaze palsy.

Over the last decade, there has been an increased appreciation of the phenotypic variability of PSP and the categorization of atypical PSP syndromes that later develop other motor features of PSP and have evidence of PSP pathology. These include PSP presenting with apraxia of speech (PSP-AOS), as Parkinson disease-like (PSP-P) with a jerky postural tremor and levodopa response, as the corticobasal syndrome (PSB-CBS), with cerebellar ataxia (PSP-C), with progressive akinesia and gait freezing (PSP-PAGF), and with the behavioral variant of frontotemporal dementia (PSP-bvFTD).

Cognitively, patients with PSP show features of a subcortical dementia with bradyphrenia, memory deficits, and predominant frontal lobe dysfunction with executive, attentional deficits, memory, visuospatial, and language and social cognitive deficits. Approximately 10 percent present with cognitive symptoms and up to 70 percent develop a dementia over the disease course. Aphasia, apraxia, and agnosia, the classic features of "cortical dementias," are not evident. Behavioral changes of apathy, irritability, childishness, impulsivity, and disinhibition are compatible with damage to orbitofrontal and medial frontal circuits and are the most common psychiatric features. There are also sleep disturbances, depression, emotional lability including pathological laughter and crying (emotional incontinence), and occasionally schizophreniform psychotic symptoms.

PATHOLOGY AND LABORATORY EXAMINATION. Neuropathologically, there is broader involvement of the basal ganglia in PSP than in PD. PSP affects the substantia nigra as well as the striatum, thalamus, subthalamus, other brainstem nuclei, and periaqueductal gray matter. The main pathological findings are high-density tau positive neurofibrillary tangles, neuronal loss and gliosis. There is pronounced dopamine depletion in the striatum, decreased density of striatal D2 receptors, widespread cholinergic neuronal damage, and effects on serotonergic and GABAergic systems. Structural imaging studies initially reveal midbrain atrophy followed by third ventricular enlargement and pontine and frontotemporal atrophy. Diffusion tensor imaging studies demonstrate white matter tract degeneration connecting the cerebellum, midbrain, thalamus, and premotor cortex. PET studies show decreased striatal dopamine binding and fluorodopa uptake and bilateral premotor hypometabolism extending into prefrontal cortex, supplementary motor area, and midbrain. A recent genome-wide association study of autopsy-confirmed PSP implicates three genes, STC6, EIF2AK3, and MOBP, as associated with PSP. However, PSP is rarely associated with a family history and is considered sporadic at this point.

DIFFERENTIAL DIAGNOSIS. After PD, which is at least 10 times more

prevalent, PSP is the second most common parkinsonian syndrome. Frontotemporal dementia (FTD) and corticobasal ganglionic degeneration manifest similar clinical and pathological features, including parkinsonism, frontal executive impairments, apathy, and tau positive inclusions. While profound apathy is the predominant psychiatric feature, it can be misdiagnosed as depression, which occurs with apathy in a minority of cases.

COURSE AND PROGNOSIS. PSP is a rapidly progressive and devastating degenerative disease. Life expectancy is generally 5 to 7 years after onset, with death often related to dysphagia.

TREATMENT. There are no specific treatments for the motor or cognitive aspects of PSP. Two recent clinical trials targeting tau represent significant progress. Both tideglusib, a glycogen synthase kinase 3 inhibitor, and davunetide, which affects microtubule stability, were safe and well-tolerated, but did not demonstrate clinical efficacy on the PSP rating scale. Levodopa is ineffective in PSP, probably because of the reduced number of postsynaptic striatal D2 receptors, although there can be an initial response. Despite extensive cholinergic damage, cholinergic enhancing agents provide little or no benefit for cognition and can be poorly tolerated. Psychiatric symptoms in PSP can be difficult to treat, but include standard targeted pharmacological and psychosocial approaches aimed at reducing apathy, irritability, agitation, mood lability, etc., in neuropsychiatric populations.

Corticobasal Degeneration

Originally termed corticodentatonigral degeneration with neuronal achromasia, corticobasal degeneration (CBD) was first described in 1968 as a syndrome of involuntary movements, and slowed voluntary movements. CBD is classified as a sporadic tauopathy characterized by deposition of hyperphosphorylated 4-repeat tau in neurons and glia, referred to as astrocytic plaques. CBD also shares clinical and pathological features with other tauopathies, including PSP and certain forms of FTD. Although numerous clinical criteria have been applied to CBD, these predict pathological CBD in 25 to 56 percent of cases and diagnostic criteria are evolving. Review of pathologically confirmed cases of CBD suggest four CBD phenotypes: the initially reported phenotype, now referred to as CBS, a frontal behavioral–spatial syndrome (FBS), nonfluent/agrammatic variant of primary progressive aphasia (naPPA), and progressive supranuclear palsy syndrome (PSPS). To facilitate development of potential neuroprotective therapies for tauopathies, two sets of clinical diagnostic criteria comprising these phenotypes for CBD have been proposed, one for probable CBD and a second broader set of criteria for possible CBD with

less specificity for CBD pathology, but still likely representing tau pathology.

CBD affects adults, with gradual onset usually after age 40 and steady progression until death within 2 to 12 years. Median survival is around 8 years. The pathological findings include ballooned achromatic neurons, frontoparietal atrophy, pronounced cortical neuronal loss, cortical astrocytic plaques, and tau inclusions in the substantia nigra, basal ganglia, and along dentatorubrothalamic tracts. Neuroimaging studies show asymmetric frontoparietal atrophy and subcortical hypometabolism contralateral to the more affected side along with bifrontal–parietal hypometabolism.

The clinical manifestations reflect both basal ganglia and cortical dysfunction. Motor signs almost invariably include asymmetric progressive motor signs of levodopa-resistant parkinsonism (rigidity, brady/akinesia, and postural or action tremor). There is also postural instability, focal reflex myoclonus, and limb dystonia. Oculomotor motility may be impaired, similar to PSP. The most prominent neuropsychiatric features are lateralized ideomotor and ideational apraxia and the "alien hand syndrome." Cognitive deficits, originally regarded as late-stage phenomena, are evident in most patients early in the disease course, along with behavioral features. These may be overlooked, however, when there is prominent motor dysfunction. As with other subcortical motor disorders, there are multiple cognitive deficits, including predominant executive, memory, and attentional dysfunction that may be severe enough to amount to dementia. Parietal lobe dysfunction, manifest as sensory or visual neglect and visuospatial deficits, is apparent on cognitive testing, as are deficits in social cognition. Aphasia may also be present. Depression is much more common in CBD relative to PSP, which is characterized by apathy. In CBD, apathy, disinhibition, and delusions are also common and may co-occur, but appear distinct from depression, suggesting different underlying mechanisms. There are no definitive treatments for the motor or cognitive symptoms in CBD, but the parkinsonism may have a limited response to levodopa. Antidepressants are used for depression and anxiety, but are not always effective. Caregiver education and assessment for depression are important aspects of patient care.

PRIMARY DEMENTIA SYNDROMES WITH PARKINSONIAN SIGNS

Frontotemporal Dementia

Social cognition is one of the six cognitive domains assessed in the diagnosis of major neurocognitive disorder in the DSM-5. FTD is a clinical syndrome characterized by prominent behavioral or personality changes involving social cognition, language dysfunction, and neuronal loss, spongiosis and gliosis in the frontal and temporal cortices. Onset of FTD is

usually after age 50, but ranges from 35 to 75 years. Life expectancy after diagnosis is 3 to 15 years.

FTD is neuropathologically, clinically, and genetically heterogenous, with approximately 30 to 50 percent of cases having a positive family history. FTD also overlaps clinically and neuropathologically with CBD, PSP, and motor neuron disease/amyotrophic lateral sclerosis (MND/ALS), which are often considered part of the FTD complex. The characterization and genetics of the FTDs have evolved rapidly over the last decade. Neuropathologic consensus criteria include 3 major subgroups based on protein pathology: tau, transactive response DNA-binding protein 43 (TDP-43), and fused in sarcoma (FUS) and there are three major FTD disease genes [microtubule-associated protein tau (MAPT), progranulin (GRN), and C9*orf*72] and many other rarer disease genes. Mutations in MAPT, located on chromosome 17, are typically associated with bvFTD and parkinsonism, and are generally referred to as FTD and parkinsonism linked to chromosome 17 (FTDP-17).

Symptom onset in FTD is insidious and many patients with FTD are initially misdiagnosed with a primary psychiatric illness. The behavioral variant (bvFTD) involves prominent behavioral and personality changes such as apathy, impulsivity, disinhibition (e.g., shoplifting or grabbing food off another person's plate), loss of personal awareness (e.g., neglect of personal grooming), stereotyped and perseverative behaviors, delusions, depression, affective lability, overeating and other signs of hyperorality, social inappropriateness, and lack of tact or empathy. A second variant presents with a progressive nonfluent aphasia (PNFA) and a third variant with a semantic dementia (SD) with receptive language impairments. Cognitive testing reveals profound impairment on tests of executive function, but relative preservation of other cognitive domains, including memory.

Motor deficits may include symptoms of MND or parkinsonism of the akinetic-rigid type that is common in bvFTD or PNFA and rarely seen in SD. Structural imaging studies may show frontal and anterior temporal atrophy; functional imaging typically reveals hypoperfusion in the same regions, though there may be pronounced asymmetry. More widespread involvement of parietal regions is suggestive of Alzheimer disease. The electroencephalogram may be normal in early stages. The differential diagnosis in younger persons includes schizophrenia, mood disorders, addictions, and personality disorders. When onset is later, the diagnosis is often confused with Alzheimer disease, late life depression, and, when there are prominent motor features, CBD and PSP. These conditions resemble one another in their advanced stages, but are distinguished by the course of the disease prior to the advanced stage. There are no treatments for the underlying disease process, and no definitive evidence supporting the use of cholinergic treatments. The behavioral symptoms may improve

with serotonin reuptake inhibitors. Aggressive behaviors can be managed with low doses of atypical antipsychotics or trazodone and antiepileptic medicines can reduce agitation and aggression. Family counseling is a critical part of treatment so that the safety of the patient and others is maintained in the setting of their poor judgment and to provide caregiver support and assist with psychosocial issues.

Dementia with Lewy Bodies

DLB is a dementia syndrome with Lewy bodies in the cerebral cortex, brainstem, thalamus, and striatum. It is now regarded as the second most common form of dementia after Alzheimer disease. Symptom onset is usually after age 50 and the disease is more common in men. Average survival is less than 10 years. There are both genetic and sporadic forms of DLB, and there is overlap in the clinical and pathological findings of DLB, Alzheimer disease, and PD and other α-synucleinopathies. About half of the patients with DLB also have Alzheimer pathology. Up to one-third of cases show microvascular ischemic pathology. Neurochemically, there is evidence of a marked cholinergic deficit.

While the defining feature of DLB is progressive cognitive impairment, its association with extrapyramidal signs and hallucinations leads to its inclusion in the differential diagnosis of parkinsonian movement disorders with psychiatric presentations. In the earlier stages, cognitive deficits often involve attentional, visuospatial, or frontal–subcortical skills, and memory impairment may only become apparent with progression. Cognition is often described as fluctuating with variations in attention and alertness, but a reliable definition of fluctuating cognition has been difficult to establish. With progression, there are the usual cortical features of dementia (aphasia, apraxia, and agnosia). The psychiatric symptoms present early in the course and often include dramatic visual hallucinations, depression, or paranoid delusions. Motor features include myoclonus and parkinsonian signs of bradykinesia, rigidity, and gait difficulties with early falls in about two-thirds of patients. Resting tremor is uncommon. There may be an early but limited response to levodopa. As a general rule, the occurrence of psychosis unrelated to medications and early dementia with spontaneous parkinsonism, especially within the first year, should lead to consideration of DLB.

The diagnosis of DLB remains clinically based and neuroimaging studies are not especially helpful in distinguishing DLB from AD. Treatment of DLB requires a targeted approach. Motor symptoms can be treated with antiparkinsonian therapies, though this may aggravate psychiatric symptoms, especially psychosis. Atypical neuroleptics can reduce psychosis, but are often not tolerated and cases of neuroleptic malignant syndrome are reported. The most promising treatments include cholinergic therapies, which can benefit cognitive and psychiatric

symptoms. Antidepressants and stimulants can also be helpful.

Multiple System Atrophy

The term MSA was introduced in 1969 and operationally defined in 1999 to describe a clinically heterogeneous neurodegenerative disorder with variably present parkinsonian, autonomic, cerebellar, and pyramidal manifestations. MSA is now used in place of three older diagnoses, olivopontocerebellar atrophy (OPCA), Shy–Drager syndrome, and striatonigral degeneration, which could not be reliably distinguished from each other. Two types of MSA are defined: autonomic dysfunction with predominant cerebellar findings (MSA-cerebellar, or MSA-C) and autonomic dysfunction with predominant parkinsonian findings (MSA-P). The disease is rare and affects both sexes equally. It typically begins in late middle-age, with a median survival of about 9 to 10 years after symptom onset. The cause is unknown, but almost all cases are sporadic. The neuropathology consists of glial cytoplasmic inclusions (containing tau, synuclein, and ubiquitin, among other proteins) in the presence of neurodegeneration in a variable combination of CNS structures, potentially including the putamen, caudate, globus pallidus, substantia nigra, inferior olives, brain stem nuclei, cerebellum, and spinal cord. MRI may reveal typical signs (hyperintense portions of the putamen, “hot cross bun” sign in the basis pontis), but diagnosis is essentially made clinically or neuropathologically.

A prodromal premotor phase is evident in up to 75 percent of cases and, consistent with multisystem pathology, includes sexual and urinary dysfunction, orthostatic hypotension, respiratory disturbances, and rapid-eye-movement sleep behavior disorder. Psychiatric and cognitive manifestations of MSA have received little systematic attention. Dementia and visual hallucinations have been considered as nonsupportive criteria for diagnosis, and suggestive of DLB. However, more recent data from patients surviving longer than 8 years shows evidence for significant cognitive impairment in up to 50 percent of cases, with a profile of cognitive deficits similar to other parkinsonian disorders and dementia in up to one-third of cases. Executive functions are most commonly seen, more so in MSA-P, whereas learning and visuospatial impairments are more prominent in MSA-C. Emotional incontinence, depression, anxiety, panic attacks, and suicidal ideation are also reported. Treatment is symptomatic. A transient response to levodopa is seen in up to 40 percent of patients in the early disease stage. Orthostatic hypotension and urinary symptoms may respond to both pharmacological agents and environmental manipulations, and such treatment may substantially improve patient and caregiver quality of life. There are no clear guidelines for treating psychiatric disorders in MSA; it is likely that strategies useful for PD, HD, or cerebellar disease will also prove of value in MSA. A small cases series

demonstrated the usefulness and safety of ECT in patients with MSA.

Encephalitis Lethargica

Encephalitis lethargica, described by Constantin von Economo, presented as a worldwide epidemic between 1917 and 1929, but isolated cases have been reported in the last decade. It is important historically as an early example of the relationship between brain disease and psychopathology. The acute phase was characterized by profound lethargy, hence its name, along with other mental status changes such as psychosis or catatonia and various movement abnormalities, including parkinsonism, choreoathetosis, and myoclonus. The mortality rate was approximately 30 percent during the acute phase. Survivors returned to normal for several months to years, then many developed a chronic post-encephalitic syndrome characterized most commonly by parkinsonism (usually without tremor), and less commonly by hyperkinetic movements resembling tardive dyskinesia and episodic oculogyric and respiratory crises. Psychiatric manifestations included disinhibition, mania, attentional deficits, apathy, narcolepsy, verbal tic-like behaviors, executive dysfunction, and bradyphrenia. Hyperactivity and oppositional behaviors developed in children. Postmortem studies of acute victims showed inflammatory changes in the substantia nigra, globus pallidus, and subthalamic, red, and brainstem nuclei. Postencephalitic cases showed degenerative changes and depigmentation of the same subcortical areas, with neurofibrillary tangles but without Lewy bodies. The etiology has never been determined.

Huntington Disease

Introduction and Definition. HD (formerly Huntington chorea) is an autosomal dominant neurodegenerative disorder characterized by mid-life onset, a relentlessly progressive course, and a combination of motor, psychiatric, and cognitive symptoms. The disease is caused by a CAG repeat expansion mutation in the huntingtin gene on chromosome 4. For 130 years, HD has been recognized as one of the most terrible yet enlightening of all medical disorders, and now, as in the past, the approach to HD provides a model for the development of clinical, scientific, and ethical strategies in other genetic diseases of the brain. HD remains an important example of subcortical dementia and the prominence of the psychiatric consequences of neurodegenerative disorders.

History. George Huntington first described the disease that now bears his name in 1872. His original paper, which noted the salient features that still define HD, remains a model of clarity and clinical observation. After serving as the textbook case of dominant inheritance for 100 years, HD again made news in 1983 as the first disorder mapped to a specific

chromosomal locus using anonymous DNA markers, and then again in 1993 when the genetic etiology was discovered to be a member of a new class of mutations referred to as repeat expansion or dynamic mutations. Current research focuses on understanding HD pathogenesis and the use of this knowledge to develop rationale therapeutics for the underlying disease process.

Epidemiology. The prevalence of HD in the United States is about 5 cases per 100,000 individuals, with similar rates found in most European countries. The prevalence in Japan is about 10-fold less, and prevalence also appears to be low in China, Finland, and Africa. As with many rare autosomal disorders, pockets of high prevalence exist. The most famous concentration is in the villages surrounding Lake Maracaibo in Venezuela; analysis of this population led to the discovery of the HD mutation and has contributed greatly to the clinical understanding of HD.

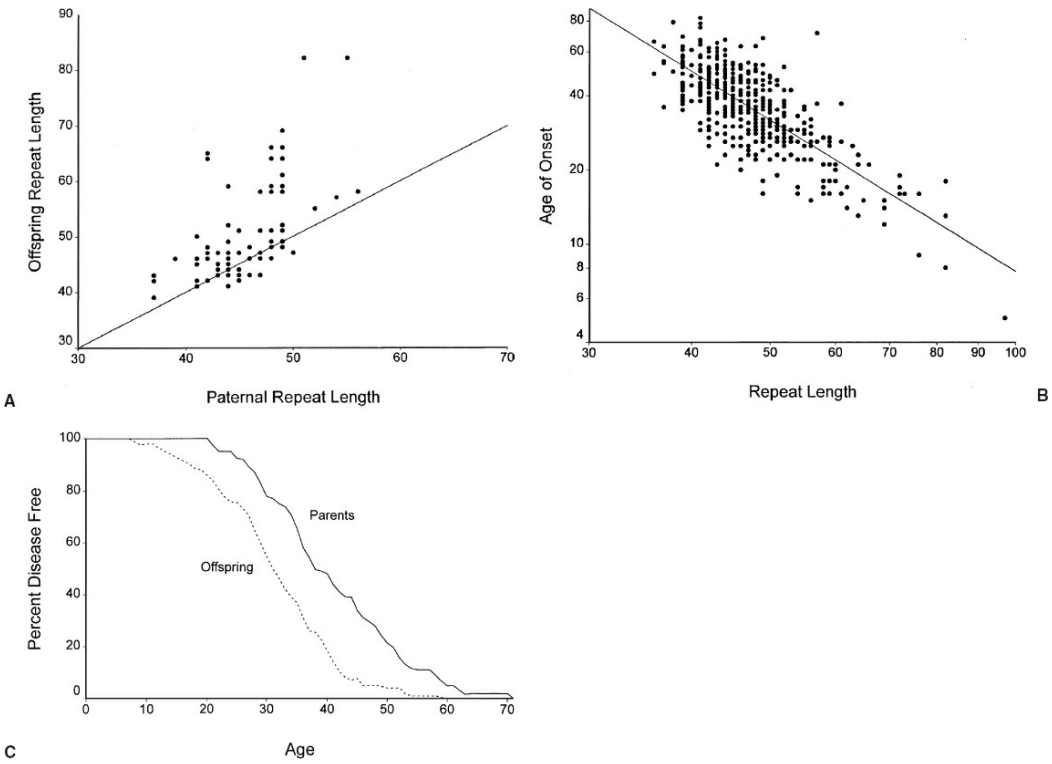

FIGURE 2.6–3. Repeat length as an important variable in Huntington disease. **A:** Repeat length increases in paternal transmission of the expanded allele. Points above the diagonal line represent cases in which the repeat length increases during paternal transmission. Expansion occurs in most cases, and in some instances the expansion is substantial. N = 84 pairs, mean +/– SD increase of repeat length = 4.2 +/– 0.8 triplets. **B:** Correlation of repeat length with a younger age of onset. As repeat length increases, age at disease onset decreases. N = 480, r^2 = 0.57. **C:** Anticipation in HD. The age at which affected parents and their affected offspring first develop disease symptoms is depicted as a survival curve. The younger generation develops the disease at a substantially younger age, a result of increasing repeat length during paternal transmission and the correlation of repeat length with younger age of onset. N = 61 parents, 82 offspring. (All data from the Baltimore Huntington Disease Center, courtesy of Drs. Adam Rosenblatt and Christopher A. Ross.)

Etiology. HD is one of a group of at least 11 neurodegenerative diseases caused by a mutation in which repeating units of DNA consisting of the nucleotides cytosine, adenine, and guanine (a so-called CAG repeat; cytosine, thymine, guanine—CTG—on the opposite strand) increase in number. This is referred to as a CAG, or CAG/CTG, trinucleotide repeat expansion mutation. In nine of these disorders, including HD, the CAG repeat is in-frame to encode for the amino acid glutamine. Genes containing this type of expansion therefore express an RNA transcript with an expanded CAG repeat, which in turn is translated into a protein with an excessively long stretch of glutamine residues. The CAG repeat expansion occurs in different and unrelated genes in each of the CAG repeat diseases. The only similarity among the genes is the repeat itself. This and other lines of evidence have led most investigators to conclude that proteins with stretches of more than about 35 to 40 consecutive glutamine residues become neurotoxins, though there is increasing recognition that the RNA transcript may also have toxic properties. The differences among the CAG repeat expansion diseases, particularly concerning which brain regions are most prominently affected and the precise threshold for toxicity, are thought to derive from other aspects of protein expression and structure.

The CAG repeat expansion in HD is located in exon 1 of the *huntingtin* gene (formerly referred to as *IT15*) on the short arm of chromosome 4. Like the genes associated with other polyglutamine disorders, huntingtin is widely expressed, and is not limited to the CNS. The normal function of huntingtin is only partially understood, though it has been associated with embryogenesis, intracellular organelle transport, and regulation of growth factors. Like many of the other repeat expansion diseases, the phenomenon of anticipation (younger age of onset in subsequent generations of a family) occurs in HD. The explanation derives from a strong correlation between repeat length and younger age of onset and a tendency for repeat length to increase during paternal transmission of the gene (Fig. 2.6–3A,B). There is also increasing recognition that understanding disease pathogenesis requires understanding of the molecular complexity of CAG/CTG repeat regions of the human genome, including the HD locus, in which an RNA transcript is encoded by the DNA strand antisense to huntingtin. The antisense transcript appears to regulate huntingtin expression, and could also independently contribute to disease pathogenesis.

The investigation into the pathway that leads from the CAG repeat expansion to clinical phenotype in HD and related diseases had led to dozens of candidate pathogenic pathways. Early pathogenic steps appear to include abnormal phosphorylation and cleavage of the huntingtin protein into small fragments by cellular enzymes, and subsequent misfolding of fragments with long glutamine tracts. These fragments eventually aggregate into structures that can be visualized with immunohistochemical stains in postmortem brain from patients with HD (Fig. 2.6–4A). Many

mechanisms have been invoked to explain subsequent steps in pathogenesis. For instance, proteins with long glutamine stretches sequester other proteins, including transcription factors. These proteins become unavailable to perform their normal cellular duties, with potentially lethal consequences. Alternatively, proteins with long polyglutamine expansions, or aggregates of these proteins, may interfere with the process of autophagy, through which cells degrade abnormal protein and recycle organelles. Failure of this system might lead to an accumulation of a variety of toxic proteins. Other proposed pathogenic mechanisms include interference with mitochondrial energy metabolism, polyglutamine-induced vulnerability to excitotoxicity and oxidative stress, and decreased brain-derived neurotrophic factor (BDNF) release.

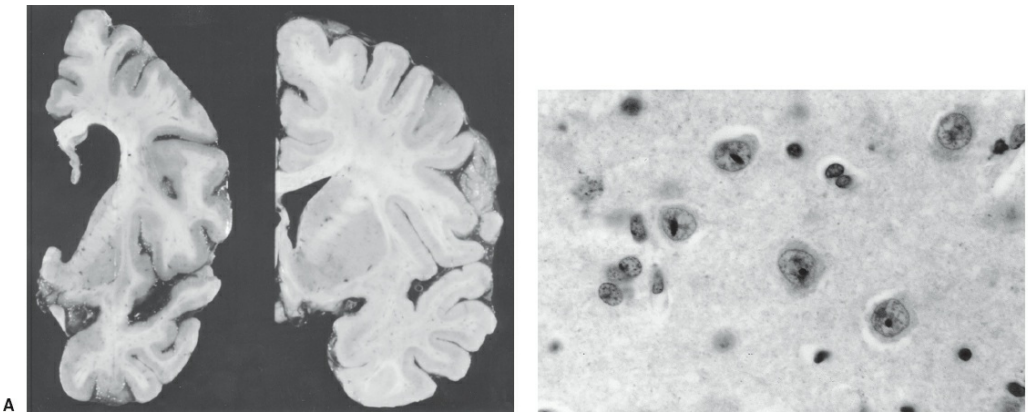

FIGURE 2.6–4. Huntington disease neuropathology. **A:** Gross pathology. HD case on the right, control brain on the left. Note the enlarged lateral ventricle, nearly absent caudate and putamen, and atrophy of the cerebral cortex. **B:** Intranuclear inclusions in HD. Immunohistochemical staining of cerebral cortex from an HD patient with an antibody that labels the huntingtin protein demonstrates prominent intranuclear inclusions in neurons. (Photographs courtesy of Dr. Christopher A. Ross.)

Diagnosis and Clinical Features. Typically, the diagnosis of HD is made by a genetic test in the setting of motor signs and symptoms in an individual with a family history of HD. The advent of genetic testing now makes diagnosis possible in individuals with a family history of HD who do not yet have symptoms themselves, or in patients with unusual presentations or HD-like symptoms in the absence of a known family history of a movement disorder.

MOTOR MANIFESTATIONS. The primary *involuntary* movement abnormality is chorea or choreoathetosis, continuous and irregular writhing and jerking movements. The limbs and trunk are most prominently affected, but respiratory, laryngeal, pharyngeal, oral, and nasal musculature may also be involved. Chorea may be among the early symptoms, but often plateaus and may even diminish during the course of the disease. Voluntary movement abnormalities typically affect visual tracking (slow hypometric

saccades and catchy pursuit), fine motor movements (slow, poorly coordinated, arrhythmic), speech and swallowing, tone (rigidity), and gait. The voluntary motor problems are most associated with functional disability and are most likely to lead to serious illness or death. Unlike chorea, voluntary movement abnormalities progressively worsen.

COGNITIVE MANIFESTATIONS. Cognitive decline generally parallels the development of movement abnormalities. Carefully controlled studies indicate that detectable cognitive abnormalities likely emerge at least 10 years prior to clinical onset of disease, though these will progress slowly and may be difficult to detect in an individual patient. Early cognitive changes, which may be subtle, often affect cognitive flexibility and speed, such that switching from one task to another, or tracking multiple tasks simultaneously, becomes difficult or slow. Work or school performance may suffer to the point that colleagues notice. A four-domain model of cognitive deficits suggests that in early to middle stage HD, visuospatial deficits are the most pronounced deficits, followed by executive dysfunction, and then less prominent memory and verbal deficits. The dementia becomes more global as the disease progresses, but aphasia and agnosia are less evident than in Alzheimer disease, as in other subcortical dementias. MMSE scores typically decline by 1 to 2 points/year.

PSYCHIATRIC MANIFESTATIONS. As many as 80 percent of patients with HD develop some form of noncognitive psychiatric disorder within 10 to 15 years of disease onset. The consequences of these disorders may be devastating to the patient and family, perhaps more so than either the cognitive or motor dysfunction. The suicide rate is higher than in the general population, with an average across studies of 4.6 percent of all deaths. Nonetheless, since it is the psychiatric aspects of HD that are most amenable to treatment, at present, psychopathology should be sought for diligently. Thirty to forty percent of patients with HD will develop major depression, and another 10 to 20 percent will have a nonmajor form of depressive syndromes. Up to 10 percent of patients with HD develop mania, though this diagnosis may be confused with disinhibition and other personality changes. Less frequently, patients with HD develop disorders that are clinically indistinguishable from idiopathic obsessive–compulsive disorder, schizophrenia, or delusional disorder. As in patients with other types of brain damage, patients with HD are vulnerable to delirium from pharmacological or medical illness.

Personality change is very common, occurring in up to 50 percent of patients with HD by mid-course. Most patients with HD with personality change also have cognitive impairment or dementia, but cognitive deficits and personality change do not always occur together. Other patients may develop a naïve, superficial, child-like personality that does not fit into

current DSM terminology. Apathy and irritability are particularly common and troublesome states, may occur prior to overt clinical disease, and may exist as either as separate syndromes or as symptoms of other disorders.

The many psychosocial issues that accompany HD increase the chance for adjustment disorders, relationship disorders, and other behavioral disorders. Caring for a cognitively impaired, irritable, or apathetic family member may result in emotional turmoil and financial hardship. Frequent additional stressors include social stigma (real or feared), parental guilt over passing the disease gene on to children, a family legacy of secrecy, and the burden faced by young gene positive adults who can predict their own future in light of the progressive decline of a parent or other relative. These issues and their consequences must be recognized and treated, but should not be used to "explain away" other psychiatric disorders for which other treatments may be more appropriate.

Pathology and Laboratory Examination

Pathological Exam. Gross pathology of HD (Fig. 2.6–4B) is limited to the brain, though microscopic or metabolic involvement has been detected in muscle, liver, adipocytes, and lymphocytes. In advanced cases, brain weight may be reduced by 25 to 30 percent. The most prominent atrophy involves the caudate and putamen, which is readily apparent on MRI scans or on gross pathology, along with accompanying hydrocephalus ex vacuo. The cerebral cortex becomes quite atrophied, with less pronounced atrophy in other regions of the basal ganglia (especially the globus pallidus and STN) and to a lesser and more variable extent in other brain regions.

On the microscopic level, selective neuronal vulnerability is striking. Striatal cell loss primarily affects medium spiny projection neurons, and in a gradient such that the earliest and most severe loss is in the dorsomedial region. Cortical cell loss consists mostly of large neurons in layers III, IV, and V. Surviving neurons in the striatum and cortex may have an increased number of dendrites and dendritic spines, but some cells will have evidence of dystrophic processes. Gliosis is present in regions with significant cell loss. Intranuclear inclusions are evident in striatum and other brain regions (Fig. 2.6–4B). The inclusions consist of amyloid-like fibrils that contain, among other proteins, mutant huntingtin, ubiquitin, and synuclein, but not tau. Inclusions may also be detected in neuronal cytoplasm and processes.

Laboratory Evaluation. It has been possible to directly test for the HD repeat expansion for the past two decades, using a simple PCR-based assay on DNA obtained from a blood sample. The test is available from a number of commercial and academic centers. Interpretation of results is shown in Table 2.6–3. While the test is generally reliable, ambiguity may arise if a repeat length is on the borderline of one of the categories, since the results

of even the best laboratories have an error of +/− 1 triplet. Also, occasionally only a single allele of normal length is detected. This usually means that both alleles contain a repeat of the same length, but the possibility exists that an expanded allele went undetected, which requires additional assays.

Ancillary Tests. MRI, CT, SPECT, and PET were previously of diagnostic importance, but are now mainly used to exclude strokes, tumors, and subdural hematomas, etc., that may coexist with HD. MRI scans should be considered if the patient's course rapidly deteriorates or in the presence of signs and symptoms not typical for HD, as falls are common in HD, and the risk of a complicating but treatable brain injury is relatively high.

Table 2.6–3.
Interpretation of HD Repeat Length (Longest of the Two Alleles)

Repeat length (CAG Triplets)	Interpretation	Comments
<29	Not at risk	Repeat length stable
29–35	Not at risk	Reports of a few HD cases
36–40	May develop HD	Increasing risk with longer repeats
41 and greater	Will develop HD	Earlier age of onset with longer expansions

Premanifest Testing. The availability of a diagnostic test for HD has made testing of individuals who are at risk but do not yet have clinical disease testing possible. So far, only a minority of individuals at risk of HD request testing, often for psychological and practical reasons related to child-bearing, long-term planning about caring for a relative with HD, and finances. Perhaps because of the self-selected nature of those seeking premanifest testing and the exclusion from testing of those who are suffering from major depression, most patients feel relieved to learn of their status, whether testing positive or negative. Careful protocols for pre- and post-test counseling have been developed to assure that patients understand the potential psychological risks and risks of discrimination (employment and insurance) that may arise from the test results. Expert counseling is mandatory before premanifest testing is obtained. For known gene carriers, it is possible to obtain prenatal diagnosis early in pregnancy. A few centers offer preimplantation testing of embryos generated by in vitro fertilization. Premanifest testing is likely to increase in the future, as the availability of clinical trials for premanifest individuals will increase incentive for learning about gene status.

Differential Diagnosis. The availability of a definitive diagnostic tests leaves two issues in HD differential diagnosis: when to consider a diagnosis of HD and what diagnoses to consider if a patient with HD-like symptoms

tests negative for HD. HD may occasionally present with rigidity, dystonia, athetosis, or tic-like movements rather than chorea, sometimes in every member of a pedigree. At other times, patients present with typical clinical manifestations, but they have no known family history of HD. This may arise when disease onset occurs earlier in a child than their parent, a parent carrying the mutation has died from other causes prior to the onset of clinical HD, there is false paternity or adoption, or, in very rare cases, a new mutation arises that is not present in previous generations. Genetic testing should be considered in either situation. Table 2.6–4 lists some of the disorders that present with HD-like symptoms. Notably, any disease or insult that affects the basal ganglia might lead to chorea or dystonia.

Course and Prognosis. HD is relentlessly progressive and fatal (Fig. 2.6–3C). Onset is typically insidious, with early complaints of clumsiness, jerks, and tremor. While the chorea tends to plateau and may even decline later in the course, the voluntary motor abnormalities progress, paralleled by cognitive decline. Patients with late-stage HD require complete care for all activities of daily living. They are mostly nonverbal, rigid, and akinetic, but may make whole body movements resembling myoclonic jerks when disturbed. Dementia is global and severe, but may be less severe than the patient's appearance would suggest. Death usually occurs 15 to 20 years after disease onset, often from the consequences of swallowing difficulties (suffocation or aspiration pneumonia) or general debilitation. Juvenile-onset HD often presents with bradykinesia, rigidity, dystonia, and sometimes seizures or myoclonus. Chorea may be minimal or absent. The course seems to be more rapid and severe than typical HD. Large-scale longitudinal studies in Europe and the United States are attempting to define the course of HD in sufficient detail that the effectiveness of interventions can be precisely measured.

Table 2.6–4.
Differential Diagnosis of Choreiform Disorder

Genetic Causes of Chorea	Comment
Dentatorubralpallidoluysian atrophy (DRPLA)	Autosomal dominant, rare in United States, more common in Japan, variable presentation
Huntington disease-like 2	Rare autosomal dominant disorder clinically indistinguishable from HD
Autosomal dominant spinocerebellar ataxias	Most commonly in SCAs 2,3, but also in others
Hereditary Creutzfeldt-Jakob disease	Autosomal dominant; 85% of cases are sporadic
Wilson disease	Autosomal recessive, distinguished by eye and liver abnormalities, abnormal copper metabolism
Ceroid neuronal lipofuscinoses	Usually childhood onset autosomal recessive, rarely adult onset autosomal dominant

Neuroacanthocytosis	Recessive and dominant forms, distinguished by abnormal red blood cells (acanthocytes)
Benign hereditary chorea	Usually childhood onset
Pantothenate kinase-associated neurodegeneration	Childhood onset, recessive
Fahr disease	Basal ganglia calcification, usually autosomal dominant, adult onset
Mitochondrial diseases	Maternal inheritance, heterogenous phenotypes
Paroxysmal dyskinesias	Induced by external factors, multiple genetic subtypes, some sporadic forms
Lesch–Nyhan syndrome	Boys only, self-mutilation
Other metabolic disorders	Phenylketonuria, guaric acidemia type I, methylglutaconic aciduria type III, various aminoacidopathies
C9ORF72 disease	Reports of prominent chorea in some mutation carrier
Neuroferritinopathy	Middle age onset, associated with brain iron accumulation
Sporadic forms of chorea	Comment
Tardive dyskinesia	Common
Sydenham chorea	Poststreptococcal complication
Systemic lupus erythematosis	May occur in 1–7% of cases, chorea also seen in other autoimmune disorders
HIV	Direct encephalitis or secondary lesions
Thyroid disease	Typically hyperthyroidism
Neurosyphilis	May also complicate HIV
Paraneoplastic syndromes	Mechanism appears to be autoantibodies
Drug-induced	Anticonvulsants, lithium, stimulants

Treatment. There is currently no treatment of the underlying neurodegeneration of HD. Recent large-scale clinical trials of creatine and co-enzyme Q10 were both terminated early based on strong evidence that the drugs had no beneficial effect. Several other agents have showed sufficient promise, or at least evidence of safety, to continue clinical trials, but at best agents currently under investigation are likely to have only a modest effect on disease progression. Surgical treatments, either via DBS or fetal striatal transplantation, remain in early stages. Perhaps the most promising therapeutic approach is to suppress expression of the mutant huntingtin gene, using one of several methods that prevent gene transcription or translation. While many hurdles to this strategy must be overcome, there is considerable optimism in the field for its ultimate success. In the meantime, more sophisticated study design, and improved measures of treatment outcome including highly sensitive cognitive, motor, and neuroimaging techniques, will likely increase the sensitivity and throughput of clinical trials.

Despite the absence of treatments to modify the fundamental course of the disease, the quality of life for the patient and their family can be substantially improved by medical intervention. Some patients are concerned about the social implications of chorea and others may find chorea functionally impairing. Chorea may be effectively treated by tetrabenazine (now FDA approved for use in HD), which blocks dopamine

release via action at the vesicular monoamine transporter 2. Chorea may also respond to low doses (0.5 to 2.0 mg/day) of typical high potency antipsychotics such as haloperidol or fluphenazine, other neuroleptics, benzodiazepines, or other dopamine-depleting agents such as reserpine or amantadine. Care must be taken, since dopamine depleting agents increase the risk of depression, and other agents may worsen voluntary movement abnormalities or apathy, or cause akathisia or excess sedation. The simple mechanical strategy of weights strapped to the wrists can often prove of considerable value without causing side effects. Environmental interventions, such as removal of throw rugs, adding grab bars in bathrooms, and avoidance of stairs, are always essential to reduce falls.

Other common symptoms can also be at least partially ameliorated. Swallowing abnormalities can be assessed by pharyngeal esophogram and treated with thickened liquid and softened solid diet. Botulinum injections may be useful for dystonia. Myoclonus has been successfully treated with valproate and clonazepam and valproate may also be effective for seizures.

There are no established pharmacological treatments for cognition. In particular, there is little evidence for an effect of anticholinesterase inhibitors. Simple education interventions may prove helpful, as in other forms of dementia. Early in the disease, patients can be taught to write down notes and to limit the number of tasks undertaken at once. Patients, and especially family members and care providers, need to be told about the progressive cognitive decline, so that they can learn to phrase questions more simply, provide cues for recall, and keep cognitive demands on the patient at a level commensurate with cognitive capacity. The topic of guardianship or power of attorney for legal and medical matters should be raised before the patient is too incapacitated to participate in the discussion.

Depression may be treated with any of the standard antidepressants, with the caveat that as in any patient with brain damage, starting doses should be low and doses should be increased slowly, with carefully monitoring for delirium. SSRIs are often useful and, given the higher suicide risk in patients with HD, are notably less toxic in overdose than tricyclic antidepressants or MAOIs. SSRIs may occasionally cause akathisia, insomnia, or dyskinesias. Other antidepressants have been used successfully. ECT has been used quite successfully to treat life threatening or intractable depression, and other than an increased risk of delirium, is generally well-tolerated. Mild mood elevations may not warrant treatment, but mania, like depression, can be treated with standard pharmacologic agents; limited data and anecdotal accounts suggest that divalproex sodium or carbamazepine is better tolerated and more effective than lithium. Benzodiazepines or antipsychotics (the atypical agents may have fewer negative effects on movement than the typical agents) may be helpful adjuncts.

Aggression and irritability stems from multiple causes, including psychiatric syndromes, cognitive and personality changes, and environmental precipitants such as hunger, thirst, pain, frustration over an inability to communicate, boredom, or unexpected change in routine. Treatment of any underlying psychiatric disorder and careful attention to possible precipitants is the first approach. When necessary, anecdotal evidence supports the effectiveness, in some patients, of SSRIs, antipsychotics, anticonvulsants, β-blockers, and benzodiazepines. Apathy probably responds best to family education combined with a structured and stimulating environment. SSRIs, stimulants, and dopaminergic agents such as amantadine and bromocriptine may be of value, though the latter agents may increase chorea.

Wilson Disease

Introduction and Definition. WD, also named hepatolenticular degeneration, is an autosomal recessive disorder in which mutations in the copper-transporting gene ATP7B result in abnormal copper accumulation in the liver, brain, and other tissues, with consequent hepatic cirrhosis and neuronal degeneration. Psychiatric and movement abnormalities are common, and patients may present to a psychiatrist first. WD is treatable, yet fatal without treatment. Recognition of the characteristic signs, symptoms, and laboratory abnormalities is therefore essential.

History. The disease was named after Kinnear Wilson, who first recognized the syndrome, including a high rate of psychiatric abnormalities, in 1912. In 1948, Cummings recognized the role of copper in disease pathogenesis; Schenberg and Gitlin, in 1952, discovered the deficiency of ceruloplasmin in patients with WD (establishing a test that is still in use), and, in 1960, Bearn definitively established the recessive nature of the disorder. The molecular basis of the disease emerged first with linkage to the esterase D locus on chromosome 13 in 1985, and then, in 1993, with the discovery of causative mutations in ATP7B.

Epidemiology. WD is found worldwide, with a frequency estimated at about 1/30,000. Approximately 1 in 90 individuals are heterozygous carriers.

Etiology. Over two hundred mutations associated with WD have been described in ATP7B. A few mutations are common (H1069Q accounts for more than 40 percent of mutations in Northern European populations), while most are rare. The result is that most individuals with WD are compound heterozygotes, carrying a different mutation on each allele. The Wilson protein, an ATPase involved in copper transport, functions to maintain copper homeostasis by facilitating excretion of copper from

hepatocytes into bile. In patients with WD, this function fails, resulting in accumulation of copper in the liver, brain, and other tissues, eventually leading to hepatic failure and neurological and psychiatric signs and symptoms. In the brain, the basal ganglia appear to be particularly affected.

Clinical Features. About 40 percent of patients first present with liver abnormalities, 40 percent with neurological abnormalities, and 20 percent with psychiatric abnormalities. Liver dysfunction, the most common childhood presentation of WD, may be manifest as mild elevations of serum transaminases, chronic active hepatitis, cirrhosis, or acute and fulminant liver failure. Copper toxicity may also lead to cardiac arrhythmias, renal disease, rhabdomyolysis, joint disease, anemia, and thyroid and parathyroid dysfunction. Neurological signs and symptoms, which strongly correlate with brain and CSF copper levels, generally reflect the multiple brain regions involved, and include dysdiadokinesis, dysarthria, tremor, dystonia, rigidity, choreoathetosis, bradykinesia, masked facies, and micrographia. Seizures occur in 6 percent of patients.

Cognition in WD has not been well studied. A relatively mild degree of cognitive impairment is thought to be present in about 25 percent of patients, worsening with time and in conjunction with accumulating neurological symptoms.

Two-thirds of patients may have psychiatric symptoms by first presentation, and 10 to 20 percent of patients seek psychiatric treatment before the diagnosis of WD is established (a stage at which, tragically, the diagnosis is often missed). Eventually, psychiatric manifestations affect nearly all patients with WD and the psychiatric profile generally resembles HD. The most common symptoms are personality changes (often described as irritability, aggressiveness, and emotional lability) and depression (ranging from subsyndromal mood changes to major depression, at times with suicidality). It is likely that anxiety disorders, psychotic disorders (and individual psychotic symptoms), and catatonia are more common than in the general population. The severity of psychiatric symptom correlates more strongly with the extent of neurological findings than to the liver disease, providing evidence that the psychiatric symptoms reflect the brain damage rather a psychological reaction to the severity of illness.

Diagnosis. Since the hepatic, neurological, and psychiatric manifestations of copper toxicity are only partially reversible, prompt diagnosis of WD is essential. The key step is considering the diagnosis. Elevated liver transaminases are nonspecific, but may be the first diagnostic clue. Kayser–Fleischer rings, copper deposits in the limbus of the cornea, may be detected by inspection or with a slit lamp examination, but are often absent early in the disease. Serum ceruloplasmin is below normal in 95 percent of patients with WD, but is also low in newborns and

20 percent of heterozygous WD carriers, and may be misleadingly normal in the setting of infection or inflammation. The absence of serum ceruloplasmin oxidase activity is a useful test in the latter cases. Urinary copper concentration is generally elevated. A liver biopsy will reveal increased hepatic copper concentration in addition to evidence of hepatitis or cirrhosis. Genetic testing for WD mutations has been limited due to the many possible mutations in ATPB7, though increasingly efficient methods for detecting mutations should make this a critical laboratory test in the near future.

In patients with neurological or psychiatric disorders, and sometimes in patients without clinical evidence of brain disease, a head MRI will reveal generalized atrophy, and increased T2-weighted signal in the basal ganglia (most typically the putamen) secondary to copper deposition. Atrophy may also be detected in the cerebral cortex (especially the frontal lobe), brainstem, and cerebellum. PET, SPECT, and MR spectroscopy have been used to reveal functional abnormalities in these regions. The most prominent pathological changes involve cavitary degeneration of the basal ganglia, with neuronal loss, gliosis, and high levels of copper.

Differential Diagnosis. Elevated hepatic copper may also result from cholestatic liver disease (normal or elevated ceruloplasmin) or idiopathic childhood cirrhosis (elevated ceruloplasmin). Aceruloplasminemia may result in neurological symptoms but results in elevated liver iron rather than copper. Urine copper excretion may be elevated secondary to metal chelation therapy or in the presence of acute liver failure or cirrhosis with cholestasis or nephrotic syndrome.

Course, Prognosis, and Treatment. Onset of WD is usually in the second or third decade with mean age of onset at 17 years. Rare cases are reported in individuals over the age of 60. Without treatment, death may occur within 5 years, and sooner for patients with CNS disease. However, WD is one of the few neurodegenerative diseases in which the fundamental pathogenic process, copper accumulation, is amenable to treatment. If caught early enough, many of the manifestations of the disease, including the neurological and psychiatric consequences, can be at least partially reversed. Treatment strategies involve either removing copper (chelation) by agents such as D-penicillamine, trientine, and tetrathiomolybdate or blocking copper absorption by zinc salts. Except for patients in areas with high levels of copper in the drinking water, it appears that dietary control of copper is not helpful. D-penicillamine, which is poorly tolerated by some patients, is prone to induce hypersensitivity reactions, and trientine has been noted to occasionally increase neurological and psychiatric symptoms acutely, as hepatic copper is mobilized and deposited in the brain. For patients with end-stage liver disease, orthotopic liver transplantation has

proven successful. All treatment modalities appear to improve neurological and psychiatric symptoms, but the response may only be partial if disease was at an advanced stage by the time therapy was instituted.

There are no controlled trials regarding the symptomatic treatment of the psychiatric manifestations of WD. Anecdotal evidence suggests that patients with WD may be particularly vulnerable to extrapyramidal side effects of antipsychotic medications. It seems prudent to treat psychiatric syndromes with standard psychiatric treatment modalities, both pharmacologic and psychotherapeutic, but to choose lower doses of medicines, increase doses slowly, and avoid high-potency neuroleptics.

CEREBELLAR DISORDERS

Introduction

The cerebellum has long been considered a modulator of motor function. Over the past 10 years, it has become increasingly clear that the cerebellum also modulates cognition and affect, and that damage to the cerebellum from many different causes can result in cognitive deficits and psychiatric syndromes.

History

A role for the cerebellum in regulating motor function was established early in the 19th century, beginning with the studies of Luigi Rolando. By the early 20th century, the effect of cerebellar degenerative disease on coordination of the limbs, gait, posture, and speech was appreciated. By 1939, the work of Holmes established much of the terminology and methods of examination of the cerebellum still in use. Simultaneously, diseases that affect the cerebellum were delineated, including hereditary disorders described by Friedreich in 1863 and Marie in 1893. Anecdotal reports of cognitive and psychiatric disorders arising from cerebellar damage have been noted since the early days of cerebellar research. However, systematic investigation of this link only began 20 years ago, stimulated by a series of neuroanatomical studies, still ongoing, that demonstrated connections between the cerebellum and multiple other brain regions, and by the frequent observations of odd behavior and mutism in patients following resection of posterior fossa tumors. There is now a growing consensus that the cerebellum plays an important role in normal cognition and emotion and that cerebellar damage leads to neuropsychiatric phenomena. In addition, substantial evidence now supports a role of the cerebellum in generating the psychopathology of such common psychiatric disorders as major depression, schizophrenia, and autism.

Epidemiology

The overall prevalence of cerebellar disorders is difficult to determine, since many different diseases may affect the cerebellum. As many as 50 percent of patients who have alcohol dependence syndromes develop clinical or histological evidence of cerebellar pathology; this may be the single most common cause of cerebellar damage worldwide. The cerebellum is affected in 1.5 to 8 percent of all strokes, so this is probably the second most common cause of cerebellar damage. In children, the annual incidence of tumors affecting the cerebellum is between 0.25 and 1/100,000/year. In adults, primary tumors of the cerebellum are more rare, but metastases are more common. The prevalence of recessive (primarily Friedreich ataxia), dominant, and sporadic neurodegenerative diseases that primarily affect the cerebellum is, in each case, about 2/100,000. Other causes of cerebellar lesions are rare.

Etiology

Table 2.6–5 lists many of the diseases that cause permanent injury to the cerebellum. With the exception of some strokes, tumors, and degenerative conditions, most of these disorders are unlikely to affect the cerebellum in isolation. Therefore, cognitive and psychiatric manifestations of these diseases may derive from damage to multiple brain regions. However, based on recently established neuroanatomical links between the cerebellum and other regions of the brain, and analysis of focal cerebellar lesions, it is now clear that damage to the cerebellum itself is sufficient to cause cognitive and emotional disturbance.

The pathway from the cerebral cortex to the cerebellum originates with projections from neurons in cortical layer Vb in motor, somatosensory, association, and paralimbic regions, including posterior parietal association cortex, portions of the superior temporal sulcus and gyrus, parastriate and parahippocampal regions, and dorsolateral, medial, and to a lesser extent, ventral lateral prefrontal regions. After synapse in the pons, where topographical distinctions among the different sources of input are maintained, the pathway continues as projections into the cerebellar cortex. Feedback to the cortex begins with axons projecting from cerebellar Purkinje cells to the deep cerebellar nuclei, where the pathway again remains topographically defined. These nuclei then project to both “motor” and “nonspecific” nuclei of the thalamus. Axons in neurons from both types of thalamic nuclei project to posterior parietal, superior temporal, and prefrontal association regions, and the intralaminar nuclei also project to limbic regions. This cerebrocerebellar pathway appears to consist of a series of discrete and relatively closed loops, similar to the loops connecting the basal ganglia and the cerebral cortex. Other pathways of potential relevance include reciprocal connections between the cerebellum and brain stem catecholaminergic (locus coeruleus) and serotoninergic nuclei (dorsal raphe), the hypothalamus, and the brainstem reticular formation. fMRI

studies now suggest that anterior cerebellar regions subserve sensorimotor processing, while more posterior regions subserve emotional and cognitive processing, with a specific functional association of different regions to relevant areas of cerebral cortex.

In addition to its connections to multiple other brain regions, the other aspect of cerebellar anatomy important to affect and cognition are the "cerebellar cortical modules." These modules exist as radially oriented, structurally identical units that comprise the cerebellar cortex. Each module consists of a Purkinje cell (numbering about 15 to 30 million in the human cerebellum), granule cells that receive input from outside of the cerebellum and innervate the Purkinje cell, and inhibitory interneurons (basket, stellate, and Golgi cells). The Purkinje cell thus serves to integrate excitatory and inhibitory input.

Table 2.6–5.
Diseases That Cause Cerebellar Damage

Hereditary Degenerative Autosomal Dominant
Spinocerebellar ataxia types 1–23
Dentatorubropallidoluysian atrophy
Huntington disease (uncommonly)
Hereditary prion diseases
Episodic ataxias types 1–7
Hereditary Autosomal Recessive (metabolic)
Abetalipoproteinemia
Niemann–Pick type C
Hexosaminidase A deficiency (Tay–Sachs)
Glutamic aciduria
Hydroxyglutaric aciduria
Cholestanolosis (cerebrotendinous xanthomatosis)
Metachromatic leukodystrophy
Galactocereberosidase deficiency (Krabbe disease)
Alpha-tocopherol deficiency (Vitamin E deficiency)
Hartnup disease
Refsum disease (phytanic acid oxidase deficiency)
Biotinidase deficiency
Hereditary Autosomal Recessive, DNA repair
Ataxia-telangiectasia
Xeroderma pigmentosum
Cocakyne syndrome
Hereditary X-linked
Pelizaeus–Merzbacher disease
Adrenoleukodystrophy
Rett syndrome
Mitochondrial
Kearns–Sayre syndrome
Myoclonic epilepsy associated with ragged-red fibers (MERRF)

Neuropathy, ataxia, and retinitis pigmentosa (NARP)
Leiber hereditary optic neuropathy (LHON)
Nutritional/Dietary
Vitamin E
Thiamine (often in association with alcoholism)
Celiac disease
Diseases That Cause Cerebellar Damage
Toxic
Ethanol
Phenytoin
Lithium
Organic solvents
Oncologic
Primary tumors
Metastases
Paraneoplastic (small cell carcinoma of lung, breast, female genital tract, lymphomas)
Infectious/postinfectious
Abscesses
Whipple disease
Subacute sclerosing panencephalitis
Progressive rubella panencephalitis
Infectious prion disease
Other
Multiple sclerosis
Sporadic late-onset cerebellar degeneration
Multisystem atrophy
Trauma
Radiation
Vascular
Embolic, hemorrhagic, or occlusive
Congenital (some Autosomal Recessive or X-linked hereditary)
Marinesco–Sjögren syndrome
Joubert syndrome
Disequilibrium syndrome
Early onset ataxia with hypogonadism, retinopathy, or deafness

It is generally believed that these units perform the same type of transform for all information coming into the cerebellum, with the specificity of information processing provided by the unique connections of each region of the cerebellum to other brain structures. Schmahmann has therefore proposed that just as the "cerebellum regulates the rate, force, rhythm, and accuracy of movements, so may it regulate the speed, capacity, consistency, and appropriateness of mental or cognitive processes." He has termed this the "dysmetria of thought" hypothesis and suggests that damage to the cerebellum will lead to an "unpredictability and illogic to social and societal interaction" just as it leads to poorly coordinated and misjudged movements, using the term cerebellar cognitive affective

syndrome to broadly define the clinical phenotype. He has suggested the cerebellar cognitive affective syndrome includes five axes: attention control, emotional control, autism spectrum disorders, psychotic disorders, and social skills. While intriguing, it remains uncertain to what extent this hypothesis can explain the multitude of cognitive and psychiatric phenomena that accompany cerebellar disease.

Diagnosis and Clinical Features

The diagnosis of cerebellar disorders involves narrowing the complicated differential: Is onset age at birth (congenital), childhood (metabolic, oncologic, Friedreich ataxia), middle-age (spinocerebellar ataxias), or elderly (oncologic, vascular)? What other organ systems or regions of the nervous system are affected? Is the process acute (vascular, toxic, infectious), subacute (infectious, oncologic,), chronic (hereditary, toxic, nutritional), or episodic? Is there progression or have symptoms remained static? Is there a family history of a similar disorder? Have there been exposures to drugs, alcohol, or toxins? The history and examination, followed by an MRI scan (and potentially other imaging modalities such as PET, SPECT, or angiography) and then specific genetic or metabolic tests will often, but by no means always, provide a final diagnosis.

Motor Features. Problems with balance and speech are often the presenting complaints in disorders with primarily cerebellar involvement. The "cerebellar gait" is wide based and staggering. Patients may complain of walking as if they were drunk, and stairs are particularly difficult. On examination, patients are unable to perform a tandem walk and may sway or display titubation, a low-frequency tremor of the body. The dysarthria is characterized by slurred speech of irregular rhythm and with incorrect emphasis of syllables. Speech may be slowed and monotonic (scanning speech) in compensation.

Cerebellar damage is also manifest in other ways. Eye movement abnormalities vary, and may include nystagmus, square wave jerks (inability to maintain fixation when gazing at a specified point), dysmetria (over- or undershoot when making rapid eye movements from one target to another), and uneven smooth pursuit. Limb incoordination may be manifest as trouble with fine motor tasks. Tremor may be either postural or kinetic; the later is often observable as an intention tremor during finger-to-nose testing. The presence of other signs or symptoms signals the possibility that other regions of the CNS are also involved, more the rule than the exception with diseases affecting the cerebellum.

Cognitive Features. The cerebellum is involved in a variety of cognitive tasks, including procedural learning, executive function, language processing, visual–spatial orientation, sensory processing, timing, and

attention. Within 10 to 15 years of disease onset, about 20 percent of patients with cerebellar degeneration will meet criteria for cognitive decline or dementia, and the rate presumably increases significantly later in the disease course. The pattern of deficits is consistent with a subcortical dementia, with deficits in verbal and nonverbal intelligence and especially executive function in the absence of aphasia, apraxia, or agnosia. Cognitive deficits in these patients, particularly when subtle, may not be detected on simple clinical mental state examinations that focus on language and memory. Executive functions impaired in cerebellar degeneration include attention, sequencing, and timing. Executive, visuospatial, verbal and visual memory, and language deficits also frequently follow focal lesions of the cerebellum. It has been suggested that cognitive deficits are more likely to follow lesions of the cerebellar hemisphere, while vermal lesions result in psychiatric disturbances.

Psychiatric Features. Many anecdotal reports, but only a few small systematic studies, have examined the frequency and nature of psychiatric syndromes in patients with cerebellar disease. Following strokes restricted to the posterior circulation (cerebellum and brain stem), about 35 percent of patient develop either a major or minor depression acutely, with the frequency declining to about 20 percent 6 months after the stroke and 10 percent 12 to 24 months later. Personality changes, particularly blunted affect and disinhibition, may be prominent in patients with large posterior circulation infarcts, and potentially other lesions, while anterior lobe or very small cerebellar infarcts elsewhere are less likely to produce personality change.

In degenerative disorders predominately limited to the cerebellum, about 75 percent of patients will develop one or more psychiatric syndromes within 10 to 15 years of disease onset. As many as one-third experience at least one episode of major depression and another third experience other forms of depression (including brief recurrent depression, minor depressive disorder, and dysthymia). About 25 percent develop persistent personality changes, similar in type to those found in HD, and, as in HD, not limited to patients with cognitive decline. The frequency of mania, anxiety syndromes, and psychotic syndromes may all be somewhat elevated. The neuropsychiatric findings in disorders in which cerebellar degeneration is accompanied by degeneration in other brain regions may be particularly severe. For instance, neuropsychiatric findings in the late onset autosomal dominant disorder spinocerebellar ataxia type 2 (loss of cerebellar Purkinje and granule cells, loss of inferior olive, pontocerebellar nuclei, substantia nigra, and motor neurons, and white matter tract degeneration) include executive dysfunction, impaired memory, apathy, depression, and potentially psychosis.

Pathology and Laboratory Examination. The pathology of cerebellar disorders depends on the underlying disease. Even with isolated cerebellar diseases, pathological examination is likely to reveal noncerebellar findings. The choice of laboratory evaluation, including genetic testing for hereditary disorders and metabolic screens for metabolic disorders, will be based on clinical and MRI findings. Neuropsychological testing may help detect subtle cognitive difficulties of considerable clinical relevance. In addition to the diagnostic workup, imaging studies are important if the nature or severity of signs and symptoms change unexpectedly, to exclude the presence of new subdural hematomas, infarcts, or tumors.

Differential Diagnosis. The differential diagnosis of the underlying disease for a patient presenting with cerebellar signs and symptoms is depicted in Table 2.6–5. Similar to other movement disorders, the differential diagnosis of affective symptoms is perhaps the most important issue facing the psychiatrist. Delirium must always be considered, regardless of whether the criteria for DSMV major depression are met. Isolated affective symptoms in the absence of delirium and major depression may reflect a subsyndromal depressive disorder, but could also be manifestations of a personality change or cognitive deficits either resulting in affective change (particularly apathy or irritability) or fueling demoralization.

Course and Prognosis. The course of cerebellar diseases varies widely by etiology, ranging from steadily progressive to intermittent to completely static. Personality and cognitive change are likely to follow the underlying disease course, while major depression may be more intermittent, suggesting differences in pathophysiology among these disorders.

Treatment. Determining the cause of the underlying disorder is essential, since cerebellar degeneration secondary to nutritional, dietary, and some metabolic disorders may be treatable and at least partially reversible. Cognitive treatments are supportive and educational, as described for HD. Little is known about the treatment of psychiatric disorders. Anecdotal accounts indicate that patients respond to standard treatments, including various classes of antidepressants and ECT. Treatable conditions, especially mood disorders, are likely under-recognized and undertreated. Caution is necessary with agents that may increase ataxia. Large-scale international registries of patients with specific disorders (e.g., European Friedreich ataxia database registry) may help elucidate disease course, establish biomarkers, and facilitate clinical trials of likely agents.

Essential Tremor

Essential tremor (ET) has been well characterized since 1887, but controversy remains over its definition, whether it represents a symptom, syndrome, or disease entity, and the basis for associated neuropsychiatric abnormalities. ET affects about 5 percent of individuals over age 65 and is the most common type of movement disorder. In 50 to 70 percent of patients, ET is a familial disorder, and genetic factors are important, but no causative gene mutations have been consistently reproduced. Growing evidence demonstrates the heterogeneity of ET with respect to its clinical presentation, age onset, progression, pharmacological response, genetics, pathology, and relationships with other diseases.

The main clinical sign in ET is the gradual evolution of a predominantly postural and kinetic tremor affecting the hands and arms, head, or both. It is either sporadic or genetic in origin and is not associated with other known causes of tremor. Activities of daily living can be significantly affected.

The etiology and pathophysiology of ET are unclear. Traditionally, ET has been regarded as a benign condition presenting as a monosymptomatic disorder without other neurological involvement or nonmotor effects. Whether ET represents a complex neurodegenerative disorder or a functional reversible disorder of neuronal oscillation is currently debated. Recent postmortem studies suggest that ET may be a cerebellar disease affecting Purkinje cells. This is supported by increasing evidence that ET is associated with development of other motor abnormalities, including comorbid diseases such as PD or dystonia and cognitive and psychiatric features. The kinetic component of ET, as well as subtle signs of ataxia, further suggests cerebellar involvement and neuroimaging studies suggest nigrostriatal damage. Other recent studies identified higher prevalence of MCI or selective cognitive deficits that suggest frontocerebellar dysfunction in association with ET, including executive dysfunction, and impairment in attention, verbal fluency, verbal learning, and verbal memory. Higher rates of depressive and anxiety symptoms are also reported in ET, but the epidemiology and features of discrete psychopathological syndromes have not been examined. No studies have addressed specific treatment of neuropsychiatric features in ET.

Dystonias

Dystonia is characterized by involuntary, sustained muscle contractions that result in twisting, torsional, repetitive movements of one or more body parts and frequently appear as abnormal and sometimes painful or awkward positions and postures. It is the third most common movement disorder after PD and ET. Like ET, it can present as a monosymptomatic condition or as a feature of broader syndrome. Any part of the body can be affected, including the limbs, trunk, neck, eyelids, face, or vocal cords. In 2013, an international consensus committee proposed a restructured

classification system for the dystonias based on two axes: clinical characteristics (age of onset, body distribution of symptoms, temporal pattern of symptoms, and associated features) and etiology based on shared biological mechanisms (e.g., neuropathology, inherited, acquired, idiopathic). Age onset is defined as childhood, adolescence, or adulthood (>28 years) and is an important predictor of whether symptoms will spread to other body parts. Distribution of dystonia is categorized as generalized, focal, segmental, multifocal, or hemidystonic and is one indicator of disease severity. The etiologic classification scheme is evolving, owing to ongoing discovery of different genetic forms of dystonia. Isolated dystonias are defined by the existence of dystonia as the major symptom and include both hereditary and sporadic forms. In combined syndromes, dystonia is accompanied by other prominent neurologic feature, such as parkinsonism or myoclonus, or other systemic features, including psychiatric symptoms. At least 10 genes or chromosomal locations are associated with isolated and combined dystonias, including Early-Onset Generalized Dystonia, Dopa-Responsive Dystonia, Paroxysmal Dystonia, X-Linked Dystonia-Parkinsonism (Lubag), Myoclonic Dystonia, and Rapid-Onset Dystonia-Parkinsonism. Acquired dystonias can be attributed to cerebral insults, including perinatal injury, trauma, toxins, or infarction or as a dopaminergic medication-related effect in PD or schizophrenia. Dystonia also occurs in neurodegenerative conditions, such as WD and PD. Neurophysiologic dysfunction in dopaminergic neurotransmission and basal ganglia output activity is thought to underlie dystonic disorders in which there is no evidence of nerve cell loss.

Dystonic movements begin with a particular motor action, but eventually generalize to multiple actions and are then present at rest. The contractions may be extreme, usually resulting in muscle hypertrophy and often pain or an oppositional tremor. Certain postures or sensory tricks can reduce dystonic contractions. Anticholinergic medications used to treat dystonia can affect memory, but cognitive impairment is not a general feature of dystonias.

Increased attention is being paid to the neuropsychiatric features of dystonia. Pain, sleep disturbances, psychiatric symptoms, fatigue, and cognitive impairment impact quality of life and aggravate dystonia-associated disability and motor symptoms. Compared to healthy controls, patients with focal dystonia have over twice the rate of psychiatric disorders, with increased lifetime and point prevalence for major depression and anxiety disorders, particularly social phobia (71 percent in a series of patients with cervical dystonia) and panic disorder. Often, the psychiatric disturbance begins many years before the onset of dystonia, indicating that the mood disturbance is not solely a reaction to the motor symptoms. Similarly, obsessive–compulsive disorder may be over-represented in families with myoclonic dystonia. A reactive element,

however, also exacerbates psychiatric distress, and can be the result of visible and painful motor symptoms. Alcohol may be misused to self-treat motor symptoms.

The prognosis of dystonia and rate of progression vary with the underlying condition and age of onset. Often, it stabilizes within 5 years of onset, but stressful circumstances provoke symptoms. Most therapy is symptomatic. Strategies include medications to reduce spasms and pain (e.g., anticholinergics, benzodiazepines, and muscle relaxants). Dopamine and dopamine blocking agents have either negative or beneficial effects on dystonias. Botulinum toxin injections into selected muscles are used extensively to treat the various forms of dystonia. Exercise programs are an important component of therapy. When medical approaches are inadequate, surgery includes thalamotomy, pallidotomy, and DBS procedures. There have been no studies for the treatment of psychiatric disorders associated with dystonia, but care must focus on the identifying and treating psychiatric disorders or psychological distress. For example, botulism injections may reduce torticollis and depressive symptoms, but self-perceptions remain impaired. SSRIs can precipitate, aggravate, or improve dystonia. ECT may benefit motor and mood symptoms. Initial studies indicate that DBS procedures for dystonia do not have adverse (or positive) effects on rates of cognitive or psychiatric disturbances.

References

Akbar U, Ashizawa T. Ataxia. *Neurol Clin*. 2015;33(1):225–248.

Armstrong MJ, Litvan I, Lang AE, et al. Criteria for the diagnosis of corticobasal degeneration. *Neurology*. 2013;80(5):496–503.

Bandmann O, Weiss KH, Kaler SG. Wilson's disease and other neurological copper disorders. *Lancet Neurol*. 2015;14(1):103–113.

Bruns MB, Josephs KA. Neuropsychiatry of corticobasal degeneration and progressive supranuclear palsy. *Int Rev Psychiatry*. 2013;25(2):197–209.

Burrell JR, Hodges JR, Rowe JB. Cognition in corticobasal syndrome and progressive supranuclear palsy: a review. *Mov Disord*. 2014;29(5):684–693.

Calleo J, Burrows C, Levin H, et al. Cognitive rehabilitation for executive dysfunction in Parkinson's disease: application and current directions. *Parkinsons Dis*. 2012;2012:512892. Erratum in: *Parkinsons Dis*. 2012;2012:907513.

Chen JJ, Marsh L. Anxiety in Parkinson's disease: identification and management. *Ther Adv Neurol Disord*. 2014;7(1):52–59.

Cummings JL, Chow T, Masterman D. Encephalitis lethargica: lessons for neuropsychiatry. *Psychiatr Ann*. 2001;31:165.

Cummings J, Friedman JH, Garibaldi G, et al. Apathy in neurodegenerative diseases: recommendations on the design of clinical trials. *J Geriatr Psychiatry Neurol*. 2015;28(3):159–173.

DeLong MR, Wichman T. Basal ganglia circuits as targets for neuromodulation in Parkinson's

disease. *JAMA Neurol*. 2015;72(11):1354–1360.

Elbaz A. Prodromal symptoms of Parkinson's disease: implications for epidemiological studies of disease etiology. *Rev Neurol (Paris)*. 2016;172(8-9):503–511.

Giugni JC1, Okun MS. Treatment of advanced Parkinson's disease. *Curr Opin Neurol*. 2014;27(4):450–460.

Gustafsson H, Nordström A, Nordström P. Depression and subsequent risk of Parkinson disease: a nationwide cohort study. *Neurology*. 2015;84(24):2422–2429.

Hermann A, Walker RH. Diagnosis and treatment of chorea syndromes. *Curr Neurol Neurosci Rep*. 2015;15(2):51413.

Josephs KA. Key emerging issues in progressive supranuclear palsy and corticobasal degeneration. *J Neurol*. 2015;262(3):783–788.

Lashley T, Rohrer JD, Mead S, Revesz T.· Review: an update on clinical, genetic and pathological aspects of frontotemporal lobar degenerations. *Neuropathol Appl Neurobiol*. 2015;41(7):858–881.

Obeso JA, Rodriguez-Oroz MC, Stamelou M, Bhatia KP, Burn DJ. The expanding universe of disorders of the basal ganglia. *Lancet*. 2014;384(9942):523–531.

Ojo OO, Fernandez HH. Current understanding of psychosis in Parkinson's disease. *Curr Psychiatry Reports*. 2016;18(10):97.

Ramirez-Zamora A, Zeigler W, Desai N, Biller J. Treatable causes of cerebellar ataxia. *Mov Disord*. 2015;30(5):614–623.

Rosness TA, Engedal K, Chemali Z. Frontotemporal dementia: an updated clinician's guide. *J Geriatr Psychiatry Neurol*. 2016;29(5):271–280.

Ross CA, Aylward EH, Wild EJ, et al. Huntington disease: natural history, biomarkers and prospects for therapeutics. *Nat Rev Neurol*. 2014;10(4):204–216.

Sawamoto N, Honda M, Hanakawa T, Fukuyama H, Shibasaki H. Cognitive slowing in Parkinson's disease: a behavioral evaluation independent of motor slowing. *J Neuroscience*. 2002;22(12):5198–5203.

Schmahmann JD. The role of the cerebellum in cognition and emotion: personal reflections since 1982 on the dysmetria of thought hypothesis, and its historical evolution from theory to therapy. *Neuropsychol Rev*. 2010;20(3):236–260.

Shakiba A. The role of the cerebellum in neurobiology of psychiatric disorders. *Neurol Clin*. 2014;32(4):1105–1115.

Shannon KM, Fraint A. Therapeutic advances in Huntington's disease. *Mov Disord*. 2015;30(11):1539–1546.

Sinoff G, Badarny S. Mild cognitive impairment, dementia, and affective disorders in essential tremor: a prospective study. *Tremor Other Hyperkinet Mov (NY)*. 2014;4:1–7.

Stankovic I, Krismer F, Jesic A, et al; Movement Disorders Society MSA (MODIMSA) Study Group. Cognitive impairment in multiple system atrophy: a position statement by the Neuropsychology Task Force of the MDS Multiple System Atrophy (MODIMSA) study group. *Mov Disord*. 2014;29(7):857–867.

van Duijn E, Craufurd D, Hubers AA, et al; European Huntington's Disease Network Behavioural Phenotype Working Group. Neuropsychiatric symptoms in a European Huntington's disease cohort (REGISTRY). *J Neurol Neurosurg Psychiatry*. 2014;85(12):1411–1418.

Verstraeten A, Theuns J, Van Broeckhoven C. Progress in unraveling the genetic etiology of Parkinson disease in a genomic era. *Trends Genet*. 2015;31(3):140–149.

Wang HF, Yu JT, Tang SW, et al. Efficacy and safety of cholinesterase inhibitors and memantine in cognitive impairment in Parkinson's disease, Parkinson's disease dementia, and dementia with Lewy bodies: systematic review with meta-analysis and trial sequential analysis. *J Neurol Neurosurg Psychiatry*. 2015;86(2):135–143.

Zurowski M, McDonald WM, Fox S, Marsh L. Psychiatric comorbidities in dystonia: emerging concepts. *Mov Disord*. 2013;28(7):914–920.

▲ 2.7 Neuropsychiatric Aspects of Multiple Sclerosis and Other Demyelinating Disorders

RUSSELL T. JOFFE, M.D.

INTRODUCTION

The demyelinating disorders, including multiple sclerosis (MS), are a heterogeneous group of disorders of the CNS. The demyelinating disorders share a common pathology, namely, the partial or complete loss of the myelin sheath surrounding axons in the white matter of brain and spinal cord. As these axons are widely disseminated throughout the brain and spinal cord, lesions may occur diffusely, resulting in complex and widespread neurological and psychiatric symptoms.

The study of the neuropsychiatric effects of the demyelinating disorders, particularly MS, has been the focus of a consistent research effort over many years. This interest has been sustained by an increasing awareness that different forms of psychopathology are very common in these disorders and that the prevalence exceeds, which would be expected to occur in other similar chronic and disabling conditions. Moreover, it has also been hoped that these disorders, especially MS, may represent a biological model for primary mood disorders, which may facilitate the elucidation of the etiology and neuroanatomy of primary mood and other psychiatric disorders.

The etiology of the demyelinating disorders is diverse and largely conforms to the traditional, clinical classification of etiology of disorders, that is, congenital and acquired. Under the congenital disorders are included presumed genetic and chromosomal abnormalities. The chromosomal abnormalities include, for example, the gangliosidoses, such as Tay–Sachs disease, which usually presents in infancy or early childhood. Another group of disorders includes the leukodystrophies, of which there are numerous subtypes, such as metachromatic, globoid cell, and adrenoleukodystrophy, as well as Hurler syndrome. The acquired demyelinating disorders include infectious disease, particularly AIDS; trauma, including open and, especially, closed brain trauma; vascular

disorders, including the vascular dementias; toxins, including alcohol and other solvents; and autoimmune disorders, the most prominent of which is MS. This list is far from inclusive, and many disease types may cause a demyelinating encephalopathy. By far, the most common and well-studied of these disorders is MS.

MULTIPLE SCLEROSIS

Epidemiology and Clinical Features

MS is a demyelinating disorder of the brain and spinal cord that has been described in the literature for more than 150 years. The French neurologist Jean-Martin Charcot first established diagnostic criteria for MS in the middle of the 19th century. His original diagnostic triad is now known not to be specific, and in fact, over the years, numerous other diagnostic criteria were developed, all of which had a lack of specificity for the illness, thus including many false positives.

The epidemiology of MS has been studied in most countries, but a clear geographic distribution remains difficult to clarify. The highest prevalence rates appear to occur in Caucasians of northern European descent and in the United States, a clear north–south gradient of prevalence has been demonstrated. The disease is particularly rare in certain Asian populations, such as Uzbekistanis, native Siberians, North and South Native Americans, Chinese, and Japanese, as well as African Americans and New Zealand Maoris. The data suggest more than just a geographical vulnerability but also strong genetic influences. The disease is far more common in women and, most commonly, is diagnosed in the age group of 20 to 50 years of age.

The diagnosis of MS is a largely clinical one, although MRI has greatly supported, but not superseded, the clinical diagnosis. MRI is not always specific for the diagnosis of MS and may lead to a false-positive diagnosis if it is relied on exclusively.

Currently, the diagnosis of MS is clinically defined as the presence of two attacks, involving different parts of the CNS, separated by a period of at least 1 month, and lasting a minimum of 24 hours.

The laboratory support for the diagnosis of MS, apart from MRI lesions, consists of demonstration in the CSF of at least two oligoclonal bands or increased CNS synthesis of immunoglobulin G (IgG). In addition to this evidence, evoked potentials, particularly visual or somatosensory evoked potentials, are slowed or absent and may contribute or support the clinical diagnosis. The clinical diagnostic criteria specified allow one to distinguish MS from disseminated encephalomyelitis, which is most commonly a single episode of illness, with all lesions occurring within a few weeks, although multiple separate lesions do sometimes occur.

The initial symptoms of MS are nonspecific and heterogeneous and may include a broad range of somatosensory, ocular, and motor symptoms. The

motor symptoms could include transient hemiparesis and various forms of seizure, among others. Symptoms of cerebellar dysfunction may occur, including scanning speech, gait ataxia, and various tremors. Although these are common in the initial presentation, almost any neurological symptom can be the first presentation of MS.

MS has a variable course over time. The most common course is the relapsing–remitting form where the patient has relapses of illness that can present with a broad range of neurological symptoms leading to remission and almost complete recovery. The secondary progressive form may evolve from the relapsing–remitting form, leading to incomplete recovery after acute exacerbations of illness. Occasionally, the secondary progressive form can occur at the outset of the illness, and the disorder then takes on a chronic progressive nature. The last type is quite uncommon, occurring in up to 10 percent of cases, and is known as the primary progressive type in which there is a steady worsening of course from the outset of the disease. This form may be rapidly progressive and is rarely seen in patients younger than 40 years of age.

The relationship between these different forms is not clearly understood, and it is uncertain to what extent they represent different types of illness with regard to etiology. The etiology of MS is becoming more fully understood as an autoimmune disorder leading to damage to the myelin with an inflammatory reaction, resolution of which leads to the formation of plaque in various parts of the brain.

Neuropsychiatric Aspects

There are several broad groups of mental symptoms that affect patients with MS. These include:

1. Fatigue
2. Psychiatric disorders
3. Emotional symptoms
4. Cognitive impairment

Fatigue

Fatigue is one of the most common presenting symptoms of patients with all demyelinating disorders as MS. It also substantially contributes to the disability and poor quality of life that these patients experience. Fatigue of MS has been extensively described. It is referred to as a physical fatigue and is especially well described in MS. Fatigue may be profound and usually includes both physical fatigue that makes the patient feel less able to perform physical activity as well as mental fatigue, which makes the patient unable to participate in his or her usual activities, including work. Fatigue is one of the most prominent and enduring symptoms of MS and it is estimated to occur in almost all patients with MS and to be the

predominant and often the presenting symptom at time of diagnosis in up to 50 percent of cases. Fatigue in MS can have a reverse diurnal rhythm in that it is worse as the day progresses so that many patients with MS alter their lifestyle to do much of their activities early in the day before the fatigue of the later day and evening becomes disabling.

The etiology of fatigue in MS is complex and multidetermined and remains an area of active research. There are no clear etiological factors identified, but neurological and endocrine factors may contribute Other factors may also worsen fatigue such as comorbid medical disorders, associated depression, sleep difficulties, and the side effects of medications used to treat the symptoms of MS and comorbid conditions. In particular, the fatigue may be related to comorbid depression, where fatigue is also a common symptom and treatment of depression may, to a varying extent, alleviate the symptom of fatigue.

Psychiatric Disorders

A broad range of psychiatric disorders are more common in patients with MS. The most prominent among them are disorders, particularly depression. A large number of patients with MS report symptoms of depression, and many are diagnosed with a major depressive disorder. Various studies have reported that between 20 and 50 percent of patients with MS have major depression at some point during the course of their illness. Point prevalence rates may be as high as 10 to 20 percent. The exact causal relationship between depression and MS remains unclear but, as described by G.M. Galeazzi and colleagues in a review, may be explained as:

1. Depression as a prodromal feature of MS
2. Depression and MS as coincidental but independent conditions
3. Depression and MS as the same condition with common etiology
4. Depression as a reaction to the chronicity and disability of MS
5. Depression as a side effect of MS treatment particularly corticosteroids and interferon
6. A combination of the above

The occurrence of a major depressive syndrome in MS patients may substantially add to the disability and the burden of illness associated with MS. Besides the classic features of a major depressive disorder, affective lability, sadness or depression, and debility may be prominent features of the disorder.

There are limited studies on antidepressant treatment of depressive syndromes associated with MS, and, in the studies to date, the efficacy of antidepressants has not been clearly demonstrated. The same applies to the use of nonpharmacological treatments such as cognitive behavior therapy (CBT). These studies generally have methodological limitations so that despite the limited and equivocal research evidence, both antidepressants

and CBT are widely used to treat depressive and anxiety disorders in MS.

In addition to depression, it has been observed that bipolar affective disorder may occur to a greater degree in patients with MS. This may be a phenomenon secondary to the MS with periods of depression, alternating with periods of emotional dyscontrol or irritability. In addition, in bipolar patients with a rapid-cycling course, the possibility of MS as a course-modifying factor for the bipolar disorder should be seriously considered. Besides bipolar and unipolar mood disorders, other psychiatric disorders such as anxiety and psychotic disorders may be more common in MS. Suicide may be twice as common in MS patients.

Emotional Symptoms

The recent evidence-based guidelines from the American Academy of Neurology describe emotional symptoms as a comorbid issue distinct from psychiatric disorders. They describe two broad categories: emotional symptoms such as depressive and anxiety symptoms that have a complex, multifactorial etiology and affective symptoms that derive directly from the MS disease process. The latter symptoms include pseudobulbar affect, euphoria, and apathy. Emotional and affective symptoms may occur simultaneously and there are limited therapeutic options for which there is good evidence for efficacy in this population.

Cognitive Impairment

Cognitive impairment is common in patients with MS. It has been estimated that up to 50 to 70 percent of patients with MS show evidence of cognitive impairment on an appropriate battery of neuropsychological tests. As with other psychiatric symptoms, there are inconclusive data to support a clear correlation between these cognitive difficulties and other clinical variables, such as neurological dysfunction, CNS lesions, and degree of disability. Nonetheless, cognitive symptoms contribute substantially to the dysfunction and to poor quality of life that occurs in patients with MS.

Cognitive dysfunction in MS may be broad-based and may include many different features. For example, all aspects of memory function may be impaired. Moreover, higher-order functions, such as information processing or executive function, may also be impaired, and there is evidence in some cases of MS of subcortical dementia. In addition, abnormalities of attention and concentration may be evident, and these may contribute substantially to work impairment. Abnormalities of language are less common, although some dysfunction may be determined on broad-based neuropsychological tests. Recent studies suggest that there is increasing consensus on the specific cognitive deficits most common in MS. These include deficits of memory, speed of information processing,

visuospatial function, and some aspects of executive function. In contrast, intellectual and language function remain relatively intact.

Most of the studies of cognitive impairment in MS are cross-sectional in nature. There are limited data on the longitudinal course of cognitive difficulty. In general, however, these studies show a gradual decline in all aspects of cognitive function in the patient with MS.

MS presents a substantial challenge as far as differential diagnosis is concerned. In those with the common relapsing–remitting course of illness, together with typical clinical, laboratory, and MRI findings, the diagnosis is more easily made. However, in atypical cases, differential diagnosis can be difficult. This applies particularly at the onset of the disease, during the first episode, when the conventional clinical criteria are hard to apply. With the first attack of MS, it is difficult to differentiate this illness from the other demyelinating disorders.

Differential Diagnosis

There are a variety of other conditions to consider in the differential diagnosis of MS. These include:

1. Disseminated encephalomyelitis: Factors that may point to this diagnosis include systemic symptoms, signs of cortical dysfunction, and, perhaps, bilateral optic neuritis. MRI may be quite characteristic of this disorder and helps distinguish it from MS.
2. Cerebral vasculitis: This has numerous causes. They include primary cerebral vasculitis, which is extremely rare, associated with systemic lupus erythematosus, AIDS, lymphoproliferative disorders, and Wegener granulomatosis. Various ancillary laboratory findings may help distinguish these various forms of disease. Rarer forms of vasculitis may include Behçet disease and also CNS sarcoidosis.
3. Leukodystrophies: These disorders are characterized by the accumulation of unchanged adrenoleukodystrophy, which may have an adult-onset form. These conditions are rare and can usually be identified by the presence of the family history.
4. Chronic fatigue syndrome: It has been previously emphasized that fatigue is one of the most predominant symptoms of MS. Chronic fatigue syndrome, also known as fibromyalgia, requires differentiation, especially if no focal neurological symptoms are found. Failure to elicit specific neurological symptoms, as well as the presence of other associated features of fibromyalgia, including sleep disturbance, myalgias, and arthralgias, among other relatively nonspecific symptoms, is supportive of this diagnosis.

FUTURE DIRECTIONS

The demyelinating disorders present a diagnostic challenge. The most

common of them, MS, requires differentiation from other disorders because of its distinct clinical presentation and course of illness. Moreover, there are now specific treatments that may reduce relapse rates and degree of dysfunction, and, therefore, a definitive diagnosis is necessary. MS can present with a variety of psychiatric issues, particularly depressive syndromes and symptoms. While the broad range of pharmacological and nonpharmacological interventions used in primary psychiatric disorders are often used, with good effect, in MS patients, large-scale, well-designed studies are needed to provide a good evidence base for the treatment of psychiatric symptoms and disorders associated with this condition, which may require specific intervention or psychotropic agents.

References

Amato MP, Ponziani G, Siracusa G, Sorbi S. Cognitive dysfunction in early-onset multiple sclerosis: a reappraisal after ten years. *Arch Neurol.* 2001;58:1602–1606.

Arnett PA, Barwick FH, Beeney JE. Depression in multiple sclerosis: review and theoretical proposal. *J Int Neuropsychol Soc.* 2008;14(5):691–724.

Bamer AM, Cetin K, Johnson KL, Gibbons LE, Ehde DM. Validation study of prevalence and correlates of depressive symptomatology in multiple sclerosis. *Gen Hosp Psychiatry.* 2008;30(4):311–317.

Bashir K, Whitaker JN. Current immunotherapy for demyelinating diseases. *Arch Neurol.* 2001;58:1611.

Benedict RH, Weinstock-Guttman B, Fishman I, Sharma J, Tjoa CW. Prediction of neuropsychological impairment in multiple sclerosis. *Arch Neurol.* 2004;61:226–230.

Bobholz JA, Rao SM. Cognitive dysfunction in multiple sclerosis: a review of recent developments. *Curr Opin Neurol.* 2003;16:283–288.

Brinar VV. The differential diagnosis of multiple sclerosis. *Clin Neurol Neurosurg.* 2002;104:211–220.

Brown RF, Valpiani EM, Tennant CC, Dunn SM, Sharrock M, Hodgkinson S, et al. Longitudinal assessment of anxiety, depression, and fatigue in people with multiple sclerosis. *Psychol Psychother.* 2008;82(Pt 1):41–56.

Feinstein A, Feinstein K. Depression associated with multiple sclerosis: looking beyond diagnosis to symptom expression. *J Affect Disord.* 2001;66:193–198.

*Feinstein A, Magalhaes S, Richard JF, Audet B, Moore C. The link between multiple sclerosis and depression. *Nat Rev Neurol.* 2014;10(9):507–517.

Figved N, Benedict R, Klevan G, Myhr K, Nyland H, Landrø N, et al. Relationship of cognitive impairment to psychiatric symptoms in multiple sclerosis. *Mult Scler.* 2008;8:1084–1090.

*Galeazzi GM, Ferrari S, Giaroli G, Mackinnon C, Merelli E. Psychiatric disorders and depression in multiple sclerosis outpatients: impact of disability and interferon β therapy. *Neurol Sci.* 2005;26:255–285.

Ghaffar O, Feinstein A. The neuropsychiatry of multiple sclerosis: a review of recent developments. *Curr Opin Psychiatry.* 2007;20:278–285.

Gold SM, Schulz H, Monch A, Schulz KH, Heesen C. Cognitive impairment in multiple sclerosis

does not affect reliability and validity of self-report health measures. *Mult Scler*. 2003;9:404–410.

*Greim B, Engel C, Apel A, Zettl UK. Fatigue in neuroimmunological diseases. *J Neurol*. 2007;254(Suppl 2):II102–II106.

Hoffmann S, Tittgenneyer M, von Cramon DY. Cognitive impairment in multiple sclerosis. *Curr Opin Neurol*. 2007;20:275–280.

Holmes G, Kaplan J, Gantz N. Chronic fatigue syndrome: a working case definition. *Ann Intern Med*. 1988;108:387–389.

Janculjak D, Mubrin Z, Brzovicz Z, Birnar V. Change in short-term memory processes in patients with multiple sclerosis. *Eur J Neurol*. 1999;6:1.

Janssens AC, van Doorn PA, de Boer JB, Kalkrs NF, van der Meche FG, Passchier J. Anxiety and depression influence the relation between disability status and quality of life in multiple sclerosis. *Mult Scler*. 2003;9:397–403.

*Joffe RT, Ippert GP, Gray T, Sawa G, Horvath Z. Mood disorder multiple sclerosis. *Arch Neurol*. 1987;44:376–378.

Kesseloring J, Miller D, Robb S. Acute disseminated encephalomyelitis, MRI findings and the distinction for multiple sclerosis. *Brain*. 1990;113:291–302.

Koch M, Uyttenboogaart M, van Harten A, Heerings M, De Keyser J. Fatigue, depression and progression in multiple sclerosis. *Mult Scler*. 2008;14(6):815–822.

Lamberg L. Psychiatric symptoms common in neurological disorders. *JAMA*. 2001;286:711–719.

McDonald W, Holliday A. Diagnosis and classification of multiple sclerosis. *Br Med Bull*. 1977;33:4–9.

Miller A, Galboyz Y. Multiple sclerosis: from basic immunopathology to immune intervention. *Clin Neurol Neurosurg*. 2002;104:172–176.

*Minden SL, Feinstein A, Kalb RC, Miller D, Mohr DC, Patten SB, et al. Guideline Development Subcommittee of the American Academy of Neurology. Evidence-based guideline: assessment and management of psychiatric disorders in individuals with MS: report of the Guideline Development Subcommittee of the American Academy of Neurology. *Neurology*. 2014;82(2):174–181.

Minden SL, Roav J, Reich P. Depression in multiple sclerosis. *Gen Hosp Psychiatry*. 1987;9:426–434.

Minden SL, Schiffer RB. Affective disorders in multiple sclerosis. Review and recommendations for clinical research. *Arch Neurol*. 1990;47:98–104.

Mohr DC, Hart SL, Goldberg A. Effects of treatment for depression on fatigue in multiple sclerosis. *Psychosom Med*. 2003;65:542–547.

Poser CM. Onset symptoms in multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 1995;58:253–254.

Poser CM, Brinar VV. Multiple sclerosis 2001. *Clin Neurol Neurosurg*. 2002;104:165–167.

Pugliatti M, Sotgiu S, Rosati G. The worldwide prevalence of multiple sclerosis. *Clin Neurol Neurosurg*. 2002;104:182–191.

Pullicino PM, Ostrow PT, Kwen PL. Cerebral white matter disease: imaging, clinical, and pathological aspects. *Neurologist*. 1996;2:288.

Rabins PV, Brooks BR, O'Donnell P, Parlson GD, Mobr GP. Structural brain correlates of

emotional disorder in multiple sclerosis. *Brain*. 1986;109:585–597.

Rao SM, Leo GJ, Burnardin L, Unvrzagt F. Cognitive dysfunction in multiple sclerosis. 1. Frequency, patterns and prediction. *Neurology*. 1991;41:685–691.

Shafey H. The effect of fluoxetine in depression associated with multiple sclerosis. *Can J Psychiatry*. 1992;37:147–148.

Triulzi F, Scotti G. Differential diagnosis of multiple sclerosis: contribution of magnetic resonance techniques. *J Neurol Neurosurg Psychiatry*. 1998;64:S6–S14.

Zarei M, Chandran S, Compston A, Hodges J. Cognitive presentation in multiple sclerosis: evidence for a cortical variant. *J Neurol Neurosurg Psychiatry*. 2003;74:872–877.

*Zephir H, De Seze J, Stojkovic T, Dellise R, Ferriby D. Multiple sclerosis in depression. Influence of interferon β therapy. *Mult Scler*. 2003;9:284–288.

Zorzon M, de Masi R, Nasuelli D, Ukmar M, Mucelli RP. Depression and anxiety in multiple sclerosis. A clinical and MRI study in 95 subjects. *Neurology*. 2007;24:416–421.

▲ 2.8 Neuropsychiatric Aspects of HIV Infection and AIDS

GLENN JORDAN TREISMAN, M.D., PH.D., JEFFREY HSU, M.D., HEIDI E. HUTTON, PH.D., AND ANDREW F. ANGELINO, M.D.

INTRODUCTION

The HIV epidemic continues to be a major public health problem more than 30 years after the initial discovery of the infection and the delineation of the routes by which it is spread. The virus has now become a chronic disease for most patients, and effective multidrug therapy (referred to as HAART—**h**ighly **a**ctive **a**nti**r**etroviral **t**herapy or more commonly now CART—**c**ombined **a**nti**r**etroviral **t**herapy) so effectively inhibits viral replication that it is undetectable in the blood of most treated patients. Psychiatric disorders play several roles in the HIV epidemic, in elements of treatment, and in comorbidity. Psychiatric conditions are associated with increased risk behaviors for infection, with decreased access to care, with decreased adherence to HIV therapies, and with increased medical comorbidity. Because of this, some authors have gone as far as saying that HIV is a psychiatric epidemic.

Prevention of HIV has been a goal since the mid-1980s when the routes of transmission were defined and testing became available. A massive effort at education resulted in a decrease in the spread of HIV in people able to modify their risk behaviors, but began to select for vulnerable populations that were either unable or unwilling to modify their behavior or unable to understand the risks. While there are still individuals who contract HIV as a result of ignorance and bad luck, the vast majority of new infections arise in individuals who are aware of HIV but continue to engage in high-risk behaviors anyway. As the media attention regarding HIV has diminished and with more effective and better tolerated antiviral cocktails, some of the people have become less cautious and there has been a resurgence of risk behaviors in some groups (see below). Since many of the behaviors that can result in HIV infection are those that are associated with the brain's reward systems, it is understandable that humans would be drawn to these behaviors, and that they would be common. Sex and substance use are indeed common behaviors. Individuals with various types of untreated mental illness are vulnerable to practice them unsafely at a higher frequency, and thus increased risk of HIV infection. One of the new concepts that has arisen in the HIV literature is the concept of "treatment as prevention," the idea that aggressive treatment of all infected individuals will result in a dramatic decrease in new cases and thus prevent the continued spread of the disease. This has resulted in several "test and treat" programs to try to accelerate the progress of patients in achieving an

undetectable viral load, a condition where they are minimally infectious to others.

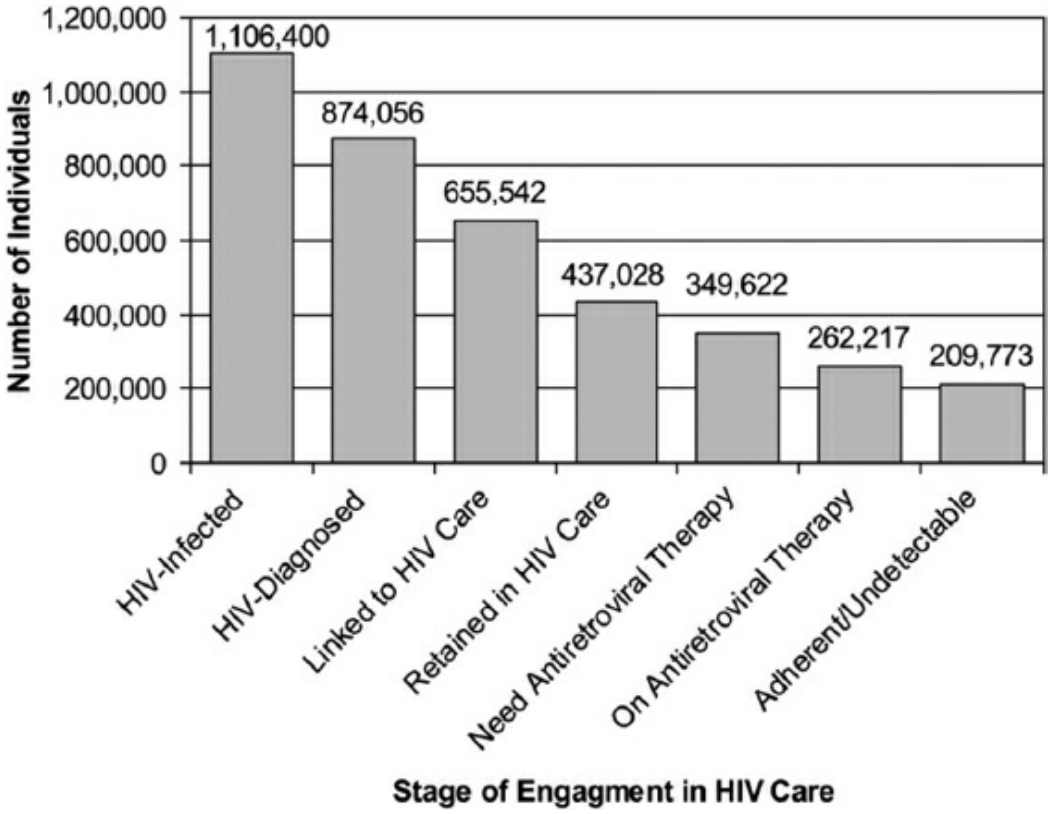

FIGURE 2.8–1. The spectrum of engagement in HIV care in the United States spanning from HIV acquisition to full engagement in care, receipt of antiretroviral therapy, and achievement of complete viral suppression. The authors estimate that only 19 percent of HIV-infected individuals in the United States have an undetectable HIV load. (From Gardner EM, McLees MP, Steiner JF, del Rio C, Burman WJ. Strategies for prevention of HIV infection. *Clin Infect Dis.* 2011;52(6):793–800.)

Once infected, there is a stepwise hierarchy accessing care (see Fig. 2.8–1). Based on the description by Eldred and Malitz, it is possible to stratify the degree of engagement in care in various populations. Vulnerable populations (such as those with psychiatric conditions) tend to lag behind in accessing care.

Further, once an individual is engaged in care, there is a need for nearly complete adherence to the antiviral regimen prescribed for the patient. Individuals with untreated mental illness are vulnerable to nonadherence, and therefore have higher viral loads, more immunosuppression, more morbidity, and a higher mortality rate. If these same individuals continue their high-risk behaviors for spreading HIV infection to others, the epidemic is perpetuated.

Individuals with untreated mental illnesses are more likely to engage in high-risk behaviors (leading to greater risk of infection) and less likely to take their medications correctly, leading to greater viral loads in their blood and an increased risk of infecting others. This makes the recognition and

treatment of mental illness crucial to putting an end to the spread of HIV and implies a critical role for psychiatry in integrated care for these patients.

This chapter is organized around those conditions commonly seen in the HIV clinic and those that are clearly involved in the risk of HIV or are clearly barriers to HIV treatment. The introductory part of the chapter is a medical overview of HIV disease. The second part considers psychiatric conditions associated with HIV, and the third part considers related neurological issues that have an impact on psychiatric function and are part of the differential diagnosis of any patient presenting with HIV and psychiatric complaints.

OVERVIEW OF HIV INFECTION AND AIDS

Epidemiology

HIV was originally recognized through a series of case descriptions involving young homosexual men with *Pneumocystis carinii* (now *Pneumocystis jiroveci*) pneumonia and other rare conditions, including an aggressive form of Kaposi sarcoma, (a condition seen prior to that time most often in older Caucasian men of middle eastern extraction) in the early 1980s in Los Angeles. The conditions seen were mostly infections associated with immune systems that had been compromised, termed "opportunistic" infections. Subsequently, it became clear that these patients had severe immune system compromise and were vulnerable to infections that had been seen in other immunocompromised individuals.

The 2012 UN AIDS fact sheet gives the number of currently infected people with HIV in the world as 35.2 million (most in sub-Saharan Africa) with an estimated total number of infections at 75 million since the beginning of the epidemic. There were 2.3 million new infections in 2012 down from 3.1 million in 2001. In the United States, most estimates suggest that about 1.2 million people are infected with HIV.

Several populations have been shown to have increased risk of infection. Early in the epidemic, transfusion-associated infection represented a significant risk, particularly with those who required pooled or high-frequency transfusions, such as those with hemophilia and clotting factor deficiencies and those with blood dyscrasias, but the transfusion-related risk has been reduced to about one in 1.5 million. Currently, the populations at highest risk are homosexual men, intravenous (IV) drug users, female partners of IV drug users, and those who trade sex for money or drugs. Mother to fetus transmission remains a serious problem in some populations, with transmission originally occurring in 25 to 30 percent of live births depending on delivery type, severity of HIV disease, but has been dramatically curtailed in the United States to less than 2 percent by the development of antiviral treatment and prophylaxis at parturition.

Efforts at promoting safer sex practices, clean needle use, bleaching of drug paraphernalia, and prenatal assessment had an impact on behaviors on a great many individuals, as evidenced by the flattening of the epidemic curve from an exponential shape to a linear shape. Unfortunately, psychiatric disorders continue to play a significant role in thwarting prevention efforts by making vulnerable patients unable to access or benefit from prevention efforts. Men who have sex with men remain the group with the highest risk of new HIV infection, and psychiatric comorbidity, including addictions, remains one of the major drivers of new infections. Internationally, psychiatric disorders have been identified in both the developed and developing countries as an important comorbidity in the risk of infection, poorer likelihood of successful treatment, and higher morbidity and mortality.

The retrovirus subfamily to which HIV belongs is called lentivirus (lenti: slow). Early in the HIV epidemic, it was thought that HIV was a "slow virus' that integrated into the cell genome and then was latent until activated much later. In the early 1990s, this idea was challenged when studies demonstrated very large rates of turnover of CD4 expressing T-lymphocyte cells during the course of HIV infection, with the production of vast quantities of virus. Approximately 10 billion viral particles are manufactured on an average day. HIV infection can be described as a war of attrition in which the virus gradually depletes the immune system. During this fight, the body mobilizes enormous immune system resources, producing up to 2 billion CD4 cells a day. HIV has a relatively high error rate when it is transcribed, and therefore produces large numbers of mutants, thus giving rise to strains of virus that can be resistant to drugs, target different cell receptors, and evade the body's strategies for defense. More remarkably, the virus has several specific mechanisms of disrupting the immune defenses of the host.

HIV infects cells expressing the CD4 surface protein (referred to as the CD4 antigen or receptor), mostly T lymphocytes and macrophages. The main disease mediator is the infection and destruction of T4 helper cells or CD4 cells critical in cell-mediated immunity. The average person has approximately 1000 CD4 lymphocytes per cubic millimeter (or microliter) of blood. Before the development of effective treatment for HIV, the average patient's CD4 cell count dropped approximately 100 CD4 cells per cubic millimeter (or microliter) per year, so over a 10-year average course the infected person would lose effective immune function and develop infections that are usually neutralized by cell-mediated immunity (Fig. 2.8–2). The immune system becomes quite compromised as the CD4 cell count drops below 200, and "cellular" immunity diminishes and is finally lost. At this point, the patient becomes vulnerable to opportunistic infections. There are also several tumors that occur in immunocompromised patients that are also found in patients with AIDS.

The term “AIDS” (Acquired Immunodeficiency Syndrome) is used to describe the syndrome in which patients develop these opportunistic medical conditions.

It is now recognized that HIV disease represents a continuum, with an increasing liability to opportunistic infection that correlates well with decreasing CD4 counts. The CD4 count can be used to track the progress of the infection and the status of the immune system. As shown in the diagram, untreated patients with HIV infection have a brief severe decline in CD4 count immediately after infection. As the immune response to HIV begins, the CD4 count returns to normal levels but then begins a slow decline. The CD4 count decreases by about 100 cells per year of infection in the average patient. Over time, the effectiveness of immune function declines and the patient may develop the conditions shown in Figure 2.8–2. In the years prior to the development of effective treatment, this graph served as a kind of grim road map that ended inevitably in the death of most patients.

There were several major breakthroughs that changed the overall understanding of the infection: the ability to measure CD4 cells, the ability to measure and quantitate the “viral load” (the quantity of viral particles in the blood), and the ability to detect HIV mutant virus and predict sensitivity to antiviral medications.

HIV disease was originally staged purely by clinical indicators, which fell into three phases: asymptomatic, AIDS-related complex (ARC), and finally AIDS. The WHO currently stages HIV by clinical symptoms 1 to 4 with increasing clinical severity. These are not based on laboratory testing but instead on clinical presentation for use in resource limited settings. The United States Center for Disease Control and Prevention (CDC) has recently adopted a staging system that uses CD4 cell number and years of infection to stage the HIV, as well as adding a Stage 0 that includes those who are not yet clearly infected but have indications of being infected. The table for current staging can be found at http://www.cdc.gov/hiv/statistics/recommendations/terms.html. Those with normal T cells and clinically essentially asymptomatic are in Stage 1, those with lower T cells than the length of their infection would suggest are in Stage 2 and may have some symptoms, and those with advanced disease as measured by low CD4 number or by opportunistic infection are in Stage 3.

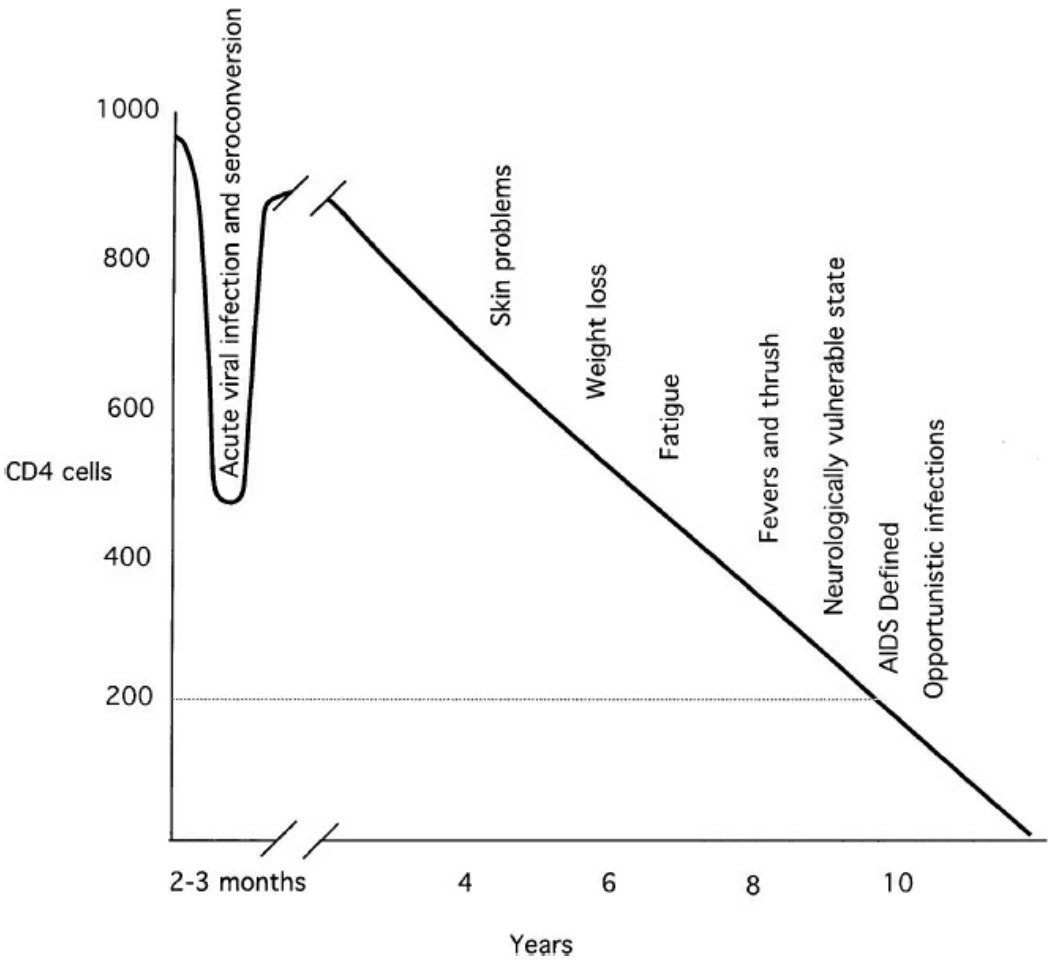

FIGURE 2.8–2. Schematic CD4 count decline in an untreated HIV-infected patient and the points at which immunosuppression and opportunistic diseases may be expected to occur. (Courtesy of John G. Bartlett, M.D.)

Virology

The HIV virus itself is thought to have originated in Africa. It is genetically related to a family of retroviruses that all infect immune cells that have been extensively researched and mapped. The generic term HIV usually refers to HIV-1, while a second genetically distinct type of human-specific immune virus called HIV-2 also targets humans. The CDC has released specific definitions required for a confirmed diagnosis of HIV-2. It is prevalent in regions in Western Africa. HIV is mostly likely a product of spread to humans from chimpanzees and gorillas. Current molecular biological evidence suggests that SIV "jumped" to humans from both gorillas and chimpanzees and that the jump happened on at least four different occasions. SIV infections in chimpanzees and other large apes are pathologic, but many monkeys are endemically infected and the virus appears not to be pathologic and is tolerated as a chronic infection by those species. Based on genetic mapping and the identification of cases of confirmed infection dating from the early 1960s suggest that HIV jumped to humans in the early 20th century. The strongly held belief that HIV was

a result of oral polio vaccine trials contaminated by infected chimpanzee kidney cells in the Belgian Congo still has tremendous political appeal for some, but is felt to have been discredited by most virologists and epidemiologists.

HIV is a single-stranded RNA virus that selectively infects immune cells, particularly T lymphocytes and macrophages. RNA viruses use several strategies to infect cells and reproduce themselves. HIV is one member of a class of RNA viruses called retroviruses, all of which carry the enzyme reverse transcriptase (RT) (RNA-dependent DNA polymerase) in the viral particle and depend on this enzyme for successful infection. This enzyme is not found in uninfected cells, and therefore represents an ideal target for antiviral treatments. RT synthesizes viral DNA from the viral RNA strand. This DNA is then transported into the nucleus of the cell and directs the production of more viral RNA using the cell's synthetic machinery. This DNA is inserted into the nuclear DNA (where it is referred to as a provirus) and can lie dormant only to be activated when that cell is "switched on" to defend the host. The mechanism by which some cells become "latently infected" while others become active producers of virus is currently a topic of great interest, as it is thought that latent infection is the key factor that prevents eradication of HIV virus even with in patients with complete viral suppression with antiviral treatment for extended periods of time. The one documented cure of HIV occurred in an infected patient referred to as the "Berlin patient" who underwent two serial bone marrow transplants for Leukemia using a donor with an HIV resistant mutation.

The HIV viral particle is an icosahedral particle that is made up of an outer envelope and a viral core. The envelope is mostly derived from the host cell membrane. It is studded with viral proteins including gp120 and the transmembrane gp41 proteins, which perform specific roles in receptor binding, internalization, and subsequent viral budding. The viral envelope is lined with the matrix protein (MA) referred to as p17. The cone shaped core is made of capsid protein p24 and contains nucleic capsid protein p7 associated with two copies of single-stranded viral RNA and the enzymes RT and integrase (IN). The two strands of RNA may undergo recombination by crossing over. Infection with HIV virus requires direct contact between viral particle containing material with immune cells. Infection takes place when viral particles are bound to CD4 receptors on immune cells. There are two primary cell types that get infected: T cells that express CD4 receptors (T4 cells or CD4 cells) and macrophage-derived cells. The usual route of initial infection is viral contact with dendritic cells in mucus membranes that then present the viral particles to T cells. The viral particles bind to the CD4 receptors on the T cell surface. For infection to occur, a second receptor (a coreceptor) must also be present. Two coreceptors that have been identified (although there is evidence for several others) are the chemokine receptor found on macrophages (CCR5) and the

chemokine receptor found on some T cells (CXCR4). The HIV virus has "strains," or common mutations, that are more infectious to cells expressing particular chemokine receptors. The CCR5 selective strains seem to predominate early in the course of infection and are considered the more infectious strains. They target macrophages and are call M-tropic. Later in the course of infection, many patients develop increasing levels of CXCR4 topic virus. These are important because specific drugs have been developed that will inhibit the binding of coreceptors for both strains, but the strain of virus present must be determined by a tropism assay to determine whether these drugs will be effective. Mutations in the CCR5 receptor have been associated with reduced susceptibility to infection because this is the most common tropism in the early part of the infection.

VIRAL LIFE CYCLE AND TREATMENT OF HIV DISEASE AND AIDS

The steps by which human cells are infected are summarized in Figure 2.8–3. Initially, the viral gp120 protein binds to CD4 receptor expressing cells. Once bound, the virus binds to a second receptor, either CCR5 (most commonly) or CXCR4, and can then be internalized. During internalization, the outer coating of the viral particle fuses with the cell membrane and is removed. The enzyme RT carried within the viral particle then begins the process of transcribing the viral RNA into DNA. This DNA moves to the nucleus of the cell and is integrated into the human DNA genome using another viral enzyme called IN, an essential step, where it directs the cell machinery to produce large quantities of viral messenger RNA. The viral RNA then uses the cells own machinery to produce the proteins required to make viral particles and other viral proteins. The viral particles are assembled around viral RNA, and the mature virions bud outward from the cells surface. Finally, viral particles are cleaved by a specific protease from the cell surface and released into the host to infect other immune cells.

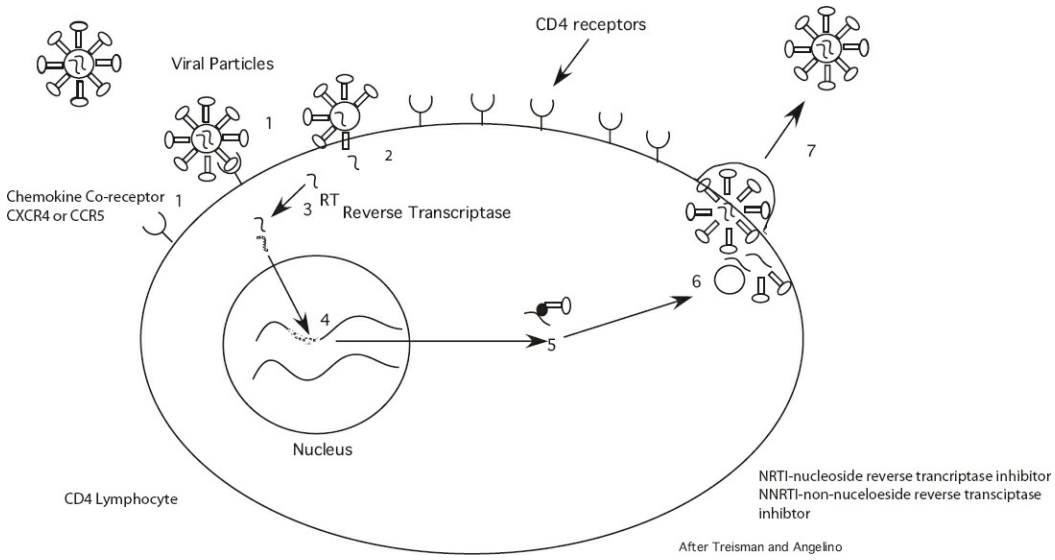

FIGURE 2.8–3. The HIV life cycle. HIV viral particles are shown attaching themselves to a CD4 expressing cell and then the steps of viral invasion and reproduction. 1. Binding of viral particle to CD4 receptor. Co-receptor (CXCR4 or CCR5) binding (site of co-receptor blocker Maraviroc-specific for CCR5). 2. Fusion of particle to cell (site of action of fusion inhibitor enfuvirtide [T-20]). 3. Reverse transcriptase RNA to DNA via viral enzyme (site of action of NRTIs and NNRTIs). 4. Integration into the nucleus by viral enzyme Integrase (site of action of integrase inhibitors). 5. Translation of viral proteins and manufacture of viral RNA strands. 6. Processing of viral contents and proteins into immature virions. 7. Release of immature (noninfectious virions) into mature virions by viral protease (site of action of protease inhibitors). (Courtesy of Glenn Treisman, M.D., Ph.D, and Andrew F. Angelino, M.D.)

Each of the steps in the process of HIV infection provides a potential target for antiviral agents (Fig. 2.8–3). Binding of gp120 to the CD4 receptor on the cell surface and binding to the coreceptor are both targets for inhibition. Maraviroc is a drug that inhibits binding to the CCR5 coreceptor and is an effective antiviral drug. Because some virus develops a predilection for the use of the CXCR4 coreceptor, a "tropism" assay has been developed to identify viral subtypes that make maraviroc less effective. The fusion of the viral particle with the cell membrane can be inhibited by enfuvirtide (also called T-20), although seldom used now because it must be injected, it remains a powerful antiviral drug.

Once the viral RNA enters the cell, it is "reverse transcribed" by an enzyme carried within the viral particle and not normally found in eukaryotic cells. This enzyme was the target for the first effective HIV drugs. The RT step is the most targeted step, and there are three classes of agents directed at this step on the market: the nucleoside RT inhibitors, the nucleotide RT inhibitors (collectively known as "nucs"), and the non-nucleoside RT inhibitors (the "non-nucs"). These were the first antiviral agents and continue to be widely used, with new agents in the pipeline for this step (Fig. 2.8–3).

The DNA copy made in this step is then integrated into the nucleus and the host genome by an enzyme called IN, now targeted by a class of drugs called IN inhibitors (raltegravir and dolutegravir). The viral DNA is transcribed to viral messenger RNA and is translated into viral proteins, assembled into virions, and finally cleaved form the cell membrane by a protease targeted by a class of drugs called protease inhibitors (Fig. 2.8–3).

Current antiviral treatment recommendations can be found at several websites, but the Johns Hopkins Guide summarizes most of the current information. The initial clinical use of these agents began in the late 1980s with AZT (originally called azidothymidine, now called zidovudine), which delayed the onset of AIDS and prolonged the lives of most patients by about 1 to 2 years, or led to the significant improvement in immune function in some patients. As new agents emerged, they were also deployed sequentially as monotherapy.

In the mid-1990s, several investigations studied combination therapies

of two RT inhibitors and a protease inhibitor. This "triple therapy" was later referred to as highly active antiretroviral therapy (referred to as HAART—**h**ighly **a**ctive **a**nti**r**etroviral **t**herapy or more commonly now CART—**c**ombined **a**nti**r**etroviral **t**herapy). CART dramatically reduced viral load (often to levels that were undetectable by the standard polymerase chain reaction [PCR] assay of the time) and often resulted in an increase in CD4 cell count. Subsequent studies have shown that extended treatment does not result in a cure, but that effective continued treatment results in years or even decades of time with little clinical disease manifestation. This is now understood to be the result of successful suppression of viral production to minimal levels, leading to the production of very few viral mutations. It takes several mutations in the viral genome to produce drug resistance; so in the presence of effective treatment, these mutations occur very slowly.

The preferred combinations used in HIV care have become very complex as more information has accumulated in clinical trials. The recommended combinations for initial therapy are mostly two "nucs" and a third drug. Ritonavir, one of the first protease inhibitors, inhibits the metabolism of other protease inhibitors, allowing for lower doses and is used in combination with other protease inhibitors to boost their effectiveness, essentially by inhibiting their metabolism. It has been joined by a second drug, cobicistat, which also boosts drug effectiveness by inhibiting metabolism, but unlike ritonavir it has no antiviral properties by itself. When combined with another protease inhibitor, the combination is referred to as a boosted protease inhibitor, but is counted as one drug. An ideal response to treatment is suppression of viral replication to a level where virus particles are undetectable. If this occurs, patients who continue to take their medications will continue to have undetectable viral loads without the emergence of resistant mutations. However, even a few missed doses were shown to be associated with viral production and thus emergence of mutations. CART combinations were initially quite burdensome for patients, as some drugs required special diets, refrigeration, very specific timing, and large numbers of pills. All of these interfered with the ability of patients to take treatment as prescribed. Newer treatments take advantage of drug combinations, drug–drug interactions that increase blood levels of some medications, and sustained release preparations. While there are very complex regimens for patients with multiple resistant HIV virus, the newest agents allow most patients to take one drug once a day or twice a day. The newer drugs provide treatment for even those who have been infected a long time and are resistant to most of the older therapies. The FDA-approved drugs for HIV are shown in Table 2.8–1.

Several controversial issues face clinicians as they begin to treat a patient for HIV. The first is when to treat patients with CART in the course

of HIV disease. The second is which antiviral agents to use and the third is when to change treatment. When CART therapy was first described, some clinicians hoped that aggressive early treatment would result in a cure. Because there has been little evidence that early treatment has improved the course for patients, and because all treatments have significant toxicity, previous recommendations for treatment suggested beginning later in the course of HIV. Treatment recommendations were usually based on CD4 cell counts tempered by viral load. Now, as studies have shown much less toxicity with the current agents and probably better survival with early treatment, most guidelines support earlier treatment. Several sets of treatment recommendations from different organizations have had minor differences. The decision may also tempered by other factors for many clinicians, including patient readiness for treatment evidenced by the ability to adhere to a regimen. In addition, the presence of comorbid diagnoses such as hepatitis C infection, substance use, psychiatric disorders, and other medical conditions influence the decision.

Table 2.8–1.
FDA-Approved HIV Drugs as of End of 2014

Brand Name	Generic Name	Abbreviation
Nucleotide and nucleoside reverse transcriptase inhibitors (NUCS)		
Combivir	zidovudine + lamivudine	AZT + 3TC
Emtriva	emtricitabine	FTC
Epivir	lamivudine	3TC
Epzicom	abacavir + lamivudine	ABC + 3TC
Hivid	zalcitabine	ddC
Retrovir	zidovudine	AZT or ZDV
Stribild	tenofovir + emtricitabine + elvitegravir	TDF + FTC + EVG
Trizivir	abacavir + zidovudine + lamivudine	ABC + AZT + 3TC
Truvada	tenofovir DF + emtricitabine	TDF + FTC
Videx	didanosine: buffered versions	ddl
Videx EC	didanosine: delayed-release capsules	ddl
Viread	tenofovir disoproxil fumarate	TDF
Zerit	stavudine	d4T
Ziagen	abacavir	ABC
Sustiva	efavirenz	EFV
Viramune	nevirapine	
Rescriptor	delavirdine	
Intelence	etravirine	
Edurant	rilpivirine	
Protease Inhibitors		
Agenerase	amprenavir	APV
Aptivus	tipranavir	TPV
Crixivan	indinavir	IDV
Invirase	saquinavir	SQV

Kaletra	lopinavir + ritonavir	LPV
Lexiva	fosamprenavir	FPV
Norvir	ritonavir	RTV
Prezista	darunavir	DRV
Reyataz	atazanavir	ATZ
Viracept	nelfinavir	NFV
Fusion Inhibitors		
Fuzeon	enfuvirtide (T-20)	T-20
CCR5 (Receptor Binding) Inhibitors		
Selzentry	maraviroc	
Integrase inhibitors		
Isentress	raltegravir	
Vitekta	elvitegravir	
Tivicay	dolutegravir	

Some of these drugs are only available as part of combination treatments; some of the combinations are not listed.

The choice of regimen must balance the complexity of the regimen, the side effects of the agents, and the type of mutations that will arise if a patient develops resistance. It is important to know if the patient has had prior antiretroviral treatment, and therefore may have developed resistance already. Specific drugs select for certain mutations, and some of the mutations confer resistance to whole classes of agents, while others may actually select for mutations that can increase sensitivity to other drugs. With currently available drugs, clinicians balance convenience and ease of a regimen, ability to preserve treatment options, toxicity and tolerability, and sometimes cost, in trying to select a regimen. In some cases, clinicians choose a regimen based on what future regimens will still work if the patient develops resistance to drugs in the first line.

Several key studies have shown that the single most important factor in successful treatment is the ability to accurately and completely take the prescribed medications. The term “treatment adherence” has replaced the older term “compliance” in the HIV literature, and refers to measurements of how accurately patients follow their prescription. Several studies show that adherence correlates highly with treatment success in terms of viral load reduction. Research in this area is complicated by difficulties in accurately assessing adherence. Clinician and patient assessments of adherence are often inaccurate. This problem was circumvented in a study of prison inmates treated with directly observed therapy and mouth checks, where 100 percent of patients were able to achieve undetectable viral loads presenting some of the strongest evidence that adherence is the key to successful treatment.

Patients who are likely to have poor adherence and therefore fail treatment may be excluded from treatment. These include IV drug users, mentally ill patients, or patients with a high risk of poor outcomes of whom

there have been relatively few studies of treatment success. Resources are often scarce in HIV clinics, and patients who are "high cost users" and "resource intensive" such as the above groups are the least likely to receive effective treatment. Therefore, the ability to provide adequate psychiatric care to HIV-infected patients is critical for effective treatment of HIV. Psychiatric disorders compromise the ability to take medications, adhere to treatment, practice safer sexual behaviors, and stop using IV drugs. While they are not the only factors driving the HIV epidemic, they have been largely overlooked as a critical factor in the continued spread of HIV.

Outside of the United States, financial issues have a greater impact on treatment decisions, and the ability to dramatically reduce perinatal transmission with even one drug has complicated public health approaches to HIV care. Also, the pharmaceutical companies have patents and attempt to recover research and development costs and profit on HIV medications, leading some governments to nationalize the manufacture of HIV drugs to reduce costs. In these settings, mental health issues may have an even greater impact on treatment selection decisions.

PSYCHIATRIC CONDITIONS IN THE HIV CLINIC

Delirium

Delirium is a state of global derangement of cerebral function. It occurs more frequently in medically ill, brain injured, or metabolically unstable patients. Prevalence of delirium in HIV-infected populations has been reported between 43 and 65 percent. The clinical presentation and significance of delirium in patients with HIV are the same as in non–HIV-infected individuals, and is characterized by inattention, disorganized thinking or confusion, and fluctuations in level of consciousness. Emotional changes are common and often unpredictable, and hallucinations and delusions are frequently seen. The syndrome has an acute or subacute onset, and remits fairly rapidly once the underlying etiology is treated. If untreated, patients have a marked increased risk of mortality, with estimates of about 20 percent mortality in hospitalized inpatients.

Aside from general risk factors such as older age, multiple medical problems, multiple medications, impaired visual acuity, and previous episodes of delirium, patients with HIV-associated dementia (HAD) are at increased risk to develop delirium. The differential diagnosis of delirium includes HAD, especially with AIDS mania, minor cognitive-motor disorder, major depression, bipolar disorder, panic disorder, and schizophrenia. Delirium can usually be differentiated from the above conditions based on its rapid onset, fluctuating level of consciousness, and link to a medical etiology.

The cause of delirium should be aggressively sought by intensive

medical examination. Vital signs and oxygen saturation, careful history and physical exam, laboratory tests, EKG, radiological exams, and critical review of all medications are essential to the workup. Possible etiologies include toxic (poisonings, drugs and new, recently changed or interacting medications, especially medications with potent anticholinergic activity), metabolic (electrolyte disturbances), infectious (central nervous system [CNS] infections or sepsis), endocrine (especially thyroid and adrenal axes), neoplastic (especially CNS), cardiovascular (myocardial infarction, arrhythmia), neurological (seizure, stroke), pulmonary (hypoxia, hypercapnia), and traumatic (head injury, burns) causes. In toxic/metabolic causes especially, the electroencephalogram may show diffuse slowing of the background alpha rhythm, which resolves as confusion clears.

Treatment consists of three parts. The first is the identification and removal of the underlying cause. The second is the reorientation of the patient by maintaining a normal diurnal variation of light cycles, providing orienting stimuli, such as calendars, clocks and a view of the outside world, and active engagement and reorientation by staff members, family, and friends. The third, if necessary, is the management of behavior or psychosis. Low doses of high-potency antipsychotic agents work well. Newer, atypical antipsychotics are currently being used with some success, but those with more anticholinergic activity may worsen the condition. Benzodiazepines should be used with caution, as they may contribute to delirium in some patients, but are of particular use in alcohol or benzodiazepine withdrawal deliria. Physical restraint may be necessary if the patient becomes violent, but all attempts should be made to use this at a minimum, not only because of the legal issues it poses but because restriction of movement or access to orienting stimuli worsens delirium.

HIV-ASSOCIATED DEMENTIA

Introduction and Neuropathology

Early in the AIDS epidemic, some patients presented with rapidly progressing neurocognitive disturbances leading to an intensive search for etiology. Several CNS opportunistic conditions were identified, including cytomegalovirus (CMV) encephalitis, progressive multifocal leukoencephalopathy (PML), cerebral toxoplasmosis, cryptococcal meningitis, and CNS lymphoma. However, a subset of patients remained for which no identifiable pathogen could be found and it was deduced that HIV itself was the causative factor behind the dementia. These patients had a subcortical dementia characterized by increasing loss of attention and concentration, marked motor slowing, and variable behavioral components generally leading to death in less than a year. Autopsy studies of demented patients with AIDS revealed characteristic white matter changes and

demyelinization, microglial nodules, multinucleated giant cells, and perivascular infiltrates, but a marked absence of HIV within neurons. This has led to the current theories of neuronal loss due to inflammation through action of macrophages and microglial cells and/or through activation of cytokines and chemokines that trigger abnormal neuronal pruning. It appears that basal ganglia and nigrostriatal structures are affected early in the dementia process, with diffuse neuronal losses following. Typical late findings show an approximate 40 percent reduction in frontal and temporal neurons. With the development of CART, there has been a dramatic reduction in HIV-induced dementia of the type seen earlier in the epidemic. It is now uncommon in populations in developed countries and less common even in underdeveloped settings. There is, however, still a high prevalence of cognitive impairment in HIV-infected patients, often more subtle and with an unpredictable course, and with some marked clinical differences.

Epidemiology, Risk Factors, and Clinical Course

With the development of CART, there was a hope that HIV dementia would become a thing of the past. Before CART, HIV dementia was presented as a rapidly progressive subcortical dementia that usually ended in death. In the era of effective antivirals, HIV-associated cognitive impairment is harder to assess, and there is considerable debate about the prevalence, causes, and methods of assessment. There is also no effective treatment if the patient is well controlled on CART. Although there is even debate about the terminology, most authors divide patients with HIV-associated neurocognitive disorder (HAND) by severity into three classes: asymptomatic neurocognitive impairment (ANI), mild neurocognitive disorder (MND), and HIV associated dementia (HAD) in order of increasing severity. One of the problems with assessment of these conditions is that accurate control groups are not well studied and much of the current prevalence data come from cohort studies. Although the prevalence of HAD has declined, cross-sectional studies show overall rates of HAND are still remarkably high. For example, in a clinically diverse cross-sectional study of 1,555 patients at six academic centers, the overall rate of cognitive impairment was 52 percent with 33 percent ANI, 12 percent MND, and only 2 percent HAD. Higher medical comorbidity resulted in increased HAND. Severity of disease such as current CD4 cell count and historical CD4 nadir as well opportunistic infections and detectable viral load and detectable virus in the CNS all appear to contribute to the risk.

The confounds identified that may explain persistent cognitive impairment include both damage from the virus, damage from immune activation, drug toxicity, and many others. Initially, identification of variable drug penetration into the CNS appeared to be an exciting and

remediable cause of continued viral injury. The expected finding that better penetrating antivirals would be associated with less HAND has not materialized, perhaps because of the confounds of poor correlation between CSF concentrations and those in the brain, direct mitochondrial toxicity of the drugs, degree of immune reactivation in patients with previous immunosuppression from HIV, and subtle ongoing viral replication in the CNS. There has also been the suggestion that the chronic immune activation associated with continued HIV infection, even with complete viral suppression, causes more rapid "aging" of patients and therefore cognitive impairments that are associated with older age groups. In fact, the aging of the entire population of patients with HIV is a confound.

HAND in the current era involves both cortical and subcortical loss. Prior to CART, patients had predominantly subcortical loss with greater impairments in motor skills, cognitive speed, and verbal fluency, whereas current patients tend to show more impairment in memory (learning) and executive function including deficits is prospective memory (the ability to "remember to remember") that has been shown to produce problems with medication adherence. Patients now have characteristics in common with cortical dementias such as Alzheimer disease. The course of HIV dementia prior to CART was clinical progression, while in the current longitudinal studies, some patients decline while others actually improve.

Treatment of HAND

The best treatment is aggressive treatment with CART in patients who are untreated or inadequately treated. Depression can be treated with standard antidepressants, and in some cases, methylphenidate (Ritalin) or other stimulants may be useful in the treatment of apathy. Safety assessments should be done when indicated and patients requiring intensive monitoring should be referred to appropriate facilities. To date, despite numerous trials, no clear effective treatment for HAND has emerged. There are patients who individually respond to interventions, sometimes dramatically, but clinical trials have not yet shown reliable results in populations.

MAJOR DEPRESSION IN PATIENTS WITH HIV DISEASE

Epidemiological Considerations

Depressive symptoms are the most common complication of chronic medical illness. Depression has a negative impact on adherence with medical treatments, quality of life, and treatment outcome. Despite education and efforts to improve treatment, depression remains underrecognized, underdiagnosed, and undertreated in medical clinics. Depression is a significant problem in HIV/AIDS. There are several long-

standing issues for patients with comorbid medical conditions and depression. Major depression as a disease state is different from the depressive symptoms in patients who are demoralized (psychologically distressed due to their circumstances), but most research uses checklists and mood questionnaires to measure depression, making it difficult to clearly assess prevalence of and associations with major depression. To add to the complexity, severe psychological stress may trigger episodes of major depression in vulnerable individuals, and the body of research suggests a role for both the stress axis and the immune axis in major depression. Also complicating this issue is the increased risk of getting infected with HIV due to behaviors associated with depression, making it difficult to determine whether the HIV or major depression came first. High rates of major depression have been found in homosexual men and patients with substance use disorders. A recent meta-analysis of 10 studies comparing HIV-positive and at-risk HIV-negative patients demonstrated a twofold increase in the prevalence of major depression in patients infected with HIV. Treatment of major depression has been shown to improve outcome of HIV treatment.

Major depression is a risk factor for HIV infection by virtue of its impact on behavior, intensification of substance abuse, exacerbation of self-destructive behaviors, and promotion of poor partner choice in relationships. In this way, depression can be seen as a vector of HIV transmission. Depression has been clearly shown to hinder effective treatment of infected individuals. Patients with major depression have been shown to be at increased risk of disease progression and mortality. HIV increases the risk of developing major depression through a variety of mechanisms, including direct injury to subcortical areas of brain, chronic stress, worsening social isolation, and intense demoralization. While direct evidence for a relationship between worsening HIV disease and the development of major depression is limited, there are several studies that support this link, particularly the study based in the Multicenter AIDS Cohort Study (MACS) showing that there is a two and a half fold increase in rates of depression as patient CD4 cells fall below 200 just before they develop AIDS.

Taken together, these lines of evidence suggest that HIV is a causal factor in depression, and that depression is a causal factor in HIV transmission and its morbidity, making the patients with these disorders a treatable vector for the HIV epidemic, and suggesting an important role for mental health care in HIV treatment and prevention.

Differential Diagnosis of Major Depression

The differential diagnosis of depression includes nonpathologic states of grief and mourning (sometimes made quite severe by the vulnerabilities of the person) and a variety of disorders related to both psychological and

physiological disturbances. Patients with complaints of depressive syndromes can have dysthymia, dementia, delirium, demoralization, intoxication, withdrawal, CNS injury or infection, acute medical illness, and a variety of other conditions. AIDS dementia and other HIV-related CNS conditions can produce a flat, apathetic state that is often misdiagnosed as depression. Cocaine withdrawal produces a depressive syndrome, and delirium can mimic many psychiatric conditions. CNS syphilis, a condition that had become quite rare, has been reappearing in medical centers with HIV specialty services and remains "the great imitator" as it was called when it was originally described.

The diagnosis of major depression in the HIV clinic is complicated by the high frequency of depressive symptoms that are associated with chronic illness, significant losses and isolation, and complex medical treatments. The diagnosis may be further complicated by the presence of comorbid neurological illness, comorbid substance use, and the use of many medications that can alter mental function. Despite this, studies of patients with major depression have shown that the response to treatment is similar to that expected in other populations.

Mood is the feature of affective life that describes the prevailing emotional tone a patient experiences. Patients express depression as feeling blue, low, sad, or sometimes flat, devoid of emotion, or empty. Some say that they are miserable, in agony, anxious, worried, or angry, upset, and irritable. Although some are apathetic, most find the state extremely uncomfortable. Moods may vary with the time of day, usually worse in the morning and better in the afternoon, but sometimes in other patterns. Pleasure (or a sense of reward) is associated with many behaviors of life, including those driven by appetite like sleeping, eating, and sex, and those associated with function, such as work, hobbies, dress, social activity, and artistic expression. This is suppressed or absent in depression, and is a cardinal feature. Hedonic responsiveness refers to the enjoyment or positive reinforcement related to activities and experiences of life. Anhedonia describes the state in which hedonic responsiveness is blunted or absent, and is a principle feature of depression.

Vital sense is the subjective sense of well-being. Depressed patients say they are sick, have a heavy pressure in their chest, have low energy, and even feel themselves dying. Self-attitude is the feelings directed at self. Depressed patients feel guilty, that they are a bad or evil, that they are undeserving of things they have, or that they have failed those they love.

Neurovegetative symptoms are commonly associated with this cardinal syndrome. These include difficulties with sleep, appetite, concentration and memory. Sleep disturbances can include insomnia or hypersomnia, but most characteristically patients describe waking early in the morning with difficulty falling back asleep. Appetite may be decreased or increased significantly, and patients often complain that food has lost its flavor.

Appetite changes result in corresponding weight changes if untreated for long periods of time. Patients report slowed thought processes, with impairments in concentration and short-term memory, and occasionally generalized confusion.

Patients with HIV suffering from major depression frequently present to internists and family practitioners with multiple somatic symptoms. These include, but are not limited to, headache, gastrointestinal (GI) disturbances, inexplicable musculoskeletal or visceral pain, cardiac symptoms, dizziness, tinnitus, weakness, and anesthesia. Given the burdens of HIV, the medical problems associated with it, and the side effects of medications, depression may be very low on the list of considered causes of the patient's complaint. Even patients complaining of depressive symptoms may have depression overlooked or discounted due to the presence of a plethora of other diagnoses. Nonspecific somatic symptoms are often the result of depression rather than HIV infection in patients who do not have concurrent medical illness.

As an example, fatigue has been found to be associated with depression and not HIV disease progression. Worsening of fatigue and insomnia at 6-month follow-up was highly correlated with worsening depression but not CD4 count, change in CD4 count, or disease progression by CDC category. These findings support the notion that somatic symptoms generally suggestive of depression should trigger a full psychiatric evaluation. In later stage HIV infection, a variety of illnesses are common, often moving depression down on the differential diagnosis list. Somatic symptoms should always be evaluated carefully, and considered in context, that is, either with other indicators of progression of HIV disease or with other indicators of depression.

Certain HIV-related medical conditions and medications can cause depressive symptoms. These include CNS disorders such as toxoplasmosis, cryptococcal meningitis, lymphoma, syphilis, and others. Some investigators have found significant rates of depressive symptoms among male patients with HIV with low testosterone levels and male patients with hypogonadism and major depression often have treatment refractory depressive episodes until the testosterone is repleted. Of the antiretroviral agents, efavirenz is most often associated with depressive symptoms, which may remit spontaneously in a number of patients but persist in a few. More recently, raltegravir has been implicated in depression as well. Efavirenz has also been reported to cause frank psychosis in some patients, necessitating its withdrawal in these cases. Patients receiving interferon are at risk of depression from this agent, as are patients receiving chronic steroids (both anabolic and corticosteroids) for comorbid medical problems. Because of high coinfection rates, hepatitis C treatment with interferon containing regimens often complicates HIV-related depression, but due to the development of hepatitis C directed antiviral agents,

interferon use for hepatitis C is rapidly becoming an historical problem. A number of other agents, including metoclopramide, clonidine, propranolol, sulfonamides, muscle-relaxants and many others, have all been reported to produce major depression or similar syndromes. While these depressive syndromes often respond to withdrawal of the offending drug, when they do not, they should be treated as major depression with appropriate antidepressant medication. The experience with the depressive syndromes caused by efavirenz and interferon indicates that treatment with antidepressant medications is well tolerated and often necessary for success.

Routine screening for psychiatric disease in HIV clinic patients can effectively preempt urgent referrals for rapidly progressing disorders. Several screening tools for depression in medical settings have been studied, and in one study specifically looking at HIV-infected patients, the combination of two brief self-administered questionnaires, the BDI and the General Health Questionnaire (GHQ), prospectively predicted a psychiatric disorder (other than substance abuse) with a sensitivity of 81 percent, specificity of 61 percent, and a positive predictive value of 71 percent.

Treatment of Major Depression

Pharmacotherapy is the mainstay of treatment for major depression. Several studies have demonstrated efficacy of various antidepressant agents in patients with HIV, but no single antidepressant has been found superior in treating HIV-infected patients as a group. In a recent study examining the overall efficacy of antidepressants in patients with major depression and HIV, technical variations among studies were found to account for the largest portion of the variance of results, and overall the use of antidepressants was supported.

Aside from how well the pharmacology of the antidepressant matches a patient's disease, the engine that drives effectiveness is patient adherence. Patients who reliably take adequate doses of antidepressants have the best chance of improvement. A general rule is to start at low doses of any medication, titrating up to a "full" dose or therapeutic serum level (when meaningful) slowly, so as to minimize early side effects that may act as obstacles to adherence.

Although some papers have suggested that SSRI antidepressants are the "first-line treatment," the data do not show any advantage of one antidepressant over another, and most recommend using a medication that best fits the other elements of the case. Table 2.8–2 lists antidepressants often used in HIV disease, their dosages, and common side effects. In many patients, side-effect profiles of certain agents may be used to an advantage, and these "good match" agents should be tried first. Second choices are medications that do not have side effects that might worsen the patient's symptoms (e.g., a medication that is not associated with insomnia for a

patient having difficulty falling asleep). Later choices include medications that have side-effect profiles that would likely exacerbate a patient's symptoms (e.g., a medication associated with increased GI motility for a patient with chronic diarrhea).

The first week of treatment with a medication usually determines whether a patient will be able to tolerate the medication at all. Following this brief period, the dosage of any antidepressant should be increased slowly to either a typical full dose or a therapeutic serum level. Once at this level, patients should be encouraged to wait as long as possible for the therapeutic effect, which may take longer than 6 weeks to achieve in some patients. Side effects should be assessed at every visit, and attempts should be made to treat any side effects the patient finds bothersome. For example, insomnia caused by SSRIs may respond well to low doses of trazodone (25 to 150 mg at bedtime). Constipation from a TCA often is relieved by increased water and fiber intake. Sexual side effects from SSRIs are common; erectile difficulty may be treated with nitric oxide synthase inhibitors (NOS inhibitors include sildenafil, tadalafil, and vardenafil) in select patients, while decreased libido and delayed orgasm may be alleviated with a drug holiday or by the addition of bupropion.

Patients who show only partial response to antidepressant medication should be offered an augmentation strategy or an alternative treatment. There are several other agents that have been used for antidepressant augmentation. The best studied is lithium, but its side-effect profile often prevents its use in the HIV setting. Thyroid preparations, especially triiodothyronine, have also been shown to be of benefit, and may be of particular advantage in patients complaining of fatigue. Atypical neuroleptics and pindolol have also been reported to be effective augmenting agents, although many others have also been used without support from the literature, including a second antidepressant, other mood stabilizers, trazodone, other antipsychotic medications, methylphenidate, benzodiazepines, sleep deprivation, and bright light therapy.

If no benefit is gained from the primary antidepressant, even after augmentation, or if it must be abandoned due to intolerable side effects, a new primary agent should be chosen, titrated slowly, and augmented as necessary. Although medications in the same class may produce similar side effects, and therefore may not be tolerated, there is evidence to suggest that a response may be seen from one drug where none was seen from another in the same class.

Finally, clinicians often wonder about the interaction of antidepressant medications and HAART medications. Some potential interactions are listed in Table 2.8–2, but two points deserve emphasis. First, since depression is associated with reductions in adherence to HAART, untreated depression may be equally or more detrimental to disease progression than any medication interaction. Second, experience in working with comorbid

HIV and depression has not yet shown clinical significance to the drug–drug interaction, that is, need for dose adjustments for either antidepressants or HAART for successful outcomes.

Psychotherapeutic Treatment. Psychotherapy is an important and integral part of the treatment of major depression. Treatment with medication plus psychotherapy has been shown to be more effective for patients than either modality alone. A major open question continues to be which type of psychotherapy is most appropriate to provide. Among the individual psychotherapies, interpersonal psychotherapy and cognitive-behavioral psychotherapy are quite popular for the treatment of depression and have the best evidence to support their efficacy.

The literature on the use of psychotherapy for the treatment of depression in patients with HIV is extensive, but clinical trials data are sparse. One study showed that imipramine with either interpersonal or supportive psychotherapy had better efficacy than those therapies used alone. Group cognitive-behavioral therapy has also demonstrated efficacy for patients with HIV used alone or in combination with medication.

Improvements have been demonstrated as well for patients with HIV treated with group cognitive-behavioral therapy either as a single treatment modality or combined with medication.

Supportive psychotherapy helps patients with major depression who interpret their suffering to be a reaction to the diagnosis or morbidity of HIV infection. These patients often believe they can pull themselves out of depression and get frustrated when they continue to expend effort with little result. They need education about the disease nature of their depression, encouragement to keep going, and therapeutic optimism that the treatments will work.

In addition, psychotherapy applied judiciously and in combination with effective antidepressant medication provides patients with a framework for the provider–patient relationship that is so crucial to success. The medical providers who keep the concept of psychotherapy in mind will structure their interactions with patients to slowly empower and enable the patients to take control of their lives, thus relying on the providers less and less.

BIPOLAR (MANIC-DEPRESSIVE) ILLNESS IN PATIENTS WITH HIV DISEASE

Bipolar disorder is a condition in which patients experience episodic alterations in mood that cause disorder. Manic episodes are associated with increased rates of substance abuse and impulsive behavior, and there has been speculation that bipolar disorder may be a risk factor for HIV infection. To date there has been no unequivocal evidence to show that bipolar illness directly increases the risk of HIV infection, but the technical difficulties in demonstrating this link are considerable.

In the classic presentation, patients alternate between extended

episodes of depression similar to major depression and briefer episodes of increased mood, increased energy, and increased confidence and well-being, often with grandiose ideas about themselves and their circumstances. The synonymous appellation "manic-depressive insanity" is a reminder that many patients suffer from auditory hallucinations and frank delusions when they are ill. Most often, patients cycle from one type of mood state to the other, at times interspersed with periods of normal mood, but occasionally show features of both depressive and elevated mood states simultaneously (mixed states) or in very rapid succession (rapid cycling). Milder forms of mania are seen in a condition termed Bipolar Illness Type II. The spectrum of bipolar illness is broad, ranging from a severely crippling and chronic mental illness to a mildly disordering alternating experience of prevailing mood. This has made it difficult to accurately measure prevalence and incidence, and to explore the relationship of bipolar Illness with HIV disease. Additionally, it is difficult to distinguish severe bipolar illness from schizophrenia even in studies that utilize rigorous research approaches. Thus, investigators looking at the relationships between HIV and mental illness often will use the term "chronically mentally ill" for patients with severe disability from either schizophrenia or bipolar disease.

Table 2.8–2.
Use of Antidepressants in HIV Disease, Their Dosages, and Common Side Effects

Drug	Start Dose	Usual Therapeutic Dose	Serum Level	Advantages	Interactions with HIV medicine
Nortriptyline (Pamelor)	10–25 mg q hs	50–150 mg q hs	100–150 ng/dL	Promotes sleep, weight gain, decreases diarrhea	Increases nortriptyline levels, flu lopinavir/ritonavir, ritonavir
Desipramine (Norpramin)	10–25 mg q hs	50–200 mg q hs	>125 ng/dL	Promotes sleep, weight gain, decreases diarrhea	Increases desipramine levels, lo ritonavir
Imipramine (Tofranil)	10–25 mg q hs	100–300 mg q hs	>225 ng/dL	Promotes sleep, weight gain, decreases diarrhea	Increases imipramine levels, lop ritonavir
Amitriptyline (Elavil)	10–25 mg q hs	100–300 mg q hs	200–250 ng/dL	Promotes sleep, weight gain, decreases diarrhea	Increases amitriptyline levels Lopinavir/ritonavir, ritonavir
Clomipramine (Anafranil)	25 mg q hs	100–200 mg q hs	150–400 ng/dL	Promotes sleep, weight gain, decreases diarrhea	Increases clomipramine levels, lopinavir/ritonavir, ritonavir

Doxepin (Sinequan)	10–25 mg q hs	150–250 mg q hs	100–250 ng/dL	Promotes sleep, weight gain, decreases diarrhea	Increases doxepin levels Lopinavir/ritonavir, ritonavir
Fluoxetine (Prozac)	10 mg q am	20 mg q am	Unclear	Activating	Increases HIV med levels Amprenavir, delarvidine, efavire lopinavir/ritonavir, nelfinavir, saquinavir Decreases fluoxetine levels
Sertraline (Zoloft)	25–50 mg q am	50–200 mg q am	Unclear		Nevirapine Increases sertraline levels Lopinavir/ritonavir, ritonavir
Citalopram (Celexa)	20 mg q am	20–60 mg q am	Unclear		Increases citalopram levels Lopinavir/ritonavir, ritonavir
Paroxetine (Paxil)	10 mg q hs	20–40 mg q hs	Unclear	Somewhat sedating	Increases paroxetine levels Lopinavir/ritonavir, ritonavir
Fluvoxamine (Luvox)	50 mg q hs	150–250 mg q hs	Unclear	Somewhat sedating	Increases HIV med levels Amprenavir, delarvidine, efavire lopinavir/ritonavir, Nelfinavir, ritonavir, saquinavir Decreases fluvoxamine levels Nevirapine
Escitalopram (Lexapro)	10 mg q am	10–30 mg q am	Unclear	Nausea, akathisia	Unknown
Venlafaxine XR (Effexor)	37.5 mg q am	75–300 mg q am	Unclear	Activating, risk of hypertension	Lopinavir/ritonavir, ritonavir inc venlafaxine levels Venlafaxine may decrease indin
Desvenlafaxine (Prestiq)	50 mg	50–100 mg q am	Unclear	Nausea, dizziness, vomiting	Not yet clear
Duloxetine (Cymbalta)	20 mg q am	60–120 mg q am	Unclear	Nausea, akathisia, sedation	Unknown
Mirtazapine (Remeron)	7.5–15 mg q hs	15–45 mg q hs	Unclear	Promotes sleep, weight gain	Unknown
Nefazodone (Serzone)	50 mg BID	300–400 mg/day in divided doses	Unclear	Somewhat sedating, risk of hepatic complications	Increases HIV med levels Efavirenz, Indinavir
Vortioxetine (Brintellix)	5 mg q day		Unclear		When used with maraviroc the dosage should be reduced to daily.
Vilazodone (Viibryd)	10 mg	20–40 mg a day 40 mg		Nausea, constipation, vomiting Diarrhea, nausea, vomiting, insomnia	Unknown Unknown
Trazodone (Desyrel)	50–100	50–150 mg q hs for	Unclear	Promotes sleep	Increases trazodone levels Lopinavir/ritonavir, ritonavir

	mg q hs	sleep; 200–600 mg q hs for depression			
Bupropion SR or XL (Wellbutrin)	100 mg q am	150–400 mg/day in divided doses	Unclear	Activating, no sexual side effects	Efavirenz reduces bupropion lev Lopinavir/ritonavir has been de reduce bupropion levels 57% Tipranavir/ritonavir reduces l 46% Elvitegravir/cobicistat/tenofovir may increase bupropion leve

The elevated mood states form a continuum from increased energy, euphoria, irritability and decreased sleep called hypomania to a more extreme condition complicated by hallucinations, delusions, disordered thinking, disorganized, and sometimes violently agitated behavior called mania. Hypomania is characterized by euphoria, an improved self-attitude and an elevated vital sense. Patients feel elated, energized and as if they are functioning better than usual. Thoughts speed up and horizons expand, such that the patient feels many brilliant things are coming to her in rapid succession. Often, there is a noticeable increase in the amount and speed of speech; interrupting these patients may be necessary and hard to accomplish. Since energy is so high, patients feel a decreased need for sleep, and occasionally do not sleep at all. When these symptoms impair judgment and function is lost, the patient is seen to be further down the spectrum, in the syndrome of mania. Manic patients not only have pressured speech, they often demonstrate a thought disorder in which ideas come so quickly that it is impossible to see the connections between them, the so-called flight of ideas. The expansive self-attitude may take on proportions outside the realm of reality, known as grandiose delusions. Paranoid delusional thoughts may be seen, and hallucinatory experiences occur in some patients.

In contrast to this “garden variety” type of bipolar disorder, there is a type of mania that appears to be specifically associated with late stage HIV infection and is associated with cognitive impairment and a lack of previous episodes or family history. This syndrome is called AIDS mania and may represent a related but different condition. Studies of this form of mania are less common as HAART has had a significant impact on both the frequency of AIDS dementia and mania.

Mania can occur anytime in the course of HIV infection for individuals with pre-existing Bipolar disorder, but AIDS mania has been described in late HIV infection, thus appearing to be a consequence of HIV brain involvement. In general, manic syndromes in patients with HIV occur with higher frequencies after the onset of AIDS. Furthermore, patients with AIDS develop mania at rates substantially greater than the general

population: in one series, mania occurred in 8 percent of all patients with AIDS seen at the HIV clinic over 17 months (more than 10 times the 6 months general population prevalence). The study grouped mania patients into those whose first manic episode came late in their HIV course with CD4 count <200 and those whose episode came early with CD4 count >200. The late-onset patients were less likely to have a personal or a family history of mania or any mood disorder, which presumably means they were less likely to have bipolar disorder or a genetic predisposition to mania. They were also more likely to have dementia or other cognitive impairment indicating brain damage.

AIDS mania seems to have a somewhat different clinical profile than bipolar mania. Patients tend to have cognitive slowing or dementia. Although without a previous dementia diagnosis this may be difficult to ascertain in the midst of an acute manic episode, the history will usually reveal progressive cognitive decline prior to the onset of mania. Irritable mood is more characteristic than euphoria. Sometimes prominent psychomotor slowing accompanying the cognitive slowing of AIDS dementia will replace the expected hyperactivity of mania, which complicates the differential diagnosis. Clinical experience has suggested that AIDS mania is usually quite severe in its presentation and malignant in its course. In one series, late-onset patients had a greater total number of manic symptoms than early-onset patients. They were also more commonly irritable and less commonly hypertalkative. AIDS mania seems to be more characteristically chronic than episodic, has infrequent spontaneous remissions, and usually relapses with cessation of treatment. Because of their cognitive deficits, patients have little functional reserve to begin with. Also they are less able to pursue treatment independently or consistently.

One clinically described presentation of mania, either early or late, is the delusional belief that one has discovered a cure for HIV, or has been cured. While this may serve to cheer otherwise demoralized and depressed patients, it may also result in the resumption of high-risk behavior and lead to the spread of HIV and exposure to other infectious entities. When euphoria is a prominent symptom in otherwise debilitated late-stage patients, caregivers may wistfully question the humaneness of robbing patients of the illusion of happiness. It is the clearly impairing, often devastating effects of the other symptoms of mania that tips the balance of the risk/benefit equation toward treatment.

The treatment of mania in early-stage HIV infection is not substantially different than the standard treatment of bipolar disorder. It relies on the use of mood-stabilizing medications, particularly lithium salts and the anticonvulsants valproic acid, lamotrigine and carbamazepine, and antipsychotic agents, now more commonly atypical agents. These medications decrease manic symptoms and may prevent recurrence.

Although seldom seen in developed countries now, AIDS mania in

advanced HIV disease with low CD4 counts, more medical illnesses, more CNS involvement, and greater overall physiological vulnerability, may require different treatment strategies. While treatment with traditional antimanic agents may be preferred, it can be very difficult in patients with advanced disease. AIDS mania patients typically respond to treatment with antipsychotic agents alone. In general, late-stage patients are far more sensitive to the therapeutic effects, but even more so to the toxic side effects, of antipsychotic agents. In late stage disease, the dose of antipsychotic needed may be much lower than customarily used for mania in other settings. The more advanced the patients' HIV and/or dementia, the more sensitive they are to dosage changes that might otherwise seem trivial. These patients can develop extrapyramidal symptoms, but will also prove very sensitive to the side effects, especially delirium, of anticholinergic agents.

In recent years, the newer atypical antipsychotics have often taken the place of the older agents. These agents have fewer extrapyramidal side effects than traditional antipsychotics, but have fewer data and less experience to support their use. The decreased likelihood of motor side effects of dystonia and tardive dyskinesia has made the use of atypical antipsychotics the first choice in most situations, and may be primary treatment in most advanced cases. Starting doses should be low and titrated to effectiveness.

There has been considerable experience with traditional mood stabilizing agents in selected AIDS mania patients, but with relatively sparse documentation. Lithium use has been problematic for several reasons, including high rates of associated delirium and cognitive difficulty, GI symptoms including nausea and diarrhea, and polyuria resulting in dehydration. Lithium is also associated with the development of diabetes insipidus in rare patients and may be difficult to use in combination with tenofovir, an nucleoside reverse transcriptase inhibitor (NRTI) with known renal side effects. The major problem with lithium in patients with AIDS has been rapid fluctuations in blood level, occurring even in the hospital on previously stable doses, causing lithium intoxication. Valproic acid has been used with success, titrating to the usual therapeutic serum levels of 50 to 100 ng/dL. Enteric-coated Depakote is better tolerated in most patients. This is sometimes limited by side effects, especially hepatotoxicity in the setting of chronic viral hepatitis. Monitoring of liver function tests is essential, but hepatic toxicity is not often a problem. In cases of severe hepatic *Mycobacterium avium* complex (MAC) infiltration, for example, with portal hypertension, valproic acid should likely be avoided, but this and related considerations have not been formally studied. Depakote can also affect hematopoietic function so white blood cell and platelet counts must be monitored. Carbamazepine may also be effective but more poorly tolerated because of sedation, and because of the presumed potential for

synergistic bone marrow suppression in combination with antiviral medications and HIV itself. Lamotrigine is often used to successfully treat bipolar disorder and may have special promise in treating patients with prominent depressive episodes. There are no studies showing an increased incidence of rash or Stevens–Johnson syndrome in patients with HIV receiving lamotrigine, but prescribers should monitor for these adverse events closely, as with any patient. Other agents such as gabapentin, oxcarbazepine, and topiramate have been tried, but the reports of success remain anecdotal.

SCHIZOPHRENIA IN PATIENTS WITH HIV DISEASE

The literature on patients with severe and chronic mental illnesses, accounted for primarily by schizophrenia and bipolar I disorder, reports prevalence rates of between 4 and 19 percent in both inpatient and outpatient samples. There is no evidence that HIV infection causes schizophrenia, but there are data to show that schizophrenia contributes to behaviors that may lead to HIV infection. Although injection drug use accounts for the majority of infections in many studies, there has also been a wealth of information published regarding the sexual risk factors for patients with schizophrenia. In particular, data reveal high rates of unprotected sex, multiple sex partners, trading sex for money or other goods, and sex while intoxicated. Further, there is evidence that patients with more positive symptoms and impulse control problems are at increased risk of high-risk sexual behavior despite demonstration of adequate knowledge of HIV risk factors.

Practitioners that see patients with schizophrenia should be sensitive to the risk of acquiring HIV and should screen patients carefully for risk behaviors in addition to inquiring about patients' knowledge of HIV transmission routes. A screening tool called the Risk Behaviors Questionnaire (RBQ) consists of 13 questions and has been validated for use in psychiatric patients.

The principles of treatment for HIV-infected patients with schizophrenia follow the same basic principles as any other patient with schizophrenia, namely control of symptoms with medications and psychosocial support and rehabilitation. In these cases, however, close ties with HIV providers are strongly suggested, so that HIV treatment can be coordinated and monitored.

In a recent survey, HIV care providers were just as likely to recommend antiretroviral therapy for patients with schizophrenia who met criteria as those without schizophrenia, but to recommend avoiding efavirenz-based regimens due to higher risk of neuropsychiatric side effects. Further, these practitioners were more likely to seek collaboration with a mental health care provider to coordinate treatment for patients with schizophrenia, suggesting that this population is recognized as outside the scope of

practice for solo management by an HIV primary care provider.

ISSUES OF PERSONALITY IN PATIENTS INFECTED WITH HIV

A disturbing trend in the HIV epidemic has been the persistence of high-risk behaviors among individuals who are HIV infected. Such individuals, who report high rates of sex and/or drug risk behaviors, include HIV-infected drug users, patients presenting at HIV primary care clinics for medical treatment, and HIV-infected men who have sex with other men. Apparently, knowledge of HIV and its transmission is insufficient to deter these individuals from engaging in HIV risk behaviors, suggesting that certain personality characteristics may enhance their vulnerability to practice such behaviors.

Traditional approaches in risk reduction counseling emphasize the avoidance of negative consequences in the future, such as using a condom during sexual intercourse to prevent sexually transmitted diseases (STDs). Such educational approaches have proved ineffective for individuals with certain personality characteristics. Most theoretical models of HIV risk behavior have not considered the role of personality factors, and few studies have examined mechanisms accounting for dispositional influences on sexual risk taking. Effective prevention and treatment programs for HIV-infected individuals must consider specific personality factors that render them vulnerable to practicing risky behaviors that further endanger their health as well as the health of others. This section will:

1. Discuss extraversion/introversion and stability/instability, measurable traits of temperament that drive personality and behavior;
2. Outline the role of personality characteristics and personality disorder in HIV risk behavior; and
3. Highlight specific interventions to reduce HIV risk behaviors that are formulated for individuals whose personality characteristics place them at increased risk.

Dimensional Traits and Personality

Personality is defined by the emotional and behavioral characteristics or traits that constitute stable and predictable ways that an individual relates to, perceives, and thinks about the environment and the self. Personality emerges as the underlying disposition or temperament of an organism is shaped and affected by development and environment. Traits that may appear to be positive or negative are in fact adaptive in one setting and maladaptive in another. Individuals vary in the degree to which they possess a given trait and in the way it influences their behavior. When traits found in certain individuals exceed the levels found in most of society, and are sufficiently rigid and maladaptive to cause subjective distress or functional impairment, a *personality disorder* is usually diagnosed.

Most personality models depict individuals along temperamental dimensions of extraversion–introversion and stability–instability. The dimension of extraversion–introversion refers to the individual's basic tendency to respond to stimuli with either excitation or inhibition. Individuals who are *extraverted* are (1) present-oriented, (2) feeling-directed, and (3) reward-seeking. Their chief focus is their immediate and emotional experience. Feelings dominate thoughts, and the primary motivation is immediate gratification or relief from discomfort. Extraverts are sociable, crave excitement, take risks, and act impulsively. They tend to be carefree, inconsistent, and optimistic. By contrast, *introverted* individuals are (1) future and past-oriented, (2) cognition-directed, and (3) consequence-avoidant. Logic and function predominate over feelings. Introverts are motivated by appraisal of past experience and avoidance of future adverse consequences. They will not engage in a pleasurable activity if it might pose a threat in the future. Introverted individuals are quiet, dislike excitement, and distrust the impulse of the moment. They tend to be orderly, reliable, and rather pessimistic. The second personality dimension, stability–instability, defines the degree of emotionality or lability. The emotions of *stable* individuals are aroused slowly and minimally and return quickly to baseline. By contrast, *unstable* individuals have intense, mercurial emotions that are easily aroused and return slowly to baseline. If these two personality dimensions are juxtaposed, four personality types emerge as described from the time of the ancient Greeks: phlegmatic (stable introverts), sanguine (stable extraverts), melancholic (unstable introverts), and choleric (unstable extraverts).

Traditional Personality Models, Their Instruments, and Findings Related to HIV Risk Behaviors

After the higher order traits of extraversion and neuroticism, personality models have elaborated other trait domains and developed personality inventories to empirically assess these traits. The major models and their inventories are as follows:

1. The three factor model (Eysenck & Eysenck) Extraversion, Neuroticism, and Psychoticism, measured by the Eysenck Personality Questionnaire (EPQ).
2. The five factor model (Costa & McCrae) Neuroticism, Extraversion, Openness, Agreeableness, Conscientiousness, measured by the NEO.
3. The alternative five factor model (Zuckerman) Neuroticism–Anxiety, Sociability, Impulsive Sensation Seeking, Aggression–Hostility, Activity, measured by the Zuckerman Kuhlman Personality Questionnaire (ZKPQ).
4. The seven-dimensional model (Cloninger) Novelty Seeking, Harm Avoidance, Reward Dependence, Persistence, Self-Directedness,

Cooperativeness, and Self-Transcendence, measured by the Tridimensional Personality Questionnaire (TPQ).

Personality traits appear to influence a variety of sexual risk behaviors, yet there is relatively little research on sexual risk taking from the major personality models. On the EPQ, Extraversion is associated with sexual promiscuity, desire for sexual novelty, multiple sex partners, and in a quantitative review of overall sexual risk taking, and shows a modest effect size. Neuroticism is related to unprotected anal sex. Psychoticism is associated with number of sexual partners and unprotected sex in several studies. On the NEO, Neuroticism is associated with unprotected sex and to a lesser extent, sex with multiple partners. Low conscientiousness is also associated with unprotected sex. Low openness to experience is associated with the denial of risk of HIV infection. The TPQ has minimal research on sexual risk taking, but one study shows that Novelty Seeking is associated with unprotected sex. The Impulsive Sensation Seeking scale of the ZTPQ has received the most research attention and predicts number of sex partners, unprotected sex, and high-risk sex encounters, such as sex with a stranger, across a variety of populations.

Research on personality traits links Extraversion and Neuroticism to drug and alcohol addiction. On the EPQ, Psychoticism and, to a lesser extent, Neuroticism has been linked prospectively to alcohol dependence in a 6-year study. On the ZTPQ, Impulsive sensation seeking is consistently associated with addiction severity as well as amount and variety of illegal drug use. While there is no specific “alcoholic” or “drug-using” personality, there appears to be a modest link between substance abuse and either impulsivity/high novelty seeking or high on Neuroticism/negative emotionality. Individuals with *both* these traits may be at the greatest risk of addiction.

Sensation Seeking. A significant empirical contribution to the understanding of the role of personality factors and HIV risk is the conceptual model connecting sexual sensation seeking, alcohol expectancies, and drinking before sex as key predictors of risk.

In a series of studies, Kalichman et al. adapted Zuckerman Sensation Seeking Scale that measures preference for exciting, optimal, and novel levels of stimulation or arousal. Sensation seeking functions as the “third variable” connecting alcohol use to sexual risk behaviors. Using path analysis in a series of experiments, Kalichman et al. tested a model that predicts the association among sensation seeking, alcohol use expectancies, alcohol use, and sexual risk behavior in both men at risk of HIV and HIV positive men. Sensation seeking is associated with alcohol outcome expectancies (or the beliefs that the individual has about the effects of alcohol on experience or behaviors). Having positive expectancies about the effect of alcohol on sexual pleasure or sexual behavior increases the

likelihood that alcohol will be used in sexual situations. In turn, having sex when under the influence of alcohol is associated with an increased likelihood of having unprotected sex. The importance of this model is identifying a marker, sensation seeking, for multiple risk practices as well as identifying alcohol expectancies that can be a point of intervention for prevention and treatment. Recently this model has been corroborated in heterosexual women and men.

Implications for HIV Risk Behavior: Clinical Observations

There has been relatively little empirical investigation of the influence of personality characteristics on HIV risk behavior; however, the authors' clinical observation suggests that of the four temperaments, unstable extraverts are the most prone to engage in HIV risk behavior. The authors estimate that in Psychiatry Service of the Johns Hopkins AIDS Service, about 60 percent of the authors' patients present with this blend of extraversion and emotional instability (unpublished observation). These individuals are preoccupied by, and act upon, their feelings, which are evanescent and changeable. Thus, their actions tend to be unpredictable and inconsistent. Most striking is the inconsistency found between thought and action. Regardless of intellectual ability or knowledge of HIV, unstable extraverts can engage in behavior associated with extreme risk of HIV infection. Past experience and future consequences have little salience in decision making for the individual who is ruled by feeling; the present is paramount. Their overarching goal is to achieve immediate pleasure or removal of pain, regardless of circumstances. Furthermore, as part of their emotional instability, they experience intense fluctuations in their mood. It is difficult for them to tolerate painful affect, such as boredom, sadness, or unresolved drive; they want to escape or avoid such feelings as quickly and easily as possible. Thus, they are motivated to pursue pleasurable experiences, however, risky, and eliminate low moods.

Unstable extraverts are more likely to engage in behavior that places them at risk of HIV infection. They are less likely to plan ahead and carry condoms and more likely to have unprotected vaginal or anal sex. They are more fixed upon the reward of sex, and remarkably inattentive to the STD they may acquire if they do not use a condom. Unstable extraverts are also less likely to accept the diminution of pleasure associated with the use of condoms or, once aroused, to interrupt the "heat of the moment" to use condoms. Similarly, unstable extraverts are more vulnerable to alcohol and drug abuse. They are drawn to alcohol and drugs as a quick route to pleasure. They are more likely to experiment with different kinds of drugs and to use greater quantities. Unstable extraverts are more vulnerable to alcohol and drug abuse. They are drawn to alcohol and drugs as a quick route to pleasure. They are more likely to experiment with different kinds of drugs and to use greater quantities. Unstable extraverts are also more

likely to become injection drug users because the experience is more intense. They are also less likely to defer this intensity in the interest of safety.

The second most common personality type that the authors have observed, which may represent about 25 percent of JHAS patients referred to psychiatry, is that of the *stable* extravert. Stable extraverts are also present-oriented and pleasure seeking; however, their emotions are not as intense, as easily provoked, or mercurial. Hence, they are not as strongly driven to achieve pleasure. Their emotional imperturbability (described by many as sanguine) may generate a kind of indifference to HIV risk more than a drive to seek pleasure at any cost. Stable extraverts may be at risk because they are too optimistic or sanguine to believe that they will become HIV-infected.

Introverted personalities appear to be less common among the authors' psychiatric patients. Their focus on the future, avoidance of negative consequences, and preference for cognition over feeling render them more likely to engage in protective and preventive behaviors. HIV risk of introverts is determined by the dimension of emotional instability–stability. About 14 percent of the authors' patients present with a blend of introversion and instability. Unstable introverts are anxious, moody, and pessimistic. Typically these patients seek drugs and/or sex not for *pleasure,* but for relief or distraction from pain. They are concerned about the future and adverse outcomes, but believe that they have little control over their fates. Stable introverts comprise the remaining 1 percent of patients. These patients with their controlled, even-tempered personalities are least likely to engage in risky or hedonistic behaviors.

Personality Disorder in HIV

Personality disorders represent extremes of normal personality characteristics and are disabling conditions. Prevalence rates of personality disorders among HIV infected (19 to 36 percent) and HIV at-risk (15 to 20 percent) individuals are high and significantly exceed rates found in the general population (10 percent). Antisocial personality disorder (ASPD) is the most common and is a risk factor for HIV infection. Individuals with personality disorder, particularly ASPD, have high rates of substance abuse and are more likely to inject drugs and share needles compared to those without an Axis II diagnosis. Approximately, half of drug abusers meet criteria for a diagnosis of ASPD. Individuals with ASPD are also more likely to have higher numbers of lifetime sexual partners, engage in unprotected anal sex, and contract STDs compared to individuals without ASPD.

Clinically, it has been useful to characterize patients along the dimensions of extraversion/ introversion and emotional stability/instability rather than in the discrete categories such as narcissistic, histrionic, and antisocial. This approach is useful for several

reasons. First, it is easier and quicker for all staff to determine where a patient falls along two dimensions than to evaluate, for example, each of the criteria to make a personality disorder. Second, a diagnosis of antisocial or borderline personality disorder can be stigmatizing, particularly in a general medical clinic where health care providers may have less experience managing such patients. Finally, a classification system based on a continuum approach is a better predictor of HIV risk behavior than the discrete categorical classification of patients.

Implications for Medication Adherence

Average medication adherence across a variety of diseases and patient populations has been consistently estimated at 50 percent. Adherence is especially challenging in HIV disease, which is associated with all of the components of low adherence: long duration of treatment, preventative rather than curative treatment, asymptomatic periods, and frequent and complex medication dosing. Average rates of nonadherence to antiretroviral therapy range from 50 to 70 percent, with adherence rates of <80 percent associated with detectable viremia in a majority of patients.

Personality factors have received little investigation in relation to adherence and antiretroviral therapy. Personality traits such as neuroticism were significantly associated with poorer quality of life, whereas conscientiousness and extraversion were associated with better quality of life. In contrast, personality traits were not directly related to HAART adherence. However, the authors' clinical experience suggests that nonadherence is more common among their extraverted or unstable patients. The same personality characteristics that place them at risk of HIV also reduce their ability to adhere to demanding drug regimens. Specifically, their present-time orientation, combined with reward seeking, makes it more difficult for these patients to tolerate uncomfortable side effects from protease inhibitor drugs whose treatment effects may not be immediately apparent. It is also difficult for feeling-driven individuals to maintain consistent, well-ordered routines. Hence, following frequent, rigid dosing schedules can also be problematic. The authors' unstable, extraverted patients are usually intent upon following the schedule, but their chaotic and mercurial emotions are more likely to interfere and disrupt daily routines. For example, a patient may report that he felt very upset and nihilistic after a fight with a family member and miss several doses of his antiretroviral medicines. Missed doses of HAART can increase the chance of HIV resistance developing. Identifying factors that influence adherence in HIV disease is important in improving overall health outcomes.

Treatment Implications

Psychiatric and medical treatment of patients with extraverted and/or emotionally unstable personalities is challenging. Such patients are often baffling or frustrating for physicians and other medical care providers because they engage in high-risk sex and drug behaviors in spite of knowing the risks, or fail to adhere to treatment regimens for HIV infection in spite of knowing the consequences. A patient may stop his antidepressant medicine because of a headache while being perfectly willing to inject heroin into multiple sites on his body. After 6 months of missed medical appointments, a patient may impulsively leave the clinic if the primary care provider is 15 minutes late for the appointment. Such personality traits reflect relatively stable, lifelong modes of responding; thus, direct efforts to change these traits are unlikely to be successful. It is possible, however, to modify the behavior that is an expression of the trait. By recognizing individual differences in risk-related personality characteristics, interventions can be better targeted and their impact maximized.

The authors have found that a cognitive behavioral approach is most effective in treating patients who present with extraverted and/or emotionally unstable personalities. Following are the five principles that guide the authors' standard care:

1. Focus on thoughts, not feelings. Individuals with unstable, extraverted personalities benefit from learning how they are predisposed to act in certain ways. Often, they do not recognize the extent to which their actions are driven by the impulse or feeling of the moment. These patients can fail to understand why they intend to stay clean but later find themselves "shooting dope." Treatment helps to identify the role that strong feelings play, so that these patients can begin the process of understanding their own chaotic, often irrational behavior.

   Simultaneously, treatment encourages the patient's cognitive, logical side. This process begins by identifying what the patient thinks *in* a given situation, as opposed to what s/he thinks *about* the situation: "I deserve some cocaine because I have had a difficult day." The influence of the patient's assumption upon feelings, behavior, and ultimately the consequences of behavior are examined. Through the treatment dialogue patients understand that the notion of a "cocaine reward" creates a feeling of urgency and entitlement to getting high. As a consequence of using cocaine, the development of other more constructive reward systems or coping methods are preempted, a relapse into cocaine dependence can occur, and/or family or employers may be alienated. Through treatment, patients learn to identify maladaptive assumptions that drive feeling and behaviors so that they can either lessen the force of these assumptions or substitute more constructive assumptions to guide their life experience.

2. Use a behavioral contract. A behavioral contract is developed with all patients. The contract outlines goals for treatment, often only a day or a week at a time. While patients and mental health professionals may develop the contract, the focus of treatment is *not* on what patients *want* or *are willing to do* to get off of drugs, but rather on established methods, such as drug treatment and Narcotics or Alcoholics Anonymous. The importance of the behavioral contract lies in the creation of a *stable* plan that supersedes the emotional meanderings of these patients. Unstable extraverts present an ever-changing array of concerns and priorities. The task of the treatment is to order the priorities with patients and help them follow through on these, regardless of changing emotions. In short, behavioral contacts provide consistent, cognitive focus to patients' bewildering emotional experience of life.
3. Emphasize constructive rewards. In developing the behavioral contract and in treatment, the purpose is cast in terms of the rewards that will follow from their behavioral change. Positive outcomes, not adverse consequences, are salient to extraverts. Most of the patients have already experienced negative consequences from their behavior, HIV, drug addiction, homelessness, and such. Exhortations to use condoms to avoid STDs are unpersuasive. More success has been achieved with extraverts by eroticizing the use of condoms or by the addition of novel sexual techniques (erotic massage, use of sex toys) into sexual repertoires. Similarly, the rewards of abstaining from drugs or alcohol are emphasized, such as having money to buy clothing, having a stable home, or maintaining positive relationships with children.

   In building adherence to antiretroviral therapies, the focus is on the rewards of an increased CD4 count and reduced viral load, rather than avoiding illness. Using the viral load as a strategy to build adherence can increase acceptance in all patients, but is especially effective in reward-driven extraverts.
4. Use relapse prevention techniques. The relapse prevention model, originally developed for treatment of substance abuse behavior, is an effective method for changing any habitual way of behaving. This intervention trains individuals to recognize and interrupt the sequence of behaviors that link to the final high-risk behavior. Behavioral therapy methods are also used to teach individuals how to recognize and avoid situations that trigger high-risk behaviors.
5. Coordinate with medical care providers. Medical care providers are often frustrated or discouraged when treating unstable, extraverted patients. It is useful to provide education about a patient's personality and how it influences behavior. Particularly effective is the development of a coordinated treatment plan, where medical care provider and mental health professional work in tandem to develop behavioral contracts to

reduce HIV risk behaviors and build medication adherence.

Conclusion

Personality characteristics and personality disorders reflect relatively stable, lifelong propensities that are difficult to change. This does not mean, however, that HIV risk reduction efforts are necessarily futile. Rather, by understanding personality characteristics and their role in HIV risk behavior and medication adherence, the mental health professional can develop more effective, specific treatment strategies. Similarly, the HIV-infected patient who can identify aspects of their personality that might interfere with intentions to practice safer behavior, and who knows strategies for dealing with these situations, is less likely to practice high-risk behaviors. Finally, the mental health professional can provide valuable assistance to medical care providers to improve health outcomes for these patients.

ISSUES OF SUBSTANCE USE DISORDERS AND ADDICTION IN HIV DISEASE

Substance use is a primary vector for the spread of HIV. This impact is directed not only at injection drug users and their sexual partners but also at those who are disinhibited or cognitively impaired by intoxication, causing them to act impulsively and engage in unsafe sexual practices. Furthermore, patients infected by HIV are often demoralized, become hopeless, and are more likely to use drugs and engage in high-risk behaviors. Patients with substance use disorders may not seek health care or may be excluded from health care because of stigmatization. In addition, intoxication and the behaviors necessary to obtain drugs interfere with adherence to medication regimens and medical appointments.

Injection drug use is obviously a primary risk factor for contracting HIV by needle sharing. In the United States, injection drug use has accounted for approximately one-third of all AIDS cases. In alcohol and noninjection drug users, substance use plays a major, albeit more subtle, role in HIV transmission. Addiction and high-risk sexual behavior have been linked across a wide range of settings. For example, female crack cocaine abusers are more likely to engage in prostitution to obtain money for drugs. Homosexual men who use crack cocaine or methamphetamine are more likely to engage in unprotected anal sex with casual male contacts. Alcohol use, which is very prevalent in the HIV population, can lead to risky sexual behaviors during intoxication by way of cognitive impairment and disinhibition.

A multifactorial matrix of influences initiates, drives, and sustains substance abuse and addiction. Many of these are intrinsic to the substances themselves, but others are characteristic of the host or the environment. Behavioral approaches to understanding addictions have

been particularly fruitful, as seen in the work of Joseph Brady and his colleagues. This approach allows the development of animal models of self-administration and the measurement of reinforcing properties of drugs, many of which are profoundly predictive of human behavior. However, environmental influences and the understanding of the individual, cultural and social forces that impact substance use are essential to translation of those models into human settings.

Psychiatric and psychological comorbidity can increase vulnerability to substance use disorders. Personality factors may lead to more risk-taking, greater likelihood to experiment with novel sensations, and increased sensitivity to rewards, therefore leading to more sensitivity to the reinforcing properties of drugs and less sensitivity to the negative consequences of drug use. Other personality types are consequence and risk avoidant, and are relatively protected from addiction. However, these risk avoidant types may become addicts because of an underlying affective or anxiety disorder and turn to the rewarding properties of drugs and alcohol to "self-medicate" dysphoria and anhedonia. For example, depression can make the ordinary experiences of life less rewarding, and sensitizes people to the positive reinforcement of drugs. Life experiences that expose people to drugs such as social acceptance of drug use within peer groups or family members, particularly during adolescence, can also increase the risk of addiction.

Individual biology is involved in several ways. In the case of alcohol, genetic factors may affect the degree to which alcohol is rewarding, so that some patients report that their first drink was so rewarding that they began a lifetime of heavy drinking immediately. Others will say they never really liked drinking all that much, and therefore are surprised as they gradually become more dependent on alcohol to control the emotional discomforts of their lives. Genetic predisposition plays a lesser role in cocaine and opioid use. These substances are intrinsically more rewarding and patients exposed to these substances have a higher use to addiction ratio. Finally, co-occurring medical problems common in HIV such as chronic pain, opportunistic infections, and surgical procedures may result in exposure to opioids and/or sedatives, which can lead to addiction in a vulnerable individual.

Research on substance use and HIV is complicated by the same problems of definition, detection, and methodology that impede research on the other areas discussed here. The DSM-5 recognizes that substance use disorders may exist on a spectrum of increasing use, physiological and psychological dependence, and increasing impairment of function that blend gradually into one another. Indeed, it is often difficult to define precisely when the transition from heavy drinker to alcoholic occurs. Some patients use heavily but never actually become disordered by their use. Others are disordered by surprisingly modest use of a substance. The key

feature of addiction is the habitual, compulsive use of substances where drug procurement and use become the main focus and goal of the person's life and continues despite negative physical, psychological, and social consequences. Physiological dependence defined by tolerance and/or withdrawal may be present but should not be confused with addiction or a substance use disorder as defined by DSM-5.

Substance Use Disorders and Their Interaction with HIV Disease and Treatment

Ongoing substance use has grave medical implications for HIV-infected individuals. Screening for alcohol and drug use disorders is essential during the initial clinical evaluation and a number of quick and easy screening tools are available to clinicians, such as the CAGE, AUDIT-C, and DAST-10. Substance use may be difficult to detect because physical symptoms of HIV infection overlap with those of substance intoxication and/or withdrawal, including malaise, fatigue, weight loss, fevers, and night sweats. In addition, the accumulation of medical sequelae from chronic substance abuse can accelerate the process of immunocompromise and amplify the progressive burdens of the HIV infection itself. Injection drug users, for example, are at higher risk of developing bacterial infections such as pneumonia, sepsis, soft tissue infections, and endocarditis. Tuberculosis, STDs, viral hepatitis infection, and coinfection with human CD4 cell lymphotropic virus occur more commonly in injection drug users who are infected with HIV. Certain malignancies, lymphomas in particular, occur more frequently in HIV-infected drug users. Alcohol users may experience faster progression of HIV disease and poorer response to antiretroviral therapy secondary to the immunosuppressive effects of alcohol. In addition to the direct physical effects caused by drugs, active substance use is highly associated with both nonadherence and reduced access to antiretroviral medication.

Neurological symptoms can overlap between HIV infection and substance abuse. For instance, both AIDS dementia and drug intoxication can present with apathy, disorientation, aggression, and an altered level of consciousness. Drug withdrawal can present with seizures and neurovegetative symptoms, as can opportunistic infections of the CNS. HIV-infected injection drug users tend to be at higher risk of developing fungal or bacterial infections of the brain and spinal cord. HIV-infected patients who drink alcohol, particularly those older than 50, may be more vulnerable to cognitive decline and structural brain changes on neuroimaging than patients who are nondrinkers.

Co-occurring psychiatric disorders are common in the drug using HIV population. In the past, the term "dual diagnosis" referred to a patient diagnosed with both a drug use disorder and another psychiatric disorder and "triple diagnosis" referred to a dual diagnosis patient who also has HIV infection. Such patients are over represented in treatment settings because

of their symptom severity and chronicity. For instance, in inner city Baltimore, as many as 44 percent of new entrants to the HIV medical clinic at Johns Hopkins Hospital had an active substance use disorder. Twenty-four percent of these patients had both a current substance use disorder and a co-occurring nonsubstance-related Axis I diagnosis.

Affective disorders, especially major depression, are the most common co-occurring psychiatric disorder in this population with studies estimating a prevalence of 15 percent to 30 percent. Diagnosing affective disorders (and other psychiatric disorders) in drug users can be difficult and even controversial. This controversy stems from the problem in determining the causal or even chronological relationship between drug disorders and affective disorders. Although some theorists have wanted to emphasize the primacy of one or the other in guiding treatment, this "chicken or egg" approach is not especially productive. Given the prevalence of overlapping addictive and affective disorders in clinical settings, as well as the poor prognosis associated with both disorders if left untreated, a treatment approach should necessarily emphasize simultaneous and equal treatment of both entities. This is not to suggest that it is easy to distinguish transient depressive symptoms such as those presenting in drug withdrawal, demoralization, or grief reactions from persistent depressive symptoms indicating a major affective disorder. It may be necessary to observe the patient over a period of abstinence in a confined treatment setting, before the presence of primary mood disorder can be established. Family psychiatric history, temporal relationship of symptoms to periods of drug use, and abstinence as well as longitudinal history of the patient's symptoms obtained from family members, friends, and other psychiatric providers can help clarify the presence of a separate affective disorder.

The importance of identifying affective disorders lies not only in their own well-known sequelae, including suicide, but also in their complex interactions with addiction and HIV disease. Depression is associated with higher severity of addiction, resistance to treatment, nonadherence to antiretrovirals, higher viral loads, and lower CD4 counts. The anhedonia of depression can be a sustaining factor in addiction. Addicts with major depressive disorder find it difficult to respond to and enjoy life's other rewards that pale in comparison to the intense, though, ephemeral rewarding effect of intoxicating drugs. Depressed patients are also more difficult to engage, invest in, and sustain treatment, given their anergy and negativism. It is essential, therefore, for the clinician to recognize and treat depression early to maximize a successful outcome and improve patient adherence and retention in treatment.

Difficulties in the realm of personality (Axis II) are among the most common psychiatric problems seen in this population. Although personality disorder diagnoses are currently described in a categorical fashion in the DSM-5, it is probably more useful to view personality as

being dimensional in nature. Using this model, personality traits exist along a continuum, which predicts habitual maladaptive approaches to life's difficulties.

Most HIV-positive substance abusers would be classified as "unstable extraverts," with personality traits of risk-taking, reward seeking, emotionally driven, and present-time oriented. These traits are commonly present in the cluster B personality disorders in the DSM-5 (antisocial, borderline, narcissistic, and histrionic) and can be found in as many as 49 percent of all substance abusers. Not only do these traits result in a vulnerability to addiction and other risky behaviors that predispose one to become infected with HIV, but they also pose significant barriers to both HIV and addiction treatment. These patients tend to act on strong, impulsive feelings rather than on carefully considered treatment instructions. Their behavior tends to be driven by the transient, immediate rewards of drugs rather than by their lasting future consequences. Such patients tend to get bored easily and treatment is often unexciting. They tend to "want what they want when they want it" rather than when it may be good for them. It is critical to identify these personality vulnerabilities because they can have a profound effect on treatment engagement and prognosis.

Because the HIV-infected patient is likely to be on a variety of antiretroviral agents and prophylactic agents for opportunistic infections, the clinician must be especially mindful of interactions between these medications and the abused substances. For example, dideoxyinosine can cause peripheral neuropathies as a side effect, which may be worsened by the neurotoxic effects of alcohol and malnutrition related to chronic substance abuse. Opioid users on methadone maintenance treatment are at particular risk of medication interactions. Rifampin, for example, increases the elimination of methadone from the body and may result in the rapid onset of withdrawal symptoms. Decreased plasma levels of methadone also occur with concurrent administration of ritonavir, nelfinavir, efavirenz, fosamprenavir, and nevirapine; necessitating adjustments in methadone dosage should withdrawal symptoms occur. This has important implications for treatment adherence in that the patient in a methadone program may be less likely to take a medication because of the fear of going into opioid withdrawal. Patients may be more likely to relapse if opioid withdrawal symptoms do occur when they are started on antiretrovirals without proper communication with their methadone program. Sublingual buprenorphine may be alternatively used for opioid maintenance therapy. Studies examining drug interactions between buprenorphine and antiretrovirals have shown little to no clinically significant interactions.

Treatment of Substance Use Disorders in Patients Infected with HIV

Although oversimplified, the steps for the treatment of substance use can

be outlined in this simple way. These steps often occur simultaneously as treatment begins, but will be described as a sequence.

1. Induction of patient role
2. Detoxification
3. Treatment of comorbid conditions
4. Rehabilitation
5. Relapse prevention

Role Induction and Motivation to Change. The initial and often most daunting task of treating the addict are engagement and induction of the patient role. The general rule is that addicts and treatment providers begin with differing agendas—addicts tend to come to treatment settings seeking comfort and crisis relief, whereas physicians and other health care providers look at long-term goals of improvement in the patient's health and overall functioning. One of the critical initial tasks of the health care provider is to engender the gradual evolution of the patient's attitudes to coincide with those of the treatment plan.

Prochaska and DiClemente have described a "transtheoretical stages of change" model to explain the addiction and recovery process. The patient is viewed as progressing through several different stages of change: (1) precontemplation—the patient has no intention to change their addictive behavior; (2) contemplation—the patient is considers change because of the negative consequences of his drug use but is ambivalent about it; (3) preparation—the patient shows intention to change and takes initial steps to seek treatment; (4) action—the patient decides to modify behavior, environment and circumstances in order to relinquish the addictive lifestyle; (5) maintenance—the patient works to prevent relapse and consolidate his changed behavior and lifestyle. The clinician's job is to assist the patient in moving from one stage to the next in order to facilitate the recovery process. A technique known as "motivational interviewing" developed by Miller and Rollnick can be used to heighten the patient's readiness to change by using empathy and gentle confrontation to amplify the discrepancy between the substance abuser's current lifestyle and long-term life enhancing goals.

Detoxification. In order for intoxicated patients to understand and process the cognitive steps needed for recovery, detoxification is the first step. Many HIV positive substance abusers benefit from a brief hospital stay to stabilize their psychiatric and medical comorbidities. Slowly tapering either the drug of dependence or using a cross-dependent drug that has a similar pharmacological mechanism of action best accomplishes detoxification.

Detoxification is often unpleasant, and there has been no evidence to support the idea that noxious withdrawal during detoxification improves

outcome. In fact, some behavioral studies suggest that patients suffering through severe withdrawal during detoxification may actually develop conditioned withdrawal, such that exposure to environments similar to the one experienced during "cold turkey withdrawal" can bring about subacute withdrawal symptoms months later leading to relapse. It should be noted that benzodiazepine, barbiturate, and alcohol withdrawal can be life threatening and clinicians should have a low threshold for admitting these patients for inpatient detoxification. Active tapers using a slow downward titration of medication from the class to which the patient is addicted is recommended for opiates and sedative-hypnotics. Some authors have used antidepressants to help patients detoxify from psychomotor stimulant classes of drugs (amphetamines and cocaine) but data to support this practice is controversial. This approach may work best for patients with clear evidence of major depression.

Sedative-hypnotics or alcohol are best detoxified through the use of a long-acting benzodiazepine with a quick onset of action such as diazepam or chlordiazepoxide. Lorazepam or oxazepam should be used in patients with liver disease to prevent the accumulation of active metabolites. Detoxification of opiates is accomplished by starting patients on a taper of sublingual buprenorphine or oral methadone. Clonidine, methocarbamol, dicyclomine, loperamide, and ibuprofen can be used adjunctively to provide symptomatic relief.

Treatment of Co-occurring Psychiatric Conditions

Many patients with HIV and addictions have co-occurring psychiatric conditions, which need to be treated in order to maximize treatment adherence and abstinence. Conditions such as major depression, bipolar disorder, and schizophrenia are best managed with pharmacological treatment. Because these patients tend to have multiple medical complications, it is important to remember to start medications at low dosages and to titrate slowly to minimize the risk of developing adverse side effects and delirium.

Disorders of personality, in particular unstable extraversion, are managed with cognitive-behavioral forms of psychotherapy. The ways that unstable extraverts may sabotage treatment include staff splitting, doctor shopping, general nonadherence and manipulative behavior. Therapy addressing these personality issues should include firm limit setting and consistency on the part of all health care providers involved. To this end, a documented treatment plan with clear goals agreed upon by all the treatment staff is essential. The treatment plan should be reviewed with the patient at the initiation of treatment and regularly during treatment so that he understands clearly what is expected of him and what he can expect from his treatment providers if he is adherent to these goals. Frequent and consistent communication among all treatment providers during the course

of treatment can minimize splitting and address issues of nonadherence as they arise.

Maintenance Treatment and Relapse Prevention

Long-term treatment is necessary for patients to begin the process of lifestyle change and recovery since addiction is a chronic disorder. Because this patient population is complicated and especially vulnerable to relapse, the most useful model of treatment is integrated care. To this end, an HIV clinic with comprehensive care is especially useful to engage and maintain patients in treatment. Ideally, a clinic treating HIV positive addicts should include medical care providers, psychiatrists, social workers, housing counselors, day care workers and substance abuse counselors. Integrated settings can help the patient access needed services, adhere to the overall treatment plan and improve provider communication. Specific strategies integrating substance abuse treatment with HIV care include the on-site provision of sublingual buprenorphine for opioid maintenance treatment at HIV clinics (see BHIVES Collaborative model below) and providing directly observed antiretroviral therapy and HIV care at methadone maintenance clinics.

It is important to remember that addiction treatment is active rather than passive and involves transforming previously held beliefs, attitudes and personal identity into a new way of life. To this end, group therapy should be included as part of all substance abuse treatment. Various group modalities are available, including 12-step meetings (Alcoholics or Narcotics Anonymous), network therapy, rational recovery, therapeutic community or SMART recovery. Group principles are similar in that more experienced members of the group provide both confrontation and support for the newly initiated member. Group support also provides the newly recovering addict with a hopeful view of the benefits to be achieved with recovery, exemplified by the lifestyle and achievements of group members with longer periods of abstinence. A commitment to a community of recovery assists the patient in severing ties from the drug community and provides the patient with new bonds that help maintain a sense of purposefulness and hopefulness. Specific HIV-positive recovery groups are now widely available that may be helpful for addicts who are uncomfortable with their HIV status in a regular group setting.

Patient individualized therapy should focus on identifying triggers to substance use, on minimizing or decreasing exposure to substances, and on defining a clear plan of action if relapse occurs. It is important to realize that relapse is often the rule and not the exception, and plans should be in place for early intervention. Monitoring measures such as urine and serum toxicologies and Breathalyzer tests can help to enforce compliance. Contingency management using a variety of positive and negative reinforcers tied to urine toxicology results has also been shown to be

effective in maintaining sobriety.

Individual and family therapies can enhance the effectiveness of treatment but should not take the place of group therapy. In individual treatment it is important that the treatment provider remain flexible in the treatment approach. While regular psychotherapy may work for some patients, others may need a more "hands-on" approach and benefit from being referred to vocational rehabilitation, occupational therapy and social skills training. Treatment failures often result when the therapist has a "one size fits all" mentality and adheres too rigidly to one model of therapy.

Pharmacological treatments can be used as adjuncts to the overall treatment plan but not as a replacement. Pharmacological treatment can be divided into the following categories:

1. Aversive conditioning
   a. Disulfiram (alcohol)
2. Blockade of positive or negative reinforcement
   a. Naltrexone (opioids, alcohol)
   b. Acamprosate (alcohol)
3. Drive suppression
   a. Bupropion (tobacco)
   b. Buprenorphine (opioids)
   c. Methadone (opioids)
   d. Naltrexone (opioids)
   e. Varenicline (tobacco)
   f. Nicotine replacement therapies (tobacco)

Relapse can occur during a single impulsive moment, and some patients find pharmacological therapy helpful as a type of "insurance policy" against cravings. Disulfiram, an inhibitor of acetaldehyde dehydrogenase, is taken once daily at dosages from 250 to 500 mg and causes an unpleasant reaction when alcohol is ingested due to the buildup of acetaldehyde in the body. Symptoms include nausea, flushing, headaches, and hypotension. Liver enzymes should be monitored because of the risk of hepatotoxicity. Naltrexone given orally once a day has been shown to reduce alcohol cravings in a number of studies, probably by reducing the pleasurable effects of alcohol in alcoholics. Naltrexone is also available in an intramuscular long-acting depot formulation given once a month, which may help patients who have problems with medication adherence. As with disulfiram, liver function tests are monitored in patients on naltrexone because of the risk of hepatotoxicity. Acamprosate reduces alcohol relapse by inhibiting excitatory glutamatergic transmissions that cause subacute withdrawal symptoms in newly abstinent individuals. This medication is excreted renally so it can be particularly useful in patients with liver disease or impairment; however, it is dosed three times a day that makes medication adherence more problematic.

In opioid dependent individuals, pharmacotherapy has been shown to be highly effective and includes the use of both opioid agonist and antagonist medications. Naltrexone is an opioid antagonist that has a high affinity for blocking mu receptors. Opioid addicts maintained on naltrexone experience little or no euphoria when opiates are used. The medication should only be started in patients when they are opioid-free because of the possibility of precipitating withdrawal symptoms. Methadone is the opioid agonist most commonly used and has been extensively studied for maintenance treatment. This medication is given to the patient at varying dosages from 60 to 120 mg daily and can only be dispensed in a federally regulated treatment facility. Patients on methadone maintenance treatment show better levels of antiretroviral adherence and reduced HIV risk behaviors. Several agents used to treat HIV and related infections induce methadone metabolism so providers should investigate drug–drug interactions before prescribing and monitor patients closely for signs of withdrawal once therapy has been initiated. Sublingual buprenorphine/naloxone (bup/nx) has been shown to be as effective as methadone in opioid maintenance treatment and has the advantage of being prescribed in an outpatient clinic setting by a qualified physician with a Drug Enforcement Agency waiver. Bup/nx is well tolerated by HIV patients and no significant elevations of liver enzymes have been noted during treatment.

Integrating pharmacological treatment into HIV care can result in improved treatment outcomes and patient satisfaction with care. The Buprenorphine-HIV Evaluation and Support (BHIVES) Collaborative was a multisite demonstration and evaluation project funded by the US Department of Health Resources and Services Administration from 2005 to 2007 that examined the integration of sublingual bup/nx treatment into HIV primary care clinics. Patients participating in integrated treatment were 52 percent less likely than nontreated patients to use illicit opioids during each treatment quarter and reported overall decreased illicit stimulant use. Patients not on antiretroviral therapy at baseline were found more likely to be started on treatment after initiating bup/nx and to achieve nondetectable viral loads over time. Improvements were also noted in their health-related quality of life. Patients receiving integrated treatment subjectively reported being highly satisfied with the effects of bup/nx and with their treatment outcomes. One study site found clear advantages to integrated treatment compared to referral for treatment to an opioid treatment program offsite: patients prescribed bup/nx in the HIV clinic were more likely to start on opioid agonist therapy, less frequently found to have positive opioid and cocaine test results, and attended significantly more HIV primary care visits.

PSYCHOLOGICAL PROBLEMS IN PATIENTS INFECTED WITH HIV

In the context of the psychological realm of living with HIV, it is difficult to distinguish which came first, the variety of psychiatric disorders the patient has, the HIV risk behaviors, the losses and psychological traumas, or the impact of having HIV itself. In fact, it is a common sentiment from patients to echo one of the patients at the authors' clinic who said "HIV is not even one of my biggest problems." Each of these factors seems to be both a consequence and an antecedent to all the others. The DSM-IV provides psychiatrist with the opportunity to choose between "Adjustment disorder" and PTSD for patients living with violence, despair and loss of unthinkable proportion. Unfortunately, the psychological difficulties of living with HIV are infinitely more complex than this.

Patients with HIV have been shown to have more severe trauma, high rates of PTSD and anxiety, economic and social disenfranchisement, and high rates of interpersonal instability. Cognizance of the psychological issues in the care of patients with HIV is essential. On the other hand, the treatment of these problems is well within the scope of most clinicians in our field. The psychotherapy for patients infected with HIV can be broadly divided three categories:

1. Psychotherapy for the problems of life encountered by patients infected with HIV
2. Psychiatric treatment for specific psychiatric conditions associated with HIV infections
3. Specific psychoeducational and psychotherapeutic interventions associated with specific types of interactions with HIV-infected patients

For numbers 1 and 2 above, the psychotherapeutic issues are largely similar to those seen in patients who are not infected with HIV. While specific papers have been written regarding coping skills, social isolation, grief, PTSD, intimate relationships, family problems, and many others, these recommendations are largely similar to those found in patients coping with other chronic medical conditions, or chronic impoverishment elsewhere in this text. The psychotherapy issues of conditions commonly seen in HIV clinics have been reiterated where appropriate, but are also similar to those issues for patients with the same conditions but without HIV infection. This chapter includes a section on PTSD because of the high prevalence and special issues for patients with HIV.

Posttraumatic Stress Disorder

Traumatic events that are life threatening provoke terror, anxiety, and stress in most people. In some individuals, the chronicity, intensity, frequency, and comorbidity of these symptoms can become psychiatrically disabling. Specifically, patients have intrusive intense recollections of the traumatic event, sometimes to the point where they feel as if they are experiencing the event again, and there is a persistent avoidance of stimuli

associated with the trauma and persistent symptoms of increased arousal not present before the trauma. When these symptoms persist for more than 1 month and interfere with social, familial, and/or occupational functioning, PTSD may be diagnosed (DSM-IV-TR). PTSD has a current prevalence of less than 1 percent and a lifetime prevalence of 1 to 9 percent with a female-to-male lifetime prevalence ratio of 2:1. In civilian populations, rape is the event most likely to produce PTSD, particularly if it occurs before or during adolescence.

PTSD increases the likelihood of engaging in destructive behaviors such as alcohol and other drug abuse, sexual promiscuity, or prostitution. PTSD is of particular concern in HIV treatment and research because it may engender or exacerbate HIV risk behaviors and worsen health outcomes.

Cross-sectional research has shown that both symptoms of PTSD and PTSD have been associated with HIV risk behaviors and markers of HIV progression. Symptoms of PTSD in adolescence have been associated with prostitution, injection drug use, and choice of a high-risk sex partner in young adults. HIV-infected adults who have a history of child sexual or physical abuse have reported engaging in more HIV risk behaviors such as drug abuse and sexual compulsivity than persons with no history of trauma. A high prevalence (42 percent) of HIV-infected women attending county medical clinics had PTSD symptoms sufficiently severe to meet a diagnosis of PTSD. In HIV treatment, traumatic stressors and PTSD symptoms have also been associated with lower CD4 T cell to CD8 T cell ratio at 1-year follow-up. The relationship between a psychiatric diagnosis of PTSD and HIV risk behavior or infection has received less attention. Veterans with a diagnosis of PTSD are at greatly increased risk of HIV infection, particularly if they are also diagnosed with a substance abuse disorder. Women prisoners with a lifetime history of PTSD are more likely to have engaged in prostitution and receptive anal intercourse prior to incarceration compared to women prisoners without PTSD. These studies of PTSD/PTSD symptoms and HIV risk behaviors have been cross-sectional, thus a causal relationship cannot be inferred. It may be that HIV risk behaviors, such as prostitution or drug abuse, increase exposure to trauma, and thus the likelihood of developing PTSD. Alternatively, PTSD that stems from early trauma may predispose an individual to engage in sex or drug behaviors that can increase the risk of HIV infection.

The presence of PTSD in an at-risk or HIV-infected positive individual is of particular concern because of high rates of comorbidity (up to 80 percent) with other psychiatric disorders. Specifically, PTSD is most often comorbid with depression and cocaine/opioid abuse—both risk factors for HIV. Prior depression may be either a risk factor for the development of PTSD following a traumatic event or a co-occurring response with PTSD to a trauma. Substance abuse may be either an attempt to “self-medicate” suffering after a traumatic experience or a lifestyle that increases exposure

to traumatic events, such as robbery or assault.

PTSD and substance abuse disorders that occur together can also adversely affect treatment. Comorbid conditions have been associated with poorer treatment adherence and motivation, quicker relapse, more inpatient hospitalizations and medical problems, and lower global functioning than either disorder alone. Thus, treatment of PTSD that does not address coexisting depression or substance abuse may be insufficient or even worsen psychiatric status. PTSD treatment that typically involves behavioral exposure and flooding has been reported to exacerbate emotional arousal and precipitate relapse, although this had not yet been shown in experimentally controlled investigations. Likewise, the AA philosophy of surrender and sharing one's story may be counterproductive to substance-abusing HIV-infected individuals with PTSD.

HIV at-risk or HIV-infected individuals should be routinely screened for PTSD. Instruments such as the Trauma History Questionnaire and the PTSD checklist have increased detection rates of the disorder. Similarly, any individuals presenting with symptoms of PTSD should routinely be screened for depression and substance abuse. For HIV-infected individuals with PTSD and another concurrent psychiatric disorder, treatment that simultaneously addresses both disorders is likely to be more effective and practical.

HIV-Specific Psychotherapeutic Issues

There are a number of specific circumstances regarding HIV-infected patients that should be discussed here. They include the following:

1. Pretest, test and posttest counseling issues
2. Risk behavior reduction in patients at risk or infected with HIV
3. Issues of partner notification in patients infected with HIV
4. Impaired patients with issues of capacity and competence
5. HAART adherence issues

Pretest/Test Counseling and Education

Patients at risk of HIV infection are often reticent to get testing. Surveys suggest they fear the results of the test or are too overwhelmed by the issues of their current life and behavior to present for testing. Prior to 1993, patients who received a positive HIV test result saw their diagnosis as essentially fatal. Additionally, many patients thought it would be burdensome to know that they were placing others at risk of infection. In more recent years with the advent of HAART, HIV has become a chronic treatable illness. In this setting, it seems more reasonable for patients to get tested and engage in treatment. Nonetheless, survey data shows that a large number of patients that are at significant risk are not getting tested. Psychoeducational psychotherapy directed at encouraging patients to get

tested has been offered to a variety of at-risk populations of patients. The outcomes of these intervention studies show that such psychotherapy does result in patients getting tested and diagnosed earlier.

Pretest counseling has been described in a number of papers. Prior to testing, patients need informed consent regarding the meaning of a positive and a negative test. It should be explained that the test looks for antibodies the individual's body makes to the HIV virus, not the virus itself. It should be stressed that a negative test does not mean that a patient is immune and cannot become infected later, and that a positive does not mean that a person has AIDS, is going to die, or will suffer from opportunistic infections. It should also be stressed that the test will remain negative for a time after infection (the time it takes for antibody to develop), and therefore after a recent exposure a patient may have a negative test, but be in fact infected. Pretest counseling should also include information on safe sex, safe needle, and other risk-reduction interventions. This is because a significant percentage of patient who obtain testing do not return for their results for extended periods of time and sometimes not at all. The development of rapid testing using cheek cell swabs has provided for a much faster turn-around time, thus enabling a quicker result for the patient, but confirmation with a blood test is usually recommended, and so a combination of pre- and posttest counseling and then another pretest counseling session may be required in short order.

For patients with HIV infection, a number of monitoring tests are necessary, including CD4 counts and HIV RNA loads, and sometimes resistance testing. In similar fashion to the counseling required before HIV antibody testing, patients should be counseled before each of these types of tests, explaining what the test is for, what the possible results will be, and what the test results could mean as to the overall treatment and prognosis. Time should be allowed for questions, and many clinics have developed take-home pamphlets on these tests.

Posttest Counseling

A number of articles have described posttest counseling psychotherapy issues and interventions. These include both psychoeducational interventions regarding the meaning of test results, recommendations for treatment, and importantly, risk reduction interventions to stem the spread of HIV infection. These posttest interventions should occur in both HIV-negative and HIV-positive patients.

At the time of test results being given to patients, it is not uncommon for patients to have a variety of intense psychological reactions including suicidal feelings, anger, homicidal thoughts directed at potentially infecting partners, overwhelming grief, and complete psychological breakdown. Patients with poor coping skills, poor impulse control, history of suicidal feelings and behaviors, substance abuse disorders, and lack of social

supports are at increased risk of impulsive behaviors and self-destructive behaviors. Because of these circumstances, availability of psychological interventions at the time of HIV testing and result provision is critical.

For established HIV positive patients, bad news regarding progression of HIV disease or detection of resistance can provoke similar responses, and should be considered in a similar way. Transition from the asymptomatic phase to the development of an opportunistic infection or to a formal diagnosis of AIDS because of a decline in CD4 cells may provoke denial, anger, depressive feelings, anxiety, hopelessness, or a myriad of other emotions. Again, the presence of psychological evaluation and intervention is extremely critical in this setting. Patients may be overwhelmed by the news that they need to start antiretroviral drugs, and therefore need emergent attention at this time. In particular, discussions around detection of viral resistance may evoke a number of defensive responses from the patient, as there may be a perception of failed adherence and guilt.

Psychotherapy to Prevent HIV Transmission in Selected Populations of Patients

Men who have sexual contacts with other men may either be exclusively homosexual, bisexual, or heterosexual men. Men with sexual contacts with other men were the largest subgroup in terms of new AIDS diagnoses in the United States in the year 2000. In states where HIV is reported, it continues to be the largest subgroup of newly reported HIV infections. In intervention studies looking at men who have sex with other men, many interventions have shown a decrease in either risk behaviors or infection. Studied interventions include stress management and relaxation techniques, education cognitive self-management training, negotiation skills training, psychotherapy directed at emotional distress reduction, relapse prevention models of high-risk behavior reduction, education directed at eroticizing safer sex, assertiveness training, and peer education in bars. Outcomes of these interventions showed that all have a modest impact on either risk behavior or HIV infection, depending on study outcome measured. Although there are fewer data, similar studies have been done targeting heterosexually transmitted HIV, substance-related risk behaviors, women, and IV drug users. It is unclear from the data what the best intervention is, and how to stratify the interventions. More important the results of these studies are quite modest, with a 25 percent reduction in risk being quite a good outcome.

Studies of rates of psychiatric disorders in at-risk population show impressively high rates of affective disorders substance abuse, personality disorders, and psychological distress. As yet no systematic study with treatment and targeted intervention methods based on psychiatric diagnosis has been reported. It is clear from the data on risk and epidemiology that this is the direction that needs to be taken to try to

improve prevention.

Partner Notification

The landmark legal decision in the Tarasoff case in California has resulted in the increased scope of responsibility for health care providers. A variety of legislation differing from state to state has afforded practitioners an increased number of options with regard to confidential situations. In some states, Tarasoff statutes (those statutes providing a duty to warn vulnerable individuals of imminent danger from a patient overriding issues of confidentiality) have completely changed the way in which mental health professionals handle confidential issues.

Numerous articles have been written about the issues of ethics, confidentiality, duty to warn, and medical/legal aspects of this element of practice. Although no clear consensus has been reached, recommendations are that patients who are sexually active and infected with HIV be counseled about potential risk to their sexual partners. Additionally, known partners should be notified of exposure risk and potential infection as well. Partner notification has been an extremely hotly debated topic. However, many states have developed legislation requiring or allowing either physicians or health department officials to notify partners of HIV-infected patients of their risk. The current standard, despite the controversy, appears to be an obligation on the part of health care professionals to ensure the notification of anyone who could be construed as clearly at risk and who may be unaware of their risk.

A particularly difficult situation is that of sex workers, known to be HIV-infected and known to be working actively as prostitutes. There are public health issues that pose a risk both for these patients and depending on the politics of the circumstances for their potential partners, clients, customers, victims, or victimizers. The responses to this problem have ranged from a sense that sex care workers and their clients can make their own decisions and should be responsible for their own behavior all the way to the sentiment that HIV-infected sex workers should be arrested and jailed for attempted murder. It has additionally been noted that some sex workers are impaired by a variety of psychiatric conditions including cognitive impairment, major mental illness, personality disorder, and substance use disorders. These may further contribute to the sense that some sex workers may be less than fully responsible for their behavior. Recommendations have been made for voluntary and involuntary interventions regarding these patients. Specific psychiatric interventions regarding competency, ability to consent, capacity, and most importantly treatment for the conditions that impair such people are critical to the mental health needs of patients with HIV.

Capacity to Consent/Competence

Patients with psychiatric disorders in HIV clinics often have a variety of difficulties with medical care provision. In some of these settings, the patient's capacity or competence to make medical decisions regarding their health care can be in doubt. Provision of mental health care evaluations to determine this is an often unmet need in HIV clinics. The issues of competence and capacity in these patients are often no different than those described elsewhere in this textbook; however, the consequences of inadequate assessment of patients' competency and capacity can have grave consequences in this clinical setting. HIV dementia and delirium can be overlooked in this population, as can intoxication with substances, creating problems in obtaining meaningful consent.

The question of capacity and consent involves several issues, but most important is the specific question that the patient is being asked. In many cases, medical care providers want a judgment about competence as a general rule, a judgment impossible to make. To have capacity to give a particular informed consent, a patient must understand that there is a decision before him/her regarding some aspect of care, and must understand the consequences not only of each option but also of refusal to make a choice. The patient should be able to repeat the benefits and risks of each possible option and must be able to clearly communicate the decision and have the ability to maintain the choice made over time. Finally, in the process of making the decision, the patient must be able to manipulate the information involved in a rational way. Thus, patients who are delirious or intoxicated may not maintain choices over time as concentration and memory are often impaired. Demented patients may have difficulty maintaining choices due to impaired memory as well. Psychotic patients may not base decisions in reality, occasionally arriving at choices that derive from delusional constructs. While physicians from many disciplines may be able to determine capacity in many situations, it is in these patients with psychotic illness that psychiatrists are often especially helpful.

In many cases, dangerousness, patterns of prior behavior, severity of illness, poor judgment, and psychiatric vulnerabilities complicate these decisions and play an important role in tempering the way in which patients are managed. The ethics of a particular case may get very complex when a patient understands the issues in a cognitive way, but their judgment is colored by their affective state, temperament, drug cravings, social situation, or simply difficulty with tolerating discomfort. These cases often divide medical teams and require consultation and a group conference to resolve. It is critical to get all providers to discuss the most difficult cases, clarify the issues, and come to a decision based on the patients' best interests, not the most expedient management.

Adherence Counseling

The single most important factor regarding outcome of HIV treatment is

the patient's ability to adhere to prescribed regiment. While this has been debated in literature, a recent study by Margaret Fischel showed looking at HIV-infected prisoners revealed that 100 percent of patients who received directly observed therapy in a prison setting developed undetectable viral loads. This strongly supports adherence as the major feature of treatment. There are compelling studies suggesting that major depression, substance abuse, personality disorder, and psychosocial disruption all affect adherence. Intervention in these conditions is presented above. More subtle factors effecting adherence include psychosocial support networks, individual coping skills, life structure, access to resources, and behavioral control. Interventions such as cognitive behavioral psychotherapy, structured psychoeducational psychotherapy, supportive psychotherapy, and group interventions have all been used to improve patient adherence to office visits and medication regimens.

In the current literature on HIV, medication adherence focuses on technical interventions such as pill box and timer reminders, less complex pharmacological interventions, decreased pill burdens, and increased access to care. A growing literature examines psychosocial interventions, relationship with health care providers, case management, and psychiatric disorders as barriers to adherence. It is in this arena that mental health care can have an enormous impact on outcome. Psychotherapy has been shown to improve clinic visit adherence, the best indirect predictor of medication adherence.

NEUROLOGICAL COMPLICATIONS OF HIV AND AIDS

Opportunistic Infections

Toxoplasmosis. *Toxoplasma gondii is* a protozoan acquired most commonly from cat feces or uncooked meat. Infection generally occurs in patients with less than 200 CD4 cells/mL. In patients with AIDS, toxoplasmosis is the most common reason for intracranial masses, affecting between 2 and 4 percent of the AIDS population. Other manifestations are possible, including hepatosplenomegaly, myositis, pneumonitis, myocarditis, and maculopapular rash. Lymphadenopathy may be present in cutaneous cases. Symptoms of CNS infection are fever, change in level of alertness, headache, focal neurological signs (approximately 80 percent of cases), and partial or generalized seizures (approximately 30 percent of cases). CT and MRI scans usually show multiple, *ring-enhancing lesions* in the basal ganglia or at the gray–white matter junction. CSF studies are normal in 20 to 30 percent of cases, but more often show a mild monocytosis. Serum *T. gondii* IgG is generally helpful in the diagnosis but has a false-negative rate of 5 to 10 percent. Brian biopsy provides the definitive diagnosis, but since this invasive procedure carries some risks, empiric treatment is often offered if the

clinical and radiographic pictures suggest infection.

Treatment consists of pyrimethamine plus sulfadiazine or clindamycin. Clinical and radiological improvement is seen in over 85 percent of patients by day 7. Since these medications are effective only against the tachyzoite form of the protozoan, they must be continued for a full 6 weeks and then prophylaxis, usually with the treating agents, must be prescribed to prevent recrudescence. The use of trimethoprim–sulfamethoxazole as prophylaxis has reduced the incidence of *T. gondii* infection. Patients with hypersensitivity to sulfa drugs may use pyrimethamine plus dapsone.

Cytomegalovirus. CMV infection is found at autopsy in about 30 percent of brains from HIV-infected patients. However, the development of clinically evident CMV encephalitis is fairly rare, and most often occurs in patients with CD4 counts less than 50 cells/μL. Of particular note, CMV infection of another tissue, such as retina, blood, adrenal glands or GI tract, is often found at the time of encephalitis.

There are two distinct syndromes of CMV CNS infection. The first and more common is encephalitis with dementia, which presents with subacute onset accompanied by periods of delirium, confusion, apathy, and focal neurological deficits. The second is a ventriculoencephalitis, in which CMV infects the ependymal cells lining the ventricles, causing a rapid progression from delirium to death, with cranial nerve deficits and ventriculomegaly developing quickly.

Investigation of CMV encephalitis begins with examination for signs of CMV infection of the retinae, electrolyte studies to look for adrenal insufficiency, and viral blood cultures. CT scan may show ventriculomegaly or decreased attenuation diffusely throughout the parenchyma. MRI may show increased signal intensity around the ventricles. CSF studies may be normal or show high protein, low glucose, and pleocytosis. CSF CMV cultures are usually negative but PCR may reveal the presence of the virus. Brain biopsy provides a definitive diagnosis.

Treatment is mostly supportive. Ganciclovir and foscarnet may be prescribed, but are of questionable benefit. Trials of a promising new medication, cidofovir, are underway.

Cryptococcal Meningitis. While meningitis caused by *Cryptococcus neoformans* is rare in immunocompetent persons, it occurs in approximately 8 to 10 percent of patients with AIDS and may be devastating. Patients generally present with fever and delirium. In contrast, meningeal signs (headache, stiff neck, photophobia and nausea) are not universally seen. Seizures and focal neurological deficits occur in about 10 percent of patients. CT scans are normal, but gadolinium-enhanced MRI may show meningeal inflammation. Intracranial pressure is elevated in 50 percent of patients. CSF studies are normal in about 20 percent, but

otherwise show mild-to-moderate monocytosis, elevated protein, decreased glucose, and positive fungal cultures. The fungus can be seen on India ink stain of CSF about 60 to 80 percent of the time. There is also a test for *C. neoformans* antigen, which is usually positive in both serum and CSF.

Treatment of cryptococcal meningitis requires amphotericin B and flucytosine. Patients who survive must receive prophylaxis against recurrence, since this is very common. Some authors suggest that patients who receive HAART for 6 months with a rise in CD4 count to >100 cells/mL may terminate secondary prophylaxis. Prophylaxis can be prescribed as oral fluconazole or intermittent IV amphotericin B. Primary prophylaxis for *C. neoformans* is not recommended.

Progressive Multifocal Leukoencephalopathy. PML is a demyelinating disease of white matter in immunocompromised patients. First described in cancer patients, the causative agent is a polyomavirus, named JC virus after a patient (not to be confused with Creutzfeldt–Jakob disease, caused by a prion). Its transmission route is unclear but may be respiratory, and there is no known clinical syndrome of acute infection. The prevalence of PML in AIDS is between 1 and 10 percent of patients, while patients with AIDS account for almost three quarters of PML cases seen in the United States. Typically, PML affects patients with AIDS with fewer than 100 CD4 cells/μL.

The pathology of PML consists of demyelination and death of astrocytes and oligodendroglia, with a multifocal presentation. The clinical syndrome consists of multiple focal neurological deficits, such as mono- or hemiparetic limb weakness, dysarthria, gait disturbances or sensory deficits, and progressive dementia, with eventual coma and death. Occasionally there may be seizures or visual losses. There is usually no fever or headache.

MRI is more useful than CT in diagnosis, displaying multiple areas of attenuated signal on T2 images primarily in the white matter of brain, although gray matter, brain stem, cerebellar, and spinal cord lesions are possible. CSF studies are generally unhelpful, except for PCR evaluation for the presence of JC virus, which is sensitive and specific. Brain biopsy provides the definitive diagnosis, but is rarely used. There is no specific antiviral therapy for JC virus. Treatment of PML includes support of the patient and HAART. There are no data suggesting higher CNS-penetrating antiretrovirals are of any particular benefit, but this logical conclusion is often followed by providers.

CNS Neoplasms

Lymphoma is the most common neoplasm seen in patients with AIDS, affecting between 0.6 and 3 percent. AIDS is the most common condition associated with primary CNS lymphoma. The patient is generally afebrile

and may develop a single lesion with focal neurological signs or small, multifocal lesions most commonly presenting with a mental status change. Seizures are present in about 15 percent of patients. CNS lymphoma is at times misdiagnosed as toxoplasmosis, HIV dementia, or other encephalopathy. CT scan of the brain may be normal or show multiple hypodense or patchy, nodular enhancing lesions. MRI generally shows enhanced lesions that may be difficult to differentiate from CNS toxoplasmosis, but thallium SPECT scanning may help differentiate the two disorders and is 90 percent sensitive and specific for lymphoma. CSF studies may be normal or show a moderate monocytosis; cytology studies reveal lymphoma cells in less than 5 percent of patients. Brain biopsy is required for the confirmation of the diagnosis of CNS lymphoma. As this procedure carries some morbidity, clinicians should weigh the clinical presentation carefully, suspecting lymphoma in afebrile patients with a negative toxoplasma IgG screening test, patients with a single lesion, and patients who fail to respond to empiric therapy for toxoplasmosis as demonstrated by clinical exam and repeat MRI at 2 weeks. The differential diagnosis of CNS neoplasm also includes metastatic Kaposi sarcoma and primary glial tumors. Lymphoma may respond in part to radiation therapy and steroids, thus alleviating high intracranial pressure and its associated symptoms. Chemotherapy is generally adjunctive for lymphoma. While CNS lymphoma had a grim prognosis with an average survival of 3 to 5 months prior to the advent of HAART, the prognosis is now dependent on the HAART response, with considerable improvement possible in patients who respond to HAART.

Direct CNS Manifestations of HIV

1. Guillain–Barré syndrome

   A small percentage of patients, usually young men, will present with Guillain–Barré syndrome associated with early HIV infection. Guillain–Barré syndrome is an inflammatory demyelinating polyneuropathy causing symmetrical paralysis and few if any sensory symptoms, usually beginning in the lower extremities and progressing upward. The condition becomes especially serious if abdominal musculature is involved, as it may impair respiration. The disorder is thought to be autoimmune in etiology and is generally self-limited. IV immunoglobulin and plasmapheresis have been used to shorten the course, but neither treatment has been studied well in HIV-infected individuals.
2. Vacuolar myelopathy

   Vacuolar myelopathy is highly prevalent among patients with AIDS, being found in approximately 50 percent of patients at autopsy. Clinical manifestation of this disease is much less common, affecting 20 to 30 percent of end-stage patients with AIDS. The presence of vacuolar

myelopathy has been associated with history of *P. carinii* and *Mycobacterium avium-intracellulare* infections, suggesting that the development of vacuolar myelopathy is related to more severe immunosuppression. The mechanism of the disease is unclear, but appears similar to the myelopathy of combined systems disease associated with vitamin B_{12} deficiency. Multinucleated giant cells are seen on histological examination, and theories about mechanism focus on immunological activation damage, direct toxicity of HIV products, and metabolic dysfunction of transmethylation processes.

The clinical manifestations of vacuolar myelopathy appear when the disease progresses to affect the lateral and posterior columns, and thus includes spastic paraparesis, loss of proprioception and vibration sense, bowel and bladder urgency or incontinence, and impotence. There are now reports of improvement with the initiation of CART.

Peripheral Nervous System Disorders in HIV

Peripheral Neuropathy. Patients with HIV infection may develop peripheral neuropathy, most often in the feet but occasionally in the hands. The neuropathy may range from paresthesia to burning pain, and patients will have a vibratory-sense gradient with decreased sensation in the distal extremity compared to more proximal points. The incidence of peripheral neuropathy increases as HIV disease progresses, but cases are found in patients with well-preserved immune function. Treatment of peripheral neuropathy may include TCAs, pregabalin, gabapentin, or other antiepileptic drugs used to treat neuropathic pain. Opiate analgesics should be used sparingly, and longer acting agents are preferable, since long-term use of opiates presents the risk of eventual tolerance and dependence. Benzodiazepines are of no use.

Special Issues in HIV

Fatigue. Fatigue is a common symptom in HIV-infected patients, which is often overlooked, improperly assessed or inadequately investigated. Several authors have commented on the high prevalence of fatigue as a symptom of HIV infection, especially in later stages. Fatigue may be mild and annoying, or it may be severe enough to impair function. Several scales have been published to assess fatigue symptoms and severity. Fatigue is a nonspecific symptom and may have a single or multifactorial etiology. Medical causes include pneumonia, bronchitis, hypothyroidism, hepatitis, heart failure, renal failure, many cancers, and myopathy. In a sample of ambulatory patients with AIDS, fatigue is significantly correlated with anemia and pain. In addition to disease causes, patients may present with fatigue as a side effect of medications, such as antihypertensives, anticonvulsants, benzodiazepines,

antidepressants, narcotic analgesics, antipsychotics, antiemetics, antihistamines, and most importantly for patients with HIV, HAART. In fact, fatigue has been found to be one of the most common side effects of protease inhibitors and may be a reason for nonadherence.

Fatigue may also be the result of psychiatric disorder. Alcohol and substance use disorders may lead to fatigue, either related to the use or withdrawal of the substance, or as a symptom of demoralization in addicts. Most importantly, fatigue is caused by major depression. Patients with HIV with major depression are much more likely to complain of fatigue than patients without depression. Many depression screening tools, such as the BDI and Hamilton Depression Rating Scale, have not been very useful in distinguishing fatigue from major depression, usually because fatigue symptoms are present on the screening tools.

In general, the evaluation of a patient complaining of fatigue should include a careful history of its temporal characteristics, severity, and associated symptoms. It should also include careful review of current and recent medications, physical examination, and a mental status examination. The latter should carefully examine for anhedonia, diminished sense of self-worth, guilty feelings, sleep disturbance—especially early morning awakening with inability to return to sleep, appetite changes—especially a recent more than 5 percent change in body weight, thoughts of death or suicide, and impairments in concentration or memory. Certain laboratory studies should also be obtained, including complete blood count, electrolytes, liver tests, oxygen saturation, and thyroid function tests.

If fatigue is thought to be related to a medical or medication cause, all attempts should be made to treat the illness or modify the medication so as to alleviate the fatigue. In this context, testosterone is a successful treatment of fatigue in HIV-infected men, even when depressive symptoms are present. Of course, a clinical major depression should be treated with standard therapies as discussed above. More activating antidepressants, such as fluoxetine, escitalopram, venlafaxine XR or bupropion SR or XL, may be better tolerated by fatigued depressed patients. Some authors have reported that dextroamphetamine may be useful in treating fatigue and depression in HIV. Care must be exercised in using stimulants as long-term use may lead to dependence or worsening depression on some occasions.

HIV/HCV Coinfection

Hepatitis C virus (HCV) is a blood-borne pathogen that is currently most commonly transmitted by injection drug use, but may be transmitted sexually, although far less commonly than HIV. Some clinics have reported that 50 percent of HIV-infected patients are also infected with HCV. The natural history of HCV infection in HIV-negative individuals is that 15 percent of patients clear the infection after the acute phase, while 85

percent progress to a chronic infection. Hepatic fibrosis develops, often requiring about 10 years to reach significant levels, with cirrhosis following about 20 years from time of infection. Chronic HCV infection is the most common etiology of hepatocellular carcinoma (HCC), which usually develops about 30 years after infection in HIV-negative patients. Unfortunately, HIV infection is likely to make individuals more susceptible to contract HCV if exposed, likely due to immunosuppression, and also to cause more rapid progression of liver disease, cutting the above approximate timetable in half (i.e., fibrosis in 5 years, cirrhosis in 10, HCC in 15). The leading causes of mortality in HIV-infected patients are hepatic diseases, most often related to HCV infection.

Chronic hepatitis C is associated with increased rates of depression. Additionally, treatment with interferon–ribavirin combinations increase major depression and can produce mania. In particular, it appears that interferon-alpha has been associated with depressive syndromes, suicide, and, on rare occasions, mania. Patients with pre-existing depression or bipolar disorder are more likely to develop affective symptoms while receiving the drug, but may not be more likely to stop treatment than patients developing these symptoms de novo. Further, depressive symptoms associated with interferon-alpha have been successfully treated with both SSRIs and TCAs. Because of the development of direct acting antivirals that target and cure hepatitis C in greater than 90 percent of patients without interferon, the authors are optimistic that this will be of historical interest, possibly by the time you are reading this chapter.

Coinfected HIV/HCV patients should be screened for the presence of psychiatric disturbance like any other patient with HIV, but special monitoring should be performed during the period of treatment with interferon-alpha for the purpose of early recognition and treatment of affective symptoms. Alcohol hastens the progression of HCV disease and should be strongly discouraged in any amount. Drug use other than alcohol may exacerbate neuropsychiatric side effects of interferon-alpha, and patients should be stabilized prior to this antiviral treatment. While there are yet no data, methadone maintenance may be a good option for patients who cannot achieve abstinence from opiates but need to start interferon therapy due to precipitous declines in liver function.

CONCLUSIONS

HIV and AIDS are conditions intimately linked to psychiatry. In a sense, psychiatric disorders can be seen as vectors of HIV transmission and additionally complicate the treatment of HIV. Also, HIV produces a number of psychiatric conditions and exacerbates many others. The authors have attempted to show the intense comorbidity and links between various types of psychiatric conditions. The way depression exacerbates addictions, the way personality disorder exacerbates addictions, and the

way in which addictions exacerbate both personality vulnerabilities and depression. HIV disease is driven by behaviors that are intimately connected with all these conditions.

The authors see HIV as a model for the way in which psychiatry needs to speak to the rest of medicine about the role of psychiatry in general medicine and health care. It is a sad symptom of the problems in US health care that there is abundant data showing the need for a psychiatric presence in the every phase of HIV care, and yet the poverty of funding and availability of psychiatric care in HIV clinics.

The authors' experience in caring for patients with HIV is that by developing a comprehensive diagnostic formulation on which to base treatment, they have significant success with even difficult patients. The authors' formulation includes disease syndromes such as major depression and schizophrenia, personality vulnerabilities such as unstable extraversion, behavioral disorders such as addictions, and problems of life experience such as trauma and trust issues. Each problem has the potential to sabotage treatment for all the remaining conditions. The treatment plan must be comprehensive in scope in order to address the whole person.

ACKNOWLEDGMENT

The authors gratefully acknowledge the assistance of Dr. Joyce King for her editorial efforts and insightful criticisms.

REFERENCES

Angelino AF, Treisman GJ. Management of psychiatric disorders in patients infected with human immunodeficiency virus. *Clin Infect Dis*. 2001;33(6):847–856.

Antinori A, Arendt G, Becker JT, et al. Updated research nosology for HIV-associated neurocognitive disorders. *Neurology*. 2007;69(18):1789–1799.

Arseniou S, Arvaniti A, Samakouri M. HIV infection and depression. *Psychiatry Clin Neurosci*. 2014;68(2):96–109.

Batkis MF, Treisman GJ, Angelino AF. Integrated opioid use disorder and HIV treatment: rationale, clinical guidelines for addiction treatment, and review of interactions of antiretroviral agents and opioid agonist therapies. *AIDS Patient Care STDS*. 2010;24(1):15–22.

Breitbart W, McDonald MV, Rosenfeld B, Monkman ND, Passik S. Fatigue in ambulatory AIDS patients. *J Pain Symptom Manage*. 1998;15:159–167.

Brooner RK, Greenfield L, Schmidt, CW, Bigelow GE. Antisocial personality disorder and HIV infection among intravenous drug users. *Am J Psychiatry*. 1993;150:53–58.

Ciesla JA, Roberts JE. Meta-analysis of the relationship between HIV infection and risk for depressive disorders. *Am J Psychiatry*. 2001;158:725–730.

Cournos F, Guido JR, Coomaraswamy S, Meyer-Bahlburg H, Sugden R, Horwath E. Sexual activity and risk of HIV infection among patients with schizophrenia. *Am J Psychiatry*. 1994;151:228–232.

Durand CM, Flexner C. HIV cure: knocking on the door. *Clin Pharmacol Ther*. 2013;93(5):382–

384.

Eldred L, Malitz F. Introduction to the supplemental issue on the HRSA SPNS Outreach Initiative. *AIDS Patient Care STDS*. 2007;21(Suppl 1):S1–S2.

Erbelding EJ, Stanton D, Quinn TC, Rompalo A. Behavioral and biologic evidence of persistent high-risk behavior in an HIV primary care population. *AIDS*. 2000;14:297–301.

Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clin Infect Dis*. 2011;52:793–800.

Golding M, Perkins DO. Personality disorder in HIV infection. *Int Rev Psychiatry*. 1996;8:253–258.

Heaton RK, Clifford DB, Franklin DR Jr, et al. HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy. CHARTER study. *Neurology*. 2010;75:2087–2096.

Himelhoch S, Medoff DR. Efficacy of antidepressant medication among HIV-positive individuals with depression: a systematic review and meta-analysis. *AIDS Patient Care STDS*. 2005;19:813–822.

Himelhoch S, Powe NR, Breakey W, Gebo KA. Schizophrenia, AIDS and the decision to prescribe HAART: results of a national survey of HIV clinicians. *J Prev Interv Community*. 2007;33:109–120.

Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*. 1995;373:123–126.

Hutton HE, Treisman GJ, Hunt WR, et al. HIV risk behaviors and their relationship to posttraumatic stress disorder among women prisoners. *Psychiatr Serv*. 2001;52:508–513.

Ickovics JR, Hamburger ME, Vlahov D, et al. Mortality, CD4 cell count decline, and depressive symptoms among HIV-seropositive women: Longitudinal analysis from the HIV epidemiology research study. *JAMA*. 2001;285:1466–1474.

Lucas GM, Chaudhry A, Hsu J, et al. Clinic-based treatment of opioid-dependent HIV-infected patients versus referral to an opioid treatment program: A randomized trial. *Ann Intern Med*. 2010;152(11):704–711.

Lyketsos CG, Hanson A, Fishman M, McHugh PR, Treisman GJ. Screening for psychiatric morbidity in a medical outpatient clinic for HIV infection: the need for a psychiatric presence. *Int J Psychiatry Med*. 1994;24:103–113.

Lyketsos CG, Hoover DR, Guccione M, et al. Changes in depressive symptoms as AIDS develops. *Am J Psychiatry*. 1996;153:1430–1437.

Lyketsos CG, Schwartz J, Fishman M, Treisman G. AIDS mania. *J Neuropsychiatr Clin Neurosci*. 1997;9(2):277–279.

Michael NL, Vahey M, Burke DS, Redfield RR. Viral DNA and mRNA expression correlate with the stage of human immunodeficiency virus (HIV) type I infection in humans: evidence for viral replication in all stages of HIV disease. *J Virol*. 1992;66:310–316.

Rutter M. Temperament, personality and personality disorder. *Br J Psychiatry*. 1987;150:443–458.

Sharp PM, Hahn BH. The evolution of HIV-1 and the origin of AIDS. *Philos Trans R Soc Lond B Biol Sci*. 2010;365(1552):2487–2494.

Strathdee SA, Palepu A, Cornelisse PG, et al. Barriers to use of free antiretroviral therapy in

injection drug users. *JAMA*. 1998;280(6):547–549.

Tobst KK, Wiggins JS, Costa PT Jr, Herbst JH, McCrae RR, Masters HL III. Personality Psychology and Problem behaviors: HIV risk and the Five-Factor Model. *J Personality*. 2000;68:1232–1252.

Treisman G, Fishman M, Schwartz J, Hutton H, Lyketsos C. Mood disorders in HIV infection. *Depress Anxiety*. 1998;7:178–187.

Wagner GJ, Rabkin JG, Rabkin R. Testosterone as a treatment for fatigue in HIV+ men. *Gen Hosp Psychiatry*. 1998;20:209–213.

Watkins CC, Pieper AA, Treisman GJ. Safety considerations in drug treatment of depression in HIV-positive patients: an updated review. *Drug Saf*. 2011;34(8):623–639.

▲ 2.9 Neuropsychiatric Aspects of Other Infectious Diseases (Non-HIV)

BRIAN A. FALLON, M.D., AND SHANNON L. DELANEY, M.D.

The link between severe neuropsychiatric disorders and infectious disease was first established in the early 1900s by the identification of a spirochete as the cause of syphilis and reinforced in the 1920s after severe neurobehavioral syndromes were observed among people affected by the viral influenza epidemic. At times, the link between an infectious agent and a neuropsychiatric disorder is strong, as in the case of rabies or the current human immunodeficiency virus (HIV) and Lyme disease epidemics. At other times, the link is less clear but strongly suspected, as has been true for chronic fatigue syndrome (CFS). The establishment of a link is best confirmed by demonstration of the organism at the time of the onset of the neuropsychiatric disorder and by the demonstration of anatomic and functional pathology. This level of confirmation is not always possible, however, as the infecting agent may be undetectable given the sensitivity limitations of current biotechnology or because the infection itself has adaptive survival mechanisms that enable it to evade detection. Alternatively the infectious agent, although no longer present to cause direct effects, may have triggered an autoimmune process, altered neurotransmitter function, or provided an aberrant influence at a critical developmental period that exerts long-lasting effects.

Infectious agents may affect the central nervous system (CNS) directly or indirectly. Direct involvement by a neurotropic agent may result from attachment of the microbe to neuronal tissue, eliciting a local inflammatory response and immediate dysfunction, or by integration of the microbial genome into the cellular DNA, resulting in long-term alterations in human brain function in the adult or in altered development of neuronal and glial cells in utero.

Alternatively, an infection outside of the CNS may work indirectly, such as by triggering the cells of the innate immune response, primarily macrophages and dendritic cells, to initiate an inflammatory response leading to the synthesis and release of cytokines or by toxins released by the peripheral infection itself. Both the circulating cytokines and the bacterial toxins for example can exert their effect on the CNS through humoral and neural routes of communication. Cytokines circulating in the blood can enter the brain by activating cytokine production at the CNS vascular endothelium, thus disrupting the blood–brain barrier. In addition, cytokines as well as other neural activators such as infection-triggered antineuronal antibodies can reach the brain directly by leaking across the blood–brain barrier via the highly vascularized circumventricular organs.

As an alternative to the humoral route of communication, peripheral pathogens may trigger the brain outside the CNS via neural pathways, such as when local inflammation leads to the activation of the afferent vagus nerve through which neural signals are communicated first to the lower brain stem and then to the higher brain centers, such as the hypothalamus and amygdala. CNS inflammation from peripheral microbes therefore can occur not by direct infection but by humoral and neural pathways that then activate central cytokine secretion by microglia and other neural cells.

That microbes may have indirect effects through their impact on the host-determined cellular, humoral, or cytokine immune responses has been well demonstrated. For example, activation of proinflammatory cytokines or the induction of nitric oxide in an adult brain may lead to neuronal and behavioral dysfunction. Behavioral dysfunction as a result of inflammation may be manifest by the typical symptoms associated with depression, such as anhedonia, sadness, fatigue, altered sleep patterns, anorexia, and unexplained pains; the cause–effect relationship has been demonstrated by the temporal relationship between the exogenous administration of proinflammatory cytokines and depressive symptoms. Inflammation that occurs in utero can harm the developing embryo leading to inhibited dendritic development in cortical neurons with later neurobehavioral consequences. The quality and intensity of the immune response, modulated by genetic factors, may be perpetuated by the continued presence of a viable organism, a piece of a nonviable organism, or a misdirected cross-reactive autoimmune process that was initiated by prior infection. The immune response in its effort to protect may thereby provoke neuropsychiatric disorders.

This chapter, however, will focus primarily on those organisms that are known to impact the CNS directly. A variety of infectious agents have been examined over the past two decades as possible causes of neuropsychiatric disorders. These include bacteria, viruses, yeast, protozoa, and protozoa. Although these organisms may induce neuropsychiatric disorders in some individuals after infection, this outcome is variable, largely determined by a complex interplay between host response genes, infectious agent, and timing in neurodevelopment. In certain heritable neuropsychiatric illnesses, such as schizophrenia and mental retardation, epidemiologic studies reveal that the clinical severity, age at onset, or treatment response may be largely determined by environmental agents, such as infection. The expression of the disease is modulated by when in the course of development the exogenous agent had its influence. For example, maternal exposure during pregnancy to poliovirus, retrovirus, influenza, measles, rubella, varicella zoster, and bacterial agents has been associated with an increased risk of schizophrenia in the offspring.

A variety of medical disorders previously thought to be outside the domain of microbial etiology have now been linked to concurrent

infections, such as *Chlamydia,* contributing to atherosclerotic heart disease, and *Helicobacter pylori,* contributing to gastric and duodenal ulcers. On the neuropsychiatric front, compelling evidence links streptococcal infection with the onset of obsessive–compulsive disorder (OCD) and tic disorders in susceptible children and borrelial infection with the onset of irritability, mood swings, and cognitive problems. The search for infectious causes of neuropsychiatric disorders is a logical enterprise given the increasing recognition of the importance of environmental factors in the development of psychiatric disorders.

SPIROCHETAL DISEASES

Under the umbrella of the order of spirochetes are three agents that are known to invade the CNS: *Borrelia, Treponema,* and *Leptospira*. *Borrelia,* which requires an arthropod vector and a mammalian or bird reservoir, is commonly known to cause relapsing fever and Lyme disease. *Treponema,* which is spread person to person and does not use an arthropod vector, is the spirochete responsible for syphilis. *Leptospira,* which is spread by contaminated water, is the agent of Weil disease, which can have CNS manifestations.

Lyme Disease (Lyme Borreliosis)

The agent of Lyme disease, *Borrelia burgdorferi,* is transmitted by the bite of an infected *Ixodes* tick and can induce a multisystemic illness in the human host. While early treatment with antibiotics at the time of the skin rash (erythema migrans) often results in rapid resolution of the illness, delayed treatment may result in a more entrenched set of symptoms and a longer course of illness. Two terms are commonly used to refer to the problem of persistent symptoms after antibiotic treatment: post-treatment Lyme syndrome and chronic Lyme disease. “Post-treatment Lyme syndrome” is used by the academic community to refer to those patients who experience persistent musculoskeletal, cognitive, or fatigue symptoms after antibiotic treatment of well-defined prior Lyme disease. “Chronic Lyme disease” is a term used most often by patients to refer to any clinically significant symptoms that remain after treatment of known Lyme disease. These terms create much distress within the community, as physicians and patients debate over the cause of the persistent symptoms. Academics tend to avoid the term “chronic Lyme disease” as to them it appears to imply the infection persists. While the exact cause of persistent symptoms after treatment is unclear and most likely multidetermined, there is compelling evidence to indicate that some patients have residual infection while others have a postinfectious disorder caused either by prior damage from the spirochete or by ongoing secondary immune effects (such as molecular mimicry or a cytokine response that continues unabated). The

natural biology of the *Borrelia* spirochete in the animal host (e.g., mouse, dog, rhesus macaque) is one of persistence; indeed, several studies have confirmed that even after aggressive antibiotic therapy, a small number of less virulent organisms persist in the animal host. These observations have led to in vitro research demonstrating that antibiotic treatment of *B. burgdorferi* may require two different sets of antimicrobials: one that treats the spirochetal growing form and the other that treats the dormant stationary "round" forms; human trials have not yet been conducted to examine whether these bench findings are directly translatable to the human situation.

Unfortunately current tests are not sufficiently sensitive or specific to clarify whether a patient has persistent infection or not. The cloud of medical uncertainty hovering over patients with persistent symptoms after treated Lyme disease leads to variety of physician reactions, ranging from dogmatic certainty that the prior Lyme infection has nothing to do with the current symptoms, to an unwillingness to accept such patients into his/her practice due to the uncertainties about how to help, or to an overly confident assertion that a patient has active Lyme disease despite lack of exposure to Lyme-endemic areas, negative blood tests, and poor response to prior antibiotic treatment.

In case reports and larger series, Lyme disease has been reported to cause a vast array of neuropsychiatric disorders, ranging from the common presentations of depression/irritability and cognitive problems (short-term memory loss, verbal fluency problems) to the much less common manifestations of psychosis, mania, or dementia. Lyme disease has been reported throughout the United States and in over 80 countries throughout the world. The spirochete, *B. burgdorferi,* is initially inoculated into the skin by an infected tick, typically inducing a local expanding erythematous rash (sometimes appearing as a bull's eye), known as erythema migrans, which is recalled by approximately two-thirds of infected patients. Rapidly disseminated by the blood stream throughout the body, *B. burgdorferi* has been found in the CNS as soon as 2 weeks after initial skin infection. Known to be neurotropic, *B. burgdorferi* may reside in the CSF or adhere to glial cells or other brain tissue. Like its spirochetal counterpart, *Treponema pallidum, B. burgdorferi* may remain latent, causing illness months to years later. Partly because of this latency in disease expression, patients may be unable to recall the initial tick bite or rash. *B. burgdorferi* may persist in the human host by evading normal immune surveillance through a number of mechanisms, including antigenic variation or up- or downregulation of its immunogenic surface proteins, lodging in less accessible areas like the extracellular matrix or brain, and/or inducing the production of both anti-inflammatory cytokines and surface proteins that enable the spirochete to resist complement mediated killing.

Diagnosis. The epidemiologic surveillance criteria for the diagnosis of Lyme disease in the United States require a history of exposure to a Lyme-endemic area and either a physician-diagnosed erythema migrans rash or serologic evidence of exposure to *B. burgdorferi* and at least one of the following three clinical features: (1) arthritis; (2) neurologic signs (cranial or peripheral neuropathy, meningitis, encephalomyelitis, or encephalitis with evidence of intrathecal antibody production); or (3) cardiac conduction defects. Although useful for epidemiologic monitoring, these criteria are unduly restrictive for clinical purposes, as approximately 18 percent of patients may present with diffuse myalgias and arthralgias but not manifest any of the objective signs of Lyme disease. Further complicating the diagnosis is the fact that serologic tests are helpful but not perfect. False-positive results, particularly on the whole cell sonicate enzyme-linked immunosorbent assay (ELISA) or immunoglobulin M (IgM) Western blot, might result because of cross-reactivity with other micro-organisms. This is particularly a problem among patients with a history of mononucleosis or syphilis. False-negative results may occur because the patient is tested too soon after infection before an appropriate antibody response is mounted or because the patient's immune response has been abrogated by early antibiotic treatment. For these reasons, a rational approach to the diagnosis of Lyme disease must be based on the clinical presentation primarily, followed by the supportive evidence supplied by laboratory tests. Laboratory tests include indirect tests such as the ELISA and Western blot and direct tests such as the polymerase chain reaction (PCR) for borrelial DNA or antigen detection assays. A newer ELISA, based on an invariant C6 region, is a highly specific test for Lyme disease. Bands of particular significance on the Western blot include the ones identified by the Centers for Disease Control (CDC) as being most frequent and specific, as well as the 31kD (OspA) and 34 kD (OspB) bands. To reduce the risk of false positives due to cross-reactive antibodies, some diagnostic labs are now using recombinant proteins that are more *B. burgdorferi*-specific. The PCR assay, although highly specific for *B. burgdorferi* DNA, rarely is positive on blood testing of individuals beyond the stage of the rash because the organism is tissue tropic and does not reside for prolonged periods in the blood stream. For similar reasons and because the organism is so slow-growing (weeks to months), culture is not often used.

Clinical Manifestations. The expanding erythema migrans rash is the hallmark feature of early Lyme disease. Antibiotic treatment at this stage often results in cure. Because patients may not recall or see the rash, the flu-like symptoms that often occur shortly after the rash may be ignored, only to be followed several months to years later by the emergence of a multisystem disease affecting the joints, the heart, the eyes, and/or the CNS or peripheral nervous system. Fifteen to forty percent of patients may

have neurologic signs as their presenting feature. Headaches may be followed by meningitis, cranial neuritis, peripheral neuritis, a radiculitis (motor weakness or inflammation of the nerve roots with lancinating radicular pain), and/or encephalitis characterized by mood lability and disturbances of memory or sleep. Although suggestive of Lyme disease, a facial palsy (cranial nerve VII) may occur in only 5 to 10 percent of a sample of patients with neurologic Lyme disease. Symptoms of peripheral nerve involvement include sharp stabbing pains, areas of numbness, burning or tingling, weakness, and fasciculations.

In patients with CNS involvement, formal neuropsychological testing may reveal impairment in short-term memory, processing speed, and/or verbal fluency. This cognitive impairment, although worsened by marked pain, severe fatigue, sensory hyperacusis, or mood disorders, exists independently of the number of physical symptoms or the severity of concurrent depression. Typical cognitive symptoms include word-finding problems, word substitutions, transient episodes of geographic disorientation, problems with fine motor control, difficulty with consistent short-term retrieval, and the sensation that one's brain is in a fog. Less commonly, the severity of the cognitive disturbance causes a global impairment, suggestive of a new onset dementia. In a recent study in Europe, 1.25 percent of patients diagnosed with dementia were found to have CSF evidence of neuroborreliosis, half of whom responded favorably to antibiotics. Although the full spectrum of psychiatric disorders has been associated with *B. burgdorferi* infection, by far the most frequent are disturbances of mood, characterized by irritability, dysthymia, and sleep loss. These are seen during the disseminated phase of the infection—not typically with erythema migrans. The majority of controlled studies in which patients with Lyme disease are compared to healthy controls or to patients with other illnesses reveal that depression occurs more frequently in the group with Lyme disease. Children with neurologic Lyme disease typically present with complaints of headaches as the most common symptom, followed by behavioral, attentional, or mood disturbance as the next most prevalent symptom. Among children with Lyme disease and headaches, a lumbar puncture may reveal elevated intracranial pressure (pseudotumor cerebri), which, in extreme cases, may result in damage to the optic nerves. Other less common neuropsychiatric aspects associated with Lyme disease include panic-like attacks associated with spontaneous palpitations, transient paranoia, illusions or hallucinations (visual, olfactory, auditory), depersonalization, OCD, agitated mania, and what appears to be personality change. Because of the multisystem involvement in Lyme disease and the frequent concurrence of anxiety and/or depression, patients may be mistakenly diagnosed as having a primary psychiatric or a somatoform disorder before Lyme disease is even considered. If Lyme disease is considered but the patient never developed

objective signs of Lyme disease previously, the somatoform label may once again be mistakenly applied.

A previously healthy 19-year-old man developed sudden onset irritability and paranoia. After medical clearance, he was hospitalized on a psychiatric unit and treated with antipsychotic medication; he had a partial reduction in paranoia but fear and irritability remained. In addition to these symptoms, he complained of marked fatigue, distal paresthesias, hypersensitivity to light and sound, and feeling that his brain was in a "fog." There was no family history of psychosis, no substance abuse, no history of learning disability, and no prodromal symptoms. Neurocognitive testing revealed an intelligent young man with deficits in short-term memory, information retrieval, processing speed, and verbal fluency but intact attention. Further medical work-up was recommended to rule out a medical cause, particularly Lyme disease as the cognitive pattern was characteristic of Lyme disease but less characteristic of depression or psychosis. The initial Lyme ELISA screening test was negative and so Lyme was considered ruled out. However, because he came from a Lyme-endemic area and had a history of tick bites, his parents pursued further workup with a doctor who conducted additional blood tests that revealed positive results on the Lyme C6 peptide ELISA and near positive results on the IgG Western blot (4 of 5 CDC-recommended bands). A spinal tap revealed mildly elevated protein, but white blood count (WBC) and Borrelia-specific antibodies were not present. The second doctor concluded that this young man clearly had prior infection with the agent of Lyme disease based on the positive C6 ELISA and the strongly suggestive IgG Western blot. Although he could not be certain there was CNS invasion given the absence of intrathecal antibody production, the mildly elevated protein and the positive serologic results led him to recommend oral doxycycline treatment followed by intravenous (IV) ceftriaxone if there was an insufficient response to the oral therapy. While oral antibiotic was not helpful, the 4-week course of IV antibiotic therapy led to an improvement in energy, clearing of the brain fog, and reduced hypervigilance. Over the course of the next 6 months, there was a gradual improvement in all of his symptoms, although symptom exacerbations occurred at a waning intensity over this time; eventually his cognition returned to normal, his psychiatric symptoms and hyperarousal resolved, and the paresthesias abated. A partial relapse of symptoms occurred about 10 months after initial symptoms at which time his physician and psychiatrist were unsure whether this represented a new nontick-related independent illness, a new *B. burgdorferi* infection from a new tick bite, or a recrudescence of the prior infection due to a small number of persistent organisms as has been demonstrated in animal models. Together the two clinicians opted to act empirically as controlled trials to address this puzzling scenario have not yet been conducted. His neuropsychiatric and physical symptoms were treated without delay with a combination of pharmacotherapy and 6 weeks of oral antibiotics with excellent response. The patient remained well and was able to taper antibiotics off after several months without further clinical relapse.

The above-mentioned case highlights several points: (1) neuropsychiatric symptoms may be an early manifestation of Lyme disease; (2) psychiatric symptoms are often accompanied by cognitive deficits as well as paresthesias, fatigue, and sensory hyperarousal; (3) the initial whole cell sonicate ELISA screening blood test may fail to detect the *Borrelia* antibodies, but the ELISA based on recombinant proteins (C6 or VlsE) and the Western blot may be informative when the whole cell sonicate ELISA is not; and (4) medical and mental health practitioners should be alert to the possibility of relapse due to a small amount of persistent infection and the potential benefit of combined psychiatric and antimicrobial treatment.

Tests for CNS Lyme Disease. Examination of the CSF is critical to rule out other possible causes of CNS disease and to identify the presence of Lyme meningitis or encephalitis. In the latter conditions, a spinal tap may reveal lymphocytic pleocytosis, mildly increased protein, and an elevated *Borrelia*-specific IgG index. PCR studies of the CSF are insensitive. Clinicians should order a Lyme ELISA on both the serum and CSF drawn at the same time so that a Lyme index can be calculated; in rare cases, Lyme-specific antibodies at diagnostically high levels may be found in the CSF but not in the serum. In about 20 percent of cases of later stage neurologic Lyme disease, however, the CSF may appear normal. MRI studies may reveal punctate white matter lesions on T2-weighted images, suggestive of a demyelinating disorder such as multiple sclerosis (MS). Electroencephalogram (EEG) studies are generally normal, although diffuse slowing or epileptiform discharges may be seen. SPECT and PET studies may have a role in late stage CNS Lyme disease, revealing a pattern of diffuse heterogeneous hypoperfusion; this pattern, however, is not specific to Lyme disease and controlled studies assessing the use of brain SPECT clinically for differential diagnosis have not yet been reported (Fig. 2.9–1). Given the difficulties facing the clinician attempting to determine whether fatigue, mood lability, and cognitive tracking problems are due to primary depression or to an underlying systemic disease, functional imaging studies provide the promise of a valuable tool to assist in the differential diagnosis.

Differential Diagnosis. In considering the diagnosis of Lyme disease, it is critical to ask about exposure to a Lyme-endemic area, history of a tick bite or unusual rash, and the presence of multisystemic involvement. A tick needs to be attached for at least 36 hours before *B. burgdorferi* transmission can occur, except in the rare circumstance when the tick had a prior partial blood meal. Called the "New Great Imitator" (after the original Great Imitator, syphilis), the broad spectrum of atypical neurologic manifestations of Lyme disease includes strokes, Guillain–Barré syndrome, cerebellar syndromes, seizures, pseudotumor-like syndrome in children, spastic paraparesis, MS-like illnesses, and progressive dementias. Similarly, other diseases that may look like neuropsychiatric Lyme disease need to be excluded, such as major depression with somatic preoccupation, panic disorder, systemic lupus erythematosus or other connective tissue diseases, CFS, endocrinologic disorders, vitamin deficiencies, other infectious illnesses, multi-infarct dementias, and other neurodegenerative disorders.

Treatment. For early Lyme disease without CNS involvement, a course of 2 to 3 weeks of oral doxycycline (Doryx; 100 mg twice a day), amoxicillin (Amoxil; 500 mg three times a day), or cefuroxime (Ceftin; 500 mg twice a

day) is recommended. For Lyme disease with CNS involvement, an initial course of 3 to 4 weeks of IV ceftriaxone (Rocephin; 2 gm per day) or cefotaxime (Claforan; 2 gm every 8 hours) is recommended. Symptoms may worsen during the first week of antibiotic treatment, much like the Jarisch–Herxheimer reaction during the treatment of syphilis. For patients who relapse, a repeated course of antibiotics may be helpful. Failure to treat Lyme disease early in its course or for a sufficiently long duration may lead to a chronic illness characterized by persistent waxing and waning neuropsychiatric disturbances, arthralgias, myalgias, sensory hyperacuities, and/or severe fatigue. In some patients, these symptoms reflect the effects of persistent infection, while in others the symptoms may reflect a residual postinfectious syndrome. Because the serologic tests for Lyme disease only reveal evidence of past exposure and do not document the presence of persistent infection, decisions regarding treatment are often based on the physician's clinical judgment. Four placebo-controlled trials of post-treatment Lyme syndrome provide conflicting results on the efficacy of repeated courses of IV antibiotic therapy for patients with persistent symptoms after having previously received the standard recommended treatment. Two showed no improvement in functional ability in either seropositive or seronegative patients. A third showed significant improvement in the primary outcome measure of fatigue, and a fourth study showed a moderate improvement in cognition overall that was not sustained after antibiotics were discontinued. The latter two studies reported cases of serious adverse events associated with the IV antibiotic therapy, highlighting that risk–benefit discussions with the patient are essential in the treatment selection. Future research needs to identify biomarkers that will guide clinicians in identifying appropriate treatments. New research on antibiotics preferentially effective against persister forms of *B. burgdorferi* may provide additional strategies for clinical care. In addition, controlled trials are needed of nonantimicrobial approaches to determine which strategies are most effective for patients previously treated with antibiotics who have persistent multisystemic symptoms.

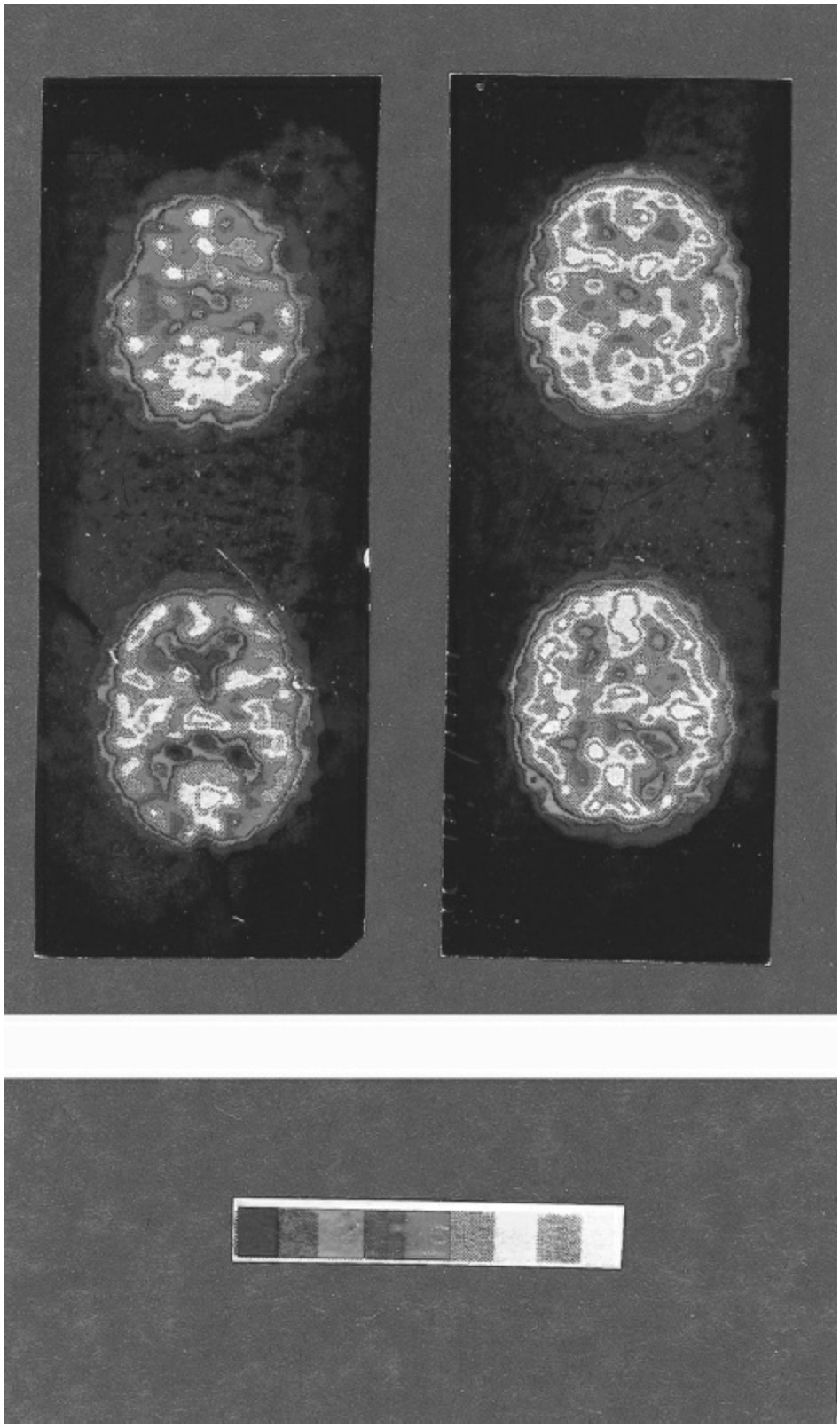

FIGURE 2.9–1. Single-photon emission computed tomography (SPECT) scan demonstrating multiple areas of decreased blood flow in a Lyme patient (L) compared to a healthy control (R). (See color atlas.)

Neurosyphilis

The cause of syphilis, *T. pallidum,* was identified in 1905. Because of the

cognitive loss and neuropsychiatric disturbances associated with tertiary neurosyphilis such as delirium, dementia, mania, psychosis, personality change, and/or depression, these patients accounted for 5 to 15 percent of psychiatric hospital admissions, labeled as "general paresis," "general paralysis of the insane," or "dementia paralytica." With penicillin treatment of primary and secondary syphilis, neurosyphilis is now an uncommon cause of hospital admissions.

Primary syphilis is manifest by a syphilitic ulcer, the chancre, at the site of inoculation. Secondary syphilis, a result of hematogenous dissemination of the spirochete, is characterized by flu-like symptoms followed by a skin rash, generalized lymphadenopathy, and mucosal lesions. Left untreated both primary and secondary syphilis resolve on their own, after which the patient enters a latent period wherein infection is present but clinical symptoms are not manifest. After months to years, about one-third of patients with untreated latent syphilis develop tertiary syphilis affecting the brain or heart. As in neuroborreliosis, invasion of the CNS by *T. pallidum* occurs early in the disease and may be asymptomatic for months to years prior to clinical expression.

Clinical neurosyphilis can be divided into four types: syphilitic meningitis, meningovascular syphilis, parenchymatous neurosyphilis, and gummatous neurosyphilis. Syphilitic meningitis, the result of direct meningeal inflammation, rarely has focal findings. Meningovascular syphilis results from the ischemic changes caused by proliferative endarteritis, causing permanent CNS damage, and presents most commonly as a stroke syndrome. At this stage there may be mild encephalitic symptoms, including personality change, emotional lability, insomnia, and decreased memory. In parenchymatous neurosyphilis (general paresis or tabes dorsalis), which generally starts 10 to 20 years after infection, there is direct neural destruction resulting in diminished neuron concentration, demyelination, and gliosis. In gummatous neurosyphilis, the mass effect causes neurologic symptoms.

General paresis, peaking in incidence 20 to 30 years after infection, represents a progressive frontotemporal meningoencephalitis with loss of cortical function. It often starts with subtle cognitive and emotional changes, such as problems with motivation, memory problems, irritability, and poor concentration, and if untreated can lead to confabulation, anomia, apraxia, or pseudobulbar palsy. The disease may mimic any psychiatric disorder as well. Half of the patients with neurosyphilis will manifest dementia, of whom one-quarter will have prominent psychiatric manifestations, such as depression, paranoia, psychosis, or mania. A worsening of symptoms during the first 24 hours after the initiation of antibiotic treatment has been termed the Jarisch–Herxheimer reaction; in rare cases, psychosis may emerge shortly after antibiotics are started. With disease progression, there is loss of muscle tone, fine motor control,

seizures, spasticity, and eventually paralysis and death. Focal neurologic findings are rare, consistent with the generalized pathophysiology. Tabes dorsalis, however, is characterized by progressive degeneration of the posterior columns and posterior roots of the spinal cord, resulting in a characteristic clinical picture of lancinating pains, sharp abdominal pains, and paresthesias. Because of progressive loss of proprioception and sensation, patients compensate by a broad-based shuffling gait. Unlike general paresis, not all patients with tabes will have CSF abnormalities. The late forms of neurosyphilis, general paresis, and tabes dorsalis are fairly uncommon due to the widespread use of antibiotics.

Tests. Serologic tests for syphilis include the nontreponemal Venereal Disease Research Laboratory (VDRL) test or RPR test and, for confirmatory purposes, the fluorescent treponemal antibody absorption (FTA-ABS) test. The FTA may be false positive in patients with *B. burgdorferi* infection, so another treponemal-based test (microhemagglutination assay-*T. pallidum*) should be used in that circumstance. Real-time PCR techniques now exist that allow for the detection of *T. pallidum* with estimates of sensitivity of 80 percent and specificity of 98 percent; real-time PCR appears also capable of discriminating between wild-type and drug-resistant strains.

CSF studies are essential for diagnosis and can also serve to detect asymptomatic involvement so that treatment can be started and efficacy followed. The diagnosis of neurosyphilis is based on a lymphocytic pleocytosis with a WBC count of greater than or equal to 20 and/or a reactive CSF VDRL and/or a positive CSF intrathecal *T. pallidum* antibody index. Higher cell counts and higher protein elevations are found in early compared to late neurosyphilis. Unfortunately, *T. pallidum* is difficult to demonstrate in the CSF and difficult to culture. CSF studies are limited by the low specificity of the elevated protein, γ-globulin, and leukocyte count and the low sensitivity (30 to 70 percent) (but high specificity) of the VDRL. The CSF FTA-ABS, however, is thought to have excellent sensitivity but less specificity than the CSF VDRL. A positive CSF VDRL or CSF RPR result from a patient with appropriate clinical history establishes the diagnosis of neurosyphilis.

Neuroimaging studies in patients with general paresis reveal frontocortical atrophy and disseminated high signal lesions in a frontal distribution; T2 white matter hyperintensities may reverse after antibiotic therapy. SPECT imaging in general paresis reveals a marked reduction in cerebral perfusion, particularly in the bilateral frontal and temporal cortices.

Treatment. The goal in clinical neurosyphilis is to reverse the manifestations or arrest the disease progression, although in some patients

antibiotic therapy may not be able to achieve these goals. Standard courses of antibiotic for 10 to 14 days consist of IV aqueous penicillin G (Pfizerpen), 12 to 24 million units daily in divided doses at 4-hour intervals, or alternatively intramuscular weekly injections of 2.4 to 4.8 million units of benzathine penicillin (Bicillin L-A), or intramuscular (IM) injections of 2.4 million units of procaine penicillin (Wycillin) four times daily. The likelihood of marked improvement for patients with general paresis is less than that for patients with syphilitic meningitis or meningovascular syphilis, reflecting the pathological process, which in the former is irreversible neuron damage and in the latter CNS inflammation. During the first year after treatment, the serum and CSF should be regularly monitored for the re-emergence of reactivity so that treatment can be reinitiated if necessary. Certain conditions, such as comorbid HIV infection, may place patients at greater risk for persistence of treponemal infection after antibiotic treatment. Most neurosyphilis patients with treatment will, however, show improvement in the cognitive, psychiatric, and functional domains.

Tick-Borne Coinfections. In about 30 percent of *Ixodes scapularis* ticks, more than one disease causing micro-organism can be found. Of particular concern are the human pathogens *Borrelia miyamotoi, Babesia microti, Ehrlichia chaffeensis* (causing Human Monocytic Ehrlichiosis [HME]), and *Anaplasma phagocytophilum* (causing Human Granulocytic Anaplasmosis [HGA]).

Borrelia miyamotoi, a relapsing fever *Borrelia* group spirochete, is a newly recognized tick-borne infection causing a flu-like illness, fevers, elevated liver function tests, leukopenia, and rarely meningoencephalitis; patients who have a Lyme-like illness after tick bite that tests negative for Lyme disease may instead have been infected by *B. miyamotoi* (which has been reported in 1 to 15 percent of *I. scapularis* ticks in Lyme-endemic regions). Because of the nonspecific nature of these symptoms, diagnosis requires confirmation by blood smear, PCR, antibody assay, and/or culture. While doxycycline is effective for uncomplicated *B. miyamotoi* infection, ceftriaxone or penicillin G may be more appropriate for the severe cases, such as meningoencephalitis.

Babesia are malaria-like protozoans that parasitize and reproduce within mammalian red blood cells, producing a clinical picture characterized by intermittent fevers, night sweats, and marked fatigue; because of the sudden dense fatigue after tick bite, patients with babesiosis are sometimes misdiagnosed as seronegative Lyme disease. In the absence of a known tick bite, they may be misdiagnosed as having CFS. The primary neuropsychiatric feature of babesiosis is profound fatigue, which may be accompanied by relapsing fever, chills, sweats, headaches, depression, myalgia, arthralgia, nausea, or vomiting. Splenomegaly or hepatomegaly

may occur. Hemolytic anemia, thrombocytopenia, and elevated liver enzymes are common. Because this may be a coinfection among patients who also have Lyme disease, persistent fatigue with sweats after receiving treatment of Lyme disease should prompt tests for babesiosis such as a blood smear for parasites, PCR for DNA, and serologic tests of antibabesial antibody. *Babesia* can be a life-threatening infection among immunocompromised or elderly patients.

Both *Ehrlichia* and *Anaplasma* are small, gram-negative bacteria that that cause a similar illness characterized by high fever, severe headaches, and myalgias as well as the laboratory markers of leukopenia, thrombocytopenia, and elevated liver enzymes. *Anaplasma* are transmitted by the *I. scapularis* tick, while *Ehrlichia* are transmitted by the Lone Star tick. *Ehrlichia* primarily invade monocytes and macrophages, while *Anaplasma* invade neutrophils. HME is a more severe disease than HGA, with a 3 percent case fatality rate and 42 percent requiring hospitalization. CNS involvement occurs in up to 20 percent of cases of HME, primarily meningitis and meningoencephalitis; less commonly, peripheral and cranial neuropathies, seizures, or coma may occur. Among patients with HME, CSF studies reveal a pleocytosis in 60 percent of cases and mildly elevated protein. Immunocompromised patients or the elderly are at greatest risk.

For all of these tick-borne coinfections, both antibody and PCR assays are available; as *B. miyamotoi* is a newly described human pathogen, commercially available tests for this agent are still in development. Doxycycline is an effective treatment of many tick-borne infections, including *Borrelia, Anaplasma,* and *Ehrlichia,* but not for *Babesia*. A key point for the psychiatrist faced with a patient with persistent fatigue after treatment of Lyme disease is that there may be a previously unrecognized *Babesia* coinfection that requires different antimicrobials, commonly atovaquone and azithromycin.

BARTONELLA. *Bartonella* species are Gram-negative alphaproteobacteria with at least 15 different species causing a *Bartonellosis* in humans. Transmitted by animal bites, scratches, needle sticks, and arthropods (fleas and ticks), *Bartonella* are highly adapted to mammalian reservoir hosts (e.g., *Bartonella henselae* in cats, *Bartonella vinsonii* in wild dogs) and have been shown in vitro to invade erythrocytes, endothelial cells, and macrophages. Cat-scratch disease is the most commonly recognized manifestation of human infection with *B. henselae,* typically producing an acute, self-limiting illness characterized typically by a cutaneous papule, fever, and lymphadenopathy. Neurologic manifestations of CSD may include encephalopathy and neuroretinitis and, less often, cranial neuropathies, expressive aphasia, polyneuropathy, and transverse myelitis. In one U.S. study, *Bartonella* was the most common bacterial cause of

encephalitis (2 percent of 334 patients). The pathogenesis of these CNS manifestations is unclear, as they may occur due to direct invasion of the CNS or to other mechanisms such as local vasculitis. Nearly half of patients with CSD encephalopathy develop seizures that can range from focal and self-limited to status epilepticus. Unilateral loss of visual acuity is a typical ocular manifestation of neuroretinitis. Although antibiotic therapy appears effective for *Bartonella* infection, spontaneous resolution without treatment can often occur within 2 to 4 months.

Recent case reports and small series, particularly from one research group that uses a novel enrichment blood culture technique followed by PCR and DNA sequencing, suggest that various species of *Bartonella* (*henselae, vinsonii, koehlerae*) may lead to a chronic intravascular infection in immunocompetent hosts lasting months to years and possibly causing a range of neuropsychiatric symptoms, including memory loss, seizures, hallucinations, insomnia, ataxia, and fatigue. In these cases, the CSF may appear unremarkable with routine studies, but demonstrate *Bartonella* using PCR amplification or enriched culture medium. Support for the hypothesis of persistent infection in immunocompetent patients came from a study from this research group in which 41 percent of 296 patients with rheumatologic symptoms were found to have a *Bartonella* species bacteremia. Whether the persistent bacteremia in most cases is related to clinical symptoms is unclear.

Tests. Antibody testing for *Bartonella* exposure is valuable but insufficiently sensitive; in one study, positive titers were found in only 30 percent of patients in whom *Bartonella* infection was confirmed by PCR and DNA sequencing. Fine-needle aspiration or biopsy of an involved lymph node can be sent for histopathology, PCR, and culture when serology is negative or equivocal. PCR assays and culture using enriched medium on the blood and CSF can also be conducted; culture using standard media, however, is insensitive. *Bartonella* infection may not always be pathogenic, as has been demonstrated in animals. Given the typically self-limiting nature of this infection, a high percentage of healthy individuals may test positive serologically but not recall having had a prior *Bartonellosis* set of symptoms.

Treatment. Antimicrobial therapy has not been well-studied, but clinical reports suggest that doxycycline with rifampin can be effective in shortening the course of the disease.

NON-HIV VIRAL INFECTIONS OF THE CNS

Numerous viruses are invasive and neurotropic, with the extent of consequent neuronal dysfunction varying widely depending on both the virulence of the virus and the immunologic response of the host. This section will focus on agents known to cause striking neuropsychiatric

diseases: herpes simplex, rabies, measles, and subacute sclerosing panencephalitis (SSPE) (Table 2.9–1 lists other viruses).

Herpes Viruses

Included under the spectrum of herpes simplex viruses are HSV-1 and HSV-2, varicella zoster virus (VZV), Epstein–Barr virus (EBV), CMV, human herpes virus (HHV)-6, HHV-7, and Kaposi sarcoma herpes virus. Herpes viruses consist of double-stranded DNA surrounded by a protein capsid that, after infection, integrates into the host cell DNA and establishes latency in the nervous system with periodic lytic cycles that could last a lifetime. CNS infection by CMV, EBV, or VZV is rare in an immunocompetent host, while HSV-1 and HHV-6 can cause CNS infections in immunocompetent human hosts. HSV-2 CNS infection tends to occur in neonates.

Table 2.9–1.
Selected Infectious Causes of Neuropsychiatric Disorders

Bacterial Infections
Acute: *Haemophilus influenzae,* Meningococcus, Pneumococcus
Subacute: Brucellosis, leptospirosis, Lyme disease, syphilis, tuberculosis, Whipple disease
Fungal Infections
Coccidioidomycosis, cryptococcosis, histoplasmosis, *Candida*
Parasitic Infections
Cysticercosis, malaria, toxoplasmosis
Prions
Creutzfeldt–Jakob disease, fatal familial insomnia, kuru
Viral Infections
Arbovirus, coxsackievirus, cytomegalovirus, enterovirus, Epstein–Barr virus, flavivirus, herpes simplex virus, human immunodeficiency virus, influenza virus, lymphocytic choriomeningitis virus, measles virus, mumps, papovavirus, poliovirus, rabies virus, rubella, togavirus

Herpes Simplex

HSV encephalitis is a dramatic disorder, characterized by the abrupt onset of fever, personality change, alteration of consciousness, and headaches, followed by cognitive changes and focal neurologic signs, such as dysphasia, autonomic dysfunction, visual field deficits, hemiparesis, or seizures. Although focality is an important feature of HSV encephalitis, other viruses may also have focal signs, such as the La Crosse virus or the nonpolio enteroviruses. Neurobehavioral aspects of HSV encephalitis such as hallucinations, memory loss, or behavioral disturbances may be the presenting feature, a consequence of the predilection of the virus for the temporal lobes. Although the course of illness is typically rapidly progressive, resulting in refractory seizures, coma, and death within 2 weeks, occasionally the progression may be slower with varied

neuropsychiatric features.

HSV-1 is usually transmitted orally, entering the CNS through sensory nerves, residing in a latent state in the trigeminal ganglia most commonly but possibly also in other tissues such as the cornea and brain. HSV-2 is usually transmitted genitally and may seed the sacral ganglia or disseminate hematogenously. HSV typically produces a lytic infection with neuronal necrosis and tissue destruction and intranuclear inclusion bodies in the neurons and glia. On MRI, the prefrontal gray matter volume may be decreased among HSV-1 seropositive first-episode antipsychotic-naive schizophrenia/schizoaffective disorder patients compared to HSV-1 seronegative first-episode psychosis patients and to HSV-1 seropositive healthy controls; this finding raises the possibility that HSV-1 infection in adults with schizophrenia may be associated with regional gray matter differences. Patients who survive HSV encephalitis may exhibit postencephalitic symptoms, such as cognitive impairment, amnesia, aphasia, and less commonly, the Klüver–Bucy syndrome or dementia. Maternal humoral immunity to HSV-2 during gestation has been linked to a propensity later in life to the development of schizophrenia-spectrum disorders; this was not found to be true for HSV-1 infection.

Routine serologic studies are of little value in suspected HSV encephalitis. The CSF usually demonstrates leukocytosis (approximately 100 cells/mm^3), a moderate protein elevation, and a normal or depressed glucose content. PCR analysis of the CSF to detect HSV DNA is at present the diagnostic procedure of choice, as the PCR assay has high sensitivity and specificity. Approximately 80 percent of patients with biopsy-proven HSV encephalitis will have focal EEG abnormalities consisting of slowing or repetitive epileptiform discharges in the frontotemporal area. MRI studies in early stages of HSV encephalitis may reveal T2 prolongation in the insular cortex and cingulate gyrus. SPECT or PET imaging may show reduced blood flow in the orbitofrontal and temporal regions. Brain biopsy in difficult to diagnose cases can be helpful, although the complication rate is approximately 3 percent.

If untreated, 40 to 70 percent of patients with HSV encephalitis will die. Antiviral therapies include acyclovir (Zovirax) and vidarabine (Vira-A); however, even with acyclovir treatment that is considered more effective, fewer than 40 percent of patients survive with minimal or no sequelae.

Epstein–Barr Virus

Most adults have evidence of past exposure to EBV, with approximately 50 percent seropositivity among children over age 5. Infection in childhood is generally mild, whereas in adolescence and young adulthood it may result in infectious mononucleosis or, rarely, a fulminant life-threatening disease. EBV enters the body by infecting oral mucosal epithelial cells. The clinical symptoms of infectious mononucleosis of sore throat, headache, malaise,

and fatigue are largely a result of the vigorous cellular immune response to EBV infection (T cell and natural killer cells) rather than direct cytotoxic effects. Significant neurologic complications of EBV infection are rare, occurring in less than 0.5 percent of cases of infectious mononucleosis.

EBV encephalitis occurs usually within 1 to 3 weeks after the onset of clinical infectious mononucleosis. Patients with EBV encephalitis may present with cerebellar ataxia, personality changes, psychosis, transient global amnesia, perceptual distortions of size, and space (metamorphopsia or "Alice in Wonderland" syndrome), focal neurologic findings, seizures, or coma. EEG usually reveals generalized slowing with occasional sharp wave activity. The diagnosis of an EBV neuropsychiatric syndrome requires an appropriate clinical history in the setting of serologic evidence of acute, or rarely chronic, active infection. In cases of EBV encephalitis, commonly there is a lymphocytic pleocytosis (atypical lymphocytes are particularly suggestive) with elevated protein. In most cases, EBV encephalitis is self-limited, with recovery occurring within weeks to months. Rarely, acute EBV infection may result in relapsing or chronic encephalitis. Treatment is generally supportive.

HHV-6

HHV-6 is neuroinvasive and has been shown to be widely disseminated in the human brain. Encephalitis attributed to HHV-6 infection has been reported most often in immunocompromised individuals but may also occur in immunocompetent hosts. Typically the CNS syndrome is of a limbic encephalitis characterized by altered mental status, confusion, anterograde amnesia, fever, hallucinations, insomnia, and emotional lability. HHV-6 has been investigated as an autoimmune trigger in MS, with inconsistent results, but several studies reveal higher prevalence of DNA for HHV-6 by PCR in patients compared to controls, particularly at a time of clinical relapse, as well as a HHV-6 mRNA detection in the serum of patients with relapsing remitting MS over time. Seizures are a primary complication of HHV-6 infection. The laboratory diagnosis of HHV-6 infection requires isolation of the virus in culture along with seroconversion. HHV-6 antibodies are common in the adult population, indicative of prior exposure. A positive PCR for HHV-6 DNA in a peripheral blood mononuclear cell (PBMC) sample does not reveal whether there is active viral replication (primary infection or reactivation) or latent infection. A newer test—reverse transcriptase PCR (RT-PCR)—is more informative as it amplifies the mRNA of the late structural proteins produced during active infection. Early suggestions that HHV-6 infection was associated with CFS, Sjögren syndrome, and certain malignancies have not been supported by more recent investigations.

Other Herpes Viruses

With herpes zoster, neuropsychiatric complications occur most frequently in immunocompromised patients, resulting in encephalitis, myelitis, or leukoencephalitis. With CMV infection, encephalitis may also occur as CMV is tropic for the CNS; however, only in rare exceptions has CMV encephalitis occurred in non-HIV infected immunocompromised individuals. CMV infection is the leading viral cause of congenitally acquired mental retardation.

Rabies

Although most cases of human rabies occur via saliva after animal bites, other modes of transmission of rabies virus infection include aerosols (risk of spelunkers) and person-to-person transmission following corneal transplants. The rabies virus is a negatively stranded RNA virus that replicates locally at the site of inoculation and subsequently spreads to the CNS by retrograde axonal transport, infecting the lower areas of the brain most prominently, particularly the limbic system, hippocampus, brain stem, and cerebellum. Limbic system involvement may result in aberrant sexual behavior and behavioral dyscontrol, whereas brain stem involvement typically results in alterations of body temperature and respiratory control. The site and amount of inoculation are associated with morbidity. For example, multiple dog bites to the face may result in a 60 percent mortality rate without prophylactic intervention, whereas multiple bites to the hand are associated with lower mortality rates of about 15 percent. The incubation period prior to symptomatic expression ranges from a few days to several years, although 20 to 90 days is the most common range. Once symptoms emerge, the course is rapidly fatal. There are two main forms: encephalitic rabies (80 percent of cases) and paralytic rabies (20 percent of cases). Most patients with encephalitis get the encephalitic “furious” form characterized by agitation, hallucinations, odd behaviors, extreme excitability, and in some cases, hydrophobia. States of generalized arousal and aggression may be offset by lucid periods. Diagnosis is based on the history of an animal bite in a patient with unexplained encephalitis that has been confirmed by the demonstration of rabies antigen on a skin biopsy of the patient or from a putatively infected animal or by rabies virus RNA in fluids or tissues using RT-PCR amplification. CSF studies typically show a mononuclear pleocytosis (<100 cells). There is no treatment of rabies virus infection. Disease prevention is critical, aided by pre-exposure vaccination in high-risk individuals and postexposure prophylaxis with rabies immunoglobulin and rabies vaccine.

Rubella

The rubella virus, a member of the *Togaviridae* family, causes an acute exanthematous viral infection, characterized by rash, fever, and

lymphadenopathy. Because postnatal rubella exposure causes only a mild illness, the main concern is with prenatal exposure, which can cause fetal death and severe congenital defects. Prenatal exposure to rubella virus has also been associated with a much higher risk of the emergence of other diseases in childhood and young adulthood, such as diabetes mellitus, progressive encephalopathy resembling SSPE, and, as suggested by recent studies, schizophrenia. Since the development of the live attenuated rubella vaccine in 1969, there have been no large rubella epidemics in countries where the vaccine is widely used.

Subacute Sclerosing Panencephalitis

SSPE is a very rare slow infection with measles virus that causes progressive inflammation and sclerosis of the brain. Primarily manifesting in children and young adults who were infected prior to age 2, the rate of SSPE decreased markedly after 1960 as a result of widespread measles vaccination, with a current rate in the United States of only one per 100 million people per year. The onset generally occurs 7 to 12 years after the measles infection and is subtle, characterized by gradual changes in behavior, personality, and school performance. Neuropsychological testing may demonstrate reduced overall intelligence and problems with reading, writing, and visuospatial processing. Neuropsychiatric symptoms may include hallucinations, apraxia, agnosia, and Balint syndrome (optic ataxia, simultanagnosia, and sticky fixation). Repetitive myoclonic jerks are common, at times accompanied by movement disorders and cerebellar ataxia. In advanced stages, dementia, mutism, cortical blindness, optic atrophy, stupor, coma, and death occur. The usual course of illness is 1 to 3 years, with rare patients surviving up to 10 years.

Serologic testing may reveal unusually high titers of antibodies to measles virus. CSF studies typically show high measles antibody titers and a greatly elevated γ-globulin fraction with oligoclonal bands in a CSF with slightly elevated protein levels. EEG studies are essential, particularly in the myoclonic stage, revealing high-amplitude bilateral and stereotyped complexes that repeat every 3 to 5 seconds. MRI studies may reveal enlarged ventricles and diffuse brain atrophy, with multifocal low-density white matter lesions and lucent areas in the basal ganglia. PET and SPECT studies may reveal early subcortical hypermetabolism followed by global cortical and subcortical hypometabolism.

No treatments are known to reverse the disease, although slightly prolonged survival has been associated with inosiplex (Isoprinosine) and with intraventricular injections of γ-interferon. Measles vaccination is the best means of prevention.

West Nile Virus

West Nile virus (WNV) is an arthropod-borne flavivirus that is usually spread by infected mosquitoes. Most infected persons are either asymptomatic or experience mild symptoms such as headache, rash, and low-grade fever. Only approximately 1 percent of patients develop neuroinvasive disease manifesting with more severe symptoms such as meningitis, encephalitis, or acute anterior myelitis. However, certain outbreaks have been associated with much higher rates of neurologic disease. In the 2010/2011 WNV outbreak in Greece, over 75 percent of patients had neuroinvasive disease (encephalitis, meningitis, or acute flaccid paralysis). In the 1998/1999 outbreaks in New York, two-thirds of hospitalized patients had encephalitis and one-third had meningitis. Studies of neuropsychiatric aspects of WNV are limited and uncontrolled, with reports of new onset depression in one-third of infected patients, as well as problems with fatigue, irritability, and cognition, particularly attention, executive functions, and motor skills. CSF exam of patients with encephalitis may show a moderate lymphocytic pleocytosis and protein elevation. IgM antibodies against WNV in the CSF (detected by IgM capture ELISA) are indicative of neurologic WNV infection. Real-time PCR detects only 55 percent of patients with WNV encephalitis.

Progressive Multifocal Leukoencephalopathy

This disease, affecting immunocompromised subjects, is a progressive infection of white matter (oligodendroglial cells) with the JC papovavirus. Typically the onset is abrupt with focal neurological or neuropsychological signs and the course is almost invariably fatal within 2 to 4 months. Definitive diagnosis requires a brain biopsy. Neuroimaging studies reveal multifocal areas of high signal intensity in the white matter. Functional imaging with PET or SPECT may reveal a heterogeneous pattern of reduced metabolic activity and perfusion.

TRANSMISSIBLE SPONGIFORM ENCEPHALOPATHIES

Included in this group of transmissible neurodegenerative diseases are Creutzfeldt–Jakob disease (CJD); kuru, a dementing disease of three New Guinea tribes most likely spread by ritual cannibalism; Gerstmann–Straussler syndrome, a familial disorder characterized by dementia and ataxia; fatal familial insomnia, a disorder causing disturbances of sleep and of motor, autonomic, and endocrine function; and, in cattle, bovine spongiform encephalopathy (BSE or mad cow disease). These are all fatal neurodegenerative disorders caused by prions. A prion is a small infectious pathogen containing protein that is resistant to procedures that modify or hydrolyze nucleic acids. Human prion diseases share several features: (1) pathology is almost exclusively confined to the CNS; (2) the diseases typically have long incubation times; (3) the course is progressive and fatal;

(4) the neuropathologic hallmarks include a reactive astrocytosis with little inflammation and typically neuronal vacuolation, leading to spongy degeneration of the cerebral cortical gray matter; and (5) each of the diseases appears to result in accumulation of the prion protein (PrP). When there is a prion infection in a human, there is a post-translational conversion of a normal host encoded prion protein to an abnormal pathogenic form (PrP^{Sc}). Unlike other naturally occurring microbes, prions do not rely on nucleic acid to replicate within the host organism. Unlike other infectious diseases, prion diseases are unique in that they can occur in three forms: sporadic, infectious/acquired, and genetic. Approximately 85 percent of human prion disease is sporadic, 10 to 15 percent is genetic, and 1 percent is acquired.

Invariably fatal, this transmissible, rapidly progressive disorder occurs mainly in middle age or older and is manifest early on by fatigue, flu-like symptoms, mild cognitive impairment, or focal findings, such as aphasia or apraxia. Psychiatric manifestations may then emerge, including mood lability, anxiety, euphoria, depression, delusions, hallucinations, or marked personality changes. Progression of disease occurs over months, leading to dementia, akinetic mutism, coma, and death. Other common neurologic findings are generalized "startle" myoclonus, cortical blindness, and extrapyramidal and cerebellar signs.

Worldwide the rates of CJD range from 0.25 to 2.0 cases per million per year. The infectious agent self-replicates and can be transmitted to humans by inoculation with infected tissues and sometimes by ingestion in food. Iatrogenic transmission has been reported via transplantation of contaminated cornea or to children via contaminated supplies of human growth hormone. Household contacts are not at greater risk than the general population, unless there is direct inoculation.

Because of the epidemic of prion disease (BSE or mad cow disease), among cattle in the United Kingdom in 1986 and because of the unexpected emergence in 1995 of cases of a "new variant" form of CJD (vCJD) among teenagers in the United Kingdom, fears emerged that transmission to humans may have occurred as a result of eating infected beef. Strong evidence now supports a causal relationship between BSE and vCVD. Since 1995, 228 human cases of vCJD have been reported, the overwhelming majority (more than 76 percent) from the United Kingdom. Patients with vCJD compared to typical sporadic CJD are considerably younger at age of onset (29 years vs. 65 years), experience a longer duration of illness (14 months vs. 4.5 months), and more frequently present with sensory disturbances and psychiatric manifestations, including psychosis, depression, personality change, and anxiety. As disease progresses, patients with vCJD develop pyramidal signs, myoclonus, rigidity, cerebellar signs, and akinetic mutism. Neuropathologically, the main distinction between vCJD and sporadic CJD appears to be the prominent involvement

of the cerebellum in nearly all cases of vCJD, with prominent PrP^{Sc+} amyloid plaques distributed throughout the cerebrum and cerebellum.

Definitive diagnosis of CJD requires pathological examination of the cortex, which reveals the classic triad of spongiform vacuolation, loss of neurons, and glial cell proliferation. Genetic susceptibility is a factor in disease risk, indicated by a common polymorphism of the human prion protein. Deep white matter T1-MRI hyperintensities in the neocortex, basal ganglia, and/or thalamus are characteristic of the disease; in fact, MRI images of the brain show a 91 to 96 percent sensitivity and 92 to 94 percent specificity. EEG abnormalities, although not specific for CJD, are present in nearly all patients with sporadic CJD revealing a slow and irregular background rhythm with periodic sharp wave complexes. CT and MRI studies may reveal cortical atrophy later in the course of disease. SPECT and PET reveal heterogeneously decreased uptake throughout the cortex. Routine CSF analyses are typically normal. While the MRI investigations have better sensitivity and specificity, CSF studies can also be helpful as markers of rapid neuronal injury. The t-tau has 86 percent sensitivity and 88 percent specificity while the 14–3–3 protein in large patient cohorts has been shown to have 86 percent sensitivity and 74 percent specificity in sCJD. There is no known treatment of CJD.

PROTOZOA

Protozoa are known to chronically infect human brain tissue and to cause behavioral changes. Of particular interest are *Toxoplasma gondii* (toxoplasmosis), *Babesia microti* (babesiosis), *Plasmodium* (malaria), and *Trypanosoma* (sleeping sickness). *T. gondii* infection during pregnancy has a known deleterious impact on the developing fetal CNS. A 2007 meta-analysis of studies examining *T. gondii* antibodies among individuals with schizophrenia reported an odds ratio of 2.73, suggesting that *T. gondii* may play some role in the etiology of schizophrenia. This association between schizophrenia and *T. gondii* antibodies has not revealed a connection with cognitive performance or psychosis severity. Although it is known that *T. gondii* is neurotropic and affects neurotransmitters, it should also be noted that most people with *Toxoplasma* do not develop psychosis and the *Toxoplasma* organism has been difficult to detect in the brains of individuals with schizophrenia.

OTHER INFECTIOUS CAUSES OF NEUROPSYCHIATRIC DISORDERS

A variety of bacterial, mycoplasmal, fungal, and parasitic infections can cause neuropsychiatric disturbances as a result of chronic meningitis or sequelae from an acute infection (Table 2.9–1).

EMERGING AREAS OF INVESTIGATION

Chronic Fatigue Syndrome

CFS, more commonly referred to as myalgic encephalomyelitis in Britain and Canada, is a common multisystem syndrome characterized by 6 months or more of severe, debilitating fatigue that is not relieved by rest, often accompanied by myalgia, headaches, pharyngitis, arthralgias, low-grade fever, sleep disturbance, cognitive complaints, gastrointestinal (GI) symptoms, postexertion malaise, and tender lymph nodes. A 2015 report from the Institute of Medicine proposed a new name—"systemic exertion intolerance disease"—to reflect the core symptoms of sustained depletion of energy after minimal physical and cognitive activity. Many patients with CFS have been ill for at least 5 years and are as functionally impaired as patients with MS or heart disease. The search for an infectious cause or trigger of CFS has been active because of the high percentage of patients who report abrupt onset after a severe flu-like illness. Excessive immune activation to an infectious stimulus may result in an overproduction of pro-inflammatory cytokines (e.g., interleukin-6 [IL-6] or tumor necrosis factor-α [TNF-α]) as has been shown in several studies of the serum and CSF of patients with CFS, possibly inducing the typical symptoms of fatigue, cognitive dysfunction, sleep disturbance, and increased sensitivity to pain. Other evidence of cytokine-mediated immune activation reported by some but not all studies of patients with CFS include increased levels of autoantibodies, activated complement, activated T lymphocytes, and decreased natural killer cell activity. A proteomic study in 2012 compared the CSF of three groups: CFS, post-treatment Lyme, and healthy controls; the analysis demonstrated an elevation of complement cascade proteins in the CSF of both patient groups compared to controls. A 2015 case-control study demonstrated that the cytokine immune signature of short duration CFS (<3 years) differs from that of later CFS (>3 years), with a prominent activation of both pro- and anti-inflammatory cytokines in early disease and a dissociation of intercytokine regulatory networks in later disease. Together these studies confirm an immune dysregulation in patients with CFS.

In the mid-1980s, the etiology of CFS was linked to infection with EBV. After EBV was shown in controlled studies to have no specific role in the etiology of CFS, reports have linked CFS to a variety of other agents, including enteroviruses, retroviruses, and new lymphotropic herpes viruses. These reports have not been replicated in well-designed studies. Certain organisms, however, can result in a CFS-like picture, such as infection with *B. burgdorferi,* which causes Lyme disease, or infection with *B. microti,* which causes babesiosis; however, most cases of CFS are not linked to these agents. Some patients with CFS-like symptoms may suffer from neurally mediated hypotension (NMH), a dysfunction of the autonomic nervous system. Checking for NMH through a tilt-table test

among patients with CFS is important as recent research indicates that medications effective for NMH may lead to relief from CFS. Population-based studies of CFS, however, do not support a large role for NMH as a cause of CFS.

Risk factors for CFS include female gender, physical and emotional stressors, and certain personality traits such as high emotional reactivity. Among those individuals for whom stressful life events, family dysfunction, and early childhood trauma have been linked to CFS, the onset of their CFS was more often gradual than acute, suggesting that the type of onset may identify differing subgroups of patients with CFS. Some studies indicate that patients who are most likely to be plagued by persistent fatigue after an acute viral illness are patients with pre-existing or comorbid psychiatric problems. However, other research has shown that the cognitive impairment in CFS exists even in the absence of pre-existing or comorbid psychiatric disorders, thus leading to the conclusion that psychiatric disorders alone cannot account for CFS. Various studies have found high rates of depressive disorders among patients with CFS, ranging from 15 to 54 percent. At present, CFS is best conceptualized as a heterogenous syndrome of uncertain etiology, most likely involving an interplay of psychiatric, infectious, neuroendocrine, and immunologic factors.

Controlled clinical trials among patients with CFS do not support the use of antidepressants, corticosteroids, or evening primrose oil. At present, the most convincing clinical trial results have come from non-pharmacological therapies. The results from numerous well-designed studies now support the use of cognitive behavior therapy and graded aerobic exercise programs to help alleviate the symptomatology and reduce the disability associated with CFS. However, as the immune signature of CFS becomes better characterized, future immune-based therapies may be an additional powerful tool for patients who respond poorly to CBT.

Group A β-Hemolytic Streptococcus

Historically Sydenham chorea (the movement disorder typically associated with rheumatic fever) was one of the first disorders associated with Group A Streptococcus (GAS) infection described as having acute onset neuropsychiatric symptoms. Poststreptococcal autoimmunity has been postulated to be a cause of certain types of childhood onset OCD and Tourette syndrome based on the observation that children who develop Sydenham chorea are often observed to have tics or obsessive–compulsive symptoms prior to the onset of the chorea.

Despite insufficiently robust diagnostic clarity in published studies, the literature has consistently described an ongoing relationship between GAS infections, OCD, and tics; in 1998 a workgroup investigated 50 case reports and designated the acronym PANDAS (*p*ediatric *a*utoimmune *n*europsychiatric *d*isorders *a*ssociated with *s*treptococcal infections). This

subgroup of patients is characterized by five clinical criteria: (1) the presence of OCD and/or tics; (2) prepubertal symptom onset; (3) abrupt onset (24 to 48 hours) and episodic course; (4) presence of neurologic signs, such as choreiform movements; and (5) symptom onset or exacerbations occurring shortly after (typically 7 to 14 days) group A β-hemolytic streptococcal (GAS) infections. Affected children also are more likely to have attentional disorders and associated symptoms such as emotional lability, deteriorating visual spatial skills, math and reading deficits, and enuresis.

In 2012, researchers and clinicians familiar with PANDAS developed a new broader diagnostic category called pediatric acute onset neuropsychiatric symptoms (PANS) to include similar clinical presentations that are more etiologically diverse, involving infectious agents that are not exclusively due to Group A Streptococcal infections. Criteria include (1) abrupt and dramatic onset of OCD or severely restricted food intake; (2) concurrent presentation of 2 out of 7 other neuropsychiatric symptoms: anxiety, emotional lability and/or depression, irritability, aggression and/or severe oppositional behaviors, behavioral (developmental) regression, deterioration in school performance, sensory or motor abnormalities including heightened sensitivity to sensory stimuli, and somatic signs or symptoms such as sleep disturbance or enuresis; and (3) these symptoms should not be better described by an otherwise known neurological or medical condition such as Tourette disorder, Sydenham chorea, or systemic lupus erythematous. Unlike PANDAS that requires a temporal relationship with GAS infection, PANS does not require a specific infection as a precursor to symptom onset, as other infections (e.g., mycoplasma pneumonia; viral infections; *B. burgdorferi*) have also been linked to new onset OCD.

The pathophysiology of PANDAS is hypothesized to be based on the concept of molecular mimicry, whereby infection with β-hemolytic streptococci triggers antistreptococcal antibodies that cross-react with epitopes on the basal ganglia of susceptible hosts, resulting in acute inflammation. Some neuroimaging studies have revealed increased basal ganglia volumes, and a recent study demonstrated that children with OCD had a significantly higher rate of antibasal ganglia antibodies compared to several control groups, a finding in support of an autoimmune component to certain types of childhood onset OCD.

Given that PANDAS affects multiple cognitive, emotional, and neurological domains, a multimodal treatment approach is often needed, consisting of pharmacological treatment of OCD or tic symptoms, and therapies such as CBT and occupational therapy. Pharmacologically the antibiotics penicillin, cephalosporin, and amoxicillin are all considered appropriate treatments of GAS infections and have been correlated with OCD symptom reduction in PANDAS youth. Prophylactic antibiotics have

also been investigated as potential PANDAS treatment; one small study showed benefit with penicillin and azithromycin in reducing recurrent streptococcal infections and neuropsychiatric exacerbations, but further studies are needed. Immunomodulatory treatments can also be helpful; a controlled trial showed that IV immunoglobulin (IVIG) therapy resulted in a reduction in OCD symptoms, while plasmapheresis resulted in both improved OCD and fewer tics, and both showed that treatment gains were maintained at 1 year.

Borna Disease Virus

Borna disease virus (BDV) is a small neurotropic RNA virus that infects various domestic animal species, causing meningoencephalitis and disturbances in behavior and cognition and rarely fatal outcome. In animals, BDV targets cells of the limbic system, replicates at low levels, persists for the lifespan of the host, and compromises their neuronal function without causing direct damage. Serologic and molecular studies on human patients have been performed to determine whether BDV may also cause neuropsychiatric disease in humans, such as major depression, bipolar disorder, or schizophrenia. Infection is typically evaluated by serology or PCR analyses of PBMCs or tissues. Serologic studies have revealed differing results, with positive antibody titers to BDV reported in 1.6 to 100 percent of subjects with specific neuropsychiatric disorders versus 1.2 to 46 percent of healthy controls. Some of this variability may reflect differences in assay specificity. Other studies have reported elevated BDV-specific antigen-antibody complexes, but this methodology also has been reported to have poor specificity. Additional studies have reported the presence of BDV RNA or BDV antigens in the peripheral blood samples as well as in autopsied brains of psychiatric patients. These data support the possibility of human infection with BDV. Other research groups, however, have been unable to replicate these findings; one recent large U.S. study of 198 patients with psychiatric illness (schizophrenia, bipolar disorder, major depression) and 198 matched controls reported a complete absence of such BDV markers from their samples. The results of these conflicting BDV studies may be due to poor interlaboratory reliability for serologic studies and the possibility of laboratory error or cross-contamination in the BDV nucleic acid studies. At present, the most cautious conclusion would be that there is insufficient cumulative evidence to indicate that BDV has a role in the pathogenesis of human psychiatric disorders.

Influenza Virus

An accepted risk factor for schizophrenia is birth in winter or spring months. The preponderance of evidence suggests that the prevalence of influenza in winter months accounts for this association. Recent research

has demonstrated that second-trimester respiratory infection increases the risk of schizophrenia in the offspring three- to sevenfold. Influenza, also known as the flu, is caused by RNA virus in the *Orthomyxoviridae* family (other RNA viruses include hepatitis C and severe acute respiratory syndrome [SARS]). Common symptoms include fever, pharyngitis, headache, cough, fatigue, and malaise. The influenza virus is transmitted from infected mammals through the air by coughs or sneezes. Influenza is common, with pandemics occurring every 10 to 20 years. Vaccinations against influenza can be helpful as can antiviral treatment if the case becomes severe. Most current animal research suggests that the risk to the developing brain is not from direct invasion by the influenza virus but through maternal immune activation and elevated cytokines (e.g., IL-6). IL-6 has been shown to have a direct effect on developing neurons and glia, including changes in proliferation, death, and gene expression.

Retroviruses

Endogenous retroviruses, well known to cause a variety of diseases including neoplasia, autoimmunity, and encephalitis, have also been reported to be expressed to a significantly greater extent in the brains of individuals affected with schizophrenia and other neuropsychiatric disorders compared to the brains of unaffected individuals. Evidence consists of the identification of viral sequences in affected brains and the increased activity of virally encoded reverse transcriptase. Because retroviruses are capable of cellular infection and integration into the host genome, the activation of these viral sequences in cells within the CNS can then lead to the transcription of adjacent genes and alterations in neural functioning. Although viral triggers or causes for neuropsychiatric disorders are compelling in their ability to help explain seasonal birth effects, the impact of perinatal complications, and discordance among monozygotic twins, much more investigation in this area is needed before conclusions can be drawn.

Antimicrobial Effects of Psychiatric Medications

That antimicrobial medications may have therapeutic effects for primary psychiatric disorders was first described in the 1950s when astute clinicians observed that when depressed tubercular patients were treated with the antibiotic iproniazid (Marsilid), a monoamine oxidase inhibitor, the depression often improved; based on this observation, a class of effective antidepressants were identified. In recent years, minocycline, a second-generation tetracycline that inhibits microglial activation and has an ability to modulate glutamate-induced excitotoxicity, has demonstrated therapeutic potential as evidenced by animal models and clinical trials, which have shown that minocycline can reduce negative symptoms in

schizophrenia. D-cycloserine (a tuberculosis antimicrobial) is also a *N*-methyl-D-aspartate (NMDA) receptor agonist and has been shown in controlled studies to enhance fear extinction when given before or after exposure to feared stimuli in animals and has also been shown to enhance antidepressant efficacy in one controlled study when used as an augmenting agent.

Antimicrobial drugs may also contribute to psychiatric disorders. For example, in a review of 82 unpublished case reports to the World Health Organization of antibiotic-induced mania, 27.6 percent were attributed to clarithromycin (Biaxin), 14.4 percent to ciprofloxacin (Ciloxan), and 12 percent to ofloxacin (Floxin).

More recently, emerging data raise questions whether the reverse may also be true, that certain psychiatric medications may have antimicrobial effects. Antipsychotics, for example, have demonstrated an inhibitory effect on several neurotropic viruses, including herpes simplex, and on several protozoans, including *Leishmania, Trypanosomes,* and *T. gondii.* In vitro research now indicates that several antipsychotics (in particular haloperidol [Haldol]) and the mood stabilizer valproic acid (Depakene) are capable of inhibiting the growth of *T. gondii,* an intracellular protozoan that can cause neuropsychiatric disorders. Because studies have reported increased levels of *T. gondii* antibodies in the serum of individuals with schizophrenia and mood disorders, the possible antimicrobial role of certain antipsychotics and mood stabilizers is of particular interest. In a similar vein, serotonin reuptake inhibitors (SSRIs), such as sertraline (Zoloft), have been reported in vitro to have antimicrobial activity against such organisms as staphylococci, enterococci, *Bacteroides fragilis, Brucellae,* and *Pseudomonas aeruginosa.* Another report suggested that sertraline engages in synergistic action, increasing the activity of some antibiotics such as tetracyclines and fluoroquinolones. Finally, certain psychiatric medications, such as antidepressants, have been shown to have anti-inflammatory effects through suppression of interferon gamma.

This line of research, while still highly exploratory, demonstrates the increasingly fruitful interdisciplinary investigations linking infectious disease, immunology, neurology, and psychiatry.

Prenatal Infections during Pregnancy

Prenatal infections have been found to play a role in the etiology of schizophrenia, and prenatal exposure to influenza, increased toxoplasma antibody, genital-reproductive infections, rubella, and other pathogens have been associated with the increased risk of developing schizophrenia. It is theorized that prenatal infection may affect the neurodevelopment of offspring (and later development of psychopathology, such as schizophrenia) through a mechanism of maternal immune activation.

REFERENCES

Abouesh A, Stone C, Hobbs WR. Antimicrobial-induced mania (Antibiomania): a review of spontaneous reports. *J Clin Psychopharm*. 2002;22:71–81.

Alter HJ, Mikovits JA, Switzer WM, et al. A multicenter blinded analysis indicates no association between chronic fatigue syndrome/myalgic encephalomyelitis and either xenotropic murine leukemia virus-related virus or polytropic murine leukemia virus. *MBio*. 2012;3(5):pii: e00266-12.

Bode L, Zimmermann W, Ferszt R, Steinbach F, Ludwig H. Borna disease virus genome transcribed and expressed in psychiatric patients. *Nat Med*. 1995;1:232–236.

Breitschwerdt EB, Sontakke S, Hopkins S. Neurological manifestations of Bartonellosis in immunocompetent patients: a composite of reports from 2005–2012. *J Neuroparasitology*. 2012;3:1–15.

Brown P, Cathala F, Castaigne P, Gajdusek DC. Creutzfeldt-Jakob disease: clinical analysis of a consecutive series of 230 neuropathologically verified cases. *Ann Neurol*. 1986;20:597–602.

*Brown AS, Susser ES. In utero infection and adult schizophrenia. *Ment Retard Dev Disabil Res Rev*. 2002;8:51–57.

Carson PJ, Konewko P, Wold KS, Mariani P, Goli S. Long-term clinical and neuropsychological outcomes of West Nile virus infection. *Clin Infect Dis*. 2006;43:723–730.

Coyle PK, Schutzer SE, Deng Z, Krupp LB, Belman AL. Detection of *Borrelia burgdorferi* specific antigen in antibody-negative cerebrospinal fluid in neurologic Lyme disease. *Neurology*. 1995;45:2010–2015.

DeLuca J, Johnson SK, Ellis SP, Natelson BH. Cognitive functioning is impaired in patients with chronic fatigue syndrome devoid of psychiatric disease. *J Neurol Neurosurg Psychiatry*. 1997;62:151–155.

Fallon BA, Keilp JG, Corbera KM, Petkova E, Britton CB. A randomized, placebo-controlled trial of IV antibiotic therapy for Lyme encephalopathy. *Neurology*. 2008;70(13):986.

*Fallon BA, Nields JA. Lyme disease: a neuropsychiatric illness. *Am J Psychiatry*. 1994;151:1571–1583.

Heim C, Wagner D, Maloney E, et al. Early adverse experience and risk for chronic fatigue syndrome: results from a population-based study. *Arch Gen Psychiatry*. 2006;63:1258–1266.

Hooshmand H, Escobar MR, Kopf SW. Neurosyphilis: a study of 241 patients. *JAMA*. 1972;219:726–729.

*Hornig M, Montoya JG, Klimas NG, et al. Distinct plasma immune signatures in ME/CFS are present early in the course of illness. *Sci Adv*. 2015;1:1–10.

*Ikuta K, Ibrahim MS, Kobayashi T, Tomonaga K. Borna disease virus and infection in humans. *Front Biosci*. 2002;7:470–495.

Institute of Medicine of the National Academies. *Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness*. Washington, DC; 2015.

Jackson GS, Collinge J. The molecular pathology of CJD: old and new variants. *J Clin Pathol Mol Pathol*. 2001;54:393–399.

Jones-Brando L, Torrey EF, Yolken R. Drugs used in the treatment of schizophrenia and bipolar disorder inhibit the replication of *Toxoplasma gondii*. *Schizophr Res*. 2003;62(3):237–244.

Kaneko M, Sugiyama N, Sasayama D, Yamaoko K, Miyakawa T. Prion disease causes less severe lesions in human hippocampus than other parts of brain. *Psychiatry Clin Neurosci*. 2008;62(3):264–270.

Komaroff A. Is human herpesvirus 6 a trigger for chronic fatigue syndrome? *J Clin Virol*. 2006;37:S39–S46.

*Lipkin WI, Hornig M. Psychotropic viruses. *Curr Opin Microbiol*. 2004;7:420–425.

Lyall M, Peakman M, Wessely S. A systematic review and critical evaluation of the immunology of chronic fatigue syndrome. *J Psychosom Res*. 2003;55:79–90.

Mays CE, Kang HE, Kim Y, Shim SH, Bang JE. CRBL cells: establishment, characterization and susceptibility to prion infection. *Brain Research*. 2008;1208:170–180.

Munoz-Bellido JL, Munoz-Criado S, Garcia-Rodriguez JA. Antimicrobial activity of psychotropic drugs: selective serotonin reuptake inhibitors. *Int J Antimicrob Agents*. 2000;14:177–180.

Murphy TK, Gerardi DM, Leckman JF. Pediatric acute-onset neuropsychiatric syndrome. *Psychiatr Clin North Am*. 2014;37(3):353–374.

Pachner AR. Borrelia burgdorferi in the nervous system: the new "Great Imitator." *Ann N Y Acad Sci*. 1988;539:56–64.

Prasad K, Watson AM, Dickerson FB, Yolken RH, Nimgaonkar VL. Exposure to herpes simplex virus type 1 and cognitive impairment in individuals with schizophrenia. *Schizophr Bull*. 2012;38(6):1137–1148.

Sanchez-Valle R, Arostegui JI, Yague J, Rami L, Llado A. First demonstrated de novo insertion in the prion protein gene in a young patient with dementia. *J Neurology Neurosurgery Psychiatry*. 2008;79(7):845–846.

Sauder C, Muller A, Cubitt B, Mayer J, Steinmetz J. Detection of Borna disease virus (BDV) antibodies and BDV RNA in psychiatric patients: evidence for high sequence conservation of human blood-derived BDV RNA. *J Virol*. 1996;70:7713–7724.

Schutzer SE, Angel TE, Liu T, et al. Distinct cerebrospinal fluid proteomes differentiate post-treatment Lyme disease from chronic fatigue syndrome. *PLoS One*. 2011;6(2):e17287.

Stewart LA, Rydzewska LH, Keogh GF, Knight RS. Systematic review of therapeutic interventions in human prion disease. *Neurology*. 2008;70(15):1272–1281.

*Swedo SE. Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections (PANDAS). *Mol Psychiatry*. 2002;7:S24–S25.

Thomas EW. *Syphilis: Its Course and Management*. New York: Macmillan; 1949.

Torrey EF, Bartko JJ, Lun ZR, Yolken RH. Antibodies to *Toxoplasma gondii* in patients with schizophrenia: a meta-analysis. *Schizophr Bull*. 2007;33:729.

Waltrip RW, Buchanan RW, Summerflet A, Breier A, Carpenter WT. Borna disease virus and schizophrenia. *Psychiatry Res*. 1995;56:33–44.

Wessely S, Chalder T, Hirsch S, Pawlikowska T, Wallace P. Postinfectious fatigue: prospective cohort study in primary care. *Lancet*. 1995;345:1333–1338.

Yolken RH, Torrey EF. Are some cases of psychosis caused by microbial agents? A review of the evidence. *Mol Psychiatry*. 2008;13(5):470–479.

▲ 2.10 Neuropsychiatric Aspects of Prion Disease

Transmissible spongiform encephalopathies (TSEs) are an unusual and uncommon group of infectious neurodegenerative disorders that are caused by conformational changes, misprocessing, and malfunction of the prion protein (*PrP*). The neurodegenerative disorders caused by prions can present as genetic (about 14 percent of cases), sporadic (85 percent), and acquired infectious disorders of human CNS (1 percent). TSEs affect humans and several animals and invariably have a fatal outcome. The human prion diseases are sporadic Creutzfeldt–Jakob disease (sCJD), the variant Creutzfeldt–Jakob disease (vCJD), iatrogenic Creutzfeldt–Jakob disease (iCJD), fatal familial insomnia (FFI), Gerstmann–Sträussler–Scheinker (GSS) disease, and kuru (Table 2.10–1). The prion diseases of animal hosts include scrapie, BSE (commonly known as mad cow disease), chronic wasting disease (CWD), related bovine spongiform encephalopathy (rBSE), and transmissible mink encephalopathy.

As an uncommon cause of dementia, in humans, prion diseases present as rapidly progressive cognitive and behavioral changes associated with visual and cerebellar impairments, extrapyramidal and pyramidal signs and symptoms, and akinetic mutism. Additionally, these disorders as a group share a spongiform degeneration of the brain and a variable amyloid plaque formation.

During the past two decades, there have been growing worldwide concerns regarding BSE outbreaks in various areas of the world, particularly the United Kingdom and its transmission to human hosts in the form of vCJD. Thus, vCJD is another major concern for scientists and physicians, although it incidence has declined significantly since the past century. vCJD presents with unique clinical and neuropathological manifestations, which set it apart from sCJD. Clinically, the vCJD form differs from sCJD in that it has early psychiatric symptoms such as anxiety, depression, apathy, and social withdrawal with dementia being a late complication. A significant recent finding is that the causative agents of vCJD in humans and BSE in cattle share a common origin and differ from the causative agent of sCJD. This discovery stimulated the scientific community's interest in TSEs in general and in the possibility of their transmission from animals to humans in particular.

Prions cause TSEs. In 1982, Stanley B. Prusiner proposed this term to distinguish this group of disorders from other conventional infectious disorders (such as viral diseases) and pointed out that prions are *pro*teinaceous *in*fectious particles that resist inactivation by most procedures that denature nucleic acids (Fig. 2.10–1). Prions have emerged as a new group of infective agents that are composed principally of abnormal isoforms of a host-coded glycoprotein, *PrP*. The *PrP* gene is

located at the short arm of chromosome 20 (20pter-p12) (Table 2.10–2).

DEFINITION

TSEs, which are also known as prion diseases, are unique infectious and invariably fatal neurodegenerative disorders of humans and animals that result from the misfolding of a normal cell protein into an abnormal protein. The abnormal protein gains toxic properties that coerce or cause "normal" PrPs to be transformed by conformational changes into the toxic forms.

Table 2.10–1.
Human and Animal Prion Diseases

Disease	Cause	Distribution/Incidence
Sporadic Creutzfeldt–Jakob disease (sCJD)	Unknown	Global
Variant Creutzfeldt–Jakob disease (vCJD)	Exposure to bovine spongiform encephalopathy (BSE)	As of 2006, more than 150 cases of vCJD have been recorded and all were associated with methionine homozygotic status at codon 129 of *PRNP* gene
Iatrogenic Creutzfeldt–Jakob disease (iCJD)	Genetic	Accidental transmission of CJD to human hosts through various medical/surgical procedures such as tissue transplantation
Fatal familial insomnia (FFI)	Familial	Rare
Gerstmann–Sträussler–Scheinker disease (GSS)	Genetic	Extremely rare
Kuru	Ritual cannibalism	Papua
Scrapie	Unknown	Europe, Iceland, United States, Canada
BSE (mad cow disease)	Animal feed with animal body parts, initially from sheep	Europe, United States
Chronic wasting disease (CWD)	Caged elk and deer	United States, Canada
Transmissible mink encephalopathy	Farm-raised	United States

From Dormont D. Prion disease: pathogenesis and public health concerns. *FEBS Lett*. 2002;529:17.

COMPREHENSIVE NOSOLOGY

TSEs are fatal infectious and neurodegenerative diseases of human and other mammals. These diseases appear as variations on a theme, PrP

conformational changes, and result in several phenotypes but have a common end point, dementia, and death. The protein-only hypothesis states that a small proteinaceous infectious particle, without any nucleic acid, is the transmissible agent that causes neurodegenerative diseases in the susceptible hosts. Prions resist inactivation procedures that denature proteins and nucleic acids. The neuropathological hallmark of prion diseases is the misfolding of the PrP in the brain of affected patients.

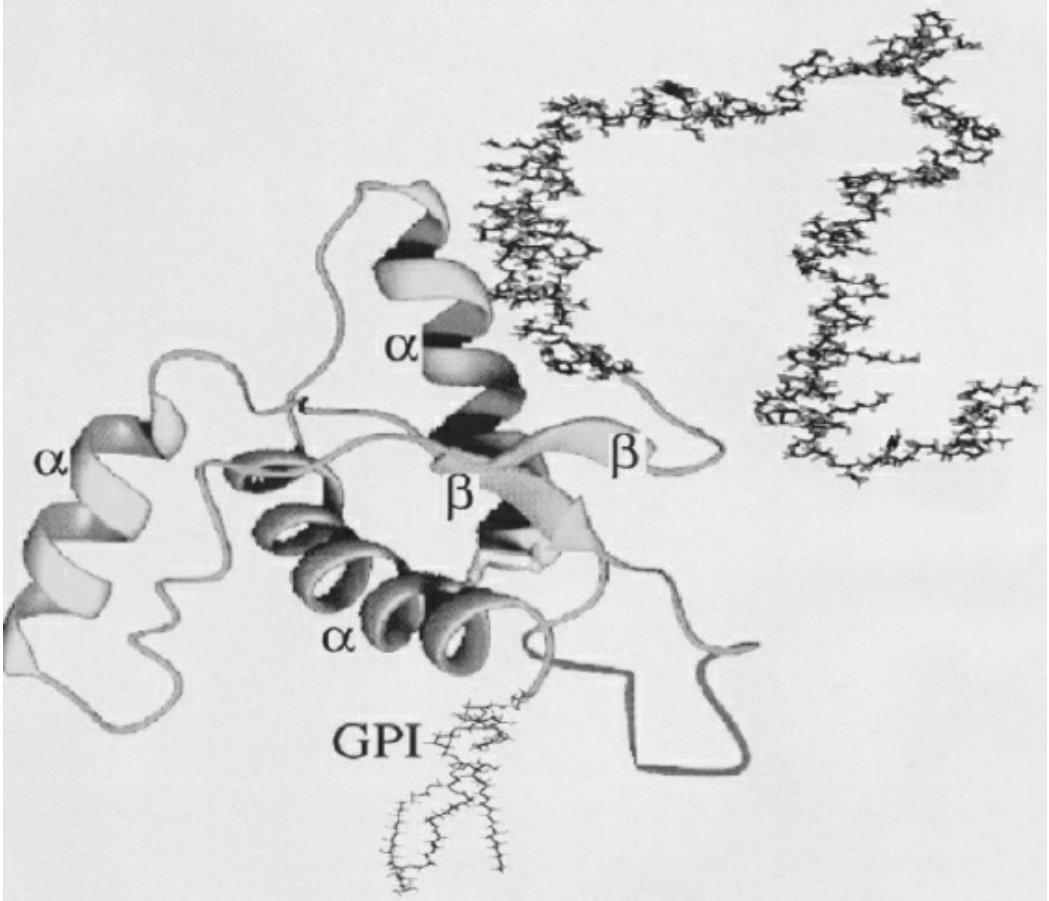

FIGURE 2.10–1. Three-dimensional structure of the cellular prion protein. (From Dormont D. Prion diseases: Pathogenesis and public health concerns. *FEBS Lett*. 2002;529(1):17 [Review]. FEBS Letter Biochemie-Zentrum Heidelberg with permission.)

NONFAMILIAL PRION DISEASES

Sporadic (Nonfamilial) CJD

sCJD represents the most frequent form of CJD that occurs around the globe at a rate of one per million individuals and mainly affects older adults.

FAMILIAL GENETIC PRION DISEASES

Familial CJD

fCJD is defined as a form of CJD that results from mutation(s) in the gene

that encodes PrP. fCJD shows an autosomal dominance pattern of inheritance. About half of the individuals carrying the mutated form of *PrP* gene will develop CJD. Although clinically the presentation of fCJD varies with site of mutation within the *PrP* gene, it most closely resembles sCJD in its manifestations.

Fatal Familial Insomnia

FFI is a hereditary prion disease that is identified by disrupted sleep, motor abnormalities, and hyperactivation of the autonomic nervous system. As with fCJD, it has an autosomal dominant pattern of inheritance with an atypical presentation and exclusive bilateral degeneration of the thalamus without the typical neuropathological changes of spongiform degeneration or amyloid deposits.

Gerstmann–Sträussler–Scheinker Syndrome

GSS is another human prion disease that typically occurs in the fourth or fifth decade and mainly manifests with cerebellar ataxia and motor deficits. It has a familial autosomal dominant pattern of inheritance.

Acquired Prion Diseases

Kuru is a prion disease that was discovered in geographically and ethnically isolated tribes of the Fore highlands of New Guinea. In native language of the Fore highlands tribes kuru means "trembling with fear" and this disease was related to endocannibalistic funeral practice where deceased family members were ritualistically cooked and consumed by other family members. During this ceremony, the closest female relatives and children usually ate the deceased individual's brain, and this led to the fact that kuru's victims were primarily women and children.

Table 2.10–2.
Prion Chromosome Locations

Name or Disease/Abbreviation	Chromosome Location	OMIM #	Mutation
Prion gene complex, downstream (Doppel) *PRND*, *DPL*	20pter-p12	604263	Has some similarity to prion protein cell (PrP^C) mutations
Sporadic Creutzfeldt–Jakob disease (sCJD)	20pter-p12	176640, 123400	met129 and met/val129
Creutzfeldt–Jakob disease (CJD)	20pter-p12	176640, 123400	met129
Variant Creutzfeldt–Jakob disease (vCJD)	20pter-p12	176640, 123400	val129 provides resistance to vCJD prion
Kuru	20pter-p12	245300	—
Gerstmann–Sträussler–Scheinker	20pter-p12	137440	—

disease (GSS)			
Familial fatal insomnia (FFI)	20pter-p12	600072	asp178-to-asn (D178N)
Spongiform encephalopathy with neuropsychiatric features	20pter-p12	—	his187-to-arg (H187R)
Huntington-like variant Creutzfeldt-Jakob disease	20pter-p12	—	8-octapeptide repeat near val129
Bovine spongiform encephalopathy (BSE)	13q17	—	Mutations similar to those found in vCJD prion infecting humans
Scrapie	OA 13q17/q18, CHI13q15	—	—

Adapted from Hall DA, Leehey MA, Filley CM, et al. PRNP H187R mutation associated with neuropsychiatric disorders in childhood and dementia. *Neurology*. 2005;64:1304; OMIM 1. http://www.ncbi.nlm.nih.gov/Omim/getmap.cgi; Wadsworth JD, Asante EA, Desbruslais M, Linehan JM, Joiner S. Human prion protein with valine 129 prevents expression of variant CJD phenotype. *Science*. 2004;306:1793.

iCJD is completely dependent on human-to-human transmission most commonly by medical procedures with a highly variable incubation period. The presentation of iCJD also varies in that iCJD typically presents with motor findings (ataxia and gait abnormalities) associated with only mild dementing signs and symptoms toward the later phase of the disease.

Nonhuman Mammalian Spongiform Encephalopathies

Scrapie is a fatal disease of sheep and goats that is characterized by chronic itching, loss of coordination, and a progressive degeneration of the CNS.

Transmissible mink encephalopathies are defined as infectious neurodegenerative diseases that affect human and animal hosts.

CWD is a TSE that affects deer, elk, and moose.

BSE is an infectious disease, affecting cattle. In humans, this is also referred to as *new variant CJD* (nvCJD), which was identified in Great Britain and shown to be due to transmission across mammalian species to man. It is derived from the same infectious agent as BSE and hence the reference to “mad cows.” The neuropathology resembles that of scrapie as opposed to sCJD (the main neuropathological feature of scrapie is the presence of vacuolated neurons, while in sCJD vacuolation of the neuropil between nerve cell bodies is more characteristic), and it can be detected in peripheral organs such as appendix, tonsils, spleen, and lymphatic tissue as well and nervous system tissue.

It should be mentioned that prion-like proteins also occur in fungi including yeast. This is important to note here because it demonstrates the importance of prion-like proteins in normal function if such proteins are present in fungi, denizens of the earth that arose early in the origin of life. Two proteins of yeast, the most studied of fungi, ure2p and sup35p, are partially protease resistant and agencies into amyloid-like filaments.

Moreover, the yeast "prion" proteins serve as a paradigm for the animal and human prion diseases in that they also undergo conformational changes in their function, propagation, and strain variation. However, the yeast prions are not infectious, but by introduction into different strains of yeast demonstrate the "seeding" phenomenon. This is thus supportive of a more widespread phenomenon than previously imagined in living systems biology. A chaperon protein may be involved in yeast prion function; yeast prion sequences have no resemblance to prions from higher orders on the evolutionary chain.

HISTORY OF PRION DISEASES

The history of prion diseases commenced in the year 1920, when Hans Gerhard Creutzfeldt reported a 22-year-old woman with a mysterious and progressive focal syndrome of the CNS that was clinically identified by psychomotor abnormalities and cortical symptoms. Autopsy of this patient revealed prominent gliosis with noninflammatory focal lesions of the cerebral cortex. In 1921, Alfons Maria Jakob observed three additional cases of chronic neurological disease, which he considered as a new entity; in these patients Jakob reported a spastic psuedosclerosis and encephalomyelitis with spread focal degeneration. A few years later, Walter Spielmeyer reviewed these cases and based on their similarities, recommended the name CJD. Other prion diseases, GSS syndrome and FFI were described in 1936 and 1986, respectively. In 1957, Carlton Gajdusek traveled to the Kuru region of the New Guinea where Walter Zigas was positioned as a medical officer. In a collaborative effort, these physician-scientists described the clinical and neuropathological features of kuru and reported that kuru may be associated with cannibalism. In 1959, Hadlow distinguished certain similarities between kuru and the spongiform encephalopathy of sheep scrapie. This astute observation led to a series of experiments to transmit kuru to nonhuman primates and further discovery of salient similarities between kuru and the spongiform encephalopathy of sheep scrapie. However, studies in the United Kingdom failed to demonstrate any form of relationships between sCJD and scrapie. Further research into the nature and pathogenesis of prion diseases was stimulated after the discovery that BSE in the United Kingdom was causally related to vCJD. It is to Gajdusek credit that he was able to reduce the incidence of kuru among the Fore people by convincing them to halt ritual cannibalism.

The precise nature of the transmissible infective agent in TSEs has been the subject of intense debate and speculations for several decades. Initially, it was hypothesized that slow viruses cause TSEs. However, a lack of any solid immune-mediated response against any viral agents and the unusual resistance to certain treatment procedures that normally denature and inactivate nucleic acids indicated that TSEs are viral disorders. These observations led the scientists to propose that "an agent devoid of nucleic

acids" or a protein may be the causative agent of TSEs. In 1982, Bolton isolated a protease-resistant sialoglycoprotein, designated as *PrP,* from brain homogenates. Later that year, Stanley Prusiner coined the term prion (from proteinaceous infectious particle). In fact, prions were described as "small proteinaceous infective particles" that resist inactivation by procedures that modify nucleic acids.

EPIDEMIOLOGY OF PRION DISEASES

Ritual cannibalism, which was practiced by Fore tribes in New Guinea, transmits kuru to females and children of both genders who eat the brains of dead individuals. Brain tissue possesses high amounts of transmissible infective agents and its consumption transmits disease to these two groups in that population. Although kuru is transmitted horizontally, with cessation of cannibalism rituals in this area of the world, transmission of kuru in individuals born after cannibalism was discontinued has not been witnessed.

The epidemiological assessment of CJD only exists in certain countries where surveillance units follow development of every new case. CJD can be sporadic, genetic, or iatrogenic. Patients with CJD die rapidly; therefore, annual statistics for incidence, prevalence, and mortality rates are analogous. Generally, the worldwide annual incidence of CJD is estimated at 1 million per year. Human prion diseases affect most populations around the globe with a frequency of 1 to 1.5 cases per million, with some variation from country to country. In certain countries where constant surveillance for prion disease is carried out, the incidence of sCJD has been reported as approximately 0.6 to 1.2 × 10^6. The United States, with a population of approximately 270 million, averages 300 diagnosed cases of prion diseases annually. Human prion diseases occur in three forms: sporadic (85 percent), genetic (14 percent), and acquired (1 percent).

GSS, similar to CJD, is not one single disease, and there are at least six dominantly inherited syndromes with dissimilar clinical features, underlying neuropathology, and linkage to different mutation of the *PRNP* gene. GSS is an extremely rare neurodegenerative disorder with an incidence of two to five cases per 100 million.

FFI is a midlife disease that affects both genders equally. The age of onset of FFI ranges from 36 to 62 years, and the disease is uniformly fatal and causes death in 8 to 72 months with an average of 18 months.

PRION MOLECULAR BIOLOGY AND PATHOGENESIS

Prion Introduction

Analysis of the biological, biochemical, and molecular activity and properties of *prion proteins* PrP^{C} (*cell*) and PrP^{Sc} (scrapie) enables one to better understand the pathogenesis of prion diseases. Research during the

past three decades supports the concept that prion diseases result from conversion of the host-encoded cell-surface glycoprotein, *PrP*^{c}, into abnormally and conformationally altered isoforms, *PrP*^{Sc}. These studies also analyze mechanisms of how the structure and function of prions of form *PrP*^{c} are altered to prions of form *PrP*^{Sc}, and how these processes lead to the development of inexorable terminal progressive neurodegenerative disease. A fundamental finding that sets prions apart from any other known life form and also is contradictory to the original Watson-Crick "fundamental dogma" of molecular biology is that prion *PrP*^{Sc} proteins are infectious and can propagate in the absence of viral nucleic acids.

Molecular Pathogenesis

During the course of prion diseases, characteristic destructive CNS damage occurs in the form of spongiform vacuolation, widespread neuronal loss, microglial activation, proliferation of astrocytes, and *PrP*^{Sc} production. Great strides are being made in scientists understanding of prion diseases, tropism, and molecular pathogenesis. There is extensive study as to the mechanisms by which *PrP*^{Sc} prions gain access to the CNS after infection. Prions are generally transmissible by inoculation into subhuman primates, hamsters, and rodents. Intracerebral inoculation of brain homogenate is the most rapid method for inducing spongiform encephalopathy; the inoculation of, for example, 10^6 infectious units will cause disease within roughly 6 months in the animal host. (Units are typically defined as 50 percent infectious dose.) Additionally, TSE can be induced through peripheral routes including intravenous (IV) and intraperitoneal injections, feeding, and via various aspects of the eye. Iatrogenic needle sticks, scalpel cuts, blood transfusions, and tissue grafts also result in human disease transmission of CJD, and as mentioned above, cannibalism transmits kuru (Fig. 2.10–2).

One model of prion brain invasion following ingestion of prions involves early infection of the Schwann cell (the glial myelin-producing cell of the peripheral nervous system) followed by cell–cell spread centripetally along peripheral nerves finally into the spinal cord and then the brain. However, this mechanism does not involve retrograde axonal flow that is too rapid a mechanism compared to the weeks to months that it takes to reach the brain. However, a neuronal synaptic spread of prions via neuron axons in a domino-like fashion has been proposed as well for streamlined prion entry into the brain from the periphery.

PrP^{c} concentration is highest in CNS (brain and spinal cord) neurons at both early stages of embryogenesis and also in the adult, specifically in association with synaptic membranes as well as nicotinic acid/acetyl choline and NMDA receptors. Tissues infected by *PrP*^{Sc} prions extend beyond the CNS, including spleen and muscle tissue of patients with sCJD. In addition, chronic inflammation can broaden the tropism of prion

infectivity to other tissues including liver, pancreas, kidney, muscle, and mammary glands, although these tissues were originally considered to be prion free. However, in 2003 Adriano, Aguzzi and colleagues found that lymph nodes have a higher prion load than spleen and that possibly lymph node infections do not involve dendritic cells that are generally present in lymph nodes.

Prions also reside within lymphoid/lymphoreticular system (LRS) compartment of the infected host. Thus, following oral administration of prions in mice, prion protein is detectable in intestinal Peyer patches using immunohistochemistry. Similarly, prions are also present in primary B-cell follicles and germinal centers of secondary lymphoid tissues including the appendix and tonsils. Possibly, then Peyer patches may function as a primary gateway for orally administered prions from where they eventually gain access to the CNS.

Apart from lymphoid tissue, PrP^{C} itself promotes PrP^{Sc} transport from the periphery to the CNS. Mice that lacked the PrP^{C} gene (termed $Prnp^{0/0}$ knockout) were resistant to PrP^{Sc} infection. Wild-type bone marrow transferred adoptively into $Prnp^{0/0}$ knockout mice restored the capability of the spleen to accumulate high titers of PrP^{Sc}. However, there continued to be an absence of CNS prion infection. Thus, B and T lymphocytes and macrophages aid prion transfer from the peripheral entry site to the secondary lymphoid tissues but not further into the CNS.

Cell Function Properties

PrP^{C} may exert antiapoptotic protective activity in mammalian cells, mice, and yeast, which is due to internal or environmental stress factors that initiate apoptosis. For example, PrP^{C} is able to protect human fetal neurons in culture against Bax-related pathway apoptosis.

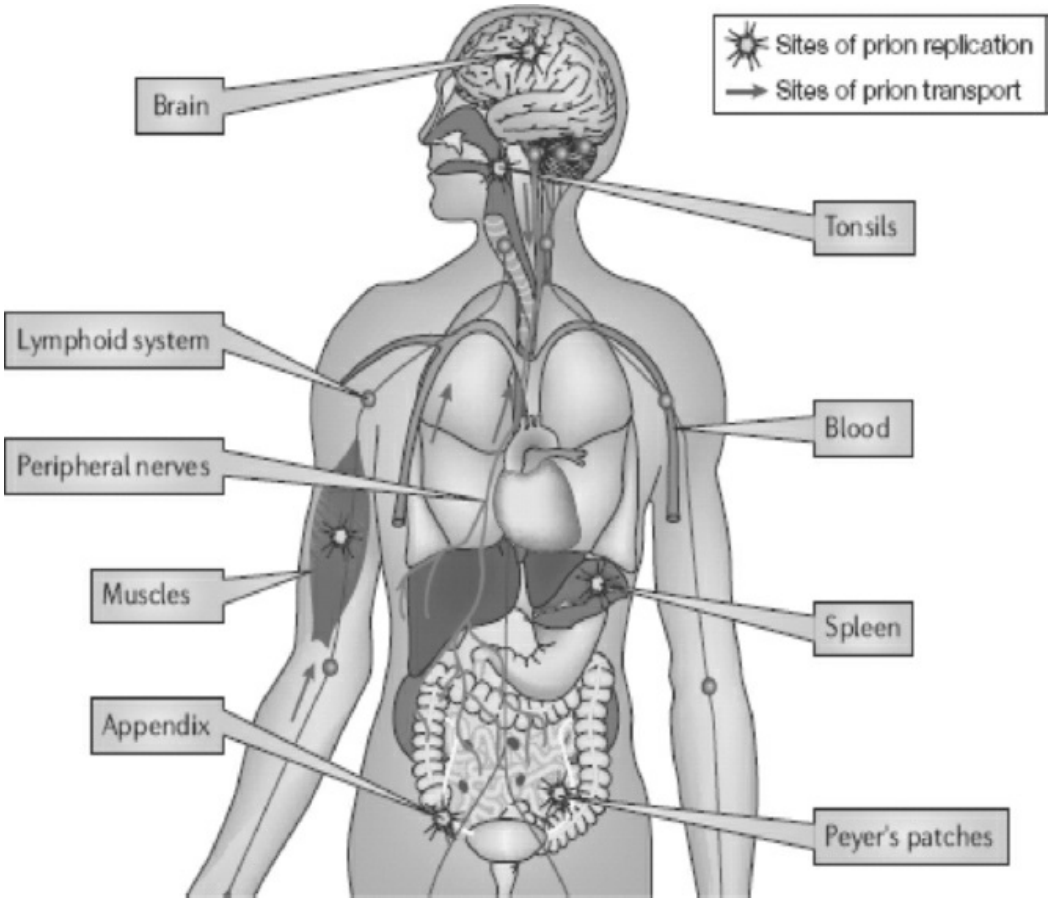

FIGURE 2.10–2. Human tissues and blood involved in propagation and transport of prions. Orally ingested prions are intestinally absorbed and transported to the blood and lymphoid fluids. After a peripheral replication step in the spleen, appendix, tonsils, or other lymphoid tissues, prions are transported to the brain primarily by peripheral nerves. Direct penetration into the brain across the blood–brain barrier is conceivable. (From Aguzzi A, Heikenwalder M. Pathogenesis of prion disease: Current status and future outlook. *Nat Rev Microbiol*. 2006;4(10):765. Review. Nature Publishing Group with permission.)

Oxidative stress affects a number of interconnected cellular pathways and can potentially cause mitochondrial dysfunction, damage to the ubiquitin protease system, aggregation of proteins, altered iron metabolism, and excitotoxicity. A number of observations support a potential cytoprotective role for PrP^{C} against oxidative stress. PrP^{C} may protect cells from oxidative stress that is associated with super-oxide dismutase (SOD) activity. It has been hypothesized that chronic oxidative stress of neurons is a major promoter of neurodegenerative disorders; this is exemplified in neuro-AIDS and drug abuse in which there is an increase in neuronal nitric oxide synthase expression. It has been demonstrated that neurons (cerebellar granular and neocortical) cultured from $Prn\text{-}p^{o/o}$ mice are more susceptible than neurons from wild-type mice to treatments with oxidative agents such as hydrogen peroxide, xanthine oxidase, and copper irons. PrP^{C} appears to have both antiapoptotic and SOD activities that are obviously crucial neuron protective functions. However, it should be noted that work by other investigators find no SOD activity associated with PrP^{C}.

The still puzzling observation is that *PrP*^{C} has SOD activity but also contains copper ions within it normal structure. Discussion of the copper association with *PrP*^{C} follows.

Table 2.10–3.
Normal and Abnormal Forms of Prion Proteins

Features	Normal	Pathological/Disease
Name	*PRNP, PrP*^{C}	*PrP*^{Sc}, CJD, GSS, FFI, etc.
Molecular weight	Monomer, 22–36 Kd	Stacked aggregates, >400 Kd
Cellular location	External, GPI-linked at cell surface, synaptic clefts	Internal, Golgi, vesicles
Sensitivity to protease K	Yes	No
Biochemical structure	40% α-helices 3% β-sheets	40% β-pleated sheets 30% α-helices
Cellular location	GPI-linked at cell surface, synaptic clefts	Internal Golgi, vesicles, extracellular
Detergent miscibility	Miscible	Fibrils and rods form
Human tissue distribution	Brain, lymphocytes, heart, lung	Brain, lymphoid tissue
Concentration in normal brain	Variable, 1–5 μg/g tissues	Absent
Sensitivity to typical autoclave	—	No
Sensitivity to 4N NaOH hydrolysis	—	Yes
Apoptosis	Protective	Not protective
Superoxide dismutase	Yes	No
Binds Cu^{+2} cations	Yes	No
Coupled to tyrosine kinase Fyn	Yes	No

PrP^{C}, prion protein cell; *PrP*^{Sc}, prion protein scrapie; CJD, Creutzfeldt-Jakob disease; GSS, Gerstmann–Sträussler–Scheinker disease; FFI, familial fatal insomnia; GPI, glycosyl phosphatidyl inositol.

Data from Irani DN, Johnson RT. Diagnosis and prevention of bovine spongiform encephalopathy and variant Creutzfeldt-Jakob disease. *Annu Rev Med*. 2003;54:305; Johnson RT. *Viral Infections of the Nervous System*. 2nd ed. Philadelphia, PA: Lippincott-Raven; 1998; Roucou X, Gains M, LeBlanc AC. Neuroprotective functions of prion protein. *J Neurosci Res*. 2004;75(2):153; Van Rheede T, Smolenaars MMW, Madsen O, de Jong WW. Molecular evolution of the mammalian prion protein. *Mol Biol Evol*. 2003;20:111.

PrP^{C} may be a copper-binding protein. The histidine-containing octapeptide repeats bind up to four Cu^{+2} ions in a pH-dependent and negatively cooperative style. This binding interaction between *PrP*^{C} and Cu^{+2} ions changes the biological and biochemical functions of the *PrP*^{C}. Endocytosis of *PrP*^{C} via clathrin-coated pits is stimulated by micromolar concentrations of Cu^{+2} ions (Table 2.10–3).

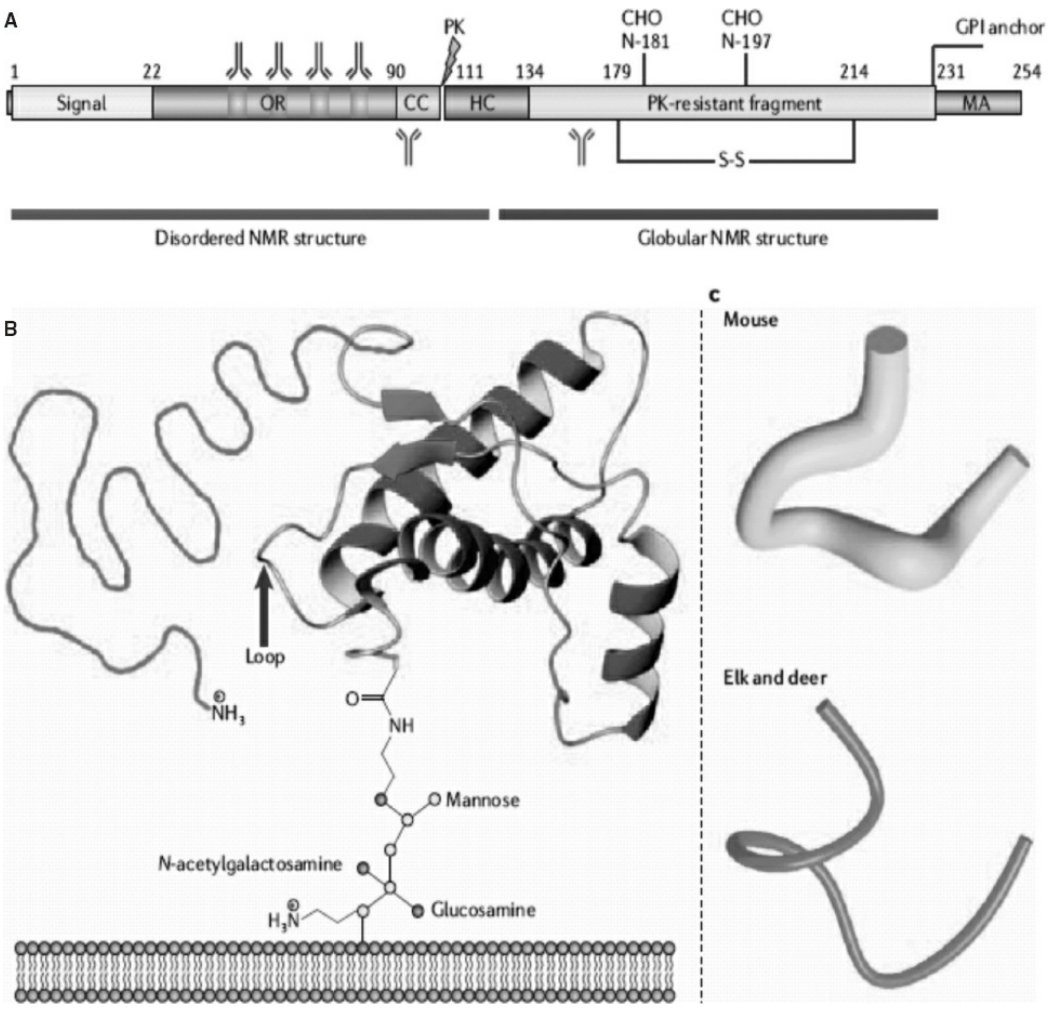

FIGURE 2.10–3. Structural features of the cellular prion protein (PrP). **A:** An outline of the primary structure of the cellular PrP including posttranslational modifications. A secretory signal peptide resides at the extreme N-terminus. The numbers describe the position of the respective amino acids. CC defines the charged cluster. HC defines the hydrophobic core. S-S indicates the single disulfide bridge. The proteinase K (PK) resistant core of PrP^{Sc} is indicated by the lightning symbol. MA denotes the membrane anchor region. The epitopes recognized by the POM antibodies, some of which have extremely high affinities, are also indicated. According to competition assays in solution in surface plasmon resonance assays, POM2 binds to residues 58–64, 66–72, 74–80, and 82–88 (QPXXGG/SW); POM3 to residues 95–100 (HNQWNK), and POM5 to residues position 168–174. **B:** Tertiary structure of the cellular prion protein inserted into a lipid bilayer, as deduced from NM spectroscopy, including the unstructured N-terminal tail (*gray*) and the glycosyl phosphatidyl inositol (GPI) anchor. The loop connecting the second β-sheet and the third α-helix is indicated by the *black arrow.* OR, octarepeat region. **C:** The loop region is extremely flexible in most species (e.g., mouse), but it is almost entirely rigid in the prion protein of elk and deer as indicated by the average of the three-dimensional space occupied during its oscillations. The figure shows amino acids 165 to 172 of the cellular prion protein of mouse, elk, and deer. (From Aguzzi A, Heikenwalder M. Pathogenesis of prion diseases: Current status and future outlook. *Nat Rev Microbiol*. 2006;4(10):765. Review. Nature Publishing Group with permission.)

Structure

Normally, PrP^{C} is expressed on the cell surface where it is anchored to the cell membrane lipid bilayer via an extraordinary carboxy-(C)-terminal,

glycosyl-phosphatidyl-inositol (GPI) anchor. *PrP*^{C}, similar to the other membrane proteins, is synthesized in endoplasmic reticulum–attached ribosomes, translocates to the Golgi apparatus, and finally attaches to the cell surface. *PrP*^{C} is a protease-sensitive sialoglycoprotein that possesses two N-linked complex oligosaccharide chains, similar to other GPI-anchored proteins; most of *PrP*^{C} is found in cholesterol-rich lipid membrane rafts (Fig. 2.10–3).

There are 23 enzymes involved in the biosynthesis of human GPI-linked proteins that are on 11 chromosomes, including two on chromosome 2: For example, one is GPIT, phosphatidyl-inositol glycan anchor biosynthetase, cl ss T, which transfers the fully assembled glyco-phosphatidyl-inositol moiety to proteins prior to membrane anchoring (GPI pathway-url). There are at least 6 GPI-linked proteins on neurons (NCBI URL). A search for GPI-anchored human proteins at NCBI indicates that there are several hundred. Thus, compared to the entire proteome, there are few GPI-linked proteins. The function of this linkage is under investigation. The *PrP*^{C} molecule contains 250 amino acids and possesses two consensus sequences for N-linked glycosylation with molecular variants that have different degrees of glycosylation: Unglycosylated, monoglycosylated, and diglycosylated. Some of *PrP*^{C} is transferred to clathrin-coated pits where it undergoes endocytosis and recycling. *PrP*^{C} from human, mouse, Syrian hamster, and cattle share common properties, including a flexible amino-terminal tail (residues 23 to 128), three α helices, and a two-stranded antiparallel β-sheet that flanks the α-helix. A flexible large loop connects the second β-sheet and the third α-helix. The N-terminal portion of the *PrP*^{C} contains two conserved regions: A region of five repeats of octameric amino acid sequence, identified as octapeptide repeat region (OR, octarepeat region), and a second region with a highly hydrophobicity (hydrophobic core), which is preceded by a hydrophilic domain known as charge cluster.

Table 2.10–4.
Prion Protein Structure

Location, Amino Acid Number[a]	Name	Comments
1–22	Signal peptide	Cleaved as part of protein maturation and migration
23–28	NLS-I	Nuclear localization signal
50	Pro	Proline hydroxylation site
56–128	Contains octapeptide repeats (variable)	Repeats of GWGQPHGG, with homology to TWIQDNGG (Bcl-2 BH2), and GWIQDNGG (Bax BH2)
151–163	Hydrophobic region	—
146/147	Lysine/histidine	Cleavage site
216 and 252	Cysteines	Disulfide bridge
218 and 234	ASN	Asparagines glycosylation site

164–168	B1	β region
180–190	H1	α helix
197–203	B2	β region
208–233	H2	α helix
237–269	H3	α helix
275	GPI	Glycolipid-phospho-inositol anchor
276–298	Signal peptide	Cleaved as part of protein maturation and migration

[a]It should be noted that there is variation in numbering among different publications. One cause is variation in the number of octapeptide repeats.

Adapted from Roucou X, Gains M, LeBlanc AC. Neuroprotective functions of prion protein. *J Neurosci Res*. 2004;75(2):153; Van Rheede T, Smolenaars MMW, Madsen O, de Jong WW. Molecular evolution of the mammalian prion protein. *Mol Biol Evol*. 2003;20:111.

The OR region of PrP^{C} is significant for a number of reasons. This region contains the histidine residues that are capable of binding copper with copper-binding locations within the OR region. Copper complexing is associated with endocytosis of PrP^{C} and through this plays a role in metabolism of this surface glycoprotein. The other reason for significance of the OR region is that the expansion of the OR domain of up to 13 total repeats is also associated with genetic CJD and GSS (Table 2.10–4).

PrP^{C} and PrP^{Sc} have the same primary amino acid sequence, and their different secondary structures impart dissimilar physiological and physicochemical properties (Table 2.10–2). These features are conserved in evolution as demonstrated in the phylogenetic tree in Figure 2.10–4. Study of their secondary structures demonstrated that PrP^{C} contains a high α-helical content (roughly 42 percent) and small β-sheet content (3 percent). The tertiary structure of PrP^{C} shows that the normal molecule possesses three α-helical domains and two β-strands. It is believed that PrP^{Sc} resistance to degradation by protease is related to alteration of its structure from a molecule with the α-helical structure into a different conformation molecule with a large β-sheet content. Accumulation of this protease-resistant isoform within neurons disrupts their normal function and results in vacuolization and widespread cell death. PrP^{Sc} through a self-promoting or -catalyzing mechanism converts normal PrP^{C} into PrP^{Sc} (Tables 2.10–3 and 2.10–4).

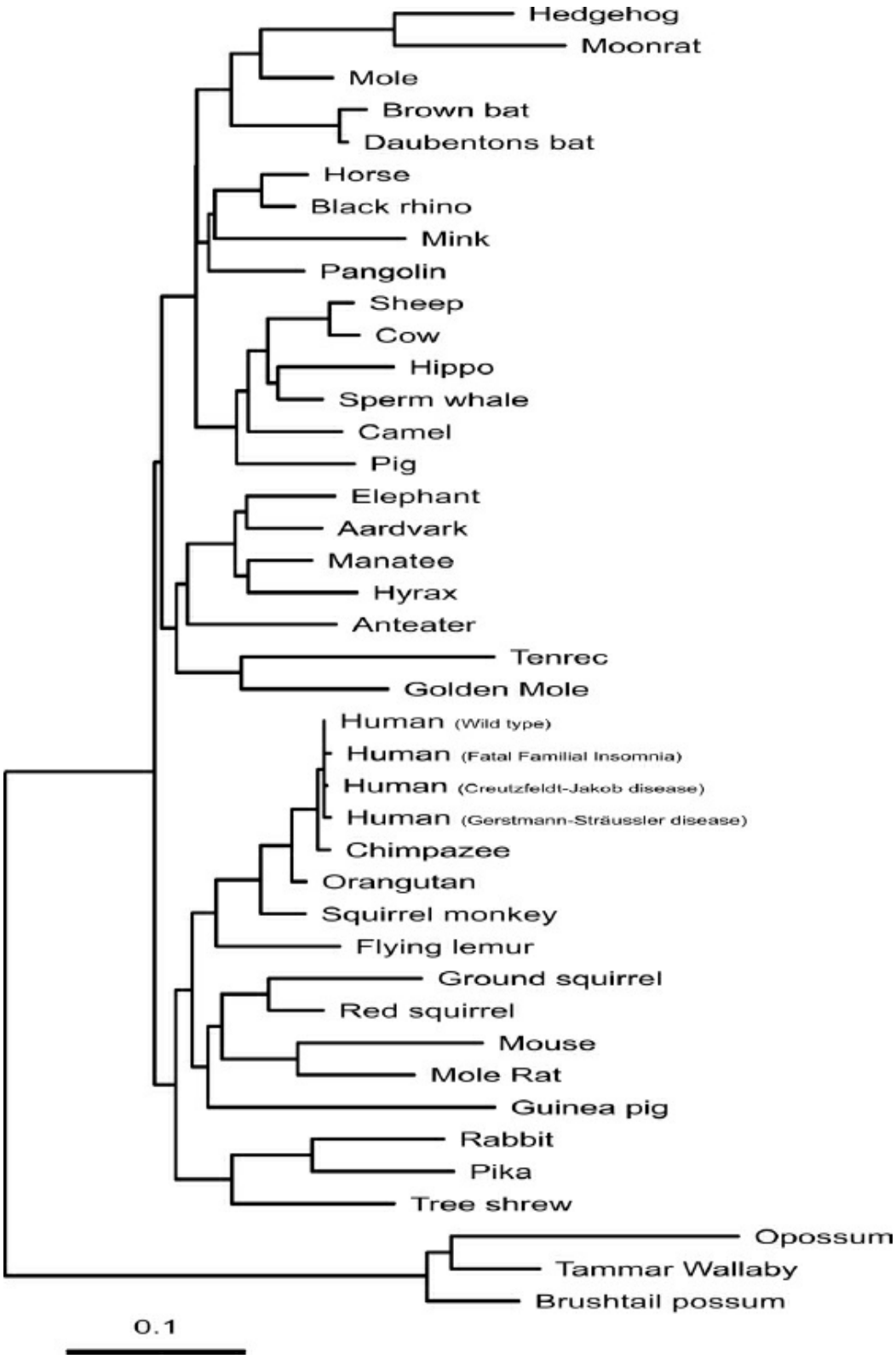

FIGURE 2.10–4. Phylogenetic tree based on prion sequences from humans and other mammals. This phylogenetic prion tree illustrates the basic evolution of prions among the different mammalian orders. This prion tree largely corresponds to the mammalian species tree, indicating that cellular prion protein has a normal function in human and mammalian cells. The close phylogenetic association of the different human disease prions shows that minor changes causes prion related diseases. (Contributed by Ole Madsen, Ph.D., Wageningen, The Netherlands, 2007.)

Function

Proposed functions for *PrP*^{C} include transmembrane signaling, cell adhesion, and synapses. A widely used mechanism for the role of *PrP*^{C} in pathogenesis of prion diseases is toxic gain of function, which occurs when *PrP*^{C} is converted to *PrP*^{Sc}, and this may be a crucial step for development of prion disease. In these terms, host *PrP*^{C} plays a significant role in determining the susceptibility of the host due to exposure to the deleterious prion *PrP*^{Sc}.

As mentioned above, mice deficient in the *PRNP* gene, homozygous knockouts (*PrP*$^{C-}$/*PrP*$^{C-}$) do not develop scrapie, implying that *PrP*^{C} is necessary for the development and propagation of prion disease. Thus, abnormal folding, protease resistance, and β-sheet properties result when normal host protein *PrP*^{C} is transformed by *PrP*^{Sc}.

PrP^{Sc} exhibits a number of toxic features that are unrelated to the physiological functions of *PrP*^{C}. *PrP*^{C} and *PrP*^{Sc} or *PrP-resistant* [*PrP*^{Res}] exhibit a number of differences that are used in their study.

Strains

Different strains of prions are associated with different prion diseases and their separate neuropathologies and the clinical signs and symptoms as described in this section. In animal model studies, biochemical traits were preserved through several passages in prion disease rodent models. Work by I.H. Pattison and G.C. Millson in 1961 provided the first evidence for the presence and production of various prion strains. Goats infected by the same preparation of infectious scrapie agent developed two different clinical syndromes: Scratching and drowsy. Prion strains are infectious isolates that show distinct prion-disease phenotypes, including incubation periods, neuropathology, and specific neuronal targets in several regions of brain, in the same hosts, and with serial transmission. Several mechanisms have been proposed, including that prion strain-specific phenotypic traits may be encoded by an ancillary genome, due to specific *PrP*^{Sc} genetics and conformations in donor *inocula,* selected postinfection, or by host genetics.

Host Protein Interactions

PrP^{C} interacts with a number of other cellular proteins, and some of these interactions may relate to pathogenesis of prion diseases (Fig. 2.10–4). Many of these potential interaction molecules are primarily or entirely cytoplasmic proteins, and their interactions with *PrP*^{C} are under study. As mentioned, *PrP*^{C} is a GPI-anchored protein and its entire polypeptide chain is external to the cytoplasmic membrane. Direct *PrP*^{C} interactions are probable with other receptors, secreted, or trans-membrane proteins. However, *PrP*^{C} contains a conserved hydrophobic sequence that can span the lipid bilayer in both directions. This results in two transmembrane

variants, ^{Ntm}PrP (N-transmembrane) and ^{Ctm}PrP (C-transmembrane). Generally, without predisposing mutations in *PrP,* these variants are present in very small quantities.

Genetics

There are three prion-related genes *PRNP, PRND,* and *SPRN* and the detailed study of these genes commenced after the cloning and sequencing of the human, hamster, and mouse prion proteins, first accomplished in 1986.

PRNP. The *PRNP* gene in humans is located on chromosome 20 and its product is PrP^c (OMIM 1). Table 2.10–2 shows chromosome locations of prion protein genes. Polymorphisms of *PRNP* have been extensively studied, and the polymorphisms in codon 129 may play a critical role in host susceptibility to prion diseases. For example, the presence of the V129 allele is not associated with vCJD, which raises the possibility that the presence of this allele is protective against vCJD superinfection. Some mutations in *PRNP,* including D178N, are associated with fCJD or FFI (depending on the residue *cis* at codon 129), and E200K is associated with fCJD. Moreover, P102L has been linked to GSS.

Prion Gene Complex, Downstream. *PRND* (prion gene complex, downstream) encodes the Doppel protein and is another gene related to the prion diseases. This is a paradigm known for the human genome that is due to gene tandem duplications. The following are terms used for the Doppel PrP: DPL, prion-like protein Doppel precursor; MGC41841, prion gene complex, downstream; PrPLP, PrP2 (doublet); and dJ1068H6.4, prion-like protein Doppel. The observation that $Prnp^{0/0}$ knockout mice exhibited late-onset ataxia led to the discovery of the *PRNP* Doppel. These mutants had an underlying neuropathology of cerebellar Purkinje cell loss. *PRNP* Doppel is located 16 kb downstream of the mouse *PRNP*. The significance of Doppel protein and prion pathogenesis is under further investigation.

Shadoo. The third prion-related gene encodes a short protein, known as Shadoo (*SPRN*), with similarities to the alanine-valine–rich central hydrophobic domain of PrP^C. The following are terms used for Shadoo (shadow precursor): FLJ41197, shadow of PrP homolog; and SHO, shadow of PrP. *SPRN* gene is not part of the PrP genomic complex and resides on the human chromosome 10. The significance of this protein and prion pathogenesis is under investigation.

Neuropathology

Neuropathological studies of patients with kuru have revealed that except for atrophy of the cerebellar vermis and flocculonodular lobe, the brain

macroscopically may appear normal. Microscopically, prominent neuropathological features manifest in cerebellum with loss of granule and Purkinje cells, fusiform swelling of the proximal portion of Purkinje cells' axons, and severe radial gliosis of Bergmann astrocytes.

Macroscopically, brains of patients with CJD may offer no specific diagnostic clues or alternatively, one may observe various levels of cerebral cortical, striatal, and cerebellar atrophy. The WHO proposed neuropathological criteria for CJD, including the presence of spongiform encephalopathy in cerebral and/or cerebellar cortex and/or subcortical gray matter, and/or encephalopathy with PrP immunoreactivity (plaque and/or diffuse synaptic and/or patchy/perivacuolar types). Microscopically, the prominent neuropathological features of CJD consist of spongiform degeneration of neurons and neuronal processes, significant neuronal loss, presence of severe astrocytosis, and formation of amyloid plaques (Fig. 2.10–5A,B). Spongiform degeneration that occurs in the context of CJD consists of development of abundant rounded vacuoles within the neuritic processes and synapses and can be observed in the cerebral cortex (invariably in all cases regardless of the clinical manifestations), the subiculum of hippocampus, caudate nucleus, thalamus, putamen, and the molecular layer of the cerebellum. Interestingly, the degree of spongiform degeneration and vacuolation in the same cortical section may vary from one area to another. The reactive astrocytosis that occurs in the course of CJD involves widespread presence of enlarged astrocytes in the cerebral cortex. In 10 percent of CJD cases, amyloid plaques that are positively immunoreactive with *PrP* antibodies and not immunoreactive with β-amyloid antibodies are present.

Neuropathology of vCJD demonstrates morphological and immunocytochemical features that set it apart from other prion diseases. The most prominent neuropathological feature of vCJD is the abundant presence of PrP^{Res} in the form of PAS-reactive, *PrP* amyloid plaques in the cerebrum and cerebellum (Fig. 2.10–5C,D). Many of the cerebral plaques are surrounded by a halo of spongiform vacuoles and form florid plaques. Other *PrP* plaques and amorphous pericellular and perivascular *PrP* deposits are prominent in the cerebellar molecular layer. The caudate nucleus and putamen are the focus of spongiform alterations, while the thalamus demonstrates severe neuronal loss and intense gliosis. These abnormalities are more pronounced in the posterior thalamic nuclei.

GSS disease is a TSE that is neuropathologically characterized by the presence of multicentric amyloid (*PrP*) plaques that are present in the molecular layer of the cerebellar cortex, cerebral cortex, and basal ganglia. The GSS-amyloid plaque characteristically contains a central larger mass surrounded by smaller satellite amyloid deposits. Similar to CJD, GSS is a conglomerate of at least six dominantly inherited syndromes, each of which is linked to a different mutation of the *PRNP* gene. These include GSS

(P102L), GSS (P105L), GSS (A117V), GSS (Y145Stop), GSS (F198S), and GSS (Q217R), and these dissimilar syndromes possess four general properties on neuropathological and molecular genetic grounds: (1) the clinical manifestations resulting from each mutation differs with the others; (2) each mutation is coupled with different subtypes of *PrP* plaques; (3) certain mutations are linked with considerable neurofibrillary tangle neuronal degeneration and generation of neuritic plaques; (4) amyloid plaques present in neuropathology of GSS syndromes consists of highly truncated *PrP* peptides (Figs. 2.10–6 and 2.10–7). Figure 2.10–7 demonstrates deposits of abnormal PrP in brain using the human vCJD (that originated in cows) in a monkey model for CJD.

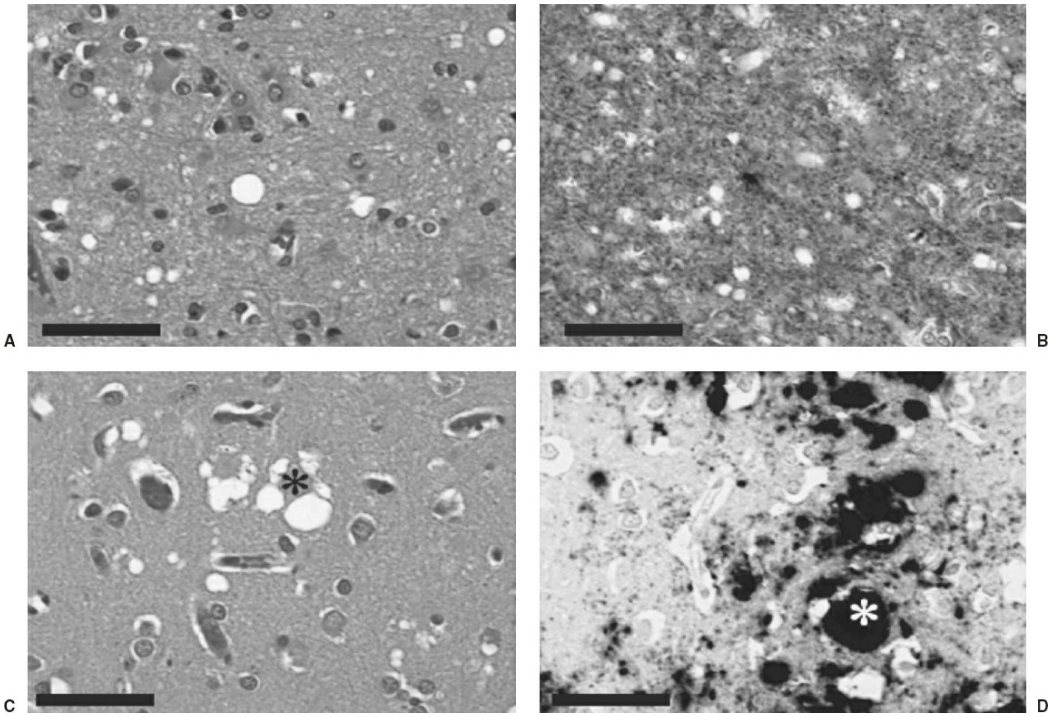

FIGURE 2.10–5. Histologic features of prion diseases. Central nervous system parenchyma of sporadic Creutzfeldt–Jakob disease (**A** and **B**) and variant Creutzfeldt–Jakob disease (**C** and **D**) showing astrogliosis and widespread spongiform changes. The protease-resistant form of host-derived prion protein depositions is synaptic (**A** and **B**) and is in the form of florid plaques (*asterisk,* **C** and **D**). **A,C:** Hematoxylin-eosin, original magnification ×400. **B,D:** Immunohistochemical stainings for prion protein, original magnification ×400. Scale bar = 50 µm. (From Glatzel M, Stoeck K, Seeger H, Lührs T, Aguzzi A. Human prion diseases: molecular and clinical aspects. *JAMA Neurol*. 2005;62(4):545. American Medical Association, with permission.)

The main neuropathological feature of FFI is thalamic atrophy. The affected thalamic nuclei include anterior ventral, dorsal medial, pulvinar, and centromedian. Although atrophy of cerebellar cortex is minimal, the inferior olives demonstrate atrophic changes. Some of the neuroscientists have argued that neuronal loss observed in neuropathology of FFI may be caused by apoptosis.

DIAGNOSIS AND CLINICAL FEATURES

Prion diseases, as a unique group of spongiform encephalopathies that affect human and animal hosts, have become the subject of intense media and popular interest. The significance of this group of invariably fatal disorders is further demonstrated by two Nobel prizes won by D. Carleton Gajdusek and Stanley B. Prusiner in 1976 and 1997, respectively.

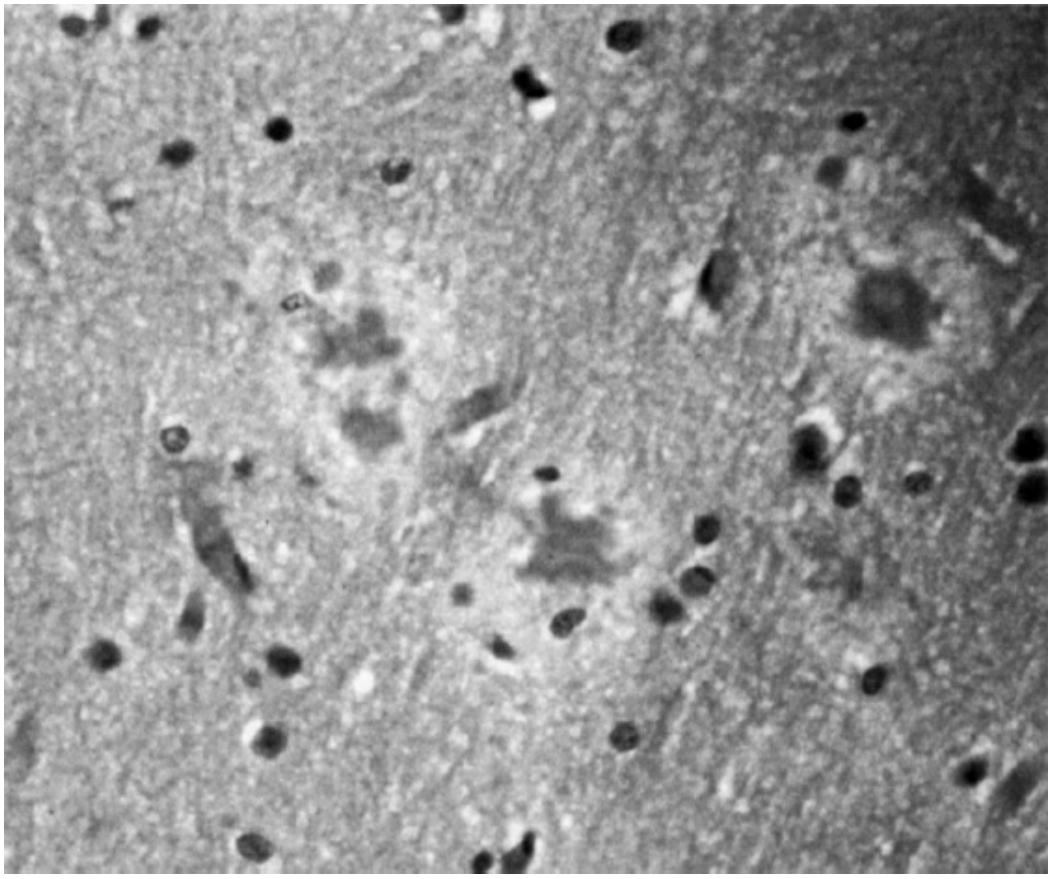

FIGURE 2.10–6. Polycentric prion protein (*PrP*) plaques in Gerstmann–Sträussler–Scheinker syndrome. This hematoxylin and eosin stain photomicrograph shows the densely eosinophilic cores of the PrP plaques seen in the cerebellum of patients with Gerstmann–Sträussler–Scheinker syndrome. Magnification ×250.

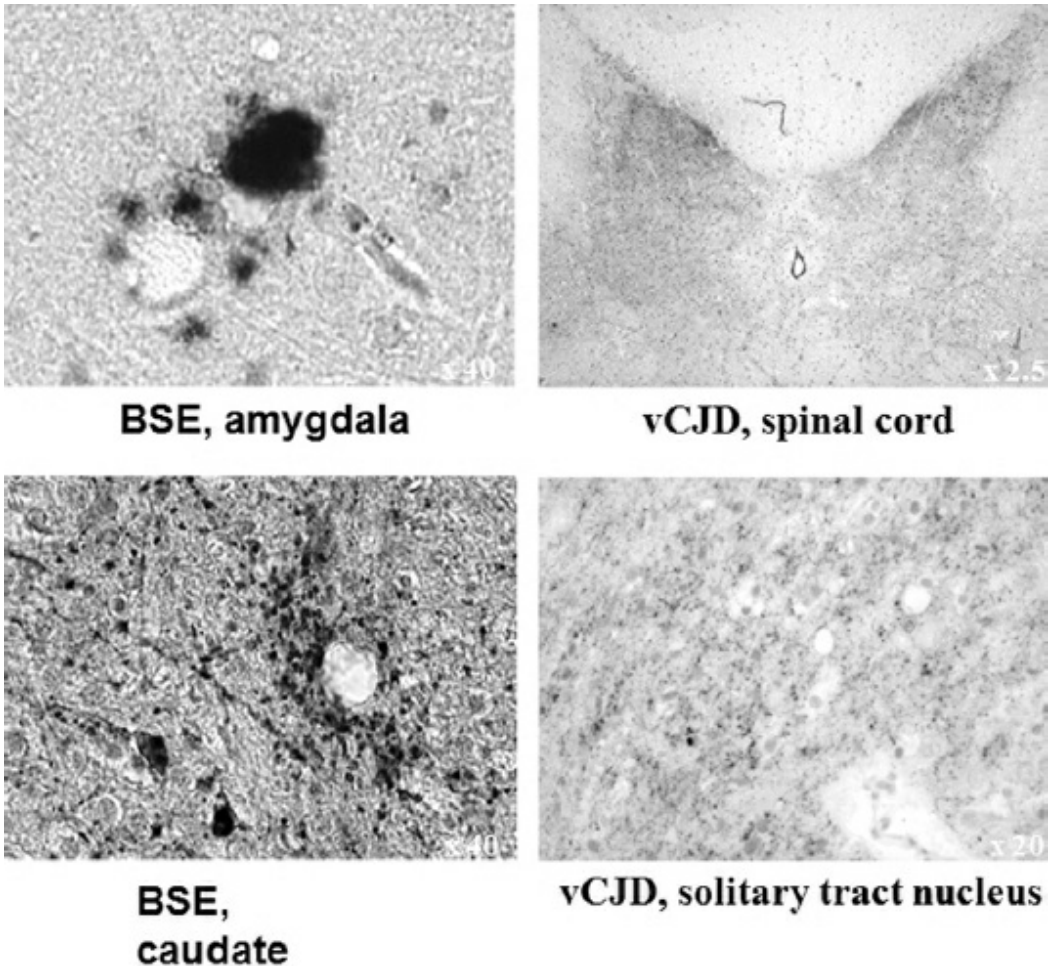

FIGURE 2.10–7. Immunohistochemistry of the deposition of abnormal prion protein in brain. Photomicrographs showing immunohistochemistry of the deposition of abnormal prion protein in the brains or spinal cord of cynomolgus macaques experimentally infected by bovine spongiform encephalopathy (BSE) or variant Creutzfeldt–Jakob disease (which is also the BSE agent accidentally transmitted to humans). (Contributed by Corinne Lasmezas, Scripps Research Institute, Jupiter, Florida.)

Prion diseases are a rare but fatal group of neurological disorders with unique pathophysiology. As a general rule, in any patient with a rapidly progressive dementia or other neurological deficits and in the absence of other reasonable diagnosis, the presence of prion diseases should be investigated.

Neurological Manifestations

A healthy 38-year-old right-handed female complained of weakness of the lower extremities, imbalance, and memory loss. Within a few weeks, this neurological syndrome was followed by increasing ataxia, startle myoclonus, and moderate dysarthria. Three months later she was admitted to the neurology service with profound lethargy, dementia, lip smacking, spasticity of the lower extremities, seizures, and hallucinations. Her electroencephalogram (EEG) demonstrated presence of periodic sharp wave complexes (Fig. 2.10–8), while her brain magnetic resonance imaging (MRI) was normal. She died within 6 months from the initiation of her neurological disease and autopsy confirmed the diagnosis of CJD.

A 63-year-old previously healthy woman was admitted with visual agnosia, visual hallucinations, and cognitive decline followed by ataxia and depression. CT scan of brain revealed mild brain atrophy and her EEG was interpreted as slow and disorganized. Within 2 months she developed tonic clinic seizures and her dementia rapidly deteriorated so that she developed akinetic mutism. MRI of brain, axial T2-weighted images, revealed the presence of hyperintense signals in the caudate nuclei and putamen (Fig. 2.10–9), and upon examination of her CSF 14–3–3 proteins were detected. Four month after the onset of her neurological symptoms she died. Postmortem examination of her brain confirmed sCJD.

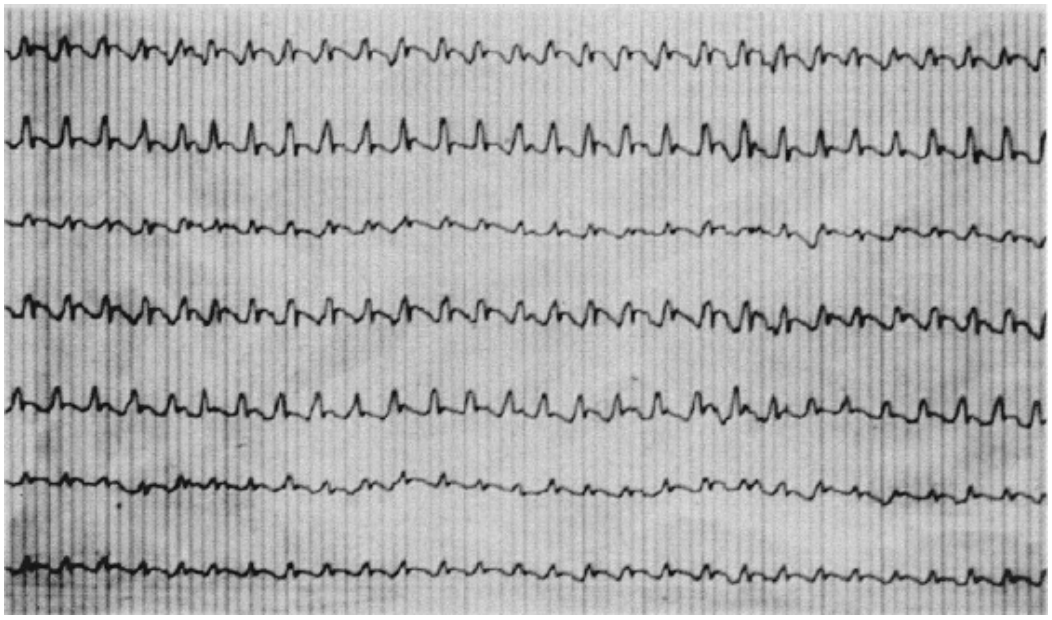

FIGURE 2.10–8. Typical triphasic periodic electroencephalogram of sporadic Creutzfeldt–Jakob disease. This shows the typical 1- to 2-Hz periodic electroencephalogram with triphasic waves that are seen diffusely in 70 percent of patients with sporadic Creutzfeldt–Jakob disease in the latter stages of illness. (Courtesy of Paul Brown.)

Sporadic CJD. sCJD is still the most common and most prominent human prion disease, which usually manifests insidiously and affects both genders equally. sCJD is a disease of late middle age that causes an invariably fatal disease. Usually, sCJD develops insidiously with nonspecific behavioral abnormalities including anxiety, asthenia, depression, loss of appetite, alteration of the sleep pattern, weight loss, fatigue, dizziness, and social regression. Neurologically, a large number of these patients develop worsening forgetfulness, progressive decline of higher cortical functions such as reasoning, abstract thinking, calculation, and judgment. Almost one third of these patients initially present with purely neurological deficits, mainly cerebellar gait ataxia. A minority of patients present with a combination of cognitive impairment and focal neurological deficits. With further progression of disease process, patients rapidly become demented, aphasic or apraxic, and demonstrate choreiform-athetoid movements, myoclonic jerks, and pyramidal and extrapyramidal signs (including parkinsonism). Initially, most patients do

not manifest myoclonus; however, with disease progression a majority of patients develop myoclonus. Startle myoclonus is a prominent feature of CJD, which is precipitated by certain stimuli such as loud noises or touch. Other symptoms that appear with disease progression include paratonic rigidity, primitive reflexes, cortical blindness, and oculomotor disturbances. Other visual symptoms include diplopia, blurred vision, and visual agnosia. Terminal patients develop akinetic mutism, collapse into coma, and die of infections or thromboembolic complications. Certain clinical features such as epileptic seizures (only in 10 percent of patients), sensory deficits, and lower motor neuron features are uncommon among these patients. In addition, the presence of certain signs or symptoms, such as early onset seizures during the clinical course, acute onset motor deficits, cranial nerves palsy (except for visual loss or diplopia), and significant ataxia without concomitant cognitive decline, raise serious questions about the accuracy of the diagnosis of CJD.

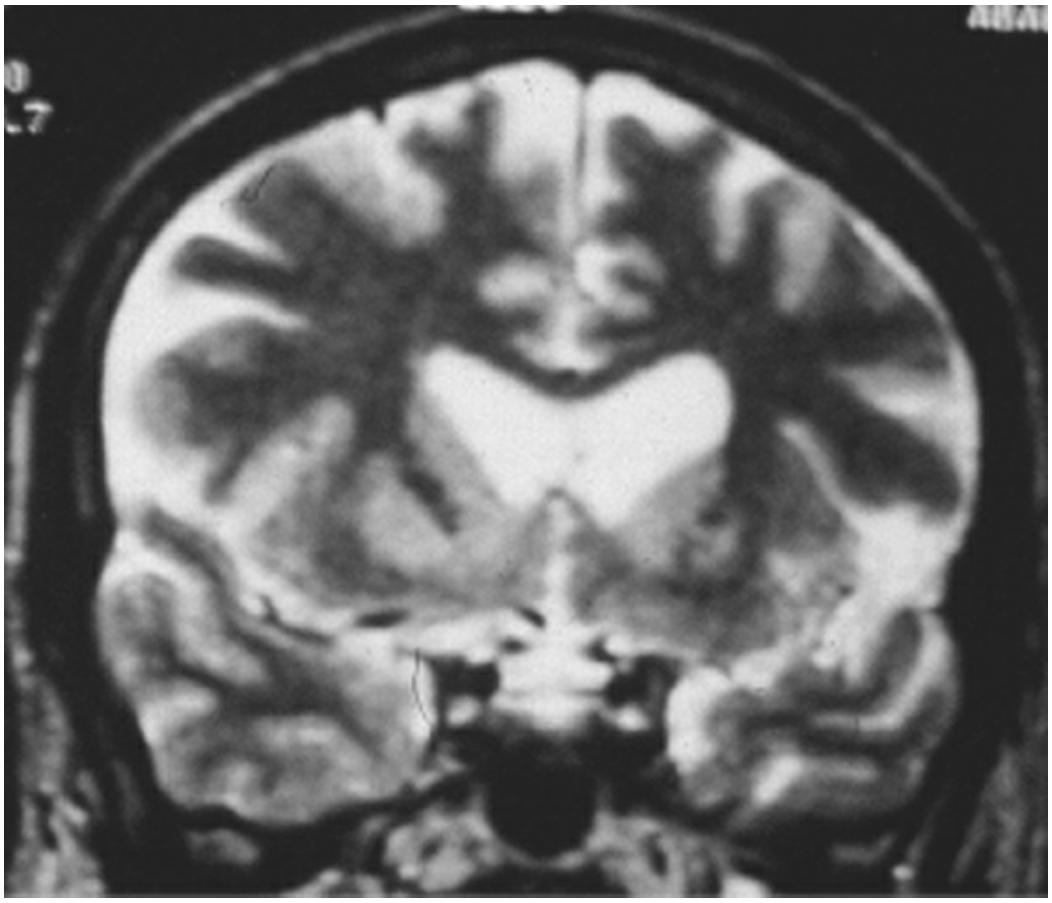

FIGURE 2.10–9. Magnetic resonance imaging coronal T2-weighted image demonstrating hyperintense signals in the basal ganglia of a male patient with sporadic Creutzfeldt–Jakob disease with rapidly progressive dementia and myoclonic jerks. (Courtesy of Dr. Eduardo Gonzalez-Toledo.)

The two rare clinical variants of CJD include Heidenhain and Brownell-Oppenheimer. In the Heidenhain variant of CJD the neuropathological

changes mainly affect the occipital cortex and result in cortical blindness, optic hallucinations, and agnosia. Brownell-Oppenheimer variant manifests with a pure cerebellar syndrome secondary to widespread neuropathological changes of the cerebellum. WHO diagnostic criteria for CJD are presented in Table 2.10–5.

Variant CJD. In 1995 the vCJD form of human CJD emerged from the United Kingdom. vCJD, as a distinct variation, differs substantially from the more typical sCJD because of the following characteristics: Disease clinical picture, development in younger individuals as compared to CJD, initial dominance of psychiatric manifestations, underlying neuropathology, the absence of any mutations in the *PRNP* gene and almost exclusive occurrence in those with methionine homozygosity at codon 129, and its linkage to BSE. The exact mechanism of transmission to humans remain unknown; however, it is hypothesized that consumption of meat or meat products that were contaminated with BSE infectious agent in the late 1980s caused BSE migration of BSE prion from other species to humans, generating the vCJD epidemic. The potential causal link between vCJD and BSE has attracted public interest. In addition, to date it remains unknown why the clinical and neuropathological abnormalities of vCJD tend to affect younger adults.

Table 2.10–5.
World Health Organization Diagnostic Criteria for Creutzfeldt–Jakob Disease (CJD)

1. CJD clinical diagnosis
   Criteria for probable sCJD: The clinical diagnosis of CJD is currently based on the combination of progressive dementia, myoclonus, and multifocal neurological dysfunction, associated with a characteristic periodic EEG. However, new variant CJD, most growth hormone-related iCJD, and up to 40% of sporadic cases are not noted to have the characteristic EEG characteristics.
   Proposed diagnostic criteria for probable CJD:
   a. Progressive dementia
      and
   b. At least two out of the following four clinical features:
      i. Myoclonus
      ii. Visual or cerebellar disturbance
      iii. Pyramidal/extrapyramidal dysfunction
      iv. Akinetic mutism
      and
2. A typical EEG during an illness of any duration
   And/or
3. A positive 14–3–3 cerebral spinal fluid assay and a clinical duration to death less than 2 years
4. Routine investigations should not suggest an alternative diagnosis

sCJD, sporadic Creutzfeldt-Jakob disease; iCJD, Iatrogenic Creutzfeldt–Jakob disease; EEG, electroencephalogram.

Adapted from World Health Organization: Human transmissible spongiform encephalopathies. *Wkly*

Epidemiol Rec. 1988;73:361.

Table 2.10–6.
Current World Health Organization Diagnostic Criteria for Variant Creutzfeldt–Jakob Disease (vCJD)

Diagnosis	Criteria
I	A. Progressive psychiatric disorder B. Duration greater than 6 months C. Routine workup does not suggest an alternative diagnosis D. No history of iatrogenic exposure E. E. No evidence of familial prion disease
II	A. Early psychiatric symptoms B. Persistent painful sensory symptoms C. Ataxia D. Myoclonus, chorea, or dystonia E. Dementia
III	A. Absence of typical periodic slow wave on EEG B. Bilateral pulvinar high signal on MRI
IV	A. Positive tonsillar biopsy
Definite diagnosis of vCJD requires the presence of IA and positive neuropathologic findings (spongiform changes and extensive accumulation of PrP^C with florid plaques throughout cerebrum and cerebellum); probable vCJD requires I and 4 of 5 of II and IIA and IIIB, or I and IVA; and possible diagnosis consists of I and 4 of 5 of II and IIIA.	

EEG, electroencephalogram; MRI, magnetic resonance imaging, PrP^C, prion protein cell.

From Will RG, Zeidler M, Stewart GE, et al. Diagnosis of new variant Creutzfeldt-Jakob disease. *Ann Neurol*. 2000;47(5):575.

In contrast to sCJD, vCJD can be transmitted through blood transfusion and presents with a relatively slow clinical course. The average age of onset of vCJD is 29 years as compared to the sCJD average age at onset of 65 years. Psychiatric symptoms such as depression, anxiety, agitation, delusions, and hallucinations are typically the initial manifestations, and within 6 months they are followed by more characteristic neurological features of CJD such as gait ataxia, dementia, dystonia, chorea, myoclonus, upgaze paresis, and sensory abnormalities. In patients with vCJD nerve conduction study is usually normal; however, somatosensory-evoked responses may demonstrate minor abnormalities, indicating central involvement of pain pathways or thalamic origin for the pain. Similar to what is observed in sCJD, the end-stage patients with vCJD develop profound dementia along with akinetic mutism. Patients die after a median duration of 13 months. WHO diagnostic criteria for vCJD are presented in Table 2.10–6.

Iatrogenic CJD. Accidental transmission of CJD to humans has been reported and well documented with corneal transplantation, implantation of contaminated EEG electrode, implantation of dura matter grafts, and

following use of contaminated human growth hormone (HGH) preparations derived from human pituitaries. In addition, one case of CJD following repair of an eardrum perforation with a pericardium graft has been reported. The disease can also be transmitted to humans by using inadequately sterilized neurosurgical instruments. Interestingly, in cases of central transmission (use of inadequately sterilized neurosurgical instruments or implanted EEG electrodes) patients mainly manifest with cognitive decline, while in peripheral cases such as transmission of CJD via injection of contaminated HGH or gonadotropins, progressive ataxia followed by dementia predominate the clinical picture. Majority of these patients develop myoclonic jerks.

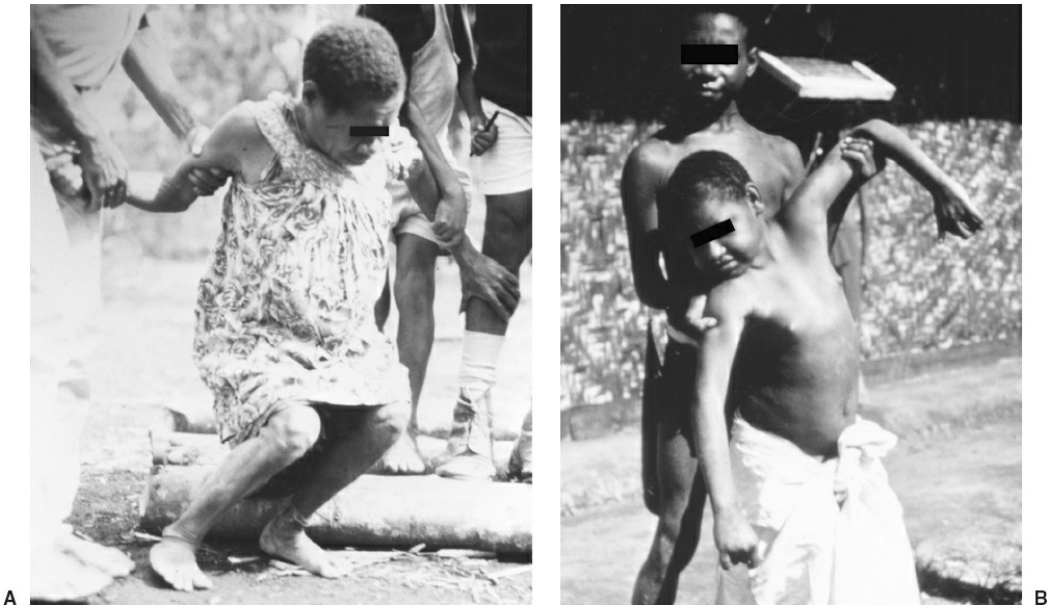

FIGURE 2.10–10. Two kuru patients. These figures show a young woman (**A**) and a child (**B**) in the late ambulatory phases of kuru. (Courtesy of Kimbra Kenney, M.D.)

Kuru. The word *kuru* originates from the Fore word for shiver and refers to a progressive cerebellar syndrome that occurs among the highlanders of Papua New Guinea. Kuru affects individuals between the ages 5 and 60 years, with equal ratio among male and female children, but prominent excess in female adults. Kuru is transmitted by ritual cannibalism. Women and children, the majority of the victims of kuru, participate in the cannibalistic feasts and consume the brain and intestines of the dead relative due to a social order within the tribe (Fig. 2.10–10). It is unclear when the first case of kuru occurred in this population; however, it is believed that the first case began as sCJD and then consumption of this infected deceased individual by other tribe members eventually led to recycling of the prions within this isolated population and occurrence of kuru at an epidemic proportion. By the late 1950s, cannibalism ceased and the frequency of kuru dropped significantly. If kuru has continued with an

even low incidence, then it is not purely due to cannibalism but may have a genetic component as well. In fact, that could be an alternative situation. One postulates that a mutant develops in this highly homogeneous and isolated set of tribes in New Guinea. Their degree of isolation from each other is supported by the very large number of unrelated languages spoken there until they were discovered. Once the mutant prion gene was transmitted, then cannibalism that may have existed prior to that or occurred as a result then propagated the disease and its further spread. In that case, one would expect a continued low level of kuru to linger after cannibalism ceased. However, secret cannibalism may still be practiced and could contribute to a very low incidence of kuru. Interestingly, prior to cessation of cannibalism each potential patient was exposed to prions many times, which in turn made it difficult to determine precisely the incubation period of kuru. The kuru disease process commences with vague prodrome characterized by malaise, arthralgia, and headache. Neurologically, kuru usually manifests with gait instability followed by balance problems and frequent falls until the individual can no longer walk independently or sit without support. Gait ataxia is accompanied by dysmetria, dysarthria, incoordination of the upper extremities, various movement disorders such as clonus, chorea, and athetosis, emotional incontinence with inappropriate laughter and convergent strabismus. Interestingly, weakness and rigidity are absent. Kuru patients' cognition is initially intact and usually in advanced stages of disease dementia manifests. As the disease advances, patients become apathetic and withdrawn. At the end, patients become completely incapacitated and dependent on another for daily life activities. Usually, within 1 year of disease onset patients die of aspiration pneumonia, sepsis from decubitus ulcers, or inanition.

Gerstmann–Sträussler–Scheinker Disease. GSS was initially described by Josef Gerstmann and colleagues in 1928 and 1936 in the members of a large Austrian family with multiple affected generations and autosomal dominant inheritance. However, GSS was categorized as TSE in 1981 when its transmissibility was discovered. Affected patients present with clinical manifestations of cerebellar involvement such as ataxia and dysarthria and later dementia manifests. In some of the reported families parkinsonism predominates, while others demonstrate cortical blindness, deafness, and gaze palsies. A codon 102 mutation of *PRNP,* which results in a substitution of a proline for leucine in the *PrP* molecule, has been linked to the ataxic form of GSS. Recently, Jones et al. reported a case of GSS associated with strongly seropositive VGKC-complex antibodies, which emphasize the point that these antibodies should be searched for in patients with GSS.

Fatal Familial Insomnia. FFI is a rare disease, which was initially reported by Medori and colleagues in 1992. FFI is the third most common

genetic prion disease with worldwide occurrence and with analogous clinical manifestations also in different genetic settings. This dominantly inherited prion disease manifests with sleep–wake cycle disturbances and insomnia followed by hallucinations and eventually patients plunge into coma. Initially, most patients complain of disturbances of vigilance, such as being unable to go to sleep, attention, visuomotor function, or having uninvigorating sleep and demonstrate personality changes such as apathy. Insomnia is an early manifestation of FFI, which in certain patients may progress to almost complete inability to sleep. The insomnia of FFI is associated with major and lasting disorganization of sleep cycles. Loss of ability to sleep is more obvious and more progressive in the early disease course in subjects who are methionine homozygous at codon 129 than in those who are methionine and valine heterozygous. Some patients also develop visual fatigue with diplopia and sympathetic activation. Other autonomic disturbances including pyrexia, hyperhidrosis, tachycardia, hypertension, and cardiac arrhythmia develop. Major motor abnormalities consist of ataxia, spontaneous and evoked myoclonus, and pyramidal signs. In contrast to sCJD where dementia rapidly develops, in patients with FFI dementia may or may not develop. Other abnormalities include loss of circadian rhythm for secretion of melatonin, prolactin, and growth hormone as well as decrease in secretion of adrenocorticotropine and increase in secretion of cortisol.

FFI rarely affects the thalamus without involvement of other regions of the forebrain, and the underlying neuropathology extends throughout much of the thalamus and cerebral cortex. Neuropsychological assessment of the original FFI families demonstrated increasing disturbances of attention and vigilance and impairment of working memory and temporal ordering of events. Patients' intelligence was preserved at least until vigilance levels permitted meaningful neuropsychology testing; however, frontal lobe functions were impaired.

A 67-year-old right-handed man developed depression, visual agnosia, and visual hallucinations. In 3 months, his neurological syndrome deteriorated and he developed progressive dementia, myoclonus, and ataxia. CT scan of brain showed only brain atrophy and EEG did not reveal any significant abnormalities. Within the next 2 months, his dementia progressed and he became bed-bound and eventually became unresponsive. Examination of CSF demonstrated presence of protein 14–3–3, and within 7 months from the onset of his first neurological symptoms he passed away. Neuropathological examination of brain during autopsy confirmed sCJD.

A previously healthy 29-year-old right-handed female developed brief episodes of forgetfulness and lapse of attention. She was initially diagnosed with major depression and was treated with antidepressants. Her neurological condition deteriorated with progressive further loss of communicating and social skills. Within the next 4 months she became severely ataxic and globally demented. The rapidly progressive dementia led to akinetic

mute condition followed by coma and death. EEG revealed rhythmic triphasic sharp wave activity more obvious on the right frontal and temporal areas on a slow background. MRI of brain on axial T2-weighted images revealed bilateral hyperintense signal in caudate nuclei and putamen (Fig. 2.10–11). During postmortem examination, typical neuropathological findings of sCJD were detected.

Psychiatric Manifestations

Psychiatric disorders commonly occur in patients with sCJD and generally manifest during the early phase of this progressive neurodegenerative disorder. Christopher A. Wall and colleagues conducted a retrospective review of 126 patients with sCJD who were assessed at the Mayo Clinic from 1976 to 2001. The investigators reviewed the clinical data for the presence of psychiatric disorders such as depression, anxiety, psychosis, behavior dyscontrol, sleep disturbances, and neurological abnormalities during the course of the disease. Based on their analysis of the obtained data, 80 percent of patients suffered from psychiatric symptoms within the first 100 days of their illness with 26 percent having them at presentation. The most frequently reported psychiatric symptoms were sleep disturbances, psychotic symptoms, and depression. They concluded that psychiatric symptoms are common in sCJD and occur early prior to formal diagnosis.

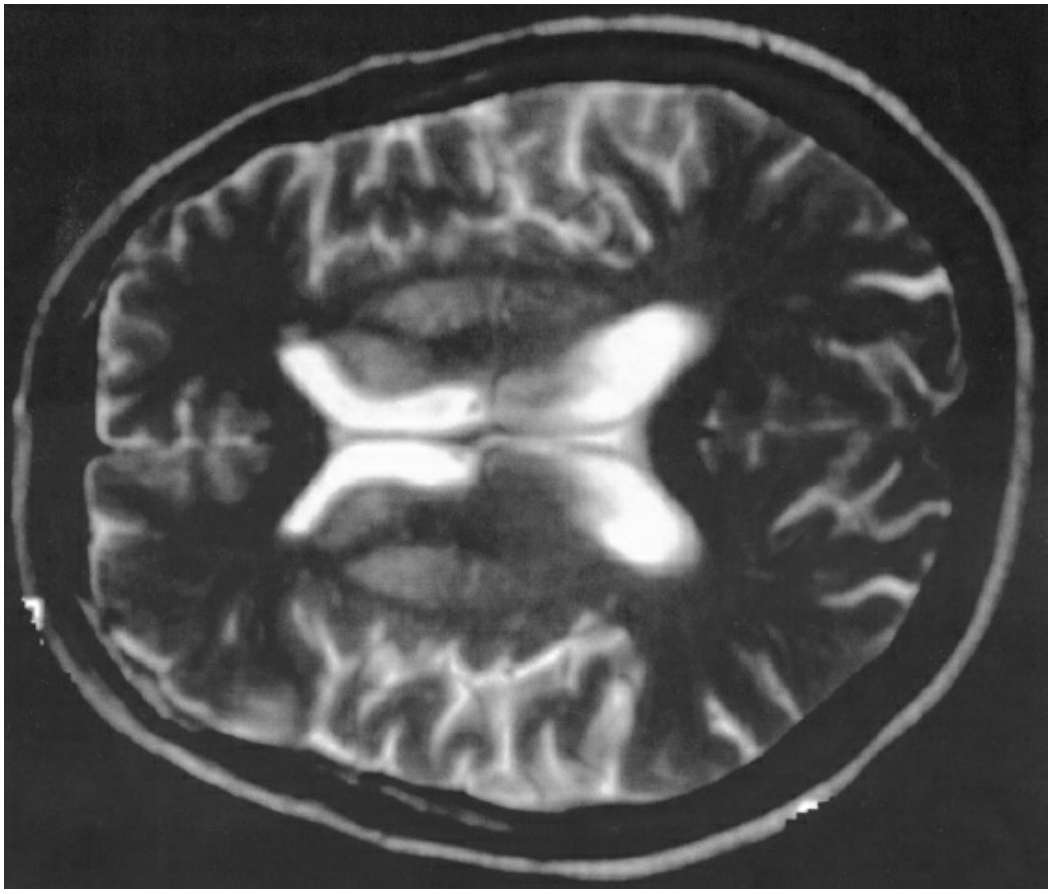

FIGURE 2.10–11. T2-weighted magnetic resonance image of a patient with sporadic

Creutzfeldt–Jakob disease. There is a hyperintense signal in the caudate nuclei and putamen. (Courtesy of Donald Collie.)

vCJD, which is less common than sCJD, affects younger individuals and has distinct clinical and neuropathological features, with more neuropsychiatric manifestations compared to sCJD (Table 2.10–7).

Martin Zeidler and colleagues identified and reviewed the first 14 cases of vCJD for psychiatric features. All patients had psychiatric symptoms during the early phase of their disease process, and the majority were diagnosed with depression or depression secondary to an organic disease. Two patients had symptoms suggestive of psychosis, while the majority manifested transient delusions and auditory or visual hallucinations. The authors concluded that psychiatric symptoms, as consistent early components of vCJD, occur in a majority of these patients. However, they do not manifest any characteristic features that can differentiate them from other common psychiatric disorders. Only the occurrence of associated persistent sensory symptoms may raise the possibility of vCJD. Nonetheless, the evaluation of any patient with subtle neurobehavioral signs and symptoms should include the systematic approach described below.

NEUROPSYCHIATRIC WORKUP: THE STRUCTURED PSYCHOLOGICAL AND NEUROPSYCHOLOGICAL HISTORY

A comprehensive cognitive history is essential for initiating the assessment and directing which tests should be used in the evaluation battery. In combination with the clinical pattern of an evolving dementia, the history can be key in the differential diagnosis of the etiologies of behavioral and cognitive dysfunction in patients with CJD with psychiatric complaints. Sometimes, the interview identifies the basis of the problem when neuropsychological testing is not possible or available. Table 2.10–8 lists the essentials of the interview for formulating this history. Answers to these questions can help associate patient complaints cognitive disorders and dementia.

Table 2.10–7.
Comparison of Clinical, Neuroradiological, and Pathological Features of and Presents sCJD versus vCJD

	sCJD	vCJD
Median age of onset	65 years (range 15–94)	26 years (range 12–74)
Median duration	4 months (range 1–74)	13 months (range 6–39)
Typical presentation	Progressive dementia, ataxia, myoclonus	Psychiatric and behavioral symptoms, dysesthesias
Cerebellar signs (% of patients)	40	100
Periodic EEG complexes (% of patients)	>90	0
CSF 14–3–3 protein (% of patients)	99	33
Neuropathologic abnormalities	Rare PrP plaques	Diffuse PrP plaques
Presence of the infective agent in the lymphoid tissue	Not readily detected	Readily detected
Presence of hyperintense signal in the caudate and putamen on diffusion-weighted and FLAIR sequences of MRI	Frequently present	Frequently absent
Pulvinar sign on brain MRI	Not reported	Frequently present
Increased glycoform ratio on immunoblot analysis of protease resistant prion protein	Not reported	Marked accumulation of protease resistant prion protein

sCJD, sporadic Creutzfeldt-Jakob disease; vCJD, variant Creutzfeldt-Jakob disease; EEG, electroencephalogram; MRI, magnetic resonance imaging; CSF, cerebrospinal fluid; *PrP*, prion protein; FLAIR, fluid-attenuated inversion recovery.

Data from Henry C, Knight R. Clinical features of variant Creutzfeldt-Jakob disease. *Rev Med Virol*. 2002;12:143; Irani DN, Johnson RT. Diagnosis and prevention of bovine spongiform encephalopathy and variant Creutzfeldt-Jakob disease. *Annu Rev Med*. 2003;54:305; Belay ED, Schonberger LB. Variant Creutzfeldt-Jakob disease and bovine spongiform encephalopathy. *Clin Lab Med*. 2002;22(4):849.

The Mental Status Examination

This introductory cognitive history taking should include an appropriate mental status examination. A standard examination, such as the Mini-Mental State Examination (MMSE), is frequently done because it is well known, and many practitioners have developed a sense of what its score means in terms of overall cognitive functioning. This test was developed to help identify the cortical type of dementia associated with Alzheimer disease (AD). It has a concentration of items associated with language and orientation and has a visuospatial task. As such, it may miss the types of memory, speed of information processing, and attention/concentration problems often associated with early manifestations of CJD. Another screening task, the High Sensitivity Cognitive Screen (HSCS), may be useful in screening for simple presence or absence of cognitive dysfunction. The creators of this measure report high correlations between this test and the overall result of neuropsychological testing. They note that the HSCS is also correlated with EEG results in medical psychiatric inpatients and with functional status. Results on this measure may then establish the eligibility of the patient for more in-depth neuropsychological assessment.

Neuropsychological Assessment

The most prevalent component of cognitive impairment related to CJD includes general cognitive decline. Early problems with abstraction, attention and concentration, learning and memory, and psychomotor speed progress to more serious difficulties with these functions, as well as impaired executive functions, nonverbal problem solving, and visuospatial integration and construction.

As with the language-heavy MMSE, neuropsychological tests for the assessment of other forms of dementia include tasks that gauge such cortical functions as complex language-associated functions (such as aphasia and apraxia), higher level cognitive functions of verbal and nonverbal abstract reasoning and problem solving, and perceptual functioning. These tasks are still necessary when one suspects or has information that the patient is experiencing dysfunction related to focal disturbances in the CNS. These can be caused by varying conditions such as infection, tumor, stroke, vasculitis, multiple sclerosis (MS), HIV, and other etiologies. Other tasks often used in comprehensive batteries for the evaluation of memory, attention, and concentration, and psychomotor speed are more useful for detecting the often-subtle impairments of the early stages of cognitive impairment as may occur in vCJD. Section 7.4 shows a neuropsychological battery recommended for dementing illness. These tests taken together yield an appreciable number of scores that individually may be difficult to interpret in isolation. Neuropsychologists have often tried to summarize or digest multiple scores into indices that

correspond to diagnostic levels (e.g., the Halstead–Reitan Impairment Index). With regard to one of the most generally used instruments, the Rey 15-Item Test, several investigators have shown that genuine conditions may yield false-positive results, and clinical judgment must ultimately determine the validity of this task's performance. Another method, the forced two-choice selection task, in which correct selections below the 50 percent level (actually, from the research, a much higher cutoff score; 69) may imply an attempt to deceptively manipulate the task, appears to be less problematic. These tasks have not been validated in the CJD population, and, unless the patient is obviously severely demented or delirious, there can be high confidence that results reflect the patient's actual effort.

Psychological Functioning

Psychiatric conditions have received much attention as they arise often in vCJD neurobehavioral disorders. The initial manifestations of emotional lability, anxiety, depression, aggression, and psychosis along with the emotional distress caused by a change in one's behavior and all the influences it has on a person, and possible pre-existing psychiatric disorders and/or neurological syndromes caused by other medical circumstances may necessitate a psychiatric consultation. A number of screenings and more comprehensive measures have been used but are not validated in CJD.

Table 2.10–8.
Cognitive History

Name
Age and birthday
Handedness
First language at home
Educational background
Best subjects, grades
Worst subjects
Occupational background
How long
Medical history
Childhood diseases or injuries
Head injuries with loss of consciousness
Strokes
High fevers
Toxin exposure
Major illness, injuries, or surgeries
Medicines: Prescription, nonprescription
Duration of changes in behavior or functioning
Current problem: Change in thinking functions: How long, or over what period of time
Any change in ability to concentrate

Any periods of confusion or mental “fuzziness”
When talking with people, or on the phone, watching TV or a movie, reading
Any problem with following the train of thought
Any difficulties with handwriting
Any word-finding problems; difficulties with slurring or stammering
Any slowing of thinking or understanding, trouble with mental arithmetic such as making change or balancing checkbook
Wear glasses
Any blurring vision, double vision, or flashing lights in eyes
Any change in understanding what is seen; do things look right in their relation to each other
Overlook things when right in front of you
Hearing any unusual sounds; see unusual things; have any strange feelings
Any changes in any other senses
Decreased hearing, ringing, or buzzing sounds
Change in smell or taste
Any numbness, “pins or needles,” loss of feeling, tingling, or burning feelings
Any severe pain
Memory
Any areas of memory that are better or worse
Memory for recent information
Information from way back in life
Any difference in memory for *situations* versus rote facts and figures
Kinds of things most easily forgotten: Names, addresses, directions, reading
How long can things be remembered, more notes written than used to
Any lapses noted
Any getting lost or forgetting where one is
Any new difficulties with thinking through problems or solving them, decisions making, staying organized—on job, at home
How is sleep: Any trouble getting to sleep; night versus daytime; any awakenings from which one cannot immediately return to sleep
Any inability to move any parts of the body
Muscle weakness, twitching, spasms, trouble walking, coordination problems, tremors or shakiness, problems with dropping things, feeling like moving more slowly, difficulty using tools or household utensils, getting dressed, telling right from left
Headaches or dizziness, instances thought to be seizures (staring off into space for a long time, uncontrollable movements, periods where one seemed “lose” time, incontinence)
Changes in mood, feelings, ideas
Mood swings, loss of patience or change in temper, increase in irritability, change in amount of worry, sense of panic

Although not formally researched in CJD, screening instruments such as the Beck Depression Inventory and the Symptom Checklist-90 (or its brief form, the Brief Symptom Inventory) have appeared frequently in the literature with medically ill patients and are simple to administer and score. These instruments and the more comprehensive Minnesota Multiphasic Personality Inventory 2 (MMPI-2) have been used to assess psychological difficulties in medically ill patients. A common problem and point of debate is the overlap of somatic or vegetative symptoms (e.g.,

fatigue) of depression and the systemic effects of any disease and/or its treatments. A newer, multifactor screening instrument, the HADS, attempts to avoid this problem by focusing on the psychological distress elements. For psychological evaluations beyond the screening level, the MMPI-2 and a structured interview would provide information required for a full psychological diagnosis.

Clinical Assessment

Psychiatrists will best diagnose CJD by recognizing its characteristic symptoms. The approach described above will assist only to exclude a number of other illnesses that mimic or overlap with CJD. Clinically, one would likely make the diagnosis of mild neurocognitive disorder (Table 2.10–9) or fully developed dementia (see Section 10.3).

Unfortunately, there are several issues with the category for mild neurocognitive disorder in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM). For example, the basis for the diagnosis is mostly focused on the presence of neuropsychological impairment instead of clinical findings. Also, there is no listing of motor signs or symptoms and thus is not very specific for prion, progressive supranuclear palsy, Parkinson disease (PD), HIV, or other etiologies.

Although early diagnosis may be made by clinical psychiatrists, it is the progressive nature of the neurological dysfunction that leads to neurological consultation and a definitive assessment and diagnosis made. In one survey in the United Kingdom, only 30 percent of psychiatrists were aware of the diagnostic criteria and the surveillance project for CJD. Although in all the cases of vCJD psychiatric symptomatology is the rule, it is nearly impossible to make the diagnosis in this early phase of the illness as there is no specific psychiatric phenotype to warn clinicians. Thus, development and progression of neurological aspects of the disease and either molecular or neuropathological examination are necessary to properly identify all CJD cases.

Table 2.10–9.
DSM-5 Diagnostic Criteria for Mild Neurocognitive Disorder

A. Evidence of modest cognitive decline from a previous level of performance in one or more cognitive domains (complex attention, executive function, learning and memory, language, perceptual motor, or social cognition) based on:
   1. Concern of the individual, a knowledgeable informant, or the clinician that there has been a mild decline in cognitive function; and
   2. A modest impairment in cognitive performance, preferably documented by standardized neuropsychological testing or, in its absence, another quantified clinical assessment.
B. The cognitive deficits do not interfere with capacity for independence in everyday activities (i.e., complex instrumental activities of daily living such as paying bills or managing medications are preserved, but greater effort, compensatory strategies, or accommodation may be required).
C. The cognitive deficits do not occur exclusively in the context of a delirium.

D. The cognitive deficits are not better explained by another mental disorder (e.g., major depressive disorder, schizophrenia)

From American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. DSM-5 5th ed. Text rev. Washington, DC: American Psychiatric Association; 2013, with permission.

Neurodiagnostic Workup

Laboratory diagnosis of CJD rests on abnormalities found in one or more of these tests: Examination of CSF, EEG, or brain MRI. Table 2.10–10 provides a summary of the initial diagnostic workup for CJD.

Routine examination of the CSF is usually within normal limits, although occasionally slightly elevated protein levels or minimal pleocytosis (less than 10 cells/mL) may be detected. However, CSF protein level is generally less than 100 mg/dL. A class of 14–3–3 proteinase inhibitor proteins is currently being measured within CSF as surrogate markers for neuronal injury that occurs in the course of CJD. Despite the fact that elevated CSF proteins 14–3–3 lend support to a diagnosis of CJD, the elevated levels may occur in other pathological conditions such as recent stroke, subarachnoid hemorrhage, hypoxic brain damage, glioblastoma, postictal, inflammation, corticobasal degeneration, or paraneoplastic syndromes. However, where CJD is suspected, the detection of protein 14–4–3 improves the accuracy and confidence of diagnosis. Recently, a new method for PrP^{SC} assay in the CSF has been developed which is known as "protein misfolding cyclic amplification" (PMCA) and is being utilized as a diagnostic test. In certain aspects, PMCA is similar to polymerase chain reaction and works as an amplification procedure to multiply misfolded prions. It is presently used for diagnosis of prion diseases A modified version of this technique has been developed as real-time quaking-induced conversion (RT-QuIC), which is capable of separating patients with sCJD from AD with 80 percent sensitivity and no false positive cases. CSF from patients with sporadic as well as inherited CJD, GSS, and FFI have been examined utilizing RT-QuIC and the results demonstrate that RT-QuIC is as sensitive as measurement of protein 14–3–3 but more specific than the latter assay.

EEG of patients with CJD characteristically demonstrates presence of periodic bi- and triphasic sharp wave complexes (PSWC) in the context of slow background. These classic EEG abnormalities are present in up to 70 percent of patients with sCJD, but may take up to 3 months to appear. Other diseases that may imitate EEG abnormalities observed in CJD include AD, HIV-associated dementia, Lewy body disease, MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes), anoxic/ischemic encephalopathy, hypoglycemia, hypo- and hypernatremia, hepatic encephalopathy, hyperammonemia, and lithium toxicity.

Table 2.10–10.
Initial Diagnostic Tests for Creutzfeldt–Jakob Disease

Blood tests: Complete blood count, comprehensive metabolic panel, serum ammonia, RPR, sedimentation rate, C reactive protein, serum vitamin B_{12}, HIV test, Lyme test, PT/INR
Urine tests: Urine analysis, urine drug screen
Cerebrospinal fluid analysis: Cell count and cytology examination for malignant cells, presence of oligoclonal bands, IgG index, protein and glucose level, VDRL
Neuroimaging: MRI of brain with and without contrast including
FLAIR and DWI images
EEG
CT scan of body with and without contrast searching for malignancy
Brain biopsy

RPR, rapid plasma reagin; HIV, human immunodeficiency virus; PT/INR, prothrombin time/international normalized ratio; IgG, immunoglobulin G; VDRL, Venereal Disease Research Laboratory test; MRI, magnetic resonance imaging; FLAIR, fluid-attenuated inversion recovery; DWI, diffusion-weighted imaging; EEG, electroencephalogram; CT, computed tomography.

CT scan of brain remains normal in most patients; however, in some it may reveal cerebral atrophy with enlargement of ventricles and cisterns or cerebellar atrophy. Brain MRI usually demonstrates distinguishing hyperintense signals in the caudate and putamen with less involvement of cortex or thalamus. These abnormal hyperintense signals can be observed on T2-weighted, proton density, and FLAIR sequences.

Compared to sCJD, studies of CSF and EEG are less useful in the diagnosis of vCJD. CSF 14–3–3 proteins are less sensitive in vCJD and are positive in only half of the patients, and EEG traces do not demonstrate the typical PSWC that is seen in sCJD.

MRI of brain in patients with sCJD demonstrates presence of hyperintense signals in the caudate and lenticular nuclei on diffusion-weighted and FLAIR images (Fig. 2.10–12). These abnormal areas on brain MRI correlated with the areas with the most spongiform vacuolation and neuronal loss than gliosis. MRI of brain in patients with vCJD shows bilateral hyperintense signals in the pulvinar of the thalamus (the pulvinar sign) on T2-weighted and proton-density images (Fig. 2.10–12). Pulvinar hyperintense signals correlates with gliosis. In patients with FFI routing, neuroimaging does not yield any characteristic features. Brain imaging with fluorodeoxyglucose positron emission tomography ([^{18}F]-FDG PET) demonstrates remarkable thalamic hypometabolism, which is sometimes associated with diffuse hypometabolism of the basal ganglia, cerebral cortex, particularly the frontotemporal cingular regions, and the cerebellum. Bilateral thalamic with less pronounced cingular cortex hypometabolism is a typical aspect of FFI. Polysomnographic studies of these patients have shown absence of sleep spindles and δ-sleep along with prominent autonomic and motor hyperactivity (also known as agrypnia

excitata). In FFI patients stage 1 nonrapid eye movement (NREM) sleep is substantially preserved. In patients with FFI, an increase in 5-hydroxytryptamine synthesizing neurons within the median raphe system has been reported and examination of CSF of patients with FFI has revealed a four- to fivefold increase in the 5-hydroxyindoleacetic acid (5-HIAA) catabolites compared to normal controls.

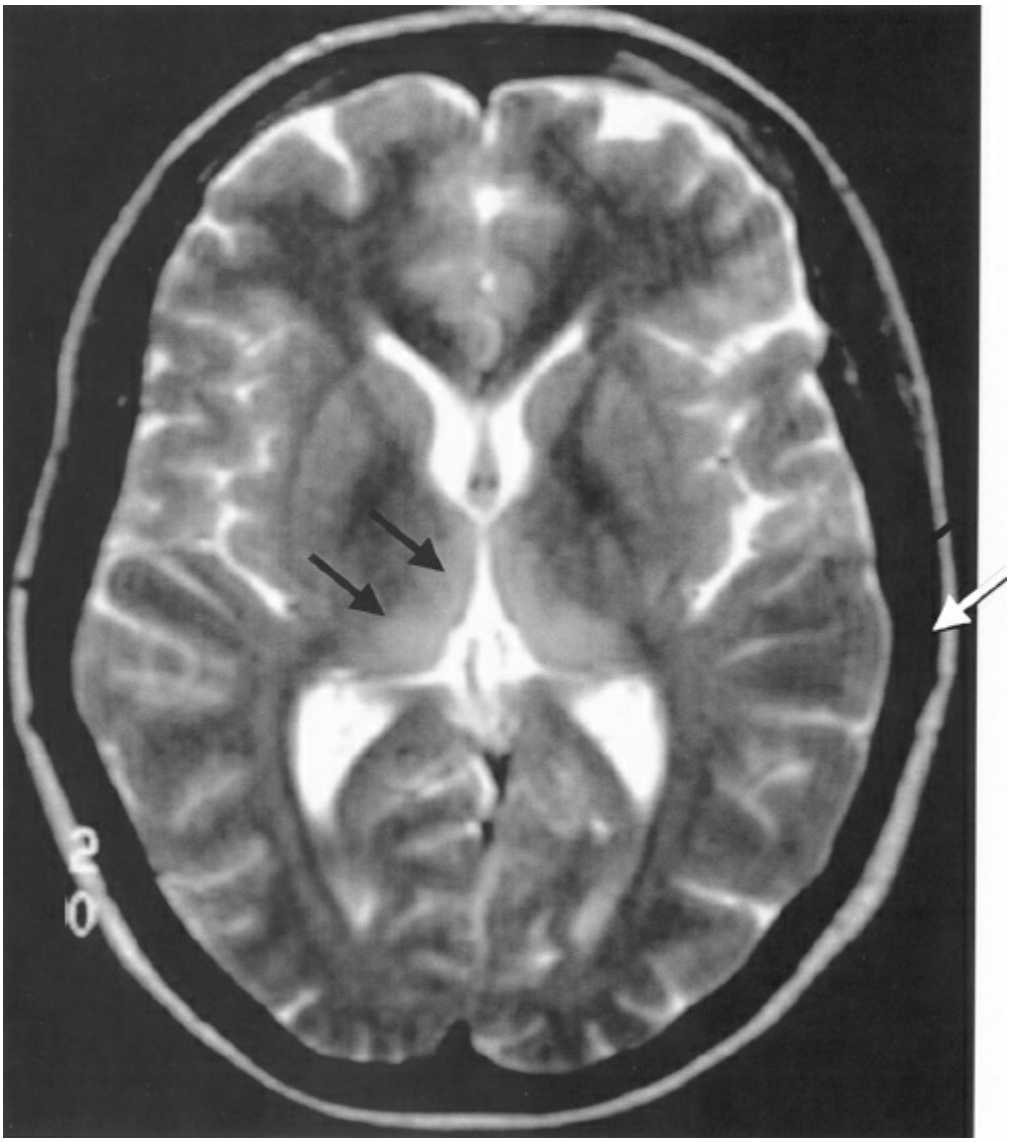

FIGURE 2.10–12. T2-weighted magnetic resonance image from a patient with variant Creutzfeldt–Jakob disease. There is a hyperintensity (*black arrows*) of the pulvinar nuclei extending into the more anterior thalamus, leading to the "hockey-stick" sign of variant Creutzfeldt–Jakob disease. (Courtesy of Donald Collie.)

Differential Diagnosis

Kuru does not pose a diagnostic challenge for clinicians since there are not too many other disorders that can cause relentlessly progressive cerebellar disease in Fore tribe members of New Guinea. Similarly, FFI is not a major diagnostic dilemma once it occurs in a member of an FFI pedigree.

However, in certain cases a positive family history is absent or insomnia develops only with disease worsening. In these cases, other prion-disease imitators such familial AD should be ruled out. However, this may not be the case for the other prion diseases, and in certain cases comprehensive diagnostic procedures should be performed.

Differential diagnosis of sCJD includes subacute encephalopathies secondary to intoxication with certain drug toxins such as lithium, mercury, or bismuth, and other diseases including CNS vasculitis, Hashimoto thyroiditis, paraneoplastic syndromes, infectious and granulomatous disorders such as neurosyphilis, sarcoidosis, HIV-1-associated disorders, CNS fungal infections, other neurodegenerative disorders such as AD, Lewy body dementia, PD with dementia, corticobasal degeneration, amyotrophic lateral sclerosis, and frontotemporal dementia. An extensive list of differential diagnosis of CJD is presented in Table 2.10–11.

COURSE AND PROGNOSIS

Despite dissimilar ages of onset, clinical manifestations, and durations of survival, CJD is a relentlessly progressive disease with almost daily compromise of neurological and psychiatric function that culminates into death. The mean duration of CJD is approximately 7 months, and disease duration of more than 2 years is rare. Similar to other prion diseases, vCJD is invariably fatal. Patients with FFI may experience worsening of their dysautonomic syndrome and eventually succumb to their fatal illness. To date, no information exists on the effect(s) of pregnancy on prion diseases.

TREATMENT

To date there is no known treatment of prion diseases, although pharmacotherapies are emerging. Despite the differences in their age of onset, clinical manifestations, and disease course and duration, prion diseases remain invariably fatal. One major stumbling block in development of effective therapy for this group of disorders is that significant PrP accumulation within CNS commences a long time before clinical manifestations develop, and currently available diagnostic procedures cannot reliably identify patients in the early stages of disease. A clinical trial of quinacrine (mepacrine hydrochloride) for human prion disease has been conducted and the results were published in 2013. The exact antiprion effect(s) of quinacrine remains unknown; however, these effects are possibly mediated by destabilization of cholesterol-rich detergent-resistant membrane (DRM) domains (also known as lipid rafts) of the PrP. Geschwind et al. conducted an NIH-funded, double-blinded, placebo-controlled, stratified trial of oral quinacrine (300 mg daily) to assess its impact on the survival rate of patients with sCJD. The investigators reported that their clinical trial failed to improve the survival

rate of patients with sCJD after 2 months.

Table 2.10–11.
Differential Diagnosis of Creutzfeldt–Jakob Disease

Toxic/metabolic disorders
- Endocrine disorders (thyroid, parathyroid, adrenal gland)
- Electrolyte imbalance (sodium, calcium, phosphate, magnesium)
- Deficiency of vitamin B_{12}, B_1, E, niacin, folate
- Uremic encephalopathy
- Wilson disease
- Porphyria
- Neuroacanthocytosis
- Metal intoxication (lithium, manganese, mercury)

Infections
- HIV-associated dementia
- Viral encephalitis
- Lyme disease
- Progressive multifocal leukoencephalopathy
- Fungal CNS infections
- Whipple disease
- Subacute sclerosing panencephalitis

Neurodegenerative disorders
- Alzheimer disease
- Huntington disease
- Pick disease
- Lewy body disorders
- Amyotrophic lateral sclerosis
- Progressive subcortical gliosis

Immune-mediated disorders
- Multiple sclerosis
- Hashimoto encephalopathy
- Systemic lupus erythematosus cerebritis
- Anti-GAD syndrome
- CNS vasculitis (primary or secondary)
- Sprue
- Acute disseminated encephalomyelitis
- Anti-VGKC syndrome

Neoplasms
- Metastatic encephalopathy
- Primary CNS lymphoma
- Glioblastoma cerebri
- Intravascular lymphoma
- Paraneoplastic syndromes
- Metastatic CNS tumors

Vascular neurological disorders
- Multi-infarct dementia
- Cerebral amyloid angiopathy

Thalamic or corpus callosum infarcts

HIV, human immunodeficiency virus; CNS, central nervous system.

Other presently investigational methods that may assist scientists to develop treatment of human prion disease include utilization of polyanionic, polycyclic drugs such as pentosan sulfate, treatments targeting the laminin receptor, use of neuroprotective agents, immunotherapy and utilization of osmotic pump delivery of antisense oligonucleotides into the ventricle.

At present, families of identified patients with CJD may inquire or request genetic testing especially in the context of familial prion disease. In light of the fact that there is no available treatment of CJD, it is debatable whether such testing is warranted. Although there are no specific treatment guidelines for genetic counseling in such situations, one can adapt recommendations that exist, inclusive of other neurodegenerative disorders (such as AD).

As with any other type of genetic testing and counseling, the risk–benefit ratio of testing should be carefully reviewed with the family and the medical team. If one proceeds with such testing, it should be done under the strictest confidentiality and privacy rules unless otherwise stipulated as with the U.S. Food and Drug Administration (FDA) ruling on blood donations and genetic analysis. Consistent with these developments, the following recommendations are provided for the clinician:

1. Testing conditions should be voluntary and follow local statutes where available. Pre- and posttest counseling should not be combined in the same session.
2. Either genetic counselors or psychiatrists should be competent in counseling patients about seeking genetic testing and should be well informed about the need for pre- and posttest counseling, provide information about the limits of confidentiality, should be aware of CJD surveillance procedures in their local jurisdictions, and should be aware of resources available to patients and their families.
3. A complete behavioral health assessment should be considered in all cases in order to fully address the individual's reaction to the results. This should include an assessment that the patient is indeed ready to be tested. Patients with high risks of untoward psychological reactions or destructive behaviors should be a special concern to anyone screening for CJD, and a referral to psychiatry should be initiated.
4. Genetic counseling involving pre- and posttest counseling of patients should include an assessment of capacity and discussion of the risks and benefits of testing, previous experience with genetic testing, the implications of a positive or negative result, the limits of confidentiality, strategies for reducing anxiety and depression, and help with the

availability of and referral to appropriate resources for further counseling and assistance.

5. Genetic testing should not be performed solely for the purpose of medical team and staff awareness.
6. All CJD testing must be done with appropriate informed consent. It is not sufficient simply to have a consent form signed; it must also be documented that the person is informed and understands the consequences of both a positive and negative result.
7. The confidentiality of information regarding genetic testing for CJD should be protected.
8. Sharing of CJD status should be in compliance with applicable state and federal statutes.
9. Patients should be made aware of program policies regarding documentation of CJD status in the medical record before initiating genetic testing.
10. Patients should be aware that when third parties pay for genetic testing, both positive and negative test results may be available to the Medical Information Bureau and can subsequently affect eligibility for future insurance.

Living in the era of molecular medicine forces all health professionals to confront difficult legal and ethical issues. Problem solving around issues of genetic testing requires an educated balancing of diverse interests and a thoughtful approach that rationally weighs the benefits and disadvantages of standard as well as new solutions. This field is rapidly changing and as such there are no easy answers in terms of whether or not asymptomatic family members should be tested for an untreatable disease.

PREVENTIVE MEASURES

Since prion diseases remain incurable, universal precautionary measures need to be applied during care and management of affected patients. Despite the fact that the infectious proteins causing CJD can be found in many organs, it appears less frequently in certain body fluids such as tears, saliva, sweat, urine, or feces. However, this infective prion can be detected in CSF and rarely in blood. In terms of infectivity, human tissues are divided into three groups: High-infectivity tissues: CNS tissues that attain a high titer of infectivity in the later stages of all TSEs, and certain tissues that are anatomically associated with the CNS; lower-infectivity tissues: Peripheral tissues that have been tested positive for infectivity and or PrP^{TSE} in at least one form of TSE; and tissues with no detectable infectivity: Tissues that have been examined for infectivity and/or PrP^{TSE} with negative results.

Health workers should wear gloves while handling these patients' biological specimens and should avoid any penetrating injuries from

contaminated sources that can potentially transmit the infected material. Any accidental exposure of intact skin to these infective proteins must be irrigated with fresh undiluted bleach or 1 N sodium hydroxide followed by washing and cleansing with copious amounts of water and soap. For medical and surgical procedures, disposable materials and instruments should preferably be used, and once they have been used they must be destroyed by incineration. In case certain instruments or devices cannot be discarded or destroyed, they should be sterilized based on proposed WHO decontamination guidelines. Autopsies and cremation of patients affected with prion diseases must be carried out based on the guidelines published by WHO.

Universal Precautions Regarding CJD Risk

It should also be noted that frequently the patient is unaware of having CJD or other infectious diseases when coming to the doctor's office or emergency room. Under these circumstances, universal precautions should be used. The psychiatry staff inpatient unit, trainees, and attending physicians are responsible for laboratory procedures that are part of the medical workup of the patients admitted to the unit, including the psychiatry emergency room. Moreover, this includes nursing personnel nurses who must provide parenteral injections to patients without knowing their CJD, HIV, or hepatitis C status. Under these conditions, universal precautions must be followed according to the Centers for Disease Control (CDC) (www.CDC.gov) as well as the *American Psychiatric Association Compendium of Practice Guidelines* (www.psych.org/AIDS/) and the WHO Infection Control Guidelines for TSE (www.who.inf.csr/resources). Pertinent information is presented in Table 2.10–12.

Table 2.10–12.
Universal Precaution Guidelines for Prion Disease

Situation	Response
CJD is considered in the differential diagnosis	Eye, nose, and mouth protection along with full-length fluid proof gowns should be used
Percutaneous exposure to blood or CSF	Washing with sodium hypochlorite (household bleach) followed by whatever care is available for infectious disease including for HIV, hepatitis C, or CJD
All clinical personnel	Vaccinated against hepatitis B and have annual checks for TB
Instruments	Disposable wherever possible
Sharps	Properly disposed of in designated sharps containers
Needle/syringes	Should be single-use and retractable

CJD, Creutzfeldt–Jakob disease; CSF, cerebrospinal fluid; HIV, human immunodeficiency virus; TB, tuberculosis.

From www.CDC.gov; www.psych.org/AIDS/; www.who.inf.csr/resources.

Biohazards: Prion Transmissibility and Infectiousness

Overall, the exercise of universal precautions related to biohazards is essential for handling and processing all specimens of human tissues and fluids. The examination of this topic in PubMed indicates that this is a topic discussed in the literature, due to ubiquitous biohazard risks. The dangers, risks, remedies, and countermeasures in the laboratory and clinical situation have been well developed through the use and institution of (biosafety level) BSL-3 and BSL-4 laboratories. This has become most apparent during the recent Ebola virus outbreaks.

There are several corrective laboratory measures in regard to human prion diseases. Thirty-five years ago it was already recognized that prions derived from spongiform encephalopathies including Kuru, CJD, transmissible mink encephalopathy, and scrapie posed biohazard risks for humans. Based upon that, laboratory design proceeded.

Tissues and fluids infected or even potentially infected (and that includes all human specimens) are treated and handled with care and universal precautions. These precautions are well known for tissues and fluids at risk of HIV and HCV infection. Universal precautions emphasize the risk of infectivity due to unanticipated biological agents. However, risks of infection by prions are an additional risk that is associated with any human specimens—derived from symptomatic or asymptomatic individuals. Similarly, animal specimens are treated with similar care. The goal is reduction in the potential exposure risk. In the United States, several accrediting agencies deal with such risks and include the NIH, CDC, and Occupational Safety and Health Administration (OSHA).

The central concept in the infectiousness of PrPs is that PrPs, which are in misfolded conformations that are self-replicative, can infect, seed, and propagate their strain characteristics in dissimilar host species as well as homologous host species. The seeding infecting PrP converts much of the host's cellular PrP (PrPC) into infectious PrP proteinase K resistant particles. This process is also termed PMCA.

Current investigations also implicate a process of "seeding" misfolded alpha-synuclein protein involvement in some aspects of PD. This possibility is related to the prion disease, CJD. In fact, it has further been proposed that there is a component of prion-related pathogenesis (gain of toxic function) in AD as well as in PD. It is suggested that these diseases should be included with TSEs or prion diseases in common with the well-known prion diseases including CJD and its variant, kuru, as well as with FFI and GSS syndrome. In addition, *amyotrophic leukospongiosis* pathogenesis may involve 18 to 20 kD prion-like globular structures.

The use of *dura mater* lyophilisates, corneal grafting, neurosurgery, and clinical treatment with pituitary-derived hormones has resulted in multiple

cases of iatrogenic CJD. Transmission infectiousness is associated with the organ used. The highest infectiousness is in the CNS tissue followed by lymph nodes and spleen. In experimental models, skeletal muscles, skin, and bone do not show infectiousness. It is concluded that in the hospital environment a wider more careful practice must be implemented concerning tissues and blood vis-à-vis CJD agent transmission.

Slaughter waste used as cattle feed results in the occurrence and spread of spongiform encephalopathy (BSE) associated with vCJD disease in humans in Europe and in the United States. It has been also detected in goats and sheep and in North America as CWD in captive elk and deer. In farm-reared mink, it was detected as transmissible mink encephalopathy.

As a post-script, it may be noted that the scientific attitude to potential infection by microbial life forms or nonviable particles, for that matter, has wide reaching implications. For example, there was, apparently, an absence of any sensible biocontainment protocols during the many years of the Apollo Space Program missions. There was also an absence of a functional particle containment protocol even though the lunar surface was composed of dust. There were findings of extensive human and mechanical contamination by lunar dust that has been well documented. Thus, it is not always obvious to the nonbiologically oriented scientist what types of exposure there may be, biological or mechanical. This is of relevance because of the unanticipated situations including prion diseases in which higher levels of biocontainment or mechanical containment are necessary though originally not perceived.

References

Aguzzi A. Prion diseases of humans and farm animals: epidemiology, genetics, and pathogenesis. *J Neurochem*. 2006;97:1726–1793.

Atarashi R, Satoh K, Sano K, et al. Ultrasensitive human prion detection in cerebrospinal fluid by real-time quaking-induced conversion. *Nat Med*. 2011;17(2):175–178.

Ayuso Blanco T, Urriza Mena J, Caballero Martínez C, Iriarte Franco J, Munoz R. [Fatal familiar insomnia: Clinical, neurophysiological and histopathological study of two cases]. *Neurologia*. 2006;21(8):414–420.

Collins SJ, Lawson VA, Masters CL. Transmissible spongiform encephalopathies. *Lancet*. 2004;363(9402):51–61. PMID: 14723996.

Cordery RJ, Almer K, Cipolotti L, et al. The neuropsychology of variant CJD: a comparative study with inherited and sporadic forms of prion disease. *J Neurol Neurosurg Psychiatry*. 2005;76:330.

Daus ML, Wagenführ K, Thomzig A, et al. Infrared microspectroscopy detects protein misfolding cyclic amplification (PMCA)-induced conformational alterations in hamster scrapie progeny seeds. *J Biol Chem*. 2013;288(49):35068–35080.

Derogatis LR. *Brief Symptom Inventory*. Minneapolis: Pearson Assessments; 1993.

Geschwind MD, Kuo AL, Wong KS, et al. Quinacrine treatment trial for sporadic Creutzfeldt-

Jakob disease. *Neurology*. 2013;81(23):2015–2023.

Grizzle WE, Fredenburgh J. Avoiding biohazards in medical, veterinary and research laboratories. *Biotech Histochem*. 2001;76(4):183–206.

Hall DA, Leehey MA, Filley CM, Steinbart E, Montine T. PRNP H187R mutation associated with neuropsychiatric disorders in childhood and dementia. *Neurology*. 2005;64:1304–1306.

Hilker R, Brotchie JM, Chapman J. Pros and cons of a prion-like pathogenesis in Parkinson's disease. *BMC Neurol*. 2011;11:74.

Jahrling PB, Keith L, St Claire M, et al. The NIAID Integrated Research Facility at Frederick, Maryland: a unique international resource to facilitate medical countermeasure development for BSL-4 pathogens. *Pathog Dis*. 2014;71(2):213–219.

Jones M, Odunsi S, du Plessis D, et al. Gerstmann-Straüssler-Scheinker disease: novel PRNP mutation and VGKC-complex antibodies. *Neurology*. 2014;82(23):2107–2111.

Korth C, Peters PJ. Emerging pharmacotherapies for Creutzfeldt-Jakob disease. *Arch Neurol*. 2006;63(4):497–501.

Lawson VA, Collins SJ, Masters CL, Hill AF. Prion protein glycosylation. *J Neurochem*. 2005;93(4):793–801.

Ligios C, Sigurdson CJ, Santucciu C, Carcassola G, Manco G. PrP^{Sc} in mammary glands of sheep affected by scrapie and mastitis. *Nat Med*. 2005;11(11):1137–1138.

McGuire LI, Peden AH, Orrú CD, et al. Real time quaking-induced conversion analysis of cerebrospinal fluid in sporadic Creutzfeldt-Jakob disease. *Ann Neurol*. 2012;72(2):278–285.

Nazor Friberg K, Hung G, Wancewicz E, et al. Intracerebral Infusion of Antisense Oligonucleotides Into Prion-infected Mice. *Mol Ther Nucleic Acids*. 2012;1:e9.

Panegyres PK, Armari E. Therapies for human prion diseases. *Am J Neurodegener Dis*. 2013;2(3):176–186.

Poleshchuk NN, Kapitulets SP, Kapitulets NN, Kvacheva EB, Eremin VF, Votiakov VP. [The infectiousness of 18- to 20-kd proteins isolated from the brain of people who have died from amyotrophic leukospongiosis]. *Biull Eksp Biol Med*. 1993;116(10):409–412.

Ross ED, Minton A, Wickner RB. Prion domains: sequences, structures and interactions. *Nat Cell Biol*. 2005;7(11):1039–1044.

Sano K, Satoh K, Atarashi R, et al. Early detection of abnormal prion protein in genetic human prion diseases now possible using real-time QUIC assay. *PLoS One*. 2013;8(1):e54915.

Toupet K, Compan V, Crozet C, et al. Effective gene therapy in a mouse model of prion diseases. *PLoS ONE*. 2008;3(7):e2773.

U.S. Food and Drug Administration, Centers for Biologics and Evaluation and Research. Guidance for Industry: Revised Preventive Measures to Reduce the Possible Risk of Transmission of Creutzfeldt-Jakob Disease (CJD) and Variant Creutzfeldt-Jakob Disease (vCJD) by Blood and Blood Products. 2002. http://www.fda.gov/cber/gdlns/cjdvcjd.htm.

Vandevelde M, Zurbriggen A, Fatzer R. Spongiform encephalopathies with special reference to bovine spongiform encephalopathy. *Schweiz Med Wochenschr*. 1992;122(23):887–892.

Wadsworth JD, Joiner S, Linehan JM, Cooper S, Powell C. Phenotypic heterogeneity in inherited prion disease (P102L) is associated with differential propagation of protease-resistant wild-type and mutant prion protein. *Brain*. 2006;129(Pt 6):1557–1569.

Wall CA, Rummans TA, Aksamit AJ, Krahn LE, Pankratz VS. Psychiatric manifestations of

Creutzfeldt-Jakob disease: a 25-year analysis. *J Neuropsychiatry Clin Neurosci*. 2005;17:489–495.

Walter ED, Chattopadhyay M, Millhauser GL. The affinity of copper binding to the prion protein octa-repeat domain: evidence for negative cooperativity. *Biochemistry*. 2006;45(43):13083–13092.

Westergard L, Christensen HM, Harris DA. The cellular prion protein (PrP(C)): Its physiological function and role in disease. *Biochim Biophys Acta*. 2007;1772(6):629–644.

Wilham JM, Orrú CD, Bessen RA, et al. Rapid end-point quantitation of prion seeding activity with sensitivity comparable to bioassays. *PLoS Pathog*. 2010;6(12):e1001217.

▲ 2.11 Neuropsychiatric Aspects of Headache

KATHLEEN RIES MERIKANGAS, PH.D., TARANNUM LATEEF, M.D., AND JAMES R. MERIKANGAS, M.D.

During the past decade, there has been growing attention to the enormous public health impact of migraine. In recognition of its high prevalence and burden, as well as the limited devotion of research resources to migraine, the World Health Organization has launched a global campaign called "Lifting the Burden" to reduce the burden of headache (www.who.int/mental_health/neurology/headache/en/index.html). There have also been advances in the treatment of both acute migraine attacks and preventative agents, and a growing interest in evaluation and treatment of the pervasive comorbidity of headache with both mental and physical disorders. Spurred by establishment of large international collaborative studies, there has also been progress in identifying the influence of genetic factors on migraine and other headaches. The goals of this chapter are to: (1) provide an overview of the diagnostic nomenclature for headache syndromes; (2) describe the magnitude, risk factors, and patterns of comorbidity of migraine and other headache subtypes; (3) summarize the clinical and laboratory examination necessary for making a diagnosis of headache and the neuropsychiatric manifestations of migraine and other headache syndromes; and (4) review current acute and prophylactic treatment strategies for the major headache syndromes.

DEFINITIONS

The International Classification of Headache Disorders-2 (ICHD-2) were introduced in 2004 to update the original set of criteria introduced in 1988. The ICHD-2 classification system was developed in order to provide specific operational criteria for the major headache syndromes and to facilitate international standardization of the diagnostic nomenclature of headache syndromes. Information obtained from a history, a physical and neurological examination, and appropriate laboratory investigations should be used to apply the ICHD-2. There are three basic subtypes of headaches or facial pain in the ICHD-2 system: primary headache syndromes (i.e., migraine without aura, migraine with aura, tension-type headache, and cluster headache); secondary headache syndromes include eight types of headaches secondary to other acute and chronic conditions; and cranial neuralgias and other causes of facial pain. The eight types of secondary causes of headache include: head or neck trauma (posttraumatic headache); cranial or cervical vascular disorder; nonvascular intracranial disorder; substance or its withdrawal; infarction; disorder of homeostasis; disorder of cranium, neck, eyes, ears, nose, sinus, teeth, mouth, or other

facial or cranial structures, and psychiatric disorders. The latter is a new category that has limited empirical basis. It is defined as a headache that occurs for the first time in close temporal association with the onset of a psychiatric disorder. This headache subtype is only considered as definite if it improves after successful treatment of the psychiatric disorder. This chapter addresses the primary headache types of migraine, tension-type, and cluster headaches, and a brief discussion of posttraumatic headache.

Migraine

Migraine is a disorder characterized by recurrent attacks or episodes of headache accompanied by other neurologic and gastrointestinal systems. Migraine presentation is multifaceted with symptoms emanating from multiple systems, including vascular, neurologic, gastrointestinal, endocrine, and visual. There is general agreement that a comprehensive neuropsychiatric evaluation is required for all patients presenting with headache complaints.

The ICHD-2 criteria for migraine with and without aura are presented in Tables 2.11–1 and 2.11–2. The core features of most definitions of migraine include recurrent headache that is often unilateral, gastrointestinal symptoms such as nausea or vomiting, and hyperesthesia manifested by photophobia or phonophobia. The headache generally has a pulsatile or throbbing quality, and the pain is exacerbated by routine physical activity involving movement of the head. The ICHD-2 criteria operationalize these features of headache to draw common thresholds and distinctions between migraine and other types of headache. Migraine was formerly divided into two major subtypes, common and classic, with the latter being distinguished by the presence of neurologic symptoms that precede the onset of the headache. The ICHD-2 no longer includes the common–classic distinction; instead migraine is subtyped according to the presence or absence of aura symptoms (reversible neurologic dysfunction).

Similar to recent controversy regarding the DSM-5 system, there are still several characteristics of headache syndromes that complicate the classification of headache. These include: the co-occurrence of multiple headache syndromes within individual persons; the tendency for headache characteristics to change across the life span; the effects of professional and self-treatment of headache that may obscure the manifestations of the underlying headache syndrome(s); and the lack of generalizability of treated samples from which the diagnostic criteria were derived. Specific areas of the classification system have also been identified as requiring additional clarification, notably: the specification of procedures for ensuring standardized methodology for the ascertainment of the diagnostic criteria; methods for assessing and coding multiple headache syndromes within individuals; and the development of standardized methods for discriminating between primary headache syndromes and those for which

the etiology is known (i.e., secondary headaches).

Table 2.11–1.
ICHD-2 Criteria for Migraine without Aura

A. At least five attacks fulfilling criteria B–D
B. Duration between 4 and 72 hours (untreated or unsuccessfully treated)
C. At least two of the following:
    1. Unilateral
    2. Pulsating pain
    3. Moderate to severe intensity
    4. Aggravation by or causing avoidance of routine physical activity (walking or climbing stairs)
D. During headache at least one of the following:
    1. Nausea and/or vomiting
    2. Photophobia and phonophobia
E. Not attributed to another disorder

Table 2.11–2.
ICHD-2 Criteria for Migraine with Aura

A. At least five attacks fulfilling criteria in B–D
B. Aura consisting of at least one of the following but no motor weakness:
    1. Fully reversible visual symptoms including positive features (e.g., flickering lights, spots or lines) and/or negative features (i.e., loss of vision)
    2. Fully reversible sensory symptoms including positive features (i.e., pins and needles) and/or negative features (i.e., numbness)
    3. Fully reversible dysphasic speech disturbance
C. At least two of the following:
    1. Homonymous visual symptoms and/or unilateral sensory symptoms
    2. At least one aura symptom develops gradually over 5 minutes and/or different aura symptoms occur in succession over 5 or more minutes
    3. Each symptom lasts 5 or more minutes and less than or equal to 60 minutes
D. Headache fulfilling criteria B–D for migraine without aura begins during the aura or follows within 60 minutes
E. Not attributed to another disorder

Tension-Type Headache

The definition of tension-type headache according to ICHD-2 criteria is presented in Table 2.11–3. Briefly, tension-type headache is characterized by episodes of stable bilateral pain lasting several days at a time. It is distinguished from migraine headache by its generally longer duration, the lack of pulsating quality of the pain, the lack of worsening with physical activity, and the absence of gastrointestinal concomitants. However, migraine and tension-type headache may often coexist, either simultaneously or alternating over time. It is no longer believed that tension-type headache results from muscle tension. Indeed, neck pain may result from head movement to reduce headache pain.

Cluster Headache

Cluster headache is a distinct syndrome characterized by frequent intense attacks (often several per day) over a 1- to 2-month period, separated by headache-free intervals for as long as 1 or 2 years. Although it is commonly grouped with migraine, current evidence including epidemiological data, treatment response, and clinical features suggests that cluster headache may comprise a distinct syndrome.

Cluster refers to a "clustering in time," with the headache bouts occurring every day to several times a day over a period of days to weeks, followed by a lengthy headache-free interval. Cluster headache is generally retro-orbital in location and is accompanied by autonomic changes such as lacrimation, rhinorrhea, erythema of the eye, and agitation. Males tend to suffer more from cluster headache than females. Patients with cluster headache do not retire to dark rooms and lie down to avoid the stimulation, but may in fact do quite the opposite, appearing almost manic in their agitation. The pain can be so intense that the sufferer may appear to be psychotic because of the screaming and thrashing that may be associated with the pain. Prior smoking and alcohol use have been associated with cluster headache, with alcohol often triggering the onset of the headache.

Chronic paroxysmal hemicrania is a type of cluster headache, specifically responsive to treatment with indomethacin and characterized by many daily focal attacks of pain lasting for short periods, generally about 15 or 20 minutes per attack.

Table 2.11–3.
ICHD-2 Criteria for Tension-Type Headache

Frequent episodic tension-type headache

A. At least 10 episodes occurring on greater than or equal to one but less than 15 days per month for at least 3 months (greater than or equal to 12 and less than 180 days per year) and fulfilling criteria for B–D
B. Headache lasting from 30 minutes to 7 days
C. Headache has at least two of the following characteristics:
    1. Bilateral location
    2. Pressing/tightening (nonpulsating) quality
    3. Mild or moderate intensity
    4. Not aggravated by routine physical activity such as walking or climbing stairs
D. Both of the following:
    1. No nausea or vomiting (anorexia may occur)
    2. No more than one of photophobia or phonophobia
E. Not attributed to another disorder

Chronic tension-type headache

A. Headache occurring on greater than or equal to 15 days per month on average for more than 3 months (greater than or equal to 180 days per year) and fulfilling criteria B–D
B. Headache lasts hours and may be continuous
C. Headache has at least two of the following characteristics:
    1. Bilateral location
    2. Pressing/tightening (nonpulsating) quality
    3. Mild or moderate intensity

4. Not aggravated by routine physical activity such as walking or climbing stairs

D. Both of the following:
   1. No more than one of photophobia, phonophobia or mild nausea
   2. Neither moderate or severe nausea nor vomiting

E. Not attributed to another disorder

Headache Attributable to Head or Neck Trauma

The key symptoms of ICHD-2 posttraumatic headache include a headache following head trauma accompanied by a loss of consciousness, posttraumatic amnesia, and abnormal laboratory tests. Posttraumatic headache is variable in symptom presentation, severity, and duration.

Although headache following a traumatic head injury has often been attributed to emotional factors, empirical evidence suggests that emotional factors are more likely to be a sequela rather than a cause of posttraumatic headache. Nevertheless, the pathogenesis of posttraumatic headache is unknown. The major hypotheses include cerebral edema, cortical spreading depression, innate vulnerability to cerebral vasospasm, and transient elevation of intracranial pressure. There is no direct relationship between the prevalence or chronicity of posttraumatic headache and several indicators of severity of head injury including duration of unconsciousness, posttraumatic amnesia, electroencephalographic abnormalities, presence of skull fracture, or the presence of blood in the cerebrospinal fluid. There appears to be an inverse relationship between the severity of the head injury and the development of posthead injury headache; posttraumatic headache is more common after injuries that do not result in skull fracture.

The onset of typical migraine attacks following acute head trauma occurs so frequently that it has been hypothesized that head trauma serves as a trigger for migraine in persons with underlying susceptibility to migraine or with a personal or family history of migraine. Moreover, relatives of posttraumatic migraine subjects have an increased prevalence of neurologic symptoms, suggesting a propensity to neurologic manifestations of migraine.

EPIDEMIOLOGY AND COURSE

Prevalence

Headache is very common in the general population, with about 70 percent of adults endorsing a history of headaches in a given year. Aggregate weighted average prevalence of definite migraine is 10.9 percent from international studies is very close to the 10 to 12 percent rates in pre-ICHD-2 studies. The aggregate prevalence of *probable migraine* defined by ICHD-2 averages 7 percent, with a range from 1.8 to 26.3 percent. The combined prevalence of probable and definite migraine is about 18 percent. By contrast, chronic migraine is quite rare, with an aggregate estimate of 0.5 percent. Recent reviews of the prevalence of tension-type headache

estimate the aggregate prevalence of tension-type headache of 32 percent.

Despite the large body of cross-sectional studies on the prevalence and correlates of migraine, there is a dearth of prospective research from community samples that provides information on the incidence, stability, and course of migraine in adults. Estimates of migraine incidence in the United States converge in demonstrating an earlier peak incidence of migraine, particularly migraine with aura in males in the mid-teens than in females in whom the peak incidence is in early twenties. However, despite the high prevalence rates, about half of those with migraine remit, about 35 percent continue to have intermittent headache, and only 20 percent continue to develop chronic migraine over time. In general, both the frequency and duration of migraine decrease at mid-life in both men and women, and the symptomatic manifestations may change substantially over time.

The severity of migraine ranges from mild to nearly total disability. Over 80 percent of those with migraine report some degree of disability. Recent community studies have underscored the enormous personal and social burden of migraine in terms of both direct and indirect costs. Although there is an increasing proportion of those with headaches seeking professional treatment, only approximately one-half of those individuals who suffer from debilitating migraine seek professional help.

Migraine in childhood is more likely to be associated with gastrointestinal complaints, particularly episodic bouts of stomach pain, vomiting, or diarrhea, and the duration is shorter than that commonly observed in adults. In women, migraine is strongly associated with reproductive system function, with increased incidence during puberty and the first trimester of pregnancy, and is associated with exogenous hormone use. After menopause, the frequency of migraine attacks generally decreases dramatically, unless estrogen replacement therapy is administered.

Aside from sex and age, a family history of migraine is one of the most potent and consistent risk factors for migraine. The results of twin studies implicate genetic factors underlying approximately one-third of the familial clustering of migraine, but the mode of inheritance is clearly complex. Despite an increasing number of candidate gene association studies of migraine, to date, no replicated linkage or associations between specific genes and migraine has emerged, except for hemiplegic migraine. Recent large-scale genome-wide association studies have begun to identify some potential susceptibility genes for migraine.

The course of migraine is highly variable. In general, both the frequency and duration of migraine decrease at mid-life in both men and women, and the symptomatic manifestations may change substantially over time. There are numerous precipitants of migraine attacks that have been consistently implicated as precipitants of acute headache attacks (including hormonal

changes, stress or its cessation, fasting fatigue, over-sleeping, particular foods and beverages, drug intake, chemical additives, bright light, weather changes, and exercise), but these agents/situations vary dramatically within and between individuals in prospective research.

The prevalence of tension-type headache is greater in women, but the sex difference is far less pronounced than that of migraine. Although tension-type headache is most common in young adults, there is a less steep decrement in prevalence with age.

By contrast, posttraumatic headache is quite rare in the general population (i.e., about 1 percent lifetime prevalence). However, it is likely that this is an underestimate because of the lack of systematic data on posttraumatic headache in population-based samples. Based on retrospective reports of those who suffer from a serious head injury, the prevalence of severe and chronic headache ranges from 28 to 62 percent. Children and young adults appear to be particularly susceptible to the development of headache after head trauma.

The results of prospective studies of the incidence of headache following severe head injury, usually defined as postconcussion headache, reveal that approximately 50 percent of each series of admissions continue to suffer from headache at the time of discharge from the index admission, with a gradual dissipation to 20 percent a year. Persistence of headache has been related to female sex, age over 45, the presence of dizziness, lack of skull fracture, intracranial hematoma, disorders of smell, hearing, or vision, depression, and impaired concentration.

The population prevalence of cluster headache is very low (less than 1 percent of the general population) and occurs nearly exclusively in males. The age at onset of cluster headache is somewhat later than that of migraine and tension-type headache; the first attack of cluster usually begins in the late twenties or thirties and may recur intermittently throughout life. Risk factors include smoking and heavy alcohol use. There is some family study and twin research demonstrating the role of genetic factors underlying the etiology of cluster headache.

Comorbidity

Although comorbidity of migraine and numerous other conditions has long been reported in clinical samples, etiologic inferences are precluded by the role of comorbid disorders in treatment-seeking for primary disorders. There is now a strong evidence from population-based surveys regarding associations between migraine and musculoskeletal conditions, cardiovascular disease, particularly ischemic stroke, asthma, allergies and other immune and inflammatory disorders, and epilepsy. Many of these studies have examined the specificity of the associations between comorbid disorders with migraine by including comparison groups of nonmigraine headache and other pain conditions. For example, analyses of the National

Comorbidity Survey Replication in the United States revealed that hypertension was the only chronic physical condition that was specifically associated with migraine rather than headache in general. Likewise, the association between migraine and obesity has been shown to be attributable to headache in general rather than migraine whereas the association with migraine and cardiovascular disease is specifically associated with migraine.

Psychiatric comorbidity in migraine has also been well-established in population samples of adults and children. Migraine is most strongly associated with anxiety and mood disorders, particularly phobic states and major depression. The presence of a preexisting physical or medical disorder may also elevate the risk of migraine. Prospective research on children demonstrates that migraine is associated with an increased risk for the development of depression rather than the converse. However, anxiety disorders, particularly phobias, are associated with an increased risk of migraine. This link may actually be a manifestation of underlying autonomic reactivity that may represent a common underlying diathesis for migraine. The elevated rates of infantile colic in children in treatment for migraine would be consistent with this explanation. As such, comorbid disorders may reflect underlying etiologic rather than environmental triggers of migraine. Several studies have now confirmed that there is a syndromic association between migraine, depression, and anxiety, with anxiety preceding the onset of migraine followed by the subsequent development of mood disorders. Studies that investigated specificity of these associations have concluded that migraine was specifically associated with phobias and hypomania/bipolar spectrum conditions, whereas the association with depression was found for headaches in general. Prospective studies of children with headache have indicated an increased risk for both headaches and somatic and psychiatric symptoms in adulthood. Research into the possible mechanisms underlying these associations, however, remains limited. Studies examining the order of onset and cross-transmission of migraine and psychiatric disorders in controlled family studies have been unable to clearly distinguish between causal and common aetiological models of association.

Comorbidity has a major impact on migraine sufferers' quality of life and disability, and is associated with a poorer course and outcome of migraine. As such, comorbidity may comprise an index of severity of migraine. For example, simultaneous investigation of somatic and psychiatric comorbidity in headaches and migraine has demonstrated greater health care utilization and negative health perception among those with comorbidity. In fact, about 65 percent of migraine-associated disability can be attributed to comorbid physical and mental disorders.

Impact and Health-Seeking of Migraine

Recent community studies of migraine have highlighted the enormous personal and social burden of migraine in terms of both direct and indirect costs. Over 80 percent of those with migraine report some degree of disability. In one US population survey, one-third of those with migraine had three or more attacks per month, and more than half reported severe impairment requiring bed rest. Children with migraine have been consistently shown to have more school absences, decreased academic performance, social stigma, and impaired ability to establish and maintain peer relationships. In fact, the quality of life in children with migraine is impaired to a degree similar to that in children with arthritis or cancer.

Severe headaches and migraine not only have substantial impact on the affected individual but they also have major economic impact due to medical expenses and employer costs. In the United States in 2012, the estimated annual US health care costs for migraine associated with: outpatient visits were $3.2 billion, ER visits were $700 million, and inpatient hospitalizations were $375 million. The bulk of the expenses attributable to migraine derive from its high population prevalence and indirect costs due to occupational disability rather than direct health care costs which are lower for migraine than for the other neurological conditions. During the past few years, there has been a concerted effort to raise awareness of the enormous public health impact of migraine. In recognition of its high prevalence and burden, as well as the limited devotion of research resources to migraine, the World Health Organization recently launched a global campaign to reduce the burden of headache (Lifting the Burden 2). This review documents the initial success in terms of the rapid growth of information on the magnitude of migraine in areas of the world that had been previously under-represented.

Despite the high magnitude of disability associated with migraine, only approximately one-half of those individuals who suffer from debilitating migraine seek professional help. The gap in treatment is remarkably similar across the world despite variation in health systems across the world. Of those who do seek treatment, many do not continue in treatment. Only a minority of those with migraine in the general population ever seek treatment with clinicians with expertise in headache. As expected, those who seek professional treatment are characterized by greater severity, longer duration, more disability, and more comorbidity. In light of the overwhelming evidence regarding the tremendous burden of migraine, leaders in the headache field have called for increased awareness of the availability of preventive efforts. Operational criteria for prevention have been developed based on headache frequency and attack-related disability, yet few of those with migraine have received preventive interventions. For example, only 25 percent of those with migraine in the general population of the United States actually seek professional treatment and receive appropriate medications.

ETIOLOGY

Although the etiology of the major types of headaches is still unknown, recent advances in brain imaging and neuroscience have advanced the understanding of the pathogenesis of migraine. Most theories of migraine focus on activation of the trigeminovascular system and its central projections. Functional imaging studies have shown that migraine is associated with brainstem activation, particularly the dorsolateral pons. These studies also provide evidence that the associated neurologic symptoms are the human homolog of cortical spreading depression. Diffusion tensor imaging studies demonstrate abnormalities in primary and modulatory components of the central nociceptive pathways such as the periaqueductal gray, an important central modulatory component of the pain pathway, the somatosensory cortex, and in the occipital lobe in subjects with migraine with aura. What remains unclear however is whether these findings represent causative abnormalities or whether they have occurred as the consequence of frequent repetitive attacks in migraine sufferers.

Psychophysiological studies have also contributed to a better understanding of mechanisms underlying migraine pathogenesis. In particular, interictal abnormalities of evoked and event-related potentials have been reported for different cortical areas in migraineurs such as the sensory cortices. Several studies suggest that the excitability of neurons in the visual cortex plays a fundamental role in the brain's susceptibility to migraine attacks. Other reports include increased amplitudes of visual evoked potentials, reduced habituation of cortical evoked responses, and an increased contingent negative variation in migraine patients.

Current theories of the etiology of cluster headache posit that hypothalamic and central pain control regions and trigger a cascade of events in the brain stem comprising afferent pain and efferent parasympathetic pathways. PET has shown vasodilatation of the major basal arteries during the acute pain attack in cluster headache, representing the first convincing demonstration of activation of neuronal vasodilator mechanisms in humans.

Far less is known about the etiology of tension-type headache. It is however clear that tension-type headache is a misnomer, since there is no evidence that muscle tension is the underlying cause of this headache subtype.

DIFFERENTIAL DIAGNOSIS AND CLINICAL EVALUATION

A very skillful workup is essential because headache is such a nonspecific complaint with an enormous number of etiologies ranging from the trivial to the acutely life-threatening (Table 2.11–4). A thorough examination should include a description of the type and location of pain, timing,

precipitants, prodromal events, and associated symptoms. Patients should be encouraged to keep a headache diary. The following factors are important to determine in order to define whether or not the headache is migrainous: (1) onset; (2) frequency; (3) location; (4) duration; (5) quality; (6) severity; (7) precipitants; (8) precursors; (9) triggers; (10) phenomena that worsen or relieve the pain; (11) warning signs; (12) prodromal events; (13) specific symptoms including visual changes, gastrointestinal symptoms, or neurologic symptoms; (14) sensitivity to light, noise, sounds, or touch; mood changes; and (15) cognitive changes. In addition, it is important to obtain a detailed family history, description of course, and a history of previous evaluation and treatment. Differential diagnosis of headache is based on a neurological examination to rule out pathognomonic signs that might indicate other brain disorders.

Migraine is more than a headache. There are a variety of manifestations of migraine that may mimic a number of neurological or psychiatric disorders including epilepsy, psychosis, and "conversion." Visual and auditory hallucinations may occur, especially in children. Migraine may have autonomic manifestations suggesting cardiac disease, irritable bowel syndrome, or even acute abdominal emergency. Migraine may be associated with irritability, mood swings, and in some cases, impulsive temper outbursts. In a sense, migraine may be considered as lying on a continuum between the rapid neurophysiological changes of epilepsy and the less rapid state changes of bipolar disorder. Basilar artery migraine, defined by a particular vascular distribution, may produce stupor and coma or paralysis, blindness, ataxia, dysarthria, or perceptual abnormalities. The extent to which such manifestations may be attributed to comorbid disorders has not been established.

In addition to a history and physical examination, laboratory studies are critical. Even if the results are negative and do not uncover a metabolic, endocrine, or autoimmune etiology, this information may serve as a baseline for subsequent drug therapy. Headache onset over age 50 is particularly ominous and requires careful investigation. Temporal arteritis, or giant cell arteritis, must be ruled out, as it has the potential to cause stroke or blindness if not urgently treated. A sedimentation rate and serum alkaline phosphatase are useful screening tests followed by corticosteroids and antiviral medication. Application of the ICHD-2 requires that all of the potential causes of headache should be considered. The diagnosis of headache requires the exclusion of other conditions, including structural lesion, vascular malformation, viral or bacterial meningitis, or encephalitis, intracranial abscess or hemorrhage, cerebral contusion, metabolic disorders (urea cycle disorders, aminoacidopathies, mitochondrial disorders), pseudotumor cerebri, vasculitis, brain tumors, sinusitis, or ocular disorders, any of which may be concurrent rather than causal. One of the most important findings of the past decade is the converging

evidence that there is an increased risk of ischemic (but not hemorrhagic) stroke among young women with migraine. This finding supports the importance of reduction of other stroke risk factors including oral contraceptive use, hypertension, and smoking among young women with migraine.

Table 2.11–4.
Headache Symptoms Indicating Further Diagnostic Workup

- ✓ First headache
- ✓ Worst headache
- ✓ Gradual worsening over days or weeks
- ✓ Vomiting prior to headache onset
- ✓ Abnormal neurologic examination
- ✓ Ongoing systemic illness
- ✓ Onset after age 50
- ✓ Accompanied by fever
- ✓ Occurs during sleep

Based on the low frequency of detection of lesions such as arteriovenous malformation or brain tumors, the American Academy of Neurology practice guidelines discourage the routine use of neuroimaging procedures in patients with headaches who have normal neurologic examinations. However, headache experts who often serve as tertiary referral sources may often ignore this recommendation because of the lack of diagnostic certainty in headache, lack of curative properties of current treatment, and unacceptable medical and legal risks of any missed diagnosis.

Although imaging procedures may not be considered necessary in the evaluation of primary headache syndromes, an image of the brain is mandatory for the evaluation of patients with severe or persistent headache, the “first” or “worst” headache or when a subdural hematoma is suspected. An abnormal neurological examination or seizure mandates brain imaging. A CT scan is indicated to rule out acute hemorrhage, while magnetic resonance imaging (MRI) is indicated when hydrocephalus, brain tumor, sinusitis, vasculitis, or posterior fossa lesions are suspected. X-rays of the jaw and cervical spine are useful to rule out malocclusions and degenerative changes of arthritis.

TREATMENT OF HEADACHE SYNDROMES

Migraine

The mainstay of migraine treatment is pharmacologic intervention. Treatment of migraine is divided into medications that prevent future attacks (prophylactic treatment), and interventions during the acute attack that provide symptom relief (abortive treatment).

Pharmacologic Prevention Treatments. There are not clear indications as to when preventative medications should be used, but in general, experts recommend that both the number of headache days per month as well as the level of impairment caused by the headaches be considered. Also, when nonpharmacologic approaches have failed and the frequency and severity of migraine attacks lead to impairment in functioning, preventative treatment is indicated. The major classes of drugs that have been investigated in the prophylaxis of migraine include the β-adrenergic blocking agents, antidepressants, anticonvulsants, calcium channel blockers, and aspirin. In the United States, five medications have FDA-approved indication for migraine prevention. In addition, flunarizine has been approved in the EU and Canada. The approved agents are two β-blockers, propranolol and timolol; two anti-epileptic drugs (AEDs) divalproex sodium and topiramate; and onabotulinumtoxinA has been approved for the treatment of chronic but not episodic migraine. Table 2.11–5 provides a summary of evidence for preventative treatments in episodic and chronic migraine, as well as a list of common and serious adverse effects. Clinical trials of migraine treatment are complicated by the high placebo response rate among subjects with migraine, the heterogeneity of diagnostic subtypes of headache, the intermittent nature of the condition, and the frequent use of additional analgesics to treat headache pain. Due to the highly variable placebo effect in prophylactic migraine RCTs the European Medicines Agency and the International headache Society have recommended the inclusion of placebo in comparative RCTs of migraine prophylaxis. However, only two of 26 such RCTs, published between 2002 and 2011, were placebo controlled.

Table 2.11–5.
Preventative Treatments for Episodic Migraine

Treatment	Recommendation[a]	Serious/Common Adverse Events
β-Blockers		
Propranolol, metoprolol, timolol	Level A: medications with established efficacy (>2 class I trials)	Lethargy, depression, dizziness, exercise intolerance, hypotension, and sleep disorders/avoid use in patients with asthma, diabetes, bradycardia, and congestive heart failure; exacerbation of angina and, in some cases, myocardial infarction, after abrupt discontinuance
Anticonvulsants		
Valproate	Level A: medications with established efficacy (>2 class I trials)	Weight gain, tremor, nausea, alopecia/pancreatitis, liver failure, thrombocytopenia
Topiramate	Level A: medications with established efficacy (>2 class I	Cognitive inefficiency, paresthesia, fatigue, weight loss/acute angle closure glaucoma, metabolic acidosis, hyperthermia

	trials)	
Gabapentin	Level U: inadequate or conflicting data to support or refute medication use	Dizziness, somnolence, fatigue, edema, suicidal behavior, and ideation
Antidepressants		
Amitriptyline	Level B: Medications are probably effective (1 class I or 2 class II studies)	Dry mouth, sedation, weight gain, constipation/cardiac dysrhythmias
Venlafaxine	Level B: Medications are probably effective (1 class I or 2 class II studies)	Nausea, vomiting, palpitations, tachycardia/suicidal thinking and behavior (suicidality), serotonin syndrome
Fluoxetine	Level U: inadequate or conflicting data to support or refute medication use	Insomnia, asthenia, tremor
Other		
OnabotulinumtoxinA 155u	No evidence for episodic migraine. Large-scale, double-blind, placebo-controlled trials in CM FDA approved for chronic migraine	Muscle weakness, pain/distant spread of toxin, effect, swallowing and breathing difficulties

[a]American Neurologic Association and American Headache Society.

PHARMACOLOGIC TREATMENTS

Drugs with FDA Approval for Migraine. Propranolol and timolol are the two β-blockers that have an indication for migraine and have been given an A level recommendation. β-Blockers have few side effects and may also treat comorbid cardiovascular diseases in people who suffer from migraine. Clinicians should be particularly cautious in prescribing this class of drugs to individuals with a history of depression, since the β-blockers are associated with the development of anhedonia, irritability, and lassitude which may occur after many months on any of these agents. In contrast, patients with high levels of autonomic anxiety may actually benefit from this class of drugs.

Divalproex sodium and topiramate are the two AEDs that have been FDA approved in the United States as migraine prophylactic agents and both have demonstrated superiority compared with placebo with Level A evidence. Topiramate has also been studied in chronic migraine and shown efficacy compared with placebo. Flunarizine is approved in the EU and Canada for migraine prevention but it is not available in the United States. OnabotulinumtoxinA has been studied in a variety of headache conditions but significant efficacy has only been shown in its placebo-controlled trials

in chronic migraine.

Drugs with Strong Evidence for Efficacy and Tolerability. Besides those that have FDA approval, numerous other therapies have been studied and used in the treatment of migraine and a number of these have been subject to well-designed clinical trials. Metoprolol, like other β-blockers, has shown efficacy in migraine treatment. Amitriptyline, a tricyclic anti-depressant, is used in a variety of pain syndromes besides migraine, and at first had received a top-tier recommendation in the original US migraine treatment guidelines. In the most recent guidelines, however, it was revised downward because of issues with percentages of patients completing the trials. The major side effects of amitriptyline of sedation and weight gain are often not well-tolerated in migraine patients. The secondary amines (e.g., nortriptyline and desipramine) appear to be efficacious in the treatment of depression, but have fewer side effects than do the parent tertiary amines (e.g., amitriptyline, imipramine). However, the relative efficacy of the various tricyclic antidepressants in migraine prevention has not been examined.

Drugs with Lesser Evidence for Efficacy and Tolerability. Agents in this category include tricyclic antidepressants other than amitriptyline, calcium channel blockers (other than flunarizine), angiotensin-converting enzyme inhibitors, monoamine oxidase inhibitors, NSAIDs, magnesium, and herbal derivatives including MIG-99 derived from the feverfew plant and the vitamin riboflavin. The selective-serotonin uptake inhibitors do not have demonstrated efficacy in migraine. In fact, many patients complain of headache as a secondary effect of the latter class of drugs. Daily aspirin may also be highly effective in migraine prophylaxis but has not been studied systematically. Finally, methysergide, an ergot alkaloid, has been used historically for the prevention of migraine. The risk of retroperitoneal fibrosis has diminished its role in the treatment of migraine.

The strong association between migraine with both depression and anxiety should be considered in the treatment of individuals with migraine. Systematic evaluation of the lifetime history of both depression and anxiety is necessary for determining optimal treatment strategies. If there is a subtype of migraine associated with anxiety and depression, it is critical to treat the entire syndrome rather than limiting the treatment goal to headache cessation. In general, comorbid depression and anxiety are more important in the selection of migraine prophylaxis than is the treatment of an acute attack of migraine. The use of prophylactic medications with side effects of lassitude, fatigue, or depression should be avoided, if possible. If not, careful clinical evaluation of the above-cited manifestations of depression including anergia, hypersomnia, and irritability should be monitored. Given the overlap of symptoms of the actual migraine episode, including acute changes in energy, appetite, mood, and level of anxiety as

well as those that occur between attacks, and those on the anxiety/depression spectrum, it is not surprising that similar pharmacologic agents have been successfully employed in the treatment of migraine and the anxiety/depression.

Drug Treatment for Children and Adolescents. There are few well-designed large clinical trials that have examined the efficacy of migraine prophylactic drugs for children. The only effective agents that have been effective in migraine prophylaxis for children in multiple controlled trials are topiramate and flunarizine. The remainder of the studies failed to show superiority of migraine treatments over placebo, or were open-label studies or retrospective reviews. Despite the lack of robust evidence, cyproheptadine is commonly used as a preventative medication in this age group, particularly for children under 12 years of age.

Nonpharmacologic Treatments. There has been substantial research on the use of behavioral treatments including biofeedback, relaxation training, and cognitive behavioral therapy for migraine prevention. Although most of the studies show greater reductions in migraine for the intervention compared to the no-treatment group, the lack of both clinician and patient blinding in these studies diminishes the extent to which such studies may provide conclusive evidence for efficacy. Likewise, herbal treatments such as feverfew (Tanacetum parthenium) and dietary supplements such as magnesium and vitamin D have been assessed in a few studies.

Acute Management of Migraine. The NSAIDs including ibuprofen, naproxen sodium, and indomethacin and the analgesics ASA and acetaminophen were commonly used as the first-line treatment of mild-to-moderate migraine. Excedrin was approved for labeling for the indication of migraine, as were the ibuprofen drugs Advil and Motrin. However practice has gradually shifted from a treatment paradigm involving slowly escalating abortive therapy toward starting with combination therapy—NSAIDs or paracetamol and triptan—from the outset. The seven triptans—almotriptan, eletriptan, frovatriptan, naratriptan, rizatriptan, sumatriptan, and zolmitriptan—are 5HT1B/1D receptor agonists with vasoconstrictive effects on blood vessels. The choice of triptan depends on efficacy, side effects, duration of headache, and coexistent vomiting and cost. Efficacy data at 2 hours favor eletriptan and rizatriptan. At 24 hours, eletriptan performed best, followed by zolmitriptan. The route of triptan administration is best tailored to the individual's migraine attack. Patients with rapid onset migraine might try a fast acting oral triptan such as rizatriptan, zolmitriptan, or eletriptan or nasal preparations that are faster (zolmitriptan and sumatriptan); alternatively subcutaneous sumatriptan has the fastest onset of action. The triptans with a longer half-life and slower onset of action have fewer side effects, for example naratriptan and

frovatriptan. Triptans should be avoided in individuals with uncontrolled hypertension, cardiovascular, or cerebrovascular disease (Table 2.11–6).

Table 2.11–6.
Acute Treatment of Migraine

Drug
5-HT1B/1D AGONISTS
Sumatriptan (subcutaneous) 6 mg
Sumatriptan (50 mg)
Rizatriptan (10 mg)
Eletriptan (80 mg)
Almotriptan (12.5 mg)
OTHER
Dihydroergotamine (intranasal)
NSAIDs
Aspirin (325 mg)
Combinations of NSAIDs and anti-emetics (Prochlorperazine Maleate; thioridazine; metoclopramide)

Other classes of drugs that are commonly prescribed for more severe attacks include ergot derivatives (ergotamine and dihydroergotamine), and narcotics. Ergotamine tartrate and dihydroergotamine are two of the most commonly prescribed ergot derivatives for moderate to severe attacks of migraine. In order to counterbalance the common side effect of nausea, metoclopramide or prochlorperazine are recommended.

Combination agents generally comprised of barbiturates, analgesics, and caffeine are also highly effective in the treatment of migraine episodes. Clinicians should be particularly alert to the dangers of abusing drugs such as ergotamine and narcotics. In general, narcotics should be restricted to severe attacks that are not responsive to other agents. Percocet and Percodan are two of the most popular drugs among opiate addicts who prefer these oral narcotics because oxycodone, their narcotic ingredient, is short acting, effective orally, and a euphoriant with a high street value. In acute use, however, narcotics do not produce addiction.

Tension-Type Headache

Pharmacological Treatments. At present Ibuprofen (800 mg) is the first choice for acute treatment of tension-type headache followed by naproxen sodium (825 mg) that has a higher risk of GI bleed (OR = 9.1) because it has the lowest risk of gastrointestinal bleeding or perforation. These recommendations are based on several drug trials that fulfilled the criteria set by the International Headache Society. Taken together, these studies suggest that although simple analgesics, aspirin (500 mg or 1,000 mg) and various NSAIDs are more effective than placebo in aborting tension headaches, the effectiveness of aspirin is comparable to that of paracetamol

(500 to 1,000 mg) while that of NSAIDs is superior to simple analgesics. The combination of analgesics and caffeine, sedatives, or tranquilizers might be more effective in some patients than simple analgesics or NSAIDs. The adjunction of caffeine (130 mg or 200 mg) significantly increases the efficacy of simple analgesics and ibuprofen in controlled trials.

Preventative treatments such as amitriptyline or topiramate can be used for people with chronic tension-type headache. The initial dose of tricyclic antidepressants should be low (10 to 25 mg amitriptyline or clomipramine before bed), and increased gradually. The average required dose of amitriptyline for patients with chronic tension-type headache, however, is 50 to 75 mg per day. Most authorities recommend the discontinuation of treatment after 6 months, regardless of efficacy. A decrease in the daily dose by 20 to 25 percent every 2 to 3 days might avoid rebound headache. The mode of action of antidepressants in chronic tension-type headache is unclear. Tricyclic antidepressants' effect on increasing serotonin and endorphin levels, or inhibiting NMDA receptors, all important chemical components of the pain pathway, might all play a role in alleviating pain in tension-type headache. However, SSRIs have as not shown to be effective for prevention of tension-type headache.

Nonpharmacological Treatments. There are numerous behavioral treatments that have been tested in the treatment of tension-type headache including relaxation and biofeedback alone or in combination. Cognitive behavioral treatments, such as stress management, are also effective in treating tension-type headache, especially when combined with biofeedback or relaxation therapies in subjects with high levels of stress.

The improvements produced by behavioral treatment might appear slowly compared with those produced pharmacologically; however, improvement is maintained for longer periods—up to several years—without monthly sessions or contact with the therapist. The combination of stress management techniques and tricyclic antidepressants has been shown to be superior to either treatment alone in treating chronic tension-type headache.

A systematic review of randomized clinical trials with physiotherapy and spinal manipulation in patients with tension-type headache showed that there is insufficient evidence to support the effectiveness of such techniques. Treatments such as massage, transcutaneous electrical nerve stimulation, the application of heat or cold, have not been shown to be effective in the long-term treatment of tension-type headache. In addition, as already mentioned in the section on nonpharmacological treatments for migraine, cervical manipulations are associated with the possibility of vertebral artery dissection.

Cluster Headache

First-line abortive therapy is subcutaneous sumatriptan with most people obtaining relief quickly. Home oxygen can be very useful at high flow, 100 percent concentration via a non-rebreathable mask at 12 to 15 L/min for up to 20 minutes. Transitional treatments, such as oral steroids, can help the patient through the acute phase while starting preventative strategies. Prophylactic medicine is almost always indicated for treating cluster headache because of the extreme severity of pain induced by an acute attack, which often occurs at night. Medications that have been shown to be effective in preventing attacks of cluster headache are lithium, the corticosteroids, methysergide, the calcium channel blockers, β-blockers, and valproic acid. Side effects can be severe, and combinations of these agents are often necessary to achieve success. Some may benefit from adjuvant topiramate. Histamine desensitization and surgical intervention are options upon exhaustion of traditional agents.

Posttraumatic Headache

There have been no randomized double blind, placebo-controlled studies of posttraumatic headache in either adults or children. Therefore, treatment should be tailored to the symptoms of the individual patient. Most people with posttraumatic headache have characteristics of either tension-type headache or migraine, and so treatment is directed toward those entities. If a peripheral mechanism, such as muscular or ligamentous changes is identified, physical therapy may be indicated. Clinicians should be aware of legal action that may diminish motivation of the patient to improve.

References

Anttila V, Winsvold BS, Gormley P, et al. Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nat Genet*. 2013;45(8):912–917.

Burch RC, Loder S, Loder E, Smitherman TA. The prevalence and burden of migraine and severe headache in the United States: updated statistics from government health surveillance studies. *Headache*. 2015;55(1):21–34.

Buse DC, Lipton RB. Global perspectives on the burden of episodic and chronic migraine. *Cephalalgia*. 2013;33(11):885–890.

Cameron C, Kelly S, Hsieh SC, et al. Triptans in the acute treatment of migraine: A Systematic review and network meta-Analysis. *Headache*. 2015;55(Suppl 4):221–235.

Charles A, Hansen JM. Migraine aura: new ideas about cause, classification, and clinical significance. *Curr Opin Neurol*. 2015;28(3):255–260.

Chiang CC, Schwedt TJ, Wang SJ, Dodick DW. Treatment of medication-overuse headache: A systematic review. *Cephalalgia*. 2016;36(4):371–386.

Davanzo R, Bua J, Paloni G, Facchina G. Breastfeeding and migraine drugs. *Eur J Clin Pharmacol*. 2014;70(11):1313–1324.

Derry CJ, Derry S, Moore RA. Sumatriptan (all routes of administration) for acute migraine

attacks in adults—overview of Cochrane reviews. *Cochrane Database Syst Rev*. 2014;5:CD009108.

Diener HC, Charles A, Goadsby PJ, Holle D. New therapeutic approaches for the prevention and treatment of migraine. *Lancet Neurol*. 2015;14(10):1010–1022.

Diener HC, Solbach K, Holle D, Gaul C. Integrated care for chronic migraine patients: epidemiology, burden, diagnosis and treatment options. *Clin Med (Lond)*. 2015;15(4):344–350.

Freitag FG, Shumate D. Current and investigational drugs for the prevention of migraine in adults and children. *CNS Drugs*. 2014;28(10):921–927.

Kacperski J, Hershey AD. Preventive drugs in childhood and adolescent migraine. *Curr Pain Headache Rep*. 2014;18(6):422.

Kalaydjian A, Merikangas K. Physical and mental comorbidity of headache in a nationally representative sample of US adults. *Psychosom Med*. 2008;70(7):773–780.

Kristoffersen ES, Lundqvist C. Medication-overuse headache: epidemiology, diagnosis and treatment. *Ther Adv Drug Saf*. 2014;5(2):87–99.

Lateef TM, Cui L, Nelson KB, Nakamura EF, Merikangas KR. Physical comorbidity of migraine and other headaches in US adolescents. *J Pediatr*. 2012;161(2):308–313.e1.

Lopez JI, Holdridge A, Rothrock JF. Hemiplegia and headache: a review of hemiplegia in headache disorders. *Curr Pain Headache Rep*. 2015;19(2):471.

MacGregor EA. Migraine management during menstruation and menopause. *Continuum (Minneap Minn)*. 2015;21(4 Headache):990–1003.

MacGregor EA. Migraine in pregnancy and lactation. *Neurol Sci*. 2014;35(Suppl 1):61–64.

Merikangas KR. Contributions of epidemiology to our understanding of migraine. *Headache*. 2013;53(2):230–246.

Merikangas KR, Cui L, Richardson AK, et al. Magnitude, impact, and stability of primary headache subtypes: 30 year prospective Swiss cohort study. *BMJ*. 2011;343:d5076.

Nagel MA, Khmeleva N, Boyer PJ, et al. Varicella zoster virus in the temporal artery of a patient with giant cell arteritis. *J Neurol Sci*. 2013;335(1–2):228–230.

Penzien DB, Irby MB, Smitherman TA, et al. Well-established and empirically supported behavioral treatments for migraine. *Curr Pain Headache Rep*. 2015;19(7):34.

Pringsheim T, Becker WJ. Triptans for symptomatic treatment of migraine headache. *BMJ*. 2014;348:g2285.

Sandrini G, Friberg L, Coppola G, et al. European Federation of Neurological Sciences. Neurophysiological tests and neuroimaging procedures in non-acute headache (2nd edition). *Eur J Neurol*. 2011;18(3):373–381.

Smitherman TA, Wells RE, Ford SG. Emerging behavioral treatments for migraine. *Curr Pain Headache Rep*. 2015;19(4):13.

Starling AJ, Dodick DW. Best practices for patients with chronic migraine: burden, diagnosis, and management in primary care. *Mayo Clin Proc*. 2015;90(3):408–414.

Weatherall MW. Drug therapy in headache. *Clin Med (Lond)*. 2015;15(3):273–279.

Weatherall MW. The diagnosis and treatment of chronic migraine. *Ther Adv Chronic Dis*. 2015;6(3):115–123.

Winther Schytz H, Olesen J. Laboratory tests of headache disorders - dawn of a new era?

Cephalalgia. 2016 Jan 13.
Yonker M, Mangum T. Migraine management in children. *Curr Neurol Neurosci Rep*. 2015;15(5):20.

▲ 2.12 Neuropsychiatric Aspects of Neuromuscular Diseases

TIMOTHY P. WUERZ, D.O., RAJEET SHRESTHA, M.D., AND ALAN J. LERNER, M.D.

INTRODUCTION

The peripheral nervous system (PNS) is a complex system of interconnected motor, sensory, and autonomic nerves and pathways that allow for rapid information passage to and from the CNS. In daily life, it subserves vital functions including sensation leading to perception, movement around joints throughout the body, and coordination of automatic functions.

A wide variety of disorders occur in the PNS, and two major classes of neuropsychiatric aspects can be considered. Some disorders, such as amyotrophic lateral sclerosis (ALS) or systemic vasculitis, may affect both PNS and CNS simultaneously. Other disorders, such as myasthenia gravis affect solely the PNS, and their neuropsychiatric aspects arise directly from the disabling injury, its treatment, or its psychosocial effects. Medications and toxic exposures may also affect both systems. This chapter will cover the structure of the PNS, and major disorders with neuropsychiatric aspects important for clinical practice.

STRUCTURE OF THE PERIPHERAL NERVOUS SYSTEM

The traditional separation of the nervous system into PNS and CNS is not always anatomically straightforward and at times results in ambiguous distinctions. Some portions of the PNS have their cell bodies located in the CNS (e.g., anterior horn cells or motor neurons that innervate skeletal muscle fibers), whereas other neurons originating in the PNS (e.g., sensory neurons located in the dorsal root ganglia) project axons into the CNS where they synapse in the dorsal horn of the spinal cord (e.g., ascending nerve tracts in the spinal cord). The autonomic nervous system can be divided into sympathetic and parasympathetic units, and has central and peripheral components. Nevertheless, this imperfect separation is well established and has resulted in distinct neurological subspecialties, including those related to peripheral neuromuscular diseases and clinical electrodiagnostic medicine.

The PNS is defined to include the motor neurons and their axons projecting into the periphery through peripheral nerves, the neuromuscular junction, and the innervated musculature. A motor unit encompasses all muscle fibers innervated by a single motor axon. This may vary from a few

fibers in an eye muscle to several hundred in the quadriceps, allowing for graded contraction responses. The peripheral sensory system consists of various modality specific receptors including touch, pain, temperature, vibration and joint position sense, afferent sensory axons, their cell bodies in the dorsal root ganglion, and the portion of the axon projected centrally that comprises the dorsal nerve roots.

The ventral (motor) and dorsal (sensory) segmental nerve roots merge to form the segmental spinal nerves, which in turn combine to form the brachial and lumbosacral plexuses before dividing into the individual peripheral nerves (sensory, motor, or mixed sensorimotor), plus their autonomic nervous system components. The cranial nerves (except olfactory and optic nerves) are similar to the segmental nerves, comprised of peripheral and central segments, but traversing or originating in the brainstem instead of the spinal cord.

Neuronopathies and Disorders of The Dorsal Root Ganglion

Poliomyelitis and Other Disorders of Anterior Horn Cells

Poliomyelitis has been eliminated throughout much of the world, although pockets of infection remain in rural areas in Pakistan, Afghanistan, and scattered cases in the Middle East and Africa. Polio is the quintessential lower motor neuron disease (MND) affecting motor neurons in the spinal cord, although bulbar symptoms may occur. Syndromes similar to polio may occur with arthropod borne diseases such as WNV infection, and with idiopathic cases of flaccid paralysis. WNV may be limited to a segment of the spinal cord, with concomitant weakness, but encephalitis with mental status changes and coma may occur. The prognosis for recovery from West Nile Virus (WNV) infection is guarded, and supportive care and counseling throughout are important components of care for patients with this condition.

Postpolio Syndrome

Postpolio muscle dysfunction disorder has an unclear etiology. A consensus workshop in 1996 proposed the following diagnostic criteria: (1) history of paralytic polio: confirmed or not confirmed with partial or fairly complete functional recovery; (2) after a period of functional stability of at least 15 years, development of new muscle dysfunction emerges; (3) muscle weakness; (4) muscle atrophy; (5) muscle pain; and (6) fatigue. Neurological examination in postpolio muscle dysfunction should be compatible with prior poliomyelitis with lower motor neuron lesion, decreased or absent tendon reflexes, no sensory loss, and compatible findings on EMG and/or MRI.

Postpolio syndrome can lead to significant distress, anxiety, and depression as the pain and fatigue cause marked change in the level of

functioning. In a sample of polio survivors, more than a quarter had depression while those with postpolio syndrome and lower family functioning were at higher risk of depression. It is important to carefully assess postpolio patients for psychiatric complications and to consider interventions including counseling and anti-depressants. Sedatives and alcohol should be avoided due to increased risk of falls and hypoventilation.

Amyotrophic Lateral Sclerosis

Long considered a very rare disease, ALS, the most common form of MND has an annual incidence of roughly 2 per 100,000 and prevalence of 6 per 100,000 persons. About 5 to 20 percent of cases are familial, and several different genes have been implicated. The typical clinical presentation is the combination of both upper motor neuron and lower motor neuron signs. Lower motor signs include weakness, muscle atrophy, presence of fasciculations, which are painless spontaneous discharges of motor units and involuntary rapid twitches in limb or trunk muscles. Electrodiagnostic testing is a critical part of the evaluation and may also reveal fibrillations, a sign of active denervation. Upper motor neuron signs include spasticity, loss of dexterity, hyperreflexia, and pseudobulbar palsy. Death from respiratory failure or aspiration typically occurs within 3 years of diagnosis, although occasionally patients show a prolonged course, particularly if a decision is made by the patient to accept respiratory support. Because of the uniformly fatal course of ALS, a wide-ranging differential diagnosis should be considered, encompassing virtually the entire neuraxis.

Table 2.12–1.
Frontotemporal Spectrum Disorders Associated with Amyotrophic Lateral Sclerosis

Frontotemporal lobar dementia (FTLD)
Pick disease (also known as behavioral variant FTD)
Frontotemporal dementia with parkinsonism, chromosome 17 (FTDP-17)
Primary progressive aphasia (PPA)
Neuronal intermediate filament inclusion dementia (NIFID)
Corticobasal degeneration (CBD)
Dementia lacking distinct histopathological features (DLDH)
Argyrophilic grain disease (AGD)
Progressive supranuclear palsy (PSP)

The pathology of ALS involves degeneration of the motor neurons of the spinal cord (anterior horn cells) and brainstem as well as the upper motor neurons of the cerebrum. Gliosis and neuronal loss occurs in superficial layers of the dorsomedial neocortex of the frontal lobes and may occur in the hippocampus and parahippocampus in the temporal lobes.

There has been increasing recognition of the neuropsychiatric aspects of ALS, most notably its association with frontotemporal degeneration (FTD). FTD is clinically heterogeneous and may present as a behavioral and dysexecutive syndrome, primary progressive aphasia, or parkinsonism, with features of progressive supranuclear palsy, or corticobasal syndrome (Table 2.12–1). ALS may develop in patients with previously diagnosed FTD or vice versa. Patients with the C9ORF72 expanded hexanucleotide repeat mutation may present with behavioral syndromes, and especially psychosis and agitation. These seemingly disparate disorders are linked by genetic and pathological commonalities, with ubiquitin-positive neuronal cytoplasmic inclusions in lower motor neurons, hippocampus, and neocortex in both conditions. Gliosis and neuronal loss in superficial layers of the dorsomedial neocortex of the frontal lobes also occurs in both syndromes.

One-third to one-half of patients with ALS perform poorly on neuropsychological testing compared with matched control subjects. The pattern of cognitive change generally includes mild memory loss, with associated word finding difficulties, and executive function impairments. Facial emotional recognition is also impaired in motor neuron patients. These patterns of cognitive impairment differ substantially from that seen in Alzheimer disease, wherein memory loss and apraxia predominate.

Pseudobulbar affect is a syndrome of disinhibited affective display in which patients demonstrate inappropriate and exaggerated laughing or weeping responses to environmental stimuli and may affect up to 50 percent of ALS patients. The behaviors are not usually associated with disturbances of the underlying mood, and the phenotype behavior tends to be stereotypical for each patient. Successful treatment of pseudobulbar affect in an ALS patient has been reported with tricyclic compounds or selective serotonin reuptake inhibitors. A combination of dextromethorphan and time released quinidine (Nuedexta) has been found to be an effective treatment in many cases.

Spinal Muscular Atrophies

Spinal muscular atrophies (SMAs) are disorders with lower motor neuron dysfunction, due to loss of anterior horn cells. The syndromes may present in infancy (Werdnig–Hoffman syndrome; SMA I), childhood (chronic infantile; SMA II), adolescence (Kugelberg–Welander disease, SMA III), or adulthood (SMA IV). Developmental arrest and limited survival are features of SMA type I and II. SMA is an autosomal recessive disorder, with an incidence of about 1 per 10,000 live births. The differentiation of these complex disorders is increasingly being done by genetic analysis. The autosomal recessive proximal SMA have been linked to mutations or deletions in the SMN (survival motor neuron) gene on chromosome 5q11.12–13.3. Variants of SMA also linked to chromosome 5q include SMA

with arthrogryposis, and a very severe form with survival measured in weeks after birth. There are also SMA syndromes, often with other neurological features, that are not linked to 5q mutations. These include Kennedy disease (X-linked bulbospinal muscular atrophy with androgen resistance), Fazio–Londe (a progressive bulbar atrophy of childhood), late-onset Tay–Sachs disease, focal atrophy syndromes, and distal SMAs ("spinal" Charcot–Marie–Tooth disease, and autosomal dominant adult-onset SMA).

Dorsal Root Ganglion Disorders

Disorders of the dorsal root ganglia present as pure sensory syndromes. Paraneoplastic sensory neuropathies due to inflammation of the dorsal root ganglion may be associated with the production of other autoimmune antibodies producing paraneoplastic encephalomyelitis. Symptoms of encephalomyelitis include cognitive and behavioral disturbances (seen in limbic encephalitis), ataxia, seizures, cerebellar ataxia, opsoclonus, myoclonus, and gut dysmotility.

PERIPHERAL NEUROPATHIES

There are a great number of ways to classify peripheral neuropathies. A clinically useful two-pronged approach is to consider (1) the onset and course as acute, subacute, or chronic; (2) whether the involvement is pure motor, mixed motor and sensory, or pure sensory, and with or without autonomic involvement. This classification applied to specific disorders is presented in Table 2.12–2.

Sensory or Sensorimotor Neuropathies

Neuropathies involving sensory nerves or their cell bodies, typically present with symptoms and signs of sensory loss involving the distal limbs. Sensory symptoms are very hard to describe in ordinary language: symptoms described as "burning," freezing," or "cutting" may represent identical pathophysiology, and indeed, may be described alternately or even simultaneously in the same patient. Despite decreased sensation, patients paradoxically show hyperesthesia and often complain of uncomfortable sensations, characterized as burning and tingling in response to light touch or at rest. Traditional modality specific sensory testing is quite crude, aimed primarily at detecting changes in sensory thresholds. Clinical testing does not address the stimulus–response relationship, alterations in which may help explain this apparent contradiction. Most neuropathies are sensorimotor in nature, as most peripheral nerves subserve both functions.

Table 2.12–2.
General Categories of Sensory or Sensorimotor Neuropathies

Acute	Subacute/Chronic	Subacute/Chronic
Guillain–Barré syndrome (Acute Inflammatory Demyelinating Polyneuropathy [AIDP])	Vitamin B_{12} deficiency	Cryoglobulinemia
Toxic	Environmental toxicity (Arsenic, lead, N-hexane, Organophosphates, Taxol, Pyridoxine, Cisplatin, Vincristine, Dapsone, Chloroquine, many other medication and toxins)	Acromegaly
HIV associated AIDP	Paraneoplastic	Graft versus Host disease
Vasculitis	Diabetes	Thiamine deficiency
Lyme disease	Hypothyroidism	Gastric bypass
Acute pandysautonomia	Monoclonal gammopathies (esp. associated with multiple myeloma or other neoplasms)	
Porphyria	Sarcoidosis	
Critical illness neuromyopathy	Connective tissue disorders (SLE, Sjogren, Scleroderma)	
	Hereditary (Charcot–Marie–Tooth, Fabry disease, Tangier disease)	
	POEMS syndrome	
	Amyloidosis	

Diagnosis of Neuropathies

The diagnostic evaluation of a neuropathy is a highly specialized process, with a bewildering number of conditions possible with overlapping clinical phenotypes (Table 2.12–2). Clinical history may suggest positive family history, the presence of early onset and slow progression, consistent with an inherited etiology. History may also suggest toxic exposures and systemic conditions (neoplasms, autoimmune disorders, medications, and toxic exposures) associated with neuropathies. The American Academy of Neurology recommends checking a serum electrophoresis for underlying monoclonal gammopathy and a vitamin B_{12} level in addition to electrodiagnostic testing for further classification. Nerve and muscle biopsy may be indicated but some neuropathies may remain undiagnosable.

Treatment of Neuropathies

Treatment of neuropathy is primarily directed at the underlying cause. Treatment of neuropathic pain is typically resistant to traditional analgesics (i.e., NSAIDs or opiates) and requires a multimodal approach, often requiring adjunctive pain medications. Antidepressants with mixed pharmacologic effects, are useful in chronic pain syndromes, including the pain associated with some peripheral neuropathies. Amitriptyline (Elavil), is inexpensive and effective for this indication, and is favored by many as a first choice agent. Doses in the range of 25 to 75 mg per day are usually sufficient to generate some symptomatic improvement. Imipramine may also be used, and has fewer anticholinergic adverse effects. Venlafaxine has also been recommended as a first-line treatment for neuropathic pain, at doses similar to those used for treatment of depression. Duloxetine (Cymbalta) is FDA approved in the United States for treatment of painful

diabetic neuropathy. Daily doses for duloxetine range from 30 mg per day up to 60 mg twice a day.

Pregabalin (Lyrica) is indicated for several neuromuscular conditions, specifically, management of neuropathic pain associated with diabetic peripheral neuropathy, management of postherpetic neuralgia (PHN), management of fibromyalgia, and management of neuropathic pain associated with spinal cord injury.

Neuropsychiatric adverse effects include mood changes, anxiety, panic attacks, sedation or insomnia, impulsivity, agitation or aggression, and suicidal ideation. Dosage ranges from 50 mg three times daily up to 100 mg three times a day in patients with creatinine clearance of at least 60 mL/min, with reduced dosage based on renal function.

Although frequently used for neuropathic pain, gabapentin (Neurontin) is indicated for usage in postherpetic neuralgia (PHN). Doses range from 100 mg/day up to 600 mg three times daily and must be adjusted for renal insufficiency. Neuropsychiatric side effects profile is similar to that seen with pregabalin.

Clinicians must be aware that the approved indications for the use of duloxetine, pregabalin, and gabapentin with regard to neuropathic pain are limited. Their usage, or the use of many antidepressant medications used for modulation of neurotransmitter systems involved in pain transmission and perception, in other pain syndromes, other neuropathies or for vague indications is encountered commonly, and may be of limited efficacy. Supportive care, correction of metabolic abnormalities, and removal of toxic exposures need to be integrated in the care model. Physical and occupational therapy can address motor weakness and provide techniques and adaptive equipment to help overcome sensory loss.

Diabetic Neuropathy

The most common cause of sensory or sensorimotor neuropathy in the United States is diabetes mellitus. In primary care settings, more than half of diabetic patients have symptoms and signs of peripheral neuropathy, and diabetes may be newly diagnosed during the evaluation of a new-onset painful neuropathy. Improved glycemic control slows progression of diabetic neuropathy, but symptomatic improvement of unpleasant sensory symptoms remains a major clinical challenge.

Guillain–Barré Syndrome

Guillain–Barré syndrome (GBS) is an acute inflammatory demyelinating polyneuropathy (AIDP). It is classically described as an ascending paralysis, frequently with sensory loss, and autonomic instability. The pattern of presentation is somewhat variable, and may be misdiagnosed as a nonspecific illness at first. Many cases follow a nonspecific illness, such as

an upper respiratory infection. Cases of GBS, especially those associated with *Campylobacter jejuni* infection may have axonal injury in addition to demyelinating features. The presence of axonal injury and loss may predict a poorer prognosis. During the acute phase, the disorder can be life-threatening with about a quarter requiring admission to an intensive care unit (ICU) for mechanical ventilation. The majority of individuals affected with GBS recover and similar degrees of improvement rate are seen with both gamma globulin infusions as with plasmapheresis.

It is important for the psychiatrist to be aware of the associated neuropsychiatric manifestations both directly and indirectly from an acute GBS syndrome with a resulting ICU admission in terms of anxiety, acute stress disorder, depressive episodes, and brief reactive psychosis. Psychotic episodes are strongly associated with severe quadriparesis, artificial ventilation use, and multiple cranial nerve dysfunctions. Patients express that experiencing loss of communication is the most difficult condition to cope with. Therapeutically, patients often feel reassured by the environment of the ICU and described contact with relatives to be most helpful. Supportive counseling is a valuable adjunct, in addition to psychopharmacological measures, in helping patients cope with this illness, which can be devastating in its effects but often carries a good prognosis.

Chronic Inflammatory Demyelinating Polyneuropathy

Chronic inflammatory demyelinating polyneuropathy (CIDP) is an immune based neuropathy of unknown cause, and may resemble GBS in its early stages. Both cellular and humoral components of the immune response appear to be important in pathogenesis, although the precipitating antigen(s) is unknown. There is breakdown of the blood–nerve barrier, and recruitment of macrophages. Secretion of toxic factors then damage myelin sheaths and may produce axonal injury. In HIV-positive patients there may be as many as 50 WBC/cubic mm in the CSF mimicking a chronic CNS inflammatory process.

CIDP may occur at any age, and may occur in connection with other autoimmune disorders, and monoclonal gammopathy of unknown significance. Successful treatments of CIDP with corticosteroids, immunoglobulin infusions, plasmapheresis, and immunosuppressants have been reported. The uncertainties regarding prognosis present a special challenge to the patients, and dealing with these iatrogenic and situational problems can benefit from psychiatric consultation.

Nutritional and Alcoholic Neuropathies

Alcohol can affect every level of the PNS and CNS. Persons with alcohol-related neuropathies frequently have a history of alcohol dependency and other co-morbid vitamin and nutritional deficiencies. Co-morbid alcohol-

related major neurocognitive disorder is complicated by the various syndromes described in individuals who abuse alcohol, as well as other possible co-morbidities contributing to cognitive dysfunction in these individuals including vitamin B_{12} deficiency, subdural hematomas and head injuries, cerebrovascular disease, and so forth. It is also unclear to what degree alcohol-related major neurocognitive disorder overlaps with Wernicke–Korsakoff syndrome. Alcohol-related major neurocognitive disorder is diagnosed much more often than Korsakoff's, but many people diagnosed with alcohol-related syndromes have a lack of thiamine that is often missed by treating health professionals. Knowledge about whether alcohol abuse may be a risk factor for other dementias such as Alzheimer disease is also sparse.

Vitamin B_{12} deficiency affects multiple levels of the nervous system, which may present as a peripheral neuropathy, myeloneuropathy, or cognitive syndrome. Vitamin B_{12} is essentially found only in animal products, and deficiency can be due to decreased intake, or failure of absorption, since intrinsic factor is a gastric protein important for facilitating dietary absorption. Sensitivity in determining vitamin B_{12} deficiency may be enhanced by measurement of intermediate metabolites such as methylmalonic acid and serum homocysteine. Rapid repletion of vitamin B_{12} is usually accomplished by intramuscular administration of 1 g daily for a week, then weekly for the first month, followed by maintenance therapy on a monthly basis. Chronic oral repletion is possible, even in patients with low intrinsic factor activity, through a low-affinity uptake mechanism.

Cranial Nerve Disorders

The twelve cranial nerves are largely analogous to the segmental peripheral nerves, and have their cell bodies within the CNS. A variety of conditions can selectively affect cranial nerves, and occasionally these are associated with neuropsychiatric complications. Examples include cranial neuropathies secondary to concomitant meningeal inflammation, which may occur with granulomatous diseases. In this context, cranial neuropathies may be associated with acute or subacute alterations in baseline mental status, hydrocephalus, or major neurocognitive disorder.

A clinical challenge may occur with cranial neuropathies in association with neurodegenerative disorders. Anosmia (olfactory nerve: CN I) may occur years before the classic motor symptoms of Parkinson disease, or memory loss in Alzheimer disease. Visual hallucinations or complex visuospatial disturbances, seen in major neurocognitive disorder secondary to Alzheimer disease or dementia with Lewy bodies often calls for ophthalmological investigation of retinal and optic nerve (CN II) function as part of the neuropsychiatric treatment plan.

Peripheral Nerve Injury

Traumatic peripheral nerve injuries range from stretch injury without rupture of the nerve sheath (e.g., foot drop after hip surgery) to complete transection of the affected nerve(s). While neural repair with re-innervation of affected nerves and their targets often occurs, the estimated daily repair rate is about 1 mm/day; thus, the time required may be weeks to months. Patient expectations and secondary mood adjustments may respond to psychiatric interventions.

Concomitant traumatic brain injury (TBI), spinal cord, or severe musculoskeletal injury greatly complicates peripheral nerve injury, particularly in the acute posttrauma period. Concurrent therapies are frequently needed as the presence of multiple injuries are often associated with greater mood effects; the range of symptoms with TBI include altered mental state, psychosis, disinhibition, mood affects, and focal neurological injuries and should be treated as appropriate.

Metabolic, Endocrine Neuropathies, and Ischemic Neuropathies

This diverse group of disorders is of interest to psychiatrists, because of concurrent psychiatric co-morbidity. Metabolic and endocrine disorders associated with neuropathy include uremia, hepatic failure, hyper- and hypothyroidism, and neuropathy associated with acromegaly. Pituitary tumors in general are often associated with psychiatric symptoms. Changes in attention, memory, and executive function have been reported to be common in acromegalic individuals.

Frostbite is an ischemic neuropathy and psychiatric illness is very pervasive in victims of frostbite. As a result, psychiatric screening for patients with frostbite has been highly encouraged.

Environmental and Toxic Neuropathies

Toxic neuropathies can be caused by exposure to substances including chemicals or heavy metals. These exposures may be accidental (in the workplace or through drinking water) or intentional (homicide or suicide). Most toxic neuropathies produce symmetric deficits, but the symptoms and signs may vary, when the intensity or duration of exposure to the substance changes. A clue to toxic exposures is that the symptoms usually involve multiple organ systems and rarely presents as an isolated peripheral neuropathy. The first step in identifying a neuropathy as toxin-related is simply having a high degree of suspicion.

Lead. Lead toxicity is a well-known cause of neurological dysfunction. Potential for exposure is usually via inorganic lead and may occur in automobile factory workers, plumbers, exposure to outdated paint products, and consuming illegal “moonshine” whiskey. Lead intoxication in

children is most commonly associated with CNS dysfunction. This may lead to chronic cognitive dysfunction (developmental delay or loss of milestones) or acute encephalopathy with ataxia, tremulousness, choreiform movements and hallucinations, which may progress to seizures, coma, or death. Acute exposure in adults can also lead to encephalopathy. Chronic lead exposure presentation usually includes gastrointestinal (GI) disturbance, anemia, weight loss, fatigue, renal dysfunction, and bluish discoloration of the gums ("lead lines") in addition to a neuropathic picture. Porphyria should be included in the differential. Treatment consists of utilizing chelating agents and preventing or minimizing exposure.

Arsenic. Arsenic is famous as a poison used for suicide or homicide, usually resulting from a massive ingestion, although it can also result from chronic low-grade exposure from smelting of lead and copper ore, contaminated well water (a common occurrence in Bangladesh), tainted illicit drugs, and burning of preserved wood. High-level acute exposure causes multiple organ dysfunction, organic psychosis, CNS dysfunction, and possible death. Peripheral neuropathy manifests itself within weeks and may continue for a period of weeks after removal of the exposure. Peripheral neuropathy symptoms include burning and numbness of the feet followed by hands. Treatment of an acute ingestion requires ICU care. Chronic exposure treatment includes chelating agents and education of those at risk.

Mercury. Mercury toxicity can be of several types: (1) elemental mercury found in outdated thermometers, barometers, and other gauges. (2) Organic mercury found in latex paints as preservative, or as an environmental pollutant. (3) Inorganic mercury salts used in the manufacture of chlorine, in dental amalgam, and in the natural gas industry.

Intoxication with mercury results in predominantly CNS dysfunction (encephalopathy ["mad as a hatter"] and dystonic movements) but associated systemic toxicity includes GI symptoms, and nephrotic syndrome. PNS dysfunction is seen in all forms of mercury poisoning. Treatment of mercury intoxication is primarily by removal from exposure, use of chelating agents, and preventive education.

Carbon Disulfide. Carbon disulfide is a clear liquid that vaporizes at room temperature and is readily absorbed by inhalation, but also absorbed via GI tract and skin. It is also a major metabolite in the breakdown of the drug disulfiram (Antabuse). Acute or subacute exposure to high levels of carbon disulfide causes primarily CNS dysfunction including memory impairment, confusion, hallucinations, and emotional lability. Pyridoxine

may have a role in treatment.

Organophosphates. Organophosphorus (OP) compounds are used as insecticides, flame retardants, various other industrial manufacturing components, and in chemical warfare. Accidental exposure from agricultural spraying is the most common cause of individual intoxication. OP compounds are a frequent agent for suicide, and in India are implicated in almost 20 percent of suicides.

The sarin gas attack in the Tokyo subways has been well studied and a small number of survivors had a short-lived neuropathy. Posttraumatic stress disorder was common in those survivors. OPs have also been implicated in the Gulf War syndrome, with various combinations of peripheral and CNS symptoms.

OPs prevent the degradation of acetylcholine (ACh) by disrupting acetylcholinesterase resulting in excessive stimulation of both muscarinic and nicotinic receptors. There are acute (type I) and intermediate (type II) cholinergic syndromes depending on which receptor is preferentially affected. Type I results from excessive stimulation of muscarinic cholinergic receptors. An acute syndrome frequently begins hours after exposure and always within 1 day. Symptoms include GI effects, micturition, sweating, and tachy- or bradycardia. CNS involvement includes emotional lability, decreased alertness, cognitive impairment, convulsions, and coma in extreme cases. Type II occurs within 12 to 96 hours after exposure and is due to nicotinic ACh receptor activation in skeletal muscle. The initial feature is usually respiratory insufficiency followed by weakness of the neck flexors and proximal muscles. The cranial muscles including the extraocular muscles may be involved. Dystonic posturing may be seen. This neuropathy has been labeled the organophosphate-induced delayed polyneuropathy (OPIDP), since there is an interval of 7 to 21 days after exposure before peripheral neuropathy symptoms are seen. Although less common than the cholinergic syndromes, OPIDP causes significant morbidity including foot drop, claw hand deformity, or atrophy in extreme cases of severe neuropathy. Administration of atropine only treats the muscarinic type I syndrome and does not prevent the intermediate type II nicotinic cholinergic syndrome. Reactivation of acetylcholinesterase may be accelerated by the administration of pralidoxime. Cathartics and gastric lavage are indicated in the case of acute ingestion. Chronic low level exposure to OP can result in mood dysregulation, increase risk of suicidality, and increase impulsivity.

Hexacarbons. Hexacarbons are components of glues and lacquers and also used as solvents. Hexacarbon exposure occurs in the production and refining of petroleum products, and high-level exposure also occurs with

intentional inhalation such as "glue sniffing." Acute exposure causes CNS depression and stupor. Repeated exposure via "glue sniffing" leads to cognitive deficits, motor neuropathy, and cranial nerve and autonomic dysfunction. Severe neuropathies can result in sensory loss, atrophy, and distal weakness. The only treatment is cessation of hexacarbon exposure. Psychiatric concerns from "glue sniffing" include emergence of violent behavior and/or self-mutilation, functional deterioration, impairment of intelligence, and possible development of psychosis.

Vasculitic Neuropathies and Neuropathies of Connective Tissue Disorders. Systemic disorders, especially primary vasculitides and connective tissue disorders (many of which are autoimmune in etiology) often affect peripheral nerves. Concomitant neuropsychiatric complications may be primary such as with vasculitis affecting CNS structures, or secondary to treatment (e.g., steroid-related mood disturbance), or adjustment to severe chronic illness. Prominent syndromes affecting the CNS are systemic lupus erythematosus, Sjögren syndrome, giant cell arteritis, and disorders of circulating autoantibodies, especially the antineutrophil cytoplasmic antibodies (ANCA) classically seen in Wegener granulomatosis.

Infectious and Granulomatous Neuropathies. The infectious and granulomatous neuropathies result from a variety of inflammatory conditions. Peripheral nerve involvement in these disorders may manifest clinically as a mononeuropathy, mononeuropathy multiplex, polyneuropathy, or polyradiculopathy. Nerve pathology may result from an inflammatory reaction induced by the infection and from the host's immunological reaction.

HIV-ASSOCIATED NEUROPATHIES. More than one-third of AIDS patients are estimated to have a peripheral neuropathy. HIV neuropathy may affect proximal and distal components of the motor, sensory, and autonomic PNS. Other AIDS-related neurological conditions, such as myelopathy, myopathy, and dementia may coexist.

The diagnosis of HIV-associated neuropathy may be particularly complicated and challenging, as wide range of potential etiologies underlies peripheral neuropathy in HIV-infected patients. Etiologies of HIV-associated neuropathy include cytotoxic effects due to direct HIV infection, cytokine dysregulation, opportunistic infection, neoplastic processes, autoimmune disorders, metabolic abnormalities, and toxic effects of medications.

AIDS patients present with a variety of PNS syndromes. Distal symmetrical polyneuropathy (DSP) is a painful neuropathy occurring in 30 percent of AIDS patients. Patients complain of numbness, burning, and paresthesias of the feet with pain usually beginning in the toes and soles of

feet. Inflammatory demyelinating polyneuropathy produces rapidly progressive muscle weakness of two or more limbs, often starting in distal muscles and may resemble GBS or CIDP. Mononeuropathy multiplex is a neuropathy that affects single or multiple nerves causing asymmetric motor and sensory complaints presenting with weakness or dysesthesias in the distribution of cutaneous nerves, mixed nerves, and nerve roots. Progressive polyradiculopathy is a serious condition presenting as a rapidly, progressive cauda equina syndrome, characterized as leg weakness, difficulty walking, bowel and bladder dysfunction, low back pain, and leg paresthesias.

Prompt recognition and treatment of neuropathy may significantly benefit a patient's quality of life. Possible etiologies range from relatively benign facial neuropathy to severe opportunistic infections including lymphomatous meningitis, CMV encephalitis, and cryptococcal meningitis. In addition, cranial neuropathies occasionally are caused by primary CNS lymphoma, non-Hodgkin lymphoma, or brain stem glioma. Treatable conditions in HIV populations include B_{12} deficiency, ethanol abuse, neurotoxic medications, uremia, or diabetes.

Evaluation of AIDS or HIV-related neuropathies consist of ancillary tests including nerve conduction studies (NCS); needle electromyography (EMG); autonomic testing; and blood, CSF, and neuroimaging studies. Treatment of HIV-related peripheral neuropathy depends on etiology, but has four main components to be considered: antiviral treatment, immunomodulation, treatment of underlying causes, and symptom reduction.

HIV-associated peripheral neuropathy is often refractory to standard treatments for pain. For example, a placebo-controlled study of amitriptyline (tricyclic antidepressant) and mexiletine (antiarrhythmic) in patients with HIV-associated peripheral neuropathy found no significant improvement in pain intensity, relative to placebo, with either treatment. Similarly, a randomized, placebo-controlled study of the efficacy of acupuncture, amitriptyline, and a combination of the two treatments for pain relief in patients with HIV-associated peripheral neuropathy found that none of the treatments worked better than placebo. However, a recent study showed that lamotrigine worked significantly better than placebo to decrease neuropathic pain. Severe rash, an important adverse effect of lamotrigine treatment, occurred more frequently than in studies of lamotrigine treatment for epilepsy.

LEPROUS NEUROPATHY. Leprosy is caused by *Mycobacterium leprae,* an acid-fast bacillus, and is characterized as a chronic infectious disease that affects skin and peripheral nerves. The prevalence of leprosy has declined greatly in developed countries and is seen only rarely in immigrants from endemic areas. Sensory loss due to intracutaneous nerve damage is the

hallmark of leprous neuropathy. The goal of treatment for leprous neuropathy is eradication of *M. leprae*. The prognosis for neuropathy depends on the type of leprosy, the time of diagnosis, chemotherapeutic response, and reactive sites.

LYME NEUROPATHY. Lyme disease, the most common tick-borne disease in the United States, is caused by *Borrelia burgdorferi,* a gram-negative spirochete, and transmitted by *Ixodes dammini* and related species. After the incubation period of 3 to 30 days, three stages emerge. Stage I is characterized by a rash at the tick attachment site, causing constitutional symptoms, although many patients do not recall a tick bite. Neurological complications develop during stages II (early) and III (late). Lymphocytic meningitis and lymphocytic meningoradiculitis (LMR) are common in early neuroborreliosis. Patients with LMR usually have variable motor and sensory involvement of multiple nerves and symptoms may wax and wane over several months. Focal and multifocal peripheral and cranial nerve involvement is observed in half of patients. Facial nerve paralysis causing facial diplegia is the most common sign.

In Europe, neuroborreliosis has a somewhat different presentation with meningoradiculoneuritis characterized by severe radicular pain with or without paresis. Late neuroborreliosis is marked by peripheral neuropathy and encephalomyelitis. The PNS abnormalities of this late stage last for weeks to months. Predominant sensory neuropathy may be seen in up to two-thirds of untreated patients. Stage III is characterized by asymmetric oligoarthritis, cardiac involvement, and a variety of CNS syndromes. Evaluation of Lyme disease consists of a laboratory two-test protocol (ELISA) and Western blot performed on serum or CSF. PNS complications of Lyme disease are treated by Borrelia eradication. Neurological manifestations carry a good prognosis after antimicrobial therapy.

HERPES ZOSTER NEUROPATHY. Approximately 95 percent of the United States population has been infected with VZV, a Herpes virus that is the causative agent for chicken pox. Latent, lifelong infection persists in sensory nerve ganglia and can reactivate to cause zoster neuropathy, or “shingles.” VZV neuropathy results from reactivation of latent virus from a primary infection. VZV infection is more commonly recurrent and chronic in AIDS and cancer patients receiving chemotherapy and radiotherapy. The most common PNS manifestations of VZV infection include PHN, radiculopathy, zoster sine herpete (zoster without rash), and cranial nerve palsies. PHN can be diagnosed when pain persists more than 4 months after VZV eruption, and is more common in older adults. Vaccines are available for the prevention of chicken pox in children and for risk reduction of zoster and PHN in adults. Antiviral treatment is also available for acute attacks, prophylaxis against recurrent attacks, and decreasing the incidence of HSV.

Antidepressants, such as amitriptyline, desipramine, and nortriptyline, may relieve pain associated with PHN, but the side effects limit their use. Anticonvulsants are of limited value in the management of neuritic pain. Opioids are the most effective agents for severe neuritic pain. In addition, topical lidocaine 5 percent gel or patch, nerve blocks, skin infiltration, transcutaneous electrical nerve stimulation, IV lidocaine, and capsaicin dermal patch have shown to provide adjunctive relief in severe PHN.

SARCOID NEUROPATHY. Sarcoidosis is a multisystem granulomatous disorder that affects young adults and is manifested clinically with bilateral hilar adenopathy, pulmonary infiltration, and reticuloendothelial system, ocular, and cutaneous involvement. Approximately 5 percent of sarcoidosis patients develop nervous system complications, and of these, the PNS is affected in approximately 85 percent of patients. Sarcoid neuropathy can affect large nerve fibers or small nerve fibers. Large fiber neuropathy may present as cranial neuropathy, mononeuropathy multiplex, radiculopathy, or symmetric polyneuropathy. Cranial neuropathy is most common and often involves the facial nerve, which may be bilateral. The optic nerve may be affected and hearing loss is not uncommon. Intradural and extradural granulomatous masses may cause radiculopathy. The diagnosis of sarcoid neuropathy is straightforward when accompanied by recognized systemic involvement. In spite of their potential side effects, corticosteroids remain the mainstay of therapy.

NEUROMUSCULAR JUNCTION DISORDERS

Myasthenia Gravis

Myasthenia gravis (MG) is the most common disease of the neuromuscular junction. MG is an autoimmune disease, with the target of autoantibodies directed against postsynaptic ACh receptors on the muscle membrane. Two major clinical syndromes are ocular myasthenia and generalized myasthenia. Most frequently, cranial nerves are affected initially, especially extraocular muscles presenting as fluctuating double vision, facial weakness, or difficulty swallowing. Ocular MG may progress to generalized MG, or remain an isolated syndrome.

Weakness in MG is classically effort dependent, so that weakness increases with exertion, and frequently worsens as the day progresses. MG severity also fluctuates over time with frequent remissions and exacerbations, and there are no objective findings such as muscle atrophy or altered reflexes. Perhaps for these reasons, patients with MG frequently are thought to have psychiatric or psychological conditions or even seemingly unlikely diagnoses such “psychogenic respiratory arrest” before the appropriate tests establish the diagnosis (e.g., elevated ACh receptor antibody levels, abnormal response to repetitive motor nerve stimulation,

or abnormal single-fiber electromyogram).

Just as the fatigable weakness of MG should not be casually attributed to psychiatric disorders, vague symptoms such as "tiredness" which can be seen in many psychiatric conditions, should not be considered a manifestation of MG. Patients with somatization disorder have been misdiagnosed as MG sometimes with disastrous outcomes.

Dysphagia is a prominent and sometimes isolated symptom in early MG. Dysphagia is a nonspecific symptom and frequently associated with many anxiety disorders, and several figures-of-speech recognize this connection between anxiety and difficulty swallowing ("I choked," "I had a lump in my throat," "I can't swallow that"). The accompanying facial weakness characterized by intermittent ptosis, disconjugate gaze, and inability to express facial expression is seemingly inconsequential compared to the life-threatening components of MG. However, for many patients, these cosmetic and communicative features are far more disconcerting, often reaching pathological levels and requiring psychiatric support. Myasthenic patients demonstrate high rates of anxiety symptoms when followed longitudinally through the course of their illnesses, especially anxiety symptoms focused on respiratory status. Research does not presently view these psychiatric symptoms as generated directly by the neurobiology of the disease.

Acute episodes of worsening (myasthenia crises) can be life-threatening if they involve respiratory paralysis or dysphagia with aspiration. MG treatment involves use of anticholinesterase medications to prolong the availability of ACh in the neuromuscular junction, increasing the likelihood that the transmitter will find an available ACh receptor, and use of long-term immunosuppressant medications, most commonly corticosteroids, with their attendant adverse behavioral side effects. Many medications can exacerbate weakness in MG, and clinicians should always check new medications to determine if they may worsen neuromuscular transmission in patients with MG.

Lambert–Eaton Myasthenic Syndrome

Lambert–Eaton myasthenic syndrome (LEMS) is sometimes confused with either MG, chronic fatigue syndrome, or fibromyalgia. LEMS is an autoimmune, paraneoplastic syndrome that results in progressive weakness frequently with autonomic symptoms, and may progress to respiratory insufficiency. The majority of LEMS patients have a circulating voltage-gated calcium channel antibody. LEMS may be a presenting symptom of a neoplasm, with small cell lung cancer and breast cancers most common.

MUSCLE DISEASES

Myotonic Dystrophy

Myotonic dystrophy, the most frequent *adult* muscular dystrophy, is often easily recognized in adults by the characteristic face, muscle myotonia with difficulty releasing the hand grip for example, or percussion myotonia, and distal weakness. Myotonic dystrophy occurs 15 of 100,000 live births, and is autosomal dominant with incomplete penetrance; heterozygotes may present at any age. The genetic basis for myotonic dystrophy type I is an expanded trinucleotide repeat in a protein kinase gene on chromosome 19. The rarer DM type 2 presents with proximal weakness, and maps to chromosome 3. *Anticipation* may occur, with the development of more severe symptoms in successive generations due to expansion of the length of trinucleotide repeat between generations. Systemic manifestations of myotonic dystrophy include cataracts, diabetes, temporal baldness, and testicular atrophy. Cardiac conduction defects, cardiomyopathy, and potentially fatal arrhythmias may occur and should be treated aggressively. Respiratory muscle weakness may develop later in the illness and is often the cause of death.

The interference in ribonucleic acid (RNA) processing resulting from expanded trinucleotide repeats accounts for the peripheral manifestations and the behavioral syndromes that appear predictably among patients with the disease. Cognitive impairment of variable severity occurs in patients with myotonic dystrophy. The pattern of cognitive impairment most commonly includes personality change, alterations of visuospatial and constructional abilities, and prefrontal or executive cognitive domains. The severity of cognitive impairment does not correlate with age or severity of muscle disease, but correlates most closely with the length of the trinucleotide repeat expansion in the kinase gene. The cognitive impairment syndrome may remain stable over a prolonged period. Along with cognitive dysfunction, a range of personality traits from avoidant to obsessive–compulsive and passive–aggressive tendencies may be found.

Patients with myotonic dystrophy frequently complain of excessive daily sleepiness (hypersomnia), sleeping 12 or more hours per day. They may have hypnagogic hallucinations upon awakening and may resemble narcolepsy. As many as 39 percent of myotonic dystrophy patients have such symptoms, and there is a positive correlation between complaints of hypersomnia and functional disability due to muscular impairment. Central and obstructive sleep apnea with chronic hypercapnia may further complicate excessive daytime somnolence in myotonic dystrophy patients. The neurophysiological basis of this complaint is somewhat similar to that found in narcolepsy, with reports of decreased hypocretin-1 in cerebrospinal fluid, suggestive of dysfunction of the hypothalamic hypocretin system. Therapy of hypersomnolence has generally relied on stimulants such as modafinil (Provigil), or methylphenidate 20 mg daily.

Dystrophinopathies

Different mutations in the dystrophin gene on the X chromosome produce Duchenne muscular dystrophy (DMD) and the milder Becker muscular dystrophy phenotype. Dystrophin is one component of a glycoprotein complex located in the sarcolemmal membrane of muscle fibers. In DMD, there is a virtual absence of dystrophin, while Becker dystrophy has a partial deficiency of dystrophin. Since dystrophin is X-linked, males are predominantly affected while female carriers are typically asymptomatic. DMD affects 1 of 3,500 live male births. Up to one-third of cases of DMD are sporadic.

Symptoms in DMD start with proximal muscle weakness and calf muscle hypertrophy developing at 3 to 4 years of age. Weakness is progressive and becomes more diffuse, with loss of ambulation by 12 years of age and death by the age of 20 due to respiratory failure. In Becker dystrophy, the clinical syndrome is less severe, with symptomatic weakness beginning between ages 5 and 15. Loss of ambulation typically occurs after 15 years of age, and death often occurs in the third or fourth decade. A mild form of Becker dystrophy may present in middle age with myalgias and exertional muscle cramping. Either disorder may be associated with a symptomatic or asymptomatic cardiomyopathy.

Although usually considered an isolated muscle disease, a variety of cognitive and behavioral syndromes among patients with dystrophinopathies that may be directly related to the neurobiology of the diseases. The dystrophin gene product is expressed in brain as well as in muscle, although its CNS function is unknown.

A range of depressive and even psychotic syndromes has been reported. Approximately 30 percent of Duchenne patients have cognitive impairments by comparison with age-matched controls, but also by comparison with matched subjects with other spinal muscular disorders. Verbal function is more affected than visuospatial functions. Prior studies reported an average intelligence quotient (IQ) of 85 in Duchenne patients. The cognitive defect is usually nonprogressive.

Glucocorticoids are the only known pharmacological agents that can change the natural history of DMD, prolonging independent ambulation, preserving cardiac and pulmonary functioning in the teen years, and preventing severe scoliosis. However, side effects may include mood swings and difficult behavior. While anecdotal accounts suggest that these negative reactions may improve over time, they continue to remain one of the primary reasons for discontinuation of therapy. The impact of steroids on learning and memory is currently unclear. Acute steroid treatment has been shown to result in reversible impairment of long-term memory functions and retrieval of negative words, suggesting that emotional material is especially sensitive to the memory modulating effects of

steroids. Some research indicates that glucocorticoid-associated memory effects may be dose dependent. Other available treatments of DMD-associated cognitive symptoms in DMD is currently limited to special education programs and may not be the primary focus of those treating an affected child as the disease progresses. No data are available concerning the use of cognitive enhancers in these disorders.

Inflammatory Myopathies

There are three major syndromes of inflammatory myopathies: Polymyositis (PM), dermatomyositis (DM) and inclusion body myositis (IBM). All may be associated with autoimmune disorders such as SLE, and DM in particular may be associated with occult neoplasms, suggesting a paraneoplastic syndrome. Neuropsychiatric complications may occur secondary to associated autoimmune disorder. The weakness in PM/DM is usually symmetric and proximal, whereas weakness in IBM may be asymmetric and distal, with prominent early involvement of quadriceps, finger and wrist flexors, and ankle dorsiflexors. The weakness is usually slowly progressive, has elevated serum creatine phosphokinase (CPK), and many patients are not diagnosed for years after onset. PM and DM are treated with immunosuppressant agents, with their attendant neuropsychiatric risks, whereas IBM responds less well to immunosuppression.

Inclusion Body Myositis. IBM is a disease of adults, most commonly over age 50 years. It affects men more than women. The nature of the inclusion bodies, and the finding of β-amyloid in muscle fibers has led to theories that IBM may be a degenerative disorder of muscle (with homologies to Alzheimer disease) or possibly of viral origin. Rarely, IBM may be familial. Familial IBM associated with Paget disease of bone, and frontotemporal dementia, is associated valosin-containing protein (VCP) gene mutations. Neuropsychiatric features of this syndrome may include behavioral variant FTD, primary progressive aphasia, and MND.

Mitochondrial Myopathies

Muscle tissue, unlike peripheral nerve, is very energetically active, and from childhood onward, a number of myopathies are the so-called mitochondrial myopathies. Mitochondria are the “powerhouse of the cell” and contain their own DNA, which is uniquely transmitted by the mother (since sperm do not contribute mitochondria), so family histories show no male-to-male transmission. However, many mitochondrial proteins are of nuclear origin, and show classic Mendelian inheritance, so not all mitochondrial myopathies are due to intrinsic mitochondrial gene mutations. Mitochondrial disorders affect bioenergetic processes with

dysfunction in other body systems, such as developmental short stature, learning deficits, visual loss and blindness, eye movement abnormalities, deafness, cardiomyopathy and conduction defects, diabetes, liver and kidney problems, GI motility defects, and stroke-like symptoms.

Psychiatric features are found within the clinical spectrum of mitochondrial syndromes. Case reports have linked depression and obsessive–compulsive disorder to specific mitochondrial mutations. It is increasingly recognized that mitochondrial dysfunction may be associated with neuropsychiatric abnormalities such as dementia, major depression, and bipolar disorder. Furthermore, several lines of evidence suggest that there is involvement of mitochondrial dysfunction in schizophrenia, including documented alterations in brain energy metabolism, electron transport chain activity, and expression of genes involved in mitochondrial function.

Endocrine Myopathies

Endocrine myopathies include weakness due to hormonal imbalance in the pituitary, adrenal, thyroid, or parathyroid glands. Examples of this class of muscles disorders include steroid myopathy with weakness caused by either endogenous or exogenous glucocorticoid; psychiatric co-morbidities of mood or thought disturbance are common. Adrenal insufficiency is associated with generalized weakness and fatigue. Treatment is aimed at normalizing hormonal levels or removal of toxic compounds as tolerated, while working with the patient and medical team to optimize quality of life.

MISCELLANEOUS NEUROMUSCULAR DISORDERS

Stiff-Person Syndrome

Stiff-person syndrome is a rare disorder of the CNS, characterized as having progressive and fluctuating muscle rigidity with painful spasms. The spasms may occur at rest and may be triggered by sensory stimuli. The etiology of Stiff-Person syndrome is autoimmune, and is often associated with Type 1 Diabetes, and with dysfunction in the GABA inhibitory system. A majority of patients have autoantibodies to glutamic acid decarboxylase (GAD).

Psychiatric symptoms in this disorder are prominent anxiety, probably mediated by changes in the GABAergic system. Treatments used include diazepam up to 100 mg/day or clonazepam up to 6 mg/day. Other GABAergic mediating drugs such as Vigabatrin, Tiagabine, and Valproic acid (the latter up to 3,000 mg/day in divided doses) may also be helpful.

Neuroleptic Malignant Syndrome

Neuroleptic malignant syndrome (NMS) is usually conceived of as a syndrome of fever, catatonic rigidity, with increased serum CPK levels in

the setting of neuroleptic administration. The etiology of this disorder is obscure, and lack of consistent case definitions have hampered fuller study of this syndrome. The initial description of NMS referred to it as pallor and hyperthermia, and isolated primary symptoms of fever, akinesia or stupor with hypertonicity, and pulmonary congestion. Since then a number of diagnostic criteria have been proposed. In essence, the syndromic definitions stress the presence of rigidity, fever, and autonomic changes in the presence of neuroleptic exposure. Alteration in sensorium is also a common feature of several of these criteria. Conditions with clinical similarity to NMS include serotonin syndrome (SS), malignant hyperthermia, drug toxicity (e.g., lithium), heat stroke, and related syndromes.

Serotonin Syndrome

SS occurs due to excessive serotonergic activity in the CNS and PNS, usually in the context of initiation or dose increase of a serotonergic agent. Although this may occur with a single agent, it is usually the result of a combination of medications, supplements, illicit drugs, or foods that are proserotonergic. The symptoms fall under three broad categories—altered mental status, neuromuscular hyperactivity, and autonomic instability. Altered mental status may range from hyperactivity, anxiety, restlessness, mood elevation, and excitement to lethargy, confusion, disorientation, hallucinations, and coma. Neuromuscular hyperactivity can manifest as myoclonus, hypertonicity, ocular clonus, hyperreflexia, and tremors. Autonomic instability may include diaphoresis, tachycardia, hypertension, and in moderate or severe cases, hyperthermia. In severe cases, symptoms can progress to generalized tonic–clonic seizures, multiorgan failure, and coma. Although clinical criteria such as Sternbach criteria and Hunter serotonin toxicity criteria exist, there is not much concordance among these criteria. A heightened awareness of signs and symptoms, obtaining a detailed medication history and conducting a careful physical examination are crucial for diagnostic discrimination of SS and other similar syndromes such as NMS, malignant hyperthermia, and anticholinergic toxicity. Treatment of SS involves cessation of the offending agent, supportive therapy, and use of the serotonin antagonist cyproheptadine. Given the widespread use of serotonergic antidepressants and the large number of substances that have been implicated in SS, it is of paramount importance that clinicians thoroughly evaluate the indication for adding a serotonergic agent to a medication regimen that might already contain serotonergic agent(s), and monitor those who are at increased risk closely for signs and symptoms of SS.

Fibromyalgia

Fibromyalgia is a syndrome characterized by chronic widespread pain localized in various muscle groups, accompanied by fatigue, sleep disturbances, mood changes, and cognitive deficits. It is thought of as a centralized pain state in that it involves central amplification of peripheral nociceptive stimuli. It is more common in females in their thirties to fifties. The provisional diagnostic criteria for fibromyalgia published by the American College of Rheumatology in 2010 is shown in Table 2.12–3.

There is significant overlap and comorbidity between symptoms of fibromyalgia and psychiatric disorders such as depression and anxiety. The shared putative mechanisms underlying depression, anxiety, and pain, namely dysregulation of serotonergic, noradrenergic, and dopaminergic neurotransmitter systems, may explain the co-occurrence of these disorders. This has important treatment implications—antidepressants that modulate the serotonin and norepinephrine neurotransmitter systems have been found to be effective in treating symptoms of fibromyalgia; furthermore, the secondary psychiatric syndromes can provide therapeutic targets for psychopharmacology. There is also significant comorbidity between patients with fibromyalgia and rheumatologic disorders, such as rheumatoid arthritis, systemic lupus, and others.

There is significant disability associated with fibromyalgia, with 35 percent of patients receiving Social Security disability income in one study. Although treatment with multiple modalities may lead to improvements in symptoms and quality of life, their impact on functional recovery is limited. Duloxetine, milnacipran, and pregabalin have been approved by the U.S. Food and Drug Administration for treatment of fibromyalgia. Duloxetine and milnacipran are both SNRIs that mediate antinociceptive effect through modulation of serotonin and norepinephrine in the synaptic cleft whereas pregabalin modulates neuronal calcium influx by binding to α_2-δ subunit of voltage-gated calcium channels. Duloxetine is also approved for other psychiatric and pain conditions, while milnacipran is only approved for treatment of symptoms in fibromyalgia. Other pharmacological agents that have been shown to improve symptoms include tricyclic compounds such as amitriptyline, nortriptyline, and cyclobenzaprine, and selective serotonin reuptake inhibitors. The choice of medication should be based on targeting predominant symptoms and side effect profile. Nonpharmacologic treatment plans include graded exercise regimens, rehabilitation programs, cognitive–behavioral therapy, and patient education.

Table 2.12–3.
2010 Fibromyalgia Diagnostic Criteria

1. A patient satisfies diagnostic criteria for fibromyalgia if the following 3 conditions are met: Widespread pain index (WPI) ≥7 and symptom severity (SS) scale score ≥5 or WPI 3 to 6 and SS scale score ≥9.

Symptoms have been present at a similar level for at least 3 months. The patient does not have a disorder that would otherwise explain the pain.			
WPI: note the number areas in which the patient has had pain over the last week. In how many areas has the patient had pain? Score will be between 0 and 19.	Shoulder girdle, left Shoulder girdle, right Upper arm, left Upper arm, right Lower arm, left Lower arm, right	Hip (buttock, trochanter), left Hip (buttock, trochanter), right Upper leg, left Upper leg, right Lower leg, left Lower leg, right	Upper back Lower back Neck Jaw, left Jaw, right Chest Abdomen
1. SS scale score: Fatigue Waking unrefreshed Cognitive symptoms	For each of the 3 symptoms, indicate the level of severity over the past week using the following scale: 0 = no problem 1 = slight or mild problems, generally mild or intermittent 2 = moderate, considerable problems, often present and/or at a moderate level 3 = severe: pervasive, continuous, life-disturbing problems	Considering somatic symptoms in general, indicate whether the patient has:[a] 0 = no symptoms 1 = few symptoms 2 = a moderate number of symptoms 3 = a great deal of symptoms	The SS scale score is the sum of the severity of the 3 symptoms (fatigue, waking unrefreshed, cognitive symptoms) plus the extent (severity) of somatic symptoms in general. The final score is between 0 and 12.

[a]Somatic symptoms that might be considered: muscle pain, irritable bowel syndrome, fatigue/tiredness, thinking or remembering problem, muscle weakness, headache, pain/cramps in the abdomen, numbness/tingling, dizziness, insomnia, depression, constipation, pain in the upper abdomen, nausea, nervousness, chest pain, blurred vision, fever, diarrhea, dry mouth, itching, wheezing, Raynaud phenomenon, hives/welts, ringing in ears, vomiting, heartburn, oral ulcers, loss of/change in taste, seizures, dry eyes, shortness of breath, loss of appetite, rash, sun sensitivity, hearing difficulties, easy bruising, hair loss, frequent urination, painful urination, and bladder spasms

Adapted from Wolfe F, Clauw DJ, Fitzcharles MA, et al. Arthritis Care Res (Hoboken). 2010 May;62(5):600–610.

FUTURE DIRECTIONS

Advances in the understanding of the molecular basis of many neuromuscular diseases and their relationship to neuropsychiatric

symptoms has ushered in a new era in psychiatry. The natural history and prognosis of these disorders is generally well understood; yet even in less well-understood conditions such as fibromyalgia, the focus should be on available clinical evidence and build excellent physician–patient relationships, a key factor in providing improved neuropsychiatric outcomes. Further advances and insights into the integrated neurobiology of these disorders raise hope that psychiatrists will find themselves intimately involved in all phases of care targeting these challenging disorders.

REFERENCES

Al-Allaf AW. Work disability and health system utilization in patients with fibromyalgia syndrome. *J Clin Rheumatol*. 2007;13:199.

Arnold LM, Hudson JI, Keck PE, Auchenbach MB, Javaras KN, Hess EV. Comorbidity of fibromyalgia and psychiatric disorders. *J Clin Psychiatry*. 2006;67:1219.

Bennion Callister J, Pickering-Brown SM. Pathogenesis/genetics of frontotemporal dementia and how it relates to ALS. *Exp Neurol*. 2014;262:84.

Boyer FC, Tiffreau V, Rapin A, et al. Post-polio syndrome: Pathophysiological hypotheses, diagnosis criteria, drug therapy. Ann Phys Rehabil Med. 2010;53:34.

*Clauw DJ. Fibromyalgia: a clinical review. *JAMA*. 2014;311:1547.

Cohen JI. Clinical practice: Herpes zoster. N Engl J Med. 2013;369(3):255.

Crofford LJ. Psychological aspects of chronic musculoskeletal pain. Best Pract Res Clin Rheumatol. 2015;29:147.

da Rocha AJ. Granulomatous diseases of the central nervous system. *Top Magn Reson Imaging*. 2005;16:155.

Davies R, Ahmed G, Freer T. Chronic exposure to organophosphates. *APT*. 2000;6:187.

Dworkin RH, Johnson RW, Breuer J, et al. Recommendations for the management of herpes zoster. *Clin Infect Dis*. 2007;44(Suppl 1):S1.

*Dworkin RH, O'Connor AB, Backonja M, et al. Pharmacologic management of neuropathic pain: evidence-based recommendations. *Pain*. 2007;132:237.

Hendriksen JG, Vles JS. Neuropsychiatric disorders in males with Duchenne muscular dystrophy: frequency rate of attention-deficit/hyperactive disorder (ADHD), autism spectrum disorder, and obsessive-compulsive disorder. *J Child Neurol*. 2008;23:477.

Ickmans K, Meeus M, De Kooning M, et al. Associations between cognitive performance and pain in chronic fatigue syndrome: comorbidity with fibromyalgia does matter. Pain Physician. 2015;18:E841.

Iqbal MM, Basil MJ, Kaplan J, Iqbal MT. Overview of serotonin syndrome. *Ann Clin Psychiatry*. 2012;24:310.

Kemp BJ, Adams BM, Campbell ML. Depression and life satisfaction in aging polio survivors versus age-matched controls: relation to postpolio syndrome, family functioning, and attitude toward disability. *Arch Phys Med Rehabil*. 1997;78:187.

Kluger BM, Krupp LB, Enoka RM. Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy. *Neurology*. 2013;80:409.

Kulaksizoglu IB. Mood and anxiety disorders in patients with myasthenia gravis. *CNS Drugs*. 2007;21:473.

Leung DG, Wagner KR. Therapeutic advances in muscular dystrophy. *Ann Neurol*. 2013;74:404.

*Meola G. Clinical aspects, molecular pathomechanisms and management of myotonic dystrophies. *Acta Myol*. 2013;32:154.

National Clinical Guideline Centre (UK). Motor Neurone Disease: Assessment and Management. London: National Institute for Health and Care Excellence (UK); 2016 Feb.

Orisakwe OE. The role of lead and cadmium in psychiatry. *N Am J Med Sci*. 2014;6:370.

Ringholz GM, Appel SH, Bradshaw M, Cooke NA, Mosnik DM. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology*. 2005;65:586.

*Serra A, Ruff RL, Leigh RJ. Neuromuscular transmission failure in myasthenia gravis: decrement of safety factor and susceptibility of extraocular muscles. *Ann N Y Acad Sci*. 2012;1275:129.

Sharshar TI, Polito A, Porcher R, et al. Relevance of anxiety in clinical practice of Guillain-Barre syndrome: a cohort study. *BMJ Open*. 2012;2(4).

Shu LR, Tsai SJ. Long-term glue sniffing: report of six cases. *Int J Psychiatry Med*. 2003;33:163.

Weiss H, Rastan V, Müllges W, Wagner RF, Toyka KV. Psychotic symptoms and emotional distress in patients with Guillain-Barré syndrome. *Eur Neurol*. 2002;47:74.

*Wolfe F, Clauw DJ, Fitzcharles MA, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. *Arthritis Care Res (Hoboken)*. 2010;62:600.

Wolfe F, Walitt BT, Katz RS, Häuser W. Social security work disability and its predictors in patients with fibromyalgia. *Arthritis Care Res (Hoboken)*. 2014;66:1354.

Zimmerman EK, Eslinger PJ, Simmons Z, Barrett AM. Emotional perception deficits in amyotrophic lateral sclerosis. *Cogn Behav Neurol*. 2007;20:79.

▲ 2.13 Psychiatric Aspects of Child Neurology

JEFFREY E. MAX, M.B.B.CH., SARAH N. MATTSON, PH.D., YVONNE ELEANOR VAUCHER, M.D., SHARON L. NICHOLS, PH.D., AND MARK P. NESPECA, M.D.

The topic "Psychiatric Aspects of Childhood Neurology" is large and for justice to be done, a textbook is required. However, the approach taken in this chapter was to succinctly review five common and important pediatric neurological conditions rather than review a much higher number of related topics in a cursory fashion. The topics (primary authors) chosen for this chapter are (1) Fetal Alcohol Spectrum Disorders (Sarah N. Mattson, Ph.D.); (2) Premature Birth and Low Birth Weight (Yvonne E. Vaucher, M.D.); (3) Human Immunodeficiency Virus (Sharon Nichols, Ph.D.); (4) Epilepsy (Mark P. Nespeca, M.D.); and (5) Traumatic Brain Injury (Jeffrey E. Max, MBBCh). The focus of the chapter is on psychopathological sequelae although associated problems in other domains of function will also be mentioned.

FETAL ALCOHOL SPECTRUM DISORDERS

Children born to women who consume alcohol during pregnancy are at risk for physical, behavioral, and cognitive changes that can affect everyday function. These functional effects are collectively termed *fetal alcohol spectrum disorders (FASDs)*. The US prevalence of FASDs is between 2.4 and 4.8 percent, though estimates vary based on sampling techniques and population of study.

Diagnostic Criteria

At the severe end of the spectrum, heavy prenatal alcohol exposure can cause characteristic facial features, resulting in a diagnosis of fetal alcohol syndrome (FAS). The diagnostic criteria for FAS are threefold: (1) characteristic facial features (two or more of the following: short palpebral fissures, smooth philtrum, thin vermillion); (2) pre- and/or postnatal growth deficiency (height or weight ≤10th percentile); and (3) evidence of deficient brain growth or abnormal morphogenesis (structural brain abnormality or head circumference ≤10th percentile). A diagnosis of partial FAS requires the same facial features in addition to one or more of the following: (1) pre- and/or postnatal growth deficiency; (2) evidence of deficient brain growth or abnormal morphogenesis; or (3) a complex pattern of behavioral or cognitive abnormalities. A diagnosis of FAS or partial FAS can be made with or without confirmation of maternal alcohol use in pregnancy. The original Institute of Medicine (IOM) report and revision by Hoyme and colleagues (2005) also include diagnostic criteria for alcohol-related birth defects (ARBD) and alcohol-related neurodevelopmental disorder (ARND), both of which require confirmation of maternal alcohol exposure. The diagnostic criteria for ARND require evidence of structural brain abnormality (brain growth or abnormal morphogenesis), as defined above and/or a complex pattern of behavioral or cognitive abnormalities. ARND is not precisely defined, but is a critical concept as the majority of children born to alcoholic women do not have enough or even any of the physical features required for an FAS or partial FAS diagnosis. ARND is intended to be used for children with normal growth and structural development and thus could be used to capture the broader population of children affected by prenatal alcohol exposure. IOM diagnostic criteria are summarized in Table 2.13–1.

Neurobehavioral Disorder Associated with Prenatal Alcohol Exposure (ND-PAE). To address the critical need to better identify children affected by prenatal alcohol exposure and the lack of precise criteria to do so, the most recent revision of the Diagnostic and Statistical Manual includes ND-PAE in the appendix under Conditions for Further Study. There are four main criteria for ND-PAE. The first criterion is that more than minimal prenatal

alcohol exposure is documented. Documentation can be via direct maternal report or through other reliable sources. The second criterion is impaired neurocognition supported by impairment in one or more specified domain (global intellectual performance, executive function, learning, memory, visual–spatial reasoning). The third criterion is impaired self-regulation in one or more specified domain (mood or behavioral regulation, attention, impulse control). The fourth criterion is impaired adaptive function and requires two or more impaired domains (communication, social communication, daily living skills, motor skills). One of the impairments must be in communication or social communication. As with other disorders specified in the DSM-5, the onset must be in childhood, impairment must cause distress or impairment, and it is not better explained by other factors. As proposed, ND-PAE fills a diagnostic gap by providing specific criteria for the neurobehavioral and cognitive effects seen in children affected by prenatal alcohol exposure. The specification of these criteria enables a clinical diagnosis for these children, which can improve identification, aid in treatment planning, as well as provide etiological context for behavioral concerns. Importantly the diagnosis of ND-PAE is independent from the diagnosis of FAS (i.e., can occur in exposed children with or without the FAS) (Table 2.13–2).

Table 2.13–1.
Summary of Criteria for Disorders Included in Fetal Alcohol Spectrum Disorders (FASD)

	Confirmed Exposure to Alcohol	Facial Anomalies	Growth Retardation	CNS Structural Abnormalities	CNS Functional Abnormalities (Cognitive/Behavioral
FAS[a]	Not required	Required	Required	Required	Not required
Partial FAS[a]	Not required	Required	1 or more required	Not required	—
Alcohol-related birth defects (ARBD)[a]	Required	Required	Not required	Not required	Not required
Alcohol-related neurodevelopmental disorder (ARND)[a]	Required	Not required	Not required	1 or more required	Not required
Neurobehavioral disorder associated with prenatal alcohol exposure (ND-PAE)[b]	Required	Not required	Not required	Not required	Required (see Table X2)

[a]From Institute of Medicine and Hoyme, et al.

[b]From the Diagnostic and Statistical Manual, 5th Edition.

Table 2.13–2.

Diagnostic Criteria for Neurobehavioral Disorder Associated with Prenatal Alcohol Exposure (ND-PAE) from DSM-5

A. More than minimal exposure to alcohol during gestation, including prior to pregnancy recognition. Confirmation of gestational exposure to alcohol may be obtained from maternal self-report of alcohol use in pregnancy, medical or other records, or clinical observation.
B. Impaired neurocognitive functioning as manifested by one or more of the following:
   1. Impairment in global intellectual performance
   2. Impairment in executive functioning
   3. Impairment in learning
   4. Memory impairment
   5. Impairment in visual–spatial reasoning
C. Impaired self-regulation as manifested by one or more of the following:
   1. Impairment in mood or behavioral regulation
   2. Attention deficit
   3. Impairment in impulse control
D. Impairment in adaptive functioning as manifested by two or more of the following, one of which must be (1) or (2):
   1. Communication deficit
   2. Impairment in social communication and interaction
   3. Impairment in daily living skills
   4. Impairment in motor skills
E. Onset of the disorder (symptoms in Criteria B, C, and D) occurs in childhood.
F. The disturbance causes clinically significant distress or impairment in social, academic, occupational, or other important areas of functioning.
G. The disorder is not better explained by the direct physiological effects associated with postnatal use of a substance (e.g., a medication, alcohol, or other drugs), a general medical condition (e.g., traumatic brain injury, delirium, dementia), another known teratogen (e.g., fetal hydantoin syndrome), a genetic condition (e.g., Williams syndrome, Down syndrome, Cornelia de Lange syndrome), or environmental neglect.

Behavioral Findings

Disordered behavior is characteristic of FASDs and may represent the most clinically relevant deficit. Parent reports of behavioral disturbance are pervasive and severe. Many studies indicate the presence of behavioral problems in children with FASDs including issues with impulsivity, rule-breaking, anxiety, and social interactions, which persist even when compared to IQ-matched controls. Adaptive behavior deficits have also been noted across all domains of adaptive function: communication, socialization, and daily living skills. In particular, deficits in socialization and communication appear to worsen with age. As a result, behavioral problems in childhood may translate into difficulties in adulthood, including legal problems, substance abuse, psychiatric disorders, and elevated suicide risk.

Cognitive Findings

Cognitive effects of prenatal alcohol exposure are well documented. Deficits occur in multiple domains including general intellectual function, verbal

learning, executive function, language, visual–spatial, attention, and academic function. Children with FASDs have a broad range of IQ scores, ranging from severely intellectually deficient to average, although the majority of children do not have IQ scores below 70. The average IQ score for children with FAS is around 72, while that for children with heavy prenatal alcohol exposure without FAS is somewhat higher at around 80, although there is considerable variability in both groups.

In addition to overall intellectual performance, children with FASDs also display a broad array of neuropsychological and cognitive deficits. Verbal learning and memory is consistently impaired in this population, with deficits occurring predominantly in encoding with relatively intact retention. Repeated exposure and implicit strategy aid in improving performance on learning and memory tasks. Nonverbal learning and memory is less well studied, but it appears that deficits occur in both encoding and retention. Even less is known about visual–spatial abilities. Deficits have been noted on both simple and complex visual–spatial tasks, although visual–spatial perception may account for some of these deficits.

Attention is a hallmark deficit of FASDs. Children in this population typically display deficits in vigilance, reaction time, and information processing across multiple tasks, although visual attention may be more impacted than auditory attention. Deficits in attention are supported by both laboratory measures and parent report. Other aspects of executive function are also impacted in children with FASDs. This domain has been particularly well-studied and deficits occur in planning, working memory, response inhibition, set-shifting, concept formation, and verbal and nonverbal fluency. Deficits in executive function appear especially useful in identification of alcohol-affected children.

Academic achievement is one domain that requires further study. Increased school failure has been documented and limited research indicates deficits in reading, spelling, and mathematic achievement. Recent studies indicated that deficits in reading and spelling may relate to working memory deficits and deficits in mathematics may relate to spatial processing.

Comorbidity

Children with FASDs are often diagnosed with psychiatric disorders, and have elevated rates of concurrent psychiatric disorders, especially in younger children and boys. Across multiple studies, the most common psychiatric diagnosis made in children with FASDs is attention-deficit/hyperactivity disorder (ADHD); the prevalence of ADHD in this population ranges from 44 to 96 percent. Other than ADHD, the most common conditions are oppositional defiant disorder (ODD), conduct disorder, and depressive and anxiety disorders.

Because ADHD is so common in children with FASDs, several studies

have compared children with FASDs to children with idiopathic ADHD (i.e., without FASDs or prenatal alcohol exposure). These studies have examined adaptive behavior, psychopathology, executive function, verbal learning and memory, attention, and neuropsychological functioning. Although some similarities between FASDs and ADHD have been noted, for example, on the Wisconsin Card Sorting Test and parent reports of attention, more studies show differences between these clinical groups. For example, group differences have been noted on overall IQ, verbal fluency, trail making, verbal encoding, attention, and adaptive behavior; generally children with FASDs perform worse than children with idiopathic ADHD. On measures of attention, children with FASDs show deficits in shifting, encoding information, and problem-solving while those with ADHD have more difficulty focusing and sustaining attention, although differences between groups may also relate to timing of stimuli and regulation of activation. On parent reports of adaptive function, children with FASDs, but not idiopathic ADHD, show arrested development of socialization and communication skills, failing to improve with age. Deficits in social cognition and facial emotion processing are also worse in children with FASDs than in idiopathic ADHD. Thus, overall, it appears that the cognitive–behavioral profiles of FASDs show more impairment than what is seen in ADHD.

Another important question is whether the presence of ADHD in many children with FASDs affects behavioral or neuropsychological outcomes. That is, do children with FASDs and ADHD have greater impairment than either condition alone? Three recent studies compared children with FASDs with and without ADHD, children with idiopathic ADHD, and children with neither ADHD nor prenatal alcohol exposure. Both studies used a 2 (exposure: yes/no) × 2 (ADHD: yes/no) factorial design. In the first study, parent reports of psychopathology and problem behaviors were compared. Results suggested that the presence of ADHD in the children with FASDs elevated rates of conduct disorder and externalizing behaviors. A second study of adaptive behavior, using the same design, indicated that children with both ADHD and FASDs had the greatest impairment in the communication domain of adaptive behavior, while the effects in socialization and daily living skills suggested independent effects of both ADHD and FASDs. In contrast, the third study of this type examined a broad neuropsychological test battery and did not support an additive or synergistic effect of FASDs and ADHD on neuropsychological functioning. Instead, neuropsychological performance was similar in the alcohol-exposed groups, regardless of ADHD diagnosis.

Summary

Research conducted over the last 40 years has demonstrated that alcohol is a potent physical and behavioral teratogen. While some children will not be

affected by the physical manifestations of this exposure, all children with histories of prenatal alcohol exposure are at risk for cognitive and behavioral impairment at great personal and societal cost. Debate continues over whether any safe amounts or periods of alcohol exposure exist while empirical data clearly show that the effects of this exposure can be devastating. Very little is known about long-term outcomes for individuals with FASDs, and more research is needed in this critical area. Limited research suggests continued negative outcomes while some anecdotal reports suggest that more positive outcomes are possible.

PREMATURE BIRTH AND LOW BIRTH WEIGHT

INTRODUCTION

Each year in the United States, approximately 11 percent or 445,000 children are born prematurely between 23 and 37 weeks of gestation, of whom 55,000 are very low birth weight (VLBW), weighing less than 1,500 g at birth, and approximately 29,000 are extremely premature (EP), born before 28 weeks of gestation (Table 2.13–3).

More than 90 percent of these premature newborns, including the majority of EP children born at 24 weeks of gestation, now survive as the result of remarkable medical advances in neonatal care over the past 30 years. However, those born prematurely are at increased risk for adverse neurologic, cognitive, neurosensory, behavioral, and psychiatric outcomes (Table 2.13–4). The risk is multifactorial, influenced by numerous antenatal, intrapartum, and postnatal factors. Major independent determinants of adverse outcomes include gestational age, birth weight, gender, intrauterine growth restriction, severity and type of neonatal illness, neonatal brain injury, socioeconomic status, and maternal education. Barring significant brain injury, demographic and socioeconomic factors become important predictors of neurodevelopmental outcome over time.

Table 2.13–3.
Common Nomenclature for Preterm Infants

Weight	
Low birth weight <2,500 g	
Very low birth weight <1,500 g	
Extremely low birth weight <1,000 g	
Gestational age	
Late preterm	$34^{0/7}$ to $36^{6/7}$ weeks
Moderate preterm	$32^{0/7}$ to $33^{6/7}$ weeks
Very preterm	$<32^{0/7}$ weeks
Extremely preterm	$<28^{0/7}$ weeks

The risk of adverse outcomes increases progressively with increasing prematurity with the lowest risk for those born nearest term at 36 weeks of gestation and the highest risk for those born at the margins of viability at 22 to 23 weeks of gestation. Although survival has improved, the incidence of adverse neurologic, cognitive, behavioral, and psychiatric outcome has remained essentially unchanged over the past 60 years, raising the question of whether preterm birth carries an underlying risk by altering normal brain development. Preterm infants, especially extremely preterm born before 28 weeks of gestation, are born into a dramatically different vestibular, motor, and neurosensory extrauterine environment. The extent to which this influences subsequent brain development is unknown. For the preterm infant, all organ systems are physiologically immature and vulnerable to injury after birth. The resulting medical problems such as chronic lung disease, patent ductus arteriosus, sepsis, necrotizing enterocolitis, intracranial hemorrhage, white matter injury, and retinopathy of prematurity are all associated with adverse neurodevelopmental outcomes. At the time of their child's preterm birth, parents are concerned about the chance of severe motor or cognitive impairment which affect approximately 20 percent of all VLBW children. However, most neurodevelopmental problems—behavioral, emotional, and educational—are not apparent until early childhood and the school years. Despite advances in neuroimaging in the neonatal period, it is not possible to accurately predict the long-term neurodevelopmental outcome for an individual child.

Table 2.13–4.
Adverse Motor, Neurosensory, and Cognitive Outcomes of VLBW Children

Cerebral palsy
Developmental coordination disorder
Minor neurologic dysfunction
Visual impairment
Hearing impairment
Cognitive impairment
Speech and language delay
Perceptual motor dysfunction
Low academic achievement
Learning disability
Attention-deficit/hyperactivity disorder
Anxiety
Depression
Poor social interaction
Autism spectrum disorder

BRAIN DEVELOPMENT

Preterm infants are born from the end of the second through the third trimester when the brain is undergoing very rapid growth and organization. Between 25 weeks and term the cerebral cortical volume increases fourfold in size. At 35 weeks of gestation the brain is one-third smaller than it will be at term. Neurons proliferate primarily in the first trimester and migrate radially outward from their origin in the subependymal germinal matrix beneath the floor of the lateral ventricles to the cerebral cortex during the second trimester. A smaller wave of neuronal proliferation and migration is now known to occur in humans from 20 to 27 weeks of gestation, overlapping the time when the most immature preterm infants are critically ill and receive medical care in neonatal intensive care units. The third trimester is characterized by active neuronal organization and exponentially increasing synaptic connectivity which is dependent upon complex interrelationship, between structural relationships, functional activity, sensory input, growth factors, hormonal, and neurotransmitter activity. This is also a critical period, between 24 and 32 weeks of gestation, when transient subplate neurons are in place just beneath cerebral cortex waiting to connect with and guide late ascending neurons from the thalamus and commissural axons to the cortex. By 36 weeks of gestation, these thalamocortical axons, callosal, and corticocortical association afferents have reached their ultimate destination in the cortex.

Neuronal survival and function is dependent upon glia, the most common cell type in the brain. Glia determine and maintain structural and functional integrity of the neurons, mediate the inflammatory response, and facilitate neuronal communication. Specifically, astrocytes transfer nutrients from capillaries to neurons, modulate the strength of synaptic transmission, and participate in synaptic repair; oligodendrocytes are responsible for axonal myelination; and microglia mediate inflammatory response. Glia create the topographic boundaries and maintain the structure of the white matter projection, commissural and association axon tracts. Immature, preoligodendrocytes, predominate from 23 to 32 weeks of gestation and are extremely vulnerable to inflammatory, excitotoxic, and free radical injury due to hypoxemia, ischemia, hemorrhage, cytokines, and the toxic by-products of neuronal injury, all of which may result in acute glial cell injury or death as well as later apoptosis, delayed or abnormal maturation, and impaired function. Lack of glial support and hypomyelination results in neuronal injury, loss, and dysfunction. A reciprocal relationship between neuronal and glial injury has been proposed by Volpe whereby injury to the glial leads secondarily to injury to the neuron and vice versa. He proposes that this “encephalopathy of prematurity,” is responsible for adverse neurodevelopmental outcome.

FIGURE 2.13–1. Types of intracerebral hemorrhage. **A:** Grade 1 IVH (**left**), Grade 2 IVH (**right**). **B:** Grade 3 IVH (**right** and **left**); Grade 4/Intraparenchymal hemorrhage (**left**).

The cerebellum is also in a state of very active development between 24 and 40 weeks of gestation, increasing fivefold in volume and thirtyfold in surface area with peak cerebellar growth, differentiation, and neuronal migration occurring from 30 to 40 weeks of gestation. Hemorrhagic injury, arising in the cerebellar germinal matrix, is more often unilateral and asymmetric whereas ischemic infarction is more likely to be bilateral. The sequence of events is similar to that of cerebral injury—acute destruction results in focal or diffuse volume loss with impaired neuronal development and hypomyelination. Trophic interaction (crossed diaschisis) between the contralateral cerebellar and cerebral hemispheres is such that injury to one is associated with decreased volume in the other which may explain why cerebellar injury is associated with a wide range of cognitive, language, and motor deficits.

PRETERM BRAIN INJURY

Intracranial or intraventricular hemorrhage (IVH) now occurs in approximately 25 percent of all VLBW infants. The widespread use of antenatal steroids to accelerate lung maturity has had the additional benefit of reducing the incidence of intracranial hemorrhage by half. Hemorrhage occurs within the first few days after birth, originating from the fragile capillaries in subependymal germinal matrix just under the floor of the lateral ventricles at the caudothalamic groove. IVH is most often unilateral and when bilateral, usually asymmetric in size. The mildest form (Grade 1 IVH) is contained within the subependymal germinal matrix (Fig. 2.13–1). The hemorrhage may rupture through the floor into the lateral ventricle (Grade 2 IVH) and may become large enough to distend the lateral ventricles with blood (Grade 3 IVH). The most severe form of IVH occurs when a large subependymal hemorrhage occludes the terminal vein draining the adjacent periventricular white matter resulting in

intraparenchymal venous infarction (Grade 4 IVH, intraparenchymal hemorrhage [IPH] or periventricular hemorrhagic infarction [PVHI]). The associated white matter destruction may evolve into many small cysts, a large porencephalic cyst, or into an irregular outpouching of the ventricle into the area of white matter volume loss (Fig. 2.13–2). Blood within the ventricular system, especially with Grade 3 IVH, may result in posthemorrhagic hydrocephalus. The majority (~85 percent) of hemorrhages are Grades 1 or 2. Approximately 10 percent are Grade 3 and 5 percent are Grade 4 IVH/IPH. As expected, severe IVH/IPH (Grades 3 and 4) occurs most frequently in the smallest and most immature infants, affecting about 20 percent of those born at 23 to 24 weeks of gestation. Fortunately, Grades 1 and 2 IVH and are not associated with an increased risk of motor impairment. For infants with Grades 3 and 4 IVH/IPH, however, the risk of cerebral palsy is much higher—35 percent and 80 percent, respectively. Neurologic sequelae of IVH are primarily due to associated white matter damage although free iron may also contribute to cellular injury. Almost all hemorrhages, except the most massive, are clinically silent, being recognized only by cranial imaging. Except for those in the posterior fossa, hemorrhages are easily visualized by cranial ultrasound. It was not until the advent of MRI that the incidence of cerebellar hemorrhage, occurring in up to 15 percent of VLBW infants, was appreciated.

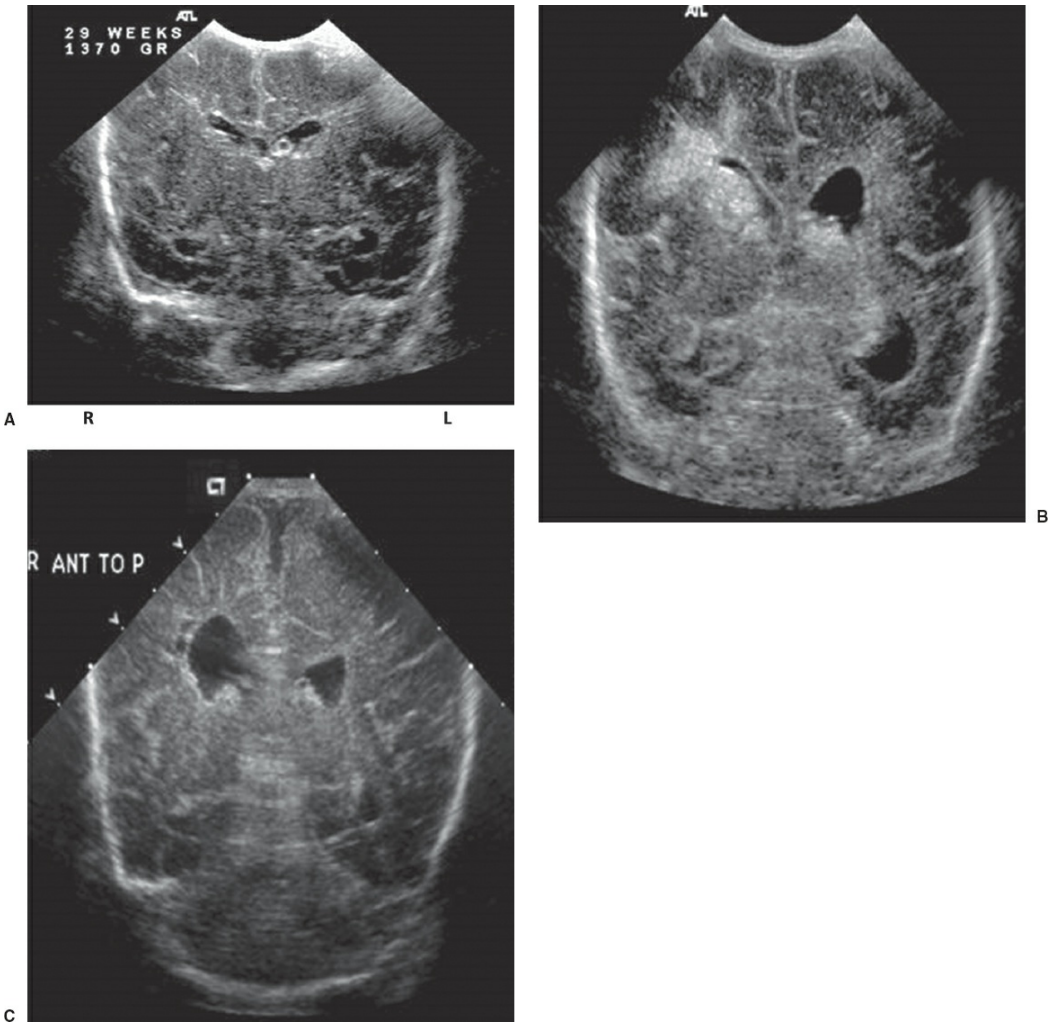

FIGURE 2.13–2 Evolving periventricular hemorrhagic infarction (IPH/PVHI). Coronal views. **A:** DOL 1: Small left Grade 1 (SEH); normal ventricular size. **B:** DOL 6: Large right Grade 1 (SEH), small left Grade 1 IVH, bilateral Grade2–3 IVH with ventricular dilatation, right intraparenchymal hemorrhagic infarction (IPH/PVHI). **C:** DOL 55: Right parenchymal cysts with right ventricular outpouching into area of white matter loss.

Periventricular leukomalacia (PVL) occurs in the white matter without associated hemorrhage and is symmetric, bilateral, and clinically silent. Watershed areas with decreased arterial blood supply in the region of the internal capsule near the angles of the lateral ventricles are especially vulnerable to ischemic injury. Acute ischemia results in focal necrosis and the formation of transient cysts; moderate ischemia results in more gradual injury due to diffuse apoptosis with varying degrees of progressive brain atrophy due to glial and axonal loss associated with low pressure, ventricular dilatation into the areas of volume loss (Fig. 2.13–3). Macroscopic cyst degeneration, visible on cranial ultrasound, occur in 5 percent of VLBW infants. However, microscopic cystic degeneration is much more common, affecting about 50 percent of VLBW infants. In one study, over 70 percent of infants born at ≤30 weeks of gestation had abnormalities in the white matter detected by MRI at term equivalent

gestation. Whereas focal necrosis with loss of all cellular elements is more likely to result in spastic cerebral palsy due to destruction of the motor tracts crossing through the area of acute injury, diffuse white matter loss is associated with more generalized cognitive and behavioral deficits. Signs of white matter and associated axonal loss include thinning of the corpus callosum, delayed myelination, immature gyral development, enlarged subarachnoid space, and volume loss in the cortical and deep grey matter, peristriatal regions, and the temporal lobes.

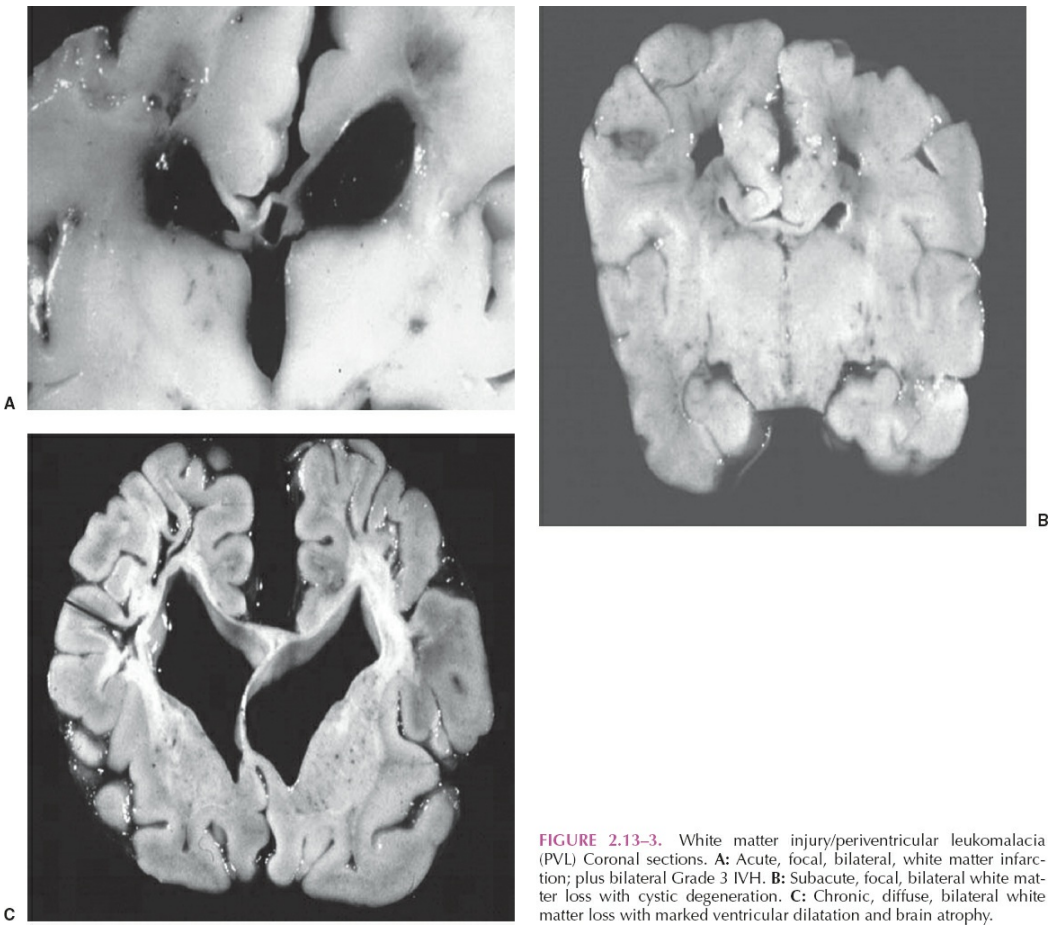

FIGURE 2.13–3. White matter injury/periventricular leukomalacia (PVL) Coronal sections. **A:** Acute, focal, bilateral, white matter infarction; plus bilateral Grade 3 IVH. **B:** Subacute, focal, bilateral white matter loss with cystic degeneration. **C:** Chronic, diffuse, bilateral white matter loss with marked ventricular dilatation and brain atrophy.

FIGURE 2.13–3. White matter injury/periventricular leukomalacia (PVL) Coronal sections. **A:** Acute, focal, bilateral, white matter infarction; plus bilateral Grade 3 IVH. **B:** Subacute, focal, bilateral white matter loss with cystic degeneration. **C:** Chronic, diffuse, bilateral white matter loss with marked ventricular dilatation and brain atrophy.

NEURODEVELOPMENTAL OUTCOME

Most neurodevelopmental outcome studies have focused on the highest-risk preterm infants—either VLBW (<1,500 g) or very premature (VP) born before 32 weeks of gestation or extremely low birth weight (ELBW, <1,000 g) or EP (<26 to 28 weeks of gestation). Adverse outcomes increase proportionately with increasing immaturity with the poorest outcomes in the EP and ELBW group. However, more than 75 percent of all preterm infants are moderate or late preterm, born at 32 to 36 weeks of gestation.

Recently more attention has been paid to neurodevelopmental outcomes of these larger, more mature preterm groups. As expected their risk of adverse cognitive and behavioral outcomes, while much lower than that of more immature preterms, is significantly higher than that of term infants and because of their greater numbers, contributes disproportionately to the burden of neurodevelopmental problems associated with preterm birth.

The best outcome studies are prospective, and comprised of large, population or geographically based cohorts with longitudinal follow-up from infancy through adulthood. Comparison of outcomes with normal birth weight (NBW), term-born control children is critical. Most such studies have been performed in Australia, Canada, Scandinavia, and Great Britain where children can be tracked easily within the health care system over long periods of time. However, reports from single centers and smaller samples are remarkably congruent with the results from larger, longitudinally followed cohorts. Whether there will be substantive differences in outcomes based on birth before and after the introduction of antenatal steroids and surfactant in the early 1990s remains to be seen. The current late adolescent and adult outcome studies are all derived from the earlier time period. Thus far, outcomes during infancy and childhood are quite similar. A myriad of variables have been associated with adverse neurodevelopmental and psychiatric outcomes. Many of these factors are also associated with preterm birth. A unifying, common path may be antenatal, intrapartum, or postnatal hypoxia and/or ischemia or inflammation, which in turn results in brain, especially white matter, injury in this physiologically vulnerable group.

The emotional effect of preterm birth, especially extremely preterm birth, on sibling and parental behavior and emotional well-being may be profound and pervasive, including PTSD, anxiety, depression, and divorce. The most immature preterm infants are critically ill for many weeks with multiple ups and downs involving difficult decisions by both the family and the medical team. Extremely preterm children are often survivors of multiple gestation, adding the burden of grieving a fetal or neonatal loss to the family dynamic. Maternal depression increases the risk of preterm delivery and is associated with internalizing behavior, anxiety, depression, and low social competence in their offspring.

Despite the physical, emotional, behavioral, and cognitive adversities facing many children and adults born prematurely, the majority self-report a satisfactory quality of life. Often self-perception is more positive than parent or teacher report particularly with regard to ADHD and depression. Whether this is lack of self-awareness or reflects resilient coping mechanisms is not clear. The majority of preterm adults compensate for their difficulties and go on to lead full, productive lives.

COGNITIVE OUTCOMES

VLBW or VP children have moderate to severe cognitive deficits with a decrement of 0.8 SD on measures of IQ. The magnitude of cognitive deficits is inversely correlated with gestational age. Moderate and late preterm children born at ≥32 weeks of gestation score an average of 8 points lower, VP children born from 28 to 31 weeks of gestation score 11 points lower, and EP children born at <28 weeks of gestation score 14 points lower—a decrease of approximately 1.5 IQ points/week if born before 33 weeks of gestation. EP/ELBW are 2 to 3 times more likely to have impaired attention than term controls. Deficits include delays in selective, sustained, encoding, divided, and shifting attention with decrements in standardized scores increasing to 0.6 SD at <26 weeks of gestation. Their processing speed, the time required to interpret and respond to incoming stimuli, is also slower. Compared to term controls VP/VLBW children are 2 to 3.5 times more likely to have deficits in all aspects of memory including short-term working memory, long-term episodic (events) memory, and semantic (facts) memory. Impaired visual acuity and visual/perceptual processing problems are common. It is therefore not surprising that executive function is also often impaired. These deficits do not appear to improve over time. Some, such as attention-related difficulties and executive function, tend to worsen as intellectual challenges become increasingly complex.

Academic performance, especially for the EP/ELBW child is moderately to severely affected, and due in part to specific learning disabilities in mathematics, reading, and spelling, often being worse than expected for a given IQ. In one study, compared with term controls, children born at <26 weeks of gestation were more likely to have impairment at 11 years of age in reading (52 vs. 11 percent) and mathematics (70 vs. 14 percent). In another study 65 percent of ELBW children at 8 to 9 years of age had at least one learning disability compared to 13 percent of those born at term. Although many VP/VLBW children may require some special educational resources at some point in their education, most VP/VLBW children are in mainstream classes and graduate from high school; many attend college and pursue postgraduate degrees.

PSYCHOPATHOLOGIC OUTCOME

Behavior problems are three to eight times more common in the preterm child compared to their term counterparts. Preterm children are more likely to have internalizing and less likely to have externalizing problems. The preterm behavioral phenotype has been described as being inattentive, anxious, and socially withdrawn. Psychiatric disorders are often comorbid with other preterm-related cognitive, motor, and neurosensory impairments making diagnosis and treatment more difficult. Being born preterm substantially increases the risk of psychopathology especially anxiety, depression, ADHD, and autism. As with other outcomes, the risk

increases with decreasing gestational age, rising sharply for those born at less than 31 weeks of gestation. Very preterm children are two to three times more likely than moderate to late preterm children to have nonaffective psychosis, depressive disorder, and bipolar disorder. The risks of autism spectrum disorder (ASD), emotional, and behavior problems are each ten times higher if born at 23 to 27 weeks compared to ≥37 weeks of gestation. Many children have some features of ADHD or ASD but not enough to meet the full diagnostic criteria. For instance, although 25 percent of low birth weight (LBW) children are reported by parents to have attention problems on the Child Behavior Checklist (CBCL), only one-quarter of these meet diagnostic criteria for ADHD.

In mid-childhood, 23 percent of EP children were reported to have a psychiatric diagnosis compared to 9 percent of term children (OR 3.2). In this same study, compared to term children, EP children were more likely to have ADHD (12 percent vs. 3 percent), emotional disorders (9 percent vs. 2 percent), and ASD (8 percent vs. 0 percent). In a meta-analysis including five studies of LBW, preterm adolescents, one in four had a psychiatric diagnosis (OR 3.66). In another study of VLBW children mid-adolescence, 11 percent had cerebral palsy and 18 percent had moderate mental retardation, 46 percent had emotional or behavioral problems and 25 percent had psychiatric symptoms fulfilling at least 75 percent of diagnostic criteria. For moderately preterm, young adults, the prevalence of a psychiatric disorder at 19 years of age was higher than their full-term counterparts (30 percent vs. 19 percent). At 19 years of age, the primary disorders were mood disorder, ADHD, and antisocial personality with males and females being equally affected. Interestingly, the risk of having a psychiatric problem increased from 11 to 19 years. More than 50 percent of those with a disorder at 19 years had not had a psychiatric diagnosis when seen at 11 years of age. From a National Registry database, the risk of hospitalization for any psychiatric diagnosis doubles below 33 weeks of gestation.

Studies of preterm young adults have noted decreased risk-taking behavior, delinquency, and incarceration as compared with their peers. The reasons for this are unclear but may reflect lack of interest or opportunity, decreased social interactions, or increased protection by the family.

CORRELATION OF PSYCHOPATHOLOGY WITH BRAIN INJURY

Structural and functional MRI studies demonstrate substantial differences in regional brain volumes and neural connectivity including smaller hippocampal, thalamic, caudate, corpus callosum, and cerebellar volumes as well as thinner cortices found in infants, children, and adolescents born preterm. Disorganization of tracts in the internal capsule and posterior corpus callosum indicates widespread alterations in neural connectivity. Functional MRI and DTI demonstrate that preterm children utilize novel

neural pathways to complete tasks suggesting the development of alternative networks to compensate for abnormal brain development or injury.

ANXIETY/DEPRESSION

In a study from the United States, the attributable risk of LBW(<2,500 g) for depression in female adolescents based on screening questionnaire followed by psychiatric assessment was 18 to 24 percent of LBW females versus 3 percent of NBW females being depressed. LBW adolescent females were also found to be less resilient than their NBW peers. Fewer risk factors were associated with depression in LBW females: for one risk factor 20percent of LBW versus 5 percent of NBW and for two risk factors 85 percent of LBW versus 20 percent of NBW females were diagnosed with depression. In this study and others the risk of depression was not increased in preterm male adolescents. Much lower rates of adolescent depression were also found in a population-based cohort from Australia where 15 percent of preterm LBW adolescent girls and 1 percent of preterm LBW adolescent males had stable depression in comparison to 2 percent of term females and 0.25 percent of term male adolescents. The OR for LBW adolescents for developing a stable depressive disorder was 11.6 after adjustment for background confounders. Moderate to late preterm young adults born at 32 to 36 weeks of gestation are 2.7 times more likely and very preterm young adults born before 32 weeks of gestation are 7.4 times more likely to develop bipolar disorder.

ATTENTION DEFICIT HYPERACTIVITY DISORDER

The prevalence of ADHD is two to four times greater in preterm children and occurs more often in the most premature children. The majority of preterm children have inattentive ADHD subtype rather than combined or hyperactive/impulsive subtypes. In one large population-based study, EP children had a high rate of generalized attention impairment at 11 years of age—77 percent were impaired in at least one area of attention and 50 percent were impaired in two areas. In the same cohort seen at 17 years of age, 18 percent had clinical ADHD compared to 7 percent of the term adolescents. Unlike in term children where there is a male predominance, preterm girls and boys are equally affected. Neonatal factors are weak predictors of attention impairment and ADHD. Poor intrauterine growth may contribute to the risk of hyperactivity. In a large twin study, the lower birth weight twin scored higher on a hyperactivity scale. In preterm children with ADHD, neuroanatomical changes are present on MRI. There are structural and functional abnormalities in specific structures and in the fronto-striatal-cerebellar circuitry with evidence of hyperactivation in some areas and decreased activation in others. Numerous antenatal, intrapartum, and postnatal factors have been associated with ADHD.

AUTISM/AUTISM SPECTRUM DISORDER

The prevalence of autism in preterm children is two to three times the prevalence in overall population. There is a gradual increase in risk of ASD associated with decreased length of gestation with the most immature and the smallest at highest risk of ASD. Clinically, preterm and LBW children have predominately social and communication problems, with less stereotypic, repetitive movements and more motor and cognitive delays.

A high proportion (21 to 41 percent) of VP children fail the Modified Checklist for Autism in Toddlers (MCHAT) screener at 18 to 24 months of age. However, these results are strongly influenced by concomitant cognitive, motor, and neurosensory impairments. The percentage of positive screens is much lower following confirmatory interviews. In later childhood and adolescence the prevalence of ASD is 5 percent in LBW children and up to 8 percent in ELBW or EP children. Although the overall prevalence of autism is four times higher in males, the risk is higher for preterm girls born at less than 33 weeks of gestation. Preterm girls are also more likely to have co-morbid conditions.

Many variables associated with ASD occur more frequently with preterm delivery including advanced maternal age, maternal hypertension, twins, assisted reproductive technology, antenatal antidepressant medication, gestational diabetes, antenatal infection, birth order, and antenatal bleeding, intrauterine growth restriction/small for gestational age, fetal distress, abnormal presentation (breech), induced labor, summer delivery, umbilical cord complications, low Apgar score, needs for neonatal intensive care, hypoxia, jaundice, respiratory distress, feeding difficulties, meconium aspiration, anemia, hemolytic disease of the newborn, congenital malformation, and lack of maternal breast milk feeding. Most of these variables are either primarily or secondarily associated with hypoxic/ischemic, inflammatory or toxic brain injury. Both genetic and environmental components of autism have been demonstrated in twin studies which have high rates of concordance in identical but not dizygotic twins. When identical twins are discordant for ASD, the smaller twin, presumably with the less optimal intrauterine environment, is more likely to be affected. Autism is associated with cranial imaging abnormalities including ventricular dilatation, intracranial hemorrhage, parenchymal white matter volume loss, isolated cerebellar injury, and hypoplasia of the cerebellar vermis.

PSYCHOTIC DISORDERS

Given the low prevalence of nonaffective psychosis, little information is available. The few studies from mental health or large birth registries associate preterm birth, hypoxia and related obstetrical complications (e.g., perinatal asphyxia), fetal undergrowth, and preeclampsia with

schizophrenia. Moderate to late preterm young adults born at 32 to 36 weeks of gestation are 1.6 times more likely and very preterm young adults born before 32 weeks of gestation are 2.5 times more likely to develop nonaffective psychosis. Brain abnormalities associated with schizophrenia include cerebellar hypoplasia, severe IVH, and ventricular dilatation.

In conclusion, psychiatric conditions including depression, anxiety, autism, ADHD, and nonaffective psychosis are all more prevalent in children, adolescents, and adults born prematurely. The risk increases steadily with increasing prematurity, such that those born at less than 28 weeks of gestation are at highest risk. Abnormalities in brain structure and function incurred as a result of preterm birth have been implicated in adverse neurodevelopmental and psychiatric outcomes.

HUMAN IMMUNODEFICIENCY VIRUS IN CHILDREN AND ADOLESCENTS

HIV in Children and Adolescents—Epidemiology

The human immunodeficiency virus (HIV) type-1 is a retrovirus acquired through exposure to infected bodily fluids, most commonly in adults via unprotected sexual contact or injection drug use; infection via blood products, a significant route of exposure early in the epidemic, has been reduced to negligible levels in the United States by donor screening and treatment of donated blood. For infants born to a woman with HIV, exposure can occur in utero, during birth, or via breastfeeding. There is regional variation in the most common route of infection and population affected; in the United States, new infections are most likely among men who have sex with men (MSM). HIV primarily infects and destroys CD4+ T-lymphocytes, and, untreated, results in progressive immunosuppression and vulnerability to opportunistic infections and neoplasms. Centers for Disease Control (CDC) surveillance classification of disease severity assigns categories according to age-specific CD4+ T-lymphocyte count and whether an opportunistic infection has been diagnosed; Stage 3, or AIDS, refers to CD4+ count below 200 or diagnosis of a Stage 3-defining opportunistic infection. The development of effective antiretroviral therapy (ART) has resulted in a dramatic decrease in AIDS incidence and increased focus on more subtle effects of HIV, such as those in the brain; long-term toxicities associated with ART; interactions of HIV with other comorbidities such as substance use, depression, concomitant infections and aging; and functional cure of HIV through elimination of viral reservoirs.

Over the more than 30 years of the HIV epidemic, children and adolescents have been significantly affected. Globally, an estimated 3.2 million children under 15 years of age were living with HIV in 2013, with 240,000 new infections and 190,000 deaths from AIDS that year. Since administration of ART during pregnancy, and to the infant after birth, was discovered to prevent most cases of mother-to-child HIV transmission, the

incidence of new infant infections has dropped dramatically in high resource environments such as the United States, with the majority of perinatal infections occurring in limited resource settings, particularly sub-Saharan Africa. However, in all settings, numbers of older perinatally infected children and youth have grown as effective ART has extended their lives; many of them are now adolescents or young adults and experiencing long-term neurodevelopmental, behavioral, and mental health issues associated with HIV as well as mental health and substance use problems common to their age group and/or demographic. Among individuals who acquire HIV infection postnatally, adolescents (age 15 to 24) have a relatively high rate of infection, accounting for about 33 years of new nonperinatal infections worldwide in 2013.

If taken appropriately, ART can suppress HIV to undetectable levels in blood for most people, preventing immunosuppression with its associated health risks, some of which remain following immune reconstitution, and dramatically reducing the risk of transmitting HIV to others. As of 2015, recommendations for initiating ART in the United States include all individuals upon diagnosis, with the strongest evidence supporting initiation of ART in adults with an AIDS-defining illness or CD4+ <350 cells/mm^3; World Health Organization guidelines call for initiation at CD4+ <500 cells/mm^3. For children, guidelines call for ART initiation in all infants under 12 months of age; for older children, initiation depends on symptoms, CD4+ and viral load.

In order to reap the benefits of ART and achieve viral suppression, however, an individual must first be identified as HIV infected, linked to and engaged in medical care, prescribed ART, and maintain high ART adherence (the so-called HIV treatment cascade). Youth age 18 to 24 with HIV have lower rates of success at all stages of the treatment cascade than other age groups and as a result only an estimated 13 percent in the United States are virally suppressed according to the CDC.

Mental Health and Psychiatric Disorders among Children and Youth with HIV

The interconnections of HIV with mental health and psychiatric disorders, including neurocognitive impairment, psychopathology, and substance use, are well established in the adult population, though less well understood among children and youth. The likelihood of sexual and injection drug use risk behaviors associated with HIV acquisition, as well as other infections, is increased in the presence of mental illness and substance use. Conversely, HIV infection increases risk for these disorders both directly in the case of neurocognitive decline and indirectly through emotional and psychosocial impacts of being infected. Furthermore, neurocognitive and psychiatric disorders complicate health and medication management for individuals with HIV, increasing morbidity and likelihood of HIV transmission, and the presence of multiple comorbidities challenges health

care systems, particularly in lower resource settings. This suggests an important role of psychiatrists and other mental health professionals in preventing HIV infection as well as increasing quality of life in those already infected. Addressing the gap in mental health care for individuals with HIV will be a necessary component of reducing and ultimately eliminating the HIV epidemic.

Adolescents with Behaviorally Acquired HIV

NEUROCOGNITIVE IMPACT OF HIV. Because studies that distinguish or focus on adolescents are scarce, findings from adults as well as adolescents will be discussed in this section. The CNS is vulnerable to HIV and neuropsychiatric sequelae of HIV have been described since the beginning of the epidemic. The virus can enter the CNS early in infection via infected monocyte-derived macrophages and lymphocytes and establish a locally productive infection that may be relatively independent of other loci, representing a viral reservoir. The effects of HIV on neural functioning are largely indirect, through exposure to viral proteins, production of proinflammatory cytokines by infected monocytes and other cells, and alteration of the blood–brain barrier rather than through direct infection of neurons. HIV can lead to a dynamic state of chronic inflammation and immune activation that can change with ART, may be influenced by other lifestyle factors such as substance abuse, and whose interaction with neurodevelopment in perinatally infected children is not well understood.

Before the advent of highly active ART, HIV-associated dementia occurred in 20 to 30 percent of adults with HIV by the time of death, typically in the context of significant immunosuppression, where it can still occur. In addition, opportunistic infections in the CNS such as progressive multifocal leukoencephalopathy can occur with immunosuppression. Among severely immunocompromised individuals, delirium, coma, seizures, and intracranial hypertension can be seen. In the era of ART, the incidence of HIV dementia and other severe sequelae has declined dramatically; however, the prevalence of milder impairments continues to be high and is particularly likely in individuals who have experienced profound immunosuppression at some point despite later immune reconstitution (i.e., “legacy” effect).

Currently, DSM-5 classifies the CNS sequelae of HIV among the major or mild neurocognitive disorders. A diagnosis of major neurocognitive disorder due to HIV requires evidence of significant cognitive decline from a previous level of performance in one or more cognitive domains, based on both concern by the individual, informant, or clinician and objective evidence from neuropsychological or other assessment; that cognitive deficits interfere with independence in everyday activities; that they do not occur exclusively in the context of a delirium; and that they cannot be

better explained by another mental disorder. It can be coded as with or without behavioral disturbance. In order to diagnose mild neurocognitive disorder, the cognitive decline should be modest and not result in interference with independence in daily activities. Other classification schemes for HIV-associated neurocognitive disorder (HAND) specify cognitive domains that should be assessed and required level of performance compared to normative data, and include an asymptomatic form with neurocognitive impairment but no interference with everyday functioning.

HIV dementia in untreated individuals has a "subcortical" presentation, with slowing, impaired retrieval from memory but more intact recognition, fine and gross motor impairment, apathy, and decrease in emotional expressivity. With ART, impairments are more subtle and have been suggested to have a more "cortical" presentation, with executive function difficulties assuming more prominence. Several mechanisms have been proposed to explain the presence of impairments in virally suppressed or early-infected individuals, including inflammation and higher viral activity in the CNS than in the periphery. Few studies of neurocognitive functioning in youth with HIV have been performed; in one of the largest to date, 220 ART-naïve behaviorally infected youth of age 18 to 24 underwent a comprehensive neurocognitive and behavioral evaluation. Striking impairments in verbal and nonverbal learning, and lesser difficulties in other areas, were observed. Limited associations with HIV disease severity markers raised the question of subtle, HIV-associated neurocognitive function in this population, but associations with demographics and psychiatric and educational covariates suggested that other causes for the impairments were likely more contributory. Distinguishing subtle HIV-associated impairments from those related to comorbidities such as depression or substance use requires careful assessment. Regardless of their origin, however, neurocognitive impairments have been found in adults to be associated with poorer daily functioning and medication adherence; monitoring is indicated in all individuals with HIV. Unfortunately, efforts to develop CNS-specific therapies have been disappointing and results of ART with greater penetrance across the blood–brain barrier into the CNS have been mixed. Although cognitive impairments have shown modest improvement with ART and systemic viral suppression in some studies, this has not yet been studied in youth; furthermore, potential neurotoxicity of ART is of particular relevance for this population facing many decades of therapy. Nevertheless, at this time, viral suppression with ART is considered the most effective treatment for HANDs.

MENTAL ILLNESS AND SUBSTANCE USE. Though prevalence estimates vary across settings and populations, in general, the literature demonstrates a higher

prevalence of mental health disorders among adults with HIV than among the general population, with rates up to 75 to 83 percent in some studies. Several classes of disorders are of particular concern due to their prevalence or potential to impact health outcomes or HIV transmission.

Depression. Depression is a common referral issue for mental health treatment for individuals with HIV, with reported prevalence rates for major depression twice as high as among the general population and depressive symptoms reported by up to 85 percent. In a survey study of 2,032 youth of age 12 to 26 years receiving care at adolescent HIV clinics, about 24 percent of both males and females with behaviorally acquired HIV reported severity of depression symptoms above the clinical cutoff on the Brief Symptom Inventory. Among youth at a subset of sites, 42.6 percent reported symptoms of distress in the clinically defined range across subscales; of these, only 39.7 percent reported receiving mental health care in the previous year, and 21.9 percent reported receiving medications for emotional issues. The rates of treatment were significantly lower for black youth, demonstrating disparities in mental health care access for black youth with HIV.

Assessment of depression is important in clinical care of youth with HIV for several reasons even in the absence of depression as the presenting complaint. The American Academy of Pediatrics notes the higher rates of depression and suicidal ideation among sexual minority (lesbian, gay, bisexual, transgender, and questioning) youth and the mental health treatment disparities resulting from homophobia and heterosexism and recommends screening for depression as part of a complete psychosocial history. Depressive symptoms can be a side effect of certain antiretroviral medications (e.g., efavirenz) and indicate the need for a change of medication regimen. Untreated depression can contribute to sexual and substance use risk behaviors and to poorer adherence and engagement in care, thus increasing morbidity and HIV transmission likelihood, and to suicide. Finally, cognitive symptoms associated with depression (e.g., psychomotor slowing, apathy) can be similar to those of HANDs and thus depression must be considered as a differential diagnosis during neuropsychological assessment.

Pharmacological treatment of depression in youth with HIV can follow principles and guidelines for other populations since there is no evidence of differential efficacy or side effect profiles in the presence of HIV; however, attention to potential interactions with ART is important (below). As in other populations, psychopharmacologic treatment can be combined with other supportive and evidence-based therapies. Evidence-based treatments such as CBT are being adapted to incorporate concerns relevant to youth with HIV and address associated issues such as medication adherence.

Anxiety Disorders. Although lifetime prevalence of anxiety disorders may be

higher in adults with HIV, anxiety disorders have not received as much focus in studies of HIV. In the survey study of youth with HIV, 16 to 17 percent reported clinically significant symptoms of anxiety. The exception is PTSD, which appears to have higher prevalence (30 to 64 percent) among adults with HIV than in the general population in North America and Africa. Youth with HIV may have been subject premorbidly to chronic stress associated with discrimination as members of racial and sexual minority groups, poverty, histories of physical or sexual abuse, and in some cases homelessness and sex work, among other stressors. Following their diagnosis, they may experience trauma related to receiving the diagnosis and anxiety related to HIV stigma, disclosure to sexual partners and others, and fear of HIV sequelae and mortality. The association of stress and trauma with inflammation and decreased immune functioning increases concern about PTSD in youth with HIV. Individual and group therapy for anxiety and PTSD should be sensitive to the acute and chronic stresses common to this population. Recommendations regarding pharmacotherapy are based largely on studies with non-HIV-infected populations but clinicians should be aware of the possibility of additional confusion or sedation in individuals with HIV-associated cognitive impairment.

Psychosis. Individuals with chronic mental illness have an increased risk of becoming infected with HIV. Those with substance use or unstable housing may exchange sex for money or drugs or have decreased access to condoms or clean needles. HIV screening, risk reduction counseling, and, where needed, substance abuse treatment should be considered for this population. Although less commonly relevant for adolescent populations, new-onset psychotic symptoms or mania can occur in late-stage disease as part of HIV disease or opportunistic infections in the CNS; these symptoms may respond to low-dose neuroleptics.

Substance Use. Injection drug use is a major vector for HIV infection, particularly in South and Southeast Asia but to varying degrees worldwide. In addition, substance use is relevant for adolescents with HIV as a common comorbidity of psychiatric disorders as well as a disorder in its own right, as a factor in sexual risk behaviors and poor medication management, and because of its potential interactions with immune and neurocognitive functioning. Among adults, some substances such as alcohol and methamphetamine have been found to exacerbate HIV-associated neurocognitive impairment. Little research on this topic has yet been done with adolescents; however, one study showed interactions of HIV and alcohol risk in associations with lower cognitive functioning in youth with HIV, and studies showing increased vulnerability of the developing adolescent brain to the effects of substance use raise concern. Furthermore, drugs of abuse may interact with some ARTs. Screening for

substance use is a critical part of psychiatric and primary care for youth with HIV.

Although estimates vary, substance use, overall, is generally found to be more prevalent among adults with HIV than among uninfected adults. In a multicenter study of 220 behaviorally HIV-infected youth engaged in care, the most common substances of abuse were cannabis, tobacco, and alcohol; use of other drugs such as methamphetamine, opioids, club drugs, and cocaine was less common but may vary regionally. Cannabis is of particular interest due to availability by prescription in some areas, its increasing acceptability, particularly among youth, immune interactions and possible exacerbation of cognitive impairment, and use by some individuals to manage HIV symptoms and ART side effects. In the aforementioned study, about 20 percent of youth used cannabis daily.

Although HIV care providers may delay ART due to concerns about the impact of substance use on medication adherence, facilitating engagement in care by youth with substance use is vitally important for clinical and public health reasons. Examples of strategies to encourage engagement in care include integrating treatment for substance use disorders with mental health treatment and HIV services where possible and reducing stigma of acknowledging substance use to adolescent providers. Where abstinence is not achievable, harm reduction (e.g., reducing behaviors with the potential to transmit HIV) is an applicable strategy. Evidence-based youth-friendly therapeutic approaches such as motivational interviewing can be implemented in a variety of settings.

Children and Youth with Perinatally Acquired HIV

NEUROCOGNITIVE IMPACT OF HIV. The current understanding of neurodevelopmental effects of HIV among perinatally infected children is complicated by the changing state of their CNS and the numerous other developmental risks present in this population. These can be environmental (e.g., economic and educational disadvantage, chronic stress related to poverty), genetic (e.g., familial psychiatric and learning disorders), prenatal or birth-related (e.g., LBW, prenatal exposure to substances of abuse and ARTs, other congenital infections such as cytomegalovirus), or related to their own and/or caregivers' chronic illness (e.g., death or changes in caregiver, missing school). Prior to the advent of ART, progressive HIV encephalopathy characterized by delay or loss of developmental milestones or cognitive functioning, pyramidal tract motor deficits, and microcephaly (or atrophy in older children) was seen in 24 to 35 percent. A plateau course with deviation from the expected developmental trajectory rather than a loss of milestones or decline in cognitive functioning was also seen. Currently, as with adults, encephalopathy is less common (2 to 15 percent) and impairments are most

likely in children or youth who have had immunosuppression sufficient for an AIDS diagnosis in the past regardless of current immune functioning, particularly if the most severe disease impact was early in life (e.g., prior to age 3) or encephalopathy was present. Although most studies have utilized measures of global cognitive functioning such as IQ, impairments in specific cognitive domains such as language, executive functions, or processing speed have been noted. According to some studies, children who have had consistently good viral control perform at a level comparable to those who were perinatally exposed to HIV but uninfected, with both groups performing in the low average range overall possibly due to the other developmental risks outlined above. However, a role of prenatal exposure to HIV and ARTs cannot be ruled out by these studies.

MENTAL ILLNESS AND SUBSTANCE USE. Children and youth with perinatally acquired HIV face numerous risks for emotional and behavioral issues. These include high rates of familial mental illness and substance use, potential or actual death of a parent, concerns about HIV-related stigma and disclosure to others, and living with the actual or potential effects of HIV itself such as illness, cognitive impairment, delayed puberty, pain, and death. Perinatal HIV differentially affects minority and economically disadvantaged populations who may as a result already be subject to high levels of chronic stress. A 2013 review (Mellins & Malee) notes higher rates of psychiatric disorders and behavioral problems in children and youth with perinatally acquired HIV than seen in the general population. Studies using diagnostic tools such as structured interviews have found that up to 70 percent have to meet criteria for a non–substance-use psychiatric diagnosis at some time point, most commonly anxiety, ADHD, disruptive behaviors, and depression. However, some studies have found similar rates among youth who were perinatally HIV-exposed but uninfected, bringing into question the relationship of the high rate of diagnoses to HIV infection. Similarly, the results of studies using caregiver- or youth-report symptom checklists vary, some demonstrating higher rates of reported mental health problems than in the general population or comparison groups, with others showing no differences or even lower ratings of mental health problems among infected youth, possibly reflecting the close monitoring of and availability of resources for youth with perinatally acquired HIV.

Only in recent years have significant numbers of youth with perinatally acquired HIV survived into adolescence and reached the age of substance use initiation. They have been found to initiate substance use at similar rates as other adolescents, with higher severity of emotional symptoms or meeting symptom criteria for ADHD, major depressive disorder, oppositional or conduct disorder associated with greater likelihood of substance use.

Regardless of the origin of mental health problems in this population, the research demonstrates a clear need for intervention. In addition to the importance of intervention for decreasing suffering and improving quality of life, as youth with perinatally acquired HIV age into adolescence, those with mental health problems are more likely to engage in substance use, sexual risk behaviors, and to have difficulty adhering to their medication regimens. Thus, intervention has the potential to improve long-term health outcomes and reduce HIV transmission. Mental health services are commonly provided by pediatric and adolescent HIV programs. However, little data exists regarding differential effects of either pharmacological or behavioral treatment in perinatally infected youth compared to uninfected youth.

Factors that have been found to increase risk for mental health problems in the context of perinatally acquired HIV include caregiver psychiatric disorder and limit-setting problems, lower child cognitive functioning, and exposure to stressful life events. Conversely, resilience to mental health problems has been associated with higher self-esteem; positive caregiver–child relationships and parental monitoring and involvement; and family, peer, and community support. Thus, points of intervention can be multisystemic and include, for example, family-based therapy to facilitate communication, including disclosure of the child's or mother's HIV infection, peer support groups, and community programs to reduce HIV stigma. At the individual level, interventions to boost resilience through increasing coping skills and self-esteem can complement the use of pharmacological and evidence-based therapeutic approaches.

Issues Related to Psychopharmacologic and Antiretroviral Treatment

Key issues for psychiatrists monitoring youth with HIV and prescribing psychopharmacologic medications are potential psychiatric side effects of ART and interactions between ART and psychotropic medications. Psychiatric side effects of some antiretroviral agents are well known; use of efavirenz, in particular, can be associated with mood disturbance, insomnia and vivid dreams, impaired concentration, and in some cases hallucinations, among other symptoms. Familiarity with the side effect profiles of patients' ART regimens is critical for the treating psychiatrist. An important issue for psychiatrists is the common pharmacokinetic and pharmacodynamic interactions between psychopharmacologic and antiretroviral agents. Some antiretroviral medications, such as protease inhibitors and nonnucleoside reverse transcriptase inhibitors, involve the CYP450 system and thus have the potential to affect drug levels of antidepressants and other medications. For example, the protease inhibitor ritonavir can increase levels of amitriptyline and its side effects; conversely, ritonavir and lopinavir can decrease the effects of bupropion. An additional consideration in prescribing psychopharmacologic treatments in the

context of HIV is that individuals using specific ARTs or who have neurocognitive impairment may be more sensitive to side effects of psychiatric medications. In general, the treating psychiatrist is advised to take a careful drug history, anticipate drug interactions and toxicities, start with lower dosages and take a cautious approach to increasing them.

EPILEPSY

Seizures and Epilepsy Syndrome in Children

It is estimated that 1 to 1.5 percent of children will have a nonfebrile seizure at some time in their life. The most common seizure condition in pediatric medicine is febrile seizures which can occur in 4 percent of infants, toddlers, and preschoolers overall with some ethnic groups such as Japanese (10 percent) and Guamanians (20 percent) having a much higher incidence. Seizures can begin at any time of a child's life from prenatally to perinatal to onset at late adolescence. Epilepsy has been defined in the past as having two or more unprovoked seizures. A slightly different operational definition of epilepsy has recently been stated by the International League against Epilepsy.

This defines epilepsy as a disease of the brain defined by any of the following conditions:

1. At least two unprovoked (or reflex) seizures >24 hours apart
2. One unprovoked (or reflex) seizure and a probability of further seizures similar to the general recurrence risk (at least 60 percent) after two unprovoked seizures, occurring over the next 10 years
3. Diagnosis of an epilepsy syndrome.

Table 2.13–5.
Most Current ILAE Proposed Classification of Seizures Types (2009)

Focal	Generalized Seizure Types
a) without impairment of consciousness type	a) Absence
With motor signs	i. Typical
With autonomic features	ii. Atypical
With sensory symptoms	iii. Absence with special features
With psychic symptoms	Myoclonic absence
b) with impairment of consciousness/responsiveness	Eyelid myoclonias
c) evolving to secondary generalized seizure	b) Myoclonic seizures
	c) Myoclonic atonic
	d) Myoclonic tonic
	e) Clonic
	f) Tonic
	g) Atonic
	h) Focal seizures
	i) Unknown
	j) Epileptic spasms

A seizure that cannot clearly be classified by one of the above categories is advised to best be considered unclassified until further information allows for accurate diagnosis.

Child neurologists and epileptologists consider in their evaluation of child with a single or recurrent seizures the historical and examination findings, seizure type, age at onset, and EEG findings. In doing so, roughly 50 percent of children can be characterized as having a particular epilepsy syndrome. Making a syndrome diagnosis can provide prognostication and often guide discussion about therapeutic options. The latter are based upon accumulated experience, clinical trials of single agents versus placebo, or rarely as in the case of Childhood Absence Epilepsy well conducted comparative clinical trial. Table 2.13–5 provides information about seizure type classification and Table 2.13–6 provides a list of Pediatric Epilepsy Syndromes.

In past decades, epileptologists and geneticists hoped for a single gene to be associated with electroclinical epilepsy syndrome, but as data have accumulated over the past 20 or more years this idea has not been supported by available data from genetic association studies. Many individual electroclinical syndromes have been found to be associated with a given single gene mutation—often de novo, while a particular electroclinical syndrome may be associated with two or three or more different genetic mutations. Knowledge of the specific gene mutation for a given epilepsy syndrome has largely failed to inform optimal choice of therapy. The two exceptions to this include SLC2A1 mutation (glucose transporter deficiency) and SCN1A mutation associated with Dravet syndrome. The former condition responds to ketogenic diet therapy and the latter is often worsened by lamotrigine therapy which is usually avoided in such patients.

Epidemiology of Psychiatric Issues in Childhood Epilepsy

The best prospective assessment of psychiatric issues in pediatric epilepsy comes from a multiyear study of new-onset seizures/epilepsy in Connecticut. In this study, 501 children were enrolled and followed for several years. One hundred fifty-nine children were judged to have a "complicated" profile with either a brain lesion evident on imaging or IQ less than 80 or both. MRI showed no etiological structural lesion and IQ was 80 or higher in 342 children, constituting the "uncomplicated" group.

	Complicated *N* = 159	Uncomplicated *N* = 342	P value
No psychiatric disorder	98 (62%)	251 (73%)	0.008
Psychiatric disorder present	61 (38%)	91 (27%)	—

Table 2.13–6.

Organization of the Epilepsies—Electroclinical Syndromes by Age of Onset

Neonatal Period
Benign familial neonatal epilepsy
Early myoclonic encephalopathy
Ohtahara syndrome
Infancy
Epilepsy of infancy with migrating focal features
West syndrome
Myoclonic epilepsy of infancy
Benign infantile epilepsy
Benign familial infantile epilepsy
Dravet syndrome
Myoclonic encephalopathy in nonprogressive disorders
Childhood onset
Febrile seizures plus (FS+)
Panayiotopoulos syndrome
Epilepsy with myoclonic atonic (previously termed myoclonic astatic) seizures
Benign epilepsy with centrotemporal spikes (BECTS)
Autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE)
Late-onset childhood occipital epilepsy (Gastaut type)
Epilepsy with myoclonic absences
Lennox–Gastaut syndrome
Epileptic encephalopathy with continuous spike-and-wave in sleep (CSWS)
Landau–Kleffner syndrome
Childhood absence epilepsy
Adolescent/Adult onset
Juvenile absence epilepsy
Juvenile myoclonic epilepsy
Epilepsy with generalized tonic–clonic seizures alone
Progressive myoclonic epilepsies
Autosomal dominant epilepsy with auditory features
Other familial temporal lobe epilepsies
Less specific age relationship
Familial focal epilepsy with variable foci (onset childhood to adult)
Reflex epilepsies
Distinctive constellations
Mesial temporal lobe epilepsy with hippocampal sclerosis
Rasmussen syndrome
Gelastic seizures with hypothalamic hamartoma
Hemiconvulsion-hemiplegia-epilepsy

Internalizing disorders (depression, bipolar disorder, obsessive compulsive disorder, schizophrenia) were reported in 17 percent of the complicated group and 15 percent of the uncomplicated group. Externalizing disorders (ADHD, ODD, and conduct disorder) were present in 31 percent of the complicated group and 18 percent of the uncomplicated

group (*p* 0.002).

When specific features were evaluated it was found that age of seizure onset influenced occurrence of any psychiatric disorder. Of those with onset <2 years old 11 percent had psychiatric disorder, with onset at 2 to 5 years old 24 percent had psychiatric diagnosis, and if onset was >10 years old 36 percent had psychiatric disorder. Gender was not a factor influencing incidence of psychiatric disorder. Of those children who had Childhood Absence Epilepsy 23 percent had a psychiatric diagnosis. Thirty percent of children who had Benign Epilepsy with Centro-temporal spikes (aka Benign Rolandic Epilepsy) had a psychiatric diagnosis. Thirty-seven of children with Juvenile Absence Epilepsy and Juvenile Myoclonic Epilepsy had a psychiatric diagnosis.

Depression and Anxiety in Pediatric Epilepsy

In the population-based study referenced above, often patients in both the complicated and uncomplicated groups had more than one psychiatric diagnosis—most often depression + anxiety, or ADHD + anxiety or depression. A surprising finding in this study was no increase in the occurrence of psychiatric disorders in children who had failed > two antiseizure medication trials versus those well controlled on the first or second medication used.

Other studies have found by semi-structured psychiatric interview assessment of children with complex partial seizures and absence seizures, 12 percent and 13 percent, respectively had features of depression. When assessed by CDI 25 percent of children in a community-based cohort of 115 had depression and this did not differ from those in a tertiary referral center.

Surveys of the general childhood population have shown 9.5 percent to have anxiety disorder and in children with epilepsy figures have ranged from 5 to 49 percent. A wide array of disorders have been identified including generalized anxiety disorder, specific phobias, social anxiety, with one study finding specific phobias to be most common. No specific data regarding optimal therapy of anxiety disorders in pediatric epilepsy have been published as of this writing, though CBT is usually recommended as first-line treatment and consideration of combined CBT with a SSRI if needed.

Data regarding the incidence/prevalence of obsessive compulsive disorder or psychosis in children who have epilepsy have not been published.

Interictal and Postictal Psychosis

Only anecdotal reports are available about psychosis in childhood epilepsy, and one must turn to the adult literature for more informative relevant

data. Up to 9 percent of adults who have temporal lobe epilepsy will have an associated interictal psychosis, and this has been observed to be more common with left versus right temporal lobe seizure origin. Such patients most often display positive symptoms with delusions and may demonstrate hallucinations in the setting of otherwise clear consciousness. They often have absence of a formal thought disorder, a better premorbid state, and less personality deterioration.

An uncommon situation in which psychosis coincides with improved seizure control often with improved EEG findings via medication, vagal nerve stimulation, or resective surgery has been termed "forced normalization" and seems to be a rare phenomenon. Great variability (<1 to 28 percent) has been reported in the occurrence of de novo psychosis after epilepsy surgery. When this occurs it usually begins within 5 months, is more likely to occur in individuals who have a family history of psychosis or are older than 30 years of age at the time of surgery.

Proposed criteria for defining postictal psychosis include the following:

1. within 1 week of a seizure or seizure cluster
2. duration >1 day and <3 months
3. disorientation, delusions, delirium, or hallucinations in clear or clouded consciousness
4. no evidence of anticonvulsant toxicity, nonconvulsive status epilepticus on EEG, prior interictal psychosis, recent head trauma, or drug/alcohol intoxication.

Postictal psychosis has been sparsely reported in children and adolescents. In the adult epilepsy literature, the risk of postictal psychosis is increased with age >30 years old, male gender, partial seizures, rapid secondary generalization of partial seizures, prior encephalitis, slowing on EEG, and prior personal or family history of psychosis. Thought insertion, commenting/command hallucinations, and negative symptoms are rare.

It is advisable to avoid as far as possible the use of clozapine, chlorpromazine, and loxapine for pharmacologic therapy of psychosis in patients with epilepsy due their risk of lowering seizure threshold and association with seizure occurrence in individuals not known to have epilepsy. It is important to be aware of potential drug–drug interactions when neuroleptic agents are used; for example, ziprasidone levels are reduced by enzyme-inducing antiseizure medications such as carbamazepine, phenytoin, and phenobarbital.

Aggression

Little data about aggression is reported in the pediatric epilepsy literature. Up to 5 percent of adults with epilepsy have problems with aggressive behavior. Risk factors include intellectual disability, onset of seizures earlier than age 10-years old, traumatic brain injury (TBI), psychosis, and

lower socioeconomic status. Medications including gabapentin, phenobarbital, lamotrigine, and levetiracetam have been associated in case series with occurrence of aggressive behavior that improves after cessation of that particular medication. Ictal and postictal aggression are a rare phenomenon and in video-EEG (VEEG) studies have been associated primarily with staff or family member physical restraint attempts.

Attention Deficit Hyperactivity Disorder and Sorting Out Staring Spells

Even prior to initiation of antiseizure medication symptoms of attention deficits are present commonly in children who develop epilepsy. Most studies have found approximately a 30 to 35 percent prevalence of attention deficit in children with epilepsy versus 5 percent prevalence in the general childhood populations. A prospective cohort study of baseline neuropsychological testing in children with new-onset seizure(s) including half not yet started on AED compared with unaffected siblings demonstrated a significantly worse performance on Conners Continuous Performance Task III in patients versus nonaffected siblings. Medication may exacerbate or precipitate attentional problems such as has been demonstrated in a large comparative blinded trial of valproate, ethosuximide, and lamotrigine in the therapy of childhood absence epilepsy. Thirty-six percent of 393 children with new-onset childhood absence epilepsy had attention deficit as gauged by Conners Performance Test II. After 16 to 20 weeks of therapy the 49 percent of the valproate group showed attention deficits compared to 32 percent of the ethosuximide and 24 percent of the lamotrigine treated groups, statistically significant for valproate versus others but no significant difference in lamotrigine versus ethosuximide group.

In the past there have been statements advising avoidance of stimulant medication in children who have seizures. However, numerous open label studies and two randomized placebo controlled studies have failed to show a worsening of seizures in pediatric epilepsy taking either methylphenidate or atomoxetine. The issue has not been well studied in pediatric epilepsy patients with unstable poorly controlled seizures.

Children who have attention deficit disorder will not uncommonly present to health care providers with a history of staring spells. When observant parents or teachers can provide a detailed history it is usually found that the staring off does not interrupt ongoing speech, play, eating, walking, or running as is usually the case in absence or complex partial seizures. When staring spells are brief on the order of seconds and not associated with versive head or eye movements or somnolence/confusion after and there is concern that a child might have absence seizures a routine awake EEG with two periods of hyperventilation will detect generalized 3 Hz spike wave activity of absence seizures in 95 percent of affected children. If there is lingering doubt, long-term ambulatory EEG or

inpatient VEEG may be required to settle the issue but will be most helpful only if the staring spells occur on a daily basis.

Psychogenic Nonepileptic Seizures

Estimates from epilepsy centers caring for children using VEEG for documentation of PNESs indicate an estimate of the prevalence of PNES to vary between 1 percent and 9 percent. This is lower than the 15 to 20 percent incidence in patients referred to adult epilepsy centers for VEEG monitoring for refractory seizures. Nonepileptic events in children vary according to age and can include syncope, parasomnias, various movement disorders, inattention/daydreaming, breath-holding spells, hypnic jerks, and sleep myoclonus. In general, the older the child the greater likelihood is that an event is a psychogenic nonepileptic event as opposed to one of the above. Older children are more likely to have motor features of their PNES and these may have a variety of manifestations—focal, upper or lower extremities alone, whole body, and so forth. Clinical videography of such events to document the habitually observed characteristics of true epileptic seizures is a critical part of the evaluation. Younger children have been found to more often simply display a behavioral pattern of staring and unresponsiveness without motor features.

The gold standard in diagnosing PNES is VEEG monitoring with proper response testing. Such patients will usually have preservation of normal EEG background rhythms and be nonresponsive to commands during PNES spells. Behavioral characteristics of PNES often include longer duration of events (>10 minutes), fluctuating course during events, side to side head or body movements, pelvic thrusting, asynchronous movements (excluding abrupt asymmetric tonic postures as are seen in frontal lobe partial seizures), crying during the event, eyes remaining closed during the event, and memory recall within event. Tongue biting, incontinence, flailing or thrashing movements, opisthotonus, and gradual onset of events may occur in either true epileptic seizures or psychogenic nonepileptic events. They usually arise from an awake state or from feigned sleep, the latter of which can be determined from analysis of the EEG. Some patients presenting to emergency rooms may receive multiple doses of anticonvulsants or even be intubated. For convulsive events a major rise in serum CPK 8 to 16 hours after a convulsive event supports that an event was a true epileptic seizure where the absence of such a rise argues against such an event being a true tonic–clonic epileptic seizure, but would not exclude such a possibility. Prolactin measurements comparing a baseline nonictal sample to a sample drawn 10 to 20 minutes after the onset of the seizure may be of help when the postictal prolactin measure is twice that of the baseline measure. Results may be unreliable in the setting of dopamine antagonist or agonist use, tricyclic antidepressant use, breast stimulation, and syncope. Unfortunately, surveys of psychiatrists reveal that a

substantial portion of psychiatrists believe that VEEG monitoring is not an accurate tool for determining if an event is a PNES. A well-explicated characterization of the level of certainty in making the diagnosis of PNES is detailed in a recent ILAE task force report and can be applied in a variety of different scenarios including health care facilities that do not have VEEG capability.

As most teens and adults who have PNES appear to have somatoform disorders as opposed to major depressive or anxiety or bipolar disorder, psychopharmacologic therapy often has a limited role in habilitating such patients. Psychiatric consultation is sought to determine in a structured psychiatric interview if such other diagnoses are present such as depression or anxiety which may respond to psychopharmacologic intervention. Such consultation however is often of great benefit in identifying the nature of precipitating factors. In published studies factors include school-related difficulties (9 to 46 percent); stressful family environment (8 to 67 percent); sexual/physical abuse (8 to 38 percent); fear of rejection/attention getting (19 to 26 percent) and no precipitant found in 12 to 29 percent of various studies. DSM-IV and ICD-10 psychiatric diagnoses have been reported in six studies. Results indicate 19 to 62 percent found to have depression or anxiety, 7 to 41 percent had separation anxiety/school phobia, 8 to 12 percent have panic disorder, 6 to 12 percent have exclusive general anxiety disorder, 2 to 8 percent have PTSD, and no psychopathologic diagnosis in 16 to 59 percent or not reported in others.

A key element in therapy is making a firm diagnosis for the habitual type of spell. This should be presented in a positive fashion to the parent or parents by the child neurologist/epileptologist. Emphasis should be placed on the “good news” that the events are not epileptic. At the same time it must be communicated to the family and child that the child should not be punished for “faking” seizures and that these events have a biopsychosocial basis.

Therapy in some situations may be as easy as simply identifying the nonepileptic nature of the events. Some series report a high resolution rate within a few weeks of diagnosis of PNES. There should be clear communication that the intervention program for the treatment program must focus on psychological as opposed to medical causes. In older children and adults multimodality therapies including CBT (once weekly for 12 weeks) have been found to be more effective, with or without sertraline compared to sertraline alone or treatment as usual. It may be important in some patients to continue to see the child neurologist for weaning off antiepileptic medication and especially when patients have both PNES and epileptic seizures which is not uncommon. Primary care physicians, school personnel, and social groups should be engaged in providing support of the treatment effort. This often requires communication about the appropriate response to the PNES spells in

environments outside the home. Psychiatrists can play a critical role in the habilitation of such children not only by providing psychopharmacological therapy monitoring but also in assisting the neurology team in identifying appropriate local resources for CBT. Interventions in addition to CBT have included other modalities namely hypnosis, occupational therapy, art therapy, music therapy, and family therapy and when available may assist in effecting a resolution or marked reduction in spells. No improvement may be found despite intervention and has been reported in 0 to 18 percent of children in various series. It is reasonable in such situations to ask the neurologist/epileptologist to reassess the patient in light of lack of improvement.

Presurgical Psychiatric Evaluation

In surveys of Epilepsy Centers, 20 percent of centers performed presurgical psychiatric evaluations on all patients prior to resective epilepsy surgery. Forty-five percent of centers referred for psychiatric evaluation only if prior psychiatric history was present or if presurgical neuropsychological evaluation advised psychiatric evaluation. One-fourth of such centers consulted with a psychiatrist focused on epilepsy while 45 percent utilized different psychiatrists for this purpose. The rationale for performing such evaluations is that postoperative psychiatric complications are not rare and identification of personal or family history of prior psychoses, depression, and anxiety can aid in the evaluation and therapy of any postoperative psychiatric disorder. Decades ago some patients were denied epilepsy surgery if there was a major psychiatric disorder, but this is no longer judged to be a specific contraindication to surgery in an otherwise good surgical candidate. However, one study found that the presence of major depression was a significant factor in reducing the likelihood of seizure freedom after temporal lobectomy. A single center study utilizing both CBCL and clinical psychiatric diagnostic evaluation reported results on 40 children (aged 6 to 17 years who also had IQ above 70) who were undergoing evaluation for surgery for refractory partial epilepsy. Eighty-two percent had at least one Axis I psychiatric diagnosis based upon the clinical interview. ADHD was found in 40 percent. Anxiety was found in 32 percent and depression found in 20 percent. Two or more diagnoses were found in 28 percent of this group. Use of CBCL found broadband categories to be in clinically significant range in 35 percent for internalizing problems and 25 percent for externalizing problems. Analysis of temporal lobe versus extratemporal foci in this group revealed that a similar proportion of both groups had attention and anxiety diagnoses, whereas CBCL data revealed that the temporal lobe foci patients had higher scores in depression-specific categories.

Adjustment Issues

Surveys have found that children with epilepsy have lower self-concept and lower self-esteem than children with other chronic illnesses such as diabetes or asthma. Lower social competence has been more closely associated with lower IQ range, prior disruptive behaviors and less so with seizure variables. Half of teens who have epilepsy state they do not feel stigmatized; however they still often "keep it a secret." Surveys of childhood and teen peers about epilepsy indicate major misperceptions supporting such stigmatization. Forty-six percent were not sure if epilepsy was contagious. Forty percent were not sure if people with epilepsy are dangerous, and only 31 percent would date a peer who is known to have epilepsy. Unfortunately one survey revealed that 24 percent of parents believed that having a child with epilepsy in their own child's classroom would be disruptive. Teachers tend to underestimate learning abilities and athletic abilities and commonly report a lack of knowledge about how to handle a seizure in the classroom. As a consequence many teachers and school personnel often advise unnecessary restrictions for students who have epilepsy. Many local chapters of the Epilepsy Foundation of America, many pediatric epilepsy centers and other organizations provide educational programs for school nurses and school staff and offer teen and child support groups, parent support groups, and summer camp programs to help children with epilepsy overcome such stigma and promote a better quality of life.

TRAUMATIC BRAIN INJURY

TBI in children and adolescents is a major public health problem. The annual incidence of TBI in this age group is 400/100,000 and the male:female incidence ratio is at least 1.8:1. Incidence rates are related to median family income independent of age, race, or ethnicity. The proportion of TBI caused by motor vehicle accidents increases with age. The mechanism of injury in infants, toddlers, and young children is related to child abuse or falls in almost 50 percent of cases. Injury severity may be classified as mild (90 percent), moderate (7 to 8 percent), and severe (5 to 8 percent).

Pathophysiology

Diffuse injury and brain swelling that result in intracranial hypertension is more common in children than adults whereas focal injuries including intracerebral, subdural, and epidural hematomas are less common in children versus adults. Children have a higher frequency than adults of lesions in frontal white matter, orbitofrontal, and dorsolateral frontal regions. The reader is referred to a review by Kochanek and colleagues (2010) of initial assessment, management, and treatment of pediatric TBI.

Relationship of Psychiatric Disorders and Pediatric TBI

Preinjury Psychiatric Disorders. Prospective studies that used standardized psychiatric interviews found 33 to 50 percent of consecutively treated children had a preinjury lifetime psychiatric disorder. Birth cohort prospective data have also supported a disproportionate increase in preinjury behavior problems in children who go on to have an injury compared with uninjured children. This finding may be related to higher impulsivity and risk taking in children with externalizing disorders.

Postinjury Psychiatric Disorders. New-onset psychiatric disorders have been documented to be related to injury variables. For example, they occurred in 54 to 63 percent of children after severe TBI, in 10 to 21 percent after mild/moderate TBI, and in 4 to 14 percent in children with orthopedic injury without TBI. The recent finding that new-onset psychiatric disorder is related to lower fractional anisotropy in bilateral frontal and temporal lobes, bilateral centrum semiovale, and bilateral uncinated fasciculi suggests that diffuse white matter damage is central to the emergence of psychiatric disorders after TBI. Noninjury variables have also been related to new-onset psychiatric disorders and these included preinjury psychiatric disorders, socioeconomic status/preinjury intellectual function, preinjury adaptive function, preinjury family function, and family psychiatric history. The relationship of new-onset psychiatric disorder or worsening of behavioral symptomatology and family function has been shown to be a bidirectional process. Across the range of injury severity, family outcome is associated with lower preinjury family function, presence of new-onset psychiatric disorder, fewer sources of support, and more stressors.

Mild TBI in children and adolescents has become an area of public concern especially in the area of sports injury. Extant psychiatric studies have yielded conflicting results and have been inadequate in design and/or power to be definitive in understanding the risk of mild TBI causing new-onset psychiatric disorder. However, it has been demonstrated that children with mild TBI are more likely than orthopedically injured controls to display transient or persistent increases in postconcussive symptoms in the first year after injury especially with more serious injuries. The families of children with mild TBI experienced more distress and burden than their orthopedic injury counterparts and this was closely related to postconcussive symptoms.

Specific New-Onset Psychiatric Disorders

PERSONALITY CHANGE DUE TO TRAUMATIC BRAIN INJURY. Personality change (PC) due to TBI is the most common and important new-onset psychiatric disorder especially after severe TBI. Despite its name it is not a personality disorder but rather a disorder of clinically significant emotional and/or behavioral problems including affective instability, aggression,

disinhibition, apathy, and paranoia. Persistent PC has been found in approximately 40 percent of consecutively hospitalized children with severe TBI an average of 2-years postinjury. PC (diagnosed by waiving a 1 year duration criterion) occurred in 5 percent of mild/moderate TBI but was always transient suggesting that the duration of symptoms criterion has validity. The most common subtypes are those with significant affective instability, aggression, and disinhibition presentations. Severity of injury and superior frontal gyrus lesions are related to PC in the first year postinjury consistent with models of affective regulation implicating dorsal prefrontal cortex. Second year postinjury correlates of PC include severity of injury, frontal white matter lesions, and preinjury adaptive function suggesting that recovery is limited by network damage and related to preinjury reserve of functioning. PC is significantly comorbid with postinjury or secondary ADHD (SADHD).

SECONDARY ATTENTION-DEFICIT/HYPERACTIVITY DISORDER. SADHD is the term given to ADHD that develops after TBI. The phenomenology includes all subtypes of ADHD although inattentive symptomatology is most pronounced. SADHD is related to increasing severity of injury and concurrent adaptive and intellectual function, memory deficits, and family dysfunction in samples ranging from severe to mild TBI. In samples limited in range from moderate to severe, the findings are mixed except for adaptive function differences still evident. SADHD and postinjury ADHD symptoms have been associated with greater psychosocial adversity and poorer preinjury family function. Imaging correlates of SADHD included orbital frontal gyrus, right putamen, and thalamic lesions suggesting that damage to different locations along the same cortico-striatal-pallidal-thalamic loop can result in a similar phenotypic expression. A recent study of neurocognitive correlates of SADHD has shown that children with SADHD exhibited deficient attention, working memory, and psychomotor speed compared to children with preinjury ADHD. In addition, children with SADHD and the injured children who had neither SADHD nor preinjury ADHD were impaired compared to the children with preinjury ADHD with regard to planning. There were no group differences related to response inhibition which is considered a signature deficit in preinjury ADHD.

OPPOSITIONAL DEFIANT DISORDER/CONDUCT DISORDER. New-onset ODD and conduct disorder and symptoms of these disorders have been associated with preinjury psychosocial adversity and concurrent emotional lability symptoms. ODD symptoms have also been found to be associated with preinjury ODD symptoms, social class, and preinjury family function. Severity of injury predicted ODD symptoms after the first year postinjury and significantly accounted for change in ODD symptoms from before

injury.

POSTTRAUMATIC STRESS DISORDER. PTSD can occur notwithstanding neurogenic amnesia. Studies have documented that new-onset PTSD occurs in 4 to 13 percent of pediatric TBI but that the presence of subsyndromal PTSD symptoms is quite common (68 percent) in the weeks after injury before declining to 12 percent by 24 months postinjury. PTSD symptoms were related to preinjury anxiety or mood disorder/symptoms, greater injury severity, postinjury anxiety and depressive symptoms, female gender, preinjury psychosocial adversity, and nonanxiety psychiatric disorders. Re-experiencing symptoms were associated with relatively intact right cingulum and hippocampus possibly because these structures are necessary in the elaboration of these particular symptoms. Hyperarousal symptoms were associated with relatively intact left orbitofrontal lesions and were associated with left temporal lesions.

OBSESSIVE COMPULSIVE DISORDER AND OTHER ANXIETY DISORDERS. New-onset obsessive compulsive disorder has been associated with frontal and temporal lobe lesions and has occurred in the absence of striatal injury. New-onset obsessive symptoms were associated with psychosocial adversity, female gender, and mesial frontal and temporal lesions. Other new-onset anxiety disorders have included specific phobia, generalized anxiety disorder, and separation anxiety disorder. Preinjury anxiety symptoms and younger age at injury were related to postinjury anxiety symptoms. In addition, postinjury level of stress and severity of brain injury were associated with mood/anxiety disorders. Superior frontal gyrus lesions have been associated with postinjury anxiety symptoms suggesting that damage to this region can result in affective dysregulation that includes anxiety as well as the affective instability of PC. One study found an association between fewer anxiety symptoms and orbitofrontal lesions.

MANIA/HYPOMANIA. Manic and hypomanic syndromes have occurred after pediatric TBI documented in case reports and in a prospective study in 8 percent of consecutively injured children. Careful consideration of the overlapping and more common syndromes of PC and SADHD should be taken before diagnosing a manic syndrome. Manic/hypomanic syndromes have been associated with frontal and temporal lesions, greater severity of injury, and a family history of mood disorder. The episodes are usually long lasting and there has been a similar frequency of irritability versus elation as the primary form of affective symptomatology.

DEPRESSIVE DISORDER. New-onset depressive disorder has occurred in 11 to 33 percent in previous studies. Children with social disadvantage were at higher risk to develop depressive symptoms. New-onset depressive disorder can be classified as nonanxious depression (no comorbid new-

onset anxiety disorder) or anxious depressive disorder (comorbid new-onset anxiety disorder). Nonanxious depression was related to left inferior frontal gyrus and left temporal tip lesions, whereas anxious depression was related to right frontal lobe white matter and left parietal lesions. Anxious depression was related to PC and a family history of anxiety disorder.

AUTISM AND PSYCHOSIS. Nonaffective psychosis has been documented in only two cases in studies of consecutive admissions of children with TBI that used standardized psychiatric interviews. Autism has not been described after pediatric TBI in contrast to other forms of early brain injury. It is important clinically to distinguish autism from qualitatively distinct aberrations of pragmatic communication and social communication that have been complications of severe TBI in young children.

Assessment

A useful model suggested by Asarnow and colleagues (1991) that has guided the clinical approach to the psychiatric assessment of children and adolescents with a history of TBI involves six pathways. (1) The psychiatric problem predates the injury and may increase the risk of its occurrence. (2) A preinjury psychiatric problem is worsened by the TBI. (3) The psychiatric problem is a direct (biological) effect of the TBI. (4) The psychiatric problem is an immediate secondary effect of the TBI (e.g., a traumatic emotional response to the incident such as PTSD). (5) The psychiatric problem is a long-term secondary effect of the TBI (e.g., disruptive behavior problems and lowered motivation from frustration resulting from new-onset cognitive problems). (6) The psychiatric problems are caused by factors other than the TBI.

Intervention

Prevention. Prevention is the most critical intervention with regard to pediatric TBI. This includes education with respect to use of bicycle helmets, improving motor vehicle safety, limiting alcohol-related motor vehicle accidents, and reducing the risk of child abuse and neglect.

Nonpharmacological Treatment

FAMILY THERAPY. Web-based family and teen problem-solving therapies have been developed to target behavioral problems after TBI. These included counselor-assisted problem solving (CAPS) which is a 6-month web-based, family-centered treatment that focused on problem solving, communication, and self-regulation. CAPS and a similar treatment called teen online problem solving (TOPS) have been effective in multiple ways including improving parent-rated behavioral executive functioning and

externalizing problems in the children, as well as enhancing problem solving and decreasing depressive symptoms in parents. Other effective approaches have included CBT and the restructuring of the child's environment to optimize behavioral supports.

SCHOOL. The Individuals with Disabilities Education Act entitle children especially with severe TBI to services. Furthermore special education interventions target poor academic function related to a variety of permutations of academic skills deficits, emotional and behavioral problems, which may compound corresponding preinjury problems. For a more detailed review of educational and neuropsychological correlates and interventions for pediatric TBI, the reader is referred to Crowe and colleagues (2015).

PHARMACOLOGICAL THERAPIES. A comprehensive review of pharmacological treatment in pediatric brain injury by Pangilinan and colleagues (2010) has concluded that evidence in support of "off label" use of many agents is extremely limited. Anecdotal guidelines only are therefore noted here. Mood stabilizing anticonvulsants and SSRIs can be helpful together with behavioral treatment for affective instability and aggression typical with PC. The disinhibited type of PC is difficult to treat. Neuroleptic medication may be helpful for the agitated or paranoid patient when the behavior is complicating rehabilitation and recovery although usage must be balanced against potential risks in the alteration of neuronal recovery. There have been several positive reports of stimulant use for children with ADHD-like symptoms after TBI although overall findings are mixed. In the absence of research data, depressive disorders after TBI may be treated with SSRIs as in corresponding syndromes in the absence of TBI.

Conclusion

Pediatric TBI is a major cause of morbidity and mortality and is thus an important public health problem. Particularly with severe TBI the sequelae may be life-long. Mental health professionals have a key role to play in this field because psychiatric disorders are common both before and after injury. As in circumstances unrelated to TBI, psychiatric disorders and family function have a close relationship. New-onset psychiatric disorders are associated with injury and psychosocial variables that can be recorded soon after injury and this opens potential options for secondary prevention and early recognition of predictable outcomes. Neurocognitive, school, and adaptive function problems often occur after severe TBI and effective treatment for the patient must address each of these domains of function. Research in psychosocial treatments is more advanced than psychopharmacological approaches.

OVERALL SUMMARY

There are several striking similarities and differences with regard to psychopathology in children and adolescents with the varying brain injuries discussed in chapter. Psychiatric disorders occur commonly in each of the brain injury groups the authors have reviewed. The rates of psychiatric disorder are at least 27 percent (in uncomplicated epilepsy) and generally higher in the range of 50 to 60 percent in children with the different kinds of brain injury reviewed. The psychiatric disorders that manifest are typically heterogeneous, that is, there is no signature disorder that occurs in all brain injured children. Comorbidity is also the rule rather than the exception. The disorder most specific to brain injury is PC due to a general medical condition which has been formally studied only in pediatric TBI. Aside from this disorder it is quite clear that the most common disorder in brain injured children is ADHD typically around 25 to 30 percent except for FASDs where ADHD occurs in at least double that rate. The very high rate of ADHD associated with prenatal alcohol exposure may be due to a confluence of factors that include a particularly virulent teratogen, genetic factors, and poverty. There is accumulating evidence that ADHD associated with brain injury has different neuropsychological characteristics compared with developmental ADHD. Other psychiatric disorders such as depressive, anxiety, and ODDs that are present occur at similar high rates across brain injury categories. Autism is most closely associated with prematurity and LBW and not the other brain disorders. While it is true that children with autism have a relatively high rate of epilepsy, only a small proportion of children with epilepsy have autism. Standardized psychiatric interviews are more sensitive in detecting psychiatric disorders compared with behavior checklists in brain injured pediatric populations possibly because parents who complete the checklists have accommodated their child's limitations and often do not keep in mind a normative expectation.

IQ tends to be low average in children who are significantly affected by the brain injuries discussed in this chapter. Intellectual disability or IQ ≤70 is unusual. Deficits occur in multiple neuropsychological domains and again there is no signature neuropsychological deficit profile across brain injury groups.

Treatment for the psychiatric disorders associated with the brain injury groups generally rely on treatments designed for children and adolescents in general or even those used in adults with brain injury. This will remain the standard of care until more intervention research is conducted. The public health impact of brain injury–related psychopathology in children and adolescents is high because of its high frequency as well as the many years of illness burden when brain injury is acquired at a young age. There is an exceptional need for mental health researchers and clinicians to understand, develop treatments, and to intervene to improve the lives of

young people with brain injury.

References

Aarnoudse-Moens CS, Weisglas-Kuperus N, van Goudoever JB, Oosterlaan J. Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics*. 2009;124:717–728.

Anderson PJ. Neuropsychological outcomes of children born very preterm. *Semin Fetal Neonatal Med*. 2014;19:90–96.

Asarnow RF, Satz P, Light R, Lewis R. Behavior problems and adaptive functioning in children with mild and severe closed head injury. *J Pediatr Psychol*. 1991;16:543–555.

*Berg AT, Caplan R, Hesdorffer DC. Psychiatric and neurodevelopmental disorders in childhood-onset epilepsy. *Epilepsy Behav*. 2011;20:550–555.

*Brown LK, Whiteley L, Harper GW, Nichols S, Nieves A. Psychological symptoms among 2032 youth living with HIV: A multisite study. *AIDS Patient Care STDS*. 2015;29(4):121–219.

Crocker N, Riley EP, Mattson SN. Visual-spatial abilities relate to mathematics achievement in children with heavy prenatal alcohol exposure. *Neuropsychology*. 2015;29:108–116.

Crowe LM, Catroppa C, Anderson V. Sequelae in children: developmental consequences. *Handb Clin Neurol*. 2015;128:661–677.

Dunn DW, Besag F, Caplan R, et al. Psychiatric and Behavioural Disorders in Children with Epilepsy (ILAE Task Force Report): Anxiety, depression and childhood epilepsy. Epileptic Disord. 2016.

Fastenau PS, Johnson CS, Perkins SM, et al. Neuropsychological status at seizure onset in children: risk factors for early cognitive deficits. *Neurology*. 2009;73:526–534.

Ferrara M, Valero IP, Moore DJ, Knight AF, Duarte NA, Atkinson JH. Treatment of psychiatric disorders in HIV. In: Joska JA, Stein DJ, Grant I, eds. *HIV and Psychiatry*. Oxford, UK: John Wiley & Sons Limited; 2014:175–193.

Fisher RS, Acevedo C, Arzimanoglou A, et al. ILAE official report: a practical clinical definition of epilepsy. *Epilepsia*. 2014;55:475–482.

Hoyme, HE, WO. Kalberg, AJ Elliott, et al. Updated Clinical Guidelines for Diagnosing Fetal Alcohol Spectrum Disorders. Pediatrics. 2016;138(2):2015–4256. doi: 10.1542/peds.

*Johnson S, Marlow N. Growing up after extremely preterm birth: lifespan mental health outcomes. *Semin Fetal Neonatal Med*. 2014;19:97–104.

Joseph RM, O'Shea TM, Allred EN, et al. Prevalence and associated features of autism spectrum disorder in extremely low gestational age newborns at age 10 years. Autism Res. 2016; doi: 10.1002/aur.1644.

Kochanek PM, Bell MJ, Bayir H. Quo vadis 2010? - carpe diem: challenges and opportunities in pediatric traumatic brain injury. *Dev Neurosci*. 2010;32:335–342.

LaFrance WC Jr., Baker GA, Duncan R, Goldstein LH, Reuber M. Minimum requirements for the diagnosis of psychogenic nonepileptic seizures: a staged approach: a report from the International League Against Epilepsy Nonepileptic Seizures Task Force. *Epilepsia*. 2013;54:2005–2018.

*Mattson SN, Roesch SC, Glass L, et al. Further development of a neurobehavioral profile of fetal alcohol spectrum disorders. *Alcohol Clin Exp Res*. 2013;37:517–528.

Max JE. Neuropsychiatry of pediatric traumatic brain injury. *Psychiatr Clin North Am*. 2014;37:125–140.

*Max JE, Wilde EA, Bigler ED, et al. Neuroimaging correlates of novel psychiatric disorders after pediatric traumatic brain injury. *J Am Acad Child Adolesc Psychiatry*. 2012;51:1208–1217.

Max JE, Wilde EA, Bigler ED, et al. Personality Change due to Traumatic Brain Injury in Children and Adolescents: Neurocognitive Correlates. J Neuropsychiatry Clin Neurosci. 2015;27(4):272–279.

May PA, Baete A, Russo J, et al. Prevalence and characteristics of fetal alcohol spectrum disorders. *Pediatrics*. 2014;134:855–866.

Mellins CA, Malee KM. Understanding the mental health of youth living with perinatal HIV infection: lessons learned and current challenges. *J Int AIDS Soc*. 2013;16:18593.

Nichols SL, Bethel J, Garvie PA, et al. Neurocognitive functioning in antiretroviral therapy-naive youth with behaviorally acquired human immunodeficiency virus. *J Adolesc Health*. 2013;53:763–771.

Pangilinan PH, Giacoletti-Argento A, Shellhaas R, Hurvitz EA, Hornyak JE. Neuropharmacology in pediatric brain injury: a review. *PM R*. 2010;2:1127–1140.

Saigal S. Functional outcomes of very premature infants into adulthood. *Semin Fetal Neonatal Med*. 2014;19:125–130.

Salpekar JA, Mishra G. Key issues in addressing the comorbidity of attention deficit hyperactivity disorder and pediatric epilepsy. *Epilepsy Behav*. 2014;37:310–315.

Vaurio L, Riley EP, Mattson SN. Neuropsychological comparison of children with heavy prenatal alcohol exposure and an IQ-matched comparison group. *J Int Neuropsychol Soc*. 17:463–473.

Volpe JJ. The encephalopathy of prematurity–brain injury and impaired brain development inextricably intertwined. *Semin Pediatr Neurol*. 2009;16:167–178.

Wainberg ML, McKinnon K, Cournos F. Epidemiology of psychopathology in HIV. In Joska JA, Stein DJ, Grant I, eds. *HIV and Psychiatry*. Oxford, UK: John Wiley & Sons Limited; 2014:175–193.

Ware AL, Glass L, Crocker N, et al. Effects of prenatal alcohol exposure and attention-deficit/hyperactivity disorder on adaptive functioning. *Alcohol Clin Exp Res*. 2014;38:1439–1447.

Ware AL, O'Brien JW, Crocker N, et al. The effects of prenatal alcohol exposure and attention-deficit/hyperactivity disorder on psychopathology and behavior. *Alcohol Clin Exp Res*. 2013;37:507–516.

▲ 2.14 Neuropsychiatry of Neurometabolic and Neuroendocrine Disorders

MARK WALTERFANG, PH.D., RAMON MOCELLIN, M.B.B.S., M.SC., M.PSYCH. (MED), F.R.A.N.Z.C.P., OLIVIER BONNOT, M.D., PH.D., AND DENNIS VELAKOULIS, DMEDSCI

Metabolic and endocrine disturbance can have wide-ranging effects on the CNS. Neurons, because of their high metabolic level of activity and demand, are often acutely sensitive to derangements in more systemic metabolic processes. As a result, disorders affecting systemic metabolism often have a very high rate of associated CNS disturbance, ranging from the severe—with gross neurodevelopmental disruption, delirium, and/or coma—to the mild, with subtle cognitive and behavioral disturbance. However, a number of these disorders cause recognizable and significant psychiatric illness, including psychotic disorders, affective disturbance, anxiety disorders, attention-deficit disturbance, and other behavioral disturbance such as apathy, dysexecutive syndromes, and catatonia. For the purposes of this section, those disorders of metabolism and endocrine function associated with significant CNS disturbance will be referred to as *neurometabolic* and *neuroendocrine* disorders, respectively, and those associated with major neuropsychiatric syndromes will be highlighted.

The relevance of recognition of neuropsychiatric comorbidity in neurometabolic and neuroendocrine syndromes extends to both diagnostic and treatment issues. A number of these disorders can produce an accurate *phenocopy* of psychiatric illness, where illness presentation carries most or all the characteristic features of a psychiatric diagnosis. For example, a schizophrenia-like psychosis may develop in metachromatic leukodystrophy (MLD) that is otherwise indistinguishable from "typical" schizophrenia, at least until other features more typical of the primary neurometabolic illness supervene. Awareness of the types of disorders that may present as psychiatric illness phenocopies, and their associated physical, cognitive, or neurological concomitants, allows for the appropriate recognition and diagnostic confirmation of an underlying metabolic or endocrine illness. In some circumstances, this may mean medical treatment of an endocrine illness, enzyme replacement, or substrate reduction therapy for a metabolic illness. Failure to recognize the underlying illness while focusing on psychiatric treatment alone can delay the institution of appropriate management and result in potentially irreversible CNS changes.

Most clinicians will undertake a routine set of investigations for underlying "organic" causes of psychiatric illness at first presentation, including screening for thyroid illness, infective disease, electrolyte

disturbance, and renal and kidney function, in addition to brain imaging and/or electroencephalography (EEG). Although these tests are relatively high yield and detect a number of illnesses producing secondary psychiatric syndromes, less common but no less significant conditions are often missed. Literally hundreds of metabolic disorders have been described, and only a very limited number of these are screened for in the neonatal period (usually less than ten in most centers). This may provide the clinician with a false assurance that any significant and/or serious metabolic disorders will already have been detected. This situation is further complicated by the fact that the age of onset and presentation, in addition to disease expression, vary greatly across and within these disorders, and disorders that present later or in a less severe form are often missed by screening processes during the neonatal and childhood periods.

Those metabolic disorders that present initially, or predominantly, with neuropsychiatric syndromes generally have their onset in the period of the life cycle associated with the onset of the majority of psychiatric illnesses, that is, adolescence or early adulthood. The disorders that present in this stage will commonly be inborn errors of metabolism, affecting cellular function in the CNS. A wide range of phenotype–genotype correlations occur in such disorders, as differing mutations in the gene for the protein serving a key function in metabolism will result in different structural or conformational changes to the protein product and differential metabolic effects. For example, some mutations in the gene encoding for the NPC1 protein in Niemann–Pick disease type C are associated with an early onset or illness, severe mental retardation, and death in childhood, whereas others are associated with a less significant protein defect, a reduced effect on NPC1's role in intracellular cholesterol metabolism (even falling into the "normal" range), and a presentation in adolescence or early adulthood. The adult-onset forms of these diseases often form a minority of cases of these disorders, although they will generally respond as favorably to primary treatments for the disorder. In addition, some carriers with autosomal recessive disorders who were previously considered asymptomatic (such as female carriers with adrenomyeloneuropathy) have been shown to have "borderline" syndromes often involving subtle psychiatric disturbance. For most of these disorders, gross metabolic disturbance results in severe impairment of CNS function (such as delirium or coma), with moderate impairments often resulting in dementia (or mental retardation in children) and movement disorders. More subtle impairments, however, are less likely to disrupt these core vegetative and functional systems of the brain, but rather disrupt higher-order functions that are much more dependent on highly synchronous corticocortical and subcortical connectivity, with the result being disturbance to higher cognitive functions and a predisposition toward major mood and psychotic disorders. As these systems tend to mature later in the brain's neurodevelopmental trajectory,

those disorders that alter or interrupt late neurodevelopment are more likely to cause neuropsychiatric syndromes. For progressive disorders, it is not uncommon for these disorders to present initially with a neuropsychiatric syndrome, which may be diagnosed and treated as a primary psychiatric illness. As the pathological effect of the metabolic derangement impinges further on the CNS and begins to result in degenerative change, frank neurological illness and dementia often supervene.

In addition to those metabolic disorders that impact on neurodevelopment, some metabolic disorders are associated with episodic but reversible metabolic disturbances (such as the acute porphyrias). As opposed to impacting late-maturing developmental networks, these disorders impact metabolically on presumably mature or normally developed brain systems. Those disorders that show a predilection for neuropsychiatric disturbance are more likely to affect brain regions (such as the frontal or temporal cortical regions) or transmitter systems (particularly dopaminergic and serotonergic systems) that are strongly associated with psychiatric illness. Similarly, the majority of major endocrine disorders that are associated with psychiatric disturbance do so in the setting of an otherwise intact CNS. These endocrine systems are crucially involved in energy metabolism, cell turnover, and downstream metabolic effects. Disorders associated with elevated rates of psychiatric disturbance act via a direct effect of the altered hormonal system on particularly vulnerable cellular populations or neurotransmitter systems such as the effects of elevated cortisol levels on hippocampal neurons, or thyroid hormone on serotonin turnover.

NEUROMETABOLIC SYNDROMES

Lysosomal Disorders

The lysosome is a subcellular organelle that is manufactured by the Golgi apparatus and contains a number of hydrolytic enzymes such as proteases, lipases, nucleases, and polysaccharidases. They function as the “garbage disposal” system of the cell, via phagocytosis (digestion of extracellular material), endocytosis (digestion of cell surface proteins), and autophagy (digestion of old or damaged intracellular organelles or structures). Macromolecules (proteins, glycoproteins, lipids, and phospholipids) transported to lysosomes are degraded by enzymatic “factories,” which then pass out their monomeric components for reutilization. Impairment in function of a lysosomal enzyme occurs if it is structurally altered, alteration occurs to a cofactor protein, or enzyme transport is affected. Each enzyme is specific for breaking a particular chemical bond, as opposed to a particular substrate macromolecule. Over 70 lysosomal enzymes are known, and more than 40 disease syndromes involving

defective enzyme function have been characterized. Many of these disorders present both in childhood and adulthood, and for those that present in adolescence or early adulthood—the period of onset of most major mental disorders—the rate of major mental illness is not unexpectedly elevated, whereas childhood presentations commonly result in major intellectual disability and older-adult presentations in dementia, in addition to frank neurological disturbance. Of those lysosomal storage disorders (LSDs) that present in adult life, those strongly associated with neuropsychiatric presentations include those involving defective breakdown of sphingolipid components (MLD, Fabry disease, GM2-gangliosidosis/Tay–Sachs disease [TSD], Niemann–Pick type C disease [NPC]), glycoproteins (α-mannosidosis), and cholesterol and lipids (neuronal ceroid lipofuscinosis [NCL] or Kuf disease).

Metachromatic Leukodystrophy

MLD is an autosomal recessive, incompletely penetrant genetic deficiency of the lysosomal enzyme arylsulfatase A (ASA). ASA hydrolyzes various sulfatides, including sulfate-containing lipids of the CNS. Lysosomal sulfatide accumulates in brain, peripheral nerves, kidney, and gallbladder, but particularly in myelinated structures, seen as metachromatic granules on histological examination and widespread loss of myelin. MLD is protean in its presentations such that in younger patients, seizures and motor symptoms predominate with psychiatric manifestations and dementia occurring frequently prior to the onset of motor symptoms in adult patients, particularly those with the I179S mutation. The adult form appears to cleave into two distinct phenotypes, one with a predominantly motor cerebellopyramidal presentation, and the other with a predominantly psychiatric presentation. Up to half of patients with illness onset between 10 and 30 years of age present with psychotic symptoms, including auditory hallucinations, systematized delusions, formal thought disorder, catatonic posturing, and inappropriate affect. As the illness progresses, other neurological symptoms supervene, such as seizures, chorea, or dystonia. Dementia is an inexorably progressive component of the illness, initially presenting as attentional disturbance, reduced speed of processing and executive impairment, and may present as a phenocopy of frontotemporal dementia; if disinhibition is a prominent feature, it may present as a phenocopy of hypomania. Diagnosis is made by demonstrating reduced enzyme activity in leukocytes or skin fibroblasts. MRI generally demonstrates typical periventricular white matter changes sparing subcortical U-fibers, and often shows pathology with a frontotemporal preponderance (Fig. 2.14–1).

Treatment is generally symptomatic, although bone marrow transplantation has shown benefit in some patients, and enzyme replacement therapies are currently being investigated. Adolescent/adult

MLD provides an intriguing model for the understanding of the neurobiology of psychosis, as it interrupts myelinative processes that occur during this critical period of neurodevelopment, in particular frontotemporal myelination. As frontotemporal connectivity is known to be impaired at both functional and structural levels in schizophrenia, MLD appears to have an almost uniquely psychotogenic propensity in this age group, and suggests that a neuropathological process that interrupts the normal development of connectivity between these cortical regions in this crucial late neurodevelopmental window can result in psychosis. It has also been shown that some treatment-refractory schizophrenia patients have ASA pseudo-deficiency or intermediate ASA levels, suggesting that this may be a disease modifier or potentially an independent risk factor for the development of psychosis.

Fabry Disease

Fabry disease is an X-linked recessive disorder of the lysosomal enzyme α-galactosidase A, resulting in accumulation of the glycolipid globotriaosylceramide (Gb3) in blood vessels and other tissues (Fig. 2.14–2). It particularly affects hemizygous males, with females varying from asymptomatic to severely affected because of X inactivation. The major clinical manifestations of Fabry disease reflect the particular impact of the disorder on vascular endothelium. Early clinical features with onset in childhood or adolescence include angiokeratomas and acroparesthesias. Angiokeratomas are dark, punctate vascular lesions most prevalent between the umbilicus and the knees but also seen on the oral mucosa and conjunctiva. Acroparesthesias are acute episodes of severe pain in the fingers and toes lasting days to weeks precipitated by exercise, fatigue, or fever, which are very disabling. Chronic pain syndromes associated with acroparesthesias may continue into adulthood and account for much of the associated psychiatric comorbidity. Corneal or lenticular opacities and hypohidrosis are also early manifestations. Disease progression is marked by vascular disease, resulting in kidney impairment, cardiovascular disease, and stroke in adulthood. MRI often demonstrates subcortical changes, in periventricular white matter, and the basal ganglia, thalamus, and brainstem. Cognitive impairment may be detected via formal neuropsychological testing, and often presents as impaired processing speed, attention, and executive functioning, likely reflecting the impact of subcortical vascular ischemia. Depressive disorders, often meeting criteria for a severe clinical depression, occur in up to half of all sufferers and is most strongly associated with the degree of peripheral pain and anhidrosis, suggesting that it reflects the burden of disease rather than being a poststroke phenomenon or associated with subcortical vascular change. Anxiety is also commonly reported; other psychiatric syndromes appear relatively rare. The degree of depression and anxiety also appears to predict

the degree of social and occupational dysfunction. Symptomatic treatment of depression is often effective; however, enzyme replacement therapy has recently become available and may prevent some of the physical manifestations of Fabry disease that appear to be causally related to depression in this disorder.

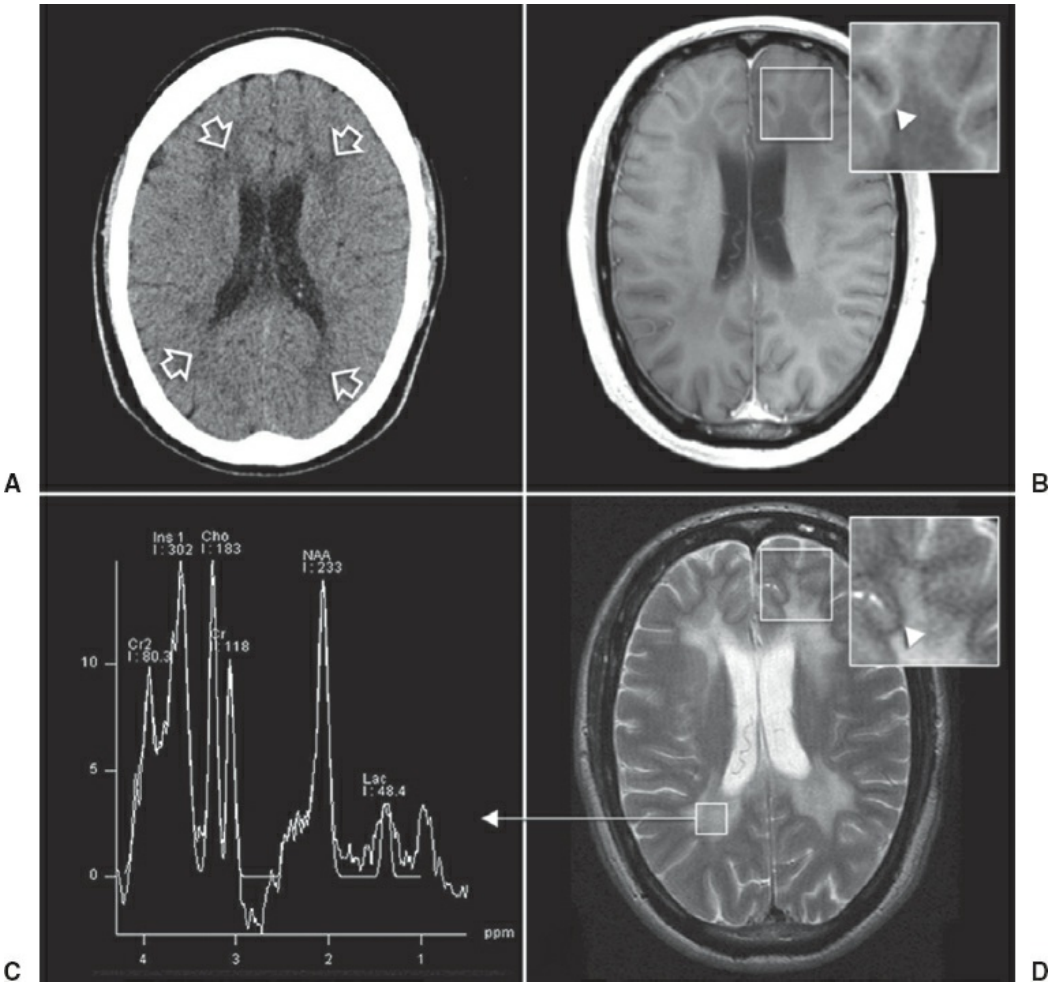

FIGURE 2.14–1. Typical imaging findings of MLD. **A:** Noncontrast CT brain demonstrates confluent bilateral periventricular deep white matter hypodensity (*arrows*) which is hypointense on T1 imaging without contrast enhancement (**B**) and hyperintense on T2 imaging (**D**). The abnormality spares the subcortical U fibers (*arrowheads,* inset **B,D**). Spectroscopy trace (**C**) demonstrates elevated myo-inositol, reduced N-acetylaspartate and increased lactate.

Gaucher Disease

Gaucher disease (GD) is the most common hereditary LSD, due to a deficiency in β-glucocerebrosidase (GBA1) which results in the accumulation of glucocerebroside in macrophage lysosomes. GD carrier status has more recently been shown to be the most common genetic risk factor for Parkinson disease, because of the molecular relationship between glucocerebrosides and α-synuclein. These are deposited in a variety of tissues, including marrow, liver, spleen, and brain. Type 1 GD is

"nonneuropathic" as the brain and spinal cord are thought to be spared; this presents with hepatosplenomegaly, anemia, leukopenia, and thrombocytopenia in addition to skeletal bone lesions. Types 2 and 3 affect the CNS, with type 2 causing progressive severe disease in infancy and type 3 being more indolent and often associated with survival into adulthood, with the development of seizures, ataxia, and gaze palsy being common neurological presentations. It is known however that even GD1 patients have some neuronal involvement, with astrogliosis and α-synuclein deposition occurring in neuronal populations, but with only type 1 patients being spared neuronal loss. GD1 patients may demonstrate subtle cognitive deficits, largely in attention and memory; in GD3 patients, cognitive function can vary widely—from severe impairment to normal functioning—according to the severity of the disease. Depression and anxiety are not uncommon in GD1 patients, at approximately three times the rate of the normal population. Depression in GD patients is known to respond to antidepressant medication. Enzyme replacement therapy in GD is associated with significant improvement in mood and psychological functioning, with similar improvements reported in quality of life.

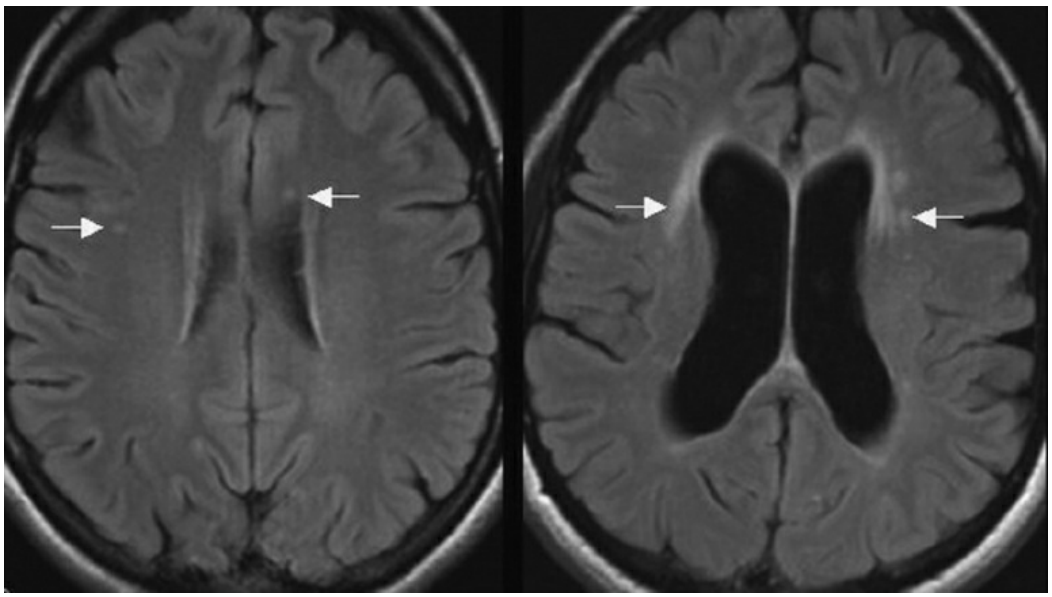

FIGURE 2.14–2. Fluid-attenuated inversion recovery magnetic resonance image demonstrating mild (**left**) white matter lesions (*arrows*) in a 39-year-old woman with Fabry disease, and more significant (**right**) white matter lesions in a 41-year-old woman with the same disease. (From Fellgiebel A, Müller MJ, Mazanek M, Baron K, Beck M, Stoeter P. White matter lesion severity in male and female patients with Fabry disease. *Neurology*. 65[4]:600, Figure 1, with permission.)

GM2 Gangliosidosis (Tay–Sachs Disease)

TSD is an autosomal recessive lipid storage disorder caused by the accumulation of GM2-gangliosides within neurons due to a deficiency in β-hexosaminidase A (HEX-A). HEX-A deficiency in lysosomes impairs the

catabolism of gangliosides from the neuronal cell membrane, resulting in accumulation of lysosomal gangliosides (Fig. 2.14–3). This leads to secondary axoneuronal changes, particularly axon hillock outgrowth to form "meganeurites" with ectopic dendritogenesis and focal axonal enlargements known as *axonal spheroids,* both of which may alter neuron-to-neuron microconnectivity. In addition, ganglioside accumulation results in direct neurotoxicity, altered neuronal electrical properties, inappropriate apoptosis, or an inflammatory response. In infantile or childhood forms, severe neurological impairment usually results in death within 3 to 10 years. A later, or adult-onset, form of TSD has been described and may be associated with having at least one allele (most commonly the G269S mutation) associated with residual enzyme activity. In this late-onset form, psychiatric symptoms may copresent with or predate the development of neurological disturbances in early adulthood. Patients present with speech disorder, gait disturbance, and tremor most commonly, with normal or near-normal cognitive function, although subtle deficits in executive function, processing speed, and memory may be present in up to half of patients. MRI often shows significant cerebellar atrophy, although cortical atrophy may also be seen. Neuropsychiatric presentations occur in up to half of late onset TSD patients, predominantly psychosis, which may occur in 30 to 50 percent of adult patients, marked by disorganization, auditory and visual hallucinations, and catatonia. Affective disorders including mania and depression have been occasionally reported. Patients only partially respond to neuroleptics or lithium (Eskalith), and are often very sensitive to motor side effects of these drugs. Importantly patients with psychotic illness appear to respond to ECT.

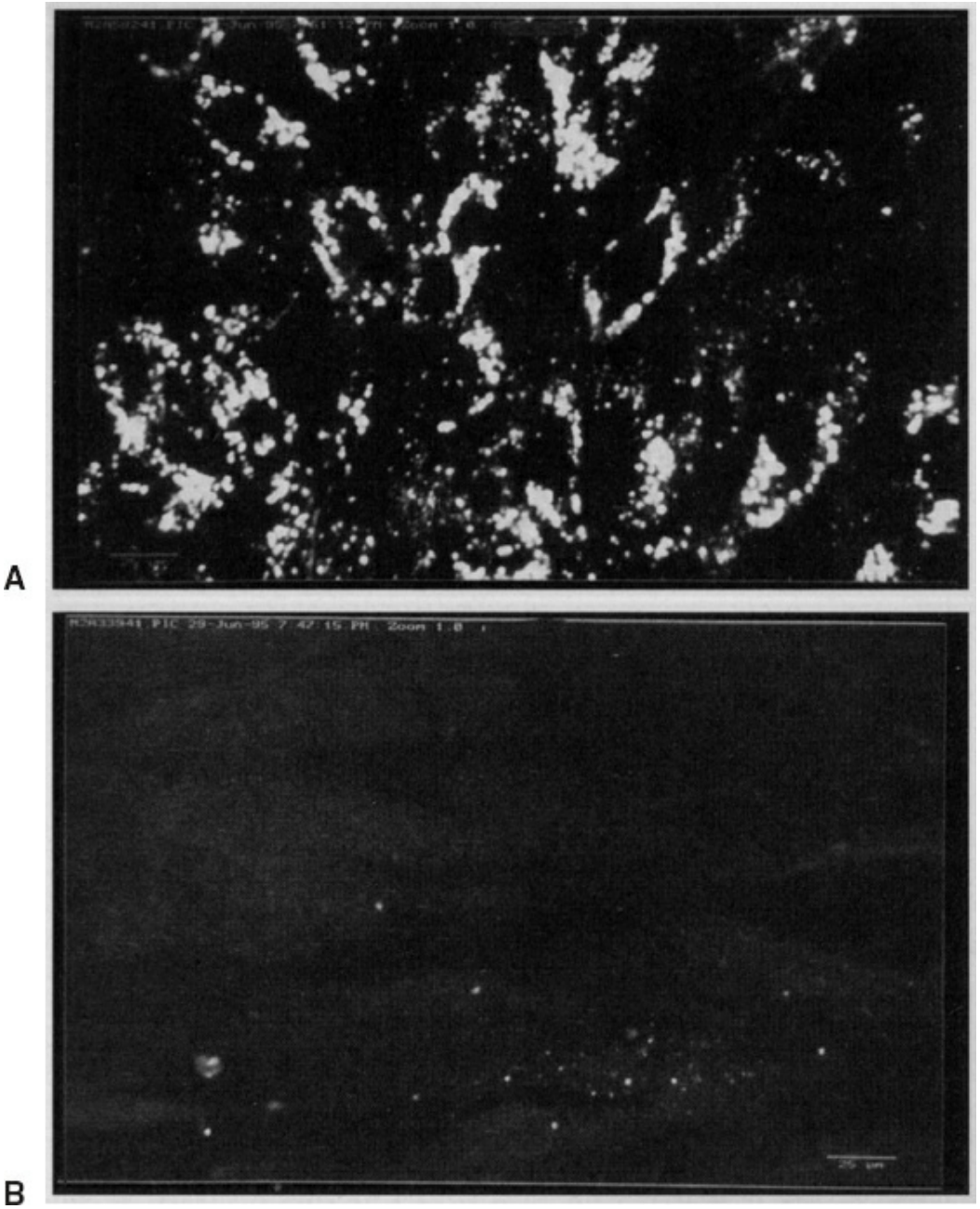

FIGURE 2.14–3. Indirect immunofluorescence staining of cultured fibroblasts with an anti-GM2 antibody in a patient with Tay–Sachs disease (**A**), showing significant accumulation of GM2 gangliosides compared to a healthy control (**B**). (From Sakuraba H, Itoh K, Shimmoto M et al. GM2 gangliosidosis AB variant: Clinical and biochemical studies of a Japanese patient. *Neurology*. 1999;52[2]:372, Figure 1, with permission.)

Neuronal Ceroid Lipofuscinosis

The NCLs are a group of neuronal storage disorders, one of which is Batten disease, the most common neurodegenerative disorder in childhood. Accumulation of lipofuscin-like material in lysosomes, particularly in neuronal tissue, characterizes NCL. Mutations in ten genes that encode for a range of proteins, all involved in lysosomal catabolism and recycling of proteins and lipids, have been shown to be responsible for this group of

disorders. Adult neuronal ceroid lipofuscinosis (ANCL), Kuf and Parry diseases may be inherited in a recessive or dominant fashion, respectively, and make up between 2 and 10 percent of all NCLs. Dominant pedigrees have been shown to have mutations in the DNAJC5 gene, and recessive pedigrees by mutations in PPT1, CLN5, and SGSH genes. Lipofuscinosis diffusely affects cortical and subcortical neurons, but particularly the hippocampus and entorhinal cortex, and is associated with significant glial activation that precedes rather than follows neuronal loss. Alteration in axonal and synaptic function due to ultrastructural change occurs across the CLNs. Symptoms most commonly appear at the beginning of the fourth decade, but may be present early in the second decade. Two clinical forms of ACNL have been described: a progressive myoclonic epilepsy form, and a neuropsychiatric/cognitive form characterized by dementia, psychiatric disturbance, and dyskinesia. Psychosis occurs in up to 20 percent of patients, particularly in adolescent and early adult patients, with characteristic hallucinations, delusions, and thought disorder, and occasionally catatonia. Mood disturbance is particularly common. Patients present cognitively with slowing, attentional disturbance, and impaired new learning; older patients may present with a frontotemporal dementia-like picture or a Klüver–Bucy-type syndrome. Diagnosis rests on identification of characteristic inclusions in skin punch biopsy or leukocytes. MRI often shows cerebral and cerebellar atrophy and callosal thinning whilst SPECT shows regional cortical hypoperfusion. Treatment is symptomatic, although these patients are very sensitive to extrapyramidal side effects such as dystonia and neuroleptic malignant syndrome, as NCL patients may have muscle membrane pathology comorbid with changes in the CNS.

α-Mannosidosis Type II

α-Mannosidosis (AM) is a recessively inherited LSD that results from deficiency of AM, which causes lysosomal accumulation or undigested oligosaccharides. AM is characterized by mild to moderate intellectual disability, hearing loss, skeletal changes, and recurrent infections, with an indolent form (type II) occurring in the minority of patients who survive to adulthood. Deficiency of AM results in the intralysosomal accumulation of mannose-rich oligosaccharides and the formation of storage vacuoles in neuronal and glial cells, which impairs myelin formation. Neuropathologically, AM appears to initially affect myelinated structures before progressing to involve the neuronal body. In type II AM, the predominant clinical features are cerebellar ataxia, hearing loss, neuropsychological impairment, and retinopathy. Diagnosis rests on the demonstration of a pattern of urinary oligosaccharides and reduced enzyme activity in leukocytes or fibroblasts. MRI scanning shows periventricular T2 hyperintensities and cortical and cerebellar atrophy (Fig.

2.14–4), and may show iron accumulation in the basal ganglia. Like many LSDs, bone marrow transplantation is the only current viable treatment option. Up to 25 percent of type II AM patients over the age of 15 develop clear mental illness, predominantly a psychotic disorder characterized by delusions, hallucinations, and confusion. Generally, psychosis presents with neurological manifestations, although it may rarely present as a prelude to frank neurologic disturbance.

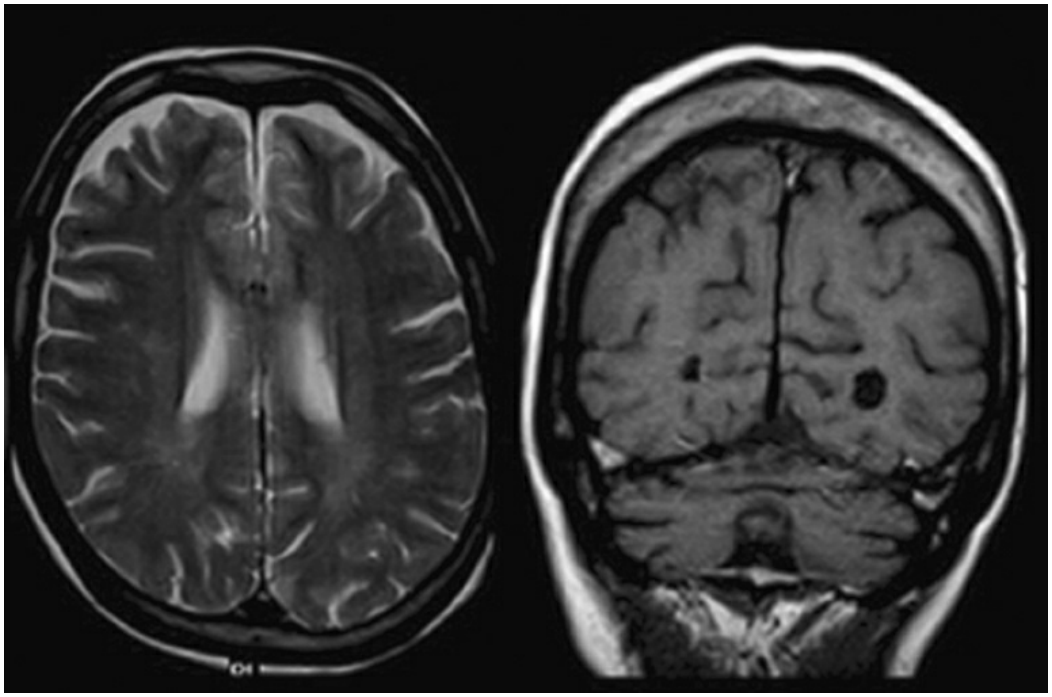

FIGURE 2.14–4. Magnetic resonance imaging scan on a patient with psychosis and α-mannosidosis. **Left:** T2-weighted imaging shows parietal and occipital white matter hyperintensities and frontal atrophy. **Right:** Coronal T2-weighted imaging at the level of the cerebellum shows atrophy of the cerebellar hemispheres. (From Gutschalk A, Harting I, Cantz M, Springer C, Rohrschneider K, Meinck H-M. Adult α-mannosidosis: Clinical progression in the absence of demyelination. *Neurology*. 2004;63[9]:1744, Figure 1, with permission.)

Peroxisomal Disorders

The peroxisome is a subcellular organelle that plays a role in the breakdown of fatty acids, the degradation of hydrogen peroxide (released via oxidation of fatty acids), membrane phospholipid and cholesterol synthesis, and the metabolism of amino acids. Unlike lysosomes, the peroxisomes bud off from the endoplasmic reticulum. Disorders of biogenesis (formation) of the peroxisome are generally fatal during infancy. Defects in single enzymes of the peroxisome may also cause disease compatible with survival well into adult life, and at least one of these, X-linked adrenoleukodystrophy (X-ALD), the most common peroxisomal disorder, is known to be associated with psychiatric illness.

X-linked Adrenoleukodystrophy. X-ALD is an X-linked recessive disorder occurring in 1 in 20,000 births and is caused by mutations to *ABCD1,* the gene for a peroxisomal membrane protein that β-oxidizes very long-chain fatty acids (VLCFAs). This leads to the accumulation of saturated VLCFAs in brain white matter and adrenocortical cells predominantly, which impairs membrane stability. It predominantly affects males, although some female carriers can be affected. In adults, it presents in either a predominantly cerebral form (5 percent of cases), marked by an inflammatory demyelinating process, or an adrenomyeloneuropathic (AMN) form (45 percent of cases), in which neuronal dysfunction is predominantly a distal axonopathy affecting the dorsal columns and corticospinal tract. The majority of other presentations are of the childhood cerebral form, which is rapidly progressive over 2 to 3 years. Both adult forms also present with adrenocortical insufficiency, often indistinguishable from primary Addison disease. The adult cerebral form shows a predilection for neuropsychiatric presentations, although the neurobiology of this is unclear. Demyelinative changes are most prominent in parietal and occipital cortex as well as the thalamus, callosum, and brainstem (Fig. 2.14–5), although more anterior cortical regions may be involved. MRI frequently shows symmetrical white matter hyperintensity that begins in the splenium of the corpus callosum and spreads into neighboring posterior regions on T2-weighted imaging. At presentation, the majority of adult-onset patients present with psychiatric disturbance, most commonly behavioral changes. Mania and affective psychosis appear to be the most common neuropsychiatric presentations, more so than schizophreniform illnesses, although the latter do occur. When white matter changes occur in a more anterior distribution, patients may show a frontotemporal dementia-like presentation. Psychiatric presentations may precede motor or cognitive changes by some years; and as in other organic psychiatric illnesses, patients often show treatment-resistance. In the AMN form, the most common form of ALD amongst adults, was thought to affect only the peripheral nervous system; however, subtle cerebral manifestations of the disorder are often present, and the rate of depressive illness appears to be elevated at least twofold. Some patients may present with mood changes subsequent to adrenal insufficiency, which reverse with appropriate corticosteroid replacement therapy. AMN patients may also show subtle neuropsychological changes that affect speed of processing and executive functioning. There is no primary treatment for X-ALD, although bone marrow transplantation has provided some stabilization in younger patients with early disease, and the oral administration of 4:1 glyceryl trioleate and glyceryl trierucate ("Lorenzo's Oil") normalizes plasma VLCFA levels but does not improve neurologic function in already symptomatic or adult patients.

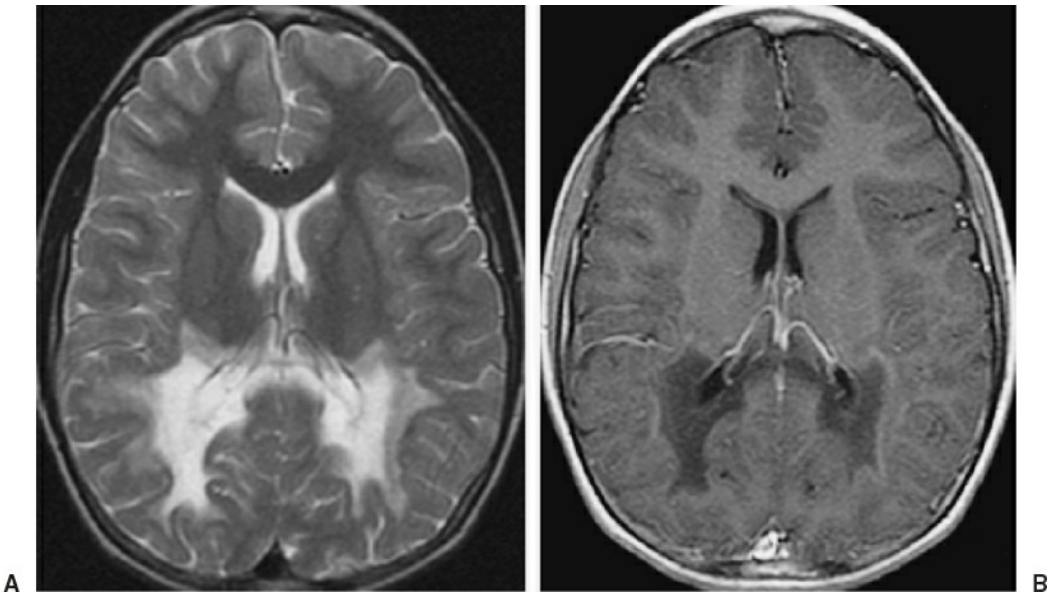

FIGURE 2.14–5. T2-weighted (**A**) and T1-weighted (**B**) axial magnetic resonance images (MRI) from an adolescent with adrenoleukodystrophy who presented with confusion and psychosis. Symmetrical and confluent hyperintensity is seen in the posterior white matter of both hemispheres, a typical MRI finding. (From Hesselink JR. Differential diagnostic approach to MR imaging of white matter diseases. *Top Magn Reson Imaging*. 2006;17[4]:243, Figure 19, with permission.)

Other Enzyme Deficiency Disorders

Acute Intermittent Porphyria. Acute intermittent porphyria (AIP) is one of the porphyria disorders group, where defects in heme biosynthesis result in excessive secretion of urinary porphyrins and their precursors; a number of heme synthesis pathway intermediates are neurotoxic. Most of the porphyrias are inherited in a dominant fashion. The incompletely penetrant, autosomal-dominant AIP results from defects in the enzyme porphobilinogen (PBG) deaminase, which speeds the conversion of PBG to hydroxymethylbilane. Deficient activity of this enzyme results in accumulation of porphyrin precursors PBG and amino-levulinic acid (ALA), which are thought to be neurotoxic. This enzymatic deficit becomes apparent in situations that require an increase in heme synthesis, including fasting, menstruation, intercurrent medical illness, and drugs that induce the cytochrome P450 system such as alcohol, estrogens, barbiturates, and sulfonamides, and presents most commonly in women of child-rearing age. Coproporphyria, a dominantly inherited deficiency of coproporphyrinogen oxidase, can present identically to AIP in addition to cutaneous findings (bullous lesions, skin fragility, and scarring).

The periodic "madness" of King George III has in recent decades been considered to be secondary to AIP, in addition to being implicated as illnesses affecting the work of van Gogh and Soren Kierkergaard. AIP has been shown to be significantly overrepresented in a sample of 4,000 psychiatric inpatients (1 in 500 compared to a community rate of 1 in

100,000). The "classical triad" consists of abdominal pain, psychiatric disturbance, and peripheral neuropathies (mostly motor, and often mimicking Guillain–Barré syndrome) during episodes, although psychiatric symptoms alone may be the single presenting feature. Of clinically symptomatic cases, psychiatric disturbance occurs in up to half of all cases, half of which are psychotic episodes, although depression, anxiety, and delirium may also be the main presenting symptoms. The intermittent "attacks" of neuropsychiatric disturbance may result in a misdiagnosis of schizophrenia; in addition, the high likelihood of abdominal pain during attacks can result in a diagnosis of a somatoform disorder. More chronic porphyrias have been associated with elevated rates of anxiety and depression. How PBG and ALA accumulation causes neuropsychiatric disturbance is unclear. Explanatory hypotheses have included oxidative stress, vascular change, and demyelination, although it may be that ALA's structural similarity to γ-aminobutyric acid (GABA) results in impaired release of GABA from synapses of GABAergic inhibitory neurons, and reductions in heme-dependent enzymes causing increased serotonin turnover and reduced nitric oxide activity.

Diagnosis of AIP rests on the demonstration of elevated urinary ALA and PBG, red cell PBG deaminase activity; for coproporphyria, urinary coproporphyrin levels should be assessed. Gross elevations of urinary ALA and PBG will often turn urine amber or purple in direct sunlight; it is imperative to collect specimens carefully (such as a 24-hour collection of urine, in light-protected containers and correctly preserved). Management involves correct identification and avoidance of precipitants if possible. During an attack, treatment includes the reversal of contributing illnesses, intravenous hydration, carbohydrate loading to inhibit heme synthesis, and hematin or heme arginate to provide negative feedback to the heme synthetic pathway. Psychopharmacological management of AIP involves judicious use of medication that will not worsen the biochemical deficit, which for psychosis includes chlorpromazine (Thorazine) and droperidol (Inapsine), fluoxetine (Prozac) for depression, lithium for mania, and lorazepam (Ativan), triazolam (Halcion), and temazepam (Restoril) for anxiolysis and sedation.

Phenylketonuria. Phenylketonuria (PKU), an autosomal recessive disorder with an incidence of 1 in 10,000 to 20,000 is caused by mutations in both alleles of the chromosome 12 gene for phenylalanine hydroxylase (*PAH*) that converts the amino acid phenylalanine to tyrosine. Mutations in both copies of the gene for *PAH* result in inactive or deficient enzyme levels, and accumulated phenylalanine (PhE) is converted to phenylketones, which are detectable in the urine. Resultant low levels of tyrosine, the precursor for the monoamines dopamine and serotonin, causes monoaminergic depletion in the CNS and severely disrupts normal

neurodevelopment. In addition, reduced large neutral amino acid (LNAA) availability causes reduced protein and cholesterol synthesis, impaired synaptogenesis and myelination, and altered glutamatergic neurotransmission. Hypomyelination occurs in untreated and early-treated patients; increased T2 signal on MRI scanning is seen in most patients in frontal and parieto-occipital zones with MRS demonstrating increased cerebral PhE that tightly correlates with serum Phe levels. Functional MRI demonstrates altered connectivity within and between prefrontal cortical regions; DTI has also demonstrated impairments in white matter integrity, and these appear to underpin the executive deficits seen in PKU patients.

PKU sufferers require an individually tailored diet of foods low in phenylalanine and supplemented with tyrosine. Routine screening of newborns and the universal initiation of dietary restriction within the first 3 months of life have essentially eliminated the severe, irreversible cognitive deficits and behavioral disturbance associated with untreated PKU. A reduced-phenylalanine diet maintained for the first decade of life is associated with essentially normal cognitive outcomes, although some controversy exists as to whether relaxing dietary restrictions in the preadolescent phase ("early treated" patients) is associated with greater psychiatric disturbance than in PKU sufferers in whom dietary control is not relaxed ("consistently treated" patients). Early treated patients are described as showing elevated rates of depression, anxiety disorders (particularly agoraphobia), attention-deficit disorder, and more nonspecific psychosocial adjustment issues in adolescence when compared to matched healthy individuals. Compared to individuals with chronic illnesses such as diabetes, they show elevated rates of anxiety disorders, suggesting that this is unlikely to be an effect of adjustment to chronic illness. Elevated rates of psychiatric and cognitive illness appear to directly negatively affect psychosocial outcomes, with a higher proportion of PKU patients remain unmarried, childless and/or living with their parents than their unaffected peers. Cognitive impairment also relates directly to the degree of past treatment, with untreated patients showing mental retardation and developmental delay; early-treated patients may show subtle reductions in IQ and executive and attentional impairment. Prefrontal dopaminergic neurons are particularly vulnerable to decreased tyrosine availability, and mild elevations in phenylalanine to tyrosine ratios (a marker of dopamine availability) in these patients in whom dietary restriction is not maintained are associated with executive impairment and may underpin the adolescent and early adulthood psychiatric disturbance seen in these patients. Acute tyrosine depletion in healthy adults is associated with depression, anxiety, and executive disturbance and may provide a model for psychiatric and cognitive disturbance as a result of a prefrontal hypomonoaminergic state as a result of elevated phenylalanine. In addition, impairment of oligodendrocyte function may impair myelination in PKU (Fig. 2.14–6).

Frontotemporal myelination, in addition to the dopaminergic innervation of the prefrontal cortex, peaks during adolescence, and the disruption of these processes may result in neuropsychiatric illness in this patient group. Animal models show that hypomyelination has been partially restored with dietary control, and limited human studies show a reversal of T2 signal abnormalities. However, perhaps less than 20 percent of adults maintain dietary control throughout adulthood, with the impact of diet on social interactions being the most commonly reported impediment for patients; discontinuation in adulthood is associated with poorer social, educational and health outcomes, but re-engagement of "lost" patients may allow for some reversal.

Maple Syrup Urine Disease. Maple syrup urine disease (MSUD) is an autosomal recessive disorder caused by a defect in the branched-chain α-ketoacid dehydrogenase enzyme complex, with resultant abnormalities in branched-chain amino acid (BCAA) catabolism. In adult-presenting disease, altered leucine metabolism results in impaired glial function, reduced synaptogenesis and dendritic arborization, impaired myelination and the induction of neuronal apoptosis. In addition, elevated leucine reduces transport across the blood–brain barrier of LNAAs, including most monoamine precursors, which results in disruptions to dopaminergic, noradrenergic, histaminergic and serotonergic transmission. Altered BCAA metabolism also results in depletion of glutamate, GABA and aspartate, which alters excitatory/inhibitory transmission throughout the CNS. Although quite rare, incidence is significant (1 in 200) in Amish and Mennonite populations. Clinical manifestations include body fluid odor that resembles maple syrup and overwhelming illness in the first week of life, beginning with vomiting and lethargy, and progressing to seizures, coma, and death if untreated. In milder forms of the disease, the illness may manifest symptoms only during stress (such as infection or following surgery). MSUD is diagnosed by elevated plasma BCAAs, particularly leucine, and profound ketosis and acidosis. Long-term management is via restriction of dietary BCAAs, although small amounts are required for normal metabolic function. Like PKU, the advent of dietary restriction has modified the illness course significantly, and MSUD patients who survive into adulthood have demonstrated subtle cognitive deficits and an elevated rate of some neuropsychiatric disorders. Like PKU, adult neuropsychological impairment appears more related to age at institution of, and adherence to, treatment rather than persistent BCAA levels. Whereas in PKU the relative tyrosine deficiency results in altered myelination and monoaminergic transmission, in MSUD the BCAA-restricted diet results in chronic cerebral valine deficiency, which itself may impair neuronal and oligodendrocyte function. Despite strict metabolic control, many children suffer from attention-deficit disorders, whereas

adolescents and adults commonly present with depression and anxiety. These are described as responding to psychostimulants and antidepressants, respectively. Cognitive impairment is usually seen in executive and attentional domains.

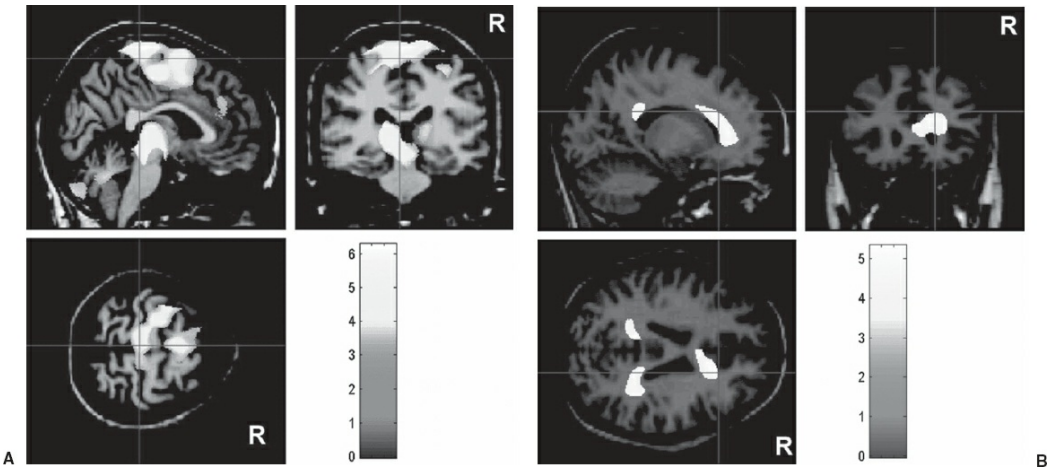

FIGURE 2.14–6. Results of voxel-based analysis of magnetic resonance image changes in phenylketonuria. **A:** Statistical parametric map of grey matter volume reduction of phenylketonuria (PKU) patients compared to controls, in motor and prefrontal cortex and thalamus. **B:** White matter volume reductions in the same comparison, in anterior and posterior forceps of the corpus callosum. R, right side. (From Pérez-Dueñas B, Pujol J, Soriano-Mas C, et al. Global and regional volume changes in the brains of patients with phenylketonuria. *Neurology*. 2006;66[7]:1074, Figure 1, with permission.)

Cerebrotendinous Xanthomatosis. Cerebrotendinous xanthomatosis (CTX) is an autosomal recessive disorder of cholestanol metabolism caused by mutations in the sterol-27 hydroxylase gene (*CYP27A1*) on chromosome 2. CTX was first described in 1937 and is also known by the eponymous name of van Bogaert disease. Since then over 300 patients with CTX have been described and about 50 mutations have been identified in the *CYP27A1* gene, just under half the mutations being missense mutations. Deficiency of the mitochondrial enzyme sterol 27-hydroxylase leads to reduced hepatic production of bile acid and reduced chenodeoxycholic acid (CDCA) production. The absence of CDCA-driven negative feedback on bile acid synthesis results in the accumulation of cholestanol precursors, increased cholestanol, and 7-hydroxycholesterol production. Accumulation of cholestanol and cholesterol in the eyes, tendons (often the Achilles, but also elbow/hand extensor tendons, patellar and neck tendons), and brain—particularly the cerebellar hemispheres (Fig. 2.14–7) leads to the classic clinical triad of juvenile cataracts, adult tendinous xanthomas, especially of the Achilles tendon, and progressive neurological impairment. The neurological manifestations are due to deposition of large granulomatous lipid deposits in basal/cerebellar nuclei and cerebral white matter or to demyelination and include peripheral neuropathy, mental retardation, seizures, cerebellar ataxia, pyramidal signs, and dementia. Psychiatric

symptoms usually accompany the dementing phase of the illness and include depression, psychosis, and personality or behavioral change. Two case series of CTX patients with psychiatric disturbance have been reported. In one, 3 of 35 cases suffered a neuroleptic-responsive psychotic disorder. A further series found psychiatric disturbance, predominantly agitation and psychosis, in seven of ten CTX patients. Depression has also been reported. A further series of 13 patients and review of all published CTX cases presenting with psychiatric syndromes suggests that younger patients present with behavioral and personality disturbance associated with learning difficulty, whereas older patients present with frontal–subcortical behavioral and cognitive syndromes associated with mood or psychotic disorders. Arteriosclerosis and osteoporotic fractures are often seen later in adult life with disease progression.

A 47-year-old woman with a 2-year history of treatment-resistant depression, unusual behavior, and memory problems was referred for neuropsychiatric assessment. On admission she fluctuated between periods of withdrawal and noncommunicativeness and periods of joviality associated with child-like comments and fatuous affect. She was unable to provide a detailed history and responded in a stereotyped way with single sentences: "I am confused," "This is what I am like at the moment." Cognitive assessment was not possible. The past history included cataracts and a hip fracture. Physical examination revealed enlarged, thickened Achilles tendons bilaterally (Fig. 2.14–7). Neurological examination revealed primitive reflexes and clonus in the right ankle, together with a right plantar response. A diagnosis of cerebrotendinous xanthomatosis was made on the basis of the clinical picture and elevated cholestanol levels. Treatment with chenodeoxycholic acid was instituted.

The diagnosis of CTX in adults should be suspected when the clinical picture includes cataracts, xanthomas, and progressive neurological or cognitive impairment. In the presence of the typical clinical picture the identification of elevated plasma cholestanol levels with low or normal cholesterol levels plus reduced urinary excretion of bile alcohols is usually sufficient to confirm the diagnosis. MRI using FLAIR sequences can identify bilateral hyperintensities of the cerebellar dentate nuclei that mirror the known neuropathological site of disease involvement (Fig. 2.14–7), usually associated with folia atrophy, and lesions in the brain stem and periventricular regions. Global grey and white matter volume is frequently reduced. Treatment consists of lifelong replacement of CDCA (750 mg per day, often combined with a statin) and is aimed at limiting the long-term damage caused by cholestanol and cholesterol deposition, which improves neurological and neuroradiological markers of disease, particularly when initiated at an early stage. Treatment with CDCA alone has also been shown to result in an improvement in psychiatric symptoms and cognition in CTX patients.

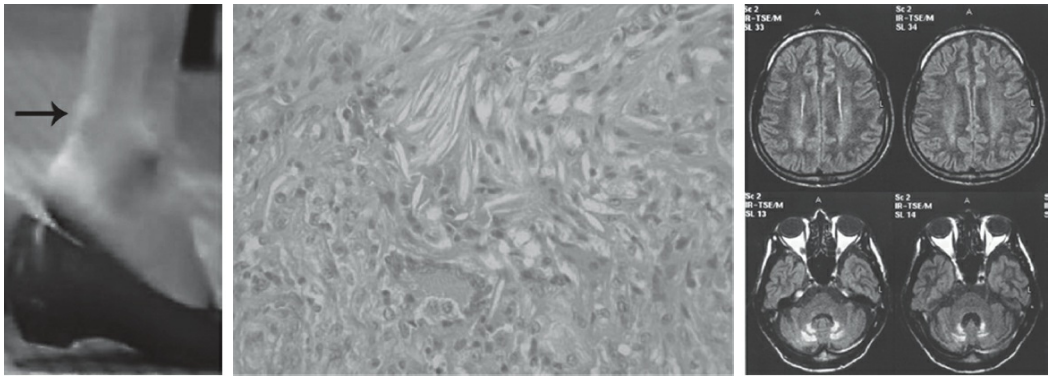

FIGURE 2.14–7. Cerebrotendinous xanthomatosis. **Left:** Image of left ankle of woman described in text demonstrating Achilles tendinous xanthoma. **Middle:** Hematoxylin and eosin stain of xanthoma of 42-year-old woman with early onset dementia and behavioral changes, showing cholesterol crystals. (From Wang Z, Yuan Y, Zhang W, Zhang, Y, Feng, L. Cerebrotendinous xanthomatosis with a compound heterozygote mutation and severe polyneuropathy. *Neuropathology.* 2007;27[1]:62, with permission.) **Right:** Transverse fluid-attenuated inversion recovery magnetic resonance imaging on two adult cerebrotendinous xanthomatosis patients showing mild hyperintensity of subcortical white matter and significant hyperintensity of the dentate nucleus of the cerebellum. (From De Stefano N, Dotti MT, Mortilla M, Federico A. Magnetic resonance imaging and spectroscopic changes in brains of patients with cerebrotendinous xanthomatosis. *Brain.* 2001;124:121, with permission.)

Tetrahydrobiopterin Deficiency. Tetrahydrobiopterin deficiency (THD) is an extremely rare recessive disease (1/500,000 to 1,000,000 newborns) caused by the mutation of several genes, in particular GCH1, PCBD1, PTS, and QDRP, which encode enzymes involved in the production and recycling of tetrahydrobiopterin. Mutations in MTHFR and DHFR genes can also cause a mild deficiency. Impaired enzymatic function leads to altered levels of phenylalanine and central monoamines. PAH, tyrosine hydroxylase, and tryptophan hydroxylase all require tetrahydrobiopterin (BH4) as a cofactor. Inherited defects that reduce the concentration of BH4, therefore, lead to PKU and to reduced CNS levels of dopamine and serotonin, as tyrosine hydroxylase and tryptophan hydroxylase are enzymes required for the synthesis of these neurotransmitters.

Classical symptoms are progressive and include intellectual disability (IQ between 60 and 90 in most patients), movement disorder, dysphagia, seizures, developmental delay, and behavioral disturbance in younger patients. BH4 deficit has been associated with depression, and reduced BH4 levels have been demonstrated in both schizophrenia and schizoaffective disorder, and supplementation with the BH4 analogue sapropterin is being investigated as a potential therapeutic agent in psychotic disorders.

Hyperprolinemia. Hyperprolinemia (HP) is a recessive disease caused by mutation of either of the two genes *ALDH4A1* and *PRODH*. These genes coding for pyrroline-5-carboxylate dehydrogenase and proline oxidase,

respectively, which are implicated in the degradation of proline. HP, which has no definitive treatment, is divided into two subtypes based on the severity of biochemical deficit; HP2 (with proline blood levels elevated ten to fifteen times) is considered more severe with patients presenting in childhood with seizures and mild intellectual disability. In HP1 (proline levels elevated three to ten times) patients have few symptoms or may be asymptomatic. There has long been a speculative association between HP1 and psychosis; the 22q11 deletion, known for its strong association with schizophrenia, results in a hemizygous deletion of the PRODH gene and up to 50 percent of patients with the 22q11 deletion have elevated plasma proline levels. In a recent study, schizophrenia patients demonstrated a mild to moderate increase in serum levels of proline when compared to sex and gender matched controls.

Propionic Acidemia. Propionic acidemia (PA), also called ketotic glycinemia and propionic aciduria, is a recessive disorder, affecting 1:100,000 births, caused by the deficiency of activity of the enzyme propionyl CoA carboxylase, a mitochondrial biotin-dependent enzyme which is essential for the catabolism of the amino acids threonine, methionine, isoleucine, valine, as well as cholesterol and some fatty acids. This results in a buildup of organic acids in blood and tissues, in particular propionic acid. Patients with early-onset PA present with acute metabolic decompensation in childhood. Late-onset PA, still largely a childhood disorder, is more heterogeneous and may fluctuate (particularly in times of increased metabolic demand), with symptoms including vomiting and feeding difficulties with impairment in psychomotor development and movement disorder such as choreoathetosis and dystonia. Acute acidosis with encephalopathy and hyperammonemia may occur. It may rarely present in the third or fourth decade of life. Diagnosis requires urine organic acid analysis and a plasma or blood spot acylcarnitine profile, and treatment is largely symptomatic. PA has been associated with adolescent or young-adult psychosis associated with metabolic decompensation; and with autism spectrum disorder (ASD) in preadolescent patients. Treatment with antipsychotics may require careful consideration as patients may have long QT syndrome (LQT).

α -Methylacyl-CoA Racemase Deficiency. α-Methylacyl-CoA racemase (AMACR) deficiency is a recently discovered recessive disorder caused by mutations in the *AMACR* gene, and has no specific treatment. The AMACR enzyme is found in mitochondria and peroxisomes. In peroxisomes, lack of AMACR leads to elevation of pristanic acid, a fatty acid, usually supplied from meat and dairy foods in the diet. The role of AMACR in mitochondria is felt to relate to the breakdown molecules derived from pristanic acid. Accumulation of pristanic acid influences functions of brain cells, leading

to neuronal and glial cell death, and leads to alterations of brain structure with encephalopathy, progressive demyelinating neuropathy and xanthomas with sterol storage in the CNS. Diagnosis is made by demonstrating elevated pristanic acid in plasma. The disorder usually presents in early adulthood with cognitive decline, seizures, and migraine; acute decompensations may present as stroke-like episodes. More chronic neurological signs, such as ataxia, spasticity, and sensorimotor neuropathy may also present, and it has been associated with a late-onset ataxia presenting in the fifth decade. MRI tends to demonstrate increased T2 signal in cerebello-thalamic pathway structures.

Disturbances in fatty acid and oxidative phosphorylation pathways have been described in brain regions of patients with schizophrenia, including the prefrontal cortex. The AMACR gene is also located on chromosome 5p13, a locus which has been shown to be a putative schizophrenia susceptibility locus in a large sample of patients from Puerto Rico with three missense variants segregating with schizophrenia.

Cerebral Creatine Deficiency Syndrome. The cerebral creatine deficiency syndromes (CCDS) encompass two autosomal-recessive creatine biosynthesis disorders: guanidinoacetate methyltransferase (GAMT) deficiency and L-arginine:glycine amidinotransferase (AGAT or GATM) deficiency, and the X-linked creatine transporter (SLC6A8) deficiency. Intellectual disability and seizures are common to all three CCDS. Creatine plays a key role in cellular energy availability through the regeneration of ATP, and is partly synthesized endogenously and supplemented by dietary intake. Creatine is synthesized from L-arginine, glycine, and methionine through a two-step process involving arginine:glycine amidinotransferase (AGAT) and GAMT. When synthesized, creatine is then transported to tissues with a high energy demand via the creatine transporter (SLC6A8). A diagnosis is suggested by a reduction or absence of the creatine peak visualized by proton MRS and increased urinary creatine excretion. Creatine monohydrate is the key treatment for GMAT and AGAT deficiency, but no treatment exists for SLC6A8 deficiency.

Almost all patients with GAMT deficiency have a behavioral disorder that may include autistic behavior and self-mutilation associated with pyramidal and extrapyramidal signs, with onset usually between the ages of 3 months and 3 years. Patients present mild intellectual disability and speech delay to severe intellectual disability, seizures, and behavioral disorder that may become more marked during the course of the disease. AGAT deficiency usually presents as mental retardation, autistic behavior and severe language disorder, without epilepsy. Autistic behavior is also a common feature in SLC6A8, although presentation tends to occur later in childhood or even in adulthood. SLC6A8 deficiency is the cause of approximately 2 percent of X-linked intellectual disability. Clinical

characteristics include psychomotor and expressive speech delay, autism, epilepsy, and muscle hypotonia and hypotrophy. Semantic pragmatic language disorders and oral dyspraxia have been described.

Glycine Encephalopathy. Glycine encephalopathy, also known as nonketotic hyperglycemia (NKH), is caused by mutation of the GLDC (80 percent of NHK) or AMT genes (10 to 15 percent); its prevalence is estimated between 1:55,000 and 1:63,000 newborns. Theses genes produce to proteins implicated in the glycine cleavage enzyme complex that is responsible for breaking down glycine into smaller pieces. A dysfunction in this complex leads to an excess of glycine reaching a level toxic to various tissues. As glycine, an amino acid, is also a neurotransmitter, the brain is a key site of this damage. The classical form of the glycine encephalopathy appears shortly after birth presenting with lethargy, feeding difficulties, hypotonia, abnormal movements, and breathing and is not infrequently fatal. Outside the neonatal period, individuals may present with profound intellectual disability and treatment resistant seizures. The chief clinical phenotype is intellectual deficiency and childhood ADHD. Due to the intrinsic role glycine plays in excitatory neurotransmission by modulating the activity of the NMDA subtype of glutamate receptors, it has been suggested that NKH theoretically could present with psychosis in older patients, although there are no published reports associating psychosis and NKH. Biological diagnosis is made by glycine measurement in the CNS and confirmed by mutation analysis of the GLDC and AMT genes.

Succinic Semialdehyde Dehydrogenase Deficiency. Succinic semialdehyde dehydrogenase (SSADH) deficiency is a rare recessive metabolic disorder with a nonspecific and variable phenotype characterized by reduced activity of the enzyme SSADH, involved in the degradation of GABA leading to an accumulation of γ-hydroxybutyric acid (GHB) and coded for by the ALDH5A1 gene. GABA is the major inhibitory neurotransmitter in the brain (although is excitatory in the embryo), and is derived from glutamate, the major excitatory neurotransmitter. Impairments to SSADH activity result in an alteration of the balance between excitatory and inhibitory neurotransmission. Diagnosis rests on demonstration of elevated urinary GABA and leukocyte SSADH assay. MRI commonly shows T2 hyperintensity in basal ganglia, dentate nucleus and white matter.

Patients with SSADH deficiency present intellectual disability with deficits in expressive language, hypotonia, choreoathetosis, nonprogressive ataxia, and hyporeflexia. Neuropsychiatric symptoms are prominent and include sleep disorders, inattention, hyperactivity, autism-like behavior and anxiety or obsessive–compulsive symptoms. Symptoms occur before the age 4 years in most cases, although diagnosis has been made as late as

the third decade. Developmental delay is present in all cases and 10 percent of patients will present movement disorders such as chorea, myoclonus, and dystonia. Delayed language is very common and frequently associated with mental retardation. A recent study on a sample of transcortical sensory aphasia (TSA) patients (n = 187) found two patients with SSADH deficiency, suggesting enrichment of this disorder in some clinical populations. In a review of 37 cases with of SSADH deficiency with psychiatric symptoms (70 percent of 53 patients), hyperactivity was present in 40 percent of patients, inattention in 50 percent, obsessive–compulsive disorder (OCD) in 35 percent, and aggressive behavior in 15 percent. Four patients in this sample also had hallucinations. The majority of neuropsychiatric presentations were seen in older children and adults, and these symptoms tended to respond to stimulants, serotonergic medication and antipsychotics. The enzyme deficiency itself has no specific treatment, although trials of vigabatrin, an irreversible inhibitor of GABA transaminase, has improved some patients but worsened others.

Serine Deficiency Syndrome. Serine deficiency syndromes (SDSs) are a group of three diseases implicating disturbance in the L-serine synthesis pathway. Three serine-deficiency syndromes have been described: 3-phosphoglycerate dehydrogenase (3-PGDH) deficiency, 3-phosphoserine phosphatase (3-PSP) deficiency, and phosphoserine aminotransferase (PSAT) deficiency. L-serine plays a key role in the development of the CNS. Its conversion to glycine provides a source of one-carbon units for methylation processes, nucleotide and phospholipid synthesis, and cell proliferation; it also facilitates the role of glycine and D-serine in excitatory neurotransmission through activation of NMDA receptors.

PSAT deficiency patients present with acquired microcephaly, psychomotor retardation, treatment-resistant seizures, and hypertonia. PSH deficiency patients have congenital microcephaly and intra-uterine growth retardation, feeding difficulties, and psychomotor retardation, and has been associated with Williams syndrome. The infantile form of 3-PGDH deficiency presents akin to PSH deficiency, whereas the juvenile form is more insidious with absence seizures, moderate developmental delay and abnormal behavior. This last clinical form may be seen by child psychiatrists in adolescence with developmental delay and aggressive behavior with no clinical neurological signs.

Diagnosis is made by demonstration of low plasma and CSF serine; treatment with L-serine (+/− glycine) improves seizures and behavioral disturbance but not developmental delay.

Cerebral Folate Deficiency. Cerebral folate deficiency (CFD) is defined as any neurological syndrome where CNS levels of 5-methyltetrahydrofolate in the presence of normal extra-CNS levels of

folate. The two main causes are autoantibodies to the folate receptor that transports folate (also known as vitamin B9) into the CSF, or a mutation in the gene encoding this receptor; it has also been described in Aicardi and Rett syndromes, and mitochondriopathies. Folate plays a key role in myelination, in the generation of biogenic amines and pterins, and is involved in the production and repair of DNA and gene expression. CFD is generally treated with folinic acid ameliorates symptoms if initiated early.

Cerebral folate transport deficiency (CFTD) is rare recessive disorder caused by a mutation in the *FOLR1* gene encoding for the folate receptor FR1 that allows folate to enter the cell, and is found on the basolateral endothelial surface of the choroid plexus. Impaired function of (or antibodies blocking) this receptor prevent the receptor-mediated endocytosis of folate and thus its entry into the CSF. Symptoms of CFTD do not begin until late infancy as until this period folate receptor FR2 is responsible for folate transport before birth and in early infancy. Affected children have classically onset of the disease around age 2 with psychomotor delay and regression, speech difficulties, seizures and ataxia; leukodystrophy is usually apparent on MRI scanning secondary to marked impairment of myelination.

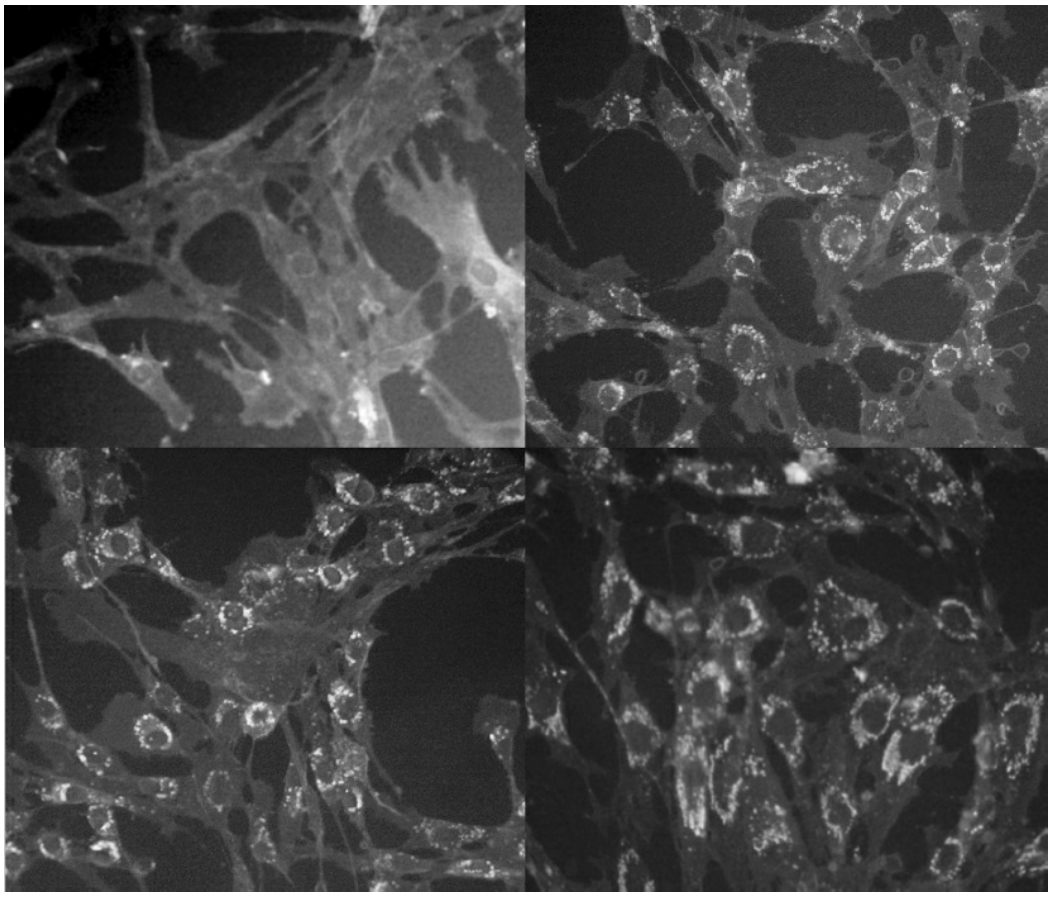

FIGURE 2.14–8. Filipin staining on fibroblasts for three patients with Niemann–Pick disease type C. **Top left:** Stain on healthy control reveals lack of filipin staining of accumulated cholesterol; on three other cases, each of whom presented with psychosis,

punctate perinuclear accumulation of cholesterol is identified.

Autoimmune cerebral folate deficiency (ACFD) is an acquired disorder, with a wider spectrum of age of onset of disease and has been reported in children, adolescents, and rarely adults, although the vast majority are infants who develop symptoms after 4 months of age with irritability, sleep disturbance and then dyskinesia, ataxia and seizures. A number of patients present with autistic behavioral features, and in adolescents it has been associated with catatonic schizophrenia. In these patients, FR1 and FR2 receptor encoding is normal, but autoantibodies develop to the receptors. It has been suggested that this may develop via induction by soluble folate-binding proteins in cow's milk in infants, or by sensitization by other similar epitopes in older patients. In one sample of ASD patients, FR1/2 autoantibodies were found in almost 75 percent of patients, and folinic acid treatment improved one-third of patients. Similarly, FR1/2 autoantibodies were detected in 80 percent of a small sample of treatment-resistant schizophrenia patients, a number of whom showed symptomatic improvement to treatment with folinic acid.

Other Neurometabolic Disorders

Niemann–Pick Disease Type C. NPC is an autosomal recessive neurovisceral disorder of lipid storage with a frequency of 1 in 100,000 live births, with 95 percent of sufferers having aberrations on the *NPC1* gene (18q11–12) and 5 percent on the *NPC2* gene (14q24.3). It is biochemically and phenotypically distinct from Niemann–Pick diseases types A and B, which result from a deficiency of lysosomal sphingomyelinase. *NPC1* and *NPC2* are involved in cyclical movement of sterols within cells and impairment results in late endosomal accumulation of cholesterol and gangliosides; excess GM2 and GM3 gangliosides cause ultrastructural change to neurons, and disrupted cholesterol metabolism results in altered τ- and β-amyloid processing (with Alzheimer-like neurofibrillary tangles), and disrupted calcium and metal metabolism. Axonal structures are particularly vulnerable and are affected early with subsequent involvement of cerebellar Purkinje cells, basal ganglia, and thalamic neurons followed by hippocampal and neocortical regions later (Fig. 2.14–8). The most fulminant neuronal loss occurs in the cerebellum, causing early cerebellar signs. Diagnosis is usually confirmed by demonstrating a low esterification rate of exogenous cholesterol in fibroblasts or by testing for lysosomal accumulation of free cholesterol by filipin staining (Fig. 2.14–9); more recently, oxidized cholesterol levels in plasma and/or direct NPC1 sequencing has been utilized. NPC may present in infancy, adolescence, or adulthood with a clinically variable picture, although its core features include dementia, dysarthria, ataxia, vertical supranuclear ophthalmoplegia, and hepatosplenomegaly. Seizures, dysphagia, and

pyramidal signs may appear with disease progression. Psychosis occurs in 25 to 40 percent of adolescent and adult-onset cases and may precede motor and cognitive disturbance by many years. Some cases have been described with onset in middle age associated with cognitive impairment alone. MRI often shows diffuse cerebral and/or cerebellar atrophy, and atrophy of striatum, hippocampus and thalamus, and thinning of the callosum. Executive function is impaired early in adults, before memory impairment and speed of processing becomes impaired. The co-occurrence of vertical gaze palsy and psychosis should prompt the clinician to consider NPC, so that appropriate treatment can be instituted. Psychosis may require higher doses of antipsychotics and may require treatment with a combination of neuroleptics and mood stabilizers or ECT. The development of substrate reduction therapy using the imino sugar miglustat (Zavesca), which penetrates the blood–brain barrier and reduces ganglioside accumulation, has shown promise in reversing the motor and cognitive deficits in NPC, and may also reduce or prevent psychiatric disturbance in this disorder.

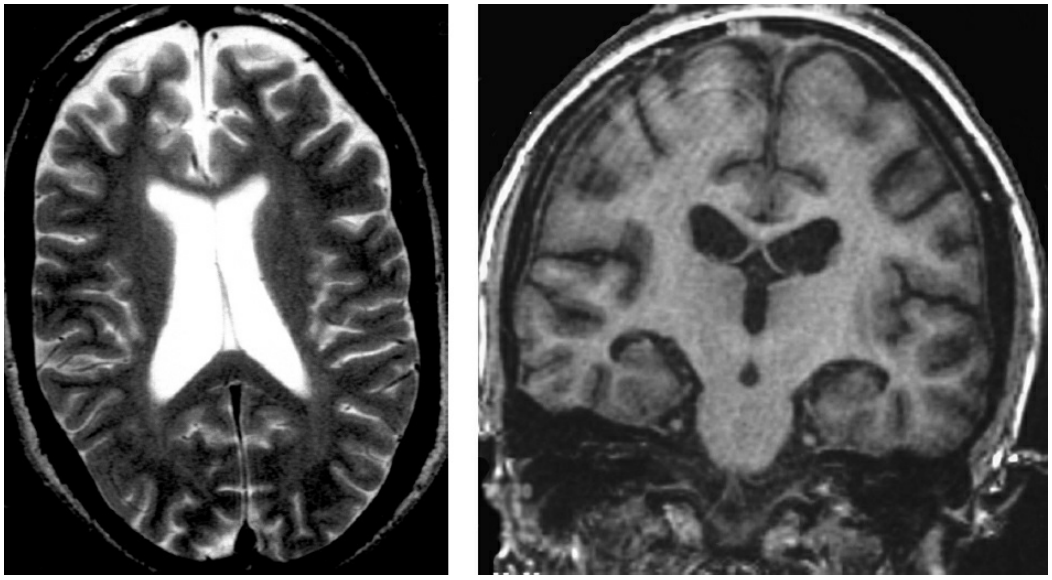

FIGURE 2.14–9. Magnetic resonance image scans on two individuals with Niemann–Pick disease type C who presented with major mental disorders. **Left:** T2-weighted axial image of a young adult male who presented with psychosis at age 16 prior to the onset of neurological disturbance at age 25, showing frontal atrophy, ventricular enlargement, and a cavum septum pellucidum. **Right:** A coronal T1-weighted image of a young male who presented with a rapidly cycling bipolar disorder in his early twenties. Note the disproportionate hippocampal atrophy in comparison to subtle global cerebral atrophy.

A 26-year-old man presented with movement disturbance and dysarthric speech following a 10-year history of a treatment-resistant psychotic disorder. He suffered persistent auditory hallucinations and referential and persecutory delusions, which only abated when treated with olanzapine (Zyprexa) 60 mg per day and valproic acid (Depakene) 2,000 g per day. On examination, he showed dysarthric speech, gait ataxia, and disturbed eye movements with jerky saccades and grossly impaired downgaze. He was cognitively rigid with significant

memory and executive impairment. Tests for TSD and other enzyme disorders were negative. Filipin staining of cultured fibroblasts showed an elevated number of cells, demonstrating perinuclear filipin staining of cholesterol (60 to 70 percent, normal less than 5 percent), and cholesterol esterification rate was mildly abnormal (2.9 pmol/h/mg; less than 2 abnormal, 2 to 3 equivocal, greater than 3 normal). Mutation analysis of the *NPC1* gene revealed a compound heterozygote status of G992R/R1186H, genetically confirming the diagnosis of Niemann-Pick's disease type C.

Pelizaeus–Merzbacher Disease. Pelizaeus–Merzbacher disease (PMD) is an X-linked recessive disorder due to abnormalities in the gene encoding for the proteolipid protein (PLP), the major structural protein of CNS myelin, resulting in patchy myelin loss as a result of oligodendrocyte apoptosis and/or axonal damage. PMD is also variable in its onset and clinical manifestations, with childhood-onset PMD resulting in mental retardation, nystagmus, and spastic paraparesis, and adult-onset cases most commonly associated with mild, adult-onset lower limb spastic paraparesis. Dementia and psychiatric dysfunction, including psychosis, are common in adult-onset cases, often associated with subtle or partial interruption to myelination (Fig. 2.14–10), although adult-onset PMD is rare. Psychosis, when reported, occurs in the fourth or fifth decade. PLP has been implicated in schizophrenia. Downregulation of the gene encoding PLP has been reported in schizophrenia and in animals treated with NMDA antagonists in experimental models of schizophrenia, implicating abnormal formation of myelin in psychosis. Treatment of PMD is symptomatic and supportive.

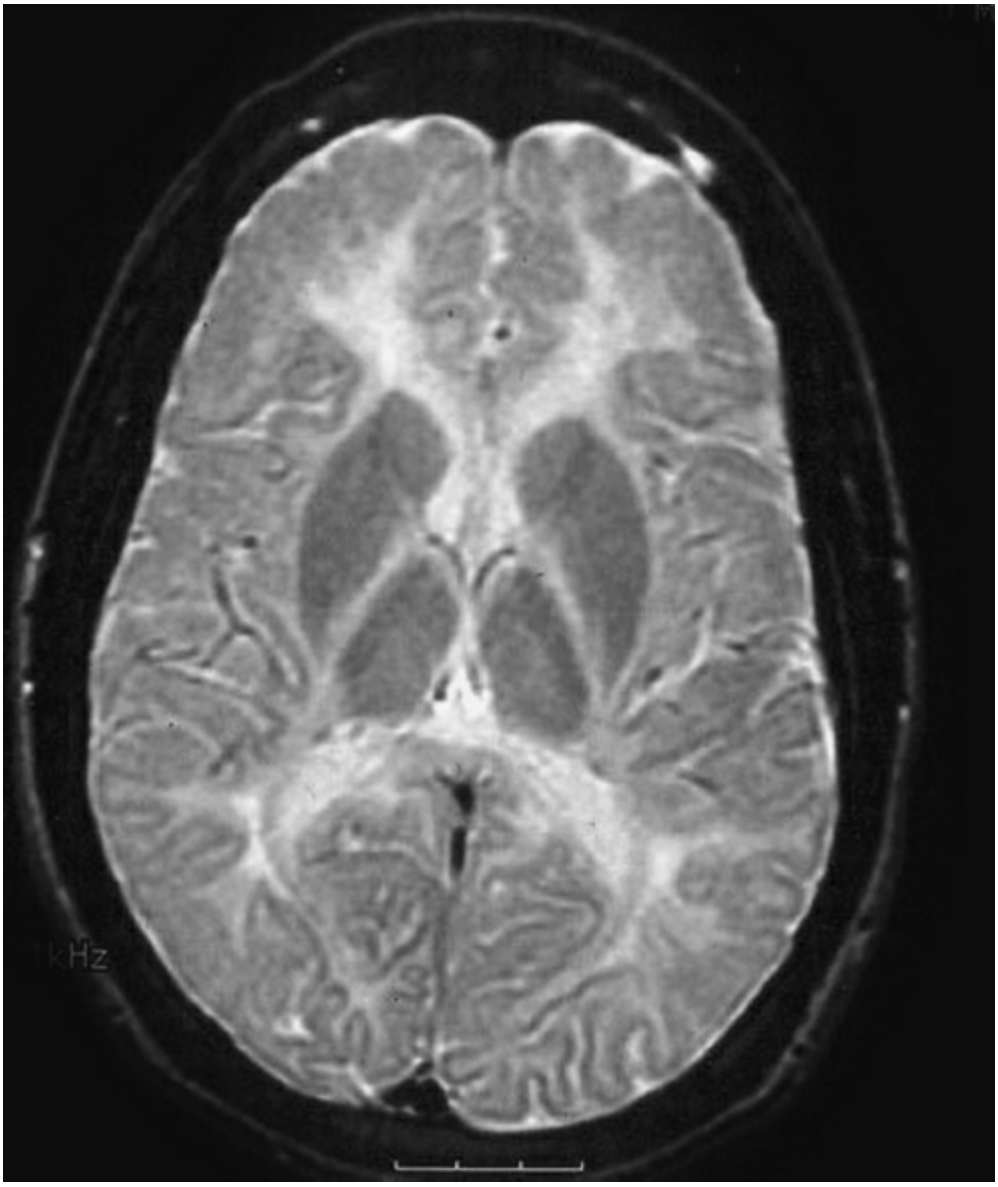

FIGURE 2.14–10. T2-weighted magnetic resonance image scan in 20-year-old individual with Perlizaeus–Merzbacher disease, demonstrating abnormally high signal in the internal capsule and posterior corpus callosum. The "mottled" frontal white matter represents "islands" of normally formed myelin. (From Koeppen AH, Robitaille Y. Pelizaeus-Merzbacher disease. *J Neuropathol Exp Neurol*. 2002;61[9]:747, with permission.)

A 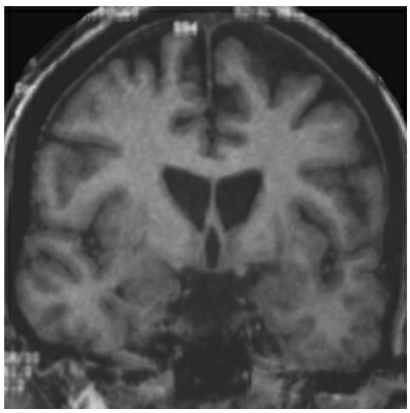 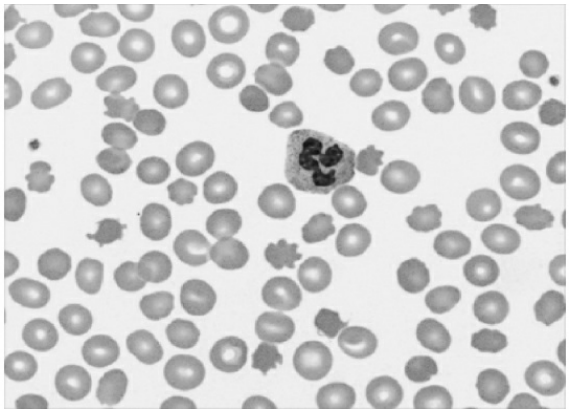 B

FIGURE 2.14–11. A 38-year-old woman with neuroacanthocytosis who first presented

with adolescent obsessive–compulsive symptoms prior to adult-onset seizures and chorea. **A:** T1-weighted coronal magnetic resonance image showing almost total loss of the caudate head. **B:** Blood smear showing characteristic red cell acanthocytes. (See color atlas.)

Chorea-Acanthocytosis. Chorea-acanthocytosis (ChAc) is a form of neuroacanthocytosis, a group of disorders that presents with neurological and psychiatric manifestations secondary to changes in the basal ganglia, and acanthocytes: spiculated red blood cells. ChAc is an autosomal recessive disorder associated with mutations or deletions in the *VPS13A* gene on chromosome 9q, which codes for the membrane protein chorein, a protein expressed in all tissues but particularly in brain, skeletal muscle, and erythroid cell precursors. The loss of function of chorein, which interacts with the membrane cytoskeletal proteins adducing and actin, causes destabilization of the membrane skeleton in red blood cells and neurons; in addition, adducing and actin are also crucially involved in synaptic function and plasticity, and synaptic vesicle trafficking. Cell loss and astrocytic gliosis then in the basal ganglia in ChAc patients, most particularly the caudate, but also the ventrolateral substantia nigra and globus pallidus (Fig. 2.14–10). The onset of neurological disturbance in ChAc is usually between the third and fifth decades, commonly with limb and orobuccal chorea that may be indistinguishable from Huntington chorea. It can be differentiated from the other main neuroacanthocytosis disorder, McLeod syndrome, by the absence of Kell antigens in red blood cells and the presence of cardiomyopathy in the latter. Patients with ChAc frequently present with mutilation of the tongue, lips, and cheeks, which is generally not a feature of Huntington chorea and can help to clinically distinguish the two disorders. Seizures, dystonia, and denervation atrophy occur in up to half of patients. Significant psychopathology is common in ChAc patients, occurring in up to two-thirds, and may precede the onset of frank neurological disturbance by up to a decade. In the original series of 19 individuals described by Hardie and colleagues (of a mixed ChAc-dominant group of neuroacanthocytosis patients), the most prominent psychiatric feature was behavioral and cognitive change (apathy, disinhibition, and poor judgment and planning) consistent with hypofrontality in more than half the patients, with OCD-like symptoms occurring in two patients. When psychiatric symptoms precede frank neurological disturbance, an overrepresentation of OCD-type disorders occurs, in addition to behavioral syndromes secondary to frontal lobe disturbance. These both appear to be the result of the predilection of neuropathology in ChAc for the head of the caudate nucleus, which is a key relay in a basal ganglia–thalamo–cortical loop known as the lateral orbitofrontal loop. This circuit integrates information from the anterior cingulate, orbitofrontal, and dorsolateral prefrontal cortex to determine behavioral and motor programs that occur to resolve conflict or facilitate decision making. Disruptions to this circuitry

lead to apathy, disinhibition, and poor judgment and planning. The motor compulsions seen in many ChAc sufferers may be secondary to behavioral dysregulation of motor acts through a loss of motor inhibition. Less commonly, patients with ChAc have been known to present with a schizophrenia-like psychosis. When compulsive disorders occur in ChAc, they have been known to respond to both SSRIs and TCAs. The disorder itself usually responds to antiseizure medication, and dopamine blockers/depleters for chorea, in addition to deep brain stimulation and botulinum toxin treatment.

A 38-year-old woman with a history of complex partial seizures with secondary generalization since the age of 21 presented with a history of movement disorder. She first presented at age 16 with contamination fears and compulsive picking at her skin, which responded to antidepressant treatment. At the age of 27, she developed involuntary movements of the limbs, head, and neck, which were severe at the time of assessment. MRI demonstrated gross bilateral caudate atrophy, particularly affecting the head of the caudate, and a blood film showed 5 percent acanthocytes (Fig. 2.14–11). A Western blot on peripheral blood demonstrated the absence of normal chorein, confirming the diagnosis of neuroacanthocytosis. Her cognitive function deteriorated, and she remained on antidepressant medication while haloperidol (Haldol) was added, which significantly improved her worsening chorea. Her cognitive and behavioral decline continued over subsequent years, and she died of an unrelated medical illness 2 years later.

Cysteine and Homocysteine Disorders

Homocystinuria. Disorders that lead to increased blood and urinary levels of homocysteine are grouped under the term homocystinurias. The classic form of homocystinuria is an autosomal recessive disorder caused by a defect in the chromosome 21 gene coding for cystathionine β-synthase, an enzyme that converts homocystine and serine into cystathionine. This leads to the accumulation of homocystine and methionine. Excessive homocystine disrupts the structure of fibirillin-1, an extracellular matrix protein, and leads to damage of collagen and elastic fibers; this leads to significant arterial and venous disease such as stroke, myocardial infarction, or venous thrombosis. In neurons it alters the release of monoamines, and may potentiate neurotoxicity through NMDA receptor agonism; it can also induce DNA strand breakage and cellular apoptosis. The primary clinical effects of the disorder reflect this pathology. The presenting features are usually with developmental delay or mental retardation associated with optic lens dislocation within the first 10 years. Optic lens dislocation is seen in almost 100 percent of patients older than 10, and other ophthalmological complications may ensue such as cataracts, glaucoma, and optic nerve atrophy. Mental retardation is seen in about half of the patients, and intellectual decline tends to be slowly progressive. Patients may appear Marfanoid with a long narrow head, arachnodactyly, kyphoscoliosis, and pectus excavatum. Patients will often have pale, pink

skin together with fine fragile, light-colored hair, and malar rash. Abnormalities in the clotting cascade and increased platelet adhesiveness lead to occlusion of and thromboembolism from arterial and venous vessels, with very high mortality. Other neurological complications include seizures and dystonias. Although early case reports identified patients with psychosis and homocystinuria, studies investigating large case series of patients have reported that personality disorder, behavioral disturbance (such as aggression), depression, and OCD are the most common psychiatric findings, and that aggression and behavioral disturbances were more common in patients with mental retardation. The diagnosis of classic homocystinuria is made on the basis of elevated plasma methionine, and elevated plasma and urine homocysteine. Brain imaging may show the effects of cerebrovascular accidents, atrophy, or venous occlusions. Treatment with methionine restriction and cysteine supplementation will generally prevent any long-term sequelae if the disorder is diagnosed at birth. Oral pyridoxine (vitamin B_6), which remethylates homocysteine to methionine, leads to a reduction in levels of methionine and homocystine in about half of patients, while the addition of folic acid and B_{12} may be of further benefit.

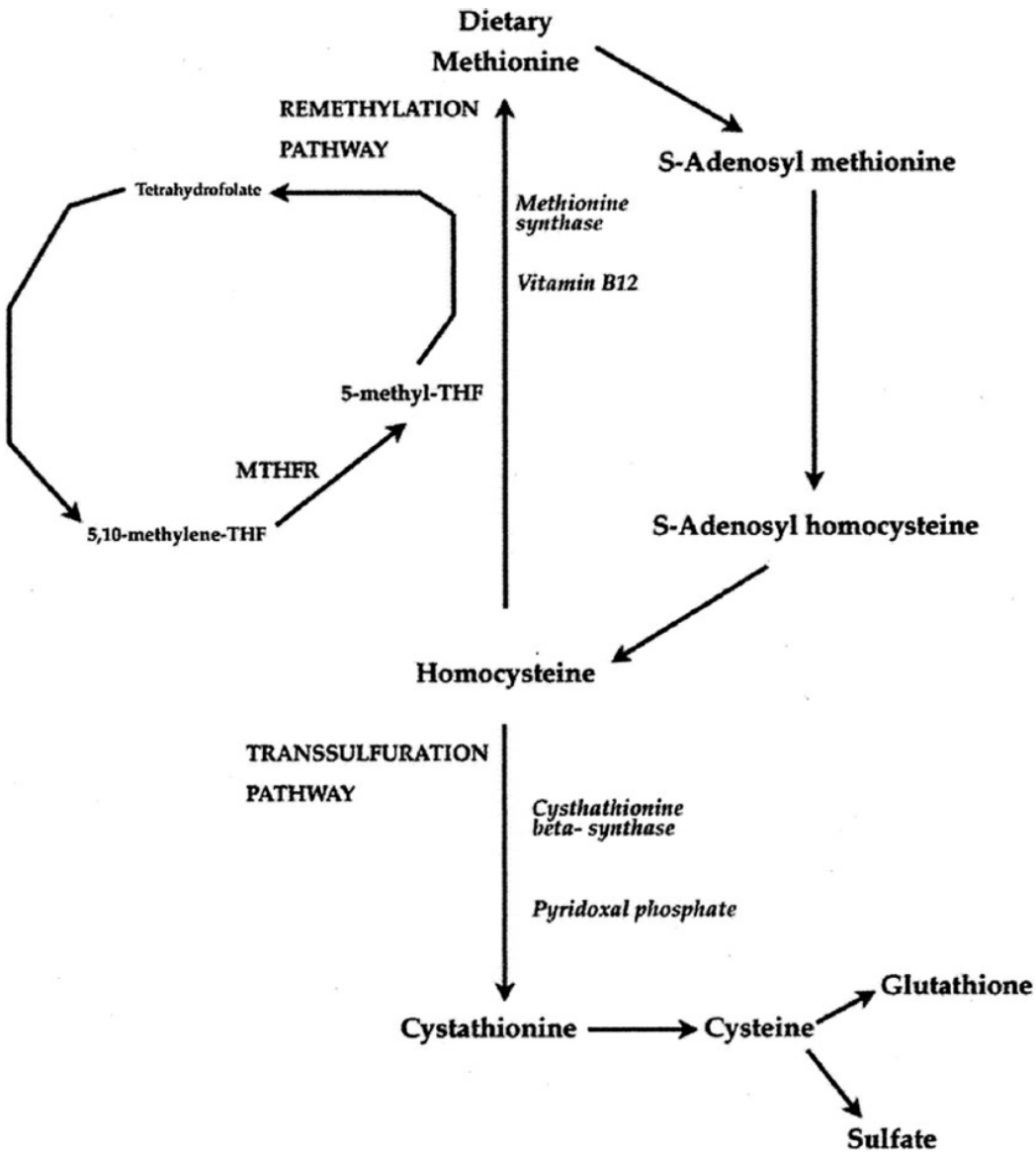

FIGURE 2.14–12. Metabolism of homocysteine in vivo. THF, tetrahydrofolate. (From Kelly PJ, Furie KL, Kistler JP, et al. Stroke in young patients with hyperhomocysteinemia due to cystathionine beta-synthase deficiency. *Neurology*. 2003;60[2]:275, with permission.)

The metabolic pathway that converts homocystine and serine to cystathionine is the most common pathway for the metabolism of homocystine (Fig. 2.14–12). Other autosomal recessive homocystinurias are caused by defects in an alternative remethylation pathway in which homocystine is catalyzed to methionine by either methionine synthetase or methylenetetrahydrofolate reductase (MTHFR). An inherited defect in methylcobalamin (methyl-B_{12}) that acts as a cofactor for methionine synthetase will also result in homocystinuria, together with methylmalonic aciduria. Deficiencies in these enzymes will lead to high levels of homocystine, but, in contrast to cystathionine synthase deficiencies, low levels of methionine. The clinical picture in these enzyme defects is one of mental retardation, seizures, and hypotonia, while megaloblastic anemia is

an additional feature of the disorder of methionine synthetase deficiency. The less severe form of MTHFR deficiency has been strongly associated with psychosis in adolescents and adults; this may be underpinned by alterations to dopamine synthesis and myelination. The C677T variant of the MTHFR gene, which causes a 35 percent reduction in enzyme function has been shown be an independent risk factor for the development of psychosis. The clinical course and MRI findings can resemble that seen in childhood leukodystrophies in these two disorders. Neuropsychiatric symptoms figure most prominently in the presenting symptoms of patients with methyl-B_{12} deficiency, with delays of up to 13 years between the onset of neuropsychiatric symptoms and the ultimate diagnosis.

Mitochondrial Disorders

Mitochondrial disorders are characteristically multisystem disorders that overlap clinically and should be considered as a differential diagnosis across a range of neurological and neuropsychiatric presentations. The genetic and phenotypic complexity of these disorders can be best understood in the context of a description of the mitochondrial genome, which is inherited maternally. Each human cell has up to several thousand copies of the mitochondrial genome, which is organized into a circular double-stranded structure. Human mitochondrial DNA was first sequenced in 1981 and consists of 37 genes, with no introns and a small noncoding portion. Thirteen proteins are translated from mitochondrial DNA and all are involved in the respiratory chain/oxidative phosphorylation system. The mitochondrial respiratory chain consists of five enzyme complexes made up of polypeptides encoded by nuclear and mitochondrial genes, except for complex II, which is entirely encoded in the cell nucleus. As a result, the mitochondrion is under the genetic control of both nuclear and mitochondrial DNA, and mitochondrial disorders can result from mitochondrial or nuclear DNA mutations. These enzyme complexes participate in a chain of metabolic processes that lead to ATP production, the overall process being referred to as oxidative phosphorylation. ATP is used in the vast majority of cellular metabolic processes as an energy source, and the respiratory chain responds to the energy needs of cells, which in some cases may be quite stable while in others (e.g., muscle), they may vary dramatically over time. Other functions of the mitochondria include cellular homeostasis, fatty acid oxidation, the urea cycle, intracellular signaling, apoptosis, and the metabolism of amino acids, lipids, cholesterol, steroids, and nucleotides.

The genetics of mitochondrial disorders are complex but the following principles can be generally applied:

1. Mitochondrial disorders may be sporadic, maternally inherited, or inherited in an autosomal pattern.

2. Due to the polyploid nature of the mitochondrial genome, the one cell may include normal and mutated mitochondrial DNA (heteroplasmy), and thus siblings may show a very broad range of clinical variability due to differences in the inheritance of such heteroplasmic mitochondria. Mitochondrial function will generally begin to fail once the proportion of mutated mitochondrial DNA has crossed a threshold of 70 to 80 percent.
3. Mitochondrial respiratory chain disorders will most affect tissues with high metabolic needs (e.g., muscle, central and peripheral nervous system, heart, endocrine, and eye).
4. The clinical expression of mitochondrial disorders may vary widely from individual to individual with the same mutation depending on the proportion of mitochondria affected in different tissues, the interaction of that individual with the environment, and the differential metabolic energy needs of different tissues within the one individual.

Diseases caused by defects of mitochondrial oxidative phosphorylation are the most common inborn errors of metabolism, accounting for 1 in 5,000 live births. For the reasons outlined above, there is no clear genotype–phenotype relationship for the mitochondrial genome disorders. Clinical features common to all mitochondrial disorders include dysfunction of endocrine (short stature, diabetes, thyroid, and adrenal disorders), neurological (deafness, myopathy, peripheral neuropathy, retinopathy, optic atrophy, ophthalmoplegia, seizures, ataxia, dementia), and cardiac (cardiomyopathy, cardiac block) systems. The patterns of clinical presentation vary significantly with regard to age of onset, the temporal order of symptoms and conditions, and the progress of the disorders. The combination of a maternal history, multisystem involvement, and a progressive course should arouse clinical suspicion of a mitochondrial disorder. In patients with atypical psychiatric presentations, physical signs such as muscle weakness, hearing loss, seizures, short stature, diabetes, Wolff–Parkinson–White syndrome, or migraines should alert clinicians to the possibility of a mitochondrial disorder.

Mitochondrial Encephalomyopathy, Lactic Acidosis, and Stroke-Like Episodes. Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS), the most common of the mitochondrial disorders, presents before early adulthood after a period of normal development. The majority (about 80 percent) of MELAS cases are due to an A to G substitution at nucleotide 3243 of tRNA leucine (A3243G $tRNA^{Leu(UUR)}$). A further 10 percent of mutations are in other regions of this same gene, while the remaining 10 percent of mutations occur in six other mitochondrial genes. The characteristic features of MELAS are stroke-like episodes whose lesions do not conform to vascular territories and may involve gray or white matter. They typically occur in tempero–parieto–occipital regions, basal ganglia, brainstem, and cerebellum and lead to

hemiparesis, hemianopia, and cortical blindness. Vomiting and migraine-like headaches are often associated clinical symptoms. Lactic acid levels are elevated and have been correlated with the level of neurological symptoms. The course of the disorder is highly variable, ranging from single stroke-like episodes through to a progressive course characterized by one or more of multiple strokes, deafness, diabetes, retinopathy, seizures, and cardiac abnormalities. Cases of schizophrenia or schizophrenia-like psychosis associated with MELAS have been reported. Such cases typically show an onset in the third decade, several years to a decade before the clinical diagnosis of MELAS is made. Similar cases of depression, OCD, bipolar disorder, and borderline personality disorder have been reported, but like the schizophrenia cases the course of the illness is not typical. Retrospective review of the patient's history will often reveal previously unappreciated features of MELAS such as short stature, diabetes, or unexplained somatic symptoms. The development of neurological signs or symptoms, cognitive decline, and evidence of strokes on imaging usually leads to the definitive diagnosis.

A 42-year-old man with deafness and diabetes was admitted with a confusional episode and found to have extensive posterior cortical and subcortical changes on MRI. The diagnosis of MELAS was confirmed by genetic testing. Five years later he was rereferred due to altered behavior, aggression, and having become very "fixed" in his ideas. He performed poorly on bedside executive function tests. Repeat MRI revealed extensive inferior frontal and parietotemporal atrophy (Fig. 2.14–13).

The diagnosis of MELAS is based on the clinical syndrome, elevated serum lactic acid levels, muscle biopsy showing ragged red fibers and molecular genetic testing. Ragged red fibers are muscle fibers, exhibiting mitochondrial proliferation in response to mitochondrial failure. MRI scanning typically shows cerebral stroke-like lesions. There is currently no available treatment for MELAS, although antioxidants, respiratory chain substrates, and cofactors have been studied in trials with varying results. Symptomatic treatments for MELAS may include cochlear ear implants for deafness, anticonvulsants, antipsychotics, and diabetic treatments. Antipsychotics with a higher risk of glucose intolerance (e.g., olanzapine) should be avoided or used with caution.

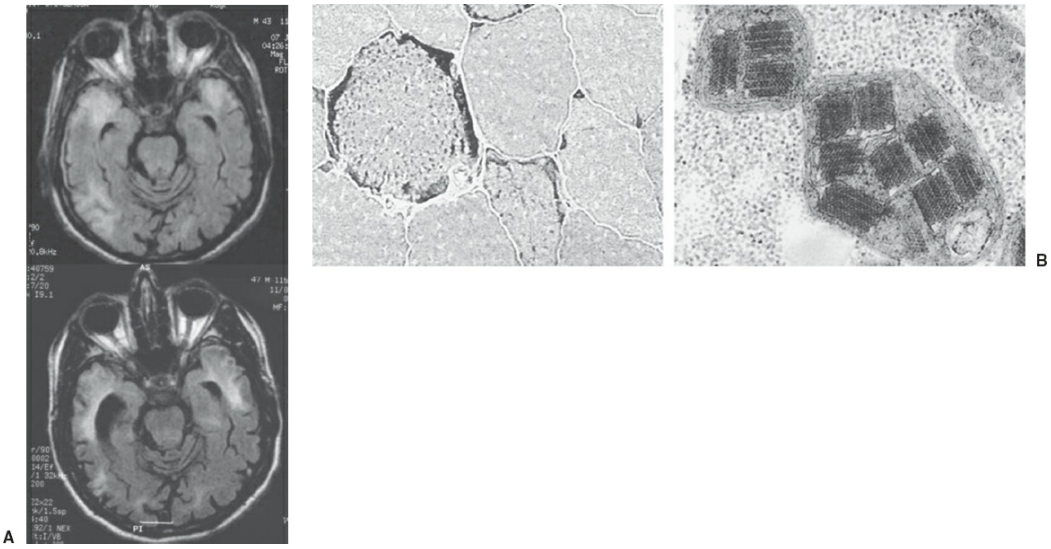

FIGURE 2.14–13. Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS). **Left:** Axial fluid-attenuated inversion recovery magnetic resonance image scans of a patient at age 43 and age 47, showing early gliosis of anterior temporal and parietal zones (**top**) with progression over 2 years to more advanced gliosis and cortical volume loss (**bottom**). **Right:** Skeletal muscle from 29-year-old man with early onset stroke. **A:** Gomori trichrome stain demonstrates the typical appearance of ragged red fibers, consisting of abnormal subsarcolemmal proliferation of mitochondria. **B:** Electron micrograph of skeletal muscle demonstrates "parking lot" inclusions, consisting of dystrophic mitochondria.

Myoclonic Epilepsy with Ragged Red Fibers. Myoclonic epilepsy with ragged red fibers (MERRF) typically begins in middle adulthood with photosensitive myoclonic seizures and is associated with limb-girdle weakness, dementia, cerebellar ataxia, cardiac abnormalities, and neuropathy. Muscle biopsy reveals ragged red fibers and the EEG is usually abnormal. In 90 percent of cases the causative mutation is an A8344G transition in the tRNA *lysine*. There is no specific treatment, but appropriate management of the epilepsy is critical to clinical outcome.

Other Mitochondrial Disorders. Kearns–Sayre syndrome (KSS) is a sporadic mitochondrial deletion disorder characterized by a triad of progressive external ophthalmoplegia, pigmentary retinopathy, plus one of the following: Heart block, cerebellar ataxia, or elevated CSF protein. Muscle biopsy reveals ragged red fibers, and serum lactate is often elevated. MRI will often show diffuse central white matter abnormalities and basal ganglia calcification. Maternally inherited Leigh syndrome (MILS) is an infantile encephalomyopathy characterized by seizures, heart block, dystonia, and optic atrophy. Leber hereditary optic neuropathy (LHON) is a maternally inherited disorder presenting as bilateral visual neuropathy in young adults. Although usually confined to the optic nerve, it has been described together with white matter pathology. "Overlap syndromes" have been described that include features typical of MELAS plus one of the other syndromes (e.g., LHON/MELAS, MELAS/MERRF, Leigh/MELAS).

Wolfram Disease Type 1 and Type 2. Wolfram disease type 1, an autosomal recessive disorder caused by a mutation on the short arm of chromosome 4 (4p16.1) in the *WFDS1* (wolframin) gene. The wolframin protein acts to regulate cellular calcium. Wolfram disease type 2 is caused by a mutation in the CISD2 gene (on 4q24) that codes for a mitochondrial membrane protein. It has been identified only in a small number of families to date. Wolfram disease is also known by the acronym DIDMOAD (*d*iabetes *i*nsipidus, *d*iabetes *m*ellitus, *o*ptic *a*trophy, *d*eafness) and often displays characteristic atrophy of the optic tract and loss of signal of the neurohypophysis on MRI (Fig. 2.14–14). Patients with CISD2 mutations do not develop diabetes insipidus but may exhibit gastrointestinal ulcers and bleeding problems. In the largest study of this disorder 68 patients with Wolfram disease were reviewed. Sixty percent (41 of 68) had had psychiatric symptoms and 25 percent (17 of 68) were classified as showing severe mental illness. Eleven (16 percent) patients had a history of psychotic symptoms. Heterozygous family members exhibited a high rate of psychiatric illness, approximately eight times greater than for noncarriers of the wolframin gene. The high rate of schizophrenia-like psychosis in this disorder is similar to that seen in some other adult-onset neurological disorders, such as velocardiofacial syndrome, MLD, and Niemann–Pick disease type C, each of which shows rates of psychosis in the 25 to 40 percent range.

FIGURE 2.14–14. Wolfram syndrome. Sagittal T1-weighted magnetic resonance image scan on a 32-year-old female with deafness and depression. Evident are frontal atrophy, thinning of optic chiasm and tracts, and atrophy of the brainstem and vermis and absence of physiological high signal of the neurohypophysis. (From Pakdemerli E, Karabulut N, Bir LS, Sermez Y. Cranial magnetic resonance imaging of Wolfram (DIDMOAD) syndrome.

Australas Radiol. 2005;49[2]:189, with permission.)

Disorders of Metal Metabolism

Wilson Disease. Wilson disease is an autosomal recessive disorder caused by mutations in a copper transporting ATPase encoded by the *ATP7B* gene on chromosome 13q14.3. About 1 percent of the population carries an *ATP7B* mutation, of which over 500 have been identified. The frequency of the disorder is estimated to be about 1 in 40,000. The gene defect leads to the accumulation of copper in the liver through impaired copper excretion and impaired binding of copper to ceruloplasmin. The subsequent catabolism of ceruloplasmin leads to low ceruloplasmin plasma levels and increased free copper. Free copper then accumulates in the brain and leads to the neurological and neuropsychiatric manifestations of this disorder. Copper deposition occurs in astrocytes but not neurons or the extracellular matrix, and is particularly evident in the basal ganglia.

The most common presentation of Wilson disease is with hepatic disease anywhere between the first and fourth decades. About 50 percent of patients are symptomatic by age 15. Up to two-thirds of patients will present with neurological disease in the second or third decade without clinical evidence of liver disease. The *dystonic* form is the commonest presenting neurological syndrome with dysarthria, dysphagia, drooling, and a rigid open mouth. In the *pseudosclerotic* form patients present with incoordination, clumsiness, unsteadiness of gait, or dysarthria. As the illness progresses patients may exhibit features of both forms with rigidity, tremor, choreic, athetoid, or dystonic movements. Such movements may be exacerbated by stress in the early stages of illness and be interpreted as functional. In particular, the tremor associated with Wilson disease, so-called wing beating, may be interpreted as a hysterical movement disorder. This tremor is characteristically absent at rest and develops after a short period of the arm extension. The arms beat in a wide violent arc, and the tremor may be altered by the position of the arms.

Wilson's original description in 1912 emphasized the importance of mental changes in the disease. Since then many authors have described the high prevalence of psychiatric symptoms with estimates ranging from 30 to 100 percent of symptomatic patients experiencing psychiatric symptoms at some point during their illnesses. Up to two-thirds of patients present with psychiatric symptoms and one-third will have received psychiatric treatment prior to the diagnosis being made. The most common reasons for psychiatric referral are behavioral and personality changes, with disinhibition, bizarre, or impulsive behavior being present in about a quarter of patients and depression in about one-fifth. Personality and behavioral changes, but not depression, correlate with the degree of neurological impairment. Although early reports suggested that psychotic presentation was a feature of Wilson disease, more recent investigations

have shown that psychotic symptoms are relatively rare. The importance of recognizing the early psychiatric presentations of Wilson disease lies in the proven benefits of early treatment intervention. Diagnostic delays may be exacerbated if extrapyramidal movements are attributed to psychotropic medications or interpreted as functional. Patients with neurological involvement show a subcortical pattern of cognitive impairment with frontal executive deficits that correlate with the extent of cerebral involvement.

The presence of Kayser–Fleischer (KF) rings (copper deposits in the outer rim of the cornea that are brown or gray-green) on slit-lamp examination is the single most important clinical diagnostic sign and is observed in about half of patients with hepatic presentation and almost all patients with a neurological or psychiatric presentation. Of lesser diagnostic value are low serum levels of ceruloplasmin, which may be normal in 10 percent of patients and high urinary copper excretion. The gold standard test for Wilson disease is liver biopsy with staining for copper. MRI shows reduced T1 signal and increased T2 signal in basal ganglia, thalamus, and brainstem (Fig. 2.14–15). The aims of treatment are to remove accumulated copper and to prevent reaccumulation through maintenance treatment. Treatment with d-penicillamine (Cuprimine) combined with pyridoxine has been the mainstay of initial treatment, but concerns that neurological symptoms are worsened by the treatment have led to calls that it be replaced with the less toxic copper chelators, trientine (Syprine) or ammonium tetrathiomolybdate. Zinc acetate therapy has been used as a preventative treatment in patients who are presymptomatic and for maintenance treatment. It acts by inducing an intestinal protein which binds copper and is not absorbed into blood. Neurological and psychiatric symptoms will improve with copper chelation over 1 to 2 years, with regression of MRI lesions on serial imaging. In patients with neurological forms of Wilson disease, the best outcomes are seen in patients with the pseudosclerotic form (82 percent symptom free), with lesser outcomes in patients with chorea (75 percent), parkinsonism (63 percent), and dystonia (53 percent).

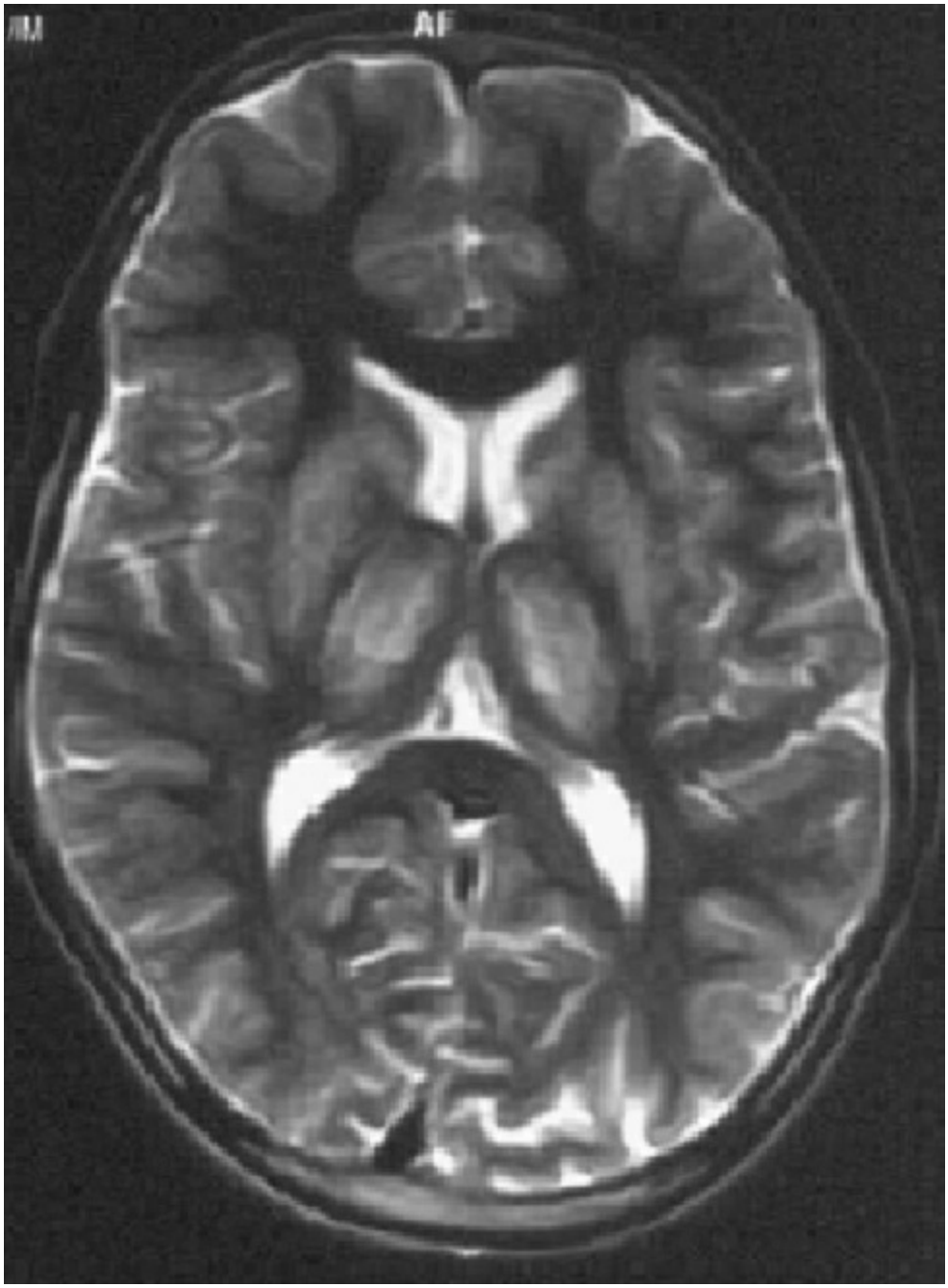

FIGURE 2.14–15. T2-weighted axial magnetic resonance image of an adolescent patient with Wilson disease, showing characteristic bilateral hyperintensity of the thalamus, caudate and putamen. (From Das M, Mirsa UK, Kalita J. A study of clinical, MRI and multimodality evoked potentials in neurologic Wilson disease. *Eur J Neurol*. 2007;14[5]:498, with permission.)

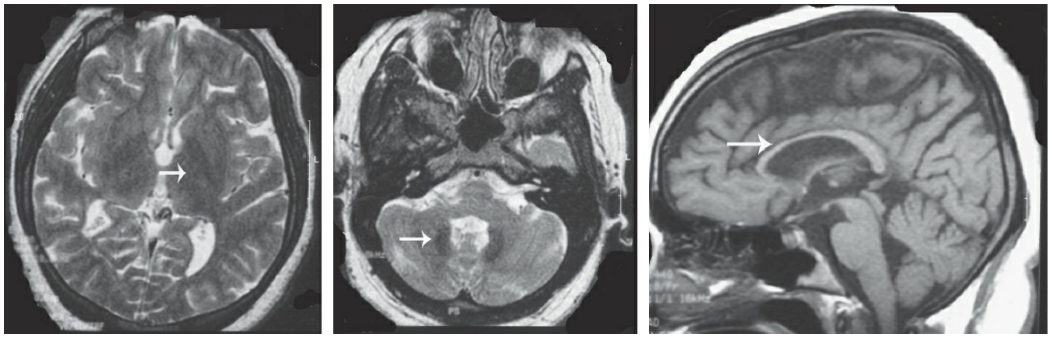

FIGURE 2.14–16. Scans of a 22-year-old woman patient with aceruloplasminemia and a schizophrenia-like psychosis. Reduced intensity of the basal ganglia (**left**) and dentate nucleus of the cerebellum (**middle**) on axial T2-weighted imaging representing iron deposition (*arrows*), and T1-weighted sagittal scan demonstrating anterior callosal thinning (**right***, arrow*).

NEURODEGENERATION WITH BRAIN IRON. The collective term neurodegeneration with brain iron (NBIA) refers to a number of genetic disorders associated with accumulation of iron in the brain. Aceruloplasminemia and neuroferritinopathy are caused by mutations in iron metabolism genes, while pantothenate kinase-associated neurodegeneration (PKAN) is caused by a mutation that disrupts several metabolic pathways, including iron metabolism.

Aceruloplasminemia. Aceruloplasminemia is an autosomal recessive disorder associated with reduced or absent levels of ceruloplasmin and tissue iron deposition. It can caused by about 40 mutations in the ceruloplasmin gene on chromosome 3q25 and occurs in 1 per 2 million births. There are two types of ceruloplasmin. Serum ceruloplasmin is predominantly synthesized in the liver but does not cross the blood–brain barrier while a second form (GPI-anchored form) is produced by astrocytes in the CNS. Ceruloplasmin promotes the loading of iron onto transferrin, allowing Fe^{3+} efflux out of cells and preventing oxidative damage caused by Fe^{2+}. GPI-anchored ceruloplasmin plays a key role in regulating iron levels in the CNS and in preventing free radical injury. The disease leads to the deposition of iron in the CNS, retina, pancreatic cells, liver, spleen, and ovaries. The major sites of CNS iron deposition in ceruloplasmin are similar to the sites of greatest iron concentration in healthy individuals: The basal ganglia, cerebellar dentate nuclei, red nucleus, thalamus, and hippocampus. Aceruloplasminemia presents with diabetes mellitus, retinal degeneration, and neurological symptoms. Neurological signs may be preceded for many years by diabetes mellitus and anemia due to inefficient iron delivery. Ataxia and extrapyramidal movements such as blepharospasm, dystonia, dyskinesia, grimacing, and parkinsonism usually develop in the fifth decade. A subcortical picture of cognitive decline then follows with personality change, amotivation, psychomotor slowing, and executive deficits. Psychosis has been reported, although psychiatric

symptoms have not been commonly described in aceruloplasminemia due to the rarity of the disorder. MRI findings typically show marked T2 hypointensity in the regions of maximal iron deposition, posterior white matter tract hyperintensity, and superficial cerebral and cerebellar cortical hypointensity.

A 21-year-old woman on treatment for aceruloplasminemia was referred for neuropsychiatric assessment. There was a family history of aceruloplasminemia and the diagnosis was made 12 months earlier after she was found to have abnormal liver function tests. She now presented with an 18-month history of schizophrenia-like psychosis and declining function in the absence of neurological signs. Neuropsychological testing showed significant dominant hemisphere deficits. MRI showed bilateral iron deposition in the cerebellar dentate nuclei and thalami, frontal atrophy, and periventricular white matter hyperintensities (Fig. 2.14–16).

The diagnosis of aceruloplasminemia can be made biochemically with findings of absent ceruloplasmin, low serum copper, normal serum total iron binding capacity, and moderately elevated ferritin. Treatment with the iron chelating agent desferrioxamine can decrease serum ferritin, reduce brain and liver iron stores, and can prevent the progression of neurological disease.

NEUROFERRITINOPATHY. Neuroferritinopathy is an extremely rare autosomal dominant disorder with fewer than 100 reported cases. It is caused by a mutation in the FTL gene on 19q13.33 with 7 mutations identified to date. The FTL gene codes for the ferritin light chain, a subunit of ferritin, which plays a role in the storage and release of iron from cells. In neuroferritinopathy ferritin is unable to store iron and excess iron is deposited in brain cells especially the globus pallidus, putamen, and dentate nuclei of the cerebellum. The clinical phenotype begins in the third or fourth decade with extrapyramidal symptoms and signs including chorea, ataxia, dystonia, tremor, dysphagia, and dysarthria. The presence of an orofacial action specific dystonia is characteristic of this disorder. A recent review of the neuropsychiatric features of neuroferritinopathy described 22 cases. This review identified that subcortical type cognitive changes began about 4 years after the onset of motor symptoms (mean onset age 47.2). Individual cases of OCD, psychosis, emotional lability, and aggression are noted.

The clinical diagnosis of neuroferritinopathy should be suspected in cases of adult onset movement disorder, MRI findings of excess iron deposition in basal ganglia or cystic degeneration on T2* MRI images. The presence of cystic degeneration may help distinguish neuroferritinopathy from aceruloplasminemia. Serum ferritin may be low. The genetic diagnosis is based on identification of a mutation in the FTL gene. There are no specific treatments available. Symptomatic treatment of the

movement disorder is indicated though conventional treatments (tetrabenazine, benzhexol, l-dopa, benzodiazepines) are variably successful. Botulinum toxin may be of benefit in dystonias.

Pantothenate Kinase-Associated Neurodegeneration. PKAN, an autosomal recessive disorder, is one of several disorders that had been previously described as Hallervorden–Spatz syndrome but are now collectively grouped under the term NBIA (neurodegeneration with brain iron accumulation). Common to these disorders is the accumulation and deposition of iron in the brain in association with clinical, radiological, and pathological evidence of neurodegeneration. The identification of mutations in the gene *PANK2* on chromosome 20p13 led to the descriptive term "pantothenate kinase-associated neurodegeneration" for what is recognized as the most prevalent form of NBIA. Pantothenate kinase 2 regulates the mitochondrial synthesis of coenzyme A (CoA), which is involved in energy and fatty acid metabolism. *PANK2* catalyses the phosphorylation of pantothenate (vitamin B_5) to phosphopantothenate, which condenses with cysteine in the next step of CoA biosynthesis. Mutations in *PANK2* lead to an accumulation of cysteine, which binds iron and leads to free radical production, which triggers cell membrane damage and death. The retina and basal ganglia appear particularly sensitive to these effects of PKAN mutations, and the primary clinical features are those of retinopathy and basal ganglia syndromes.

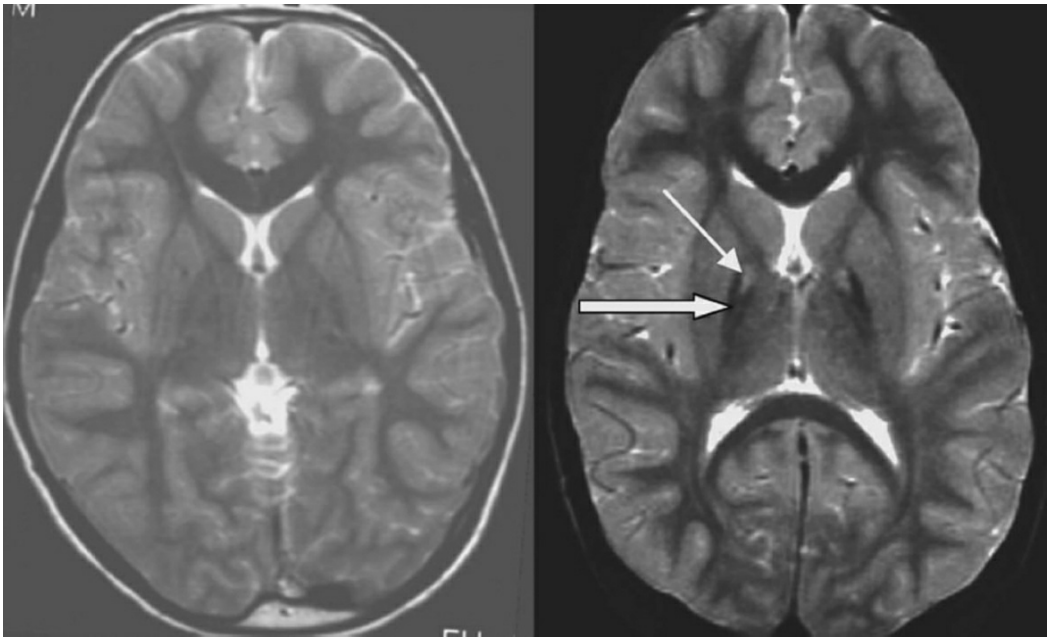

FIGURE 2.14–17. Pantothenate kinase-associated neurodegeneration. **Left:** T2-weighted axial magnetic resonance image of a normal control shows isodense globus pallidus. **Right:** *PANK2*-mutation-positive patient with neurodegeneration with brain iron accumulation (NBIA) shows hypointensity (*thick arrow*) with a central region of hyperintensity (*thin arrow*) in the medial globus pallidus, known as "the eye of the tiger" sign. (From Hayflick SJ. Unraveling the Hallervorden-Spatz syndrome: Pantothenate kinase–

associated neurodegeneration is the name. *Curr Opin Pediatr*. 2003;15[6]:572, with permission.)

The clinical phenotype of *PANK2* mutations can be divided into three types: the classic syndrome, the atypical syndrome, and HARP (hypobetalipoproteinemia, acanthocytosis, retinopathy, and pallidal degeneration). The *classic* syndrome first described by Hallervorden and Spatz consisted of the early childhood (3 to 4 years) onset of dystonia, dysarthria, rigidity, pyramidal signs, pigmentary retinopathy, and cognitive decline, and with a progressive fulminant course such that patients became nonambulatory by their mid- to late teens. All patients with this classic phenotype have been found to have a *PANK2* mutation. The *atypical clinical phenotype* (about 25 percent of PKAN) is characterized by presentation in the early to mid-teenage years with palilalia, tachylalia, dysarthria, and psychiatric symptoms, including depression, emotional lability, personality changes, and cognitive decline. Extrapyramidal rigidity, dystonia, and pyramidal spasticity develop subsequently and result in progressive loss of mobility over 15 to 40 years. About one-third of this *atypical clinical phenotype* exhibit *PANK2* mutations. Interestingly, the patients with *PANK2* mutations and late onset appear to be more likely to exhibit speech and psychiatric symptoms at onset. The third clinical phenotype, previously termed HARP, has now been identified as being caused by a *PANK2* mutation. Together with a strongly suggestive clinical picture, the finding of the characteristic MRI "eye of the tiger" sign (a low signal intensity region caused by iron deposition and a high signal area that corresponds to axonal spheroid formation, as seen in Fig. 2.14–17) is almost pathognomonic of a *PANK2* mutation and should lead to genetic testing. No treatment has been identified for PKAN and management remains symptomatic, with trials of iron chelation and antioxidants generally proving unsuccessful (Table 2.14–1).

Table 2.14–1.
Metabolic Disorders Presenting as Neuropsychiatric Syndromes, with Pattern of Inheritance, Gene Location, and Product, Prevalence, Onset, and Neuropsychiatric Features

Disorder	Inheritance	Location	Gene	Gene Product	Stor Subs
Metachromatic leukodystrophy	AR	22q13.31	ARSA	Arylsulfatase A	Cere
Fabry disease	X-linked recessive	Xq22	a-Gal	α-galactosidase A	glob
Tay–Sachs Disease	AR	15q23-q24	HEXA	β-hexosaminidase A	GM2

Neuronal ceroid lipofuscinosis (Kufs)	AD/AR	Not yet identified	CLN4	Unknown—more than one may be involved (CLN1&2 – lysosomal thioesterase & protease; CLN3&8 are membrane proteins of unknown function)	Unk
α-mannosidosis	AR	19q13.1	MANB	α-mannosidase	Oligo
X-linked adrenoleukodystrophy	X-linked recessive	Xq28	ABCD1	ALDP peroxisomal membrane protein	satu fa
Acute intemittent porphyria	AD	11q23.3	HMBS	porphobilinogen-deaminase	porp le
Phenylketonuria	AR	12q23.2	PAH	phenylalanine hydroxylase	phen
Maple syrup urine disease	AR	19q13.1–13.2 6p22-p21 1p31	BCKDHA BCKDHB DBT	branched chain ketoacid dehydrogenase	bran ac
NPC	AR	18q11—12 14q24.3	NPC1 NPC2	NPC1 NPC2	chol ga
Pelizaeus-Merzbacher Disease	X-linked recessive	Xq22	PLP1	proteolipid protein	n/a
Cerebrotendinous xanthomatosis	AR	2q33	CYP27A1	sterol 27 hydroxylase	chol ch
Cystinosis	AR	17p13	CTNS	cystonin	cysti
Homocystinuria	AR	21q22.3	CBS	cystathione synthase	hom m
Tetrahydrobiopterin deficiency	AR	Multiple	GCH1 PCBD1 PTS QDRP	Multiple	BH4
Hyperprolinemia	AR	1p36	ALDH4A1	Pyrroline-5-carboxylate dehydrogenase	Proli
Propionic academia	AR	13q32, 3q21–22	PCCA PCCB	Propionyl-CoA carboxylase	Prop
AMACR deficiency	AR	5p13	AMACR	α-methylacyl-CoA racemase	Prist
CCDS	AR & XL	19p13.3, 15q21.1, Xq28	GAMT AGAT SLC6A8	Guanidinoacetate N-methyl transferase, Glycine amidinotransferase, Sodium and chloride-dependent creatine transporter	Crea
NKH	AR	9p22, 3p21.31	GLDC AMT	Glycine dehydrogenase, aminomethyltrans-ferase	Glyc
SSADH deficiency	AR	6p22	ALDH5A1	Succinic semialdehyde	GAB

				dehydrogenase	
Serine deficiency	AR	1p12, 7p11.2, 9q21.2	PHGDH PSPH PSAT1	Phosphoglycerate dehydrogenase, phosphoserine phosphatase, phosphoserine aminotransferase	Serir
CFD	AR, autoimmune	11q13.3	FOLR1	FR1	Fola
MELAS	Maternal	mtDNA	MT-TL1	mitochondrial tRNA leucine 1	n/a
Wolfram disease	AR/Maternal	4p/mitochondrial	WFS1	wolframin	n/a
Wilson disease	AR	13q14.2-q21	ATP7B	copper transporting ATPase	copp
Aceruloplasminemia	AR	3q25	CP	ceruloplasmin	iron
Pantothenate kinase associated neurodegeneration	AR	20p13	PANK2	pantothenate kinase	cyste

Disorder	Onset	Neuropsychiatric Features
Metachromatic leukodystrophy	Childhood–adulthood	Psychosis in adolescent onset/chorea/seizures/dement
Fabry disease	Childhood–adolescence	Depression
Tay–Sachs Disease	Childhood–adulthood	Psychosis in adult onset, speech, gait disorder, tremor
Neuronal ceroid lipofuscinosis (Kufs)	Adult (Kufs)	Psychosis/depression/myoclonic epilepsy/dyskinesias
α-Mannosidosis	Childhood-adulthood	Psychosis in adult onset/ataxia/deafness/retinopathy
X-linked adrenoleukodystrophy	Childhood–adulthood	Affective psychosis in adult onset/demyelination
Acute intermittent porphyria	Adulthood	Psychosis/delirium/depression/peripheral neuropathy
Phenylketonuria	Childhood	Mental retardation
Maple syrup urine disease	Childhood–adulthood	Depression and anxiety in adults
Niemann–Pick disease Type C	Childhood–adulthood	Mental retardation in childhood onset/psychosis in adolescent onset/dementia in adult onset
Pelizaeus–Merzbacher Disease	Childhood–adulthood	Mental retardation in childhood onset/psychosis and dementia in adult onset
Cerebrotendinous xanthomatosis	Adulthood	Psychosis/agitation/depression
Cystinosis	Childhood	Cystinosis encephalopathy
Homocystinuria	Childhood	Mental retardation/seizures/dystonias/personality change/behavioral disturbance/depression/OCD
Tetrahydrobiopterin deficiency	Childhood–adulthood	Behavioral disturbance, mental retardation; adult moo psychotic disorders
Hyperprolinemia	Childhood–adulthood	Mental retardation, psychosis
Propionic academia	Childhood–	Encephalopathy, mental retardation, autism, psychosis

	adulthood	
AMACR deficiency	Early adulthood	Cognitive decline, psychosis
CCDS	Childhood	Autism, self-mutilation, mental retardation
NKH	Childhood	Mental retardation, attention deficit
SSADH deficiency	Childhood–adulthood	Mental retardation, autism, OCD, psychosis
Serine deficiency	Childhood	Aggression, mental retardation
CFD	Childhood–adulthood	Mental retardation, autism, psychosis
MELAS	Adolescence/early adulthood	Stroke-like lesions/psychosis/depression
Wolfram disease	Childhood–adolescence	Schizophrenia
Wilson disease	Adolescence/early adulthood	Extrapyramidal movements/psychosis/parkinsonism/dysphagia/dyr
Aceruloplasminemia	Adolescence/early adulthood	Ataxia, dystonia, dyskinesia, subcortical dementia
Pantothenate kinase associated neurodegeneration	Childhood–adolescence	Childhood dystonia, rigidity/adolescent palilalia, dysart depression, lability, personality change, cognitive de

AR, autosomal recessive; AD, autosomal dominant; tRNA, transfer ribonucleic acid; ATP, adenosine triphosphate; MELAS, Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes; OCD, obsessive–compulsive disorder.

Ion Channel Disorders

Disorders of ion channels, or channelopathies, have become an increasingly recognized group of disorders affecting cellular ion channels involving Na^+, Ca^{++}, K^+, and Cl^- in electrically excitable tissue, such as heart, muscle, and brain. A number of genetic ion channel diseases have now been well-described, such as the LQT, periodic paralysis, and a range of monogenic seizure syndromes. The cardinal feature of ion channel disease is the disturbance of rhythmic function, best illustrated by epilepsy, with an abnormally synchronous discharge causing a seizure; other rhythmic CNS disturbances such as ataxia, paralysis, and sensorineural deafness. Additionally, the possibility of developing an *acquired* abnormality of ion channel function has been recognized, particularly in autoimmune disorders.

Voltage-Gated Potassium Channel Encephalopathy. VGKC antibodies are linked to a group of rare disorders characterized by abnormal neuromuscular excitability and CNS manifestations. These disorders form part of a larger group of CNS autoimmune disorders involving antibodies directed against targets on the neuronal surface. Isaac syndrome (acquired neuromyotonia) is associated with thymoma and VGKC antibodies but has no CNS manifestations. Morvan syndrome is a very rare disorder characterized by neuromyotonia, severe insomnia, excessive sweating,

hypersalivation, and a subacute encephalopathy commonly accompanied by psychotic features including delusions and hallucinations. High titers of VGKC of antibodies are also detected. A group of patient with nonparaneoplastic limbic encephalitis have high titers of VGKC antibodies. This group presents with complex-partial seizures and a progressive amnestic syndrome with sleep disturbance as well as fluctuating conscious state, personality change and variable psychotic symptoms. MRI frequently demonstrates bilateral high signal in the hippocampal region. These symptoms are steroid responsive, including reversal of cognitive deficits and MRI changes, once the diagnosis has been made (Fig. 2.14–18).

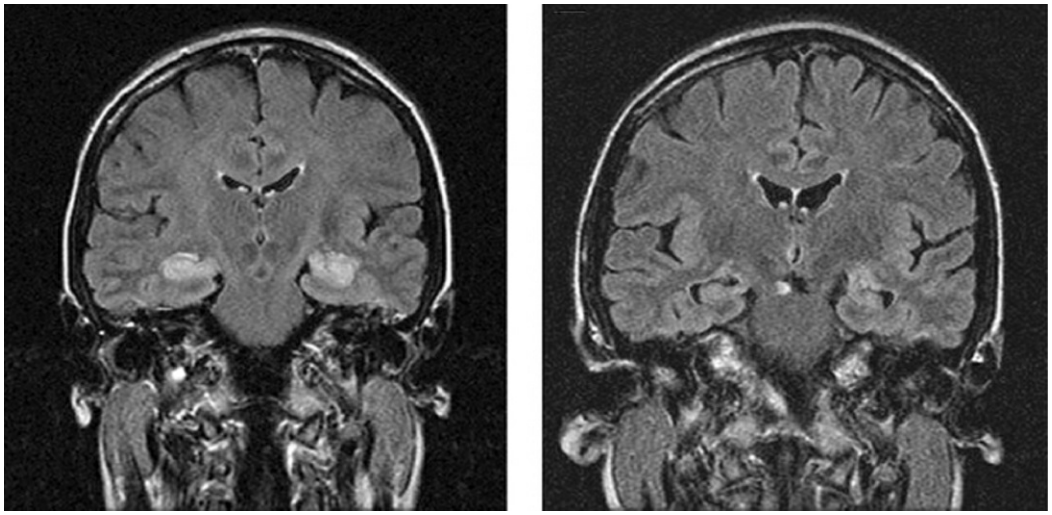

FIGURE 2.14–18. Voltage-gated potassium channel antibodies causing limbic encephalitis. **Left:** Coronal fluid-attenuated inversion recovery magnetic resonance image shows bilateral mesial temporal hyperintensity. **Right:** Identical coronal slice 3 months later, after treatment, shows resolution of hyperintensities but resultant hippocampal atrophy.

Timothy Syndrome. Timothy syndrome is a relatively recently recognized multisystem Ca^{++} channelopathy in which autism-spectrum disorders occur in 80 percent of affected individuals, alongside cardiac arrhythmias and syndactyly. The affected $Ca_v1.2$ gene is widely expressed, particularly in heart, brain, smooth muscle, and pituitary and adrenal glands. In the CNS, highest expression is in the granular layer of the dentate gyrus of the hippocampus, and cerebellum. It is not clear how changes to neuronal tissue excitability contribute to the development of autism, but it is likely to represent the endpoint of the interaction between tissue excitability and normal neurodevelopment. A greater understanding of the pathophysiology of this disorder may open up new avenues of research into the possible contribution of ion channel disturbance to polygenic neuropsychiatric disorders.

NEUROENDOCRINE DISORDERS

Hypothalamic Disorders

The relationship between the hypothalamus and pituitary gland is complex. In brief the release of six anterior pituitary hormones prolactin (PRL), growth hormone (GH), follicle stimulating hormone (FSH), luteinizing hormone (LH), adrenocorticotropin (ACTH), and thyrotropin or thyroid stimulating hormone (TSH) is under the tonic influence of hypothalamic neuropeptides which travel from hypothalamic neurons via the portal system of the anterior pituitary to influence pituitary hormone producing cells. AVP (vasopressin or ADH, antidiuretic hormone) and PRL are produced in the hypothalamus and stored in the posterior pituitary. As a consequence of this relationship endocrine abnormalities of the hypothalamus are largely manifest as abnormalities of pituitary function as described below. However, the hypothalamus also has a number of nonendocrine functions that when disturbed may manifest as neuropsychiatric disorders.

The hypothalamus contains a number of unique neurons that create two neuropeptides known as *orexins* (formerly *hypocretins*). They are synthesized only in the hypothalamic neurons, and share some homology to the gut hormone secretin. Orexinergic neurons project from the hypothalamus to a number of monoaminergic centers including the locus coeruleus, raphe nuclei, and ventral tegmentum. A number of these monoaminergic systems are involved in the regulation of sleep, in which the hypothalamus is now recognized as a key center. The contribution of the hypothalamus is via sleep-promoting GABAergic neurons of the ventrolateral preoptic area (VLPO) and the wakefulness-promoting orexinergic neurons in the lateral hypothalamus. These pathways are closely linked to the circadian pacemaker in the suprachiasmatic nuclei and the regulation of other hypothalamic functions such as temperature, food intake, metabolism, and hormone secretion.

The orexinergic system is also important in the regulation of food intake and energy expenditure. Orexin production can increase food craving, but is also inhibited by leptin, a hormone produced by adipocytes, and stimulated by ghrelin. Ghrelin is secreted by the stomach just prior to a meal and is known to stimulate caloric intake. This provides a biochemical basis for the well-demonstrated phenomenon of sleep-deprivation related catabolism despite adequate caloric intake described in animal models.

Narcolepsy

Narcolepsy is a sleep disorder characterized by excessive daytime somnolence, cataplexy (a sudden loss of muscle tone, often triggered by strong emotional reactions) and manifestations of disordered REM sleep such as hypnagogic hallucinations, automatic behavior, and sleep paralysis. Although narcolepsy in dogs is linked to mutations in hypocretin-related

genes, no such mutations have been described in humans, although in approximately 90 percent of patients, orexin levels in the CSF are low or nonexistent. Strong links to the human leukocyte antigen (HLA) system suggest an autoimmune basis. HLA-DR15 and HLA-DQ6 have been described in up to 85 percent of patients but only 20 percent of controls, and DQB1*602 allele has been detected in 98 percent of patients with cataplexy. A gene–environment interaction is likely given the <25 percent concordance rate between monozygotic twins. Depression has been described in up to 25 percent of patients in some series, although others suggest a rate no higher than that in the healthy population. Whether narcolepsy is associated with psychosis is very controversial. It has been suggested that some psychoses are "narcoleptic" in origin and respond to treatment with stimulants, and that REM-intrusion into wakefulness may be misdiagnosed as schizophrenia. The vast majority of psychoses associated with narcolepsy do appear to relate to stimulant use/misuse, as historically dopaminergic agents such as dexamphetamine have been used to promote wakefulness. Newer agents such as modafinil may be less psychotogenic. Neuroleptics however worsen narcoleptic somnolence, as dopamine outflow is central to cortical wakefulness.

A 23-year-old man was referred for assessment of psychotic symptoms occurring in the setting of narcolepsy. At age 13, narcolepsy was diagnosed on the basis of daytime somnolence, cataplexy, sleep paralysis, and hallucinations. He showed an HLA-DR15 haplotype. He had begun dexamphetamine 2 years prior to assessment and started to experience auditory hallucinations, thought broadcasting, and a complex persecutory delusional system (whereby he was being studied through a device implanted in his head) within 6 months. His family described a precipitous psychosocial decline over 12 months, as the patient became unemployed, socially withdrawn, and increasingly disorganized. On mental state, he presented as fatuous and inappropriate, with clear thought disorder. After commencing fluphenazine 4 mg and ceasing dexamphetamine, the hallucinations and delusions diminished, although he remained fatuous and amotivated. A diagnosis of narcolepsy-related psychosis secondary to dexamphetamine treatment was made, with a differential diagnosis of schizophrenia.

Hypothalamic Lesions

Hypothalamic obesity has been associated with lesions of the ventromedial nucleus, and may initially involve aggressive behavior and hyperphagia until a new set weight is reached, at which time reduced appetite and activity may manifest. Rage reactions are also well described in animals with lesions in this region of the hypothalamus, although less so in humans with such lesions. Lateral lesions have been reported to result in an apathetic state. Thirst may also be impaired if ADH production is reduced by a hypothalamic lesion. Short-term memory dysfunction has been reported, particularly in lesions of the ventromedial and premamillary areas of the hypothalamus. Extensive hypothalamic lesions may produce

features consistent with a dementing illness.

Hypothalamic disease may also result in abnormalities of thermoregulation. Temperature sensitive neurons are located in the anterior hypothalamus, whereas the posterior hypothalamus mediates heat loss mechanisms. Acute lesions such as hemorrhage, infarction or those from surgical procedures may result in acute and paroxysmal hypothermia. Conversely, posterior hypothalamic lesions may result in paroxysmal hyperthermia with associated fevers and rigors.

Childhood tumors invading the anterior and basal hypothalamus such as gliomas, midline cerebellar astrocytomas, and suprasellar ependymomas may result in the diencephalic syndrome which is manifest as motor hyperactivity, euphoria or inappropriate affect, increased alertness, and emaciation despite normal caloric intake. If death does not ensue the clinical picture may change to one of obesity and intermittent rage reactions. In adulthood, slower growing tumors usually result in a dementia syndrome, endocrine dysfunction, and food intake dysregulation. More rapid or destructive processes present with disturbances of consciousness, temperature, and autonomic dysregulation. Craniopharyngiomas, germinomas, and gliomas are the most frequently reported tumors. Craniopharyngioma, remnants of the embryonic Rathke pouch can manifest in both childhood (most common) and adulthood, usually involving the posterior hypothalamus. As most are suprasellar, patients usually present with visual abnormalities and headaches. Hypogonadism, hyperprolactinemia, diabetes insipidus, and weight gain are common as is cognitive deterioration and personality change without evidence of psychosis (Fig. 2.14–19).

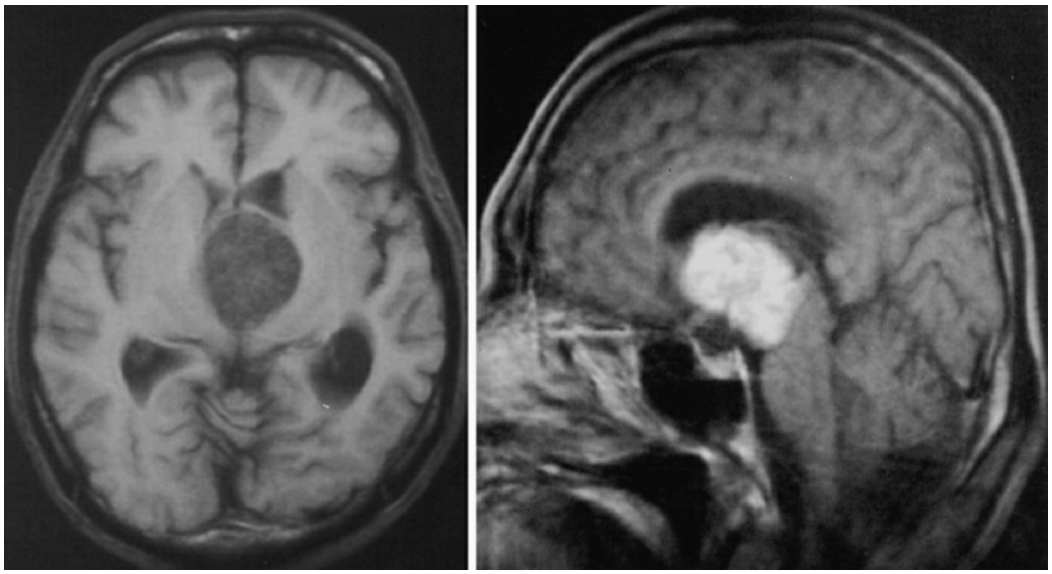

FIGURE 2.14–19. Craniopharyngioma causing the diencephalic syndrome. T1-weighted magnetic resonance imaging scans of a patient who presented with emaciation and features of the diencephalic syndrome. **Left:** Axial image demonstrating a slightly hypointense tumor within the third ventricle accompanied by moderate dilation of the

lateral ventricles. **Right:** Sagittal image obtained after the administration of gadolinium revealing that the tumor is homogeneously enhanced and is entirely confined within the third ventricle. (From Miyoshi Y, Yunoki M, Yano A, Nishimoto K, Konovalov AN. Diencephalic syndrome of emaciation in an adult associated with a third ventricle intrinsic craniopharyngioma: A case report. *Neurosurgery*. 2003;52[1]:224, with permission.)

Pituitary Disorders

The pituitary gland, or *hypophysis,* is an endocrine structure that sits in the midline sella turcica at the base of the brain in the middle cranial fossa, and is covered by a dural fold (the *diaphragma sellae*). It secretes hormones regulating homeostasis, and through the release of trophic hormones, stimulates other distal endocrine structures. Its anterior lobe, the *adenohypophysis,* is under direct functional control of the hypothalamus, via the hypophyseal-portal vascular connection in the pituitary stalk, through which stimulatory and inhibitory signals are sent to control the five distinct endocrine cell types that release pituitary hormones. The posterior lobe, the *neurohypophysis,* is predominantly a collection of axons from the supraoptic and paraventricular nuclei of the hypothalamus that secretes peptide hormones into the hypophyseal circulation. The posterior lobe is connected by the infundibulum in the pituitary stalk, and the release of oxytocin and vasopressin is controlled through the tuberoinfundibular pathway. Given the complex nature of the hypothalamic–pituitary axis, it is not surprising that clinical manifestations of pituitary disease are protean. Pituitary dysfunction may result from pituitary tumors or destructive disease processes.

Pituitary Tumors

Pituitary tumors are found in up to 20 percent of adults at autopsy and are frequent incidental findings at neuroimaging. They may result in increase or decrease of hormone levels and produce symptoms by invading surrounding structures such as the hypothalamus.

Prolactinomas. The most common secretory tumors are prolactinomas. Dysregulated secretion of PRL results in amenorrhea, infertility, and galactorrhea in women and impotence and occasionally galactorrhea or gynecomastia in men. Data regarding the neuropsychiatric manifestations of hyperprolactinemia is lacking. There is some evidence of increased aggression in lactating animals associated with elevated PRL levels, and in hyperprolactinemic human subjects. Depression and anxiety symptoms also occur with greater frequency in this group of patients, which may respond to treatment with the dopamine agonist bromocriptine. Psychotic symptoms have been described in neuroleptic naive patients with hyperprolactinemia at a case report level. Treatment with antipsychotic agents may then result in further elevation of PRL. Psychotic symptoms

have also been described in patients receiving bromocriptine for the treatment of a prolactinoma, with both delusions and hallucinations recorded. The incidental finding of a pituitary adenoma on neuroimaging can complicate the assessment of patients with a psychotic illness stabilized on an antipsychotic agent with hyperprolactinemia. Impulse control disorders such as pathological gambling have also been reported in patients receiving cabergoline (a potent Dopamine type 2 receptor agonist) therapy for pituitary adenomas.

Growth-Hormone Secreting Tumors. These are the second most common functional pituitary adenomas (Fig. 2.14–20). In adults they result in *acromegaly* with soft tissue and bone enlargement, particularly involving the hands, feet, jaw, and tongue with a characteristic overall coarsening of facial features. There may be associated hypertension, congestive cardiac failure and obstructive sleep apnea, and hypersomnolence. A number of psychiatric symptoms have been associated with acromegaly, largely at the case report level. Acromegalic patients appear to have a lifetime increased risk of affective disorders when compared with control subjects with and without chronic somatic disease. There is no evidence of a preponderance of psychotic symptoms although these are well described when acromegalic patients are treated with bromocriptine resulting in delusional symptoms, schizophrenia-like presentations, and visual hallucinations.

Adrenocorticotrophin Secreting Tumors. ACTH secreting tumors are the next most common pituitary tumor type, resulting in excessive cortisol production or Cushing disease. The neuropsychiatric manifestations of increased cortisol release are discussed below in the description of adrenal disease.

Other Pituitary Tumors. Clinical symptoms resulting from tumors producing LH or FSH are very rare. Increased TSH production can occasionally result in hyperthyroidism, the psychiatric aspects of which are discussed below. As many as 30 percent of pituitary tumors are nonsecretory, and are often larger at diagnosis because of the lack of endocrine manifestations. If large enough pituitary tumors may impinge on surrounding structures resulting in the classic visual field defect of bitemporal hemianopia, oculomotor palsies headache, and occasionally hypothalamic syndromes (see above). Visual hallucinations in the context of visual field defects related to pituitary tumors impinging the optic chiasm have been reported. More often these are of the simple, nonformed type and can be exacerbated by treatment with bromocriptine.

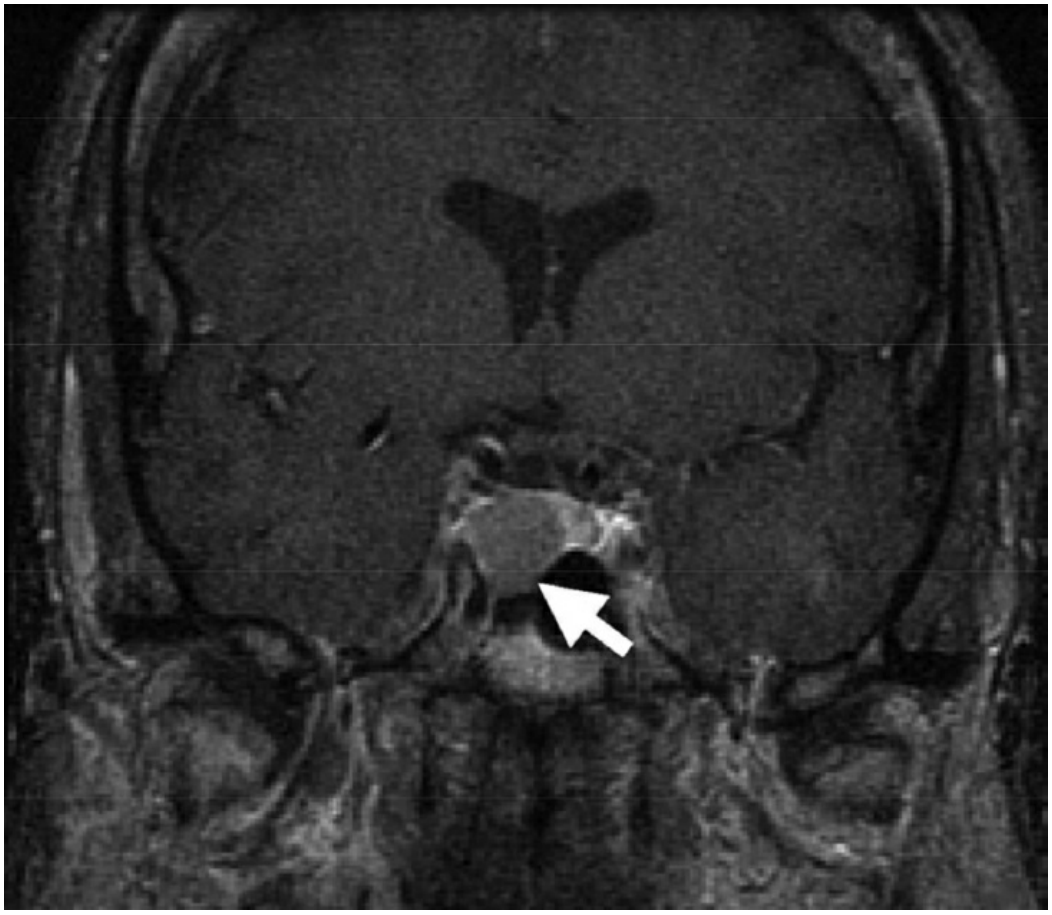

FIGURE 2.14–20. Pituitary adenoma causing acromegaly. Gadolinium-enhanced magnetic resonance imaging of the brain showing a well-defined mass in the right lateral aspect of the sella with lack of enhancement compared with the normal pituitary gland. Endocrinologic evaluation revealed elevated human growth hormone and insulin-like growth factor levels consistent with acromegaly. (From Koo CW, Bhargava P, Rajagopalan V, Ghesani M, Sims-Childs H, Kagetsu NJ. Incidental detection of clinically occult pituitary adenoma on whole-body FDG PET imaging. *Clin Nucl Med*. 2006;31[1]:42, with permission.)

Hypopituitarism

Deficiency or dysfunction of one or more of the pituitary hormones is referred to as hypopituitarism. In adulthood, this is usually caused by an acquired destructive process that may be traumatic (related to head injury), inflammatory, immune mediated, vascular, or as the result of compression from an adjacent tumor. Clinical manifestations are dependent on the hormones involved. GH is often first affected followed by the gonadotropins with associated amenorrhea and infertility in women and reduced libido and body hair loss in men. Low TSH can result in hypothyroidism, ACTH deficiency in fatigue, reduced appetite, weight loss, and impaired stress response. Vasopressin deficiency may also ensue, resulting in polyuria and thirst. A number of neuropsychiatric manifestations have been described in hypopituitarism in the absence of a delirium. Memory impairment, sleep disturbance, and personality change are commonly but variably reported. Visual and auditory hallucinations

have been described at a case report level, and several observers have reported an absence of affect on mental state examination. A systematic study of psychiatric comorbidity in hypopituitarism found more psychiatric symptoms than those expected in chronic disease such as diabetes mellitus, most particularly depression and anxiety. Psychosis is uncommon.

More specific symptoms are attributable to particular hormone deficiencies. GH deficiency may result in reduced energy, depressed mood, anxiety, emotional lability, and impulsivity or impaired self-control. Testosterone deficiency has been associated with depression irritability and insomnia. Fatigue, social withdrawal, and negativism have been reported when ACTH is low. Lastly, as in hypothyroid states fatigue, depression, insomnia or hypersomnia, and psychotic symptoms have been described when TSH levels are low. There can be considerable variation in the type of psychotic symptoms described. In general, appropriate correction of the hormone deficiency and addressing the underlying cause results in resolution of these features.

Thyroid and Parathyroid Disorders

Normal thyroid function involves the production of the two thyroid hormones, L-thyroxine (T_4) and 3,5,3′-triiodo-L-thyronine (T_3), which are required for the regulation of a number of metabolic processes. Increased thyroid hormone production results in hypermetabolism or increased caloric utilization and the other clinical features of hyperthyroidism. Hypometabolism and the features of hypothyroidism (sometimes referred to as myoedema) results from reduced hormone production. Psychiatric symptoms that accompany either of these states have been subject to more systematic study than those associated with pituitary disorders, as primary thyroid disorders have a much higher incidence in the population.

Hyperthyroidism. Hyperthyroidism is the result of excessive production of thyroid hormones. This may result from a toxic multinodular goiter, a single functioning adenoma, or from the presence of a thyroid stimulator, such as a thyroid stimulating antibody in Grave disease. Exogenous thyroid hormone can produce a similar picture as can disorders of thyroid hormone storage consequent to autoimmune thyroiditis.

Depressive symptoms are not only the most common psychiatric features seen in hyperthyroidism, occurring in up to 30 percent of patients, but frequently occur prior to the onset of the other physical features. This includes lowering of mood as well as neurovegetative disturbance such as insomnia, reduced libido, weight loss, and fatigue. However, as distinct from depression, appetite is invariably increased in hyperthyroidism. The severity of the depressive symptoms have not been found to be related to the severity of hyperthyroidism as measured by subsequent thyrotoxic features and the extent of the biochemical abnormalities.

Elderly patients with thyrotoxicosis may present with predominantly apathy, depression, and weight loss rather than increased psychomotor activity. Typically this presentation is of slower onset, the neurological and ophthalmological features are less prominent, but cardiovascular events such as exacerbation of angina cardiac failure and atrial fibrillation are more prevalent. Anxiety symptoms presenting as generalized anxiety are also common, with a prevalence of between 10 and 20 percent, but panic and agoraphobia are relatively uncommon. Like depressive symptoms, anxiety may occur prior to the other features of hyperthyroidism, but more often correlates with the severity of the thyrotoxic features. Anxiety and depressive features are often comorbid. Manic symptoms are less common in hyperthyroid states with a prevalence between 2 and 5 percent, but may be difficult to distinguish from psychomotor agitation and anxiety. Psychotic symptoms including paranoid delusions and auditory hallucinations, while historically reported as common, have a true prevalence of between 2 and 5 percent. The depressive syndrome seen in hyperthyroidism rarely requires treatment other than that which is used to restore the euthyroid state; however, some features, such as anxiety, fatigue, and loss of function may persist for as long as 12 months after a euthyroid state has been achieved.

Cognitive dysfunction is reported between 5 and 10 percent of patients with thyrotoxicosis, but to a lesser degree than that seen in hypothyroidism. Patients may present with slow processing speeds, impairments in immediate memory, defective higher-level problem solving, or frank delirium. Mild disorders of attention and concentration area common, but their severity do not always correlate with the severity of thyrotoxicosis. These deficits invariably respond to reversal of the thyrotoxic state.

Hypothyroidism. The most common cause of hypothyroidism in adults is primary autoimmune hypothyroidism related to antithyroid antibodies. Other causes include treatment for hyperthyroidism, drug-related effects, and iodine deficiency. Supra-thyroid causes (hypothalamic or pituitary dysfunction) account for less than 5 percent of cases. The symptoms of hypothyroidism may include fatigue, lethargy, weight gain, constipation, cold intolerance, stiffness and cramping of muscles, hair loss, cognitive slowing, and depression. Signs include, hypothermia, bradycardia, dry skin, sparse hair, periorbital swelling, thickening of the tongue, coarsening and deepening of the voice, and a characteristic prolonged relaxation phase of deep tendon reflexes. This clinical picture is often referred to as *myxedema*. Hypothyroid patients may present with a variety of psychiatric symptoms ranging from mild cognitive slowing and depression to frank encephalopathy which often predates other physical features. Functional neuroimaging studies have demonstrated global hypometabolism and more

specific areas of hypoperfusion in the posterior cingulate, insula, fusiform gyrus, and right parieto-occipital and primary motor cortex.

Cognitive deficits are the most common neuropsychiatric features of hypothyroidism, occurring in up to 50 percent of cases. Psychomotor speed, memory, and visual–perceptual skills are often impaired. Difficulties with constructional skills, reduced performance in trail-making and maze tasks also suggest prominent executive deficits. The severity of these disorders is correlated with the degree of biochemical abnormality, and although largely corrected by return to a euthyroid state, some cognitive deficits may remain, particularly in the elderly or those with reduced cognitive reserve. Severe disturbance of consciousness, including coma and delirium may be encountered, but may be associated with other metabolic changes related to the hypothyroid state such as hyponatremia. Depression is only slightly less common than cognitive disturbance, reported in approximately 40 percent of patients, but appears less closely related to the severity of biochemical hypothyroidism. Low mood, fatigue, anhedonia, reduced concentration, and hypersomnolence are the most commonly described features of the depressive syndrome in hypothyroidism. These features predictably respond to treatment of the hypothyroid state. The origin of depression in hypothyroidism appears to relate to the role of thyroxine in serotonergic transmission, such that reduced thyroid input reduces serotonergic tone and lowers the threshold toward the development of depressive symptoms. Conversely, thyroid replacement restores central serotonin activity in concert with improvement in depressive symptoms. This may also underpin the adjunctive antidepressant effect of thyroxine.

In contrast, manic and hypomanic symptoms have been infrequently reported in association with hypothyroidism. However, hypothyroidism may be a risk factor for the development of bipolar disorder, particularly the rapid cycling form and treatment of otherwise refractory mood disorders with thyroid hormones has occasionally been shown to be effective. Generalized anxiety symptoms are described in up to 30 percent of patients and are strongly correlated with depressive symptoms but not with biochemical severity. Psychotic symptoms including paranoid ideas, misidentification, visual and auditory hallucinations, and thought disorder, were originally thought to be common (and described as “myxedematous madness”), but likely occur in less than 5 percent of all patients with hypothyroidism and tends to emerge after the onset of physical symptoms. Although these symptoms also respond to appropriate thyroid hormone treatment, rapid titration of hormone doses may exacerbate psychosis. Careful addition of a low dose of antipsychotic to thyroxine has been reported to be well tolerated and result in an earlier remission of psychosis.

“Subclinical” Hypothyroidism. Thyroid hormone abnormalities may

occur without overt functional hypothyroidism. Designated *subclinical hypothyroidism,* these scenarios can be further classified into elevated TSH without changes in thyroid hormones (Grade II hypothyroidism), abnormal TSH response to stimulation with TRH (Grade III), and the presence of antithyroid antibodies with no thyroid hormone system abnormalities (Grade IV). Grade II hypothyroidism has been associated with depressive disorders. Patients with major depressive disorders have an increased incidence of Grade II hypothyroidism, and some studies show these patients respond poorly to conventional treatment. Grade II hypothyroidism may be a risk factor for major depressive disorders. Cognitive disturbance, particularly memory and psychotic symptoms have also been reported in Grade II hypothyroidism. Although there is some evidence of improvement in cognition, mood, and psychosis, treatment of subclinical hypothyroidism is controversial.

Hashimoto Encephalopathy. Hashimoto encephalopathy (HE), also described as steroid-responsive encephalopathy associated with autoimmune thyroiditis (SREAT) is best defined as an uncommon autoimmune encephalopathy of unknown etiology associated with high titers of serum antithyroid (usually antithyroid peroxidase ± antithyroglobulin) antibodies. The role of antithyroid antibodies is controversial. There is no direct evidence that these auto-antibodies exert direct effects on CNS tissue, and they may be epiphenomena of another autoimmune process. In addition, the base rate of elevations of these auto-antibodies in the population is as high as 10 percent. Elevated serum antithyroid antibodies are associated with other thyroid disorders (Grave disease, de Quervain thyroiditis, primary hypothyroidism, and colloidal goiter) and other auto-immune disorders (including diabetes mellitus, Addison disease, and pernicious anemia).

HE is more common in women with a mean age of onset between 45 and 50 years. There may be a history of other autoimmune disease. The clinical picture is one of encephalopathy with progressive cognitive decline although the course may also be relapsing and remitting. Common features include seizures (>95 percent of cases), stroke-like episodes (>65 percent of cases), and memory dysfunction. Neuropsychiatric features include agitation and restlessness, apathy, social isolation. Visual hallucinations are frequently reported (>90 percent), as are other disorders of perception and delusions. Patients with a more typical presentation of psychiatric illness, such as depression, in the setting of mildly elevated antithyroid peroxidase antibodies are unlikely to have HE. However, when antithyroid peroxidase levels are very high (>1,000 IU/L), the likelihood of them being associated with neuropsychiatric HE is much greater.

Routine investigations are often normal, although patients will often show slowing on EEG, with the degree of slowing being proportional to

clinical severity of the syndrome, and some patients may show triphasic waves (Fig. 2.14–21). The majority of patients respond to corticosteroid treatment of with complete resolution of the neuropsychiatric symptoms.

A 66-year-old woman presented with a four month history of cognitive decline including short-term memory and language deficits. She had a generalised tremor and developed a delusional belief that her husband wished to harm her. She became agitated and disoriented, particularly at night, and also described visual hallucinations of small animals in her room. No neurological features, including myoclonus, were described. At this time all routine blood investigations (including complete blood count, serum urea, electrolytes, creatinine, liver enzymes, thyroid hormone assays, folate and B_{12} level and CT brain) were normal. One month later she suffered a generalized tonic–clonic seizure. Routine blood investigations and examination of the CSF were normal. Angiography of the anterior and posterior cerebral circulation was normal. MRI scan was normal and EEG was unremarkable. On mental state examination she was disoriented with impaired attention and markedly poor short-term memory (particularly nonverbal memory). Her speech was characterized by impaired verbal fluency and word finding problems. There was a concrete thinking style with reduced abstract reasoning as well as poor judgment. There were no depressive or psychotic features at this time. Her then mental state rapidly deteriorated. She became virtually mute, with poor concentration and attention and required assistance with all aspects of personal care. She had bilateral grasp and pout reflexes as well as diffuse hyperreflexia. She also had bilateral and multifocal myoclonus. In the face of previously normal investigations, the diagnosis of Hashimoto's encephalopathy was considered. Thyroid hormone (TSH, T4, T3) levels were within normal limits. Anti thyroid peroxidase antibodies (anti-TPO) were raised in titre (640 IU/ml, normal <50I U/ml) consistent with this diagnosis. A course of prednisolone 60 mg was commenced with marked improvement in her mental state within 2 weeks.

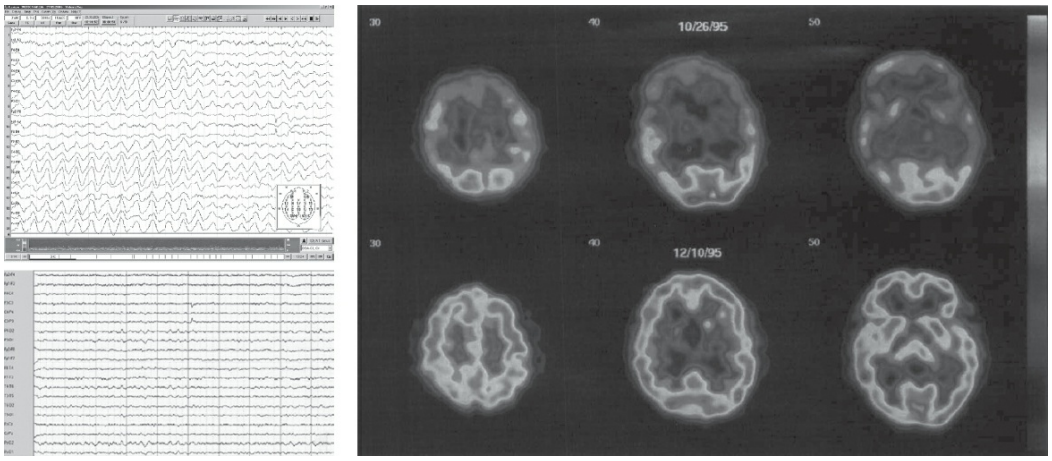

FIGURE 2.14–21. Reversibility in Hashimoto encephalopathy. **Left:** Electroencephalogram (EEG) before and after treatment in a 38-year-old woman with psychosis and an antithyroid peroxidase antibody titer of 779. **Top:** Pretreatment with corticosteroids showing general slowing with high voltage (2 to 3 Hz) δ biphasic and triphasic waves. **Bottom:** EEG after corticosteroid treatment showing α frequency, with occasional θ waves (5 to 6 Hz), mainly in posterior regions. (From Sporis D, Habek M, Mubrin Z, Poljakovic Z, Hajnsek S, Bence-Zigman Z. Psychosis and EEG abnormalities as manifestations of Hashimoto encephalopathy. *Cog Behav Neurol.* 2007;20[2]:138, with permission.) **Right:** Single photon emission computed tomography (SPECT) showing gross global hypoperfusion in all nonoccipital regions in a 59-year-old woman with rapidly progressive cognitive impairment and myoclonus, with Mini-Mental State Examination (MMSE) score of 20 (**top**) and after significant clinical improvement 6 weeks later, when

MMSE score was 27 (**bottom**). (See color atlas.) (From Forchetti CM, Katsamakis G, Garron DC. Autoimmune thyroiditis and a rapidly progressive dementia: Global hypoperfusion on SPECT scanning suggests a possible mechanism. *Neurology.* 1997;49:623, with permission.)

Parathyroid Disorders

The parathyroid glands are small accessory glands that are anatomically associated with but functionally distinct from the thyroid gland. Their sole function is to maintain serum calcium levels within a narrow range to permit optimum functioning of the nervous and muscular systems, through the release of parathyroid hormone (PTH). The release of PTH increases serum calcium via stimulating osteoclasts in bone to release calcium, and through increasing its absorption in the gut and kidneys.

Hyperparathyroidism. Hyperparathyroidism is usually diagnosed after an incidental finding of hypercalcemia on routine blood tests, and 50 percent of patient with hyperparathyroidism are asymptomatic. Increased PTH release is usually caused by a single functioning adenoma, although multiple adenomas may occur as part of multiple endocrine neoplasia syndrome (MENS). Symptoms are principally those secondary to hypercalcemia. These include fatigue, general malaise, proximal muscle weakness, renal colic, abdominal pain, and cognitive decline. There is often little to find on examination although real calculi, nephrocalcinosis, or bone changes (osteitis fibrosa cystica, now rare) may be seen on radiological investigations ordered for other reasons. Although a number of psychiatric symptoms have been described in hyperparathyroidism (as part of the symptomatic tetrad of "bones, stones, moans, and psychic groans"), the prevalence overall is likely to be less than 10 percent. These symptoms correlate with both duration and severity of the associated hypercalcemia. Affective disorders predominately of the depressive type are the most frequently recorded, with anxiety often comorbid. Psychotic disorders are reported, although infrequent, with persecutory and paranoid delusions predominating. Cognitive changes, usually related to short-term memory loss, disorders of attention and acute confusional states often occur, and severe derangement of calcium metabolism may present with somnolence and coma. Neuropsychiatric symptoms are often the initial presentation in the elderly, or those with limited cognitive reserve. The more severe the hypercalcemia, the more severe the psychiatric disturbance, but symptoms generally respond to appropriate treatment such as parathyroidectomy. When PTH levels are mildly raised in the setting of normocalcemia (most commonly due to vitamin D insufficiency, increasingly prevalent in developed countries), psychiatric disturbance is uncommon.

Hypoparathyroidism. The most common cause of impaired PTH production in adults is inadvertent surgical removal during thyroid surgery

or excessive removal for hyperparathyroidism. Clinical features relate to hypocalcemia, particularly symptoms of neuromuscular excitability including paresthesias, muscle cramps, carpopedal spasm, facial grimacing progressing to laryngeal spasm, and convulsions. Examination may reveal features of tetany, reduced or absent deep tendon reflexes, papilledema, and QT interval prolongation on ECG. Delirium is now understood to be the most common neuropsychiatric manifestation. One large study available demonstrated cognitive impairment in 39 percent of patients, affective or neurotic symptoms in 12 percent, psychotic symptoms in 11 percent, and nonspecific affective disturbance in 21 percent of patients. Again, severity of symptoms directly relate to the degree of hypocalcaemia and appropriate normalization results in resolution of these symptoms, although persistent psychosis has been noted when associated hypomagnesemia was not addressed.

Adrenal Disorders

The adrenal glands produce both adrenal steroids from the cortex and catecholamines from the medulla. The adrenal cortex produces glucocorticoids, principally cortisol under the stimulatory effects of ACTH from the anterior pituitary controlled by a negative feedback loop. Aldosterone, a mineralocorticoid, originates from the adrenal cortex but is controlled by the renin–angiotensin system influenced by body volume and potassium balance. Adrenal androgens, predominately dehydroepiandrosterone (DHEA) are also regulated by the ACTH system and undergo peripheral conversion to sex-determining androgens. Hyperfunction or hypofunction of these adrenal systems result in distinct clinical syndromes with complex physical and neuropsychiatric manifestations.

Hyperadrenalism. Hyperadrenalism (Cushing syndrome) occurs when the adrenal gland produces excess corticosteroids, usually as a result of ACTH overproduction from the anterior pituitary secondary to a pituitary adenoma (Cushing disease). Other causes include ACTH production from a nonendocrine tumor, an adrenal neoplasm or most commonly, exogenous steroids, usually prescribed for treatment of a steroid-sensitive medical condition. Patients with Cushing syndrome present with the classic features of hypertension, muscle weakness and fatigue, osteoporosis, cutaneous striae, and easy bruising. A characteristic pattern of obesity is seen involving the upper face (resulting in a “moon face”), back (“buffalo hump”), and mesentery resulting in truncal obesity. Women may experience hirsutism, acne and amenorrhea, men, decreased libido and impotence. Neuropsychiatric features are well-described, with depression the most frequently noted psychiatric symptom in up to 70 percent of patients. Anxiety is commonly comorbid with depression, in up to 50

percent of sufferers. Depressive symptoms may present prior to physical symptoms. They respond to treatment of the underlying cause and the response correlates with lowering of plasma cortisol levels. Elevated mood is infrequently reported in Cushing syndrome, reported in less than 10 percent in most case series. Delirium is also relatively uncommon and is usually a marker of a supervening infection or other metabolic disorder such as metabolic alkalosis. Although dramatic case reports of Cushing syndrome presenting with psychosis and obfuscating the underlying condition are reported, this form of presentation is rare. When present, psychotic symptoms are almost always mood congruent delusional beliefs and derogatory auditory hallucinations associated with a depressed mood. Cognitive impairment occurs in over 50 percent of patients, presenting as deficits in verbal memory, attention, and visuomotor and visuospatial function. The degree of memory impairment, and its reversibility, appears to correlate with hippocampal volume, suggesting that cognitive impairment is partially driven by the effect of excess glucocorticoids on hippocampal neurons.

With exogenous administration of steroids, manic symptoms are most frequently reported, followed by delirium, depression, and psychotic symptoms. Mixed affective states also appear over-represented. The severity of these symptoms may be dose related, reproducible in the individual, and minimized with use of divided doses of steroids. Manic symptoms may respond to both cessation of the exogenous steroid administration and pharmacotherapy with mood stabilizing agents such as lithium and valproate, or treatment with antipsychotics.

Hypoadrenalism. Primary hypofunction of the adrenocortical system may result from a primary process at the level of the gland such as destruction by autoimmune process, referred to as Addison disease (most common), other inflammatory or destructive processes, or an inborn failure of enzyme function. It may also arise secondary to dysfunction of the hypothalamic–pituitary axis or withdrawal of exogenous steroids. The onset of symptoms is usually insidious with progressive fatigue, weakness, anorexia, nausea, abdominal pain, weight loss, cutaneous and mucosal pigmentation, hypotension, and hypoglycemia. Other findings include hyponatremia, hyperkalemia, and metabolic acidosis. Reports of neuropsychiatric symptoms in Addison disease are uncommon, although the true prevalence is uncertain due to probable under-reporting. Depression, reduced motivation, and energy and behavioral changes predominate. Memory dysfunction is the most common form of cognitive disturbance. Paranoid symptoms and delusions are less common. Catatonia and self-mutilation are rare but dramatic presentations. Auditory and visual hallucinations, changes in conscious state, irritability, insomnia, and nightmares often herald an Addisonian crisis with frank delirium,

coma, and seizures. Hyponatremia and metabolic acidosis may contribute to the delirium and cognitive deficits. These symptoms appear to largely resolve, including those of adrenal crisis, when treatment with adequate doses of corticosteroids is commenced. Use of other psychotropic agents is rarely indicated.

Hyperaldosteronism. Excess of the major adrenal mineralocorticoid aldosterone can arise from the adrenal gland (primary aldosteronism) or from an extra-adrenal site (secondary aldosteronism). Primary aldosteronism is usually due to an aldosterone producing adenoma (Conn syndrome) or bilateral cortical hyperplasia. Secondary aldosteronism relates to disturbance of the renin–angiotensin system and hypovolemia. The characteristic features of Conn syndrome are hypokalemia, hypertension, muscle weakness, fatigue, polyuria, and polydipsia. Metabolic alkalosis and hypomagnesemia may also occur and edema is characteristically absent. Depression has been identified as one of the major features but more recent studies have demonstrated features of generalized anxiety with individual cases of associated depression, features of OCD and panic without specific gender differences. Untreated primary hyperaldosteronism, and resulting severe hypertension, can result in a vascular dementia if left untreated. Although treatments, such as surgical excision of a functioning adenoma are effective in reversing the physical symptoms of hyperaldosteronism, the effect on anxiety and other symptoms are unknown.

A 48-year-old man with no past psychiatric history had raised family concerns following 3 years of unusual behavior, social withdrawal, and poor self-care. He lost contact with friends, began to gamble heavily, had a number of car accidents, and was disinhibited and socially inappropriate. He had recently started a fire in his residence with one of many discarded cigarettes. On examination, he was hypertensive (240/120 mm Hg) and had signs of left ventricular hypertrophy, but no neurological abnormalities. On cognitive assessment, he had a marked dysexecutive syndrome. Serologically he had a mild metabolic alkalosis and hypokalemia (3.2 mmol/L). Serum cortisol, ACTH, and vasculitic screen were normal. ECG showed first-degree heart block and left ventricular hypertrophy, which was confirmed on echocardiography. There was failure of aldosterone suppression with saline challenge and bilateral adrenal hyperplasia on abdominal CT. Brain MRI showed extensive severe periventricular and subcortical white matter disease (Fig. 2.14–22). He was diagnosed with Binswanger's dementia secondary to hypertension associated with hyperaldosteronism.

Pheochromocytoma. Pheochromocytomas are tumors which secrete catecholamines and most commonly originate in the chromaffin cells of the adrenal medulla, but may rarely arise from similar cells in sympathetic ganglia. Familial forms are associated with MENS types 2a and 2b and von-Recklinghausen neurofibromatosis. Secretion of catecholamines may be continuous or sporadic, sometimes resulting in characteristic paroxysmal symptoms of headache, profuse sweating, palpitations, Raynaud

phenomenon, tremor, nausea vomiting, and abdominal and chest pain. The triad of palpitations, headache, and profuse sweating are the most sensitive and specific for pheochromocytoma. Clinical findings include hypertension, pallor, postural hypotension, and signs of chronic hypertension such as retinopathy. Paroxysmal symptoms occur in 40 percent of pheochromocytomas and may present as a phenocopy of anxiety symptoms, particularly a panic episode, and are often diagnosed as such. These may occur spontaneously or may be precipitated by exercise, postural change, raised intra-abdominal pressure, or emotional excitement or shock. The severity of the episodes may also vary and anxiety may persist for some time after the attack. Psychosis and cognitive deficits have not been described, although there are several case reports of a secretory pheochromocytoma causing relapse of a previously treated or quiescent psychotic disorder. Diagnosis is by detecting increased urinary secretion of catecholamines or catecholamine metabolites. Location of the underlying tumor and surgical excision with α-adrenergic blockade is effective in resolving most anxiety symptoms if paroxysmal episodes are terminated.

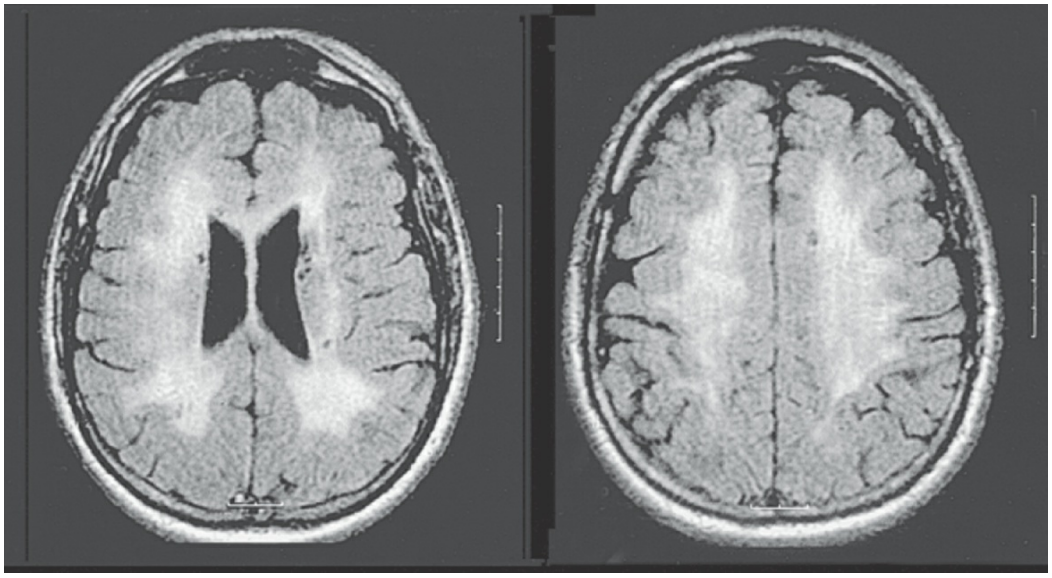

FIGURE 2.14–22. Subcortical dementia due to hyperaldosteronism in a 48-year-old man with untreated hyperaldosteronism. Axial fluid-attenuated inversion recovery images show bilateral extensive periventricular and subcortical white matter hyperintensity with multiple lacunar infarcts in the basal ganglia, consistent with long-standing untreated hypertension.

Neuroendocrine Tumors

While the term "neuroendocrine" is commonly used to refer to the interaction of the endocrine and nervous systems, histologically it refers to a particular type of cell. Neuroendocrine cells are cells that release a hormone or regulatory peptide into the circulation in response to a neural stimulus. The most extensive neuroendocrine system is in the gastrointestinal tract and associated organs, and when these tissues release

excess hormones, a range of neuropsychiatric syndromes can occur.

Carcinoid Syndrome. A carcinoid is a neoplasm of neuroendocrine cells that synthesize and secrete serotonin in the respiratory system and gastro-intestinal tract. The main symptoms are flushing (often severe facial flushing with bronchial tumors), diarrhea, wheezing, and hypotension. Although peripherally secreted serotonin does not cross the blood–brain barrier, a number of psychiatric symptoms have been described. These include depression (in up to half of patients), anxiety, sleep disorders and, rarely, psychosis. A series of 23 patients found that over half experienced personality change consisting of increased irritability and impulsive aggressive thoughts or behavior, some meeting criteria for impulse control disorder. These changes often preceded the other physical symptoms. Depression was far less common and psychotic symptoms were not detected.

Insulinoma. Insulinomas are functioning β-islet cell tumors of the pancreas which result in unregulated insulin secretion at times with abrupt fluctuations. Clinical features include fasting hypoglycemia (relieved by glucose ingestion) and weight gain. Hypoglycemia is characterized by hunger, restlessness, palpitations, flushing, and ataxia but may also include malaise, anxiety, depersonalization, and derealization. A more subacute syndrome characterized by clumsiness, disinhibited or aggressive behavior (and associated amnesia for these episodes) may mimic alcohol intoxication. Global and irreversible cognitive deficits may result if hypoglycemia is long standing. Surgical therapy is the most definitive treatment, but hyperglycemic treatment with the somatostatin analogue octreotide can be helpful.

Glucagonoma. This is a rare pancreatic islet cell which secretes glucagon. The presenting features are of impaired glucose tolerance and diabetes and a severe migratory and necrolytic erythema. Anxiety and agitation may accompany these features.

FUTURE DIRECTIONS

An awareness of the relationship between neurometabolic and neuroendocrine disease and psychiatric illness by psychiatrists and other physicians provides for more diagnostic precision, in addition to allowing for improved treatment of psychiatric comorbidity in these disorders. Psychiatric symptoms can have a profound effect on long-term quality of life, and the recognition of a comorbid mental illness in these disorders can improve health outcomes and quality of life for both patient and caregiver alike. As psychiatrists are best placed to manage major psychiatric disturbance, the involvement of a psychiatrist in the care of these patients

is crucial, whether psychiatric illness is the first or only presentation of illness, or if it develops later in the course of the illness.

Recognizing, understanding, and exploring the links between neurometabolic and neuroendocrine disorders and major psychiatric syndromes can also provide insights into the neurobiological basis of mental illness. For example, the recognition of the elevated rates of schizophrenia-like psychosis in MLD has highlighted the possible role of myelinated structures as a possible anatomical substrate for the functional disconnectivity that has been well described in schizophrenia. Following a group of publications by Hyde and Weinberger on the links between schizophrenia and leukodystrophies, a large body of research in the subsequent 20 years has produced real insights into the role of white matter structures in schizophrenia at the genetic, developmental, and structural levels. In addition, the emerging data on the exceedingly rare movement disorder ChAc suggesting significantly elevated rates of OCD and its unique predilection for neuropathology in the caudate and putamen, has further highlighted the role of disturbed prefrontal–subcortical circuitry in OCD "proper." Further delineation of the neurobiological link between the cellular and metabolic deficit in these syndromes and the psychiatric disturbance they are commonly associated with may yet yield further insights into the neurobiological basis of those disorders that present most commonly in psychiatric practice but that have only yet afforded limited insights into their underlying pathophysiology.

References

Anglin RE, Rosebush PI, Mazurek MF. The neuropsychiatric profile of Addison's disease: Revisiting a forgotten phenomenon. *J Neuropsychiatry Clin Neurosci*. 2006;18:450.

Barkhof F, Verrips A, Wesseling P, van Der Knaap MS, van Engelen BG. Cerebrotendinous xanthomatosis: The spectrum of imaging findings and the correlation with neuropathologic findings. *Radiology*. 2000;217:869.

Bonnot O, Herrera PM, Tordjman S, Walterfang M. Secondary psychosis induced by metabolic disorders. Front Neurosci. 2015;9:177.

Bonnot O, Kluenemann HH, Sedel F, Tordjman S, Cohen D, Walterfang M. Diagnostic and treatment implications of psychosis secondary to treatable metabolic disorders in adults: a systematic review. *Orphanet J Rare Dis*. 2014;9:65.

Borer MS, Bhanot VK. Hyperparathyroidism: Neuropsychiatric manifestations. *Psychosomatics*. 1985;26:597.

de Bie P, Muller P, Wijmenga C, Klomp LW. Molecular pathogenesis of Wilson and Menkes disease: Correlation of mutations with molecular defects and disease phenotypes. *J Med Genet*. 2007;44:673.

Demily C, Sedel F. Psychiatric manifestations of treatable hereditary metabolic disorders in adults. Ann Gen Psychiatry. 2014;13:27.

Gallus GN, Dotti MT, Federico A. Clinical and molecular diagnosis of cerebrotendinous

xanthomatosis with a review of the mutations in the CYP27A1 gene. *Neurol Sci*. 2006;27:143.

Gargus J. Ion channel functional candidate genes in multigenic neuropsychiatric disorders. *Biol Psychiatry*. 2006;60:177.

Hayflick SJ, Westaway SK, Levinson B, Zhou B, Johnson MA. Genetic, clinical, and radiographic delineation of Hallervorden-Spatz syndrome. *N Engl J Med*. 2003;348:33.

Heinrich TW, Grahm G. Hypothyroidism presenting as psychosis: Myxedema madness revisited. *Prim Care Companion J Clin Psychiatry*. 2003;5:260.

Knerr I, Gibson KM, Jakobs C, Pearl PL. Neuropsychiatric morbidity in adolescent and adult succinic semialdehyde dehydrogenase deficiency. *CNS Spectr*. 2008;13:598–605.

Lynch S, Merson S, Beshyah SA, Skinner E, Sharp P. Psychiatric morbidity in adults with hypopituitarism. *J R Soc Med*. 1994;87:445.

Medvei VC, Cattell WR. Mental symptoms presenting in phaeochromocytoma: A case report and review. *J R Soc Med*. 1988;81:550.

Miyajima H: Aceruloplasminemia, an iron metabolic disorder. *Neuropathology*. 2003;23:345.

Mocellin R, Walterfang M, Velakoulis D. Hashimoto's encephalopathy: Epidemiology, pathogenesis and management. *CNS Drugs*. 2007;21:799.

Nia S. Psychiatric signs and symptoms in treatable inborn errors of metabolism. *J Neurol*. 2014 261:S559–S568.

Rajasekaran A, Venkatasubramanian G, Berk M, Debnath M. Mitochondrial dysfunction in schizophrenia: pathways, mechanisms and implications. Neurosci Biobehav Rev. 2015;48:10–21.

Russo S, Boon JC, Kema IP, Willemse PH, den Boer JA. Patients with carcinoid syndrome exhibit symptoms of aggressive impulse dysregulation. *Psychosom Med*. 2004;66:422.

Sedel F, Baumann N, Turpin J, Lyon-Caen O, Saudubray J. Psychiatric manifestations revealing inborn errors of metabolism in adolescents and adults. *J Inherit Metab Dis*. 2007;30:631.

Sheth S, Brittenham GM. Genetic disorders affecting proteins of iron metabolism: Clinical implications. *Annu Rev Med*. 2000;51:443.

Sonino N, Fallo F, Fava GA. Psychological aspects of primary aldosteronism. *Psychother Psychosom*. 2006;75:327.

Swift RG, Perkins DO, Chase CL, Sadler DB, Swift M. Psychiatric disorders in 36 families with Wolfram syndrome. *Am J Psychiatry*. 1991;148:775.

Toker L, Agam G. Mitochondrial dysfunction in psychiatric morbidity: current evidence and therapeutic prospects. Neuropsychiatr Dis Treat. 2015;11:2441–2447.

Velasco PJ, Manshadi M, Breen K, Lippmann S. Psychiatric aspects of parathyroid disease. *Psychosomatics*. 1999;40:486.

Vincent A, Lang B, Kleopa KA. Autoimmune channelopathies and related neurological disorders. *Neuron*. 2006;52:123.

Walterfang M, Fietz M, Fahey M, Sullivan D, Leane P. The neuropsychiatry of Niemann-Pick type C disease in adulthood. *J Neuropsychiatry Clin Neurosci*. 2006;18:158.

Walterfang M, Velakoulis D. Storage disorders and psychosis. In: Sachdev P, Keshavan M, eds. *Secondary Schizophrenia*. New York: Cambridge University Press; 2008.

Walterfang M, Bonnot O, Mocellin R, Velakoulis D. The neuropsychiatry of inborn errors of metabolism. *J Inher Met Dis*. 2013;36:687–702.

Walterfang M, Yucel M, Walker R, Evans A, Bader B. Adolescent obsessive compulsive disorder heralding chorea-acanthocytosis. *Mov Disord*. 2008;23(3):422–425.

Yudofsky S, Hales R. *The American Psychiatric Publishing Textbook of Neuropsychiatry and Behavioral Neurosciences*. Arlington, VA: American Psychiatric Publishing; 2007.

3

Contributions of the Psychological Sciences

▲ 3.1 Contributions of the Psychological Sciences: Sensation, Perception, and Cognition

LOUIS J. COZOLINO, PH.D., AND DANIEL J. SIEGEL, M.D.

The brain is a social organ of adaptation, built and maintained through the interaction of social, psychological, and biological forces. Because of the vast complexity of this synergistic process, the study of sensation, perception, and cognition is necessarily broad and far-reaching. The rapid increase in knowledge of neuroscience and the development of new hypothetical models through which one can understand the brain's functioning make this an exciting time in the cognitive sciences.

Mental processes such as thinking, feeling, and remembering are not synonymous with brain activity, but they may be dependent upon them. At a minimum, the subjective, inner felt experience of these fundamental mental processes, the feeling of being alive and awake, is simply not the same as neural firing, even if subjective experience arises from that brain activity. Further, the mind may be more than just "an output" of what the brain does, with the possibility that it is an emergent, self-organizing process that is both fully embodied, extending beyond the skull and into the whole body, and also fully embedded in one's relationships, extending beyond the skin of an isolated individual. In this way, the element of the system of mind may be energy and patterns of energy with symbolic value called "information." Energy and information flow happens both within humans and between humans. Seen this way, mind is an emergent aspect of a complex system may arise from energy and information flow, and as a self-organizing emergent process it may regulate that flow. This regulation is reflected in mental processes from sensation and perception, to cognition, to the initiation of behavior.

Nevertheless, science is clear: Mental processes are profoundly shaped by neural activity, even if they are not limited to what goes on in the head. For this reason, understanding neural processing is an important part of comprehending information processing, one of many important aspects of

mind.

The terms s*ensation, perception,* and *cognition* are used to describe the three broadening tiers of human information processing. Think of s*ensation* as the immediate result of the stimulation of sensory neurons and *perception* as involving the organization and conscious awareness of these sensations by higher regions of the brain. *Cognition* refers to the set of interwoven processes, such as memory, language, and problem solving, that one brings to bear to generate structures and strategies to apply to perceptions that involve extensively distributed regions of the brain. Although distinguishing among sensation, perception, and cognition has a long academic history, the ability to separate them anatomically and functionally in the brain grows increasingly difficult as more is learned about their interdependence in the dynamic functioning nervous system.

COGNITIVE SCIENCE

The fields relevant to this overview are part of the interdisciplinary studies of cognitive science, which include cognitive psychology, developmental psychology, psycholinguistics, computational science, and the emerging field of interpersonal neurobiology. Biological, psychodynamic, and social psychiatry find a common home within cognitive science whereby the usual divisions of nature versus nurture and of biology versus psychology disappear on examination of the development of the brain and the origins of mental processes.

In recent years, discoveries in the neurosciences have revealed a wide range of findings relevant to psychiatry. One such discovery showed that the brain's structure and function are a result of the transaction among genetic, physiological, and experiential influences. In particular, brain development requires specific forms of experience to foster the growth of neural circuits involved in a wide array of mental processes, including attention, memory, emotion, attachment, and self-reflection. Whereas genes function as a template of information and as a mediator of transcription of the proteins that determine neural structure, experience directly shapes the selection and timing of how gene activity is translated into the structure of the brain; it is now known that experiences influence the epigenetic regulators that modulate which, when, and how genes are expressed in response to experience and thus shape the unfolding of genetically programmed development of the central nervous system (CNS).

The human brain, especially the cerebral cortex, is immature at birth. This immaturity requires that the child's brain use the caregiver's brain to help it to grow and organize its functioning. Findings from developmental neuroscience point to the centrality of interpersonal relationships in the development of the brain. The cooperative communication of infant–caregiver attachments is thought to shape neural infrastructure, not only for emotional development, but also for abstract reasoning and cognitive

abilities. The patterns of interaction between child and caregiver have a direct impact on the development of the child's brain and the functioning of the mind—the emergence of energy and information flow within and between. Consequently, cognitive processes and the internal flow of energy and information are an expression of the genetic, physiological, social, and experiential factors that shape the development and maintenance of mental function.

Hot and Cold Cognition

Traditionally, cognition was studied by experimental psychologists in university laboratories, whereas emotion was explored clinically by psychoanalysts in consulting rooms. Striving to avoid the subjective and imprecise nature of emotions, cognitive psychologists devised flowcharts and algorithms similar to those used to describe computer programs. Input was calibrated, output was measured, and theories were generated concerning what might be happening within the "black box" of the brain. The increasing complexity of these models, however, did not improve their explanatory power, and questions concerning motivation and emotion invariably arose.

As knowledge of brain functioning increased, it has become more and more evident that neural networks involved in perception and cognition are inextricably interwoven with other networks responsible for processing somatic states, survival value, emotion, meaning, and motivation. The myth of *cold cognition,* or cognitive processes devoid of affective and somatic influence, is gradually fading. Flowcharts depicting linear input–output processes are being replaced with more sophisticated models reflecting the reality of the complex neural systems being discovered. When this change is added to the notion that mental life is also shaped by the context, especially the social context, in which an individual lives, then it can be seen that the mind emerges from factors within the brain, the body, and the larger interactive environment. This is the "mind in context" that needs to be considered to understand all mental processes.

Whereas some cognition is "hot," such as recalling traumatic memories or spotting someone that is sexually attractive, some cognition may be cool, such as adding rows of figures or stacking dishes. These same "cool" tasks, however, if performed before an Internal Revenue Service auditor or in front of a strict teacher, may become warm or even hot. The fundamental principle, therefore, is that sensation, perception, and cognition all occur within the context of feedforward and feedback neural influences interwoven with and guided by complex contextual and emotional determinants.

Emotion

What is an emotion? The answer to this question is as complex as the mind itself. Although the lack of clear definitions and good animal models have hindered empirical research, it is clear that emotions play a central role in many cognitive processes. In fact, a broad view suggests that an emotion is a profoundly integrative process. Emotions connect body to brain, bring continuity to states of mind across time, and link one person to another within the "emotional connections" that create the fabric of human social lives.

One view considers emotion as a primary value system of the brain, engaging attentional resources, appraising the significance of an event, and allowing activations to be selectively elaborated and reinforced as they are encoded and stored, such that emotionally charged experiences may be more readily recalled later on than uneventful ones. According to this view, the most fundamental aspect of emotion is the arousal–appraisal system in which the brain responds to a given stimulus with the signal of "this is important—take note and pay attention now!" Emotion thus gives value to a representation by arousing attentional mechanisms and focusing a spotlight of attention on the stimulus. The second stage would then appraise the meaning of such emotional arousal by assessing its hedonic tone: "Is this good or bad? Should this be approached or avoided?" Emotion thus directs the flow of energy—the activations within specific circuits of the brain—as the arousal–appraisal system focuses cognitive processes on elements of the internal and external environments. A third level of emotional processing is the elaboration of this appraisal into a more specific form called a *categorical emotion,* which includes joy, interest, surprise, fear, anger, sadness, or shame. These categorical emotions have distinct psychophysiological manifestations and are found across cultures.

An additional view of emotion examines the way in which changes in the body's state are represented in the brain in the form of what Antonio Damasio called a *somatic marker*. According to this perspective, the bodily responses to a situation or a choice that needs to be made let the brain know how the individual feels about an experience. Such a somatic marker can then be used as a gut reaction to an experience, providing additional input into cognitive processing.

A part of the brain called the *orbitofrontal cortex* has been implicated as the site of somatic marker processing, or what is called *intuition*. Allan Schore noted the importance of early experiences with caregivers in the maturation of this region and its role during early development in coordinating self-regulatory functions with basic emotional reactions and social functioning. A closely related region, the anterior insular cortex, mediates the interoception that allows internal somatic states to be represented in the higher cortical areas and the mind to experience awareness of the interior of the body. Being open to monitoring the body's

signals permits the modulation of the somatic elements of emotions, and allows for the development of emotional regulation. Disorders in self-organization and social functioning may be better understood by examining the central role of emotion and, perhaps, the orbitofrontal and insular cortex and related regions in the development and maintenance of dysfunctional mental states.

Studies also suggest that the orbital and medial prefrontal regions are responsible for subjective experience and self-awareness, enabling the mind to reflect on the self in the past, present, and the potential future and even to consider the internal states of others. Inborn and experiential factors may play important roles in allowing this region to develop the capacity to integrate a wide range of important functions of the mind, including the appraisal of meaning, emotional regulation, social cognition, and autobiographical consciousness.

NEURAL NETWORK GROWTH AND INTEGRATION

The growth and selective connectivity of neurons is the basic mechanism of all learning and adaptation. Learning can be reflected in neural changes in a number of ways: (1) the growth of new neurons, (2) the expansion of existing neurons, (3) the changes in the connectivity among existing neurons, and (4) alterations in epigenetic regulation of specific regions of the genome. All of these changes are expressions of plasticity, or the ability of the nervous system to change.

There is now sufficient evidence for the fact that neurons demonstrate growth and changes in reaction to new experiences and learning. Existing neurons grow through the expansion and branching of the dendrites that they project to other neurons. Neurons interconnect to form neural networks, which, in turn, integrate with one another to perform increasingly complex tasks. For example, networks that participate in language, emotion, and memory interact and integrate, allowing for an emotionally meaningful story to be recalled and told with the proper affect, correct details, and appropriate words.

Association areas within the brain serve the role of bridging, coordinating, and directing the multiple neural circuits to which they are connected in what is now called "*the connectome*." Although the mechanisms of this integration are not yet known, they are likely to include some combination of (1) biochemical processing within neurons, (2) synaptic connections among neurons, (3) relationships among local neuronal circuits, and (4) interactions among functional brain systems. Changes in the synchrony of activation of multiple neural networks may also play a role in the coordination of their activity.

If everything humans experience internally is represented within neural networks, then psychopathology of all kinds, from the mildest neurotic symptoms to the most severe psychosis, must be represented within and

among neural networks. The social environment, too, plays a crucial role in well-being and can itself support healthy or unhealthy states. Healthy functioning requires the potential for proper development and functioning of neural networks that are responsible for organizing conscious awareness, behavior, emotion, and sensation. Severe trauma early in life, in situations of abuse or neglect, can alter the healthy development of the brain's integrative potential. Innate neurological conditions, unrelated to an experiential cause, may also result in impediments to the connectome functioning as an integrative whole. With integration, flexibility and harmony arise and health may ensue. With impaired integration, when the differentiation and/or linkage of areas are impeded, then chaos and/or rigidity may result.

Symptoms of the various syndromes seen through this lens may be envisioned as the chaos and rigidity of impaired integration within the brain. Psychopathology, then, can be conceptualized as an expression of suboptimal integration and coordination among neural networks—whatever the etiology. Given that experience shapes the brain's development as do nonexperiential factors such as genetics, infection, inflammation, and random events, the finding that individuals with various forms of psychiatric disorders have suboptimal integration in the brain is supportive of this proposal. Patterns of dysregulation of brain activation in specific disorders also support the theory of a brain-based explanation for the symptoms of psychopathology. In general, psychological integration suggests that the cognitive functions of the executive brain have a high degree of access to information across networks of sensation, behavior, and emotion.

Integration requires both differentiation of elements of a system and their linkage. And integration appears to be essential for optimal regulation. Anything that impedes either differentiation or linkage can disrupt integration in the moment as a function, or over the long run as a structural impediment, that may lead to impaired self-regulation. Functional dissociation among these processes can occur when biochemical changes caused by high levels of stress inhibit or disrupt the brain's integrative abilities. With brain growth influenced by a range of processes, structural impairments to integration may ensue. Physical trauma, disease processes, or genetic predispositions that disrupt the development and functioning of integrative neural networks can all result in neural dysregulation and psychiatric symptomatology.

In applying this integration and brain-based model to treatment, psychotherapy, psychopharmacology, and psychosurgery can be viewed as ways of creating or restoring integration and the resultant coordination among various neural networks. For example, research has demonstrated that successful psychotherapy correlates with changes in activation in areas of the brain hypothesized to be involved in psychiatric disorders, such as

obsessive-compulsive disorder (OCD) and depression. The return to normal levels of activation results in reestablishing positive reciprocal control among relevant neural structures and networks and a reduction of symptomology.

MIND AND BRAIN

What is this activity of the brain, and how does it give rise to such mental processes as perception and cognition? How do the human experiences of perception, thought, emotion, attention, self-reflection, and memory emerge from neural processes?

A generally accepted view of the mind is that it emanates from a portion of the activity of the brain. This perspective, however, is only part of the connection between mind and brain. The mind can be defined as a process that regulates the flow of energy and information. The human mind is both embodied and relational, meaning that the flow of energy and information occurs both within the neural firing patterns in the body as a whole and between people through their interactions. Envisioning human experience as an interaction of mind, brain, and relationships allows mental health practitioners to free themselves from the overly simplistic view that the mind is "just the activity of the brain." Conceptualizing sensation, perception, and cognition as arising from a synergy of mind, brain, and relationships provides a perspective that is broad enough to integrate findings across various fields of study.

One clinical application of this broader perspective comes from an ancient practice called "mindful awareness" that focuses attention on the present moment. Research has demonstrated the clinical utility of mindfulness-based approaches in the treatment of OCD, generalized anxiety, and borderline personality disorder and in the prevention of relapse in chronic depression and drug addiction. Controlled studies suggest that such mindful awareness–based clinical interventions improve bodily function, interpersonal relationships, and mental well-being. One way of seeing the power of such mental training is that it improves functioning across each of these aspects of life by enhancing the capacity of the mind to regulate energy and information in a more integrative way, both within and between.

ENERGY AND INFORMATION

The brain is composed of tens of billions neurons and perhaps ten times that number of supportive cells called glia. When regions such as the cortex (which may have around 20 billion neurons) and the cerebellum (which may have 100 billion) are combined with other areas, an estimate of over 120 billion neurons is derived with the distribution of glia varying in different areas. An average neuron is connected to approximately 10,000 other neurons at synaptic junctions. With hundreds of trillions of

connections within and among thousands of web-like neural networks, there are countless combinations of possible activation profiles. The term *neural net profile* is used to describe a certain pattern of activation of the complex layers of neural circuits, which is the fundamental way in which mental processes are created. These activations can lead to further neural processes in a cascade of dynamic interactions that produce a range of internal events and external behaviors.

The essential components of the mind come directly from how these neural events create the flow of energy and information. How one subjectively feels the texture of this flow from the shifts of energy and information within the matter of the brain is something no individual or theory has yet to be able to fully explain. It may simply be that an irreducible aspect of neural flow is subjective experience in ways that are not yet comprehend. Nevertheless, this situation does not impede the ability to explore how the mind, as a possible self-organizing emergent process of that flow, may actually function as a regulatory process.

The mind is a processor of patterns in the flow of energy and information within the brain (Table 3.1–1). Activations of individual neurons as well as groups, circuits, or networks of neurons all involve the flow of energy through this complex system. On one level, the system includes the brain and its body. On another level, this system may involve communications with one another. Within the brain, this energy reflects the flow of ions across membranes, the consumption of oxygen and nutrients by neural cells, and the active transport of molecules into and out of nervous tissue.

Table 3.1–1.
Basic Ideas of the Mind

The mind is a processor and regulator of energy and information.
Energy is contained within and among the activations of neural circuits as well as in the communication between people.
Information is contained within and among the patterns of activation, termed a *neural net profile* or *mental representation;* and information is embedded in energy patterns shared as communication between people.
These representations are energy patterns that serve as symbols that cause further effects in the mind, leading to the processing of information both within the nervous system and in relationships.

Information is created within the brain by a process of representation. The essential feature of information processing in the brain is that the patterns of activation of neural circuits (the neural net profile) contain information. These mental representations, in turn, produce further neural events. The location and pattern of neural activations determine the nature of what the neural net profile represents. For example, activity in the optic nerve in response to light leads to a cascade of neural responses within the

visual cortex, generating a visual sensation. Future activation of those layers in the visual cortex in that general pattern is the recollection of the visual image. Pattern and localization determine the kind of representation and the information that it specifically contains. For example, when an individual sees the Eiffel Tower, the visual system responds with the activation of a neural net profile. When the Eiffel Tower is recalled at a later time, the visual cortex activates a similar neural net pattern, and the Eiffel Tower is visualized. The activation of a particular pattern of neural firing thus contains representations of information about something, in this case, the Eiffel Tower. Examples of representations include perceptual, emotional sensory, and linguistic forms, as well as more abstract concepts and categories.

INFORMATION PROCESSING

Several elements of the brain's function as an information processor can be described (Fig. 3.1–1): At the most basic level (Fig. 3.1–1A), energy leads to neural responses. This energy can be in the form of light stimulating the rods and cones of the retina or sound waves vibrating the tympanic membrane. It may also take an internal form in which the flow of energy within neural activations produces subsequent neural responses.

A second level of conceptualizing information processing (Fig. 3.1–1B) lies in the idea that an input (internal or external) leads to a representational response (a neural net profile of activation), which, in turn, produces a downstream effect or output. This output can be internal, such as the generation of other representations, or external, such as in the form of observable behavior. Within cognitive psychology, these information-processing events can be analyzed in terms of contrasting, comparing, generalizing, chunking, clustering, differentiating, and extracting processes, all of which lead to increasingly complex mental representations.

A third level of understanding information processing in the mind (Fig. 3.1–1C) is the conceptualization of sensation, perception, attention, and memory. According to this view, external energy is sensed by the peripheral nervous system and is registered as a sensation within the brain. The selective processing of aspects of these sensations, called filtering, leads to the production of perception. These perceptions are subject to further filtering during which only a select few are placed within working memory; this is sometimes called the "chalkboard of the mind." It is within working memory that representations can be consciously manipulated, contrasted, clustered, and reassembled. Thus, consciousness may be intimately related to this aspect of mental functioning.

Sensation refers to the initial stages of the basic information-processing model (Fig. 3.1–1). In traditional experimental paradigms, sensory memory is conceptualized as lasting for approximately 0.25 second. Items in

sensory memory are then filtered into working, or short-term memory, where they last for approximately 0.5 minute. When humans consciously attempt to learn new information, working memory is able to handle approximately seven items, unless further processing creates links to other items within longer-term memory. *Rehearsal* allows these representations to remain in working memory for longer periods of time. Cognitive processes that can group bits of information into large chunks (chunking) can increase the capacity of working memory by making each unit more information rich. Representations are then processed and placed within long-term memory, where they can be retrieved for future use.

ATTENTION

Attention is the process that controls the focus and flow of information processing. Three components of attention (selectivity, capacity, and sustained concentration) have traditionally been used to describe cognitive deficits seen in psychiatric disorders, such as schizophrenia and attention-deficit/hyperactivity disorder (ADHD). All aspects of attention in nonclinical and patient populations are influenced by the emotional or motivational value of the stimulus.

Early conceptualizations of attention were based on Donald Broadbent's idea of a filter that selects a limited amount of incoming stimuli to be further processed. Limited capacity of attention was thus seen as being attributable to the inability to process the overwhelming amount of incoming stimuli. An attention "bottleneck" was described as occurring early in the sensory process (automatic) or late in the perceptual processing stage (identification and classification).

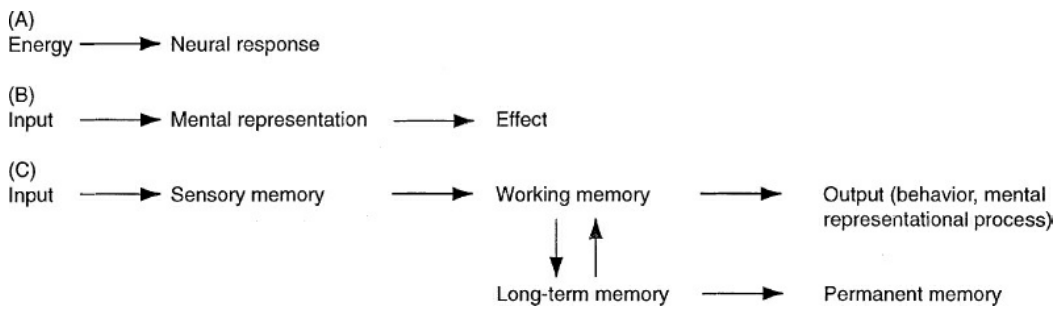

FIGURE 3.1–1. Information-processing models.

Selective Attention

Attention focuses a metaphorical spotlight on external stimuli or internal mental representations. In Broadbent's conceptualization, selectivity has three dimensions: (1) filtering—focusing on specific attributes (e.g., large squares vs. small squares); (2) categorizing—recognizing information based on stimulus class (e.g., attending to letters in whatever script they are written); and (3) pigeonholing—reducing perceptual information needed to

place a stimulus into a specified category (e.g., using only long hair to classify individuals as female). Each of these aspects of attention acts on incoming stimuli to make a determination of fit for the sought-after characteristic.

Attention Capacity

The concept of processing capacity involves the idea that a given task makes a demand on a limited pool of resources: A task with a high-processing load draws more resources from the finite pool than does a task with a low-processing load and thus will inhibit the accessibility of resources for other simultaneous functions drawing from the same pool. Focal attention requires cognitive effort, and thus has a high-processing load demand. Cognitive models describing several resource pools suggest an executive process that distributes resources to various cognitive functions.

Optimal performance is attained when there are moderate levels of arousal because this allows for the establishment of task goals, feedback, and appropriate resource allocation. Low levels of arousal impair those processes and lead to inadequate resource allocation, whereas high levels of arousal may be detrimental to the performance because of poor discrimination of stimuli and diminished efficiency of allocation, resulting in poor attention functioning.

Sustained Attention

The ability to sustain attention is called *vigilance* and can be tested with task demands for alertness and concentration over a period of a few minutes to an hour. The tests usually involve detection requirements for target stimuli that occur infrequently at random intervals. An example of such a test is the Continuous Performance Test. Defining aspects of the tests are derived from signal detection theory and include the factors of *sensitivity* and *response criterion*. Sensitivity is the distinguishing of target stimuli from nontarget stimuli; the response criterion is the amount of perceptual evidence required to support the decision regarding a target item versus a nontarget item.

SENSATION AND PERCEPTION

Forms of representations, including sensory and perceptual representations, derive from input from the external world via the peripheral sensory nervous system. The initial stage of encoding a visual representation is called an iconic image and is held within sensory memory for a brief period. Features of the initial stimulus, such as its size, direction, and color, are held within this sensory representation. This is termed bottom-up processing and is as direct as the nervous allows. As the initial

sensory activations are processed, they become influenced by higher-order, prior-experience shaped influences called top-down processes and are organized as perceptual representations. In other words, learning from the past embedded in synaptic neural patterns of memory shapes perceptual streams of representations in the moment. This can be seen as an important distinction between sensation and perception.

Attentional processes, at the level of sensory memory, act on the initial image with higher cognitive functions, such as classifications, memory, and chunking. In their essence, these top-down processes compare, contrast, and transform the initial representation from sensation to create new perceptual images within working memory. These processes are often automatic, rapid, and outside of awareness.

Perception is thus created by the top-down transformations of sensory images but does not necessarily involve the experience of consciousness. This has important clinical implications in that patients may be influenced by events and stimuli that they cannot consciously recall. Take, for example, the experience of a young girl with anorexia or body dysmorphic disorder. When she looks at herself in the mirror, she becomes consciously aware of perception of her reflection in approximately half a second (500 ms). In contrast, the sensation of her visual reflection has already reached her eyes almost instantaneously and has resulted in nonconscious neural activation in only 50 ms. As neural networks are being instantiated within her visual cortex, activation in other brain regions, such as the amygdala, is also occurring, streaming top-down emotional and appraisal information back to the visual cortex.

In the 450 ms between initial sensory activation and conscious perceptual awareness, the visual experience of her reflection is constructed through a combination of bottom-up sensations, and top-down emotion, learning history, and other visual memories to which it is compared and contrasted. What she focuses on, how she feels about it, and what it means to her—in essence, how she is seeing what she is seeing—is a blend of bottom-up and top-down constructed in automatic, nonconscious processes. At 500 ms, she experiences her brain's construction of her body, this perceptual image, as an external reality. In other words, a great deal happens on the way from sensation to perception.

MEMORY SYSTEMS

The neural networks of the brain are capable of responding to experience through the activation of neural net profiles—particular patterns of distributed neural activation. Donald Hebb described a basic principle of memory, paraphrased by Carla Shatz, that has been repeatedly supported by research: "Neurons that fire together, wire together." Neurons that are activated in a particular pattern at one time tend to fire together in a similar pattern in the future—this long-term potentiation (LTP) is the

essence of memory.

The brain has various circuits responsible for different systems of memory (Table 3.1–2): The form of memory most commonly thought of as *memory* is termed *explicit* or *declarative memory*. This form involves the conscious sensation something coming from the past, a feeling of remembering, focal attention (attention with awareness), recall of autobiographical or factual knowledge, and hippocampal activation. In these circuits, items focally attended to are placed in working memory, processed further, and then placed in long-term memory. After a period of weeks or months, items are thought to undergo a process called *cortical consolidation* that places them in permanent memory, where their retrieval no longer requires the hippocampus.

Explicit autobiographical memory becomes reliably available after the first years of life once the hippocampus and cerebral cortex have sufficiently matured. Before this, a form of memory that does not require conscious awareness and focal attention for encoding, called *implicit* or *nondeclarative memory,* is already in place and remains active throughout the life span. This form of memory involves a wide range of neural regions including behavioral, emotional, and perceptual systems. When these circuits are activated in retrieval of pure implicit memory, they do not include the conscious sensation of something being recalled. For example, when riding a bicycle, a person may not recall having learned to ride and may not even feel that anything is being recalled. Similarly, a person with a fear of dogs may be unable to explicitly recall any event that may explain such an emotional response.

Table 3.1–2.
Memory Systems

Implicit
A behavioral, emotional, and perceptual form of memory devoid of the subjective internal sense of something originating in the past. Can include schema or mental models that are summations of representations from numerous experiences.
Also known as *early, procedural,* or *nondeclarative memory.* Cannot be directly expressed in words.
Present from birth. Does not involve the hippocampus or require focal, conscious attention. Probably involves various circuits, including those of the basal ganglia, limbic system (amygdala, anterior cingulate, and orbitofrontal cortex) and perceptual cortices.
Explicit
A form of memory requiring conscious awareness for encoding and involving the subjective sense of recollection and, if autobiographical, of self and past.
Also known as *late, episodic* or *semantic,* or *declarative memory.* Can be directly expressed in words.
The autobiographical component of explicit memory does not fully develop until past the first 2 years of life, as the hippocampus and orbitofrontal cortex, on which it depends, are maturing.

The existence of intact implicit recollection in the absence of explicit memory is found in various conditions, including being under surgical

anesthesia or having neurological conditions such as Korsakoff syndrome and bilateral hippocampal strokes. Dissociation of memory systems may also occur in response to trauma. For example, patients with posttraumatic stress disorder (PTSD) may not explicitly recall a traumatic event, yet they may implicitly avoid stimuli similar to the initial trauma. They may also evidence implicit symptoms such as a startle response, anxiety, and intrusive perceptual images for that event in the absence of an explicit autobiographical memory. These latter symptoms are thought to reflect the intrusion into conscious awareness of implicit memory in the absence of the subjective sensation that something is being recalled.

CONSCIOUSNESS

The vast majority of mental processes are outside of awareness as energy and information flows without consciousness. A simple way of defining conscious awareness is the subjective sense of knowing. When someone is aware, there is a knowing. And the object that one is aware of is the known, such as visual input, a memory, or a thought. In general, many authors' views converge on the idea that there exist two fundamental forms of consciousness, enabling a knowing of two forms of what can be known in consciousness: A *here-and-now* and a *past–present–future* process of awareness. These two forms of consciousness are likely mediated via the integration of different neural circuits in the brain and the experience of the known is distinct, one in the moment, one focusing on events across time.

Two hypotheses focus on the way in which representational processes are linked or are bound together during the flow of informational transformations within the mind: One hypothesis suggests that a 40-cycle-per-second sweeping process extends from the thalamus to the neocortex. This sweep may serve to bind representational processes together in the internal experience of consciousness. Processes that are active at the time of the sweep then become linked within consciousness. A second view implicates the lateral prefrontal cortex and its role in working memory—working memory serves as the chalkboard of the mind, and representational processes that become integrated to a certain degree and linked to the activity in this region fall under the attentional spotlight of conscious awareness called focal attention.

Disturbances of Consciousness

Consciousness provides a sense of continuity—yet many psychiatric patients experience a profound sense of discontinuity and confusion that may be related to a dysfunction in the sense-making, continuity-creating process of consciousness. Some psychiatric symptoms, including derealization and depersonalization, intrusive memories and flashback

phenomena (as in PTSD), and hallucinations (as in psychotic states), may be conceptualized in terms of dysfunctions in conscious functioning.

Misidentification syndromes are other manifestations of disturbances of subjective, conscious experience. In prosopagnosia, for example, patients are unable to consciously access memories regarding persons familiar to them. Patients suffering from Capgras syndrome, on the other hand, are able to recognize a familiar person's face but cannot experience the feelings of recognition and emotional arousal associated with that person. This lack of emotional recognition then results in their conviction that their loved ones have been replaced by imposters. "Being certain," as in the process of recognition, is one aspect of consciousness that is a cognitive process. Theoretically, then, the pathological uncertainty of patients with OCD can be viewed as a disturbance in that aspect of conscious functioning.

Another example of disturbance of conscious functioning is found in cortically blind patients, who state that they cannot see visual stimuli but respond behaviorally as if they were fully sighted. Cortically blind patients describe being unaware of visual perception yet are able to make eye and hand movements that reflect the processing of information about stimulus location, shape, orientation, and direction of motion. In information-processing terms, behavioral tests reveal that these blind-sighted patients do sense and perceive visual stimuli but do not have conscious awareness of this perceptual process. Cortical blindness, then, is an example of the dissociation of the usually associated processes of perception and consciousness, or awareness of phenomena.

Mental Models and Schemata

Studies of perception and memory support the view that the mind has organizational structures that influence the interpretation of sensory data, shape the encoding of information into long-term memory, bias the retrieval of items stored in memory, and help to determine the behavioral responses. These organizing cognitive functions are called *mental models* or *schemata*.

Mental models are highly organized, implicit, top-down processes derived from past experiences that guide the interpretation of present stimuli and influence the direction of behavior. The adaptational value of a given mental model depends on an accurate reading of the survival demands of the situation. The downside of these models is seen when their unconscious and automatic activation occurs in situations in which they are inappropriate, such as in PTSD, or when they no longer sufficiently meet the demands of a given situation.

Aaron Beck's theory of depression is based on the idea that mental models or schemata can guide depressive thinking as they trigger and enhance depressed moods. John Bowlby and Mary Ainsworth used the concept of internal working models to describe the development of early

forms of schemata for relationships: Difficulties in intimate relationships and related behavioral dysregulation can be seen as derivatives of models of inadequate early attachment and the presence of multiple, conflictual models. The *inner objects* of psychodynamic theory are also examples of such mental models. Some psychiatric signs and symptoms can be seen as derivatives of conflicted schemata and situations. Classic descriptions of interpersonal patterns in some patients with personality disorders, such as idealization and devaluation, can be seen as maladaptive schema functions.

Thought and Language

Although there is no universally accepted definition of *thought,* thinking most generally involves the mental representation of some aspect of the world or of the self and the manipulation of those representations. Thinking depends on explicit and implicit memory of prior experiences and is influenced by a person's emotional state, mental models, and other unconscious determinants. The basic components of thinking include categorization, judgment, decision making, and problem solving. Cognitive processes, such as thoughts, are often directly known only through their experience in consciousness and translation into language. As a result, there is always a question of whether thought and language can be fully separated.

Cognitive science has traditionally viewed language as a dominant influence on subjective experience. Language shapes the ways in which the world is perceived, the manner in which desires are communicated and satiated, and the way in which society responds. For example, the language used to describe and define psychiatric disorders can shape how mental health professionals perceive what a patient may be experiencing as top-down filters shape and constrain the sensory input from the patient's presentation and mold a clinician's assessment and treatment planning.

Rational thought contributes to the ability to judge the probability of uncertain events and to choose among various options. These processes contribute to problem solving in which data are assessed, classified, transformed, and compared on the basis of logical rules to produce a choice that solves a problem; failures in these steps can result in limitations and distortions in normal thought processes.

Modes of Processing and Laterality

The mind is capable of distinct modes of processing mental representations: A serial mode uses sequential processing in a linear fashion but is thought to be slow and energy consuming because only a few items can be processed at a time; focal, conscious attention is believed to occur in this serial fashion. A parallel mode involves the simultaneous manipulation of large numbers of representations in a nonlinear fashion

and occurs in a rapid, low–energy-consuming process; pattern recognition is one example of parallel processing that can deal with a wide array of stimuli at the same time.

Another distinction in contrasting modes of processing has been identified in the type of mental processes primarily attributed to the right and left cerebral hemispheres. Studies in which patients whose corpus callosum have been surgically severed, who have had unilateral neurological lesions, or who have undergone brain-imaging protocols have revealed a remarkable consistency in trends of left- versus right-hemisphere functioning. Some general principles from this array of studies suggest that whereas many processes involve both hemispheres, there are also distinct patterns primarily originating from one sides of the brain or the other.

The forthcoming generalizations relate to right-handed individuals and to most left-handed people: In the right hemisphere, there are fast-acting, parallel, holistic processes, including perception and communication of nonverbal signals. The right side specializes in representations with a wide spectrum of values, such as images and sensations, and in the nonverbal meaning of words, sometimes referred to as analogic representations. The right hemisphere is thought to work as a pattern recognition center, capable of assessing the Gestalt context of a scene and providing a synthetic interpretation. In the left hemisphere, there are primarily more slowly acting, linear, time-dependent, serial processes. Left-hemispheric processes manipulate the verbal meaning of words in a logical analytical mode of processing, focusing on the digital or clearly demarcated meaning of terms categorically organized, such as up or down, right or wrong, left or right. A generalization from a number of studies is that the right hemisphere tends to note the patterns in the world and creates contextual meaning, whereas the left hemisphere notes the details of what it perceives and attempts to organized and make sense of the world, creating a cause–effect explanation of experience.

The processing of emotion appears to be biased toward the right hemisphere. Facial recognition of the affective expression of others also appears to be a specialty of the right hemisphere. Notably, the right hemisphere appears to have a more integrated representation of the body's status, or kinesthesia, information that may be essential for individuals to know how they feel. The hemispheres also seem to demonstrate biases with regard to emotional tone, with the left having a positive affect bias and approach state whereas the right has a negative bias and a withdrawal state. An integration and balance of hemispheric influences appears to be necessary for adequate affect regulation and mental health.

Jerome Bruner described the distinction between an earlier mode of thought, which he referred to as *narrative cognition,* versus a later mode, called scientific, logical, or paradigmatic cognition. Narrative thinking is a

context-dependent form of processing that incorporates the internal experiences of the teller and the perceived expectations of the listener in the production of a story. Stories also involve the subjective experiences of the characters involved in the unfolding sequence of events. Children develop narrative thinking by 2 years of age, and the co-construction of stories between parent and child is a primary mode of communication across cultures. Drawing on context, emotions, and linear explanations, such narratives help make sense experiences in ways that involve both the right and left hemispheres of each person.

Logico-scientific paradigmatic processing is said to occur in a context-independent manner that focuses on abstract concepts and their logical, cause-and-effect relationships, and thus may rely primarily on left hemisphere dominant processing. The unfolding of each of these distinct processes coincides with the specialized processing of the right and left hemispheres and the sequential process of developmental maturation. See McGilchrist (2009) for an expansive and fascinating exploration of the origins and impact of cerebral lateral specialization.

Metacognition and Self-Reflective Capacity

Metacognition concerns conscious processes that act on cognitive processes —thinking about thinking. Awareness of cognition appears to develop by approximately 6 years of age and takes various forms, including the appearance–reality distinction (things may not be as they appear). Two components of this awareness are representational diversity (the same object may appear to be different to different people) and representational change (thoughts today are different from those of yesterday and may be different again tomorrow). This form of knowledge about the person-specific meaning of cognitive representation requires some sense of the person's awareness of the separateness of minds, a theoretical domain in developmental cognitive psychology called the *theory of mind*.

The regulation of cognition, also called *metacognitive monitoring,* includes such processes as planning activities, monitoring activities, and checking outcomes. Metacognitive monitoring may involve the assessment of thinking sequences for fallacious logic, factual errors, and contradictions in the content of speech. Peter Fonagy and colleagues explored the development of reflective function, or experiencing an internal observer of mental life. Parents teach children how to be self-reflective by including their own internal state in interactions with them and by encouraging children to share their own. Children who have been taught to tell stories that include mental states usually demonstrate a greater frequency of secure attachment. Being able to understand and consider the mental states of self and others has also been shown to decrease dependency on defensive strategies.

Research suggests that what is created in parent–child narratives about

experience is not just a story; embedded within the storytelling is the selection of information to be included, how it is to be processed and understood, and whether it is egocentric or has multiple subjective centers. Children who can appreciate more than one perspective in a story may develop greater empathic capacity.

Social Cognition

Bridging the fields of social psychology and cognitive psychology, the study of social cognition focuses on the mental processes involved in social interactions. Its domains include the study of empathy, interpersonal communication (verbal and nonverbal), person perception, relationship scripts, and group processes. Other related areas include studies of attribution bias, memory for social interactions, stereotyping, mental control of social cognitive processes, and cognitive origins of a sense of self. Social cognition can be seen as a domain of social psychology that uses information processing theory to assess the components of attention, perception, encoding, memory, retrieval, and schemata. A dominant theme in social cognition research has been that top-down, theory-driven processing influences interpretations of, and behaviors in, social situations. Developmental psychologists have focused on the origins of social cognition and its deviation to reveal, for example, that children with autism spectrum disorder (ASD) have significant deficits in empathic capacity and in the ability to interpret social cues. Social cognitive deficits are also present in other psychiatric disorders, such as bipolar disorder, PTSD, and social anxiety.

One view of social cognition includes the "simulation theory" whereby mirror neurons are hypothesized to be involved in an intricate process creating resonance, sympathy, and "mirrored states" in the perceiver. Mirror neurons activated during the perception of another's intentional acts stimulate neural networks that organize the same motor actions. This work has led to several hypotheses involving a process by which these cortical maps of behaviors are communicated to the limbic areas and down into the body itself. The resulting emotional and bodily states are then registered in the middle areas of the prefrontal cortex involved in self-experience. The "resonance" of the perceiver's limbic and somatic states with that of the person being observed then serves as an unconscious template through which the prefrontal region assesses how the self is "feeling." Thus, based on simulation theory, what one experiences as self-states are highly influenced by the experiences of those around them.

These networks are hypothesized to be the substrate of sensing another person's mind—those intentions and emotions experienced by another are now felt within the observer. Maps of others' minds are thought to be created in the medial prefrontal cortex and related areas allowing for the empathic understanding of another person's experience. One hypothesis

regarding the social cognition deficits in autistic individuals suggests that there may be difficulties in the activation of mirror neuron-related areas, either from impairment in input or functioning, to these social circuits of the brain. Further research is needed to validate this preliminary view, but it serves to highlight how impairments in neural mechanisms underlying empathy and the ability to map out another person's mind may be at work in a range of disorders.

Discourse and Narrative

Discourse is communication from one person to another; it is thought to involve a sense of intention or plan. Normal discourse follows a set of rules that ensures the coherence and effectiveness of communication as what is intended to be communicated by the sender is understood by the listener or receiver. Some researchers support the idea that discourse is a cognitive function that follows the basic principles of information processing, including a schema for effective communication that takes into account the listener's perspective. Incoherent discourse can be noted by analyzing violations of the primary maxims of discourse in a process called *discourse analysis,* which examines the ways in which discourse deviates from an assumed discourse plan.

The exact method to quantify abnormalities in discourse is controversial, but clinical impressions of incoherence are important for assessing deficits in thought and social communication. The deficits may result from learned behavior, inherent cognitive abnormalities in thought or language, or deviations in social cognitive functioning, and they are clinically evident in psychosis, specifically in schizophrenia, and in individuals with unresolved trauma.

Narrative is a broad domain of research ranging from the literary study of fiction to investigations of the origin of autobiographical accounts in developmental psychology. From a cognitive point of view, narrative is important in understanding the relationships among language, memory, consciousness, mental models, self-schemata, and social cognition. *Narrative* can be generally defined as the way in which a person creates a verbal account of a sequence of events in the world and an internal experience of the characters of the story.

Studies of early monologues find that young children interpret and assign meaning to events in their world from an early age, and that narrative helps to record and make sense of the past, interpret the present, and anticipate the future. The brain has been called an "anticipation machine," and mental models, prospective memory, and narrative are all ways in which networks of top-down processing prepare for possible futures. The enactment of narrative themes guides the way in which individuals live out their lives.

Cognitive Development

Developmental theories and research can be divided into several views: Stage theories (Jean Piaget, neo-Piagetian approaches, and the sociocultural school of Alexander Luria and Lev Vygotsky) describe discontinuous periods of development, with periods of stability and consolidation alternating with periods of instability and transition. Information-processing models postulate a nonstage theory, in which the emergence of cognitive capacity is a continuous process rather than a set of invariant sequences. Both stage and nonstage views embrace the idea that hierarchical integration and ongoing differentiation are fundamental aspects of cognitive development.

A distinguishing feature of each of these theories is the relative weight given to the role of innate, biological factors on the one hand and culturally determined social learning experiences on the other. In other words, do cognitive capacities emerge from a genetically determined plan, as in the Piagetian view, or do they develop in response to experience, as in the sociocultural view? More recent conceptualizations have drawn on the functioning of complex systems, supporting the idea of a transaction between innate factors and environmental experiences in the emergence of ever-more-complex capacities.

Psychiatric disturbances in cognition may reflect arrested patterns of typical cognition (mental retardation), atypical developmental pathways (ASD), and specific cognitive impairments (schizophrenia) that may have been present early on or that only became evident after life requirements, such as school, became demanding. Investigations into the developmental features of these disorders are a major focus of the field of developmental psychopathology.

Self-Organizational Processes

An understanding of the development and subjective experience of cognitive processes has been greatly informed by insights from the fields of evolutionary neurobiology and the nonlinear dynamics of complex systems, otherwise known as *chaos theory*. With billions of neurons and trillions of synaptic connections, the brain is capable of organizing an incomprehensible number of activation patterns. In selectionist theory, the billions of neurons become organized into networks with similar functions that can be selected and reinforced by interaction with the environment in an experience-dependent fashion. In addition, the brain has value systems that selectively reinforce the activity of neuronal groups that enhance survival. In this way, the brain's neuronal groups compete and differentiate within the brain in an ever-evolving adaptational system.

Chaos theory suggests that complex systems adhere to a specific set of principles. Three of these principles, nonlinearity, self-organizational

processes, and movement toward complexity, are especially relevant to psychiatry. Nonlinear refers to the finding that small changes in input (or initial conditions) can lead to large and unpredictable changes in output. Complex systems function on the rules of probability, which predict that certain combinations of activity within the system are more likely than others, and that these combinations will tend to move the system toward self-organization. This probability also predicts that the system moves itself toward increasingly complex states of functioning.

Complexity theory may offer a foundation for proposing a useful working definition of mental health applicable to individuals, families, and larger social systems. In complex systems, self-organizational processes that move the system's states toward maximal complexity are mathematically shown to be the most flexible, adaptive, coherent, energized, and stable. The movement toward complexity lies between the extremes of sameness, with rigidity and order on the one side, and change, with randomness and chaos on the other. Complexity is achieved when the components of the system achieve a balance between the two fundamental processes of differentiation (specialization in function) and linkage (connected together as a functional whole). When a system integrates differentiated parts, it achieves a state of complexity and harmonious, adaptive functioning that is bordered on either side by chaos and rigidity.

Examination of psychiatric syndromes reveals examples of deviations from this integrated complex state in which an individual exhibits states and traits of chaos and/or rigidity. For a given individual, such a balance can be achieved when the genetically and experientially influenced growth of neural circuits combines the differentiation of specialized regions with linkage via neural fibers that connect widely distributed areas into a functional, integrated whole. According to this view from the interdisciplinary field of "interpersonal neurobiology," disorder can be seen as occurring when a system's integration is impaired and its flow toward complexity is impeded as revealed in its movement toward either rigidity or chaos.

Trauma may impair integration by blocking neural linkages within an individual, as seen in the negative effects on the integrative regions of the corpus callosum, prefrontal cortex, and hippocampus in abused and neglected individuals. A dysfunctional family system would be thought to occur when the individuals are excessively differentiated (without emotional connections) or linked (enmeshing that inhibits individuality from being expressed). Such stressed systems are limited in their movement toward complexity and, hence, in their stability, flexibility, and adaptability. Such impediments to integration make a system prone to chaos and rigidity.

Psychiatric disturbances may be conceptualized as disturbances in self-organizational processes. Both inherited and experiential internal

determinants, as well as ongoing external, environmental, and social influences that place constraints on the system, can directly affect the development and effective use of integrative self-regulatory mechanisms. Integration is the underlying mechanism of optimal self-regulation. And interpersonal integration likely supports the growth of integrative regions of the brain. Clinical interventions may thus function at the level of external constraints (psychotherapy with integrative communication within a relationship) or internal constraints (pharmacological treatments) that alter the ways in which the individual's mind is able to achieve integration and healthy forms of self-organization. Optimal self-organization is created with integration, the linkage of differentiated parts of a system. By seeing the mind as something beyond "brain activity" and "subjective experience" alone, a definition of one aspect of mind as "the embodied and relational, self-organizing emergent process that regulates the flow of energy and information" makes it possible to understand the interconnected nature of the broad range of disorders and interventions as related to the fundamental process of integration. Viewing psychiatric disturbances in this way also allows for a synthesis of the views of psychodynamic, biological, and social psychiatry.

States of Mind

The state of activation of the various parts of the system can cluster into repeated patterns called *states*. In the brain, a *state of mind* or *mental state* describes the way in which various neuronal groups may become activated at a given time. Repeated patterns of neuronal group activation—a neural net profile—can become reinforced if they occur frequently or if the value system of the brain ingrains their profile. These ingrained patterns of activation are called *attractor states*; those states that are least likely to occur are called *repellor states*. These mental states are determined by the constraints on the system, and modification of constraints allows the nature of attractor and repellor states to be altered. Constraints on the system are both external and internal: Features of the external environment, such as the way in which other people behave and relate to an individual, can directly affect which mental state is more likely to be activated within the person. Internal constraints include the synaptic strengths of association, as determined by constitutional features and genetics, and those learned from experience, as encoded within memory processes.

Repeatedly reinforced patterns of neuronal group firing link the cognitive processes of attention, perceptual bias, memory, mental models, behavioral response patterns, and emotional tone and regulation. These states of mind result in the patterns of cognitive, emotional, and behavioral symptoms seen in both mental health and various psychiatric disorders; in a depressed state of mind, for example, one may pay conscious attention to

negative aspects of experience, interpret incoming stimuli in a pessimistic manner, have greater access to depressing past experiences, activate mental models of the self as bad or guilty, have behavioral patterns of withdrawal, and have a depressed mood with difficulty regulating intense affect.

Healthy mental functioning may depend on a flow of states of mind through time, which can then allow for flexible adaptation to an ever-changing environment. This relates back to chaos theory, which suggests that nonlinear complex systems must move continually toward maximizing complexity by balancing predictability and novelty. This is how mental health can be seen to emerge from integration, both within the individual and between the individual and the environment, especially the social world.

SENSATION, PERCEPTION, AND COGNITION IN PSYCHIATRIC DISORDERS

Since the time of Emil Kraepelin and Eugen Bleuler, psychiatrists have known that certain disorders include profound disturbances in cognitive functioning. Since the 1950s, researchers have attempted to determine the exact nature of such deficits. With advances in computer technology and an increasing technical ability to analyze stimulus presentation and response times on the order of tens of milliseconds, cognitive psychologists have been able to devise research paradigms capable of testing increasingly subtle aspects of cognitive processing.

Processing research has focused on all three domains of sensation, perception, and cognition. Sensory-processing studies, using simple stimuli, focus on poststimulus events for as long as 1 second; perceptual studies, using slightly more complex stimuli, examine processing after a period of as long as approximately 5 seconds. Cognitive processing experiments can examine the early aspects of processing (e.g., phenomena occurring within the first 30 seconds), as well as long-term processes that occur over hours, days, or even years. Recent attempts have been made to correlate complex cognitive findings with clinical presentations.

A general problem with correlating cognitive processes with clinical populations lies in the symptomatic heterogeneity of patients falling within the same diagnosis. A related problem is the distinction between general and specific deficits: Do psychiatrically ill patients perform less well on a given task because they are ill or because of a deficit specific to the disorder? For example, psychomotor slowing, as measured by reaction time and response rate, is seen in schizophrenia, depression, and other psychiatric and neurobehavioral disorders. The side effects of medications can also influence processing and response speed. Researchers rely on the creative design of experimental tasks to help distinguish between general and specific deficits. A comparison of target-patient populations with matched healthy persons and other psychiatric patients can help to

determine disorder-specific cognitive dysfunction.

Another general issue is that of state markers versus trait markers; a patient with schizophrenia, for example, may have a cognitive deficit only when actively psychotic (state) or all the time (trait). These results have been found in certain cognitive tests of attention that correlate with improvement on medications. Some atypical results are also found in the non-ill first-degree relatives of schizophrenic patients. Is the marker of genetic vulnerability a coincidental finding or part of the core deficit in schizophrenia? An exploration of the implications of these cognitive differences for the daily life of the patient is an important application of the research findings to clinical psychiatry.

Schizophrenia

In the late 1890s, Kraepelin described a primary attention deficit in his elaborate clinical description of patients with schizophrenia. Some cognitive tasks have been identified as trait-linked markers of schizophrenia: reaction time crossover, backward masking, dichotic listening, serial recall tasks, vigilance (sustained attention) tasks requiring high processing loads, and span-of-apprehension tests with large visual arrays. Deficits in those areas have been explored through many studies examining various aspects of processing.

Reaction Time Crossover and Modality Shift Effects. These paradigms examine the general finding that schizophrenic patients have a slower-than-usual response on tasks that require rapid reaction times. On a test in which a stimulus is presented with varied combinations of warning signals and preparatory intervals, schizophrenic patients show an advantage only with short preparatory intervals and with long response times that have regularly spaced stimuli, a pattern distinct from that of nonclinical controls (crossover effect). In a related paradigm, when the modality of the stimulus is varied (e.g., light is interspersed with tone), the latency (delay) of the response in schizophrenic patients, when compared with controls, is longer if the preceding stimulus was of a different modality. This phenomenon, termed the *modality shift effect,* reveals a greater degree of cross-modal retardation in schizophrenic patients than in controls.

Visual Backward Masking, Sensorimotor Gating, and Habituation. In visual backward masking, a stimulus is followed by an interval of time and then a subsequent stimulus is presented. Figure 3.1–2 shows a typical masking experiment. The presentation of the secondary stimulus leads the patient with schizophrenia not to report (or mask) the initial stimulus. Lengthening of the interstimulus interval beyond 500 ms can lead to normalization, with no masking present. Some studies find that the impairment improves with treatment by medication and can be induced in

nonclinical patients given catecholaminergic agents. Other studies find that the impairment may be a marker of increased vulnerability to schizophrenia.

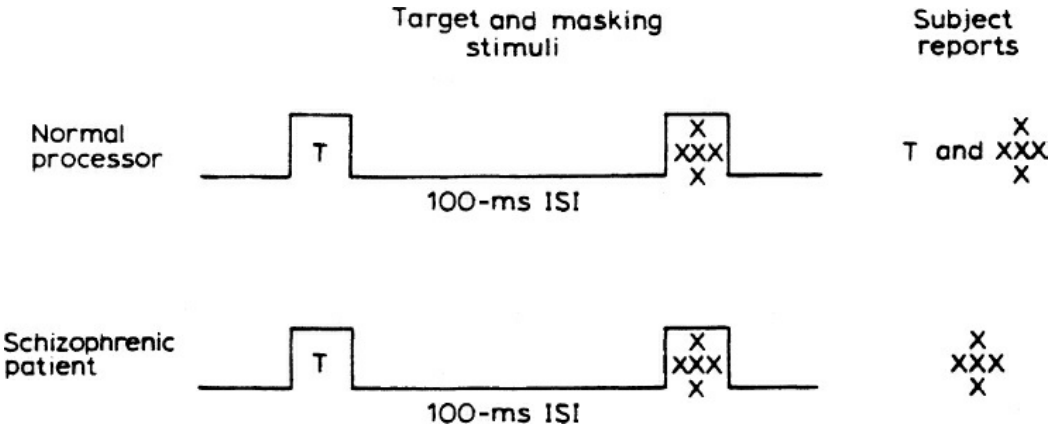

FIGURE 3.1–2. Diagram showing the difference in the verbal reports of normal versus schizophrenic subjects presented with a single backward masking trial with a 100-ms interstimulus interval (ISI). The T represents the target stimulus, and the Xs represent the masking stimulus. (Reprinted with permission from Braff DL, Saccuzzo DT, Geyer MA. Information processing dysfunctions in schizophrenia: Studies of visual backward masking, sensory motor gating and habituation. In: Steinhauer SR, Gruzelier JH, Zubin J, eds. *Handbook of Schizophrenia*. Vol. 5. Neuropsychology and information processing. New York: Elsevier Science; 1991.)

Sensorimotor gating and habituation are the processes by which the reaction to stimuli decreases with repeated presentation. Schizophrenic patients show a markedly diminished capacity to habituate, demonstrating a persistent acoustic startle reflex with repetitive tones. Lysergic acid diethylamide administration and the intracerebral injection of dopaminergic agents in rats lead to similar findings, supporting the idea that excessive dopamine activity, which is believed be central in schizophrenia, can induce those deficits.

In general, deficits in habituation and visual backward masking lend support to the idea that schizophrenic patients have a diminished capacity to regulate the flow of rapidly presented information; they experience being inundated by stimuli that are regularly filtered out in the brain of a nonclinical person. Deficits in processing may also occur with internally generated stimuli, producing an overwhelming and confusing experience of the internal and external world.

Selective Attention. In general, selective attention experiments present the person with a target stimulus and distracters. In dichotic listening tasks, the subject is asked to attend to messages presented to one ear and to ignore messages presented to the other ear. Studies of schizophrenic patients have revealed consistent deficits in their ability to repeat the message on which they were asked to focus. Analysis of these findings

suggests that schizophrenic patients have an impairment in their ability to avoid distracting stimuli (to filter) and to pigeonhole (to use category features to reduce stimulus qualities needed to respond).

These findings were explained by using the framework of an impaired filtering structure and pigeonholing process, but recent conceptualizations have also examined a generalized impairment in the information-processing capacity in schizophrenia. The capacity model examines the way in which a pool of attention is allocated across mental activities. Two components of this model are the *quantity of resources available* (capacity) and *executive allocation policy*. Schizophrenic patients may also suffer from an impaired response selection process, and thus demonstrate atypical responses on tasks.

Sustained Attention. Sustained attention, or *vigilance,* is required to process stimuli of long duration. The most common research paradigm used to examine sustained attention is the *Continuous Performance Test*. This test consists in rapidly presenting a set of tasks with varied spacing and timing of target and nontarget stimuli. The processing load for a Continuous Performance Test can be varied by blurring the stimuli presented or by changing the pace of presentation.

Two elements distinguished are sensitivity and the response criterion. Diminished sensitivity is a sign of decreased vigilance and results in a high miss rate (errors of omission); when the response criterion is diminished, a high false-positive rate (error of commission) results. The analysis of response features is important in the interpretation of results and has led to the finding that schizophrenic patients have a deficit in their ability to distinguish target stimuli from nontarget stimuli, especially when the stimuli are presented as brief signals at a rapid pace.

When compared to nonclinical controls, positron emission tomography (PET) in schizophrenic patients performing a Continuous Performance Test revealed lower metabolic activity in the prefrontal cortex bilaterally but normal or elevated activation in the occipital region. This finding is consistent with other findings that support the idea of impaired frontal functioning in schizophrenia.

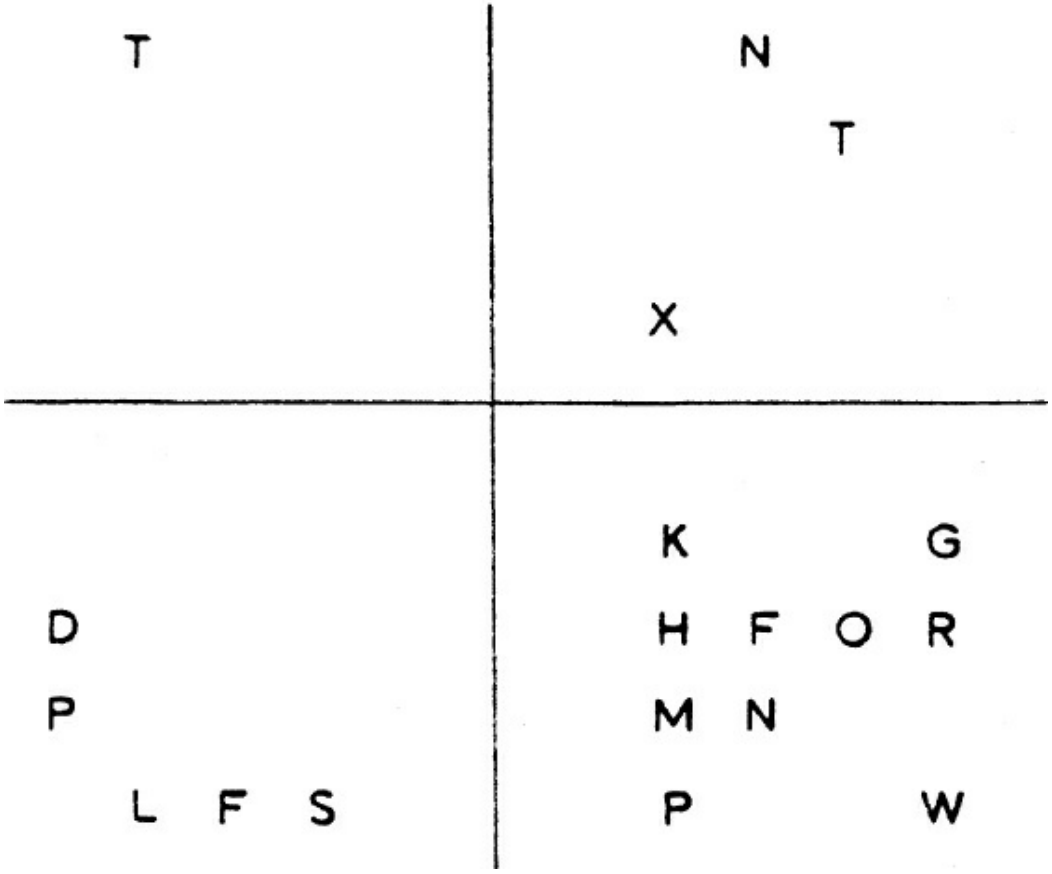

FIGURE 3.1–3. Sample arrays used in the wide–visual-angle version of the partial report span-of-apprehension task. (Courtesy of Robert F. Asarnow, Ph.D.)

Span of Apprehension. In the span-of-apprehension test, an array of letters is displayed for a brief period (from 50 to 100 ms in most studies). One of the letters is a T or an F, and the person must detect which letter is present. The number of nontarget letters is increased, and significant differences in detection are found for displays of 10 or more letters. Figure 3.1–3 provides an example of a visual display for the span-of-apprehension test.

The serial scanning process is an element of focal attention. Parallel processing in the search involves increased aspects of assessment of figure–ground and textual segregation and is thought to be an automatic process. Schizophrenic patients show significantly increased errors in the span-of-apprehension test under conditions of increased complexity of display. Their scores also worsen in psychotic conditions and improve with symptomatic improvement while they are taking medications. Increased errors also occur in nonschizophrenic mothers of children with schizophrenia. Thus, the span-of-apprehension test is a measure of state and trait in some cases of schizophrenia. Short- and long-term outcome studies have found that those patients with atypical test results whose scores improve after receiving antipsychotic medication have a good clinical response to pharmacotherapy.

Data from numerous studies show that patients with ADHD have dysfunctions on a variety of tasks, including those that involve monitoring, perception, memory, and motor control. Intact performance has been found on a number of memory tasks requiring verbal processes (digit span, word tests, and story recall) and nonverbal processes (recall, visual arrays, and block series).

In general, patients with ADHD evidence behavior that is consistent with patients suffering from frontal lobe damage, specifically with regard to deficits in the control of motor responses, in the execution of fine-motor movements, and in the inhibition of ongoing response patterns. Although memory tasks and basic aspects of information processing remain intact, these patients have impaired performance across several modalities (auditory, visual, motor, and perceptual-motor), suggesting some global deficit. These patients also appear to be unusually susceptible to boredom when the required tasks are long and repetitive.

One view examines two proposed systems, an underactive behavioral inhibition system and an impaired behavioral reward system, to explain the behavioral challenges of ADHD. A related view is that the rule-governed behavioral system is not intact, as evidenced by the fact that patients with ADHD appear to do especially poorly with delayed or nonexistent rewards, including those tasks that require sustained attention, accuracy, or task-directed activity governed by another person's direction or rules. Under these conditions, the patient's poor regulation and inability to meet functional demands are revealed. A related issue is a diminished motivational drive and, possibly, a diminished arousal regulation system.

Another perspective supported by research data is that ADHD patients have metacognitive differences. According to this view, the metacognitive processes that help to plan, monitor, and regulate performance are impaired, and the patient's ability to assess the task and determine strategies is deficient. This is an example of a top-down aspect of attention because there is impairment of the higher cognitive processes that regulate information flow. Bottom-up deficiencies, on the other hand, would involve impairment in basic aspects of attentional focus due to abnormalities in arousal, capacity, and selectivity.

The finding that numerous factors can influence the appearance of deficits led Virginia Douglas to hypothesize that ADHD is a self-regulatory disorder with pervasive effects. According to Douglas, the impairments affect each of four domains—attention, inhibition, reinforcement, and arousal—and result in deficits in several aspects of self-regulation: (1) the organization of information processing, including the making of plans, metacognition, executive functions, the adoption of appropriate cognitive sets for a given task, the regulation of arousal levels and alertness, and self-

monitoring and self-correction; (2) the mobilization, deployment, and maintenance of adequate attention; and (3) the inhibition of inappropriate responses to extraneous stimuli and reinforcers. These deficits in self-regulation imply that increased processing demands result in the diffusion of attention and the impairment of the in-depth, coherent acquisition of knowledge and understanding.

The diversity of research findings and theoretical explanations is paralleled by the clinical finding that children who are severely impaired in the classroom may have no attention problems in the confined, one-on-one setting of the psychiatrist's or psychoeducational examiner's office. Such children may also be able to attend for indefinite periods to video games and yet be unable to follow complex conceptual information. This is an important reminder that cognitive dysfunction in psychopathological conditions may be task, context, or relationship specific, depending on the nature of the cognitive impairment. The patient's clinical history and evaluation must therefore consider potentially hidden domains of atypical cognition.

Autistic Spectrum Disorder

Although early descriptions of autism delineated deficits in social functioning, later studies found cognitive dysfunctions involving abstraction, sequencing, language, and comprehension. Researchers are now focusing on the nature of the core deficit in the disorder, and how the cognitive domains relate to the social-affective deficits is of particular interest. However, there may be many "autisms" and finding a single process may prove unhelpful. In general, three broad areas appear to be atypical for those on the autism spectrum: sensory integration, emotion regulation, and social cognition.

A majority of patients with ASD are assessed as having mental or cognitive deficiencies, although cognitive impairments are especially difficult to assess if language functioning is severely limited. Mental retardation involves global, cognitive, and language impairments that may make it difficult to distinguish autistic disorder features. Studies of high-functioning patients with autistic disorder have permitted a deeper exploration of a variety of cognitive deficits. An array of dysfunctional areas, including numerous language problems, excessive or impaired responsiveness to stimuli of various modalities, different encoding of auditory stimuli, and impairment in the ability to extract important features from incoming information, have all been found; in contrast, some patients with autistic disorder have relatively intact visuospatial and Gestalt functions, musical abilities, and rote memory. Performances on standardized tests reveal relatively good results in object assembly and block design but poor results in comprehension. Language deficits vary and include syntactic, phonological, prosodic, and pragmatic domains, also

suggesting right-hemisphere involvement. The wide array of dysfunctions in autistic disorder may be due to the existence of a variety of subtypes with dysfunction in different loci in the brain. One general finding is that of a larger brain in individuals with autism, with some suggestions that differentiation has become impeded, thus impairing the capacity for integration and the achievement of complex states of cognition and behavioral social interactions.

Other studies have focused on the social cognition of patients with ASD; they have examined the nature of the patient's emotional behavior and level of understanding in order to assess the earliest manifestations of an atypical emotional connection between autistic children and their parents. Recent work in neurobiology suggests a role of the orbitofrontal cortex and the cerebellum in mediating some of these deficits. Two findings support these initial impressions: (1) Autistic children are much less likely than non-autistic children to imitate adult vocalizations and gestures, and (2) they show much less sophisticated representational play with objects than do non-autistic children. These findings have led to the suggestion that a core deficit in ASD exists in the representation of representations (meta-representations), leading to deficits in symbolic play and to the inability to understand the mental states of others. The inability to transfer cognitive representations into language symbols may be a related meta-representational deficit.

A series of studies explored the relation of these possible cognitive impairments to socioemotional behavior. In contrast to clinical lore, children with ASD were found to look at their parents; they had eye contact with their parents when social interactions were parentally elicited, and they revealed a typical distribution of behavioral patterns of attachment, even if the behaviors were not typical themselves. These studies did find a marked lack of social referencing (looking to parents for emotional cues in ambiguous situations) and protodeclarative gestures (pointing to objects and showing objects to familiar adults); three hypotheses have been proposed to explain these findings: (1) Autistic children may not have the capacity to have a representation of another person as having ideas, perspectives, or emotions that can be shared. This proposal is consistent with a "theory-of-mind" hypothesis, in which a core deficit in autism is believed to be the inability to have a sense of another's mind; (2) autistic patients may have an impaired ability to perceive or to comprehend the emotional (usually facial) signals of others; and (3) the core deficit may involve a lack of interest in others or an aversion to responding to others. Studies have found that although children with autistic disorder do express emotions, they have less positive affect during periods of joint attention with another person. Furthermore, they reveal impairments in their responsivity to the display of strong emotion by another, whether it is of distress or of pleasure.

Tests of high-functioning autistic patients reveal poor performance on emotion-recognition tasks, little comprehension of, and empathy with, depictions of social situations, and difficulty in talking about socially derived emotions, such as pride and embarrassment. Autistic patients with relatively high intelligence use adaptive cognitive strategies to interpret social stimuli to compensate for impaired emotion-processing abilities. The development of social cognition and socioemotional understanding requires complex interactions among the cognitive, perceptual, and emotional processes.

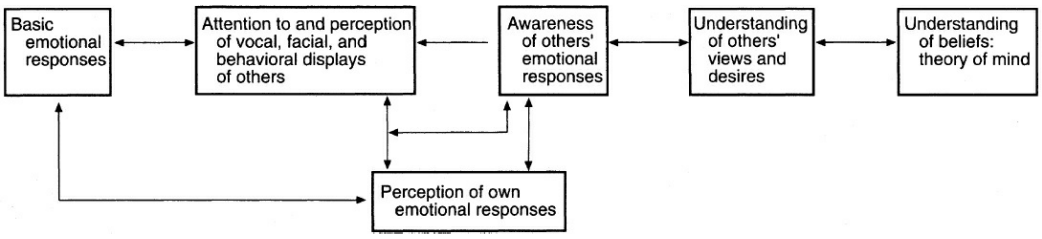

FIGURE 3.1–4. Sigman's model for the development of socioemotional understanding. (Courtesy of M. Sigman.)

A series of interactive elements essential to the development of social understanding has been proposed (Fig. 3.1–4) to describe the basic precursors of emotional responsiveness: The ability to attend to, encode, and interpret verbal and nonverbal social stimuli, the awareness of one's own and others' emotional responses, and the ability to contrast oneself with others. Out of that matrix develops the ability to understand others' views, desires, and beliefs. Accordingly, a deficit in any one of those basic elements may explain the characteristic deficits observed in the social cognition of persons with ASD. As discussed in the section on Social Cognition, one hypothesis in need of further validation is whether an atypical functioning of mirror neurons and related circuits plays a role in the social deficits of autistic individuals.

Mood Disorders

In contrast to schizophrenia, ADHD, and autistic disorder, mood disorders do not appear to have core cognitive deficits that are diagnosis specific. Instead, cognitive abnormalities appear to be related to the degree of psychopathology and the severity of the mood disturbance. In other words, they are state dependent. Most studies have examined patients with depression, whereas only a few studies have assessed cognitive functioning in patients with bipolar I disorder during a manic state. In depressed patients, the severity of the depression has ranged from mild depression in students to severe illness in hospitalized patients with major depressive disorders. The studies have primarily examined attention and memory for neutral and emotionally toned stimuli; for the majority of these studies, the

concept of self-organizational processes and state regulation has not been the primary focus of attention. These recent conceptualizations of the brain's functioning as a nonlinear complex system capable of self-organization, however, may aid in the future investigation of the primary dysregulatory aspects of mood disorders.

Depressive Disorders. Depressed patients often complain of difficulties with concentrating, learning, and remembering. Studies have documented that such patients perform poorly on tasks that require sustained attention, or effortful and elaborate rehearsal, and thus, controlled limited-capacity processes appear to be impaired in major depressive disorder. This limitation of access to capacity-demanding resources appears to be directly related to the severity of the depression, and normalizes with remission from a depressive episode.

Depressed patients have also been found to require increased supportive data from the presented stimuli before responding in a test situation. Whether this pause before responding is a psychological response to being depressed, an emotional "need to be certain," or a specific feature of the depression itself, is yet to be been determined. The reluctance to respond must be considered in interpreting research data and clinical interview findings.

The attention and memory findings have suggested several theoretical frameworks that are clinically useful: A schema theory of depression outlines a positive feedback loop in which negative self-schemata prime persons to have negative thoughts, to recall negative events in their lives, and to interpret present events with a negative bias. Whether as a cause or as a maintaining influence, these depressogenic schemata are thought to create a series of cognitive functions that produce and maintain a depressed mood.

A network theory of memory and emotion, supported by research on depressed and nondepressed persons, posits that mood leads to a spreading activation of items in memory that are congruent with the mood. Thus, emotion directly influences retrieval by a process of state-dependent learning and memory. Furthermore, while depressed, patients are likely to encode items in a form that makes them readily accessible when retrieved in a depressed state. Depressed mood, then, becomes an internal context cue that accesses depression-related memories.

The network and schemata theories of depression provide a framework for understanding how emotions and moods influence cognitive processing. Emotions in health and illness can shape the mental state instantiated at a given time by triggering the activation of previously formed depressive schemata. The activated schemata, in turn, can produce retrieval biasing and behavioral responses fundamental to a mental state, which can then further elicit a negative emotional response; this self-

reinforcing loop is established and supports the continuation of both depressed cognition and mood.

Some patients may be especially prone to marked cognitive alterations based on rapid shifts in mood, and these shifts may be a learned or constitutional feature of the individual. From the nonlinear dynamics view, rapid shifts in mental state can be conceptualized as sudden and intense changes in the constraints on the system, which determine the flow of states of mind across time. The depressed state may become deeply engrained as an attractor state that is difficult to alter. Pharmacologic interventions may aid in directly altering the synaptic constraints that maintain such a state, thereby promoting neural plasticity. Psychotherapeutic interventions, such as cognitive-behavioral, interpersonal, and mindfulness-based therapies, can all be theorized to produce changes in the internal processing of constraints, as well as in the ways in which external constraints from the social world interact to shape the overall shifts and stability of mood.

Bipolar I Disorder. Patients with manic episodes present with a spectrum of state-related cognitive dysfunctions tied to the severity of their episode. Manic disturbances have been described clinically as rapidly paced thinking and speech, quick associations to self-generated or other-generated stimuli, grandiosity, and increased distractibility.

Studies have suggested a high rate of combinatory thinking and the inclusion of loosely associated but related intrusions. The clinical impression of humor (even in the face of an underlying dysphoria) in some patients is corroborated by playful, extravagant, or flippant elaborations, as well as intrusions in speech, and seems to indicate primarily state-dependent symptoms that improve on recovery. Forced-choice, span-of-apprehension, and backward masking tasks all reveal impairments similar to those seen in actively psychotic patients with schizophrenia and yet are unlike those seen in schizophrenic patients who have persistent deficits in the remitted state. Bipolar disorders, then, can also be viewed as disorders of self-organization in which the system fluctuates between the extremes of highly activated manic states and highly deactivated depressed states of mind. Structural brain imaging of patients with both disorders has revealed impairments to integration in the form of impediments to linkage.

Anxiety Disorders

Cognitive studies of patients with anxiety disorders have focused primarily on attention or memory. Various research paradigms have been applied to assess attention bias and memory retrieval for neutral and emotionally activating stimuli. Studies find that anxious patients have an increased tendency to attend to fear- and threat-related words. One research approach includes a dichotic listening task in which an anxious patient is

more easily distracted, than are controls, by fear-related stimuli in the nonattended channel. This finding suggests that anxious patients have automatic, parallel attention processes that are primed to detect certain types of stimuli.

Another approach uses the Stroop paradigm, in which words are presented in different-colored inks and the person must look at the word and state the color of the ink. Anxious patients have a significant delay in their response times for fear-related words, a finding that suggests that increased attention capacity or cognitive processing is necessary when those "fear" words are perceived and the color of the ink must subsequently be determined.

One theory that explains these aforementioned findings posits the idea of a "fear network" that encodes fear-related information in a memory structure that is readily accessible and able to influence cognitive, motor, and psychophysiological responses. The theoretical fear networks, which may be similar to mental models, are thought to contain three related forms of information: (1) fear-eliciting stimulus cues, (2) specific response patterns, and (3) the meaning of the cues for that particular person. According to this theory, patients with anxiety disorders are believed to have fear networks that are especially coherent and stable, easily activated, and require few environmental cues.

Posttraumatic Stress Disorder

Patients with PTSD have been found to have attentional biases toward threat-related stimuli specific to the experienced traumatic event. Cognitive aspects of the disorder are characterized by intrusive processes (memories, images, emotions, and thoughts) and avoidance elements (numbing, amnesia, and withdrawal), which are driven by increased arousal. These patients are thought to have a unique configuration of fear networks containing stimulus cues (environmental stimuli), response components (cognitive, motoric, and psychophysiological), and meaning elements (the moral implications of the trauma, survivor guilt, and the meaning of intentional trauma vs. accidental trauma).

Some theories argue that states of excessive arousal during trauma impair attention capacity and memory encoding during that event. Emotional processing during and after a traumatic experience may also be hampered by the heightened state of psychophysiological arousal and altered memory storage. In fact, the part of the brain necessary for explicit memory processing, the hippocampus, has been shown to be smaller in individuals experiencing chronic PTSD. Patients who have used dissociative mechanisms during and after a traumatic experience appear to be at far greater risk for developing PTSD.

PTSD may have some aspects that are adaptive: Attention biases primed to detect fear-related stimuli may permit the early detection of threatening

situations that, if not avoided, would produce incapacitating psychophysiological arousal. Automatic, nonconscious behavioral avoidance response patterns embedded in the proposed fear networks or mental models allow these patients to minimize excessive arousal by avoiding trauma-related situations.

The psychopathological aspects of PTSD can be cognitively exemplified as follows: A combat veteran with chronic PTSD may have no direct recall (impaired explicit memory) of a helicopter crash in which his best friend, who was seated next to him, was killed. Years later, continued avoidance of airports, amnesia for combat, general apathy, and social withdrawal (avoidance elements), combined with exaggerated startle response, panic attacks, intrusive images, and nightmares (intrusive components), all suggest intact implicit memory of the combat trauma; the veteran has become emotionally, behaviorally, and cognitively impaired.

Speechless Terror. Scott L. Rauch and colleagues explored the neurobiology of intense fear, using patients with PTSD. They took eight patients with PTSD and exposed them to two audiotapes. One tape was emotionally neutral, and the other was a script of a traumatic experience. While patients listened to these tapes, measures of their heart rate and regional cerebral blood flow (rCBF) were measured via PET scans. Studies found that, during the play of traumatic audiotapes, rCBF was greater in right-sided structures, including the amygdala, posterior medial, orbitofrontal, insular, anterior, and medial temporal lobes, and anterior cingulate cortex. These are the areas thought to be involved with intense emotion. Such findings are consistent with the view of unresolved PTSD as a state of impaired integration prone to the chaotic (intrusive) and rigid (avoidant) symptoms. In developmental trauma too, those who have experienced abuse or neglect have been found to have structural differences in the brain's integrative regions, including the hippocampus, prefrontal cortex, and corpus callosum as described below. Psychotherapeutic intervention in posttraumatic patients may yield significant healing by way of promoting growth of integrative functioning which future research may reveal as including the growth of integrative structures in the brain of those who resolve their previously unresolved states.

An extremely interesting and potentially important clinical finding was the decrease of rCBF in Broca area (left inferior frontal and middle temporal cortex) in states of high arousal. These findings suggest an active inhibition of expressive language during trauma and high levels of stress. Based on these results, speechless terror, often reported by victims of trauma, may have neurobiological correlates consistent with what is known about brain–behavior relationships. This inhibitory effect on Broca area also impairs the encoding of conscious memory for traumatic events at the time that they occur. It then naturally interferes with the development of

narratives that later serve to process the experience and lead to neural network integration and psychological healing. Activating Broca area and left cortical networks may thus be essential in psychotherapy with patients experiencing PTSD and other anxiety-based disorders.

Approaches to the treatment of patients with PTSD need careful evaluation but generally include the view that the impaired emotional processing of the traumatic event requires the active recollection, in explicit terms, of the details of the experience. The process of effectively treating unresolved trauma usually involves the active cognitive processing of specific memories, including emotional responses, derived belief systems, and the psychophysiological arousal at the time of the traumatic event and during recall.

The provision of new cognitive information in the course of psychotherapy, as well as the relative safety of the therapist's office, can, in theory, widen the "windows of tolerance" for certain emotions and memories to be held flexibly in consciousness and to alter the configuration of the fear networks and allow previously inaccessible information to be explicitly processed and made available to consciousness for incorporation into memory and an ongoing autobiographical narrative. Specific techniques, such as those that facilitate cognitive processing of previously dissociated or in other ways isolated cognitions—such as beliefs, images, and sensations—may be useful in alleviating the symptoms and dysfunction after acute or chronic trauma. Such therapy and changes in mental processing may reduce the avoidant and intrusive components of the clinical syndrome of PTSD, and they also may improve social, emotional, and cognitive functioning.

The evaluation and treatment of PTSD patients can be greatly enhanced by an understanding of the relevant cognitive processes and a careful application of the assessment to presenting symptoms and signs. Several related areas that have interested clinicians for decades, and which have become a source of great societal concern and controversy, include the delayed recall of repressed memories of traumatic events and the suggestibility of patients influenced by clinicians, society, or friends to believe that they are the victims of childhood trauma.

Delayed Recall. Many scientists and clinicians believe in the cognitive capacity of patients to be unaware, for years or even decades, of severely traumatic experiences that took place in their childhoods. Other researchers disagree and emphasize the paucity of studies of corroborated cases of childhood trauma that have been followed prospectively into adulthood and have documented impaired access of consciousness to events presumably stored in memory. Two distinct mechanisms that may explain delayed recall of childhood trauma are *repressed* memories and *dissociated* memories.

A repressed memory can be thought of as originating from the active, intentional suppression of memory from consciousness. The mechanisms that underlie this process may then become automatic, and the contents of memory may become inhibited from retrieval into consciousness. This blockage may exist to avoid flooding the person's awareness with information that is associated with excessive anxiety or fear that would impair effective functioning. This is an example of knowledge isolation in which information may be layered in the nervous system and certain aspects may be kept from conscious awareness.

In contrast, a traumatic event may be so overwhelming that adaptive processing may be impaired. If focal attention is divided or if the hippocampus is shut down temporarily with high cortisol secretion, the non–focally attended (traumatic) material will only be processed implicitly. Thus, to adapt to a traumatic event, some persons may have the capacity to focus their attention on a nonthreatening aspect of the environment or on their imagination during the trauma; this may be an underlying mechanism in a process called *dissociation*. Traumatic memory that has only been implicitly encoded affects behavior and emotions and possibly contains intrusive images and bodily sensations that are devoid of a sense of past, of self, or of something being recalled. This may partly explain such symptoms of PTSD as amnesia (blocked conscious access to a memory or its origin), avoidance behaviors, hyperarousal, intrusive images, and flashbacks.

A person may use both mechanisms of repression or dissociation for different aspects of a traumatic event, and recollection of the traumatic memory may take different forms. Repressed memories may have been processed to some degree in narrative form, whereas dissociated memories probably lack integrative processing. This latter form may thus be experienced as nonpast and nonself, making the intrusive retrieval of dissociated memories a confusing and frightening experience. This is implicit memory retrieval of traumatic events.

Studies of the development of memory in children suggest that the shared construction of narratives about experienced events is often crucial in making the events accessible to long-term retrieval. The establishment of a personal memory system appears to be a function of such "memory talk" in which parents discuss with children the contents of their memory. In children who have been forced to keep traumatic events a secret, as may occur with childhood abuse, the important developmental process of narration may be blocked. This may be an additional cognitive mechanism underlying the inaccessibility of some forms of childhood trauma to consciousness in adult patients.

Suggestibility. Numerous studies have demonstrated that human thought and memory are easily influenced by suggestion from others.

Suggestibility is adaptive for a social being who relies on the experiences of others for knowledge of the world. Therefore, listening to the stories of others, reading a textbook, and being coached in athletics all require the receiver to accept data from the sender. Critical analysis of the data received is also an important component of learning; however, the metacognitive function of assessing the accuracy or usefulness of newly acquired information may be suspended under certain conditions, including hypnosis, drug-altered states, and conditions of severe threat.

Studies of human suggestibility indicate that post-event questioning can bias the meta-memory processes that help to determine the source and accuracy of a retrieved memory. The verbal and nonverbal cues given by the interviewer may influence a person to believe that actual aspects of an event never took place. A person can also be convinced of the accuracy of an event despite its lack of correspondence with actual experience. Factors that may influence the biasing of interviewees include a belief in the trustworthiness and authority of the interviewer, not being aware that "I don't know" is a permissible response, repetition of a question that has already been answered, and the interviewer's beliefs as communicated through emotional tone and nonverbal gestures.

It is crucial for clinicians to be aware of human suggestibility to avoid iatrogenic distortions. Similarly, it is important for persons who experienced severe trauma early in life to receive informed and empathic evaluations and treatment. There is a delicate balance between supportive neutrality and active advocacy in assessment and intervention; awareness of these fundamental cognitive processes may help to guide the clinician toward achieving this goal.

Complex PTSD. Complex PTSD results from prolonged and inescapable stress and trauma during early childhood. It is complex because of its extensive impact on all areas of development and functioning. Trauma shaped personality traits and coping strategies tend to increase the individual's vulnerability to future trauma through engagement in abusive relationships, cognitive distortions, poor judgment, and a lack of self-protection. Long-term PTSD has been shown to correlate with the presence of *neurological soft signs,* suggesting either a vulnerability to the development of PTSD or the long-term impact of the physiological dysregulation caused by PTSD.

Research on individuals with documented early developmental trauma in the form of severe neglect or abuse (emotional, physical, sexual) have revealed that the regions of the brain that are impaired in their growth are those that link widely separated areas to each other. These integrative regions are affected differentially depending on the timing of the trauma and include the hippocampus which links various forms of memory into explicit factual and autobiographical memory, the corpus callosum linking

left and right sides of the brain, and prefrontal cortex, which links the cortex, limbic region, brainstem, bodily signals, and the social input from other brains into a coherent whole. In many ways, developmental trauma can be thought of as a severe example of impaired interpersonal integration. With neglect, there is an impairment to linkage; with abuse, there is an impairment in honoring differential needs of child and adult.

This clinical finding of developmental trauma leading to impaired neural integration supports the following possibility: Interpersonal integration (honoring differences and promoting compassionate linkages with other people) stimulates the growth of neural integration (the linkage of differentiated regions of the brain to each other). If this proposal is true, it may mean that psychotherapy that is integrative may promote a healing process in those with complex PTSD that helps stimulate the activation and growth of neural integration. Future research will need to validate this hypothesis of a potential core mechanism in how therapeutic relationships work to improve the lives of those who have experienced severe forms of trauma early in development.

Why would developmental experiences that overwhelm an individual's ability to cope have such a profound impact on their growth? When confronted with threat under usual adult circumstances, the processes related to arousal and the fight-or-flight response become activated; the threat is dealt with and soon passes. For obvious reasons, children are not well equipped to cope with threat in this way. Because their survival is linked to dependency, fighting and fleeing with an attachment figure may actually decrease their chances for survival. When a child first cries for help but no help arrives, or when trauma is being inflicted by a caretaker, the child may shift from hyperarousal to dissociation. Traumatized children who are agitated may be misdiagnosed as having attention-deficit disorder, and the numbing response in infants can be misinterpreted as a lack of sensitivity to pain.

Dissociation allows the traumatized individual to escape the trauma via a number of biological and psychological processes. Increased levels of endogenous opioids, for example, create a sense of well-being, pain reduction, and a decrease in explicit processing of the overwhelming traumatic situation. Psychological processes such as derealization and depersonalization allow the victim to avoid the reality of his or her situation or to watch it as an observer. These processes provide the experience of leaving the body within awareness, traveling to other worlds, or immersing oneself into other objects in the environment. Hyperarousal and dissociation in childhood create an inner biopsychological environment primed to establish inflexible boundaries between different emotional states and experiences for a lifetime. If it is too painful to experience the world from inside one's body, self-identity can become organized outside of the physical self.

Early traumatic experiences shape neurophysiology and neuroanatomical networking, thus impacting experience and adaptation throughout development. The tendency to dissociate and disconnect various tracks of processing creates a bias toward unintegrated information processing across conscious awareness, sensation, affect, and behavior. General dissociative defenses result in an aberrant organization of networks of memory, fear, and the social brain and contribute to deficits of affect regulation, attachment, and executive functioning. One way of understanding such fragmentation of mental processes is in the simultaneous activation of two mutually incompatible motivational drives, mediated by two differentiated circuitries. One is the mammalian limbically mediated drive to move toward the caregiver when the infant or child is in a state of distress. The other is the brainstem-mediated drive to move away from the source of terror.

If the attachment figure is the source of terror, then the two circuits driving the body to simultaneously move both toward and away from the same individual creates a biological and psychological dilemma. This has been called "fear without solution" and is a biological paradox thought to be the mediating mechanism of disorganized attachment common in developmental trauma and often associated with clinical degrees of dissociation. Without a solution, the organized strategy to cope with stress is overwhelmed, and the mind of the child cannot achieve integration as the differentiated circuitries are unable to become linked in a coherent way to deal with the abuse or neglect. Disorganized attachment experiences may be one common developmental origin of clinical dissociation and complex PTSD.

The lack of a coherent integration of the interdependent systems within an individual, which may be due to extreme early stress, results in many disorders: Compulsive disorders related to eating or gambling, for example, and somatization disorders in which emotions are converted into physical symptoms can all be understood in this way. Some forms of borderline personality disorder and conversion disorder may also reflect complex adaptation to early trauma, as does dissociative identity disorder.

Cognition in Racial Prejudice

The modern notion of the brain as a human computer is both alluring and misleading. Unlike computers, human brains have been programmed over millions of years to process the environment in ways that enhance human survival. Adaptations which enhance survival trump accuracy nearly every time. In Ortega y Gasset's words, "*We do not live to think; we think in order that we may succeed in surviving*." Because of this evolutionary strategy, cognition is as much distortion as it is representation. One of the consequences of this distortion is racial prejudice.

The ubiquity of prejudice and racism seems to suggest that humans are

hardwired to dislike, distrust, and even hate those who are different. The fact that conflicts and slaughter based on race, tribal affiliations, and religion continue unabated in the "modern" world only seems to strengthen the argument for prejudice as a basic component of human nature. Standing in stark contrast to this view are countless examples of harmony, cooperation, and acts of altruism that seem to contradict racism as a basic human truth.

Imagining the lives of man's tribal ancestors, it is easy to see the survival benefit of developing a reflexive system for classification of unfamiliar others that triggers fear circuitry and a fight–flight response. It would then follow that these primitive reactions would lead one to distortions of sensory perceptions and social judgments that parallel those seen in social phobia. These distorted attitudes and perceptions would then become woven into culture, guiding everything from interpersonal behavior, to public policy, to international relations. Therefore, the most relevant questions facing psychiatrists in this area of study are how do human brains process social information and how can one counteract those perceptions that are incorrect and harmful to others?

In an attempt to identify the neural substrate of prejudice, a number of studies have examined brain activation during exposure to pictures of same/other group faces. It turns out that the old statement "they all look alike," where "they" are another racial group, may reflect an underlying truth of cognitive processing. Across studies, subjects demonstrate a clear advantage in recognizing faces and distinguishing among faces for same-race faces. Neural processing for the recognition of other race faces results in longer periods in frontal activation, suggesting that they are processed less automatic and with greater effort.

Supporting this idea is that fact that interracial interactions have been those that show impaired executive functions when confronted with people of other races. It is possible that the extra frontal lobe activation required for facial processing may divert energies and attention away from other frontal functions. Consistent with these findings are those that show that stereotypes are automatically activated and require conscious inhibition to produce low-prejudice responses. The overall result of these processing differences may lead one to think about individuals of other races in less complicated and more stereotyped ways.

Assuming that fear is at the heart of prejudice, a second line of research directly explored amygdala activation in same versus other race encounters. It was found that when Caucasians and African Americans are shown pictures of other-race faces, they exhibit more amygdala activation and an enhanced startle response compared to same-race faces. Although it takes approximately 500 ms for a visual stimulus to reach conscious awareness, information about race is encoded during the earliest stages of visual processing, well before conscious perceptions of who is being seen.

The rapidity of this “race response” results in the automatic and instantaneous frontloading of all past biases into the experience of each new person.

This activation of a rapid and unconscious fear response influences the conscious experience in many ways. A finding that directly parallels the experience of social phobics. For example, individuals from other groups are more readily associated with aversive stimuli. This is the likely reason why African Americans are more readily misperceived as carrying weapons and black faces are more likely to be perceived as angry by white subjects. Interestingly, both African Americans and Caucasians have greater amygdala activation to African-American faces, which may be due to the consistent pairing of African-American faces with criminal behavior in the media. This conditioning may not have as extreme of an impact on African-American viewers because of their greater exposure to other African-American faces. African Americans are also better at recognizing Caucasian faces than vice versa for this same reason. Thus, while one’s responses to people of other races are guided by bottom-up amygdala activation, they are also modulated by a variety of top-down cognitive and social factors.

Greater liberalism has been associated with greater gray matter volume in the anterior cingulate while greater conservatism is associated with increased volume in the right amygdala. Conservatives demonstrate more amygdala activation and persistent cognitive styles compared to liberals who demonstrate more sensitivity for cues to alternate responses. Along with these differences in cognitive style, greater liberalism was associated with stronger conflict-related anterior cingulate activity reflecting their continued engagement with the complexity of the problems posed to them.

To explore the issue of whether prejudice is learned or innate, faces of well-known and well-liked African Americans were shown to Caucasian subjects who demonstrated an absence of amygdala activation. These findings strongly suggest that exposure and positive experiences can counteract the effects of general cultural prejudice. Even simple exposure to other-race faces increases liking for strangers of that race, and an experience of empathy for one member generalizes to other members of that group. It is also known that through exposure and paying conscious attention to facial details, Caucasian adults can be successfully trained to distinguish among African-American faces and that this learning process results in a decrease of implicit racial bias.

Once stereotypical attitudes and beliefs are inculcated, change requires “intention, attention, and time.” Because most children are exposed to some degree of racial prejudice, open discussion and increased interracial exposure work against prejudice turning into conscious beliefs and negative behaviors. Increased focus and awareness of the handicapped has also found to decrease discrimination in the same way. It is known from studies of anxiety and fear that conscious labeling activates cortical neural

networks which can regulate and inhibit emotional activity. Verbal encoding of African-American targets results in less amygdala activation in Caucasians, reflecting the ability of conscious processing to inhibit fear.

People from other cultures, with different degrees of exposure, proximity, and prejudice need to be studied with the same depth that one has studied their own society before one can arrive at any conclusions about the origins of racial prejudice in the human brain (see Diamond, 2013 for a revealing description of cultural change). With the science of genetics revealing that the notion of race is more of a cultural than a biological construct, future study of how human brains react to one another can only help shed light not just on the nature and etiology of the neuroscience of prejudice, but on how everyone who lives on this increasingly crowded planet may learn to overcome their differences and discover the true wealth of human diversity.

Cognition in Social Phobia

It is clear that those that struggle with social phobia do not perceive the same interpersonal landscape as others. Social phobia disrupts and distorts the ability to gather and process social information with attentional and memory processes biased toward detecting, orienting toward, exaggerating, and remembering negative responses by others. Once focused on negative input, they fixate on it and find it difficult to disengage, even when it is irrelevant or harmful to their well-being.

Social phobia is most likely the result of an imbalance of the alarm and feedback systems that regulate the balance of approach and avoidance in social situations. At the same time, social phobics show decreased activation in left-hemisphere regions involved in other aspects of facial evaluation and reality testing. In essence, their brains react to others as if they are in extreme danger no matter what the situation. Fear of direct eye gaze is an evolutionary conserved submissive behavior that is woven into all human cultures but exaggerated in those with social phobia. Social phobics employ visual avoidance and elaborate processing strategies to modulate the fear evoked by being seen.

Children with social phobia experience greater anxiety while looking at faces and as adults, their memory for faces, especially those with negative expressions, is far better than average. It makes sense, then, that social phobics avoid looking directly at faces, especially eyes, preferring to focus on other aspects of those with whom they interact. Turning away from others has a number of negative consequences. First, it keeps them from viewing potentially positive expressions. Second, it keeps these children from learning to process facial expressions in a more sophisticated manner, especially during development. Finally, turning away from others correlates with using their negative self-image when developing a theory of what is on another's mind. This may be why social phobics have been found

to come to conclusions faster in social situations, recycling old information is faster than acquiring new data.

In social phobia, the amygdala shows decreased coupling with the frontal lobes and higher sustained activation in face processing networks. The combination of increased orientation to and fixation on social evaluation makes the experience of the social phobic one of constantly being on trial: Every interaction is experienced as a negative evaluation of their worth. Thus, the memories and emotional impact of negative interpersonal responses cut deeper and last longer.

The most consistent findings have tied social phobia to hyperactivation of the amygdala. Subjects with social phobia show increased metabolism in the amygdala, fusiform gyrus, insula, anterior cingulate, and the prefrontal cortex when subject to direct eye gaze. Simultaneous, they demonstrate decreased cortical activation that could inhibit and regulate their subcortically generated fear.

Decoupling of cortical and subcortical processing is likely a major contributor to the disruption of current social information processing and the tendency to become overly reliant on past experience and negative internal self-images in interpreting the present. One group of researchers found that a combination of right prefrontal and left parietal activation was specific to social phobia. This led them to hypothesizing that social phobia is related to a unique interaction between the processing of autobiographical memory and social presentation. Perhaps such neural wiring makes it difficult or impossible to overcome self-consciousness and the expectation of negative evaluation about the self.

More recent research has revealed a deactivation of the precuneus and posterior cingulate in these individuals suggestive of an inhibition of the default mode processing that is used to process information about oneself and others. Mindfulness- and cognitive-based interventions have proven helpful with social phobia because they simultaneously result in decreased amygdala activation and increased activation in brain regions involved in the deployment of attention.

While individuals who suffer with social phobia share the neurochemical profiles of other anxiety disorders (heightened adrenergic activity combined with decreased GABA activity), they appear to also have evidence of dysfunction in serotonin and dopamine systems. Hypersensitivity to negative social stimuli appears to be linked to variation in serotonin receptor pathways. Lower dopamine (D2) binding has been found in the striatum of social phobics, which in animals correlates with a lack of exploration and lower social status. Remember, dopamine is an important neurochemical in the social reward system related to approaching others, sustaining attachment, and a sense of well-being. As with so many other findings, it is not known if biological differences are a cause or effect of social phobia.

Social phobia could reflect an innate inability to compete socially or an experience-dependent shutdown of activity in the face of anxiety and shame. When parents dominate children through physical or psychological punishment, some young brains may be shaped to be of persistently lower social rank. On the other hand, social phobia may be an evolutionary continuation of the behavior of lower-ranking individuals in mammal and primate groups who "know their place" and stay in it.

FUTURE DIRECTIONS

Clinicians must keep in mind that every client strives to make sense of the world, albeit through information-processing apparatus that are negatively impacted by the acute and chronic aspects of their disorder. Cognitive science offers an array of conceptualizations for understanding the ways in which the mind functions in health and disease. This broad, interdisciplinary field provides numerous research paradigms that are helpful in further elucidating the nature of psychopathology, from the techniques of the neurosciences to the biological applications of chaos theory. It also provides ways of thinking that can be helpful in the consulting room in understanding the cognitive, emotional, and social impact of the ways in which the brains of maladaptive information processing. The cognitive understanding of emotions and consciousness may also expand psychiatry's framework for knowing about human subjective experience. Clinical tools, from medications to in-depth psychotherapy, may also find wider and more finely tuned applications as the processes of self-organization and psychological change are better understood.

Psychiatry, in turn, also has much to offer the field of cognitive science. The long history of descriptive psychopathology and the attempt to synthesize views of the mind and the brain can provide nonclinical cognitive scientists with unique data and relevant questions. This may be especially true in the understanding of the social nature of the brain and mind and the impact of both positive and negative relationships on mental health and illness. Psychiatry is invited to join in the search for understanding the cognitive processes of the human mind.

REFERENCES

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Publishing; 2013.

Baddeley A. Working memory: Looking back and looking forward. *Nature*. 2003;4:829.

Cacioppo J, Visser P, Pickett C, eds. *Social Neuroscience: People Thinking About Thinking People*. Cambridge, MA: MIT Press; 2006.

*Cozolino LC. *The Neuroscience of Psychotherapy: Building and Rebuilding the Human Brain*. 2nd ed. New York: Norton; 2010.

*Cozolino LC. *The Neuroscience of Human Relationships: Attachment and the Developing Social Brain*. 2nd ed. New York: Norton; 2014.

Cozolino, LC. *Why Therapy Works: Using Our Mind to Change Our Brains*. New York: WW Norton; 2015.

Damasio AR. *Self Comes to Mind: Constructing the Conscious Brain*. New York: Vintage Books; 2010.

Davidson RJ, Begley S. *The emotional life of your brain*. London: Hodder & Stoughton; 2012.

Diamond J. *The World Until Yesterday: What Can We Learn From Traditional Societies?* New York: Penguin Books; 2013.

Epstein R. *Attending: Medicine, Mindfulness, and Humanity*. New York: Scribner; 2017.

Harmon-Jones E, Winkielman P, eds. *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior*. New York: Guilford Press; 2007.

Kandel ER. A new intellectual framework for psychiatry. *Am J Psychiatry*. 1998;155:457.

Lane R, Nadel L. *Cognitive Neuroscience of Emotion*. New York: Oxford University Press; 2000.

Main M. Metacognitive knowledge, metacognitive monitoring, and singular (coherent) vs. multiple (incoherent) models of attachment: Findings and directions for future research. In: Marris P, Stevenson-Hinde J, Parkes C, eds. *Attachment Across the Life Cycle*. New York: Routledge & Kegan Paul; 1991.

McGilchrist I. *The Master and His Emissary: The Divided Brain and the Making of the Western World*. New Haven: Yale University Press; 2009.

Mesulam MM. Review article: From sensation to cognition. *Brain*. 1998;121:1013.

*Metcalfe J, Shimamura AP. *Metacognition: Knowing About Knowing*. Cambridge, MA: MIT Press; 1994.

Mezquita B, Feldman Barret L, Smith ER. *The Mind in Context*. New York: Guilford Press; 2010.

Poeppel D, Monahan PJ. Speech perception: Cognitive foundations and cortical implementation. *Curr Dir Psychol Sci*. 2008;17(2):80.

*Posner MI, ed. *Foundations of Cognitive Science*. Cambridge, MA: MIT Press; 1989.

*Schore AN. *Affect Dysregulation and the Damage to the Self*. New York: Norton; 2002.

*Siegel DJ. *Mindsight: The New Science of Personal Transformation*. New York: Random House; 2010.

Siegel, DJ. *The Mindful Therapist: A Clinician's Guide to Mindsight and Neural Integration*. New York: Norton; 2010.

*Siegel DJ. *The Developing Mind: How Relationships and the Brain Interact to Shape Who We Are*. 2nd ed. New York: Guilford Press; 2012.

*Siegel DJ. *Pocket Guide to Interpersonal Neurobiology: An Integrative Handbook of the Mind*. New York: Norton; 2012.

Siegel DJ. *Mind: A Journey to the Heart of Being Human*. New York: WW Norton; 2017.

Taber K, Rausch S, Lanius R, Hurley R. Functional magnetic resonance imaging: Application to posttraumatic stress disorder. *J Neuropsychiatry Clin Neurosci*. 2003;15:125.

Van der Hart O, Nijenhuis ERS, Steele K. *The Haunted Self: Structural Dissociation and the Treatment of Chronic Traumatization*. New York: Norton; 2006.

van der Kolk B. *The Body Keeps the Score: Brain, Mind, and Body in the Healing of Trauma*.

New York: Viking. 2014.
Vercammen A, de Haan EH, Aleman A. Hearing a voice in the noise: Auditory hallucinations and speech perception. *Psychol Med*. 2008;38(8):1177.
Wilson, EO. *The Social Conquest of Earth*. New York: Liveright Publishing; 2013.

▲ 3.2 Piaget and Cognitive Development: History, Theory, Extensions and Modifications

STANLEY I. GREENSPAN, M.D., AND JOHN F. CURRY, PH.D.

Jean Piaget (1896 to 1980) was, at various points in his career, considered to be a natural scientist, psychologist, sociologist, and historian of the philosophy of science. His major enterprise was what he termed "genetic epistemology," the study of the development of logic, reasoning, and higher-level thinking. He had points of contact with European psychiatry and even based his own research method in part upon the psychiatric interview. However, by and large he avoided the emphasis on individual differences or interventional psychology, as practiced by clinical diagnosticians and therapists, and focused instead on the most universal and general characteristics of children's cognitive development. His influence on education was considerable, not through a specific teaching method, but through his view of the child as an active constructor of his or her own knowledge, rather than a passive recipient of information.

Piaget was born in 1896 in Neuchâtel, Switzerland. Growing up he was influenced by social groups and unique opportunities that supported his interest in natural science, including the Swiss YMCA, the local Friends of Nature, and the director of the natural history museum, who allowed him to work with the specimens of mollusks, the topic of his 1918 doctoral thesis. For a year he then worked in Zurich combining laboratory research with clinical psychiatric interviewing supervised by Bleuler. He then moved to Paris to study at the Sorbonne and was recruited by Theodore Simon to work in Alfred Binet's school-based laboratory, standardizing a reasoning test developed by Cyril Burt. In 1921 he began a long association with the Jean Jacques Rousseau Institute, with parallel appointments at varying times at the Universities of Geneva, Lausanne, and the Sorbonne. In 1955, with support from the Rockefeller Foundation, he founded the Center for Genetic Epistemology where he continued his research and teaching with European and international colleagues, including prominent American psychologists.

Bennour and Voneche note several of the major intellectual and personal influences on Piaget. His interest in evolution stemmed from exposure to Henri Bergson's work, whereas his interest in psychiatry, psychology, and psychoanalysis stemmed from multiple sources. As

mentioned above, Bleuler supervised him in clinical interviewing; Jung and others gave psychoanalytic lectures that he attended, Simon mentored him in standardized psychometric assessment and influenced his view of cognition as adaptation rooted in biology. Janet was a strong influence on Piaget's thought, in terms of his emphasis on psychology as a science of action and representation, and affect as a driver of both action and thought.

Piaget's contacts with American psychology and social science are interesting in light of the prominence of experimental methodology and of behaviorism in the United States. In 1936 he was awarded an honorary doctorate from Harvard, but Hsueh has documented that the impetus for this came from business school and sociology faculty rather than from the psychology faculty. The nonpsychology social scientists were impressed with Piaget's method of close observation and detailed "clinical" note-taking, and his emphasis on cooperation in problem-solving. In 1969, however, Piaget became the first European to receive the American Psychological Association's Distinguished Scientific Contribution Award. The award noted his creation of an empirical science of epistemology, a field previously restricted to philosophy, as well as both his observations and his controlled data on human cognition, which by then was a major influence on all scientists concerned with cognitive development.

Piaget's Clinical Method

Susan Mayer has shown that Piaget's "Clinical Method," which had brought him to the attention of Harvard social scientists relatively early in his career, was a synthesis of three sources: naturalistic observation, psychometric evaluation, and psychiatric interviewing. Naturalistic observation was practiced by Piaget even as a youth, with his biological studies. He continued this approach when he began to be trained in the psychological testing of children. Under Simon, Piaget learned the psychometric approach, which requires standardized presentation of specific probes, tasks, and questions. Strictly speaking, psychometric assessment depends on quantitative comparisons of a given child with an adequate normative sample, and on intrasubject comparisons across tasks. It requires little interest in the reasons why children make cognitive errors.

Piaget, however, brought to bear his earlier learning of psychiatric interviewing to follow up erroneous child responses to psychometric items. This served the purpose of identifying forms of thinking that lay between clearly faulty (immature) and fully logical responses. Careful attention to children's language, explanations, and the social functions of their utterances was a further contribution of naturalistic observation to the overall clinical method.

Later in his career, as the clinical method was criticized, Piaget and Inhelder added more experimentally controlled methods to their studies of logical reasoning. However, they did not abandon the observational,

psychometric, and individualized probing aspects of the clinical method.

GENETIC EPISTEMOLOGY

Piaget's overarching task was the understanding of how one comes to know things. He took a developmental approach to this question, inquiring how knowledge becomes more adequate over time. He designated himself a genetic epistemologist—one who studies the development of knowledge from infancy to adolescence. By "genetic," he did not mean an approach based on genes, but rather an approach based on development over time. It remains important to distinguish Piaget's approach from that of nativism, on one hand, and empiricism, on the other. Nativism is a theory proposing that at least certain kinds of knowledge are innate or given. In philosophy, Kant took a nativist position regarding the categories of space, time, and causality, arguing that these categories on the data of sensory perception be imposed. In psychology, Gesell's theory was essentially based on a nativist position. According to this perspective, certain types of thinking and behavior are expected at certain ages as development unfolds in a predetermined manner.

Empiricism, at the opposite end of the spectrum, proposes that all knowledge is accumulated from the data of sensory perception. In psychology empiricism is equated with behaviorism and learning theory. Little attention is paid to predetermined (innate) factors, and the emphasis is on common processes of learning, such as classical and operant conditioning, that lead to accumulated knowledge.

Piaget's approach was, in his words, "constructivist structuralism," according to which mental structures are built through the interaction of the child and the world. Mental structures originate through the actions of the child on objects as the child strives to adapt to the environment. In the educational realm this implies that guided discovery is preferable to memorization. Only through action and interaction can the child develop a comprehensive and stable structure that enables understanding. What Piaget deemed innate was only an intelligent functioning that makes possible the construction of progressively more adequate structures of knowledge. This improvement of structures is based on abstractions from actions performed over time, and qualitative improvements mark progressive stages of development. Action takes precedence over language in Piaget's model. Abstractions from repeated actions lead to the construction of robust and adequate structures. For example, the concept of space is a fundamental mental structure developed in the earliest period of children's lives. In earliest infancy, the child is aware not of one homogeneous space, but of several heterogeneous spaces, each centered on a certain part of the child's body (e.g., visual space and tactile space). As the child acts on objects that may traverse these various spaces (e.g., a rattle occupying visual, tactile, and auditory spaces), he or she comes to

coordinate these individual spaces. Eventually, actions representing displacements in space are organized mentally into the general concept of space. That concept is a structure.

PIAGET'S THEORY

Piaget provided a succinct summary of his work in a chapter prepared for the 1970 edition of *Carmichael's Manual of Child Psychology*. The interested reader is referred to that source, which was reprinted in the 1983 edition of the *Handbook of Child Psychology*. The summary is based on this and other primary and secondary sources. Piaget's theory is broad and comprehensive with regard to cognitive development. It is analogous in this sense to Sigmund Freud's or Erik Erikson's theories of emotional or personality development. Like those comprehensive theories, it now seems to belong to an earlier era as numerous specific predictions have been modified or falsified and as a great deal of more detailed research has accumulated regarding children's cognitive capacities and processes. This topic will be discussed again after presenting the overarching theory.

Equilibration

For Piaget, the general criterion for intelligent functioning is *equilibration,* briefly defined as "a compensation for an external disturbance." Hans Furth described equilibration as "the factor that internally structures the developing intelligence. It provides the self-regulation by which intelligence develops in adapting to external and internal changes." At every level of development, the equilibration mechanism operates to further adaptation, but, as development proceeds toward the highest level of cognitive functioning, equilibration becomes progressively more adequate in enabling the organism to adapt to a wider range of internal and external disturbances. Piaget's notion of intelligence as adaptation is therefore essentially bound to an equilibration model of intelligent functioning. Equilibration can be thought of as a dynamic process of give and take as the organism strives to "take in" new and challenging information and "adjust" to this new input.

Equilibration is the final of four necessary components that lead to cognitive development. Piaget recognized the importance of biological maturation, and more specifically the maturation of the CNS, as essential for cognitive development. He also incorporated learning experiences, including learning based on conditioning, as a necessary component. The third essential component in his theory is social interaction. Not surprisingly, this component appears to be particularly essential for social intelligence. Thus, for Piaget, biological maturation, learning, and social interaction flow together to facilitate development and crystallize in the process of equilibration as the developing child reacts to cognitive

challenges generated by new input.

Assimilation and Accommodation

The biological foundation of Piaget's developmental theory is clearly evident in his discussion of assimilation and accommodation; these processes are considered functional invariants of all intelligent behavior. At every level of intellectual development, from infancy to adulthood, the processes operate in adaptation.

The assimilation–accommodation account of development stresses the interaction between organism and environment. A certain readiness within the organism is necessary for change or development to occur. In Piaget's view, associationism (empiricism) in psychology commits the fallacy of crediting only one-half of those conditions necessary for learning with all of the explanatory power. A full account of human development must include not only the influence of stimuli on respondents (S → R), but also the influence of the responding organism on incoming stimuli (S ← R). Such an account is provided by Piaget's assimilation–accommodation viewpoint: "From a biological point of view, assimilation is the integration of external elements into evolving or completed structures of an organism. In its usual connotation, the assimilation of food consists of a chemical transformation that incorporates it into the substance of the organism."

Furth referred to assimilation as "an inward-directed tendency of a structure to draw environmental events towards itself." Assimilation is the conservative side of intellectual development, ensuring continuity and coherence by incorporating new aliments (stimuli) into existing mental structures. However, assimilation alone cannot account for growth or change within those structures.

Accommodation occurs during the developmental periods when new data cannot be wholly assimilated into the child's existing mental structures and yet the data are not so foreign to those structures that they can be ignored. Furth referred to accommodation as "an organism-outward tendency of the inner structure to adapt itself to a particular environmental event." This line of thought reveals the sense in which Piaget defined *intelligence* in terms of equilibration. His equilibration is not a static, balanced system, but instead is a dynamic, or mobile, equilibration between assimilation and accommodation as the child responds to the environment.

Structuralism

The term Piaget used for a cognitive structure is *scheme* (*schema*): "A scheme is the structure or organization of actions as they are transferred or generalized by repetition in similar or analogous circumstances." Schemata exist in the infant as perceptual-motor behavior patterns (e.g., the grasping

reflex). They also exist in mature intelligence, although, as Furth pointed out, *schema* is more commonly used to refer to an early mental structure, whereas general schemata resulting from the use of higher intelligence are referred to as *operations*.

The abstraction process that leads to the formation of cognitive structures is called *reflective* or *formal abstraction*. It is an abstraction from *actions,* according to which the similarities inherent in various behavioral acts are dissociated from their particularized contexts (e.g., the earlier example of the rattle in space). These can be contrasted with *simple abstractions,* which are more familiar to psychologists and psychiatrists. Simple abstractions are similarities between or among several things. For instance, when learning colors, the child will note that "green" refers to numerous objects, all of which have the same color; when learning shapes, the child will become able to discriminate circles from squares. Although such abstractions are obviously critical for learning, they do not occur at the same depth of cognitive development as the reflective abstractions. The latter contribute not to the content of knowledge but to its progressively more adequate structure.

Theory of Stages

An integral part of Piaget's theory of genetic epistemology is a psychology of cognition that seeks to describe how knowledge develops and changes. The genetic framework for that and the process of intellectual adaptation during the major early periods of life is provided in Piaget's theory of the stages of cognitive development.

The stages of cognitive development that Piaget and his associates delineated empirically are not defined merely by the dominance of some aspect that remains present but less dominant throughout development.

Piaget was not entirely consistent concerning his stages of cognitive development, but the primary source of confusion in his later writings is whether the so-called preoperational period is to be considered apart from the period of concrete operations in which it culminates. John Flavell's 1963 study and Piaget's 1983 summary defined three major periods and one subperiod of intellectual development (Table 3.2–1). These periods contain subdivisions called *stages*. The major developmental periods are as follows:

1. The sensorimotor period, which extends from birth until approximately 1.5 years of age. The period is divided into six stages, which are described in general in the following discussion with reference to the development of the concept of the permanent object.
2. A period of preparation for, and acquisition of, concrete operations. This period extends from the appearance (at approximately 2 years of age) of the symbolic (semiotic) function to the beginning (at approximately 7

years of age) of higher mental operations applied to concrete objects.

3. The period of formal operations, which begins at approximately 11 years of age. During this period, full adult intelligence develops as the operations are extended to apply to propositional, or hypothetical, thinking.

Sensorimotor Period. The *sensorimotor period* of intelligence is so named because the child's construction of mental schemata is in no way aided by representations, symbols, or thoughts. Rather, schemata depend totally on perceptions and bodily movements.

Table 3.2–1.
Stages of Intellectual Development Postulated by Piaget

Age (Years)	Period	Cognitive Developmental Characteristics
0–1.5 (to 2)	Sensorimotor	Divided into six stages, characterized by: 1. Inborn motor and sensory reflexes 2. Primary circular reaction 3. Secondary circular reaction 4. Use of familiar means to obtain ends 5. Tertiary circular reaction and discovery through active experimentation 6. Insight and object permanence
2–7	Preoperations subperiod[a]	Deferred imitation, symbolic play, graphic imagery (drawing), mental imagery, and language
7–11	Concrete operations	Conservation of quantity, weight, volume, length, and time based on reversibility by inversion or reciprocity; operations; class inclusion and seriation
11 through the end of adolescence	Formal operations	Combinatorial system, whereby variables are isolated and all possible combinations are examined; hypothetical-deductive thinking

[a]This subperiod is considered by some authors to be a separate developmental period.

Stage 1 of sensorimotor development is marked by a relatively few organized reflexes that stand out from the spontaneous general activity of the neonate, such as the sucking reflex and the palmar reflex. These primitive reflexes demonstrate three types of assimilation: (1) reproductive (repeating the actions), (2) generalizing (repeating the actions on new objects), and (3) recognitory (performing different varieties of the actions on different objects).

Stage 2 contains the first habits and the primary circular reactions. The first habits develop out of the original schemata as they are applied to objects in the environment or parts of the infant's body without any differentiation between means and ends. In a primitive state of consciousness, the infant is aware only of action sequences and is not even aware of self. Primary circular reactions occur when, by chance, the infant

experiences a new consequence of a motor act and tries to repeat the act.

In stage 3, an initial distinction between means and ends becomes apparent, but in a primitive sense. The infant repeats a particular action pattern that achieved one end for the purpose of achieving other (unrelated) ends. For example, a baby who succeeds in shaking a rattle by pulling a string may repeatedly pull the string in an attempt to bring about other sounds or results.

In stages 4 and 5, infants use a variety of means to obtain particular goals. The distinction between stages 4 and 5 lies in the relative creativity or newness of the means. Stage 4 is marked by the use of familiar means. Stage 5 is marked by a search for new means based on further differentiations of already known schemata and by tertiary circular reactions. The latter differ from secondary circular actions in that the child no longer produces schemata that were effective in one situation to produce magically efficacious results in every situation. Instead, the child explores the environment and varies means to test for effectiveness. Discovery is a hallmark of stage 5. For example, a child may use a stick to move an object that is not within reach.

Stage 6 is transitional, leading into the preoperational subperiod. In stage 6, the child becomes capable of inventing new means, not by direct actions on objects, but by mental combination. Whereas discovery marked stage 5, insight is a characteristic of stage 6. For example, a child who has seen the father bang on a drawer to loosen it may bang on a toy box to make it easier to open.

During the sensorimotor period, a number of significant concepts are developed, including the child's concepts of space, time, and causality. These categorical concepts develop in a process parallel to the sequence of the six stages outlined in the previous discussion. Most important, during the sensorimotor phase, the child develops the schema of object permanence, the first major victory of conservation and the foundation of all future knowledge.

SCHEMA OF OBJECT PERMANENCE. The knowledge that objects in the external world have an existence independent of the child's actions on them or interactions with them is a major accomplishment of the sensorimotor period. Piaget described his observations of infants' reactions to the disappearance of interesting objects, the foundation for his theory of development of object permanence. In stages 1 and 2, for example, a child simply continues to look at the place where the object was last seen. In stage 3, if an object such as a spoon drops to the floor, the infant will look for it (e.g., by leaning over and looking at the floor). In stage 4, if an object is repeatedly hidden at point A (in sight of the child) and then hidden at point B (also in sight of the child), the child searches for it at A, not B. In stages 5 and 6, the infant is able to follow multiple displacements of the

object through points in space, even if the object is hidden within another object.

Preoperational Subperiod and Semiotic Function. The advent of the preoperational subperiod is marked by the appearance of what Piaget called the *semiotic function*. This new ability was defined by Piaget and Inhelder as follows: "It consists in the ability to represent something (a signified object, event, conceptual scheme, etc.) by means of a signifier which is differentiated and which serves only a representative purpose: Language, mental image, symbolic gesture and so on."

During the sensorimotor period, a thing could be represented in a limited sense by a part of itself (e.g., the mother's voice might represent the presence of the mother in the room). However, such signifiers are part of what they signify. Symbols and signs are signifiers that are differentiated from what they signify. A drawing of a person, for example, is not part of the person. A word is different from the object to which it refers. These become available to the child only with the appearance of the semiotic function, which makes representational thought possible.

CHARACTERISTIC BEHAVIOR PATTERNS. The semiotic function is heralded by five characteristic behavior patterns in evidence during the second year of life: (1) deferred imitation (imitation that starts after the disappearance of the model); (2) symbolic play, or the game of pretending; (3) drawing, or the use of graphic imagery; (4) the presence of a mental image, which appears as an internalized imitation and not as a function of perception; and (5) the verbal evocation of events not occurring at the time. For Piaget, the semiotic function, which so enlarges children's worlds—liberating them from the bonds of immediate space and time and enabling them to begin to manipulate symbols and to think rather than just to act on immediately present objects—is rooted in imitation.

Imitation. One can follow the development of imitation through the same six sensorimotor stages delineated for the concept of object permanence. Piaget did this in his volume *Play, Dreams and Imitation in Childhood*. A radically new form of imitation occurs during the second year of life: deferred imitation. For example, a child may put on father's hat and walk as father does, even hours after father has gone off to work.

For Piaget, intelligence is an equilibration process in which assimilation and accommodation are in balance. However, in imitation, accommodation outweighs assimilation. According to Piaget, imitation is behavior in which "the subject's schemes of action are modified by the external world without his utilizing this external world." In imitation, the child's cognitive structures undergo temporary change without simultaneously incorporating new aliment.

Symbolic Play. A second new behavior pattern that now appears is symbolic play. In imitation, the imbalance between assimilation and accommodation is weighted in favor of accommodation; however, the opposite is true in symbolic play, which is a lessening of the demand of the adaptive process.

Symbols are used in play only after the sensorimotor period, in the type of play characterized by games of pretending. For example, a little girl may pretend that she is asleep, that a box is her pet cat, or that she is a church. In each instance, symbols are generated "to express everything in the child's life experience that cannot be formulated and assimilated by means of language alone."

According to Piaget, these symbols are created by the same process of imitation that gives rise to deferred imitation at this time (a process in which accommodation outweighs assimilation). Instead of being used accurately (realistically), however, they are used in a process in which a liberating assimilation outweighs accommodation.

Drawing. A third behavior pattern associated with the rise of the semiotic function is drawing, or the use of graphic imagery. Piaget sees elements of play and imitation in this activity. In developmental terms, he considers drawing "halfway between symbolic play and the mental image," appearing at approximately 2 or 2.5 years of age. The child enjoys producing drawings for their own sake (assimilation). However, graphic play also has accommodative elements, especially as the child grows older and attempts to draw not just formless scribble but something.

Mental Image. Closely related to drawing is the presence of a mental image. Piaget tied the genesis of mental imagery to accommodation and imitation. He explicitly denied that mental images can be the product of perception; they are a construction, something the child creates. The mental image is not directly given by perceptual input; it is constructed by the process of accommodation.

Verbal Evocation of Events. The fifth behavior pattern associated with the rise of the semiotic function concerns language, the verbal evocation of events that are not present. Piaget gave the example of a little girl saying "Anpa, bye-bye" (Grandpa went away) while pointing to the path he had taken when he left. The parallel with deferred imitation is clear, but the new representational ability is supported by the social system of language.

Concrete Operations. A crucial difference between preoperational and concrete-operational thought is the presence within operative thinking of concepts of conservation. When concrete operations have been organized into a system, the child can conserve, that is, "discover what values do remain invariant ... in the course of any given kind of change or transformation." The progressive and continual structure building that

occurs in the concrete-operational period is evident in the increase, with development and age, in the scope of such concepts, such as conservation of quantity, substance, and number. In other words, conservation indicates stable cognitive structure.

CONSERVATION OF QUANTITY. If liquid is poured from a short, wide glass into a tall, narrow one, the child in the preoperational stage thinks that the amount of liquid has changed. At the level of concrete operations, however, children are no longer overwhelmed by the perceptual discrepancy between the two configurations. They begin to reason about the transformation, and their correct judgments regarding the conservation of quantity of liquid are accompanied by explanations grounded in logical properties. It is assumed that children are not aware of the logic they use. When problems of conservation begin to be solved, the child passes from the preoperational subperiod into the period of concrete operations.

CONSERVATION OF WEIGHT AND VOLUME. At approximately 7 or 8 years of age, the child can solve the conservation-of-quantity problem and can perform similar judgments of conservation when, for example, a lump of clay is transformed in shape. Between 9 and 10 years of age, the child discovers that the weight of a given object is also conserved, even when its shape is transformed. However, not until approximately 11 or 12 years of age do children have a logical comprehension that the volume displaced by a given object is conserved even after a transformation of the object's shape. Conservation entails the logical certainty that one characteristic of an object remains invariant while the object itself undergoes some type of perceived transformation.

CARDINAL NUMBERS. The concept of cardinal numbers also develops from an initially nonconserving to a conserving stage. For example, consider children in the preoperational subperiod presented with two horizontal rows of colored dots in one-to-one correspondence (i.e., imaginary vertical lines could be constructed between each red dot and its corresponding blue dot). When the experimenter destroys this optical correspondence by spreading out one of the rows of dots, the child in the preoperational period thinks that the larger row contains more dots. Only after conservation of cardinal number has been established as a logical necessity does the child maintain the numerical equivalence of the spread-out row to the other row. Clearly, preoperational concepts of number provide inadequate bases for arithmetic skills. It is possible that a lag in the development of number conservation could underlie certain types of arithmetic-related learning disabilities.

Operations. Notions of conservation are the mark of well-established concrete-operational thinking; thus, one must understand the meaning of

operation in Piaget's thought. Operations, themselves, constitute essential thinking. For Piaget, an operation is an action that is (1) interiorized, (2) reversible, and (3) part of an organized system of such actions.

The operations that form this system are interiorized actions. In the sensorimotor period, external behavior patterns give rise through a process of abstraction to the construction of sensorimotor schemata. In a similar fashion, internal thinking patterns later give rise to operations. According to Furth, the possibly generalized aspects of actions, "those which can be found in any coordination of action," enter into the construction of operations. Saying that the crucial aspect of actions in this regard is their ability to be generalized explains the importance of interiorization in the construction of operations. *Interiorization* refers to "the increasing dissociation of general form from particular content." In other words, the notions of generalization and interiorization merely point out the process of abstraction that is occurring. For example, a child adds two apples and three apples to obtain five apples. In another instance, a child adds seven blocks and one block to obtain eight blocks. In a third instance, a child combines the category of fathers with that of mothers to obtain the category of parents. The operation abstracted from these three mental actions is addition or combining, without reference to the particular content of numbers, objects, or categories.

Not only must an operation be interiorized action, it must also be reversible. The action of combining (addition) is not an operation until its relationship to the action of separating (subtraction) is comprehended. To understand reversibility is to understand the third criterion of an operation, its inclusion in a system.

The reversibility essential to operatory thought may be inversion or reciprocity. In reversibility by inversion, an action +A is reversed by −A. For example, in the conservation-of-quantity example, pouring liquid into container 2 (+A) may be mentally reversed, that is, by mentally pouring it back into container 1 (−A). In reversibility by reciprocity, a relation A < B is reversed by a relation B < A. Referring again to the conservation-of-quantity example, let A stand for container 1 and B stand for container 2. The rising height of liquid in container 2 (A < B) is offset by its narrower width (B < A).

Corresponding to these two types of reversibility are the two major categories of concrete operations: those pertaining to classes and those pertaining to relations. In the system of operations performed on classes, reversibility is by inversion; in those performed on relations, it is by reciprocity. For example, subtraction and addition relate to inversion; comparing sticks of different sizes relates to reciprocity.

CLASS INCLUSION. The concrete operation demonstrating an understanding of classes is the class inclusion task. In this task, a child is

shown, for example, an array of pets (superordinate class) consisting of dogs and cats (subordinate classes). After counting the number of dogs, cats, and pets, the child is asked whether there are more dogs or more pets. Children in the preoperational subperiod cannot keep in mind the superordinate class while perceiving only the subordinate classes. Thus, they fail the task frequently over a series of such arrays. The concrete operation of class inclusion underlies one's ability to think categorically.

RELATIONS. The concrete operation that demonstrates an understanding of relations is seriation. Children are asked, for example, to arrange a set of rods in order of increasing size. Children in the preoperational subperiod may subgroup the rods but have difficulty completing an entire array along the required dimension. They may understand the concept of smaller versus larger but have difficulty comprehending the gradual nature of change. The concrete operation of seriation underlies one's ability to think dimensionally.

Formal Operations. In the third and final stage in Piaget's conception of the intellectual development in the child, the logical structures of concrete operations are superseded by structures referred to as formal operations. The relationship between the real and the possible that characterizes adolescent thinking represents a reversal of that relationship in the thinking of the concrete-operational child. Inhelder and Piaget noted that the real has priority for the younger child, and that possibility is conceived of merely as a prolongation or extension of real operations, "as, for example, when, after having ordered several objects in a series, the subject knows that he could do the same with others." For the adolescent, however, the possible has priority, and the real is seen as a particular instance of it. "Henceforth, they conceive of the given facts as that sector of a set of possible transformations that has actually come about." This immediately presupposes that the adolescent can take a given empirical event (such as, "the long, thin rod bends") and categorize it within a system of possible combinations of events (e.g., long rods or short rods, thin rods or thick rods, bending or not bending). Three characteristics follow from this fundamental reorientation in thought: (1) adolescent thought is hypothetical-deductive; (2) it deals in propositions rather than in concrete events; and (3) it can isolate variables and examine all possible combinations of variables.

HYPOTHETICAL-DEDUCTIVE THOUGHT. As a hypothetical-deductive form of thought, formal operational intelligence proceeds from the possible to the real. In this sense, it mirrors scientific reasoning. The implications of a propositional statement are drawn and then tested against reality. Rather than building up a proposition by induction from disparate concrete examples to a loose generalization, formal intelligence operates

systematically from general statement to particular instance by means of testable hypotheses.

PROPOSITIONAL THOUGHT. That formal operations deal in propositions rather than in concrete events implies increased freedom from immediate content. At one level, this freedom implies the ability to manipulate abstractions that have been tied to concrete examples or events. The adolescent, for example, can perform a transitive inference (A < B, B < C; therefore, A < C) without any empirical demonstration of referents for the terms A and B. At another level, this freedom implies that having performed a concrete operation, the adolescent can abstract the results of that operation and can perform further operations on them. For example, an adolescent can perform the concrete operation of combining two liquids to observe the color of the resultant mix and then take the result of this operation and systematically relate it to results of all other combinations of available liquids.

ISOLATING VARIABLES AND EXAMINING COMBINATIONS. Instead of dealing with disparate concrete experiments, hypothetical-deductive adolescents can organize their investigations into a coherent pattern from the outset and then perform all relevant combinations of variables to test their hypotheses, thus isolating causal factors. Piaget's theory of formal operational cognition focused on scientific thinking. For example, the weight, speed, shape, and size of an object may all be seen to contribute to the size of a hole the object makes when it hits the ground.

A complete combinatorial system appears only during the period of formal operations. Instead of focusing on empirical givens, as a child in the concrete-operational stage does, the adolescent using formal operational thinking constructs a hypothetical system comprising the empirical givens. Whereas the younger child could classify events according to various categories, such as length, width, and weight, the adolescent uses that classification as a basis for abstracting all possible combinations of variables. Having done this, the adolescent can then test hypotheses derived from the combinatorial system. The result of this new ability is the capacity to test the causal significance of each individual factor in succession by holding all other factors constant.

Piaget interpreted the rise of formal operational thought in the context of his equilibrium model of cognitive development. Thus, he considered neurological maturation and experience of the object and interpersonal world as necessary but not sufficient to explain this qualitative improvement in thinking. In essence, the equilibration explanation is as follows: During the stage of concrete operations, a number of qualitatively heterogeneous factors are constructed by the child, resulting in the achievement of conservation of the factor in question even in the face of

perceptual transformations. Such factors include quantity, weight, volume, time, and length. Eventually, the child discovers that, in many concrete instances, the operation of these factors is interrelated. Thus, although the factors have been constructed mentally in relative isolation from one another, their presence in real objects is mixed. Through experience with impersonal and interpersonal objects, the child's concrete operational understanding of these factors is shown to be insufficient, and a more comprehensive, more intelligent understanding is stimulated.

Egocentrism

Each major period of cognitive development is characterized by a qualitative shift toward more-comprehensive and more-adaptive cognitive structures. In this sense, the adolescent is more intelligent than the infant. However, each transition to a higher level of cognitive organization is initially accompanied by a lack of full differentiation between self and object. The most relevant of these episodes for psychiatry occurs with the shift to formal operational reasoning in adolescence. The adolescent has the capacity to engage in hypothetical thinking and to understand others' points of view. However, this is initially accompanied by characteristic patterns of thought in which others are unrealistically presumed to be focusing on the adolescent. This may lead to increased social anxiety, narcissistic reactions, or hypersensitivity. The capacity for formal operations also implies that real, individual people, including parents and teachers, are one instance of what ideal people could potentially be like. This may underlie some features of adolescent rebellion and idealism.

INTELLIGENCE: PIAGETIAN MODEL AND ITS ALTERNATIVES

The Piagetian view of intelligence can be further understood by contrasting it with other contemporary models of intelligence. In the European and American traditions of clinical and educational psychology, intelligence had first of all a pragmatic meaning. Alfred Binet's work in Paris was intended to determine which children were delayed in learning and thus in need of special educational assistance. With Théodore Simon, he developed the first *intelligence test* by choosing a variety of tasks and arranging them in order according to the age at which most children master them. This enabled the clinician to assess at what mental age the child was functioning. Further developments in this tradition involved dividing the mental age by the chronological age to determine the child's intelligence quotient (IQ). This age-corrected notion of intelligence emphasized individual differences through comparisons of children with their peers instead of emphasizing the most general and progressively adaptive competencies of individuals during childhood and adolescence, as Piaget had done. Age-corrected intelligence came to be viewed as a trait, with

some children ahead of or behind their age mates in what was believed to be a basic capacity for learning.

Psychologists who espouse a trait model of intelligence differ among themselves on such key issues as to what extent the trait is heritable and whether there is a single general trait or numerous specific traits (types) of intelligence. Nevertheless, all trait theorists differ from Piaget by focusing on individual differences rather than on the most general aspects of cognitive adaptation.

Piaget's view of intelligence also differs from Lev Vygotsky's social model. In the latter, intelligence is viewed chiefly as the outcome of a process through which children internalize the cultural tools and practices represented in their social context. The contrast between Piaget and Vygotsky may at least superficially appear to be one between a theorist who emphasizes the most general cognitive processes and a theorist who emphasizes the most culturally specific cognitive content. As noted earlier, Piaget did include social experience and interaction as a factor contributing to cognitive development, along with biological maturation, learning, and equilibration. However, he did not emphasize the social and cultural content to the same extent as Vygotsky.

Finally, Piaget's construct of intelligence may be contrasted with the cognitive social learning construct of competencies. Although both constructs involve active, developing processes rather than fixed knowledge, the former pertains to the most general structures of logic derived from coordination of actions and the latter refers to more-specific, concrete sets of skills, including social and problem-solving skills. Attempts to integrate Piagetian and cognitive social learning models may hold particular promise for psychotherapeutic interventions. The latter model points to certain areas that are developmentally crucial for adjustment (goal setting, planning, problem solving), whereas Piaget model emphasizes the importance of active engagement in the development of structures.

Extensions of Piaget's Theory

Piaget's theory focuses primarily on logical-deductive reasoning. Little attention is paid to social intelligence, emotional intelligence, or alternative types of intelligence, such as artistic or athletic intelligence. Others have used a Piagetian framework to investigate some, but not all, of these realms. These efforts are referred to as "extensions" of Piaget's theory. The first and perhaps best-known attempt to extend Piaget's theory in this way was made by Lawrence Kohlberg, who studied the moral reasoning of children.

Although Piaget wrote *The Moral Judgment of the Child,* he did not further develop this area in subsequent work. Kohlberg developed a stage model in which the child's stage of moral reasoning depended on his or her

stage of (Piagetian) cognitive development. For instance, to demonstrate moral reasoning at the conventional level, the child must have already entered the concrete-operational stage of cognitive development. Likewise, principled moral reasoning demanded a foundation in formal operational thought.

Kohlberg described three major stages of moral reasoning, each divided into two substages. The three major stages included the morality of the preschool period (based on notions of avoiding punishment and striving for reward), conventional morality (based on notions of authority or mutual benefit), and principled morality (based on general internalized moral principles). As was the case for Piagetian stages, Kohlberg's stages were proposed as progressively more adequate and more highly structured, that is, a child would move only in a forward direction through the stages. However, it was not proposed that all or even most people would necessarily attain the highest level of moral reasoning. Kohlberg investigated the moral reasoning of individuals by presenting them with moral dilemmas and then observing their thinking as they attempted to resolve the dilemmas. The theory of development pertained to the form of their thinking and not to the content of their solution. The best known of the dilemmas involved a husband who needed to obtain a rare medication to save the life of his wife. The druggist who had the medication was charging an exorbitant price for it. The question is whether the husband would be justified in stealing it. In addressing this dilemma, the research participant might have recourse to such considerations as fear of punishment (preconventional), concerns for social order (conventional), or relative values of life and property (principled). Kohlberg was not concerned with the participant's final choice but only with the methods of reasoning used to reach that choice.

Kohlberg's extension of Piaget's work was subject to several criticisms. The moral subject was criticized as detached from moral content. In fact, individual differences in the predominant forms of moral reasoning did not seem to correlate highly or consistently with individual differences in actual moral behavior. The research methodology was highly verbal, so that intelligence and verbal sophistication could be confounded with moral reasoning. Finally, because his theory was originally developed from research with an entirely male sample, Kohlberg was criticized by Carol Gilligan for having proposed a theory that was gender biased. Gilligan argued that woman's moral reasoning proceeds from a primary concern with relationships, whereas man's moral reasoning proceeds from a primary concern with justice. However, subsequent empirical research has demonstrated that women score as high as, or higher than, men on measures of Kohlberg's stages of moral reasoning.

Subsequent to Kohlberg's theory, other attempts were made to relate Piaget's model to social cognition. Initial efforts followed Kohlberg's lead,

that is, stages of social cognition were proposed as being dependent on stages of intellectual development. Interpersonal perspective taking was hypothesized as being dependent on perceptual and cognitive perspective taking. However, it became clear that children's concepts about other people did not always follow such a layered structural design. James Youniss then developed a theory of children's concepts of other people that appropriated a different key element from Piaget's work. Rather than attempting to layer stages of social cognition on stages of intellectual development, Youniss proposed that social cognition has its own process of development, but that these are based on abstractions from interpersonal interactions. Integrating Piagetian psychology with the interpersonal psychiatric theory of Harry Stack Sullivan, Youniss proposed two major categories of children's social cognition: schemas about peers and schemas about authority figures. Each of these develops as a function of repeated interactions with peers or elders. As can be seen in this example, the second type of extension of Piaget's work into the interpersonal domain abandoned the position that stages of social cognition depended on broad structures of intellectual development in favor of the position that the active abstracting processes discovered by Piaget were operative across both types of cognition.

A third type of extension is directed toward the child's acquisition of social knowledge. Again, the Piagetian roots of this extension are to be found in *The Moral Judgment of the Child*. The extension is stimulated by cross-fertilization between adherents of Piaget's model and adherents of Vygotsky's theory. As Duveen argued, in *The Moral Judgment of the Child,* Piaget actually depicted two methods by which the child acquires knowledge of social and cultural content. The first is through social transmission, based on learning from an authority figure. The second is through processes of argumentation and debate with peers. Piaget's first process is quite similar to Vygotsky's notion that the child internalizes from elders a system of symbols that represents a social or collective product. Perret-Clermont pointed out that Piaget emphasized the second process much more than the first over the course of his career, and that there was a definite sense in which he viewed the first as potentially constricting. Indeed, it appeared to him to be based on authority but was lacking in the stimulation to new learning that characterized debate with peers and the social construction of knowledge.

To the extent that Piaget neglected or underemphasized the acquisition of knowledge through social transmission, his theory would be deficient in accounting for children's knowledge of social and cultural material that is passed on rather than actively constructed by the individual. Such materials would range from the obvious (socially and culturally acceptable behavior) to the subtle (mathematical inventions, such as the current number system). The implications for education are significant. Discovery

methods of education, including experimentation and peer discussion, may be more appropriate when the goal is to assist children in formulating the most general principles of reasoning or to facilitate creative processes, but other methods, including lecture, reading, and modeling, may be more appropriate when the goal is to convey specific social content.

The Post-Piaget Era

Piaget has been criticized for failing to account for individual differences among children in their intellectual development and for failing to account for the effects of cultural setting on the cognitive content of children's intelligence. Educators and clinicians may find Piaget's work of invaluable interest for its broad depiction of developmental stages and processes and yet frustrating because of its failure to provide guidance for interventions that would facilitate learning or correct deficiencies in individual children.

Cross-cultural research has called into question the generality or universality of Piaget's proposed stages. Different tasks appear to be mastered at different ages across cultures. Even within Western culture, there is evidence that modifications in training paradigms or in the way in which cognitive structures are operationalized can lead to significantly different results than those obtained by Piaget regarding children's abilities at different ages. More specifically, the criticism has been made that Piaget relied too heavily on the child's verbalizations to document understanding of a cognitive task. In fact, mastery of concrete and formal operational schema was only acknowledged in Piagetian research if the child could explain verbally why the correct answer was logically necessary.

Moreover, the superiority of formal to concrete operational thinking, in itself, appears to be culturally relative, as reviewed by Marie-Anne Suizzo. In some cultures, social intelligence, or what has been termed "emotional intelligence" is valued over technical, scholarly forms of cognition, or analogical reasoning is as valued as propositional reasoning.

Moreover, the notion of a developmental stage that may last for as long as 7 years or more, during which schemas of conservation are sequentially developed across areas as divergent as quantity and volume, may simply be too broad to represent a cohesive period. Empirical psychologists have also questioned the notion of cognitive structures by noting the low magnitude of correlations that are purported to measure the same underlying intellectual structure. Of course, similar criticisms have been made of proposed traits and of purportedly related functions in other areas of psychology as diverse as personality theory and the neuropsychology of executive functions.

As noted earlier, the era of "big theories" in psychology and psychiatry appears to be over. Given the explosion of knowledge across such diverse areas as neuroscience, cognitive science, personality and social processes, and cultural and ethnic diversity, it is unlikely that a big theory can possibly

hold in such a way as to account for such a complex arena as cognitive development. Gopnik referred to the post-Piaget era, pointing out that cognitive development appears to be specific to certain domains rather than integrated across vast stages. She advocated a developmental pluralism, with a continuing emphasis on testing the validity of each proposed theory. Two particular theories she discussed are "theory theory" and "script theory," both of which seem to have relevance for psychotherapy.

"Theory theory" proposes that cognitive development proceeds by way of the same mechanisms that scientists use to test, modify, and confirm or disconfirm theories. Like Piaget's model, this model emphasizes the active construction of intelligence on the part of the child or adolescent (or adult) and the role of evidence in leading to modifications. Theories provide predictions, descriptions, interpretations, and explanations of the ensuing evidence. Psychologists using "theory theory" have found that children's theories about phenomena are domain specific rather than general or stage specific (which Piaget proposed), and that the capacities for induction and causal thinking are present earlier in life than Piaget proposed.

Script theory is a more empirical approach to the learning of complex ideas and skills than is Piagetian theory. Scripts (or narratives) provide coherence and predictability, not on the basis of abstractions from repeated actions, but on the basis of repeated sequences of actions. Scripts are explanatory in a subjective sense, in that they represent the child's "story" about what happened in a given instance. A common example of a script is the sequence of events that one expects to occur on entering a restaurant: being greeted and seated, ordering courses in sequence, requesting the check, and making payment.

NEW CONCEPTS OF INTELLIGENCE: EMOTIONAL BASIS OF INTELLIGENCE

Human emotions have traditionally been viewed as somewhat separate from cognition and as a minor concern to overall development. However, research reviewed by Suizzo has shown that children's performance on cognitive tasks improves with higher levels of social support and is affected by the dynamics of interpersonal relationships and by feelings of self-efficacy. New clinical observations and theoretical formulations by Stanley Greenspan and emerging findings from a number of recent studies suggest that emotions are central to cognition and may actually regulate and orchestrate cognitive capacities. They may also be critical to the development of cognitive capacities. It is suggested that infants' emotional exchanges with their caregivers, rather than their ability to complete cognitive tasks, should become the primary measure of developmental and intellectual competence.

Twelve-month-old Cara sat in her mother's lap at a table, eyes locked onto the psychologist,

who tried to get her to follow the bean he was putting under the cup and search for it. Cara knocked over the cup. Is this little girl, as her mother fears, cognitively delayed? Does a 1-year-old child who never babbles like other children her age and who violently flings food and toys away from her show signs of a significant intellectual deficit? After a battery of similarly frustrating tests, the evaluator concluded that cognitive delay was the likely diagnosis in Cara's case.

For 50 years, developmental testers have expected babies to sit still in their mothers' laps, to pay attention, and to perform prescribed tasks while adults assess their basic intelligence. Traditional wisdom has long insisted that carefully scoring how well a tiny child fits pegs into boards, sorts cards by shape, or hunts beans under cups can reveal an accurate measure of intelligence and developmental competence. However, recent results from research and clinical practice by Stanley Greenspan and others suggest that this entire approach to assessing children's capacities rests on false premises and has inadvertently led to mistaken diagnoses that can stigmatize children throughout their school years.

When an evaluator schooled in this new thinking assessed Cara, he focused on her spontaneous interactions with her caregiver. He looked at each of Cara's intentional behaviors as a sign of her emotional interests. For instance, he observed her delight in yanking her mother's nose. At the assessor's suggestion, the mother permitted the tugging on her nose to continue and playfully responded, "Toot, toot." Cara smiled and pulled again. The baby was rewarded with another "Toot, toot" and a big smile from her mother. Cara soon began to copy her mother's gestures and eagerly thrust her nose towards mom. When the mother squeezed Cara's nose, the 1-year-old girl chirped, "Mo, mo," her first words.

Cara showed that she could initiate social interactions and could comprehend their consequences. That demonstrated degree of understanding put Cara at least at the 12-month level of cognitive development. Further observation revealed an extremely energetic, active, highly physical toddler who liked to have her way and to control her surroundings. With the consultant's help, Cara's mother later altered her parenting style. She learned to follow Cara's behavioral lead, then enthusiastically engaged her daughter in creative interactions while simultaneously setting firm limits. Cara's energy quickly became more focused; her babbling became richer. Before long, she was saying real words and actively cooperating with her parents.

If a series of such simple, pleasurable interactions with her mother could reveal and foster Cara's language development and organizational ability, then any conception of intellect that marked her as cognitively delayed because of an inability to search for a bean has serious flaws. Those flaws are based on a long-standing mistaken belief that the intellect is superior to and supervises the passions. Clearly, as Cara's linguistic debut demonstrates, analysis of a child's early relationships and sensory and emotional experiences is a vital key to accurate assessment of intelligence and developmental competence.

Until now, however, no one has offered an explanation of how emotions give birth to intelligence. In fact, a baby's earliest feelings play a pivotal role in all later intellectual development. Unlikely as the connection between feeling states and intelligence may seem, the emotions orchestrate a vast array of cognitive operations throughout an individual's life span. Indeed, they make possible all creative thought.

Results from four distinct lines of inquiry have recently shed new light on the importance of emotions for intelligence. In work with Arnold Sameroff, Greenspan found that children with four or more family emotional risk factors had 24 times the chance of scoring an IQ of less than

80 than children without those risks. Stephen Porges and Greenspan showed that measurements in 8-month-old infants of a part of the brain that regulates emotions correlate with these same children's IQ scores at 4 years of age.

Greenspan's work with a group of children with autistic disorder, who experience some of the most severe thinking and language problems imaginable, also confirmed the inextricable linkage of emotional and cognitive development. Therapeutic programs for these severely challenged children have traditionally concentrated on trying to stimulate their cognition and to teach them language. However, a program based on emotional cueing (like the one that revealed Cara's true abilities) proved to be more effective for a number of these children in fostering empathy, warmth, and creative thinking.

One young patient, Ashley, neither spoke nor made any response or eye contact with those around her. The 2-year-old child spent hours staring into space, rubbing persistently at the same patch of rug. Her abnormal repetition was viewed by the clinician observing her as more than just a distressing symptom of her autism. That symptom revealed an underlying interest and motivation that could be harnessed and redirected toward interacting with others. To initiate her cognitive progress, the clinician first had to motivate her to communicate with the simplest of emotional gestures—a smile, smirk, or purposeful hand movements. He suggested that her mother place her hand next to Ashley, on the favorite stretch of rug. When Ashley pushed it away, the mother gently put her hand back. Each time the child pushed, the mother's hand would return. A cat-and-mouse game ensued, and, after three sessions of these rudimentary interactions, Ashley was looking, smiling, and anticipating.

From that tiny beginning, through a comprehensive therapeutic program, grew a bridge to emotional relationships and eventual verbal exchanges. For example, as therapy progressed, the therapist helped Ashley use her imagination by repeatedly initiating pretend play. He recognized that each time Ashley repeatedly flung herself on her mother, the child was deriving sensory-based simple pleasure from her behavior. He instructed the mother to whinny like a horse each time Ashley lunged at her. Soon Ashley imitated mother's sounds and then started initiating her own sounds and words. In that way, the therapist helped the mother stretch a pleasant sensation for Ashley into a richer, more complex interaction. Over time, mother and child pretended to be neighing horses, mooing cows, and barking dogs. Their social and emotional interchange grew increasingly complex, passing through the same series of developmental stages identified in children without difficulties. At 7 years of age, Ashley enjoyed warm friendships, argued as well as her lawyer father, and scored in the low-superior IQ range.

A fourth line of inquiry—microscopic clinical observations of children's thinking—further clarifies the relationships between emotion and reason by revealing two necessary elements of thinking. The first process—creating a new idea—stems from the ability to use one's emotional experience to assign meaning and significance to daily events or concepts. The second process—reflection and logical analysis—examines the newly created idea according to whatever principles of logic the person possesses and places it in a wider frame of reference.

To understand those processes in action, the authors put a simple question to two young boys seen in therapy not long ago. When asked by the clinician, "What do you think about people

who act bossy to you?" Chris replied, "Well, teachers are bosses, baby-sitters are bosses, policemen are bosses." That articulate 7-year-old child lacked the emotional pathways that permit creative and intuitive thought. He could provide a formal classification of different types of bosses but could not relate these categories to his own life. However, 7-year-old Josh had no such difficulties. In response to the same question about bosses, he announced, "Most of the time I don't like being bossed, especially when my parents try to tell me when I can watch TV and when I have to go to sleep. I'm big enough to decide for myself. Sometimes when I'm being bad, I guess I need bossing, though. Maybe bosses are okay some of the time, and some of the time they're not." Josh found his answer in his, apparently generally irritating, brushes with bosses. Rather than simply listing categories or incidents, he could abstract a principle from the emotional core of those incidents.

How exactly did Josh's ability to think and to abstract develop? A baby's experience begins with sensations like touch and sound. Each sensation, however, also gives rise to an emotion. A toy may feel interesting or boring; a voice may feel soothing or jarring. Even young infants react to sensations emotionally. They prefer the sound or smell of their mother, for example, to any others and, by 4 months of age, can react to certain persons with fear. Furthermore, contrary to long-held assumptions, basic sensations, such as touch and sound, can be perceived differently by different people, giving rise to emotional differences.

Emotional meaning also adheres to early concepts like *big* and *little, more* and *less, near* and *far,* and *now* and *later*. *A lot* is a bit more than makes a child happy. *Near* is snuggled next to a child in bed. *Later* is a frustrating stretch of waiting. For a child without an intuitive sense of *few* and *many,* numbers have no meaning. Furthermore, a young child's experience of any sensation always occurs within the context of a relationship that gives it broader meaning. Playing with mother's hair, for example, may evoke smiles and hugs or an angry scolding.

Each sensory experience has such a dual aspect and is labeled by its physical properties and its emotional qualities. This double coding helps the child to place the memory or experience in a catalogue of experience and retrieve or reconstruct it when needed. As the child grows, emotional reactions come to operate as a sixth sense that allows the child to recognize and to understand situations.

Emotion orchestrates complex judgments as well. One of modern psychology's main enigmas is how children learn to discriminate among situations ("When can I yell and kick?") and to generalize from one to another ("Should I behave at school like I do at home?"). Consider how a child makes a seemingly simple judgment about when to say "hello." He or she does not learn a set of cognitive rules, such as greeting only those who live on his or her street or only those who wave at him or her. Rather, from countless specific encounters the child abstracts an emotional pattern; there is a feeling of warmth and friendliness in situations that rate "hello." The child's interactions create an emotional signaling system that tells him or her when to say "hello" and that it is okay to punt the football but not to

kick Sarah or Charlie in the shins. That emotional signaling system, which acts like an orchestra leader for the vast array of cognitive instruments, is a quintessentially human process. No computer, for all its apparent so-called brainpower, can ever get beyond limited elements of logical analysis and think like a person. Advocates of artificial intelligence may claim that current computational capacity limits creative, human-like thought, but the real limit is a machine's inherent inability to engage the world emotionally. No collection of microchips can ever have a child's lived emotional experience of "hello," of noses and hugs. None, therefore, can ever create the emotionally based meaning from which creative thought grows and on which it depends.

Looking at Piaget's theory from this perspective reveals the limitations of a cognitive theory that did not adequately deal with the central role of emotions. Piaget's experiments focused on how children comprehend the relationship between physical objects, developing the ability to classify them by such parameters as shape or size, but most children can classify their emotions and emotionally relevant relationships far earlier than they can classify physical objects. For example, they know members of their families from those who are not members, classifying the family as a unit. Some of Piaget's observations are limited because he depended so heavily on children's perceptual and motor performance to signal cognitive advances, even though motor skills often lag behind other skills.

More important, however, is Piaget's relative lack of focus on the role of emotions. He emphasized learning through doing but did not realize that the doing generates formative emotional reactions as well as perceptual, motor, and cognitive ones. Consider how a child learns what an apple is. You can have the child handle an apple and determine it is something red and round, bigger than a peanut and smaller than a watermelon. Alternatively, the child can observe the aspects mentioned previously, experience more satisfaction in eating an apple when he or she is hungry than when he or she is full, and know the pleasure of giving one to his or her favorite teacher.

The child can also imagine how the teacher feels when he or she gives the teacher the apple. The child may catalogue how he or she feels when the child throws one at his or her younger brother and gets him on the shoulder, as well as the child's disgust when an apple rots or when the child discovers half of a worm inside. If you ask creative adults to write an essay on apples, they will probably bring an enormous amount of personal affective experience to their reflections.

Piaget did emphasize how children's thinking comes to incorporate multiple perspectives as they grow older. A classic Piagetian experiment shows how school-age children learn to solve a problem involving weights on a seesaw. By assessing the heaviness of the weights and noticing where they are placed, they are able to figure out how a seesaw works. But what

was not appreciated by Piaget was that the children's postulated perspectives incorporated the additional enormous number of perceptions afforded by affective experiences. To neglect this element is therefore to fail to appreciate the rich array of experiences that contribute to forming abstract concepts.

TOWARD A GENERAL DEVELOPMENTAL MODEL

The diagnosis and treatment of emotional and developmental disorders in infants and young children requires that clinicians take into account all facets of the child's experience. Thus, one needs a model with which to look at how constitutional–maturational (regulatory), family, and interactive factors work together as the child progresses through each developmental phase, and each phase must be viewed from affective and cognitive perspectives.

A developmental, structuralist model formulated by Greenspan integrates cognitive and affective development and applies the types of structure Piaget described to a range of experience. Most important, the model also considers individual differences in terms of biology and interaction. New findings suggest that early interaction can alter the structure and wiring of the CNS. In this model, the biological differences express themselves in the unique way in which an infant processes sensations and organizes motor patterns. Interactions harness and change these basic processes. The model can be visualized with the infant's constitutional–maturational patterns on one side and the infant's environment, including caregivers, family, community, and culture, on the other side. Both sets of factors operate through the infant–caregiver relationship, which can be pictured in the middle. Those factors and the infant–caregiver relationship in turn contribute to the organization of experience at each of six developmental levels (consistent with cognitive and affective milestones), which may be pictured just beneath the infant–caregiver relationship.

This particular clinical and research model enables the user to look at the back-and-forth influence of highly specific, verifiable constitutional–maturational factors on interactive and family patterns and vice versa, in relationship to specific developmental processes (and to relate these processes to later developmental and psychopathological disorders).

Developmental Levels

The model contains six developmental levels, which include the ability of the infant and the child to accomplish the following:

1. Attend to multisensory affective experience and, at the same time, organize a calm, regulated state and experience pleasure.
2. Engage with and show affective preference and pleasure for a caregiver.

3. Initiate and respond to two-way presymbolic gestural communication.
4. Organize chains of two-way communication (opening and closing many circles of communication in a row), maintain communication across space, integrate affective polarities, and synthesize an emerging prerepresentational organization of self and others.
5. Represent (symbolize) affective experience (e.g., pretend play and functional use of language), which calls for higher-level auditory and verbal sequencing ability.
6. Create representational (symbolic) categories and gradually build conceptual bridges between these categories. This ability creates the foundation for such basic personality functions as reality testing, impulse control, self–other representational differentiation, affect labeling and discrimination, stable mood, and a sense of time and space that allows logical planning. This ability rests not only on complex auditory and verbal processing abilities, but also on visual-spatial abstracting.

At each level, one looks at the range of emotional themes organized (e.g., can the child play out [symbolize] only dependency themes and not aggressive ones? Is aggression behaved out and dealt with presymbolically?). One also looks at the stability of each level. Does a minor stress lead a child to lose the ability to represent, to interact, to engage, or to attend?

In their use in day-to-day clinical work, the six developmental levels can be collapsed into four essential processes that characterize development in infants and young children. These processes concern how an infant and the parents or caregivers negotiate the various phases of their early interactions, and they serve as a basis for diagnosis and treatment.

A 12-month-old infant's mother worried that, "he cries any time I try to leave him, even for a second. If I'm not standing right next to him when he is sitting on the floor, he cries and I have to pick him up. He's a tyrant. He's waking up four times at night and is a fussy eater. He eats for short bursts (breast-feeding) and then stops eating. I'm feeding him all the time."

The mother was feeling cornered, controlled, manipulated, and bossed around. Her baby was like a "fearful dictator" (therapist's term). She said, "that's the perfect way to describe him." The father was impatient with the mother; he felt that she indulged the baby too much. He was getting "fed up" because she had no time for him.

The baby was interactive and sensitive to every emotional nuance. As he came into the room, he immediately caught the clinician's eye. They exchanged smiles and motor gestures. He interacted with his parents with smiles, coos, and motor movements. Father intruded somewhat. He would roughhouse until the baby would cry, put the baby down, and then roughhouse again. Mother, in contrast, was ever so gentle, but long silences passed between her vocalizations. During her long silences, the baby would rev up, get more irritable, and start whining. He whined with his mother and cried fearfully with his father. Even before he could finish his motor gestures or vocalizations, his mother moved in and picked him up, gave him a rattle, or spoke for him. In this way, she undermined his initiative. Even while whining, however, he was interactive and contingent.

On physical examination, this baby was sensitive to loud noises and light touch on the arms, legs, abdomen, and back. He had a mild degree of low motor tone and was posturally insecure. He was not yet ready to crawl.

His constitutional and maturational patterns did not compromise his mastering the first developmental challenge of shared attention and engagement. He was an attentive, engaged baby. However, at the second developmental stage, intentional communication and assertiveness, he was a passive reactor. He was not learning to initiate two-way communication, to be assertive, and to take charge of his interactions. His low motor tone was compromising his ability to control his motor movements. His sensory hyperreactivity was compromising his ability to regulate sensation. He was frequently overloaded by just the basic sensations of touch and sound, and he was not receiving support from his mother through the nurturing and rhythmic care taking that would foster self-initiative.

This family required therapeutic work on a number of tasks simultaneously. The infant's special constitutional–maturational patterns were discussed. Hands-on practice helped the parents help their baby be attentive and calm. Those tasks included helping the mother be more patient, wait for the baby to finish what he started, and support his initiative (e.g., putting something in front of him while he was on his tummy, to motivate him to crawl and to reach); getting the mother to put more affect into her voice and to increase the rhythm and speed of her vocalizations; and getting the father to be more gentle. The parents' feelings about the interactions were explored: the father's tough-guy background, the mother's fear of her own assertiveness, her fear of her baby being injured, and their own associated family patterns.

Gradually, the baby began to sleep through the night, and he became more assertive and less clinging and fearful. He also became happier. He was slow to reach his motor milestones, so an occupational therapist worked with him and gave the parents advice on motor development and normalizing his sensory overreactivity. In 4 months, this infant was functioning in an age-appropriate manner with a tendency toward a cautious, but happy and assertive, approach to life's developmental challenges.

As developmental clinicians and researchers build on Piaget's findings and formulations, the developmental model serves as a basis for understanding social and emotional development and provides a framework for clinical and educational intervention.

Implications of Piaget's Work for Psychotherapy

Piaget was not an applied psychologist and did not develop the implications of his cognitive model for psychotherapeutic intervention. Nevertheless, his work formed one of the foundations of the *cognitive revolution* in psychology. One aspect of this revolution was an increasing emphasis on the cognitive components of the therapeutic endeavor. In contrast to classical psychodynamic therapy, which focused primarily on drives and affects, and in contrast to behavior therapy, which focused on overt actions, cognitive approaches to therapy focused on thoughts, including automatic assumptions, beliefs, plans, and intentions. By including "theory theory" and "script theory," additional applications to psychotherapy can be seen.

Cognitive development theory has influenced psychotherapeutic approaches in multiple ways. Some therapists have taken developmental notions from Piaget's work and developed intervention techniques. Others have developed cognitive models of treatment independent of Piaget but with heavy reliance on the role of cognition. Others have included Piaget's concepts in a broader set of constructs to undergird new developmental approaches to psychotherapy.

First, some psychotherapists applied Piagetian notions directly to child

interventions. Susan Harter, for example, discussed techniques for helping young children become aware of divergent or contradictory emotions and integrate these complex emotions within a more abstract or higher class of emotions. One of Harter's techniques is to ask the young child to make a drawing that shows different and conflicting feelings in one person. This technique represents an application of the concrete operation of class inclusion to the realm of the emotions. Harter's work applied Piagetian findings to the common therapeutic problem of helping children to recognize, tolerate, and integrate mixed or ambivalent affects within stable object relations. As such, it drew on cognitive theory and psychodynamic theory. Similar techniques are important in work with children who have been exposed to trauma or to sexual abuse. It is an essential component of such work to assist them in labeling, differentiating, and accepting the full range of emotions stemming from these experiences.

Second, other psychotherapists developed treatment models that, although not directly dependent on Piagetian psychology, emphasized core ideas quite similar to those Piaget discovered in his naturalistic observations of cognitive development. These models are even more closely aligned with recent developments in "theory theory." Aaron Beck, for example, developed an entire school of cognitive therapy that focuses on the role of cognitions in causing or maintaining psychopathology. Cognitive therapy has been shown to be an effective treatment for problems as diverse as depression, anxiety disorders, and substance abuse.

A core idea in cognitive therapy is that the patient has developed certain core beliefs, aspects of the self-schema, and conditional probability beliefs as a result of developmental experiences, and these contribute to emotional or behavioral problems. For example, depressed persons may have the core belief "I am unlovable." Addicted persons may have the belief "Unless I drink I cannot feel happy." In cognitive therapy, the person can be assisted to identify the negative automatic thoughts and underlying dysfunctional attitudes or beliefs that contribute to emotional distress or addictive behavior. The key therapeutic process after identification of the maladaptive thoughts is to help the patient view these thoughts more objectively, not take them in an unquestioning manner as veridical. Here, cognitive therapy emphasizes evidence, consistent both with Piagetian theory and "theory theory." The patient is assisted to seek out evidence to test negative thinking; active involvement, rather than passive listening, is required.

What the cognitive therapist accomplishes through such techniques as Socratic questioning and asking if there are other ways to look at the same event is similar to what the talented teacher does in guiding children to more adequate, more intelligent understanding of operational tasks. The notion of equilibration is relevant in both instances. By helping the individual see that previous cognitive structures are in some ways

inadequate, the therapist or teacher disturbs the old cognitive structure, and the patient or student experiences a disruption that leads to the search for more-adequate structures. The compensation for external disturbance is what Piaget termed *equilibration*. New structures can only be constructed through a process of accommodation, enabling the subject to assimilate a wider array of data, a new perspective, or more complex information.

Because it requires *thinking about thinking,* cognitive therapy seems to require formal operational thinking, although this has not been empirically tested. At the least, it requires the ability to recognize and articulate affects, to recognize and label events that give rise to affects, and to translate into a thought the mediating process that occurs rapidly between the event and the affect. Cognitive–behavioral models of psychotherapy include cognitive techniques and more-behavioral, interactive techniques, such as increasing pleasant activities and improving communication and problem-solving skills. It is possible that the less-cognitive, more-behavioral techniques, although requiring a lower level of cognitive development, can also lead to garnering of evidence and modification of specific expectancies, attributions, and self-schemata.

Because "script theory" or narrative approaches to cognition in psychotherapy are empirically based, generated by repetitive experiences rather than by reflective abstraction, and domain specific, they may have even more general application to psychotherapy than classic Piagetian theories or "theory theory." For example, in dialectical behavior therapy, patients provide a "chain analysis" of events, feelings, thoughts, situational stimuli, and interpersonal factors that led up to a negative or self-damaging behavior. This narrative provides guidance to the patient and the therapist about where and how to intervene to prevent subsequent similar behavior.

A third approach is to use Piagetian insights in psychotherapy by integrating Piaget's findings into a broader model. Greenspan, for example, has articulated a developmentally based psychotherapy that takes account of earlier, presymbolic levels of functioning that precede the ability to recognize, label, and articulate affects and their mediating cognitions.

Developmentally Based Psychotherapy

Developmentally based psychotherapy, as developed by Greenspan, integrates cognitive, affective, drive, and relationship-based approaches with new understanding of the stages of human development. Different therapies look at different aspects of the proverbial elephant, whether from a psychodynamic, object relation, self-psychology, behavioral, or cognitive–behavioral point of view. A comprehensive, cohesive developmental framework integrates elements from these approaches with a broader understanding of the developmental processes essential for emotional or mental health. It formulates a series of principles that an understanding of

human development says are prerequisite for emotional growth.

Developmentally based psychotherapy constructs its therapeutic strategies from these principles of human development and growth. The clinician first determines the level of the patient's ego or personality development and the presence or absence of deficits or constrictions. For example, can the person regulate activity and sensations, relate to others, read nonverbal affective symbols, represent experience, build bridges between representations, integrate emotional polarities, abstract feelings, and reflect on internal wishes and feelings?

After determining the developmental level, the clinician looks for constitutional and maturational contributions and difficulties with sensory processing, modulation, or motor planning. The clinician looks for interactive and family contributions. Each of these is explored in the present, the past, and the anticipated future. The patient's fantasies, sense of self and others, and conflicts are understood in the context of all of these influences. These are how patients make sense of their ego structure, physical makeup, family patterns, and interactions with others. Developmentally oriented therapists do not permit themselves the luxury of overfocusing on one set of variables, such as inner fantasies, family dynamics, biological proclivities, or prior experience. Similarly, the formulated therapeutic strategy cannot deal only with one or two factors. It must deal with all critical factors that influence the developmental process. As collaborators in the construction of experience, therapists use their understanding of the patient's development to help the patient to construct interactions that provide growth and overcome difficulties.

Often, it is assumed that critical aspects of development occur through the maturation of the nervous system along with routine, expectable experiences. It is also assumed that, from these routine, expectable maturational sequences and experiences, certain psychological structures having to do with the ability to regulate, engage, interact, represent (symbolize) experience, and reflect and compare experiences are present in most people. With these capacities in place, it is believed that the therapeutic process can focus on conflicts and anxieties and selected maladaptive behaviors or thoughts. It has been observed, however, that only a small percentage of individuals have these core capacities. For most, such capacities must be learned as part of the therapeutic process.

The developmental perspective shows how one learns these capacities during development. It suggests strategies that can be used in the psychotherapeutic process, so that adults and children who have not achieved these capacities can learn them. From a developmental point of view, the integral parts of the therapeutic process include learning how to regulate experience; to engage more fully and deeply in relationships; to read and respond to boundary-defining behaviors and affects; to perceive, comprehend, and respond to complex self- and object-defining affects,

behaviors, and interactive patterns; to represent experience; to differentiate represented experience; and to form higher-level differentiations, including the capacity to engage in the ever-changing opportunities, tasks, and challenges during the course of life (e.g., adulthood and aging) and, throughout, to observe and reflect on one's and others' experiences. Mastering these core developmental processes makes dealing with conflicts, anxieties, maladaptive behaviors, and thoughts possible.

These processes are the foundation of the ego and, more broadly, the personality. Their presence constitutes emotional health, and their absence constitutes emotional disorder. The developmental approach describes how to harness these core processes and so assist the patients in mobilizing their own growth.

REFERENCES

*American Psychological Association. Distinguished Scientific Contribution Awards. *Am Psychol.* 1970;25:65–79.

*Bennour M, Voneche J. The historical context of Piaget's ideas. In: Muller U, Carpendale J, Smith L, eds. *The Cambridge Companion to Piaget*. Available online at: http://dx.doi.org/10.1017/CCOL9780521898584.002

Gilligan C. *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press; 1982.

Gopnik A. The post-Piaget era. *Psychol Sci*. 1996;7:221.

Greenspan SI. *The Clinical Interview of the Child*. New York: McGraw-Hill; 1981.

Greenspan SI. *The Development of the Ego: Implications for Personality Theory, Psychopathology, and the Psychotherapeutic Process*. Madison, CT: International Universities Press; 1989.

Greenspan SI. *Infancy and Early Childhood: The Practice of Clinical Assessment and Intervention with Emotional and Developmental Challenges*. Madison, CT: International Universities Press; 1992.

Greenspan SI. *Developmentally Based Psychotherapy*. Madison, CT: International Universities Press; 1997.

Greenspan SI. *The Growth of the Mind and the Endangered Origins of Intelligence*. Reading, MA: Addison Wesley Longman; 1997.

Greenspan SI, Shanker S. *The Evolution of Intelligence: How Language, Consciousness, and Social Groups Come About*. Reading, MA: Perseus Books; 2003.

Harter S. A cognitive-developmental approach to children's expression of conflicting feelings and a technique to facilitate such expression in play therapy. *J Consult Clin Psychol*. 1977;45:417.

*Hsueh Y. He sees the development of children's concepts upon a background of sociology: Jean Piaget's honorary degree at Harvard University in 1936. *Hist Psychol*. 2004;7:20–44.

Inhelder B, Piaget J. *The Growth of Logical Thinking from Childhood to Adolescence*. New York: Basic Books; 1958.

Kohlberg L. The development of children's orientations toward a moral order: I. Sequence in the

development of moral thought. *Vita Humana*. 1963;6:11.

*Mayer SJ. The early evolution of Jean Piaget's clinical method. *Hist Psychol*. 2005;8:362–382.

Piaget J. *The Moral Judgment of the Child*. London: Kegan Paul, Trench, Trubner and Co.; 1932.

Piaget J. *Play, Dreams and Imitation in Childhood*. New York: Norton; 1951.

Piaget J. The stages of the intellectual development of the child. *Bull Menninger Clin*. 1962;26:120.

Piaget J. *The Early Growth of Logic in the Child*. New York: Norton; 1969.

Piaget J. Piaget's theory. In: Mussen PH, ed. *Carmichael' s Manual of Child Psychology*. 3rd ed. Vol. 1. New York: Wiley; 1970:703.

Piaget J. Piaget's theory. In: Mussen PH, ed. *Handbook of Child Psychology*. 4th ed. Vol. 1. New York: Wiley; 1983:103;128.

Piaget J, Inhelder B. *The Psychology of the Child*. New York: Basic Books; 1969.

Pinard A, Laurendeau M. Stage in Piaget's cognitive-developmental theory: Exegesis of a concept. In: Elkind D, Flavell JH, eds. *Studies in Cognitive Development: Essays in Honor of Piaget*. New York: Oxford University Press; 1969.

Shayer M. Not just Piaget, not just Vygotsky, and certainly not Vygotsky as alternative to Piaget. *Learn Instruct*. 2003;13:465.

Smith L, Dockrell J, Tomlinson P, eds. *Piaget, Vygotsky, and Beyond: Future Issues for Developmental Psychology and Education*. London: Routledge; 1997.

Sternberg RJ, Berg C, eds. *Intellectual Development*. Cambridge: Cambridge University Press; 1992.

Suizzo MA. The social-emotional and cultural contexts of cognitive development: Neo-Piagetian perspectives. *Child Develop*. 2000;71:846–849.

Tudge J, Rogoff B. Peer influences on cognitive development: Piagetian and Vygotskian perspectives. In: Lloyd P, Fernyhough C, eds. *Lev Vygotsky: Critical Assessments: The Zone of Proximal Development*. Vol. 3. New York: Routledge; 1999.

Youniss J. *Parent and Peers in Social Development: A Sullivan–Piaget Perspective*. Chicago, IL: University of Chicago Press; 1980.

▲ 3.3 Learning Theory

MARK E. BOUTON, PH.D.

The principles of learning are always operating and always influencing human activity. Because so much human behavior (including overt behavior, thought patterns, and emotion) is acquired through learning, learning principles are often deeply involved in the etiology and maintenance of psychiatric disorders. And because behavior change, the goal of all psychotherapy, is also strongly influenced by learning processes, learning principles can influence the effectiveness of a therapy. In fact, no therapy can be said to be immune to the effects of learning. Even the simple prescription of a medication can bring learning processes into play, because the patient will have opportunities to learn about the drug's effects and side effects, will need to learn to comply with the instructions and directions for taking it, and will need to learn to overcome any resistance to compliance. Learning theory is worth knowing, it has been said, because "much like the law of gravity, the laws of learning are always in effect."

HISTORICAL CONTEXT

The modern study of learning can be traced back to roughly the turn of the 20th century, when two great traditions in biology began to focus on behavior and learning. Working in the reflex tradition, the Russian physiologist Ivan Pavlov first saw the importance of what is now called *classical (or Pavlovian) conditioning*. In the famous description of his simplest experiment, Pavlov rang a bell and then gave a dog some food. After a few pairings of bell and food, the dog began to salivate to the bell, and thus anticipated the presentation of food. Today, classical conditioning is seen as an important and ubiquitous behavioral phenomenon as well as a powerful method for studying basic associative learning processes (see later discussion). In addition, from its earliest days, classical conditioning was also applied to psychiatric phenomena. For example, in the 1920s (before most of Pavlov's work was available in English), John B. Watson showed that human emotions are influenced by classical conditioning. Watson showed an infant boy a stimulus (which happened to be a laboratory rat) and then made a frightening noise. After a few pairings of the rat and the noise, the child became afraid whenever the rat was presented. This fear generalized to a rabbit, a dog, and a fur coat. Watson saw conditioning as a means by which human emotions could be elicited by an expanding range of cues.

Also working at the turn of the 20th century, but in the tradition of the early comparative psychologists (who were interested in studying the

evolution of the mind by studying the behavior of animals), Edward Thorndike ran some of the earliest systematic experiments on *instrumental (or operant) conditioning*. In his studies, cats were placed in "puzzle boxes" and required to learn to move a latch to get out of the box and find reward. Although the method was originally seen as a way to study the cognitive abilities of animals, the instrumental conditioning experiment was soon seen as a tool for studying how the effects of a behavior can strengthen or weaken it. In the late 1930s, B.F. Skinner expanded Thorndike's method by enclosing rats in a box and allowing them to operate a lever attached to the wall to earn a food-pellet reward. In this case, the animal was free to perform the behavior repeatedly, whenever it wanted to. Nowadays, and in parallel with Pavlovian conditioning, operant learning is seen as (1) an important and ubiquitous behavioral phenomenon, (2) a method for rigorously studying learning and motivation processes (see below), and (3) a phenomenon with applicability to clinical issues.

The application of learning theory to clinical issues became especially central to the *behavior therapy* movement that began in the 1950s and 1960s (e.g., Wolpe, 1958). The specific application of operant conditioning principles became known as *applied behavioral analysis*. The fundamental idea was that psychiatric disorders could be understood and treated using scientifically established principles of learning. The relationship between lab-based research on learning and the practice of clinical psychologists was viewed as analogous to the relationship between lab-based medical research and the practice of physicians. Since the 1970s, the field of behavior therapy (now called *cognitive behavior therapy*) has accepted what seems to be a wider range of explanatory principles, often grounded in social psychology and at least nominally in an information-processing ("cognitive") perspective rather than an explicitly "behavioral" perspective. At the same time, the basic science of learning theory has also advanced and expanded in ways that have not always been appreciated by clinical scientists. Beginning in the late 1960s, new findings led to conceptual changes in how learning processes were analyzed—with an increasing emphasis on cognitive processes and evolution. These developments further deepened the understanding of learning processes and widened the possible implications and applications to clinical problems and issues.

BASIC CONCEPTS AND CONSIDERATIONS

A great deal of modern research on learning still focuses on Pavlovian and operant learning. Pavlovian conditioning occurs when neutral stimuli are associated with a psychologically significant event. The main result is that the stimuli come to evoke a set of responses or emotions that may contribute to many clinical disorders, including (but not limited to) anxiety disorders and drug dependence. The events in Pavlov's experiment are

often described using terms designed to make the experiment applicable to any situation. The food is the *unconditional stimulus,* or *US,* because it unconditionally elicits salivation before the experiment begins. The bell is known as the *conditional stimulus,* or *CS,* because it only elicits the salivary response conditional on the bell-food pairings. The new response to the bell is correspondingly called the *conditional response* (*CR*), while the natural response to the food itself is the *unconditional response* (*UR*). Of course, as will be illustrated below, modern laboratory studies of conditioning use a very wide range of CSs and USs and measure a wide range of conditioned responses.

Operant conditioning occurs when a behavior (instead of a stimulus) is associated with a psychologically significant event. In the laboratory, the most famous experimental arrangement is the one in which a rat presses a lever to earn food pellets. In this case, as opposed to Pavlov's, the behavior is said to be an *operant* because it operates on the environment. And the food pellet is a *reinforcer,* an event that increases the strength of the behavior it is made a consequence of. A major idea behind this method is that the rat's behavior is "voluntary" in the sense that the animal is not compelled to make the response (it can perform it whenever it "wants" to). In this sense, it is similar to the thousands of operant behaviors that humans choose to commit—freely—in any day. Of course, the even larger idea is that even though the rat's behavior appears as though it is voluntary, it is lawfully controlled by its consequences: If the experimenter were to stop delivering the food pellet, the rat would stop pressing the lever, and if the experimenter were to allow the lever press to produce larger pellets, or perhaps pellets at a higher probability or rate, then the rate of the behavior might increase. The point of operant conditioning experiments, then, is largely to understand the relation of behavior to its payoff.

Pavlovian and operant conditioning differ in several ways. One of the most fundamental differences is that the responses observed in Pavlov's experiment are *elicited* and thus controlled by presentation of an antecedent stimulus. In contrast, the "response" observed in Skinner experiment is not elicited or compelled by an antecedent stimulus in any obvious way—it is instead controlled by its consequences. This distinction between *operants* and *respondents* is important in clinical settings. If a young patient is referred to the clinic for acting out in the classroom, an initial goal of the clinician will be to determine whether the behavior is a respondent or an operant, and then go about changing either its antecedents or its consequences, respectively, to reduce its probability of occurrence.

Effect of instrumental (operant) behavior	Positive event	Negative event
Produces event	*Reward learning* ↑	*Punishment* ↓
Prevents event	*Omission* ↓	*Avoidance Escape* ↑

FIGURE 3.3–1. The law of effect in instrumental/operant learning. Actions either produce or prevent good or bad events and the strength of the action changes accordingly (indicated by arrow). "Reinforcement" refers to a strengthening of behavior. *Positive reinforcement* occurs when an action produces a positive event, while *negative reinforcement* occurs when an action prevents or eliminates a negative event.

CS signals	Positive event	Negative event
Increase in Probability of event	*Approach CS*	*Withdraw from CS*
Decrease in Probability of event	*Withdraw from CS*	*Approach CS*

FIGURE 3.3–2. Sign tracking in Pavlovian learning. CSs either signal an increase or a decrease in the probability of good or bad events, and the CS generally engages approach or withdrawal behaviors accordingly.

Despite the academic separation of operant and respondent conditioning, they have an important common function: both learning processes are designed by evolution to allow organisms to adapt to the environment. The idea is illustrated by considering the *law of effect* (Fig. 3.3–1), which says that whether an operant behavior increases or decreases in strength depends on the effect it has on the environment. When the action leads to a positive outcome, the action is strengthened; conversely, when the action leads to a negative outcome, there is *punishment,* and the action is weakened. In a similar manner, when an action decreases the probability of a positive event, behavior also declines. (Such a procedure is now widely known as *time-out* from reinforcement.) When an action

terminates or prevents the occurrence of a negative event, the behavior will strengthen. By thus enabling the organism to maximize its interaction with positive events and minimize its interaction with negative ones, operant conditioning allows the organism to optimize its interaction with the environment. Of course, events that were once positive in the human's earlier evolutionary history are so prevalent in modern society that they do not always seem adaptive today. Thus, reward learning also provides a framework for understanding the development of rather maladaptive behaviors like overeating (in which behavior is reinforced by food) and drug taking (in which behaviors are reinforced by the pharmacological effects of drugs)—cases where reward principles lead to psychopathology.

A parallel to Figure 3.3–1 exists in Pavlovian conditioning, in which one can likewise think of whether the CS is associated with positive or negative events (Fig. 3.3–2). Although such learning can lead to a wide constellation or system of behaviors (as illustrated later), in a very general way, it also leads to behavioral tendencies of approach or withdrawal. Thus, when a CS signals a positive US, the CS will tend to evoke approach behaviors—called *sign tracking*. For example, an organism will approach a signal for food. Analogously, when a CS signals a negative US, it will evoke behaviors that tend to move the organism away from the CS. Conversely, CSs associated with a decrease in the probability of a good thing will elicit withdrawal behaviors, while CSs associated with the decrease in the probability of a bad thing can elicit approach. An example of the latter case might be a stimulus that signals safety, or the decrease in probability of an aversive event, which evokes approach in a frightened organism. In the end, these very basic behavioral effects of both operant (Fig. 3.3–1) and Pavlovian (Fig. 3.3–2) learning serve to maximize the organism's contact with good things and minimize the contact with bad things.

Perhaps because they have such similar functions, Pavlovian learning and operant learning are both influenced by similar variables. For example, in either case, behavior is especially strong if the magnitude of the US or reinforcer is large, or if the US or reinforcer occurs relatively close to the CS or operant response in time. And in either case, the learned behavior decreases if the US or reinforcer that was once paired with the CS or the response is eliminated from the situation. This phenomenon, called *extinction,* provides a means of eliminating unwanted behaviors or behavioral excesses that were learned through either form of conditioning and has led to a number of very effective cognitive-behavioral therapies.

Pavlovian and operant conditioning are usually described separately, in part because the separation has helped to simplify their analysis in the laboratory. However, it is important to bear in mind that both processes are occurring all the time, and that both are also always occurring together. For example, at the same time the drug user is learning to associate various operant behaviors with the reinforcing effects of drugs, he or she is also

learning to associate the drug with various stimuli, such as room cues or drug paraphernalia, which are also present. As another example, when a person with a weight problem learns behaviors that lead him or her to a fatty meal, he or she is also learning where that food is obtainable, and to connect the flavors of the food with their (too-desirable) caloric consequences. Behavior results because both the tendency to repeat behaviors is reinforced and because the Pavlovian signals for drugs or food can evoke approach. But there is more than this, as will be shown.

PAVLOVIAN CONDITIONING

Effects of Conditioning on Behavior

As noted in an article by Bouton in 2001, where some of the following discussion was first presented, many lay people have the mistaken impression that Pavlovian learning is a rigid affair in which a fixed stimulus comes to elicit a fixed response. In fact, conditioning is considerably more complex and dynamic than that. For example, signals for food may evoke a large set of responses that function to prepare the organism to digest food: they can elicit secretion of gastric acid, pancreatic enzymes, and insulin in addition to Pavlov's famous salivary response. The CS can also elicit approach behavior (as above), an increase in body temperature, and a state of arousal and excitement. When a signal for food is presented to a satiated animal or human, he or she may get up and eat more food. Some of these effects may be motivational; for example, an additional effect of presenting a CS for food is that it can invigorate ongoing operant behaviors that have been reinforced with food. CSs thus have powerful behavioral potential. Signals for food evoke a whole *behavior system* that is functionally organized to find, procure, and consume food.

As suggested earlier, Pavlovian conditioning is also involved in other aspects of eating. Through conditioning, humans and other animals may learn to like or dislike different foods. In animals like rats, flavors associated with nutrients (sugars, starches, calories, proteins, or fats) come to be preferred. Flavors associated with sweet tastes are also preferred, while flavors associated with bitter tastes are avoided. (These tendencies may be seen as additional examples of the approach and withdrawal tendencies summarized in Fig. 3.3–2.) At least as important, flavors associated with illness become disliked, as illustrated by the person who gets sick drinking an alcoholic beverage and consequently learns to hate the flavor. The fact that flavor CSs can be associated with such a range of biological consequences (USs) is important for omnivorous animals that need to learn about new foods. And it has some clinical implications. For example, chemotherapy can make cancer patients sick, and can therefore cause the conditioning of an aversion to a food that was eaten recently (or to the clinic itself). Other evidence suggests that animals may learn to

dislike food that is associated with becoming sick with cancer itself. On the flip side, conditioning can enable external cues to trigger food consumption and craving, a potential influence on overeating and obesity.

Pavlovian conditioning, as noted in the previous section, also occurs when organisms ingest drugs. Whenever a drug is taken, in addition to reinforcing the behaviors that lead to its ingestion, the drug constitutes a US, and may be associated with potential CSs that are present at the time (rooms, odors, injection paraphernalia, etc.). CSs that are associated with drug USs can sometimes have an interesting property: They often elicit a conditioned response that seems opposite to the unconditional effect of the drug. For example, although morphine causes a rat to feel less pain, a CS associated with morphine elicits an opposite increase, not a decrease in pain sensitivity. Similarly, although alcohol can cause a drop in body temperature, a conditioned response to a CS associated with alcohol is typically an increase in body temperature. In these cases, the conditioned response is said to be *compensatory* because it counteracts the drug effect. Compensatory responses are another example of how classical (Pavlovian) conditioning helps organisms prepare for a biologically significant US.

Compensatory conditioned responses have implications for drug abuse. First, they can cause drug tolerance, in which repeated administration of a drug reduces its effectiveness. As a drug and a CS are repeatedly paired, the compensatory response to the CS becomes stronger and more effective at counteracting the effect of the drug. The drug therefore has less impact. One implication is that tolerance will be lost if the drug is taken without being signaled by the usual CS. Consistent with this idea, administering a drug in a new environment can cause a loss of drug tolerance and make drug overdose more likely. A second implication stems from the fact that compensatory responses may be unpleasant or aversive. A CS associated with an opiate may elicit several compensatory responses—it may cause the drug user to be more sensitive to pain, undergo a change in body temperature, and perhaps become hyperactive (the opposite of another unconditional opiate effect). The unpleasantness of these responses may motivate the user to take the drug again to get rid of them, an example of escape learning, or negative reinforcement (Fig. 3.3–1), and a classic example of how Pavlovian and operant learning processes might readily interact (see below). The idea is that the urge to take drugs may be strongest in the presence of CSs that have been associated with the drug. The hypothesis is consistent with self-reports of abusers who, after a period of abstinence, are tempted to take the drug again when they are re-exposed to drug-associated cues.

USs do not always cause the conditioning of compensatory responses, as is illustrated by the fact that Pavlov's CS for food caused the dog to salivate, rather than acquire a dry mouth. Moreover, not all drugs cause the conditioning of compensatory responses either; for example, when cocaine

or amphetamine is associated with an environment, the environment tends to evoke a drug-like increase in activity rather than its opposite. Whether the CR appears similar or opposite to the unconditional effect of the drug (the UR) appears to depend on how the UR is actually mediated by the nervous system. Compensatory responses appear to develop as the conditioned response when what is labeled as the "unconditional response" is actually independent of the brain or spinal cord. For example, although an injection of insulin will cause a drop in the level of blood glucose, it does so because of its direct effect on the body's cells, and not via a reaction by the nervous system. The "response" is actually a stimulus to the nervous system, which unconditionally responds by recruiting an increase in blood glucose. The CR that consequently develops to CSs associated with insulin can therefore be a compensatory increase in blood glucose. The CR, however, is the same as the brain's actual reaction—the UR, correctly identified.

USs, however, may have multiple effects on the body, some of which may be mediated by the nervous system, and some of which may not be. For example, an injection of atropine causes both dilation of the pupils and a drying of the mouth. The first effect is mediated by the brain, but the second effect is not; atropine suppresses activity of the salivary glands, which presumably initiates a brain reaction that compensates. Consistent with the aforementioned, a CS associated with injection of atropine will elicit a conditioned dilation of the pupils as well as salivation, the latter being a compensatory response.

Pavlovian learning may potentially be involved in anxiety disorders. It is now known that CSs associated with frightening USs can elicit a whole system of conditioned fear responses, broadly designed to help the organism cope. In animals, cues associated with frightening events (such as a brief footshock) elicit changes in respiration, heart rate, and blood pressure, and even a (compensatory) decrease in sensitivity to pain. Brief CSs that occur close to the US in time can also elicit adaptively timed protective reflexes. For example, the rabbit blinks to a brief signal that predicts a mild electric shock near the eye. The same CS, when lengthened in duration and paired with the same US, elicits mainly fear responses, and fear elicited by a CS may potentiate the conditioned eyeblink response elicited by another CS or a startle response to a sudden noise. Once again, CSs do not merely elicit a simple reflex, but evoke a complex and interactive set of responses.

Classical fear conditioning can contribute to phobias (where specific objects may be associated with a traumatic US), as well as other anxiety disorders such as panic disorder and PTSD. In panic disorder, people who have unexpected panic attacks can become anxious about having another one. In this case, the panic attack (the US or UR) may condition anxiety to the external situation in which it occurs (e.g., a crowded bus) and also

internal ("interoceptive") CSs created by early symptoms of the attack (e.g., dizziness or a sudden pounding of the heart). These CSs may then come to evoke anxiety or panic responses. Panic disorder may begin because external cues associated with panic can arouse anxiety, which may then exacerbate the next unconditional panic attack and/or panic response elicited by an interoceptive CS. It is possible that the emotional reactions elicited by CSs may not require conscious awareness for their occurrence or development. Indeed, fear conditioning may be independent of conscious awareness.

As suggested earlier, in addition to eliciting conditioned responses, CSs also motivate ongoing operant behavior. For example, presenting a CS that elicits anxiety can increase the vigor of operant behaviors that have been learned to avoid or escape the frightening US. Thus, an individual with an anxiety disorder will be more likely to express avoidance in the presence of anxiety or fear cues. Similar effects may occur with CSs that predict other USs (such as drugs or food)—as already mentioned, a drug-associated CS may motivate the drug abuser to take more drugs. The motivating effects of CSs may stem from the fact that CSs may be associated with both the sensory and emotional properties of USs. For example, the survivor of a traumatic train derailment might associate stimuli that occur immediately before derailment (such as the blue flash that occurs when the train separates from its overhead power supply) with both the emotional and the sensory aspects of the crash. Consequently, when the survivor later encounters another flashing blue light (e.g., the lights on a police car), the CS might evoke both emotional responses (mediated by association with the trauma's emotional qualities) and sensory associations (mediated by association with the trauma's sensory qualities). Both might play a role in the nightmares and "re-experiencing" phenomena that are characteristic of PTSD.

The Nature of the Learning Process

Research beginning in the late 1960s began to uncover some important details about the learning process behind Pavlovian conditioning. Several findings proved especially important. It was shown, for example, that conditioning is not an inevitable consequence of pairing a CS with a US. Such pairings will not cause conditioning if there is a second CS present that already predicts the US. This finding (known as *blocking*) suggests that a CS must provide new information about the US if learning is to occur. The importance of the CS's information value is also suggested by the fact that a CS will not be treated as a signal for the US if the US occurs equally often (or is equally probable) in the presence and the absence of the CS. Instead, the organism treats the CS as a signal for the US if the probability of the US is greater in the presence of the CS than in its absence. In addition, the organism will treat the CS as a signal for "no US"

if the probability of the US is less in the presence of the CS than in its absence. In the latter case, the signal is called a *conditioned inhibitor* because it will inhibit performance elicited by other CSs. The conditioned inhibition phenomenon is clinically relevant because inhibitory CSs may hold pathological CRs like fear or anxiety at bay. A loss of the inhibition would allow the anxiety response to emerge.

Consistent with the role of "information value," the conditioning process is organized so that conditioning mainly accrues to the most valid predictors of the US. For example, in the experimental design illustrated in Figure 3.3–3, two groups of subjects received subtly different treatments that had markedly different consequences on learning and behavior. For the correlated group, on some of the trials CSs A and X were presented together (in a "compound") and always paired with a US. On other trials, B and X occurred together, and the US was never presented. After a number of such AX and BX trials, the learning process treated A as the best predictor of the US—that stimulus blocked learning to X, despite the fact that X was paired with the US on half the occasions on which it was presented. In contrast, in the uncorrelated group, X was also paired with the US half the time, but the results were very different. For this group, X was also combined with cues A and B half the time, but A and B were no more predictive of the US than X (the US was also presented on half the trials with either A or B). Because X was as valid a predictor as A or B, there was substantial learning to X. (In fact, because X was actually paired with the US twice as often as A or B, it was the best predictor of the US.) The pattern of results has been observed over a very wide range of conditioning methods, from fear conditioning in rats, category learning in humans, and sucrose conditioning in honey bees. It seems consistent with the view that, to understand conditioning, "the organism is [best] seen as an information seeker using logical and perceptual relations among events ... to form a sophisticated representation of its world." The fact that conditioning tends to accrue to the most valid predictors of a US suggests that subtle differences in conditioning history can have a profound influence on an individual's learning and behavior. Thus, if two different drug users were to take a drug in a manner analogous to that of the correlated and uncorrelated groups, one abuser would have tolerance and craving elicited almost exclusively by CS A, whereas the other would have tolerance and craving elicited strongly by A, B, and especially X.

Correlated Group	Uncorrelated Group
AX – US	AX – US
AX – US	AX – no US
BX – no US	BX – US
BX – no US	BX – no US

FIGURE 3.3–3. Types of trials administered to different groups of subjects described in the text. For the correlated group, A is the most valid predictor of the US, and it consequently controls most conditioned responding—for example, blocking conditioning of X. For the uncorrelated group, despite the overall (but superficial) similarity in types of trials, all CSs are associated with the US on 50 percent of their presentations, and all CSs therefore come to control conditioned responding. (Because X is paired with the US more times than A or B, it is actually the best-conditioned CS in this arrangement.) Conditioning depends on the relative informativeness, validity, or predictive value of the CS.

To account for Pavlovian learning, most theorists now suppose that conditioning is not determined by CS-US pairings, but by the discrepancy between (1) the US predicted by all CSs present on a trial and (2) the US that actually happens on the trial. That discrepancy is often referred to as *prediction error*. One implication is that, depending on the size and direction of the discrepancy or error, the pairing of a CS and traumatic US, for example, can cause an increase in fear conditioning, no change in conditioning, or even a decrease in conditioning. The latter idea has implications for understanding inhibition. A new CS can acquire a negative value if the US is smaller than that which other CSs present predict. Casually speaking, then, a conditioned inhibitor is really a CS that predicts "less US than expected." And a conditioned excitor is a CS that predicts "more US than expected."

These principles have been given formal power in the Rescorla–Wagner model, which asserts that the change in associative strength to a CS on any conditioning trial will be a function of the difference between the strength of the US that occurs on that trial and the extent to which the US is predicted by other CSs that are also present:

$$\Delta V_{cs} = \alpha\beta\,(\lambda - \Sigma V)$$

where ΔV_{CS} is the change in associative strength (V) to the CS, λ is the strength of the US, and ΣV is the summed associative strengths of all other CSs that are present on the trial (the extent to which the US is already

predicted). The α and β terms are merely fractions that influence the rate of learning by influencing the amount of associative change that will occur on the trial (they are based on the salience of the CS and US, respectively). In this theory, CSs can acquire either positive or negative values that correspond to excitation and inhibition, respectively. Several later theories have expanded this conceptualization and have emphasized a number of cognitive psychological processes, such as surprisingness of the US and the CS, short-term memory, memory priming, and also attention to the CS in accounting for this deceptively simple form of learning.

Classical conditioning is most robust if the CS and US are intense or salient. It is also best if the CS and US are novel. For example, in *latent inhibition,* repeated exposure to the CS alone before conditioning can diminish its ability to elicit responding when it is paired with the US. In the *US pre-exposure effect,* repeated exposure to the US before conditioning can likewise decrease the conditioning that later occurs when a CS and the US are paired. One idea is that the CS and the US must be "surprising" at the time of their pairing for learning to occur. Thus, the effects of pairing a CS with trauma or drug USs may depend in subtle ways on the individual's prior experience with the CS and the US. Conditioning theories that followed the Rescorla–Wagner model captured this idea by emphasizing that CSs or USs that were already "primed" in short-term memory by either recent presentation or by a retrieval cue were not surprising, and commanded less mental processing necessary for learning. Other theories have also emphasized a critical role for attention to the CS, which is thought to increase under circumstances where its consequence is not known and decrease when its consequence is known. Thus, with extended conditioning, the organism may respond to the CS automatically without much attention or mental processing devoted to it.

There are also important variants of classical conditioning. In *sensory preconditioning,* two stimuli (A and B) are first paired, and then one of them (A) is later paired with the US. Stimulus A evokes conditioned responding, of course, but so does stimulus B—indirectly, through its association with A. One implication is that exposure to a potent US like a panic attack may influence one's reactions to stimuli that have never been paired with the US directly; the sudden anxiety to stimulus B might seem spontaneous and mysterious. A related finding is *second-order conditioning*. Here, A is paired with a US first and then subsequently paired with stimulus B. Once again, both A and B will evoke responding. Sensory preconditioning and second-order conditioning increase the range of stimuli that can control conditioned response. A third variant worth mentioning occurs, as indicated above, when the onset of a stimulus becomes associated with the rest of itself, as when a sudden increase in heart rate caused by the onset of a panic attack comes to predict the rest of the panic or feeling the onset of a drug may predict the rest of the drug

effect. Such intra-event associations may play a role in many of the body's regulatory functions, such that an initial change in some variable (e.g., blood pressure or blood glucose level) may come to signal a further increase in that variable and therefore initiate a conditioned compensatory response.

Emotional responses can also be conditioned through observation. For example, a monkey that merely observes another monkey being frightened by a snake can learn to be afraid of the snake. The observer learns to associate the snake (CS) with its own emotional reaction (US/UR) to the other monkey being afraid. Although monkeys readily learn to fear snakes, they are less likely to associate other salient cues (such as colorful flowers) with fear in the same way. This is an example of *preparedness* in classical conditioning—some stimuli are especially effective signals for some USs because evolution has made them that way. Another example is the fact that tastes are easily associated with illness but not shock, whereas auditory and visual cues are easily associated with shock but not illness. Preparedness may explain why human phobias tend to be for certain objects (snakes or spiders) and not others (knives or electric sockets) that may as often be paired with pain or trauma.

Erasing Pavlovian Learning

If Pavlovian learning plays a role in the etiology of behavioral and emotional disorders, a natural question is how to eliminate it or undo it. As mentioned previously, Pavlov studied *extinction:* Conditioned responding decreases if the CS is presented repeatedly without the US after conditioning. Extinction is the basis of many behavioral or cognitive-behavioral therapies designed to reduce pathological conditioned responding through repeated exposure to the CS (*exposure therapy*), and it is presumably a consequence of any form of therapy in the course of which the patient learns that previous harmful cues are no longer harmful. Another elimination procedure is *counterconditioning,* in which the CS is paired with a very different US/UR. Counterconditioning was the inspiration for *systematic desensitization,* a behavior therapy technique in which frightening CSs are deliberately associated with relaxation during therapy.

Although extinction and counterconditioning reduce unwanted conditioned responses, they do not destroy the original learning, which remains in the brain, ready to return to behavior under the right circumstances. For example, conditioned responses that have been eliminated by extinction or counterconditioning can recover if time passes before the CS is presented again (*spontaneous recovery*). Conditioned responses can also return if the patient returns to the context of conditioning after extinction in another context, or if the CS is encountered in a context that differs from the one in which extinction has occurred (all

are examples of the *renewal effect*). The renewal effect is important, because it illustrates the principle that extinction performance depends on the organism being in the context in which extinction was learned. If the CS is encountered in a different context, the extinguished behavior may relapse or return. Recovery and relapse can also occur if the current context is associated again with the US ("reinstatement") or if the CS is paired with the US again ("rapid reacquisition"). One theoretical approach assumes that extinction and counterconditioning do not destroy the original learning, but instead entail new learning that gives the CS a second meaning (e.g., "the CS is safe" in addition to "the CS is dangerous"). As with an ambiguous word, which likewise has more than one meaning, responding evoked by an extinguished or counterconditioned CS depends fundamentally on the current context (Bouton, 2002).

Research on context effects in both animal and human learning and memory suggests that a wide variety of stimuli can play the role of context (Table 3.3–1). Drugs, for example, can be very salient in this regard. When rats are given fear extinction while under the influence of a benzodiazepine tranquilizer or alcohol, fear is renewed when the CS is tested in the absence of the context provided by the drug. This is an example of *state-dependent learning,* in which retention of information is best when tested in the same state in which it was originally learned. State-dependent fear extinction has obvious implications for combining therapy with drugs. It also has implications for the administration of drugs more generally. For example, if a person were to take a drug to reduce anxiety, the anxiety reduction would reinforce drug taking. State-dependent extinction might further preserve any anxiety that might otherwise extinguish during natural exposure to the anxiety-eliciting cues. Thus, drug use could paradoxically preserve the original anxiety, creating a self-perpetuating cycle that could provide a possible explanation of the link between anxiety disorders and substance abuse. One point of this discussion is that drugs can play multiple roles in learning: they can be USs or reinforcers on one hand, and CSs or contexts on the other. The possible complex behavioral effects of drugs are worth bearing in mind.

Table 3.3–1.
Effective Contextual Stimuli Studied in Animal and Human Research Laboratories

Exteroceptive context:
Room, place, environment, other external background stimuli
Interoceptive context:
Drug state
Hormonal state
Mood state
Deprivation state

Another general message is that contemporary theory emphasizes the fact that extinction (and other processes like counterconditioning) entails new learning rather than a destruction of the old. Recent psychopharmacological research has built on this idea: If extinction and therapy constitute new learning, then drugs that might facilitate new learning might also facilitate the therapy process. For example, there has been considerable recent interest in D-cycloserine, a partial agonist of the *N*-methyl-D-aspartate (NMDA) glutamate receptor. The NMDA receptor is involved in long-term potentiation, a synaptic facilitation phenomenon that has been implicated in several examples of learning. Of interest, there is evidence that administration of D-cycloserine can facilitate extinction learning in rats and possibly humans undergoing exposure therapy for anxiety disorders. In the studies supporting this conclusion, the administration of the drug increased the amount of extinction that was apparent after a small (and incomplete) number of extinction trials. Although such findings are promising, it is important to remember that the context-dependence of extinction, and thus the possibility of relapse with a change of context, may easily remain. Consistent with this possibility, although D-cycloserine allows fear extinction to be learned in fewer trials, it does not appear to prevent or reduce the strength of the renewal effect. Such results further underscore the importance of behavioral research—and behavioral theory—in understanding the effects of drugs on therapy. Nonetheless, the search for drugs that might enhance the *learning* that occurs in therapy situations will continue to be an important area of research.

Another process that might theoretically modify or erase a memory is illustrated by a phenomenon called *reconsolidation*. Newly learned memories are temporarily labile and easy to disrupt before they are consolidated into a more stable form in the brain. The consolidation of memory requires the synthesis of new proteins, and can be blocked by administration of protein synthesis inhibitors (e.g., anisomycin). Animal research suggests that consolidated memories that have recently been reactivated might also return briefly to a similarly vulnerable state; their "reconsolidation" can likewise be blocked by protein synthesis inhibitors. For example, several studies have shown that the reactivation of a conditioned fear by one or two presentations of the CS after a brief fear conditioning experience can allow it to be disrupted by anisomycin. When the CS is tested later, there is little evidence of fear—as if reactivation and then drug administration diminished the strength of the original memory. However, like the effects of extinction, these fear-diminishing effects do not

necessarily mean that the original learning has been destroyed or erased. There is some evidence that fear of the CS that has been diminished this way can still return over time (i.e., spontaneously recover) or with reminder treatments. This sort of result suggests that the effect of the drug is to somehow interfere with retrieval or access to the memory, rather than an actual "reconsolidation."

Generally speaking, the elimination of a behavior after therapy should not be interpreted as erasure of the underlying knowledge. For the time being, it may be safest to assume that after any therapeutic treatment, a part of the original learning may remain in the brain, ready to produce relapse if retrieved. Instead of trying to find treatments that destroy the original memory, another therapeutic strategy might be to accept the possible retention of the original learning, and build therapies that allow the organism to prevent or cope with its retrieval. One possibility is to conduct extinction exposure in the contexts where relapse might be most problematic to the patient, and to encourage retrieval strategies (such as retrieval cues like reminder cards) that might help remind him or her of the therapy experience.

OPERANT/INSTRUMENTAL LEARNING

The Relationship between Behavior and Payoff

As noted earlier, operant learning has many parallels with Pavlovian learning. As one example, extinction also occurs in operant learning if the reinforcer is omitted following training. Although extinction is once again a useful technique for eliminating unwanted behaviors, just as was seen with Pavlovian learning, it does not destroy the original learning—spontaneous recovery, renewal, reinstatement, and rapid reacquisition effects still obtain. Another phenomenon known as *resurgence* also occurs: If a new behavior is taught to replace an old behavior while the old one undergoes extinction, the old behavior can return ("resurge") when the replacement behavior is itself extinguished. All of these effects are consistent with the idea that extinction of an operant behavior, like a Pavlovian response, can be specific to the context in which it is learned.

Although early accounts of instrumental learning beginning with Thorndike emphasized the role of the reinforcer as "stamping in" instrumental the action, more modern approaches tend to view the reinforcer as a sort of guide or motivator of behavior. A modern, "synthetic" view of operant conditioning (see later discussion) holds that the organism associates the action with the outcome in much the way that stimulus-outcome learning is thought to be involved in Pavlovian learning.

Operant behavior has traditionally been studied in its own right, in part because so much human behavior seems voluntary, rather than directly elicited by antecedent cues. Human behavior is influenced by a wide variety

of reinforcers, including social ones. For example, simple attention from teachers or hospital staff members has been shown to reinforce disruptive or problematic behavior in students or patients. In either case, when the attention is withdrawn and redirected toward other activities, the problematic behaviors can decrease (i.e., undergo extinction). Human behavior is also influenced by verbal reinforcers, like praise, and more generally, by *conditioned reinforcers,* such as money, which have no intrinsic value except for the value derived through association with more basic, "primary," rewards. Conditioned reinforcers have been used in schools and institutional settings in so-called *token economies* in which positive behaviors are reinforced with tokens that can be used to purchase other valued items. In more natural settings, reinforcers are always delivered in social relationships, where their effects are dynamic and reciprocal. For example, the relationship between a parent and child is full of interacting and reciprocating operant contingencies in which the delivery (and withholding) of reinforcers and punishers shapes the behavior of each. Like Pavlovian learning, operant learning is always operating and always influencing behavior.

Research on operant conditioning in the laboratory has offered many insights into how action relates to its payoff. In the natural world, few actions are reinforced every time they are emitted; instead, most actions are reinforced only intermittently. There are several well-studied ways in which reinforcers can be scheduled intermittently. In a *ratio reinforcement schedule,* the reinforcer is directly related to the amount of work or responding that the organism emits. That is, there is some work requirement that determines when the next reinforcer will be presented. In a "fixed ratio schedule," every xth action is reinforced; in a "variable ratio schedule," there is an average ratio requirement, but the number of responses required for each successive reinforcer varies. Ratio schedules, especially variable ratio schedules, can generate high rates of behavior, as one observes in behavior directed at a casino slot machine. In an *interval reinforcement schedule,* presentation of each reinforcer depends on the organism emitting the response after some period of time has also elapsed. In a "fixed interval schedule," the first response after x seconds have elapsed is reinforced. In a "variable interval schedule," there is an interval requirement for each reinforcer, but the length of that interval varies. A person checking for e-mail or text messages throughout the day is being reinforced on a variable interval schedule—a new message is not present to reinforce each checking response, but the presence of a new message becomes available after variable time points throughout the day. Interestingly, on interval schedules, response rate can vary substantially without influencing the overall rate of reinforcement. (On ratio schedules, there is a more direct relationship between behavior rate and reinforcement rate.) In part because of this, interval schedules tend to

generate slower response rates than ratio schedules.

Classic research on operant behavior underscores the fact that the performance of any action always involves *choice*. That is, whenever the individual performs a particular behavior, he or she chooses to engage in that action over many other possible alternatives. When choice has been studied by allowing the organism to perform either of two different operant behaviors (paying off with their own separate schedules of reinforcement), the rate of operant behavior depends not only on the behavior's own rate of reinforcement but also on the rate of reinforcement of all other behaviors in the situation. Put most generally, the strength of Behavior 1 (e.g., the rate at which Behavior 1 is performed) is given by:

$$B_1 = K^* R_1/(R_1 + R_O)$$

where B_1 can be seen as the strength of Behavior 1, R_1 is the rate at which B_1 has been reinforced, and R_O is the rate at which all alternative (or "other") behaviors in the environment have been reinforced; K is a constant that corresponds to all behavior in the situation—and may have a different value for different individuals. This principle, known as the *quantitative law of effect,* captures several ideas that are relevant to psychiatrists and clinical psychologists. It indicates that an action can be strengthened either by increasing its rate of reinforcement (R_1) or by decreasing the rate of reinforcement for alternative behaviors (R_O). Conversely, an action can be weakened either by reducing its own rate of reinforcement (R_1) or by increasing the rate of reinforcement for alternative behaviors (R_O). The latter point has an especially important implication: In principle, one can slow down the strengthening of new undesirable behavior by providing an environment that is otherwise rich in reinforcement (high R_O). Thus, an adolescent who experiments with drugs or alcohol would be less likely to engage in this behavior at a high rate (high B_1) if his or her environment were otherwise rich with reinforcers (e.g., provided by extracurricular activities, outside interests, and so forth). Figure 3.3–4 illustrates this point by showing how the rate of an action (B_1) is predicted to increase as a function of its rate of reinforcement in environments with high or low R_O. The principle can also be used in treatment. For example, *contingency management* treatments have been successful at reducing smoking and drug-taking (among other behaviors) by providing vouchers and lottery tickets as reinforcers for alternative behaviors. One effect of such treatments is to increase R_O, although in this case the reinforcers are also usually provided only if the client abstains from the original behavior.

Choice between actions is also influenced by the size of their corresponding reinforcers and how soon the reinforcers will occur. For

example, individuals sometimes have to choose between an action that yields a small but immediate reward (e.g., taking a hit of a drug) versus another that yields a larger but delayed reward (going to a class and earning credit toward a GED). Individuals who choose the more immediate reward are often said to be "impulsive," whereas those who choose the delayed reward are said to exercise "self-control." Of interest, organisms often choose immediate small rewards over delayed larger ones, even though it may be maladaptive to do so in the long run. Such "impulsive" choices are especially difficult to resist when reward is very imminent. A leading explanation is illustrated by Figure 3.3–5. Choice is thought to be determined by the relative value of the two rewards, with that value being influenced by both the reinforcer's size and its delay. As the figure illustrates, the bigger the reinforcer the better is the value, and the more immediate the reinforcer the better too: When a reward is delayed, its value decreases or is "discounted" over time. When offered a choice, the organism will always choose the action that leads to the reward whose value is currently higher. Thus, if choice is made available at Time 1, when both rewards are coming soon, the value of the more immediate reward is higher than the value of the delayed one, and the organism may impulsively choose the immediate reward. If choice is available at Time 2, however, when the rewards are more distant in time, their relative values are reversed. The organism might now choose the delayed reward and demonstrate self-control. One implication is that "impulse control" can be achieved if choice (and commitment to that choice) can be made long before the rewards are about to happen. If the client can commit himself or herself early (further left on the abscissa), it will be easier to choose the larger delayed reward. According to theory, other techniques that would help achieve impulse control would be to further increase the value of the delayed reward (e.g., by providing additional reinforcers for choosing it) and/or decrease the value of the immediate reward (e.g., by adding a cost or punishment for choosing it).

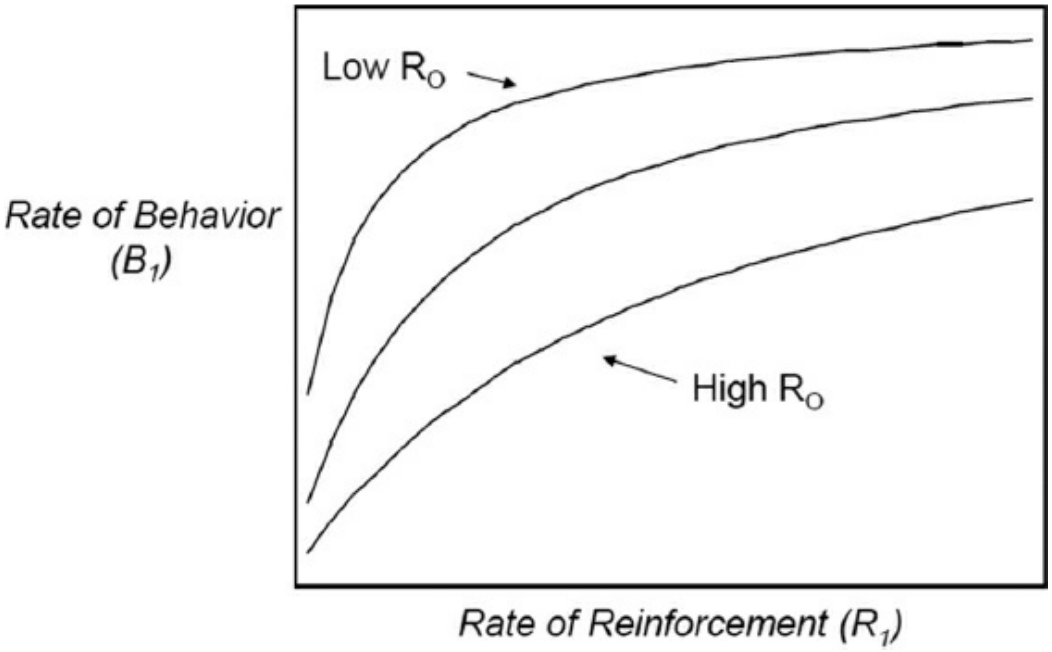

FIGURE 3.3–4. The quantitative law of effect predicts that the extent to which a behavior (B_1) is influenced by its reinforcement rate (R_1) depends crucially on the rate of reinforcement of other behaviors in the background (R_O).

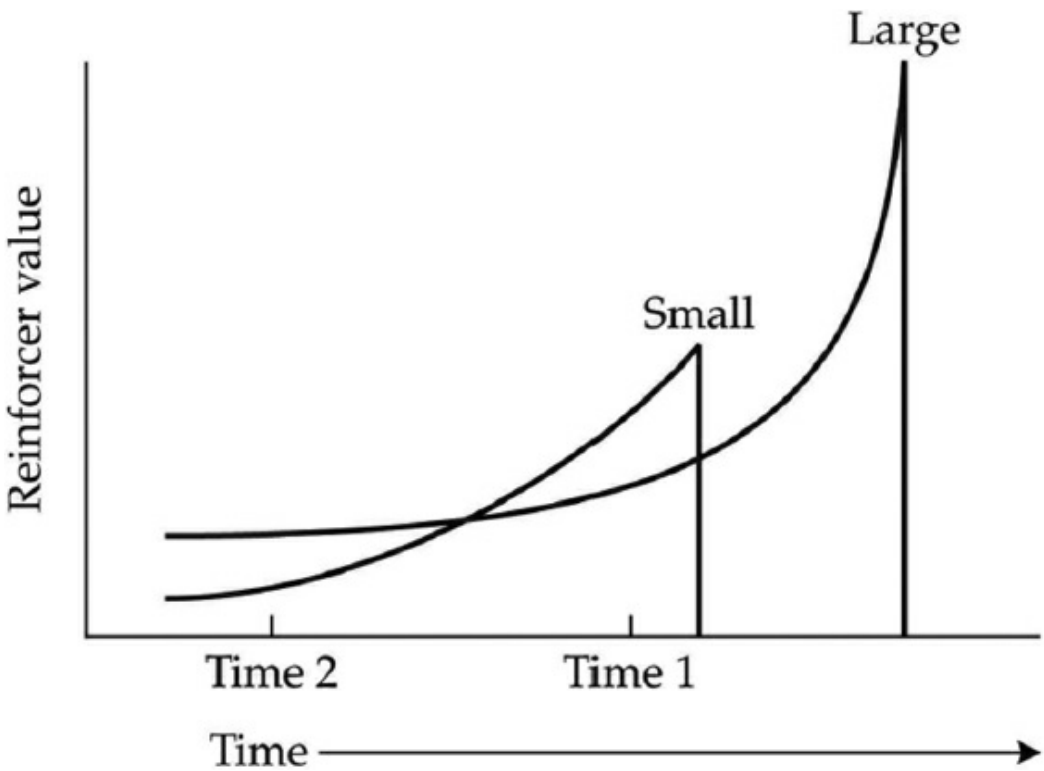

FIGURE 3.3–5. The current value of a reinforcer depends on both its magnitude and its delay. Choice is influenced by current value. Thus, at Time 1, a smaller and more immediate reward has more value than a large delayed reward. However, at Time 2, the larger reward has more value than the smaller reward, despite the same difference in their delay. Behavior is more "impulsive" if choice is made possible at Time 1 than Time 2. (From Bouton ME. *Learning and Behavior: A Contemporary Synthesis*. Sunderland, MA: Sinauer Associates, Inc.; 2007.)

Theories of Reinforcement

It is possible to use the foregoing principles of operant conditioning without knowing in advance what kind of event or stimulus will be reinforcing for the individual patient. None of the reinforcement rules described in the previous section say much about what sorts of events in an organism's world will play the role of reinforcer. B. F. Skinner was famously silent about this. He defined a reinforcer empirically, by considering the effect it had on an operant behavior. A reinforcer was defined as any event that could be shown to increase the strength of an operant if it was made a consequence of the operant. This empirical (some would say "atheoretical") view can be valuable because it allows idiosyncratic reinforcers for idiosyncratic individuals. For instance, if a therapist works with a child who is injuring himself, the approach advises the therapist merely to search for the consequences of the behavior and then manipulate them in order to bring the behavior under control. So if, for example, the child's self-injurious behavior decreases when the parent stops scolding him for doing it, then the scold is the reinforcer, which might seem counterintuitive to everyone (including the parent who thinks the scold should function as a punisher). However, it would also be useful to know what kind of event will reinforce an individual before the therapist has to try everything.

This void is filled by several approaches to reinforcement that allow predictions ahead of time. Perhaps the most useful is the *Premack Principle* (named for researcher David Premack), which claims that, for any individual, reinforcers can be identified by giving the individual a preference test in which she or he is free to engage in any number of activities. The individual might spend the most time engaging in activity A, the second-most time engaged in activity B, and the third-most time engaged in activity C. Behavior A can thus said to be preferred to B and C, and B is preferred to C. The Premack Principle asserts that access to a preferred action will reinforce any action that is less preferred. In the example above, if doing activity C allowed access to doing A or B, activity C will be reinforced—it will increase in strength or probability. Similarly, activity B will be reinforced by activity A (but not C). The principle accepts large individual differences. For example, in an early study, some children given a choice spent more time eating candy than playing pinball, while others spent more time playing pinball then eating candy. Candy eating reinforced pinball playing in the former group. In contrast, pinball playing reinforced candy eating in the latter group. There is nothing particularly special about food (eating) or any particular kind of activity as a possible reinforcer. Any behavior that is preferred to a second behavior will theoretically reinforce the second behavior.

The principle has been refined over the years. At the present point in

time, it is recognized that even a less-preferred behavior can reinforce a more-preferred behavior if the organism has been deprived of doing the low-preferred behavior below its ordinary level. In the foregoing example, even the low-preference activity C could reinforce A or B if it were suppressed for awhile below its baseline level of preference. However, the main implication is that in the long run, a person's reinforcers can be discovered by simply looking at how he or she allocates his or her activities when access to them is free and unconstrained.

Motivational Factors

Instrumental action is often said to be goal-oriented. As Edward Tolman illustrated in many experiments conducted in the 1930s and 1940s, organisms may flexibly perform any of several actions in order to get to a goal; instrumental learning thus provides a variable means to a fixed end. Tolman's perspective on the effects of reinforcers has returned to favor recently. He argued that reinforcers are not necessary for learning, but instead are important for motivating instrumental behavior. The classic illustration of this point is the *latent learning* experiment. Rats received several trials in a complex maze in which they were removed from the maze without reward once they got to a particular goal location. When arriving at the goal was suddenly rewarded, the rats suddenly began working through the maze with very few errors. Thus, they had learned about the maze without benefit of the food reinforcer, but the reinforcer was nonetheless important for motivating them to get through the maze efficiently. The reinforcer was not necessary for learning, but it gave the organism a reason to translate its knowledge into action.

Subsequent research has identified many motivating effects of reward. For example, organisms that have had experience receiving a small reward may show *positive contrast* when they are suddenly reinforced with a larger reward. That is, their instrumental behavior may become more vigorous than that in control subjects who have received the larger reward all along. Conversely, organisms show *negative contrast* when they are switched from a high reward to a lower reward—their behavior becomes weaker than control subjects who have received the same smaller reward all along. Negative contrast involves frustration and emotionality. Both types of contrast are consistent with the idea that the current effectiveness of a reinforcer depends on what the organism has learned to expect; an increase from expectation causes elation, whereas a decrease from expectation causes frustration. There is a sense in which receiving a reward that is smaller than expected might actually seem punishing.

Negative contrast is an example of a *paradoxical reward effect*—a set of behavioral phenomena given the name because they show that rewards can sometimes weaken behavior and that nonreward can sometimes strengthen it. The best known is the *partial reinforcement extinction effect,* in which

actions that have been intermittently (or "partially") reinforced persist longer when reinforcers are completely withdrawn than actions that have been continuously reinforced. The finding is considered paradoxical because an action that has been reinforced (say) half as often as another action may nonetheless be more persistent. One explanation is that the action that has been partially reinforced has been reinforced in the presence of some frustration—and is thus persistent in new adversity or sources of frustration. Other evidence suggests that effortfulness is a dimension of behavior that can be reinforced. That is, human and animal participants that have been reinforced for performing effortful responses learn a sort of "industriousness" that transfers to new behaviors. One implication is that new behaviors learned in therapy will be more persistent over time if high effort has been deliberately reinforced.

The effectiveness of a reinforcer is also influenced by the organism's current motivational state. For example, food is more reinforcing for a hungry organism and water is more reinforcing for a thirsty one. Such results are consistent with many theories of reinforcement (e.g., the Premack principle), because the presence of hunger or thirst would undoubtedly increase the organism's preference ranking for food or water. But recent research indicates that effects of motivational states on instrumental actions are not this automatic. Specifically, if a motivational state is going to influence an instrumental action, the individual needs first to learn how the action's reinforcer will influence the motivational state. The process of learning about the effects the reinforcer has on the motivational state is called *incentive learning*.

Incentive learning is best illustrated by an experimental example. In 1992, Balleine trained rats that were not hungry to lever press to earn a novel food pellet. The animals were then food-deprived and tested for their lever-pressing under conditions where the lever-press no longer produced the pellet. The hunger state had no impact on lever-press rate; that is, hungry rats did not lever-press any more than rats that were not food deprived. However, if the rat had been given separate experience eating the pellets while it was food deprived, during the test it lever-pressed at a high rate. Thus, hunger invigorated the instrumental action only if the animal had previously experienced the reinforcer in that state—which allowed it to learn that the specific substance influenced the state (incentive learning). The interpretation of this result, and others like it, will be developed further later in this section. The main idea is that individuals will perform an instrumental action when they know that it produces an outcome that is desirable in the current motivational state. The clinical implications are underexplored but could be significant. For example, persons who abuse drugs will need to learn that the drug makes them feel better in the withdrawal state before withdrawal will motivate drug seeking. Persons suffering from anxiety might not be motivated to take a beneficial

medication while anxious until they have actually had the opportunity to learn how the medication makes them feel when they are in the anxious state, and a person suffering from depression may need to learn what natural reinforcers really make them feel better while they are depressed. According to theory, direct experience with a reinforcer's effect on depressed mood might be necessary before the person will be interested in performing actions that help ameliorate the depressed state.

PAVLOVIAN AND OPERANT LEARNING TOGETHER

Avoidance Learning

From the previous discussion, it is easy to appreciate that there is more going on in operant learning than merely associating an action with a reinforcer. Indeed, theories of the motivating effects of reinforcers have usually emphasized the fact that Pavlovian CSs in the background are also associated with the reinforcer, and that the expectancy of the reinforcer (or conditioned motivational state) they arouse increases the vigor of the operant response. This is *two-factor* or *two-process theory:* Pavlovian learning occurs simultaneously and motivates behavior during operant learning.

The interaction of Pavlovian and instrumental factors is especially important in understanding avoidance learning (Fig. 3.3–1). In avoidance situations, organisms learn to perform actions that prevent the delivery or presentation of an aversive event. The explanation of avoidance learning is subtle because it is difficult to identify an obvious reinforcer. Although preventing the occurrence of the aversive event is obviously important, how can the non-occurrence of that event reinforce? The answer is that cues in the environment (Pavlovian CSs) come to predict the occurrence of the aversive event, and consequently arouse anxiety or fear. The avoidance response can therefore be reinforced if it escapes from or reduces that fear. Pavlovian and operant factors are thus both important: Pavlovian fear conditioning motivates and allows reinforcement of an instrumental action through its reduction. Escape from fear or anxiety is thought to play a significant role in many human behavior disorders, including the anxiety disorders. Thus, the obsessive-compulsive patient checks or washes his or her hands repeatedly to reduce anxiety, the agoraphobic stays home to escape fear of places associated with panic attacks, and the bulimic learns to vomit after a meal in order to reduce the learned anxiety evoked by eating the meal.

Although two-factor theory remains an important view of avoidance learning, excellent avoidance can be obtained in the laboratory without reinforcement. Specifically, if an animal is required to perform an action that resembles one of its natural and prepared fear responses—so called *species-specific defensive reactions* (*SSDRs*)—it will quickly learn to avoid.

For example, rats will readily learn to freeze (remain motionless) or flee (run to another environment) to avoid shock, two behaviors that have evolved to escape or avoid predation. Freezing and fleeing are also respondents rather than operants; they are controlled by their antecedents (Pavlovian CSs that predict shock) rather than being reinforced by their consequences (escape from fear). Thus, when the rat can use an SSDR to avoid, the only necessary learning is Pavlovian—the rat learns about environmental cues associated with danger, and these arouse fear and evoke natural defensive behaviors including withdrawal (negative sign tracking; Fig. 3.3–2). To learn to perform an action that is not similar to a natural SSDR requires more feedback or reinforcement through fear reduction. A good example is lever-pressing, which is easy for the rat to learn when the reinforcer is a food pellet but difficult to learn when the same action avoids shock. More recent work with avoidance in humans suggests an important role for CS-aversive event and response-no aversive event expectancies. The larger point is that Pavlovian learning is important in avoidance learning; when the animal can avoid with an SSDR, it is the only learning necessary; when the required action is not an SSDR, Pavlovian learning permits the expectation of something bad.

A cognitive perspective on aversive learning is also encouraged by studies of *learned helplessness*. In this phenomenon, organisms exposed to either controllable or uncontrollable aversive events differ in their reactivity to later aversive events. For example, the typical finding is that a subject exposed to inescapable shock in one phase of an experiment is less successful at learning to escape shock with an altogether new behavior in a second phase, whereas subjects exposed to escapable shock are normal. Both types of subject are exposed to the same shock, but its psychological dimension (its controllability) creates a difference, perhaps because subjects exposed to inescapable shock learn the independence of their actions and the outcome. Although this finding (and interpretation) was once seen as a model of depression, the current view is that the controllability of stressors mainly modulates their stressfulness and negative impact. At a theoretical level, the result also implies that organisms receiving instrumental contingencies in which their actions lead to outcomes might learn something about the controllability of those outcomes.

One of the main conclusions of work on aversive learning is thus that there are both biological (i.e., evolutionary) and cognitive dimensions to instrumental learning. The possibility that much instrumental learning can be controlled by Pavlovian contingencies is also consistent with research in which animals have learned to respond for positive reinforcers. For example, pigeon subjects have been widely used in operant learning experiments since the 1940s. In the typical experiment, the bird learns to peck at a plastic disk on a wall of the chamber (a response “key”) in order to

earn food. Although pecking seems to be an operant response, it turns out that the pigeon's peck can be entrained by merely illuminating the key for a few seconds before presenting the reinforcer on a number of trials. Although there is no requirement for the bird to peck the key, the bird will begin to peck at the illuminated key—a Pavlovian predictor of food—anyway. It does so even if the experimenter arranges things so that pecks actually prevent delivery of food (which is otherwise delivered on trials without pecks); here the birds will continue to peck over many trials. (Although the peck has a negative correlation with food, key illumination remains a weakly positive predictor of food.) Thus, this classic "operant" behavior is at least partly a Pavlovian one. Pavlovian contingencies cannot be ignored. When rats are punished with mild footshock for pressing a lever that otherwise produces food, they stop lever-pressing at least partly (and perhaps predominantly) because they learn that the lever now predicts shock and withdraw from it. A child might likewise learn to stay away from the parent who delivers punishment, rather than refrain from performing the punished behavior. A great deal of behavior in operant learning settings may actually be controlled by Pavlovian learning and sign tracking rather than true operant learning.

A Synthetic View of Instrumental Action

The idea, then, is that behavior in any instrumental learning situation is controlled by several hypothetical associations that are illustrated in Figure 3.3–6. As just discussed, much behavior in an instrumental learning arrangement can be controlled by a Pavlovian factor, in which the organism associates background cues (CSs) with the reinforcing outcome (*O*). As discussed earlier, this type of learning can allow the CS to evoke a variety of behaviors as well as emotional reactions (and motivational states) that can additionally motivate instrumental action.

In modern terms, the instrumental factor is represented by the organism learning a direct (and similar) association between the instrumental action (*R*) and the reinforcer (*O*). Evidence for this sort of learning comes from experiments on *reinforcer devaluation* (Fig. 3.3–7). In such experiments, the organism can first be trained to perform two instrumental actions (e.g., pressing a lever and pulling a chain), each paired with a different reinforcer (e.g., food pellet versus a liquid sucrose solution). In a separate second phase, one of the reinforcers (e.g., pellet) is paired with illness, which creates the conditioning of a powerful taste aversion to the reinforcer. In a final test, the organism is returned to the instrumental situation and is allowed to perform either instrumental action. No reinforcers are presented during the test. The result is that the organism no longer emits the action that produced the reinforcer that is now aversive. In order to perform this way, the organism must have (1) learned which action produced which reinforcer and (2) combined this

knowledge with the knowledge that it no longer likes or values that reinforcer. The result cannot be explained by the simpler, more traditional view that reinforcers merely stamp in or strengthen instrumental actions.

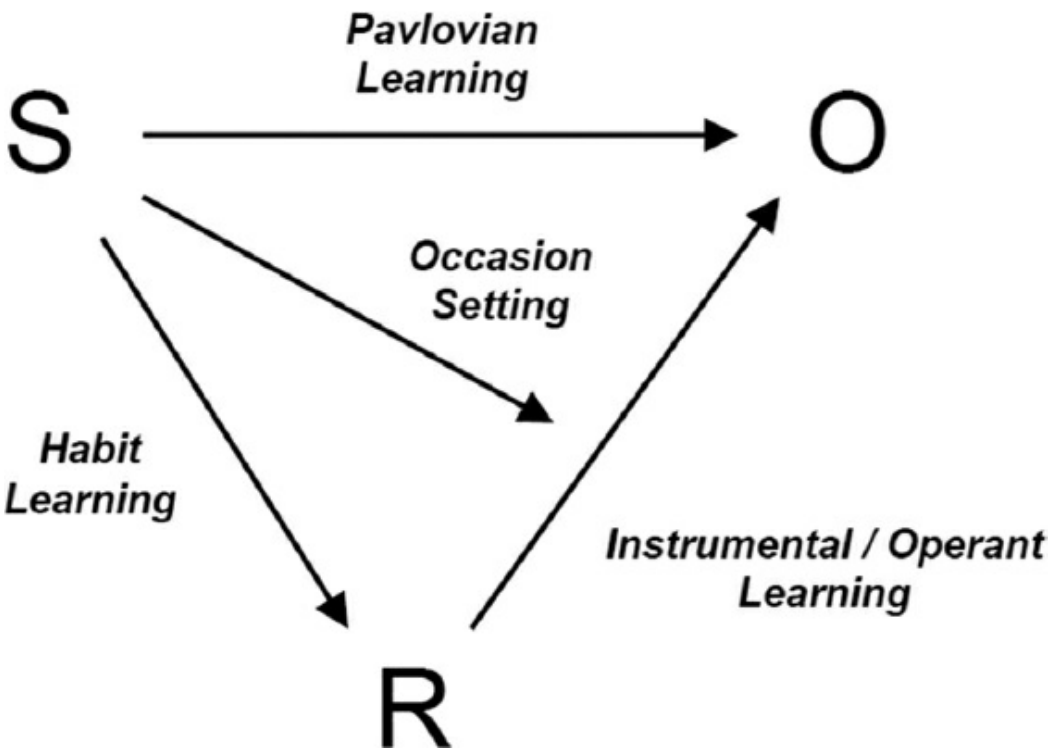

FIGURE 3.3–6. Any instrumental/operant learning situation permits a number of types of learning, which are always occurring all the time. *S,* stimulus in the background; *R,* operant behavior or instrumental action; *O,* biologically-significant outcome (e.g., reinforcer and US).

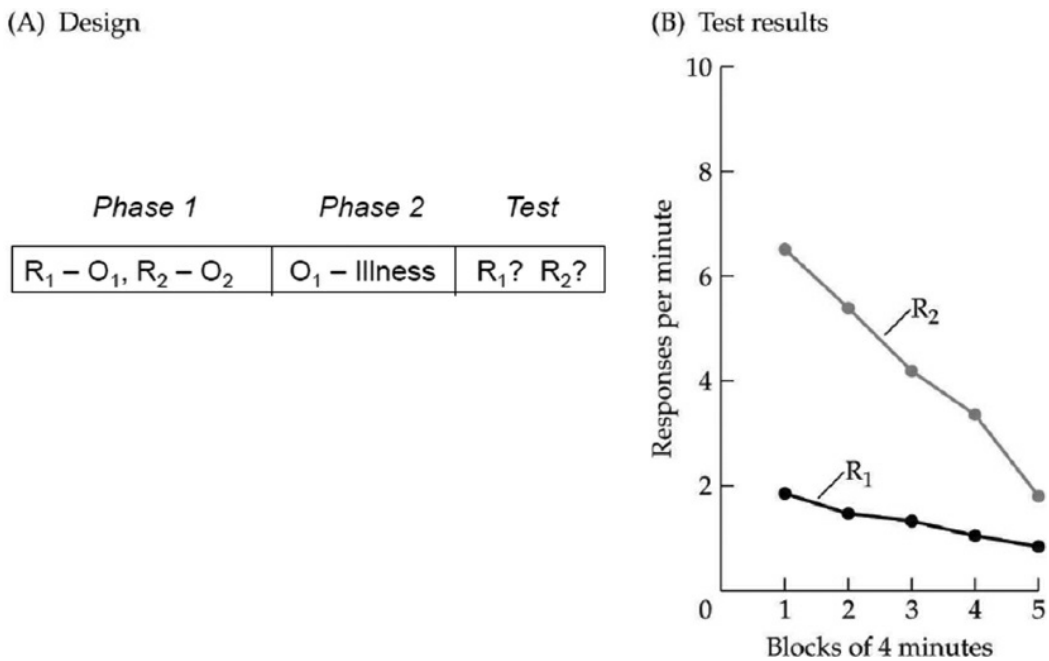

FIGURE 3.3–7. The reinforcer devaluation effect. **A:** Experimental design (see text). **B:** Results of the test session. The result indicates the importance of the response–reinforcer association in operant learning. For the organism to perform the way it does during testing,

it must learn which action leads to which reinforcer and then choose to perform the action that leads to the outcome it currently liked or valued. (Data from Colwill RM, Rescorla RA. *Associative structures in instrumental learning*. In: Bower GH, ed. *The Psychology of Learning and Motivation*, Vol. 20. Orlando, FL: Academic Press; 1986:55–104, based on Bouton ME. *Learning and Behavior: A Contemporary Synthesis*. Sunderland, MA: Sinauer Associates, Inc.; 2007.)

It has already been shown that organisms also need to learn how reinforcers influence a particular motivational state—the process called "incentive learning." Incentive learning is crucially involved in instrumental learning as a process through which the animal learns the value of the reinforcer. Thus, in the reinforcer devaluation experiment shown in Figure 3.3–7, an important thing that occurs in the second phase is that the organism must actually contact the reinforcer and learns that it does not like it. As described earlier, incentive learning is probably always involved in making outcomes (and the associated actions that produce them) more or less desirable.

The Pavlovian factor is represented by the *S-O* association in Figure 3.3–6. As noted earlier, whenever organisms earn reinforcers, he or she can also associate cues in the environment with them. Such *S-O* (Pavlovian) learning can have the many effects on behavior that were reviewed earlier: *S* may evoke approach behaviors, and may elicit many other responses that generally get the organism ready for *O*. *S* may also remind the organism of *O*—it serves as a retrieval cue. In addition, as also mentioned earlier, the presentation of *S* may excite or motivate the instrumental response. This feature of instrumental action is illustrated by *Pavlovian-instrumental transfer:* When a CS that is associated with a US (*O*) is presented while an organism performs an instrumental action that earns *O,* the rate of responding increases. CSs that predict the *O* (excitors) increase the rate of the instrumental behavior, whereas CSs that predict the nonoccurrence of *O* (inhibitors) decrease it. An *S* that predicts *O* can excite instrumental behaviors that lead to other Os that are motivationally similar—such as two different types of food. However, CSs are especially likely to increase responses that are associated with the *same O*. Thus, in addition to motivating instrumental responding in a general way, Pavlovian cues can influence choice of some actions over others. This capacity is presumably exploited by advertisers who create Ss in the form of television commercials, for example, whose presentation can motivate the consumer to buy a particular brand of pizza or a particular brand of beer.

Other experiments have illustrated the other associations to the stimulus that are represented in Figure 3.3–6. In addition to being directly associated with the reinforcer, a stimulus can signal a relationship between an action and an outcome. This is called *occasion setting:* instead of eliciting a response directly, stimuli in operant situations can set the occasion for the operant response. There is good evidence that they do so

by signaling a specific response–reinforcer relationship. For example, in one experiment, rats learned to lever press and chain pull in the presence of a background noise and a background light. When the noise was present, lever pressing yielded a pellet reinforcer and chain pulling yielded sucrose. In contrast, when the light was present, the relations were reversed: lever pressing yielded sucrose and chain pulling yielded pellet. There was evidence that the subjects (rats) learned the corresponding relationships. In a second phase, pellets were associated with illness, so the rat no longer valued the pellet. In a final test, the rats were allowed to lever press or chain pull in extinction in the presence of noise or light present during separate tests. In the presence of the noise, the animal chain pulled rather than lever pressed. When the light was present, the animal lever pressed rather than chain pulled. Thus, the noise informed the rat that lever press yielded pellet and the light informed him that the chain pull did. This is the occasion setting function illustrated in Figure 3.3–6.

It is worth noting that other stimuli besides lights and noises set the occasion for operant behavior. Modern research on learning in animals has underscored the importance of other stimuli such as contexts (see Table 3.3–1), temporal and spatial cues, and certain perception and memory processes. A particularly interesting example of research on the *stimulus control* of operant behavior is categorization. Pigeons can be shown images of different cars, chairs, flowers, and humans on a computer screen positioned on the wall of a Skinner box. Pecking one of four keys in the presence of these images is reinforced in the presence of any picture containing a car, a chair, a flower, or a human. Of interest, as the number of exemplars in each category increases, the pigeon makes more errors as it learns the discrimination. However, more exemplars create better learning in the sense that they are more ready to transfer to new test images—after many examples of each category, the pigeon is more accurate at categorizing (and responding accurately to) new stimuli. One implication is that training new behaviors in a variety of settings or ways will enhance generalization to new situations.

The final association in Figure 3.3–6 is simple *habit learning,* or a direct association between the stimulus and the response. Through this association, the background may elicit the instrumental action quite directly, without the intervening cognition of *R-O* and the valuation of *O*. Although *S-R* learning was once thought to dominate learning, the current Zeitgeist envisions it as developing only after extensive and consistent practice. In effect, actions that have been performed repeatedly (and repeatedly associated with the reinforcer) become automatic and routine. Habit learning has been recognized by psychologists dating back (at least) to William James in the 19th century. It is thought to be functional, because automatic, "mindless" responses can free the behaving individual to process other things besides her moment-to-moment behavior.

Habits can be distinguished from goal-directed actions (those that depend on processing *R-O*) experimentally. One source of evidence is the fact that the reinforcer devaluation effect—which implicates a kind of cognitive mediation of operant behavior—no longer occurs after extensive instrumental training, as if the animal reflexively engages in the response without remembering the actual outcome it produces. It seems reasonable to expect that many pathological behaviors that come into the clinic might also be automatic and blind through repetition. Interestingly, the evidence suggests that the eventual dominance of the *S-R* habit in behavior does not replace or destroy more cognitive mediation by learned *S-O, R-*, and/or *S-*(*R-O*) relations. Under some conditions, even a habitual response might be brought back under the control of the action–reinforcer association. The conversion of actions to habits and the relation of habit to cognition is currently an active area of research.

REFERENCES

Balleine B. Instrumental performance following a shift in primary motivation depends on incentive learning. *J Exp Psychol Anim Behav Processes*. 1992;18:236–250.

Balleine BW, O'Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*. 2010;35:48–69.

Bolles RC. Species-specific defense reactions and avoidance learning. *Psychol Rev*. 1970;71:32–48.

Bouton ME. Classical conditioning and clinical psychology. In: Smelser NJ, Baltes PB, eds. *International Encyclopedia of the Social and Behavioral Sciences*. Vol. 3. Oxford: Elsevier Science; 2001:1942–1945.

Bouton ME. Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biol Psychiatry*. 2002;52:976–986.

**Bouton ME. *Learning and Behavior: A Contemporary Synthesis:* 2nd ed. Sunderland, MA: Sinauer Associates, Inc.; 2016.

**Bouton ME, Mineka S, Barlow DH. A modern learning theory perspective on the etiology of panic disorder. *Psychol Rev*. 2001;108:4–32.

Bouton ME, Vurbic D, Woods AM. D-cycloserine facilitates context-specific fear extinction learning. *Neurobiol Learn Memory*. 2008;90:504–510.

Colwill RM, Rescorla RA. Associative structures in instrumental learning. In: Bower GH, ed. *The Psychology of Learning and Motivation*. Vol. 20. Orlando, FL: Academic Press; 1986:55–104.

Colwill RM, Rescorla RA. Evidence for the hierarchical structure of instrumental learning. *Animal Learn Behav*. 1990;18:71–82.

Davis DR, Kurti AN, Skelly JM, Redner R, White TJ, Higgins ST. A review of the literature on contingency management in the treatment of substance use disorders, 2009–2014. *Preventive Medicine*. 2016;92:36–46.

Davis M, Ressler K, Rothbaum BO, Richardson R. Effects of D-cycloserine on extinction: translation from preclinical to clinical work. *Biol Psychiatry*. 2006;60:369–375.

Dickinson A, Balleine BW. The role of learning in the operation of motivational systems. In: Pashler H, Gallistel H, eds. *Steven's Handbook of Experimental Psychology*. Vol. 3. Learning, motivation, and emotion. Hoboken, NJ: John Wiley & Sons; 2002.

Dworkin BR. *Learning and Physiological Regulation*. Chicago, IL: University of Chicago Press; 1993.

Eikelboom R, Stewart J. Conditioning of drug-induced physiological responses. *Psychol Rev*. 1982;89:507–528.

Eisenberger R. Learned industriousness. *Psychol Rev*. 1992;99:248–267.

Haselgrove M, Hogarth L, eds. *Clinical Applications of Learning Theory*. New York: Psychology Press; 2012.

Hogarth L, Balleine BW, Corbit LH, Killcross S. Associative learning mechanisms underpinning the transition from recreational drug use to addiction. *Ann N Y Acad Sci*. 2013;1282:12–24.

Kamin LJ. Predictability, surprise, attention and conditioning. In Campbell BA, Church RM, eds. *Punishment and Aversive Behavior*. New York: Appleton-Century-Crofts; 1969.

Lattal KM, Abel T. Behavioral impairments caused by injections of the protein synthesis inhibitor anisomycin after contextual retrieval reverse with time. *Proc Nat Acad Sci*. 2004;101:4667–4672.

Maier SF, Seligman, MEP. Learned helplessness at fifty: Insights from neuroscience. *Psychological Review*. 2016;123:349–367.

McDowell JJ. On the theoretical and empirical status of the matching law and matching theory. *Psychol Bull*. 2013;139:1000–1028.

McSweeney FK, Murphy ES. *The Wiley Blackwell Handbook of Operant and Classical Conditioning*. Chichester, UK: John Wiley & Sons; 2014.

Mineka S. Evolutionary memories, emotional processing and the emotional disorders. In: Medin D, ed. *The Psychology of Learning and Motivation*. Vol. 28. San Diego: Academic Press; 1992:161–206.

Nader K, Hardt O. A single standard for memory: the case for reconsolidation. *Nat Rev Neurosci*. 2009;10:224–234.

**O'Donohue W, ed. *Learning and Behavior Therapy*. Boston, MA: Allyn and Bacon; 1998.

Pearce JM, Mackintosh NJ. Two theories of attention: a review and possible integration. In: Mitchell CJ, Le Pelley ME, eds. *Attention and Associative Learning: From Brain to Behaviour*. Oxford: Oxford University Press; 2010.

Premack D. Reinforcement theory. In: Levine D, ed. *Nebraska Symposium on Motivation*. Vol. 13. Lincoln, NE: University of Nebraska Press; 1965.

Rachlin H. Self-control. *Behaviorism*. 1974;3:94–107.

Rescorla RA. Pavlovian conditioning and its proper control procedures. *Psychol Rev*. 1967;74:71–80.

**Rescorla RA. Pavlovian conditioning: it's not what you think it is. *Am Psychol*. 1988;43:151–160.

Rescorla RA, Wagner AR. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF, eds. *Classical Conditioning II*. New York: Appleton-Century-Crofts; 1972.

Saunders BT, Robinson TE. Individual variation in resisting temptation: implications for

addiction. *Neurosci Biobehav Rev*. 2013;37:1955–1975.

Shettleworth SJ. *Cognition, Evolution, and Behavior*. 2nd ed. Oxford: Oxford University Press; 2010.

Siegel S. Pharmacological conditioning and drug effects. In Goudie AJ, Emmett-Oglesby MW, eds. *Psychoactive Drugs: Tolerance and Sensititzation. Contemporary Neuroscience*. Clifton, NJ: Humana Press; 1989.

Spreat S, Spreat SR. Learning principles. In: Voith V, Borchelt PL, eds. *Veterinary Clinics of North America: Small Animal Practice*. Philadelphia, PA: W. B. Saunders; 1982:593–606.

Wagner AR, Brandon SE. Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In: Klein SB, Mowrer RR, eds. *Contemporary Learning Theories: Pavlovian Conditioning and the Status of Traditional Learning Theory*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 1989:149–189.

Wagner AR, Logan FA, Haberlandt K, Price T. Stimulus selection in animal discrimination learning. *J Exp Psychol*. 1968;76:177–186.

Wasserman EA, Bhatt RS. Conceptualization of natural and artificial stimuli by pigeons. In: Honig WK, Fetterman JG, eds. *Cognitive Aspects of Stimulus Control*. Hillsdale, NJ: Lawrence Erlbaum, Inc.; 1992.

Wolpe J. *Psychotherapy by Reciprocal Inhibition*. Stanford: Stanford University Press; 1958.

▲ 3.4 Biology of Memory

JOEL L. VOSS, PH.D., AND KEN A. PALLER, PH.D.

The topic of memory is fundamental to the discipline of psychiatry. Memory is the glue that binds mental life, the scaffolding for personal history. Personality is in part an accumulation of habits that have been acquired, many very early in life, which create dispositions and influence behavior. In the same sense, the neuroses are often products of learning—anxieties, phobias, and maladaptive behaviors that result from particular experiences. Psychotherapy itself is a process by which new habits and skills are acquired through the accumulation of new experiences. In this sense, memory is at the theoretic heart of psychiatry's concern with personality, the consequences of early experience, and the possibility of growth and change.

Memory is also of clinical interest because disorders of memory and complaints about memory are common in neurological and psychiatric illness. Memory impairment is also a side effect of certain treatments, such as electroconvulsive therapy. Accordingly, the effective clinician needs to understand something of the biology of memory, the varieties of memory dysfunction, and how memory can be evaluated.

FROM SYNAPSES TO MEMORY

Memory is a special case of the general biological phenomenon of *neural*

plasticity. Neurons can show history-dependent activity by responding differently as a function of prior input, and this plasticity of nerve cells and synapses is the basis of memory. In the last decade of the 19th century, researchers proposed that the persistence of memory could be accounted for by nerve cell growth. This idea has been restated many times since, and current understanding of the synapse as the critical site of change is founded on extensive experimental studies in animals with relatively simple nervous systems. Experience can lead to structural change at the synapse, including alterations in the strength of existing synapses and alterations in the number of synaptic contacts along specific pathways.

Plasticity

Neurobiological evidence supports two basic conclusions. First, short-lasting plasticity, which may last for seconds or minutes, depends on specific synaptic events, including an increase in neurotransmitter release. Second, long-lasting memory depends on new protein synthesis, physical growth of neural processes, and an increase in the number of synaptic connections.

A major source of information about memory has come from extended study of the marine mollusk *Aplysia californica*. Individual neurons and connections between neurons have been identified, and the wiring diagrams of some simple behaviors have been described. *Aplysia* is capable of associative learning (including classic conditioning and operant conditioning) and nonassociative learning (habituation and sensitization). *Sensitization* had been studied using the gill-withdrawal reflex, a defensive reaction whereby tactile stimulation causes the gill and siphon to retract. When tactile stimulation is preceded by sensory stimulation to the head or tail, gill withdrawal is facilitated. The cellular changes underlying this sensitization begin when a sensory neuron activates a modulatory interneuron, which enhances the strength of synapses within the circuitry responsible for the reflex. This modulation depends on a second-messenger system whereby intracellular molecules (including cyclic adenosine monophosphate or cAMP and cAMP-dependent protein kinase) lead to enhanced transmitter release lasting for minutes in the reflex pathway. Both short- and long-lasting plasticity within this circuitry are based on enhanced transmitter release. The long-lasting change uniquely requires the expression of genes and the synthesis of new proteins. Synaptic tagging mechanisms allow gene products that are delivered throughout a neuron to increase synaptic strength selectively at recently active synapses. In addition, the long-term change, but not the short-term change, is accompanied by growth of neural processes of neurons within the reflex circuit.

In vertebrates, memory cannot be studied quite as directly as in the simple nervous system of *Aplysia*. Nevertheless, it is known that behavioral

manipulations can also result in measurable changes in the brain's architecture. For example, rats reared in enriched as opposed to ordinary environments show an increase in the number of synapses ending on individual neurons in neocortex. These changes are accompanied by small increases in cortical thickness, in the diameter of neuronal cell bodies, and in the number and length of dendritic branches. Behavioral experience thus exerts powerful effects on the wiring of the brain.

Many of these same structural changes have been found in adult rats exposed to an enriched environment, as well as in adult rats given extensive maze training. In studies of maze training, vision was restricted to one eye, and the corpus callosum was transected to prevent information received by one hemisphere from reaching the other hemisphere. The result was that structural changes in neuronal shape and connectivity were observed only in the trained hemisphere. This rules out a number of nonspecific influences, including motor activity, indirect effects of hormones, and overall level of arousal. Long-term memory in vertebrates is thought to be based on morphological growth and change, including increases in synaptic strength along particular pathways.

Long-Term Potentiation

The phenomenon of *LTP* is a candidate mechanism for mammalian long-term memory. LTP is observed when a postsynaptic neuron is persistently depolarized after a high-frequency burst of presynaptic neural firing. LTP has a number of properties that make it suitable as a physiological substrate of memory. It is established quickly and then lasts for a long time. It is associative, in that it depends on the co-occurrence of presynaptic activity and postsynaptic depolarization. It occurs only at potentiated synapses, not all synapses terminating on the postsynaptic cell. Finally, LTP occurs prominently in the hippocampus, a structure important for memory.

The induction of LTP is known to be mediated postsynaptically and to involve activation of NMDA receptors, which permit the influx of calcium into postsynaptic neurons. LTP is maintained by an increase in the number of postsynaptic AMPA (non-NMDA) receptors and also possibly by increased transmitter release. Postsynaptic NMDA-mediated changes also support the opposing phenomenon of *long-term depression* (LTD), which involves temporary decreases in synaptic efficacy following low-frequency synaptic stimulation. The encoding of long-term memories could involve modification of synaptic connections through LTP and LTD.

A promising method for elucidating molecular mechanisms of memory relies on introducing specific mutations into the genome. By deleting a single gene, mice can be produced with specific receptors or cell signaling molecules inactivated or altered. For example, in mice with a selective deletion of NMDA receptors in the CA1 field of the hippocampus, many

aspects of CA1 physiology remain intact, but the CA1 neurons do not exhibit LTP, and memory impairment is observed in behavioral tasks. Genetic manipulations introduced reversibly in the adult are particularly advantageous in that specific molecular changes can be induced in developmentally normal animals. In particular, by using optogenetics it is possible to engineer specific populations of neurons such that certain wavelengths of light systematically modify their activity, with specific mutations allowing for light-induced excitation or inhibition. This optogenetic approach for controlling neural activity with high spatial and temporal precision has been used to provide evidence for the involvement of LTP and LTD at specific synapses in long-term memory for conditioned fear.

Associative Learning

The study of *classical conditioning* has provided many insights into the biology of memory. Classical conditioning has been especially well studied in rabbits using a tone as the conditioned stimulus and an air puff to the eye (which automatically elicits a blink response) as the unconditioned stimulus. Repeated pairings of the tone and the air puff lead to a conditioned response, in that the tone alone elicits an eye blink. Reversible lesions of the deep nuclei of the cerebellum eliminate the conditioned response without affecting the unconditioned response. These lesions also prevent initial learning from occurring, and, when the lesion is reversed, rabbits learn normally. Thus, the cerebellum contains essential circuitry for the learned association. The relevant plasticity appears to be distributed between the cerebellar cortex and the deep nuclei.

An analogous pattern of cerebellar plasticity is thought to underlie motor learning in the vestibulo-ocular reflex and, perhaps, associative learning of motor responses in general. Based on the idea that learned motor responses depend on coordinated control of changes in timing and strength of response, it has been suggested that synaptic changes in the cerebellar cortex are critical for learned timing, whereas synaptic changes in the deep nuclei are critical for forming an association between a conditioned stimulus and an unconditioned stimulus.

Fear conditioning and fear-potentiated startle are types of learning that serve as useful models for anxiety disorders and related psychiatric conditions. For example, mice exhibit freezing behavior when returned to the same context in which aversive shock was delivered on an earlier occasion. This type of learning depends on the encoding of contextual features of the learning environment. Acquiring and expressing this type of learning requires neural circuits that include both the amygdala and the hippocampus. The amygdala may be important for associating negative affect with new stimuli. The hippocampus may be important for representing the context. With extinction training, when the context is no

longer associated with an aversive stimulus, the conditioned fear response fades. Frontal cortex is thought to play a key role in extinction.

CORTICAL ORGANIZATION OF MEMORY

One fundamental question concerns the locus of memory storage in the brain. In the 1920s, Karl Lashley searched for the site of memory storage by studying the behavior of rats following removal of different amounts of cerebral cortex. He recorded the number of trials needed to relearn a maze problem that rats learned prior to surgery, and he found that the deficit was proportional to the amount of cortex removed. The deficit did not seem to depend on the particular location of cortical damage. Lashley concluded that the memory resulting from maze learning was not localized in any one part of the brain but instead was distributed equivalently over the entire cortex.

Subsequent research has led to reinterpretations of these results. Maze learning in rats depends on different types of information, including visual, tactile, spatial, and olfactory information. Importantly, neurons that process these various types of information are segregated into different areas of the rat cerebral cortex, and memory storage is segregated in a parallel manner. Thus, the correlation between maze learning abilities and lesion size that Lashley observed is a result of progressive encroachment of larger lesions on specialized cortical areas serving the many components of information processing relevant for maze learning.

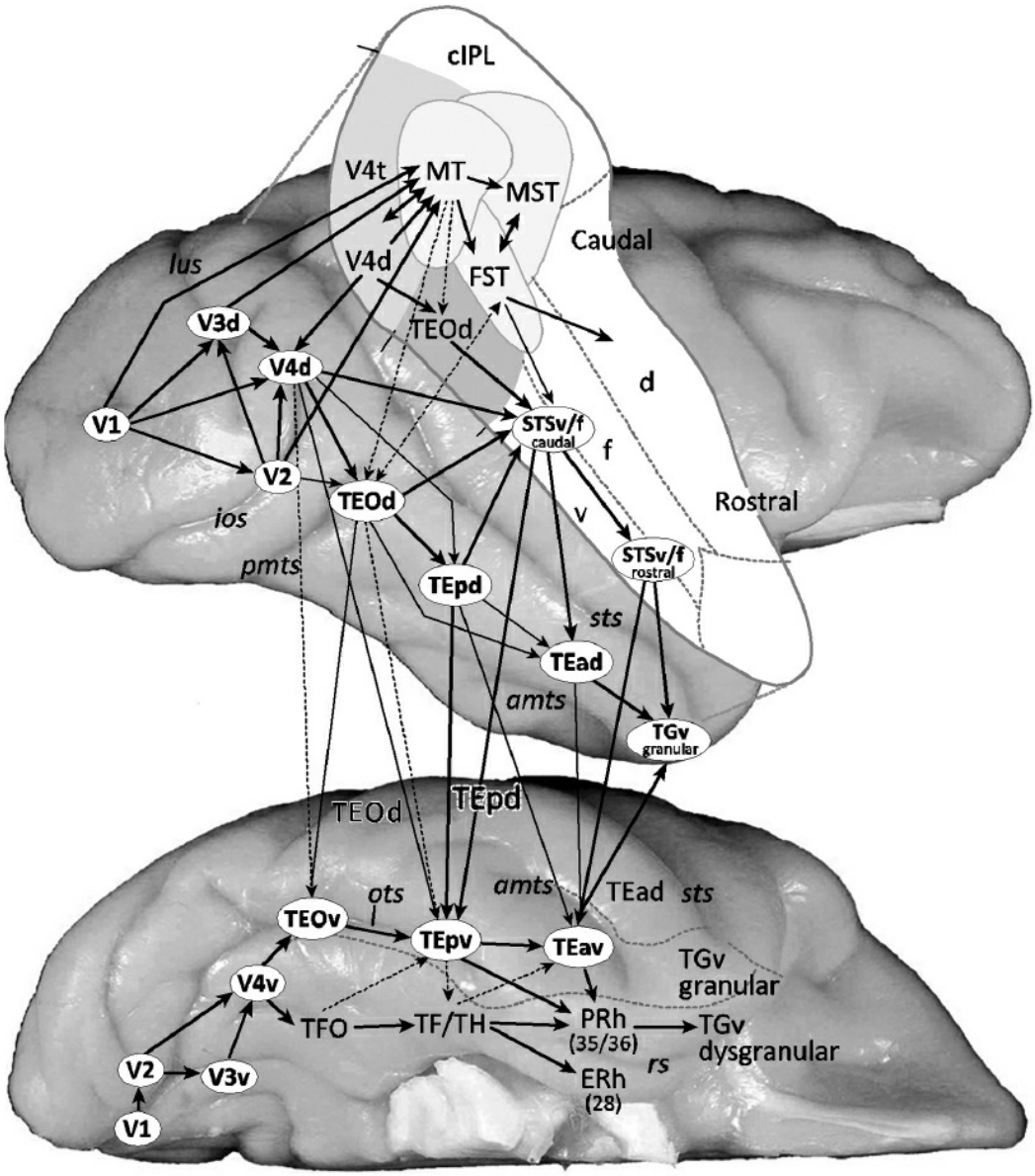

FIGURE 3.4–1. Summary of cortical visual areas and some of their connections. Two major cortical pathways begin from area V1, also known as striate cortex or primary visual cortex, and which receives visual input from the retina by way of the lateral geniculate nucleus of the thalamus. The processing stream for object vision follows a ventral route into the temporal lobe, and the processing stream for spatial vision follows a dorsal route into the parietal lobe, which then projects into the temporal lobe. The convergence of these two processing streams in the medial temporal lobe, including hippocampus, provides an anatomical rationale for the key role of this area in constructing and retrieving memories comprised different sorts of episodic and semantic information. Arrows represent reciprocal anatomical connections, with line width representing the relative strength of interconnection and dotted lines indicating very sparse projections. For anatomical abbreviations, see original article. (From Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences.* 2013;17:26–49.)

The functional organization of the mammalian cerebral cortex has been revealed through neuropsychological analyses of deficits following brain damage and through physiological studies of intact brains. The cortical

areas responsible for processing and storing visual information have been studied most extensively in nonhuman primates. Nearly one-half of the primate neocortex is specialized for visual functions.

As shown in Figure 3.4–1, the cortical pathways for visual information processing begin in primary visual cortex (V1) and proceed from there along parallel pathways or streams. One stream projects ventrally to inferotemporal cortex and is specialized for processing information concerning the identification of visual objects. Another stream projects dorsally to parietal cortex and is specialized for processing information about spatial location.

Specific visual processing areas in the dorsal and ventral streams, together with areas in prefrontal cortex, register the immediate experience of perceptual processing. The results of perceptual processing are first available as *immediate memory*. Immediate memory refers to the amount of information that can be held in mind (like a telephone number) so that it is available for immediate use. Immediate memory can be extended in time by rehearsing or otherwise manipulating the information, in which case what is stored is now said to be in *working memory*.

Regions of visual cortex in forward portions of the dorsal and ventral streams serve as the ultimate repositories of visual memories. Inferotemporal cortex, for example, lies at the end of the ventral stream, and inferotemporal lesions lead to selective impairments in both visual object perception and visual memory. Yet, such lesions do not disrupt elementary visual functions, such as acuity. Electrophysiological studies in the monkey show that neurons in area TE, which is one part of inferotemporal cortex, register specific and complex features of visual stimuli, such as shape, and respond selectively to patterns and objects. Inferotemporal cortex can thus be thought of as both a higher-order visual processing system and a storehouse of the visual memories that result from that processing.

In sum, memory is distributed and localized in the cerebral cortex. Memory is distributed in the sense that, as Lashley concluded, there is no cortical center dedicated solely to the storage of memories. Yet, memory is localized in the sense that different aspects or dimensions of events are stored at specific cortical sites—namely, in the same regions that are specialized to analyze and process what is to be stored.

MEMORY AND AMNESIA

The principle that the functional specialization of cortical regions determines both the locus of information processing and also the locus of information storage does not provide a complete account of the organization of memory in the brain. If it did, then brain injury would always include a difficulty in memory for a restricted type of information along with a loss of ability to process information of that same type. This

kind of impairment occurs, for example, in the aphasias and the agnosias. But there is another kind of impairment that can occur as well called amnesia.

The hallmark of amnesia is a loss of new learning ability that extends across all sensory modalities and stimulus domains. This *anterograde amnesia* can be explained by understanding the role of brain structures critical for acquiring information about facts and events. Typically, anterograde amnesia occurs together with *retrograde amnesia,* a loss of knowledge that was acquired before the onset of amnesia. Retrograde deficits often have a temporal gradient, following a principle known as Ribot law; deficits are most severe for information that was most recently learned.

A patient with a presentation of amnesia exhibits severe memory deficits in the context of preservation of other cognitive functions, including language comprehension and production, reasoning, attention, immediate memory, personality, and social skills. The selectivity of the memory deficit in these cases implies that intellectual and perceptual functions of the brain are separated from the capacity to maintain in memory the records that ordinarily result from engaging in intellectual and perceptual work.

Specialized Memory Function

Amnesia can result from damage to the medial portion of the temporal lobe, or from damage to regions of the midline diencephalon. Studies of a patient with severe amnesia, known as HM, stimulated intensive investigation of the role of the medial temporal lobe in memory.

HM became amnesic in 1953, at 27 years of age, when he sustained a bilateral resection of the medial temporal lobe to relieve severe epilepsy. The removal included approximately one-half of the hippocampus, the amygdala, and most of the neighboring entorhinal and perirhinal cortices (Fig. 3.4–2). After the surgery, HM's seizure condition was much improved, but he experienced profound forgetfulness. His intellectual functions were generally preserved. For example, HM exhibited normal immediate memory, and he could maintain his attention during conversations. After an interruption, however, HM could not remember what had recently occurred. HM's dense amnesia was permanent and debilitating. In HM's words, he felt as if he were just waking from a dream, because he had no recollection of what had just taken place.

Findings from human amnesia stimulated research to develop models of amnesia in experimental animals. In monkeys, many parallels to human amnesia have been demonstrated following surgical damage to anatomical components of the medial temporal lobe. Cumulative study of the resulting memory impairment eventually identified the medial temporal structures and connections that are crucial for memory. These include the hippocampus—which includes the dentate gyrus, hippocampal fields CA1, CA2, and CA3, and the subiculum—and also the adjacent cortical regions

including the entorhinal, perirhinal, and parahippocampal cortices.

Detailed study of patients with amnesia offers unique insights into the nature of memory and its organization in the brain. An extensive series of informative studies, for example, described the memory impairment of patient EP.

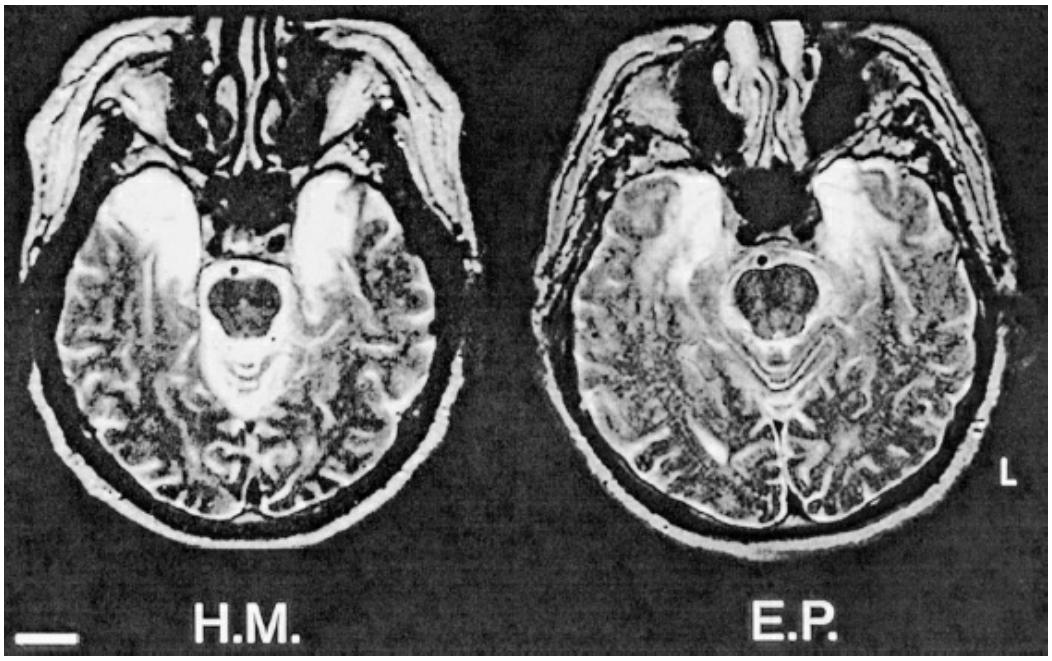

FIGURE 3.4–2. Structural magnetic resonance images of the brains of patients HM and EP through the level of the temporal lobe. Damaged tissue is indicated by bright signal in these T2-weighted axial images. Both patients sustained extensive damage to medial temporal structures, as the result of surgery for epilepsy in HM, and as the result of viral encephalitis in EP. Scale bar: 2 cm. L, left side of the brain. (H.M. image from Corkin S, Amaral EG, González RG, Johnson KA, Hyman BT. H.M.'s medial temporal lobe lesion: findings from magnetic resonance imaging. *J Neurosci.* 1997;17:3964. E.P. image from Stefanacci L, Buffalo EA, Schmolck H, Squire LR. Profound amnesia after damage to the medial temporal lobe: a neuroanatomical and neuropsychological profile of patient E.P. *J Neurosci.* 2000;20:7024.)

EP was diagnosed with herpes simplex encephalitis at 72 years of age. Damage to the medial temporal lobe region (Fig. 3.4–2) produced a persistent and profound amnesia. During testing sessions, EP is cordial and talks freely about his life experiences, but he relies exclusively on stories from his childhood and early adulthood. He will repeat the same story many times. Strikingly his performance on tests of recognition memory is no better than would result from guessing (Fig. 3.4–3A). Tests involving facts about his life and autobiographical experiences revealed poor memory for the time leading up to his illness but normal memory for his childhood (Fig. 3.4–3B). EP also has good spatial knowledge about the town in which he lived as a child, but he has been unable to learn the layout of the neighborhood where he lived only after he became amnesic (Fig. 3.4–3C).

Given the severity of the memory problems experienced by EP and other patients with amnesia, it is noteworthy that these patients nonetheless perform normally on certain kinds of memory tests, as described below. The impairment selectively concerns memory for factual

knowledge and autobiographical events, collectively termed *declarative memory*. Amnesia presents as a global deficit in that it involves memory for information presented in any sensory modality, but the deficit is limited in that it covers only memory for facts and events.

In another patient, RB, an episode of global ischemia after cardiac surgery led to moderately severe anterograde amnesia and retrograde amnesia, less severe than in patients HM and EP After his death 5 years later, detailed histological study of his brain revealed that amnesia in RB was the result of a circumscribed bilateral lesion of hippocampal field CA1 (Fig. 3.4–4). Similar histological evidence of hippocampal damage has been obtained in other patients with amnesia.

Hippocampal pathology in patients with amnesia can also be revealed using high-resolution magnetic resonance imaging (MRI). Such studies indicate that damage limited to the hippocampus results in clinically significant memory impairment. In addition to the hippocampus, other medial temporal lobe regions also make critical contributions to memory. Thus, a moderately severe memory impairment results from CA1 damage (as in patient RB), whereas a more profound and disabling amnesia results from medial temporal lobe damage that includes the hippocampus and adjacent cortex (as in patients HM and EP). Memory impairment due to medial temporal lobe damage is also typical in patients with early Alzheimer disease or amnestic mild cognitive impairment. As Alzheimer disease progresses, the pathology affects many cortical regions and produces substantial cognitive deficits in addition to memory dysfunction.

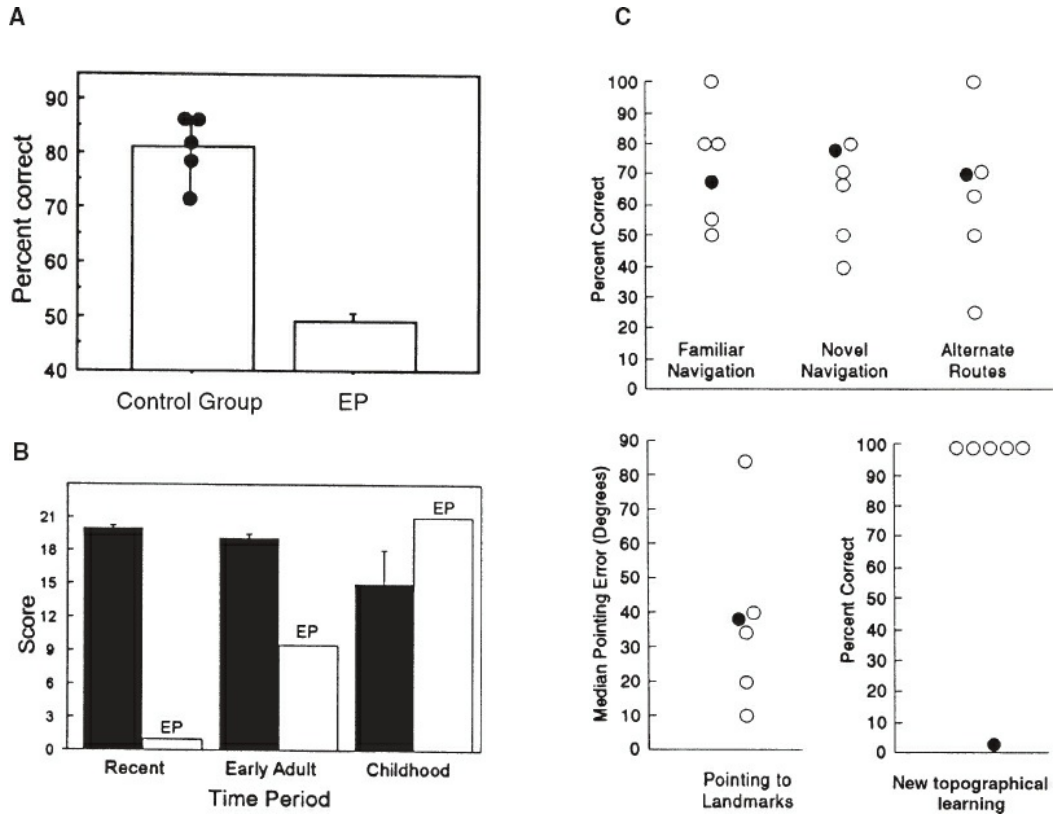

FIGURE 3.4–3. Formal test results from patient EP show severe anterograde and retrograde deficits, with intact remote memory. **A:** Scores were combined from 42 different tests of recognition memory for words given to patient EP and a group of five healthy control subjects. The testing format was either two-alternative forced choice or yes–no recognition. Brackets for EP indicate standard error of the mean. Data points for the control group indicate each participant's mean score across all 42 recognition memory tests. EP's average performance (49.3 percent correct) was not different from chance and was approximately five standard deviations (*SD*s) below the average performance of control subjects (81.1 percent correct; *SD,* 6.3). **B:** Autobiographical remembering was quantified by using a structured interview known as the *Autobiographical Memory Interview.* Items assessed personal semantic knowledge (maximum score: 21 for each time period). Performance for the recent time period reflects poor memory for information that could have been acquired only subsequent to the onset of his amnesia. For EP, performance for the early adult period reflects retrograde memory deficits. Performance for the childhood period reflects good remote memory. Similar results for semantic and episodic remembering were obtained from these time periods. (Data from Kopelman MD, Wilson BA, Baddeley AD. The autobiographical memory interview: a new assessment of autobiographical and personal semantic memory in amnesic patients. *J Clin Exp Neuropsychol.* 1989;5:724 and Reed JM, Squire LR. Retrograde amnesia for facts and events: findings from four new cases. *J Neurosci.* 1998;18:3943.) **C:** Assessments of spatial memory demonstrated EP's good memory for spatial knowledge from his childhood along with extremely poor new learning of spatial information. Performance was compared to that of five individuals (*empty circles*) who attended EP's high school at the same time as he did, lived in the region over approximately the same time period, and, like EP (*filled circle*), moved away as young adults. Normal performance was found for navigating from home to different locations in the area (familiar navigation), between different locations in the area (novel navigation), and between these same locations when a main street was blocked (alternative routes). Subjects were also asked to point to particular locations while imagining themselves in a particular location (pointing to landmarks), or they were asked about locations in the neighborhoods in which they currently lived (new topographical learning). EP showed difficulty only in this last test, because he moved to his current residence after he became amnesic. (Data from Teng E, Squire LR. Memory for places learned long ago is intact after hippocampal damage. *Nature.* 1999;400:675.) (Adapted from Stefanacci L, Buffalo EA, Schmolck H, Squire LR. Profound amnesia after damage to the medial temporal lobe: a neuroanatomical and neuropsychological profile of patient E.P. *J Neurosci.* 2000;20:7024.)

Amnesia can also result from damage to structures of the medial diencephalon. The critical regions damaged in diencephalic amnesia include the mammillary nuclei in the hypothalamus, the dorsomedial nucleus of the thalamus, the anterior nucleus, the internal medullary lamina, and the mammillothalamic tract. However, uncertainty remains regarding which specific lesions are required to produce diencephalic amnesia. *Alcoholic Korsakoff syndrome* is the most prevalent and best-studied example of diencephalic amnesia, and in these cases damage is found in brain regions that may be especially sensitive to prolonged bouts of thiamine deficiency and alcohol abuse. Patients with alcoholic Korsakoff syndrome typically exhibit memory impairment due to a combination of diencephalic damage and frontal lobe pathology. Frontal damage alone produces characteristic cognitive deficits along with certain memory problems (e.g., in effortful retrieval and evaluation); in Korsakoff syndrome the pattern of deficits thus extends beyond what is commonly found in

other cases of amnesia (Table 3.4–1).

Another important medial temporal lobe structure is the amygdala. The amygdala is involved in the regulation of much of emotional behavior. In addition to mediating associative learning of emotional stimuli (e.g., fear conditioning and extinction), modulatory effects of projections from the amygdala to the neocortex are responsible for producing enhanced memory for emotional or arousing events compared to neutral events. Furthermore, amygdala biochemical and structural abnormalities have been linked to abnormal emotional memory in conditions such as fear disorders and posttraumatic stress disorder.

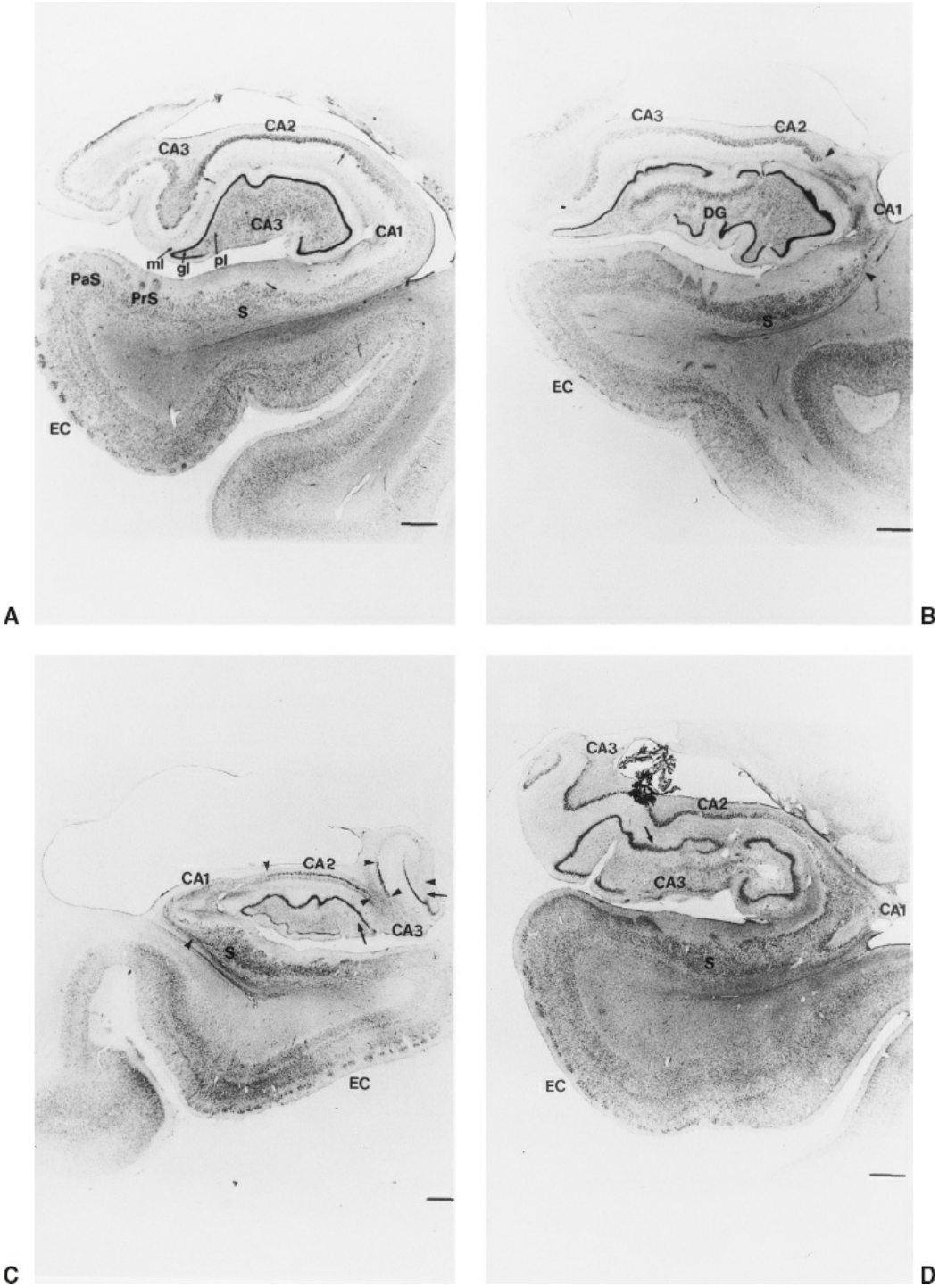

FIGURE 3.4–4. Coronal sections through the hippocampal region stained with thionin in a healthy individual (**A**) and three amnesic patients with bilateral damage to the hippocampal formation, patients GD, LM, and WH (**B–D**). The hippocampus proper can be divided into three distinct fields: designated CA1, CA2, and CA3. The CA1 field extends to the subiculum (S). Other structures include the dentate gyrus (DG), presubiculum (PrS), parasubiculum (PaS), and entorhinal cortex (EC). In **B,** damage included CA1, marked by arrowheads. In **C,** damage included CA1, CA2, CA3, DG, and EC. Arrowheads indicate the borders of the CA fields. Arrows indicate loss of polymorphic cells in DG. In **D,** damage included CA1, CA2, CA3, DG, S, and EC. Arrow indicates abnormal dispersion of granule cells. gl, granular layer; ml, molecular layer; pl, polymorphic layer. (For additional details, see Rempel-Clower N, Zola SM, Squire LR, Amaral DG. Three cases of enduring memory impairment following bilateral damage limited to the hippocampal formation. *J Neurosci.*

1996;16:5233.) (From Squire LR, Zola SM. Memory, memory impairment, and the medial temporal lobe. In: *Cold Spring Harbor Symposia on Quantitative Biology*. Vol LXI. Plainview, NY: Cold Spring Harbor Laboratory Press; 1996:185.)

Table 3.4–1.
Memory and Cognitive Deficits Associated with Frontal Damage

Test	Amnesia	Korsakoff Syndrome	Frontal Lobe Damage
Delayed recall	+	+	–
Dementia rating scale: memory index	+	+	–
Dementia rating scale: initiation and perseveration index	–	+	+
Wisconsin card sorting test	–	+	+
Temporal order memory	+	++	++
Metamemory	–	+	+
Release from proactive interference	–	+	–

+, deficit; –, no deficit; ++, disproportionate impairment relative to item memory.

From Squire LR, Zola-Morgan S, Cave CB, et al. Memory, organization of brain systems and cognition. In: *Cold Spring Harbor Symposia on Quantitative Biology.* Vol LV. Plainview, NY: Cold Spring Harbor Laboratory Press; 1990.

In summary, the ability to remember factual and autobiographical events depends on the integrity of both the cortical regions responsible for representing the information in question and several brain regions that are responsible for memory formation. Thus, medial temporal and diencephalic brain areas work in concert with widespread areas of neocortex to form and to store declarative memories (Fig. 3.4–5).

Retrograde Amnesia

Memory loss in amnesia typically affects recent memories more than remote memories (Fig. 3.4–6). Temporally graded amnesia has been demonstrated retrospectively in studies of amnesic patients and prospectively in studies of monkeys, rats, mice, and rabbits. These findings have important implications for understanding the nature of the memory storage process. Memories are dynamic not static. As time passes after learning, some memories are forgotten, whereas others become stronger due to a process of *consolidation* that depends on cortical, medial temporal, and diencephalic structures.

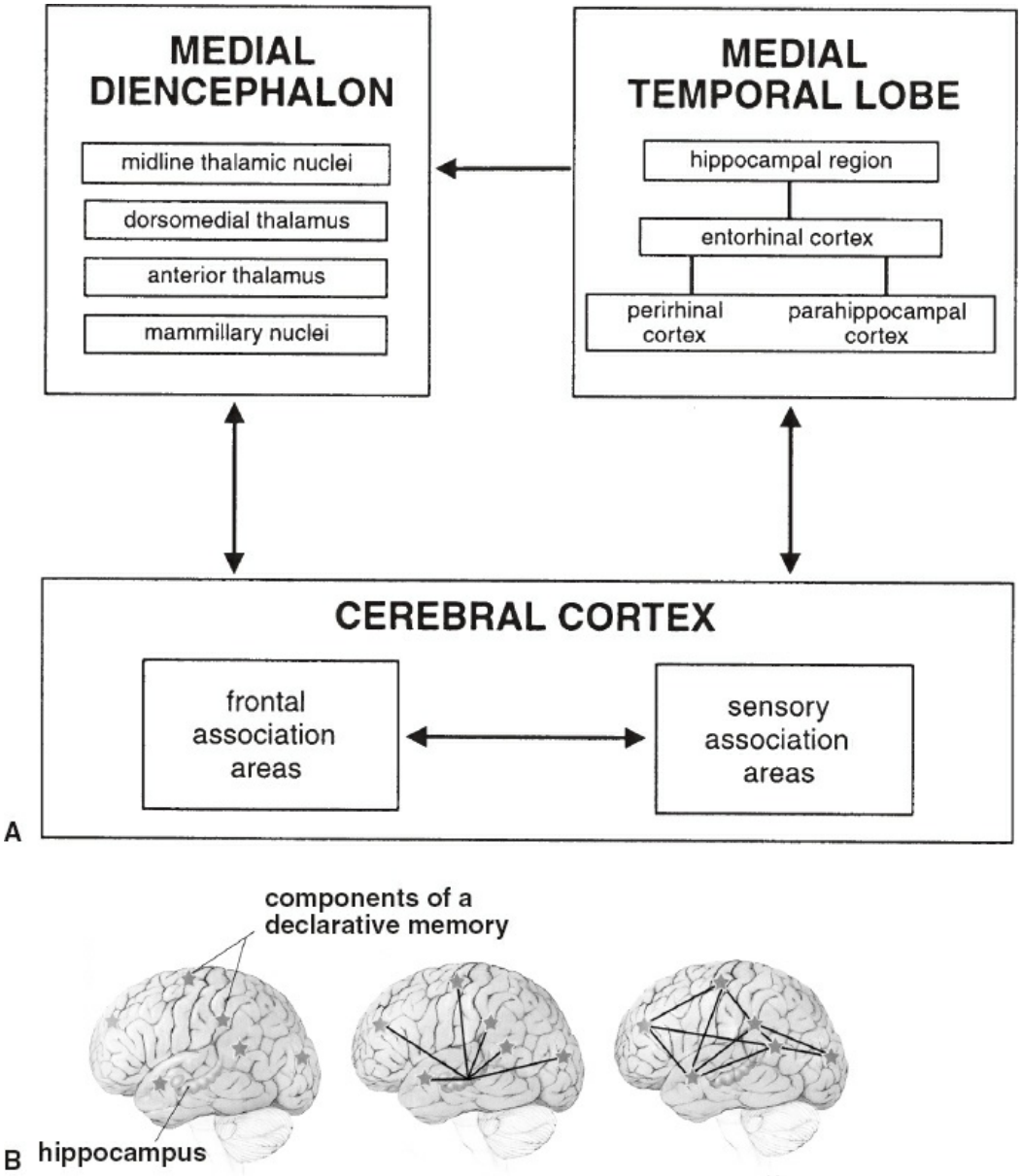

FIGURE 3.4–5. Brain regions thought to be critical for the formation and storage of declarative memory. The upper panel (**A**) lists medial diencephalon and medial temporal regions that are critical for declarative memory storage. The entorhinal cortex is the major source of projections for neocortex to hippocampus, and nearly two-thirds of the cortical input to the entorhinal cortex originates in the perirhinal and parahippocampal cortex. The entorhinal cortex also receives direct connections from the cingulate, insula, orbitofrontal, and superior temporal cortices. The lower panel (**B**) shows a schematic view of the cortical components of a declarative memory (*stars*) that are initially separate (**left**) but become linked with the hippocampus and related structures when an event occurs (**middle**) and ultimately can become intrinsically linked together (**right**). (Adapted from Paller KA. Neurocognitive foundations of human memory. In: Medin DL, ed. *The Psychology of Learning and Motivation,* Vol. 40. San Diego: Academic Press; 2008:121, and Gluck MA, Mercado E, Myers CE. *Learning and Memory: From Brain to Behavior,* Worth Publishers; 2008, fig. 3.16, p. 109.)

The study of retrograde amnesia has been important for understanding how memory changes over time. The dynamic nature of memory storage can be conceptualized as follows. An event is experienced and encoded by

virtue of a collection of cortical regions that are involved in representing a combination of different event features. At the same time, the hippocampus and adjacent cortex receive pertinent high-level information from all sensory modalities. Later, when the original event is recalled, the same set of cortical regions is activated. If a subset of the cortical regions is activated, the hippocampus and related structures can facilitate recall by facilitating the activation of the remaining cortical regions (i.e., pattern completion). When the original event is retrieved and newly associated with other information, hippocampal-cortical networks can be modified. In this way, a gradual consolidation process occurs that changes the nature of memory storage (Fig. 3.4–5). The neocortical components representing some events can become so effectively linked together that ultimately a memory can be retrieved without any contribution from the medial temporal lobe. As a result, amnesic patients can exhibit normal retrieval of remote facts and events, as well as autobiographical memories. Distributed neocortical regions are the permanent repositories of these enduring memories.

In contrast to what is observed after damage restricted to the hippocampus, extensive retrograde impairments for facts and events from the distant past can also occur. Damage to the frontal lobes, for example, can lead to difficulty in organizing memory retrieval. Accurate retrieval often begins with an activation of lifetime periods and proceeds to an identification of general classes of events and then more specific events, but this process becomes difficult following frontal damage. Damage to other cortical regions can also impair memory storage. Networks in anterolateral temporal cortex, for example, are critical for retrieving stored information, because these areas are important for long-term storage itself. Patients with *focal retrograde amnesia* exhibit substantial retrograde memory impairments together with only moderately impaired new learning ability. Some capacity for new learning remains, presumably because medial temporal lobe structures are able to communicate with other areas of cortex that remain undamaged.

MULTIPLE TYPES OF MEMORY

Memory is not a single faculty of the mind but consists of various subtypes. Amnesia affects only one kind of memory, *declarative memory*. Declarative memory is what is ordinarily meant by the term *memory* in everyday language. Declarative memory supports the conscious recollection of facts and events. The classic impairment in amnesia thus concerns memory for routes, lists, faces, melodies, objects, and other verbal and nonverbal material, regardless of the sensory modality in which the material is presented.

Patients with amnesia can display a broad impairment in these components of declarative memory while a number of other memory

abilities are preserved. The heterogeneous set of preserved abilities is collectively termed *nondeclarative memory*. Nondeclarative memory includes skill learning, habit learning, simple forms of conditioning, and a phenomenon called priming. For these kinds of learning and memory, patients with amnesia can perform normally.

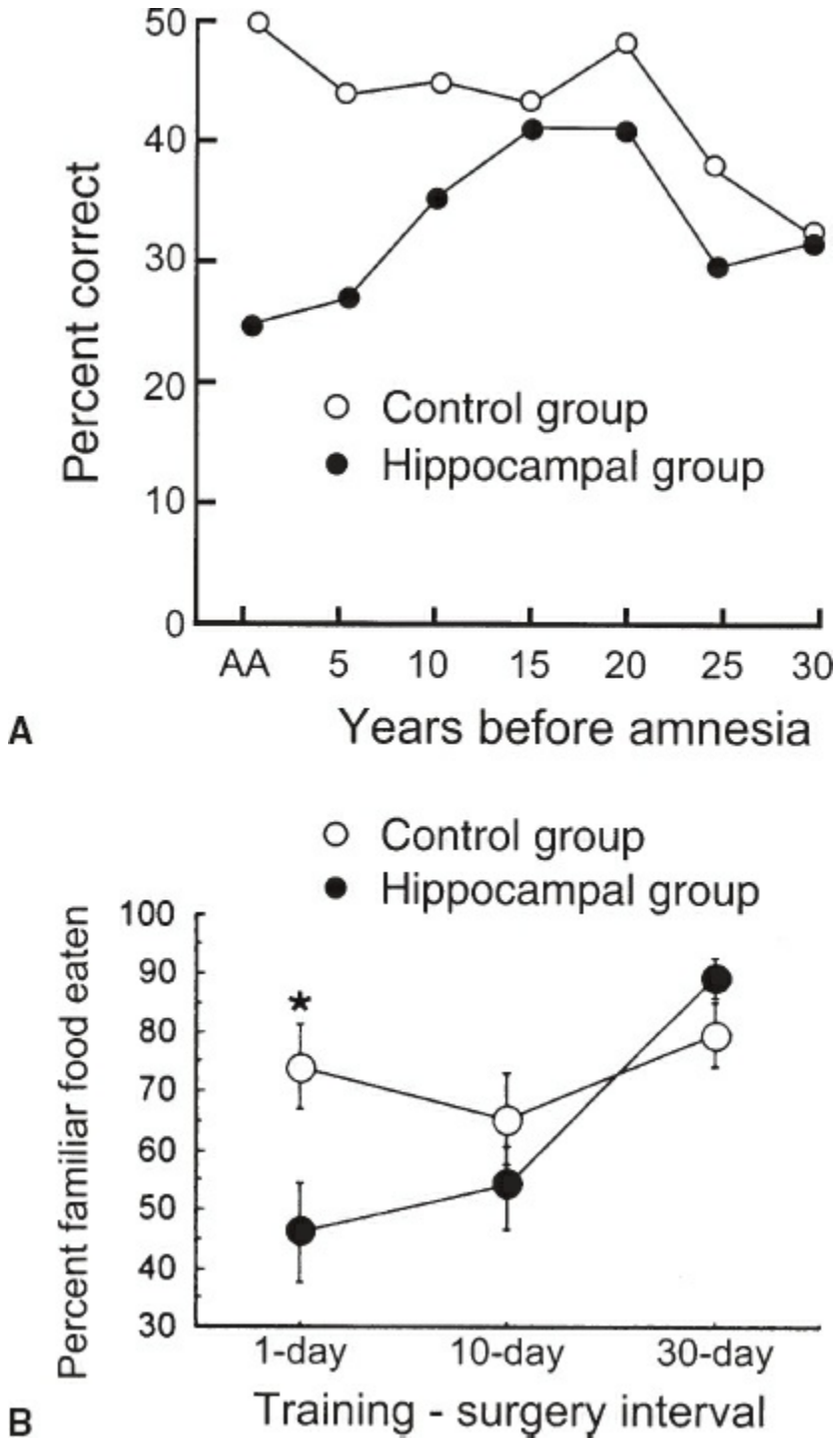

FIGURE 3.4–6. A: Temporally limited retrograde amnesia for free recall of 251 news events. Scores were aligned relative to the onset of amnesia in patients (*N* = 6) and to a corresponding time point in age- and education-matched healthy individuals (*N* = 12). The time period after the onset of amnesia is labeled *AA* (anterograde amnesia) to designate that this time point assessed memory for events that occurred after the onset of amnesia.

Standard errors ranged from 2 to 10 percent. Brain damage in the patient group was limited primarily to the hippocampal region. **B:** Temporally limited retrograde amnesia in rats with lesions of hippocampus and subiculum. Rats learned to prefer an odorous food as the result of an encounter with another rat with that odor on its breath. Percent preference for the familiar food was observed for three training-surgery intervals. At 1 day after learning, the control group performed significantly better than the rats with lesions ($P < 0.05$). At 30 days, the two groups performed similarly, and both groups performed well above chance. Error bars show standard errors of the mean. (Part A adapted from Manns JR, Hopkins RO, Squire LR. Semantic memory and the human hippocampus. *Neuron.* 2003;38:127, Part B adapted from Clark RE, Broadbent NJ, Zola SM, Squire LR. Anterograde amnesia and temporally graded retrograde amnesia for a nonspatial memory task after lesions of hippocampus and subiculum. *J Neurosci.* 2002;22:4663.)

In controlled laboratory settings, the acquisition of a variety of perceptual, perceptual-motor, and cognitive skills can be tested in isolation, and patients with amnesia acquire these skills at rates equivalent to the rates at which in healthy individuals acquire the skills. For example, patients with amnesia can learn to read mirror-reversed text normally, they exhibit the normal facilitation in reading speed with successive readings of normal prose, and they improve as rapidly as healthy individuals at speeded reading of repeating nonwords. In addition, patients with amnesia can, after seeing strings of letters generated by a finite-state rule system, classify novel strings of letters as rule-based or not rule-based. Classification performance is normal despite the fact that patients with amnesia are impaired at remembering the events of training or the specific items they have studied.

Priming

Priming refers to a facilitation of the ability to detect or to identify a particular stimulus based on a specific recent experience. Many tests have been used to measure priming in amnesia and show that it is intact. For example, words might be presented in a study phase and then again, after a delay, in a test phase when a priming measure such as reading speed is obtained. Patients are instructed to read words as quickly as possible in such a test, and they are not informed that memory is being assessed.

In one priming test, patients named pictures of previously presented objects reliably faster than they named pictures of new objects, even after a delay of a week. This facilitation occurred at normal levels, despite the fact that the patients were markedly impaired at recognizing which pictures had been presented previously. Particularly striking examples of preserved priming come from studies of patient EP (Fig. 3.4–7), who exhibited intact priming for words but performed at chance levels when asked to recognize which words had been presented for study. This form of memory, termed *perceptual priming,* is thus a distinct class of memory that is independent of the medial temporal lobe regions typically damaged in amnesia.

Another form of priming reflects improved access to meaning rather

than percepts. For example, subjects study a list of words, including *tent* and *belt,* and then are asked to free associate to other words. Thus, they are given words such as *canvas* and *strap* and asked to produce the first word that comes to mind. The result is that subjects are more likely to produce *tent* in response to *canvas* and to produce *belt* in response to *strap* than if the words *tent* and *belt* had not been presented recently. This effect, called *conceptual priming,* is also preserved in patients with amnesia, even though they fail to recognize the same words on a conventional memory test (Fig. 3.4–8).

Yet, not all types of priming are preserved in amnesia. Some priming tests have been designed to examine the formation of new associations. When priming tests are based not on pre-existing knowledge but on the acquisition of new associative knowledge, priming tends to be impaired. In other words, priming in certain complex situations can require the same type of linkage among multiple cortical regions that is critical for declarative memory.

Memory Systems

One scheme for conceptualizing multiple types of memory appears in Figure 3.4–9. Declarative memory depends on medial temporal and midline diencephalic structures along with large portions of the neocortex. This system provides for the rapid learning of facts (*semantic memory*) and events (*episodic memory*). Nondeclarative memory depends on several different brain systems. Habits depend on the neocortex and the neostriatum, the cerebellum is important for conditioning of skeletal musculature, the amygdala for emotional learning, and the neocortex for priming.

Declarative and nondeclarative memories differ in important ways. Declarative memory is phylogenetically more recent than nondeclarative memory. In addition, declarative memories are available to conscious recollection. The flexibility of declarative memory permits the retrieved information to be available to multiple response systems. Nondeclarative memory is inaccessible to awareness and is expressed only by engaging specific processing systems. Nondeclarative memories are stored as changes within these processing systems, changes that are encapsulated such that the stored information has limited accessibility to other processing systems.

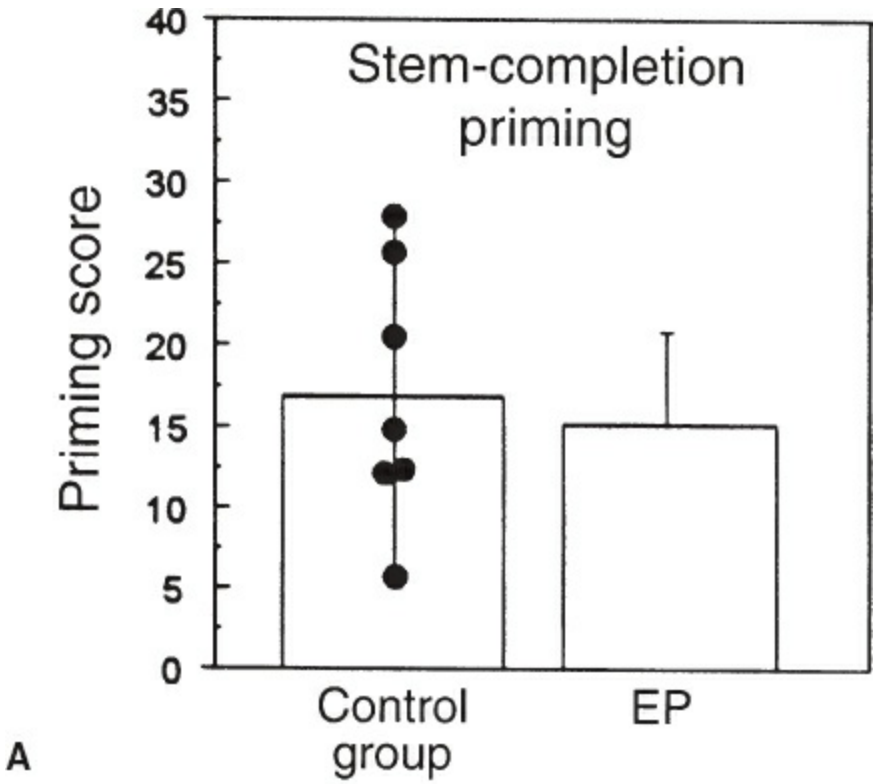

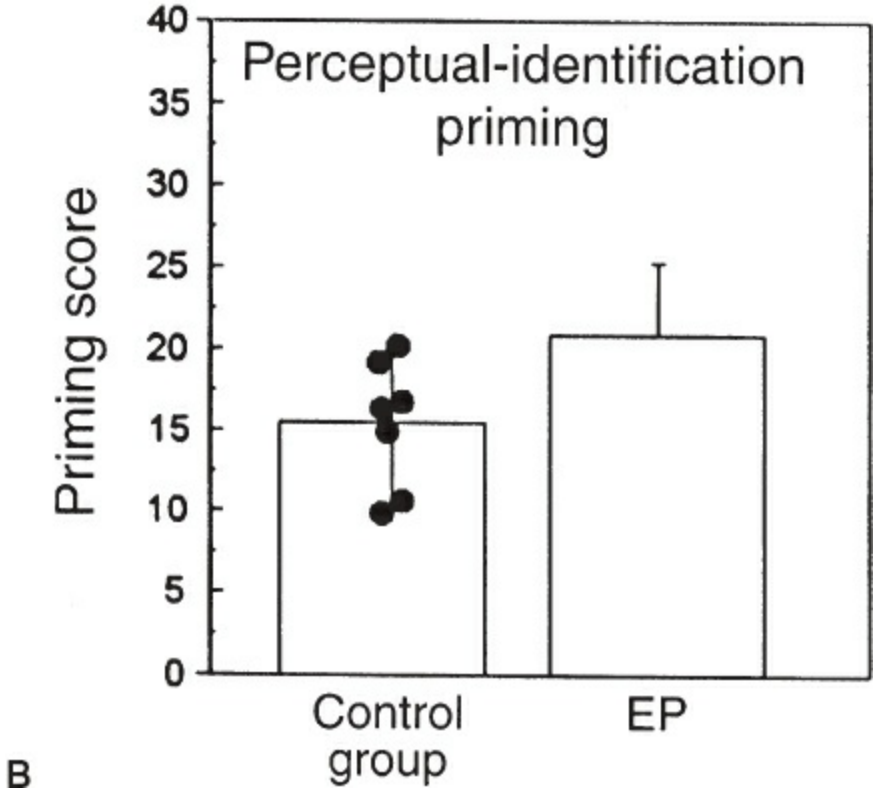

FIGURE 3.4–7. Preserved priming in patient EP relative to seven control subjects. **A:** Stem-completion priming on six separate tests. Priming reflected a tendency for subjects to complete three-letter stems with previously encountered words when they were instructed to produce the first word to come to mind (e.g., MOT___ completed to form MOTEL). Priming scores were calculated as percent correct for studied words minus percent correct for baseline words (guessing). **B:** Perceptual-identification priming on 12 separate tests. Subjects attempted to read 48 words that were visually degraded. Priming scores were

calculated as percent correct identification of previously studied words minus percent correct identification of nonstudied words. Brackets indicate standard error of the mean. (Data from Hamann SB, Squire LR. Intact perceptual memory in the absence of conscious memory. *Behav Neurosci.* 1997;111:850.) (From Stefanacci L, Buffalo EA, Schmolck H, Squire LR. Profound amnesia after damage to the medial temporal lobe: a neuroanatomical and neuropsychological profile of patient E.P. *J Neurosci.* 2000;20:7024.)

Semantic memory, which concerns general knowledge of the world, has often been categorized as a separate form of memory. Facts that are committed to memory typically become independent of the original episodes in which the facts were learned. Patients with amnesia can sometimes acquire information that would ordinarily be learned as facts, but the patients learn it by relying on a different brain system than the system that supports declarative memory.

Consider a test requiring the concurrent learning of eight object pairs. Healthy individuals can rapidly learn which is the correct object in each pair, whereas patients with severe amnesia like EP learn only gradually over many weeks and at the start of each session they cannot describe the task, the instructions, or the objects. In patients who do not have severe amnesia, factual information is typically acquired as consciously accessible declarative knowledge. In these cases, the brain structures that remain intact within the medial temporal lobe presumably support learning. In contrast, when factual information is acquired as nondeclarative knowledge, as in the case of EP's learning of pairs of objects, learning likely occurs directly as a habit, perhaps supported by the neostriatum. Humans thus appear to have a robust capacity for habit learning that operates outside of awareness and independently of the medial temporal lobe structures that are damaged in amnesia.

Frontal Contributions to Memory

Although amnesia does not occur after limited frontal damage, the frontal lobes are fundamentally important for declarative memory. Patients with frontal lesions have poor memory for the context in which information was acquired, they have difficulty in unaided recall, and they may even have some mild difficulty on tests of item recognition. More generally, frontal patients have difficulty implementing memory retrieval strategies and evaluating and monitoring their memory performance.

Patient BG suffered an infarction restricted to the right frontal lobe, resulting in substantial false remembering. He had an abnormal tendency to claim that some stimuli were familiar, even though they had not been presented for study. His false responses probably arose because he relied on a general feeling of familiarity for the kind of stimuli that had been presented, rather than on specific memories for the stimuli.

NEUROIMAGING AND MEMORY

The understanding of memory derived from studies of amnesia has been

extended through studies using various methods for monitoring brain activity in healthy individuals. For example, studies of brain activity measured using PET and functional MRI have shown that posterior prefrontal regions are involved in strategic processing during retrieval, as well as in working memory. Anterior frontal regions near the frontal poles have been linked with functions such as evaluating the products of retrieval. Frontal connections with posterior neocortical regions support the organization of retrieval and the manipulation of information in working memory. Consistent with the evidence from patients with frontal lesions, frontal–posterior networks can be viewed as instrumental in the retrieval of declarative memories and in the online processing of new information.

Neuroimaging has also identified contributions to memory made by parietal cortex. Multiple parietal regions (including inferior and superior parietal lobules, precuneus, posterior cingulate, and retrosplenial cortex) are activated in conjunction with remembering recent experiences. Although many functions have been hypothesized to explain this parietal activity, a consensus position has not yet been reached, and it may be the case that several different functions are relevant.

Neuroimaging studies have also illuminated priming phenomena and how they differ from declarative memory. Perceptual priming appears to reflect changes in early stages of the cortical pathways that are engaged during perceptual processing. For example, in the case of stem-completion priming, where subjects study a list of words (e.g., MOTEL) and then are tested with a list of stems (e.g., MOT____) and with instructions to complete each stem with the first word to come to mind, neuroimaging and divided visual-field studies have implicated visual processing systems in extrastriate cortex, especially in the right hemisphere. In contrast, conscious recollection of remembered words engages brain areas at later stages of processing. Neural mechanisms supporting priming and declarative memory retrieval have also been distinguished in brain electrical activity recorded from the scalp (Fig. 3.4–10). In sum, priming differs from declarative memory in that it is signaled by brain activity that occurs more efficiently during memory retrieval and that originates in different brain regions. Recent findings suggest that brain activity related to both priming and declarative memory often occurs together, suggesting that multiple co-occurring memory processes are normally involved in memory retrieval. The degree to which processes involved in priming versus declarative memory work together to support healthy memory function is unknown and remains an intriguing focus of contemporary research.

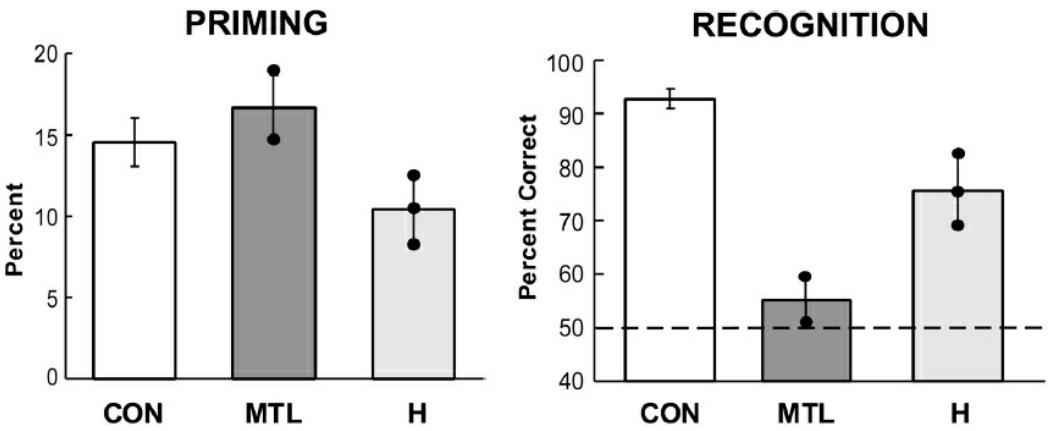

FIGURE 3.4–8. Preserved conceptual priming in amnesia. In the free-association test, subjects studied a set of words (e.g., *lemon*) and 5 minutes later viewed cue words that included associates of studied words (e.g., *orange*). Subjects were asked to produce the first word that came to mind in response to each cue word. Results are shown separately for the control group (CON, n = 12), amnesic patients with large medial temporal lobe lesions (MTL, n = 2), and amnesic patients with lesions thought to be limited to the hippocampal region (H, n = 3). The left panel shows conceptual priming scores computed as the percentage of studied words produced in the free-association test minus a baseline measure of the likelihood of producing those words by chance. All the groups performed similarly in the conceptual priming test. The right panel shows results from a yes–no recognition memory test using comparable words. Both patient groups were impaired relative to the control group. The dashed line indicates chance performance. The data points for the MTL and H groups show the scores of individual patients averaged across four separate test. Brackets show standard errors of the mean for the control group. (From Levy DA, Stark CEL, Squire LR. Intact conceptual priming in the absence of declarative memory. *Psych Sci.* 2004;15:680.)

Hippocampal activity associated with the formation and retrieval of declarative memories has also been investigated with neuroimaging. In keeping with the neuropsychological evidence, the hippocampus appears to be involved in the recollection of recent events (Fig. 3.4–11). Retrieval-related hippocampal activity has been observed in memory tests with many different types of stimuli. The hippocampus is also active during the initial storage of information. Whereas left inferior prefrontal cortex is engaged as a result of attempts to encode a word, hippocampal activity at encoding is more closely associated with whether encoding leads to a stable memory that can later be retrieved (Fig. 3.4–12). These findings confirm and extend the idea that medial temporal and frontal regions are important for memory storage and that they make different contributions.

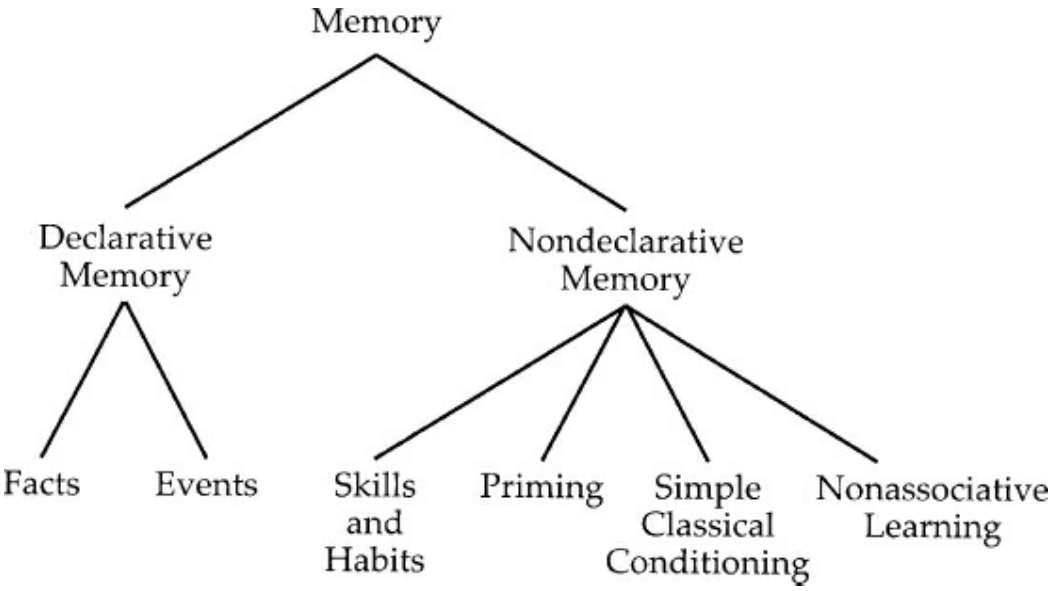

FIGURE 3.4–9. A taxonomy for subtypes of memory. *Declarative memory* refers to conscious recollection of facts and events. *Nondeclarative memory* refers to a collection of abilities wherein performance changes as the result of experience but without affording access to conscious memory of the original experience. Nonassociative learning includes habituation and sensitization. (Adapted from Squire LR, Zola-Morgan S. The medial temporal lobe memory system. *Science.* 1991;253:1380.)

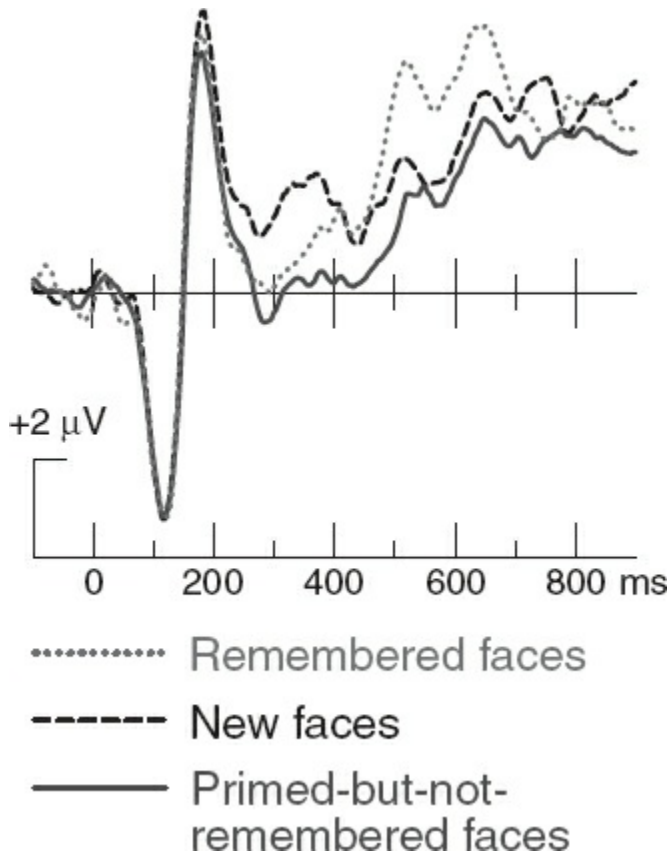

FIGURE 3.4–10. Brain potentials associated with perceptual priming versus declarative memory retrieval. Paller et al. (2003) studied sixteen volunteers who took a memory test involving three sorts of faces: new faces, faces they had seen recently and remembered well, and faces they had seen but did not remember because the faces had been presented too briefly to process effectively. In a companion experiment with a priming test, speeded

responses were found, indicative of priming. Frontal recordings of brain waves elicited by the primed faces included negative potentials from 200 to 400 ms after face presentation that differed from the brain waves elicited by new faces. These differences were particularly robust for trials with the fastest responses (data shown were from trials with responses faster than the median reaction time). Remembered faces uniquely elicited positive brain waves that began about 400 ms after face presentation. Brain potential correlates of face recollection occurred later than those for perceptual priming and were larger over posterior brain regions. (Adapted from Paller KA, Hutson CA, Miller BB, Boehm SG. Neural manifestations of memory with and without awareness. *Neuron.* 2003;38:507.)

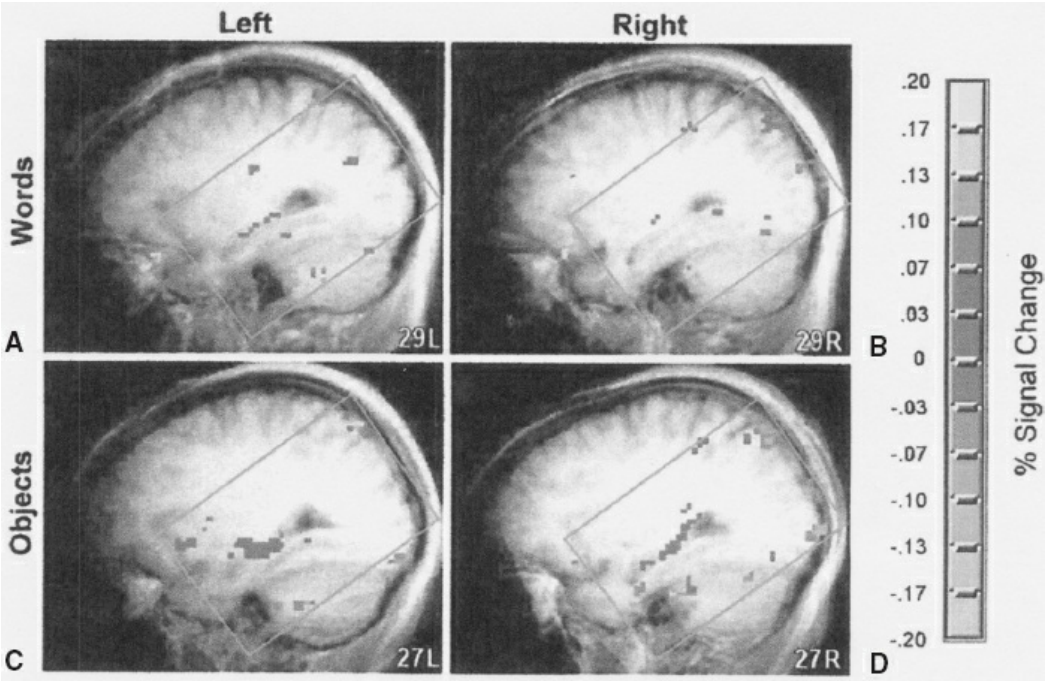

FIGURE 3.4–11. Activity in the left and right hippocampal regions measured with functional magnetic resonance imaging (fMRI) during declarative memory retrieval. Data were collected from 11 participants who saw words at study and test and from 11 different participants who saw pictures of namable objects at study and test. Recognition memory accuracy was 80.2 percent correct for words and 89.9 percent correct for objects. Areas of significant fMRI signal change (targets vs. foils) are shown in sagittal sections as color overlays on averaged structural images. The green box indicates the area in which reliable data were available for all subjects. With words, retrieval-related activity was observed in the hippocampus on the left side **(A)** but not on the right side **(B).** With namable objects, retrieval-related activity was observed in the hippocampus on both the left side **(C)** and the right side **(D).** (From Stark CE, Squire LR. Functional magnetic resonance imaging (fMRI) activity in the hippocampal region during recognition memory. *J Neurosci.* 2000;20:7776.)

MEMORY IMPAIRMENTS IN PSYCHIATRIC ILLNESS

The biological description of memory provided by studies of individuals with amnesia and by functional neuroimaging has deepened understanding of memory impairments in individuals with a variety of psychiatric illnesses. Abnormal memory is crucial to the etiology of some psychiatric illnesses, such as the pathological associations between innocuous stimuli and fear that occur in phobias as well as the intrusive and contextually inappropriate recollections of traumatic events that occur in posttraumatic stress disorder. These disorders have been associated with structural,

neurochemical, and functional abnormalities of medial temporal and frontal regions. As further evidence implicating declarative memory, such patients often have moderate impairments of declarative memory for emotional as well as neutral material, suggesting generalized disruption. It remains unclear whether declarative memory disruptions identified in these disorders resulted from premorbid individual differences that conferred risk for the disorder, versus whether brain abnormalities accrued over time as part of the disorder. Nonetheless, understanding of biological processes supporting memory, especially those operating in medial temporal structures, has led to promising developments in the treatment of these disorders. For instance, the extinction of learned fear occurs in a therapeutic setting during cognitive behavioral therapy for phobias, and is associated with NMDA-receptor signaling crucial for LTP in the basolateral amygdala. D-Cycloserine is a partial agonist of these receptors and has been shown in a number of preclinical and clinical studies to enhance fear extinction and the effectiveness of cognitive behavioral therapy for phobia. Neurochemical modulation of medial temporal structures thus holds promise for the treatment of anxiety disorders and posttraumatic stress disorder.

Memory impairments also occur in psychiatric illnesses that don't have obvious linkages to memory, such as schizophrenia and major depression. Memory impairments likely occur because these conditions are associated with disruptions of the structure and/or function of medial temporal, midline diencephalic, and frontal neocortical structures that are core components of memory systems. For instance, although schizophrenia is chiefly characterized by delusions, hallucinations, and abnormal emotion, speech, and motivation, patients can experience reduced declarative memory abilities that linger even when core symptoms are pharmacologically treated with success. Patients with schizophrenia perform poorly on memory tests that are highly sensitive to hippocampal structural abnormality, such as transitive inference and across-domain association learning, and these impairments have been associated with abnormal hippocampal shape and reduced volume and functional activity. These hippocampal structural and functional abnormalities have been shown to worsen with disease progression and are not observed as reliably in other psychotic illnesses, such as psychotic bipolar disorder. Frontal cortex structural and functional abnormality has also been reliably identified in schizophrenia, generally in association to poor working or short-term memory abilities, suggesting impairments in structuring and organization of information that could further limit declarative memory ability.

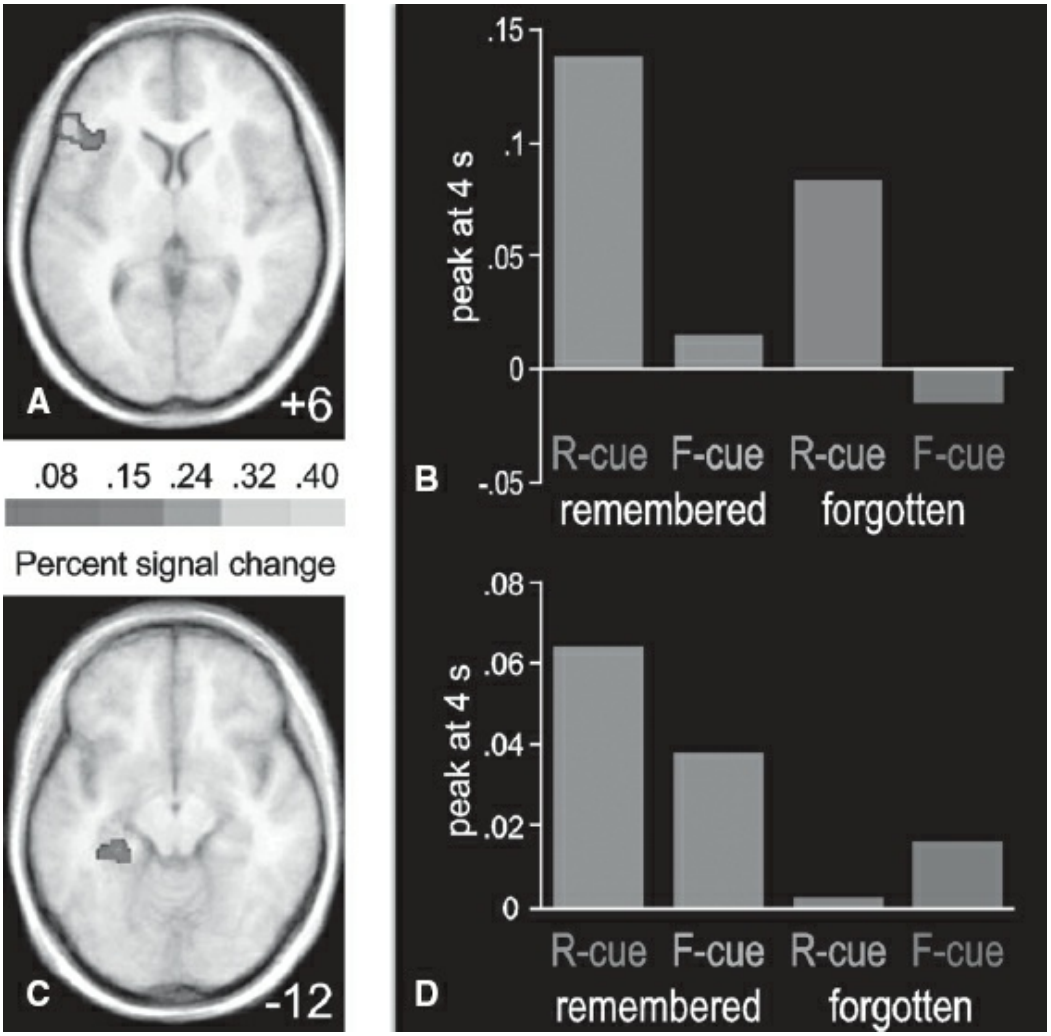

FIGURE 3.4–12. Functional activations of prefrontal and medial temporal regions that were predictive of later memory performance. Single words were presented visually, each followed by an instruction to remember (R cue) or to forget (F cue). Trials were sorted based on the remember or forget instruction and on subsequent recognition performance. Activity in left inferior prefrontal cortex and left hippocampus was predictive of subsequent recognition but for different reasons. Left inferior prefrontal activation **(A)** was associated with the encoding attempt, in that responses were largest for trials with a cue to remember, whether or not the word was actually recognized later. The time course of activity in this region **(B)** was computed based on responses that were time locked to word onset (time 0). Left inferior prefrontal activity increased for words that were later remembered, but there was a stronger association with encoding attempt, as responses were larger for words followed by an R cue that were later forgotten than for words followed by an F cue that were later remembered. In contrast, left parahippocampal and posterior hippocampal activation **(C)** was associated with encoding success. As shown by the time course of activity in this region **(D),** responses were largest for words that were subsequently remembered, whether the cue was to remember or to forget. (Redrawn from Reber PJ, Siwiec RM, Gitelman DR, Parrish TB, Mesulam MM, Paller KA. Neural correlates of successful encoding identified using functional magnetic resonance imaging. *J Neurosci.* 2002;22:9541.)

Memory impairments are similarly present in major depression. Although memory abnormalities in depression are not core symptoms of the illness and are sometimes understood as mere ramifications of negative

affect (e.g., bias towards recall of negative material and rumination on recollection of negative autobiographical episodes), depression patients also exhibit reduced recollection of specific details from previous episodes. Such patients tend to favor overly general or schematic information, even for memories with little emotional content. Again, these memory abnormalities have been associated with structural and functional alterations of medial temporal structures. Thus, schizophrenia and depression are clear examples of conditions in which memory impairments occur in the context of disrupted memory systems and present challenges for patients beyond the symptoms typically associated with the illness. Memory assessment therefore could be used to assist clinicians and caregivers in understanding and treating psychiatric patients. Furthermore, lingering cognitive challenges such as memory impairment remain salient barriers to healthy long-term outcomes in patients even when primary symptoms are successfully treated. Greater focus is therefore needed on development of treatments for memory impairments in a variety of psychiatric disorders.

SLEEP AND MEMORY

The speculation that memories are processed during sleep has a long history. Freud noted that dreams can reveal fragments of recent experiences in the form of day residues. Although many questions about how and why memories may be processed during sleep remain unresolved, recent experiments have provided new empirical support to bolster the idea that memory processing during sleep serves an adaptive function. It is now clear that memory performance can be facilitated when sleep occurs after initial learning, and that sleep-related facilitation can be observed for many different types of memory.

Memory storage appears to be specifically aided by processing during deep sleep within a few hours after learning, especially in the stage of sleep known as *slow-wave sleep,* and probably also in other sleep stages. Some results indicate that slow-wave sleep facilitates storage of declarative memories but not nondeclarative memories. Direct evidence for this proposal has been obtained using stimulation with olfactory stimuli (Fig. 3.4–13), stimulation with direct-current electrical input at the approximate frequency of EEG slow oscillations, and other methods. A fruitful approach for studying memory processing during sleep is to use sensory stimulation to activate specific memories, which then may be subject to modification. Neuronal recordings in animals have revealed a phenomenon of hippocampal replay, whereby activity patterns expressed during the day are later observed during sleep. Research is now focused on delineating the neurophysiological events that transpire when memories are reactivated and altered during sleep. A novel conjecture, supported by some recent empirical findings, is that adapting these procedures for memory

modification during sleep could be used to treat unwanted learned associations, such as those involved in nicotine craving (Fig. 3.4–14). In summary, declarative memories acquired during waking can be processed again during sleep, and this processing can influence the likelihood of subsequent memory retrieval when the individual is awake.

ASSESSMENT OF MEMORY FUNCTIONS

A variety of quantitative methods are available to assess memory functions in neurological and psychiatric patients. Quantitative methods are useful for evaluating and following patients longitudinally as well as for carrying out a one-time examination to determine the status of memory function. It is desirable to obtain information about the severity of memory dysfunction, as well as to determine whether memory is selectively affected or whether memory problems are occurring, as they often do, against a background of additional intellectual deficits. The following section presents a rationale for evaluating memory and identifies areas of memory function that should be evaluated in any comprehensive assessment. Although some widely available tests, such as the Wechsler Memory Scale, provide useful measures of memory, most single tests assess memory rather narrowly. Even general-purpose neuropsychological batteries provide for only limited testing of memory functions. A complete assessment of memory usually involves a number of specialized tests that sample intellectual functions, new learning capacity, remote memory, and memory self-report.

The assessment of general intellectual functions is central to any neuropsychological examination. In the case of memory testing, information about intellectual functions provides information about a patient's general test-taking ability and a way to assess the selectivity of memory impairment. Useful tests include the Wechsler Adult Intelligence Scale; a test of object naming, such as the Boston Naming Test; a rating scale to assess the possibility of global dementia; a test of word fluency; and specialized tests of frontal lobe function.

New Learning Capacity

Memory tests are sensitive to impaired new learning ability when they adhere to either of two important principles. First, tests are sensitive to memory impairment when more information is presented than can be held in immediate memory. For example, one might ask patients to memorize a list of ten faces, words, sentences, or digits, as ten items is more than can be held in mind. The paired-associate learning task is an especially sensitive test of this kind. In the paired-associate task, the examiner asks the patient to learn a list of unrelated pairs of words (e.g., queen-garden and office-river) and then to respond to the first word in each pair by

recalling the second word.

Second, tests are sensitive to memory impairment when a delay, filled with distraction, is interposed between the learning phase and the test phase. In that case, examiners typically ask patients to learn a small amount of information and then distract them for several minutes by conversation to prevent rehearsal. Recollection is then assessed for the previously presented material. Memory can be tested by unaided recall of previously studied material (free recall), by presenting a cue for the material to be remembered (cued recall), or by testing recognition memory. In multiple-choice tests of recognition memory, the patient tries to select previously studied items from a group of studied and unstudied items. In yes–no recognition tests, patients see studied and unstudied items one at a time and are asked to say "yes" if the item was presented previously and "no" if it was not. These methods for assessing recently learned material vary in terms of their sensitivity for detecting memory impairment, with free recall being most sensitive, cued recall intermediate, and recognition least sensitive.

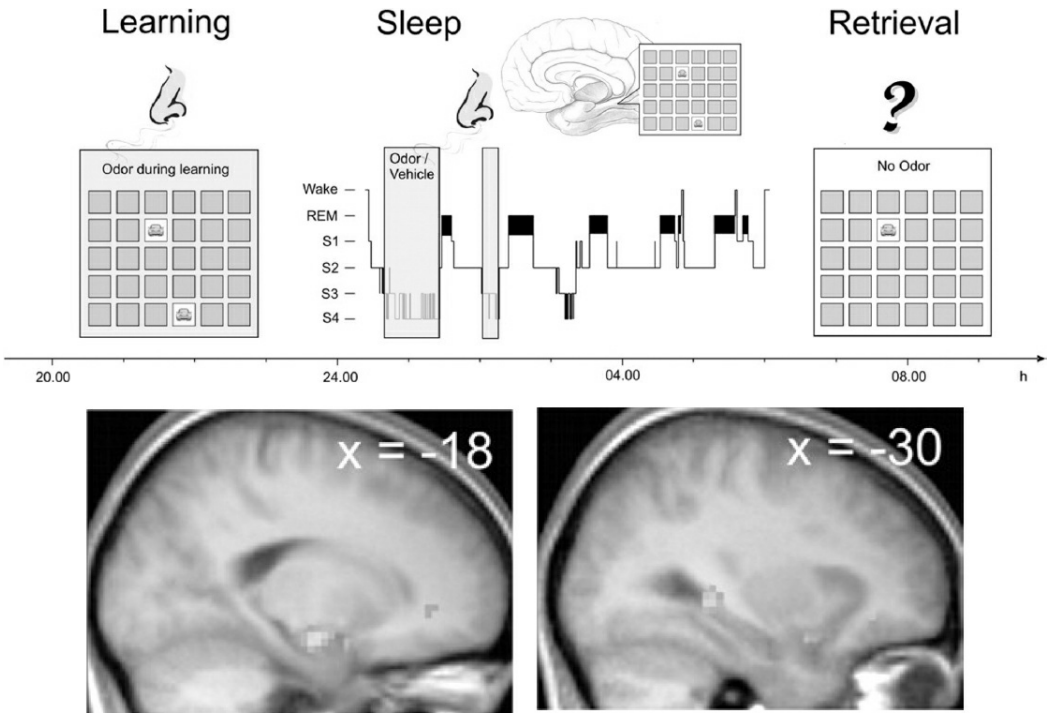

FIGURE 3.4–13. Evidence for memory processing during sleep. Subjects first learned object–location associations when a rose odor was present. Following learning, subjects slept wearing a device for delivering odors to the nose, and the rose odor was administered during the first two slow-wave sleep periods of the night (in 30-s periods to prevent habituation). Memory facilitation was observed when object–location associations were tested the following morning in the absence of odor stimulation. Facilitated memory was not found when stimulation occurred during slow-wave sleep but not during learning, when stimulation occurred during learning and then during REM sleep, or when subjects were kept awake. Moreover, odor stimulation during slow-wave sleep was found to produce anterior and posterior hippocampal activation (lower panels). (From Rasch B, Büchel C, Gais S, Born J. Odor cues during slow-wave sleep prompt declarative memory consolidation.

Science. 2007;315;1426.)

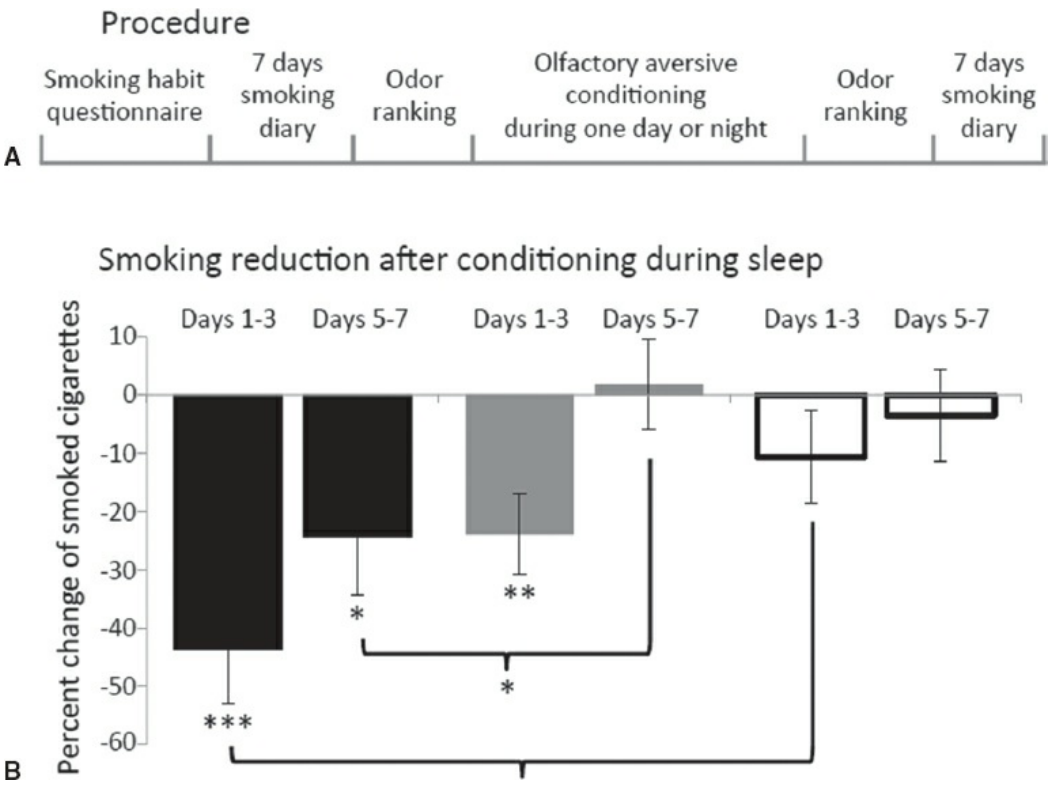

FIGURE 3.4–14. Memory processing during sleep in relation to smoking habits. **A:** Procedure used by Arzi and colleagues (2014) to study aversive conditioning during sleep in individuals who wanted to quit smoking. **B:** Results showed that measures of cigarette smoking during the subsequent week were reduced when conditioning took place during sleep. Conditioning during stage-2 sleep produced the strongest reduction in smoking, compared to conditioning during REM sleep and conditioning during wake. (Adapted from Arzi A, Holtzman Y, Samnon P, Eshel N, Harel E, Sobel N. Olfactory aversive conditions during sleep reduces cigarette-smoking behavior. *J Neurosci.* 2014;34:15382–15393.)

The specialization of function of the two cerebral hemispheres in humans means that left and right unilateral damage is associated with different kinds of memory problems. Accordingly, different kinds of memory tests must be used when unilateral damage is a possibility. In general, damage to medial temporal or diencephalic structures in the left cerebral hemisphere causes difficulty remembering verbal material, such as word lists and stories. Damage to medial temporal or diencephalic structures in the right cerebral hemisphere impairs memory for faces, spatial layouts, and other nonverbal material that is typically encoded without verbal labels. Left medial temporal damage can lead to impaired memory for spoken and written text. Right medial temporal damage can lead to impaired learning of spatial arrays, whether the layouts are examined by vision or by touch. A useful way to test for nonverbal memory is to ask a patient to copy a complex geometric figure and then, after a delay of several minutes, without forewarning, ask the patient to reproduce it.

Remote Memory

Evaluations of retrograde memory loss should attempt to determine the severity of any memory loss and the time period that it covers. Most quantitative tests of remote memory are composed of material in the public domain that can be corroborated. For example, tests have been used that concern news events, photographs of famous persons, or former one-season television programs. An advantage of these methods is that one can sample large numbers of events and can often target particular time periods. A disadvantage is that these tests are not so useful for detecting memory loss for information learned during the weeks or months immediately before the onset of amnesia. Most remote memory tests sample time periods rather coarsely and cannot detect a retrograde memory impairment that covers only a few months. Also, these tests can make it difficult to take into account what the patient might have learned years ago.

In contrast, autobiographical memory tests can potentially provide fine-grained information about a patient's retrograde memory. In the word-probe task, first used by Francis Galton in 1879, patients are asked to recollect specific episodes from their past in response to single word cues (e.g., bird and ticket) and to date the episodes. The number of episodes recalled tends to be systematically related to the time period from which the episode is taken. Most memories normally come from recent time periods (the past 1 or 2 months), whereas patients with amnesia often exhibit temporally graded retrograde amnesia, drawing few episodic memories from the recent past, but producing as many remote autobiographical memories as normal subjects (Fig. 3.4–6). Endeavors to assess remote memories are hampered by the fact that decades-old autobiographical memories tend to lose some detail, and these memories may also be subject to further modification if memory functions are intact. Patients who have amnesia for a long time may also adopt strategies or habits that promote gist-based memory retrieval that may interfere with retrieving specific details from remote events.

Memory Self-Reports

Patients can often supply descriptions of their memory problems that are extremely useful for understanding the nature of their impairment. Tests of the ability to judge one's own memory abilities are called tests of *metamemory*. Self-rating scales are available that yield quantitative and qualitative information about memory impairment. As a result, it is possible to distinguish memory complaints associated with depression from memory complaints associated with amnesia. Depressed patients tend to rate their memory as poor in a rather undifferentiated way, endorsing equally all the items on a self-rating form. By contrast, patients with

amnesia tend to endorse some items more than others; that is, there is a pattern to their memory complaints. Patients with amnesia do not report difficulty in remembering very remote events or in following what is being said to them, but they do report having difficulty remembering an event a few minutes after it happens. Indeed, self-reports can match rather closely the description of memory dysfunction that emerges from objective tests. Specifically, new-learning capacity is affected, immediate memory is intact, and very remote memory is intact. Some patients with amnesia, however, tend to have poor awareness of their memory impairment, sometimes overestimating and sometimes underestimating their ability. In patients with Korsakoff syndrome, for example, their poor metamemory stems from frontal lobe dysfunction. In any case, querying patients in some detail about their own sense of impairment and administering self-rating scales are valuable and informative adjuncts to more formal memory testing.

Psychogenic Amnesia

Patients sometimes exhibit memory impairment that differs markedly from the typical patterns of memory loss that follow brain damage. For example, some cases of amnesia present with a sudden onset of retrograde amnesia, a loss of personal identity, and minimal anterograde amnesia. These patients may even be unable to recall their own name. Given the psychological forces that prompt the onset of amnesia in these cases, they are commonly termed *psychogenic amnesia,* or sometimes *hysterical amnesia, functional amnesia,* or *dissociative amnesia*.

Differentiating psychogenic amnesia from a memory disorder that results from frank neurological injury or disease is often straightforward. Psychogenic amnesias typically do not affect new learning capacity. Patients enter the hospital able to record a continuing procession of daily events. By contrast, new learning problems tend to be at the core of neurological amnesia. The main positive symptom in psychogenic amnesia is extensive and severe retrograde amnesia. Patients may be unable to recollect pertinent information from childhood or from some part of their past. Formal neuropsychological testing has shown that the pattern of memory deficits varies widely from patient to patient. This variability may reflect a patient's commonsense concepts of memory, even when symptoms are not the result of conscious attempts to simulate amnesia. Some patients may perform poorly only when asked to remember past autobiographical events. Other patients may fail as well at remembering past news events. Some patients perform well when memory tests appear to assess general knowledge, such as remembering the names of celebrities or cities. Learning new material is usually intact, perhaps because such tests appear to concern the present moment, not traveling into the past. Occasionally, patients with psychogenic amnesia exhibit broad memory deficits such that they cannot perform previously familiar skills nor identify common objects

or common vocabulary words.

By contrast, patients with neurological amnesia never forget their names, and remote memory for the events of childhood and adolescence is typically normal, unless there is damage to the lateral temporal or frontal lobes. Patients with psychogenic amnesia sometimes have evidence of head trauma or brain injury, but nevertheless the pattern of deficits cannot be taken as a straightforward result of neurological insult. The clinician's challenge is not to distinguish psychogenic amnesia from neurological amnesia but to distinguish psychogenic amnesia from malingering. Indeed, the diagnosis of psychogenic amnesia can be difficult to substantiate and may be met with skepticism by hospital staff. Some features that argue in favor of a genuine psychogenic disorder include: (1) memory test scores that are not as low as possible, and never worse than chance levels; (2) memory access improved by hypnosis or Amytal interview; and (3) a significant premorbid psychiatric history. In some cases, psychogenic amnesia has been observed to clear after a period of days, but in many cases it has persisted as a potentially permanent feature of the personality.

Patient CB was a 29-year-old man evaluated for memory loss. He reported having severe retrograde amnesia for the past 3 months. Medical evaluation revealed swelling of the left side of his head due to recent head trauma. CT head scan was normal. Three months prior to this evaluation, he had been arrested for intoxication and taken to jail. Upon awakening the next day, he was unable to recall his past life. At the time of the evaluation, he indicated that he knew his name, but it was not clear whether he knew his name only because he had relearned it. He also had some difficulty recalling old vocabulary words. He had essentially no memories of events that had occurred before the onset of his amnesia. He was, however, able to relearn about his past life as he was told about it. He also learned new information well. He sometimes did not recognize vocabulary words and had to look them up. He reported that he had recently re-learned how to drive a car with a manual transmission.

It was notable that he described himself as having a new personality since the onset of memory problems. Previously he had been very outgoing, a heavy drinker of alcohol, and was feared by some people. He had been "obnoxious but likable." Since the onset of memory problems, he was less outgoing, more introspective, and better at getting along with other people. He did not like his personality from before the onset of memory problems, and, in fact, he had adopted his former middle name as a new first name. He hoped that he would be able to regain all of his old memories and experiences without the return of his old personality. During a 20-month follow-up period, his severe retrograde amnesia persisted. (Case history from Kritchevsky M, Chang J, Squire LR. Functional amnesia: clinical description and neuropsychological profile of 10 cases. *Learn Mem.* 2004;11:213.)

IMPLICATIONS

Memory Distortion

Current understanding of the biology of memory has significant implications for several fundamental issues in psychiatry. Given the selective and constructive nature of autobiographical remembering, and the imperfect nature of memory retrieval more generally, it is surprising that memory is so often accurate. How much can memories be trusted? Subjective feelings of confidence are apparently not perfect indicators of

retrieval accuracy. And memory distortion can clearly lead to unfortunate consequences, such as when mistaken eyewitness testimony harms an innocent individual.

The fact is that it is possible to confidently remember events that never happened. For example, it is possible to confuse an event that was only imagined, or dreamed about, with an event that actually occurred. One factor that contributes to memory distortion is that similar brain regions are important both for visual imagery and for the long-term storage of visual memories (Fig. 3.4–15). Another factor that contributes to memory distortion is that memory functions best in remembering the gist of an event, not the particulars from which the gist is derived. In a noted demonstration, people listen to a list of words: *candy, sour, sugar, tooth, heart, taste, dessert, salt, snack, syrup, eat,* and *flavor*. Subsequently, when asked to write down the words they heard, 40 percent of them wrote down the word *sweet,* even though this word did not appear in the list. Thus, many people in this demonstration failed to discriminate between the words that had been presented and a word that was strongly associated with all the words but had not itself been presented. The word *sweet* can be thought of as a gist word, a word that represents the other words and that captures the meaning of the whole list. Presumably, the words in the study list evoked a thought of the word *sweet,* either at the time of learning or during the memory test, and people then tended to confuse merely thinking of the word with actually hearing it.

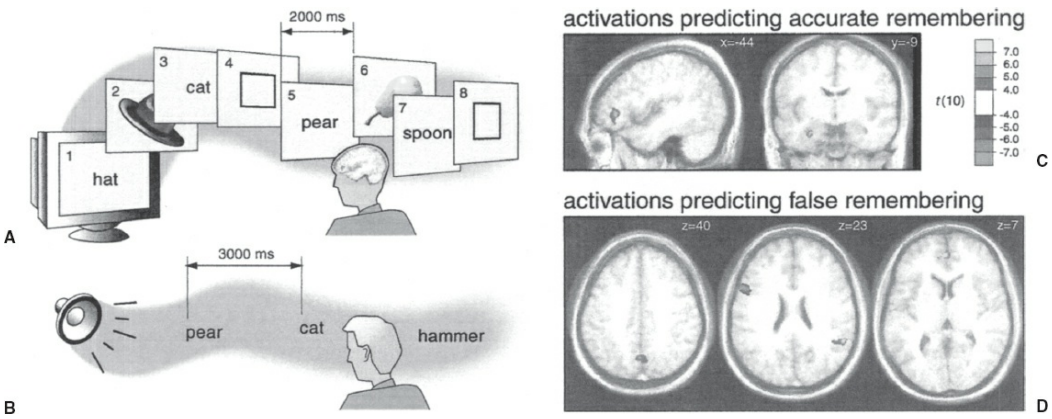

FIGURE 3.4–15. Neural substrates of false memories. **A:** Functional magnetic resonance imaging data were acquired in a learning phase, when subjects read names of objects and visualized the referents. One-half of the names were followed 2 seconds later by a picture of the object. **B:** In a surprise memory test given outside the scanner, subjects listened to object names and decided whether they had seen a picture of the corresponding object. On some trials, subjects claimed to have seen a picture of an object that they had only imagined. **C:** Results showed that left inferior prefrontal cortex and left anterior hippocampus were more active during learning in response to pictures later remembered compared to pictures later forgotten. **D:** Several different brain areas showed a greater response to words in the learning phase that were later falsely remembered as pictures, compared to words not misremembered. Activations that predicted false remembering were found in a brain network important for the generation of visual imagery in response to object names (precuneus, inferior parietal cortex, and anterior cingulate, shown in **left,**

middle, and **right** images, respectively). (**C** and **D** adapted from Gonsalves B, Reber PJ, Gitelman DR, Parrish TB, Mesulam MM, Paller KA. Neural evidence that vivid imagining can lead to false remembering. *Psychol Sci.* 2004;15:655.)

The reconstructive nature of recollection means that the interpretation of eyewitness testimony is not straightforward. Whole episodes are not available in the neocortex, but rather must be pieced together based on fragmentary components and in the context of potentially misleading influences present at the time of retrieval. Studies in adults and in children have documented that illusory memories can be created. Children are particularly susceptible to these effects, especially when subjected to leading questions and false suggestions.

In view of these features of memory, when a memory for childhood abuse is remembered after many years, it is prudent to ask whether the memory is accurate. Genuine examples of memory recovery have been documented, whereby an individual produces a veridical memory for a past traumatic event after not recalling the event for extended time periods. Numerous examples of apparent memory recovery have also been subsequently discovered to be instances of false memory. Unfortunately, there is no perfect method at present, in the absence of independent corroboration, for determining whether a recollective experience is based on a real event or not.

Infantile Amnesia

The biology of memory has also provided insights relevant to the phenomenon of infantile amnesia, the apparent absence of conscious memory for experiences from approximately the first 3 years of life. Traditional views of infantile amnesia have emphasized repression (psychoanalytic theory) and retrieval failure (developmental psychology). A common assumption has been that adults retain memories of early events but cannot bring them into consciousness. However, it now appears that the capacity for declarative memory does not become fully available until approximately the third year of life, whereas nondeclarative memory emerges early in infancy (e.g., classical conditioning and skill learning). Thus, infantile amnesia results not from the adult's failure to retrieve early memories, but from the child's failure to store them adequately in the first place.

Nevertheless, studies in young infants show that a rudimentary capacity for declarative memory is present even at a few months of age. As a child develops, memories can be retained across increasingly long intervals, and what is represented becomes correspondingly richer and more full of detail. Medial temporal and diencephalic regions seem to be sufficiently developed during these early months and years. What limits the capacity for declarative memory appears to be the gradual development and

differentiation of neocortex.

As the neocortex develops, the memories represented there become more complex, language abilities allow for more elaborate verbal descriptions of events, and a developing sense of self supports autobiographical knowledge. As new strategies emerge for organizing incoming information, declarative memories become more persistent, more richly encoded, and better interconnected with other information. It is not the case that fully formed childhood memories are stored but cannot be retrieved. The perspective consistent with current understanding of the biology of memory is that declarative memories formed very early in life are fragmentary, simple, and tied to the specific context of an infant's understanding of the world. They are unlike typical declarative memories in adults, imbued with meaning and a complex understanding of events.

Memories and the Unconscious

The existence of multiple memory systems also has implications for issues central to psychoanalytic theory, including the construct of the *unconscious*. How past experience influences current behavior depends on what view one takes of the nature of memory. By the traditional view, memory is a unitary faculty, and representations in memory vary mainly in strength and accessibility. Material that is unconscious is below some threshold of accessibility but could potentially be made available to consciousness.

The modern, biological view begins with the distinction between a kind of memory that can be brought to mind—declarative memory—and other kinds of memory that are, by their nature, unconscious. Stored nondeclarative memories are expressed through performance without affording any conscious memory content. Personalities are shaped by nondeclarative memories in the form of numerous habits and conditioned responses. In this view, one's behavior is indeed affected by events from early life, but the effects of early experience persist in a nondeclarative form without necessarily including an explicit, conscious record of the events. Learned behavior can be expressed through altered dispositions, preferences, conditioned responses, habits, and skills, but exhibiting such behavior need not be accompanied by awareness that behavior is being influenced by past experience, nor is there a necessity that any particular past experience has been recorded as a complete episode. That is, an influence from early experience does not require a memory of any specific episode. One can be afraid of dogs without remembering being knocked down by a dog as a child. In this case, the fear of dogs is not experienced as a memory. It is experienced as a part of personality. Furthermore, a strong fear of dogs carries with it no implication that the brain retains a specific record of any early experience that subsequently resulted in a fear of dogs.

Behavioral change can occur by acquiring new habits that supersede old

ones or by becoming sufficiently aware of a habit that one can to some extent isolate it, countermand it, or limit the stimuli that elicit it. However, one need not become aware of any early formative event in the same sense that one knows the content of a declarative memory. The unconscious does not become conscious. Various forms of nondeclarative memory simply influence behavior without having the additional capacity for these influences to become accessible to conscious awareness.

ACKNOWLEDGMENT

The authors gratefully acknowledge contributions to the present article from Larry Squire (author of "Biology of Memory" in prior editions of this book).

REFERENCES

Addis DR, Barense M, Duarte A, eds. *The Wiley Handbook on The Cognitive Neuroscience of Memory*. Hoboken, NJ: Wiley-Blackwell; 2015.

Baddeley AD, Wilson BA, Watts N, eds. *Handbook of Memory Disorders*. 2nd ed. New York: John Wiley and Sons; 2002.

Barrett LF, Lewis M, Haviland-Jones JM, eds. *Handbook of Emotions*. 4th ed. New York City, NY: Guilford Press; 2016.

Bayley PJ, Frascino JC, Squire LR. Robust habit learning in the absence of awareness and independent of the medial temporal lobe. *Nature*. 2005;436:550.

Bayley PJ, Hopkins RO, Squire LR. Successful recollection of remote autobiographical memories by amnesic patients with medial temporal lobe lesions. *Neuron*. 2003;38:135.

Byrne JH, ed. *Learning and Memory—A Comprehensive Reference*. 2nd ed. Cambridge, MA: Academic Press; 2017.

Clark RE, Manns JR, Squire LR. Classical conditioning, awareness, and brain systems. *Trends Cogn Sci*. 2002;6:524.

Corkin S. *Permanent Present Tense: The Unforgettable Life of the Amnesic Patient, H.M*. New York: Basic Books; 2013.

Dalgleish T, Werner-Seidler A. Disruptions in autobiographical memory processing in depression and the emergence of memory therapeutics. *Trends Cogn Sci*. 2014;18:596.

Dew IT, Cabeza R. The porous boundaries between explicit and implicit memory: Behavioral and neural evidence. *Ann NY Acad Sci*. 2011;1224(1):174–190.

Eichenbaum H. *Still searching for the engram*. Learn Behav. 2016;44:209.

Fuster JM. *Cortex and Mind: Unifying Cognition*. New York: Oxford Press; 2003.

Heckers S, Konradi C. Hippocampal pathology in schizophrenia. *Curr Top Behav Neurosci*. 2010;4:529.

Inostroza M, Born J. Sleep for preserving and transforming episodic memory. *Annu Rev Neurosci*. 2013;8:79–102.

Kandel ER. Biology and the future of psychoanalysis: a new intellectual framework for psychiatry revisited. *Am J Psychiatry*. 1998;155:457.

Kandel ER. The molecular biology of memory storage: a dialogue between genes and synapses.

Science. 2001;294:1030.

Kopelman MD. Disorders of memory. *Brain*. 2002;125:2152.

Paller KA. Consolidating dispersed neocortical memories: the missing link in amnesia. *Memory*. 1997;5:73.

Paller KA, Wagner AD. Monitoring the transformation of experience into memory. *Trends Cogn Sci*. 2002;6:93.

Parsons RG, Ressler KJ. Implications of memory modulation for post-traumatic stress and fear disorders. *Nat Neurosci*. 2013;16:146.

Rudoy JD, Weintraub S, Paller KA. Recall of remote episodic memories can appear deficient because of a gist-based retrieval orientation. *Neuropsychologia*. 2009;47:938–941.

Schacter DL, ed. *Memory Distortion: How Minds, Brains, and Societies Reconstruct the Past*. Cambridge, MA: Harvard University Press; 1995.

*Schacter DL. *Searching for Memory: The Brain, the Mind, and the Past*. New York: Basic Books; 1996.

*Squire LR, Kandel E. *Memory: From Mind to Molecules*. 2nd ed. New York: Scientific American Library; 2009.

*Squire LR, Schacter DL, eds. *Neuropsychology of Memory*. 3rd ed. New York: Guilford; 2002.

Squire LR, Zouzounis JA. Self-ratings of memory dysfunction: different findings in depression and amnesia. *J Clin Exp Neuropsychol*. 1988;10:727.

Staniloiu A, Markowitsch HJ. *Dissociative Amnesia*. Lancet Psychiatry. 2014;1:226.

Tulving E. Episodic memory: from mind to brain. *Annu Rev Psychol*. 2002;53:1.

*Voss JL, Lucas HD, Paller KA. More than a feeling: pervasive influences of memory processing without awareness of retrieval. *Cogn Neurosci*. 2012;3:193–207.

▲ 3.5 Consciousness and Dreaming from a Pathophysiological Perspective: The Thalamocortical Dysrhythmia Syndrome

RODOLFO R. LLINáS, M.D., PH.D.

A true appreciation of the interdependence between basic neurosciences and the clinical aspects of mental dysfunction has matured immensely over the last decades, and most particularly in the last few years. This rapport is both technological and conceptual. Indeed, spearheaded by imaging instrumentation with ever-increasing temporal and spatial resolution, a new picture of the normal and abnormal brain function is emerging while, simultaneously, new molecular/pharmacological and interventional tools have clarified the fact that psychiatry is but CNS pathophysiology. In place of the static phrenology of yesteryear, now the ever-changing dynamic geometry of ongoing brain function is being witnessed. For the first time, the temporal resolution of the current instrumentation matches the speed of cognition and motricity and the imaging is not restricted to superficial brain structures, as recently described by Llinás and Ustinin. Indeed, the

neuronal basis for cognition, once unattributable to any specific neuronal domain, is being assigned to the rich thalamocortical interconnectivity and the recurrent nature of the thalamocortical neuronal dynamics. This is most particularly the case for the temporal interaction between the specific and nonspecific thalamic loop activities. Such recurrence places the thalamus as a hub from which any site in the cortex can communicate with any other cortical site or sites.

COGNITION AS AN INTRINSIC FUNCTIONAL STATE OF THE BRAIN

It was suggested more than a decade ago that consciousness is an oneiric-like functional state that is modulated, rather than generated, by the senses. According to this view, internal events such as thinking, imagining, or remembering are strictly confined to intrinsic neuronal activity looping continuously through the brain's neuronal circuits. This proposal is supported by the fact that a large percentage of the brain's connectivity is recurrent and is not necessarily driven by sensory input but rather by intrinsic neuronal activity. This becomes self-evident when considering introspection or dreaming.

Thus, consciousness (i.e., being awake and able to feel, judge, respond, and remember) is but one functional state of the brain. Other conditions, such as being in one of the many dreamless sleep states or in coma, do not support consciousness or even the feeling of self-awareness. Perhaps the most spectacular difference concerning global brain states is that between wakefulness and dreamless delta sleep. It is recognized that there are no gross morphological changes in the brain during sleep that could explain the enormous disparity between sleep and wakefulness. The difference, of necessity, is functional. If a strong sensory stimulus occurs during sleep (e.g., the havoc wrought by an alarm clock), a person can awaken extraordinarily rapidly. It is also clear that, given the large number of neuronal elements involved, the only substrate capable of supporting the speed of these changes, from one state to the other, must be electrical and time coherent in nature. That is, the globally organized electrical activity of neurons and their synaptic interactions comprise the only mechanism that is fast enough to initiate or terminate abruptly wakefulness from the dream state. These, together with the speed with which humans can perceive and respond to their environment, are among the most important clues concerning the nature of consciousness.

Concerning the localization of such activity, it is known from classical neurology that damage to the cerebral cortex in mammals can cause a variety of well-defined dysfunctional conditions that modify or curtail consciousness. Visual cortex damage is accompanied by blindness that can be of different types, depending on the precise location of the cortical insult. Similar findings are also encountered in other cortical structures. Thus, injuries to the auditory, somatosensory, motor, or premotor cortices

are accompanied by well-defined conditions. In fact, the first such neurological lesion to be described in the terms used previously was that after damage of Broca area, generating dysarthria.

Given the previous statement, it has been historically accepted that consciousness can be equated with cortical function, more or less exclusively. However, this view ignores the fact that the nervous system is deeply recursive in its connectivity. As such, a dynamically recurrent corticosubcortical organization is more likely than a strictly hierarchical organization.

There are several other problems with the hierarchical proposal relating to the categorization of cortical activity into ever-increasing functional specialization. This hierarchy converges into the famous grandmother-neuron concept (i.e., the idea that specialized neurons represent each object in cognition, for example, the image of a person's grandmother). Among the problems with the hierarchical proposal is that the number of possible percepts is much larger than the total number of neurons in the cortical mantle. To make the matter worse, given the proposed hierarchical architecture, the number of such neurons would have to be a small percentage of all cortical neurons. Another difficulty with a hierarchical model is that of sampling size. Indeed, the number of specialized neurons, each representing a component of reality, would be so exorbitantly large that addressing specific neuronal elements would make the retrieval problem immense. Thus, even considering that neuronal elements integrate and transmit signals at a millisecond rate from the onset of sensory primitives, exhausting all sequential combinations of all grandmother cells would be awkwardly time intensive. This is not the case, however, concerning cognition. It takes roughly the same amount of time to recognize that a face is familiar as it takes to recognize that it is not. Yet, in any sequential strategy, it would take much longer to conclude no familiarity (as it would require comparing it with "all known faces") than it would to conclude familiarity (the search would proceed for only as long as necessary to "find" the match). From a different perspective, the grandmother-neuron hypothesis fails to explain how the unique perceptual insights of such neurons (the specific elements in a given category) are communicated to the rest of the nervous system. How do grandmother cells tell the rest of the neurons what they know, given their unique position at the top of a hierarchy?

Alternatively, if categories are generated by spatiotemporal mapping in the thalamocortical system, a dynamic representation based on temporal coherence would have the necessary speed to support cognition time constants. Thus, a simultaneity mapping may be envisioned that takes advantage of the parallel and synchronous organization of brain networks to generate cognition.

THALAMOCORTICAL GAMMA-BAND RESONANT COLUMNS

The hypotheses to be discussed here are derived from two areas of research: (1) single-neuron and network electrophysiological recordings made in vivo and in vitro in animals and (2) measurements made using noninvasive imaging of human brain function, in particular, magnetoencephalography (MEG). The central issue to be discussed is the assumption that the intrinsic electrical properties of neurons and the dynamic events resulting from their connectivity result in global resonant states known as *consciousness*.

Given this approach, two related views are present in the current literature regarding the actual mechanism for this functional condition. One view is that coherent events occur at the cortical level and that such cortical events are the primary binding substrate. A second view proposes that the binding event must be thalamocortical rather than purely cortical. Some of the reasons for supporting the latter view are sketched in the following discussion.

Even though the cortex receives a large number of nonthalamic inputs, damage to the thalamus is cognitively equivalent to damage to the cortex to which it projects. Experimental results from studies using noninvasive techniques, such as MEG in humans, as well as animal research using extracellular and intracellular recordings in vivo, indicate that cognition is supported by resonant recurrent electrical activity between thalamic and cortical structures at gamma-band frequencies (i.e., having a 25- to 50-Hz oscillatory frequency), often centered close to 40 Hz. Such findings favor the hypothesis that cognitive events depend on activity involving thalamocortical resonant columns. Indeed, the neuronal mechanisms responsible for the high-frequency thalamic oscillations that support thalamocortical synchronization and coherence are now beginning to be understood.

OSCILLATORY PROPERTIES OF THALAMIC CELLS

In the early 1980s, it was suggested that the presence of neuronal elements with intrinsic oscillatory membrane potential properties would support network oscillatory resonance, which facilitate coherence between interconnected elements. Such membrane potential oscillations were recorded in thalamic neurons as intrinsic activity and were related to the state of consciousness. Low frequency (1 to 6 Hz), being related to sleep, are generated by the activation of low-threshold calcium channels (T channels) in thalamic neurons. Middle frequencies (9 to 11 Hz) thalamic activity is associated with wakefulness with eyes closed (alpha rhythm) and fully awake with eyes open with thalamic high frequency (30 to 80 Hz) known as gamma-band (see Fig. 3.5–1A,B). Moreover, in a series of in vitro studies, the intrinsic electrical properties of thalamic neurons that support

gamma-band frequencies were characterized. These intrinsic properties proved to be subthreshold membrane potential oscillations generated in thalamus and cortex neurons at the dendritic level when thalamic neurons were depolarized beyond −45 mV (Fig. 3.5–1A,B). Indeed, the neuronal mechanism underlying the generation of gamma oscillations has been studied at the level of single neurons and of neuronal circuits at sites other than the thalamus. For example, it has been shown that the membrane potential of sparsely spiny inhibitory neurons in cortical layer IV supports gamma-frequency membrane voltage oscillations (Fig. 3.5–1C). The mechanism for the oscillation, in this case, is the sequential activation of persistent low-threshold sodium currents.

Dendritic oscillations are functionally significant because the return input from the cortex to thalamic neurons terminates in distal dendritic segments, providing a unique opportunity for resonance between intrinsic dendritic oscillation and thalamic synaptic inputs. This issue is of further significance because waking and rapid eye movement (REM) sleep states are associated with thalamic neuron depolarization, whereas slow wave sleep is associated with thalamic hyperpolarization. The two levels of membrane potential elicit a different intrinsic oscillatory property in thalamic cells. Given the previous statement, it can be concluded that the coherence of fast rhythms in thalamocortical loops depends on the pattern of synaptic inputs and on the responsive state of the neurons as modulated by their intrinsic properties.

Because sensory inputs generate but a fractured representation of universals, the issue of perceptual unity concerns the mechanisms that allow these different sensory components to be gathered into one global image. In recent years, this has been described as *binding,* and one of the mechanisms proposed for such implementation is known as *temporal conjunction*.

Synchronous neuronal activation during sensory input has been studied in mammalian visual cortical cells as light bars of optimal orientation and displacement rate are presented. The components of a visual stimulus that correspond to a singular cognitive object—for example, a line in a visual field—yield coherent gamma-band oscillations in regions of the cortex that may be as far as 7 mm apart or that may even be in the contralateral cortex. Moreover, gamma-band oscillatory activity between related cortical columns is highly correlated under such circumstances. In addition, MEG recordings in the early 1990s demonstrated coherent 40-Hz thalamocortical oscillations in an awake human. These oscillations could be reset by sensory stimuli and displayed a high degree of spatial organization. This geometrical peculiarity makes such an oscillatory frequency particularly apt to support the temporal conjunction of rhythmic activity over a large ensemble of neurons while conserving a high degree of spatial differentiation.

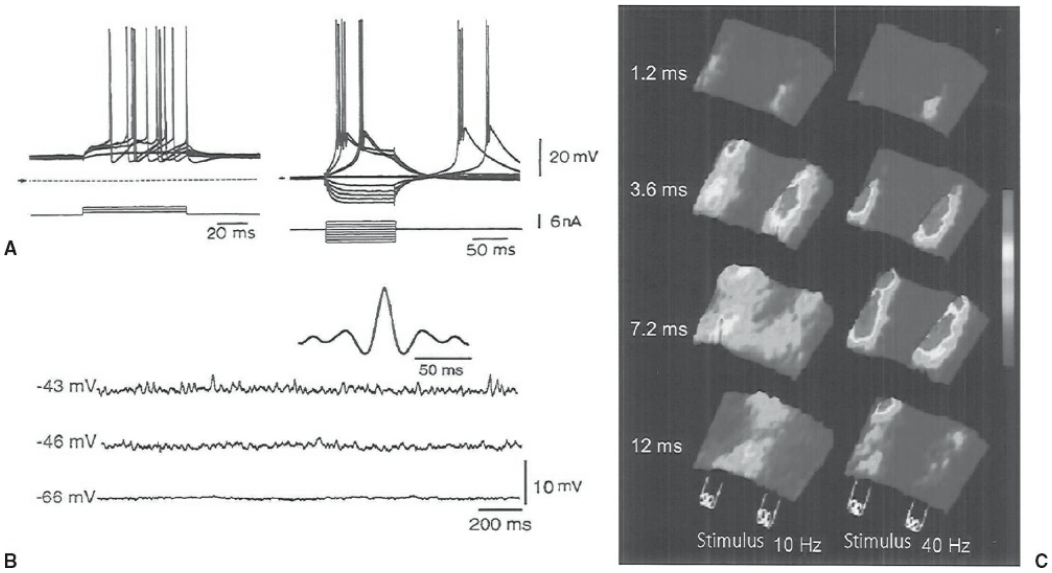

FIGURE 3.5–1. Oscillatory properties of thalamic neurons. In vitro intracellular recording from a thalamic neuron. **A:** Left, direct activation of a thalamic cell from a resting potential of –50 mV generates repetitive high threshold spike activation of the thalamic neuron. Right: The same-amplitude direct stimulation at –65 mV evoked low threshold spike activation. Hyperpolarizing current pulses were followed by rebound low threshold spikes. **B:** Gamma band frequency subthreshold oscillations were evoked by protracted outward direct current injection through the recording microelectrode (bringing the membrane from –66 mV to –46 mV and –43 mV). The autocorrelogram in the inset corresponds to the –43 mV trace. (Modified from Pedroarena C, Llinas R. Dendritic calcium conductances generate high-frequency oscillation in thalamocortical neurons. *Proc Natl Acad Sci USA.* 1997;94:624.) **C:** Voltage-gated imaging from a rodent cortical slice activated with two bipolar electrodes: Left row response to 10 repetitive stimuli at 10 Hz at which frequency the columnar cortical activation vanished quite rapidly. Note that, by contrast, at 40 Hz the columnar activity distribution is maintained and covered most of the cortex depth in the slice, demonstrating that the geometry of cortical activation is frequency dependent. (Adapted from Llinás R, Ribary U, Contreras D, Pedroarena C. The neuronal basis for consciousness. *Philos Trans R Soc Lon.* 1998;353(1377):1841–1849.)

The possibility of coherent temporal–spatial conjunction in thalamocortical circuits has been supported by the direct imaging of the spatiotemporal distribution of afferent activity using voltage-sensitive dye imaging in rodent visual cortex in vitro. Indeed, the delivery of two cortical afferent inputs repeated at low frequency (10 Hz or less) gave rise to two waves of excitation (Fig. 3.5–2A, upper left quadrant) that moved horizontally and summed over a wide cortical area (Fig. 3.5–2A, upper right quadrant) with a time course of 10 ms. This spatial distribution of activity lacks the high granularity required for a rich cognitive discrimination and does not conform to the specific columnarity observed in the somatosensory and visual cortices. In contrast, when stimuli were delivered at a higher frequency (40 Hz), the geometry of activation is rapidly restricted to submillimeter areas, that is, to cortical column size (Fig. 3.5–2A, lower right quadrant), rather than the wider geometry observed with low-frequency activation.

These results agree with data showing that slow-frequency brain

rhythms (less than 15 Hz), such as those characteristic of slow wave sleep (e.g., sleep spindles and slow oscillations), demonstrate long-range cortical coherence sometimes spanning the entire neocortex. In contrast, during activated states, such as waking or REM sleep, high-frequency oscillations (20 to 50 Hz) show a pattern of coherence that is restricted to its immediate vicinity or occurs between distant discrete areas. More fundamentally, however, a pharmacological examination of the mechanism by which the granularity of activation is engendered indicates that the areas of silence between patches of activity at 40 Hz are generated by active inhibition. Thus, in the presence of γ-aminobutyric acid (GABA) blockers, the spatial filtering of cortical activity described previously disappears. These results clearly agree with the findings that cortical inhibitory neurons are capable of high-frequency oscillation and with the view that, if such neurons are synaptically coupled and fire in synchrony, they might be formative in generating cortical gamma-band activity.

Such in vivo findings suggest yet another function for coherent inhibition at the cortical level: that of generating the *thalamocortical resonant column*. From this point of view, the thalamocortical resonant column would comprise a basic functional unit in the generation of consciousness. Thus, these columns may be viewed as a permissive network endowed with frequency-gated spatial filter properties. Accordingly, frequency forges activity into well-specified cortical activation geometry by increasing the encoding contrast as the activity areas are intensified and focused by gamma-band oscillation. By increasing the contrast between activation and inhibition, such an organization may underlie the columnar organization observed in vivo during visual stimulation of the cortical mantle. This would have the added advantage of a thalamocortical resonance that might link this expanded columnar activation into the binding patterns required for consciousness. The picture that emerges that of a thalamocortical resonance based cognition is, however, far more complicated. It becomes evident if the brain response to a simple but short visual stimulus is analyzed, then the recognition of a simple image, that is, a letter, is projected on to a screen presented for short but increasing periods of time. This has been recently studied using MEG recordings, by K. Sekar and colleagues. The findings are quite revealing as, with short lasting visual stimuli (5 to 9 ms), a clear response may be recorded 100 +/− 9 ms later from the primary visual cortex, in the absence of any visual sensation. As the visual stimulus is increased in duration of 10 to 14 ms, a secondary thalamocortical response with a latency of 238+/− 21 ms is recorded in the parietal association cortex. When that occurs the individual will report seeing something but not recognizing what. A further increase in stimulus duration (16 to 18 ms) activates the temporal cortex via the thalamus. At that point the individual will be able to say that the stimulus was a number and will name it. The

process takes 293+/−25 ms. Notice, then, that visual perception is not simply activation of the visual cortex but rather a thalamocortical resonance event involving, at least, three thalamic and three cortical structures.

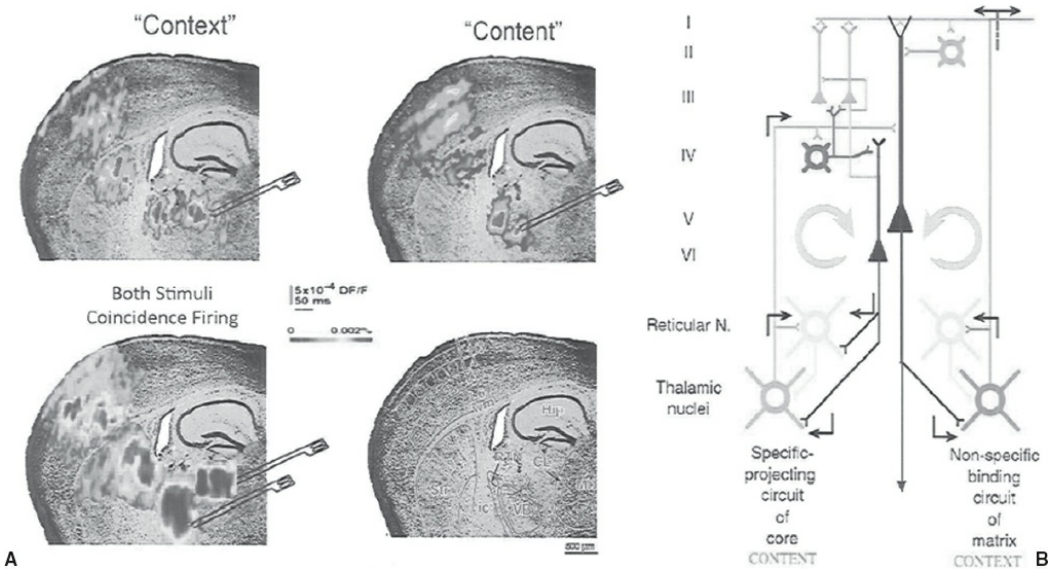

FIGURE 3.5–2. **A:** Voltage imaging of thalamocortical activation of a somatosensory brain slice demonstrating temporal binding of specific and non-specific thalamocortical inputs. Upper left quadrant, response average of a set of single pulse stimulation of centrolateral (CL) thalamic nucleus (Context). Upper right quadrant response average to a set of single stimuli to the ventrobasal (VB) thalamic nucleus (Content). Both CL and VB stimulation activated the thalamic reticular nucleus and the striatum. A different cortical activation patterns was, however, observed for each stimulation site. Thus, Cl stimulation (upper left quadrant) activated layers 6, 5, and 1, while VB activated layer 4, 2/3, and 5, 6. When both thalamic nuclei were activated simultaneously (lower left quadrant), a clear increase in activity was observed at all sites, indicating functional resonance as expected during gamma band thalamocortical binding. **B:** Diagram of two thalamocortical systems. **Content:** Specific sensory or motor nuclei project to cortex layer 4, generating cortical oscillation by direct activation of layer 4 neurons that contact dendrites of layer 5 and 6. Layer 6 pyramids generate return inhibition to specific thalamus via the reticular thalamic nucleus. **Context:** The second loop relates to the nonspecific intralaminar nuclei projecting to the most superficial layer of the cortex and giving collaterals to the reticular nucleus. Pyramidal cells in layer 5 return the oscillation to the reticular and the nonspecific thalamic nuclei, establishing a second resonant loop. The conjunction of the specific and nonspecific loops (content and context) is proposed to generate temporal cognitive binding. (See color atlas.) (Modified from Llinás R, Leznik E, Urbano IF. Temporal binding via coincidence detection of specific and non-specific thalamocortical inputs: a voltage dependent dye imaging study in mouse brain slices. *Proc Natl Acad Sci USA.* 2002;99:449.)

From this perspective, when the interconnectivity of the thalamocortical system is combined with the intrinsic properties of the individual neurons, a network for resonant neuronal oscillation emerges in which specific corticothalamocortical reentry circuits would tend to resonate at gamma frequency. According to this hypothesis, neurons at different levels, most particularly those in the thalamic reticular nucleus as well as the corticocortical interhemispheric and transcallosal connectivity and

resonance, would be engaged. Such a functional set of events would be responsible for the synchronization of gamma oscillation at distant thalamic and cortical sites that are responsible for visual perception.

BINDING OF SPECIFIC AND NONSPECIFIC GAMMA-BAND ACTIVITY

Coherent Gamma-Band Thalamocortical Activity and Temporal Binding

Given the spatially fractured nature of sensory representation over the corticothalamic realm, activity such as that described above addresses fundamental issues concerning the mechanism responsible for the unity of perception. As stated previously, one likely possibility is that such binding occurs, via temporal coincidence, of specific and nonspecific thalamocortical synaptic activation over the length the apical and basal dendrites of pyramidal cells. Such binding may also occur via interhemispheric ipsilateral and contralateral (callosal) connectivity. This would ultimately support coherence in deep brain cellular clusters, establishing a resonant thalamocortical reactivation loop. Because many such loops are activated simultaneously, coherent function can be supported in brain areas remote from each other.

These large sets of isochronous events, supported by the oscillatory properties of thalamic and cortical neurons, can conjoin all sensory inputs on the basis of temporal coincidence, resulting in one unified brain activation pattern. For this mechanism to be useful, the specific thalamic inputs must constantly update cortical structures about external events (content), whereas the nonspecific thalamic inputs serve to bind content information on the basis of internal significance (context) arising from association cortices. According to this hypothesis, the activation of the thalamus via corticothalamic recurrent activity would serve to maintain a self-supporting feedback activity loop that is continuously modified by the incoming sensory information. That such specific and nonspecific thalamocortical inputs can sum at cortical level was tested experimentally using voltage-sensitive dye imaging in in vitro thalamocortical slice preparation. In these experiments, specific (ventrobasal [VB]) and nonspecific (centrolateral [CL]) thalamic nuclei were stimulated at gamma-band frequency. The spread of activity from VB and CL nuclei into the cortex and the interaction of their activity at the cortex were imaged directly.

A representative cortical activity distribution, visualized using voltage dependent neuronal imaging, follows VB and CL stimulation. This image superimposed on a Nissl-stained thalamocortical slice, is illustrated in Figure 3.5–2B. VB stimulation was followed by synaptic activation of reticular nucleus, striatum–putamen, and the somatosensory cortex. In the cortex, layer 4 activation was followed by a radial spread of activity to layers 2/3 and 5. By contrast, although CL stimulation was also followed by

the activation of reticular nucleus and striatum–putamen, the cortical activation occurred, exclusively, at the most superficial cortical layer, and in cortical layers 5, 6. This response was quite widespread, especially in cortical layer one. These results directly demonstrate that thalamic specific and nonspecific connectivity to the somatosensory primary cortex results in quite discrete and well-defined cortical layers, which ultimately result in spatial binding of thalamic inputs at gamma-band frequency.

Temporal Binding of CL and VB Nuclear Activity at Cortical Level

The findings illustrated above address the question of whether gamma-band temporal binding of specific and nonspecific thalamic nuclei activity could be actually visualized using voltage-dependent dye imaging. The distribution of the activity during 40-Hz stimulation of the VB and CL nuclei demonstrated a supralinear summation of the individual VB and CL responses. Indeed, at gamma-band frequency, the focus of cortical activity, remained quite high for the duration of the train (Fig. 3.5–2A, lower left quadrant), as previously demonstrated for subcortical white matter stimulation.

Thus, specific and nonspecific thalamocortical inputs sum supralinearly in cortex in a time-sensitive fashion, where coincidence of layers 1 and 4 activation is integrated at the cortex, most probably over the apical dendrite of layers 5 and 6 pyramidal cells as reviewed by Spruston. This is a central issue because these two layers represent the return output to thalamus (via layer 6) and the main output to the rest of the nervous system (via layer 5). Concerning the thalamocortical projections, the fact that the nonspecific system activates cortical layer 1 is in agreement with the anatomical distributions of these terminal axons and addresses the importance of synaptic inputs to the tuft of pyramidal cell apical dendrites in that layer.

Input to the dendritic tuft has been shown experimentally, therefore, to serve as a coincidence integrator. Temporal binding was demonstrated by the summing of well-organized activation at layers 2/3 and 4 with thalamic Cl activation of layer 1. Thus, the temporal conjunction of input to layer 1 and layer 4 results in coincidence detection, as had been proposed on theoretical grounds and following results from single-cell multi-intracellular recordings as obtained by. This conjunction is significant because it represents a possible mechanism for the global temporal binding required to generate single cognitive events from the large number of sensory inputs arriving at the brain at any particular time.

The previously stated information may be summarized by using a simplified schematic diagram of a neuronal circuit that supports temporal binding (Fig. 3.5–2B). In this diagram, layers 5 and 6 pyramidal cells are represented by a single element. Gamma oscillations in neurons in specific thalamic nuclei establish cortical resonance through direct activation of

pyramidal cells and feed-forward inhibition through activation of 40-Hz inhibitory interneurons in layer 4 (Fig. 3.5–2B, red). These oscillations recur in the thalamus via layer 4 pyramidal cell axon collaterals, producing thalamic feedback inhibition via the reticular nucleus. A second system is illustrated in blue. Here, the nonspecific thalamic nuclei are shown projecting to cortical layers 1 and 5 and to the reticular nucleus. Layer 5 pyramidal cells return oscillations to the reticular nucleus and intralaminar nuclei. The cells in this complex have been shown to oscillate at gamma-band frequency and to be capable of recursive activation. Furthermore, it has also been shown by recent experiments by Llinas and colleagues that in the absence of gamma band activation, as occurs in rodents with P/Q calcium channel mutation, cognitive function is deeply impaired.

It is apparent from the literature that neither of these two circuits alone can generate cognition. Indeed, as stated previously, damage of the nonspecific thalamus produces deep disturbances of consciousness, whereas damage of specific systems produces loss of a particular modality. Although, at this early stage, it must be quite simple in its form, the previous statement suggests a hypothesis for the overall organization of brain function. This rests on two tenets: First, the *specific* thalamocortical system is viewed as encoding specific sensory and motor activity through the resonant thalamocortical system specialized to receive such inputs (e.g., the lateral geniculate nucleus [LGN] and visual cortex). The specific system is understood to comprise those nuclei, whether sensorimotor or associative, that project mainly, if not exclusively, to cortical layer 4. Second, after optimal activation, any specific thalamocortical loop would tend to oscillate at gamma-band frequency, and activity in the *specific* thalamocortical system would be easily *recognized* over the cortex by this oscillatory characteristic.

In this scheme, areas of cortex peaking at gamma-band frequency would represent those components of the cognitive world that have reached optimal activity at a particular time. The problem now is to coalesce such a fractured description into a single cognitive event. It is proposed that this could come about by the concurrent summation of specific with nonspecific 40-Hz activity along the radical dendritic axis of the pyramidal elements by coincidence detection.

In conclusion, the thalamocortical system would function on the basis of temporal coherence. Such coherence would be embodied by the simultaneity of neuronal firing based on passive and active dendritic conduction along apical dendritic core conductors. In this fashion, the time-coherent activity of the specific and nonspecific oscillatory inputs would enhance de facto 40-Hz cortical coherence and would provide one mechanism for global binding. The specific system would thus provide the content from the external world, and the nonspecific system would give rise to the temporal conjunction or the context concerned with attention to

particular incoming information. These would together generate a single cognitive experience. The actual demonstration of such principle is well illustrated by the hemineglect syndrome that follows damage of intralaminar thalamic damage or damage to association cortex in particular the parietal association circuit as described by Heilman and Watson in their hemineglect studies.

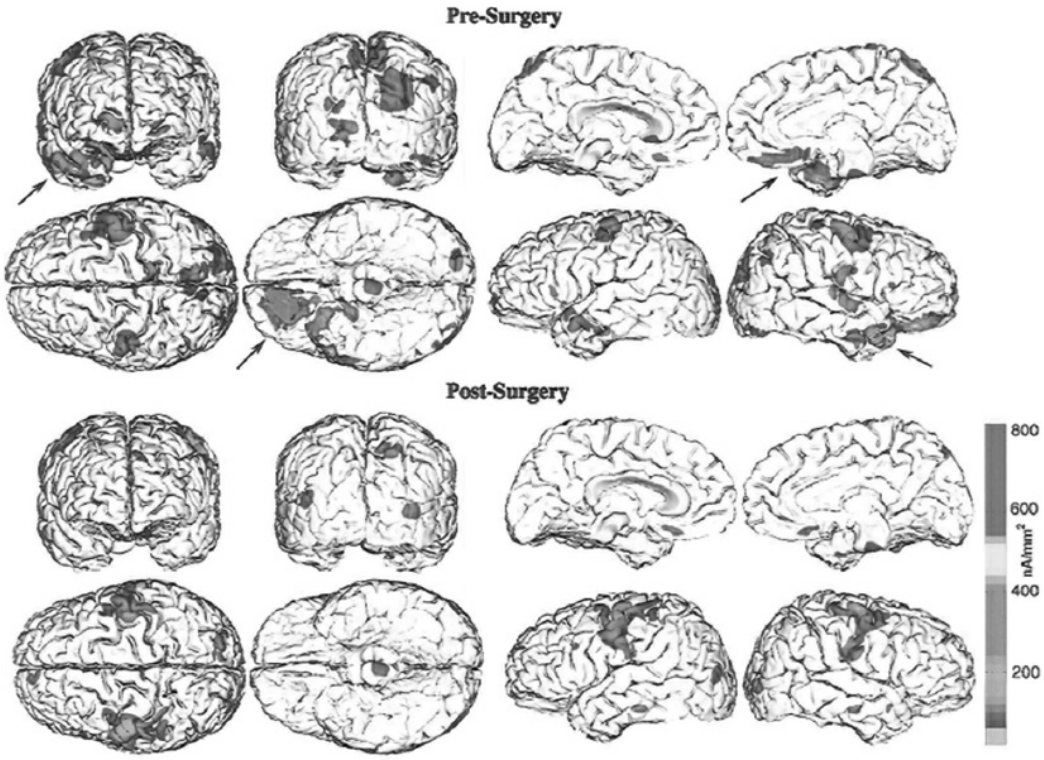

FIGURE 3.5–3. Magnetoencephalography source localization for a schizoaffective patient before and after selective miniablation at the Field of Forel. Projection of 4- to 10-Hz activity onto a magnetic resonance image of the whole brain before (top) and after (bottom) surgery. The thalamocortical dysrhythmia (TCD) pattern in this patient was localized (*arrows*) in the right-sided paralimbic domain comprising the temporopolar, anterior parahippocampal, orbitofrontal, and basal medial prefrontal areas. This low-frequency focus disappears postsurgically and the patient was integrated into society as a normal individual. (Adapted from Jeanmonod D, Schulman J, Ramirez R, Cancro R, Lanz M. Neuropsychiatric thalamocortical dysrhythmia: surgical implications. *Thalamus Related Syst.* 2003;2(2):103–113.)

Thalamocortical Resonance as the Substrate for Consciousness: Human MEG Studies. The activity at the thalamocortical level encountered in the in vitro rodent experiments described previously is relevant to understanding human cerebral function, given the similarity in neuronal intrinsic properties and connectivity among mammalian forms. Although similar studies are impossible to implement in human tissue, thalamocortical recurrent activity was demonstrated over a decade ago in the human brain using MEG.

Furthermore, such MEG studies have indicated that these may be a minimal temporal interval for sensory discrimination—a binding quantum

—on the order of 15 ms. This finding indicated that consciousness is a noncontinuous functional process determined by the dynamic properties of the thalamocortical system. Moreover, continuously recorded spontaneous magnetic activity demonstrated similar gamma-band activity during wakefulness and REM sleep but not during delta sleep. In addition, although auditory stimuli produced well-defined 40-Hz oscillatory reset responses during wakefulness (Fig. 3.5–3), no such resetting was observed during slow wave or REM sleep in all subjects examined.

These findings indicated that, although the awake state and the REM sleep state are electrically similar with respect to the presence of 40-Hz oscillations, a central difference remains: the inability of the sensory input to reset 40-Hz activity during REM sleep. In contrast, during slow wave sleep, no gamma-band activity was observed, and, as in REM sleep, no 40-Hz sensory response could be evoked. Indeed, stimuli from the external world (below a certain level) are not perceived during REM sleep because they are out of context with the functional state being generated by the brain at that time. Putting it another way, the dreaming condition is a state of hyperattentiveness to intrinsic activity in which sensory input cannot easily access the machinery that generates conscious experience. This may also be the case in hallucinatory psychiatric conditions.

An interesting possibility in considering a morphophysiological substrate is that the nonspecific thalamic system, particularly the intralaminar complex, has an important role in coincidence generation. Indeed, neurons in this complex project in a spatially continuous manner to the most superficial layers of all cortical areas, including the primary sensory cortices. This possibility is particularly attractive, given that single neurons burst at 30 to 40 Hz, especially during REM sleep. This finding is consistent with the macroscopic magnetic records observed in this study and with the fact that damage to the intralaminar system results in lethargy or coma.

Introspection, Reality Emulation, and Cognition

As reviewed previously, several lines of research suggest that the brain is fundamentally a closed system capable of self-generated activity based on the intrinsic electrical properties of its component neurons and their connectivity. According to this hypothesis, the CNS is viewed as a reality-emulating system. The parameters of such a reality are determined by sensory input, but thalamocortical iterative activity provides the mechanism for the setting of such input into a coherent context resulting in consciousness. This is also in keeping with the observation that the thalamic input from the cortex is larger than that from the peripheral sensory system.

In addition, neurons with intrinsic oscillatory capabilities that reside in the complex thalamocortical synaptic network allow the brain to generate

dynamic oscillatory states that shape the functional events elicited by sensory stimuli. In this context, functional states, such as wakefulness, REM sleep, and other sleep stages, are prominent examples of the multiple states that brain activity can support. This hypothesis assumes that, for the most part, the connectivity of the human brain is present at birth and is fine tuned during normal maturation. This neurological a priori was suggested by early neurological research. It includes the identification by Pierre Paul Broca of a cortical speech center and the discovery of point-to-point somatotopic maps in the motor and sensory cortices and in the thalamus.

A second organizing principle may be equally important—one that is based on the temporal, rather than the spatial, relationships among neurons. This temporal mapping may be viewed as a type of functional geometry.

Thalamocortical Dysrhythmic Syndrome

Neuropsychiatric conditions have long been associated with paralimbic cortical areas and the mediodorsal thalamic nucleus. Indeed, studies of patients have revealed low-frequency oscillatory activity, histopathological abnormalities, and hypometabolism, as well as magnetic resonance (MR) changes, in these brain areas.

Recent MEG studies of patients with neurological or psychiatric conditions have suggested a significant correlation between these disorders and the thalamocortical dialogue hypothesis concerning the genesis of consciousness. Spontaneous MEG activity recorded from patients with neurogenic pain, tinnitus, Parkinson disease, depression, OCD, or schizoaffective syndrome shows increased low-frequency theta rhythmicity (Fig. 3.5–2A, upper panel) in conjunction with a widespread and marked increase of coherence among high- and low-frequency oscillations (Fig. 3.5–2A, lower panel). That is, a thalamocortical dysrhythmia (TCD), characterized by a continuous abnormal oscillatory state and by hypercoherence, is present in patients with the previously mentioned conditions.

The coherent theta activity—the result of a resonant interaction between thalamus and cortex—is clearly correlated with the generation of low-threshold calcium spike bursts by thalamic cells. The presence of these bursts is directly related to thalamic cell hyperpolarization brought about by excess inhibition or by disfacilitation. The emergence of positive clinical symptoms is viewed as resulting from ectopic gamma-band activation, which is referred to as the *edge effect*. This effect is seen as increased coherence between low- and high-frequency oscillations and probably results from inhibitory asymmetry between high- and low-frequency thalamocortical modules at the cortical level.

That the neuropsychiatric ailments discussed previously correspond to

the TCD syndrome is supported by the finding of low-threshold thalamic spike activity during neurosurgical procedures and the characteristic MEG patterns described previously in such patients.

The TCD syndrome is characterized by the following sequence of events:

1. Hyperpolarization of thalamic relay or reticular thalamic nucleus (RTN) cells or both, due to disfacilitation or overinhibition or both.
2. This hyperpolarized state is the source of T-type calcium channel deinactivation, causing the production of low-threshold spike (LTS) bursts in thalamic (Fig. 3.5–1) or reticular neurons or both.
3. Neurons in such a state impose a slow rhythmicity on their thalamocortical loops and are locked in the theta low-frequency domain by their membrane conductances. Recurrent corticothalamic and RTN projections back to the projection thalamus provide the necessary coherent diffusion of these frequencies to various related cortical areas. MEG and electroencephalogram (EEG) recordings demonstrate increased theta power under these conditions, which may be directly correlated with cortical and thalamic hypometabolism seen in PET results. In psychotic disorders, the disease source may be found in the paralimbic cortical domain or in the paralimbic striatum, leading to corticothalamic low-frequency activation.
4. The final step in the description of this syndrome is the proposed existence of the activation of gamma-frequency cortical domains due to an asymmetrical corticocortical GABAergic lateral inhibition. The proposed mechanism, named the *edge effect* (analogous to the edge effect observed in the retina due to lateral inhibition), would result from the asymmetrical inhibition between a low-frequency cortical area and neighboring high-frequency domains, providing an area of disinhibition that results in the activation of the cortex surrounding the low-frequency area. This is not unlike the visual imagery encountered in migraine. MEG studies support the postulation of such an edge effect, as evidenced by large multifrequency coherence between theta and gamma domains—an event absent in the normal brain. This activation of high-frequency areas might express itself through abnormal EEG spiking activity, as seen in patients with psychosis.

The characteristics of the TCD syndrome suggest the basis for a possible treatment that goes beyond the neuroleptics now used. A promising approach will be the development of T-type calcium channel blockers that would specifically address particular thalamic groups. Such agents would control the low-frequency intrinsic thalamic activity seen in the type of patient described previously. Although these specific drugs are being developed, procedures that have been used quite favorably in patients with Parkinson disease may be implemented in these patients. These include deep brain stimulation (DBS) or selective miniablation (SMA) of unspecific

thalamic or prethalamic structures.

A 50-year-old male patient presented with a 20-year history of a schizoaffective disorder characterized by delusional, hallucinatory, and affective bipolar manifestations. Over a few weeks after the second CL thalamic and anterior media pallidal stereotaxic microlesion, the patient experienced progressive and then complete relief of all of these disease signs. Three years after surgery, he enjoyed a complete improvement in stability. Initially, he experienced two episodes of agitation, anxiety, and confusion, followed by a reactive depressive phase. These were apparently based on internal stressors related to personal and interactive family factors, with activation of powerful guilt and self-insufficiency feelings. These initial events were followed by the recovery from the aforementioned symptoms. Drug reduction was being processed very slowly, however, in view of the 20-year-long intake. Figure 3.5–2 describes the MEG power spectrum and coherence plot and localization of theta-band MEG activity before and 3 months after surgery.

The basic assumption concerning the genesis of this syndrome is that TCD is a CNS intrinsic property brought about by changes in intrinsic voltage-gated ionic conductances at the level of thalamic relay cells, namely, the deinactivation of T-type calcium channels by cell membrane hyperpolarization. LTS bursts are thus produced, which lock the affected portion of thalamocortical circuits in a low-frequency resonance attractor. Low-frequency loops disfacilitate anatomically related cortical circuits, giving rise to gamma-band activity, the *edge effect,* and to the generation of positive symptoms, such as those seen in tinnitus, neurogenic pain, Parkinson disease, and some neuropsychiatric disorders. In these cases, the dysrhythmic mechanism originates in a bottom-up fashion, that is, from the thalamus toward the cortex. In other situations, such as epilepsy, neuropsychiatric conditions of cortical origin, and central cortical neurogenic pain, a top-down mechanism may be at work. This would be triggered by a reduction in the corticothalamic input. Bottom-up and top-down situations result in excess inhibition at the thalamic level or in disfacilitation, generating thalamic cell membrane hyperpolarization and low-frequency oscillations. The proposal, then, is that the same mechanism that is responsible for the genesis of consciousness can generate neuropsychiatric conditions when its organization and timing are altered.

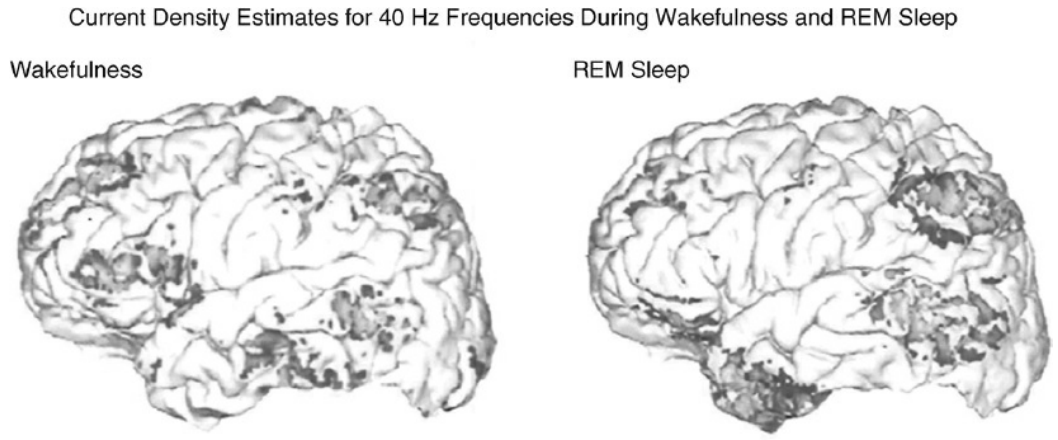

FIGURE 3.5–4. Gamma-band oscillations during dreaming. Left image: Gamma-band oscillation evoked in an awake subject by a stimulus (auditory click). Right image: Similar stimulus does not evoke a response in delta sleep (*right image*). However, in REM sleep (*right image*) while not evoking a sensory response demonstrates spontaneous gamma band activity. Note also that during dreaming frontal lobe activity at gamma band is decreased, while association cortices (both parietal and temporal and the temporal pole) show high level gamma band activity. (Adapted from Llinás R, Ribary U. Coherent 40-Hz oscillations characterize dream state in human. *Proc Natl Acad Sci USA.* 1993;90:2078.)

Cognition, a property of thalamocortical cycling, would function on the basis of temporal coherence. Such coherence would be embodied by the simultaneity of neuronal firing based on passive and active dendritic conduction along the apical dendritic core conductors. In this fashion, the time-coherent activity of the specific and nonspecific oscillatory inputs, by summing distal and proximal activity in given dendritic elements, would enhance de facto 40-Hz cortical coherence by their multimodal character and, in this way, would provide one mechanism for global binding. The *specific* system would provide the *content* that relates to the external world, and the *nonspecific* system would give rise to the temporal conjunction, or the *context* (on the basis of a more interoceptive context concerned with alertness), that would together generate a single cognitive experience. Furthermore, when this rhythmicity is altered in particular ways, neurological and psychiatric conditions ensue.

Psychotherapy and Brain Activity

Of interest at this time is the relationship between the mental and the physical. And so, in discussion with Prof. Robert Cancro, the issue came up as to whether an everyday event such as a psychotherapy event, and its result, could be in anyway measured by an unbiased methodology that would confirm that, indeed, psychological entities are amenable to direct demonstration via changes in brain activity. And thus it was decided that a patient who responds clearly to psychotherapy should be studied, using an imaging technology and most significantly, whether an MEG recording session during psychotherapy could be implemented in the recording room of the author's MEG system. As shown in Figure 3.5–4 a set of recordings were obtained from a patient responding to psychotherapy, where modulation of mood was modifiable during the session. The results obtained were what was to be expected if, indeed, psychotherapy was changing the basic functional organization of brain function. Indeed, an unambiguous change in the slow frequency oscillation was clearly observed during the psychotherapy session and is illustrated in Figure 3.5–5. Furthermore, the effect of psychotherapy induced a change of close to an order of magnitude in the amplitude of the abnormal thalamocortical activity.

FUTURE DIRECTIONS

In the immediate future, diagnosis, prognosis, and treatment are poised for fundamental changes that will not only modify the current view of psychiatry but also will help to amalgamate the neurosciences into a single hard science posture. Diagnosis, with better correlation between structure and function and the increased resolution afforded by better imaging instrumentation (high-field functional MRI, high-resolution EEG and MEG), as well as by noninvasive procedures such as transcranial magnetic stimulation, will be invaluable in defining nosology. Several invasive technologies, such as cortical and DBS for certain psychiatric conditions, are clearly of interest on the immediate horizon. Other more esoteric approaches, such as transvascular stimulation/recording paradigms using nanowire technology (either metallic or polymeric), are beginning to emerge. Although only in its initial experimental stage, such partially invasive avenues should also be considered.

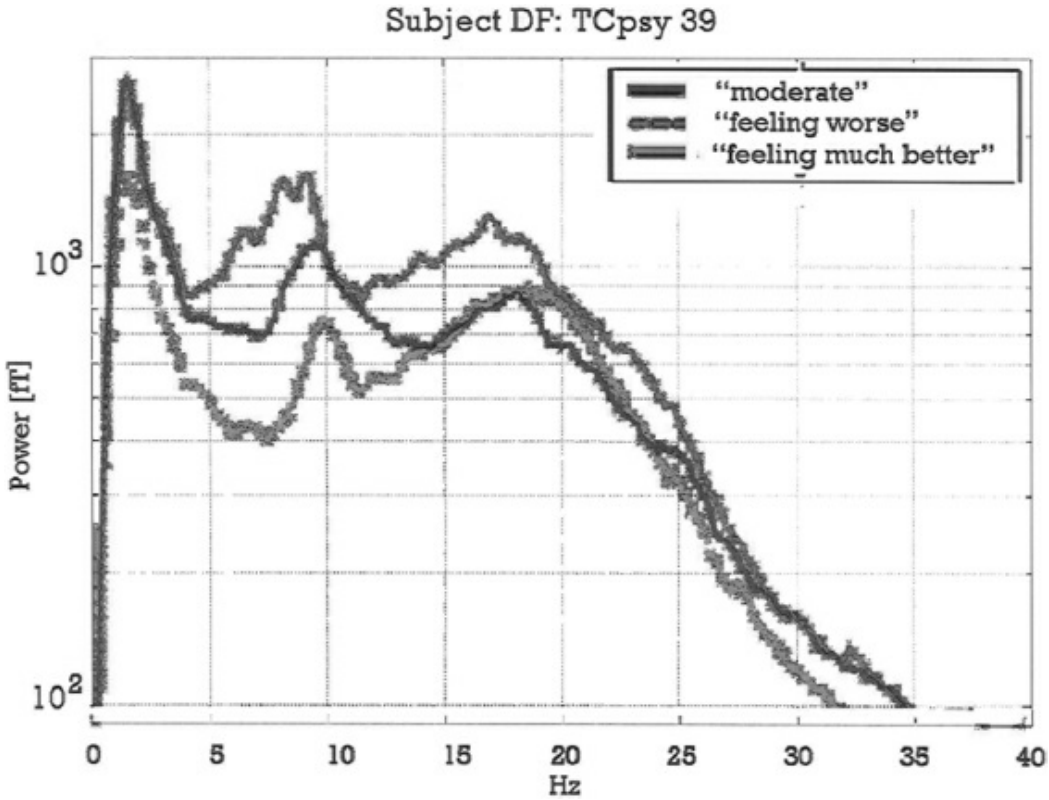

FIGURE 3.5–5. Spontaneous magnetoencephalographic recording from a depression patient during a psychotherapy session conducted in the MEG recording chamber by Prof. Robert Cancro Head of Psychiatry in 2001. The session consisted of a set of recordings in which the patient was in a moderate depressive mood, in an enhanced depressing mood, and following a psychotherapy session in the recording chamber in a reduced depression mood. Note that the low-frequency activity patterns were reduced by an order of magnitude during the psychotherapy session. (Unpublished observation.)

In addition to these instrumentation-dependent approaches, the pharmacological/molecular biological paradigm continues to be of

paramount importance. The fact that channelopathies may soon be able to be controlled, many of which are directly related to psychiatric conditions, by using specific promoter-dependent neuronal targeting, looms large on the horizon. However, rather than given technologies or specific molecular interventions, it is the ability to contextualize and combine these apparently unrelated approaches into a single arsenal to address psychiatric ailments that will ultimately help mentally compromised patients.

REFERENCES

Castaigne P. Ramollissement pédonculaire median, tegmento-thalamique avec ophtalmoplegie et hypersomnie. *Rev Neurol (Paris)*. 1962;106:357.

Crick F, Koch C. Some reflections on visual awareness. *Cold Spring Harbor Symp Quant Biol*. 1990;55:953.

Edelman GM. *Neuronal Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books; 1987.

Facon E, Steriade M, Wertheim N. Hypersomnie prolongée engendrée par des lésions bilatérales du système activateur médial; le syndrome thrombotique de la biffurcation du tronc basilaire. *Rev Neurol (Paris)*. 1958;98:117.

*Gray CM, Singer W. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc Natl Acad Sci USA*. 1989;86:1698.

*Heilman KM, Watson RT. Neglect and related disorders. In: Heilman KM, Valenstein E, eds. *Clinical Neuropsychology*. New York: OUP; 2001:243–293.

Hubel DH, Wiesel TN. Sequence regularity and geometry of orientation columns in the monkey striate cortex. *J Comp Neurol*. 1974;158:267.

Jeanmonod D. Thalamocortical dysrhythmia II. Clinical and surgical aspects. *Thalamus Relat Syst*. 2001;1:245.

Jeanmonod D. Neuropsychiatric thalamocortical dysrhythmia: surgical implications. *Thalamus Relat Syst*. 2003;2:103.

Joliot M, Ribary U, Llinás R. Neuromagnetic oscillatory activity in the vicinity of 40 Hz coexists with cognitive temporal binding in the human brain. *Proc Natl Acad Sci USA*. 1994;91:11748.

Landisman CE, Long MA, Bierelein M, Deans MR, Paul DL. Electrical synapses in the thalamic reticular nucleus. *J Neurosci*. 2002;22:1002.

Larkum ME, Zhu JJ, Sakmann B. A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature*. 1999;398:338.

*Llinás R. The intrinsic electrophysiological properties of mammalian neurons: insight into central nervous system function. *Science*. 1988;242:1654.

Llinás R. Content and context in temporal thalamocortical binding. In: Buzsaki G, ed. *Temporal Coding in the Brain*. Berlin: Springer-Verlag; 1994.

Llinás R. *I of the Vortex; from Neurons to Self*. Cambridge, MA: MIT Press; 2001.

Llinás R, Jahnsen H. Electrophysiology of mammalian thalamic neurons in vitro. *Nature*. 1982;297:406.

Llinás R, Leznik E, Urbano IF. Temporal binding via coincidence detection of specific and non-

specific thalamocortical inputs: a voltage dependent dye imaging study in mouse brain slices. *Proc Natl Acad Sci USA*. 2002;99:449.

Llinás R, Pare D. Of dreaming and wakefulness. *Neuroscience*. 1991;44:521.

Llinás R, Ribary U. Coherent 40-Hz oscillation characterizes dream state in humans. *Proc Natl Acad Sci USA*. 1993;90:2078.

Llinás R, Ribary U, Contreras D, Pedroarena C. The neuronal basis for consciousness. *Philos Trans R Soc Lond*. 1998;353:1841.

*Llinás R, Ribary U, Jeanmonod D, Kronberg E, Mitra PP. Thalamo-cortical dysrhythmia: a neurological and neuropsychiatric syndrome characterized by magnetoencephalography. *Proc Natl Acad Sci USA*. 1999;96:15222.

Llinás R, Ustinin M. Frequency-pattern functional tomography of magnetoencephalography data allows new approach to the study of human brain organization. *Front Neural Circuits*. 2014;8:1–8.

*Mountcastle VB. The columnar organization of the neocortex. *Brain*. 1997;120:701.

Pedroarena C, Llinás R. Dendritic calcium conductances generate high-frequency oscillation in thalamocortical neurons. *Proc Natl Acad Sci USA*. 1997;94:724.

Rhodes PA. Functional implications of active currents in the dendrites of pyramidal neurons. In: Ribary U, ed. Magnetic field tomography of coherent thalamocortical 40-Hz oscillations in humans. *Proc Natl Acad Sci USA*. 1991;88:11037.

Sekar K, Findlay WM, Poppel D, Llinas R. Cortical response tracking the conscious experience of threshold duration visual stimuli indicates visual perception is all or none. *Proc Natl Acad Sci USA*. 2013;110(14):5642–5647.

Spruston N. Pyramidal neurons: dendritic structure and synaptic integration. *Nat Rev*. 2008;9:206–221.

Steriade M, Curró-Dossi R, Contreras D. Electrophysiological properties of intralaminar thalamocortical cells discharging rhythms (40 Hz) spike-bursts at 1000 Hz during waking and rapid eye movement sleep. *Neuroscience*. 1993;56:1.

Steriade M, Jones EG, McCormick DA. *Thalamus*. Oxford: Elsevier Science; 1997.

Ulinski PS, Jones EG, Peters A, eds. *Cerebral Cortex*. Vol. 13. New York: Plenum Press; 1984:139.

Zonenshayn M, Mogilner AY, Rezai AR. Neurostimulation and functional brain imaging. *Neurol Res*. 2000;22:318.

▲ 3.6 Normality and Mental Health

GEORGE E. VAILLANT, M.D.

INTRODUCTION

Too often, psychiatry has been only preoccupied with mental illness. To paraphrase Mark Twain's quip about the weather, psychiatry is always talking about mental health, but nobody ever does anything about it. Science has conceptualized the building blocks of nuclear fission more readily than the building blocks of mental health. Thus, with the notable exception of the chapter by Daniel Offer and Melvin Sabshin in the early editions of this textbook, review of other recent major psychiatric textbooks reveals virtually no serious discussion of mental health. One reason for this lack of attention is that the study of positive mental health is a very new field. Only within the last three decades has mental health per se been addressed empirically instead of platonically.

There has been an implicit assumption that mental health could be defined as the antonym of mental illness. In other words, mental health was the absence of psychopathology and synonymous with "normal." Achieving mental health through the alleviation of gross pathologic signs and symptoms of illness is also the definition of mental health model strongly advocated by third-party payers. Indeed, viewing mental health as simply the absence of mental illness is at the heart of much of the debate concerning mental health policies. The great epidemiological studies of the past half-century have also focused upon who was mentally ill and not who was well. Only the Sterling County Studies by Alexander Leighton came close to defining positive mental health operationally. Mental illness, after all, is a condition that is easy to define reliably, and its limits are relatively clear. As a result, it has been argued that achieving above average mental or physical health is not the province of medicine, but of education. Such a definition ignores concepts of positive health like physical fitness and cardiac reserve.

Like physical fitness, however, positive mental health is too important to be ignored. To believe mental health is merely a GAF of 70 on Axis V of DSM-IV is to underestimate human potential. Sports medicine does more than mend skier's broken legs; sports medicine rehabilitates the skier to ski advanced trails once more. Starting early in the last century internists began studying physiology at high altitude and devise measures of positive physical health for athletes, pilots, and finally astronauts. In 1929–1930 at the University of California at Berkeley, the Institute of Human Development was founded by Howard Jones, Nancy Bayley, and Jean McFarlane. Originally, founded to study healthy child development, the

Institute was to provide a seminal influence on Erik Erikson's model of healthy adult development. A decade later, in the late 1930s, Arlie Bock, an internist trained in high-altitude physiology and interested in positive physical health began at the Harvard University Health Services, The Study of Adult Development. The Study was designed as an interdisciplinary study of both mental and physical health. Results from that study, lasting for 70 years, have informed many facets of this chapter.

Although above average mental health is more difficult to define than above average physical fitness, it is important for psychiatrists to emulate sports medicine in order to provide precise definitions and measures of positive mental health. Psychologists have already learned to quantify not only "normal" but better than normal intelligence. Thus, the antonym of mental retardation is not regarded as an IQ of 100 but as an IQ over 130. Psychiatry must follow suit.

Certainly, over the last 70 years psychiatrists have become increasingly involved in mental health consultations to agencies. Rather than merely deciding who is too sick for a job, they are called on to make decisions about who is mentally healthy enough for certain positions—like air controllers and submariners. Analogous to cardiac reserve, the measurement of resilience and of the capacity to withstand adversity are on psychiatry's psychometric wish list. In addition, interest in the empirical evaluation of psychiatric therapy outcome has also brought the issue of positive mental health into focus. Indeed, one of the weaknesses of much existing work on psychotherapy assessment has been a lack of clarity regarding outcome definitions. *For health is not the absence of negatives but the presence of positives.*

But the definition of positive mental health is not easy. Several cautionary steps are necessary. The first step in discussing mental health is to note that "average" is not healthy; it always includes mixing in with the healthy the prevalent amount of psychopathology in the population. For example, in the general population "average" weight or eyesight is unhealthy; and if all sources of biopsychosocial pathology were excluded from the population, the average IQ would be significantly above 100. For as with Garrison Keiller's description of Lake Woebegon youth, mentally healthy children "are all above average." Being at the center of a normal bell curve of distribution may or may not be healthy. In the case of red blood count, body temperature, or mood the middle of the bell curve *is* healthy. In the case of eyesight, exercise tolerance or empathy only the upper end of the bell curve is healthy; and in the case of serum cholesterol, bilirubin and narcissism only the low end of the curve is healthy.

The second cautionary step in discussing mental health is to appreciate the caveat that what is healthy sometimes depends upon geography, culture, and the historical moment. Sickle cell trait is unhealthy in New York City, but in the tropics where malaria is endemic, the sickling of red

blood cells may be lifesaving. Punctuality is a virtue in Germany, and sometimes a failing in Brazil. Indeed, some aboriginal Australian languages do not even have a word for time. General George Patton's competitive temperament was for him a psychological liability in time of peace but a virtue in two World Wars. In World War II paranoid personalities made very poor submariners but excellent airplane spotters.

A third cautionary step is to make clear whether one is discussing *trait* or *state*. Who is physically healthier: an Olympic miler disabled by simple but temporary (state) sprained ankle or a Type 1 diabetic (trait) with a temporarily normal blood sugar? In cross-cultural studies such differences become especially important. Superficially, an Indian mystic in a state of trance may resemble a catatonic schizophrenic but not over time.

In defining mental health, the fourth and most important cautionary step is to appreciate the twofold danger of "contamination by values." On the one hand, cultural anthropology teaches how fallacious any definition of mental health can be. Competitiveness and scrupulous neatness may be healthy in one culture and regarded as personality disorder in another. And, if mental health is "good," what is it good for? The self or the society? For "fitting in" or for creativity? For happiness or survival? And who should be the judge? Since every culture differs in its diet, the World Health Organization (WHO) should never be called upon to design restaurant menus. For 10 years, Adolf Hitler was a better leader of his country than Jimmy Carter; but by every definition the author knows, Carter had better mental health than Hitler.

Thus, common sense, not post modernism, must prevail. Biology trumps anthropology. Health is the activity of a living body in accordance with its specific excellences.

"What, for example, is a healthy squirrel?" asked Kass, a research professor in bioethics and neurology, "Not a picture of a squirrel, not really or fully the sleeping squirrel, not even the aggregate of his normal blood pressure, serum calcium, total body zinc, normal digestion, fertility, and the like. Rather, the healthy squirrel is a bushy-tailed fellow, who looks and acts like a squirrel; who leaps through the trees with great daring, who gathers, buries, and covers but later uncovers and recovers his acorns; who perches out on a limb cracking his nuts, sniffing the air for smells of danger, alert, cautious, with his tail beating rhythmically; who chatters and plays and courts and mates and rears his young in large improbable looking homes at the tops of trees; who fights with vigor and forages with cunning; who shows spiritedness, even anger, and more prudence than many human beings... Health is a natural standard or norm—not a moral norm, not a 'value' as opposed to a 'fact,' not an obligation but a state of being that reveals itself in activity."

It is true that cultural anthropology teaches that almost no form of behavior is considered abnormal in all cultures, but that does not mean

that the tolerated behavior is mentally healthy. WHO would be in error to ignore the universal importance to diet of vitamins and of the four basic food groups. Just because until recently Portugal does not recognize alcoholism as an illness, does not reduce the contribution of alcoholism in Lisbon to morbidity. Just because the American Constitution protected the right to own slaves as inalienable did not make slavery mentally healthy for slave or for slave owner. Just because martinis and Coca Cola outsell carrot juice does not make the former healthier. The best way to enrich the current understanding of what constitutes mental health is to study a variety of healthy populations from different perspectives, in different cultures and for a long period of time.

This chapter will contrast eight different empirical approaches to mental health. Significantly, the empirical underpinnings of all eight models have emerged only in the past 50 years. First, mental health can be conceptualized as *Above Normal* and a mental state that is objectively desirable—as in Sigmund Freud's alleged definition of mental health was the capacity to work and to love. Second, mental health can be conceptualized as *Positive Psychology*—as epitomized by the presence of multiple human strengths. Third, from the viewpoint of healthy adult development, mental health can be conceptualized as *Maturity*. Fourth, mental health can be conceptualized as *Resilience,* as the capacity for successful adaptation, "grit" and homeostasis. In such a view, mental health, analogous to a competent immune system, allows the individual to function well despite stressful or dangerous environments. Fifth, mental health can be conceptualized as *Emotional Intelligence* and wise and successful object relations. Sixth, mental health can be conceptualized as *Subjective Well-Being* (*SWB*)—a mental state that is subjectively experienced as happy, contented, and desired. Seventh, mental health can be conceptualized as ***Positive or "Spiritual" Emotions.***

MODEL A: MENTAL HEALTH AS ABOVE NORMAL

This first perspective differs from the traditional medical approach to health and illness. Until recently, being without manifest psychopathology equaled mental health. In this medical model, if one were to put all individuals on a continuum, normality would encompass the major portion of the young adults, and abnormality would be the small remainder. This definition of health correlates with the traditional role model of the doctor who attempts to free his patient from grossly observable signs of illness. In other words, in this context health refers to a reasonable, rather than an optimal, state of functioning.

Table 3.6–1.
Three Contrasting Definitions of Mental Health

Model A Mental Health/Normality	Model B Positive Psychology	Model C Maturity
Marie Jahoda as summarized by Daniel Offer and Melvin Sabshin	Christopher Peterson, Martin Seligman	William Menninger
Ability to love, work, and play Possesses empathy adequacy in interpersonal relations	**Love:** Intimacy/reciprocal attachment Kindness/generosity/nurturance Social intelligence/emotional intelligence	**The capacity for love** The capacity for a variety of mutually fulfilling and lasting relationships The need to seek for a major source of fulfillment in productive work
Efficient in problem solving Perceives reality Resistance to stress Environmental mastery	**Temperance:** Forgiveness/mercy Modesty/humility Prudence/caution Self-control/self-regulation	Free of stereotyped and nonproductive patterns of problem solving Ability to discharge hostility without harming others or oneself The capacity to adapt to change and endure frustration and loss
Invested in life Self-actualization Oriented toward future	**Wisdom and Knowledge:** Curiosity/interest Love of learning Judgment/open-mindedness Perspective Creativity/originality/ingenuity	A realistic acceptance of the destiny imposed by one's time and place in the world
Autonomy Recognition of needs In touch with their own identity and feelings	**Courage:** Valor Honesty/authenticity Industry/perseverance Zest/enthusiasm	Appropriate expectations and goals for oneself Ability to respond to the uncertainties of reality in a manner consistently free of domination by one's wishes or peers
	Justice: Citizenship/loyalty/teamwork Equity/fairness Leadership	
	Transcendence: Awe/wonder Gratitude Hope/future mindedness Spirituality/faith Playfulness/humor	**Having the capacity for hope** An altruistic concern for other human beings outside one's own group and beyond one's own time and place The capacity to suspend one's adult identity and engage in childish play at appropriate times

Yet, as already pointed out, mental health is not normal, it is above average. Indeed, some believe that true mental health is the exception not the rule. Moreover, until recently some believed mental health was imaginary. Many post-war psychiatrists had continued to agree with Sigmund Freud who had dismissed mental health as "an ideal fiction." In the late 1950s, two of the world's most distinguished psychiatrists could dismiss the term entirely. Sir Aubrey Lewis wrote in 1958, "Mental health is an invincibly obscure concept," and in 1957 Fritz Redlich asserted, "We do

not possess any general definition of normality and mental health from either a statistical or a clinical viewpoint."

Certainly, the absence of illness and the presence of health overlap but do not always coincide. In the military, the mental health for a jet pilot must exceed the draft board's 1-A. In primary school teachers and in Supreme Court justices, it is more than simply freedom from symptoms or for Freud's capacity to work and to love that are looked for. However, to avoid quibbling over what traits characterize mental health, it is helpful to adopt the analogy of a decathlon champion. What constitutes a "track star"? Is it muscle strength, or speed, or endurance, or grace or competitive grit? Does not the definition differ from nation to nation, and century to century? Not really. The salience of a given facet of track or of mental health may vary from culture to culture, but all are important. A high score on the decathlon has conveyed behavioral skill in track for millennia and in every country of the world—even as the hair splitters argue over the right words with which to describe a track athlete.

As with excellence in the decathlon, no single measure defines mental health, but all measures are highly intercorrelated. Table 3.6–1 illustrates how multifaceted and unique, each model of positive mental health can be and how very different can be the semantics. But the familial resemblance of each definition to each other is unmistakable.

In 1835 Adolphe Quetelet published the first important book on normality ever written. Rather than focus on pathology, he tried "to approach more closely to what is good and beautiful" and his goal was the statistical analysis of healthy humans. He challenged generations of future investigators with his introductory sentence, "Man is born, grows up, and dies, according to certain laws which have never been properly investigated" (p. 5).

Until World War II, however, Quetelet's challenge to mental health workers went largely unnoticed. When in 1941 John Clausen and his coworkers were commissioned to assess mental health for the draft board, they were embarking on a novel task. As a way of assessing mental health of recruits, they focused on the absence of psychosomatic symptoms. Although limited, their approach was not abandoned until the 1970s, and questions about psychosomatic symptoms still form an important part of the Hopkins SCL-90, and of scales assessing neuroticism.

After World War II influential works on normal adaptive behavior began to be published—Roy Grinker and John Spiegel's *Men under Stress,* Robert White's *Lives in Progress,* Leo Sroles' *Mental Health in the Metropolis* and Alexander and Dorothea Leightons' *Cove and Woodlot* were four of the more important. Although all four studies concentrated on the adaptation of nonpatient or normal populations, they still put their emphasis on *not pathological* rather than on *above normal mental health.*

Then, in 1958, Marie Jahoda's report to the Joint Commission on

Mental Illness and Health led to a psychiatric sea change regarding the existence of mental health. As illustrated in Table 3.6–1, Jahoda suggested that mentally healthy individuals should (1) be in touch with their own identity and their own feelings. (2) They should be oriented toward the future and over time they should remain fruitfully invested in life. (3) Their psyches should be integrated and provide them resistance to stress. (4) They should possess autonomy and recognize what suits their needs. (5) They should perceive reality without distortion and yet possess empathy. (6) They should be masters of their environment—able to work, to love, and to play and to be efficient in problem solving. Instead of emphasizing the absence of negative symptoms, Jahoda underscored a number of positive traits. Nevertheless, although the purpose of Jahoda's report was to rid the term, mental health, of "vague, elusive and ambiguous connotation," at the time she published her criteria, there was still no evidence to prove that her plausible definition was more than mere platitudes that reflected her own culture bound values.

Thus, it was with Roy Grinker's 1962 studies of "homoclites," that investigators began the empirical study of positive mental health. Grinker's homoclites were physical education majors selected for normality but studied only briefly. A second more longitudinal study was the elimination process by which out of 130 healthy jet pilots, already selected for mental health, the 7 original astronauts were selected. The process underscored both the importance and the commonsensical nature of mental health. The final seven astronauts not only enjoyed exemplary work records but also were competent at loving. All came from intact, happy, small town families; they all were married with children. Although venturesome test pilots, they all had had very few accidents during their years of flying or even earlier. They could tolerate both close interdependent association and extreme isolation. They trusted others and were uncomplaining under discomfort. They manifested great capacity for withstanding frustration; nevertheless, emotions both negative and positive were strongly experienced. Not introspective, the astronauts seldom dwelt on their inner emotions; but they were sensitive to their emotions and could describe them when asked. They were aware of the feelings of others, and they avoided interpersonal difficulties. Unlike most people their performance improved under stress (e.g., working at a simulated altitude of 65,000 feet in a poorly functioning pressure suit). Their group neuroticism (Maudsley Personality Inventory) score was the lowest of any group reported in the literature. Although each of the astronauts was very different, they all would have starred in a mental health "decathlon."

A still more influential study of mental health was the Menninger Psychotherapy Project. In order to cut through the hair splitting of projective tests and the subjectivity of pencil and paper tests Menninger psychologist, Lester Luborsky, devised a behavioral guide (Health-Sickness

Rating Scale or HSRS) to assess psychological functioning on a scale of 0 to 100. In 1976 because Luborsky's scale had been designed to evaluate candidates for psychotherapy rather than for general epidemiological studies, two of the architects of DSM-III developed a revision of the HSRS called the Global Assessment Scale (GAS). Rater reliabilities for both instruments and between the two instruments were about 0.85 to 0.95. In cross-cultural comparison investigators noted that "the usefulness of HSRS as an international thermometer of mental health is strongly supported." A modified version of the GAS was introduced in DSM-III-R as the Global Assessment of Functioning (GAF).

On Luborsky's scale a score of 95 to 100 reflected "an ideal state of complete functioning integration, of resiliency in the face of stress, of happiness and social effectiveness." On the GAF, a score of 95 to 100 equaled "no symptoms, superior functioning in a wide range of activities; life's problems never seem to get out of hand; patient is sought out by others because of his warmth and integrity." The words differ, but the melody is the same.

By 1978 *The Report to the President* by the President's Commission on Mental Health forcefully reiterated the importance of defining clearly what is meant by mental health. Nevertheless, it was not until 15 years later when evidence emerged to support the validity of Axis V, DSM-IV did psychiatry finally possess a metric for the measurement of "above average" mental health. The World Health Organization Disability Assessment Schedule replaced it in the DSM-5, but the new tool is still a measure of mental illness rather than health: all 36 questions contain the words "difficulty," "problem," "drain," or "health condition."

In order to identify mental health empirically, behavioral data are needed and a longitudinal biopsychosocial perspective must be maintained. Two very different life histories from Harvard's Study of Adult Development illustrate what is meant by "above normal" mental health. One man, "Alfred Paine," received an "average" score for mental health (GAF = 72); he never sought psychiatric care. In contrast, "Richard Luckey" received an above normal score for mental health (GAF = 95). These two (disguised) case histories help to underscore that objective and subjective mental health are not merely value ridden "ideal fictions." Both men had been selected at 18 for mental health and were prospectively observed for a lifetime. Confounding variables like gender, education, social class, ethnicity, age, and birth cohort were held constant. Neither man had ever sought psychiatric help, for nor been diagnosed with mental illness; but there is no question who was the healthier.

Alfred Paine was a master of cheerful denial. On questionnaires, he described himself as a social drinker, in quite good physical health and close to his children. It was only by interviewing him personally, talking with his wife, examining his medical record, reading the disappointed questionnaires from his children—and then, finally, by reading his obituary—that

Alfred Paine's misery could be fully appreciated

The ancestors of Alfred Paine had been successful New England clipper ship captains. All his grandparents had graduated from high school. One grandfather became a merchant banker, and the other the president of a Stock Exchange. Paine's father had graduated from Harvard, and his mother from a fashionable boarding school. When Alfred Paine was 2 weeks old, his mother died from the complications of childbirth. When he was 2 years old, his father died. Paine was bottle-fed by a variety of surrogates and raised by his grandmother. As a young boy Paine was a head-banger; in adolescence he was a "lone wolf."

In college Alfred Paine was often in love, but for Paine being in love meant having someone to care for him. His three marriages were all unhappy—in part, because of the alcohol abuse that he maintained that he did not have, and in part because he was frightened of intimacy. At 50 Paine answered "true" to the statements "Sexually most people are animals" and "I would have preferred an asexual marriage."

At 47 Paine recalled the ages from 1 to 13 as the unhappiest in his life. At age 70 he believed that the ages from 20 to 30 were the unhappiest. But there had never been a time that Paine was happy. Nevertheless, on a pencil and paper test of depression, Paine had achieved one of the best scores in the study. He had never sought psychotherapy; and none of his doctors ever called him mentally ill. His alcohol abuse never led to an arrest or to missed days of work. At 70 Paine wrote that his own physical health was "excellent;" but in fact he was seriously overweight, afflicted by hypertension and gout, and suffered from obstructive pulmonary disease—the result of life-long smoking. By 75, he had lost all his teeth; both his kidneys and his liver were failing; and he was cursed with a mild dementia, the result of a drunken fall.

Although he had made a good living over the years at a job he disliked, by 75 Paine's trust fund and his pension had been eroded through multiple divorces and self-inflicted tax troubles. His house looked as if furnished from yard sales, and little other than television now absorbed him. On his age 75 questionnaire, Paine refused to answer the part that dealt with life enjoyment. Thus, his subjective lack of joy in life had to be inferred from the fact that over the past 20 years there was no area of his life other than his religious activities in which he had expressed satisfaction. On average he saw his three children once every 2 years and he grumbled that, "They hardly let me see my grandchildren." Although Paine's third wife was protective and quite loving towards him, he was quite disrespectful and uncaring towards her. Asked how he and his wife collaborated, he replied, "We don't. We lead parallel lives." Since age 50, he had engaged in no pastimes with friends; and at 73 when asked to describe his oldest friend, he growled, "I don't have any." Often in his life Paine had found love, but he could not let it inside.

In contrast, Richard Luckey was a well-loved child who took excellent care of himself. Unlike Alfred Paine, Luckey had come from more modest beginnings. None of his four grandparents had gone beyond grade school. His father graduated from high school and went on to become a successful businessman. After college Luckey became head of two successful businesses (one of which he created)—at the same time. He loved his work.

At 70, when looked at through the eyes of his internist, Richard Luckey's objective physical health had seemed actually worse than Alfred Paine's. Luckey had high blood pressure, atrial fibrillation, a cardiac pacemaker, and pancreatitis. He was even more overweight than Paine. But if Luckey was ill, he certainly did not feel sick. Over the next 10 years Paine sickened unto death, and Luckey's health got steadily better. Thus, by 76 Lucky had completely recovered from his pancreatitis. He still wore a pacemaker; and, his blood pressure remained high, but in his doctor's words—not just his own—"Mr. Luckey continues to enjoy relatively good health… he continues to be active physically and also mentally." As Luckey, himself, expressed it, "I have done less chain sawing, but I still split wood." At age 76 Luckey spent 2 months downhill skiing in Austria.

Another crucial difference was that Luckey had never smoked, and he had used alcohol in moderation. "Almost everything we do," Luckey explained, "is family oriented. His wife amplified, "We rarely go out, but we will have groups for supper such as the church fellowship group or the basketball team for a weekend of skiing, or a vestry meeting at the house." Luckey was also commodore of his yacht club, and he had friends with whom he exercised regularly.

Richard Luckey described not only his hobbies, his religion and his income producing work

as "very satisfying," but, more important, he experienced his relationships with his wife and with his children as "very satisfying." His wife and children's questionnaires revealed a similar satisfaction with him. Luckey's daughter had described her parent's marriage as "better than my friends"; then, for good measure she had added two pluses. Luckey's wife gave her marriage a "9" out of "9," and on her questionnaire she wrote the study, "My husband is my best friend; I like looking after him. We have grown closer and fonder every year." Asked how his marriage had lasted for 40 years, Luckey replied, "I really love Chrissie and she loves me. I really respect her, highly esteem her, and she is a real person." Luckey was close to his brother with whom he fished regularly in the summer. He stayed in touch with and took great pleasure from his children and his grandchildren. It was so easy for Luckey to take in, to "metabolize," any love that he was offered.

Asked about his retirement Richard Luckey replied, "I think it is all pretty nice. I don't have a day when I don't have something to do that I want to do ... creativity is absolutely necessary for someone to be healthy. With painting," he added dreamily, "you forget everything, and that is why it is so very relaxing." In church Luckey sang both solo and in the choir. Despite an indifferent college scholastic record, at age 75 Lucky was making a serious effort to master the technology of a research institute of which he had become a trustee.

In other words, health is based on an active, joyous, energetic engagement with the world. Such a naturalistic model of health is congruent with the increasing attention to "flow," a concept recently elaborated and empirically studied by psychologist, Mihaly Csikszentimihalyi and his students. "Flow" involves the focused attention and psychic absorption which is characteristic of meditation, but unlike meditation, with "flow" the "clutch" is engaged and skilled behavior takes place. With "flow" the participant feels alive and in the world. In "flow" experience with emotions are not just contained and channeled, they are energized and aligned with consciousness of the task at hand. Action, cognition, and feeling are merged into one. Often, when manifested in intense experiences like advanced tennis, technical rock climbing or violin playing, the "flow" experience has required hours of prior practice until much of the effort involved has become second nature.

Csikszenmihalyi's concept of "flow" is distinct from Freud's libido. "Flow" occurs when a task is challenging and requires both skill and concentration, when there are clear goals and immediate involvement, when time seems to stop and one's sense of self vanishes, and when one finds themself both deeply involved—and in control. A species does not survive just by reproducing; it also has to produce and face new challenges.

MODEL B: MENTAL HEALTH AS POSITIVE PSYCHOLOGY

The fact that psychologists have approached mental health somewhat differently from psychiatry has led to the second model. Psychologists, like physiologists, look at continua (traits) rather than categories; while in medicine you either have an illness or you do not. In psychology interventions to improve adequate intelligence or social skills are common, while in medicine to meddle with adequate thyroid function, a healthy hematocrit or a normal mood is only to cause trouble. In the healthy rested individual virtually all psychopharmacological interventions will, over time,

make the brain function worse; in contrast, many psychological interventions (e.g., literacy training, stress management) will make the brain function better. Thus, the medical goal of using medication to remove pathology is different from the psychologists' goal of fostering joy, enthusiasm, curiosity, and love for others in an educative model. As a result, physicians and psychologists draw attention to quite different—if complimentary—models of mental health.

In the 20th century, medicine became increasingly concerned with concrete pathology, while psychology remained interested in education and in platonic and utopian virtue. The model of Positive Psychology conceives of mental health as "the best possible" and has provided the basis for the positive psychology movement.

The last century witnessed not only psychologist John Dewey's idealism regarding education but also the work of Carl Jung and William James (physicians by training, but psychologists in spirit) concerning the search for and discovery of deeper meanings in life. In the last 70 years, Abraham Maslow's concept of self-actualization and his emphasis on humanistic psychology have drawn attention to "full use and exploitation of talents, capacities, potentialities." But until the 21st century, such humanistic psychology has not undertaken empirical research and has ignored predictive validity and follow-up. Humanistic psychology has emphasized "happiness" and the "self" at the expense of collective well-being and has spawned myriad self-help movements that have benefited clients more through placebo effect than from controlled and replicated therapeutic interventions. As early as 1925 psychiatrist, Adolf Meyer, was already warning of the need to stop "moralizing" about utopian mental health. Mental health, he suggested should be studied through "conscientious and impartial study" and "constructive experimentation."

Recently, two popular books by a distinguished psychologist, Martin Seligman, *Learned Optimism* and *Authentic Happiness* have served notice that positive psychology is beginning to follow Adolph Meyer's rules of conscientious and impartial study and constructive experimentation. Seligman's concept of "learned optimism" and positive psychology incorporate the empirical advances in cognitive psychology that have taken place over the past three decades. Creating a positive attributional style not only serves as a cognitive behavioral treatment for depression but can lead to positive mental states.

Positive psychology wishes to learn how to build the qualities that help individuals and communities, not just to endure and survive but also to flourish. Formally introduced in the January 2000 issue of *American Psychologist,* positive psychology hopes to render the psychology of strength and well-being amenable to scientific study and intervention. In that issue Martin Seligman and Mihalyi Csikszentmihalyi wrote:

"At the individual level, it is about positive individual traits; the capacity

for love and vocation, courage, interpersonal skill, aesthetic sensibility, perseverance, forgiveness, originality, future mindedness, spirituality, high talent, and wisdom ... Psychology is not just a branch of medicine concerned with illness or health; it is much larger. It is about work, education, insight, love, growth, and play. And in this quest for what is best, positive psychology does not rely on wishful thinking, faith, self-deception, fads, or hand waving; it tries to adapt what is best in the scientific method to the unique problems that human behavior presents to those who wish to understand it in all its complexity."

Over the past 40 years cognitive therapists have demonstrated that altered cognition cannot only change behavior, it can also alter brain function. If pessimism is the dominant cognition of the depressed, optimism appears the dominant cognition of the mentally healthy. If learned helplessness leads to depression, learned optimism and self-efficacy lead to mental health. Even if the depressed often appears to view reality more realistically, such accurate reality testing avails depressed individuals nothing.

In part, the importance of optimism to positive mental health depends upon an attributional cognitive style that asserts the good things that happen to someone will last forever and are pervasive. The bad things that happen to someone are limited and unlikely to happen again. In addition, optimism includes hope, a facet of mental health as old as the Greek myth of Pandora, yet hope is a topic to which psychiatry up to now has given little formal consideration. In contrast, the explanatory style of many chronically depressed individuals is that bad events, for which one may be responsible, will last forever and will generalize, while he or she bears no responsibility for good events which not only occur only by chance but are limited and fleeting. How much better to take responsibility for the good events and conceive of them is lasting and pervasive, while the bad events are bad luck, isolated instances, and unlikely to happen again.

Longitudinal studies reveal that an optimistic attributional style improves physical health and wards off depression rather than merely "prolonging human suffering" as Nietzsche feared. In addition, the so-called illusion of optimism permits one to contemplate and plan for rather than to deny the future. As longitudinal studies have repeatedly shown, future mindedness is a critical ingredient of mental health. In addition, in major depression the hopeful illusion of placebo effect accounts for roughly 50 percent of the variance between the effect size produced by SSRIs and no treatment. Thus, experimentally and longitudinally optimism has been linked to positive mood and good morale, to perseverance and effective problem solving, to longevity and to success in a wide spectrum of activities.

A critical distinction, of course, is the degree to which optimism distorts and to which pessimism accurately perceives reality. The pessimism of

psychotic depression, like the psychotic optimism of acute mania, involves reality-defying delusion. But the comforting illusions of the mentally healthy do not interfere with learning that stoves are hot, while the pessimistic perceptions of the dysthymic, while often accurate, can still inhibit future efforts to cook. In the prevention of future mental illness, one of psychiatry's most powerful tools may be the inculcation in the young of optimism, hope, and future mindedness. But to be effective, the optimistic must pull into consciousness emotional awareness of future pain and discomfort (see *anticipation* in Model D).

Recently, advocates of positive psychology have divided positive mental health into four components: *Talents, Enablers, Strengths,* and *Outcomes*. *Talents* are inborn, genetic, and are not much affected by intervention (e.g., high IQ, or being an easy baby). *Enablers* reflect benign social conditions, interventions, and environmental good luck (e.g., a strong family, a good school system, living in a democratic meritocracy); enablers can be experimentally modified to enhance strengths. *Strengths* (see Table 3.6–1) are character traits like curiosity and openness that reflect facets of mental health that are amenable to change. The 24 *Strengths* are currently essential component of modern positive psychology and can be self-assessed via the Internet. *Outcomes* reflect dependent variables (improved GAF, social relationships, SWB) that can be used to provide evidence for the predictive validity that efforts by clinicians to enhance *Strengths* are not just wishful thinking.

Admittedly, which so-called *strengths* are most associated with mental health is open to debate. Wisdom, kindness, and the capacity to love and be loved are strengths with which few would argue. But should courage be included as a strength; and why were intelligence, perfect musical pitch, and punctuality excluded? The answer is that the 24 strengths listed in Table 3.6–1 may be subjected to a variety of tests. First, they must have been recurrent positive values across cultures and across centuries—why punctuality was excluded. Second, they must be valued in their own right and not just as a means to ends. Third, like intelligence and perfect musical pitch they are not inborn talents.

There are several pitfalls with positive psychology. First, the perspective of mental health as Utopia is one of the bogey men of national health policymakers. They are afraid, without directly expressing it, that this perspective will put a backbreaking burden on health insurance. Thus, there is considerable debate within the mental health professions about whether helping people achieve greater SWB is a process that any insurance program should reasonably be expected to cover. Helping people achieve their potential for community participation is perhaps a more laudable goal. But that, too, is highly value laden and culture bound. Over time society will have to decide who should pay for positive mental health: the individual, the educational system, third party payers, religious

organizations, or a combination of all four.

A second caution to positive psychology is the danger of culturally insensitive prescription of parochial virtues. The dangers of value judgments are enormous. "Virtues," even Aristotelian virtues, need to be distinguished from health. Keeping wounds clean is healthy, but not a virtue. Body hygiene in public places is a virtue but not necessarily healthy.

A third controversial facet of positive psychology is its emphasis on optimism. Since the late 19th century, many social scientists especially those in Europe have mistrusted optimistic cognition, especially religious optimism, as a maladaptive "American" illusion interfering with accurate perception of reality. Nietzsche, Freud, Marx, and Darwin all perceived optimism as evidence of an ingenuous cultural adolescence, not of mature mental health. Opponents of optimism believed that it was healthier to face the hard facts of life. In addition, for most of the 20th century, investigators like Pavlov and Skinner have viewed the study of cognitions of all kinds, of which optimism is one, as inferior to the study of behavior. Only in the last 20 years has psychiatry begun to see the value of positive psychology.

MODEL C: MENTAL HEALTH AS MATURITY

Unlike other organs of the body that are designed to mature in the first two decades and then remain functionally static but gradually les robust until death, the brain is different. The brain can develop new functions until the eighth decade and develop anatomically (more efficiently myelinated) until the sixth. Therefore, just as optimal brain development requires almost a lifetime so does assessment of positive mental health. A 10-year-old's lungs and kidneys are more likely to reflect optimal function than are those of a 60-year-old, but that is not true of their central nervous systems. To some extent, then, adult mental health reflects a continuing process of maturational unfolding. Statistically, physically healthy 70-year-olds are mentally healthier than they were at thirty; and prospective studies reveal that individuals are less depressed and show greater emotional modulation at age 70 than they did at age 30.

But if prospective studies of adult development reveal that the immature brain functions less well than the mature, does that mean adolescents are mentally healthier than toddlers? And are the middle-aged mentally healthier than adolescents? The answer is both yes and no, but the question illustrates why to understand mental health one must understand what is meant by maturity.

In some respects in 1950 Erik Erikson anticipated Jahoda and Grinker when he provided the first model of adult social development. He viewed each of his well-known 8 "stages" of human development as a "criterion of mental health." Subsequently, Jane Loevinger provided a model of adult ego development. Lawrence Kohlberg has provided a model of adult moral development and most recently, James Fowler has provided a model of

spiritual development. Implicit in all these models is the assumption that greater maturity reflected greater mental health. Arguably, the best definition of mental health that is available is William Menninger's definition of maturity depicted in Table 3.6–1. In his model maturity is not only the antonym of narcissism, but it is quite congruent with other models of mental health.

To confirm the hypothesis that maturity and positive mental health are almost synonymous the study of the behavior and feeling states of persons studied over a lifetime becomes necessary. Although such longitudinal studies have come to fruition only recently, all illustrate the association of maturity with increasing mental health. After 50, of course, the association between mental health and maturity is contingent upon a healthy central nervous system. The ravages of brain trauma, major depression, arteriosclerosis, Alzheimer's, alcoholism, and schizophrenia must all be avoided. But if so with good physical health, intellectual function can remain virtually unimpaired until eighty. Common folklore bemoans that millions of neurons are lost with each decade of adulthood, until the value of selective pruning is recognized. Remember 5-year-olds have twice as many neurons as adults do at 20.

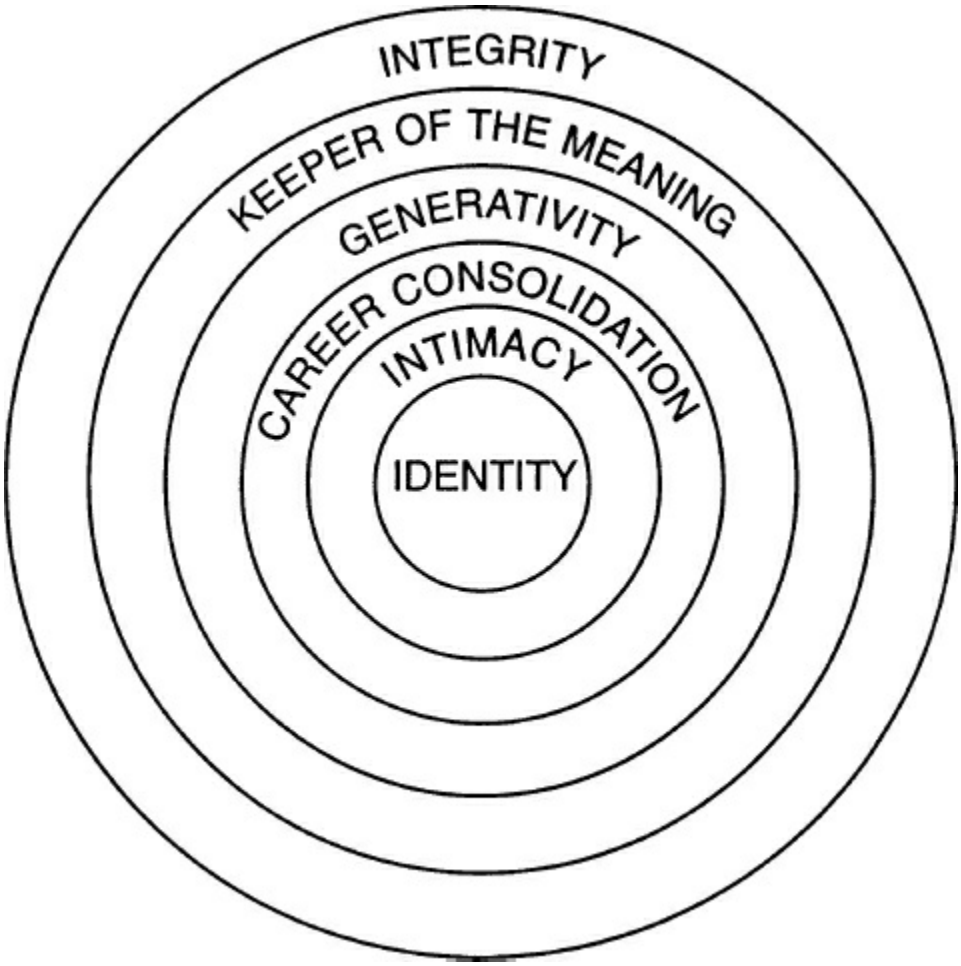

FIGURE 3.6–1. A schematic model of the expanding social radius of maturing individuals

during adulthood.

The association of mental health to maturity is mediated not only by progressive brain myelinization into the sixth decade but also by the evolution of emotional and social intelligence through experience. Erikson conceptualized that such development produced a "widening social radius." In such a view, life after age 50 was no longer to be a staircase leading downwards, as in the Pennsylvania Dutch cartoons of lifespan development, but a path leading outwards. In Erikson's model the adult social radius expanded over time through the mastery of the four tasks of "Identity vs. Identity Diffusion," "Intimacy vs. Isolation," "Generativity vs. Stagnation," and "Integrity vs. Despair." On the basis of empirical data from Harvard's Study of Adult Development, Vaillant has added two more tasks—Career Consolidation and Keeper of the Meaning or Guardian—to Erikson's four (see Fig. 3.6–1). The mastery of such tasks appears relatively independent of education, gender, social class, and probably culture.

In such a model, the social radius of each earlier adult developmental task fits inside the next. First, adolescents must achieve an *Identity* that allows them to become separate from their parents; for mental health and adult development cannot evolve through a false self. The task of *Identity* requires mastering the last task of childhood: sustained separation from social, residential, economic, and ideological dependence upon family of origin. Identity is not just a product of egocentricity, of running away from home, or of marrying to get out of a dysfunctional family. For there is a world of difference between the instrumental act of running away from home and the developmental task of knowing where one's family values end and one's own values begin. Such separation derives as much from the identification and internalization of important adolescent friends and nonfamily mentors as it does from simple biologic maturation. For example, accents become relatively fixed by age 16 and reflect those of one's adolescent peer group rather than the accents of one's parents.

Then, young adults should develop *Intimacy,* which permits them to become reciprocally, and not selfishly, involved with a partner. But living with just one other person in an interdependent, reciprocal, committed, and contented fashion for years and years may seem neither desirable nor possible to a young adult. Once achieved, however, the capacity for intimacy may seem as effortless and desirable as riding a bicycle. Sometimes the relationship is with a person of the same gender; sometimes it is completely asexual; and sometimes, as in religious orders, the interdependence is with a community. Superficially, mastery of intimacy may take very different guises in different cultures and epochs, but "mating-for-life" and "marriage-type love" are developmental tasks built into the developmental repertoires of many warm-blooded species, including the human species. Nevertheless, recent life time studies of

marriage reveal "mating" for life is not always the rule—or even desirable. Personalities often change over a lifetime.

Career Consolidation is a task that is usually mastered together with or that follows the mastery of intimacy. Mastery of this task permits adults to find a career as valuable as they once found play. On a desert island one can have a hobby but not a career, for careers involve being of value to other people. There are four crucial developmental criteria that transform a "job" or hobby into a "career": contentment, compensation, competence, and commitment. Obviously, such a career can be "wife and mother"—or in more recent times "husband and father." To the outsider the process of Career Consolidation often appears "selfish," but without such "selfishness" one becomes "selfless" and has no "self" to give away in the next stage of Generativity. Not only schizophrenics but also individuals with severe personality disorder often manifest a lifelong inability to achieve either intimacy or sustained, gratifying employment. In short, they fail to grow up.

Mastery of the fourth task, *Generativity,* involves the demonstration of a clear capacity to care for and guide the next generation. Existing research reveals that sometime between age 35 and 55, the need for achievement declines and the need for community and affiliation increases. Depending on the opportunities that the society makes available, generativity can mean serving as a consultant, guide, mentor, or coach to young adults in the larger society. Like leadership, generativity means to be in a caring relationship in which one gives up much of the control that parents retain over young children. Good mentors learn "to hold loosely" and to share responsibility. Generativity reflects the capacity to give the self—finally completed through mastery of the first three tasks of adult development—away. Its mastery is strongly correlated with successful adaptation to old age. This is because in old age there are inevitable losses and these may be overwhelming if one has not continued to grow beyond his or her immediate family.

The penultimate life task is to become a *Guardian* (in anthropological terms, Keeper of the Meaning). Like grandparenthood this task involves passing on the traditions of the past to the future. Generativity and its virtue, *care,* require taking care of one person rather than another. While keeper of the meaning and its virtues of *wisdom* and *justice* are less selective, for justice, unlike care, means not taking sides. Indeed, mastery of this fifth task is epitomized by the role of the wise judge. The focus of a *Guardian* is with conservation and preservation of the collective products of mankind—the culture in which one lives and its institutions—rather than with just the development of its children.

Just as "selfless" coaches play a different role from "selfish" athletes; just so the organizers, judges, and guardians of the Olympic Games play a very different role from the Games' more generative but still parochial

coaches. Clearly caretakers and grandparents are not mentally healthier than caregivers and parents. The distinction is only that grandparents are usually better at the tasks of Keepers of the Meaning than are 30-year-olds.

Finally, in old age it is common to feel that some life exists after death and that one is part of something greater than one's self. Thus, the last life task in Erikson's words is *Integrity,* the task of achieving some sense of peace and unity with respect both to one's own life and to the whole world. Erikson described integrity as "an experience which conveys some world order and spiritual sense. No matter how dearly paid for, it is the acceptance of one's one and only life cycle as something that had to be and that, by necessity, permitted of no substitutions." One of the Study subjects, age 88, described Integrity in more mundane terms "I need to go to the Department of Motor Vehicles to discover if the world would be safer if I no longer drove a car." How different the concerns of the elderly adult are from the adolescent seeking Identity.

Admittedly, healthy adult development does not follow rigid rules. Some individuals, often due to great stress, tackle developmental tasks out of order or all at once. On the one hand, Alexander the Great, General Lafayette, Napoleon Bonaparte, and Joan of Arc were inspirational and generative leaders in their twenties. On the other hand, William Osler, one of the most generative physicians of all time, did not confront the task of Intimacy until he married at age 40—well after he had established his career as a brilliant professor of medicine. Ludwig van Beethoven enjoyed a brilliant committed career, but never enjoyed intimacy.

Finally, it must be kept in mind that mastery of one life task is not necessarily healthier than mastery of another; for adult development is neither a foot race nor a moral imperative. Rather, these six sequential tasks above are offered as a road map to help psychiatrists make sense of where they are and where their patients might be located. One can be a happily married 20-year-old; that is usually healthy. One can be a chronically unemployed single 50-year-old; that is usually unhealthy. Acquiring a social radius that extends beyond oneself by definition allows more flexibility and SWB and thus is always healthier than self-preoccupation.

MODEL D: MENTAL HEALTH AS RESILIENCE

In 1856, Claude Bernard, a French physiologist and a founder of experimental medicine, started the current understanding of positive health when he wrote, "We shall never have a science of medicine as long as we separate the explanation of the pathological from the explanation of normal, vital phenomena." For example, coughing and pus in response to infection seemed pathologic but, in fact, were often healing. In other words, it is not stress that kills people, but healthy mastery of stress that permits people to survive.

In 1925, Adolf Meyer, a founder of modern American psychiatry, introduced the current understanding of mental health when he asserted that there were no mental diseases, there were only characteristic reaction patterns to stress. Meyer's point was that although adaptive mental "reaction patterns" like denial, phobias, and even projections can appear to reflect illness, they may in fact be "normal, vital phenomena" related to healing. Just as immune mechanisms, clotting mechanisms and callous formation heal by distorting bodily equilibrium; just so equally involuntary coping mechanisms heal by distorting mental processes.

Of course, the symptoms of organic brain damage usually reflect disease and not adaptation. Manic-depressive psychosis is due to genetic defect; the mental devastation produced by alcoholism is due to poisoning; and the negative symptoms of schizophrenia reflect brain defect and not adaptation. Nevertheless, much so-called mental illness is more like the healthy, but red and tender swelling that immobilizes a fracture so that it may heal. Much of what is called mental illness in the current diagnostic nomenclature—the symptoms of many anxiety disorders, some depressions, and most personality disorders—are the outward manifestations of homeostatic struggles to adapt to life. Admittedly, analogous to acne and autoimmune disease, such reaction patterns often reflect rigid, poorly modulated, and pathologic efforts at adaptation.

To set the stage, there are three broad classes of coping mechanisms that humans use to overcome stressful situations. First, there are the ways in which an individual elicits help from appropriate others: namely *consciously seeking social support*. Second, there are *conscious cognitive strategies* that are intentionally used to master stress. Third, there are *adaptive involuntary coping mechanisms* (often called "defense mechanisms") that distort the perception of internal and external reality in order to reduce subjective distress, anxiety, and depression.

By 1970, the term, *defense mechanisms,* like many psychoanalytic metaphors, had been largely discarded by empirical social scientists. Consistency of definition and rater reliability was lacking. But over the last 40 years, the idea of involuntary coping has entered the literature of empirical cognitive psychology under such rubrics as "hardiness," "grit," "self-deception," "emotional coping," and "illusion." In recent years experimental strategies for assessing such mechanisms, especially with videotape and the Q-sort, have also improved reliability. Several empirical studies, well reviewed by Skodol and Perry, have clarified the current understanding of healthy and unhealthy involuntary adaptive mechanisms (a.k.a. defenses). Nevertheless, no one has yet developed a method for assessing defenses, which meets conventional standards for psychometric reliability. As a result the axis of adaptive involuntary coping mechanisms has been excluded from DSM-5.

This third class of *involuntary adaptive coping mechanisms* reduces

conflict and cognitive dissonance during sudden *changes* in internal and external reality. If such changes in reality (e.g., the unanticipated crisis of 9/11) are not "distorted" and "denied," they can result in temporarily disabling anxiety and/or depression. In summary, such homeostatic mental "defenses" shield people from sudden changes in the four lodestars of conflict: affect, reality, relationships, and conscience (see Fig. 3.6–2).

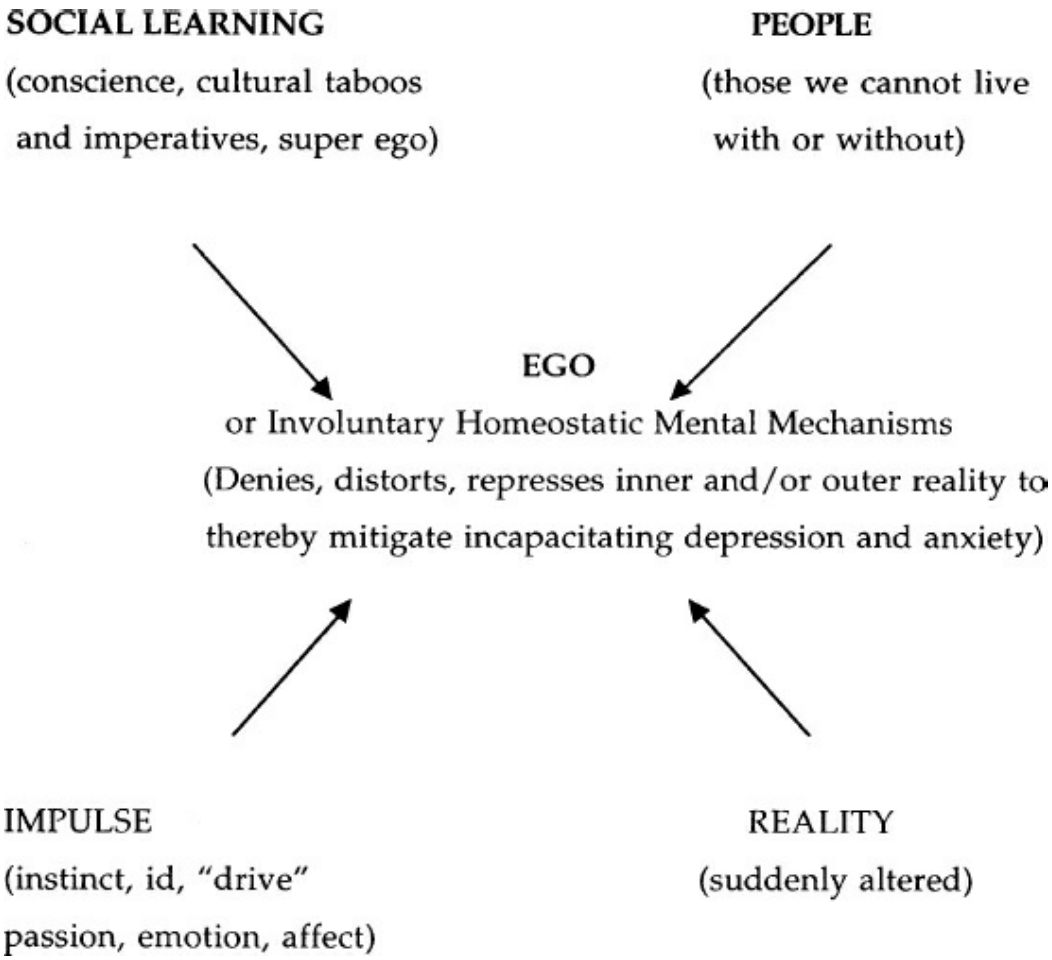

FIGURE 3.6–2. The four lode stars of human conflict.

First, such involuntary mental mechanisms can restore psychological homeostasis by ignoring or deflecting sudden increases in the lodestar of impulse, affect, and emotion. (The emotions of terror, grief, longing and dependency loom just as important as sources of conflict as do lust, fear, and rage.) Psychoanalysts call this lodestar "id," fundamentalists call it "sin," cognitive psychologists call it "hot cognition" and neuroanatomists point to the hypothalamic and limbic regions of brain.

Second, such involuntary mental mechanisms can provide a mental time out to adjust to sudden changes in *reality* and self-image, which cannot be immediately integrated. Individuals who initially responded to the television images of the sudden destruction of New York's World Trade Center as if it were a movie provide a vivid example of denial of an external

reality that was changing too fast for voluntary adaptation. Sudden good news—the instant transition from student to physician or winning the lottery—can evoke involuntary mental mechanisms as often as can an unexpected accident or diagnosis of leukemia.

Third, involuntary mental mechanisms can mitigate sudden irresolvable conflict with important *people,* living or dead. People become a lodestar of conflict when one cannot live with them and yet cannot live without them. Death is such an example; another is an unexpected proposal of marriage. And since people exist within one's self as well as without, internal representations of important people may continue to haunt someone and to cause conflict for decades after they have ceased to live with them and thus evoke involuntary mental response.

Finally, the fourth source of conflict or anxious depression is social learning or conscience. Psychoanalysts call it "super ego," anthropologists call it "taboos," behaviorists call it "conditioning," and neuroanatomists point to the associative cortex and the amygdala. This lodestar is not just the result of admonitions from one's parents that are absorbed before age five, but is formed by one's whole identification, with one's culture, and sometimes by irreversible learning resulting from overwhelming trauma.

There are many ways that one can ignore or distort one or more of these four lodestars in order to mitigate intrapsychic conflict. These coping mechanisms or "defenses" can abolish impulse (e.g., by reaction formation), or social learning (e.g., by acting out), or other people (e.g., by schizoid fantasy) or reality (e.g., by psychotic denial). They can abolish one's conscious recognition of the subject (e.g., by projection) or the object of a conflict (e.g., by turning against the self) or awareness of the idea (e.g., by repression), or of the affect associated with the idea (e.g., isolation of affect).

All classes of adaptive involuntary coping mechanisms are effective in "denying" or defusing conflict and in "repressing" or minimizing stress; but they differ greatly in the psychiatric diagnoses assigned to their users and in their consequences for long-term biopsychosocial adaptation. In the most pathological category, are found distortion and denial of external reality. These mechanisms are common in young children, in one's dreams, and in psychosis. To breach them requires altering the brain by neuroleptics or waking the dreamer.

More common to everyday life are the relatively maladaptive, "narcissistic" or "immature" defenses. They include projection, passive aggression, schizoid fantasy, hypochondriasis, dissociation, and "acting out." Defenses in these categories are common in adolescents, in immature adults, and in individuals with personality disorders. They often make others more uncomfortable than the user. Such defenses are consistently and negatively correlated with global assessment of mental health, and they profoundly distort the affective component of interpersonal relationships.

Less maladaptive mechanisms include repression, intellectualization, reaction formation (i.e., turning the other cheek), and displacement (i.e., directing affect at a more neutral object). They are often associated with anxiety disorders and with the psychopathology of everyday life. In contrast to the "immature" defenses, these intermediate defenses are manifested clinically by phobias, compulsions, obsessions, somatizations, and amnesias. Such users often seek psychological help and such "compromise formations" respond more readily to interpretation. Such defenses usually cause more conscious suffering to the user than to other people. They are common to everyone from 5 years old until death. They are neither healthy nor unhealthy.

The "mature" still distort and alter feelings, conscience, relationships, and reality; but they achieve these alterations gracefully and flexibly. These mechanisms allow the individual consciously to experience the affective component of interpersonal relationships, but in a tempered fashion. Thus, the beholder regards the involuntary mature defenses as virtues, just as the beholder may regard the prejudice of projection and the tantrums of acting out as sins.

Doing as one would be done by (altruism), keeping a stiff upper lip or deploying Grit (suppression), keeping future pain in awareness (anticipation), the ability not to take one's self too seriously (humor), and turning lemons into lemonade (sublimation) are the very stuff from which positive mental health is made. Unfortunately, like tightrope walking, without months of practice mature mechanisms cannot easily be deployed voluntarily and only then by those with innate balance.

Identification of defenses is difficult. Rarely can people identify their own defenses, and people often fail to recognize them in others or, still worse, "project" their own defenses onto them. To identify a defense an objective observer needs to triangulate past truth with present behavior and subjective report. For example, one woman wins community praise for founding a shelter for battered women (altruism), but she dismisses her behavior as "a need to rent my house." Another woman is imprisoned for breaking her toddler daughter's arm in a tantrum (acting out), but she dismisses her behavior as an accident. In fact, social agency records from 30 years before reveal that both women had been taken at age two from the care of physically abusive, alcoholic mothers. Whether one ultimately views another's coping response as healthy or psychopathic depends on the results of their involuntary efforts. Ultimately, like other facets of mental health, the reliable identification of healthy but involuntary coping mechanisms requires longitudinal study.

Healthy Involuntary Mental Mechanisms

Longitudinal studies from both Berkeley's Institute of Human Development and Harvard's Study of Adult Development have illustrated

the importance of the mature defenses to mental health.

Humor. Everyone recognizes that humor makes life easier. As Freud suggested, "Humor can be regarded as the highest of these defensive processes," for humor "scorns to withdraw the ideational content bearing the distressing affect from conscious attention, as repression does, and thus surmounts the automatism of defense." With humor, one sees all, one feels much, but one does not act. Humor permits the discharge of emotion without individual discomfort and without unpleasant effects upon others. Mature humor allows one to look directly at what is painful, whereas dissociation and slapstick distracts them to look somewhere else. Yet, like the other mature defenses, humor requires the same delicacy as building a house of cards—timing is everything.

Altruism. When used to master conflict, altruism involves getting pleasure from giving to others what that person themselves would have liked to receive. For example, using reaction formation, a former alcohol abuser works to ban the sale of alcohol in his town and annoys his social drinking friends. Using altruism, the same former alcoholic serves as an Alcoholics Anonymous sponsor to a new member—achieving a transformative process that may be lifesaving to both giver and receiver. Obviously, many acts of altruism involve free will, but others involuntarily soothe unmet needs.

Sublimation. The sign of a successful sublimation is neither careful cost accounting nor shrewd compromise, but rather psychic alchemy. By analogy, sublimation permits the oyster to transform an irritating grain of sand into a pearl. In writing his Ninth Symphony the deaf, angry and lonely Beethoven transformed his pain into triumph by putting Schiller's "Ode to Joy" to music. But like humor, sublimation does not usually occur on purpose.

Suppression. Suppression is a defense that modulates emotional conflict or internal/external stressors through stoicism. Suppression minimizes and postpones but does not ignore gratification. Empirically, this is the defense most highly associated with other facets of mental health. Used effectively, suppression is analogous to a well-trimmed sail; every restriction is precisely calculated to exploit, not hide, the winds of passion. Evidence that suppression is not simply a conscious "cognitive strategy" is provided by the fact that jails would empty if delinquents could learn to just say "No." Over the last decade Angela Duckworth has carried out cutting edge empirical research on Suppression relabeled as "Grit."

Anticipation. If suppression reflects the capacity to keep current

impulse in mind and control it, anticipation is the capacity to keep affective response to an unbearable future event in mind in manageable doses. The defense of anticipation reflects the capacity to perceive future danger affectively as well as cognitively and by this means to master conflict in small steps. Examples would be the fact that moderate amounts of anxiety before surgery promotes post-surgical adaptation and that anticipatory mourning facilitates the adaptation of parents of children with leukemia.

As with the use of altruism and suppression the use of anticipation can often be voluntary and independent of conflict. But it is in cases of "hot cognition" that these defenses can become involuntary and lifesaving.

Just as psychiatry needs to understand how a GAF of 75 might become 90; just so psychiatry needs to understand how best to facilitate the transmutation of less adaptive defenses into more adaptive defenses. One suggestion has been first to increase social supports and interpersonal safety and second to facilitate the intactness of the central nervous system (e.g., rest, nutrition, and sobriety). The newer forms of integrative psychotherapies using videotape can also catalyze such change by allowing patients to see their involuntary coping style.

MODEL E: MENTAL HEALTH AS SOCIOEMOTIONAL INTELLIGENCE

High socioemotional intelligence reflects above average mental health in the same way that a high IQ reflects above average intellectual aptitude. Such emotional intelligence lies at the heart of positive mental health. In the *Nicomachean Ethics* Aristotle defined socioemotional intelligence as follows: "Anyone can become angry—that is easy. But to be angry with the right person, to the right degree, at the right time, for the right purpose, and in the right way—that is not easy." Nevertheless, as recently as 40 years ago a textbook on intelligence dismissed the concept of such Aristotelian social intelligence as "useless."

All emotions exist to assist basic survival. While the exact number of primary emotions is arguable, seven emotions are currently distinguished by characteristic facial expressions connoting anger, fear, excitement, interest, surprise, disgust and sadness. The capacity to identify these different emotions in one's self and in others plays an important role in mental health. The benefits of being able to read feelings from nonverbal cues have been demonstrated in almost a score of countries. These benefits included being better emotionally adjusted, more popular and more responsive to others. Empathic children, without being more intelligent, do better in school and are more popular than their peers. Head Start found that early school success was achieved not by intelligence but by knowing what kind of behavior is expected, how to rein in the impulse to misbehave, being able to wait, and how to get on with other children. At the same time the child must be able to communicate their own needs and turn to teachers for help.

Ethologically emotions are critical to mammalian communication. Since such communications are not always consciously recognized, the more skilled one is in identifying his or her own emotions, the more skilled he or she will be in communicating with others and in empathically recognizing their emotions. Put differently, the more one is skilled in empathy, the more he or she will be valued by others and so the greater will be his or her social supports, self-esteem and intimate relationships.

Social and emotional intelligence can be defined by the following criteria:

- ► Accurate conscious perception and monitoring of one's own emotions.
- ► Modification of one's emotions so that their expression is appropriate. This involves capacity to self-soothe one's own anxiety and to shake off hopelessness and gloom.
- ► Accurate recognition of and response to emotions in others.
- ► Skill in negotiating close relationships with others.
- ► Capacity for focusing emotions (motivation) toward a desired goal. This involves delayed gratification and adaptively displacing and channeling impulse.

Some behavioral scientists divide emotions up into positive and negative as if negative emotions were unhealthy. This is probably an error. As with pus, fever and cough, so the negative emotions of sadness, fear and anger are also important to healthy self-preservation. Positive emotions like joy, love, hope, gratitude, and compassion are associated with subjective contentment. While negative emotions interfere with contentment, their expression can be equally healthy.

Over the last 15 years, three important empirical steps have been taken in the current understanding of the relationship of socioemotional intelligence to positive mental health. The first step is that both the functional MRI (fMRI) and ingenious neurophysiological experimentation have led to advances in the current understanding of the integration of prefrontal cortex with the limbic system, especially with the amygdala and its connections. These research advances have brought psychiatrists closer to understanding emotions as neurophysiological phenomena rather than as platonic abstractions. It has learned that the prefrontal cortex is the region of the brain responsible for working memory, and that the frontal lobes through their connections to the amygdala, hippocampus, and insula encode emotional learning in a manner quite distinct from both conventional conditioning and declarative memory.

The second step forward has been the slow but steady progress in the conceptualizing and even the measuring of "emotional intelligence." Over the last decade measures of emotional intelligence are evolving rapidly.

The third advance is the use of videotape to chart emotional interaction. Videos of sustained family interactions reveal that the most important

aspect of healthy infant development, of adolescent development, and of marital harmony is how partners or parents respond to emotion in others. To ignore, to punish, and to be intimidated or contemptuous of how another feels spells disaster. Children of emotionally attuned parents are better at handling their own emotions, and are more effective at soothing themselves when upset. Such children even manifest lower levels of stress hormones and other physiological indicators of emotional arousal.

Where the passive study of positive mental health ends and active primary prevention begins is unclear but, like the model of positive psychology, the model of socioemotional intelligence is potentially interventionist. Just as someone can have above average musical skill or physical coordination and yet can train these strengths to be even higher; just so someone can learn to enhance emotional modulation. Until late in the Cold War, empirically researched techniques for "getting to yes" in negotiations were not in any academic curricula.

There are now many exercises in handling relationships that help couples, business executives and diplomats become more skilled at conflict resolution and negotiation. In the past decade, there has also been an increasing effort to teach schoolchildren core emotional and social competencies, sometimes called "emotional literacy." The relevance of these advances in psychology to psychiatry include teaching emotion recognition and differentiation in the eating disorders, and teaching anger modulation and finding creative solutions to social predicaments for behavior disorders.

MODEL F: MENTAL HEALTH AS SUBJECTIVE WELL-BEING

In this chapter, happiness will be construed to refer to *SWB* a relatively neutral and, more important, a long-lasting state of affairs. The term "happiness," however, has multiple meanings. For example, happiness can also be construed to mean nothing more than selfish pleasure. This interpretation has given Positive Psychology the disparaging reputation of "happyology." To underscore this aspect, the author prefers to define "happiness" as selfish *drive reduction* and "joy" to refer to the equally satisfying, less autistic, *and longer lasting interpersonal connection*. The search for happiness can appear selfish, narcissistic, superficial, and banal. Happiness is often based on illusion or on dissociative states. Pleasures can come easily and be soon gone. Illusory happiness is seen in the character structure associated with bipolar and dissociative disorders, with the maladaptive denial of Pollyanna and Voltaire's Dr. Pangloss and with exaltation of "me" advocated by much pop psychology. Examples of maladaptive "happiness" can be the excitement of risk taking, from being "high" on drugs, and from "turning-on" to any unmodulated but gratifying primitive need like binge eating, tantrums, promiscuity, and revenge. Such maladaptive happiness can bring temporary bliss but has no sticking

power. It is because of such ambiguity of meaning that, throughout this section, the term *SWB* will be substituted for happiness.

National well-being has been used more appropriately than "national happiness" by political scientists describes the "health" of nations. The demographic variables assessed include women's rights, lack of corruption, percent live births, and freedom of the press.

The importance of the term *SWB* addresses the question is it better to meet some expert's definition of mental health or is it better to feel subjectively fulfilled? The answer is "both." For positive mental health does not just involve being a joy to others; one must also experience SWB. Long before humankind considered definitions of mental health, they pondered criteria for subjective happiness. For example, objective social support accomplishes little if subjectively the individual cannot feel loved. Thus, capacity for SWB becomes an important model of mental health. Put differently happiness and SWB are always congruent if they involve positive emotions. Negative emotions like Greed, Lust, Hate, and Vengeance are all about *me* and lead only to fleeting satisfactions. The eight great positive emotions: Gratitude, Love, Joy, Compassion, Forgiveness, Trust (Faith), Hope, and Awe are all about the other and can lead to lasting well-being.

SWB is never categorical. Healthy blood pressure is the objective absence of hypotension and hypertension, but SWB is less neutral. SWB is not just the absence of misery but the presence of positive contentment. Nevertheless, if SWB is an inescapable dimension of mental health, it is often regarded with ambivalence or ignored altogether. If through the centuries philosophers have sometimes regarded "happiness" as the highest good, psychologists and psychiatrists have tended to be as oblivious to SWB as they are to mental health. An electronic search of psychological abstracts since 1987 turned up 57,800 articles on anxiety, 70,856 on depression, but only 5,701 mentioned life satisfaction and only 851 mentioned joy.

On the one hand, "no man is happy who does not think himself so"; but this will be true only if the "Happiness" comes from unselfish positive emotions or "selfish" behaviors that also gratify others: e.g., self-efficacy, play, and "flow" (deep but effortless involvement). Spirituality (e.g., prayer, meditation, and belief in a power greater than the self) all rid the individual of narcissistic focus on shame, resentments, and the "poor-me's." Fostering the 24 Strengths in Table 3.6–1, model B is perhaps the most important prescription for SWB from Positive Psychology.

The mental health issues involved in SWB are complicated and clouded by historical relativism, value judgment, and illusion. Europeans have always been skeptical of American concern with happiness. Only in the last decade have investigators like Barbara Fredrickson, Martin Seligman, and David Snowdon pointed out that a primary function of positive emotional

states and optimism is that they facilitate self-care. SWB makes available personal resources that can be directed towards innovation and creativity in thought and action. Thus, SWB, like optimism, becomes an antidote to learned helplessness. Again, controlling for income, education, weight, smoking, drinking, and disease, happy people are only half as likely to die at an early age or become disabled as unhappy people.

A distinction can be made between *pleasure* and *gratification*. Pleasure is in the moment, is closely allied with happiness and involves satisfaction of impulse and of biological needs. Pleasure is highly susceptible to habituation and satiety. In contrast, gratification can be equated with what Aristotle called eudaimania and Czikszentmihalyi terms "flow." In such a distinction if pleasure involves satisfaction of the senses and emotions, gratification involves joy, the satisfaction of "being the best you can be" and of meeting aesthetic and spiritual needs. This can be manifested by a child lost in play, a mountaineer transported by rock climbing, or a father marveling at his daughter's first solo on a bicycle.

But once again the role of value judgment clouds the view of whether any given form of happiness is good or bad. Is the happiness of helping others better than the happiness of gratifying creative play? The jury is still out. In addition, all emotions exist to promote survival as well as to communicate with others. Subjective (unhappy) distress can be healthy. As ethologically minded investigators have long pointed out, subjective negative affects (e.g., fear, anger, and sadness) can be healthy reminders to seek environmental safety and not to wallow in SWB. If positive emotions facilitate faith, hope, optimism, and contentment, fear is the first protection against external threat; sadness protests against loss and summons replacement, and anger signals trespass.

Positive emotion generates exploration and promotes mastery. Unlike negative emotions, which generate zero–sum, win–lose situations, positive emotions generate win–win situations. However, tolerance of negative emotions may characterize some periods of the lifecycle more than others. The relative absence of depression and anxiety over time is healthy in grammar school children from 8 to 10 and in young submariners in their twenties. But increasing capacity to tolerate and bear subjective depression and anxiety distinguishes healthy adolescents from healthy fifth graders, and healthy (generative) submarine skippers from healthy enlisted men. There is a time to hope and a time to fear. Human development is complex.

Until recently the scientific parameters of SWB were as vague as those for objective mental health. A 1967 definition suggested that a happy person is "*young, healthy, well-educated, well paid,* extroverted, optimistic, *worry free,* religious and married with high self-esteem, a good job, morals, and *modest aspirations*." In the last 30 years, however, research has shown that such a vacuous generalization only partly correct and that the italicized adjectives are untrue or true only with qualifications.

The Nun Study by Snowdon and Danner provides perhaps the most convincing link between subjective happiness and physical health. In their twenties, 180 nuns were asked to write a two to three page autobiography. Of those who expressed the most positive emotion only 24 percent had died by 80. In contrast, by age eighty 54 percent of those who expressed the least positive emotion had died.

Only in the last three decades have investigators, especially Edward Diener, made a serious effort to attend to definitional and causal parameters of SWB and thereby address important questions. One such question is: Is SWB more a function of environmental good fortune, or is it more a function of an inborn, genetically based temperament? Put differently, does SWB reflect trait or state? If SWB reflects a safe environment and the absence of stress, it should fluctuate over time, and those individuals who are happy in one domain or time in their lives might not be happy in another.

A second question but one related to the first is what is cause and what is effect. Are happy people more likely to achieve enjoyable jobs and good marriages or does conjugal stability and career contentment lead to SWB? Or are such positive associations the result of still a third factor? For example, the absence of a genetic tendency for alcoholism, for major depression, for trait neuroticism, and even for the presence of a chronic wish to give socially desirable answers (impression management) might conflate both SWB and reports of good marriage, and career contentment.

Research over the last three decades has provided a tentative answer to the first question. Modern research has confirmed the aphorism of Duke La Roche-Foucauld, "Happiness and misery depend as much on temperament as on fortune." SWB is highly heritable, and relatively independent of demographic variables. SWB is due more to "top-down" processes—temperamental factors governing SWB—rather than due to "bottom-up factors"—for example, high income leading to SWB.

As with physiological homeostasis, evolution has prepared to make subjective adjustments to environmental conditions. Thus, one can adapt to good and bad events so that he or she does not remain in a state of either elation or despair. But they have a harder time adjusting to their genes. Studies of adopted away twins have demonstrated that half the variance in SWB is due to heritability. The SWB of monozygous twins raised apart is more similar than the SWB of heterozygous twins raised together. Among the heritable factors making a significant contribution to high SWB are: low trait neuroticism, high trait extraversion, absence of alcoholism, and absence of major depression. In contrast to tested intelligence when heritable variables are controlled, SWB is not affected by environmental factors like income, parental social class, age, and education.

For example, investigators have been startled that a significant number of AIDS victims perceive that their illness has enhanced the quality of their

subjective lives. Paraplegic victims of spinal injuries adapt so that within 2 months after injury their SWB returns to a state where positive emotion exceeds negative emotion. Similarly, after a few weeks of temporary elation lottery winners also return to baseline. Some of the homeostatic mental mechanisms underlying such adaptation were elaborated in a prior section.

If SWB were due largely to the meeting of basic needs, then there should be a relatively low correlation between SWB in work and SWB in recreational settings or between SWB in social versus SWB in solitary settings. However, in most studies the correlations between SWB in different facets of life are very high—higher than the correlation between height and weight. People who are satisfied with their lives at one point in time are more likely to be satisfied with their jobs in the future and people with high job satisfaction are more likely to report satisfaction with retirement. Admittedly, a few studies exist where correlations are low between satisfaction with different facets of life. Again, more work is needed.

Since women experience more objective clinical depression than men do, the fact that gender is not a determining factor in SWB is interesting. One explanation is that women appear to report both positive and negative affects more vividly than men. In one such study gender accounted for only 1 percent of the variance in happiness but 13 percent of the variance of the intensity of reported emotional experiences.

Consistently, relationships are more important to SWB than money. In a representative study of 800 college alumni, respondents who preferred high income, occupational success, and prestige to having very close friends and a close marriage were twice as likely to describe themselves as "fairly" or "very" unhappy. Over the last two decades the doubling of net disposable income (controlling for inflation) in the Western world (Fig. 3.6–3) did not effect SWB. Additional evidence that environmental events are much less important to SWB than might be expected comes from cross-cultural studies. Mean life satisfaction in socioeconomically challenged Brazil and China is higher than in socioeconomically blessed Japan and Germany.

In some instances environment *can* be important to SWB. Young widows remain subjectively depressed for years. Even though their poverty has been endured for centuries, respondents in very poor nations, like India and Nigeria, report lower SWB than do other more prosperous nations. The loss of a child never stops aching. Although achieving concrete goals, like money and fame do not lead to a sustained increase in SWB, social comparison, like watching one's next door neighbor become richer than you, does exert a negative effect on SWB.

Maintaining self-efficacy, agency, and autonomy make additional environmental contributions to SWB. For example, surprisingly, elders will use discretionary income to live independently even though this means living alone rather than with relatives. SWB is usually higher in

democracies than in dictatorships. Assuming responsibility for favorable or unfavorable outcomes (internalization) is another major factor leading to SWB. Placing the blame elsewhere (externalization) significantly reduces SWB. In other words, the mental mechanisms of paranoia and projection make people feel worse rather than better.

Over the past 30 years, methodological solutions have accelerated understanding of SWB just as they have accelerated understanding of objective mental health. Longitudinal studies, multivariate analysis, and causal modeling have each done much to free investigators from correlational reasoning. Nevertheless, efforts to measure SWB have been quite varied and lack a gold standard. Some investigators measure SWB as simply global life-satisfaction; other investigators assess more specific domains like work, marital satisfaction, and positive affect. However, the question "How do you feel about your life as a whole?" answered on a simple 7 point scale ranging from "delighted" to "terrible" works surprisingly well.

More refined methods of measurement of subjective states of mind have included the PANAS (The Positive and Negative Affect Scale), which assesses both positive and negative affect each with ten affect items. *The Satisfaction with Life Scale* represents the most recent evolution of a general life satisfaction scale. Most recently the widely validated SF-36 has allowed clinicians to assess the subjective cost–benefits of clinical interventions. Because short-lived environmental variables can distort SWB, consensus is emerging that naturalistic experience-sampling methods are the most valid way to assess SWB. With such sampling methods, research subjects are contacted by beeper at random times during the day for days or weeks and at each interval are asked to assess their SWB. This method provides a more stable report of SWB. Finally, in order to tease verbal self-report from actual subjective experience, physiological measures of stress (e.g., measures of galvanic skin response, salivary cortisol, and filming facial expression by concealed cameras) have also proven useful.

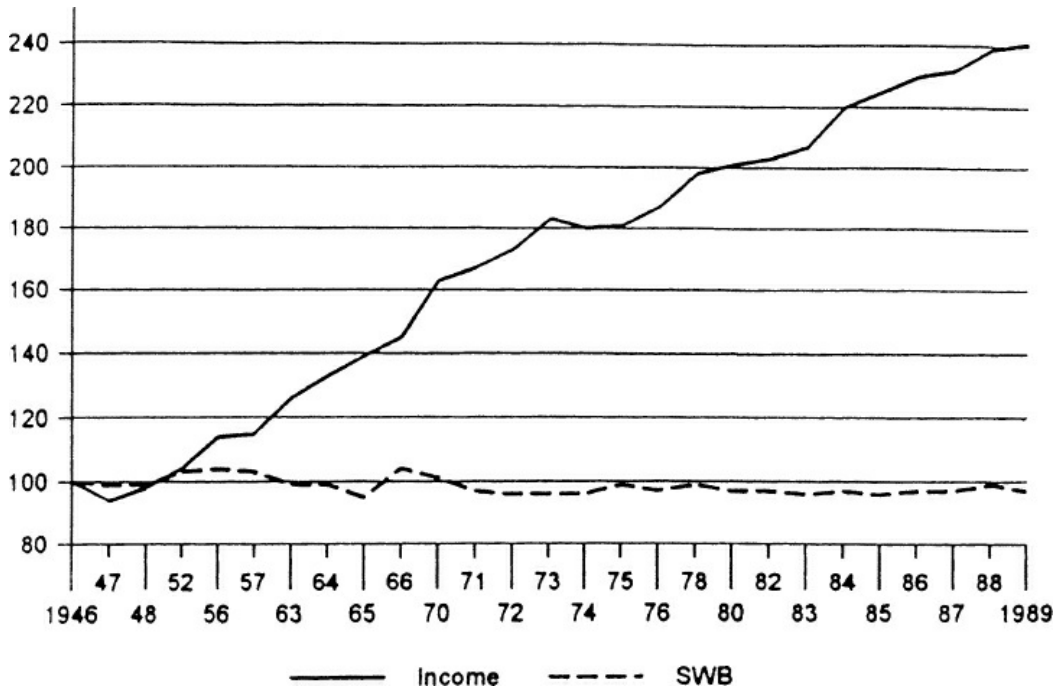

FIGURE 3.6–3. The independence of a 240 per cent increase in US individual income (adjusted for inflation) from subjective well-being over 40 years.

In the 21st century, investigators like Fredrickson, Vaillant, and Damasio have found instruments like the PANAS too cognitive and replaced "positive affects" like Joviality, Self-Assurance, Attentiveness, and Serenity with more limbic affects like Joy, Love, Compassion, Trust, and Gratitude. However, these latter more significant positive affects (described in the next model) do not lend themselves to rating scales.

MODEL G: MENTAL HEALTH AS POSITIVE OR "SPIRITUAL" EMOTIONS

In the 19th century, psychiatrists wrote about concepts like "moral insanity" and "good character," and mental health was deemed related to morality and religious observance. During the 20th century, the rise of cultural anthropology, psychoanalysis, behaviorism, molecular biology, and secularism in general all led psychiatrists to doubt there was any relationship between morality and health. Recently, with advances in neuroscience that viewpoint has begun to change.

Model G as a reflection of positive ("spiritual") emotions, however, was not included in the 8th edition of this textbook because the science was not yet in place to support its inclusion. Recent advances in neuroscience and in the biological understanding of positive emotion have necessitated psychiatry taking "positive psychology" seriously (see Peterson and Park, Chapter 44). In the last 10 years, however, positive emotions, previously relegated to popular songs, pastoral counseling, and humanistic psychology, have been rendered scientifically tangible.

Fifty years ago medical students were taught mainly about hypothalamic emotions like lust, hunger, fear, and rage. Such emotions are

common also to lunatics, alligators, and decorticate cats. Prosocial emotions like empathy, compassion, and parental love were thought to be learned behaviors and, therefore, placed in the neocortex, and in curricula of schools of education.

In the 1960s, Paul MacLean with his focus on the limbic system unique to mammals and Harry Harlow and John Bowlby, shifting attention to love as attachment rather than as Eros and lust, created the basis for neuroscience to provide substance to the platitudes of Jahoda, Abraham Maslow, and even of St. Paul with his "theological virtues" of faith hope and love.

Positive emotions arise from the inborn prosocial mammalian capacity for unselfish parental love. Thus, they are grounded in evolutionary heritage. Research over the last 20 years has placed these emotions firmly in the limbic system. Remove a mother hamster's cortex and she cannot do mazes but remains a competent mother. Damage her limbic system, however, and she can still do mazes but not parent her pups.

Model G defines both mental and spiritual health as the amalgam of the positive emotions that bind to other human beings. Love, hope, joy, forgiveness, compassion, faith, awe, and gratitude comprise the important positive and "moral" emotions included in this model. Of great importance, the eight selected positive emotions all involve human connection. None of the eight emotions listed is just about the self. These positive emotions appear to be a common denominator of all major faiths. Omitted from the list are five other positive emotions—excitement, interest, contentment (happiness), humor, and a sense of mastery, for a person can feel these latter five emotions alone on a desert island.

Negative emotions originating in the hypothalamus such as fear and anger are elaborated in the human amygdala (larger in humans than in other mammals). Of tremendous importance to individual survival, the negative emotions are all about "me." In contrast, positive emotions, apparently generated in the limbic system and unique to mammals, have the potential to free the self from the self. People feel both the emotions of vengeance and of forgiveness deeply, but the long-term results of these two emotions are very different. Negative emotions are crucial for survival in time present. The positive emotions are more expansive and help to broaden and build. In time future, they widen one's tolerance for strangers, expand one's moral compass, and enhance one's creativity. Careful experiments by neuroscientists like Jaak Panskepp and Richard Davidson and psychologists like Barbara Fredrickson and Sonia Lyubormirsky document that while negative emotions narrow attention and miss the forest for the trees, positive emotions, especially joy, make thought patterns more flexible, creative, integrative, and efficient.

The effect of positive emotion on the autonomic (visceral) nervous system has much in common with the relaxation response to meditation. In

contrast to the metabolic and cardiac arousal that the fight-or-flight response of negative emotion induces in *sympathetic* autonomic nervous system, positive emotion via *parasympathetic* nervous system reduces basal metabolism, blood pressure, heart rate, respiratory rate, and muscle tension. fMRI studies of Kundalini yoga demonstrate that meditation increases activity of the hippocampus and the right lateral amygdala, which in turn leads to parasympathetic stimulation, and the sensation of deep peacefulness.

Positive emotions have a biological basis, which means that they have evolved through natural selection. The prosocial emotions probably reflect adaptations, which permitted survival of relatively defenseless Homo sapiens and their extremely defenseless children in the Africa savannah 1 to 2 million years ago.

Evidence for Positive Emotions

However, it has taken recent developments in neuroscience and ethology to make positive emotions a fit subject for scientific study. For example, infantile autism, a not uncommon genetic disorder of emotional attachment, was not discovered until 1943 by a Johns Hopkins child psychiatrist, Leo Kanner—in his own son. Until then it was not possible for medicine to articulate a positive emotion as basic, but as cognitively subtle, as attachment. Today, the congenital lack of empathy and difficulties of attachment in childhood autism can be recognized by any competent pediatrician.

It was in the years 1945–1950 that psychoanalyst John Bowlby's ethological studies first convinced physicians that orphans needed affection as much as food. It was in the 1950s that Harry Harlow's objective findings of attachment behavior in rhesus monkeys rendered nonsexual love a tangible biologic reality for psychologists. Two decades later Paul Ekman, doubting Margaret Mead's assertion that emotions and their facial expression were culturally based, traveled as far afield as the New Guinea highlands to demonstrate that social emotions were biological and not cultural in origin.

To locate positive emotion in the mammalian limbic system has been a slow arduous process. In 1955, James Olds, an innovative neuropsychologist, observed that 35 out of 41 electrode placements within the limbic system of rats, but only 2 out of 35 placements outside of the limbic system, proved sufficiently rewarding to lead to self-stimulation. In the 1950s, also, neurobiologist Paul MacLean pointed out that the limbic structures govern mammalian capacity to remember (cognition) but also to play (joy), to cry out at separation (faith/trust) and to take care of own (love). Except for rudimentary memory, reptiles express none of these qualities.

Not until the last 20 years, however, have fMRI studies demonstrated

that when individuals subjectively experience existential states of fear, sadness, or pleasure, blood flow increases in limbic areas and decreases in many higher brain areas. Various studies have located human pleasurable experiences (tasting chocolate, winning money, admiring pretty faces, enjoying music, and orgasmic ecstasy) in limbic areas—especially in the orbitofrontal region, anterior cingulate, and insula. These diverse structures are closely integrated and organized to help to seek and to recognize all that falls under the rubric of mammalian love and human spirituality. Only in the last 20 years have neuroscientists like John Allman and Giacomo Rizzolatti identified the limbic spindle (Von Economo) cells and mirror cells that undergird prosocial human mentation.

The *anterior cingulate gyrus* links valence and memory to create attachment. Along with the hippocampus, the anterior cingulate is the brain region most responsible for making the past meaningful. In terms of mediating attachment, the anterior cingulate receives one of the richest dopaminergic innervations of any cortical area. Thus, the cingulate gyrus provides motivational salience not only for lovers but also for drug addicts. The anterior cingulate is crucial in directing whom one should approach and whom one should avoid. Maternal touch, body warmth, and odor via the limbic system and especially via the anterior cingulate, regulate a rat pup's behavior, neurochemistry, endocrine release, and circadian rhythm. Brain imaging studies reveal that the anterior cingulate gyrus is aroused neither by facial recognition of friends per se nor by sexual arousal per se. Rather, anterior cingulate fMRI images lights up when a lover gazes at a picture of a partner's face or when a new mother hears her own infant's cry.

Perhaps no area of the brain is more ambiguous in its evolutionary heritage or more crucial to mental health than *prefrontal cortex*. The prefrontal cortex is in charge of estimating rewards and punishments and plays a critical role in adapting and regulating emotional response to new situations. Thus, the prefrontal lobes are deeply involved in emotional, "moral," and "spiritual" lives.

From an evolutionary standpoint, the human frontal lobes are not different from chimpanzees in terms of number of neurons. Rather, it is the frontal lobe white matter (the connectivity between neurons through myelinated fibers), which accounts for larger frontal lobes of humans. This connectivity to the limbic system underscores its "executive" function that includes the ability to delay gratification, comprehend symbolic language, and, most important, to establish temporal sequencing. By being able to connect memory of the past to "memory of the future," the frontal lobes establish for Homo sapiens predictable cause and effect.

Surgical or traumatic ablation of the ventromedial prefrontal cortex can turn a conscientious responsible adult into a moral imbecile without any other evidence of intellectual impairment. Destroy the medial prefrontal lobes of the neocortex as was done by the tamping rod that penetrated the

skull of the legendary 19th century railroad worker, Phineas Gage; and one destroys the capacity for obedience to social mores. Phineas Gage had been a responsible foreman until an explosion drove the tamping rod that he was using to place a dynamite charge through the orbit of his left eye and out through the top of his skull. Although he survived, his medial frontal lobes were destroyed, as was his capacity for socially appropriate and empathic behavior. His intellect remained intact, but lacking empathy he was no longer mentally healthy.

The *insula* is another part of the limbic system that is only beginning to be understood. The insula is a medial cortical gyrus located between the amygdala and the frontal lobe. The brain has no sensation; humans feel emotion only in their bodies. The insula helps to bring these visceral feelings into consciousness: the pain in one's heart of grief, the warmth in one's heart of love, and the tightness in one's gut from fear all make their way into consciousness through the insula.

Both the limbic anterior cingulate and insula appear active in the positive emotions of humor, trust, and empathy. The higher apes are set apart from other mammals by a unique neural component called the spindle cell. Humans have twenty times more spindle cells than either chimps or gorillas. (Adult chimpanzees average about 7,000 spindle cells. Human newborns have four times more; and human adults have almost 200,000 spindle cells.) Monkeys and other mammals with the possible exception of whales and elephants are totally lacking in these special cells. These large, cigar-shaped spindle or "Von Economo" neurons appear central to the governance of social emotions and moral judgment. Spindle cells may have helped the great apes and humans integrate their mammalian limbic system with their expanding neocortex. Spindle cells are concentrated in the anterior cingulate cortex, the prefrontal cortex, and the insula. More recently, scientists have discovered a special group of "mirror neurons" that reside in the insula and anterior cingulate. These neurons are more highly developed in humans than in primates and appear to mediate empathy—the experience of "feeling" the emotions of another.

While the practical applications of this newest model of mental health are still many years away, these findings provide further evidence that the brain and mind are one. In several studies, the prosocial biological activity of the anterior cingulate cortex and insula was highest in individuals with the highest levels of social awareness (based on objectively scored tests). In other words, there are not only biologic individual differences for negative mental health but also for positive mental health.

FUTURE DIRECTIONS

In the 19th century, mental health was deemed related to morality and the psychiatrists wrote of "moral insanity" and "good character." But in the latter half of the 20th century, with a 19th century boost from Sigmund

Freud, religion and spirituality were expunged from psychiatry.

Religiosity is consistently and positively correlated with well-being; but there may be a chicken–egg relationship between social support and religious observance. For example, among churchgoers it is difficult to disentangle where spiritual faith ends and community support begins. Nevertheless, it makes sense that positive faith should enhance SWB. Heightened spirituality provides hope that bad events will diminish and lends wonder and joy to ordinary events. While the concept of spirituality still remains too value laden to serve as a 7th model of mental health, increasingly sophisticated neurological research on the positive emotions by investigators like Jan Panskepp, Marco Iacoboni, Barbara Fredrickson, and Andrew Newberg may find in the 21st century a role for spirituality among the models for mental health.

This chapter has so far suggested six conceptually distinct ways to assess a single construct—mental health. It would be a terrible mistake to believe any one of these six models superior to the others. From different vantage points they all measure the same thing. For example, case example, Richard Luckey not only had a GAF of 95, but another independent rater saw him as highly generative and a third rater saw him using defenses largely from level VII. He radiated SWB. Contrasting case example, Alfred Paine, had a GAF of 72, was not generative, used largely level II–V defenses, and was chronically dissatisfied with his life.

Table 3.6–2.
Cross Correlation between Four Different Models of Mental Health at Midlife (Age 47) and Their Positive Prediction of Late Life (Age 65) Mental Health and Their Independence from Parental Social Class and Childhood Environment

	A	C	D	F
Model A: Ais V	—			
Model C: Maturity	0.61	—		
Model D: Social intelligence	0.57	0.50	—	
Model F: Resilience	0.81	0.57	0.51	—
Subjective well-being age 65	0.44	0.32	0.42	0.35
Mental health at 65	0.52	0.38	0.43	0.50
Parental social class	0.08	0.05	0.12	0.06
Warm childhood environment	0.19	0.12	0.16	0.13

N = 160.

Correlations >0.25 significant at $p < 0.001$, $n = 137–143$.

Correlations >0.15 significant at $p < 0.05$.

Table 3.6–2 (from the Study of Adult Development at Harvard) provides an empirical illustration of the interrelationship of these different models in a 75-year prospective study of nondelinquent innercity men. Not

only was each of four models (measured by independent raters) significantly correlated with the other three, but also each model predicted objective global mental health and SWB assessed 20 years later. Significantly, none of the four models was well predicted by parental social class, or a warm childhood environment.

In Table 3.6–2, Axis V was assessed by Luborsky's HSRS. Maturity was determined by an assessment of presence or absence of Generativity. SWB at 55 was quantified by summing each man's report of satisfaction over the past 20 years (on a 5 point scale) in four life areas (marriage, children, job, and friends) and, then, adding his best score from one of four additional areas (hobbies, sports, community activities, and religion). Social intelligence was assessed by objective evidence of good relations with friends, workmates, wife, children, and siblings. Resilience was measured by the relative frequency with which men deployed the five **mature defenses** against the relative frequency with which they used **the five immature** (*level II–V*) defenses (in Table 3.6–3). Mental Health at 65 reflected success at work, relationships, play, and not using psychiatrists. Parental social class used the method of Hollingshead and Redlich (parental residence, education, and occupation). Warm Childhood Environment reflected good relations with mother, father, siblings, and cohesive home at age 14.

Currently, four of this chapter's models are capable of being assessed psychometrically—above average normality by the GAF (Axis V), maturity by the presence or absence of Generativity, SWB by self-report, and resilience by defense level on the optional DSM-IV axis. As already noted measures to assess psychometrically the other two models—the Strengths of Positive Psychology and socioemotional intelligence are underway.

The concept of mental health also raises the issue of therapeutic interventions to achieve it. What facets of mental health are fixed and which are susceptible to change? With Clozapine or with cognitive behavior therapy one can raise a GAF from 40 to 70, but how would one raise a GAF from 70 to 90? Chemicals can alleviate mental illness but do not improve healthy brain function. One can enhance mental health only through cognitive, behavioral and psychodynamic education.

Table 3.6–3.
Defense Levels and Individual Defense Mechanisms

I. **Level of Defensive Deregulation.** This level is characterized by failure of defensive regulation to contain the individual's reaction to stressors, leading to a pronounced break with objective reality. Examples are:
   - ► delusional projection (e.g., psychotic delusions)
   - ► psychotic denial of external reality
   - ► psychotic distortion (e.g., hallucinations)

II. **Action level.** This level is characterized by defensive functioning that deals with internal or external stressors by action or withdrawal. Examples are:

► acting out
► passive aggression
► apathetic withdrawal
► help-rejecting complaining

III. **Major Image-Distorting Level.** This level is characterized by gross distortion or misattribution of the image of self or others. Examples are:
► autistic fantasy (e.g., imaginary relationships)
► splitting of self-image or image of others (e.g., making people all good or all bad)

IV. **Disavowal Level.** This level is characterized by keeping unpleasant or unacceptable stressors, impulses, ideas, affects, or responsibility out of awareness with or without a misattribution of these to external causes. Examples are:
► denial
► projection
► rationalization

V. **Minor Image-Distorting Level.** This level is characterized by distortions in the image of the self, body, or others that may be employed to regulate self-esteem. Examples are:
► devaluation
► idealization
► omnipotence

VI. **Mental Inhibitions (Compromise Formation) Level.** Defensive functioning at this level keeps potentially threatening ideas, feelings, memories, wishes, or fears out of awareness. Examples are:
► displacement
► reaction formation
► dissociation
► repression
► intellectualization
► undoing
► isolation of affect

VII. **High Adaptive Level.** This level of defensive functioning results in optimal adaptation in the handling of stressors. These defenses usually maximize gratification and allow the conscious awareness of feelings, ideas, and their consequences. They also promote an optimum balance among conflicting motives. Examples of defenses at this level are:
► **anticipation**
► self-assertion
► affiliation
► self-observation
► **altruism**
► **sublimation**
► **humor**
► **suppression**

Adapted from American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders.* 4th ed. Text rev. Washington, DC: American Psychiatric Association; 1994:752.

One can change some facets of brain function better than others. By analogy, the most intensive educational intervention in individuals who are not severely deprived will raise their IQ only about 7 points, but sustained therapeutic intervention can change individuals utterly illiterate in Italian into fluent Italian conversationalists. Admittedly, a correct accent to go with the words is harder to teach.

In concluding this chapter it seems important to review some of the safeguards for a study of positive mental health. First, mental health must

be broadly defined in terms that are culturally sensitive and inclusive. Second, the criteria for mental health must be empirically and longitudinally validated. Third, validation means paying special attention to cross-cultural studies. In somatic medicine, criteria have been developed so that people of widely varying backgrounds and beliefs can agree upon what constitutes rational therapy for disease. The same criteria for mental health need to be developed. Fourth, although mental health is one of humanity's important values, it should not be regarded as an ultimate good in itself. One must proceed in efforts toward trying to improve mental health while maintaining due respect for individual autonomy.

There is no logically perfect method of analysis by which one can distinguish private bias and culturally determined value judgment from a culturally universal and scientifically valuable definition of mental health. Predictive validity remains the best guide. A parable offered by the protagonist in Gotthold Lessing's great 18th century play, *Nathan the Wise,* will help to illustrate why mental health can only be identified by longitudinal study. In Lessing's play an angry sultan had asked Nathan, a Jew, on pain of death, to identify the one true religion—Christianity, Islam, or Judaism. Nathan gently pointed out the need to maintain a longitudinal perspective.

> "A man lived in the East,
> Who owned a ring of marvelous worth,
> Given to him by a hand beloved.
> The stone was opal, and possessed the secret power
> To make the owner loved of God and man,
> If he but wore it in this faith and confidence..."

The ring was sought as an inheritance by each of Nathan's three sons. Since he loved him all equally, the doting father gave each son an identical ring. After their father's death, the three sons hurried off to a judge and demanded that the judge identify the lucky owner of one true ring. The judge exclaimed,

> "But stop! I've just been told that the right ring,
> Contains the wondrous gift to make its wearer loved,
> Agreeable alike to God and Man.
> That must decide, for the false rings will not have this power...
> Let each one strive to gain the prize of proving by results
> The virtue of his ring and aid its power
> With gentleness and heartiest friendliness...
> The virtue of the ring will then
> Have proved itself among your children's children."

In the future, it is incumbent upon psychiatry to understand the

implications of each of these six models of mental health. In conducting research on positive mental health, it is important for psychiatrists to recognize which model they use. Equally important, in the area of national health policy, one needs a clear understanding of not only what constitutes positive mental health, but also of who is responsible for intervention and of who is responsible for paying for it.

Finally, primary prevention is clearly superior to treating disease once it has occurred. Thus, we need to study individuals with positive mental health the way agronomists study wheat that is resistant to drought and blight. Psychiatry also needs to be able to measure and record mental health. Although room exists for improvement, Axis V, the Global Assessment of Functioning, provides the same reliability and has much greater predictive validity than the presence or absence of most Axis I and II designations. No psychiatric chart should be without Axis V. The capacities to work and to love over time are extremely important indices of mental health. They are far more important than the cross-sectional presence or absence of anxiety, depression, or illegal drug use. But capacity to work and to love must be assessed longitudinally. "How many years since age 21 have you spent employed?" is more useful than, "What is your present job?" Again, "Tell me about your longest intimate relationship" is more useful than "Are you married?"

Finally, assessment of maturational development provides with the best prediction of future clinical course. Thus, the mental status and diagnostic formulation should both reflect an assessment of social maturation and coping style. If the person is 35, have they mastered Erikson's task of intimacy? If they are 40, have they achieved competence in, commitment to, contentment with and compensation from their career? If they are over 50, have they mastered Generativity and learned to care less about themselves and more about their children. Again, when the going gets tough, do they eschew less mature mental mechanisms like projection, passive aggression and dissociation (being in denial), and do they employ mature involuntary coping mechanisms like grit, humor, altruism, and sublimation.

REFERENCES

American Psychiatric Association. *Global Assessment at Functioning (GAF) Scale, in Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington, DC: American Psychiatric Association; 1994:32.

American Psychiatric Association. *Defense Levels and Individual Defense Mechanisms, in Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington, DC: American Psychiatric Association; 1994:752.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

Csikszentmihalyi M. *Flow: The Psychology of Optimal Experience*. New York: Harper Row;

1990.
*Diener E, Suh EM, Lucas RE, Smith HL. Subjective well being: three decades of progress. *Psychol Bull*. 1999;125:276–302.
Duckworth AL, Peterson C, Mathews MD, Kelly DR. Grit: persistence and passion for long term goals. *J Pers Soc Psychol*. 2007;92:1087–1101.
Erikson E. *Childhood and Society*. New York: Norton; 1950.
Frederickson BL. What good are positive emotions? *Rev Gen Psychol*. 1998;2:300–319.
*Goleman, D. *Emotional Intelligence*. New York: Bantam Books; 1995.
Grinker RR Sr, Grinker RR Jr, Timberlake J. Mentally healthy young males (homoclites). *Arch Gen Psych*. 1962;6:405–453.
Iacobini M. *Mirroring People*. New York: Farrar, Straus and Giroux; 2008.
Jahoda M. *Current Concepts of Positive Mental Health*. New York: Basic Books; 1958.
Korchin SJ, Ruff GF. Personality characteristics of the Mercury astronauts. In: The Massachusetts Institute of Technology, ed. *The Threat of Impending Disaster*. Boston, MA: MIT Press; 1964.
Ledoux J. *The Emotional Brain*. New York: Simon and Schuster; 1996.
Loevinger J. *Ego Development*. San Francisco, CA: Jossey-Bass; 1976.
Luborsky L. Clinicians' judgment of mental health: a proposed scale. *Arch Gen Psych*. 1962;7:407–417.
Mayer JD, Caruso D, Salovey P. Emotional intelligence meets traditional standards for an intelligence. *Intelligence*. 1999;27:267–298.
Menninger WC. *A Psychiatrist for a Troubled World: Selected Papers of William C. Menninger, M.D*. New York: The Viking Press; 1967:788–804.
Offer D, Sabshin M. Chapter 66: Normality. In: Kaplan HI, Freedman AM, Sadock BJ, eds. *Textbook of Psychiatry*. 3rd ed. Baltimore, MD: Williams and Wilkins; 1980:608–613.
Panskepp J. *Affective Neuroscience*. New York: Oxford University Press; 1998.
*Peterson C. *A Primer in Positive Psychology*. Oxford: Oxford University Press; 2006.
Peterson C, Seligman ME. *The VIA Classification of Strengths*. Cincinnati, Ohio: Values in Action Institute; 2003.
Salovey P, Sluyter DJ, eds. *Emotional Development and Emotional Intelligence*. New York: Basic Books; 1997.
Seligman ME. *Authentic Happiness*. New York: Free Press; 2002.
Seligman ME, Csikszentmihalyi M. Positive psychology. *Am Psychol*. 2000;55:5–14.
Tellegen A, Lykken DT, Bouchard TJ, Wilcox KJ, Segal NL, Rich, S. Personality similarity in twins reared apart and together. *J Pers Soc Psychol*. 1988;54:1031–1039.
Vaillant GE. *Ego Mechanisms of Defense: A Guide for Clinicians and Researchers*. Washington, DC: Am Psych Press; 1992.
Vaillant GE. *Aging Well*. New York: Little Brown; 2002.
*Vaillant GE. *Spiritual Evolution: A Scientific Defense of Faith*. New York: Doubleday Broadway; 2008.
*Vaillant GE. *Triumphs of Experience*. Cambridge MA: Harvard; 2012.

4 Contributions of the Social Sciences

▲ 4.1 Evolutionary Foundations for Psychiatric Research and Practice

RANDOLPH M. NESSE, M.D.

FIVE ADVANCES

Major advances in the evolutionary explanation of behavior and emotions from the second half of the 20th century have yet to be fully applied in psychiatry. This chapter summarizes those advances and how understanding them can make clinicians and researchers more effective. Some have direct applications, such as understanding the human mating system, how negative emotions are useful, and why there are no common alleles with large effects on highly heritable diseases such as schizophrenia. However, the big payoffs in the long run will come from the integrated framework evolution provided for understanding normal behavior and psychopathology. Evolutionary approaches are not an alternative to understanding mechanisms, they are the crucial complement that explains why mechanisms are the way they are, and why they are vulnerable to failure. Evolutionary biology provides a framework that can transform the bio-psycho-social model from a catch phrase into a solid working model. It provides, for psychiatry, what physiology provides for the rest of medicine —an understanding of normal functioning as the foundation for understanding pathology.

Five major advances in evolutionary biology are especially important for psychiatry. Each is described briefly below.

Proximate and Evolutionary Explanations Are Both Essential

Until the middle of the 20th century, the study of animal behavior was mostly descriptive. Meticulous observations by ethologists described behavior in detail, but theory was absent. Everything changed in the mid-1960s, with recognition that all biological traits need evolutionary as well as proximate explanations. This insight came first from students of animal behavior, especially Nico Tinbergen. So-called “proximate questions” are

about mechanisms; answers to such questions describe how mechanisms work, and how they develop. Other "evolutionary questions" are about how traits got to be the way they are; answers to such questions describe how and why the trait has changed over evolutionary time. For instance, a proximate explanation of anxiety describes the relevant brain and psychological mechanisms, how they develop, and how they are influenced by events and environmental influences. However, even describing every detail about the proximate mechanism does not explain why the capacity for anxiety exists at all. That requires an evolutionary explanation of how the capacity for anxiety increased fitness for our ancestors.

In a famous article published in 1963, Tinbergen outlined four questions that all must be answered to have a complete understanding of any trait. Two are evolutionary questions. The first is about phylogeny, the history of a trait. The other is about the adaptive significance of the trait and what forces of selection shaped it. There are also two proximate questions. One is about the mechanism and how it works at all levels, from molecules to social psychology. The other is about ontogeny, how the mechanism develops in an individual from a zygote to an adult (Table 4.1–1). Almost all research in psychiatry has been to describe mechanisms, their development, and their malfunctions. Only now is attention turning to also address evolutionary questions such as why emotions exist.

Evolutionary Explanations for Individual Behavior

The second core advance, closely related, was the application of evolutionary thinking to behavior. Initial approaches provided evolutionary explanations for specific behaviors, such as rams butting horns in the mating season, and birds removing broken shells from nests to avoid attention from predators. It was soon recognized, however, that it was not behaviors themselves that natural selection shaped, but behavior regulation mechanisms. That is, the brain.

Table 4.1–1.
Tinbergen's Four Questions

Tinbergen's Four Questions Adapted from Nesse, 2013		**Two Objects of Explanation**	
		Developmental/Historical A sequence that results in the trait	Single Form The trait at one slice in time
Two kinds of explanation	Proximate Describes the mechanism and its ontogeny	Ontogeny **How does the trait develop in individuals?**	Mechanism **What is the structure of the trait; how does it work?**
	Evolutionary Describes adaptive significance and	Phylogeny **What is the phylogeny of the trait?**	Adaptive significance **How have variations in the trait influenced fitness?**

A general core principle soon became clear: **Brains are shaped by natural selection to give rise to behavior that maximizes Darwinian fitness.** Instead of implying rigid behavior patterns, this principle recognizes the enormous fitness benefits of flexible behaviors that allow organisms to adapt to rapidly changing situations. Learning is not an alternative to an evolutionary explanation, it is a capacity shaped by natural selection. Textbooks of animal behavior are now all grounded in evolutionary biology.

Evolutionary Explanations for Social Behavior

The evolutionary analysis of social behavior was central to transforming animal behavior research. The old idea that organisms are shaped to behave in ways that benefit their groups and species was replaced by recognition that selection generally shapes traits that maximize individual fitness, even if that harms the group. Many have heard about lemmings jumping into the sea for the good of the group when food is too scarce for all to survive. This turns out to be a myth, one promulgated by a Disney film that seems to show lemmings flinging themselves off a cliff, when in fact they were swept off by workers with brooms. An allele that induces individuals to sacrifice themselves for the good of the group will be quickly selected out except in very special circumstances. Behaviors costly to individuals that benefit groups, such as parents sacrificing for their offspring, wolves sparing those who signal defeat, or crowded populations reducing reproduction, have other explanations.

Kin Selection

Recognition that costly individual sacrifices cannot routinely be explained by benefits to others was followed closely by William Hamilton's discovery of the main exception, kin selection. Close relatives share genes that are identical by descent; 50% of a child's genes are identical to each of its parents. Alleles that decrease an individual's number of offspring can be selected for if they sufficiently benefit identical alleles in kin. Hamilton's famous formula, $C < B \times r$, is a crucial foundation for understanding family relationships. An action that costs an amount C can nonetheless give inclusive fitness benefits if the cost to self is less than the benefit to a relative, B, times the percentage of genes identical by descent. The term "inclusive fitness" incorporates actions that increase the reproduction of an individual's relatives, as well as the individual. Kin selection offers a profound evolutionary explanation for basic human behaviors, such as parents making sacrifices that benefit their children.

This discovery led to thinking in terms of "selfish genes." The metaphor

has been useful to combat naïve group-selection thinking, but it has been somewhat too potent for its own good. If genes are selfish, this seems to imply that organisms must be selfish. However, capacities for generosity and moral behavior exist because they increase inclusive fitness. Selfish genes make individuals who can be generous … selectively.

Evolutionary Medicine

Natural selection shapes traits that work well, so it would seem unable to explain diseases. That's correct—diseases are not shaped by natural selection. However, traits that leave individuals vulnerable to disease, such as the appendix, the narrow birth canal, and the capacity for low mood, have been shaped by natural selection. Evolutionary medicine tries to understand, for each disease, why selection has left bodies vulnerable. Six main kinds of explanation help to organize the search. Notice that these are not explanations for why some individuals become ill while others do not. Instead, they are about why all humans share vulnerability to a disorder.

Six evolutionary explanations for vulnerability:

1. Mismatch with modern environments
2. Constraints—things selection cannot do
3. Coevolution with fast-evolving pathogens
4. Trade-offs
5. Reproduction at a cost to health
6. Defenses that are useful but costly and aversive

Mismatch of Environment to Body. The bulk of chronic disease results from the mismatch between the human body and the very pleasant environments humans have created for themselves. Atherosclerosis, breast cancer, and autoimmune diseases are vastly more common in technological societies. So are obesity, eating disorders, alcoholism, and drug abuse. Human behavior regulation mechanisms were never shaped to cope with the ready availability of hamburgers, candy bars, whiskey, and heroin. Our distant human ancestors also never had to cope with life in bureaucracies, 9-to-5 jobs, alarm clocks, and expectations of lifetime monogamous mating.

Many alleles that predispose to major disorders are not defects, they are merely genetic "quirks" that cause harm only in modern environments. For instance, nearsightedness is highly heritable, but nonetheless rare in hunter-gatherer populations not exposed to reading early in life. Other heritable tendencies may have vastly different significance in modern environments. For instance, attention deficit disorder may not have posed serious problems for hunter-gatherers (or it may have been uncommon). Some who advocate an evolutionary approach to mental disorders have

attributed nearly all problems to a mismatch with modern environments. While such factors are significant, the five other factors are also important.

Pathogens Evolve Faster than Hosts. Infectious disease will always be present because pathogens evolve much faster than humans can. Also, every time selection shapes a host defense, pathogens evolve some way around it. The resulting arms race leaves us with costly defense mechanisms prone to cause disease.

For instance, streptococci have evolved immune markers similar to human tissues, making it risky for the immune system to attack them. Rheumatic fever, Sydenham's chorea, and some cases of obsessive-compulsive disorder are examples.

Constraints. There are many things that natural selection cannot do. It can never completely prevent mutations, although the extraordinary fidelity of DNA replication illustrates the power of selection. Evolution also cannot redesign a trait from scratch; it proceeds only by a process of small "tinkering" changes. So, while birth via the pelvis is painful and dangerous, no alternative route can evolve. Similarly, vertebrates are stuck with the blind spot where nerves and vessels enter an opening in the back of the eyeball, while cephalopods have nerves that traverse their eyeballs without causing a blind spot.

Trade-offs. No trait can be perfect. The vast majority of changes that could make a trait superior in one respect would make it inferior in other respects, or would compromise the effectiveness of other traits. Finches with larger beaks can crack tougher seeds, but they have less access to very fine seeds. Individuals with more tendency to anxiety are able to avoid dangers, but at the cost of the inability to take advantage of many opportunities.

Reproduction at the Expense of Health. It is disturbing to recognize that natural selection does not shape organisms to maximize health, happiness, or longevity. It instead shapes organisms to maximize inclusive fitness. Usually health and long life are good for fitness, however, an allele that harms health will nonetheless spread if it increases reproduction.

The best example is the short lives of men compared to women. Male mortality rates in early adulthood are three times those for women because tendencies that lead to success in mating competition are selected for in males even if they harm health. Tendencies to tissue repair and sensible caution are relatively favored in females.

Much research has sought to understand the proximate mechanisms that account for women having more anxiety disorders than men. An evolutionary framework suggests that it is men who have the problem.

Women have closer to the optimal amount of anxiety for their own health and well-being.

Some have speculated that tendencies to mania may have been selected for because they result in increased matings and increased reproductive success. If that were really the case, then the tendency would have spread and become universal. This provokes an interesting possible explanation for the high prevalence of mild bipolar tendencies. However, the rarity of manic depression suggests that the increased sexuality of individuals with tendencies to mania has not paid off with net increases in numbers of grandchildren.

Defenses. Most problems people bring to their physicians are not direct manifestations of disease; they are useful defensive responses, such as pain, fever, cough, nausea, and vomiting. These defenses, shaped by natural selection, are usually expressed only in the presence of some disease, but they are not diseases themselves. This knowledge is crucial for making good clinical decisions about when to block such defenses.

The smoke detector principle is also crucial for making such decisions well. Natural selection shapes mechanisms to regulate defenses so they are expressed in situations where their benefits are greater than their costs. However, information is often insufficient to be certain about the presence of a threat, so even an optimal system will have some false alarms, and some failures to respond when a response would be useful. This is particularly germane for understanding panic attacks. A fight–flight response will be potentially life-saving if the noise from behind a rock is made by a lion, but it will be wasted if the noise is coming from a monkey. Whether an individual should flee in a panic depends on how likely it is that the noise was from a lion. If the cost of flight is, for instance, 100 kcal, and the cost of not fleeing if the lion is actually present is 100,000 kcal, then the optimal system will set off the fight–flight response whenever the noise is loud enough to indicate a probability of a predator being present of greater than one in 1,000. This means that 999 responses of 1,000 will be false alarms, each useless, but completely normal. False alarms in the system are expected and normal. This has profound implications for understanding panic disorder, and explaining it to patients.

CORE PRINCIPLES OF EVOLUTIONARY BIOLOGY

Evolution has two major components, both discovered by Darwin. The first is the unity of all life; all individuals in all species are descendants of ancestors that go back in an unbroken chain to the origins of life on earth. This is phylogeny, now traceable via genomics. It is relevant for tracing human origins, and for identifying the history of alleles related to mental disorders, but the focus here will be on Darwin's other discovery. Darwin's second discovery is natural selection and how it explains adaptations. It is

the only viable explanation for why so many aspects of organisms are well-suited to coping with life's challenges. The rabbit's astute hearing, the woodpecker's barbed tongue, and the eagle's sharp eyes are products of natural selection. So are human empathy, emotions, cognition, and capacity for language and love. Or, more exactly, the brain, and its mechanisms for regulating thinking, emotions, and behavior.

Natural Selection

Natural selection is best understood in the context of selection in general. If members of a group differ in ways that influence their likelihood of being in the group in the future, the group will change over time. If you put all coins from your pocket into a jar each evening and select only the silver-colored coins in the morning, the jar will gradually take on a copper hue as the concentration of pennies increases. Television producers try out extraordinarily varied offerings. Those that get viewers and advertising dollars persist and are imitated, others are eliminated. Selection explains how groups change when variations among individuals influence their contributions to future groups.

Natural selection is the subtype of generic selection in which the variations are products of genetic variation that influence reproductive success. The principle is simple. If individuals in a population differ genetically in ways that influence how many surviving offspring they leave, the average characteristics of the population *will* change over the generations. This is not a theory, it is a theorem that is necessarily true. In the classic example, if previously light trees become covered with soot and moth color varies, light-colored moths are more likely to be eaten by birds, so the average moth will become darker over succeeding generations. VISTA is a useful mnemonic for the process of natural selection: Variation—Inheritance—Selection—Time—Adaptation.

Dogs offer a familiar example. From their common ancestors in wolves, artificial selection has produced dramatically diverse dog breeds in just a few thousand years. The process of natural selection is the same as breeding, except that differences in offspring numbers result not from decisions made by breeders, but from variations in ability to cope with life challenges that influence reproductive success. The products of natural selection look as if they were designed, and the sequence of species makes it seem as if there is something planned, or at least preordained, about the progression. However, there is no plan or endpoint, and traits that appear to be designed were shaped by mindless natural selection.

Standard examples usually describe how natural selection changes a trait, however natural selection mostly keeps traits the same. If individuals 7-ft tall had more children than others, the mean height would shift to that optimum. If individuals more aggressive or outgoing than others consistently had more children, the population mean would shift. However,

at least over the past few tens of thousands of years, individuals at the extremes of height, aggressiveness, and extroversion almost certainly had fewer offspring than those closer to the mean. The persisting wide variation in these traits suggests that fitness is similar across a wide range.

Whether a trait or a gene is advantageous or deleterious depends on the environment. In equatorial climates, dark skin protects against sun damage, in cloudy environments it can result in rickets from insufficient sunlight-induced vitamin D synthesis. A tendency to prefer fatty foods can be useful where fat is scarce, but in modern environments it can lead to disease.

Evolution

Evolution involves much more than natural selection. The three other main forces are mutation, migration, and genetic drift. **Mutations** happen despite being minimized to an astounding degree by natural selection; increasing evidence points to their importance for explaining schizophrenia and bipolar disorder. **Migration** may strongly influence the prevalence of an allele in a population. Most northern Europeans can digest milk because migrants brought the relevant gene from rapidly expanding dairying populations further south. Huntington chorea is especially common near Lake Maracaibo in Venezuela because Antonio Justo Doria, a Spanish sailor, brought the mutation there in the late 1800s. **Genetic drift** refers to changes in gene frequency, more correctly allele frequency, that result from random factors. An allele that increases fitness may be lost from an entire population if the individuals carrying it all die in an accident. An allele that is mildly harmful can nonetheless become more common if those who carry it have more children than others just by chance. Alleles whose effects on fitness are close to neutral can drift to fixation or be eliminated, especially in small populations. There are good reasons why population genetics relies on the so-called “null model” in which frequency changes result only from drift. This is not, however, an alternative to natural selection or a reason to think that most traits are not well-suited to their functions.

What Selection Shapes

The usual textbook examples of selection are fixed physical traits such as bird beaks and moth wing color. Far more important for psychiatry, however, is selection for systems that regulate the body and behavior. Most regulation systems keep things stable. Selection shapes mechanisms to maintain homeostasis at every level. Regulation of glucose, oxygen tension, acid–base balance, and osmolality are the usual examples.

The association of natural selection with fixed traits and systems that maintain homeostasis fosters the misperception that an evolutionary

explanation somehow implies rigidity, especially rigidity of behavior. Nothing could be further from the truth. Natural selection has shaped hundreds of mechanisms that adjust organisms to changing situations. The traits vary in scale from the molecular to large-scale behavior, and across time from instantaneous changes to those that play out over a lifetime. The term *plasticity* often refers to longer-term changes, especially those that shape an individual in relatively irreversible ways as function of early experience, such as the barrel-shaped chest of humans living at high altitudes. In the medium term, high altitude increases hemoglobin concentration, and exposure to sunlight induces skin tanning. In the short term, cold induces shivering, infection causes fever, and threat arouses fear. Some responses, such as blinking, are nearly instantaneous. All are facultative adaptations, mechanisms shaped by selection that adjust individuals to changing aspects of environments.

The brain is obviously responsible for many such responses. It monitors internal and external cues as input to mechanisms that adjust physiology, expression, behavior, cognition, and more. These responses vary across individuals, and some of the variation results from genetic variation. To make a long story short, the brain was, like every other organ, shaped by natural selection. It's function is to regulate behavior and other responses in ways that maximize inclusive fitness. Capacities for learning and relationships are especially valuable ways to accomplish that. Humans have, perhaps uniquely among all animals, capacities for abstract representation that make language and causal thinking possible. Humans also have distinctive capacities for enduring relationships that go beyond merely trading favors. These capacities, in turn, make possible social complexities and cultures that create new selection forces, that further shape brains, in a feedback process that has shaped humans' astounding social and cognitive abilities. An evolutionary approach to human nature emphasizes the remarkable plasticity and adaptability of behavior to different circumstances.

Genotypes and Phenotypes

Natural selection changes gene frequencies, more accurately, allele frequencies. Population genetics describes how selection, migration, and genetic drift increase and decrease allele frequencies in different specific situations. This sometimes gives the impression that evolutionary explanations are genetic explanations, and that gives the impression that an evolutionary approach to behavior somehow implies a commitment to genetic explanations for variations in personality and behavior. These are misconceptions. It is instructive to realize that Darwin knew nothing about genes. The principles of natural selection were discovered by observing variations in phenotypes.

Genes carry the information that is acted on by natural selection, but

that action is almost entirely via interactions of phenotypes and environments. Just as there is no such thing as a trait that is adaptive in general, there is no such thing as an allele that is useful in general; the adaptive significance of a genetic variation depends on interactions with other genes (epistasis) and interactions with environments.

Sexual Selection

The interests of an individual and his or her genes are generally the same. An allele that makes healthy long-lived phenotypes is more likely to be passed on. There is, however, an exception. Mutations that decrease health, happiness, longevity, or cooperativeness will spread nonetheless if they increase the number of offspring.

Sexual selection is the best example. It is the subtype of natural selection in which individuals compete for mates, either by displaying extreme traits that lead to being chosen by the other sex, or by traits that lead to success competing with others of the same sex. Gigantic colorful tails are a costly drag for peacocks, but the responsible genes spread because peahens prefer peacocks with big tails. The process can set up a positive feedback loop in which the advantages of extreme male traits give advantages to female preferences for those traits, because sons with extreme traits will have an advantage. The trait becomes more extreme until the costs it imposes are as large as the reproductive benefits.

The relevance of sexual selection to humans is illustrated by the excess mortality rates experienced by men compared to women. In the early reproductive years in modern societies, men are three times more likely to die than females. Much of the excess is from fights and accidents, but almost all causes of death are greater for males because they allocate effort to competitive ability at the expense of tissue repair. Even early childhood mortality rates are 50% higher for boys than girls.

Sexual selection is a subtype of social selection, the more general category of evolutionary change resulting from social choices and competitions. In sexual selection the resource at stake is mates, with its obvious connection to reproductive success. However, especially for humans, competition for status, friends and group memberships also influences reproductive success. Individuals preferred as relationship partners or group members get advantages substantial enough to shape extreme human capacities for altruism and empathy unmatched by any other species. These advantages also help to account for extraordinary social sensitivity and the prevalence of social anxiety in humans.

Many people experience extreme concern about what others think about them. The explanation is that other people's opinions can help or harm reproductive success. There are good evolutionary reasons why many psychiatric disorders are associated with internalizing or externalizing personality extremes. There are good evolutionary reasons to expect

constant conflicts between (id) impulses to get short-term advantages and (superego) constraints that sacrifice personal short-term gain for longer-term benefits of being a preferred social partner. Rigidity that emphasizes one over the other is disadvantageous and is recognized as sociopathy and neurosis. Most people benefit from mechanisms shaped by selection that allow them to shift their social strategies as a function of subtle shifts in social circumstances.

Levels of Selection

Recognition that selection works mainly at the level of the gene and the individual has profound implications for psychiatry. While few evolutionary biologists now support the idea that traits are routinely selected to benefit the group or the species, consensus about levels of selection remains elusive, and the idea of group selection remains attractive to many nonbiologists. This may be, in part, because it is so often framed as an explanation for "the mystery of altruism." In an evolutionary framework, tendencies that benefit the genes of unrelated individuals more than those of the actor will be selected against, while genes that benefit kin more than the actor can be selected for only if the cost/benefit ratio is greater than the percentage of genes in common.

What about cooperation between those who do not share genes in common by descent? Many are "mutualisms" that give benefits to both parties with no opportunity for defection, such as two people moving a heavy rock to get food. Often, however, exchanges are not immediate, leaving the possibility that one partner will fail to reciprocate a favor. Such "reciprocity" exchanges have been subject to exhaustive modeling using the prisoner's dilemma schema. In this model, the largest payoff comes if the actor defects and the other person cooperates. If both cooperate, they both get a modest benefit. If both defect, neither gets any benefit.

Because the cells in an individual's body are all genetically identical, their interests are almost always aligned with the welfare of the individual. However, some exceptions are particularly important for psychiatry. In particular, the interests of maternal and paternal genomes differ. Female mammals invest a proportion of their available resources in a pregnancy, reserving some for future offspring. The male parent may or may not be the parent of those future offspring, so the paternal genome gets an advantage if the fetus extracts a bit more nutrition from the mother. It may seem preposterous that maternal and paternal genomes should compete over relatively small differences of investment in an infant, yet intriguing genetic evidence suggests this is exactly what is happening.

IGF-II alleles tend to make offspring bigger. When they are passed on via an egg, they are imprinted, that is, methylated in a way that inhibits expression of the gene, thus making the baby smaller. IGF-IIr alleles counteract the actions of IGF-II, yielding larger offspring. IGF-IIr alleles

are imprinted if they come through the paternal line. The balance between these opposing forces yields babies of normal size, but if imprinting is missing for either gene, offspring will be especially small or large.

Bernard Crespi has suggested that this phenomenon may be related to schizophrenia and autism. They can be viewed as flip sides of the same coin, with autism reflecting unbridled influences of the paternal genome, and schizophrenia reflects excess influences of the maternal genome. Genes that increase the risk of schizophrenia and autism are imprinted in ways remarkably congruent with this theory. A recent study found, as predicted, that babies lighter than average were especially likely to develop schizophrenia, while those born heavier than average were more likely to develop autism. These findings do not confirm the theory, but they do illustrate how to test a creative idea inspired by evolutionary thinking.

A more mundane but nonetheless fascinating example is the weaning conflict. There is a period when mammal mothers try to stop nursing but infants protest vigorously and try to continue. This conflict is predictable from evolutionary principles. The mother and infant share 50% of their genes, but the infant shares 100% of its genes with itself. A mother that fails to nurse her infant loses a huge investment, so providing whatever her baby needs early in life is in the mother's genes interests. As the baby grows, it is increasingly able to get nutrition from other sources. At some point this means that the mother's reproductive success would best be served by another pregnancy. The baby's interests, however, are best served by continuing to nurse until the point when the benefits via kin selection to a younger sibling (who shares 50% of genes) are as great as the benefits from continuing to nurse. At this point cessation of nursing is in the child's interest, so the weaning conflict ceases. An extensive literature in psychoanalysis and child development investigates nursing and weaning, but much remains to be done to put it in an evolutionary context.

Misunderstandings about Natural Selection

Some common misunderstandings about evolution deserve mention. Is evolution over for humans? This idea that lower death rates end evolution reflects a deep misunderstanding of natural selection. Natural selection depends on differential reproductive success, not survival. If a heritable trait results in some individuals having fewer offspring than others, then natural selection occurs. Individuals with two different versions of a gene may have the same lifespan, but if one has more offspring, selection will take place. So long as heritable traits influence numbers of offspring, selection will continue to change the average characteristics of the human species. This does not meet that future traits can be predicted with certainty. Death from drunk driving is now a powerful selection force, but it is hard to say whether it will select for disliking alcohol or better driving while drunk.

Can selection have any impact after menopause? Many people think not, but this is because they have not recognized the role of kin selection. Postmenopausal women can have a major influence on the proportion of their genes in future generations if they do things that help their grandchildren. Erickson's generative phase of life has an easy evolutionary explanation. Some have hypothesized that menopause is an adaptation shaped by the advantages of taking care of existing children instead of undergoing the risks of reproduction at advanced age that might jeopardize the welfare of older children.

Do all humans have Stone Age bodies and minds, unchanged for the past 10,000 years? No. Substantial changes have been documented; examples include changes in skin color, ability to digest lactose, and the ability of certain human subpopulations to cope with high altitudes. Increasingly complex social structures may also have imposed strong selection forces that have shaped social and emotional traits and language abilities. Selection for behavioral traits may have sped up since the development of agriculture changed human social structures.

The opposite misconception, that selection changes traits in just a few generations, is also prevalent. However, selection is too slow to adjust the human body to environments that change drastically over just a few generations. Some people think that nearsightedness, a highly heritable trait, has become more common because the invention of eyeglasses has made it less of a disadvantage. However, a few hundred years is not nearly enough time for disadvantageous alleles to drift to a prevalence that would cause disability in about a third of the population. Myopia is a fine example of a "genetic quirk." The responsible alleles probably have disadvantageous consequences only when they interact with some modern environmental factor, possibly early reading, to cause pathology. One wonders how many alleles that predispose individuals to psychiatric disorders are also not defects, but quirks.

Does an evolutionary view of behavior imply some kind of rigidity or genetic determinism? Quite the opposite. Natural selection has shaped innumerable mechanisms that adapt bodies to current circumstances. Learning is not an alternative explanation for behavioral traits, it is but one of many behavioral regulation mechanisms shaped by natural selection to adjust behavior to different circumstances. Similarly, culture is not an alternative explanation for the existence of behavioral traits, although it is obvious that culture has massive effects on behavior. The emergence of complex cultures has created new selection forces that have given humans capacities for culture unmatched by any other species. Gene-culture coevolution escalates human capacities for culture, which in turn makes more complex culture possible, in a positive feedback cycle whose end is not yet in sight.

Much difficulty incorporating evolutionary biology as a basic science for

psychiatry arises because of perceptions that it is associated with atheism. While belief in God is less common among scientists compared to the general public, the fact that humans have been shaped by natural selection does not directly imply anything about the existence of God. It does contradict some myths that are interpreted literally by some religious traditions, however many religions, including the Catholic Church, now officially recognize that humans are products of natural selection.

BEHAVIORAL ECOLOGY

The synthesis of ethology with evolutionary biology produced the field now known as behavioral ecology. It analyzes how behavior patterns influence fitness. Aside from the core principle that behavior tends to maximize inclusive fitness, behavioral ecology focuses on trade-offs between the different kinds of efforts individuals exert to get different kinds of resources. It routinely distinguishes several categories of effort. Somatic effort is to get calories and other resources to grow, sustain, and protect the body. Reproductive effort is to get mates, and to take care of offspring. Social effort is to recruit allies and to achieve status in the group.

This framework offers a powerful tool clinicians can use to unpack the sources of stress. Most stress comes from inability to get a secure, sufficient supply of some resource, threat of loss of a resource, or because efforts to get one resource interfere with efforts to get another. Examples abound in the clinic. One person may be working 70 hours a week, and feeling, correctly, that this makes searching for a romantic partner unrealistic. Another may be stressed because a commitment to full-time childcare makes it hard to allocate effort to anything else. Many people are distressed because of inability to find a mate, insecurity of a marriage, or wanting to leave a relationship. Many others are unable to find a satisfactory job, worried about losing a job, or feeling stuck in a position that provides little respect.

Other problems arise from unbalanced allocations of effort. For some people money is the overwhelming goal, others are dedicated to being good, getting famous, having friends, or taking care of children. The task of balancing efforts to get these various resources is constant, and imbalances create pathology. Narcissism can be viewed as investing all life effort in the pursuit of self and status to the exclusion of all else. Conversely, some patients invest all effort into trying to please others. Some are preoccupied with their appearance, others with their wealth. Stress often arises from health problems that make it impossible to pursue other resources.

To identify the specific source of life stress, one can ask, for each area, whether the person's life strategy is providing sufficient and secure sources of the resource. Few people ever have everything they want, securely and to full measure, but it is nonetheless useful to distinguish patients who have generally adequate sources for most of life's core resources, from those who

are unable to get something crucial, from those who lack sources of most main life resources.

Other animals face the same challenges. A grazing white-tailed deer looks up from the grass every few minutes to see if enemies might be present. Each time the animal looks up, it takes time away from eating. The allocation of effort between foraging and defense is regulated according to the level of nutritional supplies and the danger from predators. Risk-taking is also adjusted according to circumstances. Birds given a choice between a feeder that provides one seed every visit, and one that provides six seeds every sixth visit, generally prefer the steady payoff. If, however, the temperature is lowered to a point where the steady payoff provides too few calories to survive through the night, they switch to the risky payoff because that at least offers some chance of survival.

Studies of foraging also have important implications for psychiatry. Food for most animals is distributed in patches, such as different fruit trees. Starting in a new patch often provides many calories per minute, but as the patch is depleted, payoffs slow. How quickly should the animal move to the next patch? It depends on how long it will take to find a new one. Staying too long will waste time, resulting in very few calories per minute. Moving too quickly to a new patch will waste time looking for food when it could be spent gathering food. The optimum strategy is to continue in the current patch until the rate of return declines to that averaged over many patches. This is Charnov's marginal value theorem. Animals do not have to do calculus to figure out the optimum moving time; their brains have been shaped by selection to optimize such choices to a remarkable degree. This is relevant for attention disorders and for mood. Consider the enthusiasm experienced upon first beginning to pick berries from a new bush, starting a new job, or engaging in a new relationship. Inevitably, payoffs decline with time. Deciding when to make a change is a difficult decision that often is the focus for psychotherapy.

Reproductive Effort

Brains result in behaviors that maximize inclusive fitness, a term that combines the reproductive success of an individual and his or her relatives. The simplest rule of thumb is that natural selection shapes brains that result in behavior that tended, in ancestral environments, to maximize the number of grandchildren, nieces, and nephews.

This makes it seem as if everything depends on mating often and with many partners. Differences in reproduction are evolution's engine, but the notion that reproductive success is maximized by maximizing mating success is simplistic. Compared to chimpanzees, humans have a remarkable ability to selectively inhibit sexual impulses. This ability must have given a selective advantage in the context of the mating patterns that developed during human evolution.

Sex and mating behaviors have been a natural focus for behavioral ecological study. There is no room here to expand on this topic, but it is important to recognize that the very existence of sex poses an evolutionary mystery because parthenogenetic populations grow twice as fast as sexual ones. Also important are the forces that shape many small gametes in one sex (e.g., sperm) and fewer larger ones in the other sex (e.g., ova). This is the essence of the difference between males and females. It means that different mating strategies maximize fitness for males as compared to females. Compared to males, female mammals invest much more in each offspring—many calories, and the time and effort of pregnancy—so the total number of possible offspring is quite limited. The investment of a male can be as little as a few sperm and a few minutes, so one man can have hundreds of children, an outcome uncommon in technological cultures, but reported in cultures where agricultural surpluses made it possible for one man to have many mates. Variation in reproductive success is therefore potentially much larger for men; while some have many offspring, others have none.

However, the human mating system is quite different from that of most other primates. In particular, men make major investments in rearing children. In this respect, humans are more like birds than most other mammals in that rearing offspring takes two parents. Chicks left alone in the nest without protection are likely to quickly be eaten by a predator, so a partnership is required for successfully rearing young. Human babies are so helpless, and their rearing takes so many years, that just having sex is rarely enough to successfully reproduce.

Large investments in child rearing make sense for a man only if he can be relatively certain that the children are his own. So, tendencies to jealousy and mate guarding are human universals, although the range of accepted behaviors varies enormously between cultures, from death by stoning for what would elsewhere be dismissed as mere flirtation, to extramarital sex being acceptable in some situations.

These distinctive aspects of the human niche are likely related to another human peculiarity—human ovulation is concealed instead of being advertised, as it is in many species. This may have increased paternity certainty by drastically lowering the benefits of any forced copulation by a man who is not the woman's mate.

These generalizations are too often pushed to an essentialized version that purports to be THE normal human mating pattern, despite massive evidence for huge cultural variations in mating and family structures. Nonetheless, understanding the evolutionary origins of the human niche can be helpful in understanding marital conflicts, jealousy, and the complications of affairs and divorce.

Life History Traits

Traits such as the duration of gestation, rate of growth, age at maturity, interbirth interval, duration of nursing, and rates of aging are all shaped by selection. One particularly important dimension has been called fast versus slow life strategy. In environments where life is short, reproductive success is maximized by taking risks and reproducing early and often, investing less in each offspring. In more secure environments, it is better to invest more in longer-term strategies and invest more in each offspring. The wide range of human variation on this dimension has been proposed to result from a facultative mechanism that responds to different environments by inducing a faster or slower life history strategy.

Attachment

John Bowlby's recognition of the adaptive value of mother–infant attachment marked the beginning of evolutionary approaches to psychiatry. However, he worked before the major advances that revolutionized the study of behavior, so it is not surprising that he and his followers have tended to see secure attachment as the norm, and other kinds as pathological. Secure attachment is best for mental health, but important recent work has suggested that anxious or avoidant attachment may give selective advantages in some situations.

EMOTIONS

Emotions are special modes of operation shaped by selection to cope with the challenges that arise in situations that have recurred for thousands of years. Like sweating, cough, and pain, they are facultative adaptations that are useful in certain situations. They evolve in concert with mechanisms that monitor for cues associated with the situations in which they are useful. Like all such responses, emotional responses can be pathologically excessive or deficient. Determining if an emotional response is abnormal requires knowing the situation it has been shaped to cope with, and whether or not that situation is present.

When a patient presents with pain, fever, or cough, clinicians search for what might be arousing the response, such as tissue damage, infection, or foreign material in the lungs. When a patient presents with a negative emotion, the same kind of investigation is needed, but the search is more difficult. Instead of being aroused by tissue pathology or infection, emotions are aroused by less tangible situations. Modern emotions theory emphasizes that the stimuli are not just cues; emotions arise from an individual's appraisal of what new information means for his or her ability to pursue personal goals. A thorough investigation of the context is essential to determine if an emotion is normal and useful, normal but useless, excessive, or deficient.

Low mood and anxiety are the emotions that most often cause

problems. Almost all effort to understand these disorders has focused on what is different about people who experience intense anxiety or low mood, compared with other people. An evolutionary framework suggests different questions. Why do the capacities for anxiety and low mood exist at all? In what situations are they useful? Why are the mechanisms that regulate them so vulnerable to failure?

The most fundamental question is why emotions exist at all. Controversies about emotions distract from a growing consensus that they are special states shaped by natural selection along with regulation mechanisms that express a specific emotion in the situation where it gives a selective advantage. No one aspect of emotions is primary. Emotional states shift sensation, perception, cognition, behavior, physiology, motivation, and learning. Subjective experience is just one of many aspects of an emotion.

Controversy about emotion has focused on whether they are best viewed as distinct categories (the basic emotions view), or whether they are best understood as positions on dimensions, such as positive versus negative, and aroused versus calm. Neither perspective is consistent with the evolutionary origins of emotions. Emotions are neither fully separate nor fully dimensional; they are overlapping clusters of changes that have evolved from previous emotions. See Figure 4.1–1.

Some attempts to understand the utility of emotions have tried to map specific functions to specific emotions. However, most emotions serve multiple functions including motivation, communication, and changes in physiology and behavior. Emotions were not shaped to serve specific functions, they were shaped to increase the ability of an individual to cope with the many challenges that arise in a **situation** that has recurred over evolutionary time. Being chased by a predator happened often enough that it reliably evokes a coordinated suite of physiological behavioral and emotional changes, the fight–flight response. Realizing one has broken a promise arouses guilt. Experiencing generous help arouses gratitude. The key to understanding an emotion is to understand the situation it was shaped for.

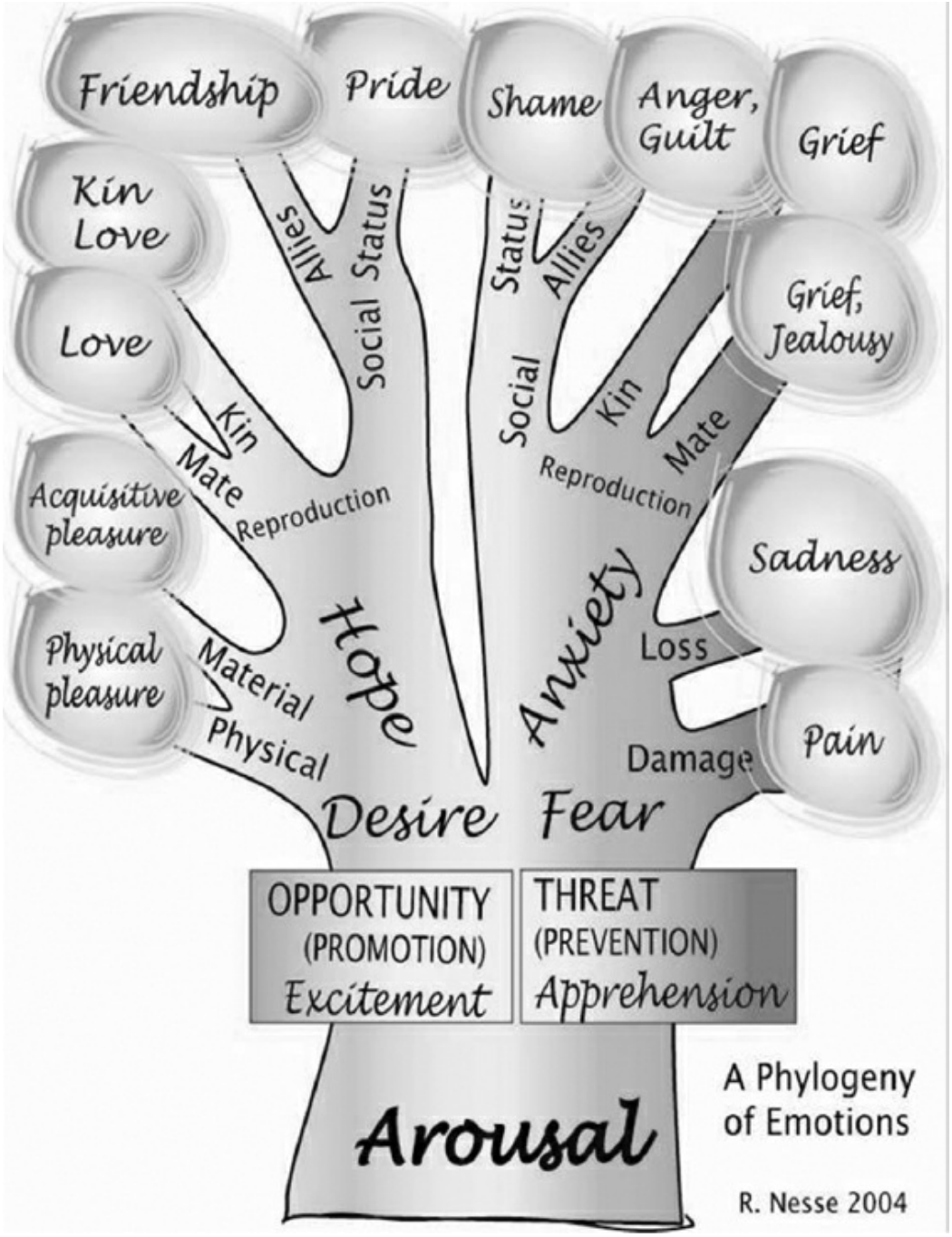

FIGURE 4.1–1. A phylogeny of emotions.

One profound implication for psychiatry is that negative emotions are useful. Their aversiveness is one aspect that contributes to their utility. Anxiety, low mood, jealousy, anger, and envy exist because our ancestors with a capacity for these emotions got a selective advantage compared with those who lacked them. Negative emotions seem pathological because they are usually associated with untoward circumstances, and because they are so painful and prone to interfere with important daily tasks. However, they exist only because they have been useful … for human genes, even if not always for human individuals.

Fear is useful in the face of threat. It mobilizes energy, cognition, and behavior to escape the danger. Like all adaptive responses, pathology can result from deficiencies as well as excesses. People with excessive anxiety crowd psychiatric clinics. People with deficiency of the fear response, hypophobia, cope poorly with dangerous situations, but they rarely come for treatment.

Anxiety is often distinguished from fear because it is aroused by less tangible dangers that are distant or social. It is especially prone to be aroused by situations where an individual is tempted to do something that might give a short-term advantage at a major social cost. Here too, pathology from excessive anxiety, especially social anxiety, is obvious, while the costs of anxiety deficiencies are far less obvious, even though sometimes more devastating. People with social anxiety tend to seek treatment, those with deficient social sensitivity suffer major social losses.

The value of low mood is much less obvious. The term *low mood* allows description of states of low motivation, low self-esteem, and pessimism without the terminological tangles that accompany the word "depression." In what situation can sadness or low mood be useful? Sadness is aroused by a loss. It helps to recover, replace, or otherwise adjust to a loss. Low mood is aroused when efforts to reach a goal are proving fruitless. That situation has recurred millions of times for individuals in the course of human evolution. While simply persisting with a positive attitude seems intuitively attractive, that is not the best strategy when foraging in winter, trapped in pursuit of an uninterested partner, or trying for the fifth year in a row to get into medical school. In such situations it is best to pause, conserve energy, consider other strategies, and, if no route appears viable after all options have been explored, give up, and put energies toward a more achievable goal.

A 20-year-old community college student requested help for moderately severe depression that was interfering with his ability to pass his courses. He said he had to stay in school or his fabulous girlfriend would leave him. When asked about the girlfriend, he said she was still in high school but planning after graduation to attend a college called Vassar. He was dead set on persisting in school despite his failing grades in order to preserve the relationship. Medication and psychotherapy were only modestly helpful over a 2-month period, but he returned at 3 months to report that his symptoms were gone and he had stopped his medication weeks ago. When asked about the girlfriend he said, "Oh, you mean the old girlfriend. She was too uppity for me, my new girlfriend is much more down to earth."

Continued pursuit of an unreachable goal can escalate low mood into clinical depression. There are many possible reasons for inability to give up a goal. Few parents are able to give up trying to get a child off of drugs or alcohol. Many individuals have identities so grounded in their careers that they persist despite years of failure. And often, in the face of impending divorce, one partner persists in trying to preserve the relationship, even as demoralization turns into depression.

Other hypotheses about the origins of capacities for depression are also under consideration. One is that depressive rumination is useful for finding solutions to complex social problems. Another is that depression motivates looking inward at what one can do to avoid being expelled from a group. Another is that depression symptoms give benefits by manipulating others. A particularly influential and plausible proposal is summarized in the phrase "involuntary yielding." After losing a status competition, persisting in challenging the winner will likely bring more attacks or expulsion from the group. Experiencing oneself as helpless and lower status than one actually is can prevent such attacks. The relevant situation is pursuing an unreachable goal in the domain of status competition. Subtypes of low mood have been shaped to cope with the pursuit of different kinds of goals.

Jealousy is aroused by a threat of loss of a mate or the mate's fidelity. Responsible for vast interpersonal problems and millions of murders, sexual jealousy is a toxic emotion. However, the situation of a mating relationship being threatened by an outsider has occurred so often over human evolutionary history, with such extreme effects on reproductive success, that it has shaped a specialized pattern of responses. The actions characteristic of jealousy are often unreasonable in the extreme, but game theory tells us that unpredictability is essential to a successful strategy. Being predictable means being manipulable. Not being fully aware of one's motives or able to control one's behavior may well be advantageous. Once again, the benefits may be for one's genes, not necessarily for oneself.

Table 4.1–2.
Emotions Aroused by Situations That Emerge in Reciprocity Relationships

	Other Cooperates	Other Defects
You cooperate	Friendship, trust	Suspicion (before) Anger (after)
You defect	Anxiety (before) Guilt (after)	Rejection, disgust

Positive emotions are equally useful, even if not, unfortunately, so common. In situations characterized by short-term opportunities, enthusiasm, risk-taking, and energetic initiative are all valuable. This suggests that short-term opportunities will give rise to more enthusiasm and effort than long-term opportunities. Excesses of positive mood rarely motivate treatment-seeking except when they go the extremes found in mania.

The resources that influence human reproductive success are overwhelmingly social, so specific emotions were shaped to cope with the situations that arise in relationships. The boxes of the prisoner's dilemma lend themselves well to understanding the origins of some social emotions. The game is named for situations in which the police offer leniency to

whichever of two suspects first confesses the details of a crime. The dilemma arises because keeping quiet results in a modest punishment, but each prisoner knows that he will get a higher sentence if the other confesses first to avoid any punishment.

Repeated cooperation by both parties arouses friendship, loyalty, and trust. If you cooperate when the other defects, anger is the reliable result. It signals that apologies and reparations are required to preserve the relationship. It also may be accompanied by spite, motivations to harm the defector even when the costs are larger for the self. This apparently senseless emotion gives advantages by its powerful ability to keep others from defecting when they otherwise would. The game-theoretic best strategy of doing what the other person did on the previous move is called tit-for-tat. It accounts for the runs of cooperation and defection seen in relationships (Table 4.1–2).

RELATIONSHIPS

Views of human relationships as fully explained by kinship and reciprocity are grossly simplistic. A small research industry has grown up to try to explain apparent anomalies of human social behavior, including tendencies to cooperation even without guarantee of payback, willingness to punish defectors even when that will harm the actor, and eagerness for spiteful revenge that can have no direct payoff.

A key to such anomalies is found in a variation on the core principle. Just as selection shapes bodies to maximize reproductive success, even at a cost to health or longevity, it shapes minds and emotions to maximize reproductive success at the expense of rationality. As Robert Trivers pointed out in his preface to *The Selfish Gene,* the idea that selection shaped us for objective thought is naïve indeed. Often, especially in social situations, commitments beyond justification, and wishes for revenge beyond measure, are more useful.

This has an important implication for psychoanalysis and psychotherapy in general. While self-knowledge is desirable and can be enormously helpful, lack of conscious access to one's own motives and emotions may sometimes be adaptive. It has been suggested that the human capacity for psychodynamic defenses that maintain active repression of memories and motives may give selective advantages, not just by making life more bearable, but also by influencing others.

The field of sociobiology studies how natural selection shapes social behaviors and resulting group behaviors. There is no room here to explore its well-developed framework; however, when a relationship is at the root of a clinical problem, as is often the case, a formal analysis of how that relationship is or is not working offers a solid starting place. This is very different from trying to understand the personality traits of the participants or the early events that lead to distortions and rigid relationship strategies.

Instead, it looks at the history and current state of exchanges and expectations in the relationship.

While this perspective is just beginning to be applied, some clinical phenomena are readily approached from this point of view. For instance, the classic borderline personality pattern of excessive immediate intimacy and adulation, followed by disappointment and distancing, reflects a rigid relationship strategy of intense personal commitment that, when not reciprocated, leads to rejection and anger. The pattern in neurosis is more one of trying always to please others, and being angry when they do not live up to expectations.

EXPLAINING VULNERABILITY TO SPECIFIC DISORDERS

Every medical disorder needs two kinds of explanation. First, a proximate explanation for what aspect of the body's mechanism has gone awry, and why. Second, an evolutionary explanation for why natural selection has left the body vulnerable to this kind of failure. The six categories from evolutionary medicine help to explain vulnerability to mental disorders, just as they do to other medical disorders.

Disorders of Emotions

As noted already for anxiety and mood disorders, emotional disorders result from dysregulation of otherwise useful responses. This means that determining whether a particular emotion is normal or abnormal requires understanding the situation in which that emotion is useful, the cues that usually regulate it, and the presence or absence of that situation. Current DSM diagnostic criteria ignore context and attend only to frequency, severity, and duration of emotional symptoms. This works well at the extremes and increases diagnostic reliability, but gives no guidance in the more common cases where it is hard to decide if an emotion is normal or not.

The tendency to automatically view negative emotions as pathological contributes greatly to this problem, as does the failure to recognize that emotional responses were shaped to benefit reproductive success, not individual welfare. The functional significance of negative emotions should in no way inhibit attempts to relieve them. The rest of medicine routinely provides relief from the suffering that accompanies normal pain, fever, cough, and other symptoms. Psychiatrists should not be inhibited about relieving anxiety, low mood, and jealousy, even when they are normal responses. Having an evolutionary perspective on the origins and functions of such emotions gives clinicians a framework for understanding the causes of such emotions, and when blocking a negative emotion might be unwise.

A professional motorcycle racer requested treatment because he was unable to sleep or keep food down the night before a race. The problem started when a friend died in a crash. It

became worse after another crash caused another friend to become quadriplegic. He denied symptoms of other disorders and denied problems with anxiety before his friend's crashes. He had predictable difficulty recognizing the potentially life-saving nature of his anxiety.

The other categories of evolutionary explanation for vulnerability also apply for emotional disorders. Useless emotions can be aroused by aspects of modern environments, such as being trapped in a hierarchical bureaucracy. Emotions routinely involve trade-offs, with both costs and benefits from excessive or deficient expression. Finally, emotions all too often motivate behavior that is good for reproductive success, but bad for an individual's health and well-being. Examples include sexual desires that cannot be satisfied, jealousy, envy, wishes for revenge, and ambitions that result in constant dissatisfaction. There are, unfortunately, good evolutionary reasons why such emotions are prevalent, aversive, and hard to control.

Schizophrenia

Most of the variation in risk for schizophrenia results from genetic variation. Why did not natural selection eliminate the responsible alleles? Many have suggested possible advantages to those with the disorder, or, more plausibly, to their relatives. However, reproductive success for people with schizophrenia is a fraction of that for other people, and there is no documented advantage for relatives, so this hypothesis fails. Others have suggested that it is a disease of modern environments, but there is limited evidence for dramatically lower rates in nonindustrial societies. New genetic evidence finds that every identified common allele that increases schizophrenia risk has only a tiny effect. Larger effects are mostly due to rare copy-number variations. This pattern is exactly what one would expect if selection were continuously and efficiently purging vulnerability alleles from the genome. Increased rates of schizophrenia associated with older paternal age and runs of homozygosity further support the hypothesis that alleles predisposing to schizophrenia are in mutation-selection balance.

This leaves open the question of why the cognitive system is so vulnerable to this particular kind of failure. One possibility is that some advantageous cognitive trait has been pushed close to a fitness "cliff edge" beyond which failure is likely, just as strong selection for speed in race horses has made their lower leg bones long, light, and thin, but vulnerable to breaking. The neural mechanism could involve neuron pruning, cortical folding, or any number of other traits that give major benefits up to some threshold beyond which failure becomes likely.

Substance Abuse

The devastation and early death resulting from substance abuse selects against the alleles that increase vulnerability. However, pure substances

and means to administer them have been reliably and readily available for far too little time for selection to have a big effect. From this perspective, vulnerability to substance abuse is a product of mismatch with modern environments. Chemicals mediate the brain's motivation and learning systems, so it should be no surprise that exogenous chemicals can hijack those systems. The amazing thing is that so many people can use drugs and alcohol without major problems. Active efforts are underway to understand why some people have brain mechanisms that make them especially vulnerable. It would be most interesting to see how people with high and low levels of vulnerability behave in environments where substances of abuse are not readily available. Do they use different strategies when foraging for food? Are their interpersonal relationships different? Answering these questions could lead to new behavioral tests that estimate vulnerability.

Eating Disorders

Anorexia and bulimia have likely become more common in the past century. The predisposing alleles are almost certainly quirks without major effects on fitness until they interact with modern environments. Developmental and personality factors strongly influence who gets an eating disorder. However, the shared human vulnerability seems to result from evolved nutrition regulation mechanisms interacting with novel environmental exposures. Good candidates include the ready availability of every imaginable kind of food, a resulting epidemic of obesity, and media that portray slimmer-than-real bodies. Most serious eating disorders are initiated by a strict diet. The resulting binge eating would be life-saving in a famine, but it causes panic and guilt in dieters, and increased resolve to more strictly restrict intake, creating a vicious cycle.

Child Abuse

It makes no evolutionary sense for parents to harm their own children. This led evolutionary psychologists to investigate whether the risk of death of a child depends on how it is related to the parents. The finding is remarkable —death rates from child abuse are 80 times higher if there is a step-parent in the home. This research was inspired by studies of vervet monkeys that found infanticide to be routine when a new male takes over a harem. Eliminating nursing infants speeds return to estrus in the females, thus increasing the reproductive success of the new male during the relatively brief window until another male displaces him. The different human mating system means that infanticide in humans gives no such advantage and was not shaped by selection. However, not being with both parents from the start seems to interfere with the bonding that usually protects children from harm.

Extended early contact also seems to be the signal that makes parents and children uninterested in each other as sex partners. For instance, children raised together from their early years in a kibbutz experience embarrassment and disgust at the idea of marrying another member of the group. Over evolutionary time, this incest inhibition mechanism has increased reproductive success by reducing the risk of having children with many pairs of identical deleterious recessive genes. In this example, cultural prohibitions amplify already existing evolved inhibitions. The relevance to psychiatry is all too obvious in cases where such inhibitions are weak because a parent has little early contact with a child. It is interesting to consider if the development of such inhibitions is related to early Oedipal wishes and their subsequent repression.

Sociopathy

The range of prosocial tendencies in humans is breathtaking. Some individuals lie awake nights wondering if some small oversight might have caused others inconvenience, while sociopaths take pleasure in deceiving others and sadistically causing them pain. In this case, selection seems not to have settled on a narrow optimal mean. Frequency-dependent selection has been proposed as a possible explanation.

If people with sociopathic tendencies do well in groups when there are many cooperators to exploit, but badly in groups where many others are also sociopathic, this could maintain a small proportion of sociopaths in the population. Their tendency toward sexual license and lack of commitment to families and children could give an advantage until the proportion of sociopathic individuals becomes high enough that they must deal with each other, and until cooperative individuals have developed strong abilities to detect and avoid them. This hypothesis is supported by the high heritability for sociopathy, but it is undermined by evidence for minor neurologic abnormalities in many such individuals, and by the role of early abuse and neglect. From a longer-term perspective, the presumed relatively recent evolution of strong prosocial tendencies may turn out to be important. Note that these evolutionary explanations say little about why some people become sociopathic and others do not. They only explain how extreme prosocial traits may give selective advantages, and why not all individuals may have such traits.

Personality Disorders

Other personality disorders are also characterized by rigid extremes of social strategies. People with obsessive-compulsive personality traits do their duty to an extreme, expect others to be equally conscientious, and are constantly disappointed. Ordinary neurotics are not so rigid, but they also invest huge proportions of life effort into trying to please others, and they

also often experience disappointment. Those with paranoid tendencies expect that others will deceive or harm them, and their lack of trust means they have difficulty finding relationship partners who can correct their misperception. Those with anxious, avoidant personalities invest vastly excessive proportions of life energy into protecting against possible harm in general. Dependent personality tendencies result in overwhelming investments in particular relationships, and in preserving those relationships. From an evolutionary point of view, personality disorders are not interpreted as adaptations, but as extremes of normal personality dimensions that reflect different social strategies.

FUTURE DIRECTIONS

The evolutionary foundations for psychiatry provide an invaluable framework for understanding behavior, relationships, and emotions. Clinicians who understand the evolutionary origins and functions of emotions have tools that allow them to better understand disorders of emotions. Those who understand the evolutionary origins of capacities for relationships are better able to deal with problematic relationships and the emotions they arouse. Those who understand how natural selection has shaped extreme prosocial tendencies in humans are better able to understand social anxiety and neurosis. Those who recognize personality characteristics as social strategies have a framework for understanding personality disorders. Most globally, clinicians who grasp the distinction between proximate and evolutionary explanations are well prepared to integrate their knowledge of factors increasing risks for individuals with their knowledge about the evolutionary reasons why all members of the species are vulnerable to a particular disorder.

The implications for research are perhaps even more important. Huge efforts to try to identify specific disorders based on specific brain abnormalities will eventually succeed for some disorders, but others will turn out to be, like epilepsy or congestive heart failure, failures in evolved systems that can have many causes. Severe brain disorders such as autism and schizophrenia may turn out to have a specific, consistent neuropathology, but an evolutionary view suggests that there should be openness to the possibility that they result from systems that have been pushed by selection close to some fitness cliff edge, or that they are otherwise intrinsically vulnerable to failure for other good evolutionary reasons.

Implications for studies of treatment are substantial. Studies of drugs to relieve emotional suffering are often framed as normalizing neurotransmitter abnormalities. However, recognizing that even negative emotions are useful suggests that drugs relieving aversive emotions are like drugs that relieve cough or fever by blocking normal mediating mechanisms. Studies of cognitive-behavioral interventions increasingly go

beyond simple learning theory to also consider social factors that influence schemas, and the advantages as well as the disadvantages of distorted thinking. They also increasingly recognize the power of social groups and relationships to get people to carry out behavioral exercises that are difficult, but essential. Interpersonal therapy analyzes social roles and conflicts that almost invariably involve struggles to get important reproductive resources, and trade-offs among conflicting strategies, roles, and relationships.

This chapter provides only the briefest sketch of a basic science whose implications for psychiatric research and treatment are just beginning to be explored. Recognizing the opportunities will lead to increased support for research that should result in major advances.

REFERENCES

*Alcock J. *The Triumph of Sociobiology*. New York: Oxford University Press; 2001.

Alcock J. *Animal Behavior: An Evolutionary Approach*. 10th ed. Sunderland, MA: Sinauer Associates; 2013.

Baron-Cohen S, ed. *The Maladapted Mind*. East Sussex: Psychology Press, Erlbaum; 1997.

*Brüne M. *Textbook of Evolutionary Psychiatry: The Origins of Psychopathology. 2nd ed.* Oxford; New York: Oxford University Press; 2016.

Brüne M, Belsky J, Fabrega H, et al. The crisis of psychiatry—insights and prospects from evolutionary theory. *World Psychiatry*. 2012;11(1):55–57.

Crespi B. An evolutionary framework for psychological maladaptations. *Psychological Inquiry*. 2014;25(3–4):322–324.

Crespi B, Foster K, Úbeda F. First principles of Hamiltonian medicine. Philosophical Transactions of the Royal Society B: *Biological Sciences*. 2014;369(1642): 20130366–20130366.

Crespi B, Stead P, Elliot M. Evolution in health and medicine Sackler colloquium: Comparative genomics of autism and schizophrenia. *Proc Natl Acad Sci U S A*. 2010;107(Suppl 1):1736–1741.

Del Giudice M, Ellis BJ, Cicchetti D. Evolutionary foundations of developmental psychopathology. *Developmental Psychopathology*. 2014;1.

Fried EI, Nesse RM. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC Medicine*. 2015;1(1):72.

Gilbert P. Evolution and depression: issues and implications. *Psychol Med*. 2006;36(03):287–297.

Gilbert P, Bailey KG. *Genes on the Couch: Explorations in Evolutionary Psychotherapy*. Philadelphia, PA: Taylor & Francis; 2000.

Horwitz AV, Wakefield JC. *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. New York: Oxford University Press; 2007.

Horwitz AV, Wakefield JC. All we have to fear: *Psychiatry's transformation of natural anxieties into mental disorders*. New York, NY: Oxford University Press; 2012.

*Krebs JR, Davies NB. *Behavioural Ecology: An Evolutionary Approach*. 4th ed. Cambridge,

MA: Blackwell Science; 1997.

McGuire MT, Troisi A. *Darwinian Psychiatry*. Cambridge, MA: Harvard University Press; 1998.

Nesse RM. Explaining depression: neuroscience is not enough, evolution is essential. In: Pariente CM, Nesse RM, Nutt DJ, Wolpert L, eds. *Understanding Depression: A Translational Approach*. Oxford; 2009:17–35.

*Nesse RM. Ten questions for evolutionary studies of disease vulnerability. *Evol Appl*. 2011;4(2):264–277.

Nesse RM. *Evolutionary Psychology and Mental Health. In D. Buss (Ed.), The Evolutionary Psychology Handbook*. 2nd ed., Vol. 2: Integrations, pp. 1007–1026. New York, NY: John Wiley and Sons; 2015.

Nesse RM, Bergstrom CT, Ellison PT, et al. Making evolutionary biology a basic science for medicine. *Proc Nat. Acad Sci U S A*. 2010;107:1800–1807.

Nesse RM, Berridge KC. Psychoactive drug use in evolutionary perspective. *Science*. 1997;278:63–66.

Nesse RM, Ellsworth PC. Evolution, emotions, and emotional disorders. *Am Psychol*. 2009;64(2):129–139.

Nesse RM, Stein DJ. Towards a genuinely medical model for psychiatric nosology. *BMC Med*. 2012;10(1):5.

*Nesse RM, Williams GC. *Why We Get Sick—The New Science of Darwinian Medicine*. New York: Times Books; 1994.

Price J, Sloman L, Gardner R, Gilbert P, Rohde P. The social competition hypothesis of depression. *Br J Psychiatry*. 1994;164:309–315.

Stearns SC, Koella JC, eds. *Evolution in Health and Disease*. 2nd ed. New York: Oxford University Press; 2008.

Troisi A. Mental health and well-being: clinical applications of Darwinian psychiatry. *Applied Evolutionary Psychology*. 2012:276.

van Dongen J, Boomsma DI. The evolutionary paradox and the missing heritability of schizophrenia. *Am J Med Genet B Neuropsychiatr Genet*. 2013;162(2):122–136.

Wenegrat B. *Sociobiological Psychiatry: A New Conceptual Framework*. Lexington, MS: Lexington; 1990.

▲ 4.2 Neurocentrism: Implications for Addiction and the Courtroom

SALLY L. SATEL, M.D., AND SCOTT O. LILIENFELD, PH.D.

In the early 21st century, neuroscience has captured the public imagination like never before. Understanding the brain is helpful, perhaps essential, to developing treatments for devastating illnesses like schizophrenia and Parkinson's. More abstract but no less compelling, the functioning of the brain is intimately tied to one's identity, memories, and aspirations. But the excitement to explore the brain has spawned a new fixation that one might call *neurocentrism*—the view that human behavior can be best explained by looking solely or primarily at the brain.

In the view of some critics neurocentrism poses a threat to psychiatry because it risks oversimplifying a number of complex issues. In its extreme form, it devalues the importance of psychological explanations and environmental factors, such as familial chaos and stressors, understanding disturbances of mood, thought, and overt behavior. In turn, neurocentrism may pose a distraction from seeking effective and perhaps more enduring solutions, including psychotherapies, and may imply that pharmaceutical approaches should be first-line intervention for addictions and other behavioral problems.

The prime impetus behind this enthusiasm is functional magnetic resonance imaging (fMRI) and the now-iconic, vibrant brain images that grace the science pages of the daily newspaper. Author Tom Wolfe was characteristically prescient when he wrote of fMRI in 1996, just a few years after its introduction, "Anyone who cares to get up early and catch a truly blinding twenty-first century dawn will want to keep an eye on it." Why the fixation? First, there is the very subject of the scans: the brain itself. More complex than any structure in the known cosmos, the brain is a masterwork of nature endowed with cognitive powers that far outstrip the capacity of any silicon machine built to emulate it. Containing roughly 85 billion brain cells, or neurons, each of which communicates with thousands of other neurons, this 3-lb universe cradled between the ears has more connections than there are stars in the Milky Way. How this enormous neural edifice gives rise to subjective feelings, a question often called the "hard problem" of consciousness, is one of the greatest mysteries of science and philosophy. Combine this mystique with the simple fact that multicolored pictures—in this case, brain scans—can be powerful. Of all the human senses, vision is the most developed and occupies the largest share of cortical space.

There are good evolutionary reasons for this arrangement: The major threats to primate ancestors were apprehended visually, as were their sources of food. Plausibly, the survival advantage of vision gave rise to man's reflexive bias for believing that the world is as he or she perceive it to be, an error that psychologists and philosophers call naive realism. This misplaced faith in the trustworthiness of one's perceptions is the wellspring of two of history's most famously misguided theories: that the world is flat and that the sun revolves around the earth. For thousands of years, people trusted their raw impressions of the heavens. Yet, as Galileo understood all too well, the eyes can deceive. He wrote in his *Dialogues* of 1632 that the Copernican model of the heliocentric universe commits a "rape upon the senses"—it violates everything the eyes reveal.

Trusting the patterns on brain scans to reveal nuanced mental contents is a form of neurorealism, a term coined by University of Montreal researcher Eric Racine. A first cousin of naive realism, neurorealism denotes the fallacious but tempting propensity to regard brain images as

inherently more "real" or valid than other types of behavioral data. As Stanford neuroeconomist Paul Zak has described his work on the neurobiology of trust, a brain scan "lets me embrace words like 'morality' or 'love' or 'compassion' in a non-squishy way. These are real things." Or take the psychological impact of combat. A researcher at the Minneapolis VA tells TIME magazine that claimed that brain imaging confirms that posttraumatic stress disorder (PTSD) is a "real disorder." In both quotes, it is not clear that brain images provide with novel information; such images are not required to conclude that love or PTSD is genuine.

As a tool for exploring the biology of the mind, neuroimaging has given neuroscience not merely a huge scientific boost but a strong cultural presence. As one scientist remarked, brain images are now "replacing Bohr's planetary atom as the symbol of science." With its implied promise of decoding the brain, it is easy to see why brain imaging would beguile almost anyone interested in pulling back the curtain on the mental lives of others: politicians hoping to manipulate voter attitudes, marketers tapping the brain to learn what consumers really want to buy, agents of the law seeking an infallible lie detector, addiction researchers trying to gauge the pull of temptations, psychologists and psychiatrists seeking the causes of mental illness, and defense attorneys fighting to prove that their clients lack malign intent or even free will. At the same time, this fascination is a double-edged sword, as it may lead eager audiences to accept dubious neuroscience claims without adequate scrutiny.

Some misapplications of neuroscience are merely amusing and may be essentially harmless. Take, for instance, the new trend of neuro-management books, such as one entitled *Your Brain and Business: The Neuroscience of Great Leader*. The latter advises nervous CEOs "to be aware that anxiety centers in the brain connect to thinking centers, including the PFC [prefrontal cortex] and ACC [anterior cingulate cortex]." The fad has, perhaps not surprisingly, infiltrated the parenting and education markets, too. Parents and teachers are easy marks for "brain gyms," "brain-compatible education," and "brain-based parenting," not to mention dozens of other unsubstantiated techniques. Although these methods may not be dangerous *per se,* they may incur both direct financial costs and opportunity costs arising from a failure to seek out more efficacious interventions.

For the most part, these enterprises merely dress up or repackage good advice with neuroscientific findings that add little or nothing to the overall program. As one cognitive psychologist quipped, "Unable to persuade others about your viewpoint? Take a Neuro-Prefix—influence grows or your money back." But reading too much into brain scans can be a problem when real-world concerns hang in the balance. Consider the law. When a person commits a crime, who is at fault: the perpetrator or his or her brain? This is a false choice. If biology teaches anything, it is that "my brain"

versus "me" is a false distinction. Still, if biological roots can be identified—and better yet, captured on a brain scan as juicy blotches of color—it is too easy for nonprofessionals to assume that the behavior under scrutiny must be "biological" and therefore "hard-wired," that is, involuntary or uncontrollable. Criminal lawyers, not surprisingly, are drawing increasingly on brain images supposedly showing a biological defect that "made" their clients commit murder.

Looking to the future, some neuroscientists envision a dramatic transformation of criminal law. They anticipate a time when scientists will be able to attach a detailed biological explanation to many types of bad behavior and eventually come to think about the decision making that gave rise to such behavior in the same way one thinks about any physical process, such as diabetes or lung disease. It is plausible that highly detailed mechanical accounts will be possible in the future, the prediction that jurors will necessarily come to see criminals as less blameworthy, if at all accountable, is far less certain, and potentially problematic, as this chapter will show.

But this may not be the correct conclusion to draw from neuroscientific data. After all, if every behavior is eventually traced to detectable correlates of brain activity, does this mean that all troublesome behavior will one day be understood on a "don't-blame-me-blame-my-brain theory" of misbehavior? Will no one ever be judged responsible? This chapter is intended to show why thinking through these profoundly important questions hinges on how one understands the relation between the brain and the mind.

BRAIN AND MIND'WHY THEY ARE DIFFERENT

The "the mind" is basically what the brain (and the rest of the central nervous system) does, but the mind and the brain are not identical.

They are different because they represent different levels of analysis. Virtually all neuroscientists, psychiatrists, and psychologists certainly do not believe that brain and mind occupy materially separate domains. Every subjective experience, from the ache of nostalgia to the frisson of a Christmas morning, corresponds to physical events in the brain. The mind—the realm of feelings, desires, ideas, memories, intentions, and subjective experience—is produced by the action of neurons and brain circuits. Yet, the mind is not identical with the matter that produces it. So, to say that the brain and the mind are different is not necessarily evidence of scientific ignorance, as some have alleged. It means simply that one cannot use the physical rules from the cellular level to completely predict activity at the psychological or behavioral level.

As an analogy, consider the words Shakespeare wrote on a page. They are, of course, "material," but one could not capture the meaning of his plays by conducting a chemical analysis of the ink in which the letters of his

words were written. The brain and the mind are different levels of explanation of the same phenomenon; the brain refers to mechanisms enabling perception, emotion, and cognition; the mind to subjective experiences such as awareness and meaning. Because brain and mind are different levels of description, they have different properties. This position is called *property dualism*. It must be distinguished from substance dualism, which implies that brain and mind are materially different.

From the standpoint of property dualism one cannot rely on the brain alone to predict or understand everything important about human subjectivity or behavior. This is because many psychological phenomena are emergent properties of lower-order constituents, such as neural circuits, neurons, proteins, and genes. "Constitutive" reductionism—reducing complex entities to the sum of their component parts to facilitate study—is not controversial in the scientific community. In contrast, radical or "eliminative" reductionists go a crucial step further. They insist that everything mental will ultimately be *explained* fully at the material level of analysis. This variant of reductionism ultimately eliminates the need for the psychological level of analysis, not to mention all other levels—social, cultural, and so on.

Adherents of eliminative reductionism speculate that such properties (the painting in full) will ultimately prove reducible to more basic elements (the paint). They may prove correct. But for the foreseeable future, valuable information is often lost when descending from higher explanatory levels, such as mental states, to lower levels, such as neuronal systems. To be sure, the neural level of explanation is appropriate for certain scientific questions, such as the nature and treatment of Alzheimer disease. Therapeutic research focuses on biology, such as the amyloid plaques that disrupt communication among neurons, and neurofibrillary tangles that degrade them.

Other times, however, a neural explanation can lead one astray. Consider depression. Although depression is clearly influenced by biological factors, including a genetic predisposition, this does not mean that it is necessarily best treated by medication. In fact, certain psychological treatments, especially cognitive-behavioral and interpersonal psychotherapies, appear to be at least as effective as medications for depression while having no known physical side effects and often showing superior prevention of future episodes.

LEVELS OF ANALYSIS, ADDICTION, AND THE BRAIN DISEASE FALLACY

Distinguishing brain from mind, or criticizing the excesses of brain imaging, is hardly a repudiation of the great insight of human neuroscience: Mental life arises from brain activity. Both levels, the brain and the mind, are essential to a complete understanding of human experience and behavior. The field of addiction psychiatry provides a good

example. Within the last decade in particular, neurocentrism has been ascendant, and not always for the better. As a consequence of heavy promotion by the National Institute on Drug Abuse (NIDA), part of the National Institutes of Health, addiction has been labeled a "brain disease." The logic for this designation, as explained by a former director of the NIDA is that "addiction is tied to changes in brain structure and function." It is certainly true that repeated use of drugs such as heroin, cocaine, and alcohol alters the neural circuits that mediate the experience of pleasure as well as motivation, memory, inhibition, and planning—modifications that are often visible on brain scans. The critical question, though, is whether this neural disruption renders the addict incapable of self-control. Evidence indicates that he or she is not.

The neurocentric model leaves the addicted person in the shadows. Yet, to treat addicts and guide policy, it is important to understand how addicts think. It is the *minds* of addicts that contain the stories of how addiction happens, why they continue to use, and, if they decide to stop, how they cope with life stressors. The answers cannot be divined alone from an examination of addicts' brains alone, no matter how sophisticated the probe. Viewing addiction as a "chronic and relapsing" condition diverts attention from promising behavioral therapies that challenge the inevitability of relapse by holding patients accountable for their choices. At the same time, because the model implies that addicts cannot stop using drugs until their brain chemistry returns to normal, it overemphasizes the value of brain-level solutions, such as pharmaceutical intervention.

Pragmatically, psychiatrists should be open to using any treatment if there is persuasive evidence that it is effective. But according to some thoughtful critics, NIDA has overemphasized the importance of pharmacological intervention in the care of addicts. In 1997, the head of NIDA ranked the search for a medication to treat methamphetamine addiction as a "top priority." A decade later, another director told *Newsweek* magazine that psychiatrists will "treating addiction as a disease [by 2018], and that means with medicine." The search for a magic bullet is at best dubious—and even NIDA eventually gave up hope of finding a wonder drug—but simplistic forms of the brain-disease narrative continue to inspire unrealistic goals. When British pop star Amy Winehouse succumbed to her high-profile alcoholism in the summer of 2011, a *Psychology Today* blogger asked rhetorically, "Could neuroscience have helped Amy Winehouse?" The author answered in the affirmative, suggesting a dopamine-altering medication of the future because addiction "may be a brain problem that science can eventually solve." One neuroscientist goes even further, asserting that "addiction can be reasonably viewed as a neurological problem that allows for medical solutions, just as pneumonia can be viewed as a lung problem." But the analogy is questionable at best. Changing a behavior such as addiction

requires addicts to work hard to change their patterns of thought and behavior.

In contrast, antibiotic cures for pneumonia work even if the patient is in a coma. The hope of a medical treatment is the logical outgrowth of placing the brain at the center of the addictive process. Overall, success to date has been genuine but modest. When motivated patients take medications—especially patients already armed with relapse-prevention strategies and the support of family and friends—they can sometimes vault into sustained recovery. Methadone, a long-acting synthetic opiate taken once a day to prevent opiate withdrawal has played a major role in treating addiction to heroin and painkillers since the 1960s. Still, up to half the patients in methadone clinics also fortify themselves with heroin, cocaine, or benzodiazepines, sold on the street.

Despite three decades of effort, there is still no medication therapy for cocaine dependence. Cocaine immunotherapy (popularly called a cocaine "vaccine") to prevent cocaine molecules from entering the brain is now in development, but initial data do not look promising for wide-scale use. Other types of medications include blocking agents, such as naltrexone for opiate addiction, which occupy receptors for endorphins and blunt a drug's effect. Aversive agents, such as Antabuse (disulfiram), cause people to feel nauseated and vomit when they ingest alcohol. They can be effective in some cases, although many individuals elect to stop taking them.

NEURAL CORRELATES AND VOLUNTARINESS COEXIST

To be critical of the neurocentric model of addiction is not to dismiss the reality and research importance of biological correlates, but rather to put them in perspective. The reward circuit is intimately involved in "cue-induced" craving. Such craving is a special species of desire that manifests itself in a sudden, intrusive urge to use drugs brought on by cues associated with use. The mere clink of a whiskey bottle, a whiff of cigarette smoke, or a glimpse of an old drug dealer on the corner can set off an unbidden rush of yearning, fueled by dopamine surges. For the addict who is trying to quit, this is typically a tense, not a pleasurable, feeling. Because the rush of desire typically seems to come out of the blue, users may feel blindsided, helpless, and confused.

In a typical demonstration of the biological correlates of craving, researchers study addicts as they lie in a brain scanner and watch videos of people handling a crack pipe or needle. In response, their prefrontal cortices, amygdalae, and other brain structures typically bloom with activity. (Videos of neutral content, such as landscapes, usually induce no such response.) Even in users who quit several months previously, neuronal alterations may persist, leaving them vulnerable to sudden, strong urges to use drugs.

The familiar late-1980s slogan "This is your brain on drugs" still echoes,

but now manifests with the brain itself substituting for the fried egg. But that egg is not always sizzling. There is a surprising amount of lucid time in the daily life of addicts. In their classic 1969 study "Taking Care of Business: The Heroin User's Life on the Street," criminologists Edward Preble and John J. Casey found that addicts spend only a small fraction of their days getting high. Most of their time is spent either working or hustling. The same is true for many cocaine addicts, as work by ethnographer Philippe Bourgois reveals. Most people tend to think of them, at their worst, frantically gouging their skin with needles, jamming a new rock into a pipe every fifteen minutes, or inhaling lines of powder. In the grip of such hunger, an addict cannot be expected blithely to get up and walk away.

These tumultuous states—with neuronal function severely disrupted—are the closest drug use comes to being beyond the user's restraint. But in the days between binges, cocaine users worry about a host of everyday matters, such as whether to find a different job, enroll their children in school, or attend a Narcotics Anonymous meeting. It is during these stretches of relative calm that many addicts could make the decision to get help or quit on their own—and many of them do. But the decision to quit can be long in coming, far too long for those who destroy their health, families, or careers in the meantime.

The paradox at the heart of addiction is this: How can the capacity for choice coexist with self-destructiveness? "I've never come across a single person who was addicted who wanted to be addicted," is a common observation. The implication is that because the state was not desired, it must have been involuntary, that is, beyond the user's control. But how many have ever come across a heavy person who wanted to be overweight? Many undesirable outcomes in life arrive incrementally. "We can imagine an addict choosing to get high each day, though not choosing to be an addict," writes psychologist Gene Heyman. "Yet choosing to get high each day makes one an addict."

Imagine a typical trajectory to see how this dynamic unfolds. In the early phase of addiction, drugs or alcohol become ever more appealing, while once-rewarding activities, such as relationships, work, or family, decline in value. The attraction of the drug starts to fade as consequences accrue—spending too much money, disappointing loved ones, attracting suspicion at work—but the drug retains its allure because it blunts psychic pain, suppresses withdrawal symptoms, and douses intense craving. Addicts find themselves torn between the reasons to use and reasons not to.

Sometimes a spasm of self-reproach or a flash of self-awareness tips the balance toward quitting. William S. Burroughs, an American novelist and heroin addict, calls this the "naked lunch" experience, "a frozen moment when everyone sees what is on the end of every fork." Christopher Kennedy

Lawford, himself in recovery from drugs and alcohol, edited a 2009 collection of essays called *Moments of Clarity* in which the actor Alec Baldwin, singer Judy Collins, and others recounted the events that spurred their recoveries. Some quit on their own; others obtained professional help. A theme in each of their stories is a jolt to self-image: "This is not who I am, not who I want to be." One recovered alcoholic described the process: "You tear yourself apart, examine each individual piece, toss out the useless, rehabilitate the useful, and put your moral self back together again." These are not the sentiments of people in helpless thrall to their diseased brains. Nor are these sentiments the luxury of memoirists. Patients have described similar experiences: "My God, I almost robbed someone!" "What kind of mother am I?" or "I swore I would never switch to the needle."

Data consistently demonstrate that quitting is the rule, not the exception—a fact worth acknowledging, given that the official NIDA formulation is that "addiction is a *chronic and relapsing* [italics added] brain disease." The Epidemiologic Catchment Area Study, done in the early 1980s, surveyed 19,000 people. Among those who had become dependent on drugs by age 24, more than half later reported not a single drug–related symptom. By age 37, roughly 75 percent reported no drug symptom. The National Comorbidity Survey, conducted between 1990 and 1992 and again between 2001 and 2003, and the National Epidemiologic Survey on Alcohol and Related Conditions, conducted between 2001 and 2002 with more than 43,000 subjects, found that 77 and 86 percent of people who said they had once been addicted to drugs or alcohol reported no substance problems during the year before the survey. By comparison, people who were addicted within the year before the survey were more likely to have concurrent ("comorbid") psychiatric disorders.

Additionally, NIDA estimates that relapse rates of treated drug-addicted patients range from 40 to 60 percent. In other words, they are not representative of the universe of addicts. They are the challenging cases—the chronic and relapsing patients. Yet these patients often make the biggest impressions on clinicians and shape their views of addiction, if only because clinicians are especially likely to encounter them. Falling prey to what psychologists term an "availability heuristic," researchers and medical professionals may err in generalizing from the sickest subset of people to the overall population of patients.

This caveat applies across the medical spectrum. Just as the clinician wrongly assumes that all addicts must be like the recalcitrant ones who keep stumbling through the clinic doors, psychiatrists and other mental health professionals sometimes view people with schizophrenia as doomed to a life of dysfunction on the basis of their frequent encounters with those whose delusions and hallucinations do not improve with treatment. The error of extrapolating liberally from these subsets of difficult patients is so common that statisticians Patricia and Jacob Cohen dubbed it the

"clinician's illusion."

At this time, behavioral and psychological approaches have the research edge in treating addiction. In December 1966, Leroy Powell of Austin, Texas, was convicted of public intoxication and fined $20 in a municipal court. Powell appealed the conviction to county court, where his lawyer argued that he suffered from "the disease of chronic alcoholism." Powell's public display of inebriation therefore was "not of his own volition," and the fine constituted cruel and unusual punishment. A psychiatrist concurred, testifying that Powell was "powerless not to drink."

Then Powell took the stand. On the morning of his trial, he had a drink at 8 a.m. that his lawyer gave to him, presumably to stave off morning tremors. Here is an excerpt from the cross-examination:

Q: You took that one [drink] at eight o'clock [a.m.] because you wanted to drink?
A: Yes, sir.
Q: And you knew that if you drank it, you could keep on drinking and get drunk?
A: Well, I was supposed to be here on trial, and I didn't take but that one drink.
Q: You knew you had to be here this afternoon, but this morning you took one drink and then you knew that you couldn't afford to drink anymore and come to court; is that right?
A: Yes, sir, that's right.
Q: Because you knew what you would do if you kept drinking that you would finally pass out or be picked up?
A: Yes, sir.
Q: And you didn't want that to happen to you today?
A: No, sir.
Q: Not today?
A: No, sir.
Q: So you only had one drink today?
A: Yes, sir.

The judge let stand Powell's conviction for public intoxication. A second appeal followed, this time to the U.S. Supreme Court. It, too, affirmed the constitutionality of punishment for public intoxication. "We are unable to conclude," said the court, "that chronic alcoholics in general, and Leroy Powell in particular, suffer from such an irresistible compulsion to drink and to get drunk in public that they are utterly unable to control their performance."

For people like Powell who are not otherwise motivated to quit, consequences can play a powerful role in modifying behavior. Powell took only a single drink on the morning of his trial because of foreseeable and meaningful consequences. Far from being unusual, his ability to curtail his drinking accords with a wealth of studies showing that people addicted to all kinds of drugs—nicotine, alcohol, cocaine, heroin, methamphetamines—can change their behavior in response to rewards or sanctions. Powell had surely experienced many alcohol-induced brain changes, but they did not keep him from making a choice that morning. If Powell came before a judge today, his lawyer might well introduce a scan of his brain "craving" alcohol as evidence of his helplessness. If so, the judge would be wise to reject the scan as proof. After all, a judge, or anyone, can ponder scans of "addicted" brains all day, but he or she would never consider someone an addict unless that person behaves like one.

Researchers have found that the brains of smokers reporting a strong desire to smoke commonly display enhanced activation of reward circuitry, as would be expected. But they also showed that subjects could reduce the brain's telltale signs of craving by considering the long-term consequences of smoking, such as cancer or emphysema, while observing videos depicting people smoking. When subjects did so, their brains displayed enhanced activity in areas of the prefrontal cortex associated with focusing, shifting attention, and controlling emotions. Simultaneously, activity in regions associated with reward, such as the ventral striatum, decreased.

Investigators at NIDA observed the same pattern when they asked cocaine users to inhibit their craving in response to cues. Subjects underwent positive emission tomography (PET) scanning as they watched a video of people preparing drug paraphernalia and smoking crack cocaine. When researchers instructed the addicts to control their responses to the video, they observed inhibition of brain regions normally implicated in drug craving. When not deliberately suppressing their cravings, the addicts reported feeling their typical desire to use drugs, and the PET scans, accordingly, revealed enhanced activation in brain regions that mediate craving.

These powerful findings illuminate the capacity for self-control in addicts. They also underscore the idea that addicts persist not because of an inability to control the desire to use but at least in part because a failure of motivation. Granted, summoning sustained motivation can be a great challenge: It takes considerable energy and vigilance to resist craving, especially urges that ambush the addict unexpectedly. Studies on the regulation of craving also help distinguish behaviors that people do not control from behaviors that they cannot control. Imagine, by way of contrast, promising a reward to people with Alzheimer's if they can keep their dementia from worsening. That would be both pointless and cruel because the kinds of brain changes intrinsic to dementia are immune to rewards or penalties.

What Powell's case showed was that even though he sustained brain changes, those changes did not prevent his behavior from being shaped by consequences. Contingency management—the technical term for the behavior therapy technique of adjusting consequences, including incentives—often succeeds with people who face serious losses, such as their livelihood, professional identity, or reputation. When addicted physicians come under the surveillance of their state medical boards and are subject to random urine testing, unannounced workplace visits, and frequent employer evaluations, they fare well: 70 to 90 percent are employed with their licenses intact 5 years later. Likewise, scores of clinical trials show that addicts who know they will receive a reward, such as cash, gift certificates, or services, are nearly two to three times as likely to submit drug-free urine samples as addicts not offered rewards.

Unfortunately, treatment programs are rarely in a position to offer cash or costly rewards. But the criminal justice system has an ample supply of incentives at its disposal and has been using such leverage for years. One of the most promising demonstrations of contingency management comes from Honolulu in the form of Project HOPE, Hawaii's Opportunity Probation with Enforcement. Project HOPE includes frequent random drug testing of offenders on probation. Those who test positive are subject to immediate and brief incarceration. Sanctions are fair and transparent: All offenders are treated equally, and participants know what will happen in

case of an infraction. The judges express a heartfelt faith in offenders' ability to succeed. These basic elements of HOPE's contingency administration—swiftness, sureness, transparency, and fairness combined with expectation for achievement—are a potent prescription for behavior change in just about anyone. Indeed, after 1 year of enrollment in Project HOPE, participants fared considerably better than probationers in a group who served as a comparison. They were 55 percent less likely to be arrested for a new crime and 53 percent less likely to have had their probation revoked. These results are even more impressive in light of the participants' criminal histories and their heavy, chronic exposure to methamphetamine, which can impair aspects of cognitive function.

These findings join a vast body of experimental data attesting to the power of incentives to override the lure of drugs. Yet because the facts contradict the idea that addiction is analogous to Alzheimer disease, some HOPE personnel objected to incentives, arguing that addicts could not be accountable for their behavior. Indeed, the brain-disease narrative misappropriates language better used to describe such conditions as multiple sclerosis or schizophrenia—afflictions of the brain that are neither brought on by the sufferer nor modifiable by the desire to be well. It offers false hope that an addict's condition is completely amenable to a medical cure. Finally, it threatens to obscure the vast role of personal agency in perpetuating the cycle of use and relapse.

Addicts embarking on recovery often need to find new, clean and sober friends, travel new routes home from work to avoid passing near their dealer's street, or deposit their paycheck directly into a spouse's account to keep from squandering money on drugs. A teacher trying to quit cocaine switched from using a chalkboard—the powdery chalk was too similar to cocaine—and had a whiteboard installed instead. An investment banker who loved injecting speedballs—a cocktail of cocaine and heroin in the same syringe—made himself wear long-sleeved shirts to prevent glimpses of his bare and inviting arms. Former smokers who want to quit need to make many fine adjustments, from not lingering at the table after meals to ridding their homes of the everpresent smell of smoke, removing car lighters, and so on.

Thomas Schelling, a 2005 Nobel laureate in economics, refers to these purposeful practices as self-binding. The great self-binder of myth was Odysseus. To keep himself from heeding the overpowering song of the sea sirens—the half-woman, half-bird creatures whose beautiful voices lured sailors to their deaths—Odysseus instructed his men to tie him to the mast of his ship. The famous Romantic English poet Samuel Taylor Coleridge, an opium addict, is said to have hired men to prevent him from entering a pharmacy to purchase opium. Today, one can hire a firm that will provide binding services. It imposes surprise urine tests on the client, collects evidence of attendance at Alcoholics Anonymous (AA) meetings or

treatment sessions, and sends a monthly status report (with the good or bad news) to another person, such as a parent, spouse, or boss.

Some addicts devise their own self-binding strategies. Others need the help of therapists, who teach them to identify and anticipate cues that trigger craving. Beyond the classic triad of people, places, and things, addicts may come to realize that internal states, such as stress, bad moods, and boredom, can prompt drug urges. Managing craving matters mightily in recovery, but it usually is not enough. Another important truth is that an addict uses drugs or alcohol because they serve a purpose, or what behavioral psychologists term an operant function. Caroline Knapp, in her powerful 1996 memoir *Drinking: A Love Story,* recounted why she spent two decades of her life as an alcoholic: "You drank to drown out fear, to dilute anxiety and doubt and self-loathing and painful memories." Knapp doesn't describe an urge to drink so much as a need to drink. She was not manipulated by an alien desire but by something woven into her being. To say that Knapp's problem was merely the effect of heavy drinking on her brain is to miss the larger picture.

Heroin and speed helped screenwriter Jerry Stahl, author of *Permanent Midnight,* attain "the soothing hiss of oblivion." But when the drugs wore off, his vulnerabilities throbbed like a fresh surgical incision. In surveying his life, Stahl wrote, "Everything, bad or good, boil[ed] back to the decade on the needle, and the years before that imbibing everything from cocaine to Romilar, pot to percs, LSD to liquid meth and a pharmacy in between: a lifetime spent altering the single niggling fact that to be alive means being conscious." Or take Lisa, a thirty-seven-year-old woman featured in a 2007 HBO documentary on addiction. When viewers meet her, Lisa is living in a rundown hotel room in Toronto and working as a prostitute. She sits on the bed and talks with the filmmaker behind the camera. Flipping her shiny brown hair and inspecting her well-kept nails, Lisa is animated as she boasts about how much she makes selling sex, how much she spends on cocaine, and the longed-for "oblivion" that drugs help her attain. When Lisa was filmed, she was healthy and engaging; she looked and talked like someone who had recently been abstinent but was back in the early stages of her next downward spiral. She had no interest in stopping things at this point. "Right now, I am in no position to go into recovery. [This way of life] is working for me ... I have money, drugs, business. I'm O.K." To say that Lisa's problem is the effect of cocaine on her brain distracts from true threat to her well-being: Lisa herself. "I always use for a reason. It's repressing what needs to be repressed," she says.

These stories highlight one of the shortcomings of the strictly neurocentric view of addiction. This perspective ignores the fact that many people are drawn to drugs because the substances temporarily quell their pain: persistent self-loathing, anxiety, alienation, deep-seated intolerance of stress or boredom, and pervasive loneliness. The brain-disease model is of little use here because it does not accommodate the emotional logic—or illogic—that triggers and sustains addiction.

NEUROCENTRISM AND THE LAW

It is only natural that advances in knowledge about the brain make people think more mechanistically about themselves. But in one venue, in particular—the courtroom—this bias can be a prescription for confusion. The brain-based defense ("Look at this fMRI scan, your Honor. My client's brain made him do it.") is being used with increasing frequency in capital defenses. The problem with these claims is that, with rare exception, neuroscientists cannot yet translate aberrant brain functions into the legal requirements for criminal responsibility—intent, rational capacity, and self-control.

What is known about many criminals is that they did not control themselves. That is very different from being unable to do so. To date, brain science cannot allow distinctions between these alternatives. Moreover, even abnormal-looking brains frequently have owners who are otherwise quite normal. Looking to the future, some neuroscientists envision a dramatic transformation of criminal law. David Eagleman of the Baylor College of Medicine's Initiative on Neuroscience and Law, hopes that "we may someday find that many types of bad behavior have a basic biological explanation [and] eventually think about bad decision making in the same way we think about any physical process, such as diabetes or lung disease."

But is this the correct conclusion to draw from neuroscience? If every troublesome behavior is eventually traced to correlates of brain activity that can be detected and visualized, will society be able to excuse it on a don't-blame-me-blame-my-brain theory? Will no one ever be judged responsible? Eagleman's way of thinking may fall prey to what law professor Stephen Morse calls the "psycho-legal error," the powerful temptation to equate cause with excuse. Morse notes that the law excuses criminal behavior only when a causal factor produces an impairment so severe that it deprives the defendant of his or her rationality. Bad genes, bad parents, or even bad stars are not an excuse.

Nothing less than the legal system's authority to hold offenders accountable hinges on a clear grasp of the relationship of mental contents and mental capacity to ascriptions of responsibility. How can neuroscientific data assist the law in ascertaining guilt? To answer this question, it is important to understand how the law determines guilt. Some brief background: American criminal law holds a person responsible for a crime if he or she intended to commit a prohibited act. This mental state is called *mens rea,* or guilty mind, which generally requires either intent or recklessness.

Without evidence of *mens rea,* the law cannot hold a person criminally responsible. For example, hitting and killing a pedestrian when one's car goes out of control does not entail *mens rea,* whereas aiming one's car at a pedestrian, stepping on the gas, and running over that person does. The question for brain imaging is whether it is able to translate information about brain function into the legal requirements for criminal responsibility. The answer is that brain scans cannot shed significant light on this question. Thus, lawyers, jurors, and judges will still need to rely on traditional methods of assessing the defendant: interviews, observations, witness reports, psychiatric history, and well-established clinical assessments. It is from these methods that a subtler, necessary appreciation of the defendant's mental state can be inferred.

NEUROCENTRISM AND FREE WILL

When defense attorneys want to convince the judge and jury that their

clients are less worthy of blame or punishment, they try to make the case that the defendant's capacity to reason or to control himself is defective. And because his brain is defective, he made a bad choice, a choice that was unavoidable, and so he should not be held responsible. But what about the argument that no one is truly responsible? What are the implications of brain science for morality? Although individuals generally think of themselves as free agents who make choices, a number of prominent scholars claim that they are mistaken. "Our growing knowledge about the brain makes the notions of volition, culpability, and, ultimately, the very premise of the criminal justice system, deeply suspect," contends Stanford biologist Robert Sapolsky.

To be sure, virtually everyone agrees that people can be held accountable only if they have freedom of choice. No one "blames" the weather, inanimate objects, the severely mentally ill (who, depending upon degree of cognitive distortion, may be excused as not guilty by reason of insanity), the very young, or the demented. But what about the rest of humanity? If otherwise healthy individuals can think coherently, and plan, and change their minds in response to information, should not that render them free enough to be blamed (or praised, for that matter)? Philosophical "compatibilists" would say yes.

The compatibilist view does not negate material explanations of the world. They concede that behavior and thoughts are the products of a vast array of influences—genes (and the evolutionary history they represent), brain mechanisms, personal upbringing, and the physical and social environments in which people live. They further acknowledge that these forces converge, presumably, to produce one and only one specific act, be it "choosing" soup over salad or murder over mercy. Yet, under this determined view of the universe, people who nonetheless possess the capacities to be guided by reason, to follow rules, deliberate, be deterred, and to conform to norms are considered by philosophers to have freedom to choose—and, thus, to be accountable and perhaps deserving of blame and punishment.

Others, like Sapolsky, disagree. They insist that individuals' deliberations and decisions do not make them free to choose because they are dictated by neuronal circumstances. They say that, as people come to understand the mechanical workings of their brains, they will be compelled by logic to adopt a strictly utilitarian model of justice in which criminals are "punished" solely as a way to change their behavior, not because they deserve blame. Philosophers call this a "hard determinist" view of free will, and some believe that neuroscience can demonstrate that it is the correct understanding of free will.

Philosopher–neuroscientist Joshua Greene and psychologist Jonathan Cohen contend that neuroscience has a special role to play in giving the age-old argument about free will more rhetorical bite. "New neuroscience

will affect the way we view the law, not by furnishing us with new ideas or arguments about the nature of human action, but by breathing new life into old ones," they write. "[It] can help us see that all behavior is mechanical, that all behavior is produced by chains of physical events that ultimately reach back to forces beyond the agent's control," they added. The compatibilist view is most common among the public and even among philosophers. But so what if most people think that they and others have "free will"? Now comes the heart of the matter: The question of whether humans can live in a material world and yet be morally responsible is not empirically testable. It is not a scientific problem. It is a conceptual and ethical impasse that has bedeviled thinkers since antiquity and is still without a resolution.

What is important to establish is that neuroscience has not resolved it either. Those who believe that the absence of ultimate free will means that moral responsibility is an incoherent notion and that society should therefore abolish blame have already staked out a philosophical claim. Simply amassing more data on the workings of the brain may strengthen their conviction that determinism is correct, but it will not make their argument against the coexistence of determinism and moral agency any stronger.

The hard determinist view—a world without "just desert"—raises a set of interesting questions regarding how the criminal justice system should work. Most momentously, are humans even capable of removing blame from the moral and legal domain? Could it somehow be pried loose from entrenched notions of how human beings operate? The prospect of a world without moral responsibility collides head-on with the innate sense that people can choose freely. By the age of five, most children perceive others' actions in terms of their intentionality and agency. In one experiment, kindergartners observed an investigator sliding open the lid of a box, putting her hand in, and touching the bottom of the box. Asked by the experimenter if she had to touch the bottom or whether she "could have done something else instead," the vast majority of children believed that she could have done something else. But when the experimenter placed a ball on the lid and slid open the lid, and the ball fell to the bottom of the box, only a few of the children said that the ball itself could have "done something else instead."

The intuition that people can do something else persists with age. Adults across cultures, religions, and countries steadfastly reject the idea that the decisions are fixed in such a way that no other actions are possible. Just as consistently, human beings are highly attuned to notions of fairness. Communities buzz with gossip about who did what scandalous thing to whom; people keep score ("I did two favors for you, but you did only one for me") and punish cheaters who fail to reciprocate. Indigenous people in the Amazonian rain forest are as diligent about detecting free

riders and are as ready to punish offenders as are college students in the United States, Europe, and Asia. In his 1991 book *Human Universals,* anthropologist Donald E. Brown presented a comprehensive survey of moral concepts shared across the globe. Among them are proscriptions against murder and rape, as well as redress of wrongs.

In an exhaustive review of anthropological data, psychologists found that within all human cultures, individuals tend to react with quick, automatic feelings of anger, contempt, and indignation when they see people causing suffering. The universality of these responses raises the distinct possibility that intuitions about fairness, harm, and respect for authority are innate to the human mind. The behavioral economics laboratory is a showcase for fairness attitudes as well. Nobel Prize–winning psychologist and behavioral economist Daniel Kahneman and colleagues documented how study participants voluntarily incurred a penalty to punish others for unfair behavior. It is even more striking that they found evidence that people will pay to punish a third person even when participants are not directly involved in the transaction if they perceive that the third person's unfair behavior was intentional.

The stirrings of moral sentiment begin early on. In a series of studies, psychologists Karen Wynn and Paul Bloom of Yale University found that toddlers who watched a puppet show in which one puppet "stole" a ball and another puppet returned the ball to its rightful owner were far more likely to give candy to the helper puppet than to the "bad" puppet and preferentially take it away from the "bad" puppet. More generally, violations of fairness arouse emotions that trigger the urge to retaliate, especially when the violator intentionally injures an innocent person.

Social psychologist Philip Tetlock of the University of Pennsylvania and his collaborators presented subjects with a scenario involving a brutal assault that left the victim permanently brain damaged. No matter what happened to the aggressor—in one hypothetical scenario he suffered a painful accident; in another he was cured with medication—subjects' desire for justice was unchanged. Merely rendering him harmless or causing pain accidentally was not sufficient. This is why certain nations, such as Israel, direct their law enforcement officials to pursue octogenarian Nazi war criminals living quietly in South America and elsewhere even though they no longer pose a danger. The deep-seated desire for justice compels that the moral ledger be balanced.

Intrinsic to the idea of retribution is that people must suffer in proportion to the suffering they inflicted. Psychologist Jonathan Haidt of New York University and colleagues showed clips from Hollywood films that portrayed injustice (one involving the rape and murder of a child, and another in which a slave's foot is mutilated by a slave catcher). They next gave subjects a variety of endings and asked which one was the most "satisfying." Among the alternative endings was the "revenge" option: The

grieving mother violently kills her daughter's rapist; the hobbled slave chops off part of the foot of the man who mutilated him. In the "catharsis" option, the mother undergoes "primal-scream" therapy; the slave chops wood while visualizing the slave catcher's foot.

In the "forgiveness" ending, the victims join a support group or become more active in church and learn to forgive the transgression that was committed. The viewers derived far less satisfaction from the scenarios in which the victims come to terms with their tragedies and forgive the transgressor. They wanted the perpetrators to pay. It was most satisfying when the punishment matched the crime. At the same time, viewers found gratuitous and less satisfying another ending in which the slave retaliated by murdering his catcher.

This result is consistent with the extensive work of psychologists, the late Kevin M. Carlsmith and John M. Darley of Princeton University. In a series of experiments, they found that when it came to sentencing wrongdoers, subjects were exquisitely sensitive to the severity of the offense—say, the theft of $100 to feed a starving child versus using the money to create the world's largest margarita—and largely ignored the likelihood that the person would offend again. Subjects punished people exclusively in proportion to the harm done and not for the harm that might be done in the future. "People want punishment to incapacitate and to deter, but their sense of justice requires sentences proportional to the moral severity of the crime," they concluded. Punishment that violated proportionality offended people's intuitive sense of fairness.

Although hard determinists reject the concept of blameworthiness, they still accept the view that censure can have practical value. As any parent knows, well-calibrated disapproval, along with encouragement, is essential to raising children who are sensitive to the rights of others, behave kindly toward the injured and vulnerable, and reciprocate when they have been helped. No society, whether modern or preliterate, can function and cohere unless its citizens exist within a system of personal accountability that stigmatizes some actions and praises others. Punishment signals to the community who should not be trusted, and its severity reflects the enormity of the transgression. But none of these punishment-related functions requires that anyone be blamed. In utilitarian fashion, punishment could be meted out simply to shape future behavior.

Retribution, however, is a thoroughly different concept. In its pristine, theoretical form, retributive practices are triggered by the simple fact that someone who is mentally competent and uncoerced has committed an offense. The point of punishment is to make perpetrators suffer in proportion to the harm that they have already caused the victim and society. Any incidental benefit to the greater society that comes from reinforcing norms or protecting against future crime is irrelevant.

But when retribution is applied in the real world, it inescapably carries

great practical value, too. For one thing, it strengthens people's shared norms of moral obligation to one another. One of those norms is that victims should be valued as human beings. Consider the following vignette: A serial rapist, John, attacks Mary, is found guilty, and is sent to prison. A few months later, John is treated with "Castrex," a fictional new anti-rape medication guaranteed to permanently eradicate sexually aggressive urges. Castrex works after just a few doses, and several weeks later, John is freed. He is no longer a danger to anyone. His rehabilitation was a success. But John's light punishment would almost surely have woeful repercussions for Mary, her family, and the community at large.

When society fails to condemn aggressors or simply slaps them on the wrist, victims often feel unavenged and therefore devalued and dishonored. If perpetrators die before they can be judged or are killed in prison before they can be adequately punished, victims and their families usually feel enraged. Criminals who do not "pay their debt" can spur victims and their families to contemplate private retaliation and sometimes even undertake it. Contrary to common perception, such feelings do not always arise out of a sense of white-hot rage or bloodlust vigilantism. The motivation to give wrongdoers what they deserve can instead be motivated by grief or by a solemn sense of duty to set things right.

Communities, too, resist what they perceive to be inadequate punishment. Consider the 2011 trial of Casey Anthony, the twenty-five-year old Florida woman who never reported her two-year-old daughter missing. Widely believed to have killed her child or at least to have abetted in her death, Anthony received death threats after she was exonerated for the crime of murder. Likewise, jurists often speak of their moral duty to satisfy the victims and their families. When a U.S. district court judge sentenced disgraced New York City financier Bernard Madoff, who swindled thousands of investors of billions of dollars, to a term of 150 years, he explained to the press that the exceptionally long sentence for an elderly man who would probably die within a decade or so was a symbolic way to help the victims heal.

Restoring the social standing of the victim is another vital function of retribution. When the law imposes inadequate punishment on wrongdoers or allows them to escape penalties altogether, the demoralizing message to the community is clear: The victim is so inconsequential that his or her rights, security, and property can be breached with impunity. This is why proportionate punishment must be carried out in full view of the community in open trials, announced in the press, and preceded by a clear message of society's moral condemnation. Everyone in the community must be reminded that such mistreatment will not be tolerated. In a series of experiments, legal scholar Kenworthey Bilz, now of the Illinois University College of Law, asked subjects to evaluate the moral worth of a female victim whose attacker was punished for rape and a similar woman whose rapists pled guilty to a lesser offense. Both women were judged by participants before their rapists were punished. When participants learned that the rapists were convicted for rape, they rated the victim as more "respected," "valued," and "admired" than they were before punishment.

When the rapists pled to a lesser nonsexual crime, subjects ranked the victim's social standing lower than in the prepunishment phase.

What happens when the community is prevented from administering justice? In the mid-1960s, sociologist Melvin Lerner developed the "belief in a just world" hypothesis. According to this influential hypothesis, people harbor a strong need to believe that the world is a place where people get what they deserve, and where actions have predictable consequences. A just-world belief is akin to a "contract" with the world regarding the outcomes of one's conduct: do the right thing and be rewarded—or, at least, know largely what to expect. Lerner understood that the notion of a truly just world is illusory, yet he posited that it helps people to plan their lives and achieve their goals.

In one of his seminal experiments, Lerner asked subjects to observe a 10-minute video of a fellow student as she underwent a learning experiment involving memory. The student was strapped into an apparatus sprouting electrode leads and allegedly received a painful shock whenever she answered a question incorrectly (she was not receiving real shocks, but believably feigned distress as if she were). Next, the researchers randomly assigned the observers into groups. One was to vote on whether to remove the victim from the apparatus and reward her with money for correct answers. All but one voted to rescue her. The experimenters told another group of observers that the victim would continue to receive painful shocks; there was no option for compensation. When asked to evaluate the victim at this point, participants in the victim-compensated condition rated her more favorably (e.g., more "attractive," more "admirable") than did subjects in the victim-uncompensated condition, in which the victim's suffering was greater. Lerner's conclusion is disturbing. "The sight of an innocent person suffering without possibility of reward or compensation motivated people to devalue the attractiveness of the victim in order to bring about a more appropriate fit between her fate and her character," he wrote. In other words, when subjects' intuitions of justice are satisfied, their belief in a just world is supported. But when subjects (or society) are prevented from restoring justice, they typically blame the victim. Somehow, the reasoning goes, she must have asked for it.

Beyond the victims, the legal system itself suffers when criminals are not punished, although not just any punishment will do. The penalty must seem proportionate to the offense. If it is perceived as skewed in either direction—too lenient or too harsh—the law can lose its moral force. Recall Rodney King, the man who led California police on a high-speed chase before being caught and beaten severely. Shortly after the 1992 acquittal of the police officers involved in the assault, half of all Californians surveyed said that they had lost confidence in the court system. In a system perceived as unfair, juries may ignore judges' instructions, and police officers may impose their own judgment on whether to arrest, trump up

charges, or abuse suspects.

For their part, witnesses may rebel by refusing to participate in investigations or to testify. Researchers have observed that subjects are more willing to commit minor offenses such as traffic violations, petty theft, and copyright violations when laws more generally do not conform to their sense of right and wrong. Jurors are more susceptible to nullification —that is, to acquitting defendants who are legally guilty—when the verdict dictated by law is contrary to their sense of justice, morality, or fairness. Examples of perceived unfairness include suspicions of judges' failure to admit evidence, prosecutorial suppression of evidence, police misconduct, or police lying on the stand.

Finally, intuitions about justice are strong motivators for social change. The victims' rights movement grew out of a powerful sense of unfinished business. Unmet justice alienates juries, judges, and prosecutors, thereby undermining the authority of the law, an institution that depends for its smooth functioning on the good faith of its participants. Pragmatic efforts aimed solely at reducing future crime (and not at all on exacting retribution for past bad deeds) would leave many victims questioning the moral credibility of the law. Those victims do not necessarily seek harsh treatment for their violators although they surely want judges to administer appropriate punishment. What they want most is for the law to acknowledge that what happened to them was wrong and morally offensive. Indeed, the idea of a victim impact statement was developed so that the sentencing judge could hear directly about the anguish that victims and their loved ones had suffered. Some victims want an apology from their violators—not in place of a penalty, but in addition to it.

In summary, some scientists say that they look forward to a day when neuroscience will explain the mechanical workings of the brain so thoroughly that it will be impossible for society to continue to ignore the "fact" that people do not choose their behavior and therefore should be absolved of blame. This vision seems to be on a collision course with the psychological and social meanings that retribution holds for victims, their loved ones, and societies. Victims of crime are exquisitely attuned to society's regard for the people who have violated their rights. Unless authorities mete out deserved punishment, victims feel devalued and suffer a loss of moral standing in their community. Society is also harmed when citizens do not see justice being served and thereby lose faith in the moral authority of the law.

Although cloaked in neuroscientific garb, this free-will question remains one of the great conceptual impasses of all time, far beyond the capacity of brain science alone to resolve. Unless, that is, investigators can show something truly spectacular: that people are *not* conscious beings whose actions flow from reasons and who are responsive to reason. True, people do not exert as much conscious control over their actions as they

think they do. Every student of the mind, including William James and Sigmund Freud, has been aware of this important truism. But it does not mean humankind is powerless.

CONCLUSION

The brain is said to be the final scientific frontier, and perhaps with some justification. Yet in many quarters brain-based explanations appear to be granted a kind of inherent superiority over all other ways of accounting for human behavior—they have exercised the "neurocentric" presumption, or the view that human experience and behavior can be best explained from the predominant or even exclusive perspective of the brain. From this popular vantage point, the study of the brain is inherently more "scientific" than the study of human motives, thoughts, feelings, and actions. By making the hidden visible, functional brain imaging has been a spectacular boon to neurocentrism. Such imaging has contributed substantially to the knowledge of the correlates of mental illness and personality. The risk is losing sight of the mind in the age of brain science. Although the scans are dazzling and the technology an unqualified marvel, it is important to remember that the brain and the mind, although not differing materially, are two complementary but different frameworks for conceptualizing human behavior.

The neurobiological domain is one of brains and physical causes, the mechanisms behind one's thoughts and emotions. The psychological domain, the realm of the mind, is one of people—their desires, intentions, ideals, and anxieties. Both are essential to a full understanding of why people act as they do.

REFERENCES

*Bloom P. Seduced by the flickering lights of the brain. *Seed Magazine*. 2006.

Bourgois P. *In Search of Respect: Selling Crack in El Barrio*. Cambridge, MA: Cambridge University Press; 2002.

Cohen P, Cohen J. The clinician's illusion. *Arch Gen Psychiatry*. 1984;14(12):1178–1182.

Dennett DC. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press; 1984.

DuPont RL, McLellan AT, White WL, Merlo LJ, Gold MS. Setting the standard for recovery: Physician's health programs. *J Subst Abuse Treat*. 2009;36(2):159–171.

Eagleman DM, Correro MA, Singh J. Why neuroscience matters for rational drug policy. *Minn J L Sci Tech*. 2010;11(1):7.

Greene J, Cohen J. For the law, neuroscience changes everything and nothing. *Trans R Soc Lond B Biol Sci*. 2004;59:1775–1785.

Hawken A. Behavioral triage: A new model for identifying and treating substance-abusing offenders. *J Drug Policy Anal*. 2010;3(1):1.

*Heyman GM. *Addiction: A Disorder of Choice*. Cambridge, MA: Harvard University; 2008.

Higgins ST, Silverman K, Heil SH, eds. *Contingency Management in Substance Abuse Treatment*. New York: Guilford Press; 2008.

Interlandi J. What addicts need. *Newsweek*. 2008.

Kagan J. *An Argument for the Mind*. New Haven, CT: Yale University Press; 2006.

Koob GF, Volkow ND. Neurocircuitry of addiction. *Neuropsychopharmacology*. 2010;35(4):217–238.

*Morse SJ, Roskies AL, eds. *A Primer on Criminal Law and Neuroscience: A Contribution of the Law and Neuroscience Project*. New York: Oxford University Press; 2013.

Richards JR. *Human Nature After Darwin: A Philosophical Introduction*. London: Routledge; 2000.

*Satel S, Lilienfeld S. *Brainwashed: The Seductive Appeal of Mindless Neuroscience*. New York: Basic; 2013.

*Schelling TC. The intimate contest for self-command. *Public Interest*. 1980;60:94.

▲ 4.3 Transcultural Psychiatry

ROBERT KOHN, M.D., RONALD M. WINTROB, M.D., AND RENATO D. ALARCÓN, M.D.

THE CULTURAL ASSESSMENT OF THE PATIENT

Culture, race, and ethnicity are concepts of crucial importance in the clinical assessment of every patient. *Culture* is defined as a set of meanings, norms, beliefs, values, and behavior patterns shared by a group of people. These values include social relationships, language, nonverbal expression of thoughts and emotions, moral and religious beliefs, rituals, technology, and economic beliefs and practices, among other items. Culture has six essential characteristics: (1) Culture is learned; (2) it can be passed on from one generation to the next; (3) culture involves a set of words, behaviors, events, and symbols whose meanings have been tacitly or objectively agreed upon by the cultural group; (4) culture acts as a template to shape and orient future behaviors and perspectives within and between generations and to take account of novel situations encountered by the group; (5) it exists and evolves in a constant state of change; and (6) it includes patterns of both subjective and objective components of human behavior. During the clinical encounter, culture shapes how and what psychiatric symptoms are expressed. Culture influences the meanings that are given to the symptoms, and impacts the interaction between the patient and his/her family and the physician and the health care system within which the latter operates.

Race is a concept that entails people being grouped primarily by physiognomy. Although the scientific validity of the concept is now considered highly questionable, its impact on individuals and groups, however, is intense, due to its reference to physical, biological, and genetic underpinnings, and because of the intensely emotional meanings and

responses it generates. *Ethnicity* refers to the subjective sense of belonging to a group of people with a common national or regional origin and shared beliefs, values, and practices, including religion. It is part of every person's identity and self-image.

CULTURE AND PSYCHIATRIC DIAGNOSIS

Culture influences what is viewed as a social, moral, spiritual, or specifically a *mental* health problem. The distinction between mind and body varies across cultures. Hence, the conceptualization and the experiencing of a mental disorder depend on cultural modes of perception, explanation, and causal attributions. Some of the disorders included in DSM-5 may be viewed as social, moral, or spiritual problems in other cultural contexts, and help may be sought from corresponding nonmedical practitioners and institutions.

Culture influences the process of diagnosis in many ways: the experience and expression of symptoms, signs, and behaviors that are criteria for diagnosis; the configuration of symptoms into culturally distinctive syndromes; the stigma attached to specific symptoms and syndromes that may prevent their clinical reporting and influence the subsequent course of illness; and the conduct of the clinical encounter through which diagnosis is made. As a result, cultural context and differences between the clinician and patient have implications for the accuracy of diagnosis as well as treatment decisions, prognostic considerations, treatment compliance, and clinical outcomes. Consideration of the social and cultural contexts of the patient's life and of the clinical encounter is essential for accurate psychiatric diagnosis.

Judgment that a symptom or behavior is abnormal and requires clinical attention depends on cultural norms that are internalized by the individual or applied by others around them. The boundaries between normality and pathology vary across cultures for specific types of symptoms and behaviors. Thresholds of tolerance for specific symptoms or behaviors differ also across cultures and social settings. Hence, the level at which an experience becomes problematic or pathological will differ. Diagnostic assessment must consider the degree to which the patient's experience and behavior differ from relevant social and cultural norms and generate conflict or difficulties in adaptation within his/her culture of origin and the current social context.

CULTURAL FORMULATION

Culture plays a role in all aspects of mental health and mental illness; therefore, a cultural assessment should be a component part of every complete psychiatric assessment. In the wake of significant criticisms of the cultural components of DSM-IV, the DSM-5 Committee determined to give culture a more prominent place in the fifth edition of the *Diagnostic and*

Statistical Manual of Mental Disorders (DSM-5). This resulted in a cultural positioning of the DSM-5 itself in the introductory chapter of the Manual, with a description of basic cultural concepts and culture-related diagnostic issues for each diagnostic category. In addition, DSM-5 includes a revision of the Cultural Formulation chapter in Section III. The text of the Outline for Cultural Formulation was revised, and the *Cultural Formulation Interview* (CFI) was developed, field tested in more than 300 patients in six countries around the world, and incorporated in DSM-5. The Appendix includes clinical conceptualizations of distress (definitions of cultural syndromes, idioms of distress and causal attributions) and the most studied cultural syndromes.

The outline for cultural formulation incorporated in the Cultural formulation of DSM-5 is intended to give clinicians a framework for assessing the role of culture in mental illness. Its purposes are (1) to enhance the application of DSM-5 diagnostic criteria in multicultural environments; (2) to provide a systematic review of the individual's cultural background; (3) to identify the role of the cultural context in the expression and evaluation of psychiatric symptoms and dysfunction; (4) to enable the clinician to systematically describe the patient's cultural and social reference groups and their relevance to clinical care; and (5) to identify the effect that cultural differences may have on the relationship between the patient and family and the treating clinician, as well as how such cultural differences affect the course and the outcome of treatment provided.

The Outline for Cultural Formulation consists of five areas of assessment: (1) cultural identity of the individual; (2) cultural conceptualizations of distress; (3) psychosocial stressors and cultural features of vulnerability and resilience; (4) cultural features of the relationship between the individual and the clinician; and (5) overall cultural assessment.

Cultural Identity of the Individual

Cultural identity refers to the characteristics shared by a person's cultural group. Identity allows for a self-definition. Factors that comprise an individual's cultural identity include race, ethnicity, country of origin, language use, religious and tradition-nourished beliefs, socioeconomic status, sexual orientation, migration history, experience of acculturation, and the degree of affiliation with the individual's group of origin. Cultural identity emerges throughout the individual's life and in social context. It is not a fixed trait of an individual or of the group of which the individual is part. An individual may have several cultural reference groups.

Evaluating the cultural identity of the patient allows identification of potential areas of strength and support that may enhance treatment effectiveness, as well as vulnerabilities that may interfere with the progress

of treatment. Eliciting these data permits identification of unresolved cultural conflicts that may be addressed during treatment. These conflicts can be between the various aspects of the patient's identity and between traditional and mainstream cultural values, as well as behavioral expectations affecting the individual. Knowledge of the patient's cultural identity allows the clinician to avoid misconceptions based on inadequate background information or stereotypes related to race, ethnicity, and other aspects of cultural identity. In addition, it assists in building rapport because the clinician is attempting to understand the individual as a person and not just as a representative of the cultural groups that have shaped the patient's identity.

Cultural Conceptualization of Distress

The conceptualizations of distress represent ways in which the individual patient "experiences, understands, and communicates his or her symptoms to others." Severity, meaning of the experiences, coping, and help-seeking patterns are part of this representation. Cultural conceptualization of distress also includes the patient's understanding of their symptoms and attempt to explain why he or she became ill. The explanatory model, which may include some cultural syndromes, defines the culturally acceptable means of expression of the symptoms of the illness or *idioms of distress,* the particular way individuals within a specific cultural group report symptoms and their behavioral response to them that are heavily influenced by cultural values. The cultural explanations of illness also may help define the sick role or behavior the patient assumes. The explanatory model of illness includes the patient's beliefs about their prognosis and the treatment options they would regard as being helpful. The patient's explanatory model may include several conceptual perspectives that could be in conflict with one another. Formulation of a collaborative model that is acceptable to both the clinician and the patient is the sought-for end point, which would include an agreed upon set of symptoms to be treated and an outline of treatment procedures to be used.

Difficulties may arise when there are conceptual differences in the explanatory model of illness between clinician, patient, family, and community. Conflicts between the patient's and the clinician's explanatory models may lead to diminished rapport or treatment noncompliance. Conflicts between the patient's and the family's explanatory models of illness may result in lack of support from the family and family discord. Conflicts between the patient's and the community's explanatory models could lead to social isolation and stigmatization of the patient.

Examples of the more common explanatory models of illness include the *moral model, religious model,* the *magical* or *supernatural explanatory model,* the *medical model,* and the *psychosocial stress model.* The *moral model* implies that the patient's illness is caused by a moral

defect such as selfishness or moral weakness. The *religious model* suggests that the patient is being punished for a religious failing or transgression. The *magical* or *supernatural explanatory model* may involve attributions of sorcery or witchcraft as being the cause of the symptoms. The *medical model* attributes the patient's illness primarily to a biological etiology. The *psychosocial model* infers that overwhelming psychosocial stressors cause or are primary contributors to the illness.

Culture has both direct and indirect effects on help-seeking behavior. In many cultural groups an individual and his or her family may minimize symptoms due to stigma associated with seeking assistance for mental disorders. Culture affects the patient's expectations of treatment, such as whether the clinician should assume an authoritarian, paternalistic, egalitarian, or nondirective demeanor in the treatment process.

Psychosocial Stressors and Cultural Features of Vulnerability and Resilience

An understanding of the patient's family dynamics, social networks, and religion in its cultural context is integral to assessing the patient's psychosocial environment and level of functioning, the exposure to real or perceived negative events and the individual and social resources to resist or overcome them. The definition of family includes an understanding of the patient's cultural group and its relationship to the mainstream culture or cultures, as well as life stressors such as racial and ethnic discrimination. For immigrants and refugees, it includes the individual's and family's perceptions of the openness of the host society toward people of their country and region of origin, their racial, ethnic, religious, and other attributes. The patient and family may identify strongly or weakly with communal sources of support familiar from their country or region of origin, or they may identify along the same gradient with communal sources of support in the host culture. For example, middle-aged and older immigrants may experience familial and communal role reversals and erosion of self-esteem in situations in which the older generation becomes dependent on the younger generation, due to the younger members' greater language fluency and conceptual ability to deal with the institutions of the host culture, such as hospitals and government agencies. Conversely, feelings of alienation may occur when young people in immigrant families feel cut off from their heritage and from the usual association of wisdom and experience with parents and other adults of their cultural group.

Cultural Features of the Relationship between the Individual and the Clinician

The cultural identity of the clinician and of the mental health team has an impact on patient care. The mental health care professional's culture influences diagnosis and treatment. Clinicians who have an understanding of their own cultural identity may be better prepared to anticipate the

cultural dynamics that may arise in interactions with people of diverse cultural backgrounds. Unacknowledged differences between the clinician's and the patient's cultural identity can result in assessment and treatment that is unintentionally biased and stressful for all. Experiences of racism and discrimination may impede establishing trust and a sense of safety in the clinical setting. Clinicians need to examine their assumptions about other cultures in order to be optimally effective in serving the culturally diverse patient populations that are now the norm in most contemporary medical facilities.

Culture influences transference and countertransference in the clinical relationship between people seeking psychiatric care and their treating clinicians. Transference relationships and dynamics are affected when the patient and clinician have different cultural background characteristics. A perceived social power differential between the patient and clinician could lead to overcompliance, to resistance in exploration of family and social conflict situations, or to the clinician being conceptualized as a cultural role model or stereotype.

When the patient and clinician are of different genders, culturally ingrained role assumptions may pose difficulties. For example, male patients from cultures where men are assumed to have higher status than women may feel that expressing their emotional problems to a female therapist is evidence of weakness and is culturally humiliating. Conversely, females may view it as culturally inappropriate to discuss with male clinicians interpersonal issues and emotions that are only considered proper to talk about with females of their age group and within the setting of their extended family.

Overall Cultural Assessment

This section summarizes the information gathered above and points out the most relevant aspects of the interview. The treatment plan should include the use of culturally appropriate health care and social services. Interventions also may be focused on the family and social levels. In making a psychiatric diagnosis, the clinician should take into account principles of cultural relativism and not fall prone to category fallacy as many mental disorders show cross-cultural variation. Objective evaluation of the multiple possible effects of culture on psychopathology can be a challenging task for the clinician. Diagnostic dilemmas may arise in dealing with patients of diverse cultural backgrounds. Some of these dilemmas may include problems in judging distortion from reality, problems in assessing unfamiliar behaviors, and problems in distinguishing pathological from normal cultural behavior.

Cultural Formulation Interview

The DSM-5 includes the *CFI* to assist in obtaining pertinent information about the impact of culture on the individual's clinical presentation, so that the subsequent care plans can be conceived and understood appropriately. The CFI is a semistructured interview that includes explicit instructions to the interviewer. It uses a person-centered approach to conduct a cultural assessment by eliciting information from the patient about his or her views and those of others in their family and social network, and thereby avoids stereotyping about the way that patients interpret their illness experience. The CFI is especially useful when (1) there are significant differences in the cultural, religious, or socioeconomic background of clinician and patient; (2) there is uncertainty about culturally distinctive symptoms, their meaning and significance to the patient and their relationship with existing, conventional diagnostic criteria; (3) there is difficulty judging illness severity or extent of impairment; (4) there is disagreement between patient and clinician about the recommended course of treatment; (5) there are concerns about the patient's acceptance, engagement, and adherence to recommended treatment.

The CFI explicitly recognizes cultural variation in symptom presentation. It recognizes cross-cultural and regional variation, as well as intracultural and intrafamilial variation in symptom presentation and prevalence. It recognizes and validates cultural assessment as a component of person-centered care in medicine and in psychiatry.

The CFI enables clinicians to operationalize concepts inherent in cultural case formulation. Interview questions can be readily applied for use with family members and other informants to provide sociocultural context. The CFI emphasizes collaborative, shared decision making concerning diagnosis and treatment between patients, families, community members, and clinicians.

The CFI comprises four sections that together include 16 questions. The four sections focus on (1) cultural definition of the presenting problem, (2) cultural perceptions of cause, context, and support available, (3) the role of cultural identity, and (4) cultural factors affecting coping and health-seeking practices, including misunderstandings in the clinician–patient relationship. In addition, the CFI includes an informant version that can be used to supplement information from the core CFI or when the patient is unable to provide sufficient information. Eleven supplementary modules can be used to help clinicians achieve a more comprehensive assessment of issues such as explanatory models of distress, an individual's level of functioning, psychosocial stressors, spirituality, religion and moral traditions, and cultural identity. Other supplementary modules focus on school-age children and adolescents, older adults, immigrants and refugees, and caregivers.

Glossary of Cultural Concepts of Distress

The Culture-bound Syndromes section of DSM-5 consists of three concepts that may assist the clinician in rounding up a culturally competent evaluation.

Cultural Syndrome defined as a cluster or group of cooccurring symptoms found in a specific cultural group, community, or context. The syndrome may or may not be recognized as an illness within the culture (e.g., it might be labeled differently) but may nevertheless occur and be recognized by an outside observer. DSM-5 includes in its Appendix, nine more or less well-delineated cultural syndromes (*Ataque de nervios, Dhat, syndrome, Khyâl cap, Kufungisisa, Maladi moun, Nervios, Shenjing shuairuo, Susto* and *Tajin kyofusho*), that are connected to conditions in other cultural contexts and specific diagnoses in DSM-5.

Cultural Idiom of Distress is a linguistic term, phrase or way of talking about suffering, shared with other people from the same culture (i.e., ethnicity, religion, community) and used to express, communicate, or comment on distress in general. An idiom of distress need not be associated with specific symptoms, syndromes, or causal explanations. It may be used to convey a wide range of uncomfortable, emotional pain or social shakiness including subclinical conditions or everyday experiences that do not necessarily constitute mental disorders.

Causal attribution is a label, an attempt at explaining or ascertaining the causes of the symptoms, illness, or distress. Causal explanations may be part of folk classifications of disease used by laypeople or healers, which, in a good number of cases, may provide temporary relief.

These three concepts are culturally based ways of understanding and describing illness experience that can be elicited in the clinical encounter. They influence symptomatology, help-seeking, clinical presentation, illness adaptation, and treatment response. The same term can serve more than one of these functions. Determining whether a mention of a term represents a syndrome, an idiom of distress, and/or an explanation depends on the information about the patient's cultural background and the cultural context of the terms' use throughout the interview.

MIGRATION, ACCULTURATION, AND ACCULTURATIVE STRESS

From the time of the first major surge of immigration to the United States in the 1870s, and for the next 100 years, the predominant national sentiment toward immigrants, as in most other host countries, was that they should accept and adapt to the normative behaviors and values of the majority or mainstream culture of the host population. Most immigrants had the same wish to assimilate, to become part of the *melting pot*. This process of acculturative changes can be seen as unidirectional when individuals who identified themselves as part of immigrant, indigenous, and other minority groups both rejected and progressively lost distinctive aspects of their cultural heritage, in favor of becoming part of the

mainstream majority culture of the host country. In countries that encouraged this outcome of acculturation, people were expected to progress from unacculturated, through the gradient of minimally, moderately, and fully acculturated. However, this model of acculturation generally overlooked the issues of prejudice and discrimination within the majority host population that stood in opposition to immigration and strongly resisted the cultural assimilation of immigrant, indigenous, and other minority groups. In reaction to such negative pressures, groups that were discriminated against often formed, or were forced into, communities of their own for mutual support and protection. Those ghettoized communities generated fear and resentment, both within the isolated community and in the perceptions and behaviors of the external community toward the minority groups.

The intensity of acculturative stress experienced by immigrant and other minority groups, and the individuals comprising those groups, has been directly proportional to the openness of the host government and population. The central issue is to what extent are immigrants' and other minority groups' customs, values, and differences from the majority population of the host country accepted, encouraged, and welcomed as an enrichment of the cultural life of the host country, as opposed to being seen as alien and unwelcome? The *acceptance position* encourages the cultural integration of immigrants, whereas the *rejection position* encourages either cultural exclusion or cultural assimilation.

During the past 50 years the societal ideals of cultural diversity and multiculturalism have favored cultural integration rather than cultural assimilation. As the societal ideals of cultural diversity and multiculturalism grew in acceptance and influence on public policy in the United States and other host countries for migrants, acculturation has been increasingly conceptualized as a nonlinear, complex, and ongoing process.

The process of acculturation and the psychological effects of acculturative stress on communities and on individuals have been increasingly recognized as affecting the majority population, as well as the immigrant, indigenous, and other minority groups that, together, comprise the changing national population. From this perspective, the acculturation process is conceptualized as progressive and dynamic, continuing over several generations. It involves acquisition and retention, as well as relinquishing of values, thought patterns, and social behaviors of both majority and minority population groups as their interactions are modified by cumulative life experience and over time.

Acculturation and the outcomes of acculturative stress for groups, as well as for the individuals who comprise those groups, are both social and psychological phenomena. The outcome of acculturative stress is not a finite end point, but rather a continuous process, with periods of greatly intensified intrafamilial, intergenerational, and individual intrapsychic

stress alternating with periods of comparative calm, insight, and successful adaptation to unanticipated change.

In order to assess the outcome of acculturative stress, for groups and their component individuals, two determining factors need to be considered. The first is the extent to which the group and its members value and wish to preserve their cultural uniqueness, including the language, beliefs, values, and social behaviors that define the group. The second factor is the mirror-image issue of the extent to which the group and its members value and wish to increase their contact and involvement with other groups, particularly the majority or mainstream culture. This conceptual framework leads to four possible *outcomes of acculturative stress* that are not conceptualized along the unidirectional gradient from unacculturated to completely acculturated.

The four possible outcomes of acculturative stress are *separation, integration, assimilation,* and *marginalization. Separation* is characterized by individuals' wishes, both conscious and intuitive, to maintain their cultural integrity, whether by actively resisting the incorporation of the values and social behavior patterns of another cultural group or groups with whom they have regular contact, or by disengaging themselves from contact with and the influence of those other cultural groups. Some religious cults are examples of separation.

Integration, as an outcome of acculturative stress, derives from the wish to both maintain a firm sense of one's cultural heritage and not abandon those values and behavioral characteristics that define the uniqueness of one's culture of origin. At the same time, such individuals are able to incorporate enough of the value system and norms of behavior of the other cultural group with which they interact closely, to feel and behave like members of that cultural group, principally the majority host culture. Accordingly, the defining feature of integration is psychological: It is the gradual process of formulation of a bicultural identity, a sense of self that intertwines the unique characteristics of two cultures. Examples are found among the large number of *hyphenated Americans:* Italian-Americans, Jewish-Americans, Mexican-Americans, Japanese-Americans, Haitian-Americans, Arab-Americans, and many other people who define their sense of self in terms of belonging simultaneously to two distinctly different cultural groups, and sometimes more than two. Psychological integration of two cultural traditions is neither easy to accomplish nor conflict-free. It involves a continuous inner, subjective struggle to balance inherently conflicting components of a bicultural identity. That is, the outcome of acculturative stress for any individual is shaped by the particular intrapsychic conflicts and coping abilities of that individual, factors that account for the great intragroup variation among members of any cultural group during the process of acculturation over decades and generations.

Assimilation is the psychological process of the conscious and

unconscious giving up of the unique characteristics of one's culture of origin in favor of the more or less complete incorporation of the values and behavioral characteristics of another cultural group, usually, but not always, the majority culture. Examples include involuntary migration, when war and social upheaval necessitate such changes for purposes of survival. However, there are many other life circumstances, including racial, ethnic, and religious discrimination, that motivate people to overlook, suppress, or deny aspects of their cultural heritage in an attempt to have a seamless fit within another group. The price of such an effort, in terms of intrapsychic conflict, can be high.

Marginalization is defined by the psychological characteristics of rejection or the progressive loss of valuation of one's cultural heritage, while at the same time rejecting, or being alienated from, the defining values and behavioral norms of another cultural group, usually that of the majority or mainstream population. This is the psychological outcome of acculturative stress that is closest to the concept of identity diffusion. As such, it is most often exemplified by the angry, lost, and anguished youth and young adults of many groups, those whose intense intrapsychic conflicts are reflections of substantive intrafamilial, intergenerational, intracommunal, and intercommunal conflict. Part of their search for psychological meaning and self-esteem is reflected in their turmoil about their ethnic identity and in their formation of a negative identity.

The literature on acculturation and acculturative stress emphasizes the need for long-term study of the process. The outcome of acculturative stress cannot be determined by cross-sectional, short-term analysis. The process is continuous throughout the life cycle and is strongly influenced by life events over which individuals, families, and cultural groups may have little control. At the same time, understanding the theoretical underpinning of acculturative stress and its possible outcomes can enable clinicians to take account of its complex influence on the clinical presentation of the very large numbers of people affected by it and thereby improve the quality of their treatment and the likelihood of compliance through the course of treatment.

The rate of acculturative change and the circumstances that influence it vary greatly both between and within groups. For these reasons, studies of groups experiencing acculturative change often divide the groups between first-, second-, and third-generation immigrants. Families within such groups have been categorized as *traditional, transitional,* or *bicultural. Traditional families* are defined largely on the basis of intrafamilial use of preimmigration language, residence in ethnic enclaves, resistance to or exclusion from interaction with majority cultural institutions, and maintenance of preimmigration thought patterns, values, and social behaviors. *Transitional families* are characterized by greater fluency in the language of the host culture, by children becoming familiar with the values

and social behaviors of the majority culture through their attendance at school and participation in school-related activities. *Bicultural families* are those with a high degree of language fluency and economic stability, living in multiethnic communities and having a parental authority structure that is more egalitarian than patriarchal.

Finding ways to operationalize the concept of bicultural families and bicultural individuals has been an increasing focus of research. Scales have been developed to measure the nature and extent of individuals' identification with their culture of origin and with the majority or host culture. This has led to current efforts to describe and validate the concept of *multicultural competence* (that some simply call *cultural competence*) including the cognitive, emotional, and behavioral characteristics of people who are capable of feeling comfortable and functioning effectively in a variety of distinct cultural contexts. The more complex issue of understanding the process by which such individuals integrate the elements of both cultural traditions into a psychologically consistent sense of self remains to be elucidated.

One approach has been the development of a framework for investigating individual differences in *bicultural identity integration,* focusing on bicultural individuals' subjective perceptions of how much their dual cultural identities intersect or overlap. Bicultural identity organization refers to the degree to which biculturally competent people perceive their mainstream and ethnic cultural identities as compatible and integrated versus oppositional and difficult to integrate. Individuals high on bicultural identity integration tend to see themselves as part of a *hyphenated culture,* or as part of an emerging third culture, and do not find it emotionally stressful to integrate both cultures in their everyday lives. These *high bicultural identity integration* individuals are described as having developed compatible bicultural identities, meaning that they do not perceive the two cultures to be mutually exclusive, oppositional, or conflicting. On the other hand, people with *low bicultural identity integration* report difficulty in incorporating both cultures into a cohesive sense of self. Although low bicultural identity integration individuals also identify with both cultures, they are particularly sensitive to specific tensions between the two cultural orientations and see this incompatibility as a source of ongoing and unresolved intrapsychic conflict.

The political ideal of *multiculturalism* or cultural diversity has been adopted in a large number of host countries. There has also been serious consideration of the concept of *binational* as well as *bicultural identity:* the notion that individuals can be active participants in and feel equally involved in the cultural, economic, community, and even the political and religious life of two countries, or of two distinctly and uniquely different cultures in one country. This has been the subject of discussions and studies among indigenous individuals in countries throughout North and

South America, as well as among the aboriginal population of Australia and the Maori people of New Zealand.

PSYCHIATRIC ASSESSMENT OF IMMIGRANTS AND REFUGEES

Migration History

Mental illness among immigrants and refugees may have been present before migration, may have developed during the immigration process, such as during months or years living in refugee camps, or presented for the first time in the country of immigration. The immigration process and premigration trauma may precipitate the manifestation of underlying symptoms or result in exacerbation of a pre-existing disorder. Obtaining a thorough migration history, such as in the *DSM-5* CFI supplement, will assist in understanding background and precipitating stressors and help guide development of an appropriate treatment plan. The components of the migration history include the premigration health and mental health status, premigration planning, the migration experience, and the resettlement phase.

The premigration history includes inquiry about the patient's social support network, social and psychological functioning, and significant premigration life events. Information about the country and region of origin, the family history in the country of origin—including an understanding of family members who may have decided not to immigrate—educational and work experiences in the country of origin, and prior socioeconomic status should be obtained. Additionally, premigration political issues, trauma, war, and natural disaster faced by the patient and family in the country or region of origin should be explored. For those who had to escape persecution, warfare, or natural disaster, what were the means of escape and what type of trauma was suffered prior to and during migration? Traumatic life events are not limited just to refugees. Immigration may result in losses of social networks, including family and friends; material losses, such as business, career, and property; and loss of the cultural milieu, including their familiar community and religious life. Premigration planning includes reasons for immigrating, duration and extent of planning, premigration aspirations, and beliefs about the host country. The type of migration experience, whether as voluntary immigrants or as unprepared refugees, can have profoundly different effects on migrants' mental health. Table 4.3–1 provides a list of common premigration traumatic experiences among refugees and factors contributing to psychological distress.

A history concerning the migration experience should include the duration, difficulties, and dangers encountered in the migration process. For example, refugees may spend years in camps where they are exposed to additional psychological disorientation, distress, and trauma. Their

traumatic experiences could include exposure, forced labor, torture, rape, starvation, and imprisonment. It could also involve psychological trauma such as prolonged and enforced dependency, exploitation, isolation, overt discrimination, and loss of hope.

The resettlement phase raises issues of acculturative stress and adaptation to the host country institutions and ways of doing things. Immigrants often start their new lives in host countries living in unfamiliar types of housing, in crowded and sometimes dangerous neighborhoods, unfamiliar with social cues and sources of support that could make their adaptation less stressful. Immigrants in the host country not only find themselves in a minority status, but often face prejudice and discrimination and their psychological consequences. The clinician needs to assess the extent to which the immigrant's premigration aspirations and expectations regarding their life in the host country are being realized or frustrated.

Table 4.3–1.
Refugee Premigration Experiences and Risk Factors for Psychopathology

Common Premigration Traumatic Experiences of Refugees	Common Risk Factors for Psychological Distress among Refugees
War and deliberate killings	Past experiences of traumatic events
Genocide	Premigration health problems
Murder of family members/friends/relatives	Lack of accommodation/overcrowding
Witnessing the murder of family members/friends	Isolation/loneliness
Prolonged mental and physical torture	Lack of supportive social networks
Rape of women and girls	Lack of supportive friends
Kidnap of children and women	Family separation
Protracted imprisonment and/or forced labor	Unemployment
Cruel amputations	Lack of access to education
Food shortage/starvation	Language problems
Lack of water	Barriers to culturally appropriate health and social services
Looting and daily robbery	Legal insecurities
Destruction of personal property	Cultural shock and adjustment problems
Destruction of public infrastructures	Substance misuse
	Experiences of racism and discrimination
	Persistent poverty

Adapted from Warfa N, Bhui K. Refugees and mental health. In: Bhugra D, Bhui K, eds. *Textbook of Cultural Psychiatry.* Cambridge: Cambridge University Press; 2007:505.

The Mental Status Examination

As with any patient, conducting a mental status examination is a central component of the psychiatric examination. However, its interpretation in culturally distinct groups and among immigrant populations requires caution, as it may be culturally biased. The patient's response is molded by

his or her culture of origin, educational level, and type of acculturative adaptation. The components of the standardized mental status examination are cooperation, appearance and behavior, speech, affect, thought process, thought content, cognition, insight, and judgment. Cultural differences are wide and varied in dress and grooming. Facial expressions and body movements used in the expression of affect may be more reflective of normal cultural manifestations than pathology. If the clinician is unfamiliar with the individual's culture and the patient's fluency in the language of the host country is limited, the clinician must use caution in interpreting disturbances of speech and thought process, perception, and affect. The presence of hallucinations, for example, can be easily misinterpreted, such as hearing encouraging or clarifying comments from deceased family members, which are normative experiences in many cultures. The clinician should not assume that the patient understands what the clinician is trying to communicate, and miscommunication involving use of interpreters is a common problem. The cognitive examination may be particularly tricky. Education and literacy have an important and biasing role. The patient may need adequate time to fully express him or herself through repeating questions and restating questions in the effort to reduce miscommunication. Asking about the meaning of proverbs unfamiliar to the patient may be an inappropriate means of determining abstract thinking. An accurate mental status examination can be accomplished when one allows additional time for clarification of the cultural equivalence of concepts.

THE CHANGING U.S. POPULATION: RACE, ETHNICITY, AND IMMIGRATION

Knowledge of population demographics enhances clinicians' ability to assess the cultural characteristics of their patients. This enables them to provide appropriate and culturally sensitive diagnosis and treatment for the culturally diverse population they are inevitably going to encounter in their everyday clinical experience.

U.S. Population Growth: 1850 to 2015

Using prior, current, and future estimates from U.S. census through the 19th and 20th centuries the U.S. population increased from 23 million in 1850 to 50 million by 1880, and by 1900 to 76 million. In 1950 the population reached 150 million and 282 million by 2000. By 2015 the U.S. population exceeded 320 million.

Immigration: The Foreign-Born Component of the U.S. Population

Immigration surged in the United States between 1860 and 1930. During those decades, the foreign-born component of the total U.S. population increased from 4.1 to 14.2 million, representing 13.2 percent of the total

population in 1860 and 11.6 percent in 1930. The peak decade of immigration in the 20th century was the first decade. After 1910 the inflow of immigrants diminished sharply. The foreign-born component of the population decreased from 14.7 percent in 1910 to 11.6 percent in 1930. During the first half of the 20th century, two world wars and tighter legal restrictions on immigration resulted in a steady decline in the number of foreign-born Americans, reaching the low point for the 20th century of 4.7 percent in 1970 (Table 4.3–2).

During the last three decades of the 20th century, and continuing through the first two decades of the 21st century, there has been an increasing inflow of immigration to the world's most highly developed countries, particularly in North America and Europe. Between 1970 and 2013, the foreign-born component of the U.S. population grew from less than 10 million to more than 40 million, representing an increase from 4.7 to 13.3 percent of the total U.S. population. Viewed from the perspective of how long ago they immigrated, 37 percent of the total foreign-born component of the U.S. population came to the country between 2000 and 2012, 27 percent between 1990 and 1999, 18 percent between 1980 and 1990, 10 percent in the 1970s, and only 8 percent before 1970. Immigration numbers have continued to rise steadily since the first decade of the 21st century. The immigrant population is expected to rise to over 78 million by 2060, comprising some 18.2 percent of the projected U.S. population of over 400 million in 2060.

Until 1960 the overwhelming majority of immigrants to the United States came from European countries. In 1900, Europeans comprised 86 percent of the 10.3 million foreign-born component of the U.S. population. Between 1900 and 1970 there was a steady, although not dramatic, decline in the proportion of Europeans among the foreign-born population of the U.S., from 83 percent in 1930 to 60 percent in 1970. Since 1980, however, the demographic changes in the foreign-born population have been both rapid and dramatic. By 2013, Europeans comprised less than 12 percent of the 41 million foreign-born U.S. residents.

Since 1970 there has been a very dramatic increase in immigrants from Latin America, most notably from Mexico, from 9 percent of the total foreign-born population in 1960 to 52.9 percent in 2013. There has been a similar surge of immigration from the Asia/Pacific region, from 5 percent of the total foreign-born population in 1960 to 30.2 percent in 2013. By 2013 the 41.3 million foreign-born included 21.4 million from Latin America, 12.2 million from Asia, 4.8 million from Europe, and 1.8 million from Africa. This dramatic increase in Latin Americans has continued during the second decade of the 21st century, along with the continuing decline in percentage of Europeans among the foreign born.

Starting in the 1960s, opportunities for temporary migration increased steadily in many industrialized countries, including the United States, for

temporary workers, students, and technicians. The greatest numbers were for seasonal agricultural workers and factory workers, followed by temporary workers in the service industry, mainly hotels and restaurants, and building maintenance. Currently there is great competition for technical and professional workers. Between 1997 and 2013 the annual inflow of foreign-born temporary residents to the United States exceeded the number of permanent immigrants in each of those years. In 2013 nearly 3 million temporary workers came to the United States.

Table 4.3–2.
U.S. Population That Is Foreign Born per 1 Million and Percentage Distribution 1850 to 2013

Year	Total Population	US-Born Population	Foreign-Born Population	Percentage of Population Foreign Born
2013	311.5	271.1	41.3	13.3
2010	304.0	265.3	38.7	12.7
2000	282.2	252.2	30.0	10.6
1990	248.7	228.9	19.8	7.9
1980	226.5	212.5	14.1	6.2
1970	203.2	193.6	9.6	4.7
1960	179.3	169.6	9.7	5.4
1950	150.2	139.9	10.3	6.9
1940	131.7	120.1	11.6	8.8
1930	122.8	108.6	14.2	11.6
1920	105.7	91.8	13.9	13.2
1910	92.0	78.5	13.5	14.7
1900	76.0	65.7	10.3	13.6
1890	62.6	53.4	9.2	14.8
1880	50.2	43.5	6.7	13.3
1870	38.6	33.0	5.6	14.4
1860	31.4	27.3	4.1	13.2
1850	23.2	20.9	2.2	9.7

Changes in Race and Ethnicity of the U.S. Population: 1900 to 2013

In 1900, 88 percent of the U.S. population was categorized as white. In 1950 the figure was 90 percent. Since 1950 there has been a gradual decline in that percentage; to 80 percent by 1980, 75 percent by 1990, and 74 percent by 2013. In 2013 the 197 million non-Hispanic white residents of the United States comprised 63.7 percent of the total population.

During the first half of the 20th century, the number of African Americans in the total population of the United States increased from 8.8 million in 1900 to 15 million in 1950. Throughout the first half of the 20th

century, the African American proportion of the total U.S. population was approximately 10 percent. By 1980 the number of African Americans reached 26.5 million, comprising 11.6 percent of the total population. Growth has continued to be steady, with the number reaching 39.2 million in 2013. However, the African American percentage of the total U.S. population increased only modestly; 11.6 percent in 1900 compared to 12.6 percent in 2013.

U.S. census data for Hispanic Americans, collected only since 1970, show a surge in population from less than 10 million in 1970, to 15.6 million in 1980, 22.6 million in 1990, and 51.8 million in 2013. That represents an increase from less than 5 percent of the total U.S. population in 1970 to 16.6 percent in 2013, making it the largest minority group in the country. In 2013, 64 percent of Hispanic Americans living in the United States were of Mexican ethnicity.

Americans of Asian and Pacific Island background increased from 4 million, 1.8 percent of the total U.S. population in 1980, to 15.7 million in 2013, 5.1 percent of the population. They have experienced the fastest growth rate of all racial-ethnic groups in the United States during those decades.

Throughout most of the 20th century, the immigrant population was concentrated mainly in the industrial cities. During the past several decades that settlement pattern has changed. There is now a much greater distribution of immigrants throughout the country.

U.S. Population Projections: 2014 to 2060

Projected figures estimate that the non-Hispanic white population of the United States will decrease from 198 million in 2014 to 182 million in 2060. This decrease in number represents a continuing decline in the non-Hispanic white percentage of the total U.S. population, to 44 percent in 2060, as the country becomes steadily more culturally diverse. By 2060 the African American population is expected to reach 59.7 million, constituting 14.3 percent of the total U.S. population. By 2060 Hispanic Americans are expected to reach 119 million, comprising 28.6 percent of the total population, thereby far exceeding the projected 14.3 percent African American component of the U.S. population in 2060. Projected through 2060, 39 million Asian Americans are expected to comprise 9.3 percent of the total U.S. population. The Native American population of the United States, consisting of Indians, Eskimos, and Aleuts, has grown over the past several decades from 1.4 million in 1980 to 3.9 million in 2014; 0.6 percent and 1.2 percent, respectively of the total U.S. population. The total Native American population in 2060 is expected to reach 5.6 million, but still will comprise only 1.3 percent of the total 2060 U.S. population.

From 2014 to 2060, the total U.S. population is expected to increase by 30.8 percent, the non-Hispanic white component of the population is

expected to decrease by 8.2 percent, the African American population is projected to grow by 42.0 percent, Hispanic Americans by 114.8 percent, and Asian Americans by 128.1 percent.

Age Distribution of the U.S. Population

As has happened in most industrialized countries in the world for several decades, the median age of the U.S. population has been rising and the birth rate falling. However, these demographic changes do not apply equally to all segments of the population. When median age is considered in relation to race and ethnicity, it becomes evident that the non-Hispanic white component of the U.S. population has a higher median age for both males and females than other groups. In 2013, the median age of non-Hispanic whites was 41.5 for males and 44.1 for females. The median age of Asian American males and females was 6.5 years younger than for non-Hispanic whites, and for African Americans was approximately 10 years less than non-Hispanic whites. For Hispanic Americans, the mean age differences are the most dramatic, males 14 years and females almost 16 years younger than non-Hispanic whites. Projecting these findings over the decades to come, the median age of the non-Hispanic white population is expected to increase at a much greater rate than that for Hispanic Americans, African Americans, and Asian Americans.

What all of these demographic changes indicate is that public health planners and clinicians throughout the United States need to be prepared to serve an increasingly diverse and multicultural population throughout the 21st century, requiring greater cultural sensitivity and cultural competence among health service administrators and clinical staff in all medical disciplines.

Poverty in the United States: Race, Ethnicity, and Immigration Status

Immigration and citizenship status are related to poverty. Between 1997 and 2013, figures for U.S.-born Americans with incomes below the poverty line varied between 10.7 and 13.9 percent. Foreign-born naturalized citizens had poverty rates of 9.1 to 12.7 percent, lower percentages than the U.S. born in each year between 1997 and 2013. By contrast, immigrants who had not yet become citizens had poverty rates twice as high as the U.S.-born and naturalized U.S. citizens ranging from 19.4 to 22.8 percent.

A substantial part of the differential rate of poverty between naturalized citizens and immigrants who have not acquired U.S. citizenship is explained by how long the noncitizens have lived in the country. Among those who immigrated prior to 2000 the poverty rate is 15.1 percent. Immigrants who came to the United States between 2000 and 2009 have a poverty rate of 22.6 percent, while 30.0 percent of those who came after 2010 are living below the poverty line. During those three time periods the

percentage for noncitizens living below the poverty line was 36.6, 74.5, and 94.5 percent, respectively. On average, it takes 10 to 15 years for immigrant families to make the transition from economic dependence to independence, from being overall recipients of direct financial support, to being financially self-sustaining and tax contributors.

Between 1990 and 2013 the overall poverty rate in the United States varied from 13.5 to 14.5 percent. There were substantial differences in poverty rates when race and ethnicity were taken into account. The poverty rates for non-Hispanic whites for 1990 to 2013 ranged from 8.8 to 9.6 percent. For African Americans, poverty rates were much higher, ranging from 31.9 to 27.6 percent. Poverty rates for Hispanic Americans are similar to those for African Americans, ranging from 28.1 to 27.2 percent. For Asian Americans poverty rates ranged from 12.2 to 10.5 percent, respectively. These data illustrate another source of emotional stress; poverty applies to immigrants and to African Americans and Hispanic Americans to a much greater degree than it does to non-Hispanic whites and to U.S.-born citizens.

IMMIGRANTS AND REFUGEES: A GLOBAL PERSPECTIVE

Throughout the last half of the 19th and into the 21st century, there were two great flow patterns of *voluntary migrants* in the world: (1) People from rural areas migrating to urban, industrial areas in their own and neighboring countries, motivated mainly by job opportunities in the industrializing cities, and (2) migration from European and Asian countries to countries in North and South America, Southern Africa, Australia, and New Zealand. There were also waves of *involuntary migrants* who left the regions and countries of their upbringing because of discrimination, persecution, revolution, and war, but followed the same paths of migration to industrialized countries. Involuntary migrants are frequently considered refugees.

Migration has been a very sensitive subject for centuries, politically, morally, and emotionally, for the source countries as well as for the receiving countries. The intensity of the host countries' internal conflict, within government and within the host population, about whether migrants should be welcomed or prohibited from entry, has swung from one pole to the other over the past 150 years. There have been periods of welcoming immigrants, followed by periods of revulsion toward immigrants that led to closing the borders and overt discrimination against the foreign born, sometimes accompanied by violence and loss of civil liberties.

Another aspect of migration that has become very evident during the past 60 years has been the large flow of migrants who do not go to other countries with immigration and permanent residence as their objective, but rather migrate for shorter-term purposes of work or study and then return

to their home countries. The increasing speed and ease of international travel and communication have made this type of migration possible. The demand for workers since the 1970s has increasingly shifted toward technical and professional personnel.

Between 1990 and 2013 the total number of migrants in the world, consisting of immigrants, temporary migrants, refugees, and asylum seekers increased 50 percent, from 155 million to 232 million people, comprising more than 3 percent of the world's population.

This large increase was mainly due to the surge in migrants from strife-torn countries in Asia and Africa migrating to Europe, North America, Australia, and New Zealand. During the past 25 years, the migrant population of the more developed countries increased from about 4 percent to nearly 10 percent of the total population of those countries.

Between 1990 and 2013, the migrant population of North America rose steadily from 9 percent to over 14 percent of the total population. The United States accepted the largest number of international migrants between 1990 and 2013: Nearly 23 million, equal to 1 million additional migrants per year. There was a comparable surge in migrant numbers in Europe; from 7 to 10 percent of the total population between 1990 and 2013. By contrast, the migrant population of Asian countries has remained steady, at less than 2 percent of the total population.

In 2013, the U.S. population included 23 million migrants, of which almost 80 percent were immigrants and 20 percent temporary workers, students, refugees, and asylum seekers. Among other highly developed countries, Germany had 5.9 million migrants in 2013, France 5.9 million, Canada 4.5 million, Australia 3.9 million, and the United Kingdom 3.6 million. In 2013, 28 percent of Australia's population was foreign born, compared with 21 percent in Canada, and 14 percent in the United States.

Refugees

In 1990 there were 18.5 million refugees in the world, comprising about 12 percent of the world's migrant population. The number of refugees decreased by 15 percent to 15.7 million in 2013, comprising 7 percent of the world's migrant population. In 2010 there were 15.4 million refugees and since the number has begun to rise again. Nearly 9 of every 10 refugees in the world live in the developing regions, Asia or Africa. In 2013 percentage of the world's refugee population was 11 percent in Europe, and less than 5 percent in North America. A small number of developing countries host the majority of refugees worldwide. Most of the host countries for these refugees bordered countries suffering from the effects of protracted internal strife, and war. In 2013, Jordan hosted the largest number of refugees, followed by Palestine, Pakistan, Syria, Iran, and Germany.

There was a sharp decrease in refugees in Asia between 1990, the peak year, and 2013; 8.4 million and 6.4 million. In Africa, in 1990 there were

5.6 million refugees, decreasing by 2013 to 3.3 million. The peak year for refugees in Africa was 1994, 6.5 million. In 1990 there were 1.4 million refugees in Europe, increasing to a peak of 3.6 million in 1996. However, in 2013, the number of refugees in Europe had decreased to 1.6 million.

A total of 69,909 refugees were admitted to the United States as refugees during 2013. The leading five countries of nationality for recent refugees were Iraq, Burma, Bhutan, Somalia, and Cuba. During 2013, 25,199 individuals were granted asylum. In 1990 the United States took in over 120,000 refugees. Prior to 1983 the United States had waves of Cuban and Indochinese refugees. Between 1985 and 1991 many of the refugees came from the former Soviet Union, Jews and evangelical Christians. From 1992 to 1998 due to the breakup of Yugoslavia refugees came from Bosnia and Kosovo and other Balkan states. Liberians also were admitted as refugees until the end of their civil war. The 1990s to the present also saw refugees come from Ethiopia and the Sudan.

In 2013 there were 284,000 asylum seekers. Sweden, United States, Germany, Italy, and Malaysia were the countries granting asylum to the largest number of persons. The leading countries of nationality for persons granted asylum were Syria, the Congo, Afghanistan, Somalia, and Iraq.

Attitudes Toward Migrants

Government policy on migration also merits consideration. In Switzerland, Australia, Portugal, and the United Kingdom about 40 percent of all immigrants are workers. In Canada, France, the United States, Denmark, and Sweden 50 to 85 percent of immigrants are admitted for the purpose of family reunification. In Australia, Norway, France, and Sweden about 20 percent of immigrants are refugees. These categories are not necessarily mutually exclusive, but suggestive of the criteria different countries use to determine their immigration policies.

During the past 25 years, countries in both Europe and Asia that had historically been the home countries for waves of emigrants have become destination countries for migrants from other countries in Europe, Asia, Latin America and the Caribbean, and Africa. This is the case for southern European countries, such as Spain, Portugal, Italy, and Greece, and for several countries in Asia, including Turkey, Jordan, and Lebanon.

Many countries have had difficulty coping with the surging numbers of migrants. This has led to greater restrictions on migrant numbers, partly in response to public sentiment that the social and cultural integrity of the host nations have become threatened, even undermined, by waves of migrants from other countries and cultures. During the past 15 years, fears of terrorist violence and civil disruption have led many countries to adopt increasingly restrictive and sometimes punitive policies toward legal and illegal migrants, refugees, and asylum seekers. This trend has been observed in the United States, in some countries of the European Union, in

Asia, and in Australia.

In the United States, the defining moment of changing attitudes toward migrants was September 11, 2001. The tide of public opinion in the United States at present favors strengthening border security, preventing inflow of illegal migrants from Mexico and Central America, and reducing the estimated 12-million migrants who are currently living in the country illegally. However, this shift in public sentiment has not reduced legal immigration to the United States, which has continued to rise steadily between 2000 and 2013, but has led to decreasing numbers of refugees and asylum seekers.

Immigrants and refugees to the United States from Muslim countries of Asia and Africa have experienced a sharp rise in vulnerability and emotional and acculturative stress, along with fear of negative stereotyping, discrimination, detention, and deportation. There has been a similar rise in emotional and acculturative stress among Hispanic Americans in recent years, not only among those who have migrated to the United States since 2001 and not yet achieved citizenship, but also among those who immigrated decades earlier and have experienced recent erosion of their confidence in being treated fairly and given the legal protections and access to the economic and educational opportunities, health care, and social services to which they are entitled as U.S. citizens.

GLOBAL MENTAL HEALTH

The Global Burden of Mental Illness

Mental disorders are responsible for little more than 1 percent of all deaths. However, mental and substance use disorders (excluding cognitive disorders) account for almost 7.4 percent of the disease burden, disability adjusted life years (DALYs), worldwide in 2010. Mental disorders are highly prevalent and begin early in life, frequently resulting in inability to work or function, and as a result engender prolonged periods of disability. DALYs are higher in developed countries compared to developing countries that are still dealing with issues that result in high infant mortality and death in young adults due to infectious diseases. In developed countries mental and substance use disorders account for 11.1 percent of the disease burden compared to 6.7 percent in developing countries. Mental and substance use disorders, however, account for a much greater proportion of the years lived with disability (YLDs). Worldwide mental and substance use disorders account for 22.9 percent of YLDs. Depressive disorders alone accounts for 1 in every 10 YLDs worldwide.

Globally across all ages, mental and substance use disorders are the fifth leading disorder category of DALYs and the first for YLDs. The role of mental disorders in causing disability comes to the forefront when young adults, age 15 to 49, are examined. In developed countries the leading cause of DALYs is low back and neck pain followed by depressive disorders, with drug use disorders, alcohol use disorders, and anxiety disorders among the top ten. Among males, of all the mental and substance use disorders, drug use disorders are the leading cause of DALYs and rank fourth among all disorders. For females, depressive disorders are the second leading cause of all DALYs. In developed countries the leading causes of DALYs is HIV, low back and neck pain and tuberculosis, followed by depressive disorders, with anxiety disorders ranking tenth. For males, among all causes of DALYs, depressive disorders ranked sixth and drug use disorders tenth, while for females depressive disorders were the second leading cause of DALYs and anxiety disorders the fifth. As countries undergo the epidemiological transition away from disorders that result in early mortality, mental and substance use disorders account for an increasing proportion of the global burden of disease.

There are large discrepancies in countries around the world, and even within countries, as to the availability of mental health care, despite the enormous amount of disability that these illnesses cause. One metric, the proportion of total health expenditures for mental health, is an indicator of the priority given to mental health within the health sector of different countries. Proportionally, lower-income countries allocate a smaller percentage of their health budget to mental health. The world median is 2.8

percent of the health budget; highest-income countries 5.1 percent and lowest-income countries 0.5%. There is a stark disparity in the availability of psychiatric beds between high-income and low-income countries, which is even more dramatic for psychiatric beds located in general hospitals which utilize an acute care rather than a chronic care treatment model; 13.6 compared to 0.6 per 100,000 population. Similarly, there is a lack of psychiatrists in low-income countries; 0.05 per 100,000 population compared to 8.6 in high-income countries.

Cross-National Differences in Mental Disorders

A number of factors have been identified that explain how culture may influence prevalence and phenomenology of mental disorders. Pathogenicity, culturally shared ideas or beliefs directly contribute to stress and lead to psychopathology. *Pathoselectivity,* culturally sanctioned reaction patterns as a means of coping are adopted in times of stress. *Pathoplasticity,* cultural beliefs influence the presentation of symptoms of a mental disorder. *Pathoelaborative effects,* cultural beliefs may reinforce certain pathological behaviors making them more exaggerated. *Pathofacilitative effects,* cultural factors influence the frequency or presentation of a mental disorder such as its prevalence. *Pathoreactivity,* cultural beliefs influence the population's reaction toward a mental illness, thereby affecting the labeling of disorder and expression of suffering due to the disorder.

The rates of mental disorders vary widely across nations. This is best illustrated by the World Mental Health Surveys that have used the same instrument, the Composite International Diagnostic Interview, in numerous countries around the world and have attempted to employ similar methodology. This approach assumes that mental disorders are universal across cultures and nations; which is controversial, as it marginalizes the role of social factors and culture on mental distress. DSM-IV disorders in national surveys in the prior year varied widely from 6 percent in Nigeria to 27 percent in the United States. This wide disparity in rates of disorders is also found in European countries, where the rates ranged from 8.8 percent in Italy to 21.4 percent in the Ukraine. Even within countries, rates varied widely. For example, in China the rates were 4.3 percent in Shanghai and 16.0 percent in Shenzhen. Overall, anxiety disorders were found to be the most common across countries.

Are rates of disorders substantially different cross-nationally? Other cross-national comparisons would suggest that they are. However, these differences could be influenced by cultural biases, by translation errors in survey instruments, and by subjects' responses to fully structured psychiatric interview schedules administered by nonmental health professionals. Concepts and phrases used in standardized diagnostic instruments to describe psychiatric symptoms, such as the Composite

International Diagnostic Interview, are less consistent with cultural concepts in many developing countries of Asia, Africa, and Latin America than in developed Western countries. Table 4.3–3 presents the 12-month prevalence rates from international epidemiological studies using the Composite International Diagnostic Interview.

Table 4.3–3.
Twelve-Month DSM-IV[a] Adult Prevalence Rates from Selected Composite International Diagnostic Interview Studies since 1990 (%)

	Anxiety	Mood	Substance	Mild	Moderate	Severe	Any
AMERICAS							
Brazil: São Paulo	19.9	11.8	3.8	9.8	9.8	10.0	29.6
Canada	4.7	5.2	11.0	8.3	6.1	4.4	18.7
Chile	5.0	9.0	6.6	3.2	6.9	6.9	17.0[a]
Colombia	14.4	6.9	2.8	8.1	8.4	4.9	21.0
Guatemala	2.3	1.4	3.8	5.1	1.4	2.4	8.9
Mexico	8.4	5.0	2.5	5.2	4.7	3.5	13.4
Peru	7.9	3.5	1.7	5.6	5.6	2.3	13.5
United States	19.0	9.8	3.8	9.4	10.7	6.9	27.0
EUROPE							
Belgium	8.4	6.1	1.3	3.8	5.1	4.3	13.2
Bulgaria	7.6	3.2	1.2	5.4	3.6	2.3	11.2
Finland	3.7	4.8	4.2				
France	13.7	6.8	0.8	7.2	8.1	3.5	18.9
Germany	8.3	3.4	1.2	3.9	4.8	2.4	11.0
Italy	6.5	3.6	0.1	3.2	4.2	1.4	8.8
Netherlands	8.9	5.5	1.7	5.2	4.2	4.2	13.6
Northern Ireland	14.6	10.6	3.5	8.7	7.7	6.7	23.1
Portugal	16.5	8.3	1.6	7.3	11.6	4.0	22.9
Romania	4.9	2.5	1.0	3.5	2.4	2.3	8.2
Spain	6.6	4.4	0.3	3.6	4.2	2.9	9.7
Turkey	5.8	4.2	0				8.4[a]
Ukraine	6.8	10.0	6.4	8.1	8.4	4.9	21.4
OCEANIA							
Australia	9.7	5.8	5.7				17.7
New Zealand	15.0	8.0	3.4	6.7	8.6	5.3	20.7
MIDDLE EAST AND AFRICA							
Iran	15.6	14.6	2.8	8.5	7.0	8.1	23.6
Iraq	10.4	4.1	0.3	5.7	4.9	3.0	13.6
Israel	3.6	6.4	1.3	2.8	3.5	3.7	10.0
Lebanon	12.1	7.0	1.3	6.2	7.7	4.0	17.9
Nigeria	4.2	1.2	0.9	4.5	0.8	0.8	6.0

South Africa	8.4	4.9	5.7	7.2	5.3	4.3	16.9
ASIA							
India: Pondicherry	10.5	5.5	5.3	7.9	7.8	4.3	1.2
Japan	4.8	2.8	1.0	2.9	3.8	1.3	8.0
PRC: Beijing, Shanghai	3.0	2.2	1.6	3.8	2.3	1.0	7.1
PRC: Shenzhen	11.4	4.8	0	9.8	5.2	1.0	16.0
South Korea	6.8	3.6	4.4				13.5
Vietnam: Mekong Delta	0.4	0.4	1.1				1.9

[a]*Diagnostic and Statistical Manual of Mental Disorders* (DSM-III-R) diagnoses.

PRC, People's Republic of China.

While recognizing the public health need for such cross-national studies, debate continues about whether such efforts take into account relevant cultural issues. Classification of depression and anxiety, in particular, continues to be a contentious issue, since diagnostic labels such as depression and phobias have no conceptually equivalent terms in many non-European languages.

Treatment Gap in Access to Psychiatric Treatment

Despite its limitations, the World Mental Health initiative has provided important insights into mental health in different national populations. Severity of disorders in nations is closely associated with access to treatment. In developed countries 35.5 to 50.3 percent of people diagnosed with some form of mental illness, and in less-developed countries 76.3 to 85.4 percent, received no treatment at all for the diagnosed condition in the previous 12 months. The percentage of people diagnosed with a mental illness that receive psychiatric treatment correlates with the countries' percentage of gross domestic product spent on health care. Seriousness of the disorder was also correlated with service utilization, but varied widely across nations. For example, for those with a serious mental disorder, 59.7 percent in the United States and 72.8 percent in Northern Ireland received psychiatric treatment, compared to only 4.3 percent in Shenzhen, China, 20.1 percent in Lebanon, and 21.3 percent in Nigeria.

The *treatment gap* is the difference between the true prevalence of a disorder and the treated proportion of individuals affected by the disorder. Data on the extent of the treatment gap between high- and low-income countries are beginning to emerge. Data from the Americas illustrate this disparity. The treatment gap for anxiety disorders, affective disorders, and substance use disorders for North America is 46.3, 57.9, and 63.4 percent, while for Latin America it is 76.3, 83.2, and 85.3 percent respectively. The treatment gap for schizophrenia in low- and middle-income countries is alarming. The treatment gap in 50 low- and middle-income countries was estimated to be 69 percent and among the low-income countries was 89

percent. In an analysis of 35 Latin America and the Caribbean countries the treatment gap for schizophrenia was 56.4 percent, with the highest rate being in Central America; 79.4 percent. In order to reduce disability, as well as the economic burden associated with mental disorders, bridging of the treatment gap and reducing the *treatment lag time* needs to occur. There are serious implications of not reducing the treatment gap for the mentally ill, which include increased poverty and lower socioeconomic status for the families of mentally ill individuals, impaired family functioning, decreased educational attainment, poorer quality of life, and increased mortality.

To reduce the treatment gap, scaling up services and task shifting needs to occur. Many more people need access to effective mental health services that are culturally appropriate taking community beliefs and attitudes into account. To address the global treatment gap 25 grand challenges have been identified by the NIMH and summarized into six goals: (1) Identify root cause, risk, and protective factors of mental disorders: identify modifiable social and biological risk factors across the life course; understand the impact of poverty, violence, war, migration, and disaster; and identify biomarkers; (2) advance prevention and implementation of early intervention strategies: support community environments that promote physical and mental well-being throughout life; reduce the duration of untreated illness by developing culturally sensitive early interventions across settings; develop interventions to reduce the long-term negative impact of low childhood socioeconomic status on cognitive ability and mental health; develop an evidence-based set of primary prevention interventions for a range of mental disorders; and develop locally appropriate strategies to eliminate childhood abuse and enhance child protection; (3) improve treatments and expand access to care: integrate screening and core packages of mental health services into routine primary health care; reduce the cost and improve the supply of effective medications; develop effective treatments for use by nonspecialists, including lay health workers with minimal training; incorporate functional impairment and disability into assessment; provide effective and affordable community-based care and rehabilitation; improve children's access to evidence-based care by trained health providers in low- and middle-income countries; and develop mobile and IT technologies, such as telemedicine, to increase access to evidence-based care; (4) raise awareness of the global burden of untreated illness: develop culturally informed methods to eliminate the stigma, discrimination, and social exclusion of patients and families across cultural settings; establish cross-national evidence on the cultural, socioeconomic, and services factors underlying disparities in incidence, diagnosis, treatment, and outcomes; develop valid and reliable definitions, models, and measurement tools for quantitative assessment at the individual and population levels for use across cultures and settings; and establish shared, standardized global data

systems for collecting surveillance data on the prevalence, treatment patterns, and availability of human resources and services; (5) build human resource capacity: increase capacity in low- and middle-income countries by creating regional centers for mental health research, education, training, and practice that incorporate the views and needs of local people; develop sustainable models to train and increase the number of culturally and ethnically diverse lay and specialist providers to deliver evidence-based services; and strengthen the mental health component in the training of all health care personnel; (6) transform health system and policy response: establish and implement minimum health care standards for mental disorders around the world; redesign health systems to integrate mental disorders with other chronic disease care, and create parity between mental and physical illness in investment in research, training, treatment, and prevention; and incorporate a mental health component in international aid and development programs.

Thus, culture has a significant role in global mental health. It contributes to significant health and social disparities in society, leading to inequalities in the distribution of health problems and resources in the population. The phenomena of *globalization,* directly affected by migration of all kinds and by technological advances, particularly in the field of electronic communication and its social network derivatives, resulted in creating new forms of hybrid identities. The move toward universal measures, diagnoses and treatments of mental illness is inconsistent with the belief that mental distress is culturally and socially mediated.

Risk for Schizophrenia in Developing Countries. That schizophrenia has a better course and outcome in countries of the developing world, than in most developed countries, has been identified by some scholars as the single most important finding of cultural differences in cross-cultural research on mental illness. The most compelling evidence comes from three studies by the World Health Organization (WHO): The International Pilot Study of Schizophrenia, the Determinants of Outcome of Severe Mental Disorder, and the International Study of Schizophrenia, as well as from several other recent studies.

The findings of the three WHO studies have been independently reevaluated for potential sources of bias: Differences in follow-up, arbitrary grouping of centers, diagnostic ambiguities, selective outcome measures, gender, and age. However, none of these potential confounds significantly altered the findings. In addition, the findings have consistently stood up over time with longer periods of follow-up, although the differences between the developing countries (in Africa, Asia, and Latin America) and the industrialized West (mainly countries in Europe and North America) have progressively decreased.

It has been hypothesized that likely factors promoting better recovery

from schizophrenic disorders in developing countries include supportive kin; alternative beliefs about causation of mental illness, including malign magic, sorcery, and witchcraft; greater social acceptance both within families and in society generally; and participation in the microeconomy, with more flexible work roles. For example, in developed countries individuals may not be allowed to work while they are receiving disability insurance payments, whereas in developing countries individuals with schizophrenia may assume a role or job that their disability allows them to perform, for example, collecting firewood that the family might sell, thereby also helping to integrate the mentally ill person into the daily life of their family and community.

A recent systematic review of the prevalence of schizophrenia once again found that rates of schizophrenia were significantly higher in developed nations compared to developing nations. This finding is congruent with the belief that the outcome of schizophrenia is more favorable in poorer, developing nations than it is in richer, developed nations. However, another meta-analysis could not demonstrate that the incidence of schizophrenia varied between nations with different economic status, in contrast to comparing countries by categorization as developed or developing.

Other investigators argue that this claim appears to fly in the face of clinical judgment and experience. A review of 23 studies independent of the WHO studies has brought those findings into question. These investigators found that wherever inadequate clinical treatment is found, lack of care is associated with relatively poor outcomes, and that being able to access effective psychiatric treatment is associated with improved outcomes. The authors argued that because the individuals diagnosed as having schizophrenia in many of the WHO studies were receiving care at the leading academic psychiatric facilities in the countries participating in the WHO studies, this might account for the finding of a better outcome. Although the WHO studies findings of a more favorable outcome for schizophrenia in developing countries has stood up to scrutiny, other studies have brought the finding into question, continuing the controversy over this association.

RACIAL AND ETHNIC DIFFERENCES IN MENTAL DISORDERS IN THE UNITED STATES

A number of community-based epidemiological studies in the United States have examined the rates of disorders across specific ethnic groups. These studies have found lower than expected prevalence of mental disorders among disadvantaged racial and ethnic minority groups in the United States. African Americans were found to have lower rates of major depression in the Epidemiological Catchment Area study. The lifetime prevalence rates of major depression for whites was 5.1 percent; for

Hispanics, 4.4 percent; and for African Americans, 3.1 percent. African Americans, however, had higher rates for all lifetime disorders combined. This finding of differential rates could be explained by adjusting for socioeconomic status; that is, greater poverty correlates with increased prevalence of mental illness.

The National Comorbidity Survey (NCS) found lower lifetime prevalence rates of mental illness among African Americans than whites, and in particular for mood, anxiety, and substance use disorders. The lifetime rates for mood disorders were 19.8 percent for whites, 17.9 percent for Hispanic Americans, and 13.7 percent for African Americans. The National Health and Nutrition Examination Survey-III also found lifetime rates of major depression to be significantly higher among whites, 9.6 percent, than African Americans, 6.8 percent, or Mexican Americans, 6.7 percent. Although African Americans had lower lifetime risk of mood disorders than whites, once diagnosed they were more likely to remain persistently ill.

NCS rates for anxiety disorders were 29.1 percent among whites, 28.4 percent for Hispanic Americans, and 24.7 percent for African Americans. The rates for lifetime substance use disorders for the three groups, whites, Hispanic Americans, and African Americans, were 29.5, 22.9, and 13.1 percent, respectively. Hispanic Americans, and in particular Mexican Americans, were found to be at lower risk for substance use and anxiety disorders than whites. The lower rate for substance use disorders was also found in the National Epidemiological Survey on Alcohol and Related Conditions (NESARC), with whites having a prevalence rate of 1-year alcohol use disorders of 8.9 percent, Hispanic Americans, 8.9 percent, African Americans, 6.9 percent, Asian Americans, 4.5 percent, and Native Americans, 12.2 percent. This study also found lower lifetime rates for major depression among Hispanic Americans, 10.9 percent, compared to whites, 17.8 percent. In 2007 the National Survey of American Life compared rates of major depression between Caribbean blacks, African Americans, and whites. Although there were no significant differences in 1-year prevalence between the three groups, lifetime rates were highest among whites, 17.9 percent, followed by Caribbean blacks, 12.9 percent, and African Americans, 10.4 percent. The chronicity of major depressive disorder was higher for both African Americans and Caribbean blacks, approximately 56 percent, while much lower for whites, 38.6 percent. This study was consistent with findings from the NCS that members of disadvantaged racial and ethnic groups in the United States do not have an increased risk for mental disorders; however, once diagnosed, they do tend to have more persistent disorders.

Although African Americans have a lower prevalence rate for mood, anxiety, and substance use disorders, this may not be the case for schizophrenia. The Child Health and Development Study found that

African Americans were about three times more likely than whites to be diagnosed with schizophrenia. The association may be partly explained by African American families having lower socioeconomic status, a significant risk factor for schizophrenia.

A more detailed examination of differences across racial groups was included in the National Comorbidity Survey Replication (NCS-R). Non-Hispanic African Americans and Hispanic Americans were at significantly lower risk than non-Hispanic whites for anxiety disorders and mood disorders. Non-Hispanic African Americans had lower rates of substance use disorders than non-Hispanic whites. More specifically, both minority groups were at lower risk for depression, generalized anxiety disorder, and social phobia. In addition, Hispanic Americans had lower risk for dysthymia, oppositional-defiant disorder, and attention-deficit/hyperactivity disorder. Non-Hispanic African Americans had lower risk for panic disorder, substance use disorders, and early-onset impulse control disorders. The lower rates among Hispanic Americans and African Americans compared to non-Hispanic whites appear to be due to reduced lifetime risk of disorders, as opposed to persistence of chronic disorders. The researchers concluded that the pattern of racial-ethnic differences in risk for mental disorders suggests the presence of protective factors that originate in childhood and have generalized effects, as the lower lifetime risk for both Hispanic Americans and African Americans begins prior to age 10 for depression and anxiety disorders. The retention of ethnic identification and participation in communal, religious, and other activities have been suggested as protective factors that may decrease the lifetime risk for mental disorders in close-knit ethnic minority communities. NESARC replicated this finding by showing that higher scores on ethnic identity measures, a sense of pride, belonging, attachment to one's racial or ethnic group, and participating in ethnic behaviors were associated with decreased lifetime diagnosis of a mental disorder. Cultural differences in response to psychiatric diagnostic survey items may be another possible explanation for these findings. However, disadvantaged ethnic groups usually overreport in studies measuring psychological distress, whereas these studies find lower rates.

Many studies group Hispanic Americans from all countries in Latin America into one category. By contrast, the National Latino and Asian American Study (NLAAS) compared the rates of mental disorder among Puerto Ricans, Cuban Americans, Mexican Americans, and other Latinos. Puerto Ricans had the highest overall prevalence rate among the Latino ethnic groups assessed. Increased rates among Puerto Ricans were noted specifically for depressive disorders, anxiety disorders, and substance use disorders. Mexican Americans were noted to have the lowest risk among the Hispanic American groups for depressive disorders. Cuban Americans were at the lowest risk for substance use disorders and, among men, for

anxiety disorders. Differences among the Hispanic American groups in overall risk for lifetime mental disorders were also noted, with Cuban American males being at the lowest risk and other Latin American females having the lowest rate. The increased risk for mental disorders among Puerto Ricans compared to other Hispanic American groups has been noted in other studies.

MENTAL HEALTH OF INDIGENOUS PEOPLE OF THE UNITED STATES

The 2.5 million U.S. indigenous populations of Native Americans, Eskimos, and Aleuts comprise 561 recognized tribes with over 200 spoken languages and represent 0.9 percent of the total U.S. population. There are no words for depression and anxiety in many indigenous languages. Indigenous people in the United States, as is true for other indigenous populations worldwide, have a median age of death lower than the majority population, and as a result, the median age of the indigenous population is younger than the national population. Native Americans are five times more likely to die of alcoholism and from liver disease, and twice as likely to have diabetes mellitus compared to whites. The indigenous population also has higher rates of accidents, violence, suicide, and homicide than the majority U.S. population. One-third of the Native American population does not receive regular health care. The Indian Health Service provides care only to about 20 percent of the Native American population.

Several studies have focused on the rates of mental disorders among specific Native American populations; the Pacific Northwest Indians, the Southwest California Indians, the Southwest Indians, and the Northern Plains Indians. The American Indian Service Utilization, Psychiatric Epidemiology, Risk and Protective Factors Project of the Southwest Indians, and the Northern Plains Indians have provided comprehensive insight into the mental health of the Native American population. Alcohol use disorders and posttraumatic stress disorder (PTSD) were found to be more common among these Native Americans compared to other ethnic groups, but they were at lower risk for major depression. Studies of American Northern Plains and Southwest Indian Vietnam veterans also found very high rates of PTSD: 31 and 27 percent for current prevalence and 57 and 45 percent for lifetime prevalence, respectively. This was significantly higher than for Vietnam war veterans of other ethnic groups including whites, African Americans, and Japanese Americans. These Native American veterans were more likely than other groups to have been exposed to war zone stress in Vietnam, including combat exposure, as well as being wounded and having noncombat injuries. This higher rate of exposure to war zone stress accounted for the high rate of PTSD among the Native American veterans compared to other ethnic groups. In addition, these veterans had higher rates of alcohol use disorders, over 70 percent current and 80 percent lifetime prevalence, in contrast to 11 to 32 percent

current and 33 to 50 percent lifetime for the other ethnic groups.

Native American children are more likely to receive mental health treatment through the juvenile justice system and inpatient facilities. Among adults, Native Americans have higher rates of seeking help for substance use problems than whites, but lower rates of seeking help for depression and anxiety disorders. Less than 30 percent of those with a mental disorder obtain help from mental health service providers.

ETHNICITY AND PSYCHIATRIC SERVICE UTILIZATION

A number of recent studies have examined differences in mental health service delivery among U.S. racial and ethnic minorities, highlighting their markedly lower rates of service utilization. The NLAAS examined mental health service utilization by Asian Americans and Hispanic Americans. Among those Asian Americans with any DSM-IV-TR diagnosis in the past 12 months, 34.1 percent utilized mental health–related services, whether through general medical services or specialized mental health care. This was a significantly lower rate in comparison to the 41.1 percent utilization rate in the general U.S. population, measured by a similar methodology in the NCS-R study.

Among Asian Americans, perceived helpfulness, subjective satisfaction, and rate of mental health service use varied by country of origin and generation in the United States. Second-generation Asian Americans with a psychiatric diagnosis had service utilization rates, 28.8 percent, similar to those of other immigrant groups, 30.4 percent. The third generation had the highest rate of mental health service utilization, 62.6 percent. Unlike other reports, which were not based on large household samples and suggested that acculturative stress was a major factor in help seeking, acculturative stress did not play as strong a role as previously believed. There were no differences in mental health service utilization based on country of origin, the number of years in the United States, age at immigration, or fluency in English. This suggested that more general factors, or possibly cultural factors, such as stigma or “loss of face” may be barriers to service use, beyond immigrant-specific factors such as lack of language fluency and lack of knowledge of how to obtain services.

Among Hispanic Americans in the NLAAS, factors such as nativity, language facility, age at migration, years of residence in the United States, and generational status were associated with whether or not Hispanic Americans had used mental health services. When those with a past-year psychiatric diagnosis were examined, these associations were no longer apparent. Among those with a mental disorder during the past year, service utilization in the past year did not greatly vary by country of origin, English language fluency, age at immigration, or number of years in the United States. Unlike Asian Americans, mental health service utilization did not vary greatly by generation since immigration. Having health insurance

coverage was associated with increased mental health care utilization. Only 19.1 percent of those with a mental disorder who had no health insurance coverage obtained mental health care.

Overall, the rates of mental health service utilization among Hispanic Americans during the past year were not much different than those of the total U.S. population (Table 4.3–4). This was indicative of increased mental health service utilization in the past 10 years among Hispanics, in which a threefold increase has been observed since the 1990s. Earlier studies of Mexican Americans showed much lower rates of mental health service utilization. The more recent increase in service utilization has been attributed to greater public awareness of mental disorders and the need to obtain treatment; the media targeting Hispanic Americans and reducing stigma; and improved screening among Hispanic Americans by primary care providers.

Table 4.3–4.
Rate of Any Mental Health Service Utilization in the Past Year among Those with a *DSM-IV* Diagnosis in the 12 Months Prior to Interview by Ethnicity and Generation since Immigration (%)

		Generation since Immigration		
Race-Ethnic Group	**Total**	**1st**	**2nd**	**3rd**
ASIAN AMERICAN[a]	34.1	30.4	28.8	62.6
Chinese	31.0			
Filipino	34.1			
Vietnamese	48.5			
Other Asians	33.0			
HISPANIC AMERICAN[a]		35.8	37.1	37.0
Puerto Rican	43.0			
Cuban	41.9			
Mexican	34.2			
Other Latino	36.8			
BLACK CARIBBEAN[b]	31.4	10.9	44.2	49.5
AFRICAN AMERICAN[b]	35.4			
U.S. GENERAL POPULATION[c]	41.1			

[a]National Latino and Asian American Study.

[b]National Survey of American Life.

[c]National Comorbidity Survey Replication.

Overall seeking mental health care among those with any mental disorder in the past 12 months was low and not significantly different between African Americans, 35.4 percent, and American-Caribbean blacks, 31.4 percent. In particular, the NASL found that most African Americans

with major depression did not seek mental health care. Only 45.0 percent of African Americans and 24.3 percent of American-Caribbean blacks received treatment of any type. The rate of utilization of mental health services did not significantly increase as symptom severity increased. Only 48.5 percent of African Americans and 21.9 percent of American-Caribbean blacks with severe or very severe symptoms received any psychiatric treatment.

Unlike Hispanic American and Asian American immigrants with a mental disorder, there were differences in acculturative stress variables among Caribbean black immigrants to the United States. Those who were foreign born used mental health services significantly less, 10.9 percent compared to 46.8 percent for U.S.-born subjects. Utilization of mental health services was lowest among those who had been in the United States for less than 13 years.

Consistent with the findings within the Hispanic American and African American groups in the above studies, the NCS-R did not find differences in mental health treatment adequacy among those who were diagnosed with a mental disorder in the past year. There is evidence that rates of mental health treatment increased almost threefold from 1987 to 1997. During that period the percentage of patients who were prescribed antidepressants rose from 37.5 percent to 74.5 percent, particularly among Hispanic American and African-American patients.

DISCRIMINATION, MENTAL HEALTH, AND SERVICE UTILIZATION

Disparities in Mental Health Services

Numerous studies, including recent ones, have shown that racial and ethnic minorities in the United States receive more limited mental health services than whites, adjusted for need. This disparity in mental health care may in part be due to living in neighborhoods with fewer mental health providers, receiving less adequate mental health care, delay in seeking care until symptoms are more acute, and misdiagnosis of mental disorders. Based on data from 2004 to 2007, African Americans had 53% lower odds and Hispanic Americans had 41% lower odds of seeking any type of mental health treatment in the past year, compared with whites. Analysis of medical expenditures in the United States from 2004 to 2010 for probable mental illness has shown that the mental health care system provides comparatively less outpatient care to African Americans and Hispanic Americans than to whites, even after controlling for income, education, and availability of health insurance: Rates of mental health outpatient expenditures were 20.8%, 22.2%, and 28.4%, respectively. Minorities also had lower inpatient expenditures than whites; African Americans 30.6% lower and Hispanic Americans 50.7% lower, but greater emergency room expenditures. Some differences in expenditures for psychotropic

medications were noted for those with probable mental illness; whites 28.4%, African Americans 25.0%, and Hispanic Americans 27.6%.

In addition, studies conducted over the past 25 years have shown that regardless of disorder diagnosed, African American psychiatric patients are more likely than white patients to be treated as inpatients, hospitalized involuntarily, placed in seclusion or restraints without evidence of greater degree of violence, and treated with higher doses of antipsychotic medications. These differences are not due to the greater severity of disorders between white and African-American patients. Disparities are also seen in the Department of Veteran's Administration; for example African American and Hispanic American veterans compared to whites were less likely to receive minimal trials of psychopharmacological treatment for PTSD. One hypothesis for these discrepancies in treatment between African American and white patients is that whites are more likely to seek out mental health care voluntarily than African Americans, and African Americans are more likely to enter the mental health care system through more coercive and less voluntary referral systems. African Americans are also more likely than whites to use emergency room services, resulting in more crisis-oriented help seeking and service utilization. Once hospitalized in an institution with predominantly white staff, African-American patients may receive differential care as a result of discrimination. That is, service personnel who are not familiar with the illness concepts and behavioral norms of non-white groups tend to assess minorities as more severely ill and more dangerous than patients of their own racial or ethnic group. Consequently, such patients tend more often than white patients to be hospitalized involuntarily, placed in seclusion or restrains, and treated with higher doses of antipsychotic medications.

African-American patients assessed in psychiatric emergency services are more likely to be diagnosed with schizophrenia and substance abuse than matched white patients. White patients are more often diagnosed with a mood disorder. The cultural distance between the clinician and the patient can affect the degree of psychopathology inferred and the diagnosis given. These differences in diagnosis by race have also been found when comparable research diagnostic instruments have been used for patient assessment. Semistructured diagnostic instruments based on explicit DSM-IV-TR criteria did not necessarily eliminate racial disparities in diagnostic outcomes. It appears that the process clinicians use to link symptom observations to diagnostic constructs may differ, in particular for schizophrenia, between African American and white patients. The pattern of psychotic symptoms that predicts a clinician making a diagnosis of schizophrenia in African American and white patients is different. Among African-American patients loose associations, inappropriate affect, auditory hallucinations, and vague speech increased the likelihood of a diagnosis of schizophrenia. Positive predictors for white patients were

vague speech and loose associations. Auditory hallucinations are more frequently attributed to African-American patients.

African Americans are less likely to have had outpatient treatment, are more likely to have had longer delays in seeking care, and to present as more severely ill. Even with expanded health insurance coverage, African Americans with behavioral health needs were 10 percent less likely to receive mental health services than whites, whereas no differences were found for Hispanic Americans. African American outpatients with schizophrenia are more likely to be treated with antipsychotic medications and are more likely to be hospitalized. The reason for hospitalization was also different between African Americans and whites. African-American patients were more likely to be admitted for some form of behavioral disturbance, while white patients were more likely to be admitted for cognitive or affective disturbances. Also, African Americans were more likely to have police or emergency service involvement, despite an absence of racial differences in violence, suicide risk, or substance use when assessed. In addition, African-American patients are more likely, even after controlling for health insurance status, to be referred to public rather than private inpatient psychiatric facilities, suggesting racial bias in psychiatric emergency room assessment and recommended treatment. Based on the 2004 to 2010 Medical Expenditure Panel Survey, following inpatient discharge African Americans were less likely to receive any treatment or begin adequate follow-up within 30 days of discharge.

African-American patients diagnosed with major depression are less likely to receive antidepressant medications than whites, and are less likely to be treated with electroconvulsive therapy. These findings cannot be explained by demographic or socioeconomic differences. One explanation may be that there are conscious or unconscious biases in psychiatrists' treatment decisions. Although both African Americans and Hispanic Americans were less likely to fill an antidepressant prescription when diagnosed with depression, once a prescription was filled, they were just as likely as whites to receive an adequate course of treatment. These findings indicate that initiating treatment for depression is the biggest hurdle in overcoming these disparities. African-American patients have been found to be more likely to be treated with depot rather than oral neuroleptics compared to whites, after controlling for the type and severity of illness. When treated with antipsychotic drugs, African Americans are less likely to receive second-generation antipsychotics than whites, placing them at increased risk for dystonia and tardive dyskinesia. These differences in antipsychotic prescribing patterns may be due to physicians' concern over an increased risk of diabetes among African Americans compared with whites, or may be due to physicians perceiving their symptoms differently. Disparities in mental health care for African Americans and Hispanic Americans have also been noted in studies of adolescents.

A disparity in prescription drug use for mental illness also has been found among Hispanic Americans and Asian Indian Americans. From 1996 to 2000, Asian Indian Americans were found to use prescription drugs 23.6 percent less than whites, while the difference between whites and African Americans and between whites and Hispanic Americans was 8.3 and 6.1 percent, respectively. Disparities in mental health service use among Asian American immigrants may be linked to language-based discrimination, although racial bias cannot be excluded. A study of Chinese Americans found a higher level of use of informal services and help seeking from friends and relatives for emotional problems. Those Chinese Americans who reported experiencing language-based discrimination had a more negative attitude toward formal mental health services.

Data on racial and ethnic differences in mental health counseling and psychotherapy are similar to the psychopharmacological studies showing disparities for minorities. A study examining visits to primary care physicians based on the National Ambulatory Medical Care survey from 1997 to 2000 found that primary care physicians provided similar or higher rates of general health counseling to African American than white patients. However, the rates of counseling for mental health issues were significantly lower for African-American patients. A Medicaid database record study from 2003 also noted that African Americans were less likely to receive combined psychotherapy and antidepressant treatment, and attended fewer psychotherapy sessions. The lower rate of mental health counseling among African Americans may be due to decreased reporting of depressive symptoms, inadequate communication between African-American patients and their primary care physicians, and decreased willingness to discuss mental health issues. On the other hand, another study utilizing the Medical Expenditure Panel Survey from 2000 found that African Americans were more likely than Hispanic Americans or whites to receive an adequate course of psychotherapy for depression. These findings suggest that initiating treatment is the biggest hurdle, and that once they are engaged in treatment, African Americans have high compliance with psychotherapy.

Self-Perceived Discrimination

It is unclear what impact the disparities in mental health care service delivery and possible discriminatory differential mental health care practices have on minority groups' mental health. Numerous studies have shown that self-perceived discrimination is associated with impaired mental health. For example, in a study of Korean immigrants in Toronto, perceived racial discrimination was associated with depressive symptoms and erosion of positive affect. Self-reported discrimination was associated with increased risk of mental disorders among Asian Americans in the NLAAS and was found to be a more important predictor than acculturative

stress. The relationship between discrimination and increased rates of mental disorders was robust and could not be accounted for by social desirability, physical health, other stressors, and sociodemographic factors. Earlier studies also found that discrimination may be an important predictor of impaired mental health status among African Americans, Hispanic Americans, and Native Americans. Discrimination may lead to affective reactions; it may shape a person's appraisal of the world; engender hopelessness; lower self-esteem; and contribute to the internalization of negative stereotypes. Discrimination has not only been shown to increase the risk of depression and anxiety disorders, but also has been shown to increase the incidence of schizophrenia in minorities.

A longitudinal study found that a change in self-reported discrimination was associated with a change in self-reported depressive symptoms. The negative effect of discrimination on mental health is not limited to adults. It has been found in studies of Puerto Rican adolescents who had lower self-esteem and higher depression and stress symptoms. However, having a strong ethnic identity, as noted in a study of Filipino Americans, helps reduce the stress of discrimination and is associated with fewer depressive symptoms. Having a sense of ethnic pride, involvement in ethnic communal activities, and having an emotional commitment to one's ethnic group may be protective factors for mental health.

Racial and ethnic discrimination can adversely affect mental health in three ways. Racism in societal institutions can lead to truncated socioeconomic mobility, differential access to desirable resources, including housing, educational and occupational resources, and overall living conditions. Second, discrimination can induce physiological and psychological reactions that can lead to adverse changes in mental health status, such as dysregulation of the hypothalamic–pituitary–adrenal axis. Third, in race-conscious societies, the acceptance of negative cultural stereotypes can lead to unfavorable self-evaluations, resulting in diminished self-esteem, which can have deleterious effects on psychological functioning.

IMMIGRATION AND MENTAL HEALTH

The possible causal relationship between immigration and mental illness has been the subject of intensive research and debate since the pioneering studies in the 1930s by Ødegaard and Malzberg. Does immigration itself have an influence on mental health status? Is the increase or decrease in psychopathology noted among immigrants a result of the immigration process and the stresses of acculturation, or is it a result of a selection bias or a reflection of psychopathology that developed in the immigrant's country of origin? These issues are of theoretical as well as practical importance and continue to be debated.

Psychological stress associated with immigration has been examined in

a number of ethnic groups. The greatest stress appears to become manifest during the first 3 years after arrival in the United States, as noted in studies of Hispanic immigrants and immigrants from the former Soviet Union. Refugees from Southeast Asia have been shown to experience initial euphoria, followed by increased demoralization in the second year, and a gradual return to well-being beginning in the third year after immigration. Similar findings have been noted for Cuban, Eastern European, and Chinese immigrants to North America. Among Korean immigrants to the United States, the highest level of depressive symptoms is observed during the first 2 years, with symptom remission in the third year following immigration. Studies of individuals originating from the same region or country and immigrating to two different countries are rare. Such a study comparing Jewish refugees from the former Soviet Union immigrating to either Israel or the United States demonstrated that the economic benefits immigrants experienced in the United States were not sufficient to overcome the effects of their sense of communal and psychological marginalization in the United States, leading to greater psychological distress than was observed among their counterparts who immigrated to Israel. Immigrants from the former Soviet Union to Israel had a higher social status upon arrival than immigrants from Asian and North African countries, represented a larger proportion of the total population, and their Jewish identity was not questioned. On the other hand, Soviet immigrants to the United States experienced diminished social status after immigration compared with their premigration expectations, and had to prove their Jewish identity in the United States through participation in communal and religious activities foreign to them during their lives in the former Soviet Union.

In the 19th century, Edward Jarvis's studies of Irish immigrants to Massachusetts suggested that they had a higher prevalence of mental illness than the general population of the state. By contrast, a number of epidemiological studies conducted since 1980 have found that foreign-born individuals are at lower risk for specific mental disorders other than schizophrenia. Table 4.3–5 provides an overview of a number of current studies that indicate a lower prevalence rate for mental disorders among immigrants, in contrast to the view held since Jarvis's seminal work, that immigrants had a higher prevalence of mental illness than the general population. The National Epidemiologic Survey on Alcohol and Related Conditions found that foreign-born Mexican Americans and foreign-born non-Hispanic whites were at significantly lower risk of DSM-IV-TR substance use and mood and anxiety disorders compared with U.S.-born subjects. The risk of specific mental disorders was noted to be similar between foreign-born Mexican Americans and foreign-born non-Hispanic whites. However, U.S.-born Mexican Americans were at significantly lower risk of psychiatric morbidity than U.S.-born non-Hispanic whites. This

finding has been replicated among Asian immigrants to the United States; Asian American immigrants had lower rates of mental disorders than their U.S.-born counterparts. The NCS also found higher rates of mental disorders in Hispanics born in the United States than in those born in Mexico, in particular for diagnoses of phobia and depression. A similar finding comparing Mexican Americans born in the United States with those born in Mexico was noted in the Los Angeles site of the Epidemiological Catchment Area, where there were lower rates in the foreign born. The findings were similar for young adult Cuban and other Hispanic immigrants to Florida. Increasing acculturative stress is believed to increase the prevalence of mental and substance use disorders among Hispanic American subjects in the NCS. The Mexican American Prevalence and Services Survey found that immigrants who had been living in the United States more than 13 years had higher rates of mental disorders than those who had been living in United States less than 13 years. This survey also found the rate of mental disorders for U.S.-born Mexican Americans to be 48 percent overall, whereas those born in Mexico had a rate of 25 percent for psychiatric and substance use disorders. These studies raise important questions as to why socialization into American culture increases immigrants' vulnerability to mental disorders.

Over the past several years there have been three major studies that have examined the mental health of immigrant populations compared to nonimmigrants (Tables 4.3–5 and 4.3–6). The National Survey of American Life examined the rates of disorders among Caribbean black immigrants to those who were second- and third-generation born in the United States. First-generation Caribbean black immigrants had lower rates of mental disorders compared to those of the second generation who were born in the United States. Among women this difference was limited to rates of substance use disorders. Among these immigrants, increasing years of residency in the United States was associated with an increased risk for mental disorders. As was found with Hispanic Americans, increasing generations who lived in the United States was associated with increased risk for disorder; third-generation Caribbean-American blacks had the highest rates.

Immigration was also found to play a role in the mental health of Asian American immigrants in the NLAAS. Asian women born in Asia were less likely than those born in the United States to have a lifetime psychiatric diagnosis. Second-generation Asian American women were at a higher risk for a psychiatric diagnosis than the first generation. This study found a different pattern, however, among Asian American men. Among Asian American men generational differences in rates of mental disorder were not noted, and the impact of time in the United States was inconsistent. Interestingly, Asian American men who were more proficient in English had lower rates of mental disorders. English-language proficiency may be a

proxy for the ability of an immigrant to interact and function effectively in the majority community, leading to expanded opportunities in educational, occupational, and social spheres of activity, thereby strengthening these immigrants' functional bicultural identity.

Table 4.3–5.
Lifetime Prevalence Rates in U.S. Surveys Examining Immigration (%)

	NEASRC						MAPSS		Florida			
	Hispanic American		Asian American		White		Mexican American		Cuban American		Other Hispanic American	
	FRG	U.S.	FRG	U.S.	FRG	U.S.	FRG	U.S.	FRG	U.S.	FRG	U.S.
MDD	7.7	15.2	8.2	11.7	12.0	18.2	5.2	14.8	17.8	19.0	18.0	18.3
DYS	1.7	2.8	2.6	4.3	3.1	4.6	1.9	5.2	0	1.7	0	0.5
BIP	3.5	7.0	2.7	3.5	4.5	5.5	1.1	2.8				
GAD	1.5	3.3	2.0	3.4	3.2	4.7			0.8	1.0	1.2	2.6
PAN	1.3	5.0	1.6	5.2	3.4	5.7	2.0	1.4	1.0	3.0	2.1	3.6
PTSD									8.9	8.4	9.4	15.9
ASP									6.8	6.4	8.7	12.8
ALC	15.3	30.5	7.3	24.5	16.2	35.0	9.3	22.2	28.1	29.0	24.8	31.2
DRG	1.7	12.0	2.3	8.3	4.8	11.6	3.9	16.2				
ANY	28.5	47.6	8.0	26.9	32.3	52.5	24.9	46.2	59.3	60.9	56.5	71.3

NEASRC, National Epidemiologic Survey on Alcohol and Related Conditions; MAPSS, Mexican American Prevalence and Services Survey; FRG, foreign born; U.S., U.S. born; White, non-Hispanic white; MDD, major depressive disorder; DYS, dysthymia; BIP, bipolar disorder; GAD, generalized anxiety disorder; PAN, panic disorder; PTSD, posttraumatic stress disorder; ASP, antisocial personality disorder; ALC, alcohol abuse/dependence; DRG, drug abuse/dependence; ANY, any mental disorder.

Table 4.3–6.
Risk among Immigrants and Mental disorders by Nativity and Number of Generations in the United States (odds ratios)

	Depressive Disorders		Anxiety Disorders		Substance Use Disorders		Overall Mental Disorders	
	Men	Women	Men	Women	Men	Women	Men	Women
NATIVITY STATUS								
Hispanic foreign born[a]	0.75	0.81	0.69	0.81	0.31[c]	0.07[c]	0.45[c]	0.62[c]
Caribbean blacks foreign born[b]	0.33	0.65	0.32[c]	0.63	0.15[c]	0.18[c]	0.23[c]	0.55
Asians foreign born[a]	0.91	0.50[c]	0.90	0.47[c]	0.31[c]	0.13[c]	0.61	0.53[c]
GENERATIONAL STATUS								
Hispanic Americans[a]								
Second	1.18	0.61	0.97	0.91	2.86[c]	—	1.33	0.93
Third or later	2.16[c]	1.64	1.93	1.26	2.85[c]	—	1.97[c]	1.75[c]

Caribbean blacks[b]								
Second	2.10	0.95	2.02	0.89	8.70[c]	4.60[c]	3.06[c]	0.93
Third or later	5.06[c]	4.05[c]	3.02	3.02[c]	6.49[c]	13.47[c]	6.83[c]	4.99[c]
Asian Americans[a]								
Second	0.56	2.50[c]	1.15	2.11	2.24	7.60[c]	1.31	2.07[c]
Third or later	2.04	1.40	1.01	2.17	3.80[c]	9.50[c]	2.13[c]	1.61

[a]National Latino and Asian American Study.

[b]National Survey of American Life.

[c]Denotes $p < 0.05$.

The reference group for nativity (OR = 1) are those born from the same ethnic-racial group in the United States. The reference group for generational status (OR = 1) are first-generation immigrants from the same ethnic-racial group to the United States.

In the analysis of Hispanic American immigrants, the NLAAS had findings similar to the earlier Mexican American Prevalence and Services Survey. Immigrant Hispanic American men and women had lower rates of substance use disorders and overall mental disorders compared to Hispanics born in the United States. The more proficient in English, the more they were biculturally adapted and the higher were the rates of mental disorders. Both second- and third-generation Hispanic Americans had higher overall rates of mental disorders than the first-generation subjects. Further analysis of NLAAS data among Hispanic Americans found that differences in rates of disorders by nativity and ethnicity might be in part explained by a number of factors relating to acculturative stress. These include family burden in the form of increased financial and emotional demands by extended family, greater frequency of intergenerational family conflicts, perceived low neighborhood safety, exposure to discrimination, disrupted marital status, underemployment and unemployment, and perceived low social status. Increased family cultural conflict and family burden were associated with increased risk for depressive and anxiety disorders. These findings showed that family dysfunction and ineffective social support among immigrants predict depression, and that family harmony is important to counter depression. Furthermore, these findings suggested that it was not nativity that protects immigrants from mental illness once they arrive in the United States, but rather family cohesion and support, the contextual environment (absence of overt discrimination, neighborhood safety, and religious participation), and social status associated with nativity and age of arrival in the United States.

The third study, the NCS-R, reported the risk of mental disorders among all immigrants compared to all individuals in the study who were born in the United States. Among immigrants, the NCS-R study found rates of 13.1 percent for mood disorders, 22.5 percent for anxiety disorders, 6.1 percent for substance use disorders, and 34.8 percent for any mental

disorder. By comparison U.S.-born subjects had rates of 21.3 percent, 28.6 percent, 15.0 percent, respectively, for mood disorders, anxiety disorders, and substance use disorders, and 46.8 percent for any mental disorder. The investigators found that immigrants had a lower lifetime risk of having a mental disorder compared to those born in the United States. This risk was inversely related to age at immigration and directly related to the duration of residence in the United States. These findings suggest that both early age of immigration and the duration of living in the United States contribute to increased risk for mental disorders among them, as they attempt to cope with the acculturative stress inherent in their effort to adapt to American society.

Two studies have been conducted to examine rates of mental disorders among immigrants compared to individuals who did not emigrate from their country of origin, and they provide additional insights into the relationship between immigration and mental illness. Both studies compared Mexican Americans in the United States to Mexicans living in Mexico, using psychiatric epidemiological surveys conducted around the same time, using similar methodologies, although by different investigators. The two studies have somewhat contrasting results. The first study, Mexican American Prevalence and Services Survey, found lifetime prevalence of mental disorders to be lowest among immigrants living 1 to 12 years in the United States. This study also found intermediate rates among residents of Mexico City and the highest rates among immigrants who have been living in the United States for 13 or more years. The results of this study were interpreted as showing that immigrants have less mental illness during their first 12 years living in the United States, in comparison with their counterparts living in Mexico City, possibly due to selective migration of people with good mental health, but that this advantage reverses with time, possibly as a consequence of acculturative stress.

The second study was a comparison between Mexican American immigrants from the NCS-R and Mexicans living in Mexico, from the Mexican NCS. Unlike the earlier study, immigrants to the United States from Mexico were found to have significantly higher lifetime and 12-month prevalence rates of mood and anxiety disorders than the Mexican sample, with a lifetime prevalence of any anxiety or mood disorder twice as high among immigrants as Mexican residents and a 12-month prevalence nearly three times as high among Mexican American immigrants as compared with subjects living in Mexico. The NCS-R study found that pre-existing anxiety disorders predicted immigration. Immigration predicted subsequent onset of anxiety and mood disorders and persistence of anxiety disorders. The results were inconsistent with the *healthy immigrant hypothesis* that mentally healthy people immigrate. Arguably, the *migration morbidity hypothesis* is not supported as U.S. immigration policies may play a role in preventing unhealthy immigrants from

immigrating to the United States. The findings were partially consistent with the *acculturative stress hypothesis;* that stresses of living in a foreign country, especially one that has negative attitudes about immigrants from particular countries and cultures, contribute to mental illness. In addition the NCS-R found that the elevated risk among immigrants in the United States was more pronounced for those immigrants who came to the United States as children, compared with those who arrived later in life.

Immigration and Schizophrenia

Ødegaard showed that the rate of first admissions for schizophrenia among Norwegian immigrants to the United States was twice as high as that for native-born Americans and for Norwegians living in Norway. Ødegaard attributed this difference largely to *social selection,* an inference that is still actively debated. Since Ødegaard's observations in the 1930s, other studies have produced robust findings that support social selection factors in the etiology of schizophrenia. Other researchers, however, increasingly have argued for *social causation* factors. The *social causation hypothesis* suggests that stressful conditions of life postmigration account for the increased risk of disorder in the lower socioeconomic class immigrants. Such conditions include discrimination, underemployment and unemployment, and stress related directly to poverty. This is in contrast to the *social selection hypothesis,* where as a result of disorder a person loses social status or fails to rise out of low socioeconomic status in upwardly mobile societies.

Much of the basis of the argument for the social causation hypothesis derives from studies that found higher rates of schizophrenia and lower rates of less severe mental disorders among West Indian immigrants to the United Kingdom. This pattern was suggestive of culturally determined patterns of response to adversity. Subsequent studies have substantiated this finding of an increased risk of schizophrenia among immigrants, especially among African Caribbean immigrants to the United Kingdom, as well as among Surinamese, Caribbean, and Moroccan immigrants to the Netherlands, and among immigrants to Denmark and Sweden. In the United Kingdom, it was also noted that second-generation African Caribbean immigrants had significantly higher admission rates for schizophrenia than both their first-generation and white counterparts, whereas rates of the disorder in their countries of origin in the English-speaking Caribbean were not unduly high. Those and other studies identified in a meta-analysis found both an increased risk for schizophrenia among first- and second-generation immigrants, with the latter being higher, and that immigrants from developing countries as compared to those from developed countries were at a higher risk. These findings have implicated the social environment as the source for a number of putative factors for schizophrenia, including discrimination.

A recent study conducted in Israel confirmed the findings of an increased risk for schizophrenia among immigrants, with the highest rates among those coming from less developed countries, who may face more acute discrimination. However, the finding of higher rates among the second generation compared to the first, as noted in the African Caribbean studies, could not be replicated. Additionally, the risk of schizophrenia among immigrants appears to be greater in countries with a longstanding versus recent history of immigration.

Mental Health of Refugees in the United States

Unlike those who emigrate voluntarily, often for primarily economic reasons, refugee populations have been highly traumatized prior to migration and are at a higher risk for mental disorders. Among Southeast Asian refugees, premigration trauma was a significant factor in predicting psychological distress even 5 years after immigration to the United States. Some studies have suggested that up to 70 percent of Vietnamese, Cambodian, and Laotian refugees have PTSD. A community survey among Cambodian refugees found that 45 percent had PTSD and 51 percent had depression. Cambodian adolescent refugees also have been found to have high rates of depression; 41 percent had depression even 10 years after their traumatic exposure. Vietnamese adult refugees to the United States have been found to be at higher risk for depression compared with Chinese refugees. Even two decades after resettlement, Cambodian immigrants have been found to have extremely high rates of PTSD (62 percent) and major depression (51 percent). In addition, over 70 percent have been exposed to violence after immigration to the United States. A dose–response relationship, with increasing degree of exposure to trauma, both premigration and postmigration, and the increased likelihood of a current mental disorder have been noted. Somewhat surprisingly, the largest study conducted on Cambodian refugees using a structured diagnostic instrument found that women were no more likely than men to develop PTSD or depression.

Studies of more recent refugee groups, such as the Bhutanese have found that being a family provider, poorer self-reported health, and illiteracy in their native language are risk factors for depression. Among Iraqi refugees living in the United States between 8 and 36 months, 27 percent were found to have depression and 31 percent PTSD. The role of mental health professionals is not well understood among recent refugees, such as those from the Congo and Somalia. A focus group study of refugees from Bhutan, Burma, Ethiopia, and Somalia found that refugees were willing to describe their symptoms in their political context, and that PTSD symptom expression varies across cultures. This study highlighted the need to ask refugees about their political history and struggles, and not only symptoms, leaving the refugee feeling stigmatized. The concern of

retraumatization is controversial. The premigration life of refugees from war-torn developing countries ill prepares them for the difficulties encountered upon immigration, due to having limited occupational skills, limited educational level, barriers to cultural integration including language, and significant premigration health and mental health problems; thereby placing them in a continuing position of psychosocial vulnerability and at risk for chronic mental illness.

FUTURE DIRECTIONS FOR RESEARCH IN TRANSCULTURAL PSYCHIATRY

There are three perspectives, among other possible approaches, that offer great promise for future research in cultural psychiatry. The first would be based on identification of specific fields in general psychiatry that could be the subject of focused research from a cultural perspective. Topics of epidemiology and neurobiology could be assessed in this way. The former would address issues primarily in the public health arena, including stigmatization, racism, and the process of acculturation. A number of cultural variables should be considered in conducting cultural psychiatry research, including language, religion, traditions, beliefs, ethics, and gender orientation.

The second option would aim at the exploration of key concepts and/or instruments in culturally relevant clinical research. There are four key concepts: *Idioms of distress, social desirability, explanatory or causal attribution models, and ethnographic data. Idioms of distress,* already defined, are specific, unique ways in which different cultures or societies report ailments, behavioral responses to threatening or pathogenic factors, and descriptive nomenclature of stress and stressors. *Social desirability* stems from the similarities or differences among cultures *vis-à-vis* the actual experiencing of stressful events. Members of some cultures may be more or less willing to show physical or emotional problems, thus showing different levels of vulnerability or resignation, resilience, or acceptance. Issues of stigma in different cultural contexts contribute to this level of desirability or rejection. Third, *explanatory models or causal attributions,* reflect he distinctive way in which each culture explains psychopathologies (or pathology of any kind). The explanation includes not only the presumptive original cause, but also the impact of the adduced factors and the interpersonal exchanges and interactions that lead to the culturally accepted clinical diagnosis. The fourth concept, *ethnographic data* should be included, together with strictly clinical data, as well as narratives of life that enrich the descriptive aspects of the condition and expand on the surrounding sociocultural and interpersonal and environmental aspects of the experience. The reemergence of the so-called narrative medicine is an effort to vindicate the true interpersonal nature of any clinical encounter, and the cultural spontaneities very much threatened by the current "technologization" of clinical practice.

The last approach in these efforts attempts to combine the first two by examining different clinical and research areas on the basis of the clinical dimensions of cultural psychiatry. This deals with conceptual, operational, and topical issues in the field, including their biocultural connections. From a strictly clinical perspective, this resembles current discussions about issues such as integrated care, behavioral medicine, personalized medicine, or global health—an imposing set of challenges aimed at improving the mental health of all people.

Conceptual Issues in Cultural Psychiatry

One of the primary issues in general psychiatry research is the conceptual differentiation between culture and environment. Although generally accepted as the conceptual opposite of genetic, environment represents a very broad, polymorphic concept. It is therefore important to establish that, while perhaps part of that environmental set, culture and cultural factors in health and disease are terms of a different, even unique, nature.

To what extent does culture apply to the clinical realities of psychiatry? Culture plays a role in both normality and psychopathology. The role of culture in psychiatric diagnosis is an excellent example of this conceptual issue, as demonstrated by the content of DSM-5. Furthermore, culture has an impact on treatment approaches, based on both conventional medical and psychiatric knowledge, and on the explanatory models. Finally, cultural variables have a role in treatment outcome and prognosis.

A conceptual debate exists between those who advocate an *evidence-based* approach to research and practice, in contrast to those who assign a *value-based* view to everything clinical, more so if influenced by cultural factors. The value-based approach invokes moral issues that touch on areas such as poverty, unemployment, internal and external migration, and natural and manmade disasters. Evidence may be found to support both positions in scientific research.

Operational Issues in Cultural Psychiatry

The dichotomy of normality and abnormality in human behavior is a crucial operational issue. Culture plays a definitive role in shaping these approaches. This raises the notion of relativism, a strong conceptual pillar in cultural psychiatry. Normality is a relative idea; that is, it varies in different cultural contexts.

Another operational issue is that of the choice of cultural variables. Each one has a specific weight and impact on the occurrence of symptoms, syndromes, or clinical entities in psychiatry. Some of them may be essential in the assessment of a clinical topic, namely, language, education, religion, and gender orientation. An additional operational factor is the description, assessment, and testing of the strengths and weaknesses of an individual

patient. Aspects of an individual's behavior, attitudes, disposition, sociability, occupational skills, and other factors have, likewise, strong cultural determinants.

Culture plays a significant role in the perception of severity of symptoms, the disruption of the individual's functionality, and his/her quality of life. The assessment of severity is also the result of the meanings attributed to causal or pathogenic factors of psychopathology. Judgments about level of dysfunction and the quality of a patient's life involve the elusive description, content, and meaning of concepts such as happiness, well-being, and peace of mind.

Research on cultural psychiatry issues needs to take into account representativeness of the study populations and generalizability of the findings. Methodological rigor must be strongly applied to the collection of demographic data, delineation of and differentiation between ethnic groups or subgroups, and measurement of demographic variables, symptoms, diagnosis, and culturally specific constructs.

Many tests and questionnaires used in clinical settings and research have been developed on English-speaking Western subjects and may not be appropriate for use among ethnic minority patients or non-English–speaking individuals due to lack of cultural equivalence. Translating items is insufficient to achieve linguistic equivalence, as the meaning and connotation changes and idioms of expression differ between languages. In addition, norms also may differ between ethnic groups, and tests need to be standardized with representative patient samples.

The complexity of translating an instrument varies depending on how much the construct being measured differs between the two cultures. There are four different approaches to translation. An ethnocentric approach is one in which the researcher assumes that the concepts completely overlap in the two cultures. This is described as the *etic* approach. The instrument is used with individuals who differ from the population in which the instrument was originally developed and normed. The pragmatic approach assumes that there is some overlap between the two cultures and attempts are made to measure the overlapping aspects of the construct. This is the *emic* approach. An emic plus etic approach goes one step further and also attempts to measure culture-specific aspects of the construct. Lastly, sometimes translation is not possible when the concepts do not overlap at all within the two cultures.

Topical Issues in Cultural Psychiatry

There are five dimensions of the clinical process that are relevant to research in cultural psychiatry. These include the consideration of culture as an interpretive or explanatory tool of human behavior, a pathogenic or pathoplastic agent, a diagnostic and nosological instrument, a therapeutic or protective factor, and a service or management element. The basic

assumption is that culture impacts each of these areas, all of which have relevance at different stages of the clinical encounter between patient and clinician.

Culture as an Interpretive or Explanatory Tool. The use of culture in the interpretation and explanation of behaviors and/or symptoms is based on the premise that it reflects aspects of the individual's relationship with his or her cultural milieu. *Explanatory models* are the ways in which individuals in different cultures see the core reasons for their suffering once they have identified their experience as abnormal, morbid, or pathological. The main instrument for researching explanatory models is the ethnographic narrative. This is the emic approach, describing elements of belief and behavior that are distinct for each cultural group, which includes the analysis of the individual's original language style, intensity of the experience, use of metaphors, and culturally distinctive stories, legends, traditions, beliefs, and rituals. In many cases, this is a valid clinical strategy to depathologize behaviors that a conventionally trained clinician may consider truly morbid.

Culture as a Pathogenic or Pathoplastic Agent. Culture may be a *pathogenic* agent in the construction of a clinical picture. It may contribute to the production of symptoms. Life events, for instance, colored by distinctive cultural characteristics in the patient's milieu or life setting, represent a significant topic of research. This is similar, but not identical, to the intrafamily transactions across different stages of the life cycle. Child-rearing practices, in addition to having a beneficial, protective effect, may also have a pathogenic potential. Contextual issues may be both pathogenic and *pathoplastic,* the latter being the way symptoms are shown or expressed, according to the historical period and social circumstances in which they occur. Culture also influences personality and temperament, human settings of normal and pathological behavior.

Culture as a Diagnostic or Nosological Tool. The DSM-5's Outline for Cultural Formulation, CFI, and Cultural Concepts of Distress have been developed not only to be frequently used clinical instruments, but also to become valuable research tools. The number of academic centers where DSM-IV's cultural formulation was regularly used both in clinical practice and in research was small. Time will tell whether the CFI will be used broadly, both clinically and in research. Research of its different components would improve its validity, emphasize intra- and intercultural variability, and avoid cultural stereotyping, thereby increasing the instrument's clinical usefulness and research value for psychiatry, as well as for other medical disciplines. The combination of research using narrative and quantitative approaches in national and international

collaborative studies could produce important results.

Furthermore, research on diagnosis should address the issue of population-based mental health agendas, the impact and removal of stigma in different population groups and subgroups, the promotion of greater diagnostic precision, and comprehensiveness. Addressing cultural issues is necessary for the success of the evolving global mental health agenda. The challenges of culture in the diagnostic and nosological areas call for close interdisciplinary collaboration, the reconceptualization of topics such as cultural and ethnic identity, and the development of longitudinal views of illness, meaning, and adaptation.

Culture as a Therapeutic or Protective Factor. This dimension centers on the positive impact of culture in the assessment and management of psychiatric conditions. It considers that culture may be, in fact, a healing force. As such, a great variety of cultural psychotherapies have emerged. The use of religion and spirituality as clinical instruments for treatment and enhancement of the so-called protective factors is acquiring growing relevance around the world.

Culture as a Service or Management Element. Issues influencing access to services—availability, accessibility, affordability, acceptability, and accountability—require better evidence of their impact, relevance, and usefulness. The use of providers who belong to the same ethnic group as the patient is a matter of continuous debate. The setting in which the clinical encounter occurs deserves more research. Issues of help-seeking patterns vary according to culturally acceptable practices, including the role of figures of authority, relationship with health providers, status of the help seeker (immigrant, second generation, refugee, victim of natural disaster, urban or rural inhabitant), all need additional study. The very broad topic of cultural competence of service settings, and of their clinical and administrative staff, and its impact on care needs further evaluation.

Biocultural Links in Psychopathology. A number of the variables mentioned, topics, and areas of research could be applied to the biological and cultural connections in different psychopathological conditions. The following are two examples of the areas in which this integrative perspective may lead to advances in the field.

The area of trauma and the clinical entity known as PTSD involve an enormous cultural component that has to do not only with issues of resilience and vulnerability, but also with the context of symptomatological or syndromic pictures. In addition to studies on war-related trauma, new areas including sequelae of domestic or sexual violence, child abuse, increased street delinquency, and criminal behavior reflect this convergence of factors in need of study. The severity, expression of

symptoms, and management of the different features of trauma and its psychopathological reflections are based on both cultural roots and biological connections. Predisposition to symptoms may well depend on genetic vulnerability, but this will not operate without triggering cultural or environmental factors of relevance. Variations in neurohormonal and neurotransmitter levels, neural circuits, and functional localization of emotional responses in the brain and other parts of the central nervous system have both clinical and cultural correlates.

A clinically relevant understanding of gene–environment interactions is important to understand the clinical and cultural link not only in the area of psychotherapy, but also from a biocultural perspective, in the areas of ethnopsychopharmacology and pharmacogenomics. The pharmacogenomic profiles for P450 enzymes, their genes and subsequent metabolic capabilities regarding psychotropic agents, as well as those of serotonin transporter and receptor genes are different among various ethnic groups. Thus, the study of evolutionary, chronological, and geographical considerations in these two fields across the world is an emerging field of research.

References

Abe-Kim J, Takeuchi DT, Hong S, et al. Use of mental health-related services among immigrant and US-born Asian Americans: Results from the National Latino and Asian American Study. *Am J Public Health*. 2007;97:91–98.

Alarcón RD. Cultural psychiatry: a general perspective. *Adv Psychosom Med*. 2013;33:1–14.

Alegría M, Mulvaney-Day N, Woo M, Torres M, Gao S, Oddo V. Correlates of past-year mental health service use among Latinos: results from the National Latino and Asian American Study. *Am J Public Health*. 2007;97:76–83.

Beals J, Novins DK, Whitesell NR, Spicer PP, Mitchell CM, Manson SM. Prevalence of mental disorders and utilization of mental health services in two American Indian reservation populations: mental health disparities in a national context. *Am J Psychiatry*. 2005;162:1723–1732.

Berry JW. Conceptual approaches to acculturation. In: Chun KM, Organista PB, Marin G, eds. *Acculturation; Advances in Theory, Measurement and Applied Research*. Washington, DC: American Psychological Association; 2003:17–37.

Breslau J, Aguilar-Gaxiola S, Kendler KS, Su M, Williams D. Specifying race-ethnic differences in risk for mental disorder in a USA national sample. *Psychol Med*. 2006;36:57–68.

Burnett-Zeiler I, Bohnert K, Illgen MA. Ethnic identity, acculturation and the prevalence of lifetime psychiatric disorders among black, Hispanic, and Asian adults in the U.S. *J Psychiatr Res*. 2013;47:56–63.

Cantor-Graae E, Selten JP. Schizophrenia and migration: a meta-analysis and review. *Am J Psychiatry*. 2005;162:12–24.

*Cohen A, Patel V, Thara R, Gureje O. Questioning an axiom: better prognosis for schizophrenia in the developing world? *Schizophr Bull*. 2007;34:229–244.

*Collins PY, Patel V, Joestl SS, March D, Insel TR, Daar AS, et al. Grand challenges in global mental health. *Nature*. 2011;475:27–30.

Comas-Diaz L, Jacobsen FM. Ethnocultural transference and countertransference in the therapeutic dyad. *Am J Orthopsychiatry*. 1991;61:392–402.

*Committee on Cultural Psychiatry Group for the Advancement of Psychiatry. *Cultural Assessment in Clinical Psychiatry*. Washington, DC: American Psychiatric Publishing; 2002.

Cook BL, Liu Z, Lessios AS, McGuire T. The costs and benefits of reducing racial-ethnic disparities in mental health care. *Psychiatr Serv*. 2015;66:389–396.

Eaton J, McCay L, Semrau M, Chatterjee S, Baingana F, Araya R, et al. Scale up of services for mental health in low-income and middle-income countries. *Lancet*. 2011;378:1592–1603.

Gee GC, Ryan A, Laflamme DJ, Holt J. Self-reported discrimination and mental health status among African descendants, Mexican Americans, and other Latinos in the New Hampshire REACH 2010 Initiative. The added dimension of immigration. *Am J Public Health*. 2006;96:1821–1828.

Horvitz-Lennon M, McGuire TG, Alegria M, Frank RG. Racial and ethnic disparities in the treatment of a Medicaid population with schizophrenia. *Health Serv Res*. 2009;44:2106–2122.

Jackson JS, Neighbors HW, Torres M, Martin LA, Williams DR, Baser R. Use of mental health services and subjective satisfaction with treatment among black Caribbean immigrants: Results from the National Survey of American Life. *Am J Public Health*. 2007;97:60–67.

Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. *Bull World Health Organ*. 2004;82:858–866.

*Lewis-Fenandez R, Aggarwal NK, Hinton L, Hinton DE, Kirmayer LJ. *DSM-5 Handbook on the Cultural Formulation Interview*. Washington, DC: American Psychiatric Publishing; 2015.

Lin K-M, Kleinman AM. Psychopathology and clinical course of schizophrenia: A cross-cultural perspective. *Schizophr Bull*. 1988;14:555–567.

*Lora A, Kohn R, Levav I, McBain R, Morris J, Saxena S. Service availability and utilization and treatment gap for schizophrenic disorders: a survey in 50 low- and middle-income countries. *Bull World Health Organ*. 2012;90:47–54.

Marshall GN, Schell TL, Elliott MN, Berthold SM, Chun CA. Mental health of Cambodian refugees 2 decades after resettlement in the United States. *JAMA*. 2005;294:571–579.

Rodriguez-Seijas C, Stohl M, Hasin DS, Eaton NR. Transdiagnostic factors and mediation of the relationship between perceived racial discrimination and mental disorders. *JAMA Psychiatry*. 2015;72:706–713.

Rost K, Hsieh YP, Xu S, Menachemi N, Young AS. Potential disparities in the management of schizophrenia in the United States. *Psychiatr Serv*. 2011;62:613–618.

Salas-Wright CP, Kagotho N, Vaughn MG. Mood, anxiety, and personality disorders among first and second-generation immigrants to the United States. *Psychiatry Res*. 2014;220:1028–1036.

Sohler NL, Bromet EJ, Lavelle J, Craig TJ, Mojtabai R. Are there racial differences in the way patients with psychotic disorders are treated at their first hospitalization? *Psychol Med*. 2004;34:705–718.

Takeuchi DT, Zane N, Hong S, Chae DH, Gong F. Immigration-related factors and mental disorders among Asian Americans. *Am J Public Health*. 2007;97:84–90.

U.S. Department of Health and Human Services. *Mental Health 2001: Culture, Race, and Ethnicity—A Supplement to Mental Health: A Report of the Surgeon General.* Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services; 2001.

Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet*. 2007;370:841–850.

Williams DR, González HM, Neighbors H, Nesse R, Abelson JM. Prevalence and distribution of major depressive disorder in African Americans, Caribbean blacks, and non-Hispanic whites: Results from the National Survey of American Life. *Arch Gen Psychiatry*. 2007;64:305–315.

5

Quantitative and Experimental Methods in Psychiatry

▲ 5.1 Quantitative and Experimental Methods in Psychiatry: Epidemiology

WILLIAM E. NARROW, M.D., AND DIANA E. CLARKE, PH.D.

BACKGROUND

Epidemiology provides a description of how often disease occurs in populations, the rates at which they change through time, and the factors that explain their occurrence. Because it is a field that studies populations as contrasted to individuals it provides useful information for public health decisions about prevention, treatment, and social costs of illnesses. This chapter focuses on psychiatric epidemiology, where the illnesses of interest are psychiatric disorders. In view of the fact that mental disorders more closely follow a chronic pattern, analytical models used in psychiatric epidemiology resemble those used in other chronic illnesses, as opposed to classic acute or infectious disease models.

In 1957, Morris described the uses of epidemiology, which are still relevant despite the passage of close to six decades, and can be readily applied to psychiatric epidemiology:

Historical study—Findings from epidemiological studies can be compared through time. A classic example is the finding of a shift in the major causes of mortality that accompanies economic development, from infectious diseases and infant and maternal mortality, to chronic diseases such as cardiovascular disease and cancer. For psychiatric disorders, such historical comparisons are difficult due to radical changes in diagnostic customs since the 1970s and a relatively recent focus on community diagnosis, as opposed to diagnosis in treated populations such as hospital discharge diagnoses. Still, active research with a few historical data sets is examining such issues as a possible increase in rates of depression

over time, particularly in younger age groups. Another use of historical data makes use of changes in population demographics by projecting future rates of disorders that are linked to such factors as age. For example, with the aging of the U.S. "baby boom" population, rates of age-related disorders such as major neurocognitive disorder due to Alzheimer disease can be projected for future years and compared to current rates.

Community diagnosis—Epidemiological studies provide a picture of the health status, morbidity, and mortality characteristics of a population and its subgroups. These data can then be used to identify and characterize vulnerable populations and critical health service needs. The burden of an illness to society can be estimated in terms of the reduced worker productivity, treatment costs, and the repercussions of premature mortality. The World Health Organization Global Burden of Disease project is discussed at the end of this chapter.

Working of health services—Epidemiological studies can provide valuable information on a population's treatment needs and its current level of treatment demand. Demand for treatment is here defined as the current level of service use in the population. Epidemiological studies performed since the 1970s have documented a "defacto mental health service system" in the United States that is characterized not only by traditional specialty treatment settings but also high levels of service use in the primary medical care, social services, school, and alternative medicine settings. The primary care setting has been an especially important and highly used source of mental health care in the United States and around the world, highlighting the need for workers in these settings to be familiar with accurate psychiatric diagnosis and treatment strategies.

Treatment need for persons with psychiatric disorders, as opposed to many other medical disorders, has not been well defined for population-level purposes. For medical disorders such as cancer or bone fracture or sepsis, the assignment of a diagnosis implies treatment need. Psychiatric disorders, on the other hand, are often on a continuum with normal feeling states such as sadness or anxiety, and the threshold between normal reaction and disorder is often unclear. Further, psychiatric conditions that do not reach full diagnostic criteria may still be associated with significant distress or disability in an individual. Cases may also spontaneously remit without treatment, particularly mild cases. For these reasons, the definition of need for treatment is an unresolved concept in psychiatric epidemiology. Various combinations of symptoms and symptom severity, levels of distress and disability, and duration of

illness have been used in such definitions. These definitions also carry implications for defining the concepts of treatment and prevention. No matter how treatment need is defined, however, it has been well documented that in the United States and other countries, a large unmet need for treatment exists.

Completing the clinical picture—Community-based epidemiological studies are able to provide data on the natural history of a disorder, including risk and protective factors, prodromal states, onset of illness, and disease progression. The large number of persons with psychiatric disorders who need treatment but do not receive it has implications for the study of these disorders. Persons who seek treatment are likely to have more severe cases of disorder, longer duration of disorder, and have more comorbid conditions. Sociodemographic and socioeconomic factors also play a role in which cases of disorder are seen for treatment. Hence, samples drawn from clinical settings are different from those drawn from community settings and may thus give biased information about the nature of a disorder. A well-known example of this difference is referred to as Berkson bias, which describes the erroneous conclusions about risk factors that may be drawn about a disorder simply due to higher rates of comorbidity in hospitalized patients compared to nonhospitalized patients.

Identification of syndromes—Psychiatric disorders lacked well-defined diagnostic categories until the publication of the third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III) in 1980. The lack of objective validating tests for the disorders listed in the DSM means that the disorders it lists are only approximations of underlying disease states, and their diagnostic criteria are the best available hypotheses for detecting these underlying diseases. Examination of the diagnostic criteria in clinical and research settings since the release of DSM-III has suggested that the categorical, symptom-based approach to diagnosis has significant limitations, both in its ability to describe the clinical complexity of patients and to facilitate the discovery of the underlying causes of mental illnesses. Therefore, efforts were made in the development of DSM-5 to incorporate cross-cutting dimensional symptom assessments as supplements to the categorical diagnostic criteria. The Research Domain Criteria (RdoC) project, sponsored by the National Institute of Mental Health, is examining the neuroscience of basic behavioral domains, with the goals of facilitating etiological research and ultimately, improving diagnosis and treatments.

Epidemiological research can support the scientific investigation of syndromes by assessing their prevalence in the community, the disabilities associated with them, and their relationship to other

disorders. Syndromes that are highly prevalent in the community, with little associated disability, are more likely to represent normal variations of human behavior than disorders. A syndrome that is highly associated with another disorder is likely to be a variant or subtype of that disorder. Much attention has been drawn to the relationship between depression, anxiety, and somatization, which are highly comorbid in the community and treatment settings.

Assessing individual risks—The results of epidemiological studies are important in assessing the likelihood that an individual will develop a disease. A well-known example is that individuals who have high cholesterol levels, hypertension, and are smokers have an increased cardiovascular disease risk. Such findings are important for treatment and preventive interventions at the individual level as exemplified by widespread acceptance of efforts to effect changes in diet, smoking cessation, and treatment of hypertension for individuals possessing these risk factors. Epidemiological findings on risk also figure into other areas, such as actuarial determinations for insurance coverage purposes and decisions on marketing of pharmaceuticals to certain population groups. Risk and protective factors for psychiatric disorders are continuing to be discovered, but translation to individual level preventive interventions is still at an early stage. However, prevention of the adverse consequences of untreated psychiatric disorders can be effected by early treatment of these disorders.

Identifying causes—One of the ultimate goals of epidemiological investigations is to elucidate the causes of disorders. Once the cause of a disorder is found, targeted treatments and preventive interventions can be developed to help reduce the burdens of the disorder in the community. For complex conditions, such as psychiatric disorders, causes are often multifactorial and interdependent, involving genetic and environmental factors. Genetic factors thought to be involved in the causation of psychiatric disorders tend to have small effects, environmental factors are often nonspecific and not well differentiated from each other, and the mechanisms by which all of these factors interact to cause disease are generally not well elucidated. One of the best-known examples of a specific gene–environment interaction in a psychiatric disorder involves a polymorphism of the promoter region of the serotonin transporter gene and stressful life events in the development of major depression. However much more research needs to be done before these findings can be used to prevent and better treat major depressive disorder in the population.

MEASURES OF DISEASE FREQUENCY

Prevalence

A useful measure to quantify the occurrence of a specific condition in a population is the prevalence estimate, a ratio describing the number of people with a disorder in a specific population in a designated time period. The three elements important to prevalence estimation are case definition, target population, and time.

$$\text{Prevalence rate} = \frac{\text{number of cases identified in a specified period of time}}{\text{number of people in the population at the same time}}$$

Commonly used prevalence rates include:

Current (Point and 1-Month) Prevalence. This rate describes people with a disorder at a specified point in time is referred to as the point prevalence. While the point prevalence is the best estimate of the "current" prevalence of a disorder, it is difficult to estimate for mental disorders in the general population since mental disorder diagnosis requires multiple symptoms to cluster together for specific time periods, and these symptoms may wax and wane on a daily basis. Therefore the 1-month prevalence rate has been more frequently used to describe the percent of people who currently have a mental disorder. This time period describes people in a specified population whose symptoms have met criteria for a disorder at any point in the past month.

1-Year Prevalence. This rate describes people who have had a disorder at any time in the past year. Because service planning is done on an annual basis, planners are most interested in how many people can be expected to have a disorder in a given year. If an assumption is made that mental disorder prevalence rates do not change greatly from year to year, and there have been no significant demographic changes in the population of interest, a 1-year prevalence rate from a recent time can be multiplied by an estimate of the total population for a given year to determine how many people in the population can be expected to have the disorder in that year.

Lifetime Prevalence. This rate describes people who have had a disorder at any point in their lives. An accurate estimate of lifetime prevalence can be a useful tool to describe the risk of developing a disorder in a population. Lifetime prevalence is also valuable for risk factor studies. In studies of risk and protective factors, it is important to determine, as accurately as possible, who has had a disorder and who has not. The relapsing and remitting course of many mental disorders, with potentially years between episodes, makes 1-month and 1-year assessments less valuable for this purpose: true cases would be counted as noncases. Lifetime prevalence is not helpful for service planning, as it will count as

cases single episodes of disorders that occurred in the remote past and have little chance of recurring in the next year. Lifetime prevalence rates are also more subject to recall bias than are prevalence estimates bases on shorter time frames, that is, remotely occurring episodes of disorder are less likely to be accurately recalled than those occurring more recently, particularly if complex diagnostic criteria are used (e.g., multiple symptoms, age of onset, duration, and disability requirements). Lifetime prevalence estimates are generally lower in older age groups, raising the question of whether older generations were more protected from developing mental disorders than younger generations. The role of recall bias cannot be discounted in this deliberation: compared to younger respondents, older respondents have lived more years in which they could have developed a mental disorder, but also, on average, have had more years elapse between their episode of illness and the present time.

Caution is needed when interpreting prevalence rate differences over time. A shift in the distribution of a demographic factor in the population can cause such observed differences. For example, when migration increases the proportion of young people in a population, disorders more frequent in the elderly will have lower prevalence rates in the overall population, even though the prevalence remains stable in the elderly population itself. For this reason, statistical adjustments are made to account for such differences in the two populations being compared.

Incidence

At times epidemiologists require estimates of how many new events occurred in a period of time; for example, the number of new cases of posttraumatic stress disorder since a natural disaster. Incidence rates are the measure of choice in such instances. The numerator of the incidence rate is the number of new events arising in a specified time period; the denominator is the population at risk of having an event during that time period.

$$\text{Incidence rate} = \text{Incidence rate} = \frac{\text{number of new cases of disorder in a specified period of time}}{\text{people at risk for the disorder in the same time period (person-time)}}$$

People who already have the disorder at the start of the time period are, by definition, taken out of the denominator for they are not at risk for developing the disorder. A new case is a person who did not have the disorder at the start of the specified time period and developed it afterward. Persons without the disorder at the start of the specified time period may or may not have the disorder previously. When the new cases are limited to people who have never before had the disorder in question, the measure is

referred to as first incidence. Because most mental disorders are relatively rare in the population, the occurrence of new cases is also relatively rare, and the occurrence of first-incident cases is even more rare. This makes estimation of incidence rates difficult, requiring large populations-at-risk and relatively long periods of observation to accrue new cases. Most incidence rates are based on a 1-year time period. Longer periods of observation are also potentially subject to recall biases.

Measurement of the population at risk over a period of time involves several considerations. In reality, populations are often nonstationary, that is, the population at risk at the beginning of the study changes during the period of observation due to people entering the population after the study period started, dropping out of the study, moving without leaving new contact information, or dying. These individuals will have been observed for different time periods, but the data they contributed is included in the denominator as person-time at risk, an estimate of the actual time at risk that all persons contributed to a study. It is estimated as the product of number of people at risk multiplied by the time elapsed until they were no longer at risk (became a case) or stopped being followed.

Relationship between Prevalence and Incidence

Prevalence, incidence, and mean duration are mathematically related. The mean duration of the disorder is an indicator of the chronicity of the disorder. It is estimated as the average of the difference between age of onset and age at the last time when criteria for the diagnosis was met. Duration is affected not only by the intrinsic characteristics of the disorder, but also external factors such as receipt of treatment. The prevalence of a disorder is proportional to incidence times duration ($P \sim I * d$). This equation illustrates another reason why caution should be used in the interpretation of prevalence rates, as prevalence rates reflect not only increases in the number of new cases (incidence), but also the length of time that cases have the disorder (duration). When a treatment is found to cure a previously incurable condition, the prevalence of the condition can be expected to decrease. However, a treatment that extends the life of a chronic, incurable condition might actually raise its prevalence rate through increased duration of the condition. For example, the prevalence of AIDS in the population increased after an effective treatment was found that prolonged the lives of affected patients and reduced mortality. In such cases it is wise to employ incidence measures; relying solely on prevalence estimates makes it impossible to tell apart an increase in new cases from an improvement in survival.

Comorbidity

In psychiatry, co-occurrence of disorders, meaning individuals diagnosed

with more than one mental disorder, is frequently observed in both community and clinical populations. Random co-occurrence of disorders is possible whenever the presence of one disorder does not preclude the presence of the other disorder. For common conditions, such random co-occurrences may be quite common. By definition, random co-occurrence does not imply underlying relationships between the co-occurring disorders.

The most commonly used concept of comorbidity refers to a nonrandom co-occurrence of disorders, in which the prevalence of condition A is higher in people with condition B, than the prevalence of condition A in persons without condition B. This type of comorbidity has implications for the underlying processes and boundaries of the two disorders. For example, the comorbid disorders may share common risk factors, or may not be separate disorders at all, rather different presentations of the same disorder. At times, the apparent co-occurrence can be explained by artifacts of the criteria used to make the diagnoses. For example, high rates of co-occurring substance use disorders were observed in persons with DSM-III antisocial personality disorder. However, substance use was one of the symptoms that could be used to meet criteria for the DSM-III antisocial personality disorder diagnosis, and persons with this symptom also had an increased probability of meeting criteria for a substance use disorder diagnoses. Sampling bias may also create a false comorbidity; for example, comorbidity is higher in treated samples than community samples, but this relationship is not related to shared etiologies.

MEASURES OF EFFECT AND ASSOCIATION

Risk, Protection, and Causation

Epidemiological studies often aim at estimating the effects of putative risk and protective factors on the occurrence of disorders in the population. A risk factor increases the likelihood that a person possessing that factor will develop a specified outcome (such as a mental disorder), compared to a person who does not possess the factor. A protective factor decreases the likelihood that a person possessing that factor will develop the outcome compared to a person who does not possess the factor. A plausible model accounting for biological and psychosocial risk and protective factors is a prerequisite for the interpretation of effect measures. Quantifying these effects relies on accurate identification and differentiation of cases and noncases, unambiguous definitions of the risk and protective factors in question, and accurate identification of the exposure of each case and noncase to the factor. The findings of a risk or protection should not be viewed as synonymous with causation, although they may provide crucial clues in the search for causes. For example, female sex is an accepted risk

factor for the development of major depressive disorder. In this case one would not cite female sex as a cause of depression, rather it is measureable characteristic with a wide variety of underlying biological and associated psychosocial factors, each of which may play a more direct causal role.

In psychiatric epidemiology, exposures can be genetic or environmental. For example, being a victim of a natural disaster is an environmental exposure and being the offspring of a depressed mother is an exposure that likely has mixed genetic and environmental aspects, even though the genetic exposure is unlikely to be precisely identified. The role of specific genetic exposures in the development of psychiatric disorders is likely to be studied with increasing frequency in the future, as candidate genes for disorders are identified and genomic science advances. Without a plausible model accounting for biological and environmental factors, observed associations cannot be justified as contributing to causation.

The factors involved in the development of chronic diseases are numerous, complex, and intertwining, and the process by which they operate has been described as a "web of causation." In this web of causation, causal pathways involving the main effects of interest might be modified by other, noncausal factors. These indirect modifying effects should be considered and taken into account in the initial explanatory model, to the extent possible, since the effects of unknown and unmeasured factors are always a threat to the validity of a risk factor study. The fields of epidemiology and biostatistics have developed study designs, sampling strategies, and analytic techniques to help overcome some of these difficulties in estimation of effects, although these strategies cannot overcome the disadvantage of a poorly specified model.

Studies of causation are logistically difficult to implement and maintain and tend to be expensive. For many disorders, including psychiatric disorders, the occurrence of a new case in a population is a rare event, requiring the monitoring of large populations in order to gain a sufficient number of new cases. Also, the time span between the exposure to risk and protective factors and the onset of a disorder may be long, requiring periodic monitoring of the population to prevent drop-outs ("loss to follow-up") and detect new exposures among the previously unexposed.

Relative Risk

The classic relative effect quantification is relative risk, the ratio of two incidence rates. Table 5.1–1 lays out the distribution of exposed and unexposed persons who will go on to develop or not develop the disease of interest. Cell a represents persons with the exposure who go on to develop the illness, while cell b represents exposed individuals who have not developed the disease. Cell c represents unexposed persons who nonetheless have developed the disease, while cell d represents unexposed persons who have not developed the disease. The risk of developing the

disease for a person with the exposure is a/a + b. The risk of developing the disease in persons who were not exposed is c/c + d. Therefore, the relative risk of developing the disease for those with compared to those without the exposure is:

$$\frac{(a/a+b)}{(c/c+d)} \quad \text{or} \quad \frac{\text{incidence}_{\text{exposed}}}{\text{incidence}_{\text{unexposed}}}$$

Table 5.1–1.
2 × 2 Table for Calculation of Relative Risk

		Disease		
		Present	**Absent**	**Total**
Exposure	Present	a	b	n3
	Absent	c	d	n4
	Total	n1	n2	N

A relative risk greater than 1 suggests that the exposure is a risk factor and a relative risk less than 1 indicates that the exposure is a protective factor. A relative risk of 1 suggests that exposure has no effect on the onset of the disorder.

Odds Ratios as Approximations of Relative Risk

For reasons mentioned above, studies using incident cases are not always feasible. Still, it is possible under some circumstances to design studies where prevalent cases are sampled and risk and protective factor data are then collected (see Analytic Studies, below). In such cases, incidence is not directly measured and risk ratios cannot be calculated. However, with carefully chosen cases (disease present) and controls (disease absent) and accurate measurement of exposures, an approximation to the relative risk can be made by using an odds ratio. Table 5.1–1 can also be used to represent the different assignments that can be made in a case-control study. Cases with the exposure present are represented in cell a, cases without the exposure are in cell c. Controls with the exposure are in cell b and controls without the exposure are in cell d. In this table, the odds of having the disease, among those with the exposure is a/b. The odds of having the disease among those without the exposure is c/d. The odds ratio compares the odds of having the disease and the exposure to the odds of having the disease without the exposure.

$$\text{Odds ratio} = \frac{a/b}{c/d}$$

Simple algebraic manipulation of this formula allows a convenient way

to calculate the odds ratio, the cross-product ratio. Care must be used when calculating an odds ratio using this method, to make sure that the rows and columns of the table are set up as in Table 5.1–1.

$$\text{Odds ratio (cross-product ratio)} = \frac{ad}{bc}$$

As with the relative risk, odds ratios vary around 1. An odds ratio greater than 1 indicates an association between the exposure and the presence of the disease and an odds ratio less than 1 indicates an association between the exposure and not having the disease. An odds ratio of 1 suggests that there is no association between exposure and having or not having the disease.

Confounding. Confounding occurs when a third factor (A), other than the exposure of interest, is independently associated with the outcome (O) as well as with the exposure of interest (E) and is not in the causal pathway.

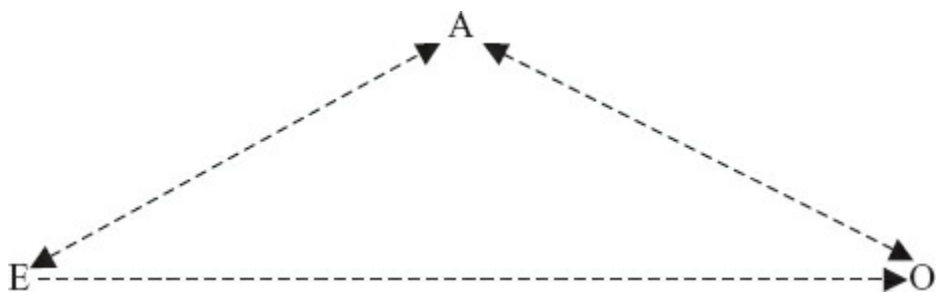

For example, studies have found that smoking is associated with suicidal behaviors. However, alcohol use is associated with smoking in individuals with and without suicidal behaviors and also is independently associated with suicidal behaviors regardless of smoking status. Therefore, alcohol use is considered a confounder of the relationship between smoking and suicidal behaviors. In order to interpret the effect of smoking on the risk of suicidal behaviors, adjustments must be made for alcohol use.

Confounding can distort the relationship between exposure and outcome in any direction, either reducing or augmenting the size of the effect. Therefore, when measuring the effect of an exposure, other factors that might explain the outcome should be measured. When potential confounders have been measured, analytic and statistical modeling techniques such as stratification and regression analysis can help disentangle the "true" contribution of the exposure from that of the confounding factors. When such techniques are used, the results of the analysis are referred to as "adjusted" or "controlled" for the confounding factors (e.g., an adjusted odds ratio for the effect of smoking on suicidal behavior controlling for alcohol use). Unmeasured confounders by definition cannot be accounted for by these techniques, and their potential

presence should always be considered when interpreting the results of an analysis.

Mediation. Mediation occurs when a third factor (A) in the causal pathway explains the mechanism through which the exposure (E) leads to the outcome (O). The mediator variable (A) must be independently related to, and occur temporally after, the exposure variable (E).

For example, cognitive therapy can reduce depression by reducing anxiety. In this example, anxiety is a mediator of the relationship between cognitive therapy and depression; it provides a mechanism by which cognitive therapy and depression are related. Mediation can be partial, where the third factor only explains some but not all of the relationship between the exposure and outcome, or complete, where the mediator explains all of the relationship between the exposure and the outcome. In mediation analysis, the relationship between exposure and outcome is broken down into direct (i.e., E directly affects O) and indirect (i.e., E affects O only through A) components.

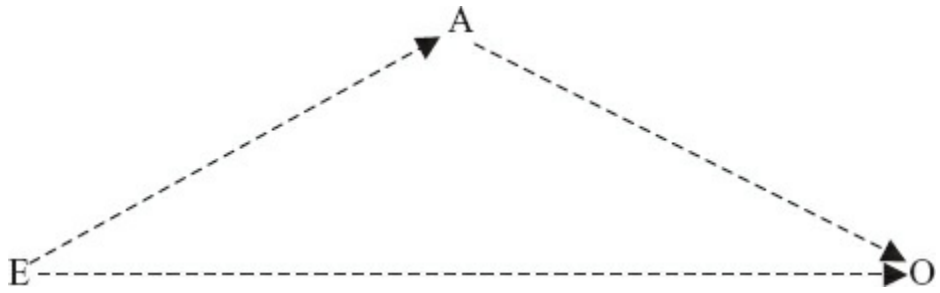

Effect Modification/Moderation. Moderation occurs when a third factor (A), at different levels (i.e., A_1 vs. A_2), impacts the relationship between the exposure and outcome differently. Furthermore, for the third factor (A) to be a moderator it must temporally precede and be independent of the exposure. For example, the relationship between antidepressant use and depression recovery is found to be different for individuals who are adherent compared to those who are nonadherent. In this example, level of adherence moderates the relationship between antidepressant use and depression recovery. Typically, stratified analyses (i.e., stratification based on different levels of A) or inclusion of an interaction term between the third factor (A) and the exposure is used to identify effect modification.

Attributable Fraction. The attributable fraction is used to study the effect of a risk factor on the incidence of a disease. Attributable fractions can be calculated for those exposed to a risk factor or for an entire population, and are also referred to as etiological fractions or population

attributable risks. The attributable fraction for those exposed to a risk factor shows how much excess disease can be attributed to the risk factor among those exposed. It is calculated by dividing the excess risk for the risk factor by the incidence rate in the exposed group.

$$\text{Attributable fraction for exposed group} = \frac{\text{incidence}_{\text{exposed}} - \text{incidence}_{\text{unexposed}}}{\text{incidence}_{\text{exposed}}}$$

The population attributable fraction tells us how much excess disease in a population can be attributed to a risk factor. It is calculated as follows:

$$\text{Population attributable fraction} = \frac{\text{incidence}_{\text{total}} - \text{incidence}_{\text{unexposed}}}{\text{incidence}_{\text{total}}}$$

If the risk factor is related to the disease in a causal way, then attributable fractions can be used to determine the amount of disease reduction that would occur if the risk factor was prevented. The prevalence of exposure to the risk factor will affect its population attributable fraction. A rare exposure, no matter how strongly related to the disease, will result in a small population attributable fraction.

EPIDEMIOLOGICAL STUDY DESIGNS

The choice of the design of an epidemiological study depends on several factors. First and foremost, the study's design must be appropriate to the research question that the study is addressing; not all types of studies can address all types of questions. At the most basic level, study designs can be divided into experimental and observational. In an experimental study, the researcher is able to control the exposure, which is usually a treatment or preventive intervention. In nonexperimental, or observational, studies the researcher is only able to observe and quantify the effect of the exposure, which cannot be manipulated as in an experimental study. Observational studies can be further characterized as descriptive or analytic.

Experimental Studies

In standard clinical trials, the exposure is a treatment (usually medication, but psychotherapies have been increasingly subjected to clinical trial designs), with the goal being to assess the *efficacy* of the treatment, that is, whether the treatment produces a desired outcome for the condition it is supposed to treat. Desired outcomes may include cure or remission of disorder, decrease in symptom levels, prevention of disability, or increased survival time. Efficacy trials tend to use highly selected subjects possessing the disorder of interest, with as few potential confounding factors as possible. The strict inclusion and exclusion criteria for enrollment into an

efficacy trial allow for a purer experimental milieu, but they do not reflect clinical reality. Over the past two decades, it has been recognized that treatment outcomes in routine clinical practice do not always mirror the results of efficacy trials, and that clinical trials should also reflect the reality of routine practice. In routine practice settings, as opposed to efficacy trials, treatments may be changed or discontinued because of nonresponse, side effects, or patient preference, and patients may have comorbid mental or general medical disorders, take multiple medications, use substances, and not follow treatment instructions to the letter. This recognition has led to the fielding of clinical *effectiveness* trials in which a wider range of subjects is included, in an effort to provide a clearer picture of how a treatment will work in clinical practice.

The standard clinical trial is controlled and randomized. Control groups receive either an inactive treatment (placebo) or a treatment already proven to be efficacious for the disorder. Control groups are necessary for comparison purposes and to account for effects that may occur regardless of the effects of experimental treatment. For example, there are often salutary effects of receiving care, even if the patient is receiving an inactive treatment; clinical improvement in the face of an inactive treatment is referred to as a placebo effect. Similarly, patients receiving inactive treatment may report side effects. The use of inactive comparison treatments is now less frequent than in previous years. Research experience has shown that there can be deleterious, long-lasting clinical consequences when psychiatric patients with a treatable disorder are not actively treated, particularly among those with severe symptoms. This raises obvious ethical issues around the use of placebos.

Random assignment of subjects to either the experimental or control condition helps to prevent bias in the outcome of the experiment, by maximizing the chance that factors related to the outcome are evenly distributed between the groups. For example, subjects with more severe symptoms may be less likely to respond to treatment; preferential assignment of them to the experimental group may bias the study results in favor of the control treatment and conversely, preferential assignment to the control group might produce a bias in favor of the experimental treatment. Another method used to increase comparability between groups is to “double blind” the study, so that neither the study staff nor the subjects themselves know to which group the assignments have been made.

The main difference between clinical trials and prevention trials is that subjects in prevention trials have not developed the disorder of interest. Subjects for a prevention trial can be selected from the community or from any other facility that is not a treatment center for the disorder of interest. Prevention trials are usually more costly than clinical trials, mainly because sample sizes required for these trials must be adequate to observe a sufficient number of outcomes, usually defined as incident cases of a

disorder. Well thought-out sampling strategies can help to maximize the number of subjects at risk for the study outcome, while still maintaining the ability to generalize the results to the population studied. Additionally, because the period of risk for developing a disorder can be lengthy, especially compared to a circumscribed treatment period, prevention trials can take longer to complete, requiring more extensive tracking of study subjects than is needed for clinical trials. Although all guidelines for subject assignment in clinical trials apply to prevention trials, at times randomization is not feasible in larger studies. For example, studies of the so-called universal interventions are applied to all members of the general population and are not feasibly randomized.

Observational Studies

There are two primary nonexperimental, observational studies in epidemiology, analytic and descriptive. Analytic studies, in which the effect of an exposure, uncontrolled by the investigator, is studied relative to a specific outcome, can be subdivided into cohort and case-control studies.

Analytic Studies. A cohort study is similar in concept to a clinical trial. In these studies, two cohorts of individuals without the outcome of interest are randomly chosen from a defined population, one cohort with an exposure and one without an exposure. These cohorts are then followed over time, and the rate at which the subjects in each cohort develop the outcome of interest is measured. By comparing the rates of the outcome in the two groups, the effect of the exposure can be estimated. Cohort studies have many advantages inherent in their design: onset of the outcome can be observed in real time (as opposed to relying on recollection) and potentially confounding variables can be measured before the onset of the outcome. However, because incident cases are being measured, and the base population prevalence of these outcomes are often quite low, the sizes of the cohorts may need to be large, increasing the cost and complexity of the study. Cohort studies are also dependent on accurate measurement of exposure: sometimes such exposures are reliably obtained through administrative records (e.g., employees working in a particular industry or in the military; individuals living in a geographic area at the time of an exposure), at other times measurement of exposure must rely on the subjects' recollections, which may be inaccurate.

In contrast to cohort studies in which selection is based on exposure, with subsequent measurement of the outcome, in case-control studies, selection is based first on having the outcome, with subsequent measurement of past exposure. Cases with the outcome of interest are selected randomly from a population and a control group without the outcome is randomly sampled from the same population that gave rise to the cases. Because the control group is used to estimate the distribution of

the unexposed and exposed individuals in the population, it is important that the controls are sampled from the source population independently of their exposure status. Frequently, given the low prevalence of many disorders and the size of the original population, all cases are chosen; statistical power of the study can also be increased by selecting more controls than cases. The rates of exposure in the case group and the control group are compared to give an estimate of the effect of the exposure in the development of the disorder. Case-control studies have the advantage of being easier to carry out than cohort studies.

Descriptive Studies. Descriptive studies, as their name implies, are aimed at describing the health status of a population. Descriptive studies typically include studies documenting the prevalence and incidence of disorders, the characteristics of the people who have or develop disorders (e.g., age, sex, race, education, income), and the use of health services in a population. They can be cross-sectional (study of a population sampled at a given point in time), repeated cross-sectional (studies of a population sampled and then re-sampled at several points in time) or longitudinal (studies of a population sampled once and then followed over time). Descriptive studies provide the basic data on which all other types of epidemiological investigations are built, as well as serving as valuable resources for service planners and policy makers. A primarily descriptive focus of an epidemiological study does not preclude analytic aims. The representative sampling of a large descriptive study can be a valuable asset for the collection of representative data for analytic studies. For example, collection of biological samples for genetic analysis could be done in a large community sample, provided that the analytic goals are well-specified, the technology for collection and analysis is available, and the study subjects provide informed consent for use of these samples.

SAMPLING AND SAMPLING STRATEGIES

Sampling refers to the process of selecting a subset (i.e., sample) of the population of interest for a research study (e.g., selecting a subset of college students from among all college students in the United States). The goal is to select a sample that is representative of the population of interest, which is determined by the specific research question. In order to select a representative sample of the population of interest, one needs to have an exhaustive list of the members in the population, called the sampling frame, from which the sample will be drawn. Ideally the sampling frame should be the list of the entire population of interest—as may be the case for a population of college students—but this may not be the case depending on the size of the population of interest. For instance, the population may be too large to have a list of every member (e.g., the U.S. population). In such an instance, the researcher may rely on a list that

approximates a representative sample of the population of interest.

There are two main methods of sampling: probability and nonprobability sampling. Probability sampling occurs when the sampling frame for the population of interest is known, and each unit/member within the population has a known chance, or probability, of being selected. When a random sample of the population of interest is selected, the chance of missing entire sections of the population is minimized, and the size and effect of sampling error is calculable. The results from a study that uses probability sampling can be generalized to the entire population if the study has an adequate response rate.

Two main types of probability sampling are simple and stratified random sampling. In simple random sampling, each unit in the population, as determined by the sampling frame, has an equal chance of being selected (e.g., randomly selecting 500 college students from the list of all college students in a particular state). However, if there is variability across subsamples of the population with respect to the variable of interest, simple random sampling cannot ensure that all subsamples will be represented within the selected study sample. For instance, if the variable of interest is expected to vary for males and females in and there are five times more female than male college students in the state, simple random sampling might not provide enough males for study. When there is expected variability across subsamples of the population with respect to the variable of interest, stratified sampling is the preferred method of choice. In stratified random sampling, the sampling frame is divided into a number of nonoverlapping strata based on a factor that may affect the variable of interest (such as sex in the example above), and individuals are randomly selected from within each stratum (i.e., selecting a random sample of 250 males from among all male college students and a random sample of 250 females from among all female college students within the sampling frame). Stratified random sampling can ensure that representative samples of all relevant subsamples of the population are selected (in this example, it would ensure enough males were selected for the study of college students). Stratified random sampling provides greater statistical precision because there is less variability within a stratum compared to the variability when dealing with the entire population, as in simple random sampling.

Nonprobability sampling is used when the sampling frame is not available. With nonprobability sampling, information on entire sections of the population may be missing, which affects the ability to estimate the size and effect of the sampling error. Results from a study that uses nonprobability sampling cannot be generalized to the entire population of interest. In nonprobability sampling, subjects are selected on the basis of their accessibility or by the purposive personal judgment of the researcher. Examples of nonprobability sampling include convenience, consecutive, quota, and snowball sampling. Convenience sampling refers to the

collection of sample units that are easiest to access until the desired sample size is reached (e.g., a volunteer sample). Like convenience sampling, consecutive sampling is not based on random sampling. However, in consecutive sampling efforts are made to include all accessible sampling units typically within a specified time period. An example of consecutive sampling would be the recruitment of all patients who present to a mental health clinic from January 1 to June 30 of a specified year. Quota sampling involves the nonrandom selection of sample units based on some fixed quota that is related to one or more specific sample characteristics. Sampling is done until the quota is met, for example, adding new clinic admissions with a diagnosis of bipolar disorder to the study sample until a specified sample size is met. Snowball sampling is the method of choice when trying to sample hidden or hard-to-reach populations, such as undocumented immigrants or homeless individuals. Under such conditions, researchers may identify one or more individuals from their populations of interest and invite them to participate. These individuals are then asked to identify and recruit additional individuals until the desired sample size is attained. For example, undocumented immigrants may notify other undocumented persons to participate.

MEASUREMENT

Measurement instruments are developed in order to identify disorders and their associated characteristics in community and clinical populations. When considering a measure for use, either in a research or clinical setting, the target population on which the instrument was tested should match the population in which the instrument will be used. An instrument cannot be guaranteed to perform well in a population for which it was not developed or tested. This is a particularly important consideration when using an instrument in populations with different sociodemographic characteristics, cultures, or languages. Similar caution must be exercised when using an instrument that has been modified from its original construction. The deletion of items from an instrument, changing of item ordering, and addition of items may all have an effect on the performance of the instrument, no matter how small or useful the changes may seem.

Reliability

Reliability measures the ability of a measurement to produce the same results on repeated administrations. Test–retest reliability assesses the ability of a measure to arrive at the same result for the same subject on repeated administrations. The interval between the test and the retest should be long enough to ensure that the person's responses are based on his or her current condition, rather than on the memory of his or her responses in the first test administration. If the interval is too long,

however, there is an increased risk that the person's condition may have changed between test and retest. Another frequently used assessment of reliability is interrater reliability. This test describes a measurement instrument's ability to produce the same result when administered by two different raters.

Table 5.1–2.
2 × 2 Table for Reliability Assessments

		Rater 1 or Test		
		Present	**Absent**	**Total**
Rater 2 or Retest	Present	a	b	n3
	Absent	c	d	n4
	Total	n1	n2	N

The importance of reliability is obvious: if a measure cannot produce the same results on repeated administrations and across various raters, then it has limited value in assessing the presence or absence of a condition. Test–retest reliability is enhanced when the construct being assessed is stable across time, when it is specifically defined, and when the assessment instrument is standardized. In addition, reliability is more easily demonstrated when the individuals being assessed are clear cases of the disorder, not cases barely meeting the threshold for caseness. The reliability of an instrument, particularly a diagnostic instrument, is therefore generally lower when measured in a community population, where untreated cases around the diagnostic threshold abound, as opposed to a more selected population, such as a clinical population. Individuals chosen from clinical populations are more likely to fall well beyond the diagnostic threshold, are more accustomed to answering questions related to their clinical status, and are less likely to change answers on retest and less likely to fall below a diagnostic threshold.

Table 5.1–2 is used to illustrate the outcomes of interrater and test–retest reliability assessments, where the outcome is dichotomous—each individual must be classified as a case or a noncase. Cells a and d represent individuals for whom there is agreement between both raters (for an interrater reliability assessment) or ratings (for a test–retest assessment). Cells b and c represent individuals for whom there is disagreement between raters or ratings. The statistic most commonly used to describe the reliability of an instrument is Cohen kappa statistic, which calculates agreement adjusted for chance. In Table 5.1–2, observed agreement is represented by $(a + d)/N$, chance agreement is represented by $[(n1 * n3) + (n2 * n4)]/N^2$.

$$\text{Kappa} = \frac{\text{observed agreement} - \text{chance agreement}}{1 - \text{chance agreement}}$$

The kappa statistic can vary between 0 and 1, with 0 representing agreement no better than chance and 1 representing perfect agreement. Interpretation of kappa scores between 0 and 1 has not been standardized, but guidelines suggested by Landis and Koch in 1977 are frequently used. These guidelines divide the kappa range into five equal categories (0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, 0.8–1.0). The range of 0.4–0.6 is termed "fair" or "moderate" agreement, and a kappa coefficient of at least 0.4 is generally considered to represent adequate reliability for a diagnostic instrument.

Validity

Validity is the extent to which an assessment instrument measures what it is purported to measure. There are several types of validity.

Content Validity. This type of validity is concerned with how well a measure covers the construct that is being measured. For a diagnostic instrument assessing a psychiatric disorder such as major depression, content validity might at a minimum assess how well the DSM criteria are covered. For example, an instrument that does not inquire about depressed mood, or that inquires about spending sprees, would have suspect content validity as a measure of major depressive disorder. Content validity also depends on the population that the instrument is meant to cover, and the method by which data are gathered. For example, one would not ask men about menstrual symptoms when assessing somatic complaints, nor would one ask a respondent to estimate his or her current blood pressure as opposed to directly measuring it. Content validity cannot be empirically measured, but rather must be assessed by expert consensus after careful examination of the characteristics of the construct that is being measured and comparison with other instruments that may be available.

Criterion Validity. This type of validity examines the measurement instrument in relationship to a factor that is external to the measurement, known as the criterion. Criteria used to assess the validity of a measurement for a mental disorder are problematic in that there is usually no "gold standard" by which to assess the definitive presence or absence of disorder. Semi-structured clinical interviews are often used as a criterion, in the absence of a gold standard, although they are also prone to some degree of error. Two types of criterion validity are commonly assessed, concurrent validity and predictive validity.

Concurrent validity is assessed when the measure being tested is administered along with another measure known to be related to the construct of interest. An example of concurrent validity was demonstrated

in the MECA study, described below, in which a highly structured lay interviewer-administered diagnostic instrument, the DISC-2.3, was compared to the results of a clinical interview administered at the same session. Predictive validity is assessed when the measure of interest is compared to a criterion that is measured in the future. In other words, the measure is assessed in terms of how well it predicts an outcome. Predictive validity is particularly important in assessing a test that predicts later development of a disorder, or a person's response to treatment. Criterion validity is often expressed in terms of the sensitivity, specificity, and predictive value of a test illustrated by Table 5.1–3.

True positives, in which the test correctly identifies a case are represented in cell a. True negatives in which the test correctly identifies a noncase are represented in cell d. A perfect test would categorize all persons in one of these cells. In practice, however, no test is totally accurate; these measurement errors should be minimized. Errors of measurement include false positives, in which the test assigns caseness to a noncase (cell b) and false negatives in which the test does not recognize a case (cell c). The degree to which a test correctly assigns cases and noncases is described by the concepts of sensitivity and specificity.

Table 5.1–3.
2 × 2 Table to Determine Measures of Criterion Validity

		Criterion		
		Positive	**Negative**	**Total**
Test	Positive	a	b	n3
	Negative	c	d	n4
	Total	n1	n2	N

Sensitivity is described as the proportion of cases that are correctly assigned as cases by the test.

$$\text{Sensitivity} = a/n1$$

Specificity can be described as the proportion of noncases that are correctly assigned as noncases by the test.

$$\text{Specificity} = d/n2$$

Sensitivity and specificity vary between 0 and 1. Maximizing the sensitivity and specificity of an instrument has important ramifications. A high sensitivity assures that a minimal number of persons with a disorder will be missed by the test, an obvious necessity in clinical or screening situations. High specificity minimizes the chance that an individual will receive unnecessary treatment or confirmatory tests.

Predictive value is another measure that can be derived from data in Table 5.1–3. Positive predictive value is the proportion of persons who test positive for the disorder who actually have the disorder. Negative predictive value is the proportion of persons test negative for the disorder who actually have the disorder.

$$\text{Positive predictive value} = a/n_3$$

$$\text{Negative predictive value} = d/n_4$$

Construct Validity. Construct validity is the most complex type of validity assessment, indicating how well the measure of interest relates to a theoretical set of concepts that, taken together, describe the condition being measured. Implicit in a determination of construct validity is a theoretical model of the condition being measured, with an explanation of how the components of the model are related. The most well-known theoretical model of the qualities of a psychiatric disorder was described in 1972 by Robins and Guze. This model posits the following components of a valid psychiatric diagnosis:

Clinical description—A clear description of the clinical syndrome is needed, which may include not only symptoms but also associated factors such as age of onset, precipitating factors, and sex ratios.

Laboratory studies—Consistent and reliable findings on tests do much to increase the validity of a diagnosis. Such tests may include psychological tests as well as biochemical, imaging, anatomical, and physiological tests.

Delimitation from other disorders—Robins and Guze argued that different disorders can have similar clinical descriptions and test results, so exclusion criteria should be specified to differentiate disorders, as well as to exclude borderline and doubtful cases.

Follow-up study—Once diagnosed, the natural history of a disorder (course and outcomes) should be consistent, and certainly the diagnosis should not change over time.

Family study—An increased prevalence of the same disorder among family members lends credibility to the validity of diagnosis.

It is clear from this list of validators that establishing construct validity is a process that involves more than a single study, unlike other tests of validity. Ideally, multiple studies should be used to replicate a validator using different measurement tools. The slow progress in establishing construct validity for psychiatric disorders indicates that further research is needed; as noted above, this may include a rethinking of the categorical model of mental disorders to include dimensional symptom assessments, as well as continued work on the neurobiology of basic mental and

behavioral processes.

CASE IDENTIFICATION

Background

As in all fields of epidemiology, case identification is a key issue in psychiatric epidemiology. In contrast to illnesses where pathological processes or lesions can be identified through imaging, cellular pathology, blood tests, and so on, making a psychiatric diagnosis relies on the self-reported history of the individual being assessed, reports of other informants, and ancillary sources of information such as past medical records. It should be noted that this state of affairs is not limited to psychiatric disorders; some neurological disorders such as migraine headaches and multiple sclerosis, for example, lack definitive diagnostic tests, and diagnosis relies heavily on the accumulation of supportive clinically reported data.

The Diagnostic and Statistical Manual of Mental Disorders

The development of the diagnostic system for mental disorders has obvious implications for the development of psychiatric epidemiology. Although various diagnostic classifications have been in use since ancient times, the development of psychiatric epidemiology has been inextricably linked to the development of the DSM. The first edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-I) was published in 1952 by the American Psychiatric Association. Both DSM-I and its successor, DSM-II, published in 1968, contained glossaries that described the diagnostic categories in the manuals. Clinicians were expected to make diagnoses based on these descriptions and their clinical judgment. Given the latitude for variations in interpretation of glossary entries, along with differences in clinical training and judgment, diagnostic reliability was problematic for DSM-I and DSM-II, and clinical uptake of the diagnostic nomenclature was inconsistent. The problems with the reliability of psychiatric diagnoses were dramatically illustrated in a landmark cross-national study of diagnostic patterns, now known as the "U.S.–U.K. Study." This study aimed to elucidate the disparities in the rates of schizophrenia and affective disorders in the United States and the United Kingdom by examining diagnostic methods. At the time, schizophrenia was diagnosed more frequently in the United States and affective disorders were diagnosed more frequently in the United Kingdom. The study trained a group of psychiatrists from the two countries in the administration of the present state examination (PSE), a semistructured diagnostic instrument. When these psychiatrists made diagnoses with the PSE, rather than relying totally on their clinical judgments, their diagnostic patterns were similar; no country-specific differences in assignment of affective disorder or

schizophrenia were found. The transition from DSM-II to DSM-III in 1980 allowed for more precise case definition in psychiatry and spurred a new generation of studies in psychiatric epidemiology. The DSM-III aimed to provide clinicians with a common language to describe psychiatric disorders, so that a diagnosis had the same meaning regardless of the diagnostician. It also aimed to be theoretical, and dispensed with many concepts in previous editions that were tied to then-prevalent and untestable psychoanalytic theories. These aims were accomplished by replacing the glossary of DSM-II with specific diagnostic criteria in DSM-III, based on symptoms that were observable or could be elicited directly from the patient. The DSM-III was accepted relatively readily and quickly became the diagnostic standard for mental health clinicians and researchers, as well as others relying on these diagnoses, such as federal, state, and local governments, regulatory agencies, the legal system, and the insurance industry. A revision of the DSM-III-R was released in 1987, to correct inconsistencies and clarify criteria that were shown to be problematic in DSM-III. DSM-IV was released in 1994, and DSM-5 was released in 2013. Both DSM-IV and DSM-5 represented major updates of the DSM, incorporating new findings from research that had accumulated since the release of the prior editions. As discussed below, changes in the DSM diagnostic criteria have had implications for the epidemiological measurement of mental disorders and for comparisons between studies performed with different diagnostic criteria.

The International Classification of Diseases

The *International Classification of Diseases* (ICD) is a diagnostic nomenclature maintained by the World Health Organization (WHO). It is in use in the United States and around the world as the common language for reporting basic morbidity and mortality statistics. Because psychiatric disorders are reported as morbidity statistics, the ICD and the DSM have been associated to varying degrees for many years. The DSM-I was a variant of the sixth edition of the ICD (ICD-6), and DSM-II was developed to be released in tandem with ICD-8. ICD-9 was published in 1975, only 1 year after work began on DSM-III, and implemented in 1978. The ICD-9 terminology was therefore more reflective of DSM-II, and its diagnostic categories did not match those in the new DSM-III. Because of dissatisfaction across all of medicine with the lack of specificity in ICD-9, it was modified for use in the United States, resulting in ICD-9-CM (*Clinical Modification*). The development of the DSM-IV and the ICD-10 was more closely coordinated than any previous efforts, and the resulting systems were highly congruent, although not identical. The WHO also, for the first time, developed "Diagnostic Criteria for Research" to accompany the ICD-10 mental disorder diagnoses. As with the disorder list and nomenclature, these diagnostic criteria were usually, but not always, congruent with DSM-

IV diagnostic criteria. Several diagnostic instruments have incorporated ICD-10 criteria along with DSM criteria, allowing examination of their differences and refinement of both sets of criteria. ICD-10 was published in 1992, but in contrast to most other countries was not quickly adopted as the official code set for the United States. The clinical modification of the ICD-10 (ICD-10-CM) is scheduled to be implemented in the United States on October 1, 2015. The development of DSM-5 was intended to be closely coordinated with the development of the ICD-11. However, because the release date for the ICD-11 was pushed back several years, to 2017, the opportunities for coordination with DSM-5 were limited by these different timetables for completion. Drafts of the ICD-11 section on mental and behavioral disorders display considerable similarities with DSM-5 in chapter structure and disorder names.

Disability. Over the past 30 years, the consequences of mental disorders have assumed an increasingly important role in the theoretical underpinnings of epidemiological studies and the interpretation of their data. The concepts of clinically significant distress or impairment in functioning have been included in the DSM definition of "mental disorder" since the publication of DSM-III. To further emphasize the importance of these concepts in establishing a threshold for diagnosis, the DSM-IV added (for most disorders) a criterion requiring clinically significant distress or impairment in an important domain of functioning before a disorder could be diagnosed. This decision was criticized on conceptual grounds for its blending of two separate constructs, symptomatology and the consequences of symptomatology, as opposed to a conceptualization of a mental disorder as a combination of symptom aggregation, severity, and duration without regard to any resulting disability. The decision was reexamined during the development of DSM-5, and the revised definition of "mental disorder" in the manual states that mental disorders are "usually associated with significant distress or disability in social, occupational, or other important activities." This change allowed the developers of DSM-5 to remove the "clinical significance criterion" of DSM-IV if they so chose. Although the multiaxial assessment was removed from DSM-5, the manual continued to emphasize the importance of the clinician's assessment of the patient's social and occupational functioning and disability for treatment planning and prediction of outcome.

The WHO has taken a leadership role in developing the theory, definitions, and classification of concepts related to health and functioning. The *International Classification of Impairments, Disabilities, and Handicaps: A Manual of Classification Relating to the Consequences of Disease* (ICIDH) was released by the WHO in 1980, and subsequently revised as the *International Classification of Impairments, Activities, and Participation: A Manual of Dimensions of Disablement and Functioning*

(ICIDH-2). The next and most current version of this classification was released in 2001 and titled the *International Classification of Functioning, Disability, and Health* (ICF). This document is part of the WHO's "family of international classifications" that includes the ICD-10 and allows coding of complementary information on health conditions and their consequences to provide a more detailed picture of the health status of an individual or a population. The ICIDH was mainly a classification of "consequences of disease" classification, while the ICF made a determined shift to classify "components of health." Part I of the ICF covers "Functioning and Disability," while part II covers "Contextual Factors." Table 5.1–4 summarizes the further organization of the ICF. The ICF uses an alphanumeric coding system for three of its components: "Body Functions and Structures," "Activities and Participation," and "Environmental Factors." Personal factors (e.g., race, gender, habits, coping styles, etc.) are not associated with codes. The categories of the ICF are nested, so that broad categories have a number of subcategories. For example, Chapter 1 of the "Body Functions" section describes mental functions, which are subcategorized as global mental functions and specific mental functions. Global mental functions contain eight subcategories, such as temperament and personality, sleep, orientation, and consciousness. There are 14 subcategories of "Specific Mental Functions," including attention, memory, emotional, perceptual, and mental functions of language. Although the ICF is comprehensive in its description of health domains and provides a common language for describing these domains, its main users continue to be in the physical medicine and rehabilitation fields, where interventions focus more on improving functioning and removing barriers, rather than diagnosing and treating illnesses. There is no doubt, however, that the use of a standard classification of disabilities would be helpful for psychiatric research and clinical practice, by drawing attention to the conceptual distinctions between disorders and their consequences.

Table 5.1–4.
Overview of the ICF

	Part 1: Functioning and Disability		**Part 2: Contextual Factors**	
Components	Body functions and structures	Activities and participation	Environmental factors	Personal factors
Domains	Body functions Body structures	Life areas (tasks, actions)	External influences on functioning and disability	Internal influences on functioning and disability
Constructs	Changes in body	Capacity to execute tasks in	Facilitating or	Impact of

	functions (physiological) Changes in body structures (anatomical)	a standard environment and Performance executing tasks in the current environment	hindering impact of features of the physical, social, and attitudinal world	attributes of the person
Positive Aspect	Functional and structural integrity	Activities and Participation Functioning	Facilitators	Not applicable
Negative Aspect	Impairment	Activity limitation and participation restriction Disability	Barriers, hindrances	Not applicable

ASSESSMENT INSTRUMENTS USED IN EPIDEMIOLOGICAL STUDIES OF MENTAL DISORDERS

Diagnostic Instruments for Adult Populations

The Diagnostic Interview Schedule. The Diagnostic Interview Schedule (DIS) was developed in 1978 in response to two major impending events: the release of DSM-III, with its criterion-based diagnostic system, and the development of the National Institute of Mental Health (NIMH) Epidemiologic Catchment Area (ECA) program, which was being planned to capitalize on DSM-III's innovative new diagnostic system. Lee Robins, Ph.D., the lead developer of the DIS, wrote that there was no existing diagnostic instrument that could be used for the ECA program, not only because the DSM-III criteria sets were totally novel, but also because of limitations with the existing instruments in meeting the needs for a large community survey. Particularly important was the need for the instrument to be administered by interviewers who did not have clinical training, a necessity given the large numbers of respondents estimated for the ECA program. Other requirements for the diagnostic instrument included the ability to cover of a broad range of major DSM diagnoses in a single administration, to assess specific DSM-III criteria from respondent reports only, without use of medical records or outside informants, and also to be acceptable and comprehensible to a general adult population. The Renard Diagnostic Interview (RDI), developed at Washington University in St. Louis and based on Feighner diagnostic criteria, was felt to represent a suitable model for the ECA program's diagnostic instrument. The RDI was developed in a clinical setting, from observations of psychiatrists' clinical interviews based on Feighner criteria. The instrument was "fully structured," meaning that its questions, response options, and administration patterns were all specified and therefore did not require clinical judgment for administration or data recording. Its reliability and concurrent validity were good in clinical populations, with lay and clinical interviewers. The task of developing a diagnostic instrument for DSM-III was assigned to the RDI's authors. The DIS covered approximately 28

DSM-III diagnoses, as well as diagnoses based on the earlier Feighner criteria and Research Diagnostic Criteria (RDC). The DIS generally assessed symptomatology first on a lifetime basis (i.e., "Have you ever had [symptom…]?") and then assessed clustering of the positive lifetime symptoms into an episode. If symptoms did cluster together at the same time, assessment of recency was made, so that period prevalence estimates could be made (e.g., 1-year, 6-month, 1-month prevalence rates). Like the RDI, the DIS attempted to distinguish "true" psychiatric symptoms or syndromes from normal reactions to life events, and from conditions stemming directly from general medical conditions or the use of substances. Questions assessing the former situation were referred to as *clinical significance* questions; questions assessing the latter situation were referred to as the *probe flow* questions, named after the "probe flow chart" from which interviewers read these questions. The highly structured nature of the DIS allowed diagnoses to be made by computer via a diagnostic algorithm. This diagnostic algorithm allowed a diagnosis to be made free from any diagnostic custom or clinical interpretation of the interviewer, and along with the operationalized DSM-III criteria and the highly structured interview format, allowed a reliable diagnosis to be made. Testing showed that the DIS achieved the DSM-III's goal of diagnostic reliability, and that concurrent validity with a clinical research interview was reasonably strong. The DIS was updated to account for changes from DSM-IIIR to DSM-IV and to take advantage of new survey techniques and lessons learned from surveys using the DIS and similar diagnostic instruments. For instance, assessment of disability in the context of mental disorders was greatly expanded, diagnostic coverage was expanded to include disorders of childhood and an updated assessment of cognitive impairment, and there is less frequent use of skip-outs (where respondents who deny a cardinal symptom of a disorder are not asked any other symptoms of the disorder). There is currently no DSM-5 version of the DIS.

The Composite International Diagnostic Interview. The Composite International Diagnostic Interview (CIDI) was first developed in the early 1980s, at the request of the U.S. Alcohol, Drug Abuse, and Mental Health Administration and the WHO, which had formed a Joint Project on Diagnosis and Classification. Originally designed to incorporate questions from both the DIS (DSM-III) and the Present State Examination, for expanded use in epidemiological studies worldwide, it was subsequently modified to cover other diagnostic classifications including DSM-III-R, DSM-IV, and ICD-10. The CIDI was based largely on the DIS and shared many of its questions and characteristics, including highly structured administration and scoring, assessment of psychopathology beginning at a lifetime level, and use of clinical significance and probe questions. Taking advantage of the increasingly rapid uptake of computers throughout the

world, a computerized version of the CIDI was developed. An expanded module for detailed examination of substance use disorders was also developed, called the CIDI-SAM. One of the intended first uses of the CIDI was the National Comorbidity Survey (NCS).

The Composite International Diagnostic Interview, University of Michigan Version. The paper-and-pencil administered, DSM-III-R/ICD-10 version of the CIDI was modified by the investigators in the NCS, which was conducted out of the University of Michigan. The resulting instrument was called the UM-CIDI. The UM-CIDI was thought to bring major improvements to the CIDI, while still maintaining the basic CIDI structure. Nonetheless, the developers of the modified instrument acknowledged that the decision to modify the instrument was not agreed on by all of the original CIDI developers. Among the modifications made in the UM-CIDI were changes in the layout of the instrument form; addition of an introductory question asking the respondent if he or she is willing to think carefully in order to provide the interviewer with accurate information; deletion of diagnoses with low reliability or of secondary importance to the goals of the NCS; using the CIDI psychosis module as a screener for follow-up clinical interviews, rather than as a self-contained diagnostic module; moving the diagnostic stem questions to their own section at the beginning of the interview; reordering questions; adding and rewriting questions for clarity; adding notes to help interviewers and reduce the amount of memorization required of them; changing the methods of administering the probe flow questions; and adding a respondent booklet with visual cues. Although many of these modifications in the UM-CIDI did address problems detected in previous usages of the DIS and the CIDI, it was not clear whether they all represented improvements when implemented in a major household survey such as the NCS. The differences between the UM-CIDI and the DIS were therefore among the primary reasons cited for the disparities in prevalence rates between the ECA and the NCS. Attempts to reconcile these prevalence rates stimulated a critical examination of the instruments and survey methods in these two landmark surveys, as discussed below.

The Composite International Diagnostic Interview, World Mental Health Surveys Version. The CIDI underwent further modification in anticipation of the WHO WMH Survey Initiative. It had been recognized that although the CIDI provided a common diagnostic assessment, it did not provide guidance on other domains critical to large-scale epidemiological surveys. Thus, it was difficult to compare risk factors for disorders, consequences of disorders, patterns of disorders, and correlates of service use across surveys. The WMH CIDI was created to address these inconsistencies in assessments. It consists of seven sections: Screening and lifetime review,

disorders, functioning and physical disorders, treatment, risk factors, sociodemographics, and a methodological section. The interview is designed to be conducted in two parts, so that interviews can be ended for a subset of subjects who do not meet criteria for lifetime psychopathology. A number of changes were made to the CIDI's diagnostic questions to improve the quality of responses, and several new diagnoses were added, such as intermittent explosive disorder, separation anxiety disorder, and attention-deficit/hyperactivity disorder. A fuller assessment of disability was instituted, including expanded diagnostic-level clinical significance questions, a modified WHO Disability Assessment Schedule (WHODAS), and detailed questions on work performance. New sections on pharmacoepidemiology, nonspecific psychological distress, family burden, and childhood experiences were also included. Research on the validity of this instrument across the diverse range of nations in the WMH Survey Initiative is ongoing. There is currently no available update of the CIDI or WMH CIDI for DSM-5 disorders.

The Alcohol Use Disorder and Associated Disabilities Interview Schedule. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS) was originally developed at the National Institute on Alcoholism and Alcohol Abuse (NIAAA) in 1989, with a focus on substance use disorders and the mental disorders commonly comorbid with them. It was subsequently updated for DSM-IV and DSM-5 diagnostic criteria. The AUDADIS-5, based on DSM-5 diagnostic criteria, was completed before the release of the DSM-5, although the developers of the instrument note that the criteria for the disorders covered in the instrument were substantially finalized by 2011. It provides an assessment of criteria for substance use disorders including alcohol, tobacco, and other drugs; depressive and bipolar disorders; posttraumatic stress disorder; and personality disorders. The AUDADIS has been used to make DSM diagnoses in the NIAAA National Longitudinal Alcohol Epidemiologic Survey (NLAES) and the three versions of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). As with other instruments used in large epidemiological surveys, it is fully structured, computerized, and can be administered by lay interviewers.

Diagnostic Interviews for Child and Adolescent Epidemiology

The development of diagnostic interviews for child and adolescent psychiatric epidemiology has paralleled that of instrument development for adult populations. However, the complexities involved in obtaining diagnoses in these populations and funding priorities placed on adult research have resulted in somewhat slower progress. In 1969 a semistructured instrument patterned after the Renard Diagnostic Interview was developed, the Diagnostic Interview for Children and Adolescents

(DICA). The development of the DIS for adults in the late 1970s spurred much interest in developing a fully structured lay interview schedule for assessing DSM-III diagnoses in children and adolescents. Under the sponsorship of the NIMH, the first version of the Diagnostic Interview Schedule for Children (DISC), covering DSM-III diagnoses, was developed in the early 1980s and work on the DISC-R, covering DSM-III-R diagnoses, was started in 1985. Refinements of the DISC-R led to the DISC-2.1 and -2.3, the latter version being extensively tested for reliability and validity in the MECA study. Findings from the MECA led to the current version of the DISC, the DISC-IV, which covers DSM-IV (and also DSM-III-R and ICD-10) disorders. The DISC-IV assesses over 30 disorders, organized into 19 diagnostic sections. The DISC-IV is a fully structured diagnostic interview that is suitable for administration by lay interviewers. A computer-assisted version (C-DISC) has been developed to simplify interviewer administration and reduce errors and costs. The DISC-IV has matching parent and child/youth versions (DISC-P and DISC-C); a partial teacher version (DISC-T) is also available. The DISC-P can be administered to parents of children ages 6 to 17 years; the DISC-C can be administered to children ages 9 to 17 years; the DISC-T is suitable for teachers of elementary school age children. Symptoms are assessed on a past-year and past-month basis. An optional "whole-life" module was developed for administration at the end of the interview. Questions on age and context of onset, receipt of treatment, and distress and disability are asked at the end of each diagnostic section if a respondent meets a predetermined number of symptoms for that diagnosis. Diagnoses are made by computer algorithm. There is currently no DSM-5 version of the DISC. Although the DISC is the most widely used epidemiological diagnostic instrument in child and adolescent populations, other instruments are in use. The Child and Adolescent Psychiatric Assessment (CAPA), developed at Duke University, was used in the Great Smoky Mountains Study. The CAPA provides DSM-IV and ICD-10 diagnoses with a high level of detail and can be administered by lay interviewers. It does require intensive training to provide the interviewers with a sufficient level of judgment to make decisions on the presence or absence of symptoms (referred to as an "interviewer-based" approach). However, with training and ongoing monitoring of interviewers, the CAPA works well in community surveys with children and adolescents ages 9 to 17 and their parents. The developers of the CAPA have also developed the Young Adult Psychiatric Assessment (YAPA) for respondents ages 18 and over, and the Preschool Age Psychiatric Assessment (PAPA) for administration to parents of 2- to 5-year-old children. The DICA was updated to provide DSM-IV diagnoses; it is a semistructured instrument that allows interviewers to probe responses. There are currently no available DSM-5 versions for the CAPA or the DICA.

Instruments to Assess Disability and Impairment

There are very few instruments that have been developed to assess disability in community-based epidemiological surveys. Most instruments require clinical training to administer and have been developed for patient populations. The development of instruments for child and adolescent populations has seen more activity than the development of instruments for adults, in part because the diagnostic threshold issues for younger populations were recognized as a particular problem.

WHO Disability Assessment Schedule. The WHODAS 2.0 was developed to be compatible with the disability model put forth in the ICF, particularly in the activity and participation dimensions. It is fully structured and can be administered to adult subjects by lay interviewers or self-administered. The core instrument is self-report, but a version has been developed that can be administered to a proxy respondent. The reporting period is the past 30 days. The full version of the WHODAS 2.0 is 36 items long and assesses six domains: cognition (understanding and communicating), mobility (moving and getting around), self-care (hygiene, dressing, eating, and staying alone), getting along (interacting with other people), life activities (domestic responsibilities, leisure, work, and school), and participation (joining in community activities). There is a short version that is 12 items long. An earlier version of the WHODAS 2.0, the WHODAS II, was substantially modified for use in the WHO World Mental Health Survey Initiative. These modifications appear to have reduced the sensitivity of the instrument to detect milder levels of disability, complicating interpretation of the results. The full version of the WHODAS 2.0 was tested in the DSM-5 field trials, and was found to have good to excellent reliability for the domains and total scores in randomly sampled adults at seven clinical sites.

Disability Assessment in Child and Adolescent Populations. The Children's Global Assessment Scale (CGAS) is based on the adult Global Assessment Scale (GAS). Although developed for clinical use, an adaptation of this scale with nonclinical language, for use by lay interviewers, was developed for use in community surveys and was administered in the MECA survey (see below). Like the GAS, the rater assigns a single number that best describes the child's functioning, from 1 (most impaired functioning) to 100 (healthiest). A score of 70 or less indicates at least mild functional impairment. The Brief Impairment Scale (BIS) grew out of work on the Columbia Impairment Scale (CIS) which was used in the MECA survey. It is a fully structured, 23-item scale that assesses the child's functioning without tying it to any specific health conditions that the child may have. The BIS was designed to allow administration by a lay interviewer in community surveys, but it is also suitable for use in clinical settings. It is administered to an adult informant. Three domains are

assessed: interpersonal relations, functioning at school or work, and self-fulfillment. Initial testing of the scale has shown that it has acceptable test–retest reliability, internal consistency, and convergent and concurrent reliability. The three domains show low intercorrelations, suggesting that they are related, but separate components of the overall global functioning measured by the scale. Several other assessment scales are available, but have limitations including unidimensionality and mixing of symptom and disability constructs, or have practical difficulties for use in community surveys, such as length of administration or a need for clinical judgment in scoring.

MAJOR EPIDEMIOLOGICAL SURVEYS

Surveys of Adult Populations

The Epidemiologic Catchment Area Program. Spurred by the release of the President's Commission on Mental Health, chaired by First Lady Rosalynn Carter, the ECA program was developed by the NIMH in conjunction with the simultaneous development of the DSM-III and the DIS. The main aims of the ECA were to estimate the prevalence of DSM-III mental disorders in community and institutional populations and determine the extent of the use of treatment services for these disorders. The ECA was carried out in five catchment areas across the United States: New Haven, Connecticut; Baltimore, Maryland; Durham, North Carolina; St. Louis, Missouri; and Los Angeles, California. In order to analyze specific subgroups of interest, oversampling of the elderly, African Americans, and Hispanics was done at selected sites. There were two samples in each site, a household sample and an institutional sample, which included jails and nursing homes; respondents were all age 18 or older. The final sample size was 18,571 household respondents, and 2,290 institutional respondents. The ECA was designed as a longitudinal survey with three waves conducted 6 months apart with the same respondents. The first and third waves collected diagnostic and service use data, while the second wave collected service use data, to reduce recall bias in service use reports. The ECA program's sampling strategy was not nationally representative, but the investigators combined the data from the five sites and adjusted them to the U.S. population to create national estimates. Prevalence and service use rates from the ECA program are shown in Table 5.1–5, columns 1 and 3. Over time, the analysis of data generated by the ECA program served to point out weaknesses in the DSM-III and its subsequent revisions, as well as providing valuable data for service planning and analytic studies. The high prevalence of diagnostic comorbidity among DSM-III disorders, particularly in the absence of DSM-III diagnostic hierarchies, was one of the first major issues uncovered by the analyses of ECA data. Surprisingly, a high prevalence of disorders previously thought to be rare, such as social

phobia and obsessive-compulsive disorder, were found in the community; low rates of treatment for these and most other psychiatric disorders contributed to this mistaken assumption. The onset of many disorders was found to originate in childhood, contrary to much contemporary thinking that children were relatively protected from psychiatric disorders. Data from the ECA survey were instrumental in educating the general public as well as policy makers and legislators that mental disorders are real, prevalent in U.S. society, disabling, and grossly undertreated. Such basic efforts provided support, through an increasingly detailed series of reports released by NIMH, for parity of mental health insurance benefits with the benefits for other medical or surgical conditions and treatments. A similar purpose was served for the landmark *Surgeon General's Report on Mental Health,* published in 1999. More recently, the DIS clinical significance criteria data provided a crucial piece of information to evaluate the role of clinically significant distress and functional impairment in diagnosis and estimation of prevalence rates, as detailed below. Longitudinal follow-ups of the Baltimore ECA sample continued, 16 and 22 years after the sample was originally drawn. Among the findings reported from these follow-ups are reports that major depressive episode is an independent risk factor for the development of hypertension, and that cluster B personality disorders and traits are independent risk factors for the development of cardiovascular disease as well as cardiovascular mortality. Finally, the ECA program served as a stimulus for similar surveys in other locations, such as Puerto Rico and the Netherlands, and set the stage for subsequent U.S. and international surveys, such as the NCS and the WMH Survey Initiative.

Table 5.1–5.
Effect of Clinical Significance Criteria on 1-Year Prevalence Rates of Mental Disorders in Adults, Age 18–54, ECA and NCS Surveys

	Prevalence, % (95% Confidence Interval)			
	Without Clinical Significance Criteria		With Clinical Significance Criteria	
Disorder	**ECA**	**NCS**	**ECA**	**NCS**
Simple phobia	#	8.6 (7.6, 9.6)	8.5 (7.9,9.1)	4.4 (3.6, 5.2)
Social phobia	#	7.4 (6.6, 8.2)	2.0 (1.6,2.4)	3.7 (3.1, 4.3)
Agoraphobia	#	3.7 (2.9, 4.5)	5.0 (4.4,5.6)	2.2 (1.6, 2.8)
Any phobia	#	14.7 (13.3, 16.1)	11.4 (10.6,12.2)	8.0 (7.2, 8.8)
Panic	#	2.2 (1.6, 2.8)	1.6 (1.2,2.0)	1.7 (1.1, 2.3)
Generalized anxiety	NA	3.4 (2.8, 4.0)	NA	2.8 (2.2, 3.4)
PTSD	NA	3.6 (2.8, 4.4)	NA	3.6 (2.8, 4.4)
OCD	#	NA	2.3 (1.9,2.7)	NA
Any anxiety	#	18.7 (17.1, 20.3)	13.3 (12.5,14.1)	12.1 (10.7,13.5)
Major depressive disorder	5.4 (4.8,6.0)	8.9 (7.7, 10.1)	4.6 (4.0,5.2)	5.4 (4.4, 6.4)

Bipolar I	1.2 (1.0,1.4)	1.3 (0.9, 1.7)	0.6 (0.4,0.8)	1.3 (0.9, 1.7)
Bipolar II	0.6 (0.4,0.8)	0.2 (0.0, 0.4)	0.3 (0.1,0.5)	0.2 (0.0, 0.4)
Dysthymia	#	2.5 (2.1, 2.9)	1.7 (1.3,2.1)	1.8 (1.4, 2.2)
Any mood	10.4 (9.6,11.2)	11.1 (9.7, 12.5)	5.7 (5.1,6.3)	7.5 (6.3, 8.7)
Schizophrenia/NAP	1.3 (1.1,1.5)	0.2 (0.0, 0.4)	1.3 (1.1,1.5)	0.2 (0.0, 0.4)
Antisocial P.D.	2.0 (1.6,2.4)	NA	2.0 (1.6,2.4)	NA
Somatization	0.3 (0.1,0.5)	NA	0.3 (0.1,0.5)	NA
Anorexia nervosa	0.1 (0.1,0.1)	NA	0.1 (0.1,0.1)	NA
Severe cognitive impairment	1.3 (1.1,1.5)	NA	0.2 (0.2,0.2)	NA
Any nonsubstance use disorder	22.0 (21.0,23.0)	23.4 (21.6, 25.2)	18.2 (17.2,19.2)	15.4 (13.6, 17.2)
Any alcohol disorder	9.1 (8.3,9.9)	9.9 (8.9, 10.9)	8.9 (8.3,9.5)	6.5 (5.7, 7.3)
Any drug disorder	4.0 (3.6,4.4)	3.6 (3.0, 4.2)	1.5 (1.1,1.9)	2.4 (1.8, 3.0)
Any substance use disorder	11.7 (10.9,12.5)	11.5 (10.7, 12.3)	9.7 (8.9,10.5)	7.6 (6.6, 8.6)
Any disorder including substance use	29.6 (28.4,30.8)	30.2 (28.2, 32.2)	24.7 (23.7,25.7)	20.6 (18.8, 22.4)

#Clinical significance for the anxiety disorders and dysthymia was assessed at the symptom level in the ECA, and rates for syndromes without clinical significance could not be calculated.

NA refers to not assessed.

The NCS and Associated Studies. The NCS was an NIMH and National Institute on Drug Abuse (NIDA)-sponsored survey designed to study the comorbidity of DSM-III-R substance use disorders and other mental disorders. Unlike the ECA, it was a nationally representative household survey, with 8,098 respondents. It was fielded from 1990 to 1992. Because of the relatively low prevalence of substance use disorders in older populations, the age of subjects sampled for the NCS was capped at 54 years to maximize its analytic goals; to minimize recall bias of early-onset disorders, the lower age range of the survey was 15 years. In addition to measuring a wide range of psychiatric disorders with the UM-CIDI and mental health service use, a plethora of other domains were included in the NCS survey, such as human immunodeficiency virus (HIV) risk behaviors, personality constructs, and social life. The initially published prevalence rates of mental disorders from the NCS were generally higher than those previously published from the ECA program, leading to intense scrutiny of both surveys and uncertainty as to how to use the NCS results for health service policy and planning. In particular, it was not evident that the prevalence estimates of the NCS reflected a true, substantial increase in the rates of mental disorders in the United States in the decade since the initiation of the ECA program. Nor was it evident that the ECA substantially underestimated the prevalence rates of mental disorders; indeed, the rates of disorders reported by the ECA program were themselves felt to be nonconservative. Methodological and diagnostic differences in the two surveys were suspected, at least in part, to underlie

the differences in prevalence rates. First, there were some differences between the diagnostic criteria in DSM-III and DSM-III-R, which were reflected in the respective diagnostic instruments, and could affect prevalence rates. Second, differences in the age ranges of the respondents in the two surveys could affect prevalence rates; although few previous comparison data were available for the prevalence of mental disorders in adolescents aged 15 to 17 years sampled in the NCS, it had been documented that in epidemiological surveys, prevalence rates of mental disorders declined with age and were particularly low in the elderly. Third, the effects of several innovations in the UM-CIDI, such as grouping of stem questions and the commitment question, were unknown but potentially could increase rates of positive responses in the NCS. Finally, it was recognized that the clinical significance questions in the NCS were not used to determine prevalence rates and were used inconsistently in the ECA prevalence estimates. In a study that aimed to reconcile the age range differences and apply the unused clinical significance questions, the overall prevalence estimates of the ECA were reduced by about 17 percent and estimates of the NCS were reduced by about one third. As a result, the estimates of the two surveys were brought closer together (Table 5.1–5).

Several surveys related to the NCS have subsequently been carried out. The NSC-R surveyed another national sample from 2001 to 2002. The sample size was 9,282. This survey allowed comparisons of rates of mental disorders and service use in the decade since the original NCS survey in a repeated cross-sectional design. Many aspects of the NCS survey instrument were kept intact for the NCS-R to promote these comparisons. However, based on knowledge gained in the intervening years, new components were also added to the NCS-R, as described in the section on the WMH-CIDI above. One-year prevalence rates from the NCS-R are found in Table 5.1–6 and rates of service use are found in Table 5.1–7. Although the NCS-R permitted another cross-sectional estimate for the U.S. population, the NCS-2, by contrast, attempted to re-interview all respondents in the original NCS survey. The NCS-2 was conducted from 2001 to 2003. The longitudinal aspect of the survey has allowed investigators to examine risk factors for the new onset of disorders and to study the course of illness in persons who did have disorders in the original NCS survey. Several articles have been published using the NCS/NCS-2 data sets. These include studies of risk factors for suicide-related outcomes, mental disorders as risk factors for substance use disorders, and the role of baseline mental disorder severity in later clinical outcomes.

Table 5.1–6.
1-Year Prevalence Rates of DSM-IV Mental Disorders in Adults, NCS-R and NESARC Surveys

Disorder	Prevalence, % (SE)	
	NCS-R	NESARC #
Specific phobia	8.7 (0.4)	7.1 (0.3)
Social phobia	6.8 (0.3)	2.8 (0.1)
Agoraphobia w/o panic	0.8 (0.1)	
Panic	2.7 (0.2)	
Panic w/ agoraphobia		0.6 (0.1)
Panic w/o agoraphobia		1.6 (0.1)
Generalized anxiety	3.1 (0.2)	2.1 (0.1)
PTSD	3.5 (0.3)	
OCD	1.0 (0.3)	
Any anxiety	18.1 (0.7)	10.4 (0.3)
Major depressive disorder	6.7 (0.3)	7.1 (0.2)
Bipolar I or II	2.6 (0.2)	
Mania		1.7 (0.1)
Hypomania		1.2 (0.1)
Dysthymia	1.5 (0.1)	1.8 (0.1)
Any mood	9.5 (0.4)	9.2 (0.2)
Any impulse control disorder	8.9 (0.5)	
Alcohol abuse	3.1 (0.3)	4.7 (0.2)
Alcohol dependence	1.3 (0.2)	3.8 (0.1)
Any alcohol disorder		8.5 (0.2)
Drug abuse	1.4 (0.1)	1.4 (0.1)
Drug dependence	0.4 (0.1)	0.6 (0.1)
Any drug use disorder		2.0 (0.1)
Any substance use disorder	3.8 (0.3)	9.4 (0.3)
Any disorder	26.2 (0.8)	

#Rates for NESARC mood and anxiety disorders exclude substance-induced disorders.

In the 1990s increasing attention began to be focused on disparities in health status between the white population and racial and ethnic minority populations in the United States. These disparities appeared across diseases and covered not only differences in prevalence, but also differences in the use of health services to treat these conditions. Three nationally representative surveys, connected with the NCS-R, were conducted to further explore these issues for mental disorders. The National Survey of American Life (NSAL) interviewed 3,570 African American subjects, 1,623 African Caribbean immigrants, and 1,006 non-Hispanic white subjects. The NLAAS conducted surveys of four Latino groups (Cuban, Mexican, Puerto Rican, and “other Latino”) with a sample size of 2,554, four Asian groups (Chinese, Filipino, Vietnamese, and “other Asian”) with a sample size of 2,095, and a small control group of non-

Hispanic whites. The surveys used the WMH-CIDI, with translations into Spanish, Chinese, Vietnamese, and Tagalog for respondents who preferred to take the interview in those languages. Many articles have been published from these data sets on a variety of topics, although 12-month prevalence and service use rates for specific disorders in specific ethnic/racial groups were published in a piecemeal fashion.

World Mental Health Initiative. The paucity and noncomparability of epidemiological data on mental disorders and their consequences in many countries, particularly in developing nations was a hindrance to the WHO's Global Burden of Disease study (described below). For this reason the WMH Initiative was conceived by the WHO to increase precision in future estimates of the societal burden of mental disorders and in turn provide a more solid basis for mental health service planning and delivery of effective interventions. The initiative aimed to study the prevalence of mental disorders, their consequences in terms of disabilities and social disadvantages, and patterns of service use for mental disorders. Twenty-six countries participated in the WMH Initiative, representing all WHO regions (6 countries in the Pan-American region including the NCS-R in the United States, 2 in Africa, 2 in the Eastern Mediterranean, 12 in Europe including Israel, 1 in Southeast Asia, and 3 in the Western Pacific). Some countries chose to perform national surveys, others chose to sample from smaller geographical units in their countries. All surveys did, however, use multistage household probability sampling to ensure that their results could be generalized to their target populations. In all, over 130,000 respondents were sampled worldwide. The WMH-CIDI was the diagnostic instrument, and both DSM-IV and ICD-10 diagnoses were assessed. To increase comparability in implementation, instrument translations were done according to standard WHO protocols. Interviewers for each country were trained centrally with the same documents and protocols. Informed consent and institutional review board approvals were required. The 12-month prevalence of mental disorders in an initial group of WMH participating countries is shown in Table 5.1–8. The wide range of rates across countries, and in the case of China, between two urban areas, is a finding that needs more exploration. A recent analysis of bipolar spectrum disorders in several WMH sites offered some potential explanations for these differences. They may reflect the effects of mental illness stigma in increasing false-negative rates in some countries. Also, the validity of the WMH-CIDI diagnoses was not tested at all sites, so the ability of the instrument to detect specific diagnoses in diverse cultures is uncertain. Idioms of distress and conceptualizations of psychiatric symptoms may vary across cultures, reducing the validity of a translated common measurement instrument. It was also suggested that variations in the translation, implementation, and quality control across sites may account

for the prevalence differences. Finally, it is not possible to adequately control for the large differences in potentially relevant social factors (e.g., income, education) between countries in the WMH surveys and these differences, when uncontrolled, could account for some portion of the prevalence differences.

Table 5.1–7.
1-Year Service Use by Persons with DSM-IV Mental and Addictive Disorders, NCS-R

	% With Disorder Using Service in Sector (SE)				
Disorder	**Specialty Mental Health**	**General Medical**	**Human Services**	**Complementary and Alternative**	**Any Service Use**
Specific phobia	19.0 (1.8)	21.2 (1.5)	8.6 (0.9)	7.0 (0.8)	38.2 (1.9)
Social phobia	24.7 (1.5)	25.3 (1.7)	7.7 (1.1)	7.7 (1.0)	45.6 (1.9)
Agoraphobia w/o panic	#	#	#	#	52.6 (7.4)
Panic	34.7 (2.6)	43.7 (3.3)	10.8 (1.9)	8.0 (2.0)	65.4 (3.3)
Generalized anxiety	25.5 (2.9)	31.7 (2.6)	14.0 (3.5)	10.1 (1.8)	52.3 (2.9)
PTSD	34.4 (2.9)	31.3 (2.5)	10.7 (2.4)	12.6 (2.0)	57.4 (3.3)
Any anxiety	21.7 (1.2)	24.3 (1.0)	8.2 (0.9)	7.3 (0.6)	42.2 (1.3)
Major depressive disorder	32.9 (1.6)	32.5 (2.3)	10.7 (1.2)	9.0 (1.3)	56.8 (2.2)
Bipolar I or II	33.8 (2.3)	33.1 (3.0)	11.7 (2.2)	12.2 (2.7)	55.5 (3.0)
Dysthymia	36.8 (4.1)	39.6 (5.1)	13.3 (3.2)	7.1 (2.3)	67.5 (4.1)
Any mood	32.9 (1.3)	32.8 (1.8)	11.0 (1.2)	9.8 (1.3)	56.4 (1.8)
Intermittent explosive disorder	13.9 (2.3)	12.6 (2.4)	7.6 (2.3)	3.7 (1.2)	29.6 (2.9)
Alcohol abuse	25.6 (2.3)	16.4 (2.1)	7.0 (1.8)	7.4 (1.9)	37.2 (2.6)
Alcohol dependence	35.1 (4.4)	19.3 (3.7)	8.2 (2.8)	14.5 (3.3)	48.4 (5.4)
Drug abuse	32.8 (4.9)	21.8 (4.1)	7.1 (3.9)	7.7 (2.7)	43.1 (4.8)
Drug dependence	42.9 (10.0)	23.9 (7.3)	0	6.0 (3.6)	51.5 (9.9)
Any substance use disorder	26.2 (2.5)	18.1 (1.7)	7.8 (2.1)	7.2 (1.7)	38.1 (2.7)
Any disorder	21.7 (0.9)	22.8 (0.9)	8.1 (0.8)	6.8 (0.6)	41.1 (1.0)

#There were too few respondents with service use and obsessive-compulsive disorder or agoraphobia without panic to make estimates.

Table 5.1–8.
12-Month Prevalence of DSM-IV Mental Disorders, World Mental Health Survey Initiative[a]

	% (95% Confidence Interval)				
Country	**Anxiety**	**Mood**	**Impulse-Control**	**Substance**	**Any**

Americas Colombia	10.0 (8.4–11.7)	6.8 (6.0–7.7)	3.9 (3.2–4.7)	2.8 (2.0–3.7)	17.8 (16.1–19.5)
Mexico	6.8 (5.6–7.9)[b]	4.8 (4.0–5.6)	1.3 (0.9–1.8)[e]	2.5 (1.8-3.3)	12.2 (10.5–13.80)
United States	18.2 (16.9–19.5)	9.6 (8.8–10.4)	6.8 (5.9–7.8)	3.8 (3.2–4.5)	26.4 (24.7–28.0)
Europe Belgium	6.9 (4.5–9.4)	6.2 (4.8–7.6)[d]	1.0 (0.3–1.8)[e]	1.2 (0.6-1.9)[j]	12.0 (9.6–14.3)
France	12.0 (9.8–14.2)	8.5 (6.4–10.6)[d]	1.4 (0.7–2.0)[e]	0.7 (0.3-1.2)[j]	18.4 (15.3–21.5)
Germany	6.2 (4.7–7.6)	3.6 (2.8–4.3)[d]	0.3 (0.1–0.6)[e]	1.1 (0.4–1,7)[j]	9.1 (7.3–10.8)
Italy	5.8 (4.5–7.1)	3.8 (3.1–4.5)[d]	0.3 (0.1–0.5)[e]	0.1 (0.0-0.2)[j]	8.2 (6.7-9.7)
Netherlands	8.8 (6.6–11.0)	6.9 (4.1–9.7)[d]	1.3 (0.4–2.2)[e]	3.0 (0.7–5.2)[j]	14.9 (12.2–17.6)
Spain	5.9 (4.5–7.3)	4.9 (4.0–5.8) [d]	0.5 (0.2-0.8)[e]	0.3 (0.0–0.5)[j]	9.2 (7.8–10.6)
Ukraine	7.1 (5.6–8.6)[b,c]	9.1 (7.3–10.9)[d]	3.2 (2.4–4.0)[f,g,h]	6.4 (4.8–8.1)[j]	20.5 (17.7–23.2)
Middle East and Africa Lebanon	11.2 (8.9–13.5)	6.6 (4.9–8.2)	1.7 (0.8–2.6)[f,h]	1.3 (0.0–2.8)	16.9 (13.6–20.2)
Nigeria	3.3 (2.4–4.2)	0.8 (0.5–1.0)	0.0 (0.0–0.1)[f,g,h]	0.8 (0.3–1.2)	4.7 (3.6–5.8)
Asia					
Japan	5.3 (3.5–7.0)[b]	3.1 (2.2–4.1)	1.0 (0.4–1.5)[f,g,h,i]	1.7 (0.3–3.0)	8.8 (6.4–11.2)
People's Republic of China Beijing	3.2 (1.8–4.6)[b]	2.5 (1.5–3.4)	2.6 (1.3–3.9)[f,g,h]	2.6 (1.2–3.9)	9.1 (6.0–12.1)
Shanghai	2.4 (0.9–3.9)[b]	1.7 (0.6–2.9)	0.7 (0.4–1.1)[f,g,h]	0.5 (0.3–0.6)	4.3 (2.7–5.9)

[a]Anxiety disorders include agoraphobia, generalized anxiety disorder, obsessive-compulsive disorder, panic disorder, posttraumatic stress disorder, social phobia, and specific phobia. Mood disorders include bipolar I and II disorders, dysthymia, and major depressive disorder. Impulse-control disorders include bulimia, intermittent explosive disorder, and reported persistence in the past 12 months of symptoms of three child-adolescent disorders (attention-deficit/hyperactivity disorder, conduct disorder, and oppositional-defiant disorder). Substance disorders include alcohol or drug abuse or dependence. In the case of substance dependence, respondents who met full criteria at some time in their life and who continue to have any symptoms are considered to have 12-month dependence even if they currently do not meet full criteria for the disorder. Organic exclusions were made as specified in the *Diagnostic and Statistical Manual of Mental Health Disorders, Fourth Edition,* but diagnostic hierarchy rules were not used.

[b]Obsessive-compulsive disorder was not assessed.

[c]Specific phobia was not assessed.

[d]Bipolar disorders were not assessed.

[e]Intermittent explosive disorder was not assessed.

[f]Bulimia was not assessed.

[g]Attention-deficit/hyperactivity disorder was not assessed.

[h]Oppositional-defiant disorder was not assessed.

[i]Conduct disorder was not assessed.

[j]Only alcohol abuse and dependence were assessed. No assessment was made of other drug abuse or dependence.

National Epidemiological Survey on Alcohol and Related Conditions. The National Institute of Alcohol Abuse and Alcoholism's NESARC is a series of population-based surveys of alcohol use disorders and other substance use and mental disorders in the United States. Among its goals were to provide prevalence and incidence rates of alcohol-related disorders and associated disabilities; examine comorbidity patterns with other mental disorders, including other substance use disorders; document the natural history of alcohol-related disorders and disabilities; examine service use rates and patterns and barriers to care; and attempt to elucidate and differentiate "alcohol-induced" physical and mental disorders from "independent" disorders. The NESARC has been fielded three times. The first survey, with an unprecedented sample size of 43,093, was fielded from 2001 to 2002. This sample was nationally representative of the adult household population of the United States, also including noninstitutional group quarters such as boarding and rooming houses, nontransient hotels, shelters, worker housing, college quarters, and group homes. The second survey aimed to re-interview all first wave respondents and was fielded in 2004 to 2005. A total of 34,653 interviews were completed, amounting to a response rate of 86.7%. The data from the second wave of the survey has allowed estimation of incidence rates and risk factors for the development of alcohol, drug and other mental disorders, as well as examination of the course and outcome of treated and untreated mental and substance use disorders. One-year prevalence rates from the first NESARC survey are listed in Table 5.1–6. One-year incidence rates based on the first two surveys are listed in Table 5.1–9. NESARC-III was fielded from 2012 to 2013. The target population was the U.S. civilian, noninstitutionalized population ages 18 and over, with an oversampling of Black, Hispanic, and Asian populations. The total sample size was 36,309, representing an overall response rate of 60.1%. Psychometric studies of the AUDADIS-5 are currently the only publications from the NESARC-III.

Table 5.1–9.
12-Month Incidence of DSM-IV Psychiatric Disorders, NESARC Wave 2

DSM-IV Disorder	Unweighted Number of Respondents at Risk	Unweighted Number of 12-Month Incident Cases	Weighted Annual Incidence per 100 Person-Years
Alcohol use			

disorder[a]	23,713	363	1.66 (0.11)[b]
Alcohol abuse	24,291	233	1.03 (0.10)
Alcohol dependence	30,220	481	1.72 (0.09)
Any drug use disorder[a]	30,830	89	0.31 (0.04)
Any drug abuse	30,887	75	0.28 (0.04)
Any drug dependence	33,639	101	0.32 (0.04)
Any mood disorder[a]	26,867	666	2.24 (0.12)
Major depressive disorder	28,859	489	1.52 (0.09)
Bipolar I	32,990	189	0.53 (0.05)
Bipolar II	34,045	74	0.21 (0.03)
Any anxiety disorder[a]	26,179	452	1.58 (0.09)
Any panic disorder	32,302	224	0.62 (0.05)
Social phobia	32,352	112	0.32 (0.04)
Specific phobia	29,502	147	0.44 (0.05)
Generalized anxiety	32,361	390	1.13 (0.08)

[a]Because the population at risk for each aggregate disorder category excludes all respondents with prior lifetime histories of any disorder in that category, the denominator of the incidence rate is invariably smaller than the denominator for any component disorder. For the same reason, and because, individuals can become incident cases of more than one disorder in each category in the same time period, the numerator for the aggregate rate is not a simple sum of the number of incident cases across the specific disorders.

[b]Numbers in parentheses are standard errors.

From Grant BF, Goldstein RB, Chou SP, et al. Sociodemographic and psychopathologic predictors of first incidence of DSM-IV substance use, mood and anxiety disorders: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *Mol Psychiatry.* 2009;14(11):1051–1066.

Surveys of Mental Disorders in Children and Adolescents

Concerted research efforts on child and adolescent mental disorders started later than on adult disorders, and the knowledge base is correspondingly smaller for epidemiological and treatment studies alike. The reasons for this are multifaceted, but as noted above, mental disorders in these populations were considered to be relatively rare until the ECA and other studies in the 1980s revealed much younger ages of onset for common mental disorders than was previously assumed. Diagnostic criteria for children and adolescents were widely criticized by clinicians and researchers for being difficult to operationalize and not reflecting

developmental considerations and clinical reality, deficiencies that the DSM-5 attempted to remedy. In addition, the paucity of treatment services for children and adolescents hid what was later found to be a high prevalence of untreated mental disorders. However, interest in the diagnosis and the epidemiology of child and adolescent mental disorders has increased since the fielding of the ECA program. In fact, researchers have been able to take advantage of, and contribute to, advances in diagnostic criteria and research methods since the DSM-III and the ECA, by developing sophisticated diagnostic instruments and study designs. Unfortunately, no epidemiological studies comparable to the ECA, NCS, and NESARC have been fielded in child and adolescent populations to take full advantage of these research developments.

Methods for the Epidemiology of Child and Adolescent Mental Disorders.

In the wake of the successes of the ECA program, the need to stimulate the field of child and adolescent psychiatric epidemiology with a "child ECA" was recognized as a crucial next step by the NIMH. Also recognized, however, was that diagnosing children and adolescents in community settings would require different methods than were used for adults in the ECA. It was assumed that, depending on the developmental level of the children surveyed, a parent or primary caregiver might be needed as an informant on diagnostic and service use variables, and that other informants such as teachers might give valuable additional information. A major issue to be clarified was how to gather this information in a large-scale survey, and then how to best combine these multiple-source reports to arrive at reliable and valid diagnoses. The reliability and validity of the available diagnostic instruments also needed to be confirmed. Developing and evaluating assessments of disability and service use also needed to be addressed. NIMH set up a two-phase program, the first phase of which was a methodological evaluation addressing the issues named above, and the second phase a full multisite survey, the "child ECA." The MECA survey was a cooperative agreement between the NIMH and four academically based sites in New York City, New Haven, Atlanta, and San Juan, which was funded to address these methodological issues in child and adolescent psychiatric epidemiology. The MECA surveys included children ages 9 to 17 years. The NIMH DISC, version 2.3, was used as the core diagnostic assessment instrument for DSM-III-R disorders, and an instrument to assess service use, sociodemographic factors, and potential risk and protective factors was developed. The CGAS and the Columbia Impairment Scale were also included in the survey to assess disability. Only parent and child reports were obtained, mainly due to a lack of funds to obtain other informant reports. Data collection was completed in 1992, with 1,285 parent–child pairs being interviewed. Reliability and validity tests were set up to assess the DISC's performance with both clinicians and lay

interviewers. The DISC 2.3 was found to be have acceptable reliability and criterion validity, and several reports on methodology, diagnosis, disability, and service use were published. Despite the successes of the MECA survey, the anticipated "child ECA" was never fielded. The results of the longitudinal Great Smoky Mountains Study have also provided extremely valuable descriptive and analytic data in a sample that has been followed longitudinally, but this study's sampling was limited to a geographical area in North Carolina.

Table 5.1–10.
Estimates of 12-Month Prevalence of DSM-IV and CIDI Disorders in Adolescents, Ages 13–17, NCS-A

Disorder	Prevalence % (SE)
Any mood disorder	10.0 (0.8)
Major depressive disorder or dysthymia	8.2 (0.8)
Bipolar I or II disorder	2.1 (0.2)
Any anxiety disorder	24.9 (0.9)
Agoraphobia	1.8 (0.2)
Generalized anxiety disorder	1.1 (0.2)
Social phobia	8.2 (0.4)
Specific phobia	15.8 (0.8)
Panic disorder	1.9 (0.2)
Posttraumatic stress disorder	3.9 (0.4)
Separation anxiety disorder	1.6 (0.2)
Any behavior disorder	16.3 (1.1)
Attention-deficit/hyperactivity disorder	6.5 (0.5)
Oppositional-defiant disorder	8.3 (0.7)
Conduct disorder	5.4 (0.8)
Eating disorder	2.8 (0.2)
Any substance disorder	8.3 (0.5)
Alcohol abuse with or without dependence	4.7 (0.3)
Drug abuse with or without dependence	5.7 (0.5)
Any disorder	40.3 (1.2)

From Kessler RC, Shelli Avenevoli S, Costello EJ, et al. Prevalence, persistence, and sociodemographic correlates of *DSM-IV* disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Arch Gen Psychiatry.* 2012;69(4):372–380.

National Comorbidity Survey Replication Adolescent Supplement. The NCS-A was a supplemental survey to the NCS-R, conducted from 2001 to 2004. As such, it sampled youth ages 13 to 17 years from the households selected for the NCS-R, and from a representative sample of schools located in the counties where the household surveys were conducted. The final

sample size from households was 904, an 86.8% response rate. From the 320 schools sampled, 9,244 youth were selected, an 82.6% response rate. The youth were administered a modified version of the CIDI that had alterations in phrasing and examples to make it more appropriate to the respondents' language level and life experiences. Some disorders were dropped from the CIDI for this population, such as pathological gambling, obsessive-compulsive disorder, and psychotic disorders. A parent or adult surrogate was given a self-administered assessment, but not directly interviewed. This assessment was relatively brief, covering the youth's symptoms in categories for which adult reports of psychopathology were felt to be particularly important for making a diagnosis (major depressive episode, dysthymia, separation anxiety disorder, attention-deficit/hyperactivity disorder, conduct disorder, and oppositional-defiant disorder). The adult also filled out the Strengths and Difficulties Questionnaire. Publications using NCS-A data have been slowly accumulating since the initial publication of findings in 2009. One-year prevalence estimates from the survey are in Table 5.1–10.

WHO Global Burden of Disease Study

The innovative Global Burden of Disease (GBD) studies are perhaps the best examples of the use of population-based epidemiological surveys to serve the needs of public health policy and planning. There have been two GBD studies, the first published in 1996 and sponsored by the WHO and the World Bank, and the second (GBD 2010) published in 2012 and sponsored by the Bill and Melinda Gates Foundation. The GBD studies recognized that the consequences of disease include not only death but also disability, and the impact of disability should be quantified and factored into any estimates of disease burden. By taking disability into account, these studies served to draw attention to chronic diseases, including mental and substance use disorders, which may ultimately cause premature death but often cause significant societal burden through reduced role functioning. Disability adjusted life years (DALYs) were used to estimate the burden of disease. DALYs represent the sum of years of life lost to premature death (YLL) and years lived with a disability of a specified severity and duration due to the condition (YLD). One DALY represents 1 year of healthy life lost. GBD 2010 was a more ambitious project compared to the first study, expanding the number of health conditions examined from 110 to 291. This expanded number included mental and substance use disorders: compared to the 1990 study, the 2010 study added six more anxiety disorders, eating disorders, pervasive developmental disorders, childhood behavior disorders, dysthymia, and bipolar disorder subtypes. For substances, cocaine, amphetamine, and opioid dependence were analyzed separately and cannabis dependence was added. GBD 2010 also added new computations that accounted for the burden of mental and

substance use disorders as risk factors for other disorders. Four risk factors were quantified: major depressive disorder as a risk factor for ischemic heart disease; cannabis use as a risk factor for schizophrenia; major depressive disorder, bipolar disorder, anxiety disorder, anorexia nervosa, schizophrenia, cocaine dependence, amphetamine dependence, and opioid dependence as risk factors for suicide; and injection drug use as a risk factor for hepatitis B and C and HIV.

To calculate the global burden of mental and substance use disorders, the study investigators performed extensive literature searches for data on prevalence, incidence, remission and duration, and all-cause excess mortality. The searches were limited to the years 1980–2011 for mental disorders and 1990–2011 for substance use disorders. All languages of publication were accepted. Publications were required to use DSM or ICD diagnostic criteria, to employ representative sampling of a community, and to provide sufficient information on study methods and sample characteristics. Prevalence was required to be reported in a current (point or past month) or past-year time frame. Incidence data had to reflect annual incidence rates or be presented as person-years, and be based on prospective cohort or longitudinal studies. For remission studies, longitudinal follow-up must have been for 2 years or more and not included an intervention. Mortality estimates were required to reflect mortality from all causes. Observational and analytic studies were accepted, but not psychological autopsy studies, case studies, and treatment trials. The GBD 2010 investigators identified almost 77,000 articles related to mental and substance use disorders, but less than 1% of these met the inclusion criteria for use in the GBD study. The main reasons for noncompliance were not using DSM or ICD diagnostic definitions, nonrepresentativeness of the samples, use of measures that did not allow comparability between studies, and insufficient information on methods. Not surprisingly, the largest number of acceptable data sources for prevalence data were from high income regions, particularly North America, Australasia, Western Europe, and high-income Asia Pacific countries, with data coverage being better for high prevalence disorders such as major depressive disorder, compared to low prevalence disorders such as schizophrenia. Prevalence data coverage from other global regions was poor to absent. Considerably fewer studies related to incidence, remission, and mortality were found; with few exceptions, these were located in high-income countries. A Bayesian meta-regression method was then used to model prevalence by disorder type, age, gender, year, region, and country. This same regression method was also used to generate prevalence estimates for the countries and regions for which no prevalence data were available.

Table 5.1–11.

Proportion of YLDs, YLLs, and DALYs Explained by the 10 Leading Causes of Total Burden in 2010

	Proportion of Total DALYs (95% UI)	Proportion of Total YLDs (95% UI)	Proportion of Total YLLs (95% UI)
Cardiovascular and circulatory diseases	11.9% (11.0–12.6)	2.8% (2.4–3.4)	15.9% (15.0–16.8)
Diarrhea, lower respiratory infections, meningitis, and other common infectious diseases	11.4% (10.3–12.7)	2.6% (2.0–3.2)	15.4% (14.0–17.1)
Neonatal disorders	8.1% (7.3–9.0)	1.2% (1.0–1.5)	11.2% (10.2–12.4)
Cancer	7.6% (7.0–8.2)	0.6% (0.5–0.7)	10.7% (10.0–11.4)
Mental and substance use disorders	7.4% (6.2–8.6)	22.9% (18.6–27.2)	0.5% (0.4–0.7)
Musculoskeletal disorders	6.8% (5.4–8.2)	21.3% (17.7–24.9)	0.2% (0.2–0.3)
HIV/AIDS and tuberculosis	5.3% (4.8–5.7)	1.4% (1.0–1.9)	7.0% (6.4–7.5)
Other noncommunicable diseases	5.1% (4.1–6.6)	11.1% (8.2–15.2)	2.4% (2.0–2.8)
Diabetes, urogenital, blood, and endocrine diseases	4.9% (4.4–5.5)	7.3% (6.1–8.7)	3.8% (3.4–4.3)
Unintentional injuries other than transport injuries	4.8% (4.4–5.3)	3.4% (2.5–4.4)	5.5% (4.9–5.9)

DALYs, disability-adjusted life-years; YLDs, years lived with disability; YLLs, years of life lost.

From Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet*. 2013;382(9904):1575–1586.

The method for calculating disability also was changed for GBD 2010. The first GBD study ranked disabilities according to a person tradeoff method that involved the judgments of 22 conditions by health care workers from around the world. Each participant in this exercise was asked a question about extending the life of people with a specified health state versus extending the life of healthy people, and a question about giving health back to persons with an impaired health state versus extending life for healthy people. For GBD 2010, population-based surveys were conducted in the United States, Tanzania, Bangladesh, Indonesia, and Peru, and an internet survey was opened for speakers of English, Mandarin, or Spanish. The participants in the surveys were asked to compare two randomly selected health conditions and state which they considered healthier, and responses to these questions provided the basis for the calculation of disability weights. For some disorders, disorder subgroups were presented separately, such as mild, moderate and severe major depressive disorder, or manic, depressive, and residual phases of bipolar disorder.

Tables 5.1–11 and 5.1–12 show some of the main results of GBD 2010. Cardiovascular, circulatory, and infectious diseases were the major causes of disease burden worldwide. However, mental disorders were a major source of burden as well, accounting for roughly the same percentage of total DALYs as neonatal disorders, cancer, and musculoskeletal disorders. Mental disorders and musculoskeletal disorders were the top two causes of disability in the world in 2010. Looking solely at mental disorders, the leading sources of DALYs were the depressive disorders, followed by anxiety disorders, drug and alcohol use disorders, schizophrenia, and bipolar disorder. This pattern was also seen for YLDs. Alcohol and drug use disorders accounted for the majority of YLLs; data were sufficient to attribute only five disorders as a cause of death. The major causes of disease burden for developed nations were chronic diseases, while those for developing nations were a mix of infectious, perinatal, and chronic disease. As developing nations become more developed and the prevalence of infectious and perinatal diseases decreases, the importance of early death can be predicted to decrease in DALY calculations, and years lived with disabilities will become more important. Accordingly, the burden of mental disorders, with their high levels of disability, can be expected to increase in these nations. Overall, the results of the Global Burden of Disease 2010 study reinforced findings of the GBD 1990 study in demonstrating the large societal burden of mental disorders in the context of other nonpsychiatric disorders. Health policies in the United States and most other countries have yet to address the magnitude of this burden.

Table 5.1–12.
Proportion of YLDs (A), YLLs (B), and DALYs (C) Explained by Each Mental and Substance Use Disorder Group in 2010

	YLDs	YLLs	DALYs
Depressive disorders	42.5 (33.3–51.7)		40.5 (31.7–49.2)
Anxiety disorders	15.3 (11.7–19.3)		14.6 (11.2–18.4)
Drug use disorders	9.4 (7.3–11.5)	41.7 (27.9–56.9)	10.9 (8.9–13.2)
Alcohol use disorders	7.9 (6.0–10.0)	44.4 (29.1–60.0)	9.6 (7.7–11.8)
Schizophrenia	7.4 (4.9–9.9)	7.1 (4.8–10.2)	7.4 (5.0–9.8)
Bipolar disorder	7.4 (4.6–10.8)		7.0 (4.4–10.3)
Pervasive developmental disorders	4.4 (3.4–5.6)		4.2 (3.2–5.3)
Childhood behavioral disorders	3.5 (2.3–4.9)		3.4 (2.2–4.7)
Eating disorders	1.1 (0.8–1.5)	2.4 (1.4–3.4)	1.2 (0.9–1.5)
Other mental disorders	0.6 (0.4–1.1)	4.3 (2.4–6.3)	0.8 (0.5–1.2)
Idiopathic intellectual disability	0.6 (0.4–0.9)		0.6 (0.3–0.9)

Data are % (95% UI). DALYs, disability-adjusted life years; YLDs, years lived with disability; YLLs, years of life lost.

From Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet.*

2013;382(9904):1575–1586.

REFERENCES

Alegria M, Takeuchi D, Canino G, et al. Considering context, place and culture: the National Latino and Asian American Study. *Int J Methods Psychiatr Res*. 2004;13:208–220.

*Caspi A, Sugden K, Moffitt TE, et al. Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT Gene. *Science*. 2003;301:386–389.

Cooper R, Kendell RE, Gurland BJ, Sharpe L, Copeland JRM, Simon R. *Psychiatric Diagnosis in New York and London: A Comparative Study of Mental Hospital Admissions. Institute of Psychiatry, Maudsley Monograph, No. 20*. London: Oxford University Press; 1972.

Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. *Arch Gen Psychiatry*. 2003;60:837–844.

Demyttenaere K, Bruffaerts R, Posada-Villa J, et al. WHO World Mental Health Survey Consortium: Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *JAMA*. 2004;291:2581–2590.

*Eaton WW, Kessler LG, eds. *Epidemiologic Field Methods in Psychiatry*. Orlando, FL: Academic Press; 1985.

Fleiss JL. *Statistical Methods for Rates and Proportions*. New York: John Wiley & Sons; 1981.

Freeman C, Tyrer P, eds. *Research Methods in Psychiatry*. 3rd ed. London: Gaskell; 2006.

Grant BF, Goldstein RB, Chou SP, et al. Sociodemographic and psychopathologic predictors of first incidence of DSM-IV substance use, mood and anxiety disorders: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *Mol Psychiatry*. 2009;14:1051–1066.

Grant BF, Stinson FS, Dawson DA, et al. Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders. *Arch Gen Psychiatry*. 2001;61:807–816.

Jackson JS, Torres M, Caldwell CH, et al. The National Survey of American Life: a study of racial, ethnic and cultural influences on mental disorders and mental health. *Int J Methods Psychiatr Res*. 2004;13:196–207.

Kahn HA, Sempos CT. *Statistical Methods in Epidemiology*. New York: Oxford University Press; 1989.

Kelsey JL, Thompson WD, Evans AS. *Methods in Observational Epidemiology*. New York: Oxford University Press; 1986.

*Kessler RC, Avenevoli S, Costello EJ, et al. Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Arch Gen Psychiatry*. 2012;69:372–380.

*Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62:617–627. Erratum in: Arch Gen Psychiatry. 2005;62:709.

Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry*. 1994;51:8–19.

Kessler RC, Üstün TB. The World Mental Health (WMH) Survey Initiative version of the World

Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res*. 2004;13:93–121.

Kraemer HC. Statistical issues in assessing comorbidity. *Stat Med*. 1995;14:721–733.

Lahey BB, Flagg EW, Bird HR, et al. The NIMH methods for the epidemiology of child and adolescent mental disorders (MECA) study: Background and methodology. *J Am Acad Child Adolesc Psychiatry*. 1996;35:855–864.

Löwe B, Spitzer RL, Williams JB, Mussell M, Schellberg D, Kroenke K. Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. *Gen Hosp Psychiatry*. 2008;30:191–199.

Morris JN. *Uses of Epidemiology*. Baltimore, MD: Williams and Wilkins; 1964.

*Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380:2197–2223.

Narrow WE, Rae DS, Robins LN, Regier DA. Revised prevalence estimates of mental disorders in the U.S.: Using a clinical significance criterion to reconcile two surveys' estimates. *Arch of General Psychiatry*. 2002;59:115–123.

Regier DA, Narrow WE, Rae DS, Manderscheid RW, Locke BZ, Goodwin FK. The de facto US mental and addictive disorders service system: epidemiologic catchment area prospective 1-year prevalence rates of disorders and services. *Arch of General Psychiatry*. 1993;50:85–94.

Robins E, Guze SB. Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *Am J Psychiatry*. 1970;126:983–987.

Robins LN, Wing J, Wittchen HU, et al. The composite international diagnostic interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch General Psychiatry*. 1988;45:1069–1077.

Shaffer D, Fisher P, Dulcan M, et al. The NIMH diagnostic interview schedule for children, version 2.3 (DISC 2.3): Description, acceptability, prevalence rates and performance in the MECA study. *J Am Acad Child Adolesc Psychiatry*. 1996;35:865–877.

*Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62:629–640.

*Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet*. 2013;382:1575–1586.

World Health Organization. *International Classification of Functioning, Disability and Health: ICF*. Geneva: World Health Organization; 2001.

▲ 5.2 Statistics and Experimental Design

EUGENE M. LASKA, PH.D., CAROLE E. SIEGEL, PH.D., MORRIS MEISNER, PH.D., AND ISAAC R. GALATZER-LEVY, PH.D.

Jerry Cornfield declared that statistics is the bedfellow of the sciences in his presidential address to the American Statistical Association in 1974. Even a casual reader of articles that appear in the medical and psychiatric literature becomes aware of arcane terms such as chi square and analysis of variance (ANOVA) and *p* values. The lexicon of statistics is now part of the common parlance of editors, reviewers, and authors, so to appraise and appreciate a scientific article, it is incumbent on the reader to have some familiarity with the lingua franca. Of course, it may be argued, the user of a computer need not understand the bits and bytes and operating system that drive the processor to use it, so why can't the reader simply accept that the referees and editors have done their jobs? Indeed, why can't the reader just peruse the abstract and the discussion section and be done with it? Well, truth be told, he or she can. A word processor can be used successfully even by one with limited computer skills. A scientific report can be skimmed to obtain the essence of its message, without needing to appreciate the nuances of the methods utilized. But that strategy will not enable an independent view of the strength of the evidence nor of the credibility of the findings. The reputation of the author and the journal may provide some comfort, and perhaps in reports related only peripherally to the reader's interests that is all that is required. But at least in one's own field of expertise, a full understanding of the techniques used to obtain the results and to reach the conclusions should be accessible. If the reader wants to really understand the material, then at least a rudimentary understanding of statistical concepts is necessary.

So what are the must-know elements of statistics? It is hard to say. It depends on the area of science or therapeutics that is of interest. The discipline of statistics has become so vast that few if any professional can pretend to be an expert in all of its parts. But a psychiatrist should have at least a basic understanding of the rudiments. This chapter will not try to replicate what countless articles, books, and abbreviated introductory courses do. The basic ideas are described without any assumptions about the mathematical skill level of the reader in many elementary texts and chapters. It is assumed that the reader already has some familiarity with the basic concepts, but nevertheless the chapter begins with a very brief review of some elementary statistical ideas. The remainder of the chapter discusses some of the many topics in statistics selected in the hope that they will enlarge the capacity of a reader to make independent judgments

and interpretations about the articles that he or she reads. Necessarily, the presentations are not complete, the choice of topics is somewhat arbitrary, and there are many more topics that could have been included in this chapter. It is the authors' hope that this presentation will stimulate further study of statistical methods particularly by individuals interested in psychiatric research. Although the authors have strived to keep technical details to a minimum, occasionally some complexities are unavoidable.

BASIC IDEAS

In this section the notions of elementary probability theory are reviewed and the concepts of a random variable, measures of central tendency, and a measure of dispersion are introduced. To describe the joint behavior of two random variables, the ideas of statistical independence, covariance, and correlation are necessary and enable the introduction of a multivariate density.

Probability Theory

The mathematical foundation on which statistical methods are built is probability theory. This theory considers an experimental situation in which the outcomes are random events. The set of all possible simple outcomes of the experiment is called the *sample space*. For example, throwing a die once gives rise to a sample space containing six elements: 1, 2, 3, 4, 5, 6. A random variable is a mapping of the sample space on to the real line. It is usually denoted by an upper case Latin letter such as X, and an actual outcome is usually denoted by its corresponding lower case letter x. Thus, X is the random outcome of the face value resulting from the roll of the die and $x = 6$ is an observed outcome. Assigned to each simple outcome is a positive number corresponding to the probability of that outcome. The sum of these probabilities over all *simple* outcomes in the sample space is 1. The assignment of a probability to a simple outcome automatically determines the probability of all subsets of the sample space. For example, the probability that the outcome of the roll of the die is an odd number is the probability that $X = 1$ plus the probability that $X = 3$ plus the probability that $X = 5$.

Probability Density Functions

A characterization of the random nature of the experiment is provided by the chance of observing a particular value x. If there are only a finite number of possible outcomes, the probability of each event x, $P(X = x)$, can be positive and as a function of x, it is called a probability density function. If the range of outcome values includes all values in an interval, sometimes called the *continuous case,* the probability density function, generically denoted by $f(x)$, has a different meaning. This is because the probability of

observing any particular single value x, possibly with a finite (countable) number of exceptions, is 0. In this case, the integral of $f(x)$ over the interval is equal to the probability that an outcome falls in that interval. The *cumulative distribution function,* frequently represented by $F(x)$, is the probability that the random variable takes a value less than or equal to x. It is obtained in the discrete case by summing the probabilities of all possible outcomes that are less than or equal to x and in the continuous case by integrating the probability density function (PDF) from minus infinity to x. All probabilities must be nonnegative and their sum over all possible events, or in the continuous case the integral over the sample space, must be 1. Therefore, as x goes to infinity $F(x)$ goes to 1. The frequentist interpretation of probability is that in the long run, the fraction of times that the outcome x is observed (or that x falls in a specific interval $[a,b]$ of the sample space) in N replications of the experiment, tends to $f(x)$ (or to the integral of the PDF from a to b) as N tends to infinity.

Conditional Probability and Independence

Many compound events can be defined from the simple outcomes in an experiment. In the roll of the die experiment, the event that an odd number occurs or the event that either a 1 or a 5 occurs are two examples. In some cases knowing that one event has occurred can impact the likelihood of another event occurring. For example, if the faces of the die are equally likely (all have probability 1 of 6), then the probability of a 1 or 5 occurring given that an odd number has occurred is 2 of 3, while the probability of a 1 or 5 occurring with no other event information is 1 of 3. The random variable X that takes the value 1 if the die falls on face 1 or 5, and 0 otherwise, and the random variable Y that takes the value 1 if the die falls on an odd numbered face are said to be dependent. Two random variables X and Y are said to be independent if $P(X = x|Y = y) = P(X = x)$. The notation $P(X = x|Y = y)$ is read "the probability that $X = x$ given (or conditional on) $Y = y$." It is the probability that the event $X = x$ occurs calculated in the outcome space defined by the event $Y = y$. Formally, $P(X = x|Y = y) = P(X = x \text{ and } Y = y)/P(Y = y)$. Also, two random variables are independent if the probability of the joint event $X = x$ and $Y = y$ is equal to the product of the probabilities of the two events, that is, $P(X = x \text{ and } Y = y) = P(X = x)P(Y = y)$. Independence and conditional probabilities play a major role in statistical thinking. For example, suppose treatment assignment and degree of improvement on a five level clinical global scale (CGI) is reported in a table with five rows (CGI) and two columns (treatment or placebo). A chi-square statistic is used to analyze such contingency tables for testing the independence of two random variables that define the rows and columns. If the statistic is significant, the data suggest that the variables are dependent, and the distribution of CGI outcomes conditioned on a value of treatment assignment differs

depending on the value of the conditioning variable. If treatment and placebo improvement rates differ, then one would expect a chi-square test to be significant because the distribution of CGI conditional on receiving the treatment is not the same as the distribution conditional on receiving placebo.

Parametric Forms

In parametric analyses, the *form* of the PDF is assumed known with the exception of the value of one or more parameters, which, once specified, completely determines the distribution. The most commonly assumed parametric form is the so-called gaussian or normal distribution. The shape of the normal PDF is the familiar bell-shaped curve, whose peak (mode), mean, and median occur at the same value. In fact the curve merely translates to the right or left but retains the same form as the mean is respectively increased or decreased. Distributions with this property are called *translation parameter families* and they have the mathematical form $f(x - \phi)$, where ϕ is the mean of the random variable X. Greek letters, such as ϕ, are often used to represent parameters of the PDF to distinguish them from random variables. The logistic, beta, exponential, F and t, for continuous random variables, and the binomial, multinomial and Poisson for discrete random variables are PDFs that frequently appear in the scientific literature.

Mean, Median, and Variance

A variety of measures that help characterized the nature of a random variable have been defined. Among the most important are the mean, median, and variance. The *mean* is the average of all possible outcomes, weighted by the probability of the outcome. The mean is thought of as a measure of central tendency of the random variable. It is also called the *expected value,* and if X is the random variable, then it is usually denoted by $E(X)$. The value m is the median if the probability that the random variable X is less than m equals the probability that X is greater than m. The *variance* is the average of the square of the difference between each outcome x and the mean, weighted by the probability of the outcome. The variance, usually denoted by σ^2, and its square root, σ, the standard deviation, are thought of as measures of the spread of the observations about the mean.

Multivariate Random Variables

Most serious investigations of scientific phenomena cannot be described by the outcome of a single random variable. A multivariate probability distribution describes the chance of various outcomes of an experiment in which k events are observed on each "unit" in a sample. Typically in trials

in mental health, a unit is a patient and observations are made on one measure collected over time, many measures collected at one point in time, or many measures collected at multiple times. Two random variables, X_1 and X_2, may be independent or dependent. As discussed above, they are independent if and only if their joint (simultaneous) probability function, $f(x_1, x_2)$, is always equal to $f_1(x_1)$ times $f_2(x_2)$. The probability density functions $f_1(x_1)$ and $f_2(x_2)$ are called the *marginal distributions* of X_1 and X_2.

Covariance and Correlation

One important measure of the relationship between dependent pairs of random variables X_1 and X_2 is the covariance, the average of the product of the difference between each outcome x_1 and its mean and the difference between each outcome x_2 and its mean, weighted by the probability of the joint outcome x_1, x_2. Closely related is the correlation, usually denoted by ρ, which is defined to be the covariance divided by the product of the standard deviation of X_1 and the standard deviation of X_2. The distribution most widely used in multivariate applications is, by far, the multivariate normal. If the mean and variance of each of the random variables and the covariance between all pairs of random variables are known, then the multivariate normal distribution is completely specified.

USING DATA TO ESTIMATE PARAMETERS

The scientific question under consideration must be represented by a statistical model, which at least approximately characterizes the set of outcomes that can be observed after the data collection is complete. If the model, represented in parametric form as the PDF $f(x, \phi)$ is fully known, there is little left to do. Neither an experiment nor statistical inference is needed. The probability distribution is the answer to the question: If the experiment is performed, what can be expected with what probability? But the true value of the distribution is rarely known and the reason for the experiment is to collect data so that something can be said about the nature of the PDF and the value of the parameters that specify it. This section discusses the problem of parameter estimation. It begins with a description of the method of maximum likelihood. Some of the properties that are used to appraise the merits of an estimator are described. Confidence interval (CI) estimators and their interpretation are introduced.

Point Estimators

In many application areas, the general form of the parametric distribution that describes the collected data is well established. After an experiment is performed, the first question an investigator should attempt to answer is

what values of the unknown parameters within the assumed PDF form best represents the data. This is the setting of the theory of statistical estimation, which has a rich and long history. Of course the principal goal is to come as close to the true value of the parameter as possible. Estimation theory is concerned with the problem of how to obtain an estimator and appraise its worth, not in an individual case but over all possible outcomes, weighted by the probability of the outcome. Perhaps the premier approach, introduced by Sir Ronald Fisher, is maximum likelihood estimation (MLE). The idea is both intuitive and simple. Suppose that the parameter ϕ is unknown and $f(x, \phi)$ is the probability of observing the event x. The likelihood is just $f(x^*, \phi)$ where x^* is the value of the observed outcome. Here x^* is fixed and f is thought of as a function that depends only on ϕ. For any given ϕ, the likelihood is a measure of how "likely" the observation x^* is had the outcome been selected randomly from the probability distribution f with parameter ϕ. If the observed random variables are independent, then the likelihood is the product of the likelihoods of each observation. It is a measure of the probability of the joint occurrence of x_1^*, x_2^*, ..., and maximum likelihood theory posits that ϕ should be chosen so that the probability of the event that was actually observed is as large as possible. Any other choice of parameters would make the observed data set less likely to have been observed. For example, if a coin is tossed 20 times, landing heads 10 of the times, setting the probability of a head, *P,* equal to 10 of 20, or 0.5, is perfectly sensible. In fact, under the assumption that the distribution of the random variable (the number of heads) is binomial, it makes the chance of 10 heads in 20 tosses maximal. Setting it equal to a lower or higher number makes the estimator less consonant with the data. In a great many cases, the properties of the MLE are very desirable.

Properties of Estimators

Since an estimator is itself the result of a random experiment, it is a random variable and so it has a mean or expected value. The theoretical calculation of its mean involves the use of the density $f(x, \phi)$, where the true value of ϕ is unknown. The estimator is referred to as *unbiased* if for each value of the parameter ϕ, the mean of the estimator is ϕ. Such estimators exhibit a kind of impartiality that is usually desirable. Another property of an estimator concerns its behavior as the sample size of the experiment increases. Imagine a sequence of estimators in which the *i*th estimator is based on a sample of size *i*. For example, the *i*th estimator in the sequence could be the sample mean based on a sample of size *i*. An estimator is said to be *consistent* if for any true value ϕ the sequence converges to ϕ. Another property of an estimator concerns its precision, as measured by its variance. An estimator is said to be *efficient* if it is unbiased and its

variance is as small as it can possibly be compared to any other unbiased estimator. In some situations a lower bound for the variance of an estimator can be determined. In this circumstance, the *relative efficiency* of an estimator can be determined by comparing its variance to the theoretical minimum variance. Consistency suggests that as the sample size increases, the variance decreases and the estimator approaches ϕ.

Confidence Interval Estimation

Even if an estimator has all of the most desirable properties, it is not likely that it will be exactly right. To leave a little wiggle room, an interval estimate, in addition to a point estimate, is desirable. These are termed *confidence interval* (CI) estimators. The amount of confidence placed on the estimator, in the sense that the parameter being estimated lies within the interval, is called the *confidence level* or *confidence coefficient*. The usual form of a CI is $P[a \leq \phi \leq b] = c$. If $c = 0.95$, this is read: $[a,b]$ is a 95 percent CI for the parameter ϕ or; $[a,b]$ is a CI for ϕ with confidence coefficient 0.95. Typically for a fixed sample size, as the confidence coefficient increases, so too does the length of the interval. An interval from minus infinity to plus infinity with confidence coefficient 1 always covers the unknown parameter, but it is completely uninformative.

Interpretation of a Confidence Interval

It is important to understand how to interpret a CI. From a frequentist point of view, the probability of an event is characterized by the frequency with which the event occurs in infinitely many replications of the experiment. The form of the interval before the experiment is conducted is usually given in terms of some not yet observed random variable. For a normal random variable with mean ϕ and variance 1, the probability statement $P[X- -1.96/\sqrt{N} \leq \phi \leq X- + 1.96/\sqrt{N}] = 0.95$, where $X-$ is the arithmetic mean of N observations, has a clear and unambiguous meaning. Suppose that a random sample of $N = 100$ observations is repeatedly drawn from a population governed by a normal distribution with mean ϕ and variance 1. As the number of repeated draws becomes very large, regardless of the true value of ϕ, the proportion of times that the interval $[X- - 1.96/\sqrt{N} \leq \phi \leq X- + 1.96/\sqrt{N}] = [X- - 0.196, X- + 0.196]$ covers ϕ approaches 0.95. Once data are collected, however, the game has changed. The statement $P[0.304 \leq \phi \leq 0.696]$ can only take the value 1 or 0, depending on whether or not the true value of ϕ falls in the interval. So, from a frequentist point of view, it is incorrect to think of the confidence coefficient of a CI estimate as the probability that the observed interval includes the true value of the parameter. Rather, it is a statement about the joint probability that the random variable that defines the lower limit falls below ϕ and the random variable that defines the upper limit of the interval

falls above ϕ (before data are collected) for any true value of ϕ.

USING DATA TO TEST HYPOTHESES

A statistical hypothesis is a statement about the possible values of the parameters of a probability distribution. In this section, test statistics are introduced and the potential errors that may arise in testing a statistical hypothesis are discussed. The roles of power and the ubiquitous *p* value in hypotheses testing are presented.

Hypothesis Testing

Sometimes interest is focused on answering a scientific question about a random variable with a (qualified) yes or no. Is drug 1 more effective than drug 2? In such cases, the goal often is to determine whether the mean of one distribution is larger than the mean of another. Many scientific questions can be framed in statistical terms as a hypothesis test about the parameters of a probability distribution. Suppose that a clinical trial is designed to test the null hypothesis that two treatments have the same mean effect. Formally, the treatments have true mean effects, denoted respectively by parameters ϕ_1 and ϕ_2, and the interest of this chapter lies in testing whether the difference of the parameters of the two treatments $\phi = \phi_1 - \phi_2$ is 0. The null hypothesis then is $Ho{:}\phi = 0$. Many alternative hypotheses are possible. The two most common in this setting are the two-sided alternative $H_1{:}\phi \neq 0$ and the one-sided alternative $H_1{:}\phi > 0$ (or $H_1{:}\phi < 0$). A test statistic must be chosen to test the null hypothesis against the alternative. Notably, different alternative hypotheses often lead to different rejection regions for the same test statistic. If the data from both treatments are assumed to follow a normal distribution and both have the same albeit unknown variance, then the test statistic usually used is a student's *t* statistic, which has the form

$$t = (\bar{x}_1 - \bar{x}_2)/\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Here, $\bar{x}_1$ and s_1, and $\bar{x}_2$ and s_2 are estimates of the mean and standard deviation based on samples of size n_1 and n_2 from the two respective populations. The *t* distribution depends on the expected value of $\bar{x}_1 - \bar{x}_2$, which is equal to $\phi_1 - \phi_2$, the variance of the random variable and sample sizes, n_1 and n_2.

Type I and Type II Errors

The decision as to whether to accept or reject the null hypothesis is

determined by the value of the test statistic, which, of course, is a random variable whose value is unknown before the data are analyzed. If the value of the statistic falls in a prespecified *rejection* or *critical* region *R,* the null hypothesis is rejected. Even if the null hypothesis is true, there is still a chance that the statistic will lie in *R,* in which case a type I error has occurred. That is, a true null hypothesis has been erroneously rejected. Alternatively, the test statistic may not fall in *R,* although H_0 is not true, in which case a type II error has occurred. That is, a false hypothesis has not been rejected. Both kinds of errors must be considered in choosing *R*. Usually the choice is made so as to ensure that the probability of a type I error, usually denoted by the Greek letter α, is small. In most clinical trials, *R* is chosen so that α is 0.05, although in some applications choosing a type I error rate of 0.10 may be reasonable. This value is called the *level* of the test. For the comparison of two means, the calculation of this probability is performed under the assumption that $\phi = 0$, a *point* or *simple* null hypothesis. If the alternative hypothesis is $H_1{:}\phi \neq 0$, and the sample sizes are large, then the hypothesis is rejected if the *t* statistic is either less than –0.168 or greater than 1.68. These are the *large sample critical values* of a *t* distribution with $n_1 + n_2 - 2$ degrees of freedom with α set equal to 0.05. This is the so-called two-tailed test. If the alternative hypothesis is $H_1{:}\phi > 0$, the null hypothesis is rejected if the *t* statistic is greater than 1.96.

If there are only two parameter values under study, that is, two simple hypotheses, which one should be labeled the null and which the alternative? The null hypothesis should be chosen to be the one whose false rejection would be considered to be the more serious error. This is because the testing paradigm that selects α, the level of the test, controls the probability of this error.

Power of a Test

The probability of a type II error, or rather one minus this probability, is called the *power* of the test. Since the alternative hypothesis contains many values of ϕ, it is called a *composite* hypothesis—power is a function of the possible values of ϕ in H_1. In the earlier example, power is an increasing function of ϕ and the sample size and a decreasing function of σ^2. During the design phase of a study, the sample size is chosen so that the power to reject a false null is relatively large, perhaps 0.8 or greater, for a value of ϕ that is sufficiently large to be clinically meaningful. The calculation is often performed under an assumed value of σ^2 that is based on data from other similar clinical trials. Suppose that the estimated value of ϕ turns out to be a number that is smaller than the clinically meaningful value chosen when the power calculation was performed and that the test statistic falls in the critical region. There is no difference in the statistical interpretation of the results of the trial. The hypothesis is still rejected at the designated level of

the test.

p Values

Even a casual reader of the psychiatric literature will encounter a study in which *p* values are reported. Unfortunately, in some articles, only *p* values appear, without estimates of the parameters themselves or the variance of the estimators. All too often both the author and the reader misinterpret the *p* values. Does a very small *p* value mean that one of the treatments is very much better than the other?

A good place to start is with an explanation of the meaning of a *p* value. Before the analysis of data commences, the value of the test statistic is, of course, random and in the example, the comparison of two treatments, it follows a *t* distribution. Suppose that the null hypothesis being tested is true, so the expected value of $\bar{x}_1 - \bar{x}_2 = 0$. In this case the distribution of the statistic is completely specified. It follows a *t* distribution with $n_1 + n_2 - 2$ degrees of freedom. So, for every possible value of the test statistic, the chance of obtaining a value more extreme can be computed or found in a table. If t^* is the value of the test statistic observed in the trial, then the corresponding probability of observing a more extreme value is defined to be the *p* value. Alternatively, the *p* value may be regarded as the smallest level (had it been chosen in advance) for which the hypothesis would be rejected with the observed data. But a small *p* value and a true null hypothesis are inconsistent, for were the trial repeated over and over again, obtaining the observed value of the test statistic or a value more extreme would be very rare. This compellingly argues against the null hypothesis and suggests that it should be rejected. Scientific tradition or perhaps convention usually defines rare as an occurrence whose frequency is less than or equal to 1 time in 20. Note well that the fact that the null is rejected does not mean that the magnitude of the difference has any clinical or scientific relevance. The sample size could be so large that even small trivial differences achieve statistical but not practical significance. Also, the inference that a large *p* value implies that the mean of the two treatments are equal, that is, proving the null cannot be made. A calculation of the power of the test would help to rule out the possibility that the sample size was not adequate to detect a meaningful difference. More helpful still would be a CI for the parameter whose upper limit falls below any values that would be scientifically or clinically meaningful. Finally, it is necessary to emphasize that in the frequentist tradition of statistical inference, the *p* value cannot be interpreted as the probability that the null hypothesis is true. The hypothesis is not a random event so it is either true or it is false. In a subsequent section the Bayesian approach will be introduced. In this framework, a method for forming the probability or perhaps degree of belief in the truth of a null hypothesis modified by the collected data will be

described.

Testing Multiple Hypotheses: The Family Wise Error Rate

In the preceding section testing a single hypothesis was considered. However, it is rare that there is only one outcome variable, or one dose of a drug, or one homogeneous population that is of interest. As the costs of even simple clinical trials increase, it is prudent to consider simultaneous investigation of a variety of issues, because gathering additional information may represent marginal additional expenditures. The result of such a strategy is many more hypotheses tests to answer the variety questions that may be asked of the data. The collection of statistical hypotheses, each derived from some comparison of interest that relates to a common aspect of the research question, is referred to as a *family of hypotheses*. For example, three doses of a drug and placebo may be simultaneously studied in a clinical trial. The doses are each to be compared to placebo. The three contrasts, high dose versus placebo, middle dose versus placebo, and low dose versus placebo are a family of hypotheses. A priori it is possible that all of the null hypotheses in the family are true. Testing each of these at a specified level α ensures that the probability of a type I error of each contrast is α. But, it is prudent to ask, what is the probability that one or more of the tests erroneously rejects the corresponding null hypothesis? The probability of erroneously rejecting one or more of the hypotheses in the family is called the *family wise error rate*. Clearly, in general, this probability is much higher than α. Under the null hypothesis, the probability of not making a type I error when testing a single hypothesis at a 0.05 level is 0.95. Suppose that a family is comprised of m individual hypotheses. If the test statistics are independent, the probability that none of the hypotheses are erroneously rejected is 0.95^m. Therefore, the probability of the complementary set, that one or more of the m hypotheses is erroneously rejected is $1 - 0.95^m$. For $m = 3$, it is 0.14 and for $m = 5$, it is 0.23. The family wise error rate (FWER) approaches 1 as m increases.

In many research settings, it is appropriate to require that the statistical procedures utilized should not result in a large FWER. For example, if multiple doses of a putative drug are studied and the drug is ineffective, it would be unreasonable to allow the chance to be high so that one of the doses is found to be statistically different from placebo if the null hypothesis is true. In this section, statistical tests that control the FWER multiple hypotheses tested are discussed. Several procedures for controlling the size of the FWER are presented and their advantages and disadvantages are discussed.

The most widely known approach to controlling the FWER is the Bonferroni method, which is based on a probabilistic inequality. In fact, the inequality is known as *Boole's inequality* and as often happens, the credit

for its utilization in simultaneous statistical inference may have been misplaced. If there are m hypotheses in the family to be tested and the type I error level of each test is set to α/m, then the probability of an FWER $\leq \alpha$. That is, each test is performed at a level that is much smaller, α/m, than α, which would be used if there were only one test. The Bonferroni procedure is conservative in that, while it controls the family wise error rate, it renders the chance of rejecting any one of the individual hypotheses considerably lower. As a consequence, the user of this approach pays a heavy price in terms of loss of power. For example, if the family has five hypotheses, then, in order to achieve an FWER of 0.05, the individual hypotheses must be tested at the $\alpha = 0.01$ level. If a t test is used for each of the individual hypotheses, then the value of the t statistic calculated after the experiment is complete must be larger than 2.457 in order to reject the corresponding null. If only a single hypothesis were being tested, then the t statistic would have to be larger than 1.697. This makes the task of rejecting any of the individual hypotheses considerably more difficult. On the other hand, the Bonferroni method can be used to control the FWER in any setting. It is easy to use and does not depend on the distribution of the test statistic. Several improvements to this method have been suggested. Sidak modified the Bonferroni approach providing only slight improvement in power. If the desired FWER is desired to be α, then instead of using α/K, the Sidak cut point is defined to be $\alpha^* = 1 - (1 - \alpha)^m$, where m is the number of hypotheses tested.

In some studies, for example, one comparing K treatments, the investigator is interested in a contrast of each treatment with every other treatment. If say $K = 3$, and the τ_i is the parameter characterizing treatment i, then the family consists of three pairwise comparisons $(\tau_i - \tau_j)$. The first step in analyzing pairwise comparisons is often an analysis of variance on the K treatments. One of the first tests to control the FWER, the least significant difference (LSD) procedure was developed by Sir Ronald Fisher. He argued that if the analysis of variance rejected the null that all three treatments have equal means (the omnibus null hypothesis), then to test the pairwise contrasts, t tests should be computed between each pair. The estimated variance used in the denominator of the t statistic is based on all the data, not just on the data for the two treatments being compared. If there are 30 subjects per group, then a two-tailed t statistic rejects the null for each contrast if it exceeds 2.042. If the analysis of variance is not significant, then neither the omnibus null hypothesis nor any other null hypothesis about differences between means can be rejected. This procedure is valid in that the FWER is controlled for $K = 3$ but it is not valid for any $K > 3$. Its use for such values does not result in tests whose FWER is protected at the α level.

This approach utilizes a staged approach to testing a family of hypotheses. The omnibus test, in the above case an ANOVA, is performed

first. Such tests have been called *gatekeeper* tests because they limit the overall FWER. Only after the omnibus test rejects the global null does further testing take place. These tests can be less demanding in terms of power than Bonferroni-like tests without affecting the FWER. The Newmann–Keuls procedure is another well-known example of such tests.

A variant of the LSD procedure is Tukey test for pairwise comparisons. It utilizes a t-like statistic called the *studentized range*. Most statistical software packages that compute ANOVAs include routines to compute this test. Tukey test compares each of the $K(K - 1)/2$ studentized t statistics with a single-tabled critical value. If the sample sizes of treatments are the same and each of the pairwise comparisons is to be tested, then many believe that Tukey method is the method of choice. It exactly achieves the desired FWER and has reasonable power. The theory for this test was developed under the assumption that the sample sizes for each treatment are equal. Various corrections have been proposed to handle the unequal sample size case. For example, the Tukey–Kramer test replaces the common sample size in the formula with the harmonic mean of the respective sample sizes, that is, $(1/n_i + 1/n_j)/2$. Though this modification is widely used, the FWER fails to attain its proscribed level in most cases.

Although a clinical trial may involve K treatments, not all pairwise comparisons may be of interest. For example, $K - 1$ new treatments may be in a trial together with one standard treatment. A new medication for an illness may only be worth further study if it is superior to the standard medication, frequently called the control. Here the comparisons of primary interest are each new treatment against the control. Thus, there are only $K - 1$ comparisons to consider. Dunnett derived a test statistic for this problem that is similar to a t statistic. However, in this test, the estimate of variance used in the formula is derived from the data from all of the K treatments. The critical value is based on K and the *degrees of freedom,* which is the total number of subjects on all medications minus K. For a given FWER α, the critical value may be obtained from published tables. Dunnett test achieves the exact FWER.

The procedures discussed above for multiple comparisons are all specified in advance, but the order in which testing is performed are irrelevant. *Adaptive* or *sequential procedures* modify the process of testing depending on the outcome of earlier tests. They can provide an increase in power while still controlling the FWER. The Holm procedure for testing a family of m hypotheses begins by testing each hypothesis individually using a statistic such as a student's t. Each test produces a p value without multiplicity adjustment, so the experiment yields $p_1, p_2, \ldots, p_m$. These values are then ordered from the smallest to the largest, $p(1) \leq p(2) \leq \ldots \leq p(m)$. The notation $p(i)$ refers to the ith smallest p value. Relabel the m hypotheses so that $H(i)$ is the hypothesis associated with $p(i)$. If $p(1) > \alpha/m$, $H(1)$ is not rejected, the process terminates and none of the m null

hypotheses are rejected. If $p(1) \leq \alpha/m$, $H(1)$ is rejected and the process proceeds to examine $p(2)$. If $p(2) \leq \alpha/(m-1)$, H(2) is rejected and the process proceeds to examine $p(3)$. Otherwise testing is stopped and $H(2)$, $H(3)$, ..., $H(K)$ are not rejected. The procedure continues in this manner until a hypothesis is not rejected or there are no further hypotheses to tests. The Holm procedure is more powerful than the Bonferroni test, which requires that each $p(i) \leq \alpha/m$ as opposed to $p(i) < \alpha/(m-i+1)$. The Holm procedure is called a *step down procedure* because the cut points are in descending order. Note that the p_i are put in ascending order, which sometimes results in confusion as to what the test actually is.

A dual of this step down testing procedure is the *step up procedure,* introduced by Hochberg. Here the quantities $p_1, p_2, ..., p_m$ are reordered in descending order $p[1] \geq p[2] \geq ... \geq p[m]$ where $p[i]$ is the ith largest among the $p_1, p_2, ..., p_K$. The hypotheses are then relabeled so that $H[i]$ is the hypotheses associated with $p[i]$. If $p[1] \leq \alpha$, $H[1]$ is rejected and all of the remaining hypotheses are rejected. If $p[1] > \alpha$, then $H[1]$ is not rejected and $p[2]$ is examined albeit against a modified threshold. The values against which $p[2]$ and subsequent. In brief the step up procedure accepts hypotheses in order and at the first nonacceptance rejects all remaining hypotheses.

A *closed testing procedure* may be used to control the FWER on a "closed" family, and the method has some remarkably desirable properties. The procedure sequentially tests members of the family of hypotheses, yet each test is performed at size α. The family of hypotheses is said to be closed under intersection if the family satisfies the condition that if *H* and *G* are any two hypotheses in the family, the hypotheses represented by *H and G* are also members of the family. The sequential nature of the closed testing procedure requires that prior to rejecting any hypotheses, *H,* one must reject all hypotheses that include *H*. For example, suppose the family contains three hypotheses, *G, H,* and *J*. Then closure under intersection requires that it must also contain *G and H, G and J, H and J,* as well as *G and H and J*. To test *H* at level α, one must first test *G* and *H* and *J,* then test *G* and *H* and test *H* and *J,* and finally test *H*. Each of these tests is carried out at level α, and if any test does not reject, then the process ends. The FWER of the closed testing procedure is α.

This procedure has the very desirable property that the first hypothesis is tested at 0.05 and the very undesirable property that if any test fails along the way, the process stops. It might very well be that one or more of the hypotheses that are not reached would have been found significant had an alternative procedure for protecting the FWER been used. However, once the choice is made, hopefully before the data are examined, there is no going back. Further examination of *p* values is of course legitimate and desirable, but they no longer represent probabilities of a type I error under

the null hypothesis.

Comparing a combination medication to each of its component medications is another example of a testing procedure where each test is performed at the α level and yet the FWER is α. Let ϕ_i denote the mean difference in efficacy between the combination and medication *i*. Federal regulations require that prior to licensing a combination, one must show that each component medication contributes to the claimed effects. One translation of this requirement is that ϕ_i is positive for every component medication. This leads to the null hypotheses that for at least one component $\phi_i \leq 0$. If the outcome variable follows a normal distribution and ϕ_i is the test statistic for $\phi_i \leq 0$, then the test statistic for the null hypotheses is min $(\phi_1, \phi_2, \ldots, \phi_K)$. Each ϕ_i is tested at the same α level. This procedure is referred to as the *min test*. It is a simple example of what is known as the union–intersection principle.

Each of these approaches to controlling the FWER has positive and negative properties. Especially if the test statistics are not independent, the Bonferroni method becomes increasingly conservative as the number of hypotheses increases. The sequential procedures are usually more powerful than the other methods described above. However, when the number of false hypotheses in the family is large, they can be problematic. Tests based on the union–intersection relationship also tend to be quite conservative, and they are dependent on the existence of a satisfactory omnibus statistic.

False Discovery Rate

The shortcomings of the FWER methods are never more apparent than when the number of statements in the family of hypotheses is very large. Imaging data present a prime example. The number of potential statements is enormous, reaching even into the tens of thousands. Consider, for example, the question of whether a given voxel has been activated by a particular stimulus, or whether in a microarray analysis particular genes are differentially expressed. The statistics related to a pair of adjacent voxels will undoubtedly be statistically dependent, and the compound null hypothesis may be partially or even completely false. Under such circumstances, the use of the methods used to control the FWER for simultaneous testing may be useless. Some have instead conceptualized the problem as one of separating the important effects in the image from those that are small or unimportant. The false discovery rate (FDR) is a relatively new approach to the problem of multiplicity. The FDR controls the expected *proportion* of false positives among the declared rejected hypotheses. Just as in FWER methods, each voxel is tested resulting in a *p* value. On the basis of the distribution of the observed *p* values, a threshold is determined. The hypothesis corresponding to those *p* values falling below the threshold are rejected. However, the interpretation of the *p*

values is quite different from the approach described above. If 100 genes with a maximum FDR of 0.20 are declared by the procedure to be differentially expressed, then the expected value of the maximum number of false positives is 20. There is no comparable interpretation that attaches to *p* values. There are four major factors that affect the characteristics of a study utilizing an FDR criteria: (1) the proportion of null hypotheses that are false, (2) the distribution of the variables being assessed, (3) the measurement variability, and (4) the sample size. The investigator has limited ability to manipulate any of these except for the size of the study. The FDR approach provides much weaker control of the type I errors in the test procedure than do the traditional methods described above. This is because the FDR criteria for false-positive outcomes (the expected *proportion* of false positives is less than α) are much weaker than the FWER criteria (the probability of one or more false rejections is less than α). Thus, FDR methods are said to have weak control of the type I error.

THE BAYESIAN APPROACH

The formal setting for incorporating prior beliefs into a statistical analysis in which the data modify such beliefs is the essence of the method presented below.

Bayes Theorem

One of the most important results that emanates from elementary probability theory is Bayes theorem, which states

$$P(\mathrm{A}|\mathrm{B}) = P(\mathrm{B}|\mathrm{A})\, P(\mathrm{A})/P(\mathrm{B})$$

Bayesian inference plays a major role in statistical thinking until the early part of the 20th century. Although the hypothesis testing ideas of Jersey Neyman, Karl Pearson, and Sir Ronald Fisher played a dominant role in the rest of the century, in the past 25 years the tide has begun to turn. The methodology and intuitive logic of the Bayesian approach, made more feasible by high-speed computers, is now pervasive in the literature of modern science and statistics.

Bayesian Inference

Bayes theorem is the basis on which Bayesian inference is built. Suppose that ϕ is a parameter and that $\phi = 0$ represents the event that the null hypothesis is true. Then according to Bayes theorem,

$$P(\phi|\mathrm{data}) = P(\mathrm{data}|\phi)\, P(\phi)/P(\mathrm{data}).$$

The parameter ϕ may be thought of as a representation of the scientific theory under investigation. The left-hand side of the equation is a

quantification of the probability, or perhaps degree of belief, of the truth of the proposition that ϕ represents the true state of nature, given the observed data. In this setting, the unknown parameters are viewed as random variables and as such they follow a probabilistic law. Remarkably, if the right-hand side can be evaluated, a numerical value for the probability can be determined for this quantity, which arguably is the main issue once data collection is complete. It reflects the current belief, given the data, about the veracity of any value of ϕ and it is called the *posterior* distribution. The first term on the right is just the likelihood of the observed data, evaluated at the parameter value ϕ. It is the next term that has made Bayesian inference controversial. What is the value of $P(\phi)$? This quantity is a function of all of the possible values of the unknown parameter ϕ, and it is called the *prior probability distribution,* or just the *prior* for short. A prior distribution that has the same functional form as the posterior distributions is called a *conjugate* prior. This is a particularly desirable situation, for it permits efficient sequential updating of the posterior distribution as more data are collected. The posterior distribution at one stage is used as the prior in the next.

Note well that the result of applying the Bayesian method is an entire posterior distribution and not a single estimate of the unknown parameter. The posterior summarizes all of the information, past and present, about the unknown parameter value. In the Bayesian framework, issues such as estimation and hypothesis testing are all based on the posterior distribution. For example, sensible estimators of a parameter are the mean or median of the posterior distribution. For hypothesis testing the posterior distribution can be used to calculate the probabilities that the null and the alternative hypotheses are true. One possible rule is to decide to accept the null as true if the ratio of the posterior probability of the null divided by the posterior probability of the alternative exceeds some constant and reject the null otherwise. The interpretation of a p value in this context is more natural than it is in the frequentist paradigm. The value of p is in fact the probability that the null hypothesis has been erroneously rejected. Contrast this with the frequentist interpretation that p is the chance of obtaining a data set that is more extreme.

The controversy about the use of Bayesian inference hinges on the problem of objectively choosing a prior. In many if not most cases, knowledge gleaned from previous experiments and perhaps scientific theory can guide the mathematical specification of $P(\phi)$. However, in situations in which little is known or there is controversy about prior experience or relevant theory, the choice of the prior is subjective. There are two schools of thought about choosing the prior. In one approach the prior reflects the individuals' real-world experience and insights. An alternative approach is mechanistic or objective. The prior is chosen to satisfy some property; for example, it is noninformative or it is the

conjugate distribution.

What price is paid for introducing subjectivity into the scientific process? Supporters of the Bayesian approach argue that Bayes theorem mathematically demonstrates the inevitability of a subjective component to scientific inquiry. Scientists are the first to admit that insight, intuition, judgment, and beliefs are part of the process that drives their practice. Regulatory bodies are less convinced, but even in such scientifically driven agencies such as the U.S. Food and Drug Administration (FDA), Bayesian approaches frequently appear. A full discussion of how the a priori distribution can be specified and the dependence of the final a posteriori distribution upon this choice is somewhat technical. Needless to say there are a wide array of formal a priori distributions, including a noninformative prior, that is, one in which values of the parameter are equally likely, that can be used. A commonly used approach is to specify a range of priors and examine the degree to which conclusions are sensitive to the choice. The more robust the finding under different priors, the more believable they are.

RANDOMIZED CLINICAL TRIALS

There are many kinds of research projects in mental health that require the use of statistical methods to appraise the results. Here the discussion is focused on randomized clinical trials because they are important and ubiquitous, and many of the issues apply in other research contexts. The methods discussed are neither exhaustive in the range of topics covered nor are they complete in the detail that is required to appreciate the nuances fully. The goal of many mental health studies is to conclude "causality," that is, some observed effect is the result of an intervention. To begin this a model that leads to conditions under which causality may validly be concluded. This may be accomplished by randomization or by using matched subjects. In the situation where matching is practically difficult an alternate approach, the propensity score is introduced. The section concludes by introducing randomization tests that are valid regardless of the underlying distribution of the random variable of interest.

Causality

The principal reason for carrying out clinical research is to demonstrate the causal effect of the treatments being studied. Donald B. Rubin described an approach by which causality can be demonstrated in a clinical trial. Suppose that a putative antidepressant is given to a well-diagnosed patient suffering from the illness. If the patient does poorly, assuming the agent does no harm, no evidence is adduced about its effectiveness. Suppose the patient does well and there is considerable reduction in symptoms. How can it be concluded that the treatment was the causal agent? If the clock

could be turned back and the same patient with the same symptoms did not benefit from placebo treatment, it would be reasonable to conclude that administration of the test agent caused clinical improvement in this patient. Each patient has two hypothetical outcomes, O_T and O_P, which would result from being treated with T or P, respectively. If this hypothetical thought experiment were repeated a large number of times, then the frequentist would conclude causality if the average effect of an individual receiving the treatment is larger than the average effect produced by placebo in the same individual. In statistical terms, it would be written $E(O_T - O_P) > 0$, and it would be recognized that the concept is the same whether the hypothetical experiment is repeated over and over with the same patient each time or with different patients. But the generalizability of the conclusion is not the same in the two cases, and from a statistical perspective the interest is in population causality rather than individual causality. Of course the clock cannot be turned back, so what other possibilities are there? If each of the patients has an identical twin with the same genetic makeup and identical clinical symptom profiles, then one of each pair could be given the test treatment and the other given placebo. If the former improved and the latter did not, a causal relationship has been demonstrated. That is, it has been demonstrated that $E(O_T) - E(O_P) > 0$. Unfortunately, in the real world, each subject has only one outcome, O. If the treatment assignment is T, then the outcome is O_T, and if it is P, then the outcome is O_P. In this case, $E(O|T) - E(O|P) = E(O_T) - E(O_P) = E(O_T - O_P)$. What is necessary to be able to conclude causality is a mechanism that would enable the conclusion that $E(O_T - O_P) = E(O_T|T) - E(O_P|P) > 0$, where the expectation on the left-hand side indicates that each individual received both treatments and the expectation on the right-hand side indicates that each individual received only one of the treatments. The notation beyond the vertical bar indicates which treatment was assigned to the patient, and the expectation, the computation of the average over the population, is restricted to those receiving the indicated treatment. Suppose that the treatment assignment is a random variable determined by some mechanism that is independent of the outcome, O. For example, suppose a patient is assigned to the test treatment if a fair coin (one whose probability of each of the two outcomes is 0.5) lands heads and to placebo otherwise. Then the treatment assignment is independent of the outcome random variable, O. But then $E(O|T) - E(O|P) = E(O_T) - E(O_P) = E(O_T - O_P)$. So use of a random assignment mechanism enables conclusions of causality.

Note that the coin did not have to be a fair coin for the treatment assignment to be random. The same conclusion on causality would be reached were the probability of assignment to the treatment group equal to,

say, 0.6. This is called a *biased coin random allocation*. The only difference is that the result of such a randomization process would be that more individuals would tend to be assigned to the treatment group (about 60 percent) than to the control group (about 40 percent). This fact is important for application of propensity score methods, which are discussed below.

In fact, the idea extends to situations in which treatments are assigned to subjects based on factors z that are observed. For example, z might be severity at baseline or some other stratification factor. Even *response adaptive designs,* where the probability that the next subject is assigned to the test treatment depends on outcomes observed on patients studied earlier in the trial, can demonstrate causal effects so long as response to treatment and assignment to treatment are independent, conditional on z.

Matching and Randomization

Suppose that a clinical trial is to be conducted to establish a causal relationship between receipt of a treatment believed to be an active drug and clinical improvement, the effect produced. The treatment is to be allocated to one group of patients and a control treatment, a placebo, is to be allocated to another. As was just seen, the control group can be comprised of one member from each set of twins and the test treatment group the other twin and causal conclusions can be reached. Although such a design can only rarely be realized, matching two groups on all variables that are prognostic indicators of the outcome may be feasible. Subjects can be matched on the basis of characteristics such as age, weight, severity, prior episodes, genetic factors, family factors, and on and on. However, in psychiatric research, it will be a rare study indeed in which the investigator and more importantly the target audience of the research believe that the two groups are truly matched. However, there are some circumstances in which this will be the best approach, and such cases need to be appraised individually. In the idealized twin case, treatment assignment need not be random. But here, since there is no guarantee that paired subjects are truly identical, for causal inference the assignment mechanism within the pair must be random.

Randomization helps to avoid systematic bias, for example, evaluating one treatment in older or more severely ill patients and the other in younger or less severely ill patients. It has no effect on selection bias, which chooses individuals to participate in the study who are not representative of the target population. For example, in most drug comparison studies, the various eligibility requirements, both inclusion and exclusion, as well as the option candidates have of declining to participate in the trial bias the sample. However, this does not compromise the validity of the treatment contrasts so long as treatments are randomly assigned to participants.

The very mechanism of assigning treatments to subjects in the clinical

trial provides the probabilistic basis for computing, under the null hypothesis, valid probabilities of the various possible experimental outcomes. But this is exactly what is needed to assess the likelihood that the null hypothesis is true. In particular, it enables calculation of a rejection region corresponding to a chosen type I error, which is the basis of hypothesis testing. Further, because randomization guarantees that outcomes for different subjects are independent, a valid estimate of the experimental error can be obtained. This quantity is required for obtaining valid CI estimates of treatment effects and for parametric testing of a null hypothesis. Also, given a sufficiently large sample size, all of the possible known and unknown confounding effects that might possibly affect outcome are, with high probability, eliminated in the sense that they are averaged out. Finally, the objectivity and repeatability of randomized assignment to treatment must be pointed out.

There is no point to conducting a clinical trial whose results will not be creditable. The ability to describe exactly how candidate patients were placed into the various treatment groups makes transparent the process and gives credence to the conclusions reached.

Randomization Tests

Randomization tests make no assumptions as to the distribution of the random variables and so are considered to be distribution free. The results of the test depend only on the data actually collected. Suppose there are two groups of patients, whose group assignment is determined by a randomization procedure, with one group receiving the test treatment and the other receiving a placebo. After a suitable period of time, an outcome is obtained on each subject. The quantity of interest is the difference between the mean outcomes of the two groups. The first step in a randomization test is to compute the statistic, the difference between means, from the sample. The second step is to rerandomize and recalculate the value of the mean difference for the new treatment assignments. Under the null hypothesis, there is no difference between the groups; so reassigning some patients from one group to the other should not change any probabilistic properties of the distribution of treatment differences. Reassigning treatments to patients by rerandomization (or, for small numbers of subjects by exhausting all possible permutations of possible treatment assignments) and recalculating the mean differences induce a probability distribution of mean differences. The percentile in the randomization-induced probability distribution of the actual mean difference observed in the trial is informative about the plausibility of the null hypothesis. If the observed mean difference lies near the middle of the distribution, the null distribution seems likely to be true. If, on the other hand, the observed mean difference falls into one of the extremes of the randomization distribution, say less than 0.05 (α) of the values are larger, then the null

hypothesis seems unlikely to be true. The mean values of the two treatments are likely to be different. The randomization p value is the proportion of rerandomizations in which the value of the statistic is more extreme than the value obtained from the original treatment assignment. Randomization tests produce valid significance levels (probability of a type I error) regardless of the usually unknown underlying probability distribution of the outcome measures.

Rank randomization tests are similar to randomization tests. However, first the observations are converted to ranks. That is, the two groups are merged and put in numerical order. The largest is assigned rank one, the second largest rank two, and so on. Next, the two groups are separated into their original configuration, and a randomization test is computed on the ranks. This is a particularly useful approach when the scale used to measure patient response has some nonlinear properties and when outliers in the data might skew estimates of mean effects. The disadvantage is that some information is lost when the actual observed values are converted to ranks. Although rank randomization tests may be somewhat less powerful than randomization tests based on the original data, the loss in power may be a small price to pay to purge the potential contamination caused by exceptional large or small outcomes.

Propensity Scores

In efforts to understand the properties of a new treatment or intervention, it is often impossible or too expensive to conduct a randomized clinical trial. On the other hand, there may be data available on the new treatment as well as on standard practice that has been collected under naturalistic conditions. In earlier sections, considerable attention was paid to the merits of randomization and the benefits that accrue from its use. Randomization of persons to treatment or control groups allows statements of causality to be made, and systematic differences in characteristics between the groups are not likely to occur if the sample size is sufficiently large. Is it possible to nevertheless obtain valid estimates of effect size from data on patients who were not randomly assigned to treatment? Many standard statistical methods of analysis are not appropriate because they implicitly (at least for the purpose of causal inference) assume that the groups were formed by random assignment.

Suppose that background characteristics, some known, such as severity of illness and age, and others that may not be known, are markers for the assignment of patients to a new treatment or to a control or usual practice group. This might occur as part of usual care procedures or by unconscious hunches designed to individualize and optimize treatment outcome. If the two groups are compared without regard to the treatment assignment mechanism, the results can be biased particularly if there is an imbalance across groups in variables that influence outcome.

If there are only a few confounding variables and they are known, it may be possible to form stratum in which the treatment and control group look alike. For example, suppose there is a difference in the proportion of males receiving treatments in the two groups. Then a group containing only males and a group containing only females can be formed. The treatments can be compared within groups and if they are similar the results can be pooled.

If there are two known confounding variables, such as age (young, middle aged, and elderly) and gender, then age/gender groups can be formed yielding six strata in which to make comparisons. However, as the number of confounding variables grows, it becomes increasingly difficult to form the multiple strata that would be required. Sample sizes in each stratum would considerably diminish, resulting in a loss of power to detect differences when they truly exist.

Propensity score methods are used to control for a collection of confounding covariates. The propensity score is a probability of treatment group membership based on patient covariates. An estimate of the propensity score for each individual in the study is usually obtained by first fitting a logistic regression and then substituting his or her covariates into the resulting model. Naturally, persons who are actually in the treatment group will tend to have higher propensity scores than those in the control group. If the range of propensity score values in the two groups overlap, then it is possible to match individuals in one group with those in the other group who have like scores. Rather than one-to-one matching, usually the propensity scores for the combined groups are ordered, and five strata, that is, quintiles, are formed. Within a stratum, all individuals have approximately the same estimated likelihood of being assigned to the treatment group, and the multivariate distribution of the covariates used to estimate the propensity score in the two groups do not differ statistically. Across strata, the likelihood of being assigned to the treatment group differs. Within each stratum, some individuals were assigned to treatment and some to the control, so biased coin randomization has in essence been mimicked. If all subjects have a propensity score of about one-half, then treatment assignment appears to have been determined by a fair coin randomization process. Treatment effects are estimated within propensity score strata, and if the results are similar, they can be combined to obtain an estimate of the overall effect size. As Philip Lavori put it, within each observational study there lives a randomized study that can be found with the help of propensity scores.

More on Causality

When possible, causal relationships are inferred by manipulating the causal agent and determining its effect on the target in an experiment. However, there are many instances in which experiments are not possible and only observational or naturalistic data can be obtained. For example, a

researcher studying the development of posttraumatic stress disorder (PTSD) may hypothesize that different trauma types (combat exposure, motor vehicle accident, sexual assault) differentially affects the likelihood of the emergence of PTSD.

Causal Graph Models, which utilize directed acyclic graphs (DAGs) to represent causal relationships between variables or events, has been developed to facilitate causal inference in such situations. A DAG is simply a mathematical model described pictorially with arrows (edges) connecting vertices (variables or events) whose chance of occurrence are given by conditional probability distributions. Typically, arrows point in the same direction from *parent* or *ancestors* nodes (the cause) to *child* or *descendent* nodes (the consequence). *Paths* refer to a sequence of edges connecting one node to another possibly distant node; *causal paths* are those leading from a causal node to an outcome node. *Colliders* are nodes at which two edges converge. Necessarily, such graphs are acyclic in that there are no paths in the graph that start at one node that eventually loops back to the same node. The goal of this technique is not discovery of an *association,* statistical dependence between variables. The goal is determination of causality. The simultaneous presence of the symptoms of streptococcus and streptococcus bacteria is a dependence relationship (they are not independent) which is substantially different from the causal chain: the presence of symptoms is caused by the presence of the bacteria. A dependence relationship cannot differentiate disease causes symptoms and symptoms cause disease.

A *Bayesian network* is a DAG in which the nodes are events together with their probability distributions that depend only on the probabilities of prior events. In this sense, a probabilistic DAG formally encodes knowledge and beliefs about causal and potentially causal relationships among variables through visual representations. Edges represent conditional dependencies while nodes that are not connected are conditionally independent. A probabilistic DAGs has two primary purposes: to represent assumptions explicitly in order to reveal testable causal hypotheses and to enable explicit data driven hypotheses testing of these causal effects. This approach sharply contrasts with traditional hypothesis testing approaches that assume a null relationship which is rejected when data supply sufficient evidence of incompatibility. In Bayesian network theory, potential relationships between nodes are specified in a DAG and data is used to sustain or reject the relationships. Many algorithmic computer programs have been written to discover causal graphs by hypothesizing edges between all nodes and removing those that are incompatible with data.

Missing or Incomplete Data

In the conduct of research in mental health settings, having missing data is

inevitable. Such data sets are said to be "incomplete," and the problem of parameter estimation or hypothesis testing is regarded as an "incomplete-data" problem. Missing data problems occur both in clinical trials and in studies that are survey or field based, and the data can be missing for many reasons. In a clinical trial in which data are collected at multiple time points, subjects may not be available for all interviews. Some may drop out of the trial before it is completed because of adverse events or death, or they may move to another home, hospital, or jail or to the streets, or they may just perceive continuing in the trial to be a nuisance. In surveys, respondents may refuse to answer particular questions because of the sensitive nature of the inquiry, or they just may not know the answer.

There is no one approach to handling all missing data situations, but ignoring holes in the data when performing a statistical analysis can have an enormous effect on the results. The possibility of substantial structural biases that might have led to the data loses can lead to analyses that may even be so fatally flawed as to nullify the results. Conceptually, the complete data set consists of two parts, only one of which is observed. The missing part can only be inferred from the observed part. The steps to take depend on the mechanisms that led to the missing data. This is often referred to as the *missingness* of the data. Roderick Little and Donald Rubin have defined three kinds of missing data mechanisms: Missing completely at random (MCAR), missing at random (MAR), and nonignorable missing.

A very simple probabilistic model can be used to conceptualize data that are MCAR. Suppose that a random variable taking the value 0 or 1 with respective probability P and $1 - P$ was observed for each observation of the variable Y. Assume further that the value of the observation on Y is independent of the binomial random variable. Such a random variable is obtained by tossing a coin whose probability of landing heads is P. If the coin toss corresponding to a particular observation of Y lands heads, the data are lost. The remaining observations are a random subsample of the original sample and so the data remain a random sample.

An example of MAR can be obtained in a similar fashion. Suppose Z is gender, and a coin is tossed for males with probability P_1 of landing heads and a second coin is tossed for females with probability P_2 of landing heads. For each observed Y, one or another coin is tossed depending on the gender as captured by Z. If the coin lands head, the data are lost. The remaining data in each subsample is still a random sample.

As an example of nonignorably missing, suppose the observations on Y have been sorted into good and bad outcomes, and two coins are tossed with respective probabilities of heads P_1 and P_2 for the two types of outcomes for each observation. If either coin lands heads, then the corresponding observation is lost. The remaining data no longer arise from

the same underlying probability distribution that gave rise to the original sample. If $P_1 = P_2$, then the effect of sorting an outcome no longer matters, and the subsample is a random subsample of the original sample. In this case the data would be MAR.

There are several methods used to handle the situation when data are MCAR or MAR. Certainly the simplest strategy is to delete from the study those records with missing data on the primary variable of interest. If the sample size is sufficiently large, this may be a viable approach. However, if the amount of information lost diminishes the ability of the study to reach confident conclusions, then this method is not recommended. If the data are multivariate and only some of the variables in a record are missing, it would be rather wasteful to dispose of the entire record. Some analysts retain those parts that have data that are present and use them in marginal analyses not involving the variables with missing data. This approach too can be problematic. For example, suppose a clinical trial is designed to compare an active treatment with placebo. Those individuals who respond to the placebo remain in the trial, while the others drop out. For a self-limiting illness, the placebo responders and the treatment responders are not likely to differ toward the end of the trial. Ignoring the early dropouts and limiting the analysis to those who remain in the trial could lead to the erroneous conclusion that the drug loses its efficacy after a while. Filling in or imputing the missing data with a conservative estimate of its value is the current favored approach. However, this strategy too can lead to erroneous conclusions if there are too many records with missing data. Neither strategy is completely satisfactory nor is there a satisfactory rule of thumb as to how much missing data is too much for imputation. Individual variations among studies require critical appraisal in each particular case as to what is an appropriate analytical approach.

Imputation of missing data can be done in several ways. One simple imputation procedure is called *mean/median substitution,* in which the mean or median value of the observed data is substituted for the missing value. This approach tends to reduce the estimated variance, which thereby provides more comfort to the researcher than may be warranted. Several methods rely on a matching strategy. Records with missing data are matched with records of subjects with similar demographics whose outcomes are present. The data from the matched records are used to impute the missing value. In *hot deck imputation* cases the missing data are matched with similar cases with complete data on the basis of defined background variables, for example, age and gender. A critical issue is how to choose the variables that define "similar." Matching criteria variables are chosen because they are prognostic indicators of the missing data variables. The value substituted for the missing value is randomly selected from data present in the set of matched cases. *Multiple imputation* uses an imputation method such as hot deck to repeatedly generate estimates for

the missing data values. For each replication, the full data set, actual and imputed, is analyzed. The results from these multiple analyses are combined into one overall analysis. This approach gives an idea of the variability of the results among the various possible imputations that could have randomly been selected.

For data collected longitudinally, many researchers have used the last observed value carried forward (LOCF) approach to fill in the missing value. If values are available before and after the missing time point, many interpolate a value. These approaches too tend to reduce the estimated variance, which can lead to a greater chance of reporting an effect when there really is none. Hierarchical linear modeling provides a more satisfactory approach for analyzing longitudinal data. In these models, a time curve is estimated for each individual in the study from the available data without imputation, which works well as a substitute for interpolation. If many patients have missing observations in the later stages of a study, there may be a considerable degree of extrapolation. In this case, the method must be used with caution.

More sophisticated methods for imputing missing values are model based. For example, regression models have been used where the variable with missing data is the dependent variable. Parameters in the regression equation are estimated based on data from complete records. The predicted value from the model is used as an estimate of the missing data value. Another approach relies on maximizing a likelihood function using the available data values while taking into account the nature of the missingness. Here great care must be taken in specifying the underlying distribution function of the random variables.

Perhaps the most widely used method of imputation is the expected/maximization (EM) algorithm. It was originally conceptualized as a method for finding maximum likelihood estimators for incomplete data. The algorithm starts by choosing initial estimates of the values of the unknown parameters. The estimates are improved iteratively by first computing the conditional probability distribution of the missing data given the observed data and then the current estimate of the parameters. The expectation of the log likelihood over the previously calculated distribution of missing values is obtained. This is called the *E step* of the algorithm. A new or updated estimate of the unknown parameters may be obtained by maximizing this expectation. This step is called the *maximization step* (M step) after which the E step is repeated. An important feature of the EM algorithm is that it gives the complete probability distribution for the missing data and point estimates for the model parameters. Under some reasonably general conditions, the EM estimates obtained in the iteration procedure converge locally to the MLE.

The methods described above handle missing data under the MAR assumption. Pattern-mixture models have been proposed for nonignorable

missing data. Records are stratified into groups with different expected patterns of response and missingness. The group then becomes a predictor variable in a method of analysis (e.g., hierarchical linear models [HLM]) that allows subjects to have incomplete data across time. In this way the degree to which missingness influences outcomes can be assessed. Group membership can also be used in an interaction term in the model. This allows the assessment of the degree to which missingness patterns interact with key effect variables such as treatment group.

In sum, there is no one simple recipe for handling missing data. An understanding of the reason for the missingness is necessary, and a critical eye must be kept on the question of whether the MAR assumption is plausible. Whatever technique is used, some form of sensitivity analysis should be used to see how the imputation carried out in different ways affects the results. The choice of technique to use to handle the missingness must be made on a case-by-case approach.

SPECIFIC STATISTICAL METHODS

In this section a variety of commonly used analytic techniques are introduced. These include linear regression and logistic regression, contingency tables, the general linear model (GLM), and survival analysis. In mental health research, outcomes are typically observed over time, and a discussion of longitudinal methods to addresses such situations is given.

Regression Models

Regression models quantify a relationship between a dependent variable and one or more independent, explanatory variables. A statistical model is complete only if both its functional and probabilistic form and its parameters are determined. The procedure for estimating the unknown parameters is known as *model fitting*. Historically, the dependent variable, say Y, is assumed to follow a normal distribution, and the hypothesized model relates its mean or expected value and a function of the independent variables, say x_1, x_2, ..., x_k. The data consist of an observation on independent and dependent variables on each of N subjects. The most common model assumes that the normal mean is a linear function of the explanatory variables. That is, it is of the form

$$E(Y) = \alpha + \beta 1x1 + \beta 2x2 + \ldots + \beta k\, xk,$$

where the $k + 1$ Greek letters, α, β_1, β_2, ..., β_k are unknown parameters to be determined from the data. In terms of the actual observations rather than the expected value of Y, the form of the model is assumed to be

$$Y = \alpha + \beta 1x1 + \beta 2x2 + \ldots + \beta k\, xk + \varepsilon.$$

Here ε represents a random error term that accounts for the deviations among the observations from their mean value. For each subject, the difference between the observation Y and the model fit evaluated at his or her specific value of the independent variables is called the *residual error*. If Y follows a normal distribution with mean $\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$, and variance σ^2, then the error term is also normal with the same variance but with mean 0. In a regression analysis, estimates of the coefficients are obtained by maximizing the likelihood of the observations. Elementary calculus is used to find the value of the estimates, which are the solutions of a set of simultaneous equations, referred to as the *normal equations*. An alternative strategy for finding estimates of the parameters is to minimize the *sum of the squared errors* as a criterion for the goodness of the fit of the model to the data. The errors or the *residuals* are the difference between the observed values and the predicted values, the latter obtained by substituting the values of the xs into the regression equation. The *ordinary least squares* (OLS) approach makes no assumption as to the distribution of the error term. The method of OLS was developed at the turn of the 18th century by the mathematician Karl Friedrich Gauss working in Germany, and some historians say by others including Adrien Marie Legendre, Pierre-Simon Laplace, and (possibly) Robert Adrain, France and America, respectively. The least squares solution and the maximum likelihood solution are identical if the dependent variable is normally distributed.

Since specification of the set of explanatory variables is somewhat subjective, regression analyses include tests of whether each coefficient is equal to 0. If the null hypothesis is rejected, then the explanatory variable is included in the model. Stepwise procedures for choosing the variables to include in the model are also commonly used. Goodness of fit may be ascertained in a variety of ways, and model validation is a critical step in building a creditable model. At a bare minimum, the value of an R^2 statistic, which is the fraction of the total variability in the response that is accounted for by the model, should be obtained. But a high value of R^2 is not a guarantee that the model fits the data well, so it is wise to examine the residuals as well. Different types of plots of the residuals provide information on the adequacy of different aspects of the model. If the model fits well, a plot of the residuals will appear to behave randomly. This approximates the random errors that explain the difference between the independent variables and the response variable. If there is a nonrandom structure evident, then the model does not fit the data well.

There is, of course, no reason to limit the function on the right-hand side of the regression equation to linear forms. Many applications use rather complex nonlinear functions to represent the hypothesized relationships. In these cases, numerical methods are the only practical means for obtaining the estimates, and hypothesis tests are usually valid

for large samples only. In the past quarter century statisticians have generalized regression models in many directions, and applications incorporating these extensions appear frequently in the psychiatric literature.

Classification: Logistic Regression and Machine-Learning Techniques

Many research studies deal with outcome variables that are discrete. The simplest and most common case is dichotomous or binary random variables. A binary variable is often used to define the occurrence of an event, such as a treatment success, a treatment emergent side effect from a study medication, the disappearance of a symptom, or death. The event of interest might also be a classification problem, that is, predicting membership in one of two groups. The random variable Y takes the value 1 with probability P if the event occurs, and 0 with probability $1 - P$ otherwise. This is called a *Bernoulli random variable,* and an experiment in which many Bernoulli random variables are observed is called a *Bernoulli trial.* An investigator may want to know how other explanatory variables (covariates or risk factors) affect the frequency of the occurrence of the event. The idea is to use data, the training set, to find a method that will identify relevant covariates and use them to correctly predict an outcome in the test set. In recent years, there has been an explosion in the amount of available biological information (genetics and epigenetics, imaging, peripheral neuroendocrinology, clinical self-report, etc.) that can contribute to the classification and prediction accuracy. The limitations of traditional data analytic methods in such circumstances, in which there may be many more potential predictors than subjects has led to the development of "big data" methods. After introducing classical logistic regression analysis, this section will introduce machine-learning techniques, algorithmic methods that build models based on sample data that do not follow strict rule based or structured statistically modeling approaches.

Logistic Regression

Were the dependent variable continuous and approximately normally distributed, multiple linear regression would likely be used to explore its relationship to the covariates. Here, linear regression is not appropriate. An applicable model for a dichotomous variable is *logistic regression.* In this approach, the model assumes that the log of the odds ratio of the event occurring is linear in the covariates. These independent variables may be continuous variables, categorical variables, or both. The *odds* is the ratio of the probability that the event occurs to the probability that it does not occur. Odds are commonly used in gambling contexts. If the odds that a team will win are three to one, then one would expect the team to win

about three times more often than it losses. In this case, the probability of winning is 3/4. Suppose q is the probability of the occurrence of the event. Then odds = $q/(1 - q)$ and q = odds/(1 + odds).

The logistic model is

$$\log q/(1 - q) = \alpha + \beta 1x1 + \beta 2x2 + \ldots + \beta_k xk.$$

The left-hand side of the equation is called a *logit*. The covariates on the right-hand side of the equation can be continuous, dichotomous, nominal, or ordinal. Tests of the null hypothesis that each parameter is 0 may be made. Interpretations of the coefficients are made in a manner similar to ordinary linear regression. However, logistic regression calculates changes in the log odds of the dependent variable, not changes in the dependent variable itself, as in linear regression. For plotting and interpreting results from logistic regression, it is usually more convenient to express fitted values on the scale of probabilities. The inverse transformation of the model equations yields the logistic function for the probability of the event q,

$$q = \exp[\alpha + \beta 1x1 + \beta 2x2 + \ldots + \beta_k xk]/$$

$$(1 + \exp[\alpha + \beta 1x1 + \beta 2x2 + \ldots + \beta_k xk]).$$

The fit of a logistic regression can be assessed by looking at a classification table that displays whether the model has correctly classified each of the observed outcomes of Y. Goodness-of-fit tests are available as indicators of model appropriateness. One method begins by dividing the subjects into 10 groups, in increasing order of estimated probability of occurrence of the event. The first group corresponds to those subjects who have the lowest predicted probability; the next group corresponds to those with the next lowest predicted probability; and so on. A chi-square statistic is used to examine whether the observed counts in each group are close to the expected count under linearity. If they are, the chi-square statistic will be small and the corresponding p value will be large. Logistic regression can be extended beyond dichotomous response variables to ordered polychotomous (more than two) categories.

There are many measures of the accuracy of a classifier. The most widely used is the area under the receiver-operating curve (ROC). The ROC, graphically displays the true positive rate (sensitivity or recall) plotted against the false-positive rate (1 – specificity, or fallout) at various thresholds. The AUC from a sample data set is an estimate of the probability that the classifier assigns a higher score to a randomly chosen member of one class than to a randomly chosen member of the other class. The p-value of a chi-squared test that this probability differs from ½, is

often reported.

Supervised Machine Learning

In the classification setting, *machine-learning* methods are shorthand for a collection of computer-driven approaches for obtaining a classification rule learned from training data. Among the many machine-learning algorithms developed, two prominent and promising approaches, *CART* and its generalization *Random Forests* and *Support Vector Machines* are introduced below. These are popular and have been implemented in a number of software packages.

CART determines a set of logical if-then conditions, instead of a linear equation as in logistic regression, on predictor variables that corresponds to a logical or decision tree leading to terminal nodes. CART, the abbreviation for the classification and regression tree algorithm, is based on *recursive partitioning,* is often able to uncover complex interactions among predictors or features that may be difficult or impossible to uncover using traditional multivariate techniques. The composition of the training set membership in a specific terminal node is used to estimate the conditional probability of class membership and to predict or classify all subjects in that node. For example, if 80% of the subjects in the node are in one group, then all subjects in the node, are classified, some erroneously, as members of that group. The feature on which to split and the value on which to split (e.g., age greater than 75 years) at each node is chosen by the software to produce the greatest improvement in predictive accuracy at each step. This is usually driven by a "node impurity" measure such as the Gini Index, which provides an indication of the relative classification homogeneity of subjects that fall in the terminal nodes. The general approach to limiting over-fitting is first to stop generating new splits when further splitting produces only a small improvement in prediction accuracy. The candidate prediction tree is subsequently parsimoniously "pruned back" to further avoid the use of more complex models than are required. An added possibility is the use of linear regression at the terminal nodes.

Random Forests. CART provides the underpinning of the Random Forests classification method. In the RF approach, a multitude of *decision trees* are constructed during the training phase. The individual trees in the "forest" are grown by taking a random sample of subjects. If there are many more candidate features than subjects, these too are sampled at random as well. The final probability that an individual with a particular vector of covariates is predicted to be a member of a class is by "majority vote," that is, the average of the probabilities of the terminal nodes into which the predictor vector falls in each tree in the forest. This *bootstrap aggregating approach* is called *bagging* and its purpose is to reduce variance and to

prevent over fitting. Each individual tree may be highly sensitive to noise, but model averaging across many trees reduces variability, under the condition that individual trees are not highly correlated. By selecting random subsets of features in growing the forest, referred to as "feature bagging," correlations among trees are reduced.

Support Vector Machines. Support Vector Machine (SVM) algorithms build classifiers by identifying a linear decision surface that separates the classes with the largest distance, also called *gaps* or *margins,* between objects at the borderline. If there are two covariates, the space is two dimensional and the decision surface is a line. Although an infinite number of lines can separate two classes, only one, called the support vector, uniquely separates classes with the largest gap between borderline objects. Often it is not possible to linearly separate sample objects belonging to two classes. When a linear decision surface cannot be identified, an SVM maps the data into a higher dimensional feature space in which a separating linear decision surface can be identified. This technique is known as the "*kernel trick,*" which extends SVMs beyond linear to nonlinear classification. This framework is particularly useful in the not uncommon situation in which the dimension of the features is extremely large. SVMs construct a hyperplane, or set of hyperplanes in high dimensional space, which separates the groups and determines the classification of a new subject. Support Vector Regression is an analogous procedure when the variable of interest is a real number instead of a class indicator.

Over-Fitting. When there are hundreds or thousands or even millions of *features* under consideration, the potential for over-fitting is profound, particularly when the number of subjects is relatively small. With so many predictors available to use in fitting the model, the algorithms of most classification programs can achieve near perfect results on the training set, but more often than not fail abysmally on the test set. Many techniques have been developed to avoid or at least detect this over-fitting difficulty while at the same time permitting utilization of the full range of potential predictors. These include evaluating the model's performance on a data that is independent of the training set, for example, cross-validation, as well as methods that utilize fit criteria that penalize use of increasing numbers of parameters and more generally, model complexity, for example, regularization and pruning. In a Bayesian framework, regularization is equivalent to imposing specific forms of priors on the parameters. Randomization tests are also used; the class membership in the training sets is randomly scrambled and the classification algorithm applied. If good sensitivity and specificity are obtained for fictitious class labels, clearly, the algorithm cannot be trusted.

Contingency Tables and Log Linear Models

Many research studies involve discrete random variables. Frequency distributions of two discrete variables tabulated simultaneously are called *contingency tables*. The levels of one variable are listed in the rows of a table, the levels of the other variables are listed in the columns, and the joint frequency is entered in each cell. The sums of the cell frequencies across both the r rows and c columns are placed in the margins. The lower right-hand corner of the table displays the sum of the row marginal frequencies, which is equal to the sum of the column marginal frequencies, which is equal to the sample size.

Reports of many research studies begin with a description of the sample population, including the distribution of patient characteristics, such as treatment and baseline severity of illness. If there are two treatment groups being compared, then it is important to assess whether the groups were comparable at baseline. Define a treatment group indicator variable for the columns and a severity level variable for the rows. The question then is: Are the two variables independent or is there an association? Effects in a contingency table are relationships between the row and column variables in which levels of the row variable are differentially distributed over levels of the column variables. If the two variables are not independent, the two treatment groups have different distributions of baseline severity level, and the row variable will need to be controlled in any subsequent treatment comparisons in the analyses. Failure to reject the null means that there is not enough evidence to conclude that the variables are dependent; observed differences in cell frequencies are likely due to chance. (Nevertheless, even if the test failed to reject the independence hypothesis, it may still be necessary to adjust for imbalances in baseline variables. Important prognostic variables may be sufficiently imbalanced to alter inference and cause statistical errors.)

Hypothesis tests of association are often based on a chi-square statistic. The test statistic examines the difference between the observed counts in the table and those counts that would be expected if the variables were independent. The statistic sums the ratio of the squared difference between the actual count and the expected count to the expected count under the assumption of independence. If the differences are small, the data support the hypothesis that the variables are independent. In the 2×2 table, the difference of proportions, the relative risk, and the odds ratio are measures of the strength of association.

Contingency tables can involve more than two variables. One of the most common situations is the $2 \times 2 \times K$ table. For example, one variable may be treatment assignment to drug or placebo, the second may be a variable that measures success or failure, and the third may represent a control variable such as research sites, gender, or initial severity. There are

three test statistics commonly used to analyze such tables. The Cochran-Mantel-Haenszel (CMH) statistic assumes a common odds ratio and tests the null hypothesis that two of the variables are conditionally independent, while controlling for the possible confounding factors of the rest of the variables. The odds ratio is a measure of the increase in odds of success for those given drug relative to those given placebo. If the odds ratio is 1, the drug has no effect. The CMH tests whether the response is conditionally independent of the explanatory variable (treatment assignment) while adjusting for the control variable (site). If the directions in some strata are opposite to the direction of those in other strata, the CMH statistics has low power for detecting an association. The CMH procedure first estimates a common odds ratio across the K strata and tests whether it is equal to 1. The Breslow-Day statistic tests the null hypothesis of homogeneous odds ratios. That is, it tests whether the odds ratio between the assignment and outcome variables are the same for the K levels of the third control variable.

Higher level multidimensional contingency tables and more complex hypotheses are often analyzed using an approach known as *log linear models*. These models are used when two or more variables are measures of response rather than explanatory or control variables. The log linear model is one of the specialized cases of the generalized linear model, discussed below, for Poisson-distributed data.

The General Linear Model

The restriction to a normally distributed dependent random variable in regression theory was removed with the introduction of the GLM by John Nelder and Robert Wedderburn. This model assumes that the dependent random variable has a distribution in what is called a *parametric exponential family*. This family includes not only normally distributed and other continuous random variables, but also many discrete random variables such as the binomial and the Poisson. The latter distributions are used for modeling count data. The classical regression approach that models a normal mean as a linear function of the explanatory variables is generalized in GLM to a regression between a function of the mean and the explanatory variables. The function is called the *link function*. Although there are many possible link functions, associated with each particular parametric distribution is a natural or *canonical link function,* which enables efficient estimation of the coefficients. Here too, estimation is accomplished by maximizing the likelihood of the observations under the assumed parametric exponential family distribution via the link function. For example in the case of a Bernoulli response variable with mean p, the canonical link function is $\log[p/(1 - p)]$, the log of the odds. Here, GLM assumes that the error term follows a logistic distribution, which is a member of the exponential family of distributions. But this is exactly what is done in logistic regression, which is therefore a special case of GLM. As

in ordinary linear regression, tests of the hypotheses that coefficients of the explanatory variables do not differ from 0 are used to determine which ones have an effect on the mean. Similarly, if count data are distributed as Poisson variables and the link function is the log of the counts, then the GLM analysis is equivalent to a log linear model analysis of a contingency table.

Longitudinal Models

In the GLM and regression context, a single dependent random variable for each subject is considered. Especially in studies of mental and other chronic illnesses, subjects are usually observed at multiple time points or occasions. A typical hypothesis in such situations is that the means at all T time points are equal. This corresponds to the statement that there is no change over time. Conceptually, data in this situation may be described as a rectangular array, comprised of N rows and T columns. The rows correspond to the N subjects, and the entries in the columns are the subjects' response at the T time points. In a pre- and postmedication trial with N subjects, there are $T = 2$ columns. In the simplest case, the question is whether the mean effect of the pre- and postmedication responses is the same. A paired t, based on the difference between the pre- and postresponses, is the usual statistic for this comparison.

Until fairly recently, the most common approach to analyzing longitudinal data was based on a *repeated measures* model. The statistical assumption for this model requires that the T random responses of each subject follow a multivariate normal distribution in which the covariance structure is completely symmetric. That is, the variances at each time point are equal and the covariance between any two responses, irrespective of their time difference, are equal. Clearly these are rather demanding assumptions. Also, missing observation presented a significant problem, and in many trials all data on some subjects have been omitted. More recently, methods such as the EM algorithm enabled incomplete repeated measures data sets to be analyzed. However, hierarchical multivariate linear models, a generalization of repeated measures model have become more popular.

In a longitudinal study, despite the best of intensions, the occasions when subjects are observed as well as the number of time points that subjects are observed may vary widely from patient to patient. HLM allow such deviations, which overcomes the missingness problem when the data are MAR. While still retaining the multivariate normal distribution assumption for each subject, these models do not restrict the structure of the covariance matrix. The name hierarchical refers to the view that conceptually the analysis proceeds on different levels: Time is nested within subjects, subjects may be nested within sites, sites may be nested within countries, and so on. The model posits that the random normal

observation is a sum of three terms: (1) a sum of the form, $\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$, (2) a sum of random effects, and (3) an independent error term. The first sum is referred to as the *fixed effects* term. Although the random effects comprising the second sum are not observed, they enter the model by contributing to the variances and covariances of the observations. Since subjects are independent, the form of the likelihood is a product, which leads to a set of extended normal equations for estimating the parameters. Parameters may also be estimated using a least squares approach, although within this model a generalized least squares criterion that weights the residuals in terms of variances and covariances is used.

In a still more recent development, the multivariate normal assumption has been relaxed through the introduction of a *semiparametric regression* method, which incorporates elements of the GLM. The method of generalized estimating equations (GEE), introduced by King-Yee Liang and Scott Zeger, focuses on the means at each time point through link functions. In this way, each marginal mean response at each observed occasion is modeled. In many psychiatric studies, hypotheses are concerned with the behavior of marginal means. In lieu of modeling the entire joint distribution of the observations over time, the GEE method models only the variance and covariances of the observations over time. The variance at a given occasion is modeled as a function of the marginal mean at that occasion. The method allows missing data provided they are MCAR, that is, they are MAR conditional on the covariates. Although the likelihood function cannot be formed because the parametric form of the distribution is not specified, estimators can be obtained by solving an analog of the normal equations. The resulting estimators are consistent. That is, as the number of subjects increases, the estimators converge to their true values. Consistency holds even if the structure of the correlation matrix has been misspecified. Also, asymptotically the distribution of the estimators is normal, so that testing whether the coefficients of the independent variables differ from 0 can be accomplished.

Latent Class Growth Analysis and Growth Mixture Modeling

A constantly occurring research question is identification of distinct patterns of longitudinal course and outcome of a disease and the differential characteristics of individuals who follow them. For example, individual's response to trauma varies; many resilient individuals rapidly remit, some are slow remitters, some are nonremitting and some "late-onset" victims display low symptom levels initially that increase over time. Latent Class Growth Analysis (LCGA) and Growth Mixture Modeling (GMM) are commonly used statistical methods for identifying trajectories and the associated covariates associated with class membership. GMM extends traditional growth curve modeling, which allows the intercept and

slope parameters to vary across individuals. However, growth curve modeling assumes a common pattern of change across the population. In the case of GMM, the latent, that is, the underlying but not directly observable, variables that characterize the time course (the growth curve model, such as the slope and intercept) are utilized to identify an additional second-order categorical latent variable, class membership. Each class has its own growth model with unique covariance structure. The distinction between LCGA and GMM is that the former assumes that the variance and covariance among the growth factors within each class can be set to zero; that is the trajectories of class members are homogeneous. Because of complexities in model identification, analysts often begin with LCGA and constraints on the variance are removed as the search for a best fitting model proceeds. A typical data set is comprised of repeated measures across multiple time points and if data are missing at random, maximum likelihood estimators are valid. A trajectory that is assumed to have a linear form only requires observations at three time points. A trajectory with higher order quadratic terms requires four or more time points. Maximum likelihood parameters estimates are obtained iteratively through the use of the EM algorithm. As such, sufficient, high-quality data is required or the algorithm may not converge and other modeling difficulties can result. Modelers must take care to use appropriate model selection methods to guard against over-fitting, which typically results in the identification of more classes than are truly present "in nature." The inclusion of appropriate covariates in the model can aid in correct model identification/specification and result in increased accuracy of classification of individuals. The most commonly utilized software package for LCGA and GMM is MPlus. However, LCGA can be executed in SAS procedure Proc TRAJ and in R. For step-by-step instructions on building and LCGA or GMM in MPlus and a further nontechnical explanation see Muthén, Muthén and Asparouhov (2016).

Survival Analysis

Many psychiatric issues are related to events in the time domain, for example, time to onset of effect of a treatment, time in hospital, and time to remission of symptoms of depression. These are generically called *time-to-event* or *survival* variables. The first survival variable that was studied was time to death. *Life table methods* were originally devised for actuarial and insurance uses to describe expected survival time. For a particular group of persons (e.g., males) and selected age intervals (e.g., decades), life tables usually exhibit the number of persons entering the age interval alive, the proportion of these who die in the interval, and the proportion for whom data became unavailable during the interval. The proportion who die in the interval is called the *hazard rate*. Other useful statistics that can be derived from such tables include the *conditional survival rate* (among those who

enter the interval, the proportion surviving beyond the interval) and the *survival function* (the cumulative proportion of individuals surviving to the end of the interval). The *survival distribution* is defined to be 1 minus the cumulative distribution function, $1 - F(x)$. The time at which the survival function is equal to 0.5, the *median survival time,* is the statistic most often used to characterize the distribution. Other percentiles are of course of interest as well. Although life tables provide a great deal of information, they are of little use for predictive purposes or for assessing the impact of other variables on survival.

Sophisticated statistical methods for estimating survival time distributions were developed primarily to model data from clinical trials of cancer treatments, and even more sophisticated methods have emerged in association with the study of acquired immunodeficiency syndrome (AIDS). Survival time methods are widely used in most of the medical and biological sciences as well as in engineering (where the method is called failure time analysis) and in the social and economic sciences. The methods for estimating survival time frequency distributions from sample data have been developed to take into account the possibility that persons may still be alive at the end of the study period. Such individuals in essence have missing data as to the time of their death, but it would be folly to exclude them from the analysis. Observations on these individuals as well as the times of those who dropped out of the study before the signal event occurred, for example, a subject has moved and contact is lost, are called *censored observations*. If a new treatment is designed to study recurrence of depression and the study lasts 2 years, a censored observation at study completion or at loss to follow-up indicates that the treatment has been successful at least for that period of time. As in the missing data situation, analysis of these latter types of censored observations is facilitated if the data are MCAR, and, are therefore, noninformative as to the distribution of the time to the event being studied. Nonignorably missingness such as systematic censoring, if erroneously assumed to be MCAR, may greatly bias the estimation of the survival time distribution.

The Kaplan–Meier survival function estimation technique is nonparametric because it makes no assumption as to the underlying shape of the distribution. Based on the observed survival times and censoring times, it defines time intervals that contain exactly one observed case (except if there are ties). This is in sharp contrast with life table methods that count the number of deaths in fixed time intervals. The probability of surviving beyond time t is the product of the probability of being alive in each interval up to time t. If a case is censored at some time point, it is used in the calculation of the estimated probabilities up to the time of censoring and not used in subsequent calculations. Survival distributions generated from two samples, such as two treatment groups, can be compared statistically. The most commonly used tests are nonparametric, including

Gehan test, the generalized Wilcoxon test, and the log-rank test.

Parametric survival methods assume an underlying shape of the survival distribution, for example, exponential, Weibull, or Gompertz. Techniques such as MLE or least squares to fit the parameters of the model from the observed data are used to estimate the value of unknown parameters.

Many research projects entail an investigation of the influence that covariates have on survival time. Cox proportional hazard models were introduced to enable regression type analyses to be conducted on time-to-event data. No assumptions are made concerning the underlying form of the survival distribution, so the method is called *semiparametric*. The model assumes that the hazard rate, the probability of dying at time *t* given one is still alive up to that time, is equal to a fixed population or underlying hazard rate multiplied by the exponential of a linear function of the covariates. The log of the exponential multiplier is a linear sum of the covariates. The multiplicative form of the model implies that the hazards are proportional. This is called the *proportionality assumption,* which explains the name of the method.

Proportional hazard models are frequently used to test the null hypothesis of equality of the survival distributions of two treatments. If there is reason to believe that there is a common underlying hazard function for the two groups, a treatment group assignment term is used as a covariate and the null hypothesis that the corresponding parameter is 0 is tested. If there is reason to believe that each group has its own underlying hazard, a stratified analysis may be used to test the null hypothesis that the same regression model applies to the two groups.

Cure models have been proposed to account for situations in which the event being studied cannot occur for some part of the population. Cure models were introduced in cancer studies in which time to death was the principal outcome variable, but some fraction of the population was cured. In essence there are two populations: those who when given the test treatment will be cured and those who will die from the disease. In a genetics context these are the susceptible and nonsusceptible groups. In psychiatry, the study of onset of treatment effect parallels this situation. In a study of the effectiveness of an antidepressant, an important parameter is the time for the treatment to take effect. But some in the population may not respond to the treatment and will never have onset. This group is analogous to the cured group in the disease model. The quantities of interest in this model are the proportion cured (proportion of nonresponders), and among the group not cured (responders), the conditional probability of surviving (having onset). The estimation of these quantities needs to be made with care. A subject may survive beyond the end of the study either because he or she is cured or because he or she has a true survival time that exceeds the time that observations cease. The

proportion not cured may be estimated nonparametrically by the proportion still surviving at the end of the trial, as estimated in the Kaplan–Meier approach. The Kaplan–Meier curve can also be used to solve for the conditional survival distribution of the population not cured. If two or more groups are being compared, parametric and nonparametric models can be used to test for the equality of the cure fraction and for the equality of the conditional survival distribution. The advantage of parametric cure models is that they can include covariates that affect either the proportion not cured, or the conditional survival distribution, or both. Thus one can learn about treatment effects as well as the characteristics of a group that predict cure as well as those that predict shorter or longer survival time.

STUDY DESIGN

Usually, in the pursuit of a research question there are many ways to design a study to estimate parameters and to test hypotheses. In all fields of inquiry, but in medical research in particular, an experiment must be designed so as to produce the most accurate possible estimates of treatment effect within the limitations of sample size and personal safety. To use a design that is less than optimal is certainly inefficient but also may very well be unethical. The study of the design of experiments is a field of inquiry in statistics that searches for designs that are as efficient as possible. This section examines only one issue, the design of a crossover trial, as an illustration of the kind of considerations that enter the search for optimality.

Crossover Designs

There has been much interest and indeed controversy about the use of crossover designs. In a crossover trial, the comparison of t study treatments is accomplished by observing N subjects on several occasions, usually called periods. A subject may receive a different treatment in each period, or one or more treatments may be repeated, so a subject acts as his or her own control. The best-known crossover design is a two-treatment, two-period, two-sequence trial in which each subject receives both treatments, some in the order AB and the rest in the order BA. Experimental designs in which subjects are crossed over to other treatments have long been recognized as having considerable potential for achieving greater precision than can be obtained from parallel groups (also called completely randomized) designs. This is because between-subject variation is eliminated from the error term. Another advantage is that a crossover design requires fewer subjects to obtain the same number of observations. In a two-period design, each subject contributes two observations. For some types of studies, it is expensive to recruit participants, and once one has been trained in the details of the study procedure, there may be high

motivation to continue to obtain observations on the response to other treatments. Although they have not been used frequently in psychopharmacology, their use may well increase as the value of crossover designs is more widely appreciated by psychopharmacologists.

The negative aspect of crossover designs is the possibility that carryover or residual effects, that is, the effects of the previous treatment, whether pharmacological or psychological, may alter subsequent responses. The standard analysis of a two-period, two-treatment, two-sequence design (AB, BA) is rife with problems. An FDA advisory committee compared this simplest of crossover designs, in terms of precision and cost, to the completely randomized one-period design. Byron Brown, a member of the committee, concluded "the crossover experiment can yield great savings in cost if the assumption of no carryover effect is valid, but the design should not be used if this assumption is in doubt." In many cases, the sample size needed to test whether there is carryover present exceeds the sample size needed for a similarly powered parallel groups trial. For many years, the crossover design (AB, BA) was in disfavor, particularly by researchers in the pharmaceutical industry, as a result of Brown analysis and the presumed concurrence of FDA statisticians. However, enlarging the number of sequences to four so that the two-period design is (AB, BA, AA, BB) avoids the carryover problems of the (AB, BA) design, albeit with diminished efficiency to separate treatments. However, expanding the number of periods results in designs without the inefficiency of the two-period designs. These results have led to considerable renewed interest in finding ways to capitalize on the known advantages of crossover designs with three or more periods without falling prey to the problem of carryover. But there are many sequences possible, even for a three-period, two-treatment design there are eight possible sequences: AAA, AAB, ABB, BBB, BBA, BAA, ABA, BAB. If there are to be N participants in the trial, how should they be allocated to each sequence to achieve the maximum possible efficiency in estimating the parameters that characterize the treatment effects?

In a p period crossover design, there are p responses from each subject, which naturally are correlated. If the $p \times p$ covariance matrix of the responses is known, then the optimal design can be exactly determined. But this is hardly ever the case, and ignorance of the covariance matrix greatly complicates the problem of specifying an optimal design. There are two practical approaches to solving the design dilemma. (1) Based on information from other similar studies, the researcher can assume that the correlation structure is of some specified form, from which an optimal design may be determined. Almost all medical research involving crossover designs has been conducted under some assumption as to the form of the covariance matrix. In this case, the design is called *fixed* because the entire design is specified before the clinical trial begins. (2) A clinical trial with p

periods can commence, and as data accumulate the information on the first few of the *N* subjects to be studied can be used to estimate the covariance matrix. Treating the estimate as if it were the true covariance matrix, the optimal sequences to which the next few subjects will be allocated are determined. Their data are added to the first set of data, and a new estimate of the covariance matrix is obtained. The process continues until all *N* subjects are studied. For reasonable sample sizes, simulated experiments suggest that the resulting trials are nearly as efficient as the optimal design that would have been used had the covariance matrix been known. Designs in which no specific form of the covariance matrix is assumed are called *response-adaptive* because the subsequent course of the design is specified only after prior responses are obtained.

Notably, the controversy over the use of crossover designs has not subsided. Optimality, results are based not only on the assumed covariance matrix but also on the assumed response model, that is, the set of assumptions as to how carryover and response are related. Different models may yield entirely different results. For example, some authors maintain that the carryover effect that occurs if two consecutive treatments are identical, for example, AA, is not the same as the carryover effect that arises from two consecutive treatments that are different, for example, AB. Recently, statisticians have begun to consider response models that reflect this and other views, and designs that are optimal under some of these models have been found.

EVALUATING TREATMENTS UNDER CONTROLLED CONDITIONS, IN THE REAL WORLD AND UNDER FINANCIAL CONSTRAINTS

Efficacy studies, designed to determine whether a treatment has an effect, are usually carried out under relatively pristine laboratory-like conditions with rigid adherence to detailed protocols. In contrast *effectiveness studies* attempt to evaluate how and if a treatment works under real-world conditions. Invariably, treatment choices must involve their costs. Statistical *cost-effectiveness analysis* (CEA) provides information to help support the decision maker.

Efficacy Studies

Randomized double-blind, controlled clinical trials are the gold standard for efficacy studies, enabling, theoretically at least, an unbiased assessment of the relative merits of the study treatments and causal inference. In an efficacy study, extraneous variables that might have an impact on outcome and therefore confound the results and their interpretation are controlled to the greatest extent possible. But when experimental trials are conducted under the best of laboratory-like conditions, it must be kept in mind that the ability to generalize the findings is limited. Such trials provide little

evidence as to how the study treatments will work under usual care conditions in the real world. Efficacy studies are usually conducted in a limited number of well-circumscribed patient populations, and many of the patients most likely to receive the treatment if the FDA approves it are excluded. Trial eligibility criteria may exclude, for example, patients with comorbidities, those whose symptoms are too severe or not severe enough, those suspected or known to be "noncompliant," or those who have a history of violence. The trials may involve care conditions that are quite specialized, delivered by highly trained staff who provide individual attention far in excess of the norm. A positive finding does prove that the experimental treatment works, but the issue of efficacy beyond the well-defined populations on whom it has been tested in the well-defined environments in which the studies were conducted remains open.

Effectiveness Studies

Effectiveness studies examine how well a treatment works under real-world conditions. Many studies designed as randomized clinical trials should actually be considered to be effectiveness studies, and the environmental deviations from controlled conditions should be taken into account. For example, clinical trials of psychotropic agents are rarely conducted under pristine experimental conditions. They may take place for a short time in hospitals, but follow-up is usually in an outpatient setting. They may be based in outpatient clinics or in multiple practitioners' offices so subjects come and go from their homes. They may have multiple exposures to various unreported risk factors that impact the conditions for which they are being treated. Such conditions cannot be completely controlled. Over the course of the study, treatment groups may be differentially affected in ways that can influence outcome to such a degree that the effect of the treatment they receive is overwhelmed.

Although randomization is preferred, this may not always be possible; many studies are conducted in which patients have been assigned to treatments in the naturalistic course of their care. Although methods for mimicking randomization (e.g., propensity scores) can be used, findings from these studies are often affected by environmental factors unrelated to the study treatment. Sometimes study treatment and locus of care are confounded because of the nature of the intervention. This is almost always true in trials that test nonpharmaceutical treatments, such as new clinical programs, screening models, or complex approaches such as supported housing. In such trials, blinding as to treatment assignment is essentially impossible so that unbiased assessment of outcomes may be questionable.

Such situations create studies in which analysis of the data and more importantly their interpretation are substantially more complex than analyses of the traditional randomized, double-blind controlled clinical trial. Nevertheless, the value of these studies can be enormous. Besides

having the potential to provide information on efficacy, a well-designed effectiveness study can discover the way to maximize the yield on use of an efficacious agent under real-world conditions. These studies test acceptability both to consumers and to the health professionals who deliver the treatment. Outcomes of interest include measures of the treatment's acceptability to recipients and to caregivers, the treatment's impact on quality of life, and its costs. Compliance issues need to be examined with respect to events in the clients' world and to the biological or structural elements of the intervention itself. To understand the factors that affect compliance, data should be collected on basic demographic descriptors and on substance use, work and housing conditions, social relations, and environmental conditions, including those of the caregiver and the treatment setting. Such data items need to be collected at baseline and also over the course of the study.

From a statistical perspective, multiple analyses of multiple outcomes are required. In the comparison of treatment groups based on an intent-to-treat analysis, all subjects, including noncompliers, treatment crossovers, and dropouts remain members of the group to which they were originally assigned. This analysis provides a view of how well the "policy" of assigning the study treatments works in the real world, blemishes and all. An efficacy analyses can be conducted using the same data set by identifying matched persons from each group based on predicted treatment assignment, compliance, or crossover behavior patterns. Treatment outcomes are compared within matched groups. Propensity score methods can be used to accomplish the matching. A third set of analyses might examine the impact of environmental factors on treatment acceptability and compliance.

Cost-Effectiveness

The high cost of health care provides a compelling reason to examine the costs of interventions, not just their effectiveness. CEA is an economic technique for guiding choices among alternative treatments, policies, or program-level interventions so as to maximize effectiveness under a constrained budget. In a cost-effectiveness (CE) study alternative interventions are compared in terms of the costs required to provide a particular level of effect.

For CE purposes, a program P is characterized by an ordered pair of positive values (ε, γ), representing respectively its mean effectiveness and its mean cost. Programs are *mutually exclusive* if only one of them may be used at one time in one patient. However, a mixture occurs when two or more mutually exclusive treatments are given to a collection of subjects. For example, if 100 patients are to receive medication treatment A or B and no person can receive both treatments because they are incompatible, then of 80 percent receive drug A and 20 percent receive drug B, a mixture of programs has occurred. Programs are *independent* if any number of them

may be used at one time and neither the cost nor effectiveness of any program is affected by a decision to use or not use any other program(s) in the set. A program is *dominated* if there is another program with greater or equal effectiveness and less cost. A *weakly dominated* program is one that is dominated by a mixture of programs. Within a class of mutually exclusive programs, only those that are neither dominated nor weakly dominated (the admissible class) should be considered for use. The *cost-effectiveness ratio* (CER) of a program, the price per unit of effectiveness, is the ratio γ/ε. If two programs are ranked in order of increasing cost, the *incremental cost-effectiveness ratio* (ICER) of the program with higher cost compared to the program with lower cost, the additional price per additional unit of effectiveness, is the incremental cost divided by the incremental effectiveness.

Subject to conditions on independence and mutual exclusivity among specified programs and *given the values of their mean cost and mean effectiveness,* the *deterministic* resource allocation problem of CEA is the selection of a set of programs for funding whose effectiveness is the maximum that can be achieved without incurring a cost greater than a specified fixed budget, C^*. For independent programs resources are allocated sequentially in order of increasing CERs until the budget is exhausted. For mutually exclusive programs, *dominated* programs are eliminated and the remaining programs ranked in order of increasing effectiveness, ICERs are computed, weakly dominated programs are eliminated, ICERs are recalculated, and the process continues until all remaining programs are admissible. The remaining programs form the *frontier,* the piecewise line segments connecting adjacent admissible programs. Given a budget constraint, the program that should be funded is the one on the frontier with that budget; if the budget lies between two frontier programs budgets, then a mixture of these two is funded.

When analyzing cost-effectiveness from the societal perspective, the above fixed budget approach is replaced by a fixed price approach. Economists have shown that the greatest feasible health gain is achieved by implementing within each independent set of mutually exclusive programs that program whose ICER is largest among those programs that do not exceed a specified threshold value. This value, denoted by λ, is interpreted as the amount society is willing to pay (WTP) for an additional unit of health gain. Programs with ICERs above the threshold value are not implemented. The fixed budget and fixed price approaches prescribe the same allocation, except that the latter does not allow partial or mixture implementations of programs.

Statistical methods specifically for CEA began to emerge in the mid-1990s centered on methods for obtaining CIs for ICERs. The inherent difficulty of dealing with a random ratio whose denominator may assume the value 0 compounded the difficulties of this approach. Recently, the

value of net health benefit (NHB) for statistical CEA has become widely accepted. The *NHB* of program k is defined as $NHB_k(\lambda) = \varepsilon_k - \lambda_k/\lambda$, $k = 1,2, \ldots, K$. At each WTP value, λ, the optimal NHB rule for mutually exclusive program funds the program with the largest positive NHB. For resource allocation, both the standard ratio-based allocation rules of CEA and the new NHB-based rule, lead essentially to the same optimal solution.

The NHB rule requires neither rank ordering of programs nor elimination of inadmissible programs, which enables statistical methods that control errors and that emulate the optimal deterministic rules of CEA. For bivariate normally distributed cost and effectiveness variables and a specified WTP, λ, the statistical procedure is based on the method of *constrained multiple comparisons with the best* (CMCB) for determining the program with the largest NHB. This method both controls the pointwise error rate at each λ and provides a confidence set for the programs that are either best or comparable to the best. The statistical frontier, plotted in the λ-NHB plane, displays the program or possibly the set of programs with the largest NHB at each WTP value. Bayesian methods for performing a CEA have also been developed.

REFERENCES

Agresti A. *Categorical Data Analysis*. New York: Wiley; 1990.

Bluman AG. *Elementary Statistics: A Step by Step Approach*. New York: McGraw-Hill; 1995.

Brown BW Jr. The crossover experiment for clinical trials. *Biometrics*. 1980;36:69–79.

*Cornfield J. A statistician's apology. *JASA*. 1975;70:7.

Dudoit S, Shaffer JP, Boldrick J. Multiple hypothesis testing in microarray experiments. *Stat Sci*. 2003;18(1):71.

Fliess JL. *The Design and Analysis of Clinical Experiments*. New York: Wiley; 1986.

Gehan EH, Lemak N. *Statistics in Medical Research: Developments in Clinical Trials*. London: Plenum; 1994.

*Gelman A. *Bayesian Data Analysis*. London: Chapman & Hall; 1995.

Hedeker D, Gibbons R. A random-effects ordinal regression model for multilevel analysis. *Biometrics*. 1994;50:933–944.

Hochberg Y, Tamhane A. *Multiple Comparison Procedures*. New York: Wiley; 1987.

Holland PW. Statistics and causal inference. *JASA*. 1986;81:945.

Hsu JC. *Multiple Comparisons, Theory and Methods*. London: Chapman & Hall; 1996.

Jones B, Kenward M. *Design and Analysis of Cross-Over Trials*. 3rd ed. Monographs on Statistics and Applied Probability 138. London: Chapman & Hall/CRC, 2014.

*Kleinbaum DG, Klein M. *Survival Analysis: A Self Learning Text*. 3rd ed. New York: Springer-Verlag; 2013.

*Kotz S, Read CB, Banks DL. *Encyclopedia of Statistical Sciences*. New York: Wiley; 1997.

Kushner HB. Optimality and efficiency of two-treatment repeated measurements designs. *Biometrika*. 1997;84:455.

Laska EM, Klein DF, Lavori PW, Levine J, Robinson DS. Design issues for the clinical evaluation

of psychotropic drugs. In: Prien RF, Robinson DS, eds. *Clinical Evaluation of Psychotropic Drugs*. New York: Raven Press; 1994.

Laska EM, Meisner M, Siegel C, Wanderling J. Statistical determination of cost-effectiveness frontier based on net health benefits. *Health Econ*. 2002;11:249–264.

Lavori PW, Laska EM, Uhlenhuth EH. Statistical issues for the clinical evaluation of psychiatric drugs. In: Prien RF, Robinson DS, eds. *Clinical Evaluation of Psychotropic Drugs*. New York: Raven Press; 1994.

Liang KY, Zeger S. Longitudinal data analysis using general linear models. *Biometrika*. 1986;73:13.

Little R, Rubin D. *Statistical Analysis with Missing Data*. New York: Wiley; 2014.

McCullagh P, Nelder JA. *Generalized Linear Models*. London: Chapman & Hall; 1989.

Muthén BO, Muthén LK. Integrating person-centered and variable-centered analysis: Growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res*. 2000;24:882–891.

Muthén BO, Muthén LK, Asparouhov T. *Regression and Mediation Analysis Using Mplus*. Los Angeles: Muthén & Muthén; 2016.

Nichols T, Hayasaka S. Controlling the familywise error rate in functional neuroimaging: A comparative review. *Stat Meth Med Res*. 2003;12(5):419–446.

Pearl J. Causal inference in statistics: an overview. *Statistics Surveys*. 2009;3:96–146.

Piantadosi S. *Clinical Trials: A Methodologic Perspective*. New York: Wiley; 1997.

Rosenbaum PR. *Design of Observational Studies. New York: Springer; 2010*.

Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41.

*Rosenberger WF, Lachin JM. *Randomization in Clinical Trials: Theory and Practice*. New York: Wiley; 2002.

Rubin D. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley; 1987.

Zeger S, Liang KY, Albert P. Models for longitudinal data: A generalized estimating equation approach. *Biometrics*. 1988;44:1049–1060.

6 Theories of Personality and Psychopathology

▲ 6.1 Classical Psychoanalysis

WILLIAM W. MEISSNER, S.J., M.D.

Psychoanalysis has existed since before the turn of the century, and in that span of years, has established itself as one of the fundamental disciplines within psychiatry. The science of psychoanalysis is the bedrock of psychodynamic understanding and forms the fundamental theoretical frame of reference for a variety of forms of therapeutic intervention, embracing not only psychoanalysis itself but also various forms of psychoanalytically oriented psychotherapy and related forms of therapy employing psychodynamic concepts. Currently considerable interest has been generated in efforts to connect psychoanalytic understandings of human behavior and emotional experience with emerging findings of neuroscientific research. Consequently, an informed and clear understanding of the fundamental facets of psychoanalytic theory and orientation are essential for the student's grasp of a large and significant segment of current psychiatric thinking.

At the same time, psychoanalysis is undergoing a creative ferment in which classical perspectives are constantly being challenged and revised, leading to a diversity of emphases and viewpoints, all of which can be regarded as representing aspects of psychoanalytic thinking. This has given rise to the question as to whether psychoanalysis is one theory or more than one. The divergence of multiple theoretical variants raises the question of the degree to which newer perspectives can be reconciled to classical perspectives.

The spirit of creative modifications in theory was inaugurated by Freud himself. Some of the theoretical modifications of the classic theory after Freud have attempted to reformulate basic analytic propositions while still retaining the spirit and fundamental insights of a Freudian perspective; others have challenged and abandoned basic analytic insights in favor of

divergent paradigms seemingly radically different and even contradictory to basic analytic principles.

One of the difficulties in presenting such a synthetic account is that it must draw its material from more than a century of thinking and theoretical development. Although there is more than one way to approach the diversity of such material, the decision has been made to organize this material along roughly historical lines, tracing the emergence of analytic theory or theories over time, but with a good deal of overlap and some redundancy. But there is an overall pattern of gradual emergence, progressing from early drive theory to structural theory to ego psychology to object relations, and on to self psychology, intersubjectivism, and relational approaches.

THE ROOTS OF PREPSYCHOANALYTIC THINKING

Psychoanalysis was the child of Freud's genius. He put his stamp on it from the very beginning, and it can be fairly said that, although the science and theory of psychoanalysis has advanced far beyond Freud, his influence is still strong and pervasive. In recounting the progressive stages in the evolution of the origins of Freud's psychoanalytic thinking, it is useful to keep in mind that Freud himself was working against the background of his own neurological training and expertise and in the context of the convinced empirical and physicalist scientific thinking of his era.

Scientific Orientation

Freud himself was a convinced empirical scientist whose early training in medicine and neurology had been in the most progressive scientific centers of his time. He shared the prevailing scientific convictions of his day that scientific methods and the systematic study of physical and neurological processes would ultimately yield an understanding of the mysteries of mental processes. When he began his study of hysteria, he believed that brain physiology was the definitive scientific approach and that it alone would yield a truly scientific understanding.

With his own increasing clinical experience, however, Freud was forced to modify that basic scientific credo, but it is significant nonetheless that he maintained it in one or other form throughout the whole of his long career. His own efforts to elaborate a scientific physiology of mental phenomena would in the end prove frustrating and disappointing. After abandoning that attempt, contained in the long lost pages of the *Project for a Scientific Psychology* (1895), he continued to believe that, although the clinical material he dealt with forced him to work on a level of psychological reflection, there was a close and intimate connection between physical and psychical processes.

On Aphasia. An early state of Freud's understanding of how the mind

worked came in his book *On Aphasia* (1891). Although a good deal of attention has been paid to Freud's *Project* as expressing his early model of the mind, more recent attention has been drawn to this earlier important neurological work, which is widely regarded as a classic. In it Freud advanced his earliest views of the relation between structure and function in the brain. Following John Hughlings Jackson's emphasis on the complex relations between thought and language, Freud challenged the prevailing notions of brain localization of function advanced by Pierre Broca, Karl Wernicke, Theodor Meynert, and others. Rather than thinking in terms of brain centers, à la Broca's speech center, Freud related the functions of speech to functional capacities in a widespread network of visual, acoustic, tactile, and even kinesthetic associations reflecting generalized changes in the functioning of the brain as a whole. Thus he viewed simple psychological functions, like perception or memory, as physiologically complex and involving multiple brain systems. In his view it was the disruption in the associative network that was responsible for various forms of aphasia rather than destruction of specific centers. It is interesting that this perspective on language functions in the brain is quite resonant with more contemporary views of the distribution of language functions in the brain.

Following Jackson's differentiation between mind and brain and his concept of functional retrogression from higher to lower levels of functional organization, Freud regarded aphasia as reflecting a form of retrogression to earlier states of speech development. He attributed speech functions to a "zone of language" that was independent of anatomic location, a position that would resonate with his later formulations regarding hysteria in which symptoms were unrelated to anatomical lesions but reflected patterns of meaning and symbolization related to associative networks. In any case, many concepts developed in his study of aphasia would later reemerge in his psychological theory, specifically concepts of association, mental representation, cathexis, symbol formation, and word and object representation. The view of retrogression from higher to lower levels of functioning seems to foreshadow his later doctrine of regression, and his comments on forms of paraphasia read like a preliminary draft of the psychopathology of everyday life.

The Project. The effort to bridge the chasm between psychological processes and neurological mechanisms came to a climax in Freud's attempt to construct a "scientific psychology"; that is, a psychology based on neurological principles. Committed as he was to the scientific ideals of the approach to physiology and psychology developed by Herman Helmholtz, he conceived the scheme of elaborating a complete psychology that would be based on the physicalistic suppositions of the Helmholtz school. For nearly 2 years, from 1895 to 1897, Freud struggled with these

ideas. Finally, in the white heat of intense inspiration, in a period of no more than 3 weeks, he wrote out what is known today as the *Project for a Scientific Psychology*. When the intensity of his inspiration had begun to wane, Freud became increasingly discouraged with what he had written and finally in disgust threw it into a desk drawer where it was to remain for years. He wrote in discouragement and despair, in 1898, to his friend Wilhelm Fliess that he was "not at all inclined to leave the psychology hanging in the air without an organic basis. But apart from this conviction [that there must be such a basis] I do not know how to go on, neither theoretically nor therapeutically, and therefore must behave as if only the psychological were under consideration." It was his intention that the *Project* should be destroyed, but after his death his papers came into the hands of those who recognized its importance and it was finally published posthumously. If it brought Freud's neurological period to a brilliant close, it also opened the way to the broad vistas of psychoanalysis and, in extremely important and significant ways, determined the shape that psychoanalytic principles were to take.

Freud's understanding of the principles of mental functioning traditionally formed the core of how the mental apparatus works and functions is grasped. But within the past half century, the central position of these principles has come into question. The *Project* was based on two principal theorems: First the idea that the nervous system was composed exclusively of neurons, separated by "contact barriers" (Freud's expression for synapses), and second a quantitative concept of neural excitation (Qn) transmitted from cell to cell in the nervous system and either stored or discharged, thus accounting for various forms of nervous activity. The energy in this early model was simply a form of quantitative excitation within a closed-system neuronal reflex model, but it quickly acquired surplus meaning as a hypothetical substance with hydrostatic properties. Out of these simple elements, in combination with a set of regulatory principles, Freud elaborated his complex and ingenious account of mental functioning. The basic model employed in the *Project* centered on a reflex apparatus whose function was withdrawal from stimuli, particularly excessive stimuli, and discharge of accumulated excitation as governed by the *constancy principle,* and the necessity of withdrawing from excessive stimulation in accordance with the *unpleasure principle.*

When Freud finally surrendered his effort to formulate his psychology in terms of a physical model, he was forced to shift to a more specifically psychological model of the mental apparatus, but without completely abandoning the ideas in the *Project*. His thinking remained tied to the physicalistic model of energy systems and their derivation from instinctual drive forces. Surrender of his objective of explaining mental life in terms of physiological and neurological processes was more a compromise than a surrender. In his view, the mind possessed certain dynamic properties so

that the psychological model had to be constructed according to the dynamic laws and principles consistent with current physical theories of the distribution and regulation of the flow of energy. Nonetheless, psychic energy was clearly distinct and different from the metabolic energy of the brain and referred specifically to purposeful striving.

In any case, the concept prevailed that psychic energy represented a purely quantitative and nonqualitative capacity for work, which the psychic apparatus carried out by the transformation, storage, discharge, or delay of discharge of psychic energy. In the classic theory, the instinctual drives impose this demand for work on the mind. Not only does the concept of energy as work potential deriving from drive activation imply the status of drives as sources of independent agency, but also there is no necessary connection or dependence of economic principles on psychic energy. Economic principles are thereby removed from the line of causal efficacy; in other words, economic principles are not principles of efficiency, they are principles of quantitative regulation; efficiency (causality, work) belongs to energic factors related to sources of agency. However, the regulatory principles, as essential to the economic perspective, were originally formulated in terms of the regulation of psychic energies and have been viewed almost exclusively in such terms ever since. They may retain some validity in economic rather than energic terms as regulatory principles specifying how the mental system works (Table 6.1–1). Other questions remain regarding the place of the drives in relation to the capacity for work or force and whether the drives are the only guarantee of connection between the mind and the physical organism.

Freud used psychic energy both as a device to describe observable phenomena and as a construct in his model of the mind. The extent to which Freud utilized the neurological and energic terminology of Helmholtz and Fechner as metaphorical devices to express his psychological constructs has only begun to be appreciated. The use of such metaphors may have utility especially in expressing issues of conflict and development dealt with by psychoanalysis, phenomena that lend themselves readily to verbal metaphoric hypotheses and less readily to mathematical quantification. As Freud's viewpoint became increasingly psychological, his use of concepts of drive and energy became increasingly metaphorical, as in the *Three Essays* (1905). Probably after 1900, Freud became increasingly aware of the limitations of his theory, especially the closed-system and hydraulic aspects of his model, which prompted further revisions. After his abandonment of the *Project,* he preferred to think of his theories as purely psychological, but even so, the energic assumptions carried over into the structural theory.

Table 6.1–1.

Energic Principles Based on Freud's *Project Revised*

Principle	Explanation
Entropy	Freud: Tendency for energy in any physical system to flow from a region of high energy to regions of lower energy. Tendency of system toward homogeneity. Tendency of system to spontaneously diminish the amount of energy available for work. Revised: A revised schema would reject the physical model for psychic processes in favor of a psychological model based on motivation and the tendency of psychic systems to seek and achieve purposeful goals. Energy is expended in action and goal attainment.
Conservation	Freud: The sum of forces (energy) in any isolated (closed) system remains constant. Revised: The psychic apparatus is not a closed system, but open, and does not operate on the basis of closed system dynamics.
Neuronic inertia	Freud: Neurons tend to divest themselves of quantities of excitation. Implies application of principles of entropy and conservation to neuronal activity. Revised: General psychic tendency to resolve situations of conflict, tension, affective imbalance (including anxiety, fear, guilt, shame, depression) in favor of greater balance and more harmonious integration and compromise with other psychic systems.
Constancy	Freud: The nervous system tends to maintain itself in a state of constant tension or level of excitation. Return to a level of constant excitation is achieved by a tendency to immediate energic discharge (through the path of least resistance). Revised: Tendency of psychic functions to return to a state of resting potential as a result of motivated action and achievement of goal satisfaction and as contextual stimulus conditions and the balance of feedback regulated conditions in the same and related psychic systems allows.
Nirvana	Freud: The dominant tendency to reduce, keep constant, or remove the internal psychic tension due to the stimulus excitation. The tendency to reduce the level of excitation to a minimum. Extension of constancy principle. Expressed in pleasure principle, ultimately in death instinct. Revised: This principle would have no place in an economically revised schema, but would be replaced by an opposite tendency to seek stimulation, even to seek complex rather than simple stimulation.
Pleasure-unpleasure	Freud: Tendency of mental apparatus to seek pleasure and avoid unpleasure. Unpleasure is due to the increase of tension or level of excitation, while pleasure is due to the release of tension or discharge of excitation. The pleasure principle thus follows the economic requirements of constancy. Revised: Indicates the degree of functional satisfaction/dissatisfaction and efficacy derived from effective/ineffective operation of psychic systems; pleasure would reflect successful transition from potential to actual operation of the psychic function(s) and achievement of purposeful and wished-for goals. Pleasure is correlative with and corresponds to goal attainment or satisfaction.
Reality	Freud: Modification or delay of energic discharge adapting pleasurable discharge according to the demands of reality. Revised: Modification or delay of psychic functioning as mutually conditioned by the internal context of intra- and intersystemic functions within the mental apparatus; limiting conditions for specific functions outside the mind are determined by reality factors. Satisfaction of motivational needs is tempered by adaptive reality considerations.
Repetition	Freud: Tendency of instinctual forces, as a result of inertial tendencies, to repeat patterns of discharge even when resulting in unpleasure.

Revised: Rather than resulting from inertial energic dynamics, repetition may relate to stabilization of psychic functions and structural integrations, thus contributing to development and the continual process of assimilating and accommodating to reality. Repetition may serve reparative needs and/or contribute to mastery of unresolved conflict, loss, or trauma.

Criticisms of Psychic Energy

All these uncertainties and ambiguities have come to roost in contemporary criticisms of psychic energy. The following objections summarize the objections and the reasons why critics conclude that the classic drive theory and the notion of psychic energy can no longer be tolerated in a contemporary psychoanalytic theory:

1. Psychic energy is not measurable, so any quantitative assumptions of the theory are unable to be tested.
2. The relationship between neural energies in the brain and psychic energy remains vague and poorly understood, so that any laws for the transformation of one to the other remain elusive.
3. The hydraulic analogy is outmoded, and the view of psychic energy was based on a simplified view of causality and a misleading equation of psychic energy with physiological energy as an explanatory principle.
4. The human organism is in some degree tension-seeking and maintaining, while the energic model is based on the principle of tension reduction.
5. Psychic energy comes in multiple forms, that is, libidinal, aggressive, narcissistic, various degrees of neutralized energies, bound energies, and fused energies, and so forth. The difficulty here is not with the varying manifestation of energies in different forms but, rather, with the idea that the differences are inherent in the energies themselves. This objection focuses on the differentiation of the energy itself, as opposed to the idea that various manifestations of psychic energy may be determined by the structures through which it is expressed; qualitative differences would be due to the patterned control, channeling, and mediation of intervening structures. Analogously to physical energy, differences of various forms (i.e., heat, light, chemical, physical) are not attributed directly to the energy as such, but rather to the physical channels through which the energy is expressed. By implication, qualitative energic differences undercut the idea of the id as unstructured chaos consisting only of energy and its modes of discharge.
6. There are also problems with the energic metaphor itself. Freud did not clearly distinguish drive and energy as biological, physiological, or psychological and connected them almost exclusively with the sexual drive, and only belatedly to aggression. The terms share both physical and psychological reference: Cathexis is both an electrochemical charge and a motive. This opened the door not only to conceptual errors, but

also to the use of an essentially nonanalytic model to explain analytic material.

7. If energy serves as a metaphor for experience, it is merely descriptive and not explanatory; if it stands for some neurophysiological function, any explanatory value rests on dualistic mind–brain assumptions.
8. The notion of psychic energy fails to meet minimal criteria of accepted scientific method. Specifically, it is internally contradictory and lacks consistency; it presents a logically closed system that misinterprets metaphor as fact; it involves a tautological renaming of observable and experienced psychological phenomena in energic terms that masquerade as explanations; it is unable to explain all of the relevant data, especially the phenomena of pleasurable tension, exploratory behavior, and stimulus hunger; it tends to lead to a false sense of explanation, particularly insofar as it offers pseudoexplanations that are inconsistent with current knowledge of neurophysiology; and it promotes a form of mind–body dualism—dualistic interactionism—that prevents integration of psychoanalytic concepts with related sciences of the mind and behavior. The metaphor is given surplus meaning, which elevates it to an objective phenomenon and introduces circularity, vitiating any verification of drive-energy concepts.
9. The energic metaphor is inconsistent with current neurophysiological understanding based on principles of selective inhibition rather than energy depletion, or the all-or-none nature of the nervous impulse rather than fluid dynamics.
10. The usefulness of the energic model for clinical purposes has been questioned. Using a quantitative model for qualitative events limits both the range and depth of explanation. Such quantitative translation persists in the name of presumed objectivity and in the belief that it offers a more scientific view of clinical conceptions. The drive-discharge model interprets aim in terms of discharge, thus blurring any distinction between drive and motive. But in the clinical context, libido and sexuality do assume significant connotations of meaning and motive. Even if meaning and motive do not exclude quantitative dimensions, the quantitative cannot substitute for the qualitative.

On the basis of such slender postulates, Freud elaborated a complex and ingenious account of mental functions. He was unable, however, to provide a satisfactory account of either defense or consciousness. In both cases he became embroiled in a continuing regress in which he seemed unable to stop. Despite a variety of ingenious feedback loops that he built into the system (Freud was many decades ahead of his time in envisioning informational servomechanisms), he was unable to complete the functioning of his system without violating the demands of his mechanical principles. He thus introduced into his system a major concession to

vitalism, an observing ego. This observing ego was able to foresee danger for the mobilization of defenses (as in signal anxiety) and was able to sense the indication of quality in conscious experiences. The ego remained as a sort of primary "willer" and ultimate "knower," a personal center within the theory that could not be reduced to the physicalistic terms of Helmholtzian postulates and that consequently enjoyed a significant degree of autonomy.

One of the persistent difficulties inherent in the discussion of psychic energy is that, because of the original mode in which Freud expressed his economic views, the economic hypothesis has become overidentified with the hypothesis of psychic energy. There is little doubt that psychoanalytic theory cannot do without a principle of economics. It would be impossible to express or understand matters of quantitative variation, degrees of intensity, levels and intensity of motivation, informational communication, or to explain how individual subjects are able to make choices among conflicting motivations and goals to bring about the resolution of conflict, or for that matter to explain the whole range of affective, motivational, and structural concepts that form the backbone of psychoanalytic understanding without invoking the very concepts and issues of quantity and intensity that led Freud from the very beginning to postulate an economic point of view. The shifting focus on informational- and communications-based concepts does not escape the need for economic governing principles.

Among theorists for whom the drive theory has fallen into disfavor, the question arises of what to replace it with. This effort has been distracted by a false dichotomy between drive and relational theories, but other possibilities may occupy a middle ground. The classic drive theory serves as the basis for one of Freud's fundamental and persistently valid discoveries, namely that the human mind operates in terms of multiple levels of motivation, some conscious and some unconscious. And further that these levels of motivation can operate in parallel at different hierarchical levels, some more or less conscious, and others operating in combination or in conflict with other motives, more or less unconscious.

The question is whether the drive theory is necessary to substantiate and preserve these essential insights into human behavior. Perhaps not. A motivational theory would shift the basis of understanding to a more specifically psychological level that has consistency and meaningful operation on its own terms without any appeal to drive sources. Such a modification of the dynamic principle requires a clear distinction between drive and motive, concepts that are fused or confused in the classic theory. Drives are sources of independent biologically derived agency and energic force in the mind; in contrast, motives are psychological in nature and take the form of wishes, desires, intentions, and purposes that respond to states of need or lack and function to draw, attract, elicit, guide, or direct the actions of the organism toward some form of goal attainment or

satisfaction.

The distinction is captured neatly in the distinction between efficient causality and finality—drives operate as causes of action, motives as forms of finality; drives are causal, motives teleological. In a revised theory, the efficiency and causality are due to the agency of the self, so that the actions in question are thus actions of the organism or person as such and not of independently operating sources of energic discharge within the self. A motivational theory would eliminate the concept of the unconscious or the id as a cauldron of instinctual energies, and replace it with a system of unconscious instinctual motivations. Such a theory, based on motivation rather than drive discharge, would have the same explanatory potential as the classic drive theory without the complexities, inconsistencies, and scientific disfavor of the drives.

BEGINNINGS OF PSYCHOANALYSIS

In the decade from 1887 to 1897, Freud immersed himself in the serious study of the disturbances in his hysterical patients, resulting in discoveries that contributed to the beginnings of psychoanalysis. These slender beginnings had a threefold aspect: Emergence of psychoanalysis as a method of investigation, as a therapeutic technique, and as a body of scientific knowledge based on an increasing fund of information and basic theoretical propositions. These early researches flowed out of Freud's initial collaboration with Joseph Breuer and then, increasingly, from his own independent investigations and theoretical developments.

The Case of Anna O

Breuer was an older physician, a distinguished and well-established medical practitioner in the Viennese community. Knowing Freud's interests in hysterical pathology, Breuer told him about the unusual case of a woman he had treated for approximately 1.5 years, from December 1880 to June 1882. This woman became famous under the pseudonym Fräulein Anna O., and study of her difficulties proved to be one of the important stimuli in the development of psychoanalysis.

Anna O. was, in reality, Bertha Pappenheim, who later became independently famous as a founder of the social work movement in Germany. At the time she began to see Breuer, she was an intelligent and strong-minded young woman of approximately 21 years of age who had developed a number of hysterical symptoms in connection with the illness and death of her father. These symptoms included paralysis of the limbs, contractures, anesthesias, visual and speech disturbances, anorexia, and a distressing nervous cough. Her illness was also characterized by two distinct phases of consciousness: One relatively normal, but the other reflected a second and more pathological personality.

Anna was very fond of and close to her father and shared with her mother the duties of nursing him on his deathbed. During her altered states of consciousness, Anna was able to recall the vivid fantasies and intense emotions she had experienced while caring for her father. It was with considerable amazement, both to Anna and Breuer, that when she was able to

recall, with the associated expression of affect, the scenes or circumstances under which her symptoms had arisen, the symptoms would disappear. She vividly described this process as the "talking cure" and as "chimney sweeping."

Once the connection between talking through the circumstances of the symptoms and the disappearance of the symptoms themselves had been established, Anna proceeded to deal with each of her many symptoms, one after another. She was able to recall that on one occasion, when her mother had been absent, she had been sitting at her father's bedside and had had a fantasy or daydream in which she imagined that a snake was crawling toward her father and was about to bite him. She struggled forward to try to ward off the snake, but her arm, which had been draped over the back of the chair, had gone to sleep. She was unable to move it. The paralysis persisted, and she was unable to move the arm until, under hypnosis, she was able to recall this scene. It is easy to see how this kind of material must have made a profound impression on Freud. It provided convincing demonstration of the power of unconscious memories and suppressed affects in producing hysterical symptoms.

In the course of the somewhat lengthy treatment, Breuer had become increasingly preoccupied with his fascinating and unusual patient and, consequently, spent more and more time with her. Meanwhile, his wife had grown increasingly jealous and resentful. As soon as Breuer began to realize this, the sexual connotations of it frightened him, and he abruptly terminated the treatment. Only a few hours later, however, he was recalled urgently to Anna's bedside. She had never alluded to the forbidden topic of sex during the course of her treatment, but she was now experiencing hysterical childbirth. Freud saw the phantom pregnancy as the logical outcome of the sexual feelings she had developed toward Breuer in response to his therapeutic attention. Breuer himself had been quite unaware of this development, and the experience was quite unnerving. He was able to calm Anna down by hypnotizing her, but then he left the house in a cold sweat and immediately set out with his wife for Venice on a second honeymoon.

According to a version that comes from Freud through Ernest Jones, the patient was far from cured and had later to be hospitalized after Breuer's departure. It seems ironic that the prototype of a cathartic cure was, in fact, far from successful. Nevertheless, the case of Anna O. provided an important starting point for Freud's thinking and a crucial juncture in the development of psychoanalysis.

Studies on Hysteria

The collaboration with Breuer brought about publication of their "Preliminary Communication" in 1893. Essentially, Freud and Breuer extended Jean Charcot's concept of traumatic hysteria to a general doctrine of hysteria. Hysterical symptoms they thought were related to psychic traumata, sometimes clearly and directly but also sometimes in a symbolic disguise. Observations based on these later cases established a connection between pathogenesis of common hysteria and that of traumatic neurosis; in both cases trauma is not followed by sufficient emotional reaction and is thus kept out of consciousness.

They observed that individual hysterical symptoms seemed to disappear when the event provoking them was clearly brought to life, with the patient describing the event in great detail and putting the accompanying affect into words. Fading of a memory or loss of its associated affect depended on various factors, including whether or not there had been an energic reaction to the event that provoked the affect. Thus, the memories could be regarded as traumata that had not been sufficiently abreacted. They noted that the splitting of consciousness, so striking in classical cases of hysteria as "double consciousness," was present to at least a rudimentary degree in

every hysteria. They described the basis of hysteria as a hypnoid state, that is, a state of dissociated consciousness. They thought psychotherapy achieved its curative effect on hysterical symptoms by bringing to an end the emotional force of the idea that had not been sufficiently abreacted in the first instance. It does this by allowing strangulated affect to gain conscious discharge through speech, thus subjecting it to associative correction, integrating it with normal consciousness.

The "Preliminary Communication" created considerable interest and was followed in 1895 by "Studies on Hysteria" in which Breuer and Freud reported on their clinical experience in the treatment of hysteria and proposed a theory of hysterical phenomena. Freud's case discussions proved to be extremely significant since they formed an early expression of his early psychoanalytic thinking and technique.

Freud concluded from these observations that an experience that had played an important pathogenic role, together with its subsidiary emotional concomitants, was accurately retained in the patient's memory, even when apparently forgotten and unrecoverable by voluntarily recall. He postulated that repression of an idea from consciousness and exclusion from any modification by association with other ideas was an essential condition for development of hysteria. At this early stage, Freud regarded repression as intentional and thought that it served as the basis for conversion of a sum of neural excitation. When cut off from more normal paths of psychic association, this sum of excitation would find its way all the more easily along a deviant path leading to somatic innervation. The basis for such repression, he argued, must be a feeling of unpleasure derived from the incompatibility between the idea to be repressed and the dominant mass of ideas constituting the ego.

Moreover, as one symptom was removed, another would often develop to take its place. The illness could be acquired even by a person of sound heredity as the result of appropriate traumatic experiences. It should be noted that Freud's view at this point was quite different from Breuer's, who ascribed the origin of hysteria to hypnoid states. In Freud's view, in the actual traumatic moment, the incompatibility of ideas forced itself on the ego so that the ego defensively repudiated the incompatible idea. This reaction brought into being a nucleus for crystallization of a separate psychic group somehow divorced from the ego. This process resulted in the splitting of the consciousness characteristic of acquired hysteria. The therapeutic process consisted in compelling this split-off psychical group to unite once more with the main mass of conscious ego ideas. In every case of hysteria based on sexual traumas, Freud felt that impressions from the presexual period, which produce little or no direct effect on the child, can attain traumatic power at a later date as memories when the girl or married woman begins to acquire an understanding and exposure to adult sexual life.

On the basis of these cases Freud reconstructed the following sequence of steps in the development of hysteria: (1) The patient had undergone a traumatic experience, by which Freud meant an experience that stirred up intense emotion and excitation, that was intensely painful or disagreeable to the individual. (2) The traumatic experience represented to the patient some idea or ideas incompatible with the "dominant mass of ideas constituting the ego." (3) This incompatible idea was intentionally dissociated or repressed from consciousness. (4) The excitation associated with the incompatible idea was converted into somatic pathways, resulting in hysterical manifestations and symptoms. (5) What was left in consciousness was merely a mnemonic symbol only connected with the traumatic event by associative links that are frequently enough disguised. And (6) if the memory of the traumatic experience can be brought into consciousness and if the patient is able to sufficiently release the strangulated affect associated with it, then the affect is discharged and symptoms disappear.

Freud's Technical Evolution

The "Studies on Hysteria" provide a valuable picture of the evolution in Freud's development of technical approaches to the treatment of cases of hysteria. His early interest in hypnosis, along with his exposure to hypnotic techniques, both in Charcot's clinic and later at Nancy, led to his extensive use of hypnosis in treating his patients when he opened his own practice in 1887. In the beginning, he used hypnotic suggestion to enable patients to rid themselves of their symptoms. It became quickly obvious, however, that although patients responded to hypnotic suggestion and were relieved of symptoms, the symptoms would nonetheless reassert themselves after a period of time.

By 1889, then, Freud was sufficiently intrigued by Breuer's cathartic method to use it in conjunction with hypnotic techniques as a means of retracing the histories of neurotic symptoms. In his early efforts, he assumed the notion of the traumatic origins of hysterical symptoms. Consequently, the goal of treatment was restricted to removal of symptoms through recovery and verbalization of suppressed feelings with which symptoms were associated. This procedure has since been described as "abreaction." However, as in the case of hypnotic suggestion, Freud was still somewhat dissatisfied with the results of this treatment approach. The beneficial effects of hypnotic treatment seemed to be transitory; they tended to last, or seemed effective, only as long as the patient remained in contact with the physician. Freud began to suspect that alleviation of symptoms was actually dependent in some manner on the personal relationship between patient and physician.

From Hypnosis to Analysis. Freud had begun to suspect that inhibited

sexuality may have played a role in producing the patient's symptoms. His suspicion of a sexual aspect in the treatment of such patients was amply confirmed one day when a female patient awoke from a hypnotic sleep and suddenly threw her arms around his neck. Freud suddenly found himself in the same position as Breuer during his earlier treatment of Anna O. Perhaps bolstered by Breuer's experience and apparently better able to maintain his objectivity, Freud did not panic or retreat in the face of this sexual advance. Rather, he was able to disengage himself sufficiently to evaluate this experience as a scientific observation.

From this point on, Freud began to understand that the therapeutic effectiveness of the patient–physician relationship, seemingly so mystifying and problematical to him until this time, could be attributed in fact to its erotic basis. These observations became the basis of the theory of transference he later developed into an explicit theory of treatment. In any event, these experiences reinforced his dissatisfaction with hypnotic techniques. He became aware that hypnosis might be masking and concealing important manifestations related to the process of cure or, in some cases, to the failure of the patient to achieve a definitive resolution of the neurosis. Later, he came to appreciate that continued use of hypnosis precluded further investigation of transference and resistance phenomena, which had become more central in his thinking about the treatment process.

In the meantime, Freud had also discovered that many of his patients were in fact refractory to hypnosis. Only gradually did he come to recognize that his inability to hypnotize certain patients might often enough be due to that patient's reluctance to remember the traumatic events. He later identified this reluctance as resistance. Thus the vagaries of the hypnotic method left Freud dissatisfied, and he felt it necessary to develop an approach to treatment that could be usefully applied regardless of whether the patient was hypnotizable or not. Consequently, although Freud continued for some time to use hypnotic techniques as a basic approach to treatment of hysteria, he began to experiment with other techniques, leading to further modifications.

CONCENTRATION METHOD. An intermediate next step in the evolution of his technique was the concentration method. One of the patients whom Freud found to be refractory to hypnotic technique was Elizabeth von R. In the treatment of this case, Freud decided to abandon hypnosis as his primary therapeutic tool. He based his decision to alter his technique on the observation of Hippolyte Bernheim that, although certain experiences appeared to be forgotten, they could be recalled under hypnosis and then subsequently recalled consciously if the physician were to ask the patient leading questions urging reproduction of these critical memories. Thus in the method of concentration the patient was asked to lay on a couch and to

close her eyes. She was then instructed to concentrate on a particular symptom and to recall any memories associated with it. The method was substantially a modification of the technique of hypnotic suggestion without actual hypnosis. At times Freud would press his hand on the patient's forehead and urge her to recall the unavailable memories.

Freud's graphic descriptions of this technique convey the unavoidable impression that he was struggling against a force he sensed in the patient and against which he found himself battling, as though in hand-to-hand combat. He came slowly, by dint of this laborious experience, to realize that the isolation of certain memory contents was a matter of the operation of mental forces generating considerable power that kept the complex of pathogenic ideas separate from the mass of conscious ideation. This substantially provided him both with the empirical basis for the notion of resistance and with the basic metapsychological perspective of the mind as operating in terms of psychic forces.

FREE ASSOCIATION. The material presented in such graphic detail in "Studies on Hysteria" reflects dramatically the evolution of Freud's technique in the direction of his more definitive approach to psychoanalysis. He became increasingly convinced by the late 1890s that the process of urging, pressing, questioning, and trying to defeat the resistance offered by the patient—all part and parcel of the "concentration" method—rather than facilitating overcoming of the patient's resistances, actually interfered with the free flow of the patient's thoughts. Piece by piece, then, Freud gave up the concentration method and replaced it with the method of free association. Through this progressive evolution, the basic rule of psychoanalysis—free association—came into focus and was established. Gradually, Freud surrendered his technique of forehead pressure, as well as the requirement that patients close their eyes while lying on the couch. The only remnant of this earlier procedure persisting in the practice of psychoanalysis is the customary use of the couch. The emergence of the central operative principle of psychoanalysis, namely the method of free association, was the end product of this gradual evolution in Freud's thinking.

The evolution of Freud's associative technique continued to progress until it had assumed the form known today. The modification evolved with increasing reliance on the patient's capacity to freely manifest mental contents without suggestive interference on the part of the therapist. By the end of the 19th century, Freud had more or less established his associative technique. In his 1900 work *The Interpretation of Dreams,* he described:

> This [technique] involves some psychological preparation of the patient. We must aim at bringing about two changes in him: An increase in the attention he pays to his own psychical perceptions and the elimination of the criticism by which he normally sifts the thoughts that occur to him. In order that he may be able to concentrate his attention on his self-observation it is an advantage for him to lie in a restful attitude and to shut his eyes. It is necessary to insist

explicitly on his renouncing all criticism of the thoughts that he perceives. We therefore tell him that the success of the psychoanalysis depends on his noticing and reporting whatever comes into his head and not being misled, for instance, into suppressing an idea because it strikes him as unimportant or irrelevant or because it seems to him meaningless. He must adopt a completely impartial attitude to what occurs to him, since it is precisely his critical attitude which is responsible for his being unable, in the ordinary course of things, to achieve the desired unraveling of his dream or obsessional idea or whatever it may be.

In just a few years more, closing of the eyes was also abandoned. Thus, free association became the definitive technique of psychoanalysis. In fact it was development of this technique that opened the door to the exploration of dreams, which became one of the primary sources of data for unconscious effects and provided further support for the nascent psychoanalytic point of view.

Theoretical Innovations

The theoretical point of view expressed in "Studies on Hysteria" was relatively complex. Breuer held the view that hysterical phenomena were not altogether ideogenic, that is, that they were not determined simply by ideas. In fact the phenomena of hysteria may be determined by a variety of causes, some brought about by an explicitly psychical mechanism, but others without it. Although the so-called hysterical phenomena were not necessarily caused by ideas alone, their ideogenic aspects were essential aspects of hysterical symptoms. The main contribution of Freud and Breuer was that they focused on investigation of these ideogenic aspects and discovered some of their psychic origins. Particularly, the concept of neuronal excitation, conceived of as subject to processes of hydraulic flow and discharge as described in the *Project,* was regarded as of fundamental importance in understanding hysteria, as well as neurosis in general.

A careful reading of Breuer's theoretical section in "Studies" makes it abundantly clear that he was proposing an essential reworking of the ingenious ideas of Freud's *Project,* with specific application to the explanation of hysterical phenomena. Breuer described two extreme conditions of central nervous system (CNS) excitation; namely, a clear waking state and the state of dreamless sleep. When the brain was performing actual work, greater consumption of energy was required than when it was merely prepared to do work. The phenomenon of spontaneous awakening could take place in conditions of complete quiet and darkness without any external stimulus. This demonstrated that development of psychic energy is based on vital processes of neural elements themselves.

Breuer also provided an explanation of hysterical conversion. His basic explanatory concept—originally proposed by French psychiatrists, particularly Pierre Janet—was the notion of hypnoid states. Such states were thought to resemble the basic condition of dissociation obtaining in hypnosis. Their importance lay in the amnesia that accompanied them and in their power to bring about splitting of the mind. The spontaneous origin

of such states through a process of autohypnosis was thought to be found fairly frequently in a number of clinically symptomatic hysterics. These hypnoid states often alternated rapidly with normal waking states. Experience of an autohypnotic state was usually related to more or less total amnesia while the patient was in the waking state. Hysterical conversion seemed to take place more easily in such autohypnotic states than in waking states, similar to more facile realization of suggested ideas in states of artificial hypnosis.

Neither hypnoid states during periods of energetic work nor unemotional twilight states are pathogenic. Such reveries, however, when filled with emotion and in states of fatigue arising from protracted affects, did seem to be pathogenic. Occurrence of such hypnoid states was important in this view of the genesis of hysterical phenomena because they somehow made conversion easier and prevented, by way of the resulting amnesia, the converted ideas from dissipating and losing their intensity.

It must be said that Freud had little sympathy with Breuer's concept of hypnoid states, although at this stage, Freud had not yet been able to bring himself to reject it. The concept, in fact, did not explain very much. The hypnoid state was appealed to as an explanation for hysterical states, but the occurrence and function of hypnoid states themselves were in no way explained or supported. They were merely postulated or attributed to a hereditary disposition to such states. Such an unproven postulate was unacceptable to Freud's scientific mind.

The section on the psychotherapy of hysteria, written by Freud, was quite different in spirit from Breuer's theoretical treatment. Freud's discussion of the treatment of hysteria in the "Studies" gives one a good sense of the extent to which Freud had moved away in his own thinking from the somewhat restrictive formulations of the *Project*. He pointed out that each individual hysterical symptom seemed to disappear more or less permanently when the memory of the traumatic event provoking it was brought into conscious awareness along with its accompanying affect. It was necessary for the patient to describe such traumatic events in the greatest possible detail and be able to express verbally the affective experience connected with it.

Freud became convinced that the basic etiology of the neuroses had to be attributed to sexual factors. He proposed that different sexual influences were operative in producing different forms of neurotic disorders. Usually the neurotic picture was mixed, and purer forms of either hysterical or obsessional neurosis were relatively rare. Freud did not regard all hysterical symptoms as psychogenic in origin, so that they could not all be effectively treated by a psychotherapeutic procedure. In the context of his theory and technique at that time, he found that a significant number of patients could not be hypnotized despite an apparently certain diagnosis of hysteria. In these patients Freud believed that he had to overcome a certain psychic

force in the patient, a force that was set in opposition to any attempt to bring the pathogenic idea into consciousness. In the therapy, he experienced himself engaging in sometimes forceful psychic work to overcome this intense counterforce.

The pathogenic idea, however, despite the force of this resistance, was often close at hand and could be reached by relatively easily accessible associations. The patient seemed to be able to rid himself of such ideas by turning them into words and describing them. Nevertheless, Freud's experience was that, even in cases in which he was able to surmise the manner in which things were connected and could tell the patient this even before the patient had actually uncovered it, he could not force anything on the patient about matters in which the patient was essentially ignorant, nor could the therapist influence the product of the analysis by arousing the patient's expectations.

Resistance. The basic question that confronted Freud and Breuer had to do with the mechanism that made and often kept pathogenic memories unconscious. The divergence in their points of view was not simply a matter of theoretical differences. Freud's own thinking underwent a definite transition, and the transition seemed to be based primarily on his experience in dealing with his patients. In the beginning he and Breuer had agreed that their hysterical patients suffered from traumatic sexual experiences and that these traumatic experiences were not available to conscious recollection. They had also agreed, at least for a time, that recovery of these forgotten experiences during an induced hypnotic state resulted in abreaction and consequently symptomatic improvement.

But Freud also discovered that his patients were often quite unwilling or unable to recall the traumatic memories. He defined this reluctance of his patients as resistance. As his clinical experience expanded, he found that, in the majority of patients he treated, resistance was not a matter of simply reluctance to cooperate—that is, the patients seemed to willingly engage in the treatment process and were willing to obey the fundamental rule of free association. The patients generally seemed to be well motivated for treatment, but frequently enough it was particularly patients who were most distressed by their symptoms who seemed most hampered in treatment by resistance. Freud's conclusion was that resistance was a matter of the operation of active forces in the mind, of which the patients themselves were often quite unaware, and which tended to maintain the exclusion from consciousness of painful or distressing material. Freud described this active force that worked to exclude particular mental contents from conscious awareness as *repression,* one of the fundamental ideas of psychoanalytic theory.

Repression. The concept of repression, together with its related notion

of defense, became the basic explanation for hysterical phenomena in Freud's thinking. The notion of repression reflects one of the basic hypotheses of Freud's theory, namely, the dynamic hypothesis according to which the human mind includes in its operation basic dynamic forces that can be set in opposition and that serve as the basic source of powerful motivation and defense. Freud described the mechanism of repression in the following terms: A traumatic experience or a series of experiences, usually of a sexual nature and often occurring in childhood, had been "forgotten" or "repressed" because of their painful or disagreeable nature; but the excitation involved in sexual stimulation was not extinguished, and traces of it persisted in the unconscious in the form of repressed memories. These memories could remain without pathogenic effect until some contemporary event, for example, a disturbing love affair, served to revive them. At this juncture, the strength of the repressive counterforce was diminished, and the patient experienced what Freud termed the *return of the repressed*. The original sexual excitement was revived and found its way by a new path, allowing it to manifest itself in the form of a neurotic symptom. Thus, the symptom results from a compromise between repressed desire and the "dominant mass of ideas constituting the ego." The whole process of repression and the return of the repressed were thus conceived of as involving conflicting forces; that is, the force of the repressed idea struggling to express itself against the counterforce of the ego seeking to keep the repressed idea out of consciousness.

Freud's development of the notions of repression and resistance were based primarily on his studies of cases of conversion hysteria. Specifically, in such cases he felt that impulses that were not allowed access to consciousness were diverted into paths of somatic innervation, resulting in such hysterical symptoms as paralysis, blindness, disturbances of sensations, and other manifestations. Despite this early emphasis on conversion hysteria as the prototype of repression, Freud extended the basic proposition that symptoms resulted from compromise between a repressed impulse and other repressing forces to obsessive-compulsive phenomena and even to paranoid ideation. The logical consequence of this hypothesis was that the treatment process during this period focused primarily on enabling the patient to recall repressed sexual experiences, so that the accompanying excitation could be allowed to find its way into consciousness and be discharged along with the revivified and previously dammed-up affects.

Seduction Hypothesis and Infantile Sexuality. These early researches into hysteria led to an additional striking aspect of psychoanalytic understanding. Invariably, when inquiring into past histories of his hysterical patients, Freud found that the repressed traumatic memories, apparently connected with the root of the pathology, had to do with sexual

experiences. His attention became increasingly focused on the importance of these early sexual experiences, usually recalled in the form of a sexual seduction occurring before puberty and often rather early in the child's experience. Freud began to feel that these seduction experiences were of central importance for understanding the etiology of psychoneurosis. Over a period of several years, he continued to collect clinical material that seemed to reinforce this important hypothesis. He even went so far as to distinguish between the nature of the seductive experiences involved in hysterical manifestations and those involved in obsessional neurosis. In the case of hysteria, he felt, the seduction experience had been primarily passive; that is, the child had been the passive object of seductive activity on the part of an adult or older child. In obsessive-compulsive neurosis, however, he felt the seduction experience had been active on the part of the child. Thus, the child would have actively and aggressively pursued a precocious and traumatizing sexual experience.

A significant aspect in this development was that Freud had taken literally the accounts his patients had given him in the form of forgotten but revived memories of such sexual involvement. The patients provided him with "tales of outrage" committed by such relatives or caretakers as fathers, nursemaids, or uncles. Freud had devoted little attention to the role of the child's own psychological experience in the elaboration of these tales. But, little by little, he began to have some second thoughts about these so-called memories. Several factors contributed to his doubt. First, he had gained additional insight into the nature of pathological processes from his clinical experience and his increasing awareness of the role of fantasy in childhood. Second, he simply found it hard to believe that there could be so many wicked and seductive adults in Viennese society. The third influence, however, which undoubtedly was of major significance in this reconsideration, was his own self-analysis.

As this important process of self-analysis progressed, Freud began to have more and more reason to call the seduction hypothesis into question. During this time, from 1893 to 1897, Freud was still using the combined technique of pressure and suggestion with relatively great assurance. Often he would insist that patients recall the seduction scene, so that much of the evidence on which the seduction hypothesis was based was open to the charge of suggestion. Consequently, as Freud became more aware of the role of suggestion in his technique, his doubts about the seduction hypothesis grew apace. In September 1897 his doubts came to a focus, and he wrote to his good friend Fliess as follows:

> And now I want to confide in you immediately the great secret that has been slowly dawning on me in the last few months. I no longer believe in my *neurotica* [theory of the neuroses]. This is probably not intelligible without an explanation, after all, you yourself found credible what I was able to tell you. So I will begin historically [and tell you] where the reasons for disbelief came from. The continual disappointment in my efforts to bring a single analysis to a real conclusion; the running away of people who for a period of time had been most gripped [by analysis]; the

absence of the complete successes on which I had counted; the possibility of explaining to myself the partial successes in other ways, in the usual fashion—this was the first group. Then the surprise that in all cases, the *father,* not excluding my own, had to be accused of being perverse—the realization of the unexpected frequency of hysteria, with precisely the same conditions prevailing in each, whereas surely such widespread perversions against children are not very probable.... Then, third, the certain insight that there are no indications of reality in the unconscious, so that one cannot distinguish between truth and fiction that has been cathected with affect. [Accordingly, there would remain the solution that the sexual fantasy invariably seizes upon the theme of the parents.] Fourth, the consideration that in the most deep-reaching psychoses the unconscious memory does not break through, so that the secret of childhood experiences is not disclosed even in the most confused delirium. If one thus sees that the unconscious never overcomes the resistance of the conscious, the expectation that in treatment the opposite is bound to happen, to the point where the unconscious is completely tamed by the conscious, also diminishes.

It is obvious that at this period Freud was struggling with his own great reluctance to abandon the seduction hypothesis. The doubts and the clarifying realization that he expressed to Fliess were depressing. After all, he had put in years of effort and had compiled a significant amount of evidence to bolster this seduction hypothesis. It was only with reluctance that he could surrender it. He also sensed, however, that in surrendering the seduction hypothesis, new possibilities for psychological exploration were opened up. In fact this juncture in the development of Freud's thinking was crucial. The so-called abandonment of the seduction hypothesis, with its reliance on actual physical seduction, forced Freud to turn with new realization to the inner fantasy life of the child.

It can be said, in the real sense, that this shift from an emphasis on reality factors to an attention to and an understanding of the influence of inner motivations and fantasy products marks the real beginning of the psychoanalytic movement. In this attempt to distinguish psychic reality and fantasy from actual external events, and psychoneurosis from perversion, psychoanalysis itself took on a new and highly significant dimension. What inevitably emerged from this shift in direction was a dynamic theory of infantile sexuality in which the child's own psychosexual life played the significant and dominant role. This notion would replace the more static point of view in which the child represented an innocent victim whose eroticism was prematurely disrupted at the hands of unscrupulous adults.

The turning point was one of extreme significance for Freud himself. Increasingly, he turned his attention to his own self-analysis and put increasing reliance on it. He wrote to Fliess: "My self-analysis is in fact the most essential thing I have at present and promises to become of the greatest value to me if it reaches its end." More and more, he became involved in the study of dreams, all the more so as he developed the technique of free association, which provided him with a tool for exploring the associative content underlying the dream experience. He concentrated increasingly on the nature of infantile sexuality and on the inner sources of fantasy and dream content, which he conceived to be unconscious

instinctual drives.

In recent years, this so-called abandonment of the seduction hypothesis has been subjected to severe criticism on the grounds that it tended to minimize the role of actual seduction, which has become a pervasive problem in contemporary society. But in Freud's defense, it should be said that he never denied that seduction was a problem; he knew well enough that it existed, as is demonstrated in some of his clinical cases; but at the time it did not provide him a path leading to deeper understanding of the dynamic aspects of instinctual infantile sexual life. The sustained attacks and criticisms of Freud in this regard seem misguided, since, although the shift in his emphasis was clear enough, he never denied or abandoned the idea of actual seduction. A more balanced view would find room for both sides of the seduction equation in which the role of actual seduction and infantile sexual fantasies can find their proportional emphasis. This is in fact the prevailing viewpoint in contemporary analysis.

By 1897, then, when the hypothesis of actual seduction, as the exclusive etiology of neurosis, had fallen in the dust at Freud's feet, he could look to a number of significant accomplishments. The fundamental concepts of psychic determinism and the operation of a dynamic unconscious were established, and concomitantly, a theory of psychoneurosis based on the idea of psychic conflict and the repression of disturbing childhood experiences had become clearly established. Sexuality, particularly in the form of childhood sexuality, had been unveiled as playing a significant but previously underplayed or ignored role in the production of psychological symptoms. More significantly, perhaps, Freud had arrived at a technique, a method of investigation that could be exploited as a means of exploring a wide range of mental phenomena that had previously been poorly understood. Moreover, the horizons of psychoanalytic interest had begun to expand rapidly. Freud's attention was no longer focused on certain limited forms of psychopathology. It had begun to reach out, reflecting the wide-ranging curiosity and interests of Freud's mind, to embrace the understanding of dreams, creativity, wit and humor, the psychopathology of everyday experience, and a host of other normal and culturally significant mental phenomena. Psychoanalysis had indeed come to life.

INTERPRETATION OF DREAMS

Currently, the whole area of sleep and dream activity is one of the most actively studied aspects of psychological functioning. The discovery of rapid eye movement (REM) cycles and the definition of the various stages of the sleep cycle and their relation to dreaming activity have stimulated an intense and extremely productive flurry of research activity into the neurobiology of dreaming. A whole new realm of fresh and important questions has been opened as a result of this activity, and the integration of psychoanalytic and neuroscientific thinking is drawing much closer to a

more comprehensive understanding of the links between patterns of dream activity and underlying neurophysiological and psychodynamic variables. As more is learned about this fascinating and complex question, one comes much closer to understanding the nature of the dream process and the dream experience itself.

In this context it is difficult to look back and to appreciate the uniqueness and originality of Freud's contribution to understanding the dream experience. Only when Freud's attention had been refocused to the significance of inner fantasy experiences, partly as a result of his abandonment of the seduction hypothesis and partly due to his developing the technique of free association, did he come to appreciate the significance and value of the investigation of dreams. Freud became aware of the significance of dreams in his experience with his patients when he realized that, in the process of free association, his patients would frequently report their dreams along with the associative material that seemed connected with them. He discovered little by little that dreams had a definite meaning, although that meaning was often quite hidden and disguised. Moreover, when he encouraged his patients to associate freely to the dream fragments, he found that what they frequently reported was more closely connected with repressed material than to events of their recent waking experience. Somehow the dream content seemed to be closer to the unconscious memories and fantasies of the repressed material, and associations to dream material seemed to facilitate disclosure of this content. He came to regard dreams as the "royal road to the unconscious."

Theory of Dreaming

The rich complex of data derived from Freud's clinical exploration of his patients' dreams and the profound insights based on his associative investigation of his own dreams were distilled into the landmark publication of *The Interpretation of Dreams* in 1900. Basing his analysis on these data, Freud presented a theory of the dreaming process that paralleled his earlier analysis of psychoneurotic symptoms. He viewed the dream experience as a conscious expression of unconscious fantasies or wishes not readily accessible to conscious waking experience. Thus, dream activity was considered to be one of the normal manifestations of unconscious processes.

The dream images represented unconscious wishes or thoughts, disguised through a process of symbolization and other distorting mechanisms. This reworking of unconscious contents constituted the dream work. Freud postulated the existence of a "censor," pictured as guarding the border between the unconscious part of the mind and the preconscious level. The censor functioned to exclude unconscious wishes during conscious states but, during regressive relaxation of sleep, allowed certain unconscious contents to pass the border, only after transformation

of these unconscious wishes into disguised forms experienced in the manifest dream contents by the sleeping subject. Freud assumed that the censor worked in the service of the ego—that is, as serving the self-preservative objectives of the ego. Although he was aware of the unconscious nature of the processes, he tended to regard the ego at this point in the development of his theory more restrictively as the source of conscious processes of reasonable control and volition. It should not be forgot that, even in the "Studies on Hysteria," repression was still envisioned in intentional and volitional terms. His deepening appreciation of the unconscious dimension of these processes led him to view the ego as in some part unconscious, one of the reasons for his later formulation of the structural theory in 1923.

Analysis of Dream Content. Freud viewed dream thoughts as containing content that had been repressed or excluded from consciousness by defensive activities of the ego. The dream material, as consciously recalled by the dreamer, is simply the end result of unconscious mental activity taking place during sleep. Freud hypothesized that the upsurge of unconscious material was so intense that it would threaten to interrupt sleep itself, so that he envisioned that one function of the censor was to act as a guardian of sleep. Instead of being awakened by these ideas, the sleeper dreams.

From a more contemporary viewpoint, it is known that cognitive activity during sleep has a great deal of variety. Some cognitive activity follows the description that Freud provided of dream activity, but much of cognitive activity in the sleep state is considerably more realistic and more consistently organized along logical lines. The dreaming activity that Freud analyzed and described is probably more or less associated with the stage 1 REM periods of the sleep–dream cycle. The so-called manifest dream embodying the content of the dream as experienced by the dreamer, which the sleeper may or may not be able to recall after waking, is the product of dream activity. Unconscious thoughts and wishes that in Freud's view threatened to awaken the sleeper were described as "latent dream content." Freud referred to unconscious mental operations by which latent dream content was transformed into the manifest dream as the "dream work." In the process of dream interpretation, he was able to move from the manifest content of the dream by way of associative exploration to arrive at the latent dream content that lay behind the manifest dream and that provided it with its core meaning.

In Freud's view, a variety of stimuli initiated dreaming activity. The contemporary understanding of the dream process, however, would suggest that dreaming activity takes place more or less in conjunction with patterns of CNS activation that characterize certain phases of the sleep cycle. What Freud thought to be initiating stimuli may in fact not be

initiating at all, but may be merely incorporated into the dream content and determine to what extent the material presents in the dream thoughts. Stimuli could arise from various sources. Although the dreaming process may reflect an underlying pattern of neurophysiological activation, it also serves as the vehicle for expressing unconscious motivations and meanings.

NOCTURNAL SENSORY STIMULI. A variety of sensory impressions, for example, pain, hunger, thirst, or urinary urgency, may play a role in determining dream content. Thus, instead of disturbing one's sleep and leaving a warm bed, a sleeper who is in a cold room and who urgently needs to urinate may dream of awaking, voiding, and returning to bed. Freud's view would have been that the activity of dreaming preserved and safeguarded the continuity of sleep. It is known now, however, that the function of dreaming is considerably more complex and cannot be regarded simply as preserving sleep, although there is still room for this process to be counted among the dream functions.

DAY RESIDUES. One of the important elements contributing to shaping the dream thoughts is the residue of thoughts, ideas, and feelings left over from experiences of the preceding day. These residues remain active in the unconscious and, like sensory stimuli, can be incorporated by the sleeper into thought content of the manifest dream. Thus, day residues could be amalgamated with unconscious infantile drives and wishes deriving from the level of unconscious instincts. The amalgamation of infantile drives with elements of the day's residues can effectively disguise the infantile impulse and allow it to remain effective as a driving force behind the dream. Day residues may in themselves be quite superficial or trivial, but they acquire significance as dream instigators through unconscious connections with deeply repressed instinctual drives and wishes.

REPRESSED INFANTILE DRIVES. Although these various elements may be determining aspects of the thought content of the dream experience, the essential elements of the latent dream content were thought to derive from one or more impulses emanating from the repressed part of the unconscious. In Freud's schema, the ultimate driving forces behind dream activity and dream formation were the wishes, originating in drives, stemming from an infantile level of psychic development. These drives and wishes took their content specifically from oedipal and preoedipal levels of psychic integration. Thus, nocturnal sensations and day residues only played an indirect role in determining dream content. A nocturnal stimulus, however intense, had to be associated with and connected with one or more repressed wishes from the unconscious to give rise to the dream content. This point of view needs some revision because it seems that in some phases of nighttime cognitive activity the mind is able to process residues of daytime experience without much indication of

connection with unconscious repressed content. However, in phases of cognitive activity during sleep that bear the stamp of dreaming activity, as Freud described and defined it, this essential link to the repressed probably still retains some validity.

Significance of Dreams. Once Freud's attention had definitively shifted to the study of inner processes of fantasy and dream formation, the study of dreams and the process of their formation became the primary route by which he gained access to understanding unconscious processes and their operation. In *The Interpretation of Dreams,* he maintained that every dream somehow represents a wish fulfillment. He bolstered this hypothesis with a considerable amount of documentation, including exhaustive analysis of his own dreams.

There is a more general tendency today to view the dream activity as expressing a broader spectrum of psychological processes, preserving the aspect of wish fulfillment as one among the dimensions of dream activity, but not as an absolute principle, as it seemed to be in Freud's thinking. The manifest dream content may represent imaginary fulfillment of a wish or impulse from early childhood, before such wishes have undergone repression. In later childhood, and even later in adulthood, however, the ego acts to defend itself against unacceptable instinctual demands or impulses of the unconscious. Wish fulfillment in the dream process is usually quite obscured by the extensive distortions and disguises brought about by the dream work, so that it often cannot be readily identified on a superficial examination of the manifest content without associative connections and reverberations.

The Dream Work

The theory of the nature of dream work became the fundamental description of the operation of unconscious processes—the basic mechanisms and the manner of their operating—that stands even today as an unsurpassed and foundational account of unconscious mental functioning. The focus of Freud's analysis was on the process by which unconscious latent dream thoughts were disguised and distorted in such a fashion as to permit their expression and translation into conscious manifest content of the dream. However, these unconscious processes also found ready application and extrapolation not only to understanding the formation of neurotic symptoms but, more broadly, to a whole range of unconscious mental productivity. The theory of dream work consequently became the basis for a wide-ranging analysis of unconscious operations that found expression in Freud's study of everyday experiences, as well as artistic creativity, jokes and humor, and a variety of culturally based activities of the human mind. Aspects of the dream work are as follows.

Representability. A basic problem in understanding dream formation is to determine how it is that latent dream content can find a means of representation in the manifest content. As Freud saw it, the state of sleep brought with it relaxation of repression, and concomitantly, latent unconscious wishes and impulses were then permitted to press for discharge and gratification. Because the pathway to motor expression was blocked in the sleep state, these repressed wishes and impulses had to find other means of representation by way of mechanisms of thought and fantasy. Activity of the dream censor provided continual resistance to discharge of these impulses, with the result that the impulses had to be attached to more neutral or "innocent" images to be able to pass the scrutiny of censorship and be allowed into conscious expression. This displacement was made possible by selecting apparently trivial or insignificant images from residues of the individual's current psychological experience and linking these trivial images dynamically with latent unconscious images, presumably on the basis of some resemblance that allowed associative links to be established. In the process of facilitating economic expression of latent unconscious contents and, at the same time, of maintaining the distortion that was essential for unconscious contents to escape the repressing action of the censor, the dream work used a variety of mechanisms, making it possible for more neutral images to both represent and transform repressed infantile components. These mechanisms, as envisioned in Freud's theory, included *symbolism, displacement, condensation, projection,* and *secondary revision.*

SYMBOLISM. *Symbolism* is a complex process of indirect representation that in the psychoanalytic usage has the following connotations:

1. A symbol is representative of or substitute for some other idea from which it derives a secondary significance that it does not possess of itself.
2. A symbol represents this primary element by reason of a common element that these ideas share.
3. A symbol is characteristically sensory and concrete in nature, as opposed to the idea it represents, which may be relatively abstract and complex. A symbol thus provides a more condensed expression of the idea represented.
4. Symbolic modes of thought are more primitive, both ontogenetically and phylogenetically, and represent forms of regression to earlier stages of mental development. Consequently, symbolic representations tend to function in more primary process or relatively regressed conditions—in the thinking of primitive peoples, in myths, in states of poetic inspiration, and particularly in dreaming.
5. A symbol is thus a manifest expression of an idea that is more or less hidden or secret. Typically, the use of the symbol and its meaning is unconscious. Thus, symbols tend to be used spontaneously,

automatically, and unconsciously. The use of symbols is a sort of secret language in which instinctually determined content can be reexpressed by other images; for example, money can symbolize feces, or windows can symbolize the female genitals.

Many questions still persist about the origins of symbolic processes, the stage of development in which they become organized, the extent to which they require altered states of consciousness such as the sleep state for their implementation, and the degree to which symbolic expression is related to underlying conflicts. Current formulations would regard the symbolic function as a uniquely human trait involved in all forms of human mental activity, from the most primitive expression of infantile wishes to the most complex creative processes of literary, artistic, religious, as well as scientific thinking.

DISPLACEMENT. The mechanism of *displacement* refers to the transfer of amounts of energy (cathexis) from an original object to a substitute or symbolic representation of the object. Because the substitute object is relatively neutral—that is, less invested with affective energy—it is more acceptable to the dream censor and can pass the border of repression more easily. Thus, whereas symbolism can be taken to refer to the substitution of one object for another, displacement facilitates distortion of unconscious wishes through transfer of affective energy from one object to another. Despite the transfer of cathectic energy, the aim of the unconscious impulse remains unchanged. For example, in a dream, the mother may be represented visually by an unknown female figure (at least one who has less emotional significance for the dreamer), but the naked content of the dream nonetheless continues to derive from the dreamer's unconscious instinctual impulses toward the mother.

CONDENSATION. *Condensation* is the mechanism by which several unconscious wishes, impulses, or attitudes can be combined into a single image in the manifest dream content. Thus, in a child's nightmare, an attacking monster may come to represent not only the dreamer's father, but may also represent some aspects of the mother and even, in addition, some of the child's own primitive hostile impulses as well. The converse of condensation can also occur in the dream work, namely, an irradiation or diffusion of a single latent wish or impulse distributed through multiple representations in the manifest dream content. The combination of mechanisms of condensation and diffusion provides the dreamer with a highly flexible and economic device for facilitating, compressing, and diffusing or expanding the manifest dream content, which is derived from latent or unconscious wishes and impulses.

PROJECTION. The process of *projection* allows dreamers to rid themselves

of their own unacceptable wishes or impulses and experience them as emanating in the dream from other persons or independent sources. Not surprisingly, the figures to whom these unacceptable impulses are ascribed in the dream often turn out to be those toward whom the subject's own unconscious impulses are directed. For example, the individual who has a strong repressed wish to be unfaithful to his wife may dream that his wife has been unfaithful to him; or a patient may dream that her analyst has sexually approached her, although she is reluctant to acknowledge her own repressed wishes toward the analyst. Similarly, the child who dreams of a destructive monster may be unable to acknowledge his or her own destructive impulses and the fear of the father's imagined power to hurt the child. The figure of the monster consequently would be a result of both projection and displacement.

SECONDARY REVISION. The mechanisms of symbolism, displacement, condensation, and projection are all characteristic of relatively early modes of cognitive organization in a developmental sense. They reflect and express the operation of the primary process. In the organization of the manifest dream content, however, primary process forms of organization are supplemented by a final process that reorganizes the absurd, illogical, and bizarre aspects of the dream thoughts into a more logical and coherent form, particularly in the context of recounting the dream in a coherent narrative to someone else. Thus the distorting effects of symbolism, displacement, and condensation, through a process of secondary revision, acquire a coherence and rationality that are necessary for acceptance on the part of the subject's more mature and reasonable ego. *Secondary revision* uses intellectual processes that more closely resemble organized thought processes governing rational states of consciousness. It is through secondary revision, then, that the logical mental operations characteristic of the secondary process are introduced into and modify dream work.

Affects in the Dream Work. In the process of displacement, condensation, symbolization, or projection, as Freud hypothesized, the energic component of the instinctual impulses was separated from its representational component and followed an independent path of expression in the form of affects or emotions. The repressed emotion may not appear in the manifest dream content at all, or it may be experienced in a considerably altered form. Thus, for example, repressed hostility or hatred toward another individual may be modified into a feeling of annoyance or mild irritation in the manifest dream expression, or it may even be represented by an awareness of not being annoyed—that is, a conversion of the affect into its absence. The latent affect may even possibly be directly transformed into its opposite in the manifest content, for example, as when a repressed longing might be represented by a manifest

repugnance or vice versa. Thus, the vicissitudes of affect and the transformation by which latent affects are disguised introduce another dimension of distortion into the content of the manifest dream. The vicissitudes of affect, then, took place in addition to and in parallel with processes of indirect representation that characterized the transformation of dream content.

Regression. In the seventh theoretical chapter of *The Interpretation of Dreams,* Freud provided a model of the psychic apparatus as he understood it at the turn of the 20th century. Not only was the model a description of the functioning of the dreaming mind, but it also represented a broader conceptualization of the psychic apparatus as it functioned in both pathological and normal human experience, as he had formulated previously in his *Project*. It seems clear that the economic model, on which Freud had expended such intense effort in the 1890s and had seemingly abandoned in frustration, had revived to reassert itself, now in a new language and in a different setting. The lines of continuity and parallels between the model of the mind in the *Project* and the model of the seventh chapter of *The Interpretation of Dreams* could not be appreciated until the manuscript of the *Project* was rediscovered after Freud's death.

The model was an elaborate construction based on a basic notion of a stimulus–response mechanism. In normal waking experience sensory input was taken into the receptor end of the apparatus and then processed through a number of mnemonic systems of increasing degrees of elaborateness and complexity. After varying degrees of processing, the impulse was subsequently discharged through the motor effector apparatus. In the dream state, however, motor effector pathways were blocked so that, instead of discharge through motor systems, excitation was forced to move in a backward or regressive direction through the mnemonic systems and back toward expression through sensory systems.

During waking hours, the path leading from unconscious levels of the apparatus through preconscious to conscious levels was barred to the dream thoughts by activity of the censor. In sleep, however, this pathway was again made more available because resistance of the censor was diminished. Consequently, unconscious memories and their instinctual determinants could press for discharge through the perceptual apparatus, as was particularly the case in hallucinatory dream experiences. Thus, dreams could be described as having a regressive character. Consisting specifically in turning back of an idea into the sensory image from which it was originally derived, regression was an effect of the resistance opposing discharge of psychic energy associated with the thought into consciousness along the normal path. Regression was also abetted by the simultaneous attraction exercised on the dream thought by associated memories in the unconscious.

In dreams, regression was further facilitated by diminution of the progressive informational current flowing from continuing sensory input during waking hours. Regression, as Freud viewed it, was essentially a regression to the originating source of an impression in the reversal he described within the mental apparatus, but it also was a regression in time. Freud thus distinguished several forms of regression, namely, a topographic regression involving a regression from conscious to unconscious systems within the mental model; a temporal regression according to which the mental process refers back to older psychical structures, particularly those deriving from an infantile level of development; and formal regression, in which more primitive methods of expression and representation take the place of the more normal ones. "All these three kinds of regression," Freud commented in *The Interpretation of Dreams,* "are, however, one at bottom and occur together as a rule; for what is older in time is more primitive in form and in psychical topography lies nearer to the perceptual end."

Primary and Secondary Process. Perhaps the most central aspect of the functioning of this mental model, closely related to the formulations of the *Project,* has to do with Freud's notions of the primary and secondary processes. To begin with, the impulses and instinctual wishes originating in infancy served as the indispensable nodal force for dream formation. The energic conception of these drives followed the basic economic principles laid down by the *Project* (Table 6.1–1). They were elevated states of psychic tension in which the energy was constantly seeking discharge, according to the constancy principle and the pleasure principle.

The tendency to discharge, however, was opposed by other psychic systems. Thus, Freud envisioned two fundamentally different kinds of psychic processes involved in the formation of dreams. One of these processes tended to produce a rational organization of dream thoughts, which was of no less validity in terms of contact with reality than normal thinking. However, another system—the first psychic system in Freud's schema—treated the dream thoughts in a bewildering and irrational manner. He felt that a more normal train of thought could only be submitted to abnormal psychic treatment if an unconscious wish, derived from infancy and in a state of repression, had been transferred onto it. As a result of the operation of the pleasure principle, the first psychic system was incapable of bringing anything disagreeable into the context of the dream thoughts. It was unable to do anything but wish. Operating in conjunction with the demands of this primary system, the secondary system could only cathect an unconscious idea if it could inhibit any development of unpleasure that might have proceeded from the coming to awareness of that idea. Anything that might evade that inhibition would be equivalently inaccessible to the second system, as well as to the first,

because it would promptly be eliminated in accordance with the unpleasure principle.

The psychic process derived from the operation of the first system is referred to as *primary process*. The process resulting from the inhibition imposed by the second system is referred to as the *secondary process,* and it reflects the operation of the system of inhibition and delay sketched in the *Project*. The secondary system thus corrects and regulates the primary system in accord with principles of logic, rationality, and reality. Among the wishful impulses derived from infantile impulses, there are some whose fulfillment would contradict the purposes and ideas of secondary process thinking. The fulfillment of these wishes could no longer generate an affect of pleasure, but would be unpleasurable. This formulation, it should be noted, formed the basis for Freud's later elaboration of the pleasure principle as opposed to the reality principle. The secondary process organization of preconscious thinking is aimed at avoiding unpleasure, at delaying instinctual discharge, and at binding mental energy in accordance with the demands of external reality and the subject's moral principles or values. Thus, functioning of the secondary process is closely connected with the reality principle and is governed, for the most part, by the dictates of the reality principle.

TOPOGRAPHIC THEORY

Beginning with abandonment of the seduction hypothesis and the concomitant turning of Freud's interests to inner processes of fantasy and dream formation, and ending with publication of *The Ego and the Id* in 1923, in which Freud propounded his structural model of the psychic apparatus, Freud's thinking was cast largely in terms of the topographic theory.

Basic Assumptions

There were a number of assumptions underlying Freud's thinking that served as lines of continuity between various stages of his investigations and helped him to organize his thinking in terms of successive models of the mental apparatus. The first assumption was that of "psychological determinism," according to which all psychological events, including behaviors, feelings, thoughts, and actions, are caused by—that is, are the end result of—a preceding sequence of causal events. This assumption derived from Freud's Helmholtzian convictions and represented application of a basic natural-science principle to psychological understanding; but it was also reinforced by Freud's clinical observation that apparently meaningless hysterical symptoms, which had been previously attributed to somatic etiology, could be relieved by relating them to past, apparently repressed, experiences. Thus, apparently arbitrary

pathological behavior could be tied into a causal psychological network. Although Freud seemed to maintain a view of hard determinism requiring a causal explanation and predetermination for every phenomenon, more contemporary views are more accepting of forms of soft determinism requiring only that psychic acts be motivated to fulfill the demands of deterministic explanation.

The second assumption is that of "unconscious psychological processes." This assumption derived from a considerable amount of evidence gathered through the use of hypnosis, but it was also consolidated by Freud's experience of the free associating of his patients in which unconscious and past experiences came to awareness. The unconscious material, which survived and was able to influence present experience, was found to be governed by specific regulatory principles, for example, the pleasure principle and the mechanisms of primary process that differed radically from those of conscious behavior and thought processes. Thus, unconscious processes were brought within the reach of psychological understanding and explanation. In a motivational perspective, the same emphasis is preserved, but unconscious processes are viewed as related to unconscious motives rather than drive forces.

The third assumption was that "unconscious psychological conflicts" between and among psychic forces formed the basic elements at the root of psychoneurotic difficulties. This assumption related to Freud's experience of resistance and the drive to repression in his patients. The full realization of this aspect of psychic functioning came only with awareness that the reports of patients represented not memories of actual experiences but, rather, unconscious fantasies. The assumption of unconscious forces accounted for the process that created those fantasies and brought them into consciousness during free association. It also accounted for the agency that opposed the coming to consciousness of such fantasies. This counterforce that clashed with the sexual drives and diverted them into fantasies or symptoms was related to the function of censorship as developed in the dream theory and, later, to the operation of ego instincts that were set in opposition to the sexual instincts. In motivational terms, conflicts arise between forms and levels of motivation, some conscious and some unconscious.

The final assumption of the topographic theory was that there existed "psychological energies" that originated in instinctual drives. This assumption was derived from the observation that recall of traumatic experiences and their accompanying affects resulted in disappearance of symptoms and anxiety. This suggested, therefore, that a displaceable and transformable quantity of energy was involved in the psychological processes responsible for symptom formation. Freud originally assumed that this quantity was equivalent to the affect, which became dammed up or strangulated when it was not appropriately expressed and, thus, was

transformed into anxiety or conversion symptoms. After he had developed his notion of instinctual drives, this quantitative factor was conceived of as drive energy (cathexis). As noted previously in the discussion of the *Project,* the assumption of psychic energies served Freud as an important heuristic metaphor. The usefulness of the metaphor and its necessity as a basic assumption of analytic theory have subsequently been questioned and found wanting. Refocusing these observations in motivational terms loses nothing in terms of explanatory potential.

Topographic Model

Freud's thinking about the mental apparatus at this time was based on the classification of mental operations and contents according to regions or systems in the mind. These systems were described neither in anatomical nor spatial terms but were specified, rather, according to their relationship to consciousness. The topographic model has essentially fallen into relative disfavor because of its limited utility as a working model of psychoanalytic processes and largely because it has been surpassed and supplanted by the structural theory. The topographic viewpoint, however, is still useful for classifying mental events descriptively in terms of the quality and degree of awareness. There is a tendency currently to revive aspects of the topographic model of the mind in viewing mental processes as descriptively more or less conscious or unconscious, rather than as a reflecting operation of a mental structure as in the systemic unconscious of classic metapsychology. This reflects a current tendency to see conscious and unconscious mentation as a continuum of levels of consciousness or the lack of it.

Conscious. The conscious system was that region of the mind in which perceptions coming from the outside world or from within the body or mind were brought into awareness. Internal perceptions could include introspective observations of thought processes or affective states of various kinds. Consciousness was by and large a subjective phenomenon, the content of which could only be communicated by language or behavior. It has also been regarded psychoanalytically as a sort of superordinate sense organ, which can be stimulated by perceptual data impinging on the CNS. It was assumed that the function of consciousness used a form of neutralized psychic energy called *attention cathexis.*

The nature of consciousness was described in less detail in Freud's early theories, and certain aspects of consciousness are not yet completely understood and actively debated by psychoanalysts. Freud regarded the conscious system as operating in close association with the preconscious. Through attention, the subject could become conscious of perceptual stimuli from the outside world. From within the organism, however, only elements in the preconscious were allowed to enter consciousness. The rest

of the mind lay outside awareness in the unconscious. Currently, there is greater awareness of the extent to which unconscious derivatives can play a role in influencing conscious behavior. Before 1923, however, Freud also believed that consciousness controlled motor activity and regulated the qualitative distribution of psychic energy.

Preconscious. The preconscious system consisted of those mental events, processes, and contents that were for the most part capable of reaching or being brought into conscious awareness by the act of focusing attention. The quality of preconscious organizations may range from reality-oriented thought sequences, or problem-solving analysis with highly elaborated secondary process schemata, all the way to more primitive fantasies, daydreams, or dream-like images, which reflect a more primary process of organization. Thus, it stands over and against unconscious processes in which the transformation to consciousness is accomplished only with great difficulty and by dint of the expenditure of considerable energy in overcoming the barrier of repression.

The preconscious has been amplified by recent findings in neuroscientific study of memory. An essential distinction is that between episodic memory and procedural memory. *Episodic memory* deals with past events in the individual's experience that are usually autobiographical or semantic in content. Other memories, however, have more to do with skills and habitual patterns of behavior, as for example, riding a bike, driving a car, playing the piano, grammatical rules, social norms of politeness and etiquette, and so forth. These are aspects of normal daily living and behavior that are rarely if ever thought about, people just do them, but the procedures are embedded in memories and are readily applied without any effort to recall them. In fact, any effort to recall them more than likely only interferes with their employment. These memory systems, along with others that can be differentiated, are apparently served by different neural circuits and have different connections with consciousness and behavior. Additional questions have arisen concerning levels of access to conscious recovery and the association of preconscious and unconscious mental processes.

Unconscious. Unconscious mental events, namely, those not within conscious awareness, can be described from several viewpoints. One can think of the unconscious descriptively, that is, as referring to the sum total of all mental contents and processes at any given moment outside the range of conscious awareness, including the preconscious.

One can also think of the unconscious dynamically, that is, as referring to those mental contents and processes that exercise a pressure for discharge on the rest of the mental apparatus but remain incapable of achieving consciousness because of the operation of counterforces of

censorship or repression. This repressive force or "countercathexis" manifests itself in psychoanalytic treatment as resistance to remembering. The unconscious mental contents in this dynamic sense consist of drive representations or wishes that are in some measure unacceptable, threatening, or abhorrent to the intellectual or ethical standpoint of the individual. This results in intrapsychic conflict between the repressed forces and the repressing forces of the mind. When repressive countercathexis weakens, this may result in formation of neurotic symptoms. The symptom is thus viewed as essentially a compromise between conflicting forces. These unconscious mental contents are also organized on the basis of infantile wishes or drives and strive for immediate discharge, regardless of the reality conditions. Consequently, the dynamic unconscious is thought to be regulated by the demands of primary process and the pleasure principle.

Finally, there is a systemic sense of the unconscious referring to a region or system within the organization of the mental apparatus that embraces the dynamic unconscious and within which memory traces are organized by primitive modes of association, as dictated by the primary process. This systemic view of the unconscious is considered, in a specifically topographic sense, as a component subsystem within the topographic model, and in the structural theory is attributed to the id. Consequently, the systemic unconscious can be described in terms of the following characteristics in Freud's view:

1. Ordinarily, elements of the systemic unconscious are inaccessible to consciousness and can only become conscious through access to the preconscious, which excludes them by means of censorship or repression. Repressed ideas, consequently, may only reach consciousness when the censor is overpowered (as in psychoneurotic symptom formation), relaxes (as in dream states), or is fooled (in jokes).
2. The unconscious system was exclusively associated with primary process thinking. The primary process had as its principal aim facilitation of wish fulfillments and instinctual discharge. Consequently, it was intimately associated with—and functioned in terms of—the pleasure principle. As such, it disregarded logical connections, permitted contradictions to coexist simultaneously, recognized no negatives, had no conception of time, and represented wishes as fulfillments. The unconscious system also employed the same primitive mental operations that Freud identified in the operation of the dream process. Moreover, the quality of motility, characteristic of primary process thinking and of unconscious energy, was also frequently linked to the capacity for creative thinking.
3. Memories in the unconscious have been divorced from their connection with verbal symbols. Freud discovered in the course of his clinical work that repression of a childhood memory could occur if the energy was

withdrawn from it and, especially, if the verbal energy was removed. When the words were reconnected to the forgotten memory traits (as during psychoanalytic treatment), it became recathected and could thus reach consciousness once more.

4. The content of the unconscious was limited to wishes seeking fulfillment. These wishes provided the motive force for dreams and neurotic symptom formation. It has already been noted that this view may be oversimplified.
5. The unconscious was closely related to the instincts. At this level of theory development, the instincts were considered to consist of sexual and self-preservative (ego) drives—aggression was added later. The unconscious was thought of as containing mental representatives and derivatives particularly of the sexual instincts.

Dynamics of Mental Functioning

Freud conceived of the psychic apparatus, in the context of the topographic model, as a kind of reflex arc in which the various segments have a spatial relationship. The arc consisted of a perceptual or sensory end through which impressions were received; an intermediate region, consisting of a storehouse of unconscious memories; and a motor end, closely associated with the preconscious, through which instinctual discharge could occur. In early childhood, perceptions were modified and stored in the form of memories.

According to this theory, in ordinary waking life the mental energy associated with unconscious ideas sought discharge through thought or motor activity, moving from the perceptual end to the motor end of the apparatus. Under certain conditions, such as external frustration or sleep, the direction in which energy travels along the arc was reversed, and it moved from the motor end to the perceptual end instead of the other way around. It thereby tended to reanimate earlier childhood impressions in their earlier perceptual forms and resulted in dreams during sleep or hallucinations in mental disorders. This reversal of the normal flow of energy in the psychic apparatus was the "topographic regression" discussed previously. Although Freud subsequently abandoned this model of the mind as a reflex arc, he retained the central concept of regression and applied it later in somewhat modified form in the theory of neurosis. The theory states that libidinal frustration results in reversion to earlier modes of instinctual discharge or levels of fixation, which had been previously determined by childhood frustrations or excessive erotic stimulations. Freud called this kind of reversion to instinctual levels of fixation *libidinal* or *instinctual regression*.

Framework of Psychoanalytic Theory: Repressed Versus Repressing.

Throughout his long lifetime and in the course of many twistings and

turnings of the theoretical developments in his thinking, Freud's mind was dominated by a tendency to describe many aspects of mental functioning in terms of contrasting polarities; some of the primary polarities were subject (ego) versus object (outer world), pleasure versus unpleasure, and activity versus passivity. The fundamental and dominant dualism was between the forces and contents of the mind viewed as repressed and unconscious and those forces and mental agencies responsible for the act of repressing. Although the persistence of such basic dualisms in psychoanalytic thinking has clear advantages and undoubtedly helps one understand some fundamental aspects of the mind, one should not forget that such paradigms may prove to be overly restrictive. There is real question in the current state of psychoanalysis as to whether some of these assumed basic dimensions may not in fact be limiting the capacity of psychoanalytic theory to grow apace with the expanding horizons of both clinical experience and experimental, especially neuroscientific, exploration. The historical role and the present vitality of the basic psychoanalytic dualisms, however, should not be undervalued since they provide powerful tools for understanding and treating clinical pathology.

INSTINCTUAL THEORY

Freud postulated that all human beings have similar instincts or drives. The actual discharge of instinctual impulses is organized, directed, regulated, or even repressed by functions of the individual ego, mediating between the organism and the external world. Historically, Freud's early theorizing was concerned primarily with the nature and functioning of instinctual drives.

Concepts of Instincts

One of the first problems in the theory of instincts is what is meant by the term instinct. The problem is made more complex by the variation in usage between a primarily biological meaning and Freud's primarily psychological concept. The difficulties are also compounded by the complexities in Freud's own use of the term. The term instinct was introduced primarily in the study of animal behavior, referring generally to patterns of species-specific behavior based mainly on potentialities in the animal determined by heredity and therefore considered to be relatively independent of learning. The term was used to explain a great variety of behavior patterns, for example, appealing to a maternal instinct, a nesting instinct, or a migratory instinct. Such usage resisted successful physiological explanation and tended to introduce strong teleological connotations, implying some sense of purposefulness built into the instinct, as in the concept of an instinct of self-preservation. Freud adopted this usage unquestioningly, but even strong proponents of instinctual theory among animal behaviorists, for whom the line between instinctual and

learned behavior has become increasingly more complex and debatable, have questioned its validity. The dichotomy of nature versus nurture can no longer be simplistically or rigidly maintained. Thus, instinctually derived patterns of behavior are seen to be increasingly modifiable in the interests of adaptation. Ethologists consequently prefer to speak simply of species-typical behavior patterns that are based on innate equipment but that mature and develop or are elicited through some form of environmental interaction.

Freud, of course, took as the basis of his thinking the older concept of instinct, but in adopting it for his purposes he transformed it. Actually, Freud's own formulation of the notion of instinctual drives underwent contextual modification so that he actually offered a variety of definitions. Perhaps the most cogent, as formulated in *Instincts and Their Vicissitudes,* was: "An 'instinct' appears to us as a concept on the frontier between the mental and the somatic, as the psychical representative of the stimuli originating from within the organism and reaching the mind, as a measure of the demand made upon the mind for work in consequence of its connection with the body." It is immediately evident that the basic ambiguity in the concept of instinct between biological and psychological aspects continued to influence Freud's thinking about instinctual drives and remains latent in subsequent psychoanalytic usage of the term. Freud himself varied in the emphasis he placed on one or other aspect of the concept, so that subsequent discussions of the concept of instinct in psychoanalysis have varied similarly, and at times confusingly, between emphasis on biological aspects and emphasis on psychological aspects.

Theory of the Instincts

When Freud began his investigation into the nature of unconscious forces, he strove consistently to base psychoanalytic theory on a firm biological foundation. One of the most important measures of his attempt to link psychological and biological phenomena came in basing his theory of motivation on instincts. Freud viewed instincts as a class of borderline concepts that functioned between the mental and organic spheres. Consequently, his use of the term "instinct" is not always consistent because it emphasizes either the psychic or biological aspect of the term in varying degrees in varying contexts. Sometimes, then, *libido* refers to the somatic process underlying the sexual instinct, and at other times, it refers to the psychological representation itself. Thus, Freud's usage is quite divergent from the Darwinian implications of the term "instinct," which implies innate, inherited, unlearned, and biologically adaptive behavior.

Characteristics of the Instincts. Freud ascribed to instinctual drives four principal characteristics: *Source, impetus, aim,* and *object*. In general the *source* of an instinct refers to the part of the body from which it arises, the

biological substratum that gives rise to the organismic stimuli. The source, then, refers to a somatic process that gives rise to stimuli, which are represented in the mental life as drive representations or affects. In the case of libido the *stimulus* refers to the process or factors that excite a specific erotogenic zone. The *impetus* or *pressure behind the drive* is a quantitative economic concept referring to the amount of force or energy or demand for work made by the instinctual stimulus. The *aim* is any action directed toward satisfaction or tension release. The aim in every instinct is satisfaction, which can only be obtained by reducing the state of stimulation at the source of the instinct. The *object* is the person or thing that is the target for this satisfaction-seeking action and that enables the instinct to gain satisfaction or discharge the tension and thus gain the instinctual aim of pleasure.

These characteristics deserve some comment. The notion of source reflects Freud's view of instinctual drives as independent sources of psychic activation. The source was presumed to be bodily and operating to demand satisfaction. From the point of view of motivational theory, there may be sources of motivation that transcend or do not necessarily depend on such physical derivation. Nonetheless, if one accepts a view of the mind–body relation as integrated such that mental acts are effectively brain actions, the body is an integral part of any instinctual process, but the connection is not necessarily with a single organ system. The impetus concerns the role of instinctual drives as independent sources of energic activation creating a constant pressure on the mind for work. In motivational terms, the pressure is not necessarily biologically driven, causal, and constant, but is elicited in response to a motive related to satisfaction of a specific need. There is an impetus to action, but it is aroused in response to a specific motive. The aim is consistent with the motive that is directed to satisfaction of a specific need state, but it can be related to a putative drive only by way of the motive. Freud's view is sustainable only if one combines or confuses motives and drives. With regard to the object, Freud commented that it was the most variable characteristic of the instinct, because it is only appropriate to the extent that its characteristics make satisfaction possible —a view that has been significantly revised in the light of a later understanding of object relations. Although this early view of the instinctual object long held sway in psychoanalytic thinking, it has come under some serious criticism recently. Considerably more weight is put on the significance of the objects of libidinal attachment, particularly by object relations theorists. Increasingly, it has become apparent that the psychoanalytic concept of instincts is meaningless, unless it includes and derives from a context of object relatedness. Moreover, it can no longer be said simply that the objects of infantile drives are the most variable characteristic of the instinct, because attachment to the primary objects, particularly the mothering object, is of the utmost significance

developmentally.

CONCEPT OF LIBIDO. The ambiguity in the term *instinctual drive* is also reflected in use of the term libido. Briefly, Freud regarded the sexual instinct as a psychophysiological process that had both mental and physiological manifestations. Essentially, he used the term libido to refer to "the force by which the sexual instinct is represented in the mind." Thus, in its accepted sense, *libido* refers specifically to the mental manifestations of the sexual instinct. Freud recognized early that the sexual instinct did not originate in a finished or final form, as represented by the stage of genital primacy. Rather, it underwent a complex process of development at each phase of which the libido had specific aims and objects that diverged in varying degrees from the simple aim of genital union. The libido theory thus came to include all of these manifestations and the complicated paths they followed in the course of psychosexual development.

INFANT SEXUALITY. It had long been supposed, as one of the favored myths of analytic lore, that Freud's thought on infantile sexuality constituted an assault on the cherished ideas of 19th century and Victorian thinking and that he was violently attacked for his views of the erotic life of young children. It seems, however, that his significant contribution, the 1905 *Three Essays on the Theory of Sexuality,* came to light not as a revolutionary work but as part of a flood of literature dealing with sexual problems.

Freud had become convinced of the relationship between sexual trauma, in both childhood traumata and the genesis of psychoneurosis, and disturbances of sexual functioning in the so-called actual neuroses—that is, hypochondriasis, neurasthenia, and anxiety neuroses. Freud originally viewed these conditions as related to misuse of the sexual function. For example, he thought anxiety neurosis to be due to inadequate discharge of sexual products, leading to the damming up of libido that was then converted into anxiety. Also, he attributed neurasthenia to excessive masturbation and a diminution in available libidinal energy. In any case, these views reflected Freud's increasing awareness of the importance of sexual factors in the etiology of psychoneurotic states and indicated his reliance on economic considerations.

PART INSTINCTS. Freud described the erotic impulses arising from pregenital zones as *component* or *part instincts*. Thus, kissing, stimulation of the area surrounding the anus, or even biting the love object in the course of lovemaking are examples of activities associated with these part instincts. The activity of component instincts or early genital excitement may undergo displacement, as for example, the eyes in looking (scoptophilia), and may consequently be a source of pleasure. Ordinarily, these component instincts undergo repression or persist in a restricted

fashion in sexual foreplay. More specifically, young children are characterized by a polymorphous-perverse sexual disposition. Their total sexuality is relatively undifferentiated and encompasses all of the part instincts. In the normal course of development to adult genital maturity, however, these part instincts are presumed to become subordinate to the primacy of the genital region.

Aggression and Ego Instincts. The aggressive drives held a peculiar place in Freud's theory. His thinking about aggression underwent a gradual evolution. Early in his thinking, his attention had been preoccupied by the problems posed by libidinal drives. He was quite aware that aggressive impulses were often expressed in the operation of libidinal factors, but he could not long avoid taking explicit account of the more destructive aspects of instinctual functioning. Undoubtedly, also, the horrors and destructiveness of World War I made a significant impression on him, so that he began to realize more profoundly the significance of destructive urges in human behavior.

By 1915, Freud had arrived at a dualistic conception of the instincts as divided into sexual instincts and ego instincts. He recognized a sadistic component of the sexual instincts, but this still lacked a sound theoretical basis. Oral, anal, and phallic levels of development all had their sadistic components. Devoid of any manifest eroticism, and covering a wide range from sexual perversions to impulses of cruelty and destructiveness, the sadistic aspects certainly had different aims from the more strictly libidinal.

Increasingly, Freud saw the sadistic component as independent of the libidinal and gradually segregated it from the libidinal drives. Moreover, impulses to control, tendencies toward the acquisition and exercise of power, and defensive trends toward attacking and destroying all manifested a strong element of aggressiveness. It seemed, then, that there was sadism associated with the ego instincts, as well as with the libidinal instincts. Freud once again followed the dualistic bent of his mind and postulated two groups of instinctual impulses, two qualitatively different and independent sources of instinctual impulses with different aims and modalities. With the publication of *The Ego and the Id* in 1923, Freud gave aggression a separate status as an instinct with a separate source, which he postulated to be largely the skeletomuscular system, and a separate aim of its own, namely, destruction. Aggression was no longer a component instinct, nor was it a characteristic of the ego instincts; it was an independently functioning instinctual system with aims of its own.

The elevation of aggression to the status of a separate instinct, on a par with sexual instincts, dealt a severe blow to any lingering romantic notions of the essentially or exclusively benign nature of man. Aggression and destructiveness were seen as inherent qualities of human nature, such that aggressive impulses were elicited whenever an individual was sufficiently

thwarted or abused. Freud's new formulation also drew attention to the specific role of aggression in forms of psychopathology, as well as to understanding of the developmental processes through which aggression could be normally integrated and controlled.

It should be noted that aggression remains a problem for psychoanalytic thinking even today. Although a great deal has been learned about the operation and vicissitudes of aggression since Freud originally struggled with it, there is still a great deal that remains to be learned about its nature, its origins, the conditions that produce and unleash it, as well as the developmental factors that contribute to its pathological deviations and to its more constructive integrations in realms of human functioning. Some more recent revisions of aggression see it less in terms of destructive or sadistic aims, but more broadly as a capacity for effective action in the face of obstacles or opposition embracing capacities for mastery and self-assertion, and as related to patterns of motivation rather than as a biologically determined drive force. More often than not, destructive aggressive outcomes tend to be motivated by threats to the narcissism of the self.

Life and Death Instincts. When Freud introduced his final theory of life and death instincts in *Beyond the Pleasure Principle* in 1920, he took what can now be seen as an inevitable and logical next step in the evolution of the instinct theory he had been developing. It was nonetheless a highly speculative attempt to extrapolate the directions in which his instinct theory was taking shape to the broad realm of biological principles. One can recall that Freud's thinking about the instincts always cast its shadow in a dual modality. In the beginning he had distinguished sexual and ego instincts. This distinction provided the basic dichotomy for the explanation of psychological conflict and the understanding of psychoneurosis.

The introduction of the life and death instincts must be seen in the course of this development and as extending the inherent duality of instinctual theory to the level of ultimate and final biological principles. Freud had not divorced his notion from the underlying economic principles, derived from principles of entropy and constancy. The constancy principle was extended to the nirvana principle, the objective of which was cessation of all stimuli or a state of total rest. It was only a small, subsequent step that led Freud from the formulation of a nirvana principle to the death instinct, or Thanatos. Freud postulated that the death instinct was a tendency of all organisms and their component cells to return to a state of total quiescence—that is, to an inanimate state.

In opposition to this instinct he set the life instinct, or eros, referring to tendencies of organic particles to reunite, of parts to bind to one another to form greater unities, as in sexual reproduction. As Freud viewed the matter, the ultimate destiny of all biological matter, driven by the

inexorable tendencies of all life to follow principles of entropy and constancy (with the exception of the germ plasm), was to return to an inanimate state. He felt that the dominant force in biological organisms had to be the death instinct. In this final formulation of life and death instincts, the instincts were considered to represent abstract biological principles, which transcended the operation of libidinal and aggressive drives. The life and death instincts represented the forces underlying sexual and aggressive instincts. Consequently, they represented a general trend in all biological organisms.

Needless to say, Freud's extravagant speculation has been subjected to severe criticism. It is impossible to argue that a general biological principle exists merely on the basis of clinical observation. If the inherent destructiveness of some states of psychopathology can permit the inference of destructive forces operating in the individual psyche, it by no means points to the existence of inherent and biologically determined forces of self-destructive potential. However one regards the argument as a biological speculation, for these thinkers it has little relevance as a psychological speculation. On the contrary, the life and death instincts are alive and flourishing in Kleinian and French analytic circles. The school of analysts following the lead of Melanie Klein constitutes the most significant group of psychoanalytic theorists who embrace the death instinct. Kleinian analysis bases a considerable portion of its understanding of intrapsychic processes on the operation of the life and death instincts. In Klein's work with severely disturbed children, she ascribed the aggressive behaviors and fantasies in such children to the operation of the death instinct. This point of view seems to collapse the intervening steps in the organization of instinctual theory and makes almost any manifestation of destructive aggression a direct expression of the death instinct. Although contributions of Klein and her followers to the psychopathology of childhood disturbances are significant, other schools of analytic thinking have not followed their lead in this conceptualization of the primary instincts.

NARCISSISM AND THE DUAL INSTINCT THEORY

The concept of narcissism holds a pivotal position in the development of psychoanalytic theory. It was Freud's dawning realization of the importance of narcissism that led him to important modifications in his understanding of libido and his instinct theory. At the same time, Freud's examination of narcissism and its related clinical phenomena led to an increasing concern with the origins and functions of the ego. It must be said that the introduction of and focus on narcissism have had broad implications and reverberations in psychoanalytic thinking since Freud's day. The whole problem of narcissism remains difficult and problematic for psychoanalysis. The problem of pathological narcissism remains a focus of active interest, thinking, and clinical concern even today. The problem has

special relevance with regard to certain forms of character pathology, which are relatively resistant to therapeutic intervention.

In exploring aspects of pathological narcissism, Freud observed that in cases of dementia praecox (schizophrenia), libido appeared to have been withdrawn from other persons and objects and turned inward. He concluded that this detachment of libido from external objects might account for the loss of reality contact so typical of these patients. He speculated that the detached libido had then been reinvested and attached to the patient's own ego, resulting in megalomaniacal delusions and suggesting that this libidinal reinvestment found expression in its grandiosity and omnipotence.

Freud also became aware at the same time that narcissism was not limited to these psychotic manifestations. It might also occur in neurotic and, to a certain extent, even in "normal" individuals under certain conditions. He noted, for example, that in states of physical illness and hypochondriasis, libidinal cathexis was frequently withdrawn from outside objects and from external activities and interests. Similarly, he speculated that in sleep libido was withdrawn from outside objects and reinvested in the person's own body. Thus, he thought it could be that the hallucinatory and emotional intensity of the dream experience might result from the libidinal cathexis of fantasy representations of the persons who composed the dream images. Freud also appealed to the basically narcissistic form of object choice in perversions, particularly homosexuality.

The introduction of narcissism into his theory played a significant role because it required that he reconcile his theory of libido with what now seemed to be a libidinal force operating within the ego. Freud originally thought of the reinvestment of libido as directed to the ego as such. This formulation has given rise to a considerable confusion in the understanding of narcissistic libido. A decisive reorganization of the concept of narcissism was provided by Heinz Hartmann when he pointed out that it was more accurate to regard narcissistic libido as attached, not to the ego as such, but to the self. The ego, as an intrapsychic construct, was opposed to the self as related to external objects extrapsychically. The proper opposition, then, between object libido and narcissistic libido was that the former is attached to object-representations, whereas the latter is attached to self-representations. This important shift in the understanding of narcissism has opened an area of theoretical reconsideration, which is still very much in flux, and has introduced into psychoanalytic thinking the concept of self as an important, albeit as yet ill-defined, intrapsychic structural component.

Narcissism and the Choice of Love Object

Reference was made earlier to the crucial role of early object relationships in the later choice of love objects. Freud had found that a deepened

understanding of the vicissitudes of narcissism made it easier to understand the basis for choice of certain love objects in adult life. A love object might be chosen, as Freud put it, "according to the narcissistic type," that is, because the object resembles the subject's idealized self-image (or fantasied self-image). Possibly the choice of object might be an "anaclitic type," in which case the object might resemble someone who took care of the subject during the early years of life.

In summary, the concept of narcissism occupies a central and pivotal position in psychoanalytic theory. With the introduction of the concept of narcissism, it became clear that further understanding and advances in psychoanalytic theory would depend on a clearer definition of the concept of self and its more adequate delineation from the concept of ego. Attempts to implement such understanding have brought into focus the ambiguities in the concept of the ego and have underscored the need for the systematic study of its development, structure, and functions. Attention to narcissistic phenomena has also enlarged the understanding of a variety of mental disorders, as well as various normal psychological phenomena. These issues will be discussed in relation to treatment issues.

STRUCTURAL THEORY AND EGO PSYCHOLOGY

The topographic theory was essentially a transitional model in the development of Freud's thinking and served an important function in providing a framework for the development of his basic instinct theory. However, the problems inherent in the topographic theory underscored, once again, the need for a more systematic concept of psychic structure. The main deficiency of the topographic model lay in its inability to account for two extremely important characteristics of mental conflict.

The first important problem was that many of the defense mechanisms that Freud's patients used to avoid pain or unpleasure, and which appeared in the form of unconscious resistances during psychoanalytic treatment, were themselves not initially accessible to consciousness. He drew the obvious conclusion that the agency of repression, therefore, could not be identical with the preconscious, because this region of the mind was by definition easily accessible to consciousness. The second problem was that he found that his patients frequently exhibited an unconscious need for punishment or an unconscious sense of guilt. According to the topographic model, however, the moral agency making this demand was allied with the anti-instinctual forces available to consciousness in the preconscious level of the mind.

From Topographical to Structural Perspective

The germination of the shifting currents of Freud's thinking finally came to fruition in his abandoning the topographic model and replacing it with the

structural model of the psychic apparatus in *The Ego and the Id*. The introduction of the structural hypothesis initiated a new era in psychoanalytic thinking. The structural model of the mind, or the "tripartite theory" as it is often called, was composed of three distinct entities or organizations within the psychic apparatus—the id, the ego, and the superego.

The terms have become so familiar and the tendency to hypostatize them so great that it is well to bear in mind their nature as scientific constructs. The terms are theoretical constructs that have as their primary referents specific groups of mental functions and operations that they are used to classify. Each refers to a particular aspect of mental functioning, and none of them expresses or represents the sum total of mental functioning at any one time. If they often are spoken of as though they functioned as quasi-independent systems, they are, nonetheless, ultimately coordinated aspects of the operation of the mental apparatus representing mental actions of the person. Attribution of agency to any one of them is, therefore, a form of misplaced concreteness since their actions and functions are basically those of the person him- or herself. In this sense, there is only one agent in the person, that is the self; the tripartite entities are equivalently substructures of the self-reflecting different forms of self-generated activity. Moreover, unlike such phenomena as infantile sexuality or object relations, id, ego, and superego are not empirically demonstrable phenomena in themselves but must be inferred from the observable effects of the operations of specific psychic functions.

Historical Development of Ego Psychology

The evolution of the concept of ego within the framework of the historical development of psychoanalytic theory parallels to a large extent the shifts in Freud's view of the instincts and can be divided into four phases. The first phase ended in 1897 and coincided with the development of the early psychoanalytic formulations. The second phase extended from 1897 to 1923, thus spanning the development of psychoanalysis proper. The third phase, from 1923 to 1937, saw development of Freud's theory of the ego and the gradual emergence to prominence of the ego in the overall context of the theory. Parallel to this development was the evolution of Freud's thinking about anxiety. Finally, the fourth phase, coming after Freud's death, saw the emergence and systematic development of a general psychology of the ego, as well as a shifting of focus from the operation of ego functions themselves to the broader social and cultural contexts within which the ego developed and functioned. This last phase has led to further delineations between the functioning of the ego as such and the self-system.

First Phase: Early Concepts of the Ego. In the initial phase, coinciding

with Freud's early theory, the ego was not always precisely defined. Rather, it referred to the dominant mass of conscious ideas and moral values, which were distinct from impulses and wishes of the repressed unconscious. The ego was concerned primarily with defense, a term Freud soon replaced with the notion of repression, so that repression and defense were regarded as synonymous. In the neurophysiological jargon of the *Project,* the ego was described as "an organization ... whose presence interferes with passages of quantity (of excitation)." Translating this into the language of psychology, the ego was regarded as an agent defending against certain ideas that were unacceptable to consciousness. These ideas were found to be primarily sexual in nature and were initially thought to have been engendered by premature sexual trauma and real seduction. Presumably, because memory of such trauma led to arousal of unpleasant and painful affects, they evoked a defensive response and repression of the original thought content. This repression, however, led to a damming up of energy and the consequent production of anxiety. Functioning of this "early ego" was contradictory to a degree because its primary purpose was to reduce tension and thus avoid unpleasant affects connected with sexual thoughts, but in the process of repression it seemed to evoke an equally unpleasant affect state, that of anxiety.

Second Phase: Historical Roots of Ego Psychology. During the years preceding publication of *The Ego and the Id* in 1923, analysis of the ego as such received little direct attention because Freud was concerned primarily with the instinctual drives—their representatives and transformations. Consequently, references to defense or defensive functions were much less frequent. The clarification of these concepts required further elucidation of the ego, its functions, and the nature of its organization. It was during this second phase that Freud grappled with these problems and gradually approached the more definitive resolution provided by the structural theory.

The ego's relationship to reality is particularly relevant in this connection. As noted earlier, the concept of a secondary process implies the ability to delay discharge of instinctual drives in accordance with demands of external reality. Introduction of the reality principle provided a principle of regulation for secondary process functioning comparable to the pleasure principle for primary process. The capacity for delay was later to be ascribed to the ego. The progression from pleasure principle to reality principle in childhood involves a similar capacity to "postpone gratification" and thereby conform to the requirements of the outside world by way of reality testing.

Finally, if neither the preconscious nor the ego instincts were solely responsible for repression or censorship, how was repression to be achieved? Freud tried to answer this question by postulating that ideas are

maintained in the unconscious by a withdrawal of libido or energy (cathexis). In the manner characteristic of unconscious ideas, however, they were constantly driven by libidinal energies to renew their attempt to reach consciousness. To prevent this, the withdrawal of libido must be constantly repeated. Freud described this process as "anticathexis" or "countercathexis." Again, however, if such countercathexis is to be consistently effective against unconscious ideas, it must be permanent and must itself operate on an unconscious basis. Understanding of psychic structure, specifically of the ego, which could perform this complicated defensive function, was clearly called for and constituted still another indication of the need for the development of ego psychology. Thus, the way was pointed toward the third phase, wherein the ego was delineated as a structural entity and separated definitively from the instinctual drives.

Third Phase: Freud's Ego Psychology. With publication of *The Ego and the Id,* the phase of introduction and development of Freud's own theory of the ego was accomplished. The ego was presented as a structural entity, a coherent organization of mental processes and functions, primarily organized around the perceptual conscious system, but also including structures responsible for resistance and unconscious defense. The ego at this stage, however, was viewed as relatively passive and weak. Its functioning was still a resultant of mediating the pressures deriving from id, superego, and reality. The ego was the helpless rider on the id's horse, adopting a Platonic image, more or less obliged to go where the id wished to go. The assumption remained that the ego was not only dependent on forces of the id, but was somehow genetically derived and differentiated out of the id. Freud had as yet to recognize any real development of the ego comparable to the phases of libidinal development.

During this period the view of the ego underwent radical transformation. Some of the details of this development took place in connection with Freud's theory of anxiety. In *Inhibitions, Symptoms, and Anxiety* in 1926, Freud repudiated the conception of ego as subservient to the id. Signal anxiety became an autonomous function for initiating defense, and the capacity of the ego to turn passively experienced anxiety into active anticipation was underlined. Here, too, the relatively rudimentary conception of the defensive capacity of the ego was enlarged to include a variety of defenses that the ego had at its disposal and could utilize in the control and direction of id impulses. Moreover, elaboration of Freud's conception of the reality principle introduced a function of adaptation that allowed the ego to curb instinctual drives when action prompted by them would lead into real danger.

The effect of this transformation of his theory of the ego was threefold. First, it brought the ego into prominence as a powerful regulatory force responsible for integration and control of behavioral responses. Second, the

role of reality was brought to center stage in the theory of ego functioning. It had been banished to the wings in the preceding quarter century, but concern with the adaptive function of the ego again brought it back to prominence. Even so, the conception of adaptation here was rudimentary and limited to the ego's capacity to avoid danger. The notions that Freud was evolving during this phase provided the foundation for the later concept of the autonomy of the ego, as developed by later theorists. Finally, it was toward the end of this period that Freud finally made explicit the assumption of independently inherited roots of the ego that were quite independent of the inherited roots of the instinctual drives. This formulation was taken over by Hartmann and served as the basis for his notion of primary ego autonomy, which consequently stimulated the developments of the fourth phase.

Fourth Phase: The Systematization of Ego Psychology. If the third phase can be thought of as culminating in Anna Freud's work on the defense mechanisms of the ego (1936), the fourth phase can be seen as taking its initiation from the publication of Heinz Hartmann's work on the ego and adaptation (1939). Hartmann's work primarily focused on two aspects of Freud's later notions of the ego; namely, the autonomy of the ego and the problem of adaptation. Discussion of the apparatuses of primary autonomy was the basis for a doctrine of the genetic roots of the ego and a development of the notion of epigenetic maturation. He also recognized that ego structures and functions, arising in conflict, could undergo a change in function to become relatively autonomous from drives in the forms of the so-called secondary autonomy. Hartmann's treatment of adaptation also brought the adaptational point of view into focus in such a way that it has become generally acceptable as one of the basic metapsychological assumptions of psychoanalytic theory.

Although this development of thinking about the ego was an important advance, many psychoanalysts began to feel that it created an imbalance in the theory and that, by increasingly focusing on the mechanical and quantitative aspects of ego functioning, it left a picture of personality functioning and dysfunctioning that seemed relatively mechanistic and inhuman. Moreover, there developed a widening split between the id, the vital stratum of the mind and the dynamic source of psychic energies, and the noninstinctual, nondynamic, structural apparatuses of the ego. Consequently, the id increasingly came to be seen as the source of instinctual energies—the image of the seething cauldron—without the representational or directional qualities that so long characterized Freud's views of the instincts and their functions.

The other extremely important aspect of the fourth phase is reemergence of the importance of reality in its broadest and most profound meanings as a significant dimension of psychoanalytic thinking. This is in

many ways a direct extrapolation of Hartmann's thinking about adaptation, because the adaptive functioning of the organism has directly to do with fitting in with the requirements of external reality and adaptively interacting with the environment, not only the inanimate, but also the personal and social environment. Correlative to the increasing concern with the relation between the individual and his or her environment, there has been a resurgence of interest in the self insofar as relations with the outside world occur between self and other (following Hartmann's distinction between self and ego). Emerging paradigms include Heinz Kohut's self psychology and other more structurally oriented views of the self in psychoanalytic terms.

Structure of the Psychic Apparatus

From a structural viewpoint, Freud divided the psychic apparatus into three groups of functions designated as id, ego, and superego, distinguished by their different functions. The *id* is the locus of the instinctual drives and drive energy and is organized in terms of primary process. It operates according to dictates of the pleasure principle, without regard for the limiting demands of reality. The *ego,* however, represents a coherent organization of functions, whose task is to avoid unpleasure or pain by opposing or regulating the discharge of instinctual drives to conform to demands of the external world. The regulation of id discharges is also contributed to by the third structural component of the psychic apparatus, the *superego,* which contains the internalized moral values, ideals, prohibitions, and standards of the parental imagoes.

The Id. Freud separated the instinctual drives into a separate compartment, the vital stratum of the mind, and in so doing reached the culminating point of the evolution of his theory of instincts. In contrast to his concept of the ego as having an organized, problem-solving capacity, Freud conceived of the id as a completely unorganized, primordial reservoir of energy, derived from the instincts and under the domination of the pleasure principle and primary process. It was not, however, synonymous with the unconscious, because certain functions of the ego, specifically certain defenses against unconscious instinctual pressures, were also unconscious; for the most part the superego also operated on an unconscious level.

The Ego. The conscious and preconscious functions typically associated with the ego—for example, words, ideas, or logic—do not account entirely for its role in mental functioning. The discovery that certain phenomena that emerge most clearly in the psychoanalytic treatment setting, specifically repression and resistance, both associated with the ego, could themselves be unconscious pointed to the need for an

expanded concept of the ego as an organization retaining a close relationship to consciousness and external reality and yet performing a variety of unconscious operations in relationship to drives and their regulation. Once the scope of the ego had been thus broadened, consciousness was redefined as a mental quality that, although exclusive to the ego, constitutes only one of its qualities or functional aspects, rather than a separate mental system itself as in the topographic model.

No more comprehensive definition of the ego is available than the one Freud himself provided toward the end of his career in his *Outline of Psychoanalysis:*

> *Here are the principal characteristics of the ego.* In consequence of the pre-established connection between sense and perception and muscular action, the ego has voluntary movement at its command. It has the task of self-preservation. As regards external events, it performs that task by becoming aware of stimuli, by storing up experiences about them (in the memory), by avoiding excessively strong stimuli (through flight), by dealing with moderate stimuli (through adaptation) and finally by learning to bring about expedient changes in the external world to its own advantage (through activity). As regards internal events, in relation to the id, it performs that task by gaining control over the demands of the instinct, by deciding whether they are to be allowed satisfaction, by postponing that satisfaction to times and circumstances favourable in the external world or by suppressing their excitations entirely. It is guided in its activity by consideration of the tension produced by stimuli, whether these tensions are present in it or introduced into it.

Thus, the ego controls the apparatuses of motility and perception, contact with reality, and, through mechanisms of defense, inhibition, and control of primary instinctual drives.

ORIGINS OF THE EGO. If the ego is defined as a coherent system of functions for mediating between instincts and the outside world, one must concede that the newly born infant has no ego or, at best, the most rudimentary of egos. Nonetheless, the neonate certainly has a rather complex array of intact ego capacities, including both sensory and motor functions. There is, however, much current discussion of the extent to which these functions are organized, but developmentalists are increasingly inclined to credit the newborn infant with surprisingly well-developed and adaptive capacities that facilitate his or her capacity to react to and relate with caretaking figures. Developmental ego psychology concerns itself with tracing the paths and stages by which ego capacities mature and increase their scope and power, leading ultimately to a mature and adaptive set of ego functions. Currently these developments are viewed more holistically in relation to the emergence and consolidation of a sense of self.

Freud believed that the ego developed out of modifications of the id, and that this occurred as a result of the impact of the external world on the drives. Pressures of external reality enabled the ego to appropriate energies of the id to do its work. In this process of formation, the ego sought to bring the influences of the external world to bear on the id, to bring the effects of the reality principle to bear on the pleasure principle, and thereby

contribute to its own further development. In summary, Freud emphasized the role of instincts in ego development and, particularly, the role of conflict. At first this conflict is between the id and the outside world, but later it is between the id and the ego itself.

DEVELOPMENT OF THE EGO. In addition to the maturation of relatively autonomous ego functions, the ego is built up on the basis of processes by which aspects of the external world are acquired and become qualities of ego functioning. These processes through which the internal world is built up and by which structure is consolidated within the self are referred to under the heading of internalization. Forms of internalization —*incorporation, introjection,* and *identification*—are variously connected with development of the ego.

Incorporation was originally conceived of as an instinctual activity derived from and based developmentally on the oral phase and was considered as a genetic precursor of identification. However, although incorporative fantasies are often associated with internalizing processes, they are by no means identical and may be quite independent. Some authors have envisioned incorporation as the mechanism of primary identification, aimed at a primary union between oneself and the maternal object. Incorporation as a mechanism of internalization seems to involve a primitive oral wish for union with an object. The union has a quality of totality and globalization, so that in the internalization of the object, the object loses all distinction and function as object. The external object is completely assumed into the person's inner world. Incorporation thus comes into play in relatively more infantile or regressive conditions in which the sense of distinction of the object as separate is lost.

Introjection is perhaps the most central process in development of the structural apparatus involving ego and superego. Introjection was originally described by Freud in *Mourning and Melancholia* as a process of narcissistic identification in which the lost object is introjected and thus retained as a part of the internal structure of the psyche. Freud later applied this mechanism to the genesis of superego, making introjection the primary internalizing mechanism by which parental imagoes were internalized at the close of the oedipal phase. The child tried to retain gratifications derived from these object relationships, at least in fantasy, through the process of introjection. By this mechanism, qualities of the person who was the center of the gratifying relationship are internalized and reestablished as part of the organization of the self. In so doing, they retain their object-connectedness and carry a degree of coloration from motivational and defensive connections with the object. Freud referred to this internalized product as a *precipitate of abandoned object cathexis*. However, because of the degree of residual object-connection, introjects are not fully integrated in the structuring of ego and superego; they retain a

quasi-independence and inner presence. They have been aptly described in Kleinian terms as "internal objects."

Identification has often been confused with introjection, partially because Freud treated the two processes in an overlapping and somewhat interchangeable fashion. Many analysts simply use the terms interchangeably. There are, nonetheless, grounds for maintaining a distinction between them. Identification is, properly speaking, an active structuralizing process that takes place within the self, by which the self constructs the inner constituents of regulatory control on the basis of selected elements derived from the model. What constitutes the model of identification can vary considerably and can include introjects, structural aspects of real objects, or even regulatory components of group structures and group cultures. The process of identification is specifically an intrasystemic structuralizing activity, attributed to the ego functions of the self and related to its synthetic function, affecting structural integration in all parts of the psychic apparatus, including ego and superego. By identification the character of these structures is internally altered, such that their qualities can be regarded as authentically aspects of the self effectively modifying aspects of autonomous self-functioning.

FUNCTIONS OF THE EGO. The ego comprises a class of self-functions that shares in common the task of mediating between instincts and the outside world. Thus, the ego is a subsystem of the personality and is not synonymous with the self, the personality, or character. Any attempt to compile a complete list of ego functions would have to be relatively arbitrary. Invariably, the list of basic ego functions suggested by various authors differs in varying degrees. This discussion will be limited to several functions generally conceded to be fundamental ego functions.

Control and Regulation of Instinctual Drives. Development of the capacity to delay immediate discharge of urgent wishes and impulses is essential if the ego is to ensure the integrity of the individual and fulfill its role as mediator between id and outside world. Development of the capacity to delay or postpone instinctual discharge, like the capacity to test reality, is closely related to the progression in early childhood from pleasure principle to reality principle.

Relation to Reality. Freud always regarded the ego's capacity for maintaining relationship to the external world among its principal functions. Although the relation to the real is primarily a self-function, the connection with reality is mediated by ego functions. The character of its relationship to the external world may be divided into three components: (1) the sense of reality, (2) reality testing, and (3) adaptation to reality.

Sense of Reality. The sense of reality originates simultaneously with the development of the ego. Infants are responsive from the first to external

stimuli and become increasingly aware of the reality of their own bodily sensations as different from outside objects. Only gradually do they develop the capacity to distinguish a reality outside of their own bodies.

Reality Testing. *Reality testing* refers to the ego's capacity for objective evaluation and judgment of the external world, which depends first on primary autonomous functions of the ego, such as memory and perception, but then also on the relative integrity of internal structures of secondary autonomy. Under conditions of internal stress, in which regressive pulls are effectively operating, introjective aspects of inner psychic structure can tend to dominate and, thus, become susceptible to projective distortions that color the individual's perception and interpretation of the outside world. Because of the fundamental importance of reality testing for "negotiating" with the outside world, its impairment may be associated with severe mental disorder.

Adaptation to Reality. *Adaptation to reality* refers to the capacity of the self in virtue of its ego functions to use the individual's resources to form adequate solutions based on previously tested judgments of reality. It is possible by way of such adaptive ego functions for the self to develop not only good reality testing, with perception and grasp, but also to develop an adequate capacity to accommodate the individual's resources to the situation thus perceived. Adaptation is closely allied to the concept of mastery, both in respect to external tasks and to the instincts. It should be distinguished from adjustment, which may entail accommodation to reality at the expense of certain resources or potentialities of the individual. The function of adaptation to reality is closely related to the defensive functions of the ego. The mechanism that may serve defensive purposes from one point of view may simultaneously serve adaptive purposes when viewed from another perspective. Thus, in the obsessive-compulsive person, intellectualization may serve important inner needs to control drive impulses, but by the same token, the intellectual activity itself may serve highly adaptive functions in dealing with the complexities of external reality.

Object Relationships. The capacity for mutually satisfying relationships has been traditionally attributed to the ego, although self–other relationships are more properly a function of the whole person, the self, of which the ego is a functional component. Significance of object relationships and their disturbance—for normal psychological development and a variety of psychopathological states—was fully appreciated relatively late in the development of classical psychoanalysis. The evolution in the child's capacity for relationships with others, progressing from initial relations with maternal and other caretaking figures to social relationships within the family and then to relationships within the larger community, is related

to this capacity. Development of object relationships may be disturbed by retarded development, regression, or conceivably by inherent genetic defects or limitations in the capacity to develop object relationships, or impairments and deficiencies in early caretaking relationships.

Defensive Functions of the Ego. As was pointed out previously, in his initial psychoanalytic formulations, and for a long time thereafter, Freud considered repression to be virtually synonymous with defense. More specifically, repression was directed primarily against the impulses, drives, or drive representations and, particularly, against direct expression of the sexual instinct. Defense was thus mobilized to bring instinctual demands into conformity with demands of external reality. With development of the structural view of the mind, the function of defense was ascribed to the ego. Only after Freud had formulated his final theory of anxiety, however, was it possible to study the operation of the various defense mechanisms in light of their mobilization in response to danger signals. By that time he had established that defenses were one of the psyche's major devices for managing threatening instinctual motivations and affects, that they operated unconsciously, that while they were characteristic of neurotic syndromes they were dynamically motivated and reversible, and finally that they could be functionally adaptive as well as pathological.

Thus, Freud's efforts opened the way to a more systematic and comprehensive study of ego defenses, which was formulated for the first time by Anna Freud. In her classic monograph *The Ego and the Mechanisms of Defense,* she maintained that everyone, whether normal or neurotic, uses a characteristic repertoire of defense mechanisms, but to varying degrees. On the basis of her extensive clinical studies of children, she described their essential inability to tolerate excessive instinctual stimulation and discussed processes whereby the primacy of such drives at various developmental stages evoked anxiety in the ego. This anxiety, in turn, produced a variety of defenses. With regard to adults, her psychoanalytic investigations led her to conclude that although resistance was an obstacle to progress in treatment, to the extent that it impeded the emergence of unconscious material, it also constituted a useful source of information concerning the ego's defensive operations.

Genesis of Defense Mechanisms. In the early stages of development, defenses emerge as a result of difficulties in the capacity of ego functions to mediate pressures of the id and the requirements and strictures of outside reality. In the classic theory, at each phase of libidinal development, associated drive components evoke characteristic ego defenses. Thus, for example, introjection, denial, and projection are defense mechanisms associated with oral-incorporative or oral-sadistic impulses; whereas reaction formations, such as shame and disgust, usually develop in relation to anal impulses and pleasures. Defense mechanisms from earlier phases of

development persist side by side with those of later periods. When defenses associated with pregenital phases of development tend to predominate in adult life over more mature mechanisms, such as sublimation and repression, the personality retains an infantile cast.

Classification of Defenses. The defensive forms of ego functioning can be categorized in various ways, none of which is all-inclusive or takes into account all of the relevant factors. Defenses may be classified developmentally, that is, in terms of the libidinal phase in which they arise or with which they are associated. Thus, denial, projection, and distortion would be assigned to the oral stage of development and to the correlative narcissistic stage of object relationships. Certain defenses, however, such as magical thinking and regression, cannot be categorized in this way. Moreover, certain basic developmental processes, such as introjection and projection, may also serve defensive functions under certain specifiable conditions. The defenses have also been classified on the basis of the particular form of psychopathology with which they are commonly associated. Thus, the obsessional defenses would include isolation, rationalization, intellectualization, and denial; however, defensive operations are not limited to pathological conditions. Finally, the defenses have been classified as to whether they are simple mechanisms or complex, in which a single defense would involve a combination or composite of simple mechanisms. Table 6.1–2 gives a brief classification and description of some of the basic defense mechanisms most frequently employed and most thoroughly investigated by psychoanalysts.

Table 6.1–2.
Classification of Defense Mechanisms

Narcissistic-Psychotic Defenses	
These defenses are usually found as part of a psychotic process, but may also occur in young children and adult dreams or fantasies. They share the common note of avoiding, negating, or distorting reality.	
Projection	Perceiving and reacting to unacceptable inner impulses and their derivatives as though they were outside the self. On a psychotic level, this takes the form of frank delusions about external reality, usually persecutory, includes both perception of one's own feelings in another, with subsequent acting on the perception (psychotic paranoid delusions). Impulses may derive from id or superego (hallucinated recriminations).
Denial	Psychotic denial of external reality, unlike repression, affects perception of external reality more than perception of internal reality. Seeing, but refusing to acknowledge what one sees, or hearing, and negating what is actually heard, are examples of denial and exemplify the close relationship of denial to sensory experience. Not all denial, however, is necessarily psychotic. Like projection, denial may function in the service of more neurotic or even adaptive objectives. Denial avoids becoming aware of some painful aspect of reality. At the psychotic level, the denied reality may be replaced by a fantasy or delusion.

Distortion — Grossly reshaping the experience of external reality to suit inner needs, including unrealistic megalomanic beliefs, hallucinations, wish-fulfilling delusions, and employing sustained feelings of delusional grandiosity, superiority, or entitlement.

Immature Defenses

These mechanisms are fairly common in preadolescent years and in adult character disorders. They are often mobilized by anxieties related to intimacy or its loss. Although they are regarded as socially awkward and undesirable, they often moderate with improvement in interpersonal relationships or with increased personal maturity.

Acting out — The direct expression of an unconscious wish or impulse in action to avoid being conscious of the accompanying affect. The unconscious fantasy, involving objects, is lived out and impulsively enacted in behavior, thus gratifying the impulse more than the prohibition against it. On a chronic level, acting out involves giving in to impulses to avoid the tension that would result from postponement of their expression.

Blocking — An inhibition, usually temporary in nature, of affects especially, but possibly also thinking and impulses. It is close to repression in its effects, but has a component of tension arising from the inhibition of the impulse, affect, or thought.

Hypochondriasis — Transformation of reproach toward others arising from bereavement, loneliness, or unacceptable aggressive impulses, into self-reproach in the form of somatic complaints of pain, illness, and so forth. Real illness may also be overemphasized or exaggerated for its evasive and regressive possibilities. Thus, responsibility may be avoided, guilt may be circumvented, and instinctual impulses may be warded off.

Introjection — In addition to the developmental functions of the process of introjection, it also can serve specific defensive functions. The introjection of a loved object involves the internalization of characteristics of the object with the goal of ensuring closeness to and constant presence of the object. Anxiety consequent to separation or tension arising out of ambivalence toward the object is thus diminished. If the object is lost, introjection nullifies or negates the loss by taking on characteristics of the object, thus in a sense internally preserving the object. Even if the object is not lost, the internalization usually involves a shift of cathexis reflecting a significant alteration in the object relationship. Introjection of a feared object serves to avoid anxiety through internalizing the aggressive characteristic of the object, and thereby putting the aggression under one's own control. The aggression is no longer felt as coming from outside, but is taken within and utilized defensively, thus turning the subject's weak, passive position into an active, strong one. The classic example is "identification with the aggressor." Introjection can also take place out of a sense of guilt in which the self-punishing introject is attributable to the hostile-destructive component of an ambivalent tie to an object. Thus, the self-punitive qualities of the object are taken over and established within one's self as a symptom or character trait, which effectively represents both the destruction and the preservation of the object. This is also called identification with the victim.

Passive-aggressive behavior — Aggression toward an object expressed indirectly and ineffectively through passivity, masochism, and turning against the self.

Projection — On a nonpsychotic level, projection involves attributing one's own unacknowledged feelings to others; it includes severe prejudice, rejection of intimacy through suspiciousness, hypervigilance to external danger, and injustice collecting. Projection operates correlatively to introjection, such that the material of the projection derives from the internalized but usually unconscious configuration of the subject's introjects. At higher levels of function, projection may take the

form of misattributing or misinterpreting motives, attitudes, feelings, or intentions of others.

Regression — A return to a previous stage of development or functioning to avoid the anxieties or hostilities involved in later stages. A return to earlier points of fixation embodying modes of behavior previously given up. This is often the result of a disruption of equilibrium at a later phase of development. This reflects a basic tendency to achieve instinctual gratification or to escape instinctual tension by returning to earlier modes and levels of gratification when later and more differentiated modes fail or involve intolerable conflict.

Schizoid fantasy — The tendency to use fantasy and to indulge in autistic retreat for the purpose of conflict resolution and gratification.

Somatization — The defensive conversion of psychic derivatives into bodily symptoms; tendency to react with somatic rather than psychic manifestations. Infantile somatic responses are replaced by thought and affect during development (desomatization); regression to earlier somatic forms or response (resomatization) may result from unresolved conflicts and may play an important role in psychophysiological and psychosomatic reactions.

Neurotic Defenses

These are common in apparently normal and healthy individuals as well as in neurotic disorders. They function usually in the alleviation of distressing affects and may be expressed in neurotic forms of behavior. Depending on circumstances, they can also have an adaptive or socially acceptable aspect.

Controlling — The excessive attempt to manage or regulate events or objects in the environment in the interest of minimizing anxiety and solving internal conflicts.

Displacement — Involves a purposeful, unconscious shifting of impulses and/or affective investment from one object to another in the interest of solving a conflict. Although the object is changed, the instinctual nature of the impulse and its aim remain unchanged.

Dissociation — A temporary but drastic modification of character or sense of personal identity to avoid emotional distress; it includes fugue states and hysterical conversion reactions.

Externalization — A general term, correlative to internalization, referring to the tendency to perceive in the external world and in external objects components of one's own personality, including instinctual impulses, conflicts, moods, attitudes, and styles of thinking. It is a more general term than projection, which is defined by its derivation from and correlation with specific introjects.

Inhibition — The unconsciously determined limitation or renunciation of specific ego functions, singly or in combination, to avoid anxiety arising out of conflict with instinctual impulses, superego, or environmental forces or figures.

Intellectualization — The control of affects and impulses by way of thinking about them instead of experiencing them. It is a systematic excess of thinking, deprived of its affect, to defend against anxiety caused by unacceptable impulses.

Isolation — The intrapsychic splitting or separation of affect from content resulting in repression of either idea or affect or the displacement of affect to a different or substitute content.

Rationalization — A justification of attitudes, beliefs, or behavior that might otherwise be unacceptable by an incorrect application of justifying reasons or the invention of a convincing fallacy.

Reaction formation — The management of unacceptable impulses by permitting expression of the impulse in antithetical form. This is equivalently an expression of the impulse in the negative. Where instinctual conflict is persistent, reaction formation can become a character trait on a permanent basis, usually as an aspect of obsessional character.

Repression	Consists of the expelling and withholding from conscious awareness of an idea or feeling. It may operate either by excluding from awareness what was once experienced on a conscious level (secondary repression), or it may curb ideas and feelings before they have reached consciousness (primary repression). The "forgetting" associated with repression is unique in that it is often accompanied by highly symbolic behavior, which suggests that the repressed is not really forgotten. The important discrimination between repression and the more general concept of defense has been discussed.
Sexualization	The endowing of an object or function with sexual significance that it did not previously have, or possesses to a lesser degree, to ward off anxieties connected with prohibited impulses.
Mature Defenses	
These mechanisms are healthy and adaptive throughout the life cycle. They are socially adaptive and useful in the integration of personal needs and motives, social demands, and interpersonal relations. They can underlie seemingly admirable and virtuous patterns of behavior.	
Altruism	The vicarious but constructive and instinctually gratifying service to others, even to the detriment of the self. This must be distinguished from altruistic surrender, which involves a masochistic surrender of direct gratification or of instinctual needs in favor of fulfilling the needs of others to the detriment of the self, with vicarious satisfaction only being gained through introjection.
Anticipation	The realistic anticipation of or planning for future inner discomfort: Implies overly concerned planning, worrying, and anticipation of dire and dreadful possible outcomes.
Asceticism	The elimination of directly pleasurable affects attributable to an experience. The moral element is implicit in setting values on specific pleasures. Asceticism is directed against all "base" pleasures perceived consciously, and gratification is derived from the renunciation.
Humor	The overt expression of feelings without personal discomfort or immobilization and without unpleasant effect on others. Humor allows one to bear, and yet focus on, what is too terrible to be borne, in contrast to wit, which always involves distraction or displacement away from the affective issue.
Sublimation	The gratification of an impulse whose goal is retained, but whose aim or object is changed from a socially objectionable one to a socially valued one. Libidinal sublimation involves a desexualization of drive impulses and the placing of a value judgment that substitutes what is valued by the superego or society. Sublimation of aggressive impulses takes place through pleasurable games and sports. Unlike neurotic defenses, sublimation allows instincts to be channeled rather than to be dammed up or diverted. Thus, in sublimation, feelings are acknowledged, modified, and directed toward a relatively significant person or goal so that modest instinctual satisfaction results.
Suppression	The conscious or semiconscious decision to postpone attention to a conscious impulse or conflict.

Data from Vaillant GE. *Adaptation to Life*. Boston: Little Brown; 1977; Semrad E. The operation of ego defenses in object loss. In: Moriarity DM, ed. *The Loss of Loved Ones*. Springfield, IL: Charles C Thomas; 1967; Bibring GL, Dwyer TF, Huntington DS, Valenstein AA. A study of the psychological principles in pregnancy and of the earliest mother–child relationship: Methodological considerations. *Psychoanal Stud Child*. 1961;16:25.

Synthetic Function. The *synthetic function* of the ego refers to the self's capacity to integrate various aspects of its functioning. This function of the ego involves the capacity to unite, organize, and bind together various

drives, motives, tendencies, and functions within the personality, enabling the individual to think, feel, and act in an organized and directed manner. Briefly, the synthetic function is concerned with the overall organization and functioning of the ego in the self-system and consequently must enlist the cooperation of other ego and nonego functions in its operation. Although the synthetic function subserves adaptive functioning in the self, it may also bring together various forces in a way that, although not completely adaptive, is an optimal solution for the individual in a particular state at a given moment or period of time. Thus, the formation of a symptom that represents a compromise of opposing tendencies, although unpleasant in some degree, is nonetheless preferable to yielding to a dangerous instinctual impulse or, conversely, trying to stifle the impulse completely. Hysterical conversion, for example, combines a forbidden wish and the punishment for it into a physical symptom. On examination, the symptom often turns out to be the only possible compromise under the circumstances.

Autonomy of the Ego. Although Freud only referred to "primal, congenital ego variations" as early as 1937, this concept was greatly expanded and clarified by Hartmann. Hartmann advanced a basic formulation about development; that is, that the ego and id differentiate from a common matrix, the so-called undifferentiated phase, in which the ego's precursors are inborn apparatuses of primary autonomy. These apparatuses are rudimentary in nature, present at birth, and develop outside the area of conflict with the id. This area Hartmann referred to as a "conflict-free" area of ego functioning. He included perception, intuition, comprehension, thinking, language, certain phases of motor development, learning, and intelligence among the functions in this "conflict-free" sphere. Each of these functions, however, as analysts have become increasingly aware, might also become involved in conflict secondarily in the course of development and usually do. For example, if aggressive, competitive impulses intrude on the impulse to learn, they may evoke inhibitory defensive reactions on the part of the ego, thus interfering with the conflict-free operation of these functions.

Primary Autonomy. With the introduction of the primary autonomous functions, Hartmann provided an independent genetic derivation for at least part of the ego, thus establishing it as an independent realm of psychic organization and functioning that was not totally dependent on and derived from the instincts. This was an insight of major importance because it laid the foundations for the emerging doctrine of ego autonomy and meant that the analysis of ego development would have to consider an entirely new set of variables quite separate from those involved in instinctual development.

Secondary Autonomy. Hartmann observed that the conflict-free sphere

derived from structures of primary autonomy can be enlarged, that further functions could be withdrawn from the domination of drive influences. This was Hartmann's concept of secondary autonomy. Thus, a mechanism that arose originally in the service of defense against instinctual drives may in time become an independent structure, such that the drive impulse merely triggers the automatized apparatus. Thus, the apparatus may come to serve other functions than the original defensive function, for example, adaptation or synthesis. Hartmann referred to this removal of specific mechanisms from drive influences and becoming relatively autonomous as a process of change of function.

The Superego. The origins and functions of the superego are related to those of the ego, but they reflect different developmental vicissitudes. Briefly, the superego is the last of the structural components to develop, resulting in Freud's analysis from resolution of the oedipal complex. It is concerned with moral behavior based on unconscious behavioral patterns learned at early pregenital stages of development. Frequently, in Freud's view, superego functions become involved in neurotic conflict by imposing demands in the form of conscience or guilt feelings. Occasionally, however, the superego may be allied with id functions against the ego. This happens in cases of severely regressed reaction, where functions of the superego may become sexualized once more or may become permeated by aggression, taking on a quality of primitive (usually anal) destructiveness.

HISTORICAL DEVELOPMENT. In his 1896 paper *Further Remarks on the Neuropsychoses of Defence,* Freud described obsessional ideas as "*self-reproaches* which have reemerged from *repression* and which always relate to some *sexual* act that was performed with pleasure in *childhood*." The activity of a self-criticizing agency was also implicit in Freud's early discussions of dreams, which postulated existence of a "censor" that did not permit unacceptable ideas to enter consciousness on moral grounds. He first discussed the concept of a special self-critical agency in 1914, suggesting that a hypothetical state of narcissistic perfection existed in early childhood; at this stage, the child was his or her own ideal. As the child grew up, admonitions of others and self-criticism combined to destroy this perfect image. To compensate for this lost narcissism, or to recover it, the child "projects before him" a new ideal, or ego-ideal. It was at this point that Freud suggested that the psychic apparatus might have still another structural component, a special agency whose task it was to watch over the ego, to make sure it was measuring up to the ego-ideal. The concept of the superego evolved from these formulations of an ego-ideal and a second monitoring agency to ensure its preservation.

Again in 1917, in *Mourning and Melancholia,* Freud spoke of "one part of the ego" that "judges it critically and, as it were, takes it as its object." He

suggested that this agency, which is split off from the rest of the ego, was what is commonly called conscience. He further stated that this self-evaluating agency could act independently, could become "diseased" on its own account, and should be regarded as a major institution of the self. In 1921, Freud referred to this self-critical agency as the *ego-ideal* and held it responsible for the sense of guilt and for the self-reproaches typical in melancholia and depression. At that point he had dropped his earlier distinction between the ego-ideal, or ideal self, and a self-critical agency, or conscience.

In 1923, however, in *The Ego and the Id,* Freud's concept of the superego again included both these functions—that is, the superego included both the ego-ideal as well as the function of conscience. He also demonstrated that operations of the superego were mainly unconscious. Thus, patients who were dominated by a deep sense of guilt lacerated themselves far more harshly on an unconscious level than they did consciously. The fact that guilt engendered by the superego might be eased by suffering or punishment was apparent in the case of neurotics who demonstrated an unconscious need for punishment. In later works Freud elaborated on the relationship between ego and superego. Guilt feelings were ascribed to tension between these two agencies, and the need for punishment was an expression of this tension.

ORIGINS OF THE SUPEREGO. In Freud's view, the superego comes into being with resolution of the Oedipus complex. During the oedipal period, the little boy wishes to possess his mother, and the little girl wishes to possess her father. Each must, however, contend with a substantial rival, the parent of the same sex. The frustration of the child's positive oedipal wishes by this parent evokes intense hostility, which finds expression not only in overt antagonistic behavior but also in thoughts of killing the parent who stands in the way, along with any brothers or sisters who may also compete for the love of the desired parent.

Quite understandably, this hostility on the part of the child is unacceptable to parents and, in fact, eventually becomes unacceptable to the child as well. The theory contends that, in the case of the little boy, his sexual explorations and masturbatory activities may themselves meet with parental disfavor, which may even be underscored by real or implied threats of castration. These threats and, above all, the boy's observations that women and girls lack a penis convince him of the reality of castration. Consequently, he turns away from the oedipal situation and its emotional involvements and enters the latency period of psychosexual development. He renounces thereby the sexual impulses of the infantile phase.

Girls, when they become aware of the fact that they lack a penis (in Freud's terms they have "come off badly") were thought to seek to redeem the loss by obtaining a penis or a baby from the father. Freud pointed out

that although the anxiety surrounding castration brings the Oedipus complex to an end in boys, in girls it is the major precipitating factor. Girls renounce their oedipal strivings, first, because they fear the loss of the mother's love and, second, because of their disappointment over the father's failure to gratify their wish. The latency phase, however, is not as well defined in girls as it is in boys, and their persistent interest in family relations is expressed in their play; throughout grade school, for example, girls "act out" the roles of wife and mother in games that boys scrupulously avoid. This was the basic outline of Freud's theory of the superego.

EVOLUTION OF THE SUPEREGO. What, indeed, is the fate of the object attachments supposedly given up with resolution of the Oedipus complex? Freud's formulation of the mechanism of introjection came into play here. During the oral phase, the child is entirely dependent on the parents. Advancing beyond this stage, the child must abandon these earliest symbiotic ties with the parents and form initial introjections of them, which, however, follow the anaclitic model—that is, they are still characterized by dependence on the parents. Thus dissolution of the Oedipus complex and the concomitant abandonment of these object ties led to rapid acceleration of the introjection process.

These introjections from both parents became united and formed a kind of precipitate within the self, which then confronted other contents of the psyche and became organized as the superego. This internalization of the parents was based on the child's struggles to repress instinctual aims that were directed toward them, and it was this effort of renunciation that gave the superego its prohibiting character. It is for this reason, too, that the superego results to such a great extent from introjection of the parents' own superegos. Yet, because the superego evolved as a result of repression of instinctual desires, it had a closer relation to the id than did the ego itself. Its origins were more internal; the ego originated to a greater extent in relation to the external world and was its internal representative.

Finally, throughout the latency period and thereafter, the child (and later the adult) continued to build on these early identifications through contact with teachers, heroic figures, and admired persons, who formed the sources of the child's moral standards, values, and ultimate aspirations and ideals. The child moved into the latency period endowed with a superego that was, as Freud put it, "the heir to the Oedipus complex." The child's conflicts with the parents continued, of course, but now they were largely internal, between his or her own ego and superego. In other words, the standards, restrictions, commands, and punishments imposed previously by the parents from without were internalized in the child's superego, which now judges and guides behavior from within, even in the absence of the parents.

CURRENT INVESTIGATIONS OF THE SUPEREGO. Exploration of the superego and its functions did not end with Freud, and such studies remain of current active interest. Recent interest has focused on the differentiating between superego and ego-ideal, a distinction that Freud periodically revived and abandoned. At present, the term superego refers primarily to a self-critical, prohibiting function bearing a close relationship to aggression and aggressive introjections. The ego-ideal, however, is a kinder function, based on a transformation of the abandoned state of perfect infantile narcissism, or self-love, which existed in early childhood and has been integrated with positive elements of introjections from the parents. In addition, the concept of an ideal object—that is, the idealized object choice—has been advanced as distinct from the ideal self. Many theorists regard the ego-ideal as an aspect of superego organization derived from good parental imagoes.

A second focus of recent interest has been the contribution of the drives and object attachments formed in the preoedipal period to the development of the superego. These pregenital (especially anal) precursors of the superego are generally thought to provide some of the very rigid, strict, and aggressive qualities of the superego. These qualities stem from projection of the child's own sadistic drives and primitive concept of justice based on retaliation, which was attributed to the parents during this period. The harsh emphasis on absolute cleanliness and propriety that is sometimes found in very rigid individuals and in obsessional neurotics is based to some extent on this sphincter morality of the anal period. Other components have been traced back to factors operating in the oral phase. One result of these developments is that the connections between oedipal dynamics and superego development have been significantly diluted in the sense that preoedipal superego precursors and preoedipal superego-like functions are better understood on one hand, and postoedipal adaptive integrations, especially with ego functions, on the other, have modified the understanding of superego functions. The understanding of superego development and functioning has become much more complex than was envisioned by Freud. The case of conscience is one such area, insofar as conscience is effectively a judgment of good or evil that inevitably involves ego functions in integration with superego functions. Similarly, ethical values, in their formation and implementation, may represent important integrations of superego and ego functions.

PSYCHIC DEVELOPMENT: INTEGRATION OF PSYCHOSEXUAL PHASES AND OBJECT RELATIONS

As his clinical experience increased, Freud was able to reconstruct to a certain degree the early sexual experiences and fantasies of his patients. These data provided the framework for a developmental theory of childhood sexuality, which, in the subsequent course of psychoanalytic developmental exploration based on direct observation of childhood behavior, has been widely corroborated and accepted in some of its essential aspects, but also further elaborated by developmental theorists. As will be seen, Freud's early views have been subjected to considerable revision and development, as well as criticism and rejection, in ensuing years. Freud's view tended to combine states of infantile functioning with supposed stages of development, an approach largely abandoned by current developmental theorists. The idea that a temporal sequence of developmental stages is also a causal sequence has a certain appeal, but may not stand up to closer scrutiny. In current nonlinear dynamic systems theory, the concept of a continuous and progressive sequence of stages, as Freud's psychosexual theory proposed, has given way to an emphasis on changes in discontinuous states or phases brought about by both internal and external conditions. Although the Freudian stages may retain a certain historical interest and descriptive validity, the perspective of stage sequencing requires further qualification.

Perhaps an even more important source of information that contributed to Freud's thinking about infantile sexuality was his own self-analysis that began in 1897. He was gradually able to recover memories of his own erotic longings in childhood and his conflicts in relationship to his parents, related specifically to his oedipal involvement. Realization of the operation of such infantile sexual longings in his own experience suggested to Freud that these phenomena might not be restricted only to the pathological development of neuroses, but that essentially normal individuals might undergo similar developmental experiences. The progressive integration of psychosexual developments and object relations has been further elaborated in Freud's phases of instinctual development, Margaret Mahler's separation-individuation process, and Erik Erikson's epigenetic sequence.

Phases of Psychosexual Development

The earliest manifestations of infantile sexuality arose in relation to bodily functions that had been regarded as basically nonsexual, such as feeding and development of bowel and bladder control. But Freud saw that these functions involved degrees of sensual pleasure, which he interpreted as forms of psychosexual stimulation, and divided them into a succession of developmental phases, each of which was thought to build on and subsume

accomplishments of the preceding phases—namely the oral, anal, and phallic phases. The oral phase occupied the first 12 to 18 months of the infant's life; next came the anal phase, lasting until about 3 years of age; and, finally, the phallic phase, from approximately 3 to 5 years of age. Urethral, latency, and genital phases were added to complete the picture.

Freud postulated that in boys, phallic erotic activity was essentially a preliminary stage for adult genital activity. In contrast to the male, whose principal sexual organ remained the penis throughout the course of psychosexual development, the female had two leading erotogenic zones, the clitoris and the vagina. Freud felt that the clitoris was preeminent during the infantile pregenital period but that after puberty erotic primacy was transposed to the vagina. Recent sexual investigations have cast some doubt on a supposed transition from clitoral to vaginal primacy, but many analysts still retain this view. The question remains a matter of debate for the time being and remains unresolved.

Freud's basic schema of the psychosexual stages was modified and refined by Karl Abraham, who further subdivided the phases of libido development, dividing the oral period into a sucking and biting phase, and the anal phase into a destructive-expulsive (anal sadistic) and a mastering-retaining (anal erotic) phase. Finally, he hypothesized that the phallic period consisted of an earlier phase of partial genital love, which was designated as the true phallic phase, and a later, more mature genital phase. For each of the stages of psychosexual development, Freud delineated specific erotogenic zones that gave rise to erotic gratification. Table 6.1–3 provides an overview of traditional, and currently more or less tentative and questioned, views on psychosexual development. Current theories, largely resulting from direct empirical and experimental observations of children in child analyses and developmental studies rather than merely relying on the reconstruction of childhood experiences based on the data from adult analyses, have tended to focus less on libidinal phase specificity, with the further supposition of programmatic progression of libidinal stages, progressing through the sequence of stages from oral to genital in prescribed order, and place greater emphasis on the complex integration of multiple developmental influences, including maturational factors, temperamental dispositions, object relations involvements and vicissitudes, affective development, cognitive development, language acquisition, and so on. There is accordingly a greater inclination to view libidinal stages as more loosely organized, intermingled, and not necessarily rigidly sequential.

Development and Object Relations

Current theories in psychoanalytic psychiatry have focused increasingly on the importance for later psychopathology of early disturbances in object relationships—that is, disturbances in the relationship between the child's

affect and the significant objects in the environment, particularly the mothering object. From the very beginning of the child's development, Freud regarded the sexual instinct as "anaclitic," in the sense that the child's attachment to the feeding and mothering figure is based on the child's utter physiological dependence on the object. This view of the child's earliest attachment would seem consistent with Freud's understanding of infantile libido based on his discovery that sexual fantasies of even adult patients were typically centered on early relationships with their parents. In any event, throughout his descriptions of libidinal phases of development, Freud made constant reference to the significance of children's relationships with crucial figures in their environment. Specifically, he postulated that the choice of a love object in adult life, the love relationship itself, and object relationships in other spheres of interest and activity were dependent to an important degree on the nature and quality of the child's object relationships during the earliest years of life.

Object Relations during Pregenital Phases. At birth, the infant's responses to external stimulation are relatively diffuse and disorganized. Even so, as recent experimental research on neonates has indicated, the infant is quite responsive to external stimulation, and the patterns of response are quite complex and relatively well organized, even shortly after birth. Even neonates of a few hours of age will respond selectively to novel stimuli and will demonstrate remarkable preferences for complex as compared to simple patterns of stimulation. The infant's responses to noxious and pleasurable stimuli are also relatively undifferentiated. Even so, sensations of hunger, cold, and pain give rise to tension and a corresponding need to seek relief from painful stimuli. At the beginning of life, however, the infant does not respond specifically to objects as objects. A certain degree of development of perceptual and cognitive apparatuses is required, as well as a greater degree of differentiation of sensory impressions and integration of cognitive patterns, before babies are able to differentiate between impressions belonging to themselves and those derived from external objects. Consequently, observations and inferences based on data derived from the first 6 months of life must be interpreted in the context of the child's cognitive functioning before self-object differentiation.

In these first months of life, human infants are considerably more helpless than any other young mammals. Their helplessness will continue for a longer period of time than for any other species. They cannot survive unless they are cared for, and they cannot achieve relief from the painful disequilibrium of inner physiological states without help of external caretaking objects. Object relationships of the most primitive kind only begin to be established when an infant first begins to grasp this fact of experience. In the beginning, an infant cannot distinguish between its own

lips and its mother's breast, nor does an infant initially associate satiation of painful hunger pangs with presentation of the extrinsic breast. Because the infant is aware only of his or her own inner tension and relaxation and is unaware of the external object, longing for the object exists only to the degree that the disturbing stimuli persist and longing for satiation remains unsatisfied in the absence of the object. When the satisfying object finally appears and the infant's needs are gratified, longing also disappears. Gradually, but also rather quickly, the infant becomes aware of the mother herself, in addition to her breast, as a need-satisfying object. It can be said that the infant is object related from the beginning of life, but that the capacity for relating to objects as such requires further development.

ORAL PHASE AND OBJECTS. This experience of unsatisfied need, together with the experience of frustration in the absence of the breast and need-satisfying release of tension in the presence of the breast, forms the basis of the infant's first awareness of external objects. Libido theory envisioned these patterns of response as driven by the need to discharge tension by seeking oral satisfaction, but later theorists have emphasized the importance of the relation to the mother and the inherent need of the infant to engage with and relate to objects, primarily the mother. This first awareness of an object, then, in the psychological sense, comes from longing for something that is already familiar, for something that actually gratified needs in the past but is not immediately available in the present. Thus, it is basically the infant's hunger in this view that in the beginning compels recognition of the outside world, but this may be superseded by the basic need for human contact with or without hunger. The first primitive reflex reaction to objects, putting them into the mouth, then becomes understandable. This reaction is consistent with the modality of the infant's first recognition of reality, judging reality by oral gratification, that is, whether something will provide relaxation of inner tension and satisfaction (and should thereby be incorporated, swallowed) or whether it will create inner tension and dissatisfaction (and consequently should be spit out).

Early in this interaction the mother serves the important function of empathically responding to the infant's inner needs in such a manner as to become involved in a process of mutual regulation, which maintains the homeostatic balance of the infant's physiological needs and processes within tolerable limits. Not only does this process keep the child alive, but it sets a rudimentary pattern of experience within which the child can build elements of a basic trust that promotes reliance on the benevolence and availability of caretaking objects. Consequently, the mother's administrations and responsiveness to the child help to lay the most rudimentary and essential foundation for subsequent development of object relations and the capacity for entering the community of human

beings.

Table 6.1–3.
Stages of Psychosexual Development

Oral Stage	
Definition	Earliest stage of development in which the infant's needs, perceptions, and modes of expression are primarily centered in mouth, lips, tongue, and other organs related to oral zone and around the sucking reflex.
Description	Oral zone maintains dominance in psychic organization through approximately first 18 months of life. Oral sensations include thirst, hunger, pleasurable tactile stimulations evoked by the nipple or its substitute, sensations related to swallowing and satiation. Oral drives consist of two components: Libidinal and aggressive. States of oral tension lead to seeking for oral gratification, as in quiescence at the end of nursing. Oral triad consists of wishes to eat, sleep, and reach that relaxation that occurs at the end of sucking just before the onset of sleep. Libidinal needs (oral erotism) predominate in early oral phase, whereas they are mixed with more aggressive components later (oral sadism). Oral aggression is expressed in biting, chewing, spitting, or crying. Oral aggression is connected with primitive wishes and fantasies of biting, devouring, and destroying.
Objectives	To establish a trusting dependence on nursing and sustaining objects, establish comfortable expression and gratification of oral libidinal needs without excessive conflict or ambivalence from oral sadistic wishes.
Pathological traits	Excessive oral gratifications or deprivation can result in libidinal fixations contributing to pathological traits. Such traits can include excessive optimism, narcissism, pessimism (as in depressive states), or demandingness. Envy and jealousy are often associated with oral traits.
Character traits	Successful resolution of the oral phase results in capacities to give to and receive from others without excessive dependence or envy, capacity to rely on others with a sense of trust as well as with a sense of self-reliance and self-trust. Oral characters are often excessively dependent and require others to give to them and look after them, and are often extremely dependent on others for maintaining self-esteem. These are readily amalgamated with narcissistic needs.
Anal Stage	
Definition	The stage of psychosexual development promoted by maturation of neuromuscular control over sphincters, particularly the anal sphincter, permitting greater voluntary control over retention or expulsion of feces.
Description	This period extends roughly from 1 to 3 years of age, marked by recognizable intensification of aggressive drives mixed with libidinal components in sadistic impulses. Acquisition of voluntary sphincter control is associated with an increasing shift from passivity to activity. Conflicts over anal control and struggles with parents over retaining or expelling feces in toilet training give rise to increased ambivalence, together with conflicts over separation, individuation, and independence. Anal erotism refers to sexual pleasure in anal functioning, both in retaining precious feces and presenting them as a precious gift to the parent. Anal sadism refers to expression of aggressive wishes connected with discharging feces as powerful and destructive weapons. These wishes are often displayed in fantasies of bombing or explosions.
Objectives	The anal period is marked by greater striving for independence and separation from dependence on and control of parents. Objectives of sphincter control without overcontrol (fecal retention) or loss of control (messing) are matched by attempts to achieve autonomy and independence without excessive shame or self-doubt

from loss of control.

Pathological traits	Maladaptive character traits, often apparently inconsistent, derive from anal erotism and defenses against it. Orderliness, obstinacy, stubbornness, willfulness, frugality, and parsimony are features of anal character. When defenses against anal traits are less effective, anal character reveals traits of heightened ambivalence, lack of tidiness, messiness, defiance, rage, and sadomasochistic tendencies. Anal characteristics and defenses are typically seen in obsessive-compulsive neuroses.
Character traits	Successful resolution of the anal phase provides the basis for development of personal autonomy, a capacity for independence and personal initiative without guilt, a capacity for self-determining behavior without a sense of shame or self-doubt, a lack of ambivalence, and a capacity for willing cooperation without either excessive willfulness or self-diminution or defeat.

Urethral Stage

Definition	This stage was not explicitly treated by Freud but serves as a transitional stage between anal and phallic stages. It shares some characteristics of anal phase and some from subsequent phallic phase.
Description	Characteristics of the urethral phase are often subsumed under phallic phase. Urethral erotism, however, refers to pleasure in urination as well as pleasure in urethral retention analogous to anal retention. Similar issues of performance and control are related to urethral functioning. Urethral functioning may also have sadistic quality, often reflecting persistence of anal sadistic urges. Loss of urethral control, as in enuresis, may frequently have regressive significance that reactivates anal conflicts.
Objectives	At stake are issues of control and urethral performance and loss of control. It is not clear whether or to what extent objectives of urethral functioning differ from those of anal period, except that they are expressed in a later developmental stage.
Pathological traits	The predominant urethral trait is competitiveness and ambition, probably related to compensation for shame due to loss of urethral control. This may instigate development of penis envy, related to feminine sense of shame and inadequacy in being unable to match male urethral performance. This may also be related to issues of control and shaming.
Character traits	Besides healthy effects analogous to those from the anal period, urethral competence provides a sense of pride and self-competence based on performance. Urethral performance is an area in which the small boy can imitate and try to match his father's more adult performance. Resolution of urethral conflicts sets the stage for budding gender identity and subsequent identifications.

Phallic Stage

Definition	Phallic stage begins sometime during third year and continues until approximately end of fifth year.
Description	The phallic phase is characterized by a primary focusing of sexual interests, stimulation, and excitement in the genital area. The penis becomes the organ of principal interest to children of both sexes, with lack of penis in females being considered as evidence of castration. The phallic phase is associated with an increase in genital masturbation accompanied by predominantly unconscious fantasies of sexual involvement with the opposite-sex parent. Threats of castration and the related anxiety are connected with guilt over masturbation and oedipal wishes. During this phase oedipal involvement and conflict are established and consolidated.
Objectives	To focus erotic interest in genital area and genital functions. This lays the foundation for gender identity and serves to integrate residues of previous stages into a predominantly genital-sexual orientation. Establishing the oedipal situation is essential for furtherance of subsequent identifications serving as a basis for important and perduring dimensions of character organization.

Pathological traits	Derivation of pathological traits from phallic-oedipal involvement is sufficiently complex and subject to such a variety of modifications that it encompasses nearly the whole of neurotic development. Issues, however, focus on castration in males and penis envy in females. Patterns of internalization developed from resolution of the Oedipus complex provide another important focus of developmental distortions. The influence of castration anxiety and penis envy, defenses against them, and patterns of identification are primary determinants of the development of human character. They also subsume and integrate residues of previous psychosexual stages, so that fixations or conflicts deriving from preceding stages can contaminate and modify oedipal resolution.
Character traits	The phallic stage provides the foundations for an emerging sense of sexual identity, of a sense of curiosity without embarrassment, of initiative without guilt, as well as a sense of mastery not only over objects and persons in the environment but also over internal processes and impulses. Resolution of the oedipal conflict gives rise to internal structural capacities for regulation of drive impulses and their direction to constructive ends. The internal sources of such regulation are the ego and superego, based on introjections and identifications derived primarily from parental figures.

Latency Stage

Definition	This is the stage of relative instinctual quiescence or inactivity of sexual drive during the period from the resolution of the Oedipus complex until pubescence (from about 5–6 years until about 11–13 years).
Description	The institution of the superego at the close of the oedipal period and further maturation of ego functions allow for considerably greater degrees of control of instinctual impulses and motives. Sexual interests are generally thought to be quiescent. This is a period of primarily homosexual affiliations for both boys and girls, as well as a sublimation of libidinal and aggressive energies into energetic learning and play activities, exploring the environment, and becoming more proficient in dealing with the world of things and persons around them. It is a period for development of important skills. The relative strength of regulatory elements often gives rise to patterns of behavior that are somewhat obsessive and hypercontrolling.
Objectives	The primary objective is further integration of oedipal identifications and consolidation of gender and sex-role identity. Relative quiescence and control of instinctual impulses allow for development of ego apparatuses and mastery of skills. Further identificatory components may be added to the oedipal ones on the basis of broadening contacts with other significant figures outside the family, e.g., teachers, coaches, and other adult figures.
Pathological traits	Dangers in the latency period can arise either from the lack of development of inner controls or an excess of them. Lack of control can lead to failure to sufficiently sublimate energies in the interest of learning and the development of skills; an excess of inner control, however, can lead to premature closure of personality development and precocious elaboration of obsessive character traits.
Character traits	The latency period is frequently regarded as a period of relatively unimportant inactivity in the developmental schema. More recently, greater respect has been gained for the developmental processes in this period. Important consolidations and additions are made to basic postoedipal identifications and to processes of integrating and consolidating previous attainments in psychosexual development and establishing decisive patterns of adaptive functioning. The child can develop a sense of industry and capacity for mastery of objects and concepts that allows autonomous functioning and a sense of initiative without risk of failure or defeat or a sense of inferiority. These are all important attainments that need to be further integrated, ultimately as the essential basis for a mature adult life of satisfaction in work and love.

Genital Stage	
Definition	The genital or adolescent phase extends from the onset of puberty from approximately ages 11–13 until young adulthood. Current thinking tends to subdivide this stage into preadolescent, early adolescent, middle adolescent, late adolescent, and even postadolescent periods.
Description	Physiological maturation of systems of genital (sexual) functioning and attendant hormonal systems leads to intensification of instinctual, particularly libidinal, drives. This produces a regression in personality organization, which reopens conflicts of previous stages of psychosexual development and provides opportunity for re-resolution of these conflicts in the context of achieving a mature sexual and adult identity. This period has been described as a "second individuation."
Objectives	Primary objectives are the ultimate separation from dependence on and attachment to parents and establishment of mature, nonincestuous, heterosexual object relations. Related are the achievement of a mature sense of personal identity and acceptance and integration of adult roles and functions that permit new adaptive integrations with social expectations and cultural values.
Pathological traits	Pathological deviations due to failure to achieve successful resolution of this stage of development are multiple and complex. Defects can arise from a whole spectrum of psychosexual residues, since the developmental task of adolescence is in a sense a partial reopening and reworking and reintegrating of all of these aspects of development. Previous unsuccessful resolutions and fixations in various phases or aspects of psychosexual development will produce pathological defects in the emerging adult personality and defects in identity formation.
Character traits	Successful resolution and reintegration of previous psychosexual stages in the adolescent genital phase set the stage normally for a fully mature personality with the capacity for full and satisfying genital potency and a self-integrated and consistent sense of identity. This provides the basis for a capacity for self-realization and meaningful participation in areas of work, love, and in creative and productive application to satisfying and meaningful goals and values.

As differentiation between the limits of self and object is gradually established in the child's experience, the mother becomes acknowledged and recognized as the source of gratifying nourishment and, in addition, as the source of the erotogenic pleasure the infant derives from sucking on the breast. In this sense she becomes the first love object. The quality of the child's attachment to this primary object is of the utmost importance, as developmental and attachment theorists have demonstrated. From the oral phase onward, the whole progression in psychosexual development, with its focus on successive erotogenic zones and emergence of associated component instincts, reflects the quality of the child's attachment to the crucial figures in the environment, as well as the strength of feelings of love or hate, or both, toward these important persons. If a fundamentally warm, trusting, secure, and affectionate relationship has been established between mother and child during the earliest stages of the child's career, then at least theoretically, the stage will be set for development of trusting and affectionate relationships with other human objects during the course of life. Attachment theory has in some degree confirmed this perspective.

ANAL PHASE AND OBJECTS. During the oral phase, the infant's role is not

altogether passive since, caught up as it is in a process of mutual interaction, the infant makes its own contribution to eliciting certain responses from the mother. The activity, however, is more or less automatic and dependent on such physiological factors as level of activity, irritability, or responsiveness to stimuli, but in addition the infant has an active role in engaging and interacting with the mothering figure. Generally speaking, however, the infant's control over the mother's feeding responses is relatively limited. Consequently, the primary onus remains on the mother to gratify or frustrate the demands of the infant.

In the transition to the anal period, however, this picture changes significantly. The child acquires a greater degree of control over behavior and in the classic view particularly over sphincter function. Moreover, for the first time during this period demands are placed on the child to relinquish some aspect of direct gratification by reason of expectations to accede to parental demands for use of the toilet and for regulation of bowel and bladder functions. This conflicts, however, with the primary libidinal aim of anal eroticism in enjoyment of the pleasurable sensations of excretion. Nonetheless, at this stage of development the demand is placed on the child to regulate gratification, to surrender some portion of the gratification at the parent's wish, or to delay gratification according to a schedule established by the parent's dictates. One of the more salient aspects of the anal period, therefore, is that it sets the stage for a contest of wills between parents and child over when, how, and on what terms the child will be allowed such gratification.

PHALLIC PHASE AND OBJECTS. The passage from the anal to phallic phase marks not only the transition from preoedipal to the beginnings of oedipal development, but also brings to a close the process of separation-individuation and, in the normal course of development, leads to the achievement of object constancy. The oedipal situation evolves during the period extending from the third to the fifth years in children of both sexes.

The Oedipus Complex. In the normal course of development, the so-called pregenital phases were regarded as primarily autoerotic. Primary gratification derived from stimulation of erotogenic zones, while the object served a significant, although secondary and instrumental, role. A fundamental shift begins to take place in the phallic phase in which the phallus becomes the primary erogenous zone for both sexes, thus laying a foundation for and initiating a shift of libidinal motivation and intention in the direction of objects. The phallic phase sets the stage for the fundamental striving to relate to a libidinal (sexual) object, a dynamic that advances the progression in establishing love relations within the oedipal context, and beyond that to more mature adult object choices and love relationships in the genital period. The phallic period is also a critical phase for the consolidation of the child's own sense of gender identity—as

decisively male or female—based in part on the child's discovery and realization of the significance of anatomical sexual differences. Freud also regarded the events associated with the phallic phase, particularly the oedipal situation, as setting the stage for the developmental predispositions to later psychoneuroses. Freud used the term *Oedipus complex* to refer to the intense love relationships, together with their associated rivalries, hostilities, and emerging identifications, formed during this period between child and parents.

Castration Complex. Freud postulated certain differentiations between the sexes in the pattern of phallic development. He explained the nature of this discrepancy in terms of genital differences. Under normal circumstances, he felt, for boys the oedipal situation was resolved by the castration complex. Specifically, the boy had to give up his strivings for his mother because of the threat of castration, resulting in castration anxiety. In contrast, the Oedipus complex in girls was evoked by castration anxiety, but unlike the boy, the little girl had already been castrated and had to seek compensation for her loss by turning to her father as bearer of the penis, out of a sense of disappointment over her own lack of a penis. Some have suggested that the little girl may thus be more threatened by a loss of love than by actual castration fears.

The Boy's Situation. In boys, development of object relations may be relatively less complex than for girls, because he remains attached to his first love object, the mother. The primitive object choice of the primary love object, which develops in response to the mother's gratification of the infant's basic physical and emotional needs, takes the same direction as the pattern of object choice in relation to opposite-sex objects in later life experience. In the phallic period, the boy develops a strong erotic interest in her and a concomitant desire to possess her exclusively and sexually. These feelings usually become manifest at approximately 3 years of age and reach a climax at 4 or 5 years of age.

With appearance of the oedipal involvement, the boy begins to show his loving attachment to his mother almost as a little lover might—wanting to touch her, trying to get in bed with her, proposing marriage, expressing wishes to replace his father, and devising opportunities to see her naked or undressed. Competition from siblings for the mother's affection and attention is intolerable. Above all, however, the little lover wants to eliminate his arch rival—the mother's husband. His wishes may involve not merely displacing or superseding the father in the mother's affection but eliminating him altogether. The child understandably anticipates retaliation for his aggressive wishes toward his father, and these expectations in turn give rise to severe anxiety in the form of the castration complex.

This somewhat simplified picture of the evolution of the Oedipus

complex is considerably more complex in the actual course of development. Usually the boy's love for his mother remains a dominant force during the period of infantile sexual development. It is known, however, that love is not free of some admixture of hostility and that the child's relationship with both parents is to some degree ambivalent. The boy also loves his father, and at times when he has been frustrated by his mother, he may hate her and turn from her to seek affection from his father. Undoubtedly, to some degree he both loves and hates both his parents at the same time. In addition, Freud's postulation of an essentially bisexual basis of the nature of the libido complicates matters further. On the one hand, the boy wants to possess his mother and eliminate the hated father rival. On the other hand, he also loves his father and seeks approval and affection from him, whereas he often reacts to his mother with hostility, particularly when her demands on her husband interfere with the exclusiveness of the father–son relationship. The *negative Oedipus complex* refers to those situations where the boy's love for his father predominates over the love of the mother, and the mother is relatively hated as a disturbing element in this relationship.

The Girl's Situation. Understanding of the little girl's more complex oedipal involvement was a later development. Because it could not be regarded as equivalent to the boy's development, it raised a number of questions that proved to be more difficult. Freud could not get beyond viewing female sexual development as a variant of male development, but later elaborations of female development have altered that picture dramatically. Similar to the little boy, in Freud's view, the little girl forms an initial attachment to the mother as a primary love object and source of fulfillment for vital needs. For the little boy, the mother normally remains the primary love object throughout his development, but, in contrast, the little girl is faced with the task of shifting this primary attachment from the mother to the father to prepare herself for her future sexual role. Freud was basically concerned with elucidating the factors that influenced the little girl to give up her preoedipal attachment to the mother and to form the normal oedipal attachment to the father. A secondary question had to do with the factors that led to the dissolution and resolution of the Oedipus complex in the girl, so that paternal attachment and maternal identification would result as the basis for adult sexual adjustment.

The girl's renunciation of her preoedipal attachment to the mother could not be satisfactorily explained as resulting from ambivalent or aggressive characteristics of the mother–child relationship, for similar elements would influence the relationship between boys and the mother figure. Freud attributed the crucial precipitating factor to anatomical differences between the sexes—specifically the girl's discovery of her lack of a penis during the phallic period. Up to this point, exclusive of

constitutional differences and depending on variations in parental attitudes in relating to a daughter in comparison to a son, the little girl's development was thought to parallel that of the little boy.

Fundamental differences, however, emerge when she discovers during the phallic period that her clitoris is inferior to the male counterpart, the penis. The typical reaction of the little girl to this discovery is an intense sense of loss, narcissistic injury, and envy of the male penis. At this point the little girl's attitude toward the mother changes. The mother had previously been the object of love, but now she is held responsible for bringing the little girl into the world with inferior genital equipment. The hostility can be so intense that it may persist and color her future relationship to the mother. With the further discovery that the mother also lacks the vital penis, the child's hatred and devaluation of the mother becomes even more profound. In a desperate attempt to compensate for her "inadequacy," the little girl then turns to her father in the vain hope that he will give her a penis, or a baby in place of the missing penis.

Obviously, the Freudian model of feminine psychosexual development has undergone, and is still currently undergoing, considerable revision. The charge has been made, and justifiably so, that masculine phallic-oedipal development was the primary model in Freud's thinking, and that feminine development was viewed as defective by comparison. Freud saw women typically as basically masochistic, weak, dependent, and lacking in conviction, strength of character, and moral fiber. He thought these defects were the result of failure in the oedipal identification with the phallic father because of female castration. The resulting internalization of aggression was both constitutionally determined and culturally reinforced.

These concepts must now be regarded as obsolete. Freud's hypotheses of a passive female libido, arrest in ego development, incapacity for sublimation, and superego deficiencies in women are outdated and inadequate. Differences in male and female ego and superego development may be defined, but there are no grounds for judging one to be superior or inferior to the other. They are simply different. In his 1976 article on *Masochism, the Ego Ideal and the Psychology of Women,* Harold Blum observed: "Female development cannot be described in a simple reductionism and overgeneralization. Femininity cannot be predominantly derived from a primary masculinity, disappointed maleness, masochistic resignation to fantasied inferiority, or compensation for fantasied castration and narcissistic injury. Castration reactions and penis envy contribute to feminine character, but penis envy is not the major determinant of femininity."

The adequate conceptualization and understanding of feminine psychology and its development are still very much in progress. There is much that is poorly understood and much more that is hardly understood at all. Current research has given partial support to and convincing

refutation of Freud's ideas. Current views emphasize the role of primary femininity and conflicts in identification with the mother as determining the course of development of feminine gender identity, rather than the outmoded views of castration anxiety and penis envy. The view of female development as following an independent and characteristic path of its own, and not one based on or derivative from or reactive against male development, has been increasingly consolidated and confirmed by contemporary feminist theorists. It has become increasingly clear, in all of this, that Freud was simply wrong about much of this whole area, but much of what he described may have simply expressed what he was able to observe in the women of his time and reflected the influence of attitudes toward women in his society and culture. Times change, however, and the culture and the place of women in it have changed and are still changing. To that extent, women are different, and much of their psychology is different too. Psychoanalytic understanding must inevitably lag behind these changing patterns of psychological experience, but a new and revised understanding of feminine development and functioning is now extant.

Mahler's Separation-Individuation Process

Autistic Phase. The separation-individuation process advanced by Margaret Mahler and her associates was a major empirically based contribution to understanding of the developmental process after Freud. Her theory, organized in terms of phases of separation and individuation, has come under severe criticism by contemporary developmentalists. The Mahler theory emphasizes the process of separation from the maternal orbit and the establishment of personal autonomy. In contrast, developmentalists, largely following the inspiration of self psychology and intersubjectivism, put the accent on maternal dependence in development and the continuing dependent reliance on selfobjects throughout the life cycle. These divergent emphases, however, may not be exclusive insofar as mature autonomy does not rule out meaningful dependence in object relations, nor does mature dependence exclude the possibility of meaningful autonomy.

The first phase of Mahler's theory of development describes the autistic phase: "During the first few weeks of extrauterine life, a stage of absolute primary narcissism, marked by the infant's lack of awareness of a mothering agent, prevails. This is the stage termed normal autism. It is followed by a stage of dim awareness that need satisfaction cannot be provided by oneself, but comes from somewhere outside the self." Choice of the term "autistic" is unfortunate and has been strongly criticized for using a pathological term to describe a normal developmental stage. In addition, Mahler's account of this earliest phase of development has been challenged by subsequent developmental studies. Nonetheless, the theory also articulates the origin of the initial differentiation of self and object, in

which infants can be said to experience something outside of themselves, to which they can relate, as satisfying their inner needs. This dawning awareness of the external object is a most significant state in the psychological development of children and involves not only cognitive and perceptual developments but also goes hand-in-hand with the organization of rudimentary infantile drives and affects in relation to emerging object experiences. This first awareness of the need-satisfying object relationship occurs during the oral phase of libidinal development, but the oral phase and the development of need-satisfying relationships are not equivalent. The oral phase is primarily concerned with libidinal development and stresses predominance of the oral zone as the main erotogenic zone. The concept of need-satisfying relationship, however, is not concerned directly with issues of drive development but, rather, with the characteristics of object involvement and object relationship. There is a shift here from drive to relational theory.

Symbiotic Phase. Mahler indicates that this awareness signals the beginning of normal symbiosis "in which the infant behaves and functions as though he and his mother were an omnipotent system—a dual unity within one common boundary." Boundaries become temporarily differentiated only in the state of "affect hunger" but disappear again when the need is gratified. Consequently, the object is recognized as separate from the self only at moments of need, so that, once the need is satisfied, the object ceases to exist—from the infant's (subjective) point of view—until a need again arises. Thus, it is useful to distinguish between need satisfaction as a stage of development in object relationships and need satisfaction as a determinant in object relationships at every level of development. The satisfaction of various kinds of psychological needs continues to play a role at all levels of object relatedness, but the satisfaction of such needs cannot be used as a distinguishing characteristic of the specific stage of need-satisfying object relationships. As objects become increasingly differentiated in the child's experience, their representations achieve increasing psychological complexity and value in a context of increasingly complex and subtle needs for a variety of input from objects. Development of object constancy implies a constant relationship to a specific object, but within that relationship the wish for satisfaction of needs and the actual satisfaction of those needs may still be a significant component of the object relationship.

Separation-Individuation. The separation-individuation process is divided into successive phases or periods—the hatching period, the practicing period, rapprochement, and the development of object constancy—describing the gradual separation from maternal dependency and the increasing autonomy of the child.

HATCHING. During the hatching period, the child with effort gradually differentiates out of the symbiotic matrix. The first behavioral signs of such differentiation seem to arise at about 4 or 5 months of age, at the high point of the symbiotic period. The first stage of this process of differentiation is described in the 1975 book *The Psychological Birth of the Human Infant* as "hatching" from the symbiotic orbit:

> In other words, the infant's attention, which during the first months of symbiosis was in large part inwardly directed, or focused in a coenesthetic vague way within the symbiotic orbit, gradually expands through the coming into being of outwardly directed perceptual activity during the child's increasing periods of wakefulness. This is a change of degree rather than of kind, for during the symbiotic stage the child has certainly been highly attentive to the mothering figure. But gradually that attention is combined with a growing store of memories of mother's comings and goings, of "good" and "bad" experiences; the latter were altogether unrelievable by the self, but could be "confidently expected" to be relieved by mother's ministrations.

PRACTICING. As "hatching" and separation from the mother gradually increases, there is a move to the second or practicing subphase of separation-individuation. The practicing period can be usefully divided into an early practicing period and a practicing period proper. The early practicing phase begins with the infant's earliest ability to move physically away from the mother by locomotion; that is, crawling, creeping, climbing, and assuming an upright sitting position. But moving away from the safe protective orbit of the mother has its risks and uncertainties. In the early practicing phase there is frequently a pattern of visually "checking back to mother" or even crawling or paddling back to her to touch or hold on as a form of "emotional refueling."

The practicing period proper is characterized by the attainment of free upright locomotion. It is marked by three interrelated developments that contribute to the continuing process of separation and individuation. These are (1) rapid bodily differentiation from the mother, (2) establishment of a specific bond with her, and (3) growth and functioning of autonomous ego apparatuses in close connection and dependence on the mothering figure.

RAPPROCHEMENT. As this testing of the freedom of individuation proceeds, by about the middle of the second year the child enters the third subphase of rapprochement; as described in *The Psychological Birth of the Human Infant:*

> He now becomes more and more aware, and makes greater and greater use, of his physical separateness. However, side by side with the growth of his cognitive faculties and the increasing differentiation of his emotional life, there is also a noticeable waning of his previous imperviousness to frustration, as well as a diminution of what has been a relative obliviousness to his mother's presence. Increased separation anxiety can be observed: At first this consists mainly of fear of object loss, which is to be inferred from many of the child's behaviors. The relative lack of concern about the mother's presence that was characteristic of the practicing subphase is now replaced by seemingly constant concern with the mother's whereabouts, as well as by active approach behavior. As the toddler's awareness of separateness grows—stimulated by his maturationally acquired ability to move away physically from his mother and by his cognitive growth—he seems to have an increased need, a wish for mother to share with him every one of

his new skills and experiences, as well as a great need for the object's love.

The crisis in the rapprochement phase is particularly that of separation anxiety. The child's wishes and desires to be separate, autonomous, and omnipotent are tempered by an increasing awareness of the need for and dependence on the mother. Ambivalence is also characteristic of the middle phase of the rapprochement subphase. Thus, the mother's availability and the reassurance of her continuing love and support become all the more important.

OBJECT CONSTANCY. As the conflicts and crisis of rapprochement are gradually resolved, the child enters the final phase of separation and individuation; namely, the phase of consolidation of individuality and the beginnings of emotional object constancy. At this stage there are significant developments in the structuralization and integration of the ego, as well as definite signs of internalization of parental demands, reflecting the development of superego precursors.

Attainment of object constancy marks a transition from the stage of need-satisfying relationships to a more mature psychological involvement with objects. Object constancy implies a capacity to differentiate between objects and to maintain a meaningful relationship with one specific object, whether needs are being satisfied or not. Such object constancy also implies stability of object cathexis and specifically the capacity to maintain positive emotional attachments to a particular object in the face of frustration of needs and wishes in regard to that object. This achievement also implies the capacity to tolerate ambivalent feelings toward the object and the capacity to value that object for qualities that it possesses over and beyond the functions that it may serve in satisfying needs and in gratifying drives.

Erikson's Epigenetic Sequence: Instinctual Zones and Modes of Ego Development

Erikson theory of epigenetic psychosexual–psychosocial development made a major integrative contribution to the psychoanalytic concept of development in linking aspects of libidinal instinctual zones with the development of specific modalities of ego functioning. His theory ingeniously links aspects of ego and psychosocial development with the epigenetic timetable of instinctual psychosexual development, clarifying the insight that each society and culture deals with and shapes respective phases of development by specific culturally enforced practices and institutions to ensure that the developing individual can become a viable member of that society and culture. During the course of libidinal development, particular erotogenic zones become loci of stimulation for development of particular modalities of ego functioning.

Zones and Modes. The first developmental mode is related to the oral

phase, specifically to stimulus qualities of the oral zone. This early stage is called the *oral-respiratory-sensory stage,* and it is dominated by the first oral-incorporative mode, which involves the modality of "taking in." Other auxiliary modes are also operative, including a second oral-incorporative (biting) mode, an oral-retentive mode, an oral-eliminative mode, and finally, an oral-intrusive mode. These modes become variably important according to individual temperament but remain subordinated to the first incorporative mode unless the mutual regulation of the oral zone with the providing breast of the mother is disturbed, either by a loss of inner control in the infant or a defect in reciprocal and responsive nurturing behavior on the part of the mother. The emphasis in this stage of development is placed on the modalities of "getting" and "getting what is given," thus laying the necessary ego groundwork for eventually "getting to be a giver."

The second stage, also focused on the oral zone, is marked by a biting modality because of the development of teeth. This phase is marked by development of interpersonal patterns, centered in the social modality of "taking" and "holding" on to things.

Similarly, with the advent of the anal-urethral-muscular stage, the "retentive" and "eliminative" modes become established. Extension and generalization of these modes over the whole of the developing muscular system enable the 18- to 24-month-old child to gain some form of self-control in the matter of conflicting impulses, such as "letting go" and "holding on." Where this control is disturbed by developmental defects in the anal-urethral sphere, a fixation on modalities of retention or elimination can be established that can lead to a variety of disturbances in the zone itself (spastic), in the muscle system (flabbiness or rigidity), in obsessional fantasy (paranoid fears), and in social spheres (attempts at controlling the environment by compulsive routinization).

Psychosocial Crises. Erikson laid out a program of ego (or perhaps better self) development that reached from birth to death: The individual passed through the phases of the life cycle by meeting and resolving a series of developmental psychosocial crises.

These phases of the life cycle and their respective crises accomplished several things. First, they made it clear that ego development was open ended and never finished. Second, the capacity to successfully resolve any one developmental crisis depended on the degree of resolution of preceding crises. One could form a mature and integral sense of identity only to the extent that one had achieved a meaningful sense of trust, autonomy, initiative, and industry. Third, they clarified the relation between the various later phases of development and earlier phases of libidinal development. Erikson's developmental schema gave a better understanding of how earlier libidinal developmental residues were carried along in the course of growth and were built into later developmental efforts of the ego.

Psychoanalysis had not previously had the conceptual tools to deal with this problem, particularly in regard to the postadolescent phases of the life cycle. Finally, Erikson's treatment of these crises as specifically psychosocial brought into focus the fact that the development of the ego or self was not merely a matter of intrapsychic vicissitudes dealing with the economics of inner psychic energies, but was more a matter of the interaction and "mutual regulation" between the developing human organism and significant persons in its environment. Even more strikingly, it is a matter of mutual regulation evolving between the growing child and the culture and traditions of society. Erikson has made the sociocultural influence an integral part of the developmental matrix out of which the personality emerges, and thus made an early contribution to the centrality of relationships in the developmental process.

TRUST VERSUS MISTRUST. This first psychosocial crisis takes place in the context of the intimate relationship between infant and mother with emphasis on the feeding situation. Depending on the quality of feeding experiences, the child learns to accept what is given by the warm and loving mother, to depend on that mother, and to expect that what she provides will be satisfying. Although the child's basically oral orientation is largely biologically determined, the mother's feeding orientation is a product not only of biological factors, but also of a complex process of personal development in which her sense of identity as a woman and as a mother plays a vital part. Any defect in her identity will thus have important consequences for the quality of the interaction between herself and her child.

Successful resolution of this initial phase of interaction entails a disposition to trust others, basic trust in oneself, a capacity to receive from others and to depend on them (to entrust oneself), and a sense of self-confidence. Unsuccessful resolution of this crisis will result in the defect of these same qualities and relative dominance of such opposite qualities as mistrust and lack of confidence. Consequently, the designations "basic trust" and "basic mistrust" stand for a complex of personality factors characterizing successful or unsuccessful resolution of this first crisis.

AUTONOMY VERSUS SHAME AND SELF-DOUBT. The second stage of anal eroticism is marked by formation of a fuller stool and maturation of the neuromuscular system to allow control of sphincter muscles governing retention and release of waste materials. Likewise, the anal zone becomes a source of erotic stimulation through pleasurable sensations of retaining or releasing. Psychosocially, this period is marked by emergence in the child of self-awareness as a separate and independent unit. Increased muscular control is accompanied by increasing capacity for autonomous expression and self-regulation, typically centered on sphincter control. The ego thus

enters into interactions of assertiveness of his newly experienced capacity to will autonomously with and/or against other wills in the social environment, particularly with the parents, and thus begins to experience the alternatives between autonomy and dependence as mediated by socially embedded concepts of the meaning of autonomy and coercion. Successfully resolved, the crisis of autonomy lays the foundation of a mature capacity for self-assertion and self-expression, a capacity to respect the autonomy of others, an ability to maintain self-control without loss of self-esteem, and a capacity for rewarding and effective cooperation with others. The corresponding defect lays the foundation of false autonomy that must feed on the autonomy of others by domination and excessive demands or of an excessive rigidity that can be identified in the fragile autonomy of the compulsive (anal) personality. Failure to achieve basic autonomy implies the lack of self-esteem reflected in a sense of shame and the lack of self-confidence implied in self-doubt.

INITIATIVE VERSUS GUILT. When the child enters the play age, the maturing subsystems serving functions of locomotion and language are sufficiently organized to permit facile use. Motor development permits a wide-ranging experimentation in locomotion. The child begins to "test the limits" of this newfound capability and his or her activity becomes vigorous and intrusive. A similar crystallization of function occurs in the use of language, which becomes an exciting new toy calling for experimentation and the satisfaction of curiosity. The mode of activity is marked by intrusion: Intrusion into other bodies by physical attack, into other's attention by activity and aggressive talking, into space by vigorous locomotion, and into the unknown by active curiosity. All this activity is accompanied by a growing sexual curiosity and a development of the prerequisites of specifically masculine or feminine initiative, which are conditioned by development of a phallic eroticism.

Successful resolution of this crisis provided positive residues for the development of conscience, along with a sense of responsibility and dependability, self-discipline, and a certain independence in the mature personality. This stage is therefore crucial for formation of the superego, based on the introjection of authoritative, and especially parental, ideals and prohibitions. Unsuccessful resolution provides the basis for the harsh, rigid, moralistic, and self-punishing superego that serves as the dynamic source of a basic sense of guilt.

INDUSTRY VERSUS INFERIORITY. This crisis corresponds to a latency period in which the child's world expands to embrace the school situation and widened interaction with peers and adults outside the family. The child takes a step forward from the level of imaginative exploration and play to involvements foreshadowing later participation in the adult world. In

Western culture, children begin to learn skills that will equip them to take their places one day in adult society. They learn the reward systems of the school society and assimilate the values of application and diligence. They also assimilate the implicit cultural values of work and productivity. They can achieve a sense of the pleasure of work, of the satisfaction of a task accomplished, and of the merit of perseverance in difficult enterprises. In other words, normally developing children add to their evolving personality a sense of industry. The danger at this stage is that a lack of success in meeting demands of the school society and failure to resolve this psychosocial crisis will produce a sense of inadequacy and inferiority.

IDENTITY VERSUS IDENTITY CONFUSION. The passage to adolescent years is marked by an intense period of physiological growth and sudden maturation of genital organs and the beginnings of genitality. Puberty is accompanied on the psychosocial level by a reorganization or crystallization of the residues of the preceding formative phases. The developmental preparations for participation in adult life now begin to take a more or less definitive shape, so the adolescent must begin to experiment with establishing a future role and function within adult society. The adolescent must develop a confident sense of self-awareness predicated on the ability to maintain inner sameness and continuity and on the confidence that this awareness is matched by the sameness and continuity of his or her meaning to others. This particular psychosocial crisis is therefore peculiarly vulnerable and sensitive to social and cultural influences.

INTIMACY VERSUS ISOLATION. Following the vicissitudes of adolescence, young adulthood is marked on the psychosexual level by achievement of genital maturity. On the psychosocial level, the establishment of significant interpersonal relationships that complement the previously formed identity in the social sphere parallels this development. Typically, the emerging sexual drive focuses on another individual of the opposite sex as its object. The elements of sexual identification, which are essential aspects of personal identity, are naturally expressed as established by the standards of intersexual behavior of the society and culture. The intimate association of male and female in a close interpersonal union is thus an extension of their own identities, as well as a culturally approved institution (marriage). This fact does not mean that the sexual act is the only path to a sense of intimacy. From the point of view of personality development, the crucial element is the capacity to relate intimately and meaningfully with others in mutually satisfying and productive interactions. The pattern of such self-fulfilling relations will depend in large measure on the identity one has accepted as his or her own and on the determinants of object choice.

The inability to achieve a successful resolution of this psychosocial crisis

results in a sense of personal isolation. The incapacity to establish warm and rewarding relationships with others is but a reflection of the failure to realize a secure and mature self-acceptance. Interpersonal relationships become strained, stiff, or formal. Even if a facade of personal warmth can be produced, there is a rigidly maintained inner wall that is never breached, a wall defended by intellectualization, distancing, and self-absorption.

GENERATIVITY VERSUS STAGNATION. This crisis extends into middle and late adulthood. "Generativity" points to a primary concern with establishing and guiding the succeeding generation (through genes and genitality). As expressed in Erikson's *Insight and Responsibility,* "Generativity, as the instinctual power behind various forms of selfless 'caring,' potentially extends to whatever a man generates and leaves behind, creates and produces (or helps to produce)." It must also be recognized, however, that other areas of altruistic effort cannot be excluded. Perhaps "productivity" or "creativity" would be better or analogous terms.

Such creativity can assume myriad forms, depending on the native endowment of the person; but realized generativity is also determined to a large extent by the identity the individual has accepted and by the extent to which one is capable of interacting maturely and cooperatively with others. Consequently, successful resolution of this crisis depends closely on the degree of success achieved in the resolution of the preceding phases of identity and intimacy. Moreover, true generativity has as its goal enrichment of the lives of others; it involves a direct concern with the welfare of others, exclusive of any concern over self-interest.

INTEGRITY VERSUS DESPAIR. Integrity marks the culmination of development of the personality in Erikson's schema, and thus looks forward to the declining adult years and the shadow of death. It implies acceptance of oneself and the life one has lived with its accomplishments, achievements, failures and disappointments, joys and sorrows. Consequently, in its optimal resolution, existence holds no fear; the ego has resigned itself to acceptance of life itself and to acceptance of the end of that life in death. Integrity thus represents the fully developed personality in its most mature self-realization. The failure to achieve ego integration often results in a kind of despair and an unconscious fear of death: The one life cycle given to every human person as one's own has not been accepted. The person who does not achieve integrity is doomed to live in basic self-contempt.

The parallel lines of development and their interrelation are indicated in Table 6.1–4.

Erikson's ingenious schema has not enjoyed universal acceptance among many analysts. This may be in part because it relies on and preserves the earlier Freudian view of phases of libidinal development that has undergone diminished endorsement by many developmentalists. On

the other hand, it presents an integrated psychosocial developmental theory that seems to look outward to social and cultural influences rather than to internal drive dispositions that many analysts were not prepared to accept. Also Erikson's views point toward a more elaborated concept of the self, in which self-object and self-reality considerations could be more clearly articulated, and in which issues of self-identity are primary but have yet to evolve or find resolution in psychoanalysis. In these terms, the integration of psychosexual with psychosocial themes provides a broadened platform for psychoanalytic understanding of personality development and functioning. In the author's view, if the notion of libidinal zones and modes were transcribed into a theory of motives, the epigenetic schema would have the advantage of pointing toward a more meaningful theory of the development and functioning of the self that would find useful application in extending developmental and adaptive considerations from one end of the life cycle to its inevitable end.

Table 6.1–4.
Parallel Lines of Development

Instinctual Phases	Separation-Individuation	Object Relations	Psychosocial Crises
Oral	Autism, symbiosis	Primary narcissism, need-satisfying	Trust vs. mistrust
Anal	Differentiation, practicing, rapprochement	Need-satisfying, object constancy	Autonomy vs. shame, self-doubt
Phallic	Object constancy, oedipal complex	Object constancy, ambivalence	Initiative vs. guilt
Latency	—	—	Industry vs. inferiority
Adolescence	Genitality, secondary individuation	Object love	Identity vs. identity diffusion
Adulthood	Mature genitality	—	Intimacy vs. isolation Generativity vs. stagnation Integrity vs. despair

Current Considerations

In recent years, a fresh current has entered the mainstream of analytic developmental theory. In some degree reacting to the work of Freud and Mahler, on the basis of observational studies of very young infants, the view of the infant emerged as active, surprisingly well organized even at the very beginning of life, and attuned to and interactive with the mothering figure in complicated and quite sophisticated ways. Particular objections to Mahler's approach to separation-individuation focused on her use of pathological terms (autism, symbiosis) to describe the developmental phases of normal infants and the fact that her extensive observations were often too overlaid with metapsychological terms that obscured the

boundary between fact and theory. The work of Daniel Stern particularly presented a different view of the infant and focused attention on other important aspects of development that had been treated only tangentially in previous studies.

Stern's observations brought into clearer focus the intense affective and interactional matrix between mother and child and directed attention more specifically to the emergence of a sense of self than had previously been available. To begin with, Stern's baby was calm, alert, cognitively well organized, and highly responsive to and interactive with the mother. The extent to which the infant was preadapted to stimuli from the mother and active in eliciting certain responses from the mother, whether caretaking or affective, cast a different light on early stages of development. The relation between mother and child proved to be more active and interactive than had previously been appreciated.

Also following in the wake of Kohut's selfobject theory and the intersubjective developments that arose from it, much of the analysis of mother–child interaction was cast in terms of the relational and intersubjective models of relating. Stern centered his interest on study of the development of a subjective sense of self separately from study of the ego, which had held center stage in Mahler's view. Stern's baby is active and interactive from the beginning, so that emphasis falls on self-functions and interpersonal involvements as central to the developmental experience of the child. Stern's analysis of the development of the subjective self specified a series of four stages leading to four senses of the self as imbricated in four corresponding "domains of relatedness." At about 8 weeks the neonate undergoes a qualitative change in behavior, leading to eye contact and the origins of smiling and other socially inclined behaviors. The succeeding 2-month interval is thought to involve forming a sense of *emergent self,* articulated within a *domain of emergent relatedness,* that organizes the beginning of a sense of self that persists throughout life. After 2 or 3 months, when social interactions seem to become more integrated, the infant is thought to develop a more integrated sense of self as a distinct and coherent entity with control over its own actions and a sense of others as separate and interacting. This basic self-experience is marked by a sense of self-agency, self-coherence, self-affectivity (integrating feeling states with the sense of self), and self-history (sense of enduring through time and having a past). Stern calls this the *sense of core self,* which arises in relation to a *domain of core relatedness*. Somewhere in the 7th to 9th month, there occurs a quantum leap in which the infant discovers that inner experiences can be shared and that others have minds of their own. This marks the development of the *subjective sense of self,* which emerges within a *domain of intersubjective relatedness*. The final step in this progression takes place somewhere in the second year when the capacity for language emerges and gives rise to the *sense of a verbal self,* reflecting the emerging patterns of

linguistic interaction characterizing the *domain of verbal relatedness*. Stern notes that the previous senses of self as emergent, core, and subjective persist and are in some ways enhanced with the onset of verbal capacity, but can be encompassed only partially by verbal means.

But in contrast to the emphasis on separation and release from the entanglement with maternal dependence in Mahler's approach, Stern emphasized intersubjective relatedness, suggesting "a deliberately sought sharing of experiences about events and things." The levels of self-development are intimately related to and reflective of patterns of interaction with the caretaking environment. All these interactions, including the highly intimate bodily interchanges between mother and child, are steeped in affective resonances. Stern carefully traced the emergence of a core sense of self out of increasingly complex bodily and self-other experiences, especially with the mother, all taking place within an intensely affective ambiance, leading to development of an affective core in emerging self-experience. This process leads ultimately to establishing a degree of libidinal self-constancy, involving a sense of self as relatively integrated, independent, durable and stable, as differentiated from others but involved in complex relationships, and as an active agent capable to causing effects and influencing the surrounding environment.

Whether and to what extent this approach to development of the self and its affective resonances is contradictory to the Mahlerian schema or is open to eventual integration remain subjects of debate. Despite the at times oppositional stance that advocates of one or other view assume, the potential exists for combining these perspectives in the hope of deepening one's understanding of the complexities of human psychological development. The relation between Stern's emergent sense of self and Mahler's view of individuation and development of the sense of identity and a separate and autonomous self remain to be explored. Certainly Mahler's concentration on the sense of identity and self-constancy would seem to resonate with certain aspects of Stern's sense of emerging and core self and cannot be discounted.

Critics have pointed out that these different approaches may be looking at children in different contexts of observation—the Stern baby in periods of relatively calm and unconflicted interaction with the mother, the Mahler baby in contexts of greater conflict and separation. It may be that the observations of both groups are valid enough within their own contexts and need to be brought together to accomplish a more complete view of the developing infant. Certainly Stern's focus on the emergence of the self adds a fresh direction of inquiry more congruent with evolving theoretical concepts of the self in psychoanalysis and is more attuned to and reflective of emerging concepts in relational and intersubjective thinking.

OBJECT RELATIONS THEORY

Object relations theory emerged more or less in parallel to the evolution of psychoanalytic ego psychology, but only gradually over the ensuing years have these parallel and somewhat independent courses of theoretical development converged into complementary rather than oppositional perspectives. The development of classical psychoanalytic theory through elaboration of a systemic ego psychology has led inexorably in the direction of better understanding of the adaptive functions of the ego, particularly the close involvement between the ego and reality in its functioning and development. The most important dimension of the problem of the relation to reality involves relationships with others, namely the whole question of object relations. The integration of these complementary currents of analytic thinking provides a more comprehensive basis for thinking about the mind as it functions not only intrapsychically but interpersonally in its relation to others as important sources of the social environment of the human person.

Origins

The origin of the object relations view can best be traced from the contribution of Klein. Klein's theorizing based itself on Freud's later instinct theory, primarily on the death instinct as the main theoretical assumption of her metapsychology. Working primarily with very young children, she described instinctual dynamics in the first years of life. Driven by the death instinct, the child was compelled to rid him- or herself of intolerable, destructive impulses (predominantly oral) and to project them externally. The earliest recipient of these projected impulses was the mother's breast, providing need-satisfying nourishment and satiation (good breast), but also often depriving and failing to satisfy (bad breast). At this stage the images of the breast are part-objects that the infant had yet to combine into a single whole object, the mother. Early frustration of oral needs, even in the first year of life, reinforced these trends so that the bad breast became a persecutory object that was hated, feared, and envied.

Experience of the bad breast and its associated persecutory anxiety formed the earliest developmental stage in Klein's theory: The paranoid-schizoid position. The bad breast withheld gratification and thus stimulated the child's primitive oral envy, provoking sadistic wishes to penetrate and destroy the mother's breasts and body. In boys, these primitive destructive impulses gave rise to the fear of retaliation (based in part on projection) in the form of castration anxiety; in girls, the primitive envy was expressed in envy of the mother's breast during the oral developmental phase, and later was transformed into penis envy during the genital phase. Klein held that by the time of weaning, the child was capable of recognizing the mother as a whole object possessing good and bad qualities. But the combination of good and bad qualities in a single object—previously separated in part-objects—created a dilemma: Destructive attacks on the bad object would

also destroy the good and needed object. This prevented the child from unleashing aggressive impulses against the object and layed the basis for the depressive position, in which aggression was turned against the self rather than against the object. The guilt associated with destructive wishes against the object was the precursor of conscience.

Klein's emphasis in the child's developmental experience fell on the processes of introjection and projection, derived from basic instinctual drives and their interactions with the important and primary objects of the child's early experience. Projection of destructive superego elements permitted acceptance of good introjects (internalization of good objects), thus alleviating the underlying paranoid anxiety. The projected superego elements were later reintrojected to become the agency of guilt and early forms of obsessional behavior. The Kleinian emphasis on good and bad introjects concentrated on vital relationships to objects at the earliest level of child development, and the delineation of the internal structuring of the child's inner fantasy world in terms of the vicissitudes of these introjects, or internal objects, provided the basis and the rudimentary content for an object relations view of development.

Thus internal objects that were either good or bad and with whom the individual was involved in intrapsychic interactions and struggles, which were in many ways as real as those carried on with the real objects outside the person, comprised Klein's "inner world." In fact, Klein saw external object relations as derived from and influenced by projective content derived from the internal object relations. Klein has been generously criticized for her almost blind interpretation of all forms of aggressive or destructive intent as manifestations of the death instinct, for her failure to distinguish among the various kinds of intrapsychic content (lumping object representations, self-representations, internal objects, fantasies, and psychic structures of various kinds together indiscriminately and treating them in a unitary fashion), for her tendency to substitute theoretical inferences for observations, and, finally, for her marked tendency to predate the emergence of intrapsychic organizations that are generally thought by other theorists to be achieved only in later developmental stages, for example, locating the origin of the superego in the first year of life rather than in resolution of the oedipal situation in latency.

In any case Klein's observations and formulations had a tremendous impact, particularly in bringing into prominence the role of aggression in pathological development, in making theorists of development much more aware of the early developmental precursors of later structural entities, and particularly in providing the basic rudiments and foundations for an emergent theory of object relations.

Wilfred Bion extended and applied Klein's ideas, especially developing the ramifications of the notion of projective identification—a process, originally described by Klein, by which a subject displaces a part of the self

into an object and then identifies with that object or elicits a response in the object corresponding to qualities of the projection. Bion applied this notion to a wide range of psychotic and cognitive operations. He developed the metaphor of the "container" and the "contained" to express the manner in which projective identification occurs, especially in the contexts of mother–child and analyst–patient interaction. The child–patient projects toxic or destructive contents onto the mother–analyst, who in turn absorbs, modifies, or "contains" it so that it becomes available in more benign form for subsequent reinternalization by the child–patient, resulting in a healthier modification of the child–patient's pathogenic introjects.

Ego and Objects

Beginning in about 1931, Ronald Fairbairn shifted the emphasis in his thinking specifically to the problem of ego analysis. Fairbairn's contribution was to bring personal object relations into the center of the theory. Whereas the ego in Freudian theory had been regarded as a superficial modification of the id, developed specifically for the purpose of impulse control and adaptation to the demands of reality, Fairbairn conceived of the ego as the core phenomenon of the psyche. Rather than an organization of functions, he conceived of it more specifically as embodying a real self, that is, as the dynamic center or core of the personality. Instead of basing his theory on the instinctual drives with their implied seeking of gratification, Fairbairn shifted the emphasis to the ego and saw everything in human psychology as specifically an effect of ego functioning.

This reorientation was accompanied by a parallel reformulation of the instinctual perspective. The libido or instincts in general, rather than mechanisms for energic discharge, was regarded as essentially object-seeking. Erotogenic zones were not the primary determinants of libidinal aims but, rather, served as channels mediating the primary relationships with objects, particularly with objects that had been internalized early in life under the pressure of deprivation and frustration. Ego development itself was characterized by a process whereby an original state of infantile dependence, based on a symbiotic union with the maternal object, was abandoned in favor of a state of adult or mature dependence based on differentiation between self and object. Thus, Fairbairn conceptualized the developmental process in terms of the vicissitudes of relations with objects, rather than the vicissitudes of instinctual dynamics.

The basis for much of Fairbairn's theorizing was his experience with schizoid patients. He contrasted the basic dilemma of the schizoid with that of the neurotic patients he felt were treated primarily by classical psychoanalysis. He developed the view that the schizoid was not primarily concerned with control of threatening impulses toward significant objects, but that the issue for this kind of patient was essentially that of having an ego capable of forming object relations at all. The relationship to objects

presented a difficulty, not because of dangerous instinctual impulses arising in connection with them, but because the ego itself was weak, undeveloped, infantile, and fragile. In the struggle to defend and protect this inner weakness, the schizoid's impulses became antisocial.

Thus, object relations theory in its bare essentials contains a number of basic points that differentiate it from classical theory. First, the ego is conceived of as whole or total at birth, becoming split or losing inner unity as a result of early bad experiences in object relationships, particularly in relation to the mothering object. This point differs quite radically from the classical theory, according to which the ego begins as undifferentiated and unintegrated and only achieves unity through the course of development. Second, libido is regarded as a primary life drive of the psyche, the energic source of the ego's search for relatedness with good objects, which is essential for ego growth. Third, aggression is regarded as a natural defensive reaction to frustration of the libidinal drive, rather than specifically as an independent instinct. Fourth, the structural ego pattern that emerges when the pristine ego unity is lost involves a pattern of ego splitting and the formation of internal ego–object relations.

The shift in emphasis toward the primacy of the external environment and the influence of objects on the course of development has established a definite trend in psychoanalytic thinking and has been advanced primarily in the work of British theorists, among whom the work of Michael Balint and Donald Winnicott stands out. Winnicott particularly has emphasized the importance of early interactions between mother and child as determining factors in laying down important components of ego development. Currently, there is ample room for overlap and integration in the approaches and formulations of both object relations theorists and more classical psychoanalytic ego theorists. Such integration has advanced to the point that it is generally regarded as forming at least complementary aspects of a common theory, if not a more comprehensive and unified theory as such.

Both Balint and Winnicott were concerned with levels of early developmental failure that were essentially preoedipal, which were manifested in forms of personality disorder that are more primitive and more difficult to treat than the usual neurotic disorders, and that seemed to involve critical aspects of the relationships with objects early in the course of development, and correspondingly do not fit well with classical psychoanalytic structural theory, with its basic focus on issues of intrapsychic conflict.

Balint envisioned several layers of psychological functioning in analysis. The first is the familiar genital level, centering on triadic relationships and concerned specifically with intrapsychic conflicts. These conflicts and the corresponding quality of relationships were usual and familiar in most analytic processes and could be treated by use of adult language in verbal

interpretations. There was, however, a second, deeper level in which the conventional meaning of words no longer had the same impact, and interpretations were no longer perceived as meaningful by the patient. This was the level of preverbal experience. He referred to this level of impairment in object relations as the *basic fault*. Balint recognized that at this level of preverbal experience any attempt to address or describe the child's experience in adult language is bound to fail. Problems arose in analysis when efforts were made to interpret events from this preverbal level in adult or secondary process terms.

Balint distinguished between forms of regression that he described as benign and malignant. Benign regression was more or less equivalent to the usual form of analytic regression to more primitive relationships with primary objects. Such regression was gradual, tempered, and modulated according to the patient's capacity to tolerate and productively integrate the resulting anxiety. During this regression, the analyst maintains the analytic structure and relation, and his empathic responsiveness made it possible for the patient to withstand this unstructured experience and to keep the anxiety within manageable limits. At the level of the basic fault, the lost infantile objects can be mourned, and the quality of the relationship with them was open to reworking so that the patient's basic assumptions governing his or her interaction with the internal and external object world could be reformed.

During phases of benign regression to this preverbal and pregenital level of object relationship in the analysis, the analyst could usually provide an adequate degree of empathic acceptance and recognition, rather than verbalized interpretations, of this level of the patient's unstructured and regressive experience, without anxiety or any need to escape or subvert this level of experience through interpretation. Balint felt that the dynamics at this level are more primitive than can be adequately expressed in terms of conflict because they derive from the basic form of dual relationship involved in early mother–child interaction, which is the basic fault.

In contrast, malignant regression tends to be precipitous and extreme; the ego is prematurely overwhelmed by traumatic and unmanageable anxiety. This anxiety prevented any effective reworking of fundamental disturbances in object relationships, recreating and reinforcing the basic fault, rather than creating the conditions for its therapeutic revision. At an even deeper level, beyond the reach of analytic resources, there lies the area of creativity; that is, an idiosyncratic, uncommunicable, and objectless area beyond any conventional or linguistic expression.

Regression to the level of the basic fault was a quite different and distinct phenomenon than the more usual oedipal regressions experienced in the analysis of adult neurotics. In the benign oedipal regression, the aim was gratification of infantile instinctual wishes. Regression to the level of the basic fault, however, sought a basic recognition by the therapist, as well

as protective support and consent to express the inner core of creativity that lies at the heart of the patient's being and accounts for the capacity to become ill or well. Balint used the notion of primary love at this deepest level to describe withdrawal of libido from a frustrating object in the effort to reestablish a certain inner harmony in which it becomes possible to recover the conditions of early care and tranquility. He referred to a "harmonious interpenetrating mix-up" to describe this early, almost undifferentiated interaction of the infant and environment. The analogy he used was that of breathing air; the organism cannot exist without air, so air and the organism are seemingly inseparable, but to cut off the supply of air reveals both the organism's need of it and the distinction between air and the organism. In terms of primary love, as it came to bear in analysis, then, the patient would seek a basic form of recognition from the analyst, as he had from significant objects in the patient's early life experience.

Winnicott, in turn, also was concerned with the earliest phases of the mother–child relationship and the importance of what he described as "good-enough mothering" for the child's psychic development. Development involved movement from an early stage of total or absolute dependence toward a more adult phase of relative dependence and independence. As he saw it, the inherited native potential for growth was strongly influenced by the quality of maternal care. This potential for development is affected even from the moment of conception. Even before birth, the child becomes invested by a strong narcissistic cathexis that allows the mother to identify with the child and to become empathically attuned to the child's inner needs, as if the child were—and indeed is—an extension of her own self. He called this early prenatal involvement of mother and child-in-the-womb a *primary maternal preoccupation*. This set the stage for development of a holding relationship, in which the mother becomes sensitively attuned to the infant's needs and sensitivities and is both physically and emotionally responsive to them, thus providing a physical, physiological, and emotional ambiance, protection, and security for the absolutely dependent infant.

As the infant moves from this early stage of absolute dependence toward a more relative dependence, awareness of personal needs and of the existence of the mother as a caretaking object grows. The optimal relationship at this stage involves a continuation of protective holding, along with an optimal titration of gratification and frustration. As a result of this optimally attuned relationship between the patterning of infantile drives and initiatives and their harmonious fitting in with maternal sensitivities and responsiveness, there is a developing sense of reliable expectation that the infant's needs will be satisfied without the threat of excessive withdrawal of the mothering object and without the threatening, overwhelming, and short-circuiting of the infant's initiatives as a result of excessive maternal impingements.

In the course of normal development, this allowed for the emergence of a certain infantile omnipotence from which the child gradually retreats as a result of the experiences of tolerable degrees of frustration by and separateness from the maternal object. Although the mother continues her holding at this phase, she must yet allow enough separation between herself and the developing infant to permit expression of the baby's needs and initiatives that form the rudiments of an emerging sense of self. If she is too distant, too unresponsive, or not sufficiently present, anxiety arises and is accompanied by the fading of the infant's internal representations of her.

The transition from a phase of absolute dependence to one of relative dependence represents a crucial development in the capacity for object relations. It is accompanied by a critical transition from total subjectivity to the capacity for objectivity in the perception of and relation to objects. The transition from subjectivity to objectivity is accomplished by development of transitional phenomena, expressed in the first instance in the emergence of transitional objects. These objects are the child's first object possessions that are perceived as separate from the emerging self—the first "not me" possessions. From the study of infant behavior, Winnicott argued that the transitional object was a substitute for the maternal breast, the first and most significant and substitute object in the environment to which the infant related. He regarded the transitional object as existing in an intermediate realm or space contributed to both by the external reality of the object (the mother's breast) and by the child's own subjectivity. This intermediate realm was at once both subjective and objective without being exclusively either.

Winnicott referred to this realm as the *realm of illusion,* an intermediate area of experience that embraced both inner and external reality and may be retained in areas of adult functioning, involving such imaginative capacities as creativity, religious experience, and art. In its primitive form, however, the transitional object commonly experienced in childhood development may take the form of a particular object, a blanket, a pillow, or a favorite toy or teddy bear to which the child becomes intensely attached and from which it cannot be separated without stirring up severe anxiety and distress. Attachment to this object was an immediate displacement from the figure of the mother and represented an important developmental step, insofar as it allowed the child to tolerate increasing degrees of separation from the mother, using the transitional object as a substitute.

The mother participated in this intermediate transitional realm of illusion by her responsiveness to the infant's need to continually create her as a good mother. In her sensitivity and responsiveness, she functions as a good-enough mother. However, her failure to provide such adequate mothering, either by excessive withdrawal or by excessive intrusion and

control, may result in emergence of a false self in the child based on compliance with the demands of the external environment, a condition that reflected a developmental failure and resulted in a variety of often severe character pathologies.

When such patients were seen as adults, they were neither neurotic nor psychotic, but seemed to relate to the world through a compliant shell that was not quite real to them or to the analyst. They were often mistrustful without being specifically paranoid, they appeared withdrawn and disengaged, and seemed able to relate only by means of the protective shell, which seemed apparently obsessive and compliant but that separated and isolated them from meaningful contacts with their fellows, even as it provided their only basis for relationship. These disturbed personality types reflected a basic impairment in very early object relations, particularly in the mutuality and responsiveness of very early mother–child interaction.

Infants who developed in the direction of a false self-mode have not experienced the security and mutual satisfaction of such a maternal relationship. Such mothers are empathically out of contact with the child and react largely on the basis of their own inner fantasies, narcissistic needs, or neurotic conflicts. The child's survival depends on the capacity to adapt to this pattern of the mother's response, which is often grossly out of phase with the child's needs. This established a pattern of gradual training in compliance with whatever the mother was capable of offering, rather than seeking out and finding what is needed and wanted. Consequently, the child's needs, instinctual impulses, wishes, and initiatives, instead of becoming a meaningful guide to satisfying growth experiences and enlarging capacities to interact meaningfully with objects, become from the very beginning threatening to the harmony of the relationship with the mother, who remains unresponsive to affective feedback from the child.

Winnicott's attempts to formulate principles of treatment for such basically impaired patients built on the model of good-enough mothering. This called for a capacity for holding, for empathic responsiveness, and for a capacity for creatively playful exchange that allowed the patient's capacities for growth to emerge and flourish. Such playful engagement and interaction between analyst and patient had a facilitating effect on expansion of the patient's authentic sense of self, which had been hidden behind an external facade of false–self-compliance.

ATTACHMENT THEORY

Attachment theory forms another more recent development in the study of relationships with objects, which seems to extend or complement object relations theory, especially in trying to characterize early forms of infant maternal attachment and its consequences for further development and for adult personality functioning. Attachment theory derives from the work of

John Bowlby, who emphasized, not unlike the insistence of self psychology on the persistence of dependency needs throughout life, the importance of attachment needs from birth to death. In his studies of infant attachment and separation, Bowlby pointed out that attachment constituted a central motivational force, and that mother–child attachment was an essential medium of human interaction that had important consequences for later development and personality functioning. Attachment theorists have subsequently studied patterns of early infant attachment and related them to patterns of adult interaction with significant objects. Using the Adult Attachment Interview (AAI), developed by Mary Main and others, they have been able to document the nature of internal working models, similar to introjections or internal objects, based on early attachment relations. Such studies document otherwise difficult phenomena to assess regarding degrees of security of attachment between infants and caregivers and the development of internal representations of such object relationships. Both attachment theorists and object relations theorists emphasize the significance of the mother's empathic responsivity to infant needs for self-development and relatedness, the importance of the mother–child involvement for personality development, and the role of the mother as catalyst for age-appropriate development. An important measure, based on AAI measures, is *reflective functioning,* describing the capacity to perceive and think about the intentionality of self and others. This capacity reflects the formation of coherent representations of the psychic world of others, especially caregivers, and of the subjects' own internal states. High levels of reflective functioning in mothers were found to contribute significantly to the secure attachment of their infants.

Attachment theorists have studied the patterns of infant responses in the so-called Stranger Situation, developed by Mary Ainsworth and her colleagues, an arrangement for observing the quality of parent–child interaction and the effects of separation on the child's emotional adjustment. These findings led to defining four categories of infant attachment behavior: Secure versus insecure, latter divided into avoidant or dismissive, resistant or anxious-ambivalent, and disorganized-disoriented. It turned out that these patterns of infantile attachment could then be related to attachment attitudes in adults based on results of the AAI. Thus secure infants were found to develop secure and relatively autonomous patterns of attachment and interaction with others in adult relationships and were more likely to become more resilient, self-reliant, empathic, and confident in forming meaningful relationships. Avoidant or dismissing infants were found to be associated with an adult pattern of dismissing object relationships as important or significant, a combination of indifference and independence, and a form of self-reliance akin to counterdependence and denial of any need for the presence or caring of others. Their investment is in maintaining aloofness and detachment in the

service of defending against loss and separation. These subjects were seemingly more vulnerable to adult disorders than others, for example, borderline, histrionic, and dependent personality disorders, and may be more difficult to treat in psychotherapy. Resistant or preoccupied infants later seemed to develop a somewhat preoccupied adult stance in which subjects are aware of attachment needs but come to expect that they will be inevitably disappointed, resulting in excessive dependency accompanied by disappointment and fears of loss and abandonment. Childhood experiences of past attachments were often connected with confused, angry, or fearful emotions, especially fears of abandonment. They seem invested in maintaining contact with need-satisfying objects but continually defending against the anxiety related to separation and loss. The disorganized-disoriented infantile pattern seemed to be related to an unresolved and disorganized pattern in adults, suggesting a more severe disturbance in self-object relations. These findings provide an extension and specification of an object relations approach, as well as providing specific empirical and observational methods for more detailed study of the development of object relations from childhood to adulthood.

Some object-relations theorists have advanced objections to the attachment theory approach. Criticisms have focused on the attachment categories, namely that they seem stagnant, lack developmental complexity, give little account of basic instinctual motivations and the role of unconscious fantasy, and provide little or no account of the progression in the evolution of object representations and patterns of internalization as development proceeds. Moreover, attachment theory pays little attention to relational dynamics between mother and child, especially changes over time as the child progresses developmentally. No mention is made of the uniqueness of each mother–child relation and its modification by surrounding influences from other relationships.

Theoretical issues that would seem to call for clarification and integration might include the relation between mental representations and the internal working models of attachment theory. They seem to be analogous and both theories remark that both representation and working models of self and others derive from experiences with early caretaking others and become guides for subsequent development of more mature object relations. They differ, however, in the quality of epigenetic differentiation, representations following a specified developmental sequence and thus becoming more complex, symbolic, and linguistically mediated in contrast with internal working models that preserve a relatively constant character throughout. This and other theoretical divergences would seem to provide a ripe field for theoretical revision and integration of these two relatively independent perspectives on early object relationships and their developmental course.

PSYCHOLOGY OF THE SELF

Over the past nearly half century, the concept of the self has been emerging with increasing emphasis and definition as a central notion in the deepening psychoanalytic understanding of the organization and functioning of the human psyche. Although the concepts regarding the understanding of the self are still very much in flux and the place of the notion of the self in psychoanalytic theory remains tentative and uncertain, the issues addressed by the psychology of the self seem to be of sufficient significance and to have gained a more or less permanent place in psychoanalytic thinking, so that a consideration of these issues is warranted in this presentation.

The issues that self psychology addresses are by no means new to psychoanalysis. Part of the problem stems from the ambiguity in Freud's use of the term *Ich,* referring both to the ego as part of the mental apparatus, a structural agency, and to the more experiential subjective and personal sense of self. The decision of the editors of the English *Standard Edition of The Complete Psychological Works of Sigmund Freud* to translate *Ich* by the term "ego" tended to shift the meaning of the term toward the more impersonal structural sense of an internal psychic agency and away from the more subjective experiential implications of the self as engaged in the world of objects. There are passages where it is quite clear that Freud uses the German term *selbst* as synonymous with the term *ich,* referring to the subjectively experienced self, the person as such.

This unresolved ambiguity and the progressive shift in implications of the term ego have left a certain vacuum in psychoanalytic metapsychology. This deficit has been attacked by a number of thinkers as reflecting a lack of personal ego or a sense of self-as-agent in psychoanalytic theory. Partly in an attempt to deal with this issue, various analytic thinkers in a variety of contexts have focused on the notion of the self. Kohut's development of self psychology entered the field with an emphasis on narcissistic issues and a focus on the self, more or less exclusively, in terms of its subjective and experiential dimension and has subsequently become a major stimulus to renewed interest in the self.

The self psychology movement came into prominence largely as the result of Kohut's efforts in the late 1960s, but there was a history of development of a self-concept in psychoanalysis well before that. Development of a concept of the self in the context of the structural theory was stimulated by Hartmann's distinction between ego and self, terms that had been left ambiguous by Freud: The ego was an intrapsychic organization of functions, while the self was cast in terms of self-representations that then became objects of narcissistic libido, but also was specifically connected with object relations. In these terms, the ego was related to and interactive with other intrapsychic structures, for example, the superego, and the self was concerned with object relations and self-

object interactions. This distinction also clarified the differentiation of object libido and narcissistic libido, since the self was the repository for secondary narcissism and as such distinct from the ego.

Hartmann's distinction between the ego and the self casts the self in representational terms. The self was conceptualized either as a complex representation, organized and synthesized as a function of the ego, or, by later theorists, in structural terms as a more complex and supraordinate integration of the tripartite structures (that is, embracing the tripartite entities as subordinate substructures). The former view regarded the self as part of the representational world, whereas the latter assigned it to the realm of internal psychic structure. The differences of emphasis and formulation regarding these two perspectives remain a persistent problem in developing a consistent concept of the self. This emerging line of thinking about the self was pushed out of the analytic limelight by the emergence of Kohut's version of self psychology.

Kohutian Self Psychology

Self psychology, as a separate movement within psychoanalysis, takes its origin from contributions by Heinz Kohut and his followers. Kohut linked the origin of the self to narcissism, viewing the self as the result of a separate line of narcissistic development that progresses through a series of archaic narcissistic structures toward establishing a mature and cohesive self-organization.

In Kohut's view of narcissistic development, the original primary narcissism differentiates in the course of development, and in response to lapses in parental empathy differentiates into two archaic configurations, the *grandiose self* and the *idealized parental imago*. The *grandiose self* involves an exaggerated and exhibitionistic image of the self that becomes the repository for infantile perfection; the *idealized parental imago,* in contrast, transfers the previous perfection to an admired omnipotent object or objects. Further normal development of the grandiose self leads to more mature forms of ambition, self-esteem, self-confidence, and pleasure in accomplishment. The idealized parental imago likewise becomes integrated into the ego ideal with the mature values, ideals, and standards it represents. Pathological persistence of the grandiose self results in intensification of grandiosity, exhibitionism, shame, envy, depression, hypochondriacal concerns, and undermining of self-esteem. Loss of the idealized object or the idealized object's love can result in narcissistic imbalance, leaving the individual vulnerable to depression, depletion, poor self-esteem, failure of ideals and values, and even self-fragmentation.

Kohut based his self psychology on the need, both during the course of development and during the course of life, for empathic interaction with *selfobjects*. The original selfobject is the mother or caretaking person who provides empathic response to selfobject needs in the infant in the form of

love, admiration, acceptance, joyful participation, warmth, and responsiveness, communicating a sense of valued and cherished existence to the child. But, Kohut contends, human beings continue to seek objects to fulfill these basic selfobject needs throughout life. Failures to fulfill such needs can result in the formation of pathological psychic structures and patterns of behavior during development and pathological character structures during adult life.

Evolving Concepts of the Self

An alternative line of development of the notion of the self in more structural terms, as previously discussed, can best be traced back to Hartmann's effort to clarify the ambiguity latent in Freud's use of the term *Ich,* distinguishing ego as an intrapsychic agency interacting with other intrapsychic entities, for example, superego and id, from the self conceived as equivalent to the concept of the person as such and primarily conceived in relationship to objects.

In an effort to clarify the theoretical implications of the self, early thinkers, following Hartmann's lead, came to define the self in representational terms—that is, as referring to the self-representation, which was then regarded as a subordinate function of the ego. Another point of view, however, would see the self as a structural organization, either envisioned as a fourth focus of organization in addition to the tripartite entities or as a supraordinate organization including the tripartite structures along with additional structural aspects.

Part of the difficulty is that the notion of the self can be looked at from a variety of perspectives. The self can be seen as agent, or as object, or even in locational terms, with respect to questions of what is inside or outside of the mind or the psychic structure and what it might mean for parts of the self to be internalized or externalized. The representational view of the self seems to lend itself most clearly to a view of the self as object; that is, as what can be cognitively and experientially grasped of the self as an object of the subject's inner experience. Such an experienced self-as-object must have representational qualities to be cognitively relevant. By the same token, the structural perspective seems to be most congruent with the view of the self-as-agent, as a source of psychic integration and activity, and as synonymous with the originating source of personal action and awareness. The structural aspect of the self-as-agent, with particular reference to its function of conscious subjectivity, comes closest to satisfying the demand for a "personal ego," or perhaps better "personal self," in the theory.

The theory of the self remains at this juncture uncertain and very much in flux. The current prevailing view of the self is Kohut's self psychology. Although the emphasis in self psychology on the subjective and experiential aspects of self-functioning has wide following, his concept of the self seems relatively weak in accounting for the self as a source of personal action and

decision; there is little sense of the self as autonomous or as an independent and creative source of self-generated productivity and responsibility. Emphasis is placed rather on the initial and continuing dependence of the self in relation to selfobjects. Developmentally, the importance of empathic parental attunement is considered vital to normal psychic growth in the interest of facilitating infantile needs for nurturance, love, and narcissistic approval. This sets self psychology apart from the separation-individuation model in which infantile dependence is viewed as yielding to the increasing independence and autonomy of the child in the course of development and thus connoting a form of gradual separation from parental dependence. At the same time, the developmental sympathies of self psychology have found increasing endorsement by current developmental theorists and have gradually been absorbed into more specifically relational and intersubjective viewpoints.

However the ultimate conceptualization of the nature and functioning of the self may be resolved, it seems apparent that the psychology of the self in some form will continue to gain a permanent place in psychoanalytic thinking and theory. It is possible at this point to specify some of the theoretical gains of the emerging role of the self-concept:

1. The self as a theoretical construct provides a focus for formulating and understanding the complex integrations of functional processes that involve combinations of functions of the respective operational components of the self. This would have specific application to such complex activities as affects, in which all of the psychic systems seem to be in one way or other represented; complex ego–superego integrations reflected in such formations as value systems; and other complex interactions of psychic systems that involve fantasy production, drive-motor integration, or cognitive-affective processes. There is room here for considerable reworking and refocusing of traditional psychoanalytic ways of looking at and understanding psychic phenomena in terms of the self as a referent system.
2. The self-concept provides a more specific and less ambiguous frame of reference for the articulation of self-object interrelationships and interactions, including the complex areas of object relations and internalizations.
3. The emergence of a self-concept provides a locus in the theory for articulating the experience of the personal self, either as grasped introspectively and reflectively or experienced as the originating source of personal activity. This sense of the self-as-subject would provide a place within the theory for an account of subjectivity and subjective meaning.

This approach raises an important metapsychological issue, namely the relationship between the experiential organization of the self and the

tripartite entities. The organization of the self and the organization of the structural tripartite entities cannot be simply identified. The self-organization operates at a different level of psychic organization than the structural entities; moreover, the structural entities in the strict theoretical sense are understood to be organizations or classifications of specific self-functions. Although the theory at various points attributes more or less personalized, anthropomorphized metaphors to describe the operation of these structures, their strict theoretical intelligibility is nonetheless given in terms of the organization of specific functions cast in the form of structural metaphors. In an authentic theory of the self, there is only one agent, the self as representing the total person, whose functions can be categorized and classified in reference to multiple structural formations conceived as substructures of the self. Thus, when it is said, for example, that the ego decides something, the deciding can be viewed as an action of the self of such type that can be classified as having the quality of an ego function. It would then be the self, functioning in its ego modality, that does the deciding.

Relational and Intersubjective Approaches

The emergence of relational and intersubjective approaches to the nature of the analytic relationship has in many ways incorporated and extended aspects of Kohutian self psychology. The relational perspective, following Kohut's initiative, began by distinguishing between drive theories and relational theories, implying that viewing the analytic relation as a function of the ongoing interaction between analyst and analysand demanded abandonment of the drive theory. This dichotomy proved false in that one could entertain the relational view without dispensing with the operation of internal dynamics and psychic processes operating within the individual participants, both before and within the development of the analytic relation. Even a relational viewpoint, for example, cannot do without a system of intrapsychic motivation, since the process of relating to another person and forming a relationship is itself motivated.

These approaches to understanding the analytic interaction employ a form of social constructionist epistemology in which transference is regarded as resulting from the interaction of analyst and patient. This constructivist view contrasts with the more objectivist view of traditional ego psychology and object relations. Thus the analyst's attention is focused on the here-and-now interaction with the patient rather than on the inner dynamics of the patient's mental life and experience. This effects a shift from a one-person to a two-person psychology in which the ongoing interactions, whether conscious or unconscious, between the participants become central and transference and countertransference are regarded as mutually cocreated by both analyst and analysand.

This approach was further developed into a view of analytic interaction

in intersubjective terms. The self psychological emphasis on self-selfobject transferences has encouraged movement away from considerations of the analyst's stance as neutral or observational, questioning the analyst's subjectivity, authority, and capacity to know any objective reality about the patient. On these terms, personality development is dependent on the interpersonal field insofar as psychic life is continually being remodeled in terms of both past and present relationships, and not determined by fixed patterns deriving from past unconscious conflicts. The concept of personality as developing within a relational matrix calls for a central focus on the intersubjective field within the relationship between analyst and patient. It is this aspect of the analytic situation that is explored and interpreted in the interest of bringing about personal growth in the patient. The analyst's technical neutrality and objectivity are rejected in this approach as illusory and little more than expressions of the analyst's authoritarian position. Within a self-selfobject or intersubjective relation neutrality is precluded as potentially traumatizing and destructive to potential consolidation of the self. An inevitable consequence of these approaches is that they do away with the traditional notion of the unconscious and undercut any sense of transference as reflecting unconscious aspects of the patient's inner psychic life, since transference is created anew in the present analytic interaction.

In its intersubjective guise, this approach poses the most radical departure from classical analytic views. In the "intersubjective field," cocreated by analyst and patient, both diagnosis and treatment are functions of the mutual interaction between them, implying that neither diagnosis nor prescribed therapy techniques have any objective validity but arise spontaneously from the workings of the intersubjective process. Any contribution from biological determinants and constitutional factors, or for that matter from prior history of object relationships and developmental vicissitudes, is discounted. Transference and countertransference are viewed as arising in the immediacy of the intersubjective interaction and as developing in virtue of mutual here-and-now influences the participants exercise on each other, rather than from displacements or projections from the past. Questions have arisen as to the meaning of the unconscious in this perspective and whether it has been replaced by interactional determinants. The emphasis on here-and-now interaction has promoted technical divergences such as the analyst's self-revelation and disclosure, even to the point of sharing countertransference elements with the patient. Preferences for spontaneity and self-expression prevail, even "throwing away the book" in the interest of escaping from any implication that the analyst has any knowledge, competence, skill, or authority that would disturb the presumed equality of the intersubjective matrix.

This approach has thus raised some significant epistemological issues, perhaps the most telling having to do with the question of what a relation

might be without a constitutive subject and object. Ordinarily relations arise between a subject and an object, connecting them in some fashion. But the relation itself arises from their mutual connection and interaction and has no meaning or existence without both subject and object. Many intersubjective theorists address the relation as though it were a third entity existing independently between subject and object, such that properties of the analytic interaction arise within the relation separately and independently of the constituting subject and object, that is, patient and analyst.

Despite the apparent conceptual difficulties and often stringent criticism, the relational and intersubjective viewpoints cannot be dismissed. They bring into focus and articulate aspects of the analytic situation that have long been assumed and taken for granted, as if in themselves they were of secondary importance and could be ignored in favor of the exclusive exploration of the patient's inner world. The contribution of these relational approaches lies in the focus on the analytic relation and interaction as such, bringing into the clear light of analytic scrutiny and exploration aspects of the analytic process that have the potential of deepening and expanding one's understanding of their therapeutic importance. Once the radical extremes of the dialectic have spent their course, there does not seem to be any reason why these currently divergent and oppositional currents in the contemporary analytic field cannot be reconciled and ultimately integrated.

Neuropsychoanalysis

Another exciting and challenging branch of contemporary psychoanalytic thinking has arisen from the neuroscientific exploration of brain functions. The advent of neuroimaging and microelectronic brain recording devices has crossed a new threshold in the understanding of the functioning of the brain, from which important discoveries emerge almost daily. Insofar as psychoanalysis is itself a scientific study of the functioning of the mind, these new developments thrust to the foreground issues related to the ways in which brain processes are involved in and related to mental functioning.

Some analysts have argued that there is a fundamental sense in which neurobehavioral discoveries do not affect analysis that much. For the individual analyst working in his or her consulting room, it does not matter, nor does it affect either his or her work or approach to the patient in any significant way. Some have also sounded a warning that efforts to integrate psychoanalysis with neurobehavioral discoveries might put analysis at risk of losing touch with its essential principles and centering on unconscious dynamics and subjectivity. Thus, the attraction of such scientific integration and the lure of linking analysis with current scientific developments may come at too high a cost, namely the loss or diminution of an authentic psychoanalytic perspective. But that view has to be qualified

in some degree insofar as some recent neuroscientific discoveries have already found a place in analytic thinking and practice. Neurobehavioral research, for example, has identified multiple and more or less independent forms of memory function, each subserved by different neural systems. The primary distinction between explicit and implicit forms of memory has begun to have an impact on analytic technique, insofar as it brings into focus the problem that free associative and interpretive methods are appropriate for the recovery and processing of explicit forms of memory, as for example autobiographical memory, but not the implicit forms. Analysts are already at work to determine what ways are possible for engaging with and what therapeutic use can be made of implicit memory systems.

But, although it may not matter which specific neural mechanisms are at work in producing mental actions and functions, it is immensely important that there are such brain mechanisms and that they can be identified. The significance, therefore, is less in terms of implications for clinical technique and practice, and considerably more in terms of the scientific status and integration of psychoanalytic theory. The critical focus in this area thus falls on the understanding of the mind–brain relation, or as more broadly conceived the mind–body relation.

Neuroscientific research makes it increasingly and abundantly clear that there is no action or function of the human mind that does not relate directly to and correspond with a parallel process in the neural net. It is not surprising that the brain is found to be involved in regulating and directing bodily functions: When one moves their hand, the corresponding activation of cells in the motor cortex seems unsurprising. But when one imagines the outline of a triangle, or when one adds a set of numbers, it somehow seems surprising that parallel activity can be identified in the brain. The conundrum here results from a dualistic assumption: People expect the physical brain to operate as part of the physical body, but they do not expect the physical brain to operate in similar fashion in the production of mental acts. The problem is how to conceptualize the observations linking particular mental functions with identifiable patterns of brain activation, a linkage that is specific and characteristic for each such function. Is a dualistic persuasion in which the corresponding patterns are regarded as simply independent and separate but running in parallel followed, or is some form of mind–brain interactionism postulated? An alternative would be to interpret the mind–brain relation as an integrated and unified function such that every mental action is the result and product of some form of brain mechanism. In this sense, mental functions would be regarded as the direct expression and product of the brain. This form of integrated theory prevails currently among neuroscientists, but has found less acceptance among psychoanalysts.

The stumbling block for many analysts is that there seems to be an inescapable gap between neuroscience on one side and psychology and

psychoanalysis on the other. Neuroscience focuses on third-person data, on observations and measurements of neural processes, patterns of electrochemical transmission, activations of patterns of neurophysiological organization traced by multiple imaging and neural recording techniques. In contrast, analysis studies the inner psychological processes in the mind; the data are subjective, first person in character, not available simply to external observation but only through subjective processes of introspection, recall, and association. The nature of the data and methods of acquiring the data are totally different and independent. The conclusion is unavoidable: No amount of detailed and technically sophisticated study of neurological process in the brain will ever reveal anything about the nature and quality of individual subjective experience resonant with motives and meanings. By the same token, no amount of analytic inquiry, exploration, and interpretation will ever yield an iota of further understanding of the brain mechanisms that underlie and generate psychic experiences. The impasse seems absolute. The path of scientific progress, therefore, lies in the pursuit of both channels of knowledge about the human phenomenon; either one without the other provides no more than a limited and incomplete understanding of what is essentially a complex, integrated, and unified mind–brain entity that constitutes the reality of the human person or self.

But the fact remains that neuroscience currently offers a kind of evidence that was not available to previous generations, including Freud, neurologist though he was. It is as though the brain going into action in performing mental acts can almost literally can be seen. The connection is no longer an inference, but more akin to an observation. One can argue that when effects follow acts, there is direct causality at work. It opens the possibility that the brain is directly causing these effects, that mental actions are equivalently actions of the brain. The conclusion would have far-reaching effects on how subjective phenomena is understood. Further, the role of psychoanalysis as a scientific study of the mind would be more clearly established and confirmed. However, these questions and the possible alternatives they open are still matters of active debate and reflection.

CLASSICAL PSYCHOANALYTIC TREATMENT

Certain aspects of the therapeutic technique that Freud developed and that were later expanded by his followers are closely related with psychoanalytic theory. One of the distinctive aspects of the psychoanalytic approach to treatment in general is its consistent attempt to integrate therapeutic usages and approaches with the understanding of psychic functioning available from psychoanalytic theory. In its origins and clinical application, psychoanalysis is primarily a theory of therapy.

Analysis versus Analytic Psychotherapy

One of the chronically recurring issues among analysts is whether and to what extent psychoanalysis is or is not distinguishable from psychotherapy. There are distinguishing features between them as more or less pure forms: The use of a couch in analysis, not in therapy; free association as a primary method in analysis, not in therapy; intensive and long-term scheduling in analysis, not in therapy; emphasis on neutrality, abstinence, and interpretation on the part of the analyst in analysis, not in therapy; the central focus on transference and countertransference in analysis, not in therapy. Over the years, however, forms of psychotherapy have evolved, modifying all of these criteria and resulting in a spectrum of psychotherapeutic interventions ranging from psychoanalysis at one end to diluted forms of supportive psychotherapy at the other. The distinction between explorative versus supportive therapy parallels this continuum, so that many variants of the analytic process have arisen in which both components are employed in varying degrees. Some of this variation has come about by reason of the expansion of analytic techniques to the widening scope of psychopathology and the corresponding challenge of adapting analytic techniques to these patient needs. Another factor, however, has been the rejection of traditional analytic approaches and methods accompanying rejection of more traditional analytic theories. Some have questioned whether the therapeutic modifications introduced by relational and intersubjective approaches have altered the therapeutic interaction to an extent that it is no longer consistent with psychoanalytic principles. In any case, there is reason to think that a better means for determining which patients are better served by what forms of therapy needs to be developed.

Analytic Situation

The origins of Freud's approach to treatment have been discussed previously, emphasizing development of his basic techniques of free association and his growing awareness of and interpretation of the transference. Early in his approach to therapy, Freud felt that recognition by the physician of the patient's unconscious motivations, the communication of this knowledge to the patient, and its comprehension by the patient would in themselves effect a cure. This was his basic doctrine of therapeutic insight. Further clinical experience, however, has demonstrated the fallacy of these expectations.

Specifically, Freud found that interpretation of the patient's unconscious wishes and attitudes was often insufficient to induce any meaningful change in the patient. Such insight might permit clarification of the patient's intellectual appraisal of problems, but the emotional tensions for which the patient sought treatment were not effectively alleviated in this way. This discovery led to a significant breakthrough. Freud began to realize that the success of treatment depended in part on the patient's

ability to understand the emotional significance of his or her experiences on an emotional level and depended on the patient's capacity to put that insight to use in bringing about change. In that event, if the experience recurred, it would elicit another reaction; it would no longer be repressed, and the patient would have undergone an internal psychic change. Freud's formula for this process was: "Where id was, there ego shall be."

Freud thus elaborated a treatment method that attached minimal importance to the immediate relief of symptoms, to moral support from the therapist, or to guidance. The goal of psychoanalysis was to pull the neurosis out by its roots, rather than to prune off the top. To accomplish this, it was necessary to break down the pregenital, deep crystallization of id, ego, and superego and bring underlying material near enough to the surface of consciousness so that it can be modified and reevaluated in light of reality. This method distinguishes the classical psychoanalytic treatment from other psychodynamic forms of psychotherapy.

The patient is unaware of the repression of the forces of conflict and the psychic mechanisms of defense the mind uses. By isolating the basic problem, the patient has protected him- or herself against what seems, from the patient's view, to be unbearable suffering. No matter how it may impair functioning, the neurosis seems somehow preferable to the emergence of unacceptable wishes and ideas. All the forces that permitted the original repression are thus mobilized once again in the analysis as a resistance to this threatened encroachment on dangerous territory. No matter how much the patient may cooperate consciously with the therapist and in the analysis, and no matter how painful the neurotic symptoms may be, the patient automatically defends against reopening of old wounds with every subtle resource of defense and resistance available.

Analytic Process. In discussing the analytic therapy, a distinction is drawn between the *analytic process* and the *analytic situation*. The *analytic process* refers to the actual conduct of the analysis in which the regressive emergence, working through, interpretation, and resolution of the transference or transference neurosis takes place. The *analytic situation,* however, refers to the framework or setting in which the analytic process takes place, not only the time and place in which the analysis occurs, but more specifically the collaborative relationship between patient and analyst based on the therapeutic alliance.

The regression induced by the analytic situation (instinctual regression) allows for a reemergence of infantile conflicts and thus optimally induces formation of a transference neurosis. In the classical transference neurosis, the original infantile conflicts and wishes, originally directed toward the parents, become focused on the person of the analyst and are thus reexperienced and relived in the analytic process. In the analytic regression, earlier infantile conflicts are revived and can be seen as a

manifestation of the repetition compulsion. Regression has a dual aspect; from one point of view it is an attempt to return to an earlier state of real or fantasy gratification, but from another point of view it can be seen as an attempt to master previous traumatic experience. The regression in the analytic situation and the development of transference are preliminary conditions for the mastery of unresolved conflicts. They can also represent regressive and unconscious wishes to return to an earlier state of narcissistic gratification. The analytic process must work itself out in the face of this dual potentiality and tension.

If the analytic regression has a destructive potentiality (ego regression) that must be recognized and guarded against, it also has a progressive potentiality for reopening and reworking infantile conflicts and for achieving a reorganization and consolidation of the personality on a more mature and healthier level. As in any developmental crisis, the risk of regressive deterioration must be balanced against the promise of progressive growth and mastery. The essential component within the analytic situation, against which regressive pulls must be balanced and by which the destructive or constructive potential of the regression can be measured, is the therapeutic alliance. A firm and stable alliance offers a buffer against excessive (ego) regression and also offers a basis for positive growth.

The Analytic Relation. The analytic or therapeutic relation can be conceived as compounded of at least three components that are coexistent, mutually interacting and influencing, and intermingled at all points in the analytic process. Although constantly interacting to influence the patterns of interaction between analyst and patient and determining the course of the analytic process, they can be usefully distinguished in that they point to differentiable issues and aspects of the therapeutic process and call for different therapeutic responses and interventions. They are the *transference and countertransference,* the *therapeutic alliance,* and the *real relation*. In theorizing about the analytic process, many analysts make little use of, or simply ignore, or more often do not directly address but nonetheless imply the operation of one or the other of these components of the analytic relation, usually the therapeutic alliance or the real relation, but a convincing case can be made for their effective realization in all analyses. Thus, while transference and countertransference play an undoubtedly central—often dominant—role in the analytic process and interaction, the influences from the alliance and real relation remain continually active and should not be ignored.

TRANSFERENCE. Through free association, hidden patterns of the patient's mental organization, which may be fixated at immature levels and refer to events or fantasies in the patient's private experience, are brought to life

and activated in the relation with the analyst. In the simplest model of transference, the analyst is gradually invested with emotions usually associated with significant figures in the past. The patient displaces or projects feelings originally directed toward these earlier objects onto the analyst, who then is perceived alternately as a friend or enemy, one who is nice or frustrates needs and punishes, one who is loved or hated, and so on, as the original objects were loved or hated. Moreover, this tendency persists, so that to an increasing extent the patient's feelings toward the analyst replicate feelings toward the specific people being talked about or, more accurately, those about whom the patient's unconscious is talking. This transference object acts as a lens through which the patient views the analyst, seeing him or her in the image of the transference representation.

Analysts differ in regard to the degree to which emphasis is placed on the role of transference in analysis. Some insist that analysis is incomplete unless a full-blown transference neurosis is developed and ultimately resolved in the course of the analysis. Others feel that transference phenomena are to be expected, but that the development of a transference neurosis as such is neither necessary nor common. Some even are of the opinion that a transference may not develop at all. Others, especially relational and intersubjectivist theorists, dismiss the classic notion of transference as irrelevant and focus on the here-and-now interaction between analysts and analysand, disregarding for the most part residues of prior object relations experience and their impact on the present relation. In the relational approach, transference is thought to be cocreated by both analyst and patient, so that the resulting interaction says nothing about the patient's past history or experience of object relations and everything about the emerging pattern of relatedness with the analyst within the analysis.

In classical terms, as unresolved childhood attitudes and feelings emerge and begin to function as displacements or fantasied projections toward the analyst, he or she becomes for the patient a phantom composite figure representing various important persons in the patient's early environment or objects represented in his or her inner world. Those earlier relationships are reactivated with some of their original affective vigor, thus exposing in some degree the roots of the patient's disturbance. The concept of transference has undergone considerable elaboration over time, resulting in multiple variants, broadening its connotations to include every emotional connection to the analyst, and extending the transference model to encompass the widening range of psychopathology addressed by psychoanalysis. Some of the variations in transference and their descriptions are listed in Table 6.1–5. Understanding transferences requires some exploration of mechanisms involved in their formation and their dynamic interactions. The basic mechanisms by which transferences are effected—displacement, projection, and projective identification—are described in Table 6.1–6.

COUNTERTRANSFERENCE. If the patient is capable of transference in the analytic interaction, the analyst is correspondingly capable of countertransference, meaning that the analyst engages in the interaction with his or her own burden of elements coming from his or her own developmental past or that may be activated in the course of interaction with the patient, especially in response to the patient's transference. Originally countertransference was seen as a matter of a response in the analyst's unconscious affecting his or her view of and reaction to the patient, but recent views tend to see it as encompassing the total affective response of the analyst to the patient, whether conscious or unconscious, and as reflecting more responses aroused in the present interaction with the patient than influences coming from the analyst's past experience or unconscious.

Table 6.1–5.
Transference Variants

Libidinal Transferences

These follow the classic model and usually take the form of milder expression as positive *transference reactions;* but in more extreme expression can take the form of more intense and disturbing *erotic transferences.* They are generally derivatives of phallic-oedipal, libidinal impulses and may be permeated in various degrees by pregenital influences. They may occur with varying degrees of intensity, and in mild forms may not even require interpretation if they contribute to and support the therapeutic relation. Freud regarded these as "unobjectionable positive transferences." He also recommended that they call for interpretation only when they begin to serve as a resistance.

Aggressive Transferences

These take the form either of negative or more pathological paranoid transferences. *Negative transferences* are seen at all levels of psychopathology, but may predominate in some borderline patients, who tend to see the therapeutic relationship in terms of power and victimization, regarding the therapist as omnipotent and powerful, threatening and hurtful, while the patient experiences him- or herself as helpless, weak, and vulnerable. However, negative transferences are identifiable in varying degrees in all analyses and usually require specific intervention and interpretation. *Paranoid transferences* represent the extreme of negative distortion of the figure of the analyst, who is perceived as destructive and severely threatening; this is usually associated with psychotic reactions such that the transference becomes delusional.

Transferences of Defense

This form of transference is opposed to *transferences of impulse;* defenses against impulses find expression in the transference rather than the instinctual impulses themselves. In this form of transference, attention shifts from threatening drives or motives to the ego's defensive functioning, so that transference is no longer due merely to the repetition of instinctual pressures but includes aspects of ego functioning as well.

Transference Neurosis

This degree of transference involvement implies that a recreation or more ample expression of the patient's infantile neurosis is reenacted anew within the analytic relation, thus at least theoretically mirroring aspects of the infantile neurosis. The transference neurosis usually develops in the middle phase of analysis, when the patient, at first eager for improved mental health, no longer consistently displays such motivation but engages in a continuing battle with the analyst over the desire to attain some kind of emotional satisfaction from the analyst, so that this becomes the most compelling reason for continuing analysis. At this point of the treatment, the transference emotions become more important to the patient than alleviation of distress sought initially, and the major, unresolved,

unconscious problems of childhood begin to dominate the patient's behavior. They are now reproduced in the transference, with all their pent-up emotion.

The transference neurosis is governed by three outstanding characteristics of instinctual life in early childhood: The pleasure principle (before any effective reality testing), ambivalence, and repetition compulsion. Emergence of the transference neurosis is usually a slow and gradual process, although in certain patients with a propensity for *transference regression,* particularly more hysterical and borderline patients, elements of transference and transference neurosis may manifest themselves relatively early in the analytic process. One situation after another in the life of the patient is analyzed and progressively interpreted until the original infantile conflict is sufficiently revealed. Only then does the transference neurosis begin to subside. At that point termination usually begins to emerge as a more central concern.

Contemporary opinion is divided as to the importance and centrality of the transference neurosis, whether it forms to the extent Freud thought, and whether it is necessary for successful analysis—for some it remains an essential vehicle for analytic interpretation and therapeutic effectiveness; for others it may never develop or, to the extent that it does, may play a less central role in the process of cure.

Transference Psychosis

This occurs rarely in ordinary analytic experience, but can develop when failure of reality testing leads to loss of self-object differentiation and diffusion of self and object boundaries. This may reflect an attempt to refuse with an omnipotent object, investing the self with omnipotent powers as a defense against underlying fears of vulnerability and powerlessness. Transference psychosis may also include negative transference elements in which fusion carries the threat of engulfment and loss of self that may precipitate a *paranoid transference reaction.* In this extreme form of transference, the as-if quality of transference experience is eroded so that the analyst is no longer seen as like or reminiscent of prior transference figures, but is experienced as though he or she was the hated or feared object from the past.

Narcissistic Transferences

These were clarified in 1971 by Kohut as variations of patterns of projection of archaic narcissistic configurations onto the therapist, both superior and inferior—the superior form reflecting narcissistic superiority, grandiosity, and enhanced self-esteem, and the inferior the opposite qualities of inferiority, self-depletion, and diminished self-esteem. The therapist comes to represent, in Kohut's terms, either the grandiose self in *mirror transferences* or the idealized parental imago in idealizing transferences. In idealizing transferences all power and strength are attributed to the idealized object, leaving the subject feeling empty and powerless when separated from that object. Union with the idealized object enables the subject to regain narcissistic equilibrium. *Idealizing transferences* may reflect developmental disturbances in the idealized parent imago, particularly at the time of formation of the ego ideal by introjection of the idealized object. In some individuals narcissistic fixation leads to development of the grandiose self. Reactivation in analysis of the grandiose self provides the basis for formation of mirror transferences, which occur in three forms: *Archaic merger transference,* a less archaic *alter-ego or twinship transference,* and *mirror transference in the narrow sense.* In the most primitive merger transference, the analyst is experienced only as an extension of the subject's grandiose self, and thus becomes the repository of the patient's grandiosity and exhibitionism. In the alter-ego or twinship transference, activation of the grandiose self leads to experience of the narcissistic object as similar to the grandiose self. In the most mature form of mirror transference, the analyst is experienced as a separate person, but nonetheless one who becomes important to the patient and is accepted by him or her only to the degree that the analyst is responsive to the narcissistic needs of the reactivated grandiose self.

Selfobject Transferences

These represent extensions of the self-psychology selfobject paradigms to the figure of the analyst beyond merely narcissistic configurations. The selfobject involves investment of the self in the object so that the object comes to serve a self-sustaining function that the self cannot perform for itself—either in maintaining fragile self-cohesion or in regulating self-esteem. The other is thus not experienced as an autonomous and separate object or agency in its own right, but as present only to

serve the needs of the self. Transference in this sense reflects a continuing developmental need that seeks satisfaction in the analytic relation.

Selfobject transferences reflect the underlying need structure the patient brings to the therapeutic relationship, based on the predominant pattern of selfobject deprivation or frustration and the corresponding seeking for the appropriate form of selfobject involvement. These configurations have been described as the *understimulated self,* the *overstimulated self,* the *overburdened self,* and the *fragmenting self.* Other descriptions of selfobject need translate patterns of transference interaction based on narcissistic dynamics into the perspective of the relationship between self and selfobject, as in *mirror-hungry personalities* and *ideal-hungry personalities.* Variations on the mirroring transference theme include the *alter-ego-hungry personality,* the *merger-hungry personality,* and in contrast, the *contact-shunning personality.* In transferences derived from such personality configurations, the classical meaning of transference has undergone radical modification. Rather than displacements or projections from earlier object relational contexts, the patient brings to bear a need based in his or her own currently deficient capacity and defective character structure—a need to involve the object in a dependent relationship in order to complete or stabilize his or her own psychic integration.

Transitional Relatedness

This transference model is based on Winnicott's notion of the transitional object. Transference in more primitive character structures is regarded as a form of transitional phenomenon, the *transitional object relation,* in which the therapist is perceived as outside the self, but is invested with qualities from the patient's own archaic self-image. The transference field in this view is envisioned as a transitional space in which the transference illusion is allowed to play itself out.

Transference as Psychic Reality

This form of transference reflects the need of each participant in analysis to draw the other into a stance corresponding to his or her own intrapsychic configuration and needs, as a reflection of the individual subject's psychic reality. On these terms, the classic view of transference, based on displacement or projection from past objects, is regarded as inadequate, resulting in further diffusion of the meaning of transference as equivalent to the individual's capacity to create a meaningful world or to inform the world with meaning. In this rendition, transference becomes equivalent to the patient's psychic reality, so that any distinction between the meanings given to reality and the meanings inherent in transference are lost. The specificity of derivatives from past developmental history is lost. Transference in these terms becomes all-encompassing, and whatever distinguishing and dynamic significance it may have had fades into obscurity. Dynamic or defensive parameters are simply those involved in the individual's current psychic reality. In this form of transference, there does not seem to be any definable mechanism at work, other than whatever is involved in the subject's psychic reality. The subject's view of his or her environment and impression of objects of his or her experience, including the analytic object, are indistinguishable from ordinary cognitive and affective processes characterizing the subject's more general involvement and responsiveness to the world.

Transference as Relational or Intersubjective

The relational or intersubjective view of transference as emerging from or cocreated by the intersubjective interaction between analyst and analysand transforms transference into an interactive phenomenon in which individual intrapsychic contributions from either participant are obscured. As cocreated the transference can be attributed to neither party. Transference in this sense is not anything individual to or intrapsychically derived from the patient, but is based on the present ongoing interaction between analyst and patient coconstructing transference. On these terms, analysis of transference has little to do with past derivatives and everything to do with the ongoing relation with the analyst primarily in the form of interpersonal enactments. Transference in this sense is no longer a one-person phenomenon, but reflects a two-person transference–countertransference interaction. The supposition is that there is no such thing as transference without countertransference and no such thing as countertransference without transference. The patient is thus relieved of any burden of a personal dynamic unconscious reflecting developmental vicissitudes and residues of a life history. Transference is created anew in the immediacy of present

analytic interaction as the product of mutual influence and communication between analyst and analysand, probably relying on some form of mutual projective identification to sustain the interactive connotation.

Views of countertransference are somewhat divided. The view of countertransference as total, largely inspired by the Kleinian orientation, tends to focus the understanding of the analytic process more or less completely and exhaustively on transference in the patient and countertransference in the analyst and the interaction between them. Thus, all reactions and responses of the analyst to the patient are included under the rubric of countertransference. Although this effectively covers a major portion of the analytic relation and interaction, it also tends to ignore, minimize, or neglect other aspects of the analytic relation, as though these other dimension were insignificant, unimportant, and hardly worthy of exploration and understanding. The assumption seems to be that these other aspects can be presumed or ignored and that they have little impact on the central issues related to transference and countertransference. A contrary view is that transference and countertransference are important aspects of the analytic interaction, but they are only part of the picture, and that other forms of interrelation and interaction take place along with the transference phenomena and have important influences on how transferential components are expressed and on what terms they can be effectively worked with in the analytic process. These other aspects would include both real relation and therapeutic alliance.

Where patient transference and analyst countertransference are caught up in an interaction the result is a transference–countertransference interaction. Early views of countertransference saw it as interfering in the work of the analysis, as it may often do, but recent revisions have emphasized the possible positive contributions to more effective analytic work arising from attention to and utilization of countertransference responses in the course of an analysis. Countertransference has thus become regarded as inevitable and as a possible form of communication within the analytic process, not necessarily destructive to the analytic process. The author's opinion is that they are useful when they can be detected insofar as they reveal unconscious factors that would otherwise remain hidden, but that their therapeutic management and correction cannot be achieved through countertransference as such, but only through the effective use made of it from the vantage point provided by the therapeutic alliance. If a therapist, for example, were to find him- or herself annoyed at a patient, it would not be therapeutically useful to express or act out the annoyance on the patient. Rather it would be useful to analyze the possible sources of the anger in his or her past experience and find a constructive way to deal with it in relation to the patient.

THERAPEUTIC ALLIANCE. The therapeutic alliance is based on the one-to-one collaborative relationship that the patient establishes in interaction with the analyst. This interaction deals with those aspects of the therapeutic relation that enable patient and analyst to engage meaningfully and productively in the analytic process with the objective of achieving therapeutic benefit for the patient. The terms of the alliance are negotiated between analyst and patient; obviously, not any terms of their working together will do, but only those that can predictably contribute to or set the stage for their effectively working together. The alliance on these terms would include at least the following elements: Empathy, trust, autonomy, responsibility, authority, freedom, honesty, and neutrality. All these elements are as pertinent to the role of the patient in analysis as to the analyst.

Table 6.1–6.
Transference Mechanisms

Displacement

The basic mechanism of classic transference paradigms in which an object representation or representations derived from any level or combination of levels of the subject's developmental experience are displaced to the representation of the new object, namely the analyst, in the therapeutic relationship. Displacement is the usually the basic mechanism for libidinally based transferences, both positive and erotic, as well as for aggressive and especially negative transferences. By and large, displacement transferences tend to play a more dominant role in neurotic disorders, in which phallic-Oedipal (and to a lesser degree pre-Oedipal) dynamics tend to play a dominant, though not exclusive, role.

Projection

Projection is the process by which qualities or characteristics of the self, specifically the self-as-object, usually involving introjections or self-representations, are attributed to the object representation of an external object, and the subsequent interaction with the object is determined by the projected characteristics. Thus the analyst-object may be seen as sadistic, that is as possessing the sadistic character of the analysand-subject, an aspect of the subject's self that is denied or disowned by the subject. Projection tends to play a more prominent, though again not exclusive, role in formation of transferences in more primitive character disorders, but can be found in variously modified forms throughout the spectrum of neuroses. Since projections derive primarily from the configuration of introjects constituting the patient's self-as-object, the effect of projective or externalizing transferences is that the image of the therapist comes to represent part of the patient's own self-organization rather than simply an object representation as such.

Projections derived from destructive introjects can provide the basis for both negative and paranoid transference reactions. Those based on the victim-introject result in the patient relating to the therapist as a victim and him- or herself assuming a correspondingly hostile or sadistic position as a destructive aggressor or victimizer in relation to the therapist's role as victim. Then again, projection based on the aggressor-introject results in the patient relating to the therapist as an aggressor and him- or herself assuming the role of a weak, vulnerable, or masochistic position in which he or she becomes a passive and vulnerable victim of the therapist's destructive aggression. Similar patterns can take place around narcissistic issues involving introjective configurations of narcissistic superiority and inferiority. From the narcissistically superior perspective, the transference takes the form of seeing the analyst as superior, idealized, valuable, and exceptionally gifted or intelligent, while the patient retains within him- or herself the image of narcissistic inferiority, worthlessness, inadequacy, and diminished self-esteem. The opposite configuration obtains from the narcissistically

inferior perspective—the analyst is then seen in the light of the projected image as inferior, inadequate, worthless, and so on, while the patient retains a view of him- or herself as superior, valued and valuable, narcissistically enhanced, even grandiose and exceptional.

However, projective dynamics in selfobject transferences seem to involve more than narcissistic projections, since these forms of transference tend to draw the analyst into meeting the pathological needs of the self. If anything is projected, it would be an infantile wished-for imago, one lacking earlier in the patient's experience, as for example an empathic and idealized parental figure. On the other hand, transitional transferences, despite their considerable overlap with selfobject phenomena, tend to involve a more explicit projective element as the self-related contribution to the transitional experience.

Projective Identification

The concept of projective identification was first proposed by Melanie Klein, arguing that the projection of impulses or feelings into another person brought about an identification with that person based on attribution of undesirable dimensions of one's own qualities to that other. This attribution also served as the basis for a sense of empathy and connection with the other. On these terms, projective identification was a fantasy taking place solely in the mind of the one projecting.

Projective identification is often appealed to as a mechanism of transference, or more exactly transference–countertransference interactions. Confusion arises from the failure to clearly distinguish between projection and projective identification. The notion of projective identification added to the basic concept of projection the further notes of diffusion of ego boundaries, a loss or diminishing of self-object differentiation, and inclusion of the self as part of the object.

Later elaborations of the notion of projective identification transformed it from a one-body into a two-body phenomenon, describing interaction between two subjects, one of whom projects something onto or into the other, whereupon the other introjects or internalizes what has been projected. Instead of the projection and introjection taking place in the same subject, the projection now takes place in one and the internalization in the other. This latter usage has led to extensive extrapolation of the concept of projective identification to apply to object relations of all sorts, including transference. The emphasis in Kleinian transferences is less on the influence of the past on the present, but rather the influence of the internal world on the external in the here-and-now interaction with the analyst.

It is worth noting that there can be important clinical differences between displacement and projective transferences, especially when the latter involve projective identification. Generally, in displacement transferences, the emphasis is on the patient's experience of the analyst object as like a previous object. The process is by and large internal with little in the way of external influences to draw the object into conformity with the image. Projective transferences have a different quality, in that part of the patient's self is experienced as in the other. This calls for differences in therapeutic handling and interpretation. Moreover, any interpretation without substantial evidence pointing to the presence of the projected qualities within the patient's self would be less than optimal. When projective identification is in question, the same kind of evidence is required, but in addition the projection carries with it the effort to influence the analyst to unconsciously internalize and thus enact the role that the patient seeks to impose on him or her. This can and often does elicit some form of countertransference reaction in the analyst, resulting in a transference–countertransference interaction, which can complicate and even derail the analytic relation and process.

The therapeutic alliance allows a split to take place in the respective selves of both patient and analyst—that is, the observing part of the patient's ego functions allow him or her to ally him- or herself with the analyst in a working relationship, which opens the way to gradually identifying positively with the analyst in the analyzing function and thus modifying pathological defenses put up by the defensive ego against the anxiety aroused by internal danger situations. Similarly, the split in the

analyst allows him or her to engage affectively and empathically with the patient, while at the same time preserving an observant, reflective, and neutral analytic stance from which he or she is able to understand and effectively interpret the meaning of the patient's experience and behavior.

On the part of the patient, maintenance of this therapeutic split, as well as the relationship to the analyst involved in the therapeutic alliance, requires maintenance of self-object differentiation, tolerance and mastery of ambivalence, and the capacity to distinguish fantasy from reality in the relationship. In many analyses, these qualities are not readily available, so that consequently, the alliance requires work and effort to establish and can thus serve as one of the objectives of the analytic work. In no case can the alliance be taken for granted or assumed, since the propensity of all patients to create various subtle forms of misalliance is pervasive. If they are not carefully looked for and attuned to, such misalliances can easily distort the course of an analysis, only becoming apparent when they reach a point of crisis or impasse. In more severely disturbed personalities, there is a greater tendency to outright disruptions of the alliance, rather than misalliances, which can destroy an analytic process and often require extreme efforts to salvage the therapy.

Maintenance of the therapeutic alliance requires that the patient be able to differentiate between the more mature and more infantile aspects of his or her experience in relationship to the analyst. The therapeutic alliance serves a double function. On one hand, it acts as a significant buffer to excessive regression of the ego in the analytic process; on the other hand, it serves as a fundamental aspect of the analytic situation, against which the wishes, feelings, and fantasies evoked by the transference and transference neurosis can be evaluated, measured, and interpreted. In many pathological conditions—some character neuroses, borderline personalities, and more severe neurotic disorders—it may be difficult to maintain a clinical distinction between therapeutic alliance and the transference neurosis.

The therapeutic alliance derives from the mobilization of specific ego resources relating to the capacity for object relations and reality testing. The analyst must direct attention toward eliciting the patient's capacity to establish such a relationship that will be able to withstand the inevitable distortions and regressive aspects of the transference and transference neurosis. It is inevitable that the fundamental features of the therapeutic alliance be carefully evaluated and understood and ultimately integrated with the analysis of the transference. When therapeutic misalliances arise, as they inevitable will, they can become matters of inquiry and interpretation—more often than not, they reflect unconscious and hidden transference dynamics that otherwise have not been previously recognized or explored in the analysis. Analysts should remain alert to such misalliances in that they provide often subtle indications that something is

awry in the analytic work that will both require readjustment and offer the potentiality for deeper transference exploration.

REAL RELATION. Reality pervades the analytic relationship. On one hand, there is the reality of the personalities and characteristics of analyst and analysand that distinguishes them as individual and unique human personalities and that they bring with them to the analytic encounter. On the other hand, there are the realities of time, place, and circumstance external to the analytic setting that constantly influence the course of the analytic relation. These include realities of the location of the analyst's office, the physical surroundings, the furniture and decorations in the room, the geographic location itself, and even how the analyst dresses; they all have their effects in the analytic situation and influence how the patient experiences the person of the analyst. The surrounding circumstances that create the framework for the analytic effort—the patient's financial situation, whether married or not, job demands, arrangements for payment of the fee, whether the patient has insurance or not and what kind, what kinds of pressures are pushing the patient into treatment, accidental factors like illness, interfering family, or business obligations—these are all reality factors extrinsic to the analysis but exercise significant influence on the analytic relationship and how it is established and maintained.

The most important and central reality for the patient in the analytic situation is the person of the analyst. Every analyst has his or her own constellation of personal characteristics, including mannerisms, style of behavior and speech, habits of dress, gender, way of going about the task of managing the therapeutic situation, attitudes toward the patient as a human being, prejudices, moral and political views, and personal beliefs and values. These are all relevant aspects of one's real existence and personality as a human being. They are realities that play a role in the therapeutic relationship and are entirely distinct from transference and countertransference and the alliance. In terms of the analytic process, none of this is lost on the patient who is comprehensively curious about, observant of, and sensitive to the smallest details of the analyst's real person and life. The same considerations operate from the side of the analyst in relation to the patient. Important aspects of the reality of the patient's person and life situation play a decisive role in the analyst's evaluation of analyzability and continue to present themselves throughout the course of the analysis as subjects of analytic concern, inquiry, and exploration since they contribute to and affect the patient's ongoing life and experience and his or her participation in the work of the analysis. If a patient, for example, is having difficulties in his or her marriage, this must have important reverberations for what happens in the analysis.

Technical Aspects. Analytic technique is always adapted to the

idiosyncrasies of the patient's developmental capacities, needs, conflicts, compromises, and defensive constellation. Analytic techniques do not stand in isolation, but are part of a living, dynamic process that is intended to induce and achieve significant internal psychic growth. A few of the more salient techniques are as follows.

FREE ASSOCIATION. The cornerstone of the psychoanalytic technique traditionally is free association. As many have noted, the process of associating is not altogether free, but is to a certain degree guided by the expectations of the analytic setting and the analyst's responses toward more meaningful or affectively loaded or conflictual material. The patient is encouraged to use this method as far as possible throughout the treatment. The primary function of free association, besides obviously providing content for the analysis, is to help induce the necessary regression and relatively passive dependence connected with establishing and working through of the transference. Thus, free association is conjoined with the other techniques that induce such regression, namely lying on the couch, not being able to see the analyst, and conducting the analysis in an atmosphere of quiet reflection and restful tranquility.

One also cannot simply regard the process of free association as something that takes place in isolation in the patient. In fact the process is more complex, more difficult to conceptualize, and increasingly must be seen in the context of and in reference to the more fundamental relationship between analyst and patient. The patient's free associating is a function of the more basic relationship and reflects the complex influences arising from transference, the real relation, and the alliance. Moreover, it is increasingly clear from a contemporary perspective that much more is required of a patient than simply free associating. It is not enough for the patient to lie back and allow self-surrender to a position of passive dependency within the analytic relationship, without at the same time being able to mobilize basic ego resources in the service of mastery, gaining insight, mobilizing executive and synthetic capacities, and ultimately being able to assume a less passive and more active and autonomous function within the analytic relationship. This is a reflection of the analytic split discussed previously. Obviously, there is a gradation in the mobilization of these capacities in the patient, which varies from phase to phase of the analytic process, optimally leading to more mature and adaptive levels as the analysis progresses, especially in the progression to termination.

RESISTANCE. The most conscientious efforts on the part of the patient to say everything that comes to mind and to engage in the work of the analysis are never completely successful. No matter how willing and cooperative the patient may be consciously in attempting to free associate, the signs of resistance are apparent throughout the course of every analysis. The

manifestations of resistance are protean. The patient pauses abruptly, corrects him- or herself, makes a slip of the tongue, stammers, remains silent, fidgets with some part of clothing, or asks irrelevant questions, intellectualizes, arrives late for appointments, finds excuses for not keeping them, offers critical evaluations of the rationale underlying the treatment method, simply cannot think of anything to say, or even censors thoughts that do occur and decides that they are banal or uninteresting or irrelevant and not worth mentioning.

The development of resistance in the analysis is quite as automatic and independent of the patient's will as the development of the transference itself. The sources of resistance are just as unconscious as the sources of transference. The emotional forces, however, that give rise to resistance usually are defending against or expressing those that produce the transference. Thus, resistances tend to emerge more in the middle phase of the analysis, in which regressive emergence of the transference is a central concern. The analysis becomes a recurring field of conflict between the tendencies toward transference and those toward resistance, manifested by the involuntary inhibition of the patient's efforts to associate freely. This inhibition may last for moments or days or may persist throughout the whole course of the analysis.

Resistance may take place in all phases of the analysis, but its quality and significance are different, depending on the analytic task at hand. In any case, the patient's resistance enables the analyst to evaluate and become familiar with the defensive organization of the patient's ego. In this way the pattern of resistance not only offers valuable information to the analyst about the patient's basic conflicts, but also offers a channel by which the patient can be approached therapeutically.

The significance of the conflicts underlying the resistance is clear. It is a repetition of the very same sexuality-guilt conflict that originally produced the neurosis itself. Transference may itself take the form of resistance, in that the wish for immediate gratification in the analysis can circumvent and postpone essential goals of treatment. Consequently, the analysis of resistance, particularly transference resistance, is one of the analyst's primary functions. It also accounts in many cases for the extended time period required for successful psychoanalytic treatment. No matter how skillful the analyst, resistance is never absent and can often absorb significant amounts of energy and time.

In the light of relational and intersubjective perspectives on the analytic process, the concept of resistance has fallen out of favor in that any such phenomenon, viewed as coming from the internal dynamics and issues of conflict-and-defense in the patient, is rejected but is rather seen as a byproduct of the interaction between analyst and patient. There is then no resistance coming from the patient, but it is cocreated and thereby contributed to by both participants. It can only be dealt with by examining

the interaction causing it and not by interpretation of the patient's defenses. This point of view remains highly controversial.

INTERPRETATION. Interpretation was long regarded as the chief tool of the analyst in efforts to reduce unconscious resistance and in transference analysis. Current views tend to give it a more modest role, as one important technical tool among a variety of other influences, including aspects of the analytic relation and the quality of the analytic interaction. As mentioned earlier, in the early stages of the development of psychoanalytic therapeutic techniques, interpretation was used primarily to inform the patient of his or her unconscious wishes. Later, it was designed to help the patient understand the resistance to spontaneous self-awareness. In current psychoanalytic practice, the analyst's function as interpreter is not limited to simply paraphrasing the patient's verbal reports, but rather to indicating at appropriate moments what is not reported or is implicit in what is reported. Consequently, as a general rule, analytic interpretation does not produce immediate symptomatic relief. On the contrary, even after correct and clarifying interpretations, there may be a heightening of anxiety and an emergence of further resistance.

If a correct (mutative) interpretation is given with proper timing, the patient may react either immediately or after a period of emotional struggle, during which new associations are offered. These new associations often confirm the validity of previous interpretations and add significant additional data, thus disclosing motivations and experiences of the patient of which the analyst could not previously have been aware. Generally speaking, it is not so much the analyst's insight into the patient's psychodynamics that produces progress in the analysis as it is the patient's ability to gain his or her own insight, whether independently or in collaboration with the analyst; the analyst can facilitate this process by reducing unconscious resistance to such self-awareness through appropriate, carefully timed interpretations. The most effective interpretation is timed so that it is given by the analyst in such a way as to meet the emerging, if hesitant and half-formed, awareness of the patient. Thus, the analyst must gauge the capacity of the patient at any given moment to hear, assimilate, and integrate the content of a given interpretation.

Another important aspect of an interpretation is that it cannot be seen in isolation from the total context of the analytic situation and the analytic process. An interpretation, both as given by the analyst and as received by the patient—which may focus on elements of either transference or therapeutic alliance or both—takes place within the context of the therapeutic relationship. Thus, the giving and receiving of interpretations are cloaked with a series of meanings that unavoidably influence both the capacity of the patient to accept and integrate interpretations and the

analyst's sense of offering and providing such interpretations. Experience has shown that, at best, the therapeutic benefits produced by virtue of the analyst's explanations or unilaterally provided insights are only temporary. Those interpretations that are most effective and of lasting therapeutic value are those arrived at by the delicate dialectic arising from the mutually facilitated and growing awareness of both patient and analyst.

Through experience, it has been learned that interpretations are best offered in a spirit of inquiry, as hypotheses suggested for the patient's consideration and remaining open to his or her further consideration, acceptance, questioning, modifying, or rejecting. Interpretations, in a sense, are synonymously invitations for the patient to exercise his or her own curiosity, reasoning, questioning, and judgment as quasiexperiments in autonomous self-reflection and analysis. Any interpretation the analyst might offer is no more than a piece of sophisticated guesswork about what the patient is thinking or meaning. Contrary to what some analysts may have thought, the analyst cannot read the patient's mind. Thus, interpretations can have little meaning or effect unless they are scrutinized, evaluated, and found to have application or relevance that the patient can acknowledge and accept. This cannot be achieved without the patient's active participation in the process of inquiry, discovery, and understanding of his or her own unconscious thinking and motivation.

MODIFICATIONS IN TECHNIQUES. There are no shortcuts in psychoanalytic treatment, although there have been enough attempts to devise them. Psychoanalytic treatment typically extends over a period of years and requires interminable patience on the part of both analyst and patient. Rigid adherence, however, to the fundamental principles of psychoanalytic technique is an impossibility. For example, the immediate environmental situation may be so serious for the patient that the analyst must pay common-sense attention to its practical implications. Those patients whose early childhood was extraordinarily deficient in love and affection, so that they suffer from a basic developmental defect in their capacity for one-to-one relationships and, consequently, in their capacity to sustain a therapeutic alliance, must be given more support and encouragement than usually advocated by psychoanalytic technical principles.

The analyst's role in the early stages of analysis in helping to establish the therapeutic alliance is of particular importance. As noted, with the primitive patients the establishment of a therapeutic alliance can be the more significant aspect of the treatment process and can even persist as a problem through most of the analysis. Even so, establishment of the therapeutic alliance for most patients is a significant aspect of the analytic process. Often serious problems in the subsequent stages of analysis of the transference can be due to a failure to establish a meaningful alliance in the initial stages of treatment. Thus, suitable interventions of the analyst in the

early stages of treatment can be of help to the patient in establishing such a meaningful therapeutic alliance.

Patients who are more borderline or very narcissistic must establish a strong personal tie and strong feelings of attachment and relationship with the analyst before they can develop sufficient interest and motivation for treatment. Moreover, such a strong object tie with the analyst for these more primitive patients is an absolute necessity if the destructive effects of excessive regression are to be avoided. Development of sufficient trust is also essential for these patients if they are to establish any meaningful alliance. These are difficult problems, however, because experience also suggests that every deviation from analytic technique that such special conditions compel tends to prolong the length of treatment and to considerably increase its vicissitudes and problems.

Such modifications in analytic technique were once regarded, somewhat pejoratively, as "parameters," but they remain a considerable source of discussion and controversy among analytic therapists. A significant trend today is the increasing tendency of analysts to treat more difficult and complex cases; thus the necessity for introducing modifications in various aspects of the treatment process correspondingly increases. As a result, what might previously have been thought of as parameters are increasingly accepted as valid technical practices. Many practices that were traditionally regarded as parameters and generally frowned upon by classical analysts are now being recommended and advocated by interpersonal, relational, and intersubjective analysts as a result of their revised view of the analytic process and interaction as focused in the here-and-now and as attempting to balance the analytic equation by casting both analyst and patient is similar or equivalent roles. Thus, for example, in extreme form, if the patient is expected to be open and self-disclosing, similar expectations should apply to the analyst as well—an echo of Sándor Ferenczi's experiments in mutuality. The resolution of such tensions and difficulties in assessing and exploring modifications of techniques must ultimately rest on the basis of clinical experience.

Results of Treatment. The therapeutic effectiveness of psychoanalysis has in past years presented problems in its evaluation. Impartial and objective evaluation in popular terms are handicapped by the fact that so many patients state that they have been analyzed when no such procedure was, in fact, undertaken or when it was undertaken by someone who used the title of analyst and who, in fact, had little understanding of analytic science and technique. Other patients have been in analysis only for a very short time and then discontinued treatment on their own initiative or were advised they were not suitable candidates for analytic treatment. Except for psychoanalysts themselves, professionals as well as lay people demonstrate varying degrees of confusion as to what psychoanalysis is and what it is not.

In current years, however, the picture has changed dramatically. A series of important and sophisticated research evaluations of therapeutic effectiveness have developed meaningful studies of analytic outcomes that show the considerable advantages of analytic therapy and with which patients analytic methods can most usefully be applied. Outcome studies seem to indicate that when analysis is conducted with sufficient intensity and duration, and with appropriate patients, the outcomes are at least as good as, if not better than, comparative therapeutic techniques. In more recent studies, many of the impediments and conceptual and methodological difficulties that plagued earlier studies have been corrected, so that the results can be taken as more valid and with greater confidence.

Despite the methodological difficulties and complexities of such outcome studies, extensive empirical evaluations from a number of centers have demonstrated the effectiveness and relative success of psychoanalysis and psychoanalytic therapy for appropriately selected cases in the range of psychoneurotic conditions, personality disorders, and forms of self-pathology. Therapeutic outcomes in cases of psychosomatic illness, more primitive levels of personality disorder, and psychoses have been more guarded.

Even so, it is clear that no analyst can ever eliminate all the personality defects and neurotic factors in any given patient, no matter how thorough or successful the treatment. One can look for certain outcomes in successful treatment. Mitigation of the rigors of a punitive superego would seem to be an essential criterion of the effectiveness of treatment. Psychoanalysts do not usually regard alleviation of symptoms as the most significant aspect in evaluating therapeutic change. The absence of a recurrence of the illness or a further need for psychotherapy is perhaps a more important index of the value of psychoanalysis. The chief basis of evaluation, however, remains the patient's general adjustment to life; that is, the capacity for attaining reasonable life satisfaction, for contributing to the happiness of others, the ability to deal adequately with the normal vicissitudes and stresses of life, and the capacity to enter into and sustain mutually gratifying and rewarding relationships with other people in the patient's life.

More specific criteria of the effectiveness of treatment include the reduction of the patient's unconscious, neurotic need for suffering; reduction of neurotic inhibitions; decrease of infantile dependency needs; and an increased capacity for responsibility and for successful relationships in marriage, work, and social relations. Other important criteria are the capacity for pleasurable and rewarding sublimation and for creative and adaptive application of the patient's own potentialities. The most important criterion of the success of treatment, however, is the release of the patient's normal potentiality, previously blocked by neurotic conflicts, for further internal growth, development, and maturation to mature personality functioning.

REFERENCES

Diamond D, Blatt SJ, eds. Attachment research and psychoanalysis. 1. Theoretical considerations. 2. Clinical implications. *Psychoanal Inquiry*. 1999;19:4.

Diamond D, Blatt SJ, Lichtenberg J. Attachment research and psychoanalysis. 3. Further reflections on theory and clinical experience. *Psychoanalytic Inquiry*. 2003;23:1.

Erikson EH. *Childhood and Society*. New York: Norton; 1963.

Fenichel O. *The Psychoanalytic Theory of Neurosis*. New York: Norton; 1945.

Freud A. *The Writings of Anna Freud*. 7 vols. New York: International Universities Press; 1965–1974.

*Freud S. *The Standard Edition of the Complete Psychological Works of Sigmund Freud. 24 vols*. London: Hogarth Press; 1953–1974.

Gay P. Freud. *A Life for Our Time*. New York: Norton; 1988.

Goodman G. *The Internal World and Attachment*. Hillsdale, NJ: Analytic Press; 2002.

*Greenberg JR, Mitchell SA. *Object Relations in Psychoanalytic Theory*. Cambridge, MA: Harvard University Press; 1983.

Hartmann H. *Essays on Ego Psychology*. New York: International Universities Press; 1954.

Hartmann H. *Ego Psychology and the Problem of Adaptation*. New York: International Universities Press; 1958.

Hinshelwood RD. *A Dictionary of Kleinian Thought*. London: Free Association Books; 1991.

Kohut HS. *The Analysis of the Self*. New York: International Universities Press; 1971.

Kohut HS. *The Restoration of the Self*. New York: International Universities Press; 1977.

Kohut HS. *How Does Analysis Cure*? Chicago: University of Chicago Press; 1984.

*Laplanche J, Pontalis J-B. *The Language of Psycho-analysis*. New York: Norton; 1973.

Loewald HW. *Papers on Psychoanalysis*. New Haven, CT: Yale University Press; 1980.

*Mahler MS, Pine F, Bergman A. *The Psychological Birth of the Human Infant*. New York: Basic Books; 1975.

Meissner WW. *The Therapeutic Alliance*. New Haven, CT: Yale University Press; 1996.

Meissner WW. *The Ethical Dimension of Psychoanalysis—A Dialogue*. Albany, NY: State University of New York Press; 2003.

Mitchell SA. *Relational Concepts in Psychoanalysis*. Cambridge, MA: Harvard University Press; 1988.

Person ES, Cooper AM, Gabbard GO, eds. *Textbook of Psychoanalysis*. Washington, DC: American Psychiatric Publishing; 2005.

Richards AD, Tyson P, eds. Psychoanalysis, development and the life cycle. *J Amer Psychoan Assn*. 2000;48(4):1045–1618.

Richards AD, Tyson P, eds. The psychology of women: Psychoanalytic perspectives. *J Amer Psychoan Assn*. 1996;44(4):i–555

Rizzuto AM, Meissner WW, Buie DH. *The Dynamics of Human Aggression: Theoretical Foundations, Clinical Applications*. New York: Brunner-Routledge; 2004.

Schore AN. *Affect Regulation and the Repair of the Self: The Neurobiology of Emotional Development*. Hillsdale, NJ: Erlbaum Associates; 1994.

Shapiro T, Emde RN, eds. Research in psychoanalysis: Process, development, outcome. *J Amer*

Psychoan Assn. 1993;41(Suppl):1–424.

*Stern D. *The Interpersonal World of the Infant*. New York: Basic Books; 1985.

Stolorow R, Atwood G. *Contexts of Being: The Intersubjective Foundations of Psychological Life*. Hillsdale, NJ: Analytic Press; 1992.

Tyson P, Tyson RL. *Psychoanalytic Theories of Development*. New Haven, CT: Yale University Press; 1990.

▲ 6.2 Erik H. Erikson

DORIAN S. NEWTON, PH.D.

"I am what survives of me."

Erik H. Erikson was an artist, teacher, anthropological field worker, and psychoanalyst. His legacy includes an original and authoritative theory of individual psychosocial development with cross-cultural and universal applications. Previous conceptions of personality development had been centered on infancy and childhood; Erikson's extended across the life span and was anchored in familial, societal, and historical contexts. His was a model of human experience which emphasized adaptation and growth over fixation and pathology.

A prolific writer, Erikson is best known for three major works. In *Childhood and Society*, Erikson advanced a theory of the life cycle which traced the trajectory of individual psychosocial development, through the maturing ego's relations with an expanding social world. In *Young Man Luther*, Erikson reconstructed and analyzed the youthful monk's identity crisis and its creative resolution in early adulthood. *Gandhi's Truth* demonstrated the union of authentic identity and calling in a man at midlife. Through his psychological biographies of these historical figures, Erikson elucidated the interrelations between individual psychodynamics and development on the one hand and social structure and history on the other, deftly avoiding a reductionism in either a psychodynamic or sociological direction.

Sixty-five years after the publication of *Childhood and Society*, Erikson's ideas on identity development, intimacy, and generativity continue to inspire research and engender debate.

LIFE AND WORK

Erik Homburger Erikson was born on June 15, 1902, in Karlsruhe Germany, an old capital of a Lutheran principality. His father was Protestant and his mother Jewish. His parents, both Danish, had separated before his birth; his mother was visiting friends in Germany when the baby was born. She stayed in Karlsruhe and a few years later married her child's pediatrician, a well-to-do Jewish doctor, Theodor Homburger, in whose house young Erik grew up. During an alienated adolescence the tall, blond boy found himself regarded as a gentile in his father's Jewish milieu and as a Jew at school. He remembered his mother as sad, bookish, and artistic; his adoptive father as professionally respected; and both as loving. The boy's adoption by his stepfather was the formative occasion of what became a lifelong pattern of getting himself adopted by kind men. His last name

remained Homburger until he was 37, when he changed it to Erikson.

Erikson attended the *Humanistische Gymnasium* in Karlsruhe where he studied Greek, Latin, philosophy, literature, ancient history, art, and science. His primary interest was art; impatient with formal study and possessed by a restlessness he never lost, Erikson chose not to go to a university, preferring to travel about the countryside reading, drawing, and making wood carvings. Back home after a year, he tried formal art study with some success, first in Karlsruhe at the Badische Landeskunstschule and then in Munich at the Kunst-Akademie, but in neither case with decisive commitment.

This sort of wandering about was not uncommon among German youth of the period, so Erikson was permitted, as his biographer Robert Coles sagely wrote, "to go through his own years of discontent and confusion without being especially singled out and thereby forced to defend behavior often best granted the limits of its own momentum." As Erikson confided to Coles, "... if ever an identity crisis was central and long drawn out in somebody's life it was so in mine." Erikson was having what he would later term a *psychosocial moratorium*; in so doing, he was also mitigating the asperities of an *identity crisis*. At about age 21 Erikson went to live in Florence where he continued his art studies informally. There he enjoyed the friendship of his old *Gymnasium* chum Peter Blos, a writer who later became a famous American child psychoanalyst.

Becoming a Child Psychoanalyst

When Erikson was 25 years of age, Peter Blos invited him to become a faculty member in a progressive grammar school in Vienna, where Blos taught language and science. Erikson's difficult transition from adolescence to early adulthood was over. The year was 1927; Sigmund Freud was 71, and his youngest child, Anna, an educator and psychoanalyst, had started a psychoanalytically enlightened school for children with an American friend named Dorothy Burlingham. Erikson joined Blos and later recalled Blos's determination to turn him into a disciplined worker: "To make a teacher of me ... the highly disciplined Peter first had to teach me to keep regular work hours, a task which was initiated every morning, no matter what time of year, by a cold shower, then the preferred shock treatment for identity confusion." At the school, "Herr Erik" taught his students art and history; together, they studied different cultures and illustrated what they learned with drawings, essays, toys, tools, and exhibits.

Before long, Erikson found himself not only a teacher of children but also an *analysand*, what is now called a *candidate*, at the Vienna Psychoanalytic Institute, in treatment with Anna Freud. Any sort of psychoanalytic approach to treating or educating children was then a radical idea, even an approach as cautious as Freud's. Both educationally and clinically, Freud's group was sufficiently deviant such that a person

with Erikson's diverse identity was able to fit in. There was in 1927 a *configurational affinity* between Erikson's personal history and the history of psychoanalysis as a profession. Still there was the worry: What was a fledgling artist without a university education to do among those high-powered theorists and intellectuals at the Vienna Psychoanalytic Institute?

Erikson recalled this epiphanic exchange with his analyst, Anna Freud: "When I declared once more that I could not see a place for my artistic inclinations in such high intellectual endeavors, she said quietly: 'You might help to make them see.'" Dreams had been Sigmund Freud's royal road to the unconscious; observing children's play (and later, anthropological field studies in the role of participant observer) would be Erikson's path to understanding the ego and its development. Looking back, Erikson thought that it was Anna Freud's "simple mandate" that enabled him to succeed in combining the artistic and the theoretical in *Childhood and Society*.

Erikson remained in Vienna for 6 years, until 1933. In those years between the ages of 25 and 31, he turned his artist's eye from the observation of nature to the analysis of children; he learned about psychoanalysis by studying children's play and his own free associations as a patient, and he learned fundamental clinical skills in the supervised treatment of others. As he explained to Coles, Erikson related art to psychoanalysis in this way, "I began to perceive how important visual configurations were, how they actually preceded words and formulations: certainly dreams are visual data, and so is children's play, not to speak of the 'free associations' which often are a series of images, pure and simple – only later put into words." Erikson trained with a Montessori group in Vienna, graduating from the Montessori teacher's association, the Lehrerinnenverein. He studied psychoanalysis with August Aichhorn, Edward Bibring, Helene Deutsche, and Ernst Kris. He formed mentor relationships with Anna Freud and Heinz Hartmann, and a more distant, admiring-inspirational relationship with Freud's sick, aging, but still-productive father.

During his own late adulthood Sigmund Freud had moved beyond clinical problems more narrowly conceived to problems of the ego, society, and history in works such as *Beyond the Pleasure Principle* (written at age 64), *Group Psychology and the Analysis of the Ego* (at 65), and *The Ego and the Id* (at 66); during Erikson's years in Vienna, Freud was working on *The Future of an Illusion* (at 71) and *Civilization and Its Discontents* (at 73). Erikson seemed to identify with both Freuds, internalizing them as mentors, and they provided him with a strong, inner basis for a lifetime of psychoanalytic treatment and psychosocial research.

In 1929, Erikson met Joan Serson, a woman of mixed Canadian and American background. She had a Master's degree in sociology, a special interest in the history and social origins of modern dance and in

psychoanalysis. They married and Serson joined the faculty of Burlingham's school, where she taught English, literature, and American history. Serson's unusual combination of interests and skills—education, psychoanalysis, and sociology—coupled with her ability as a writer, gave Erikson a skilled coworker for a lifetime of intellectual work. Several articles were coauthored, Joan Erikson helped with most of the others, and the endowed chair at Harvard that was named after Erikson carried both their first names.

With the birth of their sons, Kai and Jon (daughter, Sue, was to be born 5 years later) Erikson had by 31 become a husband, a father, and a child psychoanalyst. He had come a long way from the aimless youth of his early 20s and was productively engaged in the stage of ego development that he later characterized by the polarity *intimacy versus isolation*. He continued to work on that ego stage, but changes were now required in the life that Erikson had begun to build at 25.

Emigrating to America

By 1933 the Fascist menace in Europe was growing. When Erikson graduated that year from the Vienna Psychoanalytic Institute, he and his family prepared to leave. They considered repatriating to Denmark but encountered problems because Erikson had lost his Danish citizenship when he had been adopted and become a naturalized German. A serendipitous meeting between Erikson and Freud's disciple Hanns Sachs in which Sachs enthusiastically invited Erikson to come to Boston settled the matter. In 1933 the Eriksons emigrated to the United States. At 31, Erikson became Boston's first child psychoanalyst and a member of the faculty of the Harvard Medical School. He took a job as a consultant at Judge Baker Guidance Center, where he helped to diagnose and treat poor and delinquent children with emotional disorders. In Cambridge, Erikson met Margaret Mead, Gregory Bateson, Ruth Benedict, and Kurt Lewin, each of whom influenced his intellectual development in important ways.

Erikson joined a group of personality researchers working under the leadership of Henry Murray, who was Director of the Harvard Psychological Clinic. Murray's book *Explorations in Personality: A Clinical and Experimental Study of Fifty Men of College Age* was published in 1938. It is one of the few integrative masterpieces in American psychology —experimental and biographical, psychological and sociological, Freudian and Jungian, developmental and diagnostic—and it provided a luminous example of studying persons both in their origins and in their purposes. Murray's example proved seminal when Erikson later turned his hand to the biographical study of the world historical figures Sigmund Freud, Martin Luther, and Mohandas Gandhi.

Listed on the title page of Murray's book, among other research associates, was "Erik Homburger." The next year, at age 37, Homburger

renamed himself Erikson, retaining Homburger as his middle name. With characteristically bold creativity, Erikson solved the problem of his paternity by adopting himself.

Psychoanalytic Anthropology

In 1936 Erikson hit the road again, this time to the Institute of Human Relations of the department of psychiatry at Yale University. The interdisciplinary work of the institute further shaped Erikson's interest in cross-cultural research, and in 1938 he joined a colleague on a research expedition to the Sioux Indians in South Dakota. On the Pine Ridge reservation, he observed children, interviewed adults, and noted the impact of economic, geographical, and historical factors on child-rearing practices. Based upon this research, Erikson published his paper, "Observations on Sioux Education" in 1939.

In the same year the peregrinating Eriksons moved to California, where he joined the faculty at the University of California at Berkeley. Erikson affiliated himself with Mount Zion Hospital and the San Francisco Psychoanalytic Institute, and ultimately became a U.S. citizen. At Bekeley, Erikson continued his cross-cultural research, this time with the Yurok Indians of northern California. He also lent his research skills to the war effort in articles on submarine habitation, the interrogation of prisoners of war, and difficulties encountered by veterans in returning to civilian life.

Erikson remained at U.C. Berkeley for a decade, the longest period in one place since his pre-adulthood years in Karlsruhe. During that time, he became his own man, visible first in the outward signs of changing his name and later in refusing to sign the university's loyalty oath of anti-Communist purity. To Erikson, the contract was "... an empty gesture toward meeting the danger of infiltration into academic life of indoctrinators, conspirators, and spies." In 1950, following the dismissal of some of his peers for refusing to declare themselves noncommunist, Erikson resigned his position as professor of psychology. He wrote a protest which was subsequently read at a meeting of the American Psychoanalytic Association. In that statement, he asserted:

> Young people are rightfully suspicious and embarrassingly discerning. I do not believe they can remain unimpressed by the fact that the men who are to teach them to think and to act judiciously and spontaneously must undergo a political test; must sign a statement which implicitly questions the validity of their own oath of office; must abrogate "commitments" so undefined that they must forever suspect themselves and one another; and must confess to an "objective truth" which they know only too well is elusive.

He continued:

> If the universities themselves become the puppets of public hysteria, if their own regents are expressly suspicious of their faculties, if the professors themselves tacitly admit that they need to deny perjury, year after year—will that allay public hysteria?

At Harvard, Erikson had studied "normal" students and their use of toys and dramatic scenes to enact and illustrate internal conflicts. At the Institute of Child Welfare at the University of California, Erikson continued his research and found important sex-related differences in the use of toys and play space by "normal" adolescents. As Coles noted, "Psychological themes—what the child says he is up to, what he can be seen doing and experiencing—are related to 'spatial configurations,' namely, objects and forms that exist in a world outside the mind. The mind in turn uses those objects or forms to express and reveal its wishes, fears, and conflicts." In research which had its roots in his work at Burlinghams's experimental school, Erikson had once again combined his artistic and psychoanalytic sensibilities in the study of human behavior.

But at age 40, as Erikson was entering middle-age, he was becoming more biographical and more interested in the adult years of the life cycle. Like others who manage to continue developing, he found a way to pursue his development through the strictures of exigent reality. World War ll was under way and Erikson was deeply concerned about it. After his initial narrowly framed attempts at scholarly patriotism, he began writing his first psychobiographical essays on Adolf Hitler and the psychosocial dynamics of his appeal to young Germans—"Hitler's Imagery and German Youth" was published in 1942. It united Erikson's interests in political science, history, and anthropology. As Coles stated, "Very little that Erikson would do for the next two decades was not in some respects foreshadowed by this paper."

Midlife Transition

Childhood and Society was the creative product of Erikson's transition to midlife. He had intended it to be a contribution to the psychiatric education of clinicians from various disciplines, but the book outgrew its author's intentions and found its way into every corner of the academy and beyond.

Erikson began work on the book when he was 42 and largely completed work on it by 46. Published in 1950, it is at once a product of early research, a prospectus of what was to come, and an initial integration of both. In it he presented clinical cases in which individual psychodynamics, society, and history are interwoven with a skill not seen before or since; analyses of children's play and development in various cultures; a theoretical sketch of the entire human life cycle; pieces on the problem of identity; and biographical essays on Adolf Hitler and the Russian writer Maxim Gorky.

As in Karl Abraham's and Freud's psychosexual stages, most of Erikson's stages in ego development occur in childhood and adolescence. Unlike the Abraham-Freud conception, however, Erikson's stages are psychosocial, describing crucial steps in the maturing ego's relations with the social world rather than a biological unfolding of neurophysiological capacities for excitation. Also, whereas Freud's developmental theory falls

back on itself after adolescence, Erikson's continued with characterizations of developmental tasks during youth, middle age, and old age.

Erikson's emphasis, like Freud's, was both cross-cultural and universal. Erikson's eight stages were ineluctable parts of the human life cycle, yet each person traversed them in distinct ways determined by culture, concrete circumstance, and personality. With *Childhood and Society*, Erikson became and remained an ego psychologist, shifting the traditional psychoanalytic focus on the drives to adaptation and growth. In so doing, he was furthering the work of his analyst-mentor, Anna Freud, who in 1936 had written ego psychology's basic theoretical treatise, *The Ego and The Mechanisms of Defense*.

Eight Stages of the Life Cycle

Erikson's conceptualization of psychosocial development across the life cycle is the centerpiece of his life's work and he elaborated the conception throughout his subsequent writings. It takes as its model the epigenetic principle of organismic growth in utero. In Erikson's view, psychosocial growth occurs in phases, with individual aspects of development proceeding according to a predetermined time table. Elements of each phase are present from birth and differentiate over time. Every phase has its own period of quiescence and critical ascendancy, and each is dependent on the proper development of the other phases in the proper sequence. Work on any particular phase, Erikson theorized, is never complete and old developmental conflicts can be activated by critical life events.

The eight stages of the life cycle represent points along a continuum of development in which physical, cognitive, instinctual, and sexual changes combine to trigger an internal crisis whose resolution results in either psychosocial regression or growth and the development of specific *virtues*. In *Insight and Responsibility* Erikson defined virtue as "inherent strength," as in the active quality of a medicine or a liquor. He wrote in *Identity: Youth and Crisis* that "crisis" refers not to a "threat of catastrophe, but to a turning point, a crucial period of increased vulnerability and heightened potential, and therefore, the ontogenetic source of generational strength and maladjustment." Elsewhere, Erikson averred, "...we do not consider all development a series of crises: we claim only that psychosocial development proceeds by critical steps—'critical' being a characteristic of turning points, of moments of decision between progress and regression, integration and retardation." A child's experience of each crisis or critical turning point is affected by the attitudes of its parents and the values and customs of its culture.

Trust versus Mistrust (Birth to about 18 Months) In *Identity: Youth and Crisis*, Erikson noted that the infant "lives through and loves with" its mouth. Indeed, the mouth forms the basis of its first *mode* or pattern of behavior, that of incorporation. The infant is taking the world in

through the mouth, eyes, ears, and sense of touch. The baby is learning a cultural modality that Erikson termed *to get*, that is, to receive what is offered and elicit what is desired. As the infant's teeth develop and it discovers the pleasure of biting, it enters the second oral stage, the active-incorporative mode. The infant is no longer passively receptive to stimuli; it reaches out for sensation and grasps at its surroundings. The social modality shifts to that of *taking and holding on* to things.

The infant's development of basic trust in the world stems from its earliest experiences with its mother or primary caretaker. In *Childhood and Society* Erikson asserts that trust depends not on "absolute quantities of food or demonstrations of love, but rather on the quality of maternal relationship." A baby whose mother is able to anticipate and respond to its needs in a consistent and timely manner despite its oral aggression will learn to tolerate the inevitable moments of frustration and deprivation. The defense mechanisms of introjection and projection will provide the infant with the means to internalize pleasure and externalize pain such that "consistency, continuity, and sameness of experience provide a rudimentary sense of ego identity." Trust will predominate over mistrust, and hope will crystallize. For Erikson, the element of society corresponding to this stage of ego identity is religion, as both are founded upon "trust born of care."

In keeping with his emphasis on the epigenetic character of psychosocial change, Erikson conceived of many forms of psychopathology as examples of what he termed *aggravated development crisis*, development that, having gone awry at one point, affects subsequent psychosocial change. A person who, as a result of severe disturbances in the earliest dyadic relationships, fails to develop a basic sense of trust or the virtue of hope may be predisposed as an adult to the profound withdrawal and regression characteristic of schizophrenia. Erikson hypothesized that the depressed patient's experience of being empty and worthless is an outgrowth of a developmental derailment that causes oral pessimism to predominate. Addictions may also be traced to the mode of oral incorporation.

Autonomy versus Shame and Doubt (about 18 Months to About 3 Years) In the development of speech and sphincter and muscular control, the toddler practices the social modalities of *holding on and letting go*, and experiences the first stirrings of the virtue that Erikson termed *will*. Much depends on the amount and the type of control exercised by adults over the child. Control that is exerted too rigidly or too early defeats the toddler's attempts to develop its own internal controls, and regression or false progression results. Parental control that fails to protect the toddler from the consequences of his or her own lack of self-control or judgment can be equally disastrous to the child's development of a healthy sense of autonomy. In *Identity: Youth and Crisis*, Erikson

asserted: “This stage, therefore, becomes decisive for the ratio between cooperation and willfulness, and between self-expression and compulsive self-restraint or meek compliance.”

Where that ratio is favorable, the child will develop an appropriate sense of autonomy and the capacity to “have and to hold”; where it is unfavorable, doubt and shame will undermine free will. According to Erikson, the principle of law and order has at its roots this early preoccupation with the protection and regulation of will. In *Childhood and Society*, he concluded, “[T]he sense of autonomy fostered in the child and modified as life progresses, serves (and is served by) the preservation in economic and political life of a sense of justice.”

An individual who becomes fixated at the transition between the development of hope and autonomous will, with its residue of mistrust and doubt, may develop paranoiac fears of persecution. When psychosocial development is derailed in the second stage, delinquency and others forms of pathology may emerge. The perfectionism, inflexibility, and stinginess of the person with an obsessive-compulsive personality disorder may stem from conflicting tendencies to hold on and to let go. The ruminative and ritualistic behavior of the person who suffers from an obsessive-compulsive disorder may be an outcome of the triumph of doubt over autonomy and the subsequent development and retention of a primitively harsh conscience.

Initiative versus Guilt (about 3 years to about 5 Years) The child’s increasing mastery of locomotor and language skills expands its participation in the outside world and stimulates omnipotent fantasies of wider exploration and conquest. Here the youngster’s mode of participation is active and intrusive; its social modality is that of *being on the make*. The intrusiveness is manifested in the child’s fervent curiosity and genital preoccupations, competitiveness, and physical aggression. The Oedipus complex is in ascendance as the child competes with the same-sex parent for the fantasized possession of the other parent. In *Identity: Youth and Crisis*, Erikson wrote that “Jealousy and rivalry … now come to climax in a final contest for a favored position with one of the parents: the inevitable and necessary failure leads to guilt and anxiety.”

Guilt over the drive for conquest and anxiety over the anticipated punishment are both assuaged in the child through repression of the forbidden wishes and the development of a superego to regulate its initiative. This conscience, the faculty of self-observation, self-regulation, and self-punishment, is an internalized version of parental and societal authority. Initially the conscience is harsh and uncompromising; however, it constitutes the foundation for the subsequent development of morality. Having renounced oedipal ambitions, the child begins to look outside the family for arenas in which it can compete with less conflict and guilt. This is the stage that highlights the child’s expanding initiative and forms the basis

for the subsequent development of realistic ambition and the virtue of *purpose*. As Erikson noted in *Childhood and Society*, "[T]he 'oedipal' stage...sets the direction toward the possible and the tangible which permits the dreams of early childhood to be attached to the goals of an active adult life." Toward this end, social institutions provide the youngster with an economic ethos in the form of adult heroes who begin to take the place of their storybook counterparts.

When there has been an inadequate resolution of the conflict between initiative and guilt the individual may ultimately develop a conversion disorder, inhibition, or phobia. Those who overcompensate for the conflict by driving themselves too hard may experience so much stress as to produce psychosomatic symptoms.

Industry versus Inferiority (about 5 Years to about 13 Years) With the onset of latency, the child discovers the pleasures of production. He or she develops industry by learning new skills and takes pride in the things made. Erikson wrote in *Childhood and Society* that the child's "ego boundaries include his tools and skills: the work principle ... teaches him the pleasure of work completion by steady attention and persevering diligence." Across cultures this is a time when the child receives systematic instruction and learns the fundamentals of technology as they pertain to the use of basic utensils and tools. As children work they identify with their teachers and imagine themselves in various occupational roles.

If the child is unprepared for this stage of psychosocial development, either through insufficient resolution of previous stages or by current interference, the child may develop a sense of inferiority and inadequacy. Teachers and other role models help to transmit social values and become crucially important in the child's ability to overcome a sense of inferiority and to achieve the virtue known as *competence*. In *Identity: Youth and Crisis*, Erikson noted: "[T]his is socially a most decisive stage. Since industry involves doing things beside and with others, a first sense of division of labor and of differential opportunity, that is, a sense of the technological ethos of a culture, develops at this time."

The pathological outcome of a poorly navigated stage of industry versus inferiority is less well defined than in previous stages, but it may concern the emergence of a conformist immersion into the world of production in which creativity is stifled, identity is subsumed under the worker's role, and feelings of inferiority are warded off through a defensive preoccupation with status and compensation.

Identity versus Role Confusion (about 13 Years to about 21 Years) With the onset of puberty and its myriad social and physiological changes, the adolescent becomes preoccupied with the question of identity. Erikson noted in *Childhood and Society* that youth are now "primarily concerned with what they appear to be in the eyes of others as compared to what they feel they are, and with the question of how to connect the roles

and skills cultivated earlier with the occupational prototypes of the day." Childhood roles and fantasies are no longer appropriate, yet the adolescent is far from equipped to become an adult. In *Childhood and Society* Erikson writes that the integration that occurs in the formation of ego identity encompasses far more than the summation of childhood identifications. "It is the accrued experience of the ego's ability to integrate these identifications with the vicissitudes of the libido, with the aptitudes developed out of endowment, and with the opportunities offered in social roles." Coles explains that ego identity is an "accrued confidence" which develops over time and crystallizes in early adulthood—or does not. It is the confidence that "... somehow in the midst of change one *is*; that is, one has an 'inner sameness and continuity' which others can recognize and which is so certain that it can unselfconsciously be taken for granted."

The formation of cliques and the intolerance of individual differences are ways in which the young person attempts to ward off a sense of identity confusion. In *Childhood and Society*, Erikson noted:

> For adolescents not only help one another temporarily through much discomfort by forming cliques and by stereotyping themselves, their ideals, and their enemies; they also perversely test each other's capacity to pledge fidelity. The readiness for such testing also explains the appeal which simple and cruel totalitarian doctrines have on the minds of the youth of such countries and classes as have lost or are losing their group identities (feudal, agrarian, tribal, national) and face world-wide industrialization, emancipation, and wider communication.

Falling in love, a process by which the adolescent may clarify a sense of identity by projecting a diffused self-image onto the partner and seeing it gradually assume a more distinctive shape, and an overidentification with idealized figures are means by which the adolescent seeks self-definition. With the attainment of a more sharply focused identity, the youth develops the virtue of *fidelity*—a faithfulness not only to the nascent self-definition but to an ideology that provides a version of self-in-world. As Erikson, Joan Erikson, and Helen Kivnick wrote in *Vital Involvement in Old Age*: "[F]idelity is the ability to sustain loyalties freely pledged in spite of the inevitable contradictions of value systems. It is the cornerstone of identity and receives inspiration from confirming ideologies and affirming companionships."

Role confusion ensues when the youth is unable to formulate a sense of identity and belonging. Coles explains:

> In a state of 'acute identity diffusion' the young individual may feel isolated, empty, anxious, and unable to make any number of choices or decisions that he himself (let alone his parents or teachers) feels pending. He feels threatened by what he senses to be close at hand: the possibility of intimacy, the chance at last to choose a career or find a job, the presence of others, who are seen as competitors, as somehow 'better' or less 'troubled.' He finds himself at a loss, and he fears that the world is breathing hard down his back – ready to restrict him, type him, define him, and thus close him off from any number of possibilities he still finds attractive. He wants 'out,' he wants to be away, he wants 'time' to think and decide and only later act.

Erikson held that delinquency, gender-related identity disorders, and

borderline psychotic episodes can result from such confusion.

Intimacy versus Isolation (about 21 Years to about 40 Years) Freud's famous response to the question of what a normal person should be able to do well, "Lieben und arbeiten" (to love and to work), is one that Erikson often cited in his discussion of this psychosocial stage, and it emphasizes the importance he placed on the virtue of *love* within a balanced identity. Erikson asserted in *Identity: Youth and Crisis* that Freud's use of the term love referred to "the generosity of intimacy as well as genital love; when he said love and work he meant a general work productivity which would not preoccupy the individual to the extent that he might lose his right or capacity to be a sexual and a loving being."

Intimacy in the young adult is closely tied to fidelity; it is the ability to make and honor commitments to concrete affiliations and partnerships even when that requires sacrifice and compromise. The person who cannot tolerate the fear of ego loss arising out of experiences of self-abandonment (e.g., sexual orgasm, moments of intensity in friendships, aggression, inspiration, and intuition) is apt to become deeply isolated and self-absorbed. *Distantiation,*an awkward term coined by Erikson to mean "the readiness to repudiate, isolate, and, if necessary, destroy those forces and people whose essence seems dangerous to one's own," is the pathological outcome of conflicts surrounding intimacy and, in the absence of an ethical sense where intimate, competitive, and combative relationships are differentiated, forms the basis for various forms of prejudice, persecution, and psychopathology.

Erikson's separation of the psychosocial task of achieving identity from that of achieving intimacy, and his assertion that substantial progress on the former task must precede development on the latter have engendered much criticism and debate. Critics have argued that Erikson's emphasis on separation and occupationally based identity formation fails to take into account the importance for women of continued attachment and the formation of an identity based on relationships.

Generativity versus Stagnation (about 40 Years to about 60 Years) Erikson asserted in *Identity: Youth and Crisis* that "Generativity ... is primarily the concern for establishing and guiding the next generation." The term *generativity* applies not so much to rearing and teaching one's offspring as to a protective concern for all the generations and for social institutions. It encompasses productivity and creativity as well. Having previously achieved the capacity to form intimate relationships, the person now broadens the investment of ego and libidinal energy to include groups, organizations, and society. *Care* is the virtue that coalesces at this stage. In *Childhood and Society* Erikson emphasized the importance to the mature person of feeling needed. "[M]aturity needs guidance as well as encouragement from what has been produced and must be taken care of." Through generative behavior the individual is able to pass on knowledge

and skills while obtaining a measure of satisfaction in having achieved a role with senior authority and responsibility in the tribe.

When people are unable to develop true generativity, they may settle for pseudo-engagement in occupation. Often such people restrict their focus to the technical aspects of their roles at which they may now have become highly skilled, eschewing larger responsibility for the organization or profession. This failure of generativity can lead to profound personal stagnation, masked by a variety of escapisms, such as alcohol and drug abuse, and sexual and other infidelities. Midlife crisis or premature invalidism (physical and psychological) may occur. In this case pathology appears not only in middle-aged persons but also in the organizations that depend upon them for leadership. Thus, the failure to develop at midlife can lead to sick, withered, or destructive organizations that spread the effects of failed generativity throughout society; examples of such failures have become so common as to constitute a defining feature of modernity.

Integrity versus Despair (about 60 Years to Death) In *Identity: Youth and Crisis*, Erikson defined integrity as "the acceptance of one's one and only life cycle and of the people who have become significant to it as something that had to be and that, by necessity, permitted of no substitutions." From the vantage point of this stage of psychosocial development, the individual relinquishes the wish that important people in his life had been different and is able to love in a more meaningful way—one that reflects an acceptance of responsibility for one's own life. The individual in possession of the virtue of *wisdom* and a sense of integrity has room to tolerate the proximity of death and to achieve what Erikson termed in *Identity: Youth and Crisis* a "detached yet active concern with life." Despair may be present but it does not predominate.

Erikson underlined the social context for this final stage of growth. In *Childhood and Society*, he wrote, "The style of integrity developed by his culture or civilization thus becomes the 'patrimony' of his soul ... In such final consolidation, death loses its sting."

When the attempt to attain integrity has failed, the individual may become deeply disgusted with the external world, and contemptuous of persons as well as institutions. Erikson wrote in *Childhood and Society* that such disgust masks a fear of death and a sense of despair that "time is now short, too short for the attempt to start another life and to try out alternate roads to integrity." In thinking about the relationship between adult integrity and infantile trust, Erikson observed, "[H]ealthy children will not fear life if their elders have integrity enough not to fear death."

Becoming a Biographer of Youth

In 1949, as *Childhood and Society* was going to press, Erikson's stand on the University of California loyalty oath had rendered his position at the university untenable. News of his availability spread east, and other

institutions vied to capitalize on Berkeley's mistake. Erikson was well known at the Menninger Clinic in Topeka, Kansas, where he had lectured, so when Robert Knight left Menninger to become the director of the Austen Riggs Center in Stockbridge, Massachusetts, he took Erikson with him. The Austen Riggs Center was devoted to psychoanalytic research and to the treatment of severely disturbed adolescents and young adults—Erikson had been happily adopted once again.

Erikson stayed at Austen Riggs from 1950 to 1960, from age 48 to 58, when he returned to the faculty of Harvard University. During those years, Erikson completed the transition to biographer, forming more fully the approach foreshadowed in his incomplete essays on Hitler, Gorky, and George Bernard Shaw. The transition was not easy; Coles describes that time as constituting a second "identity crisis." Erikson had been a clinician and a theorist of youth; he was now old enough to enact the program promised in *Childhood and Society* and to become a biographer and theorist of the whole life cycle. However, before he wrote *Young Man Luther*, he once again returned to Freud and the origins of his own professional identity. The way to the future was through the past.

In his early 50s Erikson wrote three biographical essays on Freud: "The Dream Specimen of Psychoanalysis," "Freud's 'The Origins of Psychoanalysis,'" and "The First Psychoanalyst." These essays all concern themselves with the crisis Freud suffered during his own midlife transition as he struggled to leave neuropathology and to define a new professional identity as a psychoanalyst. Strengthened by that reexamination of Freud's successful transition (and perhaps having reassured himself of Freud's blessing), Erikson began writing one of the genuine masterworks of psychoanalysis, *Young Man Luther*.

As always, Erikson's clinical work and observations enriched his theorizing about the life cycle. He acknowledged his debt to Austen Riggs and the Western Psychiatric Institute at the University of Pittsburgh in his preface to *Young Man Luther*. His work there had allowed him to "study the afflictions of young patients as variations on one theme, namely, a life crisis, aggravated in patients, yet in some form normal for all youth. I could identify those acute life tasks that would bring young people to a state of tension in which some would become patients."

However, Erikson would neither make Martin Luther a patient nor reify the psychopathology of the young people who were his patients. Instead, he asserted that comparisons between Martin Luther and his patients were "not restricted to psychiatric diagnosis ... but ... oriented toward those moments when young patients, like young beings anywhere, prove resourceful and insightful beyond all professional and personal expectation. We will concentrate on the powers of recovery inherent in the young ego."

Working with the young patients at Austen Riggs helped Erikson hone

his understanding of identity formation. As he later put it in "The Problem of Ego Identity," "[I]dentity ... is gradually established by successive ego syntheses and re-syntheses throughout childhood. It is a configuration gradually integrating *constitutional givens, idiosyncratic libidinal needs, favored capacities, significant identifications, effective defenses, successful sublimations, and consistent roles*."

These "ego syntheses and re-syntheses" were about to receive great attention as Erikson attempted his biography of Luther. But to do that work, Erikson the biographer had to move further away from his own professional origins as a child psychoanalyst. He had criticized psychoanalysis for its contention that something has been explained by finding an analogy to its earliest manifestations. Yet his own conception of the life cycle, with most of its stages occurring within the pre-adult era, left him still heavily rooted in childhood. In his study of Luther, Erikson was trying to understand a life, not just a personality; from that perspective, childhood and adolescence had to be introductory rather than the story itself.

Seeing that a child-centered view was not adequate for the tasks of biography, without either formally changing the imbalance in his theory or neglecting childhood determinants, Erikson deftly devoted the greatest attention to an explication of development in the adult years. So successful was he in interrelating Luther's adult problems with those of his early development that *Young Man Luther* provided the seminal inspiration for the next generation of life-cycle investigators. Some of those investigators, such as Daniel Levinson, formally redressed in their own theories the originological bias vestigial in Erikson's theories.

In Luther's early or mid-20s, according to some of his contemporaries, the young monk had fallen to the floor of the choir of his monastery in a fit and shouted, "ich bin's nit!" or "Non Sum!" "It isn't me," or "I am not." Erikson's task as a psychological biographer was to explain how Luther got from the identity crisis of his 20s to nailing his 95 theses to the door of the church in Wittenberg at the age of 32. Yet *Young Man Luther* is also, as it is subtitled, "a study in psychoanalysis and history." Since it is not a psychoanalysis of history, some nonreductionist connecting concepts between the individual and the collectivity were needed. One of Erikson's connecting concepts was *ideology*, which he defined as "an unconscious tendency underlying religious and scientific as well as political thought: the tendency at a given time to make facts amenable to ideas, and ideas to facts, in order to create a world image convincing enough to support the collective and the individual sense of identity."

Luther had personal and psychiatric problems, but in his defiance of Catholic orthodoxy he created a new religious ideology that not only supported his own identity but the identities of emerging generations of Europeans as well. "Ich bin's nit" expressed the young monk's identity

crisis as he found himself "fatally over-committed" to what he was not, a young man authentically embarking on a future as an orthodox Catholic priest. "In some periods of history," Erikson asserted, "and in some phases of his life cycle, man needs...a new ideological orientation as surely and as sorely as he must have air and food." As such a man, Luther commanded Erikson's "sympathy and empathy" as he "faced the problems of human existence in the most forward terms of his era."

The reception of the book cemented Erikson's international reputation, initially won by *Childhood and Society*, as a leader in original, humane thought about psychological development. In no other biography have the dynamics of individual conflict and development been so seamlessly interwoven with those of society and history. With *Young Man Luther*, Erikson in his late 50s ascended from the first rank of developmental psychoanalysts to that of a cultural seer.

Becoming a Biographer of Middle Age

As Erikson moved through the transition from middle to old age he abandoned clinical work with patients to turn his attention fully to the study of the life cycle and the survival of the species. His emphasis, emergent in *Young Man Luther* and implicit in all his earlier work, was on the ego virtues that permit a person to live in a constructively critical relationship with the social institutions of his time. Earlier, Erikson had described the stages in the development of the ego across the life cycle. In 1964 he wrote more about the ego virtues in *Insight and Responsibility*. He was getting ready amidst the American moral and political crises of the mid-1960s to embark on his last major work, a study of the great moral leader Mohandas Gandhi. In the end, he dedicated it, *in memoriam*, to Martin Luther King, Jr.

Young Man Luther had been a story of personal choice and historical change wrought by a young man establishing and revising an initial identity and life structure as an adult. In his mid-60s Erikson had gained sufficient perspective on the life cycle to analyze a case of decisive crisis and change in a man who was going through a midlife transition and solving it by creating a vital life structure for middle age. Gandhi did that, as he himself put it, by realizing "his vocation in life," leading *satyagraha* truth force campaigns in his nation, in his family, and in his own soul. Gandhi was entering the stage of the life cycle in which Erikson's ego polarity of *generativity versus stagnation* becomes active; from his developmental work on this polarity, Gandhi was learning the ego virtue of *care*.

Erikson saw that Gandhi, like Luther, was both a leader and a "religious actualist." As a leader, he succeeded in articulating inner concerns in a way that struck a collective cord. Self-rule and home rule were inextricably combined in Gandhi's program. He would have endorsed in his own terms, Erikson believed, Luther's assertions, "Christ comes today; God's way is

what makes us move; we must always be reborn, renewed, regenerated; to do enough means nothing else than always to begin again."

Gandhi had begun and successfully completed his political novitiate as an Oxford-trained barrister in South Africa during his 20s and 30s. By his late 30s, he had developed passive resistance as a political strategy to defeat the Black Act, a law that required all Indians in the Transvaal to register with the government and to carry identification papers on their persons. At 40, Gandhi staked out his claim for leadership in his Indian Home Rule Manifesto, and during the next few years of his midlife transition he developed the device of the Satyagraha campaign, culminating in the Great March of striking mine workers in South Africa when he was 44.

When Gandhi returned to India in 1915 at age 45, he did so, Erikson wrote, "like a man who knew the nature and the extent of India's calamity and that of his own fundamental mission." As a mature, middle-aged man, he was in Erikson's view, a person who has determined what "he does and does not *care for*" as well as "what he *will* and *can* take *care of*. He takes as his baseline what he irreducibly is and reaches out for what only he can, and therefore, *must* do."

What Gandhi had to do was lead a labor strike at the mills in Ahmedabad and the next year, at age 50, a national strike for independence from Britain. Erikson contended that Gandhi's emergence at that time as the "father of his country" underlined "the fact that the middle span of life is under the dominance of the universal human need and strength which I have come to subsume under the term *generativity*." Erikson refers here not merely to the generativity that a parent at 20 or 30 may feel for a child. Erikson's concept of generativity is often mistakenly taken to apply to the child-rearing work of young adults. Instead, he is describing the protective concern for the generations and their retarding-facilitating social institutions, whose leadership is the essential responsibility of the middle-aged individual.

In the early 1970s Erikson returned to the San Francisco Bay Area to live and begin again. The Eriksons relocated to Cambridge, Massachusetts, some 10 years later. In 1981 the Eriksons and psychologist Helen Kivnick undertook an intensive study of 29 octogenarian parents of children whose lives had been meticulously scrutinized in the Guidance Study begun in 1928 at the Institute of Human Development at the University of California at Berkeley. The results of their research were published in *Vital Involvement in Old Age*.

In this work, Erikson, Erikson, and Kivnick revisit the eight stages of psychosocial development and their attendant virtues or strengths. They emphasize the opportunities that exist at the end of life to integrate "maturing forms of hope, will, purpose, competence, fidelity, love, and care into a comprehensive sense of wisdom." A successful integration results

from a final reworking of the earlier polarities of basic trust versus mistrust; autonomy versus shame and doubt; initiative versus guilt; industry versus inferiority; identity versus confusion; intimacy versus isolation; generativity versus self-absorption; and integrity versus despair. The authors use the term *vital involvement* to identify an involvement with the environment characterized by "actuality" and "mutuality." The elder who is able to maintain a vital involvement with his world—no matter how small its scope—is better able to tolerate the depredations of aging: the loss of sensory acuity, physical mobility, and stamina; changes in intimate relationships as children marry and move away, and friends and partners sicken and die; the loss of social status, financial security, and sources of pride in professional competence; and the realization that relatively little time remains.

Developed in infancy, hope is the basis upon which the senescent individual attempts to integrate faith in the universe and the relative predictability of its laws, with a realistic mistrust about that which is unpredictable and unreliable. The elder may experience a need to affirm a basic faith—in religion or nature—and to understand more deeply his own place in a generational progression.

Late in life the individual must struggle to redefine a sphere of autonomy even as he becomes increasingly dependent upon others. Living independently, residing in familiar surroundings, and following established routines can be crucial symbols of autonomy and the means by which signs of impairment (e.g., a walker) may be accepted without undue shame. An individual's lifelong experience of autonomy influences his or her capacity to adjust to restrictions upon freedom. The ability to allow others to help supports the generative impulses of younger adults and reinforces for both the sense of a generational cycle.

The polarities of initiative versus guilt, first evident in the youngster's exploration and play and subsequently in adult work and recreation, must be balanced anew in old age. With the passage of time the person faces a diminution of opportunities for the exercise of initiative, and must regret those instances where he or she failed to act decisively or with sufficient concern for others. Successful adaptation to the limitations of old age must draw upon the strength of purposefulness developed in childhood and tempered by maturity.

With the natural decline in physical strength and sensory acuity, the senescent individual increasingly depends upon a lifelong sense of effectiveness. External rewards are usually not as numerous as they were in young adulthood or middle age. The feeling of competence and mastery must be sustained internally. An old person may cope with growing physical limitations by changing the criteria for a sense of accomplishment, pursuing new hobbies or modifying old ones, developing new avenues of involvement with former professions, or living vicariously through the

accomplishments of children and grandchildren.

Old age provides the individual with the chance to review a life-time of beliefs, to come to terms with choices made and opportunities lost, and to make and act upon any final commitments, which now most clearly reflect what Erikson termed "the 'I' in the totality of life." In striking a final balance between identity and identity confusion, the elder concerns himself with both an external image and a personal image. The external image is the way the individual may expect to be remembered; the personal image reconciles the sense of who he or she has been with an evolving sense of who he or she may yet become. Identification with the successes of younger generations and satisfaction in one's own part in the transmission of values helps the senescent to stave off feelings of insignificance or obsolescence in the face of technological and sociological changes.

The sense of love that can emerge at the end of life is one that is built upon a foundation of love, expressed and unexpressed, throughout the life cycle. A new balance may be forged between intimacy and isolation in the face of decline. Intimate relationships may be recast in more positive terms; reminiscences provide the elder with the opportunity to draw upon and reintegrate earlier experiences of tenderness and sexuality.

In order to reconcile feelings of integrity and despair, the individual must make peace with previous choices, acknowledging what was gained and lost and accepting that it is too late to change. Relationships and values may be reviewed with more room to consider alternate points of view. In balancing the dispositions toward generativity and stagnation, the elderly person may be able to build upon years of experiences as a parent, worker, and creatively productive person from middle age, which themselves were shaped by childhood experiences of caring or its absence. Grandparenthood offers another chance at generativity and a way to undertake it more robustly and less ambivalently. Erikson coined the term "grand-generativity" to distinguish the generativity of old age in which the person, freed from the direct responsibility for family, institution, and community, is better able to incorporate care for the present with concern for the future. Such concern may be expressed through advice or financial assistance to family members or on a larger scale, through commitment to religious beliefs, political involvement, or community action. Grand-generativity contributes to a feeling of immortality; in caring for the younger generation and allowing oneself to be taken care of, the senescent secures a personal place in a generational history.

The "joint reflections of old age" contained in the book are an effort to elucidate the psychosocial process of vital involvement and to extend Erikson's observations to the outer limits of the life cycle. The book also brought Erikson back full circle to his professional beginnings—having made almost 40 years earlier, play observations on the Guidance Study children.

Application of Erikson's Concepts to Clinical Work

Erikson's view of individual experience as inexorably embedded in developmental, familial, societal, and historical contexts crucially shaped his ideas concerning mental illness and psychiatric treatment. As noted previously, Erikson was reluctant to pathologize behavior or to rush to judgment about the meaning of any given symptom. He asserted, "Perhaps there are certain stages in the life cycle when even seemingly malignant disturbances are more profitably treated as *aggravated life crises* rather than as diseases subject to routine psychiatric diagnosis."

To understand the meaning of such "disturbances," Erikson drew upon his training in classical drive theory (with its concepts of the unconscious; id, ego, and superego conflicts; repression, regression, and repetition compulsion) but emphasized the adaptive and synthesizing capacities of the ego in brokering relations between internal drives and external reality. When Erikson listened to his patients, he thought about their symptoms not simply as compromise formations between libidinal drive and superego prohibition but also as expressions of arrested or derailed psychosocial development. Toward this end, Erikson examined many aspects of his patient's current life (e.g., skills, talents, and aspirations; religious and political commitments; social and intimate relationships; roles as worker, partner/spouse, and parent) and located them within a larger social, cultural, and historical context. His focus was on the individual's entire life cycle, the development of specific virtues and those factors which facilitated or retarded psychosocial growth. As Shapiro and Fromm point out:

> He paid attention to the whole life context of any immediate situation. He asked a number of questions. What is the immediate stimulus for the patient's reaction? What is the acute life conflict, the current developmental stage, the issues that are manifest? In what developmental context did the patient's reaction first occur? Is it now manifest in the relationship to the therapist, in a repetitive conflict, in a characteristic way that the individual solved earlier developmental struggles? In what social context is the individual embedded, what roles are available?... What defenses does the individual use? What are the individual's deepest psychological investments?"

Erikson conceived of these psychological investments in broad terms. He used concepts of attachment, separation, and mutuality (defined as "a relationship in which partners depend on each other for the development of their respective strengths") to illustrate ways in which the individual is essentially a social being. He recognized that psychological crises often occurred at times of developmental separation or individuation and that fears regarding dependency and abandonment were frequently the catalyst for such crises. Though Erikson acknowledged the influence of early childhood conflict or trauma on an individual's progress in negotiating developmental tasks in adolescence and adulthood, he did not think that the issues being played out were simply reiterations of infantile conflict.

Erikson's belief that psychosocial development continues throughout

the life cycle led him to caution psychotherapists against a theoretical reductionism in which symptomatic behavior was attributed to a traumatic past and in which insufficient attention was paid to contemporary conflicts in the ego's relations with the world. He was particularly critical of the "originology" in psychoanalytic thought. "I mean by it a habit of thinking which reduces every human situation to an analogy with an earlier one, and most of all to that earliest, simplest and most infantile precursor which is assumed to be its 'origin.'" Erikson was wary of other theoretical blindspots within his profession. As Coles noted, "... psychiatric and psychoanalytic concepts emerge from and become very much suited to a particular society —which exerts its (moral and puritanical) influence on everything, even the most complicated, rarefied, and 'objective' of abstractions."

In his article, *The Nature of Clinical Evidence*, Erikson outlined the relationship between patient and therapist. He described the "clinical core of medical work" as the "encounter of two people, one in need of help, the other in the possession of professional methods. Their *contract* is a therapeutic one: in exchange for a fee, and for information revealed in confidence, the therapist promises to act for the benefit of the individual patient, within the ethos of the profession." Though he acknowledged certain parallels between medical and psychiatric consultations (e.g., the patient presents with a list of symptoms, the physician/psychiatrist asks questions designed to elucidate the problems, arrives at a tentative diagnosis, and proposes a treatment plan) Erikson departed from a traditional medical model in important ways. Though according the therapist the respect and authority appropriate to his special training and expertise, Erikson envisioned the patient as an active participant in a collaborative endeavor. Within the limitations imposed by treatment setting (e.g., inpatient vs. outpatient) and degree of psychological impairment, the patient was expected to take significant responsibility for his treatment. Erikson noted a natural transformation which occurs in the patient. "'Under observation,' he becomes self-observant. As a patient he is inclined, and as a client often encouraged, to historicize his own position by thinking back to the onset of the disturbance, and to ponder what world order (magic, scientific, ethical) was violated and must be restored before his self-regulation can be reassumed."

How does the patient participate in his own "cure?" He is invited to free associate. Erikson noted, "[W]e consider a patient's 'associations' our best leads to the meaning of an as yet obscure item brought up in a clinical encounter, whether it is a strong affect, a stubborn memory, an intensive or recurring dream, or a transitory symptom." The associations consist of whatever thoughts, feelings, sensations, or images occur to the patient during and after the mention of that item. Erikson continued, "Except in cases of stark disorganization of thought, we can assume that what we call the synthesizing function of the ego will tend to associate what 'belongs

together,' be the associated items ever so remote in history, separate in space, and contradictory in logical terms ... It is ... this basic synthesizing trend in clinical material itself which permits the clinician to observe with 'free-floating attention,' to refrain from undue interference, and to expect sooner or later a confluence of the patient's search for curative clarification and his own endeavor to recognize and to name what is most relevant, that is, to give an *interpretation*." Free-floating attention "turns inward to the observer's ruminations even as it attends the patient's 'free associations' and which, far from focusing on any item too intentionally, rather waits to be impressed by recurring themes." A message represented by a patient's memory, dream image, or symptom can be overdetermined—"a condensed code transmitting a number of other messages, from other life situations, seemingly removed from the therapy. This we call 'transference.'"

As for the therapist's role in working with the patient's free associations, Erikson averred, "I have to assume that the patient is (to varying degrees) unconscious of the meaning which I discern in his communications, and that I am helping him by making fully conscious what may be totally repressed, barely conscious, or simply cut off from communication ... I take for granted an effective wish on his part (with my help) to see, feel and speak more clearly." Erikson felt that a "therapeutic" interpretation was one which was clear and concise, yet presented the patient with a "unitary" theme, "... a theme common at the same time to a dominant trend in the patient's relation to the therapist, to a significant portion of his symptomatology, to an important conflict of his childhood, and to corresponding facets of his work and love life." How would the therapist determine if an interpretation was "right?" Erikson believed, "[T]he proof lies in the way in which the communication between therapist and patient 'keeps moving,' leading to new and surprising insights and to the patient's greater assumption of responsibility for himself."

Erikson acknowledged that there was a great deal of subjectivity involved in both the patient's complaints and the therapist's interpretations. He cautioned, "... even while facing most intimate and emotional matters, [the therapist] must maintain intellectual inner contract with his conceptual models, however crude they may be." He must evince "... a *specific self-awareness* in the very act of perceiving his patient's actions and reactions." Erikson wrote, "... there is a core of *disciplined subjectivity* in clinical work – and this both on the side of the therapist and of the patient—which it is neither desirable nor possible to replace altogether with seemingly more objective methods." The therapist was advised to undergo a personal psychoanalysis in order to reduce or eliminate potential blindspots which might compromise this disciplined subjectivity.

Erikson noted "the power of the transference, i.e., the patient's transfer to me of significant problems in his past dealings with the central people in

his life" but also conceived of what we would now call a "real" relationship between himself and his patient. In *The Nature of Clinical Evidence*, Erikson provided a lengthy clinical vignette illustrating dream analysis and his use of his emotional reactions in his interpretations to his patient. Because he did not conceive of clinical work as primarily oriented toward the development and interpretation of the transference, Erikson accorded himself and other clinicians the freedom to cultivate "real" aspects of the relationship with the patient. He felt it important and necessary at times, that the therapist educate and guide his patient. Erikson averred, "... only by playing my role as a new person in his present stage of life can I clarify the inappropriateness of his transferences from the past."

A psychotherapy informed by Eriksonian principles is essentially optimistic in nature. For it takes as a given the notion that psychosocial development is universal and continues throughout the life cycle. Natural forces propel the individual forward. Psychiatric symptoms are often best understood as manifestations of aggravated developmental crises—development which has been blocked or derailed. In such cases, the therapeutic task is to release or redirect those built-in forces for growth and adaptation.

As Erikson noted in *Identity: Youth and Crisis*, the object of psychotherapy is not to head off future conflict but to assist the patient in emerging from each crisis "with an increased sense of inner unity, with an increase of good judgment, and an increase in the capacity 'to do well' according to his own standards and to the standards of those who are significant to him."

References

Arnett JJ. Identity development from adolescence to emerging adulthood: What we know and (especially) don't know. In: McLean KC, Syed M, eds. *The Oxford Handbook of Identity Development*. New York: Oxford University Press; 2015:53–64.

*Coles R, Erik H. *Erikson. The Growth of His Work*. Boston: Little, Brown and Company; 1970.

*Erikson EH, Erikson J, Kivnick H. *Vital Involvement in Old Age*. New York: Norton; 1986.

Erikson EH. *Insight and Responsibility*. New York: Norton; 1964.

Erikson EH. Observations on Sioux education. *J Psychol*. 1939;7:101–156.

Erikson EH. Hitler's imagery and German youth. *Psychiatry*. 1942;5:475–493.

*Erikson EH. *Childhood and Society*. New York: Norton; 1950.

Erikson EH. The dream specimen of psychoanalysis. *J Am Psychoanal Assoc*. 1954;2:5.

Erikson EH. Freud's "The Origins of Psychoanalysis." *Int J Psychoanal*. 1955;36:1.

Erikson EH. The first psychoanalyst. *Yale Rev*. 1956;46:40.

Erikson EH. The problem of ego identity. *Psychol Issues*. 1959;1:379.

*Erikson EH. *Young Man Luther*. New York: Norton; 1962.

Erikson EH. *Identity: Youth and Crisis*. New York: Norton; 1968.

*Erikson EH. *Gandhi's Truth*. New York: Norton; 1969.

Evans R. *Dialogue with Erik Erikson*. New York: Harper and Row; 1967.

Freud A. *The Ego and Mechanisms of Defense*. New York: International Universities Press; 1966.

Friedman LJ. *Identity's Architect: A Biography of Erik Erikson*. New York: Scribner; 1999.

Friedman LJ. Erik Erikson on identity, generativity, and pseudospeciation: A biographer's perspective. *Psychoanal Hist*. 2001;3:179.

Levinson D, Darrow C, Klein E, Levinson M, McKee B. *The Seasons of a Man's Life*. New York: Knopf; 1978.

McAdams DP, Zapata-Gietl C. Three strands of identity development across the human life course: Reading Erik Erikson in full. In: McLean KC, Syed M, eds. *The Oxford Handbook of Identity Development*. New York: Oxford University Press; 2015:81–96.

Roth MS. Psychoanalysis and history. *Psychoanalytic Psychology*. 2016;33S:S19.

Schein S, ed. *Erik Erikson: A Way of Looking at Things*. New York: Norton; 1987.

Shapiro ER, Fromm MG. Eriksonian clinical theory and psychiatric treatment. In: Sadock BJ, Sadock VA, eds. *Comprehensive Textbook of Psychiatry*. Vol 7. New York: Lippincott Williams & Wilkins; 2000:2200–2207.

Slater C. Generativity versus stagnation: An elaboration of Erikson's adult stage of human development. *J Adult Dev*. 2003;10:53.

Stevens R, Erikson EH. *Explorer of Identity and the Life Cycle*. New York: Palgrave Macmillan; 2008.

Syed M, McLean KC. The future of identity development research: Reflections, tensions, and challenges. In: McLean KC, Syed M, eds. *The Oxford Handbook of Identity Development*. New York: Oxford University Press; 2015:562–574.

Tummala-Narra P. Cultural competence as a core emphasis of psychoanalytic psychotherapy. *Psychoanalytic Psychology*. 2015;32:275.

Wallerstein R, Goldberger L, eds. *Ideas and Identities: The Life and Work of Erik Erikson*. Madison, CT: International Universities Press; 1998.

▲ 6.3 Other Psychodynamic Schools

ADAM M. BRENNER, M.D., AND LAURA S. HOWE-MARTIN, PH.D.

The professionals discussed in this chapter—Adolf Meyer, Alfred Adler, Carl Gustav Jung, Otto Rank, Karen Horney, Franz Alexander, Harry Stack Sullivan, Wilhelm Reich, Erich Fromm, and Eric Berne—expanded psychiatric thought and practice throughout the 20th century and into the 21st. With knowledge of the neurosciences initially primitive, the ethos of psychiatry was to use clinical observation to search for an overarching theory that would encompass not only psychopathology but all of human behavior. None of these individuals doubted that the mind has a biological basis, but they all were aware of Sigmund Freud's woeful failure in his *Project for a Scientific Psychology* to develop an understanding of how experience, thoughts, and feelings are transduced to and from material states. Because of this, they were forced to stand back from the physical

organism and view humans strictly as psychological entities. What, then, is the relevance of these theories to 21st-century psychiatry? Are these theories of merely historical interest, or do they hold some wisdom for current clinical practice? Until 30 years ago, psychiatrists spent great time and energy comparing and contrasting theorists and arguing the pros and cons of various formulations. Within psychoanalysis and other "schools" of psychotherapy, such theoretical discussions still occur. Now, however, the primary contribution of these theorists is in their broad brush strokes; for example, should one embrace the frustration model of classical psychoanalysis, the actively empathic model of self-psychology, or the interactive model of interpersonal psychiatry?

There is also a historical shift that has occurred. In the past, as a clinician contemplated some new insight into the workings of patients' inner lives, it was often seen as necessary, not only to build a new theory, but also to create a new "school" of therapy. This may have been an outgrowth of Freud's personality, which tended to focus on his current line of observations and thought, thus forcing others (e.g., Adler and Jung) to move outside the inner circle to pursue their different lines of thinking. These days, psychoanalysis is a far more open intellectual enterprise, embracing a variety of different models and inputs.

These theories remain relevant in several important ways. First, regardless of what is learned about neurobiology, psychiatry is and will remain concerned with distress in behavior, thoughts, emotions, and relationships. Inevitably, these distresses will be communicated through words and nonverbal signals. Psychiatrists will always need some organizing templates at the behavioral and verbal communicational levels to help order and interpret their observations. A sad, pensive, tearful expression does not automatically translate into a serotonin or norepinephrine deficiency, or some other structural abnormality. It must first be translated into the psychological constructs of loss, depression, grief, and so forth. Second, these theorists spent enormous time listening to patients, and all were careful observers. As such, their observations remain helpful in defining key dimensions for psychiatrists to consider in understanding the inner experiences patients attempt to communicate. Finally, all of these theorists were well versed in a rigorous approach to theory construction. Every human being has a "theory of personality," acknowledged or otherwise. This theory, often intuitive and unrigorous, based on a mixture of reading, personal values, and unsystematic observation, influences much of the way people interact with one another. In the clinical situation, when a psychiatrist mixes supportive management and medication, he or she is using a personal theory of personality. The advantage of knowing of the theorists in this chapter is that models of intellectual rigor can be brought to personal intuitions.

To illustrate the different understandings and therapeutic approaches of

each theorist, the following case is used as a basis for formulation according to each theorist.

Mr. A was a 26-year-old white man who had a history of bipolar I disorder. He was brought in for treatment after not completing the last required course for his advanced degree and being arrested for disturbing the peace. He had consistently lied to his family about where he stood with his coursework and about having skipped an examination that would have qualified him to use his professional degree. He had also not told them that he had been using marijuana almost daily for a number of years and occasionally used hallucinogens. His arrest for disorderly conduct was for swimming naked in an apartment complex in the middle of the night while under the influence of hallucinogens.

Mr. A's use of marijuana began early in college but became daily during graduate school. He was diagnosed as having bipolar I disorder early in his senior year at college after a clear episode of mania. His mood disorder was well controlled on lithium (Eskalith). During graduate school, he was episodically compliant with medications, preferring to try to maintain a state of hypomania. He saw a psychiatrist every 3 to 6 months for medication checks. During his 4 years in graduate school, he had two clear episodes of depression and began taking sertraline (Zoloft), 100 mg per day, with questionable benefit. Mr. A believed that he could be a great writer. He spent most of his time reading and trying to write. He dreamed of going to New York or Paris and becoming part of a new wave of novelists. This aspiration and his marijuana abuse predated his development of bipolar I disorder. He attended class episodically, nonetheless performing adequately. His last class had no final examination but required a paper. He planned to write this paper in the form of a play, involving a dialogue between two thinkers from different times and cultures. His professor was very excited about this idea, but Mr. A kept postponing the task until he was forced to extend his schooling by a year. His other major interest during this time involved growing and photographing flowers.

Mr. A was born and raised in a large city. His father had been very successful in commercial real estate, and his mother, after raising the children, used the substantial real estate holdings she inherited from her father to set up a business to manage them. Most of the money was placed in a trust for the patient and his siblings. His mother had total financial control of the trusts and doled out the proceeds to the children as they needed them. There was no family history of any psychiatric disorders.

The patient described his mother as very loving and caring but to the point of being intrusive and controlling. For example, the mother arranged the initial treatment but then was angry that the psychiatrist had not called her regularly to report on her adult son's progress. She was also critical of various aspects of the treatment as reported to her by her son. The patient's two older siblings had attended prestigious colleges and graduate schools but had returned home to work in the mother's real estate management company. The 30-year-old sister was living in the parents' home. The 35-year-old brother had lived at home for a time but then moved out to a location a few blocks away. There was a younger brother, still in college, who also smoked marijuana excessively. He tried to minimize the patient's problems to the family and tried to protect the patient, who desperately had not wanted to return home. Of note is that none of the children were married, although the two older ones had each had a couple of serious relationships.

The children seemed to regard the mother with affectionate amusement and bemusement. The father was seen as a very caring but undemonstrative man who put much energy into keeping the mother from becoming too upset and encouraged the children to do the same. The children often wanted to provoke the mother for her judgmental, detail-oriented intrusiveness. The father discouraged them but occasionally found their provocations amusing.

The family viewed itself as very close, with strong values oriented toward community service and family loyalty. The family belonged to a religious community but expressed their involvement primarily in social service and social action volunteer work, accompanied by very generous financial contributions.

The patient had been a very successful debater in high school and recalled his development as very positive but provided few details. He tended to place himself in the role of the outsider, an observer of humanity, which he saw as consonant with the role of a writer. He was proud to have bipolar I disorder and tried to regulate his medications so that he would be hypomanic

much of the time, seeing this as enhancing his creativity. He viewed his use of marijuana in the same vein. One of the most distressing aspects to him of his depressive episodes was that marijuana no longer created a feeling of well-being but made him feel worse. His current depressive episode involved no neurovegetative symptoms. Rather, he presented as flat, numb, apathetic, ashamed, and anhedonic. He was particularly ashamed of being back in his hometown and of living with his parents.

The patient ostensibly understood and accepted his illness well and had read much about it. However, the family had responded to the information "with proper treatment, bipolars can live normal lives" as meaning that the information should be kept secret so that he should be treated normally. Mr. A, on the other hand, was very open with friends at graduate school about his illness and his pride in it and the creativity he associated with it.

The patient had two longstanding recurrent dreams. One involved him flying. The narrative line varied, but the flying theme recurred. Often, he had other magical powers in his dreams such as the ability to heal, to not be killed by bullets, to save the world or some group of people from mortal danger, and so on. The other recurrent dream was of a hotel lobby. These dreams regularly began with him entering a hotel lobby to meet a group of people, accompanied by a feeling of dread (Table 6.3–1).

ADOLF MEYER

Adolf Meyer (1866 to 1950) immigrated to the United States after training as a neuropathologist in Switzerland (Fig. 6.3–1). Uninterested in metapsychology, he espoused a commonsense psychobiological approach to understanding mental disorders, emphasizing the interrelationship of symptoms and individual psychological and biological functioning. His approach to the study of personality was biographical; he attempted to bring psychiatric patients and their treatment out of isolated state hospitals and into communities, and was also a strong advocate of social action for mental health. He began his career as a clinically oriented state hospital pathologist, becoming the second head of the Pathological Institute (later, the New York Psychiatric Institute), and initiated its affiliation with the Columbia College of Physicians and Surgeons.

Table 6.3–1.
Techniques and Terminology

Meyer	Biographical Study Approach
	Distributive Analysis and Distributive Synthesis
	Focus on patient assets
	Habit training
Adler	Individual Psychology
	Inferiority toward Mastery; "Inferiority Complex"
	Birth Order
	Dynamism (future-directed, goal-oriented)
	Removing blocks to living productively in the real world
	Reframing
	Paradoxical communication
Jung	Complexes
	Archetypes
	Projective/Introjective processes

	Persona, Shadow, Anima, and Animus
	Typology (Introversion-Extroversion, Sensation-Intuition, Thinking-Feeling)
Alexander	Psychosomatic medicine
	Corrective emotional experience
	Controversial attempts to shorten/manipulate psychodynamic therapy
Reich	Ego psychology
	Character armor and personality types
	Bioenergetics
Rank	Birth trauma
	Death fear; Life fear
	Concept of Will
	Therapy as rebirth
Horney	Gender, culture, and personality traits
	Theory of Neurosis, character neuroses, and neurotic trends
	Alienation
Fromm	Importance of historical and social context for human behavior
	Marxist view of character types
Sullivan	Therapeutic communities
	Definition of personality and emphasis on interpersonal relationships
	Cognitive modes: Prototaxic, Parataxic, Syntactic
	Social development model
	Interpersonal Psychotherapy
Berne	"Strokes" as primary motivators

He later became the president of the American Psychiatric Association and had a 32-year tenure as chair of psychiatry at Johns Hopkins. His major social contribution was helping to found the National Committee on Mental Hygiene.

Psychobiology

Despite his background as a neuropathologist, which was common among European psychiatrists of his generation, Meyer strongly opposed the Kraepelinian view of mental illness as having a predetermined course based on phenomenologically identified syndromes. Instead, he believed that individuals' habitual reaction patterns made them more susceptible to specific types of breakdown. Meyer used ***biographical study*** and strongly encouraged psychiatrists to take a thorough, detailed life history of the patient to understand each individual's reaction patterns. He saw development as lifelong, thus viewing the data from adolescence onward as equal in importance compared with data from childhood in understanding the patient's narrative.

FIGURE 6.3–1. Adolf Meyer. (Courtesy of National Library of Medicine, Bethesda, M.D.)

He observed reaction patterns, attempted to predict the conditions under which they might occur, and tested and validated methods for their modification. He acknowledged the contributions of Freud and Jung but believed they were too narrow. A thorough pragmatist, he preferred common sense to metapsychological constructs as the means to understand and deal with psychopathology.

Theories of Personality and Psychopathology

Meyer believed human beings possess a fundamental tendency toward integration, and multiple biological, social, and psychological forces

contribute to personality development. The vulnerable person uses poorly planned, ill-suited means of adaptation. Meyer used the biographical approach as a practical guide to elicit information about personality development, to organize the information, and to reevaluate the information obtained under different circumstances. His clinical examination assessed each patient's life history; physical, neurological, genetic, and social status; and the relationship between those factors and personality factors. A diagnosis and an individual treatment plan were based on this assessment.

Treatment

The aim of psychobiological therapy was to help individuals make the best possible adaptation to changing environmental circumstances. It began with the development of a collaborative relationship. Out of this collaborative relationship came ***distributive analysis***, an examination of the factors in patients' lives that contributed to their adjustment or lack thereof, and concluded with ***distributive synthesis***, helping patients to understand themselves and to develop better coping skills. The first step in distributive analysis is the patient's own exposition of the presenting problem. Assets and liabilities are then determined by eliciting the life history in terms of the memories that are immediately available and those that are later fleshed out by reconstructing past experiences.

Treatment is initiated by ***focusing on patients' assets***. It involves psychological, chemical, physical, and environmental measures as needed. In more severe cases, attention is paid first to patients' sleep habits, nutrition, and daily routines because these must be normalized before psychological work can be done. Patients are helped to describe their difficulties in detail. In addition to eliciting complaints or worries, patients are asked what eases or worsens their complaints and what significance they attach to their symptoms and concerns. In doing this, the therapist attempts to use the patient's own language and concepts to communicate suggestions and advice.

Meyer did not pay attention to unconscious mechanisms but focused on patients' functioning in reality. Both present-day and long-term adaptive patterns are considered. Therapeutic sessions proceeded from immediate, obvious problems in the present to longer-term issues and historical data. With guidance, patients investigate their own personality problems, ascertain the origin of their conflicts, and work to develop more useful behavior patterns. Meyer called this ***habit training***, a term he may or may not have borrowed from a behaviorist tradition. When unhealthy adaptive patterns are modified, proper adjustment and personal satisfaction result.

Meyerian Treatment of Mr. A

Meyerians would focus first on the adequate treatment of Mr. A's mental illness. Controlling biological forces that might disrupt his life is always a primary goal. Because he is still dependent on his parents, they are included in the treatment plan and might be seen separately by a social worker. Both Mr. A and his parents are told that failure to control the symptoms of his mental illness could cost him his life and any satisfaction that he might derive from it. Mr. A's mood swings are stabilized on an appropriate medication, and he is observed closely for restoration of normal sleeping and eating habits. With our modern knowledge of neurobiology, the vulnerability of Mr. A's dopamine reward system to addiction would be another biological force to be dealt with. He is asked to discontinue his habit of marijuana use for this reason, but also because of its tendency to aggravate the symptoms of his mental illness. Random urine screening might be recommended. Mr. A is helped to develop a daily schedule, including time for work, social interaction, and recreation. He is not encouraged to return to school because of his tendency to use school as an escape from responsibility. He is encouraged to develop his photography as a hobby after a period of stable work performance. Later, he would begin the distributive analysis phase of his treatment and would be asked to reflect on the impact of his bipolar I disorder and his avoidance of responsibility in his life. In the distributive synthesis phase of treatment, he is helped to realize that his avoidance of responsibility keeps him from achieving independence and the pleasure of attaining any realistic goal. He is asked to develop a plan for achieving both appropriate independence and for remaining within the family structure. Mr. A's parents are seen periodically for a progress report and to urge them to push Mr. A toward appropriate self-care, toward making use of his own resources. They might be asked to place Mr. A on an allowance and to hold him to it strictly to reinforce his accountability for his actions. Over time, his parents are asked to slowly withdraw financial support and to give Mr. A more independence. They might be asked to place his inheritance in a trust designed so that he could not impulsively misspend it.

FIGURE 6.3–2. Alfred Adler (print includes signature). (Courtesy of Alexandra Adler.)

ALFRED ADLER

Alfred Adler (1870–1937) was born in Vienna, Austria, and spent most of his life there (Fig. 6.3–2). A general physician, he became one of the original four members of Freud's circle in 1902. However, Adler never accepted the primacy of the libido theory, the sexual origin of neurosis, or the importance of infantile wishes. In 1911 he resigned as president of the Vienna Psychoanalytic Society and continued the development of his own socially and interpersonally focused theory of development. He posited a striving for self-esteem through overcoming a sense of inferiority, which he saw as an inevitable force in the human condition as a result of its extended

childhood. He equated psychological health with constructive social engagement, developing a system that he called ***individual psychology***, which is still vigorous in many countries. His major social contribution was the establishment of child guidance centers in Vienna that served as a model for the rest of the world.

Personality Theory

Adler saw individuals as unique, unified biological entities whose psychological processes fit together into an individual lifestyle. He also postulated a principle of ***dynamism***, which in every individual is future directed and moves toward a goal. Once the goal is established, the psychic apparatus shapes itself toward attainment of that goal. Life goals are chosen and are thus subject to change. Changes require modification of memories, dreams, and perceptions to fit the accomplishment of the new goal. Adler also emphasized the interface between individuals and their social environment, and the primacy of action in the real world over fantasy. Community mindedness, acceptance of the need to conform to the legitimate demands of society, is an important precept, but Adler also recognized a dialectic that occurs between individuals and their interpersonal environment, each constantly reacting to and shaping the other. Thus, Adler anticipated some of the modifications of Freud's theories introduced by Heinz Hartmann as well as some of Sullivan's thinking.

Normal Personality and Adaptation. The cornerstone of Adler's personality theory is the concept of ***moving from a sense of inferiority to a sense of mastery***. Early in life, everyone has a sense of inferiority resulting from realistic comparison with adults' size and abilities. Moving from this sense of inferiority to a sense of adequacy is the most important motivation in life. Thus, the ideal person strives for superiority and does so through high social engagement; the emotionally handicapped person continues to feel inferior and reinforces that position through lack of striving and social interest.

There are many obstacles to the development of self-esteem and social interest. Prominent among them are poorly developed or "inferior" organs or systems (such as poor eyesight or poor eye–hand coordination), childhood diseases, pampering, and neglect. Physical handicaps and childhood diseases may promote self-centeredness and loss of social interest; ***birth order*** is another factor. According to Adler, first-born children, having lost their position of only child, tend not to share, and become conservative. Later children are more open to change and become social activists. Youngest children feel secure because they have *never* been displaced. Adler may have been the first to anticipate recent psychological research on the importance of birth order in human behavior. We acknowledge our indebtedness to him whenever we use the term

"inferiority complex." Adler was also the first theoretician to place mastery and self-esteem at the center of personality development.

Theory of Psychopathology

Adler believed emotional disorders result from "mistaken lifestyles" that are subject to change by will and by self-understanding. Individuals subject to emotional disorders have false ideas about themselves and the world, and inappropriate goals that lead them away from constructive social interest. Individuals with a pampered lifestyle, for example, expect and demand support from others, avoid responsibility, blame others for their failures, but they also feel incompetent and insecure because their well-being depends on pressing others into service. If life poses no challenge, a mistaken lifestyle may have no consequences. When a mistaken lifestyle is ineffective, symptoms develop that protect self-esteem while helping the individual to avoid dealing realistically with the problem being confronted. The difference between neurosis and psychosis is that neurotic individuals maintain social interest but are blocked from life goals by symptoms, whereas psychotic individuals lose social interest and retreat into their own world.

Psychotherapy

Because his theory emphasized the mismatch of lifestyles with the demands of the real world, Adler focused on ***blocks to living productively in the real world***, not on exploring unconscious conflict. His aim was to point out mistaken self-views and mistaken views of the world and then, by mobilizing will, make the needed changes, including a change in life goals.

Therapeutic Process. Starting with three sessions per week and tapering off to one per week, a positive relationship with patients is established and used to lead patients to awareness of their lifestyle, how it is discordant with the demands of social reality, and ways in which it may be reoriented. Instead of striving for goals of no social value that falsely increase self-esteem, patients are pushed to work toward ameliorating their situation. Having become aware of the obstacles they have placed in their own path and the consequences of these self-defeating behaviors, they are now helped to develop constructive interests in themselves and others. As they become less self-engrossed, they find themselves better accepted by others, which reinforces their constructive efforts. People who have dedicated themselves to symbolically defeating others learn to cooperate and advance toward useful goals. Any endeavor in which patients can develop real competence is encouraged, whether social, work, artistic, or musical.

Patients are encouraged to address the concrete obstacles to developing

a useful lifestyle, including reading instruction for slow readers or contact lenses for people who are self-conscious about their appearance. Early recollections, birth order, dreams, daydreams, and present-day interactions are all used to help patients see the inappropriateness or falseness of their ideas and life goals. Actual life events or memories of events are less important than individuals' reactions to those events or memories. Because memories are likely to be retrospective falsifications justifying an erroneous lifestyle, there is little need to verify them. There is also no need to look for latent content in dreams; they are merely an expression of present-day concerns. Nor is it necessary that therapists' interpretations be correct because they need only to help patients build a useful conception of themselves and the world. This perspective anticipated what has come to be called the *hermeneutic approach* to psychotherapy.

Several of Adler's techniques, including reframing and paradoxical communication, now enjoy wide popularity. ***Reframing*** is viewing the same data from a different point of view. Indecision, for example, was reframed from a product of mixed feelings to a wish to maintain the status quo. Failure to act keeps everything the same, which is the self-fulfilling prophecy of the discouraged person. After this reframing statement, patients are pushed to act constructively. ***Paradoxical communication*** is instructing patients to do the opposite of what the therapist wishes them to do. In dealing with an indecisive person, for example, Adler might caution against doing anything rash. Adler also paid attention to the impact of his patients on their environment and recognized that individuals do much to create their own interpersonal worlds. In response to complaints about being treated unfairly by others, Adler asked patients how they dealt with the people about whom they complained. Above all, Adler treated his patients as rational and as able to learn more productive ways of living. In his emphasis on practical, constructive solutions; misconstrued goals; and misperceived views of the world and the self, Adler also anticipated important elements of cognitive therapy.

An Adlerian View of Mr. A

As seen from the Adlerian point of view, Mr. A has developed a mistaken lifestyle and an inappropriate life goal. He maintains himself in fantasy as a writer while failing at the accomplishments that would enable him to become a writer. Thus, he has been attempting to make the normal step of moving from inferiority to mastery in fantasy instead of through realistic achievement. Blocks to his development include pampering and his concomitant denial and abuse of his bipolar I disorder, the latter representing his organ inferiority. He has lost the social interest he had earlier in life and become extremely self-centered, unconcerned with others or with the consequences of his actions. He uses drugs and hypomania to avoid the pain of defeat.

An Adlerian therapist encourages Mr. A to develop a realistic self-view. He has a mental illness that requires lifelong treatment. Grandiose dreams and intoxication with drugs cannot substitute for accomplishments in the real world and always lead to defeat. He is asked to set and to strive to accomplish small, realistic goals for himself such as holding a steady job while enjoying photography as a hobby. His mental illness is reframed as a challenge to his creativity and his use of marijuana as an obstacle to mobilizing his creativity. He might initially be encouraged to accept his dependency on his family and later, as he stabilizes, to reduce his dependency on his family as a means to heighten his self-esteem and to make a transition into adulthood. He is encouraged to join the Depression and Bipolar Support Alliance or other available support and educational groups as a means to better understand and accept his illness, to develop a social conscience, and to stimulate his altruism. His dream about flying is interpreted as his wish to achieve mastery; his dream about the hotel lobby is recast as awareness that he has been trying to substitute fantasy for reality.

CARL GUSTAV JUNG

Carl Gustav Jung (1875–1961) was a lifelong resident of Switzerland (Fig. 6.3–3). He trained in psychiatry under Eugen Bleuler at the Burgholzli Mental Hospital in Zurich and was strongly involved with Freud and the psychoanalytic movement from 1906 to 1914, when he resigned as president of the International Psychoanalytic Association. He was an important figure in connecting psychoanalysis with clinical psychiatry. After a period where he experienced psychotic symptoms but made use of them as a spur to creative thinking, which lasted from 1914 to 1918, Jung became an advocate of active introspection as the means to intrapsychic change. This episode and Jung's subsequent interpretation of it, as well as that of his disciples, form the focus of much of the controversy about him and his ideas. Although Jung rejected Freud's notion of libido as sexual energy and the Oedipus complex as a universal developmental stage, he believed not only in the unconscious mind but in a shared racial and species unconscious. This interest of Jung's is the source of another controversy involving Jung's relationship to the Nazi establishment. Jung, an intuitive introvert, was not interested in the practical aspects of living in the world; his focus was instead on individuation through becoming aware of the unconscious. His theories have particular attractiveness to clinicians and nonclinicians interested in spirituality, creativity, and mystical experience.

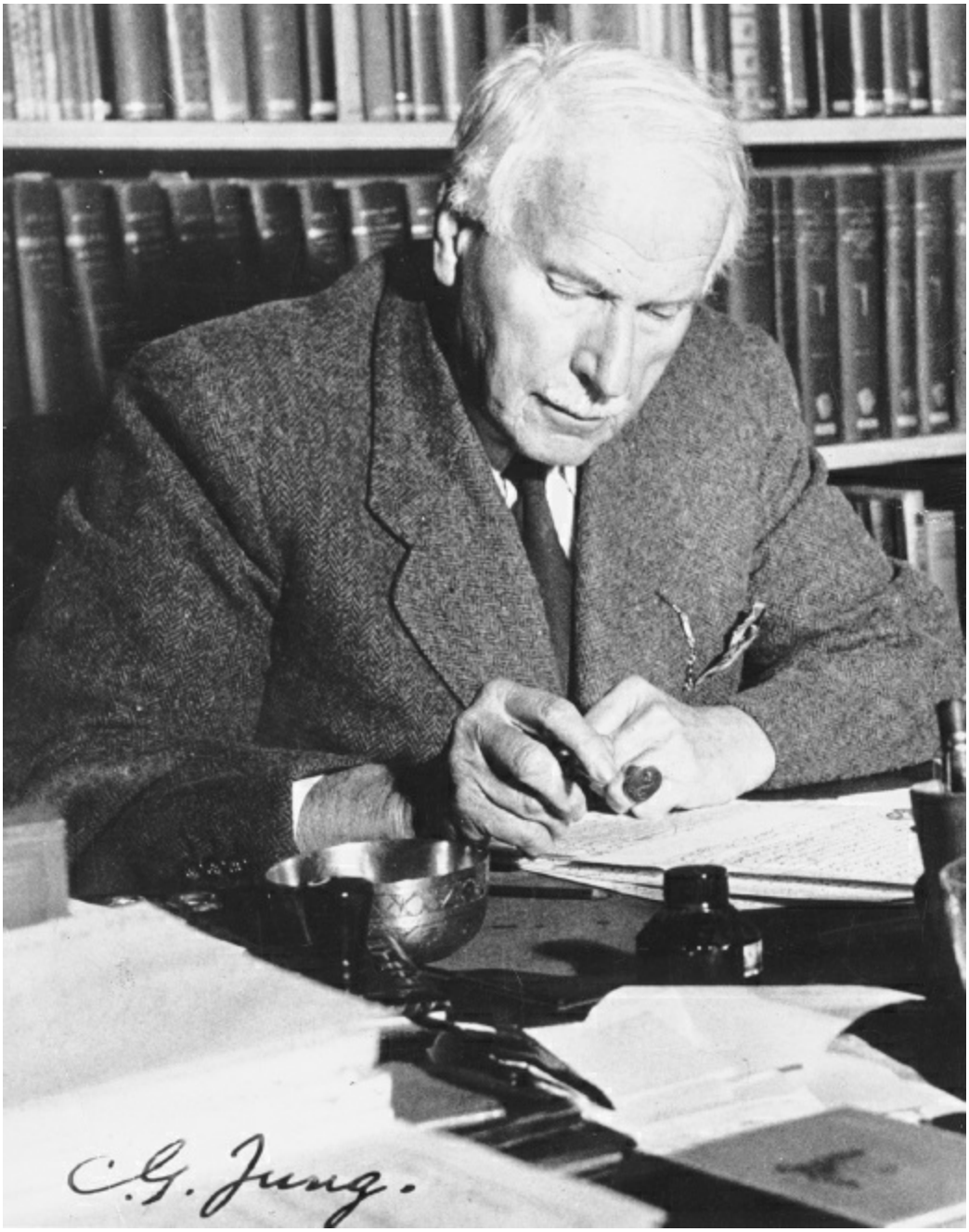

FIGURE 6.3–3. Carl Gustav Jung (print includes signature). (Courtesy of National Library of Medicine, Bethesda, M.D.)

Personality Theory

Jung developed an elaborate metapsychology that was every bit as detailed and formalized as Freud's. His theories identify most of the important fault lines in classical psychoanalysis. Jung's construct of the psychic apparatus differs from the Freudian structure of ego, superego, id, and ego ideal (Fig. 6.3–4). Below an outer rim of consciousness is the personal unconscious, which contains the complexes. Contained within the personal unconscious and connected to the complexes are the archetypes, the elements of the self, which in turn connect to the surface of the personality as the ego.

Complexes. ***Complexes*** are groups of unconscious ideas associated with particular emotionally toned events or experiences. Jung inferred them from his early word-association studies, when he noted that certain words provoked intense reactions or produced less reaction than expected. Complexes are built around genetically determined intrinsic models of the world known as ***archetypes***. Complexes are also reinforced by environmental events and by selective attention or inattention and are thus self-perpetuating. They are endowed with psychic energy from their affective tone: Positive, negative, mild, or strong, the more intense the complex is, the greater the emotion, imagery, and tendency to action. Complexes are often stimulated by interactions with others. A father complex can be stimulated by a person who symbolizes a father (such as an older friend) or by a stimulus, such as music or art, that evokes father memories. The complex, formerly dormant in the unconscious, comes to the fore and tends to dominate consciousness and to displace other complexes. This is a very different model of the dynamic unconscious from Freud's, in that the boundary between the conscious and unconscious is far more permeable, and the emphasis is on what in external reality stimulates feelings and thoughts to enter awareness rather than on what forces a complex out of awareness. Similarly, Jung's conceptualization of interpersonal boundaries was far more permeable than Freud's; Jung also placed much more emphasis on nonverbal stimulation (music, art, spiritual life) in this process than did Freud.

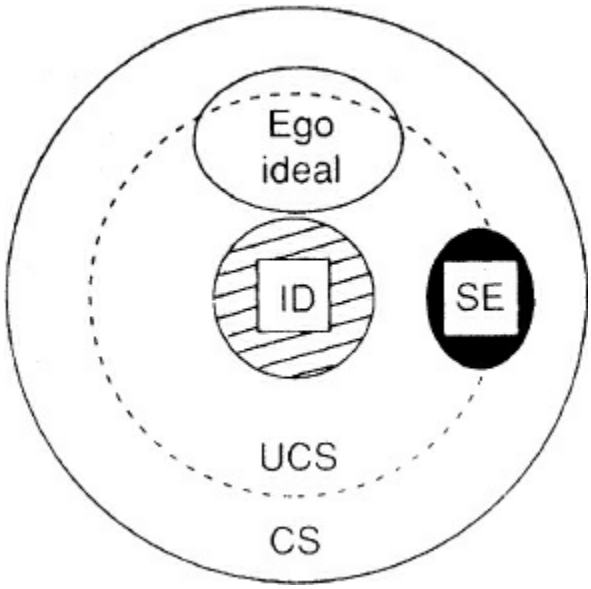

FIGURE 6.3–4. The Freudian topology of the psychic apparatus. CS, conscious; SE, superego; UCS, unconscious.

Some complexes are more conscious, better developed, and more acceptable to the individual; others are less conscious, poorly developed, and uncomfortable. The latter are projected onto the environment,

especially by children, and from this, ***projective*** and ***introjective*** processes evolve. One person may introject and identify with a complex being projected by another person. Thus, therapists may become psychologically infected by their patients. It is also possible to project a complex that is not integrated within oneself onto another person and then develop a relationship with that projected complex. One can envision an interpersonal environment charged with mutual projections available for introjection, thus offering endless potential for mutating perceptions and misperceptions. Thus, for Jung, the psychological boundary between individuals was far less clear than for Freud or Adler. In addition, one can see that Jung was much less concerned with an ultimate, objective, rational reality to be determined and compared with the individual's complexes of projections and introjections. Jung also had a parallel appreciation for the process of projective identification that Melanie Klein was developing at approximately the same time. Working as he did for many years in a long-term sanatorium with very ill patients, Jung clearly became comfortable with blurred interpersonal boundaries. His theory has been criticized for accepting as normative what might be regarded as pathological models of relationships.

Another important aspect of complexes is their ***bipolarity***, again paralleling Klein's work on splitting. Each complex has a positive and negative pole, such as good father and bad father or rewarding father and punishing father. One pole of a complex can be projected onto another person, who in turn introjects it and acts on it in a relationship. In this way, the theory of complexes is a theory of interpersonal as well as intrapsychic relationships.

In Jungian theory, the ego is also a complex. It serves the same function as the Freudian ego, controlling conscious life and bridging the intrapsychic and external worlds. The other complexes that also make up the psyche may align with or lie in opposition to the ego. Emotionally charged primitive complexes (e.g.,, the earth mother in a patient ambivalent about motherhood) have a tendency to become autonomous and may, when activated by external stimuli, behave oppose or control the ego, leading to irrational, impulsive behavior or misperceptions. They appear as images in dreams, as hallucinations, and as separate personalities in dissociative identity disorder. They also appear in séances when mediums bring forth so-called personalities from the dead. For Jung, this phenomenon also explained animism and states of possession.

Jung had a great interest in mystical experience, which Freud saw as the persistence of infantile, magical thinking, primary process, or wish fulfillment. This is another of the clear divisions in their thinking. For Freud and Freudians, the world is a harsh, demanding place that forces one to give up primitive wishes for magical experiences, symbiotic relationships, and intense gratifying emotional moments. For Jung and

Jungians, magic is alive and well in the inner world and is to be embraced.

Archetypes. Complexes are connected to structures embedded more deeply in the psychic apparatus, the archetypes. ***Archetypes*** are the inherited capacity to initiate and carry out behaviors typical of all human beings, regardless of race or culture, such as nurturing and aggression. Just as vision cannot develop without visual input during physiological critical stages, so archetypes require interactional stimulation for their elaboration into complexes. Thus, the human infant's psyche is not amorphous energy awaiting organization by the environment; it is instead a complex and organized set of potentials whose fulfillment and expression depend on environmental stimuli. There are as many archetypes as there are prototypical human situations. In positing the archetypes based on anthropological data, Jung was remarkably prescient of current understanding of the brain's organization.

The mother complex archetype illustrates the interrelationship of complex and archetype. All humans are born with a poorly formed but relatively clear model of an all-nurturing caretaker, which Jung referred to as the *earth mother archetype*. The mother complex emerges from this based on experiences with mothers or mother surrogates: Their attitudes, personalities, and relationship to the particular individual. The mother archetype is found in dreams or fantasies, often as a huge woman or an animal with many breasts. The motif of a many-breasted animal, found in many cultures, is that of unlimited nurturance.

Unconscious. The Jungian unconscious has two layers: The more superficial being the personal unconscious and the deeper layer the collective unconscious. The complexes exist in the personal unconscious, the archetypes in the collective unconscious or *objective psyche*. The *personal unconscious* is the equivalent of the Freudian unconscious, a repository of individual memories that have been repressed. The *collective unconscious* is the residue of what has been learned in humankind's evolution and ancestral past, much as human deoxyribonucleic acid (DNA) is an aggregate of the past. In this portion of the psyche reside instincts, potential for creativity, and the spiritual heritage of humankind. The potent synthesis of Darwinian natural selection theories with Mendelian genetics, which has had such a profound impact on all of biology, occurred in the 1920s. It is not clear whether Jung was aware of this intellectual thrust, but his understanding of the collective unconscious and of the archetypes is remarkably consonant in form, if not in detail, with the modern understanding of hereditary and developmental neurobiology as applied by cognitive science.

The psyche, like all other living systems, attempts to stay in balance. Jung's term for homeostasis in the relationship of conscious to unconscious

life is the *law of compensation*. For any conscious attitude or experience that is overly intense, there is an unconscious compensation. A person experiencing neglect might fantasize or dream about a huge, many-breasted mother. When interpreting dreams, Jung asked himself for what conscious attitude the dream compensated.

Personality Structure. At the center of the conscious personality is the complex called the *ego*. Several universal complexes attend the ego. The ***persona***, or public personality, mediates between the ego and the real world. The ***shadow***, a reverse image of the persona, contains traits that are unacceptable to the persona, whether positive or negative. A brave persona, for example, has its fearful shadow. The archetype of the shadow is the enemy or feared intruder. The ***anima*** is a residue of all the experiences of woman in a man's psychic heritage; the ***animus***, the residue of all the experiences of man in a woman's psychic heritage. The anima or animus connects the ego to the inner world of the psyche and is projected onto others in day-to-day or intimate relationships. When connected with the shadow, a man, for example, might see attributes of woman as undesirable and might experience guilt encountering such qualities in himself.

SELF. The *self* is the archetype of the ego; it is the innate potential for wholeness, an unconscious ordering principle directing overall psychic life that gives rise to the ego, which compromises with and is partly shaped by external reality. In Jungian metapsychology, the unconscious gives rise to integration, order, and individuation. The self appears from the unconscious in dreams, fantasies, and altered states of consciousness to give direction. In the first half of life, the ego attempts to identify with the self and to appropriate the power of the self in the service of the ego's growth and differentiation. During this time, the ego may become inflated with an unrealistic sense of power. If it is cut off from the self, there may be a sense of alienation and depression.

INDIVIDUATION. In the second half of life, the ego begins to attend more to the self than to the conscious realm of life. Jung called this developmental process *individuation*, the drive for individuals to become unique and to fulfill the spiritual propensities common to all humanity. Often, this process requires withdrawing from earlier identities and conventional definitions of success and seeking new paths. This change often has the paradoxical effect of leading to broader and more mature relationships in addition to greater creativity.

PSYCHOLOGICAL TYPES. Jung's theory of psychological types has three axes (Fig. 6.3–5). The *extroversion–introversion* polarity refers to the two basic types of object relatedness. Extroverts are oriented to others and to the

world of consciousness. Their energy flows outward first, then inward. Introverts are oriented to their inner world, their energy flowing first inward and then to outer reality. Introverts might therefore be seen as selfish and unadaptable because they attend first to their inner world and then determine how the outer world can fit them. Extreme extroverts, on the other hand, can seem insensitive to themselves and to the inner lives of others.

The *sensation–intuition* polarity concerns perception. The sensation type comes to understand a situation by assembling the details; the intuitive type grasps the overall situation before attempting to assimilate its parts. The sensation type sees the trees first; the intuitive type sees the forest first.

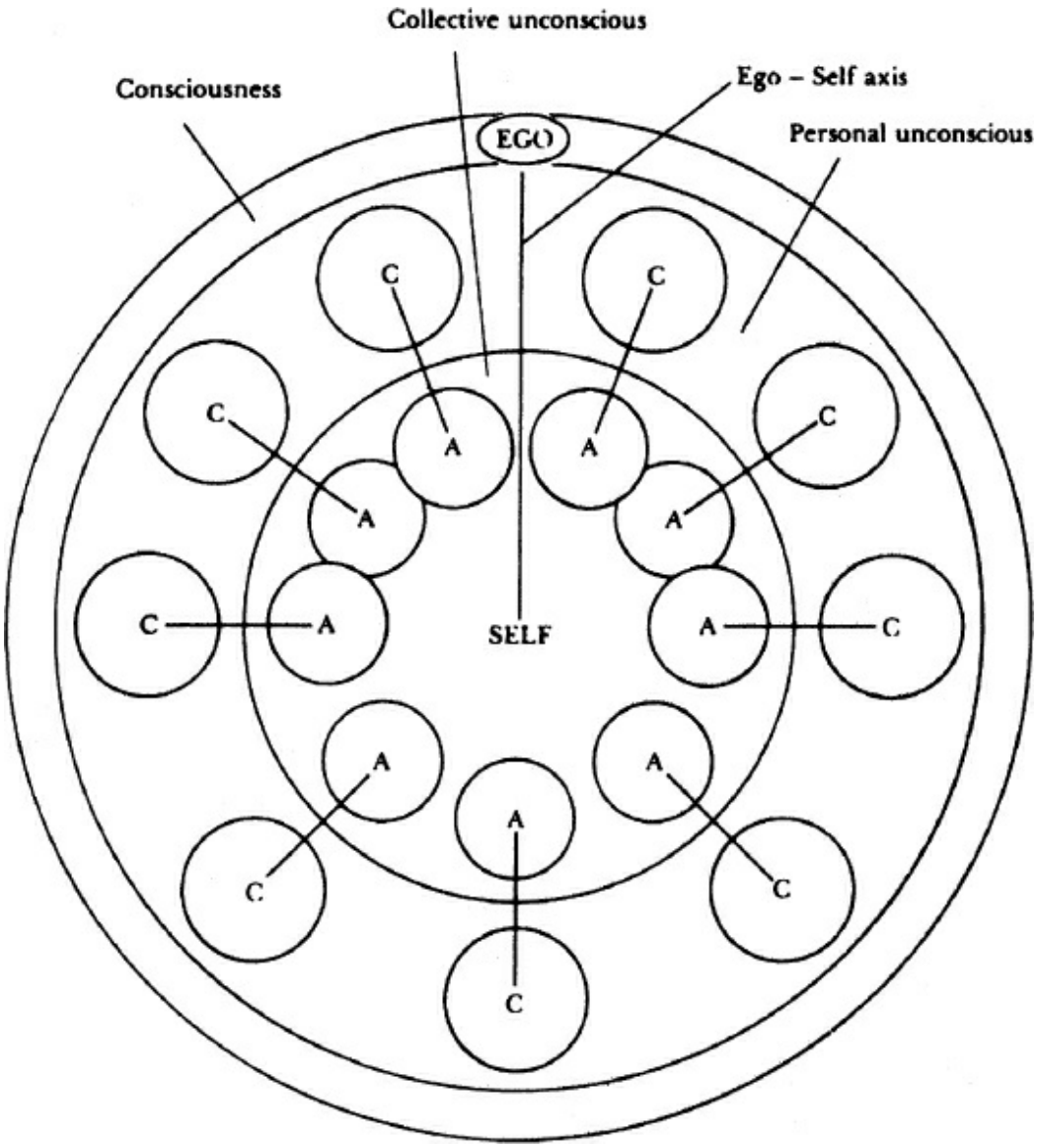

FIGURE 6.3–5. The Jungian psychic apparatus. A, archetype; C, complex. (From Stevens A. *On Jung*. London: Routledge; 1990, with permission.)

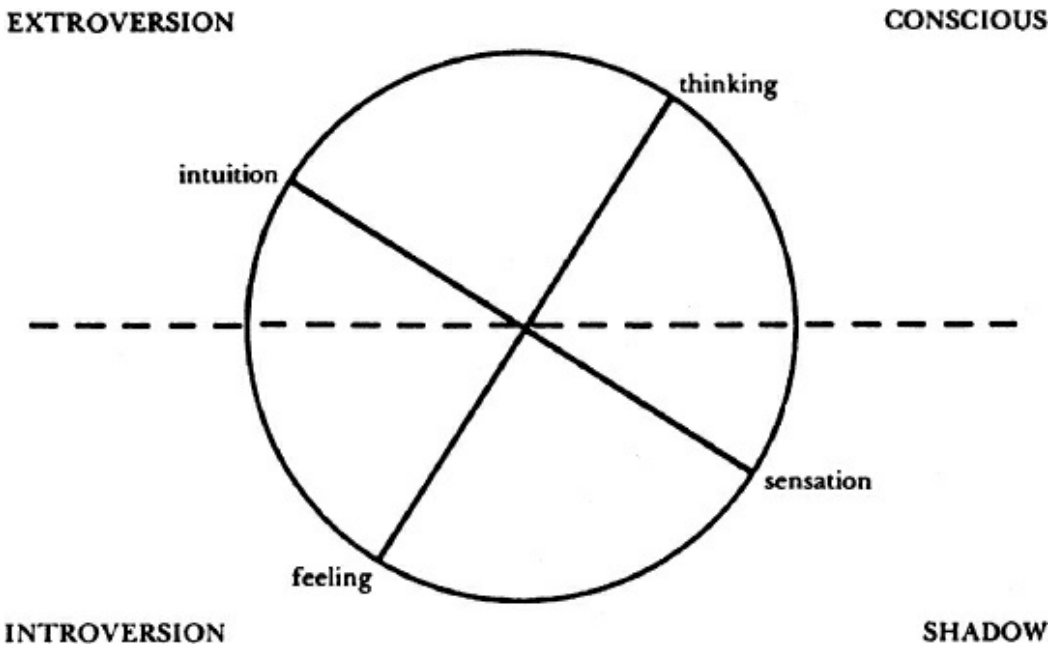

FIGURE 6.3–6. The three personality axes: Extroversion–introversion, thinking–feeling, and intuition–sensation. (From Stevens A. *On Jung*. London: Routledge; 1990, with permission.)

The *thinking–feeling* polarity deals with information processing and judgment. In the *thinking* mode, data are evaluated according to logical principle. *Feeling*, at the opposite pole, is making judgments through nonlogical processes having to do with values and understanding relationships. In social relationships, the thinking type deals with people according to their social rank or according to the tradition of etiquette; a feeling type deals with others in terms of their present social relationship or perceived emotional state.

Each individual has a preferred mode on each of these three dimensions. By placing an individual on each of the three axes indicated in Figure 6.3–6, each individual can be identified as a type. An extroverted-sensation-thinking type is oriented to the real world, tends to perceive external details, and organizes them into a logical structure. An introverted-intuition-feeling type is self-oriented, grasps situations as a whole, and is sensitive to their emotional implications.

Everyone's psyche contains all of the types. However, each person has a superior set of types that are evolved from early life and that are shaped strongly by temperamental factors. In the second half of life, adults who continue the process of individuation attempt to integrate or broaden and deepen their understanding of their inferior functions. The Myers–Briggs Type Indicator, a simple paper-and-pencil test consisting of approximately 40 questions, reliably places individuals on these dimensions. It has become a very popular instrument among psychology, management consulting, counseling, and self-improvement organizations for quickly identifying the qualities of individual participants. This typology may be

the most widely known and accepted aspect of Jung's theories, although it is often used independently of its original context.

Therapeutic Process

Jung suggested that therapy begin with four visits per week and then be spaced out to one or two per week. Present-day Jungian practitioners may work with their analysands once a week, face to face. Jung placed great emphasis on the human relationship between analyst and analysand and noted that both parties change in the course of analysis. He defined *transference* as the patient's attempt to develop psychological rapport with the doctor and held that without rapport and object relationship, the technical operations of the analyst were of no value.

With its emphasis on reintegration, symbolism, and dreams, Jungian analysis seems well suited to help educated people deal with the developmental problems of midlife. Having achieved a professional identity, material success, and established a firm family role, they often begin to ask, "Who is the real me?" "What is most meaningful to me?" and "What is my relationship to humankind and human history?" However, the complicated, often blurred intrapsychic relationships described, and the interest in mystical experience with its potential shading into magical thinking, mean that Jungian clinicians must be careful about the risk of blurring the therapeutic boundaries. The break between Freud and Jung, which had for decades been ascribed to Freud's rigidity and authoritarianism, now appears to have had a contribution from Jung: Freud disapproved of a complicated, dependent relationship Jung is believed to have had with a former patient.

A Jungian View of Mr. A

Jungians see Mr. A's symptoms of bipolar I disorder and use of marijuana as having unleashed conflicting archetypes, all attacking the ego complex he had grown up with. This could be seen as an opportunity for Mr. A to claim a fuller definition of the self. Mr. A's aspirations to be a great writer are taken as potentially healthy, and even the content of his psychotic and substance-induced states is analyzed so as to incorporate their meanings into his ego complex. Mr. A's previous psychological type is described as extroverted-sensation-thinking—an observer of life but not a full participant. The type emerging in his psychotic and intoxicated states is described as introverted-feeling-intuitive—the self-absorbed, self-experiencing lover of beauty and ideas. The dreams about flying suggest an expression of the hero archetype; the anxiety dream suggests some emergence of his shadow in association with his aspirations to be a writer. His love of flowers and creativity during his periods of psychosis and intoxication suggests that the shadow is connected to his anima.

The task of therapy is to assist Mr. A in accepting these alien parts of himself and integrating them into his ego complex. Relatively less emphasis is placed on education or specific treatment of his bipolar I disorder, except insofar as it prevented a good therapeutic alliance. His fantasies, dreams, and psychotic experiences are investigated, discussed, and analyzed in terms of the psychological functions used and the archetypes expressed. His interpersonal issues with his family are seen primarily as projections by him of other unacceptable archetypes that need identification, analysis, acceptance, and integration. For example, his relationship with his mother is seen as a projection and fear of the earth mother archetype. The overall goal is to harness the unleashed unconscious material to allow Mr. A to pursue his ambitious and creative goals.

FRANZ ALEXANDER

Franz Alexander (1891–1964) was one of the second generation of psychoanalysts. He was born in Budapest and attended medical school there, graduating in 1912 (Fig. 6.3–7). He conducted research at the Institute for Experimental Pathology in bacteriology until World War I, when he practiced clinical microbiology on the Italian front, primarily combating malaria. After the war, he joined the department of psychiatry at the University of Budapest Medical School as a brain researcher. This led to an encounter with Freud's work, and, in 1919, he became the first student of the Berlin Psychoanalytic Institute. In 1930, Alexander became a visiting professor of psychoanalysis at the University of Chicago and, in 1932, founded the Chicago Psychoanalytic Institute. He established the institute independent of the psychoanalytic societies, leading the Chicago Institute to be one of the more creative sources of psychoanalytic thought. During the same time, he began his interest in psychosomatic illnesses, helping to found the journal *Psychosomatic Medicine*. A guiding principle of his work was to make psychoanalysis an integral part of medicine.

FIGURE 6.3–7. Franz Alexander. (Courtesy of Franz Alexander.)

In 1946 Alexander became professor of psychoanalysis at the University of Southern California, where he continued his work on ***psychosomatic medicine*** and became interested in the interface of learning theory, the psychophysiology of stress, and psychoanalysis.

Theory of Personality

Alexander did not develop a unique overarching theory of personality; his contribution was his application of psychoanalytic thought to pathophysiological processes. In this he laid the groundwork for the burgeoning fields of psychosomatic medicine, behavioral medicine, and psychophysiology. Alexander created the basis for the biopsychosocial model and studied the mind and body at a time when American psychiatry, despite Freud's original ideas, had become purely psychological in its orientation. In studying and treating many patients with serious physical illnesses, he was also forced to consider creative modifications of therapeutic technique.

Alexander and his group began by intensively studying, by means of clinical interviews, patients who had one of seven illnesses that had been identified by general practitioners as having strong psychological components: peptic ulcer disease, ulcerative colitis, essential hypertension, Graves disease, neurodermatitis, rheumatoid arthritis, and bronchial asthma. Out of these clinical studies emerged the specificity hypothesis, which proposed that certain illnesses are the product of a complex interaction of specific constitutional predispositions, specific unconscious conflicts, and specific types of stressors that activate such conflicts. The group then tested these hypotheses in a series of clinical studies of patient populations. The independent variable was usually the ability of skilled clinicians to predict a patient's illness from disguised case reports.

In the field of psychotherapy, Alexander was one of many who sought to shorten the analytical process. He hypothesized that intellectual insight was not the central curative factor in therapy, a radical and revolutionary proposal at the time. Rather, he emphasized the role of ***corrective emotional experience.*** This led him to experiment with variations in technique that might facilitate such experiences. This, his most controversial stand, nearly ruptured his relations with the psychoanalytic movement.

Theory of Psychopathology

The seven diseases that Alexander studied (peptic ulcer disease, ulcerative colitis, essential hypertension, Graves disease, neurodermatitis, rheumatoid arthritis, and bronchial asthma) were studied to identify the single, specific core conflict that, in interaction with constitutional predispositions and in the circumstance of a particular stressor, would activate the disease.

The core conflict identified in peptic ulcer disease was hyperindependence as a defense against unacceptable dependency needs. The stressor that results in an acute attack can be any situation that demands that the individual openly acknowledge or ask that dependency needs be met. Ordinarily, patients use dominance and control to intimidate others into meeting their dependency needs. This particular illness model received confirmatory evidence from a study of army inductees in whom psychological profiles as described by Alexander, in combination with measurements of serum pepsinogen, were successful in predicting the development of duodenal ulcer. It is not clear what the impact of the recent recognition of a specific pathogen means for this particular hypothesis. Perhaps the identified pathogen, *Helicobacter pylori*, is the constitutional factor that had previously been believed to be serum pepsinogen concentration, or perhaps Alexander's hypothesis only applies to some subgroup of patients with duodenal ulcers.

Alexander's theory of ulcerative colitis also implicated dependency

conflicts; however, rage at unmet needs was seen as the defining feature. This rage provokes guilt and the urge to make restitution toward the object of anger by means of gifts of achievements and successes. The model here is clearly the angry child seeking to placate a parent by means of performance. The precipitating event for reactivation of the illness is the perception that the efforts at placation will be unsuccessful. Alexander claimed that this resulted in excess parasympathetic activity, leading to diarrhea.

Alexander's hypothesis about essential hypertension focused on inhibited anger and suspiciousness in an outwardly compliant, cooperative individual. A hypertensive patient often goes for long periods with blood pressure under good control and then, seemingly inexplicably, experiences a dramatic increase in blood pressure. Alexander attributed these episodes to incidents when the chronic anger is exacerbated, and intense defenses must be used, causing chronic sympathetic fight-or-flight activation. Several psychophysiological studies have suggested that this theory has some validity at least in terms of short-term changes in the blood pressure of a subset of hypertensives. Labile hypertensive patients appear to fit this model best.

The proposed specific conflict for rheumatoid arthritis postulates rebellion against overprotective parents. A compromise formation in which the conflict is discharged via physical activity, especially sports, works for a time, but the anger eventually is expressed in self-sacrifice designed to control others. Failure of this pattern results in increased ambivalent tension, directly expressed via muscular contractions that lead to joint degeneration. Confirmatory evidence for this constellation appeared in a series of psychological test studies; however, more recent research has failed to replicate much of this and has implicated more general stress issues, life change, and psychoneuroimmunological mechanisms.

Alexander proposed that the wheeze of bronchial asthma represented a symbolic cry. The specific conflict, according to him, was the wish for protection versus the fear of envelopment. This conflict leads to sensitization to separation issues, which become the events that provoke the suppressed cry of the asthma attack. In recent years it has become clear that the population with asthma is far more heterogeneous both psychologically and physiologically (in terms of vulnerability to allergens) than was recognized in Alexander's time, and that immunologic variants play a primary role in the illness etiology. The role of a vicious physiological–psychological cycle in which asthma stimulates panic, which in turn triggers pathological pulmonary psychophysiological responses, has been a focus of more recent research.

Although the role of conflicts in neurodermatitis remains widely accepted by clinicians, the specific conflict proposed by Alexander, that early deprivation leads to wishes for closeness that are opposed by a fear of

it, is no longer accepted. Finally, Graves disease (thyrotoxicosis) is no longer widely accepted as a psychosomatic illness. Alexander's hypothesis was that premature responsibility led to a martyr-like denial of dependency.

Treatment

The specificity hypothesis led Alexander to focus his psychotherapeutic efforts in a way that other analysts of his time did not. He reasoned that if he could help patients to resolve their core-specific conflict without necessarily addressing other parts of their personality structure, the medical illness would improve. Indeed, he published numerous case studies suggesting just this kind of success. In addition, he was among the first to question the value of intellectual insight as the curative agent in psychotherapy. He proposed that a corrective emotional experience is the central agent of change. A corrective emotional experience involved disconfirmation within the transference relationship of previous assumptions and projections.

Thus, Alexander felt justified in introducing a variety of techniques that would initially induce and heighten the emotional experience of the transference and, subsequently, challenge the underlying unconscious assumptions. These techniques included manipulating the frequency and length of sessions, making direct suggestions about the patient's life, self-conscious alteration of the therapist's behavior according to the patient's conflict, and behavior therapy techniques. In many ways, this was the most controversial aspect of Alexander's work. Serious questions about the validity of his suppositions and the ethics of his "manipulative stance" were raised. He was impatient with the slow, methodical process of convincing his colleagues; his intellectual energy led him to embark on ever-newer experiments while other analysts were still struggling to digest his previous suggestions. Yet today, few quarrel with the concept that emotional learning is at least as important as intellectual insight in psychotherapeutic success. Alexander's efforts to modify and shorten the analytical process have come closer to the norm in psychiatric practice than has classic analysis. Indeed, his emphasis on a specific focal conflict that could be addressed in brief therapy by modifying techniques anticipated the later work of Peter Sifneos, David Malan, Habib Davanloo, and James Mann, who systematically developed broad-based models of brief psychodynamic psychotherapy.

Ironically, although Alexander's therapeutic innovations seem prescient today, his specificity hypothesis, which was far less controversial, seems simplistic, naive, and forced. He did not have available the sense of how complicated illness causation truly is. The dominant model of the time was infectious disease: One organism, one illness. In addition, the complexity of social phenomena and stressors was unknown in his time. Yet Alexander

was the first to postulate a multicausal etiology for disease: A specific constitutional defect, a specific conflict, and a specific stressor, all necessary for a disease to occur. He was also the first to study mind–body interactions in a systematic way.

Alexander's View of Mr. A

In the absence of a clear psychosomatic illness, Alexander would have had little to say specifically about Mr. A's dynamics. However, like any good dynamically oriented psychotherapist, he would have readily recognized Mr. A's core conflict about dependence–independence. Mr. A would be understood as rebelling against his mother's nurturant control, although secretly desiring to continue basking in it. He has serendipitously found that illicit drugs and his mental illness can be harnessed such that his mother has to rescue him and take over, while he avoids becoming aware of the conflict. Alexander's unique contribution to Mr. A's case would be to propose that he could be treated with 40 or fewer sessions, focusing exclusively on this conflict (and ignoring others such as oedipal components and issues with his father). Further, Alexander's agenda would be to structure the therapeutic relationship to intensify Mr. A's experiencing of this conflict. He might do this by taking a somewhat intrusive, directive stance, thus creating the opportunity for a rapidly intense transference around this issue. Alexander might then spread out the sessions to elicit expression of the patient's yearning for that control. The key therapeutic effort is then to respond constructively to attempts by Mr. A to assert himself appropriately and take over control of his life—for example, by learning more about his illness and taking responsibility for his treatment.

WILHELM REICH

Wilhelm Reich (1897–1957) was one of Freud's most controversial disciples; his later years were marred by mental illness (Fig. 6.3–8). Reich fixed on and elaborated Freud's early, but later discarded, view that neurosis results from the damming up of sexual energy. In so doing, Reich laid the groundwork for a psychoanalytic theory of personality. Before Reich, the focus was on symptoms and psychopathology. There was occasional mention of "character" in psychoanalytic work, but there was no focus on personality as an entity unto itself. This shift is of great importance because, as neurobiology has come to explain more and more Axis I mental disorders, it has become clear that one of psychodynamic theory's enduring strengths involve understanding and treating personality traits and disorders. Together with Anna Freud and Hartmann, Reich is regarded as the originator of ***ego psychology***.

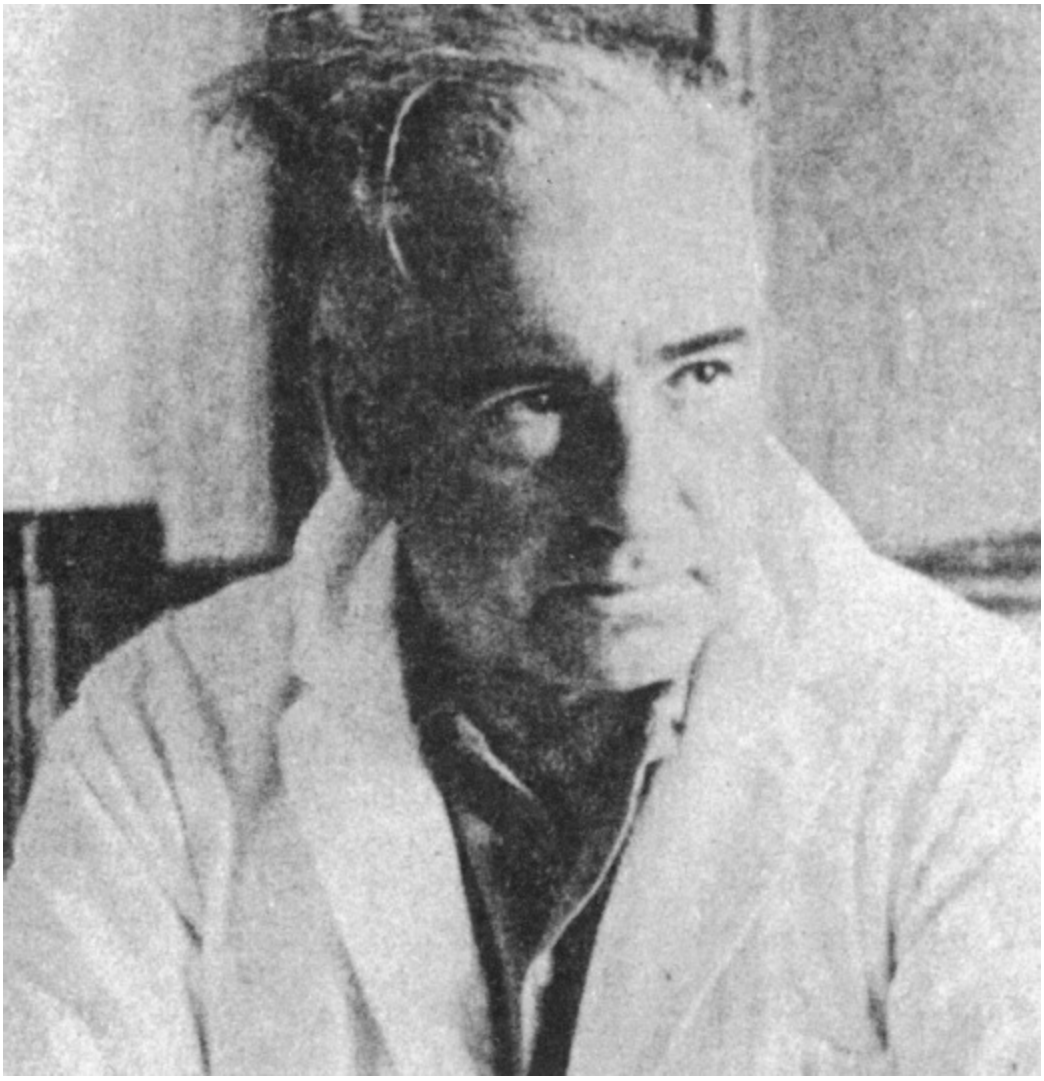

FIGURE 6.3–8. Wilhelm Reich. (Courtesy of New York Academy of Medicine.)

Personality Theory

Reich did not disagree with Freud's notions concerning personality development, including character types based on fixation at specific levels of psychosexual maturation. Reich elaborated the interpersonal and physical behavior of these personality types. Specific behavioral traits constitute ***character armor*** that defends against internal and external dangers. Character armor is comprised of involuntary, repetitive, ego-syntonic behaviors that prevent the emergence of repressed impulses. For instance, the trait of ingratiation frequently defends against hostile impulses, just as the traits of hostility or self-assertion may defend against wishes to be dependent and passive. These traits manifest physically in the voluntary musculature as characteristic postures (clenched jaw or fist, rigid or bowed back) or in excessive fluidity of movement. Although Reich's ideas of the behavioral and postural components being central features of character armor are no longer accepted, his ideas successfully transformed the focus of psychodynamic thought from individual wishes defended

against by individual defenses, expressed in specific transferential forms, into an emphasis on patterns, organization, and intertwining of all of these elements. He also added substantially to the attention paid to nonverbal cues and their implications.

Hysterical Character. The hysterical character has the least body armoring, hence the most lability of function. Body movements tend to be soft, rolling, and sexually suggestive. These individuals are superficial, excitable, flighty, fearful, highly suggestible, and easily disappointed. Their armor helps to defend against easy sexual arousability by flushing out potential sexual stimuli in the environment and then reacting to them with anger.

Compulsive Character. These individuals are tense and restrained, walk stiffly, and sit rigidly. They are overconcerned about orderliness, tend to ruminate, and are indecisive and distrusting. They experience a blockage between their thoughts and feelings. Because they have little access to their feelings, they have little ability to prioritize their actions, to make decisions, or to sense others' reactions to them. The compulsive character avoids expression of repressed impulses by rigid overcontrol. Because of this, these individuals are very threatened by trivial changes in routine.

Phallic-Narcissistic Character. Phallic-narcissistic people appear cold, reserved, and prickly. They are outspoken, provocative, and seek positions of power. Frustrated at the genital-exhibitionist stage of development, the men are identified with the penis and the women with the fantasy of having a penis. The men have strong erective potency but little capacity for intimacy; the women actively dominate men.

Masochistic Character. Masochistic individuals suffer, complain, damage, and deprecate themselves in ways that provoke and torture others. Reich differed with the analytical interpretation that these people enjoy suffering. He believed the opposite—that pleasure is painful for the masochist because of an enormous need, excessive guilt, and the resultant low tolerance for love or pleasure. Suffering allows the masochist to then indulge in a certain amount of self-gratification. Sexual intercourse can be enjoyed, for example, if the partner is inconsiderate or overly domineering.

Treatment

Reich's major contribution was in the realm of treatment. He was the first to recognize the need to deal with character resistances before attempting to recover repressed material and that interpersonal resistances need to be dealt with before free association is possible. He did this by analyzing patients' character armor—their characteristic behaviors (including tone of

voice, posture, and physical movements)—in the analytical setting before proceeding to an analysis of the unconscious. Reich worked face to face with patients and sought to relax their character armor by physical manipulation. This type of therapy, called *vegetotherapy* by Reich, was later practiced by Reich's followers as ***bioenergetics***.

Reichian View of Mr. A

A Reichian analyst might see Mr. A's seeming rebellion against conformity as armor against his fear of being away from his mother's protectionand, more importantly, a fear of domination and dependence. The Reichian might point out that the more Mr. A rebels, the more tightly he binds himself to his mother. The therapist might also suggest that the behaviors that seem so much under Mr. A's control, are in fact masochistic behaviors. His prolongation of his manic states and use of drugs, as well as difficulties with task completion and subsequent praise, help him sustain feelings of need and guilt, and avoid those of true pleasure. The "pride" he takes in embracing mental illness may be viewed as deprecation as a way of provoking others, and maintaining overall suffering. In addition, the family structure does not allow for the direct provocation of Mr. A's mother. As such, masochistic behaviors may be the only way Mr. A can engage in provocation. Of note in Reichian analysis, the therapist would specifically attend to nonverbal cues that would reflect this character armor, and potentially even work directly with the patient to physically relax their body posture.

OTTO RANK

Otto Rank (1884–1939) was a 21 year-old student when he met Freud (Fig. 6.3–9). Rank later earned a doctorate in psychology and eventually became a peer of Freud before ultimately breaking from him. Rank saw each person as an artist whose ultimate task is the creation of an individual personality. In Rank's view, the neurotic was a person whose strong creative urge was stifled by the negative use of will. His own creativity was most clearly expressed in his formulation of birth trauma and his description of a primary relationship with the mother that included hatred as well as love.

FIGURE 6.3–9. Otto Rank. (Courtesy of New York Academy of Medicine.)

Rankian Dialectic

The basis for his break from Freud was Rank's view that the ***birth trauma*** was more important than the oedipal conflict. According to Rank, the physical and psychological experience of birth gives rise to a primal anxiety that is dealt with by primal repression. The crucial intrapsychic conflict that occurs in all developmental phases is the conflict between maintaining the primal bliss of attachment and experiencing the excitement and fear associated with separation. Union stands in contrast to separation; likeness stands in contrast to difference. In adulthood, movement toward another person is possible only if one knows who one is, which can come about only through having experienced separation. Movement toward autonomy is possible only after having established the

sense of belonging and self-worth that derive from the experience of union. Although Rank is rarely cited today in the psychoanalytic literature, his writings anticipated the development of object relations theory, especially in terms of the important emphasis on the preoedipal child's ambivalent experience of the first important attachment—the mother. Rank's dispute with Freud and his circle paved the way for the perspectives that would later become mainstream; for example, that conflictual experience could precede the oedipal phase, and that the early dyadic relationship with the mother could be a source of significant pathology.

Rank viewed movement toward either union or separation not as an innate biological process but as an act of will. In moving toward and engaging with another person, all individuals experience their need for belonging. Moving away from others allows individuals to experience their uniqueness. Maturity is the triumph of will over the forces that inhibit movement both toward and away from others—guilt, death fear, and life fear.

Rank saw guilt as the price to be paid for any act of will. Moving toward union causes guilt over being needy; moving away causes guilt at abandoning another person. ***Death fear*** is the fear of losing one's identity by fusing with another person. The weaker one's personal identity, the stronger the death fear. ***Life fear***, by contrast, is fear of losing all ties in the process of becoming separate. Every person experiences the cycle of movement from union to separation and back again as part of the life process. This movement takes place at various levels: Family, societal, artistic, and spiritual. At each level, there is movement toward union and rebirth. Each person, for example, hopefully yields to a love experience in which personal differences are set aside to experience unity with another, to experience self-worth, and to be relieved of the sense of difference. This yielding to another ends when the will asserts its separateness and a new affirmation of individuality occurs.

Will, the prime mover in the Rankian dialectic, is an irreducible creative force. It is not solely an agency for the expression of Freudian sexual or aggressive impulses nor is it the will to power in the Adlerian sense. The beginning of will is in the child's "no," an assertion of what the child will not do. In maturation, will becomes a positive force. Neurotic people, however, deny will because of guilt. They deal with that guilt by using defense mechanisms such as projection and rationalization. Viewed from this perspective, neurotics are strong-willed people who cannot acknowledge what they will or even that they will. As a result, they cannot use their will constructively in the service of their greatest potential artistic creation, their own personalities.

Treatment

Rankian psychotherapy is a here-and-now interaction with the therapist

that seeks to mobilize the patient's will, resulting in a ***rebirth experience***. The treatment, which is time limited, focuses on the relationship with the therapist. In the therapist–patient relationship are reenacted earlier life struggles, especially struggles involving intimacy. After patients are strengthened through the therapist's acceptance, they begin a process of negative will assertion that is seen as resistance in classic analysis. Rank regarded this negative will assertion as indicative of growth and supported it. Now able to provide self-affirmation on their own, patients free themselves of the therapist and begin to individuate, which is ultimately expressed through terminating the treatment. They overcome the life fear by living up to their fullest potential. Therapy is not aimed at reconstructing personal history; it is a struggle in the here-and-now between the patient and therapist as a representative of transference objects and reality.

The therapeutic process parallels the process of personality growth. At first, the therapeutic relationship recapitulates prototypical early relationships. The first rebirth experience for patients is claiming their own individual personalities and their uniqueness as human beings. The second phase is their discovery of the physical universe and their likeness to it. Later, they claim their distinctness as creators of themselves. With the emergence of the self, individuals unite with ideological, philosophical, and spiritual reality and experience the final birth of the ideal person, a self-fulfilled person who no longer needs to create to justify his or her existence. In Rank's theories, the beginning of the shift to what is now called *pregenital issues* in psychoanalytic thought can be seen.

Rankian View of Mr. A

A Rankian therapist sees Mr. A as attempting to achieve separation from his family of origin but as lacking the will to achieve real separation. The prolongation of his manic episodes and his use of drugs help him to avoid the pain of separation and individuation, and his inability to complete graduate school allows him to maintain his dependence on his family. Initially, the Rankian therapist might approve of Mr. A's attempts to sustain his manic episodes and his marijuana use as a negative assertion of will against his mother's efforts to run his life. The therapist then would suggest that developing more constructive and positive means might enable Mr. A to become freer in reality. The Rankian therapist might challenge Mr. A to develop attachments outside his family and to develop skills to help free himself. Resistance to the development of these skills is interpreted as fear of separation and aloneness. Mr. A is then encouraged to develop his will in a positive direction, following a path of action that he desires instead of rebelling against a path of action seemingly dictated by others. Finally, Mr. A would work toward separation from the therapist as a prototypical separation-individuation experience, or a "rebirth."

KAREN HORNEY

Physician–psychoanalyst Karen Horney (1885–1952), who emphasized the preeminence of social and cultural influences on psychosexual development, focused her attention on the differing psychology of men and women and explored the vicissitudes of marital relationships (Fig. 6.3–10). She was one of three women whom Freud trained specifically to have female analysts who would contribute their unique perspective to the psychoanalytic theory about women; the other two were Anna Freud and Helene Deutsch. Horney's view that repression and sublimation of biological drives are not the primary determinants of personality development led to her removal as an instructor in the New York Psychoanalytic Institute and her founding, in 1941, of the American Psychoanalytic Institute.

FIGURE 6.3–10. Karen Horney. (Courtesy of the Association for the Advancement of Psychoanalysis, New York.)

Personality Theory

Horney believed that personality development results from the interaction of biological and psychosocial forces that are unique for each individual. At the core of each personality is an enduring real self. Partially equivalent to the Freudian ego and similar to Donald Winnicott's focus on selfhood and partly to Berne's child ego state, the real self combines choice, will, responsibility, and identity with spontaneity and aliveness. A natural unfolding process of self-realization leads to the development of human potential in three basic directions: Toward others, to express love and trust; against others, to express healthy opposition; and away from others toward self-sufficiency.

Although conditions during childhood may block psychological development, healthy growth is always possible if the internal blockages can be removed. Children whose family situation leads them to feel endangered concentrate on psychological survival and may do so at the cost of developing stereotyped coping mechanisms. All human beings have basic anxiety, which Horney saw as the normal response to the infant's helplessness and separateness. How families respond to this fundamental situation, guided by cultural norms, determines whether individuals spend their lives struggling with basic anxiety or pursuing self-realization. Horney believed that the attributes of passivity and suffering were not biologically specific to women, as taught by the analysts of her day, and that male and female personalities are in fact ***culturally determined***.

Theory of Neurosis

Horney defined neurosis in both intrapsychic and interpersonal terms. She noted that her patients complained not of the symptomatic neuroses, such as phobias and compulsions, but of unhappiness, blockage, lack of fulfillment in their work, and inability to establish or maintain relationships. She saw these individuals as having a complex system of self-perpetuating defensive patterns against basic anxiety—***character neuroses***.

Safety-seeking children move psychologically in three directions to relieve their anxiety, make life safe and predictable, and achieve satisfaction. They seek affection and approval, become hostile, or withdraw. Children eventually use the coping strategy that best meets their needs, but if only one basic strategy is used, children become limited in their coping repertoire and in their experience of themselves and their world. Their sense of safety is tenuous because there is now danger from within, from suppressed or repressed feelings and impulses. Given continued

unfavorable environmental conditions, conflicting feelings are driven into the unconscious, and such children are left with a sense of discomfort, anxiety, apprehension, and an insecure sense of self. At this juncture, their point of reference is externalized; patterns of behavior rigidify and increase blockages to growth develop. Horney designates these complex, relatively fixed attitudes toward self and others as ***neurotic trends***.

Character Types

Horney's three main character types are based on the predominant mode of relating to others. The *compliant, self-effacing type* results from the defensive operation of clinging to others. These individuals try to curry favor with others, subordinate themselves, and are reluctant to disagree. The *aggressive, expansive type* results from moving against others and relies heavily on power and mastery as a means to achieve security. The *detached, resigned type* results from moving away from others in an attempt to avoid both dependency and conflict. These are very private individuals who, although refusing to compete openly, see themselves as rising above others.

Idealized Image. As teenagers, individuals who grow up to be neurotic create an ideal image that, if achieved, promises to end their painful feelings and provide self-fulfillment. The idealized image covers over all the contradictions, conceals the defensive nature of their behavior, and restores a sense of wholeness. Energy formerly available for self-realization is now used in efforts to become like the idealized image. For example, an individual who has adopted the strategy of moving toward others and is consequently dependent on others for affection and approval experiences the fear of reasonable self-assertion, striving instead for saintly humility and considerateness of others.

Because the ideal self is imaginary, neurotic people are readily bruised by confrontation with reality and work excessively to prove they are in fact their ideal selves. This results in a type of perfectionism that insists on flawless excellence in which "I should" replaces "I want" or "I need." It also results in the neurotic ambition to be first and in a strong drive for revenge against those perceived as having interfered with their efforts to match the ideal self. This aspect of Horney's theories anticipates some of Heinz Kohut's ideas on the origins of narcissism.

Claims, "Shoulds," and Self-Hatred. Despite their frequent self-disparagement, neurotic individuals expect to be treated as though they were their ideal selves. These claims to special treatment, when frustrated, produce anger, righteous indignation, and resentment. The "shoulds," or self-imposed demands that they live up to their idealized selves, are irrational and unrelated to the realities of daily life. They are projected,

experienced as demands made by others, and are also demanded of others. This results in neurotic people being critical of others and very sensitive to criticism themselves. Self-hatred results when the threat arises that neurotic individuals may be unable to achieve their idealized selves. If support is not needed for the idealized self, claims, shoulds, and self-hatred are not such important parts of the psychic apparatus.

Neurotic Pride and the Pride System. Glorifying aspects of the idealized self, *neurotic pride*, substitutes for healthy self-confidence. Thus, when their pride is injured by others, neurotic individuals become enraged and seek to avenge their injury and to conceal their self-deception by achieving a vindictive victory over the offending person. Together with supporting claims and shoulds, neurotic pride and self-hatred form a defensive network or pride system that protects the idealized self. Despite the armoring of their defensive network, such individuals are not at peace because they are in inner conflict with the forces that protect them. The conflict between the forces driving toward healthy self-realization and the pride system is the *central inner conflict*.

Alienation. ***Alienation*** from self is one of the most serious consequences of neurotic development. It results from the combination of repeated denial of external reality and the repression of genuine thought, feelings, and impulses. As the process of alienation continues, neurotic individuals lose touch with the core of their being and can no longer determine or act on what is right for them. Their feelings may range from uncertainty and confusion to inner deadness and emptiness.

Therapeutic Process

Horney did not regard adult neurotic people as recapitulating childhood experiences and thus did not focus on the recovery of childhood memories; she dealt instead with the self-perpetuating neurotic process. She stressed the importance of dreams in analysis. As unconscious attempts to solve conflicts, dreams can show constructive forces at work that are not yet discernible in patients' conscious thoughts and behavior. Later, she stressed the exploration of the patient–analyst relationship. She was one of the earliest analysts to recognize and make constructive use of her own feelings toward patients. To Horney, psychoanalysis was a cooperative venture that enabled patients to free themselves from their neurotic structures and mobilize themselves toward self-realization. The analyst's responsibility was to assist in liberating patients from *blockages*, the forces that impede healthy growth.

Early in therapy, termed the *disillusioning process*, the two types of blockages are identified and examined. The first group of safety-oriented blockages, *protective blockages*, helps to avoid the anxiety caused by self-

awareness. They include silence, lateness, depreciating the analyst, the use of drugs, and even the use of self-accusation as a means to avoid further exploration. *Positive-value blockages* reinforce patients' satisfaction with themselves and support their idealized selves. In the disillusioning process, the analyst identifies both types of blockages, exposing the protective blockages before exposing the blockages that defend the idealized image. Analyzing the positive-value blockages first arouses too much fear. As the central conflict is resolved successfully, patients move into the final phase of treatment: The discovery and use of their real inner self.

Horney's View of Mr. A

From the standpoint of Horney, Mr. A's process of self-realization has been blocked in all three directions. He has not developed the ability to love and to trust; he expresses opposition in an unhealthy way, and he has made self-defeating moves toward independence. He is seen as having developed a detached style of relating to others, having substituted hypomania and the use of drugs for real relatedness. He justifies his illness by taking pride in it. His goal of becoming a writer is part of the development of an ideal self that is a detached observer. He supports his ideal self by having fantasies of involvement with other writers, by convincing his professor that he was able to write a play as his term paper, and by procrastination. He has become more and more isolated from his real self by denying the reality of the facts that his mental illness and cannabis abuse pose a danger to him and that he was failing in school.

A therapist in the Horney tradition begins by pointing out that Mr. A's drug use and his sustaining of manic episodes are blocking his ability to learn about himself and works with Mr. A toward abstinence from substances of abuse and appropriate use of mood-stabilizing agents. The therapist later begins to point out that Mr. A's pride in his manic episodes serves the purpose of sustaining a false self of boundless energy and creativity. Mr. A is encouraged to decide whom he really wants to be—a writer in fantasy or a person able to obtain real satisfaction from real accomplishments and real relationships. His dreams are examined and their themes explored. His omnipotent, messianic dreams are interpreted as evidence of his ideal self and viewed in the light of reality. His dreams of being exposed are interpreted as the fear engendered by his ideal self as a means of defending itself against exposure: "If you expose me, you will be embarrassed and humiliated." Mr. A is supported through his fear of humiliation and the pain of realizing that he has been deceiving himself. His self-loathing is interpreted as the activity of the pride system in defending his ideal self. As he begins to relinquish his ideal self, he will begin to discover and to mobilize his real inner self in directions that might not have been predictable at the beginning of treatment.

ERICH FROMM

Psychoanalyst Erich Fromm (1900–1980) was often thought of as the archetypical neo-Freudian, the leader among those who emphasized that culture and social setting influence an individual's dynamics as much as instincts do (Fig. 6.3–11). Neither physician nor biologist, Fromm, a native German, received his doctorate in philosophy, sociology, and psychology from the University of Heidelberg in 1922. There he was exposed to a Marxist emphasis on how history shapes societies and how societies, in turn, shape individuals according to economic needs. He was trained as a psychoanalyst at the Berlin Psychoanalytic Institute and then founded, with his wife, Frieda Fromm-Reichmann, the Frankfurt Psychoanalytic Institute. In 1933 he immigrated to the United States and in 1949 moved to Mexico City to found another psychoanalytic institute. In 1974 he moved to Switzerland, where he died in 1980. As much a social critic as a personality theorist, he was later claimed by the existential and humanistic psychoanalysts. Fromm's intellectual agenda was the integration of Freud's theory of a dynamic unconscious with Karl Marx's theory of history and social criticism.

Personality Theory

For Fromm, two central facts dominate human behavior: The inevitability of separateness and the ***historical and social moment*** into which each person is born. He argued that every person yearns to recapture the state of blissful union that existed prenatally. From the moment the baby begins to recognize itself as a separate human being, a titanic struggle begins, pitting the desperate anxiety of loneliness against the urge to fully express and actualize oneself, and, ultimately to transcend the self. Most individuals find the loneliness too painful to bear and suppress their striving for individuation in the service of maintaining the illusion of connectedness. They are socialized by their parents into the roles defined by their society of birth. Fromm actually used the term *symbiosis* years before Margaret Mahler employed it to describe the universal human yearning for fusion, safety, and security.

FIGURE 6.3–11. Erich Fromm. (Courtesy of Erich Fromm.)

Facing aloneness and choosing individuation lead to freedom and a productive life. However, true freedom is too terrifying for many people, who instead construct a series of illusions that engender the feeling of safety. They create a pseudo-self, think pseudo-thoughts, and experience pseudo-feelings in support of these illusions, thereby cutting themselves off from the fullness of their own inner lives. Fromm saw Freud's theory as a special case of his own more general ideas. The illusions that Victorian society offered involved sublimation of sexuality and aggression in the service of social respectability. Social respectability, in turn, provided the illusion of acceptance and security. In other places and at other times, different solutions might be offered. Early in World War II, shortly after his escape from Nazi Germany, Fromm wrote about the willingness of people to give up their freedom to serve an authoritarian society. In the 1950s he wrote of the pursuit of material acquisitions in the service of postwar productivity, leading to self-satisfied conformity.

To achieve true freedom, Fromm said four basic human needs must be met: *Relatedness*, *transcendence*, *identity*, and a *frame of orientation*.

Relatedness is the need to feel connected to other humans. *Transcendence* refers to rising above basic instincts. *Identity* is the need to feel accepted yet unique. Emphasis on the need for a *frame of orientation* led Fromm late in his career to an exploration of the constructive and destructive roles that religion may play in individual lives.

Theory of Psychopathology

As a social philosopher and critic, Fromm did not really develop a systematic theory of psychopathology. He identified three major mechanisms of retreat from individuation. Some individuals, he said, may seek an authoritarian solution, trying to live through someone or something external to themselves to achieve a sense of adequacy. Others may become destructive, attacking anything that confronts them with their separateness and aloneness. Most individuals develop a conformist attitude, warding off the anxiety of experiencing their own intentionality by accepting socially offered thoughts, roles, and attitudes.

These mechanisms result in four unproductive orientations or ***characters*** typical of modern capitalist society: *Receptive*, *exploitative*, *hoarding*, and *marketing*. The *receptive character* often appears to be cooperative and open; however, the primary agenda is to establish a passive relationship with a leader who solves problems magically. *Exploitative characters* are likewise interested in filling themselves up from the outside; however, they aggressively manipulate and usurp whatever reduces their terror, for example, power. *Hoarders* collect, store, and close in on themselves, often being cold and aloof in their efforts to feel secure. *Marketers* treat themselves as a plastic commodity to be manipulated as needed to achieve externally validated success.

Treatment

Fromm wrote nothing at all on the practice of psychotherapy; therefore, what is known is derived from anecdotal reports by those who studied with him or were treated by him. They report his emphasis on a tender and empathic inquiry into the self-deceptions and illusions created by patients in their efforts to ward off the anxiety of separateness and to maintain some sense of connectedness to significant others. He placed great emphasis on the tendency of unloved children to identify intensely with parental values to capture the magical safety they seem to offer. At a time when most psychoanalysts were preoccupied with detailed examinations of the instincts and defenses, Fromm contributed a sense of the range and richness of inner experience that underlies superficial adaptation. He contributed the sense that a new authenticity could be found by those willing to confront the truth about themselves with all its terror of aloneness.

Fromm's View of Mr. A

In some respects, Fromm's ideas are uniquely applicable to Mr. A, who is clearly maintaining the illusion of his separateness by conforming to the socially sanctioned role of rebellious artist. The bipolar I disorder becomes an unexpected weapon in his battle with his terror of aloneness. The dread in the hotel lobby dream is probably the closest this fear comes to consciousness. Fromm would gently probe this dream and Mr. A's self-image. He would point out the chains created by Mr. A's seemingly rebellious independence. The self-destructive quality of the patient's lifestyle would also be confronted and investigated. Ultimately, Mr. A has to experience his own loneliness and face his terror of it to find true freedom, true intimacy, and an authentic self-definition.

HARRY STACK SULLIVAN

Harry Stack Sullivan (1892–1949) is generally acknowledged as the most original and distinctive American-born theorist in dynamic psychiatry (Fig. 6.3–12) and most American psychiatrists make significant use of concepts and approaches he developed. For many years the primary theoretical dispute within dynamic psychiatry circles was between the classic Freudians and the Sullivanians (or interpersonal psychoanalysts). When psychiatrists use the term *parataxic distortion*, apply the concept of self-esteem, consider the importance of preadolescent peer groups in development, or view a patient's behavior as an interpersonal manipulation, they are applying concepts Sullivan first proposed.

FIGURE 6.3–12. Harry Stack Sullivan. (Courtesy of the New York Academy of Medicine.)

Sullivan graduated from medical school in Chicago in 1917. He spent from 1921 to 1930 in the Washington, DC, area working with schizophrenia patients at St. Elizabeth's and then Sheppard and Enoch Pratt Hospitals, where he developed a reputation as a remarkable clinician with an uncanny ability to communicate with floridly psychotic patients. He initiated the first of what are now called ***therapeutic communities***. Later, he entered private practice in New York and eventually returned to the Washington area, where he was involved in clinical, consulting, and teaching activities. In the 1920s and 1930s, he wrote a number of papers on schizophrenia, later collected in *Schizophrenia as a Human Process*. His other books were

compiled from his lectures by his students; most were published posthumously, explaining some of the density and seeming disorganization of his written work.

Personality Theory

Sullivan rejected the Kraepelinian dogma of his day that dominated psychiatric thinking about schizophrenia. In searching for alternative understandings of psychosis, he turned initially to Freud but rejected his theories as increasingly rigid and dogmatic. Thus, he developed his own working theory of personality, psychopathology, and therapy.

Sullivan believed language could be misleading and was therefore wary of self-reifying conceptualizations that led to rigid theories. He emphasized the psychiatrist as participant/observer in the clinical situation seeking to keep observations as objective as possible, while recognizing the difficulty this presented in dealing with private emotional experience. What can be observed is the social interaction of patients; thus, he defined ***personality*** as the "relatively enduring pattern of interpersonal relations which characterize a human life." From the outset, his focus was very different from the intrapsychic emphasis of psychoanalysis. By approaching psychopathology in this way, he necessarily created a field theory, characterized by temporal and interactive processes, rather than a structural theory.

Sullivan's theory is fundamentally one of needs and anxiety. *Needs* are divided into needs for satisfaction and needs for security. *Anxiety* occurs when fundamental needs are in danger of not being met and is the primary motivator of human behavior. Needs for satisfaction include physical needs (e.g., air, water, food, warmth), and emotional needs include need, especially, for human contact and for expressing one's talents and capacities. Because infants are utterly unable to meet their own needs, interpersonal relationships are crucial. Decades before Mahler wrote of a symbiotic stage in infant development, Sullivan spoke of the "empathic linkage" between caretaker and infant and described the complicated interaction of infants communicating tension and anxiety, arousing anxiety in the caretaker, leading to tender responses to the infant's needs. Failure to meet these needs results in loneliness and anxiety.

Sullivan defined *security* as the absence of anxiety. Thus, *needs for security* are defined as the need to avoid, prevent, or reduce anxiety. Because there is no such thing as a perfect mother or parent, anxiety is inevitable and becomes the primary driver for personality development. The *self-system* is defined by Sullivan as the dynamism that is responsible for avoiding or reducing anxiety. Sullivan equated the self, identity, or ego with the individual's developed patterns for avoiding the discomforts that arise from the inability of others to meet one's fundamental needs. It exists, like all else, purely within an interpersonal framework.

The self-system develops a set of mechanisms, called *security operations*, which effect this goal. Security operations function within Sullivan's theory much as defense mechanisms do within psychoanalytic theory. The specific security operations, however, were defined interpersonally, and Sullivan tried to link them closely to actual observation or experience. Some bore the same labels and definitions as Anna Freud's, but Sullivan is best known for three contributions that bore his distinct stamp: Apathy, somnolent detachment, and selective inattention. These were drawn from observing the way infants and young children react to painful interactions, such as scolding, with their parents.

The self-system accrues from ever-evolving interpersonal experiences—that is, fulfillment of needs for satisfaction as a result of the empathic linkage with the mother. The most difficult experiences are not necessarily those involving the inability to meet the child's needs, but the child's sensing of the caretaker's anxiety in the process of responding to those needs. This arouses anxiety in the child, promotes the need to establish a sense of security, and leads to evolution of the self-system and the development of security operations. The self-system is divided into three parts. The "good me" is a set of images, experiences, and behaviors associated with an unanxious, tender, empathic, and approving or accepting response from the environment. The "bad me" comes to be associated with ideas, actions, and perceptions that provoke anxiety and disapproval from caretakers. Some situations, however, provoke such intense anxiety that they are entirely disavowed and disowned; they become part of the "not me." Eventually, the empathic linkage becomes unnecessary and the self-system operates autonomously within the individual, developing ever more subtle and complex ways to manage the person's anxiety.

Developmental Theories. Sullivan had two theories of development, one cognitive, the other social. He postulated three developmental ***cognitive modes*** of experience whose degree of persistence into adulthood is important in understanding psychopathology. The *prototaxic* mode, characteristic of infancy and early childhood, involves a series of disconnected, brief states experienced as totalities with no temporal relationship. In later life, mystical experiences and schizophrenic fusion represent persistent prototaxic experiences. *Parataxic* experience begins early in childhood as the self-system begins its more independent functioning. It, too, involves a series of momentary experiences; however, they are now recorded in sequence and with apparent connection to one another. They may be given symbolic meanings, but rules of logic are absent, and coincidence plays a major role in how the world is perceived. The self-system uses this mode to seek effective anxiety-reducing behaviors and to repeat them, seeking sameness and predictability. Sullivan used this

mode to explain transference, slips of the tongue, and paranoid ideation. The *syntactic* mode of experiencing is based on the development of language and consensual validation. The world and the self are perceived within rules of logic, temporal sequencing, external validity, and internal consistency. Thinking about oneself as well as others becomes testable and modifiable based on rigorous analysis of experiences in a variety of different situations. *Maturity* may be defined as extensive predominance of the syntactic mode of experiencing.

Social development is somewhat based on these evolving cognitive modes. However, disturbed interpersonal relationships may cause persistence of the more primitive (prototaxic or parataxic) ways of experiencing the world. Infancy spans birth to the onset of language and is characterized by the primary need for bodily contact and tenderness. The prototaxic mode predominates, and the primary zones of interaction are oral and, to some extent, anal. Insofar as needs are fulfilled with a minimum of anxiety, the infant experiences euphoria and a sense of well-being. To the extent that some anxiety is commonly present in the caretakers, apathy and somnolent detachment are regularly used as security operations, persisting into adult life as a basic detached and passive stance. If anxiety and inconsistency are severe, intense experiences of dread persist, presenting in later life as the eerie, uncanny, bizarrely disruptive internal states seen in individuals with schizophrenia.

Childhood begins with the onset of usable language, continues until the beginning of school, and is characterized by the child's focus on the parents as the other from whom praise and acceptance are sought. The primary mode of experience shifts to the parataxic, and the most common zone of interaction is anal. The child needs an approving adult audience. This leads to a variety of learning of language, behavior, self-control, and so on. It can also be observed in a variety of trial-and-error efforts by the child to find what pleases the adult. Gratification leads to an expansive self-system with many facets of life associated with the "good me" and positive self-esteem. Moderate anxiety leads to chronic anxiety, uncertainty, and insecurity. Extreme anxiety results in giving up known successful behavior in favor of self-defeating patterns that fulfill others' expectations.

The juvenile era covers ages 5 to 8 years. The shift to syntactic cognitive modes begins, and the interpersonal focus spreads to the peer group and outside authority figures. Peers and teachers have the opportunity to approve and accept behavior previously frowned on within the family (e.g., talking dirty with one's friends). Interpersonal cooperation, competition, play, and compromise become the gratifying experiences. Juveniles learn to negotiate between their own needs and legitimate social concerns without sacrificing their self-esteem in the process. The risks of excessive anxiety are either too great a need to control and dominate social situations or they become an internalization of restrictive, prejudicial social attitudes.

Preadolescence, ages 8 to 12 years, marks the child's movement from peer group cooperation and competition based on roles toward genuine intimacy with a chum. Sullivan saw this phase as a particularly important stage in which the give and take of the special friend could repair and undo distortions that resulted from excessive anxiety at earlier stages. This is the point at which the individual truly moves outside the family and engages in a free, creative interaction with another person unfettered by the same dynamics. During this stage, the major shift toward syntactic thinking takes place, although some distortions may persist into adolescence. The preteen years see the initiation of a capacity for attachment, love, and collaboration or their inability to develop in the face of excessive anxiety. Although sexual exploration may be a part of the chum relationship, Sullivan did not see sexuality as a central element in this developmental phase.

Adolescents, beginning at puberty, are seen to have concerns similar to those of preadolescents, except that sexual desire is added to the interpersonal equation. Thus, the same needs for a special sharing relationship persist but shift to the other sex for their outlet, whereon a major opportunity for learning or severe anxiety begins. As the person faces culturally defined stereotyping, many new opportunities for social experimentation may lead to consolidation of self-esteem or self-ridicule. The struggle to integrate lust with intimacy is accomplished by painful trial and error. If this is completed with the self-system relatively intact, the later years of adolescence are an opportunity to expand the syntactic mode to such areas as a consensual view of interpersonal relations, values and ideals, career decisions, and social concerns (Table 6.3–2).

Table 6.3–2.
Sullivan's Social Development Theory

Stage	Age	Mode	Characteristics	Hurdles
Infancy	Birth–Language onset	Prototaxic	Focus on need fulfillment	Excessive anxiety can lead to detachment and sometimes later psychosis
Childhood	Usable language	Parataxic	Focus on pleasing others	Excessive anxiety can lead to chronic anxiety and possibly self-defeating patterns
Juvenile	5–8 years	Shift to Syntactic begins	Focus of pleasing spreads to peer group and authority figures	Excessive anxiety can lead to need for control/domination or internalization of prejudicial social attitudes
Preadolescence	8–12 years	Significant shift to	Initiate capacity for attachment, love, and	Excessive anxiety can block ability to

		Syntactic	collaboration. Importance of a single close friend.	expand capacity for attachment and love
Adolescence	Puberty	Syntactic	Incorporation of sexuality in relationships; later years involve solidifying ideas regarding relationships, values, career, and social concerns.	Excessive anxiety in the face of new intimate relationships and painful trial-and-error can lead to self-ridicule

Theory of Psychopathology

Sullivan abhorred diagnostic labeling for being unhelpful, overly restrictive, dehumanizing, and used primarily to impress patients and colleagues. In discussing schizophrenia, he said, "We are all much more simply human than otherwise." Thus, he sought to understand the fundamental human process within his patients, especially his sickest ones. He saw psychopathology as resulting from excessive anxiety arresting development of the self-system thereby limiting both opportunities for interpersonal satisfaction and available security operations. He viewed psychiatric patients as struggling to maintain their self-esteem with very limited means. To understand them, the developmental phase at which they operate has to be gauged, and the interpersonal needs they express have to be understood.

Sullivan believed that several factors could play a role in the particular form that these disturbances might take. The level of anxiety at particular developmental stages can lay the groundwork for a developmental arrest. Basic cognitive capacity might play a role in the choice of security operations relied on or retained. The degree of success achieved interpersonally combined with whatever capacities are used affects later success. Finally, the chance occurrence of stresses encountered during life is deemed a factor. Thus, Sullivan theorized that anyone might develop schizophrenia, even people with relatively successful developmental histories, should their chosen defenses fail dramatically and their life stresses mount in the extreme. However, it was more likely that schizophrenic patients would be highly vulnerable along all four dimensions, whereas others with greater developmental strengths might become obsessive, hysteroid, schizoid, or paranoid.

Interpersonal Psychotherapy

Sullivan's ***interpersonal psychotherapy*** approach emphasized that the psychiatrist is a participant–observer in all interactions with patients. He thought deeply and extensively about the nuances and opportunities involved in this unique situation. By interacting actively with patients, verbal and nonverbal expressions of recurrent interpersonal patterns become apparent. These observations then inform the therapist's further behavior, thereby creating the opportunity for change. This process occurs

over seconds and over months and years as the psychotherapy unfolds. Sullivan saw this perspective as an antidote to what he perceived as the wrongheaded emphasis on objective neutrality embodied by the "blank screen" model of psychotherapist behavior. He argued that parataxic distortions emerge in all interactions, not only in the classic analytical situation. This differing view of transference and of it being a universal human process was among the core debates for decades between classic analysts and interpersonal analysts.

Sullivan saw therapy as elucidating the patient's interpersonal patterns, exploring their usefulness in the service of the patient's needs, and considering alternative, more favorable possibilities. Thus, he shared the ego psychologists' understanding that even the sickest behavior was the best adaptation available at a given moment to the patient. He emphasized the experiencing of the distortions, the needs, the patterns, and the potential changes within the ongoing interaction with the therapist. He saw great power in the very entanglement of the therapist with the patient and recognized the ability of a skilled therapist to manage the interpersonal process to reveal patterns and to shape the patient's emotional experience. However, he constantly emphasized and respected the ultimate autonomy of his patients, who could still, in the end, choose not to reshape their approach to the world.

Sullivan viewed psychotherapy as divided into four distinct stages: *Inception*, *reconnaissance*, *detailed inquiry*, and *termination*. *Inception* involves the very beginning, often only a part of the first interview, during which the contract and roles are stipulated. *Reconnaissance* might go on for as many as 10 to 15 sessions, during which the therapist identifies the patient's recurring patterns and assesses their adaptive and maladaptive qualities. The *detailed inquiry* is a very lengthy process of exploring the patient's thoughts, feelings, and memories and evaluating and reevaluating data from earlier stages, seeking to recognize, clarify, and change persistent parataxic distortions. The recurrent patterns are discussed within the context of the person's developmental history, needs, anxieties, failures, and successes. There is often much ongoing interchange between patient and psychiatrist as feelings and perceptions are validated or questioned within the context of mutual emotional interchange within each session. *Termination* is a product of the evolving contract and understanding between the patient and therapist and may reflect either extensive or limited goals. Sullivan emphasized the constant reassessment of goals by the therapist and the power of the ongoing negotiation and renegotiation of the therapeutic contract as a means to reveal and change parataxic distortions. The ultimate goal of psychotherapy is to experience as much as possible within the syntactic mode and to broaden the repertoire of the self-system. To the extent that this is achieved, individuals are in a position to become responsible for their ongoing growth through subsequent

interpersonal interactions.

Sullivan's View of Mr. A

Sullivan would see Mr. A as likely arrested in childhood, when his fear of displeasing his mother led him to give up healthy self-esteem strivings for independence in favor of a distant yet dependent position. He uses drugs and psychosis as escapes to maintain some degree of self-esteem. His "good me" consists of his debating and his intelligence. His "bad me" is expressed in his rebelliousness, whereas the "not me" seems to encompass issues of closeness, independence, and constructive engagement with others and with life's tasks. In therapy, the reconnaissance phase is extremely important, identifying the moment-to-moment interactions through which Mr. A's security operations interfere with his attempts at constructive independence. His bemusement and detachment are identified, as is his escape into prototaxic thought through drugs and noncompliance with medications. His acceptance of his mother's management of his trust fund is noted as well. A Sullivanian therapist actively interacts, empathically identifying and confronting Mr. A's ways of avoiding authenticity and constructive interaction. Once identified, their meanings are gently probed, as are the feelings associated with them. The terror of displeasing his mother and its effects on Mr. A's ongoing interactions with the therapist are addressed. Interactions are examined for their consequences, with the assumption that the outcome has bearing on the motivation. Finally, Mr. A is encouraged to try out different ways of relating to the therapist as a prelude to restructuring his outside relationships. He is encouraged to think rationally about his circumstances, to seek a good peer group, develop close friendships, and gradually move away from his near-exclusive focus on his immediate family for all of his interpersonal needs.

ERIC BERNE

Eric Berne (1910–1970) was an American original in both style and substance. He worked in the San Francisco area most of his career, breaking with psychoanalysis in the mid-1950s but never becoming antianalytical as did many of his followers (Fig. 6.3–13). Like many others, he felt the need to develop briefer treatments than were offered at the time.

FIGURE 6.3–13. Eric Berne. (Courtesy of Wide World Photos.)

A group gathered around him that came to be known as the San Francisco Transactional Analysis Seminars. Through weekly discussions of clinical cases and social and political issues, Berne gradually refined his theory. Berne was wry and provocative in his approach to human behavior, contributing much to "pop psychology" in the 1970s and 1980s. Few clinicians now call themselves transactional analysis therapists; nonetheless, Berne's ideas remain useful in grasping hidden agendas in human interactions.

Personality Theory

For Berne, the primary motivator of all human behavior is the need for "***strokes***"—attention, recognition, and response from others. Early survival depends on adequate physical contact, stimulation, and nurturance. This need remains strong but later becomes more symbolic and interpersonal. Children learn rapidly what works within their families

and practice it extensively. This led to one of Berne's more widely quoted observations: "Negative strokes are better than no strokes at all." People evolve ways of interacting with their world to obtain regular strokes in whatever way possible and in whatever way they have been taught to define a stroke (e.g., sympathy in response to chronic depression may provide such gratifying attention that the depression cannot be given up). So great is the need for regular stroking that blatantly destructive actions persist in the face of insight, recognition, and enduring psychological or physical pain.

Like Adler and Sullivan, Berne suggested that hidden social needs motivate human behavior to the extent that, with rare exceptions, there is an interpersonal hidden agenda in all human activity. For Berne, the unit of observation was the transaction, the short-term process of individuals interacting with one another. He spent much energy analyzing transactions to try to discover patients' definitions of strokes and their preferred mechanism for obtaining them. He noted that most people engage in very predictable, stereotyped, repetitive transactions. The content may vary from situation to situation, but the form tends to be quite rigid. He called these transactions ***games***, and his bestseller, *Games People Play*, captured the imagination of the American public in 1964. Some games are harmless, some are socially encouraged; many have destructive elements or at least limit opportunities for more gratifying relationships (intimacy), and some are highly destructive. A common, socially accepted game is cocktail party flirtation, which is ordinarily pleasant and harmless but, depending on the intensity, frequency, and seriousness with which it is played, may cause an inability to experience intimacy, disrupted marriages, or even physical harm.

Berne divided the human psyche into three primary parts: *Child, parent*, and *adult*, with two of those further subdivided. He called these *ego states*. An *ego state* consists of characteristic body language, voice qualities, verbal productions, and affective experience. The child represents the persistence of child-like experience and expression in all people. It is divided into the *natural child*, the ego state in which the spontaneity, joy, and intuitive perceptiveness of young children persists in all adults; the *adapted child*, the part that is compliant and cooperative; and the *rebellious child*, the repository of that part of each person prone to fight authority, challenge accepted wisdom, and struggle for autonomy. The parent is the residue of internalized parental messages and injunctions. It is divided into two parts: The *critical parent* and the *nurturing parent*. The *critical parent* bears some resemblance to Freud's superego, embodying rules, values, instruction, criticism, and restrictions. The *nurturing parent* is the internalization of positive caring experience, the memory of loving interactions. The adult is a purely rational, data-processing element that is objective, calculating, and weighs options and estimates probabilities.

Mental health is the flexible availability of all ego states with no one predominating. Excess critical parent produces guilt and depression, but insufficient critical parent produces sociopathy. Excess nurturing parent produces a narcissistic laziness, whereas insufficient nurturing parent causes an inability to soothe oneself and to maintain self-esteem. Excess adult results in a cold, overly rational person, but insufficient adult leaves individuals unable to balance the various internal forces in their lives. Too much natural child may result in irresponsible behavior, but not enough depletes the ability to experience joy in living. An overabundance of rebellious child results in self-destructive battles with no constructive purpose; however, insufficient rebellious child may result in an overly conformist stance. Excess adapted child prevents appropriate striving toward autonomy; insufficient adapted child prevents participation in group or hierarchical efforts.

Berne placed great emphasis on the child's ability to intuit parental messages and instructions, especially those communicated nonverbally and unconsciously. In this way, the growing child might encode a conscious verbal instruction in the critical parent (e.g., be strong, be perfect, be smart), whereas internalizing a more powerful, unconscious message in the child (don't grow up, don't leave me, don't surpass me). Usually, a model for carrying out the instruction is preserved in the adult (be passive, drink alcohol, run around with women, act crazy). Together, these make a ***script***. Berne emphasized the active role of the child in searching for the messages and in accepting them. This was crucial for him because it emphasized the individual's responsibility in deciding to follow the script, although this might have occurred at a very young age. The entire basis for psychological change lies within the person's capacity, having once accepted the script, to later reject it. However, Berne was impressed by the intensity and persistence with which individuals play out the script throughout their lives. Scripts come in all varieties, ranging from the successful to the utterly self-destructive. Thus, for Berne, the business of human living involved carrying out one's script according to the transactions learned as a child.

Theory of Psychopathology

Given this view of human nature, Berne's understanding of psychopathology is based on the adaptiveness of a person's games and script and the capacity for adaptive use of all ego states. Written at a time when American psychiatry rejected phenomenological diagnosis, Berne's theories are difficult to accept. However, he did describe numerous clinical situations and case histories that are familiar to all psychiatrists, and he analyzed them according to their scripts. He was very impressed with the role of fantasy and fairy tales in children's development and often asked patients what their favorite story was as a child. He then searched for the

hidden identifications and messages within the story to help discern the patient's script. Insofar as he developed a nosology, it lay in the scripts encouraged by culturally sanctioned fairy tales. For example, Cinderella is about a young girl waiting to be rescued by a fairy godmother or a charming prince and unwilling to assert herself against unjust authority. Little Red Riding Hood is about a girl who likes to chat with wolves and gets into trouble trying to please everybody. Men, he found, often identified with the wolf or Prince Charming and were unable to regard women as real people.

A separate nosology of games was also developed and catalogued by Berne. Examples from this nosology include cocktail party flirtation (Rapo I), which concludes with the man and woman having experienced the stroke of feeling attractive and attracted, moving on to other conversations, seeking other people to attract. Rapo II usually ends with a line like, "What kind of girl do you think I am?" and Rapo III ends up with a painful, destructive affair. Another game is "Why Don't You ... Yes, But." In its mild, socially acceptable form, one person presents a problem, successfully enlisting the advice of many others; however, for every proffered solution the person has a ready explanation of why it will not work. The stroke (or payoff) is the attention and the feeling of superiority achieved by defeating others. The second-degree version of this game produces a chronic, helpless depression; the third-degree version may lead to the backward of a state hospital.

There was no rigid mapping of his nosology of games and of scripts, and Berne never claimed that either was exhaustive. They were intended as clinical examples to guide the psychotherapist's thinking in evaluating each patient who presented unique script and transactional issues. Individuals were assessed according to the straightness of their transactions (Were there hidden communication and invitations emanating from different ego states?), the adaptiveness of their scripts, and the predominant ego state and its effects. Thus, Berne might describe a compulsive person as having too much critical parent or a histrionic person as having too much natural child. In the final analysis, however, he saw his catalogues and nosology as poor substitutes for careful, clinical study of individual patients.

Treatment

Having been trained in psychoanalysis, Berne used many psychoanalytic techniques, but he applied them very differently from traditional psychoanalytic methods. Much of his clinical work was done in groups, with the therapist being very active in confronting and interpreting the interpersonal behavior of the patient. He emphasized the initial contract with the patient and the setting of clear, concrete goals. He encouraged therapists to inquire very early on in therapy, "How will we know when our work is finished?" If a clear and specific answer was not forthcoming, he

would inform the patient that therapy had not yet begun. The focus would then be on the patient's difficulty in defining a recognizable end point for the work. Berne's approach was heavily informed by the psychoanalytic concept of resistance, but he applied it in a very different manner. Berne examined any treatment goal defined by patients for evidence that it was a way to continue playing games or pursuing scripts more effectively rather than giving them up entirely.

Like Sullivan, Berne regarded the ongoing interaction between patient and therapist as the key ingredient of psychotherapy. The therapist's job was to interact actively with the patient; recognize the ego states, games, and scripts being enacted; and to counter them using confrontation, interpretation, or various other interpersonal maneuvers that were designed to thwart the enactment and to confront the patient with a choice and an opportunity to relate differently. He emphasized simple direct statements using everyday words to stimulate affective experience and interaction. Commonly, interactions began with his inquiry, "What do you want to work on today?" Observers sometimes described his work as individual therapy within a group context because the other patients, who were silent observers, often responded emotionally to the interaction.

Berne placed great emphasis on the role of individual responsibility for one's life and experience, which he believed had been obscured by the emphasis on psychic determinism. He constantly communicated the opportunity for choosing to continue or discontinue the patterns or habits encouraged as a child. He also showed patients how they invite others to behave in various ways, thereby creating their own interpersonal reality. Health lies in recognizing one's option for deciding which invitations to offer to others and which to accept from others, although these invitations may be communicated unconsciously and nonverbally.

Berne's View of Mr. A

Mr. A is clearly playing a very destructive game and living out a particularly unproductive script. The script seems to contain the messages: "Don't ever leave me" and "act crazy." The message "you must be special" may also exist. The games used might be called "I'm out of control," "Helpless," "Artiste," and "I'm better than you are." A transactional analyst places Mr. A in a group and works with him on these games and scripts. Many such therapists have productively combined Berne's theories with Fritz Perls's empty chair technique. Mr. A might be asked to carry on imaginary dialogues with his mother, professor, his illness, marijuana, and admired writers of fiction, with Mr. A assuming both sides of the dialogue. The transactions he uses are elucidated in this way, and he is invited to recreate the transaction more constructively. His internalization of his mother as an excessively critical parent is identified, and the fantasy dialogues are used to enable more internalization of both his mother's and father's nurturing

qualities, which contain other script messages such as "grow up" or "be successful." Similarly, his excess rebellious child is demonstrated, and efforts to bring it under more adult control are made.

In looking at the formulations and treatment plans each of these theorists might propose for Mr. A, significant, even contradictory, interpretations become apparent. Commonalities exist but are expressed in different words and concepts. Each of these great thinkers has contributed to the overall understanding of the psychodynamics of psychopathology and treatment.

References

Alexander F, French T, Pollack GH. *Psychosomatic Specificity*. Chicago: University of Chicago Press; 1968.

Ansbacher HL, Ansbacher RR, eds. *The Individual Psychology of Alfred Adler: A Systematic Presentation in Selections from His Writings*. New York: Basic Books; 1956.

Bair D. *Jung: A Biography*. New York: Little, Brown; 2003.

Berne E. *Games People Play*. New York: Grove; 1964.

*Berne E. *What Do You Say After You Say Hello? The Psychology of Human Destiny*. New York: Bantam; 1972.

Boeree CG. Karen Horney website. Available at http://www.ship.edu/~cgboeree/horney.html. Accessed October 26, 2007.

Carlson J, Watts RE, Maniacci M. *Adlerian Therapy: Theory and Practice*. Washington, DC: American Psychological Association; 2006.

Conci M. Erich Fromm, a rediscovered legacy. *Int Forum Psychoanal*. 2000;9:141.

Fromm E. *Escape from Freedom*. New York: Avon Books; 1965.

Hirsch I. Review of the interpersonal theory of psychiatry. *J Am Psychoanal Assoc*. 2004;52(1):257.

Horney K. *Neurosis and Human Growth*. New York: Norton; 1950.

International Transactional Analysis Association web site. Available at http://www.itaa-net.org/. Accessed October 26, 2007.

Jacobs TJ. The corrective emotional experience—Its place in current technique. *Psychoanal Inq*. 1990;10:433.

Jung CG. *Two Essays on Analytic Psychology*. Princeton, NJ: Princeton University Press; 1966.

Kramer R. Why did Ferenczi and Rank conclude that Freud had no more emotional intelligence than a pre-oedipal child? In: Barbre C, Ulanov B, Roland A, eds. *Creative Dissent: Psychoanalysis in Evolution*. New York: Praeger; 2003:23.

Lieberman E. *Acts of Will: The Life and Work of Otto Rank*. New York: Free Press; 1985.

Lieberman EJ. Otto Rank web site. Available at http://www.ottorank.com. Accessed October 26, 2007.

Meyer A. *Psychobiology: A Science of Man*. Springfield, IL: Charles C Thomas; 1957.

Mitchell S, Black M. *Freud and Beyond*. New York: HarperCollins; 1996.

*Mulahy P. *Psychoanalysis and Interpersonal Psychiatry: The Contributions of Harry Stack Sullivan*. New York: Science House; 1970.

Mosak HH. Adlerian psychotherapy. In: Corsini RJ, Wedding D, eds. *Current Psychotherapies*. 7th instr. ed. Belmont, CA: Thomson Brooks/Cole; 2005:52.

Paris BJ. *Karen Horney. A Psychoanalyst's Search for Self Understanding*. New Haven, CT: Yale University Press; 1994.

Paul R, Bluma S. *Heresy. Sandor Rado and the Psychoanalytic Movement*. Lanham, MD: Jason Aronson; 1995.

Rado S. *Psychoanalytic Theory of Behavior*. New York: Gone & Sumner; 1962.

Rank O. *The Trauma of Birth*. New York: Harper & Row; 1973.

*Reich W. *Character Analysis*. New York: Fount, Stores & Young; 1949.

Sharaf M. *Fury on Earth: A Biography of Wilhelm Reich*. New York: St. Martin's Press/Marek; 1983.

Stevens A. *On Jung*. London: Routledge; 1990.

Sullivan HS. *The Interpersonal Theory of Psychiatry*. New York: W.W. Norton; 1953.

Tomlinson C. Sandor Rado and Adolf Meyer. A nodal point in American psychiatry and psychoanalysis. *Int J Psychoanal*. 1996;77(5):963.

Williams D. CG Jung web site. Available at http://www.cgjungpage.org/. Accessed October 26, 2007.

▲ 6.4 Approaches Derived from Philosophy and Psychology

PAUL T. COSTA, JR., PH.D., AND ROBERT R. MCCRAE, PH.D.

Four major schools of theories of personality—psychoanalytic, behaviorist, humanistic, and trait—reflect four radically different philosophies of human nature and have very different and consequential implications for the field of psychiatry. Psychoanalytic views of human nature hold that individuals are fundamentally irrational, driven by blind animal instincts, with control and some measure of normality resulting from social structures and forces that operate by inducing guilt and anxiety. Behaviorist accounts posit that the individual's nature is driven not by psychic conflicts but by histories of reward or reinforcement of certain behaviors. Human nature is reactive; individuals are habit-bound creatures of the environment they inhabit. In the third tradition, humanistic approaches depict human nature in a far more positive and perhaps even heroic light. People's capacity for love, creativity, and transcendence are their quintessential features. Maladaptive traits and behaviors, including instances of irrationality and impairment, are attributed to society's thwarting of people's basic potential and goodness.

Contemporary trait theories of personality reflect the increasing recognition that all three of the other schools contain a kernel of truth. Are people driven by their impulses? Some are, some are not. Some are habit-bound, some are heroic; people vary in their emotional, interpersonal, experiential, attitudinal, and motivational tendencies. Trait psychology's emphasis on consistent differences among individuals can be seen as an alternative philosophy of human nature, one better suited to pluralistic and diverse societies.

The first section of this chapter raises philosophic consciousness by hinting at some of the fundamental ideas concerning human nature that may challenge the implicit tenets of modern psychiatry. It is, of course, impossible to do justice to the full scope of philosophic thought in the space of this section. Here, a few examples are selected and their relation to the empirical tradition that is central to modern psychology and psychiatry is especially considered. The remainder of the chapter provides an overview of psychological theories, grouped under the traditional rubrics of behavioral, humanistic, and trait approaches. Because of its importance in contemporary personality research and its application to psychiatric diagnosis and treatment, special attention is given to the trait perspective, especially the Five-Factor Model (FFM) of personality traits and its accompanying Five-Factor Theory of the personality system.

PHILOSOPHY

The recognition of individual differences is probably as old as human culture, but differences in personality were long confounded with differences in status, class, or caste. Individuals from a higher status were assumed to be superior human beings, more sensitive, honorable, and wise. They were endowed with these characteristics by education, by bloodline, or—in the view of Indian philosophy—by moral behavior in a previous life.

Plato

The first major account of personality in Western thought was provided by Plato (circa 428–348 BCE). In *The Republic*, he made extended comparisons between the constitutions of different states and the constitution of the soul. Just as every state must have peasants, artisans, soldiers, and rulers, so too must each individual have appetites (for food, sex, and so on), passions (for honor and advancement), and reason. The relative strength of these three determines character, as well as fitness for a particular place in society. Intelligent and thoughtful people ought to rule, and passionate people should be chosen to defend the state, whereas dull and spiritless individuals, lacking reason and passion, should be given the menial chores of agriculture and industry.

Plato assumed that these psychological characteristics, like physical strength or musical talent, were largely inborn, but he regarded them as a property of the individual, not the individual's social status. Although he assumed that high-status citizens would normally bear children of the greatest potential, he specifically acknowledged that there would be exceptions, and, in his ideal state, children of the lower classes would be promoted and those of the higher classes would be demoted on the basis of their own merit. Here, personality would be the basis of the social order, not vice versa. His most radical extension of this idea was to argue that women should be given social equality and allowed to become soldiers and rulers if they possessed the necessary mental and physical qualifications.

One of the recurring questions in personality theory has been the relative importance of nature versus nurture. Trait psychologists—particularly those interested in the study of temperament—have frequently pointed to innate differences in personality, whereas behaviorists and psychoanalysts have laid great stress on the formative influences of the environment and early childhood experiences. Ready as he was to acknowledge the importance of inborn potential, Plato was also keenly aware of the influence of education. Much of *The Republic* is devoted to his views on the effects of physical exercise, mental instruction, and poetry and music on the development of personality. Present-day concerns about the influences of video games and rap music on children follow in this tradition.

Like most philosophers, Plato was more concerned with understanding

virtue and vice than psychopathology, but, because he considered vice to be the result of weak or corrupted nature rather than free but evil choice, his discussions of flawed character can be viewed as early descriptions of what might now be viewed as personality disorders. Just as there are better and worse forms of government for states, so are there also better and worse configurations of reason, passion, and appetite. Plato described five types, corresponding to five forms of government. The ideal of mental health, corresponding to government by wise rulers, is one in which reason holds in check passions and appetites. In the arrogant and ambitious type, there is an excess of passion or pride; in the avaricious type, there is an excess of appetite. However, in both of these types, reason still retains some authority; for example, avaricious individuals can control most of their appetites to indulge their desire for money. In the fourth, self-indulgent type, appetites are undisciplined, and, in the debauched, fifth type (corresponding to a government by despot), reason is completely disregarded, and a clearly psychopathological state is reached: "Thus, when nature or habit or both have combined the traits of drunkenness, lust, and lunacy, then you have the perfect specimen of the despotic man."

Aristotle

Plato's vivid but rough typology was succeeded by Aristotle's (384–322 BCE) detailed analysis of human character in the *Nicomachean Ethics*. Courage, temperance, generosity, pride, ambition, irascibility, friendliness, boastfulness, and shame were all defined and distinguished. Aristotle attributed pathological variations on these traits to innate defects or to disease processes: "Among all the excesses of foolishness, cowardice, intemperance, and irritability some are bestial, some diseased. If, for example, someone's natural character makes him afraid of everything, even the noise of a mouse, he is a coward with a bestial sort of cowardice." He considered variation within the normal range to be the result of training and habit and thus to be subject to praise or blame.

Aristotle's basic moral precept is the *golden mean*: He argued that extreme standing on either end of a trait dimension should be avoided. Thus, stinginess and extravagance are vices, whereas generosity is a virtue; similarly, vanity and humility represent excessive or insufficient self-esteem. This conception continues to influence some modern notions of psychopathology, in which marked deviation from the norm in either direction is considered pathological. Individuals excessively concerned with social attention may be regarded as having a histrionic personality disorder; those insufficiently concerned with social attachments may have a schizoid disorder.

Aristotle carried conceptual analysis to a level that has seldom been surpassed, distinguishing, for example, between the superficially similar

qualities of temperance and self-control: Individuals are temperate if they have healthy and moderate impulses that require little control, whereas individuals have self-control only if they have immoderate appetites that are nevertheless held in check. Such considerations allowed him to form rational taxonomies of traits that anticipate the empirical taxonomies proposed by 20th-century factor analysts.

Shankara and Eastern Philosophy

Indian philosophy is a vast body of work that combines early science and psychology with Hindu theology and its offshoots (including Buddhism). Its foremost thinker (from the perspective of mainstream Hinduism) was Shankara (788–820 CE). In a brief life of 32 years, he restored the authority of the ancient Hindu scriptures, the Veda, after centuries of criticism by heterodox schools, and he articulated a worldview that has influenced Indian thought ever since. Shankara argued that the basis of existence was Brahman (the Absolute, Emerson's Oversoul), which pervades everything but is itself indivisible. The physical universe is not, for Shankara, an illusion, but a manifestation of Brahman. What is an illusion is the idea that the individual is part of the material world, when in fact the Self is nothing other than Brahman. Through study and virtuous actions it is possible to come to understand this fact, leading to a state of enlightenment in which the individual becomes blissfully detached from life, while continuing to participate in it.

Although the philosophical rationales vary immensely, much Eastern thought shares the core ideas and ideals of Shankara's system. In particular, Eastern thought emphasizes renunciation and asceticism, not as a way to mortify the flesh, but as a way to liberate the soul. Themes of self-transcendence are also seen in the reverence for all life that is common to Hinduism, Buddhism, and Jainism (and that is increasingly seen in Western concerns with the environment).

Eastern philosophy was much admired by Carl Jung, and it had some influence on psychiatric thinking during the 1950s and 1960s. However, despite recent calls for multiculturalism in psychology, there appears to be relatively little interest in the topic today. Yet Eastern thought provides an entirely different perspective on what the goal of psychiatric treatment ought to be: Not palliation of distress or more effective functioning in social groups, but liberation from the illusions that create distress and dysfunction. This goal and the large body of techniques for achieving it developed over the centuries surely merit more serious attention.

Immanuel Kant and Arthur Schopenhauer

The last great period of the Western philosophical tradition begun by Plato and Aristotle was inaugurated at the end of the 18th century by Immanuel

Kant (1724–1804), whose critical philosophy forms a transition between purely rational approaches to human nature and the empirical sciences that followed. One of Kant's last works, *Anthropology from a Pragmatic Point of View*, considered natural and moral variations in character and reintroduced the Roman physician Galen's taxonomy of choleric, phlegmatic, sanguine, and melancholic types to modern psychology.

Sciences are usually distinguished from philosophy by their reliance on empirical data, but it should not be imagined that philosophers made no use of experience. On the contrary, like many of the clinicians who offered psychological theories of personality, philosophers based their ideas heavily on their observations of human nature. A striking instance of this is provided by Arthur Schopenhauer (1788–1860), a follower of Kant, whose dark view of the world as a place of purposeless striving had an extraordinary influence on early personality theorists, including Freud and Jung.

One of the central beliefs of Western thought has been that human happiness or misery is the result of external conditions—what is now called *quality of life*. However, Schopenhauer's acute observations, reported in *The World as Will and Representation*, led him to propose "the paradoxical but not absurd hypothesis that in every individual the measure of pain essential to him has been determined once and for all by his nature.... His suffering and well-being would not be determined at all from without, but only by ... what is called his temperament."

In support of this view, he noted that wealth and power do not make people happy, "for we come across at least as many cheerful faces among the poor as among the rich." He also argued that the effects of great misfortunes or successes are short lived and that, for the most part, evaluations of the external causes of state of mind are illusory attributions: "We often see our pain result only from a definite external [cause] and ... believe that, if only this were removed, the greatest contentment would necessarily ensue. But this is a delusion .... [The pain] would appear in the form of a hundred little annoyances and worries over things that we now entirely overlook."

Recent scientific research on psychological well-being has confirmed this account in every detail: Well-being is chiefly a function of enduring personality dispositions; wealth, social class, and other markers of the objective quality of life are virtually unrelated to subjective happiness; and processes of adaptation quickly return individuals to their own characteristic baseline of happiness after favorable or unfavorable life events. Schopenhauer's observations are also consistent with recent evidence on the heritability and lifelong stability of many mental disorders.

Yet, reason and insight are not enough. On the basis of his own experience and the examples of history, Schopenhauer also concluded that "man inherits his moral nature, his character, his inclinations, his heart

from the father, but the degree, quality, and tendency of his intelligence from the mother"—a conclusion not supported by the findings of modern behavior genetics. Psychology broke from philosophy over precisely this need to seek empirical verification of hypotheses, but it carried with it concepts and insights accumulated over two millennia of profound thought about human nature.

Jacques Derrida and Postmodernism

Throughout much of the 20th century, philosophers tried to support empiricism by examining the conceptual foundations of the scientific method and inductive inference. Concepts such as operationalism, theory falsification, and construct validity are products of the philosophy of science that have been widely adopted in scientific practice. The most influential philosophical movement of recent years, however, has often been seen as antiscientific. Deconstructionism and its intellectual descendants have provided a radical critique of conventional scientific thinking that has had a marked impact, for good or ill, on the humanities and social sciences.

Jacques Derrida (1930–2004) was a French philosopher who is usually identified as the founder of deconstructionism. Through analyses of literature, history, and psychoanalysis, he questioned the basic tenets of Western thought. He drew attention away from the ostensible subject of any writing, pointing to the writing itself as the object of investigation. Following this lead, social constructionists argued that there is no objective reality for science to study; instead, reality is constructed as a product of language and culture.

Derrida has had an enormous impact on the humanities, which is seen in the scholarly journals that abound with references to *discourse*, *text*, and *narrative*. In combination with other influences, such as feminism, social constructionism, and multiculturalism, deconstructionism has led to postmodernism, a stance that is critical of conventional empirical science on several accounts. Postmodernists argue that science is a set of cultural conventions revered in the West but is ultimately no more valid than any other belief system. The evidence that a Darwinian offers in favor of evolution is no more convincing than the evidence that a Fundamentalist offers in favor of creationism. Both are discourses that have their own justification. Postmodernists are also critical of the claim that science is value-neutral. They argue that the enterprise of science is inextricably bound to value judgments, and they have pointed out numerous ways in which science has contributed to the oppression of women, minorities, and members of non-Western cultures.

Many psychoanalysts, humanistic psychologists, and anthropologists have embraced postmodernism. Within psychiatry, one of the pioneers of postmodern thinking is Thomas Szasz, who argued that mental illnesses

are not objective medical conditions, but questionable value judgments. Kenneth Gergen has urged psychologists to adopt a postmodern perspective to liberate their science from orthodox methods and theories and to make psychology more relevant outside Western culture. Mainstream psychologists, like most psychiatrists, tend to view postmodernism skeptically; indeed, many reject it outright.

However, the criticisms and alternatives offered by postmodernists can be useful to empirical scientists, even if the radical premises are not accepted. Some of the qualitative methods they advocate, such as content analysis, can provide a valuable supplement to quantitative methods. It is useful to be reminded that the constructs of science, such as the categories of the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), are not direct representations of reality, but rather imperfect maps always in need of revision. Therapists are well advised to be sensitive to the cultural background of their patients, and, surely, all scientists need to be mindful of the social consequences of their work. This is particularly true in psychiatry, which has so much power over the lives and well-being of its patients.

BEHAVIORAL AND SOCIAL LEARNING APPROACHES

Theories of Personality

Behaviorism as a school of psychology grew up in reaction to the prevailing mentalistic model, in which introspection was used to determine the contents and operations of consciousness. John B. Watson (1878–1958) proposed that scientific psychology should confine itself to an examination of observable behavior and should explain all human conduct in terms of stimuli and learned responses. Ivan Petrovich Pavlov's (1849–1936) experiments with conditioned responses offered hope that such a science could be successful, and theorists such as Clark L. Hull (1884–1952) provided elaborate mathematical models of learning.

Radical Behaviorism. The most influential behaviorist and, perhaps, the most influential psychologist of the century was B. F. Skinner (1904–1990). Skinner's basic concept was *operant conditioning*, in which behaviors are viewed as a function of the organism's history of reinforcement. The observation that animals can be taught tricks by giving them rewards and punishments is nothing new. However, behaviorists, such as Skinner, refined and systematized this idea, using elegant experimental designs to tease apart the effects of the amount and schedule of reinforcements, the use of reinforcers and punishers, and the difficulty of the discriminations required. Behaviors could be shaped, maintained, or eliminated by the judicious use of these principles.

Skinner was a radical behaviorist, a purist who denied not only the

scientific value, but also even the existence of mind. Furthermore, he avoided any neurophysiological or psychophysiological theorizing, preferring to study the *empty organism*. Individual differences were ignored in understanding basic phenomena and were explained in individuals as the result of different histories of reinforcement. Even differences between species were neglected: Skinner believed that the pigeon provided an adequate model for the study of learning in all organisms, and he and his followers were, in fact, able to replicate many of their animal findings by using human subjects. Much of Skinner's success can be attributed to his single-minded pursuit of a highly circumscribed set of variables.

Skinner's view of personality was, predictably, a reductionistic one. As he stated in *About Behaviorism*, "a self or personality is at best a repertoire of behavior imparted by an organized set of contingencies. The behavior a young person acquires in the bosom of his family composes one self; the behavior he acquires in, say, the armed services composes another." This position is rejected by humanistic psychologists, who attribute more choice and control to the individual, and by trait psychologists, who see consistencies of behavior that appear to transcend the consistencies of the reinforcing environment. Many personality psychologists have argued that controlled laboratory experimentation is a poor basis for theories of personality because individuals play a large role in selecting and shaping their own environments. Skinner's radical behaviorism was rejected or modified by many later learning theorists who acknowledged the power of conditioning but also recognized differences among species and among individuals within a species.

Social Learning Theory. One of the most distinctive features of human organisms is their use of speech, which makes possible elaborate thinking and planning and complex social interactions. In recent decades, learning theorists have increasingly emphasized social and cognitive processes. Among the most important have been Julian Rotter (1916–2014) and Albert Bandura (1925 to present), both of whom have offered versions of social learning theory.

Rotter's theory proposes that human behavior is guided not only by the actual history of reinforcement, but also by plans, goals, and expectations of success. Individuals perform a behavior if they believe it is likely to lead to a valued goal, based on their past experiences in general and in similar situations. Individuals with a history of success are likely to have a generalized expectancy that they can control their lives; they are described as having an *internal locus of control*. At the opposite extreme are those whose prior efforts have been generally unsuccessful; they come to believe that rewards and punishments are a matter of luck or the arbitrary decisions of powerful others. Such individuals are said to have an *external*

locus of control. Locus of control has been one of the most popular variables in personality research, used in numerous studies that generally are consistent with Rotter's theory.

Bandura's version of social learning theory also acknowledges the importance of internal cognitive processes. Individuals learn not only on the basis of their own experience, but also through vicarious reinforcement from observation of others. Bandura's demonstration of modeling effects in experiments conducted in the 1960s gave scientific legitimacy to the social learning perspective.

Rotter's and Bandura's theories are general theories of behavior, not specific theories of personality. However, for them, personality is something more than a collection of learned behaviors. The total pattern of experience leads to a generalized expectation of reinforcement or to a general sense of self-efficacy that can be considered the central individual difference variable. People are to be characterized primarily on the basis of their beliefs in their own ability to control their lives, because these beliefs powerfully determine the effort they make to adapt to their surroundings.

Social-cognitive approaches to personality are among the most influential for current research in personality. These approaches focus on the individual's understanding of him- or herself and how these self-appraisals shape goals, plans, and behaviors. Because of their origins in social learning theory, they tend to emphasize the role of the environment, pointing out that an individual's sense of self varies from setting to setting. Because of their ties to social theory, they usually explain personality in terms of the effects of social interactions. For example, Hazel Markus (1949 to present) has suggested that individuals have a number of *possible selves*: Conceptions of what one is or could be, which result from the messages significant others provide. For such theorists, concern has moved beyond the social learning of specific behaviors to the learning of entire identities.

Theories of Psychopathology

Psychoanalytic theories of personality grew out of attempts to understand psychopathology, and there are thus intimate connections between the two. By contrast, behavioral approaches have focused on the general principles by which behavior is acquired and maintained, and psychopathology, where it is considered at all, is usually treated as an area of application. Learning theories have had much more influence on methods of psychotherapy than on theories of psychopathology itself.

Behavioral approaches might suggest two different classes of explanations for psychopathology: Psychopathology might be related to the mechanisms of learning themselves, or it might be considered the result of learned behaviors that are maladaptive or socially unacceptable.

In the 1950s, Hans Eysenck (1916–1997), a psychologist who has figured prominently in both learning and trait schools of personality,

proposed that individual differences in the dimension of introversion versus extraversion determined the ease with which individuals could acquire conditioned responses, which in turn determined the form of psychopathology to which they were prone. Extremely introverted individuals, he proposed, were easily conditioned and thus acquired many inhibitions. They were predisposed to the development of depressive, anxious, and obsessive-compulsive disorders. By contrast, extreme extraverts were considered to be resistant to conditioning and were likely to develop hysterical and psychopathic disorders. (In later versions of Eysenck's theory, psychopathic disorders were grouped with psychotic disorders and linked to a different dimension of personality, psychoticism.)

Most behaviorists have viewed the laws of learning as universal processes and considered psychopathology to be the result of normal learning processes. In the 1920s, Irena Shenger-Krestovnika taught a dog to discriminate between a circle and an ellipse as a cue for food. When the ellipse was made increasingly circular, the dog's ability to discriminate between them was taxed, and the dog began to struggle, squeal, and bite. Pavlov dubbed this an *experimental neurosis* and proposed that human neuroses might have parallel causes.

Probably the most famous attempt to explain psychopathology in learning theory terms was provided by John Dollard (1900–1980) and Neal Miller (1909–2002). Dollard, an anthropologist, and Miller, an experimental psychologist, shared an interest in psychoanalysis. Their goal was to translate psychoanalytic concepts into the more testable terminology of learning theory. Consider, for example, the central psychoanalytic notion of *repression*. Parents might punish sexual behaviors in the child, and the child might learn to associate the behaviors with pain. By stimulus generalization, even the thought of the behaviors would elicit anxiety, and cognitive processes that blocked those thoughts would lessen anxiety and thus would be reinforced. Eventually, the thoughts would be effectively barred from consciousness.

Many behaviorists who did not share Dollard and Miller's enthusiasm for psychoanalysis followed their lead in attempting to explain psychopathology in terms of principles of learning. Phobias, in particular, were easily explained as conditioned responses reinforced by avoidant behavior. Similarly, compulsions could be understood as a kind of self-reinforcing behavior: Each time the compulsive act is performed, the anxiety associated with not performing the act is reduced, increasing the probability that the behavior will be repeated.

Social learning theorists have also noted the self-perpetuating nature of some maladaptive behavior. Individuals who lack a strong sense of self-efficacy in social situations may avoid them. As a consequence, they fail to learn the social skills that would enhance their self-confidence. Self-defeating behaviors, which may appear irrational from an outside

perspective, are often understandable in terms of the dynamics of learning.

Application of Theory to Therapy

Behaviorists have a rather rudimentary view of personality, seeing it as an assemblage of learned behaviors. They also tend to see psychopathology in superficial terms. Psychological maladjustment is considered to be the result of learned behaviors, which are called *symptoms*, but there is no underlying disorder of which they are symptomatic. Curing the symptoms cures the disorder. At worst, this position is naive and simplistic, equating the patient's presenting problem with the real source of difficulty. At best, however, it focuses attention on a specific problem that can be concretely addressed.

A large number of techniques for behavior modification have been used with considerable success in treating symptoms of psychopathology. Joseph Wolpe (1915–1997) developed *systematic desensitization* as a treatment for phobias. Patients were instructed to relax and were then presented with increasingly vivid cues of the phobic object. Eventually, they were able to face the object itself without anxiety. A more dramatic technique is *implosive therapy*, in which the individual is confronted directly with the feared object (e.g., a room full of snakes) without an opportunity to escape. Because the object itself is harmless, and because avoidant behavior cannot be performed and is therefore not reinforced, the phobic reaction is swiftly extinguished.

Therapeutic interventions may be based on eliminating the reinforcements that sustain behavior, punishing unwanted behaviors, modeling or shaping more desirable behaviors, and so on. Any variable known to affect the acquisition or extinction of behaviors may provide an opportunity for behavior change, and behavioral techniques have been applied to physiological responses, as well as voluntary behaviors, through techniques of biofeedback. Behavior therapies have been used extensively in treating phobias, controlling addictive behavior, reducing the self-destructive behavior of autistic children, and improving classroom discipline—for the behaviorist, the distinction between psychopathology and bad behavior is generally unimportant. Behavior therapies are most effective when the problem can be clearly traced to a particular set of behaviors or conditioned responses; they are much less effective in dealing with vague complaints of confusion and distress, although these are frequently the problems the patient presents to the clinician.

HUMANISTIC APPROACHES

Theories of Personality

Psychoanalysis and behaviorism are mechanistic theories that trace human behavior and experience to the gratification of instinctual impulses or to

the acquisition of learned responses. Many of the most influential personality theorists have defined themselves in terms of their opposition to these two approaches. Although they vary widely in terms of the explanations they offer of personality, humanistic approaches share a positive evaluation of human nature and emphasize its unique and distinctively human aspects. Personality produces and reflects organization, rationality, consistency, future orientation, planfulness, self-expression, cognitive complexity, and adaptability. Most humanistic theorists prize human reason and freedom of will, the capacity for growth and change, the need for love, and self-transcendence. The positive psychology movement promoted by Martin Seligman (1942 to present) falls within this tradition.

Social learning theories might be seen as a humanized form of behaviorism because they recognize the role of complex symbolization and language in human learning. In a similar way, many now-classic humanistic theories were intended as modifications of psychoanalysis. Indeed, the first major psychoanalytic revisionist was Carl G. Jung (1875–1961), who argued that human beings had spiritual, as well as sexual, needs. Henry Murray (1893–1988) also made major modifications to psychoanalysis in his view of personality. For example, he credited the mature ego with much more autonomy than Freud granted it, and he argued that the individual's sense of morality was not fixed by the superego instilled in childhood but could continue to develop into more rational and altruistic forms. Erich Fromm (1900–1980) and Karen Horney (1885–1952) minimized the instinctual origin of personality development and suggested that culture played a large role in shaping the individual. Erik Erikson (1902–1994) proposed stages of psychosocial development to parallel Freud's psychosexual development, emphasizing such distinctly human characteristics as identity, intimacy, and generativity. He also theorized that personality development continued throughout the life span, giving encouragement to research on aging and personality.

Gordon Allport. Although he is usually classified as a trait psychologist, in his general orientation to theorizing, Gordon Allport (1897–1967) was clearly a humanist. For him, man's behavior is proactive, reflecting internal, self-initiating characteristics more than situational forces. In his view, human personality possesses psychological coherence and momentary (cross-sectional) and long-term (longitudinal) organization. Allport considered personality functioning to be characteristically rational, organized, and influenced by such conscious characteristics as long-range goals, plans of action, and philosophies of life.

Perhaps the most salient and controversial feature of his approach was the extreme emphasis that he put on the uniqueness of the individual personality. Allport viewed the major task of *personology* (or personality

psychology) as the understanding and prediction of the individual case. To grasp the real personality, personal dispositions must be assessed, and this requires intensive study of an individual's past, present, and anticipated future functioning through the use of such techniques as the case history and content analysis of personal documents. In *Becoming: Basic Considerations for a Psychology of Personality*, Allport championed the view that concepts and laws must be developed to fit the individual case, creating the terms *idiographic* and *morphogenic* to symbolize his conviction that "each person is an idiom unto himself, an apparent violation of the syntax of the species."

Allport was deeply concerned with identifying personality functions, which he discussed under the concept of the *proprium*, the superordinate concept in his system. Propriate functioning not only organized and integrated actions and experience, but also provided the impetus to psychological growth. Allport described the functions of the proprium as *sense of body*, *self-identity*, *self-esteem*, *self-extension*, *rational coping*, *self-image*, and *propriate striving*. These propriate functions were vital to personality, and, although they were ongoing, they were by no means considered unchanging. Allport theorized that the propriate functions are modified throughout life, predominantly in the direction of greater differentiation and integration, or growth. The development of selfhood, away from the undifferentiated, opportunistic functioning of infancy and early childhood toward propriate functioning and striving, is part of human nature for Allport. The person guides or directs his or her life by attempting to fulfill his or her sense of self or proprium. Development continues into adulthood, with increasing signs of maturity and personal lifestyle.

Abraham Maslow. Abraham Maslow (1908–1970) interpreted personality in motivational terms. The individual's whole life—his or her perceptions, values, strivings, and goals—is focused on the satisfaction of a set of needs, and the needs themselves are arranged in a universal hierarchy. Maslow's needs serve an organizing and integrating role in life and are not to be understood as a simple and invariant set of responses to environmental pressures. They organize and create action possibilities and external reality.

At the lowest level of the motivational hierarchy are physiological needs for food, water, sex, and sleep. The second level consists of safety needs, needs for protection and security. The third level consists of needs for love and belongingness, and the fourth level consists of needs for self-respect and esteem from others. Above these lie the higher needs—for beauty, truth, justice, and *self-actualization*, the development of one's full potential as a unique human being.

Individuals live at the lowest level of motivation that is problematic for them. That is, if needs for food and shelter are not routinely met, they

become the overriding concern in life: The hungry individual risks danger and social ostracism to find food. Those who have always been well fed, however, learn to take the satisfaction of physiological needs for granted, and their attention is dominated by higher needs. Instead of examining specific behaviors and their reinforcements, Maslow's theory concerns the long-term satisfaction of broad classes of needs and thus gives a much broader depiction of the individual.

Maslow devoted much of his writing to a characterization of higher needs, including self-actualization, a drive to fulfill one's unique potential. He believed that personality psychology had become obsessed with psychopathology, and that a corrective emphasis on positive mental health was needed. His biographical studies of such exemplary people as Eleanor Roosevelt and Abraham Lincoln suggested a number of distinctive characteristics of self-actualizers, including accurate perception of reality, creativity, a need for privacy, and the frequent experience of mystical or *peak* experiences. As an exception to his general theory of motivation, he also noted that such individuals often skip the lower levels and proceed directly to self-actualization. The most creative artists and musicians never seemed to care about poverty or lack of social acceptance.

George Kelly. One of the most unconventional theories of personality was offered by George Kelly (1905–1967). It is, in some respects, a purely cognitive approach, but one with few ties to traditional learning theory. Instead of seeing them as organisms that are conditioned by their environment, Kelly argues that human beings should be seen as scientists trying to make sense of their world. In *The Psychology of Personal Constructs*, he states the fundamental postulate of his theory: "A person's processes are psychologically channelized by the ways in which he anticipates events."

The basic unit for understanding personality is the *personal construct*, a schema for classifying and interpreting experiences. For example, an individual might construe other people in terms of the contrast *strong* versus *weak*. Each new acquaintance would be categorized as a strong or weak person, and subsequent interactions with this person would be guided by the original construal. In the course of experience, it would probably be necessary to reclassify some individuals initially thought to be strong as weak, and vice versa; more importantly, some people might act in ways that were neither strong nor weak, leading the individual to develop new constructs (say, *friendly* vs. *hostile*) that were more useful in predicting other people's behavior.

This rather abstract and bloodless theory is made relevant to psychopathology by Kelly's ingenious reconstruals of some basic emotional reactions. He defines *anxiety* as the awareness that one's construct system is inadequate for construing important events. *Guilt* is the recognition that

one's behavior is inconsistent with the ways in which one construes oneself. *Hostility* is viewed as the attempt to force experience to fit one's existing constructs. Such definitions are remote from common sense and clinical notions of anxiety, guilt, and hostility, but, precisely for that reason, they offer the prospect of novel ways of treating them.

Carl Rogers. Carl Rogers (1902–1987) is probably the most influential humanistic personality theorist. He articulated a formal theory of personality; pioneered a major school of therapy, *client-centered therapy*; and encouraged rigorous research on his theory and therapy. Rogers held that all organisms tend toward their own actualization—that mental health and personal growth are the natural condition of humankind. Psychopathology is a defensive distortion of this actualization process, and psychotherapy consists of creating conditions in which defense is unnecessary. Given these conditions, patients (or *clients*, as Rogers called them) essentially cure themselves.

Under ideal conditions, people's needs, desires, and goals emerge naturally as part of self-actualization and are recognized as part of the self; individuals are fully open to experience. In real life, however, one person's needs and desires often conflict with others' needs and desires; in particular, children find themselves in conflict with their parents, who withhold love when the child (from their perspective) misbehaves. Because love is so essential, children internalize these *conditions of worth* and believe that they are good and worthwhile individuals only when their self is consistent with the ideals imposed by significant others. To maintain their sense of worth, they may distort their experience; this leads to anxiety and self-defeating behavior.

In some respects, Rogers's theory is much like Freud's: Both see psychopathology as the result of defensive distortions and see the ideal state as one in which individuals can accept conflicts and deal with them rationally. Rogers is a humanistic theorist because he assumes that human nature is essentially good and that defenses are ultimately unnecessary. For Freud, the impulses of the id are eternally primitive and selfish, and their full actualization would be socially catastrophic.

Viktor Frankl. An Austrian neurologist and philosopher, Viktor Frankl's (1905–1997) distinctive view of human nature and psychopathology was profoundly shaped by his experience in Nazi concentration camps. There he came to the conclusion that even the most appalling circumstances could be endured if one found a way of making them meaningful. He described his experience in *Man's Search for Meaning*, a book that has been read by millions around the world.

Frankl was both a humanist and an existentialist. He believed that human beings shared with other animals somatic and psychological

dimensions, but that humans alone also had a spiritual dimension that confers both freedom and responsibility. People find meaning in their lives through creative and productive work, through an appreciation of the world and others, and by freely adopting positive attitudes even in the face of suffering. Those who fail to find meaning face alienation, despair, and existential neuroses. Traditional societies provided a framework of meaning in religion and shared cultural values; in modern society, people must find their own sources of meaning, and Frankl attributed many social problems, such as drug abuse and suicide, to their failures to do so.

Because of the spiritual dimension, human beings show self-transcendence and self-distancing. The former refers to the capacity to put other values (e.g., the well-being of a loved one) above self-interest. The latter is the ability to take an external perspective, as seen in a sense of humor. These capacities form the basis for therapeutic interventions in Frankl's version of psychotherapy known as logotherapy.

Dan McAdams. In the past three decades, a number of scholars in the humanities and social sciences have turned their attention to the *life narrative* as a focus of research. These writers argue that the consciousness of self that distinguishes human beings from other animals is not a static list of personal characteristics; it takes the form of a story. In telling their life stories, people explain themselves in the context of their history and their significant relationships. Life narratives are not objective life histories; they are subjective interpretations that give personal meaning to past, present, and future events.

The personality psychologist who is most closely associated with this perspective is Dan McAdams (1954 to present). McAdams postulated that personality could be understood on three levels: Level 1 consists of *traits*, abstract and enduring tendencies seen in general styles of action and experience. Level 2 is defined by *personal concerns*, the goals, plans, and strategies that preoccupy the individual at a certain time and place; later McAdams adopted the term *characteristic adaptations* to describe a broader conception of this level. Level 3 is the *life narrative*, the story the person tells to him- or herself and to others in an effort to give a sense of unity and purpose to life. A common theme in America is the "redemptive self," in which the protagonist faces obstacles but eventually triumphs. Such a narrative appears to inspire a productive and caring life.

Psychiatrists and psychotherapists have, of course, been listening to patients' life stories for many decades, usually seeking clues to the origins of current problems. McAdams's approach is different; he wishes to understand the narratives as stories. Stories can be analyzed in terms of such features as narrative tone (e.g., tragic, comic, and ironic), imagery, theme (what goals the characters pursue), ideological setting, and nuclear episodes (crises and turning points). From this perspective, key life events

are not construed as causes of development but as symbols of identity. A patient's vivid recollection of a childhood humiliation may be important chiefly because it conveys the patient's sense of being a victim to whom fate has always been, and always will be, unfair.

Much of the scholarship on life narratives has been in a humanistic tradition that eschewed empirical rigor: A narrative reduced to a set of numbers is no longer a narrative. McAdams and his colleagues, however, have conducted psychometric studies on such basic issues as the interrater reliability and temporal stability of life story variables. These studies form the basis for a scientific evaluation of narrative perspectives on personality.

Theories of Psychopathology

In general, humanistic theories of personality stress positive aspects of human nature and discuss maladjustment in terms of failures of and blocks to the full growth and development of the individual. It is of some interest to note that Rogers and Kelly formulated their theories in the context of counseling students—individuals who presumably had relatively minor maladjustments and considerable personality strengths. The applicability of humanistic theories to patients experiencing schizophrenia or dementia is certainly questionable, but the theories have had a profound effect on routine clinical practice, in which clearly diagnosable psychiatric disorders seldom account for all of the patient's problems in living.

Humanistic psychologists differ tremendously in their views on the origins of maladjustment. Rogers pointed to internalized conditions of worth acquired chiefly during childhood. Fromm, who was influenced by Marx, blamed society as a whole for instilling nonproductive orientations in individuals, such as *hoarding* or *marketing* orientations. Kelly said little about the origins of maladjustment but thought its essence was an ineffective construct system too rigid to be corrected by experience. Frankl believed some mental disorders were somatic or purely psychological in origin, but he also pointed to spiritual causes for the class of existential disorders.

Application of Theory to Therapy

Some humanistic theories of personality—for example, Allport's—have had little impact on psychotherapy; others, such as Rogers's, have been tremendously influential. It is helpful to recall that, for decades, the dominant form of therapy was psychoanalysis, a process that might require years and in which treatment was focused on dreams, childhood memories, and the ongoing relation with the therapist (the transference) rather than on the immediate problems of the patient. Many of the standard techniques of contemporary counseling, clinical psychology, and psychiatry rest on a very different set of assumptions about human nature made scientifically

respectable by the work of humanistic psychologists.

Brief psychotherapies often consist of opportunities for individuals to express their feelings and to rethink their problems in a supportive atmosphere. The therapist may provide advice and guidance or at least may offer new ways in which patients can think about their problems. (Many psychiatrists use this general approach as an adjunct to medication, even in the treatment of serious mental disorders.) This process implicitly assumes that individuals, even those requiring psychotherapy, are basically rational and able, with some help, to solve their own problems; it also assumes that, given the right conditions, they move toward mental health: Patients are seen as scientists and self-actualizers.

Noting that Freud's "psychoanalytic cure" was based on a conscious understanding of one's life experiences, McAdams has argued that having a satisfying life narrative is a requirement for complete mental health. A good life story has coherence, credibility, a capacity for growth and change, and a generative orientation toward the world. One goal of psychotherapy, then, may be to help the patient rewrite his or her life narrative along these lines. Some (e.g., Sayer and colleagues) have claimed that writing about traumatic experiences has beneficial effects on mental and physical health, perhaps because the exercise of writing allows a rethinking of one's life narrative.

Modifying a patient's life story is a modest but, perhaps, realistic goal. Psychotherapists cannot easily alter patients' personality traits nor can they undo traumatic events of the past. They may, however, be able to help patients make sense of their dispositions and their life histories. Human tragedy can bring great suffering, but, in the hands of an artist, it can also be the source of great beauty. This was one of the major principles of Frankl's logotherapy, in which patients were invited to examine their attitudes and to reconstrue their suffering in ways that made its value apparent. A terminally ill women who despaired that she would have to leave behind the children that she loved was asked if she would rather have been childless; that question allowed her to appreciate her life and legacy.

The most celebrated technique derived from logotherapy is *paradoxical intention*, in which patients are invited to use their capacity for self-distancing to will what they fear or yield to their compulsions. A patient who stutters may be instructed to stutter as violently as possible; this reduces the pressure of trying not to stutter and that in turn makes it possible to speak clearly. A number of behavioral techniques, such as implosive therapy, work on a similar principle of blocking a vicious circle of reinforcing undesirable behavior, but it is often difficult to convince patients to hazard the experiment. Frankl makes use of the characteristically human sense of humor to frame the paradoxical command in ways that motivate the patient.

The humanistic emphasis on freedom and responsibility has often

clashed with the psychiatric tradition of regarding mental disorders as diseases. Labeling individuals as *schizophrenics* or *phobics* is held by some to be dehumanizing, and critics such as Thomas Szasz have argued that mental disorders are social and ethical judgments, not matters of medical fact. There is, of course, abundant evidence that some mental disorders have a biological basis, but the criticisms of humanistic psychologists make the point that the disorder occurs in a human being who, in many respects, may be like any other person. There may be normal and abnormal behaviors, experiences, or relationships, but there are not normal and abnormal people.

TRAIT AND FACTOR MODELS

Individual differences are peripheral concerns in many social learning and humanistic theories of personality; they are the central focus of trait theories. The study of variations in human character and temperament goes back at least to Theophrastus (371–287 BCE), a Greek, whose *Characters* depicted 30 different types. The morose type, for example, he described as follows:

> A malignant temper sometimes vents itself chiefly in ferocity of language. The man whose tongue is thus at war with all the world, cannot reply to the simplest inquiry except by some such rejoinder as—"Trouble not me with your questions": Nor will he return a civil salutation.... He has no pardon for those who may unwittingly shove or jostle him, or tread upon his toe.... He will neither wait for, nor stay with anyone long: Nor will he sing, or recite verses, or dance in company. It is a man of this spirit who dares to live without offering supplications to heaven.

Theories of Personality

The scientific study of individual differences in personality can be traced to Sir Francis Galton (1822–1911) in England, who laid the foundations of psychometrics, and to Gerard Heymans (1857–1930) in the Netherlands, who undertook the first large-scale study of rated personality traits. The first major trait theorist in the United States was Gordon Allport, whose 1937 volume, *Personality: A Psychological Interpretation*, spelled out the basic issues in trait psychology. He defined a *trait* as "a neuropsychic structure having the capacity to render many stimuli functionally equivalent, and to initiate and guide equivalent (meaningfully consistent) forms of adaptive and expressive behavior." In this view, something in the brain of the morose man makes him see even simple questions or greetings as a personal affront, and his sullen attitude is expressed in a variety of social situations.

Allport believed that traits were concrete features of individuals that uniquely described them and that might be understood by a case study of a single individual. A contrasting view is that traits are dimensions of individual difference that can only be discovered by comparing and contrasting different individuals; individuals are then described in terms of

their standing on a set of common traits. The two definitions are closely related; people who are more anxious than 99 percent of the population presumably have a neuropsychic structure that makes them so anxious.

Characteristics of Traits. Although there are many different trait theories, there is general agreement on several key features of traits: (1) Traits are tendencies to show consistent patterns of thoughts, feelings, and actions. Behaviors that are specific to a single setting or situation may better be considered habits than traits; some evidence of cross-situational consistency is necessary to infer a trait. However, concrete instances of behavior have many determinants, including learned habits, aroused needs, social contexts, role requirements, and the influence of many different and potentially conflicting traits, so the influence of a specific trait on any particular behavior may be quite modest. It is usually only by viewing behavior across many different situations that a consistent pattern can be detected. (2) Traits are relatively enduring features that characterize the individual. In this respect, they are to be distinguished from transient moods or episodes of mental disorder that affect the individual. The fact that traits are relatively enduring does not mean that they cannot change; traits are not immutable, even if they are durable. (3) Traits are continuously distributed, usually approximating a normal or bell curve. Although it is convenient to speak about *introverts* and *extraverts*, in fact, most individuals are *ambiverts*, showing some of the characteristics of introverts and some of the characteristics of extraverts. There is no consistent evidence of discrete personality types.

From time to time, there has been controversy over the reality of traits, fueled by the fact that human beings easily and readily ascribe traits to others on the basis of little or no information, with correspondingly limited accuracy. These personality ascriptions may be triggered by stereotypes of age or physical appearance or may be quite idiosyncratic. Demonstrations of this fact in laboratory experiments by social psychologists, together with the relatively loose cross-situational consistency of most traits, led a generation of psychologists in the 1970s to conclude that traits were cognitive fictions. Subsequent work—particularly, demonstrations that judges who knew the individual well agreed with one another and with the self-reports of the individual about his or her standing on a variety of traits —restored faith in the consensual validity of traits. However, the controversy does make the crucial point that some trait ascriptions are more accurate than others, and that first impressions may be quite misleading. Psychiatrists ought not to assume that their clinical judgments of a patient's personality are correct; validated personality questionnaires and rating forms completed by knowledgeable others may be needed to portray and understand personality accurately.

Personality Structure and Factor Analysis. The most important differences among trait theorists are in the specific traits that they have conceptualized and measured. Jung identified introversion and extraversion as basic personality variables, Bandura emphasized self-efficacy, and Rogers was concerned with openness to experience. Over the years, literally thousands of scales have been developed to measure traits that psychologists considered important in understanding personality. As early as 1936, Allport pointed to another source for identifying personality traits: The natural language. In a monograph he published with Henry Odbert, he listed some 18,000 terms extracted from an unabridged dictionary that could be used to describe people; some of them were mere evaluations (e.g., *swell* and *awful*), but he regarded approximately 4,000 as legitimate trait terms.

The problem for trait psychologists was how to choose a manageable set of traits from among the many possible constructs. It was obvious that trait terms were highly redundant—for example, *anxious*, *worrying*, *nervous*, *apprehensive*, and *fearful* reflect similar, if not identical, characteristics—so what was needed was a procedure for identifying major groups of traits that covaried. Factor analysis, a statistical technique that reduces the complexity of a set of correlations among variables, was first used in personality research by J. P. Guilford (1897–1987) and has remained one of its basic tools. The factors, or dimensions, identified in this process correspond to groups of closely related traits; the set of basic dimensions identified by the factor analysis constitutes a model of the structure of personality traits.

Raymond Cattell (1905–1998) developed one of the first and most influential factor models. He reasoned that, in the course of cultural evolution, any personality trait important in human social interaction would have been noticed and named; the 4,000 trait terms identified by Allport and Odbert could thus be assumed to represent an exhaustive listing of personality characteristics (this has become known as the *lexical hypothesis*). Cattell grouped synonyms and near synonyms together to obtain a set of 35 personality variables and asked respondents to rate acquaintances on each of these sets of terms. He intercorrelated the ratings and factored the correlations, identifying 12 factors. Together with four more factors found in research using self-report questionnaires, these became the basis for the 16 Personality Factor Questionnaire (16PF), a self-report instrument that has been widely used in personality research and clinical psychology for more than 50 years.

It was originally hoped that factor analysis would provide an objective solution to the question of personality structure, but for many years there was little agreement among factor analysts. Eysenck believed that Cattell's model could not be replicated and was needlessly complex. He proposed a simple and powerful two-dimensional model that identified extraversion–

introversion (E) and neuroticism–emotional stability (N) as superfactors and showed that, if the scales of the 16PF were themselves factored, the two largest factors resembled his E and N. He and his wife, Sybil Eysenck (1927 to present), developed a series of instruments to measure these factors (and, later, a third superfactor called *psychoticism*) that has also been widely used, particularly in Great Britain. Eysenck's stature as a learning theorist and a critic of psychoanalysis contributed to the importance of these dimensions in psychiatric contexts.

Five-Factor Model. Eysenck's two factors were widely replicated, but they seemed to omit many important characteristics, such as curiosity, aggression, and achievement striving. An alternative solution was offered in 1961 by two U.S. Air Force psychologists, Ernest Tupes and Raymond Christal. They began with the 35 clusters of traits identified by Cattell and obtained ratings on these clusters in eight different samples. They found that a five-factor solution fit the data in all eight samples. Two of the factors resembled Eysenck's E and N, but three other factors were new.

A small group of lexical researchers, including Warren Norman (1930–1998) and Lewis R. Goldberg (1931 to present), continued to study personality structure as represented by natural language trait adjectives, and, after 20 years, they came to the conclusion that the five-factor structure proposed by Tupes and Christal was essentially correct. Renewed interest in this model showed that the five factors could be recovered in analyses of self-reports and observer ratings; in ratings of children, college students, and older adults; and in several different languages, including non–Indo-European languages, such as Chinese, Hebrew, and Filipino. The factors appeared in analyses of trait adjectives, descriptive phrases, and questionnaire scales. Contemporary five-factor researchers differ somewhat on their conceptualizations of the factors and consequently give them somewhat different labels. The terms *neuroticism* (N), *extraversion* (E), *openness to experience* (O), *agreeableness* (A), and *conscientiousness* (C) are used here. The five factors and some of the traits, or facets, that define them are given in Table 6.4–1, along with associated adjectives.

Many psychologists were skeptical of the lexical hypothesis that had led to the FFM. These critics believed that personality theory and clinical experience would lead to the identification of important traits for which no lay terms existed, and they continued to offer alternative models. Katharine Briggs (1875–1968) and Isabel Myers (1897–1980) operationalized Jung's psychological functions in the Myers–Briggs Type Indicator, which classifies individuals in terms of the dichotomies of introversion versus extraversion, intuition versus sensing, thinking versus feeling, and judging versus perceiving. Some (e.g., Jerry S. Wiggins, 1931–2006) have argued that the traits that influence social interactions were better represented in a circular order than as a set of factors, and many instruments have been

developed to measure this model.

A particularly important system was suggested by Theodore Millon (1928–2014), who was interested in personality traits associated with psychiatric disorders. His reviews of clinical literature led to a theory of personality and psychopathology that specified 11 personality disorders as extreme variants of normal traits. For example, the histrionic personality disorder is supposed to be related to the trait of gregariousness; the schizoid personality disorder is related to the trait of detachment. Millon's theory had a profound impact on the formulation of Axis II in the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III) and in subsequent editions. His instrument, the Millon Clinical Multiaxial Inventory, has been widely used by psychiatrists and clinical psychologists.

The fact that trait theories of personality are usually tied to assessment inventories makes it relatively easy to make empirical comparisons, and, in the last decade, researchers in several countries have undertaken the task of relating different trait models. There is a growing consensus among these researchers that virtually all the traits measured by theory-based personality questionnaires—including those derived from psychodynamic, behavioral, and humanistic theories—are related to one or more of the five factors of Tupes and Christal. However, instruments vary in comprehensiveness. For example, the four scales of the Myers–Briggs Type Indicator correspond to four of the five factors, whereas measures of the interpersonal circumplex represent only two factors (extraversion and agreeableness).

These comparisons of instruments can lead to significant reconceptualizations. For example, Cloninger's Temperament and Character Inventory (TCI) consists of four factors that are intended to assess temperament and three that assess character. However, joint analysis with a measure of the FFM shows that the Harm Avoidance temperament scales and the Self-Directedness character scales are really opposite poles of a single neuroticism factor, whereas the Reward Dependence temperament scales and the Cooperation character scales are both measures of agreeableness. From the perspective of the FFM, the Temperament and Character Inventory's distinction between temperament and character appears not to be supported.

At the broadest level, then, the problem of personality structure appears to have been resolved: Personality is described by five basic factors. This does not mean that personality can be exhaustively described by standing on five dimensions. Most trait psychologists adopt hierarchical models of trait structure. They assume that the broadest factors are composed of more specific traits, which, in turn, are defined by subtraits that correspond to individual items in a questionnaire. In the NEO Personality Inventory 3 (NEO-PI-3), for example, six specific traits or facets are measured for each

of the five factors (or domains) of personality. The facet scales are listed in Table 6.4–1; the hierarchical organization is illustrated in Figure 6.4–1.

Table 6.4–1.
Five Factors of Personality, Defining Traits, and Empirically Associated Adjectives

Factor	Trait	Adjectives
Neuroticism (N)	Anxiety	Anxious, fearful, worrying, tense
	Angry hostility	Irritable, impatient, moody, not gentle
	Depression	Pessimistic, worrying, not contented, moody
	Self-consciousness	Shy, not self-confident, timid, inhibited
	Impulsiveness	Hasty, self-centered, excitable, loud
	Vulnerability	Not confident, not efficient, not clear-thinking, anxious
Extraversion (E)	Warmth	Friendly, warm, sociable, not aloof
	Gregariousness	Sociable, outgoing, talkative, not withdrawn
	Assertiveness	Assertive, forceful, aggressive, confident
	Activity	Energetic, hurried, quick, active
	Excitement seeking	Pleasure-seeking, adventurous, daring, spunky
	Positive emotions	Enthusiastic, humorous, optimistic, jolly
Openness (O)	Fantasy	Dreamy, imaginative, artistic, complicated
	Aesthetics	Artistic, original, inventive, idealistic
	Feelings	Excitable, spontaneous, affectionate, insightful
	Actions	Wide interests, versatile, adventurous, imaginative
	Ideas	Curious, original, insightful, inventive
	Values	Not conservative, not cautious, flirtatious, unconventional
Agreeableness (A)	Trust	Trusting, not suspicious, forgiving, not wary
	Straightforwardness	Not shrewd, not autocratic, not charming, not demanding
	Altruism	Soft-hearted, gentle, generous, kind
	Compliance	Not stubborn, not demanding, not headstrong, not impatient
	Modesty	Not a show-off, not clever, not argumentative, not self-confident
	Tender-mindedness	Sympathetic, soft-hearted, warm, kind
Conscientiousness (C)	Competence	Efficient, thorough, resourceful, intelligent
	Order	Organized, precise, methodical, thorough
	Dutifulness	Thorough, not careless, not distractible, not lazy
	Achievement striving	Ambitious, industrious, enterprising, persistent
	Self-discipline	Energetic, not lazy, organized, not absent-minded
	Deliberation	Not hasty, not impulsive, not careless, not immature

Adjectives are significantly correlated with traits, $N = 305$; $p < 0.001$.

Adapted from McCrae RR, Costa PT Jr. Discriminant validity of NEO-PI-R facet scales. *Ed Psychol Meas*. 1992;52:229.

Assessment on the level of specific facets provides a more detailed and personalized portrait of the individual.

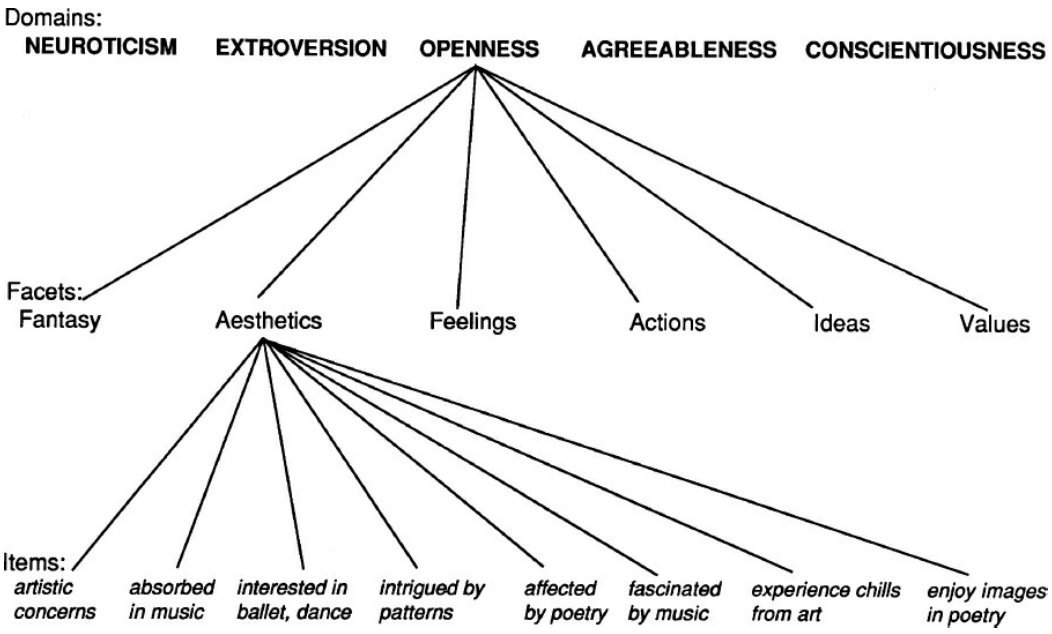

FIGURE 6.4–1. An example of hierarchical organization in the Revised NEO Personality Inventory: domains, facets of openness, and items measuring openness to aesthetics.

Origin and Development of Personality Traits. Psychoanalytic, learning, and humanistic theories have usually offered causal explanations for individual differences. Thus, differences in self-efficacy are supposed to be caused by different histories of success in pursuing goals; variations in openness to experience might be the result of differences in the conditions of worth imposed by parents. The factor analysts who scoured the dictionary for trait terms usually bypassed this issue, being content to describe the personality differences they found in adults. In principle, some traits might be inherited, some might be instilled by parents, and some might be learned from experience; whatever their origin, they could be measured and used to predict important life criteria.

Until recently, it was generally assumed that personality was shaped primarily by a variety of environmental influences, including parental love and discipline, social and economic opportunities, and life experiences through childhood and adolescence. Surprisingly, this assumption has rarely been tested: There has been little prospective longitudinal research documenting links between early childhood experiences and subsequent adult personality. A different research design, using the techniques of behavior genetics, can answer these questions by comparing personality measures in adults with different degrees of genetic and environmental similarity. For example, similarity between identical twins reared apart can be attributed to genetic influences, whereas similarity between adopted siblings reared together must be due to shared environmental influences.

Since the early 1980s, behavior genetics studies using many different samples, personality measures, and methods of data analysis have converged on the surprising conclusion that personality traits are, to a

considerable extent, heritable, and, to a considerable extent, due to unknown and idiosyncratic causes. The molecular genetics underlying trait heritability has been a topic of great interest in the past few years, but research approaches are still being refined. Early studies claiming to find single genes for major traits are giving way to newer paradigms. These more sophisticated approaches use combinations of traits to define personality styles or patterns that better define the personality phenotypes for genetic linkage and association studies. In addition, gene–environment and gene–gene interactions are leading to new insights into the role that genotypes can play in moderating environmental substrates.

An interesting and important conceptual extension concerns these G×E and G×G interactions in influencing personality traits but also clinically important biological outcomes such as cardiovascular disease (CVD). For example, Redford Williams (1940 to present) has proposed a comprehensive model whereby genetic variants and environmental stressors act to influence brain functions involved in the expression of behaviors, psychological traits (personality factors) and states, and neuroendocrine and autonomic functions that influence the expression of biological endophenotypes that, in turn, influence the development and clinical course of coronary heart disease. In an extensive body of research Williams and colleagues have documented the actions of elements hypothesized by this model. They have shown, for example, that a functional single neucleotide polymorphism (SNP, rs6318, C/G) of the 2C serotonin receptor (5HTR2C) gene on the X chromosome moderates the increase in blood cortisol levels during acute mental stress in healthy samples: men carrying the rs6318 C allele and women with the CC genotype have cortisol responses to stress that are twice as large as those in men and women carrying the G allele. In a sample of CHD patients, male C carriers and females with the CC genotype had a 38% higher 5-year incidence of death or myocardial infarction than G carriers, with the impact on adverse events larger in CC females (62%) than C males (23%). They also found in the nationally representative Add Health sample that females with the rs6318 CC genotype who reported high chronic stress levels had depressive symptom levels that were 0.3 standard deviation higher than female G carriers with similar high stress levels. As predicted by the overall model, these findings indicate that healthy persons with a gene variant that is associated with larger cortisol responses to acute mental stress are also more likely to be more depressed if exposed to chronic stress and, once CHD is present, to have a more adverse clinical course.

Less success has been found in identifying environmental origins of personality traits. In fact, they are hardly attributable to shared environmental causes at all. Socioeconomic status, family diet, religious training, parental modeling, and all the other influences that children growing up in the same household would normally share seem to have little

or no influence on adult personality. This conclusion applies as much to character traits, such as achievement striving and modesty, as to temperament traits, such as anxiety and energy level. This dramatic and counterintuitive finding will doubtless reshape theories of personality. As Sandra Scarr (1936 to present) commented, "psychology [currently] has no adequate theories to account for individual variation in behavior because our theories address the wrong sources of variation."

One new theory that does begin to address the "right" sources of variation is five-factor theory. This theory attempts to provide a conceptual framework in which the body of research on the FFM can be interpreted. Five-factor theory is a systems theory, represented schematically in Figure 6.4–2. The central components are *basic tendencies* (including personality traits and intelligence and other abilities) and *characteristic adaptations* (such as skills, attitudes, values, and roles). In this model, as suggested by behavioral genetics, the only input to basic tendencies is from biological bases. However, personality traits and *external influences* are important in shaping characteristic adaptations. For example, an individual who is temperamentally high in self-consciousness and who repeatedly experienced ridicule as a child might develop an avoidant personality disorder, which is a *characteristic (mal)adaptation*.

For many years developmental psychologists assumed that personality development ended with adolescence. A 21-year-old is legally an adult and is unlikely to grow much in physical height or intellectual capacity, but recent research has shown consistent changes in personality between college age and middle adulthood. There is a dramatic decline in excitement seeking over that interval, but, in addition, cross-sectional and longitudinal studies of American college students have shown that of the five-factor traits, N, E, and O decline during the decade of the 20s, whereas A and C increase. As a result, 30-year-olds are less emotional and better socialized than 20 year olds.

There is reason to believe that these changes are not the product of the relatively indulgent American experience of adolescence. The same pattern of age differences is seen in less-affluent countries, such as Croatia, and in non-Western countries, such as South Korea. Such cross-cultural studies suggest that there may be intrinsic maturational processes in personality development, which is consistent with the biological basis of personality traits hypothesized by five-factor theory.

There have also been important developments in understanding personality change in individuals older than 30 years of age. Traits are defined as relatively enduring patterns of behavior, but most psychologists assumed that traits would be modified by life experiences. Jung postulated that the process of individuation required each individual to express all of his or her potentials, and, thus, the young introvert would normally become an extravert in old age, and vice versa. Lay stereotypes of aging held that, as

they age, people become cranky, conservative, and depressed. Gerontologists rebutted these myths of aging, arguing that old age was more likely to bring maturity, wisdom, or detachment. Theories of adult development, popularized in Gail Sheehy's best-selling *Passages*, suggested that personality changed in stages and, in particular, that men (and perhaps women) went through a midlife crisis in their 30s or 40s.

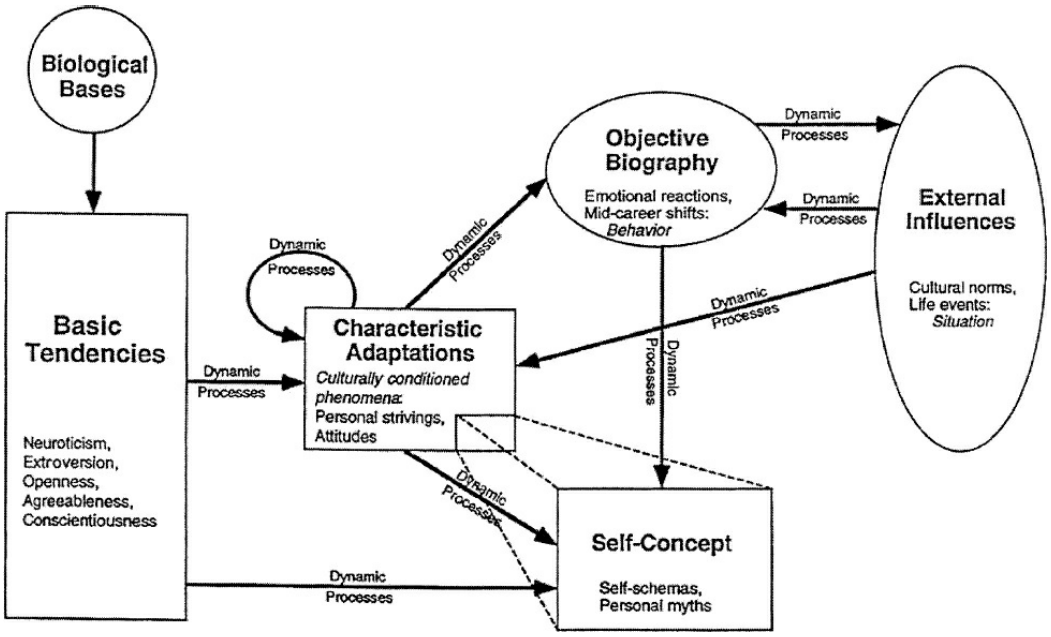

FIGURE 6.4–2. A representation of the personality system as depicted by the five-factor theory. The *ellipses* represent peripheral elements, and the *rectangles* represent central elements. Examples of content are given in each. *Arrows* indicate the direction of causal influences. (Adapted from McCrae RR, Costa PT Jr. A five-factor theory of personality. In: Pervin LA, John OP, eds. *Handbook of Personality: Theory and Research*. 2nd ed. New York: Guilford; 1999:139.)

In sharp contrast to all these theories are the findings from a number of independent longitudinal studies of personality in adulthood. These studies present a clear picture of predominant stability of the full range of personality traits. Although there are very gradual changes in the mean levels of traits across adulthood, individuals tend to maintain the same rank order. The 30 year old who is outgoing, curious, and hardworking is likely to become an outgoing, curious, and hardworking 80 year old. This conclusion has been supported by studies using self-reports and observer ratings of all five factors of personality and appears to apply to men and women. Of course, there are often dramatic changes in personality as a result of dementing disorders in old age, but normal life experience seems to have little impact on personality in individuals older than 30 years of age.

There are distressingly few data on the long-term effects of psychopathology on personality. Individuals in recovery from an episode of major depressive disorder generally score high on measures of the five-

factor trait N, and it is sometimes argued that this is a result of the experience of depression. However, the few prospective studies of the first onset of major depression typically show that these individuals scored high on N before the first episode, suggesting that high N may be an enduring feature of people prone to clinical depression.

Theories of Psychopathology

The links between psychopathology and trait psychology are at once intimate and complex. For decades, psychiatrists and clinical psychologists have administered psychological tests such as the Minnesota Multiphasic Personality Inventory (MMPI) and Cattell's measure of personality, the 16PF, to inform the diagnosis of mental disorders; these same tests have been used routinely by trait psychologists to examine relations between personality and such criteria as vocational preference, creative potential, and coping with life stress. The dividing line between normal variations in personality dispositions and psychopathology is frequently unclear.

This poses more of a problem to psychiatric nosology than it does to trait theories of personality. Psychiatric diagnoses are supposed to represent the presence of a discrete disorder, and the DSM-5 specifies the criteria that must be met to confer each diagnosis. This categorical model is appropriate for some disorders but not others. It may be difficult to separate social phobia, for example, from extreme shyness and self-consciousness or a schizoid personality disorder from marked introversion. This is entirely consistent with trait models of personality, which regard individual differences as continuously distributed variables.

Neuroticism and Psychopathology. One theory of psychopathology, therefore, is that some psychiatric disorders reflect extreme standing on personality traits, particularly those related to the five-factor trait N. Among the traits that covary in normal populations to define this factor are predispositions to experience chronic levels of negative affects, such as fear, anger, shame, and sadness. Individuals with high standing on these traits may qualify for a diagnosis of generalized anxiety disorder, borderline personality disorder, social phobia, or dysthymia.

The five-factor trait N may also be considered a risk factor for the development of psychiatric disorders that are not themselves trait-like. Several recent studies have demonstrated that individuals scoring high on measures of N are at increased risk of subsequently receiving a diagnosis of major depressive disorder. Poor control of urges and impulses and excessive concern with physical functioning are also characteristics associated with N, and it might be hypothesized that individuals high in N are predisposed to develop eating disorders and hypochondriasis. N is, in essence, a generalized disposition to experience psychological distress, and individuals seeking psychiatric treatment are almost always distressed. It is

therefore not surprising that virtually all clinical populations, from drug abusers to schizophrenics, score high on measures of N.

Historically, the term *neurosis* or *psychoneurosis* was coined to identify psychiatric disorders of functional origin that were closely related to anxiety and its control. It survived in the revised third edition of DSM (DSM-III-R) in many secondary labels (dysthymia, e.g., was also called *depressive neurosis*) but is not used in subsequent editions. Many personality psychologists object to the term *neuroticism* as a label for a dimension of normal personality because of its suggestion of psychopathology. Many psychiatrists object to the term because it appears tied to an outdated psychiatric nosology. However, the term serves a useful purpose: It is a reminder to clinicians that patients with a wide variety of diagnoses share many features related to chronic psychological distress and is also a reminder to personality psychologists that the difference between normal and abnormal functioning is often only one of degree. The cliché that people are all more or less neurotic has some scientific basis.

g

Personality Traits and Personality Disorders. DSM-5 provides a general definition of a personality disorder that applies to all 10 of the specific disorders: "A personality disorder is an enduring pattern of inner experience and behavior that deviates markedly from the expectations of the individual's culture, is pervasive and inflexible, has an onset in adolescence and early adulthood, is stable in time, and leads to distress or impairment." (p. 645)

It is reasonable to ask whether these inflexible and maladaptive traits are the same as or different from those encountered in nonpsychiatric populations. Several recent studies on this question have concurred in finding strong and replicable links between scales measuring personality disorders and the five factors in normal and clinical populations.

In the case of N, only high scores are associated with psychiatric impairment, but both poles of the other factors appear to be associated with specific forms of psychopathology. Individuals high in E tend to have histrionic and narcissistic personality disorders; those who are low in E have avoidant and schizoid disorders. High C is associated with obsessive-compulsive personality disorder; low C is associated with antisocial personality disorder. Low A, or antagonism, is characteristic of individuals with paranoid, antisocial, and narcissistic personality disorders. High A is associated with dependent personality disorder.

FFM Descriptions of Personality Disorders. Thomas Widiger and colleagues have linked the personality disorders with FFM traits. In one study, the diagnostic criteria and text descriptions of each personality disorder were coded in terms of the 60 poles of the 30 facets of the FFM

(see Table 6.4–1). In other studies, personality disorder researchers and practicing clinicians were asked to describe a prototypic case of each of the ten personality disorders using the negative and positive poles of the 30 facets of the FFM. Table 6.4–2 presents an integrative summary of these FFM descriptions.

These associations show some of the ways in which disorders can be conceptually distinguished. For example, avoidant and schizoid personality disorders are both characterized by low gregariousness, but only the former shows associations with high anxiety, depression, self-consciousness, and vulnerability. Research in several countries has generally supported this mapping of disorders onto personality traits, although the conceptual distinctions highlighted in Table 6.4–2 are usually blurred in real data because of the pervasive comorbidity of personality disorders.

It is a matter of current controversy just how these personality traits are related to the disorders: Are normal personality traits carried to an extreme inherently maladaptive, or do these traits merely predispose individuals to develop disorders under certain circumstances? Some research has also called into question the meaningfulness of the syndromes recognized by DSM-5 by showing that the symptoms used to define the disorders, when separately assessed, do not covary in ways that match the DSM-5 syndromes. Instead, the symptom clusters that do emerge are interpretable in terms of the FFM. As a result of such evidence, proposals have been made to replace the current categorical model by dimensional models that relate problems in adjustment to standing on basic dimensions of personality. This would appear to be a logical extension of trait theories of psychopathology.

Revisions in the DSM-5. More than previous editions, the DSM-5 highlights the importance of individual differences, particularly in personality functioning and personality traits, stating that "the DSM like other medical disease classifications should accommodate ways to introduce dimensional approaches to mental disorders including dimensions that cut across current categories" (p. 5) but further "The DSM-5 Task Force recognized that it is premature scientifically to propose alternative definitions for most disorders." (p. 13)

After a 12-year process (including research conferences, task force meetings, drafts and revisions posted on a public website, etc.), the Personality and Personality Disorders Work Group presented an alternative dimensional model, which was strongly endorsed by the APA DSM-5 Task Force. This alternative model is similar to the dimensional FFM of general personality structure. "These five broad domains are maladaptive variants of the five domains of the extensively validated and replicated model known as the 'Big Five' or 'Five Factor Model of Personality (FFM)' ..." (p. 773).

However, the American Psychiatric Association Board of Trustees overruled the DSM-5 Task Force and retained all 10 DSM-IV-TR personality disorder categories, placing the Alternative Model in Section III —Emerging Measures and Models, hoping that such a move would allow additional research and give clinicians time to become better acclimated to this dimensional approach to diagnosis.

An alternative way to revise the diagnosis of personality disorders has also been offered by Thomas Widiger and colleagues, who have proposed a four-step process. In the first step, domains and facets of personality would be assessed, providing all the benefits of trait assessment described in the following discussion. In the second step, clinicians would identify specific problems in living associated with salient traits. For example, it might be noted that a patient high in self-consciousness experiences intense embarrassment when around strangers. In the third step, the clinician would evaluate the degree of impairment that the problems caused, to determine if they are clinically significant. Does the patient's embarrassment interfere with holding a job or finding a mate? Finally, in an optional fourth step, the clinician could interpret the full personality profile to see if it matched the prototype for an identified personality disorder. For example, if the patient was high in anxiety, depression, and vulnerability, as well as self-consciousness, and low in gregariousness, assertiveness, and excitement seeking, he or she could be considered to have an avoidant personality pattern (Table 6.4–2). A major advantage of this system is that patients who did not meet full criteria for any recognized personality disorder could still be treated for personality-related problems in living.

Table 6.4–2.
Hypothesized Relations between DSM-5 Personality Disorders and Five-Factor Model Personality Traits

Personality Trait	Paranoid	Schizoid	Schizotypal	Antisocial	Borderline	Histrionic	Narcissistic	Avoidant	Dependent	Obsessive-Compulsive
Neuroticism										
Anxiety			H	L	H			H	H	H
Angry hostility	H			H	H		H			
Depression					H			H	H	
Self-consciousness			H	L	H	H		L/H	H	
Impulsiveness				H	H			L		
Vulnerability				L	H	H	H	H	H	
Extraversion										
Warmth	L	L	L			H			H	L
Gregariousness	L	L	L			H	H	L		
Assertiveness		L		H			H	L	L	H

Activity		L		H				L		
Excitement seeking		L		H		H	H	L		L
Positive emotions	L	L	L							
Openness to experience										
Fantasy			H		H	H	H			
Aesthetics										
Feelings		L				H				L
Actions	L	L	H	H				L		L
Ideas			H							
Values	L									L
Agreeableness										
Trust	L		L	L	L	H	L		H	
Straightforwardness	L			L	L	L	L			
Altruism				L	L		L		H	
Compliance	L			L	L				H	
Modesty				L		L	L	H	H	
Tender mindedness	L			L			L			
Conscientiousness										
Competence									L	H
Order						L				H
Dutifulness				L						H
Achievement striving							H			H
Self-discipline				L					L	H
Deliberation				L	L	L				H

High and low ratings are based on Table 6.1, DSM IV Personality Disorders from the Perspective of the Five-Factor Model of General Personality Structure in Chapter 6, Five-Factor Model Personality Disorders Research.

Adapted from Widiger TA, Costa PT, Gore Wl, Crego C. In: Widiger TA, Costa PT Jr, eds. *Personality Disorders and the Five-Factor Model of Personality*. 3rd ed. Washington, DC: American Psychological Association; 2013.

Application of Theory to Therapy

Psychodynamic, behavioral, and humanistic theories all specify the causal mechanisms by which psychopathology is created and maintained and thus imply points of intervention. If maladjustment is caused by rigidly internalized conditions of worth, then unconditional positive regard in the therapeutic setting may provide a cure; if the problem is a learned behavior, then altering reinforcements may extinguish it. Trait models do not emphasize causal or developmental explanations and thus may seem to have no implications for psychotherapy. In fact, they have many.

Recent evidence on the substantial heritability of most personality traits clearly indicates that Allport was right: Traits do have some underlying neuropsychic structure. Research on the psychophysiology of traits is a topic of growing interest and parallels the extensive work done on the neurophysiological basis of many forms of psychopathology. In this broad

sense, then, psychopharmacological approaches to psychotherapy are, in principle, consistent with trait theories of personality: If personality traits and disorders reflect brain processes, drugs that affect the brain may offer useful instruments for psychotherapy. In practice, there is still much to learn. For example, individuals with dysthymia, who score high on measures of trait depression, often do not respond to antidepressant medication.

Personality Assessment. Historically, the chief role of trait psychology in psychotherapy has been in diagnosis and assessment. A number of self-report measures, such as the MMPI and the Millon Clinical Multiaxial Inventory, were designed specifically to measure psychopathology and include primarily items that tap psychiatric symptoms. They can be regarded as measuring psychopathological traits and states (although they have often been used in college and volunteer samples as measures of personality per se). These instruments are chiefly of value as aids to psychodiagnosis. General personality questionnaires, including the 16PF, the Guilford-Zimmerman Temperament Survey, and the Edwards Personal Preference Schedule, have also been used for decades in clinical psychology and psychiatry as part of a complete psychological assessment. Several measures of the FFM have recently appeared, including the Big Five Inventory and the NEO Personality Inventory, and there has been considerable interest in the clinical application of the FFM. The primary advantage of this model is its comprehensiveness. By assessing traits from all five dimensions, the clinician can efficiently obtain a full portrait of the individual.

Some psychiatrists are concerned that personality scores obtained from acutely depressed patients are invalid. This concern is based on the empirical finding that scores, particularly on scales related to the five-factor trait N, change with remission and on the belief that personality traits cannot change. If the latter belief were true, it would be necessary to conclude that depression distorts self-reports of personality. However, as Figure 6.4–2 suggests, personality traits may change if their biological basis is altered, and acute major depression has demonstrated effects on brain functioning. Controlled studies and clinical experience suggest that trait scores obtained from acutely depressed patients are indeed informative in most cases.

Personality psychologists and psychometricians have devoted years to the development of self-report inventories, and the usefulness of this approach to assessment is beyond doubt. Self-reports, however, are by no means infallible. Patients may not understand their own personalities or may deliberately misrepresent themselves. Concerns about defensiveness and socially desirable responding have led to the use of projective tests, the development of special validity scales to detect and to correct for distorted

responses, and reliance on the clinical judgment of the psychiatrist. Each of these possible solutions introduces problems of its own, however, and most clinicians rely on multiple sources of information.

One source of information that has been underused in clinical settings is informant ratings. Research in the past decade has confirmed that ratings on standardized instruments by significant others—usually spouses or parents—can provide reliable and valid assessments of personality. These findings appear to hold for clinical, as well as volunteer, samples. Personality ratings may be particularly valuable when patients are incapacitated or are strongly motivated to present an overly favorable or unfavorable picture of themselves. A case example of the clinical interpretation of informant ratings is given below.

Uses of Trait Profiles. Traits are enduring and stable features of the patient's behavior and experience, so assessing a broad range of traits gives the clinician a sense of what the person is like, which can be useful for many purposes beyond the formulation of a diagnosis. A complete personality profile can point to the patient's strengths as well as weaknesses, can help predict the course of therapy, and can aid in the selection of an optimal mode of treatment.

Humanistic theories of personality stress human potentials for altruism, creativity, and commitment and argue that psychotherapy must use these assets. Trait theorists would qualify that stance: They believe that some people are altruistic, creative, or committed, but others are not. Assessing traits allows the clinician to identify and capitalize on the particular strengths that characterize each patient.

Standing on trait dimensions can also give clues to the probable course of therapy. Patients who are low on the five-factor trait A are distrustful and uncooperative, and it may prove difficult to form a therapeutic alliance with them; by contrast, patients who are extremely high on A may be excessively compliant and become dependent on the therapist. The five-factor trait C involves commitment and self-discipline. Patients high in C will probably adhere to treatment recommendations and work hard to solve their problems; those low in C are less persistent and dedicated, and the therapist may have to work to motivate them.

Finally, a few controlled studies and a good deal of clinical experience suggest that personality traits influence the effectiveness of various kinds of treatment interventions. Interpersonal therapies, in which patients are required to speak a great deal about themselves, appear to be most effective for extraverts; pharmacological management may be superior for introverts. Similarly, individuals who are high in the five-factor trait O benefit from such techniques as guided imagery, whereas those who are low in O prefer biofeedback. Ideally, the choice of therapies should be guided not only by the nature of the disorder, but also by the enduring

characteristics of the patient.

An in-depth case study incorporating trait data with other assessment perspectives is provided in Jerry Wiggins's volume on *Paradigms of Personality Assessment*. Madeline G was a flamboyant Native American woman. She had overcome a history of early delinquency to become a lawyer who fought for the rights of her people. She had no children, and her common-law husband left her a year after her personality assessment. This led to a period of depression, and 3 years after her assessment she still remained alone. She provided self-reports on the NEO-PI-R, and her husband rated her. They agreed that she was extraverted, open to experience, and very low in agreeableness, but disagreed on neuroticism and conscientiousness. Figure 6.4–3 plots her husband's ratings of her.

This is a dramatic profile, with a number of extreme scores. From her husband's perspective, she is high in neuroticism, especially angry hostility and impulsiveness. Although he sees her as high in most facets of extraversion, he rates her as very low in warmth. Similarly, he describes her as being uniformly low in facets of agreeableness. He acknowledges that she is high in achievement striving, but claims she is very low in dutifulness and deliberation.

The four-step process of personality disorder assessment predicts that she would be vulnerable to a specific set of problems in living. Her high angry hostility means that she is likely hypersensitive to criticism and prone to react with rage. Her low warmth suggests that she may have difficulty developing or sustaining personal, intimate relationships. Low agreeableness can lead to rudeness that alienates friends and to trouble with the law (as seen in her adolescent delinquency). Because she is very low in self-discipline, she might be expected to have difficulty holding a job, but in fact she is successful in her career. Although personality traits suggest the possibility of certain problems in living, the clinician must determine whether or not those problems are present in a particular case.

The clinician must also judge whether the problems are severe enough to warrant a diagnosis and treatment. Madeline G did not seek treatment, but both her depression and inability to develop new intimate relationships after her husband left her might have justified a psychiatric intervention.

Whether Madeline G met criteria for the diagnosis of one of the ten DSM-5 personality disorders cannot be ascertained simply by looking at her personality profile. However, she shares personality traits with patients who have paranoid, schizotypal, antisocial, borderline, histrionic, and narcissistic personality disorders. Conversely, she is unlikely, according to her husband's description, to have schizoid, avoidant, or dependent personality disorders.

Madeline G's personality profile suggests that if she were in psychotherapy, she would find it difficult to form a therapeutic alliance (low agreeableness) and might have some problems making a sustained effort at the work of therapy (low conscientiousness). But she does have a number of strengths, including high assertiveness and achievement striving. Because she is quite open to experience, she would be willing to consider interpretations and suggestions for change made by her therapist.

Psychotherapy and Personality Change. If psychopathology is an expression or an outcome of personality traits and if, as longitudinal studies demonstrate, personality traits change little over time, how can psychotherapy be effective? The pessimistic answer is that it cannot. Many psychiatric disorders are lifelong or recurrent, and treatment consists of management rather than cure. It is well known that those who have the best prognosis for recovery are those who are initially least impaired. In these cases, the disorder may represent a transient adjustment reaction among individuals who have relatively healthy personality traits. Individuals high in N and low in A and C may seem to be poor candidates for psychotherapy.

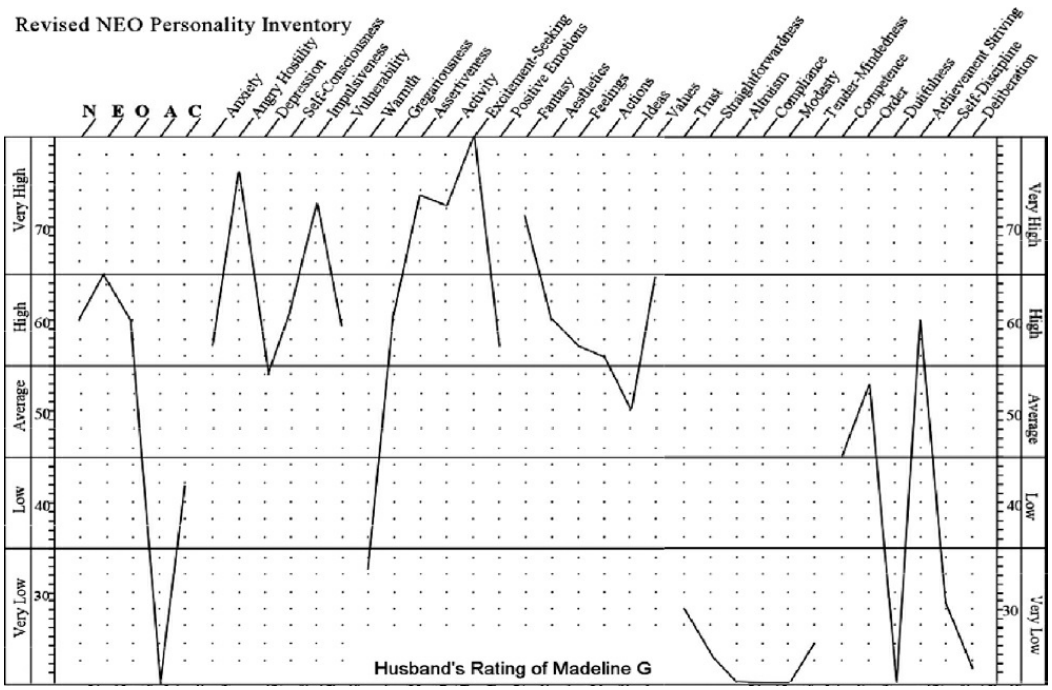

FIGURE 6.4–3. Personality profile of Madeline G as rated by her husband. *T*-scores (*M* = 50; *SD* = 10) comparing her to other adult men and women are plotted. The five factors are given on the left; the 30 facets, grouped by factor, are given toward the right. (NEO-PI-R profile form reproduced by special permission of the publisher, Psychological Assessment Resources, Inc., 16204 North Florida Avenue, Lutz, FL 33549, from the Revised NEO Personality Inventory by Paul T. Costa, Jr., and Robert R. McCrae. Copyright 1978, 1985, 1989, 1991, 1992, 2010 by Psychological Assessment Resources, Inc. [PAR]. Further reproduction is prohibited without permission of PAR.)

However, such a pessimistic conclusion is unwarranted. Longitudinal studies show that people change little in the course of normal experience, but they do not rule out the possibility that direct interventions can make changes. It is unlikely that psychotherapy will make dramatic and lasting changes in basic personality traits, but more modest improvements may be enough to allow the patient to function adequately in daily life. Interventions may be particularly effective in early adulthood, when traits are not yet fully developed. Even without changing personality, therapy may help individuals to adapt to their own nature. The chronically anxious person may learn techniques of relaxation; the disagreeable individual may learn social skills that improve interpersonal relationships.

In some sense, most forms of psychotherapy can be seen as learning experiences. Psychoanalysts promote insight into unconscious conflicts; behaviorists help their clients understand the contingencies that reinforce troubling behavior; humanistic psychologists encourage patients to discover their true potential. Trait psychology can also be a source of insight. All human beings are, in some respects, alike, and it is often comforting to learn that one is not alone in one's suffering. In other respects, however, people are different from one another, and it can also be a therapeutic experience to learn this fact. Helping patients understand

their own enduring dispositions can prepare them for some of the problems they face in life.

References

Allport GW. *Personality: A Psychological Interpretation*. New York: Holt; 1937.

Bandura A. *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall; 1977.

Brummett BH, Babyak MA, Williams RB, et al. A putatively functional polymorphism in the HTR2C gene is associated with depressive symptoms in white females reporting significant life stress. *PLoS One*. 2014;9(12):e114451.

Costa PT Jr, McCrae RR. The NEO Inventories as instruments of psychological theory. In: Widiger, TA, ed: *The Oxford Handbook of the Five-Factor Model of Personality*. New York: Oxford University Press; 2015.

Costa PT Jr, McCrae RR. A Five-Factor Model perspective on personality disorders. In: Strack S, ed. *Handbook of Personology and Psychopathology*. New York: Wiley; 2005.

Digman JM. Personality structure: Emergence of the Five-Factor Model. *Annu Rev Psychol*. 1990;41:417.

Dollard J, Miller NE. *Personality and Psychotherapy: An Analysis in Terms of Learning, Thinking, and Culture*. New York: McGraw-Hill; 1950.

Eysenck HJ, Eysenck M. *Personality and Individual Differences*. London: Plenum; 1985.

*Feist J, Feist GJ, Roberts TA. *Theories of Personality*. 8th ed. New York: McGraw-Hill; 2012.

Frankl VE. *On the Theory and Therapy of Mental Disorders: An Introduction to Logotherapy and Existential Analysis (JM Dubois, Trans.)*. New York: Brunner-Routledge; 2004.

Herbst JH, Zonderman AB, McCrae RR, Costa PT Jr. Do the dimensions of the Temperament and Character Inventory map a simple genetic architecture? Evidence from molecular genetics and factor analysis. *Am J Psychiat*. 2000;157:1285.

Hiriyanna M. *Outlines of Indian Philosophy*. London: Allen & Unwin; 1932.

Kelly GA. *The Psychology of Personal Constructs*. New York: Norton; 1955.

*Maddi SR, Costa PT Jr. *Humanism in Personology: Allport, Maslow and Murray*. Chicago: Aldine; 1972.

Maslow AH. *Motivation and Personality*. New York: Harper & Row; 1954.

McCrae RR. A more nuanced view of reliability: Specificity in the trait hierarchy. *Pers Soc Psychol Rev*. 2015;19:97.

*McCrae RR, Costa PT Jr. *Personality in Adulthood: A Five-Factor Theory Perspective*. New York: Guilford; 2003.

*McCrae RR, Harwood TM, Kelly SL. The NEO Inventories. In: Harwood TM, Beutler LE, Groth-Marnat G, eds. *Integrative Assessment of Adult Personality*. 3rd ed. New York: Guilford; 2011.

McCrae RR, John OP. An introduction to the Five-Factor Model and its applications. *J Pers*. 1992;60:175.

McCrae RR, Terracciano A, 78 members of the Personality Profiles of Cultures Project. Universal features of personality traits from the observer's perspective: Data from 50 cultures. *J Pers Soc Psychol*. 2005;88:547.

Miller T. The psychotherapeutic utility of the Five-Factor Model of personality: A clinician's

experience. *J Pers Assess*. 1991;57:415.

Mõttus R, Kandler C, Bleidorn W, Riemann R, McCrae RR. Personality traits below facets: The consensual validity, longitudinal stability, heritability, and utility of personality nuances. *J Pers Soc Psychol*. In press.

Rogers CR. *On Becoming a Person: A Therapist's View of Psychotherapy*. Boston, MA: Houghton Mifflin; 1961.

Sayer NA, Noorbaloochi S, Frazier PA, Pennebaker JW, Orazem RJ, Schnurr PP, Murdoch M, Carlson KF, Gravely A, Litz BT. Randomized controlled trial of online expressive writing to address readjustment difficulties among U.S. Afghanistan and Iraq War Veterans. *J Trauma Stress*. 2015;28:381.

Scarr S. Distinctive environments depend on genotypes. *Brain Behav Sci*. 1987;10:38.

Schopenhauer A. *The World as Will and Representation*. 3rd ed. New York: Dover; 1969.

Skinner BF. *About Behaviorism*. New York: Knopf; 1974.

Vachon DD, Lynam DR, Widiger TA, Miller JD, McCrae RR, Costa PT Jr. Basic traits predict the prevalence of personality disorder across the lifespan: The example of psychopathy. *Psychol Sci*. 2013;24:698.

*Widiger TA, Costa PT Jr, McCrae RR. Diagnosis of personality disorders using the Five-Factor Model. In: Widiger TA, Costa PT Jr, eds: *Personality Disorders and the Five-Factor Model of personality*. 3rd ed. Washington, DC: American Psychological Association; 2013.

Widiger TA, Trull TJ. Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. *Am Psychol*. 2007;62:71.

Widiger TA, Trull TJ, Clarkin JF, Sanderson C, Costa PT Jr. A description of the DSM-IV personality disorders with the Five-Factor Model of personality. In: Costa PT Jr, Widiger TA, eds: *Personality Disorders and the Five-Factor Model of Personality*. 2nd ed. Washington, DC: American Psychological Association; 2002.

Wiggins JS. *Paradigms of Personality Assessment*. New York: Guilford; 2003.

7

Diagnosis and Psychiatry: Examination of the Psychiatric Patient

▲ 7.1 Psychiatric Interview, History, and Mental Status Examination of the Adult Patient

RORY P. HOUGHTALEN, M.D., AND JOHN S. MCINTYRE, M.D.

INTRODUCTION

George Engel said that the clinical interview in medicine is "Virtually indispensable for the physician–patient interaction, the well-constructed interview truly may be regarded as the most powerful, the most sensitive and the most versatile instrument available to the physician." Despite advances in science and technology, this statement remains true today and perhaps nowhere in medicine is it more salient than in psychiatric practice, a specialty that relies almost entirely on a careful history and examination to develop a diagnosis.

Two primary aims shape the initial psychiatric interview. One is to develop a biopsychosocial database that informs diagnosis, risk assessment and person (patient)-centered treatment planning. The other is to form a positive alliance that sets the stage for effective treatment that may follow the initial encounter. This chapter will focus on the initial psychiatric interview, but much of the commentary here is relevant to patient contacts in the course of treatment as well.

Among the settings in which psychiatric interviews take place are psychiatric inpatient units, medical-surgical inpatient units, partial hospital programs, emergency rooms, outpatient offices, recovery services, nursing homes, other group homes, and correctional facilities. Emerging technologies like telemedicine are making virtual "contacts" with patients possible.

The sections that follow will cover general principles and the processes of the diagnostic psychiatric interview from the initial request for consultation to the closing of the contact. Content includes commentary about documentation practices, addresses factors that influence the

dynamics and quality of communication and provides commentary on interviewing techniques. The chapter closes with a review of difficult patient scenarios and special topics related to the psychiatric interview.

GENERAL PRINCIPLES

Pre-Interview Contacts and Interactions

Referrals for psychiatric evaluation come from many sources including patients seeking help, concerned family, other medical providers, schools, workplaces and other agencies interested the person's well-being. Some referrals are for reasons other than clinical treatment, for example, forensic evaluations. These types of referrals require special considerations that this chapter's section *Forensic Interviews* briefly reviews.

The first contact in ambulatory settings usually begins with a phone call during which demographic, insurance information is obtained, and the patient (or other caller) is briefed on what to expect in the initial encounter including the length of the initial interview, fees and other particulars of the office or service setting. Staff interacting with callers should be adept at differentiating urgent from routine referrals. Psychiatrists may speak personally to a newly referred patient before the first interview to determine the urgency of the problem, triage to the appropriate level of care and to determine if the problem is something within the scope of their expertise and capability to manage.

Most offices ask new patients to bring to the first encounter a list of current medical providers and medications, and may ask the patient to begin the process of securing records from previous providers of mental health care. Some offices send out a packet to gather information prior to the first interview containing questionnaires for the patient to return by mail, submit via a website link, or to bring in person to the first appointment.

A note confirming the reason for the referral and offering background information often accompanies referrals from other medical or mental health providers. A personal discussion with the referral source prior to the first interview is helpful to add additional clarity to the purpose of the referral and to obtain nuanced background information that may not appear in the documentary history. In consultation/liaison settings, clarifying the nature and scope of referral questions allows the consultant to tailor an evaluation that is responsive to these questions.

The patient's first impressions influence attitudes and expectations. Training office staff that interact with patients before the initial interview to appreciate and recognize the anxieties that patients bring to the first encounter and offering them strategies to help patients feel as much at ease as possible is important to successful engagement.

Self-report tools may be introduced during the pre-encounter waiting

time to aid in gathering history and make the subsequent interview more efficient. Patients are increasingly familiar with the use of such forms in medical settings. Broadband screening tools, like the Brief Patient Health Questionnaire, gather information about depressive and anxiety symptoms, social distress, traumatic exposure and for female patients, screen for the possibility of pregnancy and other reproductive health issues. Disorder-based screening tools, like the Patient Health Questionnaire-9 (PHQ-9), aid in identifying patients likely to have a specific diagnosis, in this case major depressive disorder. Others, such as the Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q), are designed to survey social and emotional functioning.

While this method of data gathering has many advantages, there are potential unintended consequences to consider when offering screening and rating tools prior to the first encounter with the clinician. The patient may inadvertently develop the impression that short, closed-ended responses are desired, that the use of a particular tool means the clinician already has a preconceived notion of the diagnosis and that treatment will have a formulaic outline. In addition, any screening tool will produce false-positive and false-negative identifications depending on its sensitivity and specificity. The clinician who uses screening tools must be prepared to take steps to define the true significance of the screening result.

Setting the Stage for the First Encounter

In contemporary behavioral health practice, the psychiatrist interacts with patients in diverse settings that may be variably suited to the psychiatric interview. With the advent of telemedicine, the doctor and patient may not even be in the same location. No matter the setting certain principles for staging a successful first interview apply including attention to optimizing comfort, safety, and privacy. Most patients come to their first encounter with a mental health professional experiencing some level of trepidation. Some arrive cajoled or forced by others to have come at all. Anxieties are reduced by early clarification of the nature and purpose of the encounter and defining a mutual agenda before embarking on the interview. To the extent feasible, the patient should experience the clinician focused on him or her for the time they have together. Toward this end, unnecessary distractions such as beeper and text alerts and phone calls ought to be minimized. Potential distractions related to note taking and the electronic medical record are discussed in the *Special Topics* section of this chapter.

Comfort and Safety

For most patients adding to the emotional distress that brings a person to the initial interview is considerable apprehension about seeing a mental health professional. Often, inadequate knowledge about mental health care

and misconceptions about psychiatric treatment arising from stigma, misinformation, and misconceptions fuel these anxieties. The skilled clinician is aware of these potential issues and interacts in a manner to decrease, or at least not increase, the distress. Maya Angelou said, "I've learned that people will forget what you said, people will forget what you did, but people will never forget how you made them feel." The psychiatrist is advised to remember this adage when setting the stage for the initial interview. The patient's experience of the first encounter is bound to affect engagement and set the tone for subsequent interviews or treatment that may follow and may affect treatment attitudes far into the future.

Not much is accomplished in any human interaction when the participants are physically uncomfortable and/or excessively anxious about personal and emotional safety. The psychiatrist is in control of the environment in some settings, like a private office, but not so much in others, like the lock-up of a county jail. Regardless of where the initial contact occurs, interviewers should do what is within their power to adjust the environment to provide a reasonable level of comfort and assure the safety of both the patient and the clinician.

Engaging with patients known or found to have active psychosis, acute intoxication, or a history of aggression requires prudence and planning. Areas used for interviewing optimally have adequate interpersonal space and arrangements that allow both patient and interviewer an easy exit without "going through" one another's space. It is wise not to interview patients like this in isolated areas. Tenuous situations may require modifications to usual interview techniques including having another staff member present in the room during an encounter or standing by outside the door.

Explicitly discussing safety can reduce uncertainty, tension, and risk. The overtly disturbed patient is often calmed by acknowledgement of their discomfort and reassurance that the first goal is for both the patient and staff to remain safe. Sometimes pointing out the clinician's own anxieties about safety can level tension. A statement like this may help: "When you swear and speak in such a loud voice it makes me anxious too. When I am very anxious, it is hard for me to pay attention to what you are saying and interferes with my ability to help you. Do you think you can tone the language down a bit so that I listen better?" "Is there anything that I can do to help you feel calmer?" Even in a properly staged and managed interview, it may be necessary to shorten or quickly terminate the encounter if the patient escalates or becomes explicitly threatening. Judicious use of a therapeutic "show of force" can actually reduce the patient's fears of loss of control by demonstrating that there are adequate staff present and a commitment to keep everyone safe.

Privacy and Confidentiality

Privacy and confidentiality are essential components of therapeutic relationships. Staging the interview to ensure that others do not overhear the content is an important consideration. In some settings, for example, a shared inpatient room or a cellblock in a jail, privacy may be difficult to establish. In less than optimal environments, judiciously steering the interview away from material that may over expose the patient is wise.

The level of confidentiality afforded the patient varies depending on the nature of the encounter making it important to define the boundaries of confidentiality at the onset so that patients can make informed decisions about the nature and extent of material they are comfortable sharing. The clinician should consider reviewing with the patient situations that may lead one to share information without expressed consent. Examples include contacting informants to evaluate concern about imminent suicide or violence risk, transfer of information in medical or psychiatric emergencies and mandated reporting of child abuse or neglect. Insurance claims and communications with pharmacies routinely require sharing of information. Minor children need to understand that a parent or guardian must be informed about diagnosis and treatment options to provide surrogate consent. Detailing the boundaries of confidentiality and exceptions to it in writing may strengthen trust and confidence in the clinical engagement.

Specific consent is first obtained (best in writing or at least carefully documented in the medical record) when information is shared outside of a genuine emergency or a mandated report. In the United States, the federal Health Insurance Portability and Accountability Act (HIPAA) rules define standards for proper release of protected health information and state-by-state regulations provide additional guidance for the management of protected health information. The details of this are beyond the scope of this chapter, but the clinician should understand and carefully follow appropriate procedure for release of information, understanding that HIPAA usually defers to the state regulation if it is more restrictive.

When it is necessary to share information without specific consent, the nature and extent of disclosure should be limited to that information essential to the task. This discretion conserves confidences to the extent possible even in situations when the clinician has a responsibility that may conflict with the patient's privacy.

Involvement of Third Parties

Often members of the patient's family, or other supports, attend a first encounter and may expect to be involved in the interview. These persons can be important sources of collateral information and ultimately everyone may agree they should be involved in treatment decisions and management. The patient and psychiatrist should discuss whether and how to involve significant others in clinical interviews and the nature and extent of information the patient wishes shared, assuming the patient is

competent to assert confidentiality and an overriding factor, such as suicide or violence risk, does not trump the patient's choice. Patients typically appreciate a principle of "no meeting about me without me." It is usually best to sequence the involvement of family or important others by first meeting privately with the patient and then, with the patient's agreement, bringing others into the encounter with the patient present.

Engagement and Relationship Building

The term "doctor–patient relationship" is a phrase used to describe that remarkable interpersonal connection that sets the stage for healing interventions and on which current culture confers special status and obligations. Recently, the term "patient–doctor relationship" was coined reversing the order of the parties to reinforce that the treatment should always be patient centered. Projecting a demeanor that is friendly and open, maintaining a respectful and nonjudgmental posture and demonstrating genuine interest in developing an empathic understanding of the patient's predicament are physician behaviors that build the type of rapport from which the patient–doctor relationship is forged. In the initial interview, successful development of a trusting relationship is what allows the patient to feel comfortable enough to share information that is often sensitive and quite private.

While the relationship between any one patient and their physician will vary depending on each of the respective personalities and experiences, as well as the setting and purpose of the encounter, there are general principles that, when followed, help to ensure that the relationship established is one from which successful treatment can occur. Jerome Frank (1974) examined therapeutic relationships from an anthropological perspective discovering common characteristics, regardless of culture, that confer on an individual the status of healer. In Western cultures, specialized training and experience are factors. In addition to training and expertise, physicians are expected to behave in ways that project caring, instill hope, and maintain focus on the patient's best interest.

In the clinical setting, rapport can be defined as the harmonious responsiveness of the physician and patient to one another. It is important that the patient experiences the evaluation as a joint effort and that the psychiatrist is truly interested in their story. Appreciating how and why the patient thinks and feels as she does requires an understanding of the patient's life experience and perspectives in some detail and depth. Early in an initial interview, the clinician is forming hypotheses about the patient's experience through what the patient says, what affects accompany the story and just as importantly, what is not said. As the interview progresses, the details of the patient's symptoms and story unfold, and patterns of thinking and behavior become evident, resulting in greater clarity about the patient's actual lived experience.

Empathy is the capacity to appreciate and understand the experience of the patient at an emotional level, achieved when the interviewer is able to imagine himself in the patient's position while at the same time maintaining objectivity. Objectivity differentiates empathy from sympathy. Sympathetic identification with the patient may distort and stunt the development of a mature therapeutic relationship. Consider the experience of interviewing a patient with borderline personality disorder. Sean Shea has noted, "One feels compelled to say a silent prayer for the poor patient with borderline features who meets a clinician who boldly proclaims, 'I can feel your pain.'" An awareness and understanding of the dynamics of the patient–physician relationship are important in helping the physician maintain objectivity while still transmitting concerned understanding and extending a helping hand. Empathic expressions ("*that must have been very difficult for you*" or "*I'm beginning to understand how awful that felt*") serve to increase rapport. Frequently a nonverbal response (raised eyebrows or leaning toward the patient) or a very brief response ("wow") will be similarly effective.

Ultimately, it must become clear to the patient that the doctor both understands and cares about his experience for effective treatment to proceed. In 1927, Francis Peabody wrote, "The secret of care of the patient is caring for the patient." Caring for the patient is not automatic and at times, it is frankly difficult. James Groves, in 1978, published a classic paper entitled, "Taking Care of the Hateful Patient." Rael Strous and colleagues updated the paper for contemporary practice. Groves described patterns of patient behaviors that made them unlikeable to physicians and warned how this disdain could interfere with optimal medical care, noting that if physicians become more aware of their own feelings and reactions, they have increased opportunities to interact in a manner that can be helpful even in difficult situations.

The development of a healthy patient–physician relationship is reinforced when the physician is perceived as genuine. Being comfortable enough to laugh in response to a humorous comment, admit a mistake, or apologize for an error that inconvenienced the patient, like being late for or missing an appointment, strengthens the therapeutic alliance. It is also important to be flexible in the interview and responsive to patient initiatives in relationship building. If the patient brings in an item, for example, a photo she wants to show the psychiatrist, it is good to look at it, ask questions, and thank the patient for sharing it. Much can be learned about the family history and dynamics from such a seemingly sidebar moment. The psychiatrist should be mindful of the reality that there are no irrelevant moments in the interview room.

At times patients will ask questions about the psychiatrist that range from inquires about training and experience to more personal matters such as marital status, sexual orientation, or religious beliefs. Depending on the

reason for the interview, the nature of the question and the particular characteristics of the setting and patient, these questions may properly be answered directly or reflected back to the patient to gain insight into the motive for the inquiry. Avoiding excessive sharing of personal information is a good general principle to adhere to, especially if the interview is likely to proceed into a psychotherapeutic treatment that relies on the relative anonymity of the therapist as central tenet, such as in the case of psychodynamic therapy. On the other hand, the psychiatrist should not be excessively stingy with information for reasons that are more self-serving than patient-centered at risk of appearing aloof or callous.

A clinical example may help to illustrate these principles. The inpatient psychiatrist is rounding on a newly admitted patient with major depressive disorder for the first time and the following conversation initiated their contact:

Psychiatrist: *Good morning, Mr. Smith. I'm Dr. Jones.*

Patient: (Interrupting) *Before you get started I have one question for you. Are you Catholic?*

Psychiatrist: *Wow. That's quite a start. May I sit down?*

Patient: *You can sit down but I'm not going to answer any questions until you tell me if you are Catholic.*

Psychiatrist: *I can see that's very important to you. Can you tell me why that is important?*

Patient: *I'm not going to say anything till you tell me.*

Psychiatrist: *I can tell you something about who I am and why I'm seeing you this morning.*

Patient: *Never mind about that stuff. I want to know if you're Catholic.*

The psychiatrist presses on attempting to understand the reason for the question. Was there a prior unsatisfying experience with a Catholic or non-Catholic psychiatrist? Does the patient worry that his religious concerns will not be understood or incorporated into care? The same resistant responses occur with each attempt to reflect on or address Mr. Smith's feelings.

Psychiatrist: *I'd answer the question but I anticipate answering will lead to be more questions about me and we won't talk about you and why you were hospitalized, or how you feel about being here.*

Patient: *No that's the only question I have, if you answer that question I'll answer whatever questions you have.*

Psychiatrist: *Okay. I'm Catholic.*

Patient: *Do you go to mass every week?*

There are occasions when, depending on the nature of the planned or

ongoing treatment, it can be helpful for the psychiatrist to share some personal information even if the patient does not inquire directly. In all cases, the motive for self-revelation should be to strengthen the therapeutic alliance in the service of the patient not to meet needs of the doctor.

Meaning and Communication

An understanding of the process and content of the psychiatric interview would be incomplete without considering the role of the unconscious. From psychodynamic and neurobiological perspectives, the reality is that the majority of mental activity remains outside of conscious awareness. The psychiatrist and the patient both operate on conscious and unconscious levels and bring to the encounter their respective personalities, personal histories, and idiosyncrasies. The psychiatrist should remain alert for manifestations of unconscious process in communication and in relationship dynamics on the part of both herself and the patient.

Unconscious process in the patient–doctor relationship reveals itself through transference and countertransference manifestations and various psychological defense mechanisms. Transference is the process by which the patient unconsciously and often habitually displaces/transfers onto individuals in his current life those patterns of behavior and emotional reactions that originated with significant figures from earlier in life, often from childhood, like those attached to a parent or authority figure. Similarly, countertransference is the process whereby the physician unconsciously and reciprocally is involved in displacement to the patient. Remaining mindful that such distortions drive affects and behaviors of the patient and the psychiatrist alike reduces the burden of unnecessary emotional distress, unhelpful behaviors, and missed therapeutic opportunities that may occur when these processes are not recognized. An example may help to illustrate these concepts:

The primary care physician (PCP) offered considerable detail about an elderly woman's background and symptoms in a letter sent to initiate a referral for psychiatric evaluation. Early in the psychiatrist's first interview it became apparent that the patient's major concern was the sexual relationship with her husband which the psychiatrist found puzzling because the PCP was psychologically minded and had mentioned nothing about this in the letter. When asked if she had discussed this issue with her PCP the patient responded, "Well, it took me awhile to even broach the topic with him because he is young enough to be my son. I tried to bring it up a couple of times, but he looked uncomfortable, and he changed the subject. I didn't bring it up again; I didn't want to bother him." Subsequently, in a conversation about this patient, the PCP spontaneously commented, "She's a very nice lady; she reminds me a lot of my mother."

Other conscious and unconscious processes represented in the so-called "defense mechanisms" of psychodynamic theory shape and sometimes

interfere with interviewing. The patient is generally unaware of the impact of these feelings and behaviors, which take many different forms including tangential references to an issue, slips of the tongue, mannerisms of speech, avoidance of topics, exaggerated emotional responses, intellectualization, generalization, missed appointments, or other acting out behaviors; each of these may be examples of the unconscious at work. For example, phrases such as "to be honest with you" or "to speak frankly" suggest that the speaker is perhaps not actually being honest or speaking frankly. Interpreting the meaning of statements like these when they occur in initial interviews is tempting but generally ill advised. It is better to note such manifestations while refraining from premature exploration of defenses until the relationship is secure enough to tolerate this manner of intervention and the patient seems ready to receive it as helpful, though a reflection such as "I would hope you will always be honest with me," may be appropriate and constructive.

Open- versus Closed-Ended Questions

An open-ended question identifies and introduces an area of inquiry, but provides minimal structure about how to respond to the question or the desirable response. This style of questioning, in theory, allows the subject to respond spontaneously with as much detail as the subject chooses and minimizes the potential for influence of the questioner on the answer. An example of an open-ended question is, "What can you tell me about this nervousness you mentioned earlier?" Closed-ended questions invite brief, "Yes" and "No" types of responses and may telegraph the desirable response. An example of a closed-ended question: "You are not nervous all the time are you?"

The flow of a typical interview moves from open-ended questions that probe a particular topic to closed-ended questions that provide definition and detail when that is required. Both types of questioning prove important to full exploration of topics and each approach has its relative strengths and weaknesses. The balance of open- versus closed-ended questions will vary depending on the patient's capacity to grasp the invitation of open-ended questions, the nature of the material explored, the patient's response style (e.g., patient's prone to circumstantial responses may have to be structured more than others) the person's openness to disclosure and the purpose and time available for the interview. No prescription defines the proper balance.

Person-Centered and Disorder-Based

The initial psychiatric interview has at least three important goals: establish a traditional bio-medical diagnosis, understand the biopsychosocial dimensions of the person and build rapport. Clarifying the

agendas of the clinician and the patient early in the encounter helps to define the boundaries of the interview and plan the best use of time. Starting an interview with a statement like this can help initiate agenda setting: "We have all or part of an hour together today. My agenda is to understand the problem or problems that bring you to this visit. I also want to learn background about your life story to understand you as a person. Before we begin, I would like to know what you hope to get out of our meeting today and what questions you have so that we also address your expectations."

The individuality of the patient's experience is a central theme. The interviewer elicits the patient's life history in more or less detail subject to the constraints of time and the patient's willingness to share details of experiences and their private mental life. Adolf Meyer's "life-charts" were graphic representations of the material collected in this endeavor and were a core component of the "psychobiological" understanding of illness. The patient's early life experiences, cultural background, family dynamics, education, work history, religious beliefs and practices, hobbies, talents, relationships, traumatic events, and losses are each areas that, in concert with genetic and biological factors, represent the field in which distress arose. An appreciation of the richness of life experiences and their impact on the person is necessary to develop a complex understanding of the patient. Gleaning all of this information in a first encounter may not be practical, but gathering enough of it to grasp the essence of both the problem and the person is an important goal.

Traditionally, medicine has focused on illness and deficits rather than strengths and assets. A person-centered approach focuses on strengths and assets as well as deficits. During the assessment, it can be helpful to ask the patient, "*Tell me about some of the things you do best*," or, "*What do you consider your greatest asset?*" A more open-ended question such as, "*Tell me about yourself*," may elicit information that focuses more on either strengths or deficits depending on a number of factors including the patient's mood and self-image.

In addition to being person-centered, it is also important that the psychiatric interview be disorder-based. Great advances in medicine have occurred over the past two centuries as psychiatrists began to recognize clusters of signs and symptoms were characteristic of disorders that have common epidemiology, features, presentations, course, and response to treatment. Developing a criteria-based nomenclature, as catalogued in the *Diagnostic and Statistical Manual of Mental Disorders,* now in its fifth edition (DSM-5), has furthered psychiatric diagnosis and effective treatment. The DSM has some limitations of course, but it transformed the esoteric methods of diagnosis that preceded it into a standardized method that improved the reliability of diagnosis and aided research as well as clinical work. Evidence-based psychiatric practice has flourished in part

because of the increased reliability of diagnoses.

A best practice interview integrates the person-centered and the disorder-based approaches, and the process of the interview weaves back and forth between the two ending in a working diagnosis, differential diagnosis and treatment plan. It is essential to incorporate the patient's needs and goals in the resulting treatment plan. Numerous studies have demonstrated that too often the patient's goals for treatment (e.g., safe housing) are not the same as the psychiatrist's (e.g., decrease in hallucinations). This dichotomy may have its origin in an initial interview wherein the focus was not sufficiently person-centered but rather was exclusively or largely symptom-based. Even when the interviewer specifically asks about the patient's goals and aspirations, the patient, likely exposed on numerous occasions to what a medical professional is interested in hearing about, may attempt to focus on "acceptable" or "expected" goals rather than her own. Clinicians should explicitly encourage patients to assert their goals and aspirations and to express them in their own words.

Time and Number of Sessions

The time devoted to the initial interview will vary considerably by the goal of the encounter and the setting. In a typical outpatient treatment environment, about an hour is generally allotted. In settings like emergency and hospital inpatient units where patients are often distressed, confused, actively psychotic or agitated, the time may be foreshortened and assessment may by default occur over multiple, brief encounters, or rely heavily on collateral historians.

The clinician must accept the reality that any history obtained is never complete or fully accurate because of the nature of human memory and the distortions of perception that each person involved in an experience makes when recounting it and variations that occur with each retelling. An interview is dynamic and some aspects of the evaluation are ongoing—new facts emerge so that the history often morphs in lesser or greater ways depending on various factors including the patient's gradual comfort with revealing and the dynamics of the relationship between the patient and doctor.

HISTORY AND EXAMINATION

Process of the Interview

The key elements of the psychiatric interview are the patient's histories and the mental status examination (MSE). The histories are based primarily on the subjective report of the patient but importantly should be enriched by and compared to available records and the reports of collateral historians. It is important to document the sources of information contributing to the

evaluation with commentary on the quality and reliability of the sources. Table 7.1–1 lists the traditional elements of the psychiatric history and examination in the conventional sequence that findings are typically documented. The following sections will further define and discuss these elements.

The MSE is in essence the "physical examination" of psychiatry; a series of observations of signs and symptoms that are gathered and documented systematically. This chapter's section, *Physical Assessment,* provides an overview of physical assessment in the context of psychiatric evaluation. Together the histories, MSE and physical assessment contribute to defining a working diagnosis.

Opening the Interview

After introductions (including clarifying how the patient prefers to be addressed), explaining the practice's processes, answering initial questions and agenda setting, the interview itself begins.

Table 7.1–1.
Sections of the Psychiatric History and Examination

- Identifying data
- Sources of information and reliability
- Chief complaint
- History of present illness
- Past psychiatric history (including suicide and violence risk)
- Substance use/abuse/addictive behaviors
- Past medical history
- Family history
- Developmental and social history
- Review of systems
- Mental status examination
- Physical assessment
- Formulation
- DSM-5 diagnosis
- Treatment plan

The psychiatrist can begin with an open-ended inquiry, "*Why don't we start by you telling me your understanding of what has led to your being here today.*" Often, the patient is uncertain as to why referral to a psychiatrist was necessary or may feel frustrated or angered by the need to see a psychiatrist. The following kind of conversation is not uncommon:

Mr. Smith: *Dr. Jones said I should see you.*

Doctor: *What's your understanding of why Dr. Jones recommended you see me?*

Mr. Smith: *Well, you're a psychiatrist, right?*

Doctor: *Yes I am.*

Mr. Smith: *Well, after my tests all came back negative, he kind of implied this nervousness —that it's all in my head and I should see you. I think he's pretty frustrated with me.*

The patient should have an opportunity to express his feelings about the need for the encounter, and his anxieties about it. The psychiatrist has an opportunity to inform about the role of the psychiatrist as a medical specialist and the nature and importance of collaborative relationships, such as those with primary care providers or other mental health clinicians in the care of the patient.

Chief Complaint

Traditionally, the chief complaint is recorded in the patient's own words, for example, "I have been depressed for months" or, "I have a lot of anxiety in public speaking." The following example illustrates the importance of recording the chief complaint as the patient offered it:

> An elderly man presented in a psychiatric emergency room with the chief complaint, "I'm melting away like a snowball." He had become increasingly depressed over the prior 3 months. Four weeks before this emergency visit, the PCP increased imipramine (an antidepressant) from 25 to 75 mg, and added hydrochlorothiazide (a diuretic) 50 mg because of mild hypertension and slight pedal edema. Over the ensuing month, the patient's condition deteriorated. In the emergency room, he was noted to have depressed mood, a hopelessness attitude, weakness, significant weight loss, psychomotor retardation, and was described as appearing "depleted." He also appeared dehydrated. Blood work found elevated BUN and creatinine levels and hypokalemia. Examination of his medication supplies revealed that the medications were dispensed in the wrong bottles such that the patient was actually taking 25 mg of imipramine, generally a nontherapeutic dose for depression, and 150 mg of hydrochlorothiazide, generally a supratherapeutic dose for diuresis. He was indeed "melting away like a snowball" from dehydration. Fluid and potassium replacement resulted in immediate improvement in his general medical status and allowed the antidepressant to be safely titrated toward a therapeutic dose to treat his depression.

Having establishing the chief complaint, the interview proceeds on into the history of present illness.

History of Present Illness

The history of present illness (HPI) is a chronological description of the evolution of the symptoms of the current episode. As with any well-developed HPI in medicine, the psychiatric history includes details about the nature, frequency, severity, and pattern of symptoms including exacerbating and alleviating factors. Stressful life events, physical symptoms, medication changes, and substance use patterns may provide clues to the context of the presentation, the primary diagnosis and comorbid conditions that can influence treatment choices.

The exploration begins, ideally, with an open-ended invitation like, "Please tell me how this problem began, what the experience has been like and how it has played out over time." The interviewer may have to structure the patient through parts of the story. It is important to try to answer to the question, "Why is the patient presenting now?" (e.g., *"I'm*

here now because my girlfriend told me if I don't get help with this nervousness, she will leave me.")

Treatments the patient utilized during the current episode are reviewed with attention to response, adherence, and attitudes about the treatment. The common use of alternative and over-the-counter treatments makes it important to inquire specifically about the use of these products.

Suicide and violence risk assessment (sometimes termed "lethality" assessment) is a standard element of the HPI and any complete psychiatric evaluation. Patients may not spontaneously announce dangerous ideation or acts for a host of reasons including shame and fear that revealing may lead to undesirable consequences including reporting and confinement. The *Suicide and Violence Risk History* section of this chapter provides additional information about risk assessment.

Rounding out a comprehensive HPI is a review of psychiatric and physical symptoms to ensure that important symptoms of the primary problem or comorbid disorders relevant to comprehensive diagnosis, treatment planning, and prognosis are not missed. Two examples may illustrate the importance of a review of symptoms. Failing to conduct a review of psychotic symptoms with the depressed patient could lead to a missed diagnosis of psychotic depression, the treatment of which varies importantly from nonpsychotic depression. Failing to conduct a medical review of symptoms, the clinician could easily miss the fact that a patient with depression and congestive heart failure (CHF) developed shortness of breath after the PCP initiated a tricyclic antidepressant (TCA). Not recognizing that the TCA was exacerbating the CHF, the psychiatrist missed the opportunity to recommend a therapeutic alternative for depression that would steer way from inotropic side effects. Table 7.1–2 offers an outline of major categories of psychiatric symptoms, including mood, anxiety, psychotic symptoms, that are part of a psychiatric review of systems. The *Review of Systems* section of this chapter and figures that accompany it offer more information about the physical review of symptoms.

Past Psychiatric History

The past psychiatric history explores psychiatric illness prior to the current presentation including the nature of symptoms, course, and treatment. Table 7.1–3 summarizes the items generally covered in a complete past history review. Details of past episodes including age of onset, context, nature and duration of episodes, the diagnosis offered, treatment applied and its setting, degree of response, treatment adherence, and attitudes toward treatment, are all important facts to gather. Understanding the details of past treatment will lead to a better understanding of which treatments are viable alternatives and which to avoid.

Most psychiatric disorders have a pattern of recurrence, and

comorbidity is the rule rather than the exception. Some episodes of illness may have gone undiagnosed, especially milder forms of illness. Establishing the patient's best functional baseline and the extent of recovery between episodes provides a perspective on the impact of the problem(s) on the trajectory of the patient's life and helps to identify elements of a comprehensive treatment plan that optimizes fullest possible recovery.

Table 7.1–2.
Psychiatric Review of Systems

Sleep
Sleep phase problems (initial, middle, terminal insomnia), total sleep time, abnormal sleep events
Mood
Depression: persistent sadness, reduced interest or pleasure in usual activities, tearfulness, reduced or excessive sleep, reduced or increased appetite, weight loss or gain, low energy, reduced concentration, low libido, excessive or inappropriate guilt, psychomotor change (slowing or agitation), negative self-appraisal, helpless and hopeless thinking thoughts of death or suicide. A common mnemonic used to remember the symptoms of major depression is SIGECAPS (Sleep, Interest, Guilt, Energy, Concentration, Appetite, Psychomotor agitation or slowing, Suicidality).
Hypomania/Mania: elevated, expansive or irritable mood, decreased need for or inability to sleep, excessive energy, marked increase in goal and pleasure directed activity, increase amount and pace of speech and thought, grandiosity, heightened libido, impulsivity and/or recklessness in behaviors such as spending and sex
Anxiety
Experience of panic attacks, somatic symptoms of anxiety, phobic, or social avoidance
Psychosis
Experience of hallucinations, delusions, disorganized behavior, speech or thought, negative symptoms
Obsessive-Compulsive
Repetitive intrusive and unwanted thoughts, compulsive behaviors to neutralize anxiety, hoarding behaviors
Trauma
Traumatic exposure; intrusive and avoidance symptoms, negative alterations in cognitions and mood, excessive arousal and reactivity
Behavioral
Substance use, gambling, impulse control problems, disordered eating, repetitive self-harm

Fastidious review of prior medication trials provides valuable information to define reasonable treatment medication options and to rule out others. Important elements of this review include the dose and duration of trials, the degree of nature and degree of effect and side effects and whether the medication was used alone or in combination with other psychotropics or psychotherapies. Exploring adherence may lead to discovery of obstacles such as negative attitudes, side effects, dosing complexities, stigma, social pressures, and financial burdens that affected prior medication trials. Normalizing problems with medication adherence helps to obtain frank responses. Consider these types of questions to invite the patient to acknowledge challenges with adherence: "Most people miss

some doses of medication. How many times a week do you miss your medication doses?" "What gets in the way of taking the medication regularly?"

Table 7.1–3.
Past Psychiatric History

1. Details of prior mental health contacts
   - ► Sites/Levels of care: school-based, primary care-based, outpatient, emergency, inpatient, partial hospital, recovery programs
   - ► Nature of prior treatment: psychotherapy, medications, technologies (e.g., ECT), recovery-based, alternative and complimentary
   - ► Adherence and response to treatment
2. Prior acts of self-harm, suicide, aggression, and violence

It is similarly important to gather details about psychotherapy trials including the type of psychotherapy (individual, group, couple, or family), the model of psychotherapy (Was it evidence-based for the diagnosis?) and the frequency and duration of the treatment (Was it delivered in a manner that one expects to be effective?).

To ensure the history is complete, the clinician should inquire about less common technological modalities that the patient may have used including light therapy, biofeedback, electroconvulsive therapy, and vagal nerve and transcranial magnetic stimulation.

Suicide and Violence Risk History

Developing the risk history is a critical component of a complete past psychiatric history. The field of suicide and violence risk assessment has matured to offer an understanding that risk is a dynamic process influenced importantly by risk-enhancing and risk-reducing factors with origins in genetics, social learning, current environments, and the particular nature of present signs and symptoms. A full discussion of risk assessment is beyond the scope of this chapter, but it is important to point out key elements of a risk assessment history. These include details of past events of intentional self-harm and aggression including the number of events, details of the context of prior dangerous behaviors, the level of intent and seriousness of behaviors, harmful outcomes, and the means used to harm self or others. A full exploration includes nonlethal behaviors that increase the risk of future lethality including intentional self-injury. In addition to behavior itself, feelings and attitudes about the behavior and emotional states that accompany or follow the behavior should also be explored, as well as the degree to which the patient has tried to conceal evidence of these behaviors. The degree of family and/or community support and religious beliefs that admonish against suicide are potential risk-reducing factors that should be assessed. Given the fact that firearms

are the leading method of suicide in men and the second leading method for women, screening for firearms access has become a routine in psychiatric evaluation practice.

Additional information about risk assessment is found in other sections of this chapter: the *Mental Status Examination* section under *Risk Assessment,* and in the *Selected Challenging Interview Scenarios* under the headings *The Patient with Depression, The Potentially Suicidal Patient* and *The Potentially Violent Patient*.

Substance Use, Abuse, and Addictive Behaviors

Substance use may be the primary reason for a referral, and often is an important secondary problem. Substance use disorders can mimic or induce psychiatric syndromes, elevate risk of suicide and violence, and have important impact on safe medication prescribing. These facts highlight the importance of understanding the patient's current and past patterns of substance use.

Starting the inquiry about substance use with a normalizing statement can help to reduce defensiveness and elicit a fuller history: "Most people drink alcohol at times and many have also used a recreational drug at some point in their life. I want to ask you some questions about your past and current use of alcohol and recreational drugs."

Various tools can be used to aid in gathering the substance use history. Examples include the commonly used CAGE questionnaire which has been modified to include other drugs (and now called CAGE-AID) though studies have found the operating characteristics of this tool wanting in primary care settings, Caucasian females, pregnant women, college students, and less severe levels of drinking. The National Institute on Alcohol Abuse and Alcoholism (NIAAA) recommends a procedure for alcohol use inquiry (available at http://www.niaaa.nih.gov/guide) that offers opportunity for education about safe levels of drinking for those patients at risk in addition to screening for current alcohol use problems using the Alcohol Use Disorders Identification Test (AUDIT).

For each substance the patient acknowledges using, the interviewer gathers basic information about age of onset, pattern of use over time, current frequency, and level of use and consequences of use (physical, mental, social, and legal). These facts help to identify problematic current or past patterns of use. Inquiry about tobacco use and problem gambling should be a part of the addictive behavior inventory.

The interviewer should identify periods of abstinence and determine what helped the patient achieve control of use including any specialized treatment for substance use with details about the setting, nature of and response to treatment, as well as engagements with peer support engagement such as Alcoholics and Narcotics Anonymous.

Past Medical History

Past medical history is a bit of a misnomer given that this section of history taking and documentation is used to review both past and present medical history. The interviewer is interested in obtaining an accounting of major medical disorders both to develop a complete history and to identify illness that could mimic a psychiatric disorder, contribute to the context of the presentation or factor into treatment planning.

Medical disorders can precipitate a psychiatric disorder (e.g., anxiety disorder in an individual recently diagnosed with cancer), mimic a psychiatric disorder (e.g., hypothyroidism presenting "as if" major depression), be precipitated by treatment of a psychiatric disorder (e.g., metabolic syndrome emerging during exposure to a second-generation antipsychotic medication), or influence the choice of treatment of a psychiatric disorder (e.g., renal insufficiency and the use of lithium carbonate).

A well-developed past medical history archives both current and past major medical disorders, surgeries, hospitalizations and significant physical trauma, such as head injuries. In psychiatry, neurological and endocrine disorders are of particular interest because of the significant overlap in symptoms and signs with psychiatric syndromes. For female patients, obtaining a reproductive and menstrual history is important, as well as a careful assessment of potential for current pregnancy and plans for future pregnancy.

A thorough review of all current medications is essential because some medications used in other fields in medicine can create side effects that mimic psychiatric disorders and to identify potential drug–drug interactions before prescribing. Patients have access to all variety of over-the-counter, herbal and alternative medications. They may not mention the use of these products unless specifically asked. Many over-the-counter and herbal/alternative products can produce psychiatric symptoms or interact with psychotropic medications in ways that range from nuisance side effects to outright hazards. Frank Ayd published results of the so-called "Brown Bag Study" that illustrated just how revealing an inspection of all the medications and products that a patient takes can be, and that serves as a reminder of the value of having the patient actually bring to the appointment their "brown bag" of bottles for inspection.

Allergies to medications must be identified including the nature of the allergic response.

Family History

Because many psychiatric illnesses have a genetic predisposition, if not cause, a careful review of family history is important to the assessment and can aid in diagnosis and establishing expected prognosis. This area of

history identifies family members with histories of known or suspected mental illness, substance use problems, and other behavioral problems, for example, criminality. Patients may not know the diagnosis of a family member, but may be able to give an accounting of the relative's signs and symptoms from which reasonable hypotheses regarding diagnosis can be made. When the patient offers a diagnosis of a family member without much information to assess the reliability of the claim, the alleged diagnosis should be held skeptically or some attempt made to clarify it. Specific inquiry about family history of completed suicide is important to suicide risk assessment because this finding elevates the index patient's future risk. Family history of medical disorders may offer clues to the patient's risk factors. For example, diabetes in first-degree relatives ought to raise concern about the potential for metabolic syndrome with exposure to second-generation antipsychotics.

Developmental and Social History

Factors such as the reason for the interview and the amount of time available for it will determine how broad and deep the inquiries about developmental and social history can be, though it is important in first encounters for the clinician to give patients a sense that one is interested in their personhood and the contexts of their distress, not just their symptoms. Often in the first interview an outline is obtained that is further developed in subsequent encounters. Table 7.1–4 provides an outline of topics for the developmental and social history.

The developmental and social history reviews the stages of the patient's life from gestation to the present with an eye toward understanding the important exposures, relationships, and events that shaped the person's life story. One is interested in understanding the nature of the person's temperament and character and the degree to which the person has achieved developmentally appropriate role functions such as academic progress, work, peer and romantic relationships, and parenting capacity. Gestational and birth history, developmental milestones and early childhood development, family of origin, cultural identifications, educational, occupational, legal and military histories are all areas to be explored. Histories of abuse (emotional, physical, and sexual), neglect (emotional and physical) and specific traumatic exposures are important areas for specific inquiry. The nature of the patient's current social environment is defined including financial status, housing, and current relationships.

It is often helpful to review the social history chronologically; doing so provides a natural flow to the questions and ensures a complete history. Depending on the interviewer's assessment of how important missing information is and the assessment of credibility of the patient's account, it may be necessary to consult collateral historians to obtain a clear

developmental and social history.

Table 7.1–4.
Social and Developmental History

1. Pregnancy and delivery
2. Developmental milestones
3. Educational (including history of special needs, in-school counseling, disciplinary problems)
4. Occupational history
5. Military
6. Legal
7. Relationships/committed relationships
8. Parenting
9. Leisure
10. Traumatic or potentially traumatic exposures (including neglect, physical and sexual abuse)

Review of Systems

As in a general medical interview, the review of systems is intended to capture any current physical signs and symptoms not already identified in the HPI or past medical history (including Table 7.1–2 and is organized by asking sentinel questions about the major systems of the body).

Mental Status Examination

The MSE is the functional equivalent of the physical examination in other areas of medicine. It is a systematic collection of the observations (e.g., signs such as blunt affect or rapid speech) and reported mental experiences (e.g., symptoms such as depressed mood or hallucinations) that produce a picture of the patient's current mental state. The interviewer makes these observations throughout an encounter and ultimately documents the findings together in the MSE section of the evaluation document.

Table 7.1–5 lists the traditional elements of the MSE along with examples of potential abnormal findings. The sections that follow offer commentary about each item of the MSE and model statements to illustrate how these findings might be documented in a medical record.

Appearance and Behavior

This section of the examination archives findings related to the patient's levels of cooperativeness and behavioral activation, manner of relatedness and interaction, and state of grooming and dress.

> The patient appeared her stated age, was dressed eccentrically and was poorly groomed. She seemed tense, was distracted and difficult to structure, but was otherwise cooperative with the interview.

Table 7.1–5.
Elements of the Mental Status Examination with Examples of Abnormal Findings

Appearance and behavior: odd or eccentric dress, poorly groomed, disheveled, unkempt, difficult to engage, evasive, indifferent, defensive, seductive, hostile
Eye contact: avoidant, excessively prolonged, intense
Motor activity: reduced activity, restlessness, agitation, posture and gait abnormalities, praxis problems, mannerisms and other stereotyped behaviors, posturing, tics, tremor, choreic, athetoid, and dyskinetic movements
Mood: sad, irritable, angry depressed, elevated, expansive, euphoric, elated, dysphoric, anxious, incongruous with affect
Affect: restricted, constricted, blunt, flat, labile, inappropriate to mood
Speech: nonspontaneous, mute, rapid or slowed, low or high volume, abnormal pitch, tone or tempo, monotonous, robotic, stilted, excessive or reduced production, paraphasia, word approximations, poverty of content, dysarthria
Thought process/form: derailment, tangential, circumstantial, flight of ideas, incoherence, blocking, perseveration, clanging
Thought content: obsession, delusion, magical thinking, overvalued ideas, ideas of reference or influence, persecutory ideas
Perception: hallucination, misperceptions (e.g., illusion), depersonalization, derealization, jamais vu, déjà vu
Risk assessment: suicidal or violent ideation, intent, plan
Cognition: impairments of orientation, attention, memory, language functions, calculation, visual spatial capacity, following commands, impairment of intelligence, impairment of abstraction
Insight: impairment along a continuum
Judgment: impairment along a continuum

Eye Contact

Humans are keenly aware of eye contact and deduce a good deal of about another's state of mind from it. Eye contact and its modulation in the psychiatric interview offer important clues to the subject's internal emotional state. For example, a depressed patient may make little or no eye contact, an anxious patient's eye contact may be intermittent whereas a patient with psychosis may stare and maintain intense eye contact. Eye contact is normally regulated to modulate intensity and to prevent inadvertent dominance displays. If a patient is especially intense or overtly paranoid, it is useful to avoid prolonged eye contact.

> Eye contact was intermittent. When she made eye contact it was inappropriately prolonged and intense.

Motor Activity

This section of the examination offers observations about the patient's general level of motor activity, posture and gait, and identifies the presence of abnormal movements such as tics, mannerisms, posturing, choreic, athetoid, and dyskinetic movements. If screening neurological examination

is conducted, for example, testing of muscle tone and deep tendon reflexes, these findings are archived here. Many clinicians use this space to document the Abnormal Involuntary Movement Scale (AIMS) examination for tardive dyskinesia.

> Her gait was slowed and unsteady. When sitting, she was restless and rubbed the top of her head repetitively. On AIMS examination, she had mild rolling movements of her tongue at rest that were accentuated when her tongue was extended.

Mood

The terms mood and affect vary in their definition such that one term often merges with the other. A number of authors have recommended combining the two elements into a new label "emotional expression." Traditionally, mood is defined as the patient's internal and sustained emotional state. Its experience is subjective, and hence it is best to use the patient's own words when describing mood. Terms such as normal, calm, sad, happy, angry, irritable, and anxious are examples of descriptors patients may offer.

> She said her mood was "calm."

Affect

Affect differs from mood in that it is the outward manifestation of a mood state visible to others. Mood is deduced to some extent from the person's emotional display. Some terms used to describe the quality (or tone) of a patient's affect include dysphoric, happy, excited, euthymic, irritable, angry, tearful, blunt, and flat. Speech provides affective information as well through volume, tone, and word choices. Quantity of affect is a measure of its intensity. Two patients both described as having depressed affect can vary greatly when one is mildly depressed and the other is severely depressed. Terms such as restricted, constricted, normal, or labile describe the range of affect. Affect that correlates with the patient's stated mood and expressed thoughts is said to be appropriate or congruent to mood or thought content, whereas affect that contrasts sharply is said to be inappropriate with or incongruent to mood or thought content.

> Her affect was labile ranging from inconsolable tearfulness to wild displays of irritability accompanying hostile speech, and was incongruent with her claim that her mood state was "calm."

Speech

Speech has an expected range of rate and volume and elements of pitch,

tone, and tempo that is recognized within a normal range though there is cultural variation in these characteristics that must be accounted for. It normally flows in a relatively linear manner conveying information efficiently. Strong emotions, such as intense anxiety, can lead to speech abnormalities that are transient. Certain psychiatric conditions, such as mania and schizophrenia, often have characteristic speech abnormalities that affect both the linear flow of speech and its effectiveness in communication. The psychiatric interviewer should be alert for evidence of aphasic speech which can be mistaken for disordered speech due to a psychiatric condition. Patients with speech disorders are often unaware that the communicative value of their speech is failing and may not show reciprocity in speech that is characteristic of the interactive nature of spoken communication.

At times, she was mute. When she did speak, her speech was excessive, difficult to interrupt and came in rapid, loud bursts that made most of her communication ineffective.

Thought Process/Form

Thought process, also referred to as thought form, is the manner in which ideas are connected one to another in conveying thoughts to a listener, or in the case of written speech, a reader. Normally, ideas string together in a relatively linear pattern obeying conventions of grammar and syntax. Disordered thought takes many forms described by terms such as circumstantial (an inordinately circuitous route from question to answer), tangential (an answer to a question that veers off from the target of the question, but the connection may still be appreciated or inferred), and derailment, also called "loose associations" (thoughts proceed from idea to idea in a manner so oblique that the listener cannot follow the train of thought). Table 7.1–6 contains some examples of named thought disorders and their typical characteristics.

Table 7.1–6.
Examples of Disordered Thought Process/Form

Clanging: use of the sounds of words rather than their meaning or conventional application to determine the connection between ideas
Circumstantial: circuitous and overly detailed responses
Derailment (loose associations): convoluted and random connections between ideas such that communicative value is lost
Flight of ideas: derailment in the context of rapid speech
Incoherence (word salad): complete disregard for conventions of word usage, grammar, and syntax resulting in incoherence
Neologism: a word created by the speaker with idiosyncratic meaning
Tangential: responses that have only partial or remote connection to the original idea
Thought blocking: loss of the goal of a communication and not being able to return to the topic

Occasionally, her speech was linear and conveyed information well, but for the most part she was quite disorganized with tangential and circumstantial speech. She was hard to interrupt during periods of flight of ideas and became frustrated when the interviewer tried to clarify what she was trying to say. At one point, she used an apparent neologism. When asked how was transported to the hospital she said, "In a convestation, you fool!"

Thought Content

Thought content refers to the general themes of the patient's ideation and is an area of the examination that identifies thought abnormalities such as obsessional thoughts and delusional ideation. The extent to which the examiner gains access to the patient's private ideational life is determined by factors such as the patient's sense of comfort with the examiner and interviewing techniques aimed at eliciting this content.

Some patients share a broad range of content spontaneously and may even require structure to restrain the field of information they offer, while others perseverate or ruminate on specific thought content and require prodding to move beyond a narrow focus. Some patients guard content carefully. Guarding of thought content has many roots including embarrassment or shame about the ideation, as is often the case with obsessional thoughts and paranoia that spills over into distrust of the interviewer.

A patient may reveal abnormal thought content freely, for example, complaining vigorously about persecution as a chief concern. At other times, gaining access to abnormal ideation is a delicate dance, especially when the patient is anxious about the consequences of revealing such material. Identification and exploration of delusional ideation often requires particular care and respectful inquiry. A delusion is a false belief that the patient holds firmly despite evidence to the contrary and that is outside the normal range of beliefs of the patient's culture or subculture. Common types of delusions are listed in Table 7.1–7. Delusions may be divided into two main groupings. One grouping distinguishes whether the idea is bizarre or not and the other whether the idea is mood congruent or incongruent. Bizarre delusions are those that no one in the culture would consider a valid belief whereas nonbizarre delusions are plausible but unlikely to be true. Mood-congruent delusions have themes consistent with the patient's dominate mood state, for example, a depressed patient's certainty that he has committed a mortal sin, whereas mood-incongruent delusions have no referent to a mood state, for example, the belief that aliens have implanted a device in one's tooth through which they are controlling thoughts.

Table 7.1–7.
Examples of Delusional Thought

Bizarre: a delusional belief that members of one's culture consider impossible
Erotomanic: belief that another person is in love with one, usually someone perceived of as having higher status
Grandiose: belief that one has special powers, influence, or a special relationship with a deity or famous person
Identity: belief that the identity of oneself or another person (usually close to the individual has been altered, replaced or replicated in some way, e.g., Capgras syndrome)
Influence: belief that someone of something outside one's self is influencing one to think, feel, or act in a certain way
Persecution: belief that one is the subject of persecution, attack, harassment, or is being conspired against
Reference: belief that an event, person, or object that is unconnected to the person has some special meaning or significance to the person
Somatic: belief that a physical sensation or physical change is occurring to one's body caused by a source other than a normal biological or pathophysiological process
Thought broadcasting: belief that the individual has thoughts that are produced by someone or something outside oneself
Thought insertion: belief that the individual has thoughts that are produced by someone or something outside oneself

A delusion is a firmly held, false belief based on an incorrect inference that is unshakeable despite evidence to contradict it. It cannot be simply a cultural belief.

Gentle exploration of the degree of conviction about a delusional belief can test the patient's capacity to consider the possibility that their conviction is flawed. If one confronts the legitimacy of a delusional idea at all, this should only be done when there is a firm enough relationship for the patient to tolerate challenge and with respect for the fact that delusions often serve an explanatory role. It may be hard for the patient to consider the possibility that the idea misrepresents reality for doing so may require confronting an intolerable alternative truth. Consider the example of a young man with new onset of schizophrenia who is failing in his college coursework. He attributes academic problems to a conspiracy among peers and faculties whom he is sure are modifying his work to ensure failure because they are threatened by his intellect. To consider the possible truth that is there is something desperately wrong with his mind now interfering with his cognitive capacities and derailing his academic progress may be too much to bear.

When I inquired about his concerns about fellow students and his professors he said, "Certainly you can see as well as I do that the other students and faculty members are threatened by me. They are colluding to undo me to push me out of the way. No, I do not think this could be a trick of my mind and the fact that you would even bring that up makes me wonder if you too are a part of this effort to ruin my academic career." This appeared to be a delusion with persecutory and grandiose themes.

Perception

Perceptual abnormalities include phenomena like hallucinations, illusions,

depersonalization, and derealization. Table 7.1–8 offers names and definitions for some of the more common types of perceptual disturbances.

Hallucinations are perceptions in the absence of external stimuli to account for them; they may occur in any of the five senses (auditory, visual, gustatory, olfactory, and tactile) and seem as real in the person's experience as a true perception. Auditory hallucinations are the most common type of hallucination in nonorganic psychiatric conditions. In North American culture, the finding of nonauditory hallucinations should peak curiosity about an organic cause rather than mental disorder. While visual hallucinations do present (usually accompanying auditory hallucinations) in North American culture, some research in other cultures finds visual hallucinations to be the most common form of hallucinations in schizophrenia. To avoid over diagnosis and over treatment the interviewer should be careful to distinguish between a true hallucination and a misperception of actual sensory stimulus (i.e., an illusion) and other phenomena, like hypnagogic and hypnopompic hallucinations, that are sleep–wake phenomena and a sign of a sleep disorder not a psychotic illness. Most everyone can point to an example of a misperception or fleeting perception, for example, hearing one's voice called when no one else is home. These phenomena should not be mistaken for a hallucination, much less evidence of psychosis.

Table 7.1–8.
Examples of Perceptual Abnormalities

- Déjà vu: perception that a present circumstance is the duplicate of an experience that has occurred in the past
- Derealization: perception that one's surrounding and events are experienced as if the person is detached from them, or that they are distorted, changed, or unreal
- Depersonalization: perception that one is standing outside oneself as a detached observer to surroundings, experiences, and events that occur
- Hallucination: the experience of a perception that seems genuine yet occurs without an actual external stimulus. The perception may occur in any of the five senses: auditory, gustatory, olfactory, tactile/somatosensory, or visual
- Illusion: inaccurate perception or interpretation of an actual perception
- Jamais Vu: perception that a present experience in entirely foreign when it actually should be very familiar

When documenting perceptual abnormalities the interviewer should provide some description of the patient's actual experience, the context(s) in which the experience occurs, its frequency, and intensity, the degree of the patient's conviction of the reality of the perception, and the degree of discomfort it causes, as well as any steps the patient may have taken to alleviate adverse effects of the experience. In the case of auditory hallucinations, details of the nature and character of the voices can be important to diagnosis: Does the patient hear one or several voices, simple statements or complex sentences? Do the voices engage in a conversation

or comment on the patient's thoughts? Generally speaking, the more complex the hallucinations, the more likely they are due to a schizophrenic spectrum condition. Whether the voices ever urge or command the patient to actions, including suicidal or violent behavior, is important to risk assessment.

As with delusions, the character of hallucinations is often divided into those that are mood congruent (e.g., a depressed patient hearing a voice chiding her for failure and urging her to suicide) and mood incongruent (e.g., a patient with schizophrenia who despite being quite paranoid hears female voices making warm, sexual comments that he finds arousing and soothing).

> The patient is experiencing hallucinations. She said, "Yes, I do hear a voice that keeps repeating statements like 'you have failed as a mother, you will lose your family, you might as well be dead.'"

Risk Assessment

Suicidal, violent, and homicidal ideation fall under the category of thought content but many interviewers document these findings in a separate section of the examination labeled risk assessment. Risk assessment is an important aspect of any initial clinical psychiatric evaluation.

Developing depth of understanding about high-risk ideation requires something more than simply asking the subject "Are you suicidal or homicidal?" Ideas about suicide and violence are among the most private of thoughts. Patients are often hesitant to divulge ideation for fear of how the revelation will be received and what actions the clinician may take. A useful approach is to begin with a normalizing statement followed by a question. For example, "In my experience, people who are struggling with the kinds of emotions and life problems you have described to me are also having ideas about death, dying, or taking their own life. Has this been the case for you?" A complete inquiry into risk addresses not only the presence of ideation but also the level of intent and plans, and preparations that the patient has considered. Questioning about suicidal and violent ideation generally flows best by moving along the continuum from ideation to action.

Because many patients will withhold specific information about recent suicidal behaviors or suicidal ideation, Shawn Shea recommends a technique called the Chronological Assessment of Suicidal Events (CASE) approach. This technique uses a sequence of specific behavioral questions to determine how close the patient was to a lethal attempt. For example, if a patient had suicidal ideation involving a gun, the psychiatrist might ask, "Is there a gun in the home?" If the answer is "yes," then the interviewer continues with a series of follow-up questions until a "no" response is

obtained: “Have you taken out the gun?” “Did you load the gun?” “Did you point the gun at yourself?” “How long did you point the gun at yourself?” “What stopped you from completing the attempt?” Any developed planning for or practicing means of suicide or violence should be viewed as a high-risk finding requiring immediate attention.

The American Psychiatric Association Practice Guideline for the Assessment and Treatment of Patients with Suicidal Behaviors is one source for more detailed information about best practices in suicide assessment. The interview methods it recommends are analogous to the CASE approach reviewed earlier in this chapter. This approach leads the patient through a full exploration of episodes of suicidal tension to develop a fuller understanding of the precipitants and contexts of these occurrences, as well as the particular feelings, thoughts, and behaviors that define the degree of intent, planning, and lethal potential of actual suicidal conduct. The guideline also highlights the importance of inquiry into whether psychotic thought content or hallucinations is promoting suicidal ideation and behavior.

Insight

Insight is a frontal lobe/executive function represented by the capacity of the individual to appraise whether one’s thoughts, feelings, behaviors, perceptions, and planned actions are appropriate and realistic, and by the capacity to reflect on how one’s presentation may be perceived and interpreted by others. In everyday clinical practice, degree of insight is typically rated on a continuum from absent to full; many patients have partial insight. The degree of insight does not track neatly with severity of the illness. A person with dementia may be painfully aware of cognitive decline while a person with a mild anxiety disorder may have little or no insight into their overreactions to a fear stimulus.

> Her insight appeared to be partial. She had some sense that her behavior was being perceived as abnormal, but could not consider the possibility that she is ill in some way that requires medical attention despite the evidence for overt psychosis.

Judgment

Judgment is another executive function mediated by the frontal lobe that represents the capacity of the individual to appraise a situation or problem, consider and decide among alternatives, plan and execute a course of action and modify the course of action when necessary based on new inputs. At times, judgment is impaired despite adequate insight. For example, a patient may recognize that he is confused about money management, including the funds in his checking account, but decides to make a large purchase without clarifying his account balance.

Tests of judgment that do not require real-world decision making generally offer little value to an accurate assessment of judgment. Take the example of a patient just arrested for dancing nude down the middle of a busy street while yelling at passing cars in the throes of a manic episode. The patient may answer a question about what to do if one found a stamped envelope on the sidewalk with a socially acceptable response such as, "mail it," but the clinician would not therefore assess judgment as good. For this reason, it is useful to distinguish between abstract application of judgment and demonstrable judgment. Most examiners rate judgment on a continuum from very poor to good.

> Her social judgment in the context of the interviewer was fair. For the most part, she behaved well in the examination, but she did become loud at one point and called me a "fool." Recent events in the community, for example her disrobing in public, demonstrate very poor judgment. She stopped her medication weeks ago and is now unwilling to consider any type of psychotropic medication.

Cognitive Assessment

During the course of the interview, the clinician will develop history and make observations that inform cognitive assessment. In general, vocabulary level and abstract reasoning capacity are rough surrogates for intellectual capacity. Use of cognitive screening tests can be quite helpful in defining areas of strength or deficit that may not be obvious in a casual conversation without using more specific probes.

Domains of cognitive function important to the initial assessment include level of alertness, orientation, attention/concentration, visual-spatial function, memory (registration and recall), calculation, receptive and declarative language functions, fund of knowledge, capacity to abstract, and executive functions including insight and judgment.

The amount of detail required in the assessment of cognitive function will depend on the purpose of the examination and findings that raise concern about the possibility of cognitive problems, like a history of learning challenges or a clear decline from a previous level of cognitive function.

Two bedside tests commonly used to screen for cognitive function are the Mini Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MOCA). Each requires about 5 to 10 minutes to administer and both are scored on a 30-point scale providing cutoffs between normal range and levels of impairment. The MOCA is supplanting the better-known MMSE in many settings because it tests a wider range of functional problems, performs better in identifying mild cognitive impairment and provides administration instructions in many languages. The MOCA is available as an open source product accessed at www.mocatest.org while

the authors of MMSE have begun to assert copyright protection.

Physical Assessment

The extent of physical assessment will depend on the nature of the referral, the setting and the findings of the history and MSE. In the outpatient, setting, routine physical examination is not generally a part of the initial consultation while it is often critical to emergency room or inpatient settings. Psychiatrists do not usually personally conduct comprehensive physical examinations but may conduct focused examinations such as neurological or thyroid examinations. In the outpatient setting, the psychiatrist generally relies on the PCP to conduct the physical examination and it is useful in the initial evaluation to record the date of the most recent physical examination and review of recent laboratories if results are available. All psychiatrists should be prepared to measure vital signs and conduct other measurements, like waist circumference, that are indicated when psychotropic medications are used. The Abnormal Involuntary Movement Scale (AIMS) is an example of a screening physical examination conducted at baseline and serially during treatment with antipsychotic medication to monitor for abnormal movements like tardive dyskinesia.

Review and interpretation of laboratory results and ordering of relevant laboratory tests or special studies, such as neuroimaging, is an important aspect of many psychiatric evaluations. Psychiatrists lacking confidence in interpretation of test results should consult with medical providers that can interpret them and offer opinions about the significance of results in the context of other findings.

Closing the Interview

Inexperienced interviewers often are naive to the importance of the last 5 to 10 minutes of the first encounter and find closing of encounters challenging. It is important to alert the patient to the remaining time: "*We have to stop in about 10 minutes.*" This signal gives the patient the opportunity to plan how to use the time, especially if she has delayed or failed to address something that she wanted to. Directly asking the patient if there is anything else he wanted to share before the end of the encounter is a useful technique to reduce the risk that the patient will bring up an important issue "at the door" that cannot be adequately addressed. Sometimes late interview disclosures are too complex to address in the remaining time. If so, this may be an indication for a second interview to complete the evaluation. It is important to leave time to share opinions and recommendations and for questions.

COMMUNICATION OF OPINIONS AND RECOMMENDATIONS

The Formulation

The psychiatric interview and examination culminates with a formulation of the case that pulls together the biopsychosocial data into a working hypothesis about diagnosis and recommendations for treatment. The formulation should include a brief summary of the relevant findings from the history and examination including the psychosocial contexts in which the problem has developed and comments on the relevant contributions to the presentation of personality function, medical problems, social stress, and other social and cultural factors.

Although the formulation typically comes at the end of an evaluation document, the hypothesis checking process is a dynamic one that continues from the first contact with the patient to the summary of opinions. Since most psychiatric diagnoses represent syndromes that often have overlapping features and presentations, a differential diagnosis of possibilities other than the working diagnosis is usually found in the formulation, as well as plans to obtain other data (history, tests, studies) to sharpen the diagnosis when this is indicated. The *Diagnostic and Statistical Manual of Mental Disorders,* Fifth Edition (DSM-5) is the standard to establish criteria-based diagnosis in the United States. Some other areas of the world use ICD-10. The two systems cross walk in coding frameworks.

Finally, the formulation should include a summary of the risk assessment with estimates of acute and long-term risk of suicidal or violent behavior and an opinion about the appropriate level of care that will lead to a safe and successful outcome.

Treatment Recommendations

Consultations initiated by the patient and those requested by other professionals for input or care collaboration will usually include treatment recommendations. Certain types of evaluations, for example, disability determination, competency assessment, and other forensic-related assessments may not.

When evaluation produces treatment recommendations, these are typically shared with the patient at the conclusion of the encounter in a manner consistent with the patient's capacity to receive and process the information and with explicit discussion of matters relevant to informed consent for recommended treatment. Treatment discussions typically involve a good deal of psychoeducation about diagnosis, the nature, risks, and benefits of recommended treatments and information that addresses stigma and adherence. It is wise to involve significant others in these conversations especially if there are concerns that the patient may need assistance in processing information and making decisions to ensure that decisions are consistent with the patient's best interest. Recommendations

may include referral to other professionals or peer supports such as the National Alliance for Mental Illness (NAMI), the Mental Health Association (MHA), or Alcoholics Anonymous (AA). Plans for crisis contacts and supports are typically addressed. All of these counseling and coordination of care efforts should be documented in the medical record.

TECHNIQUES

General principles of the psychiatric interview, such as the patient–doctor relationship, open-ended interviewing, and confidentiality have been reviewed in earlier sections of the chapter. This section of the chapter will review a number of specific techniques that aid in facilitating and expanding information acquisition that are referred to here as "facilitating" and "expanding" interventions. Also reviewed here are some interventions that may prove to be counterproductive and interfere with patients telling their stories and that can interrupt development of the therapeutic alliance.

Facilitating Interventions

Facilitating interventions are effective in enabling the patient to continue sharing their story and are helpful in promoting a positive patient–doctor relationship (Table 7.1–9). At times, some of these techniques may be combined in a single intervention.

Table 7.1–9.
Facilitating Interventions

Reinforcement
Reflection
Summarizing
Education
Reassurance
Encouragement
Acknowledging emotion
Humor
Nonverbal communication
Silence

Reinforcement. Brief phrases such as, "I see," "Go on," "Yes," "Tell me more," "Hmm," or "Uh-huh," all convey the interviewer's interest in the patient continuing. It is important that these phrases fit naturally into the dialogue.

Patient: *For the past 2 months I've been waking up about 4 AM and I can't get back to sleep. I feel anxious like something bad is going to happen. A lot of times I feel bad all day; it's only about 8 PM when I begin thinking of going to bed that I feel a little bit better.*

Psychiatrist: *I see.*

Patient: *I used to be a good sleeper. This seemed to come out of nowhere. It's a miserable feeling; I can tell you that.*

Reinforcement interventions, although seemingly simplistic, are very important in encourage the patient to continue sharing material. Without these reinforcements the interview often becomes less productive.

Reflection. By using the patient's words, the psychiatrist indicates that she has heard what the patient is saying and conveys the implicit message that he or she is interested in hearing more.

Patient: *I don't know what's happening. I don't like going in to work anymore. The other guys at work are really starting to bug me.*

Psychiatrist: *Really starting to bug you.*

This response is not a question. A question, with a slight inflection at the end, calls for some clarification (see below). The reflection is not stated with a tone that is challenging or disbelieving but rather as a statement of fact: the patient's experience is what it is and the psychiatrist clearly hears it. Sometimes it is helpful to paraphrase the patient's statement so it does not sound like it is coming from an automaton. However, in the example provided the interviewer is not clear what the patient meant by "bug me" so changing the words (other than the pronoun) may steer the patient in a different direction.

Summarizing. Periodically during the interview, it is helpful to summarize what the patient offered about a certain topic. This provides the opportunity for the patient to clarify or modify the psychiatrist's understanding and possibly add new material.

Psychiatrist: *So, as I understand it, ever since you got a new boss this Spring things began to happen. You began to feel more anxious, and some of the comments of your co-workers really bothered you.*

Patient: *Yeah, and now that I think about it, around that time my wife started complaining I wasn't fun anymore.*

New material has been introduced. The psychiatrist may decide to continue with a further exploration of the previous discussion and return to the issue concerning the patient's wife at a later point.

Psychiatrist: *I want to hear about how things were between you and your wife but before we talk about that is there anything else you can say about how you felt at work?*

Education. At times in the interview, it is helpful for the psychiatrist to educate the patient about the interview process.

Patient: (after considerable hesitation) *There are some problems at home, but I don't know if that's what I'm supposed to be talking about.*

Psychiatrist: *It's helpful to talk about whatever has been bothering you. If I think we're getting off track I'll let you know. Tell me about the problems at home.*

If this is not the first session and the patient has generally been sharing information, then it might be useful to focus on the hesitation.

Psychiatrist: *It seems difficult for you to mention that. Why do you hesitate to talk about the problems at home?*

This may lead to a discussion of confidentiality, loyalty to family members, or the nonjudgmental stance of the psychiatrist.

Reassurance. It is often appropriate and helpful to provide reassurance to the patient. For example, accurate information about the usual course of an illness can decrease anxiety, encourage the patient to continue to discuss their illness, and strengthen their resolve to continue in treatment.

Patient: *I don't think I'll ever feel better.*

Psychiatrist: *I understand how hopeless it feels for you right now. Feelings like that are common in depression, but most people with this type of depression do get better and I think it's very likely you will also.*

It is generally inappropriate to reassure the patient when the psychiatrist cannot predict with some confidence what the outcome will be. In these cases the psychiatrist can assure the patient that he or she will continue to be available and will help in whatever way he or she can.

Encouragement. It is difficult for many patients to come for a psychiatric evaluation. Often they are uncertain about what will happen,

and receiving encouragement can facilitate their engagement.

Patient: *I'm not doing very well describing this nervousness. I've never done this before.*
Psychiatrist: *I think you are doing well in describing the nervousness. As you talk I'm getting a clearer picture of what it's been like for you.*

The psychiatrist is careful not to overstate the progress in the interview. The patient is given positive feedback about their efforts, but the secondary message is that although the "picture" is getting "clearer" there is more work to be done.

Acknowledgment of Emotion. It is important for the interviewer to acknowledge the expression of emotion by the patient. This frequently leads to the patient sharing more feelings and being relieved that they can do so. Sometimes a nonverbal action, such as moving a tissue box closer can suffice, or be used as an adjunct behavior to an empathic comment.

Patient: *He was a good friend.*
Psychiatrist*: As you talk about him you look very sad.*

If the display of the emotion is clear (e.g., patient openly crying), then it is not helpful to comment directly on the expression of the emotion.

Patient: (sobbing) *I really miss him.*
Psychiatrist: *I see that you are crying.*
Patient: *No shit. You're very observant.*

It is better to comment on the associated feelings.

Psychiatrist: *You feel awful without him.*

Humor. At times, patients make humorous comments or tell a joke. It can be very helpful if the psychiatrist smiles, laughs, or even, when appropriate, adds another punch line. This sharing of humor can decrease tension and anxiety and reinforce the interviewer's genuineness. It is important to be certain that the patient's comment was indeed meant to be humorous and that the psychiatrist clearly conveys that she is laughing with the patient not at the patient.

Silence. Careful use of silence can facilitate the progression of the interview. The patient may need time to think about what has been said or to experience a feeling that has arisen in the interview. The psychiatrist who cannot bear the anxiety of silence may terminate it prematurely retarding the development of insight or the expression of feeling by the patient. As George Engel encouraged new trainee, "Don't just do something, sit there." On the other hand, extended or repeated silences can deaden an interview and lead to a struggle as to which party can out wait the other. If the patient is looking at his watch or looking about the room, then it might be helpful to comment, "It looks like there are other things on your mind." If the patient has become silent and looks like he is thinking about the subject, then the psychiatrist might ask, "What thoughts do you have about that?"

Nonverbal Communication

In many good interviews, the most common facilitating interventions are nonverbal. Nodding of the head, body posture including leaning toward the patient, body positioning becoming more open, moving the chair closer to the patient, putting down pen and folder, and facial expressions including arching of eyebrows all indicate that the psychiatrist is concerned, listening attentively, and engaged in the interview. Excessive repetition or exaggeration of nonverbal techniques risks one being perceived in the vein of popular caricatures of the psychiatrist that nods his head repeatedly regardless of the content of what is being said or the emotion being expressed.

Expanding Interventions

There are a number of interventions that can be used to expand the focus of the interview. These techniques (Table 7.1–10) are helpful when the line of discussion has been sufficiently mined, at least for the time being, and the interviewer wants to encourage the patient to talk about other issues. These interventions are most successful when a degree of trust has been established in the interview and the patient feels that the psychiatrist is nonjudgmental about what is being shared.

Table 7.1–10.
Expanding Interventions

Clarifying
Associations
Leading
Probing
Transitions
Redirecting

Clarifying. At times carefully clarifying what the patient has said can lead to unrecognized issues or psychopathology.

A 62-year-old widow describing how it feels since her husband died 14 months ago repeatedly comments, "Everything is empty inside." The resident interprets this to mean that her world feels empty without her spouse and makes this interpretation on a few occasions. The patient's nonverbal cues suggest that she is not agreeing with this view. The supervisor asks the patient to clarify what she means by "empty inside." After some avoidance, the patient states that she is indeed empty inside; all her organs are missing; they have "disappeared."

The resident's interpretation may actually have been psychodynamically accurate, but the interpretation interfered with the opportunity to discover a somatic delusion. The correct identification of what the patient was actually saying led to an exploration of other thoughts, and delusions were uncovered. This vignette of "missing" the delusion is an example of the interviewer "normalizing" what the patient is saying. The interviewer was using secondary process thinking in understanding the words of the patient while the patient was using primary process thinking.

Associations. As the patient describes symptoms, other areas that may accompany or relate to a symptom should be explored. For example, the symptom of nausea leads to questions about appetite, bowel habits, weight loss, and eating habits. In addition, experiences that are temporally related may be investigated. When a patient is talking about their sleeping pattern, it can be a good opportunity to ask about dreams.

Leading. Often, encouraging the patient to continue their story can be facilitated by asking a "what," "when," "where," or "who" questions. "Why" questions are generally not helpful early in an interview.

Patient: *And I said, "That's enough."* (pause)
Psychiatrist: *What happened then?*

Sometimes the psychiatrist may suggest or ask about something not yet introduced by the patient because the psychiatrist surmises it may be relevant. In the courtroom, leading the witness is certain to raise an objection, but in clinical work it can be helpful as long as the interviewer is not making too far an intuitive leap and is open to reframing the question pending a response by the patient.

Patient: "*...feels like my husband is always telling me what to do, criticizing my driving.*"

Psychiatrist: *You mentioned that earlier. Have there been other relationships where you have also experienced that?*

Patient: *Well ...* (pause) *actually my father used to do that to my mother all the time. I really hated that.*

Probing. The interview may point toward an area of conflict, but the patient may minimize or deny any difficulties. Gently encouraging the patient to talk more about this issue may be quite productive.

Psychiatrist: *Tell me more about how things are at home.*

Patient: *Not really a whole lot to tell; everything is cool.*

Psychiatrist: *You mentioned that weekends are difficult.*

Patient: *I didn't mean that difficult.*

Psychiatrist: *OK. How are they different from before?*

Patient: *I don't know what you mean.* (pause) *Well, I guess it's a little different. Now my wife does her things and I'm doing mine.*

Transitions. Sometimes transitions occur very smoothly. The patient is talking about her primary education major in college and the psychiatrist asks, "*Did that lead to your work after college?*" On other occasions, the transition means moving to a different area of the interview and a bridge statement is useful.

Psychiatrist: *That gives me a good idea about your nervousness, perhaps now you can tell me about your health in general.*

Redirecting. A difficult technique for unseasoned interviewers is redirecting the focus of the patient. It can be difficult to move the interview in a different direction when one is concentrating on reinforcing the patient telling his or her story. However, the reality is that time constraints and the goal of obtaining a broad overview of the patient's life as well as the current problems makes redirection necessary at times. Redirection is also used when the patient changes topics too quickly or persists in offering excessive information about a nonproductive or already well-covered area. For both conscious and often unconscious reasons, the patient may avoid certain important areas and need guidance in approaching these subjects.

Patient: (After beginning to describe some significant issues with her mother) *But enough of that, let me tell you about my job.*

Psychiatrist: *Before we move on to your job perhaps you can say more about your struggles with your mother. It sounds like that has been very upsetting to you.*

Patient: (After much discussion about her arthritis) *and then six months ago I saw a new rheumatologist…*

Psychiatrist: *I know your arthritic pains have been a big burden to you, and we can come back to that, but I would also like to hear more about the issues you mentioned with your daughter.*

Sometimes if the patient is wandering and does not respond to an attempt at transitioning, then it is necessary to be more directive.

Psychiatrist: *I'm going to interrupt at this point because I am aware of the time we have left and there are several other areas it would be good to talk about.*

Obstructive Interventions

While supportive and expanding techniques facilitate the gathering of information and the development of a positive patient–doctor relationship, some other interventions are not helpful for either task (Table 7.1–11). Sometimes these interventions are technically correct, but if introduced in an unclear manner, unconnected to content, or poorly timed they may be experienced as unresponsive to the patient's concerns or feelings.

Closed-Ended Questions. A series of closed-ended questions early in the interview can retard the natural flow of the patient's story and reinforces the patient providing one-word responses or brief answers with little or no elaboration.

Psychiatrist: *Did you have a happy childhood?*

Patient: (Pause) *Yes, I guess so.*

Psychiatrist: *Did you have friends?*

Patient: (Intuitively trying to get the interview pointed in a more productive direction) *Well, I guess it depends on what you mean by a friend.*

Psychiatrist: (Not joining the patient's attempt to enrich the discussion) *Kids you did things with. How many?*

This example illustrates that the patient can be a partner in the interview, unless blocked by the psychiatrist. Many patients, some of whom have previous experiences in therapy, come prepared to talk about even painful matters. Over the course of time, psychiatrists, especially if they have had the benefit of supervision, learn from patients and refine their

interviewing skills.

Compound Questions. Some questions are difficult for patients to respond to because more than one answer is being sought.

Psychiatrist: *How did you feel? What did you do?*
Patient: *I'm not sure what happened.*

Table 7.1–11.
Obstructive Interventions

Excessive use of closed-ended questions
Compound questions
Excessive "Why?" questions
Judgmental questions
Minimizing patient concerns
Premature advice
Premature interpretations
Abrupt transitions
Ambiguous nonverbal communication

Why Questions. Especially early in the psychiatric interview, "why" questions are often nonproductive. Very often seeking an answer to a "why" question is one of the reasons that the patient has sought help.

Patient: *I felt very depressed. I just couldn't go on.*
Psychiatrist: *Why were you so depressed?*
Patient: *I don't know. I thought you might know.*

Judgmental Questions or Statements. Judgmental interventions are generally nonproductive for the issue at hand and also inhibit the patient from sharing even more private or sensitive material.

Patient: *I felt she wasn't paying attention so I yelled at her.*
Psychiatrist: *Don't you know that yelling only makes matters worse?*

It would be better for the psychiatrist to help the patient reflect on how successful that behavior was.

Psychiatrist: *How did that go?*
Patient: *Not good. Then she got upset about the yelling.*

Minimizing Patient's Concerns. In an attempt to reassure a patient the psychiatrist makes the error of minimizing a concern.

Patient: *I worry I'm going to have a heart attack. My doctor says I have to lose weight.*
Psychiatrist: *I don't think you need to worry. You look pretty healthy and your father is still alive.*

Rather than being reassured, the patient may feel that the psychiatrist does not understand what he is trying to express and is trivializing the experience. Furthermore, the advice is counter to what the PCP has said and is confusing to the patient. It is much more productive to explore the concern; there is likely much more material not yet shared.

Premature Advice. Advice given too early is often bad advice because the interviewer does not yet know all of the variables. In addition, it can preempt the patient arriving at a plan for himself or herself.

Patient: *My boss is so demanding.*
Psychiatrist: *Why don't you get in early in the morning before he does and give him a list of what will improve everyone's performance.*

Premature Interpretation. Even if it is accurate, a premature interpretation can be counterproductive as the patient may respond defensively and feel misunderstood.

Patient: *I was so angry at my neighbor when he said that and I'm not sure why.*
Psychiatrist: *He reminds you of your brother and the way he was always trying to control you.*
Patient: (angrily) *He doesn't remind me of my brother; he doesn't even look like him.*

Transitions. Some transitions are too abrupt and may interrupt important issues that the patient is discussing.

Patient: *Ever since my father died I've been feeling anxious.*
Psychiatrist: *Tell me more about your job.*

Nonverbal Communication. The psychiatrist repeatedly looking at her

watch, turning away from the patient, yawning, or refreshing the computer screen all convey boredom, disinterest, distraction, or annoyance. Just as reinforcing nonverbal communications can be powerful facilitators of a good interview; these obstructive actions can quickly shatter an interview and undermine the patient–doctor relationship.

SELECTED CHALLENGING INTERVIEW SCENARIOS

Patients with Psychosis

Patients with psychotic symptoms are often frightened, guarded, and suspicious about the purpose of the interview and the intentions of the clinician. They may have difficulty with reasoning and clear thought and communication. Hallucinations may cause inattention and distraction. These factors often require adaptations to match the capacity and tolerance of the patient. Patients with psychosis may not perceive their behaviors, speech, delusions, or hallucinations as abnormal in any way making it important to inquire in a manner that encourages the patient to share their ideational and perceptual experience without seeming judgmental of it.

Toward this end, prefacing direct questioning about psychotic symptoms by making normalizing statements such as "Many people have ideas that others find strange or different," can increase the patient's comfort with sharing delusional ideation and similarly for hallucinations by stating that, "Many people have had the experience of hearing a voice or seeing a vision." Such statement may help the patient talk more openly.

The psychiatrist should be alert for cues that psychotic processes may be part of the patient's productions during the interview. It is usually best to ask directly about any such behaviors or comments as this example illustrates:

During an initial interview, a 28-year-old man suddenly stood up and turned 360 degrees, sat down, and after a brief pause, resumed talking. When asked, "What just happened?" the patient responded that he "had to unwind." While exploring this example of concrete, primary process thinking, a number of other psychotic features were uncovered.

By definition, patients with delusions have fixed false beliefs. With delusions, as with hallucinations, it is important to explore the symptoms in some detail. Patients are often reluctant to discuss their experience because of past occurrences of being dismissed, ridiculed, or confined when they revealed their thoughts and perceptions. They may ask directly if the clinician believes the delusion. Interviewers are advised not to endorse the false beliefs, but it is rarely helpful to challenge the delusion directly particularly in the initial examination. It can be helpful to shift the focus of attention back to the patient's beliefs while acknowledging the need for more information by responding like this: "I believe that what you are experiencing is frightening and I would like to know more about your

experiences so I can form an opinion."

With patients evidencing a high degree of suspiciousness, it is useful to maintain distance and modulate statements or behaviors that patient may interpret as promoting excessive closeness, for example, excessive use of empathic statements or prolonged eye contact. Harry Stack Sullivan recommended that rather than sitting face to face with the patient who is paranoid, the psychiatrist might sit more side by side (but not too near), "looking out" with the patient. While interviewing the psychotic patient, the interviewer should keep in mind that the patient might be incorporating the clinician into the psychotic process. It can be helpful to ask about this directly, for example, "Are you concerned that I am involved?"

Acute psychosis raises concern about the potential for violent behaviors that are further reviewed in this section under the heading *The Potentially Violent Patient*.

The Patient with Depression

The depressed patient may have particular difficulty participating in the interview due to psychomotor slowing, cognitive deficits, and lack of motivation to engage caused by the depression itself. Feelings of hopelessness may contribute to a lack of engagement.

Patient: "*What's the use [of talking]? It's not going to make any difference. Nothing is going to change.*"

Psychiatrist: "*I'm glad that you are able to share that. It sounds like you feel hopeless about getting better. Hopeless feelings are common in patients who have depression. I think your condition will improve and you will feel more hopeful.*"

The depressed patient's responses to questions are often latent and may lack spontaneous elaboration. Depending on the severity of symptoms, the interviewer may need to turn to more direct questioning and abandon an open-ended format. The interview of severely depressed patients can be a laborious process and may require more time than expected to gather even minimally necessary information, even then collateral historians may need to fill in gaps in the patient's history.

The Potentially Suicidal Patient

There are times in initial interviews when it is obvious that one is encountering a patient on the brink of suicide, but in many other instances, patients have suicidal ideation without betraying much, if anything, in the way of clues. For this reason, it is important to conduct suicide risk assessment in any initial psychiatric interview. Certain diagnoses and problems heighten suicide risk, among them major depression, borderline

personality disorder, high levels of anxiety, acute psychosis, acute intoxication, patients who have become hopeless about their substance use problems, acute grief and loss reactions, and patients with grave or burdensome medical problems. Persons recently arrested and patients who are homicidal may also be suicidal. As discussed in earlier sections of this chapter, normalizing the potential for suicidal ideation is a good way to encourage patients to share their thoughts and feelings. An invitation like this can help to initiate the conversation about suicide: "Many people who describe a life event or the feelings that you have today are also having thoughts about death, dying, or even ending their own life. Has this been the case for you?"

The science of suicide risk assessment is beyond the scope of this chapter, but there are important risk factors that elevate suicide risk to survey as the history is developed. Some relate to demographics (e.g., age, gender, living alone, race and ethnicity, sexual orientation) others to particular factors related to diagnosis and clinical features (see above) and life experiences such as childhood physical and sexual abuse. Family history of suicide, prior attempts at suicide, and acts of deliberate self-harm are important historical risk factors. Active substance use, acute psychosis, and high levels of anxiety seem to fuel suicide risk. Protective factors include the level of social support, children in the home, pregnancy, and religious beliefs that admonish against suicide.

The Potentially Violent Patient

Safety for the patient, interviewer, and others present is the main concern when engaging with hostile or agitated patients. If one knows in advance that an encounter with a hostile or agitated patient is to occur, there is opportunity to plan and marshal supports. As noted elsewhere in the chapter, thoughtful attention to managing personal space and having an attendant during the encounter can be important to reducing risks in these situations. It is good practice to alert other clinical staff in the area and security, if an officer is present.

Reducing excessive stimulation of already hostile or agitated patients is a priority. Interviewers should be aware of their own body position and avoid postures, behaviors, eye contact tone of voice, and choice of words that could perceived as threatening. Approach the interview in a calm, direct manner and take care not to make bargains or promises to elicit cooperation in the interview, for example, implying that cooperation will avoid an admission. These tactics will only likely escalate agitation and the potential for aggression when the patient discovers the examiner cannot deliver.

Managing the interview requires thoughtful choices about areas to explore and those to leave alone, as well as the phrasing and timing of questions to avoid stirring up unnecessary conflict. Taking breaks, agreeing

to leave a topic that causes tension for another time and attending to basic comforts like hunger and nicotine replacement before pressing on with the interview may reduce tension and build rapport. Avoid unnecessary displays of dominance and give patients choices where there are reasonable options so that the patient can maintain some semblance of control.

Despite best efforts, termination of some interviews must occur when the patient's agitation escalates. Generally, preceding unpremeditated violence is a period of gradually escalating psychomotor agitation such as pacing, loud speech, and threatening comments. As noted earlier in the chapter, some agitated patients will calm when the interviewer shares that the patient's behavior is making the interviewer too anxious to be helpful. There is no virtue in pressing on and taking unnecessary risk, especially when the interviewer is frightened. Consider terminating the interview and seeking assistance if the patient's behaviors escalate and do not easily respond to structure and feedback.

The hostile or agitated patient may be placing others in his community at risk. The most common victims of violence by acutely ill psychiatric patients are those that live or work in close proximity to the patient, such as family, residential staff, and coworkers. If the patient can tolerate it, explore feelings and attitudes about those close to him to discover persons who may be at particular risk. If the patient makes specific threats and/or identifies specific targets, take the threat seriously and carefully define the actual risk.

Like the assessment of suicide risk, a complete discussion of the violence risk assessment is beyond the scope of this chapter. Certain facts are worth noting here. Research reveals more about risk factors for violence than protective factors. Among the important risk factors are patient age (youth), lower intelligence, past acts of violence and general suspiciousness as well as active psychosis; individuals physically or sexually abused as children are also at greater risk of violence. Interviewing for violence risk follows much the same pattern as that used for suicide risk assessment. The interview generally moves along the continuum of ideation to plan to action exploring the nature and extent of the patient's ideation, planning, and intent. Similar to suicide assessment, the clinician seeks to define past instances of violence, the particulars of these episodes and the outcomes to understand the level of risk at the time of the interview and to identify patients who are at elevated risk over time.

The Patient Engaging in Deception

For the purpose of this discussion, deception is defined as the intentional production or exaggeration of symptoms. Psychiatrists train to maintain vigilance for deception while eliciting information but they are not immune to being deceived. Studies show that even experienced clinicians will fail to recognize malingered mental illness in some settings. Patients engage in

deceit for many different reasons. Some are motivated by secondary gain, like work excuses, disability, or to secure desired medications. Others are seeking primary gains, like the psychological benefits of assuming a sick role.

This example illustrates some points:

> A soldier in basic training presents to the military base mental health service with a long list of symptoms. He escalates the drama of his responses each time he perceives that the psychiatrist is not reacting to his productions as he expected. Finally, after describing some elaborate visual hallucinations that do not comport to those typically experienced by psychotic patients and that also failed to impress the psychiatrist, the soldier blurts out, "What's the matter, you never heard of paranoid schizophrenia?" Later, the psychiatrist learned the soldier visited the base library to read about psychiatric illnesses, and schizophrenia. The solider admitted he was feigning symptoms to avoid an undesirable deployment.

There are no infallible means to detect deceit, but there are markers in interviews that raise suspicion including vague and contradictory responses, symptom claims that are inconsistent with functional findings, claims that are far beyond the variations of psychopathology seen in certain conditions or problems, or that are frankly preposterous. View skeptically patients that draw excessive attention to symptoms and those that become irritated with attempts to clarify the nature of complaints. This said it is important to remain cognizant of the fact that there are unusual presentations and one should not assume that all that confuses or does not fit typical patterns is therefore evidence of deceit. Collateral historians become critically important when there is concern about the credibility of patient reports. Certain psychological tests are designed to detect deceit (e.g., the Word Memory Test) and others have validity scales imbedded in their structure that measure the nature of response style to identify likely exaggeration or suppression of symptoms in reporting (e.g., the Minnesota Multiphasic Personality Inventory [MMPI]).

It is important to be confident in one's opinion about deceit before confronting a patient with this conclusion. It is also important to try to maintain an empathic approach to when discussing findings, ideally helping the patient to save face while also accepting responsibility. For example, the following type of exchange can be helpful:

> "Would you agree that it is important for you to be truthful with me about your history and symptoms and for me to be truthful with you about my findings and recommendations? Okay then to be a good doctor to you, I must share that I believe that you are exaggerating some of the symptoms you have described. These are the reasons why I believe this to be so (clinician explains the reasons). I am concerned that you are trying to present yourself as ill in certain ways that in the end could lead to bad outcomes, like applying treatments that could do more harm than good. If I am right about this, let us start over and talk about what is really going on and why you might have tried to present yourself as ill in this way. Then I

can see if there are ways I can help you with your predicament that are reasonable and safe."

SPECIAL TOPICS

Note Taking and the Electronic Medical Record

Many psychiatric interviewers take handwritten notes during the interview. With the advent of electronic medical records, some advocate for a method called "contemporaneous documentation" during which the interviewer pauses either periodically during the encounter or at the end of the encounter to create all or part of the consultation document with the patient's participation. Advocates of this method point out that a patient increasingly has direct access to all or parts of their medical record through patient portals. Writing the note with the patient, so to speak, reduces the chances of documenting inaccurate information or writing statements in a manner that the patient might find offensive.

This is an example of how contemporaneous documentation might proceed in real time:

> Psychiatrist: "Now you have told me a lot about your current problem and I'd like to pause now and document what you said to summarize it in the record. I will summarize what I am writing while I type it. You listen and tell me whether I have heard you correctly."

The process of contemporaneous documentation can result in synergistic efficiency that provides opportunity to summarize, clarify, and correct information while completing documentation. For those that now offer an hour for the consultation and then do documentation after the encounter, extending the time for the consultation to include time for contemporaneous documentation an improved time economy may be realized.

Note taking and contemporaneous documentation in an electronic record during the interview are activities that present some risks. These activities can interfere with the observation of nonverbal information, like affects and body language, that are often as or more important than what the patient says. Failing to recognize these nonverbal cues can result in both loss of important data and empathic failures affecting the development of rapport. The activities can also introduce distractions for both the patient and the interviewer if they dominate the interview.

Cultural Variation

One definition of culture is a common heritage, a set of beliefs, and values that set expectations for behaviors, thoughts, and even feelings. Culture influences the patient's attribution of illness, decisions about when and where to seek help, how distress is articulated or demonstrated to others

(including when speaking with medical providers) and perceptions of necessary and acceptable forms of treatment. Often, individuals from a minority population are reluctant to seek help from a physician who is from the majority group, especially for emotional difficulties.

Just as the array of cultural syndromes and variations are beyond the scope of this chapter, interviewers cannot be aware of all culturally bound syndromes or idioms of distress but should make the effort to learn about variant presentations and syndromes that affect cultural groups commonly encountered in the scope of their practice settings. Reducing the risk of overdiagnosis (e.g., misreading a cultural belief for a delusion) and the possibility of over or undertreatment that can issue from inaccurate diagnosis are among the reasons cultural awareness is important.

By remaining humble and curious about the patient's cultural identifications and beliefs, and being open and respectful to them, the psychiatrist increases the likelihood of developing a trusting working relationship and a collaborative and effective plan.

Interviewing with an Interpreter

Translated interviews require forethought and preparation. Whenever possible, the use of a professional translator is favored over use of a person familiar to the patient. Professional translators have training and experience to interact with patients and providers skillfully, the vocabulary skills to translate medical words and concepts, operate under a code of ethics that highlights confidentiality, and understand the importance of verbatim translation and neutrality. The use of nonprofessional interpreters risks a poor quality translation fraught with all the challenges of agency, skill, and interpersonal dynamics that friends and family bring to the process of interpretation and too often lead to incomplete or frankly inaccurate communication.

It is helpful to speak with the interpreter prior to the interview to clarify the goals of the examination and the expectations. Most professional interpreters are attuned to the fact that their role is to facilitate not create communication and so understand the importance of verbatim translation, and that the interpreter does not intersperse ad-lib questions or edit the patient's responses. If the interpreter does not primarily work with psychiatric patients, it may be useful to highlight the need for verbatim translation even if the responses are disorganized or contain unusual content. During the course of the interview, the interpreter may slip into giving sparse interpretation to longer speech productions. If this occurs, the interviewer should pause and ask the interpreter to offer everything that the patient said and remind the interpreter of the request for verbatim interpretation.

To avoid the patient connecting more with the interpreter than the interviewer one should arrange the positions of the parties to allow the

patient to maintain gaze in the direction of both interviewer and interpreter during the encounter. Interviewers need to be sensitive to the pace of questioning and parse out the questions and comments into smaller segments to allow the interpreter time to facilitate communication. As a general rule, more than one or two sentences or three complete thoughts is more than the anyone, including interpreters, can hold accurately in memory to reproduce a verbatim translation once the interviewer stops speaking.

At the conclusion of the interview, the interpreter who is especially knowledgeable about the patient's cultural background may be able to provide helpful insights and opinions about the nature of the patient's communication, for example, if a formal thought disorder was present and whether ideas the patient expressed were outside the norm for the cultural group.

Observed and Recorded Interviews

One-way mirror observation and/or recording of interviews occur in some settings for clinical, educational, or forensic purposes. Before initiating an observation or recording, the patient's informed consent for these activities must be specifically established. The patient should understand the purpose of observation and/or recording (e.g., psychotherapy, supervision), how it will be used and who may have access to it (e.g., an attorney or other third parties in a forensic case). When recordings are made for clinical purposes it should be made clear how long the recording will be archived before it is destroyed. If a recording is to be used for educational purposes, there are special informed consent considerations, including the patient's right to withdraw the consent and others matters that are beyond the scope of this chapter. In treatment settings, the clinician must ensure that the patient receives the same nature and quality of care regardless of whether consent to observe or record is granted.

Observation and/or recording of an interview can perturb the clinical interaction in expected and unexpected ways. The clinician should remain vigilant for untended negative effects on engagement and the free flow of information when interviews are observed or recorded.

Forensic Interviews

For the purposes of this section, a forensic interview is defined as one taking place in a legal context for a purpose other than clinical treatment. For example, the psychiatrist may be retained by an insurance company to conduct an Independent Medical Examination (IME) to offer an opinion in a disputed disability case, or by an attorney representing or prosecuting a defendant in a criminal case to offer opinions about competency to stand trial or criminal responsibility. Even in forensic settings where psychiatrist

is acting in a clinical role, for example, providing treatment in a correctional facility, some of the issues to be discussed here will apply.

Forensic consultants, with the exception of those providing clinical care in a forensic setting, do not form patient–doctor relationships with subjects, yet should still operate in these settings observing the ethical principles that apply and aspire to maintain the subject's dignity.

The agency or duty of the psychiatrist in the forensic setting is typically owed to a third party, for example, the attorney that retained the psychiatrist. Before proceeding with the interview, the subject should understand the agency of the psychiatrist and the boundaries of confidentiality. For example, when the psychiatrist is retained by the subject's defense attorney confidential material is protected by the attorney-client privilege unless and until the attorney and client decide to use an opinion, at which point a report or testimony may be provided requiring a waiver of confidentiality to produce it. If the examiner is retained by a prosecuting attorney, an insurance company in an adversarial proceeding, or if a judicial order demands a report, the subject waives confidentiality by agreeing to the examination.

The principles of forensic interview are built on the foundation of sound clinical practices though there are some important differences. The level of detail required in a forensic setting is typically greater than that sought in a general psychiatric interview. Forensic interviewing techniques require more detailed probing of the patient's story and mental state that at times may necessitate more clarification and even confrontation than typically utilized in a general psychiatric interview. Such interventions should be conducted in an ethically sound and sensitive manner. Collateral sources of information are critical to forensic examinations. These sources help to discover important elements of the history that the subject may not provide and serve to externally verify information offered by the subject. Forensic interviews are much more likely to be recorded than general psychiatric interviews and require special considerations for recording that are beyond the scope of this chapter, but the subject must know when recording is occurring and understand who may have access to such recordings as part of the consent process. Psychiatrists engaging in forensic interviewing should have a clear sense of the boundaries of their own training and experience and recognize the implications of their activities and the opinions they offer. Knowledge of the legal system (including the legal standards that apply to clinical questions), skills in specialized interviewing techniques, the role and proper application of psychological testing, and the vicissitudes of engaging with attorneys and courts are some examples of the specialized information and skills that forensic settings demand.

Telepsychiatry

In response to concerns about access to expertise and with improvements

in technology, interviewing and examining patients through video conferencing is now emerging as a feasible alternative to live interviewing in selected circumstances. Telepsychiatry is primarily utilized in underserved areas where geographic distance from or reliable access to the provider makes it unrealistic for the patient to come to a live appointment. Typical applications are in rural areas and in the care of homebound patients. One must be cognizant of the fact that technology is not pristine enough to protect against some drop out of information normally gleaned in a live interview. Televised observation and communication can degrade nuisances of speech, subtleties of the nature of eye contact, emotional display, body language, and other forms of nonverbal communication. Causes of information degradation include the narrow perspective of the camera, the quality of the audio-visual product and the unexplained effects that humans have on one another in a live interaction that are not completely replicated in a distance medium.

The American Telemedicine Association and the American Medical Association are associations that work in collaboration with other professional organizations to develop evidence-based guidelines promoting quality and safe practices for various applications of telemedicine including telepsychiatry. Clinicians engaging in telemedicine are advised to familiarize themselves with the special considerations that this medium introduces, including medical-legal and reimbursement issues.

REFERENCES

*American Psychiatric Association. Practice guideline for the assessment and treatment of patients with suicidal behavior. *Am J Psychiatry*. 2003;(11 Suppl):1–60.

*American Psychiatric Association. Practice Guideline for the Psychiatric Evaluation of Adults. 2nd ed. *Am J Psychiatry*. 2006;(Suppl):163.

Baliant M. *The Doctor, His Patient and the Illness*. 2nd ed. New York: International Universities Press; 1972.

Chiles JA, Strosahl KD. *Clinical Manual for Assessment and Treatment of Suicidal Patients*. Washington, DC: American Psychiatric Publishing; 2005.

Cruz M, Roter DL, Cruz RF, et al. Appointment length, psychiatrists' communication behaviors, and medication management appointment adherence. *Psychiatr Serv*. 2013;64(9):886.

Dhalla S, Kopec JA. The CAGE questionnaire for alcohol misuse: A review of reliability and validity studies. *Clin Invest Med*. 2007;30(1):33.

Engel GL. The need for a new medical model: A challenge for biomedicine. *Science*. 1977;196:129.

*Epstein RM. Whole mind and shared mind in clinical decision-making: *Patient Educ Couns*. 2013;90(2):200–206.

Folstein MF, Folstein SW, McHugh PR. Mini Mental State: A practical method of grading the cognitive state of patients for the clinician. *J Psychiatry Res*. 1975;12:189.

Gabbard GO. *Psychodynamic Psychiatry in Clinical Practice*. 4th ed. Washington, DC: American Psychiatric Publishing; 2005.

Gaw AC. *Cross Culture, Ethnicity, and Mental Illness*. Washington, DC: American Psychiatric Press; 1993.

Greenfield SF, Hennessey G. Assessment of the patient. In: Galanter M, Kleber HD, eds. *Assessment of the patient. In Textbook of Substance Abuse Treatment*. 5th ed. Washington, DC: American Psychiatric Publishing; 2008.

Groves JE. Taking care of the hateful patient. *N Engl J Med*. 1978;298:883.

Irani JS, Middleton JL, Marfatia R, Omana ET, D'Amico F. The use of electronic health records in the exam room and patient satisfaction: A systematic review. *J Am Board Fam Med*. 2009;22(5):533.

Kashner TM, Rush AJ, Suris A, Biggs MM, Gajewski VL. Impact of structured clinical interview on physicians' practices in community mental health settings. *Psychiatr Serv*. 2003;54:712.

Kendell RE. Five criteria for an improved taxonomy of mental disorders. In: Helzer JE, Hudziak JJ, eds. *Defining Psychopathology in the 21st Century. DSM-V and Beyond*. Washington, DC: American Psychiatric Publishing; 2002:3.

Kraemer HC, Measelle JR, Ablow JC, Essex MJ, Boyce T. A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *Am J Psychiatry*. 2003;160:1566.

*MacKinnon RA, Michels R, Buckley PJ. *The Psychiatric Interview in Clinical Practice*. 2nd ed. Arlington, VA. American Psychiatric Publishing; 2005.

McCabe R, Priebe S. Communication and psychosis: It's good to talk, but how? *Br J Psychiatry*. 2008;192:404.

McLean M, Armstrong D. Eliciting patient's concerns: A randomized controlled trial of different approaches by the doctor. *Br J Gen Pract*. 2004;54:663.

Rodgers R. *Clinical Assessment of Malingering and Deception*. New York: Guilford Press; 2008.

Sadock BJ, Sadock VA, eds. *Kaplan & Sadock's Comprehensive Textbook of Psychiatry*. 9th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2009.

*Shea SC. *Psychiatric Interviewing; The Art of Understanding*. 2nd ed. Philadelphia, PA: W.B. Saunders; 1998.

Shore JH. Telepsychiatry. Videoconferencing in the delivery of psychiatric care. *Am J Psychiatry*. 2013;170:256.

Strous RD, Ulman AM, Kotler M. The hateful patient revisited: relevance for 21st century medicine. *Eur J Int Med*. 2006;17:387.

▲ 7.2 Outline for a Psychiatric Examination

BENJAMIN J. SADOCK, M.D.

This section complements the previous section "Psychiatric Interview, History, and Mental Status Examination," in that it provides a comprehensive outline on how to write the psychiatric report. The need to follow some sort of outline in gathering data about a person in order to make a psychiatric diagnosis is universally recognized. The one that follows calls for including a tremendous amount of information about the patient;

however, not all need be obtained depending on circumstances in the case. Beginning clinicians are advised to get as much information as possible; more experienced clinicians can pick and choose among the series of questions they might ask. In all cases, however, the person is best understood within the context of his or her life events.

The psychiatric report includes both the psychiatric history and the mental status. The mental status examination covers how the patient is feeling and thinking at the moment as he or she responds to the specific questions asked by the examiner. The psychiatric history is a record of the patient's life. Also called the anamnesis (from the Greek meaning "to remember") it describes life events from infancy to old age, to the extent it can be recalled. The clinician should attempt to elicit the emotional reaction to each life event as remembered by the patient in addition to the event itself. At times it may be of interest to report, in detail, what questions were asked and the answers that were received; but this should be kept to a minimum. When writing the report the clinician should use the patient's own words, especially when describing symptoms such as hallucinations or delusions.

The psychiatric report does not only include the psychiatric history and mental status, it also includes a summary of positive and negative findings and an interpretation of the data found, thereby helping to provide an understanding of the case. The examiner address critical questions in the report: Are future diagnostic studies needed, and, if so, which ones? Is a consultant needed? Is a comprehensive neurological workup, including an electroencephalogram (EEG) or computer tomography (CT) scan, needed? Are psychological tests indicated? Are psychodynamic factor's relevant? The report includes a diagnosis made according to the newly revised fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). A prognosis is also discussed in the report, with good and bad prognostic factors listed. The report concludes with a discussion of a treatment plan and makes firm recommendations about the management of the case.

The outline for the psychiatric examination that follows was developed by Harold I. Kaplan, M.D. and the author. It has been used by countless numbers of psychiatrists over several generations and has been revised and updated periodically over the years, especially for this edition.

PSYCHIATRIC HISTORY

A. *Identification:* Name; age; marital status; sex; occupation; language, if other than English; race; nationality; religion, if pertinent; previous admissions to a hospital for the same or a different condition; and persons with whom the patient lives.
B. *Chief complaint:* Exactly why the patient came to the psychiatrist, preferably in the patient's own words; if that information does not come

from the patient, note who supplied it.

C. *History of present illness:* Chronological background and development of the symptoms or behavioral changes that culminated in the patient's seeking assistance; patient's life circumstances at the time of onset; personality when well; how illness has affected life activities and personal relations—changes in personality, interests, mood, attitudes toward others, dress, habits, level of tenseness, irritability, activity, attention, concentration, memory, and speech; psychophysiological symptoms—nature and details of dysfunction; pain—location, intensity, and fluctuation; level of anxiety—generalized and nonspecific (free floating) or specifically related to particular situations, activities, or objects; how anxieties are handled—avoidance, repetition of feared situation, or use of drugs or other activities for alleviation.

D. *Past psychiatric and medical history:* (1) Emotional or mental disturbances—extent of incapacity, type of treatment, names of hospitals, length of illness, and effect of treatment; (2) psychosomatic disorders—Hay fever, arthritis, colitis, chronic fatigue, recurrent colds, and skin conditions; (3) medical conditions—customary review of systems, sexually transmitted diseases, alcohol or other substance abuse, or at risk for acquired immunodeficiency syndrome (AIDS); (4) neurological disorders—headache, craniocerebral trauma, loss of consciousness, seizures, or tumors.

E. *Family history:* Elicited from patient and from someone else, because quite different descriptions may be given of the same people and events; ethnic, national, and religious traditions; other people in the home, descriptions of them—personality and intelligence—and what has become of them since the patient's childhood; descriptions of different households lived in; present relationships between the patient and those who are in the patient's family; role of illness in the family; family history of mental illness; where does the patient live—neighborhood and particular residence of the patient; is the home crowded; privacy of family members from each other and from other families; sources of family income and difficulties in obtaining it; public assistance (if any) and attitudes about it; will the patient lose his or her job or apartment by remaining in the hospital; who is caring for the patient's children.

F. *Personal history:* History of the patient's life from infancy to the present to the extent it can be recalled; gaps in history as spontaneously related by the patient; emotions associated with different life periods (painful, stressful, and conflictual) or with phases of the life cycle.

1. Early childhood (through 3 years of age).
   a. Prenatal history and mother's pregnancy and delivery: Length of pregnancy, spontaneity and normality of delivery, birth trauma, whether the patient was planned and wanted, birth defects.
   b. Feeding habits: Breast fed or bottle fed, eating problems.

c. Early development: Maternal deprivation, language development, motor development, signs of unmet needs, sleep pattern, object constancy, stranger anxiety, separation anxiety.
d. Toilet training: Age, attitude of parents, feelings about it.
e. Symptoms of behavior problems: Thumb sucking, temper tantrums, tics, head bumping, rocking, night terrors, fears, bed-wetting or bed-soiling, nail-biting.
f. Personality and temperament as a child: Shy, restless, overactive, withdrawn, studious, outgoing, timid, athletic, friendly patterns of play, reactions to siblings.
g. Early or recurrent dreams or fantasies.

2. Middle childhood (3 to 11 years of age): Early school history—feelings about going to school, early adjustment, gender identification, conscience development, punishment; social relationships, and attitudes toward siblings and playmates.
3. Later childhood (prepuberty through adolescence).
   a. Peer relationships: Number and closeness of friends, leader or follower, social popularity, participation in group or gang activities, idealized figures; patterns of aggression, passivity, anxiety, antisocial behavior.
   b. School history: How far the patient went in school; adjustment to school; relationships with teachers—teacher's pet or rebellious; favorite studies or interests; particular abilities or assets; extracurricular activities; sports; hobbies; relationships of problems or symptoms to any school period.
   c. Cognitive and motor development: Learning to read and other intellectual and motor skills, minimal cerebral dysfunction, learning disabilities—their management and effects on the child.
   d. Particular adolescent emotional or physical problems: Nightmares, phobias, bed-wetting, running away, delinquency, smoking, drug or alcohol use, weight problems, feeling of inferiority.
   e. Psychosexual history.
      i. Early curiosity, infantile masturbation, sex play.
      ii. Acquiring of sexual knowledge, attitude of parents toward sex, sexual abuse.
      iii. Onset of puberty, feelings about it, kind of preparation, feelings about menstruation, development of secondary sexual characteristics.
      iv. Adolescent sexual activity: Crushes, parties, dating, petting, masturbation, wet dreams (nocturnal emissions), and attitudes toward them.
      v. Attitudes toward same and opposite sex: Timid, shy, aggressive, need to impress, seductive, sexual conquests, anxiety.
      vi. Sexual practices: Sexual problems, homosexual and heterosexual

experiences, paraphilias, promiscuity.

f. Religious background: Strict, liberal, mixed (possible conflicts), relationship of background to current religious practices.

4. Adulthood.
   a. Occupational history: Choice of occupation, training, ambitions, and conflicts; relations with authority, peers, and subordinates; number of jobs and duration; changes in job status; current job and feelings about it.
   b. Social activity: Whether patient has friends; whether he or she is withdrawn or socializing well; social, intellectual, and physical interests; relationships with same sex and opposite sex; depth, duration, and quality of human relations.
   c. Adult sexuality.
      i. Premarital sexual relationships, age of first coitus, sexual orientation.
      ii. Marital history: Common-law marriages; legal marriages; description of courtship and role played by each partner; age at marriage; family planning and contraception; names and ages of children; attitudes toward raising children; problems of any family members; housing difficulties, if important to the marriage; sexual adjustment; extramarital affairs; areas of agreement and disagreement; management of money; role of in-laws.
      iii. Sexual symptoms: Anorgasmia, impotence (erectile disorder), premature ejaculation, lack of desire.
      iv. Attitudes toward pregnancy and having children; contraceptive practices and feelings about them.
      v. Sexual practices: Paraphilias, such as sadism, fetishes, voyeurism; attitude toward fellatio, cunnilingus; coital techniques, frequency.
   d. Military history: General adjustment, combat, injuries, referral to psychiatrists, type of discharge, veteran status.
   e. Value systems: Whether children are seen as a burden or a joy; whether work is seen as a necessary evil, an avoidable chore, or an opportunity; current attitude about religion; belief in heaven and hell.

MENTAL STATUS

Sum total of the examiner's observations and impressions derived from the initial interview.

A. Appearance.
   1. Personal identification: May include a brief nontechnical description of the patient's appearance and behavior as a novelist might write it. Attitude toward examiner can be described here: cooperative, attentive,

interested, frank, seductive, defensive, hostile, playful, ingratiating, evasive, or guarded.

2. Behavior and psychomotor activity: Gait, mannerisms, tics, gestures, twitches, stereotypes, picking, touching examiner, echopraxia, clumsy, agile, limp, rigid, retarded, hyperactive, agitated, combative, or waxy.
3. General description: Posture, bearing, clothes, grooming, hair, nails; healthy, sickly, angry, frightened, apathetic, perplexed, contemptuous, ill at ease, poised, old looking, young looking, effeminate, masculine; signs of anxiety—moist hands, perspiring forehead, restlessness, tense posture, strained voice, wide eyes; shifts in level of anxiety during interview or with particular topic; eye contact (50 percent is normal).

B. *Speech:* Rapid, slow, pressured, hesitant, emotional, monotonous, loud, whispered, slurred, mumbled, stuttering, echolalia, intensity, pitch, ease, spontaneity, productivity, manner, reaction time, vocabulary, prosody.

C. *Mood and affect.*

1. Mood (a pervasive and sustained emotion that colors the person's perception of the world): How does patient say he or she feels; depth, intensity, duration, and fluctuations of mood—depressed, despairing, irritable, anxious, terrified, angry, expansive, euphoric, empty, guilty, awed, futile, self-contemptuous, anhedonic, alexithymic.
2. Affect (the outward expression of the patient's inner experiences): How the examiner evaluates the patient's affects—broad, restricted, blunted or flat, shallow, amount and range of expression; difficulty in initiating, sustaining, or terminating an emotional response; whether the emotional expression is appropriate to the thought content, culture, and setting of the examination; examples should be given if emotional expression is not appropriate.

D. Thinking and perception.

1. Form of thinking.
   a. Productivity: Overabundance of ideas, paucity of ideas, flight of ideas, rapid thinking, slow thinking, hesitant thinking; whether the patient speaks spontaneously or only when questions are asked; stream of thought, quotations from patient.
   b. Continuity of thought: Whether the patient's replies really answer questions and are goal directed, relevant, or irrelevant; loose associations; lack of cause-and-effect relationships in the patient's explanations; illogical, tangential, circumstantial, rambling, evasive, persevering statements, blocking or distractibility.
   c. Language impairments: Impairments that reflect disordered mentation, such as incoherent or incomprehensible speech (word salad), clang associations, neologisms.
2. Content of thinking.
   a. Preoccupations: About the illness, environmental problems; obsessions, compulsions, phobias; obsessions or plans about suicide,

homicide; hypochondriacal symptoms, specific antisocial urges or impulses.

3. Thought disturbances.
   a. Delusions: Content of any delusional system, its organization, the patient's convictions as to its validity, how it affects his or her life; persecutory delusions—isolated or associated with pervasive suspiciousness; mood-congruent or mood-incongruent; thought insertion.
   b. Ideas of reference and ideas of influence: How ideas began, their content, and the meaning that the patient attributes to them.
4. Perceptual disturbances.
   a. Hallucinations and illusions: Whether the patient hears voices or sees visions; content, sensory system involvement, circumstances of the occurrence; hypnagogic or hypnopompic hallucinations; thought broadcasting.
   b. Depersonalization and derealization: Extreme feelings of detachment from self or from the environment.
5. Dreams and fantasies.
   a. Dreams: Prominent ones, if the patient will tell them; nightmares.
   b. Fantasies: Recurrent, favorite, or unshakable daydreams.

E. Sensorium.

1. Alertness: awareness of environment, attention span, clouding of consciousness, fluctuations in levels of awareness, somnolence, stupor, lethargy, fugue state, coma.
2. Orientation.
   a. Time: Whether the patient identifies the day or the approximate date and the time of day correctly; if in a hospital, whether the patient knows how long he or she has been there; whether the patient behaves as though oriented to the present.
   b. Place: Whether patient knows where he or she is.
   c. Person: Whether patient knows who the examiner is and the roles or names of the persons with whom the patient is in contact.
3. Concentration and calculation: Whether the patient can subtract 7 from 100 and keep subtracting 7s; if the patient cannot subtract 7s, whether easier tasks can be accomplished, such as 4 × 9 and 5 × 4; whether the patient can calculate how many nickels are in $1.35; whether anxiety or some disturbance of mood or concentration seems to be responsible for difficulty.
4. Memory: Impairment, efforts made to cope with impairment—denial, confabulation, catastrophic reaction, circumstantiality used to conceal deficit; whether the process of registration, retention, or recollection of material is involved.
   a. Remote memory: Childhood data, important events known to have occurred when the patient was younger or free of illness, personal

matters, neutral material.

b. Recent past memory: Past few months.
c. Recent memory: Past few days, what did the patient do yesterday and the day before, what did the patient have for breakfast, lunch, and dinner.
d. Immediate retention and recall: Ability to repeat six figures after the examiner dictates them—first forward, then backward, then after a few minutes' interruption; other test questions; whether the same questions, if repeated, called forth different answers at different times.
e. Effect of defect on patient: Mechanisms the patient has developed to cope with the defect.

5. Fund of knowledge: Level of formal education and self-education; estimate of the patient's intellectual capability and whether the patient is capable of functioning at the level of his or her basic endowment; counting, calculation, general knowledge; questions should have relevance to the patient's educational and cultural background.
6. Abstract thinking: Disturbances in concept formation; manner in which the patient conceptualizes or handles his or her ideas; similarities (e.g., between apples and pears), differences, absurdities; meanings of simple proverbs, such as "a rolling stone gathers no moss"; answers may be concrete (giving specific examples to illustrate the meaning) or overly abstract (giving generalized explanation); appropriateness of answers.
7. Insight: Degree of personal awareness and understanding of illness.
   a. Complete denial of illness.
   b. Slight awareness of being sick and needing help but denying it at the same time.
   c. Awareness of being sick but blaming it on others, external factors, or medical or unknown organic factors.
   d. Intellectual insight: Admission of illness and recognition that symptoms or failures in social adjustment are due to irrational feelings or disturbances, without applying that knowledge to future experiences.
   e. True emotional insight: Emotional awareness of the motives and feelings within and of the underlying meaning of symptoms; whether the awareness leads to changes in personality and future behavior; openness to new ideas and concepts about self and the important people in the patient's life.
8. Judgment.
   a. Social judgment: Subtle manifestations of behavior that are harmful to the patient and contrary to acceptable behavior in the culture; whether the patient understands the likely outcome of personal behavior and is influenced by that understanding; examples of

impairment.

b. Test judgment: The patient's prediction of what he or she would do in imaginary situations; for instance, what patient would do with a stamped, addressed letter found in the street.

FURTHER DIAGNOSTIC STUDIES

A. Physical examination.
B. Neurological examination.
C. Additional psychiatric diagnostic interviews.
D. Interviews with family members, friends, or neighbors by a social worker.
E. Psychological, neurological, or laboratory tests, as indicated: EEG, CT scan, magnetic resonance imaging (MRI), tests of other medical conditions (e.g., human immunodeficiency virus [HIV]), reading comprehension and writing tests, test for aphasia, projective or objective psychological tests, dexamethasone-suppression test (DST), 24-hour urine test for heavy metal intoxication, urine screen for drugs of abuse, sleep studies in chronic insomnia.

SUMMARY OF FINDINGS

Mental symptoms, medical and laboratory findings, and psychological and neurological test results, if available, are summarized. Include the medications that the patient has been taking, their dosage, and their duration.

Clarity of thinking is reflected in clarity of writing. When summarizing the mental status, for example, the phrase "patient denies hallucinations and delusions" is not as precise as "patient denies hearing voices or thinking that he is being followed." The latter indicates the specific question asked and the specific response given. Similarly, in the conclusion of the report, one would write, "Hallucinations and delusions were not elicited."

DIAGNOSIS

Diagnostic classification is made according to the DSM-5.

PROGNOSIS

Opinion about the probable future course, extent, and outcome of the disorder; good and bad prognostic factors; specific goals of therapy.

PSYCHODYNAMIC FORMULATION

Causes of the patient's psychodynamic breakdown—influences in the patient's life that contributed to the present disorder, environmental, genetic, and personality factors relevant to determining the patient's

symptoms; primary and secondary gains; outline of the major defense mechanism used by the patient.

Table 7.2–1.
Common Questions for the Psychiatric History and Mental Status

Topic	Questions	Comments and Helpful Hints
Identifying data: Name, age, sex, marital status, religion, education, address, phone number, occupation, source of referral	Be direct in obtaining identifying data. Request specific answers.	If patient cannot cooperate, get information from family member or friend; if referred by a physician, obtain medical record.
Chief complaint (CC): Brief statement in patient's own words of why patient is in the hospital or is being seen in consultation	Why are you going to see a psychiatrist? What brought you to the hospital? What seems to be the problem?	Record answers verbatim; a bizarré complaint points to psychotic process.
History of present illness (HPI): Development of symptoms from time of onset to present; relation of life events, conflicts, stressors: drugs; change from previous level of functioning	When did you first notice something happening to you? Were you upset about anything when symptoms began? Did they begin suddenly or gradually?	Record in patient's own words as much as possible. Get history of previous hospitalizations and treatment. Sudden onset of symptoms may indicate drug-induced disorder.
Previous psychiatric and medical disorders: Psychiatric disorders; psychosomatic; medical, neurologic illnesses (e.g., craniocerebral trauma, convulsions)	Did you ever lose consciousness? Have a seizure?	Ascertain extent of illness, treatment, medications, outcomes, hospitals, doctors. Determine whether illness serves some additional purpose (secondary gain).
Personal history: Birth and infancy: To the extent known by the patient, ascertain mother's pregnancy and delivery, planned or unwanted pregnancy, developmental landmarks—standing, walking, talking, temperament	Do you know anything about your birth? If so, from whom? How old was your mother when you were born? Your father?	Older mothers (>35 years) have high risk for Down syndrome baby; older father (>45 years) may contribute damaged sperm-producing deficits including schizophrenia.
Childhood: Feeding habits, toilet training, personality (shy, outgoing), general conduct and behavior, relationship with parents or caregivers and peers, separations, nightmares, bed-wetting, fears	Toilet training? Bed-wetting? Sex play with peers? What is your first childhood memory?	Separation anxiety and school phobia associated with adult depression; enuresis associated with fire setting. Childhood memories before the age of 3 are usually imagined, not real.
Adolescence: Peer and authority relationship, school history, grades, emotional problems, drug use, age of puberty	Adolescents may refuse to answer questions, but they should be asked. Adults may distort memories of emotionally charged adolescent experience. Sexual molestation?	Poor school performance is a sensitive indicator of emotional disorder. Schizophrenia begins in late adolescence.

Adulthood: Work history, choice of career, marital history, children, education, finances, military history, religion	Open-ended questions are preferable. Tell me about your marriage. Be nonjudgmental: What role does religion play in your life, if any? What is your sexual preference in a partner?	Depending on chief complaint some areas require more detailed inquiry. Manic patients frequently go into debt or are promiscuous. Overvalued religious ideas associated with paranoid personality disorder.
Sexual history: Sexual development, masturbation, anorgasmia, erectile disorder, premature ejaculation, paraphilia, sexual orientation, general attitudes, and feelings	Are there or have there been any problems or concerns about your sex life? How did you learn about sex? Has there been any change in your sex drive?	Be nonjudgmental. Asking *when* masturbation began is a better approach than asking *do you* or *did you ever* masturbate.
Family history: Psychiatric, medical, and genetic illness in mother, father, siblings; age of parents and occupations; if deceased, date and cause; feelings about each family member, finances	Have any members in your family been depressed? Alcoholic? In a mental hospital? In jail? Describe your living conditions. Did you have your own room?	Genetic loading in anxiety, depression, schizophrenia. Get medication history of family (medications effective in family members for similar disorders may be effective in patient).
Mental status		
General appearance: Note appearance, gait, dress, grooming (neat or unkempt), posture, gestures, facial expressions. Does patient appear older or younger than stated age?	Introduce yourself and direct patient to take a seat. In the hospital, bring your chair to bedside; do not sit on the bed.	Unkempt and disheveled in cognitive disorder; pinpoint pupils in narcotic addiction; withdrawal and stooped posture in depression.
Motoric behavior: Level of activity: Psychomotor agitation or psychomotor retardation—tics, tremors, automatisms, mannerisms, grimacing, stereotypes, negativism, apraxia, echopraxia, waxy flexibility; emotional appearance—anxious, tense, panicky, bewildered, sad, unhappy: voice—faint, loud, hoarse; eye contact	Have you been more active than usual? Less active? You may ask about obvious mannerisms, e.g., "I notice that your hand still shakes, can you tell me about that?" Stay aware of smells, e.g., alcoholism/ketoacidosis.	Fixed posturing, odd behavior in schizophrenia. Hyperactive with stimulant (cocaine) abuse and in mania. Psychomotor retardation in depression; tremors with anxiety or medication side effect (lithium). Eye contact is normally made approximately half the time during the interview. Minimal eye contact in schizophrenia. Scanning of environment in paranoid states.
Attitude during interview: How patient relates to examiner—irritable, aggressive, seductive, guarded, defensive, indifferent, apathetic, cooperative, sarcastic	You may comment about attitude: You seem irritated about something; is that an accurate observation?	Suspiciousness in paranoia; seductive in hysteria; apathetic in conversion disorder (*la belle indifference*); punning (*witzlesucht*) in frontal lobe syndromes.
Mood: Steady or sustained emotional state—gloomy, tense, hopeless, ecstatic, resentful, happy, bashful, sad, exultant,	How do you feel? How are your spirits? Do you have thoughts that life is not worth living or that you	Suicidal ideas in 25% of depressives; elation in mania. Early morning awakening in depression; decreased need

elated, euphoric, depressed, apathetic, anhedonic, fearful, suicidal, grandiose, nihilistic	want to harm yourself? Do you have plans to take your own life? Do you want to die? Has there been a change in your sleep habits?	for sleep in mania.
Affect: Feeling tone associated with idea—labile, blunt, appropriate to content, inappropriate, flat	Observe nonverbal signs of emotion, body movements, facies, rhythm of voice (prosody). Laughing when talking about sad subjects, e.g., death, is inappropriate.	Changes in affect usual with schizophrenia: loss of prosody in cognitive disorder, catatonia. Do not confuse medication adverse effect with flat affect.
Speech: Slow, fast, pressured, garrulous, spontaneous, taciturn, stammering, stuttering, slurring, staccato. Pitch, articulation, aphasia, coprolalia, echolalia, incoherent, logorrhea, mute, paucity, stilted	Ask patient to say "Methodist Episcopalian" to test for dysarthria.	Manic patients show pressured speech; paucity of speech in depression; uneven or slurred speech in cognitive disorders.
Perceptual disorders: Hallucinations—olfactory, auditory, haptic (tactile), gustatory, visual; illusions; hypnopompic or hypnagogic experiences; feeling of unreality, *déjà vu, éjà entendu,* macropsia,	Do you ever see things or hear voices? Do you have strange experiences as you fall asleep or upon awakening? Has the world changed in any way? Do you have strange smells?	Visual hallucinations suggest schizophrenia. Tactile hallucinations suggest cocainism, delirium tremens (DTs). Olfactory hallucinations common in temporal lobe epilepsy.
Thought content: Delusions—persecutory (paranoid), grandiose, infidelity, somatic, sensory, thought broadcasting, thought insertion, ideas of reference, ideas of unreality, phobias, obsessions, compulsions, ambivalence, autism, dereism, blocking, suicidal or homicidal preoccupation, conflicts, nihilistic ideas, hypochondriasis, depersonalization, derealization, flight of ideas, *idé* fixe, magical thinking, neologisms	Do you feel people want to harm you? Do you have special powers? Is anyone trying to influence you? Do you have strange body sensations? Are there thoughts that you can't get out of your mind? Do you think about the end of the world? Can people read your mind? Do you ever feel the TV is talking to you? Ask about fantasies and dreams.	Are delusions congruent with mood (grandiose delusions with elated mood) or incongruent? Mood-incongruent delusions point to schizophrenia. Illusions are common in delirium. Thought insertion is characteristic of schizophrenia.
Thought process: Goal-directed ideas, loosened associations, illogical, tangential, relevant, circumstantial, rambling, ability to abstract, flight of ideas, clang associations, perseveration	Ask meaning of proverbs to test abstraction, e.g., "People in glass houses should not throw stones." Concrete answer is, "Glass breaks." Abstract answers deal with universal themes or moral issues. Ask similarity between bird and butterfly (both alive), bread and cake (both food).	Loose associations point to schizophrenia; flight of ideas, to mania; inability to abstract, to schizophrenia, brain damage.
Sensorium: Level of consciousness—	What place is this? What is	Delirium or dementia shows

alert, clear, confused, clouded, comatose, stuporous; orientation to time, place, person; cognition	today's date? Do you know who I am? Do you know who you are?	clouded or wandering sensorium. Orientation to person remains intact longer than orientation to time or place.
Memory: Remote memory (long term): Past several days, months, years	Where were you born? Where did you go to school? Date of marriage? Birthdays of children? What were last week's newspaper headlines?	Patients with dementia of the Alzheimer's type retain remote memory longer, than recent memory. Gaps in memory may be localized or filled in with confabulatory details. Hypermnesia is seen in paranoid personality.
Recent memory (short term): Recall of events in past day or two	Where were you yesterday? What did you eat at your last meal?	In brain disease, recent memory loss (amnesia) usually occurs before remote memory loss.
Immediate memory (very short term): Laying down of immediate information with ability to quickly recall data	Ask patient to repeat six digits forward, then backward (normal responses). Ask patient to try to remember three nonrelated items; test patient after 5 minutes.	Loss of memory occurs with cognitive, dissociative, or conversion disorder. Anxiety can impair immediate retention and recent memory. Anterograde memory loss (amnesia) occurs after taking certain drugs, e.g., benzodiazepines. Retrograde memory loss occurs after head trauma.
Concentration and calculation: Ability to pay attention; distractibility; ability to do simple math	Ask patient to count from 1 to 20 rapidly; do simple calculations (2 × 3, 4 × 9); do serial 7 test, i.e., subtract 7 from 100 and keep subtracting 7. How many nickels in $1.35?	Rule out medical cause for any defects vs. anxiety or depression (pseudodementia). Make tests congruent with educational level of patient.
Information and intelligence: Use of vocabulary; level of education; fund of knowledge	Distance from New York City to Los Angeles. Name some vegetables. What is the largest river in the United States?	Check educational level to judge results. Rule out mental retardation, borderline intellectual functioning.
Judgment: Ability to understand relations between facts and to draw conclusions; responses in social situations	What is the thing to do if you find an envelope in the street that is sealed, stamped, and addressed?	Impaired in brain disease, schizophrenia, borderline intellectual functioning, intoxication.
Insight level: Realizing that there are physical or mental problems; denial of illness, ascribing blame to outside factors; recognizing need for treatment	Do you think you have a problem? Do you need treatment? What are your plans for the future?	Impaired in delirium, dementia, frontal lobe syndrome, psychosis, borderline intellectual functioning.

COMPREHENSIVE TREATMENT PLAN

Modalities of treatment recommended, role of medication, inpatient or outpatient treatment, frequency of sessions, probable duration of therapy;

type of psychotherapy; individual, group, family or cognitive-behavioral therapy among others; symptoms or problems to be treated. Initially, treatment must be directed toward any life-threatening situations, such as suicidal risk or risk of danger to others, which require psychiatric hospitalization. Danger to self or others is an acceptable reason (legally and medically) for involuntary hospitalization. In the absence of the need for confinement, a variety of outpatient treatment alternatives are available: Day hospitals, supervised residences, outpatient psychotherapy, or pharmacotherapy. In some cases, treatment planning must attend to vocational and psychosocial skills training and even legal or forensic issues. Comprehensive treatment planning requires a therapeutic team approach using the skills of psychologists, social workers, nurses, activity and occupational therapists, and a variety of other mental health professionals, with referral to self-help groups (e.g., Alcoholics Anonymous [AA]) if needed. If the patient or family members are unwilling to accept the recommendations of treatment and the clinician thinks that the refusal of the recommendations may have serious consequences, then the patient, parent, or guardian should sign a statement to the effect that the recommended treatment was refused.

In Table 7.2–1, the reader will find sample questions to be used in eliciting data for the psychiatric history and mental status along with comments and helpful hints about their clinical significance.

References

*Ball C. Evidence-based medicine on the wards: report from an evidence-based minion. *ACP J Club*. 1999;130:15.

Callis R, Polmantier PC, Roeber EC. *A Casebook of Counseling*. New York: Appleton-Century-Crofts; 1955.

Coleman M, Gillberg C. *The Schizophrenias: A Biological Approach to the Schizophrenia Spectrum Disorders*. New York: Springer; 1996.

Cook L, Alegria M. Racial-ethnic disparities in substance abuse treatment: the role of criminal history and socioeconomic status. *Psychiatr Serv*. 2011;62:1273.

Dell'osso L, Shear MK, Carmassi C, Rucci P, Masser JD. Validity and Reliability of the structured clinical interview for the trauma and loss spectrum (SCI-TALS). *Clin Pract Epidemol Ment Health*. 2008;4(1):2.

Dwyer J, Shih A. The ethics of tailoring the patient's chart. *Psychiatr Serv*. 1998;49:1309.

Evans J. *Three Men*. New York: Knopf; 1954.

*Fischer CB, Fried AL. Internet-mediated psychological services and the American Psychological Association Ethics Code. *Psychother Theor Res Pract Train*. 2003;40:103.

Gersteneckera A, Masta B. Mild cognitive impairment: a history and the state of current diagnostic criteria. *Int Psychogeriatr*. 2015;27:199.

Gordon W, Perrott J, Dahri K. Accuracy of Medication histories: the first step of medication reconciliation. *Ann Pharmacother*. 2008;42(1):144.

Greist H. Computer-based assessment of patients. *J Clin Pharmacol*. 1998;18:359.

Groth-Marnat G. *Handbook of Psychological Assessment*. 4th ed. New York: John Wiley & Sons; 2003.

Jeffords JM. Confidentiality of medical information: protecting privacy in an electronic age. *Prof Psychol Res Pract*. 1999;30:115.

Kellera B, Heddenc T. Effects of physician awareness of symptom-related expectations and mental disorders. *Cortex*. 2015;64:271.

Krauss JB. A matter of privacy. *Arch Pract Nurs*. 2003;17:99.

Lysaker PH, Wickett AM, Campbell K, Buck KD. Movement toward coherence in the psychotherapy of schizophrenia: a method for assessing narrative transformation. *J Nerv Ment Dis*. 2003;191:538.

*Mark D. Effects of physician awareness of symptom-related expectations and mental disorders: a controlled trial. *Arch Fam Med*. 1999;8:135.

Park E, Abraham MK. Altered mental status and endocrine diseases. *Emerg Med Clin N AM*. 2014;32:367

Pierluissi E, Fischer MA, Campbell AR, Landefeld CS. Discussion of medical errors in morbidity and mortality conferences. *JAMA*. 2003;290:2838.

Shea SC. *Psychiatric Interviewing*. 2nd ed. Philadelphia, PA: WB Saunders; 1998.

Wedding D, Corsini RJ. *Case Studies in Psychotherapy*. Itaska, IL: Peacock; 1989.

Young HM, Gray SL, McCormick WC, Sikma SK, Reinhard S. Types, prevalence and potential clinical significance of medication administration errors in assisted living. *J Am Geriatr Soc*. 2008;56:1199.

Zweig RA, Turkel E. The social adjustment scale-self-report: psychometric properties for older adults. *Psychol Rep*. 2007;101:920.

▲ 7.3 Practice Guidelines in Psychiatry

LAURA J. FOCHTMANN, M.D., AND JOHN S. MCINTYRE, M.D.

Over the past half century there have been great advances in the knowledge base for all of medicine including psychiatry. Much has been learned about the etiology and epidemiology of mental illnesses and the efficacy and effectiveness of treatments. Evidence-based medicine is the process of translating those advances into the day-to-day care of patients. However, a large number of studies and reports have documented that effective interventions are not always implemented in a timely fashion, if at all. Compounding this problem, especially in the United States, is the reality that reimbursement for services is not necessarily driven by what has been shown to be optimal care. These issues can be addressed using practice guidelines, which are sets of evidence-based strategies that are aimed at assisting physicians and patients in clinical decision making and improving care.

EARLY GUIDELINE EFFORTS

Throughout the history of medicine there have been recommendations for the care of patients, which have been promulgated orally or in written form. In 1851, 7 years after its founding, the American Association of Medical Superintendents of Hospitals for the Insane (later to become the American Psychiatric Association [APA]) issued 26 propositions concerning "the construction of hospitals for the insane." Two years later 14 propositions concerning the organization of these facilities were published. These propositions were intended to provide guidance about psychiatric care but did not include documentation of the evidence behind the propositions or the process by which they were developed. Early efforts at promulgating evidence-based guidelines in the United States included the publications of the American Academy of Pediatrics on immunizations in the 1940s. Over the subsequent four decades the number of guidelines began to increase and by the mid-1980s the growth has become exponential.

AMERICAN MEDICAL ASSOCIATION AND GUIDELINES

In the late 1980s the American Medical Association (AMA) convened a group of representatives from 14 medical specialty organizations, including the American Psychiatric Association. Also attending on a regular basis were representatives from the Agency for Health Care Policy and Research, Joint Commission on Accreditation of Health Care Organizations, American Hospital Association, National Council on Quality Assurance,

and the Health Care Financing Administration. This consortium was initially named the Practice Parameter Partnership and later renamed the Practice Guideline Partnership. In 1989, the Partnership described five recommended attributes of guidelines: (1) be developed by or in conjunction with physician organizations, (2) use reliable methodologies that integrate relevant research findings and clinical expertise, (3) be as comprehensive and specific as possible, (4) be based on current information, (5) be widely disseminated. In 1996, a sixth attribute was added: Guideline development should include outcomes research, goals, and measures. The focus on measures was spurred by the increased interest in assessing whether guidelines were followed and led to the development of the AMA-led Physician Consortium for Performance Improvement (PCPI).

In 1999, the AMA and the American Association of Health Plans collaborated with the Agency for Healthcare Research and Quality (AHRQ) of the United States Department of Health and Human Service in establishing the web-based National Guidelines Clearinghouse (http://www.guideline.gov/). More than a decade later, this initiative continues to fulfill its goals of fostering dissemination and use of practice guidelines by multiple stakeholders.

INSTITUTE OF MEDICINE REPORTS

In parallel with the increasing involvement of professional organizations in developing practice guidelines, the Institute of Medicine (IOM) began to discuss and define optimal characteristics of practice guidelines. In 1990, the IOM published a monograph: *Clinical Practice Guidelines—Directions for a New Program*. Incorporating some of the recommendations included in the AMA report, "good guidelines" were identified as having eight attributes as identified in Table 7.3–1.

This IOM report describes practice guidelines as "valid if, when followed, they lead to the health and cost outcomes projected for them, other things being equal." Practice guidelines are defined as "reliable and reproducible" if a different group of developers, given the same evidence and using the same methods, arrive at the same recommendations and if the guidelines are interpreted and applied consistently by clinicians in similar settings. "Clinical applicability" requires that the population being treated with the guideline be explicitly identified. A guideline has "flexibility" if the generally expected exceptions are identified and "clarity" if terms are well defined and unambiguous language is used. The report indicates that a "multidisciplinary process" may include "[participants from different disciplines] serving on panels that develop guidelines, providing evidence and viewpoints to the panels and reviewing draft guidelines." The report also notes that practice guidelines should include specific plans about "review and revision" schedules and that there should be "meticulous

documentation and description" of the evidence used, the process followed, and the individuals involved in developing guidelines.

Two decades later, a series of IOM reports expanded on these recommendations for practice guideline development. A 2008 report, *Knowing What Works in Health Care: A Road Map for the Nation*, described the reasons that changes were needed in synthesizing of information and using it for clinical decisions. The report cited information overload and health policy challenges such as unsustainable health care costs and geographic variations in the amount and quality of care. It also noted that available systematic reviews and clinical practice guidelines had not fulfilled their original promise due to a lack of universal development standards, variations in evidence appraisal and inconsistent handling of conflicts of interest. This report served as the impetus for the Medicare Improvements for Patients and Providers Act of 2008 to request two additional IOM reports, *Finding What Works in Health Care: Standards for Systematic Reviews* and *Clinical Practice Guidelines We Can Trust*.

Table 7.3–1.
Guideline Attributes

Validity
Reliability/reproducibility
Clinical applicability
Clinical flexibility
Clarity
Multidisciplinary process
Scheduled review
Documentation

Data from Field MJ, Lohr KN, eds. *Clinical Practice Guidelines—Directions for a New Program*. Washington, DC: National Academies Press; 1990.

For systematic reviews, the IOM gives 21 recommendations to promote objectivity, transparency, and reliability in the review process. These recommendations include, but are not limited to, establishing a systematic review team with the appropriate expertise and experience, managing bias and conflict of interest among team members and those providing input into the review, obtaining public input on the review at multiple steps in the process, developing a protocol for the review that is established a priori, and using clear and transparent processes for a comprehensive search and for evidence review and synthesis. In the complementary report on clinical practice guidelines, the IOM gives eight additional recommendations that discuss the importance of transparency and management of conflict of interest, appointing a balanced multidisciplinary guideline development group, using rigorously conducted systematic reviews, providing ratings of the strength of recommendations based on the underlying quality of available evidence, and wording guideline recommendations in precise

actionable terms. The IOM also notes that practice guidelines should be subjected to public review and regular updating. Many of these precepts of guideline development were subsequently affirmed by the Council of Medical Specialty Societies in their "Principles for the Development of Specialty Society Clinical Guidelines."

PRACTICE GUIDELINES IN PSYCHIATRY

One of the first psychiatric associations to publish guidelines was the Royal Australian and New Zealand College of Psychiatrists (RANZCP) beginning in 1981. This project initially had a unique methodology for blending research findings and clinical consensus. A group of research experts developed a guideline for a particular disorder and a group of clinicians developed a guideline for the same disorder and then the two guidelines were integrated into a single document.

Other psychiatric guideline efforts soon followed. In the United States, the American Psychiatric Association has been developing guidelines since 1989 (see below) and the American Academy of Child and Adolescent Psychiatry (AACAP) published its initial Practice Parameters in 1991 (http://www.jaacap.com/content/pracparam?journalCode = jaac). The Veterans Administration also has a robust guideline development process, which includes psychiatric guidelines (http://www.healthquality.va.gov/). Other agencies and professional organizations that have developed guidelines of relevance to psychiatry include, but are not limited to, the American College of Physicians, the American Psychological Association, the Canadian Psychiatric Association, Canadian Network for Mood and Anxiety Treatments, The National Institute for Health and Care Excellence (NICE), the Scottish Intercollegiate Guidelines Network, and The World Federation of Societies of Biological Psychiatry.

APA PRACTICE GUIDELINE PROJECT

History

The American Psychiatric Association Practice Guideline project began in 1989. An action paper in the APA Assembly called for the APA to develop practice guidelines to assist psychiatrists in their care of patients. Following approval by the APA Board of Trustees and President, a Steering Committee was formed to create a process to develop guidelines and provide ongoing oversight.

Guideline Committees and Work Groups

Early in the APA project, the decision was made that no industry money would be sought or accepted for the guideline development. The cost of development has been largely offset by publication revenues. In addition, work group and practice guideline committee members are carefully

screened for potential conflicts of interest in order to limit the risk of bias in guideline recommendations. Several APA staff, including the medical editor of the guidelines, have been assigned to the project. With their participation and the volunteer efforts of countless work group members and committee members, the project has continued to flourish.

For prior practice guidelines, work groups were composed of psychiatrists who were in active clinical practice and who were appointed by the APA President on the recommendation of the Steering Committee. These work group members were chosen on the basis of clinical and research expertise relating to the guideline topic. Based upon the recommendations of the IOM, future guidelines are being written by a multidisciplinary group of individuals, some of whom have specific expertise on the guideline topic but others of whom have more general expertise in the treatment of psychiatric disorders. Consultants to the writing group include additional subject matter experts as well as patient and family representatives.

Choosing a Topic

The initial step in developing a guideline is to choose a topic. At the onset of the APA project it was decided to develop guidelines for specific illnesses/disorders as defined by the fourth revised edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR). Having a criteria-based diagnostic nomenclature was a major asset in starting the project. To date there have been two exceptions to disorder-focused guidelines: psychiatric evaluation and assessment and treatment of patients with suicidal behaviors.

Previously, the Steering Committee selected an illness/disorder that represented a significant public health problem (high prevalence and seriousness), was relevant to psychiatric practice, and for which there was a substantial body of evidence to support recommendations. Future topic selections review these factors but also consider whether rigorous systematic reviews are already available (e.g., through the AHRQ Evidence-Based Practice centers) that could inform guideline development. Topics for future guidelines may also be more circumscribed than prior guidelines; future guidelines may address one aspect of a disorder or one class of treatments for a disorder rather than a more comprehensive view of a disorder and its treatment. An additional change from prior guidelines is that topics are framed in form of a PICOTS question, with a specifically defined Population, Intervention, Comparator, Outcome, Timing, and Setting. This PICOTS question will shape the identification of evidence and subsequent steps of guideline development.

Identifying the Evidence

As noted above, guidelines must describe specific interventions and document the evidence that supports these recommendations. The literature search process, which is clearly delineated in each guideline, identifies keywords, employs a MEDLINE search, and then includes additional specific searches using specialized databases such as PsycINFO, Psychoanalytic Electronic Publishing and the Cochrane Library. Future guidelines may rely on available systematic reviews or on systematic reviews conducted by APA staff, adhering to approaches delineated by the IOM or the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews.

Table 7.3–2.
Grading of the Quality of Individual Studies

Population
Intervention
Comparators
Outcomes
Timing
Setting
Study design
Overall risk of study bias
Selection bias
Performance bias
Attrition bias
Detection bias
Reporting bias
Sponsor-related bias
Applicability

Adapted from Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. 2014. Available at www.effectivehealthcare.ahrq.gov.

With prior guidelines, literature search results were given to work group members, who then created evidence tables to facilitate review of study results, discuss the clinical implications of the research, and assign an overall rating for the strength of evidence that supported the guideline recommendations.

Beginning in 2014, systematic reviews (such as those conducted by Evidence-Based Practice centers) will be used whenever possible as a basis for guideline development. In other circumstances, literature search results will be doubly screened by members of a systematic review group based upon study design and relevance to the guideline topic. Using the methodology outlined by the AHRQ, which has many similarities to the methodology used by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) group, studies that meet inclusion criteria are assessed for aspects of study design, findings, and quality as

shown in Table 7.3–2.

Synthesis of the findings of available research also follows a defined process as outlined in Table 7.3–3.

The overall strength of research evidence is not a measure of the size of the effect of the intervention but rather reflects the degree of confidence that additional studies will show similar conclusions. Factors that enter into this determination include the consistency of findings across studies, the directness of the effect on a specific health outcome, precision of the estimate of the effect, and the risk of bias in available studies.

Table 7.3–3.
Grading of the Overall Supporting Body of Research Evidence

Overall strength of research evidence
Risk of bias
Consistency
Directness
Precision
Dose–response relationship (if applicable)
Magnitude of effect
Confounding factors (including likely direction of effect)
Applicability

Adapted from Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. 2014. Available at www.effectivehealthcare.ahrq.gov.

Developing Guideline Recommendations

The drafting of practice guideline recommendations is a natural outgrowth of the review of the available evidence. In prior guidelines, recommendations were determined by the work group and could be modified based on input obtained during reviews of the guideline drafts (see below). There were three levels of recommendation: (I) Recommended with substantial clinical confidence; (II) Recommended with moderate clinical confidence, and (III) May be recommended on the basis of individual circumstances.

In contrast, using the GRADE methodology, guideline statements are based upon the relative balance of benefits and harms of an intervention. Thus, a recommendation indicates confidence that benefits of an intervention clearly outweigh harms whereas a suggestion indicates a greater degree of uncertainty, either than the balance of benefits and harms is more difficult to judge or that either the benefits or the harms are unclear.

Based on the recommendations of the IOM, guideline statements are crafted in specific actionable terms that describe the circumstances or patient populations in which an intervention is to occur, the object of the

action (if not the identified patient) and the individual carrying out the action (if not the psychiatrist or treating clinician). Whenever possible vague terms (e.g., moderate to severe symptoms) are explicitly defined or eliminated.

Beginning in 2014, these guideline recommendations are developed by the guideline writing group using an iterative Delphi process with discussion followed by blinded voting. If consensus is not reached, additional rounds of discussing and voting occur.

Some guideline ventures have attempted to limit guideline recommendations to those supported by randomly controlled trials (RCTs). These efforts had to be modified because there is much physicians do (such as a detailed diagnostic assessment) that has not been and cannot be studied with RCTs. Also, especially in psychiatry, RCTs and other trials often are limited to acute care (6 to 12 weeks) and because of funding issues generally focus on pharmacotherapy or other somatic treatments. In addition, significant research challenges exist when studying some treatments, such as psychodynamic psychotherapy and some social therapies. For these reasons expert clinical consensus is an essential component in the development of a guideline.

In an effort to make the expert consensus process more rigorous, APA has pioneered a "snowball" sampling methodology to identify clinical and research experts based on nomination of their peers. Nominated experts then receive a web-based survey with questions that parallel the guideline PICOTS questions and answers to the expert survey are reviewed by the guideline writing group. These expert survey results do not substitute for rigorous clinical trials but they are particularly informative for situations in which clinical trials are not possible.

Format of APA Practice Guideline

Each APA practice guideline published between 1999 and 2014 followed a standardized format that included: (I) Executive Summary, (II) Formulation and Implementation of a Treatment Plan, (III) Specific Clinical Features Influencing the Treatment Plan, (IV) Disease Definition, Epidemiology, and Natural History, and (V) Review and Synthesis of Available Evidence. A discussion of Future Research Needs was also included.

In the first few APA guidelines, interventions were divided into psychotherapy and somatic treatments. However, it soon became apparent that a significant portion of the care provided by psychiatrists did not fall into either category. The term psychiatric management (not an ideal term but to date a better term has not been suggested) was introduced to describe those activities and interventions that are essential for good care but are not either a specific psychotherapy or a somatic treatment. The elements of psychiatric management were included in the section of the

guideline on formulation and implementation of the treatment plan. The specific aspects of psychiatric management varied somewhat from one diagnosis to another but the large majority of themes were consistent across diagnostic categories. For example, the elements of psychiatric management for patients with major depressive disorder included: Performing a diagnostic evaluation, evaluating the safety of the patient and others, evaluating the level of functional impairments, determining a treatment setting, establishing and maintaining a therapeutic alliance, monitoring the patient's psychiatric status and safety, providing education to patients and families, enhancing treatment adherence, and working with patients to address early signs of relapse. The implications of comorbidity for treatment planning were another common theme across guidelines and were addressed in the section on *Specific Clinical Features Influencing the Treatment Plan*.

Beginning in 2014, the format of the APA practice guidelines shifted to a different format aimed at fulfilling the recommendations of the IOM and also making it easier to deliver clinical decision support through electronic resources. The new format is centered around specific recommendations and starts with a set of related guideline statements (e.g., recommendations, suggestions) accompanied by ratings of the strength of the recommendation and the strength of the supportive research evidence. A section on potential benefits and harms of the guideline statement is aligned with the GRADE methodology and describes the pluses and minuses of the intervention. This information is useful in understanding why the guideline statement was made but also can be useful for patient-centered care in terms of shared decision making and, where relevant, informed consent. A section on the goal and rationale of the guideline review aspects of the guideline purpose, the epidemiology and public health implications of the disorder or problem, any variability of care in clinical practice, and associated morbidity/mortality if high-quality care is not received. This section is intended to show clinicians, insurers and policy makers why the guideline recommendations are important to patients, their families, health professionals, and society. A section on implementation discusses approaches for integrating the recommendations or suggestions into routine care, including discussion of barriers to integration and ways to overcome them. This section also incorporates a discussion of circumstances in which guideline recommendations may not be appropriate, due to factors such as comorbid physical or psychiatric condition. It also discusses relevant aspects of psychiatric management and ways in which patient values and preferences may shape recommended treatment approaches. To assure transparency of the guideline development process, the review of the supporting research evidence is included in the guideline in detail. A new content area in each guideline relates to quality measurement and discusses existing quality measures

that are related to the guideline topic. If existing quality measures do not exist or are insufficient, this section of the guideline also discusses whether a quality measure should be derived from the recommendation and, if so, how it should be defined.

Review and Approval Process

The writing, review, and editing of a guideline is an iterative process. The number of drafts has evolved over the history of the practice guideline project and current processes focus on allowing a thorough review without untoward delays during which evidence becomes outdated. With the current process, the guideline draft is prepared by the guideline writing group and then reviewed by a large number of individuals including 50 selected national and international expert reviewers, members of APA components and assembly, District Branches, and allied organizations. Allied organizations include other medical specialties (e.g., neurology and family medicine), other mental health specialties (nursing, psychology, social work), and advocacy groups (e.g., National Alliance on Mental Illness [NAMI]). Reviewers are asked to make specific recommendations and if possible provide evidence to support their suggestions. The large number of reviewers (each of whom is asked to identify any potential conflict of interest) is a further protection against a conflict of interest created by the bias of a work group member surviving to the final draft. Each suggestion is considered by the work group chair, the medical editor of the project, the project manager, and often the chair of the Steering Committee. The final draft is then prepared and is sent to the APA Assembly and the Board of Trustees (BOT) for approval. Hence each guideline carries the authority of the Association, and this adds to its acceptance by the field and also aids in its dissemination.

Dissemination and Implementation Strategies

Over the initial 25 years of the practice guidelines project, the APA has approved and published 14 guidelines, with an additional 7 guidelines undergoing revision on at least one occasion. Using the new development methodology, Practice Guidelines for the Psychiatric Evaluation of Adults were approved in December 2014 and additional development efforts are underway.

From the beginning of the project much attention has been directed to dissemination and implementation strategies. After approval each guideline or a key portion of the guideline is published in the *American Journal of Psychiatry*. Many of the guidelines have also been published in a *Compendium of APA Guidelines*. Quick Reference Guides were also developed to present major recommendations in algorithmic or bullet form and to improve utility of the guidelines in the day-to-day care of patients.

With the same goal, some guidelines have also been formatted as six to eight panel pocket guides following a model developed by the Veterans Administration.

Each approved guideline is published in its entirety online on the APA website (http://www.psych.org). In addition, a description of the guideline appears on the federal government's National Guideline Clearing House site (http://www.guidelines.gov) with a link to the APA site. Also online at the APA site is a continuing education course that accompanies each guideline. This continuing medical education (CME) program has been approved by the American Board of Psychiatry and Neurology (ABPN) as part of a comprehensive lifelong learning program, which is mandated by the American Board of Medical Specialties as a necessary component of maintenance of certification (MOC).

Updating of Guidelines

Because of the rapid growth of knowledge, keeping guidelines current is one of the major challenges facing this or any other guideline project. APA Practice Guidelines were scheduled for revision every 3 to 5 years. Because it takes about 2 years to develop a guideline, even if this schedule is followed (which has not consistently happened) guidelines may be 7 years old before a revision is published. In an effort to maintain currency of the recommendations, guideline Watches were developed. While they are not approved by the APA Assembly and BOT, Watches reflect new data, either published studies or clinical consensus that should be considered by the clinician in using a guideline. APA Watches are published online as soon as they are approved by the Steering Committee.

The change to a new format for guidelines was, in part, motivated by the need for continual updating of evidence and current recommendations. When guidelines encompassed all aspects of treatment for a given disorder, evidence and recommendations on one treatment might be outdated even when the remainder of the guideline was still current. When guidelines are based on more discrete topics, they can be updated independently of other topics that relate to the same disorder. APA is also exploring approaches to use informatics approaches to facilitate screening of new literature in order to support targeted updating of guidelines.

ADVANTAGES AND LIMITATIONS OF GUIDELINES

Table 7.3–4 identifies a number of potential advantages and limitations of guidelines.

Advantages of Guidelines

The most important benefit, and the major reason for developing and using guidelines, is to improve the quality of care received by patients. A growing

body of literature, particularly related to cardiovascular disease, suggests that guideline adherence results in better outcomes. However, the available evidence is less clear in psychiatry, at least in part due to limited guideline adherence. Such studies are difficult to conduct especially when the guideline recommendations are multi-faceted. Furthermore, rates of patient participation and adherence can confound interpretation of findings.

Geographic variations are well documented for rates of hospitalizations, procedures (such as hysterectomy or coronary artery bypass grafts), medication use, and other aspects of health care utilization. Within psychiatry, one study suggests that rates of death due to depression or schizophrenia also show significant geographic variation. Although this could be biased by differences in ascertainment, it may also reflect variations in care that are associated with variations in morbidity and mortality. Having nationally developed guidelines (if they are followed) can presumably decrease such variation and improve outcomes for patients.

Table 7.3–4.
Advantages and Limitations of Guidelines

Advantages of Guidelines
Improve quality of care
Decrease variation in practice
Standardize and facilitate education of professionals
Education of patients and family
Facilitates patient involvement in decision making
Stimulate research
Encourage funding of research
Support reimbursement of evidence-based care
Reduce stigma
Limitations of Guidelines
Gaps in knowledge
Reality of comorbidity
Must change with new knowledge
Cookbook rather than art
Risk of reductionism
Stifle innovation
Concern about increased liability
Cost of development and implementation
Biased by funders of development
Difficult to implement

Another major advantage of guidelines is their use in educational programs. Guidelines that include a clear and comprehensive description of the evidence and references are especially helpful. Interestingly, in many training programs it is the residents who push for increased use of the guidelines. Content related to the guidelines is also included in The Psychiatry Residents-in-Training Examination (PRITE) administered by

the American College of Psychiatrists (ACP). Use of guidelines for educational purposes extends well beyond residency training into CME and MOC programs. The ABPN has designated the APA guidelines as a major source for their MOC examinations and questions related to the guidelines are incorporated into the FOCUS self-assessment programs. Performance-in-practice modules are also available that are linked to guideline content.

Education of patients and families is also facilitated by the existence of guidelines. Many developers of guidelines also prepare a patient and family companion document that parallel the guideline but is written in a manner understandable for a lay audience. Such material also aids in involving the patient in the decision making, a major element in patient-centered care.

As guidelines are written, gaps in our evidence base become painfully clear. An important section in many guideline projects is to identify those issues that need to be further studied. This can provide direction to the field and also can be used to support requests for additional research funding.

To clinicians, decisions by third-party payers about what treatments are reimbursed frequently seem arbitrary or not driven by clinical reasoning. Practice guidelines can be used to identify those treatments that are supported by evidence and should be reimbursed and those treatments that are shown to be ineffective and should not be reimbursed. They can also delineate situations in which a number of options are appropriate and in which patient preferences and values should be taken into consideration and supported.

Guidelines are also valuable in demonstrating that mental illnesses are real and diagnosable and that specific disorders and symptoms can be ameliorated with specific treatments. This is a crucial role for guidelines because stigma and its effects on patients, families, and treatment continue to be a major challenge. Over the past three decades recognition of these realities has had a significant impact, directly and indirectly, on the funding of and access to treatment of mental illnesses.

Limitations of Guidelines

In terms of limitations to practice guidelines, gaps in the knowledge base make it challenging to write authoritative guidelines without overstating what the evidence supports. Poorly written guidelines can mislead physicians and patients and discredit the scientific underpinnings of our field.

Guidelines that focus on a single illness/disorder are limited by the reality that many patients have multiple illnesses. In psychiatry, at least 50 percent of patients have more than one psychiatric diagnosis. Physical conditions are also common and can require modifications in treatment or monitoring. Hence, single-illness guidelines must address these realities, and more guidelines should be developed for comorbid conditions.

Guidelines must be regularly revised as new knowledge becomes available. As the medical literature continues to grow at exponential rates and as the number of guidelines in a specialty area increases, this becomes a significant challenge. However, if guidelines are not updated regularly, they can be misleading to clinicians, patients, families, and policy makers.

Care is also needed to reduce the number of inconsistent recommendations among guidelines written by different groups. Often these inconsistencies result from differences in the patient population of interest (e.g., primary care vs. specialty practices) or legitimate differences of opinion in interpreting the available evidence. Nevertheless, conflicting recommendations can be confusing for patients and clinicians alike. Transparency in the synthesis of evidence and rationale for recommendations can help to reduce some of these difficulties.

A frequent criticism of guidelines, especially when they were first being introduced, is that they encourage "cookbook" medicine and interfere with the practice of the "art" of medicine by being overly prescriptive. Well-written guidelines can avoid this pitfall by presenting the existing evidence and explicitly stating the role of the physician in using the recommendations to inform decision making about care.

A related issue is concern about guidelines reinforcing reductionism. Guidelines can potentially decrease the richness and breadth of assessment and treatment strategies, and developers of guidelines must be alert to this potential. Another concern is that new treatment approaches will be discouraged and cutting-edge innovations will not be pursued. To address both of these concerns, guidelines may need to be more inclusive and less dogmatic, always describing the available evidence clearly and noting that other approaches may also constitute proper care for the disorder being considered.

Clinicians have been concerned that the presence of guidelines would lead to an increase in successful malpractice actions as plaintiff attorneys could point out discrepancies between guideline recommendations and rendered care. Although evidence is limited, it appears that guidelines have been used by defendants as well as by plaintiffs to support their cases without significant changes in the number or severity of successful malpractice claims. To mitigate the possibility for misuse of guidelines in litigation, guideline writers should avoid being overly prescriptive unless the evidence is overwhelming and other choices are not reasonable. It must also be clearly stated that the recommendations are indeed guidelines not standards of care. Instead, recommendations should be considered by the physician and decisions should be based on the needs of the individual patient, the available diagnostic and treatment options and other circumstances.

Another limitation of guideline development is cost, which can be considerable. The federal government's Agency for Health Care Policy and

Research developed 19 guidelines from 1992 to 1996, but the program was discontinued in part because of the considerable cost. As more scientific literature is available for review and as systematic review standards have become more rigorous, the costs of developing a guideline are even higher than in prior decades. While some of this cost can be offset by publication revenue from the guideline and related materials and income from educational programs, developing and disseminating guidelines is generally not a profitable undertaking. Some non-APA guidelines have been developed with the support of industry, especially pharmaceutical companies. This financial support raises a concern about the scientific integrity of the process and the independence of the recommendations.

Costs of implementation can also play a role in reducing guideline adoption with several studies noting that lack of administrative support and insufficient staff time were significant barriers. More recently, there has been increasing integration of guideline recommendations and related quality measures into electronic health records (EHRs), which can also add to implementation costs for software modification and staff time. In addition to these costs of guideline implementation, opportunity costs are often overlooked. If clinicians are focusing on delivering guideline-concordant care, they may overlook or lack the time to focus on other issues that would be of greater benefit to a particular individual.

Other difficulties in guideline implementation are also important to note. There are numerous studies indicating that often physicians do not use published guidelines. M.D. Cabana identified six primary barriers to guideline use: Lack of awareness, lack of familiarity, lack of agreement, lack of self-efficacy, lack of outcome expectancy, and inertia of previous practice. Some of these barriers can be surmounted through specific interventions. For example, some studies have shown that patient involvement in medical decisions increases the probability of receiving guideline concordant care. Other studies found that physician feedback or multifaceted educational programs including the use of opinion leaders improves the rate of physicians using guidelines. It has also been shown that use of guidelines may be enhanced by incorporating recommendations into EHRs. J.M. Grimshaw concluded that multifaceted interventions targeting different barriers to change in physician behavior are more likely to be effective than single interventions.

THE FUTURE OF PRACTICE GUIDELINES

There are now thousands of guidelines for all areas of medicine, which have been developed by professional associations, insurance companies, health care systems, and regional and national governments. The number of new guidelines will continue to increase dramatically over the next decade and will play and increasing role in the delivery of health care. With the integration of guidelines into EHRs recommendations can be presented to

clinicians at the point of service with more direct effects on clinical decisions. More research is needed on other tools and incentives to increase the use of guidelines. At the international level, efforts will increase to explore the applicability of guidelines among different countries and to identify those recommendations that should be modified to address the needs of different cultures and health care delivery systems. If guidelines are well developed, truly evidence based, and appropriately flexible to meet the needs and wishes of individual patients, their use can significantly improve the quality of care for patients.

REFERENCES

AGREE Coalition. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: The AGREE project. *Qual Saf Health Care*. 2003;12:18.

Ament SM, de Groot JJ, Maessen JM, Dirksen CD, van der Weijden T, Kleijnen J. Sustainability of professionals' adherence to clinical practice guidelines in medical care: a systematic review. *BMJ Open*. 2015;5(12):e008073.

American Psychiatric Association, Standing Committee. Report on the organization of hospitals for the insane. *Am J Insanity*. 1853;67.

American Psychiatric Association. *Practice Guidelines for the Treatment of Psychiatric Disorders: Compendium 2006*. Washington, DC: American Psychiatric Association; 2006.

Andrews G, Armstrong MS, Brodaty H. A methodology for preparing 'ideal' treatment outlines in psychiatry. *Aust N Zeal J Psychiatry*. 1982;16:153.

Boyd CM, Darer J, Boult C, Fried LP, Boult L. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA*. 2005;294:716.

Cabana MD, Rushton JL, Rush AJ. Implementing practice guidelines for depression: applying a new framework to an old problem. *Gen Hosp Psychiatry*. 2002;24:35.

Chassin MR, Brook RH, Park RE, Keesey J, Fink A. Variations in the use of medical and surgical services by the Medicare population. *N Engl J Med*. 1986;314:285.

Council of Medical Specialty Societies. Principles for the Development of Specialty Society Clinical Guidelines. 2015. Available at http://www.cmss.org/uploadedFiles/Site/CMSS_Policies/CMSS%20Principles%20for%20the%
%20September%202012%282%29.pdf.

Gaebel W, Weinmann S, Sartorius N, Rutz W, McIntyre JS. Schizophrenia practice guidelines: International survey and comparison. *Br J Psychiatry*. 2005;187:248.

Gagliardi AR, Alhabib S; members of Guidelines International Network Implementation Working Group. Trends in guideline implementation: a scoping systematic review. *Implement Sci*. 2015;10:54.

*Gagliardi AR, Brouwers MC, Palda VA, Lemieux-Charles L, Grimshaw JM. How can we improve guideline use? A conceptual framework of implementability. *Implement Sci*. 2011;6:26.

Girlanda F, Fiedler I, Becker T, Barbui C, Koesters M. The evidence-practice gap in specialist mental healthcare: systematic review and meta-analysis of guideline implementation studies.

Br J Psychiatry. 2016. [Epub ahead of print].

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–926.

Hepner KA, Rowe M, Rost K, Hickey SC, Sherbourne CD. The effect of adherence to practice guidelines on depression outcomes. *Ann Intern Med*. 2007;147:320.

Institute of Medicine, Committee on Clinical Practice Guidelines. *Guidelines for Clinical Practice: From Development to Use*. Washington, DC: National Academy Press; 1990.

*Institute of Medicine. In: Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E. *Committee on Standards for Developing Trustworthy Clinical Practice Guidelines; Clinical Practice Guidelines We Can Trust*. Washington, DC: National Academies Press (US); 2011.

Institute of Medicine. In: Eden J, Levit L, Berg A, Morton S. *Committee on Standards for Systematic Reviews of Comparative Effectiveness Research. Finding What Works in Health Care: Standards for Systematic Reviews*. Washington, DC: National Academies Press (US); 2011.

Institute of Medicine. *Knowing What Works in Health Care: A Roadmap for the Nation*. Washington, DC: The National Academies Press; 2007.

Krahn M, Naglie G. The next step in guideline development: incorporating patient preferences. *JAMA*. 2008;300(4):436–438.

*Methods Guide for Effectiveness and Comparative Effectiveness Reviews. *AHRQ Publication No. 10(14)-EHC063-EF*. Rockville, MD: Agency for Healthcare Research and Quality. 2014. Available at www.effectivehealthcare.ahrq.gov.

Polednak AP. US regional differences in death rates from depression. *Soc Psychiatry Psychiatr Epidemiol*. 2012;47(12):1977–1983.

Pronovost PJ. Enhancing physicians' use of clinical guidelines. *JAMA*. 2013;310(23):2501–2502.

Reus VI, Fochtmann LJ, Eyler AE, Hilty DM, Horvitz-Lennon M, Jibson MD, Lopez OL, Mahoney J, Pasic J, Tan ZS, Wills CD, Rhoads R, Yager J. The American Psychiatric Association Practice Guideline on the Use of Antipsychotics to Treat Agitation or Psychosis in Patients With Dementia. *Am J Psychiatry*. 2016;173(5):543–546.

Rosenfeld RM, Shiffman RN, Robertson P. Clinical Practice Guideline Development Manual, Third Edition: a quality-driven approach for translating evidence into action. *Otolaryngol Head Neck Surg*. 2013;148(1 Suppl):S1–S55.

*Silverman JJ, Galanter M, Jackson-Triche M, Jacobs DG, Lomax JW 2nd, Riba MB, Tong LD, Watkins KE, Fochtmann LJ, Rhoads RS, Yager J; American Psychiatric Association. The American Psychiatric Association Practice Guidelines for the Psychiatric Evaluation of Adults. *Am J Psychiatry*. 2015;172(8):798–802.

Van Duin D, Franx G, Van Wijngaarden B, et al. Bridging the science-to-service gap in schizophrenia care in the Netherlands: the Schizophrenia Quality Improvement Collaborative. *Int J Qual Health Care*. 2013;25(6):626–632.

*Weinmann S, Koesters M, Becker T. Effects of implementation of psychiatric guidelines on provider performance and patient outcome: systematic review. *Acta Psychiatr Scand*. 2007;115(6):420–433.

Yager J, Kunkle R, Fochtmann LJ, et al. Who's your expert? Use of an expert opinion survey to

inform development of American Psychiatric Association practice guidelines. *Acad Psychiatry*. 2014;38:376–382.
Zarin DA, Pincus HA, McIntyre JS. Practice guidelines. *Am J Psychiatry*. 1993;150:175.

▲ 7.4 Clinical Neuropsychology and Intellectual Assessment of Adults

REX M. SWANDA, PH.D., AND KATHLEEN Y. HAALAND, PH.D.

Clinical neuropsychology is a specialty in psychology that examines the relationship between behavior and brain functioning in the domains of cognitive, motor, sensory, and emotional functioning. The clinical neuropsychologist integrates the medical and psychosocial history with the reported complaints and the pattern of performance on neuropsychological procedures in order to determine whether results are consistent with a particular area of brain damage or a particular diagnosis. Although neurological syndromes are often the focus of referrals, the neuropsychological examination has a valuable place in diagnosing and identifying behavioral symptoms that can be associated with various medical, psychological, and psychiatric conditions. Neuropsychology also plays a role in medico-legal and forensic applications, such as determination of decisional capacity, disability determination, and litigation of injury.

WHO IS A CLINICAL NEUROPSYCHOLOGIST

As the field of neuropsychology has grown, clinical practitioners have come to recognize the importance of developing explicit guidelines for standards of training and specialized practice in neuropsychology. In 1996, the field of neuropsychology was formally recognized as a distinct specialty in psychology by the American Psychological Association. The 1997 Houston Conference resulted in a policy statement that formally identified training experiences that would be appropriate to the development of specialization in clinical neuropsychology. These training guidelines call for both specialized and generalized clinical training throughout a program of doctoral studies, followed by internship and postdoctoral residency, leading to eligibility for specialty board certification through the American Board of Professional Psychology (ABPP), which is the parent board of the American Board of Clinical Neuropsychology (ABCN). Although many practitioners completed appropriate training in clinical neuropsychology before these guidelines were set forth, the Houston Conference model of training is now generally expected for those persons who began their training in 1999 or later.

RELATIONSHIP TO OTHER DISCIPLINES

The unique contributions of neuropsychology can be clarified by examining its relationship to the closely allied disciplines of clinical psychology, behavioral neurology, and neuropsychiatry.

Clinical Psychology

Clinical neuropsychology is primarily differentiated from general clinical psychology by its focus on thorough and extensive evaluation of a broad range of cognitive and emotional factors and their potential relationship to brain damage. Both areas rely on the psychometric foundations of reliability, validity, and normative standards in order to objectively define behavioral symptoms and complaints. Emotional factors are also the province of both. In evaluating cognition, the clinical psychologist tends to focus on issues involving general intellectual, academic, and vocational skills rather than neurological factors. However, clinical psychologists and clinical neuropsychologists are both involved in treatment, which can include psychotherapy as well as cognitive retraining to remediate deficits. Psychoeducation is also provided by both fields in a therapeutic context that discusses the relationship between symptoms and the assessment results and provides recommendations for patients, caregivers, and other health professionals.

Behavioral Neurology and Neuropsychiatry

The medical specialties of behavioral neurology and neuropsychiatry overlap considerably with clinical neuropsychology. However, the neuropsychologist's unique contribution to the assessment of cognitive deficits is based on theoretical background that is shared with cognitive psychology and the neurosciences as well the use of rigorous, psychometrically based methods for developing standardized tests with age- and education-based normative data whenever possible. In contrast, the behavioral neurologist and neuropsychiatrist are more likely to utilize a mental status examination approach for identifying cognitive deficits. These clinicians typically rely on "internal norms," which are based upon the individual clinician's extensive experience with the examination procedures or pathognomonic signs (e.g., dysfluent speech or visual neglect). The behavioral neurologist is trained to diagnose a broad spectrum of neurological disorders, the neuropsychiatrist's focus lies in diagnosing neurological features of psychiatric patients, and the neuropsychologist focuses on the cognitive and behavioral manifestations of a broad range of disorders that can produce cognitive and behavioral impairment in medical, surgical, neurologic, and psychiatric patients. Ideally, all three work together to provide complementary perspectives to a patient's workup.

HISTORICAL INFLUENCES

Clinical neuropsychology has its roots in psychology, neurology, and psychiatry. The groundbreaking work of Broca's and Wernicke's in the 19th century first suggested that complex functions, such as speech and auditory comprehension, could be localized in particular areas in the left hemisphere. In 1909, K. Brodmann published a "cytoarchitectonic" map of the cerebral cortex (Fig. 7.4–1), based on different histological patterns of cells in various parts of the cortex. Over the years, this representational map has proved to have value in identifying functional differences across distinctive cortical regions. It has become a standard reference for identifying cortical areas, though there is now a better appreciation for individual variability. Views of functional localization were repopularized in the 1960s by Norman Geschwind, who emphasized the importance of connections among different parts of the brain in producing complex behavior. Arthur Benton's Iowa School developed a series of carefully constructed psychometric tests, based on concepts that had originally been identified by behavioral neurology, in order to assess specific deficits relative to the normal population. At about the same time, Hans-Lukas Teuber defined the concept of "double dissociation," which is regarded as the strongest evidence for localization of a particular function. This notion is based on observations of mutually exclusive brain–behavior relationships, such that damage affecting region 1 produces performance deficits in test 1 but not in test 2, while damage affecting region 2 produces performance deficits in test 2 but not in test 1. Aleksander Luria, the Russian neuropsychologist, also developed notions of functional localization but utilized a theoretical framework that associated component cognitive processes with complex skills and their neuroanatomical correlates. In the United States, Ralph Reitan applied the psychometric standards of North America to a number of instruments for the express purpose of assessing brain damage, resulting in the *Halstead–Reitan Neuropsychological Test Battery (HRNTB)*. However, the development of increasingly sophisticated imaging techniques in the neurosciences (magnetic resonance imaging [MRI], computed tomography [CT], positron emission tomography [PET], functional MRI with connectivity analyses, and diffusion tensor imaging [DTI]) have brought greatly enhanced capacity to integrate neuroimaging results with neuropsychological methods for measuring cognitive and emotional processing. Current neuropsychological assessment practices increasingly emphasize a more efficient hypothesis testing approach, in which a core screening protocol is supplemented by procedures that assess specific areas of functioning.

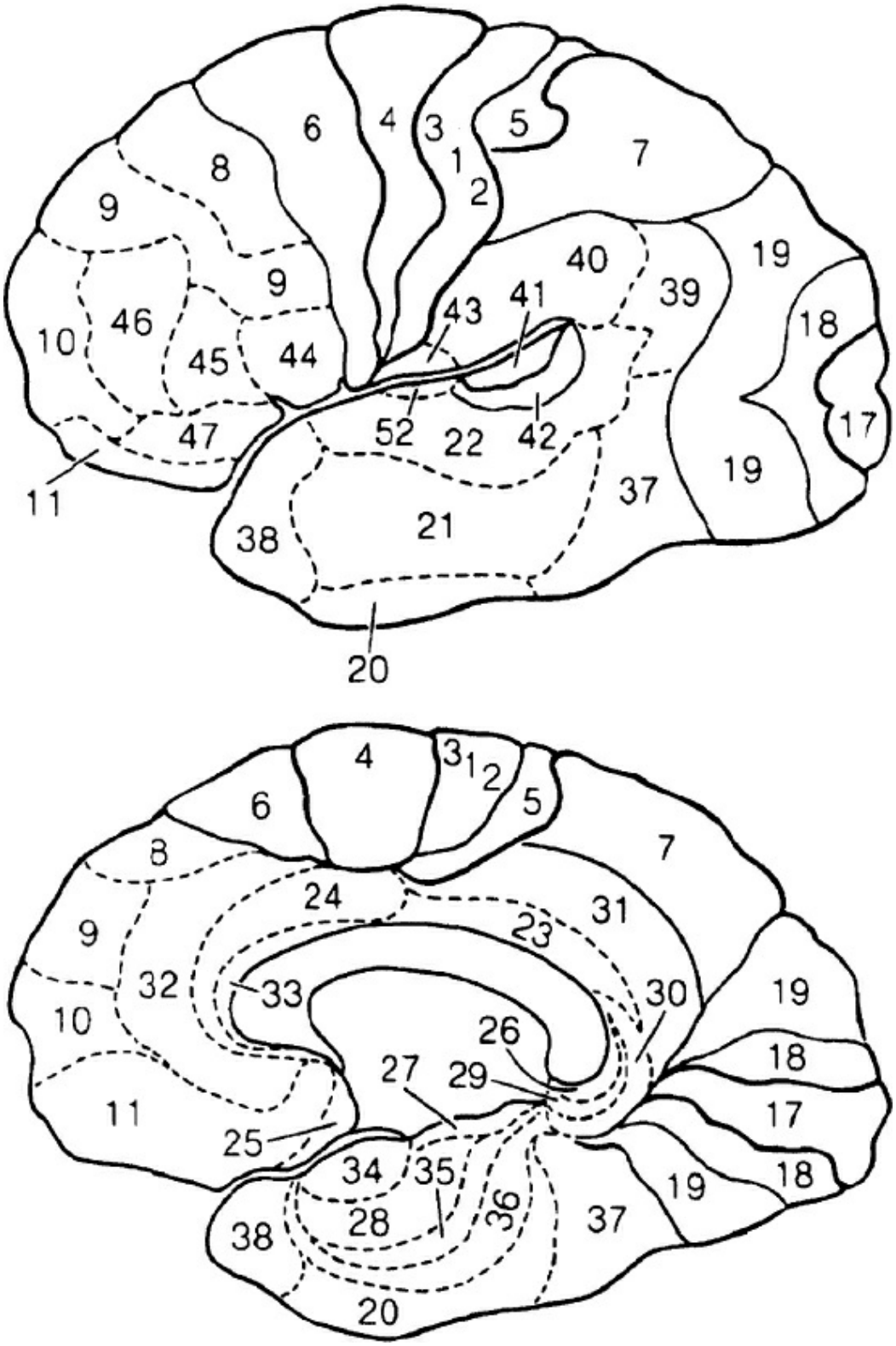

FIGURE 7.4–1. Brodmann areas of the human cortex, showing convex surface (**top**) and medial surface (**bottom**). (From Elliott HC. *Textbook of Neuroanatomy*. 2nd ed. Lippincott Williams & Wilkins; 1969.)

NEUROANATOMICAL CORRELATES

The early history of neuropsychology was driven in large part by the goal of

linking behavioral deficits to specific neuroanatomical areas of dysfunction or damage. Although this early assessment method helped to validate neuropsychological tests that are commonly used today, the localizing function of neuropsychological assessment is now considered less important in light of advances in neuroimaging techniques. Increasing knowledge in the neurosciences has also led to a more sophisticated view of brain–behavior relationships, in which complex cognitive, perceptual, and motor activities are controlled by neural circuits rather than single structures within the brain. An understanding of these brain–behavior relationships is particularly helpful when evaluating patients with focal damage. It is crucial to ensure that the neuropsychological evaluation adequately assesses relevant behavior that is likely to be associated with that area and its interconnecting pathways. The following section will briefly review some of the basic concepts of brain–behavior relationships that are used routinely by the neuropsychologist in interpreting neuropsychological results.

Hemispheric Dominance and Intrahemispheric Localization

Many functions are mediated by both the right and left hemispheres. However, important qualitative differences between the two hemispheres can be demonstrated in the presence of lateralized brain injury. Various cognitive skills that have been linked to the left or right hemisphere in right-handed individuals are listed in Table 7.4–1. Although language is the most obvious function that is largely controlled by the left hemisphere, especially among right-handed individuals, the left hemisphere is also generally considered to be dominant for limb praxis (i.e., performing complex movements, such as brushing teeth, to command, or imitation), and it has been associated with the cluster of deficits identified as Gerstmann syndrome (i.e., finger agnosia, dyscalculia, dysgraphia, and right–left disorientation). In contrast, the right hemisphere is thought to play a more important role in controlling visuospatial abilities and hemispatial attention, which are associated with the clinical presentations of constructional apraxia and neglect, respectively.

Table 7.4–1.
Selected Neuropsychological Deficits Associated with Left or Right Hemisphere Damage

Left Hemisphere	Right Hemisphere
Aphasia	Visuospatial deficits
Right–left disorientation	Impaired visual perception
Finger agnosia	Neglect
Dysgraphia (aphasic)	Dysgraphia (spatial, neglect)
Dyscalculia (number alexia)	Dyscalculia (spatial)

Constructional apraxia (details)	Constructional apraxia (Gestalt)
Limb apraxia	Dressing apraxia
	Anosognosia

Although lateralized deficits such as these are typically characterized in terms of *damage* to the right or left hemisphere, it is important to keep in mind that the patient's performance can also be characterized in terms of *preserved* brain functions. In other words, it is the remaining intact brain tissue that drives many behavioral responses following injury to the brain and not only the absence of critical brain tissue.

Language Disorders. Appreciation for the special role of the "dominant" left hemisphere in controlling language in most right-handed individuals has been validated in many studies. Some of this work is based on the use of sodium amytal testing in epilepsy surgery patients, but the incidence of aphasia following unilateral stroke to the left versus right hemisphere also informs these conclusions. Although it is rare for right-handed individuals to be right hemisphere dominant for language, it does occur in about 1 percent of cases. Hemispheric dominance for language in left-handed individuals is less predictable. About two-thirds of left-handed individuals are actually left hemisphere dominant for language, while about 20 percent each are right hemisphere dominant or bilaterally dominant.

A number of classification systems have been developed over the years for describing various patterns of language breakdown. A common method takes into account the presence or absence of three key features: (1) fluency, (2) comprehension, and (3) repetition (i.e., intact ability to repeat verbally presented words or phrases). All of the following four types of aphasia have damage within some part of the core distribution of the middle cerebral artery on the lateral surface of the left hemisphere.

BROCA'S APHASIA. *Broca's aphasia* (also called *nonfluent* or *expressive aphasia*) has traditionally been characterized by nonfluent speech but intact auditory comprehension and somewhat impaired repetition. It has long been thought to be associated with damage to Broca's area (i.e., left inferior frontal convolution) or Brodmann area 44 (Fig. 7.4–1). However, more recent neuroimaging data in stroke patients have shown that the full syndrome of *Broca's aphasia,* including *agrammatism* (telegraphic speech), is found only in the presence of more extensive damage, which encompasses the suprasylvian area from Broca's area to the posterior extent of the sylvian fissure.

WERNICKE'S APHASIA. *Wernicke's aphasia* (also called *fluent* or *receptive aphasia*) is characterized by fluent speech, impaired auditory comprehension, and somewhat impaired repetition. It has long been thought to be associated with damage to Wernicke's area in the region of

the superior temporal gyrus. However, lesion data in stroke patients has been used more recently to identify a broader range of areas that are important to Wernicke's aphasia, including the superior and middle temporal gyrus as well as the inferior frontal and parietal cortices. The impaired ability to comprehend language directly affects the individual's ability to self-monitor language output and may be related to a breakdown of the syntactic structure of language. Unlike patients with Broca's aphasia who are usually painfully and obviously aware of their communication difficulty, the patient with *Wernicke's aphasia* is typically not aware of their communication problems, because Wernicke's area is critical for comprehending their own speech as well as the speech of others. This lack of insight is similar to the condition of *anosognosia,* in which patients fail to appreciate their own deficits, and presents a particularly frustrating condition for many family members and caregivers. It would not be unusual for an affected patient to fluently produce unintelligible strings of utterances that might potentially be confused with the so-called "jargon aphasia," which is associated with schizophrenia.

CONDUCTION APHASIA. Patients with *conduction aphasia* demonstrate relatively intact auditory comprehension and spontaneous speech, due to the preservation of Wernicke's and Broca's areas. However, the ability to repeat words and phrases is specifically impaired and has traditionally been attributed to damage to the arcuate fasciculus, which interconnects Wernicke's and Broca's areas. This type of aphasia is much more subtle and tends to have less negative impact on daily functioning.

GLOBAL APHASIA. Another common classification, *global aphasia,* is characterized by impairment in all three dimensions of fluency, comprehension, and repetition due to damage to the core language areas on the lateral surface of the left hemisphere. In reality many aphasic patients cannot be neatly classified within a specific system because the pattern of deficits does not exactly fit clear descriptive categories. In fact, detailed language assessment of most aphasic patients typically demonstrates deficits in all three areas, although the degree of deficit among the three areas varies.

Limb Apraxia. Limb apraxia and other cognitive-motor skills deficits are more commonly seen with left than with right hemisphere damage. However, data reviewed by Kathleen Haaland and Deborah Harrington show that the difference in the incidence of limb apraxia after left or right hemisphere damage is not as great as with language, suggesting that left hemisphere dominance for disorders of complex movement is not as strong as that for language. While limb apraxia has not traditionally been considered to be of substantial functional importance, numerous studies demonstrate that deficits in daily functioning are common with limb

apraxia. It has also been suggested that limb apraxia significantly affects rehabilitation outcome. Different types of limb apraxia have been associated with damage to different parts of the brain, especially in the left parietal and frontal regions. For instance, *ideomotor limb apraxia* can be associated with impaired spatiotemporal execution of complex movements, which results in orientation and timing errors, such as carving a turkey by moving the knife in a "jerky" up and down motion rather than smooth back and forth movements. *Ideational apraxia* might result in using the wrong object to perform a movement, such as attempting to use a toothbrush to eat or a razor to brush teeth. Finally, sequencing errors and ideational errors can lead to disrupted activities, such as trying to light a candle before striking the match or beginning to shave before turning on the shaver.

Arithmetic. Arithmetic skills can be impaired after either left or right hemisphere damage. Left hemisphere damage, especially of the parietal lobe, produces difficulty in reading and appreciating the symbolic meaning of numbers (*number dyslexia*). Left hemisphere damage also can be associated with impaired conceptual understanding of the arithmetic problem (*anarithmetria*). In contrast, the deficits in arithmetic computation that can accompany right hemisphere damage are more likely to be observed in written problems. These emerge as problems with the spatial aspects of arithmetic, such as errors resulting from hemispatial visual neglect, poor alignment of columns, or visual misperceptions (e.g., mistaking 6 for 9), and rotations that can result in confusion of signs for addition and multiplication. These problems are obviously extremely significant from the standpoint of independent capacity to manage financial affairs.

Spatial Disorders. Right hemisphere damage in right-handed individuals is frequently associated with deficits in visuospatial skills. Common assessment techniques include drawings and constructional or spatial assembly tasks. Although spatial problems are not recognized as frequently as language disorders, they can be extremely significant from a functional standpoint, depending on their severity. The individual with visuospatial disorder might have difficulty perceiving directions, or get lost even in familiar places, and have difficulty using a map to find their way in unfamiliar places. Persons with hemispatial neglect (most frequently affecting the left hemispace) might have difficulty reading or difficulty reading an analog clock due to neglect of the left hemispace. Reading problems can also be associated with visuospatial problems that result in misperceiving letters (e.g., "b" for "d"). Persons with hemispatial neglect are at high risk for problems with driving, and in extreme cases might even leave food uneaten on the neglected side of a plate or fail to respond to persons on their neglected side. Persons with spatial disorders might also

have difficulty dressing due to problems differentiating the top or bottom of a garment, and in extreme cases, persons with visuospatial problems might have difficulty finding their way through an unfamiliar environment.

VISUOSPATIAL IMPAIRMENT. Distinctive qualitative errors in constructing block designs and in drawing a complex geometric configuration (e.g., Rey–Osterreith Complex Figure) can be seen with either right or left hemisphere damage, though the deficits are typically more extreme after right hemisphere damage. In the presence of lateralized damage to the right hemisphere, impaired performance often reflects the patient's inability to appreciate the "Gestalt" or global features of a design. In the example shown in Figure 7.4–2, this is seen in the patient's failure to maintain the 2 × 2 matrix of blocks with conversion of this matrix into a column of 4 blocks. In contrast, damage to the left hemisphere commonly results in inaccurate reproduction of internal details of the design, including improper orientation of individual blocks, but the 2 × 2 matrix (i.e., the Gestalt) is more likely to be preserved. Similar differences can be seen with drawings, as in the example of the Rey–Osterreith Complex Figure shown in Figure 7.4–3. The patient with right parietal damage draws isolated details of the design, while failing to convey the interrelationship of the details in the overall design configuration. In contrast, the patient with left hemisphere damage tends to maintain the global framework, or Gestalt, of the design but lose the details. Therefore, many neuropsychologists emphasize that a neuropsychological understanding of the impairment depends not just on a set of test scores but also on a qualitative description of the type of error. This often allows the impairment to be linked to a specific neuroanatomical region as well as enabling a better understanding of the mechanisms of the deficit for rehabilitation purposes. This qualitative focus on the type of error is similar to the *pathognomonic* approach that is often used by behavioral neurologists and is important for designing rehabilitation strategies.

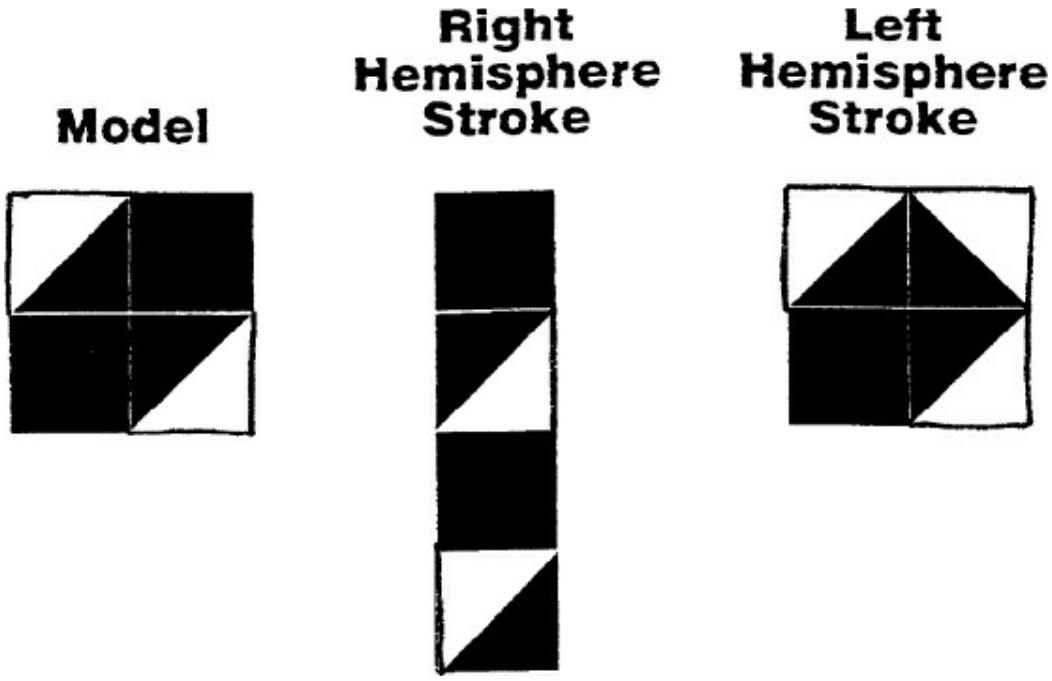

FIGURE 7.4–2. Examples of block design constructions seen in a right hemisphere stroke patient.

In another example, damage to the right hemisphere tends to be associated with decreased appreciation of *global* features of visual stimuli, while left hemisphere damage tends to be associated with decreased analysis of *local* features and detail. This notion is illustrated in Figure 7.4–4, where a left-hemisphere–damaged patient focuses on the larger Gestalt of the triangle or letter M with no regard for the internal characters that actually make up the designs. In contrast the "local" approach of a patient with right hemisphere damage emphasizes the internal details (small rectangles or letter Z) without appreciation for the Gestalt that is formed by the internal details. This example also illustrates the important point that behavioral responses (including errors) are driven as much by preserved regions of intact brain functioning as by the loss of other regions of brain functioning.

NEGLECT. *Neglect syndromes* are characterized by failure to detect visual or tactile stimuli or to move the limb in the contralateral hemispace. They are most commonly associated with right hemisphere damage in the parietal region, but damage to other areas within the cerebral cortex and subcortical areas can also produce this problem. Although neglect syndromes have a similar incidence and may cooccur with visual field cuts or somatosensory deficits, the neglect syndrome is distinct and not explained by any motor or sensory problems that may be present. Visual neglect can be assessed with line cancellation and line bisection tasks, in which the paper is placed at the patient's midline, and the patient is asked to either cross out all of the lines on the page or to bisect the single line

presented. The method of double simultaneous stimulation or visual extinction is another standard procedure for demonstrating the deficit. Neglect syndromes can have devastating functional effects on safety and the ability to live independently and should be taken into account as a standard consideration in the evaluation process. For example, when working with a patient with visual neglect, the examiner should stand in the patient's right hemispace so they won't be ignored. It is also important to position a patient in an inpatient hospital setting so that the patient's "neglected" hemispace is oriented away from the door, thereby facilitating the patient's ability to pay attention to visitors entering the room.

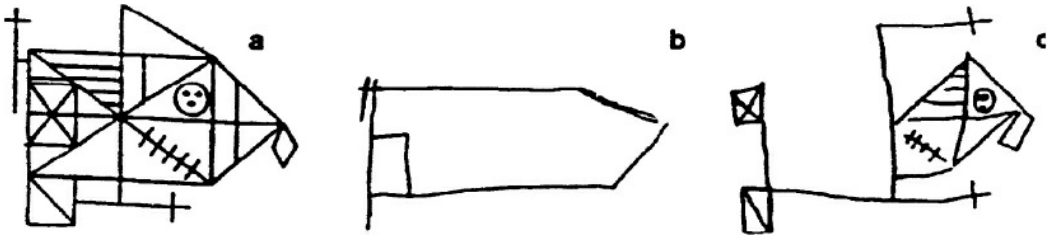

FIGURE 7.4–3. Rey–Osterreith Complex Figure (**A**) model, and drawings from memory in a (**B**) patient with left hemisphere injury, and (**C**) in a patient with right hemisphere injury. (From Robertson LC, Lamb MR. Neuropsychological contributions to theories of part/whole organization. *Cognit Psychol.* 1991;23:305, with permission from Elsevier Science.)

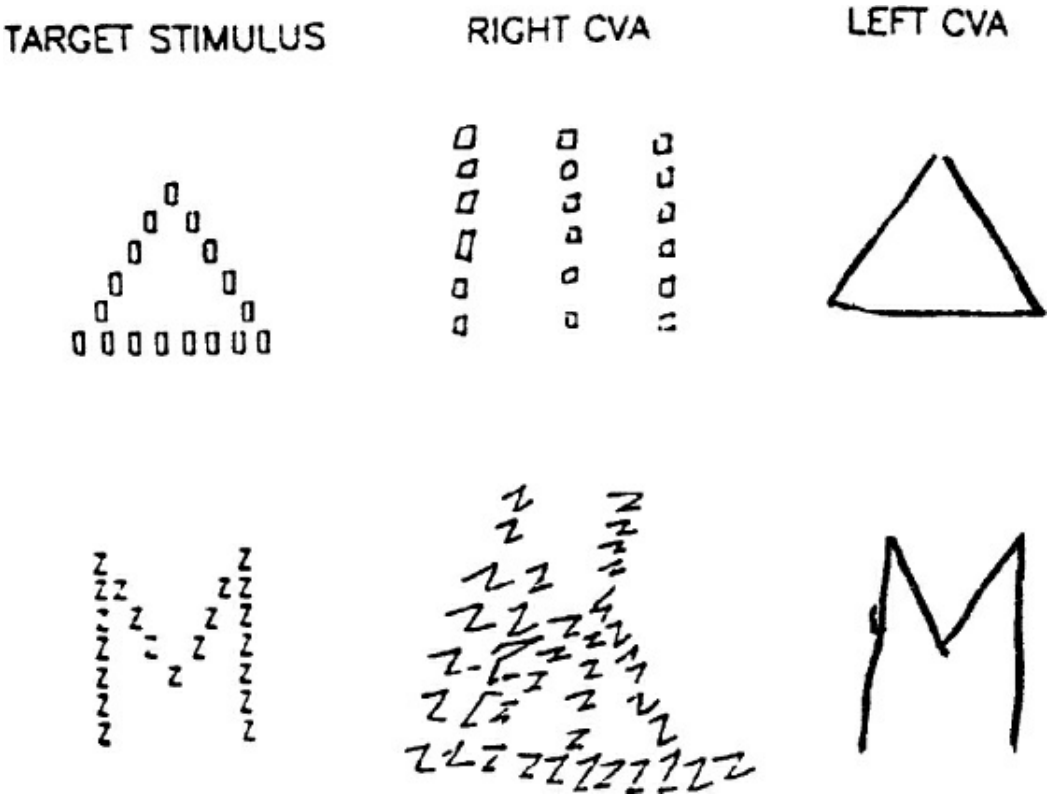

FIGURE 7.4–4. Global local target stimuli with drawings from memory by a patient with right hemisphere cerebrovascular accident (CVA) and by a patient with left hemisphere CVA. (From Robertson LC, Lamb MR. Neuropsychological contributions to theories of part/whole organization. *Cognit Psychol.* 1991;23:305, with permission from Elsevier

Science.)

DRESSING APRAXIA. The syndrome of *dressing apraxia* tends to arise in association with spatial deficits following right hemisphere damage. The resulting difficulty in coordinating the spatial and tactual demands of dressing can be seen in the patient's difficulty in identifying the top or bottom of a garment, as well as right–left confusion in inserting their limbs into the garment. As a result, dressing time can be painfully protracted, and the patient may actually present with a greater level of functional dependence than might otherwise be expected from assessment of simple motor or spatial skills alone. Although dressing apraxia might be noticed as a sudden change in functioning in the presence of an acute focal or lateralized right hemisphere lesion, it might also present as a commonly overlooked complication in the presence of many of the dementia syndromes, especially including Alzheimer disease and vascular dementia.

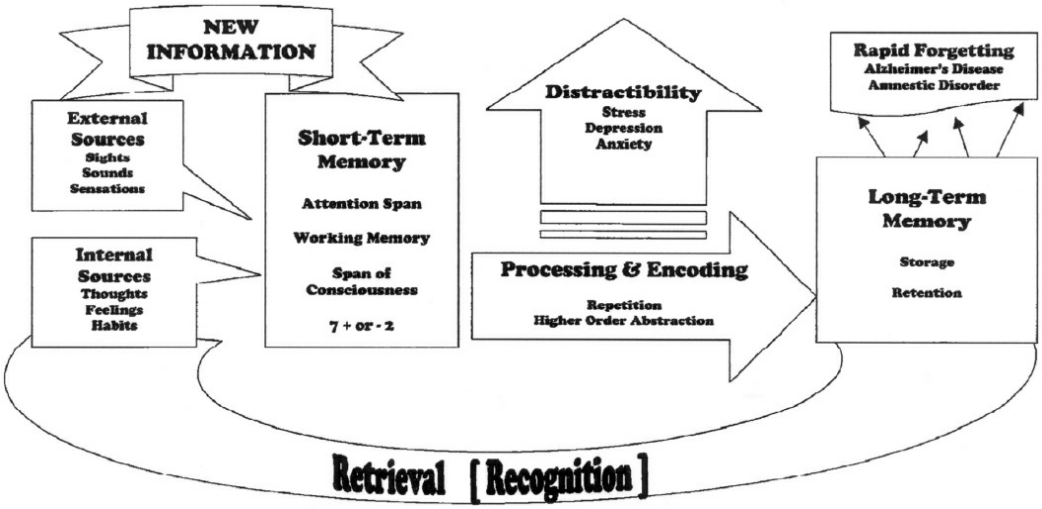

FIGURE 7.4–5. Information processing model of memory.

Memory Disorders. Memory complaints constitute the most common referral to neuropsychology. Thorough neuropsychological examination of memory considers the modality (e.g., verbal vs. spatial) in which the material is presented, as well as presentation formats that systematically assess different aspects of the information processing and storage system that forms the basis for memory. Accumulated research indicates that specialized processing of verbal and spatial memory material tends to be differentially mediated by the left and right hemispheres, respectively. In addition to interhemispheric differences in functional localization, specific memory problems can be associated with breakdown at any stage in the information processing model of memory (Fig. 7.4–5). These stages include (1) registration of the material through *attention,* (2) initial processing and encoding of the material within *short-term memory,* also known as

working memory, (3) consolidation and storage of material in *long-term memory,* and (4) *retrieval* processes, in which material moves from long-term memory storage back into consciousness. A great advantage of neuropsychological assessment is that these various types of memory problems can be specifically identified and described in the course of the examination procedures. Once identified, the specific nature of the deficit can then have important implications for diagnosis, treatment, and prognosis.

ENCODING. The initial encoding of new material can be influenced by a variety of factors, including deficits in attention, language, and spatial processing abilities. It is usually measured by immediate recall of newly learned information (e.g., narrative stories or designs) or by demonstrating the ability to learn new material that has been presented across multiple "learning trials" (e.g., word lists). Problems with the initial processing and encoding of new information can be differentiated from storage and retention problems of memory if the person is able to recall a reasonable percent of the information they initially learned following a delay interval. Attention itself is a relatively fragile cognitive function that can be affected by many factors, including neurologically based disorders (e.g., traumatic brain injury [TBI], delirium) and psychiatric disorders (e.g., depression or anxiety), so it is a crucial aspect of a proper memory assessment.

STORAGE AND RETRIEVAL. Deficits in delayed recall can be associated with, impaired storage and retention of new information, or it can be due to impaired retrieval, in which case the material is still present but not readily accessible. The best way of differentiating these problems is to examine *recognition memory,* in which a patient is typically asked to choose from a set of multiple-choice alternatives or to discriminate target words from *false-positive* foils. If the patient demonstrates accurate recognition but poor recall, then the problem most likely lies in poor retrieval. However, if recognition and recall are both impaired, then the problem is more likely to be related to impaired encoding or storage of new information. This distinction is important because the functions of retrieval and storage are subserved by different neuroanatomical structures. Impaired storage is more often associated with dysfunction of the medial temporal lobe–diencephalic systems whereas impaired retrieval can be associated with a variety of structures, especially the frontal lobes.

Executive Function. The prefrontal lobes and their interconnections to the rest of the brain are known to play an important role in *executive functions,* which are essential to the planning and organizing aspect of complex problem-solving responses, as well as self-monitoring, inhibiting inappropriate responses, and flexibly switching from one response strategy to another. Damage to the frontal lobes has also been associated with

significant personality changes. This was historically exemplified by the famous 19th century case of Phineas Gage, who became irresponsible, socially inappropriate, and unable to carry out plans after a tamping iron was blown through his frontal lobes. As conceptualized by Lezak, the executive functions include volition (i.e., formulation of a goal, motivation to achieve the goal, and awareness of one's own ability to achieve the goal), planning, purposive action (response selection and initiation, maintenance, switching, and stopping), and execution, which involves self-monitoring and self-correction as well as control of the spatiotemporal aspects of the response. Although hemispheric differences in the control of executive functions by the frontal lobes have not been as well documented as in the parietal and temporal lobes, the Montreal Neuropsychology Group led by Brenda Milner has identified asymmetry in the control of verbal and design fluency.

Jeffrey Cummings has proposed a model in which the prefrontal lobes and their interconnections are separated into the dorsolateral, orbitofrontal, and anterior cingulate divisions. In this model, various neurobehavioral functions are thought to be linked to corresponding neuroanatomical relationships between cortical and subcortical areas, especially the basal ganglia and thalamus. The *dorsolateral* prefrontal lobe circuit is associated with higher cognitive functions, such as working memory, hypothesis generation, retrieval and initiation of unique words and designs, organization of information (e.g., clustering word lists on the basis of semantic category), and development of alternating and sequential motor programs. The *orbitofrontal* syndrome is characterized by marked personality changes, such as instability, lability, and impulsivity. The *anterior cingulate* circuit is associated with impaired response initiation (i.e., akinetic mutism in the most extreme case) and response inhibition. These patients have difficulty inhibiting prepotent responses on "go-no go" tests and may be perceived by their families to be depressed, even though the patient might be content doing nothing.

Motor Skills. The neuropsychological evaluation commonly includes formal tests of motor skills, such as measures of finger tapping speed, grip strength, and fine motor dexterity. These tests, which have demonstrated validity and reliability, are useful for assessing lateralized motor impairment and have implications for functioning in daily life as well as vocational planning.

PSYCHOMETRIC ISSUES

One of the strengths of neuropsychological assessment lies in the detailed quantitative record that it provides for different areas of cognitive performance, as well as reliance on normative standards on which interpretations are based. The normative database for any given test is

particularly important to consider in order to ensure that deficits seen on individual test performances cannot be explained by demographic characteristics of the patient, such as age, education, or cultural considerations. Of course, this approach is also dependent on the care that has been expended in the development and validation of specific test techniques. The American Psychological Association has established ethical guidelines for the practice of psychological assessment that require tests to have acceptable levels of validity and reliability for the particular clinical application in which they are used. These guidelines serve as the basis for valid judgments to be made about change over time, based on a given individual's performance over successive examinations. This approach to standardization also ensures that test scores can be regarded as comparable even when administered by different examiners and in different settings.

Validity

Validity refers to a test's ability to measure what it purports to measure. Validity can be ascertained by (1) correlating scores from a particular test with other, well-established measures of the same skill, (2) relating test scores to diagnostic classifications or to measures of brain structure and function, or (3) relating test outcomes to aspects of everyday functions that are thought to be dependent on that skill (e.g., comparing pencil-and-paper tests of visuospatial skill with ability to find one's way around a shopping center). Historically, neuropsychological tests have been validated by examining their ability to differentiate brain-damaged patients from non–brain-damaged patients. More recently, they have been validated against neuroradiological or neurosurgical data and diagnostic categories (e.g., patients with multiple sclerosis vs. medical controls). Although less attention has been paid to relationships between neuropsychological test outcomes and everyday function, modest correlations have been documented between test outcomes, self-care activities, and selected complex skills, and there is a growing movement toward assessing the ecological validity of neuropsychological tests and procedures.

Reliability

Reliability refers to the consistency of test scores. Tests are judged by (1) their internal consistency, or the extent to which all items appear to be measuring the same thing, (2) their ability to yield similar scores across multiple test occasions in normal individuals, and (3) the consistency in scores that is obtained when different people administer and score a test. Individual cognitive measures often have lower reliabilities than composite scores that are derived from several measures, and even within a particular ability domain, some measures exhibit higher reliabilities than do others. For example, within the Wechsler Memory Scale-IV, some individual tests

of immediate and delayed memory have fairly low test–retest reliability (e.g., r = 0.67 and 0.64, for immediate and delayed visual reproduction, respectively), but the composite index scores for Immediate Memory Index and Delayed Memory Index have acceptably higher test–retest stability (r = 0.81 and 0.82, respectively). This underscores the importance of using more than one test to measure each type of ability in a thorough neuropsychological evaluation. Therefore, it is particularly important to take reliability into consideration if only one test measure is used.

Predictive Utility

Predictive utility takes into consideration the sensitivity, specificity, and positive and negative predictive values of a test. A test that yields positive findings for most persons with disease and negative outcomes for most persons without disease has both high sensitivity and specificity. Often, published values for sensitivity and specificity are based on ideal comparisons (e.g., clear cases of Alzheimer disease vs. carefully screened normal controls) and do not permit a clinician to estimate how useful a test will be in a general clinical population. It is important to estimate in advance the predictive utility of particular tests for particular diagnoses because the predictive value of neuropsychological tests varies considerably for different diagnostic questions. For example, as discussed above in the section on *Neuroanatomical Correlates*, some highly specific neuropsychological deficits can be linked to injury in very specific areas of the brain. However, many other neurological diagnoses involve multiple brain systems that result in generalized impairment across a battery of tests. In testing for dementia, for example, neuropsychological test results are typically more sensitive than they are specific, and so the interpretation must consider other medical and historical data in order to arrive at a more specific diagnosis. This underscores the importance of interpreting neuropsychological data in the context of the patient's history.

Base Rates

Computation of positive and negative predictive values addresses very practical questions: Given a positive test outcome, what is the probability of having disease? And, given a negative test outcome, what is the probability of not having disease? Positive and negative predictive values are computed as the ratio of accurate positive findings to total positive outcomes, and the ratio of accurate negative findings to total negative outcomes, respectively. Predictive values are judged relative to estimated base rates of illness within the referral setting. For example, if the base rate of dementia among older adults referred for psychiatric assessment is 50 percent, then any test that purports to be sensitive to dementia must have greater than 50 percent predictive value to warrant its application.

Another aspect of base rate is related to the frequency with which abnormal neuropsychological findings are obtained in normal individuals. This is a very important consideration in neuropsychology, in which large batteries of multiple tests are usually administered. It is actually more typical than not for most "normal" persons to perform in the impaired range on some measures in a large battery of tests. For example, the "Heaton" manual of normative standards points out just how common it is for non–brain-damaged individuals without psychiatric problems to perform in the "impaired" range. In fact, only 13.2 percent of healthy individuals completed a battery of 25 neuropsychological tests without any scores in the impaired range. In contrast, the same healthy group had a median value of three scores in the abnormal range, and more than 25 percent of the "normal" controls performed in the impaired range on six or more scores in the 25-score battery. Given these data, it would be very naive to conclude that an individual is impaired by simply counting the number of neuropsychological test scores in the "impaired range." Instead, these data underscore how important it is to base the neuropsychological conclusions on a *pattern of performance* that is interpreted in the *context of the patient's history*. The strongest evidence of brain injury should typically be based on a pattern of consistent impairment that is observed across several measures within the same cognitive domain, as opposed to impairment on isolated tasks that are scattered across different cognitive domains.

Selecting Appropriate Normative Data

Considerable research in recent years has been directed toward expanding the normative database for clinical neuropsychological tests with respect to important demographic factors. Normative data are now available for many cognitive measures across a broad *age* range. On the average, persons in their early 70s score about a standard deviation lower than those in their 20s on tasks that combine cognitive processing demands with requirements for speed of performance, so it is important to compare a given individual's performance on these tasks with that of age peers. *Education* is also known to have a highly significant effect on many aspects of cognitive performance. *Gender* differences are relatively small on most clinical neuropsychological tests, but there are a few types of tasks, such as measures of motor speed and strength, where gender differences can be found.

Cultural factors and language background pose special challenges in neuropsychological assessment. These factors are frequently addressed through clinical judgment and consultation with cultural informants, but current best practices would encourage a neuropsychologist to seek the best possible cultural and language match between examiner and examinee. Although normative standards based on ethnicity have been

developed for some groups, it is unlikely that appropriate normative standards will ever be developed for every specific ethnic and demographic group. In addition, reliance on norms alone outside of a cultural context would run the risk of misinterpreting important aspects of the clinical presentation. In most cases, a culturally sensitive clinician must first consider whether more appropriate clinical providers with shared language and culture are available. Any decision to proceed in assessing an individual who differs from the clinician in language and/or culture should only be made by adopting a cautious approach that clearly acknowledges the cultural limitations, especially if the examinee represents a cultural minority for whom standardized normative data might be lacking. In some cases, the context of the referral question might offer guidance about whether and how to proceed. For instance, a referral question that involves assessment of acute changes in cognitive status following TBI could conceivably be addressed by comparing a series of intra-individual test–retest measures. Such an assessment might not indicate the individual's "best" level of functioning, but could reasonably provide baseline measures for assessing change in cognitive status over the course of recovery.

A 47-year-old, right-handed woman with life-long, severe hearing impairment was referred for neuropsychological evaluation as an inpatient in a rehabilitation unit about 2 weeks after she had suffered a moderate TBI and orthopedic injuries in a motor vehicle accident. The neuropsychologist initially considered declining the case and trying to locate a neuropsychologist with more experience in assessing hearing impaired individuals. However, no other neuropsychologist could be located in the area. The neuropsychologist did consult with a local pediatric psychologist who had experience with children and adolescents with hearing impairment, and she helped him determine that he would have to be very cautious in drawing any conclusions about the woman's best level of intellectual and cognitive functioning. She also helped him identify some cultural themes related to psychosocial functioning that are common in the deaf community. He decided to take on the case after clarifying that the current referral issue involved establishing a baseline of the woman's functioning during the acute phase of her hospitalization so that the rehabilitation team could follow her cognitive recovery. Before beginning the assessment, he identified an interpreter from the local signing community who was certified as a medical and legal interpreter, and he selected a combination of measures that emphasized visual presentations of stimuli. A search of journal articles helped him identify a few tasks for which some limited norms from a hearing-impaired community had been developed, and consultation with colleagues on a professional list-serve helped him to further refine his approach to the case. He reviewed ethical guidelines regarding cultural considerations, competence and scope of practice in assessment, and he was reminded that it would be important to involve the woman's family in the process as much as possible.

Estimates of Premorbid Functioning

The degree to which an individual's intellectual and cognitive performance is thought to be impaired is always based on the degree to which the performance differs from that individual's baseline or longstanding level of functioning, usually referred to as *premorbid functioning*. Because prior

standardized baseline testing of cognitive abilities is not routinely available for most individuals, *premorbid functioning* is usually estimated, based on educational and occupational history, as well as performance on tests that are thought to be minimally affected by brain damage, such as tests of *crystallized intelligence,* including fund of vocabulary, general information, or reading recognition. Impairment can then be defined as any statistically significant deviation from that individual's estimated level of premorbid ability. As illustrated in the normal *bell curve* that is shown in Figure 7.4–6, statistically significant results on any given set of tests would fall at least one standard deviation above or below the *mean* (or average) performance that is established for that individual, either in comparison to established normative standards or in relation to his or her performance across the other tests. For example, if a 32-year-old, Anglo man from the Midwest has only 4 years of education and has worked as a laborer, his premorbid level of functioning might be estimated to lie in the borderline to low average range. This represents the population that performs at least one standard deviation below average, and therefore most test scores for that individual would be expected to hover around that level. In the absence of pathognomonic signs of impairment, such as unilateral neglect or aphasia, it is typically more challenging to identify cognitive signs of cerebral dysfunction in such an individual unless a pattern of test scores falling at least two standard deviations below average emerges that is consistent with the medical history. This example provides another example of the importance of considering the broad context of demographic and cultural factors in the interpretation of neuropsychological evaluation.

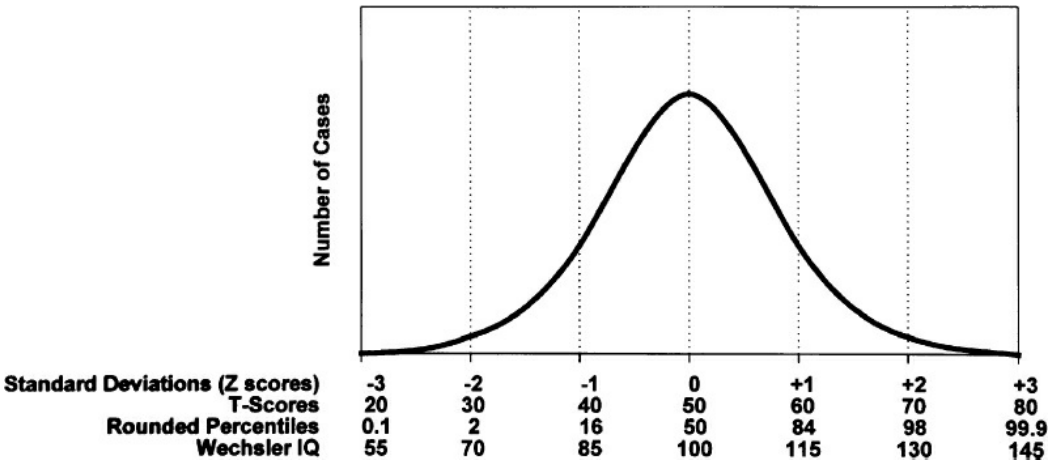

FIGURE 7.4–6. Normative (bell) curve which compares z-scores, T-scores, percentiles, and Wechsler IQ scores.

Moderating Factors

Performance on neuropsychological tests can be influenced by transient or situational factors. *Alcohol and drug use* can impair attention, memory, and speed of cognitive processing, as can many *prescription medications,*

especially in the elderly and brain injured. *Chronic pain* and the medications used to treat pain are also important to take into consideration. *Psychiatric disorders* such as mania, severe depression, or psychosis can undermine cognitive performance, rendering formal testing invalid in severe conditions or reducing scores below expected levels in milder cases. In short, the impact of drug use and psychiatric state, combined with clinical observation and input from the patient and collaterals, must be considered in interpreting the impact of these moderating variables on the neuropsychological test results. An experienced neuropsychologist should also be expected to explicitly identify relevant factors that could be expected to undermine or contribute to lower than expected performance, such as impaired vision or hearing loss, acute illness, or recent significant stressors.

GENERAL REFERRAL ISSUES

Referrants turn to neuropsychology for many reasons that include differential diagnosis, baseline measurement, and treatment planning, as well as opinions regarding causality and decisional capacity. Because many referrants have limited experience and knowledge of the scope of neuropsychology, it is both reasonable and important for the neuropsychologist to take an active role in refining the specific questions that are asked and providing realistic information about the limitations of the consultation. An early conversation with a potential referrant can go a long way toward increasing the overall level of satisfaction and improving the referrant's expectations for a satisfactory consultation. In general, it is not considered appropriate for a referrant to request specific tests or procedures, because the specific methods of neuropsychological assessment will necessarily reflect the patient's unique presentation of symptoms and complaints, history and development, the perspective of the neuropsychologist, and the referral question.

Level of Functioning

A common referral issue involves documentation of level of functioning for a variety of purposes, including assessment of change or capacity to make decisions, especially in the presence of diagnoses such as dementia, stroke, and head injury.

Differential Diagnosis

Like any other diagnostic procedure, the results of a neuropsychological examination must be interpreted in light of all available information, including the history and any associated medical factors that are documented or reported for the individual. Many neurological and psychiatric disorders have similar clusters of symptoms in common, with

complaints of concentration or memory problems being among the most frequently reported problems. For example, impaired attention and concentration are commonly found among patients who present with a history of mild closed head injury, psychiatric issues (e.g., depression, posttraumatic stress disorder [PTSD], or anxiety), or any of several types of neurological disorders that are associated with subcortical involvement (e.g., Parkinson disease, multiple sclerosis, or human immunodeficiency virus [HIV]-related cognitive decline). Therefore, a pattern of cognitive impairment across several critical areas of functioning does not necessarily provide a sufficient basis for a specific diagnostic conclusion but must be integrated with the medical and relevant psychosocial factors in order to arrive at a specific differential diagnosis. The following discussion reflects some common clinical syndromes in which psychiatrists and neuropsychologists share an interest and may work together in assessing and treating patients.

Age- or Stress-Related Cognitive Change. Many middle-aged and older adults have concerns about everyday concentration and memory failures, and with heightened public awareness about conditions such as Alzheimer disease, an increasing number of these individuals seek evaluations for these concerns. Neuropsychological testing provides a detailed, objective picture of different aspects of memory and attention, which can be helpful in reassuring healthy persons about their abilities. It also provides an opportunity for assessing undetected mood or anxiety disorders that may be reflected in cognitive concerns and for offering suggestions about mnemonic strategies that can sharpen everyday function.

A 77-year-old, left-handed man with a high school education was referred for neuropsychological assessment by his primary physician after the patient mentioned a recent episode of getting turned around while driving. He reported a 25-year work history with the postal service and current part-time work as a cashier in a community recreation facility. He also reported an active social life. Medical history was stable despite past treatment history for cataracts and prostate cancer, and MRI scan of the head was unremarkable. In interview, the patient and his wife complained that the referring physician had "made a mountain out of a molehill." They both agreed that he had always had a poor sense of direction and gave several humorous examples of "getting turned around" when driving over the years.

Results of neuropsychological assessment indicated variable performance on tests of attention and concentration. His performance was excellent on tests of memory, language, and executive problem solving abilities, but visual–spatial and constructional abilities were moderately impaired. Mild levels of grief-related depression were indicated on self-report. Follow-up discussion of the results provided an opportunity to discuss the significant visuospatial problems that were identified. Neither the patient nor his wife was surprised to hear this, and they good-naturedly pointed out that they had tried to tell the referring physician that he had always had such problems. The evaluation concluded that the findings were most consistent with a probable developmental history of visuospatial learning disability. However, the couple was advised to seek follow-up if significant changes become apparent in the future in order to compare to the objective baseline established by this examination.

Mild Traumatic Brain Injury. TBI can be classified as mild, moderate, or severe. These severity ratings are most commonly characterized by the duration of loss of consciousness or posttraumatic amnesia, as shown in the DSM-5 severity ratings in Table 7.4–2. Most neuropsychological studies of the cognitive consequences associated with TBI have consistently found a strong relationship between duration of loss of consciousness and cognitive recovery following TBI.

The vast majority of TBI cases referred for neuropsychological consultation involve mild TBI from which an excellent cognitive recovery can be expected within days to weeks after injury. However, a significant proportion of persons who have suffered a mild TBI do have subjectively experienced complaints that persist for many months after the injury, such as poor concentration and mood changes, fatigue, photosensitivity, balance problems, tinnitus, and headache or other forms of pain. This set of subjective complaints is sometimes diagnosed as post-concussion syndrome in the 10th edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10). A number of methodologically sound studies indicate that premorbid histories of psychological symptoms, including trauma and substance abuse, may be contributing factors to consider. As noted in DSM-5, since cooccurring symptoms, such as depression, irritability, fatigue, headache, photosensitivity, sleep disturbance, tend to resolve in the weeks following mild TBI, substantial subsequent deterioration in these areas should trigger consideration of additional diagnoses.

Table 7.4–2.
Severity Ratings for Traumatic Brain Injury

Injury Characteristic	Mild TBI	Moderate TBI	Severe TBI
Loss of consciousness	<30 minutes	30 minutes–24 hours	>24 hours
Posttraumatic amnesia	<24 hours	24 hours–7 days	>7 days
Glasgow Coma Scale Score	13–15 (not below 13 at 30 minutes)	9–12	3–8

In many of these cases, there is no clear evidence of brain injury on MRI or other neuroradiological tests. Therefore, neuropsychological testing often plays a crucial role in determining the extent of objective cognitive deficit and examining the possible role of psychological factors in perpetuating cognitive problems.

Understanding of the long-term consequences of TBI has improved significantly over the past 20 years. The work of Surreya Dikmen and her colleagues has been particularly significant in demonstrating a *dose–response* relationship between the severity of initial injury and the long-term cognitive consequences. For example, when loss of consciousness is

behaviorally defined as “time to follow commands” in the Glasgow Coma Scale, those persons with relatively brief loss of consciousness are indistinguishable from healthy control subjects on tests of cognitive functioning after about a year. However, contributions of premorbid or associated psychological factors have been shown to be highly significant among persons who complain of enduring subjective symptoms, such as worsening cognitive problems, emotional problems (depression, irritability, or anxiety), and somatic complaints (headache, dizziness, or chronic pain). The preponderance of evidence at this point indicates that persons who meet criteria for mild TBI can be expected to fully recover in terms of cognitive functioning. However, when people with a history of mild TBI complain of enduring symptoms that exceed expectation, they should be carefully evaluated in order to identify any contributing psychological factors that could offer potential treatment strategies, such as symptoms of depression, anxiety, PTSD, substance abuse, or other psychiatric disorder.

A 32-year-old Desert Storm veteran sought neuropsychological consultation with complaints of memory problems that were characterized by “forgetting of intentions,” such as walking into a room and forgetting why he had gone there.” He stated that he had been in the vicinity of a roadside blast 4 years before that had injured occupants of another truck about 100 ft ahead of him. He feared he was suffering from a TBI that had also caused bouts of dizziness, tinnitus, and persistent headaches. He stated that he did not suffer any actual loss of consciousness, but thought he was on “auto-pilot” when he jumped out of his vehicle to help save his comrades. The clinical interview revealed significant symptoms of anxiety, social isolation, intrusive memories, and poor sleep with nightmares. He was also drinking large amounts of alcohol and missing work. Cognitive testing showed that he was generally performing in the mildly to moderately impaired range on most measures of attention, concentration, and information processing. His memory performance showed poor initial encoding but his delayed recall showed that he had retained 90 to 100 percent of the material he had initially learned. He gave up on several measures of motor-perceptual skill, complaining that his “TBI” had ruined his life. At the conclusion of the examination, the neuropsychologist thought that, even though the test results showed impaired performances on most cognitive measures, the circumstances of the injury alone would not be expected to lead to this level of impairment. Since he had not experienced significant loss of consciousness of at least 30 minutes or more, the TBI would be classified as “mild TBI” and the cognitive symptoms would be expected to be undetectable on testing 6 months to a year following the event. Feedback to the veteran focused on reassuring the veteran that the pattern of performance did not indicate the feared “permanent brain damage” but showed that he was overwhelmed with mental distractibility due to the combination of PTSD, chronic pain, and substance use. Follow-up treatment for symptoms of PTSD was initiated, beginning with psychoeducation about the overlapping symptoms between PTSD and mild TBI.

The neuropsychologist should bear in mind that many patients with mild TBI are involved in litigation, which can complicate the neuropsychologist’s ability to identify the causes for impairment. Although outright malingering is thought to be relatively infrequent and difficult to verify, the possibility of contributing psychological factors due to somatic

symptom disorder, depression, PTSD, and substance use should be a prominent consideration when potential legal settlements or disability benefits are in question. Several psychological models, such as "symptom misattribution" or "diagnosis threat," might explain some tendency to exaggerate symptoms based on the individual's expectations about how an individual with TBI might be expected to perform. The possibility of contributing psychological factors is particularly important to consider in the case of mild head injury, when subjective complaints may be disproportionate to the objectively reported circumstances of the injury.

Poststroke Syndromes. After the acute phase of recovery from stroke, patients may be left with residual deficits, which can affect memory, language, sensory/motor skills, reasoning, or mood. Neuropsychological testing can help to identify areas of strength, which can be used in planning additional rehabilitation and can provide feedback on the functional implications of residual deficits for work or complex activities of daily living. Assessment of functional skills can also be helpful to a psychiatrist who is managing mood and behavioral symptoms or dealing with family caregivers.

Detecting Early Dementia. Conditions that particularly warrant neuropsychological assessment for early detection and potential treatment include HIV-related cognitive deficits and normal pressure hydrocephalus. When concerns about a person's memory functioning are expressed by relatives instead of the patient, there is a higher probability of a neurological basis for the functional problems, since loss of insight is a key feature of most neurodegenerative syndromes. Neuropsychological testing, combined with a good clinical history and other medical screening tests, can be highly effective in distinguishing early dementia from the mild changes in memory and executive functioning that can be seen with normal aging. Neuropsychological evaluation is particularly helpful in documenting cognitive deterioration and differentiating among different forms of dementia. An additional incentive for early diagnosis of dementia now lies in the fact that a portion of patients with early dementia may be candidates for memory-enhancing therapies (e.g., acetylcholinesterase inhibitors), and testing can provide an objective means of monitoring treatment efficacy.

Neuropsychological profiles have been used to identify patients who do not meet the full criteria for Alzheimer disease but who demonstrate memory deficits with normal performance in other cognitive domains and with no or minimal functional complaints. This pattern, termed amnestic *mild cognitive impairment* (MCI), may represent prodromal Alzheimer disease or a transitional state, in that such patients show a higher incidence of conversion to Alzheimer disease than individuals with no such cognitive

impairment. MCI conversion rates to Alzheimer disease have been reported as 15 percent annually relative to conversion rates in those without cognitive deficits of about 2 percent annually. However, these conversion rates vary from 10 to 70 percent depending on the initial MCI diagnostic criteria that are used, which emphasizes the difficulties of operationally defining cognitive and functional decline. It is also important to note that the definition of MCI has been revised since its introduction to include nonmemory deficits, especially executive dysfunction, with the introduction of nonamnestic MCI and the potential of conversion to other types of dementia. However, the clinical implications of conversion to dementia are complex. For example, the diagnosis of Alzheimer disease should only be cautiously addressed due to the possible negative emotional repercussions of such information, the fact that not all MCI patients convert to Alzheimer disease, and the fact that current medications produce only mild improvement even in patients diagnosed with Alzheimer disease. Because research data suggest that the greatest benefit of many of these medications (in terms of slowed progression) is found among patients with mild Alzheimer disease, there is hope that the same medications may provide optimal effects in amnestic MCI patients. In reality, many clinicians currently routinely prescribe acetylcholinesterase inhibitors to patients diagnosed with MCI even though there is no convincing research evidence that these medications slow progression to dementia. In contrast, the MCI diagnosis should stimulate some consideration for providing affected patients and their families with information for future planning, such as financial management oversight. There is also some evidence that cognitive retraining may improve cognitive performance as well as subjective cognitive and emotional complaints. In the future, the combination of genetic information, neuroimaging, and the behavioral evidence of neuropsychological assessment are likely to improve the ability to diagnose dementia at an earlier stage, which will be important as new treatments are developed.

Distinguishing Dementia and Depression. A substantial minority of patients with severe depression exhibit serious generalized impairment of cognitive functioning. In addition to problems with attention and slowing of thought and action, there may be significant forgetfulness and problems with reasoning. By examining the pattern of cognitive impairment, neuropsychological testing can help to identify a dementia syndrome that is associated with depression, historically known as *pseudodementia*. Mixed presentations are also common, in which symptoms of depression coexist with various forms of cognitive decline and exacerbate the effects of cognitive dysfunction beyond what would be expected from the neurological impairment alone. Neuropsychological testing in this case can be very helpful by providing a baseline for measuring the effect of

antidepressant or other therapy in alleviating cognitive and mood symptoms.

A 75-year-old gentleman with a Ph.D. in the social sciences sought neuropsychological reexamination for ongoing memory complaints, stating that, "several of my friends have Alzheimer's." In an initial examination 1 year before, he had performed in the expected range (above average) for most procedures, despite variable performance on measures of attention and concentration. His medical history was unremarkable, with no significant vascular risk factors and no known family history of dementia, but he continued to be treated for chronic depression. Results of the follow-up examination again clustered in the expected above average range with variable performance on measures of attention. On list-learning tests of memory, his initial learning of a word list was lower than expected, but delayed retention of the material was above average, with excellent discrimination of target items on a recognition subtest. He also endorsed a large number of symptoms of depression on a self-report inventory.

In a follow-up discussion of these results, when asked to evaluate his own performance, the gentleman was convinced that the results would verify evidence of dementia. Rather than directly refute the gentleman's belief, the neuropsychologist adopted a psychoeducational therapeutic approach, using an Information Processing Model of Memory to explain how memory complaints can result from several common causes (Fig. 7.4–5). He was shown how the pattern of his performance was most similar to attention-based memory problems that interfere with initial encoding of new information. He was also shown how his intact retention of material over time contrasted with the expected pattern of "rapid forgetting" in Alzheimer disease. The examination results were used to validate this gentleman's experience of "memory" problems and reinforce appropriate treatment for depression, while moving him away from an inappropriate interpretation of his problems and unrealistic expectations for his own performance.

Change in Functioning Over Time

Because many neurological diagnoses carry clear expectations regarding normal rates of recovery and decline over time, it is frequently important to reexamine a given patient with follow-up neuropsychological assessment after 6 months to a year. For example, it might be important to monitor declines in independent functioning that could be associated with a progressive dementia or to identify improvement following a stroke or tumor resection. Follow-up examinations also provide an opportunity to objectively examine complaints of longstanding or worsening cognitive sequelae following mild head trauma, even though the current literature indicates that the greatest proportion of recovery of function is likely to occur over the initial 6 months to a year post-injury. Although continuing subtle signs of recovery can continue after that period, failure to improve following the injury—or worsening of complaints—would suggest the possibility of contributing psychological factors or the existence of a preexisting or coexisting condition, such as substance abuse, dementia, or outright malingering.

Assessment of Decision-Making Capacity

Neuropsychologists are often asked to assist in determining an individual's

capacity to make decisions or to manage personal affairs. Neuropsychological testing can be useful in these cases by documenting areas of significant impairment and by identifying areas of strength and well-preserved skills. Opinions about decision-making capacity are seldom based on test findings alone and usually rely heavily on information gleaned from clinical interview, collateral interviews with family or caregivers, and direct observations (e.g., in-home assessment) of everyday function. In fact, appraisal of an individual's level of insight and capacity to appreciate his or her own limitations is typically the single most important aspect of the assessment. Standards for decision-making capacity are generally defined by state statutes, and, of course, the ultimate determination of competence rests in the authority of the presiding judge. However, the neuropsychologist or other health care professional can play a significant role in shaping the judge's ruling by providing a professional opinion that is supported by compelling behavioral data that has strong face validity.

As a general rule of thumb, consideration of decision-making capacity is best approached in the narrowest possible sense so as to infringe as little as possible on the individual's freedom to represent their own interests. The most common issues related to decisional capacity involve the areas of (1) financial and legal matters, (2) health care and medical treatment, and (3) ability to live independently. Some capacity issues involve higher standards, such as ability to drive, ability to work, or practice in a given profession (e.g., air traffic controller, surgeon, or financial advisor). In such cases, it is particularly important for the neuropsychologist to rely on appropriate normative expectations that are appropriate for the type of activity, as well as the patient's demographics. An excellent resource, entitled, "Assessment of Older Adults With Diminished Capacity," was jointly published by the American Bar Association and the American Psychological Association, with three versions available online that were written for use by attorneys, judges, and health care professionals.

A 73-year-old attorney was referred by his neurologist for neuropsychological assessment following a "slight stroke" or transient ischemic attack (TIA) that resulted in transient symptoms of left-sided upper extremity weakness. He told the neuropsychologist that he was trying to decide whether he could return to his longstanding work as an attorney. MRI examination showed extensive periventricular white matter changes as well as lacunar infarcts. On cognitive testing, the attorney performed in the above average range on standard measures of vocabulary and general information. He performed in the average range on measures of simple attention, such as digit span, but he only performed in the low average range on measures of working memory and attention that required him to mentally manipulate information, such as repeating digits in reverse order or saying months of the year in reverse. On a measure of abstract problem solving, he failed to benefit from standard cues that were available. However, the 18 "perseverative" responses that he accumulated were exactly "average" when his score was compared to age- and education-corrected norms for a 70-year-old with 18 years of education.

At this point, the neuropsychologist reasoned that the level of cognitive functioning

expected for an attorney to focus attention and engage in high-level problem-solving should be the same for any practicing attorney, regardless of age. In fact, she found that the attorney's score of 18 perseverative responses would fall nearly 1.5 standard deviations below average, into the mildly impaired range, for a 49 year old with 18 years of education. The neuropsychologist concluded that, even though the attorney's performance could technically be considered "normal" for his age, his overall performance would be considered under par in comparison to the high standards of his profession.

When the neuropsychologist offered the attorney feedback about the results, she emphasized that the attorney retained many longstanding strengths, but she explained that his performance had declined significantly on crucial measures of mental flexibility and problem-solving. The neuropsychologist pointed out that, as a result, he could probably expect to find himself at a cognitive disadvantage in comparison to his professional colleagues. The attorney's response to the discussion was initially defensive, suggesting that these measures had nothing to do with the practice of law. The neuropsychologist suggested that the attorney think it over, then return a few days later and bring his wife and adult daughter to discuss the results in more detail. In this final session, the attorney and his wife reported that they had talked over the issues and decided that, in fact, this "ministroke" served as a wake-up call that it was time to retire, and they had actually begun to discuss some of the plans for retirement that they had been delaying.

Forensic Evaluation

Neuropsychological evaluation of individuals in matters pertaining to criminal or civil law requires specialized knowledge that goes beyond general expertise in neuropsychology. Neuropsychologists are frequently called upon as experts in matters involving head injury in particular, especially in the case of mild head injury associated with a motor vehicle accident. As a distinct subspecialty, this area of practice requires integration of knowledge of statutes, laws, precedents, and legal procedures as well as expertise in identifying and describing the impact of an injury or event on cognitive, emotional, and behavioral functioning.

APPROACHES TO NEUROPSYCHOLOGICAL ASSESSMENT

The neuropsychological examination systematically assesses functioning in the realms of attention and concentration, memory, language, spatial skills, sensory and motor abilities, as well as executive functioning and emotional status. Because deficits in cognitive performance can only be interpreted in comparison to a person's longstanding or *premorbid level of functioning,* overall intellectual abilities are typically examined in order to measure the current level of overall functioning and to identify any changes in intellectual functioning. Psychological contributions to performance are also considered with regard to personality and coping style, emotional lability, presence of thought disorder, developmental history, and significant past or current stressors. The expertise of the neuropsychologist lies in integrating findings that are obtained from many diverse sources, including the history, clinical presentation, and several dozen discrete performance scores that make up the neuropsychological data.

The specific neuropsychological assessment techniques that are used in

a particular examination typically reflect the training and preferences of the neuropsychologist, as well as the referral issue. Historically, the actual practice and procedures of clinical neuropsychology developed in two general directions, which are characterized as the battery approach and the hypothesis testing approach. More detailed descriptions of these approaches are available in texts by Muriel Lezak, Diane Howieson, Erin Bigler, and Daniel Tranel, or by Igor Grant and Ken Adams.

Battery Approach

The battery approach, exemplified by the HRNTB or the *Neuropsychological Assessment Battery* (NAB), grew directly out of the psychometric tradition in psychology. This approach typically includes a large variety of tests that measure most cognitive domains as well as sensory and motor skills. Traditionally, all parts of the test battery are administered regardless of the patient's presenting problem though the NAB has a screening examination that covers all appropriate domains. The battery approach has the advantage of identifying problems that the patient might not have mentioned and that the medical history may not necessarily predict. However, it has the disadvantage of being very time consuming (i.e., 6- to 8-hour examination for the HRNTB).

Hypothesis Testing Approach

The qualitative hypothesis testing approach is historically best exemplified by the work of Luria and more recently developed as the *Boston Process Approach* by Edith Kaplan and her colleagues. It is characterized by detailed evaluation of areas of functioning that are related to the patient's complaints and predicted areas of impairment, with relatively less emphasis on aspects of functioning that are less likely to be impaired. The hypothesis testing approach has been particularly helpful in illuminating the differential roles of the two hemispheres, as discussed above. This approach has the advantage of efficiently honing in on areas of impairment and producing a detailed description of the deficits from a cognitive processing standpoint, but it has the shortcoming of potentially overlooking unexpected areas of deficits.

Screening Approaches

Many practitioners have moved away from strict battery or hypothesis testing approaches since the 1990s and developed more flexible and efficient screening approaches. These changes in clinical practice are driven by economic factors that make it less practical to evaluate patients for long periods of time, as well as the use of more efficient, empirically based clinical practice standards that draw on a widening array of well-normed clinical instruments. In this model, the neuropsychologist utilizes a core set

of screening procedures as a first step in determining whether a diagnosis can be made with less information or whether additional testing is necessary in order to identify more subtle problems. Therefore, a screening protocol that efficiently assesses the major areas of neuropsychological functioning may or may not be followed by more detailed testing in selected areas that might help to better understand the reasons for the deficits demonstrated on the screening evaluation. The *Repeatable Battery for the Assessment of Neuropsychological Status* (RBANS) and the NAB both provide examples of standardized screening approaches that have the advantage of uniform normative reference groups, although many clinicians routinely use a set of tests drawn from many sources.

Mental Status Examinations

In some cases, usually involving very acute or severe cognitive impairment, it is simply not feasible to administer extensive cognitive examination procedures, so the neuropsychologist might appropriately rely on bedside mental status examination or very brief cognitive screening procedures to address the referral issues. However, research has shown that, even with brief screening procedures, the systematic use of a structured examination format can greatly increase the accuracy of detecting cognitive impairment. Several mental status screening procedures are commonly used by a variety of health care providers in general medical and geriatric settings to provide an initial assessment of mental status. Along with a careful history, these screening procedures can serve as an indicator for more precise neuropsychological evaluation, as well as a quick baseline against which to monitor changes in cognitive functioning.

One of the most widely used screening instruments for documenting gross changes in mental status is the Mini-Mental State Examination (MMSE), although its use has declined since it is no longer in the public domain. However, two newer mental status screening approaches that are publicly available may be even more sensitive to cognitive deficits than the MMSE. Both the Montreal Cognitive Assessment (MOCA), and the St. Louis University Mental Status Examination (SLUMS) include measures of verbal fluency and clock drawing that can reflect executive functioning. The MOCA has the additional advantage of assessing abstraction ability, while the SLUMS includes items that assess logical memory and size differentiation. Like the MMSE, all of these mental status screening tools are likely to underestimate the prevalence of cognitive deficits in well-educated older persons with early Alzheimer disease or in younger adults with focal brain injury, but more likely to overestimate the presence of cognitive deficits in persons with little education. Therefore, cut-off scores should be adjusted for age and education before concluding that impairment is present. Although mental status examinations can be very useful in screening for gross signs of cognitive impairment, they do not

provide a sufficient foundation for diagnosing specific etiologies of cognitive impairment, and they are not interchangeable with neuropsychological testing.

DOMAINS OF FORMAL NEUROPSYCHOLOGICAL ASSESSMENT

The past decade has seen a virtual explosion in the growth of more sophisticated and better standardized tests and procedures for neuropsychological evaluation. A comprehensive listing of tests and techniques is beyond the scope of this chapter, but standard works that provide excellent reviews of current techniques are found in texts that are referenced in this chapter. A list of some examples of common neuropsychological tests and techniques is provided in Table 7.4–3.

Table 7.4–3.
Selected Tests of Neuropsychological Functioning

Area of Function	Comment
Intellectual Functioning	
Wechsler Intelligence Scales	Age-stratified normative references; appropriate for adults up to age 89, adolescents, and young children
Shipley Scale-II	Brief (20 minutes) paper and pencil measure of multiple choice vocabulary, open-ended verbal abstraction, and visual abstraction
Attention and Concentration	
Digit Span	Auditory-verbal measure of simple span of attention (*Digits Forward)* and cognitive manipulation of increasingly longer strings of digits (*Digits Backward*)
Digit Vigilance Test	Assesses both speed and accuracy of attention using an array of rows of randomized numbers in which a specific digit must be crossed out.
Paced Auditory Serial Addition Test (PASAT)	Requires double tracking to add pairs of digits at increasing rates; particularly sensitive to subtle simultaneous processing deficits, especially in head injury
Memory	
Wechsler Memory Scale-IV	Comprehensive set of subtests measuring attention and encoding, retrieval, and recognition of various types of verbal and visual material with both immediate recall and delayed retention; Excellent age-stratified normative comparisons for adults up to age 90 with intellectual data for direct comparison
California Verbal Learning Test-II	Documents encoding, recognition, and both immediate and 30-minute recall; affords examination of possible learning strategies as well as susceptibility to semantic interference with alternate and short forms available
Fuld Object Memory Evaluation	Selective reminding format requires patient to identify objects tactually, then assesses consistency of retrieval and storage as well as ability to benefit from cues; Normative reference group is designed for use with older individuals
Hopkins	A brief list learning paradigm in which a 12-item word list is recalled across 3 learning

Verbal Learning Test—Revised	trials. Delayed recall after 15 minutes followed by a recognition discrimination trial help isolate memory problems related to encoding, retention, and retrieval
Benton Visual Retention Test	Assesses memory for 10 geometric designs after 10-second exposures; requires graphomotor response
Brief Visuospatial Memory Test—Revised	Serial learning approach used to assess recall and recognition memory for an array of 6 geometric figures; 6 alternate forms
Language	
Boston Diagnostic Aphasia Examination	Comprehensive assessment of expressive and receptive language functions
Boston Naming Test—Revised	Documents word finding difficulty in a visual confrontation format
Verbal Fluency	Measures ability to fluently generate words within semantic categories (e.g., animals) or phonetic categories (e.g., words beginning with "S")
Token Test	Systematically assesses comprehension of complex commands using standard token stimuli that vary in size, shape, and color
Visuospatial–Constructional	
Judgment of Line Orientation	Ability to judge angles of lines on a page presented in a match-to-sample format
Facial Recognition	Assesses matching and discrimination of unfamiliar faces
Clock Drawing	Useful screening technique is sensitive to organization and planning as well as constructional ability
Rey–Osterreith Complex Figure Test	Ability to draw and later recall a complex geometric configuration; sensitive visual memory as well as executive deficits in development of strategies and planning
Motor	
Finger Tapping	Standard measure of simple motor speed; particularly useful for documenting lateralized motor impairment
Grooved Pegboard	Ability to rapidly place notched pegs in slotted holes; measures fine finger dexterity as well as eye–hand coordination
Grip Strength	Standard measure of lateralizing differences in strength
Executive Functions	
Wisconsin Card Sorting Test	Measure of problem-solving efficiency is particularly sensitive to executive deficits of perseveration and impaired ability to flexibly generate alternative strategies in response to feedback
Category Test	This measure of problem-solving ability also examines ability to benefit from feedback while flexibly generating alternative response strategies; regarded as one of the most sensitive measures of general brain dysfunction in the Halstead–Reitan battery

Trail-Making Test	Requires rapid and efficient integration of attention, visual scanning, and cognitive sequencing
Delis–Kaplan Executive Function System (D-KEFS)	Battery of measures that are sensitive to executive functions
Psychological Factors	
Beck Depression Inventory	Brief (5–10 minutes) self-report measure that is sensitive to symptoms of depression. Best for screening depression in adults up to late middle-age, who can be expected to frankly report symptoms. Available in standard (21 four-choice items) or short (13-item) form.
Geriatric Depression Scale	30-item self-report screen for symptoms of depression. The yes–no format is less cognitively demanding than other scales.
Minnesota Multi-Phasic Personality Inventory-2-RF	This recently revised psychometrically developed self-report instrument is highly useful for documenting a wide range of psychological symptoms as well as statistically determined validity indicators and information about an individual's response style. Administration time for this revised 388 version is about one-third less than the MMPI-2 (567 true–false questions, required about 1–1.5 hours).
Test of Memory Malingering (TOMM)	Effectively identifies persons who tend to exaggerate impairment using a forced choice recognition procedure based on 50 line drawings of common objects.

Interview

The clinical interview provides the single best opportunity for identifying the patient's concerns and questions, eliciting a direct description of current complaints from the patient, and understanding the context of the patient's history and current circumstances. Although the patient typically serves as the primary interview source, it is important to seek corroborating information for the patient's account from interview with caregivers or family members as well as thorough review of relevant records, such as medical and mental health treatment, educational, and employment experiences.

Intellectual Functioning

Assessment of intellectual functioning serves as the cornerstone of the neuropsychological examination. The Wechsler Intelligence Scales have represented the traditional "gold standard" in intellectual assessment for many years, based on carefully developed normative standards. The scope and variety of subtests on which the summary intelligence quotient (IQ) values are based also provide useful benchmarks against which to compare performance on other tests of specific abilities. The latest revision of this instrument, the *Wechsler Adult Intelligence Scale IV* (WAIS-IV), offers the additional advantage of greatly extended age norms (ages 16 to 89) that are directly related to normative performances on the Wechsler Memory Scale

IV. The Wechsler Intelligence Scales utilize a broad set of complex verbal and visuospatial tasks that can be summarized as Full Scale IQ, with composite scores in the domains of Verbal Comprehension, Perceptual Reasoning, Working Memory, and Processing Speed. In the context of a neuropsychological examination, the patient's performance across the procedures provides a reasonable basis for estimating an individual's level of premorbid abilities as well as the current level of intellectual functioning. Most neuropsychologists recognize that the summary Full Scale IQ provides only a ballpark range for characterizing an individual's general level of functioning, and that it is usually more appropriate and meaningful to characterize an individual's level and consistency of functioning (e.g., borderline, low average, average, high average, or superior) across the various domains that are represented by the composite scores rather than focusing on a specific IQ value.

Careful examination of the individual's performance across the various subtests within each composite area can provide information regarding the patient's pattern of strengths and weaknesses as well as the degree to which these performance characteristics are consistent with the history and performance on other aspects of the neuropsychological examination. Tests that measure longstanding, accumulated knowledge, such as fund of vocabulary or general information, provide a basis for estimating an individual's longstanding (or premorbid) level of intellectual abilities, which in turn can help to gauge the degree to which an individual may have deteriorated.

Attention and Information Processing

Attention underlies performance in virtually all other areas of functioning and should always be considered as a potential contributor to impairment on any tests that require sustained concentration and vigilance or rapid integration of new information. In a broad-based neuropsychological examination, measures of attention and concentration provide a useful basis for "previewing" the individual's ability to comprehend, process information, and otherwise engage in the assessment process. *Digit span* requires patients to repeat increasingly longer strings of digits as a way of assessing ability to process relatively simple information, while *digit span backward* reflects more complex simultaneous processing and cognitive manipulation demands or working memory. More demanding measures of attention and information processing have also been developed that are highly sensitive to subtle deficits in rapid simultaneous processing. For example, the *Paced Auditory Serial Addition Test* (PASAT) requires patients to successively sum pairs of numbers at increasingly faster rates of presentation, adding each new digit to the immediately preceding digit, in a highly demanding test of working memory. Other procedures for assessing attention and sustained concentration can be found in pencil and paper

measures of digit, letter, or symbol cancellation and various versions of computer-assisted measures of attention and vigilance.

Memory

Complaints of memory problems constitute one of the most common reasons for referral to neuropsychology. As described above, the neuropsychologist utilizes an information processing approach to assess memory problems that might involve difficulty with encoding, retrieval, or storage of new information. The *Wechsler Memory Scale IV* (WMS-IV) is the latest revision of a widely used battery of subtests that utilizes several measures of attention, memory, and new learning ability. The original version, which was developed by David Wechsler in 1945, was adapted in the 1970s by Elbert Russell to add delayed recall to the already included immediate recall. At the same time, Russell added norms for older individuals. A revised version of the battery was published in 1987, which formalized these modifications and added five summary indices, consisting of Verbal and Visual Memory, General Memory, Attention/Concentration, and Delayed Memory. The current version, published in 2009, has a modified Index structure that assesses Auditory Memory using immediate and delayed recall for a story narrative (Logical Memory I and II) and a set of word pairs (Verbal Paired Associates I and II). Visual Memory is assessed using immediate and delayed memory for the visual detail and spatial location of a set of Designs as well as Visual Reproduction of designs. Other Index domains of the WMS-IV include Visual Working Memory, Immediate Memory, and Delayed Memory. The WMS-IV offers additional advantages by extending the normative comparisons to age 90, and adding a set of mental status screening procedures (Brief Cognitive Status Exam) for older adults.

A well-standardized example of a verbal list learning technique is the *California Verbal Learning Test II* (CVLT-II). This procedure employs five learning trials of a 16-item word list, which is made up of four exemplars from each of four semantic categories. Aside from measuring several aspects of learning and memory, the CVLT-II also yields rich qualitative data about the patient's use of higher-order conceptual strategies and efficiency of verbal learning. Proactive interference is examined by use of a single free recall trial of a second list of words. Free and cued recall of the original list after both short and long delays, as well as recognition and forced choice trials, further examine memory processes of storage, retention, and retrieval. Published norms for this instrument extend from ages 16 to 89 for males and females. This revised task also offers advantages of an alternative form and a 9-item short form. Because administration time for the standard form typically requires 15 to 25 minutes (not counting a 20-minute delay interval), the short form can be particularly useful for examining more severely impaired individuals or

individuals with limited formal education.

An alternative list learning procedure to consider for older patients or those with hearing difficulty is Paula Fuld's *Object Memory Test,* in which the patient must identify (tactually or visually) 10 common objects that are hidden in a bag, then attempt to learn the list over a series of five "selective reminding" trials. In this specialized technique for measuring memory retention and storage, patients are selectively reminded at the end of each recall trial of only those words that were omitted. The normative standards for this test afford limited comparisons to older individuals who are either community-residing or who live in assisted-living arrangements. The *Hopkins Verbal Learning Test-Revised* (HVLT-R) is another useful list learning task for older or more impaired individuals in which a 12-item word list is presented in three learning trials, with follow-up measures of delayed recall after a 15-minute delay interval and a recognition discrimination trial.

A number of procedures have been standardized over the years for assessing memory for visual material. These include the *Rey–Osterreith Complex Figure Test,* which is illustrated in Figure 7.4–3, and further discussed as a visual–spatial assessment technique below. The *Benton Visual Retention Test* requires persons to draw a series of 10 geometric designs from memory, each of which is presented for only 10 seconds. Three alternate forms are available, and items are scored for errors, classified as omissions, distortions, perseverations, rotations, misplacements, and size errors. The *Brief Visuospatial Memory Test—Revised* offers six alternate forms and uses a three-trial serial learning approach to assess recall and recognition memory for an array of six geometric figures that are presented in a 2 × 3 array on a single page.

Language

Assessment of language examines both expressive abilities and comprehension. However, most neuropsychologists screen for language impairment rather than administer an extensive formal language assessment battery, such as the *Boston Diagnostic Aphasia Examination* (BDAE). Expressive language is commonly assessed by measures of *verbal fluency*. Specific procedures and normative data for several versions of verbal fluency measures can be found in the *Neuropsychological Assessment Battery* (NAB) and the Heaton norms manual. Visual confrontation naming is another commonly assessed aspect of language. The *Boston Naming Test—Revised,* provides a standard 60-item measure, and the NAB offers a shorter alternative measure of visual confrontation naming ability. Standardized evaluation of comprehension often employs procedures that systematically assess one-, two-, and three-step commands in a format, such as the *Token Test,* while the Complex Ideational Material subtest of the *Boston Diagnostic Aphasia Examination* examines an

individual's ability to comprehend more complex narrative material through responses to a series of yes–no questions.

Visuospatial Functions

Complex visuospatial abilities can be assessed through a large array of procedures that can be found in various batteries with well-developed normative data as well as stand-alone tests that vary in the quality of the norms. The perceptual reasoning measures of the WAIS-IV include Block Design, Visual Puzzles, and Matrix Reasoning, all of which can help provide understanding of an individual's ability to perform on tests of visual constructional ability and abstract concept formation. The Neuropsychological Assessment Battery (NAB) offers several standardized assessment procedures within a Spatial Module that include Visual Discrimination, based on a match-to-target paradigm, a Design Construction task that uses a set of seven geometrically shaped pieces to copy a two-dimensional design, a Figure Drawing task with immediate recall of a geometric figure, and a Map Reading task that measures visual–spatial skill, spatial/directional orientation, right–left orientation, and visual scanning. Arthur Benton's laboratory developed a set of widely used stand-alone measures, such as *Facial Recognition, Visual Form Discrimination* and *Judgment of Line Orientation*. Other measures of visual constructional ability examine the person's ability to draw spatial designs or assemble two- or three-dimensional figures (see examples in Figs. 7.4–2 and 7.4–3). In addition to the significant visuospatial component, these tasks reflect contributions of executive planning and organizational abilities. More impaired individuals can be asked to copy simple geometric forms, such as a Greek cross or intersecting pentagons in order to examine visuospatial abilities that are less influenced by planning and organization.

The widely used technique of *clock drawing* provides a surprisingly sensitive measure of planning and organization, especially for older individuals who are at risk for dementia. While problems involving poor organization, perseveration, and possible neglect are obvious in the drawing that is illustrated in Figure 7.4–7, more subtle difficulties can also be detected, especially when a patient's performance is evaluated in light of premorbid expectations. Like most neuropsychological assessment techniques, it is important to use even this relatively simple procedure in a standard manner. Most neuropsychologists use the same set of instructions each time, offering the patient a single sheet of unlined paper, and saying, "Now I would like you to draw a clock. Put all the numbers on the clock, and set the time to 10 after 11." It is best to encourage the patient to use the whole page and draw a relatively large clock, so that subtle evidence of neglect or executive deficits in organization and planning are more obvious. Although the rich qualitative information that can be gleaned from this task

may not be sufficient to stand alone for diagnostic purposes, this measure provides a useful screening technique—especially for older or more seriously impaired individuals—and can serve as a basis for deciding whether to refer the patient for more detailed neuropsychological examination.

The *Rey–Osterreith Complex Figure Test* (see Fig. 7.4–3) is a more demanding constructional task that requires the individual to copy a highly complex geometric configuration, and then draw the figure from memory. Important qualitative information about the examinee's problem-solving approach can be extracted by carefully monitoring and documenting the strategy that the examinee uses to draw the figure. In a prior discussion in this section, Figures 7.4–2 and 7.4–3 illustrate different spatial deficits that are seen on the Block Design and Rey Complex Figure constructional tasks in the presence of lateralized brain damage in either the right or the left hemispheres. It should be noted that apparent visual–spatial constructional difficulties can also be observed in the presence of frontal executive types of problems due to impaired organization and planning or poor self-monitoring.

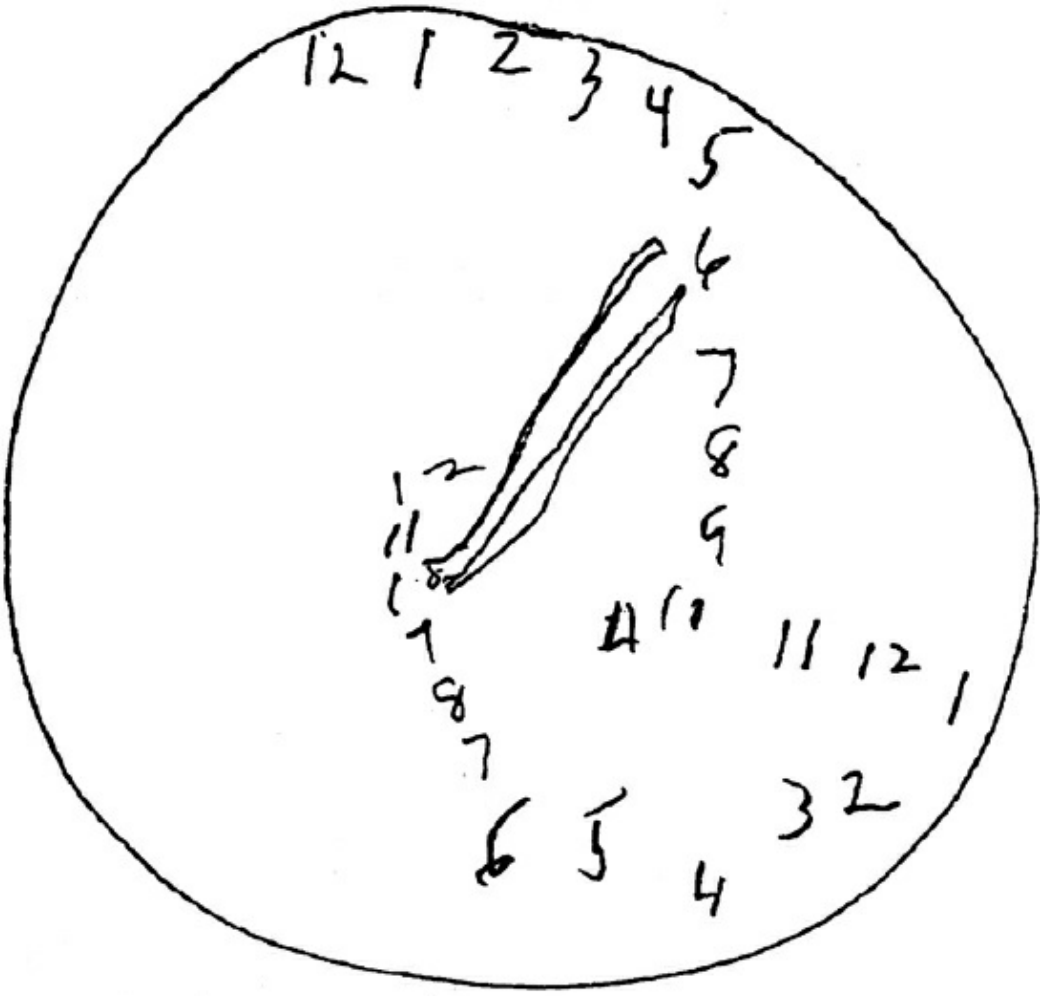

FIGURE 7.4–7. Clock drawing by a patient with vascular dementia, showing poor

planning and organization, perseveration, and possible neglect.

Sensory and Motor Functions

Double simultaneous stimulation in the visual, tactile, and auditory modalities is a standard component of the HRNTB and can be useful for assessing the integrity of basic sensory functions as well as neglect if deficits are present on one side only on the bilateral simultaneous trials and not when stimulation is unilateral. *Grip strength* and rapid *finger tapping* are commonly used measures of motor strength and speed that are sensitive to lateralized brain dysfunction. A related measure, the *grooved pegboard,* assesses the individual's ability to efficiently integrate fine finger dexterity, motor speed, and eye–hand coordination but may be bilaterally impaired following unilateral brain injury due to its high demands for cognitive integration.

Executive Functions

One of the most important aspects of the neuropsychological examination lies in the assessment of higher *executive* functions, which play an important role in the planning and initiation of independent activities, self-monitoring of performance, inhibition of inappropriate responses, switching between tasks, and planning and control of complex motor and problem-solving responses. Although the prefrontal lobes have long been regarded as an important component in mediating these functions, more recent developments in the neurosciences have also led to an increased appreciation for the essential role that is played by extensive cerebral interconnections between subcortical and cortical regions of the brain.

An underlying executive component is implied to some degree in many of the techniques discussed above. For example, complex attention, as assessed on a serial addition task, is dependent on the ability to flexibly shift focus and sustain a cognitive set. Standard expressive language measures of verbal fluency are highly dependent on the individual's ability to self-activate and maintain a productive flow of responses without losing the cognitive set. Memory functions rely on the ability to process information efficiently, including the use of higher-order processing strategies to organize incoming information (e.g., clustering words in a memory list on the basis of semantic categories). In fact, memory deficits that can be attributed to retrieval failure in the information processing model of memory are often associated with executive deficits. However, some specific techniques have gained favor over the years as primary measures of executive functions.

The *Wisconsin Card Sorting Test* (WCST) requires an individual to sort a long series of cards with designs according to various principles that can only be gleaned by incorporating feedback from the examiner. This

measure, then, assesses the individual's ability to generate and test various solutions, to flexibly respond to changing circumstances, as well as the ability to integrate new information with past experience in order to respond as efficiently as possible. The *Category Test* from the Halstead–Reitan Battery is another commonly used technique for assessing the individual's ability to flexibly incorporate feedback into problem-solving responses. Other techniques based largely on the work of Luria are more qualitative in nature and assess the individual's ability to inhibit inappropriate motor responses as well as the tendency to perseverate when asked to draw a series of repetitive graphomotor sequences. Figure 7.4–8 contains several examples of such sequences. In each case, the sample is presented with instructions for the patient to reproduce the sample and continuously repeat the segment across a page. Each prior sequence should be covered or removed from the patient's view before going on to the next, so that any tendency to reproduce elements from a prior sequence can properly be interpreted as evidence of perseveration. The *Delis-Kaplan Executive Function System* (D-KEFS) is based on a number of other techniques that are associated with executive abilities. These include variations on the *Trail-Making Test, Verbal Fluency Test, Design Fluency Test, Color-Word Interference Test, Sorting Test, Twenty Questions Test, Word Context Test, Tower Test,* and *Proverb Test*.

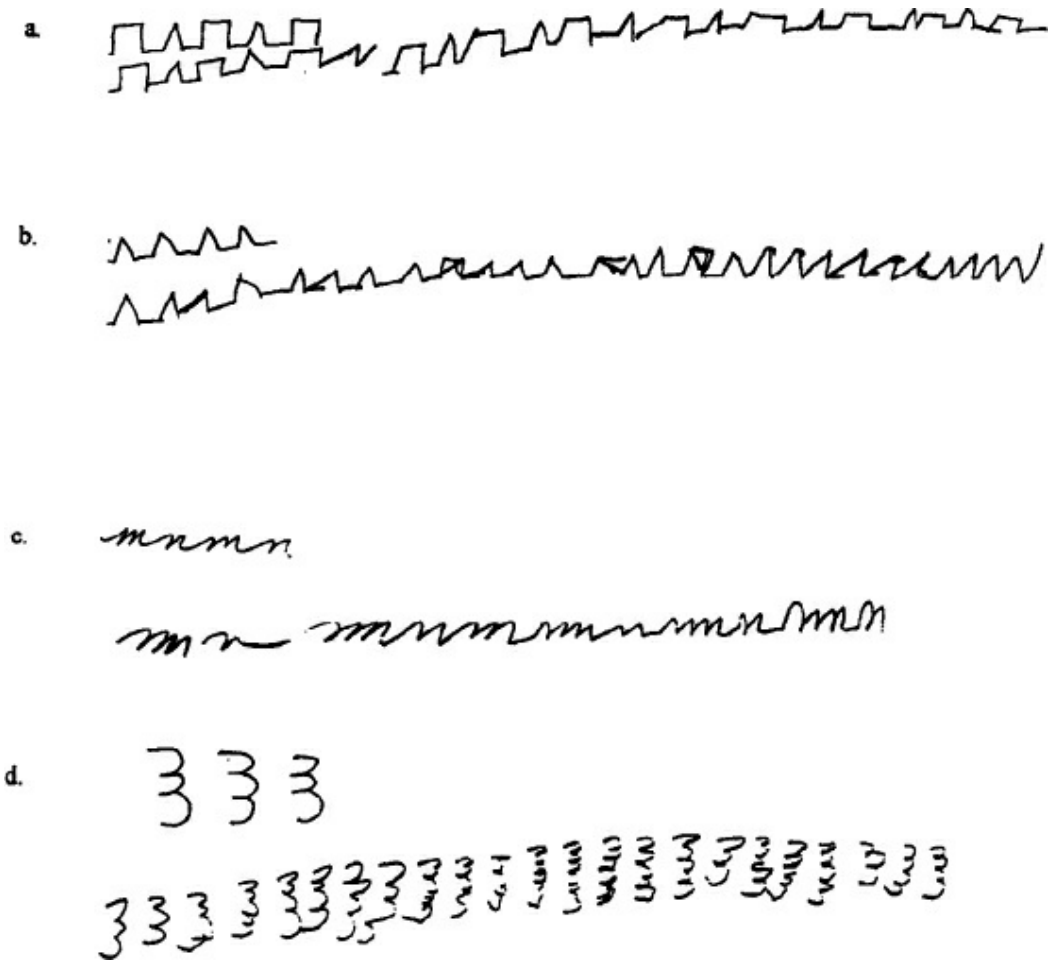

FIGURE 7.4–8. Patient reproductions of repetitive graphomotor sequences, showing perseveration of elements across designs (from **A** to **B**) and within designs (**C** and **D**).

Psychological Factors

A key component of any neuropsychological examination involves consideration of the degree to which longstanding personality or other psychological factors (including current stressors) might contribute to the patient's presentation. Common techniques for assessing personality and psychological factors include paper and pencil techniques, such as the *Beck Depression Inventory II* and the *MMPI*. The Restructured Form (MMPI-2-RF) is the most recent revision of the MMPI, and it was developed using rigorous statistical methods to generate scales with nonoverlapping items. The revised test has considerably fewer items, with only 338 items as compared to 567 items in the older version, so it takes about a third less time for most examinees to complete.

Performance Validity

Because the results of neuropsychological examinations may eventually be introduced as evidence in litigation or other forensic proceedings or be used for determining disability compensation, it is important for the neuropsychologist to address issues related to performance validity,

including the possibility of poor effort. When the test data raises concerns about performance validity, however, the neuropsychologist faces a sensitive diagnostic scenario, since it is usually not possible to differentiate between nonconscious factors that can result in poor performance versus intentional exaggeration of symptoms (e.g., factitious disorder or malingering). Slick and others have identified a set of criteria that help clinicians think about the possibility of malingering by considering discrepancies between test data and the examinee's self-report, in contrast to known patterns of brain functioning, observed behavior, reliable collateral reports and documented background history. Additional evidence of exaggerated or fabricated psychological dysfunction, as well as evidence of response bias on tests of performance validity would also add weight to the possibility of malingering.

Several instruments are available that directly assess a patient's level of effort and motivation to perform at their best. Some of these are based on memory measures, such as Green's *Word Memory Test,* and the *Test of Memory Malingering,* and others involve nonmemory tasks, such as the *Validity Indicator Profile*. Still other performance validity measures can be found embedded within other standard neuropsychological procedures. All of these procedures share a foundation based on normative research that indicates that patients with histories of bona fide brain injury or even dementia can be expected to perform at near-perfect levels on these instruments. Poor performance on any of these procedures, then, suggests that factors other than brain injury alone should be considered as possible explanations. Other indicators of performance validity are based on the pattern of an individual's performance on standard procedures in a neuropsychological examination. For example, because it is always easier to recognize than to freely recall memory material, comparing an examinee's performance on recognition and recall offers a means for assessing the possibility of poor effort. In general, a patient's level of effort is best evaluated in the context of the broader history and pattern of neuropsychological performance, so it would not make sense to administer tests of effort or "malingering" in isolation. In general, an experienced clinical neuropsychologist will avoid equating any of these measures with "malingering," however, and will instead use these results to think about factors other than brain injury or outright malingering that might influence an examinee's performance.

A 32-year-old woman with 13 years education was seen for disability evaluation, claiming current "trouble remembering things." Her account of personal history was vague, and she "forgot" information, such as her own birth date and mother's maiden name. Response latencies were extremely long, even for highly familiar information (e.g., count from 1 to 20), she could not repeat more than three digits forward consistently, and on a word list learning procedure, she was not able to correctly recognize more items (only five) than she could freely recall (also five). Despite otherwise fluent language, she was only able to generate five

examples of animals in 1 minute. On forced choice memory measure, she performed far below the chance levels of performance that would be expected even if a person had not been presented with the 50 pictures before. When asked to recall 15 items on a procedure (Rey's 15-Item Memory Test) that is presented as a challenging task, but in reality is fairly simple, her performance demonstrated exaggerated errors of commission (Fig. 7.4–9). The evaluation concluded that current levels of cognitive functioning could not be conclusively established, due to overt symptom exaggeration.

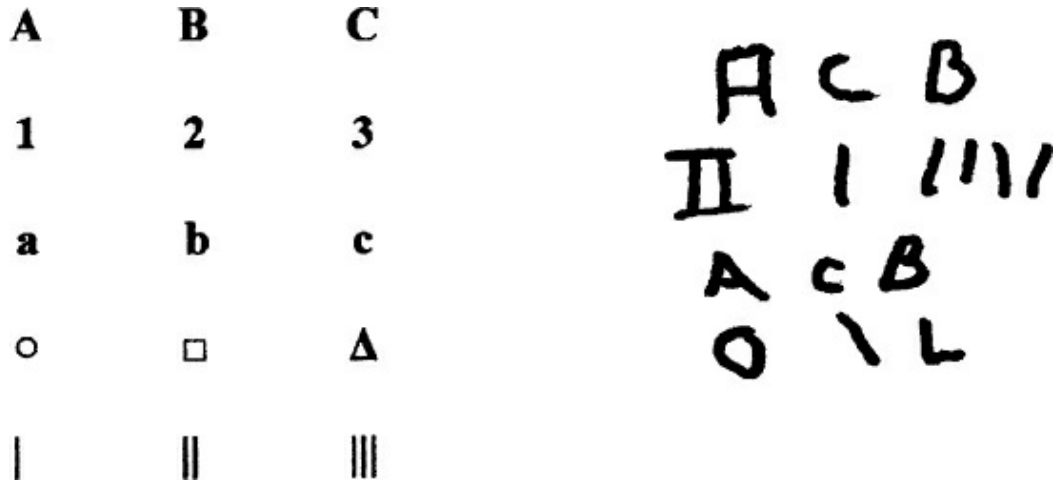

FIGURE 7.4–9. Rey's 15-Item Memory Test with example of a response that is typical of exaggerated "memory" problems.

THERAPEUTIC DISCUSSION OF RESULTS

A key component of the neuropsychological examination process is found in the therapeutic opportunity to discuss results of the examination with the patient and family or other caregivers. This meeting can represent a powerful therapeutic opportunity to educate and clarify individual and relationship issues, which can impact on the identified patient's functioning. If the patient's active cooperation in the initial examination has been appropriately enlisted, then the patient will be prepared to invest value and confidence in the findings of the examination. At the time of the results discussion, it is useful to review the goals of the examination with the patient and supportive family or caregivers and to clarify the expectations of those who are present. Typically, these sessions will include information about the patient's diagnosis, with emphasis upon the natural course and prognosis as well as compensation and coping strategies for the patient and family. Given the impact of chronic neurological disease upon the family system as well as the patient, explicit discussion of these issues is critical in maximizing adjustment to brain injury. It is equally important to discuss the impact of the results on the patient's current living circumstances, future goals, and course of adjustment. It is not unusual for strong emotions and underlying tensions within family relationships to come to light in the context of honest discussion, and so the results discussion can represent an important therapeutic opportunity to model

effective communication and problem-solving techniques.

The most challenging discussions usually arise when the objective examination results do not meet the preconceived beliefs of the patient or family members. In such cases, it is helpful to begin the feedback discussion by first asking the patient to describe their impression of their own performance in order to explicitly identify the patient's perspective and associated expectations. For those persons who appear to be strongly invested in confirming a pathological diagnosis of memory impairment, it may be particularly helpful to adopt a psychoeducational approach by using materials such as the Information Processing Model of Memory that is presented in Figure 7.4–5.

References

American Bar Association Commission on Law and Aging and American Psychological Association Assessment of Capacity in Older Adults Project Working Group. *Assessment of Older Adults with Diminished Capacity: A Handbook for Psychologists*. Washington, DC: American Bar Association and American Psychological Association; 2008.

American Psychological Association. Ethical principles of psychologists and code of conduct. *Am Psychol*. 2002;57:1060–1073.

Aupperle RL, Melrose AJ, Stein MB, Paulus, MP. Executive function and PTSD: Disengaging from trauma. *Neuropharmacol*. 2012;62:686.

Benton AL, Sivan AB, Hamsher K deS, Varney NS, Spreen O. *Contributions to Neuropsychological Assessment: A Clinical Manual*. 2nd ed. New York: Oxford University Press; 1994.

Boone KB. *Assessment of Feigned Cognitive Impairment*. New York: Guilford; 2007.

Castellani RJ. Chronic traumatic encephalopathy: A paradigm in search of evidence. *Lab Invest*. 2015;95:576.

Cummings J. Frontal-subcortical circuits and human behavior. *Arch Neurol*. 1993;30:873–880.

Dronkers NF, Ogar J. Aphasia: Overview. In: Aminoff MJ, Daroff RB eds. Encyclopedia of the Neurological Sciences. 2nd ed, Vol 1; New York: Oxford: Academic Press; 2014.

*Grant I, Adams IM. *Neuropsychological Assessment of Neuropsychiatric & Neuromedical Disorders*. 3rd ed. New York: Oxford University Press; 2009.

Haaland KY, Harrington DL. Hemispheric asymmetry of movement. *Curr Opin Neurobiol*. 1996;6:796–800.

Haaland KY, Mutha PK. Apraxia, Limb. In: Aminoff MJ, Daroff RB, eds. *Encyclopedia of the Neurological Sciences*. 2nd ed. Vol. 1; New York: Oxford: Academic Press; 2014.

Hannay HJ. Proceedings of The Houston Conference on Specialty Education and Training in Clinical Neuropsychology. *Arch Clin Neuropsychol*. 1998;13:157–158.

Heaton RK, Miller SW, Taylor MJ, Grant I. *Revised Comprehensive Norms for an Expanded Halstead-Reitan Battery*. Lutz, FL: Psychological Assessment Resources; 2004.

*Heilman KM, Valenstein E. *Clinical Neuropsychology*. 5th ed. New York: Oxford University Press; 2011.

Lamberty GJ. *Understanding Somatization in the Practice of Clinical Neuropsychology*. New

York: Oxford University Press; 2008.

Larrabee GJ. *Assessment of Malingered Neuropsychological Deficits*. New York: Oxford University Press; 2007.

*Lezak MD, Howieson DB, Loring DW. *Neuropsychological Assessment*. 5th ed. New York: Oxford University Press; 2012.

Libon DJ, Delano-Wood L, Bondi MW, Au R. Mild cognitive impairment. In: Aminoff MJ, Daroff RB, eds. *Encyclopedia of the Neurological Sciences*. 2nd ed. Vol. 3; New York: Oxford: Academic Press; 2014.

Loring DW, ed. *INS Dictionary of Neuropsychology*. 2nd ed. New York: Oxford University Press; 2015.

Luria AR. *Higher Cortical Functions in Man*. Haigh B, translator. New York: Basic Books; 1966.

McCrea MA. *Mild Traumatic Brain Injury and Postconcussion Syndrome: The New Evidence Base for Diagnosis and Treatment*. New York: Oxford University Press; 2008.

Mitrushina MN, Boone KB, Razani J, D'Elia LF. *Handbook of Normative Data for Neuropsychological Assessment*. 2nd ed. New York: Oxford University Press; 2004.

*Morgan JE, Ricker JH. *Textbook of Clinical Neuropsychology*. New York: Psychology Press; 2008.

Parsons MW, Hammeke TA. *Clinical Neuropsychology: A Pocket Handbook for Assessment*. 3rd ed. Washington, DC: American Psychological Association; 2014.

Stewart S, O'Riley A, Edelstein B, Gould C. A preliminary comparison of three cognitive screening instruments in long term care: The MMSE, SLUMS, and MOCA. *Clin Gerontol*. 2012;35:57–75.

*Strauss E, Sherman EMS, Spreen O. *A Compendium of Neuropsychological Tests*. 3rd ed. New York: Oxford University Press; 2006.

Stuss DT, Knight RT. *Principles of Frontal Lobe Function*. New York: Oxford University Press; 2002.

Suchy Y. *A Comprehensive Guide for Clinical Practice*. New York: Oxford University Press; 2015.

Wechsler D. *Wechsler Adult Intelligence Scale—Fourth Edition*. San Antonio, TX: Pearson; 2008.

▲ 7.5 Personality Assessment of Adults

RUSSELL L. ADAMS, PH.D., AND DAVID L. ADAMS, PSY.D.

INTRODUCTION

Personality has been defined in various ways over the years. For purposes of this chapter, personality is defined as an individual's enduring and pervasive motivations, emotions, interpersonal styles, attitudes, and traits. The systematic measurement of these personality traits is the purview of personality assessment. Personality tests measure such difficult to define concepts as depression, anger, and anxiety. The systematic measurement of personality traits such as suicide potential, impulsivity, and the ability to delay gratification can be challenging as well. Nevertheless, personality assessment is of utmost importance in the scientific study of psychology and psychiatry. Personality assessment is equally important in psychotherapy. In addition, psychological tests are regularly utilized in the medical/legal system. The effectiveness of various treatments or psychiatric medications is often evaluated through psychological assessment instruments. For example, to determine if a given depression treatment is effective a researcher must first quantify the degree of depression present in a treatment and control group. After a given treatment is applied, follow-up depression measures are used to assess change from the baseline. Appropriate statistical techniques are then utilized to see if the changes noted are statistically or practically significant. Similarly, in clinical practice a clinician might use a personality test such as the Beck Depression Inventory to measure the level of depression of a given patient at the beginning of treatment. The Beck Depression Inventory would again be administered periodically throughout treatment to assess change in the level of depression over time.

There are literally thousands of different personality tests available today. As a result, this section will focus on some of the more commonly used personality instruments that will be presented in both narrative and tabular format.

This section will also present the reader with a brief scientific overview of how personality tests are constructed and how they are scientifically evaluated in terms of their reliability and validity.

PURPOSES OF PSYCHOLOGICAL TESTING

Personality testing can be an expensive undertaking. A considerable amount of time is required to administer, score, and interpret psychological test results. Personality testing should not be routinely obtained from all psychiatric patients especially in today's cost-containment environment.

Nevertheless, personality testing can be helpful with selective patients from both a clinical and a cost–benefit analysis perspective.

Assisting in Differential Diagnosis

As any clinician knows, psychiatric diagnosis can be a difficult and, at times, confusing exercise. However, knowing a patient's diagnosis is essential to treatment, as a proper diagnosis can assist in understanding the etiology of the presenting psychiatric problem and the prognosis of the disorder. Consider the diagnostic dilemma in the following example:

> A 49-year-old white man had abruptly resigned his position as an accountant and decided he was going to start an oil exploration business. He had never worked in the oil business and knew nothing about the profession. The patient had received a revelation from an unknown entity through an auditory hallucination. This voice told him he would become quite wealthy in the business if he would simply follow the directions given to him. Around this time, the patient had a marked change in personality. Although his grooming was formally very neat and appropriate, he became disheveled. He began sleeping about 3 hours a night. He became somewhat agitated and talked loudly to those around him.

The differential diagnosis in this case illustration includes schizophrenia and bipolar disorder. Psychological testing might be helpful in assisting in this differential diagnosis, as well as in formulating a treatment plan.

Aiding in Psychotherapy

Psychological tests can be useful in psychotherapy. The usefulness of these tests can be even more important for short-term, problem-centered therapy where understanding the patient and his or her problem must be accomplished quickly. Psychological assessment can be used in pretreatment planning, assessing progress once therapy begins, and in evaluating the effectiveness of therapy. Patients need to have objective information about themselves at the time of therapy if they are to go about changing themselves productively. Personality tests, particularly objective tests, allow patients to compare themselves to objective norms and evaluate the extent and magnitude of their problem. Testing also can reveal areas of the patient's life that may be problematic but for which the patient may not have a full appreciation. Information about patients' willingness to reveal information about themselves can also be helpful. In some instances, the test report can be the subject of an early treatment session. Psychological tests may reveal considerable information concerning the patient's inner life, feelings, and images, which may make therapy progress faster. Psychological testing can provide baseline information at the beginning of therapy, and repeat testing can then be used to assess change that occurred during the course of therapy.

Providing Narrow-Band Assessment

Narrow-band personality tests measure a single personality characteristic or a few related characteristics. Broad-bank personality tests, on the other hand, are designed to measure a wide spectrum of personality characteristics. A psychiatrist may need answers to specific questions, such as those that arise when assessing the degree of clinical depression, measuring the intensity of the state or trait anxiety, or, possibly, quantifying the amount of a patient's anger. Such quantification can be helpful in measuring severity or in providing a baseline for future assessment.

Psychological tests can often assist with a specific problem area. For example, a young college student may be unsure what to major in. Vocational interest inventories, such as the Strong Vocational Interest Inventory and the Kuder Occupational Interest Survey, may help to better define the student's vocational interest. The Strong Vocational Interest Inventory compares a given individual's responses with those of a large group of successful individuals in a variety of occupations. Test results can show that a student's interests are similar, for example, to those of successful persons in occupations such as social work or public administration and dissimilar to those of engineers or scientists.

PSYCHOMETRIC PROPERTIES OF PERSONALITY ASSESSMENT INSTRUMENTS

The quality of personality tests varies widely. On the one hand, there are well-constructed, empirically validated instruments, and, on the other hand, there are "psychological tests" that one can find in the Sunday supplement of the newspaper or on the Internet. Evaluating the usefulness of particular psychological instruments can be challenging, even to the well informed.

Just as the *Physician's Desk Reference* helps physicians evaluate the usefulness of medications, the *Buros Mental Measurement Yearbook* helps psychologists evaluate the usefulness of personality instruments and other psychological and achievement tests. The *Mental Measurement Yearbook* is a reference source of critical reviews of individual tests. These reviews evaluate the scientific underpinnings of each test by an expert in the field. The reviews generally address such issues as the reliability and validity of the tests, the normative group used in construction, and various empirical studies utilized to check out the reliability and validity of the instrument. The advantages and disadvantages of the test are described. The reviews generally describe situations in which the test can be appropriately used. Practical information concerning the test publisher, length of time of administration, and other considerations are also included. The *Buros Mental Measurement Yearbook* is available in most medical school and university libraries and can be ordered through the book's publisher.

Normative Sample

To construct a personality test, a representative sample of subjects (normative sample) should be administered the test to establish expected performance. Basic issues, such as the size and representativeness of the sample used to construct the test, must be evaluated. A sample of 30 individuals of a given diagnosis in a single location is seldom appropriate, because there are so many moderating factors impacting results, such as age, education, sex, and racial and ethnic grouping. These factors frequently impact the norms, and appropriate controls must be made for these variables in many cases.

To illustrate this point the Minnesota Multiphasic Personality Inventory 2 (MMPI-2), a well-constructed instrument, initially tested approximately 2,900 subjects. However, approximately 300 subjects were eliminated because of test invalidity or incompleteness of needed information. The final group was comprised of 2,600 community volunteers, including 1,462 women and 1,138 men. The MMPI-2 normative group was selected using census data as a guide and was broadly representative in terms of region of the country, racial composition, age, and formal education. Despite efforts to control for moderating variables, the sample was slightly over representative in higher education levels. Obviously, the cost of obtaining large census-based normative samples is enormous, and few tests are able to meet these challenging criteria. Nevertheless, many tests that fall short of the national sample (and most do) can be useful in the hands of an experienced clinician who is sensitive to the problem and makes appropriate clinical adjustments in interpreting the results.

Test Characteristics

To be useful, any psychological test must be completed, in its entirety, by the intended test taker. If the questions are offensive or are difficult to understand, then the individual taking the test may not complete all items. These omissions can create problems, especially when normative tables are used to interpret results. To be useful, an instrument should be short enough to complete in a relatively short period of time. The Beck Depression Inventory or the Geriatric Depression Scale can each be easily completed in 5 to 10 minutes. Because of its short length, the test can be given repeatedly to a depressed patient in treatment to monitor changes over time. The MMPI-2, on the other hand, takes approximately 1 to 1½ hours to complete. This test is not appropriate for repeat testing over a short period of time. Most tests are easily scored and compared to normative tables.

Validity Issues

Perhaps the most important characteristic in evaluating the scientific merit

of a given personality test is the validity of the instrument. Does the test measure what it purports to measure? If a test is designed to measure depression, does it indeed measure depression? Although validity may seem like a simple issue to address, it can become complex, especially when attempting to measure such characteristics as self-esteem, assertiveness, hostility, or self-control. The test developer must first operationally define what is meant by the personality characteristic in question. A number of different approaches have been used to measure validity, including face validity, logical content validity, construct validity, criterion validity, and factor validity.

Face Validity. *Face validity* refers to the content of the test items themselves. In other words, do the items appear to measure what they purport to measure? For example, does a depression scale ask questions concerning sadness, guilt, sleep disturbance, weight loss, helplessness, hopelessness, or suicidal thoughts? One problem with face validity is that professionals differ in their subjective appraisal of individual items. One psychologist may think a given item does, indeed, measure the concept under consideration, while other professionals may not. To address these differences in definition, the test developers may utilize an approach known as *logical content validity*. Using this approach, the developer will describe in the manual the procedure used to select items for the test that will not depend significantly on the subjective judgment of any one professional. For example, in developing the Taylor Manifest Anxiety Scales, a number of different psychologists were asked to review items from the MMPI that seemed to measure a formal definition of manifest anxiety, and only those items agreed on by this group of psychologists were included. Many creative approaches have been used to demonstrate the logical content validity of various tests.

Criteria and Construct Validity. Although face validity refers to the degree that test items appear on the surface to measure what the instrument, as a whole, purports to measure, *criterion validity* uses data outside the test itself to measure validity. For example, if a test were designed to measure hypochondria, one would expect that a patient with high scores would have more visits to the physician's office, complain of more physical symptoms, and use prescribed and over-the-counter medications more extensively. These criteria are objective and external to the test itself and can be independently measured. Scores on the hypochondria instrument can be statistically correlated with a number of separate objective criteria, such as those mentioned previously, to determine if the test measures the theoretical construct being measured (*construct validity*).

Concurrent and Predictive Validity. To determine test *concurrent validity,* external measures are obtained at the same time that the test is given to the sample of subjects. Thus, the concurrent validity of the test reveals that, at a given point of time, high scorers on a test may be more likely than low scorers on a test to manifest the behavior reflected in the criteria (e.g., more physician visits or more medication for a hypochondriac patient). Occasionally, however, a test developer is interested in predicting future events. For example, a test designed to predict marital satisfaction among engaged couples may use the couples' future marital satisfaction and marital status as an external criterion. Thus, the external criterion would have to be obtained a number of years down the road for the test to have this predictive validity. This type of validity is referred to as *predictive validity*. Because it is difficult to obtain such information several years later, predictive validity is seldom available in most personality tests, although such data can be valuable when available.

The *discriminant validity* of a test tells whether the test is able to discriminate between known groups of patients at a given time. For example, is a test of depression able to statistically discriminate nondepressed subjects from subjects who suffer from independently diagnosed major depression, adjustment disorder with depressed mood, and dysthymia? Is a measure of depression able to statistically discriminate among mild, moderate, and severe major depression disorder?

Factor Validity. Factor validity utilizes a multivariate statistical technique known as *factor analysis* to determine if certain major groups of items on a given test empirically cluster together. For example, on a personality test measuring depression, do items concerning vegetative symptoms tend to covary together? Are there other clusters of items on the depression test measuring cognitive problems, mood disturbance, and social withdrawal?

Reliability

To be useful, a test must not only be valid, but it must also be reliable. *Reliability* refers to the degree that a test measures what it purports to measure, consistently. The key word here is *consistently*. For example, a patient with agoraphobia takes a new agoraphobia test today and takes the same test tomorrow. However, on the second testing the scores change markedly even though the patient had not received any treatment or undergone any change in his or her underlying agoraphobia. Such a test would not be useful, because it would not consistently measure the trait in question. The test would have poor reliability. There are several means of checking reliability, including test–retest reliability, internal consistency reliability, and parallel form reliability.

Test–Retest Reliability. *Test–retest reliability* is obtained by simply administering the same test on two occasions to a group of subjects and statistically correlating the results. To be useful, the correlation coefficient should be at least 0.80 if the two tests were administered within 2 weeks of each other and if the trait in question is stable.

Internal Consistency Reliability. Another approach to determine *internal consistency reliability* is to divide a given test into two equal parts and statistically correlate the two halves for the test with each other. This technique determines the *split-half reliability* of a test. The first half of the test should be highly correlated with the second half of the test if the test is consistently measuring what it purportedly measures. Alternatively, the odd-numbered items could be correlated with the even-numbered items (*odd–even consistent reliability*). When a test is split into two parts and the obtained scores correlate with each other, the resulting correlation coefficients statistically underestimate the true reliability of the test because the number of items in the test influences the magnitude of the correlation. Hence, if there were only half that number of items in the original test, then the results would be underestimated. To correct for this underestimation, the *Spearman–Brown formula* is used. Again, a reliability coefficient of 0.80 to 0.85 is needed to demonstrate usefulness in most circumstances. However, the higher the reliability as measured by the correlation coefficient, the better the test instrument. No test is perfectly reliable. The highest possible correlation coefficient is 1.0, and the lowest is 0, assuming no negative correlation.

To determine how internally consistent a given test is, a test developer also may use a statistical technique called *Cronback's coefficient alpha*. With Cronback's coefficient alpha statistic, one can determine the average correlation among all the items on the test. The higher the average correlation, the better the reliability.

Parallel Form Reliability. Sometimes, two separate forms of the same test are needed. For example, if the process of taking a test at one point in time would by itself influence a patient's score the second time that they took the same test, then parallel forms of the tests are needed. *Parallel forms* of a test measure the same construct but use different items to do so. To ensure that the test does, in fact, measure the same construct, the correlation coefficient between the two parallel forms of the same test is computed. Such parallel form reliability should be at least 0.90 or higher.

Use of Standard Error of Measurement to Assess Reliability. Another way to assess the usefulness of a given test is to examine the test's standard error of measurement (SEM), which should be included in the test's manual. The SEM is a single statistic that is used to estimate what the score

of a given patient would be on the test if the patient took the same test again within a short period of time. Consider the following example: A patient takes a given test today and obtains a score of 90 on that test. The SEM on that test is 3. An SEM value of 3 together with an individual's score of 90 on the test means that if that patient were to take that same test again tomorrow, he or she would score within ±1 SEM of their original test performance (e.g., ±3 points) 68 percent of the time. Thus, the patient in question should score between 87 (90 − 3) and 93 (90 + 3) on the second test 68 percent of the time. The same individual would score ±2 SEM values (e.g., 2 × 3 = 6) on the second test 95 percent of the time.

On the above example, the patient with an original score of 90 on a given test would score between 84 (90 − 6) and 96 (90 + 6) on the second test again 95 percent of the time. Other things being equal, the smaller the SEM relative to the standard deviation of the tests, the better.

ADULT PSYCHOLOGICAL TESTS

Objective Personality Tests

Objective personality tests are rather straightforward in approach. Patients are usually asked specific and standard questions in a structured written or oral format. Each patient is typically asked the same question. The data obtained from a given patient are compared to similar data obtained from the normative group. The degree to which the patient deviates from the norm is noted and is used in the interpretive process. The patient's responses are scored according to certain agreed-upon criteria. The obtained scores are then compared with normative tables and often converted to standardized scores or percentiles, or both. The MMPI-2 is an example of an objective personality test.

There are literally hundreds of different objective personality instruments, and discussing all of these instruments is beyond the scope of this section. Table 7.5–1 contains a representative sample of objective personality tests and concise descriptions of each test, along with a brief listing of strengths and weaknesses. Because the MMPI-2 is, by far, the most frequently used objective personality test, it is discussed in considerable detail in this section to give the reader an orientation to objective personality assessment, as well as objective personality development.

This focus on the MMPI-2 is not meant to imply that the MMPI-2 is the only valid and reliable objective personality instrument or that other objective personality tests are not as useful as the MMPI-2 for a particular purpose. There are many useful objective personality instruments.

Minnesota Multiphasic Personality Inventory. Two excellent books on the MMPI-2 are James N. Butcher's *The MMPI-2 in Psychological*

Treatment and John R. Graham's *MMPI-2 Assessing Personality and Psychopathology*. The following discussion relies heavily on information contained in these two resources. The MMPI-2 consists of 567 true/false questions concerning a wide variety of issues. The patient completes the test using a pencil. The MMPI-2 is relatively easy to administer and score and takes approximately 1 to 1½ hours for most patients to complete. It requires only an eighth-grade reading comprehension. The test provides a number of internal checks for measuring the patient's degree of cooperation. The patient is simply asked to answer true and false to such questions as "I have a good appetite," or "No one seems to understand me."

Scoring of the MMPI-2 involves adding up the number of responses on numerous scales and comparing the results to certain normative information. Interpretation of the MMPI-2 is more straightforward than with many other tests, because of the vast amount of empirical information available concerning the MMPI-2. More than 12,000 articles and books have been published on this test to date.

There are over 115 translations of the MMPI-2. The test is also cost-effective, because relatively little professional time is involved in the administration and scoring. The patient is simply given instructions, and then he or she completes the test independently in the professional's office. In certain situations, the MMPI-2 can be administered on the computer. A computer can also plot test profiles and compute various supplemental scales. The number of available specialized scales increases the usefulness of the MMPI-2 with various populations (e.g., patients with PTSD, college students who are maladjusted, and prisoners with violent tendencies).

Questions cover a wide range of different areas. When a patient takes the MMPI-2, questions are not grouped in any particular order to aid in interpretation. Various items in the MMPI-2 can be selected, sorted, and analyzed according to various criteria. The most common scales of the MMPI-2, including the validity scales, clinical scales, content scales, supplemental scales, and critical items are discussed below.

VALIDITY SCALES. The patient can approach a mental status interview in a variety of ways. The patient can be painfully honest and straightforward with the interviewer, revealing his or her innermost aspirations and fears. The patient may be somewhat distrusting of the interviewer and may try to slant the content of the interview in a certain fashion to manipulate the outcome he or she wants. The patient may be disingenuous or even attempt to be deceitful by malingering. Obviously, the information obtained by psychiatrists or other mental health professionals in an interview may well be colored by the patient's approach to the interview. Some patients are, by their nature, reserved and may use denial successfully during a clinical interview. Other patients may have little insight into their feelings and thoughts. Many times a psychiatrist may not know how candid or

straightforward a patient is during an interview.

When patients are administered personality tests like the MMPI-2, they approach the test in a similar fashion to how they approach an interview. One advantage of the MMPI-2 is that the test has been devised to assess different test-taking attitudes (e.g., fake bad, fake good, or social desirability responding). Through use of the validity scales, the clinician gleans information that helps in interpreting the remaining scales of the MMPI-2. Is the patient overly defensive? Is he or she attempting to fake good or fake bad? Does the patient maintain a reasonable balance between self-disclosure and excessive defensiveness?

Table 7.5–1.
Objective Measures of Personality

Name	Description	Strengths	Weaknesses
Minnesota Multiphasic Personality Inventory-2 (MMPI) and MMPI-2-RF	567 Items; true–false; self-report format; 20 primary scales; MMPI-2-RF is also available, which has 338 of the same items as the MMPI-2	Large research base, most known and utilized objective personality assessment instrument; provides considerable clinically useful information while requiring limited clinician time.	Required that the patient be able to read items, comprehend instructions, and follow direction; some patients attempt to fabricate responses but validity scales are available
Millon Clinical Multiaxial Inventory (MCMI-IV)	195 Items; true–false; self-report format; 25 primary scales	Brief administration time; corresponds will with DSM-IV	High degree of item overlap in various scales; may over diagnose pathology in some cases through many scale elevations.
16 Personality Factor Questionnaire (16 PF)	True–false; self-report format; 16 personality dimensions	Sophisticated psychometric instrument with considerable research conducted on nonclinical populations	Limited usefulness with clinical populations
Personality Assessment Inventory (PAI)	344 Items; Likert-type format; self-report; 22 scales	Includes measures of psychopathology, personality dimensions, validity scales, and specific concerns to psychotherapeutic treatment	The inventory is increasingly being utilized, but it does not currently have the strong research basis found in the MMPI-2
Personality Inventory for DSM-5 (PID-5)	220 Items; Likert-type format; 25 personality trait facets with 5 broad trait domains	Corresponds to the 25 personality traits outlined in the DSM-5	Additional research warranted
NEO Personality Inventory-Revised (NEO-	240 Items with 3 validity scales	Corresponds to the Five Factor Model for personality l	Requires a sixth-grade reading level

PI-R)			
California Personality Inventory (CPI)	434 questions; true–false; self-report format; 20 "folk" scales of character with 4 classes	Well-accepted method of assessing patients who do not present with major psychopathology	Limited usefulness with clinical populations
Jackson Personality Inventory (JPI–R)	300 True–false; self-report format; 15 personality scales	Constructed in accord with sophisticated psychometric techniques; controls for response sets	Unproved usefulness in clinical settings
Edwards Personal Preference Schedule (EPPS)	Forced choice; self-report format	Follows Murray's theory of personology; accounts for social desirability; measures 15 personal needs; often used in vocational counseling	Not widely used clinically because of restricted nature of information obtained
Psychological Screening Inventory (PSI:2)	139 Items; true–false; self-report format	Yields 4 Basic Screening Scales, 21 Brief Content Scales, and 6 Misrepresentation Scales; can be used as screening measures concerning the possibility of a person's need for psychological treatment	The scales are short and have correspondingly limited reliability
Eysenck Personality Questionnaire (EPQ)	Utilizes Yes/No questions (48 questions in short version; 100 in longer version); self-report format	Useful as a screening device; test has a theoretical basis with research support; measures personality traits using 4 scales	Scales are short, and items are transparent as to purpose; not recommended for other than a screening device
Adjective Checklist (ACL)	True–false; self-report or informant report	Can be used for self-rating or other rating; measures 37 traits and each of these scales are grouped into 6 Success Factors at work	Scores rarely correlate highly with conventional personality inventories
Comrey Personality Scales (CPS)	True–false; self-report format; 10 scales including the validity check scales; 180 items	Factor analytic techniques used in test construction	Not widely used; factor analytic interpretation problems
Tennessee Self-Concept Scale (TSCS:2), 2nd ed.	100 Items; true–false; self-report format; 15 scales	Brief administration time yields considerable information; measures self-concept in 15 scales which includes 4 validity scales	Brevity is also a disadvantage, lowering reliability and validity; useful as a screening device only

Adapted and revised from Robert W. Butler, Ph.D. and Paul Satz, Ph.D.

Lie (L) Scale. This scale consists of 15 items designed to measure general frankness. The items include questions concerning personal characteristics that, although highly desirable in our culture, are, nevertheless, infrequently practiced (e.g., questions such as, "I have never failed to tell the truth," or "Once in a while I think of things too bad to talk about," and

"I do not always tell the truth"). Patients who score high on the lie scale may be psychologically naïve or defensive. They may want to be seen in a positive light or they may have difficulty admitting common human frailties.

F-Scale. This is one of the longest scales, consisting of 60 items that 10 percent or fewer of the normative samples responded to in the positive direction. This scale includes such items as follows: "Evil spirits possess me at times," and "Most any time I would rather sit and daydream than do anything else."

The items on the F-Scale are quite diverse in content and may be indicative of severe psychopathology such as hallucinations or delusions. Most people in the normal sample responded to fewer than five of these items. Individuals who obtain exceptionally high scores on this scale are often either extremely pathological or trying to claim mental illness falsely. As with all personality tests, interpretation depends greatly on the setting in which the test was given, as base rates for different settings and circumstances differ. For example, patients who have been seen for court sanity evaluations have different base rates of symptom magnification than do patients seen for adjustment problems in an outpatient clinic.

Backpage Infrequency (Fb) Scale. A few patients, particularly those with passive–aggressive tendencies, will start to complete the MMPI-2 but, toward the end of the test, may respond to the items in a random fashion. The patient may simply want to finish the test quickly and leave the testing situation. The Backpage Infrequency (Fb) scale for the MMPI-2 is designed to pick up this tendency.

Variable Response Inconsistency Scale (VRIN). The VRIN consists of 67 pairs of items. Each question of the pair probes for the same or similar information in a slightly different way, for example: (1) "I sleep very poorly at night," and (2) "I wake up fully refreshed each morning." If the patient responds "true" to question 1, then he or she would normally respond "false" to question 2. Failure to do so would reflect inconsistency.

CLINICAL SCALES. The clinical scales are generally considered the meat or substance of MMPI-2 clinical interpretations. The original clinical scales have an empirical base in the sense that the items were selected because they differentiated between specific clinical groups and the normative control group. The control group for the original MMPI, developed in the 1940s, consisted of 1,500 persons who were visitors to the University of Minnesota hospitals and other normal groups. Clinical groups varied in size but, for the most part, contained at least 50 subjects. For example, the schizophrenic scale group was developed with patients with different types of schizophrenia, and the depression scale was developed with a sample of psychiatric patients with depression, including some with manic depressive

illness. Table 7.5–2 presents a brief illustration of the original sample groups. Although most clinical scales were developed using a specific clinically defined group, researchers with the MMPI soon learned that a given patient (such as one with schizophrenia) did not simply score high on the schizophrenia scale and low on all other scales. A schizophrenic patient might score high on many scales, but his or her highest score may well be the schizophrenic scale. If one were to diagnose all patients who scored high on the schizophrenic scales as schizophrenic patients, then a large number of false positives would result. To address this confusion, clinicians began to refer to the clinical scales by a number (0 to 9) rather than the diagnostic entity that underlies the scale's development.

Table 7.5–2.
Minnesota Multiphasic Personality Inventory Scale Development

Scale 1—(Hs: Hypochondriasis): The clinical group of this scale consisted of patients who had an obsessive overconcern about their physical health. This group had little or no organic basis for these complaints.

Scale 2—(D: Depression): This group consisted of a variety of patients with symptomatic depression, including reactive depression and manic depression among others.

Scale 3—(Hy: Hysteria): The hysteria group consisted of patients who had some type of sensorimotor deficit for which no organic basis could be found.

Scale 4—(Pd: Psychopathic Deviancy): This patient group had continual run-ins with legal authorities, yet they had normal intelligence and no serious neurotic or psychotic symptomatology. No major criminal types were included. These patients exhibited delinquent acts such as stealing, lying, excessive drinking, and sexual promiscuity.

Scale 5—(Mf: Masculinity–Femininity): This scale originally was developed to discriminate homosexual invert males. However, there were only a small number of items that statistically differentiated this group from the controls. As a result other items were included that differentiated men from women in the normative sample. An attempt was made unsuccessfully to identify sexual inversion in females.

Scale 6—(Pa: Paranoia): Patients in this group consisted of suspicious, mistrusting individuals who often had diagnoses other than paranoia. These persons could have a paranoid personality, paranoid schizophrenia, or an affective disorder. Interestingly, the scale had few false positives, but had a rather large number of false negatives.

Scale 7—(Pt: Psychasthenia): This scale was originally developed using a group of patients with psychasthenia. This diagnosis is no longer used today but does resemble a diagnosis of obsessive-compulsive disorder. These patients worry excessively; have major fears, compulsions, and rituals; and possess generalized anxiety and distress.

Scale 8—(Sc: Schizophrenia): The patient group used for this scale consisted of various types of schizophrenic patients. Efforts to develop separate scales for different subtypes of schizophrenia were not successful.

Scale 9—(Ma: Mania): Because patients who have full-blown mania are not testable, this scale was developed utilizing patients who were in the earlier stages of manic episodes.

Scale 0—(Si: Social Introversion): This scale was not developed on a particular diagnostic group. Rather the items were determined by contrasting selected females who scored in the high or low range on another measure of social introversion. Social Introversion–Extraversion scale of the Minnesota TSE (Thinking–Social–Emotional) Inventory.

Adapted from Graham J. *MMPI–2 Assessing Personality and Psychopathology*. New York: Oxford

University Press; 1990.

COMBINING SCALE ELEVATIONS IN THE CODE TYPES. As has been illustrated previously, interpreting single-scale evaluations on the MMPI-2 is problematic; thus, most clinicians use a combination of scale elevations in interpretation. This procedure usually involves using the patient's highest two clinical scales that exceed a *t* score of 65 to provide a configured interpretation that takes into account the relation between the scales. In this approach, for example, a 1-2 or 2-1 code type would be the same. In other words, a patient who scored highest on scale 1 (hypochondriasis [Hs]) and second highest on scale 2 (depression [D]) would be considered to have the same basic profile as an individual who scored highest on scale 2 (D) and second highest on scale 1 (Hs).

A three-point code, likewise, would list the three highest scales, all of which normally would be above a *t* score of 65. These scales also would usually be significantly higher than all of the other clinical scales.

On the MMPI-2, uniform *t* scores are used on clinical scales except in scales 5 (masculinity/femininity) and 0 (social introversion). *Uniform t scores,* briefly stated, are standardized scores in which the percentile equivalent and distribution are matched among the various clinical scales in terms of skewness and kurtosis. Thus, a uniform *t* score of 65 on scale 7 (psychasthenia) is roughly equivalent to a uniform *t* score of 65 on scale 9 (hypomania). The use of uniform *t* score rather than linear *t* scores was one of the improvements of the MMPI-2 over the original MMPI.

A review of the interpretation of common two- and three-code types can be found in most standard textbooks on the MMPI, some of which are included in the references. Certain MMPI authorities have used slightly more complex means of interpreting the MMPI-2. This approach emphasizes specific relationships about multiple scales. For example, Harold Gilderstadt and Jan Duker described criteria in their book *A Handbook for Clinical and Actuarial MMPI Interpretation* for what is referred to as a *1,2,3,4-type scale,* as follows:

1. Hs, D, Hy, and Pd greater than *t* score 70 and higher than all other scales.
2. Si less than *t* score of 70.
3. L, F, and K less than *t* score 70 unless two or more scales are greater than 100, in which case F less than *t* score 80.

The problem in using this complex method is that, in many settings, only a relatively small percentage of patients match these criteria. In other words, if one uses this method rigidly, then a substantial number of patients seen in an inpatient unit would not precisely match one of these complex criteria.

CONTENT SCALES FOR MMPI-2. Several efforts have been made to group items according to certain specified dimensions. Using a combination of rational and statistical procedures, Butcher and colleagues developed these 15 narrow-band scales. The 15 scales are internally consistent and relatively independent. Item overlap between scales is minimal. Scales include such content areas as anxiety, anger, cynicism, low self-esteem, fears, obsessiveness, and depression among many others. Norms for both males and females were simply derived using the same uniform *t* scores used in the MMPI-2 clinical scale formation. Reliability data on the scales are acceptable given the relatively few items in some of the content scales.

John R. Graham presents scoring guidelines in his book *MMPI-2 Assessing Personality and Psychopathology*. Validity data for the new scales are limited at the current time. Because efforts were made to develop reasonably homogeneous items, interpretation of these scales consists of examining high scores and low scores. For example, patients with high scores on the low self-esteem scale tend to have poor self-concept, anticipate failure, be overly sensitive to criticism, and have trouble making decisions. Individuals with low scores, on the other hand, are self-confident and decisive. A given content scale such as the anger scale, when considered alone, is very similar to the rapid assessment instruments discussed in a later section of this chapter.

HARRIS & LINGOES CONTENT SCALES. As previously stated, most of the MMPI clinical scales were developed using empirical key approaches. For example, an item was included in the depression scale because that item discriminated patients in the reference group from persons in the normal group. Although this method is scientifically sound, the approach results in the inclusion of subtle items that seemingly may have little to do with depression. For example, items such as, "I sometimes tease animals," (false) and, "It takes a lot of argument to convince most people of the truth," (false) are included in the depression scale (D) even though, on the surface, they have little to do with depression. Such items are called *subtle items*. Obvious items are those that have strong face validity, such as, "I wish I could be as happy as others," (true) and, "I brood a great deal" (true). The depression subscale reflects a heterogeneous item group coming from a variety of aspects of depression, such as subjective depression, psychomotor retardation, mental dullness, etc. The same total score might reflect a considerably different combination of items. Knowing which items in combination actually contribute to the total score can be helpful in interpretation. The Harris & Lingoes scale can help in that process.

The Harris & Lingoes scales were developed logically by grouping items in a given clinical scale into standard scales that were judged to reflect a more homogeneous trait. Thus, the Harris & Lingoes scales for depression reflect separate scales for subjective depression (D1), psychomotor

retardation (D2), physical malfunction (D3), mental dullness (D4), and brooding (D5).

The Harris & Lingoes scales can be used for two major purposes: (1) to see which group of items was elevated for a patient scoring high on a given clinical scale and (2) to see why a marginally elevated scale was elevated. For example, one patient who scored marginally high on the depression scale may be demonstrating psychomotor retardation and no subjective depression while another may be just the opposite.

SUPPLEMENTAL SCALES. Given the MMPI contained 566 items, numerous investigators have used the items to develop special purpose or supplemental scales. As early as 1972, more than 450 supplemental scales were developed, and the number has kept growing. Scales have been developed to predict success in psychotherapy (ego strength [Es]), to differentiate alcoholic patients from nonalcoholic patients (MacAndrew Revised [MACR]), and to differentiate assaultive prisoners from others not committing a crime (overcontrolled hostility [Oh]). Many dissertations for doctoral research have been conducted to uncover a new supplemental scale. The success of supplemental scales in accomplishing their stated purpose varies considerably. Many such scales were utilized for only one study, and others, such as the MacAndrew Alcoholism (MAC) scale, have been used in many studies over the years.

COMPUTER SCORING AND INTERPRETATION. A computer program can score the MMPI-2. A typical computer scoring program not only scores individual items into validity, clinical, content, and supplemental scales but also prints out certain positive critical items marked by the patient. These programs save considerable time because the program scores all of the material that may take hours to score by hand. Few clinicians have the time or patience to routinely score the vast number of scales found in the output of a typical computer program. The MMPI-2 computer programs can also be used to interpret the test clinically, or at least give clinical hypotheses concerning the patient. These computer-generated reports are in narrative format. Because the report comes from a computer, one may get the impression that the information is more valid or accurate than a clinical interpretation. Although attempts have been made to "validate" narrative statements made by the computer, efforts have only begun in this regard. One problem with computer interpretation of narrative reports is that the algorithms underlying the narrative reports are proprietary information owned by the company producing the program and are not available to the general user.

MMPI-2-RF (Restructured Form). Yossef S. Ben-Porath and Auke Tellegen have developed a version of the MMPI-2 called the MMPI-2-RF (Restructured Form). Unlike the MMPI-2, which contains 567 items, the

RF version contains 338 items; consequently it takes less time to administer. The MMPI-2-RF has 51 revised empirically validated scales. The test is designed to take 35 to 50 minutes for most patients to complete. The MMPI-2-RF is not a replacement of the well-validated MMPI-2 test but instead provides an alternative instrument. The MMPI-2-RF has the Restructured Clinical (RC) scales at its core. The 51 scales comprising the MMPI-2-RF include eight different validity scales including a new FBS-r symptom validity scale and three higher-order scales, such as the Thought Dysfunction Scale (THD).

As previously mentioned, the nine RC scales are at its core. The RC scales include an Antisocial Behavior Scale (RC-4) as well as an Ideas of Persecution Scale (RC-6).

In addition to the aforementioned scales, there are three different somatic scales, ten internalizing scales, and four externalizing scales. In order to assess interpersonal issues, five different scales are provided including the family problems scale. There are also two interest scales and five personality psychopathology scales.

Most mental health professionals use the computer for scoring because hand scoring is much too time-consuming and subject to some scorer errors. If the 567 items in the MMPI-2 are administered to a given patient, then a special software program converts the MMPI-2 scores to an MMPI-2-RF record.

The Personality Inventory for DSM-5—Adult (PID-5). As stated in DSM-5, the diagnosis of personality disorders is a difficult and confusing endeavor. The problems include the fact that a given patient frequently meets criteria for more than one personality disorder. Moreover, the patient may not formally meet the criteria for any single personality disorder, but instead receive a diagnosis of "other specified" or "unspecified personality disorder." Although this may be the correct diagnosis, it is mostly uninformative in terms of treatment or prognosis.

An alternative DSM-5 model for personality disorders is presented in DSM-5. This alternative model is found in Section III of the DSM-5 under "Emerging Measures and Models." The current approach to personality disorders still appears in Section II of DSM-5. A complete discussion of this alternative DSM-5 model is beyond the scope of this chapter on personality assessment.

Briefly, the Section III Multidimensional Personality Trait model involves assessing a given patient on 5 broad domains of personality and 25 traits. This approach uses a formal psychometric rating instrument entitled "Personality Inventory for DSM-5" (PID-5) which is a self-report inventory completed by the patient and/or another similar form completed by a significant other person in the person's life (PID-5-IRF).

This PID-5 consists of 220 questions. The Personality Disorders

Workgroup of the DSM-5 developed it in 2012. It consists of a Likert scale with the following rating system: 0 (very false), 1 (somewhat false), 2 (somewhat true), and 3 (very true). The scoring of this instrument provides a raw score on five domains (negative affect, detachment, antagonism, disinhibition, and psychoticism). In addition, the PID-5 provides information on 1 of 25 personality trait facets including: anhedonia, anxiousness, attention seeking, callousness, deceitfulness, depressivity, distractibility, eccentricity, emotional lability, grandiosity, hostility, impulsivity, intimacy avoidance, irresponsibility, manipulativeness, perceptual dysregulation, perseveration, restricted affectivity, rigid perfectionism, risk taking, separation insecurity, submissiveness, suspiciousness, unusual beliefs and experiences, and withdrawal. An average raw score is calculated for each of the 5 domains and each of the 25 personality trait facets. Higher average raw scores on a particular trait or domain indicate greater dysfunction.

The major advantage of the PID-5 is that it is directly tied to the DSM-5 model for personality disorders. Secondly, this instrument it is a very easy scale for clinicians to administer. Thirdly, it holds much promise for future research investigation.

The available research of the PID-5 is limited at the current time. In addition, the rating reliability of a practicing clinician is not known at the current time. Moreover, the interrater reliability of practicing clinicians using this measure has not yet been established.

The Millon Clinical Multiaxial Inventory IV (MCMI-IV). Published in 2015, the Millon Clinical Multiaxial Inventory IV is an objective psychological test consisting of 195 true–false questions designed to measure psychopathology. The MCMI-IV was updated in 2015 to include a new normative sample of 1547 individuals with a diversity of mental health conditions. The MCMI-IV should only be used with adults over the age of 18 with suspected mental health conditions. This instrument was standardized with psychiatric patients with mental health problems, so should not be used with nonclinical populations. The MCMI-IV is particularly effective with describing personality disorders. The MCMI-IV has 15 Personality Pattern Scales and five Validity Scales, which include three modifying indices and two random response indicators. The MCMI-IV also includes 45 Grossman Facet Scales. This instrument therefore has a large number of scales with a limited number of items (195). This high ratio of number of scales to number items does present some methodological challenges concerning scale reliability.

The advantages of this instrument are that it is quite easy to administer and takes only approximately 30 to 45 minutes to complete the test. A computerized interpretation and scoring service is available to the clinician. This computerized narrative can be helpful to the clinician;

however, the author cautions that the computerized interpretation should be considered a "professional to professional consultation" and naturally should be integrated with all other sources of information concerning a given patient. One complaint concerning this test is that it may over-diagnose psychopathology through many scale elevations.

Personality Inventory-Revised (NEO-PI-R). Developed by Costa and McCrae, the NEO-PI-R is a personality assessment that is based upon the Big Five Personality domains (openness to experience, conscientiousness, extraversion, agreeableness, neuroticism). These five personality domains are also known as the Five Factor Model (FFM). This instrument contains 240 items and is designed to take approximately 30 to 45 minutes to complete. The patient must be able to read on the sixth-grade level. The NEO-PI-R was published in 1990, and the NEO-PI-3 is the most up to date version of the NEO-PI-R.

There is a brief version of the NEO-PI-R that was revised in 2004 called the NEO-FFI. The NEO-FFI published the most recent update in 2010. The NEO-FFI only has 60 items (consisting of 12 items in each domain) and it takes 10 to 15 minutes to complete.

Personality Assessment Inventory. Another increasingly popular objective personality test is the Personality Assessment Inventory (PAI). This test consists of 344 items that are written at a fourth-grade reading level. This reading level ensures that most patients can complete it without experiencing any reading problems. The PAI takes about 45 to 50 minutes to complete for most patients. The PAI was normed on 1,000 community-dwelling individuals stratified according to gender, race, and age. There are not separate norms for male and female as there are in the MMPI. In addition, data were gathered on 1,246 clinical subjects and 1,051 college students in the normative process. The clinical subjects were drawn from a variety of different clinical settings, including inpatient psychiatric facilities (25 percent), outpatient psychiatric facilities (35 percent), correctional institutions (12 percent), medical settings (2 percent), and substance abuse treatment programs (15 percent).

Unlike the MMPI-2 where the patient is asked simply to answer whether a given item is true or false for them, the PAI allows for four alternative responses from the patient. The patient is asked if a given item for them is totally false, slightly true, mainly true, or very true.

Similarly to the MMPI-2 the PAI has four validity scales that measure the degree to which the patient is attempting to present him- or herself in an overly positive or overly negative fashion. In other words, is the patient trying to feign psychopathology (fake bad) or underreport symptoms (fake good)? If the patient is attempting to fake bad or fake good, then the resulting PAI profile would of course likely be invalid and could not be

interpreted.

The PAI has 11 clinical scales. These main clinical scales are similar to the MMPI-2 clinical scales and measure such personality issues as somatic concerns, depression, paranoia, borderline features, or alcohol or drug problems. The PAI also has five treatment-related scales that are designed to address such issues as treatment rejection, suicide ideation, or aggression. Two PAI interpersonal scales measure dominance and warmth.

Several strengths of the PAI are that the scales are nonoverlapping, the reading level is low, and four response choices to individual items are given. The MMPI-2 and PAI are not to be used as diagnostic instruments by themselves, but diagnosis determination needs to be based on a clinical interview and historical and medical information as well.

Rapid Personality Assessment Instruments. Rapid personality assessment instruments are narrow-spectrum self-report objective instruments that measure some unidimensional aspect of personality. The instrument is written in clear, simple language and provides face validity. Typically, these short instruments can be completed in less than 15 minutes. The scoring is simple and straightforward, usually just adding up the number of positive responses. On many instruments, there are cutting scores, above which an individual is considered to have a clinically significant problem in that area.

There are literally hundreds of such measures capable of assessing a wide variety of personality states and traits, such as geriatric depression, obsessive-compulsiveness, impulsivity, and state-trait anger. Table 7.5–3 presents common examples of such measures. The overall score typically provides an index of the magnitude of the patient's problems. Most of these tests can be used repeatedly to assess change over time. A number of instruments were developed using relatively small samples and have limited scientific studies documenting their usefulness in a variety of situations. There are, however, clear exceptions as a few instruments are widely used.

Some rapid assessment instruments were developed out of the behavioral assessment tradition. One advantage of short personality assessment procedures is high face validity for the domain of the behavior in question. Most instruments have standardized instructions and small normative groups. Performances of patients with the psychiatric problem in question (anger dyscontrol or geriatric depression) are frequently reported in terms of mean performance and standard deviation. One can find a rapid assessment instrument to assess many of the common psychiatric problems seen today. Another advantage of these instruments is the ease of administration and scoring. The major disadvantage is that these instruments differ greatly in their reliability and validity. Normative groups are frequently small and are obtained in one location. Most do not

have national examples. Many of these short instruments have few, if any, studies of their construct validity. Most instruments do not systematically assess response styles such as faking bad, faking good, or social desirability responding. Examples of the hundreds of rapid assessments include measures of bulimia, obsessive-compulsive behavior, sexual functioning, fear and panic, and personal beliefs.

Projective Personality Test

Projective personality tests, in contrast to objective personality instruments, are more indirect and unstructured. Unlike objective tests in which the patient may simply mark "true" or "false" to given questions, the variety of responses to projective personality tests are almost unlimited. Instructions are usually very general in nature, allowing the patient's fantasies to be expressed. The patient generally does not know how his or her responses will be scored or analyzed. Consequently, trying to feign the test becomes difficult. Projective tests typically do not usually measure one particular personality characteristic such as "type A personality" (e.g., narrow-band measurement) but instead are designed to assess one's personality as a whole (e.g., broad-band measurement).

Projective tests often focus on "latent" or unconscious aspects of personality. Obviously, psychologists and others differ in the degree to which they rely on "unconscious" information. In many projective techniques, the patient is simply shown a picture of something and asked to tell what the picture reminds them of. An underlying assumption of projective techniques (projective hypothesis) is that, when presented with an ambiguous stimulus, such as an inkblot, for which there are an almost unlimited number of responses, the patient's responses will reflect fundamental aspects of his or her personality. The ambiguous stimulus is a sort of screen on which the individual projects his or her own needs, thoughts, or conflicts. Different persons have different thoughts, needs, and conflicts and, hence, have widely different responses. The responses of a schizophrenic patient often reflect a rather bizarre, idiosyncratic view of the world.

There are literally hundreds of different projective techniques developed—most of which are not widely used today. Table 7.5–4 contains a list of common projective tests together with a description of the tests. The strengths and weaknesses of the projective tests are also presented. This section will focus on more widely used projective instruments such as the Rorschach test, Thematic Apperception Test (TAT), and sentence completion.

The literature on projective techniques is vast, and the history of projective techniques goes back for many years. Arthur Carr pointed out in the fourth edition of the *Comprehensive Textbook of Psychiatry* that Shakespeare had an appreciation of the importance of one's association to

ambiguous stimuli. In an effort to feign insanity Hamlet associated to a cloud formation certain images that presumably reflected his personality.

Rorschach Test. Herman Rorschach, a Swiss psychiatrist, developed the first major use of projective techniques around 1910. The Rorschach test is the most frequently used projective personality instrument (Fig. 7.5–1). The test consists of 10 ambiguous symmetrical inkblots. The inkblot card appears as if a blot of ink was poured onto a piece of paper and folded over—hence, the symmetrical appearance. These 6½- × 9½-in inkblot cards are the standard stimuli and, by convention, are referred to as Roman numerals I to X.

Table 7.5–3.
Rapid Personality Instruments

Name	Description	Weaknesses	Strengths
Geriatric Depression Scale	30 Items covering depression in the elderly; given orally or written; does not emphasize the vegative signs of depression	Geriatric patients frequently unwilling to admit depression because of denial; normative data limited	Has reasonably good internal consistency reliability; good concurrent validity with other measures; can distinguish between normal, mildly depressed, and severely depressed patient
Beck Depression Inventory II	21 Self-administered items for assessing depression in persons having at least a sixth-grade education; originally designed for use with trained interviewers, the test is self-administered; can assess the severity of depression; sensitive to changes due to psychotherapy and medication; measures patient change weekly during therapy; assessed depression doing specified period of 1 week; alpha coefficients were reported to be in the high 0.80s; effective in discriminating mild, moderate, and severe depression	No validity scales to assess approach to testing; is not specific to depression; patients with other disorders may score high on the test	Large number of studies available in the literature establishes both its test–retest reliability and validity; widely used
State Trait Anxiety Inventory	Anxiety is conceptualized as subjective tension and worry along with heightened arousal of the autonomic nervous system; state anxiety is the amount of anxiety a person has at a given moment in time; trait anxiety refers to the degree of anxiety a person usually experiences; there	No validity scales; patients who score high on anxiety also frequently score high on depression and other psychological	Very large normative sample of more than 5,000 in one study; inventory translated in many languages; excellent internal consistence reliability and concurrent validity

	are two separate but similarly worded instruments, one measuring state anxiety (S Anxiety) the other measuring trait anxiety (T Anxiety); both consist of 20 questions	measure instruments	
Michigan Alcoholism Screening Test (MAST-R)	Designed to screen for alcoholism and problem drinking presumable in people not manifestingly suffering from these problems; the revised version has 25 items that can be completed in 10 minutes	Extensive norms of population not available; may overdiagnose alcoholism in some cases and produced high false-positive rates particularly in college students; lack of candor on part of respondent can impact usefulness; high face validity makes it easy to fabricate responses	Most frequently used screening instrument for alcoholism; high internal consistency and good test–retest reliability; discriminates controls from other alcoholic problem groups (hospitalized alcoholics, driving under the influence offenders, persons convicted of drinking and disorderly conduct)
McGill Pain Questionnaire (MPQ)	The purpose of the MPQ is to assess quantitatively the qualitative characteristics of pain; instrument consists of 78 adjectives that the patient marks concerning the sensory, affective, and evaluative nature of pain along with other characteristics; a pain rating index and pain intensity score is calculated; There is a MPQ-2 short version available too (SF-MPQ-2)	Adjectives can be hard for individuals with low educational background and/or low IQ to understand; pain complaints can be easily malingered	Used with patients suffering from all types of pain; can be used to assess effectiveness of various pain treatments (medication, biofeedback, relaxation exercise, surgery); also has been used to discriminate between patients who have pain associated with objective physical findings and individuals who have pain without such findings

Table 7.5–4.
Projective Measures of Personality

Name	Description	Strengths	Weaknesses
Rorschach test	Ten stimulus cards of	Most widely used	Some Rorschach

	inkblots, some colored, others achromatic	projective device and certainly the best researched; considerable interpretative data available	interpretive systems have unproven validity
Thematic Apperception Test (TAT)	20 Stimulus cards depicting a number of scenes of varying ambiguity	A widely used method that, in the hands of a well-trained person, provides valuable information	No generally accepted scoring system results in poor consistency in interpretation; time-consuming administration
Sentence completion test	A number of different devices available, all sharing the same format with more similarities than differences	Brief administration time; can be a useful adjunct to clinical interviews if supplied beforehand	Stimuli are obvious as to intent and subject to easy falsification
Holtzman Inkblot Technique (HIT)	Two parallel forms of inkblot cards with 45 cards per form	Only one response is allowed per card, making research less troublesome	Not widely accepted and rarely used; not directly comparable to Rorschach interpretive strategies
Figure drawing	Typically human forms but can involve houses or other forms	Quick administration	Interpretive strategies have typically been unsupported by research
Make-a-Picture Story (MAPS)	Similar to TAT; however, stimuli can be manipulated by the patient	Provides idiographic personality information through thematic analysis	Minimal research support; not widely used

Courtesy of Robert W. Butler, Ph.D. and Paul Satz, Ph.D.

Using John E. Exner's Comprehensive System, the examiner, in essence, gives the patient the following instructions: "You'll be given a series of inkblots. Look at each inkblot and tell me what you see." The examiner gives the patient each card and asks, "What might this be?"

Minimal interaction between the examiner and the patient occurs while the Rorschach is administered, which ensures standardization procedures are upheld. The examiner writes down verbatim what the patient says during the above-described "free association" or "response proper" phase. If the patient rotates the card during his or her response, then the examiner makes the appropriate notation on the test protocol. After the patient has given responses to all ten cards, an inquiry phase of administration begins. The examiner asks the patient to go through the cards again and help the examiner see the responses he or she gave. The examiner reads the patient's initial response and asks the patient to point out what he or she saw and explain what made it look like that to him or her. An almost unlimited range of responses is possible with the Rorschach test and most projective tests. Six different scoring and interpretation systems for the Rorschach test are available for the clinician; however, Exner's

Comprehensive System is the most widely used and accepted. A structural summary of the various measures derived from the patient's response is usually computed.

FIGURE 7.5–1. Card I of the Rorschach Test. (From Hermann Rorschach, Rorschach® Test. Verlag Hans Hubar AG, Bern, Switzerland, 1921, 1948, 1994. By kind permission of the publisher Hans Huber.)

Scoring and interpretation of the Rorschach test can be a time-consuming process. Administering and scoring the Rorschach test usually takes about 1½ to 2 hours of time. Typically, examiners score the Rorschach responses according to the *location* on the blot where the patient sees the response. The factors that make patients respond the way that they did are referred to as the *determinants* (shape of blot, color, and shading). The examiner also scores the actual *content* of the patients' response (e.g., animals or human anatomy). The scoring system is elaborate and can result in dozens of different indices being computed. Although some investigators question the reliability of certain measures, many studies do support the usefulness of the Rorschach test as a personality assessment instrument in the hands of a qualified and experienced psychologist.

The limited scientific support of the Rorschach test, however, has resulted in critical reviews by some psychologists. A number of graduate schools of clinical psychology no longer offer formal coursework in the Rorschach test, although many still do.

Table 7.5–5 gives the sample Rorschach responses of a number of patients to Rorschach card I. Notice how different the responses are by various patients to this same card and how the responses reflect the

patient's underlying personality.

Table 7.5–6 shows certain cluster variables derived from Exner's Comprehensive Scoring System. The interpretive utility of these cluster variables is also described in the table.

Table 7.5–5.
Sample Rorschach Responses and Interpretation to Card 1 by Diagnostic Category

Diagnostic Category	Response Proper	Inquiry	Interpretation
Nonpatient	Well the first impression is of a wolf's head, ears here and eyes.	Overall the head, nose, eyes, ears, the whole thing.	Common response utilizing the white space as the eyes, responding to the whole blot.
"Schizo" patient	I will say the first thing that comes to mind, it would be a predator on the movie, lots of ink, predator, have you seen that movie?	The whole picture all the way, and I liked the movie and the reason why is because you said the first thing that comes to mind. I've seen the movie over and over again, ten times. The whole picture, something that looks at me mean (Looks at you mean?) Because it's a person that kills without any feeling, that's why I thought the movie was good.	The response shows difficulty staying focused on the task at hand, tangential thought process, some paranoia and distortion of the form of the card.
Depression	Could be a leaf.	Could be a leaf that has fallen off the tree and started to decay, leafs have jagged edges (You said decayed?) Because if it was a perfect leaf all this would be part of the blot, and here parts of it that have fallen off. That would have been the stem. Also you have the main vein running down the middle, this is more prevalent because it's biggest, if it was on a tree it would get nutrients to the leaf like our blood system.	The response shows morbid content to (decay) that indicates a poor view of self and the world. Leaf responses also weigh on the isolation index.
Anxiety	I don't know, I'd have to think about it. It could be trees.	Like if you are in the forest, stalks you know, when trees bland together they are nondescript. I'm a real outdoor person, I like to go to the forest and spend time there. (Can you help me see the trees?) Just this little Christmas tree here and here, the outline.	The response shows discomfort in committing to a response, and discomfort in beginning the task. Some anxiety is also seen in the rambling personalization used to justify the response.

Courtesy of Dana Foley, Ph.D.

Thematic Apperception Test. Although the Rorschach test is clearly the most frequently used projective personality test, the TAT is probably in second place. Many clinicians will include both the TAT and the Rorschach test in a battery of tests for personality assessment. The TAT consists of a series of 10 black-and-white pictures that depict individuals of both sexes and of different age groups, who are involved in a variety of different activities. For example, on Card 1, a young boy is shown sitting at a table, looking at a violin. Card 2 depicts a farm scene in which a young woman is in the foreground carrying books in her hands. A man is working in the fields nearby, and an older woman is seen in the background. Typically, a patient is shown 10 TAT cards and is asked to make up stories about them. The patient is asked to tell what is going on in the picture, what is going on before the picture was taken, what the individuals in the picture are thinking and feeling, and what is likely to happen in the future.

Table 7.5–6.
Cluster Variables (Exner System) and Basic Interpretive Utility

Cluster Variable	Interpretive Value
Control Stress Tolerance	Can assess a person's stress tolerance, ability to control their behavior, and responses in the environment.
	Can assess a person's style of processing information.
Situation Related Stress	Can assess the amount of situational stress affecting a person's capacity for control.
Affective Features	Can assess how a person manages emotional stimuli, types of emotional features such as depression and anger, and ability to modulate their affect.
Self-Perception	Can assess how a person views themselves and ability to look at themselves.
Interpersonal Perception	Can assess how a person views others, ways they interact with others, and how well they identify with others.
Information Processing	Can assess how a person organizes information in their world and how they process it.
Cognitive Mediation	Can assess the conventionality of a person's perceptions and degree of reality distortions.
Ideation	Can assess aspects of ideation such as processing style, use of defenses such as fantasy and intellectualization, and if their thoughts proceed in logical, coherent fashion or have some thought slippage.

Courtesy of Dana Foley, Ph.D.

Henry Murray developed the TAT in 1943 at the Harvard Psychological Clinic. The stories that the patient make up concerning the pictures, according to the projective hypothesis, reflect the patient's own needs, thoughts, feelings, stresses, wishes, desires, and view of the future. According to the theory underlying the test, a patient identifies with a

particular individual in the picture. This individual is called the *hero*. The hero is usually close to the age of the patient and frequently of the same sex, although not necessarily so. Theoretically, the patient would attribute his or her own needs, thoughts, and feelings to this hero. The forces present in the hero's environment represent the *press* of the story, and the *outcome* is the resolution of the interaction between the hero's needs and desires and the press of the environment. The TAT, in the hands of a skilled interpreter, can reveal considerable information about the patient's personality that can then be combined with other information obtained from the interview and other testing.

By itself, the TAT is not an effective diagnostic instrument, nor was it designed to be. Moreover, as measured by traditional methods of assessing reliability and validity, the TAT lacks the scientific merit of objective instruments. The fact that it has remained reasonably popular today, however, speaks to the instrument's usefulness in clinical practice.

Sentence Completion Test. Although a projective instrument, the sentence completion test is much more direct in soliciting responses from the patient. He or she is simply presented with a series of incomplete sentences and is asked to complete the sentence with the first response that comes to mind. The following are examples of possible incomplete sentences:

My father seldom ...
Most people don't know that I'm afraid of ...
When I was a child, I ...
When encountering frustration, I usually...

The purpose of the test is to elicit, in a somewhat indirect manner, information about the patient that cannot be elicited from other measures. Because the patient responds in writing, the examiner's time is limited. The length of time it takes to complete the sentence completion varies greatly depending on the number of incomplete sentences. Tests can range from less than 10 sentences to greater than 75.

There are many variations of sentence completion tests. Some clinicians have developed their own tests. One form of the test, developed by Julian Rotter, has some established validity and reliability, but most sentence completion tests do not. Special purpose sentence completion tests have been developed to measure different problem areas. For example, there is a sentence completion test that is used with patients who have chronic pain, and another is used to assess issues concerning transsexual patients. The sentence completion test is seldom, if ever, used alone but is combined with other appropriate instruments.

Advantages of the sentence completion are short administration time,

ease of administration, variety of instruments, and ease of construction. Disadvantages are lack of reliability and validity studies and ease of fabrication and deception.

Behavioral Assessment

Behavioral assessment involves the direct measurement of a given behavior. Rather than focus primarily on human characteristics, such as repression, ego strength, or self-esteem (vague terms to a behaviorist), strict behavioral measurement concentrates on the direct measurement that can be observed, such as a number of temper tantrums per unit of time, duration and intensity and number of hyperventilation episodes, or the number of cigarettes smoked per 24-hour period.

Although early strict behaviorists would count only behaviors that were observable, a broader definition of behavior has emerged, under which just about anything people do—whether it is overt such as crying, swearing, or handwashing or covert such as feeling and thinking—is considered behavior. To measure covert behavior, the patient is typically the informant. For example, to measure dysfunctional thinking the patient may simply carry a mechanical counter and press it every time he or she has a particular target self-disparaging thought during a given unit of time.

Direct Behavioral Measurement. The patient, a family member, or an impartial observer can measure behavior directly. Therapists use direct behavioral measurements to establish baselines of a given undesirable behavior (i.e., violent thoughts that the patient may wish to reduce). Similarly, therapists can measure behavior that the patient wants to increase (time studying or distance walked on a treadmill). Follow-up measures of the same behavior monitor progress and quantify improvement.

Different aspects of a target behavior can be counted or measured, such as frequency of behavior (e.g., number of behavioral occurrences), duration of behavior (length of time behavior occurred), and intensity. Event recording is a behavior measurement technique in which an individual counts instances of behavior. This approach is best used for discrete behaviors that have a beginning and end. For example, one can count the number of instance of nail biting, the number of times out of bed, the number of occasions one arrives late to work, or the number of dysfunctional thoughts he or she has concerning an employer during a given 18-hour period.

Interval Recording is another type of behavioral measurement technique that can be used to measure how often a behavior occurs. Sometimes it is not feasible to count each instance of behavior as the behavior may be too frequent. In addition, certain behaviors may not be discrete (with a beginning and end) such as the duration of time an

individual stays on task on a work project.

There are three types of interval recording. In momentary time sampling, one records whether or not the behavior occurred at a particular time. For example, every 15 minutes, one records whether or not the behavior is occurring at that particular moment. Momentary time sampling is best use for very high-frequency behaviors. The second type of interval recording is called whole interval recording. In whole interval recording, one divides a time segment into many intervals (e.g., 30 intervals consisting of 15 seconds for each interval). An interval is scored only if the behavior occurred during the entire interval. Whole interval recording is used for behaviors that one wishes to increase, but it tends to underestimate the duration of the behavior. Partial interval recording is set up as the same as whole interval recording, but the interval is scored if the behavior occurs anytime in the interval. This type of interval recording is best used for behaviors to decrease, and this approach tends to overestimate the occurrence of the behavior. Whole and partial interval recording will provide one with the percentage of intervals a behavior occurs (e.g., 15 percent of intervals).

The advantage of using event recording and interval recording is that it is direct, objective, and easily understood and quantifies the degree of behavioral change. The problem with counting specific behavior is that many patients have so many different symptoms it would not be possible to quantify them all. Sometimes, the process of counting a target behavior may itself alter the behavior. It can also be used to help increase awareness of behavioral patterns.

Rating Scales

Rating scales have been developed to detect various aspects of personality. On some of these scales, the patient simply rates him- or herself on a given scale dimension such as sadness, 1 = I am not sad, or 7 = I wish I were dead. On other scales, a significant other person is asked to rate the patient, 1 = the individual seldom mentions being sad, or 7 = the individual says he or she would be better off dead. Some of the more commonly used rating scales, however, are designed to quantify information obtained by the mental health professional during a mental status interview.

Brief Psychiatric Rating Scale. One of the more commonly used rating scales is a Brief Psychiatric Rating Scale (BPRS). To use the BPRS, the psychologist or psychiatrist simply completes a mental status interview with the patient. After completing the interview, the interviewer rates the patient on a series of 18 psychiatric symptoms, such as motor retardation, blunted affect, conceptual disorganization, anxiety, and guilt. Expanded definitions of each of these terms are provided to the examiner. The interviewer rates each domain on a seven-point Likert scale from “not

present," the lowest rating, to "extremely severe," the highest rating. An experienced interviewer can complete the ratings in 2 or 3 minutes. The BPRS has been used extensively in drug outcome and other studies. The advantage of the BPRS is that the interrater reliability is reasonably high for a rating scale of this nature. A summary article of over 300 studies utilizing the BPRS found interrater reliability correlative of 0.80 or higher on the total score in the majority of studies. The first author has used this instrument to teach medical students to be observant of a patient's behavior, and relatively inexperienced medical students can establish reasonably good reliability. Other advantages of the BPRS include the ease and speed of rating and well-defined symptom description.

One disadvantage of this scale is examiner subjectivity. Another disadvantage is that the interviewer can only rate the patient on what the he or she personally observes during the interview. Examiners differ greatly in their style of conducting a mental status interview, and thus, one examiner might ask many questions on interviews in one given area (i.e., somatic concern), while nearly overlooking another area (i.e., suspiciousness). Another examiner might do just the opposite.

Integrated Psychological Assessment to Answer Specific Referral Questions

Normally in the practice of medicine, a physician evaluates a patient's chief complaints, signs, and symptoms. On the basis of this information, the physician may suspect a given medical problem. In order to further investigate, the physician may order a series of laboratory tests or other studies to further delineate the problem at hand. For example, a physician may suspect that a patient has a dysfunctional liver and consequently orders liver function tests. Test results are reported back to the physician in terms of quantitative information that helps the treating physician to better understand the nature of the patient's problems.

In contrast, in requesting a psychological assessment, it is normally better for the psychiatrist or other health professional to ask the psychologist specific referral questions rather than specify particular tests to be administered. Below are listed examples of referral questions: What is the patient's DSM-5 diagnosis?

1. Is the patient psychotic?
2. Does this patient have schizophrenia or bipolar disorder?
3. Does this patient have an attention-deficit disorder or are their presenting problems due to anxiety?
4. Does the patient have the cognitive and personality resources to assist in his own defense in a criminal case?
5. Is this patient feigning a psychiatric disorder or does he in fact have a genuine disorder?

In order to answer the referral question, the psychologist will determine

which tests are most appropriate given the nature of the referral question(s). There are literally hundreds of different psychological tests that may be utilized to answer various referral questions. After conducting a biopsychosocial assessment, a mental status evaluation, and personality testing, and perhaps an interview with a significant other person in the patient's life, a psychologist will integrate this information in a comprehensive psychological assessment report in order to answer the psychiatrist's specific referral question. In this case, the goal is not simply to report back to the psychiatrist a series of isolated scores but rather the goal is to interpret the scores in relationship to a referral question that the psychiatrist posed. By using this approach to psychiatrist is most likely to have their referral question answered. In order to answer specific referral questions, the psychologist may utilize personality test but they may also utilize cognitive, achievement, and narrowband test. In some cases it may be necessary to order specialized neuropsychological testing in order to answer the referral question. When a neuropsychological test is required, the evaluation is normally completed by a neuropsychologist who has special training and experience in brain–behavior relationships.

Psychologists will often have the patient complete a comprehensive battery consisting of different tests to answer the referral question. For example, based on specific referral question concerning a psychiatric diagnosis, a psychologist may choose a battery of tests that include the MMPI-2, MCMI-III, and the Beck Depression Inventory-2. The psychologist will write a comprehensive report that presents the findings of psychological testing and a summary and recommendations section that addresses the specific referral question.

REFERENCES

American Educational Research Association. *Standards for Educational and Psychological Testing*. Washington, DC: AERA; 2014.

Archer RP, Krishnamurthy R. *Essentials of MMPI-A Assessment*. New York: John Wiley; 2001.

Archer RP, Smith SR. *Personality Assessment*. 4th ed. New York: Routledge/Taylor & Francis Group; 2014.

*Butcher J. *MMPI-2: A Practitioner's Guide*. Washington, DC: American Psychological Association; 2005.

Butcher JM. *Clinical Personality Assessment: Practical Approaches (Oxford Textbooks in Clinical Psychology)*. 2nd ed. New York: Oxford University Press; 2002.

Fox D. *The Clinicians Guide to the Diagnosis and Treatment of Personality Disorders*. Eau Claire, WI: PESI Publishing and Media; 2013.

Friedman, Alan F. *Psychological Assessment with MMPI-2/MMPI-2-RF*. 3rd ed. New York: Rouledge; 2016.

Geisinger KF. *APA Handbook of Testing and Assessment in Psychology*. Vol. 2. Washington, DC: American Psychological Association; 2013.

*Graham JR. *MMPI-2: Assessing Personality and Psychopathology*. 5th ed. New York: Oxford University Press; 2011.

Gregory RJ. *Psychological Testing: History, Principles of Application*. 7th ed. London: Allyn E. Bacon; 2016.

*Groth-Marnat G. *Handbook of Psychological Assessment*. 6th ed. Hoboken, NJ: John Wiley & Sons; 2016.

Harwood TM, Beutler L, Groth-Marnat G. *Integrative Assessment of Adult Personality*. 3rd ed. New York: Guilford Press; 2012.

Hersen M. *Psychological Assessment and Clinical Practice: A Pragmatic Guide*. New York: Taylor & Francis; 2004.

Hopwood CJ, Thomas KM, Markon KE, Wright AGC, Krueger RF. *DSM-5 Personality Traits and DSM-IV Personality Disorders*. Washington, DC: American Psychological Association; 2012.

Megargee EI. *Using the MMPI-2 in Criminal Justice and Correctional Settings*. Minneapolis, MN: University of Minnesota Press; 2006.

Melton GB, Petrila J, Poythress NG, Slobogin RK. *Psychological Evaluation for the Courts: A Handbook for Mental Health Professionals and Lawyers*. New York: Guilford Press; 2007.

Millon T, Grossman S, Millon C. MCMI-IV: *Millon Clinical Multiaxial Inventory Manual*. 1st ed. Bloomington, MN: NSC Pearson; 2015.

Nichols DS, Kaufman AS. *Essentials of MMPI-2 Assessment*. Hoboken, NJ: John Wiley and Sons; 2011.

Osberg TM, Haseley EN, Kamas MM. The MMPI-2 clinical scales and restructured clinical (RC) scales: comparative psychometric properties and relative diagnostic efficiency in young adults. *J Pers Assess*. 2008;90:81.

Paniagua FA, Yamada AM. *Handbook of Multicultural Mental Health: Assessment and Treatment of Diverse Populations*. 2nd ed. San Diego, CA: Academic Press; 2013.

Rogers R, Sewell KW, Harrison KS, Jordan MJ. The MMPI-2 restructured clinical scales: a paradigmatic shift in scale development. *J Pers Assess*. 2006;7:139.

*Rush AJ, First MB, Blacker D. *Handbook of Psychiatric Measure*. Washington, DC: American Psychiatric Publishing; 2008.

Tellegen A, Ben-Porath YS, Sellbom M, Arbisi PA, McNulty JL. Further evidence on the validity of the MMPI-2 restructured clinical (RC) scales: addressing questions raised by Rogers et al. and Nichols. *J Pers Assess*. 2006;87:148.

Urbina S. *Essentials of Psychological Testing (Essentials of Behavioral Science)*. 2nd ed. Hoboken, NJ: John Wiley & Sons; 2014.

Weiner IB, Greene RL. *Handbook of Personality Assessment*. New York: John Wiley; 2011.

*Yossef S. *Interpreting the MMPI-2-RF*. Minneapolis, MN: University of Minnesota Press; 2012.

▲ 7.6 Medical Assessment and Laboratory Testing in Psychiatry

JENNIFER L. KRUSE, M.D., MARTHA L. JAMES, M.D., J.D., AND BARRY H. GUZE, M.D.

INTRODUCTION

The goal of this section is to outline the basic imaging studies, laboratory tests, and other tools of medical assessment applicable to psychiatric practice and the strategies for using them in an effective manner. The role of imaging studies, the electroencephalogram (EEG), various biochemical, hematological, endocrine, immunological, and toxicological assessments are discussed. A logical and systematic approach to the use of these tools by the psychiatrist is vital to achieving the goals of arriving at accurate diagnoses, identifying medical comorbidities, implementing appropriate treatment, and delivering cost-effective care. With respect to the diagnosis or management of medical disease, consultation with colleagues in other specialties is important. Medical specialists also are a useful source of information about test methodology and the proper performance and interpretation of tests. Analogous consultation with appropriate specialists is important when ordering nonlaboratory tests such as the EEG, electrocardiogram (ECG), and neuroimaging. Good clinicians recognize the limits of their expertise and the need for consultation with their nonpsychiatric colleagues.

Physicians are accustomed to ordering tests. Before any test is ordered, it is always worthwhile to review how this test will alter patient management. If the test results will not alter management, then the test is unlikely to be relevant to the patient's current care.

Despite increasing pressure to deliver cost-effective care, many clinicians still order a large number of screening laboratory and other studies. The lack of consensus regarding optimal patient workup may contribute to this tendency. Increasingly, however, clinicians are more selective when ordering tests. Requested studies are chosen because of clinical hypotheses based on a particular patient's clinical data, thereby minimizing the ordering of unnecessary studies and tests. Such an approach often uses a narrow selection of routine screening tests. Later, clinicians order nonroutine, additional tests as necessary or based on the results of screening tests or particular findings on examination. The psychiatric literature contains no consensus regarding the optimal battery of screening laboratory and other studies. Few would disagree that the best screening evaluation includes a good history and physical examination, including a neurological examination and mental status examination. Excessive testing yields unnecessary costs, as well as discomfort and inconvenience to the patient. False-positive results may lead to additional unnecessary testing. Yet, ordering an insufficient number of tests is also not without risk. Significant illness may go unrecognized and untreated.

Psychiatrists use diagnostic testing to identify possible medical illnesses that may cause or contribute to a patient's clinical presentation. Such testing occurs within the context of the history of the present illness, the physical examination, and the review of systems. Other tests are specific to psychiatry. For example, the measurement of blood levels of psychotropic

medications can be useful to determine the adequacy of dosing or to monitor compliance. In addition, many tests, such as liver function tests (LFTs), white blood cell (WBC) count, and serum glucose, are performed to monitor the consequences of treatment of psychiatric illness. Laboratory detection of substances of abuse is helpful to identify patients who abuse illicit substances. For patients known to abuse substances, additional testing may be performed to assess the consequences of a patient's substance abuse.

Many medical illnesses produce psychiatric symptoms. Presentations with delirium, anxiety, depression, and hallucinations are nonspecific. To be certain that a patient's symptoms are due to a mental illness, other medical and neurological causes must first be excluded. This assessment is not performed in a vacuum but occurs in the context of the patient's medical history and physical examination. Laboratory studies supplement the history and examination. The goal of the assessment is to identify causative factors that contribute to the patient's current symptoms. When psychiatric symptoms are found to be due to a general medical condition, initial efforts are focused on the treatment of the underlying condition. For many patients, treatment of the underlying medical condition results in a resolution of psychiatric symptoms. In addition, the diagnostic criteria for many psychiatric conditions, as listed in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), exclude symptoms attributable to a general medical condition or substance abuse.

In addition to ordering diagnostic studies, psychiatrists increasingly use laboratory testing to establish that administration of a particular treatment is safe before its initiation and to monitor the ongoing safety of specific treatments. For example, baseline laboratory studies may be obtained before the initiation of particular psychotropic medications. Testing also may be used to monitor compliance, measure therapeutic concentrations, and assess toxicity. Many psychotropic medications may have significant effects on physical health. Consequently, a psychiatrist may perform a pretreatment assessment of a patient's general health before initiating treatment. This assessment commonly includes a variety of laboratory tests.

Finally, there is a widely held hope that biological markers for major mental illnesses will be identified in the future and will eventually be of assistance in the diagnosis and selection of treatment. Unfortunately, research has not yet identified biomarkers that are ready for use in routine clinical practice. Nonetheless, the ongoing search for biological markers furthers understanding of the pathophysiology of major mental illnesses and continues to enhance clinical psychiatric practice.

PHYSICAL HEALTH MONITORING

Monitoring the physical health of psychiatric patients has two goals: To

provide appropriate care for existing illnesses and to protect the patient's current health from possible future impairment. Disease prevention should begin with a clear concept of the condition to be avoided. Ideally, in psychiatry this would focus on common conditions that could be a significant source of morbidity or mortality. It is clear that in psychiatry a small number of clinical problems underlie a significant number of impairments and premature deaths.

Many psychiatric patients fail to receive periodic health maintenance examinations. Examinations and laboratory testing at the time of admission to a psychiatric hospital may be a rare opportunity to deliver preventive care. Preventive care can include monitoring for hypertension or providing counseling about healthy eating or smoking cessation.

Lifestyle-related health problems such as obesity, diabetes, hypertension, cardiovascular disease, and sexually transmitted diseases (STDs) are frequent comorbidities in some psychiatric populations. It is helpful to discuss proper nutrition, diet, and exercise as part of the prevention and management of excessive weight gain. The physician also should discuss information about preventable diseases that are more likely to occur in obese individuals and the role of monitoring fasting blood sugars and fasting lipid levels in disease prevention. Sexually transmitted infections are more common in young psychiatric patients than in the general population. In patients with risk factors for human immunodeficiency virus (HIV) infection, it is important to order screening tests for the virus.

As patients age, chronic illnesses such as cardiovascular disease become more prevalent. Patients with risk factors, such as obesity, should be screened for hypertension and dyslipidemia. All patients who use tobacco should be counseled about smoking cessation with the goal of preventing coronary artery disease, stroke, and other medical comorbidities.

DIAGNOSTIC ASSESSMENT

This section reviews studies that may be useful in the diagnostic assessment of patients presenting with psychiatric complaints. Clinicians should perform an examination that is individualized to the specific needs of the patient. There is no consensus about standard laboratory tests for psychiatric patients. Laboratory studies are usually ordered based on a clinical hypothesis about the patient's primary diagnoses. However, newly admitted psychiatric inpatients commonly receive assessments of hepatic function (alanine aminotransferase [ALT] and aspartate aminotransferase [AST]) to assess their ability to metabolize prescribed medication, assessments of renal function (blood urea nitrogen [BUN] and creatinine [Cr]), and a urinalysis to evaluate their capacity to excrete metabolized medication. Thyroid-stimulating hormone (TSH) often is ordered for patients with mood disorders. In certain psychiatric populations (e.g.,

patients with drug abuse, some personality disorders, and some mood disorders), screening for STDs is common. Test batteries such as fasting blood sugar and fasting lipid panel are ordered for patients who show evidence of metabolic syndrome or who are at risk for the syndrome due to medications prescribed for their psychiatric condition.

Role of History and Physical Examination

A thorough history, including a review of systems, is the basis for a comprehensive patient assessment. The history guides the clinician in the selection of laboratory studies that are relevant for a specific patient. Many psychiatric patients, owing to their illnesses, are not capable of providing sufficiently detailed information. Collateral sources of information, including family members and prior clinicians and their medical records, may be particularly helpful in the assessment of such patients.

The history of present illness includes the chief complaint (or the reason the patient seeks assistance), the nature of the current problems, and their current treatments. The past psychiatric history should include information about previous psychiatric problems and a history of prior treatments. It should contain the medications used to treat the patient, the doses prescribed, and the durations of each medication trial. The past psychiatric history should also contain documentation of dates and locations of past psychiatric admissions, the reason for admission, and the treatments provided. If the patient reports an allergy to a specific medication, then the nature of the allergic response should be noted.

The patient's medical history is an important component of the history. It should include notation of prior injuries and, in particular, head injuries that resulted in loss of consciousness and other causes of unconsciousness. The patient's medical history also should note pain conditions, ongoing medical problems, prior hospitalizations, prior surgeries, and a list of the patient's current medications. Toxic exposures are another important component of the medical history. Such exposures are often workplace related.

The social history contains many of the details relevant to the assessment of character pathology, including risk factors for personality disorders, as well as information relevant to the assessment of Axis I disorders. Commonly, the social history includes a legal history, information about family and other significant relationships, and an educational and occupational history. The clinician should take a thorough history of alcohol and substance use. History of the use of substances, illicit and prescribed, should contain information regarding the timing of the initiation of use, the frequency of use, the quantity used, and the date of most recent use for each substance. A thorough review of symptoms should be performed. Unrecognized medical conditions are often detected during the review of systems.

Although many psychiatrists perform physical examinations on their patients, it is not uncommon in some community and institutional settings for other physicians to perform the physical examination after the psychiatrist obtains the history. Under certain clinical scenarios, it is appropriate to postpone the physical examination. Examples include the agitated or aggressive patient. Examinations should be postponed only after consideration of the risks and benefits to the patient. Such deferments should be for as brief a time as possible. As a general rule, it is wise to have a chaperone present during the physical examination.

Physical examinations may provide data helpful in the diagnosis and management of substance abuse. For example, intravenous (IV) drug use may be identified by the presence of "track" marks—small scars following the outline of a vein produced from repeated injections. Perforation of the nasal septum may occur following chronic insufflation of cocaine. Opioid intoxication is associated with sedation, dysarthria, slowed cognition, and pinpoint pupils. In contrast, withdrawal from opiates is associated with lacrimation, rhinorrhea, yawning, piloerection, myalgias, pupillary dilation (mydriasis), and bowel discomfort. These symptoms occur 1 to 2 days following discontinuation of methadone and 6 to 12 hours after stopping shorter-acting opiates. Alcohol or sedative hypnotic withdrawal is associated with autonomic hyperactivity, nausea, tremulousness, diaphoresis, headache, anxiety, and agitation as well as perceptual disturbances and clouding of sensorium in severe cases.

In evaluating patients with cognitive impairment, the physical examination may also provide etiological clues when accompanying neurological signs are identified, such as cogwheel rigidity and tremor associated with Parkinson disease or neurological deficits suggestive of prior strokes. Standard laboratory studies commonly assessed in patients with neurocognitive disorders include a complete blood count (CBC), serum electrolytes, LFTs, BUN, Cr, thyroid function tests, serum B_{12} and folate levels, Venereal Disease Research Laboratory (VDRL) test, and a urinalysis. Currently there is no clear clinical indication for routine testing for the apolipoprotein E epsilon 4 allele. Structural neuroimaging studies are often performed, particularly if there are focal neurological findings. EEG and/or CSF evaluation may be performed in some cases. When patients are delirious the neurological examination may be complicated by inattention and/or altered level of consciousness. Delirium workup may include a similar laboratory workup, but the appropriate evaluation for a particular delirious patient depends upon historical and clinical features. Urine or blood cultures, chest radiograph, arterial blood gases, chemistry panel, endocrine evaluation, neuroimaging studies, or EEG may be appropriate, among other studies.

As is true for most presenting symptoms, the evaluation of neuropsychiatric symptoms requires consideration of a broad differential

diagnosis; the differential diagnosis for a neuropsychiatric presentation may include primary psychiatric disorders, neurological disorders, or other medical conditions. An important part of the medical and laboratory assessment of psychiatric symptoms is to determine which medical conditions to include in the differential diagnosis, based upon the patient's history and examination findings. Medical conditions falling into several broad categories should be considered. Further, comorbid medical conditions may complicate the management of psychiatric patients, and laboratory evaluation for a variety of conditions may be important in the course of clinical care, often in collaboration with other specialists.

Serum Electrolyte Evaluation

Serum electrolyte levels may be useful in the initial evaluation of a psychiatric patient. Levels of serum electrolytes often are abnormal in patients with delirium. Abnormalities also may occur in response to the administration of psychotropic medications. Low serum chloride levels may occur in eating disorder patients who purge by self-induced vomiting. Serum bicarbonate levels may be elevated in patients who purge or who abuse laxatives. Bicarbonate levels are commonly low in patients who hyperventilate in response to anxiety.

Hypokalemia may be present in eating disorder patients who purge or abuse laxatives and in psychogenic vomiting. Diuretic abuse by eating disorder patients also may produce hypokalemia. Low levels of potassium are associated with weakness and fatigue. Characteristic ECG changes occur with hypokalemia and consist of cardiac arrhythmias, U waves, flattened T waves, and ST-segment depression.

Eating disorder patients with anorexia nervosa or bulimia nervosa usually receive a fairly standard set of laboratory studies, including serum electrolytes (particularly potassium and phosphorus), blood glucose, thyroid function tests, liver enzymes, total protein, serum albumin, BUN, Cr, CBC, and ECG. Serum amylase is often assessed in bulimic patients.

Magnesium levels may be low in alcohol-abusing patients. Low magnesium levels are associated with agitation, confusion, and delirium. If untreated, then convulsions and coma may follow.

Low levels of serum phosphorus may be present in eating disorder patients with purging behavior. Phosphorus levels may also be low in anxiety patients who hyperventilate. Hyperparathyroidism may produce low serum phosphorus levels. Elevated serum phosphorus levels are seen in hypoparathyroidism.

Hyponatremia is seen in psychogenic polydipsia and SIADH and in response to certain medications, such as oxcarbazepine and carbamazepine. Low sodium levels are associated with delirium.

Serum calcium abnormalities are associated with a variety of behavioral abnormalities. Low serum calcium levels are associated with depression,

delirium, and irritability. Elevated levels are associated with depression, psychosis, and weakness. Laxative abuse, common in eating disorder patients, can be associated with hypocalcemia. Hypocalcemia secondary to hypoparathyroidism may occur in patients who have undergone surgery for thyroid disease. Hypercalcemia can be seen in association with hyperparathyroidism, which is sometimes a side effect of treatment with lithium.

Renal Function

Tests of renal function, including BUN and Cr, are serum markers used to estimate glomerular filtration rate (GFR). Other relevant laboratory studies include the routine urinalysis and Cr clearance. Renal failure may result in lethargy or delirium. BUN commonly is elevated with dehydration. Cr is the serum marker most useful in clinically estimating GFR.

Routine urinalysis is a common screening test that includes examination of the general appearance of the sample and measurement of the pH, urine-specific gravity, urine bilirubin, and urine glucose. Elevations in urine glucose can be seen in diabetes. Elevated urine bilirubin occurs in biliary disease. A high urine-specific gravity may occur in diabetes mellitus, in dehydrated states, or from SIADH. A low urine-specific gravity may occur in diabetes insipidus (DI) or from excessive water ingestion.

Elevated levels of porphobilinogen are found in the urine of symptomatic patients with acute intermittent porphyria. Symptoms of this disease include psychosis, apathy, or depression, along with intermittent abdominal pain, neuropathy, and autonomic dysfunction. If urine porphobilinogen levels are elevated when the patient is symptomatic, then collection of a 24-hour urine specimen for quantitative assessment of porphobilinogen and aminolevulinic acid is indicated.

Liver Function

LFTs commonly include the serum aminotransferases, alkaline phosphatase, γ-glutamyl transpeptidase, serum bilirubin, and tests of synthetic function, usually the serum albumin concentration and prothrombin time.

Elevations in AST may occur with diseases of the liver, heart, lungs, kidneys, and skeletal muscle. In patients with alcohol-induced liver disease, AST typically is more elevated than ALT. In viral- and drug-induced liver disease, ALT is often elevated. Serum GGT is elevated in hepatobiliary disease, including alcohol-induced liver disease and cirrhosis.

Alkaline phosphatase elevations occur in many diseases, including diseases of the liver, bone, kidney, and thyroid. Levels of alkaline phosphatase may be elevated in response to some psychiatric medications, most notably the phenothiazines.

Serum ammonia levels are often elevated in patients with hepatic encephalopathy. High levels are associated with the delirium of hepatic encephalopathy. Serum ammonia levels also may be elevated in patients undergoing treatment with valproic acid.

Serum bilirubin is an index of hepatic and bile duct function. Prehepatic, unconjugated, or indirect bilirubin and posthepatic, conjugated, or direct bilirubin are often assessed to help elucidate the origin of the elevation in bilirubin.

Total protein is the measurement of serum albumin and globulins. Many psychiatric medications are protein bound. Low levels of serum total protein result in a high fraction of an ingested drug remaining in the unbound (active) state. High levels of unbound drug increase the active effects and side effects that a patient experiences. Total protein may be low in malnourished states, such as those seen in eating disorder patients. Low total protein and serum albumin in a patient with elevated liver enzymes is an indicator of significant liver disease. Consequently, these levels are often assessed in patients who abuse alcohol.

A relatively higher elevation in serum aminotransferases compared to alkaline phosphatase suggests hepatocellular injury. The opposite pattern is seen with cholestasis. Low albumen levels are seen in chronic conditions such as cirrhosis. Alcoholic liver disease is suggested when the ratio of AST to ALT is 2:1 or greater, although this may occur with steatohepatitis or hepatitis-C–related cirrhosis. When the AST to ALT ratio is 2:1 or greater and the GGT is twice the normal level, the pattern strongly suggests alcoholic liver disease.

Infectious Disease

The diagnostic assessment for patients with psychiatric symptoms may include a workup for infectious illness, depending upon the patient's history and examination findings. Further, comorbid infections requiring clinical attention and referral may be present in psychiatric patients with high-risk behaviors such as IV substance use or high-risk sexual behavior. Below, infectious disease as a differential diagnosis in psychiatric presentations is reviewed. Additionally, infectious diseases that may be a focus of clinical attention in patients with established psychiatric illness are also discussed.

A 49-year-old homeless male with a history of bipolar I disorder was brought by ambulance to the emergency department after a witnessed seizure. He had a fever of 39°C and heart rate was in the 120s. He was severely disorganized and agitated, with unintelligible speech, and appeared to be responding to internal stimuli. He was unable to cooperate with neurological examination. Magnetic resonance imaging (MRI) of the head demonstrated T2/FLAIR hyperintensity involving the temporal lobe. Cerebrospinal fluid (CSF) examination revealed an elevated WBC of 121 cells/μL (mostly lymphocytes) and mild protein elevation (57 mg/dL). Herpes simplex virus (HSV) was identified in CSF. Following antiviral treatment for

herpes encephalitis, the patient recovered to his baseline.

In this patient's case, despite known bipolar I disorder and severe psychiatric symptoms at the time of presentation, there were indicators of a comorbid medical illness, including seizure, aphasia, and fever. An evaluation for infectious etiologies was required in order to initiate the correct treatment for his condition.

Central Nervous System Infection. Infections of the central nervous system (CNS) often present with neuropsychiatric symptoms. These symptoms may range from confusion, inattention, or altered level of consciousness, to severe agitation and psychosis. CNS infection with HSV is particularly notorious for causing psychiatric disturbances. Similarly, tertiary syphilis can have a number of psychiatric and neurological manifestations. Other CNS infections may also cause psychiatric symptoms; delirium is especially common. It is unusual for CNS infections to cause neuropsychiatric symptoms in the absence of signs such as fever, nuchal rigidity, seizure, or other neurological findings. However, there are exceptions, and it is important to know the immune status of a patient. Infected patients who are immunocompromised may be less likely to mount a fever; mental status changes among the immunocompromised host should alert clinicians to a broad differential that includes a variety of infections. Various infectious agents can infect the CNS, including bacteria, viruses, prions, fungi, and parasites. The workup initiated depends upon many factors, including the patient's travel history, occupational and recreational exposures, sexual behaviors, and immune status, all of which may put the patient at risk for particular infections.

A 58-year-old man with schizoaffective disorder, treated with stable doses of haloperidol, clozapine, and lithium, presented to the emergency department with his conservator for evaluation of fever and altered mental status. On evaluation he was febrile, tachycardic, rigid, and tremulous. Neuroleptic malignant syndrome (NMS) was suspected, and psychotropic medications were discontinued. MRI was unremarkable, but CSF revealed elevated WBC of 54 cells/μL (mostly lymphocytes), elevated protein of 71 mg/dL, and positive serology for West Nile virus. With supportive care and eventual reinitiation of his psychotropic medication regimen, he eventually recovered to baseline, but not without first requiring psychiatric hospitalization for relapse of manic and psychotic symptoms.

Other Infections. In addition to CNS infection, systemic infection may cause mental status changes. For most patients presenting with delirium, infection should be included the differential diagnosis. Patients with sepsis or septic shock commonly develop delirium. For vulnerable patients such as the elderly or cognitively impaired, an infection as mild as cystitis may be sufficient to induce mental status changes. An infectious workup for delirium generally includes a CBC, urinalysis, chest x-ray, and sometimes blood cultures, among other specific evaluations depending upon the individual patient.

In addition to considering infectious disease in the differential diagnosis when appropriate clinical and historical features are present, evaluation for certain infectious diseases is sometimes indicated in psychiatric patients with known risk factors. Below, specific infectious agents that may be relevant in patients with high-risk sexual behavior or IV substance use are discussed. Some evaluations such as HIV screening may be offered in the course of routine clinical care, regardless of patient risk factors; CDC guidelines now recommend opt-out HIV screening for all adolescents and adults, and annual HIV screening for patients at high risk for HIV infection.

Intravenous Drug Use and Infection. The IV route is used for many substances of abuse. Most commonly, heroin, amphetamines, and cocaine are used alone or in combination via the IV route. Because needles may be contaminated, IV drug users are at risk for bacterial endocarditis, hepatitis B and C, and HIV infection. It has been estimated that over 60 percent of new cases of hepatitis C occur in individuals with a history of injecting illicit drugs.

The use of contaminated needles or nonsterile injection sites places IV drug users at risk for bacterial infections, including abscesses, bacteremia, and bacterial endocarditis. Findings on physical examination suggestive of endocarditis, possible bacteremia, or abscess necessitate obtaining a CBC to rule out an elevated WBC. Blood cultures should be obtained if the patient is febrile or if findings are suggestive of bacteremia or endocarditis, and internal medicine consultation should be obtained.

Sexually Transmitted Diseases. Some psychiatric illnesses, such as mania and substance abuse, are associated with a higher risk of contracting STDs. STDs include HSV types 1 and 2, Chlamydia, hepatitis B and C viruses, gonorrhea, syphilis, and HIV. Risk factors for STD include contact with sex workers, drug abuse, prior history of STDs, meeting sex partners on the Internet, multiple sex partners, a new sex partner, and being young or unmarried.

Many patients require supportive counseling before and after STD testing. Counseling consists of providing patients with educational material, allowing patients the opportunity to ask questions, and assessing their emotional response to the process. Inquiring about a patient's motivation for testing often is a helpful way to initiate discussion of other issues. Patients should be informed about the meaning of positive and negative results. Patients who test positive may require referral to a specialist who can discuss the implications of the result and future treatment. Patients with high-risk behaviors and negative test results should be cautioned about the length of time required for seroconversion with HIV testing and the need, when appropriate, for supplemental testing.

It is useful to evaluate the patient's understanding of ways to reduce the risk of contracting or transmitting STDs. Regardless of their test results, all patients can benefit from a discussion of high-risk behaviors. For patients who test positive, in addition to treatment, it is particularly important for clinicians to provide the patient with information about how to minimize the risk of STD transmission. Clinicians should emphasize the need for protected sex with fully informed partners and the avoidance of donating sperm, blood, or organs. Clinicians should warn patients about the risks of transmitting certain STDs associated with sharing razors, toothbrushes, and any other objects that could have blood on them.

Chlamydia. Chlamydia infection can cause pelvic inflammatory disease, ectopic pregnancy, chronic pelvic pain, and infertility. It is also associated with increased risk of contracting and transmitting HIV. Reinfection is common and is usually due to lack of partner treatment. Many infections are asymptomatic. Chlamydia screening uses tests based on nucleic acid amplification (NAA). These amplification tests have high sensitivity and specificity for men and women, using urine or cervical or urethral swabs.

Gonorrhea. Gonorrhea infection can cause pelvic inflammatory disease, ectopic pregnancy, chronic pelvic pain, and infertility. It is also associated with increased risk of contracting and transmitting HIV. Reinfection is common and is usually due to lack of partner treatment. Many infections are asymptomatic. The diagnosis of a gonococcal infection can be made with NAA techniques, DNA probes, Gram stain, or a culture on Thayer–Martin media. NAA tests have high sensitivity and specificity for men and women, using urine or cervical, vaginal, or urethral swabs. In women, vaginal or cervical swabs have higher sensitivity than urine samples.

HIV. HIV-infected patients may have CNS manifestations as their predominant symptoms. CNS symptoms may occur even when other symptoms are absent. Behavioral changes in an individual patient may be due to either the direct effect of the virus on the brain or associated conditions such as infection, mass effect, or malnutrition. CNS manifestations of HIV infection include dementia, personality changes, affective symptoms of either mania or depression, and psychosis.

The enzyme-linked immunosorbent assay (ELISA) test and the Western blot test are used in the assessment of a patient's HIV status. Both tests detect the presence of antibodies to HIV. Antibodies to HIV become detectable anywhere from 3 weeks to 3 months after infection. Typically, the ELISA test is used as a screening test, and the Western blot test is used to confirm the result. Other tests are performed to assess the extent of HIV

disease burden, treatment response to antiretroviral therapy, and the degree to which the virus has affected the immune system. Measurement of viral load is used to assess HIV disease burden. Measurement of CD4 cell count provides information about the need for prophylaxis against opportunistic infections.

The Centers for Disease Control and Prevention recommend HIV screening for all adolescent and adult patients. Additionally, annual HIV testing is recommended in high-risk populations. High-risk groups include men who have sex with men, individuals who exchange sex for money or substances, IV drug users, individuals who have had sex with one or more partners since their last HIV test, and individuals who have had sex with an HIV-infected person or anyone in the aforementioned high-risk groups. HIV is transmitted via blood and certain body fluids (semen, rectal fluids, vaginal fluids, breast milk) when these fluids come into direct contact with mucous membranes, damaged tissues, or when injected directly into the blood stream.

HIV infection may be associated with delirium, dementia, mania, and atypical presentations of anxiety, schizophrenia, and depression. With antiretroviral therapies, HIV-associated dementia is less common today, but milder HIV-associated neurocognitive disorders (HAND) remain prevalent, with approximately half of HIV-infected individuals demonstrating a lower level of cognitive performance than would be predicted when compared to population norms. Depression is also common in HIV patients and may occur in those with HIV-associated dementia. Mania associated with HIV or AIDS also may occur.

Syphilis. Nontreponemal tests for syphilis, such as the rapid plasma reagin (RPR) test and the VDRL test, are common screening tests for syphilis. Quantification testing may be performed after a positive RPR or VDRL test. Quantification enables levels to be followed to assess response to treatment. After a positive nontreponemal screening test, a more specific treponemal test, such as the fluorescent treponemal antibody absorption (FTA-ABS) test, is used to confirm the positive result. The FTA-ABS test detects antibody against *Treponema pallidum* spirochetes, once positive, even with adequate treatment, specific treponemal tests like the FTA-ABS may remain positive for years. False-positive results may occur; this is true for both treponemal and nontreponemal tests. Pregnant patients commonly screen positive for syphilis. False-positive nontreponemal tests are also associated with acute febrile illnesses, and sometimes in patients with chronic conditions, especially patients with systemic lupus erythematosus (SLE).

When syphilis infection of the CNS is suspected, a VDRL is performed on CSF. Patients with neurosyphilis also commonly have abnormalities in protein levels, γ-globulins, and cell counts in CSF. Due to low sensitivity of

the VDRL, the FTA-ABS test may also be used to evaluate CSF when neurosyphilis is suspected. It has higher sensitivity, but lower specificity.

Viral Hepatitis. Several types of viruses may cause viral hepatitis. Viral hepatitis produces abnormalities in LFTs including elevation of liver enzymes. Symptoms range from mild flulike manifestations to rapidly progressive and fatal liver failure. Psychiatric manifestations include depression, anxiety, weakness, and psychosis. Viral hepatitis also may impair the metabolism of psychotropic medications that are metabolized by the liver. Impaired liver metabolism requires an adjustment of the dose of medications metabolized by the liver or consideration of agents that are less affected by alterations in liver metabolism. Viruses causing hepatitis include: hepatitis A virus (HAV), hepatitis B virus (HBV), hepatitis C virus (HCV), hepatitis D virus (HDV), and hepatitis E virus (HEV). HAV and HEV are usually transmitted via the fecal-oral route. HBV and HCV are transmitted via contamination by body fluids; this may include transmission by contact with infected blood or blood products or by sexual contact. HDV is a defective RNA virus that causes hepatitis only in association with HBV infection.

The WBC is usually normal to low in patients with hepatitis, especially in the preicteric phase. Large atypical lymphocytes occasionally are present. Rarely, aplastic anemia follows an episode of acute hepatitis not caused by any of the known hepatitis viruses. Mild proteinuria is common, and bilirubinuria often precedes the appearance of jaundice. Acholic stools frequently are present during the icteric phase. Strikingly elevated AST or ALT occurs early, followed by elevations of bilirubin and alkaline phosphatase. In a minority of patients, elevations of bilirubin and alkaline phosphatase persist after aminotransferase levels have normalized. Cholestasis may be substantial in acute HAV infection. Marked prolongation of the PT in severe hepatitis correlates with increased mortality.

Chronic hepatitis, characterized by elevated aminotransferase levels for more than 6 months, develops in less than 5 percent of immunocompetent adults with acute HBV infection. Fifty to 85 percent of all persons with acute HCV infection develop chronic HCV infection. Ultimately, cirrhosis develops in as many as 50 percent of those with chronic HCV infection and 20 percent of those with chronic HBV infection; the risk of cirrhosis is even higher in patients coinfected with both viruses or with HIV. Patients with cirrhosis are at risk, with a rate of 0 to 3 percent per year, of hepatocellular carcinoma. Even in the absence of cirrhosis, patients with chronic HBV infection—particularly those with active viral replication—are at an increased risk.

HEPATITIS B VIRUS. Psychiatric manifestations have been reported during

all stages of HBV infection. HBV is transmitted via contamination by body fluids and is usually transmitted by contact with infected blood or blood products, by sexual contact, or from mother to baby at birth. The virus is present in saliva, semen, and vaginal secretions, as well as blood. The incubation period of HBV is 1 to 4 months. Transmission of HBV commonly occurs during sexual contact and via use of shared implements for IV drug use; it also may occur through tattooing and piercing, and as a result of needle-stick injuries. After acute HBV infection, HBV infection persists in 1 to 2 percent of immunocompetent adults. Persons with chronic HBV infection are at substantial risk of cirrhosis and hepatocellular carcinoma. There is a common comorbidity between HBV infection and HIV infection, given the shared risk factors for both conditions.

HBV infection can be detected by measuring viral antigens, antibodies to the virus, and viral DNA. The appearance of HBsAg is the first evidence of infection and persists throughout the clinical illness. The detection of HBsAg establishes active infection with HBV and implies infectivity. When it is present beyond 6 months, infection is considered chronic. Hepatitis B e antigen (HBeAg) is a soluble protein found only in HBsAg-positive serum. HBeAg indicates viral replication and infectivity. Antibody to HBsAg (anti-HBs) appears after clearance of HBsAg and after successful vaccination against hepatitis B. Disappearance of HBsAg and the appearance of anti-HBs signal recovery from HBV infection, noninfectivity, and immunity. Detection of the hepatitis B surface antigen (HBsAg) indicates active or chronic infection. IgM antibody to HBcAg (anti-HBc) in the setting of acute hepatitis indicates a diagnosis of acute HBV infection. It fills the serological gap in patients who have cleared HBsAg but do not yet have detectable anti-HBs. IgG anti-HBc also appears during acute HBV infection but persists indefinitely, whether the patient recovers or develops chronic hepatitis.

HEPATITIS C VIRUS. HCV infection is often diagnosed in individuals who abuse IV drugs, as these individuals are at high risk for contracting HCV infection via use of contaminated equipment. Clinical illness is generally mild, usually asymptomatic; it is characterized by waxing and waning aminotransferase elevations and a high rate (50 to 85 percent) of chronic HCV infection. HCV infection is associated with an increased risk of cirrhosis and hepatocellular carcinoma. Diagnosis of HCV infection is based on immunoassays that detect antibodies to HCV (anti-HCV). These tests are complemented by assays for HCV RNA to confirm active versus past infection.

Endocrine Evaluation

Endocrine disease is of great relevance to psychiatry. Endocrine illness frequently has psychiatric manifestations. Management of psychiatric

illness may be complicated by comorbid endocrine disease. For these reasons, screening for endocrine disease often is of relevance to the psychiatrist.

Adrenal Disease. Adrenal disease may have psychiatric manifestations, including depression, anxiety, mania, dementia, psychosis, and delirium. However, patients with adrenal disease rarely come to the attention of psychiatrists. Assessment and management of these patients are best done in conjunction with specialists.

Low plasma levels of cortisol are found in Addison disease. These patients may have symptoms that are also common in psychiatric conditions including fatigue, anorexia, weight loss, and malaise. Patients may also have memory impairment, confusion, or delirium. Depression or psychosis with hallucinations and delusions may occur.

Elevated levels of cortisol are seen in Cushing syndrome. About half of all patients with Cushing syndrome develop psychiatric symptoms. These symptoms may include lability, irritability, anxiety, panic attacks, depressed mood, euphoria, mania, or paranoia. Cognitive dysfunction may include cognitive slowing and poor short-term memory. Symptoms usually improve when cortisol normalizes. If not, or if symptoms are severe, then psychiatric treatment may be necessary.

Cortisol levels have not been found to be useful in the assessment or management of primary psychiatric disease. In particular, the dexamethasone-suppression test (DST) remains a research tool in psychiatry that is not used in routine clinical care.

Anabolic Steroid Use. Use of anabolic steroids has been associated with irritability, aggression, depression, and psychosis. Athletes and bodybuilders are common abusers of anabolic steroids. Urine specimens can be used to screen for these agents. Because so many compounds have been synthesized, a variety of tests may be required to confirm the diagnosis, depending upon the compound that has been used. Consultation with a specialist is advised.

Generally, androgens other than testosterone can be detected by gas chromatography and mass spectroscopy. Exogenous testosterone is indistinguishable from endogenous testosterone. When the use of exogenous testosterone is suspected, the urinary ratio of testosterone to its endogenous epimer, epitestosterone, is assessed. Usually the ratio ranges from 1:1 to 3:1. When exogenous use occurs, the ratio may be greater than 4:1 and even greater than 6:1. Another method for detection of exogenous testosterone use is to evaluate the ratio of ^{13}C to ^{12}C in urinary metabolites of testosterone. Because synthetic testosterone preparations are derived from plant sterols, their presence, in comparison to the presence of endogenous testosterone alone, produces a lower ratio of ^{13}C to ^{12}C.

Antidiuretic Hormone. Arginine vasopressin (AVP), also called antidiuretic hormone (ADH), is decreased in central DI. DI may be central (due to the pituitary or hypothalamus) or nephrogenic. Nephrogenic DI may be acquired or due to an inherited X-linked condition. Lithium-induced DI is an example of an acquired form of DI. Lithium has been shown to decrease the sensitivity of renal tubules to AVP. Patients with central DI respond to the administration of vasopressin with a decrease in urine output. Secondary central DI may develop in response to head trauma that produces damage in the pituitary or hypothalamus.

About one-fifth of patients taking lithium develop polyuria, and a larger amount may have some degree of impairment in concentrating urine. Chronic treatment with lithium is a common cause of nephrogenic DI. However, there are other causes of polyuria in lithium-treated patients in addition to nephrogenic DI. Primary polydipsia is common and is often associated with the dry mouth associated with many psychiatric medications. When patients chronically treated with lithium develop altered mental status, they may become functionally fluid restricted due to inability to drink to thirst (often with polydipsia and polyuria at baseline). Close monitoring of fluid intake and urine output is required; patients are at high risk of developing hypernatremia and severe dehydration if they have impaired ability to drink to thirst, combined with impaired ability to concentrate urine.

A 73-year-old gentleman with bipolar I disorder had been stably managed on lithium for over 40 years. He presented to the hospital with altered mental status. Pneumonia was diagnosed, and the patient ultimately required intubation due to respiratory distress. While intubated, he developed hypernatremia and acute renal failure (prerenal). A review of his fluid intake and output revealed that he'd had several liters of daily urine output, but was not receiving adequate fluid replacement. Following nephrology evaluation of serum and urine osmolality studies, DI was diagnosed. IV fluids were increased and adjusted to provide free water replacement.

Excessive secretion of AVP results in increased retention of fluid in the body. This condition is called *SIADH*. Water retention in SIADH causes hyponatremia. SIADH may develop in response to injury to the brain or from medication administration (including phenothiazines, butyrophenones, carbamazepine, and oxcarbazepine). The hyponatremia associated with this condition may produce delirium. The relationship between AVP and psychogenic polydipsia remains unclear. However, it appears that in a minority of cases with psychogenic polydipsia AVP levels may be abnormal. SIADH is associated with hyponatremia, low serum osmolality, and high urine osmolality. It should be considered when patients demonstrate hyponatremia and concentrated urine (i.e., high urine osmolality or high specific gravity). This condition should be

investigated with the assistance of a specialist.

Human Chorionic Gonadotropin. Human chorionic gonadotropin (hCG) can be assessed in the urine and blood. The urine test for hCG is the basis for the commonly used urine pregnancy test. Routine tests are most accurate when performed 1 to 2 weeks after a missed period. However, there are ultrasensitive urine hCG tests that can accurately detect pregnancy 7 days after fertilization. Pregnancy tests often are obtained before initiating certain psychotropic medications, such as lithium, carbamazepine, and valproic acid, which are associated with congenital anomalies.

Parathyroid Hormone. Parathyroid hormone modulates serum concentrations of calcium and phosphorus. Dysregulation in this hormone and the resulting production of abnormalities in calcium and phosphorus may produce depression or delirium. Chronic lithium treatment may result in hyperparathyroidism and hypercalcemia in a small percentage of patients.

Neuropsychiatric symptoms associated with hyperparathyroidism and hypercalcemia include depression, anxiety, cognitive dulling, lethargy, and psychosis; delirium and stupor may occur as calcium levels rise. Symptoms usually improve with treatment of the underlying condition.

Prolactin. Prolactin levels may become elevated in response to the administration of antipsychotic agents which block dopamine receptors in the pituitary. Dopamine inhibits the release of prolactin; thus, dopamine blockade leads to an increase in prolactin release.

Elevated prolactin levels in premenopausal women are associated with menstrual abnormalities, infertility, alterations in libido, decreased bone density, and galactorrhea. Most premenopausal women who have hyperprolactinemia do not have galactorrhea. Generally, levels greater than 100 ng/mL are associated with infertility and amenorrhea, hot flashes, and vaginal dryness. Levels from 50 to 100 ng/mL may cause amenorrhea or oligomenorrhea. Values of 20 to 50 ng/mL may cause infertility even when there is no abnormality of the menstrual cycle. The normal range for serum prolactin is from 5 to 20 ng/mL.

In men, hyperprolactinemia results in decreased libido, erectile dysfunction, fatigue, gynecomastia, and infertility. Only rarely does galactorrhea occur. Chronic hyperprolactinemia may result in loss of body hair, loss of muscle mass, and osteoporosis.

Cerebral MRI is not usually performed if the patient is taking an antipsychotic drug known to cause hyperprolactinemia and the magnitude of the prolactin elevation is consistent with drug-induced causes.

Prolactin levels may briefly rise after a seizure. For this reason, prompt

measurement of a prolactin level after possible seizure activity may assist in differentiating a seizure from a nonepileptic behavioral spell.

Thyroid Hormone. Disease of the thyroid is associated with many psychiatric manifestations. Thyroid disease is most commonly associated with depression and anxiety but may also give rise to symptoms of panic, dementia, and psychosis. Thyroid studies are often evaluated in patients with depression, as thyroid disease may mimic depression. It is difficult to achieve euthymia if a patient is not euthyroid.

Measurements of TSH typically are used as a screening test for thyroid disease. If normal, then no further assessment is performed. If high, then TSH may be repeated and free T4 added to assess the degree of hypothyroidism. If serum TSH is low, then free T4 and T3 are assessed to determine the degree of hyperthyroidism. Hypothyroidism is characterized by an elevated TSH and a low free T4. Subclinical hypothyroidism consists of a high TSH and a normal free T4. Hyperthyroidism consists of low TSH and high free T4 and T3.

Usually, thyroid function is assessed in psychiatric patients with a history of thyroid disease or those with symptoms suggestive of thyroid illness. Such symptoms may include increased appetite, weight loss, and heat intolerance in hyperthyroidism or weight gain, fatigue, and cold intolerance in hypothyroidism. Alternatively, risk factors for thyroid disease may prompt examination of thyroid function. Risk factors include prior treatment with lithium or iodine-131, a history of irradiation of the neck and chest, or a history of thyroidectomy. Additionally, thyroid function commonly is assessed in patients currently receiving exogenous thyroid hormone replacement to ensure that they are euthyroid.

Interpretation of thyroid function tests may be complex and is often best done with the assistance of a specialist. Interpretation is further complicated by the presence of transient abnormalities of thyroid function in many newly admitted psychiatric patients. These abnormalities normalize over time without treatment and may be related to stress. Thyroid tests have not been found to be useful in the assessment or management of primary psychiatric illness.

Pancreatic Function. The pancreas is both an exocrine and endocrine gland. The endocrine function of the pancreas is to control blood sugar levels through release of the hormones insulin and glucagon. See the section on metabolic syndrome for a discussion of insulin. The exocrine function of the pancreas is to assist in digestion. Pancreatic function can be disturbed as a consequence of alcohol use disorders or, more rarely, as a complication of psychiatric treatment. Pancreatitis may develop in patients with alcohol use disorders. Additionally, pancreatitis is a rare complication of treatment with some psychotropic medications, including clozapine,

risperidone, olanzapine, and valproate. Pancreatitis is usually accompanied by elevations in serum amylase and lipase. Serum amylase levels may be fractionated into salivary and pancreatic components. Elevations in salivary levels are commonly observed in eating disorder patients who purge by self-induced emesis. When pancreatic disease is suspected, serum lipase typically is measured. Elevations in amylase and lipase suggest pancreatic disease. Eating disorder patients who purge usually have elevated levels of serum amylase and normal levels of serum lipase.

Autoimmune and Inflammatory Disease

Psychiatric symptoms are highly comorbid in autoimmune and inflammatory conditions. In fact, autoimmune and inflammatory conditions sometimes present with psychiatric symptoms, and thus should be included in the differential diagnosis for patients with compatible historical and clinical features.

Systemic Lupus Erythematosus. Neuropsychiatric manifestations of SLE include depression, cognitive impairment, delirium, mania, and psychosis. About 5 percent of patients with SLE experience psychosis. Similar symptoms may be produced by steroids, which often are used to treat this condition. Cognitive dysfunction is common and may include apraxia, aphasia, and agnosia along with impairments in abstract thinking, short- and long-term memory, judgment, and visual–spatial function. Depression is also highly comorbid in patients with SLE.

SLE is a chronic inflammatory disorder. Tests for SLE are based on the detection of antibodies formed as part of the disease. Antinuclear antibodies are found in virtually all patients with SLE at some point during the illness. An indirect immunofluorescence assay is used to detect the antinuclear antibodies. This test can be positive in a variety of rheumatological diseases. For this reason, a positive result usually is followed by additional tests, including a test to detect anti-double–stranded deoxyribonucleic acid (DNA) antibodies. Anti-dsDNA antibodies, when associated with antinuclear antibodies, are strongly suggestive of a diagnosis of SLE. Anti-dsDNA antibodies often fluctuate with disease activity, may predict flares, and are often followed to monitor disease activity and response to treatment.

Autoimmune or Paraneoplastic Encephalopathy. Over the past decade there has been growing recognition and identification of antibodies that target the CNS resulting in a variety of psychiatric and neurological signs and symptoms. Though presentations vary, common presenting symptoms include cognitive impairment, mood symptoms, psychosis, seizure, and abnormalities of movement. Infrequently, patients may present with psychiatric symptoms in the absence of neurological symptoms. These

syndromes occur across the lifespan. Psychiatrists should suspect these disorders when patients have atypical presentations of psychiatric illness, especially when there are comorbid neurological signs or symptoms.

There are currently a number of clinical laboratory evaluations for antibodies associated with paraneoplastic and autoimmune neuropsychiatric disorders. Antibodies may target either intracellular or extracellular antigens. Antibodies to intracellular antigens include anti-Hu, anti-Ma2, and anti-CRMP5, among others. Extracellular (synaptic) antibody targets include N-methyl-D-aspartate (NMDA) receptors, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors, gamma-aminobutyric acid (GABA) receptors, and the voltage-gated potassium channel complex, among others. Laboratory evaluations for CNS autoantibodies will certainly continue to grow as research continues; reputable laboratories should be consulted for up-to-date testing panels.

When an autoimmune or paraneoplastic encephalopathy is suspected, serum and CSF antibody testing, routine CSF studies, as well as EEG and brain MRI should be considered. These studies serve both to rule out other conditions (such as brain neoplasm), and may also provide supporting evidence for an autoimmune or paraneoplastic cause. Patients with autoimmune or paraneoplastic disorders may have mildly to moderately elevated WBC and protein in CSF. EEG may reveal seizure activity, epileptiform discharges, and/or focal or generalized slowing. MRI is often normal, but may reveal signal hyperintensities on FLAIR or T2-weighted imaging sequences. The prompt identification of autoimmune or paraneoplastic neuropsychiatric disorders facilitates rapid initiation of appropriate treatments.

An 18-year-old female with no past medical or psychiatric history presented to the emergency department with abrupt onset of mood lability, paranoia, and auditory hallucinations. She was inattentive, difficult to engage, and exhibited echolalia. She had occasional flailing movements of her upper extremities. Collateral history from her father revealed that she'd had a generalized tonic–clonic seizure 1 day prior to presentation, for which she was evaluated in the emergency department and released home, with an appointment for outpatient neurological follow-up. At her second emergency department visit in 2 days, EEG revealed epileptiform discharges. MRI was unremarkable. CSF examination revealed mildly elevated WBC (18 cells/μL) and modestly elevated protein (74 mg/dL). Infectious workup was negative. CSF antibody evaluation revealed IgG antibodies to the NMDA receptor, and autoimmune NMDA receptor encephalitis was diagnosed.

Other Autoimmune and Inflammatory Conditions. Autoimmune and inflammatory conditions ranging from multiple sclerosis (MS) to rheumatoid arthritis to inflammatory bowel disease are frequently associated with comorbid psychiatric illness, especially affective disorders. This is likely due to a variety of factors, including the psychosocial impact of chronic illness, treatment with medications that may have psychiatric

side effects, as well as potential neurobiological effects of inflammatory cytokines or direct involvement of the brain (as in multiple sclerosis). It is important to identify and treat comorbid psychiatric illnesses that cause suffering, worsen quality of life, and threaten health outcomes. Similarly, it is important for psychiatrists to identify signs and symptoms of autoimmune and inflammatory conditions in their patients, and refer to specialists for appropriate care. When managing psychiatric patients who develop symptoms concerning for autoimmune or inflammatory illness, referral to appropriate specialists who can order the appropriate medical assessment and laboratory evaluation is essential.

Neurological Disorders

Many neurological disorders have accompanying psychiatric symptoms. Cognitive disorders, cerebrovascular disease, movement disorders, seizure disorders, and demyelinating disorders are often accompanied by psychiatric symptoms. Imaging studies, EEG, and neuromuscular studies, in addition to laboratory evaluations of blood and CSF, may be part of the diagnostic evaluation of these disorders.

Neoplastic Disorders

Infrequently, psychiatric symptoms may be the result of a primary or metastatic neoplasm involving the CNS. Most often, neurological symptoms accompany psychiatric symptoms when there is a CNS tumor, but this is not always the case. Neuroimaging studies are the medical assessment of choice to rule out CNS tumors as etiological for neuropsychiatric symptoms. Some other cancers, such as pancreatic cancer, are commonly accompanied by depression, even before the cancer has been diagnosed. In the setting of psychiatric symptoms accompanied by weight loss and new or concerning somatic symptoms (abdominal pain, jaundice, bowel changes, hemoptysis), workup for malignancy may be indicated, depending upon patient demographics, risk factors, history, and examination findings. This may include positron emission tomography (PET) studies, as well as other targeted evaluations depending upon presenting signs and symptoms.

Head Trauma

In an acutely altered patient, trauma must be ruled out via history and physical examination, with head imaging supplementing the evaluation when indicated. Head computed tomography (CT) is superior at identifying bony deformities and acute bleeds, and is thus generally preferred in the acute evaluation of altered mental status when there is suspicion for head trauma.

Genetic Disorders

Though many psychiatric disorders have been demonstrated to be heritable to a certain extent, classically inherited genetic disorders are not commonly included in the differential diagnosis of psychiatric illness. Occasionally, laboratory evaluation for genetic disorders is indicated. Sometimes there is a known family history of a genetic disorder that prompts evaluation. Other times, compatible clinical features and an astute clinician lead to diagnosis of a genetic disorder, even in the absence of family history.

A 43-year-old male with history of recurrent major depression presented with new-onset psychotic symptoms. Neurological examination revealed chorea, which had reportedly developed over the past 2 years. The patient did not know his family history; he had been adopted as an infant. Genetic testing revealed an abnormally large number of trinucleotide repeats (CAG) in the *huntingtin* gene on chromosome 4. Huntington disease was diagnosed.

Huntington Disease. Huntington disease is an autosomal dominant disorder involving an excess of trinucleotide repeats in the *huntingtin* gene on chromosome 4. Laboratory evaluation for Huntington disease is most commonly indicated when there is a known family history, often in combination with compatible clinical features, or for planning purposes. Huntington disease is also known as Huntington chorea. The basal ganglia are affected, leading to a movement disorder dominated by chorea (involuntary, fidgety movements of the face and extremities). Dystonia, rigidity, and other movement abnormalities often occur as well. Cognitive deterioration and behavioral symptoms accompany the movement disorder, sometimes preceding it.

Wilson Disease. Wilson disease is a rare autosomal recessive disorder that leads to abnormal copper metabolism. The disorder leads to excessive copper deposition in the brain and liver, and may result in decreased intellectual functioning, personality changes, psychosis, liver dysfunction, and movement disorder. Symptoms usually present in the second and third decades of life. Laboratory assessment for Wilson disease includes the measurement of serum ceruloplasmin, the transport protein for copper, which is low, and urine copper, measured in a 24-hour specimen, which is elevated.

Other Genetic Disorders. Other genetic disorders may sometimes be suspected, and laboratory evaluations may be ordered when a clinical history and examination is concerning for a genetic etiology. For example, Fragile X syndrome may be suspected in a child with intellectual disability or autism who has characteristic physical features such as large ears and a long face, particularly when there is a compatible family history of similar

disability. Other genetic disorders with associated neuropsychiatric symptoms include Prader–Willi syndrome, adrenoleukodystrophy and other inborn errors of metabolism, among others. These disorders are rare; when a genetic disorder is suspected, the appropriate laboratory evaluation is determined in consultation with specialists.

Nutritional Deficiency

A number of nutritional deficiencies may have accompanying neuropsychiatric symptoms; clinical history and examination should provide clues regarding whether nutritional deficiency may be contributing to the patient's neuropsychiatric presentation. Thiamine and folate deficiency are common in patients who abuse alcohol. Thiamine deficiency can lead to Wernicke's encephalopathy and Korsakoff syndrome in severe cases. Blood thiamine concentration can be measured to assess for deficiency. However, for patients at known risk for thiamine deficiency, most clinicians simply order thiamine supplementation without first ordering laboratory evaluation of thiamine. Folate and B_{12} deficiency, among other nutritional deficiencies, may be associated with dementia, delirium, psychosis, fatigue, and personality changes. Low folate levels may be found in patients who abuse alcohol or have poor nutrition. Additionally, some medications such as methotrexate and phenytoin can lead to folate deficiency. B_{12} deficiency is usually the result of poor absorption due to pernicious anemia or gastric disease. Folate and B_{12} can be directly measured. The metabolic intermediaries methylmalonic acid and homocysteine may be useful in the assessment of folate and B_{12} deficiency, particularly when vitamin levels are in a borderline range where deficiency cannot be definitely ruled in or out. Many severe nutritional deficiencies can result in neuropsychiatric disturbances, but the aforementioned deficiencies are those most commonly assessed in psychiatric patients.

Intoxication/Toxic Exposure

Laboratory assessment for intoxication or toxic exposures may be indicated under several different circumstances. Often, a toxicology screen is essential in establishing the correct diagnosis for a patient presenting with new symptoms. For example, paranoia may be due to a primary psychiatric disorder, but it also may be due to intoxication with a number of substances. Toxicology screening may also be helpful in the evaluation and management of substance use disorders. Finally, there may be concern for environmental exposures or for accidental or intentional overdoses.

Laboratory studies in intoxicated adults may help to identify the cause or consequence of intoxication. When the patient is unable or unwilling to provide information about the source of intoxication, appropriate studies,

in addition to toxicology studies, include calcium, serum BUN, Cr, CK, glucose, ketones, electrolytes, LFTs, magnesium, serum osmolality, and a urinalysis. Reproductive-age women should undergo urine pregnancy testing.

Testing for drugs of abuse usually is performed on urine specimens. It also may be performed on specimens of blood, breath (alcohol), hair, saliva, and sweat. Urine screens provide information about recent use of frequently abused drugs such as alcohol, amphetamines, cocaine, marijuana, opioids, and phencyclidine (PCP) along with 3,4-methylenedioxymethamphetamine (MDMA) (ecstasy). Savvy patients may attempt to avoid a positive urine drug screen by providing a sample from another patient, diluting the sample, or modifying it with additives that may interfere with the drug abuse screen. Many substances may produce false positives with urine drug screening tests. When a false positive is suspected a confirmatory test may be requested.

In cases where it is unclear what agents that the patient may have ingested, it is useful to screen for salicylates and acetaminophen. In the early stages of overdoses with lethal amounts of these medications there may be few clinical signs. However, if an overdose is detected early, then effective treatments are available.

Comprehensive qualitative toxicology screening is usually performed by liquid and gas chromatography. This may require many hours to perform and is rarely done in routine clinical situations. It is usually performed in patients with unexplained toxicity and an atypical clinical picture. Quantitative toxicology assessments may be useful in managing patients who have overdosed, when combined with clinical assessment and knowledge of when the ingestion occurred. With the exception of alcohol, drug levels for substances of abuse are not usually determined. Instead, only the presence or absence of the drug is determined. There is usually not a meaningful or useful correlation between the level of the drug and clinical behavior.

Drug Abuse. Patients may be unreliable when reporting their substance use history. Substance-induced mental disorders often resemble primary psychiatric disorders. Furthermore, substance use can exacerbate preexisting mental illness. Indications for ordering a drug abuse screen include unexplained behavioral symptoms, a history of illicit drug use or dependence in a new patient evaluation, or a high-risk background (e.g., criminal record, adolescents, and history of prostitution). A drug abuse screen also is frequently used to monitor patient abstinence during treatment of substance use disorders. Such tests can be ordered on a scheduled or random basis. Many clinicians believe random testing may be more accurate in the assessment of abstinence. The tests also may help to motivate the patient.

Other laboratory data may suggest a problem with substance abuse. An increase in the mean corpuscular volume (MCV) is associated with alcohol abuse. Liver enzymes may be increased with alcohol abuse or from hepatitis B or C acquired from IV drug abuse. Serological testing for hepatitis B or C can confirm that diagnosis.

TESTED SUBSTANCES. Routine tests are available for PCP, cocaine, tetrahydrocannabinol (THC) (also known as *marijuana*), benzodiazepines, methamphetamine and its metabolite amphetamine, morphine (Duramorph), codeine, methadone (Dolophine), propoxyphene (Darvon), barbiturates, lysergic acid diethylamide (LSD), and MDMA (also known as *ecstasy*). Different laboratories have varying thresholds for detection of these substances. What constitutes a positive result varies widely. Some laboratories report a positive finding based only on the results of an initial screening test. Other facilities require a second, confirming test, often by a more sensitive or specific method.

Drug screening tests may have high false-positive rates. This is often due to the interaction of prescribed medication with the test, resulting in false-positive results, and lack of confirmatory testing. False-negative tests are common as well. False-negative results may be due to problems with specimen collection and storage. An unbroken chain of custody of the specimen is important. Patients may attempt to alter a specimen by dilution, by changing the pH (bleach and baking soda), or by using other additives (perfume and salt) that interfere with the test. Allowing a specimen to remain at room temperature too long may result in inaccurate testing.

Table 7.6–1.
Drugs of Abuse That Can Be Detected in Urine

Drug	Length of Time Detected in Urine
Alcohol	7–12 hours
Amphetamine	48–72 hours
Barbiturate	24 hours (short acting); 3 weeks (long acting)
Benzodiazepine	3 days
Cocaine	6–8 hours (metabolites 2–4 days)
Codeine	48 hours
Heroin	36–72 hours
Marijuana	2–7 days
Methadone	3 days
Methaqualone	7 days
Morphine	48–72 hours

The length of time that a substance can be detected in the urine is listed in Table 7.6–1. These are approximate times that are influenced by the frequency and duration of antecedent drug use and the presence of

illnesses that could alter excretion, such as renal or hepatic disease.

Urine drug screens may detect some opiate metabolites up to 3 to 4 days after use.

Heroin is metabolized to 6-monoacetylmorphine, a metabolite unique to heroin. Because it can be detected with urine testing, it can be used to differentiate heroin from other opioids. False-positive drug screen results for opiate may occur in patients taking quinolones or rifampin. Poppy seeds have also been reported to produce false-positive results. Confirmation testing can identify false positives from quinolones or rifampin but may not be able to detect poppy seed use.

Patients using dronabinol for chemotherapy-induced nausea or HIV wasting syndrome will test positive for THC/marijuana. It is not possible to differentiate dronabinol from marijuana on urine drug screens.

Alcohol. There is no single test or finding on physical examination that is diagnostic for alcohol use disorders. The history of the pattern of alcohol ingestion is most important in making the diagnosis. Laboratory test results and findings on physical examination may help to confirm the diagnosis. Alcohol use disorders may be associated with telangiectasias, hepatomegaly, and evidence of trauma on physical examination. In patients with acute alcohol intoxication, a blood alcohol level (BAL) may be useful. A high BAL in a patient who clinically does not show significant intoxication is consistent with tolerance. The degree of alcohol intoxication can also be assessed using the concentration of alcohol in expired respirations (breathalyzer). BAL is used to legally define intoxication in the determination of whether an individual is driving under the influence. The legal limit in many states is 80 mg/dL. However, clinical manifestations of intoxication vary with an individual's degree of alcohol tolerance. At the same BAL, an individual who chronically abuses alcohol may exhibit less impairment than an alcohol naive individual. Generally a BAL in the range of 50 to 100 mg/dL is associated with impaired judgment and coordination, and levels greater than 100 mg/dL produce ataxia.

Symptoms of withdrawal can occur with elevated BALs. It is the degree of change and the rate of change in BAL that usually are associated with symptoms of withdrawal. Alcohol withdrawal presents with an elevation in blood pressure and pulse and a tremor. These symptoms commonly occur when serum alcohol levels are low or rapidly declining. Active intervention with benzodiazepines may be necessary to prevent the condition from progressing to alcohol withdrawal delirium, which is a medical emergency that may be fatal. There is a need to assess the alcohol-using patient for withdrawal when the BAL is low or undetectable to prevent the evolution of withdrawal symptoms.

Common laboratory studies for the evaluation of a patient with the diagnosis of alcohol abuse or dependence include a CBC (to assess MCV

and red cell distribution width [RDW] for a macrocytic profile), serum electrolytes (to obtain levels of magnesium and potassium, which are commonly low), and liver enzymes (to assess GGT, which is sensitive for acute alcohol abuse, along with AST and ALT). Alcoholic liver disease is characterized by elevations in AST and ALT, typically in a ratio of AST to ALT of 2:1 or greater. GGT may be elevated. Carbohydrate-deficient transferrin (CDT) may also be helpful in the identification of chronic heavy alcohol use. Total protein, serum albumin, and PT are assessed as indices of nutritional status and liver function. Folate and thiamine deficiencies are frequently present in patients with alcohol abuse. Urine toxicology screens for other drugs of abuse may be helpful in understanding the full clinical picture.

Environmental Toxins. Specific toxins are associated with a variety of behavioral abnormalities. Exposure to toxins may occur through occupation or hobbies. Aluminum intoxication can cause a dementia-like condition. Aluminum can be detected in the urine or blood. Arsenic intoxication may cause fatigue, loss of consciousness, anemia, and hair loss. Arsenic can be detected in urine, blood, and hair. Manganese intoxication may present with delirium, confusion, and a parkinsonian syndrome. Manganese may be detected in urine, blood, and hair. Symptoms of mercury intoxication include apathy, poor memory, lability, headache, and fatigue. Mercury can be detected in urine, blood, and hair. Manifestations of lead intoxication include encephalopathy, irritability, apathy, and anorexia. A blood lead level is used to diagnose lead toxicity occurring over the past few weeks. The free erythrocyte protoporphyrin test is a test for lead toxicity over the past 3 months.

Significant exposure to organic compounds, such as insecticides, may produce behavioral abnormalities. Many insecticides have strong anticholinergic effects. Presentation with the cluster of anticholinergic effects such as dry mouth, blurred vision, constipation, and urinary retention typically alerts clinicians to the possibility of toxicity. There are no readily available laboratory tests to detect these compounds. Poison control centers may assist in the identification of appropriate testing facilities.

Volatile Solvent Inhalation. Volatile substances produce vapors that are inhaled for their psychoactive effect. The most commonly abused volatile solvents include gasoline, glue, paint thinner, and correction fluid. The aerosol propellants from cleaning sprays, deodorant sprays, and whipped cream containers may be abused. Nitrites, such as amyl nitrite ("poppers") and butyl nitrite vials ("rush"), and anesthetic gases, such as chloroform, ether, and nitrous oxide, are also abused.

Chronic abuse of volatile solvents is associated with damage to the

brain, liver, kidneys, lung, heart, bone marrow, and blood. Abuse may produce hypoxia or anoxia. Signs of abuse include short-term memory loss, cognitive impairment, slurred and "scanning" speech, and tremor. Cardiac arrhythmias may occur. Exposure to toluene, which is present in many cleaning solutions, paints, and glues, has been associated with loss of clear gray–white matter differentiation and with brain atrophy on MRI scans. Methemoglobinemia has occurred with butyl nitrite abuse. Chronic use of volatile solvents is associated with the production of panic attacks and an organic personality disorder. Chronic use may also produce impairment in working memory and executive cognitive function.

Initial treatment is to prevent access to the drug and to allow clearing of the substance through the respiratory system. When hypoxia occurs, supplemental oxygen administration may be necessary. This will not only treat the hypoxia but will also facilitate clearance of the inhaled chemical. Antiarrhythmic medication may be needed. Over the long term, patients will benefit from entering sobriety programs to prevent relapse.

Acetaminophen or Salicylate Overdose. Intentional or unintentional overdose with acetaminophen or salicylates requires medical and laboratory assessment. In an altered patient, especially those with past suicide attempts or suicidal ideation, laboratory evaluation for ingestions should be performed.

ACETAMINOPHEN. Acetaminophen is one of the most frequently used agents in intentional drug overdoses and is a common cause of overdose-related deaths. Sequelae of acute overdose most notably include hepatotoxicity and acute hepatic failure, and also acute renal failure in some cases. Chronic abusers of alcohol may be particularly vulnerable to hepatotoxicity following acetaminophen overdose. After an overdose of acetaminophen, large amounts of toxic metabolites are formed as the primary metabolic pathways of sulfate and glucuronide conjugation become saturated. The toxic metabolite can be inactivated by glutathione. Acetylcysteine (Mucomyst) treatment must occur promptly after overdose to prevent hepatotoxicity. Acetylcysteine protects the liver by maintaining appropriate concentrations of glutathione. Activated charcoal effectively binds acetaminophen. After an acetaminophen overdose, activated charcoal is often administered with acetylcysteine to prevent hepatotoxicity. The plasma acetaminophen concentration should be determined not less than 4 hours after ingestion of the overdose. If the initial determination shows a high concentration, then a full course of acetylcysteine should be administered. If a low plasma concentration is reported, then initiation or continuation of acetylcysteine treatment is not necessary. LFTs (serum AST, serum ALT, PT, and bilirubin) should be performed at 24-hour intervals for at least 96 hours postingestion if the

plasma acetaminophen concentration indicates potential hepatotoxicity. If no abnormalities are detected within 96 hours, then further determinations are not needed.

SALICYLATE TOXICITY. Aspirin frequently is ingested in overdose. Consequently, serum salicylate levels often are obtained in overdose cases. Ingestion of sufficient amounts of aspirin may be fatal. Some patients with chronic inflammatory conditions and/or chronic pain may ingest large amounts of salicylates for therapeutic reasons. Ingestion of 10 to 30 g of aspirin may be fatal in adults. Most patients will develop symptoms of toxicity when salicylate levels are greater than 40 mg/dL (2.9 mmol/L). Signs and symptoms of toxicity include acid–base abnormalities, tachypnea, tinnitus, nausea, and vomiting. In cases of severe toxicity, symptoms may include hyperthermia, altered mental status, pulmonary edema, and death.

Treatment includes administration of multiple doses of activated charcoal as well as alkalinization of systemic pH by administering sodium bicarbonate. A salicylate level greater than 7.2 mmol/L is one of several absolute indications for hemodialysis following salicylate overdose. Renal failure, altered mental status, and pulmonary edema are other indications for hemodialysis.

The Role of Brain Imaging in Diagnosis

Imaging of the CNS can be broadly divided into two domains: Structural and functional. Structural imaging provides detailed, noninvasive visualization of the morphology of the brain. Functional imaging provides a visualization of the spatial distribution of specific biochemical processes. Structural imaging includes x-ray, CT, and MRI. Functional imaging includes PET, single photon emission computed tomography (SPECT), functional MRI (fMRI), and magnetic resonance spectroscopy (MRS). With the limited exception of PET scanning, functional imaging techniques are still research tools that are not yet ready for routine clinical use.

A common clinical dilemma involves whether to obtain an MRI or a CT scan. CT scans usually are less expensive and more commonly available. Although CT scans are a relatively inexpensive screening examination, most clinicians have a hypothesis-based approach to the selection of an imaging modality: The question that they need answered drives the selection of the scan. The following subsections provide some guidance regarding the relative advantages of MRI scans and CT scans.

Currently, the use of structural imaging techniques is limited to the identification of nonpsychiatric disease. Scans are performed when a clinician suspects or seeks to rule out a nonpsychiatric illness. Attempts to use structural imaging techniques to subcategorize various psychiatric illnesses have not yielded data that are useful in clinical practice.

Consequently, structural imaging techniques are not useful for making psychiatric diagnoses.

Neuroimaging in patients with dementia is associated with some degree of controversy. This is due to the large number of scans that do not have clinically significant findings. Some argue that noncontrast MRI scans or CT scans are useful screening for all dementia patients at the time of initial diagnosis. A contrasting point of view suggests that the probability of significant findings is greater in patients younger than 60 years of age, those with focal findings on the neurological examination, or recent onset of symptoms.

Findings in patients with dementia may include atrophy, either focal or generalized, and white matter lesions. These are nonspecific changes.

Imaging in delirious patients is unlikely to alter management. This is particularly true when there is no history of head trauma or loss of consciousness and there are no focal findings on the neurological examination. Imaging may be performed if the delirious patient fails to respond to treatment or has limited responsiveness or alertness. Imaging in delirious patients may reveal underlying factors that may predispose the patient to delirium. However, it is less common that these are etiological factors.

Magnetic Resonance Imaging. MRI scans are used to distinguish structural brain abnormalities that may be associated with a patient's behavioral changes. These studies provide the clinician with images of anatomical structures viewed from cross-sectional, coronal, or oblique perspectives. MRI scans can detect a large variety of structural abnormalities. The MRI is particularly useful in examining the temporal lobes, the cerebellum, and deep subcortical structures. It is unique in its ability to identify periventricular white matter hyperintensities. MRI scans are useful in examining the patient for particular diseases, such as nonmeningeal neoplasms, vascular malformations, seizure foci, demyelinating disorders, neurodegenerative disorders, and infarctions. Advantages of MRI include the absence of ionizing radiation and the absence of iodine-based contrast agents. Contrast agents are based on gadolinium, a rare earth metal, and usually do not provoke an allergic response.

MRI scans are not always obtained with the same parameters. The nature of the scan is modified in a manner that is most likely to provide data relevant to the clinical hypothesis that generated the request for the scan. For this reason, clinicians should provide the radiologist with clear, concise indications for the scan. Scan parameters include modes referred to as T1 and T2. T1-weighted scans provide detailed visualizations of brain anatomy. Extracellular fluid, such as edema, looks dark. In contrast, T2-weighted scans reveal a white pattern for edema. These scans are useful in

examining the patient for white matter disease.

MRI scans are contraindicated when the patient has a pacemaker, aneurysm clips, or ferromagnetic foreign bodies.

Computed Tomography. CT scans are used to identify structural brain abnormalities that may contribute to a patient's behavioral abnormalities. These studies provide the clinician with cross-sectional x-ray images of the brain. CT scans can detect a large variety of structural abnormalities in the cortical and subcortical regions of the brain. CT scans are useful when a clinician is looking for evidence of a stroke, subdural hematoma, tumor, or abscess. These studies also permit visualization of skull fractures. CT scans are the preferred modality when there is suspicion of a meningeal tumor, calcified lesions, acute subarachnoid or parenchymal hemorrhage, or acute parenchymal infarction. These scans have some resolution limits in the posterior fossa and the inferior surfaces of the temporal lobes owing to artifacts created by the skull. CT scans are not as successful in detecting focal gliosis as the MRI scans.

CT scans may be performed with or without contrast. The purpose of contrast is to enhance the visualization of diseases that alter the blood–brain barrier, such as tumors, strokes, abscesses, and other infections. Contrast agents are administered by IV means. Allergic reactions to contrast agents may occur and can include an anaphylactic reaction with cardiovascular collapse. This type of reaction most commonly occurs in patients with asthma or allergies to seafood.

Indications for a CT scan are not agreed upon in the literature. Less controversial indications include focal findings on the neurological examination or abnormalities on an EEG. Slightly more controversial indications include impaired cognition on the mental status examination, first episode of psychosis (especially if the presentation is atypical), delirium, late-onset (after 50 years of age) personality disorders, psychosis, or affective illness. Prolonged catatonia is also a possible indication.

The Role of the Electroencephalogram in Diagnosis

The EEG is the assessment of regional cerebral cortical electrical activity. Clinical neuroscience has a long history of using the EEG. The EEG can be used in different ways to study specific brain states or activities by modifications to the technique of data collection or to the data itself. EEG data can be displayed on paper tracings in the manner of conventional EEG recordings. Alternatively, the data can be digitized, and the digitized data can be transformed, often using a Fourier transformation, to yield color-coded topographic brain maps of regional activity. The collection periods can be prolonged, and the data can be electronically displayed along with video monitoring of the patient to provide telemetry assessments of patients with epilepsy. Telemetry assessments are typically performed in an

effort to correlate behavioral abnormalities with brain electrical activity as part of the workup of seizure disorders. Prolonged periods of EEG recording during sleep, when coupled with recording of a limited lead ECG and facial muscle activity, result in the sleep EEG or polysomnography. Many clinicians also use the EEG to monitor ECT administration.

Clinicians use the EEG to localize seizure foci and to evaluate delirium. The EEG and its topographical descendants have not found a clear role in the diagnostic assessment of psychiatric patients. The EEG is usually used in psychiatry to rule out nonpsychiatric disease, such as seizure disorders or delirium, as a cause of psychiatric symptoms. When the differential diagnosis includes strokes, tumors, subdural hematomas, or dementia, the yield is usually higher with imaging tests. Not surprisingly, the yield is the highest in patients with a history of a seizure disorder or a clinical history that is strongly suggestive of a recent seizure or other organic illness. Such clinical features would include a history of altered consciousness, atypical hallucinations (e.g., olfactory), head injury, and automatism. In addition, the EEG is commonly obtained when there is an abnormal CT or MRI. It is important to remember that seizures are a clinical diagnosis; a normal EEG does not rule out the possibility of a seizure disorder.

When a conventional EEG does not demonstrate the dysrhythmia (abnormal bursts of spike, sharp wave, or slow activity) associated with seizures, other measures can be taken to enhance the sensitivity of the evaluation. These measures include sleep deprivation, hyperventilation, and photic stimulation. Detection also can be improved through the use of nasopharyngeal leads, computer-transformed recordings, and evoked potentials (EPs). Sleep deprivation assists in seizure detection by a simple mechanism. Patients are more likely to fall asleep during the EEG after a period of sleep deprivation. It is during the changes in consciousness associated with various stages of sleep that the spike and sharp wave activity associated with seizure disorders is more likely to be detected. Hyperventilation, usually for approximately 3 minutes, may produce slow wave activity or spike waves. Photic stimulation consists of exposing the patient to a flashing strobe light during the EEG. Abnormal wave activity not synchronized to the flashing light is pathological. Studies involving the use of nasopharyngeal leads, although uncomfortable for the patients, may increase the yield of positive findings when there is a suspicion of temporal lobe epilepsy.

The EEG can be helpful in patients with altered consciousness by identifying contributing conditions. Nonconvulsive seizures lack motor manifestations but impair consciousness. Nonconvulsive status epilepticus may also cause impairment in consciousness.

Encephalopathies due to underlying metabolic alterations may show diffuse bilateral slowing of the background rhythm along with moderate or high wave amplitude. Specific EEG findings are more commonly seen in

some conditions such as the triphasic waves found in hepatic encephalopathy or periodic sharp wave complexes in Creutzfeldt–Jakob disease.

Evoked Potential. EP testing is the measurement of the EEG response to specific sensory stimulation. The stimulation may be visual, auditory, or somatosensory. During visual EPs, the patient is exposed to flashing lights or a checkerboard pattern. With auditory EP, the patient hears a specific tone. In somatosensory EP, the patient experiences an electrical stimulation to an extremity. These stimuli occur repeatedly while the patient undergoes a routine EEG. Using a computer, the responses to these stimuli are recorded and averaged. The time frame is measured in milliseconds. These tests are useful in neurology and neurosurgery. For example, they assist in the assessment of demyelinating disorders such as MS. In psychiatry, EP testing may help in the differentiation of organic from functional impairments. A classic example is the use of EP testing to evaluate possible hysterical blindness. The usefulness of these tests in psychiatry is still under investigation.

Polysomnography. Polysomnography is used to assess disorders of sleep by concurrently assessing the EEG, ECG, blood oxygen saturation, respirations, body temperature, electromyogram, and electrooculogram. Polysomnography has demonstrated an increase in the overall amount of rapid eye movement (REM) sleep and a shortened period before the onset of REM sleep (decreased REM latency) in patients with major depression. Polysomnography is useful in diagnosing comorbid sleep disorders so that appropriate treatment can be initiated.

A 57-year-old obese male with recurrent major depressive disorder presents for follow-up evaluation of his depression. He has had worsening of his mood as well as worsening of fatigue and daytime drowsiness. His wife reports that he snores loudly each night. A sleep study is ordered and completed, and obstructive sleep apnea is diagnosed. The patient's fatigue and daytime drowsiness improve with use of a continuous positive air pressure (CPAP) machine.

TREATMENT SELECTION AND MONITORING

Medical and laboratory assessments are important not only in the diagnostic assessment of patients with neuropsychiatric symptoms or disorders, but are also important in ensuring the safe initiation and monitoring of appropriate treatments. Pretreatment assessments may be necessary to determine the safety of recommended treatments. Many treatments for psychiatric disorders require routine monitoring of drug levels and/or physical health parameters which may be adversely impacted by treatments. A summary of suggested laboratory assessments for selected

psychotropic agents is provided in Table 7.6–2. Finally, suspected complications of psychiatric treatment may also prompt medical and laboratory assessments in psychiatric patients.

Medical Clearance for Electroconvulsive Therapy

Electroconvulsive therapy (ECT) is usually reserved for patients with the most treatment-resistant depression. Typical laboratory tests obtained before the administration of ECT include a CBC, serum electrolytes, urinalysis, and LFTs. However, no specific laboratory tests are required in the pre-ECT evaluation. Usually, an ECG is also obtained. A spinal x-ray series is no longer considered routinely indicated because of the low risk of spinal injury associated with modern administration techniques that use paralyzing agents. A comprehensive medical history and physical examination are useful screening tools to identify possible conditions that could complicate treatment. Patients with a complex or concerning medical history may require more extensive evaluation prior to initiating ECT; these additional evaluations are usually guided by evaluating consultants with expertise in medically assessing patients' risk level for undergoing procedures under anesthesia.

Electrocardiogram in Treatment Selection and Monitoring

The ECG is a graphical representation of the electrical activity of the heart. Abnormalities in this activity correlate with cardiac pathology. The ECG is most commonly used in psychiatry to assess side effects of psychotropic medications.

Many psychotropic medications have been associated with prolongation of the QT interval. This includes antipsychotic agents and many antidepressant medications. Torsades de pointes, a potentially fatal arrhythmia, is associated with some medications that cause QT prolongation. For this reason, clinicians may elect to obtain an ECG before initiation of treatment with agents known to prolong the QT interval, particularly if the patient has other risk factors for prolonged QT interval, such as a known history of QT prolongation (including congenital long QT syndrome), coronary artery disease, or heart failure. Bradycardia, hypokalemia or hypomagnesemia or the concurrent use of other drugs that prolong the QT interval all increase the risk for arrhythmias. When a patient in need of psychotropic medication has a prolonged QT interval, the use of agents with low or no risk for further prolonging the QT interval are recommended, when possible. This assessment, and the selection of agents most likely to help but not harm the patient, requires a careful risk/benefit analysis.

Table 7.6–2.

Suggested Laboratory Monitoring for Selected Psychotropic Agents

Psychotropic Agent	Baseline	Routine Monitoring
Atypical antipsychotics	Fasting lipid panel, hemoglobin A1c/fasting blood glucose ECG if patient has risk factors for QT prolongation	At 12 weeks, 1 year, and then annually: fasting lipid panel, fasting blood glucose/hemoglobin A1c.
Clozapine	CBC with differential (in addition to baseline labs for atypical antipsychotics)	CBC with differential: weekly for 6 months, every other week for an additional 6 months, and then every 4 weeks, if no abnormalities in WBC or ANC.
Lithium	BMP, TSH, CBC, calcium, pregnancy test ECG if >40 or with known cardiac risk factors	Weekly lithium trough level during dose titration, until stable therapeutic drug level is achieved, then every 3–6 months. Every 6 months: BMP, TSH, and calcium. Therapeutic serum levels: 0.6–1.2 mEq/L (for acute mania: 0.8–1.2 mEq/L).
Valproate	CBC, LFTs, pregnancy test Avoid in women with childbearing potential.	Obtain trough valproic acid level 1–2 weeks after initiation and each dosing change, then as clinically indicated. At 6 months and then annually: CBC, LFTs. Therapeutic serum levels: 50–125 µg/mL (for acute mania: 85–125 µg/mL).
Carbamazepine	CBC, LFTs, BMP, pregnancy test In patients of Asian ancestry, avoid use or first obtain genetic testing for HLA-B*1502 allele	Serum carbamazepine levels: obtain weekly levels until stable, then every 3–6 months. CBC: 2 weeks after dosage changes, 2 months after initiation, and then every 6 months. LFTs and BMP: every 6 months. Therapeutic serum levels (in seizure disorders): 4–12 µg/mL.
Oxcarbazepine	BMP, pregnancy	Only as clinically indicated.
TCAs	ECG if >40 or with known cardiac disease or risk factors	Blood levels as clinically indicated. Therapeutic serum levels: Nortriptyline[a]: 50–150 ng/mL Desipramine[a]: >125 ng/mL Imipramine (desipramine + imipramine): >200 ng/mL

[a]Frequency of serious side effects may increase over 300 ng/mL.

Tricyclic antidepressants (TCAs) are associated with ECG changes. Anticholinergic effects may increase heart rate. Prolongation of the PR, QT, and QRS intervals, along with ST-segment and T-wave abnormalities, may occur. The TCAs can cause or increase preexisting atrioventricular or bundle branch block. Many clinicians obtain an ECG before beginning a TCA in a patient older than 40 years of age and in any patient with known cardiovascular disease.

Lithium therapy can cause benign reversible T-wave changes; rarely, it

can impair sinoatrial (SA) node function, and can cause heart block. ECGs are often obtained before initiation of treatment with lithium, if there are risk factors for coronary artery disease. An ECG is obtained in cases of lithium toxicity or overdose.

Psychiatrists, when treating patients with certain psychiatric diagnoses, also use the ECG. Eating disorder patients commonly have low potassium levels that may result in abnormal ECG recordings. As the serum potassium drops below normal, T waves become flat (or inverted), and U waves may appear.

Metabolic Monitoring

Prior to initiating certain psychotropic agents, particularly second-generation antipsychotics, it is important to pursue baseline medical and laboratory assessments including measurement of weight, waist circumference, blood pressure, fasting glucose or hemoglobin A1c, and fasting lipid profile.

Many psychotropic drugs, including antidepressants, mood stabilizers, and antipsychotic medications, have weight gain as a potential side effect. Weight gain can lead to obesity, a major risk factor for the development of multiple medical problems such as hypertension, diabetes, cardiovascular disease, cerebrovascular disease, and lipid abnormalities. Weight gain can also contribute to medication noncompliance and poor self-image. Waist circumference measuring greater than 35 in for a woman or 40 in for a man also is of similar concern. Initial intervention to address weight gain typically includes education and initiation of an appropriate diet and exercise program. Discontinuation of the offending medication and substitution, if possible, with a medication less likely to give rise to weight gain also may be an appropriate intervention.

Weight. Obesity is associated with a heightened risk of coronary artery disease, stroke, type II diabetes mellitus, dyslipidemia, hypertension, and osteoarthritis. Treatment with first-generation antipsychotic agents as well as atypical agents is associated with obesity, although there is a significant variation between agents in the propensity to produce obesity. Clozapine and olanzapine are the most likely to produce weight gain. Ziprasidone and lurasidone have substantially lower risk for producing weight gain. All patients with schizophrenia and patients taking atypical antipsychotic medications should have their weight monitored. Likewise, patients with anxiety or mood disorders who take antidepressants or mood stabilizers associated with weight gain should have their weight monitored. Measurement of waist circumference may be used as a risk indicator. Men with a waist larger than 40 in and women with a waist larger than 35 in are at increased health risk. Education about the role of a healthy diet and exercise in preventing weight gain are effective. It is easier to prevent

weight gain than it is to lose weight. Many methods proposed to reduce weight are ineffective.

Body Mass Index. BMI values above 25 are associated with an increased risk for the illnesses noted above. Some clinicians advocate regular monitoring of BMI by both patient and clinician for patients with schizophrenia or other patients taking atypical antipsychotic medication and recommend intervention in the form of nutritional counseling, exercise program, or change in antipsychotic medication when BMI increases by one unit.

Monitoring of weight and body mass index (BMI) is important in the management of many patients with psychiatric illness. A variety of psychiatric illnesses can affect appetite and may cause weight loss or weight gain. Measurement of weight may be useful in assessing response to treatment and recovery.

Fasting Blood Sugar. Atypical antipsychotic agents have been associated with abnormalities in serum glucose levels, including the development of diabetes mellitus. Many clinicians monitor their patients who take atypical antipsychotic agents for the development of hyperglycemia by obtaining fasting blood glucose levels and glycosylated hemoglobin levels on a quarterly or semiannual basis. In addition, extremes in serum glucose concentrations have been associated with delirium. Evaluation for diabetes or other abnormalities in glucose metabolism is usually best done by specialists.

Glycosylated Hemoglobin. Glycosylated hemoglobin (hemoglobin A1c) is a correlate of the mean serum glucose concentration over the preceding 8 to 12 weeks. New hemoglobin is formed free of attached glucose. Because red cells are permeable to glucose, the sugar becomes permanently attached to hemoglobin at a concentration proportional to the serum glucose concentration. Thus, hemoglobin A1c is a measure of the mean blood glucose concentration over the prior 8 to 12 weeks. A normal hemoglobin A1c level is below 5.7 percent. A hemoglobin A1c level of 6.5 percent or above indicates diabetes mellitus. Levels between 5.7 and 6.4 indicate an increased risk of developing diabetes. For patients with diabetes, a hemoglobin A1c level of 7 percent or less is a common treatment target.

Lipids. Hyperlipidemia, including elevated cholesterol and triglyceride levels, is associated with cardiovascular and cerebrovascular disease. Certain antipsychotic medications, including clozapine and some of the second-generation antipsychotics, have been associated with hyperlipidemia. Fasting lipid panels should be evaluated prior to initiation

of medications associated with dyslipidemia, at 12 weeks into treatment, and annually thereafter.

Psychotropic Medications: Serum Levels and Physical Health Monitoring

Testing to determine blood concentrations of certain psychotropic medications enables the clinician to ascertain whether blood levels of medications are at therapeutic, subtherapeutic, or toxic levels. Psychiatric symptoms are not uncommon when prescribed medications are at toxic levels. In the debilitated and the elderly, pathological symptoms may occur at therapeutic concentrations. The normal reference range varies between laboratories. It is important to check with the laboratory performing the test to obtain the normal reference range for that laboratory.

Blood levels of psychotropic medications often are obtained when a patient displays mental status changes in the context of a recent change in dose of prescribed medication. Levels of psychotropic medications also are obtained after an overdose or suspected overdose of psychotropic medication. Additionally, a psychotropic medication level can be useful in assessing medication compliance.

Serum concentrations of psychotropic medications are assessed to minimize the risk of toxicity to patients receiving these medications and to ensure the administration of amounts sufficient to produce a therapeutic response. This is particularly true for medications with therapeutic blood levels. Medication levels are often influenced by hepatic metabolism. This metabolism occurs via the action of enzymes in the liver.

Cytochrome P450 enzymes (CYPs) are important in hepatic metabolism of drugs. Some drug substrates induce or inhibit CYP activity. Induction results in acceleration of drug metabolism and a resultant decrease in the levels of agents metabolized by that particular CYP. Administration of drugs that inhibit CYPs may result in higher drug levels of agents metabolized by the inhibited CYP. Genetic variation also influences CYP activity. Genotypes at a particular CYP, such as CYP2D6, may lead to any of the following phenotypes: ultrarapid metabolizer, extensive metabolizer, intermediate metabolizer, or poor metabolizer. At usual doses of a given medication, individuals with the poor metabolizer phenotype may have high serum levels of CYP2D6 metabolized drugs, due to poor ability of CYP2D6 to metabolize the drug. Inversely, ultrarapid metabolizers may have very low levels of drug, due to rapid metabolism.

Antipsychotic Agents. Serum concentrations are not routinely obtained for most antipsychotic medications. Management usually consists of titration of the prescribed dose to levels that achieve reasonable symptomatic control and acceptable side effects. However, in some cases, blood levels may be helpful. These situations include suspected medication noncompliance and cases of poor response with prescription of apparently

adequate doses. In the latter scenario, the concern may include rapid hepatic metabolism.

Atypical antipsychotic agents are associated with the development of the metabolic syndrome. Therefore, appropriate laboratory monitoring should occur at regular intervals. At the initiation of treatment with an atypical antipsychotic medication, a clinician should obtain a fasting lipid profile and a fasting blood glucose or hemoglobin A1c. This assessment should be repeated in 3 months and annually thereafter. Lipid profile and blood glucose should be repeated sooner if the patient experiences significant weight gain or symptoms suggestive of diabetes. A measurement of weight, BMI, and waist circumference also should be obtained at the initiation of treatment with an antipsychotic agent. These variables should be assessed monthly for the first 3 months of treatment and then at quarterly intervals. Baseline assessment should also include a personal or family history of obesity. Many clinicians will switch to a different antipsychotic agent when a 5 percent or more weight gain has occurred.

CLOZAPINE. Clozapine (Clozaril) levels are trough levels determined in the morning before administration of the morning dose of medication. A therapeutic range for clozapine has not been established; however, a level of 100 ng/mL is widely considered to be the minimum therapeutic threshold. Although concentrations between 200 and 700 ng/mL may correlate more with response, nonresponse does occur within this range. Studies suggest that plasma levels of clozapine in the range of 200 to 400 ng/mL may lead to achievement of maximal clinical efficacy; however, there is great variability among patients in terms of response and side effects at given dosages. One study demonstrated that patients with refractory schizophrenia who responded to clozapine were more likely to have clozapine blood levels of greater than 350 ng/mL. The risk of seizures increases with clozapine dosages of 600 mg/day or greater.

Clozapine is a common cause of leukopenia in psychiatry. A baseline CBC must be determined prior to the initiation of treatment. The baseline WBC count must be at least 3,500 per mm^3 and the absolute neutrophil count (ANC) must be at least 2,000 per mm^3 for initiation of therapy. Weekly CBCs are required during the first 6 months of treatment with clozapine because of the risk of agranulocytosis. Results must be sent to the dispensing pharmacy for the patient to receive his or her medication. Clozapine should be held for a WBC of less than 3,000 per mm^3 or an ANC of less than 1,500 per mm^3. For WBC 3,000 to 3,500/mm^3 and/or ANC 1,500 to 2,000/mm^3, clozapine may be continued but the frequency of WBC and ANC monitoring is increased to twice weekly. If the WBC remains greater than 3,500 per mm^3 and the ANC remains above 2,000 per mm^3 during the first 6 months of clozapine treatment, then monitoring may be reduced to every other week for the next 6 months. If both the WBC and the

ANC remain at appropriate levels after the second 6 months of treatment, then monitoring can be further decreased to every 4 weeks.

After clozapine discontinuation, clinicians usually monitor the CBC for an additional 4 weeks, if laboratory values are within normal limits. If clozapine treatment was interrupted or discontinued due to abnormal laboratory values, the degree of severity of leukopenia or granulocytopenia determines whether daily, twice-weekly, or weekly laboratory monitoring is required. Ongoing monitoring is required until the WBC is at least 3,500 per mm^3 and the ANC is at least 2,000 per mm^3, and for at least 4 weeks from the date of clozapine discontinuation.

When moderate leukopenia or granulocytopenia develops (WBC between 3,000 and 2,000/mm^3 and/or an ANC between 1,500 and 1,000/mm^3), clozapine treatment must be interrupted, but the patient may be retreated with clozapine in the future. Repeat exposure to clozapine may occur once the WBC is at least 3,500 per mm^3 and the ANC is at least 2,000 per mm^3. These patients require weekly monitoring for a year after the reinstatement of clozapine. Subsequently, frequency of monitoring may be reduced to every other week for 6 months and then further reduced to every 4 weeks. If granulocytopenia or leukopenia are severe (WBC <2,000/mm^3 and/or ANC <1,000/mm^3), the patient should not be rechallenged with clozapine in the future, regardless of WBC and ANC recovery.

If there is a break in clozapine treatment lasting less than 1 month during the period of weekly monitoring and the WBC consistently has been greater than 3,500 per mm^3 and the ANC consistently has been greater than 2,000 per mm^3, then treatment may resume with weekly monitoring. In addition, the patient will be eligible to reduce the frequency of monitoring to every other week after 6 months of total treatment rather than after 6 months from the time of the break in treatment. However, if the break in clozapine treatment lasts more than 1 month, then the patient must have a continuous 6 months of weekly monitoring without abnormalities before reducing the frequency of monitoring to every other week.

A similar rule applies for patients on clozapine for at least 6 months. If there is a break in clozapine treatment lasting less than 1 month in a patient undergoing CBC monitoring every other week and the WBC consistently has been greater than 3,500 per mm^3 and ANC consistently has been greater than 2,000 per mm^3, then the patient must undergo weekly monitoring for 6 weeks. Thereafter, the patient may resume every other week monitoring. However, if the break in treatment is longer than 1 month, then the patient must resume weekly monitoring for 6 months.

For a patient on clozapine for more than 1 year, a similar rule applies. If there is a break in clozapine treatment lasting less than 1 month in a patient undergoing monitoring every 4 weeks and WBC consistently has

been greater than 3,500 per mm^3 and ANC consistently has been greater than 2,000 per mm^3, the patient has been taking clozapine for longer than 1 year, and the gap in treatment is for less than 30 days, then they may resume treatment with weekly monitoring for 6 weeks and then return to monthly monitoring. However, if the treatment gap has been longer than 30 days, then monitoring must be weekly for 6 months, then every other week for another 6 months, and then monthly monitoring may resume.

CLOZAPINE-ASSOCIATED MYOCARDITIS OR CARDIOMYOPATHY. Infrequently, clozapine treatment may precipitate myocarditis or cardiomyopathy. If patients taking clozapine develop new cardiac or respiratory symptoms concerning for myocarditis or cardiomyopathy, medical evaluation should ensue with the assistance of specialists. Medical and laboratory assessment may include ECG, evaluation of cardiac biomarkers such as troponin, and echocardiography. Echocardiography is the visualization of cardiac anatomy by the use of computer-transformed echoes of ultrasound.

Psychiatry was consulted to evaluate a 36-year-old woman with schizophrenia who had been started on clozapine approximately 1 month ago, and had been admitted to medicine 1 day ago with fever, tachycardia, and dyspnea. The primary team was concerned about NMS, due to her autonomic symptoms. In addition to holding the clozapine, the psychiatrist recommended a cardiology evaluation, due to the possibility that the patient's symptoms represented clozapine-induced myocarditis. The patient had nonspecific ECG changes, elevated cardiac troponins, and left ventricular dysfunction on echocardiogram; myocarditis was diagnosed.

NEUROLEPTIC MALIGNANT SYNDROME. NMS is a rare, potentially fatal, consequence of neuroleptic administration. The syndrome consists of altered mental status, generalized rigidity (often severe), hyperthermia, and autonomic instability. Sustained muscle contraction results in peripheral heat generation and muscle breakdown. Muscle breakdown contributes to elevated levels of creatine kinase (CK). Peripheral heat generation with impaired central mechanisms of thermoregulation results in hyperthermia. Leukocytosis and low serum iron concentration are common. Myoglobinuric acute renal failure may occur due to rhabdomyolysis. Liver enzymes may become elevated. Patients may die from hyperpyrexia, aspiration pneumonia, renal failure, hepatic failure, respiratory arrest, or cardiovascular collapse. Treatment includes discontinuation of the neuroleptic, hydration, and general supportive care. Administration of dantrolene (a muscle relaxant) or bromocriptine (a dopamine agonist) may be considered, along with other treatments that have anecdotal success, including amantadine, benzodiazepines, or ECT. Data is limited with regard to the efficacy of any of these specific therapies.

A typical laboratory workup of NMS includes a CBC, serum electrolytes, BUN, Cr, and CK, as well as evaluation for other serious neurological or

medical conditions that may be mistaken for NMS. A urinalysis, including an assessment of urine myoglobin, is also usually performed. As part of the differential diagnosis, blood and urine cultures are performed as part of a fever workup. Pronounced elevations in the WBC may occur in NMS. WBCs are typically in the range from 10,000 to 40,000/mm^3.

CK elevations are typically more than 1000 IU/L. Moderate elevations in CK may occur in the earliest stages of the disease. CK level may be normal or near normal if significant rigidity has not yet developed. Other causes of muscle injury leading to increases in serum CK levels may include repeated intramuscular (IM) injections, prolonged or agitated periods in restraint, or dystonic reactions from neuroleptic administration.

Mood Stabilizers

CARBAMAZEPINE. Carbamazepine (Tegretol) may produce changes in the levels of WBCs, platelets, and, under rare circumstances, red blood cells. Aplastic anemia, thrombocytopenia, agranulocytosis, exfoliative dermatitis, and hepatic failure may all occur but are rare. Hyponatremia is common, particularly in the elderly. Pretreatment evaluation typically includes a CBC with differential and platelet count, as well as LFTs and renal function tests. Given the risk for hyponatremia, baseline serum electrolytes may also be obtained, particularly in more vulnerable populations such as the elderly. The risk of exfoliative dermatitis (e.g., Stevens–Johnson syndrome) is greater in patients with the HLA-B*1502 allele. This allele is more common in patients of Asian descent, and is largely absent in individuals of other ethnic backgrounds. Prior to initiating carbamazepine, genetic testing for this allele should be pursued in patients with ancestry which puts them at higher likelihood of having the HLA-B*1502 allele. If this allele is identified, that patient should not be treated with carbamazepine. Many clinicians monitor the CBC and LFTs every 2 weeks for the first 2 months of administration. Then, if the laboratory values have been within normal limits, CBC, serum carbamazepine level, and LFTs are monitored every 3 to 6 months. Patients should be cautioned to watch for petechiae, fever or signs of infection, pallor, or undue weakness, as well as signs and symptoms of hepatic injury or dermatological reactions.

Carbamazepine is teratogenic and may produce a variety of congenital abnormalities, including neural tube defects. A pretreatment urine pregnancy test usually is obtained in women of childbearing years. Women should be cautioned to use adequate contraception when taking carbamazepine. Carbamazepine may decrease plasma concentrations of hormonal contraceptives.

When used as an anticonvulsant, target blood levels for carbamazepine are 4 to 12 μg/mL. Target blood levels have not been established for mood disorder patients, but most clinicians aim for a level in this range before

discontinuation of carbamazepine owing to lack of efficacy. Levels should be obtained before the morning dose and after the patient has taken the same total daily dose for a minimum of 5 days. Serum levels of carbamazepine can help establish an effective dose for an individual patient. Higher serum levels (8 to 12 μg/mL) are associated with an increased risk of side effects. Toxic side effects are common with levels greater than 15 μg/mL.

Manifestations of toxicity may include nausea, vomiting, urinary retention, ataxia, confusion, drowsiness, agitation, or nystagmus. At very high levels, symptoms also may include cardiac dysrhythmias, seizures, and respiratory depression.

Carbamazepine levels may be altered by coadministration of medications that inhibit metabolism of carbamazepine. Such agents include erythromycin and related antibiotics, valproic acid, cimetidine, propoxyphene, and isoniazid, among other CYP 3A4 inhibitors. Carbamazepine is a powerful CYP 3A4 inducer, and thus must be used cautiously and with careful monitoring when combined with other medications metabolized by CYP 3A4. Further, carbamazepine has the ability to lower its own serum level through the process of hepatic enzyme autoinduction. Typically, autoinduction of hepatic enzymes occurs several weeks after initiation of carbamazepine treatment. Many clinicians follow levels over time and adjust the administered dose as necessary to maintain a therapeutic concentration. In malnourished patients or those with low protein due to other medical conditions, it may be advisable to monitor serum albumin levels. A portion of the drug is bound to serum proteins. It is the free (unbound) component that is biologically active. When there are reductions in serum proteins that bind the drug, adjustment in the dose may be necessary, particularly in patients who show signs of toxicity.

OXCARBAZEPINE. Oxcarbazepine is structurally similar to carbamazepine, but has fewer drug interactions and fewer side effects, with the exception of increased frequency of clinically significant hyponatremia. Many clinicians obtain baseline serum electrolytes and repeat this testing as clinically indicated. Therapeutic serum oxcarbazepine levels have not been established. In women of childbearing age, a pretreatment pregnancy test should be evaluated; women should be advised that hormonal contraception may be less effective during oxcarbazepine treatment.

A 58-year-old woman with bipolar I disorder was recently transitioned from carbamazepine to oxcarbazepine due to poor tolerability of carbamazepine. One month after initiation of oxcarbazepine, she and her spouse reported that she was having difficulty with memory and periods of acute confusion. She appeared drowsy in the office and had poor attention. Serum electrolyte evaluation revealed a sodium of 118 mEq/L. She was admitted to internal medicine for treatment, and oxcarbazepine was discontinued.

LAMOTRIGINE. Therapeutic serum lamotrigine levels have not been established. Dosing is titrated to clinical response.

LITHIUM. Lithium has a narrow therapeutic index. Consequently, blood levels of lithium must be monitored to achieve therapeutic dosing and avoid toxicity. For acute mania, target serum levels of lithium are in the range of 0.8 to 1.2 mEq/L. Optimum maintenance serum levels may vary from patient to patient, with the goal of achieving a balance between clinical efficacy and tolerability. Many clinicians aim for maintenance levels between 0.6 to 0.8 mEq/L, especially if patients complain of side effects at levels of 0.8 mEq/L or higher. Side effects are dose dependent. Symptoms of toxicity may begin to manifest with serum levels of greater than 1.2 mEq/L and are common with levels greater than 1.4 mEq/L. Elderly or debilitated patients may show signs of toxicity with levels less than 1.2 mEq/L. With high levels (e.g., >2 mEq/L), consultation with a nephrologist or internist often occurs. In patients with high serum lithium levels, dialysis may be necessary to remove the lithium. Symptoms of toxicity include tremors, sedation, and confusion. At higher levels, delirium, seizures, and coma may occur.

Steady-state serum lithium levels are achieved after five half-lives, and lithium levels typically reach steady state 5 days after a change in dosage in patients with normal renal function. Levels are obtained as close to 12 hours after the last dose as possible, and are usually measured using samples obtained in the morning before administration of the morning dose (trough level). In cases of suspected lithium toxicity, levels are obtained sooner.

Maintenance levels are checked every 3 to 6 months and sometimes more frequently in patients with compromised renal function, general debilitation, or signs of toxicity. Close monitoring is required for patients who take medications that may alter lithium levels, such as diuretics or certain nonsteroidal anti-inflammatory drugs (NSAIDs).

Baseline pretreatment studies and periodic follow-up tests for lithium should include monitoring of renal and thyroid function. Renal function is monitored using BUN, serum Cr, and a urinalysis. Because lithium commonly affects thyroid function, a baseline screening assessment of thyroid function (TSH) is obtained. These tests are repeated once or twice during the first 6 months of treatment, and subsequently every 6 months to every year during ongoing treatment. Clinicians also should obtain thyroid studies in patients taking lithium who develop fatigue or asthenia, cold intolerance, or changes in skin, voice, or weight (symptoms of impaired thyroid function). Many clinicians measure calcium levels at baseline and every 6 months to 1 year, given the association of lithium treatment with parathyroid dysfunction and hypercalcemia. Also, because lithium commonly produces a mild leukocytosis, a pretreatment WBC is often

obtained.

A 48-year-old man with schizoaffective disorder had been stably managed with lithium and risperidone for the past 3 years. He was recently medically hospitalized for evaluation of chest pain. The primary team requested a hematology/oncology evaluation, due to concern over his persistently elevated WBC ranging from 12,000 to 16,000/mm^3. Upon medical record review, the consulting psychiatrist noted that prior to initiation of lithium, his WBC had been within normal limits. The team was advised of this common side effect of lithium, for their consideration in formulating the differential diagnosis.

Before the administration of lithium in women of childbearing age, a screening urine pregnancy test generally should be obtained. Use of lithium during the first trimester of pregnancy requires a careful risk/benefit analysis, as there is an increased risk of congenital heart disease with fetal exposure to lithium. Lithium levels should be carefully monitored during pregnancy and after labor. Rapid shifts in fluid volume can give rise to rapid changes in lithium levels during this time.

In patients over age 40, or with known risk factors for cardiovascular disease, a pretreatment ECG is recommended. Many clinicians monitor ECG on a yearly basis; guidelines vary.

VALPROATE. Valproate (valproic acid [Depakene] and divalproex [Depakote]) are used to treat acute mania and for the prophylactic treatment of bipolar disorder. Therapeutic levels for acute mania are thought to range from 50 to 125 mcg/mL; based upon analysis of clinical trial data, most clinicians aim for a valproate level between 85 and 125 mcg/mL for acute mania. During maintenance treatment, clinicians work individually with patients to achieve a dose that is within a therapeutic range but remains tolerable. Side effects are more common with levels greater than 100 mcg/mL.

Because of the risk of hepatotoxicity and hematological abnormalities during valproate treatment, baseline LFTs and hematological measures are advised. Irreversible hepatic failure and agranulocytosis are rare, idiosyncratic effects of valproate; patients should be advised to report immediately for evaluation should they develop concerning symptoms. Hepatic necrosis is rare and is more likely to occur in the young (younger than 2 years of age) and in individuals taking multiple anticonvulsants. More commonly, valproate may cause asymptomatic elevation in liver transaminase levels and mild leukopenia or thrombocytopenia. Most clinicians monitor CBC and LFTs every 6 to 12 months during valproate treatment.

Valproate increases the risk of birth defects, including neural tube defects (with rates of 3 to 5 percent), craniofacial abnormalities, limb malformations, and cardiac defects. A pretreatment urine pregnancy test should be obtained in women of childbearing years. Women should be

cautioned to use adequate contraception. Careful consideration of the known teratogenic effects of mood stabilizing agents is essential when initiating treatment for a woman of childbearing age, regardless of whether or not she is currently pregnant. Given teratogenicity, many clinicians avoid prescribing valproate to any woman who has childbearing potential.

Treatment with valproate may increase serum ammonia levels, even in the absence of LFT abnormalities. It is prudent to obtain an ammonia level in a patient undergoing valproate treatment who presents with altered mental status or lethargy. Acute pancreatitis may also occur, but is rare.

Antidepressants

MONOAMINE OXIDASE INHIBITORS. Treatment with monoamine oxidase inhibitors (MAOIs) can cause orthostasis and, rarely, hypertensive crisis as a result of drug–drug or drug–food interactions. Baseline blood pressure measurement should be obtained before the initiation of treatment, and blood pressure should be monitored during treatment as clinically indicated. There are no meaningful blood levels for MAOIs, and direct monitoring of MAOI blood levels is not clinically indicated.

TRICYCLIC AND TETRACYCLIC ANTIDEPRESSANTS. Tricyclic or tetracyclic antidepressants (TCAs) can affect cardiac conduction. Thus, patients should be screened for a history of cardiac disease or cardiac risk factors that may lead to conduction abnormalities, prior to TCA administration. Patients with risk factors for cardiac disease or who are over the age of 40 should have a baseline ECG to rule out existing conduction abnormalities. Heart rate may increase as a result of anticholinergic effects of the medication. In an overdose, a QRS interval of longer than 100 milliseconds may predict development of seizures and ventricular arrhythmias, and is an indication for bicarbonate therapy.

Baseline vital signs with orthostatic measurements should be obtained before the initiation of TCAs, especially in the elderly in whom TCA treatment can significantly increase the risk of hip fracture owing to TCA side effects of orthostasis. Subsequent orthostasis measurements should be obtained to assist with dose titration in this population.

Plasma levels of TCAs may be useful when there is a need to rapidly achieve a therapeutic dose of the medication, when patients need the lowest effective dose (e.g., sensitivity to side effects), when questions exist about a patient's compliance, and when there is a poor response to a standard dose. Other indications for obtaining plasma TCA levels include concomitant administration of medications that may alter TCA blood levels or the reemergence of depression in a patient who had an initial therapeutic response to TCA treatment.

TCA levels usually are obtained 12 hours after the administration of a dose. Typically, this is in the morning before administration of the morning

dose. Levels are most meaningful when steady state has been achieved, at least five half-lives after the last change in dose. Meaningful serum levels have been determined only for nortriptyline (Aventyl) (50 to 150 ng/mL), desipramine (Norpramin) (>125 ng/mL), and imipramine (Tofranil) (total of imipramine and desmethylimipramine [desipramine] >200 ng/mL).

SENSITIVITY, SPECIFICITY, AND PREDICTIVE VALUE

Not every patient with a particular illness manifests all symptoms, including all laboratory abnormalities. Moreover, some patients may manifest laboratory abnormalities associated with a specific illness without having the illness. Clearly, the presence or absence of a specific laboratory abnormality does not necessarily indicate whether a patient has a particular illness. It is useful to think in terms of the probability of a patient having a particular illness and its associated laboratory abnormalities.

Sensitivity

Sensitivity is the percentage of positive test results found in patients who have a particular disease. In essence, it is the probability of a test being positive in a patient who has the disease. A perfectly sensitive test detects 100 percent of patients who have the illness. The sensitivity can be modified by defining the population that has the illness (e.g., sensitivity in inpatients as opposed to all patients with the disease).

Specificity

Specificity is the percentage of negative results in patients who do not have the disease in question. In essence, it is the probability that an individual who does not have the illness will have a negative test result. A perfectly specific test is negative in 100 percent of the individuals who do not have the illness. Specificity can also be influenced by the population being studied (e.g., carefully screened disease-free individuals as opposed to a random sample of the general population).

Predictive Value

The positive or negative predictive value is the proportion of results that are accurate. Positive predictive value is the probability that a patient with a positive test result actually has the illness, and negative predictive value is the probability that a patient with a negative result does not have the illness. A test with a positive predictive value of 100 percent is only positive in those who have the illness. The positive predictive value helps assess the probability that a patient with a positive result actually has the illness. The negative predictive value helps assess the probability that a patient with a negative finding does not have the disease.

REFERENCES

Allen MH, Hirschfeld RM, Wozniak PJ, Baker JD, Bowden CL. Linear relationship of valproate serum concentration to response and optimal serum levels for acute mania. *Am J Psychiatry*. 2006;163(2):272–275.

*American Diabetes Association, American Psychiatric Association, American Association of Clinical Endocrinologists, North American Association for the Study of Obesity. Consensus Development Conference on Antipsychotic Drugs and Obesity and Diabetes. *J Clin Psychiatry*. 2004;65:267.

*American Psychiatric Association. *APA Practice Guidelines for the Treatment of Psychiatric Disorders: Compendium 2006*. Arlington, VA: American Psychiatric Association; 2006.

Barnes TR, Paton C, Cavanagh M-R, Hancock E, Taylor DM. A UK audit screening for the metabolic side effects of antipsychotics in community patients. *Schizophr Bull*. 2007;33:1397.

*Beach SR, Celano CM, Noseworthy PA, Januzzi JL, Huffman JC. QTc prolongation, torsades de pointes, and psychotropic medications. *Psychosomatics*. 2013;54(1):1–13.

Branson BM, Handsfield HH, Lampe MA, et al. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. *MMWR Recomm Rep*. 2006;55(RR-14):1–17; quiz CE1–4.

Byerly M, Suppes T, Tran Q, Baker RA. Clinical implications of antipsychotic-induced hyperprolactinemia in patients with schizophrenia spectrum or bipolar spectrum disorders: recent developments and current perspectives. *J Clin Psychopharmacol*. 2007;27:639.

Carr RB, Shrewsbury K. Hyperammonemia due to valproic acid in the psychiatric setting. *Am J Psychiatry*. 2007;164:1020.

CDC. Testing for HCV infection: an update of guidance for clinicians and laboratorians. *Morbidity and Mortality Weekly Report (MMWR)*. 2013;62(18):362–365.

Clifford DB, Ances BM. HIV-associated neurocognitive disorder. *Lancet Infect Dis*. 2013;13(11):976–986.

Crismon ML, Argo TR, Bendele SD, Suppes TS. *Texas Medication Algorithm Procedural Manual: Bipolar Disorder Algorithms*. Texas Department of State Health Services; 2007.

Goldberg RW, Kreyenbuhl JA, Medoff DR, Dickerson FB, Wohlheiter K. Quality of diabetes care among adults with serious mental illness. *Psychiatr Serv*. 2007;58:536.

Hannuksela M, Liisanantti MK, Nissinen AET, Savolainen MJ. Biochemical markers of alcoholism. *Clin Chem Lab Med*. 2007;45:953.

Kayser MS, Dalmau J. The emerging link between autoimmune disorders and neuropsychiatric disease. *J Neuropsychiatry Clin Neurosci*. 2011;23(1):90–97.

LeFevre ML. Screening for hepatitis B virus infection in nonpregnant adolescents and adults: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2014;161(1):58–66.

*Levenson JL, ed. *Textbook of Psychosomatic Medicine: Psychiatric Care of the Medically Ill*. 2nd ed. Arlington, VA: American Psychiatric Publishing, Inc.; 2011.

*Marder SR, Essock SM, Miller AL, Buchanan RW, Casey DE. Physical health monitoring of patients with schizophrenia. *Am J Psychiatry*. 2004;161:1334.

*National Institute for Health and Care Excellence 2014. *Bipolar Disorder: The Assessment and*

Management of Bipolar Disorder in Adults, Children and Young People in Primary and Secondary Care. London: National Institute for Health and Care Excellence; 2014.

Newcomer JW, Hennekens CH. Severe mental illness and risk of cardiovascular disease. *JAMA*. 2007;298:1794.

Olfson M, Marcus SC, Corey-Lisle P, Tuomari AV, Hines P. Hyperlipidemia following treatment with antipsychotic medications. *Am J Psychiatry*. 2006;163:1821.

Pfeiffer RF. Wilson's disease. *Semin Neurol*. 2007;27:123.

Smith FA, Wittmann CW, Stern TA. Medical complications of psychiatric treatment. *Crit Care Clin*. 2008;24(4):635–656, vii.

Wu L, Pilowsky DJ, Sclenger WE. Inhalant abuse and dependence among adolescents in the United States. *J Am Acad Child Adolesc Psychiatry*. 2004;43:1206.

▲ 7.7 Principles and Applications of Quantitative Electroencephalography in Psychiatry

LESLIE S. PRICHEP, PH.D., AND E. ROY JOHN, PH.D.

INTRODUCTION

This section is intended to provide a primer for the underlying neurophysiological theory, basic concepts, and nomenclature of quantitative electroencephalography (QEEG) and an overview of clinical applications of QEEG in neuropsychiatry. This chapter reviews the signal processing methods used to extract quantitative features from the raw EEG, converts them to standard scores, displays them as QEEG-based tomographic brain images and uses them as input to classification algorithms, which together reduce the profuse electrical phenomenology of the brain to more objective, anatomically localized, and clinically useful descriptors of abnormal brain function. A clinically oriented section describes statistical methods used to seek distinctive patterns or profiles of QEEG features for patients with specific disorders. This section (1) explains the use of statistical classifier methods for identification of distinctive profiles of abnormality for groups of patients with particular fourth revised edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) diagnoses; (2) gives examples of the performance and clinical utility of QEEG-based classification algorithms; (3) explains cluster analytic methods used to identify pathophysiological heterogeneity (subtypes) within clinically homogeneous cohorts of patients with the same DSM diagnosis; (4) summarizes applications of the performance and clinical utility of QEEG subtyping methods in clinical optimization of treatment decisions; and (5) describes the uses of QEEG source localization in expanding the understanding of underlying pathophysiology of disorders. Finally, research on neurophysiology of anesthesia and of information processing is integrated to review the current understanding of brain mechanisms that bind sensations, emotions, motivation, and memories into perceptual experiences that become the momentary content of conscious awareness, to describe processes that mediate how affect and mood modulate subjective reactions to these experiences, and to explore how dysfunction of these systems can contribute to psychiatric illnesses.

From EEG to QEEG

Background and History. In 1928 the human EEG by the German psychiatrist Hans Berger, who believed that its study would reveal information about psychiatric disorders and cognitive functions, and

revolutionize the practice of psychiatry. However, for over half a century, its major clinical application was in the arena of detection and treatment of seizure disorders. By 1938, Berger's discovery had aroused so much interest among neurologists and psychiatrists that an International Federation of Electroencephalographers agreed upon a standard method for positioning 19 electrodes upon the head, the International 10/20 Electrode Placement System. After extensive neuroanatomical studies, this proportional method locates each electrode approximately above a specified functional region or Brodmann cytoarchitectural area of the brain independent of the absolute size of the head, defining positions as percentages of arcs measured on the head. The locations of this system are shown schematically in Figure 7.7–1.

Contrary to Berger's hopes, insights from the EEG into the underlying causes of psychiatric disorders or how to improve the clinical treatment of psychiatric patients were not readily evident. However, its utility for the detection, diagnosis, and treatment of a wide variety of neurological disorders became rapidly apparent. Using visual examination of oscilloscopic analog recordings of voltage fluctuations recorded from the scalp, generations of electroencephalographers who were focused primarily on neurological illnesses examined tracings and sought to identify distinctive wave shapes and spatial patterns related to their clinical experience. After the introduction around 1970 of computer methods to record and analyze the EEG with readily available digital computers, digital electroencephalography (DEEG) began gradually to replace the conventional analog recording instruments, with great advantages of compact and economic digital media replacing voluminous paper recordings.

QEEG was introduced in the next decade. QEEG augmented visual analysis by computation of the power spectrum to analyze the EEG oscillations from each region on the scalp, replacing tracings by an objective set of numbers that quantified the distribution of power as a function of frequency. During the next 35 years, psychiatrists and psychologists began to study the power spectrum, and the relevance of QEEG to the study of psychiatric and neuropsychiatric disorders, as well as of brain functions and psychophysiological processes, became more and more evident. QEEG measures were expressed as numbers that quantified variables extracted from recordings at each electrode position, such as the absolute amount of power within specified frequency ranges or "bands," the percentage of the total power that falls in each band (relative power), the synchronization or symmetry of such measures between corresponding or "homologous" regions on the two hemispheres, and the relationships of electrical transactions among regions within each hemisphere ("connectivity"). Over the last decade methods derived from signal processing advances and analytic methods (e.g., from genomics, proteomics, and machine learning) were also applied to the EEG, yielding

further feature descriptors and expanding the underlying processes that could be described. QEEG methods have yielded objective data about brain electrical correlates of information processing, cognitive processes and cognitive impairments of aging and in childhood, and alterations of consciousness such as during anesthesia and sleep as well as in psychiatric illness and revealed distinctive QEEG effects of various classes of psychotropic drugs. Such research provided important insights into neurophysiological processes mediating attention, sensation, perception, behavior, and cognition.

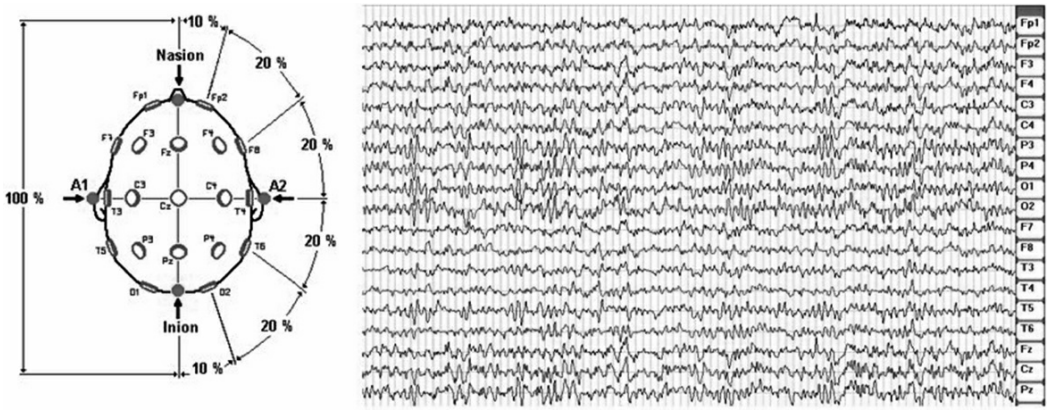

FIGURE 7.7–1. (**Left**) A schematic of the standardized electrode positions of the International 10/20 Electrode Placement System that locates electrodes at distances proportional to 10 or 20 percent of the circumference and several arcs measured from side to side and from nasion to inion on the head, treated as an ellipsoid is shown. (**Right**) Eight seconds of EEG data from the screen of the computer, showing EEG data from each electrode location referenced to linked ears.

Distinctive patterns of QEEG power spectra have been described for a wide variety of developmental, neurological, and psychiatric disorders, akin to the clinical significance of particular abnormal rhythms and patterns in the ECG. Some of these findings have been reported from the relational database of the Brain Research Laboratories of the Department of Psychiatry at the New York University School of Medicine, one of the largest QEEG quantitative electrophysiological databases in the world, which contains “raw” unprocessed EEG and event-related potential (ERP) recordings, extracted features and clinical data from over 15,000 examinations of patients and normal functioning individuals, collected over the past 35 years. This database includes EEG and multimodal ERP recordings from patients with attention-deficit/hyperactivity disorder (ADHD), alcoholism, autism spectrum disorder (ASD), and Asperger syndrome, cognitively impaired elderly patients (with staging of cognitive deterioration ranging from mild to severe in Alzheimer-type dementia, AD), depression, drug dependence, intractable chronic pain, learning disabilities, obsessive-compulsive disorder (OCD), schizophrenia, stroke, tinnitus, and traumatic brain injury. Also included is data from comatose

patients, patient's in the intensive care unit, from patients undergoing many different neurosurgical and cardiovascular surgical procedures using a wide variety of anesthetic agents, and from healthy normally functioning children and adult control subjects. These records are stored digitally in their original, unprocessed form, are readily retrieved for analysis with new mathematical algorithms as they are developed or to test new theoretical hypotheses, and constitute a major scientific resource. Multiple collaborative projects with leading research groups in many countries steadily dynamically increase the scope of this database.

For several of these disorders, profiles of deviant QEEG findings have been found that are either distinctive for patients diagnosed with a particular DSM disorder or predictive of evolution of the condition or differential response to treatment, discussed in more detail below. Certain deviations from this baseline have been found to occur invariantly and reversibly with loss and return of awareness during anesthesia or coma that offer insights into brain mechanisms related to the production and maintenance of consciousness, further discussed later in this section.

Neurometric QEEG: Homeostatic Regulation of the EEG Frequency Spectrum

A major enhancement of the clinical utility of QEEG derives from the well-confirmed discovery and recognition that, much like the ECG, the power spectrum of the EEG of the brain is regulated by a complex physiological system. This power spectrum is extremely stable and highly replicable, distinctive for each brain region in a healthy, normally functioning individual, and independent of cultural, ethnic, or racial factors. Therefore, this predictable, known power spectrum can be used as a normative baseline that specifies the expected electrical activity of the healthy, normally functioning human brain across the human life span. If the procedures described below are followed, then deviations from these norms can be described as standard scores, providing a reliable and objective metric for abnormalities of brain electrical activity. Further, since all features are expressed in the same units, they can be combined and expressed as multivariate descriptors of brain function. Like the ECG, deviations from the normal pattern of electrical activity are indicative of pathophysiological processes and may have diagnostic implications and utility. Particular profiles of deviant standard scores have been found to be reliably correlated with certain categories of brain dysfunction or disorders or with positive as well as negative responses to specific treatments or evolution of illness.

The electrical activity of the brain of a person of any age, alert but at rest in a quiet dimly lit environment, is dynamically regulated by an anatomically extensive, genetically based neurophysiological homeostatic system. The composition of the power spectrum of the EEG of every brain region in healthy, normally functioning individuals is mathematically

predictable across a wide frequency range as a function of age. The interactions between different brain regions, such as synchronization or coherence (nonrandomly phase-locked oscillations) and power symmetry between corresponding locations within and between the two hemispheres, are similarly regulated.

For each of these classes of QEEG features, the distribution of electrical measures within different portions of this range has conventionally been evaluated in so-called "wide-frequency bands," namely, the delta (1.5 to 3.5 Hz), theta (3.5 to 7.5 Hz), alpha-1 (7.5 to 10 Hz), alpha-2 (10 to 12.5 Hz), beta (12.5 to 25 Hz), and gamma (30 to 50 Hz) frequency bands. It is noted that the beta and gamma bands can also be further divided. These seemingly arbitrary definitions of frequency bands, initially made by electroencephalographers on the basis of repeated visual inspection of apparently independent variations of rhythmic oscillations, correspond to the findings of independent sources of variance in the digitized EEG subsequently obtained by computerized principle component analyses. In current QEEG practice with advances in signal processing methodology and instrumentation, there is a steady expansion of the number of electrodes to as many as 256 positions on the head and of the frequency range to examine much faster oscillations as high as 1 kHz (1,000 counts per second) and a shift toward the use of the so-called very narrow band (0.39- or 1.0-Hz divisions) spectral analysis, which has enabled the use of source localization.

Normative values of these various QEEG descriptors across a wide age range (6 to 90 years) have been established and repeatedly independently replicated in multiple countries, indicating that they vary as a function of age but are independent of race or ethnicity. High test-retest reliability has also been published. This regulation reflects activity within and interactions among many brain regions, reflected in local field potentials (LFPs) as well as the spontaneous EEG, mediated by many different chemical neurotransmitters whose synthesis and metabolism is similarly dynamically regulated, and defines the ground state of the brain. In Figure 7.7–2, the evolution of the EEG ground state for relative power (percent of total power at each frequency) at the midline parietal electrode (Pz), is depicted as a three-dimensional developmental surface; an analogous surface has been computed for the ground state of each electrode in the 10/20 system. While not shown here, normative values have now been extended to the first year of life.

The concept of the ground state of the brain being an anatomically extensive, genetically based neurophysiological homeostatic system is consistent with current hypotheses of the default mode network (DMN). The default network or DMN can be defined as the "brain system" which is active when individuals are not focused on or receiving input from the external environment. The scientific literature describes the anatomy of the

DMN as involving a distributed set of regions, including in particular, the medial frontal regions, the inferior parietal lobule, the hippocampus, the medial temporal lobe, and the posterior cingulate cortex. The interaction between the regions involved in the DMN have been described as hubs, each including interactions between subsystems which are highly correlated with each other and weakly correlated with the others. Further, just as deviations from the age-expected normal values of the ground state have been demonstrated to be characteristic of different psychiatric disorders, the DMN has been postulated to be abnormal in many psychiatric states, including autism, schizophrenia, ADHD, PTS, and AD, all of which can be thought of conditions what as associated with cognitive dysfunction in domains linked to the DMN, supporting the link between the ground state and DMN.

Steps in Neurometric QEEG Analysis

QEEG Features. "Neurometric" QEEG analysis refers to the computer extraction from the visually or computer-edited, artifact-free EEG, recorded from 19 scalp electrodes located by the 10/20 system, of a large number of quantitative "univariate" features that assess different attributes of the electrical activity of the brain, and compared to age expected normal values. The steps presented in Figure 7.7–3 describe in detail the two main steps in neurometric analysis. The first step, seen in Panel A of the figure, shows how all variables are transformed to obtain more Gaussian distributions using $\log_{10}$ transforms. The ability of these transforms to improve Gaussianity of the distributions has been confirmed. Some investigators have more recently suggested transforms which may further improve the Gaussianity of the features, but such changes do not result in a significant change in feature values.

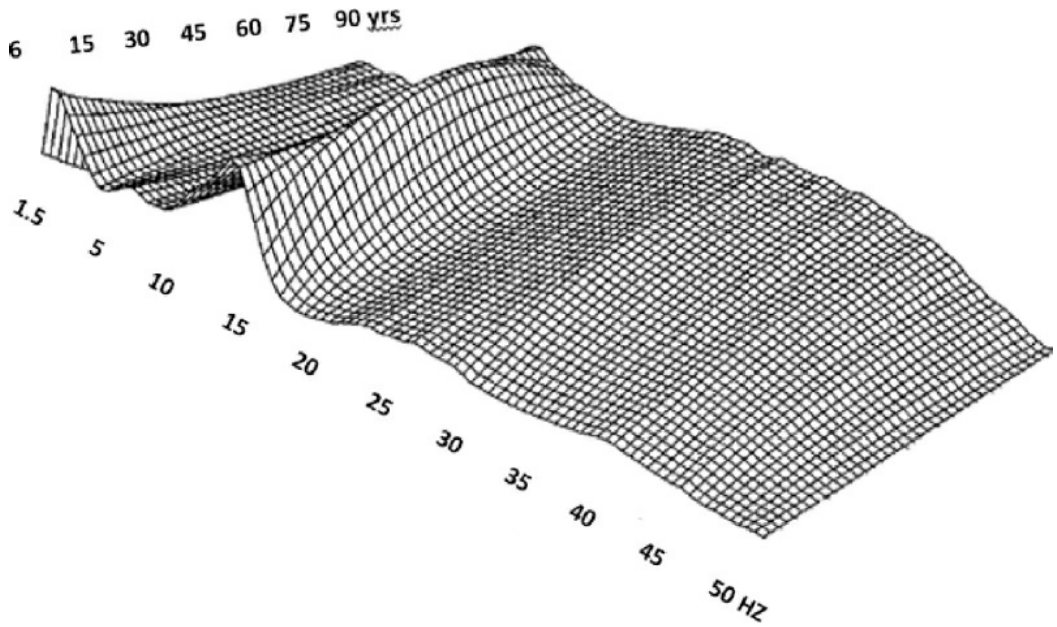

FIGURE 7.7–2. A schematic of the normative (n = 456) EEG "relative power" ground state of the brain, for the 10/20 electrode at the midline parietal region (Pz) location, represented as a three-dimensional surface: *x axis,* frequencies in 0.39-Hz intervals from 1.5 to 50 Hz; *y axis,* age shown in 1-year intervals from 6 to 90 years; and *z axis,* magnitude of percent power of the EEG at each frequency, 0 to 100 percent, represented by the height of the surface. Each "slice" of the image from left to right side of the image is the normal frequency composition of the EEG for that age. As you go from front to back of the image you are viewing the composition of frequency stacked across all ages shown. Note also that while shown here in 1-year intervals, the actual normative regression equations are for continuous age, so that each individual is compared with norms for their exact age.

The second step is the computation of the Z score for the difference between predicted normative value and the value obtained in the patient, which permits the estimation of the probability that such a value might be obtained by chance from a healthy peer the exact age of the patient. As can be seen in Panel B of figure, the mean and standard deviation of the normal population for the age of the patient is used to derive the Z score of the patient. The use of Z scores for all features validates the use of parametric statistical methods for further analysis. This method is standardly used by clinical laboratories to report findings from blood tests.

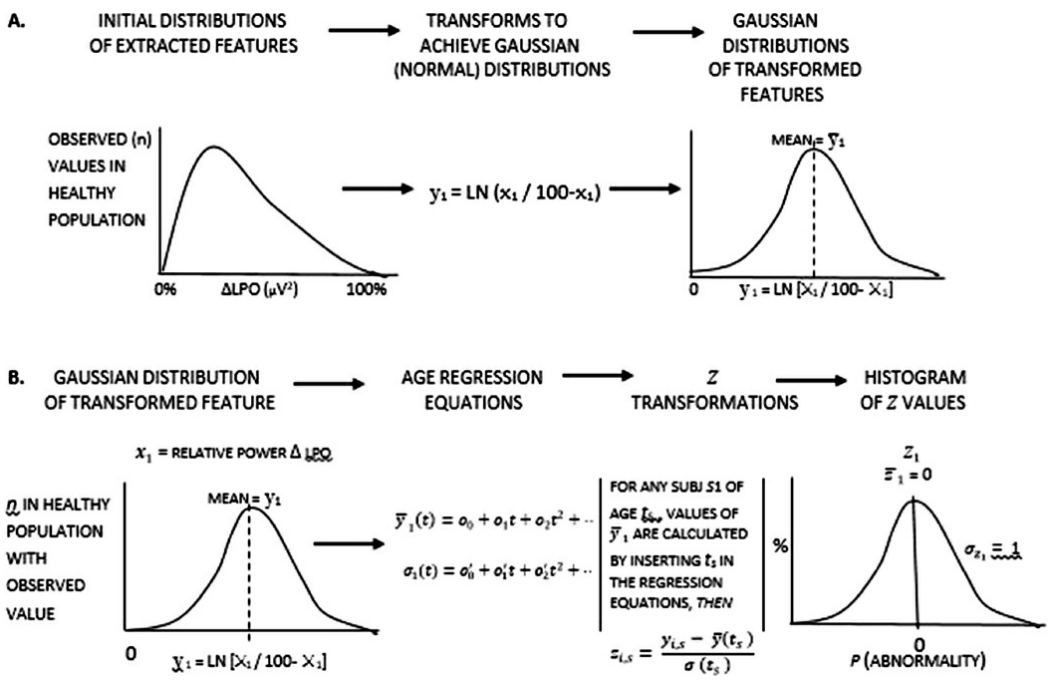

FIGURE 7.7–3. Steps in "neurometric" processing of univariate EEG features to multivariate QEEG: (Step A, top panel) transformations to achieve Gaussianity (normalization) of the distributions; (Step B, bottom panel) age regression and Z transformation to obtain univariate standard scores. Step C, not shown in this figure is the combination of selected univariate Z scores into multivariate estimators (Mahalanobis distances).

Conversion of all variables to the common metric of probability enables multivariate combination of variables, with initially different physical dimensions (e.g., uV for absolute power, percentage for relative power, and

correlation for coherence), into a multivariate distance or brain state vector (BSV), to obtain a quantitative estimate of the overall severity of dysfunction within a region, a set of regions, or between regions within a measure. One such multivariate method, used in Neurometrics, is the Mahalanobis distance. Computation of the Mahalanobis distance, removes the contribution of intercorrelations among the set, providing a corrected estimate of the overall abnormality of the combined set. Such multivariate descriptors can be thought of as characteristics of the system.

A particularly informative use of the Mahalanobis distances is in developmental disorders, in which deviations can be characterized as "maturational lags" or "developmental deviations." Maturational lags are those multivariate patterns of brain function which would be normal if the child were a younger age, that is, such activity does occur in normal brain function, but is not age appropriate. On the other hand, developmental deviations are multivariate patterns which would not be considered normal regardless of the age of the individual. Such characterization of dysfunction can have significant impact on the prognosis and treatment of the individual. Unpublished findings using this characterization of abnormalities, suggested that those who were responded to pharmacological intervention were more likely to have been described as having developmental deviations.

Reporting Results. In analyses of individuals, extracted features include an extensive set of features, each expressed as deviations relative to the age-appropriate normative values. These results can be reported in numerical tables, with highlighting of statistically significant results. For each of the 19 monopolar derivations, with the steps described above, the following measure sets are derived from the frequency spectra: the absolute power (uV2), and relative power (percent) and mean frequency (the point above and below which half of the power falls in a band) are computed for the delta (1.5 to 3.5 Hz), theta (3.5 to 7.5 Hz), alpha (7.5 to 12.5 Hz), beta (12.5 to 25 Hz), and gamma (30 to 50 Hz) frequency bands and the total spectral power. In addition, relationships between and within hemispheres are derived, including measures of coherence (synchronization) and asymmetry in all frequency bands are computed for all electrode pairs in corresponding (homologous) positions on the two hemispheres. These same measures are computed for bipolar derivations of each electrode versus every other electrode within each hemisphere (connectivity).

Interpolated topographic *Z*-score images can depict the statistical significance of local deviations relative to age-expected normal values, for all univariates in each measure set as well as for a selected set of multivariate measures. Such maps are color coded using a palette in which different hues represent the probability of the findings at each brain region. Because each *Z* value has a known statistical probability associated with it

(e.g., $Z = \pm 1.96$ has $p < 0.05$), the results of such analyses are objective as well as self-interpreting. Figure 7.7–4 shows a sample of group average QEEG Z images (nose up, left hemisphere on the left) for different populations of psychiatric patients in each row. Each column represents a selected measure set. Significant local deviations from age-expected values can be seen as colored regions on the maps. In the color coding for significance in this figure, the center of the scale, black, indicates no deviations from expected normative values, moving up the scale through red to orange and yellow goes toward significant excess of the measure and down through blue to light green indicates significant deficit of the measure. Looking at the top row on the left of this figure, all maps appear dark to black, showing no deviations from age-expected values in this large group of normal adults. On the other hand, in all other rows, significant excesses and deficits can be seen in all groups, and differences in the patterns of these deviations are also evident between the groups.

It is interesting to note that the progressive magnitude of abnormalities can be seen when comparing the images from the mild cognitive impairment (MCI) and Alzheimer disease (AD) groups (especially in relative power in theta, relative power in beta, and mean frequency of the total spectra, columns 2, 3, and 5 of the right panel). Also noted are the similarities between the schizophrenic patients off medication for 90 days and those on medication at the time of the test (bottom two rows in the left panel), suggesting the presence of the medication. The presence of distinctive patterns of QEEG abnormalities and differences between groups will be further discussed in more detail below.

Measures of connectivity have been increasingly demonstrated to be of importance in the description of brain dysfunction in traumatic brain injury (TBI), postconcussive syndrome, PTSD, autism, and AD. Figure 7.7–5 shows an example of the images for coherence connectivity measures in the theta frequency band in a 64-year-old with early signs of cognitive decline. Each electrode location shows the results for the coherence between the labeled region and each other regions, in the labeled band (Delta, Theta, Alpha, Beta, Beta2). The significance of the hypo- or hypercoherence is color coded.

In addition to the conventional measures described above, measures of complexity can also be computed, including chaotic/fractal measures, as well as information theory–based measures of entropy and measures of mutual information. These measures have been shown to be particularly important in the study of anesthetic affects (loss and return of consciousness), traumatic brain injury, and PTSD and expand the measure of disturbances in the EEG signal to include nonlinear measures beyond the frequency domain.

Homeostatic System Generating and Regulating the EEG

The neuroanatomy and neurochemistry of the homeostatic system generating and regulating the EEG is schematized in Figure 7.7–6. The alpha rhythm arises from pacemaker neurons in the thalamus, which have an intrinsic oscillatory rhythm in the EEG alpha frequency range (7.5 to 12.5 Hz) and synchronize similar activity from neurons dispersed throughout the neocortex. When "instructed" by the cortex, the nucleus reticularis (encapsulating much of the thalamus like the skin of a pear) exerts differential GABAergic influences that hyperpolarize these pacemakers, slowing their frequency down to as low as 5 or 6 Hz (cycles per second) to enter the theta range.

When arousing afferent input reaches the midbrain reticular formation and specific sensory thalamocortical relay (TCR) nuclei of the thalamus, cholinergic outflow from the ascending reticular activating system (ARAS) is sent to the nonspecific nuclei of the thalamic diffuse projection system and to the nucleus reticularis, antagonizing the inhibitory modulation of γ-aminobutyric acid (GABA), depolarizing pacemaker neurons, and allowing them to accelerate their slowed oscillations even into the 12 to 15 Hz range, usually referred to as beta. These interactions can serve to focus attention by modulating excitability of TCR gates.

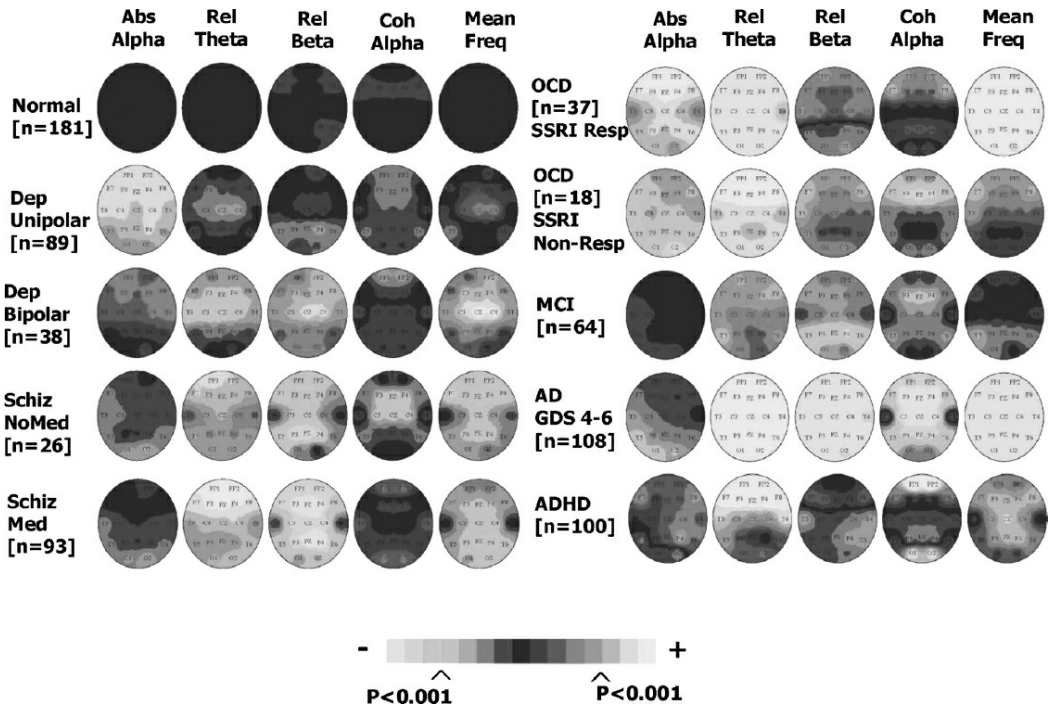

FIGURE 7.7–4. Group average topographic images of selected neurometric variables (columns) across different clinical populations (rows). Columns, left to right: (1) absolute power (microvolts squared) in the alpha frequency band; (2) relative power (percent) in the theta band; (3) relative power (percent) in the beta band; (4) interhemispheric coherence (percent synchronization) in the alpha band; and (5) mean frequency of total spectra of the EEG (Hz). (**Left**) Row 1, normal individuals not used to construct the norms (*n* = 181); Row 2, unipolar patients in depressed state (*n* = 89); Row 3, bipolar patients in depressed state (*n* = 38); Row 4, chronic schizophrenic patients, nonmedicated >3 months (*n* = 26); Row 5, chronic schizophrenic patients, currently on medication (*n* = 93). (**Right**) Row 1, drug-free

baseline data in obsessive-compulsive disorder (OCD) patients who go on to be selective serotonin reuptake inhibitor (SSRI) responders (*n* = 37); Row 2, drug-free baseline data in OCD patients who go on to be SSRI nonresponders (*n* = 18); Row 3, mild cognitively impaired (MCI) elderly patients with a GDS staging of 3 (*n* = 64); Row 4, Alzheimer disease (AD), with GDS staging of 4 to 6, (*n* = 108); Row 5, children with attention-deficit/hyperactivity disorder (ADHD) (*n* = 100). The color scale is proportional to the significance of the deviation from age-expected normal values, with black, no departure; blue to light green, deficits; and red to yellow, excesses. Statistical significance of entry for any group can be estimated by multiplying the value of the indicated hue on the *Z*-score color bar by the square root of the group *N*.

Oscillations in the beta range (12.5 to 25 Hz) and in the gamma range (30 to 50 Hz) arise in the EEG to reflect thalamocortical and corticocortical transactions concerned with information processing and consciousness. Delta rhythms (0.5 to 3.5 Hz) arise from neurons in the cortex and often reflect functional deafferentation or neuropathological influences such as ischemia or increased intracranial pressure but can be depressed by certain pharmaceutical agents or substance abuse. Hypothetical functional roles and processes attributed to regions in this diagram will be discussed further in the text.

Additional EEG rhythms arise from the activation of the dorsal striatum (sensorimotor integration) and ventral striatum and amygdala (emotional state and affect), interacting with major components of the limbic system for storage and retrieval of memories. Activation of this system results in outflow, via the nucleus medialis dorsalis of the thalamus and the anterior and posterior cingulate cortices, of rhythmic theta waves (3.5 to 7.5 Hz) carrying phase-locked gamma oscillations (30 to 50 Hz) reflecting retrieval of memories via the hippocampus. This is discussed in further detail later in this section.

Brain State Vectors

Rescaling all neurometric features as *Z* scores, providing a common dimension proportional to probability, allows any arbitrary set of QEEG measures to be compressed into a multivariate estimate of the total overall abnormality across the set. With one such method, the Mahalanobis distance (see above), the contribution of intercorrelations among the set is removed, providing a corrected estimate of the overall abnormality of the combined set.

With Mahalanobis distances, results of a QEEG evaluation can be represented as a BSV in a brain signal space. The length of the BSV corresponds to the severity of the deviation of the patient's brain electrical activity from the origin defined by age-appropriate normative values, and its orientation in the physiological space represents the quality or nature of the pathophysiology. By thus representation of the BSVs of patients with different diagnoses and normal individuals as points in the signal space, two questions relevant to diagnosis can be answered quantitatively and will

be addressed later in this section: (1) Can stepwise classifier functions (such as discriminant analysis) be defined that separate patients with different diagnoses from each other or from normal individuals? (2) Can cluster analytical methods reveal subtypes with distinctive QEEG profiles but with the same clinical diagnosis, unmasking pathophysiological heterogeneity in that diagnostic category?

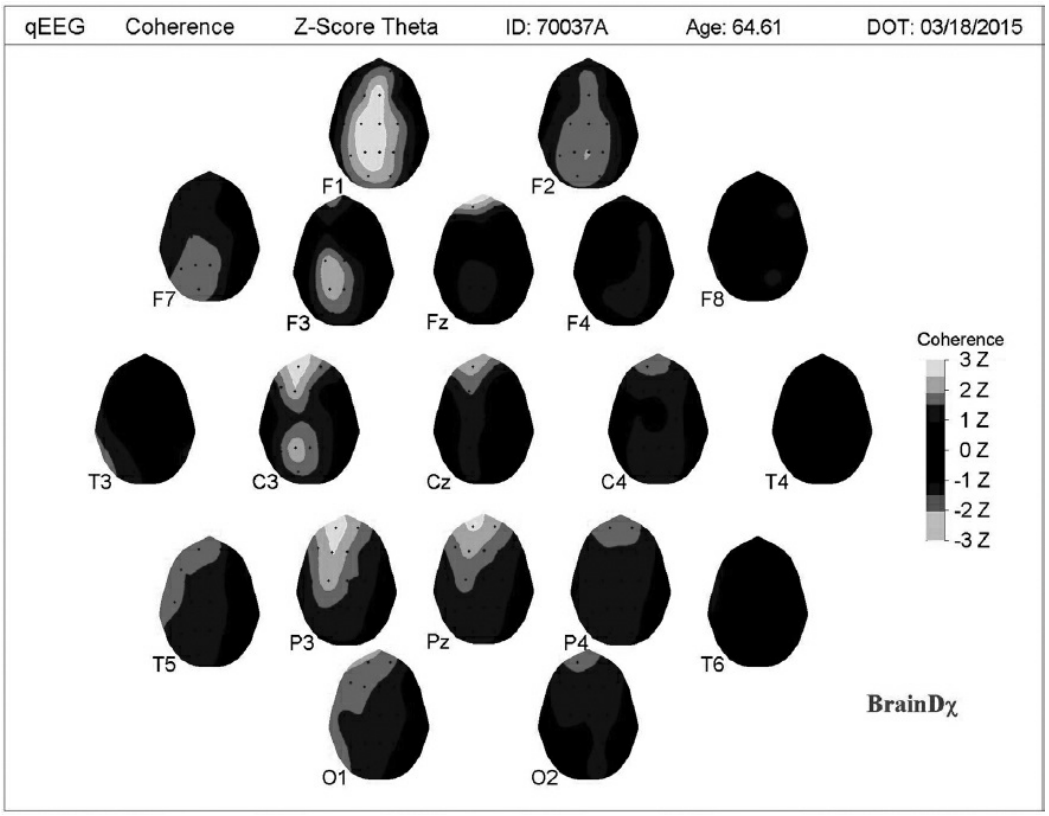

FIGURE 7.7–5. Each component image (individual head map with nose up and left on the left) shows the results for the coherence between the labeled region and all other regions, in the labeled band (in this case the theta frequency band), color coded by significance level using the z-scale shown. Regions in black are within normal limits, all colored regions are significant at $p < 0.05$ with the extremes at $p < 0.001$. The top left image shows the connectivity between the left frontopolar regions (FP1) and all other regions in the theta band, a band which has been shown in the literature to be correlated with degree of cognitive decline in the elderly. It can be seen that there is a significant hypercoherence between this regions and both short and long distance relationships on the left hemisphere and midline regions. This hypercoherence is also reflected in the coherence between FP1 and the left central (C3), left parietal (P3) and midline parietal (Pz) regions. This hypercoherence between frontal and more distant regions in a patient with early cognitive decline might serve as a baseline for a trajectory of increasing abnormality when tracked longitudinally in this patient. (See color atlas.)

QEEG Source Localization Tomographic Brain Images

A powerful expansion of the contribution of QEEG to the understanding of the underlying pathophysiology of disorders and the clinical utility of QEEG has been achieved over the past decade by the development and application of source localization analytics. QEEG low-resolution source

localization methods perform computation of an "inverse solution" to localize the mathematically most probable intracranial sources of the scalp-recorded EEG signal. With such methods voltages or *Z*-scores of QEEG measurements from the scalp can be imaged, providing three-dimensional brain images that have been demonstrated to be coregistered with other neuroimaging methods. The standardized electrode positions of the proportional 10/20 system are placed in geometrical registration with a proportional Probabilistic MRI Brain Atlas, constructed at Montreal Neurological Institute based upon the average MRI anatomy of 300 normal brains. Referring the computed sources to an average MRI Atlas averts the need to obtain an MRI of the individual patient. The positions of sources at any EEG frequency are depicted upon images of transaxial, sagittal, or coronal slices from the MRI Atlas, with each voxel color coded in hues that quantify the deviation in each voxel of the computed source from the expected age-appropriate normative value, as in neurometric QEEG. There are two source localization methods that implement voxel norming, variable resolution electromagnetic tomographic analysis (VARETA) and low-resolution electromagnetic tomographic analysis (sLORETA and eLORETA); examples of both will be given below (see Figs. 7.7–7, 7.7–10 through 7.7–12). The anatomical accuracy of these brain images computed from the QEEG has been repeatedly confirmed by coregistration studies using MRI, fMRI, PET, and SPECT.

Neurometric QEEG source localization provides a unique and revolutionary technique of functional brain mapping, since no other functional imaging method, neither fMRI, PET, nor SPECT, provides access to age-appropriate normative voxel data and only electromagnetic data (EEG or magnetoencephalopgraphy [MEG]) has the time resolution in milliseconds, allowing a more precise identification of neuronal processes. This method provides the further advantage of the ability to obtain accurate functional brain images with economical and portable equipment which can be used at the point of care without the need for movement restraint within the equipment, camera or imposition of a burden of radiation. It should be noted that advances in this method allows real-time imaging of sources, which has many interesting possibilities for unique ways to study psychiatric disorders.

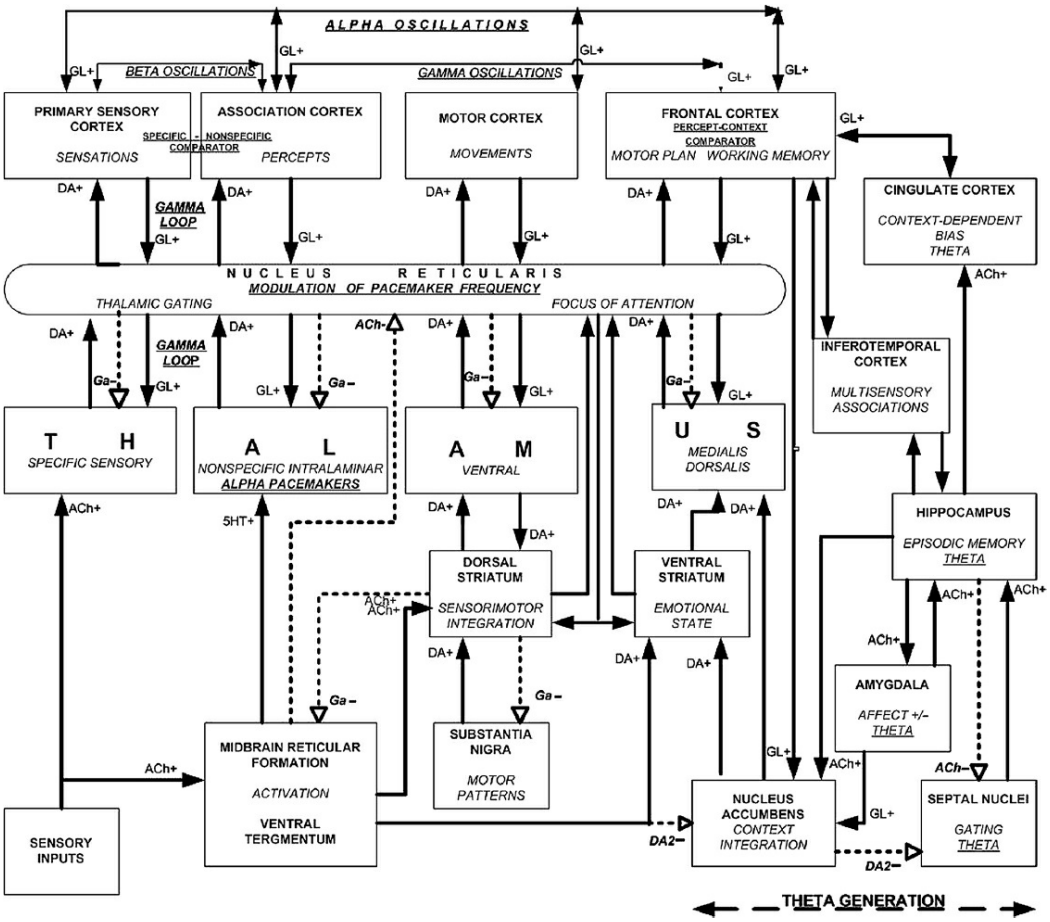

FIGURE 7.7–6. Circuits serving a major role as generators of indicated electroencephalogram (EEG) frequency bands, interrelationships, and mediating neurotransmitters proposed to comprise the homeostatic system regulating the EEG ground state of the brain. Salient hypothesized functional roles have been assigned to regions and putative transmitters assigned to relationships indicated by *arrows:* ACh, acetylcholine; DA/DA2, dopamine; Ga, γ-aminobutyric acid; Gl, glutamine; 5HT, serotonin; +, excitatory actions; –, inhibitory actions. A more complete explanation of this figure can be found in the text. In brief, pacemaker neurons in the thalamus generate an intrinsic oscillatory rhythm in the EEG alpha frequency range (7.5 to 12.5 Hz) propagated to the cortex that can be slowed into the theta range (6 to 7 Hz) by GABAergic influences from the nucleus reticularis, tending to close thalamic gates to the cortex. Incoming sensory information is transmitted by an "exogenous system" via thalamocortical relay (TCR) neurons to the cortex and also enters the midbrain, from which cholinergic arousal influences arise to antagonize γ-aminobutyric acid (GABA) and accelerate the EEG rhythms, opening the thalamic gates, and serving to focus attention. Oscillations in the beta range (12.5 to 25 Hz) and in the gamma range (30 to 50 Hz) reflect thalamo-cortical and corticocortical information processing. Delta rhythms (0.5 to 3.5 Hz) arise from neurons in the cortex and often increase with neuropathology but can be depressed by substance abuse. Afferent input also activates the dorsal striatum (sensorimotor integration) and ventral striatum and amygdala (emotional state and affect), interacting with major components of the limbic system for storage and retrieval of memories. Readout from this "endogenous system" results in outflow, via the nucleus medialis dorsalis of the thalamus and the anterior and posterior cingulate cortices, of rhythmic theta waves (3.5 to 7.5 Hz) carrying phase-locked gamma oscillations (30 to 50 Hz) reflecting retrieval of memories via the hippocampus. Thus, theta rhythms in the EEG can arise from two quite different sources.

Selection of Anatomical Regions of Interest. Since the statistical

significance of deviations from expected normative brain electrical activity can be quantitatively encoded for every voxel within a source image, this input can be used to quantify the significance of deviations for specified anatomical regions of interest (ROIs), allowing the evaluation of the maximum abnormalities in the brain regions of an individual patient. This can be done in a variety of ways, one method maps sets of voxels which are within specific anatomical regions (using an atlas such as Talairach), quantifying and displaying the mean of all voxels in the region or the maximum abnormality in the region. The method the authors have found most informative and representative of the deviation of an ROI computes the mean of the voxels which containing the highest 10 (or 20 percent) of the Z-scores within the set of voxels that define that region. Another way to select ROIs can be done by search algorithms that can be conceptually understood as methods to systematically remove from the three-dimensional representation, all voxels that have an abnormality less than some selected threshold, for example, the $p < 0.05$ level ($Z = \pm 1.96$). The remaining voxels will lie within some three-dimensional region within the brain, which an algorithm automatically identifies from the coregistered corresponding anatomical neurosurgical stereotaxic atlas. In this way, ROIs are defined by the significance of deviations rather than anatomical localization. A numerical estimate of the degree of severity of the abnormality can be calculated, for example, by integrating the Z-scores of all voxels within the volume of the ROI thus defined or counting the number of voxels above the threshold. These methods for quantifying abnormalities could be applied to the evaluation of treatment efficacy, whereby the magnitude of the abnormality before and after treatment is compared. Other methods for statistical analysis of source images include statistical probability mapping (SPM) or statistical nonparametric mapping (SnPM) as used in other brain imaging methods, especially fMRI.

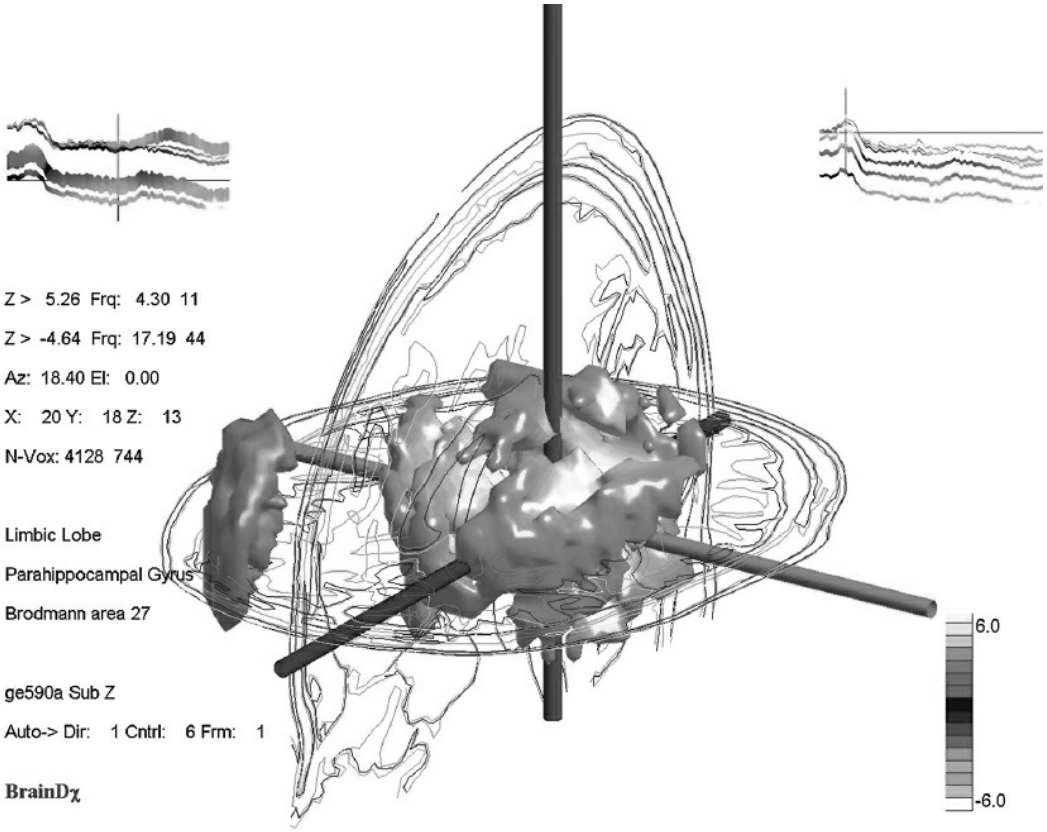

FIGURE 7.7–7. Three-dimensional low-resolution electromagnetic tomographic analysis (LORETA) image of a "dissolved brain" of a patient with AD, after removal of all voxels with a *Z* score less than 3.29 ($p < 0.001$). The maximum theta excess (*dark red shading,* 4.3 Hz) relative to age-expected normal values is in the hippocampal and parahippocampal regions, while the frontal cortex (Brodmann area 27) reveals a minimum beta deficit (*blue,* 17 Hz). (See color atlas.)

Figure 7.7–7 presents an image of the brain of a patient with Alzheimer disease (AD) after all voxels with a *Z*-score less than 3.29 ($p < 0.001$) have been removed from the brain volume. It can be seen in this figure that the area of maximum abnormality relative to age-expected normal voxel values is in the hippocampal and parahippocampal regions.

Brain ROIs identified by neurometric source imaging, scaled for degrees of abnormality quantified by deviation from norms or alteration of function relative to some reference state, can be used as an adjunct to diagnosis, the selection of treatments, evaluation of bioavailability, and pharmacodynamic assessment of efficacy or even in drug development research. ROIs thus identified might then be interrogated by noninvasive assay using MRS. MRS already has the capability to quantify the concentration of a steadily increasing number of brain metabolites, neurotransmitters, and critical ionic electrolytes. Such multidisciplinary assessments might eventually become routine in evidence-based psychiatric practice.

Clinical Applications of QEEG Departures from Ground State or Default Network

Psychopathology: QEEG Correlates (Biomarkers) of Diagnoses. Deviations from normative values imply dysregulation that may have functional and treatment implications. In numerous peer-reviewed articles, there is evidence of the high proportion of psychiatric patients in any diagnostic category, who consistently display significant deviations from the expected normative QEEG values. This is clearly shown in Figure 7.7–4 for large populations of many different diagnostic groups. Many of the articles present evidence that these profiles are sufficiently distinctive not only to indicate the most probable DSM diagnosis for an individual patient but, in some instances also can be used to predict differential response to treatments or different evolution of illness; reviews of the abundant relevant literature are available. It should also be noted that the degree of dysfunction, such as the progressive stages of dementia, can be accurately reflected in the QEEG, demonstrating that increasing QEEG abnormality is correlated well with increasing degree of cognitive impairment. Thus, QEEG can represent a biomarker, be useful as an adjunct to diagnosis and staging, as well as to aid in the optimal selection of treatment. An example of such, the relationship between degree of cognitive decline and magnitude of QEEG abnormalities, can be seen in Figure 7.7–4, comparing MCI (third row on right panel) to AD (fourth row on right panel), especially as seen in increased excess of relative power in the theta band (second column), exaggerated deficit of relative beta activity (third column), and significant slowing of the mean frequency of the spectra (fourth column).

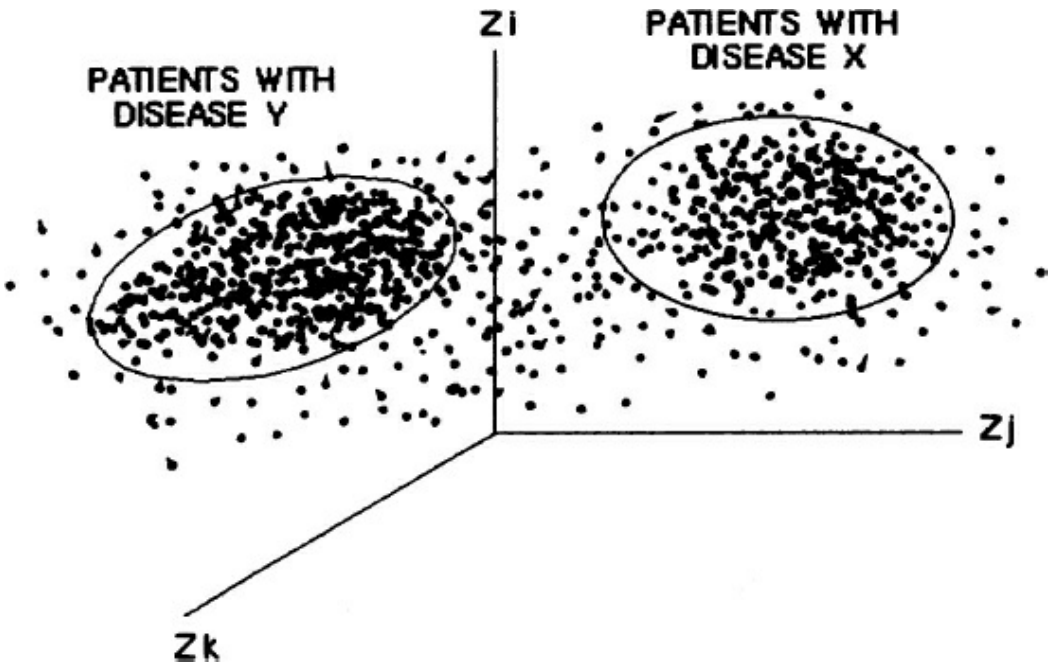

FIGURE 7.7–8. Point swarms representing the hypothetical scatter of brain state features in two different regions of a brain signal space, suggesting that discriminant classification analyses might separate patients diagnosed with disease X from others diagnosed with disease Y.

If the abnormal QEEG features detected in each member of a group of patients with different DSM-diagnosed disorders are represented as BSVs, then they can be envisaged as swarms of points (features) dispersed in a multidimensional brain electrical signal space. To the extent that particular profiles of QEEG abnormalities (sets of features or biomarkers that covary in the disease state) are in fact distinctive for patients with one specific symptomatic, DSM-based disorder, patients with the same diagnosis would be expected to lie close together within such a swarm in a region of the signal space, while patients with different DSM diagnoses would constitute swarms in different regions. This concept is shown schematically for patients with disease X and patients with disease Y in Figure 7.7–8.

Application of statistical segregation methods such as multiple stepwise discriminant functions, logistic regression, neural networks, support vector machine (SVM) and others could then be applied to construct QEEG-based classifiers that would accurately identify individuals with specific psychiatric disorders and distinguish those with one DSM diagnosis from those with a different one or from normal. Since the initial successful application of this method, numerous QEEG discriminant functions have been reported with sensitivity, specificity, and positive and negative predictive values equal or superior to the widespread use of such well-established diagnostic imaging tools as, for example, CT for detection of stroke or traumatic brain injury (TBI).

Subtyping: QEEG Heterogeneity within DSM Diagnoses using QEEG Biomarkers. As is apparent from the highly variable efficacy of a given psychotropic agent in patients with the same DSM diagnosis, patients grouped together as displaying a symptomatically homogeneous psychopathology may suffer from a variety of pathophysiological dysfunctions responsive to different treatments or displaying different evolution of illness. This variable responsivity suggests that individuals within the same swarm of points in a region of electrophysiological signal space, representing first-order similarities among BSVs from a group of patients with the same DSM diagnosis, might actually not be homogeneous. Within the region containing such a swarm or cloud, there might be sectors containing individuals with second-order differences in features that also contribute to their electrophysiological profile or pattern of disorder. These differences may reflect the existence of subtypes of pathophysiology related to biomarkers that are critical for response to a given treatment, in addition to the first-order influences that produce sufficiently similar global symptoms to warrant the same DSM diagnosis, as shown schematically in Figure 7.7–9. Within the two-point swarms, the figure depicts individual members of the distinct subtypes A and B within the cloud diagnosed with disease Y and of subtypes C and D within the other cloud diagnosed with disease X. It is important to note that although referred to as “features” a

feature can be a multivariate composite of more than one feature (e.g., the ratio of two features or a measure of connectivity between regions such as cordance).

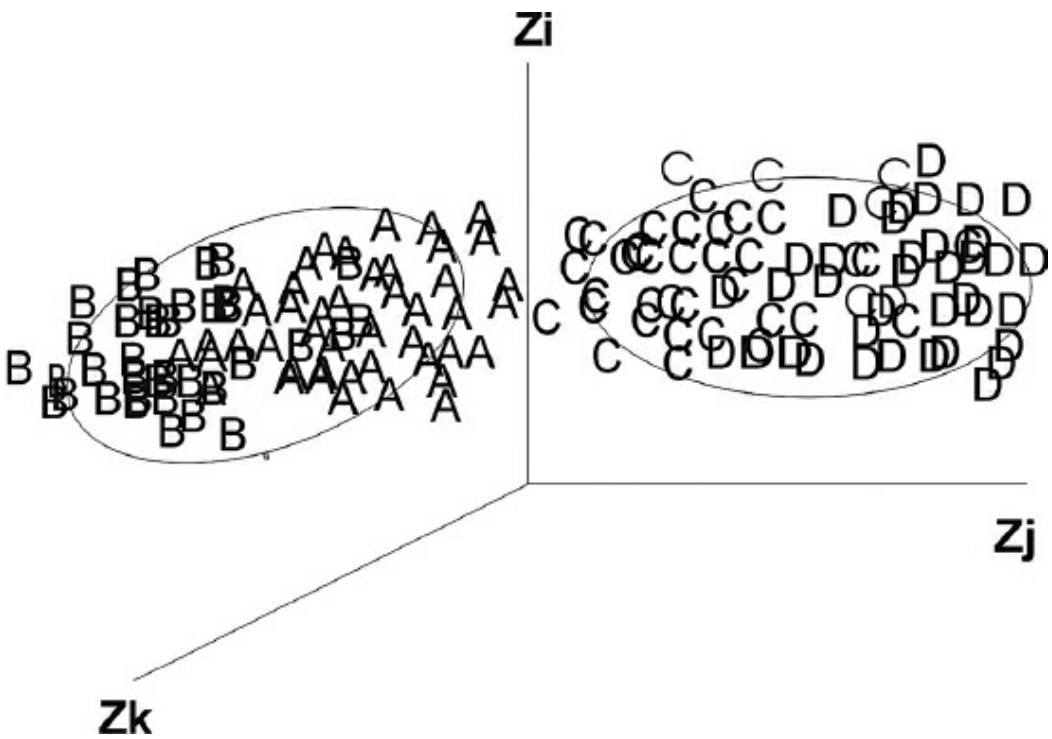

FIGURE 7.7–9. Point swarms representing the hypothetical identification of subtypes of patients in four different regions of the signal space, with different pathophysiological profiles A and B within those diagnosed with disease Y and with different profiles C and D within those diagnosed with disease X.

In most instances, diagnosis is largely determined by clinical evaluations and DSM symptom–based categorization except for cases that present dual-diagnostic problems. Potentially, the most important and useful clinical applications of QEEG lie not just as an adjunct to diagnosis but in identifying subtypes whose members, while relatively homogeneous in terms of displaying similar symptoms and receiving like diagnoses, are statistically probable to have clinically relevant differences (heterogeneous underlying QEEG profile) with respect to response to particular treatments, prognosis of evolution of illness, or prediction of treatment outcome. Investigations using methods of QEEG-based classification or cluster analyses have been useful in unmasking such heterogeneity in a number of disorders, such as: (1) negative as well as positive responsiveness to or selection of stimulant therapy in ADHD; (2) response to selective serotonin reuptake inhibitors (SSRIs) in OCD (see Fig. 7.7–4, right panel, rows 1 vs. 2); (3) probability of relapse in crack-cocaine dependence; (4) pathophysiological profiles of psychotic states or schizophrenia (described in detail below); (5) prediction of future cognitive decline in normal elderly; (6) and response to antidepressants among patients presenting in a state of

depression (see Fig. 7.7–4 left panel, row 2 vs. 3); and (7) the response to nonstimulant or stimulant medication in ADHD patients. Reviews of this literature are available.

QEEG Applications in Attention-Deficit/Hyperactivity Disorder (ADHD), Alzheimer Disease (AD), Chronic Pain, and Psychosis. The scientific literature applying QEEG methods to the study of psychiatric and neurological patients, their diagnosis, underlying pathophysiology, and treatment response is vast and includes numerous studies in ADHD, depression, dementias, drug dependence, learning disabilities, OCD, panic disorder, schizophrenia, traumatic head injury, PTSD, chronic pain, and many other disorders. This subsection provides an overview of the literature in three areas: ADHD, dementia, and chronic pain.

Attention-Deficit/Hyperactivity Disorder. Extensive studies demonstrate the clinical utility of QEEG as an adjunct to the diagnosis, evaluation, and treatment of children, adolescents and young adults with learning and attention problems. Using selected QEEG features (e.g., frontal theta activity or theta/beta ratio, alpha reactivity, and cordance) high accuracy, sensitivity, and specificity, has been achieved in distinguishing normal controls from ADHD/attention-deficit disorder (ADD) children, children with ADHD/ADD from those with learning disabilities, ADHD/ADD patients with positive response to specific medication from those with negative response. Augmenting the QEEG features with behavioral measures further enhanced the sensitivity for predicting long-term treatment response. A review of this expanding literature is available.

We will briefly summarize meta-analyses that research the large number of studies including diagnostic capabilities, subtype identification, and treatment response in the ADHD population. One such meta-analysis of QEEG studies in ADHD (DSM-IV-TR diagnosis) reported on a total of almost 1,500 patients and concluded that increased theta/beta ratios (theta power increase with a beta power decrease) is a characteristic trait of ADHD when compared to normal controls. With this ratio, a sensitivity and specificity of 94 percent was obtained across all studies. Under DSM-5, ADHD has been expanded to more accurately characterize the disorder in adults as well as children and adolescents. Symptoms are still divided into categories of inattention, hyperactivity, and impulsivity. Several studies using QEEG have explored the ability to identify different patterns of abnormalities reflecting these different symptom profiles have reported suggestive results. The existence of such electrophysiological subtypes may be related to different treatment response.

The existence of electrophysiological subtypes within the ADHD population has been further demonstrated in studies applying cluster analysis of QEEG features within a large population of ADHD patients.

Such studies have shown the existence of subtypes within the population with different underlying sources for the different subtypes. While preliminary studies are suggestive of a relationship between subtype membership and treatment response, much work in larger populations remains to be done in this area.

A nonpharmacological treatment of ADHD patients which is growing considerably is QEEG, and QEEG source, neurofeedback (NFB). A meta-analysis including five QEEG NFB studies and a total of approximately 300 patients supports improvement of some aspects of the ADHD symptom profile following NFB. In addition, preliminary studies have indicated the possibility of selecting NFB protocols based on individual QEEG biomarkers to optimize efficacy. More research is needed to support NFB as a first-line, standalone treatment modality and for the identification of the subtypes of ADHD that might most benefit from this treatment modality. An extensive meta-analysis and Bayesian network meta-analysis of the comparative efficacy of pharmacological and nonpharmacological treatments (including psychological, alternative medicine, and other treatments) of ADHD children and adolescents is underway in 2015.

Another example of the use of QEEG in the ADHD population for a better understanding of the underlying pathophysiology includes studies of alpha reactivity. This measure reflects the magnitude of the differences between alpha activity in the eyes closed resting condition and that with eyes open with increased external stimulation. These studies suggest that ADHD patients show less alpha reactivity, supporting abnormalities in arousal mechanisms. Application of cluster analysis to the ADHD population revealed the existence of neurophysiological subtypes and showed a relationship between cluster membership and response to treatment. QEEG source localization (VARETA) applied to individual subtypes demonstrated different underlying pathophysiology. A cluster characterized by excess alpha activity (maximum abnormality at 11 Hz, in the high alpha range) displayed primarily cortical abnormalities, maximal in right parietal cortical regions, whereas VARETA images of a cluster with theta excess (maximum abnormality at 5.4 Hz, in the theta range) revealed primarily temporal cortical and hippocampal abnormalities.

The utility of incorporating QEEG as an adjunct to the evaluation and treatment of ADHD was highlighted in a 2004 report of the American Academy of Pediatrics, which stated that, "… new brain wave analysis techniques like quantitative EEG will help experts more clearly document the neurological and behavioral nature of ADHD, paving the way for better understanding and treatment." Now, a decade later, a growing literature supports this thesis.

Alzheimer Disease (AD). A vast literature spanning more than three decades reports QEEG changes in the pattern of brain electrical activity

associated with aging and documents relationships between specific changes in the QEEG and degree of clinical deterioration in the elderly. Space does not permit a review of these numerous QEEG studies that consistently demonstrate the importance of the following QEEG features related to the severity and progression of dementia: Increased delta and/or theta power, decreased mean frequency of the total spectra, anteriorization of the alpha peak, decreased frequency of the dominant occipital rhythm, and changes in coherence, synchronization, and connectivity patterns. As additional signal-processing methods are applied to the EEG signal, additional features have been proposed as QEEG biomarkers of disease severity in AD, these include: features based on canonical correlation coefficients reflecting synchrony between multiple regions and theta current source density. A meta-analysis of portions of this literature is available.

Sensitivity to the earliest preclinical stage (i.e., the first symptomatic manifestation pre-MCI) of AD, currently described as subjective cognitive impairment (SCI), has also been demonstrated. Multiple publications have used QEEG features to achieve accurate differentiation of early AD from elderly normal controls, discrimination of dementia from depression in the elderly, distinction between early versus late-onset AD, discrimination between AD and vascular dementia and identification of mild dementia and MCI and can increase diagnostic accuracy when used with other imaging methods.

The use of QEEG features observed in initial evaluations for the prediction of cognitive decline of normal elderly patients and those with MCI or early-stage mild dementia to more severe dementia has been demonstrated in multiple studies. In one such longitudinal study the exquisite sensitivity of QEEG at the earliest stage of SCI, normal elderly subjects with only subjective complaints about memory, but no deficits on neurocognitive testing, were followed longitudinally for 7 years. On the basis of only the QEEG data from the initial QEEG evaluation, future cognitive decline was predicted with over 90 percent accuracy based upon subsequent clinical staging. Significant abnormalities in QEEG features were seen in the baseline evaluations of those who went on to decline and were absent in those who showed no change over time. The most salient abnormalities were seen in theta power, mean frequency of theta and of the total spectrum, and intrahemispheric coherence across all bands. With source localization in those who go on to decline, the underlying pathophysiology identified at baseline included excessive theta activity in the hippocampus, amygdala, and temporoparietal cortex, all regions identified by other neuroimaging technology as important in dementia and related to subsequent decline to dementia. Figure 7.7–10 shows the sLORETA source localization images from the initial baseline QEEG data for the abnormal activation seen in the theta frequency band for three

groups (n = 5 each group) of SCI subjects. The top row of this figure show those who upon long-term follow-up were found to show no cognitive decline, the middle row shows those who are found to decline to MCI and the bottom row shows those who convert to dementia. Such data demonstrate the high sensitivity of QEEG to the earliest signs of functional brain impairment, prior to the presence of structural abnormalities. It is noteworthy that theta power has been shown to be negatively correlated with regional perfusion, especially in temporoparietal and central regions, and with PET glucose metabolism in temporoparietal and frontal regions, and is positively correlated with hippocampal atrophy.

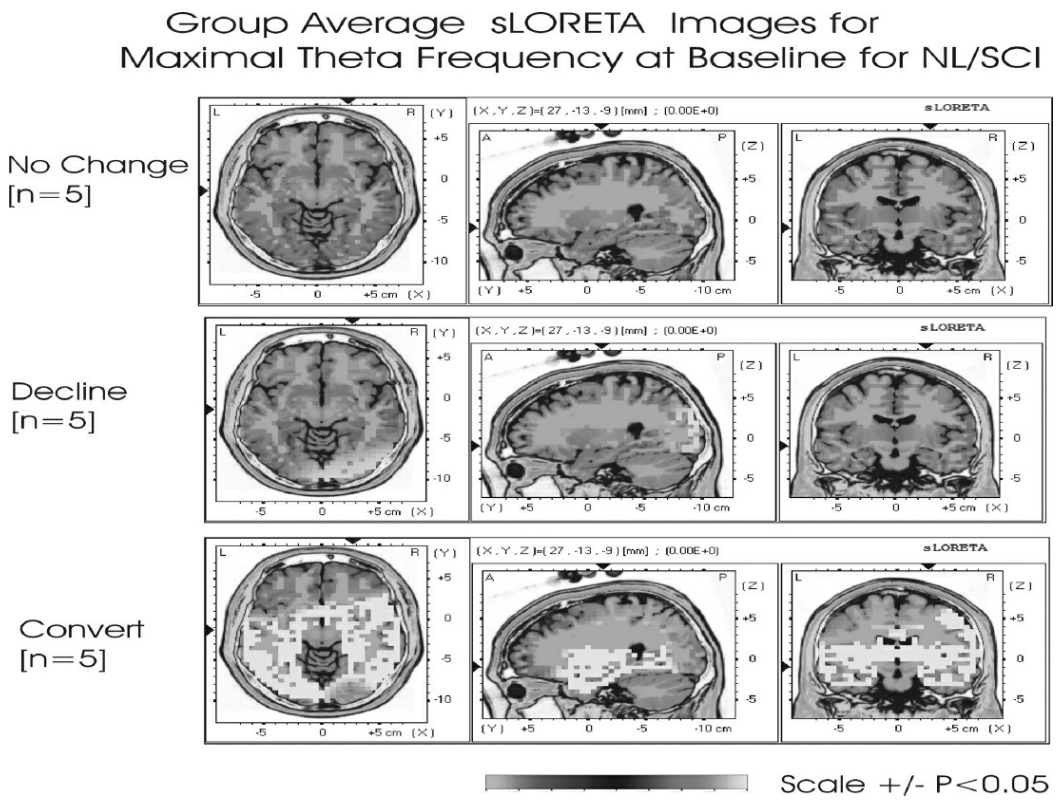

FIGURE 7.7–10. Group average low-resolution electromagnetic tomographic analysis (sLORETA) images (n = 5 in each group) of the regions displaying a maximal excess of theta activity at baseline for three different groups (n = 5 in each group) of elderly patients with only subjective complaints of impairment (SCI), all diagnosed as GDS 2, with no demonstrable neurocognitive deficits, who upon long-term follow-up (5 to 7 years) were found to show: **Top row**, no change; **middle row**, decline to mild cognitive impairment (MCI); **bottom row**, conversion to Alzheimer disease (AD), with GDS 4 to 6 and substantial impairment. The images in each row, from left to right, depict transaxial, sagittal, and coronal slices automatically selected to intersect at the coordinates of the voxel that had the most deviant standard score for excessive theta activity (4.3 Hz). Each voxel is color coded for the statistical significance of excess theta in standard (Z) scores, with hues ranging from red to yellow as significance of abnormality (over-activation) increases. Regions of increasing abnormalities include hippocampus, parahippocampus, and parietal-temporal cortex, as also indicated in other brain imaging methods.

Chronic Pain and Thalamocortical Dysrhythmia (TCD). Chronic pain affects over 35 percent of the US adult population, increasing greatly in

those over the age of 65. Furthermore, chronic pain is reportedly comorbid with depression in the vast majority of the patients, making chronic pain and its effective treatment a problem for both neurology and psychiatry. Contributing to the high prevalence of chronic pain and both the over and under treatment of the condition, is the subjective standard of care for rating pain, which can lead to inappropriate and suboptimal treatment. The use of QEEG features for the objective quantification of pain has led to an increasing number of studies which have demonstrated the coregistration between functional neuroimaging studies and QEEG studies of the relationship between activation of the brain's "pain matrix" and the experience of pain intensity. Findings of QEEG source localization in pain patients support the potential to derive a quantitative measure of the severity of pain using information extracted from a multivariate descriptor of the abnormal over-activation of regions of the pain matrix. The frequency specific (theta band) over-activation reported in the regions including the thalamus in the chronic pain patients, has been demonstrated in a number of studies and supports a model of thalamocortical dysrhythmia (TCD) as the mechanism of the neuropathic painful condition. This hypothesis is further supported by studies in chronic pain patients following successful central lateral nucleus thalamotomy (CLT), showed significant diminution in over-activation in the previously over-activated regions.

Using QEEG source localization significant differences between those chronic pain patients with and those without clinically significant depression have been shown. No significant correlation was found between the frequency-specific theta maximum abnormality in the ROIs of the Pain Matrix pain patients and the Beck depression scores in these patients. Thus, although both disorders share underlying sources in the limbic system, results support the existence of distinctive profiles of electrophysiological abnormalities.

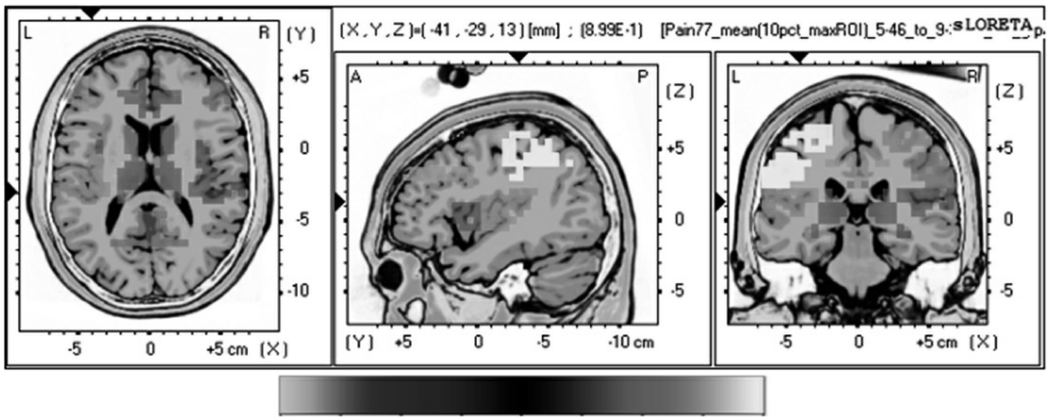

FIGURE 7.7–11. Source localization (sLORETA) images (axial, saggital, and coronal views, repsectively) for group average voxel *Z*-scores for ROIs in the band 5.46 to 9.36 Hz for the chronic pain population (n = 77). The images are color coded for significance of difference according to the color scale (below figure) where extremes of the color scale are

$p < \pm 0.01$. Over-activation is shown in red/yellow and deactivation in blue/aqua (none seen in this image). [Left is on left in these images.] (See color atlas.)

Figure 7.7–11 applies the methods described above to image the ROIs in the chronic pain patients. A clear relationship can be seen between those regions with significant over-activation and those identified using conventional, more costly, time consuming and not readily available neuroimaging tools. These regions include the anterior cingulate, anterior and posterior insula, parietal lobule, thalamus, S1 and dorsolateral prefrontal cortex (DLPFC). While consistent with those reported with conventional functional imaging, the use of QEEG source localization also include the mid and posterior cingulate not previously reported, suggesting that the increased temporal resolution of electrophysiological measures allows a more precise identification of the pain pathway.

Psychosis. EEG abnormalities in schizophrenic patients have been the subject of a large number of publications. Consistent descriptions of increased slow activity, especially in anterior and frontal locations, increased fast beta activity and interhemispheric asymmetries (R > L) appear in this literature. The influence of variations of medications, chronicity of medication, and length of illness may explain some of the inconsistencies between studies. Another problem in interpretation of these studies is the existence of electrophysiological heterogeneity within the populations. Some studies have specifically addressed this issue, with illuminating findings that reveal subtypes with the same diagnoses but with different profiles of salient QEEG abnormalities, as well as different responses to various psychotropic medications.

The results of a QEEG subtyping study of a large population of psychotic patients are summarized here. With multivariate cluster analysis of QEEG features for a large number of psychotic patients (n = 390), including medicated, unmedicated, and never-medicated first-episode patients diagnosed with the DSM-IV or ICD-10 with schizophrenia, psychotic depression, alcohol psychosis, and other psychoses, from a multisite international study, six clusters were identified, replicating and extending prior findings in schizophrenic patients. The distribution of patients across the six subtypes found in this study was not uniform, with subtypes containing as few as 24 or as many as 110 members out of the total cohort of 390 cases. Patients bearing a diagnosis of psychotic depression or schizophrenia either on or off medication, as well as first-episode medication-naive patients, were represented in at least four of the six clusters.

All clusters showed power asymmetry in all or almost all bands, and the power was greater over the right hemisphere, especially in regions of the frontal lobe. The uniform strong right hemispheric dominance of every

frequency band in each cluster would suggest a general impoverishment of language. It has long been reported that left hemispheric impairment is characteristic of "schizophrenia." In particular, abnormalities of the left parahippocampal gyrus, the left medial temporal lobe, the loss of synaptic proteins from the left thalamus, etc. have been repeatedly reported in schizophrenic patients. All clusters also displayed "hypofrontality," with pronounced deficits of beta activity in the dorsolateral and mesial prefrontal cortical regions. These findings are concordant with reports of decreased metabolism in the frontal cortex consistently noted in imaging studies using various modalities.

This data suggest that the left-sided deficiency generalizes to psychosis, with the right hemisphere becoming dominant because of a lack of interhemispheric suppression, possibly related to the pervasive disturbance of coherence in one or another frequency band between various homologous regions, to the extent that such coherence may reflect the activation of a feedback loop between the two sides of the brain. The power dominance of the right hemisphere together with the hypocoherence found in this study may indicate a deficit of homeostatic feedback between homologous regions. The severe hypocoherence in both short fiber and long fiber tracts within each hemisphere found in every cluster in this study provides strong support for the presumption that information processing is disturbed in psychotic patients. There have been repeated suggestions that the personality and thought disorder observed in schizophrenics reflect failure to integrate activity across distributed neural circuits.

Neuropsychological demonstrations of deficits in learning, memory, and executive tasks were interpreted to indicate widespread generalized frontolimbic dysfunction. Of particular interest in the context of the authors' pervasive findings of disturbed coherence are recent neurobiological studies of the binding process, which attribute a major role in perception and consciousness to the synchronization of EEG oscillations. Repeated reports have called attention to phase-locked oscillations in the gamma (30 to 100 Hz) frequency band. Phase-locking is a quantitative index of the temporally synchronized interactions in distributed neuronal systems. Decreased functional connectivity has also been reported in first-episode neuroleptic-naive schizophrenic patients.

It is also of note that five of the six clusters showed excess activity of various regions in the theta to low alpha frequency range. This slow wave activity is known to be frequently associated with hypoperfusion, repeatedly observed in PET studies of schizophrenic patients. Such slow activity has sometimes been thought to reflect neuroleptic effects, but studies have noted increased slowing in unmedicated patients compared with medicated patients. Note that in the present study, medicated, unmedicated chronic, and never-medicated first-episode patients were represented in each of the six clusters.

VARETA source localization images of the brain regions displaying the computed intracerebral sources of the most excessive and most deficient activities are shown in Figure 7.7–12. These underlying sources of the QEEG clusters indicate pathophysiological heterogeneity among patients diagnosed as psychotic, arising from two different causes of dysfunction in six basic subtypes, namely, over-excitability of one set of brain regions due to different underlying causes in each subtype, together with under-activation of a second set of brain regions perhaps due to the same cause in each subtype.

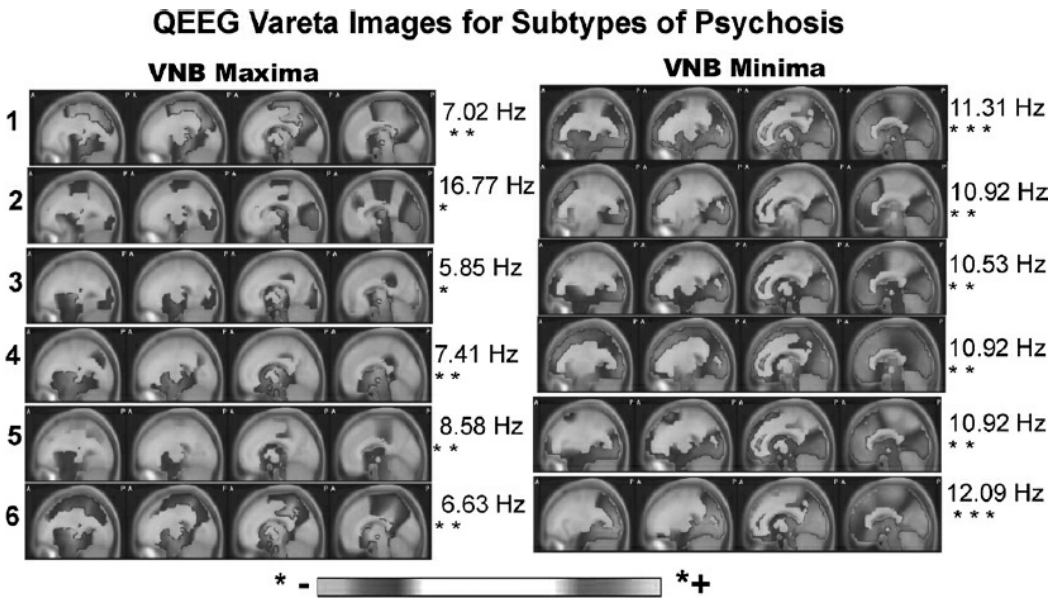

FIGURE 7.7–12. From top to bottom, each row depicts group average quantitative electroencephalography (QEEG) variable resolution electromagnetic tomographic analysis (VARETA) images of selected sagittal slices from the six subtypes (1 to 6) of psychotic patients identified by cluster analysis of a population of 377 psychiatric patients, all of whom were psychotic. Each cluster contained patients who had received a variety of fourth revised edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR)/tenth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) diagnoses, including alcoholism, depression, and chronic medicated, unmedicated, and never-medicated first-episode schizophrenia (see text). (**Left**) Sagittal slices depicting the anatomical regions that had the maximum excessive activity are shown, and (**right**) the corresponding slices with the most deficient activity, with the corresponding most deviant very narrow-band (VNB) spectral frequencies of each subtype indicated at the right side of each set of slices, are shown. Note that the most deviant excessive frequencies (Hz) of the six clusters (**left**) cover a wide range, 7.0, 16.77, 5.85, 7.41, 8.58, 6.63 Hz, and were found in a diverse set of brain regions, suggesting a variety of putative upregulated neurotransmitters, while the most deficient frequencies were within a very narrow range, 11.31, 10.92, 10.53, 10.92, 10.92, 12.09 Hz, and were found in a very consistent set of brain regions, suggesting a very small set or possibly even a single downregulated neurotransmitter. The deviation from normative values of every voxel in each slice is color-coded proportional to standard scores, with the appropriate hues corresponding to the color bar that is shown below. Note that the statistical significance of the hues of each voxel can be estimated by multiplying the *Z* score of the corresponding value from the color bar (and the palette range is indicated for each row by: *, ±1.0; **, ±2.0; ***, ±2.5) by the square root of the cluster size *n,* where: Clus 1, *n* = 53; Clus 2, *n* = 93; Clus 3, *n* = 65; Clus 4, *n* = 24; Clus 5, *n* = 110; Clus 6, *n* = 32. Thus, all voxels colored other than gray in these slices are extremely deviant from normative values, with a minimum *p* <

0.001.

All six clusters displayed some regions which were *hyperactive,* four displaying excessive activity (at maxima shown in figure below) in the theta range (6 to 7.5 Hz), one in the low alpha range (8.6 Hz), and one in the beta range (16.8), suggesting that somewhat different neurotransmitter abnormalities might exist among the members of the cohort. All detected hyperactive regions were located either bilaterally or only on the right side. The most widespread hyperactivity was noted in the thalamus in all six subtypes, in one or another region of the basal ganglia of five or all six of the six subtypes (caudate, globus pallidus, putamen, and subsantia nigra), in one or another region of the frontal lobes and limbic system of three up to five of the subtypes (DLPFC, hippocampus, amygdala, and posterior cingulate) and in various regions of the cortex of four to six of the subtypes (medial and superior temporal cortices, para- and precentral gyri, and parietal cortex). In addition, all six subtypes displayed some regions that were extremely *hypoactive,* found in some region of the prefrontal cortex of four subtypes (DLPFC, gyrus rectus, orbitofrontal, and medial frontal cortices), in one or another region of the limbic system of five to six (amygdala, anterior cingulate, parahippocampus, and hippocampus), in one or another region of the basal ganglia (caudate or putamen) and thalamus of four, and in the occipital lobe of all six.

These data suggest that the optimal management of the psychotic patient will entail separately upregulating one neuroanatomical system while down regulating a different system, although it is conceivable that the distributed hypoactivity reflects inhibitory influences exerted by the hyperactive system. The regions in which homeostatic regulation is most severely disturbed across all six subtypes, displaying both hyperactivity at one frequency as well as suppression of activity at a different frequency, are the amygdala, caudate, cingulate, DLPFC, hippocampus, and thalamus, suggesting that psychotic symptoms may primarily reflect a dysfunction arising from subcortical dysregulation.

Recent studies have used eLORETA to evaluate the EEG Source Functional Connectivity (EEG-SFC) in schizophrenic patients. One such study in stable medicated schizophrenia patients and compared them with healthy controls. Results demonstrated abnormalities in functional connectivity especially in the theta frequency band where synchronization was increased (hypercoherent) in schizophrenic patients, which was increased in frontal regions with increased duration of the disease. Alpha band connectivity was seen to be decreased (hypocoherent) in frontal regions in schizophrenics and independent of length of illness. The higher frequencies (beta1, beta2, and gamma) also showed disturbances in connectivity, involving frontal regions and patrietotemporal regions. The interactions found with chronicity need to be taken into consideration

when interpreting such findings, but results lend support to the hypothesis of schizophrenic involving disturbances in connectivity.

QEEG-Based Depth of Anesthesia Monitor: Level of Consciousness

Instruments that use QEEG to provide dependable real-time monitors of the depth of anesthesia during surgery, measuring the level of consciousness, have demonstrated distinctive and quantitative changes from the ground state. Intraoperative studies using such instruments revealed that with loss of consciousness (LOC) there was a marked anteriorization of alpha activity into all regions of the frontal cortex, followed by a shift to a power spectrum dominated by activity in the regions of the power spectrum at the margins between delta and theta activity (2 to 4 Hz), which gradually disappeared as the patients emerged from the levels of surgical anesthesia through deep sedation to arousal. It is of particular interest that gamma coherence between prefrontal and posterior brain regions was invariantly sharply diminished precisely at LOC during anesthesia induced by any anesthetic agent studied and returned to preinduction levels just prior to the return of consciousness. While anterior–posterior coherence similarly dropped sharply at LOC for activity in the delta, theta, alpha, and beta bands, it was noteworthy that coherence for anterior–posterior neural transactions in these frequency ranges did not return to preinduction levels until as late as 10 to 15 minutes following return of consciousness, as gauged by overt responsivity to verbal commands and speech. This suggests that gamma coherence may be of unique relevance to the maintenance of consciousness per se, which will be further elaborated in a later subsection of this section.

Application of source localization algorithms (VARETA, described above) to the EEG data collected during anesthesia revealed extensive inhibition of the frontal lobes, thalamus, basal ganglia, amygdala, and hippocampus coincident with the LOC. Return of consciousness was always preceded by reversal of the QEEG changes and disinhibition of these structures. Such findings have led to a new neurophysiological theory of the mechanisms of action of general anesthesia.

The intraoperative studies of the invariant and reversible QEEG changes that accompanied the loss and return of consciousness are compatible with five neurophysiological steps during the induction of surgical depth of anesthesia with any combination of agents, as follows: Step 1: Depression of the brainstem reduces the influences of the ARAS on the thalamus and cortex. Step 2: Depression of mesolimbic–DLPFC interactions leads to blockade of memory storage. Step 3: Further depression of the ARAS releases its inhibition of the nucleus reticularis of the thalamus, resulting in closure of thalamic gates (especially in the diffuse projection system) by hyperpolarizing GABA-mediated inhibitory action of the nucleus reticularis (theta increase), thereby blocking (Step 4)

thalamo-cortico-thalamo reverberations and perception (gamma decrease), so that (Step 5) parietal–frontal transactions are uncoupled (gamma coherence decreases), blocking cognition, and (Step 6) prefrontal cortex is depressed to reduce awareness (increase of frontal delta and theta).

Since its appearance, a number of publications have confirmed various aspects of this theory, particularly with respect to the central role of the changes in coherence between anterior–posterior oscillatory activity in the gamma band. More specifically, studies of QEEG symbolic transfer entropy have shown that directional connectivity between anterior and posterior regions is inhibited by anesthesia; that there is reorganization of hub networks from parietal to frontal regions under anesthesia; and that QEEG directional connectivity before and after LOC show decreases between the frontoparietal and frontooccipital regions. In summary, changes in QEEG features with loss and return of consciousness suggest an "anesthetic cascade" involving a blockade of thalamocortical reverberations and uncoupling of normal transmission between frontal and parietal/occipital regions.

Therapeutic Correction of Departures from the Ground State

There are a number of clinical areas in which the validity of the proposition that therapeutic maneuvers that restore previously QEEG-deviant patients to the ground state will have a positive clinical outcome has been demonstrated. These studies suggest the possible utility of serial QEEG or BSV evaluations as an adjunct to the selection of treatment, evaluation of efficacy, and optimization of dosage.

Pharmaco-EEG. Since pioneering work began in the 1960s, QEEG has been used as a tool in the development of new psychopharmaceutical agents. It was early observed, using only a few electrodes, that drugs in a particular class caused a distinctive alteration of the frequency composition of the spontaneous resting EEG. For example, anxiolytics typically cause an increase in power in the beta band, especially in EEGs recorded from electrodes over the prefrontal regions. These observations gradually became systematized so that data were collected from the full 19-channel array of the 10/20 system.

With the advent of QEEG methods, dose–response curves were calculated to correlate the amount of agent, time after administration, and quantitative changes in specific QEEG variables. It was noted that these changes followed typical time courses for different agents, indicating the potential utility of QEEG to study pharmacodynamics and bioavailability. These relationships are now sufficiently well understood for some agents that serial QEEG evaluations can be used to obtain an early indication of whether or not a specific drug and dose will be efficacious for an individual patient and dose–response curves can be constructed using serial

computations of BSVs across successive doses.

Since under normal conditions the QEEG is highly reproducible, in some instances it is possible to use a simple test–retest paradigm to evaluate whether a specific drug is appropriate for an individual patient, with the two tests separated by a time interval sufficiently long to enable an effective concentration to be established in the brain. For example, in a child with ADHD/ADD who will be a good responder, within 45 minutes after a single 10-mg dose of methylphenidate, pharmacodynamic effects can be seen as a significant decrease of premedication QEEG abnormalities.

The importance of the potential contribution of QEEG to pharmacotherapy and increasing use led to the formation of a committee of thought leaders by the IPEG (International Pharmaco-EEG Society), who recently published Guidelines for the Recording and Evaluation of Pharmaco-EEG in man.

Key-Lock Principle. As pharmaco-EEG knowledge accumulated, researchers discerned that certain patterns of QEEG effects were distinctive for certain classes of drugs. It gradually became apparent, at least for some conditions and some drugs, that the so-called "key-lock" principle appeared to provide a guide for the selection of medication. For example, the optimum medication for a patient whose QEEG abnormality was a deviation of the power spectrum from normative baseline with excess theta and deficit of alpha would be the pharmaceutical substance that, when administered to normal individuals, caused appearance of the opposite power spectral deviation: Decrease of theta and increase of alpha. A body of evidence has been published that provides substantial support for this proposition.

A clinical example of this principle is demonstrated in the relationship between response to SSRIs and QEEG cluster membership shown in a study of OCD patients, all of whom were considered clinically to be candidates for SSRI treatment. Two clusters were found in this study, with differences in QEEG features that can be seen in Figure 7.7–4, where the pretreatment baselines of the two different OCD clusters can be seen in the right panel, rows 1 (responders) and 2 (nonresponders). One of these OCD clusters displayed an excess of theta activity (Fig. 7.7–4, right panel column 2, second row), while members of the other cluster displayed an excess of alpha power (Fig. 7.7–4, right panel column 1, first row). The effect of SSRIs on the normal EEG is to decrease alpha and increase theta activity. As can be seen in Figure 7.7–4 consistent with what would be predicted by the key-lock theory, the cluster whose baseline (pretreatment) QEEG was characterized by excess of alpha and deficit of theta activity is the one which contained the SSRI responders, and members of the cluster with an excess of theta were nonresponders.

Neurofeedback. NFB is essentially operant conditioning of brain electrical activity, first applied to the treatment of epilepsy in the early 1970s. NFB treatments postulate that: (1) continuous measurement and representation of some selected target correlate of an undesirable activity or physiological process (autonomic measure, electromyogram [EMG]), ERP component, or composition of the QEEG power spectrum) with provision for visual or auditory feedback upon reduction of the representation or some alternative type of reinforcement to the subject whenever a decrease of the target is detected can be used to help a patient learn how to reregulate some abnormal physiological process or aspect of brain electrical activity; and (2) that normalizing the target signal will result in a therapeutic benefit. During the last decade, a large number of practitioners have begun to use NFB to treat a wide variety of behavioral disorders and dysfunctions, most widely for ADHD but including such applications as remediation of neurological deficits resulting from traumatic brain injury or stroke. Among the several hundred papers reporting the utility of NFB are explorations of its use in addiction, ADHD, especially the suppression of undesired motor activity while remaining attentive, anxiety disorders, autism, autoimmune disorders, chronic fatigue syndrome, conduct disorder, epilepsy, learning disabilities, migraine, mood disorders, and, most recently, performance enhancement of skilled actions. In several of these areas, there have been replicated reports of statistically significant, successful reduction of the targeted symptoms.

The basic method of NFB is to record a measure that reflects or is correlated with some physiological process or condition considered undesirable, pathognomonic, or pathophysiological, which is being continuously monitored. Some reinforcement is provided whenever the subject/patient succeeds in changing the undesirable process in a direction defined to be therapeutic, increasing or decreasing the measure. The manifestation of the undesirable condition is usually reflected in some stimulus parameter such as the intensity, pitch, or temporal pattern of an auditory stimulus or the size or brightness of a visual image, and the reinforcement is delivered by a change in the parameter, defined as change of the undesirable process in the desired direction, such as reduction of the intensity of sound or size of a flame.

A variety of measures have been reduced successfully by NFB, including: (1) measures reflecting the autonomic nervous system, such as anxiety, galvanic skin response (GSR), high heart rate, high blood pressure, or pupillary dilatation; (2) conditions such as anxiety, attention, or obsessive behavior; (3) ERP measures considered putatively to reflect attentional processes, such as mismatch negativity (the occurrence of "different" in a sequence of paired stimuli S1–S2 that are "same–same" or "same–different"), the amplitude of P300 in the "oddball" paradigm (in which an occasional "rare" event is interspersed in series of "common"

events producing an expected late positive component to the rare event in normal cognition); (4) ratios between QEEG measures in different bands (e.g., the "theta/alpha ratio" of the power spectrum recorded at a particular electrode position). A study in chronic pain demonstrated that the abnormal electrical activity at a particular frequency detected using real-time QEEG source localization, within the anterior cingulate, could be used for NFB to reduce intractable pain. Examples in the ADHD population are described above.

A particularly novel application of NFB has been its use to enhance artistic performance, a domain in which successful intervention is a priori quite unexpected. A series of well-controlled studies sought evidence of operant control of the alpha/theta ratio, with the rationale that such control might induce a meditative state of deep relaxation. This parameter was then used to enhance performance, using criteria established by the Associated Boards of the Royal Schools of Music of England to assess variables of artistic performance, including several in the domains of instrumental competence, musicality and musical understanding, communicative ability, and overall quality. Performances were videotaped and randomized and then evaluated by experienced assessors external to the institution and blinded as to the order of performances and experimental group membership of the students. Highly significant results were obtained and successfully replicated.

An extensive review article, with a survey of evidence that validates NFB in diverse applications but particularly in ADHD and performance enhancement, has recently appeared. Although NFB was initially disparaged by many neuroscientists, neurologists, psychologists, and psychiatrists, the success and extent of this therapeutic activity is indicated by increasingly widespread utilization, and reports of carefully designed scientific studies reported in the peer-reviewed literature suggest the clinical utility of the method.

Consciousness

Definition of Consciousness. Webster's dictionary defines consciousness as, "… the quality or state of being aware, especially of something within one's self." Webster also defines awareness as, "… the condition of being conscious of an external object, internal state, or fact." This tautology reflects the lack of scientific attention to this problem that persisted for much of the last century, while the question of how subjective experience arises from neural tissue was considered not a fit topic for neurobiological research. With advances during the last decade in functional brain imaging and methods better adapted to the study of systems, neurophysiological and psychological evidence has been forthcoming that supports a definition of consciousness as: "the subjective experience resulting from the continuous interaction of three influences:

[1] incoming information about the external environment and the internal state, melded with [2] information retrieved from a dispersed memory system of stored information about previous working memory or episodic experiences that are the most currently relevant, and [3] state factors modulating both our perceptions of and our reactions to the outer world, by affecting arousal levels, emotional states and moods."

By "relevant" is meant those experiences for which the environmental and physiological state-generated inputs are most evocative, because of inherent salience and situational correspondence or their multisensorial similarities. By "modulated" is meant the biasing of apperception and memory selection due to more global, direct or indirect influences of humoral factors to alter excitability of components of both the responsive and the reactive systems.

Information Is Synchronous Deviations from Ground State

If the billions of neurons whose activity is sensed by an EEG electrode were discharging completely at random, then there would be no net voltage detectable by the electrode. Voltage excursions in the resting EEG or ERPs that cause the QEEG power spectrum to deviate from the ground state represent temporal fluctuations in underlying regions of the synchronized activity of a significantly great number of neurons relative to the proportion that would be expected to discharge together solely by chance. Analogously, the images of sensory, perceptual, or cognitive processing detected by fMRI represent the spatially distributed decrease in oxyhemoglobin caused by metabolic requirements of massively synchronized ensembles of brain cells, relative to reference background activity. All such ensembles are intimately interconnected and in constant bidirectional transactions with numerous other ensembles of neurons at varying distances. The homeostatically regulated ground state of the resting EEG power spectrum of every brain region, as well as the remarkable stability of transactions reflected in the coherence and symmetry among these regions within and between the hemispheres, indicates that the spontaneous synchronization of nonrandom firing within and among regional populations of neurons is dynamically restricted and constrained within well-defined limits and regulated by positive and negative feedback processes. The constancy of these quantifiable variables of brain electrical activity over sustained time periods, in spite of the known fluctuations in excitability of individual neurons, indicates that this stability reflects a basic property of the large-scale organization of the central nervous system. Consistency with the concept of the DMN is noted.

Such predictable background brain activity contains no information, because information is that which cannot be predicted. Significantly nonrandom or synchronous brain electrical activity beyond these constrained limits constitutes a deviation from the homeostatically

regulated ground state that defines the information content of that perceptual moment and causes the EEG voltage to deviate from the regulated range of random fluctuations. Spatially dispersed nonrandom synchrony is the key signature of significant informational activity within the brain. Information, retrieval of memories, and the generation of the perceptual content of moments of consciousness are encoded as temporal patterns of synchronized nonrandom activity in very large ensembles of neurons distributed spatially in the exogenous and endogenous systems causing dispersed neuronal ensembles to deviate from the ground state. One might speculate that brief excursions around the constrained baseline of the ground state reflect the sporadic thoughts that fleet through the mind of a person resting quietly, which rapidly converge to the ground state unless they are long sustained.

Discontinuity of Consciousness. Subjectively, consciousness appears to be continuous and unitary, that is, seamless. However, psychophysiological studies indicate that perceptual processes within the brain are discontinuous, parsing consciousness into successive "perceptual frames," or moments of subjective experience. Under controlled experimental conditions, if the interval between two stimuli becomes as short as 80 to 100 milliseconds, then the pair merge into a single event. Although the EEG at every electrode position on the head appears to be an apparently continuous time series of voltages, the topography of the variance of the local deviations from the average voltage over the entire scalp, or global field power (GFP), is also discontinuous. A limited number of topographies actually wax and wane at rather regular intervals that have been called "microstates" or "atoms of thought," found to have an average duration about 82 ± 2 milliseconds within a normally functioning healthy population of 500 subjects ranging in age from 6 to 90. Similar discontinuities have been found in magnetoencephalographic (MEG) research. Such evidence suggests that the apparent continuity of subjective time is an illusion produced by neurophysiological processes parsing information into successive chunks about 80 to 100 milliseconds in duration.

On the average, four spatial topographies account for 94 percent of the variance of the scalp EEG, mapping the field in four basic modes: Activation of right anterior hemisphere, left anterior hemisphere, frontal cortex, and motor cortex. These may, respectively, represent encoding of spatial information in the environment, verbal encoding of that information, executive interpretation and evaluation of the verbal–spatial representation of the environment in view of relevant memories especially of the recent past, and the organization of adaptive language and motoric responses. The transitions among modes tend to follow a particular sequence in normal subjects but are somewhat different in patients with

some psychiatric disorders.

Parallel Multiplexed Information Processing Channels. Although evidence shows that perceptual frames last about 80 to 100 milliseconds, other experiments have established that temporal discriminations can be made at substantially shorter intervals, on the order of tens of milliseconds or less. Thus, the neural receptor systems mediating each sensory modality may be organized into multiple parallel channels, containing perceptual frames offset from one another by a fixed time, like commonly used multiplexing of information. Using averaging techniques time-locked to presentation of a specific sensory stimulus is equivalent to demultiplexing, enabling the electrical activity reflecting the processing of information in only one of these parallel channels to be extracted as an ERP. It is important to stress that dispersed neuroanatomical structures containing the other parallel channels in the afferent pathways are not only processing that information but are simultaneously engaged in many other neural transactions, continuously active during the time required for the averaging process. As the time-locked activity of one channel converges to a stable ERP wave shape, considered as "*signal,*" the ongoing activity in all of the other channels converges to zero because it is in random phase relative to the time of stimulus presentation and is considered as "*noise.*" The signal-to-noise ratio in the ERP waveshape extracted by the averaging process is improved proportional to the square root of the number of samples that have been averaged. If subjective experiential continuity is mediated by consecutive inputs from these hypothesized multiplexed channels, then *consciousness must reflect a global statistical steady state maintained within a spatially extensive brain system that is independent of any individual neurons.*

Construction of Consciousness. The brain is organized to assign an affective and experiential label on every moment of input from the environment that causes a significant nonrandom discharge. When such a discharge is provoked by external or exteroceptive stimuli, it activates a network of neuronal ensembles linking the sensory receptors to subcortical structures in the brainstem and the thalamus, causing volleys of oscillatory activity to be propagated to dispersed regions of the primary sensory cortex and to cortical association areas, providing a fractionated representation of the *present.* These volleys propagate in parallel to structures such as the amygdala, entorhinal cortex, hippocampus, and septum comprising the limbic system, merging with and modulating the relevant activated readout from the *past,* as well as to subcortical structures such as the basal ganglia and hypothalamus reflecting *internal state,* that together mediate emotional valence and motivation as well as motoric organization. A spatially extensive set of nonrandom electrical cortico-limbic-

subcorticocortical oscillations links these structures, *binding* them into a dynamic functional network. That which is linked becomes phase locked and coherent, producing an instant of subjective experience that merges the present with selected aspects of the past: The "*remembered present*." These coherent discharges define all of the information in the brain at any moment, which must subsume the content of consciousness. Therefore, the content of consciousness logically must arise from this spatially extensive nonrandomness.

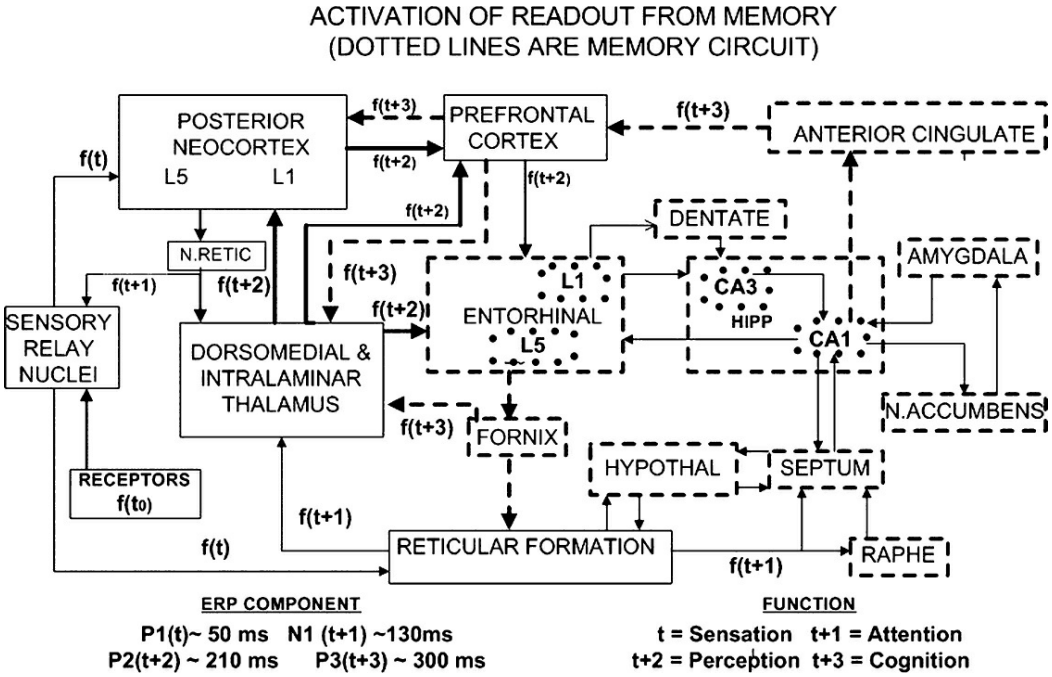

FIGURE 7.7–13. Sequential steps of "mental chronometry" constructing a perceptual frame are postulated to occur simultaneously in numerous parallel processing channels, combining information from the exogenous and endogenous system into a seamless subjective moment of the "remembered present." Incoming information from sensory receptors is propagated via the thalamocortical relay (TCR) nuclei to basal dendrites of feature-extracting pyramidal neurons dispersed throughout the neocortex in layer 5, producing the first positive (P50) event-related potential (ERP) component (at time *t,* about 50 milliseconds) and simultaneously enters the reticular formation, registering the first elements of sensation. Activated cortical regions instruct the nucleus reticularis to inhibit selected regions of thalamic neurons to enhance the relevant signal-to-noise ratio, producing negative deflections (N150) (at time *t* + 1, about 130 milliseconds) related to attention, while the ascending reticular arousal system (ARAS) sends information to the intralaminar nuclei and the raphe nuclei, hypothalamus, and septum. From the intralaminar nuclei, activity propagates to the apical dendrites of pyramidal cells of the posterior and associative cortical regions, producing the second positive (P200) ERP deflection (at time *t* + 2, about 210 milliseconds) related to perception, and is sent as well from the dorsomedial thalamus to the basal dendrites of the prefrontal cortex and into the entorhinal cortex. Such incoming information from the exogenous system then circulates through the hippocampus and related structures of the limbic system, activating the most relevant past experiences by autoassociative mechanisms. Retrieved memories and their associated valence and context are projected via the entorhinal cortex and the anterior cingulate to the prefrontal cortex and thence to the posterior cortex, producing P300, the third positive ERP component (at time *t* + 3, about 300 milliseconds) related to cognition and subjective interpretation of the "meaning" of the percept. A late negative component, N400, reflects

semantic encoding. Dispersed elements of sensations are linked into perceptions and interpreted cognitively in the context of past experience by binding mechanisms, reflected by the appearance of widespread gamma oscillations among dispersed anatomical regions.

The discrete parsing of information into sequential segments in each modality, integrated across modalities during the microstates, is experienced as a series of apparently continuous perceptual frames. The remainder of this chapter attempts to describe how activity in exogenous and endogenous systems is combined to construct the remembered present, first taking advantage of knowledge about the processing of information within only one of the hypothesized channels, a snapshot artificially visualized by the ERP averaging strategy, as described in Figure 7.7–13. The authors consider the EEG to be an envelope integrating simultaneous time-shifted voltage sequences produced by multiple processes similar to this example, occurring in numerous parallel networks offset by the multiplexing time delay as the environment is continuously sampled by the brain.

Neurophysiological, psychophysiological, and fMRI research, for which reviews are available elsewhere, provides evidence supporting the concept that "mental chronometry" is reflected in the successive peaks of the cortical ERPs during processing of information. The first positive component (P1) is associated with arrival of information reporting the sensation of the stimuli, the negative–positive complex N1P2 is associated with attending to the input, the second positive cortical component (P2) is related to perception of the stimuli, and a later positive component (P3) reflects the cognitive interpretation or "meaning" of the events.

Modulated Exogenous–Endogenous Interaction. As the described construction of consciousness proceeds, each perceptual frame or moment of subjective experience is constructed from interactions between three systems, analogous to a *mental molecule* rather than an atom of thought.

1. An "*exogenous system*" disperses spatiotemporal patterns of activity representing multimodal encoding of the environment to *basal dendrites* of pyramidal neurons throughout the neocortex, modulated by influences of internal state. Discrete attributes are distributed to many different brain regions to be analyzed before they are recombined. Information from different sensory modalities requires disparate times to be dispersed in pathways of different lengths and propagation velocities, so the sustained persistence of the perceptual frame may be essential for information from disparate modalities to be synthesized into one simultaneous multimodal percept.
2. In parallel, an "*endogenous system*" disperses spatiotemporal patterns of input to the *apical dendrites* of the pyramidal neurons representing the readout of memories about the most relevant and salient past

experiences. Evidence suggests that an "autoassociative" readout process, based upon multimodal similarities between previous patterns of activity and the exogenous report of the present, retrieves a facsimile of stored experiences from extensive networks in regions of the limbic system and the basal ganglia, as well as throughout the association regions of the cortex.

3. A "*state, modulation system,*" mediated by interneurons that reflect the state of sympathetic–parasympathetic balance of the autonomic nervous system and endocrine and adrenal activity, acts upon large populations of pyramidal neurons at levels of basal or apical dendrites to weight the relative contributions of the exogenous and endogenous systems, "reality" and "the past," to the content of consciousness. These interneurons, by nonsynaptic modulation of membrane thresholds as well as synaptic input to pyramidal neurons, have the capability to differentially shift or bias the global thresholds of huge populations of neurons. They can desensitize or hypersensitize the exogenous system to certain kinds of afferent input and increase or diminish the salience of past experiences released from the endogenous systems.

Neuron–Interneuron Interactions. Each interneuron may contact as many as tens of thousands of pyramidal neurons, synchronizing theta and gamma oscillatory neuronal networks engaged in cognitive and emotional processing. Fast spiking (FS) basket cells surround the basal regions, which receive the major inputs from the exogenous system, while low threshold spiking (LTS) basket cells can modulate apical dendrites, which receive the major inputs from the endogenous system and potentially affect mood and recollections relevant to depressive and schizophrenic pathophysiology. Long-range interneurons are "interregional" and facilitate rapid synchronization. Thus, interneuron influences continuously modulate the input and output to pyramidal cells serving as coincidence detectors between elements of the past and the present, as schematized in Figure 7.7–14.

In particular, the release by the locus ceruleus of serotonin and dopamine can differentially inhibit the release of GABA from FS interneurons to modulate attentional and arousing effects and salience of inputs from the exogenous system onto the basal dendrites of the pyramidal neurons, while inhibiting the input of the LTS interneurons to the apical dendrites. Conversely, release of adrenal cortical hormone (ACH) and norepinephrine by the nucleus basalis can modulate excitability of the LTS interneurons and selectively block GABA, disinhibiting thalamic neurons from the inhibitory influences of the nucleus reticularis.

LTS interneurons can display self-induced slow long-lasting inhibition (SSI) mediated by endocannabinoids, decreasing their inhibitory influences. The resulting heightened excitability of the apical dendrites

increases the influences of the endogenous system. Intense activation of the LTS interneurons would lead to SSI, giving an avalanche of positive feedback to the pyramidal neurons due to disinhibition. Imbalance in a specific neurotransmitter might activate a subset of GABAergic interneurons, synchronizing remote ensembles of pyramidal neurons into closely phase-locked rhythmic activity. Such synchronization could enhance or diminish the impact of the environment and experiences upon the exogenous and endogenous systems and their interactions, changing the flavor of the "remembered present."

Coincidence Detection Constructs Fractionated Elements of Perception. Incoming sensory activity is dispersed by the TCR nuclei to the basal dendrites of pyramidal neurons in layers 4 and 5 of receiving regions throughout the sensory cortex, specialized to detect specific features of the multimodal focus of attention, as shown by red and blue arrows in the upper panel of Figure 7.7–15. Outputs from an "endogenous system" activated by this information and reflecting recent events, combined with signals reflecting internal state, are sent via the ARAS and the diffuse projection system to apical dendrites of the pyramidal neurons in the association and prefrontal cortices, where they are combined with input from primary receiving areas. This convergence, shown by mauve arrows in the middle panel of Figure 7.7–15 has been shown to be critical for perception of a sensation to be accomplished.

Binding, Resonance, and Consciousness. The "binding problem" is the question of how attributes of multiple sensory modalities comprising these dispersed fragments of perception are recombined into a holistic perceptual moment of subjective experience, augmented by the associated affect, emotion, and motivational valence. Convergence of similar temporal patterns, arriving via the exogenous inputs upon the basal synapses and via the endogenous system upon the apical synapses of selected neurons in the layers of pyramidal cells dispersed throughout the cortex, results in a coherent corticothalamic discharge of synchronous volleys from this population, with a back-propagated oscillation in the gamma frequency range (~40 Hz). This oscillation causes these regions to depart from the ground state and establishes dispersed fragments of *local negative entropy,* that is, *improbably* synchronized spatiotemporal activity patterns that represent fractionated elements of perception.

When enormous numbers of pyramidal neurons dispersed throughout numerous brain regions thus oscillate, reverberation links these regions together by phase-locked transcortical gamma oscillations. At the same time, neuronal ensembles comprising the networks of the endogenous system release memories retrieved by the autoassociative processes, which are reflected by gamma oscillations phase locked to theta waves propagated

to the prefrontal and association areas of the cortex. Cortical pyramidal cells in these regions appear to serve as coincidence detectors that coregister the present with the past. These spatially dispersed oscillations in the same high-frequency range, engaging the cortex, thalamus, subcortex, and limbic system, are physically brought into phase by electronic coupling, perhaps mediated by gap junctions and glial cells, and are thereby bound together. This system of temporally patterned, spatially dispersed but phase-locked, high-frequency oscillations of electrical energy begins to resonate, represented by the green arrows in Figure 7.7–13. Coherent gamma oscillations between prefrontal and parietooccipital areas of the cortex are essential for consciousness, as shown by their invariant reversible disappearance and reappearance with loss and return of consciousness in QEEG studies of anesthesia.

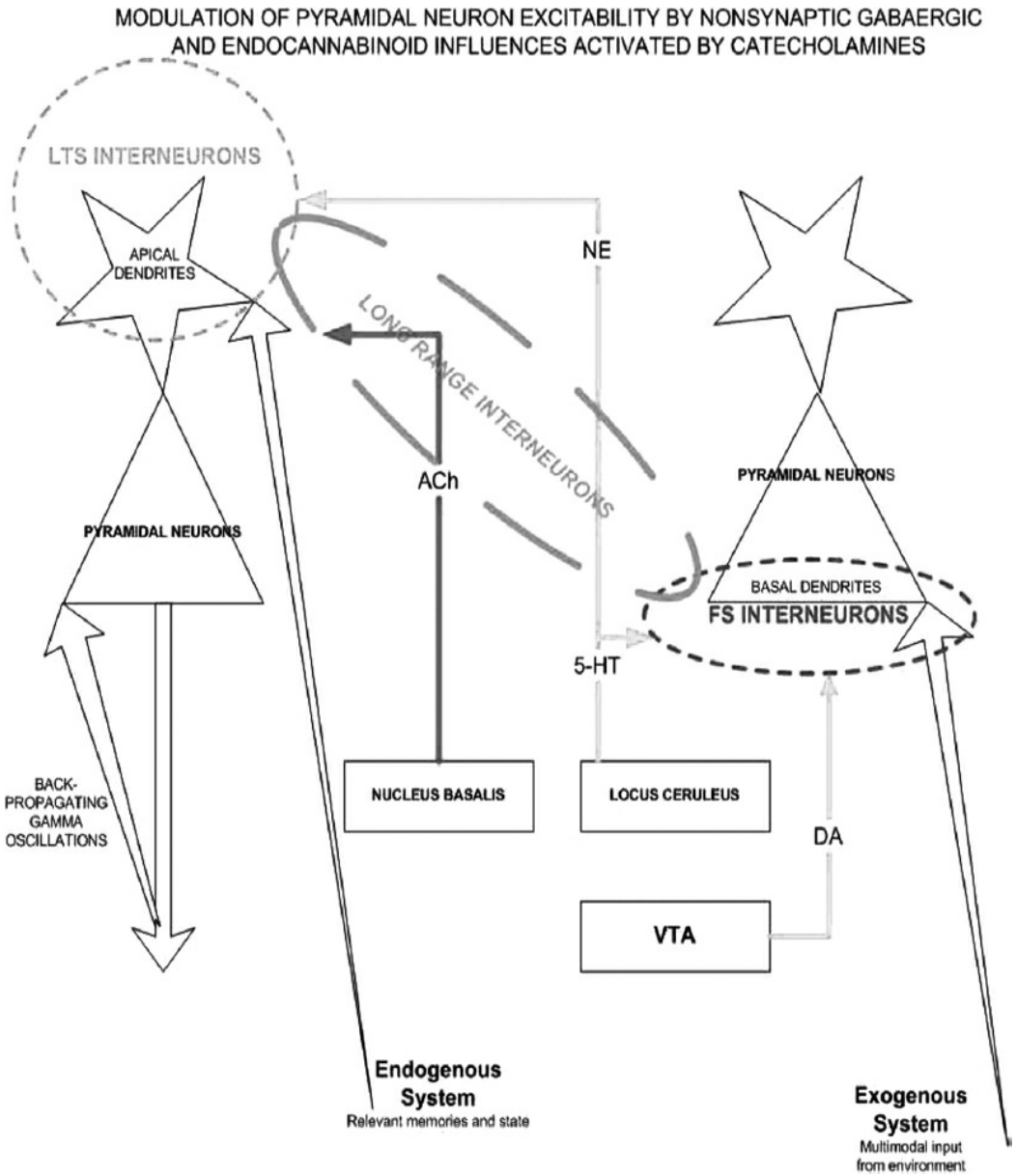

FIGURE 7.7–14. Modulation of excitability of basal dendrites sensitive to exogenous

influences by fast spiking (FS) interneurons releasing dopamine (DA) and serotonin (5HT), reflecting activity of the locus ceruleus and the ventral tegmental area (VTA), and of the excitability of apical dendrites sensitive to endogenous influences by low threshold spiking (LTS) interneurons releasing acetylcholine (Ach), reflecting activity of nucleus basalis of Meynert and depicting the synchronizing influences on huge, anatomically dispersed ensembles of populations of cells in all categories that can be exerted by long-range interneurons, are depicted.

This resonance among spatially dispersed neuronal masses results in a shift of the brain from the homeostatically regulated ground state, which creates an extremely improbable, informational field of global negative entropy. Phase-locking and resonance bind the exogenous and endogenous systems together into a continuously updating, globally coherent, anatomically dispersed yet cohesive functional network, whose operation is independent of any dedicated neuron or specific set of neurons; the posterior regions of the brain (the sensory-specific, association, and motor cortices), dominated by inputs reporting the present state from the exogenous, autonomic, and musculoskeletal systems, are coupled together in global resonance with the anterior regions (the precentral and frontal cortices), dominated by the input of readout from the most relevant endogenous emotional, motivational, and motoric state and memory systems distributed among subcortical structures, including the basal ganglia, thalamus, hypothalamus, and limbic system.

Consciousness is postulated to be a physical property of this organized informational field of global negative entropy dynamically constructed in this spatially extensive resonating system, not a localized process mediated by discrete neuronal transactions. While neuronal activity generates this property, consciousness is not directly a neuronal process but rather a consequence emerging from mass action of enormous numbers of neurons, just as a nuclear explosion is not a property of fission of any atom. Disturbances in the regulation of these complex transactions among masses of neural tissue are very relevant to the emergence of psychiatric disorders. Such disturbances might variously arise from dysfunction in the exogenous encoding mechanisms, the endogenous autoassociative retrieval mechanisms, or biasing of thresholds due to phasic or tonic fluctuations of the modulatory interneuron meshwork and can be envisaged potentially to generate psychopathological symptomatology. Malfunction of either or all of the exogenous, endogenous, and modulatory systems may distort the subjective experience to produce misperceptions and maladaptive behavior. Not only may each analytical apparatus be defective but the binding process itself may be askew. The symptoms of many psychiatric disorders may reflect disorganization of integrative processes dependent upon these binding mechanisms.

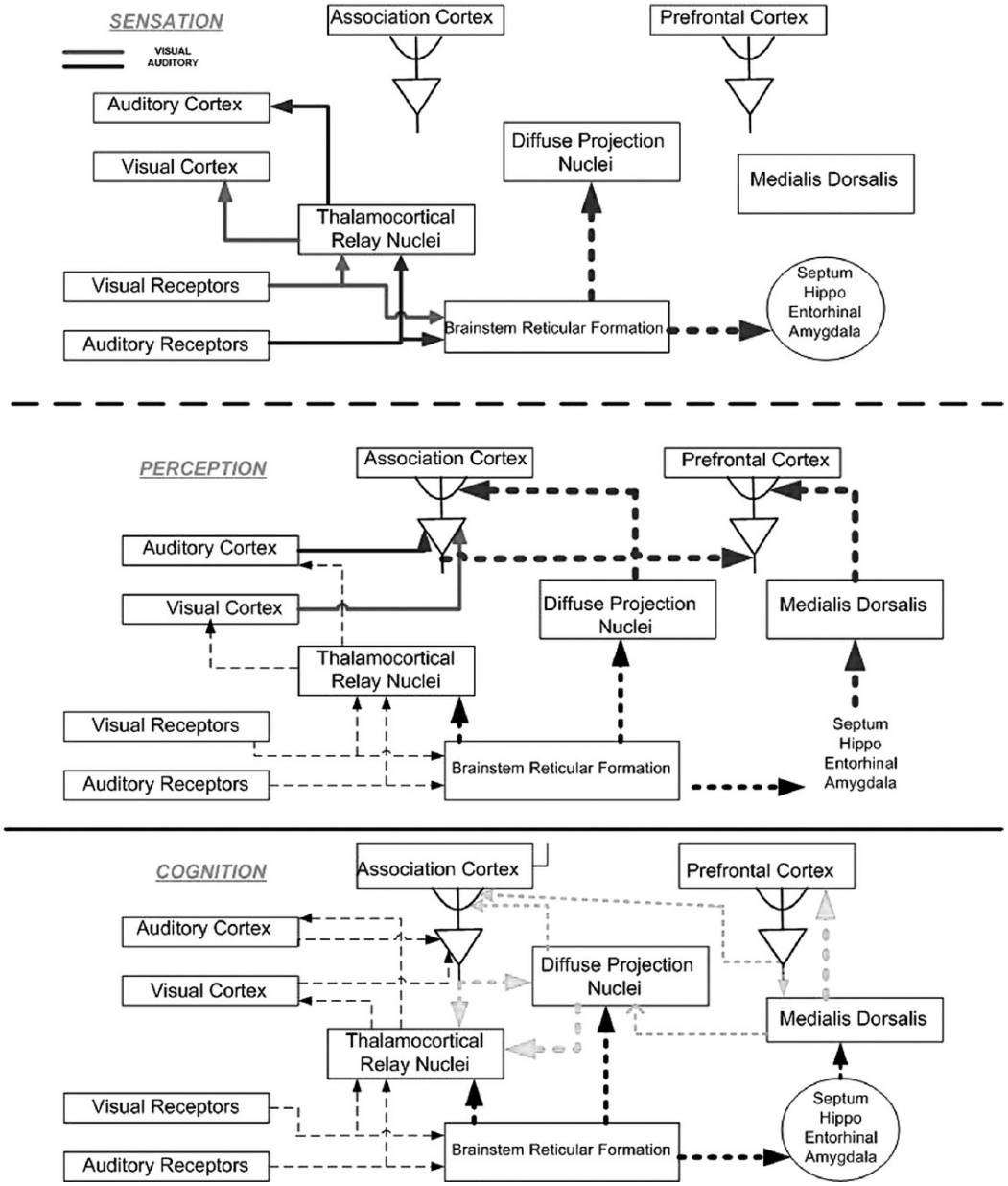

FIGURE 7.7–15. Schema of the stages of construction of conscious awareness and evaluation of percepts as derived from sensations are depicted. (**Top**) Sensations are produced by unimodal information (*dotted lines*) propagating from specific thalamocortical relay nuclei to dispersed sensory-specific cortical receiving areas, with distribution (*bold, broken lines*) into nonspecific thalamic and limbic system components of the endogenous system via the reticular formation. (**Middle**) Perception requires the coincidence of exogenous inputs (*dotted lines*) to basal dendrites with inputs from the endogenous system (*bold, broken lines*) to apical dendrites of pyramidal cells in association and frontal cortical regions. (**Bottom**) Cognition is mediated by the binding together of activated dispersed cortical regions with activated specific and nonspecific thalamic nuclei in a coherent, phase-locked, reverberating gamma oscillation (*light gray, broken lines*).

Conclusions

This section has provided an overview of the methods of QEEG and a summary of evidence of its sensitivity and specificity for detecting pathophysiological conditions in a variety of psychiatric disorders and for

identifying treatment responsive and outcome-related subtypes within disorders. It has also called attention to the powerful brain imaging capabilities using QEEG source localization, allowing imaging of the mathematically most probable underlying sources of the scalp-recorded data, greatly expanding the potential to contribute to understanding of the pathophysiology of psychiatric disorders and the clinical utility of QEEG in psychiatric populations. Space did not permit a review of the contribution of sleep EEG, however reviews of this literature exist. This chapter has also presented a theoretical framework that the authors believe is pertinent to an understanding of how the measurements available with QEEG can be used to gain insights into the underlying neurophysiological factors that may be contributory to the maladaptive behavior and emotional distress suffered by patients with a psychiatric/neuropsychiatric disorder. These symptoms can be thought of as reflecting perturbations of one or another aspect of the system that the authors have proposed to mediate the subjective content of conscious experience. The authors believe that systematic interactions between the symptoms experienced by the patient, the observations of the psychiatrist, and the careful evaluation and interpretation of functional assessments, obtained using readily available QEEG and source localization imaging technology evaluated relative to normative data, can focus attention upon disturbances of equilibrium in the brain. This in turn can lead to evidence-guided hypotheses about appropriate most optimal modes of treatment, with iterated quantitative neurobiological as well as psychiatric assessment of the appropriateness and efficacy of the treatment.

References

Arns M, Conners CK, Kraemer HC. A decade of EEG theta/beta ratio research in ADHD: a meta-analysis. *J Atten Disord*. 2013;17(5):374–383.

Bosch-Bayard J, Valdes-Sosa P, Virues-Alba E, Aubert-Vazquez E, John ER. 3D Statistical parametric mapping of EEG source spectra by means of variable resolution electromagnetic tomography (VARETA). *Clin Electroencephalogr*. 2001;32:47–61.

Bruckner RL, Andrews-Hann J, Schacter D. The brain's default network anatomy, function and relevance to disease. *Ann N Y Acad Sci*. 2008;1124:1–38.

Buzsaki G, Schomburg E. What does gamma coherence tell us about inter-regional neural communication? *Nat Neurosci*. 2015;18:484–489.

Chabot RJ, di Michele F, Prichep LS. The role of quantitative electroencephalography in child and adolescent psychiatric disorders. *Child Adolesc Psychiatr Clin N Am*. 2005;14:21–53.

Coburn KA, Lauterbach EL, Boutros N, Arciniegas D, Black K. The value of quantitative electroencephalography in clinical psychiatry. *J Neuropsychiatry Clin Neurosci*. 2006;18:460–500.

Di Lorenzo G, Daverio A, Ferrentino F, et al. Altered resting-state EEG source functional connectivity in schizophrenia: the effect of illness duration. *Front Hum Neurosci*. 2015;1–11.

Galderisi S, Vignapiano A, Mucci A, Boutros N. Physiological correlates of positive symptoms in schizophrenia. *Electrophysiol Psychophysiol Psych Psychopharm*. 2014;21:103–128.

Goossens J, Laton J, Van Schependom J, Gielen J, Struyfs H, Van Mossevelde S, Van den Bossche T, Goeman J, De Deyu PP, Sieben A, Martin JJ, Van Broeckhoven C, van der Zee J, Engelborghs S, Nagels G. EEG dominant frequency peak differtiates between Alzheimer's Disease and Frontotemporal Lobar Degeneration. *J Alz Dis*. 2017;53–58.

Hughes JR, John ER. Conventional and quantitative electroencephalography in psychiatry. *J Neuropsychiatry Clin Neurosci*. 1999;11:190–208.

Jobert M, Wilson FJ, Ruigt GSF, Brunovsky M, Prichep LS, Drinkenburg WHIM. Guidelines for the recording and evaluation of pharmaco-EEG data in man: the International Pharmaco-EEG Society (IPEG). *Neuropsychobiology*. 2012;66(4):201–220.

John ER. Consciousness: from synchronous neuronal discharges to subjective awareness? *Prog Brain Res*. 2005;150:143–171.

John ER, Prichep LS. The anesthetic cascade: how anesthesia suppresses awareness. *Anesthesiology*. 2005;102:447–471.

John ER, Prichep LS, Friedman J, Easton P. Neurometrics: computer-assisted differential diagnosis of brain dysfunctions. *Science*. 1988;293:162–169.

John ER, Prichep LS, Winterer G, Hermann WM, di Michele F. Electrophysiological subtypes of psychotic states. *Acta Psychiatr Scand*. 2007;116:17–35.

Lehmann D, Faber PL, Galderisi S, Herrmann WM, Kinoshita T. EEG microstate duration and syntax in acute, medication-naive, first-episode schizophrenia: A multi-center study. *Psychiatry Res*. 2005;138:141–156.

Leuchter A, Cook I, Jin Y, Phillips B. The relationship between brain oscillitory activity and therapeutic effectiveness of transcranial magnetic stimulation in the treatment of major depressive disorder. *Front Hum Neurosci*. 2013;7 (37):1–12.

Llinas R, Ribary U, Contreras D, Pedroarena C. The neuronal basis for consciousness. *Philos Trans R Soc Lond B Biol Sci*. 1998;353:1841–1849.

Llinas R, Urbano FJ, Leznik E, Ramirez RR, van Marle HJ. Rhythmic and dysrhythmic thalamocortical dynamics: GABA systems and the edge effect. *Trends Neurosci*. 2005;28(6):325–333.

Loo SK, Makeig S. Clinical utility of EEG in attention-deficit/hyperactivity disorder: a research update. *Neurotherapeutics*. 2012;9(3):569–587.

Lopes Da Silva FH. Dynamic of electrical activity of the brain, networks, and modulating systems. In: Nunez PL, ed. *Neocortical Dynamics and Human EEG Rhythms*. New York: Oxford University Press; 1995:249.

Pascual-Marqui RD, Esslen M, Kochi K, Lehmann D. Functional imaging with low resolution brain electromagnetic tomography (LORETA): review, new comparisons, and new validation. *Jpn J Clin Neurophysiol*. 2002;30:81.

Pizzagalli D, Pascual-Marqui R, Nitschke JB, Oakes TR, Larson CL. Anterior cingulate activity as a predictor of degree of treatment response in major depression: Evidence from brain electrical tomography analysis. *Am J Psychiatry*. 2001;158:405–415.

Prichep LS, Jacquin A, Filipenko J, et al. Classification of traumatic brain injury severity using informed data reduction in a series of binary classification algorithms. *IEEE Trans Neural Syst*

Rehabil Eng. 2012;20(6):806–822.

Prichep LS, John ER, Ferris SH, Rausch L, Fang Z. Prediction of longitudinal cognitive decline in normal elderly with subjective complaints using electrophysiological imaging. *Neurobiol Aging*. 2006;27:471–481.

Prichep LS, Mas F, Hollander E, Liebowitz M, John ER. Quantitative electroencephalographic (QEEG) subtyping of obsessive compulsive disorder. *Psychiatr Res*. 1993;50:25–32.

Prichep LS, Naunheim R, Bazarian J, Mould WA, Hanley D. Identification of hematomas in mild traumatic brain injury using an index of quantitative brain electrical activity. *J Neurotrauma*. 2015;32:17–22.

Saletu B, Anderer P, Saletu-Zyhlarz GM, Pascual-Marqui RD. EEG mapping and low-resolution brain electromagnetic tomography (LORETA) in diagnoses and therapy of psychiatric disorders: evidence for a key-lock principle. *Clin EEG Neurosci*. 2005;36:108–115.

Saletu B, Anderer P, Saletu-Zyhlarz GM. EEG topography and tomography (LORETA) in diagnosis and pharmacotherapy of depression. *Clin EEG Neurosci*. 2010;41(4):203–210.

Thatcher RW, North DM, Biver CJ. LORETA EEG phase reset of the default mode network. *Front Hum Neurosci*. 2014;8(529)41–14.

Thibodeau R, Jorgensen RS, Kim S. Depression, anxiety, and resting frontal EEG asymmetry: a meta-analytic review. *J Abnorm Psychol*. 2006;115(4):715–729.

---

*Throughout the work on this chapter, Leslie Prichet was Professor of Psychiatry and Director of the Brain Research Laboratories, Department of Psychiatry, New York School of Medicine, New York, New York.

▲ 7.8 Psychiatric Rating Scales

SAMUEL JUSTIN SINCLAIR, PH.D., MARK A. BLAIS, PSY.D., AND DEBORAH BLACKER, M.D., SC.D.

The term *psychiatric rating scales* encompasses a variety of questionnaires, interviews, checklists, outcome assessments, and other instruments that are available to inform psychiatric practice, research, and administration. Although new instruments and alternative versions of older instruments (especially shorter, self-report, or telephone versions) are developed regularly to improve quality and efficiency, to update diagnostic criteria, and to measure emerging constructs, they tend to be adopted slowly, as evidence of their usefulness accumulates. Older measures slowly fall out of favor as they become outdated and are supplanted by newer measures. Psychiatrists must keep up with major developments in rating scales for several reasons. Most critically, many such scales are useful in psychiatric practice for monitoring patients over time or for providing information that is more comprehensive than what is generally obtained in a routine clinical interview. In addition, health care administrators and

payors are increasingly requiring standardized assessments to demonstrate the efficacy of psychiatric services or to assess quality of care. Last, but equally important, rating scales are used in research that informs the practice of psychiatry, so familiarity with them provides a deeper understanding of the results of that research and the degree to which it applies to psychiatric practice.

POTENTIAL BENEFITS AND LIMITATIONS OF RATING SCALES IN PSYCHIATRY

In its latest release of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), the American Psychiatric Association (APA) advocates for the use of psychometric tools in clinical practice as a means of improving quality of care at multiple levels. These new "emerging measures" are presented in Section III of the DSM-5, and include tools for both adults and children that are designed to assess an array of clinical domains (e.g., depression, anxiety, mania, etc.), which may then call attention to potential areas of difficulty and guide the clinical evaluation process. Likewise, more in-depth psychometric tools are also included to provide for more comprehensive assessments of specific domains (e.g., depression, mania, substance use), which then can be used to track symptom profiles over time. APA discusses a number of advantages of using these types of assessment tools, including the improvement of clinical precision and decision-making at the point of initial evaluation, a standardized method/metric for evaluating treatment efficacy over time, and finally as a framework for tracking patient reported outcomes (PRO) more globally.

The key role of rating scales in psychiatry and elsewhere is to standardize the information collected across time and observers. This standardization ensures a consistent and comprehensive evaluation that may aid treatment planning by establishing a diagnosis, assuring a thorough review of symptoms, identifying comorbid conditions, and characterizing other factors affecting treatment response. In addition, the use of a rating scale can establish a baseline for following the progression of an illness over time or in response to specific interventions. This is particularly useful when more than one clinician is involved—for instance, in a hospital, group practice or research setting.

In addition to standardization, most rating scales provide users with empirical information on the measure's performance characteristics. This allows the clinician to know to what extent a given scale produces reproducible results (reliability) and how it compares to more established ways of measuring the same construct (validity).

Rating scales also offer some practical advantages. First, they offer carefully worded, pretested methods for obtaining information that can save valuable physician time—self-administered rating scales can be completed in the waiting room, or a nurse or technician can administer an

interview before a session with the physician. In addition, rating scales may make it easier to obtain information about sensitive areas, such as cognitive decline or sexual side effects, in which direct questioning may be experienced as intrusive.

It is important to realize, however, that rating scales are not a panacea. They can provide erroneous measurements due to difficulties in administration, measurement error (or statistical imprecision), or limitations in the underlying construct. In this respect, they are no different than clinical assessments, but may appear to provide more definitive information and thus a spurious sense of security. At the practical level, they take time that might be better devoted to other clinical activities. The critical decision when deciding to use a formal assessment tool in clinical practice is whether, on balance, it contributes useful information in an efficient manner. This decision depends on the specific clinical setting and goal, the practical attributes of the scale, and its psychometric properties.

TYPES OF SCALES AND WHAT THEY MEASURE

Psychometric tools are used in psychiatric research and practice to achieve a variety of goals. They also cover a wide range of areas and use a broad array of procedures and formats.

Measurement Goals

Most psychiatric rating scales fall into one or more of the following categories: Making a diagnosis (e.g., the Structured Clinical Interview for the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* [DSM-5] [SCID] or the Diagnostic Interview Schedule for Children [DISC]); measuring severity and tracking change in specific symptoms (e.g., the Hamilton Anxiety Rating Scale [HAM-A] or the Mini-Mental State Examination [MMSE]), in general functioning (e.g., the Short Form 36 [SF-36]), or in overall outcome (e.g., the Behavior and Symptom Identification Scale [BASIS-24]); and screening for conditions that may or may not be present (e.g., the CAGE).

Constructs Assessed

Psychiatric practitioners and investigators assess a broad range of areas, referred to as *constructs,* to underscore the fact that they are not simple, direct observations of nature. These include diagnoses, signs and symptoms, severity, functional impairment, quality of life, and many others. Some of these constructs are fairly complex and are divided into two or more domains (e.g., positive and negative symptoms in schizophrenia, or mood and neurovegetative symptoms in major depression). Many scales yield separate scores, or *subscales,* for each domain. In cases in which these domains are seen as substantially

independent, they may be referred to as *dimensions* (e.g., the DSM-5 hybrid dimensional-categorical model proposed in Section III).

Categorical versus Continuous Classification. Some constructs are viewed as *categorical* or classifying, whereas others are seen as *continuous* or linear. Categorical constructs describe the presence or absence of a given attribute (e.g., whether someone meets criteria for depression or not), or the category best suited to a given individual among a finite set of options (e.g., assigning a diagnosis). Continuous measures provide a quantitative assessment along a continuum of intensity, frequency, or severity (e.g., the magnitude or degree of someone's level of depression). In addition to symptom severity and functional status, multidimensional personality traits, cognitive status, social support, and many other attributes are generally measured continuously.

The distinction between categorical and continuous measures is by no means absolute. *Ordinal* classification, which uses a finite, ordered set of categories (e.g., unaffected, mild, moderate, or severe) stands between the two. In addition, a cutoff is frequently used with a continuous or ordinal scale to indicate a threshold for membership in a corresponding category. For instance, individuals with MMSE scores below 24 may be considered to have a dementia, or those with Hamilton Rating Scale for Depression (HAM-D) scores above 8 may be considered to have an episode of major depression.

Measurement Procedures

Rating scales differ in measurement methods. Issues to be considered include format, raters, and sources of information.

Format. Rating scales are available in a variety of formats. Some are simply checklists or guides to observation that help the clinician achieve a standardized rating. Others are self-administered questionnaires or tests. Still others are formal interviews that may be *fully structured* (i.e., specifying the exact wording of questions to be asked) or *semistructured* (i.e., providing only some specific wording, along with suggestions for additional questions or probes). Whether fully structured or not, instruments may be written such that all questions are always included, or they may have formal skip-out sections to limit administration time.

Individual items also vary in format. Most commonly, scales use questions with yes–no or multiple-choice response options. Often, answers will be graded on a *Likert* scale, an ordinal scale with three to seven points that measures severity, intensity, frequency, or other attributes. Likert scales are most often partially or fully *anchored,* assigning a meaning to each numeric level. The same anchors can apply to all of the items, or the instrument may provide specific anchors for each. Occasionally,

questionnaires include open-ended questions, especially at the beginning, which may be used to help establish rapport. In semistructured or unstructured interviews, this information also serves to guide the rest of the interview and to aid in forming a clinical impression about the patient.

Raters. Some instruments are designed to be administered by doctoral-level clinicians only, whereas others may be administered by psychiatric nurses or social workers with more limited clinical experience. Still other instruments are designed primarily for use by lay raters with little or no experience with psychopathology. In general, more training is required to administer less structured scales. In addition, some scales require extensive training, even for experienced clinicians, to master the appropriate procedures and achieve a good result. Virtually all scales perform better when raters are familiar with their format and specific content.

Source of Information. Instruments also vary in the source of information used to make the ratings. Information may be obtained solely from the patient, who generally knows the most about his or her condition. In some instruments, some or all of the information may be obtained from a knowledgeable informant. When the construct involves limited insight (e.g., cognitive disorders or mania) or significant social undesirability (e.g., antisocial personality or substance abuse), other informants may be preferable. Informants may also be helpful when the subject has limited ability to recall or report symptoms (e.g., delirium, dementia, or any disorder in young children). Some rating scales also allow or require information to be included from medical records or from patient observation.

ASSESSMENT OF RATING SCALES

In clinical research, rating scales are mandatory to ensure interpretable and potentially generalizable results, and are selected based on coverage of the relevant constructs, expense (based on the nature of the raters, purchase price if any, and necessary training), length and administration time, comprehensibility to the intended audience, and quality of the ratings provided. In clinical practice, one considers these factors and, also, whether a scale would provide more or better information than what would be obtained in ordinary clinical practice or would contribute to the efficiency of obtaining that information. In either case, the assessment of quality is based on *psychometric,* or mind-measuring, properties.

Psychometric Properties

The two principal psychometric properties of a measure are *reliability* and *validity*. Although these words are used almost interchangeably in

everyday speech, they are distinct in the context of evaluating rating scales. To be useful, scales should be *reliable,* or consistent and repeatable even if performed by different raters at different times or under different conditions; and they should be *valid,* or accurate in representing the true state of nature.

Relationship between Reliability and Validity. Establishing a measure's reliability is generally considered primary, as it is difficult to reach valid judgments without first achieving consistency. However, problems with reliability can be overcome to an extent by combining information from several assessments. Unfortunately, improved reliability does not guarantee improved validity, and some efforts to improve reliability may actually limit validity. For example, a personality disorder instrument might focus on overt behaviors rather than inner thoughts and feelings to achieve higher reliability, but at the cost of losing some of the most valid information about personality. Even with clinically trained raters, it is particularly difficult to achieve reliability on items requiring subjective clinical judgment (e.g., feelings evoked in the examiner). Nonetheless, when used by experienced diagnosticians, such items may contribute substantially to valid diagnoses.

Reliability. *Reliability* refers to the consistency or repeatability of ratings and is largely empirical. In the categorical context, it refers to whether there is agreement on the classification of each individual. In the continuous context, it refers to whether there is agreement on the assignment of a given score. It can also be seen as precision (i.e., whether a measure yields a "ballpark" estimate or a finely graded score). An instrument is more likely to be reliable if the instructions and questions are clearly and simply worded, and the format is easy to understand and score. There are three standard ways to assess reliability: *Internal consistency, interrater,* and *test–retest.*

INTERNAL CONSISTENCY. Internal consistency assesses agreement among the individual items in a measure. This provides information about reliability, because each item is viewed as a single measurement of the underlying construct. Thus, the coherence of the items suggests that each is measuring the same thing. Internal consistency is measured most often with *coefficient alpha* (also known as *Cronbach's coefficient alpha*), which ranges between 0 and 1—values of 0.70 are typically recommended when evaluating or making decisions about groups of people, whereas more stringent values (usually exceeding 0.85 or 0.90) are suggested when evaluating individuals. However, it should be noted that the internal consistency of a *measure* depends on the internal consistency of the *construct* that the measure purports to assess and will be higher for

unidimensional constructs than for those with two or more relatively independent domains. For multidimensional constructs, it is generally preferable to calculate coefficient alpha on individual subscales.

INTERRATER AND TEST–RETEST RELIABILITY. Interrater (also called *interjudge* or *joint*) reliability is a measure of agreement between two or more observers evaluating the same subjects using the same information. Estimates may vary with assessment conditions—for instance, estimates of interrater reliability based on videotaped interviews tend to be higher than those based on interviews conducted by one of the raters. Interrater reliability tends to be higher than *test–retest reliability,* a measure of agreement between evaluations at two points in time in which there may be differences in the information obtained (e.g., associated with differences in interviewer skill, interviewer mood, room conditions, or subject's attitude). In addition, test–retest evaluations measure reliability only to the extent that the subject's true condition remains stable in the time interval, which is problematic for many conditions but virtually impossible for rapidly fluctuating conditions like state anxiety. However, the test–retest situation more closely reflects the clinical problems associated with serial evaluations by multiple clinicians, so to the extent that concerns about interval change can be eliminated, it is generally a more useful indicator of reliability in practice.

Interrater reliability and test–retest reliability of continuous constructs are measured with the intraclass correlation coefficient (ICC), whereas those of categorical constructs are measured with the kappa coefficient (κ). A weighted version of κ is available to give greater penalties for large disagreements than for small ones (e.g., between schizophrenia and psychotic depression, as compared to schizophrenia and schizoaffective disorder). Both κ and the ICC are measures of agreement corrected for the agreement expected by chance alone, and both range from 0 to 1. As a rule of thumb, a κ or ICC above 0.8 is considered excellent, those in the 0.7 to 0.8 range are considered good, and those in the 0.6 to 0.7 range are considered fair. However, the degree of reliability required varies with the clinical purpose: Extremely reliable ratings are required to measure small changes in critical values (e.g., to detect small effects in a clinical trial or to decide when to stop a potentially toxic therapy), whereas more modest reliability may be satisfactory for other purposes (e.g., to detect large effects in a clinical trial or to estimate population means or rates).

ISSUES IN INTERPRETING RELIABILITY DATA. When interpreting reliability data, it is important to bear in mind that reliability estimates published in the literature may not generalize to other settings. Factors to consider are the nature of the sample, the training and experience of the raters, and the test conditions. Issues regarding the sample are especially critical. In particular,

reliability tends to be higher in samples with high variability in which it is easier to make discriminations among individuals. Thus, for continuous measures, reliability tends to be higher when the sample includes individuals with a wide range of scores; and for categorical measures, reliability tends to be higher when the prevalence of the attribute being measured approaches 50 percent. Reliability estimates also depend on the fraction of difficult cases (e.g., individuals near a diagnostic threshold or those resistant to being interviewed), as large numbers of these tend to diminish observed reliability.

Validity. *Validity* refers to conformity with truth, or a gold standard that can stand for truth. In the categorical context, it refers to whether an instrument can make correct classifications. In the continuous context, it refers to accuracy, or whether the score assigned can be said to represent the true state of nature. Although reliability is an empirical question, validity is partly theoretical—for many constructs measured in psychiatry, there is no underlying absolute truth. Even so, some measures yield more useful and meaningful data than others do. Validity assessment is generally divided into face and content validity, criterion validity, and construct validity.

FACE AND CONTENT VALIDITY. *Face validity* refers to whether the items appear to assess the construct in question. Although a rating scale may purport to measure a construct of interest, a review of the items may reveal that it embodies a very different conceptualization of the construct. For instance, an insight scale may define *insight* in either psychoanalytic or neurological terms. However, items with a transparent relationship to the construct may be a disadvantage when measuring socially undesirable traits, such as substance abuse or malingering. *Content validity* is different than face validity in that it describes whether the measure provides good balanced coverage of the construct, and is less focused on whether the items give the appearance of validity. Content validity is often assessed with formal procedures such as expert consensus or factor analysis.

CRITERION VALIDITY. *Criterion validity* (sometimes called *predictive* or *concurrent validity*) refers to whether or not the measure agrees with a gold standard or criterion of accuracy. Suitable gold standards include the long form of an established instrument for a new, shorter version; a clinician-rated measure for a self-report form; and blood or urine tests for measures of drug use. For diagnostic interviews, the generally accepted gold standard is the *L*ongitudinal, *E*xpert, *A*ll *D*ata (LEAD) standard, which incorporates expert clinical evaluation, longitudinal observations, medical records, family history, and any other sources of information.

When continuous measures are compared to a gold standard, a correlation coefficient is the statistic most often reported. For categorical

variables, such as diagnoses (or continuous measures with a cutoff), various types of diagnostic efficiency statistics are employed depending on the goal of the study—including *sensitivity* and *specificity* (which are most commonly used). *Sensitivity* refers to the test's ability to identify true cases, or its *true-positive rate*. Specificity is the test's accuracy in identifying noncases, or one minus the *false-positive rate*. In general, the more sensitive a test, the less specific it is. If the threshold for diagnosis, for example, is lowered, then more cases will be detected but at the expense of some false positives; if the threshold is raised to decrease the false positives, then true cases will inevitably be missed. The optimal threshold depends on the consequences of false positives and false negatives. Calculation of predictive value positive and negative, which report, respectively, the proportion of positive and negative assessments that are correct *in a given population,* may help to answer this question. Predictive value positive, for instance, can be low even with excellent specificity in a low prevalence population, for instance, when measuring psychotic disorders in the general population. Where sensitivity is low, predictive value negative may be reduced, but less so if prevalence is also low.

CONSTRUCT VALIDITY. When an adequate gold standard is not available—a frequent state of affairs in psychiatry—or when additional validity data are desired, construct validity must be assessed. To accomplish this, one can compare the measure to *external validators,* attributes that bear a well-characterized relationship to the construct under study but are not measured directly by the instrument. External validators used to validate psychiatric diagnostic criteria and the diagnostic instruments that aim to operationalize them include course of illness, family history, and treatment response. For example, when compared to schizophrenia measures, mania measures are expected to identify more individuals with a remitting course, a family history of major mood disorders, and a good response to lithium.

Two special cases of assessing validity using external validators have particular relevance for clinical psychiatry. One is *discriminant validity,* which examines a measure's ability to discriminate between populations that are expected to differ on the construct of interest. For example, does a sociopathy measure correctly separate individuals in jails from those living in the community? Although such discriminations are important in clinical practice, it is important to remember that the true test of a measure is its ability to discriminate at the margins. A study of discriminant validity is more clearly relevant if it includes the types of cases encountered in clinical practice (e.g., psychotic depression versus schizoaffective disorder) rather than more easily discriminated populations (e.g., psychotic depression versus normal). Another special case is *sensitivity to change:* The fact that a measure shows expected changes (e.g., an improvement with an efficacious treatment or a decrement with a progressive disease) can be a

strong validator.

When validity is assessed in areas in which there are few established measures and a gold standard or criterion of accuracy cannot be established, the assessment of the validity of the measure is limited by the validity of the construct itself. Nonetheless, by triangulation between a better definition of the construct, better ways to measure it, and better exploration of how it operates in clinical practice and research, the field moves to greater validity over time. Cronbach and Meehl (1955) referred to this process as defining the *nomological net* of constructs to be measured and ways in which they are thought to be interrelated, which then provides a framework for evaluating and establishing construct validity.

SELECTION OF PSYCHIATRIC RATING SCALES

In this section, a group of rating scales used in psychiatric practice and research is presented. The scales are grouped by topic, beginning with general issues, including diagnosis, functioning, symptom severity, and side effects, and then proceeding to specific diagnostic groups, organized according to the sections of the DSM-5. Selections were made based on coverage of major areas and common use in clinical research or current (or potential) use in clinical practice; instruments specifically recommended in DSM-5 were included whenever possible. For each area, there is a brief discussion of measurement issues followed by a description of each instrument, its psychometric properties, and its potential uses. In addition, whenever possible, a brief, clinically useful instrument is provided in each area. References for each measure, organized by topic, are listed in Table 7.8–1. These references include more detailed information about each measure and its psychometric properties and may also provide either the measure itself or instructions regarding how to obtain a copy. Access to instruments via the World Wide Web is increasing rapidly, but Internet addresses evolve still more rapidly, so they are not included here. Instead, a quick search using Google or a similar search engine is recommended to locate a website that may offer the instrument for purchase or download, an up-to-date list of references, and other information.

Functional Status, Impairment, and General Symptom Severity

This broad area cuts across a variety of diagnoses and is thus useful for grading patients by functional status or overall severity without reference to specific symptomatology. Most of the instruments presented here have a strong mental health focus and often include items on psychiatric symptomatology.

The World Health Organization Disability Assessment Schedule. The World Health Organization Disability Assessment Schedule (WHODAS 2.0) was published in 2010 by the World Health Organization (WHO), and

is a 36-item, self-administered measure of disability across six distinct domains including: understanding and communicating, getting around, self-care, getting along with people, life activities (i.e., household, work, and/or school activities), and participation in society. An informant (proxy) version of the tool is also available. Each of the 36 items is rated along a five-point Likert scale (1. None, 2. Mild, 3. Moderate, 4. Severe, 5. Extreme), and then items are summed within the six domains. Scores are linearly transformed to a 0 to 100 metric, where 0 represents no disability and 100 represents maximum disability. Based on DSM-5 field trials of the WHODAS 2.0 in adult populations, the APA has recommended that *average* scores for each domain be used as a means of helping clinicians to think about disability status in terms of the five anchor (scaling) points noted above (no disability, mild, moderate, severe, and extreme disability). This scoring method was found to be reliable, clinically useful, and with promising validity. It also provides clinicians with an alternative method for quantifying level of functioning (or degree of impairment) related to psychiatric symptoms given that DSM-IV's Global Assessment of Functioning (GAF) scale was not retained in DSM-5.

Behavior and Symptom Identification Scale. The Behavior and Symptom Identification Scale (BASIS-32) was developed in the early 1990s to provide a broad but brief overview of psychiatric symptoms and functional status from the patient's point of view for use in assessing the outcome of psychiatric treatment. A revised version, the BASIS-24 (Table 7.8–2), aimed at making the instrument shorter and more broadly applicable across demographic and diagnostic groups and in a wider variety of treatment settings, was released in 2004. The instrument assesses a wide range of areas, including family and work relationships; ability to complete regular tasks at home, work, or school; and symptoms of anxiety, depression, psychosis, and substance abuse. Each item is rated on a five-point scale focused on the degree of difficulty the individual has had in a given area during the preceding week. The BASIS-24 can be completed as a paper-and-pencil test (requiring 5 to 20 minutes), or the questions can be read aloud with the patient selecting the best answer from a laminated card (requiring 15 to 20 minutes). The BASIS-24 is scored by computer and generates an overall score, and five subscales: relation to self and others, daily living and role functioning, depression and anxiety, impulsive and addictive behavior, and psychosis. Good reliability and validity have been demonstrated. Its simple administration, brevity, and broad coverage make it well suited to its original task, and it is frequently used at baseline, during treatment, and after treatment to monitor progress. It is capable of providing valid ratings across a wide range of psychiatric impairment but is not generally suitable for individuals with substantial cognitive impairment. It is also not suitable for children under 14 years of age.

Short Form 36. The SF-36 was developed in 1993 as a general measure of health status from the patient's point of view that would be independent of specific diseases and equally applicable to medical and psychiatric patients and to the general population, and is part of a broader family of instruments designed for this purpose. The SF-36 consists of 36 items, a mixture of yes–no questions and short Likert scales. It is focused on the individual's functional status as it relates to physical problems, pain, and emotional difficulties over the last 4 weeks. An acute version, focused on the past week, is also available, as are 8- and 12-item versions; however, the 4-week, 36-item version is the most widely used. The items are used to rate eight separate subscales: physical function, physical role function, pain, general health, vitality, social function, emotional role function, and mental health. Computer programs are available for administration and scoring. It is designed as a self-report instrument but can easily be read by an interviewer in person or over the phone. The instrument can be completed in 10 to 15 minutes. It has been subjected to extensive psychometric testing in clinical and general population samples, with extensive norms available. There is demonstrated high internal consistency for the eight subscales and good test–retest reliability over a 2-week period. Validity has been demonstrated by correlation with other functional scales and with various severity-of-illness and burden-of-care scales. The SF-36 is in wide use around the world, especially in health services research, and the instrument has special utility for comparing health outcomes across disparate diagnostic groups. It can also be used in clinical and health administration practice to track outcomes for individual patients or health systems and offers the benefit of patient-centered assessment of treatment and its impact.

Table 7.8–1.
Key References for Measures Included

Rating Scale	Reference
Functional status, impairment, and general symptom severity	
WHODAS	The World Health Organization (WHO) manual for WHO Disability Assessment Schedule (WHODAS 2.0), 2010, Geneva.
BASIS-24[a]	Eisen SV, Normand SLT, Belanger A, Spiro A, Esch D. The revised Behavior and Symptom Identification Scale (BASIS-24): reliability and validity. *Med Care.* 2004;42:1230.
SF-36	McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short Form Health Survey (SF-36), II: Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care.* 1993;31:247.
Side effects	
SAFTEE	Levine J, Schooler NR: SAFTEE. A technique for the systematic assessment of side effects in clinical trials. *Psychopharmacol Bull.* 1986;22:343.
AIMS[a]	Lane RD, Glazer WM, Hansen TE, Berman WH, Kramer SI. Assessment of

	tardive dyskinesia using the Abnormal Involuntary Movement Scale. *J Nerv Ment Dis.* 1985;173:353.
Simpson-Angus Rating Scale for Extrapyramidal Side Effects	Simpson GM, Angus JWS. A rating scale for extrapyramidal side effects. *Acta Psychiatr Scand.* 1970;46(Suppl 212):11.
Psychiatric diagnosis and screening measures	
SCID	First MB, Spitzer RL, Gibbon M, Williams, JBW. *Structured Clinical Interview for DSM-IV Axis I Disorders, Clinician Version (SCID-CV).* Washington, DC: American Psychiatric Press, Inc.; 1996.
DIS	Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule. *Arch Gen Psychiatry.* 1981;38:381.
CIDI	Wittchen H-U. Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *J Psychiatr Res.* 1994;28:57.
PHQ	Spitzer RL, Kroenke K, Williams JBW. Validation and utility of a new self-report version of PRIME-MD: The PHQ primary care study—Primary Care Evaluation of Mental Disorder, Patient Health Questionnaire. *JAMA.* 1999;282:1737.
DSM-5 Self-Rated Level 1 Cross-Cutting Symptom Measure—Adult	American Psychiatric Association (APA), 2013. Retrieved from http://www.psychiatry.org/practice/dsm/dsm5/online-assessment-measures
DSM-5 Parent/Guardian-Rated Level 1 Cross-Cutting Symptom Measure—Child Age 6–17	American Psychiatric Association (APA), 2013. Retrieved from http://www.psychiatry.org/practice/dsm/dsm5/online-assessment-measures
DSM-5 Self-Rated Level 1 Cross-Cutting Symptom Measure—Child Age 11–17	American Psychiatric Association (APA), 2013. Retrieved from http://www.psychiatry.org/practice/dsm/dsm5/online-assessment-measures
Broadband measures of psychological functioning	
MMPI-2	Butcher JN, Dahlstrom WG, Graham JR, Tellegen A, Kaemmer B. *MMPI-2: Minnesota Multiphasic Personality Inventory–2: Manual for administration and scoring.* Minneapolis: University of Minnesota Press; 1989.
PAI	Morey LC. Personality Assessment Inventory: Professional Manual (2nd ed.). Odessa, FL: Psychological Assessment Resources, 2007.
PSI	Blais MB, Sinclair SJ. *Initial Construction, Psychometric Evaluation, and Validation of the Psychopathology Spectra Inventory (PSI): A Measure Based on Quantitative Models of Psychopathology.* Boston, MA: Massachusetts General Hospital; 2014.
Schizophrenia spectrum and other psychotic disorders	
BPRS[a]	Overall JE, Gorham DR. The Brief Psychiatric Rating Scale (BPRS): A comprehensive review. *J Oper Psychiatry.* 1991;148:472.
PANSS	Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull.* 1987;13:261.
SAPS/SANS	Andreasen NC, Arndt S, Miller D, Flaum M, Nopoulos P. Correlational studies of the Scale for the Assessment of Negative Symptoms and the Scale for the Assessment of Positive Symptoms: An overview and update. *Psychopathology.* 1995;28:7.

Depressive disorders	
HAM-D[a]	Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatr.* 1960;23:56.
MADRS	Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry.* 1979; 134:382.
BDI-II	Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation; 1996.
PHQ-9	Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals.* 2002; 32:509–521.
IDS	Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): Psychometric properties. *Psychol Med.* 1996; 26:477.
Bipolar and related disorders	
YMRS	Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: Reliability, validity and sensitivity. *Br J Psychiatry.* 1978;133:429.
ASRM	Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman self-rating mania scale. *Biological Psychiatry.* 1997; 42:948–955.
Anxiety disorders	
HAM-A[a]	Hamilton M. The assessment of anxiety scales by rating. *Br J Psychol.* 1959;32:50.
GAD-7	Spitzer RL, Kroenke K, Williams JBW, Lowe B. A brief measure for assessing generalized anxiety disorder. *Arch Inern Med.* 2006;166:1092–1097.
The PROMIS Emotional Distress—Anxiety Short-Form	The Patient Reported Outcomes Measurement Information System (PROMIS: http://www.nihpromis.org/). The instrument and scoring/interpretation instructions are also available for use at: http://www.psychiatry.org/practice/dsm/dsm5/online-assessment-measures.
PDSS	Shear MK, Brown TA, Barlow DH, Money R, Sholomskas DE. Multicenter collaborative Panic Disorder Severity Scale. *Am J Psychiatry.* 1997;154:1571.
Trauma and stressor related disorders	
CAPS	Blake DD, Weathers FW, Nagy LN, Kaloupe DG, Klauminzer G. A clinician rating scale for assessing current and lifetime PTSD: The CAPS-1. *Behav Therapist.* 1990;18:187.
Obsessive-compulsive and related disorders	
YBOCS	Goodman WK, Price LH, Rasmussen SA, et al. The Yale-Brown Obsessive Compulsive Scale: I. Development, use, and reliability. *Arch Gen Psychiatry.* 1989;46:1006.
RTBS	Goodman WK, Storch E. LEVEL 2—Repetitive Thoughts and Behaviors—Adult (Adapted from the Florida Obsessive-Compulsive Inventory [FOCI] Severity Scale [Part B], 1994. Retrieved from http://www.psychiatry.org/practice/dsm/dsm5/online-assessment-measures
Substance use disorders	
Level 2 Substance Use Module from DSM-5	National Institute on Drug Abuse (NIDA). Adapted from the NIDA-modified ASSIST. Retrieved from http://www.psychiatry.org/practice/dsm/dsm5/online-assessment-measures
CAGE	Ewing JA. Detecting alcoholism: The CAGE questionnaire. *JAMA.* 1984;252:1905.
AUDIT[a]	Saunders JB, Aasland OG, Babor TF, de le Fuente JR, Grant M. Development

	of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction.* 1993;88:791.
DAST	Skinner HA. The Drug Abuse Screening Test. *Addict Behav.* 1982;7:363.
ASI	McLellan AT, Kushner H, Metzger D, Peters F, Peters R. The fifth edition of the Addiction Severity Index. *J Subst Abuse Treat.* 1992;9:199.
Eating disorders	
EDE	Cooper Z, Fairburn CG. The Eating Disorder Examination: A semistructured interview for the assessment of the specific psychopathology of eating disorders. *Int J Eat Disord.* 1987;6:1.
EDA-5	Sysko R, Glasofer DR, Hildebrandt T, et al. The Eating Disorder Assessment for DSM-5 (EDA-5): development and validation of a structured interview for feeding and eating disorders. *Int J Eat Disorder* (in press).
Cognitive disorders	
MMSE	Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12:189.
MOCA	Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment (MoCA©): A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc.* 2005;53:695–699.
CDR[a]	Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*. 1993;43:2412.
NPI	Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA. The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. *Neurology.* 1994;44:2308.
CAM	Inouye SK, VanDyck CH, Alessi CA, et al. Clarifying confusion: The Confusion Assessment Method. A new method for detecting delirium. *Ann Intern Med.* 1990;113:941–948.
Personality disorders	
SCID-II	First MB, Spitzer RL, Gibbon M, Williams JBW. The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II): Part I: Description. *J Personal Disord.* 1995;9:83.
PDQ	Hyler SE, Rieder RO, Williams JBW, Spitzer RL, Hendler J. The Personality Diagnostic Questionnaire: Development and preliminary results. *J Personal Disord.* 1988;2:229.
PID-5	Krueger RF, Derringer J, Markon KE, Watson D, Skodol AE. Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine,* 2012;42:1879–1890.
Childhood disorders	
CBCL[a]	Achenbach TM. *Manual for the Child Behavior Checklist/4-18 and 1991 Profile.* Burlington: University of Vermont; 1991.
DISC	King C, Katz S, Ghaziuddin N, Brand E, Hill E. Diagnosis and assessment of depression and suicidality using the NIMH Diagnostic Interview Schedule for Children (DISC-2.3). *J Abnorm Child Psychol.* 1997;25:173.
Conners rating scales	Conners CK, Wells KC, Parker JD, Sitarenios G, Diamond JM. A new self-report scale for assessment of adolescent psychopathology: Factor structure, reliability, validity, and diagnostic sensitivity. *J Abnorm Child Psychol.* 1997;25:487.
Autism diagnostic interview—revised	Lord C, Rutter M, Le Courteur A. Autism Diagnostic Interview Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord.* 1994;24:659.

Note: For full name of scales denoted by these abbreviations, please see relevant section of this chapter.

Side Effects

The instruments described below are used to detect and quantify side effects from medications. One is a general interview that is applicable to all medications. The remaining two (the Abnormal Involuntary Movement Scale [AIMS] and the Simpson-Angus Rating Scale for Extrapyramidal Side Effects) are focused on the motor symptoms commonly observed with antipsychotics and, for this reason, include a brief focused physical examination, as well as questions posed to the patient.

Table 7.8–2.
BASIS-24® (Behavior and Symptom Identification Scale)

Instructions to Staff: Please fill in the following information completely.

Client ID: _ _ _ _ _ _ _

HCO ID: _ _ _ _ _ _ _

Admission / Intake Date: _ _/_ _/_ _

Time Point:

1□ Admission/Intake
2□ Mid-treatment
3□ Discharge termination
4□ Post-treatment follow-up

Level of Care:

1□ Inpatient
2□ Outpatient
3□ Partial/day hospital
4□ Residential

Program Type or Unit: _ _

Instructions to Respondents:
This survey asks about how you are feeling and doing in different areas of life. Please check the box to the left of your answer that best describes yourself during the ***PAST WEEK***. Please answer every question. If you are unsure about how to answer, please give the best answer you can.

EXAMPLE: ***During the past week,*** **how much difficulty did you have sleeping?**

0□ No difficulty
1□ A little difficulty
2□ Moderate difficulty
3□ Quite a bit of difficulty
4□ Extreme difficulty

During the PAST WEEK, how much difficulty did you have…

1. **Managing your day-to-day life?**

0□ No difficulty
1□ A little difficulty
2□ Moderate difficulty

$_3$□ Quite a bit of difficulty
$_4$□ Extreme difficulty

2. **Coping with problems in your life?**

$_0$□ No difficulty
$_1$□ A little difficulty
$_2$□ Moderate difficulty
$_3$□ Quite a bit of difficulty
$_4$□ Extreme difficulty

3. **Concentrating?**

$_0$□ No difficulty
$_1$□ A little difficulty
$_2$□ Moderate difficulty
$_3$□ Quite a bit of difficulty
$_4$□ Extreme difficulty

During the PAST WEEK, how much of the time did you…

4. **Get along with people in your family?**

$_0$□ None of the time
$_1$□ A little of the time
$_2$□ Half of the time
$_3$□ Most of the time
$_4$□ All of the time

5. **Get along with people outside your family?**

$_0$□ None of the time
$_1$□ A little of the time
$_2$□ Half of the time
$_3$□ Most of the time
$_4$□ All of the time

6. **Get along well in social situations?**

$_0$□ None of the time
$_1$□ A little of the time
$_2$□ Half of the time
$_3$□ Most of the time
$_4$□ All of the time

During the PAST WEEK, how often did you…

7. **Feel close to another person?**

$_0$□ None of the time
$_1$□ A little of the time
$_2$□ Half of the time
$_3$□ Most of the time
$_4$□ All of the time

8. **Feel like you had someone to turn to if you needed help?**

$_0$□ None of the time
$_1$□ A little of the time
$_2$□ Half of the time
$_3$□ Most of the time
$_4$□ All of the time

9. **Feel confident in yourself?**

$_0$□ None of the time
$_1$□ A little of the time
$_2$□ Half of the time
$_3$□ Most of the time
$_4$□ All of the time

During the PAST WEEK, how much of the time did you…

10. **Feel sad or depressed?**

$_0$□ None of the time
$_1$□ A little of the time
$_2$□ Half of the time
$_3$□ Most of the time
$_4$□ All of the time

11. **Think about ending your life?**

$_0$□ None of the time
$_1$□ A little of the time
$_2$□ Half of the time
$_3$□ Most of the time
$_4$□ All of the time

12. **Feel nervous?**

$_0$□ None of the time
$_1$□ A little of the time
$_2$□ Half of the time
$_3$□ Most of the time
$_4$□ All of the time

During the PAST WEEK, how often did you…

13. **Have thoughts racing through your head?**

$_0$□ Never
$_1$□ Rarely
$_2$□ Sometimes
$_3$□ Often
$_4$□ Always

14. **Think you had special powers?**

$_0$□ Never
$_1$□ Rarely
$_2$□ Sometimes
$_3$□ Often
$_4$□ Always

15. **Hear voices or see things?**

$_0$□ Never
$_1$□ Rarely
$_2$□ Sometimes
$_3$□ Often
$_4$□ Always

16. **Think people were watching you?**

$_0$□ Never
$_1$□ Rarely
$_2$□ Sometimes
$_3$□ Often
$_4$□ Always

17. Think people were against you?

$_0$□ Never
$_1$□ Rarely
$_2$□ Sometimes
$_3$□ Often
$_4$□ Always

During the PAST WEEK, how often did you…

18. Have mood swings?

$_0$□ Never
$_1$□ Rarely
$_2$□ Sometimes
$_3$□ Often
$_4$□ Always

19. Feel short-tempered?

$_0$□ Never
$_1$□ Rarely
$_2$□ Sometimes
$_3$□ Often
$_4$□ Always

20. Think about hurting yourself?

$_0$□ Never
$_1$□ Rarely
$_2$□ Sometimes
$_3$□ Often
$_4$□ Always

During the *PAST WEEK,* how often…

21. Did you have an urge to drink alcohol or take street drugs?

$_0$□ Never
$_1$□ Rarely
$_2$□ Sometimes
$_3$□ Often
$_4$□ Always

22. Did anyone talk to you about your drinking or drug use?

$_0$□ Never
$_1$□ Rarely
$_2$□ Sometimes
$_3$□ Often
$_4$□ Always

23. Did you try to hide your drinking or drug use?

$_0$□ Never
$_1$□ Rarely

2□ Sometimes
3□ Often
4□ Always

24. Did you have problems from your drinking or drug use?

0□ Never
1□ Rarely
2□ Sometimes
3□ Often
4□ Always

ABOUT YOU

25. How old are you?—

26. What is your sex?

1□ Male
2□ Female

27. Are you …

1□ Hispanic or Latino
2□ NOT Hispanic or Latino

28. What is your racial background?

1□ American Indian or Alaskan native
2□ Asian
3□ Black or African-American
4□ White/Caucasian
5□ Native Hawaiian or other Pacific Islander
6□ Multiracial or other (specify)

29. How much school have you completed?

1□ 8th grade or less
2□ Some high school
3□ High school graduate/GED
4□ Some college
5□ 4-year college graduate or higher

30. Are you now …

1□ Married
2□ Separated
3□ Divorced
4□ Widowed
5□ Never married

31. Outside of your treatment providers, what is your main source of social support?

1□ Wife, husband, or partner
2□ Other family (parents, children, relatives)
3□ Friends/roommates
4□ Community/church
5□ Other
6□ No one

32. Where did you sleep in the past 30 days?

1□ Apartment or house

2□ Halfway house/group home/board and care home/residential center/supervised housing
3□ School or dormitory
4□ Hospital or detox center
5□ Nursing home/assisted living
6□ Shelter/street
7□ Jail/prison
8□ Other (fill in)————————

33. At any time in the past 30 days, did you work at a paying job?

1□ No
2□ Yes, 1–10 hours per week
3□ Yes, 11–30 hours per week
4□ Yes, more than 30 hours per week

34. At any time in the past 30 days, did you work at a volunteer job?

1□ No
2□ Yes, 1–10 hours per week
3□ Yes, 11–30 hours per week
4□ Yes, more than 30 hours per week

35. At any time in the past 30 days, were you a student in a high school, job training, or college degree program?

1□ Yes
2□ No

36. Do you now receive disability benefits; for example, SSI, SSDI, or other disability insurance (Check one or more)

1□ No
2□ Yes, I receive disability for medical reasons
3□ Yes, I receive disability for psychiatric reasons
4□ Yes, I receive disability for substance abuse

37. Today's Date: _ _ / _ _ /_ _

THANK YOU VERY MUCH!

Copyright McLean Hospital, 2001.

Reprinted with permission of Mclean Hospital, Belmont, MA; 2002–2014.

Systematic Assessment of Treatment-Emergent Events. The Systematic Assessment of Treatment-Emergent Events (SAFTEE) is a systematic tool developed in 1986 to assess side effects in clinical trials. It has two versions, a general inquiry form (SAFTEE-GI) and a much more detailed specific inquiry form (SAFTEE-SI) that incorporates a formal review of symptoms. The SAFTEE-GI is rated based on three open-ended questions regarding new symptoms experienced in the past week: (1) Have you had any physical or health problems during the past week? (2) Have you noticed any changes in your physical appearance during the past week? (3) Have you cut down on the things you usually do because of not feeling physically well during the past week? To avoid bias on the part of the examiner or patient,

the questions are focused on the novelty of the symptoms rather than on any relationship to the drug. The SAFTEE form provides specified terms (listed by organ system) for reporting the answers to the three questions and suggests specific ways to follow up on any positive answers regarding (1) the duration, temporal pattern, severity, and associated functional impairment for each; (2) any contributory factors other than the drug (e.g., other medication changes or intercurrent illness) and any hypothesized relationship of the drug to the event (e.g., dose response or noted in other subjects); and (3) any observed changes in behavior and appearance (e.g., shortness of breath or rash). The SAFTEE-GI takes 5 to 10 minutes to administer if there are few symptoms but could potentially take much longer. The SAFTEE-GI has fair to good reliability in use (somewhat better for the presence of symptoms than for detailed ratings such as severity). The SAFTEE-GI has good validity based on a comparison with similar measures and on its ability to detect expected side effects. The SAFTEE-GI is more sensitive to side effects than is typical clinical practice, and the SAFTEE-SI is more sensitive still. Although the SAFTEE-SI is only suitable for use in research settings, the SAFTEE-GI could be used in clinical practice when a new drug is administered as a complement to specific questions about expected side effects of a particular drug.

Abnormal Involuntary Movement Scale. The AIMS (Table 7.8–3) is a clinical examination and rating scale developed in the 1970s to measure dyskinetic symptoms in patients taking antipsychotic drugs. The AIMS has 12 items, each of which is rated on an item-specific five-point severity scale ranging from 0 to 4. Total scores are not generally reported. Instead, changes in global severity and individual areas can be monitored over time. Ten items cover the movements themselves, divided in sections rating global severity and those related to specific body regions; two items concern dental factors that can complicate the diagnosis of dyskinesia. In the presence of extended neuroleptic exposure and the absence of other conditions causing dyskinesia, mild dyskinetic movements in two areas or moderate movements in one area suggest a diagnosis of tardive dyskinesia. The AIMS was developed for clinician raters, but lay raters can be trained to use it. It can be completed in less than 10 minutes. Excellent reliability has been demonstrated, especially for experienced raters, and the instrument appears valid. The AIMS is considered standard clinical practice for patients receiving long-term neuroleptic drugs in many clinical settings and is essential in clinical practice and research, both for monitoring patients for the development of tardive dyskinesia and for tracking changes in tardive dyskinesia over time.

Table 7.8–3.
Abnormal Involuntary Movement Scale (AIMS) Examination Procedure

Patient identification: _____ Date: _____

_____

Rated by: _____

Either before or after completing the examination procedure, observe the patient unobtrusively at rest (e.g., in waiting room).

The chair to be used in this examination should be a hard, firm one without arms.

After observing the patient, he or she may be rated on a scale of 0 (none), 1 (minimal), 2 (mild), 3 (moderate), and 4 (severe) according to the severity of symptoms.

Ask the patient whether there is anything in his/her mouth (i.e., gum, candy, etc.) and if there is to remove it.

Ask patient about the current condition of his/her teeth. Ask patient if he/she wears dentures. Do teeth or dentures bother patient now?

Ask patient whether he/she notices any movement in mouth, face, hands, or feet. If yes, ask to describe and to what extent they currently bother patient or interfere with his/her activities.

0 1 2 3 4	Have patient sit in chair with hands on knees, legs slightly apart, and feet flat on floor. (Look at entire body for movements while in this position.)
0 1 2 3 4	Ask patient to sit with hands hanging unsupported. If male, between legs; if female and wearing a dress, hanging over knees. (Observe hands and other body areas.)
0 1 2 3 4	Ask patient to open mouth. (Observe tongue at rest within mouth.) Do this twice.
0 1 2 3 4	Ask patient to protrude tongue. (Observe abnormalities of tongue movement.) Do this twice.
0 1 2 3 4	Ask the patient to tap thumb, with each finger, as rapidly as possible for 10–15 seconds; separately with right hand, then with left hand. (Observe facial and leg movements.)
0 1 2 3 4	Flex and extend patient's left and right arms. (One at a time.)
0 1 2 3 4	Ask patient to stand up. (Observe in profile. Observe all body areas again, hips included.)
0 1 2 3 4	Ask patient to extend both arms outstretched in front with palms down. (Observe trunk, legs, and mouth.)[a]
0 1 2 3 4	Have patient walk a few paces, turn and walk back to chair. (Observe hands and gait.) Do this twice.[a]

[a]Activated movements.

Simpson-Angus Rating Scale for Extrapyramidal Side Effects. The Simpson-Angus Rating Scale for Extrapyramidal Side Effects was developed to monitor the effects of antipsychotic drugs. It has ten items, each of which is rated on an item-specific five-point severity scale ranging from 0 to 4. Scores are reported as the mean on all ten items, with 0.3 considered the upper limits of normal. It is strongly focused on parkinsonian symptoms, particularly rigidity, but includes one akathisia item. It is designed for clinician use but can be administered by trained lay raters and takes approximately 10 minutes to administer. Good reliability has been reported, and validity is supported by the correlation of scores with antipsychotic drug dose. The scale is useful in a wide variety of clinical and research settings to monitor parkinsonian side effects and the impact of interventions to treat these effects.

Psychiatric Diagnosis and Screening

Instruments assessing psychiatric diagnosis are central to psychiatric research and may be useful in clinical practice as well. However, they tend to be rather long, especially when administered to individuals reporting many symptoms, potentially requiring many follow-up questions. When such instruments are evaluated, it is important to be sure that they implement current diagnostic criteria and cover the diagnostic areas of interest. For instance, few cover personality disorders (with the exception in some cases of antisocial personality), and not all cover disorders that typically onset in childhood.

Structured Clinical Interview for DSM. The SCID was initially developed in the early 1990s to provide a standardized DSM-III-R Axis I diagnosis based on an efficient but thorough clinical evaluation and is sometimes referred to as the SCID-I to distinguish it from the SCID-II, a similar measure for personality disorders. It was subsequently updated for the DSM-IV, and the APA plans to release the latest clinician version for DSM-5 in Spring 2015. The semistructured diagnostic interview begins with a section on demographic information and clinical background. Then there are seven diagnostic modules focused on different diagnostic groups: mood, psychotic, substance abuse, anxiety, somatoform, eating, and adjustment disorders; the modules can be given separately. Both required and optional probes are provided, and skip outs are suggested where no further questioning is warranted. All available information, including that from hospital records, informants, and patient observation, should be used to rate the SCID. The SCID is designed to be administered by experienced clinicians and is generally not recommended for use by lay interviewers. In addition, formal training in the SCID is required, and training books and videos are available to facilitate this. For individuals without symptoms, the interview takes approximately 1 hour to complete but may take up to 3 hours for individuals with extensive symptomatology. Although the primary focus is research with psychiatric patients, a nonpatient version (with no reference to a chief complaint) and a more clinical version (without as much detailed subtyping) are also available. Reliability data on the SCID suggest that it performs better on more severe disorders (e.g., bipolar disorder) or alcohol withdrawal than on milder ones (e.g., persistent depressive disorder [dysthymia]). Validity data are limited, as the SCID is more often used as the gold standard to evaluate other instruments. It is considered the standard interview to verify diagnosis in clinical trials and is extensively used in other forms of psychiatric research. Although its length precludes its use in routine clinical practice, the SCID can sometimes be useful to assure a systematic evaluation in psychiatric patients—for instance, on admission to an inpatient unit or at intake into

an outpatient clinic. It is also used in forensic practice to assure a formal and reproducible examination.

Diagnostic Interview Schedule and Composite International Diagnostic Instrument. The Diagnostic Interview Schedule (DIS) and Composite International Diagnostic Instrument (CIDI) are fully structured diagnostic interviews designed for lay administration. The DIS was developed in the 1980s for use in the Epidemiologic Catchment Area (ECA) study in the United States, which aimed to assess rates of current and lifetime psychiatric illness according to the DSM-III in a large and diverse community sample and has since been updated for DSM-III-R and DSM-IV criteria; efforts are now underway to update and validate these tools based on DSM-5 criteria. In 19 diagnostic modules, it covers a broad range of psychiatric conditions in adults, plus several childhood-onset disorders and antisocial personality, with fairly extensive information about symptoms, impairment, and treatment. The CIDI was developed from the DIS for international use and covers both ICD and DSM-IV criteria in 11 diagnostic modules; the CIDI does not cover antisocial personality or childhood-onset disorders. Both instruments begin with a screening question for each module, followed by extensive follow-up questions for each yes answer. All questions are read verbatim with little or no rewording allowed; only specified probes may be used for follow-up. The current version of the CIDI is the CIDI-3, which puts all of the screening questions at the beginning so that respondents do not answer negatively to avoid the follow-up questions. The DIS takes 90 minutes to 2 hours to complete; the CIDI may be somewhat shorter. Both can be scored by computer, yielding diagnoses and symptom profiles. A computerized, self-administered version of the DIS is also available, as are a short form (the DIS-SF) and an even shorter screening version (DIS-S). The instruments are designed for use by lay interviewers with extensive training in their use (formal training is recommended). Reliability appears to be good for both, at least for more severe disorders. Validity appears problematic for the DIS: studies of agreement with clinician diagnoses have been inconsistent, with marked discrepancies often observed for psychotic disorders. Both instruments have been used extensively in psychiatric research, particularly in epidemiological settings, and provide valuable data. However, some caution is warranted in interpreting these data given concerns about the instruments' validity, particularly for psychotic disorders.

Patient Health Questionnaire. The Patient Health Questionnaire (PHQ) is a self-administered screening instrument designed primarily for primary care. It was developed in 1999 based on the earlier Primary Care Evaluation of Mental Disorders (PRIME-MD). It focuses on more common disorders likely to be seen in primary care settings: major depressive disorder, panic

disorder, somatoform disorder, binge-eating disorder, alcohol abuse, and subthreshold depression and anxiety. It has 54 questions and takes 5 to 10 minutes to complete and thus can easily be completed in the waiting room. The primary care provider then reviews the positive responses and derives diagnoses in 1 to 5 minutes from a review of the patient's answers and an accompanying checklist. Training in the use of the instrument or careful review of the instruction manual is recommended. Reliability appears to be fair to good, better for more severe diagnoses. Validity judged against psychiatrist evaluations is quite good. The PHQ appears to be useful for primary care settings but lacks appropriate coverage for the range of disorders seen in psychiatric settings.

The American Psychiatric Association's DSM-5 Cross-Cutting Symptom Measures. With the release of the DSM-5, the APA has proposed several screening measures that can be used at the point of initial evaluation in order to guide the process of diagnosis and treatment, or to track a range of symptoms that cut across diagnostic boundaries. These instruments are referred to as the DSM-5 Self-Rated Level 1 Cross-Cutting Symptom Measures, and self-administered forms are available for adults and children (aged 11 to 17). There is also a parent/guardian rated form for children (aged 6 to 17). The adult form contains 23 items tapping 13 domains (Depression, Anger, Mania, Anxiety, Somatic Symptoms, Suicidal Ideation, Psychosis, Sleep Problems, Memory, Repetitive Thoughts and Behaviors, Dissociation, Personality Functioning, and Substance Use); whereas the Child form contains 25 items tapping 12 domains (Somatic Symptoms, Sleep Problems, Inattention, Depression, Anger, Irritability, Mania, Anxiety, Psychosis, Repetitive Thoughts and Behaviors, Substance Use, and Suicidal Ideation). All items are rated along the same five-point Likert scale (None, Slight, Mild, Moderate, Severe), and the highest scores are recorded for each domain. These measures were developed using a combination of items drawn from existing validated measures (e.g., the two depression items were taken from the PHQ 2-item screening measure for depression) and novel items developed by the DSM-5 working group (e.g., the two personality functioning items and one dissociation item). While the DSM-5 field trials of these measures indicate good test-retest reliability and convergent validity, they have not undergone full psychometric evaluation. Based on experience to date, if an individual receives a score of "2" (Mild) or greater on any scale (or 1 or greater for suicidal ideation, substance use, and psychosis), the APA recommends further evaluation through the administration of more in-depth Level 2 scales (available in DSM-5).

Broadband Measures of Psychological Functioning

The field of psychometrics has evolved considerably over the last 50 years, and has given rise to a number of broadband measures of psychological

functioning that have been shown to be psychometrically robust and strong predictors of important clinical outcomes. Among the most well-known include the MMPI, the Millon Clinical Multiaxial Inventory (MCMI), and the PAI. These types of measures allow clinicians to evaluate a wide array of important clinical domains (which may be particularly useful at the point of initial evaluation or in cases where diagnosis may be complex and/or ambiguous), while also assessing for other important factors such as test validity (i.e., inconsistent and/or idiosyncratic responding), response style (e.g., whether someone may be presenting in an overly positive or negative manner), interpersonal style, and suicide and violence risk, among others. The strengths of these measures include their comprehensive and wide-reaching scope, and more systematic evaluation of an array of clinically relevant domains. At the same time, such instruments may also be time-consuming and overly taxing for the patient, and thus careful consideration should be given a priori as to whether and under what types of conditions this type of assessment would prove most useful. For example, these types of assessments may be more inappropriate in acute settings, where patients may lack the cognitive resources to complete such a long instrument, and effort may be better spent in administering more targeted (and briefer) assessments instead.

Over the last decade, there has also emerged an accumulating body of research revealing the presence of multiple higher-order domains of psychopathology, which explain the covariance (i.e., comorbidity) often observed across many types of psychiatric disorders. These higher-order dimensions have commonly been referred to as: Internalizing (unipolar depression, generalized anxiety, social phobia, panic disorder, posttraumatic stress); Externalizing (alcohol abuse, drug abuse, impulsivity, aggression, and adult antisocial behavior); and Reality-Impairing pathology (e.g., psychotic experiences, paranoid ideation, grandiose ideation, and manic activation). Using this *quantitative model* of psychopathology as a framework, there have also recently emerged a number of broadband measures which evaluate these constructs at multiple levels—including the Psychopathology Spectra Inventory (PSI).

The Minnesota Multiphasic Personality Inventory. The standard MMPI-2 is a 567-item, self-administered instrument that assesses a broad range of psychopathology, while also providing indicators of test validity, response style, and interpersonal functioning. It takes roughly 45 to 60 minutes to complete, and uses a standard computer scoring program that produces a profile of scores which clinicians are then able to interpret relative to established population norms. MMPI items are rated along the same True/False scale, which map onto 10 clinical scales (Hypochondriasis, Depression, Hysteria, Psychopathic Deviate, Masculinity/Femininity, Paranoia, Psychasthenia, Schizophrenia, Hypomania, Social Introversion).

Likewise, the MMPI also includes a number of validity scales evaluating for response consistency, idiosyncratic responding, and response style (i.e., presenting as overly favorable or pathological). A wide range of other scales have also been developed *post-hoc,* and are included in the current scoring programs available by the test publisher. The original MMPI was initially developed in the early 1940s and restandardized in 1989 (MMPI-2), and there now exists several different versions of the tool, including one for adolescents (MMPI-A) and a reformatted version (MMPI-2-RF) that was published in 2008 based on more modern psychometric methods. The original MMPI contained 566 items, whereas the 1989 restandardized version (MMPI-2) made some additional modifications resulting in a total item pool of 567. In contrast, the MMPI-2-RF contains only 338 items. It is among the most widely researched broadband measures of psychological functioning, and its validity and reliability have been well-established across a wide array of clinical and nonclinical settings.

The Personality Assessment Inventory. The PAI was published in 1991, and is a 344-item, multi-scale, self-report assessment tool that assesses a wide range of psychopathology, not just personality. The instrument underwent multiple rounds of rigorous psychometric testing in order to maximize scaling properties (e.g., internal consistency reliability, item-to-scale convergence and divergence, etc.) and has been shown to possess strong construct validity. All 344 items are rated by the patient along the same four point Likert scale (False, Slightly True, Mainly True, Very True), and are then summed (within scales) and converted to T scores relative to a general US population sample. Another advantage of the PAI is that scores are also plotted relative to a large clinical sample, which allows clinicians to make additional interpretations about profile elevations relative to other psychiatric samples. The PAI includes four validity indices assessing inconsistent (INC) and idiosyncratic (INF) responding, as well as negative (NIM) and positive response styles (PIM). It also measures eleven clinical constructs (scales) including: Somatic Complaints (SOM), Anxiety (ANX), Anxiety-Related Disorders (ARD), Depression (DEP), Mania (MAN), Paranoia (PAR), Schizophrenia (SCZ), Borderline Features (BOR), Antisocial Features (ANT), Alcohol Problems (ALC) and Drug Problems (DRG). Finally, it contains five scales evaluating suicide risk and factors that may impact the course of treatment (e.g., treatment amenability, availability of social supports, etc.), as well interpersonal style (i.e., dominance and warmth). The PAI has been well validated across an array of clinical and nonclinical settings, and is frequently used in clinical and nonclinical (e.g., fertility clinics, employee evaluations) settings as a more comprehensive screening tool.

The Psychopathology Spectra Inventory. The PSI is a brief, self-

administered multi-scale measure of psychopathology and functioning developed in 2015. Guided by recent advances in quantitative models of psychopathology, it is comprised of 96 items which map onto 12 lower-order clinical scales: Depression, Anxiety, Social Anxiety, PTSD, Alcohol Problems, Drug Problems, Antisocial behavior, Severe Aggression, Psychosis, Paranoid Ideation, Manic Activation, and Grandiose Ideation. These lower-order scales then roll up into higher-order spectra (Internalizing, Externalizing, and Reality Impairing disorders) that explain the observed relationships among DSM disorders (using factor analytic methods). The PSI takes roughly 15 minutes to complete and has been found to demonstrate strong psychometric properties, and early research on the validity and reliability of the PSI has been promising. It employs a computer scoring program that generates patient score profiles relative to a large and representative US population sample. The PSI is unique among existing assessment instruments in that it measures the degree of pathology present across the three primary Spectra, while also identifying prominent clinical conditions. This allows clinicians to determine a patient's primary area of vulnerability for the expression of psychopathology (i.e., Internalizing, Externalizing, or Reality Impairing), as well as the specific clinical conditions currently being expressed. By combining brevity, advanced empirical conceptualization of psychopathology, and clinical utility, the PSI is well suited for use as a primary or supplemental measure across a broad range of clinical and research settings.

Schizophrenia Spectrum and Other Psychotic Disorders

A variety of instruments are used for patients with psychotic disorders. Those reported here are symptom severity measures. A developing consensus suggests that the distinction between positive and negative symptoms in schizophrenia is worthwhile, and more recently developed instruments implement this distinction. Because patients with psychotic disorders often lack insight and are sometimes agitated, observation of the patient is required in addition to direct questioning. Thus, most instruments in this domain must be administered by psychiatrists or others with clinical training.

Brief Psychiatric Rating Scale. The BPRS (Table 7.8–4) was developed in the late 1960s as a short scale for measuring the severity of psychiatric symptomatology. It was developed primarily to assess change in psychotic inpatients and covers a broad range of areas, including thought disturbance, emotional withdrawal and retardation, anxiety and depression, and hostility and suspiciousness. Its 18 items are rated on a seven-point item-specific Likert scale from 0 to 6, with the total score ranging from 0 to 108 (in some scoring systems, the lowest level for each

item is 1, so the range is 18 to 126). Because the ratings include observations, as well as patient reports of symptoms, the BPRS can be used to rate patients with very severe impairment. It is intended for use by experienced clinicians and can be administered in 30 minutes or less, including patient interview and observation. Reliability of the BPRS is good to excellent when raters are experienced, but this is difficult to achieve without substantial training; a semistructured interview has been developed to increase reliability. Validity is also good as measured by correlations with other measures of symptom severity, especially those assessing schizophrenia symptomatology. The BPRS has been used extensively for decades as an outcome measure in treatment studies of schizophrenia; it functions well as a measure of change in this context and offers the advantage of comparability with earlier trials. However, it has been largely supplanted in more recent clinical trials by the newer measures described below. In addition, given its focus on psychosis and associated symptoms, it is only suitable for patients with fairly significant impairment. Its use in clinical practice is less well supported, in part because considerable training is required to achieve the necessary reliability.

Positive and Negative Syndrome Scale. The Positive and Negative Syndrome Scale (PANSS) was developed in the late 1980s to remedy perceived deficits in the BPRS in the assessment of positive and negative symptoms of schizophrenia and other psychotic disorders by adding additional items and providing careful anchors for each. The PANSS includes 30 items on three subscales: seven items covering positive symptoms (e.g., hallucinations and delusions), seven covering negative symptoms (e.g., blunted affect), and 16 covering general psychopathology (e.g., guilt or uncooperativeness). Each item is scored on a seven-point, item-specific Likert scale ranging from 1 to 7; thus, the positive and negative subscales each range from 7 to 49, and the general psychopathology scale ranges from 16 to 112. The PANSS requires a clinician rater because considerable probing and clinical judgment are required. A semistructured interview guide is available. The ratings can be completed in 30 to 40 minutes. Reliability for each scale has been shown to be fairly high, with excellent internal consistency and interrater reliability. Validity also appears good based on correlation with other symptom severity measures and factor analytic validation of the subscales. The PANSS has become the standard tool for assessing clinical outcome in treatment studies of schizophrenia and other psychotic disorders and has been shown to be easy to administer reliably and sensitive to change with treatment. Its high reliability and good coverage of both positive and negative symptoms make it excellent for this purpose. It may also be useful for tracking severity in clinical practice, and its' clear anchors make it easy

to use in this setting.

Scale for the Assessment of Positive Symptoms and Scale for the Assessment of Negative Symptoms. The Scale for the Assessment of Positive Symptoms (SAPS) and Scale for the Assessment of Negative Symptoms (SANS) were designed to provide a detailed assessment of positive and negative symptoms of schizophrenia and may be used separately or in tandem. The domains assessed include hallucinations, delusions, bizarre behavior, and thought disorder for the SAPS and affective flattening, poverty of speech, apathy, anhedonia, and inattentiveness for the SANS. Each instrument consists of 30 fully anchored items each scored from 0 to 5; thus, the total score ranges from 0 to 150 for each. Each must be rated by an experienced clinician and requires approximately 30 minutes to complete. Good to excellent interrater reliability has been demonstrated if trained interviewers are used, and each scale has high internal consistency as well. Validity is supported by correlation with other symptom severity instruments. The SANS and SAPS are principally used to monitor treatment effects in clinical research and have also been used to help characterize positive and negative symptoms in studies of schizophrenia phenomenology. The SAPS has been to some extent supplanted by the PANSS, but the SANS still has a place in studies aimed at negative symptoms. The instruments are longer and more difficult to learn to use than the PANSS and thus generally too detailed for use in clinical practice. However, the comprehensive characterization of symptomatology provided might be useful as part of a formal evaluation.

Table 7.8–4.
Brief Psychiatric Rating Scale

Patient Name _____ Today's Date _____

Please enter the score for the term that best describes the patient's condition.

0 = Not assessed, 1 = Not present, 2 = Very mild, 3 = Mild, 4 = Moderate, 5 = Moderately severe, 6 = Severe, 7 = Extremely severe

Score	
☐	**1. SOMATIC CONCERN** Preoccupation with physical health, fear of physical illness, hypochondriasis.
☐	**2. ANXIETY** Worry, fear, over-concern for present or future, uneasiness.
☐	**3. EMOTIONAL WITHDRAWAL** Lack of spontaneous interaction, isolation deficiency in relating to others.
☐	**4. CONCEPTUAL DISORGANIZATION** Thought processes confused, disconnected, disorganized, disrupted.
☐	**5. GUILT FEELINGS** Self-blame, shame, remorse for past behavior.
☐	**6. TENSION**

Physical and motor manifestations of nervousness, over-activation.

☐ **7. MANNERISMS AND POSTURING**
Peculiar, bizarre, unnatural motor behavior (not including tic).

☐ **8. GRANDIOSITY**
Exaggerated self-opinion, arrogance, conviction of unusual power or abilities.

☐ **9. DEPRESSIVE MOOD**
Sorrow, sadness, despondency, pessimism.

☐ **10. HOSTILITY**
Animosity, contempt, belligerence, disdain for others.

☐ **11. SUSPICIOUSNESS**
Mistrust, belief others harbor malicious or discriminatory intent.

☐ **12. HALLUCINATORY BEHAVIOR**
Perceptions without normal external stimulus correspondence.

☐ **13. MOTOR RETARDATION**
Slowed, weakened movements or speech, reduced body tone.

☐ **14. UNCOOPERATIVENESS**
Resistance, guardedness, rejection of authority.

☐ **15. UNUSUAL THOUGHT CONTENT**
Unusual, odd, strange, bizarre thought content.

☐ **16. BLUNTED AFFECT**
Reduced emotional tone, reduction in formal intensity of feelings, flatness.

☐ **17. EXCITEMENT**
Heightened emotional tone, agitation, increased reactivity.

☐ **18. DISORIENTATION**
Confusion or lack of proper association for person, place or time.

Instructions for the Clinician:

The Brief Psychiatric Rating Scale (BPRS) is a widely used instrument for assessing the positive, negative, and affective symptoms of individuals who have psychotic disorders, especially schizophrenia. It has proven particularly valuable for documenting the efficacy of treatment in patients who have moderate to severe disease.

It should be administered by a clinician who is knowledgeable concerning psychotic disorders and able to interpret the constructs used in the assessment. Also considered is the individual's behavior over the previous 2–3 days and this can be reported by the patient's family.

The BPRS consists of 18 symptom constructs and takes 20–30 minutes for the interview and scoring. The rater should enter a number ranging from 1 (not present) to 7 (extremely severe). 0 is entered if the item is not assessed.

First published in 1962 as a 16-construct tool by Drs. John Overall and Donald Gorham, the developers added two additional items, resulting in the 18-item scale used widely today to assess the effectiveness of treatment.

BPRS Scoring Instructions:

Sum the scores from the 18 items. Record the total score and compare the total score from one evaluation to the next as the measure of response to treatment.

[a]Ratings based primarily on verbal report.

©2008–2012 PROMIS Health Organization (PHO) and PROMIS Cooperative Group.

Overall, JE, Gorham DR: The Brief Psychiatric Rating Scale (BPRS): recent developments in ascertainment and scaling. *Psychopharmacology Bulletin* 1988;24:97–99.

Depressive Disorders

Rating depression depends, to a substantial extent, on subjective

assessment of mood states, so interviews and self-report instruments are both common. Because depression is common in the general population and involves significant morbidity and even mortality, screening instruments—especially those using a self-report format—are potentially quite useful in primary care and community settings. A number of interviewer and self-administered instruments are described below.

Hamilton Rating Scale for Depression. The HAM-D was developed in the early 1960s to monitor the severity of major depression, with a focus on somatic symptomatology. The version in most common use has 17 items, although versions with different numbers of items (including the 24-item version) have been used in many studies as well. The 17-item version does not include some of the symptoms of depression in the DSM-III and its successors, most notably the so-called reverse neurovegetative signs (increased sleep, increased appetite, and psychomotor retardation). Items on the HAM-D are scored from 0 to 2 or from 0 to 4, with total score ranging from 0 to 50. Scores of 7 or less may be considered normal; 8 to 13, mild; 14 to 18, moderate; 19 to 22, severe; and 23 and above, very severe. The HAM-D was designed for clinician raters but has been used by trained lay administrators as well. Ratings are completed by the examiner based on patient interview and observations. A structured interview guide has been developed to improve reliability. The ratings can be completed in 15 to 20 minutes. Reliability is good to excellent, particularly when the structured interview version is used. Validity appears good based on correlation with other depression symptom measures. The HAM-D has been used extensively to evaluate change in response to pharmacological and other interventions and thus offers the advantage of comparability across a broad range of treatment trials. It is more problematic in the elderly and the medically ill, in whom the presence of somatic symptoms may not be indicative of major depression.

Montgomery-Asberg Depression Rating Scale. The Montgomery-Asberg Depression Rating Scale (MADRS) was developed in 1979 explicitly for clinical trials and thus was aimed at maximizing sensitivity to change in established depression. It can be administered by a psychiatrist, other physician, or psychiatric nurse, and a structured interview guide and self-report version are also available. It includes 10 items on seven-point, partially anchored, Likert scales. It can be administered in 10 to 15 minutes and is easy to use with minimal training. Very good reliability and validity have been demonstrated. Because of its sensitivity to change, the MADRS is frequently used in clinical trials of antidepressants and may also have a role in assessing the impact of treatment in clinical practice.

Beck Depression Inventory. The Beck Depression Inventory (BDI) was

developed in the early 1960s to rate depression severity, with a focus on behavioral and cognitive dimensions of depression. The current version, (the BDI-II) has added more coverage of somatic symptoms to be compatible with more current versions of the DSM and covers the most recent 2 weeks. Earlier versions are focused on the past week or even shorter intervals, which may be preferable for monitoring treatment response. The BDI includes 21 self-report items, each of which has four statements describing increasing levels of severity, and the total score ranges from 0 to 84. Scores of 0 to 9 are considered minimal; 10 to 16, mild; 17 to 29, moderate; and 30 to 63, severe. The scale can be completed in 5 to 10 minutes. Internal consistency has been high in numerous studies. Test–retest reliability is not consistently high, but this may reflect changes in underlying symptoms. Validity is supported by correlation with other depression measures. The principal use of the BDI is as an outcome measure in clinical trials of interventions for major depression, including psychotherapeutic interventions. Because it is a self-report instrument, it is sometimes used to screen for major depression. However, in trials of various cutoffs for a diagnosis of major depression, the best seemed to be 9, which had only fair sensitivity, at a cost of considerable nonspecificity, suggesting that the instrument has limited use for screening. The instrument's strength lies in measuring the depth of depression and in its comprehensive coverage of the cognitive dimension of depression.

The Patient Health Questionnaire—9-item Screener for Depression. The Patient Health Questionnaire-9 (PHQ-9) is a nine-item, self-administered screening measure that was developed in 1999 (as part of the larger PHQ) as a brief tool for assessing depressive symptomatology in primary care settings. The PHQ-9 items are rated along a four-point Likert scale (0. Not at all, 1. Several Days, 2. More than half of the days, and 3. Nearly every day) and then summed, with total scores ranging from 0 to 27. Although some have suggested cut-scores of 10 (or between 8 and 11) for a diagnosis of Major Depressive Disorder, the APA has also provided a system for interpreting different PHQ-9 score ranges in terms of severity: None = 0 to 4; Mild Depression = 5 to 9; Moderate Depression = 10-14; Moderately Severe Depression = 15 to 19; and Severe Depression = 20 to 27. The PHQ-9 in particular has generally performed well in terms of its underlying psychometric properties, and studies evaluating validity have also been promising. The APA has recommended the PHQ-9 as a brief screening tool for evaluating and tracking depressive symptomatology over time in DSM-5.

Inventory of Depressive Symptomatology and Quick Inventory of Depressive Symptomatology. The Inventory of Depressive Symptomatology (IDS) was developed in 1982 to reflect a broader conceptualization of major

depression than the HAM-D and other earlier instruments and to provide a comprehensive assessment of depressed inpatients and outpatients. It has better coverage of the full range of DSM-5 symptoms of depression and covers mild as well as severe symptoms of depression. A shorter version, the 16-item Quick Inventory of Depressive Symptomatology (QUIDS), was released in 2000. Both clinician-rated (IDS-C and QUIDS-C) and self-report (IDS-SR and QUIDS-SR) versions of each instrument are available. The clinician versions come with structured interview guides to improve the reliability of the ratings. Each item has four to five scaled multiple-choice answers, ranging from no difficulty to significant difficulty in that area. The IDS-SR requires 15 to 20 minutes, and the IDS-C 20 to 30 minutes; the corresponding versions of the QUIDS can be completed in half the time. The IDS has been subjected to rigorous reliability and validity assessment. It shows good reliability in a variety of paradigms and excellent validity, as judged by comparison with other depression scales. It has also been shown to have good validity in the classification of depressive subtypes and to be more sensitive to the milder end of the spectrum than older depression instruments. The QUIDS also has very good reliability and validity but does not attempt to measure subtypes. Both instruments have considerable utility in clinical research and clinical practice for the evaluation of depression and assessment of its response to treatment.

Bipolar and Related Disorders

In evaluating mania, as for psychotic disorders, limited insight and agitation may hinder accurate symptom reporting, so clinician ratings including observational data are generally required. Likewise, whether someone is actively experiencing symptoms of mania/hypomania at the point of evaluation may also affect whether and how current and past manic symptoms are endorsed on psychometric tools, and thus a careful retrospective analysis via clinical interview is also recommended to more systematically rule this out. A patient who accurately reports *no* symptoms of mania at the time of evaluation may in fact have a history of mania in the past and meet criteria for Bipolar Disorder—reflecting the episodic nature of these symptoms. A number of instruments that assess for mania/hypomania are summarized below.

Young Mania Rating Scale. The Young Mania Rating Scale (YMRS) is a checklist developed in the late 1970s to provide a brief but thorough evaluation of the severity of mania that could be used to monitor treatment response or detect relapse. It consists of a checklist of 11 items rated either from 0 to 4 (seven items) or from 0 to 8 (four items). Each item has five item-specific anchors. The total score ranges from 0 to 60. Ratings include clinical observation, so it must be rated by a clinician, but reliable ratings have been obtained by nurses on inpatient units and beginning psychiatric

residents. Reliability is good based on interrater reliability and internal consistency studies. Validity also appears to be good based on limited data on correlation with other mania measures. The YMRS is useful for evaluating response to treatment in clinical research, and it has been found to be sensitive to change in this setting. It might also be used to assess treatment response or to monitor for relapse in treated or untreated patients, although extensive experience with this use has not been reported.

The Altman Self-Rating Mania Scale (ASRM). The Altman Self-Rating Mania Scale (ASRM) was initially developed in the late 1990s by Altman and colleagues, and consists of five items which assess the presence and/or severity of manic symptoms. Items are rated along unique five-point Likert scales, where patients are asked to rate their symptoms with respect to the last 7 days. In cases where patients may be too impaired to complete the tool, an informant version is also available. Total scores range from 5 to 25, where higher scores are suggestive of greater pathology. Scores of 6 or higher have been shown to be associated with a higher probability of a manic or hypomanic condition, and may signal the need for further evaluation and/or treatment. The ASRM has been shown to possess strong psychometric properties (including test–retest and internal consistency reliability), as well good convergent and discriminant validity (e.g., discriminating patient groups who are manic from other kinds of psychiatric disorders). The ASRM has also been found to be sensitive to change as a function of pharmacotherapy. The APA has recommended the ASRM as a brief measure for assessing mania/hypomania in the Emerging Measures section of the DSM-5.

Anxiety Disorders

The anxiety disorders addressed by the measures below include panic disorder and generalized anxiety disorder. When anxiety measures are examined, it is important to be aware that there have been changes over time in how anxiety disorders are defined. For example, Post Traumatic Stress Disorder and Obsessive Compulsive Disorders are no longer included in the DSM-5's Anxiety Disorders section, and generalized anxiety disorder has shifted somewhat to focus more on the affective (worry, fear) dimension of the condition. Thus, older measures have somewhat less relevance for diagnostic purposes, although they may identify symptoms causing considerable distress. Whether reported during an interview or on a self-report rating scale, virtually all measures in this domain, like the measures of depression discussed above, depend on subjective descriptions of inner states.

Hamilton Anxiety Rating Scale. The HAM-A was developed in the late 1950s to assess anxiety symptoms, both somatic and cognitive. Because the

conceptualization of anxiety has changed considerably, the HAM-A provides limited coverage of the "worry" required for a DSM-5 diagnosis of generalized anxiety disorder and does not include the episodic anxiety found in panic disorder. There are 14 items, each of which is rated from 0 to 4 on an unanchored severity scale, with the total score ranging from 0 to 56. A score of 14 has been suggested as the threshold for clinically significant anxiety, but scores of 5 or less are typical in individuals in the community. The scale is designed to be administered by a clinician, and formal training or the use of a structured interview guide is required to achieve high reliability. A computer-administered version is also available. Reliability is fairly good based on internal consistency, interrater, and test–retest studies. However, given the lack of specific anchors, reliability should not be assumed to be high across different users in the absence of formal training. Validity appears good based on correlation with other anxiety scales but is limited by the relative lack of coverage of domains critical to the modern understanding of anxiety disorders. Even so, the HAM-A has been used extensively to monitor treatment response in clinical trials of generalized anxiety disorder and may also be useful for this purpose in clinical settings.

The Generalized Anxiety Disorder-7. The Generalized Anxiety Disorder-7 (GAD-7) is a seven-item, self-administered measure that was published in 2006 as a general screening measure for symptoms of generalized anxiety. It consists of seven items that inquire about general worry and fear symptoms in the last 2 weeks, all of which are rated along a four-point Likert scale (0. Not at all, 1. Several Days, 2. More than half the days, and 3. Nearly every day). Scores range from 0 to 21 (where higher scores indicate greater levels of pathology), and cut-scores of 5, 10, and 15 have been suggested for mild, moderate, and severe GAD, respectively. The GAD-7 has been found to possess strong psychometric properties, and its use in clinical practice has increased in recent years. That being said, although it remains a useful tool for evaluating more generalized symptoms of anxiety, it may not provide the same type of correspondence to a DSM-5 diagnosis given differences in timeframe (the GAD-7 asks about symptoms in the last 2 weeks, whereas DSM-5 requires that symptoms be present more days than not for 6 months) and DSM-5's greater emphasis on the affective dimension of worry/fear.

The Patient-Reported Outcomes Measurement Information System Emotional Distress—Anxiety Short-Form. The Patient-Reported Outcomes Measurement Information System (PROMIS) Emotional Distress—Anxiety Short-Form is a seven-item measure of generalized symptoms of anxiety and worry that was developed by the larger PROMIS. PROMIS was launched in 2004 by the National Institute of Health (NIH) in an effort to

develop psychometrically sound measures of patient-reported measures of physical health, mental well-being, and social functioning. Items from the PROMIS Emotional Distress—Anxiety Short-Form are rated along the same five-point Likert scale and then summed, with total scores ranging from 7 to 35. These raw (sum) scores are then converted into T scores using a conversion table provided by PROMIS, which allows for immediate comparison relative to the general population. There are measures available for both adults (18+) and children (6 to 17). The PROMIS Emotional Distress—Anxiety Short-Form has demonstrated strong preliminary reliability and validity and was recommended for use in assessing anxiety in DSM-5, although given its more recent development, additional research is needed to substantiate this research base before widespread adoption.

Panic Disorder Severity Scale. The Panic Disorder Severity Scale (PDSS) was developed in the 1990s as a brief rating scale for the severity of panic disorder. It was based on the Yale-Brown Obsessive-Compulsive Scale and has seven items, each of which is rated on an item-specific, five-point Likert scale. The seven items address frequency of attacks, distress associated with attacks, anticipatory anxiety, phobic avoidance, and impairment. The items are scored from 0 to 4, and the total score ranges from 0 to 28. The instrument was designed for use by clinicians, but a self-report version is also available. Reliability is excellent based on interrater studies, but, in keeping with the small number of items and multiple dimensions, internal consistency is limited. Validity is supported by correlations with other anxiety measures, both at the total and item level; lack of correlation with the HAM-D; and, more recently, by brain imaging studies. Growing experience with the PDSS suggests that it is sensitive to change with treatment and is useful as a change measure in clinical trials or other outcome studies for panic disorder, as well as for monitoring panic disorder in clinical practice.

Trauma and Stressor-Related Disorders

With the reorganization of DSM-5, Trauma and Stressor-Related Disorders were moved out of the Anxiety Disorders section and given their own distinct section—which now focuses on an array of conceptually similar disorders, in which exposure to a stressful or traumatic event is a necessary criterion. These include Reactive Attachment Disorder, Disinhibited Social Engagement Disorder, Post Traumatic Stress Disorder, Acute Stress Disorder, and Adjustment Disorders. PTSD remains the most commonly studied of these disorders, and a PTSD instrument is described here.

Clinician-Administered PTSD Scale. The Clinician-Administered PTSD Scale (CAPS) was developed in 1990 to operationalize the DSM-III-R

diagnosis of PTSD and can easily be modified to obtain a diagnosis by DSM-5 criteria. The interview includes 17 items required to make the diagnosis, covering all four criteria: (1) the event itself, (2) reexperiencing of the event, (3) avoidance, and (4) increased arousal. For each of the 17 items, the interviewer rates both the frequency and the intensity on a five-point scale ranging from 0 to 4. The diagnosis requires evidence of a traumatic event, one symptom of reexperiencing, three of avoidance, and two of arousal (typically, an item is counted if frequency is at least 1 and intensity is at least 2). The items can also be used to generate a total PTSD severity score obtained by summing the frequency and intensity scales for each item. The CAPS also includes several global rating scales for the impact of PTSD symptomatology on social and occupational functioning, for general severity, for recent changes, and for the validity of the patient's report. The CAPS must be administered by a trained clinician and requires 45 to 60 minutes to complete, with follow-up examinations somewhat briefer. It has demonstrated reliability and validity in multiple settings and multiple languages, and performs well in the research setting for diagnosis and severity assessment, but is generally too lengthy for use in clinical practice.

Obsessive Compulsive and Related Disorders

Like the Trauma and Stressor-Related Disorders, Obsessive Compulsive Disorder and Related Disorders were also removed from the Anxiety Disorders section and given their own section in DSM-5. This new family of disorders now includes: Obsessive Compulsive Disorder, Body Dysmorphic Disorder, Hoarding Disorder, Trichotillomania, and Excoriation, among others. However, OCD remains the most common and researched of these disorders, and several rating instruments for OCD are summarized below.

Yale-Brown Obsessive-Compulsive Scale. The Yale-Brown Obsessive-Compulsive Scale (YBOCS) was developed in the late 1980s to measure the severity of symptoms in OCD. It has 10 items rated based on a semistructured interview. The first five items concern obsessions: The amount of time that they consume, the degree to which they interfere with normal functioning, the distress that they cause, the patient's attempts to resist them, and the patient's ability to control them. The remaining five items ask parallel questions about compulsions. Each item has a set of item-specific anchors scored from 0 to 4, so total scores for obsessions and compulsions each range from 0 to 20, and overall total score ranges from 0 to 40. Typical scores for patients with OCD are in the 16 to 30 range, and a threshold of 16 is typically used for inclusion in drug trials. The semistructured interview and ratings can be completed in 15 minutes or less. A self-administered version can be completed in 10 to 15 minutes. Computerized and telephone use have also been found to provide

acceptable ratings. Before the first use of the YBOCS, an associated 64-item checklist is administered to provide a more detailed assessment of the specific content of the patient's obsessions and delusions. Reliability studies of the YBOCS show good internal consistency, interrater reliability, and test–retest reliability over a 1-week interval. The YBOCS has become the standard instrument for assessing OCD severity and is used in virtually every drug trial. It may also be used clinically to monitor treatment response.

The Repetitive Thoughts and Behaviors Scale. The Repetitive Thoughts and Behaviors Scale was adapted from the Florida Obsessive Compulsive Inventory (FOCI) and published in 1994, and consists of five items rated by patients along item-specific five-point (0 to 4) Likert scales that are summed for a total score ranging from 0 to 20 – with higher scores suggestive of more intense repetitive thoughts/behaviors. A cut-score of 8 has been suggested as a threshold for more intensive inquiry into potential obsessions/compulsions, and clinicians have also been encouraged to average the five items as a means of obtaining an overall estimate of severity (where 0 = mild; 1 = moderate; 2 = severe; and 3 = extreme). The Repetitive Thoughts and Behaviors scale has shown good initial reliability and validity, but, as with other DSM-5 emerging measures, future research will be helpful in further elucidating the scale's properties and clinical application.

Substance-Related and Addictive Disorders

DSM-5 now organizes substance-related disorders in a way that combines previous distinctions of "abuse" and "dependence" along a common continuum from mild to severe, and focuses on 10 separate classes of substances as separate use disorders: alcohol; caffeine; cannabis; hallucinogens; inhalants; opioids; sedatives, hypnotics, and anxiolytics; stimulants; tobacco; and other (unknown) substances. These disorders, particularly those involving alcohol, are common and debilitating in the general population, so screening instruments are particularly helpful. Because these behaviors are socially undesirable, underreporting of symptoms is a significant problem; thus, the validity of all substance use measures is limited by the honesty of the patient. Validation against drug tests or other measures is of great value, particularly when working with patients who have known substance abuse.

The Level 2 Substance Abuse Module from the DSM-5. The Level 2 Substance Abuse Module from the DSM-5 is a modified version of the National Institute on Drug Abuse's ASSIST evaluation system and was developed in as a brief screener to augment the clinical interview when evaluating for potential substance misuse. The ten-item measure is

completed either by the patient or informant prior to meeting with a clinician, takes roughly 5 to 10 minutes, and asks about how frequently the patient has used various classes of substances in the last two weeks on a five-point Likert scale (0, not at all; 1, 1 or 2 days; 2, several days; 3, more than half the days; 4, nearly every day). Clinicians are advised to interpret each item separately given that each assesses use of a different substance. Although no formal scoring program exists, higher scores are indicative of greater complexity and severity of substance use, and suggest the need for further evaluation and treatment. The APA has recommended this measure in the DSM-5's Emerging Measures section as one method for evaluating and tracking substance misuse over time, although validity and reliability data is somewhat limited given its more recent release.

CAGE. The CAGE was developed in the mid-1970s to serve as a very brief screen for significant alcohol problems in a variety of settings, which could then be followed up by clinical inquiry. CAGE is an acronym for the four questions that comprise the instrument: (1) Have you ever felt you should *c*ut down on your drinking? (2) Have people *a*nnoyed you by criticizing your drinking? (3) Have you ever felt bad or *g*uilty about your drinking? (4) Have you ever had a drink first thing in the morning to steady your nerves or to get rid of a hangover (*e*ye-opener)? Each "yes" answer is scored as 1, and these are summed to generate a total score. Scores of 1 or more warrant follow-up, and scores of 2 or more strongly suggest significant alcohol problems. The instrument can be administered in a minute or less, either orally or on paper. Reliability has not been formally assessed. Validity has been assessed against a clinical diagnosis of alcohol abuse or dependence, and these four questions perform surprisingly well. Using a threshold score of 1, the CAGE achieves excellent sensitivity and fair to good specificity. A threshold of 2 provides still greater specificity but at the cost of a fall in sensitivity. The CAGE performs well as an extremely brief screening instrument for use in primary care or in psychiatric practice focused on problems unrelated to alcohol. However, it has limited ability to pick up early indicators of problem drinking that might be the focus of preventive efforts.

Alcohol Use Disorders Identification Test. The Alcohol Use Disorders Identification Test (AUDIT) was developed by the WHO in the late 1980s as a brief screening instrument designed for the early detection of hazardous (i.e., involving the risk of harm) and harmful (i.e., involving the presence of harm) alcohol use in a variety of settings. It focuses on both the past year and current drinking. It includes a ten-item core screening instrument covering alcohol consumption, drinking behaviors, and alcohol-related problems. Each item is rated using item-specific anchors scored from 0 to 4 and summed for a total score from 0 to 40. The AUDIT can be

administered and scored in less than 5 minutes and does not require professional training. The AUDIT also offers a clinical screening procedure involving a physical examination and blood tests, which adds no more than 5 to 10 minutes to a routine medical examination. Reliability of the AUDIT appears good based on internal consistency data. Validity judged against a clinical diagnosis of alcoholism is also good: Using a threshold score of 8 is quite sensitive, at least in men, but somewhat nonspecific; less stringent cutoffs may perform better for women and adolescents. A score of 10 is more specific but at a cost in specificity. Validity also appears good based on correlation with other alcohol self-report measures and risk factors for alcoholism. The AUDIT provides an excellent brief screen for alcohol problems and is particularly good for detecting problem drinking at a fairly mild stage. However, its focus on early detection of hazardous and harmful drinking makes it less suitable to use as a diagnostic instrument.

Drug Abuse Screening Test. The Drug Abuse Screening Test (DAST) was developed in the early 1980s to serve as a screening and assessment instrument for drug abuse. The DAST is an adaptation of the Michigan Alcohol Screening Test (MAST) used to screen for alcoholism. It focuses on lifetime drug use, so it is not designed to measure changes over time. The current version of the DAST has 20 yes–no items, and can be given orally or as a paper-and-pencil questionnaire. An earlier version had 28 items, so the 20-item version is sometimes called the *Brief DAST*. The positive items can be summed to form a 20-point scale. The DAST can be administered and scored in less than 10 minutes. Reliability, based on internal consistency, is very good. Validity, based on ability to detect drug abuse disorder, also appears high, with excellent sensitivity and fairly good specificity using a threshold score of 5. The DAST is useful as a screening tool for drug abuse problems in patients with other mental disorders, particularly alcohol abuse. It also provides an overview of problem severity that may be useful to guide treatment choices. However, because it takes a lifetime approach, it is not useful as a change measure in clinical practice or clinical trials.

Addiction Severity Index. The Addiction Severity Index (ASI) was developed in the early 1980s to serve as a quantitative measure of symptoms and functional impairment due to alcohol or drug disorders. It covers demographics, alcohol use, drug use, psychiatric status, medical status, employment, legal status, and family and social issues. Frequency, duration, and severity are assessed. It has 142 items in varying formats, including yes–no, multiple-choice, and scaled items. It includes both subjective and objective items reported by the patient and observations made by the interviewer. In each area, the ASI yields the rater's global assessment of severity, along with a computer score on a 0 to 1 scale. The

142-item version includes information on the past 30 days and lifetime status, but a shorter version is available for use at follow-up. The instrument is designed for clinician administration but has been used successfully by trained lay raters. Training is recommended, and both manuals and formal training programs are available. A computerized version is also available. The standard ASI requires 45 to 75 minutes to complete, but the follow-up version can be completed in 15 to 20 minutes. Very good to excellent reliability has been demonstrated for the overall composite score, with somewhat lower reliability for severity ratings in each area. Validity has also been demonstrated based on correlation with other measures and discrimination of patient from nonpatient populations. Normative data are available for a range of populations of alcohol and drug abusers, including alcohol clinic patients, drug abusers, homeless persons, and prisoners. The principal use of the ASI is as an aid to treatment planning and the assessment of treatment outcome in clinical and law enforcement settings. It is relatively time-consuming to administer but performs well for this purpose. It has also been used in clinical research as a sensitive indicator of baseline severity and change over time, which allows comparison between clinical research and clinical practice.

Eating Disorders

Although feeding and eating disorders share a number of psychological and behavioral features, the DSM-5 has expanded the framework for conceptualizing and diagnosing these disorders based on differences in classification, clinical course, and treatment outcome; and as a means of reducing the frequency of the Eating Disorder NOS diagnosis. DSM-5 includes the following: Pica, Rumination Disorder, avoidant/restrictive food intake, disorder, anorexia nervosa, bulimia nervosa, and binge-eating disorder. A wide variety of instruments, particularly self-report scales, are available. Because of the secrecy that may surround dieting, bingeing, purging, and other symptoms, validation against other indicators (e.g., body weight for anorexia or dental examination for bulimia) may be very helpful. Such validation is particularly critical for patients with anorexia, who may lack insight into their difficulties.

Eating Disorders Examination. The Eating Disorders Examination (EDE) was developed in 1987 as the first interviewer-based comprehensive assessment of eating disorders, including diagnosis, severity, and an assessment of subthreshold symptoms, and a 2014 revision (version 17.0) was developed to be compatible with DSM-5. A self-report version (the EDE-Q), as well as an interview for children, has since been developed. The EDE focuses on symptoms during the preceding 4 weeks, although longer-term questions are included to assess diagnostic criteria for eating disorders. Each item on the EDE has a required probe with suggested

follow-up questions to judge severity or frequency or both, which are then rated on a seven-point Likert scale. For the self-report version, subjects are asked to make similar ratings of frequency or severity. The instrument provides both global severity ratings and ratings on four subscales: Restraint, eating concern, weight concern, and shape concern. The interview, which must be administered by a trained clinician, requires 30 to 60 minutes to complete, whereas the self-report version is somewhat faster. Reliability and validity data for both the EDE and EDE-Q are excellent, although the EDE-Q may have greater sensitivity for binge-eating disorder. The EDE performs well in both the diagnosis and the detailed assessment of eating disorders in the research context. It also has the sensitivity to change required for use in clinical trials or monitoring of individual therapy. Even in the research setting, however, the EDE is fairly lengthy for repeated use, and the EDE-Q may be preferable for some purposes. Although the EDE is too lengthy for routine clinical practice, the EDE or EDE-Q might be helpful in providing a comprehensive assessment of a patient with a suspected eating disorder, particularly during an evaluation visit or on entry into an inpatient facility.

The Eating Disorders Assessment-5 (EDA-5). The Eating Disorders Assessment-5 (EDA-5), developed in 2014, is a semi-structured interview that was designed to assist clinicians in evaluating for DSM-5 Feeding and Eating Disorders in children and adults. The tool is administered by clinicians with expertise in the evaluation of eating disorders; there is also a computerized/internet version. The EDA-5 provides a series of probes to inquire about symptoms corresponding to DSM-5 criteria for each of the Feeding and Eating disorders. Clinicians ask follow-up questions as needed, and rate each criterion using all other available clinical information (e.g., behavioral observations, information from family, medical records, objective measures of height/weight, etc.). The EDA-5 employs an algorithm whereby subsequent questions are administered based on responses to previous questions, and thus the number of items administered varies across patients. The instrument takes roughly 20 to 30 minutes to complete. Although the tool appears to be a promising measure for DSM-5 feeding and eating disorders, validity and reliability information remains limited given its recent release.

Neurocognitive Disorders

This subsection reviews measures of Major and Mild neurocognitive disorders, which primarily affect the elderly, where they are more commonly known as dementia and MCI; in younger patients, for instance those with acquired cognitive disorders resulting from head injury or HIV, the term neurocognitive disorder is more typically used. A wide variety of measures are available, and the focus here on those commonly used among

older patients. Most involve cognitive testing and provide objective, quantifiable data. However, scores vary by ability and by educational level in subjects in the general population, so these instruments tend to be most useful when the patient's own baseline scores are known. Other measures focus on functional status, which can be assessed based on a comparison with a description of the subject's baseline function; these types of measures generally require a knowledgeable informant and, thus, may be more cumbersome to administer but tend to be less subject to educational biases. A third type of measure focuses on the associated behavioral symptoms that are frequently seen in demented patients. Last, this subsection includes a measure of delirium, an acute or subacute confusional state resulting from a general medical condition.

Mini-Mental State Examination. The MMSE is a 30-point cognitive test developed in the mid-1970s to provide a bedside assessment of a broad array of cognitive function, including orientation, attention, memory, construction, and language. It can be administered in less than 10 minutes by a busy doctor or a technician and scored rapidly by hand. The MMSE has been extensively studied and shows excellent reliability when raters refer to consistent scoring rules. Validity appears good based on correlations with a wide variety of more comprehensive measures of mental functioning and clinicopathological correlations. Although the MMSE is used commonly to screen for dementia, in both clinical practice and epidemiological or clinical research, it is not optimal for this purpose, and even worse as a screen for MCI: it has limited memory items and no executive function items and thus tends to miss the areas most likely to be affected early in the course of disease. If the MMSE is used for screening, then a cutoff of 24 for identifying cases of dementia has been suggested, but it is probably more accurate to use age- and education-adjusted norms to interpret the results, and higher cutoffs are better for detecting MCI. For patients with extensive education, who may score 30 out of 30 despite clear evidence of functional decline, a more difficult cognitive test, full neuropsychological battery, or clinical interview about functional change may be required to detect dementia. The other principal use of the MMSE is in following the progression of dementia over time in the clinic or in clinical trials. As a rule of thumb, patients with mild dementia tend to score from 20 to 24, moderate from 11 to 19, and severe from 0 to 10. However, it should be noted that these figures do not take the educational differences noted above into account. In addition, the MMSE does not do as well tracking progression of dementia at later stages; scores of 10 or less are difficult to obtain, as many patients become untestable.

Table 7.8–5.
The Patient Health Questionnaire-9

Severity Measure for Depression—Adult*

*Adapted from the Patient Health Questionnaire–9 (PHQ-9)

Name: _____ Age: _____ Sex: Male □ Female □ Date: _____

Instructions: Over the **last 7 days,** how often have you been bothered by any of the following problems? (Use "✓" to indicate your answer)

						Clinician Use
						Item score
		Not at all	**Several days**	**More than half the days**	**Nearly every day**	
1.	Little interest or pleasure in doing things	0	1	2	3	
2.	Feeling down, depressed, or hopeless	0	1	2	3	
3.	Trouble falling or staying asleep, or sleeping too much	0	1	2	3	
4.	Feeling tired or having little energy	0	1	2	3	
5.	Poor appetite or overeating	0	1	2	3	
6.	Feeling bad about yourself—or that you are a failure or have let yourself or your family down	0	1	2	3	
7.	Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3	
8.	Moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3	
9.	Thoughts that you would be better off dead or of hurting yourself in some way	0	1	2	3	
					Total/Partial Raw Score:	
					Prorated Total Raw Score: (if 1–2 items left unanswered)	

Adapted from Patient Health Questionnaire–9 (PHQ-9) for research and evaluation purposes.

PHQ9 Copyright © Pfizer Inc. PRIME-MD® is a trademark of Pfizer Inc.

Table 7.8–6.
The Altman Self-Rating Mania Scale (ASRM)

LEVEL 2—Mania—Adult*

*Altman Self-Rating Mania Scale (ASRM)

Name: _____ Age: _____ Sex: Male □ Female □ Date: _____

If the measure is being completed by on informant, what is your relationship with the individual receiving care? _____

In a typical week, approximately how much time do you spend with the individual receiving care? _____ hours/week

Instructions: On the DSM-5 Level 1 cross-cutting questionnaire you just completed, you indicated that *during the past 2 weeks* you (the individual receiving care) have been bothered by "sleeping less than usual, but still having a lot of energy" and/or "starting lots more projects than usual or doing more risky things than usual" at a mild or greater level of severity. The five statement groups or questions below ask about these feelings in more detail.

1. **Please read each group of statements/question carefully.**
2. Choose the one statement in each group that best describes the way you (the individual receiving care) have been feeling for the past week.
3. Check the box (✓ or ✗) next to the number/statement selected.
4. **Please note:** The word "occasionally" when used here means once or twice; "often" means several times or more and "frequently" means most of the time.

		Clinician Use
Question 1		**Item score**
□ 1	I do not feel happier or more cheerful than usual.	
□ 2	I occasionally feel happier or more cheerful than usual.	
□ 3	I often feel happier or more cheerful than usual.	
□ 4	I feel happier or more cheerful than usual most of the time.	
□ 5	I feel happier of more cheerful than usual all of the time.	
Question 2		
□ 1	I do not feel more self-confident than usual.	
□ 2	I occasionally feel more self-confident than usual.	
□ 3	I often feel more self-confident than usual.	
□ 4	I frequently feel more self-confident than usual.	
□ 5	I feel extremely self-confident all of the time.	
Question 3		
□ 1	I do not need less sleep than usual.	
□ 2	I occasionally need less sleep than usual.	
□ 3	I often need less sleep than usual.	
□ 4	I frequently need less sleep than usual.	
□ 5	I can go all day and all night without any sleep and still not feel tired.	
Question 4		
□ 1	I do not talk more than usual.	
□ 2	I occasionally talk more than usual.	
□ 3	I often talk more than usual.	
□ 4	I frequently talk more than usual.	
□ 5	I talk constantly and cannot be interrupted.	
Question 5		
□ 1	I have not been more active (either socially, sexually, at work, home, or school) than usual.	
□ 2	I have occasionally been more active than usual.	
□ 3	I have often been more active than usual.	
□ 4	I have frequently been more active than usual.	
□ 5	I am constantly more active or on the go all the time.	

Total/Partial Raw Score:	
Prorated Total Raw Score:	

From Altman EG, Hedeker D, Petersen JL, Davis JM. The Altman Self Rating Mania Scale. *Biol Psychiatry.* 1997;42:948–955, with permission from Elsevier.

The Montreal Cognitive Assessment. The Montreal Cognitive Assessment (MoCA) is a brief, 30-point cognitive screening measure that was developed in the late 1990s to address many of the shortcomings of the MMSE, and to assist in detecting MCI and early Alzheimer dementia (AD). The MoCA evaluates a number of cognitive domains including memory (immediate and short delayed recall), attention and working memory, visuospatial abilities, executive functioning, language, and orientation. MoCA scores range from 0 to 30 (with higher scores indicating better functioning), and a cut-score of 26 has been suggested as differentiating normal from abnormal scores. It can be administered by a busy physician or technician in as little as 5 to 10 minutes. Initial validation research indicates that sensitivity and specificity of the MoCA in detecting patients with MCI and AD have generally been promising and superior to the MMSE, although research in other settings has been somewhat mixed. As noted above with the MMSE, the MoCA may provide for a useful bedside evaluation and/or briefer assessment of cognitive functioning in settings where more in-depth testing is challenging, although more comprehensive neuropsychological testing may ultimately be necessary in making a differential diagnosis and/or providing a more nuanced picture of someone's level of functioning. This is particularly true in highly educated populations, where cognitive decline may be more difficult to detect due to ceiling effects in screening measures like the MMSE and MoCA.

Clinical Dementia Rating Scale. The Clinical Dementia Rating Scale (CDR) (Table 7.8–10) is a standardized staging system for dementia based on all available information about the patient. It was developed in 1992 and in 1993 was revised slightly in the context of a standard assessment protocol. Patients are rated in each of six domains (memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care) on a scale from 0 to 3 (0 = no dementia, 0.5 = questionable dementia [approximately equivalent to MCI], 1 = mild dementia, 2 = moderate dementia, 3 = severe dementia), with detailed descriptions to anchor the ratings in each area. The ratings in each domain are typically combined into an overall CDR rating on the same scale following a weighting system that favors memory. However, they can also be summed to generate a CDR Sum of Boxes score. The staging system can be used on its own or based on the CDR protocol, a semistructured interview with the patient and an informant, plus brief bedside cognitive testing. Ratings can be made relatively quickly by a clinician who is very familiar with the

patient, but the full protocol requires 1 hour or more to administer. CDR ratings of 4 (profound dementia) and 5 (terminal dementia) are based on overall status. Whether based on the CDR protocol or on clinical experience with the patient, the CDR must be rated by a trained clinician. Additional training in the use of the scale and protocol are strongly recommended. The reported reliability of CDR scores based on the full protocol is high, but achieving this level of reliability requires extensive training. Validity of the CDR has also been shown to be high. Although the CDR *protocol* is clearly intended for research, the CDR *scale* is of potential use as a stage marker for clinical communication and charting decline over time. It is also helpful for understanding the dementia research literature, as the CDR is increasingly used as a stage marker for clinical trials and other research. However, the CDR's lack of detail makes it relatively insensitive to change over short periods (the average annual decline is approximately one half-stage per year), and the CDR Sum of Boxes is only modestly better. In addition, the CDR was developed based on the typical course of Alzheimer disease, and tends to perform less well in patients with other dementing illnesses.

Table 7.8–7.
The PROMIS Emotional Distress—Anxiety Short-Form

LEVEL 2—Anxiety—Adult*

*PROMIS Emotional Distress—Anxiety—Short Form

Name: _____ Age: _____ Sex: Male □ Female □ Date: _____

If the measure is being completed by an informant, what is your relationship with the individual?

In a typical week, approximately how much time do you spend with the individual? hours/week

Instructions to patient: On the DSM-5 Level 1 cross-cutting questionnaire that you just completed, you indicated that *during the past 2 weeks* you (individual receiving care) have been bothered by "feeling nervous, anxious, frightened, worried, or on edge", "feeling panic or being frightened", and/or "avoiding situations that make you anxious" at a mild or greater level of severity. The questions below ask about these feelings in more detail and especially how often you (individual receiving care) have been bothered by a list of symptoms **during the past 7 days. Please respond to each item by marking (✓ or ✗) one box per row.**

							Clinician Use
In the past SEVEN (7) DAYS….							**Item Score**
		Never	**Rarely**	**Sometimes**	**Often**	**Always**	
1.	I felt fearful.	□ 1	□ 2	□ 3	□ 4	□ 5	
2	I felt anxious.	□ 1	□ 2	□ 3	□ 4	□ 5	
3.	I felt worried.	□ 1	□ 2	□ 3	□ 4	□ 5	
4.	I found it hard to focus on anything	□ 1	□ 2	□ 3	□ 4	□ 5	

	other than my anxiety.						
5.	I felt nervous.	□ 1	□ 2	□ 3	□ 4	□ 5	
6.	I felt uneasy.	□ 1	□ 2	□ 3	□ 4	□ 5	
7.	I felt tense.	□ 1	□ 2	□ 3	□ 4	□ 5	
						Total/Partial Raw Score:	
						Prorated Total Raw Score:	
						T-Score:	

©2008–2012 PROMIS Health Organization (PHO) and PROMIS Cooperative Group. Reproduced with permission.

Neuropsychiatric Inventory. The Neuropsychiatric Inventory (NPI) was developed in the mid-1990s to assess a wide range of behavioral symptoms that are often seen in Alzheimer disease and other dementing disorders. The current version rates 12 areas: Delusions, hallucinations, dysphoria, anxiety, agitation/aggression, euphoria, disinhibition, irritability/lability, apathy, aberrant motor behavior, nocturnal disturbances, appetite, and eating. The standard NPI is an interview with a caregiver or other informant that can be performed by a clinician or trained lay interviewer and requires 15 to 20 minutes to complete. There is also a nursing home interview version, the NPI-NH, and a self-report questionnaire, the NPI-Q. For each area, the NPI asks whether a symptom is present and, if so, assesses frequency, severity, and associated caregiver distress. The instrument has demonstrated reliability and validity and is useful to screen for problem behaviors in both clinical and research settings. Because of the detailed frequency and severity ratings, it is also useful to monitor change with treatment.

Table 7.8–8.
The Level 2 Substance Use Module from DSM-5

LEVEL 2—Substance Use—Adult*

*Adapted from the NIDA-Modified ASSIST

Name: _____ Age: _____ Sex: Male □ Female □ Date: _____

If the measure is being completed by an informant, what is your relationship with the individual receiving care?

In a typical week, approximately how much time do you spend with the individual receiving care? hours/week

Instructions: On the DSM-5 Level 1 cross-cutting questionnaire that you just completed, you indicated that *during the past 2 weeks* you (the individual receiving care) have been bothered by "using medicines on your own without a doctor's prescription, or in greater amounts or longer than prescribed, and/or using drugs like marijuana, cocaine or crack, and/or other drugs" at a

slight or greater level of severity. The questions below ask how often you (the individual receiving care) have used these medicines and/or substances **during the past 2 weeks. Please respond to each item by marking (✓ or ✗) one box per row.**

During the past **TWO (2) WEEKS,** about how often did you use any of the following medicines ON YOUR OWN, that is, without a doctor's prescription, in greater amounts or longer than prescribed?							**Clinician Use**
		Not at all	**One or two days**	**Several days**	**More than half the days**	**Nearly every day**	**Item Score**
a.	Painkillers (like Vicodin)	□ 0	□ 1	□ 2	□ 3	□ 4	
b.	Stimulants (like Ritalin, Adderall)	□ 0	□ 1	□ 2	□ 3	□ 4	
c.	Sedatives or tranquilizers (like sleeping pills or Valium)	□ 0	□ 1	□ 2	□ 3	□ 4	
Or drugs like:							
d.	Marijuana	□ 0	□ 1	□ 2	□ 3	□ 4	
e.	Cocaine or crack	□ 0	□ 1	□ 2	□ 3	□ 4	
f.	Club drugs (like ecstasy)	□ 0	□ 1	□ 2	□ 3	□ 4	
g.	Hallucinogens (like LSD)	□ 0	□ 1	□ 2	□ 3	□ 4	
h.	Heroin	□ 0	□ 1	□ 2	□ 3	□ 4	
i.	Inhalants or solvents (like glue)	□ 0	□ 1	□ 2	□ 3	□ 4	
j.	Methamphetamine (like speed)	□ 0	□ 1	□ 2	□ 3	□ 4	
						Total Score:	

Instructions to Clinicians

The DSM-5 Level 2—Substance Use—Adult is an adapted version of the NIDA-Modified ASSIST. The 15-item measure is used to assess the pure domain of prescription medicine, and illicit substance use in adults age 18 and older. It is completed by the individual prior to a visit with the clinician. If the individual receiving care is of impaired capacity and unable to complete the form (e.g., an individual with dementia), a knowledgeable informant may complete the measure. Each item asks the individual receiving care (or informant) to rate the severity of the individual's use of various substances **during the past 2 weeks**.

Scoring and Interpretation

Each item on the measure is rated on a 5-point scale (i.e., 0 = not at all; 1 = 1 or 2 days; 2 = several days; 3 = more than half the days; 4 = nearly every day). The clinician is asked to review the score of each item on the measure during the clinical interview and indicate the raw score for each item in the section provided for "Clinician Use." Scores on the individual items should be interpreted independently because each item inquires about the use of a distinct substance. The rating of multiple items at scores greater than 0 indicates greater severity and complexity of substance use.

Frequency of Use

To track change in the severity of the individual's use of alcohol, tobacco/nicotine, prescription or illicit substance over time, the measure be may completed at regular intervals as clinically indicated, depending on the stability of the individual's symptoms and treatment status. For individuals of impaired capacity, it is preferred that completion of the measures at follow-up appointments is by the same knowledgeable informant. Consistently high scores on the measure may indicate significant and problematic areas that might warrant further assessment, treatment, and follow-up. Your clinical judgment should guide your decision.

Courtesy of National Institute on Drug Abuse.

Table 7.8–9.
The Montreal Cognitive Assessment (MoCA)

MONTREAL COGNITIVE ASSESSMENT (MOCA)
Version 7.1 Original Version

NAME :
Education : Date of birth :
Sex : DATE :

VISUOSPATIAL / EXECUTIVE

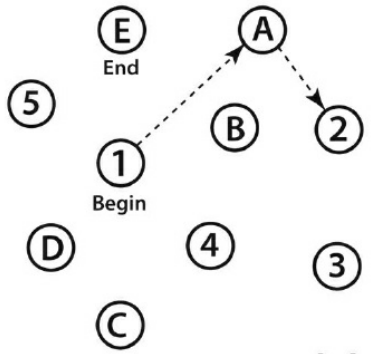

[ ]

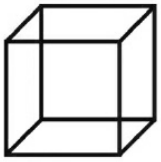

Copy cube [ ]

Draw CLOCK (Ten past eleven)
(3 points)

[ ] Contour [ ] Numbers [ ] Hands

POINTS __/5

NAMING

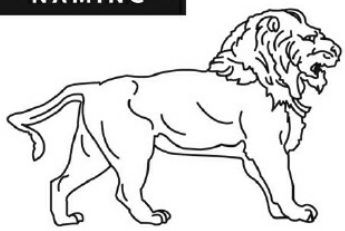

[ ]

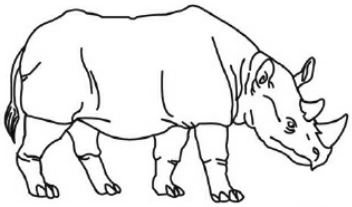

[ ]

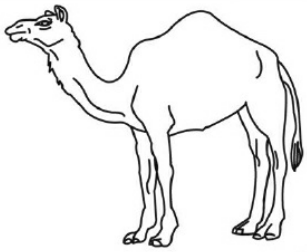

[ ]

__/3

MEMORY Read list of words, subject must repeat them. Do 2 trials, even if 1st trial is successful. Do a recall after 5 minutes.

	FACE	VELVET	CHURCH	DAISY	RED
1st trial					
2nd trial					

No points

ATTENTION Read list of digits (1 digit/ sec.).

Subject has to repeat them in the forward order [ ] 2 1 8 5 4
Subject has to repeat them in the backward order [ ] 7 4 2

__/2

Read list of letters. The subject must tap with his hand at each letter A. No points if ≥ 2 errors
[ ] F B A C M N A A J K L B A F A K D E A A A J A M O F A A B

__/1

Serial 7 subtraction starting at 100 [ ] 93 [ ] 86 [ ] 79 [ ] 72 [ ] 65
4 or 5 correct subtractions: **3 pts**, 2 or 3 correct: **2 pts**, 1 correct: **1 pt**, 0 correct: **0 pt**

__/3

LANGUAGE Repeat : I only know that John is the one to help today. [ ]
The cat always hid under the couch when dogs were in the room. [ ]

__/2

Fluency / Name maximum number of words in one minute that begin with the letter F [ ] _____ (N ≥ 11 words)

__/1

ABSTRACTION Similarity between e.g. banana - orange = fruit [ ] train – bicycle [ ] watch - ruler

__/2

DELAYED RECALL

	FACE	VELVET	CHURCH	DAISY	RED
Has to recall words **WITH NO CUE**	[ ]	[ ]	[ ]	[ ]	[ ]
Optional Category cue					
Optional Multiple choice cue					

Points for UNCUED recall only

__/5

ORIENTATION [ ] Date [ ] Month [ ] Year [ ] Day [ ] Place [ ] City

__/6

© Z.Nasreddine MD **www.mocatest.org**

Normal ≥26 / 30

TOTAL __/30
Add 1 point if ≤ 12 yr edu

Administered by: _____

From Montreal Cognitive Assessment (MoCA). Reproduced with permission.

From Montreal Cognitive Assessment (MoCA). Reproduced with permission.

Table 7.8–10.
Modified Clinical Dementia Rating (CDR): 0, 0.5, 1, 2, 3

Impairment	None (0)	Questionable (0.5)	Mild (1)	Moderate (2)	Severe (3)
Memory	No memory loss or slight inconstant forgetfulness	Consistent slight forgetfulness; partial recollection of events	Moderate memory loss; more marked for recent events; defect interferes with everyday activities	Severe memory loss; only highly learned material retained; new material rapidly lost	Severe memory loss; only fragments remain
Orientation	Fully oriented	Fully oriented or slight difficulty with time relationships	Moderate difficulty with time relationships; oriented for place at examination; may have geographic disorientation elsewhere	Severe difficulty with time relationships; usually disoriented in time, often to place	Oriented to person only
Judgment and problem	Solves everyday problems and handles business and financial affairs well; judgment good in relation to past performance	Slight impairment to solving problems, similarities, differences	Moderate difficulty in handling problems, similarities, differences; social judgment usually maintained	Severely impaired in handling problems, similarities, differences; social judgment usually impaired	Unable to make judgments or solve problems
Community affairs	Independent function at usual level in job, shopping, volunteer and social groups	Slight impairment in these activities	Unable to function independently at these activities though may still be engaged in some; appears normal to casual inspection	No pretense of independent function outside of home; appears well enough to be taken to functions outside of family home	No pretense of independent function outside of home; appears too ill to be taken to functions outside a family home
Home and hobbies	Life at home, hobbies, intellectual interests	Life at home, hobbies, intellectual interests	Mild but definite impairment of function at home; more	Only simple chores preserved; very	No significant function in home

	well maintained	slightly impaired	difficult chores abandoned; more complicated hobbies and interests abandoned	restricted interests, poorly maintained	
Personal care	Fully capable of self care	Fully capable of self care	Needs prompting	Requires assistance in dressing, hygiene keeping of personal effects	Requires much help with personal care; frequent incontinence

Note: Score only as decline from previous usual level due to cognitive loss, not impairment due to other factors.

Modified Clinical Dementia Rating Scale (CDR). Reproduced with permission.

Confusion Assessment Method. The Confusion Assessment Method (CAM) was developed in 1990 as a standardized method for identifying delirium. It is typically administered by a clinical interviewer (or lay interviewer with specialized training), and takes roughly 5 to 10 minutes to complete. Full and short-forms are also available. The CAM short-form specifically assesses for the presence/absence of the following in making a determination about possible delirium: (1) mental status altered from baseline (acute onset or fluctuating); (2) inattention; (3) disorganized thinking; and (4) altered level of consciousness. Delirium is identified when both conditions 1 and 2, and either 3 or 4 (or both) are met. The CAM long form assesses the same four domains above, as well as the following: (5) disorientation; (6) memory impairment; (7) perceptual disturbance; (8) psychomotor agitation and retardation; and (9) altered sleep–wake cycle. There has been a considerable amount of research on the CAM over the last few decades in support of its reliability and validity, and the tool has since been translated into 14 languages. It may have utility in hospital settings to detect and track delirium.

Personality Disorders and Personality Traits

Personality may be conceptualized categorically as personality disorders or dimensionally as personality traits, which may be viewed as normal or pathological. The focus here is on personality disorders and the maladaptive traits generally viewed as their milder forms. Although the most recent DSM-5 retained the diagnostic system used in DSM-IV for diagnosing personality disorders, an alternative hybrid DSM-5 model for personality disorders is also included in Section III. The APA noted that it would include both systems as a means of preserving current clinical

practice and continuity, while also allowing for additional research to build that will evolve these models and address many of the shortcomings in the current system. The DSM-5 continues to define ten personality disorders in three clusters. Patients tend not to fall neatly into these personality categories; instead, most patients who meet criteria for one personality disorder also meet criteria for at least one other, particularly within the same cluster. This and other limitations in the validity of the constructs themselves make it difficult to achieve validity in personality measures. Personality measures include both interviews and self-report instruments. Self-report measures are appealing in that they require less time and may appear less threatening to the patient. Because many of the symptoms suggesting personality problems are socially undesirable and because patients' insight tends to be limited, clinician-administered instruments, which allow for probing and patient observation, may provide more accurate data.

Structured Clinical Interview for DSM Personality Disorders. The Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II) is the counterpart of the SCID for making DSM diagnoses of personality disorders. The initial version was developed for the DSM-III-R in the mid-1990s, and the current version makes diagnoses according to the DSM-IV. As noted above, the SCID interviews are now being revised to be compatible with DSM-5, with expected release in 2015. However, given that DSM-5 retains the same diagnostic system for personality disorders used in DSM-IV, the SCID-II remains a useful tool in evaluating for and diagnosing personality disorders. The SCID-II is organized by disorder and includes all ten DSM personality disorders. A 119-item, self-report screening questionnaire is generally given first to eliminate sections not needing further exploration; each of the items corresponds to a specific criterion for a DSM personality disorder. The SCID-II proper includes one or two yes–no items for each criterion, with each affirmative answer to be followed up by examples from the person's life. On the basis of these answers, each criterion is scored 1 for false, 2 for subthreshold, and 3 for present, allowing criteria scores to be summed for a dimensional measure of each disorder or combined following the DSM diagnostic rules for a categorical approach. The screening questionnaire can be completed by the patient in approximately 20 minutes; the interview generally requires approximately 1 hour. The SCID-II must be administered by doctoral-level clinicians, and training in the SCID-II is also required. A computerized administration and scoring program is available. Reliability is good for the presence or absence of any disorder but only fair for specific personality disorders; the reliability of dimensional assessment is somewhat better. Validity is somewhat harder to determine. Agreement with clinician assessment tends to be modest, but given the instrument's

comprehensiveness and strict adherence to DSM criteria, the SCID-II may actually be more valid than clinician ratings. The SCID-II is most useful to provide a standardized, comprehensive assessment of personality disorders, whether in research, forensic, or clinical settings.

Personality Disorder Questionnaire. The Personality Disorder Questionnaire (PDQ) was developed in the late 1980s as a simple self-report questionnaire designed to provide categorical and dimensional assessment of DSM-III-R personality disorders and was subsequently revised for the DSM-IV and renamed the PDQ-4. Another alternate version allows for ratings within the last few weeks and is designed to serve as a change measure. The PDQ-4 includes 85 yes–no items, designed primarily to assess the diagnostic criteria for DSM-IV personality disorders. Within the 85 items, two validity scales are embedded to identify underreporting, lying, and inattention. There is also a brief clinician-administered Clinical Significance Scale to address the impact of any personality disorder identified by the self-report PDQ. The PDQ can provide categorical diagnoses, a scaled score for each, or an overall index of personality disturbance based on the sum of all of the diagnostic criteria. Overall scores range from 0 to 79; patients with personality disorders generally score above 30, psychotherapy outpatients without such disorders tend to score in the 20 to 30 range, and normal controls tend to score below 20. The PDQ can be completed in less than 30 minutes. Computerized administration and scoring are available. Reliability is fair to good for dimensional assessment and quite variable for categorical assessment, with good reliability for obsessive-compulsive and antisocial personality disorders and inadequate reliability for many disorders. Validity judged against semistructured clinician-administered interviews is also variable. The PDQ, like other self-report instruments, tends to overdiagnose personality disorders, with many false positives and few false negatives. Its brevity, good sensitivity, and poor specificity make it most useful as a screening device, with patients screening positive being followed up with a detailed clinical evaluation or administration of a semistructured interview.

The Personality Inventory for DSM-5 and Personality Inventory for DSM-5 Brief Form. The Personality Inventory for DSM-5 (PID-5) was developed in 2012 to measure the five higher-order domains (or dimensions) that are included in the alternative Section III of DSM-5, and have been shown to account for a significant amount of the variance observed across the personality disorders: Negative Affect, Detachment, Antagonism, Disinhibition, and Psychoticism. These five domains encompass twenty-five lower order "facets," which themselves represent different features of the higher-order construct. For example, the higher-order dimension of Negative Affect may include lower-order features of anxiety, depressivity,

affective instability, and anhedonia. The PID-5 is a 220-item measure that yields scores at both the domain and facet levels, while the Personality Inventory for DSM-5 Brief Form (PID-5-BF) is a briefer 25-item measure designed to yield scores at the domain level only. Initial psychometric testing of these measures has been promising, and in many ways these tools reflect more recent conceptualizations of personality functioning—where the combination and severity of these traits are seen to have more predictive power than a simple criteria-based approach alone. Regardless, these instruments were published only in 2012, so further research is needed to better understand their contribution and clinical utility.

Childhood Disorders

A wide variety of instruments are available to assess mental disorders in children. Despite this rich array of instruments, however, the evaluation of children remains difficult for several reasons. First, the child psychiatric nosology is at an earlier stage of development, and construct validity is often problematic. Multiple changes in diagnostic criteria from DSM-III to DSM-III-R to DSM-IV to DSM-5 complicate the choice of measures. Second, because children change markedly with age, it is virtually impossible to design a measure that covers children of all ages. Last, because children, particularly young children, have limited ability to report their symptoms, other informants are necessary. This often creates problems because there are frequent disagreements among child, parent, and teacher reports of symptoms, and the optimal way to combine information is unclear.

Child Behavior Checklist. The Child Behavior Checklist (CBCL) is a family of self-rated instruments that survey a broad range of difficulties encountered in children from preschool through adolescence. One version of the CBCL, designed for completion by parents of children between 6 and 18 years of age, is shown in Table 7.8–11. Another version is available for parents of children between 2 and 3 years of age. The Youth Self-Report is completed by children between 11 and 18 years of age, and the Teacher Report Form is completed by teachers of school-age children. The scale includes not only problem behaviors, but also academic and social strengths. Each version includes approximately 100 items scored on a three-point Likert scale. Scoring can be done by hand or computer, and normative data are available for each of the three subscales: Problem behaviors, academic functioning, and adaptive behaviors. A computerized version is also available. The CBCL does not generate diagnoses but, instead, suggests cutoff scores for problems in the "clinical range." Parent, teacher, and child versions all show high reliability on the problem subscale, but the three informants frequently do not agree with one another. The CBCL may be useful in clinical settings as an adjunct to

clinical evaluation, as it provides a good overall view of symptomatology and may also be used to track change over time. It is used frequently for similar purposes in research involving children and, thus, can be compared to clinical experience. The instrument does not, however, provide diagnostic information, and its length limits its efficiency for tracking purposes.

Table 7.8–11.
Child Behavior Checklist for Age 6–18

Please print

CHILD BEHAVIOR CHECKLIST FOR AGES 6–18

For office use only
ID #

CHILD'S FULL NAME First Middle Last

CHILD'S GENDER ☐ Boy ☐ Girl

CHILD'S AGE

CHILD'S ETHNIC GROUP OR RACE

TODAY'S DATE Mo. ____ Day ____ Year ____

CHILD'S BIRTHDATE Mo. ____ Day ____ Year ____

GRADE IN SCHOOL _____

NOT ATTENDING SCHOOL ☐

Please fill out this form to reflect your view of the child's behavior even if other people might not agree. Feel free to print additional comments beside each item and in the space provided on page 2. ***Be sure to answer all items.***

PARENTS' USUAL TYPE OF WORK, even if not working now.
(Please be specific — for example, auto mechanic, high school teacher, homemaker, laborer, lathe operator, shoe salesman, army sergeant.)

FATHER'S TYPE OF WORK _____

MOTHER'S TYPE OF WORK _____

THIS FORM FILLED OUT BY: (print your full name)

_____

Your gender: ☐ Male ☐ Female

Your relation to the child:
☐ Biological Parent ☐ Step Parent ☐ Grandparent
☐ Adoptive Parent ☐ Foster Parent ☐ Other (specify)

I. Please list the sports your child most likes to take part in. For example: swimming, baseball, skating, skate boarding, bike riding, fishing, etc.

Compared to others of the same age, about how much time does he/she spend in each?

Compared to others of the same age, how well does he/she do each one?

	Less Than Average	Average	More Than Average	Don't Know	Below Average	Average	Above Average	Don't Know
☐ None								
a. _____	☐	☐	☐	☐	☐	☐	☐	☐
b. _____	☐	☐	☐	☐	☐	☐	☐	☐
c. _____	☐	☐	☐	☐	☐	☐	☐	☐

II. Please list your child's favorite hobbies, activities, and games, other than sports. For example: stamps, dolls, book, piano, crafts, cars, computers, singing, etc. (Do ***not*** include listening to radio or TV.)

Compared to others of the same age, about how much time does he/she spend in each?

Compared to others of the same age, how well does he/she do each one?

	Less Than Average	Average	More Than Average	Don't Know	Below Average	Average	Above Average	Don't Know
☐ None								
a. _____	☐	☐	☐	☐	☐	☐	☐	☐
b. _____	☐	☐	☐	☐	☐	☐	☐	☐
c. _____	☐	☐	☐	☐	☐	☐	☐	☐

III. Please list any organizations, clubs, teams, or groups your child belongs to.

Compared to others of the same age, how active is he/she in each?

	Less Active	Average	More Active	Don't Know
☐ None				
a. _____	☐	☐	☐	☐
b. _____	☐	☐	☐	☐
c. _____	☐	☐	☐	☐

IV. Please list any jobs or chores your child has. For example: paper route, babysitting, making bed, working in store, etc. (Include both paid and unpaid jobs and chores.)

Compared to others of the same age, how well does he/she carry them out?

	Below Average	Average	Above Average	Don't Know
☐ None				
a. _____	☐	☐	☐	☐
b. _____	☐	☐	☐	☐
c. _____	☐	☐	☐	☐

Be sure you answered all items. Then see other side.

SAMPLE

Copyright 2001 T. Achenbach
ASEBA, University of Vermont
1 South Prospect St., Burlington, VT 05401-3456
www.ASEBA.org

UNAUTHORIZED COPYING IS ILLEGAL

6-1-01 Edition - 201

Please print. Be sure to answer all items.

V. 1. About how many close friends does your child have? (Do *not* include brothers & sisters)

☐ None ☐ 1 ☐ 2 or 3 ☐ 4 or more

2. About how many times a week does your child do things with any friends outside of regular school hours? (Do *not* include brothers & sisters)

☐ Less than 1 ☐ 1 or 2 ☐ 3 or more

VI. Compared to others of his/her age, how well does your child:

	Worse	Average	Better	
a. Get along with his/her brothers & sisters?	☐	☐	☐	☐ Has no brothers or sisters
b. Get along with other kids?	☐	☐	☐	
c. Behave with his/her parents?	☐	☐	☐	
d. Play and work alone?	☐	☐	☐	

VII. 1. Performance in academic subjects. ☐ Does not attend school because _____

Check a box for each subject that child takes	Failing	Below Average	Average	Above Average
a. Reading, English, or Language Arts	☐	☐	☐	☐
b. History or Social Studies	☐	☐	☐	☐
c. Arithmetic or Math	☐	☐	☐	☐
d. Science	☐	☐	☐	☐
e. _____	☐	☐	☐	☐
f. _____	☐	☐	☐	☐
g. _____	☐	☐	☐	☐

Other academic subjects–for example: computer courses, foreign language, business. Do ***not*** include gym, shop, driver's ed., or other nonacademic subjects.

2. Does your child receive special education or remedial services or attend a special class or special school?

☐ No ☐ Yes—kind of services, class, or school:

3. Has your child repeated any grades? ☐ No ☐ Yes—grades and reasons:

4. Has your child had any academic or other problems in school? ☐ No ☐ Yes—please describe:

When did these problems start?

Have these problems ended? ☐ No ☐ Yes–when?

Does your child have any illness or disability (either physical or mental)? ☐ No ☐ Yes—please describe:

What concerns you most about your child?

Please describe the best things about your child.

SAMPLE

Be sure you answered all items.

Please print. Be sure to answer all items.

Below is a list of items that describe children and youths. For each item that describes your child ***now or within the past 6 months***, please circle the ***2*** if the item is ***very true or often true*** of your child. Circle the ***1*** if the item is ***somewhat or sometimes true*** of your child. If the item is ***not true*** of your child, circle the ***0***. Please answer all items as well as you can, even if some do not seem to apply to your child.

0 = Not True (as far as you know) **1 = Somewhat or Sometimes True** **2 = Very True or Often True**

0 1 2 1. Acts too young for his/her age
0 1 2 2. Drinks alcohol without parents' approval (describe): _____
_____
0 1 2 3. Argues a lot
0 1 2 4. Fails to finish things he/she starts
0 1 2 5. There is very little he/she enjoys
0 1 2 6. Bowel movements outside toilet
0 1 2 7. Bragging, boasting
0 1 2 8. Can't concentrate, can't pay attention for long
0 1 2 9. Can't get his/her mind off certain thoughts; obsessions (describe): _____
_____
0 1 2 10. Can't sit still, restless, or hyperactive
0 1 2 11. Clings to adults or too dependent
0 1 2 12. Complains of loneliness
0 1 2 13. Confused or seems to be in a fog
0 1 2 14. Cries a lot
0 1 2 15. Cruel to animals
0 1 2 16. Cruelty, bullying, or meanness to others
0 1 2 17. Daydreams or gets lost in his/her thoughts
0 1 2 18. Deliberately harms self or attempts suicide
0 1 2 19. Demands a lot of attention
0 1 2 20. Destroys his/her own things
0 1 2 21. Destroys things belonging to his/her family or others
0 1 2 22. Disobedient at home
0 1 2 23. Disobedient at school
0 1 2 24. Doesn't eat well
0 1 2 25. Doesn't get along with other kids
0 1 2 26. Doesn't seem to feel guilty after misbehaving
0 1 2 27. Easily jealous
0 1 2 28. Breaks rules at home, school, or elsewhere
0 1 2 29. Fears certain animals, situations, or places, other than school (describe): _____
_____
0 1 2 30. Fears going to school
0 1 2 31. Fears he/she might think or do something bad
0 1 2 32. Feels he/she has to be perfect
0 1 2 33. Feels or complains that no one loves him/her
0 1 2 34. Feels others are out to get him/her
0 1 2 35. Feels worthless or inferior
0 1 2 36. Gets hurt a lot, accident-prone
0 1 2 37. Gets in many fights
0 1 2 38. Gets teased a lot
0 1 2 39. Hangs around with others who get in trouble
0 1 2 40. Hears sound or voices that aren't there (describe): _____
_____
0 1 2 41. Impulsive or acts without thinking
0 1 2 42. Would rather be alone than with others
0 1 2 43. Lying or cheating
0 1 2 44. Bites fingernails
0 1 2 45. Nervous, highstrung, or tense
0 1 2 46. Nervous movements or twitching (describe):
_____
_____
0 1 2 47. Nightmares
0 1 2 48. Not liked by other kids
0 1 2 49. Constipated, doesn't move bowels
0 1 2 50. Too fearful or anxious
0 1 2 51. Feels dizzy or lightheaded
0 1 2 52. Feels too guilty
0 1 2 53. Overeating
0 1 2 54. Overtired without good reason
0 1 2 55. Overweight
56. Physical problems ***without known medical cause:***
0 1 2 a. Aches or pains (***not*** stomach or headaches)
0 1 2 b. Headaches
0 1 2 c. Nausea, feels sick
0 1 2 d. Problems with eyes (***not*** if corrected by glasses) (describe): _____
0 1 2 e. Rashes or other skin problems
0 1 2 f. Stomachaches
0 1 2 g. Vomiting, throwing up
0 1 2 h. Other (describe): _____
_____

SAMPLE

Please print. Be sure to answer all items.

0 = Not True (as far as you know) **1 = Somewhat or Sometimes True** **2 = Very True or Often True**

0 1 2 57. Physically attacks people
0 1 2 58. Picks nose, skin, or other parts of body (describe): _____
_____
0 1 2 59. Plays with own sex parts in public
0 1 2 60. Plays with own sex parts too much
0 1 2 61. Poor school work
0 1 2 62. Poorly coordinated or clumsy
0 1 2 63. Prefers being with older kids
0 1 2 64. Prefers being with younger kids
0 1 2 65. Refuses to talk
0 1 2 66. Repeats certain acts over and over; compulsions (describe): _____
_____
0 1 2 67. Runs away from home
0 1 2 68. Screams a lot
0 1 2 69. Secretive, keeps things to self
0 1 2 70. Sees things that aren't there (describe): ___
_____
_____
0 1 2 71. Self-conscious or easily embarrassed
0 1 2 72. Sets fires
0 1 2 73. Sexual problems (describe): _____
_____
_____
0 1 2 74. Showing off or clowning
0 1 2 75. Too shy or timid
0 1 2 76. Sleeps less than most kids
0 1 2 77. Sleeps more than most kids during day and/or night (describe): _____
_____
0 1 2 78. Inattentive or easily distracted
0 1 2 79. Speech problem (describe): _____
_____
0 1 2 80. Stares blankly
0 1 2 81. Steals at home
0 1 2 82. Steals outside the home
0 1 2 83. Stores up too many things he/she doesn't need (describe): _____
_____
_____
0 1 2 84. Strange behavior (describe): _____
_____
0 1 2 85. Strange ideas (describe): _____
_____
0 1 2 86. Stubborn, sullen, or irritable
0 1 2 87. Sudden changes in mood or feelings
0 1 2 88. Sulks a lot
0 1 2 89. Suspicious
0 1 2 90. Swearing or obscene language
0 1 2 91. Talks about killing self
0 1 2 92. Talks or walks in sleep (describe): _____
_____
0 1 2 93. Talks too much
0 1 2 94. Teases a lot
0 1 2 95. Temper tantrums or hot temper
0 1 2 96. Thinks about sex too much
0 1 2 97. Threatens people
0 1 2 98. Thumb-sucking
0 1 2 99. Smokes, chews, or sniffs tobacco
0 1 2 100. Trouble sleeping (describe): _____
_____
0 1 2 101. Truancy, skips school
0 1 2 102. Underactive, slow moving, or lacks energy
0 1 2 103. Unhappy, sad, or depressed
0 1 2 104. Unusually loud
0 1 2 105. Uses drugs for nonmedical purposes (***don't*** include alcohol or tobacco) (describe): ____
_____
_____
0 1 2 106. Vandalism
0 1 2 107. Wets self during the day
0 1 2 108. Wets the bed
0 1 2 109. Whining
0 1 2 110. Wishes to be of opposite sex
0 1 2 111. Withdrawn, doesn't get involved with others
0 1 2 112. Worries
113. Please write in any problems your child has that were not listed above:
0 1 2 _____
_____
0 1 2 _____

 Please be sure you answered all items.

From Achenbach TM, University of Vermont, 1 S. Prospect St., Burlington, VT 05401. © by T. M. Achenbach. Reproduced with permission.

SAMPLE

From Achenbach TM, University of Vermont, 1 S. Prospect St., Burlington, VT 05401. © by T. M. Achenbach. Reproduced with permission.

Diagnostic Interview Schedule for Children. The DISC was originally developed in the early 1990s as a fully structured diagnostic interview for making DSM-III diagnoses in children. It has since been revised for the DSM-III-R and DSM-IV, and efforts are now underway to make it compatible with DSM-5. The current DISC, the DISC-IV, covers a broad range of DSM-IV diagnoses, both current and lifetime. It has nearly 3,000 questions but is structured with a series of stem questions that serve as gateways to each diagnostic area, with the remainder of each section skipped if the subject answers no. Subjects who enter each section have very few skips, so complete diagnostic and symptom scale information can be obtained. Child, parent, and teacher versions are available. Computer programs are available to implement diagnostic criteria and generate severity scales based on each version or to combine parent and child information. A typical DISC interview may take more than 1 hour for a child, plus an additional hour for a parent. However, because of the stem question structure, the actual time varies widely with the number of symptoms endorsed. The DISC was designed for lay interviewers. It is fairly complicated to administer, and formal training programs are highly recommended. Reliability of the DISC is only fair to good and generally better for the combined child and parent interview. Validity judged against a clinical interview by a child psychiatrist is also fair to good—better for some diagnoses and better for the combined interview. The DISC is well tolerated by parents and children and can be used to supplement a clinical interview to ensure comprehensive diagnostic coverage. Because of its inflexibility, some clinicians find it uncomfortable to use, and its length makes it less than optimal for use in clinical practice. However, it is used frequently in a variety of research settings.

Conners Rating Scales. The Conners Rating Scales are a family of instruments designed to measure a range of childhood and adolescent psychopathology but are most commonly used in the assessment of ADHD. There are teacher, parent, and self-report (for adolescents) versions and both short (as few as 10 items) and long (as many as 80 items, with multiple subscales) forms. Extensive normative data drawn from an ethnically diverse population are available for each sex across a broad age range. Even the longer forms can be completed in 15 to 20 minutes, and scoring can be accomplished rapidly. Training for raters is not required. Reliability data are excellent for the Conners Rating Scales. However, teacher and parent versions tend to show poor agreement. Validity data suggest that the Conners Rating Scales are excellent at discriminating between ADHD patients and normal controls. It has more difficulty separating ADHD from other disruptive behavioral disorders, such as conduct disorder, but this may be partly related to the genuine clinical difficulties separating these syndromes. Newer versions of the Conners

Rating Scales have been developed that aim to improve these discriminations but have not yet been subjected to extensive testing. The principal uses of the Conners Rating Scales are in screening for ADHD in school or clinic populations and following changes in symptom severity over time; sensitivity to change in response to specific therapies has been demonstrated for most versions of the Conners Rating Scales.

Autism Diagnostic Interview—Revised. The Autism Diagnostic Interview (ADI) was developed in 1989 as a clinical assessment of autism and related disorders. The revised ADI-R was developed in 2003 with an aim to providing a shorter instrument with better ability to discriminate autism from other developmental disorders. The instrument has 93 items, is designed for individuals with a mental age greater than 18 months, and covers 3 broad areas, consistent with the DSM-5 criteria for autism: Language and communication; reciprocal social interactions; and restricted, repetitive, and stereotyped behaviors and interests. There are three versions: One for lifetime diagnosis, one for current diagnosis, and one for patients under 4 focused on considering an initial diagnosis of autism. It must be administered by a clinician trained in its use and takes about 90 minutes to complete. When clinicians are properly trained, it has good to excellent reliability and validity but performs poorly in the setting of severe developmental disabilities. It is generally intended for the research setting, when a thorough assessment of autism is required but may have use in clinical practice as well.

REFERENCES

Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman self-rating mania scale. *Biological psychiatry*. 1997;42:948–955.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. 5th ed. Washington, DC: American Psychiatric Publishing Incorporated, 2013.

Blais MB, Sinclair SJ. *Initial Construction, Psychometric Evaluation, and Validation of the Psychopathology Spectra Inventory (PSI): A Measure Based on Quantitative Models of Psychopathology*. Boston, MA: Massachusetts General Hospital, 2014.

Buka SL, Monuteaux M, Earls F. Epidemiology of child and adolescent psychiatric disorders. In: Tsuang MT, Tohen M, eds. *Textbook in Psychiatric Epidemiology*. 2nd ed. New York: Wiley-Liss; 2002.

*Burt T, Sederer L, Isgak WW, eds. *Outcome Measurement in Psychiatry: A Critical Review*. Washington, DC: American Psychiatric Press; 2002.

Butcher JN, Dahlstrom WG, Graham JR, Tellegen A, Kaemmer B. *MMPI-2: Minnesota Multiphasic Personality Inventory–2: Manual for administration and scoring*. Minneapolis: University of Minnesota Press; 1989.

Carmines E, Zeller R. Reliability and validity assessment. Beverly Hills: Sage Publications; 1979.

*First MB, ed. *Standardized Evaluation in Clinical Practice: Review of Psychiatry*. Vol. 22.

Washington, DC: American Psychiatric Press; 2003.

Hulley SB, Cummings SR. Planning the measurements: precision and accuracy. In: Hulley SB, Cummings SR, eds. *Designing Clinical Research: An Epidemiologic Approach*. Baltimore, MD: Williams and Wilkins; 1987.

Kendell R, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. *Am J Psychiatry*. 2003;160:4–12.

Kessler RC. The categorical versus dimensional assessment controversy in the sociology of mental illness. *J Health Soc Behav*. 2002;43:171–188.

Krueger RF. The structure of common mental disorders. *Archives of General Psychiatry*. 1999;56:921–926.

Krueger RF, Derringer J, Markon KE, Watson D, Skodol AE. Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine,* 2012;42:1879–1890.

McDowell I, Newell C. *Measuring Health: A Guide to Rating Scales and Questionnaires*. London: Oxford University Press; 2006.

Morey LC. *Personality Assessment Inventory: Professional Manual*. 2nd ed. Odessa, FL: Psychological Assessment Resources; 2007.

*Murphy J. Symptom scales and diagnostic schedules in adult psychiatry. In: Tsuang MT, Tohen M, eds. *Textbook in Psychiatric Epidemiology*. 2nd ed. New York: Wiley-Liss; 2002.

Murphy JM, Berwick DM, Weinstein MC, Borus JF, Budman SH. Performance of screening and diagnostic tests: application of receiver operating characteristic (ROC) analysis. *Arch Gen Psychiatry*. 1987;44:550–555.

Myers K, Winters NC. Ten-year review of rating scales. I: Overview of scale functioning, psychometric properties, and selection. *J Am Acad Child Adolesc Psychiatry*. 2002;41:114–122.

Regier DA. Dimensional approaches to psychiatric classification: refining the research agenda for DSM-V: an introduction. *Int J Methods Psychiatr Res*. 2007;16 Suppl 1:S1–S5.

*Regier DA, Kaelber CT, Rae DS, et al. Limitations of diagnostic criteria and assessment instruments for mental disorders: implications for research and policy. *Arch Gen Psychiatry*. 1998;55:109–115.

Robins LN. Epidemiology: reflections on testing the validity of psychiatric interviews. *Arch Gen Psychiatry*. 1985;42:918–924.

Robins E, Guze SB. Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. *Am J Psychiatry*. 1990;126:983–987.

*Rush J, First MB, Blacker D, eds. *Handbook of Psychiatric Measures*. 2nd ed. Washington, DC: American Psychiatric Press; 2007.

Schutte NS, Malouff JM. *Sourcebook of Adult Assessment Strategies*. New York: Plenum Press; 1995.

Shrout PE, Spitzer RL, Fleiss JL. Quantification of agreement in psychiatric diagnosis revisited. *Arch Gen Psychiatry*. 1998;744:172–177.

Spitzer RL. Psychiatric diagnosis: are clinicians still necessary? *Compr Psychiatry*. 1983;24:399–411.

Spitzer RL, Fleiss JL. A re-analysis of the reliability of psychiatric diagnosis. *Br J Psychiatry*.

1974;125:341–347.

Sudman S, Bradburn NM, Schwarz N. *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco, CA: Jossey-Bass; 1996.

Winokur G, Zimmerman M, Cadoret R. Cause the Bible tells me so. *Arch Gen Psychiatry*. 1988;45:683–684.

Zachar P, Kendler KS. Psychiatric disorders: a conceptual taxonomy. *Am J Psychiatry*. 2007;164:557–565.

Zarin DA, Earls F. Diagnostic decision making in psychiatry. *Am J Psychiatry*. 1993;150:197–206.

▲ 7.9 Electronic Media in Psychiatry

ZEBULON TAINTOR, M.D.

INTRODUCTION

The widespread use of mobile devices, electronic media and microprocessors for communication, information, and as auxiliary thinking aids is revolutionizing society and the practice of medicine, including psychiatry. This subsection considers the use of remote communications with patients and developments in uses of medical record uses because of the applications of microprocessors in both areas. It does not include use of devices that stimulate the brain through any means other than the five senses.

DEFINITIONS

The prefix "tele," denoting transmission over distance, is used to form *telegraph,* involving the transmission of text; *telephone,* involving the transmission of sounds; and *television,* the transmission of images, usually with sound. *Telemedicine* is the practice of medicine using the electronic transmission of images, and *telepsychiatry* is the practice of the medical specialty of psychiatry using the electronic transmission of images between two locations not in other visual contact. The prefix *radio* is used to refer to radiation, whereas *wireless* is used to refer to the transmission of radio waves. *Video,* a potentially more accurate prefix, is used in *videotherapy* to indicate the use of electronic images without transmission to distant points. *Virtual reality* therapy surrounds the patient with a simulated environment, using helmets or wraparound glasses/goggles. *Holographic* therapy will develop as projection of holographic images becomes widespread, blending reality with virtual reality. *E-therapy* is therapy done by electronic communication. At present this means exchanging text messages (via e-mail or texting with a handheld device), although images are being added with increasing ease. *Apps* (applications) are computer programs developed for handheld devices that interact with the user. *Online* refers to the subset of e-therapy done in real time, aiming for pauses between messages that are as brief as those between two people in the same room. *Asynchronous online* denotes responses with time lapses in between. *Unmet e-therapy* is e-therapy in which the client and the therapist have not met in person. There is no definition of treatment otherwise, and websites offering unmet e-therapy uniformly state, probably to reduce malpractice liability, that counseling rather than therapy is what is being offered. A *webcam* is a camera used in conjunction with a personal computer for video communication via the Internet.

Other definitions have emerged with electronic storage and retrieval of medical records. The Health Information Portability and Accountability Act (HIPAA), amended several times since enacted in 1996, defines most terms for electronic records in health care. *Protected health information* (PHI) is any information that can be linked to a specific individual. It can be shared as needed among *covered entities* (providers, health plans, insurers, etc., anyone transmitting data in such a way as regulated by HIPAA) and their business associates. Without a person's consent, it can be revealed as required for payment and, as required by law, to certain noncovered entities by law or regulations: law enforcement, in cases of suspected child abuse, etc. whether a covered entity or not, anyone receiving PHI is expected to maintain privacy and security. While a person requesting access to his/her PHI must be able to see it within 30 days and can request correction of inaccurate data, HIPAA does not state who owns PHI, just that it must be protected. Each covered entity is required to have a Privacy Officer, have privacy policies available on request, and to keep track of disclosures of PHI. Individuals may complain to the Department of Health and Human Services (HHS) of suspected privacy violations. Under Administrative Simplification (a title seen by many as ironic) HHS set rules for Privacy, Security, Transactions and Code Sets, Unique Identifiers, and Enforcement.

The movement to electronic health records (EHRs) has required standards that must be met for certification. Initially, certification was done solely by the *Certification Commission for Health Information Technology* (CCHIT; available at http://www.cchit.org), which for several years was to electronic medical records (EMRs) what the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) is to hospitals and clinics. CCHIT accredited inpatient and ambulatory EMRs, requiring certain defined data elements and internal operations. Workgroups set standards for privacy and compliance, security, *networking* (connecting electronic devices), *interoperability* (to enable electronic records to be uploaded to, and used via, other electronic devices), and agreement on a basic standard data package that one physician should expect from another, the *continuity of care record* (CCHR). *Granularity* enables the item-by-item retention or release of any item of PHI to any specific person or agency. Granularity is accepted as a goal, but is not yet offered widely. Usually patients are offered an opt-in or opt-out choice as to whether all of their health information will be shared. EMRs include all of the traditional components of a patient's medical chart but not administrative and billing data used in the practice management software (PMS) already in use in many facilities and practices. A patient may keep a *personal health record* (PHR), available at no cost in various unstandardized formats from dozens of vendors, often combined with wellness programs and allowing storage on the Internet, memory sticks, etc. There is no doubt that a person owns

his/her PHR and is responsible for reviewing its accuracy and maintenance. A 2014 controlled study showed improved outcomes for persons with serious mental illness who were empowered by using their PHRs.

HISTORY

Therapeutic uses of electronic media have evolved along with the media themselves and uses of their predecessors.

Written Words: From Letter Writing to Instant Messaging

An e-mail basically is a letter, the most ancient regular form of communication across distance. People use letters to develop and give a particular form to relationships. Various published correspondences often include the writers' agreement that the letters were more important than actual meetings. Numerous examples show the process of psychotherapy in traditional letter writing. Sigmund Freud developed several therapeutic correspondences. The first "e-mail" device was the telegraph, which was first used between cities in 1844. Telegraphic styles survive today in instant and text messaging, aided by inventive shorthand that allows numeric keypads to be used. Text messaging has become so ubiquitous, intrusive and attention-diverting that accidents and regulations have increased.

E-mail has been increasingly accepted in general medicine, especially for anything a patient might have to write down if told on the telephone, such as test results, written instructions, directions, and phone numbers. Cost per communication is insignificant, usually a monthly charge for Internet usage that provides unlimited e-mail. E-mail use has skyrocketed under this pricing structure, which does not depend on the number of units or distance traveled. Increased use was recommended by the Institute of Medicine's Committee on Quality of Health Care in America in their 2001 report *Crossing the Quality Chasm: A New Health System for the 21st Century*. Almost all e-mail is retrievable and discoverable in litigation.

E-mail messages often are linked to websites for patient education. The professional and self-help organizations operating these sites have developed dozens of apps for patients to use in times of need, for example, coping with anxiety, suicidal thoughts, detecting depression. The apps are programmed for interactive texting without involving another person. Some patients have found such apps very supportive and describe how tragedies were averted.

Spoken Words: Telephone

Telephone calls have been used to help people feel better since the invention of the telephone in 1885. They are immediate and, in addition to conveying verbal information, can convey nonverbal data through pauses,

tone of voice, and inflections. When costs were computed by time and distance, telephone calls tended to be brief. Now many mobile plans offer high or unlimited call times and therapists routinely schedule long sessions with patients when one or the other is traveling. Unless recorded, they are not retrievable, which may or may not be an advantage. If they are recorded, playback speedup is essential to achieve evaluation times that can match the ease with which written exchanges can be scanned. Telephone rings tend to be intrusive and may not catch the recipient in a receptive mindset. Leaving messages on answering machines frustrates the very urgency that led to the call and can lead to complaints about "telephone tag," feeding the rise of texting.

Crisis hot lines are widely used in emergencies and are accepted as one part of the range of services. Medical curricula typically do not include how to use telephones in patient care, an issue receiving belated attention. There is extensive experience with telephone therapy in the form of suicide prevention centers and other hot lines. Answerers are always available. The caller and answerer have not talked previously. Suicide prevention centers once were publicized extensively and were busy, answering many calls and serving many people in a crisis mode. Studies showed that suicide prevention centers did not reduce suicide rates in the cities they served, a phenomenon attributed to the likelihood that more seriously suicidal patients did not call but instead committed suicide. Generally, callers used hot lines once or twice. They received support and were urged to go to services where they were seen and helped. Some used crisis hot lines and emergency rooms for months or years before getting involved in more regular treatment, and others did not progress beyond the use of such lines. Some proponents of e-therapy argue from the hot line experience that some people in need of therapy will get involved only via the Internet because of its convenience and anonymity.

Use of the telephone to transmit images became widely available with the advent of facsimile machines, which have been used in published correspondence but minimally in therapy. Although videophones have been available for decades to transmit moving images in real time, they have become popular only in the past few years for photos, video recording of events and face-to-face conversations. Yet many with this capability choose to text, so as not to show their image spontaneously. People in the same room who could talk to one another may sometimes text. Webcams and smart phones with unlimited low-cost time via Internet teleconferencing have merged the telephone with telepsychiatry and made it possible to provide home health care monitoring as frequently as needed. Physiological variables, such as blood pressure and electrocardiograms, can be monitored exactly.

Spoken Word Recording

Recording of psychotherapy sessions has been used by both patients and physicians since audiotape became available in the 1940s, especially with the development of audio-cassettes that allow for playback in automobiles and elsewhere. The technology lends itself to easy storage and retrieval. State laws vary as to what permission must be obtained from the other party(ies) in a conversation before recording it, whether in telephone or in person. It is best to err on the side of caution and ask permission.

Images: Television

Although television was invented prior to World War II, its use in therapy awaited the arrival of recording media, specifically videotape. Milton Berger demonstrated the use of videotaping group therapy and other psychotherapy sessions to help patients see themselves as others saw them, and especially to show how their body language conveyed unintended messages. Videotapes were used for storing data before forwarding them to others. This technology gave way to digital video discs (DVDs), flash memory sticks, and storage on the Internet.

Telemedicine

The first recorded use of telepsychiatry was in 1957. Decreasing costs for equipment and transmission (typically by dedicated telephone lines) and improved image quality have resulted in more widespread use for particular populations. Telepsychiatry has benefited from being seen as part of telemedicine. Inevitably some data are lost off camera. Data reduction is less of an issue for some other specialties, such as radiology, pathology, and, to a lesser extent, dermatology. A patient visiting a specialist expects concentration on the specialty area of interest, given a general reassurance that the patient has a primary care physician who has performed a comprehensive physical examination.

Reimbursement. Telemedicine reimbursement for health professionals located in areas where there are shortages of clinicians began in 1999 via Medicare. Coverage was expanded via the Telehealth Improvement and Modernization Act of 2002, for which regulations appeared in 2002. Some changes included the requirement that "the physician or provider" at a distant site be paid the same (parity) as if the services were provided face to face, the elimination of fee-sharing requirements, the expansion of eligible locations to any county not part of a standard metropolitan statistical area, and the expansion of the number of Current Procedural Terminology (CPT) codes that are eligible for Medicare reimbursement. Medicare reimbursement has broadened to cover a wider array of services (defined at http://www.cms.hhs.gov/manuals): Face-to-face communication via live videoconferencing, non–face-to-face communication through live video or

store-and-forward telecommunication services, and home telehealth services. Requests for Medicare reimbursement have, however, been a small fraction of the amounts budgeted. Although optimistic predictions were for rapid growth, total reimbursements from Medicare totaled only $20,000 in the first 2 years and reached only $5.2 million by 2011. Medicaid reimbursement for telemedicine is an option in some form by 48 states. The American Telemedicine Association (http://www.americantelemed.org/) offers updates on public policy, licensure, reimbursement, and research findings (http://www.americantelemed.org/policy/state-policy-resource-center#.VUt9LpN5K1c). Parity of coverage for telemedicine services is required in 16 states, with some strides, some stagnation, and some reverses. Four states prohibit use of cell phone video for telemedicine. Overall ratings of the "telemedicine climate" combine grades for the many variables that can be regulated, such as qualifies providers, services covered, mileage requirements between originating and distant sites, etc. In early 2015, the Texas and Maryland legislatures placed additional restrictions on telemedicine. Both the Association of Telehealth Service Providers and the Telemedicine Information Exchange no longer operate websites, which may indicate acceptance of the technology and its mainstreaming or some loss of momentum.

The rationale for telemedicine consultations comes down to the view that something is better than nothing. Medical specialists are unavailable in many areas and various locations, such as prisons. Second opinions are recognized as contributing to high-quality care and lowering medical costs. Professional organizations, especially specialty and subspecialty groups, support telemedicine as an answer to charges of maldistribution that led to broadening the scope of practice of nonphysicians (although many are now eligible for Medicare reimbursement, and pending legislation would add to the list). Apart from reimbursement, a clinician can charge for time, however spent. Increased health insurance coverage through the Affordable Care Act (ACA) is likely to lead to worsen physician shortages and increase use of telepsychiatry.

Government Support. Government support has been evident nationally through the establishment of the Office of Telehealth in the Department of HHS and in various laws on rural health care. The WHO regularly promotes telemedicine through publications and statements of regional priorities. Telehealth has become a valuable component in recovery from disasters.

Licensure Issues. Issues of licensure do not arise when the physician and patient are in the same state. When the patient and the physician are in different states, the physician must have a license for the state in which the

patient is located. The Federation of State Medical Boards developed a model licensing act in 1996 to ensure that a physician who is fully licensed in one state could obtain a special license for the states in which patients to be seen in telemedicine encounters would be located. In 2015, nine states allowed a limited license to out-of-state physicians to practice telemedicine while one state withdrew such a license. Telemedicine licensing may be submerged into the uniform licensing requirements and streamlining license acquisition envisioned by the Interstate Medical Licensure Compact (www.licenseportability.org).

Malpractice. The increasing use of telemedicine, now reckoned as more than 10 million patients has been accompanied by malpractice suits. Teleradiology specialists are said to be the most frequently sued medical specialists, just as radiologists are otherwise. Law firms are advertising for telemedicine victims. "The New Medical Malpractice" book describes litigation opportunities. Malpractice coverage for telepsychiatry occasionally is included in one's policy or can be bought as separate policies or as additional coverage on existing policies. Any clinician engaging in telepsychiatry should ensure coverage in advance. Liability is reduced if there has been at least an initial face-to-face meeting and establishment of a defined doctor-patient relationship. A Colorado physician was sued after electronically prescribing an antidepressant for a California patient who later committed suicide. A California appeals court ruled that felony charges (i.e., practicing medicine without a license) could be pressed against a physician who had never set foot in the state but provided telemedical services to some of its residents, at least one of whom had an unfortunate outcome. Some Medicare fraud convictions have resulted from false telemedical service claims.

Cost Benefits/Effectiveness/Outcomes. Cost benefits have been confirmed in multiple studies. A 2007 study found that nationwide use of telemedicine could save $4.28 billion, actually a relatively modest saving compared to projected savings from the widespread use of other technologies and a 2006 national health expenditure of $2.1 trillion. Hilty's 2013 review found it as effective and greatly improving access to care. Costs have decreased as equipment costs have plummeted. Patient and clinician satisfaction has been reported as high, with some patients preferring telepsychiatry for its convenience and availability. But many prefer face-to-face, possibly because such encounters are more intense. In one study more oxytocin was produced in those meeting in the same room than those meeting online.

Internet

The Internet was established in 1969 at the University of Southern

California as a way of linking computers for national defense uses. It came into civilian use as a way of linking personal computers and electronic devices through larger computers that function as servers, switching and routing devices that choose very rapidly and automatically how to route bursts of data. The data transmitted are stored in the computers through which they pass and can be retrieved. Electronic mail was used for years among users of terminals of mainframe computers. Specialized offers to certain customers via e-mail began in 1973, and the first online service among users started in 1979. Use of e-mail for therapy was documented in the 1980s, and simulated patients were developed (e.g., "Eliza" and "Parry") to demonstrate typical psychopathology to those who signed on to interact with them.

As of May 2015, 98 percent of the American population have access to the Internet, and 80 percent are frequent users. Of the regular users, 10 percent use only mobile devices for their broadband and Internet access. The Internet is a road to content and interactions in these areas that affect psychiatrists and their patients: (1) information on disease, diagnosis, treatments, and therapists, (2) support and self-help groups (moderated or not), (3) advice, diagnosis, and counseling, in which the person being helped has not met the helper except over the Internet, (4) prescription drug sales, (5) source of gratification, such as sex, games, and gambling, (6) personal experimentation/validation, such as different identities or reinforcing beliefs.

Information on Disease, Diagnosis, Treatments, Therapists. The Internet has become a major source of data, with myriad websites offering pages of information on many topics, including health care and psychiatric treatments. Web searches yield additional pages that can be searched effectively. For example, almost 270 million web pages mention diabetes, a more than threefold increase in 5 years. Patients tend to search several sites for information, which also compensates for the frequent inaccuracies found in health information websites. Patients present with a mixture of information, misinformation, and questions. Visit times lengthened at first, but physicians have developed and use their own information sources, and more accurate and informative websites have emerged. The better informed patient is more cooperative and much more of a partner in treatment. However, Internet users can find almost anything they want, and almost any medication can be alleged to produce almost any effect, wanted or unwanted. Many practitioner-operated websites discuss philosophical and practical aspects of treatment and serve as gateways for further contact. Commercially operated websites have been found, frequently in violation of their own privacy policies, to track those who visit the site for follow-up marketing. Physicians and therapists have had to contend both with their ratings from various grading websites and with

offers to improve such rating. Routinely patients Google their caretakers and vice versa.

Support and Self-Help Groups. The Internet provides access to information and people, some of whom may be therapists, and others with and without problems (acknowledged or not). Support groups have evolved for most illnesses and are used by patients both for finding face-to-face meetings and Internet chat groups. Often these provide support and knowledge, with occasionally advice of varying degrees of helpfulness. Mental health organizations also provide patients with educational materials, referral services, and support groups centered on problems, if not diagnoses. These groups, such as those run by the National Alliance for the Mentally Ill (http://www.nami.org), involve families and significant others nationally and through many local chapters. Voluntary associations are not subject to regulation and are not considered to be medical practice. A large number of groups have evolved online along with the widespread use of e-mail, centered on such topics as breast cancer survival, spousal battering, and so on. Some groups are hybrid, meeting occasionally in person and keeping in touch frequently, sometimes daily, by e-mail. When a professional agrees to be "present" online (if the group is conducted in real time) or review and comment on e-mails, the group is called a "moderated support group." Sometimes professional moderation is achieved through a bulletin board format, with professional comments on messages posted by users or messages sent to the professional for posting. Many of these are sponsored by health websites or commercial firms that make products that members of the group can use. Several hundred mental health professionals moderate such groups, carefully wording terms and conditions of participation to reduce potential liability. Members can group themselves by diagnosis—a recent search yielded thousands of schizophrenia support groups (both moderated and unmoderated)—or by other criteria.

Self-help groups are unmoderated support groups that emphasize traditional self-help mechanisms such as following steps to wellness. These groups offer low-cost, usually available support. Members often agree to have instant messaging, so anyone online can be reached. E-mail technology lends itself to mailing lists because addressees can be added to a message with ease and without costs, and e-mail programs typically include a "reply to all" option that makes it easy to respond to everyone who received the original message at once. Variations in format allow user-friendly "chat rooms" in which an e-mail user can sign in and participate in a conversation with the others who have signed in for that particular session, which typically lasts minutes or hours, although asynchronous communication makes extension to days and weeks possible. Still other groups, such as Psychiatric Survivors, have a national organization (see

http://www.power2u.org/index.html). Many patient groups use their websites to oppose certain treatments, especially medication. Mind Freedom opposes involuntary treatment. All support groups emphasize recovery and empowerment. "Recovery" does not mean "cure" but rather being able to move on with one's life while dealing with whatever illness remains. Empowerment means taking charge of one's illness and treatment, exercising choice and the protection of civil rights. Anything that smacks of being told what or how to think is objectionable. Coercive treatment is anathema, and medication is at least suspect.

Online Advice, Diagnosis, and Therapy. It is estimated that there are about thousands of mental health professionals (mostly social workers, psychologists, and counselors) who are willing to pick up advice-seekers online and work with them only by e-mail. There are also many nonprofessionals or pseudoprofessionals who justify this practice on the grounds that those being helped would not otherwise come for therapy for various reasons, including geographic isolation, loss of privacy, and disability. Fortified by an infusion of venture capital, interest in online therapy peaked in 1997 to 2000, with several well-funded websites, numerous articles in the popular press, and presentations at scientific meetings. By 2001, some disillusion was evident, and commercial websites had largely disappeared. However, the number of practitioners has increased slightly, operating their own, inexpensive websites. There is scholarly activity in journals, and the International Society for Mental Health Online (http://www.ismho.org) was founded in 1997, with 400 members in 2015. However, there have been relatively few scientific meeting presentations since 2002, suggesting that online therapy has entered the mainstream and become as unremarkable as the telephone.

In October 2002, state medical licensing boards began to implement a consensus statement ("Model Guidelines for the Appropriate Use of the Internet in Medical Practice") adopted by the Federation of State Medical Boards earlier in the year. It said, in part:

> Treatment and consultation recommendations made in an online setting, including issuing a prescription via electronic means, will be held to the same standards of appropriate practice as those in traditional (face-to-face) settings. Treatment, including issuing a prescription, based solely on an online questionnaire or consultation does not constitute an acceptable standard of care.

Actions included Illinois closing down MyDoc.com, an online consultation service that offered assessment and prescriptions for "minor illnesses." Online pharmacies banded together to form the Council for Responsible Telemedicine to prevent the online filling of prescriptions for patients who had not been seen by the prescribing physician, and several pharmacies had to pay substantial fines. Thus, the use of the Internet has been redefined to emphasize that there should be no unlicensed practice of

medicine. The guidelines recommend the use of electronic media "to supplement and enhance" the physician–patient relationship. MyDoc's guidelines now clearly state that the online service "is not meant to substitute for establishing a regular physician–patient relationship with a physician in your community." This may have been insufficient for regulators, as MyDoc.com is now in "stealth mode." Online therapy sites caution that "if something goes wrong you likely have no legal redress."

Traditional malpractice carriers have dropped their objections to covering routine physician–patient communication by e-mail, which HIPAA requires be encrypted if PHI is included. Case law has led courts to assign increasing liability to sites functioning in ways ranging from passive through intermediate to business. Patricia Recupero urged psychiatrists operating passive websites (typically home pages with a biography, licensure, and others) to post notices discouraging e-mails because even a brief, casual, unreimbursed answer can be seen as establishing a doctor–patient relationship, citing Nicholas Perry's review of tightening licensing laws in response to rogue pharmacies. A consortium of national medical societies and malpractice carriers (the eRisk Working Group for Healthcare), with the participation of state medical licensing boards, developed guidelines to limit the liability risks of communicating via e-mail. The guidelines, revised in January 2007, are laden with cautions about privacy, security, authentication, and liability: "Online communications of any kind are best suited for patients previously seen and evaluated in an office setting." The guidelines urge physicians to obtain specific affirmative informed consent for specific services. Some suggested principles for providing services online are available at http://www.dr-bob.org/psi/suggestions.current.html. Guidelines are available from malpractice carriers, for example, http://www.thedoctors.com/KnowledgeCenter/PatientSafety/CON_ID_00 The Clinical Social Work Federation's opposition to unmet e-therapy is unchanged since 2001:

> This area is totally unregulated and potentially very dangerous for clients and therapists alike .... This new, very powerful medium blurs all the usual boundaries. Most organizations believe that it cannot or should not be evaluated by well-accepted professional standards. That just is not the case. The standards developed by the U.S. Department of Health and Human Services and Coordinated by the Office for the Advancement of Telehealth, are used by the federal government to assess federal policy on an ongoing basis .... We have yet to see the first law suits in this area, but we know they're coming. Our concern with establishing a position on the delivery of online therapy services is in absolute alignment with the mission of state licensing boards ... the protection of the consumer. The standards used to analyze the growing area of text-based counseling include principles related to confidentiality, informed consent, quality of treatment, competence of the therapist, and basic ethical and professional requirements. These standards cannot be ensured when the client and therapist know each other only from a text on a screen. Assessment is the first phase of psychotherapy, and frequently significant information about the client is based on nonverbal cues. Psychotherapy has at its heart a profoundly human connection, a connection that is, in itself, the major vehicle for change. Healing and restoration occur when the therapist and the client together find the bridge leading back, and forward at the

same time, to the true self. Alienation from others and the self will not be healed through a virtual connection in cyberspace, a connection that is fraught with risks and hazards for both clients and clinicians.

Prescription Drug Sales. Physicians increasingly use electronic prescribing to send patients to reputable pharmacies with valid prescriptions. Health insurers increasingly require patients to use favored online pharmacies, often avoiding contact with pharmacists. But people seek medications on their own. Some online pharmacies put prospective customers in touch with physicians who will prescribe only on the basis of an online questionnaire. The Federation of State Licensing Boards has barred the use of online questionnaires as the sole basis for prescribing medication. Two federal laws, the Controlled Substances Act and the Food, Drug and Cosmetic Act, have been used to prosecute pharmacies that don't bother with prescriptions. Vigorous national and international efforts have led to the reporting and shutting down of more than 200,000 websites. Despite these, the sale of medications over the Internet without valid prescriptions has been a persistent problem. Come-ons for a wide variety of substances remain a major portion of e-mail spam. In 2002, patients who previously had fluoxetine (Prozac) prescriptions received unwanted samples of a longer-acting preparation, with letters mentioning their previous treatment for depression. A 2011 survey of online pharmacies found "consumers can get anything they want on the Internet. As pharmacies increasingly move offshore to countries with lax enforcement, what comes in the mail may be poor quality or counterfeit. A report of the International Narcotics Control Board found that 34 illegal online pharmacies in the United States dispensed 98 million doses of hydrocodone in 2006. In a survey of 187 websites, 84 percent sold drugs without requiring a valid prescription, despite regular fines. The number of websites selling controlled prescription drugs increased 70 percent from 2006 to 2007 and currently is estimated to be in the thousands. A 2010 study found thousands of sites, of which 3,160 sites were chosen for study; only 4 were accredited as Verified Internet Pharmacy Practice Sites. The business is sufficiently lucrative to support daily marketing calls to many physicians and customers.

Gratification: Sex, Games, and Gambling—Internet Use/Abuse and "Addiction". Although "Internet Addiction" can be found on 28 million websites, the APA chose not to include it in DSM-5 in 2014, focusing on content rather than the technology used to convey it. A study found most apparent addicts were new users who decreased their use over the following year. There are some people whose fascination with the technology may be problematic, but they may meet criteria for obsessive compulsive or other disorders. Internet gambling is excluded, as it can be diagnosed as part of Gambling Disorder. Compulsive viewing of

pornography also is excluded as a variety of sexual dysfunction disorders can be diagnosed to include it. An exception was made for Internet Gaming Disorder, which was included in Section III for further study. A diagnosis requires that repetitive use of games on the Internet is seen as leading to "significant issues in functioning as shown by having five of these nine symptoms: (1) preoccupation or obsession with the games, (2) withdrawal symptoms when not playing, (3) increasing amount of time playing games on the Internet (tolerance), (4) failed attempts to stop, (5) loss of interest in other activities, (6) continued overuse even knowing the ill effects of continuing, (7) lying to others about Internet game use, (8) use of the games to relieve stress, anxiety or other problematic feelings—an escape, (9) loss or putting at risk an opportunity or relationship. Although the criteria mention no particular qualities of games, just the time spent thinking about and using games while doing other things, patients coming for treatment dwell on the high they experience entering a world in which they can show increasing skill and power, vanquishing foes in wonderful, different worlds. The games are an interactive virtual reality, likely to be enhanced by improved, affordable headsets. Studies use differing severity criteria, so incidence and prevalence figures are unreliable, but the most vulnerable age group is males ages 15 to 25. Incidence among Korean youth has been found as high as 50 percent.

Despite its noninclusion in DSM-5, "Internet addiction" has been accepted by public consensus. The term preferred here is "problematic Internet use." Although electronic media do not involve substances, they can be very stimulating, with sight, sound, touch, and, increasingly, smell, always available, nonjudgmental, consistent, and inexpensive.

The combination of anonymity, convenience, and escape (the ACE model) promotes the Internet as a focus of psychopathology. In 2015, it is estimated that people spend about half their waking time looking at a screen: smartphone or computer. "Internet Addiction" is mentioned on more than 26 million websites, with those at risk suffering from depression, bipolar disorder, anxiety, low self-esteem, or addiction to substances, at least previously. Online surveys find that 4 to 10 percent of users meet criteria for "Internet addiction," defined as having at least five of the following signs and symptoms: (1) preoccupation with the Internet, (2) increasing amount of time spent online, (3) failure to cut back use with concomitant restlessness, (4) moodiness or depression, (5) staying online longer than originally intended, (6) running the risk of losing a job, relationship, or other opportunity because of Internet use, and (7) lying to conceal the extent of Internet use and/or using the Internet to escape negative feelings. General population surveys show a prevalence of 0.7 to 3 percent, with higher rates when family members are queried. As Internet use varies widely by employment and other needs, diagnostic criteria proposed in 2010 include at least 6 hours of "nonessential" Internet use

weekly for at least 3 months, and five of the symptoms listed above for Internet Gaming disorder. Impairment was evident in increased divorce rates, vocational impairment, legal problems, and personal distress. Subgroups on internet use include (1) cybersex addiction (viewing pornography), (2) cyber-relational addiction (online relationships become more important than those in one's physical world), online gaming (gambling, stock trading), compulsive, debt-inducing shopping, and others, (3) information overload, (4) net compulsivity, and (5) computer (non-Internet) addiction (e.g., computer games). About 30 percent of those "addicted" reported using the Internet to escape negative feelings and because it was always available at low cost.

PERSONAL EXPERIMENTATION/VALIDATION, SUCH AS DIFFERENT IDENTITIES OR REINFORCING BELIEFS. Potential sexual partners have many venues in which they can meet and provide mutual gratification, either by using via e-mail, telephone, or webcams or by setting up face-to-face meetings. On the other hand, opportunities for choosing a false, misleading, or incomplete identity abound. Sometimes this is done maliciously for exploitation. Sexual predators use the Internet, often with disastrous results when face-to-face meetings turn out to cause harm. Couples have been known to meet on the Internet, fall in love, and arrange to marry, only to discover they had not verified one another's sex. The movie *You've Got Mail* showed the use of e-mail as a way of expressing kinder, nobler feelings than Tom Hanks' character could in real life as a cutthroat businessman. However, studies on e-mail show that feelings are more likely to be disinhibited, angry ("flaming"), and paranoid, phenomena thought to be associated with the ease of response and the lack of social cues that more data input would provide.

The Internet can offer not only anonymity, but also a choice of other identities. Second Life (http://www.secondlife.com) describes itself as "a 3-D virtual world entirely created by its Residents [who socialize, create, and chat using voice and text chat]. Since opening to the public in 2003, it has grown explosively and today is inhabited by millions of Residents from around the globe." Its world is filled with creations of the Residents, often elaborately and colorfully costumed, and often with special powers. Second Life creations can include cities or even countries. The Second Life marketplace uses Linden Dollars, which can be earned in Second Life or bought with, or exchanged for, real-world currency. Several other sites (i.e., Metaverse, Open Life) offer similar fare. All are developing additional features to make their offerings more attractive and enveloping. Participation in alternate realities offers rich material for considering identity choices and psychodynamics. Some people have reported having more success in Second Life than in their real lives, whereas some patients have entered therapy because of Second Life relationships gone awry.

Besides from such sites, it is easy to participate in e-mail and chat groups using a false identity, often for a specific purpose. Changes in technology have outstripped the legal system's ability to provide basic protections. Even where economic damages can be claimed—such as false information engendering stock price swings—prosecution has been scant.

The Internet has been used by violent extremists to find and enlist fighters for their causes. In 2015, this problem is best seen as youth flow from developed countries to fight for the Islamic State. Newspaper accounts describe this process in detail for individuals. In the United Kingdom more than 1,000 have made the trip, a number growing by about 100 every few months; half are estimated to have been killed. Recent lone-wolf killers in Canada and the United States found their radicalizing groups on the Internet. These represent developing an identity via the Internet with disastrous consequences in the real world.

Internet Victims. Deception can take a malignant turn as sexual predators deceive their victims with false identities only to exploit and harm them when they meet. These contacts are unregulated and difficult to detect except by monitoring and checking the computers used. There are weekly reports of minors having been lured into sometimes lethal situations by sexual predators. Occasionally there is a report of a couple that met to marry only to discover they had missed verifying crucial details, such as each other's sex.

Some people who make little use of the Internet nonetheless become victims and may enter treatment. The suicide of one teenager after reading untruths entered by a peer's malicious mother has inspired laws to criminalize such behavior. Internet identity theft is also rampant. An underreported and growing problem, medical identity theft is harder to detect and remedy, often requiring painstaking record correction, currently estimated as costing $26.4 billion with 1.5 million victims in 2013.

"ADDICTS"

Treatment for Internet Addicts. A subset of web pages offer a chance to evaluate one's Internet use as possibly pathological and offer both education and online counseling, with some urging face-to-face counseling as a way of becoming less involved with the Internet. A rough idea of the ratio of what is offered as possible sources of help online is the number of sites mentioning "Cybersex" (1,310,000) compared to those mentioning "cybersex addiction" (about 275,000, increased from 19,000 six years ago). There are many mentions and variants on "treatment for Internet Addiction" (7,780,000, up from 338,000 six years ago), often practitioners with some experience. Treatment is similar to other addictions, using motivational interviewing, cognitive behavior and other psychotherapies. Medication occasionally is used for associated anxiety and depression.

Twelve Step groups modeled on Alcoholics Anonymous offer face-to-face support.

Virtual Reality

The development of large-scale, increasingly inexpensive computing power enabled the storage of sounds and images in digital form. MP3 files and flash memory sticks have largely displaced CDs and DVDs, and new formats probably will have appeared by the time this in print. Construction of images that never existed in reality has become widespread and easier in computer games, movies, and on personal computers. Playback equipment in helmets provides for a sense of total immersion with 360-degree sound and a complete occupation of the visual fields. The first use of virtual reality for a psychiatric condition was in 1992 for fear of flying, followed by its use for agoraphobia, acrophobia, claustrophobia, social phobia, anxiety related to public speaking, and other situations involving anxiety. Virtual reality headsets currently cost about $100 with increasing sales volume and technical improvements. Virtual reality therapy is used for phobias, social phobias, social anxiety, interpersonal skills training, and other problems that involve the need for desensitization.

Medical Records

Medical recording accompanied the developments of writing and medical care in ancient civilizations. The earliest known medical record is the Edwin Smith Surgical Papyrus, dated to between 5,000 and 4,500 years ago, a treatise that includes 48 typical cases, mostly head injury, with histories ranging from head trauma to infection. Although format and content have greatly expanded over the years, the medical record remained a description of past events recorded on paper. Carole Siegel and Susan Fisher's unique study of the psychiatric record can be summarized as applying generally to the use of records. Practitioners with access to patients go to the patient, whereas those who cannot get to patients have to use the record. Many records are illegible, and finding data can be difficult.

Early efforts to convert to electronic formats were inspired by the ease of retrieval and legibility but suffered from overexpectation, high cost, nonstandardization of data elements, and other key factors that impeded easy communication among physicians. Despite a lack of political support, some institutions developed excellent records, but there was little spread. The present effort began in 2004 with federal initiatives: The Office of the National Coordinator for Health Information Technology and an HHS report laying out a 10-year plan for federal, state, and local leadership. It has resulted in multiple professional–public–private partnerships, which include the Health Information Technology Standards Panel to define elements and standards and the CCHIT, which has certified inpatient and

ambulatory medical record software from hundreds of vendors. EMRs are expected to improve patient care through reduction in medical errors and clinical decision support. They also are expected to reduce care costs by reducing overutilization and promoting continuity of care as clinicians communicate with each other individually or in groups through RHIOs and HIEs. For example, the Veterans Administration's excellent EHR makes it possible to retrieve a patient's record (with permission) anywhere, including all diagnoses, medications, and treatment. This has proven to be of great benefit after local records are destroyed, such as happened due to Hurricane Katrina. During the last half century other nations have greatly improved their health care systems, making more use of HIT, while the American situation has worsened. Development of HIT is the only issue that has universal political support. Physicians also are seeking tort reform to lower malpractice premiums; some carriers offer discounts for risk reduction. Some studies show that EMR use reduces risk because the record is completely legible and retrievable, often involving the use of software that prompts for completeness and consistency.

CCHIT ceased operations in November 2014, having been unable to develop a sustainable business model. Certification has become more widespread and flexible, including the possibility that an individual can develop an EHR and have it certified. The Office of the National Coordinator for Health information Technology lists EHR certifying agencies (http://www.healthit.gov/policy-researchers-implementers/authorized-testing-and-certifications-bodies). Physicians participating in Medicare and Medicaid programs can receive incentives for certifying EHR Meaningful Use or be penalized (through reduced reimbursements) for nonuse. The Center for Medicare/Medicaid Services was setting standards for Stage 3 participation as such incentives were legislated to be included in other federal efforts to shape medical practice (Medicare Access and CHIP Reauthorization Act of 2015).

Meanwhile the regulations carry a substantial documentation burden, both for an overall plan for protecting PHI and for documenting informed consent, with additional requirements that each individual practice document how it will protect PHI in electronic media. The more secondary sources of PHI exist, the more difficult it is to determine where confidentiality may have been lost. Such secondary sources can be useful to individuals seeking data on a particular patient. Such people may be criminals, unethical private investigators, and others. They proceed by hacking into databases and retrieving PHI that might bear on competence, custody, divorce, and other adversarial situations. They, rather than organizations, may be interested in the content of specific e-mails. In the past, such individuals have broken into offices to get individual patient records. Computer break-ins can be easier, depending on the relative skills of the hacker versus the burglar, and may leave no trace. Although current

guidelines for the transmission and storage of PHI recommend password protection and encryption, it would be unwise to expect that such safeguards would prevail against a specific, determined, individual attack. Penalties for criminal acts related to medical records have existed for years with variable enforcement. The risk of violation of a competently protected PMR probably is less than that of a paper record, whereas a mass of EHRs may be subject to attack by a group seeking money rather than specific information on one individual.

Over the years complaints of privacy violations, data breaches and lax enforcement have resulted in tighter rules and regulations and increased fines and penalties. Breaches involving more than 500 individuals must be reported to HHS, which must then report actions taken in response. Under HIPAA, the Health Information Technology for Economic and Clinical Health (HITECH) of 2009, and the ACA of 2010, voluntary associations of covered entities were set up and variably funded: Regional Health Information Organizations (RHIOs) and Health Information Exchanges (HIEs). RHIOs and HIEs are viewed as the major hope for increasing quality and safety and decreasing overutilization and fraud. Participation is purely voluntary and requires sharing each patient consent according to state and federal laws. Apart from RHIOs and HIEs, providers who know they are taking care of a particular patient can exchange information according to traditional consent methods (as a covered entity if transmitted electronically).

Despite almost weekly reports of confidentiality lapses involving thousands of people, there have only been a few hundred criminal complaints and just six convictions for privacy violations under the Health Insurance Portability and Accountability Act (HIPAA) from the first in 2004 through April 2015. Twenty covered entities have been fined (the highest was $3.3 million) and two lawsuits settled (see http://www.hipaasecurityandprivacy.com/p/hipaa-convictionsfinessettlements.html). HHS is required to report breaches annually to Congress, which show increases in the numbers of people whose data were breached: 2009 (2.1 million); 2010 (5.4 million); 2011 (11.4 million); 2012 (22.5 million) (see http://www.hhs.gov/ocr/privacy/hipaa/administrative/breachnotificationr 2012.pdf). As of March 31, 2015 HHS had received 112,785 privacy complaints and resolved 94 percent, with 23,476 requiring changes in privacy practices or corrective actions (http://www.hhs.gov/ocr/privacy/hipaa/enforcement/highlights/index.htn Numbers for subsequent years are likely to be much higher. The 2015 breach of Anthem health involved the records of 80 million people. Despite warnings since 2012, U.S. Government data bases with extensive medical and psychiatric data on 24 million people were hacked in 2015 as complacency and sloppiness prevailed.

THEORETICAL ISSUES

These relate to (1) the use of electronic media per se, (2) the selection of data to be recorded, (3) acceptable and reimbursable medical practice, (4) identity and the Internet, (5) treatment of pathological Internet use, (6) uses of HIT, (7) privacy and security, (8) development of an interconnected information society, (9) HIT and health care change, and (10) possible therapeutic uses of games.

Electronic Media

Microprocessors are the only major commodity that continues to do more and more for less and less cost. In 2015 Moore's law, which dictates that the number of microprocessors that can be fit on a chip will double every 2 years had held true for 50 years. Individuals are surrounded by devices using microprocessors that can be linked by varied communication modes that have become faster and cheaper while carrying more data. Sensors in everyday devices can be hacked. In 2014, a refrigerator was found to be sending fundamentalist Christian messages to its owners. People live in a world of interconnected smart things. Many of these devices have multiple uses, including Internet access, and are simultaneously converging in multiple uses while being differentiated according to user preference. Technology is evolving toward constant use of mobile devices. Individuals often are unaware of using microprocessors and the Internet as auxiliary thinking devices, but the differences in memory retrieval from one's brain versus the Internet are immense. Calculators and spreadsheets outstrip merely human efforts. For communication, people vary widely in the amount and frequency of the use of media, but there is a general increase as communication becomes cheaper and faster. New communication formats have arisen, such as instant messaging and blogs.

Apps have developed rapidly for use on mobile device (mHealth), more than 40,000 are available in 2015. Some are perceived as medical devices and thus subject to FDA approval, for example, blood pressure monitoring. Many have been developed by not-for-profits, such as the Veterans Administration and the National Alliance for the Mentally Ill. They offer self-testing for diagnosis, treatment progress, supportive contact, reminders, and do many other jobs for the user, who is thus more autonomous and empowered. By late 2015, problems were evident as one app on suicide was found to provoke anxiety and to be misleading, so it was withdrawn. The British National Health Service closed down its app rating website and withdrew all apps it previously had offered, citing privacy, efficacy, and security problems. The APA Task Force on apps was discouraged as it examined depression and smoking apps from the iTunes store, suggesting no apps for patients be endorsed until peer review was more reassuring.

Selection of Data to Be Captured

Any electronic storage of data involves not capturing other data, and no use of electronic media is ever the same as being in the same room with another person. Some data will be lost, even with the capacity to magnify images and move cameras around. Who controls the camera makes for crucial differences. Whereas some data reduction is desirable, as one's attention inevitably focuses on one thing or another, some choices may miss the key points. E-therapy captures only words without tone of voice, inflections, and the social cues that voices carry. Telephones do not capture facial expressions. Telepsychiatry may focus on the smiling face of an apparently open, cooperative person while missing the squared-off shoulders and rigid posture of controlled hostility. Use of multiple cameras and observers may increase data input, but increased equipment expenses and observer time still do not close the gap between being there and dealing with the reality of the situation. No one claims that all data can be captured. The decision to use electronic media depends on whether the advantages of such use outweigh the disadvantages of data reduction.

Adding Images and Other Data to E-mail. The current bright line between e-mail alone as not acceptable medical practice for consultation and evaluation and telemedicine as licensable and reimbursable probably will be questioned as images, including video clips, are added to e-mails. The increasing use of webcams has resulted in images, voices, and other data being transmissible. Equipment is available for home monitoring of physiological measures. Webcams also mean that telemedicine has expanded from the controlled environment of the telemedicine studio at the referring site to the patient's desktop computer, although it is unable to match the studio's production values.

Acceptable and Reimbursable Medical Practice

Acceptability is determined mostly by professional associations on the basis of evidence. Reimbursability in the United States is increasingly determined by for-profit insurers on the basis of cost savings and may lag until these are recognized by all. Thus this standard will be less useful. EHRs may progress quickly to being acceptable, reimbursable, and required in some settings, but so far many physicians see them as user-unfriendly, overly reliant on templates and documentation requirements for billing and reimbursement. Psychiatrist interest in using EMRs varies according to practice setting, volume of patient visits, and patient and psychiatrist concerns about privacy. In 2014, the American Medical Association called for EHR "overhaul," noting that 85 percent of physicians use EHRs, only 24 percent reporting increased efficiency, while 47 percent said they detracted from patient care and 43 percent were slowed down,

with less time for patients. Data requirements (set by nonclinicians) keep physicians glued to their keyboards and have engendered the rise of the medical scribe industry.

Identity and the Internet

Increasingly, people prefer to work with computers and tend to disclose more facts to impersonal electronic media than they do face to face with live interviewers. Detailed personal information often is made available to everyone on social websites such as Facebook or MySpace by individuals who are shy and reserved in person. Several dozen studies show that patients give more detailed and accurate substance abuse histories in computer interviews than to live interviewers, although feelings about the substances used are not conveyed well. This has led to the widespread practice of having patients fill out forms and screening instruments before they see a physician in person or while developing PHRs. The use of electronic media for communication avoids having to deal with reactions from real people one is facing.

Virtual reality practitioners note that each person brings his or her background into a virtual reality experience. They describe a "virtual presence," which increases as patients spend more time in the virtual environment. Doing so involves giving up some of the sense of presence in the real world. Although initially virtual reality therapy (VRT) was patterned after systematic desensitization as practiced without electronic assistance, some of its practitioners now speculate that it is in fact oriented more toward neurophysiological information processing theory and the accelerated integrative information processing paradigm presented by Frances Shapiro. Patients in VRT typically would not communicate back to their therapists in the physical world, but instead seemed to be reliving their previous anxiety-provoking experiences. Other therapists point to the flight simulator analogy as one in which the evaluation instrument is also the means of remedying the skill deficits that it elicits.

As political polarization and radicalization have become societal concerns while Internet use has become widespread, one may wonder if there is an association. Some persons seek validation rather than information, reinforcement of their views rather than new opinions or data. The Internet has no self-correction and often is used as a mirror rather than a window, a versatile tool indeed.

Treatment of Pathological Internet Use

Online therapy or anything else online cannot deal adequately with an addiction to being online. Remedying flight from the real world for whatever reason must involve reinvolvement in the real world, including the experience of treatment, as described as follows.

Breaks from the Internet are recommended as a public health mechanism, including checking one's e-mail only a few times/day.

Uses of HIT. Physicians will have to adjust to having dynamic medical records that tap automatically into various scientific databases. Clinical decision support is available in real time as decisions are made, not in the form of excessive warnings that guidelines are being violated, but in a positive, efficient, and supportive way. Many time-consuming transactions are performed automatically. Safety and quality can be achieved with EMRs, but cost savings depend on EHRs. Smart cards, presently used in France and Taiwan, offer a technology that transmits data as a service is given, records the service, and updates the patient EHR. Yet these systems do without RHIOs and HIEs. The record is accessible only when a patient is being seen and has put the smart card in the physician's card reader. Public health and public policy will be much better informed by having almost instantaneous aggregate patient-deidentified data on patient care. In 2015, big data analyses yielded some results, such as pinpointing genetic susceptibility variants in bipolar disorder. In October 2015, New York State announced data on 30 million (medical records ranging from a few lines to more than 300 pages) would be available through RHIOs to clinicians who needed it for patient care.

Privacy and Security. HIT proponents argue that some sacrifices must be made for the common good of informing policy, public health, and research, whereas privacy advocates argue that patients should give affirmative consent to sharing every piece of personal data. It is not clear how much these opposing views will conflict in specific cases or how conflicts will be resolved. The rejection of British government proposals by overwhelming majorities of primary care physicians and their patients suggest that HIT proponents can become overly enthusiastic and lose a sense of public preferences. Only competent auditing can help to reduce the fraud that consumes about 10 to 15 percent of American health care costs. It is not clear whether this will be accomplished by clinician organizations or left to the insurers with whom there already is a power struggle.

"Privacy is over," runs one argument in the debate, while others argue it can be maintained. The reality is that different data matter more or less to different people. Some will post their entire genomes, but not how to access their bank accounts.

Development of an Interconnected Information Society. The use of extracorporeal processing has created a new culture with alternate realities and identities, affecting communication, the problems that patients bring to psychiatrists, and what psychiatrists can do for them.

CULTURAL CONTEXT. A generation that has grown up with microprocessors and the Internet has incorporated them into their way of life. They use all manner of devices with ease and comfort and are driven to take up new ones. They are more concerned with communication than privacy and spend a lot of time in touch with each other. Indicators of time spent sitting indoors show that Americans are spending less time outdoors and are facing an epidemic of obesity.

The prevailing cultural model of for-profit medical care has eroded privacy in favor of marketing. Every patient can be viewed as a potential customer. Others, outside the for-profit world offer a top-down consensus that many problems in U.S. health care can be eased through the use of electronic health records, especially electronic prescribing programs that check for safety against other medications and patient data. Although such efforts have sputtered in the past, the current drive benefits from the continuing price decline and quality improvement of electronics generally, as well as better planning, funding, and standard setting.

HIT and Health Care Change. It is not clear whether HIT can bring the transparency and public policy consensus required for health care reform. Before agreeing on HIT, all sides agreed that costs were excessive, primary care doctors were underpaid, and all citizens should be able to have medical bills covered. Yet these three and other indicators of health care system viability have worsened despite intense fact finding and discussion. The United States has allowed health insurance profiteering, which is regarded elsewhere as the result of not having a system, but rather a market. Pay-for-performance programs have been initiated, offering small bonuses compared to those in other countries. It is important not to make promises for HIT that will lead to disappointment. However, the extreme portability of electronic records and the global reach of telemedicine may result in outsourcing competition that will raise a host of issues.

Possible Therapeutic Uses of Games. While Internet Gaming can be so intense and time-consuming as to be on the threshold of being a treatable diagnosis, some studies have shown that playing computer games is therapeutic. Test instruments, such as the Wisconsin Card Sorting test have been turned into games, with the idea that repeated use will improve performance—an example of the assessment test becoming the rehabilitation tool. Results have been mixed and somewhat disappointing. However, games can be linked to neural circuitry to show improvement, as was shown for alcoholism in 2015 by Weirs et al. This study shows an interaction of internal, biological processing circuits synchronizing with external, silicon-based circuitry.

Techniques

The process of becoming adept in the use of electronic devices is eased by awareness of that some cognitive and communication functions can be delegated to microprocessors and the Internet. Formerly, information was hard to obtain; now the problem is sifting through too much of it. Former Czech President Václav Havel posed the new problem as knowing more but being less wise. Whereas telepsychiatry in itself does not provide information overload, the electronic records to be reviewed before a consultation may do so. Words and images are captured and available for both the patient and the therapist to examine. Electronic media can be used to provide feedback and communication, regardless of distance and time, or to present a different reality.

Clinicians should always be aware that people using electronic media make choices in the use of words and images, often involving an individual sense of self, effectiveness, and identity that may or may not be clinically significant. Clinicians should also choose consciously what data not to capture. For example, HIPAA protects psychotherapy notes, but e-mails and telepsychiatry sessions may mix material entitled to such protection with other data for which less sequestration is possible. Use of electronic media is not a new form of therapy, but a different way of carrying out a variety of therapies, some of which lend themselves better than others to the use of electronic media. However, there are some general considerations, which are described as follows.

E-mail

Even with encryption, e-mail offers ease of composition by typing or dictating onto a screen, editing on screen, speed of delivery, ease of retrieval, legibility, ease of response, and storage. Retrieval can be so quick as to allow a conversation in real time or at one's leisure, which means that the recipient can be in a state of mind favorable for communication. Initial resistance to e-mail has been replaced by positive feelings of both patients and physicians. Unlike unrecorded telephone calls, however, all e-mail is retrievable and potentially discoverable.

Recording Techniques

All use of electronic media is a production of some sort, requiring attention to whatever is chosen for recording. Issues in microphone placement, lighting, and other production values often are neglected and contribute to less data being captured.

Storage

Data storage is a function of costs and convenience. E-mail tends to be stored wherever it goes, requiring active steps to limit storage. E-mail messages between doctor and patient are now regarded as part of the

patient's record and should be stored with it. Generally, the lower the cost of storing e-mail, the greater is the convenience, but with less security. Storage of telemedicine sessions is typically local, using videotape or DVDs. Backup of EMRs is essential.

Transmission

Similar considerations apply to the transmission of data, in that low-cost applications are less secure but more convenient. Wireless transmission is particularly insecure. Data in digital format get stored in ways that promote both encryption and retrieval. Typically, for example, digital data are transmitted through Internet servers in bursts or batches that are stored as they pass through the servers. Analog data are not stored as they are transmitted, typically on telephone lines, sometimes by satellite. Analog data can be recorded only in analog terms, unsusceptible to hacking. Organizations typically are not interested in analog data. With the 2009 move to having all television signals in digital format, there may have to be encryption for all television signals.

Virtual Reality

A person's experience of a situation in a virtual environment tends to be the same as in a similar real-world situation. It is possible to add and subtract elements of the virtual environment to determine what specifically causes the anxiety a patient feels. Generally the patient is sitting rather than standing, even if standing in virtual reality, and begins with brief sessions, a narrow field of view, and the ability to see his or her physical body partially. As tolerance develops, longer sessions with fuller immersion can be used. Use of 360-degree cameras has led to the capture of social situations in which, say, a patient with social phobia can be surrounded by people at a party, one of whom may then speak to him or her. The program can branch in one of several directions, depending on the patient's reply.

CLINICAL ISSUES

Indications

These include symptoms, diagnoses, and defense mechanisms. Although e-mail and telepsychiatry share some of the same indications, the mode of initial interaction may vary. With e-mail the patient and therapist generally have met on their own, either face to face or online. Telepsychiatry presently consists mostly of consultations, with referrals having been arranged by someone where the patient is actually located. Availability of consultant expertise usually is the key indication. Home health care monitoring can be very useful for situations in which a home visit would have been used. Both may be useful as a way of overcoming limitations of access, distance, and disability. Both are useful as an adjunct to therapy in

situations in which the patient and therapist meet face to face but want more frequent contact. Its format lends itself, for example, to the sort of written daily homework expected in cognitive-behavioral therapy (CBT). It may be useful as a way of getting patients into treatment that they otherwise would not seek but that should then progress to more definite involvement via telepsychiatry or face-to-face contact. As webcams have come into wider personal use, some patients have found telepsychiatry to be more convenient and of lower cost than travel to a therapist's office.

Virtual Reality. Experiments showing efficacy in various forms of situational anxiety are convincing. Most phobias and fear-provoking situations can be simulated in VRT. Investigations of VRT's usefulness in OCD, ADD, and PTSD are underway.

Limitations

E-mail. The principal limitations of e-mail are related to the data not transmitted or available. There are potential problems in identification of both the patient and the therapist. Investigations of those offering online services found that many did not give their professional credentials or were incomplete (e.g., "M.S." in a context that implies that the degree is in counseling when it may not be). Similarly, potential clients often give themselves fictional names. Although this is seen as a way of preserving an anonymity without which help might not be sought, therapy cannot progress far without more disclosure. Not all deception is deliberate. As noted earlier, patients may feel that they have a different identity on the Internet, which may lead to their portraying problems consistent with that identity rather than their identity in real life. Although controlled studies show that patients divulge facts more rapidly and completely to computers, most of the studies were done with substance abusers before the Internet and rise of confidentiality concerns. Computer-based assessments have not replaced live interviews because communication of facts still does not convey how a person thinks or feels about them. On the other hand, the disinhibition of emotion noted in unstructured e-mails is both an indication for therapy where defenses must be overcome and a limitation if the feeling is impulsive or marginally related to the issues for which therapy is sought. These same limitations apply to support and self-help groups. Therapist limitations must also be noted. Extensive use of e-mail is not recommended for therapists who have not mastered expressive writing, typing, use of computers and the Internet, encryption, security of electronic data, and how to deal with computer glitches. Therapists must be aware of how they may seem online, be able to correct miscommunications in both directions, and be able to shift modalities (e.g., from online to face-to-face communication) when indicated.

Telepsychiatry. Limitations are related to lost data, equipment expense, and discomfort with the medium. Personnel savings are reduced by how much of a local team is needed.

Virtual Reality. Some patients complete courses of education or desensitization and find limited generalization (e.g., just as anxious in real life as before treatment).

Complications

A person may get very good at e-therapy and working on the Internet without this result generalizing to the physical world. So-called "cyborg theory" has been invoked to describe the mechanisms whereby the animal–man boundary blurs, along with the man–computer boundary.

E-mail. Reports have described predators taking advantage of victims they met on the Internet. Online therapists report having to refer patients for face-to-face therapy, but they often view this as a desirable progression rather than a complication. Self-help groups can be taken over by chat room addicts. Patients and therapists may present false identities. Predators can pick up vulnerable people and develop private correspondences with them that may lead to disaster.

Telepsychiatry. There have been no reports of complications from telepsychiatry. It is possible that complications would develop around data not conveyed in the consultation, but these usually are averted by the local team.

Virtual Reality. No complications have been reported. It is possible that a patient would continue to have anxiety after treatment, especially if he or she is subject to flashbacks.

Contraindications

Both the use of e-mail and that of telepsychiatry are less desirable in situations in which the patient and psychiatrist can meet face to face, although the patient may continue the use of a support or self-help group otherwise. Psychiatrists should always be sensitive to the possibility that VRT is contraindicated for patients with serious medical problems and patients taking medications with major physiological or psychological effects. Flickering images, especially at 8 to 12 Hz, may produce symptoms that range from headache to epileptic seizures, and so no frame replacement speeds that approach these frequencies should be used.

Case Examples

E-mail

An intelligent, well-educated banker found that reporting daily intake of alcohol was more effective than interpretations about setting up the therapist as a parental figure. (2) A depressed person searched for "depression + help" on an Internet search engine and found and took a self-assessment scale via e-mail. The person concluded that face-to-face help was needed and later benefited from treatment. (3) An apparently anxious person seeking help was actually a person using the Internet's anonymity, convenience, and possibilities for escape by trying out a new identity. The involvement with online counseling resulted in more and more time online and confusion about and diminution of real-world relationships.

Telepsychiatry

A patient was referred for a second opinion. The consultant elicited new symptoms, substantiated a different diagnosis, and recommended treatment that proved to be effective. (2) A well-known psychiatrist performed what he considered to be an excellent interview but missed the fact that the patient was increasingly resentful about the picture-in-picture showing her. She motioned to have the local crew on-site cover it, which they did, with a sheet of paper that also covered part of the consultant's image. None of this was conveyed to the consultant, who was now somewhat walled off. The patient continued to smile while the consultant suggested that her previous head trauma could benefit from a neurological work-up and reduce her need for 25 mg of diazepam (Valium) a day, but observers in the room noted that her shoulders squared and her posture conveyed belligerence. The consultant's first clue that the interview had gone badly was at the end, when patient responded to his asking if she had any questions. She glared and said, "Yeah, what do I do the next time I'm feeling suicidal and don't want to call anyone?"

Virtual Reality

A patient with social phobia was treated in her psychiatrist's office with a virtual reality program that provided a series of social stimuli, adding and subtracting as she coped successfully with each. She realized she had a particular problem with older men, and developed skills in dealing with them. (2) A similar patient was never able to accept that virtual reality had anything to do with his problems, and instead benefited from medication, which was decreased as he developed new skills in handling real-life challenges.

Goals of Treatment

E-mail. The goal of any exchange of e-mails should be to develop, supplement, and enhance a real-world therapeutic relationship. If such is unavailable, e-mail should be used for developing a relationship with appropriate safeguards. E-mail can be included as part of treatments that involve correcting cognitive distortions, reporting events, or providing frequent support.

Telepsychiatry. The goal is to provide services that otherwise are not available.

Virtual Reality. The goal of VRT is desensitization by replicating anxiety-provoking situations and training for unfamiliar, feared challenges.

ETHICAL ISSUES

Ethical issues abound. The novelty of technological advances, regulations, and practice ensure that years will pass before standards are generally accepted and followed. The path will be marked by malpractice and anecdotes from which there will be attempts to draw lessons.

The ethical issues will also depend on theoretical considerations and practicality of maintaining confidentiality. It is likely that the current trend of improving quality and dropping price will lead to more widespread use of electronic media.

A therapist referring patients to an online support group will face an added ethical burden to maintain contact with the group and to have some sense of what is occurring in it, given that such groups can change rapidly. Ethical issues around identity and disclosure will continue as patients will want to know what is stored about them and where and how such information can be accessed.

In all cases, the therapist is required to explain the procedure in detail, how data are stored, safeguarded, and transmitted, and the limitations and potential complications of the treatment. Informed consent must be obtained. It is wise to record the consent both in writing and electronically.

RESEARCH AND EVALUATION

There is little funding for research on electronic media. Government and professional associations have decided to favor use, as described earlier. Equipment manufacturers are preoccupied with producing new models for the general population and remaining price competitive. There is no special equipment for anything described here. EMR vendors sponsor exhibits with little research, but government and the American Medical Records Institute provide some funding.

Problematic Internet Use

In a telephone survey of 2,513 adults, 69 percent were regular Internet users, 6 percent felt that their relationships suffered as a result of excessive Internet use, and 8 percent used the Internet as a way to escape problems or relieve negative mood. Other potential markers for problematic use included preoccupation with the Internet when offline, concealing nonessential use, failed attempts to cut back use, and frequently using the Internet longer than intended. Interviews of 216 college students in Taiwan found 87 students to be Internet addicts, with associated MDD, ADHD, dysthymic disorder, and social phobia, with most addicts being male students with MDD or ADHD. Only seven men and eight women had no

comorbidity.

Telepsychiatry

Twenty-eight inmates of New York prisons receiving services from their on-site treatment team and from visiting consultants in person or via telepsychiatry preferred the former for most services. None preferred telepsychiatry to face-to-face communication, but telepsychiatry was accepted and tolerated, especially for contact with experts who would not come into the prison. Diagnostic reliability using the Structured Clinical Interview for DSM Disorders (SCID) in 53 Native American veterans was found to be comparable in person and with telepsychiatry except for past-year substance use and abuse/dependence combined. Telepsychiatric consultations and brief follow-up were rated as producing outcomes equivalent to face-to-face communication in 495 Canadian patients referred by family physicians, with a 10 percent cost saving. A comparable study of 119 U.S. veterans found no differences in outcome, patient satisfaction and adherence, and cost. Other studies have generated controversies about costs in rural areas. Video feedback has been shown to be helpful in several studies in which it was used in the context of an ongoing face-to-face relationship.

There are no comparisons between online therapy and face-to-face communication because it is generally accepted that the former excludes body language, facial expressions, eye movement observation, tone of voice, and others. Studies confirm the commonsense view that evaluative feedback (not in therapy) is given more effectively in face-to-face encounters than by computer. Anecdotal evidence supports the widespread view that online availability and e-mail contact can be helpful between meetings in therapeutic sessions.

Published reports on telemedicine generally are brief, with an emphasis on how to do it and minimization of any differences between the telemedicine experience and face-to-face communication. Evaluative data describe satisfied patients and savings on travel time, costs, and so on. Although it is understood that some data are excluded, there has been no systematic investigation of how missing data can influence outcome. Despite more than five decades of use, a Cochrane review in 2010 (http://www.cochrane.org/CD002098/EPOC_telemedicine-versus-face-to-face-patient-care-effects-on-professional-practice-and-health-care-outcomes) comparing telemedicine with face-to-face communication found only variable and inconclusive results. Subsequent reviews have shown benefits in specific areas such as health monitoring of high-risk patients.

As with most forms of therapy and new treatments, there have been phases of zealotry, exaggeration of possible benefits, some disillusionment, and loss of interest. Annual meeting programs of the APA, American Psychological Association, and World Psychiatric Association show

decreasing reports on the use of electronic media from a peak in 1999. However, the basic decision to use more telepsychiatry has been adopted by those who pay for public and private care. In addition to reimbursement through Medicare and Medicaid, the Office for the Advancement of Telehealth of the Health Resources and Services Administration (part of the Department of HHS), among many funding sources, provides grants and contracts and publishes a directory of supported projects. The WHO promotes telemedicine annually in its regional advisory committee meetings, especially for developing countries. The growth in teleconferencing reflects the evaluation of those who pay for it. Rather than comparisons with face-to-face meetings, telepsychiatry should be compared to no or inexpert service. The *Journal of Telemedicine and Telecare* (http://jtt.sagepub.com) was started in 1995 and is published monthly. It mostly describes the extension of telemedicine into remote settings and use in various diagnoses, such as dementia. It includes papers on information infrastructure and the situation in developing countries but no controlled studies.

Virtual Reality Therapy. VRT has been evaluated positively versus standard desensitization for phobias, especially fear of flying, using standard outcome measures. Wiederhold and Wiederhold summarized numerous positive studies. There are no studies demonstrating its superiority, although it is likely that clinicians who invest in the equipment and master the technique eventually will attract patients who have not improved with standard desensitization. A few studies dealing with other diagnoses such as schizophrenia, body image distortion, and eating disorders include new diagnostic methodologies and showed positive outcomes persisting for up to 1 year after treatment.

Electronic Medical/Health Records. Studies of EMRs and EHRs have described the use of databases to get accurate information about patients (e.g., arrests, medications, diagnoses), including refusals and concerns (among those agreeing) of patients to have their paper records entered into electronic media. In general medicine, physicians report a decline in productivity while adapting to EMRs, followed by practicing medicine differently with higher quality and safety and noticing a decline in paperwork and wasted communication time. Only 10 percent of physicians using electronic records abandon them to return to paper records. There are many reports of electronic records being harvested for previously uneconomic investigations (e.g., substance abuse-related mortality in 169,051 veterans over 7 years).

References

Aboujaoude E, Koran L, Gamel N, Large MD, Serpe R. Potential markers for problematic

Internet use: a telephone survey of 2513 adults. *CNS Spectrums*. 2006;11(10):750.

Brailer DJ. *The Decade of Health Information Technology: Delivering Consumer-centric and Information-Rich Health Care*. Washington, DC: U.S. Department of Health and Human Services; 2004.

Currel R, Urquhart C, Wainwright P, Lewis R. Telemedicine versus face to face patient care: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev*. 2010 (20 January): CD002098.

Darkins AW, Cary MA. *Telemedicine and Telehealth: Principles, Policies, Performance, and Pitfalls*. New York: Springer; 2000.

Druss BG, Xu J, Glick G, von Esenwein SA. Randomized trial of electronic personal health record for patients with serious mental illness. *American Journal of Psychiatry*. 2014;171(3):360–368.

Freeman D. Studying and treating schizophrenia using virtual reality: A new paradigm. *Schizophrenia Bull*. 2008;34(4):605.

Gustafson DH, McTavish FM, Ming YC, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism. *JAMA Psychiatry*. 2014;71(5):566–572.

Hilty DM, Ferrer DC, Burke Parish M, Johnston B, Callahan EJ, et al. The effectiveness of telemental health: a 2013 review. *Telemedicine and e-Health*. 2013;19(6):454–454.

*Hsiung RC. *E-Therapy: Case Studies, Guiding Principles, and the Clinical Potential of the Internet*. New York: WW Norton; 2002.

Katz SJ, Nissan N, Moyer C. Crossing the digital divide: evaluating online communication between patients and their providers. *Am J Managed Care*. 2004;10(9):593.

Ko CH, Yen JY, Chen CS, Chen CC, Yen CF. Psychiatric comorbidity of Internet addiction in college students: an interview study. *CNS Spectrums*. 2008;13(2):147.

Kripalani S, LeFevre F, Phillips CC, Williams MV, Basaviah P. Deficits in communication and information transfer between hospital-based and primary care physicians. *JAMA*. 2007;297(8):831.

Maheu M, Whitten P, Allen A. *E-Health, Telehealth, and Telemedicine: A Guide to Start-Up and Success*. San Francisco, CA: Jossey Bass; 2001.

O'Reilly J. *The New Medical Malpractice*. Chicago, IL: American Bar Association Press; 2014.

Orizio G, Meria A, Schulz PJ, Gelatti U. Quality of online pharmacies and websites selling prescription drugs: a systematic review. *Journal of Medical Internet research*. 2011;13(3):e74.

Perry NP. Prescriptions sans frontieres (or how I stopped worrying about Viagra on the Web but grew concerned about the future of health care policy). *Yale J Health Policy Law Ethics*. 2004;4:183.

Raimer BG, Stobo JD. Health care delivery in the Texas prison system. *JAMA*. 2004;292:485.

Recupero P. Law & psychiatry: legal concerns for psychiatrists who maintain web sites. *Psychiatric Services*. 2006;57(4):450.

Rosen CR, Kuhn E, Greenbaum MS, Drescher KD. Substance abuse–related mortality among middle-aged male VA psychiatric patients. *Psychiatric Services*. 2008;59(3):290.

Rothstein MA, Talbott MK. Compelled disclosure of health information: Protecting against the greatest potential threat to privacy. *JAMA*. 2006;295(24):2882.

Ruskin PF, Silver-Atlaian M, Kling MA, Reed SA, Bradham DD. Treatment outcomes in

depression: comparison of remote treatment through telepsychiatry to in-person treatment. *Am J Psychiatry*. 2004;161(8):1471.

Shore JH, Savin D, Orton H, Beals J, Manson SM. Diagnostic reliability of telepsychiatry in American Indian veterans. *Am J Psychiatry*. 2007;164(1):115.

*Shortliffe EH, Cimino JJ. *Biomedical Informatics: Computer Applications in Health Care and Biomedicine*. 3rd ed. New York: Springer; 2006.

Siegel C, Fisher S, eds. *Psychiatric Records in Mental Health Care*. New York: Brunner Mazel; 1981.

Singh SN, Wachter RM. Perspectives on medical outsourcing and telemedicine—rough edges in a flat world? *N Engl J Med*. 2008;358:1622.

Steinbrook R. Personally controlled online health data—the next big thing? *N Engl J Med*. 2008;358:163.

Taintor Z, ed. *Computers and the Psychiatric Patient*. Washington, DC: American Psychiatric Press; 1997 [also *Rev Psychiatry*. 1997;16(Section 6)].

Tao R, Huang X, Wang J, Zhang H, Zhang Y, Li M. Proposed diagnostic criteria for internet addiction. *Addiction*. 2010;105:556–564.

Tucker W, Olfson M, Simring S, Goodman W, Bienenfeld S. A pilot survey of inmate preferences for on-site, visiting consultant, and telemedicine psychiatric services. *CNS Spectrums*. 2006;11(10):783.

Wachter R. Expected and unanticipated consequences of the quality and health information technology revolutions. *JAMA*. 2006;295(23):2780.

*Weirs CE, Stelzel C, Gladwin TE, Park SQ, Pawelczack S, Gawron CK, et al. Effects of cognitive bias modification on neural alcohol cue reactivity in alcohol dependence. *American Journal of Psychiatry*. 2015;172(4):335–343.

Wiederhold BK, Wiederhold MD. *Virtual Reality Therapy for Anxiety Disorders: Advances in Evaluation and Treatment*. Washington, DC: American Psychological Association; 2004.

▲ 7.10 Cultural Competency in the Organization and Delivery of Care

CAROLE E. SIEGEL, PH.D., GARY HAUGLAND, M.A., AND HELENA B. HANSEN, M.D., PH.D.

INTRODUCTION

Cultural Competency (CC) is a notion borne from considerations of nondiscrimination that has been applied in various contexts including national and international affairs, education, the workplace, and health care. In all these venues, it speaks of productive ways to achieve equitable ends in situations where a mainstream or majority group—sometimes a power group—interacts with "other" cultural groups—usually minority or possibly disadvantaged groups—who have values and views and sometimes languages that are appreciably distinct from the "mainstream." The specific contextual issues and the positions taken can substantially differ from venue to venue. For example, while debates can ensue as to whether or not immigrant children in the United States should be educated in their native languages and in specially designed educational courses, there is no analogous debate in health care. In the United States, medicine is predominantly delivered in English in modalities of care well accepted by Western groups, but there is greater agreement that language and treatment accommodations are required for persons from cultural groups who have limited English proficiency (LEP) and who may require special treatment accommodations.

In this chapter, recent work is presented that moves the discussion of CC in health care—with a focus on mental health care—beyond a rehearsal of compliance issues and the rhetoric of political correctness to a description and justification of the CC activities that can be used to promote patient-centered care. While it is critical to apply these activities in all health areas, it is particularly crucial in mental health care where diagnoses are made, and many treatments delivered strictly based on verbal communication. Operationalizing the activities of CC has remained challenging, but the landscape of the 80s, when these notions were first introduced, has considerably changed based on the broad recognition that for CC to move forward, greater specificity was required on how to do this.

DEFINITIONS: CULTURAL GROUP, CULTURAL COMPETENCY

A cultural group has shared activities, ideas, and traditions, which are reinforced by its members. The group is a dynamic entity as the views and values of its members can change over time, and are influenced by where people live as well. Thus Chinese and Chinese American cultures in 2014 differ from each other, and each group differs from itself pre- and post-Chinese economic reforms. Federal policies for health care to date have

been framed to ensure nondiscrimination only for minority racial/ethnic groups, but CC is intended to apply to any cultural group that encounters limitations in service access or in participation in services. While clearly including underserved racial/ethnic groups, it also encompasses groups disadvantaged in health care such as lesbian, gay, bisexual, and transgendered communities, the born Deaf, and even rural communities each having distinct lifestyles. Anthropologists have recently questioned the value of culture in understanding a person's behaviors, as monolithic constructs fail to account for variations in expression and experiences within cultures. The notion of a cultural identity attempts to address this deficit by recognizing that persons belong to multiple cultural groups and the intersectionality of these groups is what shapes a world view. Cultural groups are defined by more than race, ethnicity, country of origin and go beyond to encompass socially defined groups for which membership impacts daily living such as groups defined by socioeconomic status, professional roles, gender, and stage of life roles. Thus a black, unmarried, inner city male will differ in world view from a black, father who lives in suburbia. CC training materials for clinicians take cognizance that persons' views of their cultural identity are fundamental to understanding how they makes sense of the world and hence how they respond to and interact with a health care provider.

Many definitions of CC in health care have been proposed, of which the most commonly cited in mental health care is that of Terry Cross. He defines CC as "a set of congruent behaviors, attitudes, and policies that come together in a system, agency, or among professionals and enables that system, agency, or those professionals to work effectively in cross cultural situations." This definition first appeared in a monograph on ways to work with severely and emotionally disturbed minority youth written for the Georgetown University Child Development Center (now the National Center for Cultural Competency), an early promoter of CC in the delivery of childrens' services. A less proscriptive definition defines Cultural Competence in an individual or organization as having the capacity to function effectively within the context of the cultural beliefs, behaviors, and needs presented by consumers and their communities. While these definitions are all encompassing, to better understand their implications, it is helpful to parse them into the distinct ways they apply to organizations and to caregivers.

When considering healthcare organizations, that broadly include structural entities such as departments of mental health, managed care companies, hospitals, agencies, and programs, CC is referential to the policies and procedures that should be endorsed and promoted throughout an organization to enable caregivers and other staff under the purview of the organization to work effectively and efficiently with different cultures and within multicultural situations to provide high-quality services. On the

clinical level, CC refers to the desirable personal attributes of clinicians encompassing their attitudes (affective domain), knowledge (cognitive domain), skills and actions (behavioral domain) that enable them to work effectively and efficiently across cultures and in multicultural situations to produce clinically and culturally appropriate service outcomes. This latter definition addresses clinician values, overt and covert biases as well as the knowledge and skills guiding the ways to act.

There has been notable clinical resistance to the concept of CC as many view that it is already addressed or implied in clinical competency, or feel it is already adequately addressed when a clinically competent clinician practices "patient-centered" care. The Institute of Medicine (IOM) defines clinical competency as the ability to use clinical skills and knowledge to identify each consumer's unique health state and diagnosis, individual risks and benefits of potential interventions, and personal values and expectations. In working with persons with severe mental illness, some curricula endorse core competencies to address recovery, rehabilitation, and empowerment. The argument goes, if these notions are fully operant, CC will be embedded in clinical competency. But for now, CC is not treated as a distinct construct in clinical training or adequately represented in its curriculum so that CC needs to be viewed as a distinct competency.

It is well documented that clinicians' misunderstandings of cultural sensibilities can lead to failed treatment alliances. Many psychiatric training programs around the country have instituted CC didactics and curriculum that explicate one or more of the attributes of CC identified by Derald Sue and others of attitudes, knowledge and skills. For example, the New York University psychiatric cultural competency curriculum has three modules, two of which address attitudes (and awareness): Perceiving the other and Perceiving Self. In the former residents learn how sociocultural influences impact clinical presentations and in the latter residents experience their own cultural backgrounds, and examine their influence on the doctor–patient relationship. Action skills are taught through a third practicum module where residents practice conceptualizing culture in the doctor–patient encounter using a culturally interpretive clinical approach. Other models have been proposed for guiding clinical competency and one such other is discussed later.

Helena Hansen and Jonathan Metzl have recently proposed a construct called structural competency. It addresses the societal influences on care that can cause disparities in treatment and care outcomes. These influences are reflected in embedded institutional policies and ways in which systems of care do business as well as in the social determinants of health. Its influences trickle down to the clinician. They argue that "clinical training must shift its gaze from an exclusive focus on the individual encounter to include the organization of institutions and policies, as well as of neighborhoods and cities, if clinicians are to impact stigma-related health

inequalities." Figure 7.10–1 displays the hierarchical levels of a health care system in which CC activities need to be introduced. Structural competency is overarching serving as a meta-level of CC. CC operating at all levels can make a clinician's job easier and more effective when working with members from cultural groups.

WHY CC IS IMPORTANT: DIVERSITY, DISPARITIES

The Diversity of the US Population Is Growing

According to the 2010 Census, 27.6 percent of the US population is nonwhite. 16.3 percent report themselves as Hispanic/Latino, 12.6 percent as Black/African American, and 4.8 percent as Asian. The fastest growing groups are Hispanics and Asians both increasing in size 43 percent from 2000. The fastest growing Latino subgroups are Mexicans (up 54 percent) and Dominicans (up 85 percent). Racial/ethnic groups that require special attention from psychiatric practitioners comprise an ever-increasing proportion of all persons served. For example, in NYS in its public mental health system (which excludes private practice practitioners), in 2013, 24.3 percent were black non-Hispanic, 23.3 percent were Hispanics, and 1.9 percent were Asian American.

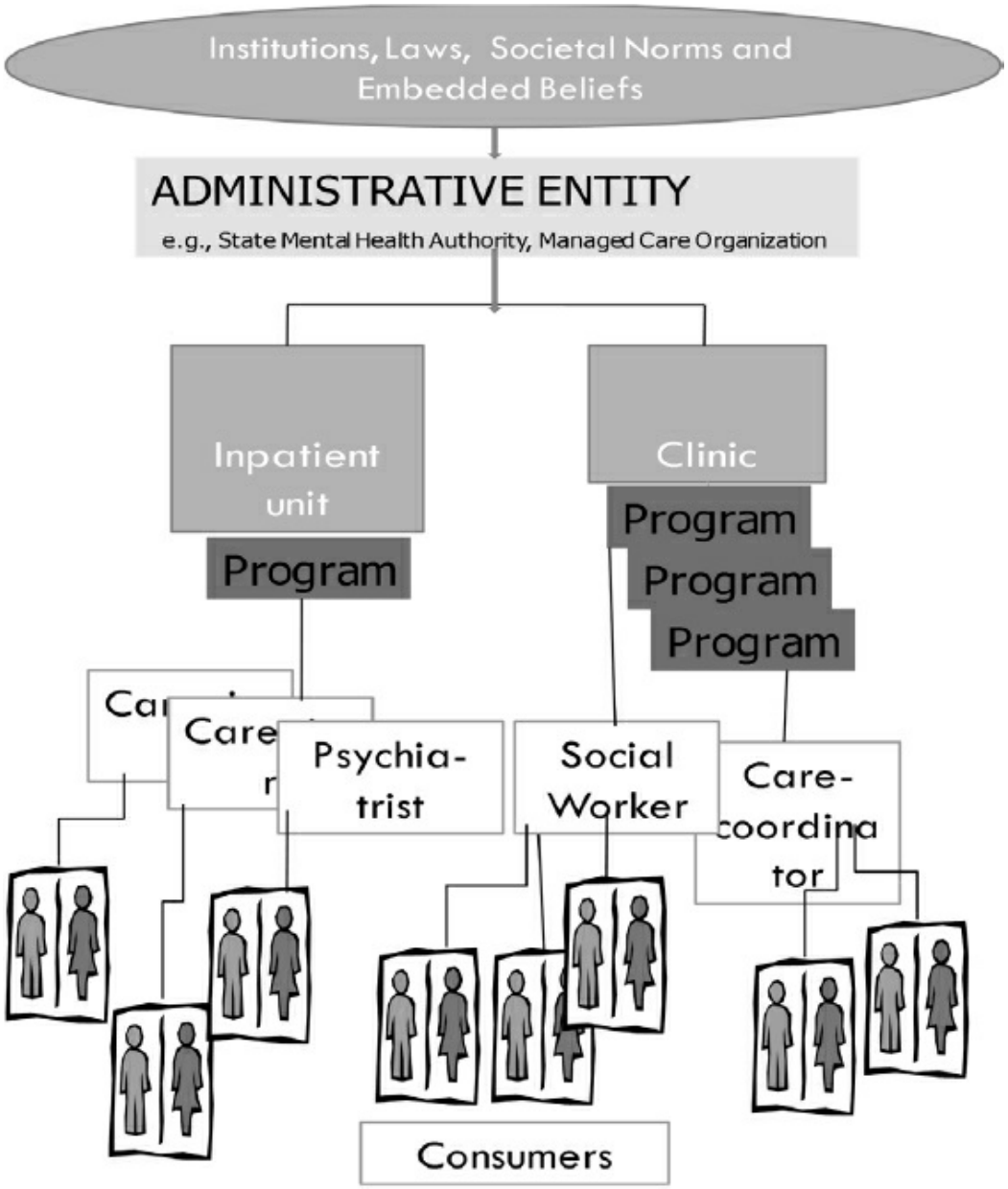

FIGURE 7.10–1. Hierarchical levels of health care system.

Continuing Evidence of Disparities

In 1978, the first Report of the President's Commission on Mental Health explicitly recognized critical disparities among racial and ethnic minorities. President Jimmy Carter and especially his wife Rosalynn became strong advocates for rectifying these disparities. A series of reports originating with the Western Interstate Commission on Higher Education commissioned by the Substance Abuse and Mental Health Services Administration focused on the four broadly defined racial/ethnic groups identified as underserved and underrepresented in managed mental health care—Americans of African Descent, Asian and Pacific Islanders, Latinos and Native American Populations—and compiled into a comprehensive volume a great depth and breadth of knowledge about the mental health

needs of these cultural groups. The Surgeon General's 2001 Supplemental Report reported on the mental health care of these major underserved racial/ethnic groups finding disparities in availability and access to needed services. The report endorsed CC in care delivery as a key approach for the reduction of these disparities, although there was no elaboration on how to do this.

There is strong evidence of sustained disparities in mental health. Researchers continue to find disparities for racial/ethnic minority groups in service accessibility, in service utilization, as well as in patient satisfaction. Asians remain less likely to access care than other groups. There are also still notable disparities in the reported prevalence of mental health conditions among groups with reasons for this at best only speculative. Several authors have linked prevalence differences to a misunderstanding of cultural presentations of symptoms resulting in both over and under diagnoses. Papers still speculate on reasons for higher rates of diagnoses of psychosis in black adults, of ADHD in black children and lower rates of diagnoses of depression in Latinos. Mechanisms of disparities are often linked to social determinants, for example, poverty, and to structural factors such as federal and economic health policy, operations of health care system, provider factors, and environmental contexts. Residential segregation of services and providers and provider shortages may also account for access and use disparities. For example, for Native Americans, it is noted that the Indian Health Service (IHS) remains understaffed and underfunded. An update on disparities across the country based on some recent studies conducted from 2010 is presented in Table 7.10–1.

GROWING RECOGNITION OF THE ROLE OF CULTURAL COMPETENCY IN CARE DELIVERY

Some authors locate the roots of a concern with cultural awareness in clinical training in 1950s–1960s post-war social science. A group of medical sociologists and anthropologists, Talcott Parsons among them, studies the professional acculturation and socialization of doctors during clinical training. What had been taken for granted as a primarily technical education was revealed to be a rich cultural milieu, replete with rituals, key symbols, rites of initiation, and social dramas, which led researchers and clinicians to a growing awareness of the "culture of medicine" as opposed to the culture of patients. There is a consensus, however, that the term "cultural competence" was not widely used in medical journals or clinical training until the 1980s, when the Association of American Medical Colleges and the Accreditation Council of Graduate Medical Education called upon medical schools and residency programs to train clinicians to address their patients' "culture," which largely referred to their patients' ethnic and racial identity. The term "cultural competency" and the concern

with outcome disparities came from the field of education, which had used the term with regard to addressing educational attainment disparities by ethnicity and race.

In psychiatry, this awareness of cultural differences between practitioners and patients as relevant to clinical practice was codified in the DSM-IV, published in 1999. Psychiatrist-anthropologist Arthur Kleinman led a special committee of the DSM-IV Task Force devoted to cross-cultural psychiatry that introduced new diagnostic frames and clinical tools to the DSM IV appendix. These included "culture bound syndromes" and the Cultural Formulation, a guide to interviewing and diagnosing patients using questions that took the patients' cultural understanding of their problem into account, using Kleinman's Explanatory Model. This contrasted with the symptom checklists of the DSM-III, which had aspired to use universally applicable symptom descriptions, descriptions that soon proved to be interpreted differently across ethnic and language groups, and which some practitioners argued led to misdiagnosis. At about the same time, Derald and Stanley Sue were advocating for the cultural competence of clinicians particularly in their work in counseling and psychotherapy. They emphasized the need for a greater understanding of psychiatric symptoms that could be misunderstood across the cultural divide between Western medicine and Chinese patients. They emphasized the need for addressing cultural beliefs, attitudes, knowledge and skills and developing new counseling competency standards when working with multicultural populations.

By 2000, as the DSM-IV Cultural Formulation gained visibility, and as increasing numbers of medical schools adopted some form of training in cross-cultural patient care, competing models of training in cultural issues emerged. The first might be called cultural expertise, in reference to the goal of having medical trainees gain knowledge of the beliefs and behaviors of their patients' cultural groups. In practical terms, this led medical schools to emphasize language translation services and, in some cases, "cultural consultants" who are members of minority ethnic groups employed by medical centers to assist healthcare providers in communicating with their patients. The second approach might be called cultural sensitivity, in which the goal is not to master bodies of knowledge about the cultural groups from which patients come, but rather to understand the broader psychosocial context of each patient, to ask each patient's guidance in learning what is relevant to him or her, and to heighten the cultural self-awareness of clinical practitioners so that they can take their own biases into account in their exchanges with patients. These two approaches, cultural expertise and cultural sensitivity, are not mutually exclusive but many medical training programs emphasize only one.

Table 7.10–1.
Selected Disparity Literature Review: 2010 to 2014

Authors (year)	Title	Objective	Sample	Conclusion
Service Accessibility and Utilization				
Wu et al. (2010)	Mental Health Service Use Among Suicidal Adolescents: Findings From a U.S. National Community Survey	Assess patterns of mental health service use among adolescents who had attempted suicide and examine service at individual and community level	877 adolescents (12–17) who attempted suicide in past 12 months	► Mental health needs of minority, lowe income familie are unmet ► Larger race/ethnic disparities in use of in and outpatient services than i school-based services ► School-based services more accessible to suicidal adolescents who may experience barriers in receiving standard care
Alegria et al. (2011)	Disparities in Treatment for Substance Use Disorders (SUD) and Co-Occurring Disorders for Ethnic/Racial Minority Youth	Review literature on racial/ethnic disparities, focusing on SUD services and comorbid behavioral health disorders in children and adolescents	N/A	► Compared wit non-Hispanic whites, black adolescents with SUD reported receiving less specialty and informal care ► Latinos reported receiving less informal care ► Mechanisms c disparities identified in federal and economic health policy, operations of health care system, provider factors, environmenta

				context, patie factors, etc.
Cummings et al. (2011)	Racial/Ethnic Differences in Mental Health Services Use Among Adolescents with Major Depression	Examines difference in mental health services use in NH-whites, black, Hispanic and Asian adolescents experiencing major depression	7,704 adolescents (12–17) with major depression in past year—nationally representative	► Black, Hispani and Asian adolescents significantly le likely than nor Hispanic white to receive prescriptions medication, tc receive treatment for depression from specialist or provider an to receive outpatient treatment, aft adjusting for family income and insurance status
Akincigil et al. (2012)	Racial and Ethnic Disparities in Depression Care in Community-Dwelling Elderly in the United States	Investigated racial/ethnic disparities in diagnosis and treatment among community-dwelling elderly	Medicare Current Beneficiary Survey, *n* = 33,708	► Blacks significantly le likely than nor Hispanic white to receive depression diagnosis from provider and i diagnosed less likely to be treated
Cook et al. (2013)	Assessing Racial/Ethnic Disparities in Treatment across Episodes of Mental Health Care	Investigating race and ethnic disparities in decisions to start and terminate care and decisions during treatment	Whites, blacks, and Latinos with probable mental illness	► Blacks and Latinos initiate care less frequently ► Black and Latino episode were shorter ► Black and Latino had fewer psychotropic drug fills
Kim et al. (2011)	Limited English proficiency (LEP) as a barrier to mental health service use: A study of Latino and Asian immigrants with psychiatric disorders	Examine the effect of LEP on mental health service among immigrant Latino and Asian adults with psychiatric disorders	372 immigrant adults (18–86) with any instrument-determined mood, anxiety, and SUD	► LEP of Latinos but not of Asians significantly reduced odds mental health service use ► Asians less

				likely to acces care
Jimenez et al. (2012)	Disparities in Mental Health Service Use of Racial and Ethnic Minority Elderly Adults	Apply the IOM definition of health disparities, to measure disparities in different episodes of mental health care	1,658 adults >60 with likely mental health care need	► Treatment initiation is significantly lower and the adequacy of treatment significantly poorer for bla and Latino elderly than white counterparts ► Latinos had significantly longer treatment duration, mor visits and high expenditures
Merikangas et al. (2012)	Service Utilization for Lifetime Mental Disorders in U.S. Adolescents: Results of the National Comorbidity Survey–Adolescent Supplement (NCS-A)	Examine rates and sociodemographic correlates of life-time mental health service use by severity, type and number of mental health disorders	6,483 nationally representative adolescents (13–18)	► Black and non Hispanic black were less likel to receive mo and anxiety disorder services than white counterparts
Siegel et al. (2013)	Access to and Use of Non-Inpatient Services in New York State Among Racial-Ethnic Groups	Report on differences in use of noninpatient services in Hispanics, Asians and non-Hispanic blacks and whites using NYS public mental health system	NYS OMH 2009 Patient Characteristic Survey	► Higher treated prevalence for Latinos and blacks compared to whites ► Lower prevalence rates for Asian ► Lower treatment intensity amo black youth with disruptiv disorders, Hispanic adult with anxiety and Asians <65 with depressic compared to white counterparts ► In upstate metropolitan

				areas, black youth and Hispanic adult had lower treatment intensity than whites ► In NYC, whites >65 had highe treatment intensity than other groups
Patient Satisfaction				
Huang et al. (2011)	The Effects of Discrimination and Acculturation to Service Seeking Satisfaction for Latina and Asian American Women: Implications for Mental Health Professions	Provide full description of Latina and Asian American women's experiences in mental health seeking service and identify ways to improve satisfaction	4,649 women from national survey of Latino and Asian American (NLAAS)	► Perceived gender and race/ethnicity discrimination in Latina and Asian America women negatively predict satisfaction in mental health service seekin ► Older age, higher education leve duration in the United States, and better mental health positively predicted satisfaction
Risk for Disorder or Condition				
O'Donnell et al. (2011)	Increased Risk of Suicide Attempts Among Black and Latino Lesbians, Gay Men, and Bisexuals	Assess whether major depressive and substance use disorders mediates higher prevalence of suicide attempts among LGB minority youth than white counterpart	388 Latino, Black and White LGB New York City residents (18–59)	► Black and Latino LGB youth had significantly increased risk suicide even without depression an SUD markers
Huang et al. (2012)	Physical and Mental Health Disparities among Young Children of Asian Immigrants	Examine physical and mental health functioning among Asian American children of U.S.-born and immigrant	7,726 Asian and U.S.-born white children	► Asian children of foreign-bor families at greater risk than children U.S.-born whit families of poo

		parents		physical healtl internalizing problems and inadequate interpersonal relationships
Morgan et al. (2013)	Racial and Ethnic Disparities in ADHD Diagnosis From Kindergarten to Eighth Grade	Identify racial/ethnic disparities in diagnosis and treatment of ADHD in children kindergarten to eighth grade	17,100 national representative sample from Early Childhood Longitudinal Study	► Minority children less likely to receiv ADHD diagnos than white children ► Among childre with ADHD diagnosis, minorities less likely to be taking prescription medication
Siegel et al. (2014)	Why are racial/ethnic rates of childhood Attention-Deficit/Hyperactivity Disorder in national studies highest among whites?	Examine race/ethnic differences in ADHD treatment of children in large public mental health system and compare with national rates	Children 3–17 receiving care in NYS public mental health system	► Contrary to national findings, ADHI prevalence rates in public mental health system are higher in whit children excep for Hispanic females ► Authors reaso that white children have greater access to private providers through privat insurance and make greater use of general practitioners
Provider Shortage				
Gone et al. (2012)	American Indian and Alaska Native Mental Health: Diverse Perspective on Enduring Disparities	Investigate mental health disparities unique to AI/An community	N/A	► Little oversigh of mental health service in Indian Healt Service (IHS) ► IHS remains understaffed and underfunded
Cook et al. (2013)	The role of provider supply and	Identify factors such as county-level	12,241 respondents	► Significant racial/ethnic

	organization in reducing racial/ethnic disparities in mental health care in the United States	provider organizations and SES characteristics associated with MH services and difference across race/ethnicity	from the National Latino and Asian American Study, National Comorbidity Survey Replication and National Survey on American Life	disparities in use of any anc specialty MH care after controlling for geographic variation ► Latino residential segregated areas had shortages of psychiatrists and Black residential segregated areas had shortages for medical specialties
Dinwiddie et al. (2013)	Residential segregation, geographic proximity and type of services used: Evidence for racial/ethnic disparities in mental health	Examines role of residential segregation in differential access and mental health service utilization	9,738 whites; 3,362 blacks, and 5,053 Latinos living in Metropolitan Statistical Areas	► High Latino isolation and centralization areas had shortage of psychiatrists ► High black concentration areas had access to psychiatrists ir their area ► Respondents i black neighborhood treated by nonpsychiatris and general doctors ► Respondents i Latino neighborhood treated by general docto ► Respondents living in high Latino isolatio and centralization areas more likely to utilize nonpsychiatris

Expanding on the cultural sensitivity approach, the term “Cultural

Humility" was introduced by pediatricians, who offered "humility" instead of "competency" because it conveys the attitude a doctor would need to adopt in order to learn from the unique circumstances and thinking of each patient. The practical elements they proposed were (1) self-evaluation of the health care provider to root out cultural assumptions that might reinforce their stereotypes of and discrimination against patients, (2) empowering patients to have a voice in their own decision making, and (3) forming partnerships with communities to advocate for more locally responsive health care.

Another expansion on the idea of provider self-evaluation was introduced by psychiatrist-social scientists Elizabeth Bromley and Joel Breslow in an approach to clinical education that develops "Critical Thinking Skills." They propose heightening trainees' self-awareness of the culture of medicine by teaching them the history of medicine and sociocultural research on medicine. They argue that this prepares trainees for leadership in medicine through understanding their own role in the larger institutions that are constraining patients' care and options. For example, trainees learn from this material that the U.S. health care system locates disease in the individual, rather than the community or the environment, which explains its focus on "compliance/adherence" rather than on the patient's milieu, their resources and supports. Another example, clinicians are trained to tell diabetics to limit their intake of sugar and processed foods, but not to ask if they live near grocery stores with fresh foods or if they know how to cook.

GOVERNMENT RESPONSE: STANDARDS, REGULATIONS, AND LEGAL REQUIREMENTS

The civil rights movement of the sixties heightened concerns about discrimination in health care. Title VI of the Civil Rights Act of 1964 states that "no person in the United States shall, on the grounds of race, color or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance." A primary focus at the federal level has been to remove language barriers that can prevent access to both general and mental health care. The Americans with Disabilities Act of 1990 mandated accessibility to facilities and programs for persons with disabilities, and this included psychiatric conditions. It also provided an impetus for introducing principles of CC into federally funded mental health programs. In 2000, President Bill Clinton signed Executive Order 13166 "Improving Access to Services for Persons with Limited English Proficiency" which requires federal agencies and those receiving federal funding to identify service needs of persons with LEP and implement systems to ensure persons they have meaningful access to services without unduly burdening the mission of the agency. Guidance issued by Department of Justice spells out

compliance standards to be followed to ensure programs and activities provided in English are accessible to persons with LEP. Ensuring that language is no barrier to health care, this action addresses discrimination on the basis of national origin which was prohibited by Title VI. In 2000, the U.S. Department of Health and Human Services, through the Office of Minority Health, published National Standards for Culturally and Linguistically Appropriate Services in Health Care, referred to as the CLAS standards. However, while provisions for culturally competent care and administrative and organizational supports were encouraged, only the provisions for linguistic assistance were made a mandated requirement. Neither the metrics to measure compliance to these standards nor enforcement guidelines accompanied the standards, although these came at a later date. The CLAS standards were revised in 2013 and augmented with an extensive set of guidelines for an organization to employ these standards (see Table 7.10–2). The new CLAS standards are organized into overarching themes. Communication and language access; Engagement; and Continuous improvement and accountability encompass the original standards, and a new standard Governance, Leadership and Workforce is included to ensure that CLAS standards are integrated and implemented throughout an organization.

The CLAS standards have been tied to the review criteria of the Joint Commission (formerly Joint Commission Accreditation of Hospitals/Healthcare Organizations) for the accreditation of hospitals and other health care organizations that have been expanded to include more explicit cultural and linguistic elements. The Joint Commission is essentially the gatekeeper for hospitals to obtain federal funds for inpatient care and have focused their reviews on an organization's safety and quality of care standards. Accreditation of hospitals and medical schools require adherence to these standards. Additionally, the National Committee for Quality Assurance has standards for cultural competence that health care organizations and managed care plans must address as part of their accreditation process.

GETTING BEYOND THE RHETORIC: DEFINING THE ACTUAL ACTIVITIES OF CC

Overview

On a clinical level, there is skepticism as to whether or not CC can be defined precisely enough to operationalize its aspects into clinical training and best practices. On an organizational level, many feel that CC is achieved once language accommodations are made such as having interpreters available when needed and having key forms in translated versions. Hospital signs in multiple languages are now commonplace. While certainly necessary these steps are not sufficient. But substantial progress has been made on isolating specific activities and defining and

developing tools and products to promote activities of CC. It is clear that CC activities need to occur on the multiple levels of a care organization in order to achieve quality outcomes.

Figure 7.10–2 indicates that the policies and procedures of an organization impact the kind of staff employed and the limits placed on their therapeutic work. The composition of the clinical team, the kind of supervision that team members receive, and the approaches taken to coordinate care define the human resource boundaries of the program. The program's work is conducted within a palpable milieu which patients encounter and react to. Clinical work uses particular clinical tools and service delivery modalities that are constrained by reimbursement issues that impact service choices and the duration of delivery. The care that is ultimately delivered by clinicians is subject to their competencies.

Table 7.10–2.
National Standards for Culturally and Linguistically Appropriate Services in Health Care (HHS, Office of Minority Health, 2000)

Principal Standard:

1. Provide effective, equitable, understandable, and respectful quality care and services that are responsive to diverse cultural health beliefs and practices, preferred languages, health literacy, and other communication needs.

Governance, Leadership, and Workforce:

2. Advance and sustain organizational governance and leadership that promotes CLAS and health equity through policy, practices, and allocated resources.
3. Recruit, promote, and support a culturally and linguistically diverse governance, leadership, and workforce that are responsive to the population in the service area.
4. Educate and train governance, leadership, and workforce in culturally and linguistically appropriate policies and practices on an ongoing basis.

Communication and Language Assistance:

5. Offer language assistance to individuals who have limited English proficiency and/or other communication needs, at no cost to them, to facilitate timely access to all health care and services.
6. Inform all individuals of the availability of language assistance services clearly and in their preferred language, verbally and in writing.
7. Ensure the competence of individuals providing language assistance, recognizing that the use of untrained individuals and/or minors as interpreters should be avoided.
8. Provide easy-to-understand print and multimedia materials and signage in the languages commonly used by the populations in the service area.

Engagement, Continuous Improvement, and Accountability:

9. Establish culturally and linguistically appropriate goals, policies, and management accountability, and infuse them throughout the organization's planning and operations.
10. Conduct ongoing assessments of the organization's CLAS-related activities and integrate CLAS-related measures into measurement and continuous quality improvement activities.
11. Collect and maintain accurate and reliable demographic data to monitor and evaluate the impact of CLAS on health equity and outcomes and to inform service delivery.
12. Conduct regular assessments of community health assets and needs and use the results to plan and implement services that respond to the cultural and linguistic diversity of populations in the service area.
13. Partner with the community to design, implement, and evaluate policies, practices, and services to ensure cultural and linguistic appropriateness.

14. Create conflict and grievance resolution processes that are culturally and linguistically appropriate to identify, prevent, and resolve conflicts or complaints.
15. Communicate the organization's progress in implementing and sustaining CLAS to all stakeholders, constituents, and the general public.

From U.S. Department of Health & Human Services. Office of Minority Health. National CLAS Standards. Available at https://www.thinkculturalhealth.hhs.gov/content/clas.asp

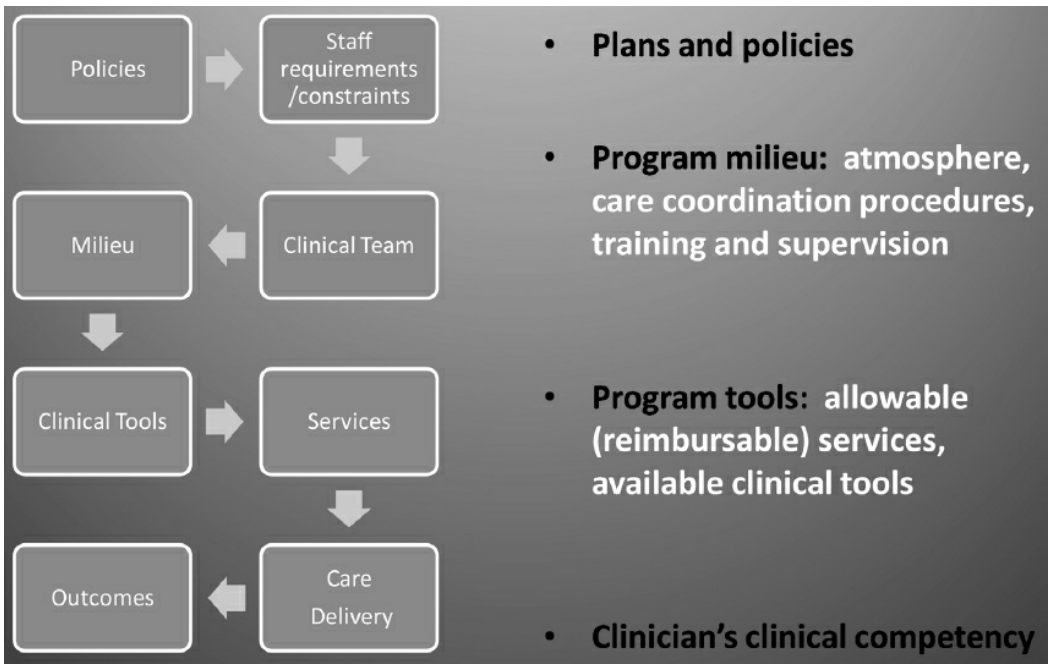

FIGURE 7.10–2. Factors in a care organization impacting care outcomes.

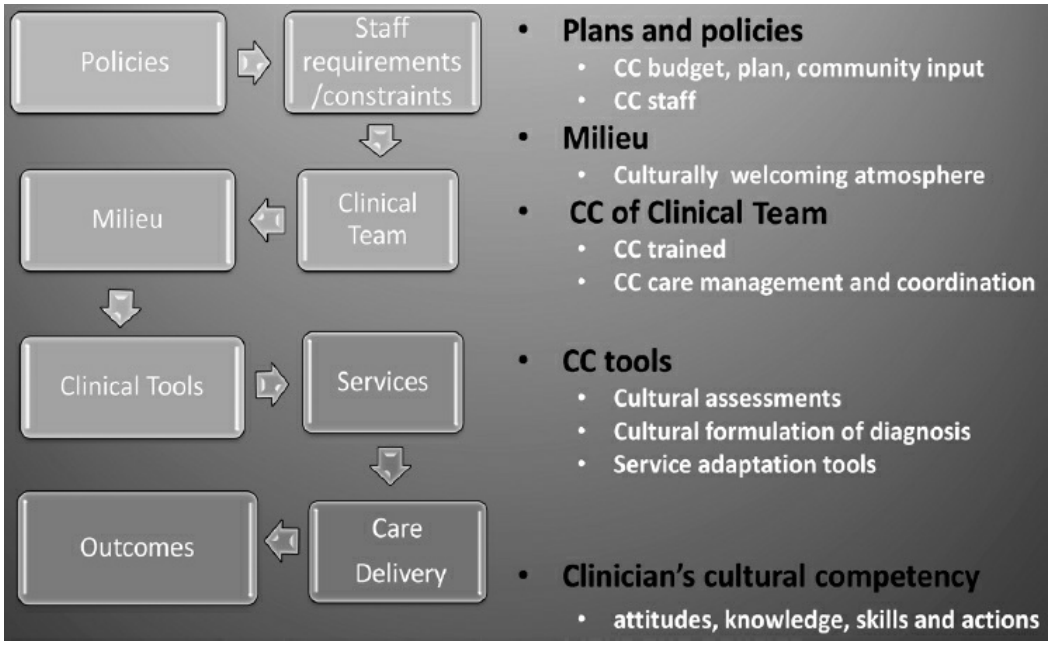

FIGURE 7.10–3. CC activities impacting care outcomes.

Figure 7.10–3 illustrates how CC needs to guide these activities. The policies of an organization that works with multicultural populations need

to address CC deliberately. The starting point is the incorporation of CC into a mission statement of the organization, allocation by the organization of a budget to enable a person/group to be assigned to spearhead CC and the development of an organizational plan for CC. The plan requires community, provider, and client input. The desirability of staff diversity needs to be addressed in hiring policies, and training activities need to be organized for all levels of the organization. The clinical team needs to address cultural strengths and challenges when managing a client's care and coordinating services. The milieu of the program needs to be welcoming to cultural groups and accommodate their special needs, for example, by having convenient hours of operation, policies on privacy and child friendly areas. When delivering services, cultural assessments need to be made in order to appropriately diagnose the client, and develop a service plan that will act to engage and retain the client in services. Services need to be chosen that are culturally acceptable to a client, and if sufficient numbers warrant it, services specifically chosen or adapted for cultural groups. In service delivery, clinical competency needs to be augmented with a requirement for CC as the behaviors of clinicians, based on their attitudes and skills can impact outcomes.

Specific Clinical Activities

Psychiatrists play a key role in engagement and retention of persons from cultural groups in care. They often are the gate keepers for entry into specific types of services, determine specific treatments and serve as arbiters in treatment team meetings. In the clinical hierarchy, they have a distinct power role thus impacting other clinicians as well as clients. CC is, however, not a concept discussed in most training programs and not yet a focus of clinical care, particularly in programs that strongly endorse medical models as explanatory of illness and symptoms and medications as the primary modalities of psychiatric care. Being clinically competent and practicing patient-centered care is endorsed and could include CC, but to date is only addressed in a hand full of training programs. But buy-in to CC by psychiatrists is critical to good care, as in order to engage persons from cultural groups in care and to have them remain in care, they need to trust their psychiatrists. Lack of trust and feelings of stigma can impact engagement in services and adherence to treatment regimens. Trust can only be engendered if clinicians are cognizant of cultural views on mental illness that can impact treatment.

Cultural Competency Curricula for Clinicians

There are several curricula that address ways to enhance the CC of psychiatrists, counseling practitioners, psychologists, and social workers. A comprehensive review of such training guides has been conducted by

researchers at the Canadian Centre for Addiction and Mental Health Centre de Toxicomanie et di Santé Mentale. They found a consistency across guides in the use of the Sue and colleagues framework of CC encompassing personal level awareness, knowledge, and skills. Awareness modules address understanding and conciliating a client's and one's own biases, beliefs, attitudes, and world views. Anti-racist and anti-oppression conceptual models may be invoked. Knowledge attainment addresses understanding both the generic cultural factors that can impact care, for example, social determinants of illness, level of acculturation, immigration status as well as specific culture factors such as a cultural view of illness causation and expectations for recovery. Skill component modules in the more practical guides suggest approaches for engagement and problem solving based on cultural views of clients, as well as on the tangibles of working with interpreters and seeking consultation from clients natural care givers such as clergy and traditional healers. Case studies and scenarios are often introduced here. In addition, several guides also address clinician–client power relationships and even system–client interactions. In the former, sensitivity to issues of power differentials that can engender either trust or mistrust is addressed. In the latter, the curricula cover areas overlapping with recent ideas on structural factors impacting relationships.

We review in greater detail one curriculum developed by psychologists Deborah Altschul and Jeff Brooks-Harris for the broad swath of counseling practitioners. The Multicultural Skills Training (MST) model is a practice-based method of CC that aims to enable providers to adapt their current practices to meet the needs of culturally diverse populations. It is applicable across a broad array of service modalities including psychotherapy, psychoeducation, and skills training. It also draws upon the Sue and Sue tripartite seminal model of training offering practical approaches for understanding, in addition to the cultural views and biases of clients, one's own views and suggests ways to conciliate clinician–patient views. It addresses the need for practitioners to integrate into their therapeutic approach ways to empower clients to deal with the high likelihood of impacting social determinants of health such as housing needs.

MST is arranged as six clusters of culturally based activities to help better match the individuals' needs to therapeutic care. To illustrate, the module Becoming Aware of Your Own Multiple Worldviews a prerequisite to working with multicultural clients recommends confronting your own personal culture-based worldview as it influences views of mental health and the etiology of mental illness and appreciating your own multiple cultural identities and their impact on your thinking and treatment decisions. When working with clients, Creating a Culturally Sensitive Working Relationship includes knowing the cultural norms and customs of

a cultural group as they manifest in a particular individual and determining the style of collaboration that is most comfortable for the consumer. Understanding the Individual's Worldviews and Facilitating Cultural Identity Development parallels the work of the first caregiver module for the client and suggests exploring with clients the historical backgrounds of their cultures, as well as the current placement of the client in that culture. There is focus on understanding cultural norms about help-seeking, problem attribution, and expected outcomes. On the client's part, it is helping the client to value their own cultural background but also to become open to other worldviews as well. Open discourse is suggested to help clients deal with experiences that they may have had that they perceived as racist or biased in social and work interchanges. In Responding to Societal Structures and Value counselors are encouraged to work with consumers to confront their feelings of being undervalued because of cultural identity and social status. This could be based on having received slurs, felt racial discrimination, or having felt undervalued as transmitted to others in ambient societal messages. Empowering consumers through social activation and encouraging them to make greater use of cultural resources such as religious affiliations is recommended. A final cluster Utilizing Complementary Roles beyond Psychotherapy suggests ways to incorporate nontraditional healing methods of their cultures into the therapeutic work.

Clinician Self-Assessment of Cultural Competency

There are several practitioner level cultural competency self-assessment scale (CCAS). While self-assessments can be pointless because of their subjective structure, if appropriately used they can serve further as educational tools. One with established psychometrics that is widely used is the Multicultural Counseling Inventory. It addresses the three domains of awareness, knowledge, and skills. It is comprised of 40 items on which the provider rates his/her self-perceived cultural competence across the subscales Multicultural Counseling Skills (Sample Item: "When working with minority clients, I am able to quickly recognize and recover from cultural mistakes or misunderstandings"), Multicultural Awareness (Sample Item: "I am involved in advocacy efforts against institutional barriers in mental health services for minority clients"), Multicultural Counseling Relationship (Sample Item: "When working with minority clients, I find that differences between my worldviews and those of the clients impede the counseling process"), and Multicultural Counseling Knowledge (Sample Item: "When working with minority clients, I make referrals or seek consultations based on the client's minority identity development).

Domains of Organizational and Program Activities

Except for that provided by psychiatrists or other professionals working entirely independently, most mental health care is organized as programs, bundles of available services and treatment strategies geared toward ameliorating psychiatric distress. Programs in turn are components of mental health care agencies, centrally administered entities that provide policy and procedural guidelines for the programs under their purview, and obtain and allocate resources. From an organizational or program perspective, CC activities should take place in the domains illustrated in Figure 7.10–4. Needs assessments refer to gaining an understanding of the diverse cultural groups using a care system, their possible service gaps and specific service needs. Information exchange refers to mechanisms for receiving and transmitting information on community needs to persons responsible for planning and delivering services. Reflexively, deliberate efforts are required for communities to receive information on services available and ways to access these services. CC tools need to be disseminated to caregivers to improve their ability to provide CC services. Services that are offered need to be appropriate for community cultural groups and modifications to existing services made and/or new services introduced. Language accommodations may need to be made. Organizations need to promote and sustain recruitment, hiring, and retention procedures of culturally competent bilingual and bicultural staff as well as to educate and train existing staff in CC. Linkages to other systems need to take into account the cultural needs of clients. Outcomes related to disparities in access and quality of care to cultural groups, desirable cultural group outcomes, and satisfaction with services need to be distinctly monitored. These activities are expected to improve outcomes and need to be monitored to see if they are taking place and the impact they may have. They need to be endorsed through the policies and procedures of the organization.

Both an organizational level instrument and a program level instrument to assess CC have been developed by the Nathan Kline Institute Center of Excellence in Culturally Competent Mental Health. The guiding principles were to have objective measures that had demonstrated importance to the concept of CC, for which it was feasible to collect appropriate data that were reliable and accurate. While clinicians can be assessed on their self-perceived attitudes and beliefs, organizational and program level items were selected to be verifiable.

The organizational level scale items were selected based on a review of recommendations, standards, and guidelines. A 27-site study succeeded in establishing the predictive validity of the several language accessibility items of the scale showing they correlated with better engagement and retention. The program level scale was based on an ethnographic study of three long-standing programs defined cultural groups in which deliberate and culturally grounded steps had been taken to make them acceptable.

Key and common elements were observed across these programs, and used to define domains of CC. It is informative to review these instruments as they were designed in form not only for assessment purposes but intended to serve as roadmaps for introducing CC activities into organizations and programs. Each criterion is scored with an ordered scale of whether or not an activity is accomplished reflecting an increasingly complex degree of activity related to CC. A score indicates that the listed activity and all those below it are carried out. The scale items are used to step through specific activities that should be undertaken to promote CC.

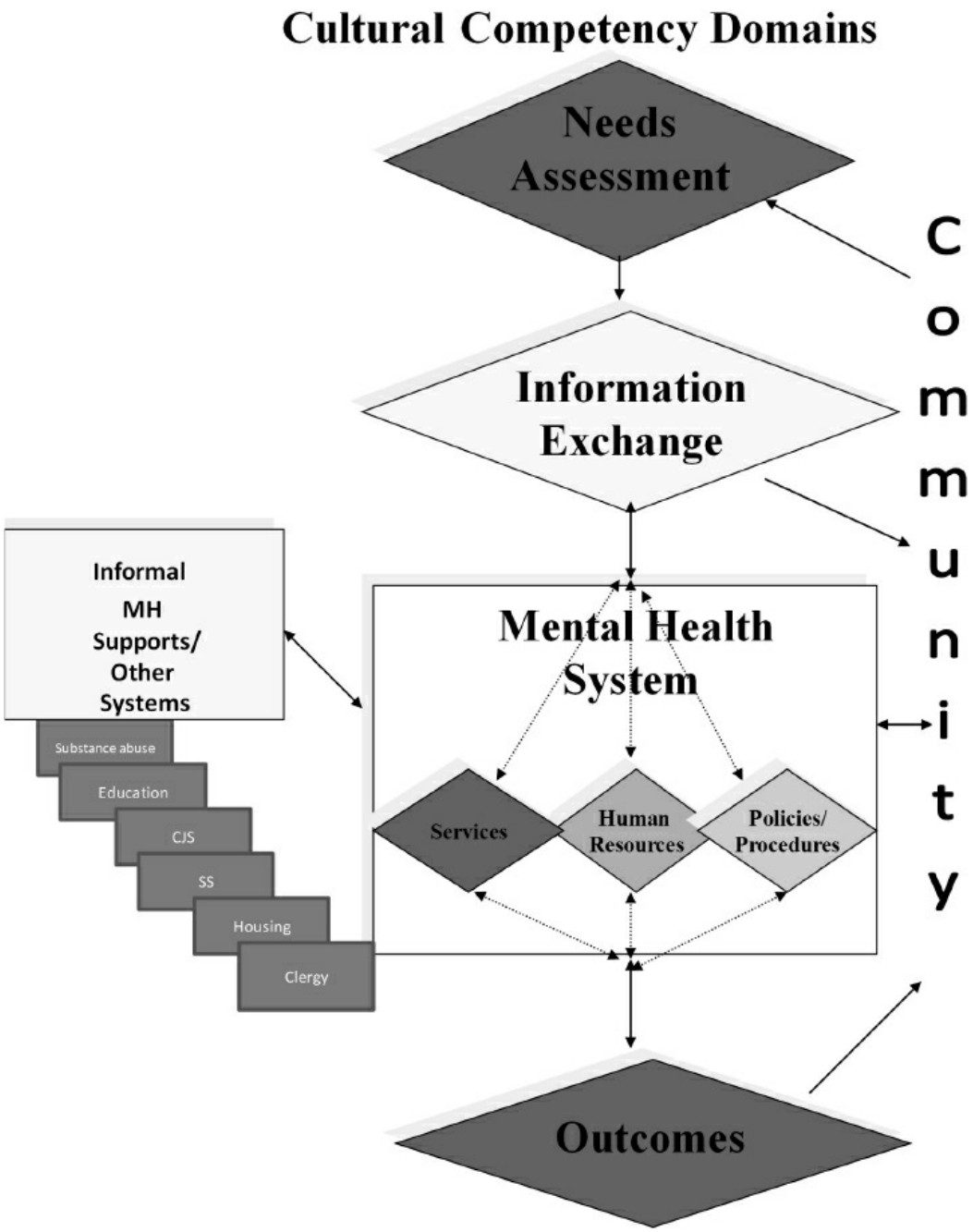

FIGURE 7.10–4. Cultural competency domains.

Organizational Activities

The first set of criteria is meant to ensure that a procedure for setting up a CC presence in an organization is established. Fundamental is that an organization have an articulated commitment to CC evident in its mission statement. By creating a position with an assigned budget and having this person/group part of the governance of the organization, there are tangible assurances that CC activities can take hold. Strategies for obtaining information about and working with members of diverse cultural groups include establishing a committee that includes members of the largest cultural groups in the community, and empowering that group to review current CC work and communicate its concerns and ideas directly with management. The organization needs to identify which cultural groups are being served, what their language needs are and which members of the community are not being seen in order to target its outreach efforts. Particular data items need to be collected for all clients in a consistent manner so that they can be aggregated to assess the cultures and language needs of the population being served by the agency. Information on the target community allows the agency to tailor its outreach and services to the needs of its cultural groups.

Several criteria relate to staff. Hiring, recruiting, and retaining staff that are from prevalent cultural groups or who have experience working with these groups is often noted as ideal. This diversity should extend beyond clinical staff to include administrative and supervisory staff as well. Organizations can have policies in place that seek culturally representative employees and strategies to identify suitable candidates. Staff training if related to the work of staff might be an effective way to enhance the CC of all, both members of cultural and mainstream groups.

Literacy level and English proficiency can be barriers to accessing services. Several criteria address language accommodation. This could be accomplished through hiring bilingual staff or if not able to do so having interpreter services available for all service users from the point of first contact with the organization. Additionally, organizations need to recognize that key forms and documents need to be translated into commonly encountered languages in order to work effectively with persons having LEP. Furthermore, service descriptions and educational materials also need to be translated, and for persons with limited reading skills alternate formats can be introduced (e.g., novellas, videos, and items in simplified English). Finally, services that do not appear to be effective for persons from certain cultural groups may need to be adapted for use with that cultural group, or new services introduced that have been developed specifically for prevalent cultural groups (Table 7.10–3).

Program Activities

The program-level CC assessment scale identifies activities that are more directly associated with delivering care than organizational activities. The criteria address strategies for engaging consumers from cultural groups into a program, measures program capacity for meeting language needs of clients including gaining an understanding of and working with communication styles of clients. The appropriateness of the services offered to the cultural groups served, the use of support groups including peers, family and kin and utilizing cultural community resources to sustain a consumer's recovery are also measured. The scale examines program outcome goals and prods a program to include into outcome measurement cultural groups' understanding of and desirability for different modes of recovery. Finally keeping in mind the health care reform goal of population health, community outreach to cultural groups is assessed.

Engagement in a program means that the client will receive a sufficient dose of the program so that the maximum benefit of the services can be realized. If appropriately engaged, it is likely that participants will adhere to treatment regimens and complete the work of a program. To engage clients, programs need to understand the cultural needs of their clients. A welcoming milieu in the program is essential as well. The scale measures whether assessments go beyond the collection of standard demographic information such as age, gender, race/ethnicity, and preferred language to include more detailed information on cultural group identity, level of acculturation, and religious/spiritual beliefs. Primary to engagement is addressing language needs and ensuring effective communication with clients through an understanding of and adjustment to their possibly different styles of communication. Examining ways to increase trust in the staff and addressing the experience of stigma from the perspectives of different cultures also increase engagement.

The scale examines whether services have been culturally modified or new culture-specific services have been offered. Certain service modalities may not be suitable for cultural groups. This is true even for some highly endorsed evidence-based practices. Exploring the possibility of making modifications to services to make them more culturally relevant can be accomplished with the input from members of the cultural groups and further ensures a connectedness of the program with the cultural groups it serves. For example, group therapy that requires sharing of personal information with strangers is not an acceptable mode of care for elderly Chinese populations. However, a mahjong group run by an experienced therapist can combine the positive aspects of both a culturally enjoyable experience and a profitable therapeutic encounter.

Utilizing various support groups available to the client can enhance the possibility of positive outcomes. The scale examines the extent to which the program includes peer staff from the cultural group. These are both paid employees and volunteers, who have mental health experience. Staff from

the same cultural groups of clients enrolled in the program who have mental health experience can facilitate engagement of cultural clients in the program by enhancing consumer trust in the program and providing role models of recovery. They can also provide other program staff with information on cultural views impacting mental health treatment and acceptable treatment strategies. Family and social networks—as identified by the consumer—can play a large role in the lives of members of cultural groups, particularly in immigrant groups. Some consumers may rely heavily on particular family members for their supports and for maintaining their ties to their communities while others may rely more heavily on close friends and members of their social network. The extent to which these persons participate in assessment, care planning and providing ongoing supports is assessed. The scale also measures the extent to which a program identifies community resources that are acceptable and responsive to cultural values and preferences of consumers as referrals to promote healthy community living.

The scale identifies three ways to measure outcomes, a critical aspect of any assessment of health care programs. Service delivery outcomes such as patterns of use and program completion rates, while not measuring quality of care, should not differ among cultural groups and act as basic markers of disparities. Satisfaction surveys need to be formulated with members of cultural groups as their desirable recovery outcomes may differ from those of clinically comparable consumers from other groups because of different beliefs, relationships, and resources to achieve goals. Finally a program's outreach efforts are measured. In order to increase the enrollment of cultural groups into programs that might be of benefit to them, outreach efforts should be undertaken that include identifying these groups in the community and informing them about what the program offers to people from these cultures (Table 7.10–4).

STRUCTURAL COMPETENCY

The importance of organizational and program level activities for tailoring mental health interventions to diverse cultural groups, and for redressing inequalities in mental health care, raises the question of how individual mental health practitioners can contribute to organizational and program level change. This question inspired a concept that is being developed and implemented by a growing group of clinical practitioners, trainees, and health-oriented social science and community researchers, called "structural competency." "Structure" in structural competency refers to the community, institutional and policy level factors that clinicians should learn to see as drivers of their patients' health, while "competence" refers to the need for clinical practitioners to develop the skills to intervene on these levels. A structural intervention is a means to foster organizational and program change by changing community, institutional, and policy factors.

A growing body of research locates mental and physical health inequalities in the disparate social conditions of nondominant cultural groups, including institutional and policy driven factors such as the quality of housing, education, public parks and recreation, law enforcement patterns, and exposure to violence. International comparisons of health outcomes reinforce the view that social conditions are a crucial target for health interventions. This finding inspired the Rio Political Declaration on Social Determinants of Health that was adopted by 125 WHO member states in October 2011. In the United States, in a 2014 report of the IOM, it is noted that the United States has the poorest health outcomes among the industrialized nations despite spending the most on health care per capita and argued that institutional and social inequalities explained these patterns. These include unequal healthcare provision; behavioral health risks that are related to social conditions such as smoking, unprotected sex and the use of firearms; the lack of safety net programs for the poor; and the low quality of public services in poor neighborhoods.

Clinical practitioners have a direct lens on the health impact of social conditions on their patients, and, as health experts, have unique credibility in the public sphere when speaking about them and participating in collaborative efforts to address them. Yet there is little in clinical training to prepare clinicians to think and act at institutional levels. It is this gap between knowledge of the importance of social conditions to health, and action by clinical practitioners to improve those conditions, that structural competency seeks to address.

Table 7.10–3.
Organizational CCAS Criteria and Their Rationale

Criterion	Description	Rationale
1. **Agency's Commitment to Cultural Competence**	The agency (or its parent organization [PO]) has a management level person responsible for CC and a dedicated budget for CC activities, a CC plan and procedures for updating the CC plan.	A management level person who has primary responsibility for CC ensures that CC will be addressed. Without a dedicated budget, only limited activities can be conducted. A written plan concretizes the agency's commitment to CC. Review and updating ensures that there is a feedback loop and that corrective actions can be taken.
2. **Assessment of Service Needs**	The agency obtains current data on its service users and its target community that enable identification of their cultures and language needs	Particular data items need to be collected for all clients in a consistent manner so that they can be aggregated to assess the cultures and language needs of the population being served by the agency. Information on the target community allows the agency to tailor its outreach and services to the needs of its cultural groups.
3. **Cultural Input into Agency**	The agency (or its PO) has a CC Committee or	A committee dedicated to CC will enhance the likelihood that activities appropriate to the cultures

Activities	other group that addresses cultural issues and has participation from cultural groups of the target community	served are introduced and carried out. The committee can go under many names or be part of the functions of existing boards, advisory groups, and committees. Input is sought from representatives of the most prevalent cultural group of the target community. Having more than one representative from a cultural group makes active participation more likely.
4. Integration of CC Committee or Other Group with Responsibility for CC Within Agency	The CC Committee or other group with responsibility for CC is integrated within agency evidenced by the following activities: ► Reviews services/programs with respect to CC issues at the agency ► Reports to Quality Assurance/Quality Improvement program of the agency/PO ► Participates in planning and implementation of services at the agency ► Directly transmits recommendations to executive level of agency/PO	The extent to which the functions of the CC Committee are reported and used in the agency provides a measure of the likelihood of change with respect to CC. Four key committee functions are expected to take place, but they may be introduced at different stages.
5. CC Staff: Training Activities	The agency (or PO) offers to staff educational activities in which cultural issues are addressed and requires staff to have an adequate amount of specific training on CC	Training and educating staff in CC enhances the likelihood of the delivery appropriate services to cultural groups. Ideally, educational activities should take place every year and that all staff members who have face-to-face contact with and provide direct clinical care to agency clients receive CC training. Administrative staff should also be knowledgeable about CC.
6. CC Staff: Recruitment, Hiring and Retention of Staff from/or Experienced with the most Prevalent Cultural Groups of Service Users	The agency is committed to hiring and retaining CC staff who are from or who have had experience working with the most prevalent cultural groups of its service users	Having direct service, supervisory and administrative staff with relevant experience with the most prevalent cultural groups enhances the likelihood of the acceptability and use of CC practices. Hiring and retaining professional staff members who are from the cultures of service users provides positive role models for clients of the agency and affords additional opportunities to increase knowledge about the cultures.
7. Language Capacity: Interpreters	The agency (or PO) accommodates persons who have limited English proficiency (LEP) by using	Language needs of persons with limited English proficiency must come to the attention of the agency at the point of first contact in order to identify their needs and link them to the appropriate services. Once a person has agreed to

	interpreter services or bilingual staff	receive the services offered by the agency, language needs must continue to be addressed. Ideally interpreters are formally trained and certified or are bilingual staff members who have received CC training.
8. Language Capacity: Bilingual Staff	The agency has staff who speak the language of the most prevalent cultural group of service users with members who have LEP	Having bilingual staff members who can work directly with persons with language needs is ideal—and likely to be cost effective as well.
9. Language Capacity: Key Forms	The agency has translated versions of key documents and forms in the language of the most prevalent cultural groups of its service users	Legal documents and critical care information must be understandable to persons with limited English proficiency to ensure that such information is provided equally to all service users.
10. Language Capacity: Service Descriptions and Educational Materials	The agency accommodates persons with LEP by translating service descriptions and accommodates persons with limited reading skills (LRS) by providing service descriptions or educational materials in formats they can understand.	Using different formats and media for descriptions of available services and for educational materials ensures that literacy level is not a barrier to accessing services. Service-related materials translated into the most prevalent non-English language will be especially helpful to overcome treatment barriers for service users and family members.
11. Assessment and Adaptation of Services	The agency (or PO) reviews and adapts or introduces services suitable to the most prevalent cultural groups of service users	A comprehensive CC-based assessment of services may result in the adaptation of existing services and the introduction of new services to better serve members of the most prevalent cultural groups.

Table 7.10–4. Program Level CCAS

Criterion	Description	Rationale
1. Cultural Assessment of Consumer	Program obtains a detailed cultural assessment of each consumer that is additional to collecting standard (age, gender, race/ethnicity; preferred language) demographic information	Information that provides a view of beliefs, values, and world views of the consumer that can be used to improve diagnosing and the therapeutic alliance of staff with cultural groups. Information is recommended to be collected at the initial clinical assessment but can be updated throughout the course of the program. Cultural information should be made available to all clinical staff and care workers.
2. Engagement:	Program has	If staff can communicate with consumers/families in

Spoken Language Capacity	procedures and staff to facilitate working with consumers/families with limited English proficiency or with consumers who are deaf	their preferred language, the therapeutic encounter will be enhanced.
3. Engagement: Communication Styles	All levels of program staff communicate with persons from cultural groups in ways that are respectful to them	Barriers between staff and consumer are reduced when modes of communication are respectful, acceptable and familiar to the consumer.
4. Engagement: Trust Building	Program intentionally and actively works to build and maintain the trust of consumers and family members from cultural groups	Retention and engagement in the program is likely to increase if consumers feel their personal information is held confidential, and that staff care about them by being responsive to their values and preferences as they relate to the program.
5. Engagement: Culturally Framed Stigma Reduction	Program intentionally conducts activities to reduce stigma of mental illness as experienced by a cultural group	Stigma can impede engagement in services. Persons from cultural groups who might benefit from mental health services often delay seeking treatment or drop out of services because of the cultural group's views of mental illness including their attribution of its causes, its impact on other family members and their social standing in their communities.
6. Culturally Friendly Service Delivery Environment	Program is conducted in physical spaces that are made appealing and familiar to the cultural groups served	Features of surroundings that are familiar, welcoming to families and celebratory of cultural group enhance consumer and family feelings of comfort and a sense of connectedness while participating in the program.
7. Culturally Modified or New Services	Program has strategies for making their services responsive to values and preferences of persons from cultural groups with the aim of improving their effectiveness	Services need to be reviewed and modified for applicability and their responsiveness to values and preferences of the cultural group. New services may need to be developed.
8. Inclusion in Program of Staff From Cultural Groups Who Have Mental Health Experiences (Consumer Peers)	Program includes staff, both paid and volunteers, who have mental health experience and are from the cultural groups served by the program	Staff from the same cultural groups of clients enrolled in the program who have mental health experience can facilitate engagement of cultural clients in the program by enhancing consumer trust in the program and providing role models of recovery. They can also provide other program staff with information on cultural views impacting mental health treatment and acceptable treatment strategies.

9. Family Member Involvement	Program engages persons that are identified by the consumer from their family including from social networks to take part in care planning and providing ongoing supports	Family and social networks can play a large role in the lives of members of cultural groups, particularly in immigrant groups. Some consumers may rely heavily on particular family members for their supports and for maintaining their ties to their communities while others may rely more heavily on close friends and members of their social network.
10. Community Resources	Program utilizes community resources acceptable and responsive to cultural values and preferences of consumers to promote health outcomes	Consumers from cultural groups benefit from interventions that involve the use of community resources.
11. Satisfaction With Program	Program conducts activities to measure consumer and family member satisfaction with the services provided	Program improvement for cultural groups needs to be based on consumer and family member appraisals of the program.
12. Service Delivery Outcomes	The program examines disparities in service delivery outcomes related to engagement and retention in the program	Service delivery outcomes for cultural groups should not differ from those of clinically comparable consumers from other groups. Consumers from cultural groups may have particular difficulties in engagement into and retention in services. Differences in service delivery measures on engagement and retention could reflect these difficulties. These need to be identified and monitored by the program.
13. Recovery Outcomes	The program examines disparities in service process outcomes and identifies and monitors recovery outcomes that are desired by consumers from cultural groups	Recovery outcomes for cultural groups may differ from those of clinically comparable consumers from other groups because of different beliefs, relationships, tools, and resources to achieve goals. These need to be identified and monitored by the program.
14. Program Outreach	Program identifies cultural groups in the community/service area who underutilize services and might benefit from the program, and conducts outreach to increase their access	To increase the enrollment of cultural groups into programs that might be of benefit to them, outreach efforts should be undertaken that include identifying these groups and informing them about what the program can offer.

"Seeing" Social Structure

Before clinical practitioners can act on social conditions, they have to see and understand the connection between their patients' presentation and the institutional and social environment in which their patients live. The first step toward structural competence is recognizing those structures that shape clinical interactions. These then become the health relevant focal points at which to intervene on behalf of patients' health in order to work toward system changes. These focal points are referred to as "structures," and in the broadest sense include the public policies that have a disproportionate impact on health indicators. On a direct impact level structures include (1) community-based organizations including centers for youth, elders, and cultural affiliation, churches, and other faith-based organizations; and (2) health-relevant public sectors such as education, law enforcement, and housing. Structural competency for clinical professionals requires them to (1) become familiar with the structural determinants of clinical presentations that includes the social history of organizations, institutions, and other systems to health status; (2) rearticulate "cultural" presentations in structural terms, which means, in addition to incorporating the role of cultural beliefs and customary practices into treatment decisions, understanding the role of organizations, institutions, and other systems that impact health behavior; and (3) develop structural humility, which requires clinicians to avoid presenting themselves as sole "experts" on the structures that impinge upon their patients' health, but instead enlist the expertise of patients themselves, as well as that of potential collaborators in local community organizations, other health relevant sectors, and public policy. Clinical training is an important window of opportunity to enhance the awareness of health-relevant structures among practitioners. Understanding the social determinants of health is part of a growing number of medical school curricula. And for physicians already in practice, one study showed that simply prompting them to ask about key social indicators in EMR systems led to increased referrals to social services and to integration of care among service providers.

Structural Interventions

Structural interventions can take place on the individual, community, or policy level. On an individual level, clinicians can become more aware of social determinants of clinical presentations, for example, by including questions about social needs in clinical assessments. By doing so, an unstably housed patient could get appropriately referred to a low-income housing program. On the community level, clinicians can collaborate with community organizations, schools, law enforcement, parks, and recreation to address the needs of their patient population in a holistic way. They

might, for example, get involved in integrated community-based care by enhancing their clinical program to address multiple health and social needs. They might address chronic trauma/violence in poor neighborhoods by offering services such as support groups for the sequelae of trauma, or prevention efforts such as parental support and education to prevent child abuse, or youth centers to prevent gang involvement. On a policy level, they can educate policy makers and the public about the health impact of drug laws, housing, education, and employment policies.

One contemporary example of a structural intervention is that of the Brownsville Partnership, a New York City collaboration of 27 agencies and community-based organizations that started as a homelessness prevention effort, and then grew into a comprehensive community development effort involving, local businesses, employment, public space, youth engagement, education. The Partnership soon saw that mental health difficulties were preventing many residents from engaging with their programs and taking advantage of public benefits. The community development partnership needed help integrating mental health into their programs. In response, an elective for New York University psychiatry residents was created in which residents helped to develop new mental health initiatives in Brownsville, including home visits to people who could not make clinic appointments, parent education about child development, and learning difficulties for parents whose custody is at risk, a violence prevention campaign for formerly incarcerated youth, and group therapies for people experiencing chronic trauma as a result of pervasive neighborhood and domestic violence.

For the first time in their training, the residents were partnering with community-based organizations and thinking about how mental health could be improved at the level of neighborhoods and institutions as part of broad-based community development. The first five residents who pioneered this elective went on to fellowships in public psychiatry, forensic psychiatry, and to work in the public mental health system. Some of these residents also joined a parallel group called Punishment to Public Health, a New York City–based collaboration of clinicians, public health researchers, formerly incarcerated people, and disenchanted corrections officers. The group hosted workshops and research groups to identify institutional and policy interventions to reduce inequalities in arrests, sentencing, and recidivism. This group is providing psychiatric trainees and professionals an opportunity to act as part of an interdisciplinary collaboration to change public policy and reduce disparities in health.

CONCLUSION

The mental health field has grappled with the complexities of psychiatric disabilities in an increasingly multicultural population. Clinicians need to be aware of and respond to more and more inclusive federal legislation

focusing on nondiscrimination and the national standards for the promotion of CC in hiring practices and health care delivery. Current health care reform efforts are resulting in new models of integrative care that address Donald Berwick's triple aims for the improvement of population health, higher-quality care, and cost reduction. Clinicians can work from the ground up to incorporate CC into the new models providing an opportunity to better fulfill the triple aims. Population health requires knowing and understanding the needs of diverse cultural groups; higher-quality care requires CC of organizations, programs, and clinicians. Cost reductions will follow if care is more appropriately targeted to cultural groups resulting in reduced use of expensive and inappropriate services. Systemic changes as alluded to in organizational CC and structural competency provide an outline for the necessary systemic changes. Clinicians including psychiatrists and other mental health professionals are the critical lynchpin in successful treatment encounters, so that along with systemic change, clinician changes in attitudes and behaviors may act to shift the entire (fragmented) system toward the goals of CC.

REFERENCES

Alegria M, Carson NJ, Goncalves M, Keefe K. Disparities in treatment for substance use disorders and co-occurring disorders for ethnic/racial minority youth. *J Am Acad Child Adolesc Psychiatry*. 2011;50:22–31.

Anderson LM, Scrimshaw SC, Fullilove MT, Fielding JE, Normand J. Culturally competent healthcare systems: a systematic review. *Am J Prev Med*. 2003;24:68–79.

Bromley E, Braslow J. Teaching critical thinking in psychiatric training: a role for the social sciences. *Am J Psychiatry*. 2008;165:1396–1401.

*Cross T, Bazron B, Dennis K, Isaacs M. *Towards a Culturally Competent System of Care*. Vol. I. Washington, DC: Georgetown university child development center. Child and Adolescent Service System Program Technical Assistance Center; 1989.

Gone JP, Trimble JE. American Indian and Alaska native mental health: diverse perspectives on enduring disparities. *Annu Rev Clin Psychol*. 2012;8:131–160.

Gottlieb LM, Tirozzi KJ, Manchanda R, Burns AR, Sandel MT. Moving electronic medical records upstream: incorporating social determinants of health. *Am J Prev Med*. 2014.

Hansen H, Metzl J. Structural Competency in the U.S. healthcare crisis: putting social and policy interventions into clinical practice. *Journal of Bioethical Inquiry*. 2016;13(2):179–183.

Hansen H, Siegel C, Wanderling J, DiRocco D. Buprenorphine and methadone treatment for opioid dependence by income, ethnicity and race of neighborhoods in New York City. *Drug and Alcohol Dependence*. 2016;164:14–21.

Kirmayer LJ, Groleau D, Guzder J, Blake C, Jarvis E. Cultural consultation: a model of mental health service for multicultural societies. *FOCUS: The Journal of Lifelong Learning in Psychiatry*. 2006;4:125–125.

*Kleinman A. Patients and healers in the context of culture: an exploration of the borderland between anthropology, medicine, and psychiatry. Berkeley: University of California Press;

1980.

Marmot M, Friel S, Bell R, Houweling TA, Taylor S. Closing the gap in a generation: health equity through action on the social determinants of health. *Lancet*. 2008;372:1661–1669

*Metzl JM, Hansen H. Structural competency: theorizing a new medical engagement with stigma and inequality. *Soc Sci Med*. 2014;103:126–133.

Mezzich JE, Caracci G, Fabrega H Jr, Kirmayer LJ. Cultural formulation guidelines. *Transcult Psych*. 2009;46:383–405.

Mian Al, Al-Mateen CS, Cerda G. Training child and adolescent psychiatrists to be culturally competent. *Child Adolesc Psychiatric Clin N Am*. 2010;19:815–831.

Miranda J, McGuire T, Williams D, Wang P. Mental health in the context of health disparities. *Am J Psychiatry*. 2008;165:1102–1108.

Qureshi A, Collazos F, Ramos M, Casas M. Cultural competency training in psychiatry. *Eur Psych*. 2008;23:49–58.

Siegel C, Davis-Chambers E, Haugland G, Bank R, Aponte C, McCombs H. Performance measures of cultural competency in mental health organizations. *Adm Policy Ment Health*. 2000;28:91–106.

Siegel C, Haugland G, Chambers ED. Performance measures and their benchmarks for assessing organizational cultural competency in behavioral health care service delivery. *Adm Policy Ment Health*. 2003;31:141–170.

Siegel CE, Haugland G, Laska EM, et al. The Nathan Kline Institute cultural competency assessment scale: psychometrics and implications for disparity reduction. *Adm Policy Ment Health*. 2011;38:120–130.

*Siegel C, Haugland G, Reid-Rose L, Hopper K. Components of cultural competence in three mental health programs. *Psych Serv*. 2011;62:626–631.

Siegel C, Laska E, Wanderling JH, Hernandez J, Levenson R. Racial/ethnic prevalence and diagnostic rates of childhood attention-deficit/hyperactivity disorder in a public mental health system. *Psychiatr Serv*. 2016;67(2):199–205.

Siegel CE, Reid-Rose L, Joseph AM, Hernandez JC, Haugland G. Cultural Activation of Consumers. *Psychiatr Serv*. 2016;67(2):153–155.

Siegel CE, Wanderling J, Haugland G, Laska EM, Case BG. Access to and use of non-inpatient services in New York state among racial-ethnic groups. *Psych Serv*. 2013;64:156–164.

Sodowsky GR, Taffe RC, Gutkin TB, Wise SL. Development of the multicultural counseling inventory: a self-report measure of multicultural competencies. *J Couns Psych*. 1994;41:137.

Sue DW, Arredondo P, McDavis RJ. Multicultural counseling competencies and standards: a call to the profession. *J Couns Dev*. 1992;70:477–486.

*Sue DW, Sue D. *Counseling the Culturally Diverse: Theory and Practice*. John Wiley & Sons; 2012.

Tervalon M, Murray-Garcia J. Cultural humility versus cultural competence: a critical distinction in defining physician training outcomes in multicultural education. *J Health Care Poor Underserved*. 1998;9:117–125.

Tirado MD. *Tools for Monitoring Cultural Competence in Health Care*. San Francisco: Latino Coalition for a Healthy California; 1996.

Tucker CM, Herman KC, Pedersen TR, Higley B, Montrichard M, Ivery P. Cultural sensitivity in

physician-patient relationships: perspectives of an ethnically diverse sample of low-income primary care patients. *Med Care*. 2003;41:859–870.

U.S. Department of Health and Human Services. *HHS Action Plan to Reduce Racial and Ethnic Health Disparities: A Nation Free of Disparities in Health and Health Care*. Washington, DC: US Department of Health and Human Services; 2011.

▲ 7.11 Medical Error

NORA C. OBERFIELD, M.D., AND BENJAMIN J. SADOCK, M.D.

INTRODUCTION

Doctors and the medical systems within which they practice strive for perfection in patient care. Nevertheless, as humans and technology are imperfect, medical errors occur in spite of this dedication to excellence. A landmark 1991 Harvard Medical Practice Study estimated that as many as 1.3 million injuries occurred nationwide to patients receiving hospital care. In 1999, the IOM warned the public that 50,000 Americans may die each year as a result of medical errors. Given that the IOM statistics were based on data from over 30 years ago, in 2013, in the *Journal of Patient Safety*, John T. James gave updated estimates of deaths associated with medical error during hospitalization based on studies published between 2008 and 2011. He estimated that more than 400,000 premature deaths occurred in hospitals per year related to preventable patient harm. That number is far greater than the estimate previously reported by the IOM, suggesting that there continues to be much room for improvement in this area.

Although a number of congressional actions led eventually to legislation requiring mandatory national or state reporting by hospitals of medical errors, reports of such documentation have been scarce. Issues to contend with are the need to protect doctors and hospitals from frivolous lawsuits resulting from public disclosure and failure to maintain confidential information about patients.

The topic of medical error has been complicated by the use of confusing and often conflicting terminology. The IOM separates medical error into two main categories: errors of execution and errors of planning. An error of execution occurs when there has been a "...failure to complete a planned action as intended ..." and an error of planning occurs when there has been "... the use of a wrong plan to achieve an aim." Thus, by definition, medical errors require that there be a component of negligence. When they do occur, there is a desire for attribution, to try to understand why the errors have happened. Did the error occur because the physician departed from an accepted standard of treatment? Was the error avoidable, or the result of poor training? Or, was it unintentional—an accident—and better considered a close call or a borderline case?

It is important to distinguish medical error from adverse events. The latter are related to a particular disease process and are most often predictable or, if not predictable, are known risks. Medical errors usually lead to adverse events and may harm the patient temporarily or permanently, may increase the time for recovery from disease, or may result in patient death. Adverse events may occur in the absence of medical errors, and medical errors may not be associated with an adverse event.

At Johns Hopkins University, the term medical mistake is used instead of medical error and is defined as follows:

A medical mistake is a commission or an omission with potentially negative consequences for the patient that would have been judged wrong by skilled and knowledgeable peers at the time it occurred, independent of whether there were any negative consequences. This definition excludes the natural history of disease that does not respond to treatment and the foreseeable complications of a correctly performed procedure, as well as cases in which there is reasonable disagreement over whether a mistake occurred.

Types of Error in General Medicine

The current literature highlights five main types of medical error: errors of commission, errors of omission, errors of communication, errors of context, and diagnostic errors.

1. An error of commission occurs when something done by a member of the medical team leads to patient harm either because it was the wrong thing to do or because it was the correct thing to do but it was not performed correctly.
2. An error of omission occurs when something that was clearly needed for the patient's healing process was not done.
3. An error of communication occurs when information is not conveyed effectively, either between at least two providers of care, or between a provider of care and the patient.
4. A contextual error occurs when a member of the health care team does not recognize social factors in the patient's life which could create barriers to implementation of effective postdischarge treatment.
5. Diagnostic errors occur when a wrong diagnosis is made, which could have sequelae such as inappropriate or delayed treatment.

Errors may also be categorized as systematic or individual. Systematic errors are caused by flaws in the hospital system or in the transfer of information from one unit of the hospital or from one group of doctors to another. In general, no one person is to blame. An example would be the unavailability or inability to retrieve a patient's medical records or a mismatched blood transfusion. The failure is in the health care delivery system as a whole rather than in the individual.

Individual errors are attributed to a particular physician or to a specific member of the treatment team. Examples include prescribing the wrong

medicine or operating on the wrong leg. The physician bears the responsibility, even if the error is made by a person under his or her supervision. Research at John Hopkins Medical Center involving an examination of 25 years of U.S. malpractice claim payouts from 1986 to 2010 indicated that diagnostic errors were the most common type of medical error. While the predominant focus in recent years has been on reducing medication and surgical errors, the importance of improving the error rate in the diagnostic area is equally important.

Types of Error Specific to Psychiatry

While there is increased discussion of medical error generally, errors occurring in patients with primary psychiatric diagnoses is an area often excluded from the literature. S.W. Cullen and colleagues separated errors on inpatient psychiatry units into four categories: treatment errors, diagnostic errors, preventive errors, and other errors.

1. Treatment errors are defined as "… failures in the formulation or delivery of an appropriate intervention, therapy, or procedure."
2. Diagnostic errors are not only inappropriate psychiatric diagnosis, but also mistaking a physical illness for a psychiatric condition or vice versa. Preventive errors were the most common and have been defined as "any failure to offer indicated prophylactic treatment, as well as inadequate or inappropriate monitoring of an initiated treatment plan."
3. Preventive errors include inadequate or disruption of safety protocols "… such as fall precautions, locked unit doors, break-away bars in closets, or 15-minute checks for patients with suicidal ideations." The "other errors" category included problematic transfers and discharges.

In the Cullen study mentioned above it was found that nonpsychiatric staff members felt ill-prepared to deal with referrals of patients with a history of psychiatric treatment or hospitalization or of patients who were in a state of emotional distress. There was often an inadequate or cursory medical examination in patients who were disruptive or otherwise uncooperative. In order to reduce the error rate among such patients, it will be necessary to decrease stigma about mental illness and educate all health care providers about the complicated presentation of some psychiatric patients with active symptoms such as psychosis. In addition, compliance issues may be a factor in the management of psychiatric patients with medical illness which may cause difficulties implementing safe and appropriate discharge plans.

Causes of Error in General Medicine

As mentioned above, causes of medical error are multifactorial. In general medicine they include medication errors, transitions of care error, errors

stemming from the use EMRs, errors related to the challenges of treating a growing geriatric population, and errors related to job overload, fatigue, and physician burnout among others discussed below. In most instances, medical errors correlate with inexperience. For example, first-year house officers make more errors than do second- and third-year residents. Junior attending physicians make more errors than senior attending physicians. Clearly, the aphorism of experience being the best teacher seems to apply in these instances, but experience or the lack thereof is not the only factor.

Medication. According to statistics from the IOM, medication errors in the United States contribute to 1 out of every 131 outpatient deaths and 1 out of every 854 inpatient deaths. The U.S. Food and Drug Administration (USFDA) reports that over 770,000 patients are injured each year due to medication errors. In France, a 2007 study on adverse events in inpatient healthcare noted that 50 percent of serious adverse events were caused by medication errors. At the Mayo Clinic, medication errors are separated into three categories: medication factors, patient factors, and health care professional factors. Medication factors include medications with names that sound similar or medications with low therapeutic indexes. Patient factors include metabolic challenges related to poor hepatic or renal function, patient cognitive status, and polypharmacy. And health care professional factors include poor handwriting and use of abbreviations in prescriptions and progress notes. Several studies found that interruptions during rounds and pressure to complete rounds were also potential contributors to medication and other errors.

Transitions of Care. It is important to note the role that transitions of care can play in medication error. One study found that nearly 50 percent of patients had at least one medication error when discharged from inpatient units.

Working Hours. Concerns about resident overwork, which may contribute to medical error, generated a new set of duty hour regulations recommended by The Accreditation Council for Graduate Medical Education (ACGME) that took effect in July of 2011. Those recommendations limited the maximum shift length for first-year residents (also known as interns) to 16 hours a day. Although the particulars of duty hours are discussed further in the section *Regulating Working Hours of Interns and Residents,* they are mentioned here in order to examine how this change effected medical error. The change in duty hours was enacted by the ACGME based on research which demonstrated that long duty hours harm both patients and residents. Decreased shift lengths for interns led to a decrease in the incidence of serious medical errors in the intensive care units. Longitudinal studies of duty hour restriction, however, did not show

that house staff used that time to sleep more hours than previously; nor was there a lessened incidence of burnout. Working shorter shifts may have had the unintended consequence of increasing stress by requiring residents to complete the same amount of work but in less total time. Increased work compression has been associated with poorer clinical performance and decreased satisfaction among residents.

HANDOFF ERROR. Shorter shifts have also led to more handoffs (when hospital staff who are going off shift communicate information about their patients to the hospital staff coming on shift), and there is a growing literature indicating an elevated rate of medical errors with more handoffs. Different hospitals perform handoffs in different ways and even individual departments within the same hospital perform handoffs differently. While some hospitals may use a handoff system combining verbal and paper sign out, other hospitals use an electronic document which is continuously updated. But there is room for error in these frequent transactions. For example, one study found that there were 24 unique handoff mnemonics used in one hospital. In addition, many doctors have never been given formal training in handoff communication. With the rise of errors attributed to handoffs, there is a growing call for a need for a systematic algorithm to be followed by all departments in all hospitals. Efforts are currently being implemented in some hospitals across the United States, focusing on training both medical students and residents, to improve handoff communication in order to decrease medical error.

Electronic Medical Records. The use of EMRs has generally been met with a favorable response, particularly in terms of how electronic documentation decreases medical errors related to illegible handwriting in progress notes and prescriptions. However, some physicians are concerned that EMR may lead to an increase in medical error because the EMR has replaced face-to-face or phone communications.

James Bernat was among the first to point out the ethical and quality problems associated with the use of EMR. He pointed out the hazards of

> ... copying and pasting data obtained from other clinicians, authorship ambiguities, inadvertent inclusion of unobtained data in templated notes, misleading history and physical exams, failure to review prepopulated data, inadequate discharge summaries, impairments to patient-physician communication, and the transformation of the purpose of the medical record to billing documentation.

In addition, many clinicians have commented on the fact that EMR is making patient care less about the patient and more about the record keeping. One study found that medical interns spent only 12 percent of their time with patients and 40 percent of their time at computers. Patient encounters have become experiences where the expectation of eye contact between doctor and patient can no longer be guaranteed, with the

physician often having to fill out automated check boxes on the computer screen while speaking with the patient. Paying less direct attention to the patient could lead to medical error by missing pertinent details of the history or missing subtle nonverbal communication signs, which could be diagnostic clues. Physicians should meet the patient before entering information into the computer.

Population Demographics

In 2009, there were 39.6 million people in the United States who were 65 years old or older. Estimates suggest that by 2030, there will be roughly 72.1 million people in this age category. In the United Kingdom, there were 4.8 million people listed as 75 years old or older in 2011, with an expected rise to 8.7 million by 2033. Through improvements in scientific understanding of disease pathology, people are living longer than they did previously, often due to taking medications that slow or halt the progression of several comorbid medical conditions. Given the growing size of elderly populations internationally, in order to reduce medical errors, there must be increased understanding into the types of medical errors that occur in the elderly population. Polypharmacy, taking several medications daily, can be a significant contributor to medical error, given that medications may have negative drug–drug interactions, and increased care must be taken to administer particular medications at specific times of day to avoid such interactions. In one study from the United Kingdom examining 13 long-term residential homes, polypharmacy is evident, with the average resident receiving 7 or more medications. Age-related metabolic changes in how the body responds to medications must also be taken into account, with a need for greater education on renal and hepatic dosing of medications.

Job Overload, Fatigue, and Physician Burnout. Job overload, in which doctors are required to care for too many patients, is a major contributor to medical error. In the hospital setting, interns are assigned too many patients with too few supervisors. In the outpatient setting, too many patients are cared for by too few doctors. Private practitioners are under pressure by managed care companies who often reward doctors financially for productivity, which translates into a high caseload of patients that leads to overwork. Similarly, some managed care and insurance companies reward doctors who order few laboratory tests, thus lowering cost; however, this may result in misdiagnosis, especially in complex cases with atypical presentations. Fatigue and burnout interfere with judgment, and lead to mistakes. While studies report rates of burnout in medical students from 31 to 49.6 percent, among medical residents this percentage increases, with some specialties reporting rates as high as 76 percent (internal medicine). Among practicing physicians, the rates vary by

specialty, with emergency medicine physicians reporting burnout at a rate of 51 percent, while psychiatrists are on the lower end of the spectrum, reporting burnout at 33 percent.

Causes of Error in Psychiatry

As within general medicine, the causes of error in psychiatry are multifactorial. Areas to examine include the specifics of the inpatient psychiatric care setting, medication errors, suicide, and restraint and seclusion practices.

Inpatient Psychiatry Unit Characteristics. One of the most striking ways in which inpatient psychiatry units are distinct from most other inpatient units in hospitals is that they are locked units. While the primary purpose of this locked structure is to maintain the safety of the patients and those around them, the fact that the units are locked has the potential to make emergency treatment logistically more challenging. Additionally, again with safety in mind, most inpatient psychiatry units do not allow for any sharps on the unit (with the exception of very small needles used for blood draws, which are often kept off the unit until needed). This means that if a patient is at a level of medical distress requiring use of IVs, delays in getting the patient to a secure environment, which allows for IV access, are inevitable. These delays may lead to patients not receiving emergent care in the most appropriate time frames, which could lead to an increased error rate in this patient population. Overall, more medication errors occur in hospitalized psychiatric patients than in psychiatric outpatients.

Medication. Most mistakes in psychiatry occur in the area of medication. A wrong prescription may be written, the wrong medication for a disorder may be prescribed, or the wrong dosage of a drug may be administered. A common cause of medication error is poor handwriting by physicians, which is then misread by a pharmacist or nurse. A potential remedy is to insist that all prescriptions be typed and transmitted to pharmacies via computer (electronic prescribing has been increasingly adopted by medical practices throughout the United States). Similar-sounding drug names have also been to blame for medication mistakes, for example, citalopram (Celexa) and celecoxib (Celebrex) and buspirone (BuSpar) and bupropion (Wellbutrin) (Table 7.11–1). Drugs with similar phonemes lend themselves to miscommunication in writing or in speech.

Table 7.11–1.
Examples of Drug Name Confusion

Psychoactive Medication	Confused with
Metadate CD (methylphenidate)	Methadone (Dolophine)

Levoxine[a] (levothyroxine) (thyroid preparation)	Lanoxin (digoxin)
Serzone (nefazodone)	Seroquel (quetiapine)
Lamictal (lamotrigine)	Lamisil (terbinafine) (for nail infection)
Zyprexa (olanzapine)	Zyrtec (cetirizine) (for allergies)
Celexa (citalopram)	Celebrex (celecoxib) (for arthritis)

[a]Name changed to Levoxyl after U.S. Food and Drug Administration report in 1994.

Wrong drug administration, that is, giving a drug to a patient that was not the drug prescribed for that patient, or giving the correct drug but at the wrong dose, is the most common type of medication error. Wrong drug administration occurs more frequently on units with either fewer registered nurses or two or more patients with the same (or similar) name staying on the same unit. Most prescription errors in psychiatry are with antipsychotic medications followed by anxiolytics and antidepressants. Increasingly, studies are showing that the use of computerized medication orders decreases errors related to drugs. There is a growing consensus that the current dispensing of drugs as needed (so-called prn medications) ought to be reexamined. One recent study found an increased risk of error by 15 percent each time a PRN medication was given on medication rounds. As needed directions are often ambiguous and do not always indicate when or how often or at what dosages they are to be used. While doctors are placing the orders, nurses are often on the front lines of decision making as to when and how these medications should be dispensed, and in some instances are not trained sufficiently to thoroughly assess how these medications may interact with other medications the patient is taking. Therefore, in an effort to decrease medication error rates, PRN orders should always be written with clear identification of indication, frequency, and dosage.

Suicide. Unlike most physicians, psychiatrists rarely have to deal with the death of patients. Exceptions may be made for geriatric psychiatrists who deal with the aged or for consultation–liaison psychiatrists who deal with the seriously medically ill; however, when death occurs in those instances, it is often expected and part of the disease progress.

Suicide is unexpected. It is also viewed by many persons as preventable. On the basis of that expectation, family members may sue for malpractice after the event. Suicide is not always preventable, however. Some persons who have a strong desire to kill themselves will do so despite the many advances that have been made in the treatment and management of suicidal patients. When mistakes occur, they are usually the result of the psychiatrist having underestimated the suicidal risk or, if the psychiatrist has appraised it accurately, the nurses and attendants not taking the necessary steps to ensure the patient's safety (e.g., constant observation).

Litigation is more common for suicides that occur in a hospital than for

those that occur in an outpatient setting, presumably because one-to-one coverage is the standard of care in the actively suicidal hospitalized patient. Of all the suicides that occur in the United States approximately 5 percent occur during hospitalization. The most common methods of suicide in the hospitalized patient is jumping or hanging. In most cases, safety protocols such as one-to-one monitoring or 15-minute checks were not followed assiduously.

Courts have not penalized psychiatrists for honest errors of judgment in clinical practice. One court stated: "The accurate predication of dangerous behavior, and particularly suicide and homicide, are almost never possible. Thus, an error of prediction or even of judgment does not necessarily establish negligence." In spite of this, however, psychiatrists are expected to evaluate patients properly for suicide risk and to establish a rational treatment plan to ensure that the risk of a patient successfully killing him- or herself is minimized as much as possible.

The effect of having a patient commit suicide can have profound effects on the treating psychiatrist. There is an initial sense of disbelief and shock upon being informed of the event often followed by feelings of guilt. Common questions asked are: Did I overlook anything? How could this have been prevented? What should I have done differently? Some psychiatrists are so traumatized that they give up practicing altogether or never again work with patients who are suicidal. More commonly, however, they seek to talk about the experience with other clinicians, especially clinicians who have gone through a similar experience of losing a patient to suicide. This process can be very salutary, allowing the psychiatrist to realize that he or she is not alone, that others have gone through the same experience, and as in all branches of medicine, including psychiatry, not all illnesses are curable and some suicides, even when care and treatment is exceptional, can neither be predicted nor prevented.

Restraint and Seclusion

Errors in the area of restraint result from not using these devices correctly; this results in physical injury or in some cases death by strangulation or cardiac arrest. Patients in seclusion rooms may injure themselves by banging their heads into the walls or exhaust themselves as a result of agitated behavior that can lead to dehydration and death. In both situations, constant vigilance is necessary. In 2006, the Centers for Medicare and Medicaid Services (CMS) mandated more training for health care workers in the area of restraint and seclusion, in an effort to promote safer practices and greater acknowledgement of patients' rights. They later released a revised set of guidelines related to restraint and seclusion which included "face-to-face examination by a physician, nurse, or physician's assistant within one hour of placement in restraints or seclusion." Some hospitals have initiated policies in which patients in restraints and

seclusion must be on one-to-one constant observation with vitals checks every hour, with physician review for necessity of restraints every 4 hours. Seclusion rooms must be designed to be environmentally protective, for example, padded walls, ventilation. In New York State nursing homes, efforts have been made to reduce the use of physical restraints, with a decrease from 6.14 percent in 2011 to 1.86 percent in 2014.

PREVENTING MEDICAL ERROR

Practice Guidelines

Practice guidelines have been developed by the APA for a variety of disorders. These include delirium, dementia, substance use disorder, depressive disorder, panic disorder, borderline personality disorder, suicide, and schizophrenia. Guidelines for other psychiatric disorders will be forthcoming. According to the APA, the term *practice guideline* refers to a "set of patient care strategies developed to assist physicians in clinical decisions." They are careful to point out that guidelines are not meant to be standards of care:

The APA practice guidelines are not intended to be construed or to serve as a standard of medical care. Standards of medical care are determined on the basis of all clinical data available for an individual case and are subject to change as scientific knowledge and technology advance and patterns evolve. These parameters of practice should be considered guidelines only. Adherence to them will not ensure a successful outcome in every case, nor should they be construed as including all proper methods of care or excluding other acceptable methods of care aimed at the same results. The ultimate judgment regarding a particular clinical procedure or treatment plan must be made by the psychiatrist in light of the clinical data presented by the patient and the diagnosis and treatment options available.

Although few studies have been done, common sense indicates that guidelines will improve quality of care and diminish medical errors in psychiatry. They are works in progress, however. As new methods of treatment develop, they have to be revised and updated, lest they become fossilized and irrelevant. Other risks are their possible misuse by third parties such as managed care companies and health maintenance organizations (HMOs). Many guidelines are prepared by third-party payers and may be compromised by conflicts of interest of cost versus quality of care. Some HMOs pay bonuses to their doctors based on high scores on how closely they adhere to a set of guidelines in the treatment of a specific disorder. This may influence psychiatrists in the choice of therapy or the length of the patient's hospital stay if they are not in keeping with the set of guidelines being used. The APA practice guidelines are not the only ones that relate to psychiatry. There is a web-based resource in a National Guideline Clearinghouse (NGC) that lists more than 500 practice

guidelines with some relevance to psychiatry (http://www.guideline.gov).

There is no indication that there has been an increase in malpractice suits against psychiatrists since their introduction; however, published guidelines are often brought forward in litigation as standards of care. Psychiatrists and other physicians should not be intimidated by this practice. Psychiatrists may legally and ethically deviate from guidelines; however, the decision to do so and its rationale should be documented in the chart.

As mentioned previously, the APA guidelines do not include all proper methods of care for a particular disorder, nor do they allow for ambiguity, controversy, or dissent. For this, the practitioner must turn to a textbook, such as the *Comprehensive Textbook of Psychiatry,* in which such differences are discussed.

Electronic Medical Record, Medication Reconciliation, and Electronic Prescribing

The EMR is fast becoming an integral part of most hospital and outpatient clinic settings, with several notable advantages over paper records. According to Dr. Michael Houston these advantages include, "... the ease with which information can be shared among providers, thereby improving the efficiency of health care delivery by reducing redundancy, simplifying billing, and improving quality and safety." EMR became more commonplace following the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act, part of the American Recovery and Reinvestment Act of 2009, as well as the 2010 Patient Protection and Affordable Care Act, national legislation which included bonuses for medical sites participating in EMR.

The advent of EMR led to improvements in the medication reconciliation, defined by the Institute for Healthcare Improvement as "a process of identifying the most accurate list of all medications a patient is taking—including name, dosage, frequency, and route—and using this list to provide correct medications for patients anywhere within the health care system." One study polling patients who were about to be discharged from inpatient hospitalizations noted that 27.9 percent of the patients were unable to list their discharge medications, and even a smaller percentage had an understanding of what their medications were for. Acknowledging that medication-related errors are frequent contributors to medical error, the Joint Commission declared medication reconciliation a 2005 National Patient Safety Goal. Stephanie Mueller and colleagues have found medication reconciliation to work most effectively when two particular conditions are met: close communication between physicians and pharmacists, and ensuring that patients at higher risk of medication error are more closely monitored. This latter category includes the elderly, individuals taking multiple medications, and individuals with multiple medical comorbidities.

In an effort to eliminate medication errors occurring due to illegible handwriting, many medical centers throughout the United States have increasingly been engaging in the use of electronic prescribing (frequently referred to as e-prescribing). This practice allows prescriptions to be electronically transmitted directly to pharmacies. New York State currently requires all prescribers to participate in electronic prescribing. The start date had been postponed, however, likely related to a recognition that—in order to avoid errors—inpatient and outpatient medical care sites needed more time to integrate electronic prescribing programs into their current hospital systems. There appears to be a growing consensus among physicians that while electronic records and prescribing will contribute to a decrease in medical error, the pace of the change may need to occur more slowly than current legislation allows for. Issues arising related to the introduction of new software and necessary training for staff often take more time to reconcile than expected.

Blame and Medical Error Reporting Structures

The Royal Australian and New Zealand College of Psychiatrists raised an interesting argument when they suggested that for improvements in medical error to occur, modern medical culture must undergo a significant frame shift. In particular, they recommended moving away from a model of individual blame toward one of acknowledging systems failures. They found that "a 'no blame' culture is essential to voluntary reporting, dispassionate analysis and correction of systems failures." Although blame is diffused in this model, accountability is not. Every member of the medical community has a responsibility to their colleagues and their patients to take part in telling the story of what went wrong, as the first step in determining how to ensure this error is not repeated. Assurances of confidentiality and ease of reporting are the most pertinent factors leading individuals to freely report errors. In one study it was found that clinicians and other members of hospital staff were most likely to tell a member of the medical error report team about an error in person (53 percent), followed by phoning in the error (35 percent), using the online data reporting system (6 percent), submitting a written form (4 percent), and lastly, by e-mail (2 percent).

Regulating Working Hours of Interns and Residents and Improving Handoffs

Teaching hospitals often rely on interns and residents to perform services such as phlebotomies and intravenous (IV) therapy and to serve as messengers and transporters, tasks that are more appropriately performed by ancillary personnel. In addition, house staff are often required to work long hours, and sleep deprivation can impair their judgment and clinical skills. The average resident works more than 80 hours per week, with a

shift of 30 hours every third or fourth night. As mentioned above, a limit on the number of hours interns and residents may work was set forth by the CMS. Their work rules were (1) residents were limited to no more than 12 consecutive hours per assignment in emergency services, (2) residents could not work more than 80 hours per week over a 4-week period and could not be scheduled to work more than 24 consecutive hours, and (3) scheduled rotations had to be separated by not less than 8 hours of nonworking time, provided for each week. Nonworking time is time away from training and patient care activities. In 2000, 8 percent of residency training programs nationwide were cited by the Accreditation Council for Graduate Medical Education for violating the CMS work hour requirements.

In response to the death of a patient at a New York hospital in 1989, the New York state legislature enacted the Bell Commission Report, which implied that resident work hours contributed in part to that accidental death. The state responded by establishing rules for resident work hours, including 80-hour week limits and 24-hour consecutive call limits.

The impact of these limits on the continuity and quality of care remains to be seen. Legitimate concerns have been raised about shorter shifts leading to more handoffs, which could lead to more opportunities for error.

The Joint Commission has reported that, "Communication and handoff failures are among the root causes in nearly two-thirds of 'sentinel events,' which are serious, often fatal, preventable adverse events in hospitals." Acknowledging the need to address handoff failures with a greater scope, the ACGME now mandates that all "...residency training programs provide formal instruction about patient handoffs and that faculty monitor handoffs and ensure adequate handoff skills through direct observation." Unfortunately, despite this mandate, many residency training programs still do not consistently train their residents in how to perform proper handoffs. In one recent study, roughly half of the residents polled believed that critical patient information was lost during resident shift changes. From June 2010 through February 2014, a comprehensive handoff curriculum, known by the mnemonic I-PASS (Illness severity, Patient summary, Action list, Situation awareness and contingency plans, and Synthesis by receiver) was piloted in nine different pediatric residency programs throughout the United States. In developing this curriculum, it became evident that the key ingredients for an effective handoff include: dedicated time for verbal exchange in a HIPAA compliant location, the active use of templates for accessing and recording information, priority given to discussing ill patients, and highlighting action items. Use of the I-PASS curriculum was associated with a 23 percent relative reduction in the medical error rates across all nine study sites, as well as a 30 percent relative reduction in the rate of preventable adverse events. Importantly, the study found that the use of a comprehensive handoff curriculum did

not elongate the handoff process in any way, nor did it cause residents to spend less time with patients. As this program has only been used in pediatric residency programs, it remains to be seen if the results of the study can be generalized to other medical disciplines.

In an effort to eliminate mistakes that might occur related to medical student involvement in patient care, the CMS defined medical student responsibilities as follows: (1) Students can take histories and perform physical examinations with the approval of the patient's attending physicians; (2) they may write in the patient's chart, but all entries must be countersigned by the attending physicians; and (3) medical or surgical procedures may be performed if they are under direct inpatient supervision.

Many medical educators believe that the current CMS rules do not go far enough and that interns and residents are not used properly by many teaching hospitals. It is not uncommon for a resident in cardiac surgery to assist at an operation for 14 hours and then to stay on duty for an additional 10 hours. Similarly, a pediatric resident may be in the emergency room for 24 hours without sleep. Even though their hours fall within the CMS guidelines, they are not conducive to the resident's education in view of the inevitable fatigue when a person is without sleep for 24 hours. For residents with families, especially those who are parents, current work schedules disrupt relationship harmony and child rearing. This added stress interferes with optimal functioning. It has been suggested that chronic sleep deprivation may put residents at increased risk for motor vehicle crashes, medication errors, and decreased mental clarity. There are no easy answers to these problems, but the current situation clearly requires resolution.

Emergency physicians commonly experience sleep deprivation because of the need to work shifts during evening and late night hours. The negative effects of this problem are compounded by job stress and traditional methods of scheduling work shifts. Sleep deprivation may be reduced by schedules designed to lessen interference with normal sleep patterns and circadian rhythms. Pharmacological treatments for sleep deprivation exist in the form of alertness-enhancing agents such as caffeine. Sleep-promoting agents may also help treat the problem by helping physicians to sleep during daytime hours. Minimizing sleep deprivation may help prevent job burnout and prolong the length of an emergency physician's career.

Coping with Medical Errors and Adverse Events

Doctors have a long tradition of examining adverse events and of learning from their mistakes. The case presentation (also called the *clinicopathologic conference* [CPC] or morbidity and mortality conference [MMC]) is a group exercise attended by clinicians, researchers, educators,

and students who examine all aspects of patient diagnosis, treatment, and outcome. Ideally, CPCs/MMCs do not attempt to cast blame on a practitioner; however, if negligence has occurred, then it becomes evident in the open discussion of those in attendance. A 2012 study examining how MMCs are used to improve the quality of medical care found that a lack of structure in the approach to MMCs frequently led to few clear action plans with definitive timelines. While CPCs/MMCs also could be a useful time to address the emotional impact of having made a medical error, there appears to be less comfort spending time on this topic. As the perception of having made a medical error can lead residents to experience significant personal distress and decreased empathy, it is perhaps time to incorporate elements of self-care and coping strategies as part of CPCs/MMCs. In addition, most hospitals have risk management teams whose duty is to identify potential or actual instances of malpractice by physicians in the facility and to approach the patient and to discuss the situation openly and honestly. Also, offices of quality assurance function to assess and to prevent risks to patients in a continuous effort to eliminate adverse effects as described in Table 7.11–2.

Ethics of Disclosure. There is an ethical duty on physicians to disclose medical errors, accidents, injury, or bad results to their patients who have an absolute right to understand what occurred during their course of treatment. An explanation is mandatory, and an apology is appropriate if an error was made. The American Academy of Pediatrics believes, "An atmosphere of openness and honesty leads to a culture of quality and safety." If, as a result of disclosure, legal proceedings ensue, then the physician has the opportunity to offer an explanation—but appropriate compensation for harm suffered as a result of malpractice is a patient's legal right.

Table 7.11–2.
Medical Documentation Requirements for Patient Adverse Events

These guidelines apply to instances of *adverse patient events,* which are defined as follows: "*A significant adverse patient event is an unplanned or unexpected occurrence that results in patient death, harms the patient, temporarily or permanently, or significantly increases the planned length of hospitalization.*" These guidelines are applicable, however, to all medical documentation and should be universally followed.

It is most important that the responsible attending physician (or covering attending) write a note as soon as possible after an adverse event, certainly within 24 hours. This note should clearly document his or her active participation in the care of the patient. Subsequently, an attending note must be written daily, for as long as appropriate to the patient's status. This is a supplement to the standard Medical Board documentation requirement, especially crucial in the first few days after an adverse event occurs.

All notes must be dated, timed, and signed.

The purpose of medical documentation is to promote patient care and safety, not to cast blame on an individual for an adverse patient event. When describing a patient adverse event, extreme care must

be exercised to avoid implied blame for such an event by another caregiver. The note should clearly document the attending physician's participation in the care of the patient, and the discussion should be an objective analysis of the events that transpired, an assessment of the patient status, and a plan for subsequent management. An open disclosure approach to adverse event documentation must be followed.

Courtesy of New York University Medical Center, New York, NY.

Patient–Doctor Relationship. A good patient–doctor relationship is characterized by a sense of trust, respect, and honesty between the two parties. The better the patient–doctor relationship, the less chance for misunderstanding leading to litigation. Studies have shown that, when a medical error or adverse event occurs within the context of a good patient–doctor relationship, litigation is rare. In its absence, however, litigation is likely in as much as 70 percent of cases of medical error, especially if a major mistake was made, regardless of whether or not death ensues. In surveys of patients, almost 100 percent desire that doctors report and discuss medical errors with them. Acknowledging medical error, minor or major, may actually reduce the risk of malpractice action.

Therapeutic Privilege. Therapeutic privilege is the physician's right to withhold information from a patient, if, in his or her opinion, such disclosure would cause irreparable harm to the patient. Some patients do not wish to know the nature of their illness and that right needs to be respected. Others may be so prone to anxiety or panic that even a hint of medical error (especially if no adverse outcome occurred) would be potentially dangerous to the patient's overall sense of well-being.

Duty to Disclose. There is a patient–consumer movement supported by some members of the legal profession that would require physicians by law to report medical errors. As mentioned previously, disclosure is currently an ethical, not a legal, obligation. Proponents of duty to disclose argue that self-policing by the medical profession is inefficient, and a legal duty to disclose malpractice would provide a remedy. Proponents of such laws would also require that they apply not only to the person committing the error, but also to those who witness the error, for example, nurses and members of the treatment team.

Opponents of the duty to disclose law argue that it would harm the patient–doctor relationship and, in particular, weaken the trust and teamwork among health care professionals who would be required to be informants on one another. Adopting a rule of disclosure also suggests to the public that doctors, as a group, are not to be trusted. Opponents of the law also note that quality assurance and risk management programs that bring negligent acts to light are already in place and that further enforcement is unnecessary. Most experts take the position that physicians are the ones who ultimately must decide to disclose errors based on their

professional and ethical beliefs.

Barriers to Disclosure. Some suggest that, rather than focusing on whether disclosure should be an ethical or legal obligation, the greater question to examine is why is there often a fear of speaking up when medical error has been committed. A 2013 study out of Minnesota reports, "Medical students and residents may be concerned about retaliation in the form of poor grades and evaluations. Confronting an attending physician can be intimidating, and…the 'hidden curriculum' (a set of influences that function at the level of the organizational structure and culture) may encourage…students and residents to remain silent." As stated earlier, a medical culture that embraces nonjudgmental systems of error reporting may be critical to decreasing medical errors from occurring in the first place.

Role of Apology. It is becoming standard practice for physicians to apologize to patients when disclosing unanticipated outcomes that fall within the category of medical error. For an apology to be successful, however, there must be a real sense of regret communicated to the patient or family. It is not sufficient to say "I am sorry" pro forma. There must be an empathic communication of concern and personal responsibility. Apologies have generally not been admitted in evidence in malpractice suits because it would have a chilling effect on the entire process. Juries could interpret an apology as admission of guilt and be less objective about facts in the case that might otherwise be exculpatory.

Regulation of Privacy, Delivery, and Quality of Health Information Distribution in the Age of Electronic Medical Records

Health Insurance Portability and Accountability Act. The HIPAA was passed in 1996 to address the medical delivery system's mounting complexity and its rising dependence on electronic communication. The act orders that the federal Department of HHS develop rules protecting the transmission and confidentiality of patient information, and all units under HIPAA must comply with such rules.

Two rules were finalized in February, 2003: the Transaction Rule and the Privacy Rule. The Transaction Rule facilitates transferring health information effectively and efficiently by means of regulations created by the HHS that established a uniform set of formats, code sets, and data requirements. The Privacy Rule, administered by the Office of Civil Rights (OCR) at HHS, protects the confidentiality of patient information. This means that a patient's medical information belongs to the patient and that the patient has the right to access it, with the exception of psychotherapy notes, which are deemed as property of the psychotherapist who wrote them.

In 2002, the Transaction Rule was implemented, meaning that all entities covered under HIPAA must comply. The health care information involved and code sets are listed in Table 7.11–3. In 2003, the Privacy Rule was executed. Under this rule, patients have new statutory rights (Table 7.11–4).

Table 7.11–3.
Transaction Rule Code Sets

Health care information: The Transaction Rule defines standards and establishes code sets and forms to be used for electronic transactions that involve the following health care information:
Claims or equivalent encounter information
Eligibility inquiries
Referral certification and authorization
Claims status inquiries
Enrollment and disenrollment information
Payment and remittance advice
Health plan premium payments
Coordination of benefits
Code sets: Under the Transaction Rule, the following code sets are required for filing claims with Medicare:
Procedure codes
American Medical Association Current Procedural Terminology codes
Healthcare Common Procedure Coding System codes
Diagnosis codes
International Classification of Disease codes
Drugs and biological
National Drug Codes
Dental codes
Code on dental procedures
Nomenclature for dental services

Adapted from Jaffe E. HIPAA Basics for psychiatrists. *Psychiatr Pract Manag Care.* 2002;8:15.

Under the Privacy Rule, there are certain guidelines by which every practice must abide:

1. Every practice must establish written privacy procedures. These include administrative, physical, and technical safeguards that establish who has access to the patient's information, how this information is used within the facility, and when the information will and will not be disclosed to others.
2. Every practice must take steps to make sure that its business associates protect the privacy of medical records and other health information.
3. Every practice must train employees to comply with the rule.
4. Every practice must have a designated person to serve as a privacy officer. If it is an individual practice or private practice, this person can be the physician.

5. Every practice must establish complaint procedures for patients who wish to ask or to complain about the privacy of their records.

Table 7.11–4.
Patient's Rights under the Privacy Rule

Physicians must give the patient a written notice of his or her privacy rights, the privacy policies of the practice, and how patient information is used, kept, and disclosed. A written acknowledgment should be taken from the patient verifying that they have seen such notice.
Patients should be able to obtain copies of their medical records and to request revisions to those records within a stated amount of time (usually 30 days). Patients do not have the right to see psychotherapy notes.
Physicians must provide the patient with a history of most disclosures of their medical history on request. There are some exceptions. The APA Committee on Confidentiality has developed a model document for this requirement.
Physicians must obtain authorization from the patient for disclosure of information other than for treatment, payment, and health care operations (these three are considered to be routine uses, for which consent is not required). The APA Committee on Confidentiality has developed a model document for this requirement.
Patients may request another means of communication of their protected information (i.e., request that the physician contact them at a specific phone number or address).
Physicians cannot generally limit treatment to obtaining patient authorization for disclosure of their information for nonroutine uses.
Patients have the right to complain about Privacy Rule violations to the physician, their health plan, or to the Secretary of HHS.

APA, American Psychiatric Association; HHS, Department of Health and Human Services.

Adapted from Jaffe E. HIPAA Basics for psychiatrists. *Psychiatr Pract Manag Care.* 2002;8:15.

The OCR at HHS is responsible for making sure that the Privacy Rule is enforced. This enforcement is conducted by the government through a *complaint-driven* system in which the OCR responds to complaints made by patients concerning confidentiality violations or denials of access to records, all of which are covered under HIPAA. In such cases, OCR may follow up and audit compliance.

In 2009, the Health Information Technology for Economic and Clinical Health (HITECH) Act was signed into law as a subsection of the American Recovery and Reinvestment Act. The purpose of the HITECH Act was to "... promote the adoption and meaningful use of health information technology," through incentive payments from Medicare and Medicaid for technology use. Examples of "meaningful use" include e-prescribing, completing electronic medication reconciliation, and updating electronic records on vaccination status. This act also led to strengthened breach notification requirements. In 2013, the final HIPAA Omnibus Rule was enacted, which increased penalties for HIPAA noncompliance to as much as $1.5 million per violation.

The APA has created a HIPAA educational packet, which can be obtained on the APA website (http://www.psych.org/). On the site, there

are also recommendations for enabling physicians to comply with HIPAA.

In 2010, the IOM formed the Committee on Patient Safety and Health Information Technology to gain a greater understanding of the impact of health information technology (HIT) on patient safety. In 2012, this committee submitted a report, *Health IT and Patient Safety: Building Safer Systems for Better Care,* in which a disturbing trend was noted, which may negatively impact this goal. Specifically, it appears that some vendors—contracted by medical facilities to create EMR platforms—may be restricting the release of information regarding safety issues by citing contractual confidentiality clauses. This may, however, be leading to difficulties in determining where the problems currently are in HIT, and therefore, limits room for improvement.

Hospital Standards and Performance. A group of agencies, such as the Joint Commission of Accreditation of Healthcare Organizations (JCAHO) and the Liaison Committee on Medical Education (LCME), influence the standards of hospital care and performance. In addition, hospitals must comply with government regulations (city and state health rules). The JCAHO inspects hospitals at least once every 39 months and is also responsible for determining the requirements for hospital accreditation. Hospital reimbursements from Medicare and Medicaid are contingent on meeting these standards, but the accreditation is done on a voluntary basis. The LCME and the Liaison Committee on Graduate Education are charged with accrediting medical schools and residency training programs, respectively. The two accrediting committees review education and training programs every 4 years; the procedure is voluntary.

The JCAHO is also involved in preventing medical errors and developed guidelines to help surgical teams standardize safety systems. The guidelines, called "The Universal Protocol for Preventing Wrong Site, Wrong Procedure, Wrong Person Surgery," address the verification of documents, marking operative sites, and other issues related to patient safety.

Table 7.11–5.
Confusing Medical Notations

Prohibited Abbreviation	Potential Problem	Preferred Term
U (for unit)	Read as a zero (0) or a four (4), causing a 10-fold overdose or greater (4 U seen as "40" or 4 u seen as "44").	Write "unit."
IU (for international unit)	Misread as "IV" (intravenous) or 10.	Write "international unit."
Q.D. and Q.O.D. (Latin for once daily and every	Mistaken for each other. Also, the period after the Q can be mistaken for an I and the O can be mistaken for	Write "daily" or "every other day."

other day)	an I.	
Trailing zero (X.0 mg) and lack of leading zero (.X mg)	Decimal point is missed.	Never write a zero by itself after a decimal point (X mg) and always use a zero before a decimal point (0.X mg).
MS, MSO_4, and $MgSO_4$	Confused for one another. Can mean morphine sulfate or magnesium sulfate.	Write "morphine sulfate" or "magnesium sulfate."

The JCAHO issued a list of dangerous, unapproved abbreviations, acronyms, and symbols that accredited organizations must discontinue using as part of their National Patient Safety Goal requirements. The list of unapproved abbreviations pertains to all handwritten and electronic, patient-specific communication (Table 7.11–5).

The Institute for Safe Medication Practices (ISMP) released an updated List of Confused Drug Names in April 2014. To avoid medication errors, ISMP recommends strategies such as: "... using both the brand and generic names on prescriptions and labels; including the purpose of the medication on prescriptions; configuring computer selection screens to prevent look-alike names from appearing consecutively; and changing the appearance of look-alike product names to draw attention to their dissimilarities." The ISMP, along with the FDA, endorse the use of "tall man" (mixed case) letters as visual aids to avoid medication error (Table 7.11–6). The FDA has begun working with drug companies to select names for new products that differ significantly from those already on the market.

Hospital Utilization Review. This in-house evaluation process was created to make sure that institutions provide efficient, quality health care that meets patients' needs. The members of the utilization review committee consist of hospital administrators, physicians, and nurses. The committee reviews each patient's chart within a specified number of days of admission. The appropriateness of the admission, treatment strategies, and the length of the hospital stay are reviewed to facilitate the patient's discharge. Through this process, the utilization review committee determines whether a particular admission was really indicated and whether the hospital stay was longer than necessary. A hospital must conduct utilization reviews to be eligible for JCAHO accreditation.

Professional Standards Review Organization. The Professional Standards Review Organization (PSRO) was set up by the federal government to review and to monitor care received by patients whose care is paid for with government funds. PSROs are made up of physicians elected by local medical societies that have been established by local medical associations and serve several functions: They attempt to ensure high-quality care, to control costs, to determine maximum lengths of stay

by patients in hospitals, to conduct utilization reviews, and to censure physicians who do not adhere to established guidelines. A PSRO may conduct a medical audit to evaluate the quality of care retrospectively by carefully examining charts.

Table 7.11–6.
List of Confused Drug Names Illustrating Tall Man Lettering Practices[a]

Drug Name	Confused Drug Name
MetaDATE CD (methylphenidate)	methaDONE (Dolophine)
LEVOTHYroxine	LANoxin (digoxin)
Serzone (nefazodone)	SEROquel (quetiapine)
LaMICtal (lamotrigine)	LamISIL (terbinafine) (for nail infection)
ZyPREXA (olanzapine)	ZyrTEC (cetirizine) (for allergies)
CeleXA (citalopram)	CeleBREX (celecoxib) (for arthritis)

[a]Note how the same drugs from Table 7.11–1 are listed, but using tall man lettering format that promotes more accurate drug identification.

Adapted from ISMP's List of Confused Drug Names, updated April 2014.

Peer Review Organization. In the early 1980s, the peer review organization (PRO) replaced PSROs as the federal review organization for hospitals receiving Medicare funds. To promote compliance with federal guidelines for health and hospital care, PROs conduct independent utilization reviews and quality-of-care studies, validate diagnose-related group (DRG) assignments, and review hospital admissions and readmissions.

Federally mandated and funded, the PROs had greater authority than the PSROs. PROs were able to impose sanctions on hospitals for inadequate care; they could even recommend the termination of federal funding to hospitals that consistently violated federal standards. In addition, PROs could adjust or refuse payment for health services that they considered unnecessary.

Quality Improvement Organization. In May of 2002, Quality Improvement Organizations (QIOs) replaced PROs. Funded by the federal government through the CMS, the QIOs are state-based organizations with three main aims: to provide safer and more effective care for patients, to improve the overall health of patients, and to make healthcare more affordable. IPRO, New York's QIO, recently released their Quality Improvement Report for New York State focused on the years 2011–2014. One goal which was achieved and highlighted in the report related to IPRO's success in helping medical practices integrate HIT for meaningful use and improved care coordination.

Centers for Medicare and Medicaid Services. In an attempt to improve

patient safety and reduce complications, Medicare stopped paying for some hospital-acquired conditions. These include errors such as an object left in the patient during surgery, blood incompatibility, decubiti, and air embolism among others. In some states, such as New York, Medicaid has stopped paying for "never events," once described by the former New York State Health Commissioner, Richard Gaines, as "… a list of things that, frankly, should never happen." The New York State never events list includes medication errors, patient disability from use of restraints or bedrails, and surgery performed on the wrong body part. CMS believes that financial disincentives will influence hospitals to provide more aggressive quality review and be better able to address issues effectively. Critics of the program claim that some errors are really adverse events citing bed sores as an example of a condition that cannot always be avoided, even with the most highly skilled nursing care.

REFERENCES

Allen S, Caton C, Cluver J, Mainous AG 3rd, Clyburn B. Targeting improvements in patient safety at a large academic center: An institutional handoff curriculum for graduate medical education. *Acad Med.* 2014;89(10):1366–1369.

Aspden P, Wolcott J, Bootman JL, Cronenwett LR. Institute of Medicine Committee on identifying and preventing medical errors. *Preventing Medical Errors.* Washington, DC: National Academies Press; 2007.

Bari A, Khan RA, Rathore AW. Medical errors; causes, consequences, emotional response and resulting behavioral change. *Pak J Med Sci.* 2016;32(3):523–528.

Bal G, Sellier E, Tchouda SD, François P. Improving quality of care and patient safety through morbidity and mortality conferences. *J Healthcare Qual.* 2014;36(1):29–36.

*Bell SK, Mann KJ, Truog R, Lantos JD. Should we tell parents when we've made an error? *Pediatrics.* 2015;135(1):159–163.

*Bernat JL. Ethical and quality pitfall in electronic health records. *Neurology.* 2013;80(11):1057–1061.

Cottney A, Innes J. Medication-administration errors in an urban mental health hospital: a direct observation study. *Intern J Ment Heal Nurs.* 2015;24(1):65–74.

*Cullen SW, Nath SB, Marcus SC. Toward understanding errors in inpatient psychiatry: a qualitative inquiry. *Psychiatr Q.* 2010;81:197–205.

DeRienzo CM, Frush K, Barfield ME, Gopwani PR, Griffith BC, Jiang X, et al. Handoffs in the era of duty hours reform: A focused review and strategy to address changes in the accreditation council for graduate medical education common program requirements. *Acad Med.* 2012;87(4):403–410.

Heaton HA, Campbell RL, Thompson KM, Sadosty AT. In support of the medical apology: The nonlegal arguments. *J Emer Med.* 2016;51(5):605–609.

Houston M. The psychiatric medical record, HIPAA, and the use of electronic medical records. *Child Adol Psych Clin N Am.* 2010;19:107–114.

Institute of Medicine. *Health IT and Patient Safety: Building Safer Systems for Better Care.*

Washington, DC: The National Academies Press; 2012.

*James JT. A new, evidence-based estimate of patient harms associated with hospital care. *J Patient Saf*. 2013;9(3):122–128.

Kass JS, Rose RV. Medical malpractice reform—historical approaches, alternative models, and communication and resolution programs. *AMA J Ethics*. 2016;18(3):299–310.

Kavoussi SC, Huang JJ, Tsai JC, Kempton JE. HIPAA for physicians in the information age. *Connecticut Medicine*. 2014;78(7):425–427.

Klein CA. Cloudy confidentiality: Clinical and legal implications of cloud computing in health care. *J Am Acad Psych Law*. 2011;39:571–578.

Kohn LT, Corrigan J, Donaldson MS. Institute of Medicine Committee on Quality of Health Care in America. *To Err is Human: Building a Safer Health System*. Washington, DC: National Academies Press; 2000.

Morgenthaler T, Harper CM. Getting rid of "never events" in hospitals. *Harvard Bus Rev*. Oct 2015.

Mueller SK, Sponsler KC, Kripalani S, Schnipper JL. Hospital-based medication reconciliation practices: a systematic review. *Arch Intern Med*. 2012;172(14):1957–1969.

Murray TL, Calhoun M, Philipsen NC. Privacy, confidentiality, HIPAA, and HITECH: implications for the health care practitioner. *J Nurs Prac*. 2011;7(9):747–752.

Norman ID, Aikins MK, Binka FN. Ethics and electronic health information technology: challenges for evidence-based medicine and the physician-patient relationship. *Ghana Med J*. 2011;45(3):115–124.

Ober KP, Applegate WB. The electronic health record: are we the tools of our tools? *The Pharos, Alpha Omega Alpha Honor Medical Society*. 2015: Winter ed. 9–14.

O'Connor N, Kotze B, Wright M. Blame and accountability 1: understanding blame and blame pathologies.*Australas Psychiatry*. 2011;19:113.

Procyshyn RM, Barr AM, Brickell T, Honer WG. Medication errors in psychiatry: a comprehensive review. *CNS Drugs*. 2010;24(7):595–609.

Recupero PR, Price M, Garvey KA, DAly B, Xavier SL. Restraint and seclusion in psychiatric treatment settings: regulation, case law, and risk management. *J Am Acad Psych Law*. 2011;39(4):465–476.

Romani M, Ashkar K. Burnout among physicians. *Libyan J Med*. 2014:1–6.

Sen S, Kranzler HR, Didwania AK, et al. Effects of the 2011 duty hour reforms on interns and their patients: a prospective longitudinal cohort study. *JAMA Intern Med*. 2013;173(8):657–662.

Soerensen AL, Lisby M, Nielsen LP, Poulsen BK, Mainz J. The medication process in a psychiatric hospital: are errors a potential threat to patient safety? *Risk Management and Healthcare Policy*. 2013;6:23–31.

Staender S. Incident reporting in anaesthesiology. *Best Prac Res Clin Anaes*. 2011;25:207–214.

Starmer AJ, O'Toole JK, Rosenbluth G, et al. Development, implementation, and dissemination of the I-PASS handoff curriculum: A multisite educational intervention to improve patient handoffs. *Acad Med*. 2014;89(6):876–884.

Starmer AJ, Spector ND, Srivastava R, et al. Changes in Medical Errors after Implementation of a Handoff Program. *NEJM*. 2014;371(19):1803–1812.

Sternberg S. Diagnosis burnout. *U.S. News & World Report*. Sept 2016.

Szczepura A, Wild D, Nelson S. Medication administration errors for older people in long-term residential care. *BMC Geri*. 2011;11(82):1–10.

Thomas L, Cordonnier-Jourdin C, Benhamou-Jantelet G, Diviné C, Le Louët H. Medication errors management process in hospital: a 6-month pilot study. *Fund Clin Pharm*. 2011;25:768–775.

Topazian R, Hook C, Mueller P. Duty to speak up in the health care setting: A professionalism and ethics analysis. *Minn Med*. 2013;96(11):40–43.

*Wittich CM, Burkle CM, Lanier WL. Medication errors: an overview for clinicians. *Mayo Clin Proc*. 2014;89(8):1116–1125.

8

Clinical Manifestations of Psychiatric Disorders

ANU A. MATORIN, M.D., ASIM A. SHAH, M.D., AND PEDRO RUIZ, M.D.

INTRODUCTION

Manifestations of psychiatric disorders express themselves as variations of normal functioning from subtle to obvious making it confounding for the clinician to integrate data into an accurate diagnostic picture. Most recently, the updated fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) has placed a strong focus on integration of symptom data, including patients' self-assessments of symptom severity. This dimensional approach is a major change in history taking and psychiatric practice with a goal to improve patient's overall satisfaction and treatment outcomes. In comparison to the previous categorical model of evaluation, this shift in approach aims to provide an even more thorough evaluation and comprehension of the patient's interpersonal world, work role, family life, and culture than is typical in general medical practice. Furthermore, this dimensional model will introduce a greater amount of information for the clinician to assess, evaluate, and measure, including the severity of psychiatric conditions.

The nature and expression of psychiatric signs and symptoms are profoundly altered by the patient's strengths, coping capacities, psychological defenses, and situational context so that the clinical picture ultimately represents a balance between psychopathology and psychological strengths. The clinical symptomatology must also be understood as part of the individual's attempts to cope with the biological, psychological, and environmental challenges that they face. When attempts to cope overwhelm the individual's capacity to respond, less adequate, more disorganized, and ineffective thoughts and behaviors emerge, and these comprise the impairments that present in clinical situations.

Need for a Comprehensive Clinical Perspective

DSM-5, like other medical disease classifications, uses multidimensional approaches to mental disorders, as the previous too-rigid categorical system was unable to capture clinical experience or important scientific

observations. The goal of this approach is to capture a more accurate description of patient presentations and for the diagnosis to be more valid. (*Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. American Psychiatric Association.)

The DSM-5 dimensional model addresses the longstanding concerns that clinicians often tend to limit their fields of vision and appreciate only part of the available information; the clinician's theoretical orientation and other personal and cultural factors also limit what is perceived. Research has demonstrated that clinicians tend to perceive primarily those signs and symptoms that are most in accord with their theoretical points of view and with the tools that they have available to treat psychiatric disorders, a phenomenon known as concept-driven perception. Such differences may lead one clinician to see a major mood disorder to be treated with medication, whereas another sees a pervasive personality problem with depressed mood to be treated with psychotherapy and to use different technical terms to label roughly the same phenomena. Psychodynamic-oriented psychiatrists might see psychomotor retardation, whereas a neuropsychiatrist sees bradykinesia; a psychodynamicist might see depressed affect and muted speech, whereas a neuropsychiatrist sees masklike facies and aprosodic speech.

The intermittent nature of many psychiatric signs and symptoms; the potential unreliability, selective recall, and false remembering of patients and others in reporting symptoms and events; differing interpretations of elicited information or observations; and subjective theoretically driven biases that influence the clinician's perception of signs and symptoms all contribute to potential errors in data collection. To help guard against misinformation and simplistic understandings and formulations, whenever possible, complete assessment of a psychiatric patient requires consultation with family, friends, coworkers, and other professional observers to enrich the history and to provide supplemental observations of the patient over time. Finally, via the DSM-5 dimensional assessment model, the comprehensive gathering of clinical symptoms will be the foundation for greater standardization, integration, and monitoring of treatment planning and outcomes.

THE CLINICAL CHALLENGES

DSM-5 development involved physicians, psychologist, social workers, nurses, counselors, epidemiologists, statisticians, neuroscientists, and neuropsychologist. This version even had feedback from patients, families, lawyers, consumer organizations, and advocacy groups. However, one of the costs of this increasing reliability has, in many instances, been the narrowing of the field of clinical vision. That is, clinicians who rely predominantly on structured interviews and checklists may become somewhat closed-minded and risk ignoring clinical phenomena that are

very important but that may not be part of the structured interview framework or mental set. An additional related difficulty can be seen in the use of nonclinical research staff that are trained to interview individuals using structured interviews (as in large epidemiological studies). In these instances, the need for interrater reliability is often achieved at the expense of a more nuance-based appreciation of the different meanings of subjective sensations, especially when the interviewer has no clinical experience with which to interpret the subject's descriptions.

However, in spite of these practical challenges, a thorough assessment of the clinical history and description of the psychopathology and detailed account of the patient's subjective experiences are important for the following reasons:

1. Significant diagnostic distinctions are made primarily on the basis of the historical information and elicited phenomenology. The more detailed, complete, and correct the diagnosis, the more rational and precise the treatment planning and the more reliable the prognosis. Consider, for example, the importance of accurately distinguishing between neuroleptic-induced akathisia versus anxiety symptoms related to psychotic thinking. On the basis of which of these diagnoses the clinician selects, the therapeutic strategy might be diametrically opposite.
2. The clinician's capacity to fully hear and communicate a comprehensive understanding of the patient's internal experiences helps to diminish the patient's sense of isolation so characteristic of many of these disorders and fosters the growth of a therapeutic alliance, increasing the likelihood of treatment adherence.

The Reliability Dilemma

As mentioned, for a given clinical manifestation of a psychiatric disturbance, different clinicians may arrive at different conclusions regarding the same clinical symptoms in the same patient. This reliability dilemma may confound an objective clinical assessment. Many factors contribute to this phenomenon. For example, some discrepancies may be due to differences in the patient's status or in information imparted by the patient from examination to examination, in the clinician's definitions of the symptoms or signs in question, and in perceiving and interpreting the patient's responses to general presentation or questions within the interview. A substantial amount of information variance in clinical practice reflects the impossibility of asking all questions within the time available for clinical evaluation. As an example, in one study, the rate of sexual side effects increased from 14 percent when depending on spontaneous patient report to 58 percent when the interviewer directly asked about these side effects.

A 52-year-old schoolteacher with recurrent major depression presented to her treating psychiatrist with her fiancé of 6 months. As part of the visit, the patient requested if she could stop her antidepressant medication, even though she had multiple severe relapses in the past. With her fiancé's encouragement, the patient disclosed arousal difficulties and anorgasmia after she started the medication. The patient had disclosed many painful memories related to past trauma to her psychiatrist, but reported feeling shy and embarrassed to mention the sexual side effects. Upon further questioning, the patient reluctantly disclosed that she was terrified she would lose her fiancé if she continued taking her current antidepressant and did not know how to address this issue impacting her personal life.

Even when simply responding to direct questions about symptoms, patients may respond differently depending on the interviewer's manner, how the questions are asked, their personal sense of trust or safety, whether they have answered these questions before, the amount of cuing that may signal the "desired" response, their fatigue, or a host of other variables.

Most clinicians still rely heavily on their own clinical intuition and subjective responses to patients as part of a diagnostic assessment. However, these clinical inferences, whether accurate or not, are often based on nonconscious assumptions, comparisons with other patients not well remembered, or distortions based on the clinician's own personal experiences. When the bases for these intuitions can be identified and described clearly, they may prove to be reliable and valid. However, intuitions are often wrong—simple trust in intuition alone is not sufficient. Thus, a clinician's sense that a patient is angry and potentially violent may result from the patient's subtle (but verifiable) body language and tone of voice, or it may represent a countertransference distortion that is not prompted by any observable patient behavior. At times, clinicians often too quickly label behaviors as inappropriate when they do not appreciate and understand contextual or cultural considerations. Appropriateness depends heavily on context, and definitions of what is proper in a given context may also be highly subjective. Appropriate behavior (or dress) in some parts of California may be inappropriate in Boston. Furthermore, the rapid change and growth of the current multicultural society places cultural understanding and context in psychiatric history taking and assessment in the forefront. Even manifestations of low intensity of emotional expression leading to a clinical description of "constricted affect" may reflect cultural norms or a psychopathological state.

Ambiguity in the Assessment of Psychiatric Signs and Symptoms

Since, at the present time there are no clear cut diagnostic, radiological, or laboratory tests to make a psychiatric diagnosis, virtually all psychiatric symptoms can be ambiguous or nonspecific and are usually seen in many different disorders. Depressed mood, for example, occurs in a wide variety

of diagnostic groups, including major depressive disorder, schizophrenia, some personality disorders, mood disorders secondary to medical or substance related states, and so on. Even the so-called first-rank symptoms of schizophrenia described by Kurt Schneider are diagnostically nonspecific. They are seen with some frequency in otherwise classic depressive and bipolar disorders. Apathy offers another good example of an important nonspecific phenomenon. Although apathy is often part of depression, research has shown this symptom to be a clinically distinct syndrome. Apathy is marked by lack of spontaneity, initiation, and emotionality as well as lack of activity and interest in friends, family, and hobbies. Apathy and the associated symptom of amotivation may exist in their own right without the presence of significant depression, as in initial phases of abstinence from cocaine and in other syndromes characterized by reduction in dopaminergic tone, such as frontotemporal dementias, Parkinson disease, or progressive supranuclear palsy.

In general medicine, symptoms not recognized as part of a clearly defined syndrome are often described as being of unknown origin. Thus, a fever that cannot be ascribed to a known disorder, such as pneumonia, is described as a fever of unknown origin. Given the nonspecific nature of psychiatric symptoms, it seems wise to use similar conventions, referring to hallucinations of unknown origin or depressed mood of unknown origin when a symptom cannot be clearly linked to a well-described syndrome. In DSM-5, the "not otherwise specified" (NOS) is replaced by *other specified disorder* and *unspecified disorder*. In the first category, clinician has to specify why the diagnosis doesn't fall under any specific category within a diagnostic class. If the clinician chooses not to specify the reason, then he or she will check on unspecified class. Even though individual signs and symptoms may be organized into syndromes and disorders, they often have courses of their own. Thus, in the appearance or the resolution of a disorder, certain associated signs and symptoms may appear very early or may persist after all the others have waned. For example, in the restricting form of anorexia nervosa, excessive exercise is often the first symptom to appear and the last to abate even after dieting has stopped. In some cases, certain signs and symptoms that are commonly associated with a given disorder may not appear. Each sign and symptom may have its own pattern and variable response to treatment. In the treatment of schizophrenia, for example, some patients experience rapid resolution of hallucinations but have persistent delusions without ever having any other thinking disturbances, whereas others may have no residual hallucinations or delusions but still have prominent thinking disturbances.

Predisposition and Vulnerability in Psychiatric Symptomatology: Genetic and Intrauterine Considerations

Genetic factors are key in determining the expression of most if not all

psychiatric disorders. Prominent among these are dementias of the Alzheimer type, schizophrenia, unipolar and bipolar mood disorders, anxiety disorders, alcohol dependence, and some personality traits. For virtually all of these disorders and traits, what is largely unknown at this point is the nature of the inherited vulnerabilities.

Intrauterine factors contribute too many psychiatric disorders. For example, maternal starvation and influenza infections during the second trimester of pregnancy have been implicated in the pathogenesis of schizophrenia. Maternal smoking and low birth weight may be risk factors in the pathogenesis of attention-deficit disorders in children. Maternal alcoholism may lead to fetal alcohol syndrome, a major cause of developmental disability.

Fundamental Neurobiological Factors

Investigational studies have shown that shortly after birth infants vary widely in temperament. For example, differences have been noted in spontaneous activity levels and thresholds, intensity, and duration of their reactions to external stimuli; the regularity or irregularity of certain biological rhythms such as sleep; tendencies to approach or withdraw from new stimuli; the speed and degree of adaptation; attention span and distractibility; the persistence of behavior; and qualities of mood. On the basis of such early behaviors, children may be described as having easy or difficult temperaments, quick or slow to warm up. Nevertheless, these temperamental qualities correlate somewhat with behavioral problems, at least through early childhood.

Aside from temperament, other persistent normal variations in personality development seem to be constitutionally related and may influence subsequent resilience or vulnerability. Traits, such as introversion, extroversion, and neuroticism, appear to be relatively enduring and stable personality dimensions. Other temperamental qualities that endure include novelty seeking, being relatively open to new experiences, and stick-to-it-iveness. Subtypes of intelligence, such as those related to conceptual, mathematical, musical, kinesthetic, and interpersonal abilities, have been postulated as having separate genetic determinants and patterns of development. The type A and B personality patterns, hardy and resilient personalities, high-strung, sensitive, fussy, irritable, and pessimistic characteristics have all been described as generally lifelong qualities that originate in early childhood. Even dimensions of character, referring to concepts about the self in relation to others that develop over time through social learning and maturation of interpersonal behavior, have been shown to have moderate heritability. These include such qualities as self-directedness, the ability to engage cooperatively with others, and the capacity to “transcend” the self by developing a sense of one’s place or purpose in the larger social context.

Other characteristics relevant to psychopathology but not diagnostically specific may also be inherited. These include psychosis proneness, cognitive styles (such as obsessionality or detail orientation), and emotional reactivity (which is part of the dimension of neuroticism). The relationship between the inheritance of these traits and the patterns of transmission of specific psychiatric disorders is unknown.

Impact of Physical Illness

All of the metabolic, toxic, infectious, and other causes of physical illness produce increased vulnerability to psychiatric disturbance. Studies have shown higher use of psychiatric services by those who are physically ill and higher than expected prevalence of physical disease among the psychiatrically impaired.

Some children with obsessive-compulsive disorder (OCD) and tic disorders with symptoms that appear during the prepubertal period and whose course of illness becomes episodic have been found to have pediatric autoimmune neuropsychiatric disorders associated with streptococcal (group A β-hemolytic streptococcus) (PANDAS) infections. Accompanying symptoms during episodes of exacerbation are emotional lability, separation anxiety, nighttime fears and bedtime rituals, cognitive deficits, oppositional behaviors, and motoric hyperactivity. In PANDAS, patients' flare-ups of behavioral problems are commonly associated with documented group A β-hemolytic streptococcus infections or symptoms of pharyngitis and upper respiratory infections.

Human immunodeficiency virus (HIV) infection leading to acquired immunodeficiency syndrome (AIDS) vividly illustrates the multiple and complex ways in which stressors can lead to psychiatric disturbances. These patients' psychiatric symptoms may represent central nervous system (CNS) changes that are the direct effects of the virus on the CNS producing changes in cognition, personality, and mood, expectable psychological adjustment responses of the patients in response to an overwhelming life-threatening disorder, or the emergence of latent or quiescent primary psychiatric problems provoked by the psychological stress of the viral illness. Adding to the complexity, some individuals with cluster B personality traits may engage in excessive risk-taking behaviors, increasing the likelihood of viral exposure; they may then be at higher risk for developing psychiatric symptoms or disorders in response to being HIV positive because of their preexisting cluster B traits.

A 42-year-old physician noticed the onset of decreased stamina, fatigue, disinterest in work, and weight gain without a feeling of increased appetite. Because his wife initially believed he was getting depressed, he consulted a psychiatrist. After the initial interview, the psychiatrist ordered a standard chemistry screening panel that revealed a thyroid-stimulating hormone of 48 mIU/L. He was referred to an endocrinologist and treated for his hypothyroidism, and

all of his depressive symptoms disappeared.

Effect of Life Events

The impacts of life events, particularly those that are negative and outside of a given individual's control, have complex effects on the development and persistence of psychiatric symptomatology. In general, the likelihood of developing psychiatric manifestations is increased in those who have a preexisting psychiatric disorder. After exposure to the same negative stressors, such as a serious accident or act of violence, individuals who previously had anxiety disorders are more prone than those without such histories to subsequently develop symptoms of posttraumatic stress disorder (PTSD). Other factors, such as gender, may also predict risk of developing PTSD, with women more likely to develop the disorder given the same exposure to stress. Although individual responses vary widely, truly catastrophic events, such as incarceration in a concentration camp, cause enduring psychiatric disturbances in a high percentage of survivors regardless of whether they had prior psychiatric problems. Similarly, the stress-related consequences of combat also vary widely, so that some heavily combat-exposed veterans develop long-lasting PTSD, whereas others develop very few persistent symptoms. The death of a parent or spouse, divorce, and major physical injury affect some people profoundly and others hardly at all in the long run. Significant stressors are likely to be more traumatic during early development rather than later or at certain critical developmental periods, compared with other times. For example, the loss of a parent at a very young age is likely to be more traumatic and have more profound and lasting effects than the loss of a parent as an adult.

The combined impact of negative life events and poor emotional and practical social support is important in predicting the emergence of at least some psychiatric disturbances. People who are ordinarily very competent in all role functions may fall apart completely when a supportive spouse who has bolstered them and taken care of many of their needs suddenly dies. Patients presenting with a major depressive episode have experienced more uncontrollable actual and threatened losses such as the death of a spouse in the year before onset.

A 65-year-old woman, presented to her primary care physician with her daughter after her neighbors told her daughter that they had noticed that she was more withdrawn and staying in her house for days without checking the mail or watering her favorite flower garden. The patient reported fatigue, "heavy leaden" arms and body, poor appetite, and poor sleep. The patient's husband of almost 40 years had suddenly passed away 1 year ago. The patient plaintively stated, "we had made so many plans in retirement and now they are all gone."

Nevertheless, not all psychiatric disturbance is attributable to easily

identified, provoking negative life events; indeed, some major negative life events that at first glance appear to have preceded the onset of a serious psychiatric disturbance may, in fact, have occurred only after the psychiatric disturbance actually began. For example, someone who attributes the onset of depression to having been fired from a job several months previously may already have been functioning suboptimally at that time and may have been fired as a consequence of a depression-induced decline in role function.

Certain environmental features can counter the effects of environmental stressors and protect against psychiatric disturbances. Stable families and friends, good financial circumstances, and supportive religious and social communities offer some protection. Research has shown that individuals with psychiatric disturbances have fewer social supports than normal controls. This may be due to friends' and relatives' withdrawal from deviant behaviors or to the disturbed individual's withdrawal from deleterious family and social relationships. In contrast, physically ill people have more social supports than others, perhaps reflecting their ability to recruit help in times of need. Of course, the quality, as well as the quantity, of social supports is important. As has been demonstrated in schizophrenia and mood disorders, for example, negative relationships even in close families may have deleterious effects, both in developing and in sustaining psychiatric disturbance.

The negative impact of a physiological or environmental stressor is closely related to its personal meaning. For example, the loss of a spouse who has been chronically demented, disabled, and burdensome ordinarily has a very different impact than the loss of a vital, supportive, loving spouse.

Finally, a relevant study has demonstrated the interaction of genetic vulnerabilities with adverse life events. In a study involving several hundred young adults followed prospectively, genetic vulnerability to depression was conferred by certain polymorphisms of the serotonin receptor. Over a period of several years, individuals with one or two copies of the short allele of the 5-hydroxytryptamine T promoter polymorphism exhibited more depressive symptoms, diagnosable depression, and suicidality in relation to stressful life events involving finances, housing, employment, and relationships than individuals homozygous for the long allele. This research provides strong evidence for gene–environment interactions, in which a person's responses to environmental insults are moderated by his or her genetic makeup.

DISTINGUISHING FEATURES OF PSYCHIATRIC SIGNS AND SYMPTOMS

Circumstance

Psychiatric signs and symptoms are usually not static entities; depending

on the circumstances, they often vary in intensity or even in their existence. The depressed mood of a melancholic depression may persist regardless of external situations, whereas the depressed mood in milder, reactive depression may vanish completely during certain situations—including a psychiatric interview—only to reappear at other times. Signs and symptoms that occur only in specific settings or with certain internal states are referred to as state dependent. For example, certain hallucinations or memories may be present only during states of drug or alcohol intoxication; in some patients, hives may erupt as a psychophysiological response only during states of anger. Interpersonal context is also important. Some people become violent only when involved in sadomasochistic relationships or in certain group settings such as adolescent gangs. In gangs, social pressures for conformity and expectations for aggressive behavior may provoke pathological behaviors that might otherwise never be expressed by gang members individually.

Cultural Context

Psychiatric signs and symptoms cannot be assessed independent of an individual's background and culture. Many phenomena often considered to be symptoms of psychiatric disorders may not be experienced as psychiatric problems by patients. Hearing an angel's voice may represent a manifestation of a psychotic disorder, yet the patient may vigorously dispute that the experience is a psychopathological symptom. In mainstream American culture, for instance, any experience of thoughts being broadcast out loud is considered pathological. These signs and symptoms must all be considered in context: Exactly what constitutes normal varies from culture to culture and from situation to situation. As mentioned earlier, behavior or subjective experience that may be defined as symptomatic in one context may be perfectly acceptable and within normal bounds in another. A phenomenon should be considered abnormal only if it seems deviant within the patient's specific culture after its full physiological and environmental context is taken into account and if it causes personal or interpersonal impairment. Too often, phenomena prematurely mislabeled as psychopathology turn out to be perfectly understandable and nonpathological once the whole situation is appraised. Conversely, some examiners are loath to label certain phenomena as psychopathological even when they clearly are, for fear of stigmatizing the patient.

A 22-year-old young man had recently moved to the United States from Mexico. He was admitted to a local hospital for the treatment of a leg infection. He was certain that his infection was the result of his "spirit" being possessed and refused to participate in his medical treatment. With his physician's understanding, his family brought in a community folk healer or "curandero" who performed a brief ceremony of prayers and rituals. Soon

thereafter, the young man agreed to receive and adhere to medical treatment.

Within cultures, most interpersonal interactions are carefully regulated by tight sets of rules and controls and constrained by reasonably well-defined sets of expectations and acceptable limits. Even slight deviations from these acceptable limits are quickly perceived by laypeople, as well as professionals, because behavioral deviances are often experienced as threats. Deviations in amplitude, duration, and intensity can occur in facial expressions, gestures, postures, vocalizations, language, and other expressions of emotion and thought. A small increase in the rate of speech, an intrusion into one person's conversation by another who does not allow proper pauses, a gesture that comes just a bit too close to a face, an excessively rigid or distant stance, or a gaze that is too staring or too avoidant each signals social insensitivity and alerts the observer to deviant behavior.

Sign and Symptom Classifications

The classification of clinical signs and symptoms in psychiatry has been accomplished using several different schemas: State versus trait, primary versus secondary, and form versus content. The state versus trait distinction refers to whether the sign or symptom is an enduring characteristic of the person ("traits") or time-limited phenomena associated with specific Axis I disorders, which are usually state phenomena. While DSM-5 does not use a multi axial system for diagnosis, and while personality disorders are still included as they were in DSM-IV-TR, an alternative model in DSM-5 has been included as Section III in the new DSM-5. As in clinical practice distinctions are often fuzzy, and comorbidity among the so-called categorically distinct disorders is often the rule rather than the exception, DSM-5 uses a single axial approach.

DSM-5 has changed the classification of schizophrenia by deleting all subtypes, but has kept all other types, like schizoaffective in the diagnosis. Additionally, a dimensional view of psychopathology fits much recent data better than the categorical view that is inherent in the DSM-IV-TR. That is the basis of DSM-5 changes for personality disorders also. For instance, personality disorders fit poorly into a categorical scheme, and the frequent "comorbidity" of these disorders likely reflects the descriptive overlap rather than the patient having two distinct disorders. Similarly, in the DSM-IV-TR, dysthymia and major depression are seen as two different mood disorders, while recent studies indicate that they are more likely manifestations of one disorder that differs in course and intensity. DSM-5 not only deleted the diagnosis of dysthymia, but combined it with major depression, as persistence major depression, with partial or full remission, early or late onset, mild moderate or severe as specifiers.

Impairments and Adversities

In addition to the traditional signs and symptoms of psychiatric disorders, recent focus has been on the impairments and adversities that psychiatric signs, symptoms, and disorders create in affecting specific role functions and causing social and economic burdens for the patient and others. These problems and impairments often cut across traditional sets of signs and symptoms of which categorical diagnoses are comprised, affecting, for example, basic abilities to care for oneself and one's family, marital functioning, child rearing, wage earning, school performance, and social behavior. They constitute the issues with which patients and families contend, and they need to appear on the problem lists that treatment plans and specific interventions target. Studies reveal, for example, that the impairments imposed by major depression are considerable with regard to physical functioning, role limitations, and social functioning. Furthermore, the relationship between symptoms and disorders on one hand and functional impairments on the other is not always straightforward. For example, in both bipolar and unipolar mood disorders, many patients recover symptomatically from episodes but achieve premorbid psychosocial function either months later or not at all. In one review of previously published studies on major depression, there was a consistent lag time between symptomatic and functional recovery. Whether this disparity between symptomatic and functional recovery reflects subtle residual symptoms, unrecognized cognitive disturbances, personality difficulties, or a combination of factors is unknown.

A 30-year-old woman with postpartum depression was reluctant to return to work as a fashion buyer. She felt she just "could not return to work this fast," in spite of marked improvement in her sleep, mood, energy level, and remission of her crying spells. She was treated with cognitive behavior therapy and antidepressant which she felt helped her very much. Yet, she felt "paralyzed" to perform in a job that she competitively had sought out prior to her pregnancy.

THINKING DISTURBANCES

Normal Thinking

Thinking is defined as the mental activity and processes used to imagine, appraise, evaluate, forecast, plan, create, and will. Most thought involves complex rules that are probably best currently approximated (albeit inexactly) by fuzzy-logic decision-making algorithms that use neural net technology, increasingly applied by scientists and engineers in vague situations in which all-or-none, black-or-white thinking does not apply but in which multifaceted, contradictory, and competing possibilities and biases are the rule. Most of what is known about thinking derives from the study of language as the product (and reflection) of thought, yet a great

deal of thinking takes place preverbally and nonverbally. Thinking occurs in images, music, and kinesthetic sensations and in symbols other than linguistic ones. Attempts to transmit preverbal and nonverbal thought using only words are frustrating and unsatisfactory. Creative artists have considerable difficulty describing the inner states of tension and inchoate awareness from which ideas are distilled.

Ordinary thought is far from logical. Streams of conscious thought are intruded on by competing thoughts and associations and by outside stimuli, and attention is easily distracted. Ordinary conversation is marked by recurring asides, interruptions, delays, and the loss of ideas. Decisions are often made on the basis of very few cues and inadequate evidence: People jump to conclusions. Beliefs are zealously held that are not supported by evidence. Thinking in stereotypes is more common than thinking in logical categories; from an evolutionary perspective, thinking in stereotypes and by approximation has probably been more adaptive than thinking in strictly defined categories. This tendency helps account for clinicians' tendencies to make diagnoses by approximation and intuition based on prototypes and to feel less comfortable using formal lists of criteria found in statistical manuals such as the DSM-IV-TR. Consistent with this notion, a recent study found that more than 60 percent of patients being treated for personality disorders by clinicians had diagnosable personality disorders when strict DSM-IV-TR criteria were applied, and the most common personality disorder diagnosed based on DSM-IV-TR was NOS. In DSM-5, the NOS is replaced by *other specified personality disorder* and *unspecified personality disorder*. In the first category, clinician has to specify why the diagnosis doesn't fall under any specific personality disorder category. If the clinician chooses not to specify the reason, then he or she will check on unspecified personality disorder. This is done to minimize the vague and confusing diagnosis of NOS.

Considerable variation exists among individuals regarding predominant cognitive styles, and an individual's style of thinking also shifts considerably from time to time. Cognitive style refers to one's predominant manner of information processing and decision making; the particular biases and distortions thinking processes make by means of augmenting, elaborating, or minimizing incoming information; and the extent to which people use careful and deliberate logic versus intuition versus thoughtless, anxiety-induced impulsivity to guide decision making. In some individuals, a particular cognitive style may come to dominate that person's repertoire so completely as to interfere with the flexible, adaptive responses required to deal with the usual variety of daily needs. An obsessional style of thinking is marked by attention to detail and hypervigilance concerning the possible implications of a particular thought or event. This may take the form of preoccupation with strict adherence to established rules, values, or beliefs. An obsessional style may be highly adaptive in certain situations, as

in professions requiring meticulous detail such as librarians, computer programmers, and surgeons. However, excessively rigid obsessionality may be maladaptive, as when someone scrupulously sticks to the rules even when such adherence is self-destructive and short sighted. A hysterical style of thinking is characterized by global, diffuse, impressionistic, emotionally laden evaluations of situations in which lack of attention is given to details and nuances. This style is poorly adaptive to detail-oriented work but may be useful in the arts, certain aspects of marketing and sales, and some social situations.

A 35-year-old pharmaceutical sales manager had difficulty in managing work responsibilities in her new job. She reported spending long hours focusing on minute details of each task herself. She struggled to delegate responsibilities to her staff for or to prioritize her workload for fear of missing any small detail. She was unable to meet deadlines for job reports or assigned tasks. Frustrated, she eventually tasked for a job transfer.

Types of Thinking

Because of the different ways in which both normal and abnormal thinking expresses itself, differences that are apparent to even a casual observer, attempts have been made to subtype thinking by the extent to which logical versus nonlogical thought is used. Although less commonly used than before, Sigmund Freud's division of thought into primary and secondary process thinking provides a useful descriptive classification.

Primary Process. Primary process thinking, the more primitive type, is typically seen in dreams but is also prominent in young children and in psychotic states. This type of thinking disregards logic, permits contradictions to exist simultaneously, disregards the linear notion of time, and is dominated by wish and fantasy. It uses symbol, metaphor, imagery, condensation, displacement, and concretism in its organization, creating the jumbled and incoherent style of thinking characteristic of dreams. Primary process thinking represents what has been loosely and metaphorically called right brain thinking, associated with visual images and creative thought.

Secondary Process. Secondary process thinking is characterized by logic. In contrast to primary process thinking, the secondary process uses linear notions of time, clearly delineated abstract categories, and deductive rules of logic. The abilities to think abstractly and to think in detail about future plans are characteristic of secondary process thinking. Normal secondary process thinking is also characterized by predictability, coherence, and redundancy. Words, vocal inflections, and gestures provide important contextual cues and create a sense of overall coherence to the communication. Ideas follow one another in a sequence that is

understandable to the listener.

A non-Freudian typology of thought divides thinking into three types: Fantasy thinking, imaginative thinking, and rational or conceptual thinking. Fantasy thinking allows the person to escape from, or deny, reality and can be seen in normal as well as pathological thinking. Everyone occasionally uses fantasy thinking when daydreaming. Some dissociative and psychotic phenomena illustrate the most pathological manifestations of fantasy thinking. Imaginative thinking merges fantasy and memory to generate plans for the future. Rational or conceptual thinking uses logic to solve problems.

Regardless of how one categorizes thought, people can fluidly shift from linear/secondary, process/rational thought to fantasy/primary, process/nonlogical thought, as in the free associative method used in psychoanalysis. During this process, individuals willfully surrender the controls that maintain secondary process thinking and switch to the less controlled modes of primary process thinking in which thoughts are loosely associated by emotional associations or based on peripheral, concrete, coincidental, loosely similar, or trivial aspects of a thought. Additionally, the fact that increases in primary process thinking can be induced in normal people under experimental conditions or with fatigue suggests that more primitive thought processes, such as those seen in psychosis, are usually inhibited by higher-order processes and that their appearance may be release phenomena—that is, nonlinear or psychotic thinking may indicate the functional absence of those overriding control systems that ordinarily sift, evaluate, and regulate the form and flow of thought before it reaches consciousness.

Flow and Form Disturbances

Because the underlying processes that govern thought are not understood, current systems for classifying thought abnormalities are primarily descriptive. Conventional classification separates form and flow from the content of thought, yet many types of abnormal thinking include both form and content abnormalities. Thus, although delusions are usually classified as thought content disturbances, they are also marked by formal abnormalities, such as rigidity and imperviousness of thought to external influence or to information that clearly contradicts the delusional idea.

Although formal thought disorder typically refers to marked abnormalities in the form and flow or connectivity of thought, some clinicians use the term broadly to include any psychotic cognitive sign or symptom.

Normal variations in the flow and form of thought are considerable. For some people, thinking appears to be effortless, rapid and productive, exhibiting linear, goal-directed thoughts and creativity, with digressions and occasional leaps but always controlled and comprehensible. For others,

thinking is a difficult exercise, a slow, painstaking process with low output, compared with other people, or "scattered," with difficulty staying with a topic or finishing a single thought. Most people experience admixtures of these extremes. Disturbances in the flow and form of thought occur with regard to rate, continuity, control, and complexity.

Thinking can be unusually slow or accelerated. Slowed (or retarded) thought, such as noted in depression, is typically goal directed but characterized by little initiative or planning. Patients experiencing retarded thought often describe feeling that even simple thought requires monumental effort, as if molasses were cluttering their thinking. These difficulties are expressed as slowness in decision making and as long latency of response, increased pause times when speech is initiated and during speech. Thought blocking, seen in schizophrenia, is experienced as the snapping off or as a sudden break in a train of thought, as if a wall suddenly comes down, interrupting thinking (and speaking) in midsentence. To an outside observer, without further explanation from the patient, thought blocking may appear identical to thought withdrawal, a disturbance in the control of thought in which the patient feels as if some alien force has intentionally withdrawn the thoughts from consciousness. The patient's further description and explanation of the inner experience is necessary to distinguish these two symptoms.

A 27-year-old man was brought to the emergency room by ambulance after being found at a gas station making bizarre hand signals at passing motorists. The man spoke in a slow, halting manner. During the interview, he would complete his sentence partially, then he would stare blankly with an odd smile. Eventually, he would complete the remainder of the sentence as if he had never stopped midway. He reluctantly disclosed that his thoughts were being programmed and controlled by powerful computer signals.

Accelerated rates of thinking, typically accompanied by fast talking, can be seen as a normal variant. Rapid rates of speech, influenced heavily by cultural and situational factors, only sometimes reflect truly rapid thought. Pressure of speech—speech that is rapid, excessive, and typically loud—is characteristic of mania (or hypomania), stimulant intoxication, and, occasionally, anxiety. Flight of ideas occurs when the flow of thought increases to the point at which the train of thought switches direction frequently and rapidly. The associative links between conceptual topics during flight of ideas are comprehensible to the listener, although not without considerable effort at times. Listening to a flight of ideas that is not overwhelmingly fast can be both a dizzying and enjoyable experience for the listener, as exemplified by the successful performance style of certain contemporary comedians.

Continuity

Disturbances in the continuity of thought may take several forms. In circumstantiality, the flow of thought includes many digressive turns and associations, often including a great deal of unnecessary detail. Transcripts of circumstantial thought or speech are marked by multiple commas, subclauses, and parenthetic asides. Nonetheless, in circumstantial thought or speech, the speaker eventually returns to the point that was initially intended without having to be prompted by the listener.

In contrast, in tangentiality, the person's thought wanders further and further away from the intended point, without ever returning, so that the person may not even remember what the original point was supposed to be. Tangentiality is a mild form of derailment in which there is a breakdown in associations. Loose associations exemplify more severe derailment, in which the flows of ideas are no longer comprehensible to the listener because the individual thoughts seem to have no logical relation to one another. Loose associations are classically a hallmark feature of schizophrenia. In extreme cases, the associations of phrases and even individual words are incomprehensible, and syntax—the rules of grammar by which phrases are organized into sentences and words into phrases—may be disrupted. Word salad describes the stringing together of words that seem to have no logical association, and verbigeration describes the disappearance of understandable speech, replaced by strings of incoherent utterances.

Clang association refers to a sequence of thoughts stimulated by the sound of a preceding word. For example, a manic patient said, "I'll kill with a drill or a pill—God, I'm ill—what swill." In echolalia, the patient repeats a sentence just uttered by the examiner. Repetition of only the last uttered word or phrase is called palilalia, a symptom found most often in chronic schizophrenia.

Perseveration and stereotypy are two other associative abnormalities in which the flow of thought or speech appears to get stuck. In perseveration, a sentence or phrase is repeated, sometimes several times over, after it is no longer relevant. Perseveration is commonly seen in delirium and other medical, substance related, and/or CNS disorders. Stereotypy refers to the constant repetition of a phrase or a behavior in many different settings, irrespective of context.

Disturbances in the control of thought include delusional passivity experiences and obsessional thinking. In delusional thought passivity, patients experience their own thoughts as being under the control of other forces. Thought passivity may take several forms: In thought insertion, thoughts are experienced as having been placed within the patient's mind from the outside; in thought withdrawal, thoughts are whisked out of the mind; in thought broadcasting, patients experience their thoughts as escaping their minds to be heard by others. These experiences are often combined with specific delusions of control, seemingly to explain the

passivity experiences. Several of these phenomena were included by Schneider among the so-called first-rank symptoms of schizophrenia. Today, these symptoms are viewed more broadly as nonspecific psychotic symptoms, more likely to be seen in schizophrenia but not pathognomonic of the disorder.

A 36-year-old man with schizophrenia had placed silver duct tape tightly around every window in his home. He was convinced that his neighbors had a device connected to their TV satellite dish that could control his thoughts and actions. He calmly explained that the duct tape interfered with the satellite signal. He never ventured out of his home without headphones on so that no one could hear his thoughts or place bad thoughts in his head to control him.

Obsessional thinking is stereotyped, repetitive, persistent thinking that is recognized as one's own thoughts. Obsessional patients do not experience their thoughts as being controlled by outside forces. Nonetheless, they experience only partial control over the obsessional thoughts. They can, with great effort, stop thinking the obsessional thoughts but cannot prevent them from recurring. Thus, characteristic of obsessions are the subjective experience of compulsion and the resistance to it. In classic obsessional thinking, insight is retained, and, as bizarre as some obsessions are, patients know that these thoughts are irrational and their own. However, more recent studies have revealed that insight into obsessional thinking is more variable than had been previously believed, at times becoming delusional. At times, obsessions may be pervasive enough to dominate the patient's consciousness. Obsessions may be simple, a sequence of words, or elaborate, such as enumerating the possible consequences of a past behavior and elaborating a cascading sequence of typically catastrophic events. Typical obsessional themes in OCD involve preoccupations with dirt and contamination, fear of harming others, symmetry, and those related to health and appearance.

Whenever she drove her car over a bump in the road, a 42-year-old woman with OCD worried that she had hit a pedestrian. Because she could not see the body, she drove around the block to look for it. After seeing no body in the road, she was transiently reassured but, within moments, became concerned that the body may have rolled under a parked car. She drove around the block for at least 30 minutes searching for the body when this obsessional thought possessed her.

Obsessional thoughts are usually seen in conjunction with compulsive behaviors, which are rituals linked to the obsessions typically constructed to undo the effects of the thought. Thus, contamination obsessions are linked to cleaning rituals, fears of harming others lead to checking rituals, and so forth.

The most prominent disturbance of thinking complexity is an impaired capacity to think abstractly. Abstract thinking is the ability to assume a mental set, to keep simultaneously in mind all of the aspects of a complex situation, to move from feature to feature as indicated by the situation, and to abstract common properties. Complex thinking also concerns the ability to simultaneously consider many different, vague, and subtle aspects of situations; to appreciate differing and contradictory points of view; and to integrate these multiple dimensions to form opinions that are marked by differentiatedness and nuance. Normal individuals vary greatly in their abilities to engage in abstract thinking—geniuses in mathematics and theoretical physics leave most mortals far behind. Concrete thinking is a disturbance in the ability to form abstract concepts, generally illustrated by literal mindedness and the inability to abstract the commonality of members of a group, for example, the fact that a flea and a tree are similar in that they are both living things. Concrete thinkers seem unable to free themselves from the literal or superficial meanings of words. Concrete thinkers may be more prone to prejudice and stereotypical thinking and more likely to manifest unidimensional or all-or-none reactions to complex situations. Concrete thinking can be seen in individuals with lower intelligence, organic mental disorders, and schizophrenia. Schizophrenic patients may also exhibit highly selective disturbances of abstraction.

Thought Content

The normal content of thought, the buzzing, booming stream of consciousness that constitutes the stuff of everyday life, comprises awareness, concerns, beliefs, preoccupations, wishes, and fantasies occurring with various degrees of clarity, vividness, differentiation, imagination, and strength. Normal thought is often illogical, containing many beliefs and prejudices that may be clearly contradictory but are nevertheless held with passion and conviction.

Belief systems are the scaffolding of thought, chains of impressions, and expectations around which plans and behaviors are organized. Belief systems may be attitudinal, setting general expectations and biases about the world that inform how incoming information is processed; examples are optimism, pessimism, and paranoia. Some beliefs are evanescent, whereas others are pervasive, tenacious, enduring, and influential. The enduring belief systems are associated with behaviors consistent with the belief, at times dominating interpersonal relationships and lifestyles. Some beliefs are unique and private, whereas many are shared by others.

A 48-year-old man had been deeply pessimistic since childhood. Although he had no vegetative features of depression, he believed and behaved as if the world was an awful place dominated by dishonest and selfish individuals. Having any goals seemed foolish to him, given the nature of the world. He expressed these beliefs regularly to his children, to the

dismay of his wife, who did not share his belief system.

Imaginative fantasy is an important component of normal thought. The vivid, eidetic imaginations of young children can produce vivid fantasies in which children become fully immersed, almost as if in hypnotic states. During latency, many children develop imaginary companions as playmates. In later years, imaginative thinking in which previously separate streams of thought playfully interact with one another to produce new ideas may be the essence of the creative reverie. Artists, writers, and creative scientists may retain access to these forms of thinking more readily than others. Meditative states of mind may facilitate the emergence of imaginative insights. Such thinking may also occur in dreams. Intrusive reveries are normal and common components of the usual adult stream of consciousness. During periods of specific deprivation, such as starvation or sexual deprivation, elaborate wish-fulfilling daydreams frequently occur.

Ideas are the contents of the stream of thought. Those that are consistent with one's sense of self, compatible with the individual's self-image, are called ego-syntonic. Other thoughts that conflict with one's central values are called ego-alien or ego-dystonic. An ego-dystonic impulse to kill someone, inconsistent with one's predominant value systems, may generate a counteractive ego-syntonic thought such as "you really don't mean it."

Disturbances in Thought Contents. Abnormal beliefs and convictions form the core of thought content disturbances. Considerations of abnormality regarding beliefs and convictions must take the person's culture into account. Beliefs that may seem abnormal in one culture or subculture may be commonly accepted in another. For example, religious hallucinations, attributed to psychological or biological factors by contemporary Western societies, are routinely attributed to religious and spiritual causes by many other cultures. With regard to intensity of conviction, distorted beliefs range on a continuum from overvalued ideas to the determined, unshakable belief that is characteristic of fixed delusions. Abnormal beliefs and delusions are, in most circumstances, diagnostically nonspecific. Delusions are commonly seen in mania, depression, schizoaffective disorder, delirium, dementia, and substance-abuse-related syndromes, as well as in schizophrenia and delusional disorders.

Overvalued ideas are unreasonable and sustained abnormal beliefs that are held beyond the bounds of reason. Patients with overvalued ideas have little or no insight into the fact that their ideas are very unlikely to be valid; however, the ideas themselves are not as patently unbelievable as most delusions. The distorted body images of body dysmorphic disorder exemplify overvalued ideas. Morbid jealousy and preoccupation with a spouse's possible infidelity may constitute an overvalued idea if no real

evidence has ever existed to warrant suspicion.

> A 42-year-old man who had a brief relationship with a woman was unable to accept the fact that she no longer wanted to see him. He ruminated about every encounter that they had and interpreted small gestures in the past as indicating her undying love for him. His infatuation led him to follow her repeatedly to work and school, and he pursued her relentlessly, to the point that she brought charges against him for stalking.

Ideas of reference are false personalized interpretations of actual events in which individuals believe that occurrences or remarks refer specifically to them when in fact they do not. Ideas of reference may be less firmly held than delusional beliefs.

> A 20-year-old college student was brought to the student health clinic because she was pacing outside her dormitory building in an agitated manner. She stated that she could not sleep for several days because when she heard a cell phone ring, it was a signal that other students were talking about her and laughing at her.

Delusions. Delusions are fixed, false beliefs, strongly held and immutable in the face of refuting evidence, that are not consonant with the person's educational, social, and cultural background. Thus, delusional thoughts can only be understood or evaluated with at least some knowledge of patients' interpersonal worlds, such as their involvements with religious or political groups. One of the mind's functions is to generate beliefs, including myths and meaning systems. These beliefs provide the individual with a sense of personal and group identity and with ways of understanding reality. They are most noticeable when shared untestable beliefs form the basis for group cohesion, as in religions and cults. Some groups adhere to their cherished beliefs despite the abundance of plausible contrary evidence —for example, the fundamentalist sects that take the biblical creation story literally. In the face of contrary evidence or grave personal threat, individuals often cling to their primary beliefs as matters of faith (i.e., alternative, nonrefutable bases for understanding). The strong faith with which religious, political, and nationalistic convictions are held, even at the cost of death, shows the power that untestable beliefs can have on behavior. Potential mental health advantages of religious beliefs have been demonstrated in epidemiological studies showing that those with a sense of personal devotion report fewer depressive symptoms.

Subjectively, delusions are indistinguishable from everyday beliefs. Therefore, the subjective experience of a delusion is no different from the subjective experience of believing that the earth is round or that one's spouse is the same person that one married on his or her wedding day. Because of the identical experience of delusions and other strongly held beliefs, it is generally impossible to argue a patient out of a delusional

belief. The content of delusions is highly influenced by culture. Centuries ago, delusions of persecution often concerned persecution by the devil and had religious connotations; persecutory delusions today often take on contemporary technological, political, and social perspectives.

Table 8–1.
Characteristics of Delusions

Simple vs. complex
Complete vs. partial
Systematized vs. nonsystematized
Primary (autochthonous) vs. secondary
Persecutory vs. nonpersecutory
How they affect behavior

A 25-year-old man diagnosed with schizophrenia adamantly believes that he has discovered a new program that will transform the internet. He explains this transforming discovery is due to his ability to receive messages directly from his computer. He would not elaborate further on the program because of his belief that other people would steal the program from his brain or his thoughts would be broadcast out loud, destroying his secret discovery.

Although delusions are diagnostically nonspecific, some types of delusions are more prevalent in one disorder than another. For example, although delusions of control and delusional percepts are often seen in schizophrenia, they also occur, albeit less frequently, in psychotic mood disorders. Similarly, classic mood-congruent delusions, with grandiose themes seen in mania or delusions of poverty characteristic of depression, may also be seen in schizophrenia.

Table 8–1lists some characteristics by which delusions have been classified. Simple delusions contain relatively few elements, whereas complex delusions may contain extensive elaborations of people, spirits, motives, and situations.

Systematized delusions are usually restricted or circumscribed to well-delineated areas and are ordinarily associated with a clear sensorium and absence of hallucinations.

They are often isolated from other aspects of behavior. In contrast, nonsystematized delusions usually extend into many areas of life, and new data—new people and situations—are constantly incorporated to further support the presence of the delusion. The patient usually has concurrent mental confusion, hallucinations, and some affective lability. Whereas the patient with a closed systematized delusional system may go about life relatively unperturbed, the patient with a nonsystematized delusion frequently has poor social functioning and often behaves in response to the delusional beliefs.

Complete delusions are those held utterly without doubt. In contrast,

partial delusions are those in which the patient entertains doubts about the delusional beliefs. Such doubts may be seen during the slow development of a delusion, as the delusion is gradually given up, or intermittently throughout its course.

During an acute schizophrenic episode, a 23-year-old man was completely convinced that the subtle pattern of superficial veins on his legs was evidence of "astral domination" by extraterrestrial beings. After 1 week of antipsychotic medication, he believed the veins were caused by others with special powers but was less sure. Two weeks later, he denied believing that the patterns on his legs were other than superficial veins and dismissed his previous belief with a shrug.

An autochthonous delusion is one that takes form in an instant, without identifiable preceding events, as if full awareness suddenly bursts forth in an unexpected flash of insight like a bolt from the blue. These delusions may be quite elaborate.

After a 2-year period of gradual academic and interpersonal decline, a 21-year-old man had the sudden conviction that certain songs played on the radio used his voice in the role of lead singer. He could not explain why this would be so nor why this belief emerged suddenly when it did.

Aside from the autochthonous types, three other types of delusions have been described as primary. Delusional percept refers to the experience of interpreting a normal perception with a delusional meaning, one that has enormous personal significance to the patient. Delusional atmosphere or delusional mood is a state of perplexity, a sense that something uncanny or odd is going on that involves the patient but in unspecified ways. Ordinary events may take on heightened significance, but the delusional interpretations are fleeting, although the uncanny feeling stays. Typically, after a period, full-blown delusions develop, replacing the delusional mood. Delusional memory is the memory of an event that is clearly delusional. As an example, a patient "remembered" that his fourth-grade teacher slipped lysergic acid diethylamide (LSD) into his apple juice; this memory served to explain his psychotic disorder. The elaboration of false memories and their subsequent fixed belief may assume delusional proportions.

A young schizophrenic woman attended a "trauma group" with her roommate and gradually came to believe that she had been repeatedly sexually assaulted by her father from the time she was in the crib. What started out as vague dreamlike images gradually coalesced into a series of "sensed" memories that then took on specific visual images of her father's fingers penetrating her and of his looking down at her in the crib leering. Her parents were horrified by these accusations, and there was not a shred of evidence to corroborate her increasingly venomous accusations.

Patients vary considerably in the extent to which they take action in response to delusional thoughts. Just as patients can experience delusions of their thoughts being controlled (thought passivity), they may similarly experience their feelings, behaviors, and will as controlled by outside forces. These delusions of control (or passivity experiences) occasionally, albeit uncommonly, result in dramatic self-destructive or aggressive behaviors, as illustrated by the murderer who called himself Son of Sam. This psychotic killer murdered a series of people in New York and claimed that he was the powerless agent of a force that required him to commit the acts. To defend themselves and others against delusional anticipated events, some patients may take bold and occasionally destructive actions.

A 38-year-old man with untreated schizophrenia was transferred from jail to hospital for evaluation and treatment after shooting (but not killing) his neighbor. He explained that the neighbor had been "hassling" him for 6 months by making noise in the middle of the night, laughing at him, and listening to his conversations through the wall. The patient explained that he shot the neighbor as a warning to back off.

Table 8–2.
Some Classic Types of Delusions

Delusions of persecution
Delusions of grandeur
Delusions of influence
Delusion of having sinned
Nihilistic delusions
Somatic delusions
Delusion of doubles (doppelganger)
Delusional jealousy (Othello syndrome)
Delusional mood
Delusional perception
Delusional memory
Delusions of erotic attachment (Clérambault syndrome)
Delusions of replacement of significant others (Capgras syndrome)
Delusions of disguise (Frégoli phenomenon)
Shared delusions (folie à deux, folie à trios, folie à famille)

Table 8–2lists some classic types of delusions. Although less common than those involving paranoia, grandiosity, and influence, delusions of misidentification are prominently reported because of their inherently intriguing nature. In Capgras syndrome, the patient believes that someone close to him or her has been replaced by an exact double. In Frégoli phenomenon, strangers are identified as familiar people in the patient's life. In the delusion of doubles, patients believe that another person has been physically transformed into themselves.

A 23-year-old young woman, recently diagnosed with schizophrenia was brought to the emergency center by ambulance after she called 911 screaming that "give me back my mother, I know what you have done." In the hospital, she accused her mother of being a double, in spite of the mother's attempts to show her old family photographs and her birth certificate.

Olfactory delusions that one emanates a foul odor are common in social anxiety syndromes, in which individuals are particularly concerned about potentially embarrassing themselves and others.

Delusions are not only seen in isolated individuals. Shared delusions may occur in couples (folie à deux) and in families (folie à famille). Many psychiatrists consider group delusions to be present in some cults as well, but exactly where the cutoff points occur between delusions and other zealous beliefs held by larger, more traditional and well-organized religious, political, and other groups is arguable.

A husband and wife were brought by police for a psychiatric evaluation. They both wore a large colorful talisman made of metal on their head which had a particular significance to them. The reason for the talisman, they concurred, was to ward off cell phone signals which could control their actions and project their thoughts to others. The husband was a dominant personality who had originated this fixed delusion. His wife, a mild-mannered, passive person, seemed to have accepted her husband's interpretation of events without challenging them at all.

DISTURBANCES OF JUDGMENT

Judgment involves a complex and diverse group of mental functions that includes analytical thinking, social and ethical action tendencies, and depth of understanding or insight. Analytical thinking includes the capacity to discriminate and to weigh the pros and cons of potential alternative actions. Social and ethical action tendencies are closely related to culture and upbringing. The evidence for genetic factors in antisocial personality disorder (defined primarily by judgments that lead to criminal behaviors) points to the additional role of constitutional factors. Insight may reflect intelligence, learning, cognitive style, and the capacity to integrate intellectual knowledge with emotional awareness.

Impairments of judgment occur in many psychiatric disturbances. Anxiety states, intoxications, fatigue, and even group pressures may cause temporary impairments of judgment in otherwise normal individuals. In mood disorders, judgment may be impaired by either an exaggerated evaluation of risk or failure in depression or, conversely, of inadequate appreciation of risk or danger in mania. Brain damage caused from medical conditions, including persistent substance-related conditions and psychotic disorders may chronically impair any aspect of judgment in any person, regardless of premorbid character. Poor role models and deviant social

backgrounds may lead to social and ethical action tendencies quite different from those of the examiner. Thus, someone raised in a criminal environment may have superb analytical judgment and self-awareness, which are, however, put to illegal use.

> A 48-year-old attorney missed multiple visits with his established psychiatrist and subsequently ran out of his mood stabilizer, in part due to increased travel related to a prominent out of town legal case. His wife of 20 years became increasingly concerned as he erratically began flying first class, spending $40,000 on airfare and not disclosing his whereabouts. After not communicating with his wife for 2 days, he was found by security at an out of town television station claiming he had been hired their national legal correspondent to provide legal commentary.

Judgment may be impaired in one dimension and spared in others. Individuals may retain sound ethical judgment when their analytical capacities fail or may retain excellent analytical abilities for nonpersonal matters, although lacking insight into personal situations or behaviors. Thus, some people who can provide socially appropriate responses to traditional mental status examination questions, such as what one would do in a movie theater if fire broke out, might at the same time be incapable of accurately assessing crucial clinical or more personal matters specifically related to one's capacity to provide informed consent, such as the pros and cons of receiving treatment; regarding judgments necessary to provide oneself with food, clothing, and shelter; or insight into one's state of health or illness.

The term insight, usually in the context of self-awareness, has been used in a variety of ways. Basic insight refers to a superficial awareness of one's situation. In evaluating insight into one's psychiatric condition, basic insight allows an individual to acknowledge the presence of an illness. A deeper level of insight is operating when the patient has an intellectual appreciation of what is going on (e.g., "I have hallucinations and delusions, and my doctors have told me that I have schizophrenia and must take medication"). Still deeper levels of insight reflect more complete cognitive and emotional appreciation of a situation (e.g., "I realize that I have schizophrenia, that it impairs my judgment and social function at times, and that I will have to take medications if I am to minimize my symptoms and try to make the most of my life. I feel profoundly disappointed about this affliction because it prevents me from achieving some of the goals I've always wished for. Nevertheless, I have do my best to get over my disappointment and hurt feelings so that I can get whatever I can out of life.")

Of course, different depths of insight as self-awareness can be evaluated in many other situations, such as physical illness, quality and nature of relationships, an appreciation of strengths and weaknesses in professional

situations, and so forth. In formal studies of insight using standardized instruments, lack of insight correlates with poor outcome in schizophrenia and bipolar disorder, medication noncompliance, and suicidality. Improvement of psychosis does not necessarily correlate with improved insight. Impaired insight may be associated with frontal lobe abnormalities. Insight may be as seriously impaired in mania as in schizophrenia and, contrary to earlier beliefs, may be lacking in OCD.

Judgment may be impaired by several factors, including cognitive clouding (as in disturbances of consciousness, e.g., intoxication, so that one's usual analytical abilities are impaired), self-deception, and impulsivity.

Self-deception refers to the almost universal tendency to hide certain issues about the external world or about oneself from various levels of awareness. Self-deception functions as a coping strategy, fostering or maintaining comfortable perspectives about the world and avoiding confrontation with issues and realities that inevitably stir up painful conflicts or the need for difficult actions, thereby preserving emotional calm. In addition, studies suggest that self-deception enables people to act and to be perceived as more convincing in the service of particular goals, as in romantic relationships or business dealings. Therefore, although "kidding ourselves" may sometimes reflect impaired judgment, it may at times also yield certain important strategic advantages.

Impulsive judgment describes a tendency to avoid taking the time and thought to fully understand and integrate all of the facts and levels of awareness required for optimal decision making. Impulsive judgment may occur only with certain issues or situations (such as how one picks investments), signal an impaired state (such as intoxication), or reflect a pervasive character trait.

Rapidly made judgments, and even the so-called snap judgments, may not be maladaptively impulsive, even when they involve very important areas of life. Rapid decisions can be very accurate, highly adaptive, and even life-saving, especially if made against a background of great experience, wisdom, and forethought concerning the area requiring the decision.

Because, in clinical practice, the terms insight and judgment are often applied to individuals' awareness and decision making about their psychiatric status, complex motivational states that incorporate insight and judgment related to how one is dealing with one's problems, the so-called stages of readiness for change, bear mention at this point. Initially described in relation to substance abuse, including alcoholism and smoking, these stages have received considerable attention and form an important aspect of clinical assessment across other diagnostic categories such as eating disorders. Several stages have been described: (1) precontemplation, the person expresses no intention to change (may be in

denial); (2) contemplation, the person acknowledges a problem and states an intention to change within several months but not right away; (3) preparation, the person intends to do something about changing in the near future and may have already made some false starts; (4) action, the individual has engaged in making sustained behavior changes; (5) maintenance, the individual has been engaged in changed behavior for more than 6 months; and (6) termination, the individual has succeeded in the change and is unlikely to ever return to the original behavior.

DISTURBANCES OF PERCEPTION

Normal perception first requires that the individual be capable of receiving information as sensations. The data must then be organized to make them meaningful and comprehensible such as distinguishing figure from ground or focusing attention selectively on some part of the sensory field. The organized entities are called percepts. In states of sensory deficit, such as blindness, deafness, and anesthesia, perception is impaired, but perception is still possible because individuals generally perceive information about an object through several sensory modalities concurrently. The intensity of sensation and perception is affected by vigilance and attention. Highly focused attention, as in intense concentration or hypnosis, may result in unusually acute sensation and perception—hyperesthesia, hyperacusis, or extraordinary visual acuity. Focused attention may also result in the inability to sense or perceive: Deep anesthesia and negative hallucinations induced by hypnosis are simply induced failures to perceive what exists in the world.

Humans usually operate in an "average expectable environment" in which certain types and levels of sensory input are expected and for which the nervous system is primed. Excessive or inadequate stimulation in any sensory modality, levels of input that are extraordinarily intense, or the presentation of novel stimuli that are entirely different from anything previously experienced by the individual can provoke distorted perceptions in most normal people. For example, total sensory deprivation produced in carefully controlled artificial environments may elicit visual and auditory illusions and hallucinations.

Individuals generally exhibit selective perception of the world, depending on what is salient at the moment and on his or her individual memories, emotions, fantasies, and values. Pregnant women are more likely to perceive babies around them than are people who are not as preoccupied with childbearing.

The intensity of perceptions depends on individual sensitivities as well as on mood, anxiety, and substance use. Unmedicated patients with schizophrenia have deficits in olfactory acuity. Depressed patients often describe that colors look faded and that the world looks washed out or gray, even though their capacity to recognize specific colors is unchanged.

Similarly, mania is often characterized by heightened perceptions, hyperesthesia. When extreme, these intense perceptions are uncomfortable. Hyperesthesia can also be seen during benzodiazepine withdrawal, hallucinogen use, and, occasionally, as part of an epileptic aura.

The intensity of perception may vary with cognitive style and other psychological and neurological factors. Some individuals tend to be augmenters and others minimizers of bodily experiences. Chronic pain and some hypochondriacal syndromes may occur more commonly among somatic augmenters.

Selective deficits in the perception of emotions may occur. Emotional aprosodies have been described in which patients with specific neurological deficits or depression are selectively unable to recognize the expression of facial emotion. These have been linked by positron emission tomography (PET) scan to blunted activity in the right prefrontal cortex and insula.

Illusions

Perceptual distortions in estimating size, shape, and spatial relations are common even in the absence of psychiatric disorders, especially when one is fatigued or excessively aroused. Illusions are misinterpretations of real sensory stimuli, as when a child in a dark bedroom at night sees monsters emanating from shadows on the walls. Pareidolia are playful and whimsical voluntary illusions that can be seen when one looks at ambiguously defined or evanescent images such as clouds or flames in a fireplace. Both the onset and termination of these perceptions are entirely voluntary. Trailing, another visual illusion, is the perception that an object moving steadily in space is followed by temporally distinct, after-images of itself. The effect is that of a series of stroboscopic photos. This phenomenon may occur with fatigue and is typically seen with marijuana and mescaline intoxication, during withdrawal from SSRIs, or, less commonly, in association with nefazodone (Serzone).

Hallucinations

Hallucinations are perceptions that occur in the absence of corresponding sensory stimuli. Phenomenologically, hallucinations are ordinarily subjectively indistinguishable from normal perceptions. Hallucinations are often experienced as being private so that others are not able to see or hear the same perceptions. The patient's explanation for this is typically delusional. Hallucinations can affect any sensory system and sometimes occur in several concurrently. When perception is altered, combinations of illusions and hallucinations, and often delusions as well, are frequently experienced together. In some studies, 90 percent of patients with hallucinations also have delusions, and approximately 35 percent of

patients with delusions also have hallucinations. Children and early adolescents, however, are more likely to have hallucinations in the absence of delusions. Approximately 20 percent of patients have mixed sensory hallucinations (mostly auditory and visual) that may accompany functional, as well as organic, conditions. A given external stimulus may evoke very different perceptual distortions in different people.

Three scientists floated in sensory deprivation tanks for long periods. One experienced a few illusions and no hallucinations; the second had many illusions and a few faint auditory and visual hallucinations; and the third had vivid, dramatic, and complex visual and auditory hallucinations.

Hallucinations are experienced by many normal people under unusual conditions. It has been estimated that between 10 and 27 percent of the general population has experienced memorable hallucinations, most commonly visual hallucinations. The large majority of self-reported hallucinations in community studies, particularly auditory hallucinations, have been associated with depressive and substance use disorders rather than frank psychotic disorders.

Hypnagogic and hypnopompic hallucinations are common, predominantly visual hallucinations that occur during the moments immediately preceding falling asleep and during the transition from sleep to wakefulness, respectively. Hypnagogic and hypnopompic hallucinations both occur in normal people and are also characteristic symptoms of narcolepsy. In acute bereavement, up to 50 percent of grieving spouses have reported hallucinating the voice or presence of the deceased, and after amputations, phantom limb hallucinations are common. Patients who become visually impaired often develop complex hallucinations (i.e., visual hallucinations with preserved insight) with preserved cognitive status, the so-called Charles Bonnet syndrome.

A 77-year-old woman presented to her psychiatrist with mild anxiety symptoms. During the past 1 year she had experienced severe loss in her visual acuity due to age-related macular degeneration. She reported experiencing unexpected vivid hallucinations and described seeing complex shapes of geometric design and animal figures. She had awareness that these "visions" were not real, and her cognitive screen, including the Mini-Mental Status Examination, was normal.

A parallel phenomenon is the emergence of hallucination, including musical hallucinations in individuals with acquired deafness. These observations suggest a "supersensitivity deprivation" hypothesis that, when deprived of important and anticipated perceptual stimuli, the mental apparatus may overinterpret any sensory stimulation as evidence of the presence of the needed objects.

A perceptual release theory suggests that hallucinations emerge from the combined presence of intense states of internal arousal and diminished

sensory input (including poor attention and poor capacity to sort out relevant from irrelevant input). Thus, diminished input from the environment (as in sensory deprivation) or reduced capacity to attend to and take in the input (as in delirious states) heightens the likelihood that internal sensations, images, and thoughts are interpreted as originating in the outside environment.

Hallucinations vary according to sensory modality, degree of complexity of the hallucinated experience, the levels of conviction about their reality, the clarity of their contents, the location of their sources of origin, the degree of volitional control over them, and the degree to which the hallucination influences the person's behavior.

Auditory hallucinations range in complexity from hearing unstructured sounds, such as whirring noises or muffled whispers, to ongoing multiperson discussions about the patient. The simple auditory hallucinations are more commonly associated with organic psychoses such as delirium, complex partial seizures, and toxic and metabolic encephalopathies. Auditory hallucinations are classically associated with schizophrenia (seen in 60 to 90 percent of patients) but are also frequently seen in psychotic mood disorders. Twenty percent of manic patients and less than 10 percent of depressed patients experience auditory hallucinations.

Three types of auditory hallucinations commonly associated with schizophrenia (which, however, are also seen less commonly in patients with psychotic depressions and mania) are (1) audible thoughts described as hallucinated voices that speak aloud what the patient is thinking, (2) voices that give a running commentary on the patient's actions, and (3) hearing two or more voices arguing with each other, often about the patient, who is referred to in the third person.

A 37-year-old man with chronic schizophrenia reported hearing multiple voices, both male and female commenting on his daily activities. He described the male voice as "critical" and commenting in a negative way about his grooming, clothing, and appearance. The female voice was calming and soothing. At times, he stated he could "block the voices out by putting on loud music." Other times he stated "they would take over," leading him to feeling despondent and hopeless. He stated during these "terrible times," he considered jumping off a bridge to end the "voices torturing me."

While auditory hallucinations in schizophrenia are frequently mood neutral, hallucinations in patients with mood disorders are characteristically consistent with their mood. In psychotic depression, the voices may be unrelievedly critical and sadistic, whereas in mania the voices often refer to the patient's specialness.

A 36-year-old woman with bipolar disorder described auditory hallucinations correlating

with episodes of severe mania or depression. When she was depressed, she reported hearing voices that were demeaning and called her "fat, ugly, and worthless." The voices were so unremitting that she became suicidal. However, during an intense euphoric manic episode, she described hearing joyful angels singing and declaring her as the next savior sent to earth by God.

Command hallucinations order patients to do things. Often, the commands are benign reminders about everyday tasks: "Pick up your shoes" or "Clean off the table." However, the voices may also be frightening or dangerous, commanding acts of violence toward the self or others such as, "Jump off the roof; you're not worth anything," or, "Pick up the knife and kill your mother." These voices vary in insistence and persistence, and patients differ in their capacities to ignore these commands. Patients with marked passivity may be helpless in the face of command hallucinations and may feel impelled to carry out the orders. Even though one study did not find command hallucinations to be associated with a higher risk of harm to the patient or others, the presence of command hallucinations and the patient's ability to resist must be assessed carefully.

A 46-year-old woman with schizophrenia heard command hallucinations on a daily basis for the last 20 years. When her symptoms were relatively quiescent, the voices were benign, and the commands referred to daily behaviors (such as "say thank you"). Under stress, her psychosis exacerbated, and the voices became louder and more insistent and commanded more complex and dangerous behaviors, once precipitating a self-stabbing when the voices demanded it.

Visual hallucinations occur in a wide variety of neurological and psychiatric disorders, including toxic disturbances, drug withdrawal syndromes, focal CNS lesions, migraine headaches, blindness, schizophrenia, and psychotic mood disorders. Although visual hallucinations are generally assumed to characteristically reflect organic disorders, they are seen in one-fourth to one-half of schizophrenic patients, often—but not always—in conjunction with auditory hallucinations.

Visual hallucinations range from simple and elemental, in which hallucinations consist of flashes of light or geometrical figures, to elaborate visions such as a flock of angels.

Stimulation of one sensory modality sometimes evokes perceptual distortions in another. Marijuana and mescaline intoxication, for example, have been associated with synesthesia, an experience in which sensory modalities seem fused. This is also a normal experience for many people. Music may be experienced visually, the sound fusing with visual illusions; a tactile sensation may be experienced as a color (e.g., a hot surface may "feel red").

In certain religious subcultures, visual hallucinations may be experienced as normal. In one fundamentalist Pentecostal church,

worshipers danced themselves into a frenzy, and, without using any drugs, several participants shared visions of the Virgin Mary at the altar.

A 62-year-old Native American man was raised on an Indian reservation and was deeply connected to his heritage. He openly acknowledged that the spirit of his elders and ancestors surrounded him. He reported that he could feel and sense their presence around him no matter what he was doing, whether he was working or resting. The presence of these spirits was a guiding force that he felt was calming to him.

Autoscopic hallucinations are hallucinations of one's own physical self. Such hallucinations may stimulate the delusion that one has a double (doppelganger). Reports of near-death, out-of-body experiences in which individuals see themselves rising to the ceiling and looking down at themselves in a hospital bed may be autoscopic hallucinations. In Lilliputian hallucinations, the individual sees figures in very reduced size, such as midgets or dwarfs. They may be related to the perceptual distortions of macropsia and micropsia, respectively, the perceptions of objects as much bigger or smaller than they actually are.

Haptic hallucinations involve touch. Simple haptic hallucinations, such as the feeling that bugs are crawling over one's skin (formication), are common in alcohol withdrawal syndromes and in cocaine intoxication. When unkempt and physically neglectful patients complain of these sensations, they may be due to the presence of real physical stimuli, such as lice. Some tactile hallucinations—having intercourse with God, for example—are highly suggestive of schizophrenia but may also occur in tertiary syphilis and other conditions and may, in fact, be stimulated by local genital irritation. Olfactory and gustatory hallucinations, involving smell and taste, respectively, have most often been associated with organic brain disease, particularly with the uncinate fits of complex partial seizures. Olfactory hallucinations may also be seen in psychotic depression, typically as odors of decay, rotting, or death.

A 32-year-old woman with schizophrenia described that, over the 15 years of her illness, she had experienced a wide array of hallucinations. She initially noticed celestial voices, auditory hallucinations that included angelic music and comforting words. However, over the years, in conjunction with a pessimistic turn of mind, the voices became more demonic and threatening. At the same time, she began to experience somatic hallucinations, consisting of sharp, intermittent electric currents being run through her skin and genitals, which, she was convinced, were the work of evil neighbors intent on driving her out of her apartment. With medication, she was partly able to perceive these experiences as hallucinatory; however, most times, particularly when she neglected to take medications, she never doubted their authenticity.

The term pseudohallucination has been used in two ways. First, pseudohallucination refers to perceptions experienced as coming from

within the mind (i.e., not at the boundary or outside the mind). Using this definition, loud voices that are alien, ascribed to other beings, but which the patient knows are actually within the mind rather than out in space, are pseudohallucinations. The term has also been used to describe hallucinatory experiences whose validity the patient doubts. A better term for this second phenomenon is partial hallucination, analogous to partial delusion. Functional hallucinations are rare hallucinations that occur only in connection with a specific external perception, for example, in the presence of a sound, such as running water, a color, or a particular place. However, unlike illusions, the hallucinated sounds are not elaborations of the perception but are simply triggered only in that specific context.

Ictal hallucinations, occurring as part of seizure activity, are typically brief, lasting only seconds to minutes, and stereotyped. They may be simple images, such as flashes of light, or elaborate ones, such as visual recollections of past experiences. While the hallucinations are being experienced, the patient ordinarily experiences altered consciousness or a twilight sleep.

Migrainous hallucinations are reported by approximately 50 percent of patients with migraine. Most are simple visual hallucinations of geometrical patterns, but fully formed visual hallucinations, sometimes with micropsia and macropsia, may also occur. This complex has been called the Alice in Wonderland syndrome after Lewis Carroll's descriptions of the world in *Through the Looking Glass,* which mirrored some of his own migrainous experiences. In turn, these phenomena closely resemble visual hallucinations induced by psychedelic drugs such as mescaline.

A flashback is an intense visual re-experience of highly charged past events, which are often replays of hallucinations. They are typically associated with heavy use of hallucinogens, such as LSD and mescaline, and often occur months after the last drug ingestion. The images may be simple or complex geometrical patterns, or they may consist of previously experienced elaborate drug-induced hallucinations. Flashback phenomena may be state dependent. For example, visual hallucinations initially experienced with hallucinogens are more likely to be subsequently experienced as flashbacks when the subject is smoking marijuana. In PTSD, some complex, intrusive flashbacklike images may attain a hallucinatory vividness. Images often include horrifying memories of traumatic events that may force themselves repeatedly into consciousness until they are acknowledged and worked through.

Hallucinosis is a state of active hallucination occurring in someone who is alert and well oriented. This condition is seen most often in alcoholic withdrawal, but it may also occur during acute intoxications and other drug-mediated states.

A 21-year-old man used LSD daily for 3 days, after which the visual hallucinations that first occurred while under the influence of the drug waxed and waned. He described pareidolias with distracting images of animals when he looked at traffic lights, the sky, or billboards. Trailers were common after he turned his head. Finally, he saw moving colors when he stared at any object for more than a transient time period. He assumed that these phenomena represented a lingering effect of the LSD and was not frightened by them. The hallucinations disappeared within 1 week.

Body Image Distortion

Body image includes both perceptual and ideational components and may reflect primarily perceptual distortions or combinations of disturbed perception and self-appraisal, or both.

Body image disturbances can occur as normal responses to abrupt changes in the body (e.g., after amputation), in brain disease, and in psychiatric disorders. Phantom limb phenomena are classic body image problems in which an amputated limb is still felt to be present. The sensation may diminish gradually over time; the phantom feels as if it is receding into the stump.

A 34-year-old man presented to his VA hospital mental health clinic appointment for persistent insomnia and anxiety. While on his second tour of active combat duty, he sustained a traumatic injury requiring amputation of his left leg. He reluctantly disclosed that he could feel his amputated leg "going in and out" and thought he was "going crazy." He described many sensations, sometimes that the area felt warm and tingling, other times, he felt throbbing, itching and sharp electrical piercing feelings.

Agnosias, lack of awareness of some parts of the body, may accompany brain damage, most often of the nondominant parietal lobe. Patients with obvious motor or sensory deficits may deny that any deficit exists at all (anosognosia), or the denial may be limited to one-half of the body (hemiagnosia), usually the left side. In hemidepersonalization syndromes, a less common disorder (hemiasomatognosia), patients feel that one of their limbs is missing, again usually on the left side. Body image distortions in which a limb feels too heavy (hyperschemazia) or weightless (hyposchemazia) can occur as a consequence of neurological conditions such as infarction of the parietal lobe. In duplication phenomena, patients feel as if part or all of them has doubled (e.g., they have two heads or two bodies). These rare phenomena may occur in schizophrenia, complex partial seizures, and migraine.

Dysmorphophobia refers to conditions in which patients distortedly perceive and intensely dislike the shape of a particular body part. As such, these symptoms are misnamed because there is no true phobic component such as fear or avoidant behavior. Fine lines exist between perceptual distortions and realistic but unhappy appraisals of one's body, given the

high social values placed on physical appearance. Dysmorphophobia may occur in the context of some personality disorders or as an isolated disorder called body dysmorphic disorder. In some ways, dysmorphophobia resembles an overvalued idea. Patients may develop dysmorphophobias in relation to any body part; common concerns are hair, breasts, penis size, the shape of the nose, or the shape of the entire body. For some, changing the body part, as in rhinoplasty for those who do not like their nose, seems to effect a lasting positive change in body image, with patients becoming happier with themselves and feeling more attractive for years or lifetimes. In severe dysmorphophobia, patients may undergo multiple plastic surgeries and feel dissatisfied with every result. At times, the condition forms part of a larger and more pervasive syndrome, such as anorexia nervosa.

An attractive 19-year-old woman was preoccupied and perplexed by the fact that she perceived her eyelids as being unequal in size and shape and by the fact that the cleft in her chin was too deep. She spent hours each day in front of the mirror, trying to pose in ways that, from her perspective, did not expose these facial blemishes. She sought consultation from a plastic surgeon, who referred her for psychiatric consultation. After 2 months of treatment with an SSRI, her preoccupations with appearance diminished considerably, and she was able to turn her attention to other issues in her life.

Hypochondriacal complaints also combine perceptual and ideational distortions. Selective hypervigilance to bodily sensations may result in a higher likelihood of perceptions of unpleasant and potentially pathological body experiences among the "worried well," hypochondriacal populations, patients with somatization disorder (Briquet syndrome), and some patients with panic disorder.

Body image distortions may, at times, be severe or bizarre. Some psychotic patients, either with schizophrenia or depression, develop somatic delusions. In depression, this often expresses itself as a delusion that part of or the entire body is rotting or filled with cancer. Some culture-bound syndromes in non-Western cultures express themselves with body image distortions such as koro, in which the man fears that his penis is shrinking into his abdomen.

DISTURBANCES OF CONSCIOUSNESS

Consciousness can be defined as subjective awareness of the self and environment. Biologists increasingly believe that a continuum of consciousness exists, extending from lower animals through *Homo sapiens*. However, consciousness is subject to conflicting definitions and conceptualizations, and exactly where consciousness begins in evolution remains unclear. Philosophers agree that it is the subjectivity of experience, the so-called qualia of consciousness, that clearly distinguishes living

consciousness from at least this generation's best versions of self-regulating automata, elegant computers, or robots. All current attempts to even approach an understanding of consciousness are very unsatisfying, and consciousness remains unexplained and as yet unexplainable from a scientific point of view. Consciousness has been viewed as an emergent property of complex biological nervous systems, as a poorly understood general property of an even more mysterious and complex universe, or as a phenomenon to be understood only in religious and spiritual terms. One of the best analyses to date of possible relationships between the construction of a brain and the possibility of consciousness has been set forth by Gerald Edelman, who believes that reflective consciousness cannot occur until complex higher-order brain systems evolve whose major functions are to monitor the experiences, activities, and results of activities of those lower-order brain systems that deal directly with appraising and responding to the external and internal environments. Such higher-order metasystems require the presence of memory so that current and immediate impressions can be checked and compared against past experiences. These metasystems may use a variety of sensor mechanisms to detect and signal their sensations or perceptions of various events. Some of these sensors may correspond to feeling states, and some may correspond initially to preverbal thoughtlike mechanisms that contain the capacity to develop and recognize abstract categories and, ultimately, conceptual language-based thought.

Clinically, consciousness can be considered from both qualitative as well as quantitative viewpoints. Qualitatively, consciousness does not seem to be an all-or-none phenomenon. Rather, conscious experiences may gradually and phasically shift in focus, intensity, and clarity; altered states of consciousness may occur in which some aspects of consciousness, such as sensation, perception, memory, orientation, and judgment, are enhanced or impaired relative to other aspects. Quantitatively, crude divisions can be made between states depending on the relative presence, impairment, or total absence of consciousness. Even within the single individual, consciousness is not a unitary phenomenon. Multiple streams of thought, operating at multiple levels of preconsciousnesses, appear to exist in all people almost all of the time, with various elements in these coexisting streams constantly shifting into higher or lower levels of conscious awareness. In pathological states, even more remarkable properties of consciousness are seen, for example, the existence of coconsciousness in humans who have had commissurotomies and of seemingly multiple discrete consciousnesses in patients with dissociative identity disorders.

Experiments involving patients with commissurotomies of the corpus callosum have shown the existence of two virtually separate systems of consciousness that seem to operate side by side. For example, when in the

course of an experiment, the picture of a nude woman was flashed only to the right brain (the left visual field) of a commissurotomized patient, the subject verbally denied being aware of anything unusual (i.e., the left brain —the verbal brain—was unaware). But, at the same time, he started to squirm and blush, blurting out, "Oh, you have some machine!" Similarly, when a cup was presented to the right brain (left visual field) only, the patient denied seeing anything (left brain was unaware, and language output of the left brain indicated no awareness), but he was able to pick out the cup from an assortment of objects with his left hand (right brain control). That literal splitting of verbal awareness from visual–spatial awareness in the brain produces behavior that is at least superficially similar to that of patients who deny being consciously upset by an event but who react with strong visceral responses. Although this formulation is simplistic, the separate consciousness for logical–verbal and for visual–spatial awareness demonstrated in split brain experiments may be crude analogs for more highly differentiated and discrete types of awarenesses and modes of information processing. Furthermore, the very fact that there are separate and, to some extent, competing modes of consciousness may increase the likelihood of psychological distress because the various modes are capable of yielding internally conflicting views of reality.

Psychological and Physiological Factors

In ordinary states of alert consciousness, individuals are able to deploy adequate amounts of attention to their surroundings and to reflective thought. Normal people vary enormously in their ability to pay careful attention in different settings without being distracted; individual variations may reflect temperamental and cognitive style differences as well as physiological shifts within the individual. Many functions of attentive consciousness, including attention, planning, and the capacity to appropriately switch between mental tasks, the so-called distractor-resistant memory, have been linked to the activity of specific neurons in area 46 of the prefrontal cortex.

A sense of increased consciousness with heightened alertness, awareness, and sharper thinking may be experienced in states of highly aroused emotional states such as threat, sexual attraction, falling in love, or other high-stakes events such as hunting among primitives, sporting competitions, or performing in front of an important audience. High levels of arousal do not necessarily guarantee effective attention because optimal consciousness depends on optimal arousal. Too little arousal due to illness or fatigue may result in insufficient stimulation and mental lethargy, diminishing the sense of alertness and attentiveness, whereas too much arousal may result in hyperintense alertness but distractibility and scattered attention.

Consciousness involves, among other things, the experience of a

continuous sense of self and of the environment, existing coherently in time and space. The experience of time and its passage may be altered by shifts in the level of awareness and by emotional states such as boredom, concentration, pain, and discomfort. The experience of time and space may be altered by hypnosis, marijuana, psychoactive and psychedelic drugs, and other events directly affecting brain physiology.

Disturbances of Orientation

Orientation refers to one's awareness of time, place, and person. Accurate orientation requires the integrity of attention, perception, memory, and ideation. Impairments occur primarily in organic mental disorders (i.e., structural and toxic metabolic brain abnormalities) and occasionally in dissociative and psychotic states.

Normal individuals vary tremendously in their attention to the details of time and in the extent to which their bodies automatically keep time. Some people have reliable built-in clocks by which they can awaken themselves at precise times or accurately gauge the passage of time with uncanny accuracy, even in the absence of external cues—in a psychotherapy session, for example. Others have difficulty making judgments about time and may develop pathological lateness or habitually schedule more activities than could ever be accomplished in the available time. Poor time judgments may be seen in a variety of psychiatric disorders, such as ADHD, or as an independent problem. Benign disorientation to time is common. After a few days in a hospital bed, most people do not know exactly what the day or date is because they are not attending to or receiving their usual cues.

Pathological time disorientation can be mild or severe, with inaccuracies of estimation ranging from days to years. The dates reported by disoriented individuals may have personal significance such as those of important births, marriages, or deaths.

Because spatial cues are generally more available and obvious than temporal cues, disorientation to place often signifies a greater degree of cognitive impairment than disorientation to time and, therefore, rarely occurs in the absence of time disorientation. Disoriented people may know, more or less, the type of place that they are in without knowing the specific place—patients may recognize that they are in a hospital without being able to name the hospital.

An 81-year-old retired engineer with severe dementia was pleasant, even chatty when asked about routine daily activities. He seemed to follow conversations, nodding his head and smiling politely. When his daughter came to visit him on her birthday, he smiled pleasantly, and kept referring her as "Aunt Matilda." He looked confused and anxious when she adamantly insisted that was not her name and could not recall the year and date of his daughter's birthday.

Disorientation to person, a lack of awareness of one's own identity, is typically seen only in advanced dementias such as primary degenerative dementia of the Alzheimer type or in dissociative states. In organically induced postconcussion amnesia, transient global amnesia, and, presumably, psychogenic fugue states, knowledge of one's own identity may disappear, and a person may remain unidentified for an indefinite period until the memory for self-returns.

Disturbances in the Level of Consciousness

Levels of consciousness (i.e., alertness, awareness, and attentiveness) may be pathologically increased or decreased. Such changes are diagnostically nonspecific and can occur in many different disorders. When levels of arousal and alertness are mildly elevated, as in hypomania or with the ingestion of small amounts of psychostimulants, subjective experiences are typically positive. In these situations, the person experiences intense alertness, prolonged concentrating ability, and hyperesthesias in which perceptual vividness is heightened: Colors are brighter, sounds are sharper, and touch is more intense than usual. With further increases in arousal and consciousness as seen in mania, more severe intoxications with amphetamines and cocaine, and catatonic excitement, attention fragments. Heightened alertness transforms into hypervigilance and paranoia, and hyperesthesias become unpleasant.

Diminished levels of consciousness can be described on a continuum. Clouding of consciousness is marked by diminished awareness of sensory cues and diminished attentiveness to the environment and to the self. Secondary process thinking is most notably compromised, and more primary process thinking emerges into consciousness. In this state, one's ability to appreciate subtleties and to think in a nuanced manner is diminished and is replaced by more dichotomous all-or-none, stereotypical thinking. The level of consciousness may fluctuate rapidly in relation to the internal physiological state or to the degree of external stimulation. In alterations of consciousness, confusion may occur with disorientation to time, place, or person. The patient is usually highly distractible and unable to pay sustained attention to a single stimulus.

Torpor is a condition in which the patient is drowsy, falls asleep easily, and shows a narrowed range of perception and slowed thinking. Stupor is a state of diminished consciousness in which the patient remains mute and still, although the eyes are open and may follow external objects. In the most extreme impairment of consciousness, coma, there is no evidence of mental activity at all. The patient essentially appears to be functioning on a decorticate or decerebrate level. In akinetic mutism, or coma vigil, patients with profound brainstem lesions appear to be awake with their eyes open, but there is in fact no evidence of consciousness. The commonly used Glasgow Coma Scale incorporates these dimensions.

Delirium, the acute confusional state, is usually characterized by a relatively abrupt onset and short duration of clouded, reduced, and fragmented attention; impaired memory and learning; perceptual and cognitive abnormalities, such as hallucinations and delusions; disrupted sleep; and other autonomic dysfunction. The level of consciousness may be consistently diminished or may fluctuate. The electroencephalogram (EEG) usually shows diffuse slowing. Typical motor abnormalities include an increase in general restlessness, fine and course tremors, and myoclonic jerks. Autonomic disturbances commonly include tachycardia, fever, elevated blood pressure, diaphoresis, and pupillary dilation. The causes of delirium are legion, including systemic medical disorders, such as metabolic imbalances or infections, intracranial disorders due to traumatic, structural, and electrical causes, and drug intoxication and withdrawal states.

A 59-year-old truck driver at work collapsed while complaining of severe chest pain. He was rushed to the hospital, and it was determined that he needed emergency coronary bypass surgery. His surgery was successful, and he was slated for discharge the next morning. That night, the resident physician was urgently called to the floor by the nurse taking care of the patient who stated that the patient was sweating profusely, confused, agitated, scratching his arms, and shouting that "spiders were crawling all over him."

Attentional difficulties are manifested by impairments in the person's ability to deploy, focus, and sustain attention. Some attentional difficulties first appear in early childhood as developmental problems of uncertain cause and are described as attention-deficit disorders (inattentive or hyperactive types). Secondary attention-deficit difficulties may appear de novo in adulthood due to a variety of exogenous agents, psychiatric disorders, and late-life developmental and degenerative factors.

In narcolepsy, characterized by sudden lapses into sleep, one's usual ability to stay alert and maintain consciousness is impaired. At times, the onset of profound sleepiness is gradual, accompanied by hypnagogic phenomena, in which dreamlike images invade consciousness, and at other times the shift in consciousness appears to be almost instantaneous. This syndrome, occurring in approximately 1 in 10,000 persons, is thought to be the second most frequent cause of automobile collisions after alcohol intoxication.

A 24-year-old, recently engaged woman presented with her fiancé for evaluation. Her fiancé reported that while shopping for her engagement ring, she was very excited and had an "attack" of being weak in her knees and falling down lasting almost 1 minute. She reported fully being aware of her surrounding and hearing the clerk and her fiancé calling her name. During the evaluation, she acknowledged significant daytime sleepiness and that her mother suffered from narcolepsy.

Altered States of Consciousness

Consciousness may also be qualitatively changed, with the production of altered states. Drugs such as scopolamine (Transderm) with strong central anticholinergic properties, some seizures, and, on occasion, other conditions associated with delirium can induce twilight states, dreamlike states of wakeful consciousness in which attention is poor, an admixture of primary and secondary process thinking appears, and patients fade in and out of alertness. Dreamlike experiences intrude into the stream of conversation. Emotional outbursts or violent acts may occur during twilight states.

Mystical states of consciousness may occur in normal and pathological conditions. Intense meditation and peak or epiphanic experiences, reported by more than 10 percent of normal individuals in community surveys, may produce a sense that the self dissolves or expands, that the self fuses mystically with the cosmos, that time stops, and that universal meaning becomes clear. These perceptions may be accompanied by a sense of rejuvenation and renewed personal identity, ineffability, intense emotionality, and concurrent perceptual changes. Such experiences do not ordinarily last more than a few minutes. Many people have achieved these states through the use of psychedelic agents such as mescaline and LSD. Reports of a white light at the end of a tunnel, described by individuals using psychedelics and in near-death experiences, have been linked to specific neurophysiological pathways believed to be stimulated under these conditions.

A 43-year-old female accountant with no prior history of spiritual or religious practice sustained a severe head injury in an automobile accident that left her comatose for approximately 6 months. When she awoke, she reported that at some point during her coma, she estimated in the third month, she experienced a profound mystical experience, which included a glorious union with God and knowing that she would not only awaken to return to life but that all was “as it should be” with the universe. After recovery, with ongoing motor impairment, she became profoundly spiritual, started attending church and teaching the Bible, and constantly demonstrated a beatific glow, which she attributed to her mystical experience.

Although hypnosis lacks a consensually accepted definition, its hallmarks are selective attention, suggestibility, and dissociation. Most, but not all, people can be hypnotized to some degree. Up to 90 percent of people are capable of achieving a light trance, whereas 10 to 20 percent are capable of entering a deep trance and exhibiting remarkable hypnotic phenomena. Hypnosis occurs when the subject is in a state of heightened, not diminished, attention. EEG studies have shown hypnotized subjects to be fully awake and alert. The heightened concentration probably accounts for the unusual levels of sensory and motor performance often seen under

hypnosis and self-hypnosis.

Hypnotic phenomena include hypnotically induced hallucinations (including negative hallucinations in which the subject selectively does not perceive sights, sounds, or other stimuli), anesthesia, sustained motor behaviors and acts of strength ordinarily beyond the individual's capacity, and distortions of memory (both hypermnesias and amnesia). Several phenomena that reveal the multiple nature of consciousness, for example, coconsciousness, are also demonstrable. Experiments have shown that even when a subject in deep trance has achieved profound hypnotic anesthesia and can, for example, keep a hand submerged in ice water for longer periods of time than usual, part of the hypnotized subject's consciousness continues to register exactly how painful the experience actually is and can signal the researcher about the pain, by finger movements, without the subject having any conscious awareness or disturbance. This phenomenon, called the hidden observer, has also been seen in postsurgical patients who, under hypnotic trance after surgery, have been able to accurately recall conversations in the operating room that occurred while they were under general anesthesia. Dissociative and psychosomatic phenomena have also been induced with hypnosis. With posthypnotic suggestion, for example, subjects may carry out complex actions without any hint that they are actually doing so because they were previously instructed to act that way under hypnosis. When asked why they are carrying out these activities, such subjects usually make up various reasons, although seemingly unaware of the real reasons for their actions. It has been suggested, not entirely facetiously, that many normal daily activities are conducted in a trancelike posthypnotic state, and although these activities are attributed to conscious intention, they may in fact be carried out due to previous suggestion. Advertisers know this well. Urticaria (hives) can be hypnotically induced and hypnotically made to disappear. Plantar warts have been successfully treated with hypnosis, and, in these conditions, diminished blood supplies to their bases have been demonstrated. It has recently been appreciated that yoga masters can exert remarkable control over basic bodily functions through self-hypnosis. As yet, little is known of the full extent to which heightened concentration may influence physiological regulation.

Suggestibility

Pathological suggestibility may be seen in several clinical conditions. Automatic obedience has been described in echolalia (the automatic repetition of a sentence or phrase just uttered by another person), echopraxia (the automatic mimicking of a movement performed by another person), and waxy flexibility (maintaining for a prolonged period of time a posture in which one is placed), symptoms common in catatonic states. In situations of group delusions, and sometimes in cults, passive individuals

adopt the delusional beliefs of stronger ones. In epidemic hysteria, as described among young women at the Salem witch trials in Arthur Miller's *The Crucible,* distorted and even delusional perceptions and beliefs may sweep over a group that has been highly aroused by a charismatic leader.

Autosuggestibility can be seen in the constructions of false memories in which an individual progressively comes to believe that something that never happened in fact occurred. Such false memories may be held with such great conviction that they are indistinguishable from the memories of real events. Because of this, memories recovered during therapy (in which there is often great pressure to "remember" certain events) cannot be taken at face value without corroboration from other sources. Various types and degrees of self-deception may be more common in individuals who are more suggestible.

Dissociative Phenomena

Dissociation refers to the splitting off from one another of what are ordinarily closely connected behaviors, thoughts, or feelings. Dissociative states are those in which there is a disturbance or alteration in the normally integrated functions of identity, memory, or consciousness and include trances, fugues, blackouts, multiple personalities (dissociative identity disorder), and dissociative frenzies. Although dissociative states are ordinarily believed to be functional in nature, arising as an adaptive defense in individuals subjected to a great deal of trauma, particularly at early ages, they occur regularly with a variety of neurological disorders, particularly those with partial complex seizures. In one series, one-third of patients with complex partial seizures had dissociative phenomena, including multiple personality. In these patients, the dissociative phenomena were not related to the seizure activity but to interictal alterations.

As in posthypnotic amnesia, elaborate activities can occur in dissociative states for which the subject has no conscious memory. This amnesia is functional in nature and may be reversed by hypnosis or drug-facilitated disinhibition, for example, with amobarbital (sodium amytal) infusion. In many of the functional dissociative states, amnestic episodes may occur for years or decades before the patient seeks medical or psychiatric attention. Blackouts are periods of amnesia in alcoholism, other intoxications, or after head trauma. An alcoholic blackout period may last for hours to days, after which the person has no recollection of what transpired, although other observers attest to the fact that, during this period, the individual carried out multiple complicated behaviors. Although memory of the blackout is lost to the predominant consciousness, during subsequent reintoxication, memories of events occurring during the previous blackout may be reawakened. This phenomenon, known as state-dependent memory, occurs in many other conditions as well, signifying

that one's ability to retrieve specific memories may be highly influenced by specific physiological alterations due to external intoxicants or other unusual physiological states.

A 43-year-old man with a 20-year history of sustained alcohol abuse was unable during a period of sobriety to recall for hospital staff the names, phone numbers, and addresses of his close friends and associates. Several weeks later, when he returned to the emergency department in an intoxicated state, he was able to recall all of those details in a reasonably straightforward manner.

Psychogenic fugue is characterized by prolonged periods in which individuals carry out very complex activities without having any recollection for their previous lives, identities, or even names. They often travel away from customary locales and assume entirely new identities. By definition, psychogenic fugue cannot be due to a neurological disorder. In comparison, the discontinuity of experience in psychogenic amnesia is typically more circumscribed and does not involve assuming an entirely new identity. The dissociated memories and affects often reveal themselves in disguised form such as nightmares, intrusive visual images, and conversion symptoms. Typically, in psychogenic amnesia, an individual may not be able to recollect what transpired during a specific period, for example, before the age of 9 or 10 years in the context of a traumatic childhood; during catastrophic events, such as traumatic, gruesome combat; or during less momentous events that a person prefers to forget to preserve self-esteem by denying shameful, immoral, or illegal activities.

A 28-year-old infantryman could not recollect the events of the combat situation in which he had been wounded and several of his comrades killed. Several months afterward, he participated in a series of amytal-facilitated interviews, during which the events of combat came back to him. He recalled becoming panicky during the firefight, turning, and running away while several other soldiers in his squad, some of whom were killed, yelled to try to get him to stay in place and fight. After the interviews, he cried inconsolably and blamed himself, somewhat irrationally, for their deaths, wondering if his remaining in place and fighting might have saved their lives.

Dissociative Identity Disorder. Dissociative identity disorder, previously known as multiple personality disorder, is a chronic, dissociative state in which two or more separate ongoing identities or personalities alternate in consciousness. It usually occurs in people who, as young children, were severely and repeatedly brutalized. The number of identities is variable, with some cases reporting 25 or more identities. The development of dissociated alter personalities is believed to be a last-ditch, primitive psychological defense against inescapable and unbearable traumatic situations. The personalities may be of different ages and even different

sexes. Typically, the presenting identity is dysphoric, anxious, and constricted, may have headaches and periods of blackout or amnesia, and is not aware of the other personalities. A second identity is typically vivacious and uninhibited. Another identity may know all about the other personalities and has a wise perspective on the life events leading to the problems and of possible solutions. A classic case is described in the popular book and film *The Three Faces of Eve*.

In so-called channeling, dissociated, complex part personalities are produced in trance states in which fictitious past lives or spirit lives are created.

Ganser Syndrome. In Ganser syndrome, the patient responds to questions by giving approximate or patently ridiculous answers; for example, in response to the question, "what sound does a dog make?" the patient answers, "moo." Additional features of the syndrome include alterations in consciousness, hallucinations (or pseudohallucinations), conversion phenomena, and amnesia for the episode during which these symptoms are manifested. This syndrome has most commonly been reported in prisoners and is generally believed to be a dissociative state, although organic features may be contributing factors.

A 23-year-old man was arrested for auto theft. He had a long history of brushes with the law, including the juvenile justice system for fire-setting, shoplifting, and physical fights. While in jail, he stated he was confused, could not remember what led to his arrest, appeared distracted at times, pleasant and smiling at other times, and would answer simple questions such as 1 + 1 as equal 3.

Depersonalization and Derealization. Depersonalization refers to an alteration in one's experience and awareness of the self, leading to feelings of being unreal or detached from one's own body or of feeling like an automaton; it is often accompanied by the complaint that the individual lacks all feelings or sensory experiences. Those experiencing depersonalization frequently fear that they are going crazy. Because of this fear, patients often endure depersonalization experiences for long periods before describing them to a mental health professional. Depersonalization is also characterized by frequent internal nonaudible dialogues between the participating self and the observing self but with full awareness that both parties are the same person (a feature that distinguishes it from hallucinations). Mild sensory distortions—but not hallucinations—are commonly associated with the experience. Depersonalization is seen in a variety of neurological and psychiatric disorders and is common in complex partial seizures. It may occur in the context of depression, anxiety disorders, or certain personality disorders, or it may occur as an entity by itself. In derealization, individuals feel themselves to be real but feel that

the world around them has suddenly become unreal. Derealization often, but not always, accompanies depersonalization. Transient episodes of depersonalization and derealization occur frequently in normal people, particularly during states of fatigue, sleep deprivation, or stressful situations such as bereavement, learning of a terminal diagnosis, or sudden awareness that one is about to be in an inescapable vehicle collision.

A 42-year-old school bus driver was unable to see a large object in the road in a severe thunderstorm and had an unavoidable collision. Although she was able to swerve the school bus at the last minute, the bus sideswiped a vehicle before coming to a screeching halt. When the police arrived, she could only say that she felt that she was "floating outside her body" and that the "world did not seem real."

DISTURBANCES OF THE SELF

At the most basic level, the key components of self-awareness are the reality and integrity of the self (that I am one person), the continuity of self (that I am the same person now that I was in the past and that I will be in the future), the boundaries of self (that I can distinguish between myself and the rest of the world as not-self), and activity of self (that it is I who is thinking, doing, and feeling). Additional components of a sense of self include body image and various self-evaluations, including self-esteem and ego-ideal (ideal self). Body image is an individual's mental representation of his or her own body. Self-esteem is believed to reflect how one measures up to the desired self-image. To the extent that what one sees in oneself approximates what one would like to be, self-esteem is positive. Ego-ideals are fantasies of the optimum person one could ever wish to be. Any of these qualities may be disturbed in psychiatric disorders.

Within each individual is a group of social selves comprised of the roles and identities that a person assumes and that are evoked in various contexts. The presenting "self" varies depending on the people with whom one interacts, such as parents, romantic partner, child, friend, or employer, and depending on what role one assumes, such as child, parent, colleague, or lover. Accompanying each of these "selves" are various levels of objective and subjective self-awareness and self-understanding.

Disturbances in Sense of Self

Disturbances of the basic elements of self-awareness may be seen in a variety of disorders. Discontinuity phenomena are characteristic of dissociative states such as psychogenic amnesias and psychogenic fugue. Depersonalization reflects a mild disturbance in the awareness of self as the agent of activity. More severe disturbance is characteristic of the psychotic passivity phenomena seen in schizophrenia. Boundary disturbances may be considered characteristic of all psychotic states regardless of diagnosis.

Disorders of self-integrity are characteristic of both dissociative identity disorders and severe borderline personality disorders in which a person's self-concept and expression of this concept to others is erratic, leading to a sense of unstable identity.

False self describes a persona or a faulty and limited superficial aspect of the personality that an individual builds up as a mechanism for adapting to a hostile world—to please, control, or negotiate with others and with him- or herself. However, the false self does not incorporate, integrate, or validate important fundamental needs, wants, values, and beliefs. Through self-deception and denial, the individual may consciously believe that this "self" constitutes his or her entire being. However, the false self is a relatively fragile construction that has usually warded off and denied fundamental strivings, which may include needs for autonomy; acting with integrity; expressing certain desires, beliefs, or talents; or other unacknowledged aspects of the self. When these warded-off needs finally break through and demand expression at various points in development, the defective false self may collapse, leading to a period of distress and identity confusion, which individuals sometimes describe as a "nervous breakdown."

Patients with pseudologia fantastica and the impostor syndrome demonstrate extreme examples of inconsistency in the sense of self. In pseudologia fantastica, patients compulsively spin out webs of lies, ordinarily self-aggrandizing ones, and also appear to be trying very hard to deceive themselves into believing that they are true. In the impostor syndrome, such fantasies are acted out by liars and impostors who seem to fervently wish that these fantasies were their reality, as if they cannot accept themselves and would be overwhelmingly ashamed to be known for who they actually are. The impostor compulsively adopts the identities of others and may, for example, show up properly attired at diplomatic functions and society galas and interact with the other guests under the assumed identity. Some famous impostors have repeatedly insinuated themselves into inner circles of high society and government.

Transsexualism is a syndrome characterized by the feeling that one was born into a body of the wrong sex and marked by the desire, starting at an early age, to be a person of the opposite sex. Male to female transsexualism is reported most often. Both psychodynamic and biological theories have been advanced to explain these unusual phenomena. Of note, studies following patients who undergo successful male to female transgender surgery have not revealed higher rates of associated psychopathology than in comparison groups.

Self-esteem is a measure of one's self-appraisal in relation to one's values and ego-ideals. Negative self-esteem is characteristic of depressive disorders, many personality disorders, and situational failures. Superficially inflated self-esteem may be seen in mania (or hypomania) or,

in a fluctuating manner, with narcissistic and other personality disorders.

Although some individuals regard their ego-ideals as unattainable and are content to live as imperfect human beings, others strive to approximate their ideals. People who feel driven to achieve unattainable ideal goals or to become unrealistically perfect ideal selves are likely to be chronically dysphoric and have poor self-esteem because their attempts to become their ego-ideals are doomed to failure.

DISTURBANCES OF MEMORY

Memory is not a unitary phenomenon. Capacities to remember vary for the different senses and perceptions. One person may have prodigious musical memory, with the capacity to remember and reproduce whole musical pieces after one hearing, but be incapable of remembering people's names or telephone numbers. Exceptionally detailed verbal memories have been associated with obsessional cognitive styles. When individuals with extraordinary memories complain of memory loss, ordinary memory tests may be inadequate to detect their deficits, as their relative memory loss may have reduced their capacities to a point within the range of most normal people.

Memory functions have been divided into three stages: Registration, retention, and recall. Registration (or acquisition) refers to the capacity to add new material to memory. The material may be sensory, perceptual, or conceptual and may come from the environment or from within the person. For new material to be acquired, the person must attend to the information presented, and it must then be registered through the appropriate sensory channels and then be processed or cortically organized. Retention is the ability to hold memories in storage. Large numbers of neurons are believed to be involved in the storage of a specific memory, and it is believed that reverberating circuits are formed in which memory traces are held by means of changes in proteins or synaptic connectivity, or both. Recall is the capacity to return previously stored memories to consciousness.

Newly registered material is transferred incrementally from immediate to short-term memory to long-term memory. Immediate memory lasts for 15 to 20 seconds, short-term memory (or recent memory) for several minutes up to 2 days (the time involved in new learning and its early consolidation), and long-term (or remote) memory for longer periods. Different physiological processes mediate each of these stages of memory. Because of this, processes that affect immediate or short-term memory often spare long-term memory. The processes by which memories are transferred from short- to long-term stores are unknown.

Cognitive scientists now refer to short-term memory as working memory, the system that briefly stores and processes information needed for planning and reasoning. Recent studies suggest that the working memory system consists of at least two short-term memory buffers, one for

verbal and another for visual memories, plus a central executive that manipulates and coordinates information stored in the two buffers for problem solving, planning, and organizing activities. Parallel processing systems involving specific areas of the prefrontal cortex and other brain areas appear to operate separately with respect to various processes concerned with working memory. For example, separate prefrontal areas appear to be involved in working memory functions concerned with object identity and spatial locations.

Other studies suggest that different types of memories are stored and retrieved by different brain systems, so there is at least a dual memory system. The first system, sometimes called a conditioned-emotional system, system for implicit memory, perceptual memory, or nondeclarative memory, is present from birth, operational through life, and addressable by situational, sensory, or affective cues. Past experiences are expressed through images, behaviors, or emotions. These memories need not involve any conscious memories of a past experience. Conditioned fear responses represent examples of memories elicited in this system. The second system, sometimes called narrative–biographical memory, or explicit memory, or reflective memory, or declarative memory, emerges during the preschool years and includes information significant to the self. Memories are addressable through intentional retrieval efforts, apart from the original learning conditions. They are identified as representing personally experienced events and compose the individual's life history, approximately equivalent to memory with consciousness or memory with awareness. Clinical studies suggest that, in at least some amnesias, implicit and explicit memory functions may be dissociated.

Disturbances in memory occur through the interruption of registration, retention, or recall.

Disturbances in Registration

Registration and short-term memory retention are usually impaired in disorders that affect vigilance and attention, such as head trauma, delirium, intoxications, psychosis, spontaneous or induced seizures, anxiety, depression, and fatigue. A variety of other metabolic and structural brain disturbances can affect short-term memory as well, particularly lesions affecting the mammillary bodies, hippocampus, fornix, and closely associated areas. Patients with impaired attention and concentration who are able to demonstrate immediate recall may not be able to retain or recollect these items from short-term memory. Benzodiazepine use has been associated with working memory difficulties, especially in the elderly. Some short-acting, high-potency benzodiazepines used as sleeping pills may be particularly troublesome in this regard.

Disturbances in Retention

The retention of memories is impaired in posttraumatic amnesia as well as in a number of cognitive disorders such as dementia of the Alzheimer type and Wernicke–Korsakoff syndrome. The latter, which ordinarily results from chronic thiamine deficiency seen with alcoholism, is associated with pathological alterations in the mammillary bodies and thalamus.

Disturbances in Recall

Disturbances in recall can occur even when memories have been registered and are in storage. At times, inability to recall may signify that the memory traces themselves have disappeared and are no longer retrievable. However, difficulties in recall can occur separately, as in the everyday event of forgetting the name of a person or object, only to spontaneously remember it hours or days later. In normal forgetting, more remote events are less well remembered than recent ones, and important events are most vividly retained in memory. Some demented patients may lose memories for all events occurring after a specific date or event, as if the slate has been wiped clean, but retain earlier memories. Some individuals may progressively erase memories so that they recall only earlier and earlier events. Specific types of memory functions may be subject to selectively different control mechanisms in the brain. For example, recent animal research suggests that endocannabinoids may facilitate the extinction of aversive memories, but not their acquisition or consolidation, through their selective inhibitory effects on local inhibitory networks in the amygdala.

A 27-year-old woman with PTSD symptoms related to a history of childhood physical, emotional, and sexual abuse developed a significant and persistent habit of cannabis abuse. She had tried a variety of street drugs but settled on cannabis as her drug of choice. She was certain that marijuana was far better than other drugs at enabling her to “forget” and not dwell on the past while permitting her to get about her usual day-to-day tasks and requirements.

Under usual conditions, forgotten events can be recalled with prompting, associative memories, or other forms of stimulation such as hypnosis. As described earlier, state-dependent memories are recall failures, reversed by reinstituting the context in which the memory was originally formed.

Amnesias are syndromes in which short- and long-term memory are impaired within a state of normal consciousness. Thus, memory disturbances in delirium should, strictly speaking, not be considered amnestic syndromes. Anterograde amnesia is the inability to register or learn new information (and therefore to form new memories) from a specific event onward; it typically follows head trauma, states of cerebral physiological imbalance, or drug effects. Patients who receive electroconvulsive therapy (ECT) frequently have anterograde amnesias

during the course of the treatments; the amnesia gradually fades over numbers of weeks. Retrograde amnesia is an impairment in recalling memories that were established before a traumatic event, extending backward in time for variable periods. As memory is regained, the more remote memories usually return first. A patient originally amnestic for the 3-month period before an accident may ultimately be left with amnesia for events only a day or an hour just before the accident. In organically caused retrograde amnesias, remote memories are usually intact, although amnesia may exist for more recent events. This contrasts with psychogenic (functional) amnesia, in which the periods of forgotten events may be more spotty or selective.

Hypermnesia, unusually detailed and vivid memory, may occur in gifted people, in association with obsessive-compulsive and paranoid personality traits, and in hypnotic trances. Intrusive memories may occur in PTSD, signaling failure of the mechanisms that usually keep unwanted memories and information out of working memory.

A 28-year-old firefighter and his partner were the first to respond to a five alarm fire at a popular local business. The firefighters were valiant in their efforts, however, one employee and his partner perished. Immediately thereafter, the firefighter reported inability to fall asleep, vivid dreams of the event, sweating, feeling his heart beating fast, smelling smoke and hearing his partner's voice calling for help.

Although many forgotten memories can be recalled in hypnotic trance, retrospective falsification and distortion may also occur under hypnosis. (Memories recalled under hypnosis usually are not accepted as evidence in court.) Retrospective falsification of memory, the development of false memories, is called paramnesia, also known as *fausse reconnaissance*. Confabulation is another common form of paramnesia in which the patient fills in memory gaps with inaccurate information. The responses given to questions by patients who confabulate may reflect past experiences or bizarre, fantastic stories. Confabulation correlates poorly with memory deficit and is believed to reflect frontal lobe dysfunction and a failure of self-monitoring. Confabulation is prominent in certain alcohol amnestic syndromes, such as Wernicke–Korsakoff syndrome, as well as other disorders of the mammillary bodies, thalamus, or frontal lobes.

Déjà vu is the sense that one has previously seen or experienced what is transpiring for the first time; it is a false impression that the current stream of consciousness has previously been recorded in memory. Related phenomena are *déjà entendu,* a sense that one has previously heard what is actually being heard for the first time, and *déjà pensé,* a feeling that one has at an earlier time known or understood what is being thought for the first time. Experiences of *jamais vu, jamais entendu,* and *jamais pensé* involve feelings that one has never seen, heard, or thought (respectively)

things that, in fact, one has. These phenomena are all common in everyday life but may increase in states of fatigue or intoxication and in association with complex partial seizures or other psychopathological states.

> A 28-year-old music teacher drove 12 hours with minimal break in order to attend a music festival for the first time. Upon entering the music hall, he reported suddenly feeling sure that he has been to this music hall before, sat in the same seat and listened to the same music.

Dementia is a syndrome in which the essential feature is an acquired impairment of short- and long-term memory with associated impairments of abstract thinking, judgment, personality changes, and other cortical disturbances. The symptoms always involve more than one sphere of function. In later stages, demented patients may become helpless, too confused to use a stove, and incapable of remembering the names of close relatives. They may wander into dangerous situations, oblivious of their surroundings. Dementias are caused by a variety of pathogenic processes, some of which are reversible, such as hypothyroidism and subdural hematoma, whereas others are irreversible, such as dementia of the Alzheimer type and multi-infarct dementia. Although the characteristic cognitive disturbances seen in severe major depressive episodes are usually called pseudodementias, many neuropsychiatrists believe that profound cognitive dysfunction meeting criteria for dementia associated with depression should properly be labeled as a reversible dementia syndrome. The presence of pseudodementia, however, is predictive of the development of Alzheimer disease (but not in all cases), indicating that the depression was a prodrome of the primary cognitive disorder or that the presence of pseudodementia marks those with an underlying progressive cognitive disorder.

LANGUAGE DISORDERS

Communication difficulties may be due to disorders of thinking as previously described, abnormal speech patterns in mood disturbances and schizophrenia, or primary speech fluency disorders, such as stuttering and stammering, disorders of the articulation and speech apparatus, and CNS disturbances involved in hearing and speech generation (aphasias).

Manic patients typically exhibit pressured speech, in which the speed of the word stream is accelerated. If severe, then the speech may be garbled, imprecise, and difficult to understand. Patients with psychomotor retardation depression speak slowly and monotonously and have a long speech latency in response to questions. Schizophrenic patients may exhibit a variety of speech abnormalities, including poverty of speech and poverty of content of speech (in which the amount of speaking is normal, but

understanding the central message is difficult or impossible). Schizophrenic patients may also be difficult to understand because of the dysarthric effect of antipsychotic medication.

Allusory speech is vague, imprecise, and hard to comprehend because too few cues and details are provided for the listener. Such speech may be heard from some patients with schizophrenia or certain personality disorders or even normal individuals who wish to convey a sense of mystery by just being suggestive, whose suspiciousness causes them to be reluctant to spell things out clearly, or who believe that the listener is more aware of their private codes, meanings, and allusions than is the case.

Stuttering and stammering (ordinarily synonymous) refer to disturbances in the rhythm and fluency of speech due to blocking, convulsive repetition, or prolongation of sounds. This disorder affects males two to three times as often as females, and there is a high rate of familial transmission.

Aphasias, impairments of language produced by brain dysfunction, are ordinarily described as being fluent and nonfluent. In fluent aphasias, which generally reflect dysfunction in the left temporal and parietal area, patients have a normal or even elevated verbal output, sometimes with logorrhea, but they ignore the social conventions of conversation. They produce many well-articulated phrases with normal prosody, but there is little informational content. The fluent aphasias are further divided according to the extent of comprehension by the patient and the ability of the patient to repeat what the examiner says. The principal fluent aphasias are Wernicke aphasia, conduction aphasia, anemic aphasia, and transcortical sensory aphasia.

Nonfluent aphasias are characterized by slow and poor verbal output, difficulty with spontaneous speech, omission of grammatical connecting words, and poor prosody. Patients may produce one-word replies or very short phrases. Brain lesions that cause nonfluent aphasias typically tend to occur in the anterior left hemisphere. The principal nonfluent aphasias are Broca aphasia, transcortical motor aphasia, global aphasia, and the mixed transcortical aphasias.

In aprosodias, the nonverbal aspects of speech, the melody, pauses, timing, stress, accent, and intonation are impaired. Damage to the right prefrontal region has been associated with expressive aprosodias, and damage to the right temporal region and insula has been associated with receptive aprosodias.

DISORDERS OF THE WILL

Central to the sense of self is the concept of will or volition. Psychologically, will is linked to the concepts of intentionality and of transforming awareness and knowledge into initiating action, as the bridge between desire and action. For individuals to manifest normal will, they must be

aware and feel desires—and these desires must arise from within themselves. Concepts related to will that may become the focus of clinical attention when disturbed include motivation and decision making (i.e., the capacity to make choices).

Pathologically heightened will, seen primarily in manic states, is characterized by excessively intense desires and an overly facile capacity to make decisions, with complex questions being decided on in an instant. With heightened psychological energy, these individuals can start new courses of action with astonishing rapidity. Closer examination of these actions in more extreme cases, however, reveals that they share much in common with decreased will in that the intense desires and quick decisions often reflect impulsiveness, which can be considered an escape from true willing and decision making, rather than enduring desires or thoughtful decision making.

The term abulia has been used to describe the loss, lack, or impairment of the power to will or to execute what is in mind. Abulic individuals show a diminished sense of motive or desire and impairment in making the transition from motive and desire to execution of action. Deficiencies in will may be seen in a variety of psychiatric disorders and at the end of life, when patients have surrendered their will to live and are simply waiting to die. In schizophrenia, a diminished sense of will can be seen in passivity phenomena, as well as in other negative (or deficit) symptoms that may affect thoughts, feelings, and behaviors. These include lack of drive, impersistence at tasks, and a general inner flatness. Depressed patients also describe volitional disturbances in general apathy and anhedonia. Patients who chronically inhale solvents (e.g., glue, gasoline, or toluene), smoke marijuana very heavily, and chronically use hallucinogens have a characteristic amotivational syndrome. The extent to which this lack of motivation results from or contributes to the chronic substance abuse is a matter of debate.

In OCD, both the obsessional thoughts and the compulsive rituals are experienced as ego dystonic and not consonant with the patient's conscious desires and will. Similarly, although patients with anorexia nervosa initially have the conscious experience of willing and controlling their intake of food, during the course of the disorder, the sense of willfulness is replaced by one of passivity of being subjugated by obsessional thoughts and compulsive behaviors that assume control of the eating behavior.

A 28-year-old woman with anorexia nervosa and OCD described having as mental background noise a loud audible thought that constantly tallied the caloric value not only of everything she ate but also of all the foods others in her vicinity ate or that she saw on store shelves. This other stream of thinking occurred nonstop, and, like a calculator, the thoughts constantly did sums and updated the tally for the day. These thoughts were experienced in parallel to but separate from any other thinking activities in which she happened to be engaged.

Disturbances of volition are among the more common complaints of patients with personality disturbances who request psychotherapy. Individuals with dependent personalities are characterized by difficulties in making decisions by themselves and often engage in courses of action contrary to their own desires. Similarly, individuals with passive–aggressive personalities obscure their own desires by being excessively involved in the demands made on them by others. Their courses of action do not reflect their own decisions so much as the thwarting of others' desires. People with compulsive personalities use inflexible rules, thereby precluding courses of action based on independent evaluation, individual desires, and decisions. In other situations, they are indecisive, sometimes making impulsive decisions at the last minute when forced to decide. Finally, many individuals seek treatment because of self-designated disturbances of willing: They do not know what they want, they cannot make choices among several options, or they procrastinate excessively. Often, these problems may mask other fears—of wanting, commitment, taking initiative, hard work, success, making a mistake, being criticized, angering others, and all the consequences related to such actions.

DISTURBANCES OF MOOD

Defining, describing, understanding, and categorizing moods have long been among the most important and difficult tasks in psychiatry. The language of feelings is filled with terms that seem to have mostly idiosyncratic meanings, as patients, phenomenologists, and psychiatrists all struggle both to describe inner emotions and to correlate them with external behavior. Even basic terms, such as mood, affect, emotion, and feelings, lack universal definition. The most common convention, which is used here, defines mood as a sustained or prevailing subjective feeling tone or range of tones. Affect is the moment-to-moment feeling state, sometimes rapidly shifting in response to a variety of thoughts and situations, that the clinician can observe. Emotions have been defined as moods and affects that are connected to specific ideas or to the physical concomitants of moods and affects. Feelings are the most poorly defined of all, leading Karl Jaspers to ultimately describe them as everything for which there is no other name. In common parlance, and often professionally as well, these words are sometimes used interchangeably. The term mood disorders, adopted for the DSM-III-R and DSM-IV-TR, replaced "affective disorders" to describe the same group of psychiatric syndromes, and in DSM-5 it has been sub divided into bipolar and related disorders and depressive disorders.

Moods, affects, and emotions can be described by a number of important qualities: Intensity (shallow to deep), range (broad to narrow [or flat]), stability (rigid to labile), reactivity to external events (none too

much), periodicity (periodic to aperiodic), congruence with thought content (congruent [or appropriate] to incongruent), speed of resolution (rapid to slow), and viscosity (evanescent to persistent). The individual's lifelong predominant mood is one component of temperament. Thus, for example, one may be described as having a calm, buoyant, irritable, depressive, anxious, or sensitive temperament.

Moods, affects, and emotions serve as internal and external signal systems. They signal the state of the individual to others and often elicit necessary help and support from the environment. A baby's face communicates its state of need, tension, or contentment, thereby recruiting appropriate maternal interventions. As adults, much of the most important interpersonal communications are transmitted nonverbally through cues that signal the observer about our moods. Positive words communicated by a scowling or sullen face lead listeners to perceive an angry message regardless of the spoken words. Moods also have an infectious quality and serve as important ways of influencing others. Thus, when people act cheerfully toward others, they, in turn, are more likely to feel cheerful and to reciprocate that cheerfulness.

Internally, moods, affects, and emotions let individuals know how well or how poorly they are doing, allowing them, for instance, to gauge the distance between actual self-appraisal and desired self-expectations. For example, individuals who desire to master important goals and feel that they have a reasonably good chance of doing so ordinarily experience pleasant emotional states in relation to these goals. If something intervenes to prevent them from reaching these goals so that there is an insurmountable gap between their desires and the likelihood of success, then they may feel hopeless. In addition to serving as signal systems, emotional states of nonspecific tension, arousal, or anger usually imply that some action is necessary to secure their discharge or release.

Emotional states and their expression are regulated by biological, psychological, and cultural influences. For example, emotional or affective lability, characterized by rapidly shifting emotions that seem unattached to the situation, typically occurs premenstrually in some women, with varying periodicity in cyclothymic individuals and in those with cluster B personality disorders, and in relation to need states such as hunger, sleep deprivation, and sexual frustration. Mood shifts have also been related to environment-related physiological influences such as seasonal changes in light. Psychological regulation of emotions may be related to specific coping mechanisms and the ability to self-soothe, which are developmentally determined. Conscious and preconscious psychological mechanisms, including varieties of self-talk, may help calm or inflame emotions. Culture factors significantly regulate emotional expression. Although the facial expressions for basic emotions are similar in all cultures studied, the range and style of emotional expression permitted in relation

to specific contexts varies greatly from culture to culture and from family to family. Some cultures and families are stiff lipped and inhibit the open expression of emotion, whereas others encourage emotional display. Marked differences exist among cultures in the emotional expression of acute grief, fear, pain, and affection.

Depression

The term depression has been used variously to describe an emotional state, a syndrome, and a group of specific disorders. When seen as part of a syndrome or disorder, depression has autonomic, visceral, emotional, perceptual, cognitive, and behavioral manifestations. As a nonpathological, ubiquitous mood state lasting hours to days, but sometimes longer, feelings of depression are synonymous with feeling sad, blue, down in the dumps, unhappy, and miserable. Depressed mood is common and appropriate after a disappointment or loss. For most people, innate psychological resilience, alternative coping options, and supportive social networks help to alleviate these brief depressive states and prevent them from becoming chronic. Some individuals have a chronically depressed mood, tend to view the world as a difficult place filled with obstacles and burdens, see themselves as victimized, and lack hope for the future. The extent to which constitutional, developmental, and ongoing aversive life events contribute to this pervasive world view is unknown. People who, in early life, were deprived and traumatized may be less resilient and more prone to chronic depressive features than others. Repeated failures and the impact of unrelenting, uncontrollable, and unpredictable negative life events may set the stage for learned helplessness in humans just as they do in animals. A subset of chronically depressed individuals may also have temperamental, biologically driven depression, often seen in conjunction with strong genetic loading for severe mood disorders.

Some depressive states are normal and common reactions to major unwelcome and undesirable life events. Normal bereavement best exemplifies this. In bereavement after major losses, such as the death of a parent, spouse, or child, people experience sadness, pining, and yearning but do not ordinarily have the feelings of guilt, unworthiness, and self-reproach that characterize depressive disorders. Feelings of helplessness and hopelessness may be temporarily present in bereavement, but they ordinarily pass with time. In uncomplicated cases, the process of bereavement takes 3 to 6 months in the acute phase and up to 1 year for complete resolution. Bereaved people are more likely to feel physically ill and seek general health care than at other times, and older widowers are more likely to die than age-matched nonbereaved controls. Pathological grief reactions, bereavements that last more than 1 year, may be seen when the surviving spouse was excessively dependent on the deceased and is unable to obtain emotional and practical (e.g., financial) support elsewhere

or when the survivor is unable to grieve fully because of markedly ambivalent feelings toward the deceased. Distinguishing between pathological grief and grief triggering a depressive episode may not always be possible.

A 52-year-old Hispanic woman was referred to see a psychiatrist by her longstanding family physician who had been treating her for multiple medical problems. Eight months ago, her husband of approximately 55 years passed away after a lengthy illness. She subsequently moved to a different town to live with her daughter and her family. The family became increasingly concerned that she had lost 30 lb, was withdrawn, and was continuously chanting on her rosary beads and "begging to die" in order to join her husband in heaven and take care of him.

The inadequate expression of grief due to incomplete bereavement is believed to be pathogenic in many subsequent psychiatric disorders. For example, impulsive acting out among adolescents who previously lost a parent is often assumed to be due to unresolved grief.

A variety of medical disorders may cause depressive syndromes. Most common among these are endocrine abnormalities, such as hypothyroidism and hyperparathyroidism, and CNS disorders such as cerebrovascular diseases and Parkinson disease. Depressions are more common in strokes affecting left anterior lesions than other locations. Some medications, especially antihypertensive agents affecting adrenergic tone, such as reserpine (Serpasil) and possibly β-blockers, may also trigger depressions. The importance of a genetic diathesis in these iatrogenic depressions is not yet known. Depressive syndromes and disorders in general, however, are unquestionably familial and are likely to have genetic contributions, especially in depressions associated with bipolar disorder.

Cognitive features of depression are prominent. Characterizing the exact nature of the memory impairment using standardized tests has been difficult. The cognitive tasks requiring sustained effort and elaborate cognitive processing may be more disrupted in depression than are those tasks that can be accomplished more automatically. The so-called cognitive triad of depression consists of pervasive cognitive schema related to feelings of worthlessness, helplessness, and hopelessness—expectations that no one and nothing can, or is likely to, help now or in the future, "I'm not OK, the world is not OK, and it's never going to get any better."

Suicidal phenomena are of particular concern. Suicide is common in severe depressive disorders. Depressed patients compose the largest diagnostic group of all completed suicides. However, suicide occurs at high rates in many other conditions as well, notably substance abuse disorders, schizophrenia, and severe personality disorders. Suicide may occur in these conditions with or without a diagnosable comorbid depressive disorder. Depressed patients with comorbid alcohol abuse may be at particularly

high risk for suicide.

Although consistent, useful, validated predictors of suicide do not exist, certain demographic features are associated with higher risk. These include being white, male, and older and living alone. In the psychiatric history, the single most important factor is that of past suicide attempts. A history of violent behavior may also predict suicide. Murderers have a very high suicide rate, especially those who murder family members during episodes of domestic violence. Among clinical signs, hopelessness, anhedonia, and severe anxiety may predict increased suicide risk. Serious physical illness in association with other risk factors, such as depression, may place a patient at higher risk. A genetic predisposition toward suicidal behavior cuts across diagnostic lines and plays a role in suicide risk. This may reflect a tendency toward impulsive behavior, correlating with low CNS levels of 5-hydroxyindoleacetic acid (5-HIAA), the major metabolite of serotonin.

Suicidal gestures and various acts of self-mutilation are also common among impulsive, dependent, and self-hating depressed people, for whom they serve as tension-releasing behaviors and as cries for help that may enlist desired social support. Because suicidal gestures have been associated with an increased risk for subsequently completed suicide, they should not be taken lightly. Such self-harm behaviors have been shown to persist in patients with borderline personality disorder well into the sixth decade of life. Contrary to earlier thinking, they do not appear to "burn out" with age. Subintentional suicide may result when suicidal gestures go awry or when reckless behavior, such as taking unnecessary risks in combat or driving while drunk, prove fatal.

Elated Moods

Elated moods include euphoria, elation, exaltation, and ecstasy. They are marked by feelings of well-being and expansiveness, optimism, capability, pleasure, and grace. Such moods are normally experienced when life is going very well, when long-sought-after goals are achieved, and in states of love, religious fervor, and spiritual transcendence. Peak experiences and experiences of mystic fusion are often accompanied by feelings of exaltation and ecstasy. Sexual pleasure and some chemically mediated states of altered consciousness may also induce these feelings.

Abnormal elated moods are primarily seen as part of manic states and from the effects of certain medications and street drugs. When subtle, as in hypomania, the mood can be ebullient and brimming with self-confidence but with occasional irritability. Other characteristic symptoms of hypomania are increased energy, decreased need for sleep, rapidly flowing thoughts, excessive talking, inflated self-esteem with a demanding nature toward others, and diminished judgment. Mania is a more extreme state in which judgment and sleep are impaired to the point of marked functional disruption. As the mania exacerbates, irritability and anger increase,

alternating rapidly with a brittle expansiveness. Cognitions become increasingly disorganized. Psychotic symptoms, usually involving themes of grandiosity or specialness, occur in 50 percent or more of manic patients. With increasing escalation of the manic state, thinking becomes very fragmented, psychotic symptoms are more prominent, and the syndrome may appear indistinguishable from acute schizophrenia. These three manic states—hypomania, mania, and the psychotic, fragmented manic state—have been referred to as stage I, II, and III mania, respectively. Manic states occur in bipolar disorder, bipolar disorder NOS, and cyclothymia and as a secondary mania caused by a variety of physical and toxic conditions. Such secondary manias may follow specific cerebral insults, accompany systemic disorders, or occur after ingestion of some drugs, including amphetamines, antidepressants, bromocriptine (Parlodel), decongestants, and corticosteroids, among others. Mania is the second most common neuropsychiatric disturbance induced by steroids, occurring in 30 to 35 percent of patients who develop steroid-induced behavioral disorders. Up to 12 percent of patients treated with levodopa (Larodopa) and bromocriptine for parkinsonism develop mania. Right hemispheric brain lesions are specifically associated with secondary mania.

Although not a DSM-IV-TR diagnosis, hyperthymic personality refers to personality characteristics similar to hypomania, DSM-5 has no mention of this in its personality section. These include unusual energy, ebullience, confidence, intensity, and so forth but without either the episodic course or the functional impairment of hypomania. Whether hyperthymic personality should be considered a bipolar spectrum disorder or refers simply to high-energy individuals is still unclear.

ANXIETY DISORDERS

Like depression, the term anxiety refers to a number of different entities: A normal transient feeling, often with adaptive functions; a symptom seen in a wide variety of disorders; and a group of disorders in which the symptom of anxiety forms a dominant element. As a transient, disagreeable emotional state, anxiety may be adaptive, signaling anticipated or impending threat and motivating necessary action. In contrast to fear, the emotional state that exists when a source of threat is precise and well known, anxiety occurs when the threat is not well defined.

Patients often find it difficult to describe feelings of anxiety precisely; at its core, however, anxiety is characterized by intense negative affect, associated with an undefined threat to one's physical or psychological self. Patients use terms such as tense, panicky, terrified, jittery, nervous, wound up, apprehensive, and worried to describe their sensations. Anxiety is additionally characterized by somatic, cognitive, behavioral, and perceptual symptoms. The somatic symptoms of anxiety are legion and often dominate; a partial list includes twitching, tremors, hot and cold flashes,

sweating, palpitations, chest tightness, difficulty swallowing, nausea, diarrhea, dry mouth, and decreased libido. Cognitively, anxiety is characterized by hypervigilance, poor concentration, subjective confusion, fears of losing control or of going crazy, and catastrophic thinking. Behavioral symptoms include fearful expressions, withdrawal, irritability, immobility, and hyperventilation. Perceptual disturbances, including depersonalization, derealization, and hyperesthesia (especially hyperacusis), are also common.

Trait anxiety refers to a lifelong pattern of anxiety as a feature of temperament. Individuals with trait anxiety are skittish, are hypersensitive to stimuli, are psychophysiologically more reactive than others, and exhibit catastrophic thinking. In contrast, state anxiety refers to episodes of anxiety that are tightly bound to specific situations and do not persist after the provoking situation has abated. Free-floating anxiety is characterized by a persistently anxious mood in which the cause is unknown and in which large numbers of diverse thoughts and events all seem to trigger and compound the anxiety. In contrast, situational anxiety occurs only in relation to specific occasions or external stimuli, as in phobias.

Anxiety symptoms can result from numerous physical conditions as well as from other psychiatric disorders. Many endocrine, autoimmune, metabolic, and toxic disorders, as well as medication side effects, are known to generate anxiety. The psychiatrist must differentiate the response of the patient to an underlying condition (e.g., secondary anxiety) from symptoms generated by the primary disorder itself. In psychiatric populations, anxiety symptoms are prevalent among patients with psychoses, organic mental disorders, depression, and substance abuse disorders as well as in the specific anxiety disorders. In patients with schizophrenia, anxiety must be differentiated from akathisia, a common and often overlooked syndrome of subjective restlessness, anxiety, and agitation resulting from antipsychotic medication. The coexistence of anxiety symptoms and depression in major depressive disorder is substantial; anxiety symptoms, such as anxious mood and irritability, are seen in the majority of depressed patients. Additionally, one-half to two-thirds of patients with panic disorder experience a major depressive episode during their lifetime. Medication and drug effects—from acute use, side effects, or as part of withdrawal phenomena—are also common causes of anxiety. Many patients with severe anxiety become dependent on anxiolytic drugs, including benzodiazepines, other sedatives, and alcohol, for symptom relief. During attempts to discontinue these substances, or sometimes during their ongoing use, confusing admixtures of anxiety symptoms, medication effects, and withdrawal symptoms may occur.

Although all of the anxiety symptoms caused by drug use are also seen in primary anxiety disorders, perceptual disturbances, such as depersonalization and hyperesthesia, may be more common in sedative-

hypnotic withdrawal syndromes than in primary anxiety disorders.

From a psychological point of view, anxiety may signal conflict between opposing desires, wishes, or beliefs on the one hand and major disequilibria generated by negative life events on the other hand. Role strains, conflicts among the major social roles that form a person's identity —spouse, parent, child, wage earner, professional, or community member —are common sources of anxiety. The more important the conflict and the less obvious the resolution, the greater the associated anxiety. For example, anxiety symptoms may first emerge when an individual is confronted with an unavoidable, unhappy choice, such as between sustaining a marriage or accepting a career advancement requiring a major move that is unacceptable to the spouse. At times, these conflicts may escape conscious awareness: Someone may feel anxious but not know why. Anxiety syndromes frequently result from a combination of several factors. A person in a work conflict facing an important deadline may try to alleviate initial anxiety symptoms by overwork, then ingest caffeine or amphetamines to keep alert, then become exhausted and fatigued, and ultimately use alcohol excessively to calm down, with each of these elements contributing separately to an anxiety state.

Certain developmental life situations are associated with anxiety. Stranger anxiety develops when infants 6 to 8 months of age begin to recognize the difference between mother and others. When children first go to school, mild anxiety symptoms are common; if the anxiety is excessive, then separation anxiety or school phobia may result. During adult life, anxiety often centers around issues of mastery and accomplishment, in both personal and work life. Performance anxiety, or stage fright, is a specific type of pathological anxiety in which anxiety escalates to panic when public performance is required. In later life, the deterioration of one's body may engender anxiety related to feelings of helplessness and death anxiety.

An 8-month-old baby, described by her mother as friendly and playful suddenly started clinging tightly to her mother when introduced to a new babysitter. The mother was shocked when she tried to leave and place the baby in the babysitter's arms as the baby started loudly wailing and crying, very different than her usual outgoing and friendly infant.

Panic attack is a circumscribed episode of severe state anxiety lasting minutes to hours, with symptoms escalating in a crescendo pattern. The subjective experience is one of utter terror, fears that one will die, go crazy, or lose control, accompanied by many of the somatic symptoms of anxiety mentioned above, including severe chest pains, marked shortness of breath, and exhausting fatigue. Individual isolated panic attacks are common, with up to 30 percent of the general population experiencing at least one attack each year. Panic attacks occur more regularly, and typically

more severely, as part of panic disorder or in association with other anxiety disorders. In classic panic disorder, the attacks are spontaneous; that is, they are not triggered by a specific, predictable environment. In contrast, panic attacks associated with other anxiety disorders are triggered by specific situations such as in social anxiety disorder by social situations or in specific phobia by confronting the feared object or situation. Patients with other psychiatric disorders may experience limited-symptom panic attacks, with episodes characterized by less intense anxiety and by fewer and milder physical symptoms such as isolated paresthesias or difficulty breathing. These limited-symptom attacks may represent aborted full-blown panic episodes that are not further exacerbated by secondary psychological reactions to the initial symptoms.

A 33-year-old man with crippling shyness and social phobia since adolescence started to experience increasing episodes of shortness of breath and chest pain in his late 30s. On numerous occasions, he called 911 for ambulances to take him to the local emergency room, fearing that he was having a heart attack. He underwent multiple cardiac evaluations and had a variety of diagnostic procedures, all of which were negative. These episodes increased in frequency at the point that he was fearful to leave his house or go to work. At this point, at the insistence of his fiancé and parents he sought a psychiatric evaluation. He gradually came to understand these episodes to be panic attacks and decreased the frequency with which he called for emergency services.

Although American Psychiatry has segregated panic attacks from other forms of anxiety, assuming categorical, phenomenological, and biological differences, these distinctions are far from universally accepted. Much of European psychiatry views panic as simply an extreme form of anxiety to be understood as part of a continuum of intensity. DSM-5 made it more simpler by deleting the association of agoraphobia from panic, and making agoraphobia a separate diagnosis.

Phobias are irrational fears. In an effort to reduce the intense anxiety attached to phobic objects and situations, patients do their best to avoid the feared stimuli. Thus, phobias consist of both the fears and the avoidance components. The fear itself may include all of the symptoms of extreme anxiety, up to and including panic. In specific phobias, persistent, irrational fears are provoked by specific stimuli. Table 8–3 lists some illustrative phobias. Common specific phobias include fear of dirt, excreta, snakes, spiders, heights, and blood (also termed blood-injury phobias).

Behavioral, psychodynamic, and biological theories have all been advanced as causes of phobias. Some well-known phobias, such as fear of animals, may result either from early traumatic events (developing along the paradigm of classic Pavlovian conditioning) or from displacements of early psychodynamic conflicts. Genetic influences may also play a role in the development of phobias. For example, some individuals with blood-injury phobias, which strongly cluster among biological relatives, may be

genetically predisposed by vagal responses to certain stimuli. Animal models also indicate possible biological vulnerability. Some monkeys that have never previously been exposed to snakes become panicky when placed in the presence of a snake. Because such fear responses obviously have adaptive value, it has been suggested that some human phobic responses also represent exaggerations of adaptive behaviors shaped by evolutionary biology. Complex phobias, more elaborate than specific phobias, involve fears related to a broader range of situations. Agoraphobia, the best known, means fear of the marketplace, symbolizing a fear of open spaces. Current thinking suggests that agoraphobia is a secondary reaction to panic attacks. In this view, individuals who have become terrified of having panic attacks in public retreat to the safety of their own homes, hoping to reduce the likelihood of panic attacks by avoiding places where they were once triggered and where they may feel exposed and embarrassed.

Table 8–3.
Specific Phobias

Phobia	Definition
Acrophobia	Fear of heights
Agoraphobia	Fear of open spaces
Alektorophobia	Fear of chickens
Amathophobia	Fear of dust
Amaxophobia	Fear of riding in a car
Apiphobia	Fear of bees
Arachibutyrophobia	Fear of getting peanut butter stuck on the roof of the mouth
Astrapophobia	Fear of lightning
Aviophobia	Fear of flying
Blennophobia	Fear of slime
Claustrophobia	Fear of enclosed spaces
Cynophobia	Fear of dogs
Decidophobia	Fear of making decisions
Didaskaleinophobia	Fear of going to school
Electrophobia	Fear of electricity
Ephebiphobia	Fear of teenagers
Eremophobia	Fear of being alone
Gamophobia	Fear of marriage
Gatophobia	Fear of cats
Gephyrophobia	Fear of crossing bridges
Gynophobia	Fear of women
Hydrophobia	Fear of water
Kakorrhaphiophobia	Fear of failure
Katagelophobia	Fear or ridicule
Keraunophobia	Fear of thunder
Musophobia	Fear of mice
Nyctophobia	Fear of night
Ochlophobia	Fear of crowds

Odynophobia	Fear of pain
Ophidiophobia	Fear of snakes
Pnigerophobia	Fear of smothering
Pyrophobia	Fear of fire
Scholionophobia	Fear of school
Sciophobia	Fear of shadows
Spheksophobia	Fear of wasps
Technophobia	Fear of technology
Thalassophobia	Fear of the ocean
Triskaidekaphobia	Fear of the number 13
Tropophobia	Fear of moving or making changes

In social anxiety disorder (social phobia), patients become overwhelmingly anxious and fear situations in which they may be observed. In the limited type, only a few specific situations evoke the fear, such as speaking in public or using a public lavatory. In the general type, broad-based fears of social situations globally hamper the person's interpersonal life.

SOMATIC DISTURBANCES IN PSYCHIATRIC ILLNESS

Nearly all psychiatric disorders manifest some change in basic physiological function. These alterations in somatic function may be quite nonspecific, but the severity of the disruption in normal function may provide clues to the amount of physiological malfunction seen in the primary underlying psychiatric illness. Although frequently nonspecific in nature, the severity of these somatic signs and symptoms provides markers as to the amount of biological disruption seen in the disorders that cause them. Furthermore, somatic symptoms can also cause exacerbations of some disorders. If untreated, then these processes can create destructive feedback loops in which the disorder causes symptoms, which then exacerbates the disorder, which causes increases in the symptoms, and so on. As examples, the insomnia of manic or hypomanic states, if untreated, causes a marked worsening of the mania; similarly, the weight loss of anorexia nervosa causes starvation effects, such as a preoccupation with food, thus exacerbating one of the hallmark features of the underlying disorder.

Sleep Disturbances

Considering the fact that about one-third of each day's activities are spent sleeping, it is not surprising that psychiatric disturbances are frequently manifested by an alteration in normal sleep. These alterations may be in the timing, the amount, and the quality of sleep as well as the presence of abnormal events occurring while sleeping. In addition, the alterations may be associated with the primary psychiatric disturbance, medications used

in psychiatric treatment, or the use of other prescribed agents or substance use.

Insomnia is usually defined by its subjective component as the sensation of not sleeping well or enough. Most, but not all, patients complaining of insomnia demonstrate some sleep abnormality if examined in a sleep laboratory. Insomnia is a common, often chronic, symptom or sign of many different psychiatric disorders, including substance abuse, depression, generalized anxiety disorder, panic, mania (in which the diminished sleep does not always provoke a complaint), and acute schizophrenia. It may also occur as a consequence of aging or as a symptom or disorder not associated with other psychopathology. Insomnia may also result from the ingestion of substances that alter the normal sleep–wake cycle, including alcohol or stimulants, and by the discontinuation of sedative-hypnotics. Although much attention is often paid to distinguishing patterns of insomnia, such as difficulty falling asleep versus middle or terminal insomnia (early morning awakening), or linking specific patterns to a specific disorder (e.g., melancholic depression with terminal insomnia), the clinical usefulness of these distinctions is sometimes unclear.

A 32-year-old computer engineer described herself as a "worry wart." She reported insomnia "my whole life" including initial insomnia and multiple awakenings throughout the night. She reported her mind could not relax as she tried to sleep and found herself worrying about "everything that happened during the day." Inevitably, she would wake up in the middle of the night, to "make a list of things I forgot, or ruminate on possible mistakes at work."

Hypersomnia, characterized by either excessive nighttime sleep or excessive sleepiness during the day, is less common than insomnia. It, too, however, may reflect a number of different pathological states. Some depressed patients, especially those with a history of mania or hypomania, may exhibit hypersomnia. Hypersomnia may also be seen during stimulant withdrawal, with excessive use of sedatives or tranquilizers, or in conjunction with a variety of medical disorders. In narcolepsy, the patient has sudden attacks of irresistible sleepiness, a symptom that may be part of a broader syndrome that includes cataplexy (sudden attacks of generalized muscle weakness leading to physical collapse in the presence of alert consciousness), sleep paralysis (waking from sleep with a sensation of being totally paralyzed that may persist for minutes), and hypnagogic hallucinations (vivid visual hallucinations that occur at the point of falling asleep). Narcoleptic attacks are often precipitated by unusual states of arousal (e.g., cataplexy may immediately follow unrestrained laughter or orgasm). Daytime sleepiness may reflect sleep apnea. In this disorder, typically middle-aged patients demonstrate severe snoring—often first reported by their bed partners—and periods when breathing stops. The

condition results from soft palate abnormalities that cause intermittent airway obstruction throughout the night; patients awake repeatedly to find themselves gasping for air. Associated daytime fatigue is common in sleep apnea. Periodic hypersomnia also occurs in Kleine–Levin syndrome, a condition typically affecting young men in which periods of sleepiness alternate with confusional states, ravenous hunger, and protracted sexual activity. Intervals of days, weeks, or months may pass between these episodes.

Somnambulism, or sleepwalking, and sleep terror disorder (night terror) are two sleep disorders characterized, respectively, by aimless wandering with incomplete arousal and by acute anxiety and physiological arousal without awakening. Although both disorders typically begin in childhood, sleepwalking may be also initially precipitated by some psychotropic medications.

A 54-year-old man with complex PTSD was treated with a sedative hypnotic by his primary care physician. During this period, his wife noted that he often went to the refrigerator in the middle of the night, sound asleep, and ate indiscriminately. When she woke him during these episodes, he appeared dazed and had no awareness of his purposive behaviors. To the best of his recollection, confirmed by his wife, before treatment with the sedative hypnotic agent, he had no history of somnambulistic phenomena.

Nightmares are a common complaint, often associated with traumatic events, anxiety disorders, and mood disorders, but not uncommon as an occasional event in otherwise healthy individuals. Vivid dreams and nightmares may also be a medication side effect.

A 32-year-old woman was treated for a mild depression with an SSRI. Her sleep pattern before treatment had been characterized by hypersomnia, sleeping between 10 and 12 hours nightly. Soon after the initiation of her SSRI antidepressant, she noticed that her sleep was not only shorter and more fragmented but that she had remarkably intense dreams that she remembered vividly. She made clear that these were not nightmares, just dreams with intense colors and clarity, which she rather enjoyed.

Sensory symptoms during sleep, typically described by patients as peculiar feelings in their legs causing an irresistible need to move around, are characteristic of restless legs syndrome. The motor abnormality of repetitive myoclonic jerking of the legs, awakening both patients and their partners, is known as nocturnal myoclonus.

Appetite and Weight Disturbances

A significant change in appetite and weight can be seen in both medical and psychiatric illness. Anorexia may occur in the latter stages of chronic medical illness and is frequently seen in depression, grief, and some anxiety

disorders. Additionally, anorexia may be secondary to alterations in taste sensation, as a function of psychiatric disturbance or medication side effect. In eating disorders, such as anorexia nervosa, patients may resist hunger to restrict food intake to achieve a physiologically unrealistic low weight. Hyperphagia (increased appetite) occurs in some depressed patients, both with and without a history of mania or hypomania. Binge eating, of up to several thousand calories per episode, may occur as an attempt to self-soothe and emotionally self-regulate during times of increased tension and anxiety and as a key feature of bulimia nervosa or of binge-eating disorder. Increased appetite may be seen as a side effect of many psychotropics or, albeit rarely, in some hypothalamic disorders or in bilateral temporal lobe dysfunction such as Klüver–Bucy syndrome, in which it occurs in association with emotional placidity, hypersexuality, hyperorality, and other related symptoms.

A 45-year-old man with schizoaffective disorder had been intermittently erratic with his follow-up visits and had requested several telephone refills over the last several months. He was instructed by his psychiatrist to come in to see him before additional refills could be called in to his pharmacy. Subsequently, he presented for an office visit with his psychiatrist. In the interview it became apparent that the patient had gained a significant amount of weight, approximately 40 lb, in the last several months. He admitted to having a "hunger" for sweets and fatty foods and acknowledged that he had a marked increase in his appetite after he was switched to an atypical antipsychotic.

Energy Disturbances

Normal energy levels vary considerably among people. Some people fatigue easily and are perceived by themselves and others as having "weak constitutions," whereas others appear to have almost boundless energy and much less need for sleep.

Fatigue is a common nonspecific symptom that occurs in both medical and psychiatric disorders. It is also frequently seen as an unexplained complaint in primary care practices; in one study, 24 percent of patients complaining of fatigue received no medical or psychiatric diagnosis. Historically, fatigue not due to another disorder, typically in association with "nervousness," has been described by terms such as asthenia, neurocirculatory asthenia, neurasthenia, and psychasthenia. Consistent with this tradition, many fatigued patients, having been labeled depressed or neurotic by their physicians, are referred to psychiatrists after routine workup has ruled out anemia, hypothyroidism, sleep apnea, and other frequent somatic causes.

Recently, those patients with primary complaints of tiredness have been most commonly diagnosed as having chronic fatigue syndrome (CFS) (previously and incorrectly labeled Epstein–Barr viral syndrome), a disorder characterized by fatigue lasting months to years, typically

beginning soon after a viral syndrome. In addition to the fatigue, CFS is characterized by myalgias and cognitive changes such as forgetfulness and poor concentration. Although controversy still exists over the extent to which cases of CFS represent discrete postviral diagnostic syndromes, mislabeled cases of depression, or modern versions of psychasthenia, evidence continues to mount to suggest that these syndromes are discrete postinfectious entities and are not simply variants of or disguised mood or anxiety disorders.

A highly accomplished, energetic 40-year-old woman was referred for psychiatric consultation after her physicians were unable to offer a definitive physiological diagnosis despite extensive medical workup after the acute onset of profound fatigue occurring in the wake of a mild viral illness. This fatigue caused her to be totally incapacitated and bedridden for many months and left her feeling quite helpless and distraught. It was exacerbated by even small amounts of alcohol. Two independent psychiatric evaluations concluded that the patient had no psychiatric disorder. The patient developed this syndrome after jogging near Lake Tahoe. Some epidemiological studies found a series of other cases of profound fatigue originating in this particular geographical area after vigorous exercise. Various antidepressants seemed to provide partial relief, but tolerance to their positive effects developed. The fatigue slowly improved over many years; 7 years after the original symptoms, the patient continued to experience waxing and waning fatigue but at a lesser degree than originally.

Disturbances in Sexual Drive

As with energy, the normal range of sexual drives is common. Some individuals are naturally lusty, whereas others have limited sexual desire. Diminished sexual drive with impotence or decreased libido is seen in a wide variety of neurological, metabolic, and other somatic conditions. Among neurological disorders, complex partial seizures are commonly associated with hyposexuality, occurring in 50 percent of patients. Psychiatric disorders known for diminished sexual drive include depressive disorders, schizophrenia, and substance abuse disorders. Diminished libido, erectile dysfunction, and anorgasmia are also common sequelae of many psychotropic agents, especially those with a strong serotonergic action.

A depressed 36-year-old man had an excellent antidepressant response to an SSRI but developed decreased sexual interest and anorgasmia with relative preservation of erectile function. When the sexual side effects did not diminish after 4 weeks, his psychiatrist tried a number of potential remedies. Although sildenafil (Viagra) improved his erections somewhat, his libido and orgasmic function continued to be markedly impaired. Reluctantly, he and his psychiatrist agreed to switch to a different antidepressant class that was not associated with sexual side effects.

Increased sexual activity may be seen in some neurological, psychiatric, and drug-induced disorders. Manic patients frequently exhibit hypersexual

interests and behaviors to an unusual degree, compared with their euthymic interests and behaviors. Hypersexuality is occasionally seen in conjunction with epileptic syndromes or in patients who have had diencephalic injuries.

Altered sexuality, including fetishism sadomasochism, pedophilia, and other paraphilias, may be seen as isolated psychiatric syndromes. In individuals whose previous sexual behaviors were within the bounds of social propriety for their groups, inappropriate sexual behaviors may signal early brain disease or psychosis. Cross-dressing may occur in transvestites, transgenderists, transsexuals, or, occasionally, in other psychiatric conditions.

A 54-year-old married engineer became preoccupied with viewing pornographic sites on the Internet. He had always been excited by the visual stimulation of early adolescent girls. He had never acted on these feelings, had never had an affair since being married, and, even when younger, had always had consensual sex with girls of his own age. With the advent of computers, he had become increasingly preoccupied by finding and spending significant amounts of time on websites devoted to photos of young girls. Eventually, he was spending 3 hours daily at these websites. At that point, he sought treatment for a pattern of behavior that he believed he could not control at all.

Appearance

Studies show that clinicians often formulate an initial psychiatric diagnosis within 30 seconds of seeing a patient. Although approximately one-half of such initial impressions prove to be incorrect, the remainders, however, are validated by psychiatric history and mental status examination, revealing just how much information is communicated by appearance and body language.

Among the physical disorders whose appearances suggest coexistent psychiatric conditions are acromegaly, Cushing syndrome, Down syndrome, systemic lupus erythematosus, fetal alcohol syndrome, Klinefelter syndrome, and Wilson disease to name a few. The general appearance of the skin may suggest the presence of occult psychiatric problems. The general condition and flush of the skin may reveal hypervascularity and ruddiness, suggestive of alcoholism, abscesses indicative of hypodermic needle abuse, tattoos indicative of certain group affiliations, or weathering and wasting indicative of self-neglect and malnutrition. Healed scars on the wrists and arms suggest a pattern of self-mutilation. Patchy baldness, especially in conjunction with torn or infected cuticles, indicates trichotillomania, a syndrome of compulsive hair pulling. Psychophysiological symptoms reflecting psychiatric disturbance include urticarial reactions and neurodermatitis, the latter resulting, in part, from self-excoriation, destructive scratching secondary to compulsions, and unrelenting sensations of discomfort.

A 23-year-old woman who had recovered from a severe episode of anorexia nervosa as an adolescent had recurrent depressions and persistent personality disturbances characterized by odd and eccentric ideas, shyness, and avoidance. She became enchanted with gothic themes, spent hours on the Internet in chat rooms featuring dark and morbid poetry, and started to have herself tattooed. Within 1 year, she had tattooed every inch of her face, eyelids, lips, earlobes, upper extremities, scalp, and torso with complex, intricate, gothic designs. She acknowledged that tattooing her skin in such an extreme manner served to mark her as deviant and to warn people away from her.

Examination of the head and neck may reveal exophthalmos or puffy eyelids, suggesting thyroid disease, marked pupillary dilation with anxiety or stimulant abuse, miosis with narcotic abuse, abnormal pupillary pigments in Wilson disease, salivary gland enlargement in bulimia nervosa, or necrosis of the nasal septum in cocaine abuse, among other signs. Frequent sighing is a common respiratory sign in depression. Simple sighing must be distinguished from respiratory dyskinesia in psychotic patients who have been treated with neuroleptics. The latter may occur as an acute dyskinesia due to antipsychotic medication, or it may be a late manifestation and component of tardive dyskinesia.

Deviant appearance is quickly perceived by laypeople and professionals and may contribute to the frightened and stigmatizing social withdrawal by strangers, as well as acquaintances, so often experienced by psychiatric patients. Akathisia and dystonic movements and parkinsonian shuffling gait in patients taking neuroleptics (especially the older, conventional antipsychotics), as well as the dilapidated and unkempt appearance of some psychiatric patients, can immediately signal psychiatric patient status to observers. The term *Diogenes syndrome* has been used to describe old people who have filthy personal appearance and demonstrate severe self-neglect about which they have no shame.

DISTURBANCES OF BEHAVIOR

Aggression, Hostility, Impulsiveness, and Violence

The spectrum of aggressive emotions and behaviors is characterized by heightened vigilance in response to a sense of threat and enhanced readiness to attack. Physiological tone may be geared for a fight. Assertiveness, the adaptive aspect of these emotions, includes sensing that something needs to be done and feeling willing and competent to take constructive action. The manner and extent to which aggressive emotions can be expressed varies from society to society and from situation to situation. These emotions are among the most carefully regulated because of their potential destructiveness. Acts of aggression are on a continuum, beginning with irritability, progressing to verbal threats and intimidation, and extending to physical bullying and assault to homicide, sometimes including acts of calculated violence and sadism.

Irritability is an unpleasant feeling state characterized by inner unease. In contrast to anger, irritability does not lessen after an outburst. Often, others are more aware of an individual's irritability than is the person him- or herself. It is diagnostically nonspecific, seen in a variety of physiological states—psychotic, anxiety, and mood disorders—and as a lifelong temperamental quality. Hunger, sleepiness, sexual frustration, and pain are among the physiological triggers commonly associated with irritability.

Individual differences in the tendency toward experiencing and expressing anger and violence are biological, developmental, and cultural in origin. Some infants are irritable from birth. Subtle early birth injuries and brain anoxia may increase the susceptibility of some people to be violent. Furthermore, studies of EEG patterns in violent people show increased abnormalities, especially in those with repeated violence and violence with little or no obvious motive. Soft neurological signs are also seen in violent criminals. Biochemically, low cerebrospinal fluid (CSF) 5-HIAA has been associated with a variety of impulsive behaviors, such as violent crimes, recurrent fire setting, and violent suicide attempts.

Consistent with the hypothesis that an inverse relationship exists between central serotonergic system function and impulsive, aggressive behavior, a few double-blind, placebo-controlled studies in patients with cluster B personality disorders have demonstrated that enhancing central serotonergic function by using SSRIs reduces irritability and impulsive aggressive behavior.

A 32-year-old man came for psychiatric treatment after a breakup with his fiancé of 5 years related to his anger outbursts when he would curse, yell, slap, and throw things at her. Once, during a rage attack, he slashed the tires of his fiancée's car. Another time, he impulsively took an overdose of pills that he found at home. He responded to treatment with an SSRI, which improved his mood lability and anger outbursts.

The pathological childhood triad of bed-wetting past the age of 6 years, setting fires, and torturing animals has been associated with subsequent violent behavior in adults. Interpersonally, studies show that violence-prone individuals require more personal space around their physical person than do others. Violent individuals feel threatened when approached too closely, particularly from the rear.

Psychological and social contributions are also strong. Violence in families breeds violence, and battered children often grow up to be battering adults. Cultural norms for the expression of violence differ considerably. In some socioeconomic and ethnic groups, violent gangs organize the energies of many adolescent youth. For some, violent behavior is an adolescent socialization pattern necessary to prove one's manhood or womanhood. Like other social organizations, violent gangs have detailed rules that inhibit and govern the expression of violence. Some

unpredictable and unsocialized violent people, loners, are too violent to be contained even in gangs.

Aggressive and violent behavior is diagnostically nonspecific. Violence in schizophrenia may occur as a consequence of paranoid delusions, in response to command auditory hallucinations, or secondary to passivity experiences. Manic patients and those in mixed states may be violent, often in response to minimal provocation. Violent behavior commonly occurs in patients with antisocial and borderline personalities (in the latter often self-directed as well as other directed). Violent behavior may occur in epilepsy, although rarely during true ictal periods; in frontal lobe syndromes as a "release phenomenon"; and in association with abused substances, particularly disinhibiting sedatives, such as alcohol, or stimulants, such as amphetamines and cocaine, which increase irritability, aggressiveness, and paranoia.

Impulsive violence may be provoked by a number of stimuli and situations. Alcohol is perhaps the most common disinhibitor of violence. Intrafamilial violence, the most common setting for homicide, is frequently related to alcohol intoxication. In episodic dyscontrol and intermittent explosive disorder, violent behaviors typically erupt after a person has ingested alcohol, a phenomenon known as pathological intoxication. In these often ferocious outbursts, the individual may confront or provoke any potential target for violence, including total strangers and police, but girlfriends, wives, and parents are frequent victims. Patients with episodic dyscontrol commonly have histories of violent sexual behavior, including rape, and, often while intoxicated, speeding and reckless driving, sometimes chasing down, stopping, and attacking other motorists who they believe "get in their way."

An attractive 40-year-old woman had been highly sought after by rich and powerful men, several of whom she had married and several of whom divorced her because of her uncontrollable rages. She described herself as becoming "insanely jealous" in a flash whenever she saw her husband glancing at other attractive women, and her behavior flared out of control, particularly when she had been drinking alcohol. On one occasion, she bit off one husband's ear. She attacked another with pots and knives, nearly killing him. She recognized how serious her problems were and sought psychotherapy, medication (SSRIs were helpful), and anger management training to help her control her emotions and behaviors.

Temper Tantrums

Immature individuals with persistent personality problems may not develop mechanisms to inhibit temper tantrums that they displayed as children. Particularly, if childhood tantrums produced the desired result, then learned tantrum behaviors may persist into adult life. Although such individuals may be pleasant and sociable when life is going well, they lack

the capacity to tolerate frustration and are easily provoked by threats to self-esteem and self-image and by not having their own way. When frustrated or threatened, they may act like bullies, glare, snarl, yell, shout, intimidate, pout and sulk, and sometimes be physically violent.

Displaced Rage

When circumstances prevent the expression of rage directly against those people or institutions provoking frustration, other outlets for aggression are often found. Acts of violence that are either calculated or wanton may result. Cruelty to animals and fire setting may persist as adult forms of destructive behavior. Rape, an act of control, intimidation, terror, and humiliation may also displace frustrations that are not expressed more adaptively.

Sadism may occur with or without explicit sexual gratification. Calculated cruelty conducted seemingly without anger or emotional arousal may reflect inadequate development of social morality or individual conscience, as in the conduct of torturers and some cold-blooded murderers. In some societies and under specific circumstances at certain times in history, such activity has been socially sanctioned, suggesting at least that some people lack inborn inhibitions against cruelty or violence.

A 50-year-old man with antisocial personality disorder was seen in a psychiatric emergency room for violent behavior. During the course of the assessment, he described how he had been emotionally and physically abused by a sadistic father who whipped him repeatedly until he bled and who chained him to a tree in the backyard. This man grew up feeling callous. He enjoyed intimidating and physically hurting others, as he often did in prison. With a laugh and smirk, he reflected on how he turned out to be "just like my old man."

Self-Mutilation

For a variety of reasons, in many different cultures and in many different disorders, people commit acts of violence against themselves, ranging from body piercing to cutting and burning to autoamputation. Psychotic patients may perform extremely self-destructive acts short of actual suicide that often have symbolic import, such as enucleating their eyes or castrating themselves. Patients with borderline personality disorders or borderline traits may cut themselves repeatedly with broken glass or razor blades or burn themselves with cigarettes on arms, legs, breasts, or other body parts. Patients typically deny that these acts are meant to be suicidal but describe the need to feel external pain to mirror internal suffering, to release tension, or to counteract dissociativelike numbness. Recently, self-cutting has become more common in teenagers, especially girls, as a tension-relieving mechanism.

A 20-year-old college student was taken to the emergency room by her friends, who were alarmed by the recent transverse cuts on her arms. None of the cuts were deep enough to require sutures, but, even when they had healed, they left obvious small red scars on her skin. In explaining her behavior, the student described that when she was upset, such as after a fight with her boyfriend or a negative telephone call with her parents, she became agitated, and cutting herself consistently relieved the agitation and anxious feelings.

Trichotillomania is a syndrome of compulsive hair pulling, resulting in bald patches. It is often associated with other self-mutilatory behavior, such as picking the face, nails, or cuticles to the point of infection and bleeding. Trichotillomania may sometimes be related to OCD.

Children with Lesch–Nyhan syndrome, a developmental disability syndrome caused by a congenital metabolic abnormality, bite and pick at themselves so compulsively as to do themselves great harm and routinely require restraint. Occasionally, patients with Tourette syndrome demonstrate compulsive self-harming behavior.

Other Disturbances of Feelings

Diminished levels of emotional intensity may be seen in anxiety disorders, mood disorders, and schizophrenia. Mild emotional flattening with blunted ability to feel joy is common in dysthymia. Some patients with narcissistic and borderline personality disorders complain of inner emptiness and pervasive boredom and ennui without demonstrating diminished affect in interviews. Similarly, patients with prominent depersonalization describe numbed emotions. Pathological levels of blunt or flattened affect, indicating markedly diminished affective expression in relation to specific thought content, may be seen in chronic schizophrenia (as part of the deficit syndrome), some organic mental syndromes, and severe depressions and as a side effect to many medications, including SSRI antidepressants, antipsychotic medications, or excessive tranquilizer use. Although the term blunted affect is not classically used to describe the affective flatness of severe depression, it is not always easy to distinguish between schizophrenic and depressive flatness on phenomenological grounds. Anhedonia, the lack of pleasurable feelings from activities that ordinarily provide pleasure, is also seen as part of severe depressions or schizophrenia. Chronically psychotic patients often exhibit emotional deterioration in which affective experience and expression are entirely unrelated to thought content. Inappropriate affect is the incongruency of affective expression and thought content. The patient may display loud and raucous laughter or giggling in relation to bland or sad thoughts or may show grief without apparent reason. Inappropriate affect sometimes indicates that the thoughts have private meanings for the patients; the emotional expression might make better sense if the private meanings were understood. Inappropriate affect must be distinguished from affective

expressions that may actually be appropriate in a given subculture or ethnic group that is unfamiliar to the observer and from defensive affect, such as the nervous laughter used to alleviate tension or ward off crying. Affective (or mood) lability is characterized by rapid emotional shifts, often within seconds to minutes. It is commonly seen during hypomanic states, late luteal phase dysphoric disorder (premenstrual syndrome), postpartum blues, other states of physiological instability, and in cluster B personality disorders.

A 51-year-old man with lifelong dysthymic disorder was treated with an SSRI. Although he had an excellent antidepressant response, 2 months after beginning treatment, he noted an affective flatness. He described not becoming as easily upset as before treatment (which both he and his wife interpreted positively), but also he believed his sense of joy, in being with his children or in nature, was markedly blunted. He enjoyed positive events but with a marked diminution of affective intensity.

Alexithymia is difficulty identifying, describing, and differentiating feelings or distinguishing between feelings and physical sensations. Alexithymic individuals often have constricted imaginations and fantasies, are preoccupied with objects and events in the outside world, and have little private, personal internal life. When distressed, the patients are simply aware of not feeling well and usually complain of somatic symptoms, leading to frustrating interactions with their physicians who cannot find physical causes for the presenting physical complaints. Some view alexithymia as a condition in which affect is communicated through somatic language.

With regard to other feeling states that may contribute to pathological conditions, recent writers have focused on the role of “shame” as a potentially important emotional mediator of specific types of narcissistic injury that may provoke states of depression, rage, PTSD, and suicide, among others.

DISTURBANCES OF INTERPERSONAL RELATIONSHIPS

Normal interpersonal relationships include relationships with parents, children, spouses, siblings, extended family members, friends, colleagues, coworkers, and members of the larger community. These relationships ordinarily help provide for the satisfaction of basic drives, for affiliative needs, and for finding purpose and meaning in life. Through stable and satisfying relationships, human needs are met for intimacy, including love, sex, and affection, to be cared for and nurtured, provide care, learn, play, relax, dominate, and be productive through mutual effort. Interpersonal relationships are carefully regulated by means of interpersonal signs and signals. The extent to which deviance from these patterns is tolerated in a given relationship varies from behavior to behavior, relationship to

relationship, family to family, and culture to culture.

Disturbances in interpersonal relationships may be viewed as characteristics attributable to a single person or as characteristics of an interpersonal system. Individual disturbances are considered to be undesirable or maladaptive personality traits. When these traits are present to a significant extent and interfere with social functioning or cause distress, they may compose a personality disorder. Disturbances of interpersonal relationships have also been described at a systems level (e.g., as dyadic and family patterns of system disturbance).

Personality Traits and Disorders

Personality, variably defined, is the characteristic pattern of an individual's attitudes, behaviors, beliefs, feelings, thoughts, and values—the sum of a person's emotional, cognitive, and interpersonal attributes. Personality traits are the prominent and characteristic features of an individual's personality and do not imply psychopathology. Aspects of personality are present from early life, and personality traits are relatively stable from adolescence onward, consistent across different environments, and recognizable by friends and acquaintances. The term personality disorder should be reserved for those consistent patterns of thought, feeling, and behavior that are maladaptive. Personality disturbances manifest primarily in interpersonal contexts and in this way can be viewed as interpersonal behavior disorders.

The determinants of personality are multiple and varied—innate and early biological, developmental, and environmental factors inside and outside the home. Through learning and the environment, temperamental factors (genetic or constitutional) are shaped into character.

The dimensional approach to personality and personality pathology characterizes individuals along a continuum of traits. Five dimensions of temperament have been described that appear to be somewhat independent and to have strong genetic contributions: Neuroticism (highly emotional, reactive, and thin skinned, contrasting with emotional stability), extroversion (contrasting with introversion), openness (contrasting with discomfort with novel experiences), agreeableness (contrasting with contrariness), and conscientiousness (contrasting with fickleness). These temperamental attributes may have implications for the course of psychotherapies that cut across diagnostic categories.

Another dimension of personality not adequately dealt with in the DSM-IV-TR and now DSM-5 concerns moral behaviors such as honesty and integrity. The extent to which individuals behave honestly and with integrity differs considerably across individuals and in different situations. Deception and lying are common behaviors that occur in benign forms (e.g., in "white lies") and in pathological forms, psychiatrically important in antisocial and sociopathic disorders, pathological liars, and malingerers.

Deception and lying may be difficult to assess clinically in the absence of additional informants. Studies of nonhuman primates indicate that, at least among chimpanzees, deception (equivalent to lying and dishonesty) is relatively common and, in some situations, adaptive.

Another personality typology characterizes personality along three dimensions related to temperamental characteristics presumed to be strongly influenced genetically: Harm avoidance, novelty seeking, and reward dependence. High scores on the three dimensions characterize inhibition and pessimism, impulsive and exploratory behavior, and dependency and sentimentality, respectively. Different personality types can be described according to patterns of scores on the three dimensions. For example, antisocial personalities are characterized by high novelty seeking, low harm avoidance, and low reward dependence, whereas dependent characters have low novelty seeking, high harm avoidance, and high reward dependence.

The DSM-IV-TR, by contrast, uses a categorical approach, yet both the large overlap among the DSM-IV-TR personality disorders and the clustering of these personality disorders into three broad groups imply a lack of clear boundaries to the currently defined categories. The three DSM-IV-TR clusters describe odd or eccentric types (cluster A); dramatic, emotional, and erratic types (cluster B); and anxious and fearful types (cluster C). While the DSM-5 task force recommended an alternative model and approach to personality disorders, the APA board of trustees decided to keep current nomenclature to preserve continuity. The alternate approach is included as Section III in DSM-5. In this model a greater emphasis on personality functioning and trait-based criteria increases the stability and empirical basis of the disorders.

The odd or eccentric group includes paranoid, schizoid, and schizotypal personality disorders. Patients with these personality disorders have the core traits of being interpersonally distant and emotionally constricted. Paranoid personalities are quick to feel slighted and jealous, carry grudges, and expect to be exploited and harmed by others. Schizoid personalities lack friendships or close relationships with others and are indifferent to praise or criticism by others. Schizotypal personalities display odd beliefs, engage in odd and eccentric gestures and practices, and exhibit odd speech.

The dramatic, emotional, and erratic group includes borderline, histrionic, narcissistic, and antisocial personality disorders. Patients with these personality disorders characteristically have chaotic lives, emotions, and relationships. Borderline personalities are impulsive, unpredictable, angry, temperamental, unstable in relationships, compulsively interpersonal, and self-damaging with regard to sex, money, and substance use. Histrionic personalities are attention seeking, exhibitionistic, seductive, self-indulgent, exhibit exaggerated expressions of emotions, and overconcerned with physical appearance. Narcissistic personalities tend to

be hypersensitive to criticism, exploitative of others, egocentric with an inflated sense of self-importance, feel entitled to special treatment, and demand constant attention. Antisocial personalities are described almost exclusively by behavioral rather than affective or relational terms. They are truant, lie, steal, start fights, break rules, are unable to sustain work or school and shirk everyday responsibilities.

The anxious and fearful group includes avoidant, dependent, and obsessive-compulsive personality disorders. Patients with these disorders are characterized by constricting behaviors that serve to limit risks. As examples, avoidant people avoid relationships, dependent personalities avoid being responsible for decisions, and obsessive-compulsive people use rigid rules that preclude new behaviors. Avoidant personalities are hypersensitive to rejection and are reluctant to enter close relationships in spite of strong desires for affection. Dependent personalities show excessive reliance on others to make major life decisions, stay trapped in abusive relationships for fear of being alone, have difficulty initiating projects on their own, and constantly seek reassurance and praise. Obsessive-compulsive personalities exhibit restricted expressions of warmth, tenderness, and generosity and also exhibit stubbornness with a need to be right and to control decisions; indecisive at times, they often use overly rigid application of rules and morals to the point of being inflexible.

A characteristic personality disturbance seen with frontal lobe damage is referred to as organic personality disorder in the tenth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) and personality change due to a general medical condition in the DSM-IV-TR. In DSM-5 it is included as Personality change due to another medical condition. Its features include irritability, inappropriate jocularity with euphoria, inappropriate socially disinhibited behavior, and impulsiveness. Other patients with damage to different areas of the frontal lobe, in contrast, exhibit apathy and indifference.

Interpersonal Systems

Couples and families have been studied as systems in their own right, and many qualities of these systems have been identified as being clinically important. A scheme for categorizing relational disorders has been proposed for future editions of the DSM-IV-TR, but, as yet, no single generally accepted typology of family psychopathology or interactional types has been established. However, elements of marital discord and harmony have been operationalized in several standard marital inventories. Characteristics of couples and families that have received the most attention include the rules of communication, such as those governing the directness or indirectness with which disagreement and conflict are addressed; the manner (organized or chaotic) in which communications are conducted; taboo topics and secrets about which no one can openly

communicate; the nature and degree of emotional expression, including affection and anger; the cohesiveness, loyalty, and compatibility of members; the nature of the members' shared identities on the one hand and their autonomous development and separateness on the other; the extent to which members treat one another respectfully or take one another for granted and use one another; the distribution of power and decision making among members; the maintenance of generational boundaries (e.g., age-appropriate performance of life roles); and the members' orientation toward and concurrence and disagreement about important values involving moral, religious, intellectual, cultural, financial, occupational, and childrearing issues, as well as aspirations, health practices, leisure activities, and other belief systems.

A characteristic family environment, called high expressed emotion, has been identified that defines a relapse-prone family environment in which one individual has schizophrenia, bipolar disorder, anorexia nervosa, or major depressive disorder. This interactional pattern includes demeaning, intense personal criticism ("You are rotten and lazy") and emotional overinvolvement with the identified patient. Aspects of overinvolvement can be measured by quantifying the numbers of hours of face-to-face contact and by the extent to which relatives categorically assert how the patients feel without ever bothering to ask the patients. Despite these descriptive generalizations, the specific elements of high expressed emotion families that make patients vulnerable to relapse are still obscure.

Couple and family system difficulties are most likely to erupt during predictable stressful events in the normal family life cycle, such as during the newlywed period; pregnancy and childbearing; difficult or contentious childrearing; difficulties with parents, in-laws, and other extended family; insurmountable and unanticipated financial or career problems; serious illness or death of a child or relative; the children's adolescence; departure of children from the home; infidelity; and separation.

Interpersonal attachment styles, based on cognitive schemas that have been linked to earlier repeated experiences with caregivers, influence how individuals perceive and act within interpersonal relationships. Specific types that have been described include insecure attachments (several subtypes include preoccupied, fearful, and anxious-ambivalent attachment), avoidant attachment (including angry-dismissive and withdrawn attachment), enmeshed attachment, unresolved attachment (in which significant losses have not been dealt with), and secure attachment styles. Specific types and subtypes have now been linked to certain types of health behaviors and clinical outcomes. For example, diabetic patients with dismissive attachment styles, a style found in approximately 25 percent of the general population and characterized by low trust of others and excessive self-reliance, show a decreased ability to collaborate with providers and have poorer glycemic control than patients with secure

attachment styles. In a study at a primary care clinic, patients with insecure attachment styles reported more unaccounted for medical symptoms than others. Among these, patients with preoccupied attachment had the highest primary care costs and usefulness, whereas patients with fearful attachment had the lowest, reflecting their respective tendencies to overuse or underuse medical services.

Interpersonal Disturbances in Illness Behavior

Abnormal illness behavior (dysnosognosia) is a persistently pathological mode of experiencing, evaluating, and responding to one's own health status despite lucid and accurate appraisal and management options provided by a health professional. These behaviors can be considered as interpersonal disorders between patients and health care professionals. Central to all of these behaviors is the adoption of the sick role by the patient, who then engages in characteristic interactions with health care providers, which typically leave both the provider and the patient dissatisfied. Patients with abnormal illness behavior typically seek repeated medical evaluations from a multitude of physicians, often undergoing a series of expensive laboratory tests. At times, the level of complaints provokes unnecessary invasive laboratory examinations or surgeries, which, in turn, thereby place the patient at genuine medical risk.

Abnormal illness behaviors may be unconscious or conscious. Unconscious abnormal illness behaviors are those in which the patient believes the symptoms reflect some genuine illness. These behaviors may occur in somatization disorder (in which multiple symptoms and organ systems are affected), conversion disorders, somatoform pain disorder (in which no cause for the subjective level of pain can be found), and hypochondriasis (in which the primary fear is of having a serious disorder). Abnormal illness behaviors in which patients act sick when they are fully aware that they are not include malingering (in which external incentives—usually financial—are the motivating factors) and factitious disorder with physical or psychological symptoms (Munchausen syndrome). In Munchausen syndrome, patients repeatedly and compulsively present themselves for medical care with feigned or self-induced illness. These self-induced conditions may be so serious as to ultimately cause death: Some patients inject themselves with feces to cause systemic infections that then warrant hospitalization and intensive care. When the self-induced natures of the illnesses are discovered, medical staffs often become outraged at these patients. The patients rarely accept or cooperate with psychiatric care, so few have been adequately studied. Most do not appear to be psychotic, but there seems to be a disturbance in personality structure. In Munchausen syndrome by proxy, a caregiver, usually a parent, induces illness in a child.

DISTURBANCES IN MOTOR ACTIVITY

Motor behavior is normally finely coordinated, purposeful, and adaptive, and necessary activities are usually carried out efficiently. In psychiatric disturbances, motor abnormalities can involve generalized overactivity or underactivity or manifest in a wide range of specific disorders of movement. Many motor disturbances are seen in psychiatric disorders. Some form part of the core symptoms of the disorders; some occur in disorders that, by their nature, bridge neurology and psychiatry (such as Tourette syndrome); others are acute or chronic medication side effects.

Overactivity

Restlessness and agitation are diffuse increases in body movement, usually noted as fidgeting, rapid and rhythmic leg or hand tapping, and jerky start-and-stop movements of the entire body, accompanied by inner tension. Restlessness accompanies psychiatric conditions of high emotional arousal or confusion, such as toxic states, deliria, mania, agitated depressive disorders, and anxiety disorders, as well as many medical disorders such as hyperthyroidism. In some depressive states, agitation is often accompanied by pacing and hand wringing.

Generalized overactivity, in which patients seem to have increased physical energy, is distinguished from agitation by its lack of inner tension and by more purposeful movements. It is commonly seen in mania, hypomania, and anorexia nervosa and as part of ADHD and in response to stimulating drugs and medicines.

> A 25-year-old extraordinarily restless man in a psychiatric emergency room, whose glances darted incessantly around the room, paced agitatedly, picked at his skin, shifted constantly on his feet, and jiggled his arms and legs. His speech was rapid, and his mood was irritable. He had been using intravenous (IV) methamphetamine on a "run" for more than 1 week and was brought to the emergency room by relatives who were concerned that he was starting to get paranoid.

In catatonic excitement, much less common now than in the preneuroleptic era, patients exhibit disorganized and overactive behaviors, including frantic jumping, thrashing of limbs, and seemingly senseless menacing or attacking behaviors. Such excitement is seen in mania, periodic catatonia, catatonic forms of schizophrenia, and some culture-bound syndromes such as amok. Confusional excitement is a state of restlessness and generalized purposeless activity seen in ictal states, some acute intoxications, and deliria.

Decreased Motor Activity

Global reductions in motor activity—motor retardation—are seen in a

variety of physical disorders, such as hypothyroidism, Addison disease, some infectious and postinfectious conditions, including CFS and postpolio syndrome, and other fatiguing conditions, as well as in some organic mental disorders, intoxications, schizophrenias, and depressive disorders. Poverty of movement (akinesia, or more properly, hypokinesia) may occur in schizophrenia and as a neuroleptic side effect. Changes in the voice frequently accompany the reduced motor activity in schizophrenia and depression, with normal inflection replaced by monotonous tone and prolonged speech latency. In stuporous states, patients remain immobile, although their eyes are open, and they are apparently awake.

Conversion reactions are functional, nonphysiological, psychogenic impairments in sensory or motor functions. Common motor forms include various paralyses and pareses, including limb paralyses, ataxias, and aphonias. In globus hystericus, the patient is unable to swallow. Patients with astasia-abasia have marked unsteadiness of gait. Sensory conversion reactions include blindness, deafness, anesthesia, and analgesia. Some hyperesthesias and pain syndromes may also originate as conversion symptoms.

An 18-year-old new army recruit from a rural area was brought to the emergency room at an Army post during his period of basic training because his gait had become very unsteady. He was stumbling into walls and seemed unable to maintain an erect posture. He found it necessary to use crutches to ambulate. There was no history of substance use. On examination, the soldier was found to be naïve, tremulous, anxious, and homesick. He had just graduated from high school, never been away from his family before and missed them deeply. When he was reassured that he would be issued a pass to see his family, his demeanor changed, and his gait instantly improved. To the amazement of his bunkmates and drill sergeant, he was able to leave his crutches at the door. The diagnosis was astasia-abasia.

Catatonic Activity

Catatonia refers to a broad group of movement abnormalities usually associated with schizophrenia but also found in other disorders such as mania, depression, many neurological disorders (especially those involving the basal ganglia, limbic system, diencephalon, and frontal lobes), systemic metabolic disorders, toxic drug states, and periodic catatonia. Catatonic stupor and excitement have already been noted. Stereotypies are repetitious, bizarre, seemingly non–goal-directed, complex organized gestures or postures that are believed to have private meanings to the patient. Examples include continuously and repeatedly crossing oneself or blessing others in a religious gesture, waving in a stylized manner, and making profane gestures. The stereotypical behaviors commonly seen in autistic children (constant spinning or rocking) may provide self-soothing, steady sensory input that helps the patients reduce the degree to which they are disturbed by the ordinarily unpredictable and uncontrollable

stimulation coming from the environment. Bizarre posturing may also be seen in catatonia. One chronic schizophrenic patient routinely stood for hours on one leg with his arms in the air like a crane. In echopraxia, the patient imitates the examiner's movements and in echolalia imitates speech, as if in mimicry. Some catatonic patients exhibit waxy flexibility, maintaining unusual postures in which they have been posed for prolonged periods of time. Negativism may take the form of refusing to behave in a prescribed manner or resisting passive movement.

Mutism

Mutism may result from a variety of peripheral muscle and CNS conditions and from functional disorders. Mutism may occur in profound depression, catatonic states, and conversion reactions. Elective mutism is occasionally seen in acute adjustment disorders and some personality disturbances.

MOTOR DISTURBANCES AND MOVEMENT DISORDERS

Tremor

Tremors, involuntary oscillating movements of the limbs or head, may occur at rest or with movement. Physiological tremors, which are minimal at rest and increase with activity, are characterized by small amplitudes and high frequencies. They are characteristic of anxiety, fatigue, and toxic or metabolic disorders, such as caffeinism or hyperthyroidism, and are commonly seen in patients taking a number of different psychiatric medications, including lithium, valproate, and stimulating antidepressants. Coarse tremors, with larger amplitudes and lower frequencies, are seen in Parkinson disease and cerebellar disease. Asterixis is a large-amplitude flapping tremor of the hands seen in hepatic disease. Parkinsonian symptoms and signs may be seen in psychiatric disorders, particularly in patients taking antipsychotic medications. Symptoms include akinesias with a marked decrease in normally spontaneous fidgeting, stiff gait with diminished arm swing, pill-rolling nonintention tremors (which seem to be less common in drug-induced parkinsonism, compared with the idiopathic type), expressionless soft and monotonous speech, micrographical handwriting, and cogwheel rigidity.

Blepharospasm

Blepharospasm is a rapid and violent repetitive, spasmodic movement of the eyelids. These movements are often a side effect of antipsychotic or other medications but are also common in a variety of neurological disorders, including Meige syndrome and Tourette syndrome.

Tics

Tics are rapid, repetitive, often spasmodic jerking involuntary movements that serve no apparent purpose. The person may try to disguise or hide the tic in a seemingly purposive movement, and the movement may ultimately be shaped into a mannerism. Tics are the central feature of tic disorders, are associated with other disorders, and may occur as a consequence of stimulant use. Tourette disorder is characterized by a chronic shifting array of motor and vocal tics. The tics may include grunts, coughs, clicks, or sniffs, whereas motor symptoms may include eye blinking, tongue protrusions, facial grimacing, hopping, and twitches. Complex tics may merge into complex compulsive behaviors such as squatting, deep knee bends, and retracing steps. Coprolalia, characterized by sudden verbal outbursts of obscenities, occurs in less than one-third of Tourette's patients. Mental coprolalia is an associated feature in which obscene words or phrases suddenly intrude into consciousness in an ego-dystonic manner. Obsessive-compulsive symptoms, as well as attention-deficit symptoms, are also common in Tourette syndrome.

A 21-year-old young man with Tourette syndrome presented to his psychiatrist's office after missing his last two appointments and being off his medications for 1 month. He had recently started a new job delivering pizza. He reported he could no longer control his compulsion to count stairs, tap his fingers, and check his car tire pressure multiple times. During the appointment, while talking, he began blinking his eyes, making guttural noises to clear his throat, followed by a coughing sound, nose picking, and loudly squealed in a manner that sounded like a cat's meow.

DRUG-INDUCED MOVEMENT DISORDERS

Dystonic Movements

Although dystonic movements are seen in many neurological disorders, in psychiatric patients, they are almost always secondary to the use of antipsychotic medications. Dystonic reactions consist of intermittent or sustained muscle spasms, typically of the head or neck. Common varieties include tongue spasms causing dysarthria, torticollis (neck spasm), and oculogyric crisis in which there is a forced upward gaze. Opisthotonus (spasms of paraspinal muscles leading to an arched posture) is seen less often. These reactions are most common in young males and typically occur soon after beginning or increasing the dose of a conventional antipsychotic medication.

Akathisia

Akathisia is a syndrome of motor restlessness seen predominantly in the context of antipsychotic and some antidepressant medication use. It has subjective, as well as motor, components. Subjectively, patients experience muscle tension, difficulty finding a comfortable body position, and inability

to stop moving; they feel as though they are "jumping out of their skin." Objectively, akathisia classically manifests by rocking from foot to foot while standing, frequently crossing and uncrossing the legs when seated, and pacing. Sleep may be disturbed because of physical discomfort. Subjective components of akathisia may be difficult to distinguish from anxiety caused by the primary disorder (typically schizophrenia). Rarely, the restlessness and inner agitation become sufficiently uncomfortable to provoke acts of violence. In pseudoakathisia, objective signs of akathisia are present, but the patient denies feeling restless.

A 45-year-old man with OCD and Tourette syndrome being treated with fluvoxamine (Luvox), 200 mg per day, inadvertently took 400 mg per day for 3 days. On the third day, he started to feel increasingly agitated, his legs felt restless and jittery, and he started to uncharacteristically fidget. He also noted frequent sighs in his breathing, not his usual respiratory pattern. He recalled feeling these sensations a decade earlier when a previous psychiatrist treated his Tourette syndrome with haloperidol (Haldol) and that his psychiatrist had diagnosed akathisia. After discontinuing the fluvoxamine for 4 days, these symptoms all abated.

Tardive Dyskinesia

Tardive dyskinesia is a movement disorder that occurs only in the context of antipsychotic medication use, occasionally after many months but more commonly after years. The abnormal movements may persist with or without continued medication use or may diminish or disappear over time. The dyskinetic movements occur at rest and can usually be temporarily suppressed volitionally or by purposeful action, distraction, or sleep. The movements are varied. In the most common type, which affects the face, especially the mouth and lips, tongue thrusting, chewing movements, lip smacking, and eye blinking are seen. Another common type is characterized by choreoathetoid movements, such as writhing finger motions. In the less common but more severe truncal dyskinesias, the torso moves in thrusting motions, and respiratory dyskinesia is characterized by grunting and irregular breathing patterns. Other tardive (late) syndromes include tardive akathisia and tardive dystonia, in which the abnormal movements emerge late in treatment or on medication discontinuation.

Neuroleptic Malignant Syndrome

Neuroleptic malignant syndrome (NMS), a potentially fatal complication of antipsychotic medication, is characterized by muscle rigidity, fever, diaphoresis, delirium, mutism, and blood pressure abnormalities. Some view NMS as the most severe end of a spectrum that starts with antipsychotic-induced parkinsonism and progresses to extrapyramidal syndrome with fever and then to fulminant NMS.

Rabbit Syndrome

This uncommon drug-induced extrapyramidal syndrome is often misdiagnosed as tardive dyskinesia. It most closely resembles a limited expression of a parkinsonian tremor. Patients make rapid chewing movements similar to those made by rabbits, ordinarily faster and more regular than the orofacial tic of tardive dyskinesia. The tongue is spared.

Serotonin Syndrome

Serotonin syndrome is a disorder caused by excessive serotonergic input in the CNS, probably within the hypothalamus. The most common cause of serotonin syndrome is the combination of two or more medications with serotonin-enhancing properties, usually by different mechanisms. It is characterized by restlessness, myoclonus, hyperreflexia, diaphoresis, shivering, tremor, autonomic changes, including fever, and mental status changes, such as confusion.

MOTOR DISTURBANCES OF SCHIZOPHRENIA

Many of the abnormal movements ascribed to tardive dyskinesia and other antipsychotic-induced extrapyramidal syndromes have been described in chronically psychotic patients before the introduction of antipsychotic medications. In one series of 100 patients, the large majority of whom were diagnosed as schizophrenic, a review of medical records before 1955 revealed that abnormal purposive movements were found in 83 percent, mannerisms and tics in 71 percent, abnormal eye movements in 27 percent, abnormal postures or facial movements in 42 percent, and gait abnormalities in 10 percent. These findings suggest that many patients with schizophrenia have neurological symptoms not due to medications and that severe psychiatric disorders may have a neurological component as well.

OTHER MOVEMENT DISTURBANCES

Gait disturbances in patients with psychiatric disorders include a variety of neurogenic gaits consistent with brain disease, intoxications, and medication side effects. These include the festinating gait of parkinsonism, spastic and ataxic gaits of neurological disease and psychiatric medications, waddling and reeling gaits associated with intoxications, and the nonphysiological gait disturbances seen in astasia-abasia, a form of conversion disorder. Gait mannerisms include clowning, prancing, and military gaits.

Bruxism, chronic jaw clenching, may occur involuntarily during tension states, as an isolated occurrence during delta sleep, in which it has sometimes been associated with benzodiazepine or alcohol use, or in association with SSRI use. In severe cases, serious damage to dental enamel and temporomandibular joint pain may occur.

Myoclonus, characterized by focal muscle jerking, can be caused in psychiatric patients by certain medications, such as SSRIs or monoamine oxidase inhibitors (MAOIs). Myoclonic jerks may be difficult to distinguish from tics, but the latter often represent larger muscle groups and more highly organized motor patterns. Myoclonus may be seen at rest but is more obvious during motor activity.

Seizurelike Behaviors

In addition to the generalized, petit mal, and complex partial seizures seen in some psychiatric patients, a number of nonepileptic seizurelike behaviors must be distinguished. Breath-holding spells, generally innocuous, impulsive, and tantrumlike phenomena, usually occur in small children who hold their breath during moments of oppositional rage and who may faint as a result. Jerking or twitching motor movements may occur. Temper tantrums in young children may look like seizures, especially to the uninformed observer. The children may lie on the floor, screaming and kicking, and do not respond to the environment. Conversion seizures (hysterical seizures or pseudoseizures) must be differentiated from genuine epileptic seizures. Patients retain consciousness, lack abnormal reflexes, and are not incontinent. However, because so many conversion seizures occur in patients who have genuine epilepsy and who know a good deal about the condition, the differential diagnosis is sometimes difficult. Compulsive behaviors may occur in relation to everyday activities, such as gambling, sexual conquest, shopping, and watching television, or in relation to substances such as alcohol, cocaine, narcotics, and food. Other compulsions involve reckless risk-taking behaviors that provide stimulation and dispel dysphoric moods. Sexual compulsive perversions, such as exhibitionism and sadomasochism, may serve similar purposes. Compulsions are seen in a variety of psychotic and nonpsychotic psychiatric disorders. The cravings that underlie compulsive behaviors are strong motivating forces, and the compulsive behaviors may regulate emotions. Unknown similarities may underlie all compulsive and addictive mechanisms. Current controversies surround the relationships of compulsive, impulsive, and addictive behaviors. Some authorities have proposed that all three should be subsumed under the rubric of a so-called reward deficiency syndrome, resulting from various combinations of gene polymorphisms governing dopamine and opioid mechanisms and including serotonin, cannabinoid, and other transmitter systems as well. From this perspective, the compulsive, impulsive, and addictive behaviors all increase the amount of dopamine and potentially other transmitters in specific brain areas. Studies in the past decade have shown that, in various subject groups, the Taq 1 a1 allele of the DRD2 gene is associated with alcoholism, drug abuse, smoking, obesity, compulsive gambling, and several personality traits. Several other candidate genes are also under

investigation. Studies have also shown that certain impulselike behaviors may be associated with very specific genetic polymorphisms. In studies of patients with severe obesity, binge-eating behavior was seen only in those individuals who had specific mutations in genes controlling the expression of melanocortin 4 receptors.

In OCD, the compulsions are ritualized, repetitive behaviors that are performed with the goal of satisfying, neutralizing, and undoing obsessional thoughts. Although intended to decrease anxiety, rituals are never more than transiently successful. The most common compulsions involve checking to make certain that gas jets and faucets have been turned off and that windows and doors are locked, hand washing, repeating certain phrases, counting objects, and placing objects in a prescribed order.

A 47-year-old woman had always been fussy about order and cleanliness but, since the birth of her child 14 years before, had become increasingly obsessed by cleanliness. She spent 3 hours daily ritualistically cleaning the apartment, using three to four rolls of paper towels daily. If a family member left even a fingerprint on a table, she became agitated and spent much time recleaning the surface and admonishing the family member.

FUTURE PERSPECTIVES

Particularly in the last decade and in the near future, there is hope that new research and advances in the neurosciences will yield an even greater understanding about psychiatric illness. However, clinical acumen in the practice of psychiatry is still as relevant today as during the origins of psychiatry. A firm conceptual grasp of the clinical descriptors of psychopathology is an important part of a clinician's armamentarium. As yet, nothing has replaced the understanding of psychiatric signs and symptoms integrated in the context of the diagnostic interview and mental status exam in formulating an accurate differential diagnosis and ultimately a comprehensive management plan. Even as psychiatrists move toward further improving and modernizing the categorical diagnostic systems vis-à-vis the DSM-5 diagnostic manual, clinicians will continue to strive to hone and refine their understanding of the clues provided in the clinical manifestations of psychiatric disorders in order to best care for their patients.

ACKNOWLEDGMENTS

The authors wish to acknowledge the work of Dr. Joel Yager and Dr. Michael J. Gitlin who wrote this chapter in the eighth edition of the *Comprehensive Textbook of Psychiatry,* part of whose work was adapted for this chapter.

REFERENCES

Alarcón RD. Culture, cultural factors and psychiatric diagnosis: review and projections. *World Psychiatry*. 2009;8(3):131–139.

Arnsten AF, Li BM. Neurobiology of executive functions: Catecholamine influences on prefrontal cortical functions. *Biol Psychiatry*. 2005;57:1377–1384.

Bakker SC, Hoogendoorn ML, Hendricks J, Verzijlbergen K, Caron S, Verduijn W, et al. The PIP5K2A and RGS4 genes are differentially associated with deficit and non-deficit schizophrenia. *Genes Brain Behav*. 2007;6:113–119.

Bhati MT. The brain, language, and schizophrenia. *Curr Psychiatry Rep*. 2005;7:297–303.

Brebion G, David AS, Jones HM, Ohlsen R, Pilowsky LS. Temporal context discrimination in patients with schizophrenia: associations with auditory hallucinations and negative symptoms. *Neuropsychologia*. 2007;45:817–823.

Cancelli I, Marcon G, Balestrieri M. Factors associated with complex visual hallucinations during antidepressant treatment. *Hum Psychopharmacol*. 2004;19:577–584.

Chamberlain SR, Sahakian BJ. The neuropsychiatry of impulsivity. *Curr Opin Psychiatry*. 2007;20:255–261.

Choong C, Hunter MD, Woodruff PW. Auditory hallucinations in those populations that do not suffer from schizophrenia. *Curr Psychiatry Rep*. 2007;9:206–212.

Feusner JD, Phillips KA, Stein DJ. Olfactory reference syndrome: issues for DSM-V. *Depress Anxiety*. 2010;27(6):592–599.

Frank GK. Recent advances in neuroimaging to model eating disorder neurobiology. *Curr Psychiatry Rep*. 2015;17(4):559.

Girard TA, Martius DL, Cheyne JA. Mental representation of space: Insights from an oblique distribution of hallucinations. *Neuropsychologia*. 2007;45:1257–1269.

Grano N, Keltikangas-Jarvinen L, Kouvonen A, et al. Impulsivity as a predictor of newly diagnosed depression. *Scand J Psychol*. 2007;48:173–179.

Hawton K, Harriss L. How often does deliberate self-harm occur relative to each suicide? A study of variations by gender and age. *Suicide Life Threat Behav*. 2008;38:650–660.

Karg K, Burmeister M, Shedden K, Sen S. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. *Arch Gen Psychiatry*. 2011;68(5):444–454.

Keefe RS, Perkins DO, Gu H, Zipursky RB, Christensen BK, Lieberman JA. A longitudinal study of neurocognitive function in individuals at-risk for psychosis. *Schizophr Res*. 2006;88:26–35.

Kounali D, Zammit S, Wiles N, et al. Common versus psychopathology-specific risk factors for psychotic experiences and depression during adolescence. *Psychol Med*. 2014;44(12):2557–2566.

Kreek MJ, Levran O, Reed B, Schlussman SD, Zhou Y, Butelman ER. Opiate addiction and cocaine addiction: underlying molecular neurobiology and genetics. *J Clin Invest*. 2012;122(10):3387–3393.

Mann JJ, Currier DM. Stress, genetics, and epigenetic effects on the neurobiology of suicidal behavior and depression. *Eur Psychiatry*. 2010;25(5):268–271.

Mocellin R, Walterfang M, Velakoulis D. Neuropsychiatry of complex visual hallucinations. *Aust N Z J Psychiatry*. 2006;40:742–751.

Raymond NC, Coleman E, Miner MH. Psychiatric comorbidity and compulsive/impulsive traits

in compulsive sexual behavior. *Compr Psychiatry*. 2003;44:370–380.

Riemann D, Kloepfer C, Berger M. Functional and structural brain alterations in insomnia: implications for pathophysiology. *Eur J Neurosci*. 2009;29:1754–1760.

Riper H, Andersson G, Hunter SB, de Wit J, Berking M, Cuijpers P. Treatment of comorbid alcohol use disorders and depression with cognitive-behavioural therapy and motivational interviewing: a meta-analysis. *Addiction*. 2014;109(3):394–406.

Schwartz CE, Wright CI, Shin LM, Kagan J, Rauch SL. Inhibited and uninhibited infants "grown up": adult amygdalar syndrome response to novelty. *Science*. 2003;300:1952–1953.

Ulloa RE, Birmaher B, Axelson D, et al. Psychosis in a pediatric mood and anxiety disorders clinic: phenomenology and correlates. *J Am Acad Child Adolesc Psychiatry*. 2002;39:337–345.

Van't Wout M, Aleman A, Bermond B, Kahn RS. No words for feelings: alexithymia in schizophrenia patients and first-degree relatives. *Compr Psychiatry*. 2007;48:27–33.

Volkow ND, Fowler JS, Wang GJ, Swanson JM. Dopamine in drug abuse and addiction: results from imaging studies and treatment implications. *Mol Psychiatry*. 2004;9(6):557–569.

Vollm B, Richardson P, Stirling J, et al. Neurobiological substrates of antisocial and borderline personality disorder: preliminary results of a functional fMRI study. *Crim Behav Ment Health*. 2004;14:39–54.

Wagner G, Koch K, Schachtzabel C, Schultz CC, Sauer H, Schlösser RG. Structural brain alterations in patients with major depressive disorder and high risk for suicide: evidence for a distinct neurobiological entity? *Neuroimage*. 2011;54(2):1607–1614.

Walder DJ, Faraone SV, Glatt SJ, Tsuang MT, Seidman LJ. Genetic liability, prenatal health, stress and family environment: risk factors in the Harvard Adolescent Family High Risk for Schizophrenia Study. *Schizophr Res*. 2014;157(1-3):142–148.

Zanarini MC, Frankenburg FR, Reich DB, Silk KR, Hudson JI, McSweeney LB. The subsyndromal phenomenology of borderline personality disorder: a 10 year follow-up study. *Am J Psychiatry*. 2007;164(6):929–935.

9 Classification in Psychiatry

▲ 9.1 Present and Future of Classification Systems for Mental Disorders

JAVIER I. ESCOBAR, M.D., AND HUMBERTO MARIN, M.D.

Introduction

This chapter will first discuss the importance of classification systems in the medical field with particular emphasis on mental disorders. Next, this chapter will expand on general diagnostic and practical issues concerning the Diagnostic and Statistical Manual (DSM) of the American Psychiatric Association (APA), the current classification in the United States. Because North American and International Classifications are getting closely aligned, these general comments will also apply to the section on Mental Disorders of the International Classification of Diseases of the World Health Organization (WHO). Finally, this chapter will discuss in detail DSM-5, the latest version of the DSM, highlight major changes from previous editions and highlight practical aspects for its clinical use.

Classification in Medicine

Classification systems are essential in all sciences. In the case of the medical field, the main functions for a diagnostic classification are communication and prediction. In medicine, "diagnosis is prognosis," as it was eloquently articulated decades ago by Donald W. Goodwin and Samuel Guze, pioneer psychiatric nosologists in North America. In the specific case of psychiatry, a critical problem in the development of classification of diseases is the fact that for most mental disorders there are no biological markers that allow identifying or separating them. Thus, in the field of mental health, the functions of a classification are even more relevant than in other medical specialties, even if these classifications are currently present only in the form of tentative charts or maps, or at best, distant, blurry "lighthouses" that guide practitioners in the open sea of clinical practice. These systems, no matter how primitive, bring practitioners a

modicum of security for dealing with the stormy issues of human behavior, thought processes, cognitions and other "subjective" phenomena, as psychiatry tries to "imitate" what other medical specialties do. In the absence of objective markers for specific disorders, several possible methods can be used to develop a classification of mental illness. The historical example is the authority-based method used for classifying mental illness until the mid-20th century, in which the descriptions of potential disorders made by some authorities were gradually accepted until they became common currency. The method adopted for current classifications (the DSM and the section on mental disorders in the ICD), had the main purpose of filling gaps in scientific evidence by using a consensual, collegiate mechanism, not very different from that of legislative bodies in a democratic state. This process, progressively engaged more and more people in the discussion and eventually got virtually everyone in the planet to utilize the same language when discussing mental illness. However, as medicine, including psychiatry, becomes more and more evidence-based, a tension has surfaced between the need to adopt the growing body of scientific evidence and the wish to satisfy the needs and interests of the constituent groups. The latter includes researchers, clinicians, patients, advocates, public and private health institutions, pharmaceutical companies, data managers, and others. Seeking more specificity, the specialty keeps building diagnosis after diagnosis in pursuit of scientific rigor, more precise treatment interventions, and better prediction of outcomes. However, as it will be seen ahead, this proliferation of disorders supported by limited scientific evidence has become troublesome. As virtually all of those practicing in the mental health today were trained during the DSM tradition, the system is generally taken for granted, used mechanically, with little thought about its origins and evolution, its benefits and limitations. Here, some background on origins, evolution, strengths, and weaknesses of the DSM are provided.

Brief Glossary

First, it is important to define key terms and concepts used herein. A *syndrome* is a cluster of symptoms that can result from different disease processes. For example, cough and fever can result from bacterial, fungal or viral infections, or autoimmunity, all processes implying different interventions and outcomes. A *disorder* is something of a medical concern in a patient, an abnormality, injury, disability, or aberration. This term is often used instead of the term *disease,* which implies a known pathological process, for medical conditions in which the causes or pathophysiology remain unknown. A *diagnosis* is simply the opinion, by someone with expertise in the matter, that a given disorder is present (or absent) in a patient, or the procedure to decide whether or not a certain disease or disorder is present (or absent). A diagnosis can be "*categorical*" (it is either

present or absent), or it may be "*dimensional*" (a certain point in a continuum). *Diagnostic Classification* is the listing of diagnoses grouped by relatedness (for example, infectious, autoimmune diseases, cancer, or injuries in medicine, or anxiety, affective or psychotic disorders in psychiatry). *Diagnostic Criteria* are the rules that need to be followed for making a diagnosis. Once made, a diagnosis is *reliable* if the same conclusion about its existence can be reached by two independent observers (inter-rater reliability) or if the patient receives the same diagnosis when examined more than once within a reasonable time period (test–retest reliability). Finally, a diagnosis is *valid,* if it picks out a "real" entity, based on etiology and pathophysiology. Obviously, the latter status is uncertain for most psychiatric conditions. It is important to note that *reliability* and *validity* are separate dimensions. In mental health, for example, the fact that different observers are able to recognize a set of symptoms when they appear together (reliability) does not necessarily mean that they correspond to a real entity (validity). In fact, the emphasis on diagnostic reliability may limit the progress on diagnostic validity, especially in the case of disorders with a complex or fluctuating symptomatology.

Definition of Mental Disorder (DSM-5)

As defined in the DSM-5 Manual, "A mental disorder is a syndrome characterized by clinically significant disturbance in an individual's cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological or developmental processes underlying mental functioning. Mental disorders are usually associated with significant distress or disability in social, occupational, or other important activities. An expectable or culturally approved response to a common stressor or loss, such as the death of a loved one, is not a mental disorder. Socially deviant behavior (e.g., political, religious, or sexual) and conflicts that are primarily between the individual and society are not mental disorders unless the deviance or conflict results from a dysfunction in the individual, as described above."

The above definition is an "ideal" one that may apply only to a limited number of conditions, as it is rare that psychiatrists can be completely assured that each syndrome in the manual reflects a specific "dysfunction in mental processes."

Brief History of Psychiatric Diagnosis

While the concept of mental illness dates back at least as far as the Greek civilization or even further back (as suggested by the finding of cranium trepanations in prehistoric times), there has always been interest and effort placed on the concept of mental disease or infirmity from ancient times to

current day, as healers, physicians, and thinkers from different centuries described mental conditions that are recognizable by today's standards. It is interesting to note that even the Greek philosophers might have entered the debate between "categorical" and "dimensional" approaches to diagnosis. Thus, the platonic ideal of classification, would consist of a system that is exhaustive and mutually exclusive, while the Aristotelian approach would conceive overlapping entities, separated from normality by arbitrary thresholds.

Some of the old, "historic," designations for mental disorders are still in use, or have been used until recent times, as is the case with hysteria, hypochondria, mania, and melancholia, the four "classic" mental syndromes of yesteryear according to Michel Foucault's influential book of the early 1960s.

However, systematic "official" classifications and nomenclatures for mental disorders are rather recent undertakings that convey additional features to the process of naming mental conditions, transforming it into a social enterprise, with implications in many domains of life (e.g., fostering scientific dialogue and progress, public discussion and education, policy decisions, reimbursement for services, data management, treatment, and others).

The introduction, by the middle of the last century, of classification systems for mental disorders that were comprehensive and practical enough to gain wide acceptance, was a critical factor contributing to the development of modern psychiatry as a branch of the medical sciences, allowing the specialty to go well beyond a simple "humanistic" endeavor. Thus, homage should be paid to the people, from multiple disciplines, that contributed to create, implement and, perhaps more importantly, reform the classification systems, so that they remain relevant up until this day.

Classical European Psychopathology

In the 19th century and beginning of the 20th century, key figures of classical European psychopathology were Emil Kraepelin, with his descriptive approach, Carl Wernicke, the main proponent of a biological approach, and Karl Jaspers who applied his phenomenological approach to psychiatric assessment and diagnosis. During that same period, North American psychiatry followed the lead of European psychiatry, but focused mainly on the Kraepelinian line. In the following, this chapter will again contrast these approaches as debates on current and future diagnostic systems in psychiatry are outlined.

The US Classification Systems Leading to DSM

The intent to record and catalogue mental illness in the United States started with the 1840 census that included, besides physical illnesses, a

category for "idiocy/insanity." In the 1880 census, "mania, melancholia, monomania, paresis, dementia, dipsomania, and epilepsy" were added and these categories were adopted by the "Association of Medical Superintendents of American Institutions for the Insane," formed in 1844, which in 1892 changed its name to "American Medico-Psychological Association," and in 1921 to "American Psychiatric Association." In 1917, together with the National Commission on Mental Hygiene, the precursor of the APA developed a "Statistical Manual for the Use of Institutions for the Insane," and also, with support from the New York Academy of Medicine, the APA precursor developed the "Mental Subsection of the United States General Medical Guide," the Standard Classified Nomenclature of Disease, usually known as the "Standard." During World War II, US psychiatrists in the armed forces were involved in the selection and treatment of soldiers, and the Office of the Surgeon General issued a classification called "Medical 203," which was adopted with modifications by different branches of the armed forces and the Veterans Administration. These different versions of "Medical 203" continued to be used, along with the Standard, contributing to the "confusion" mentioned in the foreword to the first edition of the DSM that was introduced in 1952. "By 1948, then, the situation in psychiatric nomenclature had deteriorated almost to the point of confusion which existed through medical nomenclature in the twenties. At least three nomenclatures ("Standard," Armed Forces, and Veterans Administration) were in general use, and none of them fell accurately into line with the International Statistical Classification. One agency found itself in the uncomfortable position of using one nomenclature for clinical use, a different one for disability rating, and the international one for statistical work. In addition, practically every teaching center had made modifications of the "Standard" for its own use and assorted modifications of the Armed Forces nomenclature had been introduced into many clinics and hospitals by psychiatrists returning from military duty."

International Classification of Diseases

The ICD is the descendant of the "Bertillon Classification of Causes of Death," introduced by the French physician Jacques Bertillon in 1893, at the congress of the International Statistical Institute, in Chicago. Multiple countries adopted this classification, including those in North America, and in 1900 an international conference revised what was then called the "International Classification of Causes of Death," and agreed to hold revisions of the system every 10 years. In 1948, the recently created WHO, an agency of the United Nations, assumed responsibility for the ICD. The sixth revision, published in 1949, involved significant changes, as it included morbidity in addition to mortality, so that the title was changed to International Statistical Classification of Diseases, Injuries, and Causes of

Death. It is important to point out that the sixth revision of the ICD was the first one to include a section on mental disorders. The ICD is now in its 10th official edition and they are currently working on the 11th that will be released in the next couple of years. For a while now, the DSM and the ICD section on mental disorders have worked in a collaborative way and the two systems are now very similar. Thus, the general comments that follow, on the DSM, may be applied to the mental disorders section of the ICD.

Evolution of Modern Psychiatric Nosology in the United States

In debating the merits and limitations of the DSM (and obviously, the mental disorders section of the ICD system as well), it is important to remember the historical context operating at the time when these systems originated and how the classifications have evolved (or regressed) with the passage of time. The first edition of the DSM was published in 1952, and the first ICD including a mental nosology, the ICD-6, was published in 1949. This was long before the emergence of "evidence-based" medicine and psychiatry, at a time when mental nosology in the United States was in a relatively "dormant" state. Indeed, those were the 1940s and 1950s, the decades of the clinical and theoretical hegemony of both the psychoanalytic theory and the biopsychosocial approach of Adolf Meyer. For both of these theories, mental illness was a reaction of the mind to the environment, the family environment for psychoanalysis, and psychological, social, and biological factors for Meyer and his followers. Thus, in clinical practice, diagnosis made little difference and explanatory systems based on the laws of thermodynamics guided clinical practice and brought comfort to practitioners. Most clinicians practicing psychiatry then, learned their trade and applied it to their clinical practice with little regard to classifications, research evidence, theories of disease or potential treatments. A few enclaves in the United States, particularly in the mid-Atlantic area (St. Louis, Baltimore, New York) maintained at least part of the European tradition (again, primarily the Kraepelinian one) and were influenced by the medical model approach and philosophical realism (positivism, empiricism) more than the phenomenological model (Jaspers).

It was only in the 1950s, that evidence-based psychiatry began to emerge with the development of the first sets of psychotropic drugs backed by systematic research, such as lithium, the phenothiazines, tricyclic antidepressants, MAO inhibitors, and benzodiazepines. Evidence-based psychotherapies such as cognitive behavioral therapy (CBT) had to wait much longer, until the 1970s, to be confirmed as "effective" by well-controlled clinical trials.

This was the environment in which the first edition of the DSM (without numeral) was produced, in 1952, by a group of seven people commissioned by the APA. The 1952 DSM lists 108 disorders in a descriptive section that takes only 32 small pages of text. However, a significant number of these

items correspond to neurological and other medical disorders that have since migrated to other specialties. Thus, the actual amount of mental disorders in the first DSM is in the two-digit range, a number that has grown to over 400 in DSM-5 (Fig. 9.1–1). This explosion in the number, and the detail, of the psychiatric diagnoses started with the DSM-III, and will be analyzed in more detail ahead in this chapter.

It is important to observe that, despite the dominance of the DSM in the United States since the introduction of its first edition in 1952, other methodologies and proposals of psychiatric classifications have periodically been offered, differing from the DSM mainly because of the emphasis on scientific evidence. Some of these proposals have influenced the DSM, without making it completely evidence-based yet, as will be discussed later.

The United States/United Kingdom study, demonstrating that training and collaboration across the Atlantic would make compatible the diagnostic practices of North American and British psychiatrists and the International Pilot Study of Schizophrenia (IPSS), showing that concordant schizophrenia syndromes could be reliable identified in nine countries of the world, had both an important impact on modern diagnostic developments. The Present State Examination (PSE) developed at Maudsley hospital under the leadership of John Wing, a psychiatrist adhering to Jaspers' phenomenological approach, provided excellent symptomatic descriptions of the major psychiatric syndromes, and to this date remains a unique instrument to assess psychopathology. Its rich descriptions of symptoms and signs make it ideal to teach psychiatrists in training about the range of psychopathology.

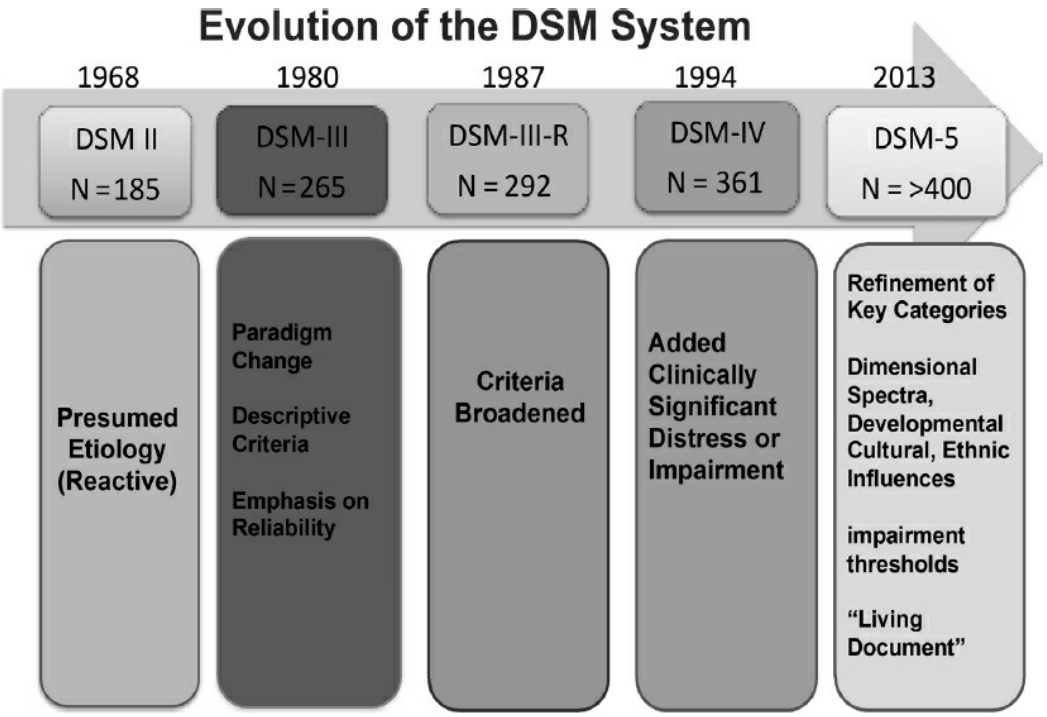

FIGURE 9.1–1. The DSM system. Key issues and number of categories.

Modern Psychiatric Classifications in North America

The St. Louis Group. Modern psychiatric classifications in North America were given a "jump start" in the 1960s, when a group of investigators at Washington University and Renard Hospital in St. Louis, MO (one of the "neo-Kraepelinian" enclaves mentioned above), led by Eli Robins and Samuel Guze, started using a process inclusive of successive steps, such as clinical description, laboratory studies, delimitation from other disorders, follow-up studies and family history, for validating psychiatric syndromes customarily diagnosed in clinical practice. This led to a list of 12 disorders that were deemed "valid" at that time, as they went through the various validation steps listed above. These "valid" entities included hysteria (somatization disorder), panic disorder (anxiety neurosis), obsessive-compulsive disorder, phobic disorder, affective disorders (depression, mania), schizophrenic disorder, alcoholism, drug dependence, sociopathy (antisocial personality), brain syndrome, anorexia nervosa, and sexual problems. While the parsimonious Eli Robins and Samuel Guze were not rushing to publicize this quite yet, an ambitious psychiatric resident, the late John P. Feighner, convinced his mentors to prepare a paper for publication. The paper, published in the Archives of General Psychiatry in 1972, has become one of the most quoted papers in the history of psychiatry. Curiously, Goodwin and Guze never mention Feighner's paper in any of the editions of the classical textbook "Psychiatric Diagnosis" they coauthored.

Research Diagnostic Criteria. Next, in the 1970s, came the Research Diagnostic Criteria (RDC), an effort led by Robert Spitzer at Columbia University (another of the US enclaves where classical psychopathology had a niche). The "neo-Kraepelinian" RDC, largely based on the St. Louis criteria, set the stage for a "paradigm change" in the official APA classification, as Robert Spitzer was named to lead the APA task force on DSM-III. Spitzer would later state: "With its intellectual roots in St. Louis instead of Vienna, and with its intellectual inspiration drawn from Kraepelin, not Freud, the task force was viewed from the outset as unsympathetic to the interests of those whose theory and practice derived from the psychoanalytic tradition." From that point on, most everyone knows the rest of the story. However, this chapter will return to the later iterations of the DSM ahead.

Heuristic Formulations for Psychiatric Diagnosis. In the late 20th century, in Europe, the "psychological dysfunction" model proposed by the British scientist Sir Aubrey Lewis, theorized that constructs such as "phobias," "obsessions," or "thought disorder" were probably linked to biological mechanisms. He also supported the concept of single depression

(rather that the melancholic/reactive dichotomy), as a continuum, with "melancholia" in one extreme and "neurasthenia" in the other. Lewis was a key figure in the ICD-8 classification of mental disorders, the first one containing a glossary. According to Fulford and Sartorius, while the ICD-8 and the glossary were based on a report to the WHO in the early 1960s by British psychiatrist Erwin Stengel (which in turn was based on a logical empiricist analysis of psychiatric classification developed by an American philosopher of science, Carl Hempel), it was Lewis who initiated the key move from the theoretical (etiological) to the descriptive (symptom based) classification of mental disorders.

Similarly, DSM-IV-R was significantly influenced by the theories of Jerome Wakefield, another philosopher of science, at the time based at Rutgers University in New Brunswick, NJ. Wakefield proposed a "harmful dysfunction analysis" as theoretical backup for psychiatric classifications. According to this theory, there is a subjective element that implies that the condition is "harmful" or "undesirable" to the individual and an objective element that assumes that the condition is due to a "malfunctioning internal mechanism." Together with Allan V. Horwitz, a social scientist, also at Rutgers, Wakefield published in 2007 a book entitled "The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder." The authors argue in the book that sadness is a normal human emotion that appears in front of losses and misfortunes and they criticize the almost exclusive biological focus of psychiatry in interpreting and managing this heterogeneous disorder. To bring into play the contribution of the social sciences, these authors proposed, following the path of DSM-III and IV, the use of bereavement as an exclusionary criterion for depression, adding an additional clause to incorporate and expand relevant contextual factors that would rule out depressive disorder. These would include not only bereavement but also other major life stressors or losses such as marital dissolution, loss of a valued job, and other life events. Little they knew (as will be described ahead under depression), that just a couple of years later, DSM-5 would also strike at the bereavement exclusion. Dominique Murphy, another philosopher interested in psychiatric diagnosis (who also spent time at Rutgers), in his book "*Psychiatry in the Scientific Image*." criticizes the "atheoretical" view of the later editions of the DSM and advocates for a "causal" nosology. While acknowledging limitations of current theories about causes of mental disorders, Murphy believes it is still possible to move forward even if the theory is not completely worked out. This may be the philosophical or theoretical underpinning for new diagnostic initiatives such as the Research Domain Criteria (RDoC).

Classifications in Medicine. Despite their pretension of "scientific objectivity," classification systems in the applied sciences (especially the

health sciences) are not created in a social vacuum by completely neutral scientists or practitioners. At the end, these systems represent essentially a consensus or compromise between the needs, beliefs, and expectations of different stakeholders operating at a particular time and place. The complexity of the deliberations is compounded by the fact that the interests and convictions of the various stakeholders may coincide, disagree, or be indifferent to those of the others, and may also change over time. As mentioned early in this chapter, diagnostic criteria for diseases in other areas of medicine can be often confirmed with objective, verifiable procedures, akin to a "gold standard." However, there are numerous medical conditions, particularly those entities termed "functional" (chronic fatigue, fibromyalgia, irritable bowel, and many others), for which objective, verifiable, and measurable "gold standards" are not available and whose "validity" is as questionable as that of many mental disorders. Several other "physical" conditions, such as asthma, chronic obstructive pulmonary disease (COPD) and diabetes among others, continue to present significant challenges for their objective assessment and proper management. For example, criteria for diagnosing COPD includes first a set of "key indicators" such as "dyspnea," "chronic cough and sputum production," "wheezing and chest tightness," "history of exposure to risk factors" and "family history." Some of these indicators (e.g., dyspnea) are quite subjective, may be hard to assess or detect, particularly in early stages of the disease, and can be shared by a multitude of conditions. "Objective" assessments such as physical examination or chest x-rays, convey limited information in early COPD, and results from "the confirmatory test," the true "gold standard," spirometry, are often difficult to interpret unambiguously. Even for disorders with clear, measurable markers such as diabetes, diagnostic thresholds change or fluctuate as classifications are updated. For decades, the diagnosis of diabetes has been based on glucose criteria either the fasting plasma glucose (FPG) or the 75-g oral glucose tolerance test (OGTT). In 1997, the "expert committee on the diagnosis of diabetes" (the equivalent group to the DSM task force) revised the criteria on the basis of retinopathy assessed with new technologies and recommended to lower the cutoff point of FPG from 140 mg/dL to 126 mg/dL. Thus even in the case of measurable markers, frequent inspection and revision of criteria seems warranted. To further complicate matters, some authors insist that if the "gold-standard" for medical diagnosis is defined as the "measure or precise equivalent that is verifiable," autopsy and biopsy remain the ultimate "gold standards" in the field. This is a main reason why the "consensus of experts" continues to be frequently used as the "gold standard" in medicine, as many medical conditions are not consistently associated to an essential lesion that is universally accepted. Thus, in many instances, quoting from the neurological textbook authored by A. Lerner in 2010, "diagnosis in medicine remains a deductive exercise."

Of course, the situation of psychiatric diagnosis is much more problematic, as the "symptoms" for most psychiatric syndromes consist of verbal reports and probes obtained directly from the patient, and the "signs" are the observed behaviors, detected through the traditional methods of anamnesis and mental status examination. In most instances, the current method for defining mental disorders involves a "willing suspension of disbelief," as is accepted, without conclusive proof, that the set of symptoms ("the syndrome") originate from a common source or that they are at least part of some type of unitary functional unit, as in the case of anxiety, depression, mania, psychoses, or dementia.

Advantages of the DSM System. The hybrid (scientific/collegiate/consensus) mechanism adopted by the APA (especially for more recent versions of the DSM) was based on focused discussion by groups of experts under the umbrella of the organization. Since the APA, in general, has had good social and scientific standing, this mechanism has been a good practical solution throughout the years, both from the scientific and political points of view. The initial impact of clearly and succinctly defining key syndromes such as schizophrenia, panic anxiety, melancholic depression, manic-depressive illness or dementia, backed by solid historical descriptions in the scientific literature, observations on heritability and other features, supported the unitary character of the key syndromes. These were the clinical entities acknowledged by the St. Louis Group in the 1960s, whose existence and validity received a powerful confirmation with the therapeutic advances of the 1950s, and 1960s, showing that there were pharmacological agents that specifically targeted each of these syndromes and that apparently, "treatment-response" to a certain agent was a "marker" for a certain clinical condition. This was of course an illusion that gradually dissipated with the passage of time, the accrual of evidence from randomized clinical trials and cumulative clinical experience.

Shortcomings of the DSM System. The most important intrinsic weaknesses attributed to the DSM include:

- ► The proliferation of new disorders in each new version of the DSM classification with only limited support in scientific evidence
- ► The problem of the arbitrary thresholds, for example, number of symptoms or duration in time required for a diagnosis
- ► The frequent need to use nonspecific diagnostic codes such as the "not otherwise specified" (NOS).
- ► The problem of comorbidity, as a significant number of patients qualify for a high number of DSM diagnoses simultaneously.

All these problems are intimately related, and will be briefly discussed

below:

As mentioned before, only a handful of diagnoses from the first two versions of the DSM have survived in recent editions, while the total number of diagnoses now exceeds 400 in DSM-5.

Reasons for this explosion in number include:

- The fragmentation of well-accepted diagnoses, such as, for example, anxiety, into a number of "subtypes" of decreasing validity. In order to create these new disorders, the detail of each diagnostic definition keeps growing, incorporating complex and arbitrary thresholds. With the increase in detail and restrictiveness, the clinical representativeness of each definition tends to diminish and trivial disorders may be reified.
- The persistence and the creation in each DSM edition of "mental disorders" defined on the basis of their discordance with culturally sanctioned rules of behavior (such as the personality disorders and a number of sexual disorders).
- The proliferation, in the absence of scientific evidence, of disorders defined by circumstances external to the individual, such as the addictive disorders, separated according to individual substances, the behavioral addictions, or the phobic disorders.
- The need for practitioners to use vague, ambiguous codes such as the NOS categories in DSM-IV, may arise to a large extent from the conflict between the increasing detail and restrictiveness of the DSM classification and the heterogeneity of clinical populations regarding number and assortment of symptoms, their duration or severity and sequencing or fluctuation over time. Thus, the more "precise" and restrictive the definition of a mental disorder becomes, the lower the proportion of clinical patients that would meet the criteria for that specific disorder and the larger the number of patients that may have to be allocated into nonspecific categories.
- Unfortunately, this increased detail of the definitions on paper, is generally interpreted by administrators and data managers as equivalent to an improved diagnostic ability in clinical practice, which is clearly not the case. An emerging reality is that insurers and data managers may end up rejecting nonspecific categories and demand formulation of precise diagnoses at the risk of losing reimbursement. Moreover, arbitrary thresholds such as the 6-month duration rule may exclude many cases that would go on later to develop full-blown syndromes.

Increased Complexity of the DSM Process. The first DSM was prepared by the "Committee on Nomenclature and Statistics of the American Psychiatric Association" and published in 1952. This first DSM listed seven people as members of the "APA Committee on Nomenclature and

Statistics" that produced the manual and this remained unchanged for DSM-II (ten people, including two who were "consultants" and one "representing the council"). However, the explosion in the size and complexity of the DSM system starts with DSM-III. The number, functions, and size of committees increased exponentially and the decision-making was organized in various layers. The DSM-III included a "Task Force on Nomenclature and Statistics" with 19 members, 12 Advisory Committees (each targeting specific types of disorders) with up to 18 people each, as well as other committees such as the "Multiaxial Diagnosis Committee," the "Glossary of Technical Terms Committee," the "Assembly Liaison Task Force," and the "Board of Trustees Ad Hoc Committee" plus numerous consultants. DSM-III, with its operational structure and descriptive criteria presented in a "Chinese menu" format, quickly became the predominant paradigm, at a time of limited treatment resources, insufficient research, and scarce evidence about causality. An enhancement of communication was the paramount initial consideration at the time, and DSM-III seemed quite successful for accomplishing this goal, as attested by its worldwide use. Responding to the publication of DSM-III in 1984, Gerald Klerman, an influential psychiatrist, also a "neo-Kraepelinian," stated categorically "the problem of diagnostic reliability has been solved."

The ensuing DSM-III-R simply broadened the criteria and DSM-IV adding a requirement for "Clinical Significant Distress or Impairment."

DSM-5. The DSM-5 task force was officially convened in August 2007. It originally consisted of 26 members, 20 of them US academic psychiatrists with high visibility in different areas of the specialty, one representative for psychiatry residents, one representative for patients, two psychiatrists representing APA governance (board and assembly) and one psychiatrist representing the World Psychiatric Association. As in the case of DSM-III and IV, the DSM-5 task force was the steering body for the whole process. The first step was to name several workgroups with responsibility for specific diagnostic areas, each led by a member of the task force. Each of these workgroups had on the average, about 10 members, as well as long lists of consultants. For the ensuing 5-year period (2008 to 2013) the task force and workgroups met frequently both in person and via teleconference, reviewing the evidence base related to current diagnoses as well as that for potential new ones. The initial goal was to come up with a new system that capitalized on the gains of the "decade of the brain" with the tremendous progress made in basic neuroscience and the enthusiasm generated by the sequencing of the human genome. Because the risk for key disorders such as schizophrenia, bipolar and major depressive disorder has a major familial component, it was hoped that by the time of the new DSM edition, specific genetic markers such as polymorphisms or structural abnormalities would have

been identified, which would "tag" certain disorders and help dissecting them into meaningful phenotypes. Another expectation was that the increased research using objective tools such as structural and functional neuroimaging would yield reliable "markers" for several psychiatric disorders. Also, the initial hope was that DSM-5 would become a second "paradigm shift" for psychiatric diagnosis (DSM-III has been viewed as the first), by incorporating a dimensional approach as alternative to a categorical one, considering new models of classifying psychiatric syndromes (e.g., brain "circuitry"), and hopefully decreasing rather than increasing the number of diagnostic categories in the manual. Upon gathering the task force and the workgroups and reviewing the available evidence, the unbridled enthusiasm on finding "markers" from neuroimaging or genetic studies and the possible use of novel strategies such as "brain circuitry" to classify mental disorders, quickly dissipated and "reality" would soon set on the horizon. It thus became clear that the main task of DSM-5 should be an effort to enhance the current descriptive, operational approach, by eliminating dubious syndromes, improving wordings and definitions, examining the evidence for new diagnostic entities and the practical utility of a dimensional approach. As part of this process, the task force carefully examined previous versions of the DSM, particularly DSM-IV, to identify key areas in need of revision or elimination. There was awareness on the fact that in defining mental disorders, previous versions of DSM had not described or mapped natural phenomena accurately, implying—using the traditional analogy of carving a turkey—that the diagnostic system "did not carve nature at the joints." It became clear that the multiple categories with doubtful boundaries forced clinical practitioners (at least the most scrupulous ones) to overutilize the "NOS" or "NOS" diagnoses. This was confirmed by a survey of general use, showing that the most common diagnosis made by psychiatrists in North America was "Depressive Disorder NOS." The discussions also acknowledged that while DSM-III (and also DSM-IV) responded to a strong desire for inter-rater reliability by developing operational criteria, they arbitrarily treated every disorder, as a category that was discontinuous from "normal" and from other disorders, thus avoiding dimensional approaches for descriptions of disorders. It was also felt that when applied rigidly, DSM-IV criteria appeared to limit research and therapeutic developments. Also, important components of key syndromes, such as cognitive problems in schizophrenia were not emphasized in the previous DSM criteria and were largely ignored until NIMH took the initiative with the MATRICS project.

As the process of DSM-5 evolved, the revision of extant categories of mental disorders was also being proposed for the ICD by leading psychiatrists from outside the United States. For example, the British psychiatrist David Goldberg highlighted that the three main problems of

DSM-IV and ICD-10 classifications were the high rates of "comorbidity" produced by current rules, the increased use of nonspecific (NOS or NEC) diagnoses by practicing clinicians and the fact that each edition was longer and more complex than the one preceding it.

The Process of Approving/Disapproving Changes or New Diagnoses in DSM-5. In contrast with the simple structure of the first DSM, in which one committee made all the decisions, DSM-5 had to deal with several levels or layers for clearance and approval of the new proposals. The first premise was that any changes to DSM-IV or proposals for new diagnoses had to be evidence-based. Thus, on top of the already complex structure of the DSM-III and DSM-IV, a consulting body charged with veto power was added. Instead of going directly to the task force, committee recommendations went to an "independent" body called the "Scientific Advisory Committee" (SAC), created by APA leadership as an important step to review the scientific justification for proposed changes and assure the rigor and consistency of the DSM review process. This committee functioned outside the DSM organizational structure and reported, in an advisory role, to the APA President and Board of Trustees overseeing the DSM process. A second committee, called the "Clinical Public Health Committee" or CPHC, also named by APA leadership outside DSM's organizational structure, reviewed proposals keeping in mind potential practical or public-health implications of proposed changes and new diagnoses and also made specific recommendations. The CPHC used a large list of advisors from academia and clinical practice as "peer reviewers." The SAC and CPHC recommendations then went to the full task force for final debate and a decision on whether or not new diagnoses and significant changes in existing diagnoses should be implemented. These final recommendations were then forwarded to the APA Board of Trustees for their review. The Board had the final say about changes, revisions, and new diagnoses that were printed in the DSM-5 manual. The number of people involved in the DSM-5 process was far higher than that of previous editions. More importantly, about 30 percent of workgroup members were international colleagues from countries outside the United States.

DSM-5 Field Trials. These trials specifically examined the reliability of DSM-5 diagnostic changes (e.g., depression, PTSD, schizophrenia) and new diagnoses (e.g., hoarding disorder, somatic symptom disorder). The trials were rigorously designed and controlled. People with specific disorders were matched to people without the disorders and examined independently by two clinicians, who did not know the patient or his/her diagnosis, in two separate occasions. These trials took place at several academic centers in the United States and included both children and adult populations. As the major purpose of these trials was to assess reliability, a set of expectations

was formed based on revisions of the literature on current reliability estimates in the medical field. The key measure of reliability was the intraclass kappa coefficient, a measure of how well a first diagnosis predicts a second independent one. A literature review by Helena Chmura Kramer, PhD and collaborators in 2002 had noted that κ > 0.8 would be an "extraordinary" result; κ 0.6 to 0.8, an "excellent" result; κ 0.4 to 0.6, a "good" result; κ 0.2 to 0.4, an "acceptable" result, and κ < 0.2 or lower, an "unacceptable" result. This review found "excellent" to "good" reliabilities with such procedures as coronary arteriography, x-rays, inspection of conjunctives for anemia and liver scans. Clinical assessments such as the "Lasègue sign" and estimations of reliability to judge "quality of medical care" tended to have only an "acceptable" level of reliability, while procedures such as pelvic examinations in the ER had "unacceptable" levels of reliability.

Results of the DSM-5 field trials showed "excellent" agreement for major neurocognitive disorder, posttraumatic stress disorder, somatic symptom disorders, autism spectrum disorder (ASD), and most of the self-rated "cross-cutting" symptom-based measures. Schizophrenia and bipolar disorder showed "good" reliability. Major depression had a disappointingly low level of reliability, and subtypes such as "anxious depression" were completely unreliable. These field trials were much more rigorous than those of DSM-III and DSM-IV, and gave confidence on the robustness of most of the changes and the few proposed new diagnoses that were tested (Tables 9.1–1 and 9.1–2).

Key Changes in DSM-5

THE DSM-5 MANUAL. The Official Manual of the fifth edition of the DSM has 947 pages. The manual is divided into three main sections.

Section I includes "the DSM-5 basics" such as introduction, use of the manual, and a cautionary statement about forensic use of the manual.

Table 9.1–1.
Reliability of Major DSM-5 Diagnoses

Disorders	kappa
PTSD	0.67
Neurocognitive disorder	0.78
Somatic symptom disorder	0.61
Hoarding disorder	0.59
Bipolar	0.56
Schizophrenia	0.46
Major depression	0.27
Mixed anxiety depression	−0.004

Table 9.1–2.
Reliability of Dimensional Diagnoses and Crosscutting Symptom Measures

	kappa
Dimensional Measures	
Autism spectrum disorder	0.69
New (alternate) personality disorders	0.41
Crosscutting Symptom Measures	
Psychosis	0.80
Depression	0.79
Anxiety	0.75
Somatic symptoms	0.73

Section II, the key section of the manual, includes all the diagnostic criteria and codes for the officially approved diagnoses.

Section III includes new diagnoses not ready for "prime time," an alternate category of personality disorders, instrumental tools for dimensional measurements as well as an instrument for assessing functional disability (WHODAS).

There is also an "Appendix" that includes highlights of changes from DSM-IV to DSM-5, a glossary of technical terms, cultural concepts of distress and alphabetical listing of DSM-5 diagnoses and ICD 9-CM and ICD-10 codes.

DELETION OF THE MULTIAXIAL SYSTEM. The five-axis frame of DSM-III and IV was eliminated in DSM-5. This makes DSM-5 closer to the structure of the ICD-10 diagnostic system that has traditionally used a single axis. However, regarding the number of diagnoses or their quality, this reduction in the number of axis may have been inconsequential, as the personality disorders, one of the most questioned diagnostic groups, which had their own axis, Axis II, in DSM-IV, did not disappear but may have gained more relevance by deserving a listing among Axis I disorders.

DIMENSIONAL APPROACH. In the introduction to DSM-5, it is stated that on the basis of early DSM-5/ICD-11 collaborations in diagnostic developments, clustering of disorders into "externalizing" versus "internalizing" perspectives represents an empirically supported framework. While strongly insinuated early in the process, a true dimensional perspective did not actually appear in DSM-5. Perhaps, the only noncategorical approach in DSM-5 are the "crosscutting" symptom dimensions (depression, anxiety, somatic symptoms, and others) that tend to be distributed across many diagnoses. These are based on self-rated measures, a majority taken from the Patient Reported Outcome Measurement Information System (PROMIS), developed under the sponsorship of NIH. These are dynamic tools to measure health outcomes

from the patient's perspective. They include symptom dimensions for pain, fatigue, physical functions, depression, anxiety, social function, etc., and are available for children and adults. These symptom dimensions are included in section III of the manual, with the hope that clinicians may include them in the clinical assessments.

STRUCTURE AND GROUPING OF DISORDERS. Major changes in DSM-5 included the removal of PTSD and acute stress disorders from the anxiety disorders category in DSM-IV and their placement in a new stand-alone chapter called Trauma and Stress Related Disorders (TSRD). Obsessive-compulsive disorder was also moved out of the anxiety disorders chapter and also placed in a new chapter called Obsessive-Compulsive and Related Disorders. This group also includes "body dysmorphic disorder," "trichotillomania" (hair pulling disorder), a new diagnosis "hoarding disorder," as well as excoriation or "skin-picking" disorder.

NEW DISORDERS. Key new disorders included in the main section (section II) of DSM-5 are "Autism Spectrum Disorder," "Somatic Symptom Disorders," "Hoarding Disorder" (a diagnosis of major public health significance, as every department of health in the country has had to deal with the hoarding issue, either of animals or things), "Skin-Picking Disorder," "Intellectual Developmental Disorder" (formerly Mental Retardation), "Disruptive Mood Dysregulation Disorder," "Mild Cognitive Impairment," "Insomnia Disorder," "Nightmare Disorder," "Suicide Disorder." New Disorders, judged not to be ready for "prime time" were included in section III of DSM-5. These include the Alternate Category of Personality Disorders, Persistent Complex Bereavement Disorder, Caffeine Use Disorder, Internet Gaming Disorder, Neurobehavioral Disorder Associated with Prenatal Alcohol Exposure, Suicidal Behavior Disorder, and Nonsuicidal Self-Injury.

NEURODEVELOPMENTAL DISORDERS. Mental Retardation was renamed Intellectual Disability (Intellectual Developmental Disorder). Autism Spectrum Disorder replaces Autism, Asperger Syndrome and Pervasive Developmental Disorder NOS in DSM-IV.

NEUROCOGNITIVE DISORDERS. These disorders were previously referred in DSM-IV as "dementia, delirium, amnestic, and other cognitive disorders." In DSM-5, they are headed by "delirium." A new category in DSM-5 is "mild neurocognitive disorder," added as a lesser form of cognitive deficit, to contrast it with "major cognitive disorder" and the "major-minor" dichotomy is used for the remaining disorders in this category including "major or minor neurocognitive disorder due to Alzheimer disease," "major or mild frontotemporal neurocognitive disorder," "major or mild neurocognitive disorder with Lewy bodies."

SCHIZOPHRENIA SPECTRUM AND OTHER PSYCHOTIC DISORDERS. Included under this heading are "schizotypal personality disorder," "delusional disorder," "schizophreniform disorder," "schizophrenia," "schizoaffective disorder," "substance/medication-induced psychotic disorder," "psychotic disorder due to another medical condition" and "catatonia," including several subtypes. In the case of DSM-5 schizophrenia, all subtypes were eliminated and criteria for the disorders were slightly modified. This included removal of special consideration for "bizarre" delusions and "special" hallucinations under Criterion A (characteristic symptoms). Rewording of negative symptoms in efforts to provide more clarity ("affective flattening" was changed to "restricted affect"; "alogia or avolition" were changed to "avolition/asociality"). Also, in order to meet criterion A for the disorder, at least one of two required symptoms must be delusions, hallucinations, or disorganized speech.

BIPOLAR AND RELATED DISORDERS. This group includes "bipolar I disorder," "bipolar II disorder," "cyclothymic disorder," "substance/medication-induced bipolar and related disorders," "bipolar and related disorder due to another medical condition," "other specified bipolar and related disorders," and "unspecified bipolar and related disorder." In the case of bipolar disorder, inclusion of "increased energy/activity" as a criterion a symptom of mania/hypomania. Also addition of "mixed features" specifiers for mania, hypomania, and depression (see below). The complete list of "specifiers" for bipolar and related disorders include "anxious distress," "manic or hypomanic episode with mixed features," "depressive episode with mixed features," "with rapid cycling," "with melancholic features," "with atypical features," "with psychotic features," "with catatonia," "with seasonal pattern."

DEPRESSIVE DISORDERS. Included here are "Disruptive Mood Dysregulation Disorder," a new disorder intended to decrease the "excess" of diagnoses of bipolar disorder observed in children and adolescents. "Major Depressive Disorder," remains virtually identical to the DSM-IV category, except for the elimination of the bereavement exclusion (this was done, despite statement under definition of mental disorder in the introduction to DSM-5 —see definition for a mental disorder in page 1152 of this chapter). "Persistent Depressive Disorder (Dysthymia)"; "Premenstrual Dysphoric Disorder"; "Substance/Medication-Induced Depressive Disorder"; "Depressive Disorder Due to Another Medical Condition"; "Other Specified Depressive Disorder"; and "Unspecified Depressive Disorder." Also, there is list of "specifiers," including a new one, "with anxious distress," in addition to "with mixed features," "with melancholic features" and "with atypical features."

ANXIETY DISORDERS. With the transfer of PTSD and obsessive-compulsive

disorder to stand-alone groupings, the disorders remaining in this category are "separation anxiety disorder," "selective mutism," "specific phobia" (with several subtypes), "social anxiety disorder," "panic disorder," "agoraphobia," "generalized anxiety disorder," "anxiety disorder due to another medical condition," "other specified anxiety disorder," and "unspecified anxiety disorder."

TRAUMA AND STRESSOR-RELATED DISORDERS. "Posttraumatic Stress Disorder (PTSD)" and "Acute Stress Disorder" were moved out of the anxiety disorders group (where they were in DSM-IV) to this new, stand-alone category. The criteria for PTSD were carefully revised. In DSM-IV, PTSD included three clusters of symptoms: (1) reexperiencing fear memories (intrusive memories), (2) avoidance of trauma reminders (amnesia, withdrawal, avoidance of situational reminders), and (3) hyperarousal symptoms (sleep disturbance, heightened arousal response). In DSM-5, the definition of a "traumatic event" was clarified. One extra criterion, D3 was added (persistent negative emotional state). Criteria C now focus on "avoidance" while criteria D focuses on negative cognitions and mood. Language and descriptions were enhanced for precision. Basically, DSM-5 has expanded the definition of PTSD beyond the "fear" construct, adding another cluster of symptoms to acknowledge negative alterations in cognition and mood (pervasive negative emotional states, negative appraisals about oneself or the world and excessive blame). In total, DSM-5 increased PTSD symptoms to 20, up from 17 in DSM-IV. Potential combinations of symptoms got increased more than 10-fold. According to some experts, this makes the disorder more heterogeneous and may complicate the research agenda. "Adjustment Disorders," are also listed under TSRD together with "Reactive Attachment Disorder" and a new condition, "Disinhibited Social Engagement Disorder." The latter emphasizes the recognition of symptoms associated with neglect during childhood. Regarding the effects of these changes in clinical practice, experts have stated, "there is potential for improved assessment of stressor-related conditions in children."

SUBSTANCE USE DISORDERS. Substance abuse and substance dependence were consolidated into a single disorder called "substance use disorder." A continuum or dimension was created that includes "mild," "moderate," or "severe" substance use. One of the DSM-IV abuse criteria was removed ("legal consequences such as multiple arrests"), and a new criterion, "craving," was added.

SOMATIC SYMPTOM AND RELATED DISORDERS. This new category replaced the somatoform disorders category from DSM-IV. All diagnoses that had been subsumed under somatoform disorders category in DSM-IV, were merged into a common category, significantly revised, or placed in another section

of the diagnostic manual. The DSM-IV requirement for symptoms to be "medically unexplained" was removed from DSM-5. Somatization Disorder, Pain Disorder, and Undifferentiated Somatoform Disorder were all merged into a new diagnosis called "Somatic Symptom Disorder." Hypochondriasis became "Illness Anxiety Disorder." "Conversion Disorder" was significantly revised and the term "Functional Neurological Symptom Syndrome" was added to the title of this disorder in parentheses. "Psychological Factors Affecting Other Medical Conditions," "Factitious Disorder," "Other Specified Somatic Symptom and Related Disorder," and "Unspecified Somatic Symptom and Related Disorders" were the new disorders adopted in the new category of somatic symptom disorders. "Body Dysmorphic Disorder" was moved to the new obsessive-compulsive and related disorders group. In terms of specific criteria, Somatic Symptom Disorder requires one or more somatic symptoms to be present, but unlike DSM-III and IV, in DSM-5, the emphasis is placed on the behavioral and psychological elements of the syndrome such as preoccupation and concern with the physical symptoms. It is also possible for Somatic Symptom Disorder to have special features (e.g., predominantly pain symptoms) or different severity dimensions, depending on the number of A and B criteria present beyond the threshold for diagnosis.

CHILDHOOD DISORDERS. A single spectrum diagnosis, Autism Spectrum Disorder (ASD), replaces "Autistic Disorder," "Asperger Disorder," and "Pervasive Developmental Disorder not otherwise specified." ASD is now differentiated from social communication disorder by the presence of the restrictive and repetitive behavior criterion. Addition of "specifiers" for all neurodevelopmental disorders, "associated with known medical or genetic condition or environmental factors." As stated by Michael Rutter, a leading expert in this area, "it was necessary to abandon the DSM-IV classification because it had been ignored or was sidestepped by clinicians and researchers."

PERSONALITY DISORDERS. These remained unchanged in DSM-5, as the original DSM-IV personality disorders were all retained in section II of the DSM-5. However, an "alternate" category of Personality Disorders has been included in section III. This consists of a trait-based diagnostic approach that separates functional impairment (A criteria) and the expression of characteristic traits (B criteria) for the six personality disorders that were tested in the field trials. A more flexible trait-specified profile can now be used for patient presentations that do not fit well the six subtypes.

SLEEP–WAKE DISORDERS. This group includes a large number of categories and subcategories, starting with Insomnia Disorder, Hypersomnolence Disorder, Narcolepsy (various subtypes), as way as a long list of "Breathing-Related Sleep Disorders" (Obstructive Sleep Apnea Hypopnea, Central

Sleep Apnea, Idiopathic Central Sleep Apnea, Cheyne–Stokes Breathing, and other subtypes) and "Parasomnias" (including nonrapid eye movement sleep arousal disorders, Nightmare Disorder, Rapid Eye Movement Sleep Behavior Disorder, Restless Leg Syndrome, and other subtypes).

OTHER CONDITIONS THAT MAY BE A FOCUS OF CLINICAL ATTENTION. This includes a long list of headings and subheadings, a majority tagged with "V" codes. The main headings are "Relational Problems," "Abuse and Neglect," "Child Physical Abuse," "Child Sexual Abuse," "Child neglect," "Child Psychological Abuse," "Adult Maltreatment and Neglect Problems," "Educational and Occupational Problems" and several others adding to a total of 133 separate codes, making up about one-third of the total categories in DSM-5.

Specific Areas of Concern in DSM-5

PROLIFERATION OF MENTAL DISORDERS. The most common public criticism of later editions of the DSM manual has to do with the growth in both the number of mental disorders and the scope of their definitions. The expressed or tacit implication of this is that it implies a growing and unjustified medicalization of multiple areas of daily life. As mentioned above, the number of disorders included in the first edition of the manual (DSM-I) has been multiplied at least fourfold in the subsequent editions. Reviewing the various editions of the DSM, it is obvious the growth in the number of disorders does not reflect the inclusion of many new disorders, but that it is due mainly to the subdivision or "splitting" of traditional disorders such as anxiety, affective, addictive, and psychotic disorders into countless subtypes (Table 9.1–3). The numbers speak eloquently: DSM-5 lists 438 diagnostic codes, while DSM-IV had 361, and DSM-III 265. However, it has been reported, "there are fewer specific categories in DSM-5 than in DSM-IV."

PERSONALITY DISORDERS. From the perspective of medicine, the concept of "personality" has always been a rather vague sociocultural construct, hard to define even in psychological terms, and thus quite alien to the medical sciences. Indeed, most of the "personality disorders" are value judgments regarding eccentricities, peculiarities, or extremes in the behavior of an individual when he/she is compared with culturally predominant norms. The personality disorders that remain in DSM-IV and DSM-5 (there were no changes in the latest edition) largely represent such "value judgments," often designated using critical, condemnatory, or pejorative language. Since DSM-III, personality disorders have remained virtually untouched except for the elimination of the multiaxial system in DSM-5 and the proposal for an "alternate category" with dimensional rather than categorical overtones, that was relegated to section III of the DSM-5.

SEXUAL DISORDERS. In parallel with the personality disorders, the inclusion of certain "sexual disorders" in a section originally headed by homosexuality has also been a point of contention within the DSM. "Sexual deviations" were included among the personality disorders in the first DSM and remained there in DSM-II. From DSM-III on, sexual disorders were listed under a separate group and have continued to proliferate in each new edition of the manual. Thus, DSM-5 added several new disorders, a number of them among the paraphilias, a particularly problematic set of disorders whose definition may even conflict with the definition of a mental disorder as included in section I of the DSM-5.

THE MEDIA "STORM" SURROUNDING DSM- 5. Neither the task force and workgroups nor the APA in general, did presage the avalanche of criticisms, complaints, passions and lobbying unleashed during the DSM-5 process, starting soon upon release of the initial plans to revamp DSM-IV and continuing until the release of the printed manual. Nothing like this had happened (at least not to this extent) in the case of DSM-III, DSM-IV, or even ICD-10. A notable observation was the broad news media coverage that went far beyond the United States and Canadian boundaries, a clear proof of the national and international interest generated by the DSM classification and psychiatric diagnosis in general.

Global Issues Related to Diagnosis. Although the DSM system is primarily designed for use in North America, it is also used in many other countries of the world, particularly for research purposes. Starting with DSM-5, international input was actively sought, both in the preparatory phase and also as part of the workgroups. Critiques related to DSM psychiatric diagnoses include their "globalization." For example, ADHD, a disorder first recognized in the United States in the 1960s, spread to the United Kingdom, Germany, France, Italy, and Brazil, possibly influenced by the transnational pharmaceutical industry, the role of the Internet, and the activism of advocacy groups. A relevant example is the one provided by Professor Sing Lin from Hong Kong, a member of the somatic disorders work group in DSM-5. It relates to what has happened with the diagnosis of Anorexia Nervosa in Hong Kong. In the 1980s and early 1990s, the disorder was very rarely diagnosed, as patients only had somatic complaints ("bloated stomachs"). In 1994, a teenage anorexic girl collapsed and died in a busy downtown street in Hong Kong. This quickly caught attention of the media, who become aware of the US diagnostic manuals, and reported that anorexia nervosa in Hong Kong was the same disorder as that in the United States and Europe. Obviously, the disorder began to appear with increasing frequency, the major complaints of sufferers shifting from somatic complains to "fat phobia."

Table 9.1–3.
The Increase in the Number of Diagnoses between the 1952 DSM I and DSM-5: The Case of Anxiety

DSM I 1952	Includes anxiety reaction and six related "reactions," dissociative, conversion, phobic, obsessive-compulsive, depressive and psychoneurotic. **(Total = 6 Disorders)**
DSM-II	Eliminates the qualification of "reaction" and includes all of the anxiety-related diagnoses from DSM I, and some related new ones, such as hypochondriacal, neurasthenia or depersonalization, in the general category of "neuroses." **(Total = 8 Disorders)**
DSM-III	The category anxiety disorders appears for the first time, including a total of 10 disorders in four groups: ► Phobic disorders (including agoraphobia with panic attacks, agoraphobia without panic attacks, social phobia, simple phobia) ► Anxiety states (including panic disorder, generalized anxiety disorder, obsessive-compulsive disorder) ► Posttraumatic stress disorder (including acute PTSD, chronic PTSD) ► Atypical anxiety disorder To the above disorders, two categories should be added as they were included in DSM-II among the anxiety neuroses. They are: ► Somatoform disorders (including somatization, conversion, psychogenic pain, hypochondriasis, and atypical) ► Dissociative disorders (including psychogenic amnesia, psychogenic fugue, multiple personality, depersonalization, and atypical dissociative) In addition, several ambiguously defined "adjustment disorders" either expressly or implicitly are merged into the anxiety disorders category. **(Total = 21 Disorders)**
DSM-IV-R	The category anxiety disorders includes, besides generalized anxiety disorder, eight specific anxiety disorders, and two unspecific anxiety disorders; to these, the following must added: five specific and two unspecific somatoform disorders, and four specific and one unspecific dissociative disorders and again, the adjustment disorders **(Total = 23 Disorders)**
DSM-5	Anxiety disorders includes 10 specific and 2 unspecific disorders (the first two were previously accepted only for children): ► Separation anxiety disorder ► Selective mutism ► Specific phobia ► Social anxiety disorder (social phobia) ► Panic disorder ► Panic attack specifier ► Agoraphobia ► Generalized anxiety disorder ► Substance/Medication-induced anxiety disorder ► Anxiety disorder due to another medical condition ► Other specified anxiety disorder ► Unspecified anxiety disorder Obsessive-compulsive gets its own group with seven specific disorders and two unspecific categories. Trauma gets its own group (trauma- and stressor-related disorders) which includes reactive attachment disorder and disinhibited social engagement disorder (previously used only in children), posttraumatic stress disorder, acute stress disorder, adjustment disorders (including five specific and one unspecific disorders), and two unspecific categories.

Somatic symptom and related disorders loses pain disorder and body dysmorphic disorder (which migrates to obsessive compulsive) but still lists five specific disorders and two unspecific categories.
Dissociative includes three specific disorders and two unspecific categories.
(Total = More than 45 disorders)

Dealing with Conflict of Interest. A most salient source of public criticism of DSM-IV and DSM-5, has been the potential "conflict of interest" of those involved in the process of revising and preparing the diagnostic manuals. While having ties with the pharmaceutical industry is the most frequently mentioned conflict of interest, there are several other potential conflicts, such as ties to patient advocacy groups, health insurers, organizational memberships, and other groups that may have a vested interest in the definition of mental disorders. These latter conflicts are rarely addressed. The inclusion or elimination of a mental disorder or a change toward a more restrictive or comprehensive definition of a disorder, all affect the size of the population covered (most commonly, increases it), and with it, the work opportunities for professionals, the potential income for drug companies, and the potential profit for public or private health insurers all grow larger. Also, given the financial benefits accorded by several diagnoses such as the coverage/subsidies for educating autistic children in several states, the worker compensation laws and the veterans administration service-connected disabilities, it is not surprising that the redefinition of disorders such as autism in DSM-5 led to a heated debate not only in mental health journals but also in the press and mass media, with organizations like "Autism Speaks" publishing the names and affiliations of members of the work force, so that they could be lobbied. Another vested interest implicit in the DSM diagnostic process is obviously that of the APA, the owner of the copyrights for DSM. The APA has a large stake in the commercial success of a manual that has gradually become one of its main sources of income. Creating expert committees completely devoid of conflicts of interest seems quite unrealistic and it may be an impossible task. However, at a minimum, a "transparency policy" for full disclosure of all possible conflicts of interest should and has become the standard.

Changes in Medical Practice. Another factor that may play a role in the proliferation of disorders is the increased access of populations to health information in the age of the Internet, resulting in a growing medicalization of daily life. The increased access to treatment occurs all across medicine and the tendency of people to look for treatment for conditions that were previously nondiagnosable or viewed as mild or even trivial, is now a worldwide phenomenon. The thresholds for treatment for elevated cholesterol, blood glucose, or blood pressure seem to be continuously

getting lower (e.g., the uniform 100 threshold rule). A growing number of individuals seek the genetic or pharmacological answers for obesity, a condition that some years ago was not even considered to be a disease. In addition, many individuals currently seek treatment not because of any defect or illness, but for cosmetic or hedonic reasons, not restricted to cosmetic surgery, such as the growing number of males with normal sexual function who ask for prescriptions of phosphodiesterase inhibitors in order to have more frequent sex (the so called "recreational use"), or patients with no attention disorders who ask for prescriptions of stimulants to gain an edge at school or work.

Alternatives to Traditional Classifications. The DSM-5 task force discussed a number of alternatives to the DSM-IV classification. One initial model was that proposed by Dennis Charney, a biologically oriented psychiatrist, then at NIMH, that had suggested a multiaxial "diagnosis for the future," in which Axis I represented Genotype, Axis II, Neurobiological Phenotype, Axis III, Behavioral Phenotype, Axis IV, Environmental Modifiers or Precipitants, and Axis V, Treatment Targets and Therapeutic Response. While this model appeared quite promising and scholarly, the scientific evidence and the needed "markers" were not there yet.

BRAIN CIRCUITRY. Progress in basic neuroscience during the decade of the brain has been unraveling, with good precision, brain circuits related to specific behaviors or emotions. Emerging techniques such as optogenetics, nurture the "circuit dynamics of psychiatric disease" receiving ample coverage in medical journals (JAMA) and the news media (The New Yorker). While most of this basic research is in the preclinical phase (animal models), the circuit-level perspective makes it possible to generate hypotheses on how the effect of multiple genes may lead to specific psychiatric syndromes. Some of the main circuits now under scrutiny include:

1. A "frontothalamic-striatal circuitry," that may be involved in cognitive control of behavior and thought to be of relevance to schizophrenia, ADHD, addictions, and impulse control disorders;
2. A "fear circuitry," with neuronal connections based on the amygdala, involved in conditioned fear, of possible relevance to anxiety and depressive disorders.
3. A "reward circuitry," consisting of meso-telencephalic dopaminergic projections, that may be relevant to addictions, impulse control disorders, schizophrenia as well as the hedonic component of mood disorders.
4. A "regulation of mood states" circuitry, involving the medial prefrontal cortex, amygdala, and hypothalamus, and possibly relevant to depression, mania, and mood disorders in general.

A key question posed during the development of DSM-5 was whether or not at that time, the specialty (psychiatry) and practicing clinicians (psychiatrists, psychologists, social workers, and others) were ready for classifications based on hypothetical pathophysiology, lacking solid, consistent "markers" and the quick answer was "not quite yet."

Table 9.1–4.
Research Domain Criteria (RDoC): Key Components

Negative valence systems:

- "Fear" and its opposite pole, "fearlessness" = amygdala, hippocampus, and ventromedial prefrontal cortex
- "Potential threat" = hypothalamic–pituitary axis, hippocampus, cortisol releasing factor, and cortisol

Arousal/regulatory processes:

- "Stress regulation" = raphe nuclei circuits; serotonin
- "Facilitated stimulus processing" = locus coeruleus circuit; norepinephrine
- "Readiness for stimulus processing and responding" = brain resting state network

Cognitive systems:

- "Working memory": dorsolateral PFC, other areas in PFC
- "Cognitive control" and opposite poles, "impulsivity, externalizing behavior" = anterior cingulate gyrus, various areas of medial and lateral PFC

NIMH RESEARCH DOMAIN CRITERIA. This alternative diagnostic approach, developed by Bruce Cuthbert and Thomas Insel at NIMH, came into play in the NIMH Director's blog, coinciding with the final release of DSM-5. It quickly spread to the news media that "NIMH had disapproved of DSM-5," at the same time that it (NIMH) urged (or mandated) as a condition for future funding, the use of RDoC, their alternative and complex strategy for classification that at the time appeared quite incipient indeed. As described in the NIMH website, RDoC is conceived as a dimensional system (reflecting, brain circuit-level measurements, behavioral activity, etc.) spanning the range from normal to abnormal (Table 9.1–4). As it happens with dimensions in medicine such as hypertension or levels of cholesterol in plasma, this approach has to define cutoff points for the definition and extent of psychopathology (e.g., mild, moderate, and severe). It is of interest to observe, that the NIMH document acknowledges, "to the extent that DSM-5 introduces dimensions in addition to classes, the crosswalks to RDoC dimensions may be enhanced." However, RDoC does not adhere to current DSM categories, as its intent is "to generate classifications stemming from basic behavioral neuroscience." Rather than starting with an illness definition and seeking its neurobiological underpinnings, RDoC begins with brain-behavior-brain relationships and links them to clinical phenomena based on several units of analysis such as neuroimaging, physiological activity, behavior, and self-reports of symptoms. Allegedly, one of the key goals of RDoC, is to allow researchers to choose independent variables from among several different units of analysis. The RDoC

framework is explicitly agnostic with respect to current definitions of disorders. For instance, depression as a clinical syndrome has been related to abnormal activity in the amygdala, anterior cingulate cortex, nucleus accumbens, and multiple monoamine systems, while also strongly comorbid with multiple anxiety disorders, eating disorders, etc. The idea is that studying the individual mechanisms may lead to better understanding of current disorders, or perhaps new and novel definitions of disorders, but in either case, improved information about treatment choices. The RDoC research framework is described as a matrix whose rows correspond to specified dimensions of function, called "constructs," that represent the fundamental unit of analysis in this system. The hope is that these functional behavioral dimensions and related genes and circuits would be continually refined as science advances. The expectation is that most research studies will focus on a single "construct" or compare two separate "constructs" on relevant measures. "Constructs" are grouped into "domains of functioning." These include "negative and positive valence systems," "cognitive systems," "systems for social processes," and "arousal/regulatory systems." The various columns in the matrix represent different classes of variables (or units of analysis) used to study the domains/constructs. These variables include genes, molecules, cells, neural circuits, physiological elements (e.g., cortisol, heart rate, startle reflex), behaviors, and patient self-reports. Neuronal or brain circuits represent the core of these variables. The inclusion of a "construct" in this system depends on the specification that a particular brain circuit or brain area implements that dimension of behavior. It is also acknowledged in the document that given the complexity of brain and behavior, there can be various degrees of ambiguity in this regard. "Constructs" included in the matrix are based on current literature reviews that have provided a neurobehavioral research base for each of the entries. The document also emphasizes that "these particular domains and constructs are simply starting points that are not definitive or set in concrete. It is expected that these will change, dynamically, with input from the field, and as future research is conducted." A major aim of RDoC is to create a framework for grouping participants in research studies, in order to create a foundational research literature that informs future versions of nosologies based upon genetics and behavioral neuroscience. RDoC is not intended for clinical diagnosis at the current time. In the future, research supported by RDoC could inform diagnostic approaches using new laboratory procedures, behavioral assessments, and novel instruments to provide enhanced treatment and prevention interventions. It is also hoped that RDoC will support enhanced development of new pharmacological and psychosocial interventions based upon neurobiological and behavioral mechanisms. The intent of the RDoC is to accelerate the pace of new discoveries by fostering research that translates findings from basic science into new treatments addressing

fundamental mechanisms that cut across current diagnostic categories. In a recent release, the NIMH Director's Thomas Insel argues that the RDoC project is a move forward in "precision medicine." He states that "it is possible that specific mutations in relevant brain circuits can be found that explain some cases of schizophrenia, bipolar disorder, or autism, just as mutations in the tumor, explain cancer, but that precision medicine for mental disorders will not come from a single genomic glitch. Rather, like many other areas of medicine, many genes each contribute only a small amount of vulnerability as part of an overall risk profile that includes life experiences, neurodevelopment, and social and cultural factors. RDoC assumes that many kinds of data are needed to reach precision, more like triangulating to find your position on a map. These data will draw from many sources, including symptoms, genotype, physiology, cognitive assessment, family dynamics, environmental exposures, and cultural background." However, to be clear, this dimensional approach represents a dramatic shift in the classification of psychiatric disorders. It theorizes that behavioral consequences of brain functioning, such as for example, working memory or reactivity of the amygdala, as measured during tests such as functional magnetic resonance imaging (fMRI), would define clinical characteristics that reflect with greater precision the biology of the illness, possibly leading to more specific therapeutic targets compared to the broad overlapping categories of the DSM. Despite the authority backing it, reception of this new system has been mixed, even among leading researchers, many of whom remain skeptical about this approach. According to some of the experts, the existing evidence makes the RDoC approach highly questionable. For example, Charles Nemeroff and Daniel Weinberger in a recent publication argue that "patients with ADHD and patients with psychosis may show similar abnormalities in frontal brain function just as patients with PTSD and patients with psychosis may show similar over activation of the amygdala in fMRI images, but the underlying reasons for these superficial similarities are different and so are the treatments. The treatment of frontal lobe dysfunction in ADHD (e.g., stimulants) will make psychosis worse, just as the treatment for amygdala over activation in psychosis (antidopaminergic drugs) is likely to have adverse effects in patients with PTSD."

DSM- 5 AND RDOC 100 YEARS AFTER JASPERS' GENERAL PSYCHOPATHOLOGY. Wernicke believed that all mental disorders were brain disorders. This belief appears to be the main context for new diagnostic systems such as the RDoC that could be then designated as "Wernickian."

Kraepelin's main premise was that "all psychiatric disorders should follow the medical model of disease." As it has been articulated above in this chapter, starting with DSM-III, American nosology became uniquely "Kraepelinian," with an emphasis on descriptive psychopathology,

following the medical model approach. Jaspers, who published his "Allgemeine Psychopathologie" (General Psychopathology) in 1913, maintained that "psychiatry is a hybrid discipline that must combine the methods of the natural and the social sciences." This "middle ground" approach that incorporates a "social context" had great influence in the United Kingdom but not in the United States. A leading proponent of this approach was Professor John Wing, at Maudsley Hospital in London, following the path of Sir Aubrey Lewis and other pioneers. Wing, developed the PSE, a semistructured instrument rich in clinical phenomenology, that provides detailed symptomatic and syndromic descriptions with elaborate probes, thus requiring significant training and clinical expertise. Previous editions of the PSE were successfully used in two pivotal international studies, "the United States/United Kingdom Diagnostic Project" and the "IPSS" coordinated by the WHO. The US/UK study demonstrated the extent to which a diagnosis of "schizophrenia" in the United States was broader and less definable compared to its use in England, and demonstrated that through training and the use of a structured instrument (the PSE in this instance), wide differences in diagnoses of schizophrenia across the Atlantic virtually disappeared. The IPSS showed that a "concordant" group of patients with schizophrenia could be identified in all countries by trained clinicians using the PSE. Incidentally, the new version of the PSE, called the Structured Clinical Assessment in Neuropsychiatry—Present State Examination (SCAN-PSE) continues to be used by a few teams of investigators around the globe including a few sites in the United States. This approach emphasizes the need to gather rich clinical data (clinical phenomenology) that takes into account social and environmental influences in addition to biology. The assumption is that biological systems interact closely with the environment and that social influences can generate or amplify biological impairments. Unfortunately, this "middle ground" approach to psychiatric diagnosis has been abandoned in modern classifications. Moreover, clinical and epidemiological studies incorporate questionnaires that are often administered by lay interviewers and contain brief or abridged instruments to gauge psychopathology. This leads to rather dry, impoverished, clinical assessments.

DSM- 5 AND CLINICAL PRACTICE. A key objective in publishing DSM-5 was to improve upon previous versions of the manual by revising categories and rigorously assessing (and proving) their reliability. A secondary goal of DSM-5 was to incorporate new strategies (dimensional approach) and provide potential new disorders in need of further evaluation and testing. The idea was to create an open, "living" system that could be easily updated as new evidence accrued. DSM-5 is unlikely to change clinical practice and many practicing psychiatrists largely remain unaffected or indifferent in front of nosological debates as recently highlighted by psychiatrist-

correspondent Sally Satel in a leading editorial. Satel adds, however, that DSM-5 diagnoses will "become passports to insurance coverage, keys to special education services, and tickets to disability benefits." Thus, it seems quite relevant for practicing clinicians to get to know and use DSM-5 criteria properly. Despite the shortcomings detailed above, the DSM-5 manual provides an excellent description for most of the key syndromes encountered in clinical practice.

Regarding the DSM-5 versus RDoC controversy, the NIMH director and the President of the APA produced a joint statement that settles the debate for now. It reads "all medical disciplines advance through research progress in characterizing diseases and disorders. DSM-5 and RDoC represent complementary, not competing, frameworks for this goal" and also highlights that "for practicing clinicians, DSM-5 represents the best available evidence at this time."

Therefore, while psychiatrists wait for future diagnostic developments, the proper, "informed" use of DSM-5 will guide clinical practitioners in assessing understanding and cataloguing complex mental health phenomena. The expectation—assuming the manual is truly a "living document"—is that experiences accruing from its clinical use may lead to the reinforcement, revision, or even the elimination of some diagnostic labels.

The next paradigm change for classification—and it is doubtful RDoC will represent such—has to wait for a clearer understanding of the matrix of brain structures, such as neuronal circuits, the "omics" repertoire (genomics, proteomics, metabolomics), and an understanding of how social experiences affect and influence such systems. A return to true Jasperian phenomenology, with its rich and meaningful analysis of symptom presentations, might be welcomed by many scholars, practicing clinicians, and researchers, as an indispensable step in this process. This may help toward the delineation of various pathophysiological hypotheses that may enhance diagnostic validity by providing etiological knowledge, a promising alternative to the "reliable" but imprecise classifications based on symptoms and subjective reports.

References

American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2014;37(Supp 1):S8–S16.

Chmura Kraemer H, Periyakoil VS, Noda A. Tutorial in biostatistics: Kappa coefficients in medical research. *Stat Med*. 2002;21:2109–2129.

Cosgrove L, Krimsky S. A comparison of DSM-IV and DSM-5 panel members' financial associations with industry: a pernicious problem persists. *PLoS Med*. 2012;9(3):e10001190.

Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Med*. 2013;11:126.

Deisseroth K, Etkin A, Malenka RC. Optogenetics and the circuit dynamics of psychiatric disease. *JAMA*. 2015;313(20):2019–2020.

De Leon J. DSM-5 and RDoC 100 years after Jaspers' General psychopathology. *Am J Psychiatry*. 2014;171(5):492–494.

Feighner JP, Robins E, Guze SB, Woodruff RA Jr, Winokur G, Muñoz R. Diagnostic criteria for use in psychiatric research. *Arch Gen Psychiatry*. 1972;26:57–63.

Foucault M. *Folie et deraison: Historie de la folie*. Paris: Librarie Plon; 1961.

Global Initiative for Chronic Obstructive Pulmonary Disease, GOLD. Updated 2014, www.goldcopd.org/uploads/users/files/GOLD_Report_2014_Jan223.pdf.

Goldberg D. Plato versus Aristotle: categorical and dimensional models for common mental disorders. *Compr Psychiatry*. 2000;41(2 Suppl 1):8–13.

Goldberg D. Should major classifications for mental disorders be revised? *Br J Psychiatry*. 2010;196:255–256.

Goodwin DW, Guze SB. *Psychiatric Diagnosis*. New York: 1996.

Horwitz AV, Wakefield JC. *The Loss of Sadness*. New York: 2007.

Hyman SE. The diagnosis of mental disorders: the problem of reification. *Ann Rev Clin Psychol*. 2010;6:155–179.

Johnson RA, Barrett MS, Sisti A. The ethical boundaries of patient and advocate influence on DSM-5. *Harv Rev Psychiatry*. 2013;21(6):334–344.

Jaspers K. *General Psychopathology*. Vols 1 and 2. In: Hoenig J, Hamilton MW, eds. Baltimore, MD/London: Johns Hopkins University Press; 1997.

Kendler KS. A history of the DSM-5 scientific review committee. *Psychol Med*. 2013;43:1793–1800.

Klerman G. The advantage of DSM-III. *Am J Psychiatry*. 1984;141(4);84.

Kramer HC, Kupfer DJ, Narrow WE, Clarke D, Regier DA. Moving towards DSM-5: the field trials. *Am J Psychiatry*. 2010;167(10):1158–1160.

Lerner A. *Diagnostic Criteria in Neurology*. Totowa, NJ: Humana Press; 2006.

Murphy D. *Psychiatry in the Scientific Image*. Cambridge, MA: MIT Press; 2006.

Nemeroff CB, Weinberger D, Rutter M, et al. DSM-5: a collection of psychiatric views on the changes, controversies and future directions. *BMC Med*. 2013;11:202.

Satel SL. Why the fuss over the DSM-5? *The New York Times*. May 11, 2013.

Wing JK, Sartorius N, Ustun TB. *Diagnosis and Clinical Measurement in Psychiatry, A reference manual for SCAN*. New York/Melbourne: Cambridge University Press; 1998.

▲ 9.2 The Classification of Mental Disorders in the International Classification of Diseases (ICD-11)

JAVIER I. ESCOBAR, M.D., AND CARLOS LóPEZ-JARAMILLO, M.D., PH.D.

Introduction

The International Classification of Diseases (ICD) of the WHO is broadly utilized across the world for clinical, epidemiological, and administrative

purposes in the area of health and disease and in the implementation of health policies. The ICD is used by physicians, nurses, and other health professionals, as well as administrators, scientists, researchers, and health insurance companies, among others. The systematic use of this classification allows the examination of the incidence and prevalence of diseases and other health problems. The proper and systematic use of ICD diagnostic codes allows the creation of reliable registries, which provide a solid base for assessing morbidity and mortality statistics globally. It therefore makes it possible to have a global perspective on health and health-related issues.

To American psychiatrists, ICD is of utmost current importance as coming very soon (October 2015) ICD-10 will become the official system for coding and billing purposes in the United States. Fortunately, the DSM-5 developers took this into account and in the fifth revision published in 2013 each DSM-5 diagnosis has corresponding ICD-9 and ICD-10 Clinical Modification (CM) codes. In the 1970s, the US government decided that ICD-9 codes were not specific enough for clinical use in the United States and developed the CM version. In DSM-5 ICD-9 codes are presented as numerals (e.g., 307.1 for anorexia nervosa) and ICD-10 codes are headed by the letter F. The corresponding ICD-10 codes for anorexia nervosa are F50.01 for the restrictive type and F50.02 for the binge-eating/purging type. The fact that the ICD-10 codes are already in DSM-5 will facilitate the task of shifting from ICD-9 to ICD-10. However, as will be seen ahead, this does not necessarily mean that there is a perfect diagnostic correspondence between DSM-5 and ICD-10.

The WHO is currently working on the 11th version of the ICD classification (ICD-11). The official release of this classification was expected for 2015, but the process has been extended and the new target date is 2017. As the goal is to include as much of the available information on the ICD-11 classification as possible, this chapter will be divided into two major sections: one providing a brief outline of ICD-10, and the other expanding on what is known, thus far, about the classification of mental disorders in ICD-11.

Taking into account that international input on DSM-5 and consultation between APA and WHO have been going on for several years, the expectation is that the final version of ICD-11 will resemble the DSM-5 to a larger degree than the ICD-10 classification resembled the DSM-IV. However, as will be seen ahead, the convergence of the two classification systems is not yet assured and there remain complex, ongoing debates on how to achieve this goal.

Excellent detailed descriptions of the ICD and the “WHO Family of International Classifications” (WHO-FIC) can be found elsewhere. Therefore, this chapter will provide only a general description of the ICD classification system including a brief outline of the other members of the

WHO-FIC. This brief introduction is intended to highlight the most salient features of the ICD classification to serve as background that may help frame the discussions on the ICD system. The main body of the chapter will summarize and discuss the ICD-11 revision process, its likely content, and the controversies surrounding the section on mental disorders. It is hoped that this will provide the readers of this chapter with a current and comprehensive picture of what is happening with the ICD revision process, following the launching of DSM-5, and the difficult task of harmonizing both classifications.

The International Classification of Diseases

The ICD is the evolving core of a set (or a "family" as it has been called) of disease and health-related classifications. It was the product of an international effort to improve the statistics first of mortality and then of morbidity, basically for public health purposes. With time, the actuarial and epidemiological focus of ICD classification was extended, allowing its use as an instrument to help diagnostic and assessment processes in order to improve clinical, research, and administrative tasks.

The ICD within the WHO Family of International Classifications

The WHO-FIC is a set of classifications that are used to describe and produce information about different aspects of health, disease, and health systems around the world. The strong endorsement WHO-FIC has received from the United Nations reflects its alignment with the core values of this international organism. The core function of the WHO-FIC is to aid the production of reliable statistical information and assist statistical systems at local, national, and international levels. These statistics on their turn will be employed in diverse ways by different kinds of stakeholders as input for actions devised to improve the health status and health care of populations around the world. In order to achieve these goals the WHO-FIC uses a standardized language in order to improve communication between groups and individuals on health and health care matters across the world.

The WHO-FIC has been subdivided into reference, derived, and related classifications.

Reference classifications deal with basic parameters of health like death, disease, functioning, disability, health, and health interventions. They can be used at both individual and population levels. The reference classifications of the WHO-FIC include:

- ► The International Classification of Diseases (ICD)
- ► The International Classification of Functioning, Disability and Health (ICF 2) classifies functioning and disability
- ► The International Classification of Health Interventions (ICHI)

Derived classifications, as the name suggests, are developments from reference classifications for more specialized purposes. They are usually created using elements from reference classifications that can be rearranged, but also by adding new items (concepts, descriptions, etc.) to the original structure. Derived classifications in the WHO-FIC include:

- The ICD-10 for Mental and Behavioral Disorders, Diagnostic Criteria for Research
- The ICD-10 for Mental and Behavioral Disorders, Clinical Descriptions and Diagnostic Guidelines
- International Classification of Diseases for Oncology, Third Edition (ICD-O-3)
- Application of the International Classification of Diseases to Neurology (ICD-10-NA)
- Application of the International Classification of Diseases to Dentistry and Stomatology, Third Edition (ICD-DA)

Related classifications contain descriptions of other aspects of the health experience and the health system that are deemed relevant for the purposes of health care. They may come from other sectors and/or working groups of the WHO (e.g., Classification of External Causes of Injury or ICECI, the Anatomical Therapeutic Chemical Classification System with Defined Daily Doses or ATC/DDD). Also they can come from other organizations (e.g., the International Classification of Primary Care, Second edition or ICPC-2, the International Classification for Nursing Practice or ICNP, etc.).

Table 9.2–1 shows the general structure of the types of classifications of WHO Family International Classification.

Development of the International Classification of Diseases

To classify is to cluster a group of objects (in this case disorders), according to logical rules that guide the clustering activity. These rules specify what criteria would determine that a given disorder belongs to a particular category (e.g., the presence or absence of delusions and/or hallucinations; a longer duration vs. a brief duration, etc.). Also, the relationships between the members of a category must be specified. Usually, an important purpose of a classification of diseases is to bring information to the users about the health problems included within its categories. The information conveyed by the classification will then be used according to the needs of the user (e.g., diagnosis, research, policy, etc.). Thus, classifications are crucial to generate certain types of information that will be stored and then used for comparison purposes, exchanges, analyses, disease monitoring, and of course, policy decisions about health, disease, and health care.

Table 9.2–1.
Types of Classifications of WHO Family International Classification

1. Reference Classifications
International Classification of Diseases (ICD)
International Classification of Functioning, Disability and Health (ICF)
International Classification of Health Interventions (ICHI)
2. Derived Classifications
International Classification of Diseases for Oncology, Third Edition (ICD-O-3)
The ICD-10 for Mental and Behavioral Disorders Clinical Descriptions and Diagnostic Guidelines
The ICD-10 for Mental and Behavioral Disorders Diagnostic Criteria for Research
Application of the International Classification of Diseases to Neurology (ICD-10-NA)
Application of the International Classification of Diseases to Dentistry and Stomatology, Third Edition (ICD-DA)
3. Related Classifications
International Classification of Primary Care, Second Edition (ICPC-2)
International Classification of External Causes of Injury (ICECI)
Technical aids for persons with disabilities—Classification and terminology (ISO9999)
The Anatomical Therapeutic Chemical Classification System with Defined Daily Doses (ATC/DDD)
International Classification for Nursing Practice (ICNP)

This practical aim of classifications has led to the statement that classifications are not true or false, but only useful or not useful depending on their purpose. Thus, a classification of diseases can be very useful for diagnosis if it contains diagnostic features. Then it can be used in epidemiological research on a particular health problem, if the diagnosis of the disease (the purpose of the classification) is accurate enough to permit the measuring of prevalence or incidence of the disorder in populations and so on. However, a diagnostic classification can be less helpful or even useless when utilized for other purposes. The preeminence of practical utility over other features of a classification such as truthfulness, ease of use, heuristic value, methodological and/or scientific rigor, social consequences, etc., have been a matter of debate. As mentioned before, one of the most important uses of disease classifications is to improve communication across a heterogeneous group of users like clinicians, researchers, or statisticians, but also between organizations and institutions that manage international or global information on diseases and health-related matters. If communication is understood as being an exchange of information between users, it presupposes a common language. The organized collection of certain terms and meanings and a clear description of the rules for their use, are the functional characteristics of contemporary international classifications. Developers of classifications must make decisions about which terms, concepts, and descriptions are to be included or excluded. They must outweigh the pros and cons of this selection. They have to justify a certain degree of standardization of language (and thought) to improve communication at the price of

suppressing some amount of the variability inherent to human experiences of health and disease. Although some of these decisions are supported by empirical research, often developers and reviewers of nosology have to resort to nonempirical solutions (e.g., practical reasons).

It must be remembered also that users can come from different disciplines, pertain to different organizations and thus have different knowledge backgrounds, goals, and interests. The heterogeneity of users is not a trivial matter. One of the main challenges for a multipurpose classification (i.e., a classification that will be used to provide epidemiological information to public health authorities but at the same time must help clinicians to make accurate diagnoses) is to offer to an ample variety of users the information they need without modifying the basic structure of their practices. Otherwise, a classification should be created and tailored for each group of users, something that is not deemed practically feasible and could lead to communication problems.

A Chronology of ICD Development. It can be said that classifying diseases according to a set of specific principles (e.g., symptoms) has been the rule since the beginnings of medicine. The objects, principles, types, methods, and purposes of classification have been changing according to the historical period and change has been driven by diverse factors such as scientific, philosophical, practical, cultural, political, economic, and legal issues. The mid-19th century, more specifically the year 1853, is generally identified as the one when interest arose to start cataloguing the causes of death that could be used in "all countries" as part of the first International Statistical Congress. Two years later, William Farr presented a five-group classification of diseases: epidemic diseases, constitutional diseases, local diseases (arranged according to their anatomical site), developmental diseases, and diseases that are a direct consequence of violence. Revisions using Farr's principles were made in 1865, 1874, 1880, and 1886. Principles of classification were changed in 1893, when Bertillon submitted to the newly created International Statistical Institute (ISI) a comprehensive classification of causes of death based on a synthesis of English, German, and Swiss classifications. Causes of death were divided into two main groups: those due to generalized diseases and those localized to a particular organ or anatomical site. Bertillon's classification proved successful and was adopted by several countries and the ISI recommended officially that this classification (the International List of Causes of Death [ILCD]) should be adopted by all those offices that were not yet using it. Later in 1929 and 1938, the revisions of the ILCD were made by commissions composed of representatives of both the ISI and the Health Organization of the League of Nations. In 1946, the aim of the international efforts on classification was expanded beyond death causes and the WHO was commissioned to create an International List of Causes of Morbidity. After a meeting in

1948, the work on revision of the classification was from then carried out by the WHO with the support by national committees on vital and health statistics. The constitution of the WHO includes the development of the classification as one of its functions. The seventh revision conference was held in Paris; the eighth, ninth, and tenth revision conferences were held in Geneva in 1965, 1975, and 1989.

It can be stated, unambiguously, that the most distinctive characteristic of the history of the ICD enterprise has been to foster international participation in efforts to create a classification of causes of death and disease that would be both useful and used by countries all around the world.

The ICD-9

The ninth revision (1975) deserves at least a brief mention, because important changes at different levels of the classification were implemented during this revision. This marked the beginning of the ICD family, as several additional documents were presented including "the classification of procedures in medicine" and the "classification of impairments and handicaps." Also the ICD was adapted to the needs of particular medical specialties like oncology.

Recommendations and norms were issued about such matters as lay reporting of disease, mortality coding rules, and the selection of a single cause for death statistics. Also the well-known convention of using a dagger (†) marking the underlying disease and an asterisk (*) to mark its manifestations was issued in this conference. However, the most important single event for classification in psychiatry was the inclusion of a glossary and brief descriptions of the categories included in the fifth chapter (mental disorders), a procedure that was not part of the other chapters. The intent of the glossary and the clinical descriptions was to assist clinicians in making the diagnosis. Thus a new expansion of the purposes of classification was implemented and therefore a new source of difficulties and complexities arose. The classification of mental disorders as conceived and elaborated by the ICD had not only the goal of maintaining its usefulness for public health statistics, but at the same time, it should prove useful for daily clinical work.

The Tenth Revision of the International Classification of Diseases (ICD-10)

Approved in 1992, the ICD-10 is divided into three volumes. In volume 1 "the tabular list," all the medical conditions included in ICD-10 are listed. The listing is organized alphanumerically ranging from A00 to Z99. In this version of the ICD-10, diagnostic codes do not go beyond the fourth character (e.g., F31.0). Each character corresponds to the hierarchical level of the classification. For example, in the code F321, the "F" corresponds to

Mental and Behavioral Disorders; "F3" corresponds to Mood (Affective) Disorders, the "F32" to Depressive Episode, and "F32.1" to Moderate Depressive Episode.

Table 9.2–2 shows the list of chapters of ICD-10 and Table 9.2–3 shows the List of Blocks of Chapter V "Mental and Behavioral Disorders" from ICD-10.

It must be kept in mind that ICD is a "multipurpose" classification, intended to be useful to a heterogeneous group of users. Thus, a balance must be kept between validity, reliability and clinical utility, simplicity, and depth of coverage, to facilitate ease of use. At the same time, cross-national and cross-cultural adaptability must be assured. In order to succeed in this task, a great amount of consultation had to be done not only including researchers or health authorities, but also different groups of health professionals (general practitioners, nurses, and psychiatrists).

The ICD-10 Version for Use in Primary Care. The ICD-10-PHC is a version of the ICD-10, which includes only those mental disorders commonly seen in primary care. The development of this version involved significant communications and exchanges of information between general practitioners and psychiatrists. The ICD-10-PHC includes 28 mental disorders common in primary care (dementia and delirium, drug, alcohol and tobacco use disorders, acute and chronic psychosis, depression, bipolar disorder, generalized anxiety disorder (GAD), panic disorder, phobic disorders, mixed anxiety and depression, adjustment disorder, dissociative disorder, unexplained somatic symptoms, eating disorders, sleep problems, sexual disorders, child psychiatry issues, e.g., mental retardation, hyperkinetic disorder, conduct disorders, enuresis, bereavement). The PHC provided information important for general practitioners such as clinical presentation/descriptions, differential diagnosis, indications for referral, etc. Field trials found that ICD-10-PHC had satisfactory interrater reliability and good concurrent validity. The results of studies about the clinical use and the impact of ICD-10-PHC have been mixed however. While general practitioners considered that it increased their interest on mental disorders and their confidence in making diagnoses, the system appeared to have little or no impact on the overall detection of mental disorders, the accuracy of diagnosis, or the prescription of antidepressants. Interestingly, use of the PHC led to a significant increase in the number of patients diagnosed with depression or unexplained somatic symptoms. General practitioners also made increased use of psychological interventions when using the PHC. The impact of the ICD-10-PHC was particularly seen in developing countries where it was used for the main purpose of training nurses, general health workers, and medical officers.

Table 9.2–2.

List of Chapters of the ICD-10

I	Certain infections and parasitic diseases (A00–B99)
II	Neoplasms (C00–D48)
III	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism (D50–D89)
IV	Endocrine, nutritional and metabolic diseases (E00–E90)
V	Mental and behavioral disorders (F00–F99)
VI	Diseases of the nervous system (G00–G99)
VII	Diseases of the eye and adnexa (H00–H59)
VIII	Diseases of the ear and mastoid process (H60 × H95)
IX	Diseases of the circulatory system (I00–I99)
X	Diseases of the respiratory system (J00–J99)
XI	Diseases of the digestive system (K00–K93)
XII	Diseases of the skin and subcutaneous system (L00–L93)
XIII	Diseases of the musculoskeletal system and connective tissue (M00–M99)
XIV	Diseases of the genitourinary system (N00–N99)
XV	Pregnancy, childbirth, and the puerperium (O00–O99)
XVI	Certain conditions originating in the perinatal period (P00–P96)
XVII	Congenital malformations, deformations, and chromosomal abnormalities (Q00–Q99)
XVIII	Symptoms, signs, and abnormal clinical and laboratory findings, not elsewhere classified (R00–R99)
XIX	Injury, poisoning, and certain other consequences of external causes (S00–T98)
XX	External causes of morbidity and mortality (V01–Y98)
XXI	Factors influencing health status and contact with health services (Z00–Z99)

Table 9.2–3.
List of Blocks of Chapter V: Mental and Behavioral Disorders from ICD-10

F00–F09	Organic, including symptomatic, mental disorders
F10–F19	Mental and behavioral disorders due to psychoactive substance use
F20–F29	Schizophrenia, schizotypal, and delusional disorders
F30–F39	Mood (affective) disorders
F40–F48	Neurotic, stress-related, and somatoform disorders
F50–F59	Behavioral syndromes associated with physiological disturbances and physical factors
F60–F69	Disorders of adult personality and behavior
F70–F79	Mental retardation
F80–F89	Disorders of psychological development
F90–	Behavioral and emotional disorders with onset usually occurring in childhood and

F98	adolescence
F99–F99	Unspecified mental disorder

The Research Version of ICD-10. Consistent with the idea of a multipurpose classification for multiple kinds of users, the ICD-FIC includes a version for research called "Diagnostic Criteria for Research" (ICD-10 DCR). The basic difference with the other versions is the less descriptive but more precise wording, more restrictive definitions, description of inclusion and exclusion criteria, and definition of fixed symptom thresholds. This of course has the main purpose of selecting more homogeneous populations that may share common underlying mechanisms thus facilitating clinical research. Homogeneity in selection of patients for clinical trials was also thought to be helpful for identifying treatment response or nonresponse patterns. The draft of ICD-10 DCR was produced in 1990 and the field trials started the same year. Interrater agreement, confidence in diagnosis, ease of use, concordance with ICD-10 CDDG, DSM-III-R, and DSM-IV were all empirically assessed. Results of the field trials showed a good interrater reliability for most categories and also users reported satisfaction and ease of use. Field trials in the United States and Canada had similar results to those of other countries. ICD-10-RDC has been used in different epidemiological studies, although it has been recognized that its use for research is less than that of the DSM system.

Comparing ICD-10 and DSM-IV

The existence of two main classification systems with widespread international use requires an understanding of differences and similarities. This is particularly important nowadays, as ICD-10 is about to become the official system in the United States for recording and billing purposes. Key issues to be examined include how mental disorders are portrayed in each system and the ease of use of each of these classifications for busy clinicians around the world. Comparisons between the two systems have proven difficult, as they represent different conceptualizations of purposes and methods for classifying mental disorders. Often, it is forgotten that the specific aims that the developers of a particular classification system had in mind, will heavily influence the methods for its development and the final written document. Therefore it is not surprising that ICD and DSM classifications have several differences, some more substantial than others. These differences do not imply that one of the classifications is better than the other. The judgment of what is better is essentially a value one, and henceforth relative. For example, a classification X can be considered better than classification Y if clinical utility is the point of reference for judgment. However, classification Y could be considered better than X if it is decided

that research utility is the main criterion for judging superiority.

For the purposes of this chapter, "comparison" means simply to review and summarize differences and similarities in order to understand better the characteristics of both classification systems for mental disorders. The main differences across instruments will be divided below into general and specific.

General Differences between ICD-10 and DSM-IV

The most important differences between Chapter V of ICD-10 and DSM-IV stem from the goals of the umbrella organizations that created them. Thus, as an international tool, ICD-10's key aim is offering clinically useful descriptions for users pertaining to different national and cultural backgrounds across the world, without negatively affecting other important features like reliability, validity, and clinical utility. As a national classification, DSM is built to make diagnostic recommendations of interest primarily to members of a national professional organization. The ICD is a system for illness reporting, whose use is mandatory in the countries that agreed to use it. National classifications (DSM) do not require governmental approval, but are under the authority of the professional associations (in this case, the APA) that regulate their construction, revision, and release.

ICD has to undergo a series of steps prior to submission and approval. Approval comes first from reference centers and then it is submitted to the health ministers of the countries members of the WHO that will discuss its acceptance. This international process involves a great degree of consultation and negotiation, where the views of different countries (and their health traditions) have to be understood, respected, and taken into account. The international character and orientation of ICD is a different process than the international dissemination and acceptance of the DSM system. The former (ICD) has to do with a conceptual and guiding principle for its development while the later (DSM) has to do with other factors, like the widespread use of the classification for research. Having different versions for different needs and users, means that ICD has to offer clinical descriptions that are useful and practical for users from different professional backgrounds (not only psychiatrists or physicians), without negatively affecting other important features like its international character and reliability. Consequently, different versions of ICD have been made available for psychiatrists, primary care practitioners, and researchers. More recently, the DSM system also developed a version for use in primary care.

National classifications like the DSM tend to be "prescriptive" in the sense they regulate what terms can or cannot be used to report diagnoses.

The international classifications take a different approach, and rather than "prescribing," they tend to provide the practical categories that health

professionals may need for their daily clinical work. A frequent example has to do with the presence in a patient of both anxiety and depressive symptoms without a clear predominance of one of them. For this, ICD would provide a category called "mixed anxiety and depression" mostly because clinicians feel they need it. However, the DSM has not accepted such category and therefore there is no coding for it, although it is acknowledged that it deserves further study. This difference between ICD and DSM has been framed as the "bottom-up" versus "top-to-bottom" approach to classification.

The use of clinical descriptors instead of diagnostic ("operational") criteria is undoubtedly one of the most important differences between ICD and DSM, reflecting concepts and values underlying the ICD enterprise. ICD developers decided to include brief clinical descriptions—a narrative account of the main features of the disorders—to help diagnosticians. DSM developers took the approach of articulating a list of criteria accompanied by instructions about the number and the combinations of criteria needed to make a diagnosis. Both approaches have strengths and limitations, mostly related to the specific purposes for using the classification (e.g., clinical work vs. research).

Table 9.2–4 shows a summary of the differences between DSM and ICD.

Specific Differences between ICD and DSM

Empirical Comparisons. Although there is a general impression that ICD-10 and DSM-IV are more similar than different, empirical data suggests that concordance between both is not as good as expected. In an important comparison published in 2009, Michael First, found that of the 176 criteria sets he compared, only one disorder (transient tic disorder) had identical DSM-IV and ICD-10 definitions. Twenty-two percent of the nonidentical sets had definitional differences that were conceptually based and the other 78 percent had nonconceptually based differences. Several other studies have also reported significant differences between DSM-IV and ICD-10.

Table 9.2–4.
Summary of Differences between DSM and ICD

DSM	ICD
National classification of the United States	Official World classification
Focus on psychiatrists and psychologists	Designed to be used by and useful for the different health professionals
Developed to fulfill the particular information needs of the US health system	Created with the idea of being useful for different health systems around the world

Operational criteria	Clinical descriptions
Property of the APA	Free of charge and open access
US, Anglophone perspective	Multilingual and multicultural
Advantages for research	Advantages for clinical use
Different formats, one version	Multiple versions and formats for multiple users

Some of specific examples are presented below:

DEMENTIA. ICD-10: The range of deficits that are required in addition to memory impairment is narrower, that is, deterioration in thinking, judgment, and general information processing (thought to be equivalent to executive function). Aphasia, apraxia, and/or agnosia are considered as supporting the diagnosis of dementia, but not at the same level than executive function deficits. Criteria are provided for rating the severity of the dementia as mild, moderate, or severe. Emotional disturbances are included in the definition of dementia. ICD-10 requires at least 6-month duration for a "confident clinical diagnosis."

DSM-IV: In DSM-IV, the range of deficits in addition to memory is broader and aphasia, apraxia, and/or agnosia are included. There are no criteria for rating the severity of dementia. Emotional disturbances are not included as part of the definition although they are cited as "associated features." Moreover, the DSM-IV definition does not include duration of symptoms.

HARMFUL SUBSTANCE USE/ABUSE. ICD-10: Requires evidence that substance use is causing physical or psychological problems.

DSM-IV: Emphasizes recurrent use, role failure, use in spite of risks for physical health, recurrent substance-related legal problems and use is persistent even if it is causing social and/or interpersonal problems.

SCHIZOPHRENIA. Duration required for diagnosis is significantly longer in DSM-IV (6 months minimum) compared to ICD-10 (1 month).

ACUTE AND TRANSIENT PSYCHOTIC DISORDERS. ICD-10: Four disorders are included in this category. The duration of symptoms is up to 3 months and significant importance is attached to "symptomatic polymorphism."

DSM-IV: Only one disorder is included in the category, regardless of type and intraepisode pattern. Duration has to be more than 1 day but less than 1 month.

SCHIZOAFFECTIVE DISORDER. ICD-10: An "approximate balance" is sought between schizophrenic and affective symptoms (both in terms of severity and duration).

DSM-IV: Criterion A, delusions or hallucinations, has to occur for at least 2 weeks in the absence of prominent mood symptoms and criterion B, mood symptoms must be present during a "substantial" part of the episode.

HYPOMANIA. Unlike ICD-10, DSM-IV does not consider a single hypomanic episode to qualify as a mental disorder.

MANIA WITH PSYCHOTIC SYMPTOMS. ICD-10 excludes delusions or hallucinations noncongruent with mood and/or first rank symptoms.

DSM-IV allows delusions or hallucinations of any kind.

BIPOLAR AFFECTIVE DISORDER. ICD-10 does not make the bipolar I versus bipolar II distinction.

SEVERE DEPRESSIVE EPISODE WITH PSYCHOTIC SYMPTOMS. ICD-10 excludes psychotic phenomena that are not congruent with mood state.

AGORAPHOBIA. ICD-10 defines it according to specific situations that are feared or avoided (phobic content and context).

DSM-IV conceptualizes it as secondary to panic attacks or panic-like symptoms.

OBSESSIVE-COMPULSIVE DISORDER. ICD-10: Outlines criteria to distinguish between obsessions and compulsions depending on the form, that is, obsessions as thoughts, ideas, or images and compulsions as acts.

DSM-IV: Emphasis is placed on the relationship of the symptoms with distress (obsessions) or its prevention/reduction (compulsions).

ACUTE STRESS REACTION. ICD-10: Considers a wide range of symptoms including dissociation after exposure to the traumatic event.

DSM-IV: Assumes that a dissociative response is necessary.

BULIMIA NERVOSA. ICD-10 requires "intrusive dread of fatness."

DSM-IV: Considers fear of fatness as a feature only of anorexia nervosa.

The above examples show that these differences are important enough to influence clinical diagnosis and research because they have an impact on the criteria for "caseness," that is the threshold that separates a "case" from a "noncase" of a given disorder.

Overall Frequency of Use and Satisfaction with ICD-10

If one accepts that both classifications have their strengths and limitations, that some categories may be better portrayed in DSM and some in ICD, and that both systems have been created with different aims in mind, acceptability and satisfaction of users is an important additional factor to take into account when trying to compare these two diagnostic systems. In general, users consider ICD-10 clinically useful and many view it as more appealing than DSM-IV, even among US psychiatrists. In the United States, DSM is most frequently used by psychiatrists, but less so by psychologists, a trend that might be changing. It is also probable that clinicians outside the United States have a preference for ICD-10, but this can vary from

country to country. Surveys of international clinicians indicate that they use DSM more than ICD, possibly because DSM is what they were primarily taught during their psychiatric training. Moreover, researchers and other international scholars have also adopted DSM as their preferred diagnostic system. Researchers appear to feel more comfortable with DSM because of the "operational" criteria, which allows a better "homogenization" of the populations under study.

In international surveys, clinicians value clinical usefulness more than research utility and enhancing of clinical communication more than other characteristics and goals of classifications such as statistics and prognosis. Psychiatrists prefer classifications with fewer categories. In the case of other cultures, Asian psychiatrists found that cultural application of the two systems was difficult, and there were differences between China and Japan (using ICD-10) compared to Korea and Taiwan (using routinely DSM-IV). Another international survey showed that although ICD-10 is widely used, many of its diagnostic categories are underused by psychiatrists. In a survey of practicing psychologists, clinical utility and facilitating communication were perceived as the most important purposes of classification. Flexibility in criteria was considered an advantage. Concerns about problematic diagnosis and cultural bias were also expressed, especially outside USA and Europe. A key WPA-WHO-sponsored survey on ICD-10 was conducted in 2011. In this survey, ICD-10 was found to be widely used throughout the world, when compared with previous surveys using smaller and selected samples. As with other surveys mentioned before, clinical utility and communication were rated as more important than other purposes of classifications. Also consistent with other studies, psychiatrists seem to use only a few categories. Research and primary care versions of ICD-10 have received less attention and are not well known by potential users, a situation that has been attributed to a failure in dissemination by the WHO. In fact, in an international survey done among general practitioners, ICD-10 and DSM-IV were found to be "complicated" and this seemed to confirm that nonpsychiatrist users would prefer diagnostic systems adapted to their individual needs. It seems highly desirable that psychiatrists and psychologists should be familiar with both classifications. Finally, empirical research on the area of satisfaction/acceptability and perceived utility, among different groups of users should receive more attention once a classification has been launched for public use.

Mental Disorders in ICD-11

As has already been mentioned, ICD-11 is currently under development and is not expected to be released until 2017. Because of this, most of what will be examined here has to be considered tentative and provisional. This section will focus specifically on Chapter 7 of the ICD-11 where mental

disorders will be classified.

Three main topics are reviewed herein:

a. A general overview of mental disorders in ICD-11.
b. Proposals that have been made for changes in specific mental disorders.
c. General problems that this process will have to face and brief analysis of the harmonization between ICD-11 and DSM-5.

Previously it had been outlined the key features of an international classification that need to be taken into account and should become "conceptual signposts" for monitoring the development of ICD-11.

- ► It originates and expands from a set of basic points and agreements between mental health professionals.
- ► It requires simplicity and understandability to allow its easy use by those dealing with common mental disorders.
- ► It should be attractive to users as a tool of information exchange.
- ► It must be easily translatable into any already established national classifications.
- ► It should be conservative and theoretically familiar in order to be acceptable to diverse users with different levels of knowledge, disciplinary backgrounds, and interests.
- ► Changes should be implemented only when there is sufficient knowledge available to justify innovations.
- ► Because of its international character, the languages in which it will be translated must be kept in mind.
- ► Successive revisions must retain continuity with the previous versions.

Why Change ICD-10?

None of the extant classifications are considered good enough, either from the taxonomic point of view or from the perspective of yielding reliable, practical diagnoses for clinical work and/or research. Some of these difficulties have been linked to irregularities in terms, concepts, and descriptions leading to limitations in the diagnostic information contained in the descriptions of disorders. The environmental components and the context in which classification is used have undergone profound changes with the passage of time. Therefore, major reasons for reviewing ICD-10 include efforts to improve the consistency and uniformity of the diagnostic information provided by the clinical descriptions and to adapt to new demands from users and conditions of use. It is important in this context to generate new research data about mental disorders that is backed up by scientific developments (evidence-based) and that satisfies new demands from clinical users about utility and ease of use in various political, social, and economic environments. It must also be user-friendly, adapt to new technological environments and be amenable to frequent updates.

Improve the Classification without Losing What Has Been Gained. A revision process needs to keep a balance between conservatism and innovation. In other words, it should improve the classification as much as possible without losing what is already good in it. The process of revision of ICD-10 has been compared to the release of new "software" that besides offering new features, they have to be compatible with previous versions. Similarly to what was the goal with DSM-5, new technological advances should help keep ICD-11 as a "living document," experiencing transformations when new significant information coming from different environments and stakeholders arrives.

Maintain the Basic Purposes While Increasing the Diversity of Stakeholders. The three basic purposes of the classification (clinical, public health, and research use) have to be maintained. The main stakeholders of ICD-11 include the governments of the WHO member countries, health care professionals (not only mental health professionals or psychiatrists), and researchers. However, patients and "users" of mental health services have been also included among the stakeholders, although the mechanisms for their participation have not been yet clearly stated. Other important stakeholders are legal systems, health policy makers, and payers of health care services.

Clinical Utility: The Core Concept Guiding the 11th Revision of ICD. It was stated that clinical utility backed up by research evidence should be the guiding principle. Michael First and collaborators note that "clinical utility deserved to be an organizing priority for the revision so long as it did not sacrifice validity as established by the best available science." However due to the scarcity of data coming from "best available science" that can be used to inform classification, it seems very possible that utility, reliability, and some equivalents to validity will be the core guiding principles of ICD-11.

The development and refinement of classifications in psychiatry is a reminder that the discourse on science, scientific rigor, objectivity, etc., must be taken with caution, because very often the practical needs and social pressures are the drivers of diagnostic decisions. This is not a negative issue per se. What is undesirable is the "reification effect" of classifications, despite the absence of reliability and validity.

The Definition of Mental Disorder Will Not Change in ICD-11

Major changes in the definition of mental disorder are not expected in ICD-11 and therefore, the ICD-10 definition will be maintained: "A clinically recognizable set of symptoms or behaviors associated in most cases with distress and with interference with personal functions."

ICD-11's Prototypes, Criteria, and Clinical Descrip tions. The question

on which would be the best approach to portray symptoms and other clinical features of mental disorders in such a way that they are easy to identify and can be readily used by clinicians, has been carefully considered during the process of ICD-10 revision. As part of the ICD-11 approach, the clinician would decide whether or not the diagnostic features outlined in the clinical description match those of the "real" patient. The ICD-11 approach conveys a considerable amount of clinically relevant information about the disorder, including both essential and typical features, differential diagnosis, the boundary with normality, typical course, and developmental and cultural-specific features. A procedure called "prototype matching" consists of clinicians estimating to what extent the patient's clinical presentation matches a paragraph-length description of the disorder. The clinician would then have to estimate the extent of matching using a 5-point scale, from 1 ("little or no match") to 5 ("very good match"). The expected result would be both a dimensional diagnosis (capturing the extent to which the patient has the disorder) and a categorical one, because higher ratings (4 and 5) correspond to the presence of the disorder. A rating of 3 would indicate "sub threshold" or "clinically significant features." However, the international debate on this approach has outlined serious or even unsolvable shortcomings. Despite the limitations, the different critics of the prototype matching procedure considered it is a potentially important development and there is consensus to keep refining this approach.

Dimensions and Categories in ICD-11. The use of the term "dimensional" has become fashionable in psychiatry in recent times as DSM-5 included this at least in the initial debates and provided some potential dimensions in section III of the manual. Dimensional approaches are viewed as novel, as they may free psychiatric classifications from the conceptual, empirical, and practical restraints of operational categories. While the emphasis will be on categories, ICD-11 may also consider dimensions at least in a few situations, where simple measurements may be feasible.

"Specifiers" and "Qualifiers" in ICD-11. ICD will not use coded "specifiers" because of the limits of the coding system. Therefore, if some features of a disorder are deemed important to be highlighted, noncoded "qualifiers" will be added (see for example schizophrenia below).

The Process of Revision of the Mental Disorders Chapter of ICD-11

The ICD-11 chapter on mental disorders is Chapter 7, thus being located differently than it was in ICD-10. The following categories have been published in the Beta Draft and their codes go from 7A00 to 7A43.2. The proposed blocks are the following: neurodevelopmental disorders, schizophrenia and other primary psychotic disorders, catatonia, mood

disorders, anxiety and fear-related disorders, obsessive-compulsive and related disorders, Disorders specifically associated with stress, dissociative disorders, bodily distress disorder, feeding and eating disorders, elimination disorders, disorders due to substance use, impulse control disorders, disruptive behavior and dissocial disorders, personality disorders, paraphilic disorders, factitious disorders, neurocognitive disorders, psychological and behavioral factors affecting disorders or diseases classified elsewhere, mental and behavioral disorders associated with disorders or diseases classified elsewhere, sleep–wake disorders and symptoms, findings and clinical forms of mental and behavioral disorders.

Table 9.2–5 shows the proposed blocks of Chapter 7 "Mental and behavioral disorders" of ICD-11.

Consultation (with individuals, institutions, and governments) and a thorough review of the literature relevant to classification of mental disorders were initially implemented as part of this revision. The technical preparation and final products have been assigned to a WHO special unit. This complex task has turned into a long and tortuous process marked by delays and postponement of deadlines.

Table 9.2–5.
List of Blocks of Chapter 7 Mental and Behavioral Disorders of ICD-11

Neurodevelopmental disorders
Schizophrenia and other primary psychotic disorders
Mood disorders
Anxiety and fear-related disorders
Obsessive-compulsive and related disorders
Disorders specifically associated with stress
Dissociative disorders
Bodily distress disorder
Feeding and eating disorders
Elimination disorders
Disorders due to substance use
Impulse control disorders
Disruptive behavior and dissocial disorders
Personality disorders
Paraphilic disorders
Factitious disorders
Neurocognitive disorders
Mental and behavioral disorders associated with disorders or diseases classified elsewhere

WHO's Topic Advisory Group (TAG-MND) was created in 2007 in the production of the proposals for the classification of mental disorders within ICD-11. The TAG-MND created several subgroups: one epidemiological and one in charge of harmonization with DSM-5 (that was replaced by a

different mechanism). New special groups were also created for mental disorders seen in primary care, health problems related to substance misuse, childhood mental disorder, intellectual disability, personality disorders.

The WHO also created a coordinating group to establish the Global Scientific Partnership Network (GSPN), involving experts representing different disciplines and traditions in psychiatry. Its role is to help in the gathering and analysis of research data relevant to classification, the design of field trials, and the assessment of the translation process among its principal tasks.

General Structure of the Categories and Clinical Descriptions

In order to improve uniformity in the ICD-11 descriptions, a template had to be used by the working groups, known as the "content form." This template will standardize the general structure of information included. Variability of content and format in ICD-10 was thought to reflect that different groups of experts worked on the various sections, but without standardization in the development of the different sections. Content form includes items arranged in a special way and guided by rules of application that yield technical, clinical, and administrative information. Also a standard format for clinical descriptions and diagnostic guidelines was prepared. Importantly, diverse authors have insisted that diagnostic guidelines should be more realistic from a cognitive perspective, that is, using a structure that represents the way clinicians think and use clinical judgment.

Table 9.2–6 illustrates the standard format for ICD-11 Clinical Descriptions and Diagnostic Guidelines.

ICD-11 Field Trials

A series of field trials both clinic-based and internet-based are being conducted, focusing on clinical utility and reliability. The web-based field studies use standardized case materials in order to test if the clinical guidelines for diagnosis of ICD-11 produce more consistent diagnostic behavior than ICD-10. The methodological rigor and predictive power of the vignette's methodology cannot be underestimated. Vignettes are designed to produce a behavioral response (diagnosis) by clinicians, which in turn may help predict how useful the classification can be. The sample of assessors is multidisciplinary and includes around 12,000 mental health professionals from more than 130 countries. Clinic-based field studies will involve pairs of clinicians applying the diagnostic guidelines to patients. It is expected that important data on interrater reliability, usefulness of the guidelines, and other relevant questions (e.g., issues on comorbidity or reduction of unspecified diagnosis) will be gathered with clinic-based field

studies.

Table 9.2–6.
Standard Format for ICD-11 Clinical Descriptions and Diagnostic Guidelines

Category Name
Brief Definition: containing a summary statement (100–125 words) of the common essential features of the disorder
Inclusion Terms
Exclusion Terms
Essential (Required) Features: relatively explicit guidance regarding features needed to confidently make the diagnosis
Boundary with Normality (Threshold): specifying those aspects of the disorder that are indicative of its pathological character
Boundary with Other Disorders (Differential Diagnosis): this section provides guidance to the clinician about how to make this differentiation
Coded Qualifiers/Subtypes
Course Features: information on temporal patterns like age of onset, persistence, duration, likely progression overtime, etc.
Associated Clinical Presentations: information about features that are important and help to recognize variations in presentation but also highlights areas where clinical intervention might be important
Culture-Related Feature: brief information regarding cultural considerations for diagnosis
Developmental Presentations
Gender-Related Features

The ICD-11-Primary Health Care (ICD-11-PHC)

Similar to what ICD-10 did, the 11th edition of ICD will also have a revision of the classification for use in primary care settings, the ICD-11-PHC. The working group has anticipated that although many of the disorders included in the ICD-10-PHC will be retained, some "radical" proposals are being made that in case they survive the field tests, they would be included in the new version. Some of these proposals include:

- To abandon the distinction between anxiety and mood disorders and create a category called "dysphoric disorders." This will allow clinicians to diagnose depression, anxiety, or the combination of both as "anxious depression."
- Anxiety disorder would be largely equivalent to generalized anxiety disorder, but without the longitudinal criteria duration, to facilitate detection and intervention by general health professionals.
- The new category of "bodily distress disorders" will include bodily distress syndrome (formerly unexplained somatic complaints), health preoccupation, and conversion disorder.

The disorders tested in field trials include:

Childhood disorders (intellectual development disorder, autism spectrum disorder, specific learning disability, attention-deficit hyperactivity disorder (ADHD), conduct disorder, enuresis, encopresis)

Psychotic disorders (acute psychosis, chronic psychosis, bipolar disorder)

Dysphoric disorders (anxious depression, depressive disorder, anxiety disorder, distress disorder, posttraumatic stress disorder (PTSD), panic/agoraphobia)

Body distress disorders (bodily distress syndrome, health preoccupation, conversion disorder)

Bodily function disorders (sexual function disorder—male, sexual function disorder—female, sleep disorder, eating disorder)

Substance use disorders (alcohol use disorders, drug use disorders, tobacco use disorders)

Personality disorder (borderline personality)

Acquired neurocognitive disorders (dementia, delirium)

Proposals for Specific Disorders in ICD-11

Because ICD-11 is not yet "officially" available, this section is inevitably a tentative and speculative perspective on what could happen with this revision, and it is based on a selection of literature on the topic. The main purpose is to give a broad idea about what to expect with respect to some disorders in ICD-11.

Psychotic Disorders. The name of the category will be changed from "Schizophrenia, schizotypal and delusional disorders" to "Schizophrenia spectrum and other primary psychotic disorders." This category will include seven disorders: schizophrenia, schizoaffective disorder, acute and transient psychotic disorders (ATPDs), schizotypal disorder, delusional disorder, other primary psychotic disorders, and unspecified primary psychotic disorders. Symptomatic profile and duration will be the main basis for categorization. Schizophrenia, schizoaffective disorder, and ATPD seem to be the categories that will undergo more significant changes.

SCHIZOPHRENIA. The specificity and diagnostic value of "first rank symptoms" has been challenged and it is likely that they will be "de-emphasized" in ICD-11. The diagnosis of schizophrenia will consist of two out of eight symptoms, one of which must be "core symptoms" (delusions, thought insertion/withdrawal, hallucinations, and thought disorder). The duration requirement will be the same as in ICD-10 (i.e., 4 weeks). If a patient presents with the required symptomatic profile but does not fulfill the duration criteria the disorder would be classified as "unspecified primary psychotic disorder." As in DSM-5, traditional clinical subtypes

(paranoid, catatonic, and so on) will be dropped, because of their lack of longitudinal stability and prognostic value. Also research shows that these subtypes are seldom use in clinical practice.

In general, what the reviewed literature suggests is that in ICD-11 schizophrenia will be quite similar and compatible with the DSM-5 category except for the issue of "functional impairment." The ICD's general position of not including functional impairment as part of the definition of a disease or a disorder will be respected in ICD-11. Consequently, it will be not part of the diagnostic guidelines for schizophrenia, although for ICD, "functional impairment" is still an important clinical "qualifier." For DSM-5 functional impairment is a criterion of clinical significance and thus necessary to diagnose schizophrenia.

Qualifiers, which are in essence equivalent to DSM-5 "specifiers," are an important resource to better define the clinical situation of patients with diagnosis of schizophrenia or schizoaffective disorder, especially for treatment formulation, planning, and implementation. It seems probable that if they prove useful in field trials this approach will be used for other disorders as well. There will be qualifiers for symptoms, course, cognitive function, and functional impairment.

- Symptom qualifiers: including the presence of positive, negative, depressive, manic, and psychomotor symptoms.
- Course qualifiers: first versus multiple episodes, acute versus chronic, full versus partial remission, and acute versus insidious onset.
- Cognitive qualifier: it is hoped that this would increase attention to cognitive impairment for therapeutic and rehabilitation purposes.

A functional impairment qualifier that should be consistent with the WHO's International Classification of Functioning, Disability and Health (ICF).

SCHIZOAFFECTIVE DISORDER. Schizoaffective disorder continues to pose difficulties because it is not clear how to improve validity and reliability of this problematic construct for clinical and/or research purposes. What has been published so far suggests that in ICD-11, this diagnosis will be made only when a patient fulfills criteria for both schizophrenia and a mood disorder of moderate/severe degree. Total duration of both components should be of 4 weeks. Apparently the issue about the need and/or utility of longitudinal criteria for diagnosis has not yet been resolved.

ACUTE AND TRANSIENT PSYCHOTIC DISORDER. The European perspective on ATPD is different from that of DSM-5, and the former perspective has always been the main influence for ICD classification of these disorders. In DSM-5, these disorders tend to be viewed more pragmatically, as being clinically similar to schizophrenia except for duration. The European

tradition sees in these disorders a heterogeneous group of psychoses different from schizophrenia, both at the level of the clinical presentation and the pathophysiology. This is very likely to be reflected in ICD-11. The main clinical features will remain including sudden onset, brief duration and most importantly, a polymorphic clinical presentation. Focusing on simplicity and ease of use, it is likely that the "acute polymorphic psychotic disorder without symptoms of schizophrenia" (F23.0) in ICD-10 will be the clinical template for ATPD in ICD-11. Duration up to 3 months will be allowed, as 2 to 4 months is the modal duration for ATPD and other psychoses of sudden onset. When an "acute polymorphic disorder with features of schizophrenia," lasts less than 4 weeks, the diagnosis will be of "unspecified primary psychotic disorder." If duration is more than 4 weeks, then a diagnosis of schizophrenia must be made. "Schizophreniform disorder" will not be introduced as a category in ICD-11.

Mood Disorders. In the last decades, the field of mood disorders has been the scenario of changes in conceptualization and there are many discussions related to their classification and diagnosis in ICD-11. How to improve the detection of bipolar disorders, specifically, how to diminish delayed diagnosis, underdiagnosis and overdiagnosis has been at the center of active work in this field. Also, there have been attempts to approach the difficult question on how to detect bipolar disorder in front of a first "unipolar" depressive episode. Hypomania has also been at the center of discussions, due to the inherent problems to decide on its presence or absence and the difficulties for retrospective detection. The bipolar distinction between types I and II (and other possible subtypes) also is part of ongoing debates. Finally, there is the issue of what to do (nosologically speaking) with clinical presentations that include both manic and depressive features (formerly called mixed states in DSM) that were dropped in DSM-5 and conceptualized as a "specifier." However, the presence of symptomatic mixtures in unipolar depression has created a problem, as it permits the "intrusion" of a bipolar presentation as part of a unipolar mood disorder. Passionate debates on this topic are fueled by concerns that the use of antidepressants in patients with bipolar disorder could be worsening the clinical course of these patients. There are several articles and editorials devoted to what "ICD-11 should do" with different mood disorders. Some of the key ideas extracted from this literature include:

- Although the intention is to achieve a higher rapprochement between ICD-11 and DSM-5, there are several important points in which divergence can be expected.
- The general structure of the category appearing in the beta version of the ICD-11 offers a single diagnostic category, "mood disorders," subdivided

into bipolar disorders and depressive disorders.

- ► The adoption of dimensional measurements and/or modeling of mood disorders seem to be out of consideration for the future ICD-11.
- ► Insofar as the decisions that will be made for the ICD-11, some authors involved in the process of review doubt the scientific validity of the concept of "spectrum" and reject including it in the classification.
- ► There are proposals directed toward changes on subjects like melancholia, psychotic and catatonic depression, cyclothymia, dysthymia, rapid cycling, etc. However, at present is difficult to know if these proposals will be incorporated into the written document.
- ► The controversial category mixed depression and anxiety disorder will be retained, because of its clinical utility as it has been shown in surveys and field trials.
- ► Definite positions of ICD-10 reviewers include subjects like bereavement and the diagnosis of depression, the nosological status of mixed states, and hypomanic symptoms in major depression.

THE ISSUE OF BEREAVEMENT AS EXCLUSION CRITERIA. An important debate has arisen after the removal of bereavement as an exclusion criterion for diagnosing a major depressive episode in DSM-5. The position of the ICD-11 developers is that bereavement should continue to be an exclusion criterion for the diagnosis of depression. For example, a key leader of WPA has stated that this change in DSM-5 "is likely to be perceived by the general public as a further step in psychiatry's attempt to pathologize normal human processes." In several articles on the subject, some of the authors involved in the ICD-10 revision process declared that "this could be one of the points of disagreement that make it difficult to harmonize the two classifications" and have argued that there is no scientific evidence at present to justify the removal of grief.

COURSE SPECIFIERS AND MIXED STATES. One of the most-commented novelties in the mood disorders category in DSM-5 refers to the disappearance of the category of mixed episode in bipolar disorders and its new conception as a specifier. The ICD-10 workgroup members question the diagnostic usefulness of this change and are considering keeping the criteria of mixed episode unchanged. There seems to be little support to the possibility of including depression with hypomanic symptoms as part of depressive disorders in ICD-11.

Anxiety Disorders

GENERALIZED ANXIETY DISORDER. GAD is one of the diagnostic categories with significant reliability/validity problems, and it is considered a residual category because it is very frequently found along other disorders. In spite

of this, GAD retains clinical relevance and is considered to have important negative consequences for health and quality of life. Proposals for GAD in ICD-11 include:

- Better ways to emphasize worry about everyday matters as the core feature of GAD. If apprehensive expectation, which is related but different construct from worry, should be also a core feature (or even the core feature) is open to discussion.
- Deciding if "excessive worry" (which is difficult to operationalize and not always recognized by patients) and/or its frequency/duration should define the boundaries with "normal worrying."
- Emphasize restlessness, tension, and gastrointestinal symptoms as the most frequent/important somatic complaints in patients with GAD.
- Change the status of the residual category of GAD by allowing the diagnosis to be made in presence of other disorders in which worrying is not the central feature.

PANIC DISORDER AND AGORAPHOBIA. The main current proposal for ICD-11 is to separate panic disorder and agoraphobia as DSM-5 has also done. This may increase artificially the level of comorbidity, but at the same time, the separation of these two disorders is consistent with research results. An important issue about anxiety disorders is the importance of contextual and culturally mediated influences on clinical presentation like the contents of worry, the importance attached to worrying in certain cultures, the places and situations where agoraphobia can occur, the fear of offending others versus being criticized and other instances.

SPECIFIC PHOBIAS AND SOCIAL PHOBIA. The subtyping of specific and social phobias in DSM is an important difference with ICD. ICD workgroup members argue that naming specific phobias has no clear clinical utility.

Obsessive-Compulsive Disorder Spectrum Disorders. The main issue with respect to Obsessive-Compulsive Disorder (OCD) is whether or not ICD-11 will follow DSM-5 in creating a separated category for OCD and related disorders. Although this question has not been officially answered, it seems likely that ICD-11 will also separate OCD from anxiety disorders. Other suggestions for OCD in ICD-11 include the notion of urge in addition to the one of impulse and some new clinical qualifiers (especially the so called "sensitive phenomena").

Posttraumatic Stress Disorder. Probable decisions regarding diagnosis and classification of PTSD in ICD-11 will include:

- The creation of a separate diagnostic category for stress-related disorders

- ► The discussions about the quality of the traumatic/distressing event related to PTSD and other clinical problems in this new category
- ► Increased attention to the distinction between PTSD and normal "adaptive fear reactions" to ongoing trauma (e.g., continuing conflict, forced migration, and natural disasters).
- ► Inclusion of complex PTSD
- ► Identifying acute stress reaction as a normal reaction to an abnormal event, and
- ► Emphasizing the advantage of ICD-11 over DSM-5 because of greater simplicity, greater clinical utility and greater feasibility in "low resource and humanitarian settings."

Intellectual Disability/Mental Retardation. The working group on mental retardation has published several papers on the rationale and process for replacing the ICD-10 conceptualization of mental retardation (MR) with the concept of intellectual developmental disorders (IDD) in ICD-11.

At least during the last two decades, there have been very important discussions on how to properly name, define, and assess IDD. In general, the term "intellectual disability" (ID) has widely replaced MR for policy, administrative, and legislative purposes in many industrialized countries and in an increasing number of low- and middle-income countries. However, the question as to whether IDD represent a disability or a health condition remains unanswered. There was a consensus in the working group about the need to reorganize IDD as part of the larger grouping of neurodevelopmental disorders. The term "intellectual" was favored because in most countries it is widely used, and is broadly acceptable in many contexts. The term "developmental" refers to the period of time during which the brain and its functions are developing and thus can be impaired. The term "developmental" also implies a process and a lifespan perspective and emphasizes the dynamic nature of IDD. The proposed definition of IDD is "a group of developmental conditions characterized by significant impairment of cognitive functions, which are associated with limitations of learning, adaptive behavior, and skills." The main descriptors of IDD include:

- ► Marked impairment of core cognitive functions that can have different patterns. This impairment affects the development of knowledge, reasoning, and symbolic representation, according to the level expected for age peers and the cultural and community environment.
- ► Problems with verbal comprehension, perceptual reasoning, working memory, and processing speed.
- ► The cognitive impairment is associated to difficulties in learning, academic, and practical knowledge.

- ► Problems with adaptive behavior with limitations in relevant conceptual, social, and practical skills.
- ► Difficulties in managing behavior, emotions, interpersonal relationships, and motivation.
- ► IDD is a life span condition requiring consideration of developmental stages and life transitions.

It is important to signal that the proposed approach distinguishes between IDD (a clinical meta-syndrome) and intellectual disability (the functioning/disability counterpart), which have different scientific, social, and policy applications.

Alcohol-Related Disorders. It is expected that ICD-11 will include the following alcohol-related disorders: acute alcohol intoxication, harmful use, dependence syndrome, withdrawal, mental disorders, and alcohol-induced disorders arising from the use of alcohol. Similar codes are to be found in the current version of ICD-10. The new coding of alcohol intoxication defines it as a transient state following alcohol consumption resulting in disturbances in level of consciousness, cognition, perception, affect, behavior, or other psychophysiological functions and responses. Typical symptoms include mood instability, false judgment, impaired social or professional functioning, and improper sexual or aggressive behavior. Harmful use of alcohol will probably be defined as a pattern of alcohol use that is causing damage to health following repetitive episodes of intoxication, regular intake of large quantities of alcohol, or harmful use of alcohol. The damage may be physical or mental, including violence and self-harmful acts of bodily damage requiring medical intervention. Alcohol dependence core features would include:

- ► Difficulties in controlling substance-taking behavior often in combination with strong desire or sense of compulsion to take alcohol
- ► Alcohol use plays a dominant role in the individual's life
- ► Physiological symptoms: tolerance and characteristic physical withdrawal state or syndrome (not just hangover)

Personality Disorders. The conceptualization and classification of personality disorders in psychiatry and by psychiatrists is one of the areas of persistent debate among mental health clinicians and researchers. In general, there is a profound dissatisfaction with the psychiatric approach to this problem. Tyrer (2015) in a very recent paper summarizes what could be expected from ICD-11 in this regard:

- ► Abolition of type-specific categories of personality disorder, apart from the presence of personality disorder itself.
- ► Different degrees of severity are defined and this will be helpful to

decide on the individual's personality functioning at the time of the assessment.

- Five domains would be used as qualifiers: (1) negative affective features (tendency to manifest a broad range of distressing emotions); (2) dissocial features (disregard for social obligations, conventions, and the rights and feelings of others); (3) features of disinhibition (persistent tendency to act impulsively); (4) anankastic features (narrow focus on the control of one's own and others' behavior to ensure that it conforms to the "individual particularistic ideal"); and (5) features of detachment (emotional and interpersonal distance).

Table 9.2–7 shows the category names and essential features of the proposal for personality disorders in ICD-11.

Feeding and Eating Disorders. This is another area of controversy and dissatisfaction for clinicians and researchers. In efforts to improve the state of affairs, a number of proposals have been made as summarized below:

- Feeding and eating disorders should be merged into a single grouping.
- Regarding anorexia nervosa, dropping some requirements (e.g., amenorrhea) and extending some of the criteria (e.g., the weight criterion to any significant underweight, including developmentally and culturally relevant presentations, among others will broaden the category. Very importantly, a severity qualifier "with dangerously low body weight" will be included to distinguish the most severe cases of anorexia nervosa.
- Bulimia nervosa should be extended to include subjective binge-eating.
- Binge-eating disorder would be included as a specific category. The definition will include subjective or objective binge eating without regular compensatory behaviors.

ICD-11 AND DSM-5: SOME CONTROVERSIES

Psychiatric classification is undoubtedly one of the most polemic issues surrounding psychiatry as a medical specialty. The controversies about classification can be summarized as follows:

- Epistemological, for example, what kind of knowledge is gained from classifications and what are their knowledge limits?
- Metaphysical and ontological, for example, what is a mental disorder? Are mental disorders natural kinds or constructs?
- Ethical, for example, what are the consequences of using a classification system in terms of the good that can be achieved for patients, society, and psychiatry?
- Practical (usefulness). The usual debate is if classification systems serve the purposes for which they were created, that is, improving

communication among clinicians, helping clinical practice, research, and administrative/legal issues?

- ► Methodological. What is the best way to classify mental disorders and how this way can be carried out?
- ► Sociological, for example, what are the politics behind classification of mental disorders? How do clinical and political uses of classifications relate to each other? What relationships exist between classification of mental disorders and vested interests like those of pharmaceutical companies? What is the sociology of the groups developing classifications?
- ► Legal, for example, how do classifications influence the legal system? How do the needs of legal systems influence psychiatric classification?
- ► Social, for example, do classification systems promote (or defuse) stigma?
- ► Personal, for example, do classifications create self-stigma in the patients receiving aid from legal or governmental systems based on a psychiatric diagnosis?

The existence of two main classification systems that have important similarities but equally important differences adds up to these complex controversies. The launching of the DSM-5 has revitalized several discussions both at theoretical and practical levels, one of the most important being the issue of how much ICD and DSM should resemble each other in terms of theory of mental disorders, aims, scope, methodology, and content, among others. This section will focus on the debates about the issue of harmonization, because some think that both classifications must resemble each other as much as possible but at the same time both are needed, while others on the contrary maintain that there must be a certain distance and/or independency between both, and finally, there are authors who think that at the end one of them should replace the other.

Table 9.2–7.
Category Names and Essential Features of the Proposal for Personality Disorders in ICD-11

Personality disorder	A pervasive disturbance in how an individual experiences and thinks about the self, others, and the world, manifested in maladaptive patterns of cognition, emotional experience, emotional expression, and behavior. The maladaptive patterns are relatively inflexible and are associated with significant problems in psychosocial functioning that are particularly evident in interpersonal relationships. The disturbance is manifest across a range of personal and social situations (i.e., is not limited to specific relationships or situations). The disturbance is relatively stable over time and is of long duration. Most

	commonly, personality disorder has its first manifestations in childhood and is clearly evident in adolescence.
Late-onset qualifier	If the disturbance has its origin in adulthood, the qualifier for "late onset" may be added. The "late-onset" qualifier should be used for cases in which, by history, there is no evidence of personality disorder or its early manifestations prior to age 25 years.
Mild personality disorder	There are notable problems in many interpersonal relationships and the performance of expected occupational and social roles, but some relationships are maintained and/or some roles carried out.
	Mild personality disorder is typically not associated with substantial harm to self or others.
Moderate personality disorder	There are marked problems in most interpersonal relationships and in the performance of expected occupational and social roles across a wide range of situations that are sufficiently extensive that most are compromised to some degree.
	Moderate personality disorder often is associated with a past history and future expectation of harm to self or others, but not to a degree that causes long-term damage or has endangered life.
Severe personality disorder	There are severe problems in interpersonal functioning affecting all areas of life. The individual's general social dysfunction is profound and the ability and/or willingness to perform expected occupational and social roles is absent or severely compromised.
	Severe personality disorder usually is associated with a past history and future expectation of severe harm to self or others that has caused long-term damage or has endangered life.

From Tyrer P, Reed GM, Crawford MJ. Classification, assessment, prevalence, and effect of personality disorder. *Lancet*. 2015;385(9969):71.

Harmonization

The issue of how much DSM and ICD should resemble each other has recently been discussed under the label of "harmonization." Despite the aesthetic and rhetorical appeal that this term conveys, the concept needs to be carefully defined and articulated. The Oxford Advanced Learner's Dictionary defines "harmonize" as follows:

- "To harmonize with something" = "if two or more things harmonize with each other or one thing harmonizes with the other, the things go well together and produce an attractive result."
- "Harmonize" = "something to make systems or rules similar in different countries or organizations."

Therefore, at least from the perspective of current English language, harmonization of the two systems is not simply to combine them but to make them similar and/or equivalent. Because of the international and multicultural character of ICD, the meaning of harmonization was also checked in leading Spanish, French, Italian, and German dictionaries. This reveals semantic differences that should be kept in mind for the harmonization enterprise. In Spanish, for example, emphasis is made on

the fact that the parts or objects to be harmonized should not reject each other. In French, the definition emphasizes agreement between elements and the importance of the goal to be achieved. Thus, in all of these languages, harmonization does not simply mean unification but some sort of "shared understanding" and complementarity. However, harmonization is also a technical term used for the main goal of reaching a high degree of similarity between international systems of rules, laws, definitions, or specific criteria.

The Direction of ICD-DSM Harmonization

One of the most sensitive issues about this harmonization project is whether ICD should resemble DSM or DSM resemble ICD. Probably, in some instances, the harmonization should go in one direction (e.g., DSM to ICD) while in others, the opposite would be the case (e.g., from DSM to ICD). The necessary comparison between both systems and the decision that one would prevail over the other adds sensitivity to the process. For example, if it is decided that anxiety disorders are better approached by ICD than DSM, DSM should make the effort to resemble ICD in regards to this group of disorders. Conversely if the criteria for mood disorders in DSM are deemed better, thus ICD should adopt them. No matter how logical this sounds, problems and debates will arise among proponents or detractors of either system.

Based on the above considerations, criteria for deciding the direction of harmonization should take into account:

- Methodological issues: Which of the two systems can ensure that its rationale and criteria are less subject to undesirable and avoidable biases?
- Empirical support: Which of the two systems has been tested by means of well-designed studies (including field trials) whose results support the usefulness, validity, or reliability of the diagnostic categories?
- Ease of use: Which system is easier for clinical use and well-liked by busy clinicians?
- Participation of different stakeholders: Which system has duly incorporated the opinions and/or feedback of patients, families, clinicians, and health providers?
- Transnational sensitivity and precision: Which classification is better suited for use across different countries, cultures, and languages?
- Cross-cultural sensitivity and precision: Which classification can be better adapted to different cultural environments?
- Development and advance potential: Which classification has a structure that allows rapid incorporation of changes or scientific advances, especially now that both aspire to be "living documents"?

Clearly, these matters cannot be resolved using only empirical data and at the end, harmonization becomes a political issue. Here, negotiation and the ability to reach agreements are at the core of the harmonization projects. Thus, both empirical and nonempirical issues have to be addressed in the process of harmonization.

"Harmonization" in the Psychiatric Literature

To gain some insight on this issue, the authors reviewed papers addressing directly (in their title, abstract and/or key words) the issue of "harmonization" of the two diagnostic systems. Most of the articles merely state the general importance of "harmonization" but do not provide additional details or recommendations. The emphasis to harmonize clearly stems from the perceived differences across the two systems that may interfere with communication among clinicians and international research efforts. Clinicians, researchers, and developers of both classifications have all called for harmonization. For most of these parties, harmonization entails diminishing and if possible eliminating significant differences between ICD and DSM. Also there is a general consensus on the importance of harmonization, even in the case of authors who are critical about DSM and/or ICD.

The WHO addressed this issue directly by creating a "DSM-ICD Harmonization Coordinating Group" that includes members of the DSM-5 Task Force and the International Advisory Group revising ICD-10. The charge for this group was "to facilitate the achievement of the highest possible extent of uniformity and harmonization between ICD-11 mental and behavioral disorders and DSM-5 psychiatric disorders and their diagnostic criteria." Also it was recommended that "The WHO and APA should make all attempts to ensure that in their core versions, the category names, glossary descriptions, and criteria are identical (author's emphasis) for ICD and DSM."

The major tasks of this group were as follows:

- To identify effective mechanisms to share information on developmental processes of the two diagnostic systems.
- To find strategies to narrow the differences between the two diagnostic systems.
- To coordinate efforts toward harmonization of the two diagnostic systems by making specific recommendations to the respective decision-making bodies.
- To provide a forum in which difficult issues around harmonization can be discussed and common approaches identified.

As articulated before, harmonization and unification are different goals although they could be part of a sequence, starting with harmonization and

once empirical data is available and conceptual agreements and political consensus reached, unification could follow. Clearly, there are system differences that are more difficult to harmonize, for example, conceptual and ideological issues, but others, for example, operational ones, that can be approached in a straightforward fashion. The process should start by listing and describing the main differences between both classification systems. A second step would entail a detailed analysis about the nature of these differences, for example, which of them are operational, which of them are terminological, and which of them are conceptual. A third step would be to tackle first terminological and operational differences, and later on focus on the conceptual ones.

The American nosologist, Michael First, has proposed three basic harmonization strategies:

- ► To select one of the two systems as the starting point for a future single and unified classification.
- ► To create a new common set of DSM-5 and ICD-11 criteria by selecting the best elements from each system in terms of reliability, validity, and clinical utility.
- ► A third useful approach in the case of each disorder would be to pick the definitions and criteria from either system to become the "harmonized" DSM-5 and ICD-11 definition for each disorder. For a particular disorder, the decision to choose either the DSM-IV or ICD-10 criteria set could be based on various mutually agreed criteria, such as empirical evidence of superior validity, reliability, or clinical utility.

At this time, it is not possible to know exactly how discordant or concordant the final version of ICD-11 will be with respect to DSM-5. Harmonization is an important task, but a very difficult one, especially because of the differences in the origins, purposes, scope, concepts, and ideology underlying both classifications.

SUMMARY AND CONCLUSIONS

- ► The process of revising the ICD-10 chapter of mental disorders and producing the new ICD-11 is currently underway. It seems quite likely that before the end of 2017 the full new version of ICD-11 will be released.
- ► It appears that ICD-11 will remain faithful to the principles and purposes underlying this international classification.
- ► It seems clear that changes to be included in ICD-11 will be driven mostly by clinical utility.
- ► There are proposals for changes both in the general structure of the classification of mental disorders and also the criteria for several individual disorders. However, it is difficult to predict at this time the

final changes included in the official version.

- ► Although harmonization between ICD-11 and DSM-5 is one of the main goals articulated by the developers and potential users of each of these systems, it is likely that significant differences across the two systems will remain at the end. However, it is hoped that these differences would not be as marked as those reported for ICD-10 and DSM-IV.

REFERENCES

Adamowski T, Kiejna A. Classifications of mental disorders from a historical and contemporary perspective. *Adv Clin Exp Med*. 2007;16(2):189–195.

Andrews G, Slade T, Peters L. Classification in psychiatry: ICD-10 versus DSM-IV. *Br J Psychiatry*. 1999;174:3–5.

Biedermann F, Fleischhacker WW. Psychotic disorders in DSM-5 and ICD-11. *CNS Spectr*. 2016;21(4):349–354.

Bornstein RF. Toward a firmer foundation for ICD-11: On the conceptualization and assessment of personality pathology. *Personal Ment Health*. 2016;10(2):123–126.

Cooper JE, Sartorius N. *A Companion to the Classification of Mental Disorders*. Oxford: Oxford University Press; 2013.

Croudace T, Evans J, Harrison G, et al. Impact of the ICD-10 Primary Health Care (PHC) diagnostic and management guidelines for mental disorders on detection and outcome in primary care. Cluster randomised controlled trial. *Br J Psychiatry*. 2003;182:20–30.

De Dios C, Goikolea JM, Colom F, Moreno C, Vieta E. Bipolar disorders in the new DSM-5 and ICD-11 classifications. *Rev Psiquiatr Salud Ment*. 2014;7(4):179–185.

**First MB. Harmonisation of ICD-11 and DSM-V: opportunities and challenges. *Br J Psychiatry*. 2009;195(5):382–390.

First MB, Westen D. Classification for clinical practice: how to make ICD and DSM better able to serve clinicians. *Int Rev Psychiatry*. 2007;19(5):473–481.

**First MB, Reed GM, Hyman SE, Saxena S. The development of the ICD-11 clinical descriptions and diagnostic guidelines for mental and behavioural disorders. *World Psychiatry*. 2015;14(1):82–90.

Goldberg D. A revised mental health classification for use in general medical settings: the ICD-11-PHC. *Int Psychiatry*. 2011;8:1–3.

Goldberg DP, Prisciandaro JJ, Williams P. The primary health care version of ICD-11: the detection of common mental disorders in general medical settings. *Gen Hosp Psychiatry*. 2012;34(6):665–670.

Grant JE, Chamberlain SR. Expanding the definition of addiction: DSM-5 vs. ICD-11. *CNS Spectr*. 2016;21(4):300–303.

Hyland P, Shevlin M, McNally S, et al. Exploring differences between the ICD-11 and DSM-5 models of PTSD: Does it matter which model is used? *J Anxiety Disord*. 2016;37:48–53.

International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders. A conceptual framework for the revision of the ICD-10 classification of mental and behavioural disorders. *World Psychiatry*. 2011;10(2):86–92.

International Advisory Group for the Revision of ICD–10 Mental and Behavioural Disorders.

Summary Report of the 3rd Meeting of the International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders. WHO; 2008.

Jablensky A. Towards ICD-11 and DSM-V: issues beyond "harmonisation." *Br J Psychiatry*. 2009;195(5):379–381.

Maercker A, Brewin CR, Bryant RA, et al. Proposals for mental disorders specifically associated with stress in the International Classification of Diseases-11. *Lancet*. 2013;381(9878):1683–1685.

**Maj M, Wolfgang G, López-Ibor JJ, Sartorius N. *Psychiatric Diagnosis and Classification*. Chichester: John Wiley & Sons; 2002.

Marras A, Fineberg N, Pallanti S. Obsessive compulsive and related disorders: comparing DSM-5 and ICD-11. *CNS Spectr*. 2016;21(4):324–333.

Miklowitz DJ, First MB. Specifiers as aids to treatment selection and clinical management in the ICD classification of mood disorders. *World Psychiatry*. 2012;11(Suppl 1):11–16.

Moriyama IM, Loy RM, Robb-Smith AH. In: Rosenberg HM, Hoyert DL, eds. *History of the Statistical Classification of Diseases and Causes of Death*. Hyattsville, MD: National Center for Health Statistics; 2011.

Paykel E, Andrade LH, Njenga F, Phillips ML. Changes needed in the classification of depressive disorders: options for ICD-11. *World Psychiatry*. 2012;11(Suppl 1):37–42.

Peralta V, Cuesta MJ. The nosology of psychotic disorders: a comparison among competing classification systems. *Schizophr Bull*. 2003;29(3):413–425

Reed GM, Mendonça Correia J, Esparza P, Saxena S, Maj M. The WPA-WHO global survey of psychiatrists' attitudes towards mental disorders classification. *World Psychiatry*. 2011;10(2):118–131.

Sartorius N. Revision of the classification of mental disorders in ICD-11 and DSM-V: work in progress. *Adv Psychiatr Treat*. 2010;16(1):2–9.

**Sartorius N. The classification of mental disorders in the international classification of diseases. In: Sadock BJ, Sadock VA, Ruiz P, eds. *Kaplan & Sadock's Comprehensive Textbook of Psychiatry*. 9th ed. Philadelphia, PA: Wolters Kluwer Lippincott Williams & Wilkins; 2009:1139–1149.

**Tyrer P. A comparison of DSM and ICD classifications of mental disorder. *Adv Psychiatr Treat*. 2014;20(4):280–285.

Tyrer P, Reed GM, Crawford MJ. Classification, assessment, prevalence, and effect of personality disorder. *Lancet*. 2015;385(9969):717–726.

Ustün TB, Goldberg D, Cooper J, Simon GE, Sartorius N. New classification for mental disorders with management guidelines for use in primary care: ICD-10 PHC chapter five. *Br J Gen Pract*. 1995;45(393):211–215.

World Health Organization Family International Classification Network. Production of ICD-11: The overall revision process. 2007. Available on: http://www.who.int/classifications/icd/ICDRevision.pdf.

10 Neurocognitive Disorders

▲ 10.1 Cognitive Disorders: Introduction

ROBERT A. SWEET, M.D.

Cognition includes attention, memory, language, orientation, praxis, executive function, judgment, and problem solving. Disorders of cognition, therefore, can impact personal, occupational, and social functioning. The demographic imperative with regard to cognitive disorders is clear. The population is rapidly aging, and the elderly are the demographic group most at risk of cognitive disorders. The joint risk of all of these disorders, provided by aging, challenges clinicians as individual presentations are characterized by multiplicity, comorbidity, and unclear boundaries.

Delirium is particularly problematic in this regard. Delirium is often superimposed on an existing, although often unrecognized, dementia. Tanya Fabian and LalithKumar Solai (Section 10.2) review the burden generated by delirium in medical settings. Delirium is ubiquitous in acute medical settings, making the elderly and, in particular, those with other cognitive disorders vulnerable. Nevertheless, early recognition and correct differential diagnosis of delirium remain a challenge due to the variable nature of the presenting signs and symptoms. The development of effective interventions for delirium has been slow to emerge, perhaps because of its multifactorial nature. Recent work, which indicates the effectiveness of multicomponent behavioral and supportive interventions in prevention of delirium, provides some hope and a model for reengineering acute hospital care for at-risk patients.

Conversely, certain dementias, such as dementia with Lewy bodies or late stages of Alzheimer disease, may have chronic clinical presentations virtually indistinguishable from delirium except for temporal onset and the lack of an identifiable acute source. Similarly, the course of nearly all subjects developing a progressive dementia is complicated by the onset of one or more distinct behavioral syndromes, including anxiety, depression, sleep problems, psychosis, and aggression. These symptoms can be as distressing and disabling as the primary cognitive disorder. Some of these

behavioral syndromes, such as psychosis, may themselves result from independent underlying biologies and may be additive with the primary neurodegenerative process. Unfortunately given the prevalence of these behavioral changes, efforts at the development of new treatments targeting behavioral disruption in dementia have been very limited but are now emerging as a new focus for investigation.

The boundaries between types of dementia can be similarly diffuse. The most common neuropathologic presentations associated with dementia reveal mixtures of Alzheimer disease, vascular, and Lewy body pathologies. Pure syndromes are relatively less common, although often the dementia is ascribed to one of the coexisting pathologies. Strategies regarding how to understand or reconcile the potential contributions of multiple pathologies to disease onset and progression in the clinic are needed. Some of these are now emerging as the result of the approval of several agents for in vivo imaging of amyloid plaques, the defining pathology of Alzheimer disease. Meanwhile, agents for imaging tau pathology have entered clinical trials. Tau pathology, while present in Alzheimer disease, also extends to define another set of neurodegenerative dementias that include frontotemporal dementia, corticobasal degeneration, and progressive palsy. Finally, advances in the genetics of frontotemporal dementia have been quite rapid. Mutations in the *C9ORF72* gene are now established as the cause of more than 20 percent of familial and 5 percent of sporadic cases of frontotemporal dementia (and higher proportions of motor neuron disease). The combination of mutations in *C9ORF72*, *GRN*, *MAPT*, and other genes can now account for the majority of cases of familial frontotemporal dementia (and up to half of all cases of frontotemporal dementia are familial), and it has provided another mechanism to differentiate among neurodegenerative disorders.

Julie Hugo and Robert Sweet review recent advances in the current understanding of the neurodegenerative dementias (Section 10.3). In addition to the advances in imaging and genetics, great strides have been made in understanding Alzheimer disease, for example, identifying the mechanisms by which accumulation of amyloid beta protein can interfere with cognitive function, even prior to the emergence of the defining neuropathologic lesions, plaques, and tangles. This elaboration of the "amyloid hypothesis" of Alzheimer disease has translated into the development of a number of antiamyloid therapies. However, trials to treat patients with dementia due to Alzheimer disease with these agents have disappointed. The study of antiamyloid agents has now largely shifted to attempts to delay or prevent dementia onset in individuals at risk of Alzheimer disease due to genetic mutations or biomarker evidence of amyloid accumulation.

With increased understanding of mechanisms of dementia, the clinical imperative has been to develop approaches for early case identification.

Daniel Varon and Carmen Andreescu, in Section 10.4 on amnestic disorders, review the increasing knowledge about a clinical presentation that often indicates the presence of a neurodegenerative disorder, mild cognitive impairment (MCI). When MCI presents as an isolated amnestic symptom in late life, it predicts subsequent onset of Alzheimer disease. Although other conditions, including other neurodegenerative syndromes, focal brain lesions, metabolic disturbances, and alcohol use, may lead to isolated amnestic disorders, in practice MCI due to underlying Alzheimer disease is more common. Recognition that neurodegeneration may precede clinical symptoms by 10 years or more has further led to the proposal of stages for preclinical Alzheimer disease, characterized by the presence of biomarkers of amyloid accumulation and neurodegeneration in the absence of overt cognitive and behavioral symptoms.

The recognition of a long, preclinical stage of neurodegeneration also indicates that in some elderly individuals who have neurodegenerative pathology and cognitive impairments, the cognitive symptoms might actually result from a comorbid condition, independent of or additive with the degenerative lesions. In Section 10.5, Jennifer Prins and Ellen Whyte review other cognitive disorders and mental disorders due to a general medical condition. Their section covers additional medical and substance-induced causes of dementia and amnestic disorders that form the differential diagnoses for many of the conditions reviewed in the earlier chapters. Importantly, their review extends to mood and psychotic disorders that may emerge as consequences of general medical conditions, presentations likely to result in psychiatric consultation and referral. Mood syndromes are particularly likely to emerge in brain disorders such as multiple sclerosis, and stroke, and their detection and management in these circumstances may differ substantially from that of idiopathic disorders. Perhaps more than any other of the disorders described in this section, these conditions place a premium on providing care collaboratively with other medical, neurologic, and rehabilitation specialists.

References

Geda YE, Schneider LS, Gitlin LN, et al. Neuropsychiatric symptoms in Alzheimer's disease: past progress and anticipation of the future. *Alzheimers Dement*. 2013;9(5):602–608.

Rademakers R, Neumann M, Mackenzie IR. Advances in understanding the molecular basis of frontotemporal dementia. *Nat Rev Neurol*. 2012;8(8):423–434. Erratum in: *Nat Rev Neurol*. 2013;9(5):240.

Schneider JA, Arvanitakis Z, Bang W, Bennett DA. Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology*. 2007;69:2197–2204.

Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging- Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7:280–292.

▲ 10.2 Delirium

TANYA J. FABIAN, PHARM.D., PH.D., AND LALITHKUMAR K. SOLAI, M.D.

Delirium as defined by the DSM-5 is characterized by an acute decline in both the level of awareness (changed from consciousness in DSM-IV-TR) and cognition with particular impairment in attention. Delirium often involves perceptual disturbances, abnormal psychomotor activity, and sleep cycle impairment and can be life threatening. Due to a high prevalence among inpatient (particularly geriatric) populations, poor recognition, and a high mortality rate, delirium significantly complicates prognosis and the delivery of health care (Table 10.2–1).

Delirium is fairly common on admission to hospitals ranging in rates between 14 and 24 percent, which further increase during hospitalization to anywhere between 6 and 56 percent. This prevalence increases with other factors such as age, medication use, and comorbidities. Despite such a common occurrence in hospitals, delirium is often unrecognized by medical professionals. Estimates of this misdiagnosis have been between 40 and 60 percent depending on the setting studied. In most cases, the development of delirium is multifactorial, including stress or trauma to the central nervous system (CNS), drug toxicity or withdrawal, and metabolic disorders arising from organ failure. Despite the numerous etiologies that have been noted, the underlying pathophysiology of delirium is still not clearly understood.

Table 10.2–1.
Burden of Delirium

Increased nursing care
Increased length of stay
Increased risk of cognitive decline
Increased risk of functional decline
Increased mortality
Delay in postoperative mobilization
Prevention of early rehabilitation
Increased rate of nursing home placement
Increased need for home care services
Increased distress to caregivers
Barrier to psychosocial closure in terminally ill patient

DEFINITION

The term delirium originates from the Latin verb *deliro*—to be crazy, which is taken from *de* + *lira*, a furrow (i.e., to go out of the furrow). The core symptoms of delirium include a disturbance of awareness that is accompanied by a change in attention that develops rapidly, usually hours

to days and tends to fluctuate during the course of the day. The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) defines delirium in the following manner: Foremost a disturbance of awareness and attention. It is a common clinical syndrome that typically heralds an increase in morbidity and mortality. Typically there is an identified symptom(s) or sign(s) that could explain the change observed in the cognition. When a specific general medical condition has been identified as the cause of the delirium, it is diagnosed as delirium due to general medical condition, such delirium due to pneumonia. At times the symptoms of delirium could be secondary to drug intoxication (e.g., cocaine or alcohol) or due to drug withdrawal (e.g., alcohol or benzodiazepine). But in reality most of the delirious episodes happen in the context of multiple factors, such as dehydration, pneumonia, pain in a patient who is abusing alcohol. These deliriums are defined as delirium secondary to multiple etiologies. International Classification of Diseases and Related Health Problems (ICD-10) classifies delirium as either due to (1) alcohol or other psychoactive substances (2) not induced by alcohol or other psychoactive substances. For consistency reasons, this chapter will focus on the delirium, not induced by alcohol or other psychoactive substances. ICD-10 describes delirium (similar to DSM-5) as an etiologically nonspecific organic cerebral syndrome with features impairing consciousness and attention, perception, thinking, and memory. But it also includes other symptoms and signs as part of the diagnostic criteria for delirium. These include (but not limited to) the following: psychomotor disturbance, disturbance of the sleep–wake cycle, and emotional disturbance (such as irritability).

Patients with delirium remain in the hospital longer and are more commonly discharged to long-term care facilities. Behavioral manifestations of delirium may interfere with treatment compliance and are often precipitants for psychiatric consultation. A psychiatrist should provide and/or advocate for the appropriate treatment beyond simple medical expedience.

HISTORY

The earliest known references to delirium in medical literature are found in the writings of Hippocrates. Some 2,400 years ago, in his Book of Epidemics, Hippocrates described the following case:

Erasinus, who lived near the Canal of Bootes, was seized with fever after supper; passed the night in an agitated state. During the first day, he was quiet, but in pain at night. On the second day, symptoms all exacerbated; at night [mad]. On the third day, he was in a painful condition; great incoherence. On the fourth day, he was in a most uncomfortable state; had no sound sleep at night, but dreaming and talking; then all the appearances worse, of a formidable and alarming character; fear, impatience. On the

morning of the fifth day, he was composed, and quite coherent, but long before noon was furiously mad, so that he could not constrain himself; extremities cold, and somewhat livid; urine without sediment; died about sunset.

Although Hippocrates predated the inception of the Latin term *delirium*, his astute observations clearly outline many of the syndrome's essential features. Tradition holds that the word delirium was first used in the formal medical context by Celsus in the first century AD. However, Celsus used the term delirium to describe a spectrum of mental disorders ranging from general insanity to acute transient states of mental disturbance, including *phrenitis*, *lethargus*, hysteria, melancholia, and mania.

Aretaeus of Cappadocia, a contemporary of Celsus, was perhaps the first to classify mental disorders into acute and chronic categories. Aretaeus identified both *phrenitis* and *lethargus* as acute manifestations of disease. The ancient Greek terms *phrenitis* and *lethargus* have given rise to the modern English words frenzy and lethargy, respectively. This observation is significant as it introduces the concepts that delirium is both an acute syndrome and that it consists of different variants or subtypes. This distinction by Aretaeus may represent the first recorded description of both the hyperactive and hypoactive motor elements of delirium.

Several centuries later, Phillip Barrough advanced and further clarified the concept of delirium in his 1583 textbook *The Method of Physick*. Barrough proposed that delirium constituted a derangement of some combination of three main internal senses: including imagination, cognition and memory. This concept of deranged senses was elaborated by Thomas Willis in his 1672 treatise De Anima Brutorum which established that delirium was in fact a specific set of symptoms and not a disease.

In the 18th century, Erasmus Darwin and John Hunter made significant contributions to the evolving theory of delirium. Darwin was the first to compare delirium with the dream state, noting that both states constituted an interruption of "voluntary power" and a suspension of the ability to attend to one's external environment. Darwin notion of the interrelation of delirium and dreams was expanded even further by John Hunter, who defined delirium as "a cessation of consciousness" of one's own existence.

Later in the same century, James Sims was the first to argue that delirium was distinctly different from general insanity and that it constituted an "alienation of the mind." Sims also distinguished two variants of delirium, which he referred to as low or raving, that correspond almost perfectly with the modern concepts of hypoactive and hyperactive motor subtypes of delirium.

The 19th century saw critical advances in the development of the concept of delirium. A medical dictionary published by Rees in 1818 argued strongly for a clarification and unification of the concept of delirium within

the medical community. He even went so far as to scold other medical writers who persisted in using the term delirium in an ambiguous sense. Rees built on the work of his predecessor Sims, insisting that delirium be recognized as a distinct condition with two specific variants and a unique etiology that was localized to the brain:

This is the sense in which delirium is commonly understood: but by some writers ... delirium is the term employed to denote the whole class of mental derangements, including the different form of mania, melancholy and febrile delirium.

Arguably the most important contribution in development of the concept of delirium took place in the 20th century when George Engel and John Romano demonstrated that delirium was due to a reduction in the metabolic activity of the brain. Through the use of electroencephalograms (EEGs), Engel and Romano illustrated decreased background activity corresponded proportionately with decreased in cognition, memory, and attention. Interestingly, these findings of reduced metabolic activity were observed in all subtypes of delirium, regardless of the level of psychomotor activity.

In introducing a reliable diagnostic tool and demonstrating the underlying mechanism of this mental disorder, Engel and Romano finally confirmed the observations of countless astute clinicians. The history of delirium stretches back over two millennia, yet it is only in the later part of the past century that psychiatrists have managed to unify its concept and elucidate its mechanisms.

COMPARATIVE NOSOLOGY

Since delirium is usually due to multifactorial causes, there are many synonyms for delirium that have been introduced and adopted by medical practitioners. Various specialties have used different terms such as encephalopathy and acute brain syndrome. Unfortunately, these multiple terms falsely suggest an element of variation that further complicates the recognition and management of delirium within and across the medical specialties. In psychiatry, the term *acute confusional state* has been used as an equivalent designation for delirium, despite the ambiguity this additional term creates.

In general medicine and many of the medical subspecialties, the terms *encephalitis* and *encephalopathy* prevail in both medical literature and clinical practice. *Encephalitis* refers to an inflammation of the brain parenchyma, as distinct from the meninges, which is usually associated with an infectious state. The term *encephalopathy* is less specific and refers to any disorder of the brain parenchyma, yet is most often associated with organ failure and metabolic disorders. These examples illustrate the diversity of terms that are employed to identify delirium, and many other examples abound (Table 10.2–2).

Although many of the aforementioned terms accurately describe individual elements of delirium, they fail to present a comprehensive summary of the patient. To be accurately diagnosed and managed, all aspects of the delirium, from etiology to resultant symptoms, must be included in the assessment. This comprehensive evaluation requires a consistent nosologic system. Great strides have been made in the establishment of such a system since the American Psychiatric Association's first attempts to clarify the concept of delirium in the third edition of the DSM-III. The tenth edition of the ICD-11 classification of delirium (F05) now has many similarities with DSM-5.

The confusion surrounding the use and definition of the term delirium is further compounded by the existence of clinical subtypes referred to as the hyperactive, hypoactive, and mixed states of psychomotor activity. It should be noted that these designations are strictly clinical and represent different manifestations of the same pathophysiologic state. Misconceptions regarding all of these terms are a significant concern as they may result in a failure to correctly diagnose and manage the delirious patient.

Table 10.2–2.
Other Terms Describing Delirium

Acute confusional state
Acute brain failure
Encephalitis
Encephalopathy
Intensive care unit psychosis
Toxic metabolic state
Central nervous system toxicity
Paraneoplastic limbic encephalitis
Sundowning
Cerebral insufficiency
Organic brain syndrome

EPIDEMIOLOGY

Incidence and Prevalence

Delirium is a common disorder in older patients, with the highest incidence and prevalence rates reported in older people. These rates of delirium vary based on the setting patients are in. In community studies, 1 percent of the elderly population aged 55 or more have delirium (13 percent in the 85 and older group in the community). Studies among elderly emergency room subjects have reported prevalence rates of 5 to 10 percent. Some of these studies have reported under recognition of delirium in up to 60 percent of these subjects. At the time of admission to medical wards, between 15 and

20 percent of older patients meet criteria for delirium-prevalent cases. Most of these studies report a subsequent incidence during hospitalization (among patients free of delirium at admission) of 5 to 10 percent, although higher rates (20 to 30 percent) have been reported. Delirium has been reported in 10 to 15 percent of general surgical patients, 30 percent of open heart surgery patients, and over 50 percent of patients treated for hip fractures. Delirium occurs in 70 to 87 percent of those in intensive care units (ICUs) and in up to 83 percent of all patients at the end of life care. Sixty percent of patients in nursing homes or postacute care settings have delirium (Table 10.2–3).

Mortality related to delirium can be as high as rates among patients with acute myocardial infarction or sepsis. Typically it can range between 22 and 76 percent in hospitalized patients. The 1-year mortality rate associated with cases of delirium is 35 to 40 percent.

Risk Factors for Delirium

It is useful to conceptualize risk factors for delirium into two categories: predisposing and precipitating factors (Tables 10.2–4 and 10.2–5). Current management approaches to delirium focus primarily on the precipitating factors and do little to address the predisposing factors. Managing predisposing factors of delirium becomes essential in preventing delirium and decreasing risk of future episodes of delirium and the morbidity/mortality associated with it.

Protective Factors

It is clear that good premorbid functioning before delirium predicts better outcomes. Given the significant morbidity and mortality associated with delirium, interventions that prevent the occurrence of delirium or that foster early recognition have a significant impact on improving patient outcome and reduce the cost of caring for delirious patients. A variety of interventions may improve outcome and reduce risk. For example, educational programs that target clinicians or placement of liaison psychiatrists on orthopedic units improved coordination of care between consultants and primary care providers. In several recent studies in general medical units, targeting education and training of clinical staff demonstrated benefit in terms of decreased frequency or shorter duration of delirium and decreased length of stay for patients compared to the usual care. Other interventions that have been studied and that demonstrate benefit include a focus on nutrition, increased rehabilitation, and attention to visual and hearing impairment. Clearly additional interventional studies are needed concerning the content, duration, and method of educational programs that may modify the frequency, duration, and severity of delirium in various patient cohorts.

Table 10.2–3.
Incidence and Prevalence of Delirium in Various Settings

Population	Prevalence Range (%)	Incidence Range (%)
General medical inpatients	10–30	3–16
Medical and surgical inpatients	5–15	10–55
General surgical inpatients	N/A	9–15 postoperatively
Critical care unit patients	16	16–83
Cardiac surgery inpatients	16–34	7–34
Orthopedic surgery patients	33	18–50
Emergency department	7–10	N/A
Terminally ill cancer patients	23–28	83
Institutionalized elderly	44	33

N/A, not available.

Table 10.2–4.
Predisposing Factors for Delirium

Demographis
- Age 65 and Older
- Male sex

Cognitive status
- Dementia
- Cognitive impairment
- History of delirium

Functional status
- Functional dependence
- Immobility
- History of falls
- Low level of activity

Sensory impairment
- Hearing
- Vision

Decreased oral intake
- Dehydration
- Malnutrition

Drugs
- Treatment with psychotropic drugs
- Treatment with drugs with anticholinergic properties
- Alcohol abuse

Coexisting medical conditions
- Severe medical diseases
- Chronic renal or hepatic disease
- Stroke
- Neurological disease
- Metabolic abnormalities
- Infection with human immunodeficiency virus
- Fractures or trauma

Not all interventions have demonstrated a benefit. For example, a randomized study of general medical patients older than 65 years of age compared usual care to care by a geriatric consultant and a protocol for systematized review of patients' records. The groups, followed for as long as 8 weeks, did not differ in time to improvement from the delirium, and less than half of the patients in both groups demonstrated an increase in the Mini-Mental State Examination (MMSE) of at least two points. Additionally, there were no differences between the groups in measures of delirium severity, length of stay, discharge rate to the community, living arrangements after discharge, or survival.

One study showed little benefit in using haloperidol (Haldol) in older patients (perioperatively) who were undergoing elective hip surgery in preventing development of delirium. Interestingly, the group treated with haloperidol did benefit from decreased morbidity and length of hospitalization. Other studies have reported similar findings in post-coronary artery bypass grafting (CABG) and general orthopedic surgery patients using atypical antipsychotics.

Table 10.2–5.
Precipitating Factors for Delirium

Drugs
- Sedative hypnotics
- Narcotics
- Anticholinergic drugs
- Polypharmacy
- Alcohol or drug withdrawal

Primary neurologic diseases
- Stroke, nondominant hemispheric
- Intracranial bleeding
- Meningitis or encephalitis

Intercurrent illnesses
- Infections
- Iatrogenic complications
- Severe acute illness
- Hypoxia
- Hyponatremia
- Shock
- Anemia
- Fever or hypothermia
- Dehydration
- Poor nutritional status
- Low serum albumin levels
- Metabolic derangements

Surgery

Orthopedic surgery
Cardiac surgery
Prolonged cardiopulmonary bypass
Noncardiac surgery

Environmental

Admission to intensive care unit
Use of physical restraints
Use of bladder catheter
Use of multiple procedures
Pain
Emotional stress
Prolonged sleep deprivation

PATHOPHYSIOLOGY

The pathophysiology of delirium remains poorly understood. Although there has been research to find a final common pathway that explains all delirium, given the heterogeneity of the etiologies and the presentations of delirium, there may not be one mechanism that encompasses the entire syndrome. Disturbance in brain oxygen supply versus demand has been one of the theories proposed for delirium. Impaired oxidative metabolism appears to be a predisposing factor for later development of delirium. EEG studies have demonstrated diffuse slowing of cortical background activity, which does not correlate with underlying causes. Neuropsychologic and neuroimaging studies reveal generalized disruption in higher cortical function, with dysfunction in the prefrontal cortex, subcortical structures, thalamus, basal ganglia, frontal and temporoparietal cortex, fusiform cortex and lingual gyri, particularly on the nondominant side. There are two leading hypotheses that lend some understanding into the complex pathophysiology of delirium. The first focuses on neurochemical imbalances that affect neurotransmission. The second focuses on inflammation, specifically the role of cytokines on the blood–brain barrier and the impact of chronic stress on cytokine and cortisol levels.

Neurotransmission

Patients with delirium have shown evidence of neurochemical changes in acetylcholine, dopamine, glutamate, γ-aminobutyric acid (GABA), and serotonin systems.

Acetylcholine. Extensive evidence supports the role of cholinergic deficiency in delirium as it is involved in rapid eye movement (REM) sleep, attention, arousal, and memory. Administration of anticholinergic drugs can induce delirium in humans and animals, and serum anticholinergic activity is increased in patients with delirium. Physostigmine (Antilirium), a cholinergic agent, reverses delirium associated with anticholinergic drugs. Interestingly, physostigmine has also demonstrated benefit in

nonanticholinergic delirium, such as delirium from alcohol withdrawal, ketamine anesthesia, H_2 receptor antagonist delirium, and γ-hydroxybutyric acid withdrawal. This suggests the possibility of interactions between other systems and the cholinergic system in the pathophysiology of delirium. Recent reports indicate that cholinesterase inhibitors appear to have some benefit even in cases of delirium that are not induced by drugs.

Dopamine. Dopaminergic excess also appears to contribute to delirium, possibly owing to its regulatory influence on the release of acetylcholine. One suggested mechanism is the involvement of dopamine in maintaining and shifting attention. Dopaminergic drugs (e.g., levodopa [Dopar] and bupropion [Wellbutrin]) are recognized precipitants of delirium, and dopamine antagonists (e.g., antipsychotic agents) effectively treat delirium symptoms.

Glutamate. Glutamate, through its excitatory neurotoxicity effects (mediated via the *N*-methyl-D-aspartate [NMDA] receptor), may cause neuronal death and can be associated with delirium. Drugs that are NMDA antagonists, such as ketamine and phencyclidine (PCP), are associated with delirium. One proposed mechanism of Wernicke encephalopathy is through glutamate abnormalities.

γ-Aminobutyric Acid. GABA, an inhibitor of brain activity, has been implicated in contributing to delirium secondary to benzodiazepine and alcohol withdrawal. Hepatic encephalopathy is caused by many factors and has been associated with increased GABA and serum ammonia levels. Ammonia and several other factors are known to induce and aggravate astrocyte swelling that initiate a cascade of events leading to delirium. In addition, elevated ammonia levels can contribute to increased glutamate and glutamine levels, which are precursors to GABA.

Other Neurotransmitters. Perturbations of other neurotransmitters, including norepinephrine and serotonin, may also have a role in the pathophysiology of delirium, but the evidence is less well developed. These neurotransmitters may exert their influence through complex interactions with the cholinergic and dopaminergic pathways.

Inflammation

Another hypothesis of delirium is that it is a CNS response to systemic inflammation during a state of blood–brain barrier compromise.

Cytokines. Cytokines, including interleukin-1 (IL-1), IL-2, IL-6, tumor necrosis factor-α (TNF-α), and interferon, may contribute to delirium by

increasing the permeability of the blood–brain barrier and altering neurotransmission. In postcardiac surgery patients, proinflammatory chemokines have been associated with delirium by disrupting the blood–brain barrier. Similar alterations have been found in patients developing delirium after trauma, primary hyperparathyroidism and delirium tremens.

Chronic Stress. Chronic stress brought on by illness or trauma activates the sympathetic nervous system and hypothalamic–pituitary–adrenocortical axis, resulting in increased cytokine levels and chronic hypercortisolism. Chronic hypercortisolism has deleterious effects on hippocampal serotonin (5-hydroxytryptamine) 5-HT_{1A} receptors, which may contribute to delirium. Given the clinical heterogeneity and multifactorial nature of delirium, it is likely that multiple pathogenic mechanisms contribute to the development of delirium.

ETIOLOGIES

Substance Intoxication

The diagnosis of a substance-induced delirium is made when a medication or other chemical agent is thought to be etiologically related to the condition. The responsible substance may be determined from clinical history, physical examination, and laboratory tests, such as a toxicology screen. The goal of treatment is to remove the offending agent. A familiarity with the common toxicologic syndromes (also known as *toxidromes*) is useful, as it may guide the investigation to include screening for additional agents. Awareness of local prevalence of various drugs of abuse can also give some indication to the etiology of the syndrome. Clinicians need to keep current with drugs of abuse, newly approved drugs, herbal preparations, and teas as potential causes of delirium.

Drugs of Abuse

Intoxication with a variety of drugs of abuse is known to cause delirium. Examples include cocaine, PCP, heroin, alcohol, nitrous oxide, amphetamine and its derivatives (e.g., speed and Ecstasy), and marijuana. The ever-changing choice and availability of recreational drugs obligate the clinician to keep up to date about the agents that are the newest street drugs and the cohorts that use them. Assays to determine these drugs are often not available in routine toxicology screens. Experimentation with drugs of abuse extends to synthetic compounds known as *raves*, which are often used in dance clubs. Drugs such as 3,4-methylenedioxymethamphetamine (MDMA), also known as *Ecstasy*, a methamphetamine analog, have been associated with delirium and fatalities. Special K, or ketamine, an anesthetic agent often used by veterinarians, and PCP (also known as *angel dust*) are NMDA receptor

antagonists and are other examples of street drugs associated with delirium. An example of the symptoms of a substance-induced delirium is described for PCP. At low-to-moderate doses of PCP, respirations may become rapid and shallow, with flushing, diaphoresis, numbness of the extremities, and muscular incoordination. At high doses, blood pressure, pulse rate, and respiration may drop, accompanied by nausea, vomiting, blurred vision, drooling, ataxia, dizziness, seizures, and coma. The patient may have prominent illusions, hallucinations, delusions, paranoia, disordered thinking and catatonia, often with sparse and garbled speech.

Flunitrazepam (Rohypnol) and γ-hydroxybutyrate (GHB) are other common choices of recreational drug users with the potential for delirium. Flunitrazepam is a benzodiazepine that causes anterograde amnesia, disinhibition, and lack of muscle control. It has been associated with violent behavior when combined with alcohol and other drugs of abuse in male juvenile delinquents who may have been experiencing substance-induced delirium. Additionally, the agent has been found in persons who have been raped, probably used by the attacker because of its clinical effects on memory. GHB is a sedative hypnotic used to treat narcolepsy. It is a common cause of stupor, and there is a risk of respiratory suppression when it is combined with other sedative agents. Additionally, withdrawal from GHB presents in a similar manner as alcohol or benzodiazepine withdrawal.

It is important to realize that patients who abuse illicit substances are at increased risk of HIV and hepatitis B and C due to sharing needles or unprotected sexual encounters. These illnesses may increase the risk of or complicate delirium.

Assistance in determining the underlying etiology of delirium due to substance intoxication is available through the local poison control center. These centers have a database that can identify unknown tablets and can describe likely patient responses to intoxication from most prescription and nonprescription agents, herbs, chemicals, and plants. Once the offending agent is identified, appropriate treatment can be initiated and guided through the local poison control center. Established procedures are available for the treatment of overdose of substances such as acetaminophen (Tylenol), aspirin, organic solvents, ethylene glycol, opioids, benzodiazepines, barbiturates, and anticholinergic agents.

Substance Withdrawal–Induced Delirium. Any acute change in mental status can herald a substance withdrawal delirium. The diagnosis may be supported by autonomic changes or the detection of a substance or its metabolite in the urine or serum. When reviewing the history and chart of a patient presenting with delirium, a thorough review of medications and substances should be conducted to include medications that are prescribed and over-the-counter as well as use of illicit substances and any recent

changes. This is especially true when a patient is admitted emergently into the hospital and is unable to give a reliable history.

Alcohol Withdrawal. The classic agitated withdrawal delirium is known as delirium tremens. This syndrome, synonymous with severe alcohol withdrawal, presents with delirium, autonomic hyperactivity, and frequent visual and tactile hallucinations and carries a significant risk of seizures and death if untreated. Patients admitted to the hospital may present several days into admission with alcohol withdrawal and may be at risk of delirium tremens. Histories of complicated alcohol withdrawal, medical comorbidity, and elevated blood pressure are associated with an increased risk of delirium tremens.

Benzodiazepine Withdrawal. Benzodiazepine withdrawal may appear similar to alcohol withdrawal and may also be accompanied by seizures. The onset of symptoms is dependent on the half-life of the particular benzodiazepine. For example, withdrawal from alprazolam (Xanax) that has a short half-life may present in 1 to 2 days, whereas withdrawal from diazepam (Valium) that has a long half-life may present 5 to 7 days after the last ingested dose.

Opiate Withdrawal. Opiate withdrawal presents with severe flu-like syndromes, gastrointestinal (GI) cramping, diarrhea, diaphoresis, autonomic hyperactivity, and craving. Delirium may also occur with opioid withdrawal and has been reported when patients are switched from transdermal fentanyl (Duragesic) to morphine (Duramorph). The presentation may occur because of unknown prior use of opiates in the context of abuse or prescription or methadone (Dolophine) maintenance.

Postoperative

Coronary Artery Bypass Graft. The incidence of delirium after CABG procedure has ranged from 3 to 35 percent. Several factors play a role in the development of delirium in patients with post-CABG.

In one study delirium was detected in 42 (3.07 percent) of 1,367 patients undergoing CABG. Although this rate is low, multiple factors (age greater than 65, peripheral vascular disease, preoperative intra-arterial blood pressure support, Euroscore 5 or more, postoperative blood product usage, and low cardiac output syndrome) independently predicted delirium development after coronary artery bypass procedures. Postoperative delirium was associated with significantly higher mortality rate, prolonged mechanical ventilation time, and increased length of ICU stay.

In one other study, delirium was detected in 74 of 220 (33.6 percent) patients undergoing CABG procedure. Once again multiple risk factors were identified for delirium in this study. Increasing age, blood urea level,

cardiothoracic index, hypertension, smoking habits, blood replacement during bypass, atrial fibrillation (AF), pneumonia, and blood balance in the postoperative period were all associated with delirium. Interestingly, no specific factor associated with the CABG procedure itself (perfusion pressure, number of grafts) was correlated with an increased risk of delirium postoperatively. The length of stay was twice as long in the delirious group.

Hip/Joint Replacement. Delirium is a common adverse event in patients undergoing hip replacement, with estimates in the range of 15 to 60 percent. Several studies have indicated that posthip replacement delirium has been associated with increased morbidity and mortality. In one recent study, the 603 hip surgery patients were stratified in the low-, intermediate-, and high-risk groups for delirium posthip replacement. The incidence of delirium was 3.8 percent in the low-risk group, 11.1 percent in the intermediate-risk group, and 37.1 percent in the high-risk group. In this study, cognitive impairment at admission had the highest predictive value for postoperative delirium, and contrary to previous findings, age was an independent predictive factor for delirium. Moreover, postoperative delirium was four times as frequent in acute patients as in elective hip replacement patients.

DIAGNOSIS

Delirium, despite its common presentation in the elderly, is frequently missed in various clinical settings. Fluctuation in the presentation of symptoms further complicates the identification of delirium. Some studies have indicated that up to two-thirds of delirium cases have been missed by clinicians. Obtaining a good history from the patient's caretaker is an essential step to diagnose delirium. In fact, neglecting this step is the leading reason for missing the diagnosis of delirium. If no history can be obtained, the patient should be assumed to be delirious until proven otherwise. Older patients should be aroused during rounds and evaluated daily for the hypoactive form of delirium, which is more subtle and often overlooked.

Approaches to Diagnosis

When clinicians search for the underlying cause of delirium, they need to be aware of the possibility of occult or atypical presentations of many diseases in the elderly, including myocardial infarction, infection, and respiratory failure as delirium is often the sole manifestation of serious underlying disease. All preadmission and current medications should be reviewed for indication and risk:benefit; even long-standing medications can contribute to delirium and should be reevaluated. If changes in long-term medications are appropriate, the hospital represents the ideal venue

for making these changes. A complete medical history must be meticulously obtained to detect occult alcohol or benzodiazepine use, which can contribute to delirium.

Rating Scales in Delirium

Every older hospitalized patient should undergo brief but formal cognitive testing with the use of instruments such as the MMSE and the Confusion Assessment Method (CAM). CAM focuses on four major aspects on delirium including acute onset, inattention with fluctuations, disorganized thinking, and altered level of consciousness. CAM also looks for the evidence of disorientation, memory impairment, psychomotor changes, altered sleep–wake cycle changes, and perceptual disturbances. CAM has been validated in a number of settings, including the ICU. In elderly, mechanically ventilated patients, the instrument retained high levels of validity when compared to expert clinicians using DSM-IV-TR criteria with 95 percent sensitivity and 88 percent specificity. The administration time is approximately 2 minutes. The instrument detected delirium in 83.3 percent of patients in the ICU, the highest published rate in the ICU setting.

PATHOLOGY AND LABORATORY EXAMINATION

The utility of extensive laboratory testing has not been demonstrated. In many instances, the cause of delirium is obvious and often predictable. In such patients, only such testing is necessary to confirm the diagnosis might be performed. In other situations, delirium may be the initial manifestation (or the only sign) of an underlying illness. The clinician must decide whether the known acute illness is sufficient to cause the delirium, often a difficult decision. Some investigators recommend extensive enough testing to exclude the common causes of delirium. A chest roentgenogram should be considered because pneumonia is one of the most common causes of delirium in elderly patients, and typical symptoms may be absent. An electrocardiogram (ECG) may be helpful in detecting cardiac ischemia. If hypoxia is suspected, pulse oximetry could be done, especially in postoperative patients. Arterial blood gases are most useful in patients with suspected acid–base disturbances, hypercapnia, and severe hypoxia.

Blood Chemistry, Urinalysis, and CSF Examination

Routine testing should be individualized and limited initially to blood chemistry studies that assess for electrolyte, glucose, calcium, renal, and hepatic abnormalities and complete blood cell counts that assess for anemia or leukocytosis (indications of possible infection). Additional tests should be performed as directed by findings from the history and physical examination or if the cause remains elusive after initial assessment. Tests

for thyroid function, analysis of magnesium levels, and syphilis or HIV serologic studies may be considered.

Results of urinalysis that suggest the presence of infection should be followed with cultures and sensitivity studies. Determination of drug levels and toxicology screens may be necessary in selected patients.

A cerebrospinal fluid (CSF) examination is important in the patient with delirium and fever in whom encephalitis or meningitis is suspected. Most hospitalized older patients with fever and delirium, however, have primary causes outside the CNS and do not require routine CSF studies. Infection, dehydration, and metabolic disturbances accounted for 97.5 percent of cases of fever and delirium in one study, which found only one case each of bacterial meningitis and aseptic meningitis.

Electroencephalography and Neuroimaging

When the diagnosis of delirium is in doubt, an EEG may be diagnostic; however, caution must be used as EEG has false-negative rate of 17 percent and false-positive rate of 22 percent. Most patients with delirium will demonstrate diffuse slow-wave or low-voltage activity. The EEG may be useful in differentiating delirium from psychiatric disorders and identifying certain diagnoses, including hepatic encephalopathy, uremia, CNS infections, and seizures. In some patients, the EEG findings may indicate a focal rather than a global impairment and may thus suggest a localized cerebral event.

Neuroimaging studies should be reserved for patients with new focal neurologic signs, those with a history or signs of head trauma, those with fever and acute changes in mental status in whom encephalitis is suspected, or those with no other identifiable cause of the delirium. Neuroimaging should also be considered when the history cannot be obtained or the neurologic examination cannot be completed (e.g., when the patient is combative) so as not to miss uncommon conditions that are life-threatening but treatable, such as subarachnoid hemorrhage and encephalitis. In elderly patients with a history of falling, neuroimaging such as computed tomography (CT) may be important for ruling out subdural hematoma (even if a head injury was not reported). If signs of focal neurologic impairment are present, CT or magnetic resonance imaging (MRI) of the head is indicated. Otherwise, the yield in routine neuroimaging is low, and its use may be limited to patients with atypical manifestations or prolonged symptoms. Functional imaging studies such as blood flow, single-photon emission (SPECT), or positron emission tomography (PET) have yielded important findings about the pathophysiologic features of delirium; however, their role in evaluation of delirium is unclear.

DIFFERENTIAL DIAGNOSIS

The clinical history, physical examination, and laboratory studies are helpful in distinguishing delirium from other causes of global cognitive impairment, such as dementia, depression, and functional psychosis. Dementia has a gradual onset, persists for more than 1 month, is usually progressive, and is not associated with reduced alertness until a terminal stage. Dementia, however, predisposes to delirium, and the sudden deterioration of cognitive or functional ability in a demented patient suggests that a delirium has developed. A delirium that is superimposed on a dementing condition may be difficult to distinguish from atypical dementia, progressive cognitive impairment, or a catastrophic reaction. Considerable overlap exists between the two conditions, and the diagnosis of dementia should be deferred until an appropriate evaluation has been done and several months have passed.

Depression can resemble a hypoactive delirium. Withdrawal, slowed speech, apathetic mood, and poor results on cognitive testing may occur in both conditions. Nevertheless, the clouding of consciousness and the variability present in delirium are absent in patients with depression. Acute psychosis can mimic a hyperactive delirium; however, hallucinations tend to be auditory, delusions are more systematic, and the symptom fluctuation so characteristic of delirium is not noted. EEG findings are usually normal in subjects with psychosis, in comparison with the diffuse slow-wave activity seen in patients with delirium. Similarly, anxiety, a state of profound apprehension with increased motor activity and autonomic arousal, can resemble a hyperactive delirium. Anxiety is common among hospitalized patients, particularly postoperatively and in ICUs, where fears of pain, disability and death can be exacerbated by sensory impairment, sleep deprivation, current pain, and medications. Delirium and anxiety must be distinguished because treatment of one may adversely affect the other. For example, an agitated patient with hepatic encephalopathy can experience worsening if treated incorrectly for anxiety with benzodiazepines.

Diagnosing Delirium in Special Populations

Delirium in Patients with Dementia. Delirium and dementia are highly interrelated, yet the nature of their interrelationship remains poorly examined. Although a cause-and-effect relationship has not been established between delirium and dementia, investigation of their intersection may yield important insights that will advance the current understanding of both conditions. Dementia is the leading risk factor for delirium, and fully two-thirds of cases of delirium occur in patients with dementia. Thus, the underlying vulnerability of the brain in patients with dementia may predispose them to the development of delirium as a result of insults related to acute medical illnesses, medications, or environmental perturbations. Studies also suggest that delirium persists much longer than

previously believed, with symptoms in many patients lasting months or years. The existence of the well-described entities of persistent delirium and reversible dementia blurs the boundaries between these conditions. Moreover, studies have shown that delirium and dementia are both associated with decreased cerebral metabolism, cholinergic deficiency, and inflammation, reflecting their overlapping clinical, metabolic, and cellular mechanisms. In fact, delirium and dementia may represent points along a continuum of cognitive disorders, rather than two entirely separate conditions.

The presence of prior cognitive impairment makes detecting the delirium particularly challenging. One study looked at the effect of severity of prior cognitive impairment on delirium prevalence and symptom presentation among patients who were older and were newly admitted to an acute care hospital. A total of 104 patients were included in this descriptive study and screened for delirium. The results showed that the prevalence of delirium increased according to the severity of the patient's prior cognitive impairment. Except for disorganized thinking, all symptoms of delirium were similar among patients with mild, moderate, and severe prior cognitive impairment. The study concluded that training nurses to recognize subtle changes in mental status among those patients who were older with prior cognitive impairment may prevent the underdetection of delirium.

Does delirium contribute to dementia? Although it is not likely that the delirium itself causes the pathologic changes of dementia, there is no question that delirium contributes to worsening functional status, loss of independence and poorer outcomes among patients with dementia. The long-standing traditional view is that delirium and dementia are two separate conditions; however, emerging evidence has highlighted their overlap. Previously it was thought that treating the underlying reversible cause of delirium brings the patient's functional and cognitive status back to premorbid levels. Emerging bodies of literature indicates otherwise. First, epidemiologic studies have documented long-term cognitive decline in patients with delirium, after controlling for relevant covariates. Second, several causes of delirium may not be completely reversible, particularly those resulting in neuronal injury and permanent cognitive sequelae, such as prolonged hypoxia or hypoglycemia. Third, neuroimaging studies demonstrate regions of hypoperfusion in patients with delirium. Thus, delirium may herald the onset of dementia in many instances. Fourth, dementia with Lewy bodies, which includes fluctuating cognition and visual hallucinations as core signs, illustrates the overlap of delirium and dementia.

Delirium can alter the course of an underlying dementia, with dramatic worsening of the trajectory of cognitive decline, resulting in more rapid progression of functional losses and worse long-term outcomes. This

phenomenon has been well recognized clinically in elderly patients with dementia: Clinicians and family members have noted that the patients "never returned to baseline" after an episode of delirium. In follow-up studies, patients in whom delirium develops have worse outcomes than those with dementia alone, including worsened cognitive function and increased rates of hospitalization, institutionalization, and death. Delirium superimposed on dementia in one study showed a twofold increase in mortality as compared to delirium alone, dementia alone, or with neither dementia nor delirium.

Delirium may serve as an important model for research by offering a unique approach to advance the general understanding of cognitive disorders and dementias. The development of delirium in certain persons may help to identify those who are vulnerable to cognitive decline through genetic predisposition or through the presence of early dementia or mild cognitive impairment that may otherwise remain unidentified. Moreover, a better understanding of the pathogenesis of delirium may help elucidate factors that lead directly to neuronal injury and, thus, to permanent cognitive sequelae. Studies investigating the pathogenesis of delirium with the use of neuropsychologic testing, neuroimaging methods, electrophysiologic methods, laboratory markers, genetic studies, and neuropathologic approaches are greatly needed. Investigation of delirium provides an important opportunity to clarify the link between brain pathophysiology and behavioral manifestations, which might hold broader implications for other cognitive and psychiatric disorders. New prospects for therapy include strategies to increase acetylcholine activity in the brain (e.g., through the use of procholinergic agents and avoidance of highly anticholinergic drugs), the use of selective dopamine antagonists that affect D_1, D_2, D_3, and D_4 receptors differently, and the use of drugs to enhance cerebrovascular flow (e.g., anti-inflammatory or antiplatelet agents). Finally, targeting delirium with new therapeutic approaches may offer opportunities for early intervention, preservation of cognitive-reserve capacity, and prevention of permanent cognitive damage, which may potentially delay or abate the ultimate development of dementia.

Delirium in Patients with Serious Mental Illness. Schizophrenia and bipolar disorder are lifelong illnesses that usually have multiple comorbid conditions. Patients with these diagnoses are on medications that have significant deleterious cognitive effects that place them at high risk of developing delirium. In one study of prevalence of delirium in psychiatric inpatients, the overall incidence of delirium was 14.6 percent. Delirium was most common among patients with schizophrenia and bipolar disorder; patients with bipolar disorder had the highest incidence (35.5 percent). Only 48 percent of delirious patients were actually recognized as having delirium at the time it occurred. Antiparkinsonian medications were the

only medications significantly associated with delirium; lithium (Eskalith) was not an independent risk factor. Older age was the only significant demographic risk variable. The hospital stays of patients with delirium were 62.1 percent longer than those of patients without delirium.

Delirium in Patients with Terminal Illness. Delirium is a common syndrome in terminal patients. Terminal cancer patients commonly suffer from severe pain, are on pain medications, and may have infections leading to delirium. In one recent study, 228 participants (49.9 percent) among 457 terminally ill inpatients were screened. The prevalence of delirium was 46.9 percent (n = 107). Of these, the most common subtype was hypoactive (68.2 percent). The mortality rate of inpatients with delirium (77.6 percent) was higher than those without delirium (50.9 percent). The overall detection rate by any member of the palliative care team was 44.9 percent. The detection rate of the hypoactive subtype was only 20.5 percent, which was significantly lower than that of the hyperactive/mixed subtypes. Therapy for delirium was prescribed in 42.1 percent, with haloperidol being the most common medication.

To date, empirical studies have identified opioids, hypercalcemia, renal failure and dehydration, and chronic subdural hematoma as the underlying etiologies of reversible delirium in terminally ill cancer patients. Many of these conditions can be reversed by appropriate interventions. It has been estimated that approximately 50 percent of cases of delirium occurring in patients receiving specialist palliative care are potentially reversible.

Delirium in Children. The characteristic symptoms and course of delirium are well known in adults, but little is known about delirium in children. One study was undertaken to describe the clinical presentation, symptoms, and outcome of delirium in children and adolescents. Eighty-four patients age 6 months to 18 years were identified with delirium, from 1,027 consecutive psychiatric consultations during a 4-year period. Mortality was high (20 percent), and length of stay was prolonged. Symptoms of psychosis and disorientation were less characteristic, but overall the presentation and course of delirium were similar to adults, and the DSM-IV-TR criteria were found applicable in the pediatric population.

In another study the authors compared pediatric and adult studies and found many of the same symptoms reported, but often at significantly different rates. Sleep–wake disturbance, fluctuating symptoms, impaired attention, irritability, agitation, affective lability, and confusion were more often noted in children; impaired memory, depressed mood, speech disturbance, delusions, and paranoia more often in adults; impaired alertness, apathy, anxiety, disorientation, and hallucination occurrence were similar. These may represent true differences in the presentation of delirium across the life cycle or may be attributable to inconsistent

methodologies. Prospective studies are needed to resolve these questions.

CLINICAL FEATURES

Typically, the onset of delirium is rapid—over a few hours or days—and the symptoms can be highly variable and intermittent. Physicians, nurses and family members may witness completely different behaviors within a brief time span. Clinical features develop rapidly and tend to fluctuate over time. For example, drowsiness, hypervigilance, normal wakefulness and agitation may occur within minutes to hours of one another. Delirious patients can present with manic-like symptoms (poor sleep, pressured speech, high energy, agitation), depression-like symptoms (withdrawn, isolative, poor motivation, low energy), anxiety-like symptoms (anxious, poor sleep, restlessness, agitation), sleep disorder–like symptoms (reversed sleep–wake cycle, insomnia, hypersomnolence), and schizophrenic-like symptoms (hallucinations, delusions, paranoia, thought disorganization). As a result, the early detection diagnosis of delirium may be overlooked. Alternatively, caregivers might recognize that the patient is confused but fail to appreciate the significance of this change in condition. Consequently, the problem may not be addressed until further deterioration ensues. In addition, delirium may mimic or be superimposed on other conditions; thus, its existence may be obscured. Delirium can be superimposed on dementia or other neurologic disorders associated with global cognitive impairment. In such situations, the symptoms of delirium are often mistakenly attributed to the underlying neurologic disorders.

A cardinal feature of delirium is reduced alertness, manifested by difficulty maintaining attention and focusing concentration. The confused patient cannot ignore irrelevant stimuli and is easily distractable. The frequent finding of impaired memory usually represents an inability to register recent information because of inattentiveness, but retrieval of older information can also be disturbed. Higher integrative functions are similarly affected; the result is a reduced ability to plan, solve problems, or guide actions. These disturbances in attention often increase toward evening, a characteristic known as *sundowning*, which may result from fatigue and reduced sensory input. The confused patient cannot plan cognitive functions appropriately. Consequently, subjects demonstrate incoherent speech and a diminished ability to sequence actions properly (e.g., rising from a bed or walking), which can lead to injury. Typically, sleep becomes brief and fragmented, and fatigue tends to aggravate the symptoms of delirium. Subjects often demonstrate a reversal of sleep patterns, with daytime napping and nocturnal wakefulness.

Several clinical subtypes of delirium have been described. Commonly, delirium is classified on the basis of the level of psychomotor activity. Although certain types of delirium may frequently occur in patients with particular disease states, they are neither exclusive to nor diagnostic of

specific underlying medical conditions. Similarly, the manifestation of delirium cannot be fully predicted by the presence of a particular etiologic toxin or illness. The hyperactive (increased psychomotor activity) variant of delirium is most commonly recognized and tends to be readily apparent even to the casual observer. It is often associated with the adverse effects of anticholinergic drugs, drug intoxication and withdrawal states. Characteristically, patients may exhibit agitation, psychosis and mood lability and may refuse to cooperate with medical care, may demonstrate disruptive behaviors (such as shouting or resisting) or may sustain injuries from falling, combativeness, or pulling out catheters and intravascular lines.

The hypoactive (decreased psychomotor activity) variant of delirium is more common than hyperactive delirium in elderly patients, but it is less frequently recognized. Often it may be misattributed to depression or dismissed as a transient, insignificant problem. Metabolic conditions such as hepatic or renal encephalopathy are frequently associated with hypoactive delirium, but other causes are also common. Patients with hypoactive delirium may appear sluggish and lethargic as well as confused. The absence of disruptive, bizarre, and injurious behaviors contributes to the lack of recognition. This quiet state of withdrawal, apathy, and clouded inattention differs from the drowsiness seen in persons dozing or sleeping in that these latter subjects can be quickly aroused to (and remain in) normal consciousness by mild stimuli. In hypoactive delirium, however, strong stimuli (e.g., vigorous shaking or shouting) are often needed for arousal, which is incomplete and transient at best.

Because of the multiple etiologic factors, the fluctuating course, and the individual medical comorbidities, many patients who experience delirium have a mixture of both hypoactive and hyperactive variants. Such patients are thought by some researchers to be at highest risk of substantial morbidity and mortality. In one recent study conducted in critically ill patients, mixed type was most common (54.9 percent), followed by hypoactive delirium (43.5 percent), and purely hyperactive delirium (1.6 percent). In a posthip surgery population, 47 percent were hyperactive type of delirium, and both hypoactive and mixed type were each seen in 26 percent of the group. Similar rates were reported in a patient group admitted to an inpatient geriatric ward.

Hallucinations, misperceptions, illusions, and delusions are reported to occur in at least 40 percent of cases of delirium and can accompany both hypoactive and hyperactive subtypes. Hallucinations are usually visual, ranging from dream-like experiences to terrifying visions, and may, in fact, represent “sleep without muscle atonia.” Auditory hallucinations, as well as those involving taste and smell, also occur, although much less frequently. Delusions are often paranoid or persecutory (e.g., suspicion of poisoned food or fear of intended harm by nursing staff).

Although these are some of the most distressing symptoms, the presence or absence of perceptual disturbances is not helpful in determining the cause or level of severity of delirium. Their presence, however, does confirm the diagnosis of an acute delirious state. Of importance, these sensory features tend to be less common in elderly than in young patients. Rather, somatic features such as urinary incontinence, gait impairment, tremor, and language disorders (including receptive and expressive aphasias) tend to predominate.

Focal neurologic signs can also occur but usually suggest an underlying cerebrovascular event. Although focal strokes are seldom considered among the causes of confusional states, localized insults in structures that subserve attention may result in delirium. Focal infarcts in the inferior temporo-occipital region or the parietal, prefrontal, or subcortical region of the right hemisphere can cause a diffuse attention disorder.

COURSE AND PROGNOSIS

Symptoms of delirium usually last about 7 days in most patients, but there is much variability, with slow resolution of symptoms contributing to persistent symptoms of delirium at 6 to 8 weeks for severely ill patients. Symptom resolution is frequently incomplete by hospital discharge, with as many as 15 percent of patients remaining symptomatic of delirium at 6 months. There have been numerous methodologic limitations of some studies examining long-term follow-up. The presence of dementia and the level of medical comorbidity were not always controlled for, and the sample sizes were frequently small. In general, studies suggest that the increased mortality risk associated with delirium was maintained at 12, 24, and 36 months, with a risk ratio of at least 2 at all time points. Additionally, at 24 months, the increased risk of cognitive and functional impairment remained. One study followed elderly patients discharged from an emergency room who were delirious. The mortality outcome at 6 months postdischarge for delirious patients not identified by the emergency room treating team was three times higher than the delirious patients who were identified and treated in the emergency room. The mortality rate was similar for the treated delirious patients as the nondelirious patients.

Early recognition of delirium is hypothesized to improve prognosis by allowing clinicians to address underlying medical conditions and substances earlier in the course. Surprisingly, preventive measures to improve prognosis have demonstrated mixed results. This may be due to insufficient exposure to staff, poor retention of information presented, or other patient and clinician factors. The outcome measure used in these trials included time to delirium resolution, increase in independence in activities of daily living, decreased length of stay, and increased rate of discharge to the community rather than institutional settings. For example, one positive study demonstrated that the presence of a liaison psychiatrist

on an orthopedic unit reduced the frequency of delirium and decreased the length of stay for the patients in the intervention group. Another study, which looked at whether a geriatric consultation followed by involvement of a liaison nurse with the consultant and primary team, failed to demonstrate that the intervention had any benefit. Educational intervention to physicians and nurses by means of distribution of guidelines for recognizing delirium and formal educational in-services about the guidelines failed to improve outcomes in acute care hospitals in England.

Recent reports of multidisciplinary educational interventions have shown some benefit in decreasing the new onset of delirium and the length of stay in delirious patients in medical units. In other report, in an academic medical center in the United States, a multicomponent interventional strategy failed to reduce the rate of delirium, placement into nursing homes, need for home care visitation, readmission rates into the hospital, and self-rated health measures, and failed to improve functional status. Adherence to nonpharmacologic interventions for delirium was directly related to patient outcome in this multicomponent intervention in the United States.

PREVENTION

Preventing delirium is the most effective strategy for reducing its frequency and complications. Successful preventive strategies include multicomponent approaches to reduce risk factors. Because delirium has many causes, multicomponent approaches represent the most effective and clinically relevant ones. The Yale Delirium Prevention Trial demonstrated the effectiveness of intervention protocols targeted toward six risk factors: orientation and therapeutic activities for cognitive impairment, early mobilization to avert immobilization, nonpharmacologic approaches to minimize the use of psychoactive drugs, interventions to prevent sleep deprivation, communication methods and adaptive equipment (particularly eyeglasses and hearing aids) for vision and hearing impairment, and early intervention for volume depletion. A randomized clinical trial involving patients who had had hip fractures demonstrated the effectiveness of a multicomponent strategy for geriatric consultation targeted toward 10 domains: oxygen delivery to the brain, fluid and electrolyte balance, pain management, reduction in the use of psychoactive drugs, bowel and bladder function, nutrition, early mobilization, prevention of postoperative complications, appropriate environmental stimuli, and treatment of symptoms of delirium.

TREATMENT

Nonpharmacologic Management

Nonpharmacologic interventions should be first line in the management of a patient with suspected delirium. Such interventions include modifying the patient's surroundings to maximize the safety and calmness of the environment to provide reassurance and decrease fear and agitation associated with delirium. Environmental changes may also enhance the patient's ability to interpret the surroundings appropriately; for example, patients with eyeglasses and hearing aids should have these devices when appropriate. Paper and pencil should be provided to improve communication when the patient cannot respond verbally because of intubation. Rooms with a window view will provide orientation. Other measures include avoiding extremes of sensory input; for example, stimulus overload in a loud and busy hospital unit or stimulus deprivation in an isolated, dark, windowless room. Because symptoms of delirium can easily be aggravated by placing two affected patients in the same hospital room, this arrangement should be avoided. The patient should be near a nursing station to facilitate close observation.

Maintenance of patient comfort is also important in managing delirium; however, adequate control of pain may necessitate the use of medications that can worsen delirium. Determining the appropriate level of analgesia in a patient with delirium is sometimes difficult. Physical activity should be initiated as soon as possible in order to minimize the adverse effects of immobility (e.g., pressure sores), avoid deconditioning and enhance orientation. Normal sleep–wake cycles can be promoted by the use of daytime activity and environmental cues (such as windows and clocks). Interruptions of sleep should be minimized when possible. Adequate nutrition is important for recovery from the delirium and the underlying illness, as well as for providing useful orientation.

Psychosocial support can be provided by both staff and family or friends. Continuity in both nursing and medical staff can help decrease patient fears and paranoia because of unfamiliar faces. Modification of staff behavior around patients with delirium can be helpful; for example, one report described exacerbation of delirium in patients when staff personnel wore holiday costumes. Similarly, staff conversations within earshot of a patient with delirium can easily be misinterpreted and produce a worsening of the patient's symptoms. In addition to the care provided by the staff, family or close friends can be extremely important in the management of the patient's symptoms. If regularly informed of the situation, family members can reassure the patient, provide reorientation, and reduce anxiety and agitation. Familiar items from home may be useful for selected patients. Staff or family members may be available to provide constant observation, which can minimize the need for physical restraints. Even when restraints are used properly, patients can fall or be injured. When appropriate, the provision of these measures by staff, family, or friends can minimize patient symptoms and perhaps decrease the cost of

care.

Pharmacologic Management

Strong evidence to support the pharmacologic management of delirium is lacking. Thus, the use of psychoactive medications should be reserved for the management of behaviors associated with delirium that pose a safety risk for the patient and others or for delirium due to drug or alcohol withdrawal. Few controlled trials have been conducted to determine which agents are most beneficial in managing the symptoms of delirium; however, two general classes of medications—antipsychotics and benzodiazepines—are most frequently used.

For the agitated patient with perceptual disturbances that are associated with sleep–wake cycle abnormalities and behavioral dyscontrol, antipsychotic agents can be beneficial. Haloperidol is commonly used to treat delirium despite weak evidence to support its efficacy and lack of U.S. Food and Drug Administration (FDA) approval for this indication. This potent agent may be preferred over other antipsychotics because it has fewer anticholinergic and hypotensive side effects. However, the increased potency of haloperidol is associated with increased frequency of extrapyramidal side effects including dystonic reactions, akathisia, tardive dyskinesia, and malignant catatonia. Therefore, the patient should be monitored closely not only for effectiveness but also for adverse events, especially in a patient with Parkinson disease who has delirium.

Haloperidol is commonly administered orally and intramuscularly in various clinical settings. Use of intravenous haloperidol (which also not FDA approved for treatment of delirium or any other indication) is most commonly seen in critical care settings and is associated with increased risk of QTc prolongation. ECG monitoring is recommended when haloperidol is administered via the intravenous route. Clinical observations have noted a possible decrease in the serious side effects in comparison with the oral and intramuscular routes. Other older antipsychotic drugs including thioridazine (Mellaril), droperidol (Inapsine), and chlorpromazine (Thorazine) have been compared with haloperidol and found to be equally effective. However, because of a higher incidence of anticholinergic side effects (which may aggravate delirium), haloperidol is preferred.

Table 10.2–6.
Memorial Delirium Rating Scale (MDRS)

INSTRUCTIONS: Rate the severity of the following symptoms of delirium based on current interaction with subject or assessment of his/her behavior or experience over past several hours (as indicated in each time).

Item 1: *Reduced level of consciousness (awareness):* Rate the patient's current awareness of and interaction with the environment (interviewer, other people/objects in the room; for example, ask patients to describe their surroundings).

0: None (patient spontaneously fully aware of environment and interacts appropriately)

1: Mild (patient is unaware of some elements in the environment, or not spontaneously interacting appropriately with the interviewer; becomes fully aware and appropriately interactive when prodded strongly; interview is prolonged but not seriously disrupted)

2: Moderate (patient is unaware of some or all elements in the environment, or not spontaneously interacting with the interviewer; becomes incompletely aware and inappropriately interactive when prodded strongly; interview is prolonged but not seriously disrupted)

3: Severe (patient is unaware of all elements in the environment with no spontaneous interaction or awareness of the interviewer, so that the interview is difficult-to-impossible, even with maximal prodding)

Item 2: *Disorientation*: Rate current state by asking the following 10 orientation items: date, month, day, year, season, floor, name of hospital, city, state, and country.

0: None (patient knows 9–10 items)

1: Mild (patient knows 7–8 items)

2: Moderate (patient knows 5–6 items)

3: Severe (patient knows no more than 4 items)

Item 3: *Short-term memory impairment:* Rate current state by using repetition and delayed recall of 3 words (patient must immediately repeat and recall words 5 min later after an interviewing task. Use alternate sets of 3 words for successive evaluations [e.g., apple, table, tomorrow, sky, cigar, justice]).

0: None (all 3 words repeated and recalled)

1: Mild (all 3 repeated, patient fails to recall 1)

2: Moderate (all 3 repeated, patient fails to recall 2–3)

3: Severe (patient fails to repeat 1 or more words)

Item 4*: Impaired digit span:* Rate current performance by asking subjects to repeat 3, 4, then 5 digits forward and then 3, then 4 backwards; continue to the next step only if patient succeeds at the previous one.

0: None (patient can do at least 5 numbers forward and 4 backward)

1: Mild (patient can do at least 5 numbers forward, 3 backward)

2: Moderate (patient can do 4–5 numbers forward, cannot do 3 backward)

3: Severe (patient can do no more than 3 numbers forward)

Item *5: Reduced ability to maintain and shift attention*: As indicated during the interview by questions needing to be rephrased and/or repeated because patient's attention wanders, patient loses track, patient is distracted by outside stimuli or over-absorbed in a task.

0: None (none of the above; patient maintains and shifts attention normally)

1: Mild (above attentional problems occur once or twice without prolonging the interview)

2: Moderate (above attentional problems occur often, prolonging the interview without seriously disrupting it)

3: Severe (above attentional problems occur constantly, disrupting and making the interview difficult-to-impossible)

Item *6: Disorganized thinking*: As indicated during the interview by rambling, irrelevant, or incoherent speech, or by tangential, circumstantial, or faulty reasoning. Ask patient a somewhat complex question (e.g., "Describe your current medical condition.")

0: None (patient's speech is coherent and goal-directed)

1: Mild (patient's speech is slightly difficult to follow; responses to questions are slightly off target but not so much as to prolong the interview)

2: Moderate (disorganized thoughts or speech are clearly present, such that interview is prolonged, but not disrupted)

3: Severe (examination is very difficult or impossible due to disorganized thinking or speech)

Several recent studies have evaluated the use of atypical antipsychotics

in the management of delirium and have shown some promise in terms of both efficacy and safety. Similar to low-dose haloperidol, evidence supports the use of low-dose risperidone (Risperdal), olanzapine (Zyprexa), and quetiapine (Seroquel) in controlling the aggression associated with delirium. While these agents are now available as generics, they are more expensive than haloperidol and are not available for intravenous use. Olanzapine is available for intramuscular use at this time. Although these new medications are relatively less likely to cause extrapyramidal side effects, they are not without adverse effects. The FDA has issued a safety class warning regarding the use of all antipsychotics in patients with dementia as they are associated with higher incidence of mortality.

Benzodiazepines have also been used historically in the management of delirium to sedate the agitated patient. In light of evidence to suggest that benzodiazepines can increase both the risk and duration of delirium, particularly in the elderly, the use of benzodiazepines should be reserved for the management of agitation associated with sedative–hypnotic withdrawal (e.g., alcohol, benzodiazepines, barbiturates, meprobamate, and carisoprodol). Patients experiencing alcohol withdrawal should be given high dose thiamine via the intramuscular or intravenous route as oral absorption is very poor to prevent progression to Korsakoff syndrome (Table 10.2–6).

Current evidence suggests that dexmedetomidine (Precedex) may be effective in preventing and treating delirium in the ICU setting; however, larger, well-designed trials are warranted to more clearly define the place in therapy for dexmedetomidine.

Electroconvulsive Therapy. Electroconvulsive therapy (ECT) is also a treatment of delirium when other approaches have failed. It has been used as a last resort for delirious patients with severe agitation who are not responsive to pharmacotherapy, such as high doses of intravenous haloperidol. The ECT is usually given en bloc or daily for several days, sometimes with multiple treatments per day. ECT use in delirium should be monitored closely as ECT can cause delirium on its own.

Sleep–Wake Cycle. Delirium is frequently complicated by changes in the sleep–wake cycle. Attempts to restore sleep integrity may include moving the schedule of existing sedating medications to the hour of sleep or reducing or moving activating medications and stimulants such as caffeine to the morning. Brief, judicious use of sedating agents, such as zolpidem (Ambien) or trazodone (Desyrel), to reset the sleep–wake cycle may be appropriate. Care should be taken to avoid excess sedation because of risk of falls, aspiration, and inability to perform or assist with activities of daily living.

Table 10.2–7.
Delirium Rating Scale (DRS)

Item 1: *Temporal onset of symptoms*

This item addresses the time course over which symptoms appear; the maximum rating is for the most abrupt onset of symptoms—a common pattern for delirium. Dementia is usually more gradual in onset. Other psychiatric disorders, such as affective disorders, might be scored with 1 or 2 points on this item. Sometimes delirium can be chronic (e.g., in geriatric nursing home patients), and unfortunately only 1 or 2 points would be assessed in that situation.

0 No significant change from long-standing behavior, essentially a chronic or chronic-recurrent disorder

1 Gradual onset of symptoms, occurring within a 6 month period

2 Acute change in behavior or personality occurring over a month

3 Abrupt change in behavior, usually occurring over a 1 to 3 day period

Item 2: *Perceptual Disturbances*

This item rates most highly the extreme inability to perceive differences between internal and external reality, while intermittent misperceptions such as illusions are given 2 points. Depersonalization and derealization can be seen in other organic mental disorders like temporal lobe epilepsy, in severe depression, and in borderline personality disorder and thus are given only 1 point.

0 None evident by history or observation

1 Feelings of depersonalization or derealization

2 Visual illusions or misperceptions including macropsia, micropsia; for example, may urinate in wastebasket or mistake bedclothes for something else

3 Evidence that the patient is markedly confused about external reality; for example, not discriminating between dreams and reality

Item 3:

The presence of any type of hallucination is rated. Auditory hallucinations alone are rated with less weight because of their common occurrence in primary psychiatric disorders. Visual hallucinations are generally associated with organic mental syndromes, although not exclusively, and are given 2 points. Tactile hallucinations are classically described in delirium, particularly due to anticholinergic toxicity, and are given the most points.

0 Hallucinations not present

1 Auditory hallucinations only

2 Visual hallucinations present by patient's history or inferred by observation, with or without auditory hallucinations

3 Tactile, olfactory or gustatory hallucinations present with or without visual or auditory hallucinations

Treatment of Specific Etiologies of Delirium

Anticholinergic Intoxication. Anticholinergic medications are associated with delirium. There may be a cumulative effect of the anticholinergic load from multiple medications. Anticholinergic poisoning almost always results in delirium and is often accompanied by physical agitation and visual hallucinations. Physical signs resulting from antimuscarinic action include widely dilated, poorly reactive pupils; warm, dry skin; dry mouth; fever; tachycardia; elevated blood pressure; constipation; and urinary retention. Use of cholinesterase inhibitors, such as physostigmine, has been shown to reduce the severity of the delirium but requires repeated dosing because of a short half-life. In a comparison study of physostigmine and benzodiazepines for anticholinergic poisoning, the physostigmine was

effective, whereas the benzodiazepine was ineffective and was associated with a higher complication rate. Anticholinergic intoxication also requires cardiac monitoring.

The cholinesterase inhibitors that are FDA approved for Alzheimer disease (donepezil [Aricept], rivastigmine [Exelon], and galantamine [Reminyl]) have not been well studied, are not without safety concerns and thus, are not recommended as a therapeutic modality in the management of delirium.

Wernicke Encephalopathy. Wernicke encephalopathy is an acute, neuropsychiatric syndrome that is common relative to other neurologic disorders. It is characterized by nystagmus and ophthalmoplegia, mental-status changes, and unsteadiness of stance and gait, although this triad is seen in only 16 percent of patients. The disorder results from a deficiency in vitamin B_1 (thiamine), which in its biologically active form, thiamine pyrophosphate, is an essential coenzyme in several biochemical pathways in the brain. Wernicke encephalopathy is a medical emergency, and in patients in whom the disorder is suspected, thiamine should be initiated immediately, either intravenously or intramuscularly, to ensure adequate absorption.

Substance Intoxication. Substance intoxication is the first essential step in treating substance abuse–related delirium in the cessation of that substance. The primary concern in managing a patient with substance intoxication is supportive, that is, ensuring that the patient does not have significant respiratory depression and that there are no cardiovascular abnormalities. For benzodiazepine ingestions, the benzodiazepine receptor antagonist flumazenil (Romazicon) may be helpful; however, repeated administrations might be required. In mixed ingestions that include benzodiazepines and tricyclic antidepressants or carbamazepine (Tegretol), flumazenil may precipitate cardiac dysrhythmias or seizures. The usefulness of flumazenil in patients with hepatic encephalopathy and alcohol overdose is questionable. Opiate intoxication may cause respiratory suppression in addition to delirium. Reversal with naloxone (Narcan) or naltrexone (ReVia) may be considered with adequate support and monitoring of the cardiovascular and respiratory status. Naloxone acts through competitive binding at opioid receptors, can reverse all the receptor-mediated actions of opioids, and is indicated for patients who have significant CNS or respiratory depression. Naloxone must be administered intermuscularly or intravenously, has a short half-life, and may require repeated administration as the effects wear off. Naloxone administration can precipitate acute withdrawal in chronic opioid users. There are many other agents that are used in the context of emergency care as antidotes to substances **ingested.** (Please refer to the detailed

description of these issues in the Substance Use Disorders section.)

Table 10.2–8.
Pharmacological Treatment

Pharmacological Agent	Dosage	Side Effects	Comments
Typical antipsychotic			
Haloperidol (Haldol)	0.5–1 mg po twice a day (may be given every 4–6 hours as needed too)	Extrapyramidal side (EPS) effects Prolonged QTc	Most commonly used (can be given intramuscularly)
Atypical antipsychotics		All can prolong QTc duration	
Risperidone (Risperdal)	0.5–1 mg a day	EPS concerns	Limited data in delirium
Olanzapine (Zyprexa)	5–10 mg a day	Metabolic syndrome	Higher mortality in patients with dementia
Quetiapine (Seroquel)	25–150 mg a day	More sedating	
Benzodiazepine			
Lorazepam (Ativan)	0.5–3 mg a day and as needed every 4 hours	Respiratory depression, paradoxical agitation	Best use in delirium secondary to alcohol/benzodiazepine withdrawal (can worsen delirium)

Substance Withdrawal. Benzodiazepines are used widely for the treatment of alcohol withdrawal, with the goals of reducing the severity of withdrawal, preventing delirium, and reducing the incidence of seizures. A systematic review identified 57 relevant randomized controlled trials with a total of 4,275 participants. The quality of the included studies, many of which date from the 1970s and 1980s, was fair, and most were small. Only nine studies clearly concealed allocation of patients to treatment or control groups, and only 21 adequately described the randomization procedures, although all but three were double-blinded.

Not surprisingly, the authors found that benzodiazepines were much more effective than placebo at preventing alcohol withdrawal seizures (relative risk 0.16; 95 percent confidence interval, 0.04 to 0.69). However, there was no significant difference in prevention of seizure between benzodiazepines and antiseizure drugs. There also was no difference in control of symptoms (as measured by standard scales) among benzodiazepines and other drugs such as clonidine (Catapres) or carbamazepine. There were trends in favor of benzodiazepines, particularly longer-acting drugs, for prevention of delirium. Thirteen studies, with a total of 571 patients, compared different benzodiazepines but found no differences in effectiveness among them. Only three studies, with a total of

262 patients, compared fixed-dose with symptom-triggered dosing schedules, and no clear conclusions could be drawn.

An evidence-based guideline from the American Society of Addiction Medicine recommends benzodiazepines as a first-line agent for the treatment of alcohol withdrawal. The guideline notes that although agents with a longer duration of action may provide fewer breakthrough symptoms, those with a shorter duration of action, such as lorazepam (Ativan), may be preferred when there is concern about prolonged sedation (e.g., in patients with significant comorbidities or liver disease).

Treatment in Special Populations

Parkinson Disease. Patients with Parkinson disease are predisposed to delirium. In Parkinson disease, the antiparkinsonian agents are frequently implicated in causing a delirium. In delirious patients with Parkinson disease, an intensive evaluation of causes of the delirium, other than antiparkinsonian medications, is necessary. Decreasing the dosage of the antiparkinsonian agent has to be weighed against a worsening of motor symptoms. If the antiparkinsonian agents cannot be further reduced, or if the delirium persists after attenuation of the antiparkinsonian agents, clozapine is recommended as it is the best studied of the second-generation antipsychotic medications. If a patient is not able to tolerate clozapine or the required blood monitoring, alternative antipsychotic agents should be considered. Quetiapine has not been as rigorously studied as clozapine and may have parkinsonian side effects, but it is used in clinical practice to treat psychosis in Parkinson disease.

Terminally Ill Patients. Not all causes of delirium are reversible, and realistic treatment expectations should be set after a discussion with the patient and caregivers. When delirium occurs in the context of a terminal illness, issues about advanced directives and the existence of a health care proxy become more significant. This scenario emphasizes the importance of early development of advance directives for health care decision making while a person has the capacity to communicate the wishes regarding the extent of aggressive diagnostic tests at life's end. The focus may change from an aggressive search for the etiology of the delirium to one of palliation, comfort, and assistance with dying.

References

Barr J, Fraser GL, Puntillo K, et al. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. *Crit Care Med.* 2013;41(1):263–306.

Bellelli G, Frisoni GB, Turco R, Lucchi E, Magnifico F. Delirium superimposed on dementia predicts 12-month survival in elderly patients discharged from a postacute rehabilitation

facility. *J Gerontol A Biol Sci Med Sci*. 2007;62(11):1306–1309.

Brown TM. Basic mechanisms in the pathogenesis of delirium. In: Stoudemire A, Fogel BS, Greenberg DB, eds. *The Psychiatric Care of the Medical Patient*. 2nd ed. New York: Oxford University Press; 2000.

Camus V, Gonthier R, Dubos G, Schwed P, Simeone I. Etiologic and outcome profiles in hypoactive and hyperactive subtypes of delirium. *J Geriatr Psychiatry Neurol*. 2000;13(1):38–42.

Cole M, McCusker J, Dendukuri N, Han L. The prognostic significance of subsyndromal delirium in elderly medical inpatients. *J Am Geriatr Soc*. 2003;51:754–760.

Collinsworth AW, Priest EL, Campbell CR, Vasilevskis EE, Masica AL. A review of multifaceted care approaches for the prevention and mitigation of delirium in intensive care units. *J Intensive Care Med*. 2016;31(2):127–141.

Fang CK, Chen HW, Liu SI, Lin CJ, Tsai LY. Prevalence, detection and treatment of delirium in terminal cancer inpatients: a prospective survey. *Jpn J Clin Oncol*. 2008;38(1):56–63.

Fick D, Foreman M. Consequences of not recognizing delirium superimposed on dementia in hospitalized elderly individuals. *J Gerontol Nurs*. 2000;26:30–40.

Gagnon P, Allard P, Mâsse B, DeSerres M. Delirium in terminal cancer: a prospective study using daily screening, early diagnosis, and continuous monitoring. *J Pain Symptom Manage*. 2000;19(6):412–426.

Häussinger D, Schliess F, Kircheis G. Pathogenesis of hepatic encephalopathy. *J Gastroenterol Hepatol*. 2002;17(Suppl 3):S256–S259.

Inouye, SK, RDG Westendorp, Saczynski JS. Delirium in elderly people. *Lancet*. 2014;383(9920):911–922.

Jackson JC, Gordon SM, Hart RP, Hopkins RO, Ely EW. The association between delirium and cognitive decline: a review of the empirical literature. *Neuropsychol Rev*. 2004;14:87–89.

Kalisvaart KJ, de Jonghe JF, Bogaards MJ, Vreeswijk R, Egberts TC. Haloperidol prophylaxis for elderly hip-surgery patients at risk for delirium: a randomized placebo-controlled study. *J Am Geriatr Soc*. 2005;53:1658–1666.

Lawlor PG, Gagnon B, Mancini IL, Pereira JL, Hanson J. Occurrence, causes, and outcome of delirium in patients with advanced cancer: a prospective study. *Arch Intern Med*. 2000;160(6):786–794.

Levkoff SE, Marcantonio ER. Delirium: a major diagnostic and therapeutic challenge for clinicians caring for the elderly. *Compr Ther*. 1994;20:550–557.

Liptzin B, Levkoff SE. An empirical study of delirium subtypes. *Br J Psychiatry*. 1992;161:843–1845.

Marcantonio ER, Flacker JM, Michaels M, Resnick NM. Delirium is independently associated with poor functional recovery after hip fracture. *J Am Geriatr Soc*. 2000;48:618–624.

McCusker J, Cole M, Dendukuri N, Belzile E, Primeau F. Delirium in older medical inpatients and subsequent cognitive and functional status: a prospective study. *Can Med Assoc J*. 2001;165:575–583.

McCusker J, Cole M, Dendukuri N, Han L, Belzile E. The course of delirium in older medical inpatients: a prospective study. *J Gen Intern Med*. 2003;18:696–704.

Naughton BJ, Saltzman S, Ramadan F, Chadha N, Priore R. A multifactorial intervention to

reduce prevalence of delirium and shorten hospital length of stay. *J Am Geriatr Soc*. 2005;53(1):18–23.

Peterson JF, Pun BT, Dittus RS, Thomason JW, Jackson JC. Delirium and its motoric subtypes: a study of 614 critically ill patients. *J Am Geriatr Soc*. 2006;54(3):479–484.

Ritchie J, Steiner W, Abrahamowicz M. Incidence of and risk factors for delirium among psychiatric inpatients. *Psychiatr Serv*. 1996;47(7):727–730.

Rockwood K. The occurrence and duration of symptoms in elderly patients with delirium. *J Gerontol Med Sci*. 1993;48:M162–M166.

Rockwood K, Cosway S, Carver D, Jarrett P, Stadnyk K. The risk of dementia and death after delirium. *Age Ageing*. 1999;28:551–556.

Rockwood K, Cosway S, Stolee P, Kydd D, Carver D. Increasing the recognition of delirium in elderly patients. *J Am Geriatr Soc*. 1994;42:252–256.

Rudolph JL, Ramlawi B, Kuchel GA, McElhaney JE, Xie D. Chemokines are associated with delirium after cardiac surgery. *J Gerontol A Biol Sci Med Sci*. 2008;63(2):184–189.

Rummans TA, Evans JM, Krahn LE, Fleming KC. Delirium in elderly patients: evaluation and management. *Mayo Clin Proc*. 1995;70:989–998.

Santana Santos F, Wahlund LO, Varli F, Tadeu Velasco I, Eriksdotter Jonhagen M. Incidence, clinical features and subtypes of delirium in elderly patients treated for hip fractures. *Dement Geriatr Cogn Disord*. 2005;20(4):231–237.

Seaman JS, Schillerstrom J, Carroll D, Brown TM. Impaired oxidative metabolism precipitates delirium: a study of 101 ICU patients. *Psychosomatics*. 2006;47:56–61.

Tabet N, Hudson S, Sweeney V, Sauer J, Bryant C. An educational intervention can prevent delirium on acute medical wards. *Age Ageing*. 2005;34:152–156.

Voyer P, Cole MG, McCusker J, Belzile E. Prevalence and symptoms of delirium superimposed on dementia. *Clin Nurs Res*. 2006;15(1):46–66.

Yokota H, Ogawa S, Kurokawa A, Yamamoto Y. Regional cerebral blood flow in delirium patients. *Psychiatry Clin Neurosci*. 2003;57:337–339.

▲ 10.3 Dementia

JULIE A. GRAZIANE, M.D., AND ROBERT A. SWEET, M.D.

This chapter will provide an overview of dementia and review features in common with all dementias. The most common subtypes of dementia will then be presented and specific features will be discussed in detail. This chapter will discuss neurodegenerative dementias not caused by general medical conditions or substances: Alzheimer disease, frontotemporal dementia (FTD), dementia with Lewy bodies (DLB), dementia due to Parkinson disease (PDD), corticobasal syndrome, progressive supranuclear palsy (PSP), dementia due to Huntington disease, and the prion diseases. Please refer to Chapter 10.5 for a discussion of vascular dementia, dementia in human immunodeficiency virus (HIV), and substance-induced dementia.

DEFINITION

Dementia refers to a disease process marked by progressive cognitive impairment in clear consciousness. The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) describes dementia, now subsumed under the term major neurocognitive disorder, as significant cognitive impairment in one or more of the domains of complex attention, executive function, learning and memory, language, perceptual motor ability, and social cognition. These deficits represent a decline from a previous level of functioning, so dementia does not refer to low intellectual functioning or mental retardation that are developmental and static conditions. This decline is often noted by the individual affected, a family member or other caretaker, or the clinician, and should be demonstrated on standardized neuropsychological testing, and if that is not possible, another measurable clinical assessment. These cognitive deficits interfere with independent functioning in daily activities, and cannot occur exclusively in the context of a delirium, or be better explained by another mental disorder. In DSM-5, major neurocognitive disorder is distinguished from mild neurocognitive disorder by the amount of cognitive decline. Mild neurocognitive disorder requires modest cognitive decline, as well as a lack of interference in activities required for independent living, although greater effort or compensatory mechanisms may be necessary to maintain this independence (see Chapter 10.4, Amnestic Disorders and Mild Cognitive Impairment). The etiologies of major neurocognitive disorders are diverse, and include degenerative diseases, general medical illnesses and infection, those that are substance induced, and combinations of the above.

HISTORY

The word dementia derives from the Latin word *dementatus*, meaning out of one's mind, and, as such, was potentially applicable to any state of psychopathology. Celsus probably first used the term *dementia* in the 1st century AD, although one of the first attempts to describe an etiology beyond old age was in the 4th century AD by Oribasius, physician to the Emperor Julian. Oribasius wrote of a disease of cerebral atrophy that caused loss of intellectual capacity and weakness of movement. It was not until the 19th century that the distinction between cognitive impairment due to dementia was separated from that caused by mental illness by Jean Etienne Dominique Esquirol, as described in his classic work *Mental Maladies: A Treatise on Insanity*. Esquirol identified three varieties of dementia: Acute, chronic, and senile. Senile dementia was defined as a "cerebral affection . . . characterized by a weakening of the sensibility, understanding and will" and also by marked impairments in memory, reasoning, and attention. Esquirol also described hallucinations, delusions,

aggressive behavior, and motor impairments in patients with dementia.

In 1845, Wilhelm Griesinger was the first to describe senile dementia as a disease of the cerebral arteries. However, it was one of Griesinger's students, Emil Kraeplin, who later narrowed the scope of dementia by differentiating senile dementia from psychoses with cerebral arteriosclerosis, which later came to be known as dementia praecox and, finally, as schizophrenia. Also in the late 19th century, reports of plaques and atrophy in the brains of patients with memory deficits and mental confusion began to emerge. In 1907, Alois Alzheimer was the first to identify specific histopathological changes associated with progressive degenerative dementia. He described two cases of dementia (in 1907 and 1911) characterized by symptoms of aphasia, apraxia, agnosia, and the histopathological finding of neurofibrillary tangles and milar foci (plaques) that distinguished it from dementia associated with cerebral arteriosclerosis. Today, efforts continue to focus on elucidating the etiology of dementia and to further delineate the associated pathological and physiological abnormalities observed across the range of clinical presentations with specific emphasis on genetics, prevention, and treatment.

COMPARATIVE NOSOLOGY

The conceptualization of dementia has evolved through successive editions of the DSM. In DSM-III, dementia was categorized as a syndrome and as a disorder. Organic brain syndrome included dementia and distinguished dementia from cognitive impairment syndromes caused by general medical conditions or substances. Dementias were also listed separately as specific disorders and included primary degenerative dementia of the Alzheimer type, multi-infarct dementia, dementia due to alcoholism, and dementia not otherwise specified. In DSM-IV, specific criteria for dementia due to Alzheimer disease, vascular disease, specific medical conditions, and substance-induced conditions were presented. The DSM-IV-TR further divided dementia into two more categories, adding dementia due to multiple etiologies, where each individual etiology is listed separately, and dementia not otherwise specified, which encompassed dementia that did not fall into one of these categories or where the etiology had not been established.

The DSM-5 utilizes the terms major neurocognitive disorder and mild neurocognitive disorder, which broadly correspond with the commonly used terms of dementia and mild cognitive impairment (MCI), respectively. The term neurocognitive disorder is broader than "dementia"; it extends beyond late-life, to include neurocognitive impairments at all ages including, for example, those caused by traumatic brain injury or HIV infection. The addition of mild neurocognitive disorder is a change from DSM-IV. Another change is that the diagnosis major neurocognitive

disorder can, rarely, be based on substantial impairment in even a single cognitive domain, if sufficient to interfere with independence in everyday activities (e.g., profound isolated memory or language deficit caused by a strategically located stroke or head trauma). Major and mild neurocognitive disorders are the basic clinical syndromes; in a change from previous DSM editions, at the syndromic neurocognitive disorder level, memory does not have to be one of the impaired domains. DSM-5 provides further diagnostic criteria for the etiological subtyping of the syndromes, for example, with criteria for neurocognitive disorders due to Alzheimer disease (where memory is typically impaired), frontotemporal lobar degeneration (where executive function is typically impaired), and further criteria are provided in accordance with current expert criteria for those diseases. Definitions and examples are provided for the different cognitive domains including social cognition that was not included in earlier editions of DSM. Specifiers exist for the presence of behavioral symptoms, as well as for the level of severity in major neurocognitive disorder, with mild severity indicating difficulties in instrumental activities of daily living, moderate severity indicating difficulties with activities of daily living, and severe for those who are fully dependent.

The etiology specific criteria within the DSM-5 are fairly compatible with the criteria set forth by many of the major expert workgroups and consensus guidelines. According to the National Institute on Aging-Alzheimer's Association workgroups for an individual to meet the criteria for dementia in general, he or she must show cognitive or behavioral impairment that interferes with functioning at home or work. This impairment is a decline from a previous level, is not explained by a delirium or major psychiatric disorder, and must be diagnosed through history taking and objective assessment of cognitive functioning. Unlike the DSM-5, there must be a minimum of two cognitive domains affected (memory, reasoning and judgment, visuospatial abilities, language, change in personality, or behavior).

Another classification system for dementias is offered in the *International Statistical Classification of Diseases and Related Health Problems*, tenth revision (ICD-10). This still uses a syndromic approach and offers criteria for general dementia and specifically for four types, Alzheimer disease, vascular dementia, dementia in other disease classified elsewhere, and unspecified dementia. These criteria require a decline in activities of daily living.

There are additional criteria for specific dementia types: the National Institute on Aging-Alzheimer's Association Workgroup Criteria for Alzheimer disease, the International Work Group on Frontotemporal Dementia and Pick's Disease criteria for Frontotemporal Dementia, Consortium on Dementia with Lewy Bodies International Workshop Criteria for DLB, the Movement Disorder Society Task Force criteria for

Dementia Associated with Parkinson Disease, the National Institute of Neurological Disorders and Stroke and the Society for Progressive Supranuclear Palsy (NINDS-PSP) Criteria for PSP, and an International Consortium's Criteria for Corticobasal Degeneration. These will be discussed in detail in later sections.

EPIDEMIOLOGY

With the aging of the population, the prevalence of dementia is rapidly increasing. Age is the leading risk factor for dementia and prevalence doubles about every 5 years from about 1.9 percent at age 65 to 70 to 6.3 percent at age 75 to 80, 21.7 percent at age 85 to 90, and 47.5 percent over age 90.

It is estimated that about 35.6 million people were living with dementia worldwide in 2010. At that time in the United States, 4.7 million Americans were estimated to be living with dementia of the Alzheimer type, with approximately 11 percent of people over the age of 65 being diagnosed with this. By 2050, the number of people in the United States with dementia of the Alzheimer type is projected to be 13.8 million, and worldwide the number of those with dementia is projected to reach 115.4 million.

The prevalence differs depending on the population. The prevalence rates reported above apply to the general population. However, in a primary care population, the prevalence rates are higher, and higher still in the general hospital setting. In the skilled nursing facility population, rates are closer to 50 to 60 percent.

Prevalence rates and anticipated increases in these are variable depending on the world region being studied. A 2013 systematic review and meta-analysis estimates slightly higher prevalence rates for those over the age of 60 in Latin America (8.48 percent) and the Caribbean (8.12 percent), and lower rates for the four sub-Sahara African regions (2.07 to 4 percent). Latin America, North Africa and the Middle East, as well as China, India, South Asia, and the Western Pacific are expected to have a rapid increase in the prevalence of dementia. The lower rates of dementia as well as a smaller expected rise in prevalence in the sub-Sahara regions may be due to the differing projections of aging there, secondary to that region's higher rates of child mortality and HIV.

Risk and Protective Factors

The leading risk factor for dementia is older age. Historically, female gender was believed to be a risk factor for dementia, but the increased number of women with dementia is now believed to be due to the fact that women have longer life expectancies than men.

Vascular risk factors increase risk not only for vascular dementia but also for Alzheimer disease and dementia in general. Hypertension,

particularly long-term hypertension starting in mid-life, increases risk for dementia. Other vascular risk factors present in mid-life including elevated body mass index, high cholesterol, and diabetes are also risk factors for dementia. Of note, hypotension in old age is a risk factor for Alzheimer disease. Cardiovascular disease, particularly when combined with peripheral vascular disease, also confers greater risk. Cerebrovascular disease is also a risk, not just for vascular dementia but also for Alzheimer disease. Congestive heart failure and atrial fibrillation are also risk factors. Obstructive sleep apnea has also been associated with an increased risk of dementia as well.

Early life depression has been associated with increased risk of Alzheimer disease, and long-term depressive burden has also been shown to be associated with an increased risk of cognitive decline, dementia, and MCI. The relationship between late life depression and dementia is more nuanced. Late life depression is associated with increased dementia but at least some of this risk is due to late life depressive symptoms occurring as part of the prodrome or early signs of dementia.

The effects of environmental factors and exposures on the development of dementia are controversial. Excessive alcohol consumption is associated with dementia, while mild to moderate alcohol intake may actually be protective. Smoking is considered to be a risk factor for the development of Alzheimer disease and vascular dementia, but there has been an inverse relationship shown between smoking and the development of Parkinson disease. Higher adherence to a Mediterranean diet has been shown to be associated with reduced risk of developing Alzheimer disease, and higher rates of physical activity have been associated with reduced risks of dementia.

There is some evidence to suggest that more years of formal education may confer some protection against dementia. Whether this is related to having a higher cognitive reserve necessitating a greater decline to cross a threshold for the recognition of dementia or whether higher education actually protects the brain from decline is unclear. Engagement in social networks and cognitively stimulating activities may also be protective against the development of dementia.

While there has been some research looking at protective factors in the development of dementia, clinical trials of putative protective factors including vitamin E, folate, vitamin B_6, and vitamin B_{12} to decrease incidence of dementia have been negative. There has also been mixed epidemiological evidence regarding statins decreasing the risk of developing dementia; however, a 2009 Cochrane review found no effect from the use of statins in late life on the prevention of Alzheimer disease or dementia.

ETIOLOGY

There are many different etiologies for dementia. Neurodegenerative disorders account for the vast majority of dementia cases. The most common is Alzheimer disease, which primarily involves neurodegeneration of the cholinergic system, among others. Alzheimer disease, alone or in combination with other etiologies, accounts for about 70 to 80 percent of dementias. On neuropathological examination of individuals with dementia, the majority of diagnoses are of mixed pathology, with Alzheimer disease comorbid with vascular infarcts or cortical Lewy body disease. In one neuropathology study of individuals diagnosed with probable Alzheimer disease, on autopsy, only 42.4 percent were found to have pure Alzheimer disease pathology, and more than 45 percent of individuals were found to have mixed pathology. DLB is characterized by the same neuropathological changes as Alzheimer disease, with neurofibrillary tangles and plaques, plus Lewy bodies in cortical and brain stem regions and accounts for about 15 to 35 percent of dementias. Vascular disease, both small vessel and large vessel disease affecting cortical or subcortical regions, accounts for about 5 to 20 percent of dementias. Cognitive deficits depend on which brain regions are affected or disconnected. The FTDs involve neurodegeneration and neuropathological changes in frontal lobe and temporal lobe regions. In early onset dementias (those occurring before the age of 65), frontotemporal lobe degeneration or vascular disease is the second most common etiology, following Alzheimer disease. Other neurodegenerative disorders affecting primarily the basal ganglia also cause dementia, specifically Parkinson disease, corticobasal degeneration (CBD), PSP and Huntington disease. These account for a smaller proportion of dementia cases. Finally, prion diseases (e.g., Creutzfeldt–Jakob disease [CJD], Gerstmann–Sträussler syndrome [GSS]) include both spontaneous and transmissible forms and can cause rapidly progressive dementias.

Other neurological disorders that can cause impairment in multiple cognitive domains include multiple sclerosis, brain tumors, normal-pressure hydrocephalus, subdural hematoma, and traumatic brain injury. Other medical conditions can also cause dementia, such as endocrinopathies (e.g., hypothyroidism), vitamin B_{12} deficiency, uremia, autoimmune disease (systemic lupus erythematosus), and others. Infectious diseases such as HIV and syphilis can cause dementia; HIV infection will be discussed in Chapter 10.5 (Other Cognitive and Mental Disorders due to a General Medical Condition). See Table 10.3–1 for a list of etiologies of dementia.

Table 10.3–1.
Etiologies of Dementia

Neurodegenerative	Alzheimer disease

	Dementia with Lewy bodies
	Frontotemporal dementia
	Parkinson disease
	Huntington disease
	Corticobasal syndrome
	Progressive supranuclear palsy
	Prion disease (Creutzfeldt–Jakob disease, bovine spongiform encephalitis, Gerstmann–Straüssler syndrome)
Vascular	Infarction
	Binswanger disease
	Hemodynamic insufficiency
Neurological disease	Multiple sclerosis
	Normal-pressure hydrocephalus
	Brain tumor (primary or metastatic)
Endocrine	Hypothyroidism
	Hypercalcemia
	Hypoglycemia
Nutritional	Vitamin B_{12} deficiency
	Thiamine deficiency
	Niacin deficiency
Infectious	Human immunodeficiency disease
	Neurosyphilis
	Cryptococcus
Metabolic	Hepatic insufficiency
	Renal insufficiency
	Wilson disease
	Metachromatic leukodystrophy
	Neuroacanthosis
Traumatic	Subdural hematoma
	Chronic traumatic encephalopathy/dementia pugilistica
Exposure	Alcohol
	Heavy metals
	Irradiation
	Anticholinergic medications
	Carbon monoxide

DIAGNOSIS AND CLINICAL FEATURES

Clinical Features

The hallmark feature of dementia is cognitive impairment. Patients must demonstrate impairment in one or more cognitive domains (see Table 10.3–2 for DSM-5 criteria). The onset, duration, course, presentation, and associated symptoms vary depending on the type of dementia. A general overview is provided here, while a more detailed discussion of the clinical features of each type of dementia will be presented in each individual section.

Table 10.3–2.
DSM-5 Criteria for Major Neurocognitive Disorder

A. Significant cognitive decline from a previous level of functioning in one or more cognitive domains, as evidenced by:
   1. Concern from the individual, the clinician or another knowledgeable informant that there has been significant cognitive decline
   2. Significant impairment in cognitive performance, ideally seen on standardized neuropsychological testing, or if that is not available, another clinical assessment
B. Cognitive deficits interfere in independence in daily activities
C. The deficits do not occur exclusively during the course of a delirium
D. The deficits cannot be explained by another mental disorder

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Cognitive Impairment. Memory impairment refers to the inability to learn new information or to recall previously learned information. Generally, short-term memory is impaired before long-term memory. Short-term memory impairment is often manifest as forgetting recent events and conversations, repetitive questions, repetitive retelling of stories, forgetting the date, forgetting appointments, misplacing objects, losing valuables, and forgetting that food is cooking on the stove. Patients are unable to learn new information or new skills and may become overwhelmed, disoriented, or confused in new environments (i.e., traveling or moving). Later in the illness progression, long-term memory will become impaired and patients will no longer recall their personal past history and will forget previously highly learned material. Procedural memory for previously learned skills may be retained longer.

Language impairment can manifest as difficulties in receptive or expressive language. Aphasia may be evident as word finding difficulties resulting in reduced fluency of productive speech, word substitutions, or mispronouncing words. Patients may have difficulty remembering names of distant family members or friends or acquaintances they see infrequently. As language impairment progresses, patients may forget names of close family and friends. They often have difficulty maintaining conversations, or may use generalized, vague terms to try to maintain the conversation; eventually they stop initiating conversation. Individuals may appear more withdrawn and disengaged as their ability to engage with others through conversation becomes more impaired. They may develop paraphasic errors and have poverty of content with empty speech. Eventually, speech may become incomprehensible and unintelligible.

Perceptual–motor impairments include difficulties in visual perception and visuoconstruction, as well as apraxia and agnosia. There may be difficulties in navigating in familiar environments, and individuals may

have more trouble during times of dusk, when lighting changes may affect perception. Apraxia, or impaired ability to carry out motor activities despite intact motor function, may manifest in the earlier stages as difficulty using familiar tools or machines (i.e., microwave oven, washing machine, lawn mower, drill) or trouble performing previously acquired skills (i.e., knitting, woodworking). In later stages, there may be dressing, bathing, or feeding apraxias that contribute to impairment in ability to perform activities of daily living. Agnosia, or failure to recognize or identify objects despite intact sensory function, may be reflected in the inability to recognize familiar objects, familiar faces or in later stages, one's own reflection in the mirror. It also manifests as lack of insight into one's own impairment and unrealistic assessment of one's abilities.

Individuals with impairments in complex attention typically have difficulty when there is competing stimuli, such as multiple conversations or ambient sound from the television or radio. They are easily distracted and have trouble registering new information; thinking may also appear to take longer.

Impairment in executive functioning refers to disturbances in planning, organizing, sequencing, and abstracting. Impairment in this area is reflected in difficulty performing complex tasks or problem solving. This may be manifest as trouble preparing a meal, managing medications, or managing finances. Judgment is often impaired and bad decisions are made because of deficits in executive functioning combined with lack of recognition of difficulties.

Social cognition includes the ability to recognize emotions and the ability to consider another individual's mental state. Impairments in social cognition may initially be manifested as difficulty recognizing social cues or emotions in others, apathy, disinhibition, and changes in levels of extraversion or introversion. Family members of the individual may describe a change in personality. As these impairments progress, behavioral changes become more overt, with individuals making sexually or politically inappropriate comments and showing disinhibited and intrusive behaviors that lack awareness for another's personal space.

Table 10.3–3.
Basic and Instrumental Activities of Daily Living

Basic Activities of Daily Living	Instrumental Activities of Daily Living
Ambulation	Transportation
Dressing	Using the telephone
Grooming	Meal preparation
Bathing	Medication management
Feeding	Financial management
Toileting	Housekeeping
	Laundry

Functional Impairment. According to DSM-5, each of the cognitive deficits must result in a significant impairment in functioning that represents a decline from a previous level of functioning. If patients are still working, there must be difficulties with job performance. Difficulties in managing finances or medications may also be evident. Initially there may be difficulty performing instrumental activities of daily living, and eventually basic activities of daily living (see Table 10.3–3for a description of these). There may also be difficulties in social functioning. This may manifest as disengagement from usual activities and interests, or trouble maintaining social relationships and social roles.

Screening

The role of dementia screening is dependent on the medical setting and history presented by the patient or concerned family members. The U.S. Preventative Services Task Force does not recommend for or against screening for dementia in the primary care setting. However, in 2011, the Center for Medicare and Medicaid Services began to require screening for cognitive impairment during the Annual Wellness Visit. Given the high prevalence of dementia, especially with advancing age, screening for dementia in high-risk groups may offer great clinical benefit. Dementia is associated with higher rates of delirium, adverse drug reactions, and falls, so that recognizing dementia earlier in its course might raise awareness and target prevention efforts to try to avoid these potential complications. In the psychiatric setting, screening for cognitive impairment in the older adult is essential given dementia's relationship to depression. Also some of the earliest signs of certain dementias may be behavioral or personality change that may lead individuals to initially present for psychiatric evaluation and treatment. Early recognition would also allow for initiating potentially disease modifying agents, when they become available, earlier in the course of illness that may potentially delay progression to more advanced stages.

Evaluation

A complete evaluation for dementia begins with a thorough history from both the subject and a reliable informant. Given the nature of memory loss in dementia, an accurate history cannot be expected to be obtained from the patient alone. Relevant aspects of the history include the onset of cognitive impairment (insidious or sudden), course (gradual or step-wise, progressive or episodic, or fluctuating), and duration of impairment. An assessment of each cognitive domain is critical. For memory, enquire about short-term, long-term, and remote memory. For language, enquire about

word finding difficulties and remembering names of family members and friends. For complex attention, enquire about difficulties maintaining attention when there are competing stimuli; for example, an individual may have difficulty following conversation in large groups of people, or when the television or radio is also on. For perceptual motor abilities and praxis, enquire about difficulties in navigating previously familiar environments, use of familiar tools or machines, maintenance of previously acquired skills and dressing or feeding apraxias. For agnosia, enquire about recognition of familiar objects and insight into their condition and limitations. Enquiring into executive function involves assessing ability to perform complex tasks or solve problems, such as the ability to manage finances or medications. For social cognition, changes in behavior or personality may be subtle or more overt, depending on the severity of the dementia.

Evaluation of Cognitive Domains

The degree of functional impairment should be ascertained. Assessment of executive functioning can be helpful in evaluating functional status (see Table 10.3–4). Asking about driving, management of finances, medication management, and activity planning can provide some information as to an individual's day to day functioning level. If there is suspicion for greater levels of impairment, enquiring about the ability to perform basic and instrumental activities of daily living is important (see Table 10.3–3). In some settings, evaluation with a structured measure, such as the Kohlman Evaluation of Living Skills (KELS) test, may be available to assist in gauging an individual's safety in the home.

In addition to the cognitive symptoms and related decline in functioning, patients should be evaluated for the behavioral and neuropsychiatric manifestations of dementia; these are described in more detail below. The Neuropsychiatric Inventory (NPI) is a useful tool that was developed to assess psychopathology in dementia patients and evaluates 12 neuropsychiatric disturbances: delusions, hallucinations, agitation, dysphoria, anxiety, apathy, irritability, euphoria, disinhibition, aberrant motor behavior, night-time behavior disturbance, and appetite and eating abnormalities.

Medical and neurological history is relevant, especially any changes that may have a temporal relationship to the onset or worsening of cognitive dysfunction. Cerebrovascular events can coincide with the onset or worsening of vascular dementia or dementia of any etiology. Certain medical illnesses can contribute to the onset of dementia and these may even be reversible causes of dementia that, when treated, can result in improvement in cognitive symptoms and functioning. Vascular risk factors, history of head trauma, and HIV risk factors may be relevant.

Psychiatric history is also relevant as certain psychiatric conditions may

overlap with or predispose to dementia. Depression can cause cognitive impairment that may mimic dementia, and may be associated with increased risk for Alzheimer disease. Patients with Down syndrome will develop Alzheimer disease at an early age. Alcohol and substance abuse history should be ascertained. Family history of dementia, early onset dementia, Huntington disease, or other neurological diseases is important.

Several aspects of the social history are particularly relevant. Education level is important as higher education may be a protective factor, and level of education is relevant in interpreting findings on cognitive testing as normative values are dependent on age and education. Exposure to certain chemical agents may also be relevant as some chemical exposures may be associated with particular types of dementia.

Table 10.3–4.
Evaluation of Cognitive Domains

Complex attention	Does the individual have difficulty paying attention?
	Is the individual easily distracted in environments where there are competing stimuli?
	Does the individual have difficulty retaining newly presented information?
	Are routine tasks taking longer than usual?
Executive function	Does the individual have more difficulty with multitasking?
	Is the individual still able to manage their own finances or medications?
	Is more assistance required to plan out activities or instrumental activities of daily living?
	Is more assistance required to make decisions?
	Is the individual having more trouble in large social situations, or finding them less enjoyable?
Learning and memory	Is the individual having difficulty recalling recent events and activities?
	Is the individual repeating himself or herself more often?
	Is the individual needing to rely more on lists or reminders?
	Is the individual misplacing or losing items?
	Does the individual need more reminders to attend to a particular task?
Language	Is there word finding difficulty?
	Is the individual making more grammatical errors in conversation?
	Is the individual forgetting people's names?
	Is the individual making more generalized, vague statements in response to questions?
	Is the individual using more generalized pronouns as substitutes for particular names of items?
Perceptual motor	Is he or she getting lost when travelling or in familiar environments?
	Is he or she having more difficulty using previously familiar tools and appliances?
	Is he or she having difficulty driving a car?
Social cognition	Has the individual or family members noticed a personality change?
	Does the individual appear to be less empathic or more disinhibited in their speech or behavior?
	Is the individual making sexually, politically, or religiously inappropriate

comments?
Does the individual show a lack of awareness of another's interpersonal space?

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

A complete physical and neurological examination is necessary. The physical examination is targeted toward medical conditions that may cause dementia. The neurological examination is necessary to identify focal neurological signs that may point to specific brain lesions or pathology.

Mental status examination is useful to assess self-care abilities (appearance and grooming), behavior (degree of engagement in the interview, demeanor, inappropriate behavior), affect (apathy, depression, anxiety), speech (language function, fluency, comprehension), perception (hallucinations, paranoia, delusions), and insight and judgment.

Bedside cognitive testing is performed to characterize the pattern and extent of cognitive impairment. The most widely employed tool for basic screening and evaluation is the Montreal Cognitive Assessment (MoCA). The MoCA is a 30 point test that can be administered in the office setting in about 10 minutes. The test is easily accessible, and available in a number of languages. It is broken down by different cognitive domains and assesses visuospatial and executive functioning, confrontation naming, memory by immediate and short-term recall, attention, language, abstraction, and orientation. The MoCA has a 90 percent sensitivity of detecting MCI and an 87 percent specificity in excluding those cognitively intact. Consideration should be given for referral to a neuropsychologist for detailed cognitive testing on one of several standardized test batteries used commonly in the elderly with dementia. These more detailed tests may be beneficial in more clearly delineating areas of impaired function and in providing a baseline against which illness progression can be measured.

Severity Measures. There are several widely employed scales to assess severity of dementia in terms of level of cognitive and functional impairment. The Clinical Dementia Rating Scale (CDR) measures several domains of function including memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. Another staging scale is the Functional Assessment Staging (FAST) that has been developed to track the course of Alzheimer disease and is also useful for other dementias. Staging is based on the ability to perform certain skills or functions and is dependent on baseline skills, acquired deficits, and the social environment. Therefore, if a patient is not performing a certain skill because they never did or because their current environment does not demand it, it is not counted as a deficit. Patients with a CDR score of 0.5 or a FAST stage 2 to 3 have mild deficits in memory and sometimes other areas but have only mild or questionable functional difficulties. This

generally corresponds with a diagnosis of MCI or mild dementia. Patients with a CDR score of 1 and a FAST stage 4 have some difficulty with instrumental activities of daily living, but they can still remain fairly independent with some assistance and prompting and corresponds to "mild" dementia. Patients with a CDR score of 2 and FAST stage 5 or 6 have impairment with more basic aspects of maintaining a home and themselves and will require some assistance with instrumental activities of daily living. Patients with a CDR score of 3 or FAST stage 6 or 7 have "severe" dementia and require considerable to total assistance with personal care and activities of daily living. In the terminal phase, patients develop contractures and are bed-bound and require total care.

Neuropsychiatric Manifestations

Neither the DSM-5 nor ICD-10 requires neuropsychiatric or behavioral symptoms for the diagnosis of dementia. However, these symptoms commonly co-occur with dementia and are often the cause of the greatest caregiver distress.

Behavioral disturbance is common in many types of dementia, although prevalence rates vary widely depending on the population and dementia subtype studied. The prevalence of behavioral disturbance increases with advancing dementia. Behavioral disturbances include mood changes, anxiety, personality changes, psychosis, sleep disturbances agitation/aggression, and other behaviors, each of which is described below.

Mood Changes. Major depression is common early in the course of dementia, and prevalence rates are estimated at about 10 percent. Depressive symptoms are even more common, occurring in about half of patients with Alzheimer disease, and may be a consequence of aspects of neurodegeneration. Depressive symptoms correlate with greater temporal lobe and cingulate gyrus hypometabolism. In addition, depressive symptoms, emotional lability, and irritability may be related to psychological reactions to the awareness of losing one's memory and anticipation of increasing functional impairment. Suicidal ideation and behavior may also occur through the course of dementia but are especially concerning in mildly impaired individuals who have insight into their deficits and still have the executive capacity to formulate and carry out a plan. Depression tends to decrease in prevalence with advanced dementia, so treatment for early depression may be withdrawn as the illness progresses. Elation and even a frank manic episode can also occur in the context of dementia but is rare.

Anxiety. Anxiety is fairly common throughout the course of dementia and is estimated to occur in about 60 percent of patients. Anxiety is often

manifest as fear of being alone, and patients will search for their caregivers so as not to be alone. Some patients exhibit "catastrophic reactions" when confronted with their cognitive limitations or when facing changes in routine or environment.

Personality Changes. A range of personality changes may be evident. FTD often presents with disinhibition and impulsivity, which often predate obvious cognitive deficits. Examples include making inappropriate jokes, being overly familiar with strangers, disregarding social norms, inappropriate sexual remarks or behaviors, impulsive buying, or succumbing to salespeople or exploitative schemes. Alternatively, an apathy syndrome with amotivation and withdrawal may also be associated with various types of dementia including Alzheimer disease and vascular dementia. In some cases, pre-existing personality traits may become stronger or exaggerated during the course of a dementing illness.

Psychosis. Psychotic symptoms in dementia generally occur in the middle stages of illness and often co-occur with behavioral disturbance. A common early psychotic symptom is paranoia and the belief that belongings have been stolen. This may develop as a result of poor short-term memory and misplacing objects. Inability to find a lost or misplaced object is interpreted in a paranoid way with the conviction that the object must have been stolen. This paranoia can become more pervasive as the illness progresses. In later stages, patients may have hallucinations in any modality, but most commonly visual hallucinations. They may see deceased relatives. The hallucinations are often congruent with delusions of the same theme. People often believe they are still working or engaging in previous activities. Psychotic symptoms can sometimes lead to agitation, aggression, and behavioral disturbance as patients act out on their hallucinations, paranoia, or delusional thinking.

Sleep Disturbance. With advancing dementia, altered sleep–wake cycles can result in disrupted and fragmented sleep in about half of patients. Patients may have phase-shifted sleep, going to bed late, and sleeping late in the morning. They may also take frequent naps, which may be more prevalent in less stimulating environments where there is less to keep them engaged and awake.

Agitation and Aggression. Agitation and aggression become more prominent with increasing levels of dementia severity. There is a significant spectrum of behaviors that fall under the label of "agitation." These behaviors are not always aggressive or directed toward a target. Agitation may also include nondirected behaviors such as continually pacing back and forth, or purposeless hyperactivity such as continuously rummaging

through drawers. Agitation and aggression can come in both physical and verbal forms, with some individuals hitting or striking out, while others may swear, make threats, or have persistent disruptive, but nonthreatening vocalizations. The behaviors can be environmentally or task dependent, such as with personal care, or may be more generalized. These behaviors often lead to acute hospitalization, placement in higher levels of care, cause significant caregiver burden, and even cause disability in caregivers due to injury.

Other Behavioral Disturbances. As cognitive impairment progresses, some individuals become at risk of wandering. This can be particularly dangerous in situations of inclement weather and for those individuals with altered sleep–wake cycles who still reside at home with a caregiver, as wandering may occur at nighttime. Disinhibited social behavior may also be seen, with individuals making inappropriate comments or gestures in public places.

LABORATORY EXAMINATION AND PATHOLOGY

The following are a description of the different studies available to aid in the diagnosis of dementia. The decision to order these to aid in the identification of a specific etiology of dementia is guided by the history and clinical examination, as well as the differential diagnosis.

Laboratory Studies

The purpose of obtaining routine laboratory studies in the evaluation of dementia is to rule out other general medical causes of dementia including potentially reversible causes (see Table 10.3–5). These include complete chemistry panel, complete blood count, hepatic function, thyroid stimulating hormone, vitamin B_{12}, and folate. Other laboratory studies that may be indicated depending on the history might include Venereal Disease Research Laboratory (VDRL) for syphilis, erythrocyte sedimentation rate, and HIV.

Table 10.3–5.
Laboratory Evaluation of Dementia

Assessments	Rationale
Labs: Complete blood count, serum electrolytes, renal and hepatic function, glucose, albumin and protein, vitamin B_{12} and folate, rapid plasma reagin (syphilis), thyroid-stimulating hormone, urinalysis	Rule out correctable or contributory causes of dementia
Structural imaging: Computed tomography or magnetic resonance imaging	Rule out infarcts, mass lesions, tumors, and hydrocephalus

Molecular and functional imaging: PET-amyloid imaging, SPECT, fluorodeoxyglucose-PET	Identifies certain pathological biomarkers (Aβ accumulation) or patterns of cerebral blood flow and/or metabolism that may be characteristic of specific neurodegenerative disorders
Cerebrospinal fluid	Detect altered levels of amyloid beta and tau protein in Alzheimer disease, or elevated 14-3-3 protein and neuron-specific enolase in Creutzfeldt–Jakob disease

Structural Neuroimaging

Structural neuroimaging studies are obtained to rule out abnormalities in the brain that can cause or contribute to dementia including normal pressure hydrocephalus, subdural hematoma, brain tumor or metastases, or cerebrovascular events. Generally a computed tomography (CT) scan is sufficient. However, magnetic resonance imaging (MRI) is preferred if looking for small or deep brain infarcts as this imaging modality is better for visualizing vascular changes. Atrophy may be seen in patterns suggestive of specific dementia causes (see review of individual causes of dementia, in section on Dementia Subtypes, below). An MRI with gadolinium contrast is useful if there is a high index of suspicion for tumor.

Molecular Neuroimaging

In an effort to identify biomarkers of the differing dementia processes, much work has gone into the development of molecular and functional imaging. During the progression of a particular neurodegenerative process, these markers may have differing utility. Positron emission tomography (PET) amyloid imaging can be used to identify Aβ accumulation. The absence of amyloid deposits indicates the dementia is not likely to be due to Alzheimer disease. Single-photon emission computed tomography (SPECT) or fluorodeoxyglucose (FDG)-PET can be used to demonstrate blood flow or metabolic rate in various brain regions to try to elucidate the etiology of dementia. For example, in Alzheimer disease, there is a characteristic pattern of decreased uptake in posterior cingulate, temporal, and parietal cortex. In vascular dementia, there will be patchy areas of decreased uptake correlating to areas of hypoperfusion or infarct. In FTD, there will be a pattern of diminished uptake in primarily frontal as well as anterior temporal lobe regions, and in DLB, there may be a pattern of occipital hypoperfusion. Dopamine transporter imaging uses ligands specific for the dopamine transporter to examine the nigrostriatal dopamine system and is abnormal in Parkinson disease, DLB, multisystem atrophy, and PSP.

CSF Studies

These studies may be used to measure amyloid-β (Aβ) protein, tau protein,

and phospho-tau, all of which may be altered in Alzheimer disease. In sporadic CJD, there may be elevations in 14-3-3, total tau, or neuron-specific enolase.

Electroencephalogram

Electroencephalograms (EEGs) may have a role in dementia evaluation if there is a concern about a seizure disorder or delirium. In CJD, EEG can show periodic sharp wave complexes. However, EEG is not necessary in the routine evaluation of dementia.

Neuropsychological Testing

Neuropsychological testing may be useful in elucidating the pattern of cognitive deficits in cases where the presentation is not typical of any one type of dementia. In Alzheimer disease, there is a pattern of early short-term memory and language impairment followed by the development of executive dysfunction and then visuospatial dysfunction. In FTD, memory is relatively spared, while either language impairments or executive dysfunction predominate. In vascular dementia, there is a pattern of patchy deficits depending on where vascular lesions have occurred in the brain. Neuropsychological testing may also be helpful in identifying when cognitive impairment is due to depression or a factitious disorder.

Neuropathology

In individuals whose dementia does not arise from a previously characterized familial autosomal dominant mutation, a definitive diagnosis of dementia type can only be made with neuropathological investigation of the brain through either brain biopsy or autopsy. There is usually not a clinical indication for brain biopsy so postmortem evaluation of the brain is the only way to provide a definitive diagnosis of dementia type.

DIFFERENTIAL DIAGNOSIS

In the evaluation of cognitive impairment, the psychiatrist must first determine that this is not occurring in the sole presence of a delirium and that the cognitive impairment is not the result of a mild neurocognitive disorder. Medical conditions that can cause cognitive impairment and medication-induced impairment should be ruled out. Other mental disorders, such as a schizophrenia spectrum or psychotic disorder, a depressive disorder, a substance related disorder, and factitious disorder must be ruled out as well. Once these causes have been ruled out, the focus switches to identifying the diagnosis of the specific underlying neurodegenerative disorders. These are described briefly below.

Delirium

Delirium can be differentiated from dementia by the presence of an altered level of consciousness with the cognitive impairment, whereas dementia by definition is in clear consciousness. In delirium, the onset is generally sudden and acute and represents a significant change from baseline. There is typically a fluctuating level of alertness. Attention is the hallmark cognitive domain that is impaired. Delirium can occur in patients with dementia, and in fact, dementia is a risk factor for the development of delirium. Dementia accounts for the baseline cognitive decline that is punctuated by an acute worsening related to the delirium. Delirium is usually precipitated by one or more, and usually a combination of, medical factors or medications or substances (see Chapter 10.2, Delirium for further discussion).

Mild Neurocognitive Disorder

Mild neurocognitive disorder, more commonly known as MCI, is impaired cognition in one or more domains, but to a lesser degree than in dementia. There may be mild impairments in the more complex aspects of daily functioning, or an individual may need to rely on more compensatory strategies; however, there is no significant impairment on activities of daily living. Because the pathophysiological process of neurodegenerative dementias begins years before there is evidence of cognitive or functional impairment, MCI may represent a prodromal state of these disorders. However, not all individuals with MCI will go on to develop dementia. Research to recognize and quantify biomarkers to identify those both with and without MCI who will go on to develop dementia is underway.

Medical Conditions

General medical conditions that can cause or exacerbate cognitive impairment must be ruled out. Treatable factors that can cause or contribute to dementia include hypothyroidism, hypercalcemia, hypoglycemia, vitamin B_{12}, thiamine or niacin deficiency, renal or hepatic impairment, or autoimmune disease (i.e., lupus erythematosus, vasculitis, neurosarcoidosis, Hashimoto encephalopathy; see Table 10.3–1 and Chapter 10.5).

Medications

Many medications, alone or in combination, can cause cognitive impairment that may mimic dementia. Medications with anticholinergic properties are major culprits (see Table 10.3–1).

Schizophrenia Spectrum and Other Psychotic Disorders

Schizophrenia is associated with multiple cognitive deficits. These cognitive deficits may emerge prior to the onset of the psychotic syndrome in young

adulthood, and often progress during the early years of subsequent illness, but are relatively constant thereafter. Some individuals with schizophrenia will develop a progressive deterioration of cognitive functions beginning in late life, which may represent a superimposed neurodegenerative process.

Depressive Disorders

Cognitive impairment frequently co-occurs with depressive episodes in late-life. The impairment is substantial (though less so than in dementia) and is present in about 50 percent of depressed elderly, which is more frequent than in younger patients. Cognitive impairment in the context of depression was formerly termed "pseudodementia." In the past, it was widely believed that cognitive impairment mimicked the symptoms of dementia, but reversed upon resolution of the depressive episode. Numerous studies have now shown that when cognitive impairment co-occurs with depression in late life, it generally persists into depression remission, particularly in the memory and executive function domains. Thus, in many cases both depression and cognitive impairment in the elderly likely have shared neurobiological underpinnings. Moreover, depression seems to increase the risk for future dementia (both Alzheimer disease and vascular dementia), although it is not known whether depressed mood represents prodromal dementia or is etiologically related to future progressive decline. The varied findings in depression suggest that there are likely multiple phenotypes and pathways with different long-term trajectories and cognitive outcomes.

Alcohol Use Disorder

Substance use disorders, particularly alcohol abuse and dependence, can cause dementia (see Chapter 11: Substance Related and Addictive Disorders). Cognitive impairment is usually worse when patients are still actively using alcohol or other substances, and there may be some improvement with abstinence.

Factitious Disorder

Malingering or factitious disorder may present with dementia-type symptoms. The cognitive deficits are inconsistent over time and generally uncharacteristic of the pattern and course seen in dementia. There is some secondary gain from appearing cognitively impaired.

Neurodegenerative Disorders

Once the diagnosis of dementia, or major neurocognitive disorder, is identified, and the above etiologies have been ruled out, the differential diagnosis shifts to identify the most probable specific etiology of the neurodegenerative process. Information and data obtained from the

history, examination, neuropsychological testing, and imaging studies aid in the identification of the specific etiology, and are summarized for each disorder in the individual sections below. It is also possible for an individual to have more than one neurodegenerative process occurring at the same time (e.g., both Alzheimer disease and Lewy body disease); this can blur the diagnostic picture, and unfortunately a definitive diagnosis cannot be made until autopsy.

COURSE AND PROGNOSIS

The course and prognosis of dementia is variable depending on the etiology.

Generally, onset is insidious, although abrupt onset due to a cerebrovascular or traumatic event is possible. Duration is generally years although it may vary from 6 months in the case of CJD to 15 years with Alzheimer disease. In neurodegenerative dementias there is a progressive cognitive decline and progressive functional impairment that may be of variable rate and pattern depending on the etiology and type. Dementia eventually leads to death unless another disease process supersedes.

Identifying the correctable causes of cognitive impairment such as vitamin B_{12} deficiency, hypothyroidism, or normal-pressure hydrocephalus can result in some improvement in symptoms. Also, modifying vascular risk factors can ameliorate the course of dementia, even of nonvascular etiologies.

Gradual functional decline renders patients increasingly more dependent over the course of the illness. They slowly lose the ability to function independently. They lose the ability to perform instrumental activities of daily living such as shopping, meal preparation, managing finances, and managing medications. Then they have difficulty bathing and maintaining hygiene and dressing and feeding themselves. They require more and more prompting and then assistance with these tasks as the disease progresses, and they will eventually require total care.

Dementia increases the risk for complications from intercurrent medical illness including delirium. Any medical illness can precipitate rapid decline in cognition and functioning and markedly increase the level of care required. Dementia is also associated with an increased risk of falls and adverse reactions to medications.

TREATMENT

Treatment of dementia requires a multimodal approach not only targeting cognitive deficits, but also addressing neuropsychiatric symptoms and overall level of functioning as well as caregiver issues, legal issues, safety, community support, and levels of care. Treatment modalities and strategies will vary with the stage and type of dementia and must evolve over time to

address changing issues. Educating patients and their families about the course of illness and providing individualized treatment plans that ensure safety and adequate levels of care are essential.

General principles of dementia management will be provided here. The evidence and rationale for specific treatments for each specific type of dementia will be discussed in that section.

Nonpharmacological Management

Psychological, behavioral, and psychosocial interventions are an important component of the management of the patient with dementia of any type. Educating patients and families about dementia and the course and what to expect at each stage is important for general treatment and anticipatory planning. Maintaining safety, ensuring adequate supervision and support, connecting patients and their families with community supports and resources, providing support for caregivers, and encouraging legal and financial planning, all play an important role in the management of the patient with dementia.

Psychotherapeutic Approaches. Psychosocial interventions are an important component of the management of the patient with dementia of any type. While these interventions are important, there is no conclusive data on what type of patient or disease process to best use them for. Evidence supporting many of these therapies is limited as few have been subjected to randomized, double-blind, controlled trials and many of the reported studies are often limited by small sample size, poorly characterized patient populations, incompletely described interventions, poorly defined outcomes, or short duration of follow-up.

Psychotherapy. Psychotherapy may be useful in earlier stages of dementia before short-term memory is too impaired. A 2014 Cochrane review of 6 studies looking at the effects of therapy on depression and anxiety in individuals with dementia and MCI demonstrated a reduction of depressive symptoms; however, there was no effect on self-rated or carer-rated anxiety. The psychotherapeutic approaches in these studies ranged from therapeutic counseling, to cognitive behavioral therapy, to brief interpersonal psychodynamic therapy. In the latter, emphasis was placed on improving interpersonal relationships and included the caregiver for a brief portion of each session.

Reminiscence therapy has also been studied in dementia. This therapy allows patients to recall and relive past life events, stimulating memory and mood within the context of their life history. This therapy has resulted in modest short-term gains in mood, behavior and cognition in “confused” elderly persons.

Behavioral Management. Behavioral management has an important role in the management of dementia. Behavioral treatments that identify antecedents and consequences of problem behaviors and then effect changes in the environment to alter the behaviors have been shown to be beneficial in reducing disruptive behaviors in small trials and case studies, and are widely employed in clinical practice. For example, scheduled toileting can reduce the frequency of urinary incontinence. Behavior management training programs for caregivers in the community or in residential or nursing homes have shown modest reductions in problem behaviors and, in some studies, improvement in mood, cognition and physical role function. Limited follow-up data suggest that the benefits do not persist beyond the intervention.

Cognitive Retraining. Cognitive skills training exercises focused on maintaining specific cognitive skills and reality orientation focused on improving orientation, have limited support from reported studies. The benefits for orientation, cognition, behavior, function, and social interaction are generally short-lived. In addition, these approaches may actually be detrimental to patients as several studies have reported anger, frustration and depression in patients and depression in caregivers with these cognitive-oriented approaches.

Stimulation-Oriented Therapies. Stimulation-oriented therapies aimed at enhancing pleasurable activities have some support from clinical trials. Examples include recreational therapy (i.e., crafts, games), art therapies, pet therapy, multisensory stimulation, simulated presence, aromatherapy, and exercise. Results with most of these modalities are mixed, with some studies demonstrating improvements in mood and behavior and reductions in agitation and problem behaviors and others showing no significant benefit. Despite the lack of convincing evidence for these modalities, their inclusion in a multimodal treatment plan for patients with dementia is humane and makes sense.

Psychoeducation. Providing education to patients and their families about dementia, its course and consequences, and treatment is important in helping them to plan for future disability. While anticipating the upcoming stages can be difficult, it affords patients the opportunity to make arrangements while they are still able to be involved in the decision-making process. There are many excellent resources, including books and community agencies, available to help guide families through the caregiving process.

Safety Awareness and Management

HOME SAFETY. Another important aspect of dementia management is

maintaining safety in the community and at home. Patients with dementia are at greater risk for accidents. Forgetting to turn off the stove while cooking can lead to fires. Individuals may have more difficulty preparing meals, and so assistance for this from family or supplemental meal programs may be helpful. Forgetting to lock doors can increase risk for burglary or home invasions. As dementia becomes more severe, patients are at higher risk of wandering, and so door alarms or identification bracelets may be necessary. Dementia patients are more vulnerable to predatory scams through door to door, phone or mail solicitation. Assessing gait and risk of falls as well as well as a patient's insight into what to do during an emergency is also important. Addressing threats to patients' safety is necessary to protect patients from inadvertently harming themselves or others.

MEDICATION MANAGEMENT. As patients have more difficulty with memory and executive functioning, it may be more difficult for them to manage their own medications, particularly if they are required to take multiple medications or medications multiple times a day. Assistance of a family member is often necessary, and at times may be needed during the MCI stage.

DRIVING. Driving safety is an important issue. Dementia patients are at greater risk for motor vehicle accidents**.** Patients should be advised to limit their driving in the mild stages of dementia and to stop driving at the first signs of impaired driving ability or accidents, or when they have reached the moderate stage of illness. Referral for driving evaluations is recommended to assess driving skills when in doubt, or to confirm to patients their limitations. Many states have mandatory reporting laws. Driving restrictions can be very difficult for patients to accept due to the loss of independence. Families often need to get involved to enforce these restrictions, removing keys or disabling or removing vehicles.

Legal and Financial Planning. Patients and their families should be encouraged to pursue legal planning such as living wills, power of attorney, and guardianship for when they no longer have the capacity to make decisions for themselves. Careful financial planning for in home or facility-based care is often needed, and is best accomplished in advance.

Caregiver Support. Caregiver burden associated with caring for a loved one with dementia is great. As demented patients decline, they become more dependent and require help with more activities of daily living. Caregiving is physically as well as emotionally demanding, because the patient has less ability to communicate, eventually may not recognize the caregiver, and frequently exhibits inappropriate and difficult behaviors. In fact, caregivers are at risk of physical injury by the patient during periods of

providing direct assistance with activities of daily living, or due to agitation or acting on delusional beliefs. Behavioral disturbance in patients is a major source of caregiver burden and also increases risk for caregiver depression. Finally, caregivers also frequently become isolated from their social supports as patients can no longer engage in their usual social functions. The lack of social support is another risk factor for developing depression.

Caregiver stress and depression can also impair the caregiver's ability to provide adequate and appropriate care and is a risk factor for abuse and neglect of the demented patient. Therefore, treatment of the demented patient must also include addressing the needs of, and when appropriate, treating the caregiver. Interventions shown to reduce caregiver burden include psychoeducation about dementia, and comprehensive support including individual and family counseling and support groups. Problem solving therapy for caregivers of individuals with dementia and MCI has recently been shown to be effective in reducing depression and anxiety. Multidisciplinary and individualized approaches particularly aimed at managing problem symptoms and behaviors may be especially beneficial.

Abuse and Neglect. Patients with dementia are at greater risk for elder abuse and neglect. The strains of caregiving are a risk factor for caregivers to abuse those in their care. In addition, due to the demented patient's poor memory and inability to recall or solve problems, they may be less likely to report abuse or the veracity of their reports may be questioned.

There are several instruments used to screen for elder mistreatment, but none has been widely accepted. In situations where abuse or neglect of the demented patient is suspected, further evaluation is indicated. Reporting is mandated in some states when elder mistreatment is suspected so that a complete investigation can be conducted and interventions performed when necessary.

End-of-Life Issues. Dementia is a terminal illness. While patients often die of intercurrent illnesses before the end stage of dementia, many patients succumb in the end stage of dementia, often due to pneumonia or other infections. Complications of end-stage dementia include loss of appetite, inability to feed oneself, and swallowing difficulties, making it difficult to ensure adequate nutrition. When dementia progresses to this stage, families will be faced with decisions regarding many life-prolonging interventions, including placement of feeding tubes and treatment of emergent infections. Ideally, these issues will have been addressed with the patient while in early illness stages, or via a preexisting advance directive, so the patient's wishes can be used to guide the decision-making process. Alternatively, involvement of hospice or palliative care services can be considered.

Pharmacological Treatments

Pharmacological treatments available for dementia are limited to symptomatic relief. To date, with the exception of interventions to reduce the risk of recurrent stroke in vascular dementia, there are no treatments available that can alter the disease course or produce substantial or long-lived improvements in symptomatology.

Treatment of Neuropsychiatric Manifestations

Treating the neuropsychiatric manifestations is an important part of the management of the patient with dementia. The neuropsychiatric manifestations, particularly the agitation, behavioral disturbances, and psychosis, contribute substantially to caregiver stress and often prompt transitions to higher levels of care. Adequate control of these symptoms can markedly improve the quality of life of the patient and caregiver alike.

Psychosis and Behavioral Disturbance. Psychotic symptoms in dementia include hallucinations, delusions, and paranoia, and are often associated with agitation or aggression. Agitation can be physical or verbal, aggressive or nonaggressive, directed toward a target or not, and can sometime be dependent on the situations (e.g., with caregiving). Agitation is present in many outpatients with dementia, and may be even more common in those residing in nursing facilities.

NONPHARMACOLOGICAL TREATMENTS. Nonpharmacological treatments of agitation are first-line, and identifying the context of problem behaviors can aid in the development of strategies for intervention. When a target behavior is associated with a particular activity, either modifying the activity, or avoiding it if possible, can eliminate the behavior. Simplifying activities or modifying expectations to the cognitive and skill level of the patient can eliminate frustration. Modifying the environment to compensate for the patient's deficits can eliminate triggers for problem behaviors. Ensuring adequate sleep, nutrition and stimulating activities can also help reduce agitation.

When the behavior poses a threat to the patient's or caregiver's safety, more restrictive measures must be considered. In a nursing home or hospital setting, one-on-one supervision may be required. The use of physical restraints should be avoided if possible as it may increase the risk of falls and mortality.

When nonpharmacological measures are ineffective, or safety is an issue, pharmacological approaches should be considered.

PHARMACOLOGICAL TREATMENTS

Antipsychotics. The second-generation antipsychotics risperidone

(Risperdal), olanzapine (Zyprexa), quetiapine (Seroquel), and aripiprazole (Abilify) have been studied in placebo and active comparator-controlled trials of psychosis and agitation in dementia in both nursing home and community outpatient settings. Studies of risperidone, olanzapine, and aripiprazole have shown modest but clinically significant improvements in measures of agitation, aggression, psychosis, and behavioral disturbances. These agents are indicated as first-line treatments in patients with dementia when agitated, aggressive, or psychotic behaviors are present and either the patient's function is impaired or they are endangering themselves or others.

A meta-analysis of these studies demonstrates greater benefit for symptoms of agitation than psychosis. There is also a significant placebo response rate in these studies suggesting the importance of nonpharmacological measures in the treatment of agitation and psychosis. The largest comparison study of three different second-generation antipsychotics—risperidone, quetiapine, and olanzapine—in patients with Alzheimer disease and psychosis or agitation treated for up to 36 weeks, the CATIE-AD trial, demonstrated no differences between any of the treatment groups and placebo in the primary outcome measure, time to discontinuation of treatment for any reason. However, time to discontinuation for lack of efficacy favored olanzapine and risperidone, while time to discontinuation due to adverse events or intolerability favored quetiapine and placebo.

First-generation antipsychotics have also been extensively studied for treatment of psychosis and agitation in dementia. Two meta-analyses of these trials showed clinically significant improvement with several first-generation antipsychotics of 59 percent and 64 percent compared to 41 percent and 38 percent with placebo, respectively. There were no significant differences between agents. However, adverse effects were common and drop-out rates were high.

Clozapine (Clozaril) has been shown to be useful in controlling psychotic symptoms in patients with Parkinson disease and DLB due to the absence of extrapyramidal side effects. There is not currently available data to support the efficacy of ziprasidone (Geodon), iloperidone (Fanapt), lurasidone (Latuda), asenapine (Saphris), and paliperidone (Invega) in treating behavioral disturbances in dementia.

Antipsychotic medications are associated with serious side effects and morbidity, particularly in the elderly. The first-generation antipsychotics, particularly the high potency agents such as haloperidol, are associated with a higher risk of tardive dyskinesia, neuroleptic malignant syndrome, akathisia and parkinsonism. The low-potency agents, such as chlorpromazine, are associated with central and peripheral anticholinergic effects, sedation, and postural hypotension. The second-generation antipsychotics are associated with an increased risk for the metabolic

syndrome including weight gain, diabetes mellitus and hyperlipidemia, but can also cause sedation, fatigue and extrapyramidal side effects as well.

Both first- and second-generation antipsychotics have been associated with increased risk of cerebrovascular events and mortality, with first-generation medications having a slightly higher risk. In discussing the risks and benefits of these medications, it is important to recognize the above risks, as well as the risks of untreated or undertreated psychosis and agitation, which could lead to inadvertent harm to self or others, restraint use, poor quality of life, and placement in a higher level of care. Therefore, use of antipsychotic medications for the treatment of psychosis and agitation must be carefully considered after nonpharmacological measures have been implemented, weighing the evidence for the potential benefit for target symptoms, against the risks of the medications themselves and the risks of alternative approaches to maintaining safety.

Benzodiazepines. Benzodiazepines have been compared to placebo as well as to antipsychotics for the treatment of behavioral disturbances. In general, they are more effective than placebo but less effective than antipsychotics in controlling behavior symptoms. However, they may be associated with paradoxical effects, disinhibition, worsening cognition, and falls. Cumulative toxicity tends to be greatest with long half-life agents, which should be avoided. At present, the use of benzodiazepines is largely limited to short-term control of acute agitated behavior with short-acting agents such as lorazepam (Ativan) or oxazepam (Serax).

Anticonvulsants. Data on the use of carbamazepine (Tegretol) is mixed. It has been shown to have modest benefit for agitation in short-term trials; however, its risk of side effects and drug–drug interactions are limiting. Valproate (Depakene) has not been shown to be efficacious in randomized controlled trials, and higher doses are also associated with high rates of adverse events. The newer anticonvulsants have not been studied for behavioral disturbances in dementia.

Selective Serotonin Reuptake Inhibitors. There have been several studies of selective serotonin reuptake inhibitors (SSRIs) in the management of behavioral disturbance. A 2011 Cochrane review of 9 trials examining the effects of antidepressants and trazodone on agitation found there to be limited evidence to support their use. A more recent randomized controlled trial using citalopram for the treatment of agitation demonstrated reductions in agitation and caregiver distress; however, treatment was also associated with a worsening of cognition and prolongation of the QTc interval. As the tolerability of these medications may be greater than that of the antipsychotics, they may help in the management of agitation, especially in patients in whom the risk benefit ratio for use of antipsychotic medication is not favorable.

Others. There is some evidence to suggest that trazodone may be helpful in decreasing irritability, anxiety and restlessness and affective disturbance especially in patients with FTD. Acetylcholinesterase inhibitors and memantine have been shown to have little benefit in the treatment of agitation and psychosis. Minimal evidence supports the use of beta-blockers or lithium (Eskalith), and both classes of agents can have significant adverse events that preclude recommending them for use in a frail elderly population. There is minimal evidence that buspirone (Buspar) may be efficacious in the treatment of agitation or anxiety in dementia.

Depression. Several antidepressants have been studied for the treatment of depressive symptoms in patients with Alzheimer disease. SSRIs have been the most studied, as there are clinical concerns regarding the tolerability and risk profile of the tricyclics and monoamine oxidase inhibitors. A 2002 Cochrane review found there to be only weak support for their effectiveness in the treatment of depression in dementia. Similarly, a 2011 review of 7 studies with 330 participants found that while 2 studies found antidepressants to have benefit over placebo, overall, efficacy was only suggested, and not confirmed. There is no data clearly favoring any specific antidepressant; however, because of its very long half-life in the elderly, fluoxetine (Prozac) is not generally recommended. There are no controlled trials of electroconvulsive therapy in depression of dementia.

Sleep Disturbance. Sleep disturbance is common in dementia, with individuals having decreased nocturnal sleep time, increased sleep fragmentation and increased daytime sleepiness. A dual approach of enhancing sleep at nighttime, through improved sleep hygiene and decreasing nocturnal interruptions, like light and sound, coupled with increasing daytime wakefulness though social and physical activity is important. One small study looking at the effects of trazodone showed improvements in nocturnal sleep time. Melatonin and ramelteon have not been found to be efficacious. There have been no randomized controlled trials of benzodiazepines or the nonbenzodiazepine sedative hypnotics for the elderly with dementia, but because of their potential for cognitive and gait impairments, they are not recommended.

DEMENTIA SUBTYPES

Alzheimer Disease

Definition. Dementia of the Alzheimer type or Alzheimer disease is the most common form of dementia. The onset and steady progression of Alzheimer disease follows a characteristic pattern, and a probable diagnosis can be made on the basis of clinical history. There must be decline in memory and learning and in one other cognitive domain for the diagnostic criteria to be met. Alzheimer disease is associated with specific neuropathological changes seen in the brain on histological examination and a definitive diagnosis can be made only at autopsy. While there is a better understanding of Alzheimer disease than many other forms of dementia, there is still much to understand to be able to provide preventative strategies and treatments that can alter the disease course. There have been tremendous advances in the last decade in the understanding of the pathophysiology and genetics of Alzheimer disease. However, treatments are limited to symptomatic relief and modest delays of progression. Given the high prevalence in older age and advancing age of the population, Alzheimer disease is a growing problem affecting the lives of not only the sufferers, but also their families.

History. Alois Alzheimer first described the illness in 1906 when he reported the case of a 51-year-old woman with a progressive cognitive disorder marked by memory loss, confusion, topographical disorientation, aphasia, and prominent neuropsychiatric manifestations of persecutory delusions, misidentifications, and behavioral disturbances. At her death 4 years later, autopsy revealed profound atrophy, amyloid plaques, and neurofibrillary tangles.

Comparative Nosology. According to the DSM-5 a diagnosis of probable Major Neurocognitive Disorder due to Alzheimer Disease is made when the criteria for major neurocognitive disorder is met, and there is either evidence of a causative genetic mutation or the individual shows impairment in memory and learning, in addition to one other cognitive domain. There must also be a steadily progressive decline, and there is no evidence of another etiology of the cognitive impairment. Diagnosis is further classified into subtypes as with or without behavioral disturbance.

Specific criteria for Alzheimer disease were also defined by the National Institute of Neurological and Communicative Disorders and Stroke—Alzheimer's Disease and Related Disorders Association Work Group (NINCDS-ADRDA) in 1984. These diagnostic recommendations were revised in 2011 by the National Institute on Aging-Alzheimer's Association (NIA-AA) workgroups. In revising these guidelines, three stages of

Alzheimer disease were identified: the preclinical stages of Alzheimer disease, MCI due to Alzheimer disease, and dementia due to Alzheimer disease. The preclinical and MCI stages are addresses in Chapter 10.4. For a diagnosis of probable Alzheimer disease dementia, an individual must first meet the criteria for dementia described above; additionally the symptoms should be of insidious onset, and there must be report of worsening cognition. The diagnosis of probable Alzheimer disease has additional descriptors for the incorporation of documented decline or when it is diagnosed in a carrier of a causative Alzheimer disease genetic mutation. A diagnosis of possible Alzheimer disease dementia can be made when there is an atypical course or an individual is showing signs of a dementia with mixed etiologies. Finally the NIA-AA guideline describe a third category of diagnosis for when the criteria for probable Alzheimer disease dementia is met, and there is biomarker evidence (further described below) supporting the diagnosis. However, at this time, the use of biomarkers is not recommended for routine clinical diagnosis.

Epidemiology. The incidence of Alzheimer disease increases with age from about 0.5 percent per year at ages 65 to 69, 1 percent per year from age 70 to 74, 2 percent per year from age 75 to 79, 3 percent per year from age 80 to 84 and 8 percent per year after age 85. About 11 percent of the US population over age 65 is affected by Alzheimer disease. In 2013, 5.2 million in the United States had Alzheimer disease. With the aging of the population, projections estimate that, by the year 2050, 13.8 million people over the age of 65 are estimated to have Alzheimer disease.

Etiology. The etiology of Alzheimer disease is multifactorial. Mutations or copy number variations within several genes that result in the cleavage of amyloid precursor protein (APP) into Aβ invariably lead to early-onset Alzheimer disease within affected families. The contributions of these genetic findings have been essential to the development of the current understanding regarding the pathophysiological cascade leading from Aβ to Alzheimer disease. In contrast, late-onset Alzheimer disease is a complex disorder, in which risk is influenced by multiple genes, by environmental factors, and by their interactions. A number of putative environmental risk and protective factors for late-onset Alzheimer disease have been identified by clinical studies. Ultimately, determination of the extent to which these putative factors are causally related to, and not just correlated with, Alzheimer disease risk will hinge on how they may interact biologically with the Aβ-induced cascade of pathology in Alzheimer disease.

GENETICS

Early-Onset Alzheimer Disease. Mutations at three genetic loci associated with early-onset Alzheimer disease have been identified. The first identified

Alzheimer disease gene was the APP gene located on chromosome 21. Currently, over 30 different mutations in APP that can cause Alzheimer disease have been identified, in fewer than 100 families worldwide. More recently, families with duplications of APP, leaving the mutation carrier with a total of three copies of the gene on his or her two chromosomes, have been reported. This is similar to the increased risk of Alzheimer disease associated with trisomy 21. A second Alzheimer disease locus was found on chromosome 14, presenilin-1 (PS1). More than 180 mutations in the PS1 locus have now been reported in several hundred families throughout the world. Onset of symptoms for those with PS1 mutations is often before age 50, though later ages are possible. Like the APP mutations, PS1 mutations appear to act as autosomal dominant traits with nearly complete penetrance (i.e., inheriting a single copy of the mutated gene is sufficient to cause disease in all individuals). A third Alzheimer disease gene has been localized to chromosome 1 and termed presenilin-2 (PS2). These mutations appear to be rare; only a few families of different ethnicity have been identified with PS2 mutations. Although the PS2 mutations are autosomal dominant, they may not be fully penetrant. PS2 mutations generally lead to onset of Alzheimer disease symptoms before age 65. There has been conflicting information on the proportion of early-onset cases attributable to mutations in APP, PS1, and PS2. Some reports have suggested that PS1 mutations may account for most early-onset familial Alzheimer disease cases. Others have reported that a more realistic estimate is 20 percent or fewer of such cases, with PS2 and APP accounting for 1 percent and 5 percent, respectively. For the remainder of individuals with early-onset familial Alzheimer disease, it is unclear whether other causative single gene mutations will be identified, or if these will prove to be multigenic in origin.

Late-Onset Alzheimer Disease. The first gene to have an established relationship to late onset Alzheimer disease was apolipoprotein E, with increased risk of Alzheimer disease found in individuals carrying the epsilon 4 allele (APOE4) in both familial and sporadic cases. The risk of Alzheimer disease increases, and the mean age of onset of Alzheimer disease is earlier, as the number of epsilon 4 alleles an individual carries increases from 0 to 2. A single copy of the APOE4 allele confers an approximate two- to fourfold increased risk while two APOE4 alleles increase the risk by approximately four- to eightfold. The magnitude of the risk associated with APOE4 may vary by age, and ethnicity, with the risk attenuated in older individuals, and in African American and Hispanic individuals (relative to Caucasians). In contrast, the risk conferred by APOE4 may be increased in individuals of Japanese ethnicity. It is generally estimated that APOE4 accounts for up to 50 percent of the genetic contribution to late onset Alzheimer disease. Nevertheless, carrying one, or even two APOE4 alleles, is neither necessary

nor sufficient for the development of Alzheimer disease. It has been estimated that for those who are disease-free at age 65, at most 50 percent of APOE4 homozygotes will develop Alzheimer disease within their lifetimes, thus limiting the use of APOE4 genotype as a predictive test. Alternatively, when APOE4 testing is added to the evaluation of individuals clinically diagnosed with cognitive symptoms consistent with a dementia, it tends to lower the sensitivity, but enhance the specificity of the diagnosis of Alzheimer disease.

Genome wide association studies have recently identified 20 additional genes/loci associated with late onset Alzheimer disease; these include *CR1, BIN1, CLU, PICALM, MS4AA/MS4A6E, CD2AP, CD33, EPHA1, ABCA7, HLA-DRB5–DRB1, SORL1, PTK2B, SLC24A4/RIN3, INPP5D, MEF2C, NME8, ZCWPW1, CELF1, FERMT2,* and *CASS4*. Other genetic approaches have identified rare mutations in TREM2 and PLD3 that can cause late onset Alzheimer disease. These recently identified loci have confirmed the importance of previously known pathological pathways in Alzheimer disease (e.g., APP and tau). In addition, they have also re-emphasized the role of inflammation. Much further research will be needed however, before the full implications of these findings for the onset, diagnosis, and management of Alzheimer disease are realized.

AMYLOID- B PROTEIN AND SYNAPSE LOSS. The identification of causative mutations in early-onset familial Alzheimer disease has been a key piece of evidence indicating the causal role of Aβ protein in Alzheimer disease. As noted above, extra copies of APP via mutation or trisomy of chromosome 21 result in Alzheimer disease. It is noteworthy that, as assessed in animal and in vitro systems, the expression of the identified mutations in APP, PS1, and PS2 that lead to early-onset familial Alzheimer disease all result in overproduction and/or increased aggregation of Aβ. These genetic findings have also provided the key to understanding how APP is metabolized in the brain into Aβ. In brief, APP protein is cleaved by three enzymes, α-, β-, and γ-secretase (Fig. 10.3–1). Cleavage by β-secretase then γ-secretase generates Aβ, while cleavage by α-secretase yields a nontoxic peptide. The identification of the Swedish mutation in APP, which selectively enhanced metabolism of APP by β-secretase, provided the needed assay for the identification of β-secretase. Similarly, the identification of Alzheimer disease–causing mutations in the novel proteins PS1 and PS2 led to a search for their function. These efforts ultimately converged on their identification as components of the proteolytic site in γ-secretase, an enzyme that is now understood to be a complex of multiple proteins.

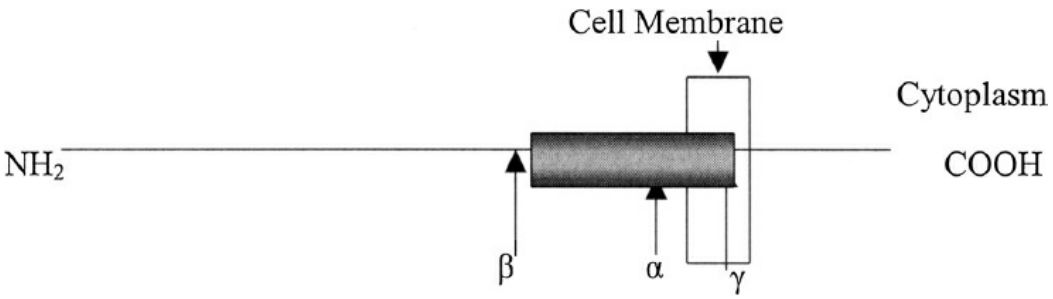

FIGURE 10.3–1. Amyloid precursor protein is a transmembranous protein. Approximate location of cleavage sites for α-, β-, and γ-secretase is indicated by the respective arrows. Cleavage of amyloid precursor protein by α-secretase results in a soluble end product. Cleavage by β- and γ-secretase forms the 40- to 42-amino acid, amyloidogenic, Aβ protein.

More recently, the pathway from overproduction of Aβ to the development of cognitive deficits has been substantially elaborated. It has been appreciated for a number of years that the strongest correlate of cognitive impairment in individuals with Alzheimer disease is loss of synapses across cortical regions. Substantial evidence now indicates that self-aggregation of Aβ into soluble low-n oligomers is a primary source of synaptotoxicity in Alzheimer disease. In vitro observations of the deleterious effects of soluble Aβ oligomers on synapses have been complemented by findings from animal and human postmortem studies. This includes evidence from animals transgenic for mutant human APP that deficits in synaptic structure and function precede deposition of insoluble Aβ into plaques and correlate with cortical soluble Aβ levels, but do not correlate with plaque numbers or APP levels. Similarly, human studies indicate that cortical synapse loss is an early pathological event and that cognitive impairments and synapse loss correlate most strongly with soluble Aβ, even in subjects with early disease.

The above cascade has important implications for many current efforts at development of disease-modifying treatments for Alzheimer disease. Efforts are currently being directed at reducing Aβ production via inhibition of β- and γ-secretase. Similarly, enhancement of α-secretase would be expected to be protective, a mechanism that may underlie reports of possible statin-induced reductions in Alzheimer disease risk. Other approaches influenced by this pathological cascade have focused on increasing clearance of Aβ from the brain. One noteworthy (and newsworthy) approach used immunization against Aβ. Vaccines were developed and found to be effective in transgenic mice at decreasing plaque. Human studies were stopped, however, because of serious CNS inflammatory reactions and encephalopathies. The recognition of the importance of soluble, rather than insoluble, Aβ has enabled an alternate strategy. Attempts to treat individuals with dementia via infusion of Aβ antibodies, which by sequestering of peripheral soluble Aβ can lead to diffusion of Aβ out of the brain, were largely unsuccessful. However, this

approach is now being tested for prevention of cognitive decline in individuals at risk for dementia due to Alzheimer disease, either because they carry a disease-causing mutation, or because of MCI in the setting of positive Aβ biomarkers.

In addition to the direct effects of Aβ on synapse loss, Aβ may lead to synapse loss and neuronal death via other downstream effectors. For example, Aβ contributes to the hyperphosphorylation of microtubule-associate protein τ (MAPT), which enhances its aggregation. Overproduction and aggregation of MAPT have been shown to cause synapse loss and neuronal death independent of the effects of Aβ. It is also clear that soluble Aβ oligomers exist in equilibrium with fibrillar deposits of Aβ in plaques. Plaques themselves form substantial space-occupying lesions in the cortex of individuals with Alzheimer disease and serve as a site for inflammatory responses. Both these mechanical and inflammatory processes may further contribute to synapse loss and neuronal death. As such, the development of inhibitors of aggregation of Aβ into fibrils might prevent this secondary process, and are currently under development. Aggregation of α-synuclein also occurs in some individuals with Alzheimer disease, and can contribute independently to cognitive impairments. At present, however, whether the aggregation of α-synuclein is enhanced by a mechanism related to Aβ is not known.

As the pathological cascade in Alzheimer disease progresses, other brain changes may become evident. For example, numerous neurochemical abnormalities have been demonstrated on autopsy and with functional neuroimaging in Alzheimer disease in later disease stages. These include deficits in acetylcholine (ACh), norepinephrine, serotonin, dopamine, γ-aminobutyric acid (GABA), glutamate, corticotrophin-releasing hormone, and somatostatin. The most robustly described neurochemical deficits are in the cholinergic system. The activity of choline acetyltransferase, the enzyme responsible for the final step of ACh synthesis, is substantially reduced in patients with Alzheimer disease. This has been confirmed on autopsy of patients with at least a moderate stage of the illness but not in patients with mild illness. This is particularly evident in the nucleus basalis of Meynert and several neocortical regions. This finding has led to the successful development of efficacious therapies directed at the enhancement of cholinergic function.

As excitatory synapses onto dendritic spines are particularly affected by Aβ-induced synaptic loss, it is not surprising that reductions in glutamate receptors, which are predominantly located at postsynaptic sites on dendritic spines, have been described in Alzheimer disease. The relationship between Aβ-induced synapse loss and glutamate receptors is complex. Aβ appears to destabilize dendritic spines, resulting in retraction and synapse loss, via modulation of the function of both α-amino-3-hydroxy-5-methyl-4-isoazole-proprionic acid (AMPA) and *N*-methyl-D-

aspartate (NMDA) classes of postsynaptic glutamate receptors. It is therefore not surprising that memantine (Namenda), a low-affinity NMDA antagonist that functions to modulate activation of NMDA receptors, has been shown to be of benefit in Alzheimer disease.

RISK AND PROTECTIVE FACTORS. A number of other factors have been related to Alzheimer disease risk in clinical case-control and epidemiological studies. In all cases it is unclear to what extent these putative factors represent true causal risk factors, consequences of the Aβ-driven pathological cascade, or correlated epiphenomena. Even those factors for which there is substantial evidence of biological plausibility of a possible causal role (e.g., the effect of oxidative stress on neuronal survival), still lack either convincing evidence that they occur early enough in the pathological cascade to be contributory or convincing evidence that they are specifically related to the Alzheimer disease process and not a consequence of some alternate, correlated process.

Age. Age is the greatest risk factor for developing Alzheimer disease. Whereas approximately 3 percent of people older than 65 years of age have Alzheimer disease, about one third of people older than 85 years of age have Alzheimer disease. Whether this association represents a correlated observation or whether there are specific age-related processes that enhance pathological processes in Alzheimer disease is not known.

Education. Epidemiological studies have suggested that people with less education are at increased risk of developing Alzheimer disease. Functional neuroimaging findings of patients with Alzheimer disease indicate that people with higher educational attainment or higher premorbid intellectual functioning but similar levels of dementia severity have more severe findings on neuroimaging. This suggests a protective effect of education, in that greater neuropathology is needed to produce the same clinical changes. Those with higher levels of cognitive reserve may be able to utilize other alternative neural networks as levels of neuropathology increase, thus compensating for the disease burden, up to a point. Studies have reported that there is a relationship between cognitive decline and neocortical synaptic density in Alzheimer disease. There is also some evidence that education can increase synaptic density.

Estrogen. A number of observational studies have suggested that estrogen replacement after menopause reduced the risk of Alzheimer disease. On a neuronal level, estrogen increases density of a synaptic population lost in Alzheimer disease, excitatory synapses onto dendritic spines, in both hippocampus and cortex. It also may interact with APP processing, to increase brain Aβ concentrations. Estrogen inhibits oxidation of lipids, lipoproteins, and nucleic acids in vitro and is protective of cells in culture against a number of insults, including Aβ. However, prospective studies of

estrogen treatment in patients with Alzheimer disease have yielded little evidence of cognitive enhancement or reduced cognitive decline. Similarly, results from the single largest randomized clinical trial of hormone replacement therapy, the Women's Health Initiative, indicated no prevention of cognitive decline with estrogen replacement therapy in women aged 65 years or older and evidence of increased incidence of dementia in those women receiving combined estrogen/progesterone replacement therapy. The argument has been made that the timing of the hormone replacement (i.e., more effective during the perimenopausal or very early postmenopausal periods) may be the critical variable in this; however, current recommendations are that hormone therapy not be used in the prevention of dementia.

Head Trauma. Traumatic brain injury can increase the rate of cognitive decline in the elderly, and is associated with an increased risk of developing dementia, and in particular Alzheimer disease. There is evidence that head trauma can lead to overexpression of APP, increase inflammatory mediators, and cause Aβ deposition in the brain. The effects of traumatic injury on Aβ may be mediated by APOE genotype (see Chapter 10.5 for a discussion of chronic traumatic encephalopathy).

Inflammation. Several lines of evidence suggest a role of the immune system in Alzheimer disease, including findings from recent genetic studies described above. Neuropathological studies demonstrate that the brains of patients with Alzheimer disease have increased concentrations of acute phase reactants, cytokines, and complement protein, when compared to aged-matched controls. Whether these changes are present in earlier disease states has not been well investigated. Population studies have demonstrated that nonsteroidal anti-inflammatory drug (NSAID) or corticosteroid use is associated with a decreased risk of developing Alzheimer disease; however, a 2005 meta-analysis of 25 case-control and cohort studies found that many of the positive results reported may have been due to a combination of recall, prescription, and publication bias. A recent review evaluating the role of NSAIDs in the prevention of dementia also showed an association in observational studies between NSAID use and a reduced dementia risk; however, the one randomized controlled trial that was included showed no effect of NSAID use on the prevention of dementia. Results from one randomized controlled trial indicate that timing of NSAID use may be important, with NSAID use possibly having an adverse effect on individuals in later stages of Alzheimer disease pathogenesis, but who have not yet clinically progressed to dementia. Use of NSAID for prevention of Alzheimer disease is most likely precluded by offsetting risks of bleeding and cardiac toxicity.

Nicotine. The literature on nicotine is unclear. Initial case-control studies

suggested a reduced likelihood of Alzheimer disease in smokers. In contrast, cohort studies evaluating Alzheimer disease incidence have found the opposite relationship, with smoking associated with an increased risk of developing Alzheimer disease dementia. There have been conflicting results of neuropathological studies in human smokers, while animal studies have shown increased Aβ burden. Smoking may also indirectly modulate Alzheimer disease neuropathology via oxidative stress and indirect effects on neuroinflammation. It is noted, however, that indirect stimulation of nicotinic receptors via use of acetylcholinesterase inhibitors has benefit in the treatment of Alzheimer disease. There is also *in vitro* evidence that nicotine may protect against Aβ toxicity. Because nicotine treatment carries substantial cardio- and cerebrovascular risks, it is unclear how these observations may translate into therapies.

Oxidative Stress. Oxidative stress has been postulated to contribute to Alzheimer disease. For example, Aβ can damage and kill neurons by inducing oxidative stress, as can the microglial inflammatory reaction surrounding plaques. Increased markers of oxidative stress have been found in the brains of subjects with MCI, suggesting that these processes may be altered relatively early in the disease course. Thus, antioxidants have been evaluated for evidence that they are neuroprotective. However, a 2011 Cochrane review examining the effects of the antioxidant vitamin E in three studies found no convincing evidence of its beneficial effects on the treatment of Alzheimer disease or MCI. Studies have also not provided sufficient evidence to support treatment with the antioxidant gingko biloba.

Diagnosis and Clinical Features. As described above, Alzheimer disease dementia has a gradual and insidious onset, and there must be two cognitive domains affected, as well as functional impairment. The most common presentation is an amnestic one, in which the primary deficits are in the learning and memory domain. There are also nonamnestic presentations, where the prominent impairment could be in the language, visuospatial or executive functioning domains. Individuals will have increasing difficulty functioning in their daily activities as the disease progresses.

Pathology and Laboratory Examination. There are no standard laboratory tests that are sensitive and specific enough to use in diagnosing Alzheimer disease. Rather, the usual battery of tests indicated in the evaluation of dementia (refer to Table 10.3–5) to exclude other diagnoses should be used. Because of the interest in developing disease modifying treatments, with the ultimate goal of prevention, identification of biological markers has been at the forefront of research.

Routine clinical diagnostic testing includes the use of qualitative structural imaging (CT or MRI), largely to identify or exclude other causes

of cognitive impairment. Quantification of change in global (whole brain and ventricles) and regional (entorhinal cortex, hippocampus, corpus callosum) volumes in patients with Alzheimer disease frequently demonstrate group differences in comparison to cognitively intact elderly patients, with MCI patients intermediate (see Fig. 10.3–2 for an example MRI in a patient with Alzheimer disease). Currently, SPECT and FDG-PET are not established as stand-alone diagnostic tests. However, the addition of SPECT or FDG-PET can increase the diagnostic accuracy of an expert evaluation, and can be particularly useful in differentiating individuals with Alzheimer disease from those with an FTD.

Several agents for PET-amyloid imaging have recently been approved for use. The absence of tracer retention during PET-amyloid imaging indicates that dementia is unlikely to be due to Alzheimer disease (Fig. 10.3–3). Though not well standardized across laboratories, decreased cerebrospinal fluid levels of the 42 amino acid species of Aβ ($A\beta_{1-42}$) is supportive of a diagnosis of Alzheimer disease as are elevated cerebrospinal fluid levels of total and phosphorylated tau protein.

The use of the genetic marker APOE4, as a diagnostic or predictive test, was discussed above in the section on Genetics of Late Onset Alzheimer disease.

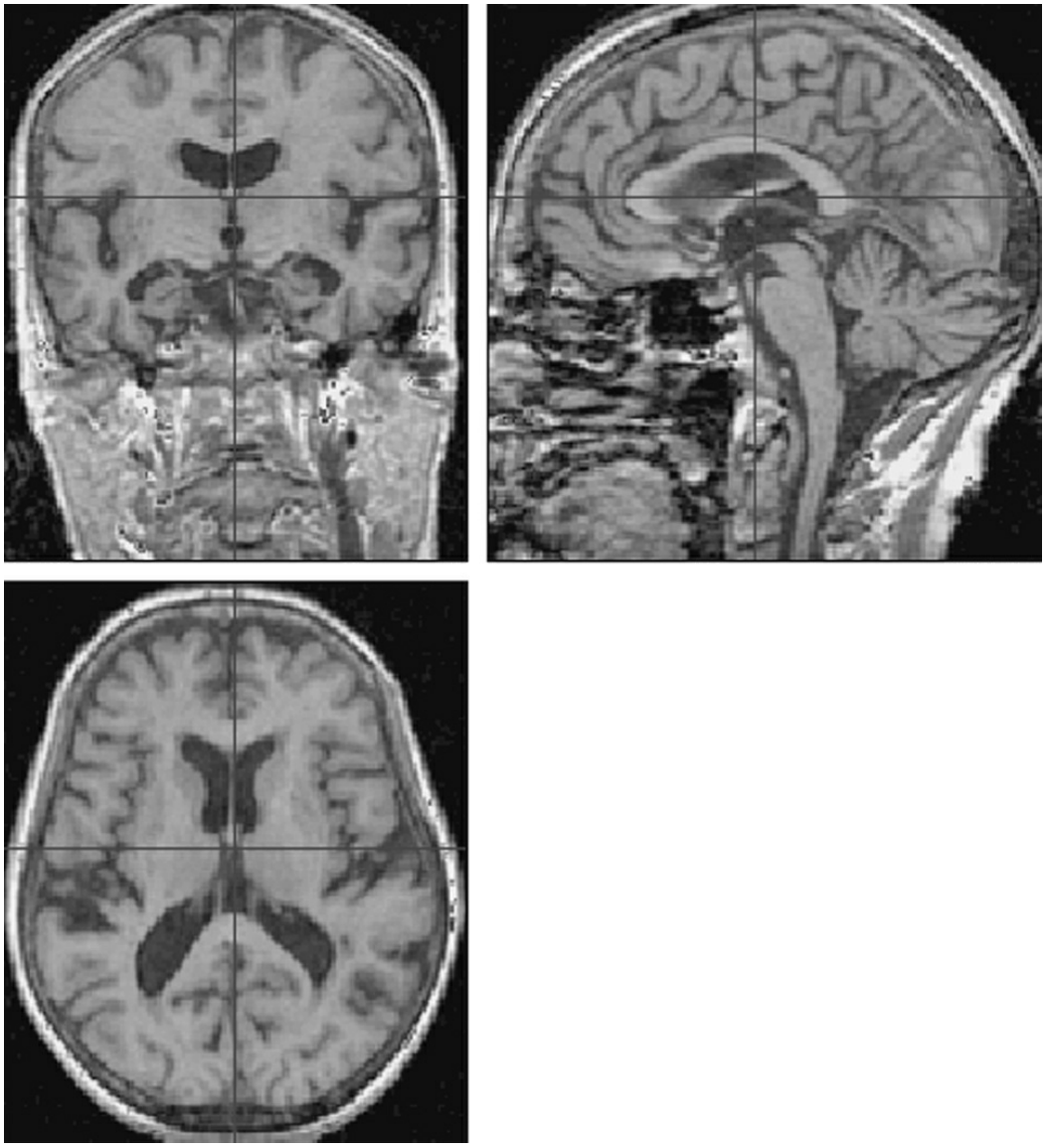

FIGURE 10.3–2. Magnetic resonance image of an individual with Alzheimer disease. Three plane views of the brain of a patient with probable Alzheimer disease at the time of their entry into the University of Pittsburgh Alzheimer Disease Research Center. Note there is significant atrophy in the hippocampus and related structures in the mesial temporal lobe (seen in the upper left image), and cortical atrophy in the frontal and temporal neocortex.

PATHOLOGY. Macroscopically, the brains of early Alzheimer disease patients may appear grossly normal, or with atrophy confined to the hippocampus, and often most easily recognized by increased volume of the inferior horn of the lateral ventricles. As the disease progresses, widened sulci throughout the cortex and increased ventricular size are seen. Atrophy in the posterior temporal, parietal, and frontal lobes is often most prominent.

Microscopically, the characteristic lesions of Alzheimer disease are amyloid plaques and neurofibrillary tangles. Amyloid plaques are the defining lesion of Alzheimer disease, composed of insoluble deposits of fibrillar Aβ (Fig. 10.3–4). Surrounding this central core are dystrophic

neurites, microglial cells, and reactive astrocytes. Amyloid plaques are first found throughout the neocortex, followed by entorhinal cortex and hippocampus, then subcortical regions.

The second major histopathological feature is the neurofibrillary tangle. Neurofibrillary tangles are intracellular inclusion bodies, which contain paired helical filaments composed of aggregates of hyperphosphorylated MAPT (Fig. 10.3–4). Neurofibrillary tangles are first found in the entorhinal cortex and spread from there to the hippocampus and lateral temporal lobe, then more widely through the neocortex. The presence of neurofibrillary tangles is not pathognomonic for Alzheimer disease, as they are the principal lesion in MAPT-related FTDs.

It has long been appreciated that a limited degree of Aβ deposits and neurofibrillary tangles can also be found in many elderly subjects with normal cognition. It is possible that these represent cases in which individuals died during the preclinical stage. Alternatively, protective factors may have mitigated against these individuals ever expressing dementia. As a result, current neuropathological diagnostic criteria for definite, or a high likelihood of, Alzheimer disease require the presence of a clinical history of progressive dementia during life and the presence of both amyloid plaques and neurofibrillary tangles on postmortem examination.

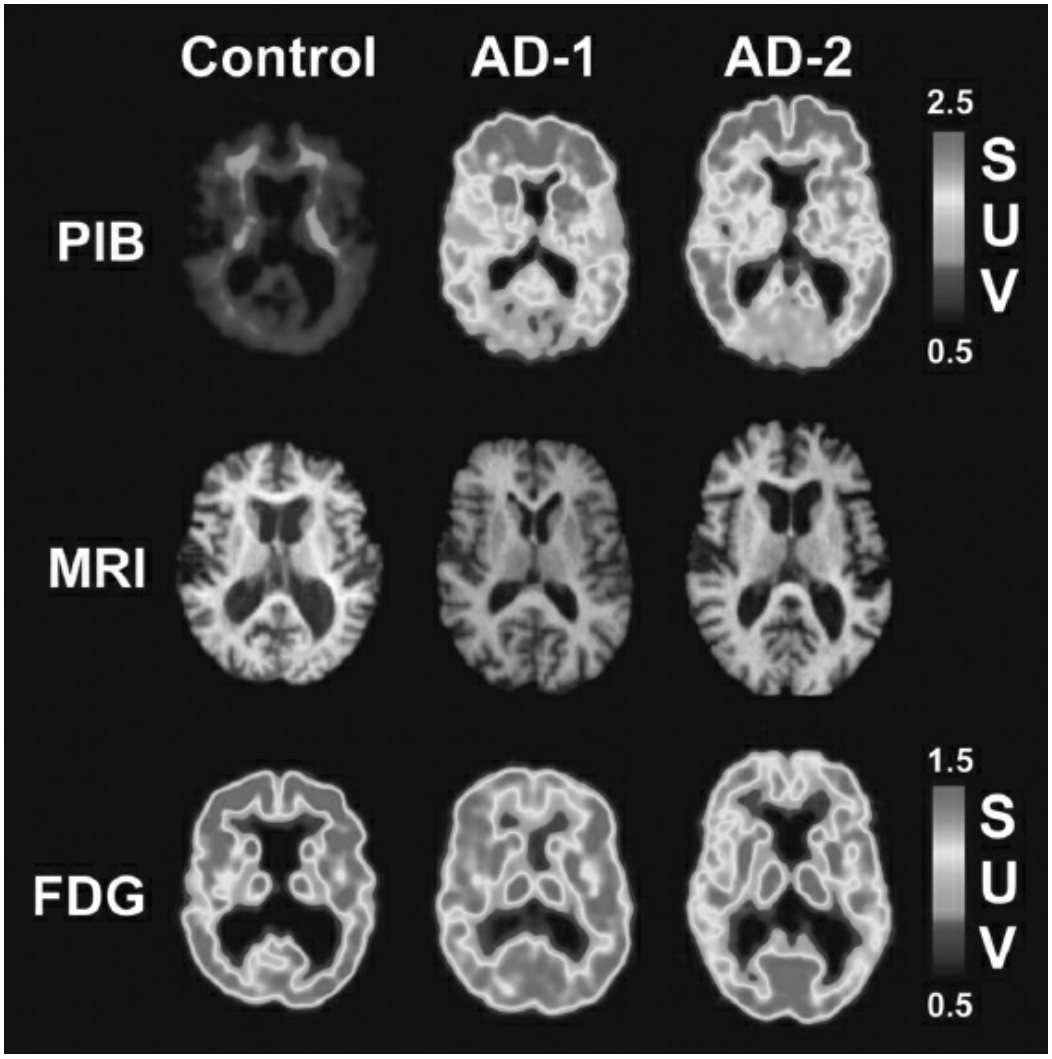

FIGURE 10.3–3. Pittsburgh Compound B (PIB) binding in normal aging and Alzheimer disease (AD). Comparison of PIB images (*top row*), magnetic resonance images (MRIs) (*center row*), and fluorodeoxyglucose (FDG) images (*bottom row*) from a cognitively normal subject (71-year-old patient; Mini-Mental State Examination [MMSE] score of 30) and two mild AD cases. AD-1 (*center column*) was 68-year-old patient, had an MMSE score of 26, showed marked PIB accumulation compared to controls, but a normal FDG scan. AD-2 (*right column*) was 69-year-old patient, had an MMSE score of 21, also showed marked PIB accumulation, and an FDG scan typical for AD showing temporoparietal and frontal hypometabolism. Scale bars indicate standardized uptake values (SUVs) for PIB and FDG. Note that the dynamic range of PIB is twice that of FDG. (Courtesy University of Pittsburgh Amyloid Imaging Group.)

Additional pathological features of Alzheimer disease include abnormal aggregation of α-synuclein, which may be detectable in as many as one half of individuals with Alzheimer disease. These aggregates most commonly are found as thread-like deposits or as intraneuronal inclusions (Lewy bodies) in the amygdala and transentorhinal cortex, though Lewy bodies may also be present in the neocortex. Transactive response DNA-binding protein of 43 kDa (TDP-43) is the most frequently aggregated disease protein in frontotemporal lobar degeneration. Recently it has been found that TDP-43 proteinopathy is also present in many cases of Alzheimer

disease. When present, abnormal TDP-43 aggregates are found predominantly in the mesial temporal lobe, and to a lesser degree in neocortical areas. The major pathological protein aggregates found in Alzheimer disease, and the overlap with other neurodegenerative dementias is summarized in Table 10.3–6.

Other neuropathological abnormalities described in this disease include neuronal loss, most markedly in the entorhinal cortex and hippocampus. Later in disease, loss of cholinergic neurons in the nucleus basalis, as well as loss of brain stem noradrenergic and serotonergic neurons, occurs. Granulovacuolar degeneration and the presence of Hirano bodies, most commonly affecting pyramidal neurons in the hippocampus, may be found.

Course and Prognosis. Alzheimer disease generally presents with a gradual decline initially in short-term memory and language function. Initially, there is MCI, which may not be recognized by family members or be clinically relevant as there is no significant functional impairment. Gradually these impairments in short-term memory and language worsen until they start to interfere with ability to perform instrumental activities of daily living.

Typically, several years pass between onset of early symptoms and recognition of deficits and presentation to medical attention. There is often a precipitating factor that alerts families to the cognitive impairment and prompts them to seek medical attention. The precipitating event can be an intercurrent medical illness that causes delirium, at which time families will look back and in retrospect recognize early symptoms of decline in the past. There may be a loss, such as death of a spouse, that unmasks deficits in functions that had previously been performed for the patient by the deceased. Sometimes symptoms are identified during a change of residence, after which the affected individuals are required to learn their new environment and are unable to meet these demands, calling their families' attention to their impairment.

After initial presentation, a gradual progressive decline in cognitive function ensues and extends eventually to all cognitive domains including motor functioning (apraxia) and perception (agnosia). The rate of decline is not necessarily steady as there may be periods of plateau followed by periods of accelerated decline.

As the illness progresses, there will be greater functional impairment and greater dependency needs. The attendant loss of independence may be very difficult for the Alzheimer disease sufferer to accept; depressive symptoms may be evident if their level of insight is not fully impaired. Alternatively if there is agnosia and a lack of insight into their deficits and impairment, the acceptance of additional help for their functional impairments can be quite difficult. Safety issues become of great concern, especially such things as leaving on the stove and burning food or pots or

causing a fire, or driving and the increased risk for traffic accidents. Patients may forget to eat or take their medications, thus becoming malnourished and exacerbating their concurrent medical problems. The level of care and supervision required will increase over time as the illness progresses. Initially daily checking in may suffice. Later, help several hours per day, particularly around meals, medications, bathing and dressing is often required. As the disease progresses, delusions, hallucinations, and agitation become more prevalent. Eventually total care and constant supervision for assurance of safety is needed.

In the final stages of Alzheimer disease, parkinsonian symptoms may be present, mobility and gait are impaired, and the risk of fall is great. Eventually patients will be bedridden. Swallowing impairments are common. Patients will often succumb to pneumonia or other infections. Palliative care is helpful in the end stage to provide comfort measures and support for the family to ensure a peaceful and dignified death.

Treatment

CHOLINESTERASE INHIBITORS. The first class of medications approved by the U.S. Food and Drug Administration (FDA) for the treatment of Alzheimer disease was the cholinesterase inhibitors: Tacrine (Cognex), donepezil (Aricept), rivastigmine (Exelon), and galantamine (Reminyl). These agents are reversible inhibitors of the enzyme acetylcholinesterase, which degrades ACh in the synaptic cleft, thereby increasing the availability of intrasynaptic ACh.

These agents have all been shown in multiple randomized, double-blind, placebo-controlled trials to slow cognitive decline in mild to moderate Alzheimer disease. There are limited trials in severe Alzheimer disease. Donepezil is the only cholinesterase inhibitor FDA approved for the treatment of severe Alzheimer disease. All of these agents show similar efficacy, and there is insufficient evidence comparing these agents to each other. There are no data on whether or how to switch from one agent to another.

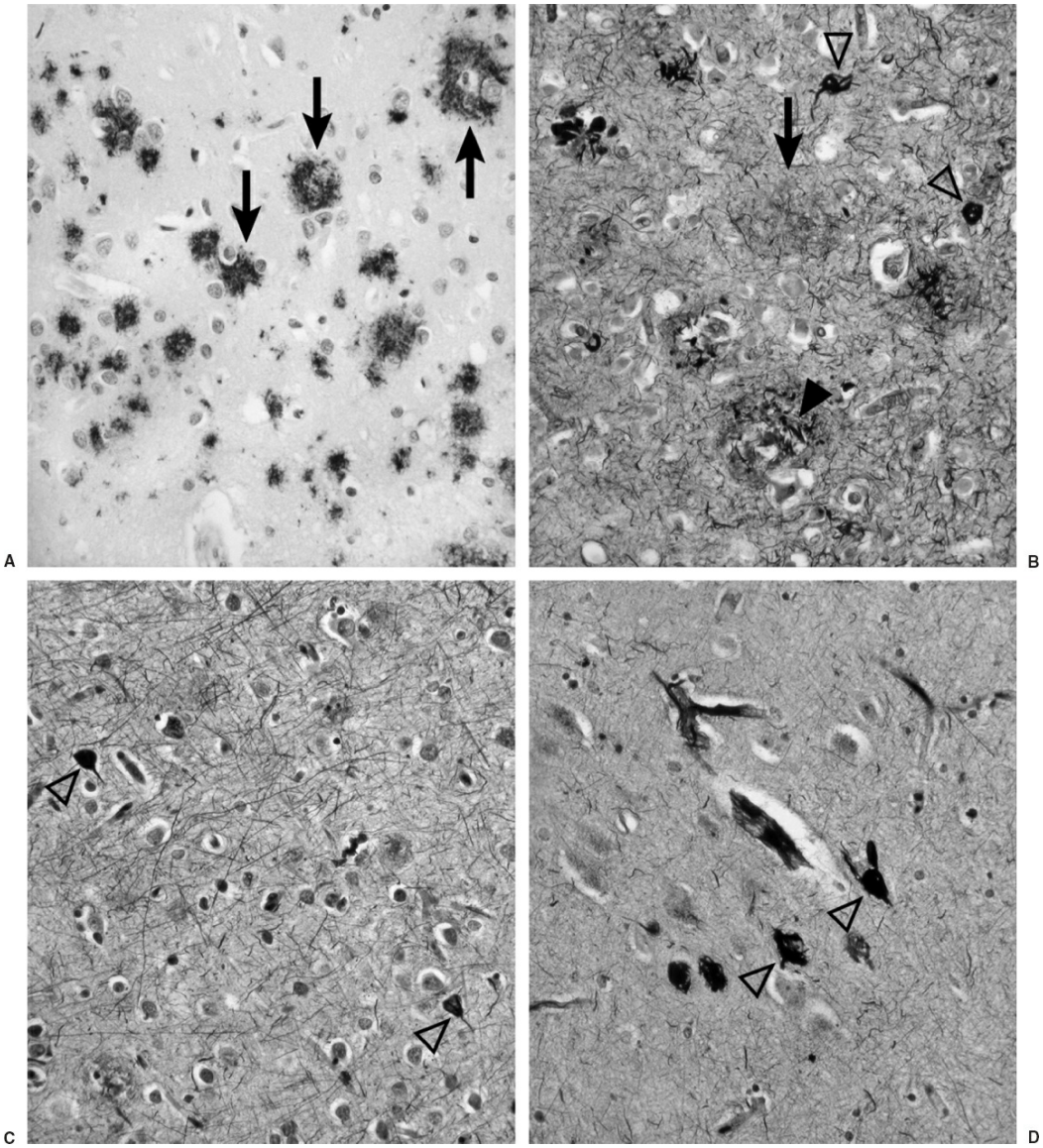

FIGURE 10.3–4. Photomicrographs of Alzheimer disease neuropathology. **(A)** Deposition of insoluble fibrillar Aβ into plaques begins in the neocortex, labeled here using an antibody against Aβ and appearing as reddish brown deposits (*arrows*). **(B)** Bielchowsky stain of neocortex from an individual who died in advanced stages of Alzheimer disease (Braak stage VI). The Aβ plaques appear as dark brown in this preparation (*arrows*), and can be seen to be associated to varying degree with dystrophic neuronal processes (*arrowheads*) in which insoluble microtubule-associate protein τ (MAPT) aggregates appear as black deposits. This neurofibrillary pathology also appears extensively throughout the neuropil, and several neurofibrillary tangles can be seen (*open arrowheads*). **(C)** Bielchowsky stain of neocortex from an individual who died in a less advanced disease stage (Braak stage IV). Though some neurofibrillary tangles are still evident (*open arrowheads*), the degree of neurofibrillary pathology in the neuropil is substantially diminished. **(D)** Isolated neurofibrillary tangles (*open arrowheads*) in entorhinal cortex that can be seen in normal aging (Bielchowsky stain). Notice the lack of Aβ plaques and limited neuropil involvement. (All images obtained at 200× magnification, and provided courtesy of Dr. Ronald L. Hamilton, Department of Pathology, Division of Neuropathology, University of Pittsburgh School of Medicine.)

Table 10.3–6.
Pathological Proteins in the Neurodegenerative Dementias

Dementia Type	Pathological Protein			
	Amyloid Beta	Tau	Alpha-Synuclein	TDP-43
FTD due to *C9ORF72*, *GRN*, *VCP*, *TARDBP*, and *FUS* mutations	No	No	No	Yes
FTD—Pick disease and due to *MAPT* mutations	No	Yes	No	No
Corticobasal Degeneration	No	Yes	No	No
Progressive Supranuclear palsy	No	Yes	No	No
Alzheimer disease	Yes	Yes	Present in up to 60% of cases	Present in 19–57% of cases
Dementia with Lewy bodies	Yes	Yes	Yes	No
Parkinson disease dementia	Yes	Yes	Yes	No

Common side effects are related to cholinergic excess and include nausea and vomiting, decreased appetite and weight, increased gastric acid secretion, muscle cramps, bradycardia with syncope, and sleep disturbances. They may induce or exacerbate urinary obstruction, worsen asthma or chronic obstructive pulmonary disease, and cause seizures. Tacrine can cause reversible hepatotoxicity; because of this common risk and the four times per day dosing, tacrine is now rarely used.

Cholinesterase inhibitors should be used with caution in patients with gastric ulcer disease, sick sinus syndrome or conduction defects, asthma or chronic obstructive pulmonary disease, or cerebrovascular disease. The only contraindication is hypersensitivity. The tolerability of these different agents is variable, so if patients do not tolerate one agent, an alternative should be tried. If gastrointestinal side effects are particularly problematic, individuals can be switched to transdermal patch preparations (see Table 10.3–7 for dosing guidelines).

Patients should be titrated gradually to the maximally tolerated dosage as there is a dose-dependent response. If there is a continued rapid decline despite the medication, it can be discontinued. However, if there is stabilization of symptoms or slowing of decline, which is a desired response, the medication should be continued. Duration of treatment is unclear. Placebo-controlled studies demonstrate efficacy up to 6 months to 2 years depending on the agent. Open-label extension studies suggest ongoing efficacy beyond that time, but there are methodological flaws to these studies. Discontinuation of cholinesterase inhibitors has been associated with marked deterioration and regression of cognition to levels of the placebo control group. There is some suggestion that after cholinesterase treatment has been interrupted for 6 weeks or more before being restarted, patients do not return to the cognitive level they were at before the medication was discontinued. Case reports have described intense hallucinations, difficulty sleeping and concentrating, agitation, changes in mood, and delirium with abrupt discontinuation of donepezil.

MEMANTINE. Memantine is a noncompetitive NMDA receptor antagonist that may modulate excitatory signaling and have neuroprotective actions.

Memantine has been studied in mild, moderate, and severe stages of Alzheimer disease. It is FDA approved for use only in moderate to severe Alzheimer disease and has demonstrated efficacy for both cognitive and functional improvement over placebo. It has been studied as monotherapy and in addition to stable doses of donepezil and is effective and well tolerated when added to the cholinesterase inhibitor. There are no data for use longer than 6 months duration. It is generally well tolerated with drop-out rates comparable to placebo. Infrequent adverse events include confusion, dizziness, headache, sedation, agitation, falls, and constipation (see Table 10.3–7 for dosing guidelines).

OTHER MEDICATIONS. Other medications, including vitamin E, gingko biloba, NSAIDs, hormone replacement therapy, selegiline, ergoloids, and anti-amyloid therapies, have been studied for the prevention and/or treatment of cognitive decline in Alzheimer disease but placebo-controlled trials have not shown sufficient efficacy.

Table 10.3–7.
Pharmacological Treatments for Cognitive Impairment

Class	Agent	Starting Dose	Titration	Effective Dose	Maximum Dose
Cholinesterase inhibitor	Donepezil	5 mg	5 mg after 4–6 weeks	5–10 mg	23 mg
	Galantamine	4 mg b.i.d.	4 mg b.i.d. every 4 weeks	16–24 mg daily divided b.i.d.	24 mg daily divided b.i.d.
	Galantamine extended release	8 mg daily	8 mg every 4 weeks	16–24 mg daily	24 mg daily
	Rivastigmine	1.5 mg b.i.d.	1.5 mg b.i.d. every 2 weeks	3–12 mg daily divided b.i.d.	12 mg daily divided b.i.d.
	Rivastigmine transdermal patch	4.6 mg/24 hours daily	Increase to 9.5 mg/24 hours daily after 4 weeks	4.6–9.5 mg/24 hours daily	9.5 mg/24 hours daily
NMDA-receptor antagonist	Memantine	5 mg	5 mg weekly divided b.i.d.	20 mg divided b.i.d.	20 mg divided b.i.d.
	Memantine extended release (Namenda-XR)	7 mg	7 mg weekly	28 mg	28 mg

NMDA, *N*-methyl-D-aspartate.

As discussed in the section on etiology above, a number of treatments that modify the generation or clearance of Aβ, and may therefore modify disease progression, have been tested in individuals with dementia due to Alzheimer disease, and have yet to establish a clear benefit in this stage of disease, and/or were associated with unacceptable toxicities. Additional studies of these agents are in various phases of clinical investigation.

Frontotemporal Dementia

Dementia associated with atrophy of the frontal and temporal lobes encompasses both sporadic and familial diseases. The hallmark of FTD is a progressive change in personality and behavior, with variable degrees of language and other cognitive impairment. Especially when it presents in midlife, it can be a significant clinical challenge to distinguish FTD from primary psychiatric disorders. In late-life presentations, it can be difficult to distinguish from a frontal-predominant presentation of Alzheimer disease.

Definition. FTD as used here is meant to identify the predominant clinical syndrome associated with the group of dementias characterized by progressive circumscribed atrophy of frontal and temporal lobe cortices, identified as frontotemporal lobar degeneration. FTD includes two broadly accepted subgroups: behavioral-variant FTD (bv-FTD); and language predominant FTD (including progressive nonfluent aphasia and semantic dementia). The spectrum of presentations of frontotemporal lobar degeneration also includes the motor syndromes of CBD, PSP, and motor neuron disease, as there are both clinical and pathological overlaps among these.

History. Arnold Pick in the late 19th and early 20th century provided the initial descriptions of the clinical syndrome of a dementia with associated circumscribed atrophy of the frontal and temporal lobes. Alzheimer subsequently described associated neuropathological abnormalities including the presence of intraneuronal inclusions and ballooned neurons. The inclusions were later named Pick bodies, and the syndrome of dementia with frontotemporal atrophy and Pick bodies was named Pick disease.

Comparative Nosology. Consensus criteria for the diagnosis of FTD were first published in 1998. This group recognized that the prototypic clinical presentation of *FTD* is disturbances of personality and behavior or impairments of language (primary progressive aphasia or semantic dementia).

The diagnostic criteria for the bv-FTD were updated most recently in 2011, by the International Behavioral Variant FTD Consortium. The updated criteria are meant to provide greater sensitivity and more flexibility in the diagnosis of FTD in its earlier stages. For a diagnosis of possible bv-FTD, three of six behavioral and cognitive symptoms must be demonstrated (behavioral disinhibition; apathy or inertia; loss of sympathy or empathy; perseverative, stereotyped or compulsive/ritualistic behavior; hyperorality or dietary changes; and executive or generation deficits with relative sparing of memory and visuospatial functions). A diagnosis of probable bv-FTD, requires the criteria for possible bv-FTD, as well as significant functional decline, and imaging consistent with bv-FTD. In 2011, an international work group published criteria for the language predominant forms of FTD, progressive nonfluent aphasia (also known as nonfluent/agrammatic), semantic dementia, and added a third variant, logopenic. Descriptions of these variants can be found later in this section.

The DSM-5 includes both a behavioral variant and a language variant in its description of Major Frontotemporal Neurocognitive Disorder. The behavioral variant is characterized by decline in social cognition or executive functioning, as well as behavioral disinhibition, apathy or inertia, loss of sympathy or empathy, perseverative, stereotyped or compulsive behavior, and hyperorality. The language variant is characterized by prominent decline in language abilities, with difficulties seen in speech production, word finding, object naming, grammar, or word comprehension. In both variants, there is an insidious onset, with gradual progression, and a relative sparing of learning, memory, and perceptual motor functioning.

Epidemiology. In early age of onset dementia, the prevalence and incidence of FTD has been reported to be similar to that of early onset Alzheimer dementia. Incidence rates have been reported to be 3.5 per 100,000 person years in a 45 to 64 years age group and 3.3 per 100,000 person years in a 50 to 59 years old age group. In population based studies conducted in the United Kingdom and the Netherlands, prevalences of FTD of 3.6 to 15 cases per 100,000 individuals in mid- to late-life have been reported. Age at presentation is most commonly in the sixth decade, but this can range from the third to the ninth decade. The data on the role of gender is mixed. A 2005 study examining the demographic characteristics of FTD did find a male predominance in the behavioral variant and semantic dementia, and a female predominance in progressive nonfluent aphasia.

Etiology. FTD occurs in both sporadic and familial forms. A family history is present in up to approximately 50 percent of cases, typically with an autosomal dominant pattern of transmission. The etiology of the

sporadic forms is currently unknown. Mutations in one of several genes can result in FTD. The most commonly affected gene is *C9ORF72*, in which a hexanucleotide repeat expansion can result in FTD and amyotrophic lateral sclerosis. Approximately 26 percent of familial, and 5 percent of sporadic, FTD cases may carry *C9ORF72* expansions. FTD due to *C9ORF72* most often present as the behavioral variant, and up to 30 percent of cases may include concomitant motor neuron disease. Mutations in *GRN*, found on chromosome 17 that result in loss of function may account for up to 5 to 10 percent of FTD cases. Mutations in the gene for *MAPT*, also located on chromosome 17, similarly result in FTD. *MAPT* mutations are relatively rare as a cause of FTD, accounting for disease in only slightly over 100 families worldwide. Though less common, FTD, has also been reported due to mutations in *CHMP2B, VCP, FUS, and TARDBP*.

Diagnosis and Clinical Features. The temporal evolution of symptoms and their neuroanatomical correlates are important in the diagnostic clarification of the differing variants.

In the bv-FTD, apathy, disinhibition, and loss of self and other awareness are early signs. Family members and loved ones may be concerned that the individual is experiencing a midlife crisis of some sort, and often a psychiatric disorder is suspected as the underlying etiology to this change in behavior early on.

Behavioral disinhibition can be characterized as socially inappropriate behavior or a loss of manners or decorum. An individual may begin to make impulsive, rash or careless actions. Apathy or inertia can also be seen. There may also be an early loss of sympathy or empathy, with an individual displaying diminished reactivity to other's needs or feelings, or diminished social interest, interrelatedness, or personal warmth. Individuals may be less supportive of a spouse's health problems. Perseverative, stereotyped, compulsive or ritualistic behaviors can also be seen. These may range from simple repetitive movements, to more complex or compulsive behaviors, like collecting specific objects, or hoarding. Stereotypy of speech, as well as mental rigidity can also be seen. Individuals may display hyperorality or dietary changes, including altered food preference, such as eating foods that are just one color, or binge eating. Neuropsychological testing reveals executive function deficits, with relative sparing of memory and visuospatial functioning. Individuals with the behavioral variant often have very limited insight into their alterations of behavior.

About 20 percent of individuals with the behavioral variant also show signs of parkinsonism, with an akinetic rigid form present in about 60 percent of these individuals, and tremor predominant present in 40 percent. Five to ten percent of individuals with FTD have co-occurring motor neuron disease as well, with symptoms including muscle wasting, weakness, and fasciculations. The motor neuron disease may occur prior to,

during, or after a diagnosis of FTD has been made. These individuals have ubiquitin and TDP-43 positive inclusions in the lower motor neurons of the spinal cord and the granular cells of the hippocampus. FTD and motor neuron disease are felt to be part of the same clinical spectrum of TDP-43 pathology.

Three different language variants have been described.

Core features of the semantic variant, also known as semantic dementia, include impaired confrontation naming and impaired single word comprehension. Other diagnostic features include impaired object knowledge and surface dyslexia (difficulty with recognition of irregular sounding words, like yacht) or dysgraphia. Individuals may substitute categories for words ("food" for "apple"), and eventually may refer to all nouns as "things." As the disease progresses speech becomes empty of content. Repetition and speech production are usually spared. Those individuals with the semantic variant, but who have right sided predominant atrophy, may show some early behavioral features, including emotional detachment or lack of empathy, or may also have impairments in facial recognition.

The nonfluent/agrammatic variant, also known as progressive nonfluent aphasia, requires one of the two core features for a diagnosis: agrammatism in language production or apraxia of speech. Speech is often described as slow, labored, effortful or halting. Speech sound errors are made and there is abnormal prosody. There may also be impaired comprehension of syntactically complex sentences. Single-word comprehension and object knowledge are spared. Individuals with progressive nonfluent aphasia tend to have more insight into their deficits and maintain social norms; if atrophy progresses to the right frontal regions, behavioral symptoms are likely to follow.

Core features of the logopenic variant include impaired single-word retrieval in spontaneous speech and naming and impaired repetition of sentences and phrases. Speech is slow due to word finding problems. There may be phonological errors in speech and naming. Single-word comprehension, object knowledge, and motor speech are spared, and there is a lack of agrammatic errors, with preservation of articulation and prosody. As some of these individuals go on to show amyloid pathology on autopsy, this variant may actually be a less common variant of Alzheimer disease, particularly if these individuals show a longitudinally stable logopenic pattern.

Pathology and Laboratory Examination. Depending on the clinical variant, there are distinct patterns of atrophy visible on structural imaging, which correlate with the predominant symptoms. In the behavioral variant, there is early degeneration of the paralimbic structures of the ventromedial prefrontal cortex, anterior cingulate cortex and anterior insula. In

progressive nonfluent aphasia, there is degeneration of the dominant frontal operculum through the premotor area, and the insular cortex. In the semantic variant, there is atrophy of the anterior temporal pole, which eventually spreads to the frontal areas, insula and anterior hippocampus.

Currently there are no established biomarkers for the diagnosis of FTD. In practice, use of MRI to establish the presence of circumscribed frontal and/or temporal atrophy (Fig. 10.3–5) is helpful in confirming a suspected diagnosis. Support for a diagnosis of FTD may similarly be obtained from functional imaging using PET or SPECT to demonstrate selective frontal and/or anterior temporal reduction in blood flow or metabolism. This approach may be particularly helpful in differentiating FTD from Alzheimer disease, as in the latter condition reductions in blood flow and metabolism are most prominent in posterior temporal and parietal regions.

Pathological examination has been essential in revealing the underlying processes resulting in FTD. Macroscopically, FTD is characterized by focal atrophy of the frontal cortex, the temporal cortex, or both. Atrophy is usually symmetric, though asymmetry is sometimes present. The Consortium for Frontotemporal Lobar Degeneration recently elaborated guidelines for the neuropathological assessment of cases with frontotemporal lobar degeneration. Systematic survey of frontal, temporal, and parietal cortices and the hippocampus with histological stains and immunocytochemistry using antibodies directed against τ, ubiquitin, or P62 (sequestosome-1) and α-synuclein is recommended to identify lesions associated with the frontotemporal lobar degenerations and to exclude other neurodegenerative dementias. Further evaluation of additional brain regions including the basal ganglia, midbrain, brain stem, and spinal cord, with the above staining approaches, immunocytochemistry for TDP-43 and for τ isoforms, and molecular genetic identification of mutations may also be necessary to achieve specific diagnoses.

Using these approaches, the frontotemporal lobar degenerations can be divided into two major groups, one with τ-positive inclusions, another with τ-negative and ubiquitin/P62-positive inclusions. These two presentations do not account for the behavioral- versus language-predominant presentations of FTD, which are presumed to relate more to the location of pathology. The associated motor syndromes, however, do demonstrate selective associations with these pathological groupings, as do genetic etiologies.

The frontotemporal lobar degenerations with τ-positive inclusions include classically described Pick disease, CBD, PSP, and dementia due to MAPT mutations. These subtypes can to some extent be distinguished by the pattern of lesion location and morphology, as revealed through the use of antibodies that nonselectively recognize phosphorylated τ. However, distinguishing among subtypes may be more reliably achieved through the use of antibodies selective for τ isoforms.

The second major group is the frontotemporal lobar degenerations with τ-negative ubiquitin/P62-positive inclusions. The most frequent pathological presentation in this group and the most frequent pathological presentation among all cases of frontotemporal lobar degenerations is the presence of intracellular inclusions positive for TDP-43. It is now recognized that many cases formerly classified as dementia lacking distinctive histological features are TDP-43 positive and that if dementia lacking distinctive histological features exists, it is quite rare. FTDs with TDP-43 lesions can occur with or without associated motor neuron disease. FTDs due to *GRN* and *VCP* mutations are TDP-43 positive, while those due to *CHMP2B* mutations are ubiquitin/P62 positive and TDP-43 negative. FTD due to *C9ORF72* presents with both TDP-43 positive and ubiquitin/P62 positive/TDP-43 negative inclusions, the latter most often found in the cerebellum.

Differential Diagnosis. Not surprisingly, given the frequent initial presentation of progressive alterations in behavior, psychiatric syndromes are often initially suspected in individuals with FTDs. Depending on whether the initial behavioral symptoms are disinhibited, withdrawn and apathetic, or inattentive and disorganized, bipolar illness, major depression, and/or late-onset schizophrenia may be suspected. Distinguishing FTD from these possibilities may be aided by structural and functional brain imaging, and by use of neuropsychological tests. Ultimately, differentiation from these psychiatric disorders may rely on longitudinal observation of progressive cognitive decline or the identification of associated motor or extrapyramidal symptoms.

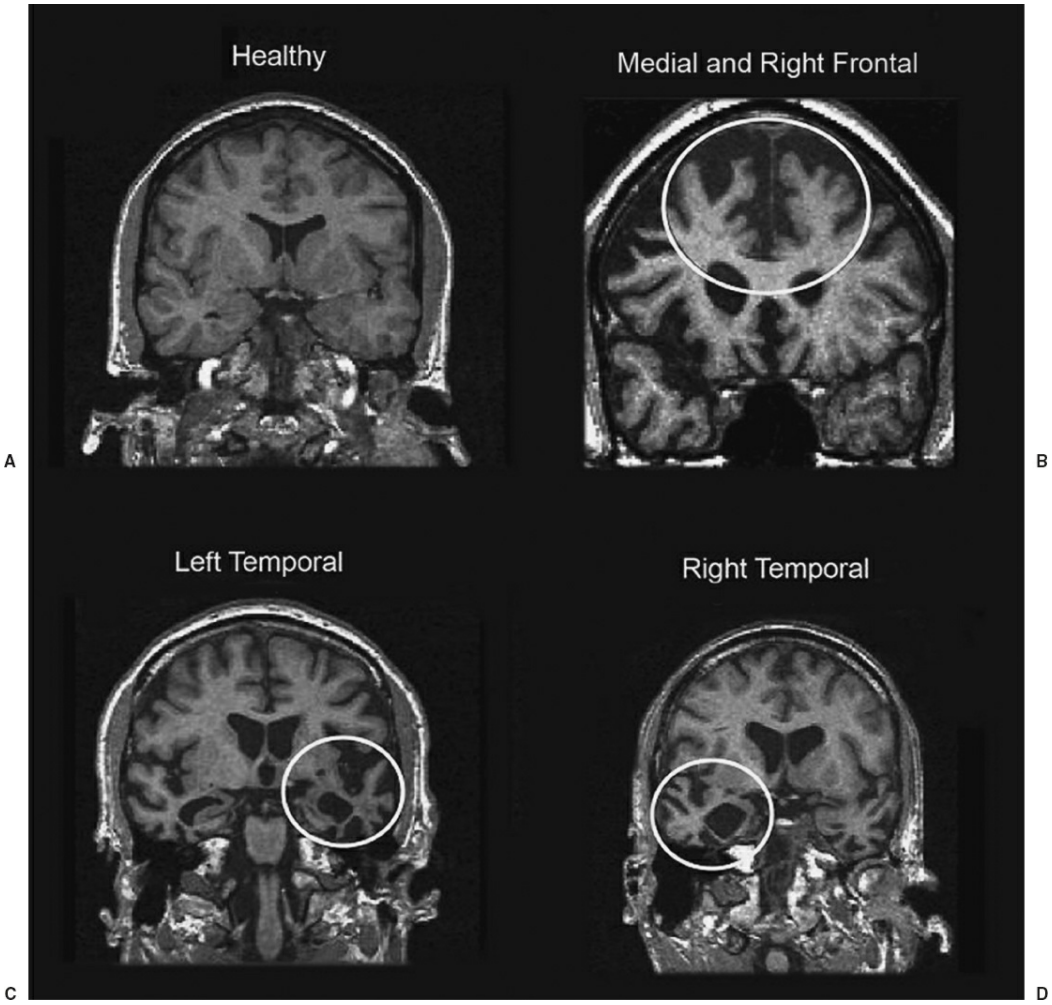

FIGURE 10.3–5. Patterns of atrophy (frontotemporal lobar degeneration [FTLD]) in patients with frontotemporal dementia. T_1-weighted coronal magnetic resonance images of **(A)** a healthy age-matched control participant, **(B)** a patient with frontal-variant FTLD, **(C)** a patient with temporal-variant FTLD showing predominantly left-sided atrophy, and **(D)** a patient with temporal-variant FTLD showing predominantly right-sided atrophy. (Images provided by Indre Viskontas, Ph.D., and Bruce Miller, M.D., the Memory and Aging Center, University of California San Francisco.)

Differential diagnosis may also include the other causes of dementia, commonly frontal predominant presentations of Alzheimer disease, particularly if the logopenic variant of primary progressive aphasia is diagnosed, or the combination of cognitive impairment and parkinsonism found in DLB. The presence of parkinsonism requires that one also consider the diagnoses of corticobasal syndrome and PSP. For example, the additional symptoms of progressive parkinsonism, axial rigidity, and supranuclear gaze palsy would indicate an FTD with PSP. Motor symptoms of asymmetric rigidity and apraxia may signal an associated CBD. Both PSP and CBD will be discussed in further detail below.

FTD should be suspected in younger patients and when there is a relative preservation of memory and visuospatial functions. A diagnosis of FTD is also supported by evidence of circumscribed frontotemporal

atrophy and selective reductions in frontotemporal blood flow or metabolism.

Course and Prognosis. The course of FTD is one of progressive deterioration. With progression, the differing initial presentations described above frequently merge, and ultimately loss of other cognitive functions may occur. Motor symptoms can precede, emerge concurrently, or follow the onset of the changes in behavior, personality, and language. On average, survival is 8 years from symptom onset, and 4 to 5 years from diagnosis. Those individuals with FTD and motor neuron disease have a significantly shorter survival time compared to those without motor neuron disease. Overall, individuals with FTD show more rapid cognitive and functional decline than those individuals with Alzheimer disease.

Treatment. Currently there are no treatments for the cognitive deficits associated with FTD, and no treatments to prevent progression of the underlying pathologies. Insufficient data exists to recommend cholinesterase inhibitors, and a recent randomized double-blind placebo controlled study showed no benefit of memantine in the treatment of frontotemporal lobar degeneration. Symptomatic use of trazodone, second-generation antipsychotics, SSRIs, and anticonvulsants may help agitation, disinhibited, and aggressive behavior. Typically, SSRIs are first-line pharmacotherapy treatments for FTD. Second-generation antipsychotic medications should be used with caution and close observation, as individuals with FTD may be more sensitive to extrapyramidal side effects. Nonpharmacological interventions for management of these behaviors should also be prioritized.

Corticobasal Degeneration

CBD can present as a syndrome with both clinical and pathological heterogeneity, with the clinical syndrome being more reflective of the neuroanatomical distribution of pathology than a particular histology.

Definition. The clinical syndrome of CBD is defined as a clinical entity with the core features of progressive asymmetric rigidity and apraxia, as well as cortical and basal ganglia dysfunction.

History. This clinical entity was first identified in 1967 in three individuals with a progressive asymmetrical akinetic-rigid syndrome and apraxia. It was later renamed CBD in the early 1990s.

Comparative Nosology. The term corticobasal syndrome is sometimes used to describe the clinical features seen, reserving the term CBD for the histopathological diagnosis. Other names for this syndrome that have been

used in the past include corticonigral degeneration and corticobasal ganglionic degeneration. In 2013, an international consortium published diagnostic criteria.

Epidemiology. The age of onset has been reported to range from 45 to 77.

Etiology. CBD is for the most part a sporadic disorder. The h1 haplotype of the MAPT gene has been associated with the development of this disease.

Diagnosis and Clinical Features. Progressive asymmetric rigidity and apraxia are considered to be the core clinical features of this syndrome. Additional symptoms include the alien limb phenomenon, cortical sensory loss, myoclonus, mirror movements, and dystonic posturing. Tremor is a common presenting symptom, but in contrast to tremor in Parkinson disease, tremor in CBD is minimal at rest and amplified with action. Furthermore, symptoms are not responsive to L-dopa therapy. Neuropsychological testing may reveal impairments in attention and concentration, executive function, language, and visuospatial function; learning and memory are often spared. Aphasia and apraxia of speech may also be evident, as can be depression and apathy. Features of an FTD with behavioral disturbance can also be seen. These frontal behaviors and speech impairments are common manifestations early in the course. Ocular motor apraxia is also seen, with individuals having difficulty initiating saccades and voluntary gaze. Dementia is common in pathologically confirmed cases of CBD.

Pathology and Laboratory Examination. On neuropathological examination, there is focal cortical neuronal loss, typically in the frontal, parietal or temporal lobes, as well as neuronal loss in the substantia nigra. The pathological features of CBD include Gallyas/tau-positive astrocytic plaques and threads, particularly in the superior frontal gyrus, the superior parietal gyrus, pre- and postcentral gyri, and the striatum. Tau-positive oligodendroglial coiled bodies may be evident. Ballooned neurons are considered to be a hallmark feature, but are not required for a diagnosis.

On structural imaging, asymmetrical cortical atrophy can be seen, particularly in the frontoparietal regions; atrophy is more prominent in the hemisphere contralateral to the side of the symptoms. There may also be asymmetrical atrophy in the basal ganglia, and cerebral peduncles. SPECT and PET can have asymmetrical hypoperfusion in the parietal and frontal cortex and basal ganglia. In dopamine transporter imaging, there is an asymmetric reduction in tracer uptake in the nigrostriatum, with greater reductions seen contralateral to the more clinically affected side.

Differential Diagnosis. The presence of parkinsonism on examination causes one to include the other Parkinson plus syndromes in the differential. There is a significant overlap between CBD and PSP. Individuals with CBD may have underlying PSP pathology, and those with the classical symptoms of PSP may have CBD pathology. Differences in the cognitive impairments may be helpful in distinguishing PSP and CBD. In PSP there is predominant axial rigidity, as opposed to the progressive asymmetric rigidity seen in CBD. In contrast to Parkinson disease, the tremor typically seen in CBD is minimal at rest and amplified with action, and the parkinsonian symptoms are typically not responsive to dopaminergic medications. Visual hallucinations should cause one to consider DLB. FTD should also be considered. Finally, the focal nature of deficits in CBD should also lead to the consideration of a vascular dementia, but one would expect to see infarcts on imaging in this case.

Course and Prognosis. Symptoms are usually asymmetric at first, but most often the rigidity and apraxia begin in an upper limb and then spread about 2 years into the disease course to either the ipsilateral lower limb or the contralateral upper limb. Median survival time has been reported to be about 6 to 7 years.

Treatment. Symptomatic treatments of the motor symptoms include botulinum toxin for focal dystonia, propranolol or primidone for tremor, and clonazepam for myoclonus, as the parkinsonism present is not typically responsive to L-dopa therapy. The importance of physical, occupational, and speech therapies has been emphasized.

Progressive Supranuclear Palsy

Definition. PSP is a gradually progressive disorder with progressive supranuclear gaze palsy, postural instability and falls, and parkinsonism.

History. In 1963, a case series on PSP was published; initially described as "heterogeneous system degeneration," this report detailed eight patients who had impairments in ocular gaze, spasticity of facial musculature with dysarthria and at times dysphagia, nuchal rigidity, and dementia. This was officially termed PSP in 1964.

Comparative Nosology. Formal diagnostic criteria were initially proposed in 1987. In 1996, the National Institute for Neurological Disorders and Stroke and the Society for PSP (NINDS-SPSP) refined the diagnostic criteria. Core features of this disease include an age of onset over the age of 40, with vertical supranuclear palsy or both vertical supranuclear palsy and prominent postural instability with falls in the first year of symptoms.

Epidemiology. The average annual incidence rate of PSP for ages 50 to 99 has been estimated to be 5.3 per 100,000 person-years. The incidence rate rises with increasing age, with estimated incidence ranging from 1.7 for ages 50 to 59 to 14.7 person-years for ages 80 to 99. Incidence has been noted to be slightly higher in men.

Etiology. PSP is considered to mostly be a sporadic disease. The H1 haplotype of the MAPT gene has been associated with PSP; inheritance of two copies of the H1 haplotype is considered to be a genetic risk factor. Recently an area on chromosome 11 containing the DNA damage-binding protein 2 (DDB2) and the lysosomal acid phosphatase 2 (ACP2) genes has also been identified as a locus associated with PSP.

Diagnosis and Clinical Features. The median age of onset is 63. Common symptoms in the initial stages include impaired balance, slowing of movement, impaired eye motion, bulbar symptoms, and personality changes such as apathy and disinhibition. Prominent axial rigidity, imbalance, and gait impairment become more apparent in later stages. Oculomotor impairment is described as a vertical gaze impairment; oculocephalic reflexes are preserved. Downgaze limitations (in contrast to upgaze) are more specific for PSP. Dysphagia and high pitched dysarthria are present as well. Frequent falling occurs. In cognitive function, there is slowing of processing, as well as impairment in letter fluency and cognitive flexibility; and this is often described as a subcortical dementia. Language comprehension, memory, and visuospatial functioning are usually preserved. Apathy and executive dysfunction, disinhibition, dysphoria, and anxiety may also be present. Agitation is less frequent. Atypical presentations can include progressive aphasia, FTD, or obsessive compulsive features. Some cases may lack gaze palsy, gait impairment, or parkinsonism. Dopaminergic drugs provide little benefit for the motor symptoms.

Pathology and Laboratory Examination. On MRI, atrophy of the midbrain tegmentum, sometimes referred to as the morning glory sign because of its concave appearance, and the hummingbird sign, in which the atrophy of the rostral midbrain as seen on midsagittal images looks like a hummingbird, have been associated with PSP. On SPECT and PET imaging, there is frontosubcortical hypoperfusion and hypometabolism; these changes are usually symmetric.

On neuropathological examination, there are neurofibrillary tangles and neuropil threads in the basal ganglia and brain stem. The presence of tau-positive astrocytes in these areas is supportive of the diagnosis.

Differential Diagnosis. There are a number of different clinical

presentations of PSP. These include Richardson syndrome, which is described as the classic phenotype, PSP-parkinsonism, pure akinesia with gait freezing, and cerebellar ataxia. Differential diagnosis of PSP from FTD and CBD is discussed above (see sections on FTD and CBD). Given the progressive parkinsonism present in PSP, Parkinson disease should also be considered on the differential.

Course and Prognosis. Survival time after symptom onset is about 5 to 9 years.

Treatment. Similar to CBD, there is poor response to L-dopa. Physical, occupational, and speech therapy may be helpful for symptomatic management throughout the course of the disease (see Chapter 2.6, Neuropsychiatric Aspects of Movement Disorders for a more detailed discussion of Progressive Supranuclear Palsy).

Dementia with Lewy Bodies

Introduction. DLB is now considered to be the second most common cause of dementia after Alzheimer disease in the elderly. Operationalized diagnostic criteria for DLB have been defined and revised over the past decade. DLB shares some clinical and pathological features with both Alzheimer disease and Parkinson disease. As a result, complete discrimination among these syndromes, even at autopsy, may not be possible.

Definition. DLB is a neurodegenerative dementia characterized by progressive cognitive decline of sufficient magnitude to interfere with normal social or occupational function. Memory impairment needs not be prominent or persistent early in the course of the illness. Deficits of attention and of frontal executive skills and visuospatial ability may be more prominent early in the course. The core features of DLB include (1) fluctuating cognition with pronounced variations in attention and alertness, (2) recurrent visual hallucinations, which are typically well formed and detailed, and (3) spontaneous motor features of parkinsonism. The presence of features that are suggestive of the diagnosis of DLB are rapid eye movement (REM) sleep behavior disorder, severe neuroleptic sensitivity, and low dopamine transporter uptake in basal ganglia demonstrated by SPECT or PET imaging. Features supportive of the diagnosis include repeated falls and syncope, transient unexplained loss of consciousness, severe autonomic dysfunction, hallucinations in other modalities, systematized delusions, depression, relative preservation of medial temporal lobe structures on CT/MRI scan, generalized low uptake on SPECT/PET perfusion scan with reduced occipital activity, abnormal (low uptake) metaiodobenzylguanidine myocardial scintigraphy, and

prominent slow wave activity on EEG with temporal lobe transient sharp waves.

History. Lewy bodies, the intracytoplasmic, spherical, eosinophilic neuronal inclusion bodies, identified originally in subcortical nuclei as one of the hallmarks of idiopathic Parkinson disease, have also been identified in the brain stems and cortex of elderly demented patients. Lewy body disease was first described in 1984 by Kosaka, who distinguished "diffuse" Lewy body disease, where Lewy bodies are present throughout cortical regions, from "brainstem-predominant" Lewy body disease, where Lewy bodies are present primarily in brain stem nuclei, and a "transitional type," which is intermediate between the two. Clinical and neuropathological diagnostic criteria were published by the Consortium on Dementia with Lewy bodies in 1996 and revised in 2005.

Comparative Nosology. Several historical terms have been used to describe what is now referred to in general as DLB. These terms reflect the overlap clinically and pathologically with Alzheimer disease and Parkinson disease and include Lewy body variant of Alzheimer disease, Alzheimer disease with Parkinson disease changes, dementia associated with cortical Lewy bodies, diffuse Lewy body disease, and senile dementia of the Lewy body type.

The first consensus clinical criteria for DLB were published in 1996. These criteria defined the core clinical features and those clinical features that were supportive of the diagnosis. The presence of one core feature was required for a diagnosis of possible DLB and at least two core features for probable DLB. The distinction between DLB and Parkinson disease with dementia was operationalized by requiring evidence of dementia within 1 year of the motor symptoms for a diagnosis of DLB. The criteria did not exclude patients who simultaneously met clinical and/or pathological criteria for a diagnosis of Alzheimer disease. The criteria for DLB were revised in 2005, with efforts made to clarify the overlap of DLB with Parkinson disease and Alzheimer disease. Though the 1-year rule for differentiating DLB and Parkinson disease with dementia was recommended in research use, clinically the criteria were clarified to indicate that DLB should be diagnosed when dementia precedes or is concurrent with parkinsonian symptoms, and should not be diagnosed when dementia arises in the "context of well-established Parkinson disease." It was also noted that any such distinction, however, may be seen as arbitrarily dividing a continuum of presentations due to underlying Lewy body disease.

The distinction between neuropathologically defined DLB and Alzheimer disease was also addressed in the revised criteria. This issue is complicated by the awareness that most individuals with cortical Lewy

bodies have evidence of cortical amyloid deposition as well, and that Lewy bodies may be present even in individuals with familial early-onset Alzheimer disease due to mutations in APP, presenilin 1, and presenilin 2. The 2005 revisions to the criteria for DLB made a shift to conceptualize the neuropathological diagnosis of DLB and Alzheimer disease using probabilistic assessments as to whether the clinical syndrome had a low, intermediate, or high likelihood of being due to DLB.

Other revisions in 2005 included the addition of the presence of REM sleep behavior disorder and functional imaging evidence for reduced striatal dopamine transporter activity, and increasing the diagnostic weighting of severe neuroleptic sensitivity as suggestive features. In addition, new features considered supportive of the diagnosis of DLB were added.

The DSM-5 has incorporated the above described criteria in the diagnosis major neurocognitive disorder with Lewy bodies. For a diagnosis of major neurocognitive disorder with Lewy bodies, an individual must have two core features or one suggestive feature with one or more core features; if an individual only has one core feature or one or more suggestive features then a diagnosis of possible neurocognitive disorder with Lewy bodies is made.

Epidemiology. The prevalence of DLB is dependent on whether clinical or pathological criteria are used and what patient population is studied. As a consequence of these confounds and the unclear clinical and neuropathological boundaries between DLB, Alzheimer disease, and Parkinson disease dementia in early diagnostic criteria, there remains uncertainty regarding the incidence or prevalence of DLB. The prevalence of DLB in populations over 65 has been found to range from 0.1 to 5 percent. Clinical studies suggest that DLB likely accounts for 15 to 35 percent of all dementia patients. One population-based study calculated the incidence of DLB to be 3.5 per 100,000 person years, with peak incidence in the 70- to 79-year age group. Neuropathological studies conducted to date are best understood as describing the frequency of Lewy body pathology in community samples, where it is typically the third most common chronic neuropathological change after lesions of Alzheimer disease and vascular lesions, and often in combination with both of those.

Etiology. Lewy bodies are comprised of abnormal fibrillar deposits of α-synuclein. Normally, α-synuclein is a protein located in presynaptic axon terminals, and presumed to have a role in synaptic function. Mutations and copy number variations in the gene encoding α-synuclein, *SNCA*, result in disorders with Lewy body pathology, including DLB and Parkinson disease. Mutations in β-synuclein, *SNCB*, and in *GBA* resulting in DLB have also been described. Mutations in other genes (*LRRK2, EIF4G1, VPS35,*

PARK2, PINK1, and *PARK7*) are associated with autosomal dominant or recessive Parkinson disease. Their relationship to DLB, however, is not clear. The APOE ε4 allele has also been associated with increased risk of DLB. The causes of most cases of sporadic DLB remain unknown.

Diagnosis and Clinical Features. The central features of DLB include a progressive cognitive decline of sufficient magnitude to interfere with normal social or occupational function. Memory impairment may not necessarily be prominent or persistent in early stages but becomes more evident with progression of the illness. The cognitive profile includes both cortical and subcortical features, with deficits of attention, executive function, and visuospatial ability especially prominent.

The three core features of DLB are (1) fluctuating cognition with pronounced variation in attention and alertness, (2) recurrent visual hallucinations that are typically well formed and detailed, and (3) spontaneous motor features of parkinsonism.

Fluctuation in cognitive function and pronounced variation in attention and alertness is common in DLB. Fluctuation may occur rapidly in minutes to hours or slower with weekly or monthly variation in level of attention or alertness. Patients may exhibit daytime drowsiness and lethargy with lengthy naps, followed by periods of confusion upon awakening. They may have reduced awareness of the environment, manifest as staring into space. They may even appear to "switch off" or go blank at times. There may be episodes of disorganized speech. Then there may be periods of marked lucidity during which patients appear to spontaneously remit. Global performance may fluctuate with cognition, attention, and alertness with periods of pronounced impairment in functioning alternating with periods of normal or near-normal functioning.

Visual hallucinations occur early in the course of DLB, whereas in Alzheimer disease, hallucinations are generally a later complication. The visual hallucinations are usually recurrent, formed, and detailed, and tend to persist. A common theme is animals or people intruding into the home. Seeing inanimate objects or writing on the walls and ceiling is not infrequent. Patients may have delusional explanations of the hallucinations or they may have insight into their unreality. Emotional responses to the hallucinatory experiences can range from indifference to amusement to fear. Visual hallucinations may be more common during periods of reduced consciousness and exacerbated by visual impairment. Patients with visual hallucinations demonstrate more profound visuoperceptual dysfunction compared to those without hallucinations. Visual hallucinations correlate with a higher density of Lewy bodies in anterior and inferior temporal lobe and amygdala as well as reduced occipital perfusion in primary and secondary visual cortex.

Spontaneous motor features of parkinsonism are another core feature.

If dementia onset precedes or is concurrent with the onset of parkinsonian symptoms, a diagnosis of DLB is warranted. The most common extrapyramidal symptoms noted in DLB are rigidity and bradykinesia, followed by hypophonic speech, masked facies, stooped posture, and a slow and shuffling gait. Postural instability is also common. The parkinsonism seen in DLB tends to include greater postural instability, gait impairment, and facial immobility compared to those patients with Parkinson disease without dementia. Resting tremor may occur, but is less common in DLB than in Parkinson disease.

Other clinical features suggestive of a diagnosis of DLB include REM sleep behavior disorder, severe neuroleptic sensitivity, and low dopamine transporter uptake in basal ganglia demonstrated by SPECT or PET imaging.

REM sleep behavior disorder is characterized by vivid and often frightening dreams during REM sleep accompanied by lack of muscle atonia allowing for "acting out of dreams." Patients respond to vivid visual images with vocalizations, thrashing about, arising from bed, and often moving about violently. Bed partners are therefore at risk of injury. Patients usually have little or no recall of their dreams or these episodes. The diagnosis can be confirmed by polysomnography. REM sleep behavior disorder may predate the emergence of parkinsonism or dementia by many years; that is, REM sleep behavior disorder may be a manifestation of an evolving synucleinopathy.

Severe neuroleptic sensitivity is present in about half of patients with DLB. They will demonstrate acute onset or exacerbation of parkinsonism and impaired consciousness with even a low dose of a first- or second-generation antipsychotic agent. Even though some patients may tolerate antipsychotics, this does not preclude a diagnosis of DLB. Sensitivity to antipsychotics has profound treatment implications in DLB. Visual hallucinations, hallucinations in other modalities, and delusions are common, but the standard treatment for these psychotic symptoms, namely antipsychotic medications, can cause severe adverse effects. Therefore, in DLB, antipsychotics should generally be avoided in the treatment of psychotic symptoms.

Functional imaging of the dopamine transporter system with SPECT or PET may be used to assess the nigrostriatal dopaminergic system. Low dopamine transporter uptake in the basal ganglia supports the diagnosis of DLB.

Clinical features that are supportive of a diagnosis of DLB, but that do not have diagnostic specificity, are repeated falls and syncope, transient unexplained loss of consciousness, severe autonomic dysfunction, hallucinations in other modalities, systematized delusions, and depression. Imaging and other diagnostic findings that are supportive of a diagnosis of DLB are relative preservation of medial temporal lobe structures on CT or

MRI scan, generalized low uptake on SPECT/PET perfusion scans with reduced occipital activity, abnormal (low uptake) metaiodobenzylguanidine myocardial scintigraphy, and prominent slow wave activity on EEG with temporal lobe transient sharp waves.

Repeated falls are common in DLB, although they are not limited to this subtype of dementia. In DLB, falls are generally not associated with focal neurological signs and symptoms as would be present in balance disorders or transient ischemic attacks. Syncope may be associated with brain stem and autonomic nervous system Lewy body pathology. Transient loss of consciousness may represent an extreme form of fluctuation in attention and alertness. Systematized delusions may arise to provide an explanation for the hallucinations or other perceptual disturbances. They are often complex and bizarre as compared to the less well-developed delusions in Alzheimer disease. The most common delusions in DLB are persecution and theft, phantom boarders, television characters in the room, spousal infidelity, and Capgras syndrome (delusional misidentification syndrome). Patients may also have hallucinations in other modalities such as auditory, olfactory, or tactile.

Apathy and anxiety are also seen in DLB.

Pathology and Laboratory Examination. Lewy bodies are the only essential finding in the pathological diagnosis of DLB. Lewy bodies are spherical intracytoplasmic eosinophilic neuronal inclusion bodies (see Fig. 10.3–6). "Classic" Lewy bodies are inclusions with a hyaline core and pale halo and are typically seen in the brain stem nuclei, substantia nigra, and locus coeruleus. Cortical Lewy bodies are less-well-defined spherical inclusions. Lewy bodies are composed predominantly of fibrillar deposits of α-synuclein, and can also include neurofilament proteins and ubiquitin. Though Lewy bodies can be detected with hematoxylin and eosin stain and anti-ubiquitin antibodies, they are most sensitively and specifically labeled with antibody against α-synuclein. In addition to forming intracytoplasmic inclusions, aggregates of α-synuclein occur in neuronal processes, called Lewy neurites. The number of cortical Lewy bodies correlates significantly with cognitive impairment in DLB. There are also other pathological changes that may be seen in DLB but that are not required for diagnosis. Alzheimer disease pathology, particularly amyloid-containing plaques, is also a common feature in most cases of DLB. In contrast, the presence of cortical neurofibrillary tangles indicates that DLB is less likely. Microvacuolation or spongiform changes may also occur in some cases of DLB.

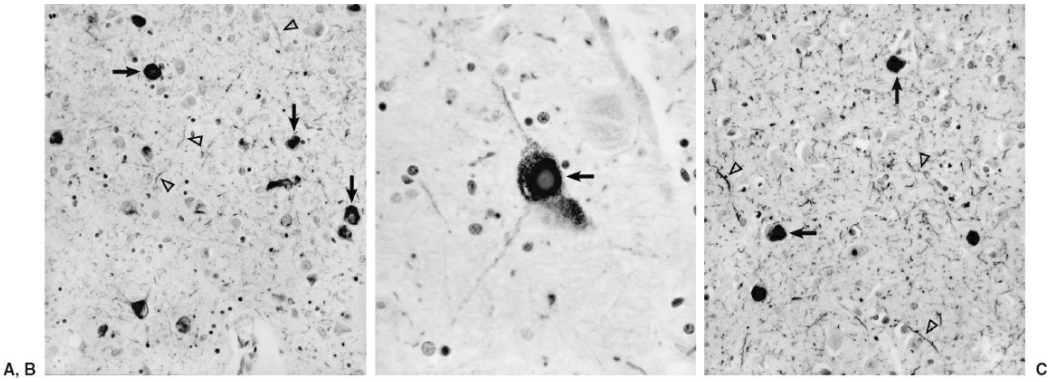

FIGURE 10.3–6. Photomicrographs of Lewy body pathology. **(A)** Abnormal accumulation of α-synuclein aggregates demonstrated by immunocytochemistry in the amygdala of a subject with dementia. Lewy bodies appear as dense intracellular inclusions (*arrows*), but staining of neuronal processes can be seen throughout the neuropil (*arrowheads*). In individuals in whom Lewy body pathology occurs concurrently with Alzheimer disease, the amygdala is often the only region affected. **(B)** Classic appearance of a Lewy body (*arrow*) in a large pigmented neuron of the substantia nigra. **(C)** Lewy body pathology in the neocortex. Both Lewy bodies (*arrows*) and substantial labeling of neuronal processes in the neuropil (*arrowheads*) are evident. (Magnification for [A] and [B] 200×, for [C] 400×. All images provided courtesy of Dr. Ronald L. Hamilton, Department of Pathology, Division of Neuropathology, University of Pittsburgh School of Medicine.)

Structural imaging studies, CT and MRI, are not able to effectively differentiate Alzheimer disease from DLB. Medial temporal lobe atrophy is seen in a slightly higher percentage of patients with Alzheimer disease than DLB, although this is not enough to differentiate the two conditions.

Functional imaging of cerebral perfusion and cerebral metabolism using SPECT and PET imaging, respectively, is a better tool for distinguishing between Alzheimer disease and DLB. In some studies, patients with DLB have greater hypoperfusion with FDG-PET in occipital regions than Alzheimer disease. Imaging the presynaptic dopaminergic terminals using ^{123}I-FP-CIT SPECT or ^{18}F-dopa PET reveals a marked loss of presynaptic dopaminergic terminals in the corpus striata in DLB, whereas there is no change in these terminals in normal individuals and in patients with Alzheimer disease (see Fig. 10.3–7).

EEG shows early generalized background slowing with abnormal transient sharp waves in the temporal lobes or frontally dominant burst patterns.

Differential Diagnosis. DLB overlaps with Alzheimer disease both in terms of clinical presentation and neuropathological changes. The distinction between DLB and Alzheimer disease is complex. The clinical symptoms are overlapping, and no specific cutoff based on number or pattern of core, suggestive, and supportive features present is definitive, though the presence of at least two of the core features in the early disease course (fluctuating attention, prominent visual hallucinations, and parkinsonism) indicates DLB with high specificity. However, even patients

who present with a typical insidious amnestic syndrome characteristic of Alzheimer disease may have Lewy body pathology at autopsy. There may be overlap with the two conditions not being mutually exclusive and patients meeting criteria for both.

There is also considerable clinical and pathological overlap and comorbidity between DLB and Parkinson disease. The primary diagnosis should be made based on the temporal relationship of the onset of cognitive and motor symptoms. DLB should be diagnosed when dementia occurs with or before parkinsonism. Parkinson disease with dementia should be diagnosed when dementia develops in a patient with well-established Parkinson disease. Dementia develops in at least 30 percent of patients with Parkinson disease, and perhaps even up to 78 percent when patients are followed over 8 years. In cases where there is not a clear temporal relationship between the motor and cognitive symptoms, the term *Lewy body disease* may be useful. For research purposes, the "1-year rule" should be used between the onset of parkinsonism and dementia to distinguish between these two entities. While the distinction is useful clinically, both diagnoses likely represent a common spectrum of Lewy body pathology with a common pathophysiology of abnormalities in α-synuclein metabolism.

Multiple system atrophy (MSA), which also has α-synuclein involvement, may also be considered in the differential. Clinical symptoms of MSA include parkinsonism, cerebellar ataxia, pyramidal symptoms, and autonomic dysfunction. Autonomic dysfunction is necessary for a diagnosis of MSA, and includes orthostatic hypotension, erectile dysfunction in males, constipation, and urinary symptoms (frequency, urgency, incontinence, and incomplete bladder emptying), and decreased sweating. Dementia is much less common in MSA compared to the other neurodegenerative illnesses with parkinsonism, and according to the second consensus statement is a nonsupporting feature of the diagnosis. However, there is recent evidence that certain cognitive domains are affected, with frontal executive dysfunction being the most common, and memory and visuospatial dysfunction also showing impairments.

Distinguishing between delirium and DLB can be challenging. Delirium and DLB share several common symptoms. The fluctuating level of alertness and consciousness in DLB is also a hallmark of delirium. Visual hallucinations are also common in both conditions, although they may be less well formed and less detailed in delirium. DLB, like other dementias, may also increase the risk for a superimposed delirium due to medications, medical illness, and after hospitalization for any reason. Delirium, however, should typically have an acute or subacute onset, and a duration limited to the period of exposure to the underlying causal agent. The contribution of medications or other comorbid medical conditions to worsening cognitive, hallucinatory, motor, or behavioral symptoms should

be investigated (see Chapter 10.2 for a further discussion of delirium).

Vascular dementia should also be considered in the differential diagnosis. A history of stroke as evident by focal neurological signs or brain imaging makes the diagnosis of DLB less likely. However, white matter lesions in the periventricular and deep white matter regions, microvascular changes, and lacunes may be present in up to 30 percent of autopsy-confirmed DLB cases, suggesting that a diagnosis of both DLB and cerebrovascular disease may be appropriate in many cases.

CJD may be considered in patients who develop myoclonus, which is also characteristic of a rapidly progressive form of DLB.

Course and Prognosis. DLB has an insidious and gradual onset, with the mean age at onset being 75 years old with a range from 50 to 80 years. There is a slight male predominance. DLB usually presents with one of the core features. It may have a more rapidly progressing course than Alzheimer disease or vascular dementia. Survival time is usually about 10 years, although it may be as little as 1 to 2 years. The course is often marked by psychosis and behavioral disturbances, which makes management more complicated.

Treatment

NONPHARMACOLOGICAL TREATMENT. Nonpharmacological interventions have not been systematically studied in DLB. Strategies to increase levels of arousal and attention by social interaction and environmental novelty may reduce the visual hallucinations and fluctuations in cognition and function.

PHARMACOLOGICAL TREATMENT

Cholinesterase Inhibitors. The cholinesterase inhibitors are the mainstay of treatment for the cognitive impairment of DLB. A randomized placebo-controlled trial showed that donepezil was associated with improvements in cognition, behavior, and global functioning, as well as reductions in caregiver burden. A randomized placebo-controlled trial of rivastigmine demonstrated significant improvement in apathy, anxiety, delusions, hallucinations, and cognitive functioning.

Memantine. Data in regard to the usage of memantine is mixed. While there are some studies that have shown that memantine is associated with improved quality of life and improvements in global clinical status, a recent systematic review found that there was no benefit on motor function, cognitive function, or activities of daily living. There have also been case reports of worsening hallucinations and delusions with memantine as well.

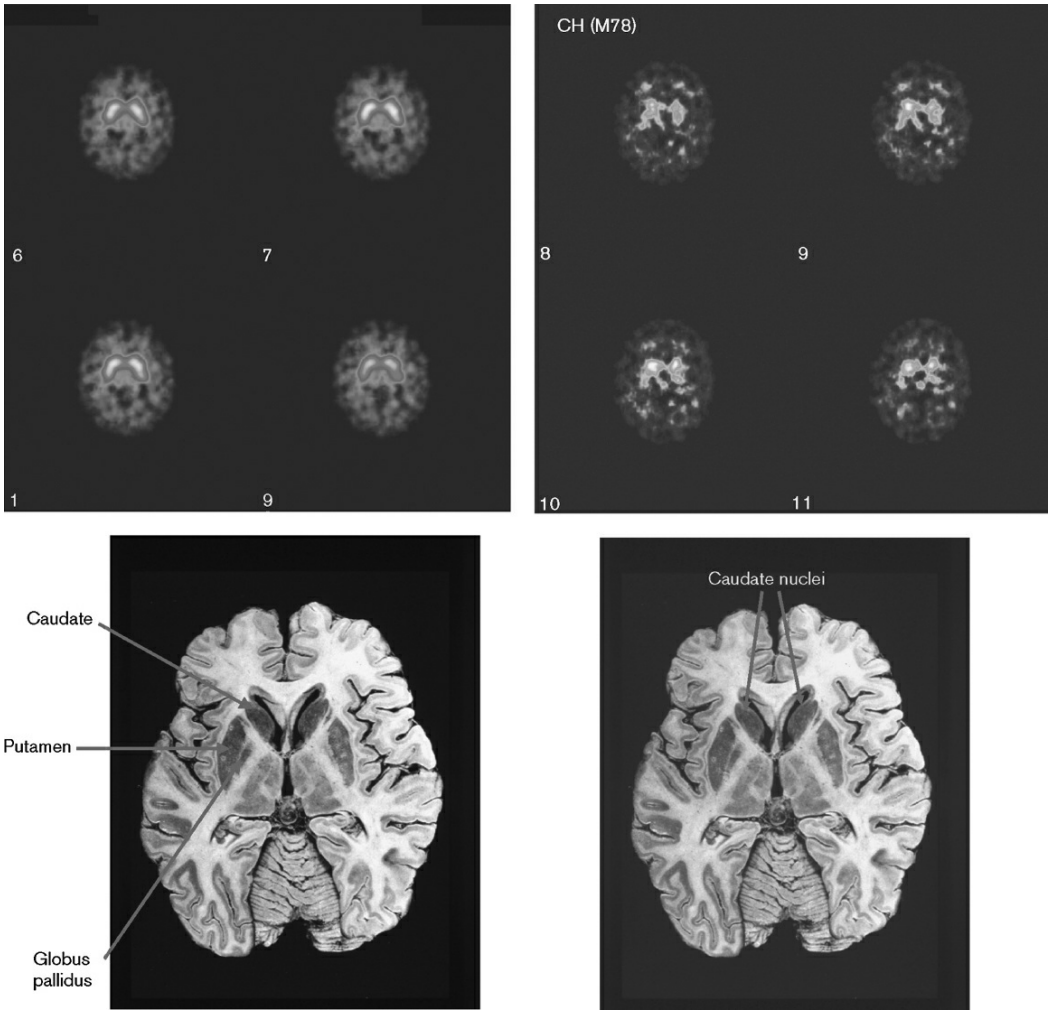

FIGURE 10.3–7. ^{123}I-FP-CIT single-photon emission computed tomography (SPECT) images of normal brains and brains affected by dementia with Lewy Bodies. *Left images:* The normal appearance of a ^{123}I-FP-CIT image. There is a symmetrical appearance of the presynaptic dopaminergic terminals in the corpus striata (caudate, globus pallidus, and putamen). This normal appearance is also seen in patients with Alzheimer disease as there is no significant loss of these dopaminergic terminals. *Right image:* The abnormal appearance of a ^{123}I-FP-CIT image in a 78-year-old patient with fluctuating cognitive impairment and recurrent, vivid, and detailed visual hallucinations. There is a marked loss of the presynaptic dopaminergic terminals in the corpus striata as seen in patients with dementia with Lewy bodies. Only a few terminals remain in the caudate nuclei. This appearance is also seen in Parkinson disease. Note that this patient with dementia with Lewy bodies had a highly abnormal scan even though there were no clinical features of parkinsonism.

Antipsychotics. First-generation antipsychotics are generally not well tolerated in DLB and should be avoided. There is little data to guide the use of second-generation antipsychotics, for which DLB patients may also have low tolerance. A placebo-controlled study of olanzapine in psychosis in Alzheimer disease looked at a subgroup of patients retrospectively diagnosed with DLB. These patients demonstrated a decrease in positive symptoms of psychosis without worsening of parkinsonism or cognition with olanzapine. A small open-label study of quetiapine for psychosis and

agitation in DLB found a clinically significant reduction in symptoms in about half the patients, but several had to discontinue treatment due to somnolence and orthostatic hypotension. There is a lack of data for the other second-generation antipsychotics. Although there are no controlled trials involving the use of clozapine in the treatment of neuropsychiatric symptoms of DLB, given this medication's utility in treating psychosis in Parkinson disease, there may be some clinical value in its treatment of psychotic symptoms in DLB as well.

Antidepressants. Antidepressants have not been systematically studied in DLB. SSRIs and selective and norepinephrine reuptake inhibitors (SNRIs) are probably preferred as antidepressants with anticholinergic properties should be avoided.

Benzodiazepines. Clonazepam may have benefit for the treatment of REM sleep behavior disorder, but like all long half-life benzodiazepines must be used cautiously in the elderly.

Antiparkinsonian Medications. Antiparkinsonian medications including L-dopa and direct dopamine agonists may be tried in DLB. Responses are variable, but generally less robust than in Parkinson disease. Given their increased potential for neuropsychiatric complications, initial doses should be small, with gradual increases while monitoring behavioral symptoms. Anticholinergic medications are generally not recommended, because of the associated cognitive side effects.

Dementia in Parkinson Disease

Cognitive impairment in Parkinson disease is common, with almost a quarter of newly diagnosed individuals displaying some impaired performance.

Definition. In 2007, the Movement Disorder Society task force published criteria to guide the diagnosis of dementia associated with Parkinson disease. For an individual to be diagnosed with probable Parkinson disease dementia, there must be a dementia syndrome that developed within the context of Parkinson disease. Similar to other dementia syndromes described in this chapter, there must be impairment in more than one cognitive domain, this must be a decline from a previous level of functioning, and there are impairments in daily life independent of those impairments due to the motor or autonomic symptoms. Cognition is impaired in two out of the following four domains (attention, executive function, visuospatial function, and free memory recall) and there must also be the presence of at least one behavioral symptom.

History. Parkinson disease was first written about in 1817, by James

Parkinson. In a treatise on the "shaking palsy" he described six cases he had observed. In 1888, the French neurologist, Jean Martin Charcot, named this clinical entity after Parkinson.

Comparative Nosology. In the DSM-5, for a diagnosis of major neurocognitive disorder due to Parkinson disease, the criteria for major neurocognitive disorder must be met, the impairment occurs in the setting of well-established Parkinson disease, there is insidious onset and gradual progression, and the impairment cannot be attributable to another medical condition or mental disorder.

Epidemiology. The prevalence of Parkinson disease increases with age. Up to 80 percent of individuals with Parkinson disease will develop dementia, with roughly 10 percent of individuals with Parkinson disease developing dementia each year. On average, Parkinson disease dementia is diagnosed 8 to 10 years from the onset of Parkinson disease, but there is a wide range. The age of onset of Parkinson disease is between the sixth and ninth decades; increased duration of illness is a risk factor for the development of dementia.

Etiology. Risk factors for the development of dementia in Parkinson disease include age, more severe parkinsonism, and MCI at initial diagnosis. While APOE ε4 is a known risk factor of Alzheimer disease and has been associated with DLB, evidence is mixed as to whether this plays a role in the development of Parkinson disease dementia. The hereditary *SCNA* and *GBA* forms of Parkinson disease have been associated with dementia.

Diagnosis and Clinical Features. As this clinical entity occurs within Parkinson disease, the cognitive impairment should occur at least 1 year after symptoms of Parkinson disease were first noted. On neuropsychiatric examination, impairments in attention, executive function, visuospatial function, language, and memory may be seen. Impaired mental speed may be evident. On memory testing, recognition is better than free recall. Regarding language impairment, word finding difficulties, and impaired comprehension of complex sentences may be seen. Associated neuropsychiatric symptoms that may be present include apathy, depression, anxiety, hallucinations, delusions, and personality changes. Hallucinations have been reported in 45 to 65 percent of individuals with PDD, and delusions in 25 to 30 percent. Similar to DLB, individuals may have REM sleep behavior disorder, and excessive daytime sleepiness may also be seen.

Pathology and Laboratory Examination. Parkinson disease is associated

with neuronal loss in the substantia nigra. Lewy bodies, made up of α-synuclein, are present in the cell bodies of the surviving neurons. α-Synuclein pathology is more extensive in Parkinson disease dementia compared to Parkinson disease without dementia, with more global cortical and limbic involvement. Individuals may also have β-amyloid plaque burden as well as the presence of tau neurofibrillary tangles.

MRI imaging may show thinning in the frontal and anterior cingulate regions. Dopamine transporter imaging is abnormal in Parkinson disease; however, this abnormality does not differentiate Parkinson disease from Parkinson disease with dementia. Cardiac metaiodobenzylguanidine scintigraphy is reduced in Parkinson disease; of note, this test is normal in MSA and PSP. On EEG, those with dementia in Parkinson disease have slower baseline EEG activity than those without cognitive impairment.

Differential Diagnosis. The one year rule can be useful in distinguishing Parkinson disease dementia from DLB. The presence of parkinsonian symptoms and cognitive impairment should also cause one to think about other Parkinson plus syndromes, including CBD, PSP, and MSA. The tremor in CBD differs from that typically seen in Parkinson disease, in that it is typically minimal at rest and amplified with action, and CBD is often associated with an asymmetric rigidity. The vertical gaze impairment as well the prominent axial rigidity in PSP can aid in distinguishing this from Parkinson disease; however, the gaze impairment can be absent in up to 50 percent of individuals with PSP. The parkinsonism in both CBD and PSP is typically not responsive to dopaminergic medications. In MSA, individuals may display bradykinesia, rigidity, postural instability, and postural tremor; the cerebellar and autonomic dysfunction may be helpful in distinguishing this from Parkinson disease. Of note, in multisystem atrophy, up to 30 percent of individuals may show some response to dopaminergic medications at some point during the disease process.

Course and Prognosis. The course of dementia in Parkinson disease is progressive over time, although individuals may have some plateaus, as well as fluctuations from day to day. Hallucinations may occur, as can other psychiatric symptoms, particularly apathy. Dementia has been associated with increasing functional impairment, nursing home placement, and decreased survival in Parkinson disease.

Treatment. Evidence suggests that acetylcholinesterase inhibitors can improve cognition, attention, executive function, and global status, and may have beneficial effects on overall activities of daily living. Memantine has not been shown to be beneficial. Should psychotic symptoms be problematic, attempts to reduce or discontinue antiparkinsonian medications, balancing psychotic and motor symptoms are warranted, with

consideration of reduction in L-dopa as a last resort. Antipsychotic medication should be used very cautiously. Clozapine is the only antipsychotic medication that has been deemed clinically useful for the treatment of psychosis in Parkinson disease by the Movement Disorder Society Task Force on Evidence-Based Medicine Review of Treatments for Parkinson disease. There is insufficient evidence for the use of quetiapine, and olanzapine is not felt to be useful given its significant worsening of motor symptoms (see Chapter 2.6, Neuropsychiatric Aspects of Movement Disorders for more detailed descriptions of the behavioral complications of Parkinson disease and their treatment).

Dementia in Huntington Disease

Definition. Huntington disease is a rare disorder that is characterized by motor, cognitive, and psychiatric symptoms. The motor symptoms include chorea, or involuntary dance-like movements, and ataxia; the disease inevitably leads to dementia.

History. The clinical entity of Huntington disease was first described by George Huntington in the late 1872. In 1993, the mutant gene was identified.

Comparative Nosology. For an individual to be diagnosed with the DSM-5 diagnosis of major neurocognitive disorder due to Huntington disease, there must be clinically established Huntington disease, or risk of this based on family history or genetic testing, in addition to meeting the criteria for major neurocognitive disorder, and there being a gradual and insidious onset.

Epidemiology. Worldwide prevalence of Huntington disease has been reported to be 2.71 per 100,000, with higher prevalence seen in Europe, North America, and Australia.

Etiology. Huntington disease is an autosomal-dominant inherited disease transmitted through a triplet CAG repeat mutation located on chromosome 4 in the *HTT* gene. This mutation results in a polyglutamate expansion in the Huntington protein. Expression of the abnormal protein ultimately results in neuronal cell death, primarily in the caudate nucleus and frontal lobes.

Diagnosis and Clinical Features. In addition to the chorea and ataxia, other motor symptoms include bradykinesia, dystonia, and ideomotor apraxia. Individuals may also develop personality changes, depression, blunting of affect, irritability, psychosis, and aggressive–compulsive behavior. Cognitive changes include visuospatial deficits, psychomotor

slowing, impairments in executive functioning and attention, and memory deficits. Dysarthria and dysphagia are present in more severe stages.

Pathology and Laboratory Examination. Atrophy of the caudate is evident on structural imaging, and striatal atrophy may be seen on imaging before a clinical diagnosis of Huntington disease is made.

A family history is important in the diagnosis of this disease, and it can be confirmed by genetic testing.

Differential Diagnosis. Given the autosomal dominant pattern, obtaining a family history is important. If family history is unknown or if genetic testing is negative, CJD (discussed below) should be considered in the differential, given the ataxia and myoclonus that may be present. The differential for chorea includes both other hereditary and acquired causes.

Course and Prognosis. Huntington disease is inevitably fatal.

Treatment. Antidepressants and sedatives can be used for associated mood disorders. Dopamine depleting agents may be used for the treatment of chorea. Antipsychotic medications may be used for the treatment of agitation and psychosis. The more potent antipsychotic medications may also be beneficial for the chorea; however, these may worsen any bradykinesia that is present. Cholinesterase inhibitors have not been shown to have benefit in the treatment of the cognitive impairment (see Chapter 2.6, Neuropsychiatric Aspects of Movement Disorders for more detailed descriptions of the behavioral complications of Huntington disease and their treatment).

Prion Diseases

Definition. The prion disorders encompass both sporadically occurring and genetically induced neurodegenerative diseases. They include Kuru, sporadic CJD, familial CJD, iatrogenic CJD, variant CJD, GSS disease, and fatal insomnia.

History. CJD was first termed in 1922, and was used to describe a clinical entity of rapidly progressive myoclonus, ataxia, and dementia, described in earlier case reports. In the 1950s, kuru was discovered in Papau New Guinea, and was characterized by truncal ataxia, tremor, and quick progression to dementia and death. Kuru was eventually found to share similar histopathology and transmission properties with CJD. Later it was determined that these infectious agents were proteins, and the term prion was created in 1982 to characterize these proteinaceous infectious particles. In the late 1990s, there was a series of variant CJD cases in the United Kingdom and France, which resulted from the transmission to

humans of bovine spongiform encephalopathy, commonly known as mad cow disease, through consumption of contaminated meat products.

Comparative Nosology. The DSM criteria for major neurocognitive disorder due to prion disease include an insidious onset with rapid progression of impairment, and motor features of prion disease or biomarker evidence must be present, and the presentation cannot be attributable to another medical or mental disorder.

Epidemiology. Incidence of sporadic CJD worldwide is 1 per 10^6 and of familial CJD is 1 per 10^7 per year. In sporadic CJD, the peak age of onset is 67 years, while the familial and variant form typically occurs in younger individuals. Fatal insomnia usually presents between the ages of 35 and 55.

Etiology. The prion diseases can be sporadic, inherited, and also horizontally transmissible. The gene for the prion protein is found on chromosome 20. Underlying these diseases is a misfolding of the prion protein, which occurs via a spontaneous or genetically induced mutation. Once this mutation occurs, the pathological form of the prion propagates. CJD and fatal insomnia are both sporadic and familial. GSS is a familial disease. Iatrogenic CJD has been associated with medical or surgical equipment contaminated with the pathological prion protein. Transmission through ingestion has also been demonstrated, as seen with Kuru and the practice of ritual endocannibalism in the Foré linguistic group in Papa New Guinea, and also variant CJD through the ingestion of contaminated meat products.

Diagnosis and Clinical Features. Sporadic and familial CJD initially present with memory impairment and confusion, which quickly progress to a cortical dementia like picture, as well as ataxia and myoclonus. Iatrogenic forms of CJD typically show more cerebellar ataxia at symptom onset. GSS first presents with ataxia, and then dementia develops. Variant CJD often presents with psychiatric symptoms, usually apathy or depression, and painful sensory symptoms in distal extremities. Often GSS presents initially with ataxia or dysarthria, and then progresses to dementia; other symptoms that have been reported include memory impairment, spastic paresis, motor impairment, and behavioral disturbances. In fatal insomnia, there is insomnia, which may be present for months until other symptoms present themselves, including dysautonomia and ataxia; cognitive impairment and dementia does not appear until much later in the course.

Pathology and Laboratory Examination. The hallmark clinical feature of sporadic and familial CJD is spongiform degeneration of the cortical gray matter; reactive gliosis is often evident as well. Pathology is usually focused

in the neocortex, subiculum of the hippocampus, caudate, putamen, and the thalamus. In GSS, there are extensive prion protein plaques, with less spongiform degeneration. There is focal thalamic gliosis and minimal spongiform degeneration seen in fatal insomnia. The neuropathology seen in variant CJD includes prion protein plaques surrounded by a halo of vacuolar change with diffuse spongiosis. Neuronal loss and gliosis is seen focally in the thalamus and the inferior olives in fatal insomnia.

On MRI, there are high signal abnormalities in the caudate nucleus and/or putamen on diffusion-weighted imaging (DWI) and fluid-attenuated inversion recovery (FLAIR) in CJD. Hypometabolism in the thalamus on PET scan points toward fatal insomnia. Elevations in 14-3-3, total tau, or neuron-specific enolase in CSF studies have been associated with sporadic forms of CJD. EEG may be useful in the diagnosis of prion diseases. In sporadic and familial CJD, EEG shows bilateral periodic discharges; iatrogenic CJD has a diffuse slow wave pattern. Slow waves can also be evident in GSS.

Differential Diagnosis. Given the progressive cognitive impairment and motor features of the prion diseases, the differential is quite broad and includes many of the neurodegenerative conditions that have been discussed in the chapter, including Alzheimer disease, the varying frontotemporal lobar dementias, DLB, CBD, and Huntington disease. The spinocerebellar ataxias as well as Hashimoto encephalopathy can also have similar symptom presentations.

Course and Prognosis. Sporadic CJD is most commonly diagnosed in the seventh or eighth decade. Course is rapidly progressive, with the prognosis somewhere between 4 and 6 months. The familial form of the disease has a more extended course, typically of 1 to 5 years. The duration of GSS is typically 2 to 6 years, and of fatal insomnia is 1 to 2 years.

Treatment. At this time, treatment is limited to symptomatic therapies, including antiepileptic medications for seizure control, clonazepam for myoclonus, and small doses of antipsychotic medications for distressing psychiatric symptoms including hallucinations and delusions.

References

Armstrong MS, Litvan I, Lang AE, et al. Criteria for the diagnosis of corticobasal degeneration. *Neurology*. 2013;80:496–503.

Bains J, Birks J, Dening T. Antidepressants for treating depression in dementia. *Cochrane Database Syst Rev*. 2002;4:CD003944.

Ballard CG, Waite J, Birks J. Atypical antipsychotics for aggression and psychosis in Alzheimer's disease. *Cochrane Database Syst Rev*. 2006;1:CD003476.

Bloom GS. Amyoid-β and tau: the trigger and bullet in Alzheimer disease pathogenesis. *JAMA Neurol*. 2014;71(4):505–508.

Brown K, Mastrianni JA. The prion diseases. *J Geriatr Psychiatry Neurol*. 2010;23:277–298.

Cairns NJ, Bigio EH, Mackenzie IR, Neumann M, Lee VM. Neuropathologic diagnostic and nosologic criteria for frontotemporal lobar degeneration: Consensus of the Consortium for Frontotemporal Lobar Degeneration. *Acta Neuropathol*. 2007;114:5–22.

Cohen AD, Klunk WE. Early detection of Alzheimer's disease using PiB and FDG PET. *Neurobiol Dis*. 2014;72:117–122.

Coyle-Gilchrist ITS, Dick KM, Patterson K, et al. Prevalence, characteristics, and survival of frontotemporal lobar degeneration syndromes. *Neurology*. 2016;86(18):1736–1743.

Dubois B, Hampel H, Feldman HH, et al. Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria. *Alzheimers Dement*. 2016;12(3):292–323.

Emre M, Aarsland D, Brown R, et al. Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov Disord*. 2007;22(12):1689–1707.

*Geda YE, Schneider LS, Gitlin LN, et al. Neuropsychiatric symptoms in Alzheimer's disease: past progress and anticipation of the future. *Alzheimers Dement*. 2013;9(5):602–608.

Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, et al. Classification of primary progressive aphasia and its variants. *Neurology*. 2011;76(11):1006–1014.

Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2012-2050) estimated using the 2010 census. *Neurology*. 2013;80(19):1778–1783.

Ittner LM, Götz J. Amyloid-β and tau—a toxic pas de deux in Alzheiemer's disease. *Nat Rev Neurosci*. 2011;12(2):65–72.

*Karageorgiou E, Miller BL. Frontotemporal lobar degeneration: a clinical approach. *Semin Neurol*. 2014;34(02):189–201.

Kemp PM, Holmes C. Imaging in dementia with Lewy bodies: a review. *Nucl Med Commun*. 2007;28:511–519.

Ling H. Clinical approach to Progressive Supranuclear Palsy. *J Mov Disord*. 2016;9(1):3–13.

Manoochehri M, Huey ED. Diagnosis and management of behavioral issues in frontotemporal dementia. *Curr Neurol Neurosci Rep*. 2012;12(5):528–536.

McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT. Diagnosis and management of dementia with Lewy bodies: Third report of the dementia with Lewy bodies Consortium. *Neurology*. 2005;65:1863–1872.

*McKeith IG, Galasko D, Kosaka K, Perry EK, Dickson DW. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): Report of the consortium on Dementia with Lewy bodies international workshop. *Neurology*. 1996;47:1113–1124.

McKhann GM, Albert MS, Grossman M, Miller B, Dickson D, Trojanowski JQ. Clinical and pathological diagnosis of frontotemporal dementia: report of the Work Group on Frontotemporal Dementia and Pick's Disease. *Arch Neurol*. 2001;58(11):1803–1809.

McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging- Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement*. 2011;7:263–269.

Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief

screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695–699.

Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and meta-analysis. *Alzheimers Dement*. 2013;9:63–75.

*Rabins PV, Rovner BW, Rummans T, Schneider LS, Tariot PN. Guideline watch (October 2014): Practice guideline for the treatment of patients with Alzheimer's disease and other dementias. *American Psychiatric Association*. Accessed on-line October 28, 2016. http://psychiatryonline.org/guidelines

Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioral variant of frontotemporal dementia. *Brain*. 2011;134(9):2456–2477.

Seitz DP, Adunuri N, Gill SS, Gruneir A, Hermann N, Rochon P. Antidepressants for agitation and psychosis in dementia. *Cochrane Database Syst Rev*. 2011;2Art. No: CD008191.

*Walsh DM, Selkoe DJ. A beta oligomers—a decade of discovery. *J Neurochem*. 2007;101:1172–1184.

Wilson RS, Leurgans SE, Boyle PA, Schneider JA, Bennett DA. Neurodegenerative basis of age-related cognitive decline. *Neurology*. 2010;75 (12):1070–1078.

▲ 10.4 Amnestic Disorders and Mild Cognitive Impairment

CARMEN ANDREESCU, M.D., AND DANIEL VARON, M.D.

In the past, amnestic disorders referred specifically to "pure" memory deficits, most often defined in the context of substance abuse/dependence or in the context of a general medical condition. However, the past 15 years have witnessed the emergence of a new concept: *MCI*. As defined by the Report of the International Working Group on MCI, it requires the presence of cognitive decline not warranting the diagnosis of dementia and/or evidence of decline over time on objective cognitive tasks but with preserved basic activities of daily living. Thus, MCI differs from classic amnestic disorders both by its etiology and by the domains of its clinical presentation.

In the revised fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR), MCI is hosted by various provisions (cognitive disorder not otherwise specified [NOS], amnestic disorder NOS). In DSM-5, the term mild neurocognitive disorder now embraces cognitive disorder NOS and MCI (Table 10.4–1). According to DSM-5, amnestic disorder can be classified under major neurocognitive disorders if the severity of the dysfunction is enough to grant this diagnosis.

Table 10.4–1.
DSM-5 Diagnostic Criteria for Mild Neurocognitive Disorder

A. Evidence of modest cognitive decline from a previous level of performance in one or more cognitive domains (complex attention, executive function, learning and memory, language, perceptual-motor, or social cognition) based on:
   1. Concern of the individual, a knowledgeable informant, or the clinician that there has been a mild decline in cognitive function; and
   2. A modest impairment in cognitive performance, preferably documented by standardized neuropsychological testing or, in its absence, another quantified clinical assessment.
B. The cognitive deficits do not interfere with the capacity for independence in everyday activities (i.e., complex instrumental activities of daily living such as paying bills or managing medication are preserved, but greater effort, compensatory strategies, or accommodation may be required).
C. The cognitive deficits do not occur exclusively in the context of a delirium.
D. The cognitive deficits are not better explained by another mental disorder (e.g., major depressive disorder, schizophrenia).

Specify whether due to: Alzheimer disease, frontotemporal lobar degeneration, Lewy body disease, vascular disease, traumatic brain injury, substance/medication use, HIV infection, prion disease, Parkinson disease, Huntington disease, another medical condition, multiple etiologies, or unspecified.

Coding note: For mild neurocognitive disorder due to any of the medical etiologies listed above, code **331.83 (G31.84).** Do *not* use additional codes for the presumed etiological medical conditions. For substance/medication-induced mild neurocognitive disorder, code based on type of substance; see

"Substance/Medication-Induced Major or Mild Neurocognitive Disorder." For unspecified mild neurocognitive disorder, code **799.59 (R41.9).**

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Table 10.4–2.
Mild Cognitive Impairment Original Criteria

1. Memory complaint, preferably qualified by an informant
2. Memory impairment for age and education
3. Preserved general cognitive function
4. Intact activities of daily living
5. Not demented

MILD COGNITIVE IMPAIRMENT

Definition

Although the term MCI has been in use since 1982 as a descriptor of stage 3 on the Global Deterioration Scale, more recently it was developed by the Mayo Clinic Group as a diagnostic category designed to fill the gap between cognitive changes associated with aging and cognitive impairment suggestive of dementia. The criteria proposed by the Mayo Clinic Alzheimer's Disease Research Center (MCADRC criteria) are (1) memory complaint, preferably confirmed by an informant; (2) objective memory impairment for age and education; (3) preserved general cognitive function; (4) intact activities of daily living; and (5) not demented (Table 10.4–2). However, at this time there are no international diagnostic criteria for MCI. Given the emergent nature of the concept, there is a profusion of terms in the field.

In 2003 the concept was refined (Table 10.4–3) and expanded to include amnestic and nonamnestic subtypes, which in turn can be subdivided in single and multiple domain subcategories (Fig. 10.4–1). These divisions can be used when trying to approximate an etiology to the MCI syndrome. In 2011 a work group assembled by the National Institute on Aging (NIA) and the Alzheimer's Association (AA) delineated specific criteria for the diagnosis of MCI due to Alzheimer disease (AD) (Table 10.4–4) and more recently criteria for MCI due to Parkinson disease were formulated as well. Similarly the DSM-5 criteria for mild neurocognitive disorder include specifiers for multiple etiologies (Table 10.4–1). As the availability of biomarkers increases, criteria for each etiology will become progressively more specific.

Historical Perspective

The imprecise border between normal aging-related cognitive decline and

dementia-related cognitive impairment has been described for several decades. Thus, in 1962, Kral introduced the terms "Benign Senescent Forgetfulness" (forgetfulness for less important facts and awareness of problems) and "Malignant Senescent Forgetfulness" (memory problems for recent events and lack of awareness). In 1986, the National Institute of Mental Health (NIMH) recommended the term "Age Associated Memory Impairment" for age-related normal memory changes. However, the validity of this concept was disputed due to diagnostic tests based on performances of younger adults (Table 10.4–5).

Table 10.4–3.
Recommendations for the General Criteria for MCI

General Criteria for MCI
Not normal, not demented (does not meet DSM-IV-TR/ICD-10 criteria for a dementia syndrome)
Cognitive decline
Self and/or informant report and impairment on objective cognitive tasks
And/or
Evidence of decline over time on objective cognitive tasks
Preserved basic activities of daily living/minimal impairment in complex instrumental functions

MCI, mild cognitive impairment; ICD-10, *International Classification of Diseases*, 10th ed. From Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment: Beyond controversies, toward a consensus. Report of the International Working Group on Mild Cognitive Impairment. *J Intern Med*. 2004;256(3):240–246.

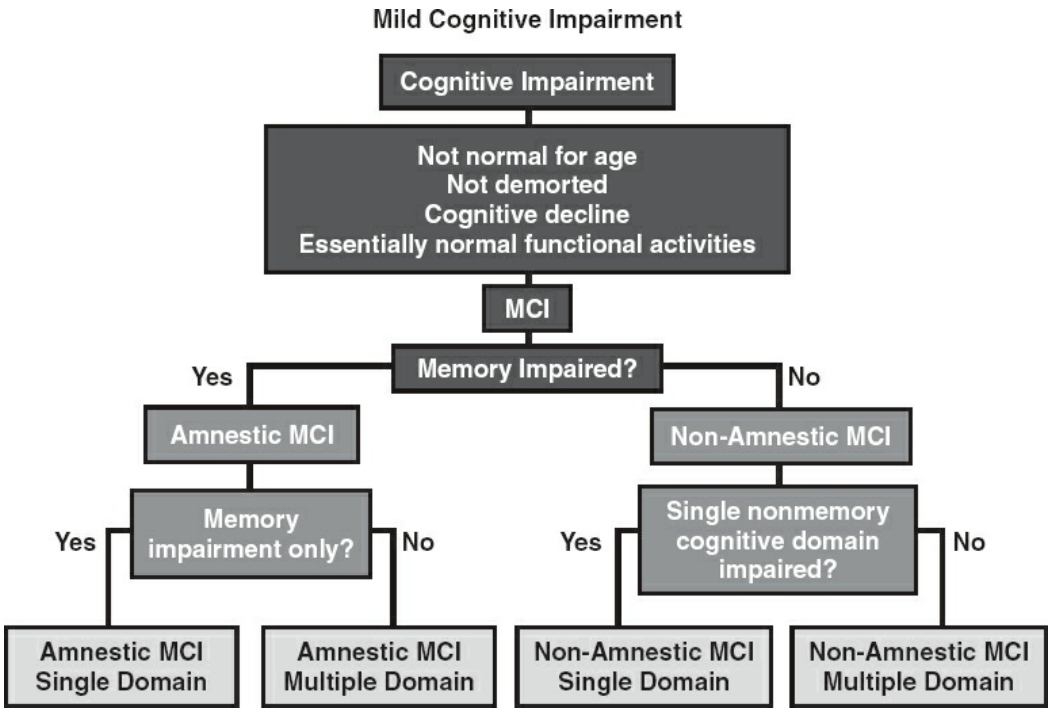

FIGURE 10.4–1. The current algorithm for diagnosing mild cognitive impairment (MCI) and its subtypes. (Reprinted with permission from Petersen RC. Mild cognitive impairment. *Continuum Lifelong Learning Neurol*. 2004;10(1):9–28.)

In 1994 the International Psychogeriatrics Association presented the concept of "Age-Associated Cognitive Decline," which described cognitive deficits, including but not limited to memory impairment in the absence of dementia or other affecting cognitive conditions.

"Cognitive Impairment No Dementia" was introduced in 1997 by the Canadian Study of Health and Aging to describe the presence of nondemented cognitive impairment regardless of the underlying process (neurologic, psychiatric, medical). It is thus a more inclusive concept (encompasses those with static encephalopathies who are not necessarily having a progressive worsening cognitive decline).

Several other classifications, including age-consistent memory impairment and late-life forgetfulness, are defined on the bases of performance on various cognitive tests.

Table 10.4–4.
Criteria for the Diagnosis of MCI NIA-AA Workgroup

A. Core Clinical Criteria:
   1. Concern regarding change in cognition
   2. Impairment in one or more neurocognitive domains
   3. Preservation of independence in functional abilities
   4. No evidence of dementia should be noted

B. Research criteria incorporating biomarkers:
   Diagnostic category……..Biomarker probability of……Aβ……Neuronal injury…
   etiology……(PET or CSF)…….(tau, FDG, sMRI)
   …… MCI-core clinical criteria…….Uninformative……
   Conflicting……
   Conflicting……Indeterminant……Indeterminant……..Unte
   Untested MCI due to AD-intermediate likelihood……..Intermediate……..Positive……Untest
   ……Untested……..Positive
   MCI due to AD-highlikelihood……Highest……Positive……..
   Positive MCI-unlikely due to AD……Lowest…….Negative……Negative

MCI, mild cognitive impairment; NIA, National Institute on Aging; AA, Alzheimer's Association; PET, positron emission tomography; sMRI, structural magnetic resonance imaging; FDG, fluorodeoxyglucose; Aβ, amyloid beta; CSF, cerebrospinal fluid.

From Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's Disease. *Alzheimers Dement*. 2011;7(3):270–279. (Public Access).

The concept of MCI is currently defined in DSM-5 as mild neurocognitive disorder. Just as functional impairment is an exclusion criterion for MCI, in mild neurocognitive disorder cognitive deficits should not interfere with capacity for independence in everyday activities. Further developments in finding biological markers for MCI will probably contribute to solidify the conceptualization and, hopefully, treatment of

prodromal dementia.

Epidemiology and Etiology of MCI

Various studies on the prevalence of MCI are presented in Table 10.4–6. The recognition that AD pathology may exist in the brain long before the presence of clinical symptoms led to the focus on preclinical stages, with the purpose of characterizing initial impairments that are associated with an increased risk of progression to AD.

The clinical expression of MCI can be viewed as a result of the interaction between several risk factors and several protective factors. The most significant risk factors are related to the different types of neurodegeneration witnessed in dementias. These are clinically expressed in different subtypes of MCI (e.g., amnestic MCI [a-MCI], single domain, is considered to be related to underlying AD pathology; a-MCI, multiple domains, is considered to be related to either AD or vascular dementia type of pathology; while nonamnestic MCI [na-MCI] has been related more closely to DLB or frontotemporal neurodegeneration). Other risk factors include the apolipoprotein E 4 (APOE4) allele status and cerebrovascular events in the form of either cerebrovascular accident or lacunar disease. The role of chronic exposure to high levels of cortisol, as seen in late life depression, is also hypothesized to increase the risk of cognitive impairment through hippocampal volume reduction. The notion of "brain reserve" suggests that effects of brain size and neuron density may be protective against dementia despite the presence of neurodegeneration (a larger number of neurons and a bigger brain volume would protect against clinical manifestations of AD despite the presence of neurodegeneration) (Fig. 10.4–2).

Table 10.4–5.
Terms Related to Mild Cognitive Impairment

Term	Author(s)	Year	Inclusion Criteria	Observations
Malignant senescent forgetfulness (MSF)	VA Kral	1962	Memory difficulties for recent events Lack of awareness regarding the memory deficit	Two-year follow-up showed a faster evolution of patients with MSF toward dementia
Age-associated memory impairment (AAMI)	NIMH (T Crook, RT Bartus, SH Ferris)	1986	Age-related memory disturbances leading to (1) subjective concern; (2) functional problem No underlying neurological illness	Memory tests were validated on young populations, leading to high rates of AAMI in the elderly
Age-associated cognitive	International Psychogeriatric	1994	Cognitive deficits not meeting the criteria	Does not include prognosis regarding

decline (AACD)	Association and World Health Organization (R Levy)		for dementia	evolution to dementia Includes several kinds of cognitive decline (not exclusive memory decline)
Cognitively impaired no dementia (CIND)	Canadian Study of Health and Aging	1997	Age 65 and older Multiple cognitive problems not meeting DSM-III-R criteria for dementia	Includes static encephalopathies

NIMH, National Institute of Mental Health; DSM-III-R, *Diagnostic and Statistical Manual of Mental Disorders*, third revised edition.

MCI Clinical Presentation

The clinical picture of MCI is a function of the criteria used to define it. The MCADRC criteria emphasize the memory impairment, but do not specify how it should be evaluated. The second criterion (objective memory impairment for age and education) is often interpreted as objective deficit of memory or another cognitive domain that is more than 1.5 standard deviations below the mean for individuals of similar age and education. The inherent heterogeneity driven by the novelty of the concept and the lack of consensus between experts has been recognized by the current concepts in MCI conference, which classified MCI in different subgroups, each associated with an increased risk of developing a type of dementia. Thus, a-MCI usually progresses to AD, multiple domains MCI (md-MCI) may progress to either AD or another type of dementia, while MCI-single nonamnestic domain develops predominantly into non-Alzheimer dementia like frontotemporal dementia or DLB. With the introduction of the NIA-AA criteria for MCI due to AD two sets of criteria were outlined: The core clinical criteria are meant to be used broadly, in any setting and without requiring any highly specialized tests or procedures and The clinical research criteria, which incorporate the use of biomarkers, are intended to be used only in research, academic centers, and clinical trials (Table 10.4–4). Biomarkers are not included in the DSM-5 criteria for mild neurocognitive disorder due to AD; however, both the NIA-AA working group and DSM-5 task force acknowledge these biomarkers will most likely make their way into clinical practice once they are adequately validated and widely available.

The correlation between a clinical presentation of MCI and dementia outcome varies. The annual “conversion” rate was 12 percent in the Mayo MCI sample (compared with 1 to 2 percent for controls), while the numbers vary from 8 to 100 percent depending on the setting (community vs. clinical) and the study design (prospective vs. retrospective). Thus, in the Monongahela Valley Independent Elderly Survey (a community-based study), although 10 to 17 percent of subjects with MCI progressed to

dementia, 33 to 56 percent became non-MCI and half of those returned to normal (Table 10.4–7).

MCI Assessment

Neuropsychological Assessment. Most experts agree that earlier deficits are noted in episodic (vs. semantic) memory. There is no consensus among experts with regard to which memory tests and which cutoffs to use. There is a lack of norms, test scores do not have normal distributions, and test performance is influenced by multiple demographic characteristics. Several experts have proposed that a scale like the delayed recall task from the Consortium to Establish a Registry for Alzheimer's Disease might be useful in detecting AD in the earliest stages. Brief mental status instruments (like mini-mental state examination) are relatively insensitive for the detection of memory problems in MCI. The Montreal Cognitive Assessment (MoCA) has been used in recent years as a more adequate way to screen patients with MCI. It has been validated multiple times, in different pathologies, populations, and languages. It includes more adequate testing for frontal lobe dysfunction when compared to the MMSE and is readily available.

Table 10.4–6.
Prevalence of MCI in Different Studies

Study	Year	Terminology	Prevalence	Observations
Canadian Study of Health and Aging (Graham et al.)	1997	CIND	16%	CIND with "circumscribed memory loss"—prevalence of 5.3%
Religious Orders Study (ROS) (Bennett et al.)	2002	MCI	26.4%	The MCI definition in this study is similar to CIND They used a volunteer-based sample
Monongahela Valley Independent Elderly Survey (Ganguli et al.)	2004	MCI	2.9–4%	11–17% progressed to dementia 33–56% reversed to non-MCI
Indianapolis Study of Health and Aging (Unverzagt et al.)	2001	CIND	23.4%	
Cardiovascular Health Study (Lopez et al.)	2003	MCI	6% amnestic MCI 16% multiple domains MCI	21.5% of the md-MCI did not have a memory impairment
Leipzig Longitudinal Study of the Aged (Busse et al.)	2003	MCI	3–20%	Multiple definitions of MCI used
Eugeria Study (Ritchie et al.)	2001	MCI AACD	3.2% for MCI 20.9% AACD	AACD included MCI

CIND, cognitively impaired no dementia; MCI, mild cognitive impairment; md-MCI, multiple domain mild cognitive impairment; AACD, age-associated cognitive decline.

MCI Subtypes

Clinical classification		Etiology: Degen-erative	Vascular	Psychiatric	Medical conditions
Amnestic MCI	Single domain	AD		Depr	
	Multiple domain	AD	VaD	Depr	
Non-amnestic MCI	Single domain	FTD			
	Multiple domain	DLB	VaD		

FIGURE 10.4–2. Outcome of clinical phenotypes of mild cognitive impairment (MCI) according to presumed etiology. AD, Alzheimer disease; Depr, depression; VaD, vascular dementia; FRD, frontotemporal dementia; DLB, dementia with Lewy bodies. (Adapted from Petersen RC. Aging, mild cognitive impairment, and Alzheimer's disease. *Neurologic Clinics*. 2000;18(4):789–805.)

The NIA-AA workgroup suggests tests of episodic memory such as the Free and Cued Selective Reminding Test, the Rey Auditory Verbal Learning Test, the California Verbal Learning Test, or the Logical Memory I and II of the Wechsler Memory Scale Revised can be used. Tests to assess other cognitive domains are also mentioned including the Trail Making Test for executive function, the Boston Naming Test, letter and category fluency for language, figure copying for spatial skills, and digit span forward for attention.

Biomarkers. Several markers of progression from MCI to AD have been studied in the past decade. The most well-established biomarkers have been divided into two categories, those that measure amyloid deposition and those that relate to neurodegeneration. Low levels of $A\beta_{42}$ (the 42 amino acid form of β-amyloid) and increased uptake of amyloid PET tracers are associated with fibrillar Aβ deposits. Increased levels of cerebrospinal fluid (CSF) total tau (t-tau) and phosphorylated tau (p-tau), hypometabolism on FDG PET, and atrophy on magnetic resonance imaging (MRI) are measures of neurodegeneration. The ratio of CSF Aβ levels to t-tau and phosphotau-231 (ptau-231) appears to be the best biomarker in CSF. Locating alterations in the expression of proteins involved in the pathogenetic pathways of AD (proteomic approach) is another approach

used to help early detection of AD. Several proteins (cystatin C, β-2 microglobulin, and BEGF polypeptides) have been detected through new techniques, and currently there are a number of proteins from both CSF and blood that are implicated in AD pathology.

Genetics. As MCI is regarded as the prodromal stage for several disorders (AD, frontotemporal, Lewy body disease or vascular dementia), different genes are probably related to MCI. Three genes involved in rare autosomal dominant forms of AD have been identified: the amyloid precursor protein (*APP*) gene, presenilin-1 (*PSEN1*), presenilin-2 (*PSEN2*). Screening for each of these mutations will have very limited value for the diagnosis of MCI in the general population unless a clear family history is present. Multiple other genes have been associated with sporadic AD. Among those APOE has the greatest effect in subjects with MCI. Longitudinal studies have demonstrated that subjects with MCI who carry a copy of the E4 allele are at higher risk of progressing to dementia and demonstrate greater degree of hole brain and hippocampal atrophy. However, testing for APOE genotype is not currently recommended since 75 percent of people with one copy of the E4 allele remain free of AD and up to 50 percent of people with AD do not carry the E4 allele. As the etiology of MCI is heterogeneous, it is likely that a very large number of different genes underlie the pathology of MCI.

Table 10.4–7.
Conversion Rates from MCI to Alzheimer Disease in Different Studies

Study	Conversion Rate	Observations
Petersen et al. (1999)	12% per year	Clinical sample, *N* = 76
Ganguli et al. (2004)	10–17% every 2 years	Community sample
Monongahela Valley Independent Elderly Survey	27% at 10 years	10-year follow-up Fourfold increased risk for progression to Alzheimer disease in MCI group compared with controls
Boston Study-Fisk et al. (2005)	18.7% over 3 years	Community volunteers
St. Louis Study-Morris et al. (2001)	100% over 9.5 years	Clinical sample
Eugeria Study-Ritchie et al. (2001)	11.3% over 3 years	Primary care based clinical sample
Larrieu et al. (2002) QUID Study (Personnes Agées)	8.3% per year	Community-based France 40% reverted to normal after 2 years

MCI, mild cognitive impairment.

Neuroimaging

Advances in neuroimaging studies aim to develop measures allowing the differentiation between MCI and healthy aging as well as within MCI between subjects who will convert to AD or will remain stable over time. This objective has led to a worldwide effort that started in the United States with the Alzheimer's Disease Neuroimaging Initiative (ADNI) and has expanded to multiple other countries including Japan, China, Taiwan, Argentina, and Australia. These studies include a large number of subjects with subjective cognitive impairment, MCI, and mild dementia who are being followed over time. A similar project has been developed in Parkinson disease.

Structural volumetric studies of subjects with MCI showed early changes in the medial temporal structures, reflecting neuronal atrophy, decreased synaptic density, and overall neuronal loss seen in postmortem studies. Atrophy of the hippocampal volume and entorhinal cortex has been described in MCI. Atrophy of the hippocampal formation and entorhinal cortex using both volumetric methods and visual rating scales has been reported to predict the rate of progression from MCI to AD. Three-dimensional modeling techniques have localized shape alteration and specific regions of atrophy within the hippocampus. Other methods like tensor-based morphometry allow tracking brain changes in detail, quantifying tissue growth or atrophy throughout the brain, and indicating the local rate at which tissue is being lost. Other innovations in neuroimaging include MR relaxometry, imaging of iron deposition, and diffusion tensor imaging that have been used to assess homeostasis of iron in the gray matter and disruption of myelin in the white matter both of which correlate with incipient pathological changes in subjects with AD.

Perhaps the most promising development has been the advent of PET tracer compounds that visualize amyloid plaques and neurofibrillary tangles. Pittsburgh Compound B (carbon-11-PIB) opened the door to imaging amyloid deposition in vivo, but limitations due to cost and half-life of the radiotracer restricted its use to research. The development of fluorinated radiotracers addressed some of these concerns and allowed the approval of these methods by the U.S. Food and Drug Administration (FDA) for clinical use. To date three compounds have been approved and guidelines for their use in clinical practice have been outlined (Table 10.4–8). More recently tracers to visualize neurofibrillary tangles have been developed as well and are being studied but are not yet available for clinical use. Specific tracers that visualize amyloid and tangle deposition allowed investigators to evaluate the pathological process in vivo and are being used to monitor progression from MCI to AD. However, the burden of β-amyloid plaques does not always correlate with the clinical stages, as some subjects with MCI can present with minimal burden similar to healthy controls,

while others have amyloid burden comparable to subjects with AD (Fig. 10.4–2). Currently no single biomarker appears sufficient to identify incipient AD. Thus, the combination of several markers further increases the accuracy of the prediction and will probably become the norm as a more specific outline of the disease course is established (Fig. 10.4–3).

MCI Diagnostic Differential

The Cognitive Continuum. The cognitive continuum describes the subtle pathway from age-related cognitive decline to MCI to dementia. Per this model, there is an overlap at both ends of MCI, which indicates that it can be quite challenging to identify the transition points (Fig. 10.4–4). In practice, differentiating MCI from age-related cognitive decline resides mainly on neuropsychological testing, showing a cognitive decline more severe for age and less education. The main differentiation between MCI and AD resides in the lack of functional impairment in MCI. An attempt to correlate the cognitive changes within this continuum with the underlying pathological processes has been underway in recent years.

Information gathered from longitudinal studies suggests different pathways during MCI can lead to a similar presentation and pathological findings when progression to dementia has occurred. In the case of AD, an initial model presented in 2009 suggested the major AD biomarkers became abnormal in a temporally ordered manner with amyloid accumulation (CSF and PET amyloid measures) occurring early, followed by CSF tau, FDG PET, and structural MRI occurring later and leading to clinical changes. The MCI syndrome was thought to become evident once a large portion of amyloid had been accumulated and only minimal neurodegeneration had occurred. A more recent model suggests biomarkers do not occur in sequence but rather become simultaneously more abnormal. Within this model β-amyloid is considered necessary but not sufficient to produce the AD syndrome and while its accumulation appears to loosely correlate with the rate of cognitive decline, neurodegeneration biomarkers appear closely associated with decline in both typical and atypical forms of the disease. The time lag between biomarker evidence and appearance of clinical symptoms would be moderated by factors such as cognitive reserve and comorbid disorders.

Table 10.4–8.
Criteria for the Use of Amyloid PET Imaging

Appropriate use would be considered in individual with the following characteristics:

1. A cognitive complaint with objectively confirmed impairment
2. AD as a possible diagnosis, but the diagnosis is uncertain after a comprehensive evaluation by a dementia expert
3. When knowledge of the presence or absence of Aβ pathology is expected to increase diagnostic certainty and alter management

a. Patients with persistent or progressive unexplained MCI
b. Patients satisfying core clinical criteria for possible AD because of unclear clinical presentation, either an atypical clinical course or an etiologically mixed presentation
c. Patients with progressive dementia and atypically early age of onset (usually defined as 65 year or less in age)

Amyloid imaging is inappropriate in the following situations:

1. Patient with core clinical criteria for probable AD with typical age of onset
2. To determine dementia severity
3. Based solely on a positive family history of dementia or presence of APOE ε4
4. Patients with a clinical complaint that is unconfirmed on clinical examination
5. In lieu of genotyping for suspected autosomal mutation carriers
6. In asymptomatic individuals
7. Nonmedical use (e.g., legal, insurance coverage, or employment screening)

AD, Alzheimer disease; MCI, mild cognitive impairment; Aβ, amyloid beta; APOE, apolipoprotein E.

From Johnson KA, Minoshima S, Bohnen NI, et al. Appropriate use criteria for amyloid PET: a report of the Amyloid Imaging Task Force, The Society of Nuclear Medicine and Molecular Imaging, and the Alzheimer's Association. *Alzheimers Dement*. 2013;9(1):e-1–e16.

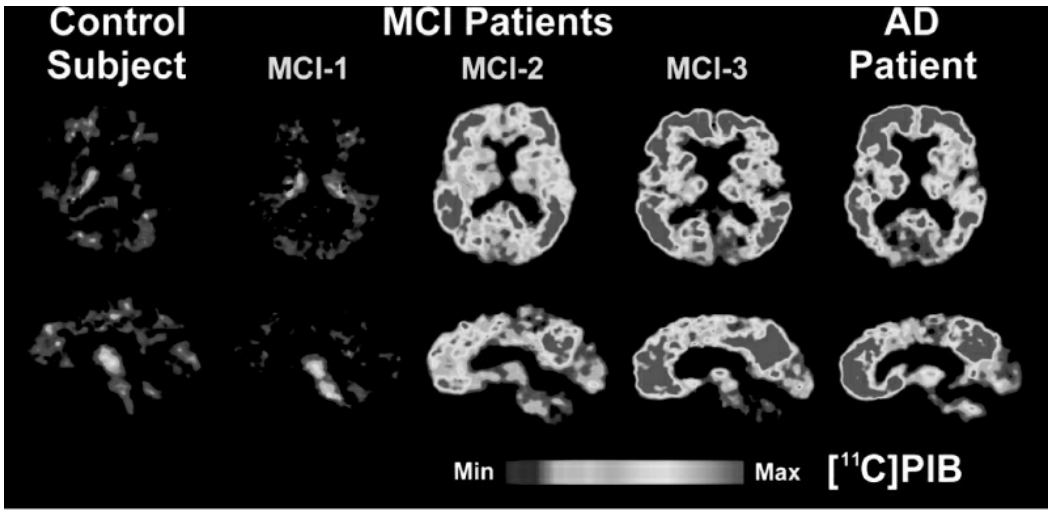

FIGURE 10.4–3. Positron emission tomography images obtained with the amyloid-imaging agent Pittsburgh Compound-B ([carbon-11]-PIB) in a normal control (*far left*), three different patients with mild cognitive impairment (MCI; *center images*), and a mild Alzheimer disease (AD) patient (*far right*). Some patients with MCI have control-like levels of amyloid, some have AD-like levels of amyloid, and some have intermediate levels. (Courtesy of William E. Klunk, M.D., University of Pittsburgh, Department of Psychiatry, Pittsburgh, PA. All rights retained.)

More recent longitudinal data from ADNI and Mayo Clinic has demonstrated some puzzling findings. Patients with MCI followed over time can be subdivided according to their positivity in biomarkers. Some will have only amyloid deposition, others only evidence of neurodegeneration, a third group both amyloid deposition and neurodegeneration and lastly a fourth group with no evidence of either one. Subjects with evidence of both amyloid deposition and neurodegeneration appear to progress at the fastest rate towards dementia, while subjects with

no evidence of either one progress very slowly if at all. However, those with only amyloid deposition do not progress as quickly as those with only neurodegeneration, in fact those with neurodegeneration progress almost as quickly as the ones with both amyloid and neurodegeneration in the Mayo sample. These findings have speared the evolution of a new concept, SNAP (Suspected Non-Alzheimer Pathology), suggesting that some subjects with MCI have an entity that resembles AD but behaves differently at the pathological level. More recently others have suggested alternative pathways to AD initiating with amyloid deposition or neurodegeneration but finally leading to the same end result. The integration of data generated from following subjects who are asymptomatic but have presence of biomarkers, people with subjective complaints who do not meet criteria for MCI, and very early MCI subjects will allow a much better characterization of all the possible pathological pathways within the cognitive continuum and their clinical correlations.

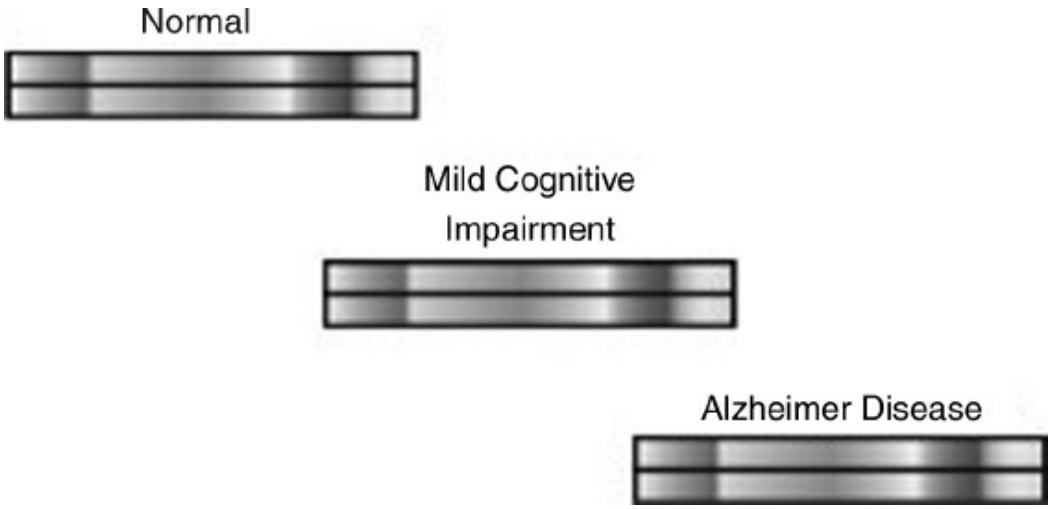

FIGURE 10.4–4. Cognitive continuum showing the overlap in the boundary between normal aging and the mild cognitive impairment and Alzheimer disease. (From Petersen RC. Normal aging, mild cognitive impairment, and early Alzheimer's disease. *Neurologist*. 1995;1326–344.)

Figure 10.4–5 outlines the various etiological pathways associated with an initial MCI clinical presentation.

Course and Prognosis

The typical rate at which patients with MCI progress to AD is 10 to 15 percent per year and is associated with progressive loss of function (Fig. 10.4–6). However, there are several studies indicating that the diagnosis is not stable in both directions; patients can either convert to AD or revert back to normal. This variability in course is related to the heterogeneous source of the subjects (clinical vs. community) as well as to the heterogeneous definition criteria used by different studies. a-MCI has been

associated with increased mortality (greater in multiple-domains a-MCI than in single-domain a-MCI) (hazard ratio 1.7) when compared with reference subjects.

Epidemiological studies that followed over time large numbers of community-dwelling subjects, who eventually developed dementia, consistently demonstrate a slow change in cognitive measures with a sharp increase in rate of decline seen few years before the clinical diagnosis of MCI or dementia is made. The point in time when this inflection occurs varies from study to study. In some, the rate of cognitive decline increases sharply 5 to 6 years before the AD diagnosis, while in others the change is seen only 3 years before. Subjects diagnosed with a-MCI appear to present with cognitive changes earlier and declined more rapidly when compared to non-amnestic MCI (na-MCI) subjects. The clinical manifestations and rate of decline in MCI are likely not only dependent on the underlying neuropathology, but also on a host of factors that can either exacerbate or ameliorate the biological impact of the underlying disease. Some of these factors include but are not limited to the expression of specific genes, the age of disease onset, baseline brain and cognitive reserve, comorbid medical and psychiatric illness, cerebrovascular burden, and lifestyle characteristics (e.g., diet, physical and cognitive activity). The manipulation of some of these factors has been considered as a potential way to decrease the likelihood of progression from MCI to dementia.

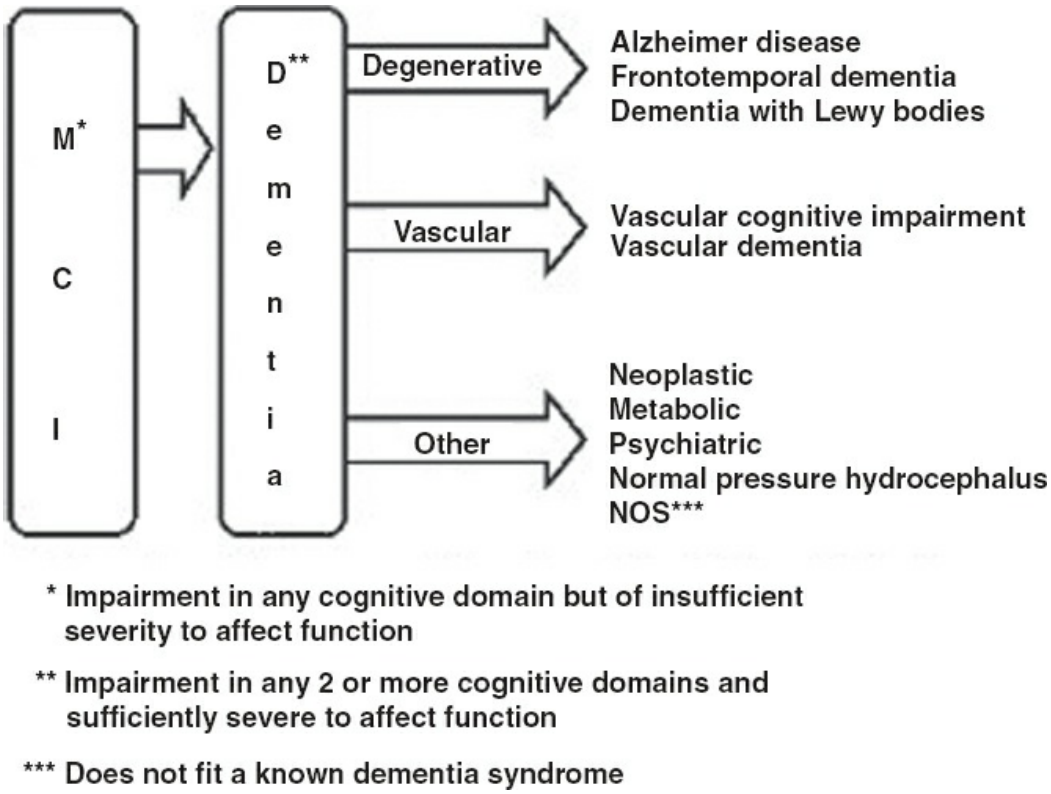

FIGURE 10.4–5. Overview of mild cognitive impairment preceding the diagnosis of dementia. NOS, not otherwise specified. (Adapted from Petersen RC. Aging, mild cognitive

impairment, and Alzheimer's disease. *Neurologic Clinics*. 2000;18(4):789–805.)

MCI Treatment

There are no FDA approved treatments of MCI at this time. MCI treatment involves adequate screening and diagnosis. Ideally, MCI treatment would also include improvement in memory loss, together with prevention of further cognitive decline to dementia. Cognitive training programs have been reported as mildly beneficial for compensating memory difficulties in MCI. Controlling for vascular risk factors (high blood pressure, hypercholesterolemia, diabetes mellitus) may be a preventive method for those MCI cases underlying vascular pathology. Currently, sensitive tools (imaging techniques or biomarkers) are not available for MCI screening in the general population. In primary care setting, clinicians should maintain a high suspicion for subjective cognitive complaints and should corroborate these complaints with collateral information whenever possible. Also, identifying reversible causes of cognitive impairment (hypothyroidism, vitamin B_{12} deficiency, medication-induced cognitive impairment, depression) can further benefit some of the prodromal dementia MCI cases.

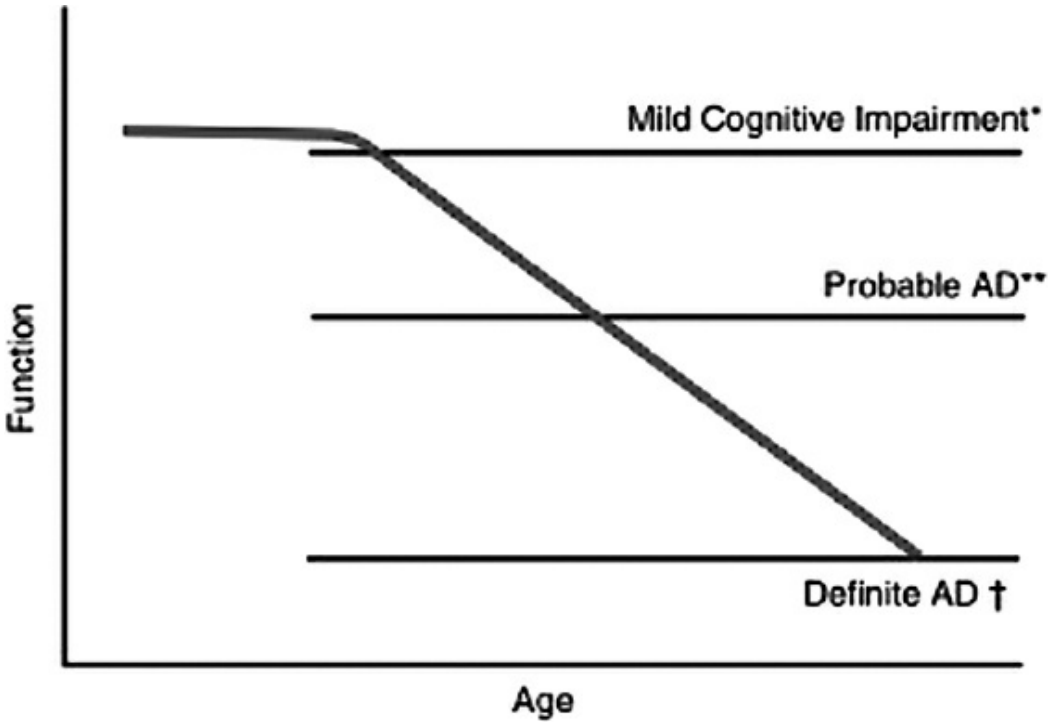

FIGURE 10.4–6. Hypothetical change in function as an individual develops Alzheimer's disease. (Adapted from Petersen RC. Aging, mild cognitive impairment, and Alzheimer's disease. *Neurologic Clinics*. 2000;18(4):789–805.)

Currently, there is no evidence for long-term efficacy of pharmacotherapies in reversing MCI. Several epidemiological studies indicated a reduced risk of dementia in persons taking antihypertensive

medications, cholesterol-lowering drugs, antioxidants, anti-inflammatory, and estrogen therapy, but no randomized controlled trials verify these data. With regard to cognitive enhancers, so far, multiple trials designed for a-MCI have reported mostly negative results (Table 10.4–9). Most of these studies were confronted with several problems: (1) obtaining homogeneous samples and identifying potential beneficiaries of treatment; (2) treating a wider population, which led to large percentages of negative responses and problematic side effects; and (3) translation of the MCI construct into multiple cultures and languages and using AD diagnosis as the primary outcome, given the variability of this diagnosis in different countries. Clinical trials using immunological methods including active immunization, monoclonal antibodies and intravenous immunoglobulin (IVIG) have been used in an attempt to reduce amyloid burden and in turn slow or stop disease progression in patients with mild or moderate AD. Unfortunately these attempts have been unsuccessful suggesting that targeting amyloid in AD once amyloid burden is established and neurodegeneration is progressing might be too late. Although some of these agents (IVIG, gantenerumab, ACC001) are being studied currently in MCI subjects, a new series of trials to treat *asymptomatic* subjects with evidence of amyloid deposition are now on their way as well. Both subjects with autosomal dominant inherited forms of AD and elderly subjects without genetic predisposition are being studied. If positive, these trials could lead the way in treating neurodegenerative disorders during the preclinical stages of disease preventing people at risk from ever becoming symptomatic.

FUTURE DIRECTIONS

Advances in MCI detection will be paramount for early detection and treatment of AD, as experts agree that disease-modifying treatments of AD will focus on cognitively intact individuals at increased risk. The field of identifying sensitive and specific biomarkers (biological and neuroimaging markers) will probably witness an exponential development in the coming years.

Table 10.4–9.
Treatment Trials for MCI

Study	*N*	Duration	Primary Outcome	Results	OBS	Sponsor
Donepezil + vitamin E (Thall et al., 1999)	769	3 years	Conversion to AD	Partially positive (reduced risk of developing AD in the active arm	Amnestic MCI status and the presence of APOE4 allele–predictive of rate of progression to	ADCS

				group for the first 12 months)	AD	
Donepezil (Salloway et al., 2004)	269	24 weeks	ADAS-Cog total score; NYU PTIR	Negative	Positive results in secondary outcome measures (ADAS-Cog13)	Pfizer (The Donepezil "401" Study Group)
Donepezil (Doody et al., 2009)	821	48 weeks	ADAS-Cog cognitive subscale, CDR-SB	Minimally positive with small decrease in ADAS-Cog score in donepezil group. No significant change in CDR-SB	On secondary measures only global assessment differed significantly in favor of donepezil	Eisai and Pfizer
Rivastigmine (Feldman et al., 2007)	1,018	48 months	Conversion to AD	Negative		Novartis
Galantamine (Winbald et al., 2008) Two studies	2,048		Progression of CDR score (from .5 to 1) Incident dementia	Negative	Attention assessed by DSST favored galantamine in both studies	Janssen-Cilag and Johnson & Johnson
Piribedil (Nagaraja et al., 2001)	60	12 weeks	MMSE score	Favored piribedil	Mean MMSE change from baseline also favored intervention group	NIMH
Rofecoxib (Thall et al., 2005)	1,457	3–4 years	Conversion to AD	Negative	Primary outcome favored placebo while secondary outcomes (ADS-cog, CDR) did not differentiate between rofecoxib and placebo	Merck
Nicotine patch (Newhouse et al., 2012)	74	6 months	Attention on CCPT and global functioning on CGIC	Nicotine improved attention but not global functioning	Secondary outcome measures showed improvements in attention, memory, and psychomotor speed.	

Piracetam	675	12 months	Composite score extracted from eight tests	Negative		UCB Pharma
Ginkgo biloba (DeKosky et al., 2008) (Snitz et al., 2009)	3,069	6.1 years	Incident dementia	Negative	No effect noted in normal cognitive either	NCCAM and NIH
		6.1 years	Rates of change in MMSE, ADAS-Cog, and other cognitive domains	Negative		
B Vitamins (B_{12}, B_6, folic acid) (van Uffelen et al., 2008)	152	1 year	Cognitive function at 6 and 12 months	Negative	Moderate intensity walking improved memory in men and memory and attention in women	TNO-VU University Medical Center

AD, Alzheimer disease; MCI, mild cognitive impairment; ADCS, Alzheimer Disease Cooperative Study; CDR, Clinical Dementia Rating; NYU PTIR, New York University Paragraph Test Immediate Recall; DSST, Digit Symbol Substitution test; CCPT, Connors Continuous Performance Test; NCCAM, National Center for Complimentary Medicine.

AMNESTIC DISORDERS

Definition

The DSM-IV-TR would provide four diagnostic entities under the category of amnestic disorder: (1) amnestic disorder due to cerebral or systemic medical condition, (2) substance-induced amnestic disorder, (3) amnestic disorder due to unknown etiology, and (4) amnestic disorder NOS. The recently developed DSM-5 criteria for mild and major neurocognitive disorder include all these categories (Table 10.4–1).

In tenth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10), amnestic disorder is defined as "Organic amnestic syndrome, not induced by alcohol and other psychoactive substances." For substance-induced amnestic disorder, ICD-10 uses the term "Amnestic syndrome." The research criteria of ICD-10 are presented in Table 10.4–10.

Epidemiology

As discussed in previous editions, data regarding the overall prevalence of amnestic disorders is lacking. The prevalence of alcohol-induced amnestic disorder varies from 0.2 to 4 percent, while a study of the incidence of transient global ischemia is reporting 5.2 cases per 100,000 per year.

Table 10.4–10.
ICD-10 Research Criteria for Organic Amnestic Syndrome

A. Memory impairment manifest by both:
   1. A defect of recent memory (impaired learning of new material), to a degree sufficient to interfere with daily living; and
   2. A reduced ability to recall past events

B. Absence of:
   A. A defect in immediate recall (as tested, for example, by the digit span)
   B. Clouding of consciousness and disturbance of attention, as defined in delirium
   C. Global intellectual decline (dementia)

C. Objective evidence (physical and neurological examination, laboratory tests) and/or history of an insult to or a disease of the brain (especially involving bilaterally the diencephalic and medial temporal structures but other than alcoholic encephalopathy) that can reasonably be presumed to be responsible for the clinical manifestations described under A

Comments: Associated features, including confabulations, emotional changes (apathy, lack of initiative), and lack of insight, are useful additional pointers to the diagnosis but are not invariable present.

From World Health Organization. *International Statistical Classification of Diseases and Related Health Problems*. 10th ed. Geneva: World Health Organization; 1992, with permission.

Etiology

The causes of amnestic disorders are listed in Table 10.4–11. The neuroanatomical structures involved in the etiology of amnestic memory have been described inferentially from matching various clinical syndromes with lesions in specific brain areas. These areas include the mammillary bodies, the hippocampus, the amygdala, the dorsomedial, and the midline thalamic nuclei. Bilateral lesions are considered as the main culprit, but there are studies reporting unilateral (left regions being more often involved) damage as the cause of amnestic disorders. Some experts differentiate between hippocampal amnesia and diencephalic amnesia, although both manifest with disproportionate impairment of anterograde memory and time-related impairment of retrograde memory (worse closer to illness's onset), while maintaining an intact implicit memory, immediate recall, and attention.

Table 10.4–11.
Causes of Amnestic Disorders

Primary Brain

Conditions	Notes
Seizure disorders	
Traumatic brain injury	The fastest growing cause of Amnestic disorder in the United States
Cerebral tumors	Mainly localized around the third ventricle
Stroke	In the posterior cerebral artery territory
Paraneoplastic syndromes (limbic encephalitis)	Focal inflammation of the temporal lobes
Infections (herpes simplex encephalitis)	Preferentially attacks the temporal lobes and produces a hemorrhagic necrosis
Metabolic (hypoxia)	Several causes (respiratory, cardiovascular, metabolic compromise, or carbon monoxide intoxication) attacking the sensitive CA1 hippocampal region
Transient global amnesia	Profound anterograde and retrograde amnesia improving over the course of several days
ECT	Retrograde amnesia that usually remits after days/weeks
Systemic conditions	
Alcohol-induced amnestic syndrome (Korsakoff syndrome)	
Thiamine deficiency	
Hypoglycemia	
Substance-induced amnestic disorder (excluding alcohol) sedative-hypnotics	

ECT, electroconvulsive therapy.

Symptoms

Memory impairment is the core symptom of all amnestic disorders. It includes both anterograde amnesia (inability to learn new information) and retrograde amnesia (inability to recall previously learned information). Short-term memory can also be impaired to a variable degree. Typically, the long-term retrograde memory impairment is temporally graded (more remote memories are better preserved). Immediate recall is usually not affected. Attention and implicit learning (which includes procedural learning as well as other forms of learning without explicit awareness) are also intact.

Associated symptoms include confabulations, changes in personality (apathy, lack of initiative, impulsivity), as well as neurological symptoms corresponding to the underlying illness (stroke/multiple sclerosis/alcohol-induced amnestic syndrome). The confabulations are more frequent in diencephalic amnesia (prototypical Korsakoff syndrome) than in hippocampal amnesia.

The following vignette presents the case of Clive Wearing, who developed severe anterograde amnesia following herpes encephalitis. His wife, Deborah Wearing, described his condition in her 2005 book *Forever Today*:

> His ability to perceive what he saw and heard was unimpaired. But he did not seem to be able to retain any impression of anything for more than a blink. Indeed, if he did blink, his eyelids parted to reveal a new scene. The view before the blink was utterly forgotten. Each blink, each glance away and back, brought him an entirely new view. I tried to imagine how it was for

him . . . Something akin to a film with bad continuity, the glass half empty, then full, the cigarette suddenly longer, the actor's hair now tousled, now smooth. But this was real life, a room changing in ways that were physically impossible. . . . It was as if every waking moment was the first waking moment. Clive was under the constant impression that he had just emerged from unconsciousness because he had no evidence in his own mind of ever being awake before . . . "I haven't heard anything, seen anything, touched anything, smelled anything," he would say. "It's like being dead."

Korsakoff Syndrome. In 1889, Serghei Korsakoff described a syndrome of polyneuritis, anterograde amnesia, and confabulations in subjects with chronic alcohol use. He did not, however, differentiate between the acute stage of the illness (the acute encephalopathy had been described 20 years earlier by Carl Wernicke) and the chronic memory impairment. That connection between the acute and the chronic condition was made at the beginning of the 20th century by Karl Boenhoffer. Later, in the eighth edition of his classic textbook, Emil Kraepelin distinguished between alcohol-related Korsakoff syndrome and Korsakoff syndrome related to other nonalcoholic processes (e.g., neurosyphilis). The etiological correlation between Korsakoff syndrome and thiamine deficiency caused by poor nutrition associated with chronic alcohol use was described in the early 1930s. Maurice Victor described the corresponding diencephalic lesion in the 1960s. Korsakoff syndrome usually develops secondary to Wernicke encephalopathy, and frequently they are described together (Wernicke–Korsakoff syndrome) (for a description of Wernicke encephalopathy clinical presentation, see Section 10.2 on Delirium). However, some cases of Korsakoff syndrome develop in the absence of the initial stage of acute Wernicke encephalopathy. Throughout the years, the terms Korsakoff psychosis, Korsakoff dementia, and Korsakoff syndrome have been used interchangeably to describe the clinical presentation of alcohol-induced amnestic disorder.

The overall rates of both Wernicke encephalopathy and Korsakoff syndrome are as low as 0.4 percent in France (although they have one of the highest per capita consumptions of alcohol) and as high as 2 percent in the United States, Scandinavia, and Australia. It is usually seen in subjects over age 40 (as it follows alcohol use of 20 or more years). Beyond the malnutrition, alcohol use induces thiamine deficiency through folate deficiency (which reduces thiamine absorption from the gastrointestinal tract) and through direct damage of the intestinal mucosa.

Primary thiamine deficiency is caused by general malnutrition and other conditions like anorexia nervosa, acquired immune deficiency syndrome (AIDS), hyperemesis gravidarum, thyrotoxicosis, metastatic cancer, hypomagnesemia, long-term dialysis, and congestive heart failure treated long-term with diuretics (renal tubular reabsorption of filtered thiamine is exacerbated by furosemide).

Thiamine pyrophosphate (TPP), the active form of thiamine, is involved in three enzymatic reactions essential for the metabolism of glucose and

the production of several neurotransmitters, including acetylcholine.

Numerous mechanisms have been suggested to explain the selective loss of neurons found with thiamine deficiency. These include cerebral energy dysfunction, breakdown of the blood–brain barrier, altered glutamate neurotransmission caused by extracellular glutamate and its impaired transport function, together with *N*-methyl-D-aspartate (NMDA) receptor-mediated excitotoxicity, the accumulation of amyloid precursor-like protein, increased free radical production, increased expression of superoxide dismutase consequent to increased microglial response, induction of nitric oxide, and oxidative stress. Changes in a number of central nervous system (CNS) neurotransmitter systems have been reported in thiamine deficiency, suggesting that thiamine is important for normal neurotransmitter function. Acetylcholine, γ-aminobutyric acid (GABA), glutamate, and aspartate are produced primarily through the oxidative metabolism of glucose. Serotonin metabolism is also influenced by thiamine deficiency. CSF levels of 5-hydroxyindoleacetic acid (5-HIAA) were reduced in patients with Wernicke-Korsakoff syndrome. Recent work suggests that thiamine may inhibit serotonin uptake, providing further evidence that thiamine may be acting as a neurotransmitter/neuromodulator as well as a cofactor in brain metabolism, and that its deficiency may be associated with behavioral and mood changes. The anatomical lesions described in association with thiamine deficiency include the medial dorsal thalamic nuclei, the mammillary bodies, the fornix, and the cerebellum. The anterior thalamic nucleus has also been associated with thiamine deficiency; it receives input from the mammillary complex and has major reciprocal connection with the subiculum-presubiculum portion of the hippocampal complex and with the cingulate gyrus. Lesions to these regions (anterior thalamic nuclei, mammillary complex) are related to impaired explicit memory.

Clinically, subjects with Korsakoff syndrome, while maintaining an overall intact intelligence, have a striking inability to form new memories (anterograde amnesia). The retrograde amnesia is less severe, although present to a certain degree (usually for events that occurred shortly before the onset of the illness). Immediate recall and implicit learning (e.g., learning procedures) are preserved. Individuals with Korsakoff syndrome classically present with confabulation, which is explained by most experts as attempts to cover the memory deficit.

Assessment

Neuropsychological testing usually shows a discrepancy between the performance on memory scales and the performance on general cognitive scale (a discrepancy between memory performance and overall intelligence).

Laboratory tests show evidence of chronic alcohol use, malnutrition, vitamin deficiency, or systemic illness. Neuroimaging tests are helpful to differentiate between several causes of amnestic syndrome (mammillary bodies atrophy in Korsakoff syndrome, strategic localized strokes, anoxic lesions of the hippocampal structures, T2-weighted hyperintensities in bilateral temporal poles in herpes encephalitis).

Differential

In contrast to delirium, amnestic disorders evolve on a background of clear consciousness and intact attention, with no perceptual disturbances or mood lability. Amnestic syndrome has no impairment in abstract thinking, judgment, or executive functions. Cognitive impairment in dementia extends to semantic memory, language, and praxis.

Psychogenic amnesia can be differentiated by its presentation (onset after traumatic event, no evidence of substance/general medical condition, amnesia for personal identity, which is conserved in amnestic disorder, and for a circumscribed event, but preserved memory for new events and preserved ability for learning) and its course (abrupt onset and resolution, no residual impairment).

Course and Prognosis

About 25 percent of patients with Korsakoff syndrome recover; the rest continue to suffer from various degrees of memory impairment, a significant minority needing permanent placement. Most amnestic syndromes have a stable course. A few would recover (e.g., transient global amnesia or postelectroconvulsive therapy amnesia) or improve (e.g., amnestic disorder secondary to traumatic brain injury).

Treatment

Treatment with high doses of thiamine (1 to 2 g daily) is likely to partially or completely reverse Wernicke encephalopathy, but its benefits in the chronic stages (Korsakoff syndrome) are not established. There are case reports but no controlled studies regarding the benefits of donepezil (Aricept), rivastigmine (Exelon), and memantine (Namenda) in treating Korsakoff syndrome. For the other types of amnestic syndrome, the treatment is tailored accordingly to the etiology (acyclovir for herpes encephalitis, hyperbaric oxygen for carbon monoxide poisoning). However, once the amnesic deficit is installed, there are very few pharmacological options. Most amnestic syndromes have a stable and some even a reversible course (global transient amnesia). Cognitive rehabilitation programs can help improve the patient's level of functioning, but so far none have indicated benefits regarding restoring/improving memory in the impaired domain.

REFERENCES

Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):270–279.

Birks J, Flicker L. Donepezil for mild cognitive impairment. *Cochrane Database Syst Rev*. 2006;3:CD006104.

Crook T, Bahar H, Sudilovsky A. Age-associated memory impairment: diagnostic criteria and treatment strategies. *Int J Neurol*. 1987;21–22:73–82.

Ganguli M. Mild cognitive impairment and the 7 uses of epidemiology. *Alzheimer Dis Assoc Disord*. 2006;20:S52–S57.

Ganguli M, Dodge HH, Shen C, DeKosky ST. Mild cognitive impairment, amnestic type: an epidemiologic study. *Neurology*. 2004;63:115–121.

*Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K. Mild cognitive impairment. *Lancet*. 2006;367:1262–1270.

Haas RH. Thiamin and the brain. *Annu Rev Nutr*. 1988;8:483–515.

Harding A, Halliday G, Caine D, Kril J. Degeneration of anterior thalamic nuclei differentiates alcoholics with amnesia. *Brain*. 2000;123(Pt 1):141–151.

Harper C, Fornes P, Duyckaerts C, Lecomte D, Hauw JJ. An international perspective on the prevalence of the Wernicke-Korsakoff syndrome. *Metab Brain Dis*. 1995;10:17–24.

Heister D, Brewer JB, Magda S, Blennow K, McEvoy LK. Alzheimer's Disease Neuroimaging Initiative. Predicting MCI outcome with clinically available MRI and CSF biomarkers. *Neurology*. 2011;77(17):1619–1628.

Jack CR Jr, Knopman DS, Jagust WJ, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol*. 2013;12(2):207–216.

Jack CR Jr, Wiste HJ, Weigand SD, et al. Amyloid-first and neurodegeneration-first profiles characterize incident amyloid PET positivity. *Neurology*. 2013;81(20):1732–1740.

Jelic V, Kivipelto M, Winblad B. Clinical trials in mild cognitive impairment: lessons for the future. *J Neurol Neurosurg Psychiatry*. 2006;77:429–438.

Katzman R, Terry R, DeTeresa R, Brown T, Davies P. Clinical, pathological, and neurochemical changes in dementia: a subgroup with preserved mental status and numerous neocortical plaques. *Ann Neurol*. 1988;23:138–144.

Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann Neurol*. 2004;55:306–319.

Larrieu S, Letenneur L, Orgogozo JM, Fabrigoule C, Amieva H. Incidence and outcome of mild cognitive impairment in a population-based prospective cohort. *Neurology*. 2002;59:1594–1599.

Lopez OL, Jagust WJ, Dulberg C, Becker JT, DeKosky ST. Risk factors for mild cognitive impairment in the Cardiovascular Health Study Cognition Study: Part 2. *Arch Neurol*. 2003;60:1394–1399.

Loy CT, Schofield PR, Turner AM, Kwok JB. Genetics of dementia. *Lancet*. 2014;383(9919):828–

840.

Petersen RC, ed. *Mild Cognitive Impairment: Aging to Alzheimer Disease*. New York: Oxford University Press; 2003.

*Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med*. 2004;256:183–194.

*Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC. Current concepts in mild cognitive impairment. *Arch Neurol*. 2001;58:1985–1992.

Petersen RC, Aisen P, Boeve BF, et al. Mild cognitive impairment due to Alzheimer disease in the community. *Ann Neurol*. 2013;74(2):199–208.

Rademakers R, Baker M, Gass J, et al. Phenotypic variability associated with progranulin haploinsufficiency in patients with the common 1477C→T (Arg493X) mutation: an international initiative. *Lancet Neurol*. 2007;6:857–868.

Smith G, Ivnik RJ, Petersen RC, Malec JF, Kokmen E. Age-associated memory impairment diagnoses: problems of reliability and concerns for terminology. *Psychol Aging*. 1991;6:551–558.

Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):280–292.

Taber KH, Wen C, Khan A, Hurley RA. The limbic thalamus. *J Neuropsychiatry Clin Neurosci*. 2004;16:127–132.

Thompson PM, Hayashi KM, Dutton RA, et al. Tracking Alzheimer's disease. *Ann N Y Acad Sci*. 2007;1097:183–214.

van Duijn CM. Prospects of genetic research of mild cognitive impairment. *J Intern Med*. 2004;256:235–239.

Wenisch E, Cantegreil-Kallen I, De Rotrou J, et al. Cognitive stimulation intervention for elders with mild cognitive impairment compared with normal aged subjects: preliminary results. *Aging Clin Exp Res*. 2007;19:316–322.

*Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment—beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. *J Intern Med*. 2004;256:240–246.

▲ 10.5 Other Cognitive and Mental Disorders due to Another Medical Condition

JENNIFER CLARE PRINS, M.D., AND ELLEN M. WHYTE, M.D.

INTRODUCTION

Historically, mental disorders were divided into those with a clearly identifiable organic cause (e.g., personality change due to a brain tumor) and those without an identifiable pathophysiological basis (e.g., personality disorder). The former illnesses, considered to be based in abnormal biology, were labeled "organic" in nature, whereas the latter, considered to be a consequence of deranged upbringing or environment, were labeled "functional."

However, it is now understood that all mental disorders reflect abnormal brain function. Hence, all mental disorders can be considered organic. Yet, it is also increasingly recognized that the environment influences the expression of underlying biology; hence, all mental disorders can be considered functional as well. Furthermore, treatments of mental disorders, whether psychological or biological, change brain function.

Viewed in this manner, the traditional separation of mental illness from physical illness appears increasingly illogical and unscientific. Yet, the separation of psychiatric illness into organic or secondary (i.e., due to an identifiable medical or neurological anomaly or psychoactive substance) and functional or primary (i.e., due to unidentified factors) has been a necessary capitulation to the limited understanding of the roots of mental illness.

In the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) and its text revision (DSM-IV-TR), the nosology shifted from the term *organic mental illness* to the term *due to another medical condition* to denote a secondary mental disorder. In DSM-5, the nosology shifted further to the term *due to another medical condition*. This shift in terminology underscores psychiatry's evolution in thinking about etiology. Furthermore, this shift is transitional, as the labels given to mental disorders will change as the understanding of the etiology of mental disorders improves. Hence, while the current DSM nosology likely represents an improvement over the previous classification, it too will evolve as the understanding of the underlying pathophysiology of mental illness evolves.

Definition

DSM-5 uses the term *due to another medical condition* to define a mental disturbance that "is caused by physiological effects of another medical

condition." DSM-5 coding requires listing the medical condition causing the mental disorder first (e.g., hypothyroidism), followed by listing the diagnosis of the mental disorder (e.g., depressive disorder due to hypothyroidism).

History

Physicians and scientists interested in mental illness have struggled with issues of etiology and classification. Nineteenth-century physicians had few tools for examining the cellular and chemical basis of mental illness and typically worked in asylums separate from other physicians. Psychiatric units began to appear in general hospitals in growing numbers in the 1920s. This, along with the development of a variety of somatic interventions, including electroconvulsive therapy (ECT) and psychosurgery, led to increased focus on the biological basis of mental disease. In the 1930s and 1940s, individuals including George Henry and Morris Kaufman lent impetus to an evolving interest in the interface of medical and mental disorders. Psychosomatic medicine and its practitioners, consultation-liaison psychiatrists, evolved to assist nonpsychiatrists in the evaluation and management of individuals in whom psychological factors were perceived to be relevant to the development or maintenance of a host of physical disorders. These psychiatrists also were enlisted to treat the psychiatric manifestations of a variety of medical conditions. In the 1960s and 1970s, liaison activity, involving formal and informal didactic training of nonpsychiatrists by their psychiatric colleagues, was aimed at increasing awareness of the interplay between mental and medical conditions.

The middle of the 20th century witnessed competing trends in psychiatry. Psychological theories and therapies, including psychoanalysis, grew to explain and manage even the most serious psychotic illnesses, including schizophrenia. Concurrently, psychopharmacological developments, especially the introduction of antipsychotic agents (e.g., chlorpromazine [Thorazine]), served as harbingers of sustained efforts to understand and manage the biochemical basis of psychopathology. By the close of the 20th century, psychiatry shifted from the psychological and psychodynamic theories to neuropsychiatric approaches, which have been viewed as embracing the most rigorously scientific methodologies, like those used in other areas of medicine. The neuropsychiatric era has coincided with renewed emphasis on incorporation of psychiatric care into the medical mainstream, with the general hospital and general medical outpatient setting usually serving as favored settings for delivery of such care.

Comparative Nosology

The evolution of psychiatric nosology paralleled these historical developments. The first edition of the DSM (DSM-I), introduced in 1952, divided mental illness into two categories: those without clear physical or structural cause and those with identifiable acute or chronic brain impairment. The second edition of the DSM (DSM-II), introduced in 1968, maintained this dichotomy, with the latter group of disorders believed to be associated with cognitive disturbance, including memory and judgment impairment. Introduced in 1980, DSM-III defined organic disorders as those "associated with transient or permanent dysfunction of the brain." DSM-III and DSM-III-R recognized that organic disorders could include the role of substances, both medications and drugs of abuse. Compared to their predecessors, DSM-III and DSM-III-R expanded the array of subtypes of organic illness, adding categories for hallucinosis, delusions, mood symptoms, anxiety symptoms, and personality disturbances, in addition to cognitive disturbances such as dementia and delirium.

DSM-IV and DSM-IV-TR abandoned the term organic to avoid the misconception that nonorganic, functional disorders were perhaps not due to biological or chemical brain aberrations. Instead, functional disorders were described as primary mental disorders, whereas organic disorders were described as secondary mental disorders, which were further divided into those induced by identifiable medical conditions (*due to another medical condition*) and those induced by identifiable substances, prescribed or abused (*substance-induced*).

Thus, DSM-IV-TR created a diagnostic distinction between mental disorders due to general medical conditions and those due to substances, a departure from the more unified approach adopted by DSM-III, which generally allowed for a single organic diagnosis defined by symptom constellation, whether due to a medical condition or substance. The dichotomy introduced by DSM-IV represented an effort at diagnostic refinement, although the distinction between substance- and medical-induced mental disorders was not meant to exclude the likelihood that evolving knowledge eventually will point to common biochemical pathways explaining mental disorders with similar phenomenological presentations.

In DSM-5, the term *due to another medical condition* has been renamed *due to another medical condition* and includes a broad array of psychiatric disorders attributable to the physiological manifestations of specific medical conditions (Table 10.5–1). In contrast to psychiatric states resulting from the emotional distress or psychological reaction to a medical condition, this category of illness includes only those disorders in which a medical condition, through physiological means, has directly impacted the functioning of the brain as manifested by the appearance of psychiatric symptoms. These same considerations apply to substance-induced mental disorders (Table 10.5–1).

Table 10.5–1.
List of DSM-5 Diagnoses due to Another Medical Condition or Substance/Medication-Induced

Diagnostic Category	Due to Another Medical Condition	Substance/Medication-Induced
Neurocognitive disorder	X	X
Psychotic disorder	X	X
Bipolar and related disorder	X	X
Depressive disorder	X	X
Anxiety disorder	X	X
Obsessive-compulsive and related disorder	X	X
Sleep disorder		X
Sexual dysfunction		X
Personality change	X	

In contrast, the tenth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) has not abandoned the *organic* nomenclature and continues a classification scheme similar to that of DSM-III and DSM-III-R. ICD-10 describes organic disorders as having a "demonstrable etiology in cerebral disease, brain injury, or other insult leading to cerebral dysfunction. The dysfunction may be primary, as in diseases, injuries and insults that affect the brain directly and selectively; or secondary, as in systemic diseases and disorders that attack the brain only as one of the multiple organs or systems of the body that are involved." Diagnoses fall within one of two categories: (1) organic mental disorders, which include nonpsychoactive drug effects and (2) mental and behavioral disorders due to psychoactive substance use.

Etiology

A vast array of medical conditions and substances can cause cognitive and mental disorders. In an attempt to organize these conditions for ease of diagnosis and study, a common strategy is to group these conditions based on their primary mechanism of action, such as neurological, substance-induced, and other medical causes (Table 10.5–2).

Diagnosis and Clinical Features

Like any medical or psychiatric illness, the diagnosis of a cognitive or mental disorder due to another medical condition is dependent on comprehensive history taking, detailed physical examination, and appropriate laboratory and imaging studies. For the history, specific details of the mental symptoms, including presenting features, onset and course, and associated symptoms, must be obtained. More than any other part of

the assessment, the description of the presenting features along with the timing of onset is the key to correct diagnosis. Past medical, psychiatric history, and family history must also be explored, as this can increase the pretest probability of a specific diagnosis and help dictate diagnostic testing. A comprehensive review of all prescription and nonprescription medications, vitamins, dietary and herbal supplements, drugs of abuse, and environmental toxin exposures must be undertaken.

Physical examination should include, but is not limited to, cognitive assessment, vital signs, and a comprehensive neurological examination. Several frequently used tests of general cognition and tests of specific cognitive domains are outlined in Table 10.5–3. Neuropsychological testing refers to the administration of a comprehensive battery of multiple tasks evaluating a range of cognitive domains. Examination of specific organ systems is indicated by history and presenting symptoms. For example, in patients with a suspected vascular dementia, detailed examination of the heart and arterial system is warranted. Specific findings on the physical examination may narrow the list of potential causative medical conditions (Table 10.5–4). An array of laboratory, brain imaging, and additional diagnostic studies may be helpful in diagnosing the causative condition (Table 10.5–5). However, performing all possible tests on all patients would be inappropriate and a waste of medical resources. The choice of diagnostic studies should be guided by history and presenting symptoms, as well as physical examination findings.

Table 10.5–2.
Medical Conditions and Substances/Medications Causing Neurocognitive Disorders and Other Mental Disorders

Neurological	Substance Induced	Other
Degenerative	**Medications**	**Metabolic**
Alzheimer disease	Anticholinergic	Anemia
Frontal–temporal dementia (Pick disease)	Anticonvulsants	Hypo- and hypercalcemia
Lewy body disease	Antihypertensive	Wilson disease/hepatolenticular degeneration
Parkinson disease	Anxiolytic	End-stage organ dysfunction
Huntington disease	Corticosteroids	**Endocrine**
Corticobasal degeneration	Immunosuppressant	Adrenal
Progressive supranuclear palsy	Opioid	Pancreases
Prion diseases	Sedative-hypnotic	Thyroid
Demyelinating disease	Sympathomimetic	**Nutritional**
Multiple sclerosis	**Toxins**	Folate deficiency
Acquired brain injury	Lead	Niacin deficiency
Cerebrovascular disease	Mercury	Vitamin B_1 deficiency
Traumatic brain injury	Carbon monoxide	Vitamin B_{12} deficiency
Chronic traumatic encephalopathy (e.g., dementia pugilistica)	**Other**	**Infectious/Inflammatory**
	Alcohol	Human immunodeficiency virus

		(HIV)
Space-occupying lesion	Amphetamine/other stimulants	Slow viruses (e.g., Creutzfeldt–Jakob disease)
Tumor	Cannabis	Neurosyphilis
Subdural hematoma	Cocaine	Lyme disease
Normal pressure hydrocephalus	Hallucinogens	Systemic lupus erythematosus
Abscess	Inhalants	Giant cell arteritis
	Phencyclidine	

In order to diagnose a neurocognitive or mental disorder due to another medical condition, DSM-IV-TR delineated three general criteria that must be met:

1. There is evidence from the history, physical examination, or laboratory findings that the disturbance is the direct physiological consequence of another medical condition.
2. The disturbance is not better accounted for by another mental disorder.
3. The disturbance does not occur exclusively during the course of a delirium.

Table 10.5–3.
Commonly Used Cognitive Assessment Tools

Cognitive Domain	Examples of Specific Cognitive Tests
Global cognition	Montreal Cognitive Assessment, Mini-Cog, Dementia Rating Scale
Learning and memory	California Verbal Learning Test, Rey–Osterrieth Complex Figure Test
Visuospatial/construction	Clock Drawing Test, Block Design Test
Language	Boston Naming Test, Controlled Word Association (Fluency)
Executive function	Trail Making Test B, Stroop Test
Attention	Digit Span-Forward, Continuous Performance Tests
Psychomotor speed	Digit Symbol Substitution Test, Trail Making Test A

Evidence of Physiological Cause. See Table 10.5–6 for a summary of the general principles for diagnosing disorders due to another medical condition or substance/medication-induced disorders.

Evidence that would support a physiological link between another medical condition or substance and a cognitive or mental disorder:

A. Presence of the general medical condition or use of a substance, based on history, laboratory studies, imaging studies, or physical examination.
B. Temporal association between the onset, exacerbation, or remission of the general medical condition or use of a substance and the cognitive or mental disorder. A temporal link can be difficult to determine, especially when the symptoms of the mental disorder occur prior to or distant to the onset of a possible causative medical condition. Specific examples include depression preceding the onset of Parkinson disease and dementia presenting after previous or prolonged exposure to an

environmental toxin. Ideally, treatment of the underlying medical condition will result in improvement or resolution of the cognitive or mental disorder, but unfortunately this is not always the case (e.g., dementia that persists after B12 deficiency is corrected) and not all medical conditions are treatable.

C. The presentation of unusual features (compared to what would be expected in the presentation of the primary mental disorder) suggests a nonprimary mental disorder. DSM-IV-TR cites specific examples including atypical age of onset (e.g., first appearance of Alzheimer-like symptoms in a 25-year-old individual), unusual associated features (e.g., visual or tactile hallucinations accompanying a major depressive-like episode), diagnostic features that are more severe than would be expected based on the overall presentation, lack of family history of mental disorders, and cognitive loss out of proportion to that typically encountered in the primary mental disorder.

D. Evidence from the scientific literature of an established or frequent link between the medical condition or substance and the cognitive or mental disorder (e.g., urinary tract infection and delirium, alcohol abuse and dementia). Although this evidence is useful, it is not sufficient in itself for making a determination of causation. Also credible scientific evidence is frequently lacking regarding these associations.

Table 10.5–4.
Specific Neurological Signs and Symptoms and Associated Medical Conditions

Neurological Sign or Symptom	Possible Related Medical Conditions
Change in consciousness	Infection, hepatic encephalopathy, substance induced
Headache	Meningitis, tumor, giant cell arteritis
Tremor	Parkinson disease, alcohol abuse, human immunodeficiency virus-associated dementia, Wilson disease
Gait disturbance	Normal pressure hydrocephalus, stroke, Parkinson disease
Meningeal signs	Meningitis
Involuntary movements	Huntington disease, Parkinson disease
Pathological reflexes	Creutzfeldt–Jakob disease, hypothyroidism
Motor defect	Amyotrophic lateral sclerosis, Huntington disease
Sensory defect	Stroke, B_{12} deficiency
Gaze palsy	Thiamine deficiency, alcohol abuse
Cranial nerve defect	Stroke
Language disturbance	Stroke, Alzheimer disease, neurosyphilis

Table 10.5–5.

Common Diagnostic Tests for Cognitive and Mental Disorders due to a General Medical Condition

Laboratory studies
Complete blood count[a]
Electrolytes[a]
Fasting glucose[a]
Blood urea nitrogen and creatinine[a]
Calcium, magnesium, and phosphorus[a]
Liver function tests[a]
Thyroid function tests[a]
Rapid plasma reagin/Venereal Disease Research Laboratory
Vitamin B_{12} and folate[a]
Erythrocyte sedimentation rate
Arterial blood gases
Toxicology
Urinalysis with or without culture/sensitivity
Human immunodeficiency virus
Ceruloplasmin
Catecholamines
Porphobilinogens
Specific drug level (e.g., medications and drugs of abuse)
Specific toxins (e.g., lead, mercury)
Brain imaging
Structural brain imaging: computed tomography or magnetic resonance imaging[a]
Functional magnetic resonance imaging
Positron emission tomography
Single photon emission computed tomography
Additional tests
Electroencephalography
Cerebrospinal fluid for biomarkers, infectious agents
Chest radiograph
Electrocardiogram

[a]Indicates tests performed routinely during evaluation of mental disorders; all other tests are performed if suggested by history or physical examination.

Table 10.5–6. General Principles for Diagnosing Disorders due to Another Medical Condition or Substance/Medication-Induced Disorders

Disorders due to Another Medication Condition:

- The symptoms of the psychiatric disturbance may be indistinguishable from that of a primary psychiatric disorder.
- Symptoms cannot be better accounted for by a primary psychiatric disorder or reflect an adjustment disorder due to the stress of having a medical illness.
- The disturbance does not occur exclusively during the course of a delirium.
- A temporal link between the course of the psychiatric disturbance and the onset, exacerbation, or remission of the medical disorder supports a causal link.
- Evidence from the medical literature supports a link between the psychiatric disturbance and the

medical illness.

- ► Evidence of atypical features (e.g., late onset, lack of prior episodes, and lack of family history) support a link between the psychiatric disturbance and the medical condition.
- ► The goal of a targeted workup is to "rule in" a specific medical diagnosis suggested by the patient's somatic symptoms.

Substance/Medication-Induced Disorders:

- ► The symptoms of the psychiatric disturbance may be indistinguishable from that of a primary psychiatric disorder.
- ► Symptoms cannot be better accounted for by a primary psychiatric disorder or psychiatric disorder due to another medical condition.
- ► The disturbance does not occur exclusively during the course of a delirium.
- ► The symptoms of the psychiatric disturbance arise in the context of initiating or changing the dose of a medication or during intoxication with or withdrawal from a substance, are in excess of those usually associated with the medication's use or the substance's intoxication or withdrawal state, and warrant independent clinical attention.
- ► A link between the psychiatric disturbance and the substance/medication is supported by the absence of psychiatric symptoms prior to use of the substance/medication, resolution of symptoms within 4 weeks of last substance/medication use, lack of prior similar episodes that occurred independent of substance/medication and the presence of psychiatric symptoms consistent with the extent of substance/medication use.
- ► Evidence from the medical literature supports a link between the psychiatric disturbance and the substance/medication.

Disturbance Is Not Better Accounted for by Another Mental Disorder. Of obvious importance is the exclusion of a primary mental disorder when diagnosing a medical or substance-induced cognitive or mental disorder. However, this can be complicated by several factors:

1. Coexisting but physiologically unrelated physical and mental disorders. For example, an individual with recurrent major depressive disorder may develop Parkinson disease. Whether the current (index) episode of depression is the result of Parkinson disease can be difficult to determine especially when the potentially causative medical condition cannot be cured. Remission of symptoms of the cognitive or mental disorder with treatment of a medical illness would support a relationship between the two conditions.
2. The influence on mental state by medical conditions or substances through a nonphysiological mechanism. For example, a new diagnosis of Huntington disease may act as a psychosocial stressor that results in depression. In this example, the resulting mental disorder (depression) is not diagnosed as being *due to* the medial condition (Huntington disease). Patients with mental disorders exacerbated by medical illness generally should receive diagnosis of a primary mental disorder (e.g., adjustment disorder with depressed mood).
3. Coincidental occurrence of both a medical condition and a mental disorder, which can further confound diagnosis.

The diagnosis of substance-induced cognitive or mental disorder should

be considered when recent or prolonged substance intoxication or withdrawal has occurred. It is important to point out that mild neuropsychiatric symptoms can routinely occur with substance use. However, the diagnosis of a substance-induced cognitive or mental disorder implies that the symptoms are greater than those typically associated with the intoxication or withdrawal state. For example, mild depression in the context of cocaine withdrawal would not be coded as a substance-induced mental disorder as mild depression commonly occurs with cocaine withdrawal. However, when the symptoms are in excess of what is typically expected, such as depressed mood associated with psychosis or active suicidal ideation, a diagnosis of cocaine-induced depressive disorder with onset during withdrawal should be coded.

Specific diagnostics tests such as electrolytes and urine toxicology studies can aid in the differentiation between medically induced and substance-induced cognitive and mental disorders. In some instances, a medical condition and a substance can both be causal agents in the disorder. For example, congestive heart failure can cause delirium by means of cerebral hypoperfusion. The anticholinergic effects of digoxin (Lanoxin) used to treat the congestive heart failure may also cause delirium. In this case, a diagnosis of a neurocognitive disorder (NCD) due to multiple etiologies would be the correct diagnosis. The diagnosis of an unspecified NCD is used when the primary condition cannot be determined.

Disturbance Does Not Occur Exclusively during the Course of a Delirium. Delirium is a common diagnosis in the acute care setting, with hospital admission prevalence of 14 to 24 percent. Over 50 percent of postsurgical patients experience delirium. Delirium is more common in older patients especially those with underlying cognitive impairment. However, delirium is under-recognized and under-diagnosed. The fluctuating nature of delirium and the common presentation of hypoactive delirium contribute to this lack of recognition. Since delirium is more common in patients with dementia, the cognitive and behavioral changes associated with delirium can mistakenly be attributed to the dementia. Awareness of the prevalence of delirium associated with acute medical illness or substance/medication intoxication or withdrawal should promote diligence on the part of clinicians in recognizing this disorder. It is essential to rule out delirium prior to making any other NCD.

Recording Procedures

In DSM-5, for a *neurocognitive or mental disorder due to another medical condition*, the presumed causative medical condition is coded first, followed by neurocognitive or mental disorder. For *substance/medication-induced mental disorders*, the substance should be named followed by the

name of the mental disorder. Generally, the diagnosis is usually clarified to either be *with onset during intoxication* or *with onset during withdrawal*.

Course and Prognosis

The course and prognosis of neurocognitive and mental disorders due to another medical conditions and substance/medication are dependent on the specific medical condition and its response to the treatment or the specific substance/medication involved. The course can range from remission of all symptoms to progressive decline, ultimately leading to death. When the neurocognitive or mental disorder has been of brief duration and the medical condition is easily identified and responds to treatment, the neurocognitive or mental disorder may improve as well. When the underlying causative condition is long-standing and/or progressive or the causal substance cannot be discontinued or its effects are permanent (e.g., previous environmental toxin exposure), a positive outcome is less likely even when the causative medical condition can be treated (e.g., neurosyphilis).

Treatment

The mainstay of treatment is the correction or amelioration of the underlying medical condition or discontinuation of the offending substance or exposure. Adjuvant treatments include psychopharmacological interventions and psychotherapeutic interventions.

Psychopharmacological interventions are typically used to treat associated mood or behavioral symptoms especially when correction of the causal medical condition or withdraw of the causal substance does not result in remission of these symptoms. Given the medically compromised nature of the patient, the clinician should be mindful of the possibility for altered pharmacokinetics (e.g., decreased clearance of drugs) as well as potential drug–drug interactions in addition to the frequent lack of empiric evidence supporting the use of psychopharmacological interventions in these secondary neurocognitive and mental disorders.

Psychotherapeutic interventions in neurocognitive and mental disorders due to medical conditions include behavior modification and cognitive retraining. Other specific interventions promote life coping strategies, educate patients about illness course, and bolster supportive networks with family and friends. For mental disorders due to substance use, detoxification and a comprehensive rehabilitation program are frequently necessary but relapse rates are high.

COGNITIVE DISORDERS DUE TO ANOTHER MEDICAL CONDITION

NCDs (major/mild) and delirium are the recognized cognitive disorders due to general medical conditions.

Major NCD, or dementia, is a complex syndrome resulting from the accumulation of cognitive impairments to a degree that functional changes occur. The prevalence of NCDs increases with age. NCDs due to degenerative diseases, such as Alzheimer disease and Parkinson disease, are discussed separately (see Section 10.3) (see Section 10.2 for a discussion of delirium). The focus of this section will be on cognitive impairment secondary to other medical disorders, such as stroke, traumatic brain injury (TBI), and HIV disease.

Definition

Major NCDs involve impairment in at least one cognitive domain. Impairment in learning and memory is a key feature of many, but not all, NCDs. Other cognitive domains that can be impaired include language, complex attention (including cognitive processing speed), executive functions, perceptual-motor skills, and social cognition. These impairments must be new and distinct from the person's previous level of functioning. Furthermore, these changes in cognitive functioning must be of sufficient degree to affect daily functioning. Timing of onset and nature of initial cognitive changes as well as time course gives information as to the etiology of the dementia. Behavioral and personality changes may also occur.

Comparative Nosology

In DSM-III and DSM-III-R, dementia was included in the classification organic mental syndromes and disorders. However, the terms organic and nonorganic were dropped from DSM-IV and DSM-IV-TR as these terms implied that nonorganic disorders did not have a biological basis. In DSM-5, the term dementia was replaced by term *major NCD*. Although the criteria for dementia required functional impairment due to impairment in memory plus one additional cognitive domain, the criteria for major NCD are broader in that impairment in only one cognitive domain sufficient enough to result in functional impairment is required for diagnosis. Furthermore, the term is often preferred to dementia when dealing with younger patients.

In addition to the neurodegenerative diseases, DSM-5 provides specific criteria for the diagnosis of NCDs due to the following medical conditions: vascular, TBI, HIV infection, due to another medical condition, and substance/medication induced. Specifiers are used to indicate the presence of a behavioral disturbance, severity of impairment, and the level of confidence in the presumed etiology (possible or probable).

ICD-10 continues to use the word dementia to describe irreversible cognitive impairment causing functional impairment. The options for noting the underlying medical disorder causing the cognitive impairment are broader than in DSM-5 and include dementia (F03.90) related to

vascular causes (F01.50), substance induced (including alcohol [F10.97]), and psychoactive drugs (F19.97), inhalant (F18.27), sedative/hypnotic/anxiolytic (F13.97), infectious (including neurosyphilis [A52.17]), HIV (B20), endocrine disorders (hypothyroidism, acquired [E03.9]), metabolic disturbances (hypercalcemia (E83.52), Wilson disease/hepatolenticular degeneration (E83.01), nutritional deficiencies (including deficiencies in niacin [E52]), vitamin B_{12} deficiency (E53.8) and autoimmune disorders (systemic lupus erythematosus [M32]), polyarteritis nodosa (M30.0), (idiopathic) normal pressure hydrocephalus (G91.2), and multiple etiologies (F03). In addition, many of these diagnoses include a specifier to indicate the presence of behavioral symptoms.

Epidemiology

Alzheimer disease and cerebrovascular disease account for the vast majority of NCD cases in the developed world. In America and Europe, Alzheimer disease is considered the leading cause of NCD, whereas in Japan cerebrovascular disease accounts for 50 percent of NCD cases. The prevalence of other acquired NCDs is discussed in Section 10.3.

Etiology

A common pathway leading to NCD is disrupted neurotransmission. This can be due to cell death, synapse loss, or deficits in neurotransmitter function. In particular, a cholinergic deficit has been implicated in many NCDs, and the importance of acetylcholine in memory formation has been established.

Many mechanisms that could contribute to cell death and disrupted neurotransmission have been postulated. These include inflammation, cerebral hypoperfusion and hypoxia, metabolic derangements, and specific infectious agents. Table 10.5–2 lists specific medical conditions that may cause an NCD.

Table 10.5–7.
DSM-5 Criteria for Major Neurocognitive Disorder

A. Evidence of significant cognitive decline from a previous level of performance in one or more cognitive domains (complex attention, executive function, learning and memory, language, perceptual-motor, or social cognition) based on:
   1. Concern of the individual, a knowledgeable informant, or the clinician that there has been a significant decline in cognitive function; and
   2. A substantial impairment in cognitive performance, preferably documented by standardized neuropsychological testing or, in absence, another quantified clinical assessment.
B. The cognitive deficits interfere with independence in everyday activities (i.e., at a minimum, requiring assistance with complex instrumental activities of daily living such as paying bills or managing medications).
C. Cognitive deficits do not occur exclusively in the context of a delirium.

D. The cognitive deficits are not better explained by another mental disorder (e.g., major depressive disorder, schizophrenia).

Specify:
Without behavioral disturbance
With behavioral disturbance
Specify current severity: mild, moderate, severe

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Diagnosis and Clinical Features

Regardless of etiology, cognitive impairment in at least one of the following domains—complex attention, executive functions, learning and memory, language, perceptual-motor skills, or social cognition—with associated impairment in day-to-day function, is the hallmark of NCDs (Table 10.5–7). Changes in personality and behavior are common. Associated neuropsychiatric symptoms can include apathy, agitation, anxiety, depression, overt psychosis, and sleep disturbance.

Although somewhat controversial, NCDs can be divided into cortical and subcortical NCD based on initial location of pathological brain changes and/or presenting symptoms. Cortical NCDs include Alzheimer disease and frontal–temporal dementia. Subcortical NCDs include those due to Parkinson disease dementia, acquired immunodeficiency syndrome (AIDS) dementia, and subcortical vascular lesions. Subcortical vascular NCD is associated with pathological changes involving the thalamus, basal ganglia, and frontal–striatal white matter tracts with relative initial sparing of the cerebral cortex. Subcortical vascular NCD are characterized by slowness of mental processing, apathy, personality changes, and depression. With regard to memory deficits, cortical NCDs are initially associated with an encoding deficit, while subcortical NCDs are typically associated with a retrieval deficit. Unlike cortical NCDs, impairment of language is much less common in the early stages of subcortical NCDs.

Laboratory Examination

The workup of an NCD should include general blood work to determine any potential causal medical condition. This blood work includes a complete blood count, electrolytes, blood urea nitrogen, creatinine, fasting glucose, liver function tests, thyroid function tests, and calcium and vitamin B_{12} levels. If neurosyphilis is suspected, a screening reactive plasma regain (RPR) can be performed with confirmatory VDRL test if the RPR is positive.

Given the extent of medical conditions that can cause cognitive impairment and dementia, other testing should be tailored to the history and the presenting symptoms. For example, an erythrocyte sedimentation

rate (ESR) can be performed if an inflammatory disorder is suspected, although ESR is a very nonspecific test. Copper and ceruloplasmin levels can be measured if Wilson disease is suspected. A detailed history of the proband as well as the family is important to obtain to guide this workup. Structural brain imaging should be performed and in specific cases, functional brain imaging may help differentiate neurocognitive syndromes. Lumbar puncture is not performed routinely, but advances in the understanding of specific cerebrospinal fluid (CSF) markers may help in the differential diagnosis, especially if CJD, neurosyphilis or other infectious etiology is suspected. Toxicology tests may be appropriate in selected cases.

Neuropsychological testing can aid in the diagnosis, especially in differentiating cortical from subcortical NCDs and also to evaluate the pattern and severity of cognitive impairments.

Differential Diagnosis

Medical conditions and substances that can cause NCDs often occur in patients with Alzheimer disease or cerebrovascular disease. Multiple medical causes may also coexist. Therefore, differentiating the factors actually causing the cognitive impairment may be difficult. In addition, depression and delirium are also part of the differential diagnosis and can occur concomitantly with the NCD. Further complicating the diagnosis is the overlap of clinical symptoms between these conditions. For example, delirium may be subclinical and chronic especially if the underlying cause of the delirium is not addressed. Both depression and NCD can be associated with apathy, weight loss, and social withdrawal. Detailed history taking and physical examination as well as appropriate testing are paramount in identifying the causal condition of the NCD.

Course and Prognosis

NCD due to some acquired etiologies (e.g., TBI, stroke, vitamin B_{12} deficiency) may be characterized by cognitive and functional deficits that are stable over time if the underlying etiology is controlled or treated. However, progressive decline is seen if the underlying disease is not successfully controlled (e.g., recurrent strokes).

Treatment

Obviously, correction or optimization of the underlying medical condition or discontinuation of the responsible substance is the primary treatment of NCDs that are secondary to a medical condition or are substance induced. For example, treatment of B_{12} deficiency may halt or reverse some cognitive and neuropsychiatric symptoms. Ventriculoperitoneal shunting in normal pressure hydrocephalus may benefit motor functioning, but its

effect on reversing or delaying cognitive impairments is less certain.

The use of cholinesterase inhibitors such as donepezil [Aricept], rivastigmine [Exelon], and galantamine [Razadyne] as well as the *N*-methyl-D-aspartate [NMDA]-receptor antagonist memantine [Namenda] for NCDs that are secondary to a medical condition or are substance induced is controversial and scientific evidence is lacking. However, since distinguishing Alzheimer disease from an NCD due to a general medical condition can be difficult, a trial of a cholinesterase inhibitor may be reasonable.

As behavioral symptoms, such as agitation and psychosis with hallucinations and delusions, occur frequently with NCDs, antipsychotics agents are frequently used. Antipsychotic medications have been associated with increased risk of death, oversedation, and secondary parkinsonism and therefore need to be used sparingly and with caution. Antidepressants such as SSRIs (e.g., sertraline [Zoloft]) and mood stabilizers may be effective in treating associated mood disorders and labile affect. In general, antidepressants with less anticholinergic effects should be used. Benzodiazepines can worsen cognitive functioning and cause delirium and should be avoided in the treatment of anxiety.

MAJOR OR MILD VASCULAR NEUROCOGNITIVE DISORDER

Introduction

Vascular NCD refers to cognitive decline caused by ischemic, hemorrhagic, or oligemic injury to the brain as a consequence of cerebrovascular or cardiovascular disease. It is recognized as a heterogeneous group of disorders with different etiologies.

History

Kraepelin first described “arteriosclerotic dementia” in 1896. This described cognitive impairment as a direct result of arteriosclerotic disease in the brain. Hachinski described “multi-infarct dementia” as dementia related to a series of multiple cerebral infarcts. He published the Hachinski Ischemic Score in 1975 as a clinical tool for evaluating the extent of vascular disease, and an elevated score suggested vascular dementia. Subsequently, additional criteria have been proposed including State of California Alzheimer’s Disease Diagnostic and Treatment Centers (ADDTC) criteria in 1992 and the National Institute of Neurological Communicative Disorders and Stroke–Association Internationale pour la Recherche’ et l’Enseignement en Neurosciences (NINDS-AIREN) criteria in 1993.

Comparative Nosology

There is a general lack of consensus on the definition and criteria for

vascular dementia as evidenced by the use of several clinical diagnostic tools in the literature. Previously, the DSM-IV-TR criteria for dementia due to vascular disease required impairment in memory and at least one other cognitive domain. In DSM-5, only one cognitive deficit, of sufficient severity to affect function, is required in the context of a clinical presentation consistent with cerebrovascular disease. Appropriate context would be either the onset of neurocognitive symptoms temporally related to known cerebrovascular events or a gradual decline in complex attention (especially processing speed) and executive functions in the context of evidence of cerebrovascular disease. The later context would be consistent with small vessel ischemic disease disrupting subcortical–cortical pathways through either extensive white matter changes or lacunar infarcts.

The ICD-10 recognizes an unequal distribution of deficits with possible relative sparing of memory. The ICD-10 further requires one of several specific focal neurological signs and evidence of cerebrovascular disease that is related to the deficits to differentiate vascular dementia from other types of dementia. The ICD-10 further specifies vascular dementia as acute onset, multi-infarct, subcortical, mixed cortical and subcortical, other, and unspecified.

An elevated Hachinski Ischemic Score has been used to identify patients with a vascular etiology of dementia. While this has a high sensitivity, there is a low specificity for detecting vascular dementia.

Table 10.5–8.
Criteria for the Diagnosis of Vascular Dementia from the NINDS-AIREN

I. The criteria for the clinical diagnosis of probable vascular dementia include all of the following:

1. Dementia defined by cognitive decline from a previously higher level of functioning and manifested by impairment of memory and of two or more cognitive domains (orientation, attention, language, visuospatial functions, executive functions, motor control, and praxis), preferably established by clinical examination and documented by neuropsychological testing; deficits should be severe enough to interfere with activities of daily living not due to physical effects of stroke alone.

   Exclusion criteria: Cases with disturbance of consciousness, delirium, psychosis, severe aphasia, or major sensorimotor impairment precluding neuropsychological testing. Also excluded are systemic disorders or other brain diseases (such as Alzheimer disease) that in and of themselves could account for deficits in memory and cognition.
2. Cerebrovascular disease, defined by the presence of focal signs on neurological examination, such as hemiparesis, lower facial weakness, Babinski sign, sensory deficit, hemianopia, and dysarthria consistent with stroke (with or without history of stroke), and evidence of relevant CVD by brain imaging (CT or MRI) including multiple large-vessel infarcts or a single strategically placed infarct (angular gyrus, thalamus, basal forebrain, or PCA or ACA territories), as well as multiple basal ganglia and white matter lacunes, or extensive periventricular white matter lesions, or combinations thereof.
3. A relationship between the above two disorders, manifested or inferred by the presence of one or more of the following: (1) onset of dementia within 3 months following a recognized stroke; (2) abrupt deterioration in cognitive functions, or fluctuating, stepwise progression of cognitive deficits.

II. Clinical features consistent with the diagnosis of probable vascular dementia include the following:
   a. Early presence of gait disturbance (small-step gait or marche a petits pas, or magnetic, apraxic–ataxic, or parkinsonian gait)
   b. History of unsteadiness and frequent, unprovoked falls
   c. Early urinary frequency, urgency, and other urinary symptoms not explained by urological disease
   d. Pseudobulbar palsy
   e. Personality and mood changes, abulia, depression, emotional incontinence, or other subcortical deficits including psychomotor retardation and abnormal executive function

III. Features that make the diagnosis of vascular dementia uncertain or unlikely include:
   a. Early onset of memory deficit and progressive worsening of memory deficit and progressive worsening of memory and other cognitive functions such as language (transcortical sensory aphasia), motor skills (apraxia), and perception (agnosia), in the absence of corresponding focal lesions on brain imaging
   b. Absence of focal neurological signs, other than cognitive disturbance
   c. Absence of cerebrovascular lesions on brain CT or MRI

IV. Clinical diagnosis of possible vascular dementia may be made in the presence of dementia (Section I-1) with focal neurological signs in patients in whom brain imaging studies to confirm definite CVD are missing; or in the absence of a clear temporal relationship between dementia and stroke; or in patients with subtle onset and variable course (plateau or improvement) of cognitive deficits and evidence of relevant CVD.

V. Criteria for the diagnosis of definite vascular dementia are as follows:
   a. Clinical criteria for probable vascular dementia
   b. Histopathological evidence of CVD obtained from biopsy or autopsy
   c. Absence of neurofibrillary tangles and neuritic plaques exceeding those expected for age
   d. Absence of other clinical or pathological disorder capable of producing dementia

VI. Classification of vascular dementia for research purposes may be made on the basis of clinical, radiological, and neuropathological features, for subcategories or defined conditions such as cortical vascular dementia, subcortical vascular dementia, Binswanger disease, and thalamic dementia.

The term *Alzheimer disease with CVD* should be reserved to classify patients fulfilling the clinical criteria for possible Alzheimer disease and who also present clinical or brain imaging evidence of relevant CVD. Traditionally, these patients have been included with vascular dementia in epidemiological studies. The term *mixed dementia*, used hitherto, should be avoided.

ACA, anterior cerebral artery; CVD, cerebrovascular disease; CT, computed tomography; MRI, magnetic resonance imaging; NINDS-AIREN, National Institute of Neurological Communicative Disorders and Stroke–Association Internationale pour la Recherche' et l'Enseignement en Neurosciences; PCA, posterior cerebral artery.

From Román GC, Tatemichi TK, Erkinjuntti T, et al. Vascular dementia: Diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology*. 1993;43(2):250.

The NINDS-AIREN criteria (Table 10.5–8) are the most widely accepted and are commonly used in research studies. These criteria require deficits in memory and at least two other cognitive domains, the presence of focal neurological signs and symptoms, neuroimaging evidence of cerebrovascular disease, and a relationship between the two. They also identify clinical features consistent with a diagnosis of probable vascular dementia including early gait disturbance, falls, urinary symptoms,

pseudobulbar palsy, personality and mood changes, and subcortical deficits including executive dysfunction. These stringent criteria are quite specific but less sensitive for detecting vascular dementia. They effectively exclude 91 percent of cases with Alzheimer disease pathology; hence, these criteria may be more sensitive for pure vascular dementia but will not detect cases of mixed Alzheimer disease plus cerebrovascular disease cases.

In contrast, the ADDTC criteria (Table 10.5–9) are more inclusive and hence include more mixed cases than the NINDS-AIREN criteria. These criteria do not operationalize the definition of dementia and do not require specific focal neurological signs, but only the presence of stroke and evidence of cerebrovascular disease.

Epidemiology

Vascular dementia accounts for about 10 to 20 percent of dementia when only a pure vascular etiology is considered. Isolated vascular dementia was considered the second most common form of dementia until dementia with Lewy bodies was recognized as a more common pathology, so isolated vascular dementia is now considered the third most common form of dementia in North America. However, if one includes the "mixed" cases where Alzheimer disease and cerebrovascular disease or vascular NCD coexist, the number is considerably higher, and some estimates put it greater than the combination of dementia with Lewy bodies with Alzheimer disease.

Age and evidence of vascular disease (e.g., peripheral or cardiac vascular disease) are among the biggest risk factors for vascular dementia and mixed Alzheimer disease plus cerebrovascular disease. Other vascular risk factors (e.g., hypertension and diabetes) also contribute to vascular dementia and mixed Alzheimer disease plus cerebrovascular disease.

Table 10.5–9.
Criteria for the Diagnosis and Prevalence of Probable Ischemic Vascular Dementia from the ADDTC

	No.	%
A. 1. Dementia (as defined in the text)	85	
2. Evidence of two or more ischemic strokes by	43[a]	50.6
History, neurological signs, and/or	38[a]	44.7
Neuroimaging studies (CT or T_1-weighted MRI), or	22[a]	25.9
Occurrence of a single stroke with a clearly documented temporal relationship to the onset of dementia	0[a]	0
3. Evidence of ≥1 infarct outside the cerebellum by CT or T_1-weighted MRI	33[a]	38.8
B. Diagnosis of probable IVD is supported by		
1. Evidence of multiple infarcts in brain regions known to affect cognition (as defined by NINDS-AIREN criteria)	22[a]	25.9

2. History of multiple transient ischemic attacks	25[a]	29.4
3. History of vascular risk factors (e.g., hypertension, heart disease, diabetes mellitus)	67[a]	78.8
4. Elevated Hachinski Ischemia Scale score (≥7)	31[a]	36.5
C. Clinical features that are thought to be associated with IVD but await further research		
1. Relatively early appearance of gait disturbance and urinary incontinence	46[a]	54.1
2. Periventricular and deep white matter changes on T_2-weighted MRI that are excessive for age[b]	. . .	. . .
3. Focal changes in electroencephalographic studies[c]	27[a]	31.8
D. Other clinical features that do not constitute strong evidence either for or against a diagnosis of probable IVD		
1. Periods of slowly progressive symptoms	46[a]	54.1
2. Illusions, psychoses, hallucinations, delusions	26[a]	30.6
3. Seizures	2[a]	2.4
E. Clinical features that cast doubt on a diagnosis of probable IVD		
1. Transcortical sensory aphasia in the absence of corresponding focal lesions on neuroimaging studies	0[a]	0
2. Absence of central neurological symptoms/signs other than cognitive disturbance	21[a]	24.7
Criteria A and E fulfilled	23[a]	27.1

ADDTC, Alzheimer's Disease Diagnostic and Treatment Centers; CT, computed tomography; IVD, ischemic vascular dementia; MRI, magnetic resonance imaging; NINDS-AIREN, National Institute of Neurological Communicative Disorders and Stroke–Association Internationale pour la Recherche' et l'Enseignement en Neurosciences.

[a]Of 85 subjects with dementia.

[b]MRI scans were not available in this study.

[c]Further investigations as proposed in the ADDTC criteria (evoked potentials, single-photon emission CT, positron emission tomography, nuclear magnetic resonance spectroscopy) were not performed in this study.

From Wetterling T, Kanitz RD, Borgis KJ. Comparison of different diagnostic criteria for vascular dementia (ADDTC, DSM-IV, ICD-10, NINDS-AIREN). *Stroke*. 1996;27:30.

Etiology

The etiology of vascular dementia is heterogeneous. Stroke is considered the leading cause and can be divided into three subtypes: (1) cortical, (2) subcortical, and (3) strategic infarct. Some level of cognitive impairment develops within 3 months in 35 to 62 percent of individuals who suffer a clinically apparent stroke. Multiple cortical infarcts, historically termed "multi-infarct dementia," involve multiple small infarcts throughout cortical regions. This is generally a result of large vessel disease. The location and size of the infarct determines the type and extent of cognitive impairment. Subcortical infarcts (usually affecting basal ganglia, thalamus, or deep or periventricular white matter) disrupt strategic cortical–

subcortical circuits, creating disconnection syndromes. These generally reflect small vessel disease. In addition to frank infarcts, small vessel ischemic disease leading to extensive white matter damage may disrupt cortical–subcortical pathways. Binswanger disease, or subcortical vascular encephalopathy, refers to atherosclerotic changes of small vessels resulting in a subcortical dementia. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) is an inherited disorder caused by mutation in the *NOTCH3* gene. It results in osmiophilic deposits in the basal lamina of small intracerebral arteries and arterioles. Patients develop lacunar infarcts and white matter changes causing a subcortical dementia. A single strategically located infarct can also produce dementia. Key regions resulting in cognitive impairment after a single stroke include the thalamus, caudate nucleus, lenticular nucleus, angular gyrus, genu of the internal capsule, and inferomedial temporal lobe. After any stroke, additional brain damage may occur over time through Wallerian degeneration. Other cerebrovascular diseases are also associated with vascular dementia. Hemorrhage, including hemorrhagic stroke, subdural hematoma, epidural hematoma, or intraparenchymal hemorrhage from aneurysm or trauma, can also produce cognitive impairment and dementia. In addition, chronic hypoperfusion resulting from hypotension or other cardiopulmonary etiologies can also cause a dementia syndrome. Cardiovascular contributing factors to vascular dementia include congestive heart failure and cardiac arrhythmias such as atrial fibrillation. Cerebral amyloid angiopathy is another cause of cerebrovascular disease and vascular dementia.

The co-occurrence of vascular NCD and Alzheimer disease reflects the overlap in risk factors and pathophysiology of these two disorders. Brain tissue damage caused by stroke induces a cerebral amyloidogenesis. This leads to amyloid changes similar to those seen in Alzheimer disease including soluble amyloid accumulation, elevated CSF-τ, and elevated synaptophysin levels indicative of axonal degeneration. The apolipoprotein ε4 allele, which is a risk factor for Alzheimer disease, is also a risk factor for the development of dementia following stroke. Apolipoprotein E is involved in cholesterol metabolism and is associated with high levels of low-density lipoproteins. The association of apolipoprotein E with both Alzheimer disease and vascular dementia may help to explain some of the interaction between the two forms of dementia and high incidence of comorbidity. Vascular risk factors that contribute to the risk of stroke and small vessel ischemia include hypertension, diabetes, hyperlipidemia, peripheral artery disease, and smoking. These factors also play a role in the etiology of Alzheimer disease, suggesting some common pathophysiological mechanisms. Cerebrovascular disease also causes a cholinergic deficit similar to that seen in Alzheimer disease. Cholinergic pathways are affected in vascular dementia, and several studies have shown low levels of

acetylcholine in the CSF and reduced levels of acetylcholinesterase transferase in the cortex, hippocampus, and striatum. The Nun Study and the Cognitive Function in Ageing Study both highlighted the interaction between Alzheimer disease pathology and cerebrovascular disease. In the Nun Study, in patients with Alzheimer disease, brain infarction was a better predictor of dementia severity than was density of plaques and tangles. The overlap between Alzheimer disease and cerebrovascular disease is great, with about one-third of patients with Alzheimer disease exhibiting significant cerebrovascular disease and one-third of patients with vascular dementia also having typical Alzheimer disease neuropathology.

Diagnosis and Clinical Features

According to the NINDS-AIREN, diagnosis of probable vascular dementia is made if dementia is associated with focal neurological signs and symptoms and neuroimaging evidence of cerebrovascular disease is present. It is important to establish a temporal and causal relationship between the brain lesion and the cognitive impairment. The cognitive symptoms leading to a presentation of vascular dementia are heterogeneous and are dependent on the type and location of the vascular lesions and etiology.

Cortical strokes are associated with typical neurological signs and symptoms including lower facial weakness, upper motor neuron lesions, and gait abnormalities. Cognitive deficits after cortical strokes reflect losses that map to the region of stroke (e.g., aphasia after stroke of Broca area, left hemi-neglect after right parietal stroke). Memory impairment and loss of executive control and information processing are common.

Subcortical vascular dementia may present with subtle focal signs including gait disorder, imbalance, and urinary incontinence. The lesions may be lacunar infarcts or extensive periventricular white matter changes. These lesions particularly affect the prefrontal–subcortical circuit, producing cognitive deficits such as a dysexecutive syndrome, mental slowing, and reduced ability to set and reach goals. Memory loss may be minor. Subcortical disease is often associated with depression, emotional lability, and personality change. This presentation may be similar to that of frontotemporal dementia.

The onset of cognitive impairment after stroke is often acute, coinciding with a single vascular event. Indeed, one strategic infarct produces “poststroke dementia” in at least 30 percent of individuals over age 70 who suffer a single stroke. However, cognitive impairment can also have a slow insidious onset similar to Alzheimer disease or frontotemporal dementia, reflecting the accumulation of cognitive impairments over a series of strokes; this is the more common pattern for subcortical vascular dementia. Vascular cognitive impairment may remain static if there are no further vascular events. However, cognitive impairment may worsen if

there are subsequent strokes or due to Wallerian degeneration.

In general, compared to Alzheimer disease, the neuropsychological impairment of vascular dementia is associated with relatively preserved long-term memory but greater executive dysfunction, which can occur with disruption to any part of the frontal–subcortical circuitry.

Pathology and Laboratory Examination

Structural neuroimaging reveals vascular damage, either ischemic or hemorrhagic lesions. These may range from one strategically placed large vessel lesion to multiple scattered lesions to extensive small vessel ischemic changes in the white matter regions. The vascular damage evident will vary depending on the etiology but must correspond to the cognitive deficits observed.

There are no clear pathological changes diagnostic for vascular dementia. Evidence of stroke is required, and some schemes have suggested quantification of total volume of infarction, though these approaches have not reached the level of standardized criteria. Evidence of amyloidopathy may be present but is not required. The presence of neuropathological changes characteristic of Alzheimer disease does not exclude vascular dementia, as the two processes are commonly comorbid and may both be diagnosed.

SM is an 88-year-old man with a history of hypertension and a remote myocardial infarction, who lives alone in the community. He was brought to the hospital by his daughter who had found him lying on the floor, too weak to get up. He was able to provide a coherent history of the current illness. Medical workup revealed uremia secondary to prostate obstruction, and he was immediately sent for transurethral prostate resection under general anesthesia. Prior to surgery, he was alert and oriented to person, place, and time but was noted to have mild deficits in short-term recall as he could register three objects but could only recall two objects at 5 minutes. However, SM was able to provide valid informed consent for surgery as he was able to verbalize an understanding of the procedure as well as the risks and benefits.

The day after surgery, SM was noted to have significantly impaired short-term memory. Although he was able to register three objects, he could not recall any objects at 5 minutes. SM was not delirious as his attention was intact, which he demonstrated by spelling WORLD forward and backward without difficulty.

It was discovered that during the operation, SM's systolic blood pressure was 30 to 40 mm Hg lower than was his norm. The psychiatry consultant diagnosed major vascular NCD (due to cerebral hypoperfusion). Specifically, his blood pressure was kept much lower than what was typical for this elderly man. The hippocampus is particularly prone to hypoperfusion injury; hence, the impairment is short-term memory.

Differential Diagnosis

Alzheimer Disease with Comorbid Cerebrovascular Disease. Alzheimer disease frequently co-occurs with cerebrovascular disease. Differentiating the contribution of each to cognitive impairment can be difficult. In the setting of NCD due to Alzheimer disease, the presence of a stroke or of significant white matter hyperintensities on brain imaging, the presence of

focal neurological signs and symptoms, and/or a clinical history of an abrupt cognitive deterioration would suggest that the appropriate diagnosis is of an NCD due to multiple etiologies (Alzheimer disease and cerebrovascular disease).

Frontotemporal Dementia. Frontotemporal dementia may have a similar presentation as subcortical vascular dementia with slow insidious onset, frontal lobe executive deficits, relatively preserved memory early in the course, and the presence of neuropsychiatric manifestations such as personality change. Neuroimaging studies demonstrating a frontal and/or temporal pattern of circumscribed atrophy, blood flow, or metabolic reductions may be helpful in the differential diagnosis.

Course and Prognosis

The course of vascular dementia is dependent on the nature and course of the underlying. Onset is often sudden if the cognitive impairment is associated with a particular large or strategic vascular event. Onset will be more insidious and progression will be slower when associated with subcortical or small vessel changes. The course may be static if there are not further vascular events or progressive often with a fluctuating stepwise decline coinciding with further vascular events or possible Wallerian degeneration.

Treatment

Treatment of vascular dementia can be divided into primary prevention and symptomatic treatment.

Primary prevention involves controlling or ameliorating vascular risk factors to prevent vascular damage from occurring in the brain. There are several modifiable risk factors for vascular disease that can be addressed. Hypertension should be adequately controlled. Hypotension should be avoided in patients who have already had stroke. Treatment of diabetes should focus on optimal glycemic control to try to prevent small vessel damage and ischemia. Hyperlipidemia should be treated with statins and related medications to reduce risk of stroke. Atrial fibrillation should be treated with anticoagulants to prevent thrombus formation. Severe carotid stenosis should be treated with endarterectomy or angioplasty. Patients with a history of transient ischemic attacks or stroke should be prescribed antiplatelet therapy to reduce the risk of recurrent ischemic events. Patients should be encouraged to make healthy lifestyle changes such as diet, exercise, weight loss, stress reduction, decreased salt intake, and cessation of smoking. Patients with sleep apnea should also be adequately treated with positive airway pressure to optimize cardiopulmonary functioning. While these recommendations seem obvious to try to prevent

stroke, they have not been studied in prevention trials of vascular dementia.

Secondary prevention to modify vascular risk factors after vascular dementia has developed also seems logical. However, there is no evidence to support that it prevents further deterioration in cognitive functioning. Intensive rehabilitation after stroke may help restore cognitive function that was impaired during a cerebrovascular event.

There are no FDA-approved pharmacological treatments for vascular NCD. Given that the cholinergic deficit in vascular dementia is similar to that seen in Alzheimer disease, cholinesterase inhibitors have been evaluated for benefit in this group. Donepezil has been the best studied and its use has been associated with improvement in cognitive function, activities of daily living, and global functioning, as well as less functional deterioration compared to placebo. Galantamine and rivastigmine have been less rigorously studied, but reports suggest a benefit on cognition as well. Memantine, an NMDA receptor antagonist, has demonstrated cognitive benefit in less severe vascular NCD and cognitive and functional benefit in severe disease.

NEUROCOGNITIVE DISORDERS DUE TO TRAUMATIC BRAIN INJURY

Definition

In recent years, the impact of TBI has garnered increasing public attention. In addition to the emerging understanding of the long-term sequelae of moderate-to-severe TBI, there is a growing consensus that mild TBI or concussion injuries are not necessarily “mild” or benign and are often temporarily accompanied by significant impairment. Psychiatry is uniquely situated to help affected individuals deal with the cognitive, psychiatric, and emotional burden of TBI, due to the field’s focus on functionality and training in psychopharmacology.

Comparative Nosology

In DSM-IV-TR, head injuries were classified under the cognitive disorders not otherwise specified, under the subheadings of mild NCD and postconcussional disorder. These disorders have been eliminated from the DSM-5. In the DSM-5, cognitive impairment due to TBI is encompassed by the diagnoses of the expanded syndromes major NCD due to TBI, with and without behavioral disturbance, and mild NCD due to TBI. DSM-5 does not specify a mechanism of injury but does require evidence that the traumatic event disrupted brain function by requiring that the patient experienced, postinjury, either loss of consciousness (LOC), postinjury amnesia, disorientation and confusion, or other (clinical or neuroimaging) sign of neurological injury.

Table 10.5–10.
Severity Ratings for Traumatic Brain Injury[a]

	TBI Severity		
	Mild	**Moderate**	**Severe**
Neuroimaging[b]	Normal	Normal or abnormal	Normal or abnormal
Loss of consciousness (LOC)	<30 minutes	30 minutes–24 hours	>24 hours
Alterations of consciousness/mental state (AOC)	≤24 hours	>24 hours	
Post-traumatic amnesia (PTA)	<24 hours	24 hours–7 days	>7 days
Glasgow Coma Scale Score[c]	13–15[d]	9–12	3–8

[a]Department of Veterans Affairs, Department of Defense. VA/DoD Clinical Practice Guideline for Management of Concussion/Mild Traumatic Brain Injury (mTBI), Version 1.0 2009

[b]Possible findings on neuroimaging include petechial hemorrhages, subarachnoid hemorrhage, or evidence of contusion on CT scan, and additionally hyperintensities suggestive of microhemorrhages on MRI.

[c]Evaluation of disorientation and confusion at initial assessment.

[d]Glasgow Coma Scale Score should not be below 13 at 30 minutes.

The most commonly used TBI severity classification system (see Table 10.5–10), as outlined by the US Department of Defense (DoD), stratifies TBI injuries into mild, moderate, and severe TBI based on structural imaging, LOC, alteration of consciousness/mental state (AOC), posttraumatic amnesia (PTA), and Glasgow Coma Scale scores. While currently the accepted standard, this classification does not have strong clinical validation. For example, the DOD classification system does not distinguish between a mild TBI with LOC nearing 30 minutes and PTA nearing 24 hours versus mild TBI in which there was brief or no LOC and PTA even though these two injuries will produce different clinical courses (American Academy of Neurology 1997). The DSM-5 does not outline specific criterion to relate major or mild NCD due to TBI to the DOD criteria, but suggests that major NCD due to TBI corresponds to moderate and severe TBI, whereas mild NCD corresponds to mild TBI (mTBI)/concussion injuries.

In ICD-10, cognitive impairment due to TBI is reflected in either *dementia in other specified diseases classified elsewhere with or without behavioral disturbance* (F02.81 or F02.80) or *mild cognitive impairment, so stated* (G31.84).

Epidemiology

TBI rates vary by age, with the highest prevalence among individuals younger than 4 years, older adolescents, and individuals older than 65 years. In the United States in 2010 there were 2.2 million TBI-related

emergency department (ED) visits, 280,000 TBI-related hospitalizations, and more than more than 50,000 TBI-related deaths, with TBI contributing to 30 percent of all injury related deaths. From 2001 to 2010, TBI-related ED visits increased by 70 percent, TBI related hospitalizations increased by 11 percent, and TBI related deaths decreased by 7 percent. In 2009, there were an estimated 248,418 pediatric (age 19 or younger) TBI-related ED visits, and from 2001 to 2009 ED visits for sports or recreation-related TBI rose 57 percent in this population. Men had higher rates than women of TBI-related ED visits, hospitalization, and death.

As illustrated above, surveillance to estimate rates of TBI relies heavily on ED visits, hospitalization, and mortality data after an acute injury. These rates are likely a gross underestimate for mild TBI and concussion injuries as these surveillance methods fail to capture events where medical attention was not sought, or sought later in a primary care or specialty setting for lingering symptoms.

Etiology

A TBI occurs when the brain collides with inside of the skull with enough force to cause cerebral contusions and axonal injury at either the external site of impact or at the opposite pole due to rapid acceleration and deceleration of the brain. Possible mechanisms of injury include direct blows to the head, the head striking an object (including falls), acceleration/deceleration movement causing internal contact between the brain and the skull without external contact to the head (e.g., whip lash type injury during a motor vehicle accident), external forces including explosions and blasts, and penetration of the brain with a foreign body. Shearing and rotational forces result in neuronal membrane damage, microvascular changes, blood–brain barrier disruption, swelling, and neuronal death. Concurrent subdural or epidural hematoma or intracranial hemorrhage can also cause compression of surrounding brain tissue, and when visible on imaging suggests a more severe injury. Any brain structure can be injured and specific cognitive deficits and symptoms may suggest areas of damage.

Overall the leading causes of TBIs resulting in ED visits, hospitalizations, and death from 2006 to 2010 were falls (40 percent), unintentional blunt trauma to the head (15 percent), motor vehicle accidents (14 percent), and assaults (10 percent). Falls accounted for a greater percent of TBIs among children aged 0 to 14 (55 percent) and elderly older than 65 years (81 percent). Of assault-associated TBI, 75 percent occurred in young adults 15 to 44 years of age.

Diagnosis and Clinical Features

The DSM-5 criterion for major and mild NCD due to TBI requires evidence

that the traumatic event disrupted brain function by requiring that the patient experienced, postinjury, either LOC, postinjury amnesia, disorientation and confusion, or other (clinical or neuroimaging) sign of neurological injury. Furthermore, the DSM states that to be attributable to TBI, the NCD must present immediately after the brain injury (and the individual has recovered consciousness) and must persist past the acute postinjury period. These features are in addition to core cognitive symptoms that can include impairment in speed of information processing and complex attention, executive abilities, and learning and memory, as well as disturbances in social cognition. Aphasia, neglect, and constructional dyspraxia may be seen in more severe TBIs.

There are numerous other somatic and psychological postconcussive symptoms that can be clinically significant in NCD due to TBI. The specific features are dependent on what area of the brain has been affected. Common somatic symptoms include headaches (typically migraine, occipital neuralgia, or cervicogenic), cognitive fatigue, physical fatigue, sensitivity to light/noise, insomnia and other sleep disturbances, drowsiness, dizziness, nausea and vomiting, vision problems, transient neurological abnormalities, seizures, and balance problems. Of particular concern to mental health professionals, the psychological symptoms following a TBI can include personality change (e.g., disinhibition, apathy, suspiciousness, aggression) as well as disturbances in emotional function (e.g., irritability, easy frustration, tension and anxiety, affective lability).

Pathology and Laboratory Examination

Beyond neuropsychological testing, CT scanning may reveal petechial hemorrhages, subarachnoid hemorrhage, or evidence of contusion. Magnetic resonance image (MRI) scanning may also reveal hyperintensities suggestive of microhemorrhages.

Differential Diagnosis

In some instances, severity of neurocognitive symptoms may appear to be inconsistent with the severity of the TBI. After previously undetected neurological complications (e.g., chronic hematoma) are excluded, the possibility of diagnoses such as somatic symptom disorder or factitious disorder needs to be considered. Posttraumatic stress disorder (PTSD) and depressive symptoms can co-occur with the NCD and have overlapping cognitive symptoms (e.g., difficulty concentrating). Among individuals with substance use disorders, the neurocognitive effects of the substance can contribute to or compound the TBI-associated neurocognitive change.

Course and Prognosis

The severity rating of the TBI itself does not necessarily predict severity of

the resulting NCD. The course of recovery from TBI is variable, depending not only on the specifics of the injury but also on cofactors, such as age, prior history of brain injuries, or substance abuse, that may favor or impede recovery.

Nevertheless, the typical recovery course in mild TBI/concussion injuries is for complete or substantial improvement in cognitive, neurological, and psychiatric symptoms. Cognitive symptoms associated with mild TBI typically resolve within days to weeks after the injury with complete resolution typical by 3 months, although some patients experience a more protracted course before making a full recovery. Psychiatric symptoms (e.g., depression, irritability, sleep disturbance) and neurological symptoms (e.g., headache, photosensitivity) also tend to resolve in the weeks following mild TBI. However, the presence of dizziness immediately after injury and cognitive fog present 3 days postinjury are associated with longer recovery times; substantial subsequent deterioration in these areas should trigger consideration of additional diagnoses. Unfortunately, repeated mild TBI may be associated with persisting neurocognitive disturbance. For moderate and severe TBI, there can be persistence of neurocognitive deficits along with persistence of neurological (e.g., seizures—particularly in the first year, photosensitivity, hyperacusis) and psychiatric symptoms (e.g., irritability, aggression, depression, sleep disturbance, apathy), which taken together, can lead to long term or permanent disability resulting in the inability to resume occupational and social functioning at preinjury level, and deterioration in interpersonal relationships. Worse outcome from moderate-to-severe TBI is associated with older age (older than 40 years) and initial clinical parameters, such as low Glasgow Coma Scale score, worse motor function, pupillary nonreactivity, and computed tomography (CT) evidence of brain injury (e.g., petechial hemorrhages, subarachnoid hemorrhage, midline shift, obliteration of third ventricle). Moderate and severe TBI have also been associated with increased risk of neurodegenerative diseases such as Alzheimer disease or chronic traumatic encephalopathy.

The features of persisting major or mild NCD due to TBI will vary by age. Persisting TBI-related impairment in an infant or child may be reflected in delays in reaching developmental milestones (e.g., language acquisition), worse academic performance, and possibly impaired social development. Among older teenagers and adults, persisting symptoms may include various cognitive deficits, irritability, hypersensitivity to light and sound, easy fatigability, and mood changes, including depression, anxiety, hostility, or apathy. Older individuals with depleted cognitive reserve may experience incomplete recoveries even after a mild TBI.

Treatment

In the acute treatment of TBI immediately after the injury, studies of

multiple agents (progesterone, monoaminergic agonists, mannitol, magnesium, corticosteroids or aminosteroids, calcium channel blockers, bradykinin β-2 receptor antagonists, and barbiturates) have not identified a helpful, safe pharmacological treatment in the acute phase following a TBI. In addition, there is insufficient evidence to suggest that other proposed interventions such as cooling, hyperventilation, or hyperbaric oxygen provide any benefit in acutely treating TBI.

Both nonpharmacological and pharmacological treatments can be employed in addressing cognitive impairments that persist beyond the acute TBI injury. Cognitive rehabilitation consists in assisting patients in developing strategies to minimize the impact of impairment to memory, attention, executive functioning, and interpersonal skills. Progressive physical exercise has also shown to have a beneficial impact on overall cognitive functioning in patients recovering from an mTBI/concussion injury (Silverberg et al., 2013).

Evidence suggests that some pharmacotherapies may be a helpful adjunct to cognitive rehabilitation in treating persisting cognitive impairment. Stimulants (of which methylphenidate is the most studied) augment dopaminergic and noradrenergic levels in the brain, may improve processing speed, and may also increase arousal and help with attention and memory. Other stimulating agents, including amantadine that has a "light" stimulant action, also appear to be helpful. Studies also suggest that donepezil may be helpful in treating memory impairment post moderate to severe TBI. For the patient with TBI, management will also include treatment of other associated symptoms such as depression and other mood disorders, headaches, insomnia and other sleep disturbances, drowsiness, dizziness and other vestibular symptoms, nausea and vomiting, vision problems, seizures, and balance problems. For these features, management typically involves interventions including both standard pharmacological or behavior treatments.

MAJOR OR MILD NEUROCOGNITIVE DISORDER DUE TO HIV INFECTION

Definition

HIV type 1 (HIV-1) is commonly associated with a spectrum cognitive impairments known collectively as HIV-associated neurocognitive disorder (HAND). HIV-associated dementia (HAD) is the most severe form, whereas asymptomatic neurocognitive impairment (ANI) is the least severe form, and mild NCD lies in between these two extremes. The presence of even mild cognitive impairment has real clinical consequences in the HIV+ population as even subtle deficits are associated with reduced survival, decreased quality of life, reduced medication adherence, and worse quality of life.

Comparative Nosology

The *HAND* nomenclature was originally proposed by the AIDS Task Force of the American Academy of Neurology in 1997 and revised by the National Institute of Mental Health and the National Institute of Neurological Disorders and Stroke in 2007. HAND is further broken down into three subcategories, based on the extent of neuropsychological testing deficits and impairment in functional performance.

HAD, the most severe form, is also known as AIDS dementia complex, HIV encephalitis, or encephalopathy. Individuals with HAD score at least two standard deviations below the mean on at least two cognitive domains on of standardized neuropsychological testing and have marked impairment in activities of daily living.

Individuals with HIV-associated mild NCD score of one standard deviation below the mean on at least two cognitive domains on standardized neuropsychological testing and demonstrate at least mild impairment of daily functioning. Both HAD and mild NCD would align with major NCD due to HIV infection in DSM-5 terminology and dementia in HIV disease in ICD-10. Individuals with ANI are HIV-seropositive individuals who score of at least one standard deviation below the mean on at least two cognitive domains on standardized neuropsychological testing but do not demonstrate functional impairment. This category would be aligned with mild NCD due to HIV infection in DSM-5 terminology and mild cognitive impairment, so stated (G31.84) in ICD-10.

Epidemiology

The introduction of effective combination antiretroviral treatment (cART) has not only improved survival of patients with HIV but also changed the relative prevalence of the three HAND categories. The prevalence of HAD pre-cART was 20 percent. With the introduction of cART, the prevalence has decreased to less than 5 percent. However, the prevalence of the milder forms of cognitive impairment, ANI, and mild NCD has increased and is estimated to be 20 to 50 percent. In addition, the prevalence of other non-HIV dementias is increasing in the aging HIV population in the Western world. Unfortunately, these positive changes are primarily seen in countries where HIV treatment is available. In sub-Saharan Africa and other regions where HIV treatment is not widely available, the prevalence of HAD remains elevated at levels consistent with what was seen in the pre-cART era in the Western world.

Etiology

HIV enters the CNS early upon initial infection and causes neurological injury through increased levels of proinflammatory cytokines. On neuroimaging, this neurological injury is seen as degeneration of the white

matter tracks with gray matter loss in the basal ganglia and posterior cortex. Neuropathological studies suggest that HAND is a single syndrome in which neuropathological findings are shared across the spectrum of cognitive impairment (ANI to mild NCD to HAD). Individuals who achieve viral suppression through cART may continue to demonstrate cognitive impairment, suggesting that whatever neurological damage that initially occurs is permanent. However, some evidence suggests that CNS inflammation continues and becomes chronic, even when viral replication is suppressed. This inflammation may contribute to current and future cognitive impairment. The potential brain toxic effects of cART on long-term cognitive state are unclear.

Diagnosis and Clinical Features

The diagnosis requires assessment of both neuropsychological performance and functional abilities. While several validated screening measure exist, formal neuropsychological assessment is the gold standard. All patients who are HIV positive should be screened for HAND. Ideally, patients would be tested as soon as possible after HIV diagnosis and before initiation of cART therapy. A recent consensus statement suggests that those patients at higher risk of cognitive deterioration should be reassessed every 6 to 12 months, whereas those at low risk should be reassessed every 12 to 24 months. Reassessment should also occur whenever there is a change in clinical status, HIV treatment, or a diagnosis of a psychiatric disorder.

Adult patients for whom the HIV infection is diagnosed early are assumed to progress from ANI to NMD to HAD, although there may be years of intervening stability. In children, HAD may present as neurodevelopmental delay.

The pattern of cognitive deficits seen in HAND has changed since the introduction of cART. In the pre-cART era, HAND was characterized primarily by subcortical deficits (e.g., impaired concentration, reduced psychomotor speed) and behavioral disturbances (withdrawal). Only in advanced disease were cortical deficits (e.g., severe memory impairment, language impairment) seen. Also notable, in the pre-cART era, motor deficits were commonly observed. With the advent of cART therapy, HAND is now primarily characterized by cortical deficits, specifically impaired executive functions and impaired learning and memory. The reason behind this shift is unclear but may represent the contribution of other causes of cognitive impairment (e.g., vascular injury) to the aging population of HIV+ individuals. Motor deficits are now rare.

While there is no clear profile of the patient at risk for HAND, there are several factors that raise the risk of developing HAND (Table 10.5–11).

Pathology and Laboratory Examination

Assessment of a HIV+ individual with cognitive impairment should include a broad workup for underlying medical causes other than HIV infection. Patients with HIV remain at risk for cerebrovascular disease, Alzheimer disease, and other causes of cognitive impairment. Additionally, HIV+ patients are at risk for opportunistic CNS infections such as cryptococcal meningitis, and other CNS disease, such as CNS lymphoma.

The clinical utility of quantifying CSF HIV RNA load for the purposes of diagnosing HAND is unclear. CSF HIV RNA load correlated with severity of impairment in the pre-cART era, but this no longer holds true in the cART era. However, a lumbar puncture with sampling of the CSF may be useful in order to assess for opportunistic infections.

MRI of the brain typically demonstrates diffuse white matter hyperintensities on T2-weighted images. In addition, cerebral atrophy is common. These findings are nonspecific. However, MRI is needed to rule out other causes of cognitive impairment, including secondary CNS diseases and infections. Diffusion tensor imaging (DTI) typically demonstrates more specific impairments, specifically abnormal signal in the corpus callosum, frontal white matter, internal capsule, and basal ganglia. The extent of abnormalities detected on DTI correlate with severity of cognitive impairment.

While not typically included in a diagnostic workup, on MR spectroscopy, decreased *N*-acetyl aspartate (NAA), reflecting neuronal injury, and increased choline (Cho) and myo-inositol (MI) levels, likely reflecting glial cell proliferation is seen in the frontal white matter tracks and the basal ganglia. Improvement in cognition in response to adequate cART therapy correlates with improvement in NAA levels.

Differential Diagnosis

Of particular importance, the differential diagnosis includes secondary CNS diseases including progressive multifocal leukoencephalopathy (PML) and CNS lymphoma as well as opportunistic infections including cytomegalovirus (CMV) encephalitis, cerebral toxoplasmosis, and cryptococcal meningitis. With the increase in life expectancy in the cART era, HIV+ individuals are now at risk of cognitive impairment due to other causes such as cerebrovascular disease, Alzheimer disease, substance use, and other causes. At this time, the contribution of cART to cognitive dysfunction through neurotoxic effects cannot be excluded.

Course and Prognosis

There are few studies examining the rate of progression from ANI to mild NCD to HAD. However, individuals with ANI are at increased risk of progression to worse cognitive function than HIV+ individuals without neurocognitive impairment. Risk factors for progression to more severe

cognitive impairment are listed in Table 10.5–11 and include higher burden of HIV virus, lack of robust cART treatment, comorbid medical, psychiatric and neurological illness, and demographic factors such as advanced age, low education and female gender.

Table 10.5–11.
Risk Factors for HAND

Disease-Specific Factors	Treatment Factors	Neurological and Psychiatric Factors
Low nadir CD4+ T-cell count (<200 cells/µL)	Poor treatment adherence	Psychiatric disorders, e.g., MDD, PTSD
High plasma HIV RNA	Treatment interruption	Illicit drug/alcohol abuse/dependence
High CSF HIV RNA	Nonoptimal cART regimen	Syphilis
Prior HIV-related CNS diseases	cART neurotoxicity	Alzheimer disease
Longer HIV duration	cART with lower CNS penetration effectiveness (controversial)	Cerebrovascular disease
		Traumatic brain injury and seizure
		Vitamin or hormone deficiency
		Prior HCV coinfection
Comorbidities	**Demographic factors**	
Positive HCV serostatus with high HCV RNA	Older age	
Stroke	Low education level	
Cerebrovascular disease risk factors (hyperlipidemia, elevated blood pressure, diabetes)	Female gender	
Anemia	Limited access to care	

HAND, HIV-associated neurocognitive disorder; HCV, hepatitis C virus; HIV, human immunodeficiency virus; CNS, central nervous system; CSF, cerebrospinal fluid; PTSD, posttraumatic stress disorder; MDD, major depression disorder; cART, combination antiretroviral therapy.

Treatment. Prevention of HAND is important. Some evidence suggests that early initiation of cART reduces incidence of HAND. However, there is no evidence at this time as to whether the early use of cART regimes that have higher CNS penetration is any better of preventing HAND than regimes with lower CNS penetration. Given the increased life expectancy, prevention also involves reducing the risk of cerebrovascular disease through management of metabolic parameters.

The mainstay of HAND treatment is cART. cART regimes not only slow or stop cognitive decline but also modestly improve cognitive function. Modest benefits in cognitive status, especially in the domains of attention, processing speed and executive function are seen after 1 to 12 months of cART. At this time, the evidence is inconsistent as to whether the use of cART regimes that have higher CNS penetration are any better at

improving cognition than regimes with lower CNS penetration. Nevertheless, an expert consensus guideline recommends use of a cART regimen with high CNS penetration in patients with HAND. The possibility of neurotoxicity from the cART regimen must be considered if a patient experiences cognitive decline. Stopping cART is not recommended; however, a change in drug regimen may be considered. Evidence also suggests that cognitive impairment can improve if medical and psychiatric comorbidities are aggressively treatment.

MAJOR OR MILD NEUROCOGNITIVE DISORDER DUE TO ANOTHER MEDICAL CONDITION

Definition

The DSM-5 category of NCD due to another medical condition is used for NCDs not included into one of the previous diagnostic categories. There are many potential mechanisms for the association between specific medical conditions and NCD, although current evidence is limited. Direct neurotoxicity and disruption in cholinergic neurotransmission have been postulated as likely common pathways.

According to DSM-5 criteria, in order to be classified as an NCD due to another medical condition, evidence must suggest that the NCD is the pathophysiological consequence of another medical condition. In addition, DSM-5 requires that another NCD (e.g., Alzheimer disease) should not better explain the cognitive deficits. A temporal relationship between the onset of the medical disorder and the onset of the cognitive impairment is strong evidence supportive of an association as is a reduction or stabilization of the cognitive impairment upon treatment for the causative medical condition.

Comparative Nosology

The DSM-IV-TR diagnoses of amnestic disorder due to general medical condition and dementia due to general medical condition have been replaced in DSM-5 with major/mild NCD due to another medical condition. The diagnosis is modified by the specifier of *with behavioral disturbance* or *without behavioral disturbance*. For major NCDs, the presumptive etiology is recorded. The related ICD-10 diagnosis for major NCD due to another medical condition would be *dementia in other specified diseases classified elsewhere*. For mild NCD, the related ICD-10 diagnosis would be *mild cognitive impairment*.

Epidemiology

The prevalence of major or mild NCD due to another medical condition is unknown.

Etiology

Etiology is dependent on the underlying causative medical condition. One example within this category is Wilson disease, an autosomal recessive disorder caused by an enzymatic defect in copper metabolism, resulting in toxic buildup of copper in the brain. The clinical course is characterized by NCD associated with neuropsychiatric symptoms and motor changes (e.g., tremor).

Structural brain lesions (e.g., tumor, subdural hematoma, and normal pressure hydrocephalus) can cause cognitive impairment by compression of specific neuronal tissue, disruption of blood–brain barrier, or impedance of cerebral blood flow. These lesions cause specific cognitive deficits based on location of lesion. Normal pressure hydrocephalus classically presents with cognitive impairment, urinary incontinence, and gait abnormality.

Specific metabolic abnormalities include organ dysfunction (e.g., hypothyroidism, diabetes mellitus, hypercalcemia, heart failure, and kidney disease) and vitamin deficiencies. Heart disease can result in cerebral hypoperfusion, and both hyperglycemia and hypoglycemia have been associated with disruption of cholinergic systems. Kidney disease can cause buildup of toxic metabolites.

The classic vitamin deficiency associated with cognitive impairment is Wernicke–Korsakoff syndrome, which is due to neurological damage secondary to B vitamin deficiency, and in particular, deficiency of thiamine (vitamin B_1). While frequently associated with chronic alcohol misuse, Wernicke–Korsakoff syndrome can occur in any clinical situation that is caused by or causes malnutrition (e.g., Crohn disease, chemotherapy). The estimated prevalence of Wernicke encephalopathy in the general population is 2 percent. Thiamine is involved in glucose and lipid metabolism as well as neurotransmitter metabolism. Wernicke encephalopathy (or Wernicke disease) is classically characterized by ophthalmoplegia, ataxia, and confusion, although this classic presentation is uncommon and a range of neurological symptoms are possible. Wernicke encephalopathy is considered to be reversible if adequate parenteral thiamine is provided in addition to other nutritional support. When Wernicke encephalopathy occurs in the setting of chronic alcohol misuse, a majority of patients progress to Korsakoff syndrome. Korsakoff syndrome is characterized by a dense amnesia with confabulation. Recovery from the cognitive impairment of Korsakoff syndrome is less common and less complete. Patients with Korsakoff syndrome typically demonstrate atrophy of mammillary bodies and thalamus on brain imaging.

Inflammatory and infectious states (e.g., neurosyphilis, chronic meningitis, and systemic lupus erythematosus) as well as specific medications (e.g., anticholinergics), substances (e.g., alcohol), and toxins (e.g., lead) have also been linked with cognitive impairment and dementia.

Damage or disruption of neurotransmission in areas involved in memory (e.g., hippocampus) can result in persistent amnesia. Specific infectious diseases such as herpes simplex can result in encephalitis that primarily targets the medial temporal lobes of the cortex, including the hippocampus.

AG is a 76-year-old woman with congestive heart failure with systolic dysfunction, hypertension, coronary artery disease and atrial fibrillation, osteoarthritis, hypothyroidism, and diabetes mellitus. She had been hospitalized repeatedly over the past year for acute decompensation of heart failure resulting in hypotension, dyspnea, and leg edema. She admitted to poor compliance with her medications, which was a contributing factor in the decompensation of her heart failure. She said that she could not remember to take them and that she began to have problems with her memory about 1 year ago when she was first hospitalized for congestive heart failure. She was also noncompliant with dietary and fluid restrictions. Prescribed medications included lisinopril, furosemide, metoprolol, aspirin, glipizide, acetaminophen, and levothyroxine.

On initial examination, congestive heart failure seemed well compensated without pulmonary edema and only mild peripheral edema. Heart rate was 68 and blood pressure was 95/45. Given difficulties with memory, cognitive testing was administered. AG scored 26/30 on the Montreal Cognitive Assessment, with difficulty on serial sevens. Attention was otherwise intact but clock drawing indicated poor planning and inability to indicate the correct time. She also displayed difficulty with problem solving and abstract reasoning. Brain imaging revealed mild small vessel disease but no infarcts. Thyroid-stimulating hormone level was normal and no other lab abnormalities were noted other than a creatinine of 1.5 and a glycosylated hemoglobin of 7.9.

Two weeks later, she again was admitted to the hospital with hypoxia and pulmonary edema. She was hypotensive and required pressor support for the management of her decompensated heart failure. During this time, AG experienced an abrupt change in mental status with poor attention, agitation, and visual hallucinations. At the time of discharge to subacute rehabilitation, these symptoms had improved but not resolved completely.

AG was given the diagnosis of major NCD due to congestive heart failure with cerebral hypoperfusion without behavioral disturbance. At the time of her hospitalization, she was also diagnosed with delirium due to congestive heart failure with hypotension and hypoxia.

Diagnosis and Clinical Features

According to DSM-5, major or mild NCD due to another medical condition is defined by the development of new cognitive impairment independent of a delirium or another NCD. A temporal and causative link between the cognitive impairment and the medical condition or substance must exist (Table 10.5–6).

Laboratory Evaluation

A thorough workup should be undertaken to identify the underlying causative medical condition as well as to rule out any of the more common NCDs (e.g., cerebrovascular disease). Given the extent of medical conditions that can cause cognitive impairment and dementia, other testing should be tailored to the history and the presenting symptoms. For example, an ESR can be performed if an inflammatory disorder is suspected, although ESR is a very nonspecific test. Copper and

ceruloplasmin levels can be measured if Wilson disease is suspected. A detailed history of the proband as well as the family is important to obtain to guide this workup. Structural brain imaging should be performed and in specific cases, functional brain imaging may help differentiate neurocognitive syndromes. Lumbar puncture is not performed routinely, but advances in the understanding of specific CSF markers may help in the differential diagnosis, especially if CJD, neurosyphilis, or other infectious etiology is suspected. Toxicology tests may be appropriate in select cases. Neuropsychological testing can aid in the diagnosis.

Differential Diagnosis

Delirium and all other NCDs should be included in the differential diagnosis.

Course and Prognosis

Onset may be sudden or slowly progressive. Course is variable depending on etiology. Recovery may occur over both the short- and long-term depending on the severity of initial injury and the persistence of the underlying medical condition.

Treatment

Treatment of the underlying medical condition is the most important therapy. Prevention of cognitive impairment is possible in individuals at risk of thiamine deficiency (e.g., secondary to long-standing alcohol misuse or severe malnutrition) through the administration of vitamin B1 as this will prevent the development of Wernicke–Korsakoff syndrome. Memory retraining or use of memory aids may have some benefit.

SUBSTANCE/MEDICATION-INDUCED MILD/MAJOR NEUROCOGNITIVE DISORDER

Definition

Cognitive impairment is commonly seen during intoxication and the acute withdrawal phase from substances and some medication. However, the categories of major or mild substance/medication-induced NCD only applies when the cognitive impairment persists beyond what is normally seen with use. DSM-5 also clarifies that the suspected substance/medication must have a known link to persistent cognitive impairment and that the onset of the cognitive impairment must be temporally associated with exposure to the substance/medication.

Comparative Nosology

DSM-IV-TR offered specific substance-induced persisting dementia (e.g., inhalant-induced persisting dementia) and specific substance-induced

persisting amnestic disorders (e.g., alcohol-induced persisting amnestic disorder) as diagnostic choices. These disorders have been replaced in DSM-5 with the general *major/mild substance/medication-induced NCD*, to which the clinician would add information regarding the substance/medication involved (e.g., alcohol-induced major NCD). Additionally, the diagnosis can be modified by a specifier to indicate persistence of the cognitive impairment. The related ICD-10 diagnoses are organized akin to the DSM-IV-TR; that is, for each substance of abuse there is a cognitive diagnosis available.

Epidemiology

The prevalence of substance/medical-induced NCD is unknown. Risk factors for developing this disorder include older age, longer period of exposure to the substance or medication, poor nutrition, liver disease, and cerebrovascular disease.

Etiology

Persistent cognitive impairment likely reflects slow recovery of underlying brain function. However, for some patients, the NCD is irreversible. For example, chronic alcohol abuse resulting in thiamine deficiency can cause Korsakoff syndrome that is irreversible as it is the result of permanent damage of the mammillary bodies and thalamus.

Diagnosis and Clinical Features. Chronic users of sedative, hypnotic, or anxiolytic drugs can develop an NCD, which is typically mild and which consists of primarily of impairment in memory. Studies demonstrate that chronic users continue to demonstrate impairment in multiple cognitive domains, especially verbal memory and processing speed, even after 6 months of abstinence.

Chronic users of cocaine, methamphetamine, opioids, and phencyclidine are also at risk of developing a mild persistent NCD. For methamphetamines, NCD primarily involves impairment in learning and memory as well as executive functions. Some evidence suggests that prolonged abstinence (i.e., 14 months) from methamphetamine is associated with recovery in some but not all cognitive domains. Use of methamphetamine and of cocaine is associated with a risk of stroke, which can result in additional persistent cognitive impairment.

Chronic inhalation of volatile hydrocarbons, for example, inhalant abuse, can produce impairment in memory, attention, visuospatial, and executive functions as well as motor abilities. Studies have demonstrated that abstinent individuals can recover over a 2-year period. However, permanent disability has been seen in those individuals with severe abuse or those who had abused leaded gasoline products and developed

concomitant lead encephalopathy with cerebellar damage.

Chronic alcohol use can lead to confusion (Wernicke encephalopathy) and progress to a dense amnestic state (Korsakoff syndrome). (See section "Major or Mild Neurocognitive Disorder due to Another Medical Condition/Etiology" for further discussion of Wernicke–Korsakoff syndrome.)

Other substances that can cause an amnestic syndrome are cannabis, lead, carbon monoxide, and some insecticides.

Other medications associated with cognitive impairment include steroids, chemotherapy agents, some nucleoside reverse transcriptase inhibitors (NRTIs), L-Dopa and dopamine agonists, select antiepilepsy drugs (e.g., topiramate), and medications with anticholinergic properties. Cognitive deficits induced by medications typically resolve upon discontinuation of the offending drug. Corticosteroid-induced cognitive deficits, whether during acute or chronic treatment, typically affect declarative (verbal) as well as working memory, suggesting that the hippocampus and dorsolateral prefrontal cortex are sensitive to the adverse effects of steroids.

Pathology and Laboratory Examination

Other than toxicology studies to help in identifying the causative substance/medication, there are no other pathognomonic lab findings or imaging findings for major or mild substance/medication-induced NCD. A thorough workup should be undertaken in order to identify any underlying causative medical condition as well as to rule out any of the more common NCDs (e.g., cerebrovascular disease).

Differential Diagnosis

Delirium and all other NCDs should be included in the differential diagnosis. It is important to remember that persons who misuse substances are more prone to other events that are also associated with cognitive impairment such as TBI and certain infections, for example, HIV. They may also be more likely to have comorbid psychiatric diagnoses, such as PTSD and mood disorders, which can include cognitive dysfunction. In addition, some substances of abuse, such as cocaine and methamphetamine, are associated with an increased risk of stroke, which can also produce cognitive impairment.

Course and Prognosis

For many substances and medications, abstinence leads to improved cognitive function. However, the time course of recovery is highly variable and is dependent on not only the substance/medication used but also the severity and duration of use as well as the presence of other risk factors for

cognitive impairment.

Treatment

For NCD due to substances and medications, the treatment of choice is cessation of that substance or medication. Unfortunately, this may not be possible for some medications. Nutritional support and other medical support may be needed in both the short and long term. For example, prevention of cognitive impairment is possible in individuals at risk for thiamine deficiency (e.g., secondary to long-standing alcohol misuse) through the administration of vitamin B_1 as this will prevent the development of Wernicke–Korsakoff syndrome. In the case of substance misuse, a comprehensive treatment program is indicated.

MOOD DISORDERS DUE TO ANOTHER MEDICAL CONDITION AND SUBSTANCE/MEDICATION-INDUCED MOOD DISORDERS

Mood disorders, particularly depression, frequently accompany a range of medical problems. In addition to emotional suffering, depression leads to disability over and above that caused by the medical illness. The relationship between the medical illness and the comorbid mood disorder is frequently unclear. The mood disturbance can be the result of the underlying pathophysiology of the medical condition. However, the mood disturbance can also be a psychological reaction to the medical illness and its associated disability. Furthermore, two conditions (medical and psychiatric) can be the sequelae of another factor.

Definition

In the classification mood disorder due to another medical condition, the mood symptoms are felt to be the direct physiological product of a general medical condition. The mood disturbance may be depressive, manic, hypomanic, or mixed in presentation but does not need to meet the full criteria for a major depressive, manic, mixed, or hypomanic episode. However, to meet the criteria for mood disorder due to another medical condition, the mood symptoms must not be better accounted for by a primary mental disorder. For example, an adjustment disorder with depressed mood would be the more appropriate diagnosis for mild depressive symptoms that are a psychological response to having a medical condition. The diagnosis of mood disorder due to another medical condition is not made if mood symptoms occur during the course of delirium, a clinical syndrome commonly accompanied by transitory shifts in mood.

A diagnosis of substance/medication-induced mood disorder relies on the appearance of mood symptoms after exposure to a medication or during or soon after intoxication or withdrawal from a substance of abuse

that is known to potentially impact mood. These mood symptoms must be in excess of those usually associated with the substance/medication or must be severe enough to warrant independent clinical attention.

History and Comparative Nosology

The corresponding term used in DSM-III-R was organic mood syndrome. This term was somewhat more restrictive as the mood symptoms needed to meet the criteria for either a manic or major depressive episode. In DSM-IV, organic mood syndrome was bifurcated into mood disorder due to a general medical condition and substance-induced mood disorder. Apart from the etiological difference, the diagnostic criteria for the disorders were nearly identical. In DSM-5, this category has been separated into two categories: *bipolar and related disorders* and *depressive disorders*. Reflecting this general separation, the two categories now available for mood disorders secondary to a medical condition include bipolar and related disorder due to another medical condition with the specifiers of *with manic features, with manic- or hypomanic-like episode* and *with mixed features* and depressive disorder due to another medical condition with the specifiers of *with depressive features, with major depressive-like episode* and *with mixed features*. Similarly, two categories are now available for substance/medication-induced mood disorders: *substance/medication-induced bipolar and related disorder* and *substance/medication-induced depressive disorder*. For both disorders, the clinician should specific whether *with onset during intoxication* or *with onset during withdrawal*.

ICD-10 offers the diagnosis *organic mood (affective) disorder*s with specifiers *for manic, bipolar, depressive, or mixed affective symptoms* due to a medical condition or nonpsychoactive substance. For psychoactive substances, the clinician initially notes the presence of abuse of a specific substance and then modifies that diagnosis with the specifier *with substance-induced mood disorder*.

Epidemiology

In general, mood disorders due to another medical condition more commonly present with depressive features and less frequently with manic features. Mood disorders due to another medical condition with depressive features appear to affect men and women equally, in contrast to major depressive disorder, which predominates in women.

Etiology

Numerous medical illnesses are capable of inducing a mood syndrome (Table 10.5–2), including neurological disorders, metabolic and endocrine conditions, hematological and oncological disorders, and infectious and

inflammatory states.

As in primary mood disorders, the mechanism by which another medical condition and substance use lead to mood syndromes likely involves the noradrenergic, dopaminergic, serotonergic, and cholinergic neurotransmission. Substances (e.g., reserpine [Serpalan]) and medical conditions (e.g., hypoxia) that deplete noradrenergic function may induce depression, whereas agents (e.g., amphetamines) that stimulate noradrenergic activity may cause mania. Any disorder or substance that affects the hypothalamic–pituitary–adrenal axis can lead to mood symptoms. Damage to the brain can lead to depression, be it through demyelination as in multiple sclerosis or in vitamin B_{12} deficiency, through head trauma, or through cerebrovascular disease; however, the precise mechanism or neuroanatomical correlates that lead to depression are unknown.

It should be noted that the strength of evidence linking a specific medical disorders to a mood syndrome is quite variable. Further complicating the diagnosis is that the mood symptoms may have multiple etiologies, for example, a physiological sequela of the medical illness, an effect of the medication used to treat the medical illness, and a psychological reaction due to being ill.

An example of a neurological disease strongly associated with depression is Parkinson disease. Parkinson disease is a relatively common degenerative neurological disorder (occurring in 20 of 100,000 persons) that is caused by gradual loss of dopaminergic neurons in the substantia nigra pars compacta. Forty to 70 percent of patients with Parkinson disease suffer from a mood disorder, most commonly depression. Depression in Parkinson disease is more common in women and in left-sided predominant disease and is often atypical in nature. Although a reaction to the stress caused by this disabling illness must be considered, evidence supports that the depressive illness is due to the actual disease process itself. Specifically, depression can be the first or among the first symptoms presenting in Parkinson disease. Second, the depression in Parkinson disease is poorly responsive to typical antidepressant therapy. Mood fluctuations are often noted in association with changes in plasma dopamine levels. It is hypothesized that the reduced levels of monoamines (dopamine, noradrenaline, 5-hydroxytryptamin) seen in Parkinson disease lead to the depression.

Another example is mood disorder due to stroke (e.g., poststroke depression). Major depression after stroke is extremely common (up to 35 percent of cases); furthermore, up to 70 percent of stroke patients demonstrate significant depressive features. Onset of depression is usually between 3 and 24 months following the stroke. The mechanism linking stroke with depression has been much studied. However, no one clear etiology has been identified. Possible biological mechanisms include the

direct physiological effects of the brain injury at a specific brain location. Specifically, left frontal stroke lesions have been associated with an increased incidence of depression after stroke. Other potential biological mechanisms linking stroke and depression include increased CNS and systemic inflammation as well as chronic elevation of hypothalamic–pituitary–adrenal axis function. Of course, while physiological etiologies have been postulated, it is reasonable that depression after stroke is a psychological reaction to stroke-related disability.

Mania after stroke has been reported but is much less common. Again, the exact mechanism linking stroke with subsequent mania has not been firmly established, although stroke lesions disrupting subcortical pathways of the right hemisphere are strongly associated with poststroke mania.

Notable endocrine causes of mood disorders include hypothyroidism causing depression and hyperthyroidism causing mania. In addition, in the elderly, apathetic hyperthyroidism mimics depression. Appropriate treatment that leads to a euthyroid state usually leads to the resolution of the mood symptoms.

Vitamin B_{12} deficiency is also considered capable of inducing a mood disorder. However, this association is controversial.

Pancreatic cancer is strongly associated with depression. Frequently, the depression is the first symptoms of this illness. Nineteen to 75 percent of patients with pancreatic cancer have significant depressive symptoms. The mechanism linking pancreatic cancer to depression is poorly understood. Several theories have been proposed including a hypothesized paraneoplastic syndrome in which the availability of serotonin is reduced and/or serotonin receptors are antagonized.

Among infectious causes, HIV is strongly associated with depression. The estimated point prevalence of major depression is 9.4 percent, which represents an approximately two times greater prevalence of major depression than in HIV negative controls. Clearly depression in this context has multiple causes, including the effect of the HIV infection in the brain, psychological reaction to the illness, and the potential side effects from highly active antiretroviral treatment regimens. Manic symptoms may develop in the context of HIV or as a result of treatment with antiretroviral agents.

Table 10.5–12.
Substance-Induced Psychiatric Disorders Listed by Substance and State

		Psychotic Disorders	Bipolar Disorders	Depressive Disorders	Anxiety Disorders	Obsessive-Compulsive Disorders	Sleep Disorders
Intoxication State	Alcohol	+	+	+	+		+
	Caffeine				+		+
	Cannabis	+			+		+

	Hallucinogens	+	+	+	+		
	Inhalants	+		+	+		
	Opioids			+			+
	Sedative, hypnotics, or anxiolytics	+	+	+			+
	Stimulants (including cocaine)	+	+	+	+	+	+
	Tobacco						
Withdrawal State	Alcohol	+	+	+	+		+
	Caffeine						+
	Cannabis						+
	Hallucinogens						
	Inhalants						
	Opioids			+	+		+
	Sedative, hypnotics, or anxiolytics	+	+	+			+
	Stimulants (including cocaine)		+	+	+	+	+
	Tobacco						+

Adapted from American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Text rev. Washington, DC: American Psychiatric Association; 2013.

As reviewed in the earlier edition, prescribed substances capable of inducing mood syndromes include agents from the following categories: stimulant and sympathomimetic (e.g., methylphenidate and theophylline [Asmalix]); corticosteroid; antiparkinsonian (e.g., L-Dopa and bromocriptine [Parlodel]); antidepressant (inducing manic symptoms); immune modulators (e.g., cyclosporine [Neoral] and interferon-α [Roferon]); antihypertensive (e.g., β-blockers and methyldopa [Aldomet]); cancer chemotherapy (e.g., vincristine [Oncovin], vinblastine [Velban], and procarbazine [Matulane]); oral contraceptives; CNS depressant (e.g., benzodiazepines and barbiturates); heavy metals; and toxins (e.g., paint and carbon monoxide).

Substances of abuse that may produce mood symptoms are listed in Table 10.5–12.

Diagnosis and Clinical Features

As reviewed in the earlier edition, patients with depression may experience psychological symptoms (e.g., sad mood, lack of pleasure or interest in usual activities, tearfulness, concentration disturbance, and suicidal ideation) as well as somatic symptoms (e.g., fatigue, sleep disturbance, and

appetite disturbance). Diagnosis in the medically ill may be confounded by the presence of somatic symptoms that may be purely related to the underlying medical illness. In an effort to overcome the underdiagnosis of depression in the medically ill, most practitioners favor counting all somatic symptoms toward the diagnosis of a mood disorder regardless of likely etiology.

Individuals with mania present with irritable or euphoric mood, heightened energy, reduced need for sleep, and accelerated rate of thoughts. Occasionally, manic patients may engage in reckless and inappropriate behavior with little consideration of the consequences; this behavior may interfere with their medical care. For example, nursing staff may refuse to work with a manic patient who has made unwanted sexual advances.

The reader is referred to Table 10.5–6 for a discussion of general principles regarding diagnosis.

Laboratory Evaluation

A targeted workup is required when a mood disorder due another medical condition is being considered as part of the differential diagnosis. If possible, tests should be selected to rule in specific diagnoses suggested by the patient's somatic symptoms (if present). Commonly ordered tests include a complete blood count and chemistries, thyroid function tests, and serum vitamin B_{12} level. Urine toxicology is appropriate if abuse of substances is suspected. In selected cases, added laboratory studies, such as ESR, rapid plasma reagin, urinalysis, or lumbar puncture may be indicated. EEG or structural brain imaging, or both, should be considered in the absence of an overt medical or substance cause, the presence of cognitive impairment, or the report or presence of neurological symptoms (e.g., headache, motor or sensory loss, visual disturbance, or lethargy).

Differential Diagnosis

Diagnosis in the medically ill may be confounded by the presence of somatic symptoms related purely to medical illness, not to depression. However, it is frequently impossible to rule out depression as either a cause of the somatic symptom or as something exacerbating the somatic illness. Therefore, most practitioners favor including somatic symptoms, such as sleep disturbance, anorexia, and fatigue, in identifying mood syndromes in the context of medical illness.

To best inform treatment decisions, diagnostic accuracy is important. In particular, care must be given to separate a mood disorder due to another medical condition from a substance/medication-induced mood disorder secondary to the medications used to treat the primary medical disorder. Identifying the most likely etiology requires taking a thorough history that

carefully traces the time course of symptoms and the response to correction of suspect medical conditions or discontinuation of substances, and, occasionally, by urine or blood toxicology tests. Mood fluctuations that occur in the context of a delirium are not coded as a mood disorder due to another medical condition. Finally, syndromes that may lead to depressed mood but that do so through primary psychological and not physiological means, such as pain syndromes, are not coded as mood disorder due to a another medical condition; in these cases, the more appropriate diagnosis may be a primary mood disorder or an adjustment disorder.

Course and Prognosis

The course of mood disorder due to another medical condition is mainly dependent on the course of the underlying medical illness as well as the extent of concurrent psychiatric intervention. Similar considerations apply to substance/medication-induced mood disorder. Prognosis for mood symptoms is best when etiological medical illnesses or medications are most susceptible to correction (e.g., treatment of hypothyroidism and cessation of alcohol use). Patients should receive formal psychiatric treatment whenever it is not possible to alter the underlying medical illness (as in poststroke depression), the mood symptoms do not resolve after the underlying medical illness is successfully treated, or the severity of symptoms indicate a need for treatment.

Treatment

Clearly, an important first step in the treatment of the mood disorder is adequate management of the underlying medical condition. If substance abuse is confirmed, a comprehensive rehabilitation program should be offered.

Pharmacotherapy should be designed to address mood symptoms while minimizing adverse interactions with other concurrently prescribed medications. SSRIs with little impact on the cytochrome P450 enzyme system, such as sertraline, citalopram [Celexa], and escitalopram [Lexapro], are preferred. Bupropion [Wellbutrin] has stimulant properties that may be useful in the depressed, debilitated, or medically ill individual, but the patient's seizure risk must be considered prior to initiating this drug. Mirtazapine [Remeron] may be especially useful in stimulating appetite in the anorectic or debilitated individual with depression. Similar benefit may be derived from psychostimulants, such as methylphenidate.

Medically ill individuals often are exquisitely sensitive to medication side effects and toxicities. Therefore, tricyclic antidepressants, which typically have a higher incidence of orthostatic hypotensive, anticholinergic, and cardiac conduction effects, are rarely used in the medically ill.

In manic or mixed mood states, divalproex [Depakote] is preferred to lithium [Eskalith]; the latter has a low therapeutic index, with major toxicity considerations in medically ill patients at risk for dehydration or electrolyte imbalance. Other anticonvulsants (e.g., carbamazepine [Tegretol]) and the atypical antipsychotics can also be considered for mood stabilization. Drug interactions may complicate the use of anticonvulsants, particularly divalproex and carbamazepine.

In the anxious, depressed individual, caution must be exercised in the use of benzodiazepines, as these agents can exacerbate cognitive impairment commonly present in the medically ill and elderly. If needed, agents with short half-lives (e.g., lorazepam [Ativan] and alprazolam [Xanax]) are preferred. In individuals with liver disease, agents that only require conjugation with glucuronide for elimination (e.g., lorazepam, oxazepam [Serax], and temazepam [Restoril]) are preferred as liver disease does not generally affect this process. In the elderly and in individuals with hepatic or renal impairment, psychotropic medications should be initiated at low dosages. If mood symptoms persist in the absence of serious toxicity or side effects, medications should be titrated until symptom relief is achieved. Interestingly, medically ill individuals often require psychopharmacological doses similar to those used in medically healthy individuals.

ECT, useful in severe depression (especially if associated with intense suicidal ideation), mania, and catatonia, has the advantage of being safe, highly effective, and arguably more rapid in onset than pharmacological modalities, a point of considerable importance in the debilitated, bedridden, anorectic patient for whom a feeding tube otherwise may be indicated. ECT may also be useful in patients whose gastrointestinal tract is nonfunctional, thereby preventing the administration of antidepressant medications. Relative contraindications to ECT include conditions associated with increased intracranial pressure (e.g., brain tumor) or cardiac conditions associated with high risk of hemodynamic compromise (e.g., malignant arrhythmia).

Psychotherapies in the medically ill with depressed mood are geared to assisting with life transitions and loss of autonomy. When pain is a factor in depressing mood, this should be aggressively treated.

A 45-year-old business executive was admitted to the hospital after attempting suicide by acetaminophen overdose. The patient had been experiencing depressed mood, anhedonia, anergia, and hypersomnia for 6 months. She presented to her primary care physician with these complaints about 4 weeks prior to the attempt. At that time, she was diagnosed with hypothyroidism, started on appropriate therapy, and told that the depressive symptoms would resolve. The patient reported that she felt increasingly hopeless when her mood did not improve with the initiation of thyroid replacement therapy. This hopelessness led her to attempt suicide in order to end her suffering.

In this case, provisional diagnosis is hypothyroidism and depressive disorder due to hypothyroidism, with major depressive-like episode. This diagnosis would be supported if the

mood disturbance resolves after the thyroid abnormality is corrected. If it fails to do so, major depressive episode unrelated to hypothyroidism would also merit diagnostic consideration, especially if the individual had experienced such episodes earlier in life or had a family history of depression. This case raises the question of when to initiate psychiatric treatment when a mood disorder is felt to be secondary to an underlying medical condition. At the very least, a patient with a mood disorder needs to be monitored for worsening psychiatric symptoms as well as for the emergence of suicidal ideation while the underlying medical disorder is addressed. While antidepressant treatment would be appropriate to treat mood symptoms that fail to resolve after the underlying medical illness is successfully treated, initiating psychiatric treatment earlier should be considered especially if the mood symptoms are severe or if the time it will take to correct the underlying medical illness is lengthy.

PSYCHOTIC DISORDER DUE TO ANOTHER MEDICAL CONDITION AND SUBSTANCE/MEDICATION-INDUCED PSYCHOTIC DISORDER

Psychosis is a departure from reality testing. Psychotic experiences may either be delusions (i.e., fixed, false beliefs) or hallucinations (i.e., perception of stimuli not actually present). Psychotic symptoms can be induced by medical conditions (psychotic disorder due to another medical condition) or by substances (substance/medication-induced psychotic disorder). Psychotic symptoms also may be seen in primary psychotic disorders as well as in mood disorders, cognitive disorders, and, occasionally, in severe personality disorders.

Definition

DSM-5 states that, in psychotic disorder due to another medical condition, hallucinations, delusions, or both are present and must be judged to be due to the physiological effects of a medical condition. The disturbance may not occur exclusively within the context of delirium and must not be better accounted for by a primary mental disorder.

Hallucinations may occur in any modality: visual, olfactory, gustatory, tactile, or auditory, or a combination of these. Delusions may involve simple or elaborate themes. Of note, a fixed false belief does not need to be bizarre in order to qualify as a delusion.

History

The term psychotic was used broadly in the past, often reserved for the most severe psychiatric conditions with the greatest impact on functioning. Thus, chronically disabling depressive illness and cognitive disorders (e.g., Korsakoff psychosis) at one time were considered psychotic illnesses. DSM-5 adheres to a narrower definition of psychosis that includes only hallucinations and delusions.

Multiple theories, including those with religious, cultural, psychological, and biological aspects, have been advanced to explain the origin of psychosis. Over the course of centuries, demonic possession, moral decay, environmental stress, poor mothering, and, more recently, derangements

of dopamine and serotonin metabolism have been offered.

Comparative Nosology

DSM-III-R offered two diagnoses: organic delusional syndrome and organic hallucinosis. Each diagnosis included medically induced and substance-induced psychoses. In DSM-IV-TR, psychoses were described as due to a general medical condition or substance induced. In DSM-5, the conditions are known as psychotic disorder due to another medical condition or substance/medication-induced psychotic disorder. For psychoses stemming from general medical conditions and nonpsychoactive substances, ICD-10 provides coding for organic delusional (schizophrenia-like) disorder and organic hallucinosis. Also available is the specifier *with substance-induced psychotic disorder* that is used to modify a primary substance use diagnosis.

Epidemiology

The prevalence of this disorder is unclear and is dependent on the underlying medical condition. It is estimated that up to 40 percent of patients with temporal lobe epilepsy experience psychosis. Visual hallucinations are common in later stage Parkinson disease and may occur in up to 40 percent of patients. Psychosis occurring in the earlier stages of Parkinson disease is less common, but when it does occur, is likely due to the medications used to treat the Parkinson disease including anticholinergic medications, L-Dopa, and dopamine agonists.

Etiology

Diseases categories that are associated with the diagnosis of psychotic disorder due to another medical condition include neurological (e.g., Parkinson disease, Huntington disease, multiple sclerosis, visual and auditory sensory loss, epilepsy, cerebrovascular accident, and head trauma); metabolic and endocrine (e.g., hypo- and hyperthyroidism, hypo- and hyperadrenocorticism, hypo- and hyperglycemia, hypoxia, hypercarbia, renal failure, hepatic failure, and Wilson disease); infectious and inflammatory (e.g., systemic lupus erythematosus and HIV); and nutritional deficiency (e.g., vitamin B_{12} and thiamine) (Table 10.5–2).

Medication categories capable of inducing psychosis include opioid (especially meperidine [Demerol]); anticholinergic (e.g., benztropine [Cogentin] and diphenhydramine); cardiovascular (e.g., digoxin, procainamide [Promine], methyldopa); cancer chemotherapy (e.g., procarbazine); corticosteroid (e.g., prednisone and dexamethasone [Decadron]); immunosuppressant (e.g., cyclosporine and tacrolimus [Prograf]); antiparkinsonian (e.g., L-Dopa and bromocriptine); antitubercular (e.g., isoniazid [Laniazid]); sympathomimetic (e.g.,

theophylline and phenylephrine [Rhinall]); sedative-hypnotic, anxiolytic, and disulfiram [Antabuse].

Stimulants (e.g., amphetamine and cocaine), hallucinogens, PCP, inhalants, cannabis, opioids, and alcohol are all capable of inducing psychosis during intoxication, whereas alcohol can cause psychosis during withdrawal (Table 10.5–12). Toxins such as heavy metals, nerve gases, organophosphate insecticides, carbon monoxide, and volatile substances such as gasoline and paint can induce psychosis.

As in primary psychoses, a presumed pathophysiological mechanism underlying is alterations in dopamine and other neurotransmitters. This link is underscored by the frequent association between psychosis and medical conditions (e.g., Parkinson disease) and substances (e.g., amphetamine and L-Dopa) that affect dopamine.

Diagnosis and Clinical Features

Two DSM-5 subtypes exist for psychotic disorder due to another medical condition *with delusions*, to be used if the predominant psychotic symptoms are delusions, and *with hallucinations*, to be used if hallucinations of any form comprise the primary psychotic symptom. The reader is referred to Table 10.5–6 for general principles regarding diagnosis. Of note, this diagnosis is not applied if the psychosis is occurring exclusively during a delirium.

Laboratory Examination

A targeted workup is required when a psychotic disorder due to another medical condition is being considered as part of the differential diagnosis. If possible, tests should be selected to rule in specific diagnoses suggested by the patient's somatic symptoms (if present). Laboratory tests to be considered include a complete blood count and chemistries, thyroid function tests, vitamin B_{12} level, RPR, and urinalysis. Urine toxicology usually is indicated. Structural brain imaging, EEG, and lumbar puncture should be considered for new-onset psychotic symptoms, lack of an overt etiological medical condition or substance, the presence of cognitive impairment, or the report or presence of neurological symptoms (e.g., headache, motor or sensory loss, visual disturbance, or lethargy). Specialized studies may be required to rule out unique disease (e.g., serum ceruloplasmin in suspected Wilson disease).

Differential Diagnosis

Primary psychotic disorders, such as schizophrenia, and primary mood disorders with psychotic features may present with symptoms identical or similar to psychotic disorder due to another medical condition or substance/medication-induced psychotic disorder; however, in primary

disorders, no medical or substance cause is identifiable, despite laboratory workup. The chronology of symptoms, response to removal of suspect substance/medication or alleviation of medical illnesses, and toxicology results may all be helpful in making the diagnosis.

First-rank symptoms of psychosis (i.e., thought broadcasting, thought withdrawal, and hallucinations of voices commenting or arguing) are nonspecific and do not clarify the diagnosis. However, the presence of olfactory, tactile, or visual hallucinations should increase suspicion of a psychotic disorder induced by an underlying medical disorder or a substance/medication as they are more commonly found in secondary psychotic disorders. Auditory hallucinations are nonspecific as they are found in both primary and secondary psychotic disorders. Other specific psychotic symptoms that are strongly associated with a specific etiology include formication (a perception of bugs crawling under the skin) secondary to stimulant (e.g., amphetamine and cocaine) intoxication; olfactory hallucinations and religious delusions due to temporal lobe epilepsy; and visual hallucinations due to any lesion (e.g., tumor, stroke) of the occipital lobe. In addition, anosognosia (or failure to attend to the left side of one's own body) can reach delusional proportions (e.g., denial of ownership of one's own left arm) and is due to right parietal lobe lesions.

The presence of psychotic symptoms in the context of an NCD (e.g., dementia) would be indicated by adding the specifier *with behavioral disturbance* to the diagnosis of the underlying cognitive disorder (e.g., major NCD due to Alzheimer disease, with behavioral disturbance [delusions]). Psychotic disorder due to another medical condition is not made if the psychosis is attributable to a delirium.

Course and Prognosis

The course of the underlying medical illness or substance/medication use influences the course of psychosis with several notable exceptions. Psychosis due to certain medications (e.g., immunosuppressants) may gradually subside even when use of those medications is continued. Minimizing dosages of such medications consistent with therapeutic efficacy often facilitates resolution of psychosis. Certain degenerative brain disorders may be characterized by episodic lapses into psychosis. If abuse of substances persists over a lengthy period, psychosis (e.g., hallucinations from alcohol) may fail to remit even during extended intervals of abstinence.

Treatment

Aggressive efforts should be made to eradicate psychosis at the earliest stages, as symptoms tend to become more elaborate and ingrained over time if untreated. The first steps in treatment are to remove or minimize

exposure to the offending substance/medication (if possible) and actively treat the underlying medical condition. If substance abuse is confirmed, inpatient or outpatient treatment should be offered. In addition, psychopharmacological and psychotherapeutic options may be considered.

Antipsychotic agents are the primary pharmacological intervention. Atypical agents have a more favorable side effect profile than do typical agents and so are preferred. Medically ill individuals are especially susceptible to antipsychotic medication side effects (e.g., extrapyramidal signs [EPS], sedation, gait disturbance, and tardive dyskinesia). In patients with Parkinson disease, first-line antipsychotic medication choices should include atypical agents with low incidence of EPS, such as quetiapine, olanzapine, or ziprasidone [Geodon]. Aripiprazole [Abilify] may represent another option. The atypical agent clozapine [Clozaril] has low incidence of EPS but is rarely selected owing to risk of agranulocytosis, which requires frequent monitoring.

For individuals who cannot or will not take oral antipsychotics, haloperidol, olanzapine, and ziprasidone are available by IM route. Long-acting IM depot formulations of haloperidol, fluphenazine [Prolixin], and risperidone may be useful for chronically medication-noncompliant individuals. Haloperidol may be administered intravenous, although this should be avoided in patients with prolonged QTc interval on electrocardiogram (ECG) owing to risk of the malignant arrhythmia torsades de pointes. QTc prolongation also must be considered in using the atypical agents, especially ziprasidone, particularly when used in combination with other medications (e.g., quinidine [Cardioquin], dofetilide [Tikosyn], sotalol [Betapace], and moxifloxacin [Avelox]) that may prolong the QTc interval or in medical states associated with prolongation of the QTc interval (e.g., hypomagnesemia, hypokalemia, and congenital QT interval prolongation).

Antipsychotic agents should be dosed low when initiated, with gradual upward titration until symptoms are relieved. Anticholinergic agents (e.g., benztropine) as prophylaxis against EPS should be avoided as medically ill individuals often are sensitive to the cognition-impairing aspects of such agents. If EPS occur, antipsychotic dose reduction or shift to an agent with a more favorable EPS profile (e.g., risperidone to quetiapine) may be considered before initiating an anticholinergic agent.

Psychotherapeutic interventions are primarily supportive and psychoeducational in nature. Appropriate goals for psychotherapy include coping strategies to handle the social and occupational limitations imposed by psychosis, education regarding early symptom identification, and education regarding the importance of medication compliance. Psychosis can alter behavior to the point that patients are unsafe. A safety plan should always be in place and explained to both the patient and their family.

ANXIETY DISORDER DUE TO ANOTHER MEDICAL CONDITION AND SUBSTANCE/MEDICATION-INDUCED ANXIETY DISORDER

Anxiety is a ubiquitous symptom. In addition to being seen in primary or secondary anxiety disorders, it can also be present in other psychiatric disorders including mood and psychotic disorders. New-onset anxiety complaints always demand a search for underlying physiological factors.

Definition

In an anxiety disorder due to another medical condition, the individual experiences anxiety that causes clinically significant distress or impairment in functioning. This anxiety must be caused by the direct physiological effect of a general medical condition and does not solely represent a psychological reaction to having a medical illness. In the diagnosis substance/medication-induced anxiety disorder, the anxiety symptoms are the product of intoxication or withdrawal from prescribed medication or from a nonprescribed substance, typically a drug of abuse and are in excess to what is typically expected during intoxication and withdrawal.

Comparative Nosology

DSM-III-R offered the diagnosis of organic anxiety syndrome to describe panic attacks or generalized anxiety stemming physiologically from either medical conditions or substances. The comparable DSM-IV-TR anxiety diagnoses were either secondary to a general medical condition or substance induced. In addition, DSM-IV-TR broadened the diagnosis of anxiety disorder due to a general medical condition to include obsessive-compulsive symptoms and, in the case of substance-induced anxiety, phobic symptoms. In contrast, in DSM-5, obsessive-compulsive disorder (OCD), including that due to another medical disorder or substance induced, has been separated from the anxiety disorders based on evidence supporting a distinct pathophysiology. In DSM-5, two secondary anxiety disorders are offered: *anxiety disorder due to another medical condition* and *substance/medication-induced anxiety disorder*.

ICD-10 offers the diagnosis of *organic anxiety disorder*, with generalized anxiety or panic aspects, for anxiety syndromes due to medical illnesses and nonpsychoactive substances. For psychoactive substances, the clinician may use specifier *with substance-induced anxiety disorder* to modify the primary substance use diagnosis.

Epidemiology

The prevalence of this disorder is unclear and is dependent on the underlying medical condition. The prevalence of panic and generalized anxiety are especially high in neurological, endocrine, and cardiology patients, although this finding does not necessarily prove a physiological

link. Approximately one-third of patients with hypothyroidism and two-thirds of patients with hyperthyroidism may experience anxiety symptoms.

Etiology

Medical disorders that may induce anxiety include cardiopulmonary (e.g., arrhythmias, mitral valve prolapse, pulmonary embolism, chronic obstructive pulmonary disease, asthma, and congestive heart failure); neurological (e.g., seizure disorder, head injury, stroke, vestibular disease, Parkinson disease, Huntington disease); metabolic and endocrine (e.g., hyper- or hypothyroidism, hyperparathyroidism, hypoglycemia, adrenal dysfunction, pheochromocytoma, and vitamin B_{12} deficiency); and inflammatory and infectious (e.g., systemic lupus erythematosus and HIV).

Medications capable of inducing anxiety include stimulants and sympathomimetic agents (e.g., theophylline, pseudoephedrine [Sudafed], and albuterol [AccuNeb]); antiparkinsonian and cardiovascular medications, antidepressants (especially SSRIs); anxiolytic agents (i.e., as a withdrawal symptom); corticosteroids; insulins; thyroid preparations; and caffeine preparations.

Substances of abuse that may induce anxiety symptoms during intoxication include stimulants (e.g., amphetamines, cocaine, caffeine), ecstasy, cannabis, PCP, inhalants; hallucinogens, whereas alcohol, barbiturates, benzodiazepines, opioids, and nicotine can produce anxiety symptoms during withdrawal (Table 10.5–12). Carbon monoxide, paint, and gasoline fumes may also induce anxiety.

Pathophysiologically, medical conditions and substances capable of inducing anxiety likely do so by affecting noradrenergic, serotonergic, or γ-aminobutyric acid neurotransmission, or a combination of these. Supporting this link is evidence that medical conditions (e.g., hyperthyroidism, hypoxia, and pheochromocytoma) and agents (e.g., sympathomimetic agents and stimulants) that stimulate noradrenergic function increase anxiety. Postconcussive anxiety has been hypothesized to stem from injury to limbic structures.

Diagnosis and Clinical Features

The reader is referred to Table 10.5–6 for general principles regarding diagnosis. Similar to primary anxiety disorders, anxiety stemming from a medical condition or substance/medication may present with physical complaints (e.g., chest pain, palpitation, abdominal distress, diaphoresis, dizziness, tremulousness, and urinary frequency), generalized symptoms of fear and excessive worry, outright panic attacks associated with fear of dying or losing control, or phobia with associated avoidant behavior.

For anxiety disorder due to another medical condition, the etiological medical condition is listed first, followed by Anxiety Disorder due to

another medication condition (e.g., hyperthyroidism, anxiety disorder due to hyperthyroidism).

For substance/medication-induced anxiety disorder, the specifiers include *with onset during intoxication, with onset during withdrawal,* and *with onset after medication use*. If the clinician believes that medical and substance/medication factors contribute to the anxiety presentation, both diagnoses should be made. If the clinician is uncertain as to whether the anxiety disorder is primary or due to another medical condition or substance, a diagnosis of other specified anxiety disorder or unspecified anxiety disorder may be appropriate.

An 86-year-old retired chemical engineer sought help for the onset of a series of attacks over the preceding 4 months in which he experienced marked apprehension, restlessness, a sense that the "walls were caving in," and the need to "get air" in order to relieve his sense of discomfort. These events typically occurred during the night and awaken him from sound sleep. In order to feel better, he would need to stick his head out of an open window, regardless of how cold it was outside. His symptoms would gradually improve over 15 to 20 minutes, but complete resolution of these symptoms took a full day. In response to pointed questioning, the patient reported sweating, dizziness, and shortness of breath during these episodes. He imagined that he would die if he could not open the window. He denied palpitations, choking sensations, paresthesias, or nausea. The patient recalled a similar series of attacks almost 30 years earlier during a period of time in which he frequently needed to travel and hence, was away from home because of work obligations. The patient denied depressed mood, anhedonia, recent sleep dysfunction, change in appetite or weight, decreased energy, or feelings of worthlessness. His medical history was notable for a right basal ganglia stroke 6 months earlier. He has had a history of hypertension, borderline diabetes, and benign prostatic hypertrophy. Laboratory studies were unremarkable.

A diagnosis of anxiety disorder due to stroke was made. The patient was prescribed alprazolam 0.5 mg orally twice a day as needed for panic attacks and started on escitalopram 10 mg per day. After 4 weeks, the escitalopram was increased to 20 mg per day. At follow-up visit, the patient reported complete resolution of his anxiety symptoms. He remained on the escitalopram but no longer required the alprazolam.

Laboratory Examination

A targeted workup is indicated with tests selected to rule in specific diagnoses suggested by the patient's somatic symptoms (if present).

Tests to consider include complete blood count, electrolytes, glucose, blood urea nitrogen, creatinine, liver function tests, calcium, magnesium, phosphorus, thyroid function tests, and urine toxicology. Occasionally, additional studies may be indicated to rule out a pheochromocytoma (e.g., urinary catecholamine), a seizure disorder (e.g., EEG), cardiac arrhythmia (e.g., Holter monitoring), and pulmonary disease (pulse oximetry, arterial blood gases). Brain imaging may be useful in ruling out demyelinating disorder, tumor, stroke, or hydrocephalus and is especially important if the anxious individual reports neurological symptoms (e.g., headache, motor or sensory changes, and dizziness), although such complaints may represent somatic manifestations of primary anxiety disorders. Lumbar puncture may be appropriate if an inflammatory or infectious cause is

suspected.

Differential Diagnosis

The reader is referred to Table 10.5–6 for general principles regarding diagnosis. Adjustment disorders with anxiety arising within the context of reaction to medical or other life stressors should not be diagnosed as anxiety disorder due to another medical condition. Anxiety symptoms arising solely in the context of a psychotic or mood disorder are not coded as an anxiety disorder. Patients with psychosis of any origin may experience anxiety commonly related to delusions or hallucinations. Depressive disorders often present with anxiety symptoms; therefore, the clinician should assess for depressive symptoms in any patient reporting anxiety. Anxiety occurring in the context of individuals with delirium should not be diagnosed as anxiety disorder due to another medical condition. Dementia often is associated with agitation or anxiety, especially at night (e.g., sundowning), but the presence of such symptoms is encompassed by the *with behavioral disturbance* specifier for NCD.

In distinguishing anxiety due to medical illness from that due to substances, differential diagnosis may be assisted by examination of anxiety symptom chronology, response to correction or removal of suspect medical conditions and substances, and, occasionally, examination of toxicology results.

Course and Prognosis

Anxiety disorder due to another medical condition usually fluctuates in direct relation to the course of the provoking factor. Medical conditions responsive to treatment or cure (e.g., correction of hypothyroidism and reduction in caffeine consumption) often provide concomitant relief of anxiety symptoms, although such relief may lag the rate or extent of improvement in the underlying medical condition. Chronic, incurable medical conditions associated with persistent physiological insult (e.g., chronic obstructive pulmonary disease) or recurrent relapse of substance use may contribute to seeming refractoriness of associated anxiety symptoms. In medication-induced anxiety, if complete cessation of the offending factor (e.g., immunosuppressant therapy) is not possible, dose reduction, when clinically feasible, often brings substantial relief.

Treatment

Anxiety disorder due to a general medical condition may improve with correction of the underlying medical condition. If substance abuse is confirmed, a comprehensive treatment program is indicated. Pharmacological interventions for anxiety symptoms include SSRI and serotonin and norepinephrine reuptake inhibitor (SNRI) antidepressants

and buspirone [BuSpar]. When using SSRIs or SNRIs to treat anxiety disorders, the starting dose should be low in order to minimize transitory anxiogenic side effects. Tricyclic antidepressants should be avoided due to their anticholinergic, orthostatic, and cardiac side effects. As it may take weeks before these drugs positively impact the anxiety symptoms, short course of benzodiazepines may be helpful in controlling anxiety symptoms. However, benzodiazepines should be used cautiously in persons prone to substance abuse, those with advanced pulmonary disease (due to risk of respiratory depression), and those with pronounced cognitive defects.

Antipsychotic medications occasionally are used for time-limited management of severe anxiety, often in the context of fleeting cognitive or psychotic symptoms, as may be seen in patients treated with corticosteroids. With brief use, the risk of tardive dyskinesia is negligible. Atypical agents, owing to lower risk of EPS compared to typical agents, are preferred.

Psychotherapeutic interventions to consider include cognitive and relaxation therapies, which may be useful in diminishing phobic, panic, and avoidant symptoms. Insight-oriented exploration of stressful life circumstances may be helpful once severe anxiety symptoms are contained pharmacologically.

OBSESSIVE-COMPULSIVE AND RELATED DISORDER DUE TO ANOTHER MEDICAL CONDITION AND OBSESSIVE-COMPULSIVE AND RELATED DISORDER SUBSTANCE/MEDICATION-INDUCED

Definition

OCDs classically involve the presence of obsessions, compulsions or both. Obsessions are recurrent and persistent thoughts, urges, or images that are intrusive and unwanted by the patient. Compulsions are repetitive behaviors or mental acts that a patient feels driven to perform in response to an obsession or according to a rigid set of rules. Within the expanded rubric of related disorders in DSM-5 are those disorders characterized by preoccupations and repetitive behaviors or mental acts in response to these preoccupations.

In the diagnosis obsessive-compulsive and related disorder due to another medical condition, the individual experiences OCD symptoms that cause clinically significant distress or impairment in functioning that is thought to be caused by the direct physiological effect of a medical condition and not solely represent a psychological reaction to having a medical illness. In the diagnosis substance-induced obsessive-compulsive and related disorder, the symptoms are the product of intoxication or withdrawal from prescribed medication or from a nonprescribed substance, typically a drug of abuse.

Comparative Nosology

DSM-III-R offered the diagnosis of organic anxiety syndrome to describe panic attacks or generalized anxiety stemming physiologically from either medical conditions or substances. The comparable DSM-IV-TR anxiety diagnoses were described as due to a general medical condition or substance induced. In addition, DSM-IV-TR broadened the diagnosis of anxiety disorder due to a general medical condition to include obsessive-compulsive symptoms and, in the case of substance-induced anxiety, phobic symptoms. In contrast, in DSM-5, OCD, including that due to another medical disorder or substance induced, has been separated from the anxiety disorders based on evidence supporting a distinct pathophysiology. Two secondary causes of OCDs are offered: obsessive-compulsive and related disorder due to another medical condition and substance/medication-induced obsessive-compulsive and related disorder.

ICD-10 offers the diagnosis of *other specified mental disorders due to brain damage and dysfunction and to physical disease*. For psychoactive substances, the clinician may use the specifier *with unspecified substance-induced disorder* to modify the primary substance use diagnosis.

Epidemiology

The prevalence of obsessive-compulsive and related disorder due to another medical condition is unknown. The prevalence of pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections (PANDAS) (see below) is considered to be rare with 200,000 or fewer cases in the United States.

Etiology

Etiology is specific to the underlying medical condition. Evidence suggests that OCD symptoms can occur as an autoimmune reaction after Group A streptococcal infection. Two disorders with prominent OCD features that occur after Group A streptococcal infection are Sydenham chorea and PANDAS. It is hypothesized that the infection triggers an autoimmune reaction against the basal ganglia. The assumption is that antigens on the cell wall of the streptococcal bacteria mimic the proteins of the brain, resulting in the autoimmune reaction. In Sydenham chorea, in addition to motor symptoms, patients can also experience obsessions, compulsions, cognitive impairment, and emotional lability. In childhood, PANDAS presents with sudden onset of obsessions and compulsions, which may also be accompanied by tics and other neuropsychiatric symptoms. However, the association of this syndrome with a streptococcal infection remains somewhat controversial; hence this disorder is also referred to as pediatric acute-onset neuropsychiatric syndrome (PANS) or idiopathic childhood acute neuropsychiatric symptoms (CANS).

Stimulants, including over-the-counter preparations (e.g., presumably high doses of pseudoephedrine), prescription medications (e.g., amphetamine, methylphenidate), and drugs of abuse (e.g., cocaine), may induce OCD symptoms in both intoxication and withdrawal.

Diagnosis and Clinical Features

The reader is referred to Table 10.5–6 for general principles regarding diagnosis.

For obsessive-compulsive and related disorder due to another medical condition, five DSM-5 specifiers exist: *with OCD-like symptoms, with appearance preoccupation, with hoarding symptoms, with hair-pulling symptoms, and with skin-picking symptoms*. For substance/medication-induced obsessive-compulsive and related disorder, the three specifiers include *with onset during intoxication, with onset during withdrawal, and with onset after medication use*.

If the clinician believes that both a medical condition and a substance/medication contributed to the clinical presentation, then both diagnoses should be made. If the clinician is uncertain as to whether the presentation is primary or due to a general medical condition or substance, a diagnosis of other specified obsessive-compulsive and related disorder or unspecified obsessive-compulsive and related disorder is indicated, based on whether the clinician can document the clinical reason that a different diagnosis was not made.

Laboratory Examination

A targeted workup should include tests selected to "rule in" specific diagnoses suggested by the patient's somatic symptoms (if present).

In regard to both PANDAS and Sydenham chorea, the only laboratory test that is helpful is to establish that the patient has or recently had a Group A streptococcal infection.

Differential Diagnosis

The reader is referred to Table 10.5–6.

Course and Prognosis

Classically, children with PANDAS experience an abrupt onset of symptoms followed by a slow improvement over weeks to months. However, in the context of another Group A streptococcal bacterial infection, the symptoms can suddenly worsen. Stimulant-induced OCD symptoms typically resolve after resolution of the intoxication and withdrawal phases.

Treatment

Obsessive-compulsive and related disorder due to another medical condition may improve with correction of the underlying medical condition. If substance abuse is confirmed, appropriate treatment should be offered. Treatment of obsessive-compulsive and related disorder symptoms that do not improve spontaneous would involve standard treatment approaches, including cognitive behavioral therapy, and medications such as SSRIs or clomipramine [Anafranil].

SUBSTANCE/MEDICATION-INDUCED SLEEP DISORDER

Sleep disturbances can be primary, arise in the context of other mental disorders (e.g., depression, anxiety, psychosis, and dementia), or can stem from the direct physiological effects of a medical condition or substance/medication.

Definition

Substances and medications can impact sleep and can cause insomnia, hypersomnolence, and/or parasomnia.

Comparative Nosology

DSM-III-R provided for the diagnosis insomnia related to a known organic factor, which could be due to either an underlying medical illness or substance use and which could include physiological disturbances unique to sleep (e.g., sleep apnea) which were subsequently excluded from the corresponding DSM-IV and DSM-IV-TR diagnoses. DSM-III-R also provided for the diagnosis hypersomnia related to a known organic factor (i.e., medical illness or substance), a category that included physiological disturbances unique to sleep (e.g., sleep apnea and narcolepsy) that were also subsequently excluded from the corresponding DSM-IV and DSM-IV-TR diagnoses. For parasomnias related to general medical conditions or substances, DSM-III-R offered the diagnosis parasomnia not otherwise specified. In DSM-5, the relationship between medical conditions and sleep disorders has been reconceptualized. In DSM-5, it is now acknowledged that that there is a bidirectional relationship between medical illness and sleep quality/quantity. Hence, there is no longer a unique category of sleep disorder due to another medical condition. Instead, many sleep disorder diagnoses, including insomnia disorder, hypersomnolence disorder and nightmare disorder, now have a specifier that allows the clinician to indicate that there is a comorbid medical diagnosis that is related to the sleep disorder. A specific diagnosis for *substance/medication-induced sleep disorder* remains in DSM-5. In ICD-10, the clinician may use the specifier *with substance-induced sleep disorder* to modify the primary substance use diagnosis.

Epidemiology

The prevalence of sleep disorders that are substance/medication-induced is unknown. However, elderly individuals with health issues and hence multiple medications and younger individuals with substance use disorders are likely at heightened risk for the development of these disorders.

Etiology

Categories of medications that may induce sleep disorders include adrenergic (e.g., clonidine), cholinergic (e.g., donepezil), dopaminergic (e.g., L-Dopa, methylphenidate, and bupropion), and serotonergic (e.g., SSRIs) agents as well as antihistamines, corticosteroids, sympathomimetic agents (e.g., theophylline), and thyroid preparations. Substances that may induce sleep disorder during intoxication or withdrawal include alcohol, caffeine, cannabis, opioid, sedative hypnotics, amphetamines, and other stimulants as well as cocaine.

Diagnosis and Clinical Features

Basic considerations for diagnosing a substance/medication-induced sleep disorder are outlined in Table 10.5–6. For drugs of abuse, the clinician must find clinical or laboratory evidence of abuse, dependence, withdrawal, or intoxication. For drugs with long half-lives (e.g., diazepam [Valium]), the onset of sleep disturbance may occur as long as several weeks after abstinence, although, in general, onset of sleep disturbance more than 4 weeks after cessation of substance use suggests that the disturbance is more likely accounted for by a primary sleep disorder or a general medical condition. Substance/medication-induced sleep disorder criteria include specifiers for type (insomnia type, daytime sleepiness type, parasomnia type, and mixed type) and onset (with onset during intoxication and onset during discontinuation/withdrawal).

Laboratory Examination

A focused workup may be useful in differentiating substance/medication-induced sleep disorders from those more closely linked to an underlying medical disorder. Laboratory examination may include complete blood count, serum chemistries, thyroid function tests, and urine toxicology. Arterial blood gas measurements and/or nocturnal oximetry should be considered when sleep apnea, another sleep breathing disorder, or an underlying pulmonary condition is suspected. A nocturnal ECG and/or a formal polysomnography examination can detect sleep-related seizures and sleep-related breathing disorders as well as help to identify other medical conditions. For example, patients with fibromyalgia often have alpha brain-wave activity during non-rapid eye movement (NREM) sleep. ECG

findings, such as decreased slow wave sleep, and a decrease in REM sleep or phasic REM density, are nonspecific.

Differential Diagnosis

Sleep disturbance are common in mood disorders as well as in NCDs.

Often the distinction between sleep disorder comorbid with medical condition and substance-induced sleep disorder is complicated. For example, an individual with chronic obstructive pulmonary disease taking sympathomimetic agents may experience sleep disturbance due to these agents or to the underlying medical condition. Examination of chronology of sleep symptoms, especially as they relate to correction of associated medical conditions or removal of suspect substances, frequently proves diagnostically useful.

Course and Treatment

Sleep disorders induced by a substance or medication may improve upon cessation of use of the offending agent. If not possible, a reduction in dose of the offending medication may result in improvement in sleep. If substance abuse is confirmed, a comprehensive rehabilitation program is indicated. Patients with insomnia often have some degree of sleep-state misperception. Hence, when tracking response to treatment, patients should be asked to record their sleep activity. In addition, informants should be used when available. Individuals with sleep disorders should be educated about the importance of sleep hygiene. Good sleep hygiene practices include use of the bed for sleep and sex only (e.g., no television watching or reading in bed), avoiding "watching the clock" while in bed, practicing relaxation techniques at bedtime, and avoiding stimulating activities during the 3 hours before bedtime including exercise. Relaxation techniques are particularly important for patients who are anxious about having insomnia. Patient should be encouraged to maintain a regular schedule for bedtime and wakening and to avoid daytime naps. In addition, patients should be advised not to spend the night in bed if awake. Instead, the sleepless patient should be advised to get out of bed and spend the time on quiet activities until they feel sleepy enough to warrant returning to bed. Patient should be advised to avoid stimulating agents such as nicotine and caffeine later in the day. In addition, patient should avoid alcohol, which can affect sleep architecture.

Patients frequently try over-the-counter preparations to improve sleep. These drugs typically block the histamine type-1 receptor. However, in view of the anticholinergic properties of these agents, they should be avoided in the elderly or medically debilitated patients, especially those with prostatic hypertrophy, cognitive disorders, and constipation. In addition, many of these drugs have a long duration of action, and their sedative effects may

persist well in to the following day.

Judicious use of benzodiazepines may be warranted to break a cycle of disordered sleep and the resultant anxiety that some patients experience. Short courses of treatment are advised and the risk of abuse needs to be considered. Tolerance and dependence tend to develop with extended use, compounding sleep disturbance and often inciting resistance at sedative–hypnotic taper by individuals reluctant to experience temporary withdrawal phenomena. Caution must be exercised in the use of benzodiazepines, as these agents can exacerbate cognitive loss commonly present in the medically ill and elderly. If needed, agents with short half-lives (e.g., lorazepam and alprazolam) are preferred. In individuals with liver disease, agents that only require conjugation with glucuronide for elimination (e.g., lorazepam, oxazepam, and temazepam) are preferred, as liver disease does not generally affect this process. Caution also must be used in patients with pulmonary disease due to risk of respiratory depression.

Short-acting nonbenzodiazepine hypnotic agents, such as zolpidem [Ambien], zaleplon [Sonata], or eszopiclone [Lunesta] are preferred in the elderly and in those experiencing initial insomnia. These medications are associated with memory loss as well as abnormal nocturnal activities (e.g., sleep-walking, sleep-eating, sleep-driving) during which the person is not fully awake and is at risk of injury. The lowest possible dose should be used for the shortest period of time possible. Longer-acting agents, such as zolpidem [Ambien CR], are now available and should be considered in patients who also have difficulty maintaining sleep. Trazodone [Desyrel] and gabapentin [Neurontin] do not cause physiological dependence. Caution should be used when prescribing trazodone in the older or medically complicated patient as it has the potential to cause orthostatic hypotension. Occasionally, brief courses of atypical antipsychotic agents, such as olanzapine and risperidone, are useful in regularizing sleep in the elderly and in hospitalized patients; these agents pose little risk of inducing delirium or exacerbating cognitive impairment, to which these populations are highly susceptible. If depression is contributing to sleep disturbance, antidepressant pharmacotherapy should be considered. SNRIs (e.g., duloxetine [Cymbalta]) and tricyclic agents (e.g., nortriptyline [Aventyl]) may be especially useful if neuropathic pain contributes to sleep difficulties. Older sleep-inducing agents, such as chloral hydrate (Aquachloral Supprettes) or meprobamate (Miltown), are rarely used.

DSM-5 CRITERIA FOR SUBSTANCE/MEDICATION-INDUCED SEXUAL DYSFUNCTION

Definition

Sexual function is influenced by numerous factors include medical/physiological, psychological, and cultural factors. Sexual

dysfunction is attributed to the use of a substance/medication only when the sexual symptoms are felt to be fully explained by the direct physiological effects of the chemical.

Comparative Nosology

DSM-III-R did not include any diagnoses for sexual dysfunction attributable "exclusively to organic factors." A formal DSM-III-R diagnosis was provided only if psychological factors were at least partially contributing to the sexual dysfunction. Diagnostic specifiers could be used to indicate that the dysfunction was psychogenic only, or both psychogenic and biogenic (i.e., organic). Furthermore, DSM-III-R sexual dysfunction diagnoses could not be made if the dysfunction occurred exclusively within the context of another Axis I mental disorder.

In contrast, DSM-IV-TR includes the diagnosis of sexual dysfunction due exclusively to physiological factors (i.e., "organic"), which is recorded on Axis I. DSM-5 has reverted back to the stance of DSM III R, namely, that there is no psychiatric diagnosis for sexual dysfunction that is thought to be fully explained by another medical condition. However, unlike DSM III R, DSM-5 does not offer a psychiatric diagnosis for sexual dysfunction that is at least partially attributable to another medical disorder. DSM-5 continues to recognize *substance/medication-induced sexual dysfunction* with the specifiers *with onset during intoxication, with onset during withdrawal,* and *with onset after medical use*. In the context of substance abuse, ICD-10 codes the induced sexual dysfunction as a specifier to the primary diagnosis of the comorbid substance use disorder. In the case of the absence of misuse (e.g., sexual dysfunction as a side effect of the appropriate use of a medication), ICD-10 provides the diagnosis of *substance-induced sexual dysfunction*.

Epidemiology

Medication can cause sexual dysfunction; for example, delayed orgasm may affect as much as 50 percent of individuals taking SSRIs.

Etiology

Medications that may cause diminished sexual interest and orgasmic dysfunction include antihypertensive agents, antidepressant agents (particularly SSRIs), anticonvulsants, and antipsychotics. Additionally, antipsychotic agents, such as haloperidol and fluphenazine, may contribute to painful orgasm. Trazodone occasionally causes priapism.

Other substances that contribute to sexual dysfunction in intoxication include alcohol, amphetamine and other stimulants, and cocaine nicotine, opioids, and sedative hypnotics or anxiolytics (Table 10.5–12).

Diagnosis and Clinical Features

The reader is referred to Table 10.5–6 for general principles regarding diagnosis. For substance/medication-induced sexual dysfunction, the clinician should specify both the relationship to substance use (i.e., with onset during intoxication, with onset during withdrawal, and with onset after medication use) as well as the severity (mild, moderate, severe).

Laboratory Examination

Urine toxicology may be relevant if substance abuse is suspected.

Differential Diagnosis

The diagnosis of substance-induced sexual dysfunction is made if a substance is fully explanatory. When the clinician encounters difficulty in distinguishing sexual dysfunction that may be substance induced from that which may be due to a concurrent medical condition, differential diagnosis may be facilitated by observation after discontinuation of the suspect substance.

Course and Prognosis

The course of sexual dysfunction due to a substance/medication is highly variable and typically depends on etiology.

Treatment

The first step in treatment is to remove or minimize exposure to the offending substance/medication (if possible) and actively treat the underlying medical condition. If substance abuse is confirmed, inpatient or outpatient treatment should be offered.

Psychotherapy issues may also be appropriate especially to address interpersonal or character issues exacerbating sexual dysfunction, including depression, anger, and relationship problems.

Pharmacological therapy for men includes the use of sildenafil [Viagra] or vardenafil [Levitra] to treat erectile dysfunction stemming from primary and medically induced or substance-induced sexual dysfunction. Sildenafil and vardenafil can lead to severe hypotension and death in patients with coronary artery disease who use nitrates and hence should be avoided in patients taking nitrates. Penile prosthetic devices are an option for men who do not benefit from or who cannot take sildenafil or vardenafil. Testosterone replacement is appropriate for men with hypoactive sexual desire and low testosterone levels.

Anecdotal evidence suggests that sildenafil or vardenafil may also be useful in women with sexual arousal difficulties. Women with dyspareunia may benefit from lubricating gels and hormone creams. Women with

androgen deficiency may benefit from testosterone therapy.

CATATONIC DISORDER DUE TO ANOTHER MEDICAL CONDITION

Catatonia is occasionally seen in severe psychotic and mood disorders. It can also be caused by an underlying medical condition.

Definition

Catatonia is a clinical syndrome characterized by striking behavioral abnormalities, which may include motoric immobility or excitement, profound negativism, or mimicry (of speech of movement). A diagnosis of catatonic disorder due to another medical condition can be made if there is evidence that the condition is due to the physiological effects of a general medical condition. The diagnosis is not made if the catatonia is better explained by a primary mental disorder, such as schizophrenia or psychotic depression, or if catatonic symptoms occur exclusively within the course of delirium.

Comparative Nosology

Catatonia due to a general medical condition first appeared in DSM-IV-TR and is also in DSM-5. There is not a formal category for substance/medication-induced catatonia. The comparable ICD-10 diagnosis is *organic catatonic disorder*.

Epidemiology

Catatonia is an uncommon condition mostly seen in advanced primary mood or psychotic illnesses. Among inpatients with catatonia, 25 to 50 percent are related to mood disorders (e.g., major depressive episode, recurrent, with catatonic features), and approximately 10 percent are associated with schizophrenia. The prevalence of catatonia due to medical conditions or substances is unknown.

Etiology

Medical conditions that can cause catatonia include neurological disorders (e.g., nonconvulsive status epilepticus, and head trauma), infections (e.g., encephalitis), and metabolic disturbances (e.g., hepatic encephalopathy, hyponatremia, and hypercalcemia).

Medications that can cause catatonia include corticosteroids, immunosuppressants, and antipsychotic (i.e., neuroleptic) agents. Catatonic symptoms may be seen in extreme forms of neuroleptic-induced parkinsonism or neuroleptic malignant syndrome, a rare, potentially life-threatening disorder associated with fever, autonomic instability, impaired consciousness, and rigidity.

Diagnosis and Clinical Features

DSM-5 criteria for the diagnosis of catatonic disorder due to a general medical condition include behavioral changes characteristic of catatonia, evidence of a physiological basis for the symptoms, and exclusion of primary mental disorders and delirium (see Table 10.5–6). Lethal catatonia is a rare advanced stage of the disorder that features fever and autonomic instability and may be fatal.

Laboratory Examination

There are no pathognomonic laboratory findings in catatonia. The laboratory evaluation should be used to identify any underlying medical condition. Appropriate medical tests may include complete blood counts, electrolytes, brain imaging, and EEG (if seizures are suspected). In addition, serum creatinine phosphokinase, white blood cell count, and serum transaminases should be checked as these laboratory tests are elevated in neuroleptic malignant syndrome.

Differential Diagnosis

Differential diagnoses include hypoactive delirium, end-stage dementia, and akinetic mutism, as well as catatonia due to a primary psychiatric disorder. It is important to identify cases of catatonia occurring in the setting of neuroleptic malignant syndrome as the latter diagnosis can be fatal. Features suggesting neuroleptic malignant syndrome include autonomic instability and delirium in addition to elevated serum creatinine phosphokinase, white blood cell count, and serum transaminases.

Course and Treatment

Catatonia impairs a person's ability to care for him- or herself and therefore requires hospitalization. In an excited state, the catatonic patient may represent a danger to others; hence, close supervision is needed. Fluid and nutrient intake must be maintained, often with intravenous lines or feeding tubes. The catatonic individual must be assisted with hygiene.

The primary treatment modality is identifying and correcting the underlying medical or pharmacological cause. Offending substances must be removed or minimized.

Benzodiazepines can provide temporary improvement in symptoms and their use may improve the patient's ability to communicate and to care for themselves. ECT is appropriate for catatonia due to a general medical condition, especially if the catatonia is life threatening (e.g., inability to eat) or has developed into lethal (malignant) catatonia. The mechanism behind the efficacy of ECT is unknown.

PERSONALITY CHANGE DUE TO ANOTHER MEDICAL CONDITION

Personality traits are formed by early adulthood, persist throughout life, and affect every aspect of day-to-day behavior including social and occupational function. Although personality is usually a stable construct, certain medical conditions and substances can alter personality.

Definition

In personality change due to another medical condition, the individual's previous characteristic personality pattern is persistently altered by direct physiological consequence of a medical condition (e.g., stroke). Of note, this diagnosis is not made in the presence of delirium.

History and Comparative Nosology

Historically, the term *frontal lobe syndrome* has been used to describe change in premorbid personality due to injury to the frontal lobes. DSM-III-R included the diagnosis of organic personality syndrome. Three patterns of behavior were recognized: (1) involves labile affect and poor social judgment, (2) marked apathy, and (3) religiosity or paranoia with verbosity. The specifier *explosive type* could be used if outbursts of aggression or rage were the predominant feature.

The comparable DSM-IV-TR and DSM-5 diagnosis provides for a much broader range of personality change subtypes including labile type, disinhibited type, aggressive type, apathetic type, and paranoid type. The comparable ICD-10 diagnosis is *organic personality disorder*.

Epidemiology

Few data exist on the prevalence of personality change due to a general medical condition. Head injury is a common cause of personality change. Young men and the users of substances are at elevated risk of head injury and, hence, likely represent a disproportionate percentage of patients with personality change due to another medical condition.

Etiology

Neurological conditions, especially those affecting the frontal or temporal lobes or subcortical structures, may lead to personality change. Damage to the frontal lobes can lead to impaired social judgment and impaired initiative or induce apathy and disengagement. Orbitofrontal lesions tend to induce disinhibited behaviors and affective labiality. Lesions in the temporal lobe can affect memory, emotional response, and interpretation of events, thus contributing to the excessive seriousness and religiosity often observed in the temporal lobe epilepsy. Diseases associated with personality change include seizures, trauma, chronic traumatic encephalopathy, neoplasm, encephalitis, cerebrovascular injury, and demyelinating diseases. Subcortical degenerative disease, such as

Huntington disease and Wilson disease, can lead to personality change. Endocrine conditions, such as hypo- or hyperthyroidism or hypo- or hyperadrenocorticism, also may alter personality. Chronic alcohol use appears to have marked adverse impact on frontal lobe function, with associated personality change.

Diagnosis and Clinical Features

The diagnosis of personality change due to another medical condition requires that the personality change must be persistent and must represent a change from the individual's previous characteristic personality pattern, is physiologically due to a medical condition, and cause significant distress or impairment in functioning. The diagnosis is not given if symptoms are better accounted for by another mental disorder or if they occur in the context of delirium (see Table 10.5–6).

In DSM-5, subtypes of personality change due to another medical condition exist according to the predominant feature of the personality change. These include labile type, disinhibited type, aggressive type, apathetic type, and paranoid type. Other type is used if the predominant feature is not captured by the previous subtypes. Combined type is used if the presentation has features of more than one defined subtype. Unspecified type serves as a residual specifier.

MS is a 67-year-old divorced female, a retired hair stylist, with no past psychiatric history who was referred for the treatment of depression by her primary care physician. The patient has diabetes and hypertension. During the interview, the patient was surprised that she was referred for depression. She stated that she felt fine and was doing well. She specifically denied low mood or lack of interest or enjoyment. She reported good sleep and normal appetite. She denied loss of self-esteem. She was not suicidal.

With the patient's permission, the daughter, who accompanied the patient to the appointment, was interviewed separately. The daughter reported that over the preceding 3 years, the patient had gradually ceased most of her usual activities. She was no longer walking daily or visiting friends (as had been her practice); she was neglecting her housework, and at times, was no longer attentive to personal hygiene. The daughter reported that the patient slept 10 to 12 hours a night and spent most of the day lying on the couch watching television. The daughter noted that her mother had not seemed sad or distraught and that she would engage in activities but only after much encouragement.

Laboratory workup with normal brain imaging demonstrated extensive small vessel ischemic changes. Neuropsychological testing demonstrated some mild impairment, especially in the domains of speed and executive function.

The patient meets criteria for personality change due to cerebrovascular disease, apathetic type.

Laboratory Examination

A targeted workup is required when a personality change due to a general medical condition is being considered as part of the differential diagnosis. If possible, tests should be selected to rule in specific diagnoses suggested by the patient's somatic symptoms (if present). Structural imaging studies,

such as MRI or CT, are indicated. EEG should be obtained if seizure disorder is considered. Laboratory tests to be considered include a complete blood count and chemistries, thyroid function tests, vitamin B_{12} level, RPR, and urinalysis. Urine toxicology usually is indicated. Special studies occasionally may be sought (e.g., serum ceruloplasmin in suspected Wilson disease).

Differential Diagnosis

Chronic severe illness, and in particular chronic pain syndromes, may alter a patient's personality through the illness's impact on the person's social or occupational functioning, mobility, and level of independence; however, these changes reflect the individual's mechanisms for coping and are not a direct physiological effect of the underlying medical illness. Therefore, in this situation, the diagnosis of a personality change due to another medical condition is not appropriate. This diagnosis of personality change due to another medical condition is also not appropriate if the change in personality is better accounted for by either a primary or secondary cognitive, mood, anxiety, or psychotic disorder or to delirium. The diagnosis of personality change due to dementia should be confined only to individuals in whom personality change is an especially prominent aspect of the clinical picture as alternation of personality is intrinsic to dementia.

Personality change due to another medical condition is not to be confused with personality disorders, which are primary mental illnesses, are stable and enduring, and are not due to the physiological effects of a general medical condition. Finally, personality change is an intrinsic aspect of substance dependence and should not receive specific diagnosis, unless such personality change forms the basis for clinical attention independent of the substance disorder.

Course and Prognosis

Course depends on the nature of the underlying medical illness. Progressive or irreversible brain damage (e.g., multiple sclerosis or severe head trauma, stroke) will yield permanent alternations in personality. However, personality change associated with treatable medical conditions (e.g., hypothyroidism) should improve or resolve with successful treatment of the underlying medical condition.

Treatment

The first step in the treatment is to actively treat the underlying medical condition. The addition of psychopharmacological and psychotherapeutic interventions is designed to address behavioral and mood disturbances (e.g., aggression and irritability). Anticonvulsants (e.g., valproic acid [Depakene], lamotrigine [Lamictal], topiramate [Topamax], or

carbamazepine) or SSRI antidepressants may be useful in managing labile affect, anger, and poor impulse control. Antipsychotic medications may be useful in managing severe aggression and paranoia. Stimulants (e.g., methylphenidate) may have a role in managing apathetic patients. For patients who are impulsive, appropriate caution is required when prescribing medications that can be lethal in overdose (e.g., tricyclic antidepressants) or that may exacerbate disinhibition or cognitive deficits (e.g., benzodiazepines).

Psychotherapy is appropriate to help the patient deal with the losses associated with the illnesses (e.g., loss of employment, loss of friends) and to help the patient adjust to his or her limitations. Vocational therapy should be considered. Psychoeducational and insight-oriented strategies designed to enhance awareness of the adverse effects of personality change may also have value. Family members may benefit from psychoeducational interventions. In addition, since altering a person's environment may have a positive impact on behavior, the family should be counseled on how to circumvent problematic behaviors and associated frustrations on behalf of the patient.

REFERENCES

American Geriatrics Society. 2012 Beers Criteria Update Expert Panel. American Geriatrics Society updated Beers Criteria for potentially inappropriate medication use in older adults. *J Am Geriatr Soc*. 2012;60:616–631.

*Chan P, Brew BJ. HIV associated neurocognitive disorders in the modern antiviral treatment era: prevalence, characteristics, biomarkers, and effects of treatment. *Curr HIV/AIDS Rep*. 2014;11:317–324.

Judd LL, Schettler PJ, Brown ES, et al. Adverse consequences of glucocorticoid medication: psychological, cognitive, and behavioral effects. *Am J Psychiatry*. 2014;171:1045–1051.

Kavirajan H, Schneider LS. Efficacy and adverse effects of cholinesterase inhibitors and memantine in vascular dementia: a meta-analysis of randomised controlled trials. *Lancet Neurol*. 2007;6:782–792.

*Kim E, Lauterbach EC, Reeve A, et al. ANPA Committee on Research. Neuropsychiatric complications of traumatic brain injury: a critical review of the literature (a report by the ANPA Committee on Research). *J Neuropsychiatry Clin Neurosci*. 2007 Spring;19:106–127.

McCrea M, Iverson GL, McAllister TW, et al. An integrated review of recovery after mild traumatic brain injury (MTBI): implications for clinical management. *Clin Neuropsychol*. 2009;23:1368–1390.

*Mind Exchange Group. Assessment, diagnosis, and treatment of HIV-associated neurocognitive disorder: a consensus report of the mind exchange program. *Clin Infect Dis*. 2013;56:1004–1017.

*Neurobehavioral Guidelines Working Group, Warden DL, Gordon B, et al. Guidelines for the pharmacologic treatment of neurobehavioral sequelae of traumatic brain injury. *J Neurotrauma*. 2006;23:1468–1501.

Román GC, Tatemichi TK, Erkinjuntti T, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology*. 1993;43:250–260.

Sacktor N, Robertson K. Evolving clinical phenotypes in HIV-associated neurocognitive disorders. *Curr Opin HIV AIDS*. 2014;9:517–520.

Wetterling T, Kanitz RD, Borgis KJ. Comparison of different diagnostic criteria for vascular dementia (ADDTC, DSM-IV, ICD-10, NINDS-AIREN). *Stroke*. 1996;27:30–36.

11

Substance-Related Disorders

▲ 11.1 Substance Use Disorders: Introduction

ERIC CAMERON STRAIN, M.D.

INTRODUCTION TO THE SECTION

This section provides a series of chapters addressing substance use disorders (SUDs) for each of the major classes of abused substances. The selection of chapter topics is generally based on the categorization of abusable substances as presented in DSM-5, and has resulted in some consolidation of topics that reflect the revisions in the new edition of DSM (e.g., cocaine and amphetamines, which were previously addressed in separate chapters of this book, are now in a single chapter that addresses stimulants). As was true in the last edition of this book, the exception is that there is again a chapter addressing steroids in this edition, despite there not being a separate steroid use disorder section in DSM-5. Given the cultural interest and concern with use of this class of drugs, the topic remains here. Another major change in DSM-5 is that the diagnosis of pathologic gambling is now in the SUD section (but is not addressed in this section of this textbook). As one considers the broader categorization of medical disorders that have a core feature that defines them by a dysregulation of choice (an inability by the person to control choice in a healthy and reasonable manner), then the inclusion of gambling and drug use disorders is a highly appropriate move on the part of the field. It seems likely that the underlying phenomenology as well as biology of disorders that are characterized by this dysregulation of choice shares common features, logics, and assumptions that can help psychiatrists and others in the mental health field, to better understand these disorders, their etiology, and their treatment. Other disorders that also have a fundamental feature of choice dysregulation, such as eating disorders and sexual disorders, should also be considered in this categorization (in future iterations of DSM), as their biology and social circumstances can have features that overlap with the SUDs as well as pathologic gambling.

The remainder of this introduction will briefly review some general topics related to SUDs—highlighting the diagnosis of these disorders, the extent of use, and public policy issues. Details of specific disorders, including their treatment can be found in the subsequent chapters.

The Definition of Substance Use Disorders

A significant change with DSM-5 is that there no longer are separate diagnoses of substance abuse and substance dependence. DSM-5 essentially merges the criteria for these two diagnoses, and has added another criterion related to craving (which was not in either set of diagnostic criteria for DSM-IV). These changes have resulted in essentially eleven criteria that can be used to make a diagnosis of an "SUD," although it is worth noting that the specific criteria can vary slightly from one class to the next. Modifiers then indicate if the severity is mild (2 to 3 criteria), moderate (4 to 5 criteria), or severe (6 or more criteria). While at the time of this writing it is still somewhat early in the use of DSM-5, it appears that clinical programs, insurers and researchers are using a cutoff of moderate or severe as the necessary severity for study protocols and the authorization of more intensive levels of treatment services. Additional modifiers are also available, for example, on maintenance therapy for opioid use disorders if the patient is being treated with a medication such as methadone or buprenorphine.

While clinicians are most familiar with the DSM-5 criteria, other criteria and coding for diagnoses are used in clinical practice and for billing purposes—most notably, the ICD-10 coding. ICD-10 differs from DSM-5 in several ways, including the inclusion of problematic use categories of "harmful use" and "dependence syndrome." ICD-10 also has more drug use categories—for example, separating cocaine from "other stimulants, including caffeine" (which then creates the somewhat unusual inclusion of caffeine with amphetamines). Finally, while DSM-5 SUDs tended to consolidate diagnostic categories relative to DSM-IV, ICD-10 tends to create a large number of subcategories for conditions. For example, there are eight different modifiers that can be used with acute intoxication for a substance.

Finally, while not diagnostic, the field of addictions research and clinical care continues to use various forms of quantity and frequency of substance use as a metric for determining risk of use, treatment needs, and study outcomes. Such work has primarily focused on alcohol, the identification of standard drinks of alcohol, and daily numbers of such drinks that increase the risk of adverse consequences to use. This work has helped to move the alcohol field forward in its use of the concept of heavy drinking days (HDDs), which can be a useful marker in studies of treatment outcome. While there are limitations to extending these concepts to other drugs of abuse (e.g., the problems associated with creating a definition of a standard

amount of cocaine or cannabis that is used), the overall concept of quantity and frequency of use may have utility for both clinical screening and research purposes, and this is a potential area of development in the coming years.

The Extent of Substance Use Disorders

In the United States, annual surveys of substance use are conducted under the sponsorship of a federal agency, the Substance Abuse and Mental Health Services Administration (SAMHSA). These surveys date back to the 1970s and for many years were titled the "National Household Survey on Drug Abuse," although the name of these reports was changed to the "National Survey on Drug Use and Health" (NSDUH) as of 2002. Subsequent chapters will address the prevalence of specific types of substances used; however, some overall findings will be noted here based on the 2015 NSDUH results. (While the focus here in this brief review is on NSDUH data, other surveys in the United States provide complementary data that can be useful for the reader interested in gaining a greater knowledge on drug abuse epidemiology. These include the Monitoring the Future [MTF] study of students and young adults, the Treatment Episode Data Set [TEDS], which examines patterns of substance abuse treatment admissions, the National Survey of Substance Abuse Treatment Services [N-SSATS], which assesses treatment programs, and other large studies of drug abuse such as the National Epidemiologic Survey on Alcohol and Related Conditions [NESARC].)

In 2015, 130,610,000 persons in the United States (12 years and older) reported a lifetime history of illicit drug use. The most commonly used illicit drug is marijuana and hashish (117,865,000 persons with lifetime use). While these findings demonstrated that there is substantial experimentation with illicit drugs, past month use of any illicit drug (27,080,000), cannabis (22,226,000), and nonmedical use of psychotherapeutics (6,365,000) generally represents less than 20 percent of those who have ever used an illicit drug. While this should by no means minimize the extent of drug use, it does demonstrate that there are a large number of persons who try a drug, but either do not go on to problematic use or who are able to surmount problematic use and stop their drug use.

When comparing findings to earlier surveys, there is evidence that past month use has increased for some substances (e.g., cocaine, cannabis), although all increases are not statistically significant, and that past month heroin use has decreased.

The extent of tobacco and alcohol use is markedly higher than illicit drug use, according to the 2015 NSDUH (as well as earlier years of the survey). There were 171,120,000 persons with lifetime tobacco use, and 63,959,000 with past month tobacco use (23.9 percent of the population aged 12 or older). For alcohol, there were 216,839,000 persons with

lifetime use, and 138,322,000 with past month use (51.7 percent of the 12 year and older population). Notably, 66,690,000 persons (24.9 percent of the population) had an episode of binge alcohol use in the prior 30 days (defined as 5 or more drinks on one occasion).

Other national surveys are conducted by the European Union, and Australia; the former conducts these under the auspices of the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), and the latter is conducted by the Australian Institute of Health and Welfare (AIHW). In general, findings from Australia, Europe, and the United States are similar (e.g., that there are high rates of cannabis use as the most commonly used illicit drug, that rates of prescription drug use have been a growing concern, and that tobacco and alcohol use are more extensively used than illicit drugs). When making cross-regional comparisons (which are not easily done for a variety of reasons, such as the tools that are used), it is notably that the United States reports on drug use for persons 12 and older, the EU for persons 15 and older, while Australia uses a cutoff of 14 and older.

Finally, while the number of persons or rates of use capture one aspect of substance use, the impact of this use can be lost in such numbers. The Global Burden of Disease study looked at a variety of medical conditions, worldwide, and found that two of the top five leading causes of disease were substance abuse related—tobacco (fourth) and alcohol use (fifth). Furthermore, in developed parts of the world, illicit drug use was the eighth leading cause of disease, according to this survey. While the extent of use is one measure of impact, this evidence of disease burden is a telling measure of the significant effect that drug use has on individuals, their families, and societies.

Treatment

It is beyond the scope of this introduction to address the treatment of SUDs in any substantial manner. However, a few general points regarding substance abuse treatment are worth emphasizing here.

First is that treatment is effective. This efficacy impacts use of the drug itself, as well as showing an impact on other aspects of lifestyle and health. Second, there is some evidence to suggest that treatments that combine medications with nonpharmacologic interventions are more effective than medications alone, although there can be significant efficacy for medications alone compared to no treatment or a placebo. Finally, there is evidence that contingency management can be a particularly useful and effective treatment for addictive disorders, with a good evidence base for its use.

Public Policy Issues Related to Substance Use Disorders

Perhaps the most dramatic change in recent years, as it relates to drug abuse public policy, is the rapidly evolving move of cannabis from an illicit drug to either a potential therapeutic ("medical marijuana"), or to legalization of any use. This changing legal and social landscape has been rapidly occurring in the United States, and varies considerably at the state level with some states maintaining cannabis as an illegal drug, and others allowing varying degrees of use. The medical use of cannabis has occurred with little of the traditional evidence base that would provide the foundation for determining the efficacy and safety of a new medication, while the greater availability of legal use for pleasurable purposes may have unintended consequences (e.g., medical complications from acute overuse, greater use by school-aged children, etc.). Issues of cannabis product purity and consequences to cannabis use are current topics of high interest.

Other policy issues related to substance abuse have persisted for several years, and include topics such as the balance of treatment availability versus quality, the role of harm reduction as a social intervention, the balance between social disapproval of illicit drug use (as one aspect of prevention) and encouragement that a drug user enter treatment (the issue of stigma), and the balance between medicalization of substance abuse treatment and the role of self-help groups in making a major impact on helping those in need.

Summary

SUDs remain highly prevalent disorders that exact a social burden as well as a personal burden on the individual. This topic area spans aspects of public policy, neurobiology, ethics, and behavior, and it is likely that addictions will continue to be a major topic area of concern for years to come (as they have been for thousands of years in the past). Of note, the growing body of knowledge about the biological and environmental aspects of these disorders, and how these aspects impact experimentation and problematic or pathologic use, have helped to bring SUDs under a medical umbrella and moved forward prevention and treatment efforts. However, much more work needs to be done in our understanding of treatment of these disorders. The following chapters in this section provide an overview to these disorders, their neurobiology and social circumstances to use, the behavioral aspects of reinforcement, the extent of substance use, and the treatment of these disorders.

References

Agboola S, McNeill A, Coleman T, Leonardi Bee J. A systematic review of the effectiveness of smoking relapse prevention interventions for abstinent smokers. *Addiction*. 2010;105:1362–1380.

Amato L, Minozzi S, Davoli M, Vecchi S, Ferri M, Mayet S. Psychosocial combined with agonist

maintenance treatments versus agonist maintenance treatments alone for treatment of opioid dependence. *Cochrane Database Syst Rev*. 2008;(4):CD004147.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC; 2013.

Australian Institute of Health and Welfare. National Drug Strategy Household Survey detailed report 2013. *Drug statistics series no. 28. Cat. no. PHE 183* Canberra: AIHW. Available at http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129549848. Accessed October 22, 2016.

Ball JC, Ross A. *The Effectiveness of Methadone Maintenance Treatment*. New York: Springer-Verlag; 1991.

Bogenschutz MP, Johnson MW. Classic hallucinogens in the treatment of addictions. *Prog Neuropsychopharmacol Biol Psychiatry*. 2016;64:250–258.

Dutra L, Stathopoulou G, Basden SL, Leyro TM, Powers MB, Otto MW. A meta-analytic review of psychosocial interventions for substance use disorders. *Am J Psychiatry*. 2008;165(2):179–187.

EMCDDA. 2012 Annual report on the state of the drugs problem in Europe. Accessed at http://www.emcdda.europa.eu/publications/annual-report/2012. Lisbon: European Monitoring Centre for Drugs and Drug Addiction; 2012.

Ezzati M, Lopez AD, Rodgers A, Vander Hoorn S, Murray CJ. Selected major risk factors and global and regional burden of disease. *Lancet*. 2002;360:1347–1360.

Falk D, Wang XQ, Liu L, et al. Percentage of subjects with no heavy drinking days: evaluation as an efficacy endpoint for alcohol clinical trials. *Alcohol Clin Exp Res*. 2010;34:2022–2034.

Gowing L, Ali R, White JM. Buprenorphine for the management of opioid withdrawal. *Cochrane Database Syst Rev*. 2009;(3):CD002025.

Hartman RL, Brown TL, Milavetz G, et al. Cannabis effects on driving lateral control with and without alcohol. *Drug Alcohol Depend*. 2015;154:25–37.

Mattick RP, Breen C, Kimber J, Davoli M. Methadone maintenance therapy versus no opioid replacement therapy for opioid dependence. *Cochrane Database Syst Rev*. 2009; (3):CD002209.

Mattick RP, Breen C, Kimber J, Davoli M. Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. *Cochrane Database Syst Rev*. 2014; (2):CD002207.

NIAAA. *Helping Patients Who Drink Too Much: A Clinician's Guide* Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism. NIH Publication No. 07–3769; 2007.

Rehm J, Taylor B, Room R. Global burden of disease from alcohol, illicit drugs and tobacco. *Drug Alcohol Rev*. 2006;25:503–513.

SAMHSA. *Results from the 2015 National Survey on Drug Use and Health: Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2013. Available at http://www.samhsa.gov/data/sites/default/files/NSDUH-DetTabs-2015/NSDUH-DetTabs-2015/NSDUH-DetTabs-2015.htm#tab2-1a. Accessed on October 21, 2016.

Schwartz RP, Highfield DA, Jaffe JH, et al. A randomized controlled trial of interim methadone maintenance. *Arch Gen Psychiatry*. 2006;63(1):102–109.

Stitzer ML, Vandrey R. Contingency management: utility in the treatment of drug abuse

disorders. *Clin Pharmacol Ther*. 2008;83:644–647.

Strain EC, Stitzer ML, Liebson IA, Bigelow GE. Methadone dose and treatment outcome. *Drug Alcohol Depend*. 1993;33:105–117.

Strain EC, Stitzer ML, Liebson IA, Bigelow GE. Buprenorphine versus methadone in the treatment of opioid dependence: self-reports, urinalysis, and addiction severity index. *J Clin Psychopharmacol*. 1996;16:58–67.

Vandrey R, Raber JC, Raber ME, Douglass B, Miller C, Bonn-Miller MO. Cannabinoid dose and label accuracy in edible medical cannabis products. *JAMA*. 2015;313(24):2491–2493.

Volkow ND, Swanson JM, Evins AE, et al. Effects of cannabis use on human behavior, including cognition, motivation, and psychosis: a review. *JAMA Psychiatry*. 2016;73(3):292–297.

▲ 11.2 Alcohol-Related Disorders

MARC A. SCHUCKIT, M.D.

INTRODUCTION

This chapter reviews conditions that affect 20 to 30 percent of middle-class men and about 15 percent of women seeking medical care. The typical individual with these diagnoses has a job, lives with their family, and is relatively high functioning, but is likely to die a decade or more prematurely. Almost half will develop temporary, but clinically significant, anxiety or depression, and they carry an enhanced risk for serious medical problems along with a heightened potential for adverse reactions to medications. The condition that is the focus of this chapter costs the United States almost $225 billion a year. Despite these characteristics, most physicians have received less than 10 hours of training regarding recognizing and treating these disorders, and many physicians, including psychiatrists, feel inadequate in making the diagnosis or treating these alcohol-related conditions.

Alcohol use disorders (AUDs) have special relevance to psychiatry. Alcohol is a potent drug that causes both acute and chronic changes in almost all neurochemical systems, with the result that heavy drinking can produce serious *temporary* psychological symptoms including depressions, anxiety, and psychoses. Several preexisting psychiatric diagnoses, including antisocial personality, bipolar, and schizophrenic disorders, increase the risk for later AUDs. In light of the importance of alcohol-related conditions to psychiatric practice this chapter offers a broad and clinically focused overview of recent developments in the field of AUDs. The following case facilitates the discussion of alcohol-related problems.

JP is a 47-year-old ophthalmologist who was referred by his internist for psychiatric evaluation. The presenting complaints included depressive and anxiety symptoms, concerns by his wife that he was spending more time by himself in his study in the evenings, and that there were times he appeared to not think clearly. The differential diagnosis included a major

depressive episode, generalized anxiety disorder, and the early stages of a dementing process resembling Alzheimer disease. The problems had become more prominent in the past year for this man who had graduated from both college and medical school with distinction, and who was still married and respected by his wife of 28 years and their young adult children. It soon became apparent that heavy drinking was a major contributor to the observed problems, and that in many ways JP was a typical middle-class person with an AUD.

JP will be used as an example of several major lessons in this chapter.

DEFINITION AND COMPARATIVE NOSOLOGY

Alcohol Use Disorders

Diagnoses in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) are intended to identify clinically significant conditions through criteria that are both reliable and valid. The two major current psychiatric diagnostic systems (DSM-5 and the *International Classification of Diseases* tenth edition [ICD-10] developed by the World Health Organization) agree that an AUD diagnosis requires repeated impairments in multiple life areas as a result of alcohol use. In the DSM-5 criteria developed in 2013, an AUD is diagnosed as the repeated presence of at least two of 11 major areas of life impairments related to alcohol that cluster together in a syndrome over approximately the same 12-month period. The criteria are very similar to those listed for abuse and dependence in DSM-IV, with the exception that one DSM-IV item was dropped (legal problems) and one new item was added (craving).

In ICD-10 the diagnosis of alcohol dependence requires three or more items clustering together. These include the same general concepts as expressed in DSM-5, although they are worded a bit differently. ICD-10 also lists a second potentially less severe diagnosis of *harmful use*, defined as repeated interference with psychological and physical health functioning. Not surprisingly, reflecting differences in the criteria items in the two systems, persons with harmful use are not likely to meet criteria for DSM-5 AUD, and vice versa.

Several approaches have been tested as potential alternative AUD criteria and as efforts to divide AUDs into clinically meaningful subgroups. For example, a more severe early onset of an AUD, accompanied by criminality and problems with other drugs, has been labeled as *type 2* or *type B alcoholism*. This approach reflects the fact that most early-onset medical and psychiatric disorders have a more severe course. However, much of the prognostic significance of type 2 or type B alcoholism rests with concomitant conduct disorder and antisocial personality disorder in early onset substance use disorders. As mentioned elsewhere in this chapter, few advantages have been found for diagnostic items related to drinking patterns to a family history of AUDs.

Severity and Remission

DSM-5 asks clinicians to record the prior year severity of an AUD as *mild* (two to three criteria endorsed), *moderate* (with four to five criteria met), or *severe* (with 6 or more of the 11 criteria), while ICD-10 has no formal notation of severity. Remission criteria in DSM-5 are based on the fact that the highest risk for relapse is in the first 3 to 12 months of recovery. Thus, early remission occurs if none of the 11 AUD criteria (other than craving) were experienced during the prior 3 months, and the notation of sustained remission indicates that no such problems were experienced for at least 12 months.

Although JP did not fit the erroneous stereotype of a person with an AUD as homeless and jobless, he was fairly typical of the middle-class person with this disorder. The history revealed that alcohol problems had become prominent in his mid-20s, with an age of onset of the AUD (i.e., when two or more of the AUD criteria occurred within the approximate same 12 months) of about 28 years. Problems included tolerance; spending a great deal of time drinking (when JP retired to his office at home and spent most of the evening drinking); and repeatedly exceeding preset amounts and periods of time spent drinking (e.g., setting a limit of two glasses of wine after dinner but consuming two bottles). Consistent with his relatively young age and general good health, JP had never experienced a withdrawal syndrome.

A BRIEF HISTORY OF ALCOHOL AND AUDs

Beverage alcohol (ethanol) is a product of fermenting fruit or grain and was probably around before humans walked the earth. Evidence of human consumption of these beverages dates back at least 12,000 years, and alcohol was used as part of ceremonies in Babylon, Greece, and the Roman Empire over 5,000 years ago when this drug was also used for pleasure, nutrition, and for its medicinal properties.

Repeated drunkenness and associated problems have been noted in almost all civilizations, and related stories are seen in the Bible. Most early consumption involved beer and wine, but distillation of beverages with higher alcohol contents, such as whiskey, dates back to about 1400 AD in Ireland. Laws actually encouraging the consumption of spirits as a way to raise tax revenues appeared in England in the 1600s.

The increased use of ethanol and the enhanced prevalence of drunkenness by the early 1800s became a public health concern and a threat to maintaining a reliable workforce, especially in the context of the industrial revolution. Later in the 1800s England tried to reverse course and control alcohol use, and a prominent movement toward total abstinence grew in strength in both Europe and America. Following World War I there were calls for prohibition of alcohol consumption that culminated in national drinking bans, including the 18th amendment to the US constitution (the Volstead Act) in 1920. However, by this point in history alcohol had become so ingrained in most societies that prohibition was difficult to enforce, and beverage alcohol became legal again in the United States in 1933.

Following the end of Prohibition, alcohol-related problems quickly escalated. Currently, problems associated with alcohol cost about $225 billion each year in the United States through lost productivity in the workplace, healthcare expenses related to alcohol, crime, and motor vehicle crashes. The number of people affected and the total costs are much higher than figures associated with any other drug.

EPIDEMIOLOGY

Prevalence of Drinking, Related Problems, and AUDs

About 90 percent of Americans have consumed alcohol at some time in their lives, with most beginning in their middle teens (Table 11.2–1). By the end of high school, 80 percent of students have imbibed alcohol, and more than 60 percent have been intoxicated. During any year, 65 percent of men were drinkers, with a ratio of the proportions of men versus women who consumed alcohol of approximately 1.3 men to 1.0 women. The highest drinking frequency and quantity usually occurs from the late teens to the mid-20s.

Men and women with higher education and income are the ones most likely to imbibe, and, among religious denominations, Jews have among the highest proportions who consume alcohol but also have among the lowest rates of AUDs. Other ethnicities, such as the Irish, have higher rates of severe alcohol problems, but are also likely to have higher rates of abstention. Very high levels of alcohol problems are found among most, but not all, Native American and Inuit tribes, with estimates that more than 60 percent of men and women in some tribes have at some time had an AUD. In the United States, in 2010 the World Health Organization reported that the average person aged 15 or older consumed 2.4 gallons of absolute alcohol in a year, a figure considerably less than the five gallons of absolute ethanol yearly intake in the 1700s. In 2010, per capita consumption ranged from 4.6 gallons in Belarus to 0.05 gallons in Saudi Arabia.

Almost half of drinkers experience temporary alcohol-related problems. One survey noted that the impaired memory for events occurring during a heavy drinking evening (i.e., an alcohol-related blackout) was reported by almost half of drinking men and women, including about 15 percent who had three or more such experiences in any year. Among relatively well-educated middle-class men, about 30 percent admitted to having driven while impaired by alcohol, and about 20 percent had missed school or work because of a hangover or a desire to “party.”

Table 11.2–1.
Alcohol Epidemiology

Condition	Population (%)
Ever had a drink	90

Current drinker	60–70
Temporary problems	40+
Alcohol use disorder (AUDs) lifetime	Male: 15+
	Female: 10+
AUDs in psychiatric patients	30

The percentages listed here are approximations generated across studies.

The lifetime risk for AUDs is approximately 15 percent for men and 10 percent for women, with 1-year prevalence rates of about 6 percent, figures that are applicable across most socioeconomic and educational levels. Rates among individuals coming to a hospital or visiting a physician's office are almost double that. The peak ages of onset of an AUD are from the early 20s to about age 40. Despite multiple difficulties, most people with AUDs have jobs, families, and relatively high levels of functioning, and, thus, few fit the stereotype of a homeless, unemployed person.

The course of AUDs is as predictable as most medical and psychiatric conditions, with an early onset likely to be associated with greater severity. Therefore, if an AUD begins in the mid-teens, it is often associated with conduct disorder, concomitant drug difficulties, and antisocial behaviors that are likely to predate the alcohol problems. At the other extreme, an onset after age 40 tends to be associated with less severe social difficulties and more subtle AUD signs and symptoms, but a greater likelihood of medical problems.

Comorbid Conditions

Temporary psychiatric symptoms are common during intoxication and withdrawal from alcohol. Therefore, to use a diagnosis to indicate prognosis and guide treatment, clinicians need to distinguish between temporary substance-related psychopathology likely to disappear after abstinence versus independent psychiatric syndromes that may require long-term medications.

Up to 80 percent of men and women with AUDs report *temporary* symptoms of sadness or anxiety during the course of their disorder. These syndromes become intense and persistent enough to meet criteria for major psychiatric conditions, such as major depressive episodes or panic disorder, in at least 40 percent. When present, these alcohol-induced conditions need to be recognized and treated in the short term, usually with cognitive-behavioral therapy (CBT)-based approaches. However, when the psychiatric syndrome is only observed in the context of intoxication or withdrawal from a depressant such as alcohol, those symptoms are likely to diminish and disappear within a week to a month of abstinence. Alcohol-induced mood or anxiety disorders do not have the same long-term prognoses nor require the same pharmacological treatments as independent major depressive episodes or anxiety disorders.

PHARMACOLOGY AND ALCOHOL'S EFFECTS

Properties and Metabolism of Alcohol

Ethanol (beverage alcohol) is a simple molecule that is well absorbed through the mucosal lining of the digestive tract in the mouth, esophagus, and stomach. The most prominent area of uptake is in the proximal small intestine, which is also the site of absorption of many B vitamins. Ethanol rapidly enters the bloodstream and, reflecting its high solubility in water, is distributed throughout the body.

Wine, beer, and distilled spirits (e.g., whiskey, gin, and vodka) differ in their content of components other than alcohol. These *congeners*, responsible for much of the taste of the beverage, include methanol, butanol, aldehydes, phenols, tannins, lead, cobalt, and iron.

A standard alcoholic drink in the United States is defined as containing 10 to 12 g of ethanol, which is roughly the amount in 12 oz of beer (containing ~3.6 percent ethanol), 5 oz of table wine (~12 percent ethanol), and about 1.5 oz of 80-proof spirits (~40 percent ethanol). For an average 70-kg (155-lb) person who has average body fat, one drink is likely to raise the blood alcohol level by approximately 15 to 20 mg/dL (the same as 0.015 to 0.020 g/dL), with roughly the same amount metabolized in about an hour. Faster absorption from the digestive tract occurs on an empty stomach and if alcohol is taken as a carbonated beverage.

Between 2 and 10 percent of the alcohol is excreted unchanged through the lungs, kidneys, or sweat, some is broken down in the lining of the esophagus and stomach (especially in men), but most is metabolized in the liver through several pathways that result in the production of acetaldehyde. Most of the process occurs through the actions of alcohol dehydrogenase (ADH) in the cell cytosol, although at high blood-alcohol levels some alcohol is also broken down in the microsomes of the liver's smooth endoplasmic reticulum (the microsomal ethanol oxidizing system [MEOS]). The ADH process is the usual rate-limiting metabolic step, occurring relatively slowly because of the liver's need to handle the produced hydrogen ions through the actions of a cofactor that is in relatively short supply, nicotinamide adenine dinucleotide (NAD).

Acetaldehyde is destroyed by aldehyde dehydrogenase (ALDH) both in liver cell cytosol and mitochondria. This occurs rapidly, so that the average person has low levels of this substance after drinking. This is fortunate because, while low levels of acetaldehyde can be both stimulating and reinforcing, at high levels this substance can release histamines and catecholamines that contribute to rapid changes in blood pressure along with nausea and vomiting. Variants of ALDH and ADH are related to a higher reaction to alcohol and lower risk for AUDs, especially in Asian (e.g., Japanese, Chinese, and Korean) men and women.

Alcohol as a Depressant Drug

Acutely, all intoxicants pleasurably change how a person feels, thus enhancing the drive to continue to take the drug despite potential consequences. Such drugs have divergent effects that are best understood by placing them into categories based on their most prominent effects at the usual doses taken. In this scheme, depressants are drugs that produce somnolence and decreased neuronal activity, but are not powerful in attenuating pain. Thus, alcohol, benzodiazepines, barbiturates, and related drugs are considered depressants and produce similar intoxications, are potentially lethal in overdose (especially when multiple depressants are taken together), have cross-tolerance with other depressants, and can produce physical dependence with similar withdrawal syndromes.

Neurochemical Effects of Ethanol

Alcohol has prominent effects on almost all neurochemical systems, with opposite actions during intoxication versus withdrawal. Perhaps the most potent effects are on γ-*aminobutyric acid* (GABA) complexes in the brain, especially on the GABA-A receptor ($GABA_A$), contributing to the sedating, sleep-inducing, anticonvulsant, and muscle-relaxing properties of alcohol. Ethanol also impacts glutamate-gated ionophoric receptors, especially those that bind *N-methyl-d-aspartate* (NMDA), with dampened stimulatory effects during intoxication and heightened activity during withdrawal. Alcohol acutely increases *dopamine* and its metabolites, and chronic drinking changes dopamine receptor numbers and sensitivity, with related effects on intoxication and craving, in part through changes in the pleasure centers in the ventral tegmental area of the brain. Consuming beverage alcohol also increases *serotonin* in the synapses and upregulates serotonin receptors. Lower brain levels of serotonin may be associated with a less intense response to alcohol and with consuming more alcohol per occasion. Finally in this brief review, alcohol acutely enhances the functioning of the *opioid-related* brain systems and impacts *adenosine*, *acetylcholine*, and *cannabinoid* 1 (CB1) receptors.

Tolerance

With repeated administration, larger doses of alcohol are required to produce the same effects. This phenomenon of *tolerance* also produces an ability to tolerate higher doses of the substance through at least three processes. In *behavioral tolerance* a person learns through practice how to perform tasks effectively while experiencing the effects of alcohol. *Pharmacokinetic tolerance* involves adaptations of alcohol metabolizing systems to rid the body of alcohol more rapidly. Most importantly, *pharmacodynamic* or *cellular tolerance* reflects an adaptation of the nervous system so that it can function despite high blood alcohol

concentrations by resisting the actions of alcohol on the cell.

Once tolerance has developed to one depressant, an individual is likely to demonstrate a similar reaction to a second drug of that class (*cross-tolerance*). Therefore, a person who has been drinking heavily, has tolerance to alcohol, and then stops drinking can be expected in the short run to require a higher dose of other depressants such as benzodiazepines for sleep induction. However, if individuals take two depressant drugs concomitantly, the effect of one drug can magnify, or potentiate, the effects of the other, with a possible lethal outcome.

FT is a 73-year-old emeritus professor at a small liberal arts college in the Midwest. He had no persistent alcohol-related problems prior to his early 60s, and used to consume one or two standard drinks (scotch) before dinner. At age 62 he moved to a retirement community in Arizona where he had more time to play golf, interact with his wife, and socialize with his new friends. In the context of a lack of work obligations and the freedom to increase his hours of socializing, FT slowly increased his alcohol intake. Subsequently, the number of standard drinks required to produce a mild buzz increased from two to four during an evening, and his daily consumption increased to five or six standard drinks most nights.

Blackout

A *blackout* is a memory impairment (*anterograde amnesia*) for the period when a person was drinking heavily but remained awake. This phenomenon is the result of the ability of high doses of any brain depressant (e.g., alcohol or benzodiazepines) to interfere with the acquisition and solidification of memory. As many as 50 percent of drinkers have at some time had a blackout, and repeated experiences of this effect of high doses of alcohol is a warning sign that an AUD may be present.

Sleep Impairment

Alcohol can help a person fall asleep, but after two or more drinks, the sleep pattern can be impaired. A heavy drinker is likely to awaken after several hours and have problems going back to sleep. Alcohol suppresses rapid eye movements (REM) sleep, inhibits stage 4 sleep late in the night, and is associated with frequent alternations between sleep stages (*sleep fragmentation*), with more intense and disturbing dreams late in the night as the blood alcohol levels fall. More intense forms of these problems are seen in individuals with AUDs for whom sleep stages might not return to normal until three or more months of abstinence. More sleep problems following abstinence are associated with a greater risk for relapse.

Cerebellar Degeneration

Characterized by unsteady gait, impaired standing steadiness, and mild nystagmus, cerebellar degeneration results from combinations of effects of

ethanol, acetaldehyde, and vitamin deficiencies. Treatment requires total abstinence and vitamin supplementation, although complete recovery is not usual.

Other Effects on the Central Nervous System

A thiamine deficiency, especially in the context of a pre-existing vulnerability such as a transketolase deficiency, can present as any of several neurological syndromes now known as alcohol-induced major neurocognitive disorder, amnestic confabulatory type, persistent. Using the more widely accepted prior terminology, this can include a sixth cranial nerve palsy (Wernicke's) and a severe anterograde amnesia that is out of proportion to the general level of confusion (Korsakoff's). These rare conditions are discussed in greater depth in the section "Diagnosis and Clinical Features."

Effects on the Rest of the Body

One to two drinks per day of any alcoholic beverage might decrease the risks for myocardial infarction and thrombotic stroke, probably through decreasing platelet aggregation and enhancing high-density lipoprotein (HDL) cholesterol levels. An additional potential cardioprotective action, unique to red wine, may occur through antioxidant flavonoids or the inhibition of a vasoconstrictor (endothelin-1). Low doses of alcohol may also decrease the risks for some dementias, peripheral arterial disease, and gallstones. However, any amount of alcohol may be harmful to a developing fetus, and the risks of even one drink per day are prohibitive for individuals in recovery from a moderate or severe AUD, people taking medications that may adversely interact with alcohol, and persons with any of several medical or psychiatric syndromes. Consuming more than three drinks a day increases low-density lipoprotein (LDL) cholesterol, triglycerides, and blood pressure, thus increasing the risk of cardiac disorders and cancers.

Peripheral Neuropathy

Approximately 10 percent of people with AUDs develop a deterioration of functioning of nerves to the hands and feet, called *peripheral neuropathy*. The symptoms include numbness of the hands and feet, often bilateral, frequently accompanied by tingling and paresthesias. Although the condition is usually relatively mild and often improves with abstinence, in some individuals the pain and numbness may be permanent.

Gastrointestinal Problems

The gastrointestinal (GI) system can be severely affected by heavy drinking, including an acute and potentially severe inflammation of the esophagus or

stomach, often accompanied by vomiting and bleeding. If vomiting occurs in the presence of dilated esophageal veins, as seen with cirrhosis, it can induce a potentially lethal blood loss.

The liver and the pancreas are also especially vulnerable to alcohol. In the liver, frequent heavy drinking results in the accumulation of fats and proteins in the cells, producing a reversible swelling of the liver often described as a *fatty liver*. Inflammation of liver cells accompanied by increases in liver function tests is called *alcoholic hepatitis*, which can lead to deposition of excessive amounts of hyalin and collagen near blood vessels in the liver, an early stage of *cirrhosis*, with the latter seen in approximately 15 percent of patients with AUDs. As damage progresses and liver scar tissue increases, the normal flow of blood through the liver is impaired, dilated veins (or varices) develop from the increased abdominal venous pressure, and fluid seeps from the liver capsule, accumulating in the abdomen as *ascites*. Subsequent liver failure can cause a cognitive impairment known as hepatic encephalopathy.

Perhaps 10 percent of people with AUDs develop an inflammation of the pancreas or *acute pancreatitis*. This can lead to a chronic irreversible condition with associated impaired sugar metabolism (a form of diabetes).

Cardiovascular Problems

Heavy alcohol intake and associated withdrawal symptoms can increase blood pressure and elevate both LDL cholesterol and triglycerides, enhancing the risk for myocardial infarction and thrombosis. At high doses, alcohol is a striated-muscle toxin with a resulting deterioration in the heart muscle that manifests itself as beating irregularities and signs of heart failure (*alcoholic cardiomyopathy*). Thus, the leading cause of early deaths associated with AUDs is heart disease.

Blood-Producing Systems

Consumption of four to eight drinks or more per day decreases the production of white blood cells and impairs the ability of those cells to migrate to sites of infection. Such drinking can also affect the stem cells that produce the red blood components, increasing the average size of red cells (the mean corpuscular volume [MCV]), and can impair the production of blood platelets.

Cancer

AUDs are associated with high rates of most cancers, especially tumors of the head, neck, esophagus, stomach, liver, colon, lungs, and breast. The association with cancer probably reflects alcohol-related immune system suppression and the direct effects of ethanol and acetaldehyde on mucosal membranes. The heightened rates of malignant tumors remain significant

even when the impact of smoking and poor nutrition is considered, and cancer is the second cause of premature death in individuals with AUDs.

Fetal Alcohol Effects

Alcohol and acetaldehyde have deleterious effects on the developing fetus. Both substances cross the placenta, and in high enough doses can produce fetal death and spontaneous abortion. At least 5 percent of surviving infants of heavy-drinking mothers evidence mixtures of a *fetal alcohol syndrome* that can include low IQ, a small head, low birth weight, facial abnormalities (e.g., flat bridge of the nose, absent philtrum, and epicanthal eye folds), an atrial septal heart defect, and syndactyly. None of these are reversible, and the cognitive defects remain throughout life. Additional less obvious problems are part of a range of *fetal alcohol spectrum disorders* (FASD), including mild to moderate cognitive difficulties (e.g., impaired executive functioning in attention and planning, problems in letter and word fluency, difficulties with impulse control, etc.), and may carry an enhanced risk for SUDs.

Other Problems

High doses of alcohol adversely affect almost all body systems, cutting a decade or more off the lifespan even for people with higher socioeconomic and educational status. Additional body effects related to AUDs include testicular atrophy; bone fractures; a host of eye effects, including cataracts; dental difficulties; skeletal muscle wasting; and greater risk for accidents.

ETIOLOGY

Many factors affect the decision to drink, the development of temporary alcohol-related difficulties, and the development of AUDs. The initiation of drinking depends largely on social, religious, and personality characteristics, although genetic influences linked to impulsivity and risk taking also contribute. However, the factors that influence the decision to drink or those that contribute to temporary problems might be a bit different from those that add to the risk for AUDs.

Genetic factors are responsible for about 60 percent of the proportion of risk for AUDs, and operate through several intermediate characteristics that interact with the environment. While each risk component is discussed separately in this chapter, these divisions are heuristic, as it is the combination of psychological, sociocultural, biological, and other factors that are responsible for the development of severe, repetitive alcohol-related life problems.

Psychological and Sociocultural Theories

Plentiful data support a relationship between *stress* and drinking, but the

strength of the interaction and the directions of the correlations are complex. Retrospective studies focusing on prominent adverse life events often report that higher stress is associated with heavier drinking and related problems. However, prospective studies often find less impact of stress during any one day on alcohol use that same day or in the subsequent 48 hours. In addition, these relationships are bidirectional, with drinking on day 1 also associated with more stress on day 2. At higher doses and especially at falling blood alcohol levels, objective measures of tension in muscles indicate that muscle tension can actually increase. Thus, tension-reducing effects of alcohol may be modest and are most obvious in light-to-moderate drinkers, and stress may play a moderate role in causing AUDs.

A person's *expectations* of the effects of alcohol also play a role in patterns of alcohol use. Positive expectations of the impact of excessive drinking on social interactions, alleviating sadness and anxiety, and on sexual performance are associated with higher probabilities of heavier drinking. Similarly, the more a person expects that alcohol will interfere with relationships, sexual performance, and other negative outcomes, the less the risk for heavy alcohol intake.

A third characteristic closely linked to attitudes toward alcohol and subsequent drinking patterns is the *perceived pattern of drinking among peers* where individuals tend to overestimate the alcohol consumption of friends. Thus, they exaggerate the amount of drinking that is acceptable, healthy, and safe, and may be more likely to engage in heavier drinking themselves.

Additional psychological and sociocultural factors related to heavy drinking include the neighborhood in which one lives and the ease of availability of alcohol, a person's income and the money available to buy alcohol, and their education and religious beliefs. Religious beliefs and practices affect alcohol intake through moderating the effects of other characteristics, with less likelihood that poverty, for example, predicts heavier drinking among individuals with closer religious affiliations.

Sociocultural theories are sometimes based on extrapolations from social groups that have high and low rates of AUDs. For example, Jews tend to introduce young children to modest drinking in religious rituals and eschew drunkenness, and also have low rates of AUDs. However, other groups have similar practices, such as French Catholics, but have higher rates of AUDs. Recent studies have found little evidence that introducing children to modest drinking in the home impacts the risks for heavy drinking and alcohol problems. So, cross-cultural comparisons are useful in generating theories, but it can be difficult to establish key cultural practices contributing to drinking-related disorders.

Psychodynamic Theories

Some psychodynamic theorists hypothesize that some people may drink heavily as part of dealing with self-punitive harsh superegos and to decrease unconscious stress levels. Classic psychoanalytical theory hypothesizes that at least some people with AUDs may have become fixated at the oral stage of development and use alcohol to relieve their frustrations by taking the substance by mouth. However, hypotheses regarding arrested phases of psychosexual development, although heuristically useful, have had little effect on the usual treatment or prevention approaches for AUDs and are not the focus of extensive ongoing research.

Similarly, few studies support an "addictive personality" present in most individuals with AUDs and associated with lack of control of intake of most substances and some foods. Although pathological scores on personality tests are often seen during intoxication, withdrawal, and early recovery, many of these characteristics do not predate AUDs, and most disappear with abstinence. In addition, prospective studies of the offspring of parents with AUDs but no personality disorders still demonstrate high risks mostly for AUDs.

Background on Biological Contributors

Sixty percent of the proportion of risk for AUDs relates to genes that impact on a range of different characteristics. Using a parallel example from medicine, myocardial infarctions are genetically influenced, but in some families the genes relate to high blood pressure; in others the heightened risk is related to genes contributing to high triglycerides; and yet others develop heart attacks through a genetically influenced vulnerability toward cardiomyopathy. Each genetic mechanism is moderated by environmental characteristics such as diet, exercise, and smoking. The following sections highlight data supporting the importance of genetic influences in AUDs, and describe biological mediators (Table 11.2–2).

Table 11.2–2.
Data Supporting Genetic Influences in AUDs

Close family members of individuals with an AUD have a fourfold increased risk.
The identical twin of a person with an AUD is at higher AUD risk than is a fraternal twin.
Adopted-away children of individuals with AUDs have a fourfold increased AUD risk, even if not raised by parents with alcohol problems.

Importance of Genetic Influences in AUDs

Three lines of evidence support genetic influences in AUDs. First, there is a three- to four-fold increased risk for severe alcohol problems in close relatives of individuals with AUDs, with rates increasing with a greater number of such relatives, a greater severity of their AUDs, and a closer genetic relationship. Second, the rate of similarity for severe alcohol-

related problems is higher in identical twins of persons with an AUD than in fraternal twins. Third, in adoption studies the risks for AUDs in the offspring of AUD parents is still high even if they were adopted away and knew nothing of the problems in their biological family.

Biological Mediators of the AUD Risk

As shown in Table 11.2–3, a series of independent characteristics impact the alcoholism risk.

Alcohol Metabolizing Enzymes. Variations in genes on chromosome 12 that control ALDH and genes on chromosome 4 related to ADH can decrease the alcoholism risk. The most potent polymorphism occurs in the mitochondrial low Km ALDH2 gene reflecting the *ALDH2*2* genotype that is seen in about 40 percent of Japanese, Chinese, and Korean individuals. An *ALDH2*2*, *2*2* homozygote inherits a disulfiram- (or Antabuse)-like aversive reaction to alcohol because their enzyme cannot efficiently metabolize acetaldehyde, with resulting nausea, vomiting, and unstable blood pressure after one or two drinks and an AUD risk close to zero. Heterozygotes for the ALDH mutation (e.g., with *ALDH2*2, 2*1* alleles) have a mild to modest facial flush after drinking, a higher heart rate, and a more intense (although not necessarily aversive) response to alcohol that contributes to a modest decreased risk for AUDs. Variations in ADH genes (e.g., ADH1B and ADH1C) are more prevalent among Asian, black, and Jewish individuals who experience a slight increased rate of breakdown of alcohol with a possible increase in acetaldehyde and a modest protective effect for AUDs. These gene variations do not affect the risk for other SUDs.

Impulsivity and Disinhibition. A second genetically influenced mechanism increases the risk for both AUDs and SUDs through genes that affect impulsivity, sensation seeking, and disinhibition. These characteristics are more prevalent in individuals with AUDs, are observed in a substantial minority of their children, and reflect the action of a range of genes, including those coding for $GABA_A$ and CHRM2 receptors. Extreme impulsivity and disinhibition characterize antisocial personality disorder where the risk for AUDs and SUDs is about 80 percent.

Independent Psychiatric Disorders. A third genetically influenced mechanism that increases both AUD and SUD risks operates through genes that enhance the vulnerability to several psychiatric disorders. In addition to antisocial personality disorder, two genetically influenced psychiatric conditions closely tied to AUDs and SUDs are schizophrenia and bipolar disorders. These relationships might reflect the poor judgment associated with these major psychiatric disorders and the impulsivity associated with mania. The link might also relate to an overlap of genes contributing to the

predisposition toward substance-related problems and those that impact on schizophrenia and bipolar disorder. Some of the relationship might also occur as an individual attempts to moderate his or her psychiatric symptoms or the side effects of medications through alcohol or illicit drugs.

Table 11.2–3.
Some Genetically Influenced Characteristics That Affect the AUD Risk

Alcohol metabolizing enzymes
Impulsivity and disinhibition
Some psychiatric disorders (e.g., schizophrenia and bipolar disorders)
A low level of response (low LR) to alcohol

The Level of Response (or Sensitivity) to Alcohol. A fourth genetically influenced risk factor for AUDs (but not SUDs) involves how a person responds to alcohol. This includes enhanced risks for people who experience greater stimulation at rising blood alcohol levels and those who demonstrate a low level of response (a low LR) at peak and falling blood levels regarding feelings of intoxication and sedating effects of alcohol. The most thoroughly studied of these characteristics is the low LR where the need for higher doses to achieve alcohol's effects contributes to larger quantities consumed per occasion. This subsequently affects the choice of heavy drinking friends, enhances positive expectations of the effects of heavy drinking, and is associated with the use of alcohol to deal with stress. Several genes that may contribute to the low LR have been identified, and functional magnetic resonance imaging (fMRI) studies have highlighted a possible cognitive inefficiency that might make it more difficult for some people to recognize the effects of low alcohol doses. The data indicating that this characteristic is genetically influenced and the identification of some environmental mediators of the effect of a low LR on heavy drinking are being used to develop programs to prevent heavy drinking and AUDs in individuals with low LRs before they develop.

DIAGNOSIS AND CLINICAL FEATURES

Alcohol use disorders are among the most common of the serious life-threatening behavioral or psychiatric syndromes, and their diagnosis requires a high index of suspicion for the disorder *in every patient*. The average man or woman presenting with severe and repetitive alcohol problems is likely to be neatly dressed, show no signs of severe alcohol withdrawal, have a job and a family, and to complain of a variety of physical conditions or temporary but potentially severe psychiatric complaints. Thus, the clinician must gather a history of alcohol-related life problems from every patient and, if possible, an additional informant, and try to determine whether alcohol has contributed to the presenting complaint. Table 11.2–4 lists the alcohol-related disorders in DSM-5 and a

comparable listing from ICD-10.

Diagnosing Alcohol-Related Disorders

This generic term includes alcohol use disorders (previously abuse and dependence in DSM-IV), as well as alcohol intoxication, withdrawal, and alcohol-induced conditions.

Alcohol Use Disorders

The key to a diagnosis of an AUD is heavy and repetitive use of alcohol to the point of developing recurrent problems. Quantities and frequencies of alcohol intake, while useful, are by themselves not sufficient to establish the diagnosis because of the large variation in the amounts of alcohol required for problems across men versus women, larger versus smaller individuals, and older versus younger people. Repetitive problems indicate that a person keeps using alcohol despite problems, which, in turn, predicts the probability that a return to drinking is likely to be associated with the re-establishment of problems.

Table 11.2–4.
DSM-5 Alcohol-Related Disorders and Corresponding ICD-10 Disorders

DSM-5	ICD-10 Corollary
Alcohol use disorders	**Alcohol dependence syndrome**
	Alcohol harmful use
Alcohol-induced disorders	
Alcohol intoxication	**Acute intoxication due to use of alcohol**
	Uncomplicated
	With trauma or other bodily injury
	With other medical complications
Alcohol intoxication delirium	With delirium
	With perceptual distortions
	With coma
	With convulsions
No DSM-5 equivalent	**Pathological intoxication**
Alcohol withdrawal	**Alcohol withdrawal state**
Specify if: With perceptual disturbances	Uncomplicated
	With convulsions
Alcohol withdrawal delirium	**Alcohol withdrawal state with delirium**
	Without convulsions
	With convulsions
Alcohol-induced persisting neurocognitive disorder	**Dementia**[a]
Alcohol-induced major neurocognitive disorder, amnestic type, persisting	**Amnesic syndrome**
Alcohol-induced psychotic disorder	**Psychotic disorder**
Specify if: With onset during intoxication/with onset	Schizophrenia-like

during withdrawal	Predominantly delusional Predominantly hallucinatory Predominantly polymorphic Predominantly depressive symptoms[b] Predominantly manic symptoms[b] Mixed Or **Organic delusional (schizophrenia-like) disorder**[a] Or **Organic hallucinosis**[a] Or **Residual and late-onset psychotic disorder**
Alcohol-induced depressive disorder *Specify* if: With onset during intoxication/with onset during withdrawal	**Organic mood (affective) disorder**[a,b] Or Residual affective disorder
Alcohol-induced anxiety disorder *Specify* if: With onset during intoxication/with onset during withdrawal	**Organic anxiety disorder**[a]
Alcohol-induced bipolar disorder *Specify if:* With onset during intoxication/with onset during withdrawal	**No ICD-10 equivalent**
Alcohol-induced sexual dysfunction *Specify* if: With onset during intoxication or during withdrawal	**Unspecified alcohol-related disorder**
Alcohol-induced sleep disorder *Specify* if: With onset during intoxication/with onset during withdrawal	**Unspecified alcohol-related disorder**
Unspecified alcohol-related disorder	**Other mental or behavioral disorder induced by alcohol**[c] **Unspecified mental or behavioral disorder induced by alcohol**

DSM-5, Diagnostic and Statistical Manual of Mental Disorders, 5th edition; ICD-10, International Classification of Diseases, 10th edition.

[a]These ICD-10 diagnoses are found in the section on organic mental disorders.

[b]These syndromes are only roughly similar in the two systems.

[c]Includes alcohol sexual dysfunction and alcohol sleep disorder.

As noted above, there are 11 elements in the DSM-5 definition of AUDs. Ten of these items are the same as those listed in DSM-III-R and in DSM-IV. The 11th item, craving, is linked to the pattern of continuing to use alcohol despite past problems with the drug. Physical dependence (i.e., developing a withdrawal syndrome during abstinence) is a diagnostic item, but is not required for a diagnosis.

JP's referring clinician was concerned that little progress was being made with the patient's mood swings and somatic complaints. Therapy, primarily using a cognitive and behavioral

approach, had been instituted 6 months previously for the depressive symptoms and irritability that were present most days for the prior year. JP had recently undergone a routine physical examination at work and was noted to have modest hypertension (blood pressure of 145/95) along with blood test results that included a mean corpuscular volume (MCV) of 92.5 μm^3 and a γ-glutamyltransferase level of 53 U/L.

On interview, the patient admitted that his wife had been complaining for several years about his drinking pattern. She had emphasized that the mornings after drinking he was more irritable than usual, and complained about his restless sleep. Dr. P had recognized potential alcohol problems and tried on several occasions to cut back on his drinking. These efforts usually resulted in abstinence of between 2 and 4 weeks, followed by several months of setting limits on his intake, but these always gave way to increased drinking during celebrations, business trips, vacations, or to "deal with" stress. Once ad-lib drinking began, he had to increase the amount consumed per day to achieve the desired effects.

Alcohol Intoxication

The DSM-5 criteria for alcohol intoxication require recent ingestion of ethanol, associated impairment, and at least one of six signs or symptoms. The ICD-10 criteria are similar, listing seven signs of intoxication, not all of which are seen in DSM-5.

As a conservative approach to identifying blood levels that are likely to have major effects on driving abilities, the legal definition of intoxication in the United States requires a blood concentration of 80 mg of ethanol per dL of blood (mg/dL), which is the same as 0.08 g/dL. Many other countries define legal intoxication regarding driving as 40 mg/dL, and the limits in Poland are 20 mg/dL. For most people, blood levels as low as 20 to 30 mg/dL are associated with some behavioral changes, a slowing in motor performance, and a decrease in the ability to think clearly, as shown in Table 11.2–5. Blood concentrations between 80 and 200 mg/dL produce more severe problems including severely impaired coordination (ataxia), lability of mood, and progressively greater levels of cognitive deterioration. Blood alcohol concentrations of 200 to 300 mg/dL produce severe slurring of speech and en bloc *anterograde amnesia* (*alcoholic blackouts*) regarding much of the evening. Further increases in blood levels result in the first level of anesthesia, and the nontolerant person who reaches 400 mg/dL risks respiratory failure, coma, and death.

Table 11.2–5.
Impairment Likely to Be Observed at Different Blood Alcohol Concentrations

Level	Likely Impairment
20–30 mg/dL (0.02–0.03 g/dL)	Slowed motor performance and decreased thinking ability
30–80 mg/dL	Increases in motor and cognitive problems
80–200 mg/dL	Increases in incoordination and judgment errors
	Mood lability
	Deterioration in cognition
200–300 mg/dL	Nystagmus, marked slurring of speech, and alcoholic blackouts

>300 mg/dL	Impaired vital signs and possible death

Alcohol Withdrawal

If a person has been drinking heavily over a prolonged period, a rapid decrease in blood alcohol levels can produce a *withdrawal syndrome* characterized by symptoms that are primarily the opposite of what was initially experienced with intoxication. Here, withdrawal includes a coarse tremor of the hands, insomnia, anxiety, and increased blood pressure, heart rate, body temperature, and respiratory rate, as described in Table 11.2–6. In ICD-10, the criteria for alcohol withdrawal are similar, with some differences in the specific items listed and the number of signs required (i.e., three) to make a diagnosis.

Withdrawal phenomena usually begin within about 8 hours of abstinence, peak on the second or third day, and diminish by the fourth or fifth day. The symptoms persist in a more mild form for as long as 3 to 6 months as part of a protracted withdrawal syndrome, the symptoms of which might contribute to relapse.

Mr. T is a 64-year-old lawyer with a 35-year history of an AUD. Alcohol-related problems included tolerance, spending a great deal of time using alcohol (despite a relatively successful career), repeatedly giving up important business and family events because he was too intoxicated or hung over, and significant interference with his relationship with his wife. She complained about his sarcastic wit and lack of care about her interests while drinking heavily.

Mr. T had been able to stop drinking on multiple occasions in the past, at which times he experienced a tremor, anxiety, and problems sleeping for several days. He typically self-medicated these symptoms with five or more 10-mg capsules of chlordiazepoxide (Librium) per day, using pills that had been prescribed for him by his general practitioner in response to his complaints of anxiety and insomnia. Recently, Mr. T was diagnosed with moderate hypertension and adult-onset diabetes, and his family practitioner urged him to stop drinking. In recognition of his age and his associated medical problems, it was believed that the current withdrawal would be best treated in a medical setting.

Table 11.2–6.
Key Elements of DSM-5 Criteria for Alcohol Withdrawal[a]

A. Recent decreasing or stopping regular heavy drinking
B. Accompanied by the development of at least two of the following:
   1. Increased pulse, blood pressure, and/or sweating
   2. Tremor of the hands
   3. Problems sleeping
   4. Gastrointestinal upset
   5. Hallucinations
   6. Inability to sit still
   7. Nervousness
   8. Seizures

[a]A short version of the criteria are offered here. See DSM-5 for complete criteria.

An examination confirmed the withdrawal syndrome along with a mildly decreased hematocrit and platelet count. The evaluation, carried out about 12 hours after his most recent drink, revealed the smell of alcohol on his breath, and a blood alcohol level of 50 mg/dL. The patient demonstrated a prominent tremor of both hands along with a pulse of 110 beats per minute, a respiratory rate of 25 breaths per minute, and a blood pressure of 150/96. Mr. T complained of feeling agitated and noted that while he was very tired he was unable to sleep, but was otherwise alert and oriented. These symptoms improved greatly by day five of abstinence, with mild levels of autonomic overactivity and insomnia gradually disappearing over the next 2 months.

Alcohol Withdrawal Delirium

Less than 5 percent of alcohol intoxications and withdrawals are accompanied by severe cognitive impairments known as alcohol intoxication delirium and alcohol withdrawal delirium. These conditions require that the patient meet criteria for both delirium and DSM-5 alcohol intoxication or withdrawal.

FT, the retired professor discussed elsewhere in this chapter, went on to experience alcohol withdrawal delirium. He was in relatively good health when he entered the hospital for an elective hernia repair. The brief history in the chart offered no comments regarding his drinking pattern and made no mention of his γ-glutamyltransferase value of 55 U/L and MCV of 93.5 μm^3. Eight hours post-surgery the nursing staff noted an increase in his pulse to 110, a blood pressure of 150/100, prominent diaphoresis, and a bilateral hand tremor just before the patient had a grand mal seizure. He awoke agitated and disoriented. Re-evaluation of the history documented an AUD with a consumption of approximately six standard drinks per night. Over the following 4 days the patient's autonomic nervous system dysfunction decreased and his cognitive impairment disappeared.

Alcohol-Induced Psychiatric Disorders

Alcohol-induced psychiatric disorders are clinically meaningful (e.g., interfere with functioning and/or cause severe distress) psychiatric syndromes (e.g., depression) that resemble independent psychiatric disorders (e.g., major depressive disorders), and are in excess of symptoms usually associated with alcohol intoxication or withdrawal. They develop during, or soon after, intoxication or withdrawal from alcohol.

For men and women presenting with psychiatric symptoms (e.g., anxiety, depression, or psychoses) with concomitant alcohol-related problems, the first step in establishing a diagnosis of an alcohol-induced disorder is to obtain a chronological history of the conditions. The focus should be on full psychiatric disorders, not subsyndromal sadness or nervousness, because the full syndromes have the clearest prognostic implications. In this process, a timeline can be useful by noting: (1) the age of onset of alcohol problems severe and repetitive enough for a diagnosis of an AUD; (2) periods of abstinence of a month or more since the AUD onset to determine if a full psychiatric syndrome ever occurred outside the context of problematic alcohol use; and (3) the ages when the patient met criteria for the major psychiatric disorder. If a review of the timeline

reveals no evidence that the additional psychiatric syndromes either clearly antedated the severe alcohol problems or persisted for a month or more after abstinence, an AUD is the major disorder. The additional psychiatric syndromes are likely to be temporary conditions that occurred during alcohol intoxication or withdrawal (i.e., are alcohol-induced). DSM-5 lists substance-induced conditions within the clinically relevant sections (e.g., within depressive, anxiety, and psychotic disorders sections). Independent psychiatric conditions in individuals with AUDs are discussed elsewhere in this chapter, including in the section on Differential Diagnoses.

Alcohol-Induced Depressive and Bipolar Disorders. In the context of heavy and repetitive intake of any brain depressant, symptoms of severe depression are common and may be labeled as an *alcohol-induced depressive disorder* (see Table 11.2–4). ICD-10 describes this and most substance-induced syndromes in the section on organic mental disorders, labeled as an organic mood [affective] disorder. This condition includes the same symptoms observed in major depressive disorders, but the alcohol-induced intense sadness markedly improves within several days to 1 month of abstinence without antidepressant medications. Eighty percent of people with AUDs report intense depressions, including about 40 percent who were depressed for ≥2 weeks at a time, but only about 15 percent of such men and women have ever had a depression that meets the criteria for a major depressive disorder when they have not been drinking heavily.

Few data are available regarding alcohol-induced bipolar disorder. This is a new diagnosis in DSM-5 for which the clinician is asked to determine if the onset of a mania was during intoxication or withdrawal. Here the criteria are basically the same as for mania, and, similar to bipolar diagnoses in general, this diagnosis should not be made if the mania developed in the context of an antidepressants or electroconvulsive therapy. It is best to limit this label to bipolar I diagnoses because hypomanic behaviors are very similar to the usual symptoms and signs of repeated intoxication and withdrawal.

A consultation was requested on a 42-year-old woman with an AUD who complained of persisting severe depressive symptoms despite days of abstinence. In the initial stage of the interview, she stated that she had "always been depressed" and believed that she "drank to cope with the depressive symptoms." Her current complaint included a prominent sadness that had persisted for several weeks, difficulties concentrating, initial and terminal insomnia, and a feeling of hopelessness and guilt. In an effort to distinguish between an alcohol-induced mood disorder and an independent major depressive episode, a timeline-based history was obtained. This focused on the age of onset of a DSM-5 AUD, periods of abstinence that extended for at least a month since the onset of the AUD (periods where independent depressions might have been observed), and the ages of occurrence of clear major depressive episodes lasting 2 or more weeks. Despite this patient's original complaints, there was no evidence of a major depressive episode before her mid-20s when her AUD began, and that during a 1-year period of abstinence related to the gestation and neonatal period of her son, she was euthymic. A provisional diagnosis of an alcohol-induced mood disorder was made. The patient was offered

education, reassurance, and CBT to help her deal with the depressive symptoms, but no antidepressant medications were prescribed. The depressive symptoms remained at their original intensity for several additional days and then began to improve. By 3 weeks of abstinence, the patient no longer met criteria for a major depressive episode, although she demonstrated mild mood swings for several additional weeks. This case is a fairly typical example of an alcohol-induced depressive disorder in an individual with an AUD.

Alcohol-Induced Anxiety Disorders. Anxiety symptoms fulfilling the diagnostic criteria for alcohol-induced anxiety disorders are also common in the context of acute and protracted alcohol withdrawal. In ICD-10, these are listed as organic anxiety disorders resembling generalized anxiety or panic disorders. Almost 80 percent of people with AUDs report prominent anxiety during an acute withdrawal episode, and their complaints can be intense enough for the clinician to consider diagnosing a panic disorder, generalized anxiety disorders, social phobia, and other anxiety conditions. However, when psychological or physiological symptoms of anxiety are observed in people with AUDs only in the context of heavy drinking or within the first month or so of abstinence, the symptoms are likely to diminish and subsequently disappear with time alone. Only a few pre-existing anxiety or stress disorders (e.g., panic disorder, social phobia and posttraumatic stress disorder [PTSD]) have been clearly shown to modestly increase the risk for later AUDs.

A 48-year-old woman was referred for evaluation and treatment of her recent onset of panic attacks. These episodes occurred two to three times per week over the past 6 months, with each lasting typically between 10 and 20 minutes. Panic symptoms occurred independently of levels of stress and could not be explained by the side effects of current medications or by medical conditions. The work-up included an evaluation of her laboratory test values, which revealed a carbohydrate-deficient transferrin (CDT) level of 3.5 percent, a uric acid level of 7.1 mg, and a γ-glutamyltransferase value of 47. All other blood tests were within normal limits.

The atypical late age of onset of panic attacks along with the blood results encouraged the clinician to probe further regarding the pattern of alcohol-related life problems with both the patient and, separately, her spouse. This documented a history of an AUD with an onset at approximately 35 years of age, with no evidence of a panic disorder before that time. The patient denied having repetitive panic attacks beyond a month of abstinence during her frequent periods of nondrinking that often lasted for 3 or 4 months. A working diagnosis of an AUD with an alcohol-induced anxiety disorder characterized by panic attacks was made, and the patient was encouraged to abstain. She was offered CBT to help decrease feelings of panic and anxiety and was appropriately treated for moderate withdrawal symptoms. Over the subsequent 3 weeks after a taper of benzodiazepines used for the withdrawal, the panic symptoms diminished in intensity and subsequently disappeared.

Alcohol-Induced Psychotic Disorders. About 3 percent of people with AUDs experience auditory hallucinations and/or paranoid delusions in a clear sensorium that develop in the context of heavy drinking or withdrawal. In ICD-10, these conditions are listed as organic delusional disorders in the organic section, and as a psychotic disorder in the substance use disorders section. Many of the symptoms resemble those seen in schizophrenia, but when the psychotic features develop only in the

context of alcohol problems they are likely to clear spontaneously within a few days to a month of abstinence. The syndromes may recur if heavy alcohol intake resumes.

Alcohol-Induced Major Neurocognitive Disorders, Amnestic Confabulatory Type, Persistent (Wernicke–Korsakoff Disorder)

One of the more dramatic alcohol-related CNS syndromes is the rare DSM-5 diagnosis of what in the past was referred to as an alcohol-induced persisting amnestic disorder, which is the result of a relatively severe CNS deficiency of thiamine. Similar criteria are offered in ICD-10 as an amnesic syndrome. Some people are at higher risk than others for this condition because of a genetically influenced transketolase deficiency. The syndrome has been historically subdivided into: (1) Wernicke encephalopathy, with prominent ataxia and palsy of the sixth cranial nerve, a condition that tends to reverse fairly rapidly with vitamin supplementation; and (2) - Korsakoff syndrome, which is permanent in at least a partial form in about 50 percent of the people affected. Korsakoff syndrome is characterized by a pronounced anterograde and retrograde amnesia along with potential impairment in visuospatial, abstract, and other types of learning. In most cases, the level of loss of recent memory is out of proportion to the global level of cognitive impairment. The patients who recover respond to 500 mg of IV thiamine, two to three times a day for 3 to 5 days, or 100 mg of oral thiamine a day that is usually administered for many months.

Alcohol-Induced Major Neurocognitive Disorders (Persisting Dementias)

A poorly studied, heterogeneous long-term cognitive problem that can develop in the course of AUDs was called *alcohol-induced persisting dementia* in DSM-IV and in ICD-10. Global decreases in intellectual functioning, cognitive abilities, and memory are observed, but recent memory difficulties are no more prominent than other elements of this condition. Brain functioning tends to improve with abstinence, but perhaps half of patients have long-term and even permanent disabilities. More than 50 percent of these patients have increased brain ventricle sizes and shrinkage of the cerebral sulci, although these changes are usually reversible with abstinence.

IDENTIFICATION IN CLINICAL SETTINGS AND RELATED LABORATORY TESTS

To establish an AUD diagnosis, a history of alcohol-related life problems must be obtained from the patient and, when available, an additional reliable informant. Because any patient, regardless of their education and income, might have an AUD, everyone should be screened for histories of alcohol-related accidents, interpersonal difficulties, problems at work, and encounters with the law. When a problem is apparent, the clinician should

ask more about the situation and about the complaints voiced by others. If an AUD is probable, the diagnostic criteria and drinking history should be reviewed with the patient.

Several simple questionnaires can help screen for problems. The Alcohol Use Disorders Identification Test (AUDIT, Table 11.2–7) uses ten items to review the pattern of life problems related to alcohol. More simple instruments, such as the CAGE ([need to] *C*ut down [on drinking], *A*nnoyance when asked about drinking, *G*uilt [about drinking], [need for] *E*ye-openers the morning after heavy drinking), are too short to be optimally sensitive or specific. Although the AUDIT is useful, such questionnaires do not diagnose an AUD, but only highlight individuals who are especially appropriate for a more detailed clinical interview.

Identifying heavier drinkers vulnerable to AUDs is facilitated by the several blood tests noted in Table 11.2–8. These *state markers of heavy drinking* reflect physiological alterations likely to be occur after ingesting about five or more drinks a day over several weeks. One marker with a sensitivity and specificity of about 60 percent is a level of at least 35 U/L of γ-glutamyltransferase (GGT), an amino acid–transporting enzyme that is temporarily induced by repetitive heavy drinking. Enzyme levels are likely to return to normal after 2 to 4 weeks of abstinence, after which further increases over 20 percent are useful in identifying a return to drinking. Similar results are seen for the deglycosylated form of the protein, transferrin, known as CDT. A value over 3 percent has sensitivities and specificities of 60 percent for the consumption of five or more drinks per day for 1 to 2 weeks. With a biological half-life of about 16 days, this test is also useful in monitoring abstinence. The combination of GGT and CDT is better than either test alone for identification of heavy drinking and monitoring abstinence.

Table 11.2–7.
Key Elements of the Alcohol Use Disorders Identification Test (AUDIT)

Item	5-Point Scale (least to most)
1. How often do you have a drink containing alcohol?	Never (score 0) to 4+ per week (score 4)
2. How many drinks containing alcohol do you have on a typical day?	1 or 2 (0) to 10+ (4)
3. How often do you have six or more drinks on one occasion?	Never (0) to Daily or almost daily (4)
4. How often during the last year have you found that you were not able to stop drinking once you had started?	Never (0) to Daily or almost daily (4)
5. How often during the last year have you failed to do what was normally expected from you because of drinking?	Never (0) to Daily or almost daily (4)
6. How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?	Never (0) to Daily or almost daily (4)
7. How often during the last year have you had a feeling of guilt or	Never (0) to Daily or almost

remorse after drinking?	daily (4)
8. How often during the last year have you been unable to remember what happened the night before because you had been drinking?	Never (0) to Daily or almost daily (4)
9. Have you or someone else been injured as a result of your drinking?	No (0) to Yes, during the last year (4)
10. Has a relative, friend, doctor, or other health worker been concerned about your drinking or suggested that you should cut down?	No (0) to Yes, during the last year (4)

The AUDIT is scored by simply summing the values associated with the endorsed response.

Adapted from Reinert DF, Allen GP. The Alcohol Use Disorders Identification Test (AUDIT): a review of recent research. *Alcohol Clin Exp Res*. 2002;26:272.

An MCV over 91 μm3 has a sensitivity of about 50 percent in identifying heavy drinking, but the 120-day lifespan of red cells does not allow MCV to be useful as an indicator of a return to drinking. Other indicators that can be helpful in identifying patients who are regularly consuming heavy doses of alcohol include high normal values of uric acid (e.g., >6.4 mg/dL), and elevations in the usual liver function tests. Useful physical findings include modest elevations in blood pressure, frequent bruising, cancer of the head, neck, or upper digestive tract, an enlarged liver, evidence of cirrhosis, and pancreatitis.

Table 11.2–8.
State Markers of Heavy Drinking Useful in Screening for AUDs

Test	Relevant Range of Results
γ-glutamyltransferase (GGT)	>35.0 U/L
Carbohydrate-deficient transferrin (CDT)	>3.0%
Mean corpuscular volume (MCV)	>91.0 μm^3
Uric acid	>6.4 mg/dL for men >5.0 mg/dL for women
Serum glutamic oxaloacetic transaminase (aspartate aminotransferase)	>45.0 IU/L
Serum glutamic pyruvic transaminase (alanine aminotransferase)	>45.0 IU/L

DIFFERENTIAL DIAGNOSIS

Once a pattern of alcohol-related life problems has been identified, the diagnosis of an AUD may be fairly obvious. However, the differential diagnosis is more challenging when a patient presents with psychiatric symptoms that may indicate an independent psychiatric disorder that requires concomitant treatment. As described in more detail above in Diagnosis and Clinical Features, to recognize an *alcohol-induced psychiatric condition* the clinician needs to focus on the full diagnostic criteria (e.g., for a major depressive disorder), not just subsyndromal problems likely to be part of alcohol intoxication and withdrawal. Using the timeline approach described in that section, if the patient only meets criteria for the psychiatric condition during intoxication or withdrawal (i.e.,

appears to have a substance-induced condition), the next step is to confirm the diagnosis by monitoring the symptoms over a month of abstinence and observing that the condition no longer meets full criteria for the full diagnostic syndrome.

It is equally important to identify *independent psychiatric disorders* that predate the AUD and continue after four abstinent weeks. As substance-induced conditions are described in more detail above in the section on Diagnostic and Clinical Features, this section offers more detail regarding independent comorbid conditions seen with AUDs.

Antisocial Personality Disorder

Antisocial personality disorder in DSM-5 is based on severe antisocial behaviors beginning before age 15 as conduct disorder and continuing into adulthood. People with antisocial personality disorder (ASPD) are likely to be impulsive, sometimes violent, and to be risk takers who do not easily learn from mistakes. About 80 percent of people with this disorder develop severe alcohol problems and/or SUDs, they are likely to discontinue treatment prematurely, and have relatively poor prognoses in AUD treatment, where they represent about 5 percent of women and 15 percent of men. Symptoms of other personality disorders are often observed during intoxication and withdrawal, but, other than ASPD and borderline personality disorders, few have been documented to predate the AUD. Thus, it is advisable to defer a diagnosis of additional personality disorders until at least a month after abstinence unless the syndrome clearly antedated the AUD. The treatment of AUDs in patients with independent ASPD involves the same general approaches as for alcohol-related disorders overall, with common sense variations to deal with the interpersonal problems.

Independent Schizophrenia

A second independent disorder associated with elevated risks for both AUDs and SUDs is schizophrenia. Typically characterized by a slow onset of paranoid delusions and auditory hallucinations in a clear sensorium and usually beginning in the late teens to the 20s, schizophrenia is likely to be severe, long lasting, and debilitating. Possibly reflecting a lack of long-term psychiatric treatment facilities, people with schizophrenia are sometimes homeless, a situation that, along with impaired judgment associated with their psychoses, may contribute to alcohol- and drug-related problems. There may also be overlap in genetic mechanisms (e.g., related to dopamine) that contribute to AUDs and psychotic disorders. Heavy drinking is likely to undercut the effectiveness of antipsychotic medications, increase mood swings, intensify psychoses, and contribute to a higher probability of repeatedly revolving into and out of inpatient care.

AUDs are observed in about 40 percent of individuals with schizophrenia.

A 34-year-old unemployed, divorced Anglo man entered treatment for an AUD. His history revealed seven hospitalizations since age 23, with discharge diagnoses including schizophrenia and AUDs. By age 23, the patient had experienced auditory hallucinations and paranoid delusions for the prior year, but had not yet demonstrated severe enough problems related to alcohol or other drugs to fulfill criteria for an AUD or SUD. A hospitalization at age 28 was followed by 6 months of abstinence while living in a recovery home, during which time the patient continued to have auditory hallucinations and paranoid delusions without insight. The current hospitalization involved the treatment of alcohol withdrawal while re-evaluating and stabilizing the antipsychotic medications for this man with an independent schizophrenia.

Independent Depressive and Bipolar I Disorders

A third group of conditions that are part of the differential diagnoses of AUDs involve mood-related diagnoses. Major depressive disorders are discussed more extensively above under alcohol-induced psychiatric disorders. However, when a major depressive episode occurs before the onset of an AUD or during a period of abstinence, it needs to be treated as an independent condition while the clinician also addresses the alcohol-related problems. While some studies indicate that independent major depressions do not increase the risk for later AUDs, pre-existing bipolar I disorder is more closely linked with the AUD risk. This may reflect heavy drinking that began during mania-related hyperexcitability, impulsiveness, and associated poor judgment. As with schizophrenia, there also may be overlap in genetic mechanisms contributing to both bipolar and substance-related disorders. Bipolar II, on the other hand, is difficult to evaluate in patients with AUDs, because their intoxication, withdrawal, and reactions to life problems can mimic hypomania. Bipolar II labels should be reserved for AUD patients with clear hypomanic episodes antedating the AUD or directly observed by the clinician after at least a month of abstinence.

Independent Major Anxiety, Trauma, and Stress-Related Disorders

There is at least a modest association between the risk for AUDs and independent panic disorder, social phobia, and PTSD. Regarding the latter, a delay of many months between the stressful event and the development of PTSD might indicate an impact of alcohol withdrawal or intoxication related depression or anxiety on subclinical trauma-related symptoms. In these instances, the PTSD might markedly improve with continued abstinence, perhaps falling below a diagnostic threshold for PTSD. In other individuals, however, PTSD was clearly present before the AUD or remained severe during abstinence, in which case once the acute alcohol-related problems are addressed, the patient should consider longer-term CBT as well as medication-based therapies appropriate for PTSD.

Other Disorders

Attention-deficit/hyperactivity disorder (ADHD) with a childhood onset involves learning problems, difficulties focusing attention, and problems sitting still and following through on tasks. Because these children can be disruptive at home and school, and because many are treated with stimulant medications such as methylphenidate (Ritalin), it is possible that they might carry a heightened future risk for AUDs and SUDs. However, studies to date do not uniformly support a close tie between ADHD and later substance-related problems unless the child also met criteria for conduct disorder.

There are also complex relationships between AUDs and *SUDs*. Men and women with ASPD have a high risk for multiple SUDs, including AUDs. It is also probable that individuals with dependence on opioids and stimulants (cocaine and amphetamines) have an increased risk for AUDs by turning to alcohol to moderate the intoxication or side effects from their preferred drug. However, while many individuals with AUDs have used other substances, most do not meet criteria for an SUD. An exception is nicotine use disorders, which are elevated among individuals with AUDs. This finding may relate to attempts to use nicotine to moderate some of the effects of high doses of alcohol or withdrawal, or a possible genetic relationship between nicotine- and alcohol-related syndromes.

COURSE AND PROGNOSIS

Most people with AUDs, regardless of their backgrounds, exhibit general similarities in the time course and prevalence of alcohol-related life difficulties. The differences that do exist mostly reflect characteristics of individuals in society in general, and may not indicate unique aspects of their AUD. For example, women in the general population and among those with AUDs have higher rates of independent depressive episodes and are less likely to be violent, have a full time job, or have police records compared to comparable men. However, some gender differences do exist (e.g., women are likely to begin drinking a few years later than men, and may demonstrate a faster escalation of AUD symptoms). Similarly, older individuals in the general population and with AUDs are more likely to have medical problems, to take multiple medications, and have fewer extensive social support systems compared to younger individuals.

Early Course

Patients with ASPD who develop an AUD are likely to have an early onset of drinking, intoxication, and alcohol-related problems, but that scenario is not applicable to the other 80 percent of men and 95 percent of women with AUDs. As shown in Table 11.2–9, the usual person with an AUD has their first drink (other than taking a sip from a parent's glass) between the ages 13 and 15 years, the first intoxication usually occurs at 15 or 16, and

the first minor alcohol-related problem is usually observed in the late teens, milestones similar to what is expected for people in the general population. More severe alcohol-related difficulties begin to cluster together in the early 20s to the middle 30s.

Later Course

Once a moderate to severe AUD has developed, unless the person abstains, there are likely to be repeated future periods of drinking problems that alternate with periods of abstinence and *temporary* controlled drinking. The usual person with an AUD uses the temporary cessation of drinking problems to convince themselves that alcohol is not really a cause for concern. Those periods of abstinence, lasting days to months, are usually followed by times when drinking rules are established and at least temporarily followed. For example, the person might avoid spirits (ignoring that a glass of beer or wine and a shot of whiskey have similar amounts of alcohol) or limit drinking to certain times of the day and under specified conditions. This period of temporary control usually leads to an escalation of alcohol intake, followed by new problems that precipitate new abstinent periods, and the cycle begins again.

Table 11.2–9.
Clinical Course of AUDs[a]

Age at first drink[b,c]	13–15 years
Age at first intoxication[b,c]	15–17 years
Age at first problem[b]	16–22 years
Age at onset of an AUD	23–40 years
Age at death	60 years
Fluctuating course of abstention, temporary control, alcohol problems	NA
Spontaneous remission in 20%	NA

NA indicates that age of onset does not apply to this item.

[a]The data presented here are based on DSM-IV abuse and dependence, and it is likely that mild AUD (2 or 3 of 11 criteria endorsed) is similar to DSM-IV abuse, while moderate or severe AUD (4+ criteria endorsed) is more similar to DSM-IV dependence.

[b]Same as general population.

[c]While these ages reflect the usual course of alcohol use in individuals without conduct disorder or antisocial personality disorder, an earlier drinking or intoxication is associated with an increased AUD risk.

An important development in at least 20 percent of people with AUDs is a spontaneous remission. Here, without recent treatment or self-help group participation the person spontaneously stops drinking and remains abstinent for extended periods. Such outcomes are most likely for those with fewer alcohol-related problems, more social supports, jobs, greater life

stability, and no concomitant ASPD or SUD diagnoses.

However, if drinking continues, the person potentially decreases their lifespan by at least 10 years as a result of the alcohol-related pathologies described earlier in this chapter. This early death also reflects a 10 percent risk for suicide, usually during a depressive episode.

Prognosis

Fewer than half of people with AUDs ever receive treatment for their disorder. However, once they complete a program, a majority do well and develop long-term, often permanent, abstinence. Good outcomes are most likely in patients without ASPD, and in the context of having a job and close family contacts. These attributes predict about a 60 percent chance for >1 year of abstinence, with 1-year rates associated with a good chance for continued abstinence over an extended period. Individuals with independent major psychiatric conditions, such as ASPD, schizophrenia, and bipolar I disorder, are likely to run the course of their independent psychiatric illness as well.

TREATMENT

The core of treatment for AUDs involves maximizing motivation for abstinence, helping the person restructure a life without alcohol, and minimizing relapse to substance-using behaviors. This CBT-based approach is similar to efforts appropriate for any long-term disorder that requires changes in lifestyles, such as diabetes or hypertension. Much of the challenge for clinicians rests with recognizing the high prevalence of AUDs, how often these conditions involve temporary symptoms of other psychiatric syndromes, and how to use clinical clues, physical findings, and laboratory tests to identify an AUD.

Three general steps are involved in treating the person with an AUD: intervention; detoxification; and rehabilitation, including relapse prevention. These approaches can begin after life-threatening medical conditions are addressed and psychiatric emergencies treated. Thus, suicidal plans, psychotic features without insight, and alcohol withdrawal delirium require inpatient hospitalizations until the patient is stabilized enough to begin alcohol rehabilitation.

Intervention

The goal is to use the principles of motivational interviewing and brief interventions to help patients recognize the adverse consequences likely to occur if they do not stop drinking. A discussion of the presenting complaint (e.g., insomnia or depression) is a useful way to both show empathy and enhance motivation for change. The emphasis is then placed on discussing how alcohol either created or contributed to these problems and reassuring

the patient that abstinence can be achieved. If the person does not respond to the first brief intervention or motivational interview, the same nonjudgmental but persistent approach can be used each time an alcohol-related problem is identified. Many individuals with AUDs take time to realize how alcohol contributed to problems and to be ready and willing to change.

For the physician, the intervention often begins with a "brief intervention," a cost-effective approach that can be used in the clinician's office. It is summarized by the acronym FRAMES: giving Feedback on alcohol risks; emphasizing a person's Responsibility for change; offering Advice; giving a Menu of options to consider; employing Empathy; and emphasizing Self-efficacy, or the ability to change. The goal is to establish an alliance with patients by demonstrating an understanding of their viewpoint while encouraging them to think through consequences associated with alcohol and the way that changing behaviors might yield benefits. During these discussions it is important to recognize the patient's ambivalence toward abstinence and to monitor the person's readiness to change. Resistance is best handled through discussion and problem solving rather than direct confrontation.

Two 15-minute sessions, along with a workbook available from the National Institute on Alcohol Abuse and Alcoholism (pubs.niaaa.nih.com), can be used to review drinking problems and help the patient understand how life cues contribute to the drinking pattern. Between two visits separated by about 1 month, patients keep notes of their drinking patterns. The second session reviews the same issues and is followed by several short telephone interviews.

Detoxification

Most people with an AUD have relatively mild, if any, withdrawal symptoms when they stop drinking. If the patient is in relatively good health and has a good social support system, the withdrawal syndrome associated with alcohol usually resembles a mild flu or bad hangover. The essential first step in detoxification is a thorough physical examination. In the absence of a serious medical disorder or a comorbid SUD, severe alcohol withdrawal is unlikely. The second step is to offer rest, nutrition, and oral multiple vitamins, including thiamine.

Mild or Moderate Withdrawal. Alcohol withdrawal develops because the brain adapted to the presence of this drug and cannot function adequately in the absence of alcohol. Treatment can include giving enough of a brain depressant on the first day to diminish symptoms, and then weaning the patient off the medication over the next 5 days. If the symptoms are mild and the person is in good health, this can be done as an outpatient with daily brief visits. Any depressant can work, but most

clinicians choose benzodiazepines for their relatively low cost and higher safety. Adequate treatment can be given with either short-acting drugs, such as lorazepam (Ativan), or long-acting medications such as chlordiazepoxide or diazepam (Valium).

An example is administering 25 mg of chlordiazepoxide by mouth three or four times a day on the first day, with a notation to skip a dose if the patient is sleepy. Up to two additional 25-mg doses during the first day can be used if tremor or autonomic dysfunction increases. The first day's dose can be decreased by about 20 percent each subsequent day, with a resulting need for no medication after 4 or 5 days. However, the prior day's dose should be restored if withdrawal symptoms increase and further decreases in doses delayed a day or two. When using a long-acting agent, such as chlordiazepoxide, the clinician must avoid producing excessive sedation; if the patient is sleepy, the next scheduled dose should be omitted. When using a short-acting drug, such as lorazepam, the patient must not miss any dose because rapid changes in benzodiazepine concentrations can precipitate a seizure. Other depressants, including anticonvulsants (e.g., gabapentin [Neurontin]) can also be effective but have added dangers and costs compared to benzodiazepines. Some clinicians recommend β-adrenergic receptor antagonists (e.g., propranolol [Inderal]) or α-adrenergic receptor agonists (e.g., clonidine [Catapres]), although these medications do not decrease the risk of seizures or delirium and can mask the intensity of some withdrawal signs that are used to monitor withdrawal severity.

Severe Withdrawal. For the less than 3 percent of AUD patients with extreme autonomic dysfunction, agitation, and confusion (alcoholic withdrawal delirium or delirium tremens [DTs]), there is no perfect treatment. The key is to determine why such severe withdrawal symptoms occurred, especially whether there is a concomitant medical problem that needs immediate treatment. The withdrawal symptoms in DTs can then be minimized with benzodiazepines (high doses are sometimes required), along with the adjunctive use of an antipsychotic (e.g., haloperidol [Haldol]) to control severe agitation and hallucinations. If the patient does not respond to benzodiazepines, a clinician experienced in treating withdrawal delirium in an ICU can consider inducing a state of arousable deep sleep with propofol (Diprivan) or the α-2-receptor agonist dexmedetomidine (Precedex). Here, higher doses are used on the first or second day to control behavior, and the patient is then weaned off medications by about the fifth day.

Another 1 percent of patients in alcohol withdrawal have a single grand mal convulsion, with a peak incidence on day two. This requires a neurological evaluation, but in the absence of an independent seizure disorder, a single seizure does not indicate using anticonvulsant drugs.

Protracted Withdrawal. Finally, regarding withdrawal, mild symptoms of anxiety, insomnia, and autonomic overactivity are likely to continue for 2 to 6 months following acute withdrawal. Although no pharmacological treatment for this syndrome has been identified, a medication discussed below for the rehabilitation phase, acamprosate (Campral), may diminish some of these symptoms. It is important to warn patients that some sleep problems or feelings of nervousness might remain after acute withdrawal, and consider efforts to help them feel more comfortable, because these protracted withdrawal symptoms enhance the probability of relapse.

Rehabilitation

Rehabilitation includes three major components: (1) increasing and maintaining high levels of motivation for change; (2) helping patients readjust to a lifestyle that maximizes functioning and decreases the risk for drinking; and (3) relapse prevention. These steps often begin in the context of acute and protracted withdrawal syndromes and life crises.

No single major life event or psychiatric disorder is a unique cause of AUDs. The effects of earlier causes are likely to have been joined by reinforcing effects of the drug, especially regarding relief of hangovers and mild withdrawal symptoms, and the years of a stressful lifestyle in the context of heavy drinking. This is true even though many people with AUDs believe that the unique cause was depression, anxiety, difficult life events, prior stress, or pain syndromes. However, the alcohol was as likely to have contributed to the mood disorder, accident, or life stress as the other way around.

The same general treatment approach is used in inpatient and outpatient settings. The selection of the more expensive, but more intense and effective, inpatient mode often depends on evidence of additional severe medical or psychiatric syndromes, the absence of nearby outpatient groups, and/or the patient's non-response to outpatient care. The treatment process in either setting involves 2 to 4 weeks of care as an intensive period of help, followed by at least 3 to 6 months of less frequent outpatient care. The latter uses a combination of individual and group counseling, avoiding psychotropic medications unless needed for independent disorders, and involvement in self-help groups such as Alcoholics Anonymous (AA).

Counseling. Counseling in the first several weeks to months of recovery focuses on day-to-day life issues to help patients maintain a high level of motivation for abstinence and to enhance their functioning. Psychotherapy techniques that provoke anxiety or that require deep insights may not be beneficial during the early phases of treatment, and, at least theoretically, the added anxiety may impair the ability to abstain.

One goal of these CBT-based counseling sessions is to help the patient

learn ways to cope with life stresses. The clinician can use role rehearsal, modeling, and role playing while encouraging patients to practice these skills between sessions. At the same time, individuals are encouraged to identify areas of problems in day-to-day functioning, paying special attention to how they react to these challenges and the impact that substance use might have on the outcomes.

To optimize motivation, treatment sessions explore the consequences of drinking, the likely future course of alcohol-related life problems, and the improvement that can be expected with abstinence. Counseling also deals with how to build a lifestyle free of alcohol. Discussions cover the need for a sober peer group, a plan for social and recreational events without drinking, and approaches for re-establishing communication with family members and friends. Whether in an inpatient or an outpatient setting, individual or group counseling is usually offered about three times a week for the first 2 to 4 weeks, followed by less intense efforts, perhaps once a week, for the subsequent 3 to 6 months.

Relapse Prevention. The third component of rehabilitation, relapse prevention, begins with identifying situations with high risks for relapse. The counselor helps patients develop ways to identify and cope with craving and with challenging events or emotional states that make drinking more likely. It is also important to remind patients that short-term relapses are not an excuse for returning to regular drinking. Recovery is a process of trial and error, and slips can be clues to help identify high-risk situations and to develop more appropriate coping techniques.

Importance of the Family. AUDs affect many people in the patient's life, and an important aspect of recovery involves helping family and friends understand more about AUDs and how rehabilitation is an ongoing process. Counseling and support for spouses, offspring, and friends can help them work to rebuild key relationships and learn how to avoid protecting the patient from the consequences of drinking, a step likely to undercut the patients motivation to abstinence.

Treatment of Independent Psychiatric Disorders . Regarding some of the independent psychiatric disorders discussed under Differential Diagnoses above, clinicians must treat the patient with bipolar I disorder who has an associated AUD with appropriate psychotherapy and lithium (Eskalith) or valproic acid (Depakene); use relevant psychological and behavioral techniques for the patient with antisocial personality disorder; and offer long-term antipsychotic medications to the patient with schizophrenia. The goal is to minimize the symptoms of the independent psychiatric disorder in the hope that a greater level of life stability will be associated with a better prognosis for the patient's alcohol problems. Initial

treatment should work to control mania or severe depression along with treatment of withdrawal, and alcohol rehabilitation should be considered once the independent disorder is under control.

Possible Medications for the AUD. Patients with AUDs who do not have an independent psychiatric disorder have no clear need for antidepressants or antianxiety drugs after withdrawal in the treatment of their AUDs. Moderate anxiety and insomnia that linger as part of a protracted abstinence syndrome should be treated with CBT, education, and reassurance. Medications, including benzodiazepines, for anxiety and insomnia are likely to lose effectiveness as tolerance develops, and patients sometimes increase the dose and have subsequent problems related to the prescribed drug.

Similarly, although low levels of sadness and mood swings can linger for months, controlled clinical trials indicate little benefit to prescribing antidepressant medications or lithium for the average person with an AUD who has no independent psychiatric disorder. The substance-induced mood disorder is likely to clear before the medications can take effect, and patients who resume drinking while on the medications face potential dangers of alcohol–drug interactions. With little evidence that such medications are routinely effective, the dangers outweigh potential benefits from their routine use.

As is true for most clinical disorders, the use of medications for the treatment of AUDs requires remembering that all treatments have both benefits and risks. Double-blind, controlled studies are important to establish an acceptable ratio of assets to dangers, while controlling for the impact of spontaneous remission and fluctuations in patients' conditions that occur over time. The random variation in the course of AUDs, along with the probability that patients are most likely to seek treatment when their problems are most intense, means that the only way to be confident that a medication contributed to a good outcome is through controls for spontaneous remissions and fluctuating intensities of drinking problems by also observing outcomes in an appropriate comparison group.

Data from double-blind trials support modest effects for two medications offered in the context of CBT. The first drug is acamprosate, an analog of the amino acid neurotransmitter taurine with a structure that resembles GABA. This medication antagonizes neuronal overactivity related to the excitatory neurotransmitter glutamate, at least in part by acting as an antagonist to NMDA receptors. Thus, acamprosate may diminish mild anxiety, mood swings, and other sleep difficulties associated with the protracted withdrawal syndrome observed after the first 4 to 5 days of abstinence. Using about 2,000 mg divided into three doses per day, the majority of studies report a 15 to 20 percent better outcome with active drug compared to placebo, with relatively mild side effects of GI problems

such as diarrhea. Although there are no clear guidelines, clinicians often prescribe this drug for 6 months, during which time it is hoped an alcohol-free lifestyle can be developed.

The second promising medication is the long-acting opioid antagonist naltrexone (ReVia). This agent has been marketed for many years as a treatment for maintaining opioid abstinence, and is effective for rehabilitation with highly motivated opioid-dependent individuals because this opioid blocker drug attenuates an opioid high. Naltrexone works by blocking opioid receptors in the brain with associated decreased activity in the ventral tegmental reward area of the brain, and is hypothesized to decrease the rewarding effects of a drink and to diminish craving. The majority of double-blind, placebo-controlled trials support the superiority of naltrexone with about a 15 to 20 percent better outcome than for placebo. The side effects for 50 mgs per day include mild GI upset and lethargy. Several studies indicate that individuals with a gene variation related to the Asn40Asp location in the gene for the mu opioid receptor (OPRM1) have better AUD outcomes. There is also a 380-mg, once-per-month naltrexone injection (Vivitrol), which might be particularly useful for non-compliant patients.

The combination of acamprosate (666 mg three times a day) plus naltrexone (50 mg per day or three times a week) may be more effective than either drug alone, while not increasing the side effect profile. However, not all studies agree and further research is needed.

A third drug of possible interest is the alcohol-sensitizing agent disulfiram (Antabuse), which is usually prescribed at 250 mg per day. The goal is to warn patients that drinking alcohol while taking disulfiram precipitates nausea, vomiting, and changes in blood pressure. However, few data convincingly prove that disulfiram is more effective than placebo, probably because patients on placebo are given the same caveat. Many clinicians have stopped prescribing this agent, partly because of the dangers associated with the drug itself, including mood swings, rare instances of psychosis, a possible increase in peripheral neuropathies, and a rare but potentially fatal hepatitis. Moreover, patients with pre-existing heart disease, stroke, diabetes, and other conditions cannot be given disulfiram because an alcohol reaction could be fatal.

Several additional medications are worth brief mention. First, while not yet approved for the treatment of AUDs, the anticonvulsant topiramate (Topamax) has the usual GABA-boosting activity associated with anti-seizure treatments along with a glutamate receptor blockade and effects on dopamine. Perhaps related to these combined effects, several studies using topiramate (300 mg per day) in AUD treatment reported improvements in drinking patterns. Second, some data indicate that 4 micrograms per kilogram body weight of the serotonin 3 receptor antagonist ondansetron (Zofran) might be beneficial in AUD treatment, particularly in early onset

AUDs with comorbid drug dependence and criminality. Additional medications that have been evaluated for the treatment of AUDs include several selective serotonin reuptake inhibitors (SSRIs, e.g., sertraline [Zoloft] up to 200 mg per day), with some evidence that patients with the SS genotype in the promoter region of the serotonin transporter gene may be more likely to respond; the GABA-B receptor agonist baclofen (Lioresal); and some antipsychotic medications, although these medications have not consistently proven to be better than placebo in controlled trials.

Self-Help Groups. Clinicians should consider the potential importance of "12-step-based" self-help groups such as AA. Similar to what occurs in CBT and relapse prevention, AA participants are encouraged to change the way they view their drinking while building a lifestyle that optimizes sober friends and uses the support of family members. Help from AA is available 24 hours a day, and participants learn they can still enjoy social functions without drinking, while observing models of recovery through the accomplishments of sober members of the group.

Clinicians play important roles by encouraging AA attendance while helping patients select the most appropriate group. Some meetings are comprised only of men or women, and others are mixed; some are mostly attended by blue-collar men and women, whereas others are mostly for professionals; some groups place great emphasis on religion, and others are more eclectic. Patients with independent psychiatric disorders may need additional education about AA. The clinician should remind them that some members of AA may not understand their special need for medications, and should suggest ways to cope with inappropriate suggestions that important medications be stopped. Although difficult to evaluate using double-blind controls, most studies indicate that participation in AA is associated with improved outcomes.

Following the family intervention and at the insistence of his wife and partners, JP agreed to enter treatment. He and his family opted for a more intensive inpatient-based alcohol and drug treatment program of 4 weeks' duration. The physical examination on entrance to treatment revealed no major medical problems. Vital signs after 3 days of sobriety were within normal limits except for a slight increase in heart rate (80 beats per minute) and blood pressure (145/90). To treat a mild withdrawal syndrome, 25 mg chlordiazepoxide up to three times per day as needed was prescribed for systolic blood pressure over 140/90, tremor, or other signs of autonomic dysfunction. JP received two doses on his first day of hospitalization, one dose on his second day, and no further benzodiazepines were required.

The core of therapy involved group counseling twice per day, emphasizing the patient's responsibility for the decision to abstain, discussing problems associated with continued substance use, and describing steps that can be taken to help rebuild one's relationship with family, colleagues, and friends. JP cooperated well with treatment, and the discharge planning included a discussion of the possible medications during the rehabilitation phase. After consulting with the patient and the family, both acamprosate (666 mg by mouth three times a day) and naltrexone (50 mg by mouth daily) were prescribed, with the goal of continuing the medications for approximately 6 months.

In addition, couples counseling was carried out with JP and his wife to optimize her understanding of the recovery process and offer her the opportunity of dealing with questions and feelings of frustration, anger, and anxiety that developed during JP's drinking years. The subsequent aftercare outpatient group met twice a week. Steps were taken to help JP identify an AA group with which he might be most comfortable in terms of shared background with members as well as time and location. An Al-Anon group was identified for his wife as well.

JP's course involved two brief relapses in the first 6 months of treatment, each involving a return to alcohol lasting less than 2 days. After each, the relapse prevention training and the education received by his spouse, as well as his return to both AA and aftercare, helped him return to sobriety. After the first 6 months, no further relapses were observed over the next year. Currently, JP continues with one or two AA meetings per week and has become involved in outreach from AA, visiting a local hospital and helping to lead an AA group for individuals in early recovery.

References

Alati R, Betts KS, Williams GM, Najman JM, Hall WD. Generational increase in young women's drinking: A prospective analysis of mother-daughter dyads. *JAMA Psychiatry*. 2014;71:952.

Areesantichai C, Chapman R, Perngparn U. Effectiveness of a tailored goal-oriented community brief intervention (TGCBI) in reducing alcohol consumption among risky drinkers in Thailand: A quasi-experimental study. *J Stud Alcohol Drugs*. 2013;74:311.

Barrowclough C, Eisner E, Bucci S, Emsley R, Wykes T. The impact of alcohol on clinical outcomes in established psychosis: a longitudinal study. *Addiction*. 2014;109:1297.

Batki SL, Pennington DL, Lasher B, et al. Topiramate treatment of alcohol use disorder in veterans with posttraumatic stress disorder: a randomized controlled pilot trial. *Alcohol Clin Exp Res*. 2014;38:2169.

*Compton WM, Dawson DA, Goldstein RB, Grant BF. Crosswalk between DSM-IV dependence and DSM-5 substance use disorders for opioids, cannabis, cocaine and alcohol. *Drug Alcohol Depend*. 2013;132:387.

Conrod PJ, O'Leary-Barrett M, Newton N, et al. Effectiveness of a selective, personality-targeted prevention program for adolescent alcohol use and misuse: a cluster randomized controlled trial. *JAMA Psychiatry*. 2013;70:334.

Del Re A, Maisel N, Blodgett J, Finney J. The declining efficacy of naltrexone pharmacotherapy for alcohol use disorders over time: a multivariate meta-analysis. *Alcohol Clin Exp Res*. 2013;37:1064.

Edwards AC, Kendler KS. Twin study of the relationship between adolescent attention-deficit/hyperactivity disorder and adult alcohol dependence. *J Stud Alcohol Drugs*. 2012;73:185.

*Garbutt JC, Greenblatt AM, West SL, et al. Clinical and biological moderators of response to naltrexone in alcohol dependence: a systematic review of the evidence. *Addiction*. 2014;109:1274.

*Hasin DS, O'Brien CP, Auriacombe M, et al. DSM-5 criteria for substance use disorders: recommendations and rationale. *Am J Psychiatry*. 2013;170:834.

Hasler BP, Martin CS, Wood DS, Rosario B, Clark DB. A longitudinal study of insomnia and other sleep complaints in adolescents with and without alcohol use disorders. *Alcohol Clin Exp Res*. 2014;38:2225.

Heron J, Maughan B, Dick DM, Kendler KS, Lewis G. Conduct problem trajectories and alcohol

use and misuse in mid to late adolescence. *Drug Alcohol Depend*. 2013;133:100.

Ilgen MA, Jain A, Lucas E, Moos RH. Substance use-disorder treatment and a decline in attempted suicide during and after treatment. *J Stud Alcohol Drugs*. 2007;68:503.

Johnson B, Seneviratne C, Wang X, Ait-Daoud N, Li M. Determination of genotype combinations that can predict the outcome of the treatment of alcohol dependence using the 5-HT3 antagonist ondansetron. *Am J Psychiatry*. 2013;170:1020.

Johnston LD, O'Malley PM, Bachman JG, Schulenberg JE. *Monitoring the Future National Survey Results on Drug Use, 1975–2013: Volume III, College Students and Adults Ages 19–50*. Ann Arbor, MI: Institute for Social Research, The University of Michigan; 2014.

Jorgensen C, Pedersen B, Tonnesen H. The efficacy of disulfiram for the treatment of alcohol use disorder. *Alcohol Clin Exp Res*. 2011;35:1749.

Kolla B, Mansukhani M, Schneekloth T. Pharmacological treatment of insomnia in alcohol recovery: a systematic review. *Alcohol Alcohol*. 2011;46:578.

Kranzler H, Feinn R, Armeli S, Tennen H. Post-treatment outcomes in a double-blind, randomized trial of sertraline for alcohol dependence. *Alcohol Clin Exp Res*. 2012;36:509.

Maisel NC, Blodgett JC, Wilbourne PL, Humphreys K, Finney JW. Meta-analysis of naltrexone and acamprosate for treating alcohol use disorders: when are these medications most helpful? *Addiction*. 2013;108:275.

Mann K, Lemenager T, Hoffman S, et al. PREDICT Study Team: Results of a double-blind placebo controlled pharmacotherapy trial in alcoholism conducted in Germany and comparison with the US COMBINE Study. *Addict Biol*. 2013;18:937.

Nomura H, Matsuki N. Ethanol enhances reactivated fear memories. *Neuropsychopharmacology*. 2008;33:2912.

*Ornstein S, Miller P, Wessell A, Jenkins R, Nemeth L. Integration and sustainability of alcohol screening, brief intervention, and pharmacotherapy in primary care settings. *J Stud Alcohol Drugs*. 2013;74:598.

Paulus MP, Schuckit MA, Tapert SF, et al. High versus low level of response to alcohol: Evidence of differential reactivity to emotional stimuli. *Biol Psychiatry*. 2012;72:848.

Pridemore WA. The impact of hazardous drinking on suicide among working-age Russian males: an individual-level analysis. *Addiction*. 2013;108:1933.

*Schuckit MA. Ethanol and methanol. In: Brunton LL, Chabner BA, Knollkmann BC, eds. *Goodman & Gilman's the Pharmacological Basis of Therapeutics*. 12th ed. New York: McGraw-Hill; 2011.

Schuckit MA. A brief history of research on the genetics of alcohol and other drug use disorders. *J Stud Alcohol Drugs*. 2014;Suppl 17:59.

Schuckit MA. Recognition and management of withdrawal delirium (delirium tremens). *N Eng J Med*. 2014;371:2109–2113.

Schuckit MA, Kalmijn JA, Smith TL, Saunders G, Fromme K. Structuring a college alcohol prevention program on the low level of response to alcohol model: a pilot study. *Alcohol Clin Exp Res*. 2012;36:1244.

Schuckit MA, Smith TL, Kalmijn JA. The patterns of drug and alcohol use and associated problems over 30 years in 397 men. *Alcohol Clin Exp Res*. 2014;38:227.

Yip SW, Doherty J, Wakeley J, et al. Reduced subjective response to acute ethanol administration

among young men with a broad bipolar phenotype. *Neuropsychopharmacology*. 2012;37:1808.

▲ 11.3 Stimulant-Related Disorders

ROCCO A. IANNUCCI, M.D., AND ROGER D. WEISS, M.D.

INTRODUCTION

Stimulants in their natural forms, such as the plants *Erythroxylon coca*, *Ephedra sinica*, and *Catha edulis*, have been used by humans for millennia. In the United States, cycles of widespread stimulant misuse and associated problems have occurred for more than one hundred years. Recent decades have witnessed waxing and waning epidemics of cocaine and methamphetamine use, as one drug supplants the next. Stimulant-related disorders remain an important public health concern today, with continued high rates of legal, psychiatric, medical, and social problems related to psychostimulant misuse.

DEFINITIONS

The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) describes 10 stimulant-related diagnoses, which are further categorized into stimulant use disorder and the stimulant-induced disorders. Stimulant use disorder is defined in DSM-5 by a cluster of cognitive, behavioral, and physiological symptoms indicating continued use of stimulants despite significant problems related to or caused by their use. In the 10th revision of the *International Classification of Diseases* (ICD-10), the corresponding diagnoses of amphetamine and cocaine dependence syndromes are defined by a cluster of cognitive, behavioral, and physiological phenomena reflecting persistent use of these stimulants despite harmful consequences and to the exclusion of other activities and obligations. The hallmark of a severe stimulant use disorder is a pattern of compulsive substance use and loss of control over use that interferes with other aspects of life and health. It may involve the development of tolerance and/or withdrawal symptoms, but neither is necessary or sufficient to meet diagnostic criteria for stimulant use disorder. Stimulants themselves are classified as amphetamine-type substances (those with a substituted phenylethylamine structure), structurally dissimilar agents with similar effects (such as methylphenidate), and cocaine.

Stimulant tolerance is defined in DSM-5 as the need for an increased amount of amphetamine, cocaine, or other stimulants to achieve the same effect, or a diminished response to continued use of a given amount. Tolerance may occur as a symptom of an SUD, but it is not included as a disorder itself. On the other hand, stimulant withdrawal is identified as a stimulant-induced disorder in DSM-5. It is defined as a constellation of symptoms opposite to the effects of stimulant intoxication, including

dysphoric mood and characteristic physiologic changes.

Cocaine and amphetamine dependence were previously included as diagnoses in DSM-IV and were used to describe illnesses involving 3 of 7 possible symptoms, very similar to those listed for stimulant use disorder as described in DSM-5. These diagnostic entities are defined here due to widespread use in clinical studies to date.

In DSM-5, the diagnoses of amphetamine and cocaine abuse were eliminated, but most of the symptoms of these disorders are included in the DSM-5 stimulant use disorder. ICD-10 includes the diagnosis of harmful use, as a less severe psychoactive SUD. This term denotes a pattern of cocaine use damaging physical or mental health but not meeting criteria for dependence syndrome. Other stimulant-related disorders are listed in Table 11.3–1.

Table 11.3–1.
DSM-5 Stimulant-Related Disorders

Stimulant use disorder
Stimulant intoxication
Stimulant withdrawal
Stimulant-induced psychotic disorder
Stimulant-induced bipolar disorder
Stimulant-induced depressive disorder
Stimulant-induced anxiety disorder
Stimulant-induced obsessive-compulsive disorder
Stimulant-induced sleep disorder
Stimulant-induced sexual dysfunction
Stimulant intoxication delirium
Unspecified stimulant-related disorder

Another important aspect of stimulant use disorders is craving for the drug. Individuals with stimulant use disorders describe powerful urges to use the substance. These may arise *de novo,* or may be triggered by access to stimulants, by environmental cues, or by seemingly unrelated stresses.

HISTORY

Cocaine

Cocaine in all its forms is a product of the leaves of the plant *Erythroxylon coca*. This plant is indigenous to the high mountain ranges of South America, where coca leaves have been chewed for at least 15 centuries. With European expansion into the Western hemisphere, the Spanish conquistadores first came into contact with coca use during the 16th century. While at first prohibiting the use of coca leaves, the Spanish subsequently used control over coca distribution to rule local inhabitants,

including paying workers in coca leaves. The Spaniards found in this practice the additional economic benefit of energizing the colonized Indian laborers so they could work long hours in gold and silver mines with little rest.

In turn, the Spanish brought coca leaves back to Europe. However, widespread use of cocaine products did not begin until the active constituent of coca leaves was extracted, first by Friedrich Gaedecke in 1855 and subsequently by Albert Niemann, who coined the term cocaine in 1859. Early experimentation with cocaine peaked around the turn of that century, with luminaries including Sigmund Freud expounding the virtues of the drug. In the United States, cocaine made its way into medicinal tonics and even soft drinks such as Coca-Cola, of which it remained an ingredient until 1903. In the early part of the 20th century, increased awareness of exaggerated and often patently false claims by the manufacturers of cocaine-containing medicinal products raised public concern about spreading cocaine use. At the same time, tales of cocaine addiction began to circulate, and newspaper reports linked the drug with crime. In 1910, President William Taft proclaimed cocaine a national threat. Shortly thereafter, the Harrison Narcotic Act of 1914 was passed, banning the nonmedical use of cocaine and highly regulating its medicinal use. Notwithstanding that cocaine is a stimulant rather than a narcotic, the effect of the Harrison Act was to severely curtail cocaine use, essentially driving it "underground." It remained there, the purview of small subgroups of people such as jazz musicians and the wealthy and privileged until the 1970s.

The reasons for the widespread use of cocaine after 1970 remain somewhat unclear. At the time, cocaine was largely perceived as a nonaddictive, short-acting, "soft" drug. Furthermore, attitudes among people who had experimented with other drugs in the 1960s left them open to trying a substance then seen as relatively safe. Increased production in the 1980s drove down the price of the drug, and cocaine use spread to groups of people who previously could not afford it, reaching epidemic proportions in the United States in the early-to-mid 1980s. Indeed, between 1976 and 1985, cocaine-related medical emergencies rose to nine times their previous level, and cocaine-related deaths increased 11-fold. The mid-1980s also brought about the advent of an inexpensive, ready-to-smoke cocaine alkaloid called "crack," which causes a very powerful high because of its rapid onset and offset.

Estimates of active cocaine use peaked in 1985 at 5.7 million, or 3 percent of the US population. After that point, cocaine use began to fall dramatically, with a 20 percent drop in use between 1986 and 1987 alone, and a 66 percent decrease by 1992. While the precise factors leading to decreased cocaine use after its peak in the 1980s remain unclear, it coincided with a number of possibly related events. In part due to a far-

reaching awareness campaign along with well-publicized celebrity deaths that were attributed to cocaine, the perceived danger of cocaine use began to increase. The previously glamorous position of cocaine changed as the drug became associated with addiction, overdose deaths, crime, fetal injury, and the problems of the urban poor. At the same time, enforcement of anti-drug laws became more aggressive, with increased arrests and penalties for cocaine-related offenses.

Cocaine use in the United States has often followed cycles of increased and decreased prevalence. At times, dramatic drops in use have been driven by government actions such as the Harrison Act of 1914 or the police and public education campaigns of the mid-to-late 1980s. Less dramatic fluctuations have been evident over shorter periods, including year-to-year variations. Since its nadir in the early 1990s, the prevalence of cocaine use and related disorders has risen at times, but it has not yet approached the peak levels seen in the 1980s.

Amphetamine-Type Stimulants

Like cocaine, the use of amphetamine-type stimulants dates back many centuries. Ma-Huang, a form of the stimulant ephedrine, has been used as a Chinese herbal remedy for thousands of years. In 1887, the active ingredient ephedrine was isolated by Nagayoshi Nagai. At the same time, amphetamine was synthesized in Germany by Lazar Edeleanu. Methamphetamine was first synthesized in Japan in 1918. Therapeutic applications of amphetamines in Western civilizations began in the 1930s with the use of amphetamine (Benzedrine) as a bronchodilator. At that time, the cognitive and emotional effects of amphetamines were recognized, including reinforcing properties such as euphoria. Indeed, amphetamines were widely used during World War II to counteract soldiers' fatigue. Following the war, the first amphetamine epidemic took place in Japan, where surplus military supplies of methamphetamine led to widespread use.

An increasingly abundant supply of amphetamines in the decades that followed was associated with peaking rates of misuse in the United States during the 1960s. At the end of the decade, regulatory agencies responded with efforts to decrease the manufacture of amphetamines to more closely match medicinal need. During this period, amphetamines were reclassified as Schedule II medications, reflecting closer government oversight and regulation than in previous decades. Despite these efforts, American misuse of methamphetamine increased dramatically beginning in the mid-1990s as illicit manufacturing sources took over. Methamphetamine use remains a significant problem in certain geographic regions of the United States, including Hawaii and some western and midwestern states. Prescription amphetamines continue to be used for therapeutic purposes including the treatment of attention deficit hyperactivity disorder (ADHD)

and some sleep disorders.

COMPARATIVE NOSOLOGY

In diagnosing stimulant-related disorders, clinicians are guided by two sets of clinical criteria. These are delineated in the American Psychiatric Association's DSM and in the World Health Organization's ICD. As is the case for other SUDs, the advent of DSM-5 heralded a consolidation of previously separate stimulant abuse and dependence diagnoses into one category: stimulant use disorder, with level of severity specified on a spectrum of mild, moderate, or severe. Reflecting the high degree of overlap in drug effects, patterns of use, and symptoms, the previously distinct diagnostic entities for cocaine and amphetamine-related disorders were combined to one comprehensive category, stimulant-related disorders.

The updated criteria for stimulant use disorder according to DSM-5 eliminate substance-related legal problems (previously a criteria for stimulant abuse) and add craving as a symptom. Other criteria that comprised the DSM-IV Stimulant Abuse diagnosis (i.e., failure to fulfill major role obligations, substance use in situations in which it is physically hazardous, and use despite persistent social and interpersonal problems) have been subsumed by the broader DSM-5 Stimulant Related Disorder diagnosis.

ICD-10 classification more closely reflects DSM-IV-TR than the new DSM-5 criteria. Cocaine and other stimulants are treated separately, and distinct diagnoses of "dependence syndrome" (similar to DSM-IV-TR substance dependence) and "harmful use" have been retained. Field trials demonstrated a good level of agreement between DSM-IV and ICD-10 diagnoses of cocaine and amphetamine dependence. It is therefore expected that there would be some level of agreement between ICD-10 and DSM-5. An updated version of ICD (ICD-11) is expected in 2017.

Previously separate Cocaine- and Amphetamine-Induced Disorders were likewise consolidated into the Stimulant Use Disorder category. DSM-5 criteria for Stimulant Intoxication, Stimulant Withdrawal, Stimulant-Induced Sleep Disorder, and Stimulant-Induced Sexual Dysfunction remain largely unchanged from their DSM-IV Cocaine- and Amphetamine-Induced counterparts. Stimulant-Induced Psychotic Disorder criteria became somewhat more stringent, with a requirement that psychotic symptoms be both temporally correlated with stimulant exposure/withdrawal and characteristic of stimulant-related symptoms; DSM-IV had required only one such criteria. DSM-IV Cocaine- and Amphetamine-Induced Mood Disorders were shifted to categories of Stimulant-Induced Bipolar Disorder and Stimulant-Induced Depressive Disorder. Stimulant-Induced Anxiety Disorder no longer includes obsessive or compulsive symptoms as criteria, with the creation of a new diagnosis

for Stimulant-Induced Obsessive-Compulsive Disorder. Stimulant Intoxication Delirium exists in both DSM-IV and DSM-5, and should be diagnosed when symptoms are in excess of those typical for stimulant intoxication. Stimulant-Induced Neurocognitive Disorders replace Dementia and Amnestic Disorders. In DSM-5, Unspecified Stimulant Use Disorder corresponds to DSM-IV Cocaine and Amphetamine Use Disorder Not Otherwise Specified.

EPIDEMIOLOGY

Lifetime cocaine use among Americans in 2013 was estimated at nearly 37 million people, 1.5 million of whom had used in the past month. The number of Americans who have used cocaine in the past year has remained relatively stable over the last 10 years, at 4.7 million people in 2013 according to available data. Estimates of current users (those who reported use in the past month) have also remained stable or decreased since the early 1990s. In 2013, it was estimated that almost 22 million Americans had ever engaged in nonmedical use of amphetamine-type stimulants, with 1.3 million having used in the past month. The worldwide burden of psychostimulant use speaks to an even more significant problem. The United Nations Office on Drugs and Crime (UNODC) World Drug Report 2013 estimated that nearly 34 million people worldwide had used amphetamine-type stimulants in the past year. North America remains an important area of concern, while use appears to be rising in Asia and Africa. Likewise, one percent of the Australian population currently uses cocaine according to the UNODC World Drug Report 2013.

The number of individuals using a substance for the first time in the past year offers a useful measure of potential trends in drug use. Initiation of cocaine remained relatively stable in the past decade, with 601,000 new users in 2013. The average age at onset of cocaine use was 20.4 years. Likewise, new onset of stimulant use has remained stable, with 603,000 new nonmedical users of psychotherapeutic stimulants and 144,000 new users of methamphetamine in 2013. Estimated prevalence of cocaine use disorders (DSM-IV cocaine abuse or dependence) decreased from 1.5 million individuals in 2003 to 855,000 in 2013. Also in 2013, an estimated 469,000 individuals in the United States suffered from an amphetamine-type stimulant use disorder.

The National Household Survey on Drug Use and Health (NSDUH) is one of the primary sources of knowledge of patterns and prevalence of drug use throughout the United States, including the use of stimulants. NSDUH information on the prevalence and distribution of cocaine use has been available since 1991, and measures of SUDs have been included in the questionnaire since 2000. Males are about twice as likely to use cocaine as females, while methamphetamine use is divided more evenly over gender lines. Cocaine is used across ethnicities, but inhaled “crack” cocaine use is

more common among African Americans.

The annual prevalence of cocaine-type stimulant use disorder in the United States in 2013 was estimated at 0.3 percent of the population, or 855,000 Americans. The age group most heavily affected was 18- to 25-year-olds, who remain twice as likely as the general population to have met criteria for cocaine use disorder in the past 12 months. A total of 469,000 Americans met criteria for amphetamine-type stimulant use disorder in 2013; again, the 18- to 25-year-old age group was the most heavily affected. This may reflect higher exposure to the drug, as this group is also the most likely to currently use cocaine or engage in nonmedical use of prescription-type psychotherapeutics including stimulants. Men who have sex with men represent another group with disproportionately high levels of methamphetamine use. In this population, methamphetamine is sometimes used to enhance sexual activity, and can be associated with risky sexual practices and sexually transmitted diseases.

The Treatment Episode Data Set (TEDS) is a component of the Drug and Alcohol Services Information System (DASIS), a national monitoring network of drug and alcohol treatment in the United States. TEDS provides information on the numbers and characteristics of patients admitted for treatment of SUDs. Data from this survey indicate that the use of cocaine and other stimulants accounted for 14 percent of all admissions for SUD treatment in 2012, following alcohol (21 percent), marijuana (17.5 percent), and heroin (16.3 percent). The Drug Abuse Warning Network (DAWN) survey of emergency department treatment presents a somewhat more alarming picture. The DAWN report for 2011 revealed that cocaine remains the most common drug other than alcohol involved in drug-related emergency treatment. Furthermore, emergency department visits related to illicit use of stimulants as a group increased substantially from 2007 to 2011, driven largely by an increase in medical problems caused by misuse of methamphetamine and prescription psychostimulants. Over the same time period, the purity of "street" methamphetamine increased by 130 percent according to the U.S. Department of Justice Drug Enforcement Administration National Drug Threat Assessment. The DAWN report for 2011 noted synthetic stimulants labeled as "bath salts" to be involved in over 20,000 drug-related emergency department visits. This was the first measureable number of such treatment episodes to be noted, and it is expected that bath salts will continue to be a problem in the future.

"Bath Salts"

Catha edulis, or "Khat" is a plant indigenous to Eastern Africa and Southern Arabia. Psychoactive chemicals released from the plant include cathinone and cathine, compounds with stimulant properties. Khat is used like chewing tobacco, held between the cheek and gums after chewing. NIDA estimates that as many as 10 million people worldwide use Khat,

including many in Eastern Africa and the Middle East, as well as an unknown number of immigrants to the United States.

In the mid-2000s, synthetic cathinones began to appear, marketed as "legal high" chemicals. They were packaged under labels including "screen cleaner," "jewelry cleaner," and most notoriously "bath salts." They are generally produced as powders marked "not for human consumption" and can be taken by oral, intranasal, inhaled (smoked) or intravenous routes. These compounds are substituted cathinones that have been chemically modified to avoid regulatory control. There are a large number of specific compounds in the class, and a given package often contains a mix of agents, further increasing risk of toxicity. Generally, substituted cathinones increase synaptic catecholamine levels by inhibiting the dopamine, serotonin, and norepinephrine reuptake transporters in a fashion similar to cocaine. Most also facilitate catecholamine release, similar to methamphetamine. Animal models indicate strong reinforcing properties of these compounds, which are considered to have high addictive potential.

Drug screens for methamphetamine and PCP may cross-react with some substituted cathinones, but this is not a reliable test. While specific assays for some cathinones are commercially available, gas chromatography is the most reliable means of detecting synthetic cathinones at this time.

ETIOLOGY

Drug use disorders are believed to result from the complex interplay of cultural, environmental, familial, and neurobiological influences. Environmental factors such as accessibility and perceived danger or stigma related to use may influence whether an individual initially tries psychostimulants or other drugs. For individuals with stimulant use disorders, environmental precipitants may increase cortisol and other stress-related hormones, increasing risk for relapse or persistent use of drugs even when there is a desire to stop. Genetic factors may both increase the likelihood of initially using illicit substances (e.g., when impulsive personality characteristics are inherited), and confer genetic vulnerability for developing SUDs.

Mechanism of Action

Neurotransmitter Systems

DOPAMINE. It is well established that stimulants directly produce changes in dopaminergic neurotransmission. In vitro animal studies using radiolabeled cocaine or cocaine analogues demonstrate high amounts of binding in dopamine-rich brain regions including the caudate-putamen and the ventral tegmental area (VTA). Moderate levels of cocaine binding

also occur in the substantia nigra, amygdala, hypothalamus, and locus coeruleus. In humans, cocaine rapidly permeates the striatum, then quickly redistributes out of it with a half-life of about 20 minutes. The time course of self-reported "high" closely follows this pattern of rapid uptake and clearance. Likewise, amphetamines have profound effects in brain reward regions including the mesolimbic regions (VTA and nucleus accumbens). Furthermore, methamphetamine is distributed even more diffusely and persists longer in the brain than does cocaine.

Dopaminergic projections from the VTA to the nucleus accumbens are most strongly identified with drug-related reinforcement. Cocaine administration causes transient increases in extracellular dopamine levels in these regions by binding to the presynaptic dopamine transporter (DAT), thereby inhibiting dopamine reuptake. Amphetamines not only inhibit dopamine reuptake, but are also believed to reverse the function of the DAT. By this mechanism and through promoting release of dopamine stored in vesicles, amphetamines potently increase synaptic dopamine levels. Transient changes in dopamine levels play a primary role in the reinforcing properties of all psychostimulants. Subjective euphoria appears to be a function of both DAT occupancy (with a threshold of 50 percent occupancy to be detectable) and rapidity of onset, with the strongest responses occurring when dopaminergic neurotransmission is increasing dynamically.

OTHER NEUROTRANSMITTER SYSTEMS. While it is widely accepted that dopamine plays a primary role in the effects of stimulants, more complex models are required to completely understand drug-related behaviors and disorders. By binding to monoamine transporters, psychostimulants can inhibit reuptake of serotonin and norepinephrine as well of dopamine. Other neurotransmitters implicated in stimulant use disorders include glutamate and the inhibitory neurotransmitter GABA. Glutaminergic projections to the nucleus accumbens modulate regional dopamine neurotransmission, and glutamate antagonists can block the behavioral effects of stimulants in laboratory animals. Similarly, GABA affects dopamine neurons in the nucleus accumbens and VTA. Reduced glutamate levels during abstinence that follows psychostimulant exposure lead to upregulation of AMPA receptors in the nucleus accumbens. Such changes are associated with drug-related behaviors in lab animals, and glutamate antagonists can block reinstatement of drug-seeking behavior.

Learning and Memory

Positive Reinforcement Mechanisms. Stimulant use disorders and related behaviors can be understood as forms of behavioral conditioning. *Operant conditioning* is a type of learning in which the likelihood of a given behavior is controlled by its consequences. *Reinforcers* are those stimuli

that increase the likelihood of the behavior in question. As such, stimulants and other substances of abuse operate as powerful reinforcers, increasing the likelihood of behaviors that result in opportunities to obtain drugs (i.e., "drug-seeking behaviors").

Principles of operant conditioning explain why animals can learn to self-administer reinforcing substances like cocaine. Other drug-associated responses can be accounted for by *respondent conditioning*. In this form of learning, an unconditioned stimulus (one that produces a response without training) can be paired with a stimulus that normally does not produce a response. Pairing the two stimuli for a period of time produces a conditioned response, in which the animal responds to the initially neutral stimulus as though the unconditioned stimulus were also present. In this way, drug-associated cues such as paraphernalia, people, or places associated with use, and even times of day, can elicit urges to use stimulants. This phenomenon may also account for cocaine-dependent individuals who experience a subjective "high" while preparing to take drugs, as the practice of obtaining or handling drugs leads to conditioned euphoric responses.

During the early stages of a stimulant use disorder, drug use is reinforced by increased dopamine transmission from the VTA via the medial forebrain bundle to the nucleus accumbens. This fits a classic model of drug addiction based on positive reinforcement, in which cues associated with drug use gain emotional valence through learned associations involving brain regions such as the hippocampus and limbic structures (in particular, the amygdala). Abnormalities in these brain areas, which are involved in essential pathways of learning and memory, are associated with development of stimulant and other SUDs. Such dysfunction may explain persistent use of psychostimulants to the exclusion of other, natural reinforcers such as food. Conventional reinforcers produce a dopamine response in the VTA and amygdala that attenuates over time. In contrast, the response to stimulants does not diminish in this way, and some laboratory animals will self-administer cocaine in preference to food and water to the point of death.

Respondent conditioning explains the activation of drug cravings by stimuli associated with drug use. Cravings that are induced by priming doses of stimulants involve a brain circuit projecting from the VTA to the medial prefrontal cortex and the nucleus accumbens. Cravings that are induced by associated cues ("people, places, and things") activate this circuit under influence of the basolateral amygdala. It is now widely recognized that stress can also elicit drug craving. This is mediated by expression of corticotropin-releasing factor in limbic areas that include the amygdala, modulating the activity of dopaminergic projections from VTA to nucleus accumbens.

Negative Reinforcement Mechanisms. In addition to reward mechanisms described above, a pattern of sustained stimulant use through negative reinforcement (avoidance of the aversive withdrawal/abstinence state) may augment and even supplant positive reinforcement. This is mediated by what may simplistically be termed a "dopamine-deficient state" caused by dopamine receptor downregulation along with neurotoxic effects on dopaminergic neuron systems. Other neurotransmitters such as glutamate also show altered function with chronic stimulant misuse. This likely contributes to states of dysphoria and drug craving that occur with cessation of use. Indeed, it has been demonstrated that abstinent methamphetamine users who experience persistently low levels of striatal dopamine activity may be at increased risk for relapse. Lower availability of D2 and D3 receptors in methamphetamine users has also been posited to be associated with increased impulsivity and "steep temporal discounting," the tendency to disproportionately value small, immediate reinforcers over delayed rewards of greater magnitude.

Compulsive Use. Beyond stimulant use compelled by positive and negative reinforcement, stimulant use disorders may also be marked by what is sometimes termed "compulsive" or "habitual" use. This involves apparently automatic, perseverative drug use even in the face of severe consequences, and long after many patients describe loss of previous euphorigenic drug effects. Such behavior appears to be associated with changed function in the following pathways associated with chronic stimulant use.

CAMP RESPONSE ELEMENT–BINDING PROTEIN. Dopamine receptors act through second-messenger systems that alter levels of cyclic adenosine monophosphate (cAMP). Dynamic interplay between dopamine receptor subtypes leads to alterations in cAMP, which in turn can increase activity of protein kinase A (PKA). PKA levels in the nucleus accumbens are increased with chronic stimulant administration. Elevations in PKA lead to increased activity of cAMP response element–binding protein (CREB), itself implicated in cocaine-related behaviors including place preference and sensitization. Downstream targets of CREB involve regulation of glutamate receptors and of dynorphin.

DYNORPHIN. Dynorphin (DYN) is the endogenous ligand for kappa-opioid receptors (KOPr). It is heavily expressed in brain regions associated with drug-mediated reward, and its activity is regulated by CREB. Stimulant administration leads to increased activation of dopamine D1 receptors by elevating synaptic dopamine levels. This leads to activation of CREB by the cAMP-PKA-CREB cascade. CREB binds to regions of the prodynorphin promoter, thereby enhancing DYN production. Chronic

activation of the DYN-KOPr system downregulates dopamine function. Hence, increased activity of the DYN-KOPr pathway may initially serve a homeostatic function during heavy psychostimulant use by helping counteract some effects of abnormally elevated levels of extracellular dopamine. However, cessation after chronic use could lead to high levels of unopposed activity within this system. Such an imbalance may help produce the sustained changes in behavioral and affective states seen with chronic stimulant use.

NEUROENDOCRINE DYSREGULATION AND STRESS. The role of environmental stress in exacerbating or precipitating episodes of mental illness has long been recognized. Individuals with SUDs may experience relapse or increased symptoms in the presence of psychosocial stressors. Animal models of stress and addiction demonstrate that stress-inducing situations (e.g., a shock to the foot) or administration of corticosteroids can increase the likelihood of drug self-administration. Conversely, antagonists of corticotrophin-releasing hormone (CRH) or glucocorticoids decrease stress-induced drug self-administration in laboratory animals by blocking CRH activity in the VTA. Drug withdrawal itself is associated with increases in ACTH, corticosterone, and CRH in the amygdala. These changes are associated with the anxiety and craving often seen in early abstinence from stimulants.

COGNITIVE/EXECUTIVE DYSFUNCTION. Habitual drug use involves a pattern of executive dysfunction, poor impulse control, and cognitive inflexibility resulting in difficulty attributing appropriate reward values to stimuli, changing behavior in response to environmental feedback, and inhibiting impulsive or habitual behavior. Chronic stimulant use is associated with anatomic, metabolic, and vascular changes in areas of the brain responsible for such higher-order cognition. Structurally, cocaine and methamphetamine use disorders are associated with reduced gray matter density in the frontal cortex. Other findings include gray matter deficits in limbic areas and low hippocampal volume. Conversely, enlarged basal ganglia (areas important for neurotransmission of dopamine) have been demonstrated with stimulant use disorders. A dose-dependent decrease in gray matter volume of the orbitofrontal cortex is associated with increased compulsivity of drug use. Abnormal connectivity in prefrontal regions is evident in both individuals with stimulant use disorders and their unaffected siblings, suggesting a condition that preexists drug exposure. Diminished connectivity between the anterior cingulate cortex (an area active during inhibition of impulses) and the amygdala is associated with increased risk of relapse in those with cocaine use disorders. Stimulant users manifest decreased blood flow in the prefrontal cortex, with frequent users showing the largest deficits. Decreased metabolic activity is also seen

in frontal regions, and may persist for months after cessation of use. Metabolic dysfunction in frontal regions, including the orbitofrontal cortex, may be related to the development of compulsive aspects of drug addiction. Individuals with stimulant use disorders show deficits in verbal, spatial, and recognition memory, functions that are generally attributed to the hippocampus. Other prominent deficits include impaired attention, decreased cognitive flexibility, impaired decision making, impulsivity, and poor delayed reward discounting. This pattern of neurobiological vulnerabilities clearly contributes to persistence of maladaptive drug use even among those who state that they want to change.

Routes of Administration and Metabolic Factors

The pharmacokinetic properties of stimulants are strongly influenced by how they are taken. In particular, the time it takes a drug to initially appear in the blood and to reach its peak serum concentration is strongly related to its mode of administration. Smoking and intravenous use of cocaine produce the most rapid pharmacologic and subjective onset of action, as well as faster termination of subjective effects. Cocaine-induced euphoria is more strongly correlated with rate of increase in serum levels than with absolute amount of cocaine ingested. As such, the use of cocaine by these rapid routes is more likely to be associated with progression to compulsive use and cocaine dependence. Similarly, injecting or smoking methamphetamine is associated with much faster penetration of the blood–brain barrier, on the order of seconds, in comparison to use by oral or intranasal routes. Injecting methamphetamine is typically associated with more severe SUDs and with poorer prognosis.

Cocaine is rapidly metabolized by liver and serum esterases, with a variable half-life on the order of 50 to 90 minutes. In humans, cocaine is catabolized by hydrolysis of its ester groups to benzoylecgonine and ecgonine methyl ester in approximately equal proportions. Neither has biological activity in humans. In the presence of alcohol, cocaine is converted by liver esterases into ethyl cocaine, or cocaethylene. Cocaethylene produces cocaine-like pharmacologic and subjective effects, has a longer half-life, and may be more toxic than cocaine. Only about 1 to 5 percent of cocaine is excreted unchanged in the urine, with the rest excreted as metabolites. Urinary benzoylecgonine can generally be detected for 2 to 3 days after a binge, although extremely heavy users may have measurable urinary benzoylecgonine levels as long as 10 to 14 days after last use.

The half-life of amphetamine in adults is on the order of 13 to 14 hours. Amphetamine undergoes oxidation by liver enzymes (primarily cytochrome P450 2D6) into the active metabolites 4-hydroxyamphetamine and norephedrine and the inactive metabolite alpha-hydroxy-amphetamine. Further oxidation of active metabolites produces the inactive metabolite 4-

hydroxy-norephedrine. About 30 percent of amphetamine is excreted unchanged in the urine, with higher rates of excretion with more acidic urine. The rest is excreted as metabolites. Amphetamine is typically detectable in urine for 3 to 5 days after binge use.

Genetic Factors

Twin studies have allowed an estimate of the genetic component of cocaine dependence. In these studies, the heritable component of SUDs has been measured by comparing monozygotic and dizygotic twin pairs. Concordance rates for cocaine dependence among monozygotic twins are substantially higher than for dizygotic twins. Estimates of the heritability of cocaine dependence based on these studies are high (60 to 80 percent). Some of the apparent heritability of cocaine and other SUDs may be explained by inheritance of traits associated with drug use. In particular, heritable personality characteristics such as a high degree of novelty seeking and low harm avoidance have been linked with initiation of drug use. Both individuals with stimulant use disorders and their unaffected siblings show deficits in executive function, high levels of trait anxiety, and poor response control. Those with stimulant use disorders show more problems across these domains than do their siblings, suggesting both a genetic state predisposing to drug addiction and exacerbation of deficits with persistent stimulant use.

Given the polygenic nature of drug addiction, and the contribution of environmental factors such as availability of drugs, identifying particular genes directly involved in the etiology of stimulant use disorders is challenging. Furthermore, associations between stimulant use disorders and particular genetic loci in one population may not apply to another. Nonetheless, a picture of genetic contributions to stimulant use disorders is gradually beginning to emerge.

Genetic variations of the D2 dopamine receptor and in RGS17, which regulates G-protein signaling, have been implicated in a number of SUDs, including cocaine. A genome-wide association study identified the locus FAM53b as associated with risk for cocaine use disorders. Genes associated with serotonin neurotransmission, including the serotonin receptors 5-HT3A, 5-HT3B, and HTR2A and the serotonin transporter SLC6A4 have also been implicated in cocaine use disorders. Butyrylcholinesterase genetic variants and the dopamine transporter SLC6A3 are specifically associated with crack cocaine use. Other genes associated with cocaine use disorders include variants of nicotinic receptor subunits CHRNB3, CHRNA3, and CHRNA5; the dopamine receptor-interacting protein NCS-1; the delta-opioid receptor OPRD1; prodynorphin (PDYN); and the cannabinoid receptor CNR1. The circadian regulator gene PER2 may confer risk for cocaine use disorders through lower dopamine D2 receptor levels in the striatum.

Risk of methamphetamine use disorder is associated with variants of adenosine A2A receptor, brain-derived neurotrophic factor (BDNF), prokineticin 2 receptor gene (PROKR2), and glial cell line-derived neurotrophic factor (GDNF). In Japanese populations, methamphetamine use disorder is associated with variants of the protein kinase AKT1, mu-opioid receptor interacting protein beta-arrestin (ARRB2), the catecholamine-metabolizing enzyme COMT, glutathione-S-transferase (GTSP1), the mu-opiate receptor OPRM1, and the neuropeptide 1 receptor NPYR1. Other genes associated with methamphetamine use disorders include the D4 dopamine receptor, GABA receptors GABRA1 and GABRG2, cytochrome P2D6, the glycine transporter GLYT1, PDYN, PICK1, and the cation transporter SLC22A3.

DIAGNOSIS AND CLINICAL FEATURES

Stimulant Use Disorder

Stimulant use disorder is characterized by a maladaptive pattern of psychostimulant use accompanied by multiple drug-related problems in different domains. Stimulant use disorder is diagnosed when at least two of 11 possible criteria are met. As with other substances, tolerance and withdrawal are neither necessary nor sufficient for a diagnosis of stimulant use disorder. Criteria include: (1) stimulant use in greater quantities or durations than intended, (2) persistent desire to curtail use, (3) significant time spent engaging in drug-related behavior, (4) craving for the desired stimulant, (5) recurrent use to the exclusion of other obligations, (6) persistent use despite social or interpersonal problems caused by use, (7) curtailing other important activities due to stimulant use, (8) use when it is physically hazardous, (9) use in spite of related physical or psychological harm it causes, (10) tolerance, and (11) withdrawal.

The category of substance that is abused (amphetamine-type, cocaine, or other/unspecified) is named. The severity of stimulant use disorder is graded as mild (two to three symptoms), moderate (four to five symptoms), or severe (six or more symptoms). Early remission is defined by the absence of all symptoms for between 3 and 12 months, and sustained remission by the absence of all symptoms for more than 12 months. The additional specifier “In a controlled environment” is used when the individual is in an environment where access to stimulants is limited.

While DSM-5 criteria for stimulant use disorder are identical to those for alcohol and opioids, patterns of use may be quite different. Although some individuals with stimulant use disorders are daily users who experience tolerance and withdrawal, many others abstain intermittently between periods of heavy use, either to recover from the effects of the drug or as a result of depleted finances that prevent access to the desired drug. During periods of use, frequency of administration varies based on the

duration of drug effects. For instance, cocaine is generally taken more frequently than is methamphetamine, reflecting cocaine's shorter half-life.

Stimulant Intoxication

Stimulant intoxication is marked by the appearance of characteristic symptoms shortly after use of an amphetamine-type substance, cocaine, or other stimulant. It may occur within minutes after ingestion of the substance. The duration of intoxication is typically briefer for cocaine than other stimulants, due to its shorter half-life. Characteristic symptoms are divided into psychological or behavioral and physiological effects. Psychological symptoms may include changes in mood state, social interactions, or judgment. Physical symptoms include adrenergic effects (changes in heart rate, blood pressure, thermoregulation, psychomotor changes, pupillary dilation), GI disturbances, motor changes, cardiopulmonary symptoms, and neurological problems. Both a significant psychological disturbance and two or more physical symptoms are required for diagnosis. Severe intoxication is marked by evidence of toxicity, which may include grand mal seizures, cardiac arrhythmia, hyperpyrexia, and death. The presence of perceptual disturbances is specified when drug-induced hallucinations or delusions occur during intoxication. If such perceptual disturbances are present and there is no insight that they are induced by the substance, the diagnosis of stimulant-induced psychotic disorder should be considered instead.

In susceptible individuals, paranoia tends to worsen during the course of a binge, and to appear earlier in the course of subsequent binges as use continues. Psychotic symptoms are common with both cocaine and methamphetamine. They may prompt attempts to "self-medicate" with alcohol or opioids. While cocaine users who experience delusions may complain of symptoms persisting beyond peak plasma cocaine levels, the duration is generally brief, lasting hours in most cases. Typically, paranoia consists of worries that police, drug dealers, or others are pursuing or planning to attack the cocaine user. Perceptual disturbances are usually limited to misinterpretations of existing stimuli (e.g., mistaking a car alarm for a police siren). Psychotic symptoms associated with methamphetamine use may be more difficult to distinguish from primary psychotic disorders. Symptoms caused by methamphetamine generally remit within a week of cessation of use. Persistent symptoms suggest the presence of an underlying psychotic disorder.

Stimulant Withdrawal

Stimulant withdrawal is common after periods of prolonged or intense use, and is marked by dysphoric mood accompanied by two or more characteristic physiologic symptoms: fatigue, vivid and unpleasant dreams,

and disturbances of sleep, appetite, and psychomotor activity. Bradycardia, drug cravings, anhedonia, and other depressive symptoms frequently occur.

There is substantial symptomatic overlap between characteristic physiologic symptoms of stimulant withdrawal and the neurovegetative symptoms of depression. At times, depressive symptoms would meet criteria for a major depressive episode were they not the direct consequence of drug use. In such cases, the diagnosis of stimulant-induced depressive disorder should be considered. Suicidal ideation or behavior can also occur in association with stimulant withdrawal or stimulant-induced depressive disorder.

Stimulant-Induced Psychotic Disorder

Brief periods of stimulant-induced paranoia are common in heavy users, occurring in 50 to 70 percent of individuals with cocaine use disorders and nearly 30 percent of those who use methamphetamine. While paranoia is common with heavy use of cocaine or methamphetamine, the diagnosis of stimulant-induced psychotic disorder is reserved for cases in which symptoms exceed the quality or severity characteristic of stimulant intoxication.

Increased risk of cocaine-induced paranoia is linked to heavy exposure to the drug, as measured by duration and quantity of use. Increased risk is also associated with routes of administration other than intranasal, and with co-occurring ASPD or cannabis use disorder. Likewise, methamphetamine-related psychotic symptoms are more likely with heavy use. Psychotic symptoms are also associated with increased risk of violence among individuals who use methamphetamine. People with ADHD appear at increased risk for either cocaine-induced or methamphetamine-induced psychotic symptoms. There may also be a genetic predisposition to stimulant-induced paranoia. Studies have identified increased levels of cocaine-induced paranoia in association with variations in dopamine-beta-hydroxylase and the dopamine transporter. Methamphetamine-induced psychosis in Japanese populations is associated with variants of the glutamate receptor (GRM2) and the serotonin receptor (5HT1A). Persistent psychotic symptoms related to methamphetamine are associated with a polymorphism of the dopamine transporter.

Other Stimulant-Induced Disorders

Mood changes and anxiety are common with stimulant use, intoxication, and withdrawal. Stimulant-induced depressive, bipolar or anxiety disorders are diagnosed when the severity of these symptoms warrants treatment and exceeds that usually seen with intoxication or withdrawal. Stimulant-induced mood disorders can resemble depressive, manic, hypomanic, or

mixed episodes. Symptoms in the manic spectrum are usually present in the setting of intoxication, while depressive symptoms often occur during withdrawal. Suicidal ideation or behavior can also occur in association with stimulant withdrawal or stimulant-induced mood disorders. Compulsive symptoms such as stereotyped behavior and skin picking are also associated with stimulant intoxication.

Stimulant-induced mood or anxiety disorders are distinguished from mood and anxiety symptoms expected during the typical course of intoxication or withdrawal by their intensity, onset, and time course. While neurovegetative symptoms suggesting a major depressive episode may be present, full criteria for major depressive disorder or hypomania need not be met. Similarly, the diagnostic criteria for stimulant-induced anxiety or obsessive-compulsive disorder require only the presence of prominent anxiety, obsessions, compulsions, or panic attacks in the setting of recent use of a psychostimulant and out of proportion with typical signs of intoxication.

Other stimulant-induced disorders include stimulant-induced sexual dysfunction, sleep disorder, and unspecified stimulant-related disorder. Stimulants such as cocaine or methamphetamine are at times used for their purported enhancement of sexual function. Chronic use of stimulants is associated with sexual dysfunction either during use or within 1 month of use. Psychostimulants may also produce a sleep disorder, typically characterized by periods of insomnia during intoxication and hypersomnia during withdrawal. Insomnia during stimulant withdrawal is also commonly described. Sleep cycles may be shifted when binges extend into the night or when withdrawal-induced hypersomnia leads to daytime sleeping.

Unspecified Stimulant-Related Disorder

The category of Unspecified Stimulant-Related Disorder is used for disorders that do not meet the full criteria for a specific stimulant-related disorder, but that nonetheless cause significant impairment or distress.

COMORBIDITY

Studies of the prevalence of non–substance-related psychiatric disorders among individuals with stimulant use disorders have consistently found elevated levels of psychiatric disorders in this population. Results from the community-based Epidemiologic Catchment Area (ECA) Study demonstrated elevated rates of non–substance-related mental disorders in those with cocaine use disorders. In this study, the estimated lifetime prevalence rate for mental disorders was 76 percent, or 11 times that found in the general population; nearly 85 percent of those with cocaine use disorders also had co-occurring alcohol abuse or dependence. Studies that

have compared treatment-seeking cocaine-dependent individuals to those who do not seek treatment have demonstrated higher rates of major depression and of ADHD among the former group. Likewise, most studies report high levels of co-occurring other SUDs in individuals with amphetamine-type stimulant use disorders. Mood disorders, psychotic disorders, and anxiety disorders are also common. While psychotic symptoms are most commonly methamphetamine-induced, mood and anxiety disorders are often not substance-induced.

TOXICITY AND MEDICAL COMPLICATIONS

Stimulant use can cause medical complications both as a result of its direct toxic effects and because of complications related to methods of preparation and administration. Stimulants are potent sympathomimetic drugs, causing physiologic effects including vasoconstriction, increased heart rate, and elevated blood pressure through adrenergic stimulation. The combination of increased myocardial demand with constriction of coronary blood vessels can precipitate angina pectoris and even myocardial infarction. Acute hypertension associated with stimulant use can also predispose to arterial hemorrhage, such as hemorrhagic stroke or aortic dissection. Vasoconstriction can involve damage to other end-organs, including the brain, kidneys, or intestines. With intranasal use, decreased blood flow to the nasal mucosa frequently causes nasal symptoms (e.g., rhinorrhea, nosebleeds) and mucosal ulceration. A combination of vasoconstriction and bruxism associated with the use of methamphetamine causes markedly poor dentition, or "meth mouth." Grand mal seizures are common with stimulant overdose, and may occur with routine use as well. Furthermore, individuals who misuse stimulants regularly may develop neuropsychological deficits, including problems with attention, learning, and memory. In many cases, these abnormalities persist during periods of abstinence.

Stimulant misuse can also precipitate behaviors that place an individual at increased risk of harm by violence or accidental injury. Unsafe sexual behaviors may lead to transmission of HIV or other sexually transmitted diseases. Driving while intoxicated and other hazardous behaviors may lead to accidental physical harm. Finally, the associated illegal behaviors involved in obtaining cocaine or methamphetamine increase the likelihood of being a victim of violence.

Pathology and Laboratory Examinations

The physical and laboratory assessment of patients with stimulant-related disorders focuses on clarifying the diagnosis, excluding other possible diagnoses, and assessing for co-occurring problems or related complications. Metabolites of cocaine and amphetamine can be detected in

blood, hair, sweat, saliva, and urine. In clinical practice, urine and blood testing are most useful. Blood testing can be helpful to assess for acute intoxication when very recent ingestion is suspected. Urine testing is often used as a means of monitoring patients for relapses and detecting concurrent use of other illicit substances.

Relative to other substances of abuse, the short half-life of cocaine renders it more difficult to discover occasional use with urine drug screening. On the other hand, with heavy use of cocaine, its breakdown products accumulate in fat tissue. As they gradually diffuse out of this body compartment, cocaine metabolites may be detectable for up to 2 weeks after cessation in some extremely heavy users. Methamphetamine is generally detectable for longer periods of time than is cocaine, on the order of three to five days after brief use. Synthetic cathinones present more of a challenge, as many are not detectable by routine drug screening. Gas chromatography may be necessary when cathinone ingestion is strongly suspected.

Other examinations are guided by historical or physical evidence of stimulant-related complications. Seizures or other acute neurological deficits may prompt assessment by brain computed tomography (CT) or magnetic resonance imaging (MRI) to detect physical abnormalities such as hemorrhagic or ischemic strokes. The presence of chest pain or other cardiovascular abnormalities may prompt assessment with an electrocardiogram (ECG) or an assay for serum cardiac enzymes to evaluate for coronary ischemia. Similarly, respiratory symptoms may indicate need for a chest X-ray to rule out pneumonitis. Likewise, signs of infection in the setting of injection use may require blood cultures to exclude bacteremia or endocarditis. A history of any drug use by injection places individuals at risk for hepatitis (type B or C) and HIV; testing for these co-occurring medical disorders is thus often indicated.

As described above, chronic stimulant misuse produces changes in levels of stress-related hormones, including CRH and cortisol. Stimulant use can also elevate levels of prolactin, luteinizing hormone, and follicle-stimulating hormone. However, there is little clinical utility to assessing circulating hormone levels in individuals with stimulant use disorders, and such assessments generally are not part of routine care.

DIFFERENTIAL DIAGNOSIS

There is substantial overlap between symptoms of stimulant intoxication or withdrawal and other psychiatric diagnoses. Symptoms suggestive of a manic episode or of a psychotic or anxiety disorder are common during stimulant intoxication. Likewise, stimulant withdrawal frequently involves symptoms of depression and anxiety. Symptoms that are more severe than usually observed suggest a stimulant-induced psychiatric disorder rather than simple intoxication or withdrawal. Typically, stimulant-induced

symptoms will attenuate or remit over several weeks with cessation of use. Symptoms that persist following a one-month period of abstinence or that precede initiation of stimulant use suggest an independent psychiatric disorder.

Cycles of binge use followed by variable periods of abstinence may lead to mood instability; periods of decreased need for sleep, agitation, and other symptoms suggestive of mania may alternate with periods of low energy, dysphoria, and hypersomnia. When a patient presents with such a symptom pattern, which resembles bipolar spectrum illnesses, a careful history and laboratory evaluation to exclude use of cocaine or other psychostimulants is necessary for accurate diagnosis.

To some extent, the quality of symptoms may provide clues regarding whether they are manifestations of stimulant toxicity or a co-occurring psychiatric disorder. For instance, psychotic symptoms caused by cocaine use generally involve nonbizarre delusions or misperceptions. Tactile hallucinations ("cocaine bugs") also suggest cocaine-related psychosis.

History and laboratory testing can also distinguish stimulant-related symptoms from those induced by other psychoactive substances. In particular, phencyclidine (PCP) and stimulant intoxication share common features.

COURSE AND PROGNOSIS

Surrogate markers of more severe cocaine use disorders are often associated with poorer prognosis. Among patient characteristics, a positive urine drug screen prior to initiation of treatment and the presence of withdrawal symptoms have frequently been associated with worse treatment outcomes. Heavy baseline use of cocaine may also predict worse response to treatment. Although alcohol use at baseline may not indicate worse prognosis, continued alcohol intake during treatment is associated with relapse to cocaine use during and after treatment.

Co-occurring mental illness may also present particular challenges to patients who struggle with stimulant use disorders. Among individuals with severe mental illness treated in a contingency management program, better prognosis was associated with lower levels of psychiatric severity, negative drug screens at onset of treatment, and use of treatment services prior to entering the study. Depression appears to denote a worse prognosis for cocaine use disorder, though some results have contradicted this conclusion. While PTSD may not predict more cocaine use in treatment, it is associated with more persistent impairment overall, and with worse outcomes on psychosocial measures related to drug use. Two personality characteristics, impulsivity and novelty-seeking, are associated with early treatment dropout, as is mild cognitive impairment.

Higher self-efficacy and coping, older age, greater commitment to abstinence, more social support, and more connection to self-help have all

been associated with better prognosis, as has less baseline use of cocaine. Longer retention in residential programs (greater than 90 days) is associated with improved outcome for cocaine use disorders. Impaired cognition, nicotine dependence, lower educational attainment but most significantly, baseline methamphetamine use, predict worse treatment outcome for methamphetamine use disorders. Early onset of methamphetamine use (itself associated with high levels of use) is also associated with a worse outcome. In a 10-year longitudinal study, male gender, white race, and being under legal supervision were associated with decreased odds of quitting while involvement in treatment or self-help groups were associated with increased odds. Injection as a route of administration is also associated with poorer prognosis among those with methamphetamine use disorders.

TREATMENT

Intensity of Treatment

Often, the treatment of alcohol, benzodiazepine, and opioid dependence may be initiated in hospital or detoxification units, where severe or complicated withdrawal can be closely monitored and managed. In contrast, the treatment of individuals with stimulant-related disorders is often begun in outpatient settings, in part because the withdrawal syndrome rarely requires specific medical intervention.

In addition to clinical considerations, treatment setting and intensity may be influenced by a host of other factors such as the availability of treatment resources, financial considerations, and patient preference. Factors that may prompt initiation of treatment in an inpatient or residential setting include medical or psychiatric conditions independently warranting intensive treatment, a lack of response at less intensive levels of care, or a perceived need to remove an individual from environmental influences that may sustain drug use.

The severity of co-occurring problems must also be taken into account when selecting a treatment setting. Patients with persistent psychotic or severe depressive symptoms involving suicidal ideation may require inpatient treatment to stabilize their co-occurring condition before outpatient care can be initiated. Inpatient treatment may also be appropriate for patients with multiple drug dependences, to provide safe detoxification from alcohol, sedatives, or opioids. Finally, some patients may require medical stabilization for complications of stimulant use such as systemic infection, cerebrovascular accidents, or myocardial infarction and arrhythmia.

Treatment of Intoxication and Withdrawal

There are no reliably effective specific pharmacologic treatments for

cocaine or amphetamine intoxication or withdrawal, and supportive care is the mainstay of treatment for these conditions. While self-limited, stimulant intoxication can be complicated by toxic reactions including paranoia, agitation, seizures, hypertension and its sequelae, and end-organ ischemia. Agitation or paranoia may call for intensive monitoring in a controlled setting and short-term treatment with benzodiazepines or antipsychotic medications. However, most patients will recover spontaneously in a matter of hours-to-days without the need for pharmacologic therapy. Unlike withdrawal syndromes described for alcohol, benzodiazepines, and barbiturates, stimulant withdrawal is generally milder, self-limited, and not typically associated with delirium or other medical complications.

Psychosocial Treatments

Psychosocial treatments remain the cornerstone of treatment for stimulant use disorders, particularly given the lack of a clearly effective medication to treat these disorders. Three behavioral models (contingency management, CBT, and twelve-step facilitation [TSF]) have received the strongest empirical grounding for the treatment of stimulant use disorders. Treatments based on 12-step approaches appear particularly useful for the treatment of cocaine use disorders.

Contingency Management. Contingency management is a treatment firmly grounded in basic behavioral principles. In this treatment approach, stimulant use disorders are conceptualized as the result of the drugs' powerful reinforcing effects, which supplant other natural reinforcers and aversive conditions that might otherwise halt drug use. In contingency management, competing reinforcers are created and applied to increase the likelihood of desired, nondrug behaviors. In most cases, the behaviors that are reinforced include treatment participation and the provision of drug-free urine samples during testing. Most commonly, contingency management procedures are combined with other active treatments such as comprehensive behavioral therapy programs and/or pharmacotherapy for co-occurring opioid use disorders.

An intermittent reinforcement schedule is commonly employed, in which an element of chance is introduced so that rewards are provided on a variable schedule. The most thoroughly studied such method for the treatment of stimulant use disorders is commonly referred to as the "fishbowl approach." Study participants who engage in targeted nondrug behaviors (such as providing drug-free urine samples) are provided chances to draw for prizes that vary in value. This strategy has the advantage of providing robust reinforcement of desired behaviors and of decreasing the overall cost of prizes (to an average cost of $200 or less per patient for 12 weeks of treatment). Beneficial results of such an approach

have been demonstrated for patients with co-occurring cocaine and opioid use disorders in methadone-maintenance programs as well as for patients with cocaine use disorders treated in community settings. Positive effects of intermittent, abstinence-based reinforcement include increased likelihood of drug abstinence and improved treatment retention. Contingency management has also shown efficacy in other populations with cocaine use disorders, including those with co-occurring ASPD, pregnant women, those exposed to tuberculosis, and those with high levels of psychiatric symptoms. Similarly, contingency management approaches have led to improved treatment outcomes in methamphetamine use disorders, including those with co-occurring psychiatric disorders. A trial of gay and bisexual men with methamphetamine use disorder demonstrated improved retention and abstinence rates with a contingency management intervention. Several other clinical trials have confirmed similar benefits of contingency management in the treatment of methamphetamine use disorders in other populations. Finally, there is evidence that contingency management techniques provide added benefit when applied to other active treatments such as CBT.

Twelve-Step-Oriented Treatment and Self-Help Groups. Treatment that is based on the principles of 12-step recovery is common in clinical practice, and has been studied in several well-designed trials. The 12-step model serves as the basis for self-help (often called mutual-help) groups such as AA, Narcotics Anonymous, Cocaine Anonymous, and Crystal Meth Anonymous. In addition to these peer-led groups, treatments based on 12-step principles are used by professionals in treating SUDs; the two most widely studied forms of professional treatment based on 12-step principles are Individual Drug Counseling (IDC) and TSF; both encourage patients to attend 12-step meetings and to accept the disease model of addiction that serves as a core principle of 12-step philosophy. Both treatments have strong empirical support for the treatment of cocaine use disorders, particularly for people without co-occurring psychiatric illness. Indeed, IDC (in conjunction with Group Drug Counseling, or GDC) outperformed cognitive therapy plus GDC, supportive-expressive psychodynamic therapy plus GDC, and GDC alone in the National Institute on Drug Abuse Collaborative Cocaine Treatment Study; this was the largest randomized controlled trial to compare different counseling approaches for cocaine-dependent patients.

The body of literature regarding the impact of self-help groups in those with cocaine-related disorders is much smaller than that for alcohol-related disorders. There is also very little empirical data regarding efficacy of 12-step and self-help approaches for amphetamine-type stimulant use disorders.

Cognitive-Behavioral Therapy. CBT for SUDs is predicated on the theory that maladaptive thought and behavior patterns sustain drug use in substance-dependent individuals. CBT theoretically interrupts such patterns by helping patients to identify thoughts, feelings, and events that lead to drug use. CBT techniques for treating drug use disorders generally include providing psycho-education and practice on skills such as avoiding or coping with high-risk situations, monitoring and challenging drug-related thoughts and beliefs, and scheduling non–drug-related activities. Often, CBT treatments include written homework and other practice assignments between scheduled sessions. CBT has a strong evidence base for the treatment of SUDs. Several studies have demonstrated the efficacy of CBT for treating cocaine use disorders, with some evidence that it is most useful for more seriously ill patients, including those with co-occurring alcohol dependence or depression. The Matrix model is a manualized treatment employing elements of several evidence-based approaches, including CBT, family therapy, motivational interviewing, and other behavioral techniques. It has demonstrated efficacy in improving treatment participation and retention as well as in increasing the proportion of methamphetamine-free urine samples provided by patients with methamphetamine use disorders.

Pharmacological Agents

Medications for Cocaine Use Disorders. To date, there are no FDA-approved medications to treat cocaine use disorders. Many agents that have shown promise in animal studies, open-label trials, or small double-blind studies have then had mixed or disappointing results in larger clinical trials. Among other approaches, knowledge of cocaine's effects on brain dopamine has led to the investigation of many medications that modulate dopaminergic neurotransmission, including direct and indirect agonists as well as dopamine antagonists. Unlike the success of agonist therapy for opioid dependence, the use of dopamine agonists has not shown similar efficacy in the treatment of cocaine dependence.

Trials of direct dopamine agonists (e.g., amantadine, bromocriptine [Parlodel], L-dopa/carbidopa [Sinemet], pramipexole [Mirapex], pergolide [Permax]) have largely been negative. Methylphenidate (Ritalin), which increases presynaptic dopamine release, has also not shown benefit for cocaine-dependent patients in the absence of co-occurring attention deficit hyperactivity disorder (ADHD). Dextroamphetamine (Dexedrine) has had mixed reports of therapeutic response, but has not demonstrated sufficiently robust evidence of efficacy to justify its widespread use as a treatment for cocaine use disorders. The atypical psychostimulant modafinil (Provigil) has shown little benefit. A sustained-release formulation of methylphenidate did show benefit in one study. Concerns

regarding the potential abuse and diversion of amphetamine-type stimulants have further dampened enthusiasm for their use in the pharmacotherapy of cocaine use disorders.

Analogous to the use of the opioid antagonist naltrexone (ReVia) in the treatment of opioid dependence, dopamine antagonists have been proposed as potential medications for cocaine dependence. Typical and atypical antipsychotics and the D1- receptor antagonist ecopipam have all been investigated as possible treatments for cocaine dependence. As yet, none have shown clear evidence of benefit in placebo-controlled trials.

In addition to its effects on dopaminergic brain pathways, cocaine is a powerful inhibitor of presynaptic serotonin reuptake. Serotonergic agents, including antidepressants, appear to blunt some of cocaine's reinforcing properties in animal models. However, clinical trials of predominantly serotonergic agents (serotonin-specific reuptake inhibitors, buspirone [BuSpar], mirtazapine [Remeron], venlafaxine [Effexor]) have been disappointing. Among the antidepressants, there is some evidence that the tricyclic medication desipramine (Norpramin) may be helpful for treating co-occurring depression in patients with cocaine use disorders.

Other negative cocaine trials to date include D-cycloserine (Seromycin), citicoline, aripiprazole (Abilify), fluoxetine (Prozac), acamprosate (Campral), naltrexone (ReVia) for co-occurring cocaine and AUDs, memantine (Namenda), atomoxetine (Strattera), baclofen, olanzapine (Zyprexa), bupropion (Wellbutrin), risperidone (Risperdal), tiagabine (Gabitril), vigabatrin (Sabril), reserpine, and selegeline (Eldepryl). More encouraging but preliminary data exist on sertraline (Zoloft) as a means of delaying relapse, N-acetylcysteine (NAC) in less severely ill individuals, doxazosin (Cardura), varenicline (Chantix), citalopram (Celexa), and galantamine (Razadyne). There is as yet insufficient evidence to recommend any of the above as effective clinical interventions for cocaine use disorders.

Disulfiram (Antabuse) has demonstrated potential benefit in the pharmacotherapy of cocaine dependence. Disulfiram has long been used in the treatment of alcohol dependence because it inhibits ALDH, an enzyme involved in the metabolism of ethanol. ALDH converts the toxic compound acetaldehyde to its less toxic metabolite, acetic acid. If ethanol is ingested by a patient who has taken a therapeutic dose of disulfiram, the resulting build-up of acetaldehyde precipitates a noxious reaction, including nausea, headache, and flushing.

Initial interest in disulfiram for the treatment of cocaine dependence was stimulated by the frequent co-occurrence of AUDs with cocaine use disorders. Early studies suggested that disulfiram had an indirect effect on cocaine use, mediated by its effects on alcohol use. However, disulfiram also appears to have a more direct pharmacologic effect on the metabolism of dopamine and cocaine itself; it acts as an inhibitor of dopamine beta-

hydroxylase, thereby slowing the breakdown of synaptic dopamine and increasing dopamine levels. Disulfiram has also been shown to increase serum cocaine levels several-fold in pre-treated persons. Subjects who are co-administered cocaine and disulfiram describe more negative responses to cocaine, including anxiety, restlessness, and paranoia. These reactions may in part explain the effects of disulfiram in decreasing cocaine use.

Following initial studies of patients with co-occurring alcohol and cocaine use disorders, more recent clinical trials have shown improved cocaine-related outcomes in groups of patients without co-occurring AUDs. In fact, one randomized, double-blind, placebo-controlled trial of disulfiram for cocaine dependence found that patients without co-occurring alcohol dependence were more likely to benefit from disulfiram than were those with co-occurring alcohol dependence. Hence, alcohol-mediated pathways cannot fully account for disulfiram's effects on cocaine use.

Traditional clinical trials attempt to determine whether a treatment is efficacious in a given population, which is assumed to be homogeneous. However, it has long been apparent that individuals may have very different responses to a given treatment. In this manner, sub-groups of patients with cocaine use disorders appear to respond differently to treatment with disulfiram. For instance, patients with co-occurring AUD and cocaine use disorder did not respond as well to disulfiram as did those with only cocaine use disorder. Likewise, women as a group may not respond as well to disulfiram as do men. Other issues affecting treatment response to disulfiram include challenges with treatment nonadherence, and the issue of achieving an adequate dose-to-mass ratio in larger individuals.

Pharmacogenetics is the study of genetic factors that affect medication response. Significant work has been accomplished to characterize the effects of genetic differences on responsiveness to disulfiram for treatment of cocaine use disorders. Genes implicated in differential treatment response to disulfiram include the alpha-1-adrenoreceptor gene (ADRA1A), ankyrin repeat and kinase domain-containing 1 (ANKK1), the D2 dopamine receptor (D2DR), the methylation enzyme encoded by MTHFR/C677T, variants of the serotonin transporter (5-HTTLPR), dopamine-beta-hydroxylase (DβH), and enzymes involved in serotonin synthesis (TPH2). Variation of each of these genes appears able to distinguish those who will and those who will not respond to disulfiram for treatment of cocaine use disorders.

Other recent studies have focused on medications with pharmacologic actions on the inhibitory neurotransmitter GABA. Baclofen (Lioresal), tiagabine (Gabitril), and modafinil (Provigil), which likely affects both GABA and glutamate neurotransmission, have failed to show benefit in clinical trials. Topiramate (Topamax) at a 200-mg-daily dose was

associated with increased rates of abstinence in a 13-week clinical trial. In a subsequent trial, 300-mg-daily topiramate in conjunction with CBT was associated with increases in number of abstinent days and the likelihood of cocaine-free weeks. In yet another study, the combination of mixed amphetamine salts and topiramate was associated with twice the rate of continuous abstinence over a 3-week period, in comparison to placebo. A similar finding was seen in a study of topiramate for individuals with co-occurring cocaine and alcohol dependence. Despite these encouraging findings, two negative trials of the use of topiramate to treat cocaine use disorders have been published: one with methadone-maintained patients and one with individuals using inhaled "crack" cocaine.

Medications to Treat Amphetamine-Type Stimulant Use Disorders. Serotonergic medications including imipramine (Tofranil), fluoxetine, and sertraline have yielded disappointing results when evaluated for the treatment of amphetamine use disorders. The 5-HT3 antagonist ondansetron (Zofran) has not shown any effect on methamphetamine-related outcomes. Mirtazapine, an antidepressant with novel mechanisms including serotonergic effects, reduced both methamphetamine-positive urine samples and risky sexual behaviors in methamphetamine-use-disordered men who have sex with men. The opioid-receptor antagonist naltrexone and the dopaminergic antidepressant bupropion have both shown promise in those with less-severe amphetamine use disorders (based on baseline use), but little effect on more dependent individuals. In similar fashion, topiramate did not increase abstinence rates overall, but was associated with decreased methamphetamine use in those who were abstinent at the start of the trial. The partial dopamine agonist aripiprazole is not helpful in the treatment of methamphetamine use disorders, and may cause escalations of drug use. Regarding medications affecting GABA neurotransmission, neither baclofen nor gabapentin (Neurontin) have proven useful. Likewise, modafinil has not shown benefit. Agonist therapy has shown the most promise, although there is still inadequate evidence to support routine clinical use.

Immunotherapy. One novel approach to treating stimulant use disorders is through immunization, using a therapeutic vaccine to induce antibodies to psychostimulants. Such a vaccine has been developed for cocaine; in a small, open-labeled study, subjects treated with the vaccine reported decreased subjective euphoric response to cocaine self-administration. It has generally proven safe in clinical trials. Two clinical trials have been completed to assess efficacy of the cocaine vaccine. One trial demonstrated a significant increase in cocaine-free urine when adequate antibody titers were achieved. A following clinical trial showed improved treatment retention with higher antibody titers, but no

significant differences in other clinical outcomes were found. Two genetic variations, low-level expression of dopamine-beta-hydroxylase and a polymorphism of the OPRK1 kappa opioid receptor, were associated with better response. A vaccine against methamphetamine remains in the pre-clinical stages. Creating an effective vaccine against either methamphetamine or cocaine is a challenge, given the small size of the compounds (too small to create an immune response independent of immunogenic carriers). It is hoped that more potent vehicles for delivering vaccines will yield better results in the future.

References

*Brensilver M, Heinzerling KG, Shoptaw S. Pharmacotherapy of amphetamine-type stimulant dependence: an update. *Drug Alcohol Rev*. 2013;32(5):49–460.

Crunelle CL, Veltman DJ, Booij J, van Emmerik-van Oortmersen K, van den Brink W. Substrates of neuropsychological functioning in stimulant dependence: a review of functional neuroimaging research. *Brain Behavior*. 2012;2(4):499–523.

Drug and Alcohol Services Information System (DASIS), Results from the 2012 Treatment Episode Data Set (TEDS). *National Findings*. *Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality*. Available at http://wwwdasis.samhsa.gov/webt/tedsweb/tab_year.choose_year_web_table?t_state=US. Accessed February 21, 2015.

Ersche KD, Turton AJ, Chamberlain SR, Muller U, Bullmore ET, Robbins TW. Cognitive dysfunction and anxious-impulsive personality traits are endophenotypes for drug dependence. *Am J Psychiatry*. 2012;169(9):926–936.

Fowler JS, Volkow ND, Kassed CA, Chang L. Imaging the addicted human brain. *Sci Pract Perspect*. 2007;3(2):4–16.

Fowler JS, Volkow ND, Logan J, et al. Fast uptake and long-lasting binding of methamphetamine in the human brain: comparison with cocaine. *Neuroimage*. 2008;43(4):756–763.

Gelernter J, Sherva R, Koesterer R, et al. Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Mol Psychiatry*. 2014;19(6):717–723.

Hser YI, Evans E, Huang D, Brecht ML, Li L. Comparing the dynamic course of heroin, cocaine, and methamphetamine use over 10 years. *Addict Behav*. 2008;33(12):1581.

Kampman KM, Pettinati HM, Lymch KG, Spratt K, Wierzbicki MR, O'Brien CP. A double-blind, placebo-controlled trial of topiramate for the treatment of comorbid cocaine and alcohol dependence. *Drug Alcohol Depend*. 2013;133(1):94–99.

Kendler KS, Karkowski LM, Neale MC, Prescott CA. Illicit psychoactive substance use, heavy use, abuse, and dependence in a US population-based sample of male twins. *Arch Gen Psychiatry*. 2000;57:261–269.

*Kim JH, Lawrence AJ. Drugs currently in Phase II clinical trials for cocaine addiction. *Expert Opin Investig Drugs*. 2014;23(8):1105–1122.

*Koob GF, Volkow ND. Neurobiology of addiction: a neurocircuitry analysis. *Lancet Psychiatry*. 2016;3(8):760–773.

Kosten TR, Domingo CB, Horter D, et al. Vaccine for cocaine dependence: a randomized double-

blind placebo-controlled efficacy trial. *Drug Alcohol Depend*. 2014;140:42–47.

McKay JR, Van Horn D, Rennert L, Drapkin M, Ivey M, Koppenhaver J. Factors in sustained recovery from cocaine dependence. *J Subst Abuse Treat*. 2013;45(2):163–172.

NSDUH. *Results from the 2013 National Survey on Drug Use and Health: National Findings. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies*. Available at http://www.samhsa.gov/data/population-data-nsduh/reports?tab=32. Accessed February 21, 2015.

*Paillet-Loilier M, Cesbron A, Le Boisselier R, Bourgine J, Debruye D. Emerging drugs of abuse: current perspectives on substituted cathinones. *Subst Abuse Rehabil*. 2014;5:37–52.

Pierce CR, Wolf ME. Psychostimulant-induced neuroadaptation in nucleus accumbens AMPA receptor transmission. *Cold Spring Harbor Perspect Med*. 2013;3(2):a012021.

Rawson RA. Current research on the epidemiology, medical and psychiatric effects, and treatment of methamphetamine use. *J Food Drug Anal*. 2013;21(4):S77–S81.

Roncero C, Daigre C, Grau-Lopez L, et al. An international perspective and review of cocaine-induced psychosis: a call to action. *Subst Abus*. 2014;35(3):321–327.

Salo R, Flower K, Kielstein A, Leamon MH, Nordahl TE, Galloway GP. Psychiatric comorbidity in methamphetamine dependence. *Psychiatry Res*. 2011;186(2–3):356–361.

SAMHSA, Substance Abuse and Mental Health Services Administration. The DAWN Report: Highlights of the 2011 Drug Abuse Warning Network (DAWN) Findings on Drug-Related Emergency Department Visits. Available at http://archive.samhsa.gov/data/2k13/DAWN127/sr127-DAWN-highlights.htm. Accessed February 21, 2015.

Shorter D, Nielson DA, Huang W, Harding MJ, Hamon SC, Kosten TR. Pharmacogenetic randomized trial for cocaine abuse: disulfiram and alpha-1-a-adrenoceptor gene variation. *Eur Neuropsychopharmacol*. 2013;23(11):10.

Stitzer ML, Petry NM, Peirce J. Motivational incentives research in the National Drug Abuse Treatment Clinical Trials Network. *J Subst Abuse Treat*. 2010;38(Suppl 1):S61–S69.

UNODC, United Nations Office of Drugs and Crime, World Drug Report. 2013. Available at http://www.unodc.org/wdr2013. Accessed February 21, 2015.

Vocci FJ, Montoya ID. Psychological treatments for stimulant misuse, comparing and contrasting those for amphetamine dependence and those for cocaine dependence. *Curr Opin Psychiatry*. 2009;22(3):263–268.

Wang GJ, Smith L, Volkow ND, et al. Decreased dopamine activity predicts relapse in methamphetamine abusers. *Mol Psychiatry*. 2012;17(9):918–925.

Weiss RD, Mirin SM, Bartel RL. *Cocaine*. 2nd ed. Washington, DC: American Psychiatric Press; 1994.

World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva: World Health Organization; 1992.

▲ 11.4 Caffeine-Related Disorders

Caffeine is the most widely used mood-altering drug in the world. Caffeine (1,3,7-trimethylxanthine) is found in more than 60 species of plants and belongs to the methylxanthine class of alkaloids, which also includes theobromine (found in chocolate) and theophylline (often used in the treatment of asthma). Eighty-five percent of the United States population consumes foods and beverages containing caffeine. Caffeine affects various neurobiological and physiological systems and produces significant psychological effects. Although caffeine is not associated with any life-threatening illnesses, psychiatric symptoms and disorders can result from its use. The ubiquitous habitual use of caffeine and its widely accepted integration into daily customs can lead to an underestimation of the role that caffeine may play in caffeine user's daily subjective experiences and can make the recognition of caffeine-associated disorders particularly challenging. It is important for clinicians to be familiar with the sources of caffeine and psychiatric symptoms and disorders that can result from its use.

DEFINITION

Caffeine use can be associated with five discrete syndromes (Table 11-4–1). Four of these conditions—caffeine intoxication, caffeine-induced anxiety disorder, caffeine-induced sleep disorder, and caffeine withdrawal—are diagnoses in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). A fifth syndrome—caffeine use disorder—is a research diagnosis in the appendix of DSM-5, but a very similarly defined disorder is in the tenth edition of the *International Classification of Diseases* (ICD-10). Caffeine use disorder can be diagnosed using the category unspecified caffeine-related disorder, which is included in DSM-5 as a substance-related disorder.

Table 11.4–1.
DSM-5 Caffeine-Related Disorders

Caffeine intoxication
Caffeine-induced anxiety disorder
Caffeine-induced sleep disorder
Caffeine withdrawal
Caffeine use disorder (diagnosis for further study)
Unspecified caffeine-related disorder

HISTORY

Caffeine-containing foods and beverages have been consumed for thousands of years. Tea has been cultivated and consumed in China since at

least 350 AD, and coffee cultivation spread from Ethiopia to Arabia in the 15th century and subsequently became a popular beverage in Arabic cultures during the 16th century. Initially, most caffeine consumption occurred in restricted geographical regions where caffeine-containing plants were indigenous. It was not until the 17th century that caffeine use began to spread, eventually achieving its current, virtually universal availability. The Dutch imported caffeine in the form of coffee from Arabic countries to Europe, where it first became popular around the middle to end of the 17th century (along with tea, tobacco, and chocolate). Medical literature during this time period described coffee as a useful beverage for a wide variety of conditions, although it is notable that coffee use to induce sobriety was a particularly attractive feature. However, considerable controversy about whether the use of coffee was beneficial or detrimental to health accompanied its expansion into Western cultures, a controversy that is mirrored in contemporary concerns over the use of caffeine.

The widespread use of coffee during the 17th and 18th centuries is reflected in the estimate that there were between 2,000 and 3,000 coffeehouses in London by the early part of the 17th century (about one coffeehouse for every 200 to 300 people). The popularity of coffeehouses in Europe also highlights how coffee was initially a beverage consumed in public; it was only later that coffee became a beverage consumed in the home. Perhaps the most famous of the coffeehouses in London was operated by Edward Lloyd. Like most coffeehouses, Lloyd's was a center of business, especially for insurance agents, and by the end of the 18th century Lloyd's became the well-known Lloyd's of London insurance company. However, in England coffee use was eventually supplanted by tea consumption, with the shift in beverage preference occurring in the first half of the 18th century.

In the United States a shift from tea to coffee use occurred in 1773 when colonists protesting British taxes threw cargoes of tea overboard in the Boston harbor—the Boston Tea Party. The repercussions of this event continue to the present day; the United States is now the major consumer of coffee in the world.

At the present time caffeine is integral to the economic activity of several countries. For example, only crude oil produces more foreign exchange earnings than coffee for developing countries. In the United States coffee is the major agricultural import and in value is second only to oil among all imports. This economic and trade activity underscores the extent to which caffeine use has spread over the past four centuries, so that caffeine is now available and accepted virtually everywhere in the world.

In the late 19th century caffeine began appearing in various soft drink beverages. While a fraction of caffeine contained in cola soft drink beverages is derived from the kola nut, most caffeine in soft drinks is added —reportedly as a flavoring agent; however, research has found little

evidence that caffeine functions as a flavoring agent at levels found in soft drinks. In recent years, there has been a tremendous increase in the variety of beverages and foods available to consumers with added caffeine. Most notable are energy drinks and energy shots, which now represent a multibillion-dollar market in the United States. These products are banned or regulated in some countries due to concerns about the high caffeine levels. Caffeine is also presently added to a wide variety of food products, dietary supplements, and medications.

COMPARATIVE NOSOLOGY

The first edition of DSM (DSM-I), published in 1952, made no mention of caffeine-related disorders. In the DSM-II (1968) it is explicitly stated that caffeine (and notably, tobacco) should be excluded from consideration for a diagnosis of drug dependence. DSM-III (1980) was the first to include caffeine (and tobacco) use disorders. At that time only caffeine intoxication was included as a discrete syndrome with specific diagnostic criteria. Caffeine intoxication has remained in all subsequent editions of the DSM, with no substantive changes in the diagnostic criteria. Caffeine-induced anxiety disorder and caffeine-induced sleep disorder first appeared in the DSM-IV (1994). Caffeine withdrawal was included in the DSM-IV and DSM-IV-TR as a research diagnosis and was included as a formal diagnosis in the DSM-5 with some modifications to the criteria based on a large body of empirical data. Caffeine use disorder is a diagnosis for further study in the DSM-5 as there is growing evidence that some individuals do indeed have a profile of symptoms consistent with problematic caffeine use. Furthermore, a population of individuals seeking treatment for their use of caffeine has been identified.

The ICD-10, published by the World Health Organization, contains criteria for caffeine intoxication ("Acute intoxication due to the use of other stimulants, including caffeine") and also includes several other diagnostic categories that can be applied to caffeine such as harmful use, dependence syndrome (similar to caffeine use disorder), and withdrawal states from other stimulants, including caffeine.

EPIDEMIOLOGY

Caffeine is contained in beverages (e.g., coffee, tea, soft drinks, energy drinks), foods (e.g., chocolate, coffee yogurt), and medications (both prescription and over-the-counter drugs), although most caffeine consumed in the United States is from coffee, soft drinks, and tea. Approximately 70 percent of all soft drinks consumed in the United States contain caffeine. Some sources of caffeine such as coffees and colas are readily identifiable, while others are less easily recognized. For example, many types of noncola soft drinks contain caffeine, and some over-the-counter analgesics also contain caffeine. In recent years, caffeine-

containing energy drinks and energy shots have been made widely available to consumers. This represents a potentially large source of caffeine, as energy drinks and, in particular, energy shots often contain significantly more caffeine than soft drinks. Furthermore, certain manufactures have begun to add caffeine to foods that have traditionally not contained caffeine (e.g., candy, potato chips).

Determining exactly how much caffeine an individual consumes can be challenging because there are various potential sources of caffeine. Also, there can be large variability within categories. For example, 6 oz of weak instant coffee may contain as little as 20 mg of caffeine, while the same serving size of strong drip coffee may contain as much as 150 mg. Likewise, different caffeinated soft drinks may range in caffeine content from 22 to 69 mg for 12 oz. Energy drinks can vary from 50 to 375 mg for a can or bottle. Energy shots, which are growing in popularity (i.e., currently representing 11 percent of the energy market) contain high concentrations of caffeine with some containing an alarming 500 mg per 2-oz dose. Furthermore, a number of over-the-counter and prescription medications have formulations both with and without caffeine. Table 11.4–2 shows the caffeine content of selected caffeine-containing products.

Surveys of the general population over the past two decades consistently find a high rate of regular caffeine use among children and adults in the US. A report was commissioned by the U.S. Food and Drug Administration (FDA) that summarized the US population's consumption of caffeine between 2003 and 2008. The report was based on the National Health and Nutrition Examination Survey and NPD Group's Food Consumption surveys, which collected self-reported food consumption using daily diaries for 14 consecutive days from 2,000 households representative of the population. It was concluded that more than 95 percent of caffeine consumed among children and adults comes from beverage intake. The mean caffeine intake among adults (over age 22) was estimated to be 300 mg. Among coffee drinkers, the mean daily caffeine intake was estimated to be 375 mg. Another recent study estimated the caffeine intakes of the US population using the Kantar Worldpanel Beverage Consumption Panel, which consists of a nationally representative sample of 42,851 individuals (ages 2 years and older) who completed 7-day diaries of beverage consumption. They estimated that 85 percent of the US population consumes at least one caffeinated beverage per day, with daily rates of 43 percent among 2-to-5-year-olds, and these daily rates progressively increasing with age to more than 99 percent among those 65 years and older (see Fig. 11.4–1). Total caffeine exposure for all ages combined was estimated to be 165 mg with the greatest intake among consumers ages 35 years and older who consume an estimated 200 to 225 mg caffeine per day. The percentage of energy drink consumers across all age groups was <10%. As stated in these reports, determining actual exposure to caffeine remains

a challenge due to difficulties in obtaining current and accurate data on the actual caffeine content of beverages.

Coffee is the primary source of caffeine among adults in the United States, whereas soft drinks are the primary source of caffeine among children and adolescents. More than half of adults consume coffee every day, with drinkers consuming an average of 3.3 cups per day. About 50 percent of caffeine users consume multiple types of caffeine containing beverages (e.g., coffee drinkers who also consume soft drinks). Figure 11.4–2 shows trends over the past 40 years in consumption of the major caffeine-containing beverages in the United States. Energy drinks have also recently become a large potential source of caffeine. Since the introduction of energy drinks into the US market in 1997, the energy drink market has grown substantially. The FDA report cited above concluded that reliable data on caffeine exposure among regular users of energy drinks is not yet available. However, sales figures suggest that use of energy drinks is widespread and on the rise. In 2008, 366 million gallons of energy drinks were sold in the US, equivalent to nearly 3 million servings. Individuals between the ages of 13 and 35 years old comprise 65 percent of the energy drink market. The use of energy drinks could have important implications for the incidence of caffeine-related disorders. Whereas the caffeine content of soft drinks is limited by the FDA, no such restrictions apply to energy drinks, with some energy drinks containing more than ten times the amount of caffeine as the most popular soft drinks. At present it is common for consumers to mix energy drinks with alcohol. This is also potentially problematic, as research has shown that individuals consuming alcohol and caffeine underestimate their levels of alcohol intoxication and impairment and may be more prone to injury.

Table 11.4–2.
Caffeine Content of Common Foods and Medications

Product	Serving Size (Volume or Weight)	Typical Caffeine Content (mg)	Range (mg)
BEVERAGES			
Coffee			
Brewed/Drip	12 oz	200	107–420
Instant	12 oz	140	40–260
Espresso	1 oz	70	60–95
Decaffeinated	12 oz	8	0–20
Starbucks drip	12 oz	260	
Starbucks cappuccino	12 oz	75	
Starbucks espresso	1 oz	75	
Starbucks bottled frappuccino	9.5 oz	90	
Starbucks decaffeinated	12 oz	20	
Tea			

Brewed	6 oz	40	30–90
Instant	6 oz	30	10–35
Canned or Bottled	12 oz	20	8–32
Soft drinks			
Typical caffeinated soft drink	12 oz	40	22–69
Mountain Dew/Diet Mt. Dew	12 oz	54	
Diet Coke	12 oz	47	
Sunkist/Diet Sunkist	12 oz	41	
Dr. Pepper/Diet Dr. Pepper	12 oz	41	
Pepsi	12 oz	38	
Diet Pepsi	12 oz	36	
Pepsi Max	12 oz	69	
Coke Classic	12 oz	35	
A & W Cream Soda	12 oz	29	
Barq's Root Beer	12 oz	23	
A & W Diet Cream Soda	12 oz	22	
A & W Root Beer	12 oz	0	
7UP/Diet 7UP	12 oz	0	
Fanta Orange	12 oz	0	
Sprite/Diet Sprite	12 oz	0	
Canada Dry Ginger Ale	12 oz	0	
Cocoa/Hot Chocolate	6 oz	7	2–10
Chocolate Milk	6 oz	4	2–7
Caffeinated water			
Typical amount	16.9 oz	60	60–200
Water Joe	16.9 oz	60	
Buzzwater	16.9 oz	100 or 200	
Energy drinks			
Typical amount	varies	varies	50–500
AMP	16 oz	142	
Red Bull	8.4 oz	80	
Full Throttle	16 oz	160	
Monster	16 oz	160	
Rockstar	16 oz	160	
Spike Shooter	8.4 oz	300	
VPX Redline Energy	8 oz	316	
Wired-X-344	16 oz	344	
Rage inferno	24 oz	375	
Energy Shots			
Typical	varies	varies	
5-hour energy	2 oz	200	
5-hour energy extra strength	2 oz	230	
10-hour energy shot	1.93 oz	422	
FOODS			
Chocolate			
Hershey's Chocolate Bar	1.55 oz	9	
Hershey's Special Dark	1.45 oz	18	
Kit Kat Wafer Bar	1.5 oz	6	
Reese's Peanut Butter Cups	1.6 oz	4	

Miscellaneous Foods			
Penguin Peppermints	1 mint	7	
Starbucks Classic Coffee Ice Cream	4 oz	30	
Dannon Coffee Yogurt	6 oz	30	
Jolt Caffeinated Gum	1 stick	33	
Powerbar Tangerine Powergel	41 g	50	
Stay-Alert Caffeinated Gum	1 stick	100	
PRESCRIPTION MEDICATIONS			
Headache/Migraine/Pain			
Fiorinal	2 capsules	80	
Fioricet/Esgic/Many others	2 tablets	80	
Cafergot	2 tablets	200	
Norgesic	2 tablets	60	
OVER-THE-COUNTER MEDICATIONS			
Stimulants			
Typical	1 tablet	100 or 200	100–200
Vivarin	1 tablet	200	
No-Doz/No-Doz Maximum Strength	1 tablet	100 or 200	
Ultra Pep-Back	1 tablet	200	
Analgesics			
Goody's Headache Powder	1 powder packet	32.5	
BC Fast Pain Relief	1 powder packet	33.3	
BC Arthritis Pain and Influenza	1 powder packet	38	
Anacin Advanced Headache	2 tablets	130	
Excedrin Extra Strength	2 tablets	130	
Menstrual Pain Relief/Diuretics			
Diurex Water Pills	2 tablets	100	
Midol Menstrual Complete	2 caplets	120	
Pamprin Max	2 caplets	130	
DIETARY SUPPLEMENTS/WEIGHT LOSS PRODUCTS			
Typical	1 or 2 tablets	varies	50–300
Dexatrim Max	1 caplet	50	
Metabolife Weight Management	2 tablets	101	
Metabolife Ultra	2 caplets	150	
Hydroxycut Weight Loss Formula	2 caplets	200	
Leptopril	2 capsules	220	
Stacker 2	1 capsule	253	
Stacker 3	1 capsule	254	
Twinlab Ripped Fuel	2 capsules	220	
Swarm Extreme Energizer	1 capsule	300	
Xenadrine EFX	2 capsules	200	

Data from Juliano LM, Ferre S, Griffiths RR. The pharmacology of caffeine. In: Ries RK, Fiellin DA, Miller SC, Saitz R, eds. *ASAM Principles of Addiction Medicine*, 5th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2014:180–200; McCusker RR, Fuehrlein B, Goldberger BA, Gold MS, Cone EJ. Caffeine content of decaffeinated coffee. *J Anal Toxicol*. 2006;30(8):611; McCusker RR, Goldberger BA, Cone EJ. Caffeine content of specialty coffees. *J Anal Toxicol*. 2003;27(7):520. Caffeine values for all brand name

products were obtained directly from product labels, the manufacturer's website, or customer service department.

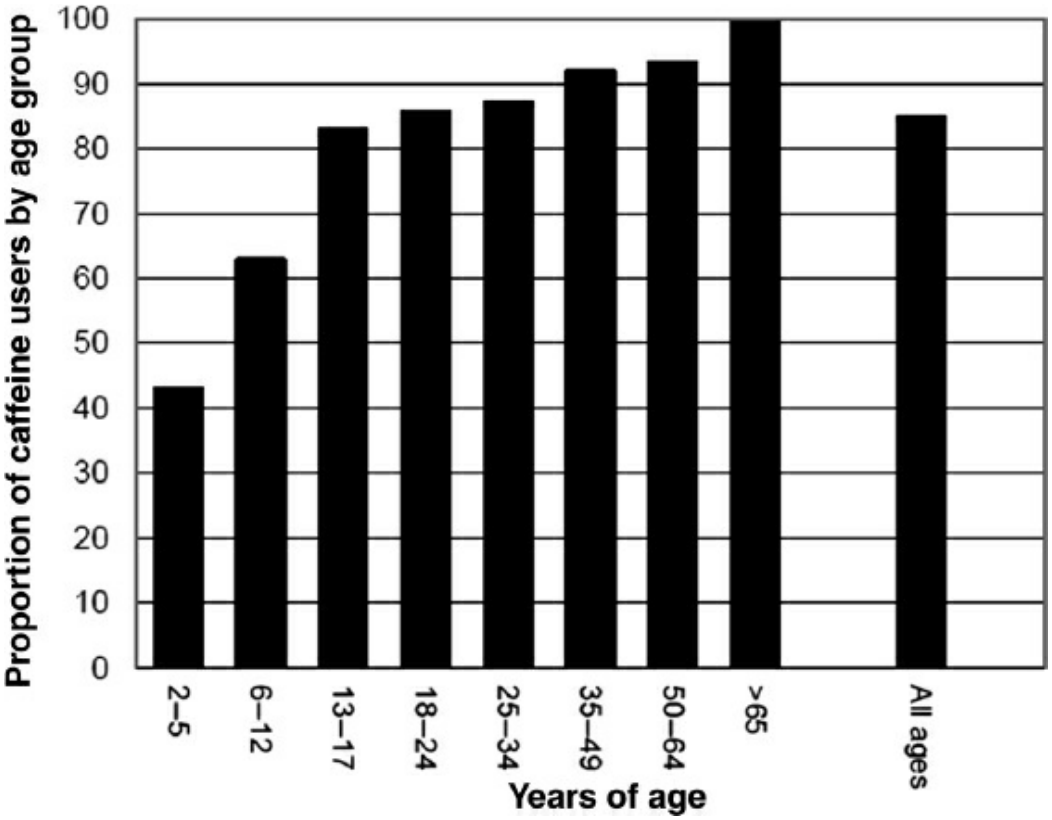

FIGURE 11.4–1. Percentage of caffeine consumers (past 7 days) by age group in the United States. (From Mitchell DC, Knight CA, Hockenberry J, Teplansky R, Hartman TJ. Beverage caffeine intakes in the US. *Food and Chemical Toxicology*. 2014;63:136–142.)

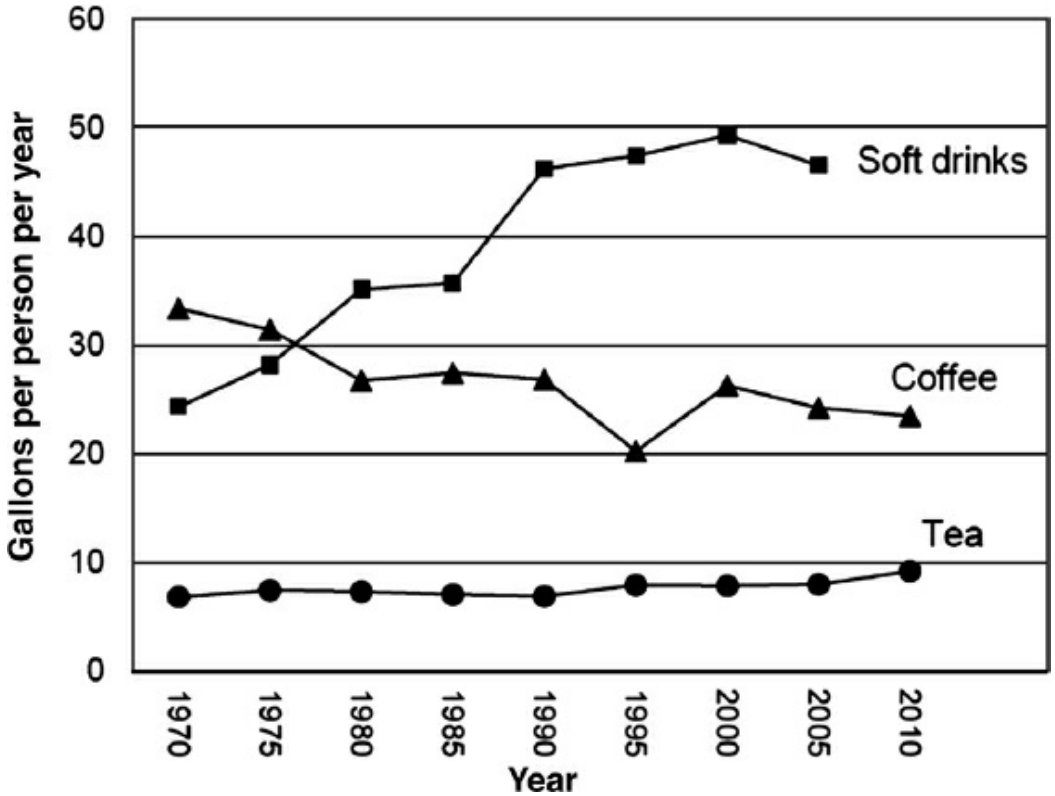

FIGURE 11.4–2. Annual per-person consumption of caffeine from coffee, tea, and soft drinks in the United States from 1970 to 2010. Soft drink data is only available until 2003. (From US Department of Agriculture.)

ETIOLOGY

In the United States, it is not uncommon for the first exposure to caffeine, usually as coffee or soft drinks, to occur during childhood. Onset of regular use of soft drinks occurs at an earlier age than regular use of coffee or tea. Continued caffeine consumption may be influenced by several different factors such as the pharmacological effects of caffeine, caffeine's reinforcing effects, genetic predispositions, caffeine withdrawal, characteristics of the consumer (e.g., age), and social norms. About 14 percent of former consumers of caffeine report stopping caffeine completely, in response to health concerns or unpleasant side effects.

Pharmacology and Effects in the Central Nervous System

Caffeine is well absorbed from the GI tract, with peak plasma concentrations typically occurring within 45 to 60 minutes after ingestion. Caffeine is readily distributed throughout all body fluids and is metabolized by the liver. Although there are several metabolic pathways for caffeine, it is primarily metabolized by the P450 1A2 system, and is frequently used as a metabolic probe to assess this system's activity. The half-life of caffeine is 4 to 6 hours on average, with large variability across individuals. The rate of caffeine elimination is increased by smoking or taking cimetidine (Tagamet) or fluvoxamine. Caffeine inhibits the metabolism of the

antipsychotic clozapine (Clozaril) and the bronchodilator theophylline (Bronkodyl) to an extent that might be clinically significant. Caffeine metabolism is slowed among women using oral contraceptives and in the second and third trimesters of pregnancy.

Caffeine exerts effects throughout the body, including bronchodilation (hence the therapeutic application of caffeine and the related methylxanthine theophylline in the treatment of asthma); modest increases in blood pressure (which are attenuated in caffeine-tolerant individuals); increased production of urine; increases in gastric acid secretion; and increases in plasma epinephrine, norepinephrine, renin, and free fatty acids. Centrally, caffeine affects turnover or levels of various neurotransmitters and also functions as a central nervous system (CNS) stimulant.

Numerous investigations have examined the effects of caffeine in the CNS in order to understand the neurobiological basis for caffeine's effects. The principal cellular site of action of caffeine is the adenosine receptor, where caffeine functions as an antagonist at both A_{2A} and A_1 receptors. Adenosine serves an important modulatory role in many brain processes and functions (e.g., reduces spontaneous electric activity in the brain, dilates blood vessels in the brain). As an adenosine antagonist, many of caffeine's effects are opposite to those produced by adenosine (e.g., CNS stimulation, constricts blood vessels in the brain). Figure 11.4–3 shows the similarities in the molecular structures of caffeine and adenosine. There is evidence that adenosine receptors form functional relationships, or heteromerize, with other types of receptors, including dopamine and glutamate receptors. For example, some of the behavioral stimulant properties of caffeine are believed to be due to an increase in dopamine activity as a result of antagonism of adenosine at the A_{2A}–D_2 receptor heteromer. It has also been reported that caffeine increases dopamine release in the shell of the nucleus accumbens, an effect believed to be a neuropharmacologic mechanism underlying the addictive potential of drugs. Other mechanisms proposed as mediating caffeine's effects have included the inhibition of phosphodiesterases and effects on intracellular calcium, although these effects may only occur at doses of caffeine that exceed typical human consumption. However, it is not known if these other mechanisms may mediate some of the clinical effects produced when caffeine concentrations are elevated, such as may be the case during caffeine intoxication.

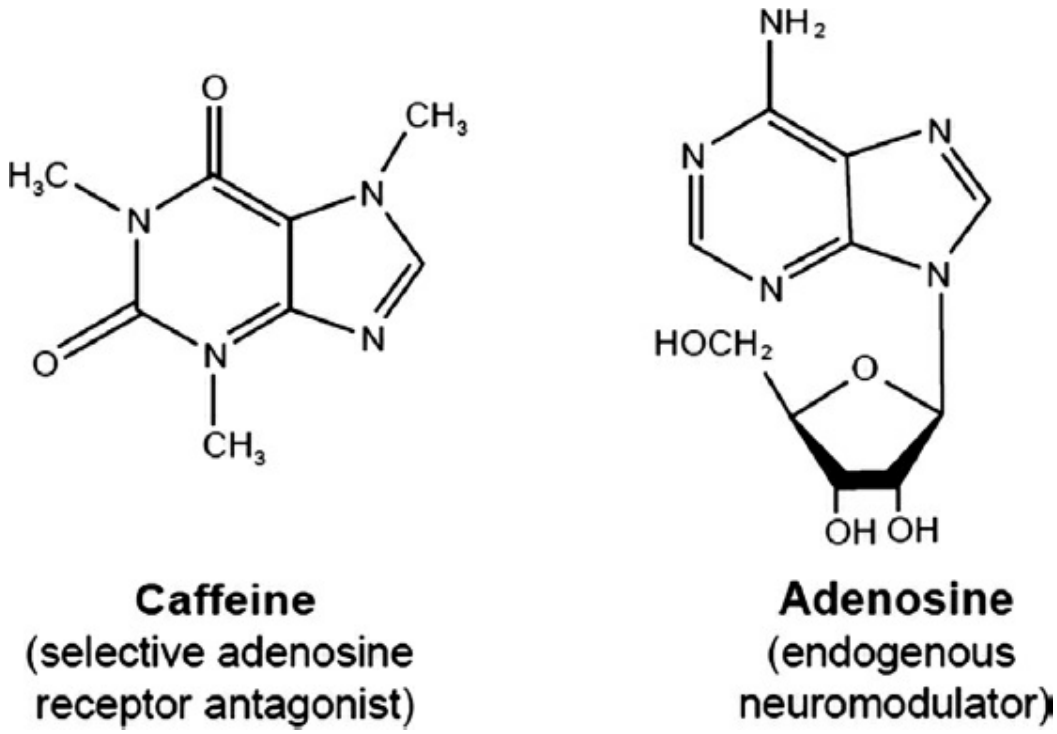

FIGURE 11.4–3. The chemical structures of caffeine (1,3,7-trimethylxanthine) and adenosine. Caffeine is structurally similar to adenosine, an endogenous neuromodulator. Most of the physiological effects of caffeine, including the central nervous system stimulant effects, are believed to be mediated through adenosine receptor antagonism.

Subjective Effects, Performance Effects, and Reinforcement

Subjective reactions to caffeine are dose-dependent and may vary depending on individual differences in sensitivity and caffeine tolerance. Single low-to-moderate doses of caffeine (i.e., 20 to 200 mg) generally produce pleasurable subjective effects. Studies have shown that such doses of caffeine result in increased ratings on measures such as well-being, alertness, energy, concentration, sociability, and motivation to work. In addition, these doses of caffeine produce decreases in ratings of feeling sleepy or tired. In general, high doses of caffeine (e.g., 400 to 800 mg) can produce negative subjective effects such as anxiety, nervousness, and jitteriness.

A large number of studies have examined the effects of caffeine on human performance. Caffeine appears to reliably improve performance especially when performance has been degraded by fatigue, sleep deprivation, prolonged vigilance, or caffeine abstinence. Compared with placebo, caffeine may improve reaction time, tapping speed, sustained attention (or vigilance), and perhaps memory. A rich literature on the effects of caffeine on exercise performance suggests that, relative to placebo, caffeine enhances performance during long-term (30 to 60 minutes) aerobic exercise, reduces ratings of perceived exhaustion, and improves power output or speed. Some studies, but not all, have also demonstrated caffeine-induced enhancement during short-term, high-

intensity exercise.

Although animal studies have generally found it difficult to demonstrate that caffeine functions as a reinforcer, well-controlled studies with humans have shown that people choose caffeine over placebo when given the choice under controlled double-blind experimental conditions. In habitual users, the reinforcing effects of caffeine are potentiated by its ability to suppress low-grade withdrawal symptoms after overnight abstinence. Thus, the positive subjective effects of caffeine and its ability to ward off the negative effects of caffeine withdrawal contribute to the regular use of caffeine. Studies have also shown that repeatedly pairing caffeine with a novel-flavored drink produces increased preference ratings for the drink, while pairing a novel-flavored drink with a placebo (in abstinent habitual caffeine consumers) decreases preference ratings for the flavor. These findings suggest that caffeine may play an important role in the development of preferences for specific types of beverages.

Genetics and Caffeine Use

There may be some genetic predisposition to continued caffeine use following exposure to caffeine. Investigations comparing coffee or caffeine use in monozygotic versus dizygotic twins show higher concordance rates for monozygotic twins for total caffeine consumption, heavy use, caffeine tolerance, caffeine withdrawal, caffeine intoxication, and coffee and tea intake, with heritabilities ranging between 34 and 77 percent. Multivariate structural equation modeling of caffeine use, cigarette smoking, and alcohol use suggests that a common genetic factor—polysubstance use—underlines use of these three substances. An analysis of monozygotic and dizygotic male and female twin pairs demonstrated that dependence on caffeine and nicotine were substantially influenced by genetic factors unique to these licit drugs and distinct from genetic factors found to be common among illicit drugs. Furthermore, caffeine dependence was shown to have the greatest substance-specific genetic variance among the five drug dependencies that were evaluated. The *ADORA2A* gene, which codes for the adenosine A2A receptor, and the *CYP1A2* gene, which codes for the primary enzyme responsible for caffeine metabolism, have been shown to be associated with the effects of caffeine, caffeine consumption, and health outcomes. ADORA2A receptor gene polymorphisms are associated with caffeine consumption, self-reported caffeine sensitivity, and caffeine's effects on psychomotor vigilance, anxiety, sleep, and blood pressure. Adenosine A_{2A} receptor knockout mice show a reduced preference for caffeine compared to controls. Polymorphisms in the *CYP1A2* gene are associated with variability in caffeine consumption. Moreover, individuals who carry the slow metabolism variant of the *CYP1A2* gene appear to be at increased risk for coffee associated hypertension and myocardial infarction,

and decreased risk for Parkinson Disease. Meta-analyses have found genome wide associations between caffeine use and variants of the ACH receptor gene (aryl hydrocarbon receptor, which regulates CYP12A) and *CYP12A* gene.

Age, Sex, and Race

The relationship between long-term chronic caffeine use and demographic features such as age, sex, and race has not been widely studied. Population-based surveys reveal that daily caffeine exposure increases with age up until middle-age and then begins to slightly decline. Some studies show that men use more caffeine than women, but this effect disappears when accounting for differences in body weight (i.e., mg/kg). There are no data that specifically address caffeine use for different races. There is some evidence suggesting that, for both children and adults in the United States, whites consume more caffeine than African Americans.

Special Populations

Cigarette smokers consume more caffeine than nonsmokers. This observation may reflect a common genetic vulnerability to caffeine use and cigarette smoking. It may also be related to increased rates of caffeine elimination in cigarette smokers. Preclinical and clinical studies indicate that regular caffeine use can potentiate the reinforcing effects of nicotine. Heavy use and clinical dependence on alcohol and other SUDs are also associated with heavy use and clinical dependence on caffeine. One study found that among pregnant women, a family history of alcohol dependence was associated with heavy use of caffeine both before and during pregnancy, despite physician advice to cease caffeine use during pregnancy. Another study also showed that a greater proportion of heavy caffeine consumers use benzodiazepines.

High daily caffeine consumption has also been found among individuals with a variety of psychiatric disorders, particularly among those with schizophrenia. One twin study found significant positive relationships between caffeine use, caffeine toxicity and dependence, and lifetime history of all seven psychiatric disorders that were assessed, including major depression, generalized anxiety disorder, panic disorder, alcohol dependence, antisocial behavior, and cannabis and cocaine abuse. Interestingly, the authors concluded that shared familial factors accounted for the relationships between caffeine use and the psychiatric disorders, rather than a direct link.

In terms of personality traits, some studies have shown that heavy caffeine use (more than 200 mg per day) and caffeine dependence is positively associated with novelty seeking or sensation seeking.

CAFFEINE INTOXICATION

Caffeine intoxication has long been recognized as a discrete syndrome associated with the use of a significant amount of caffeine. DSM-5 diagnostic guidelines state that the diagnosis is dependent on recent consumption of at least 250 mg of caffeine per day, but higher doses (e.g., more than 500 mg) are usually associated with the syndrome. Intoxication can be the result of an ingestion of a large amount of caffeine by a person who has not regularly consumed caffeine and thus may represent an overdose of caffeine. Alternatively, caffeine intoxication can occur in a person who chronically consumes large amounts of caffeine, which produces a more complicated clinical picture. Interestingly, while intoxication of some psychoactive substances, such as alcohol, is sought out by some people, it appears that most individuals do not actively seek repeated episodes of caffeine intoxication (Table 11.4–3).

Epidemiology

There is an important need for data on the incidence and prevalence of caffeine intoxication in the general population. To date most studies investigating caffeine intoxication symptoms have involved selected populations (e.g., psychiatric inpatients, college students) using ambiguous or incomplete criteria. Notably, one random-digit telephone survey among 162 caffeine users in Vermont found that two thirds of respondents reported at least one caffeine intoxication symptom and 7 percent met the DSM criteria for caffeine intoxication in the past year. One study that assessed more than 3,600 twins found that 29 percent reported having experienced feeling ill or shaky or jittery after consuming caffeinated beverages. Findings from a recent study suggest that psychiatric patients may have higher rates of caffeine intoxication than healthy controls. Reports from the poison control center's national poison data system provide some information about the rate of adverse effects from caffeine. For example, in 2012, in the category of single-substance exposures (i.e., not involving caffeine contained in dietary supplements or energy drinks), there were 2,466 cases involving caffeine. Thirty-three percent of these exposures were treated in a health care facility. Of all cases, 37 percent involved individuals 5 years and younger and 57 percent involved individuals 19 years and younger. This same year there were also 2,186 cases involving energy drinks and 296 involving energy products. Of these exposures, 25 percent were treated in a health care facility. Forty-five percent of cases involved individuals 5 years and younger and 73 percent involved individuals 19 years and younger. The American Association of Poison Control Centers has issued an alert for energy drinks stating that they can be dangerous for children or adolescents.

Table 11.4–3.
Caffeine Intoxication Symptoms

DSM-5 Symptoms
Restlessness
Nervousness
Excitement
Insomnia
Flushed face
Diuresis
Gastrointestinal disturbance
Muscle twitching
Rambling thoughts and speech
Tachycardia or cardiac arrhythmia
Inexhaustibility
Psychomotor agitation
Additional Symptoms
Fever
Irritability
Tremors
Sensory disturbances
Tachypnea
Headaches
Nausea/Vomiting
Seizures
Dizziness
Delirium

Diagnosis and Clinical Features

Signs and symptoms of caffeine intoxication include restlessness, nervousness, excitement, insomnia, flushed face, diuresis, GI disturbance, muscle twitching, rambling thoughts and speech, tachycardia or cardiac arrhythmia, inexhaustibility, and psychomotor agitation. To receive a DSM-5 diagnosis, patients must present with five or more of the aforementioned symptoms along with clinically meaningful distress or functional impairment. In addition, patients with caffeine intoxication can have fever, irritability, tremors, sensory disturbances, tachypnea, headaches, vomiting, and seizures. A series of case reports based on poison control center data found that a particular energy drink containing 250 mg of caffeine was implicated a number of times in recent caffeine intoxication incidents. The most common symptoms reported were hypertension, jitteriness, agitation, tremors, nausea, vomiting, and dizziness. Finally, although rare, it may be possible for a patient to develop delirium from ingesting an acute, extremely high dose of caffeine, and cases of fatal overdose and suicide from excessive caffeine ingestion have been documented.

Pathology and Laboratory Examination

Blood serum levels greater than 100 μg/mL may be lethal, although there have been reports of individuals surviving caffeine serum levels up to 200 μg/mL. However, it has also been noted that serum caffeine levels may be a poor indicator of toxicity due to the variable relationship between blood concentration and toxic effects. Urine or blood drug screens can also be used to rule out the presence of other drugs in the body.

Differential Diagnosis

There are a number of disorders that should be included in the differential diagnosis of caffeine intoxication. These include other substance-related disorders, such as intoxication from other stimulants (e.g., cocaine or amphetamines), and withdrawal from drugs such as sedative hypnotics (e.g., benzodiazepines), alcohol, or nicotine. In addition, other psychiatric disorders (e.g., anxiety disorders, mania, sleep disturbances), somatic disorders (e.g., arrhythmia, hyperthyroidism), and medication side effects (e.g., akathisia) should be considered in the differential diagnosis of caffeine intoxication. The final diagnosis of caffeine intoxication depends on confirming the ingestion of a significant quantity of caffeine prior to the onset of symptoms.

Course and Prognosis

Caffeine has a relatively short half-life (4 to 6 hours), and thus caffeine intoxication will typically resolve quickly and often with no significant sequelae. Patients with caffeine intoxication generally have a good prognosis, although there have been isolated reports of overdose or completed suicide by ingesting a massive amount of caffeine. The lethal dose of caffeine is generally considered to be about 10 g or the equivalent of 80 to 100 cups of brewed coffee.

Treatment

The first step in treating a patient presenting with symptoms suggesting caffeine intoxication is to confirm the diagnosis. It is important to consider all possible sources of caffeine when considering a diagnosis of caffeine intoxication, including medications, dietary supplements, energy drinks, and beverages that may not be initially recognized as containing caffeine. Because caffeine has a relatively short half-life, short-term management of the patient can be supportive while the syndrome resolves spontaneously. For the patient who habitually consumes large amounts of caffeine and repeatedly experiences episodes of caffeine intoxication, it may be helpful to assist the patient to recognize his or her total daily caffeine consumption through the use of daily diaries, teach the patient about the adverse effects associated with caffeine, and provide treatment assistance (see Table 11.4–

4). In the case of significant caffeine overdose, the patient should receive intensive monitoring, symptomatic treatment (e.g., for tachycardia), stomach aspiration, and assessment of serum caffeine level. Caffeine overdose has been successfully treated with hemodialysis or hemoperfusion.

CAFFEINE-INDUCED ANXIETY DISORDER

In addition to anxiety symptoms that are a feature of caffeine intoxication, caffeine can also cause an anxiety disorder. A caffeine-induced anxiety disorder can be panic disorder, generalized anxiety disorder, social phobia, or obsessive-compulsive disorder, although the patient does not need to fulfill all criteria for one of these disorders to quality for a diagnosis of caffeine-induced anxiety disorder. There has been no research examining this caffeine-related disorder, although there have been numerous investigations examining the anxiogenic effects of caffeine. Patients with anxiety disorders or anxiety sensitivity generally report greater anxiety than control subjects after consuming caffeine. Even among individuals who do not have an anxiety disorder, a sufficiently high dose of caffeine can produce panic attacks and increased anxiety ratings. First-degree relatives of patients with panic disorder also show an anxiogenic response to a high dose of caffeine. Polymorphisms in the *ADORA2A* gene are associated with anxiogenic responses to caffeine. Some studies have found that individuals with anxiety disorders, such as panic disorder, use lower levels of caffeine than healthy controls. However, some studies have failed to show a relationship between anxiety disorders and caffeine use, while others have found a relationship between greater caffeine use and anxiety disorders. It seems reasonable to conclude that some but not all individuals with anxiety disorders will spontaneously limit caffeine exposure.

Table 11.4–4.
Caffeine Modification Guidelines

1. *Education*
   Patients should be educated about potential sources of caffeine. Provide patients with a list of common caffeinated products (c.f. Table 11.4–2). It may be helpful to point out that caffeine is present in some noncola soft drinks, foods, and over-the-counter medications and to emphasize the wide variability in caffeine levels within and across product types.
2. *Self-monitoring*
   Caffeine consumption should be self-monitored for 1–2 weeks to determine a baseline level. Calculate daily caffeine exposure in milligrams, taking into account the caffeine content of specific products, the serving sizes, and the number of servings. If a self-monitoring period is not practical, treatment providers can determine a rough estimate of total caffeine consumption via patient report. Patients should continue to self-monitor during the caffeine reduction phase of treatment.
3. *Determine a caffeine modification goal*
   Some individuals may be interested in eliminating caffeine entirely while others may wish to reduce their caffeine consumption. Patients should be advised to set a goal of consuming no more than 50 mg per day or to use caffeine only occasionally (e.g., no more than 2 consecutive days) if

they want to avoid physical dependence on caffeine.

4. *Determine a dose reduction schedule*
   A gradual reduction schedule should help to minimize or prevent caffeine withdrawal symptoms. A reasonable decrease would be 25 percent of the baseline dose every week or so until the caffeine reduction or cessation goal is achieved. Caffeinated beverages can either be gradually omitted or can be mixed with decaffeinated beverages with the ratio of caffeinated product decreasing over time.
5. *Offer treatment advice and follow-up*
   Patients may benefit from behavior modification techniques that are effective in the treatment of dependence on other substances (e.g., nicotine). Such techniques may include self-monitoring, coping response training, reinforcement for abstinence, identifying barriers of change, social support, and framing withdrawal as a temporary inconvenience. Follow up with patients at a later date to check on their progress.

The diagnosis of caffeine-induced anxiety disorder depends on linking the use of caffeine to the anxiety symptoms of concern. For a patient with a suspected caffeine-induced anxiety disorder, a trial of caffeine abstinence may aid in clarifying the diagnosis. Caffeine-induced anxiety disorder should be diagnosed in patients with caffeine intoxication only if the anxiety disturbance is in excess to that which would be expected from intoxication. Other diagnoses to consider include caffeine intoxication, a primary anxiety disorder, and an anxiety disorder due to a general medical condition.

Mr. B was a 28-year-old single African American male graduate student who was in good health and had no history of previous psychiatric evaluation or treatment. He took no medications, did not smoke or consume alcohol, and had no current or past history of illicit drug use.

His chief complaint was that he had begun feeling mounting “anxiety” when working in the laboratory where he was pursuing his graduate studies. His work had been progressing well, he felt his relationship with his advisor was good and supportive, and he could not identify any problems with staff or peers that might explain his anxiety. He had been working long hours, but found the work interesting and had recently had his first paper accepted for publication.

Despite these successes, he reported feeling a “crescendoing anxiety” as his day would progress. He noted that by the afternoon he would be experiencing palpitations, bursts of his heart racing, tremors in his hands, and an overall feeling of “being on the edge.” He also noted a nervous energy in the afternoons. These experiences were occurring daily and seemed confined to the laboratory (although he admitted he was in the laboratory every day of the week).

When reviewing Mr. B’s caffeine intake, it was found he was consuming excessive amounts of coffee. Staff made a large urn of caffeinated coffee each morning, and Mr. B routinely started with a large mug of coffee. Over the course of the morning he would consume three to four mugs of coffee (the equivalent of about six to eight 6-oz cups of coffee), and continued this level of use throughout the afternoon. He occasionally had a single can of caffeinated soft drink, and used no other forms of caffeine on a regular basis. Mr. B estimated he drank a total of six to eight or more mugs of coffee per day (which was estimated to be at least 1,200 mg of caffeine per day). Once pointed out to him, he realized this level of caffeine consumption was considerably higher than at any other time in his life. He admitted he liked the taste of coffee and felt a burst of energy in the morning when he drank coffee that helped him start his day.

Mr. B and his physician developed a plan to decrease his caffeine use by tapering off caffeine. (Details of such a tapering schedule can be found in the section on treatment below.) Mr. B was successful in decreasing his caffeine use and had good resolution of his anxiety symptoms once his daily caffeine use had been markedly decreased.

Treatment

There are no studies on the treatment of caffeine-induced anxiety disorder. However, one study reported significant improvements in anxiety symptoms among patients seeking treatment at an anxiety disorders clinic who had been instructed to cease caffeine use for one week. Notably, some patients required no additional treatment and had successful discontinuation of anxiolytic medications. In cases of suspected caffeine-induced anxiety disorder, clinicians may want to avoid the use of anxiolytic medications until caffeine use has been eliminated. Clinicians should also be aware that some patients may be hesitant about eliminating caffeine and may express skepticism about the role of caffeine in their anxiety symptoms. In such instances clinicians should try to encourage patients to take an empirical approach and attempt a caffeine abstinence trial. General guidelines for caffeine modification can be recommended (see Table 11.4–4).

CAFFEINE-INDUCED SLEEP DISORDER

Caffeine-induced sleep disorder is characterized by a prominent sleep disturbance (e.g., insomnia, hypersomnia, parasomnia, or mixed) that is etiologically related to caffeine use. Caffeine use is most frequently associated with insomnia, although there are case reports of patients who have hypersomnia in response to caffeine. Caffeine-induced sleep disorder should be diagnosed in patients with caffeine intoxication only if the sleep disturbance is in excess to that which would be expected from intoxication. Other diagnoses that should be considered in the differential diagnosis of caffeine-induced sleep disorder include caffeine intoxication and caffeine withdrawal, sleep disturbances due to another mental disorder, a primary sleep disorder, and a sleep disorder due to a general medical condition.

Although there are essentially no studies directly examining caffeine-induced sleep disorder, there has been a great deal of research on the relationship between caffeine use and sleep. The closer caffeine is taken to bedtime, the more likely it is to produce disruptive effects on sleep. It has been shown that consuming 200 mg of caffeine before bedtime (the equivalent of about two cups of brewed coffee) delays sleep onset and reduces sleep efficiency and total sleep time. One study also showed that 200 mg of caffeine taken early in the morning disrupted sleep that evening.

Sleep disturbances secondary to caffeine use are more likely in people who are not regular consumers of caffeine. Regular caffeine consumers develop some tolerance to the sleep-disrupting effects of caffeine; however, because this tolerance is incomplete, they still may be vulnerable to caffeine-related sleep problems. Genetic factors may also underlie individual differences in the sleep disruptive effects of caffeine. A

polymorphism in the *ADORA2A* gene has been identified as being associated with differential responses to caffeine's effects on both subjective and objective measures of sleep.

It should be noted that caffeine can help to prevent unwanted sleep (e.g., such as when driving long distances) and can have beneficial effects when a person is sleep deprived. Studies have shown that sleep deprivation is associated with decrements in performance on a variety of tasks, and caffeine can reduce these performance decrements (although it does not completely reverse them).

Treatment

There are no studies on the treatment of caffeine-induced sleep disorder. In cases of suspected caffeine-induced sleep disorder, clinicians may want to discourage the use of sleep medications until caffeine use has been eliminated. Although patients may be reluctant to eliminate caffeine, clinicians should try to encourage patients to attempt a caffeine abstinence trial to rule out an etiological role of caffeine in sleep problems (see Table 11.4–4).

UNSPECIFIED CAFFEINE-RELATED DISORDER

This category is included in DSM-5 as a substance-related disorder and is to be used when the patient has a condition that is caffeine-related but not diagnosable as caffeine intoxication, caffeine-induced anxiety disorder, caffeine-induced sleep disorder, or caffeine withdrawal. Caffeine use disorder could be coded as an unspecified caffeine-related disorder.

CAFFEINE WITHDRAWAL

Caffeine withdrawal is now a diagnosis in the DSM-5. The research literature on caffeine withdrawal has more than doubled since caffeine withdrawal first appeared as a research diagnosis in the DSM-IV, and there is now a strong empirical basis for recognizing caffeine withdrawal as a discrete and clinically significant diagnostic category.

Epidemiology

Studies have examined the incidence of caffeine withdrawal under conditions in which it is experimentally induced in a sample of subjects. In a study of normal community volunteers whose average daily caffeine consumption was similar to that of the general population, 52 percent of the subjects had moderate or severe headache when caffeine withdrawal was experimentally induced under double-blind conditions. Other similar studies have found that half or more subjects will experience symptoms of caffeine withdrawal when it is experimentally induced (i.e., subjects undergo a double-blind discontinuation of their daily caffeine use).

In the United States it is estimated that more than 85 percent of the adult population are regular consumers of caffeine, suggesting that there are large numbers of people who potentially could experience caffeine withdrawal. However, there have been relatively few attempts to conduct population-based assessments of caffeine withdrawal. Such studies are usually surveys that find a wide range of rates of reports of caffeine withdrawal symptoms (i.e., from 8 to 58 percent). In one population-based survey, 71 percent of individuals who stopped their caffeine use in an attempt to quit permanently reported experiencing DSM-IV-TR defined withdrawal symptoms, and 24 percent reported having headache plus other symptoms that interfered with performance. In a sample of individuals seeking treatment for problematic caffeine use, 96 percent reported experiencing caffeine withdrawal upon abstinence. Headache was the most frequently endorsed criterion (i.e., 89 percent) and 43 percent of the sample reported that caffeine withdrawal caused functional impairment (e.g., missing work). Caffeine withdrawal has been documented among individuals fasting for medical procedures or religious observances (e.g., fasting during Ramadan, giving up coffee for Lent). Caffeine withdrawal has also been documented in children and adolescents and in neonates born to mothers who consumed large amounts of caffeine during pregnancy. Those who chronically use caffeine-containing analgesics may also be at increased risk of experiencing caffeine withdrawal upon abstinence.

Comorbidity

According to DSM-5 caffeine withdrawal may be associated with major depressive disorder, panic disorder, general anxiety disorder, antisocial personality disorder, alcohol use disorder, and cannabis and cocaine use.

Diagnosis and Clinical Features

A review of the research literature on caffeine withdrawal summarized caffeine withdrawal signs and symptoms reported in 66 studies (Table 11.4–5). Symptoms that were deemed to have the strongest empirical basis included headache, fatigue/drowsiness, difficulty concentrating, mood disturbances (e.g., dysphoric mood, depressed mood, or irritability), and flu-like symptoms (e.g., nausea, vomiting, and muscle pain/stiffness). These five symptom categories are those that appear in the DSM-5. In the DSM-5, a diagnosis is made when a patient experiences symptoms in three or more of these symptom categories, along with clinically significant distress or impairment, after abrupt cessation or a reduction in caffeine use. Patients can experience signs and symptoms even without the concurrent presence of headache, suggesting that the other features of caffeine withdrawal are not simply secondary to the headache. Although

anxiety was a proposed symptom in the DSM-IV-TR, anxiety is rarely observed in controlled studies of caffeine withdrawal and was omitted from DSM-5. Other withdrawal signs and symptoms that have been reported in the scientific literature are listed in Table 11.4–5. It is important to note that patients may be unaware of their physical dependence on caffeine and under conditions of caffeine abstinence may misattribute symptoms of caffeine withdrawal to other causes (e.g., think they are coming down with an illness such as the flu or that they did not get enough sleep).

Table 11.4–5.
Empirically Validated Signs and Symptoms Resulting from Caffeine Abstinence

Strong Evidence	Suggestive Evidence
Headache	Decreased desire to socialize
Tiredness/fatigue	Unmotivated for work
Decreased energy/activeness	Decreased self-confidence
Decreased alertness/attentiveness	Yawning
Drowsiness/sleepiness	Heavy feelings in arms and legs
Decreased contentedness/well-being	Craving/strong desire to use
Depressed mood	Analgesic use
Difficulty concentrating	Increased nighttime sleep quality/duration
Irritability	Impaired behavioral and cognitive performance
Muzzy/foggy/not clearheaded	Decreased motor activity
Flu-like symptoms	Increased cerebral blood flow
Nausea/vomiting	Electroencephalogram (EEG) changes
Muscle pain/stiffness	

From Juliano LM, Griffiths RR. A critical review of caffeine withdrawal: Empirical validation of symptoms and signs, incidence, severity, and associated features. *Psychopharmacology*. 2004;176:1, with permission.

Caffeine Withdrawal Headache. Headache is the most common feature associated with caffeine withdrawal. It is usually described as a generalized, throbbing headache that is sensitive to movement and worsened by exercise and Valsalva maneuver. It typically begins 12 to 24 hours after the last ingestion of caffeine, although the onset can occur as long as 40 hours after the last dose. It usually resolves in 2 to 4 days, although it has been suggested that caffeine withdrawal headaches can continue for up to 21 days. Caffeine consumption usually rapidly (30 to 60 minutes) and completely suppresses a caffeine withdrawal headache.

Caffeine Withdrawal and Postoperative Headache. Headache is a frequent postoperative symptom. Patients undergoing an operative procedure are sometimes required to abstain from all oral intake prior to the procedure, including daily dietary caffeine they would typically consume. Studies indicate that caffeine withdrawal plays an etiological role

in postoperative headache. For example, greater use of caffeine has been shown to be associated with a greater risk of postoperative headache in a dose-related fashion. Furthermore, caffeine users given caffeine on the day of the surgical procedure have lower rates of postoperative headache than those given placebo.

Caffeine Withdrawal and Caffeine Dose. The probability and severity of withdrawal increases as a function of increases in the daily dose of caffeine consumed. However, it should be noted that caffeine withdrawal can occur after ingestion of relatively low dosages of caffeine, such as 100 mg per day (the equivalent of about one cup of brewed coffee). Caffeine withdrawal can also develop after as few as 3 days' exposure to moderate doses of caffeine (e.g., 300 mg per day). Low doses of caffeine (e.g., 25 mg) are capable of suppressing the symptoms of caffeine withdrawal in individuals who stop consuming moderate doses of caffeine.

Severity of Caffeine Withdrawal. Although the severity of withdrawal tends to increase with increases in daily caffeine dose consumed, there are marked differences within and between individuals with regard to the incidence and severity of caffeine withdrawal. When caffeine withdrawal does occur, its severity can vary from mild to severe. A review of caffeine withdrawal studies found that the incidence of clinically significant impairment in functioning (e.g., unable to work) during caffeine withdrawal ranged from 10 to 55 percent depending on the study, with a median rate of 13 percent. A recent report describing individuals seeking treatment for problematic caffeine use found that 43 percent reported having experienced functional impairment during caffeine withdrawal.

Differential Diagnosis

Several disorders should be considered in the differential diagnosis of caffeine withdrawal. Caffeine withdrawal can mimic migraine and other headache disorders, viral illnesses, sinus conditions, tension, other drug withdrawal states (e.g., from amphetamines, cocaine), and medication side effects. The final determination of caffeine withdrawal should rest on a determination of the pattern and amount of caffeine consumption, the time interval between the last ingestion of caffeine and the onset of symptoms, and the particular clinical features presented by the patient. Attenuation of symptoms by a dose of caffeine may also be useful in clarifying the diagnosis.

Course and Prognosis

Caffeine withdrawal typically begins 12 to 24 hours after the last dose of caffeine, and usually resolves in 2 to 9 days. Most symptoms reach their

maximal intensity within the first 48 hours after cessation of caffeine use. However, it should be noted that there can be considerable variability among people in the manifestations, severity, and time course of caffeine withdrawal. Furthermore, it has been shown that the same person can experience different symptoms and different degrees of severity of symptoms across different episodes of caffeine withdrawal.

Treatment

There has been little work done in systematically investigating the optimal treatment for caffeine withdrawal. The symptoms of caffeine withdrawal typically resolve with the ingestion of a dose of caffeine. A person with a provisional diagnosis of caffeine withdrawal who continues to be symptomatic after caffeine consumption has been terminated for more than 2 weeks should be carefully re-evaluated. Caffeine withdrawal headaches may also respond to aspirin or other analgesics.

If caffeine withdrawal occurs because a person is attempting to abruptly stop all caffeine use, the clinician may want to advise the patient to taper caffeine consumption over time. Caffeine modification strategies are outlined in Table 11.4–4.

Ms. E was a 32-year-old single white woman employed full time at a local factory. She occasionally used nonsteroidal anti-inflammatory drugs, but was taking no regular prescription medications. She had a history of alcohol dependence, in remission for 9 years, and was otherwise in good health.

She first began consuming caffeine when she started college, and her current beverage of choice was coffee. She typically drank four to five mugs of coffee each day, and preferred to drink it without cream, milk, or sugar. She estimated that 5 minutes elapsed between the time she got up in the morning and the time she had her first cup of coffee; her roommate made a pot before Ms. E got up, and Ms. E immediately poured a mug when she got out of bed. She spaced her mugs over the course of the day, with her last mug either after lunch or with dinner.

Physicians had recommended she cut down or stop her coffee use because of complaints of mild indigestion, but she had been unable to do so. Her roommate had also complained about her coffee use at times. Ms. E routinely drank hot coffee in her car and had spilled it and burned herself on one occasion.

When she had stopped caffeine abruptly, Ms. E experienced marked irritability; poor concentration; and a severe, generalized headache. When asked to rate the severity of the headache, she replied that "on a scale of 1 to 10 it's a 12." She also had muscle aches, low energy, lethargy, and a craving to drink a mug of coffee. On the day that she had stopped coffee use abruptly, she left work 2 hours early because of problems with concentrating on the job and went to bed several hours earlier than usual. She then returned to her usual pattern of coffee use.

CAFFEINE USE DISORDER

An SUD diagnosis (referred to in prior editions of the DSM as "substance dependence") consists of a constellation of diagnostic criteria that are linked by the theme of problematic use of the drug. Studies have identified adults and adolescents with problematic caffeine consumption. Caffeine use disorder was added as a diagnosis for further study in the DSM-5. Prior

editions of the DSM did not recognize a disorder pertaining to problematic caffeine use, albeit a similar diagnosis of substance dependence due to caffeine has been long recognized by the World Health Organization's *International Classification of Diseases* (ICD). The DSM-5 working research criteria for caffeine use disorder contain a much more restrictive set of criteria than the generic DSM-5 SUD criteria that apply to other drugs. Consistent with a research diagnosis classification, additional research is needed to more fully characterize the features and prevalence of caffeine use disorder, its clinical significance and prognosis, associated features, and effective treatment strategies.

Epidemiology

Caffeine use is common throughout the world. More than 85 percent of adults in the United States consume caffeine regularly, and average daily caffeine consumption has been estimated to be upwards of 300 mg (the equivalent of about three cups of brewed coffee). Given the substantial number of people who consume caffeine regularly and their use of daily dosages that produce mood-altering effects, even a very low prevalence rate of caffeine use disorder could represent a nontrivial number of people. There has only been one study that has examined the prevalence of caffeine dependence in the general population—a random-digit telephone survey that used the generic criteria from DSM-III-R. In that study, 30 percent of caffeine consumers fulfilled diagnostic criteria for caffeine dependence. The most commonly endorsed symptom was desire or unsuccessful efforts to cut down or control use (56 percent).

Comorbidity

Comorbid conditions associated with high levels of caffeine consumption are discussed in the caffeine intoxication section above. There are only a few published reports that describe caffeine-dependent individuals and comorbid conditions. In one study of 16 caffeine dependent adults, about two-thirds had a past history of a psychiatric disorder in remission, most commonly another substance abuse disorder. However, nearly half had a mood disorder in remission, one quarter had an anxiety disorder in remission, and one fifth had an eating disorder in remission. Notably, there was a clustering of nicotine dependence, caffeine dependence, and a history of an AUD in several subjects, and this clustering of these three disorders has been noted in other studies. In a study of pregnant women, those who met the criteria for caffeine dependence were nine times more likely to report a history of daily cigarette smoking compared to those who did not meet dependence criteria. Furthermore, in this same study women with a family history of alcoholism and a caffeine dependence diagnosis were six times more likely to have a lifetime history of alcohol abuse or dependence.

In a sample of individuals seeking treatment for problematic caffeine use, the most prevalent concurrent psychiatric disorders were anxiety and/or mood disorders (17 percent met criteria for each). In addition, 12 percent of the sample had a past history of alcohol dependence. Studies comparing caffeine-dependent and caffeine-nondependent adolescents have not demonstrated significant associations between caffeine dependence and psychiatric comorbidities.

Diagnosis and Clinical Features

Caffeine use disorder is defined as a cluster of physiological, behavioral, and psychological symptoms indicative of problematic caffeine use along with difficulty controlling use despite a desire to do so. This is a condition for further study in DSM-5, and at the present time could be considered an unspecified caffeine-related disorder in DSM-5. Caffeine use disorder includes nine criteria; however, endorsement of the following three specific criteria is necessary for a diagnosis: (1) a persistent desire or unsuccessful efforts to cut down or control caffeine use; (2) continued caffeine use despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by caffeine; and (3) withdrawal, as manifested by either of the following: (a) the characteristic withdrawal syndrome for caffeine, or (b) caffeine (or a closely related substance) being taken to relieve or avoid withdrawal symptoms. Notably, a DSM-5 diagnosis of SUD for drugs other than caffeine requires the endorsement of only 2 of 11 criteria. The restrictive set of criteria for caffeine use disorder serves to limit the diagnosis to the most clinically important cases of problematic caffeine use. Individuals meeting criteria for caffeine use disorder have been found to be consumers of coffee, soft drinks, or both, and have demonstrated a wide range of daily caffeine intake. In one study involving individuals seeking treatment for caffeine use (84 percent diagnosed with caffeine use disorder) caffeine use ranged from 200 to 2,667 mg daily.

Differential Diagnosis

A clinical syndrome of caffeine use disorder could be easy to overlook because it is not widely recognized and acknowledged as a substance capable of producing problematic outcomes. When considering a possible diagnosis of caffeine use disorder, it is useful to include other substance-dependence syndromes in the differential diagnosis. Conditions that overlap with other caffeine-related disorders (e.g., intoxication, withdrawal) should also be considered.

Course and Prognosis

The course and prognosis for patients with a diagnosis of caffeine use

disorder is not known. Individuals with caffeine use disorder have reported continued use of caffeine despite repeated efforts to discontinue their caffeine use. Relapse to caffeine use after treatment has also been observed, suggesting that, like other drugs, caffeine dependence may be a disorder prone to relapse.

Treatment

There have been only a handful of studies describing treatment for problematic caffeine use. The most recent and largest scale study to date recruited 67 individuals seeking treatment for problematic caffeine use (84 percent meeting criteria for caffeine use disorder). A one session therapist-guided manualized treatment for problematic caffeine use including cognitive-behavioral strategies along with 5 weeks of progressively decreased consumption was evaluated relative to a waitlist control group. Treatment resulted in a significant reduction in self-reported caffeine use at the 6-week follow-up (i.e., a mean reduction from approximately 600 mg to 50 mg), and a significant reduction in salivary caffeine levels. Caffeine use was tracked for one year after treatment and on average no significant increases in caffeine use were observed.

Although medical specialists often recommend that patients modify caffeine use for various conditions (e.g., anxiety, insomnia, tachycardia, pregnancy), there may be a lack of recognition that caffeine modification may be difficult for some people. In fact, individuals diagnosed with caffeine use disorder have reported being unable to follow their doctors' recommendations to modify caffeine use. Thus, clinicians should not assume that the simple suggestion to modify caffeine use will be sufficient for successful behavior change. Effective strategies may include education, self-monitoring of daily use, coping response training, social support, and reinforcement for decreased use. SUD treatment approaches such as relapse prevention and motivational interviewing could also be applied to the treatment of caffeine use disorder. There is some evidence that caffeine withdrawal may interfere with quit attempts and thus a gradual tapering of caffeine dose over time may prove to be most efficacious. Table 11.4–4 provides suggested guidelines for caffeine use modification.

Ms. G is a 38-year-old female who is married with four children. She works part-time as a dental assistant. Ms. G presented to the primary care clinic with flu-like symptoms. She complained of nausea, heavy feelings in her arms and legs, muscle aches, and extreme lethargy. She also said she had a pounding headache that was one of the worse headaches she had ever experienced. Ms. G said that her symptoms began around 2 p.m. the day prior and she left work early because she could no longer function. She called her mother to help care for her school-aged children as her husband was away on a business trip. Upon examination Ms. G's heart rate and blood pressure were within normal limits and she was afebrile. An assessment of Ms. G's caffeine use revealed that she consumes three cups of black drip coffee each morning. Ms. G reported that she had not had caffeine the last two mornings because she had given it up for Lent. Ms. G said that she was unaware of her physical dependence on caffeine because she

hadn't gone more than a day without caffeine since her last pregnancy 9 years ago. Ms. G was diagnosed with caffeine withdrawal syndrome. Ms. G chose to drink a small cup of coffee and her symptoms resolved after approximately 45 minutes.

Ms. S is a 45-year-old Caucasian female who presented for assistance with problematic caffeine consumption, or what she described as her "caffeine addiction." Ms. S is married and has two elementary school aged children and works full-time as a nurse for a large research hospital. She takes no prescription medications, does not smoke, and has no past history of psychiatric problems. She reports occasionally drinking alcohol on the weekends. She reports having tried marijuana a few times in college and has no other history of illicit drug use. She began regularly consuming caffeinated beverages while in college and her consumption increased to its present levels during nursing school. Her primary source of caffeine comes from diet cola, which she consumes daily. She reports drinking coffee or tea only on occasion (e.g., when dining with friends). Ms. S has her first soft drink within 5 minutes of waking up in the morning and then continues drinking diet colas throughout the day until around 6 to 8 p.m. in the evening. She typically consumes a total of eight 20-oz diet cola bottles each day, the equivalent of about 600 mg of caffeine. She and her husband have argued about her soft drink use in the past, especially during her pregnancies. However, she continued to consume diet colas during both of her pregnancies. Ms. S described having attempted to give up diet cola many times, with her most recent attempt a few years prior. She said that after going cold turkey on the first day of Lent she experienced excruciating headaches, extreme fatigue, irritability, and an inability to function properly at work. She said that she returned to drinking diet colas after 3 days of "misery." Ms. S reports that her motivation for ceasing her diet cola consumption at this time is health-related. She is having increasing difficulties with urinary incontinence and furthermore feels that her diet cola consumption is generally not healthy. She is also presently complaining of sleep difficulties and anxiety that she attributes to the ailing health of her mother, although she acknowledged that her caffeine use could also be contributing to these symptoms. Ms. S reported that she was interested in completely eliminating diet cola from her diet rather than reducing her consumption. Ms. S and the therapist developed a schedule to gradually reduce her soft drink consumption over a period of 4 weeks and completely eliminating soft drinks on week 5 (i.e., 6 bottles to 4 bottles to 2 bottles to 1 bottle to 0 bottles). The therapist provided a booklet that described the effects of caffeine and caffeine withdrawal, and that contained tips for how to cope with temptations to consume caffeine and caffeine withdrawal symptoms should they occur. Ms. S was advised to substitute noncaffeinated drinks during the times that she normally consumed diet colas, and she identified carbonated water as a desirable alternative. A follow-up appointment was scheduled for 5 weeks later. At the 5-week follow-up Ms. S reported that she had reduced her consumption to two 12-oz diet colas per day. At the 6-month follow-up Ms. S reported having "been off" diet colas for the past 3 months. She reported having headaches for the first few days after completely eliminating caffeine but that they were tolerable, especially if she took ibuprofen. Ms. S reported that her problems with urinary incontinence and poor sleep had both improved and she felt confident that she could continue to refrain from diet cola consumption.

Caffeine Use and Nonpsychiatric Illnesses. Despite numerous studies examining the relationship between caffeine use and physical illness, significant health risk from nonreversible pathological consequences of caffeine use, such as cancer, heart disease, and human reproduction, has not been demonstrated. However, caffeine use is often considered to be contraindicated for various conditions, including generalized anxiety disorder, panic disorder, primary insomnia, gastroesophageal reflux, urinary incontinence, and pregnancy. In addition, the modest ability of caffeine to increase blood pressure and the documented cholesterol-

elevating compounds in unfiltered coffee have raised the issue of the relationship of caffeine and coffee use to cardiovascular disease. Finally, there may be a mild association between higher daily caffeine use in women and delayed conception, higher rates of miscarriage, and slightly lower birth weight. However, there are studies that have not found such associations, and effects when found are usually with relatively high daily doses of caffeine (e.g., the equivalent of five cups of brewed coffee per day). For a woman who is considering pregnancy, especially if there is some difficulty in conceiving, it may be useful to counsel eliminating caffeine use. Similarly, for a woman who becomes pregnant and has moderate-to-high daily caffeine consumption, a discussion about decreasing her daily caffeine use may be warranted. Caffeine may possibly offer protective effects against some diseases, most notably Parkinson Disease.

References

Clark I, Landolt HP. Coffee, caffeine, and sleep: a systematic review of epidemiological studies and randomized controlled trials. *Sleep Med Rev*. 2016.

Evatt DP, Juliano LM, Griffiths RR. A brief manualized treatment for problematic caffeine use: a randomized control trial. *J Consult Clin Psychol*. 2016;84(2):113–121.

Ferré S. Mechanisms of the psychostimulant effects of caffeine: implications for substance use disorders. *Psychopharmacology*. 2016;233(10):1963–1979.

Griffiths RR, Bigelow GE, Liebson IA. Human coffee drinking: reinforcing and physical dependence producing effects of caffeine. *J Pharmacol Exp Ther*. 1986;239:416.

Hughes JR, Oliveto AH, Liguori A, Carpenter J, Howard T. Endorsement of DSM-IV dependence criteria among caffeine users. *Drug Alcohol Depend*. 1998;52:99.

*James JE. *Understanding Caffeine: A Biobehavioral Analysis*. Thousand Oaks, CA: Sage; 1997.

*Juliano LM, Griffiths RR. A critical review of caffeine withdrawal: empirical validation of symptoms and signs, incidence, severity, and associated features. *Psychopharmacology*. 2004;176:1.

Juliano LM, Huntley ED, Harrell PT, Westerman AT. Development of the caffeine withdrawal symptom questionnaire: caffeine withdrawal symptoms cluster into 7 factors. *Drug Alcohol Depend*. 2012;124(3):229–234.

Kamimori GH, McLellan TM, Tate CM, Voss DM, Niro P, Lieberman HR. Caffeine improves reaction time, vigilance and logical reasoning during extended periods with restricted opportunities for sleep. *Psychopharmacology*. 2015;232(12):2031–2042.

Liu Rui, Guo X, Park Y, et al. Caffeine intake, smoking, and risk of Parkinson disease in men and women. *Am J Epidemiol*. 2012;175(11):1200–1207.

Meredith SE, Juliano LM, Hughes JR, Griffiths RR. Caffeine use disorder: a comprehensive review and research agenda. *J Caffeine Res*. 2013;3:114–130.

Mitchell DC, Knight CA, Hockenberry J, Teplansky R, Hartman TJ. Beverage caffeine intakes in the US. *Food Chem Toxicol*. 2014;63:136–142.

Mowry JB, Spyker DA, Cantilena LR Jr, Bailey JE, Ford M. 2012 annual report of the American association of poison control centers' national poison data system (NPDS): 30th annual report.

Clin Toxicol. 2013;51(10):949–1229.

Musgrave IF, Farrington RL, Hoban C, Byard RW. (2016). Caffeine toxicity in forensic practice: possible effects and under-appreciated sources. *Forensic Sci Med Pathol*. 2016;12(3):299–303.

Nardi AE, Lopes FL, Valenca AM, Freire RC, Veras AB. Caffeine challenge test in panic disorder and depression with panic attacks. *Compr Psychiatry*. 2007;48(3):257.

*Nawrot P, Jordan S, Eastwood J, Rotstein J, Hugenholtz A. Effects of caffeine on human health. *Food Addit Contam*. 2003;20(1):1.

Palacios N, Gao X, McCullough ML, et al. Caffeine and risk of Parkinson's disease in a large cohort of men and women. *Mov Disord*. 2012;27(10):1276–1282.

*Reissig CJ, Strain EC, Griffiths RR. Caffeinated energy drinks—a growing problem. *Drug Alcohol Depend*. 2009;99(1):1–10.

Retey JV, Adam M, Khatami R, Luhmann UF, Jung HH. A genetic variation in the adenosine A2A receptor gene (ADORA2A) contributes to individual sensitivity to caffeine effects on sleep. *Clin Pharmacol Ther*. 2007;81(5):692.

Sepkowitz KA. Energy drinks and caffeine-related adverse effects. *JAMA*. 2013;309(3):243–244.

Seifert SM, Seifert SA, Schaechter JL, et al. An analysis of energy-drink toxicity in the National Poison Data System. *Clin Toxicol*. 2013;51(7):566–574.

Silverman K, Evans SM, Strain EC, Griffiths RR. Withdrawal syndrome after the double-blind cessation of caffeine consumption. *N Engl J Med*. 1992;327:1109.

Somogyi LP. *Caffeine Intake in the U.S. Population*. Available at http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofFoods/CFSAN/CFSANFOL Accessed May 3, 2015.

Solinas M, Ferré S, You ZB, Karcz-Kubicha M, Popoli P. Caffeine induces dopamine and glutamate release in the shell of the nucleus accumbens. *J Neurosci*. 2002;22:6321.

Strain EC, Mumford GK, Silverman K, Griffiths RR. Caffeine dependence syndrome: evidence from case histories and experimental evaluations. *JAMA*. 1994;272:1043.

Van Dongen HP, Kerkhof GA. Effects of caffeine on sleep and cognition. Human Sleep and Cognition, Part II. *Clin Appl Res*. 2011;2:105.

Weber JG, Ereth MH, Danielson DR. Perioperative ingestion of caffeine and postoperative headache. *Mayo Clin Proc*. 1993;68:842.

Yang A, Palmer AA, de Wit H. Genetics of caffeine consumption and responses to caffeine. *Psychopharmacology (Berl)*. 2010;211(3):245–257.

▲ 11.5 Cannabis-Related Disorders

WAYNE D. HALL, PH.D., AND LOUISA DEGENHARDT, PH.D.

Cannabis is the most widely used illegal drug in the world, with an estimated 178 million users worldwide in 2012 (3.8 percent of the global adult population). Over the past 40 years cannabis use has become a common part of youth culture in most developed societies, with first use now occurring in the mid- to late teens. Cannabis is the fourth most commonly used psychoactive drug among adults in the United States, after caffeine, alcohol, and nicotine.

HISTORY

Cannabis preparations are obtained from the plant *Cannabis sativa*, which has been used in China, India, and the Middle East for approximately 8,000 years, primarily for its fiber and secondarily for its medicinal properties. Cannabis was introduced to Europe in the early 19th century by Napoleon's army returning from Egypt, and later in the same century to Britain for medical use by a surgeon who served in India. Cannabis was most often used in Europe for fiber and, to a lesser extent, for therapeutic purposes. It was first used recreationally by Parisian bohemians in the late 19th century.

Recreational cannabis use was introduced to the United States in the 1930s from Mexico and spread via jazz musicians to cities in the northeastern United States. Its use was banned in the United States in 1938 and in most other developed countries in 1961 under the Single Convention on Narcotic Drugs. Recreational cannabis use among bohemians in the United States in the 1940s and 1950s spread to the wider US youth population in the late 1960s. Use increased dramatically in the United States during the 1970s before peaking in the early 1980s, declining until the early 1990s, and stabilizing at a lower level since then. Cannabis use was disseminated to the international youth culture in the 1970s and 1980s via movies, music, media, and popular culture.

In the 1990s and 2000s, citizen-initiated referenda legalized the use of marijuana for medical purposes in around half the states of the United States. Recreational cannabis use was legalized by state referenda in Colorado and Washington State in 2012 and in Alaska and Oregon in 2014. It remains to be seen what impact these policy changes will have on cannabis use and cannabis-related harm in the United States.

CANNABIS PREPARATIONS

The cannabis plant occurs in male and female forms. The female plant

contains the highest concentrations of more than 60 cannabinoids that are unique to the plant. Delta-9-tetrahydrocannabinol (THC) is the cannabinoid that is primarily responsible for the psychoactive effects of cannabis. It is found in the resin that covers the flowering tops and upper leaves of the female plant. Most of the other cannabinoids are either inactive or only weakly active, although some, such as cannabidiol, may modulate the effects of THC (Fig. 11.5–1).

The most common cannabis preparations are marijuana, hashish, and hash oil. Marijuana is prepared from the dried flowering tops and leaves of the plant. Its potency depends on the growing conditions, the genetic characteristics of the plant, the ratio of THC to other cannabinoids, and the part of the plant that is used. The flowering tops have the highest THC concentration with much lower concentrations in the leaves, stems, and seeds. Varieties of cannabis cultivated for hemp fiber usually contain very low levels of THC. Cannabis plants may be grown to maximize their THC production by the sinsemilla method, in which only female plants are grown together.

The concentration of THC in marijuana may range from 0.5 to 5 percent. The sinsemilla variety may contain as much as 7 to 14 percent THC. The potency of marijuana preparations sold in the United States has increased from 0.5 percent in 1970 to 10.6 percent in 2010. Hashish or hash consists of dried cannabis resin that may contain 2 to 8 percent of THC. Hash oil is obtained by extracting THC from hashish (or marijuana) in oil and it may contain between 15 and 20 percent THC.

FIGURE 11.5–1. Marijuana (*Cannabis sativa*).

METHODS OF USE

Cannabis has traditionally been smoked as marijuana in hand-rolled cigarettes, or joints, which may include tobacco to assist burning. Using a water pipe, or bong, is also a popular way to smoke all cannabis

preparations, probably because it delivers a larger bolus dose of THC and maximizes the effect of a given amount of cannabis. Cannabis can also be administered using a vaporizer rather than by smoking to avoid the respiratory risks of smoking. Hashish may be mixed with tobacco and smoked as a joint or in a pipe, with or without tobacco. Because hash oil is extremely potent, a few drops may be applied to a cigarette or joint or the oil may be heated and the vapors inhaled. Cannabis users typically inhale the smoke or vapor deeply and hold their breath to maximize the absorption of THC by the lungs.

The oral route of administration may also be used by eating hashish baked in brownies or cookies. In experimental studies, THC is often swallowed in gelatin capsules containing THC dissolved in sesame oil. In India, cannabis may be consumed in the form of *bhang*, a tea brewed from the leaves and stems of the plant. THC is not water soluble and so is not readily injected. Almost all cannabis users in most developed societies smoke cannabis in joints or use bongs or vaporizers because the chemistry and pharmacology of cannabis make smoking or inhaling the most efficient ways to obtain THC. Given this, the reader should assume hereafter that cannabis is smoked or inhaled as a vapor, unless otherwise stated.

Comparative Nosology

A cannabis dependence syndrome and problematic cannabis use have been topics of recent research interest. The most recent editions of the two major diagnostic classification systems, the tenth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10), includes a diagnosis of cannabis dependence characterized by marked distress resulting from a recurring cluster of problems related to cannabis use that reflects impaired control over use, in the sense of continuing to use despite harms caused by use. The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) includes cannabis use disorders within the diagnosis of drug use disorders. ICD-10 includes a category of *harmful use* that attempts to capture problematic cannabis use that does not satisfy criteria for dependence but nonetheless causes the user harm. In DSM-5 drug use disorders vary in severity only. Research suggests that the ICD-10 and the old DSM-IV-TR classification systems agreed well in identifying cases of cannabis dependence, but showed poorer agreement in the classification of problem users. It remains to be seen how well DSM-5 and ICD-10 diagnoses will agree.

Epidemiology of Cannabis Use

The best data on patterns of cannabis use in the United States are provided by two major surveys of illicit drug use that have been undertaken since the early 1970s. These are the Monitoring the Future nationwide survey of high

school seniors, college students, and young adults conducted annually since 1975; and the National Household Survey on Drug Use and Health (NHSDUH), which has surveyed household samples of adults annually throughout the United States since 1972.

In 2012, 41.9 percent of a US national sample of persons 12 years and older reported that they had used cannabis at some time in their lives, 12.1 percent had used in the past year, and 7.3 percent in the past month (Table 11.5–1). Lifetime use increased from 17 percent among those aged 12 to 17 years to 52.2 per cent in those aged 18 to 25 years and 44.4 percent among those aged 26 and over. Rates of discontinuation of use were high: 71.5 percent of adults who had used cannabis at some time in their lives had not done so in the past year. Past-month cannabis use was more common among men (9.6 percent) than women (5 percent) and most common among those aged 18 to 25 years (18.7 percent).

The NHSDUH series from 1974 to 2000 show that rates of past-month cannabis use increased throughout the 1970s, peaked in 1979, and declined steadily throughout the 1980s to reach their lowest level in 1992 before increasing again after 1995. In the past few years (2007 to 2013), cannabis use in the past year and past month has declined among young adults.

Table 11.5–1.
Prevalence of Cannabis Use by Age and Gender According to the US National Household Survey on Drug Use and Health (2012)

	Lifetime	Past 12 Months	Past Month
12–17 years	17.0	13.5	7.2
18–25 years	52.2	31.5	18.7
26+ years	44.4	8.6	5.3
Males	7.7	15.2	9.6
Females	38.1	9.2	5.0
Total 12+ years	42.8	11.5	7.0

Note: Numbers are per 100 per year.

From SAMHSA. *Results from the 2012 National Survey on Drug Use and Health, Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Administration, Office of Applied Studies; 2014.

The Monitoring the Future project has estimated the prevalence of cannabis use among secondary school students, college students, and young adults since 1975. These surveys have revealed long waves of cannabis use among US adolescents since 1975. Among 18-year-olds, lifetime prevalence peaked at 65 percent in 1980, before falling to nearly half by the early 1990s, but increasing again after that time to 42 percent in 2006. Use in the past year peaked at 51 percent in 1979 and fell to 22 percent by 1992 before increasing to 32 percent in 2006. The discontinuation of use also declined, with fewer of those who had used

cannabis 10 or more times, ceasing use by age 18. Most of those who ceased cannabis use had not had a great deal of experience with cannabis. Although most users of other illicit drugs also had used cannabis, trends in the use of other illicit drugs were independent of the cannabis-use trends.

Changes in beliefs about the risks of cannabis use seem to be responsible for both the reduction in use between 1979 and 1991 and the rise in use after 1992. There is a strong negative correlation over time between the rates of cannabis use and the perceived risk of using cannabis and perceived peer disapproval of use. Between 1979 and 1992, a marked increase in perceived risk and a smaller increase in personal disapproval preceded a large decline in use. This study attributed the sharp upturn in cannabis use after 1992 to a decline in perceived risk in the preceding year.

Cannabis use is also common among the youth of most developed countries. Cannabis has been tried by many young adults in Europe and by the majority of young Australians. Most cannabis users in these countries, as in the United States, stop in their mid- to late 20s. Very few engage in daily cannabis use over a period of years. In the United States and Australia, about 10 percent of those who ever use cannabis become daily users, and another 20 to 30 percent use weekly. The pattern of use in developed countries differs from that in traditional cannabis-using countries, such as Egypt and India, where recreational cannabis use is uncommon and heavy cannabis use is confined to small, socially marginalized groups in the population.

Only a small proportion of cannabis users use the drug for several years or more. The daily or near daily use pattern over a period of years is the pattern with the greatest risk of producing adverse health and psychological consequences. Daily cannabis users are more likely to be male and less well educated; they are also more likely to regularly use alcohol and to have experimented with other illicit drugs such as amphetamines and other psychostimulants, hallucinogens, sedatives, and opioids.

Correlates of Cannabis Use

Age. First use of cannabis typically begins in the mid- to late teenage years, and the heaviest use typically occurs in the early 20s. Rates of cannabis use remain relatively high during the early 20s but decline during the late 20s. The majority of young adults who experiment with cannabis have done so by age 18 and rates of use decline steadily from the mid-20s into the early 30s.

Gender. Rates of cannabis use in the user's lifetime, the past year and the past week are consistently higher among males than females, as are daily use and long-term daily use.

Income. There is a positive relationship between income in adolescence and early adult life and cannabis use. In the United States, daily cannabis use is positively correlated with income and hours worked in a paid job.

Socioeconomic Status. The relationship between cannabis use and socioeconomic status is weak. Higher rates of cannabis use are sometimes found among lower socioeconomic status individuals, but in the past two decades there has been no relationship between parents' education and cannabis use among 12th-grade students in the United States, with the exception that those with lowest parental education had slightly lower cannabis use. That difference may be explained by differences in income during adolescence.

Ethnicity. Information on the relationship between ethnicity and cannabis use is limited. Even in the very large Monitoring the Future survey, samples from several years have to be combined to make reliable comparisons between the three largest ethnic groups. These show that African American students have lower rates of use in all grades than white or Hispanic students. Hispanics, on the other hand, tend to have the highest rates of use in the early grades, before the rates of school dropout increase.

Availability. In general, the more freely available a drug is, the more often it will be used. This has been found for alcohol consumption: the larger the number of licensed outlets and the longer the hours of trading, the higher the levels of community alcohol consumption and alcohol-related problems. There is limited evidence that this is the case for cannabis use. Self-reports from surveys on how easy it is to obtain cannabis have shown very little change over long periods of time during which rates of cannabis use have risen and fallen in the United States.

Pharmacology of Cannabinoids

Laboratory research with animals and humans has demonstrated that the primary psychoactive constituent in cannabis is THC. THC and other cannabinoid agonists produce their effects by interacting with an endogenous cannabinoid signaling system. There are at least two types of cannabinoid receptor (CB1 and CB2). The CB1 receptor is found primarily in the brain and mediates the psychological and behavioral effects of THC. The CB2 receptor is associated with the immune system and appears to modulate inflammatory responses. CB1 and CB2 receptors belong to a large group of G-Protein-coupled receptors membrane proteins that are involved in chemical signaling between nerve cells.

CB1 cannabinoid receptors are found in brain regions known to be

involved in mood, motor control, memory formation, regulation of food intake, pain modulation, and aspects of immune and reproductive functions. The fact that cannabis disrupts short-term memory is consistent with the abundance of CB1 receptors in the hippocampus, the brain region most closely associated with memory. The CB1 and CB2 receptors respond to endogenous cannabinoid ligands that are naturally produced in the brain. These ligands include anandamide and 2-arachidonyl-glycerol (2AG).

Some active compounds found in cannabis have effects that are independent of cannabinoid receptors and endogenous cannabinoids. For example, both THC and cannabidiol (CBD) can be neuroprotective through their antioxidative activity. Other likely examples of receptor-independent cannabinoid activity are modulation of activity in membrane-bound enzymes (such as adenosine triphosphatase [ATPase]), arachidonic acid release, and perturbation of membrane lipids. An important caveat is that concentrations of THC or CBD used in studies of cellular effects are generally much higher than the concentrations of THC or CBD that would be achieved in humans who smoke marijuana.

A typical joint of between 0.5 and 1.0 g of cannabis contains between 5 and 150 mg of THC. Twenty to 70 percent of the THC is found in the smoke that reaches the lungs; the rest is lost in sidestream smoke. Only 5 to 24 percent of the THC in a joint reaches the bloodstream when cannabis is smoked. Two to 3 mg of THC will produce a brief high in an occasional user, and a single joint may provide enough THC for two or three such individuals. A heavy cannabis smoker may use five or more joints per day, while heavy users in Jamaica, for example, may consume up to 420 mg of THC per day.

Different methods of using cannabis lead to differing absorption, metabolism, and excretion of THC. When cannabis is smoked, THC is absorbed from the lungs into the bloodstream within minutes. It is first metabolized in the lungs, and then in the liver. The metabolite 9-carboxy-THC can be detected in blood within minutes of smoking. Peak blood levels of THC are usually reached within 10 minutes of smoking, and decline to about 5 to 10 percent of their initial level within an hour. This rapid decline reflects the rapid conversion of THC to its metabolites and the distribution of THC to fatty tissues, including the brain.

When swallowed, THC takes 1 to 3 hours to enter the bloodstream, delaying the onset of psychoactive effects. Another metabolite, 11-hydroxy-THC, which is 20 percent more potent than THC, is found in high concentrations after cannabis is swallowed.

THC and its metabolites are highly fat soluble, so they may remain in the fatty tissues of the body for long periods of time. THC and its metabolites may accumulate in the body of chronic users because of their slow rate of clearance. They may be detected in the blood for several days,

and traces may persist for several weeks. THC may be stored in body fat for more than 28 days.

It has been argued that the medical literature underestimates the adverse health effects of cannabis because it is based on research conducted on less potent forms of cannabis than those that have become available in the past decade. The evidence suggests that the average potency of cannabis has increased but probably not to the extent often claimed in the popular media. Changes in patterns of cannabis use, with earlier age of first use and more regular use of more potent forms of cannabis, have probably been as important as increases in the average THC content of cannabis plants.

DIAGNOSTIC AND CLINICAL FEATURES

Cannabis Dependence

For much of the 1960s and 1970s, cannabis was not regarded as a drug of dependence because it did not seem to produce tolerance or a withdrawal syndrome as marked as that seen in alcohol and opioid dependence. The adoption of a broader concept of dependence in the late 1970s and early 1980s reduced the importance of tolerance and withdrawal in defining dependence and placed more emphasis on the compulsion to use, a narrowing of the drug-using repertoire, rapid reinstatement of dependent use after abstinence, and the high salience of drug use in the user's life.

Drug Dependence in DSM-5. Changes in the DSM-5 criteria for cannabis and other drug use disorders raise questions about relevance of epidemiological studies using earlier DSM criteria. The Epidemiological Catchment Area (ECA) study, for example, estimated that 4.4 percent of the US population met criteria for a diagnosis of cannabis abuse or dependence according to DSM-III. A third of those with lifetime cannabis abuse or dependence (38 percent) reported problems with cannabis use in the past year. The National Comorbidity Survey (NCS) conducted in the United States between 1990 and 1992 also found that 4.2 percent of adults met DSM-III-R criteria for cannabis dependence at some time in their lives. The proportion of people who had ever used cannabis who met criteria for cannabis dependence was 9 percent. In comparison, 32 percent of nicotine, 23 percent of heroin, 17 percent of cocaine, 15 percent of alcohol, and 11 percent of stimulant users met lifetime criteria for dependence. The NESARC estimated that in 2001 to 2002, 1.5 percent of US adults met DSM-IV criteria for past year cannabis use disorders.

The Risk of Cannabis Dependence. People who use cannabis daily over weeks to months are most likely to become dependent. Epidemiological data on disorders as defined in DSM-III-R and DSM-IV suggest the

following rules of thumb about the risks of cannabis dependence with varying frequencies of use. The risk of developing dependence is around 1 in 10 for anyone who ever use cannabis. It increases with increasing frequency of use: among those who use the drug more than a few times, to between 1 in 5 and 1 in 3; among those with a history of daily use, the dependence risk is around 1 in 2.

As a rule, the earlier the age of first use, the more often cannabis has been used, and the longer it has been used, the higher the risk of dependence. The risk of dependence is around one in six for those who start to use cannabis during mid-adolescence. Other factors that predict a higher risk of regular involvement with cannabis include: poor academic achievement, deviant behavior in childhood and adolescence, nonconformity and rebelliousness, personal distress and maladjustment, poor parental relationships, earlier use, and a parental history of drug and alcohol problems.

Clinical Populations. Cannabis-dependent persons seek help from treatment services in Australia, the United States, and Europe. In 2012 cannabis was the primary drug of abuse for 17.5 percent of clients at treatment agencies in the United States. There have also been increased numbers seeking treatment for cannabis problems in the Netherlands, where the personal use of cannabis has been decriminalized. This suggests that the increased rates of treatment seeking in the United States are not solely a by-product of increased use of legally coerced cannabis treatment. Common symptoms among persons seeking help to cease cannabis use include an inability to stop using, feeling bad about using cannabis, procrastinating, loss of self-confidence, memory loss, and withdrawal symptoms.

Intoxication

Most young people use cannabis in order to experience a "high," characterized by feelings of mild euphoria, relaxation and perceptual alterations, including time distortion, and the intensification of ordinary experiences, such as eating, watching films, listening to music, and engaging in sex. When used in a social setting, the high may be accompanied by infectious laughter, talkativeness, and increased sociability. Cognitive changes include impaired short-term memory and attention that make it easy for the user to become lost in pleasant reverie and difficult to sustain goal-directed mental activity. Motor skills, reaction time, motor coordination, and many forms of skilled psychomotor activity are impaired while the user is intoxicated, increasing the risks of accidents if users drive an automobile.

Intoxication Delirium

Psychotic symptoms, such as delusions and hallucinations, are very rare experiences that may occur at very high doses of THC and perhaps in lower doses in susceptible individuals with a pre-existing vulnerability to psychosis.

High doses of THC have been reported to produce visual and auditory hallucinations, delusional ideas, and thought disorders in normal volunteers. In cannabis-using cultures, such as India, a "cannabis psychosis" has been reported among heavy users of cannabis. In these cases, psychotic symptoms are reported to be preceded by unusually heavy cannabis use and to remit after a period of abstinence.

The existence of a cannabis psychosis in Western cultures is still a matter for debate. In its favor are case series of cannabis psychoses and a small number of controlled studies that compare the characteristics of cannabis psychoses with those of psychoses in individuals who were not using cannabis at the time of hospital admission. Critics of the hypothesis emphasize the fallibility of clinical judgments about etiology, the poorly specified criteria used in diagnosing these psychoses, the dearth of controlled studies, and the striking variations in the clinical features of putative "cannabis psychoses."

Cannabis and Schizophrenia

There is consistent epidemiological evidence that schizophrenia and dependent cannabis use are associated in community surveys and clinical populations. There is also substantial evidence from longitudinal studies suggesting that regular cannabis use can precipitate schizophrenia in vulnerable individuals or exacerbate its symptoms in those who have already developed the disorder. The hypothesis that cannabis use is a *contributory* cause of schizophrenia (one factor among others that increase the likelihood of developing the disorder) is more plausible than the alternative hypotheses that (a) persons with schizophrenia may use cannabis as a form of self-medication, or that (b) cannabis use is a marker of either other psychotogenic drug use (e.g., amphetamines) or of an increased genetic vulnerability to schizophrenia.

There is good clinical and epidemiological evidence that cannabis use may precipitate symptoms of schizophrenia in vulnerable individuals. This includes prospective studies in four countries that have found dose–response relationships between frequency of cannabis use, early age of onset, and the risk of developing psychotic disorders. These relationships have persisted after controlling for confounding variables, such as other drug use and pre-existing risk of psychosis.

A contributory causal relationship is biologically plausible. Psychotic disorders involve disturbances in the dopamine neurotransmitter systems, since drugs that increase dopamine release produce psychotic symptoms when given in large doses, and neuroleptic drugs that reduce psychotic

symptoms also reduce dopamine levels. Cannabinoids, such as THC, increase dopamine release. Psychotic disorders have been reported in elderly patients in clinical trials of cannabinoids for neuropathic pain.

There is good prospective evidence from a 35-year follow-up Swedish conscript study that cannabis use precipitates schizophrenia in persons who are vulnerable because of a personal or family history of schizophrenia. This hypothesis is consistent with the stress–diathesis model of schizophrenia in which the likelihood of developing schizophrenia is the product of stress acting on a genetic "diathesis" to develop schizophrenia. The Swedish findings have been confirmed in prospective studies in France, Germany, the Netherlands, and New Zealand that have been able to control for the most plausible alternative explanations of the relationship between cannabis use and schizophrenia. A meta-analysis of these studies provided strong support for the hypothesis that cannabis use, especially early onset use, can precipitate schizophrenia in susceptible individuals.

Genetic factors appear to increase the risk that cannabis users will develop psychosis. It has been difficult to identify a specific genetic diathesis in the majority of cases of schizophrenia, since 81 percent of persons with schizophrenia will not have a first-degree relative with the disorder, and 63 percent will not have an affected first- or second-degree relative. Recent GWAS studies suggest that very large numbers of genes with small effects may be implicated in schizophrenia.

It seems most likely that cannabis use can precipitate schizophreniform disorders in vulnerable individuals, but it is uncertain whether cannabis use can produce psychotic disorders that would not otherwise have occurred. Most of the 274 Swedish conscripts who developed schizophrenia had not used cannabis, and only 7 percent of cases of schizophrenia could be attributed to cannabis use. Trends in the *treated* incidence of schizophrenia, particularly of early onset, acute cases, during the 1970s and 1980s when cannabis use increased among young adults in the United States and Australia, have been unclear. Recent studies in Britain and Switzerland have suggested that the incidence of psychoses may have increased in recent birth cohorts with high rates of cannabis use.

Cannabis-Induced Anxiety Disorder

Some users report increased anxiety, panic, a fear of going mad, and depression after using cannabis. These experiences are most often reported by users who are unfamiliar with the effects of the drug. This has included some patients who have been given THC for therapeutic reasons. More experienced recreational users may report these effects if they use more potent forms of cannabis than usual, or if they use the oral route, whereby the effects of cannabis may be more difficult to titrate, more pronounced, and of longer duration than usual.

The most immediate effect of smoking cannabis is to increase heart rate by 20 to 50 percent within a few minutes to a quarter of an hour of smoking cannabis. Changes in blood pressure that depend on posture also occur: Blood pressure is increased while the person is sitting and decreases on standing. A sudden change from lying down to standing up may produce postural hypotension and feelings of light-headedness and faintness that are often the earliest indications of drug effects in naive users. The cardiovascular risks of cannabis smoking are probably highest in older adults, but younger adults, possibly those with undiagnosed cardiovascular disease, may be at risk. A French study, for example, found that 9.5 percent of 200 cannabis-related hospitalizations in the Toulouse area between January 2004 and December 2007 were for cardiovascular disorders. This included several cases of myocardial infarction and a fatal stroke in young adults who had no other risk factors for these disorders.

Withdrawal and Tolerance

Regular human users and chronically dosed animals develop tolerance to many of the behavioral and physiological effects of THC. The precise mechanisms are unknown, but they probably involve changes in cannabinoid receptor functioning.

Early case reports of cannabis withdrawal symptoms in humans have been supported by controlled studies of abstinence symptoms in regular users observed under laboratory conditions. Studies in clinical and nonclinical samples of long-term cannabis users have reported withdrawal symptoms that include anxiety, insomnia, appetite disturbance, and depression.

In laboratory studies, regular cannabis users who have been abruptly withdrawn from cannabis after 2 weeks on high doses of oral THC have complained within 6 hours of inner unrest, and after 12 hours have reported irritability, insomnia, and restlessness. These symptoms were correlated with THC dose and frequency of use and they were relieved by using cannabis. Similar symptoms have been reported during the first week of abstinence in subjects who had received 210 mg of smoked cannabis a day for 4 weeks. Recent laboratory studies have reported withdrawal symptoms at much lower doses of THC given orally and by smoking. The most common symptoms were anxiety, depression, and irritability.

A controlled prospective study has been completed on withdrawal symptoms among chronic cannabis users who were assessed daily on various withdrawal symptoms while in a hospital ward for 28 days. During the course of the 28 days the chronic cannabis users showed decreases in mood and appetite and increases in irritability, anxiety, physical tension, and physical symptoms, and increased scores on the Hamilton Anxiety and Depression scales. These symptoms appeared within 24 hours and were most marked in the first 10 days, although the increase in irritability and

physical tension persisted throughout the 28-day observation period. A recent double-blind trial of a cannabis extract, Sativex, in cannabis-dependent persons showed that these withdrawal symptoms could be relieved by a cannabis extract that contained equal amounts of THC and CBD.

Research using the cannabinoid antagonist SR141716A (which immediately reverses the effects of THC) has produced a withdrawal syndrome in rats, mice, and dogs that have been maintained on regular doses of THC. This antagonist produces compressed and accentuated withdrawal symptoms that are much more pronounced than the milder, more prolonged symptoms that occur after cessation in human users. The relatively long half-life and complex metabolism of cannabis probably explains the fact that the cannabis withdrawal syndrome is less intense than that for alcohol or the opiates.

Disorders Not Otherwise Specified

Amotivational Syndrome. The evidence that chronic heavy cannabis use produces an "amotivational syndrome" consists largely of case studies. Controlled field studies in the 1970s did not support such a syndrome, although their value was limited by the small sample sizes and restricted sociodemographic characteristics of the participants. Laboratory studies of sustained heavy daily use did not support the syndrome although these typically studied short periods of drug use in healthy youth and made minimal demands on volunteers. If there is such a syndrome, it is most common among heavy, chronic cannabis users and may be better described as due to the effects of chronic intoxication in dependent cannabis users.

Cognitive Impairment. The fact that cannabis use acutely impairs cognitive functioning has raised the reasonable concern that its chronic use may chronically impair cognition. There is clinical, electrophysiological and recently epidemiological evidence that long-term cannabis use produces impairment in the higher cognitive functions of memory, attention, organization, and integration of complex information. Consistently found deficits in verbal learning, memory, and attention in regular cannabis users have generally but not always been related to the duration and frequency of cannabis use, the age of initiation, and the estimated cumulative dose of THC received. It remains unclear whether cognitive function fully recovers after cessation of cannabis use.

A longitudinal study from the Dunedin birth cohort has suggested that sustained heavy cannabis use over several decades can produce substantial declines in cognitive performance that may not be wholly reversible. This study assessed changes in IQ between age 13 (before cannabis was used) and age 38 in 1,037 New Zealanders born in 1972 or 1973. It found that early and persistent cannabis users who used throughout adulthood

showed an average decline of eight-points of IQ, compared with those who had not used cannabis at all, and cannabis users who had not used cannabis in this sustained way. It is still unclear to what extent these impairments are reversed after an extended period of abstinence from cannabis.

Effects on Adolescent Development. Longitudinal studies of adolescents indicate that chronic heavy cannabis use may adversely affect adolescent development in a number of ways. Interpretation of this evidence has been complicated by the fact that many of the indicators of adverse developmental outcomes that have been attributed to cannabis use also precede its use and make it more likely that a young person will use cannabis. These include minor delinquency, polydrug use, poor educational performance, nonconformity, and poor social adjustment.

Among US adolescents in the 1970s and 1980s, the typical sequence of initiation into drug use was that the use of alcohol and tobacco preceded the use of cannabis, which, in turn, preceded the use of hallucinogens, amphetamine, and the later use of heroin and cocaine. Generally, the earlier the age of first use of any drug, and the greater the involvement with any drug in the sequence, the more likely a young person was to use the next drug in the sequence.

The explanation of the role cannabis plays in this sequence remains contested. The evidence for the hypothesis that cannabis use has a pharmacological effect that increases the risk of using later drugs in the sequence is not the most compelling. Equally plausible hypotheses are that the sequence of drug involvement reflects a combination of (a) the early recruitment into cannabis use of nonconforming and deviant adolescents who are more likely to use alcohol, tobacco, and illicit drugs; (b) a shared genetic vulnerability to become dependent on a range of substances; and (c) socialization of cannabis users in an illicit drug-using subculture that increases the exposure, opportunity, and encouragement to use other illicit drugs. Recent prospective studies that have controlled for many of these factors and studies of twin pairs discordant for cannabis use have failed to eliminate the apparent “gateway effect” of cannabis. This leaves open the possibility of some form of pharmacological sensitization in addition to selection and socialization.

In surveys of adolescents, cannabis use is associated with failing to complete a high school education and with job instability in young adulthood. The complication is that those who are most likely to use cannabis have lower academic aspirations and poorer school performance *before* they begin to use cannabis than do those who do not use. When these differences are taken into account, the relationship between cannabis use and educational and occupational performance is more modest. Even so, the effects of cannabis and other drug use on educational performance

are important, because impaired educational performance and underachievement limits later choice of occupation, level of income, choice of mate, and quality of adult life.

There is also suggestive evidence that heavy cannabis use has adverse effects on family formation and leads to involvement in drug-related (but not other types of) crime. However, the strong associations revealed in cross-sectional data are more modest in longitudinal studies that control for associations between cannabis use and other variables that predict these adverse outcomes.

Flashbacks. There are case reports of cannabis flashbacks in which users experience symptoms of cannabis intoxication days or weeks after they last used cannabis. Because of their rarity, and the fact that many affected individuals have also used other drugs, it has been difficult to decide whether these are (a) rare events that coincide with cannabis use; (b) the effects of other drugs that are often taken together with cannabis; (c) rare consequences of cannabis use that only occur at doses that are much higher than those used recreationally; (d) the result of unusual forms of personal vulnerability; or (e) the results of interactions between cannabis and other drugs.

ADVERSE EFFECTS OF CANNABIS USE

Psychomotor Effects and Driving

Cannabis intoxication impairs a wide range of cognitive and behavioral functions that are involved in driving automobiles or operating machinery. The effects are generally larger, more consistent, and more persistent in tasks that require sustained attention. Recreational doses of THC produce similar impairments in performance on laboratory tests and standardized driving courses to blood alcohol concentrations (BACs) of between 0.07 and 0.10 percent. Studies of the effect of cannabis use on driving performance on driving courses and on the road have found only modest impairments, probably because cannabis-intoxicated drivers drive more slowly, and take fewer risks, than alcohol-intoxicated drinkers. Cannabis users seem to be more aware of their psychomotor impairment than alcohol users.

Recently, better controlled epidemiological studies have found modest associations between self-reported cannabis use and hospitalization for accidental injury, which have persisted after controlling for alcohol and tobacco use in men (relative risk [RR] = ~2). A recent case-control study in France also found a higher rate of THC in the blood among persons involved in accidents (overall risk = 2.5).

A meta-analysis of the epidemiological studies suggests that cannabis use increases the risk of motor vehicle crashes around two to three times.

Studies that have measured THC in blood (that reflects intoxication rather than inactive metabolites that reflect recent use) have found a dose–response relationship between THC in blood and the risk of accident. The combination of THC and alcohol produces more marked impairment and increased accident risk. The size of the risk of accidents among cannabis users is more modest than that for alcohol, but it is of potential public health significance given the high prevalence of cannabis use among young males.

The Cardiovascular System

The acute toxicity of cannabis and cannabinoids is very low. There are no cases of fatal cannabis poisoning in the human medical literature. Animal studies indicate that the dose of THC required to produce 50 percent mortality in rodents is extremely high by comparison with other pharmaceutical and recreational drugs. The lethal dose also increases as one moves up the phylogenetic tree, suggesting that the lethal dose in humans could not be achieved by smoking or swallowing cannabis.

Within 3 to 15 minutes after cannabis is smoked or swallowed, THC increases heart rate by 20 to 50 percent for up to 3 hours. Blood pressure also increases while the person is sitting and decreases on standing. In healthy young users these cardiovascular effects are unlikely to be of any clinical significance because tolerance develops to the effects of THC, and young, healthy hearts will only be mildly stressed. These changes in heart rate and blood pressure may be less benign in older adults with hypertension, cerebrovascular disease, and coronary atherosclerosis. Several epidemiological studies (including a case-crossover study of the relationship between cannabis use and myocardial infarction) suggest that cannabis use may provoke a myocardial infarction in middle-aged adults with underlying heart disease. The public health implications of these effects may increase as the cohort of continuing cannabis users from the late 1970s and early 1980s enters old age, the period of maximum risk for atherosclerosis in the heart, brain, and peripheral blood vessels.

The Respiratory System

Regular cannabis smoking impairs the functioning of the large airways and causes symptoms of chronic bronchitis such as coughing, sputum, and wheezing. It is likely that chronic cannabis smoking increases the risks of respiratory cancer because tobacco and cannabis smoke contain similar carcinogenic substances and tobacco smoke has adverse effects on the respiratory system. There is old evidence that chronic cannabis smoking produces histopathological changes in lungs of the type that precede the development of lung cancer. Concern about the possibility of cancers caused by chronic cannabis smoking has been raised by case reports of

cancers of the aerodigestive tract in young adults with a history of heavy cannabis use. Case-control studies of aerodigestive tract cancers and cannabis smoking have produced conflicting and unclear results because most cannabis smokers also smoke tobacco. More research is needed to clarify the cancer risks of long-term cannabis smoking. A recent meta-analysis of these studies concluded that there was at most a weak association between regular cannabis smoking and lung cancer.

Cellular Effects and Cancers

There is very weak evidence that THC alters cell metabolism and deoxyribonucleic acid (DNA) synthesis in the test tube. There is much stronger evidence that cannabis *smoke* produces mutations in cells in the test tube and in live animals and hence is a potential cause of cancer. Cannabis smoke contains many of the same carcinogenic substances as cigarette smoke, and hence if cannabis smoking causes cancer it is most likely to be cancers of the lung and upper aerodigestive tract.

Aerodigestive tract cancers have been reported among young adults who have been daily cannabis users, and a case-control study has found an association between cannabis smoking and head and neck cancer. A prospective cohort study of 64,000 adults did not find an increased incidence of head and neck or respiratory cancers, but it found increased rates of prostate cancer. The relative youth of the participants and their low rates of regular cannabis use may have reduced the ability of this research to detect an increase in respiratory cancers. Further studies are needed to clarify the issue.

There is much weaker evidence for an increased risk of cancers among children born to women who smoked cannabis during pregnancy. Three studies of very different types of cancer have reported an association with maternal cannabis use. None of these was a planned study of the role of cannabis use in these cancers, so replication of their results is required. There have not been any increases in the rates of these cancers that parallel increased rates of cannabis use over the past three decades.

There is more cause for concern about recent reports of an increased risk of testicular cancer among cannabis users. This finding is of more concern because a case control study that found a doubling of the risk of nonseminoma testicular cancer has been replicated in two case-control studies. Both replications found a doubling of risk for the same nonseminoma testicular tumor and evidence that risk increased with earlier initiation and more frequent use of cannabis. The effect is also biologically plausible since cannabinoid receptors are found in the male reproductive system.

Immunological Effects

Cannabinoids impair cell-mediated and humoral immunity in rodents and reduce resistance to infection by bacteria and viruses in animals. Cannabinoid receptors are expressed in cells of the immune system in animals and humans, although their clinical significance is unclear. Cannabis smoke also impairs the functioning of alveolar macrophages, the first line of the body's immune defense system in the lungs. The clinical relevance of these findings is uncertain because the doses required to produce these effects have been very high, and extrapolation to the doses used by humans is complicated by the fact that tolerance may develop to these effects.

The limited experimental and clinical evidence in humans suggests that the adverse effects seen in animals are not replicated in humans. There is no conclusive evidence that cannabinoids impair immune system function in humans, as measured by T-lymphocytes, B-lymphocytes, macrophages, or immunoglobulin levels. There is suggestive evidence that THC impairs T-lymphocyte responses to mitogens and allogenic lymphocytes. The clinical and biological significance of these possible effects in chronic cannabis users is uncertain because there is as yet no epidemiological evidence of increased rates of infectious disease among chronic heavy cannabis users.

Reproductive Effects

Chronic administration of THC disrupts male and female reproductive systems in animals, reducing testosterone secretion and sperm production, motility, and viability in males, and disrupting the ovulatory cycle in females. It is uncertain whether cannabis use has these effects in humans because of the inconsistency in the limited literature on human males and the lack of research in the case of human females. There is uncertainty about the clinical significance of these effects in normal healthy young adults.

It is likely that cannabis use during pregnancy produces babies with smaller birth weights, perhaps as a consequence of shorter gestation, and probably by the same mechanism as cigarette smoking. There is no clear evidence that cannabis use during pregnancy increases the risk of birth defects due to the fetus's exposure to cannabis in the uterus. There is some evidence that infants exposed to cannabis in the uterus may show transient behavioral and developmental effects during the first few months after birth. These effects are small by comparison with those caused by tobacco use during pregnancy and have not been observed in all studies.

LABORATORY EXAMINATIONS

Cannabinoids can be detected in the head hair, pubic hair, urine, sweat, saliva, and blood of users. Cannabinoids are stored in the fat cells of the

body and so can remain in the body for an extended period compared to other drugs. In some cases, THC may be detectable in urine for up to 11 weeks after use. Cannabinoids can be detected in hair, and some research has suggested that higher concentrations of cannabinoids may be found in pubic than head hair. Cannabinoids may be detected in saliva and sweat, but the concentration of cannabinoids in these fluids tends to be lower compared to urine, and in some cases, cannabinoids may not be detected in these fluids.

Cannabinoid levels in the blood vary between individuals and depend on the dose received and the individual's history of cannabis use. Blood levels of THC may range between 0 and 500 ng/mL. The level depends on the potency of the cannabis and the time since it was smoked. Levels of THC in blood above 10 to 15 ng/mL indicate recent use, although it is difficult to be precise about how recent. A more precise estimate of time since last use is provided by the ratio of THC to 9-carboxy-THC. When the blood concentrations of THC and 9-carboxy-THC are similar, it indicates that cannabis has been used in the past 20 to 40 minutes, and hence suggests a high probability of intoxication. This relationship is less clear in regular users.

Cannabis intoxication impairs skills required to drive a motor vehicle, so it would be desirable to have a measure of cannabis intoxication similar to the breath test for alcohol intoxication. The major obstacle is the lack of a simple relationship between THC metabolites in blood and degree of psychomotor impairment. Recent experimentation suggests that saliva tests for THC may provide evidence of recent use that can be confirmed by analyses of THC levels in blood. Consensus is yet to be reached on the definition of a level of THC in blood that provides per se evidence of impairment analogous to the use of BACs of 0.05 or 0.08 percent. Some Australian states have nonetheless introduced roadside salvia testing for recent use of cannabis and other illicit drugs. The US states of Colorado and Washington have defined driving with a level of THC in blood greater than 5 ng/mL as cannabis-impaired driving.

With repeated frequent use of cannabis, THC accumulates in fatty tissues in the human body, where it may remain for considerable periods of time. The health significance of this storage is unclear. The storage of cannabinoids *would* be serious cause for concern if THC were a highly toxic substance that remained physiologically active while stored in body fat. THC is not, however, a highly toxic substance, and it is in fact inactive while it is stored in fat.

TREATMENT

Cannabis Dependence

There have been a number of randomized controlled trials comparing

different ways of assisting cannabis-dependent persons to stop using cannabis. The earliest study reported that at 1 month after treatment only 30 percent of their patients were still abstinent, and by the end of 1 year only 17 percent remained abstinent.

Later studies compared group relapse prevention intervention, individualized advice, and motivational interviewing adapted from Miller's Drinker's Check-up with a delayed treatment condition, in which participants did not receive any treatment for 4 months. At the 4-month follow-up, all three groups had reduced their cannabis use, but the two treatment groups showed the largest reduction and did not differ from each other. In the treatment groups, 37 percent were abstinent, compared with only 9 percent in the delayed treatment group. The level of cannabis use also declined by 70 percent in the treatment groups and by 30 percent in the delayed treatment groups. Abstinence rates declined over time, but the two treatments did not differ at 7, 13, and 16 months after treatment. Twenty-two percent of participants were abstinent throughout the 16-month study, and their partners and family members corroborated their abstinence.

There may be a role for specialist treatment for cannabis dependence in problem users in whom self-help and brief interventions fail to produce cessation. These clinical trials have now established that cannabis dependence can be treated on an outpatient basis using psychosocial approaches such as cognitive-behavioral therapy that substantially reduce rates of cannabis use and cannabis-related problems, even if the proportion who achieve enduring abstinence is modest, much like that in the treatment of problem alcohol use. Much more research is needed before advice can be given about the best ways to achieve abstinence from cannabis. In the absence of better evidence of treatment effectiveness, people offering treatment for cannabis dependence should avoid routine use of inpatient treatment, given the absence of any evidence that it is more effective than outpatient forms of treatment.

There is increasing interest in the use of antidepressants to treat dependent cannabis use because of the high rates of depression reported on presentation for treatment and after cessation. Small studies have been conducted to examine the effectiveness of such a treatment, but no large randomized controlled studies have been conducted to date. A controlled trial found that Sativex (GW Pharmaceuticals, Salisbury, UK) reduces the severity of withdrawal symptoms more than placebo but there was no difference in the proportions achieving abstinence at 6 months.

Cannabis Intoxication

The adverse acute effects of cannabis (such as anxiety symptoms) can be prevented by advising users about the cardiovascular effects they may experience. If these symptoms develop, they can usually be managed by

reassurance and support.

Therapeutic Effects of Cannabinoids

When cannabinoids and cannabis are advocated for medical uses, it is primarily to relieve symptoms. Cannabis is most often advocated for the symptomatic relief of nausea and vomiting caused by cancer chemotherapy, appetite loss in acquired immunodeficiency syndrome (AIDS), and to relieve muscle spasticity and chronic neuropathic pain in neurological disorders such as MS.

Analgesia. Animal studies and the biology of cannabinoid receptors suggest that cannabinoids may be useful analgesics with mild to moderate efficacy. The few controlled studies in humans have suggested that THC and other cannabinoids have modest analgesic effects on acute postoperative and chronic pain including neuropathic pain (compared to placebo). The efficacy of typical doses of cannabinoids was equivalent to that of 60 mg of codeine. Since patients often report adverse psychotropic effects, their use for these purposes is likely to be limited. The development of synthetic cannabinoids with fewer psychotropic effects may be a more promising way to proceed than using THC or cannabis products.

Nausea and Vomiting. Most research on the antiemetic effects of cannabinoids in patients receiving cancer chemotherapy was done in the 1980s using THC, nabilone (Cesamet), and levonantradol. Many of these studies were small and not well controlled. They showed *some* antiemetic efficacy by comparison with the relatively ineffective antiemetic agents that were then available, namely, prochlorperazine (Compazine). Clinical use of cannabinoids has declined with the advent of selective serotonin type 3 receptor agonists, such as ondansetron (Zofran), that dramatically reduce nausea and vomiting. These drugs provide complete control over nausea induced by cisplatin (Platinol) in 75 percent of cases and up to 90 percent for less emetogenic forms of chemotherapy, whereas THC provides control in only a third of patients.

Wasting Syndrome and Appetite Stimulation in Human Immunodeficiency Virus (HIV)/AIDS. THC stimulates appetite and assists in weight gain in HIV/AIDS patients in short-term trials. It has been registered for medical use for this purpose in the United States. Some patients do not like dronabinol (Marinol) because of its psychoactive side effects, the difficulty of titrating the dose, the delayed onset, and the prolonged duration of its effects. There are anecdotal reports that smoked cannabis is also effective in the treatment of HIV/AIDS-associated anorexia and weight loss, but there have not been any controlled studies. The advent of highly effective antiretroviral therapy means that very few HIV-infected persons now

develop AIDS.

Muscle Spasticity. Muscle spasticity is the increased resistance to passive stretching of muscles and increased deep tendon muscles. Involuntary contractions may occur, which can be painful and debilitating. About 90 percent of multiple sclerosis (MS) patients eventually develop muscle spasticity, as do a substantial proportion of patients with spinal cord injuries. Randomized placebo-controlled clinical trials of Sativex (GW Pharmaceuticals, Salisbury, UK), a cannabis extract, in treating muscle spasticity in MS have led to its being approved for this use in Canada, Europe, and the United Kingdom.

Movement Disorders. Movement disorders are caused by abnormalities in areas of the brain that are connected to the motor cortex, resulting in abnormal skeletal muscle movements in the face, limbs, and trunk. The movement disorders most often mentioned as candidates for medical use of cannabis are dystonia, Huntington disease, Parkinson disease, and Tourette syndrome. There is limited evidence that cannabis is useful for treating any of these disorders. The health risks of the regular cannabis smoking that would be required in persons already suffering from these neurological conditions are a major limitation on its use in these disorders.

Epilepsy. There are case reports suggesting that cannabis can control epileptic seizures, and one observational study that suggested that cannabis use was protective against seizures, but it has major limitations. Most of the anticonvulsant properties of cannabinoids appear to be attributable to CBD rather than THC. Since CBD (which has no psychoactive effects) is not a controlled substance, there are no obstacles to its clinical use if it is shown to be safe and effective in controlled clinical trials.

Glaucoma. Elevated intraocular pressure (IOP) is a chronic condition that produces blindness if untreated. IOP must be controlled continuously to reduce the risk of blindness. Cannabis and THC taken orally or intravenously reduce IOP by 25 percent, but the effect lasts only 3 to 4 hours. The high doses of THC required to produce these effects also produce adverse side effects that preclude the lifelong use of cannabis or cannabinoids to treat glaucoma. Water-soluble cannabinoids that may be topically applied and have no psychoactive effects may be a better prospect.

THC and other cannabinoids have not been widely used therapeutically or investigated in clinical trials. This is because in the United States, clinical research on cannabinoids has been discouraged by regulation and the fact that THC, the most therapeutically effective cannabinoid, is the one that produces the psychoactive effects sought by recreational users. The

discovery of a cannabinoid receptor and the cannabinoid-like substance anandamide may encourage more basic research into the therapeutic uses of natural and synthetic cannabinoids.

References

Anthony JC, Warner L, Kessler R. Comparative epidemiology of dependence on tobacco, alcohol, controlled substances and inhalants: basic findings from the National Comorbidity Survey. *Exp Clin Psychopharmacol*. 1994;2:244.

Asbridge M, Hayden JA, Cartwright JL. Acute cannabis consumption and motor vehicle collision risk: systematic review of observational studies and meta-analysis. *BMJ*. 2012;344:e536.

Bachman JG, Wadsworth KN, O'Malley PM, Johnston LD, Schulenberg JE. *Smoking, Drinking, and Drug Use in Young Adulthood: The Impacts of New Freedoms and New Responsibilities*. Hove: Psychology Press; 2013.

Degenhardt L, Hall W, Lynskey M, Coffey C, Patton G. The association between cannabis use and depression: a review of the evidence. In: Castle D, Murray RM, D'Souza DC, eds. *Marijuana and Madness*. 2nd ed. Cambridge, UK: Cambridge University Press; 2012:114–128.

Denis C, Lavie E, Fatseas M, Auriacombe M. Psychotherapeutic interventions for cannabis abuse and/or dependence in outpatient settings. *Cochrane Database Syst Rev*. 2006;(3):CD005336.

Farrell M, Buchbinder R, Hall W. Should doctors prescribe cannabinoids? *BMJ*. 2014;348:g2737.

Fergusson DM, Horwood LJ, Ridder EM. Tests of causal linkages between cannabis use and psychotic symptoms. *Addiction*. 2005;100:354.

Freisthler B, Gruenewald PJ. Examining the relationship between the physical availability of medical marijuana and marijuana use across fifty California cities. *Drug Alcohol Depend*. 2014;143:244–250.

Hall WD. The mental health risks of adolescent cannabis use. *PLoS Med*. 2006;3:159.

Hall W, Degenhardt L. The adverse health effects of nonmedical cannabis use. *Lancet*. 2009;374:1383–1391.

*Hall W, Degenhardt L. What are the policy implications of the evidence on cannabis and psychosis? In: Castle D, Murray RM, D'Souza DC, eds. *Marijuana and Madness*. 2nd ed. Cambridge, UK: Cambridge University Press; 2012:55–65.

*Hall WD, Pacula RL. *Cannabis Use and Dependence: Public Health and Public Policy*. Cambridge, UK: Cambridge University Press; 2003.

*Institute of Medicine. *Marijuana and Medicine: Assessing the Science Base*. Washington, DC: National Academy Press; 1999.

Jouanjus E, Lapeyre-Mestre M, Micallef J. Cannabis use: signal of increasing risk of serious cardiovascular disorders. *J Am Heart Assoc*. 2014;3(2):e000638.

Kedzior KK, Laeber LT. A positive association between anxiety disorders and cannabis use or cannabis use disorders in the general population–-a meta-analysis of 31 studies. *BMC Psychiatry*. 2014;14:136.

Keyes KM, Schulenberg JE, O'Malley PM, et al. The social norms of birth cohorts and adolescent marijuana use in the United States, 1976–2007. *Addiction*. 2011;106(10):1790–800.

Marshall K, Gowing L, Ali R, Le Foll B. *Pharmacotherapies for Cannabis Dependence: Cochrane Database of Systematic Reviews*. Hoboken, NJ: John Wiley & Sons, Ltd; 2014.

Mittleman MA, Lewis RA, Maclure M, Sherwood JB, Muller JE. Triggering myocardial infarction by marijuana. *Circulation*. 2001;103:2805.

Moore TH, Zammit S, Lingford-Hughes A, Barnes TR, Jones PB. Cannabis use and risk of psychotic or affective mental health outcomes: a systematic review. *Lancet*. 2007;370:319.

Pertwee RG, Howlett AC, Abood ME, et al. International Union of Basic and Clinical Pharmacology. LXXIX. Cannabinoid receptors and their ligands: beyond CB(1) and CB(2). *Pharmacol Rev*. 2010;62(4):588–631.

Quickfall J, Crockford D. Brain neuroimaging in cannabis use: a review. *J Neuropsychiatry Clin Neurosci*. 2006;18:318.

*Roffman RA, Stephens RS. *Cannabis Dependence: Its Nature, Consequences and Treatment*. Cambridge, UK: Cambridge University Press; 2006.

Solowij N, Battisti R. The chronic effects of cannabis on memory in humans: a review. *Curr Drug Abuse Rev*. 2008;1(1):81–98.

▲ 11.6 Hallucinogen-Related Disorders

MICHAEL P. BOGENSCHUTZ, M.D., AND STEPHEN ROSS, M.D.

INTRODUCTION

Classes of Drugs Considered to Be Hallucinogens

The hallucinogens comprise several classes of substances which are so diverse that it is challenging to provide an adequate definition of the term "hallucinogen." For purposes of this chapter, the term will be used to refer to substances whose subjective effects are dominated by prominent alterations in perception, cognition, affect, sense of meaning, and/or sense of self. Hallucinogen intoxication differs from delirium in that orientation and memory are relatively spared, although this distinction is not absolute. Most of the commonly abused substances of abuse (e.g., cannabis, opioids, and stimulants) can have effects that overlap with those of hallucinogens, but other effects are much more prominent. Many prescribed medications can also cause hallucinations as an unusual side effect or a symptom of medication-induced delirium.

The term "hallucinogen" is a misnomer in that true hallucinations are atypical for most of the commonly used hallucinogens. Many other labels have been proposed. The term "psychotomimetic" has been used to stress the similarity between hallucinogen intoxication and psychotic illness. Indeed both arylcyclohexamines and serotonergic hallucinogens such as lysergic acid diethylamide (LSD) have provided useful models of psychosis, approximating the negative and positive symptoms of schizophrenia, respectively. However, hallucinogens have important effects that are not typical of psychotic states. The term "psychedelic" (meaning "mind-manifesting") was coined by Humphrey Osmond to capture the altered

state of consciousness produced by classic serotonergic hallucinogens such as LSD. PCP and related arylcyclohexamines (such as ketamine) are also termed dissociatives or dissociative anesthetics. The terms "empathogen" and "entactogen" have been proposed for 3,4-methylenedioxymethamphetamine (MDMA) and related compounds whose effects are dominated by feelings of emotional closeness to other people. The term "entheogen" has been used for hallucinogens when the spiritual or religious aspect of the experience is considered to be of primary importance. The DSM-5 recognizes two categories of hallucinogen: PCP and related arylcyclohexamines, and "other hallucinogens." However, there are several distinct and sometimes overlapping classes of "other hallucinogens." Some of these drugs have been safely used by people for thousands of years. Others are newly synthesized (new hallucinogens are constantly being described), and their effects in humans are entirely unknown. Because of the great diversity of drugs under the hallucinogen label, each of the main classes of hallucinogen will be discussed briefly before moving on to describe the hallucinogen-related disorders. Characteristics of the classes of hallucinogens are summarized in Table 11.6–1. Chemical structures of representative hallucinogens in each class are illustrated in Figure 11.6–1.

NMDA Receptor Antagonists

The DSM-5 recognizes a class of PCP-type substance explicitly including PCP, ketamine, cyclohexamine, and dizocilpine. The primary effects of drugs of this class are mediated by noncompetitive *N*-methyl-D-aspartate (NMDA) receptor antagonism. Other drugs that share this mechanism of action include dextromethorphan and nitrous oxide (the latter usually being categorized as an inhalant because it is a gas). PCP and ketamine bind to PCP binding sites on the NMDA receptor and block the ionophore channel, decreasing influx of calcium and sodium ions. The highest densities of PCP receptors are found in the dentate gyrus, hippocampus, and anterior forebrain. Downstream effects of inhibition of NMDA receptor-mediated glutamatergic neurotransmission include a secondary increase in glutamate release in the medial prefrontal cortex and an increase in pyramidal cell firing, probably due to blockade of NMDA receptors on GABAergic inhibitory interneurons. In rats, noncompetitive NMDA receptor antagonists also increase extracellular prefrontal and mesolimbic dopamine levels, probably by stimulating glutamate release in the ventral tegmental area and dorsal raphe nuclei. The effects on dopamine release are thought to mediate the reinforcing effects of these drugs, as well as hyperactivity and stereotypic movements. At very high doses, PCP blocks monoamine transporters, but ketamine does not—possibly accounting for greater stimulation and reward with PCP than

ketamine. At even higher doses, PCP blocks sodium and potassium channels, which may contribute to the incidence of seizures following PCP overdose. In animal models, behavioral effects of NMDA antagonists are attenuated by agents that enhance NMDA receptor-mediated neurotransmission, such as glycine and D-serine. Some effects are also attenuated by agents that block D2 receptors and 5HT2A receptors.

Effects of PCP and Related Drugs. These drugs are reinforcing in animal models and can produce compulsive use in humans, although not all people like their effects. These drugs are sometimes termed "dissociatives" or "dissociative anesthetics." A prominent feature of intoxication is a subjective sense of separation of mind from body, accompanied by anesthesia. The drugs can also produce a full range of hallucinogen effects including visual illusions or true hallucinations, other sensory changes, altered thought process, labile or inappropriate affect, anxiety, and paranoia or other delusional ideation. PCP intoxication in particular is associated with unpredictable and potentially violent behavior. PCP intoxication is also associated with nystagmus and ataxia, hypertension and tachycardia, and muscular rigidity.

Table 11.6–1.
Classes of Hallucinogens

	Primary Mechanism	Characteristic Compounds (Typical Dose)	Characteristic Features of Intoxi cation	Dependence Liability	Ph De
PCP and related compounds	NMDA receptor antagonism	PCP, ketamine, dextromethorphan	Dissociative state, depersonalization, derealization, cognitive impairment	Moderate to high	Po
Classic hallucinogens	5HT2A receptor agonism/partial agonism	LSD, psilocybin, mescaline, DMT	Visual illusions, closed-eye visual images, mystical experience, anxiety.	Low	No
MDMA and related compounds	Serotonin release, reuptake inhibition	MDMA, MBDB	Feelings of love, acceptance, closeness to others	Moderate	No

Iboga alkaloids	Multiple mechanisms	Ibogaine	Very long-acting, can be very intense (complete break with reality)	Low	No
Kappa opioid agonists	Kappa opioid receptor agonism	Salvinorin A (Salvia divinorum)	Very brief, can be very intense (complete break with reality)	Low	No
Anticholinergic deliriants	Muscarinic receptor antagonism	Datura spp., scopolamine, diphenhydramine	True hallucinations possible	Low	Po

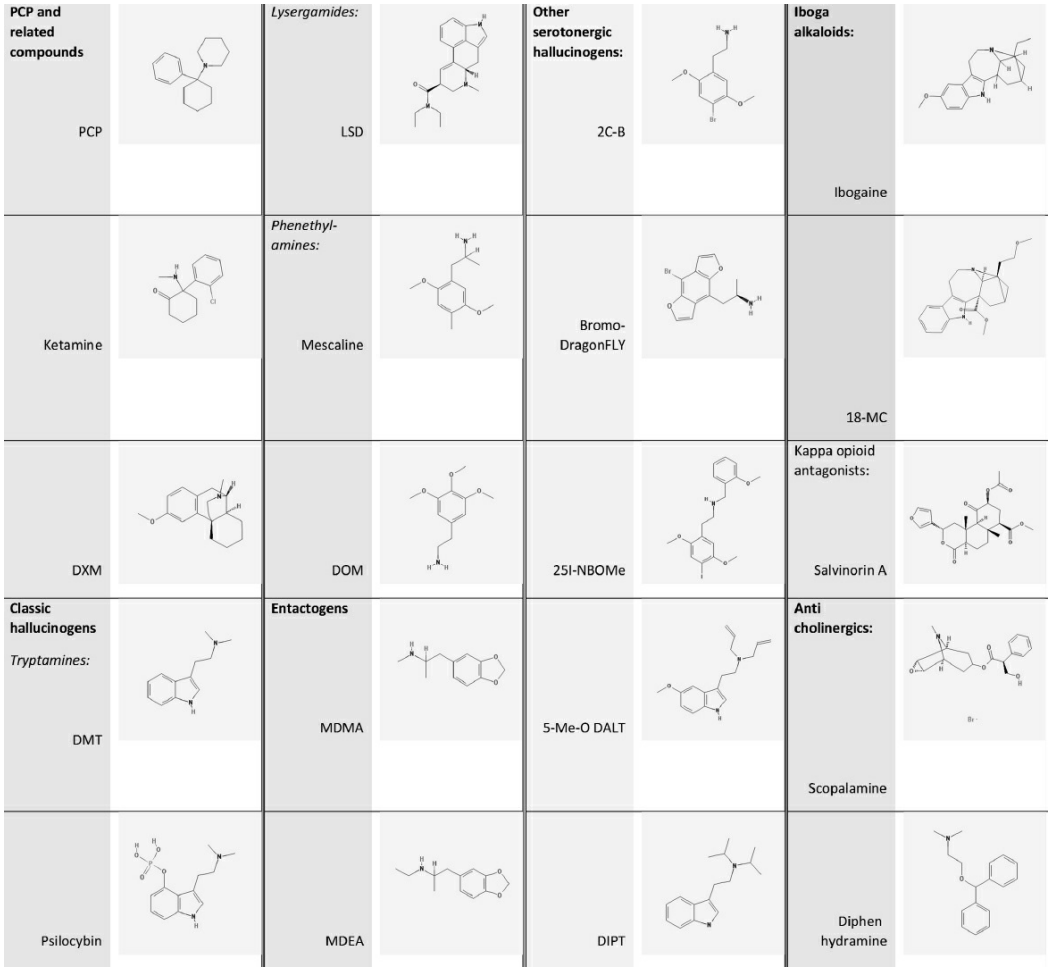

FIGURE 11.6–1. Chemical structures or representative hallucinogens of each class.

The duration of acute effects of these drugs depends on the specific drug, the dose, and the route of administration. The serum half-life of PCP in humans is highly variable, averaging about 20 hours. Ketamine has a considerably shorter half-life, about 2 hours. PCP is highly lipid soluble and accumulates in fat. Metabolism is primarily hepatic for both drugs.

PCP and ketamine produce compulsive use in animal models. Monkeys given free access to PCP will self-administer to maintain near-constant intoxication. Patterns of compulsive use of PCP and ketamine are seen in

humans as well. PCP and related drugs also lower the threshold for brain self-stimulation reward, a characteristic shared with most other addictive drugs. Pronounced tolerance has been demonstrated with PCP in animal models, and with ketamine in humans. A pronounced withdrawal syndrome is seen in primates, including vocalizations, bruxism, oculomotor hyperactivity, diarrhea, piloerection, somnolence, tremor, and seizures.

NMDA Receptor Antagonist Model of Psychosis. The use of PCP intoxication as a model for psychosis was proposed in 1959. Although PCP is no longer used in human studies due to its toxicity, ketamine is still used extensively in the study of psychosis, particularly in relation to the role of altered glutamatergic neurotransmission in the pathogenesis of psychosis. The NMDA antagonist model of psychosis more closely mimics symptoms of schizophrenia than other drug models (e.g., amphetamine or classic hallucinogen). Both PCP and ketamine induce symptoms of thought disorder including withdrawal, autism, negativism, inability to maintain cognitive set, and abnormal responses to projective testing. Catatonic posturing may occur. Disturbances in attention, perception, and symbolic thinking are also very similar to those seen in schizophrenia. Both drugs inhibit prepulse inhibition, replicating a well-established neurobiological marker of schizophrenia. Well-formed hallucinations and delusions are less common, and may be absent even in the presence of severe thought disorder. A single relatively low dose of PCP can trigger a full-blown psychotic episode in patients with schizophrenia.

Serotonergic Hallucinogens

There are several distinct and overlapping groups of naturally occurring and synthetic hallucinogens that act primarily by affecting serotonergic pathways but vary greatly in their structure, effects, and mechanisms of action. The "Other hallucinogen" class defined by DSM-5 includes principally these drugs, but also a variety of other types of hallucinogens as described below.

Classic Hallucinogens. The "classic hallucinogens" are those drugs whose effects are thought to be primarily mediated by agonism or partial agonism at the 5HT2A receptor. The prototypical drug of this class is LSD. Other representative drugs of this large group of drugs are psilocin and its prodrug psilocybin (the primary active ingredients in hallucinogenic mushrooms), mescaline (the primary active ingredient of hallucinogenic cacti such as peyote and the San Pedro cactus), and dimethyltryptamine (DMT). Structurally, these drugs fall into three main categories. The tryptamine hallucinogens (including psilocybin/psilocin and DMT) are structurally related to serotonin (5-hydroxytryptamine), containing the indole nucleus and various substituents that affect receptor affinities and

drug metabolism. The ergoline class includes the indole structure within a much more complicated 4-ring molecule. LSD is the most commonly used member of this group. LDS was originally derived from ergot fungus that infects grain and causes ergotism or St. Anthony's fire, outbreaks of which occurred frequently in Europe in the Middle Ages. Other ergolines are found in several species of the morning glory family including *Rivea corymbosa*, *Ipomoea violacea*, and *Argyreia nervosa*. The third group of classic hallucinogens is the phenylalkylamines. The structure of these drugs is very similar to that of amphetamine, but some of the drugs in this class have classic hallucinogen effects virtually identical to those of LSD and are therefore thought to be mediated primarily by activity at the 5HT2A receptor. Mescaline and 2,5-dimethoxy-4-methylamphetamine (DOM, also known as STP) are representative classic hallucinogens of this group.

The classic hallucinogens have long provided evidence implicating the serotonergic system, as one of several neurotransmitter systems (the other main ones being the dopaminergic, glutamatergic, and endocannabinoid systems), in the pathophysiology of schizophrenia and related psychotic disorders. In the early 1920s, similarities were first described between mescaline intoxication and symptoms of schizophrenia. When Albert Hoffman accidentally discovered the psychoactive effects of LSD in 1943, he reported on the first ever LSD experience and described frightening psychotic-like positive symptoms, including thoughts that he was possessed by a demon and that his neighbor was a witch. This prompted research into studying LSD as a model to understand the biological basis of psychosis with several research groups reporting that LSD, mescaline, and psilocybin produced effects similar to the early phases of schizophrenia. In fact, when LSD was first marketed by Sandoz in the 1940s, it was recommended that psychiatrists and medical students take LSD as a way to gain insight into the nature of the schizophrenic experience. Some researchers speculated that schizophrenia might be caused by an excess production of an endogenous psychotogenic compound with a particular focus on DMT, with studies showing it produced psychotic like positive symptoms in normal volunteers and other studies reporting differences in urine and plasma DMT levels between patients with schizophrenia and controls. Although these findings were not consistently replicated, it was recently shown that DMT is produced in the pineal gland of rats and possibly that of humans as well.

CLASSIC HALLUCINOGEN EFFECTS. The effects of all of the classic hallucinogens are similar, differing principally in time course, which in turn depends on the particular substance, the dose, and the route of administration. An important feature of all classic hallucinogens is that the psychological content and emotional tone of the experience are unpredictable, but are thought to be influenced strongly by the mental state, preparation, and

intention of the person taking the drug (mental "set"), and the environment in which the effects are experienced ("setting") as well as the dose and the particular drug that is taken. Somatic effects are typically mild, including chills, tremor, unsteadiness, nausea with or without vomiting, xerostomia, paresthesias, and blurred vision. Pulse and blood pressure may be mildly to moderately elevated. Anorexia is usual. Significant physiological toxicity is not seen in the doses of LSD, psilocybin, mescaline, or DMT that are typically used. Sensory effects are often prominent, including alteration in perception of shape, size, and color, and the illusion of movement. With eyes closed, vivid imagery is often perceived, ranging from elementary geometric or fractal patterns to vivid representational images of all kinds. Other sense modalities may be involved, and synesthesia can occur. True hallucinations are uncommon. Affective effects are extremely variable, and can change rapidly and frequently during a single episode of use. Anxiety is common, as are strong cathartic emotional experiences, often related to past or current life experience. Feelings of bliss, joy, and peace may also occur. Depersonalization, derealization, and paranoia can occur. The sense of the passage of time may be distorted, most commonly in the sense that time seems to pass slowly or not at all. The content of the subjective experience may be dominated by personal psychological concerns (conflicts, relationships, grief and loss), symbolic representations of archetypal themes (heroic quests, interactions with mythological figures), or religious or spiritual matters. Particularly in high doses, mystical type experiences may occur, typically including an ineffable experience of unity and transcendence of space and time which is experienced as deeply positive, sacred, and meaningful.

Tolerance to LSD develops very rapidly, with significant decreases in effect of subsequent doses after a single dose. This effect may be due to down-regulation or internalization of 5HT2A receptors, and is completely reversed within a week of abstinence. DMT may be an exception to this effect, as marked tolerance is not observed. Cross-tolerance exists among LSD, mescaline, and psilocybin, further evidence of their shared mechanism of action.

DIFFERENCES AMONG CLASSIC HALLUCINOGENS. Classic hallucinogens differ principally in potency, duration of action, and traditional patterns of use. LSD is one of the most potent drugs known, with noticeable effects beginning at a dose of about 25 μg. Because of its high potency, LSD is often distributed and sold on small squares of blotter paper, which are held in the mouth to allow buccal absorption prior to swallowing. The onset of effects usually occurs within 60 minutes, with peak effects around 3 to 4 hours, and total duration of action 8 to 10 hours. A typical street "hit" (dose) of LSD is currently on the order of 50 μg, much lower than the doses of 100 to 250 μg that were typical in the 1960s, and in part accounting for

the lower incidence of adverse reactions in recent years. Recently, it has been common for newer, and less expensive, potent research chemicals to be found in material sold as LSD (see below).

Mescaline is among the least potent of hallucinogens, with a dose of 300 to 400 mg estimated to be roughly equivalent to 100 μg of LSD. The onset of action is somewhat slower with mescaline, and the duration of action is slightly longer (10 to 12 hours). Mescaline is primarily used in the form of peyote or other mescaline-containing cacti, though purified or synthetic mescaline is sometimes found. Peyote has been used by humans for at least 5,000 years, and is currently used sacramentally within the Native American Church and by tribes in northern Mexico including the Huichol. There is a long history of ritual use of the San Pedro cactus in South America as well. Mescaline, at least when consumed as cactus, is associated with more nausea and vomiting than LSD or psilocybin. The low potency of mescaline results in a lower index of safety than other classic hallucinogens. Autonomic hyperstimulation can become problematic in doses greater than 500 mg.

Psilocybin is intermediate in potency between LSD and mescaline, with a dose of 10 to 15 mg being roughly equivalent to 100 μg of LSD. Psilocybin is almost always consumed in the form of mushrooms. Material sold on the street as psilocybin is usually something else. The effects of psilocybin are very similar to those of LSD, but of shorter duration (4 to 6 hours). Mild-to-moderate headache is common for up to 24 hours following psilocybin use.

DMT is normally inactive orally due to rapid metabolism by monoamine oxidase A (MAO A). It is normally smoked, or occasionally injected or insufflated. In smoked doses of around 25 mg, DMT produces a brief (15 min) and extremely intense experience that can involve complete loss of contact with reality, intense visual experiences, terrifying disorientation, loss of ordinary sense of reality and self, mystical experience, and/or the experience of contact with other worlds or beings. Marked tolerance does not seem to develop to DMT, but it is unusual for people to use it very frequently in any case because of the intense and sometimes overwhelming nature of the experience. DMT can be rendered orally active by administration of an MAO A inhibitor, taken either before or at the same time as the DMT. Ayahuasca is an herbal preparation, used traditionally in the western Amazon basin, which typically contains both monoamine oxidase inhibitors (MAOIs) and DMT. In the typical brew, the MAOIs are harmala alkaloids (harmine, harmaline, and tetrahydroharmine) that come from the *Banisteriopsis caapi* vine, and the DMT comes from the leaves of *Psychotria viridis*. The brew has effects similar to those of classic hallucinogens, although it should be noted that the harmala alkaloids are probably psychoactive to some degree, and they, not the DMT carrier, are traditionally considered the defining and essential ingredient to ayahuasca. Ayahuasca has been used in shamanic traditions in South America for

hundreds of years, and is used as a sacrament in several syncretic religions that developed in Brazil in the 20th century. More recently, it has been used with increasing frequency in North America and Europe for therapeutic purposes or spiritual exploration. The effects of ayahuasca typically last about 4 hours and are generally similar to those of the classic hallucinogens described above. Nausea and vomiting are common, probably due to the harmala alkaloids, and purging is believed by some practitioners to be a valuable part of the experience.

MDMA and Related Compounds (Entactogens). A large group of phenylalkylamines, although structurally similar to classic hallucinogens such as mescaline and DOB, have effects that are quite distinct from those of classic hallucinogens, and therefore will be described here separately. MDMA, 3,4-methylenedioxy-*N*-ethylamphetamine (MDEA), and a number of other compounds have subjective effects characterized by feelings of closeness and loving acceptance toward other people, with little or none of the alterations in sense perception or thought process that are prominent in the classic hallucinogen experience. To varying degrees, these drugs have stimulant effects as well. Because of the effects on feelings toward others, these drugs have been termed "entactogens" of "empathogens." MDMA has attracted interest as a potential adjunct to psychotherapy because it appears to increase tolerance of negative affect and painful memories (e.g., in posttraumatic stress disorders [PTSDs]), possibly allowing them to be dealt with more directly and effectively. Unlike the classic hallucinogens, these drugs are moderately reinforcing in animal models and in humans, and therefore have greater potential for compulsive use. MDMA is taken up presynaptically by the serotonin transporter and induces release of serotonin from presynaptic stores. It also acts as a reuptake inhibitor, and, particularly at high doses, directly stimulates 5HT2A and other serotonin receptors. Overdose can occur easily because MDMA inhibits its own metabolism by CYP 2D6 and CYP 3A4, resulting in blood levels that increase ever more steeply as the dose increases.

ENTACTOGEN EFFECTS. Somatic effects of MDMA include moderate increase in blood pressure and pulse and increase in body temperature. Although the latter averages less than ½ degree Fahrenheit in clinical research settings, it can be much greater when the drug is used in all-night dance parties (raves), particularly if the user is dehydrated. Trismus and bruxism are common side effects. In high doses, muscle stiffness and rhabdomyolysis can occur. In overdose, seizures and cardiorespiratory collapse can occur. Psychological effects include a heightened sense of closeness to others and a deep sense of well-being or acceptance. The emotional tone tends to be very positive and may be euphoric at times. However, strong releases of negative emotion can also occur. Anxiety and

paranoia can occur, although much less commonly than with classic hallucinogens.

Other Serotonergic Hallucinogens. Hundreds of serotonergic hallucinogens have been characterized and new ones are synthesized every year. Many of these "research chemicals" find their way to users, not always accurately labeled or dosed. The effects of most of these drugs in humans have not been studied scientifically. 2C-B and related compounds (e.g., 2C-E, 2C-I) appear, based on self-reports of users, to have effects that combine aspects of classic hallucinogen effects and entactogen effects. Other drugs are reported to have unique effects that have not been characterized adequately in the scientific literature. Recently, a number of highly potent hallucinogens with novel structures have begun to be marketed that have strong LSD-like effects but considerably greater toxicity. For example, Bromo-DragonFLY, an agonist at 5HT2A, 2B, and 2C receptors, is active in doses under 1 mg, with effects that last for several days. 25I-NBOMe is a highly potent full agonist at the 5HT2A receptor (LSD is a partial agonist), which also is psychoactive in doses of under 1 mg. These drugs appear to be much more toxic than classic hallucinogens, and have caused a number of fatalities in cases of overdose, an unfortunately common occurrence owing to the high potency and novelty of these compounds.

Iboga Alkaloids

Ibogaine is a hallucinogenic alkaloid found in several plants, most notably in the root of the *Tabernanthe iboga*, used for medicinal and spiritual purposes by the indigenous people of Western Africa. It has hallucinogenic effects in doses of 400 to 1500 mg, producing an intense and often unpleasant experience that can last for 2 to 3 days. Although ibogaine acts at 5HT2 and 5HT3 receptors, it also inhibits the serotonin transporter and acts at NMDA, mu, kappa and sigma opioid, and muscarinic and nicotinic cholinergic receptors. Ibogaine attracted the interest of addiction researchers when heroin addicts reported that they lost the desire to use heroin after taking ibogaine. Although initial preclinical and clinical results in the treatment of opioid withdrawal were encouraging, and preclinical work suggested that ibogaine may have antiaddictive effects across several classes of drugs, this line of research was eventually discontinued due to toxicity concerns, particularly QTc prolongation and sudden death due to cardiac arrhythmias. Currently, there is interest in developing related molecules that are less toxic, such as 18-methoxy-coronaridine (18-MC). Although ibogaine is not likely to be encountered as a substance of abuse, it continues to be used in some treatment centers outside of the United States, and within underground treatment settings in the United States.

Kappa Opioid Agonists

Salvia divinorum is a species of sage plant, the leaves of which are used traditionally by the Mazatecs in Oaxaca, Mexico. When purified, the main active component, Salvinorin A, is an extremely potent hallucinogen, active in doses as small as 4.5 μg/kg when smoked. Effects are intense but brief, with a time course comparable to smoked DMT, peaking within 2 minutes and dissipating within about 15 minutes. Effects not only overlap with those of classic hallucinogens but also include dissociative effects and impaired recall of the experience. Users report that the effects are distinct from those of other classes of drug. Physiological adverse effects have not been noted, although the experiences can be unsettling or even traumatic. Smokable salvia extracts have been available for a decade or more, and are still not scheduled by the DEA, although they have been outlawed by many states. Many analogues have now been synthesized. The most likely dangers from Salvinorin A appear to be the possibility of psychologically damaging experiences and the possibility of unintentional self-harm during the intoxication state. Users of high doses have been reported to become agitated and perform violent, purposeless actions, for example, running into walls.

Anticholinergic Deliriants

Anticholinergic alkaloids from plants of the nightshade family have been used in shamanic practice and witchcraft for millennia. Although one could argue that the hallucinatory states produced by these drugs should be classified as an anticholinergic delirium, the DSM-5 explicitly lists jimsonweed as an ethnobotanical hallucinogen. There is some logic to this, since plants like jimsonweed, henbane, and belladonna have been used specifically for their hallucinogenic properties. These solanaceous plants contain antimuscarinic alkaloids, principally atropine and scopolamine. Synthetic anticholinergics such as such as benztropine and diphenhydramine (an H1 inverse agonist which also has strong antimuscarinic effects) can produce similar effects in high doses. Anticholinergic effects include dry mouth, mydriasis, blurred vision, increased heart rate, decreased sweating, hyperthermia, seizures, and, occasionally, death. These drugs can produce true hallucinations and other features of delirium including disorientation, agitation, grossly impaired cognition and memory, and fluctuating mental status.

Differences and Similarities with Other Substances of Abuse

In spite of the great diversity among the classes of hallucinogenic drugs, there is significant overlap in the effects of most of these drugs and differences from other major classes of drugs of abuse. The DSM-5 says that “all drugs that are taken in excess have in common direct activation of the brain reward system.” However, this is not true of classic hallucinogens,

ibogaine, or kappa agonists; their effects are neither reliably reinforcing in animal models nor reliably pleasurable to humans. Thus, it is generally accepted that these drugs do not produce compulsive use or true addiction even though they can certainly be the object of a substance use disorder. In general people use these drugs because they value the effects for reasons other than primary reward, for example, because they find the experience meaningful, informative, interesting, or transformative. With NMDA antagonist drugs and serotonergic hallucinogens with entactogen and/or stimulant properties, primary reward pathways are activated, and the experience is reliably pleasurable to many users, though less so than most other substances of abuse. Thus, the usual addictive processes can also motivate continued use of these substances.

HISTORY

PCP and Related Substances

PCP and related compounds were introduced as anesthetics in the 1950s, and ketamine is still used extensively for that purpose. PCP came to attention as an illicit drug of abuse in 1965. PCP is soon becoming infamous due to the association of PCP intoxication with bizarre and violent behavior. Street use of PCP started in 1965 and its popularity peaked in late 1970s. At one time it was considered the most serious drug problem in the Unites States, causing alarm because of its extreme behavioral effects. In 1976, it was estimated that 9.5 percent of persons aged 18 to 25 years had used PCP, and this increased to 14.5 percent in 1979. In 1988, the rate was 4.4 percent for those between 18 and 25. 12.8 percent of high-school seniors reported use in 1979, and this decreased to about 3 percent from 1987 to 1995. Emergency department (ED) visits related to PCP increased fivefold between 1974 and 1979, and then subsequently declined in the following years. Ketamine became increasingly popular in the rave scene in the 1980s. Ketamine was not federally scheduled as a controlled substance until 1999, when it was added to schedule 3.

Other Hallucinogens

As discussed in the previous section, use of serotonergic hallucinogens was virtually unknown prior to the 1960s, when it rapidly became more prevalent. Hallucinogen intoxication became a major concern in that era, in part due to highly publicized tragic outcomes that were rare even at that time (e.g., a person jumping out of a window while intoxicated with LSD), but also because clinically significant problems with intoxication were much more common during that era. Classic hallucinogen intoxication now appears much less often in emergency rooms and psychiatric treatment venues than in the 1970s. In addition to decreases in overall frequency of use since the 1960s and 1970s, social learning and changes in patterns of

use account for this decrease in adverse effects of intoxication. The typical dose of illicit LSD currently taken (25 to 50 μg) is smaller than the 100 to 250 μg doses more commonly used in the 1970s. In addition, today many users have much more information about the possible effects of various doses of hallucinogens and the importance of mental set and environmental setting in minimizing negative effects and outcomes. Still, illicit use poses significant acute risks. Although some of the drugs (such as classic hallucinogens) are safe in carefully screened and prepared participants, in controlled doses, and under ideal conditions, all have some inherent risks. These risks are greatly magnified in illicit settings. People at high risk of adverse outcomes (e.g., people with psychotic disorders or a family history of psychosis) may choose to use these drugs, and preparation and support may be lacking. In addition, under typical conditions of uncontrolled use, the identity and dose of the substance are rarely known with certainty, so overdose or unexpected effects can easily happen. Risks are multiplied when new substances are introduced to the illicit market, due to ignorance as to the potency and effects of the compounds as well as to the inherent risks of the drugs. A recent example is the high frequency of lethal and other serious adverse outcomes with highly potent phenylalkylamines such as Bromo-DragonFLY and NBOMe series compounds.

Hallucinogen use disorder was unknown in the United States prior to the 1960s, when these drugs first became popular. Misuse of hallucinogens has not been described in cultures where hallucinogens are traditionally used for religious, healing, or divinatory purposes. With the sudden increase in hallucinogen use in the 1960s, particularly in the United States and Europe, problematic patterns of use emerged, and were the focus of considerable concern, in part because hallucinogen (particularly LSD) use was intertwined with the cultural upheaval of those times. MDMA became popular in the 1980s. Although initially used for therapeutic purposes or self-exploration predominated, MDMA soon became associated with all-night dance parties (raves), and patterns of excessive harmful use became increasingly common.

COMPARATIVE NOSOLOGY

Table 11.6–2 summarizes the relationships among hallucinogen-related disorders in DSM-5, ICD-9-CM, and ICD-10-CM. As is the case with other classes of substance, DSM-IV substance abuse and dependence were combined into substance use disorder in DSM-5, with elimination of the criterion relating to legal problems and addition of a criterion for craving. DSM-IV treated hallucinogens and PCP-like drugs as distinct categories, whereas DSM-5 includes PCP and related compounds as a category of hallucinogens. ICD-9 and -10 include PCP and related substances under "other psychoactive substances" rather than providing a separate category

or including them under hallucinogens. None of the four systems recognizes a hallucinogen withdrawal syndrome. The hallucinogen-induced disorders recognized by DSM-5 and DSM-IV are identical, except that DSM-5 created separate categories for substance-induced depressive disorder and substance-induced bipolar disorder, while DSM-IV grouped both under substance-induced mood disorder. Hallucinogen intoxication signs specified in DSM-IV are identical to "other hallucinogen" intoxication signs in DSM-5, but DSM-IV also mentions impaired social or occupational functioning within criterion B. Hallucinogen intoxication symptoms listed in ICD-10 likewise are similar to those in DSM-5, but not exactly the same: in ICD-10, -11 specific behavioral and perceptual changes are listed (corresponding to criteria B and C in DSM-5), at least one of which must be present. Finally, the criteria for hallucinogen persisting perception disorder are virtually identical in DSM-5 and DSM-IV. ICD-10 codes hallucinogen persisting perceptual disorder (HPPD) as "Flashbacks," a subtype of residual and late-onset psychotic disorder.

Table 11.6–2.
DSM-5 Hallucinogen-Related Disorders and ICD-9-CM and ICD-10-CM Billing Codes

DSM-5 Disorder	ICD-9-CM	ICD-10-CM
Phencyclidine use disorder	305.90	F16.10[a]
Mild	304.60	F16.20[a]
Moderate	304.60	F16.20[a]
Severe		
Other hallucinogen use disorder	305.30	F16.10[a]
Mild	304.50	F16.20[a]
Moderate	304.50	F16.20[a]
Severe		
Phencyclidine intoxication	292.89	F16.X29[b]
Other hallucinogen intoxication	292.89	F16.X29[b]
Hallucinogen persisting perception disorder	292.89	F16.983
Other phencyclidine-induced disorders[c]	292.9	F16.X59[a]
Phencyclidine-induced psychotic disorder	292.84	F16.X4[a]
Phencyclidine-induced bipolar disorder	292.84	F16.X4[a]
Phencyclidine-induced depressive disorder	292.89	F16.X80[a]
Phencyclidine-induced anxiety disorder	292.81	F16.X21[a]
Phencyclidine-induced intoxication delirium		
Other hallucinogen-induced disorders[c]	292.9	F16.X59[a]
Other hallucinogen-induced psychotic disorder	292.84	F16.X4[a]
Other hallucinogen-induced bipolar disorder	292.84	F16.X4[a]
Other hallucinogen-induced depressive disorder	292.89	F16.X80[a]
Other hallucinogen-induced anxiety disorder	292.81	F16.X21[a]
Other hallucinogen-induced intoxication delirium		
Unspecified phencyclidine-related disorder	292.9	F16.99
Unspecified hallucinogen-related disorder	292.9	F16.99

[a]Used only if there is no co-occurring PCP/other hallucinogen-induced disorder (see below).

[b]If there is co-occurring mild PCP/other hallucinogen use disorder, then *X* = 1. If there is co-occurring moderate or severe PCP/other hallucinogen use disorder, then *X* = 2. If there is no co-occurring PCP/other hallucinogen use disorder, then *X* = 9.

[c]Onset is specified as during Intoxication, not during withdrawal. These disorders are diagnosed instead of intoxication only when the symptoms are sufficiently severe to warrant independent attention.

PHENCYCLIDINE AND OTHER HALLUCINOGEN INTOXICATION

Definition

The DSM-5 criteria for PCP intoxication and other hallucinogen intoxication are similar but differ in some specifics. The diagnosis of intoxication for both classes of drugs requires recent use of a substance in the relevant class, clinically significant psychological or behavioral disturbance, and the lack of a better explanation for the symptoms (e.g., another mental disorder). Although there is overlap in the symptoms of intoxication for PCP and other hallucinogens, as defined in DSM-5, there are significant differences. Criterion B of PCP intoxication requires disruptive behavior changes, which are more frequent and prominent with PCP use than with other types of hallucinogen. The criteria for other hallucinogen intoxication mention behavioral disturbances in criterion B, but the criteria emphasize problematic psychological changes (criterion B) and perceptual changes (criterion C). Both diagnoses require two or more specific signs and symptoms (criterion C of PCP Intoxication and criterion D of Other Hallucinogen Intoxication). These signs and symptoms have no overlap. Signs of PCP intoxication include horizontal, vertical, or rotatory nystagmus, hypertension or tachycardia, numbness or diminished responsiveness to pain, ataxia, dysarthria, muscle rigidity, seizures or coma, and hyperacusis. Rotary nystagmus is a highly specific sign of NMDA antagonist intoxication. For other hallucinogen intoxication, the signs and symptoms include pupillary dilation, tachycardia, sweating, palpitations, blurring of vision, tremors, and incoordination. These signs and symptoms are characteristic of classic hallucinogens and MDMA but may not be seen with types of “other” hallucinogens.

Epidemiology

There are no data that provide accurate estimates of the incidence of hallucinogen intoxication. The rate of DSM-5 hallucinogen intoxication is necessarily lower than the rate of use because the former requires significant adverse behavioral or psychological consequences. Data from the Drug Abuse Warning Network (DAWN) provide estimates of drug-related ED visits that provide a relevant measure of cases of drug use that come to medical attention, although the ED visits are not necessarily directly related to the use of the drug. Table 11.6–3 shows the frequency of

drug-related ED visits for several classes of hallucinogens based on 2011 data. PCP is the hallucinogen most commonly associated with ED visits, with 24.2 ED visits per 100,000 population. MDMA is associated with 7.2 visits per 100,000 population, and all other hallucinogens are much lower. These numbers are low relative to other substances: the corresponding numbers for other classes of drug are stimulants 51.3, marijuana 146.2, heroin 82.7, cocaine 162.1, and alcohol 112.0. Rates are substantially higher for males than for females for each class of hallucinogen. Interestingly, rates are higher in the under 21 age group for the serotonergic hallucinogens, but are higher in the over 20 age group for PCP and ketamine, suggesting that perhaps harmful patterns of use are more persistent for the NMDA antagonist drugs than for the serotonergic drugs, consistent with their greater dependence liability and the greater intrinsic danger of their acute effects. In 2004, dextromethorphan was associated with 5.7 ED visits per 100,000 population, with higher rates in lower age groups (14.7/100,000 among those aged 12 to 20).

Table 11.6–3.
DAWN Data 2011: ED Visits per 100,000 Population

	Total	Male	Female	Under 21	Over 20
PCP	24.2	33.9	14.9	6.5	33.1
Ketamine	0.5	0.7	—	—	0.4
MDMA	7.2	10.2	4.3	11.6	5.5
LSD	1.5	2.6	0.5	2.7	1.1
Other hallucinogen (including mushrooms)	2.6	4.4	0.8	5.3	1.5

Substance Abuse and Mental Health Services Administration, *Drug Abuse Warning Network, 2011: National Estimates of Drug-Related Emergency Department Visits*. HHS Publication No. (SMA) 13–4760, DAWN Series D-39. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2013.

Etiology

Diagnosable hallucinogen intoxication is due to an interaction between the biological effects of the drug and individual and environmental risk factors. As noted above, risks vary from drug to drug, and are greater for higher doses of a given drug, for biologically vulnerable and/or unprepared individuals, and with use in stressful environments lacking in support. The mechanisms of action of the several classes of hallucinogenic drugs are discussed in the introductory section of this chapter.

Diagnosis and Clinical Features

Hallucinogen intoxication presents as acutely altered mental status and behavioral changes following ingestion of a drug. Patients are often

agitated and fearful and may fear that they are “losing their minds” or be experiencing panic or paranoia. Occasionally, users have become acutely depressed, which can lead to suicidal behavior during or soon after intoxication. In other cases, patients may present in an expansive, grandiose, or euphoric state resembling mania. Adverse reactions are more common in inexperienced users, with unexpectedly large doses, or under stressful circumstances, but can occur at any time. The fact of recent drug ingestion is critical to the diagnosis, but may be difficult to ascertain if the patient is severely intoxicated or unwilling to admit to drug use. It is important to carefully assess for a history of other drug ingestion or exposure and psychiatric and somatic illness and to perform an adequate physical examination including neurologic evaluation. All classes of hallucinogens can produce acute anxiety, agitation, inappropriate and labile affect, visual illusions and hallucinogens, and paranoia or other delusions. The characteristics features of the intoxication state vary by class of substance.

PCP and Related Substances. PCP and ketamine may be taken by any route, but PCP is more commonly smoked or snorted, while ketamine is frequently injected, snorted, or ingested. A typical street dose of PCP is about 5 mg, while a typical dose of ketamine is 100 to 200 mg. PCP is currently manufactured only illicitly, while ketamine is derived both from diverted licit (injectable formulation) material and illicit manufacture. Dextromethorphan is usually taken in the form of over-the-counter cough syrup or tablets. These drugs produce a dissociative state with a greater degree of cognitive and neurological impairment (including the signs listed in the criteria for PCP intoxication), than that produced by the classic hallucinogens. Symptoms become more severe with higher doses, and may warrant the diagnosis of PCP or ketamine-induced delirium if disturbances of attention and cognition are prominent. With PCP intoxication, typical features include impaired cognition, nystagmus, hypertension, tachycardia, dysarthria, hyperacusis, and increased deep tendon reflexes. Thought disorder is prominent, but hallucinations may also be present. Behavioral symptoms include behavioral disinhibition, agitation, panic, rage, aggression, and exaggerated reactions to environmental stimuli.

Medical complications are much more common than with classic hallucinogens. Hyperexcitability, hypertension, or hyperthermia may indicate significant toxicity constituting a medical emergency. Hypertensive crisis can lead to cerebral hemorrhage. Severe hyperthermia can occur. At higher doses of PCP, symptoms of dissociative anesthesia emerge, with numbness and diminished responsiveness to pain. At still higher doses, other receptor effects become prominent, as PCP interacts with monoamine uptake pumps, sigma receptors, sodium and potassium channels, nicotinic and muscarinic acetylcholine receptors, and mu opioid

receptors.

Unpredictable, agitated, and violent behavior occur more commonly with PCP than with other hallucinogens. Formal thought disorder, delusions, and frank hallucinations can occur, as well as confusion, impairment of motor control, impaired speech, and catatonic-like symptoms. Overdoses of PCP and related compounds can have serious medical sequelae and have led to fatalities. PCP overdose can cause seizures, hyperthermia, hypertensive crisis, cerebral hemorrhage, respiratory arrest, coma, and death.

Ketamine intoxication generally has a less malignant presentation than PCP intoxication. Medical complications are less common than with PCP due to the lower potency and shorter half-life of ketamine, and because ketamine has fewer significant actions at sites other than the NMDA receptor. Presenting complaints include anxiety, chest paint, and palpitations. Tachycardia and hypertension are common findings on physical exam, but are generally much less severe than the effects seen with PCP intoxication. Rotary nystagmus may be seen, but is present in only a minority of cases. Death from ketamine intoxication is very rare, except when it is used in combination with other drugs.

PCP or ketamine intoxication is often a sign of moderate or severe substance use disorder, so that should be diagnosed and treated if present.

Classic Hallucinogens. Although patients with classic hallucinogen intoxication can show a wide range of symptoms similar to those of PCP intoxication, physiological effects are generally very mild, and unlikely to require specific intervention. Patients with classic hallucinogen intoxication are more likely to remain fully oriented and less likely to be confused or show severe cognitive, verbal, or motor impairment. Anxiety and panic are frequently observed, as patients may feel that they are "losing their minds." Marked mydriasis is a prominent feature of classic hallucinogen intoxication. Mild-to-moderate elevation in blood pressure and pulse may be observed. Overdoses of LSD and psilocybin have not been observed to cause death.

MDMA. MDMA intoxication does not usually cause thought disorder, paranoia, disorganized behavior, or aggression. It usually comes to attention in cases of overdose, in which case symptoms are essentially those of serotonin syndrome: hyperthermia, rigidity, hyperreflexia, seizures, rhabdomyolysis, renal failure, etc. Symptomatic intoxication is more likely to occur when high doses are taken and in the context of dehydration and/or hyperthermia. Overdose can easily occur with MDMA not only because of the issue of unknown potency and purity common to most illicit drugs but also because of the relatively low index of toxicity and the nonlinear pharmacokinetics of MDMA. MDMA inhibits its own

metabolism, with the result that a relatively small increment in dose can lead to much greater increase in blood levels and effects. In high doses, MDMA can cause classic hallucinogen-like effects. Delirium may be present, in which case it would be diagnosed instead of intoxication.

Other Serotonergic Hallucinogens. The effects of many of these drugs have not been adequately characterized. The effects of Bromo-DragonFLY and 25I-NBOMe are reported to be similar to those of LSD, but serious adverse effects appear to be much more common, including hypertension, tachycardia, seizures, hyperpyrexia, elevated creatine kinase, vasoconstriction leading to tissue necrosis, and a number of deaths when users mistakenly took very large doses. The high potency combined with a low index of safety makes these drugs much more dangerous than the classic hallucinogens and entactogens discussed above.

Ibogaine. Ibogaine intoxication resembles classic hallucinogen intoxication but may include a greater degree of dissociation and confusion similar to PCP and related drugs.

Salvia. The effects of *Salvia* intoxication are similar to those of PCP or ketamine intoxication, but the duration (15 minutes or less) makes it unlikely that they will be observed in a clinical setting.

Anticholinergics. Anticholinergic drugs produce true hallucinations in the context of delirium as described in the section on hallucinogen-induced delirium below. Peripheral signs may be less prominent with agents that are relatively potent centrally (e.g., scopolamine).

Pathology and Laboratory Examination

Urine toxicology may be useful to establish the acute ingestion of hallucinogens and/or other drugs, although hallucinogens other than PCP and MDMA will not be detected with routine urine drug screens. False-negative screens for PCP are not uncommon. Toxicologic screen for other drugs of abuse is useful since more than one drug may have been ingested. Information from friends or others with the patient during the previous few hours is useful. Serum, urine, or gastric content samples can be sent to analytic laboratories for specific mass spectrographic analysis, but initial diagnosis and treatment must be from clinical assessment. Users may not be knowledgeable about the actual hallucinogen dose or drug ingested. The dose delivered from plant material is always uncertain.

MDMA intoxication usually comes to the attention of clinicians in cases of overdose, in which case symptoms are similar to those of serotonin syndrome and/or stimulant toxicity: hyperthermia, rigidity, hyperreflexia, seizures, rhabdomyolysis (elevated creatine phosphokinase), etc.

Disseminated intravascular coagulation and acute renal failure have been reported. Deaths have occurred secondary to cerebral edema and seizures secondary to hyponatremia. Hyponatremia occurs much more frequently in women than in men. Cardiac arrhythmias and cardiovascular instability are also possible.

Differential Diagnosis

The most important immediate task in evaluation of cases resembling hallucinogen intoxication is to rule out medical and psychiatric emergencies or to treat any that are present. Since the symptoms of hallucinogen intoxication can overlap with those of psychosis and delirium, it is critical to distinguish hallucinogen intoxication from these disorders. Hallucinogen-induced psychosis may be diagnosed if the symptoms of psychosis dominate the clinical presentation and are beyond (in severity or duration) those typically associated with intoxication. Particularly for PCP, diagnostic boundaries between intoxication, delirium, and psychosis are blurry, and the state of the patient can fluctuate.

Intoxication with some drugs (e.g., stimulants) and withdrawal from others (e.g., alcohol and sedative hypnotics) can also cause symptoms mimicking those of hallucinogen intoxication. The principal diagnostic tools include a history if available, urine toxicology and additional laboratory tests to rule out metabolic derangements, and physical, neurological and mental status exam, with observation over time. In most cases, acute hallucinogen intoxication will resolve within hours in the absence of complicating factors. Except for very long-lasting drugs (DOM, ibogaine), symptoms that last longer than 12 hours may be due to other causes or to hallucinogen-induced psychotic disorder. PCP intoxication can resemble the effects of other toxic drug ingestions or deliria. The combination of nystagmus (especially rotary), ataxia, and hypertension is relatively specific for PCP intoxication.

Course and Prognosis

Uncomplicated hallucinogen intoxication normally resolves without sequelae when the acute effects of the drug wear off, within 12 hours except for very long-acting hallucinogens such as DOM or ibogaine. Barring medical complications, complete recovery from PCP intoxication within 24 to 72 hours is usual. Uncommonly, hallucinogen-induced disorders may follow hallucinogen intoxication (see below).

Treatment

PCP (or PCP-like Hallucinogens). Treatment of acute ingestion is supportive, aimed at minimizing medical emergencies (seizures, hypertension, hyperthermia, and rhabdomyolysis) and behavioral

consequences. PCP levels and symptoms may fluctuate over many hours (especially with oral ingestion). Ionized PCP may be trapped in the stomach. Activated charcoal may be administered if oral administration is suspected. Sensory input should be minimized. Physical restraint can prevent violent behavior but can also increase agitation and the probability of rhabdomyolysis. Agitation, seizures, and hypertension should be treated primarily with benzodiazepines. Antipsychotics may be useful in cases of relatively mild intoxication, but in more severe cases they can lower seizure threshold, cause dystonia, and exacerbate hyperthermia and anticholinergic symptoms. Hyperthermia should be treated with standard cooling measures. Rhabdomyolysis should be treated with intravenous fluids and diuretics. Alkalization of the urine is controversial because there is a theoretical risk of causing increased reabsorption of PCP.

Classic Hallucinogens. Except in cases of delirium or overdose, treatment of acute ingestion of a classic hallucinogen is conservative. Reassurance and a calm and safe environment are the most useful interventions. Patients who are convinced they are going crazy or being harmed may benefit from reminders that they are under the influence of a drug and that these feelings will soon pass. Some degree of insight is often preserved. If agitation presents risk to the patient or others, or if the patients appears to be in severe distress, oral or parenteral benzodiazepines are the first line of treatment. Antipsychotics should be used only if behavior can not otherwise be managed. Antipsychotics with 5HT2A antagonist effects are preferred. D2 blockers such as haloperidol may worsen symptoms. Restraints are rarely necessary, and should be used only if there is immediate danger of self-harm or assaultive behavior. Patients should be monitored closely, and the mental status should be monitored at frequent intervals since mood and thought content can change rapidly during intoxication. Support from a family member or close friend can be helpful following the resolution of the acute intoxication state. Patients who have had very disturbing or traumatic experiences may benefit from short-term therapy to help them to integrate the experience and to identify and treat any psychological sequelae.

MDMA. MDMA intoxication often comes to medical attention in cases of overdose or medical complications. Such patients should be evaluated emergently, and ICU admission may be required. Cardiac monitoring and frequent monitoring of vital signs and mental status should be conducted. Activated charcoal should be administered if recent ingestion is suspected. Cooling measures should be undertaken in cases of hyperthermia. Dantrolene is sometimes used to treat hyperthermia due to MDMA ingestion, but its efficacy is not proven. Seizures, hypertension, and agitation are treated with benzodiazepines. In cases of rhabdomyolysis,

intravenous fluids, alkalization of the urine, and furosemide are recommended.

PHENCYCLIDINE USE DISORDER AND OTHER HALLUCINOGEN USE DISORDER

Definition

The DSM-5 criteria for PCP use disorder and other hallucinogen use disorder are identical to those for all other classes of substances, with the exception that the criterion for withdrawal does not apply: there is no recognized syndrome of withdrawal in humans for any class of hallucinogens, although there is evidence of clinically meaningful withdrawal following heavy MDMA use, and evidence of PCP withdrawal in animal models. As with other substance use disorders, the SUD severity is specified as mild (2–3 criteria), moderate (4–5 symptoms), or severe (6 or more criteria).

Epidemiology

Hallucinogen use disorders are uncommon. The lifetime prevalence of hallucinogen use disorder has been estimated to be 0.6 to 1.7 percent. Approximately 5 percent of those with a lifetime history of hallucinogen use are estimated to develop hallucinogen dependence. About 7.8 percent of adult current users and 17 percent of adolescent current users meet criteria for past year hallucinogen use disorder. A problem with most of the epidemiologic surveys is that they do not always distinguish between the various classes of hallucinogen. Therefore, little is known about rates of dependence for most of the individual hallucinogens.

Table 11.6–4 summarizes hallucinogen use and hallucinogen use disorder estimates from the US NSDUH, as well as estimates for 12th graders from the US Monitoring the Future (MTF) survey. Overall, 15.1 percent of the US population aged 12 and over (more than 20 million people) have used hallucinogens at least once. Rates of past year/past month use in the NSDUH survey are highest in the 18 to 25-year-old age group, and lowest in the >25-year-old group. Peak age for lifetime use is age 30 to 34 (24.4 percent), while for past year use the peak is at age 21 (8.4 percent), and for past month use at age 18 (2.6 percent). Among 12th graders in the MTF study, rates of any hallucinogen use are a bit lower than the rates reported for the 18 to 25 age group in the NSDUH survey. The rates are considerably lower for the younger age groups (8th and 10th graders). Rates of past year and past month use drop off steadily after the early 20s. Overall rates of hallucinogen use have been fairly stable since 2002. The rate of past month use for age 12 and greater has been between 0.4 percent and 0.5 percent. For those in the age group of 12 to 17 years old, it has dropped from 1.0 percent to 0.5 percent. Among 18 to 25 year olds it has ranged from 1.5 percent to 2.0 percent. Since 2006, significant

increases have been seen in the use of DMT, AMT, or 5-MeO-DIPT, and *Salvia*. Rates of hallucinogen use are somewhat higher in the Western United States, and in urban counties. Regarding race and ethnicity, rates of hallucinogen use are highest in Native Americans, Whites, and those declaring multiple races. Rates of current hallucinogen use disorder are low in all age groups, with an overall rate of 0.1 percent for any substance use disorder, and 0.0 for DSM-IV hallucinogen dependence (roughly equivalent to moderate or severe hallucinogen use disorder according to DSM-5). Hallucinogen use disorders are most common in people under 25, with a peak in the 18 to 25-year-old age group (0.4 percent for any hallucinogen use disorder, 0.1 percent for hallucinogen dependence).With respect to incidence of use, 1.1 million persons aged 12 or older used hallucinogens for the first time in the 12 months prior to being interviewed in 2013. Since 2002 this number has ranged from roughly 900,000 to 1.3 million. In the 2013 NSDUH survey, among individuals under age 50 who reported initiation of hallucinogen use in the past 12 months, the average age of initiation was 19.9 years.

An analysis of data from the 2005 NSDUH survey examined characteristics of past year MDMA users in contrast to past year users of other hallucinogens (including LSD, PCP, mescaline, peyote, and psilocybin). About 1.4 percent of the adult sample had used one or more hallucinogens in the past year. The overall prevalence of hallucinogen use disorder was 0.11 percent. Both groups were approximately two thirds male and aged 25 or younger. Users of other hallucinogens were more likely to be white than were users of MDMA. The mean age of first use was 18 to 19 in both groups. Most users in both groups used infrequently, with 34 percent of MDMA users and 22 percent of other hallucinogen users having 12 or more occasions of use in the past year. Among the past year users, 4.9 percent of both groups met DSM-IV criteria for abuse, but 3.6 percent of the MDMA users versus 2.2 percent of the other hallucinogen users met criteria for dependence, suggesting that MDMA use may confer greater risk of moderate to severe hallucinogen use disorder than does use of other hallucinogens. Frequent use and past year marijuana use disorder were associated with hallucinogen use disorder among past-year hallucinogen users.

Etiology

As with other drugs, hallucinogen use disorder likely arises out of multiple interacting biological, psychological, and social factors. Hallucinogen use disorders differ from other classes of substance use disorder in that classic incentive reward-driven mechanism does not play a central role except for PCP and related compounds, and possibly in the cases of MDMA-type drugs. PCP is avidly self-administered in animal models, and human users frequently show patterns of compulsive use, relapse, etc. Classic

hallucinogens are not strongly reinforcing, but rather aversive in animal models. Hallucinogens can cause euphoria, but this effect is not reliable for most users, and dysphoric or mixed experiences are very common. A withdrawal syndrome is not recognized for any of the hallucinogens. Although some degree of physical dependence is possible with NMDA receptor antagonist hallucinogens and MDMA-like compounds, it is unlikely that continued use is driven by negative reinforcement through relief of withdrawal symptoms. Strong cravings also do not appear to be common in hallucinogen users. Exposure, other drug use, expectancies, and perceived meaningfulness of the experience probably play more important roles. Hallucinogens are readily available and relatively inexpensive. Hallucinogens may be used for recreational purposes, out of curiosity or for excitement or thrills, for personal or spiritual growth, or with therapeutic intent.

Table 11.6–4.
NSDUH and Monitoring the Future Survey Data on Prevalence of Hallucinogen Use and Hallucinogen Use Disorder

	NSDUH 2013,[a] Age 12 and Up					MTF 2014,[b] 12th Grade		
Substance	Lifetime	Past Years	Past Months	Dependence	SUD	Lifetime	Past Years	Past Months
Any hallucinogen	15.1	1.7	0.5	0.0	0.1	6.3	4.0	1.5
M	18.7	2.2	0.7					
F	11.8	1.2	0.3					
12–17	2.8	2.2	0.6	0.1	0.3			
18–25	17.6	6.8	1.8	0.1	0.4			
>25	16.2	0.8	0.3	0.0	0.0			
LSD	9.4	0.4	0.1			3.7	3.7	1.0
M	12.3	0.6	0.1					
F	6.8	0.3	0.1					
PCP	2.5	0.0	0.0			1.3	0.8	0.4
M	3.2	0.0	0.0					
F	1.7	0.0	0.0					
MDMA	6.8	1.0	0.3			5.6	3.6	1.4
M	7.9	1.2	0.3					
F	5.7	0.8	0.2					
Peyote	2.2							
Mescaline	3.1							
Psilocybin	8.7							
Cough or cold med.	3.3	0.6	0.2					
M	4.3	0.7	0.2					
F	2.4	0.5	0.1					
Ketamine	1.0	0.1	0.0					
M	1.3	0.1	0.0					
F	0.8	0.1	0.0					

DMT, AMT, 5-MeO-DIPT	0.7	0.1	0.0
M	1.0	0.2	0.0
F	0.4	0.1	0.0
Salvia	2.1	0.2	0.1
M	3.1	0.3	0.1
F	1.1	0.1	0.0

[a]Substance Abuse and Mental Health Services Administration, *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings*, NSDUH Series H-48, HHS Publication No. (SMA) 14–4863. Rockville, MD: Substance Abuse and Mental Health Services Administration, 2014.

[b]Johnston LD, Miech RA, O'Malley PM, Bachman JG, Schulenberg JE. *Use of Alcohol, Cigarettes, and Number of Illicit Drugs Declines Among U.S. Teens*. University of Michigan News Service: Ann Arbor, MI; 2014. Retrieved 01/20/2015 from http://www.monitoringthefuture.org.

Diagnosis and Clinical Features

As with other classes of substance, the SUD criteria for hallucinogens appear to represent a single underlying dimension of severity, justifying the elimination of the abuse/dependence distinction for hallucinogens in DSM-5. Although the diagnostic criteria for hallucinogen use disorder are identical to those of any substance use disorder, the criteria perform somewhat differently for hallucinogens than for other substances. It seems probable that there would be differences between the various types of hallucinogens as well. In general, hallucinogen use disorders appear to be associated with fewer adverse consequences than most other substance use disorders. Frequency of use is low relative to other substance use disorders. The two criteria most linked to loss of control (inability to cut down or control use, and using more than intended) are by far the most difficult items (last to be endorsed) in item response theory studies of hallucinogen use disorder criteria. This suggests that people often continue to use hallucinogens because they enjoy or value the effects, rather than that they lose the ability to regulate their use. There is no recognized withdrawal syndrome for any of the hallucinogens, although animal data suggest that physical dependence to arylcyclohexamines can occur.

True physical dependence does not occur with serotonergic hallucinogens. The SUD criteria most commonly endorsed are spending a lot of time using or recovering from the substance, and tolerance. The tolerance criterion is of questionable validity for classic hallucinogens because of the very rapid and reversible development of tolerance. Marked tachyphylaxis lasting no more than a few days can develop following a single dose of a classic hallucinogen such as LSD. This phenomenon does not imply tolerance in the sense of the DSM-5 SUD criterion.

Pathology and Laboratory Examination

There are few physical signs or symptoms of hallucinogen use disorders,

independent of the signs and symptoms of intoxication. MDMA use may be associated with decreased numbers of serotonin transporters and 5HT2A receptors on PET scan, but this has not been used diagnostically. Urine, blood, or hair analysis can be used to detect ingestion of hallucinogens. PCP is detectable in urine for up to 8 days after use. However, other hallucinogens are not included in routine tests, and the ever-growing number of hallucinogens means that there will always be hallucinogens that cannot be detected with any commercially available test. Chronic ketamine use has been associated with disorders of the GI tract (gastritis, cholestasis) and the urinary tract (ulcerative cystitis, impaired renal function).

Neurotoxicity. MDMA is neurotoxic in animal models, and there is evidence of adverse effects on the brains of people who have used the drug extensively. There is fairly consistent evidence of reductions in serotonin transporters and 5HT2A receptors in MDMA users. Neurocognitive deficits in memory and inhibitory control have been found in some studies, although most of these studies are confounded by other drug use. Based on animal data, it seems likely that the neurotoxicity of MDMA may be dose-related, and mediated in part by hyperthermia. Hyponatremia also could play a role. No adverse cognitive effects have been observed in clinical studies of MDMA in which controlled doses (up to 150 mg) are given on a few occasions. In clinical research settings, maximum temperature increase is about ½ degree Celsius.

Deficits in memory and executive function have been observed in heavy users of PCP and ketamine. Vacuolization is seen in high doses in animal models with PCP and ketamine, possibly contributing to cognitive impairment in association with chronic use in humans. In rodents and primates, the most sensitive brain regions are the posterior cingulate and retrospenial cortices.

There is no evidence of significant neurotoxicity with classic hallucinogens. However, well-controlled trials with large samples have not been conducted.

Differential Diagnosis

Use of hallucinogens frequently occurs in the context of other substance use. In such cases, it is important to be clear which symptoms are caused by which substance in order to arrive at the correct SUD diagnosis/es.

Course and Prognosis

Hallucinogen use disorders tend to be of short duration and usually follow a self-limited course (another way in which they differ from other substance use disorders). Rates drop to near zero in people over the age of

Treatment

There are few clinical trials of treatment for any hallucinogen use disorder, and no pharmacotherapy studies. Since there are no specific treatments or programs for treatment of hallucinogen use disorders, treatment-seeking patients with these disorders are treated within addiction treatment programs designed for other substance use disorders or for multiple addictions. The efficacy of psychosocial treatments that have empirical support for treatment of other substance use disorders is unknown in the case of hallucinogen use disorders.

HALLUCINOGEN PERSISTING PERCEPTION DISORDER

Definition

A psychopathologic condition first appeared in DSM-III-R as "Post-Hallucinogen Perception Disorder"; the criteria were minimally modified for DSM-IV under the term "Hallucinogen Persisting Perception Disorder (HPPD)," and the criteria were kept essentially the same for DSM-5. With HPPD, users re-experience one or more perceptual effects (similar to those experienced during previous hallucinogen use) which cause clinically significant distress or impairments in functioning, and are not due to a general medical condition (i.e., migraine with aura, temporal or occipital lobe epilepsy, and temporal lobe infectious illness), better ascribed to another mental condition (i.e., schizophrenia, dementia, and delirium), or due to hypnopompic hallucinations. The perceptual symptoms mentioned as examples in the diagnostic criteria are visual (e.g., geometric hallucinations, image "trails" behind moving objects), but perceptual abnormalities from other sensory modalities and re-experiencing of other aspects of hallucinogen-induced altered states have been reported. The symptoms can occur spontaneously or be triggered by stress, anxiety, exercise or use of another drug (i.e., cannabis). They can occur episodically or take the form of constant perceptual disturbances without an identifiable trigger. The term "flashback" has sometimes been used synonymously with HPPD but is distinguished from HPPD in connoting brief persisting perceptual effects, similar to those experienced during recent hallucinogen use, that do not necessarily cause clinical distress or impairment and can even be experienced as pleasurable by the individual. Unlike flashbacks, HPPD tends to be recurrent with a continuous or intermittent/paroxysmal pattern of symptoms that is either a slowly reversible or a chronic permanent condition experienced as distressing by patients.

History

Persisting perceptual abnormalities ("flashbacks") following the use of hallucinogens were first described over 100 years ago with mescaline. The next reports of this condition occurred in the 1950s and 1960s, mostly in relation to LSD ingestion, and described prolonged changes in normal perception that were similar in nature to the perceptual effects experienced under the influence of the hallucinogen, persisting subacutely (weeks to months) and chronically after use of the hallucinogen. It should be noted that in addition to the classic hallucinogens, a smaller cohort of case reports of persisting perceptual disturbances have been reported and attributed to the use of MDMA, ketamine, and cannabis.

Epidemiology

Although the exact prevalence of HPPD is unknown with little in the way of rigorous epidemiologic study, it is thought to be a rare (but highly distressing) condition given the relatively few cases reported out of the tens of millions of doses of hallucinogens used over the last 50 years, and it appears to be significantly less common in research settings with careful screening and preparation. Data from a population study derived from the 2001 to 2004 NSDUH survey found no association between lifetime use of psychedelics and past year use of LSD with past year symptoms of HPPD-type visual phenomena. HPPD appears to be most prevalent in males with symptoms developing in late adolescence. In a systematic review of HPPD in 2003, HPPD symptoms ranged from less than 5 percent in research participants receiving LSD to approximately 50 percent in polysubstance users. In two recent web-based surveys of HPPD symptoms in individuals with histories of hallucinogen use, 24 to 40 percent of respondents reported persistent or constant visual distortions and 4 to 11 percent reported symptoms distressing enough to meet some of the criteria for HPPD or to seek treatment of the distress.

Etiology

The pathophysiology of HPPD has yet to be conclusively elucidated. However, given the predominance of visual symptoms, their typology (i.e., geometric hallucinations, trailing, and positive afterimages), and that patients typically do not believe the visual phenomenon are real, lower levels of the visual pathway in the brain (including the visual cortex) have been implicated. However, multiple brain regions may be affected and the thalamus may be a particular area of interest to examine since other perceptual stimuli (i.e., auditory and sensory) can be abnormal in patients with HPPD in addition to the visual system. Abraham and Aldridge initially proposed that damage to lower level visual system inhibitory cortical interneurons (which express 5HT2A receptors and release GABA upon activation) were primarily responsible for the mechanism of action of

HPPD and could account for visual disturbances by altering the signal to noise ratio of incoming visual stimuli. A broader conceptualization of the pathophysiology of HPPD includes a misbalance or dysregulation between excitatory versus inhibitory input in the visual system where both increases in excitation and/or decreases in inhibition could result in the variety of visual disturbances seen in HPPD.

Diagnosis and Clinical Features

The most common perceptual distortions (illusions or hallucinations) in HPPD are in the visual system (i.e., geometric hallucinations, flashes of color, intensification of colors, trails of images of moving objects, positive afterimages, halos around objects, false perceptions of movement in the peripheral visual fields, micropsia, and macropsia), although auditory and sensory distortions have been reported as well. In addition to visual illusions and hallucinations, patients commonly experience depersonalization and derealization symptoms. It is important to note that HPPD is not a psychotic spectrum illness and patients have intact reality testing and awareness that the illusions and hallucinations are not real.

Differential Diagnosis

Alternative explanations for unusual perceptual experiences must be considered before diagnosing HPPD, for example, visual epilepsy, migraine, delirium, dementia, schizophrenia, and hypnopompic hallucinations, PTSDs, and other hallucinogen-induced disorders. HPPD is easily distinguished from psychotic disorders because symptoms are sensory (primarily visual), delusions are absent, and insight, reality testing, and though process are intact.

Course and Prognosis

The course of HPPD is variable. About half the patients with HPPD stop having symptoms within 5 years. However symptoms persist for decades in some cases. Symptoms can be worsened by a variety of stimulant drugs, marijuana use, fatigue, exercise, and infections.

Treatment

Supportive psychotherapy and cognitive-behavioral treatment are warranted to help individuals cope and adapt to their symptoms. There is no established evidence-based pharmacologic algorithm to treat HPPD with most of the trials coming from case reports, case series, and open label trials with little in the way of randomized controlled trials. Benzodiazepines (e.g., alprazolam, clonazepam) have been shown to reduce some symptoms of HPPD without leading to disease remission. Their efficacy may be related to their action at the above-mentioned cortical serotonergic

inhibitory interneurons with GABAergic outputs. Along these lines, there is some evidence that phenytoin (which in addition to inhibiting sodium channels also increases CNS GABA levels and potentiates GABA-mediated postsynaptic inhibition) decreases symptoms of HPPD. Clonidine in an open label trial was also shown to alleviate some of the symptoms of HPPD and suggests a possible role of sympathetic nervous system activation in the pathophysiology of HPPD. There is some evidence supporting the use of the opioid receptor antagonist naltrexone and typical antipsychotics (haloperidol, trifluoperazine, and perphenazine) in treating symptoms of HPPD, although the adverse side effect profile of typical antipsychotics limits their use. The data on the utility of SSRIs to ameliorate symptoms of HPPD are mixed, with some data supporting their efficacy and some suggesting a worsening of symptoms. Risperidone (and possibly the atypical antipsychotics in general) should be avoided as a treatment option in HPPD as there is evidence that risperidone worsens HPPD symptoms. More controlled trials are needed to establish better pharmacologic treatments for HPPD that could help further elucidate the pathophysiology of the illness.

PCP- AND OTHER HALLUCINOGEN-INDUCED PSYCHOTIC DISORDER

Definition

DSM-5 defines hallucinogen-induced psychosis as a syndrome meeting criteria for a psychotic episode that has its onset during or shortly following use of a hallucinogen, which persists beyond the usual period of action of the drug, and is not better accounted for by another psychotic disorder or other psychiatric, neurological, or medical disorder. The disorder must be sufficiently severe to cause significant impairment and/or distress.

History

PCP. Although acute intoxication with PCP can closely resemble a psychotic state and has served as useful model of schizophrenia, persisting psychosis following PCP use was not recognized until 1974, during an epidemic of PCP use in the Washington, DC area. Numerous episodes of substance-induced psychosis were noted in patients who had been smoking "angel dust." These episodes closely resembled severe schizophrenic episodes, often included violent or aggressive behavior, and often persisted for weeks in spite of aggressive antipsychotic treatment. They eventually resolved completely. Although the prevalence and attention to this phenomenon have receded as the prevalence of PCP use has decreased in recent years, such cases are still seen fairly frequently in areas where PCP use remains prevalent.

Other Hallucinogens. After LSD became widely used as a recreational

drug in the 1960s, it was increasingly recognized that its use was sometimes associated with persisting psychosis, a common example being first use by a teenager or young adult (during the age of typical onset of schizophrenia) who developed the onset of schizophrenia and went on to develop a chronic course of the illness. However, it has never been conclusively shown that classic hallucinogens can produce a persistent psychosis in someone without an underlying psychotic disorder or diathesis (see "Epidemiology" section below).

Epidemiology

PCP. PCP psychosis is a well-known presentation in emergency rooms and acute psychiatric hospitals, but the prevalence of this condition is unknown. Increased delusional ideation has been reported among frequent users of ketamine. Ketamine and DXM have been reported to cause persisting psychotic episodes, but much less frequently than PCP. Risk factors also are not well known. Patients with pre-existing psychotic disorder are very sensitive to the psychotogenic effects of PCP, but such individuals would not generally receive the diagnosis of PCP-induced psychosis.

Other Hallucinogens. Although classic hallucinogen use can cause transient psychotic-like positive symptoms in normal volunteers and can provoke sustained psychosis in vulnerable people with psychotic spectrum illnesses (i.e., schizophrenia, schizoaffective disorder, and bipolar disorder with psychotic features), there is little evidence linking classic hallucinogen use to prolonged psychosis in individuals without a psychotic diathesis, and the type of research necessary to establish a causal link has never been conducted. New cases of schizophreniform psychoses and related disorders following hallucinogen exposure were frequently reported during the time when hallucinogen use first became widespread in the United States. However, the prevalence of schizophrenia has remained relatively stable for as long as it has been measured even though the prevalence of use of classic hallucinogens (particularly LSD) has fluctuated widely over time. It is likely that hallucinogen use sometimes precipitates psychosis in people who would inevitably develop it later, or that previously unrecognized psychotic disorders may become apparent in the context of hallucinogen use. But it remains possible that hallucinogen use occasionally causes a persistent psychotic disorder in individuals who otherwise would never have been psychotic. Estimates of the prevalence of LSD-induced psychosis as assessed by early psychedelic researchers and clinicians (many working with and administering LSD to psychiatric inpatients) were reported as follows in two major reviews of the topic: 0.8/1,000 research volunteers (single case out of 1,250 where the volunteer was the identical twin of an individual with schizophrenia), and 1.8/1,000 psychiatric patients (7 cases

out of approximately 3,850 patients); 0/170 research volunteers and 9/1000 among psychiatric patients (37 out of 4,300 patients). A recent cross-sectional study evaluating data taken from the years 2001 to 2004 of the NSDUH with a sample of 130,152 (representing a random sample of the US population living in households) found that 21,967 (13.4 percent) reported lifetime psychedelic (LSD, psilocybin, mescaline, peyote) use and did not find any significant associations between lifetime use of any psychedelic, or past year use of LSD, and increased rates of any psychiatric symptoms (including psychosis) or mental health outcomes. In the re-emergence of psychedelic research since the early 1990s, close to 1,000 doses of psilocybin (ranging from low to high doses), approximately 200 doses of intravenous DMT, and at least 2 dozen doses of LSD have been administered safely in Europe and the United States at major academic medical centers with no reported cases of prolonged psychosis, HPPD, or any other treatment-related serious adverse events.

Etiology

PCP. Although PCP and ketamine intoxication (a transitory state) have been used extensively as a model of psychosis, etiology and pathogenesis of PCP-induced psychosis are not well studied. In animal models, a single dose of MK801 (an NMDA receptor antagonist with particularly strong psychotogenic effects) has been used as a model for first-episode psychosis. A single injection can cause persisting changes in hippocampal glutamatergic neurotransmission, resulting in loss of synaptic plasticity and deficits in learning and cognitive function, although other behavioral effects gradually resolve.

Other Hallucinogens. In many cases, the onset of psychosis following hallucinogen ingestion may be due to the expression of underlying diatheses for schizophrenia or bipolar disorder. There is no established etiology for other hallucinogens inducing persisting psychosis, but persistent dysregulation of serotonergic systems is possible. Also, because classic hallucinogens cause downstream effects on glutamatergic neurotransmission similar to those produced by PCP and ketamine, it is possible that persistent changes in glutamatergic systems could play a role in some cases of persisting psychosis induced by classic hallucinogens.

Diagnosis and Clinical Features

A history of recent hallucinogen use is necessary to establish the diagnosis. However, it may be difficult to confirm or rule out recent hallucinogen ingestion or to be certain which hallucinogen was ingested. PCP in particular is often misrepresented as other drugs. Patients may deny drug use, and drug screens will often be negative.

PCP. PCP-induced psychosis is often phenomenologically indistinguishable from schizophrenia. Thought disorder and negative symptoms may be prominent, and delusions and hallucinations are not always present.

Other Hallucinogens. Patients with hallucinogen-induced psychosis are said to be younger at onset, have more visual hallucinations, more depression, and a greater prevalence of a family history of mood disorder than nondrug-using patients with schizophrenia. Clinically, they are said to have more mood swings, euphoria, grandiosity, hyperreligiosity, and multimodal hallucinations, and more prominent positive than negative symptoms. Thus, the psychosis tends to have elements of mania. If mood symptoms are prominent, the diagnosis of a hallucinogen-induced mood disorder should be considered.

Pathology and Laboratory Examination

Urine drug screens may be useful to confirm recent ingestion of hallucinogens, but routine tests are available only for PCP and MDMA. Physical exam and laboratory findings will be normal following resolution of acute intoxication or delirium and any medical sequelae.

Differential Diagnosis

Hallucinogen-induced psychosis usually resembles the psychosis of schizophrenia, but sometimes has manic features as well (in which case the diagnosis of hallucinogen-induced mania should be considered). The symptoms of hallucinogen-induced psychosis may be indistinguishable from those of a primary psychotic disorder. Sudden onset in the context of recent history of hallucinogen use strongly suggests the diagnosis. Signs and symptoms of hallucinogen intoxication overlap substantially with those of hallucinogen-induced intoxication, so the diagnosis of hallucinogen-induced psychosis should be used only if the symptoms significantly exceed the expected duration and/or intensity of the intoxication state. The diagnosis of hallucinogen-induced delirium should be used if symptoms occur solely in the course of a delirium. If symptoms persist beyond 4 to 8 weeks, the patient likely has a primary psychotic disorder that was exacerbated or precipitated by hallucinogen use.

Course and Prognosis

Hallucinogen-induced psychoses have a variable course. At one extreme are cases that resemble intoxication with the substance, resolving rapidly and completely. In other cases, symptoms persist for weeks, in spite of aggressive treatment. This pattern is relatively common with PCP-induced psychosis that has been observed to persist for up to 6 weeks but eventually

resolves completely. Finally, there are cases in which the psychosis becomes chronic. In these cases, a diagnosis of primary psychotic disorder is appropriate.

Treatment

Hallucinogen-induced psychoses are treated much like any other acute psychosis. Hospitalization may be necessary to ensure safety. Since the course of illness is unpredictable, treatment should aim to prevent irreversible consequences of illness and preserving relationships and functional capacity. Chronic psychotic disorders with onset during hallucinogen use are treated like any other case of first-break psychosis, except that particular care should be taken to minimize future harms associated with substance use.

PCP. Antipsychotic medications are frequently used to treat PCP-induced psychosis, but their efficacy in this situation has not been established. Benzodiazepines have also been used to control agitation. In animal models, atypical antipsychotics are more effective than typicals in reversing the effects of PCP. The use of drugs that modulate NMDA receptor mediated neurotransmission, such as lamotrigine, glycine, D-serine, and D-cycloserine, as well as group II metabotropic glutamate receptor agonists such as the compound LY354470, has been proposed based on preclinical data. However, there are no clinical trials of these agents in PCP- or ketamine-induced psychosis.

Other Hallucinogens. Treatment of other hallucinogen-induced psychosis does not differ substantially from the treatment of other psychoses. In addition to antipsychotic medications, lithium, anticonvulsants, ECT, benzodiazepines, and antidepressants have been used with reported success, but controlled trials are lacking.

PHENCYCLIDINE- AND OTHER HALLUCINOGEN-INDUCED BIPOLAR DISORDER

The diagnosis of substance/medication-induced bipolar and related disorder is used if the patient develops a manic, hypomanic, or mixed episode during or soon after hallucinogen intoxication. Hallucinogen effects on mood are common but usually transient. Mood may fluctuate rapidly. Symptoms resembling those of major depression, mania, or hypomania may appear, but rarely persist beyond the period of intoxication. Although this syndrome is not well-characterized, manic features are sometimes seen following hallucinogen intoxication particularly in association with psychosis. Since there are not specific data on treatment of this disorder, treatment should ensure safety, and clinicians should consider medications that are normally effective for similar symptoms (e.g., atypical antipsychotics, benzodiazepines, and

possibly mood stabilizers). As with most substance-induced disorders, symptoms are likely to remit within days to weeks if the patient stops using the offending substance. If symptoms persist beyond 4 to 8 weeks, a primary bipolar disorder diagnosis should be considered.

PHENCYCLIDINE- AND OTHER HALLUCINOGEN-INDUCED DEPRESSIVE DISORDER

The diagnosis of substance/medication-induced depressive disorder is used if the patient develops a depressive episode during or soon after hallucinogen intoxication. Depressive symptoms can develop during dysphoric hallucinogen experiences ("bad trips") and usually resolve with the acute effects of the drug. Evaluation and management of suicide risk are imperative, since suicides have occurred under these conditions. Hallucinogen use can also exacerbate a pre-existing mood disorder, but this is unlikely during or immediately following PCP or ketamine intoxication since NMDA receptor antagonists have antidepressant effects.

PHENCYCLIDINE- AND OTHER HALLUCINOGEN-INDUCED ANXIETY DISORDER

The diagnosis of substance/medication-induced anxiety disorder is used if the patient develops an anxiety episode during or soon after hallucinogen intoxication. Clinically significant anxiety (distinct from paranoia or agitation) is rarely if ever reported with PCP or ketamine intoxication, but is common during intoxication with *Salvia* and many of the classic hallucinogens. Full-blown panic attacks can occur. These symptoms are self-limited, but may be very distressing. Patients usually respond well to reassurance and being in a calm and safe environment. Oral or parenteral benzodiazepines may be useful in extreme and/or persistent cases. There is little or no evidence for persisting hallucinogen-induced anxiety disorders.

PHENCYCLIDINE- AND OTHER HALLUCINOGEN-INDUCED INTOXICATION DELIRIUM

The diagnosis of substance intoxication delirium should be used instead of substance intoxication if disturbances in attention and cognition predominate in the clinical picture and are severe enough to warrant clinical attention. Because such symptoms are an integral part of the intoxication syndrome for PCP and related substances, the diagnosis of delirium depends primarily on the severity of the neurocognitive symptoms. This is also true for anticholinergic agents, whose hallucinogenic effects are usually considered to be symptoms of anticholinergic delirium with features typically including disorientation, agitation, grossly impaired cognition and memory, and fluctuating mental status. For most other hallucinogens, delirium is uncommon, but may occur in cases of severe intoxication, or may represent a response to serious medical consequences of intoxication (e.g., in cases of overdose of MDMA

or other nonclassic serotonergic hallucinogens).

Management of substance-induced delirium is primarily focused on correcting any dangerous medical situations and keeping the patient safe while he or she recovers from the effects of the substance. Activated charcoal may be administered in cases of ingestion. Specific issues in the treatment of delirium related to PCP and related hallucinogens are discussed in the section on PCP and related substance intoxication above.

Treatment of anticholinergic delirium involves supportive care and administration of benzodiazepines. Antipsychotics should be avoided. Use of physostigmine is controversial and should be reserved for cases where it is clear that anticholinergics and no other substances have been ingested.

UNSPECIFIED PHENCYCLIDINE-RELATED DISORDER AND UNSPECIFIED HALLUCINOGEN-RELATED DISORDER

This residual category is used to describe situations in which symptoms characteristic of a PCP or other hallucinogen-related disorder predominate and cause clinically significant distress or impairment, but do not meet full criteria for any specific disorder.

REFERENCES

Abraham HD, Aldridge AM. Adverse consequences of lysergic acid diethylamide. *Addiction*. 1993;88(10):1327–1334.

Abraham HD, Aldridge AM, Gogia P. The psychopharmacology of hallucinogens. *Neuropsychopharmacology*. 1996;14(4):285–298.

Alper KR. Ibogaine: a review. *Alkaloids Chem Biol*. 2001;56:1–38.

Bokor G, Anderson PD. Ketamine: an update on its abuse. *J Pharm Pract*. 2014;27(6):582–586.

Boutros NN, Bowers MB Jr. Chronic substance-induced psychotic disorders: state of the literature. *J Neuropsychiatry Clin Neurosci*. 1996;8(3):262–269.

Capela JP, Carmo H, Remiao F, Bastos ML, Meisel A, Carvalho F. Molecular and cellular mechanisms of ecstasy-induced neurotoxicity: an overview. *Mol Neurobiol*. 2009;39(3):210–271.

Cohen S. Lysergic acid diethylamide: side effects and complications. *J Nerv Ment Dis*. 1960;130:30–40.

Drug Abuse Warning Network. *2011: Selected Tables of National Estimates of Drug-Related Emergency Department Visits*. Rockville, MD: Center for Behavioral Health Statistics and Quality, SAMHSA; 2013.

Fantegrossi WE, Murnane KS, Reissig CJ. The behavioral pharmacology of hallucinogens. *Biochem Pharmacol*. 2008;75(1):17–33.

Gable RS. Risk assessment of ritual use of oral dimethyltryptamine (DMT) and harmala alkaloids. *Addiction*. 2007;102(1):24–34.

Griffiths RR, Johnson MW, Richards WA, Richards BD, McCann U, Jesse R. Psilocybin occasioned mystical-type experiences: immediate and persisting dose-related effects. *Psychopharmacology (Berl)*. 2011;218(4):649–665.

Griffiths RR, Richards WA, McCann U, Jesse R. Psilocybin can occasion mystical-type experiences having substantial and sustained personal meaning and spiritual significance. *Psychopharmacology (Berl)*. 2006;187(3):268–283; discussion 284–292.

Gudelsky GA, Yamamoto BK. Actions of 3,4-methylenedioxymethamphetamine (MDMA) on cerebral dopaminergic, serotonergic and cholinergic neurons. *Pharmacol Biochem Behav*. 2008;90(2):198–207.

Halberstadt AL. Recent advances in the neuropsychopharmacology of serotonergic hallucinogens. *Behav Brain Res*. 2015;277:99–120.

Halberstadt AL, Geyer MA. Serotonergic hallucinogens as translational models relevant to schizophrenia. *Int J Neuropsychopharmacol*. 2013;16(10):2165–2180.

Halpern JH, Pope HG Jr. Do hallucinogens cause residual neuropsychological toxicity? *Drug Alcohol Depend*. 1999;53(3):247–256.

Halpern JH, Pope HG Jr. Hallucinogen persisting perception disorder: what do we know after 50 years? *Drug Alcohol Depend*. 2003;69(2):109–119.

Johnson M, Richards W, Griffiths R. Human hallucinogen research: guidelines for safety. *J Psychopharmacol*. 2008;22(6):603–620.

Johnson MW, MacLean KA, Reissig CJ, Prisinzano TE, Griffiths RR. Human psychopharmacology and dose-effects of salvinorin A, a kappa opioid agonist hallucinogen present in the plant Salvia divinorum. *Drug Alcohol Depend*. 2011;115(1-2):150–155.

Johnston LD, Miech RA, O'Malley PM, Bachman JG, Schulenberg JE. *Use of Alcohol, Cigarettes, and Number of Illicit Drugs Declines Among U.S. Teens*. Ann Arbor, MI: University of Michigan News Service; 2014.

Kerridge BT, Saha TD, Smith S, et al. Dimensionality of hallucinogen and inhalant/solvent abuse and dependence criteria: implications for the Diagnostic and Statistical Manual of Mental Disorders—Fifth Edition. *Addict Behav*. 2011;36(9):912–918.

Krebs TS, Johansen PO. Psychedelics and mental health: a population study. *PLoS One*. 2013;8(8):e63972.

Lerner AG, Gelkopf M, Skladman I, et al. Flashback and Hallucinogen Persisting Perception Disorder: clinical aspects and pharmacological treatment approach. *Isr J Psychiatry Relat Sci*. 2002;39(2):92–99.

Litjens RP, Brunt TM, Alderliefste GJ, Westerink RH. Hallucinogen persisting perception disorder and the serotonergic system: a comprehensive review including new MDMA-related clinical cases. *Eur Neuropsychopharmacol*. 2014;24(8):1309–1323.

Maciulaitis R, Kontrimaviciute V, Bressolle FM, Briedis V. Ibogaine, an anti-addictive drug: pharmacology and time to go further in development. A narrative review. *Hum Exp Toxicol*. 2008;27(3):181–194.

Morgan CJ, Muetzelfeldt L, Curran HV. Consequences of chronic ketamine self-administration upon neurocognitive function and psychological wellbeing: a 1-year longitudinal study. *Addiction*. 2010;105(1):121–133.

Murray RM, Paparelli A, Morrison PD, Marconi A, Di Forti M. What can we learn about schizophrenia from studying the human model, drug-induced psychosis? *Am J Med Genet B Neuropsychiatr Genet*. 2013;162B(7):661–670.

Nichols DE. Hallucinogens. *Pharmacol Ther*. 2004;101(2):131–181.

Nutt DJ, King LA, Phillips LD. Drug harms in the UK: a multicriteria decision analysis. *Lancet*. 2010;376(9752):1558–1565.

Reissig CJ, Carter LP, Johnson MW, Mintzer MZ, Klinedinst MA, Griffiths RR. High doses of dextromethorphan, an NMDA antagonist, produce effects similar to classic hallucinogens. *Psychopharmacology (Berl)*. 2012;223(1):1–15.

Studerus E, Kometer M, Hasler F, Vollenweider F. X. Acute, subacute and long-term subjective effects of psilocybin in healthy humans: a pooled analysis of experimental studies. *J Psychopharmacol*. 2011;25(11):1434–1452.

Substance Abuse and Mental Health Services Administration. *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014.

Tyls F, Palenicek T. Psilocybin—summary of knowledge and new perspectives. *Eur Neuropsychopharmacol*. 2014;24(3):342–356.

Vollenweider FX, Kometer M. The neurobiology of psychedelic drugs: implications for the treatment of mood disorders. *Nat Rev Neurosci*. 2010;11(9):642–651.

Wu LT, Ringwalt CL, Weiss RD, Blazer DG. Hallucinogen-related disorders in a national sample of adolescents: the influence of ecstasy/MDMA use. *Drug Alcohol Depend*. 2009;104(1-2):156–166.

▲ 11.7 Inhalant-Related Disorders

MATTHEW O. HOWARD, PH.D., SCOTT E. BOWEN, PH.D., AND ERIC L. GARLAND, PH.D.

Inhalants, sometimes called volatile substances or solvents, are volatile hydrocarbons, which become gases at room temperature. These compounds are used in many household products as solvents, propellants, thinners, and fuels. For example, toluene, one such volatile hydrocarbon, is a key ingredient in some glues. When a glue is pushed from its tube, the toluene volatilizes into a gas and the glue will set or harden; conversely, to remove a hardened glue a solvent (containing toluene or a similar compound) may be used. The term *inhalants* includes a wide range of volatile hydrocarbons, such as toluene, xylene, hexane, trichloroethylene, trichloroethane, and gasoline, among many others. These compounds are grouped together under the term inhalants because they are believed to share certain pharmacological properties.

The strengths and limitations of different approaches to classifying inhalants have been examined recently, with taxonomies based on product type or form being discouraged and classification based on some not-yet-determined combination of pharmacological and chemical properties and shared patterns of misuse being encouraged. At present, however, data are too scant to support an evidenced-based classification that is more refined than that embodied in *Diagnostic and Statistical Manual of Mental Disorders,* Fifth edition (DSM-5). Because of epidemiological and pharmacological differences, DSM-5 excludes from the inhalant-related disorders conditions related either to anesthetic gases (e.g., nitrous oxide) or to amyl, butyl, and isobutyl nitrites, classifying these as other (or unknown) substance-related disorders; some disorders associated with those compounds are briefly discussed at the end of this chapter. Although fumes of combustible drugs such as crack cocaine and tobacco also are inhaled, DSM-5 similarly places disorders related to those drugs in separate categories.

Individuals, especially adolescents, inhale these products for their intoxicating effects. Among US eighth graders, inhalants are among the most commonly used illicit drugs, with about 11 percent reporting lifetime use, a prevalence rate higher than those of all other illicit drugs except marijuana. However, with increasing age, prevalence of past-year inhalant use declines; most adolescents discontinue their inhalant use and few people ever meet criteria for inhalant use disorder. However, inhalants are associated with multiple problems, including conduct disorder, noninhalant SUDs, mood, anxiety and personality disorders, suicidality, and a history of physical/sexual abuse and neglect victimization; for some

adolescents, an early time-limited use of inhalants may signal a lifelong problem with externalizing behaviors, novelty seeking, impulsivity, and risk-taking propensity. A smaller subgroup use inhalants chronically; such use has been associated with multiple sequelae, including major behavioral problems and toxicity-related organ pathology.

DEFINITION

The section on inhalant-related disorders in the DSM-5 includes three relevant categories. The first category is inhalant use disorder, which is characterized by clinically significant impairment and/or distress due to a problematic pattern of volatile hydrocarbon inhalation. The second category, inhalant-induced disorders, includes disorders that result from the toxic effects of inhaled substances. Inhalant-induced disorders include inhalant intoxication, inhalant intoxication delirium, and inhalant-induced psychotic, depressive, anxiety, and major and mild neurocognitive disorder. Inhalant-induced disorders are described below.

The third category of inhalant-related disorders in DSM-5 is unspecified inhalant-related disorder. This diagnosis is reserved for persons presenting with signs of an inhalant-related disorder who are experiencing clinically significant distress and/or functional impairment as a result, but who do not meet full diagnostic criteria for any specific inhalant-related disorder or any other DSM-5 substance-related and addictive disorder.

COMPARATIVE NOSOLOGY

The diagnostic criteria for and description of inhalant use disorder, inhalant intoxication, and unspecified inhalant-related disorder are included in the inhalant-related disorders section of the substance-related and addictive disorders chapter of DSM-5. The diagnostic criteria and descriptions of inhalant-induced disorders (other than inhalant intoxication) are specified in the DSM-5 sections that specifically address the major symptoms. For example, inhalant-induced depressive disorder is included in the substance/medication-induced depressive disorders section of the DSM-5 depressive disorders chapter.

The 10th revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) refers to inhalants as "volatile solvents." ICD-10 describes the following "mental and behavioral disorders due to use of volatile solvents": acute intoxication, harmful use, dependence syndrome, withdrawal state, psychotic disorder, amnestic syndrome, residual and late-onset psychotic disorder, and other and unspecified mental and behavioral disorders. DSM-5 provides no diagnosis of inhalant withdrawal, whereas ICD-10 includes that diagnosis, but gives no diagnostic criteria.

EPIDEMIOLOGY

The impact on public health from inhalant misuse may be indicated by the prevalence of each of these: inhalant use, DSM-5-defined inhalant use disorder, inhalant-related visits to hospital emergency departments, inhalant-related admissions to drug treatment programs, and deaths associated with inhalant use.

Prevalence of Inhalant Use

The US government sponsors three large nationally representative studies of drug use. The Monitoring the Future (MTF) Survey assesses 8th, 10th, and 12th graders and the Youth Risk Behavior Survey (YRBS) assesses 9th to 12th graders, whereas residents 12 years of age and older living in US households are assessed in the National Survey on Drug Use and Health. Because of differences in sampling and interview methods, the two school surveys often show higher prevalence rates of inhalant use among adolescents compared to the household survey, although the two survey methods usually agree well on trends over time and in their ranking of different drugs' prevalence of use. The surveys provide important estimates regarding patterns and prevalence of drug use.

More US eighth graders report use of inhalants than report use of any other illicit drug except marijuana. In 2014, about 1 in 9 US eighth graders reported inhalant use at least once in their lifetime; in contrast, about 1 in 50 reported cocaine use. In 2012, an estimated 650,000 youth between 12 and 17 years of age reported past-year inhalant use and more than 21 million US residents 12 or older reported lifetime inhalant use. Between 2005 and 2014, lifetime and annual inhalant use prevalence rates in the MTF survey declined notably. The perceived harmfulness of regular inhalant use increased between 2010 and 2014. The prevalence of inhalant use among adolescents is generally similar between males and females; Caucasians, American Indians/Alaska Natives, and bi- or multiracial youth tend to have higher prevalence of use, whereas African Americans and Asians tend to have lower rates.

YRBS findings in 2013 indicated that 8.9 percent of youth had used inhalants and that rates of lifetime use were higher among girls than boys and higher among Hispanic than white and black students. Between 1995 and 2013, a significant linear decline in the prevalence of lifetime inhalant use was noted. Inhalant use prevalence rates varied significantly across the states and large urban school districts participating in the YRBS.

Obviously, if most people who try a drug continue to use it frequently, the past-month prevalence rate should approach the lifetime rate, and that does not happen with inhalants. Inhalant use occurs mainly in younger age groups, with the prevalence of past-year inhalant use falling from a peak of around 3.5 to 4.5 percent at age 14 to less than 1 percent for all adult inhalant users. Most new users do not endorse regular inhalant use, and more than half report only 1 to 3 days of use in the prior 12 months.

The prevalence of inhalant problems in small, isolated groups or communities and in high-risk populations including juvenile and criminal justice system-involved persons may deviate from larger national trends. For example, certain Native American communities experience a high prevalence of these problems, with reported lifetime rates as a high as 34 percent by eighth grade. However, rates of use may vary across communities. For example, in one large recent general population study of two Native American communities assessing persons 15 and older, prevalence of inhalant use among females from a Southwest tribe was 3.6 percent and for females from two Northern plains tribes was 10.6 percent.

Studies of juvenile probationers, incarcerated youth in short- and long-term residential treatment, and youth and adults treated for behavioral problems or other noninhalant substance dependencies have reported lifetime inhalant use prevalence rates ranging from 18 percent to nearly 40 percent.

In some countries, many children live on the street in loose-knit juvenile gangs with little or no attachment to relatives. Among such children, inhalant use is very common with prevalence rates reportedly exceeding 70 percent in some cities.

Prevalence of Inhalant Use Disorder

Prevalence of inhalant *use* should not be confused with prevalence of DSM-5-defined inhalant use disorder. Most users never progress to formal inhalant use disorder. Nationally representative surveys suggest that inhalant use disorder is relatively rare. For example, a national household survey assessed DSM-IV-TR-defined past-year inhalant use disorders. From this survey about 0.1 percent of the US population (12 years of age or older) met criteria for an inhalant use disorder in any given 12-month period. As might be predicted from information available regarding inhalant use patterns, the prevalence of past-year DSM-IV-TR inhalant abuse and dependence is somewhat higher (i.e., 0.4 percent) among youth 12 to 17 years of age. In 2012, the household survey noted that an estimated 34,000 persons in the United States met DSM-IV-TR criteria for inhalant dependence disorder.

Another survey of about 8,000 randomly sampled Americans (15 to 54 years of age) found that 0.3 percent had been dependent on inhalants in their lifetimes (as defined by DSM-III-R); again, rates of inhalant dependence (0.6 percent) were higher among younger age groups (15 to 24 years of age). In that study, most people who had tried inhalants had not become dependent on them; only about 1 in 27 respondents who reported trying inhalants met criteria for lifetime dependence (1 in 13 of individuals aged 15 to 24); in contrast, about 1 in 3 individuals reporting nicotine-use met criteria for lifetime dependence on that substance. Although prevalence of inhalant use is similar across genders in most studies,

inhalant dependence was about four times more likely among males, compared with females.

A more recent nationally representative study of over 43,000 respondents 18 years of age or older found similarly low past-year (0.02 percent) and lifetime (0.3 percent) prevalence rates of inhalant use disorders (DSM-IV-TR-defined), with most of those individuals meeting criteria for abuse; inhalant abuse and inhalant dependence prevalence rates were the lowest among the drug categories reported (i.e., alcohol, sedatives, tranquillizers, opioids, amphetamines, hallucinogens, cannabis, and cocaine). In that study about one-in-five lifetime inhalant users met criteria for a lifetime inhalant use disorder, although prevalence of lifetime inhalant use was relatively low (1.7 percent) compared to most other studies. Lifetime inhalant use disorders were about five times more prevalent among men than women.

One large study of two Native American communities recently found lifetime prevalence of inhalant use disorders of 1.7 percent for males and 0 percent for females for a Southwest tribe and 0.4 percent for males and 0.5 percent for females among two Northern plains tribes. These results underscore the variability in prevalence of inhalant use disorders by gender and across Native American communities.

Limited data are currently available that help to identify which inhalant users are likely to progress to formal inhalant-related disorders. One national cross-sectional survey found that living in a large metropolitan area, previous use of mental health services, number of delinquent behaviors, multiple other drug use disorders, and number of different inhalants used were associated with a DSM-IV-TR diagnosis of inhalant abuse among past-year inhalant users. Use of mental health services, history of foster care placement, multiple other drug use disorders, weekly use of inhalants, and early age at first inhalant use were associated with inhalant dependence among past-year inhalant users.

Several studies of high-risk populations have identified high rates of inhalant use disorders among inhalant users. One statewide survey of incarcerated youth reported that 47 percent of inhalant users met criteria for lifetime DSM-IV-TR inhalant abuse or dependence. Similarly, a study of 162 young adult inhalant users in St. Louis, Missouri reported that 47 percent met criteria for DSM-IV-TR inhalant abuse or dependence. Another large study of youth in treatment of substance use and behavioral disorders, reported that 18.3 percent of inhalant users met criteria for DSM-IV-TR-defined inhalant use disorders. Thus, while inhalant use disorder may be uncommon in the general population, high levels of inhalant use and inhalant use disorder may prevail in correctional and institutional settings serving at-risk youth and adults.

Prevalence of Inhalant-Related Emergency Department Visits

The US government monitors drug-related emergency department visits in hundreds of non-Federal, short-stay, general medical and surgical hospitals around the country. In 2011, among approximately 2.5 million drug-related emergency department visits, 51 percent which involved illicit drugs, inhalants (including nitrous oxide and nitrites and excluding accidental exposures) were implicated in an estimated 10,032 visits (0.8 percent of ED visits involving illicit drugs); in contrast, there were 505,244 cocaine- and 455,668 marijuana-related emergency department visits in 2011.

Prevalence of Inhalant-Related Treatment Admissions

The US government collected data on 1,800,704 admissions to US substance abuse treatment programs during 2012. The number of admissions with inhalants listed as the primary substance of abuse was 1,137, constituting 0.06 percent of treatment admissions that year. Approximately 61 percent of inhalant admissions were male, 39 percent were female. Caucasians accounted for 75.3 percent of admissions, African Americans for 6.0 percent, American Indians/Alaska Natives for 6.9 percent, Asians/Hawaiians for 0.7 percent, "other" races for 8.6 percent, and unknown race for 2.5 percent. Approximately one-third of all admissions were for persons age 20 or younger, whereas 39.7 percent of admissions occurred in persons age 31 or older. These findings indicate that, although inhalant use is generally considered a problem of childhood or adolescence, most persons treated for primary inhalant use disorders are adults.

Prevalence of Inhalant-Related Deaths

Inhalant-related deaths have been reported from direct toxic effects of the drug, aspiration, suffocation, flash burns, and accidental injury, among many other causes. Inhalant use is thought to increase myocardial cell depolarization, reducing conductivity and predisposing to arrhythmia, especially when paired with increased blood concentrations of epinephrine; this has sometimes been termed *sudden sniffer's death*.

Two sources help in understanding the extent of inhalant-related mortality in the United States. First, medical examiner reports are available from numerous jurisdictions and metropolitan areas and a handful of states. An investigation of 39 inhalant-related deaths in Virginia between 1987 and 1996 found that 95 percent of decedents were male, 70 percent were age 22 or younger (median age = 19), and 95 percent were Caucasian. The most prevalent inhalants identified were gas fuels including butane and propane (46 percent), chlorofluorocarbons (26 percent), and chlorinated hydrocarbons and alkylbenzenes (21 percent). A review of 144 inhalant-related deaths in Texas between 1988 and 1998 similarly found that most decedents were male (92 percent), Caucasian (81 percent), and

relatively young (mean age = 25.6 years). In 35 percent of cases, the cause of death was inhalation of Freon, followed in frequency by inhalation of chlorinated hydrocarbons (implicated in 25 percent of deaths), and alkylbenzenes (i.e., toluene) identified in 17 percent of deaths. A recent study of 30 inhalant-related deaths in North Carolina between 2000 and 2008 also found that most decedents were males (81 percent) and Caucasian (93 percent). However, decedents were somewhat older compared to previous studies (mean and median ages were 31.2 and 28.0, respectively); further, deaths due to "air duster" abuse (i.e., 1,1-difluoroethane) accounted for nearly one-third of all cases and this agent was the second most commonly identified inhalant following toluene. Finally, the Florida State Medical Examiners Commission identified 46 inhalant-related deaths in 2011 and reported that 72 percent of these cases involved the agent 1,1-difluoroethane. In 2010, 38 persons in Florida were identified as having inhalant-related deaths, 22 due to 1,1-difluoroethane use. Taken together, these medical examiner reports suggest that adult Caucasian males are at greatest risk of inhalant-related fatalities. The North Carolina study cited above also noted that most decedents were heavy polydrug users with significant health and mental health impairments.

Second, information is available from calls to poison control centers nationally in the United States. The most comprehensive investigation examined all inhalant-related reports made to 60 poison control centers between 1983 and 2008. The 35,453 inhalant cases reported involved more than 3,400 different products. The highest rate of cases occurred among youth ages 12 to 17, peaking at age 14. Nearly three-quarters of all cases were boys. The most frequently identified inhalants were propellants, gasoline, and paint. Propellant-related cases (e.g., "air dusters") increased significantly over the study period, particularly since 2003, and constituted the largest-single category of inhalant-related cases. The three inhalants most associated with fatal outcomes were butane, propane, and air fresheners.

In the United States, there are no national estimates of inhalant-related mortality, and the available data likely contain numerous biases. For example, in many instances poison control centers may never be contacted regarding an inhalant-related death, and the sometimes nonspecific findings following an inhalant-related death may make proper identification of the cause of death difficult.

In the United Kingdom, a systematic study of inhalant-related mortality (i.e., 2,093 deaths) between 1983 and 2007 concluded that the number of deaths during this interval had declined substantially, particularly among women, and that the mean age at death had increased by more than a decade in males and females. Particularly notable decreases in inhalant-related deaths were observed for boys and girls following a 1992 federal

advertising campaign that warned parents that their children might be misusing inhalants. The number of inhalant-related deaths peaked in the early 1990s, and declined to their nadir in the mid-2000s. In 2005, 45 inhalant-related deaths (9 which were suicides) were reported. This was the lowest number in the prior 20 years, and this reduction may relate to a parent-focused advertising campaign and/or a ban on the sale of cigarette lighter refills to those under the age of 18. The percentage of glue-related inhalant deaths declined during this 25-year-interval, whereas the percentage of deaths due to gas fuels (i.e., butane and propane) increased significantly. In the United Kingdom, butane was implicated in 80 percent of such deaths. In four of ten of these deaths, there was no evidence of previous inhalant use, and the majority of deaths occurred at home. Inhalant-related deaths were predominantly male. Among 10- to 14-year-olds, there were more inhalant-related deaths than deaths associated with other drug misuse (as defined by the Drug Misuse Act of 1971, including cocaine, cannabis, LSD, and opioids, among others).

A 15-year-old Caucasian boy was admitted to an adolescent day treatment program. His biological father was identified as a drug dealer, was abusive to his biological mother, and went to prison for robbery; his parents separated when he was only a few months old, and the patient did not recall ever meeting his father. His mother remarried when he was 12. Reportedly, his stepfather would "get smashed," hit his mother, get very close to the patient's face, and "yell at me for no reason." His mother was in the process of divorcing his stepfather. At age 10, the patient began using marijuana, reporting that it helped him "get away" and reduce his anger, and his use rapidly progressed to using "all day every day." He subsequently tried alcohol, cocaine, and prescription medications (i.e., benzodiazepines and opioids).

The patient reported a history of delinquent behaviors. These behaviors included stealing from his mother's purse, stealing her prescription medicines, pawning his stepfather's possessions, truancy, and initiating fights. In sixth grade, he received a drug paraphernalia ticket from the police, although this charge was later reduced to minor in possession of alcohol. He was later charged with accessory to aggravated robbery.

The patient was placed on probation, and because of his history of substance involvement, was admitted to a short-term outpatient, mainly group-based treatment program. He was discharged from that program after a few sessions for an inability to remain drug free and suspected use of his mother's prescription pain medication.

After 3 weeks in juvenile detention he was admitted to a day-treatment program. A thorough intake evaluation revealed a history of cannabis, cocaine, alcohol, and prescription drug use disorders, no other mental health concerns, and no relevant medical issues. The client did well in treatment, attending regularly, receiving above average scores for his behavior, completing his weekly school contracts and all of his initial treatment-related assignments, submitting three times a week to observed urine samples (tested for alcohol, opiates, benzodiazepines, PCP, cocaine, cannabis, and amphetamines), and attending individual and family therapy. Reports from all observers indicated the client was making progress in treatment. The patient expressed pride in being drug free and staying away from negative peers and began discussing future-oriented prosocial goals. After 1 month in treatment, the patient was progressed to higher privileges, recognizing his treatment progress.

The following week, the client was found by his mother submersed in his bathtub. He had been "huffing" (i.e., inhaling) a common spray used to remove dust from computers (i.e., "air duster," containing 1,1-difluoroethane) and other surfaces. Although attempts were made to revive him, he subsequently died.

This case demonstrates several important points about inhalant use: an association with other substance use and conduct disorder, adolescent onset, and the risk of inhalant-related mortality.

Comorbidity

Inhalant users have a high prevalence of multiple comorbid problems. For example, in nationally representative studies of the US population, respondents reporting any inhalant use were significantly more likely to have injected drugs; injection, in turn, is a risk factor for infection with human immunodeficiency virus (HIV) and hepatitis B and C viruses (HBV/HCV). Inhalant use, especially among persons also reporting cannabis use, is a marker for the elevated risk of other SUDs, and inhalant users are more likely to report delinquent behaviors and meet criteria for conduct disorder than inhalant nonusers. In one nationally representative study of adults, of those reporting inhalant use before age 18, more than 40 percent met criteria for antisocial personality disorder. In that same study, inhalant users showed high rates of mood (48 percent) and anxiety (36 percent) disorders, especially among women. A study of 664 nationally representative adult inhalant users found that 40 percent met criteria for 3 to 5 DSM-IV-TR SUDs, whereas nearly one quarter (i.e., 22 percent) met criteria for 6 to 11 such disorders. Virtually all (94 percent) adult inhalant users met DSM-IV-TR criteria for an alcohol and/or noninhalant drug use disorder. Studies of nationally representative and juvenile justice samples support a link between inhalant use/inhalant use disorders and suicidality. Several studies also support an association between inhalant use and a history of child abuse and/or neglect victimization. Although multiple explanations are possible, the etiology underlying this comorbidity has not been well studied; exploratory analyses suggest that the link between early inhalant use and later injection drug use may be explained by a general predisposition to early-onset drug use.

Summary

Trying inhalants at least once is fairly common around puberty and into adolescence, but continued frequent inhalant use is uncommon and the prevalence of at least monthly use decreases with increasing age. Desistance rates by young adulthood are higher for the inhalants than for any other psychoactive drug class. Accordingly, the limited data suggest that inhalant use disorder is rare in the general population. Reflecting that low prevalence in the general population, inhalant-related problems are the primary contributor to relatively few substance abuse program treatment admissions, account for a small proportion of drug-related emergency department contacts, and are implicated in a small proportion of all drug-related deaths. One caveat, however, is that elevated rates of inhalant use

disorders are often characteristic of incarcerated and institutionalized youth and adult populations and these populations are typically not reflected in epidemiological studies. Similarly, homeless youth and adults, persons living on the street or transiting between different geographic areas, and other marginalized and "hidden" populations may evidence higher levels of inhalant use and inhalant use disorder than are reflected in epidemiological studies of general and clinical populations. Inhalant users also report numerous barriers to treatment utilization and are commonly coping with multiple comorbid psychiatric and SUDs and impoverished social circumstances that may make it difficult for them to access health care.

The available data suggest that most adolescents who survive a brief experimentation with inhalants and who do not have conduct disorder soon abstain from inhalants and avoid severe pathological outcomes. However, among adolescents with conduct disorder, perhaps half develop adult antisocial personality disorder. Inhalant problems in these youths often herald serious alcohol and polysubstance use in adulthood, and a small percentage of these youths become chronic, deteriorated, inhalant-dependent adults.

PHARMACOLOGY AND TOXICOLOGY

National surveys have consistently reported that the inhalants most frequently used by individuals within the United States are products containing toluene (e.g., glue and shoe polish), gasoline, lighter fluid (which usually contains butane), spray paints, aerosols, degreasers, and correction fluids. Administration of these inhalants is simple and unsophisticated and results in rapid delivery to the lungs and brain of high concentrations of the volatile product. Among the most common methods of administration are breathing vapors from products placed in a paper or plastic bag or poured into a balloon; inhaling from solvent-soaked clothing, rags or cloth; sniffing the product after placement on the fingernails; huffing vapors directly from a container (e.g., paint can); inhaling vapors from cotton balls or cotton swabs placed into the nose or mouth; or breathing vapors from an inverted aerosol/spray can. Abuse usually involves 15 to 20 deep breaths, which results in a rapid onset of intoxication. Rebreathing the exhaled air may enhance the effect if the breath is into a bag containing the inhalant. While the exact solvent concentration inhaled varies by compound, estimates for one of the most commonly abused inhalants, toluene, vary from 5,000 to 15,000 ppm. This range is several magnitudes above the industrial maximum of 100 ppm for occupational exposure, a concentration that can produce temporary performance decrements after 6 to 8 hours of exposure. The intoxication achieved has been reported to continue from 15 to 60 min and can be sustained for hours by either continuous or repeated vapor inhalation.

Tolerance appears to develop to the intoxicating effects as it has been reported that individuals will increase the number of tubes of glue they inhale across days of abuse.

Toxicokinetic data for inhalants like toluene indicate that inhaling these substances results in very rapid absorption through the lungs that produce a rapid rise in blood concentrations. Toluene and other inhalants can also be absorbed through the skin and the GI tract that increases total body burden. Both animal and human studies have shown a positive correlation between toluene levels in alveolar air and blood levels. Toluene is absorbed by both red blood cells and plasma that enhance toluene transportation to all areas of the body, especially to tissues that are high in lipids and well vascularized (e.g., brain). In one study in which humans were continuously exposed to 80 ppm of toluene for 4 hours, toluene blood levels peaked (6 to 7 μmol/L) at approximately 2 hours. Data obtained from an autopsy of a 16-year-old boy who died after inhaling glue show that blood and brain toluene concentrations were 20.6 μg/mL and 297 μg/mg, respectively. In another case, a man died after falling while painting and being exposed to toluene. The reported toluene levels in his blood and brain were 48 and 80 μg/g, respectively. Toluene blood levels can increase significantly with coadministration of ethanol, which can dramatically increase levels of toxicity.

Autoradiography studies have demonstrated high levels of radioactivity in body fat, the spinal cord and spinal nerves, as well as the white matter of the brain in mice exposed to toluene via inhalation. Lower levels were found in the blood, liver, and kidneys, which indicate that toluene metabolites are rapidly formed and excreted. Both human and rat studies have shown that approximately 80 percent of inhaled toluene is absorbed by the body with increased activity increasing the amount of toluene absorbed. Unabsorbed toluene is expelled unchanged in breathe and urine with the liver metabolizing absorbed toluene into hippuric acid and other metabolites that are excreted in the urine. While a person's age, sex, race, and health status influence the rate of toluene elimination from the body, toluene concentrations in breath begin falling within minutes of exposure, while concentrations in the blood decrease at a much slower rate taking anywhere from 4 to 10 hours before being undetectable. Elimination of hippuric acid in urine takes longer and anything greater than a 1 g hippurate per gram creatinine ratio has been suggested as a marker for inhalant abuse involving toluene, although occupational exposure to toluene and some food preparations may lead to hippurate/creatinine ratios greater than 1.

Understanding of the cellular mechanisms and targets of abused inhalants has recently improved because of increased in vitro/in vivo investigations of these compounds. Unfortunately, the number of inhalants investigated remains small with toluene receiving the most study. Toluene,

like many abused drugs, changes dopamine levels in the brain. Several studies have shown that toluene exposure increases dopamine in the striatum and temporarily increases neuronal activity within the VTA. Toluene exposure also dramatically increases dopamine levels in the prefrontal cortex. Repeated toluene exposure increases the amount of dopamine within the caudate and nucleus accumbens. While the neurochemical mechanisms responsible for the increased dopaminergic activity are unclear, both the GABA and glutamatergic systems have been explored. Toluene affects both ligand-gated and voltage-gated ion channels at concentrations comparable to ranges that are voluntary inhaled during abuse episodes. Although toluene has a basic chemical structure (i.e., methylbenzene), it appears to be rather selective in its neurochemical actions. In vitro studies have shown that toluene will concentration dependently inhibit the *N*-methyl-D-aspartate (NMDA) receptor, a specific type of ionotropic glutamate receptor, but have little effect on the α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPA), a closely related non-NMDA ionotropic receptor for glutamate. Increasing NMDA or glycine concentrations could not overcome toluene's inhibition, suggesting that toluene effects are noncompetitive. Continuous long-term toluene exposure has been reported to produce neuronal damage to the hippocampus, whereas repeated exposure to toluene has been shown to produce an upregulation of NMDA receptors within certain parts of the brain. Other abused inhalants, such as 1,1,1-trichloroethane (TCE) and several alkylbenzenes (i.e., m-xylene, ethylbenzene, and propylbenzene), have been reported to inhibit NMDA-induced currents in both a concentration- and subunit-dependent fashion. Thus, acutely, NMDA receptor inhibition seems to be a shared mechanism of action for a number of abused inhalants with continued exposure changing the expression of these glutamatergic receptor subtypes.

Several in vitro reports have demonstrated that, acutely, toluene, TCE and trichloroethylene increase α1β1 $GABA_A$ and α1 glycine receptor function. Conversely, chronic toluene exposure decreases spontaneous GABA currents in hippocampal neurons (as well as increasing $GABA_A$ α1 subunit levels in medial prefrontal cortex while decreasing them in the substantia nigra), suggestive of the development of tolerance to the acute enhancement of $GABA_A$ receptor function. In contrast, flurothyl, a nonabused proconvulsant inhalant, inhibits $GABA_A$ receptors and acts as a noncompetitive antagonist at $GABA_A$ receptors with no effect on α1 glycine receptors. These results suggest that acute behavioral inhibitory or excitatory effects of inhalants associate well with in vitro $GABA_A$ receptor activation or inhibition.

Toluene's neurochemical selectivity has also been reported for nicotinic acetylcholine receptors (nAChRs), where toluene inhibits α4β2 receptors to

a greater degree than α7 nAchRs. Interestingly, another widely abused solvent, perchloroethylene, has been reported to be even more potent than toluene in inhibiting these types of nAChRs. However, it is unclear whether inhibition of nAChRs is a common mechanism of action for other abused inhalants.

Toluene also increases currents through P2X3 receptors of ATP-gated channels but inhibits the function of P2X2 and P2X4 receptors within these same ATP-gated channels. Finally, several in vitro studies have also shown that toluene, TCE, and trichloroethylene will increase 5-hydroxytryptamine ($5\text{-}HT_3$) receptor function. Taken as whole, these reports indicate that abused inhalants have distinct sites of action on individual channel subunits and the differences in regional and anatomical subunit expression appear to be important determinants of inhalant action.

While behavioral actions of inhalants vary in magnitude and duration depending on the inhalant being investigated, research in animals examining effects of exposure to abuse levels of inhalants suggests that many inhalants have a profile of effects similar to CNS depressants like ethanol, barbiturates, and other depressants. At concentrations similar to those in abuse situations, behavioral effects progress from motor excitation to motor impairment, ataxia, loss of righting reflex, and finally anesthesia. Continued exposure to high concentrations results in coma, seizures, and ultimately death by depression of respiration. While many inhalants appear to be depressant in nature, not all abused inhalants share this profile. For instance, several inhaled isoparaffins produce dose-dependent increases in locomotor activity, whereas isoamyl nitrite has no effect on locomotor activity, and other inhalants, such as flurothyl, and amyl-, ethyl, and butyl-acetate decrease locomotor activity. Repeated exposure to toluene and TCE results in sensitization to the motor increasing effects of these inhalants and repeated toluene exposure increases cocaine's motor effects suggesting a similar neuropharmacological target.

While clinical studies rarely report seizures with inhalant abuse, animal models have shown both convulsant and anticonvulsant effects resulting after abuse pattern exposures. Continuous or repeated exposure to several inhalants produces handling-induced convulsions in mice upon termination of exposure. In one study, mice continuously exposed to TCE over several days developed handling-induced seizures and increased sensitivity to pentylenetetrazol-induced convulsions upon termination of TCE exposure. When these mice were exposed to TCE or toluene vapors, or given ethanol, midazolam (Versed), or pentobarbital (Nembutal), the number and severity of seizures were decreased. Hence, physical dependence develops to abuse inhalants and they display cross-dependence with other inhalants as well as CNS depressants.

Other nonabused inhalants, like flurothyl, a $GABA_A$ antagonist, appear to have proconvulsant activity. While the mechanisms of these convulsant

actions are unknown, high concentrations of any inhalant might have nonspecific and global effects on CNS reactivity. Conversely, the anticonvulsant effects of several inhalants at low concentrations have also been reported. When given as a pretreatment, a single toluene injection has been shown to protect against convulsions brought on by several chemical agents including bicuculline, nicotine, NMDA, and picrotoxin but not strychnine. Protective effects have also been demonstrated for inhaled toluene against NMDA-induced seizures. These anticonvulsant effects may be associated with toluene's potentiating effects on $GABA_A$ receptors or its ability to inhibit NMDA and nicotine receptors, and/or to increase dopamine release.

Because several inhalants act as positive modulators on $GABA_A$ receptors, it should come as no surprise that animal research has also shown that acute and repeated inhalant exposure results in anxiolytic and sensory effects as well as changes in learning and memory. In one behavioral paradigm, toluene and TCE increased the number of entries and time spent in the open arms, similar to diazepam (Valium), which suggests that these two inhalants reduce anxiety. Additional support comes from the defensive burying paradigm where toluene and other inhalants decreased the latency to bury an increasing aversive stimulus. Conversely, flurothyl appears to be anxiogenic in that it decreased both the number of open-arm entries and the time spent in open arms in the elevated plus maze paradigm while having no effect in the defensive burying paradigm at concentrations up to 800 ppm. Whether other inhalants share the anxiolytic properties of toluene and whether tolerance or sensitization develops with repeated use remains to be determined.

The reinforcing properties of abused drugs, like inhalants, have been demonstrated in animals. In one study, animals were trained to self-administer toluene and TCE administered intravenously. In addition to self-administration, several conditioned place preference (CPP) studies with toluene have demonstrated that animals will spend more time on the side of a chamber that had been formerly paired with toluene. Similar findings have been reported for mixtures of toluene, benzene, ethyl acetate, and methyl chloride. Interestingly, one recent study demonstrated that γ-vinyl GABA, a selective GABA-transaminase inhibitor, would block the CPP that had been previously established with toluene. Other studies have expanded on these drug-like effects by demonstrating that animals trained to respond on one lever after a saline injection and on another lever after a CNS depressant (ethanol or pentobarbital injection) will respond on the lever paired with the depressant implying that the subjective effect of toluene may be depressant-like. Other animal studies using this assay have shown that toluene shows partial substitution for PCP and amphetamine suggesting that abused inhalants may share subjective effects with a wider range of drugs that was previously expected.

Organ Pathology

Inhalant use significantly increases the potential for injury and/or death. As a number of inhalants have neurochemical effects comparable to other abused drugs with mechanisms of action involving NMDA antagonism and/or GABA stimulation, very high concentrations of the inhalant may result in severe CNS depression. These effects may be magnified in combination with other CNS depressants such as ethanol. In addition to CNS effects, several inhalants have been reported to increase sensitization of the heart to endogenous catecholamines such that any abrupt physical effort or unexpected excitement may lead to ventricular arrhythmias and death. One common practice of putting substances into plastic bags and then putting the bag over the user's head may result in unconsciousness and/or death from suffocation. Death may also result indirectly from aspiration of either the compound or vomit. Fatalities have also been reported as a direct result of the inhalant. These include anoxia, asphyxiation, severe injury to the lungs or allergic reactions, or cardiac dysfunction. Inhalation of gas fuels like butane has been shown to produce laryngeal edema as well as laryngospasm. The severe behavioral and physical impairments associated with inhalant intoxication have resulted in trauma (i.e., falling and car accidents) that produced fatalities. Death has also been reported from fire-related injuries from inhalant combustion as some inhaled compounds are heated before inhalation or a cigarette was used in close proximity to where the inhalant was being misused. Chronic inhalant use has been linked to renal and haptic failure, as well as liver tumors, which can also lead to death.

Chronic abuse of inhalants can have serious repercussions for the functioning of the nervous system, especially the brain. Additionally, longer durations and larger magnitudes of inhalant abuse are positively correlated with neuroimaging and neurobehavioral abnormalities. Significant cognitive and neurological impairments can be observed in these individuals and include personality changes, hallucinations, paranoid psychosis, memory and information-processing deficits, decreased IQ, concentration problems, speech abnormalities, impaired judgment, increased impulsivity, abnormal electroencephalogram (EEG) activity, temporal lobe epilepsy, convulsions, tremors, as well as gait abnormalities. Several imaging and neuropathological techniques have provided evidence of brain abnormalities in inhalant abusers, such as demyelination, brain shrinkage, and white matter damage (leukoencephalopathy). Magnetic resonance imaging (MRI) has shown significant cerebral, cerebellar and brain stem atrophy, diffuse CNS demyelination, sulcal widening, and enlargement of the ventricles in long-term toluene abusers. The greatest amounts of atrophy were located in lipid-rich white matter versus gray matter areas of the brain with several cases displaying a loss of

differentiation between these distinct areas. Additional reports using T2-weighted MRI demonstrated that repeated toluene abuse led to demyelination that was accompanied by significant axonal damage and loss. A number of rodent studies corroborate these findings with chronic high toluene concentrations producing deterioration of the frontal cortex, the hippocampus, and brain stem as well as impaired dendritic outgrowth and white matter damage. Finally, several single-photon emission computed tomography (SPECT) studies have shown that chronic inhalant abusers have significant decreases in regional blood flow to the bilateral prefrontal cortices as well as perfusion abnormalities to other brain areas. Other investigations of individuals occupationally exposed to toluene have demonstrated abnormal cerebral blood flow as well as atrophy in a number of brain areas.

In addition to CNS effects, inhalant abuse can have many other detrimental effects. Long-term toluene abuse can result in tinnitus and hearing loss. Inhalant misuse has been documented to produce cardiovascular damage as well causing cardiovascular sensitivity to endogenous catecholamines. Several studies have reported kidney failure with inhalant misuse, resulting from tubular necrosis. In addition, liver damage, from centrilobular tumors and necrosis, has also been documented. Repeated inhalation of inhalants such as isobutyl nitrite and naphthalene can suppress the production of bone marrow, leading to anemia and leukemia. Chronic inhalation of isopropyl nitrite or toluene can result in nystagmus and vision loss. Pulmonary system damage has also been documented with a number of inhalants resulting in chronic coughing, chest pain, shortness of breath, wheezing, and pneumonia. Other harmful effects include rhabdomyolysis and GI irregularities such as nausea, pain, and vomiting.

Because of the scarcity of studies, it is difficult to know whether individuals can recover neurophysiological functioning after chronic inhalant exposure. As previously described, many of the neuroimaging investigations would suggest that any neurobehavioral impairments observed in these individuals would be most likely be irreversible. Indeed, several studies examining former long-term inhalant abusers were unable to find any significant improvements in MRI measures of brain pathology or neuropsychological measures after 5 to 18 months of abstinence. The lack of improvement in these studies may be due, in part, to the timing of measurement because a few studies have reported that neurobehavioral measurements taken 2 years after users ceased petrol sniffing were significantly improved compared to previous measurements. While full recovery was not attained, those individuals with the most impairment exhibited the greatest improvements.

With prevalence rates for inhalant use being nearly equivalent for teenage males and females, there is concern that females may be abusing

inhalants during their child-bearing years. Indeed, a recent review of drug misuse among women of reproductive age included toluene-based solvents among commonly used psychoactive substances during pregnancy. An increasing number of reports of fetal developmental disorders and malformations in the offspring of women who abused inhalants like toluene during pregnancy are found in the medical literature. Several of these clinical embryopathy reports contend that these infants display symptoms of *fetal solvent syndrome (FSS),* so named because of their similarities to infants born to women abusing alcohol during pregnancy and exhibiting signs of fetal alcohol syndrome. Women who abused inhalants during their pregnancies reportedly had infants who were born prematurely and delayed in growth. Additionally, some of these infants displayed physical abnormalities such as reduced head sizes, hypoplastic midfaces and jaws as well as lower set ears, shortened palpebral fissures, thin upper lips, and deep-set eyes. Other anomalies reported were down-turned corners of the mouth, abnormal hair patterns on the scalp, small fingernails with blunted finger tips, abnormal palmer creases, and urinary tract deformities. Fetal death was also reported in women with inhalant abuse histories where maternal exposure was very high during pregnancy. Follow-up studies with these FSS infants have demonstrated continued developmental delays and impairments such as marked growth retardation, cerebellar dysfunction, hyperactivity, language impairments, and elevated levels of aggression. It is unclear whether these delays and deficits will last into adulthood as there are no documented clinical remedies for the effects of FSS exposure nor are there follow-up studies of these affected children as they grow into adults. The accompanying discrimination may prevent many inhalant-misusing women from accessing needed health care services while pregnant. Preclinical studies that mimic the high-dose binge-type inhalant use pattern during pregnancy are lending support to an FSS. These animal studies have shown that the high toluene levels that are found in maternal tissues and organs are also occurring in the placenta and fetal tissue including the brain. These high levels of toluene in rodents reportedly produce a number of adverse outcomes including physical malformations, growth retardation, delayed skeletal ossification, deficiencies in coordination and forelimb strength, impairment in memory recall tasks, and alterations in amphetamine-induced locomotor behavior. Whether these delays and impairments last into adulthood or whether it is possible to mitigate these delays and impairments is unclear.

ETIOLOGY

In qualitative research, inhalant users report three main use scenarios: (1) using on rare occasions to have fun, (2) using because alternative substance(s) could not be used (because of no access or urine drug

monitoring), and (3) using in day-long binges. Recent efforts have been made to classify different subtypes of inhalant users. In one study, some inhalant users used many different inhalants, had high levels of psychopathology (i.e., depression, anxiety, externalizing behaviors, and substance use), suffered serious family problems, had high levels of affective instability, and utilized inhalants to help cope with those feelings or problems. Other inhalant users appeared to have lower levels of these problems, and most of those individuals had only used one type of inhalant; these individuals were hypothesized to be "experimenters." A third group appeared to use inhalants because they were easy to obtain. Another taxonomic investigation identified three subtypes of adolescent inhalant users on the basis of the social contexts in which they used inhalants. Youth in the class that most tended to use inhalants with friends to celebrate and have a good time and who had difficulties in refusing invitations to engage in inhalant use, or who socialized with groups of youth who used inhalants, had the highest levels of substance use and related problems, psychiatric distress and psychiatric symptoms, and tended to be more fearless and impulsive, than inhalant users whose use was less dependent on peer and group contextual factors.

Inhalant-related disorders have a multifactorial etiology. First, availability is important in determining the prevalence of drug use disorders. Inhalants are cheap, available in several forms in most households, easily concealed, legal to possess, and simple to take. Second, inhalant use apparently is rewarding, both through direct pharmacological action and through the drugs' social effects. As discussed above, under certain circumstances animals repeatedly self-administer inhalants, showing that these substances have inherently reinforcing properties. In addition, adolescents usually gather in small groups to use inhalants, and being a user facilitates entry into the group, socially reinforcing the use. Third, inhalant users often can evade detection or punishment by parents or school authorities, since the drugs quickly produce a high that passes within a few hours. Fourth, inexpensive inhalants may be one of the few exciting and novel experiences available to youths in impoverished and/or isolated communities that provide few other stimulating experiences. This circumstance may help to explain the high prevalence of inhalant use on some US and Canadian Native American reservations and among aboriginal Australians.

In addition to those extrinsic factors, intrinsic factors also contribute to inhalant-related problems. An impulsive and risk-taking temperament may lead some persons to seek the excitement and danger of inhalant intoxication. Persons with adolescent conduct disorder or antisocial personality disorder are prone to taking extreme risks, and many inhalant users have those disorders. Several studies suggest an association of inhalant use and conduct problems. Among youths in grades 7 through 12,

inhalant users (compared with others who used no drugs or who used only cannabis or alcohol) had many characteristics suggesting conduct disorder. They accepted cheating more readily, admitted to more stealing, perceived less objection to drug use from their families, liked school less, and reported more sadness, tension, anger, and a feeling of being blamed by others. In addition, school surveys showed that inhalant users were more likely to be involved with other drugs. Similarly, among youths referred to court-mandated education for minor alcohol offenses, those who also had used inhalants reported fewer school honors and more expulsions, truancy, academic failures, criminal offenses, run away behavior, and associations with troubled peers, as well as many more drug and alcohol problems. More of them also had mothers or siblings with alcohol- or drug-related problems. Studies of juvenile probationers, incarcerated youth, and other samples of delinquent and antisocial adolescents report lifetime inhalant use prevalence rates between 30 and 40 percent. One recent investigation of a large national sample of adult inhalant users concluded that persons with inhalant use disorder had significantly elevated levels of antisocial conduct, including a variety of early-onset and interpersonally violent behaviors. Some families are burdened by antisocial personality disorder and substance dependence in the adults and by conduct disorder and SUDs (sometimes including inhalant use disorder) in the adolescent children, and there is growing evidence that genetic factors play a role in these familial disorders. Less respect for parents, less familism (i.e., personal values related to family), less parental monitoring, and experiences of trauma within the family unit have all been related to increased risk of initiation of inhalant use in recent investigations.

Finally, two studies suggest that, even more than other drug users, inhalant users report histories of childhood abuse and/or neglect victimization.

INHALANT USE DISORDER

Diagnosis and Clinical Features

A diagnosis of inhalant use disorder should be considered in persons showing intermittent changes compatible with substance intoxication, together with an odor of organic solvents, inhalation paraphernalia, or the occasionally present perioral or perinasal papular glue-sniffers' rash. The DSM-5 inhalant use disorder criteria require recipients of the diagnosis to have exhibited a problematic pattern of use (which implies recurrent use) and to display related clinically significant functional impairments and/or distress. To receive an inhalant use disorder diagnosis, patients must meet at least two of ten diagnostic criteria. Inhalant use disorder criteria assess whether the patient has used inhalants over a longer period and/or in larger amounts then intended; whether there has been an ongoing desire or

actual attempts to quit or reduce inhalant use that have proved unsuccessful; and whether there are indications of craving for, tolerance to, or recurrent physically hazardous inhalant use. Other diagnostic criteria assess for inhalant-related social role impairments, use of inhalants despite social, interpersonal, physical, or psychological problems that are caused or worsened by inhalant use, and increased salience of inhalant use reflected in growing amounts of time spent using or recovering from the effects of inhalants. A number of specifiers may accompany an inhalant use disorder diagnosis. Individuals who meet 2 or 3 diagnostic criteria are diagnosed with mild inhalant use disorder, whereas those meeting 4 or 5 and 6 or more criteria receive diagnoses of moderate and severe inhalant use disorder, respectively. Diagnostic specifiers can also be used to indicate whether the person with inhalant use disorder is in early or sustained remission. If a specific inhalant of abuse is identified, this agent should be specified in the diagnosis (e.g., toluene use disorder instead of inhalant use disorder), but abused inhalants are often mixtures of psychoactive chemicals and it is frequently difficult to identify a specific inhalant. Thus, the diagnosis is often the generic "inhalant use disorder."

Although DSM-5 provides no diagnosis for a characteristic inhalant withdrawal syndrome, and inhalant-related withdrawal signs and symptoms are not included in the diagnostic criteria for inhalant use disorder, several studies have identified inhalant withdrawal signs/symptoms in some inhalant users; inhalant users with withdrawal signs/symptoms reportedly experience mild restlessness, inattentiveness, anxiety, and other physical symptoms. One national study of persons with DSM-IV-TR inhalant dependence disorder found that nearly 48 percent evidenced three or more clinically significant inhalant withdrawal symptoms and called for the inclusion of these symptoms in DSM-5. Thus, despite the absence of a separate inhalant withdrawal diagnosis, patients' complaints of withdrawal symptoms probably should be counted toward a diagnosis of inhalant use disorder. ICD-10 uses the category of volatile solvent dependence syndrome to describe a disorder similar to moderate/severe inhalant use disorder; it also includes the less severe category of harmful use of volatile solvents, which is defined as a pattern of use causing damage to health. Harmful volatile solvent use appears similar to mild inhalant use disorder.

Several studies of the DSM-IV-TR inhalant abuse and inhalant dependence diagnostic criteria specifically, and multiple studies of DSM-IV-TR substance abuse and substance dependence disorder criteria generally, indicated that SUDs was most properly conceptualized as unidimensional disorder of varying levels of severity. Hence, DSM-5 includes only one "inhalant use disorder" diagnosis, but allows specification of severity based on the number of diagnostic criteria a patient meets.

Differential Diagnosis

Several conditions should be considered in the differential diagnosis. First, most adolescents who experiment with inhalants stop spontaneously after one or a few episodes of use and never meet criteria for inhalant use disorder. Although such use is dangerous and occasionally fatal (often on the first try), the many spontaneous resolutions support recommendations of minimalist, nonalarmist interventions for adolescent users who do not meet criteria for diagnoses of inhalant use disorder; parents also should be involved. Firm advice to remain abstinent, coupled with education of parents on recognizing future problems, may be sufficient.

Second, polysubstance use is common in adolescent patients, and dependence on drugs other than inhalants is established through history, physical findings, and toxicological screens. Such disorders may exist in addition to, or instead of, inhalant use disorder.

Third, uncontrolled and impulsive behavior during repeated episodes of inhalant intoxication may mimic aspects of, or be comorbid with, conduct disorder or antisocial personality disorder. Antisocial behavior before the initiation of inhalant use disorder, or during periods of abstinence from inhalants, suggests the presence of these disorders.

DSM-5 includes the following conditions for consideration in the differential diagnosis of inhalant use disorder: unintentional (e.g., occupational) exposure to inhalants, intentional inhalant misuse that does not meet inhalant use disorder criteria, intentional inhalant intoxication in a patient who does not meet inhalant use disorder criteria, inhalant-induced psychiatric disorder (e.g., inhalant-induced depression) in a patient who does not meet criteria for inhalant use disorder, and a number of additional substance-related (particularly sedative-related) and medical disorders that may present similarly to inhalant misuse and its consequences.

Course and Prognosis

The relatively high prevalence of inhalant use in high school surveys, and its relatively low prevalence in adulthood, led one expert to state that inhalant use "should be regarded as a passing phase or fad." Although most inhalant users probably do not progress to inhalant use disorder, the risk of such progression is much greater for those who have used inhalants than for those who have not. Studies indicate that inhalant use is associated with increased risk of future diagnoses of a broad range of psychiatric disorders including antisocial personality disorder and other SUDs. Among adult substance-dependent patients, a history of inhalant use indicated a significantly enhanced risk of antisocial personality disorder, depression, anxiety disorders including social phobia, polysubstance use, and injection drug use. Among a group of adult heroin addicts, those who had been

inhalant users appeared to be a "marginal group with particularly unfavorable developmental conditions and a specific course of addiction."

Most adolescents who progress from inhalant use to inhalant use disorder eventually shift to use of, and frequently dependence on, other drugs, but some continue active inhalant use into adulthood. Such chronic patients may use inhalants for extended periods each day for many years; they often demonstrate moderate criminal activity, weight loss, significant medical disease, slow and slurred speech, and impaired attention and memory. In a national sample of inhalant users, those with an AUD and/or mood/anxiety disorder antedating their inhalant use were most likely to transition to inhalant use disorder, typically within 1 year of inhalant use initiation.

Tolerance, and less commonly, mild withdrawal signs, and symptoms including sleep disturbance, irritability, shakiness, sweating, fleeting illusions, and nausea can occur. One observer reported tachycardia, delusions, and hallucinations during withdrawal.

Medical problems in chronic users may include (1) muscle weakness, sometimes with myoglobinuria and rhabdomyolysis; (2) GI problems, such as pain, nausea, vomiting, or hematemesis; (3) renal dysfunction, often with severe electrolyte imbalance; (4) cardiomyopathy; (5) hepatotoxicity; (6) pulmonary disorders (e.g., pulmonary hypertension, increased airway resistance, and acute respiratory distress); and (7) hematopoietic disorders (including elevated carboxyhemoglobin levels, methemoglobinemia, hemolytic anemia, aplastic anemia, and even acute myelocytic leukemia). Neurological problems include (1) headache, (2) paresthesias with peripheral neuropathy, (3) reversible cerebellar signs or cerebellar degeneration, (4) radiological abnormalities of widened sulci and basal cisterns, and (5) dementia (e.g., lead encephalopathy from leaded gasoline or white matter dementia from toluene). A large national survey reported significant associations between longer duration of inhalant use and greater likelihood of experiencing tuberculosis, bronchitis, asthma, sinusitis, tinnitus, positive HIV/AIDS status, sexually transmitted disease infections, anxiety, and depression. As reviewed above, maternal inhalant use during gestation has been associated with an FSS similar to fetal alcohol syndrome.

Treatment

A Cochrane Collaboration systematic review published in 2010 identified no randomized or controlled clinical trials of inhalant use disorder treatments. Similar conclusions were reached in a systematic review of psychosocial interventions for inhalant misuse published in 2012. Nearly all evaluations of psychosocial treatments for inhalant use disorders were case series and the overall methodological quality of this literature was poor. One recent Turkish study of adolescents evaluated a brief cognitive-

behavioral intervention for inhalant misuse and found significantly greater rates of continuous abstinence from inhalants at 1-year follow-up for treated youth relative to controls (i.e., 58 percent vs. 38 percent).

The absence of evidence-based treatments of inhalant use disorder is problematic for several reasons. Recent surveys of mental health practitioners indicate that they feel unprepared to work with inhalant misusers, although they believe this patient group requires particularly intensive and specialized rehabilitative efforts. Further, research has documented high rates of health services utilization and numerous perceived barriers to treatment among inhalant users. Some clinicians have advocated for specialty inhalant use disorder treatment programs (although there are no data supporting this recommendation) and/or clinics and extended treatment periods given the manifold psychiatric, neurocognitive, and health disorders evidenced by many, if not most, persons with inhalant use disorder.

Case studies of pharmacological interventions have evaluated drug use to reduce craving (e.g., aripiprazole [Abilify], lamotrigine [Lamictal]), inhalant-related withdrawal symptoms (e.g., baclofen [Gablofen]), neuropsychiatric symptoms caused by or co-occurring with inhalant misuse, such as paranoia, psychosis, depression, and anxiety (e.g., carbamazepine [Tegretol], haloperidol [Haldol], and risperidone [Risperdal]), and behavioral symptoms due to acute or chronic inhalant misuse such as increased locomotor activity (e.g., vigabatrin [Sabril], mGluR 2/3 agonist LY 379268). Although many of these case studies report positive findings, the literature is simply too scant and lacking in rigor to support codified treatment guidelines for pharmacological treatment of persons with inhalant-related disorders.

As noted above, the scientific literature relating to psychosocial interventions for inhalant use disorder is equally undeveloped, although holistic, indigenous-led residential treatment, family-based, activity-oriented, and broadly-focused biopsychosocial interventions (with aftercare) have been recommended in light of the variety and number of impairments typically evidenced by inhalant users.

Appropriate medical care is required for the medical sequelae of inhalant use. Given that risk of inhalant-related disorders is greatest in youth, vigorous treatment is needed for adolescent patients who progress from experimentation to inhalant-related disorders. Most of these youths have comorbid conduct disorder and are at high risk of serious adverse outcomes.

One authority recommends that a comprehensive treatment plan should include eight aspects: detoxification; a peer advocate system; assessment of physical, cognitive, and neurological deficits; building on existing strengths; developing new strengths; therapists trained in inhalant abuse; attention to personal and family issues; and assistance in returning

to the community.

Many inhalant-using youth are treated in outpatient, day treatment, and nonhospital residential programs. However, treatment of these delinquent youths, many who are dependent on multiple substances, is usually not specific to drug category (e.g., inhalants), but instead targets substance use and conduct problems generally. Referrals are often from social service and juvenile justice agencies, which pay for the treatment.

Treatment in these programs begins with detailed interviews, addressing use and assessing for SUDs for each drug category in DSM-5. Most patients meet criteria for multiple SUDs. Interviews also address diagnoses of disorders commonly comorbid in this group: conduct disorder, attention-deficit/hyperactivity disorder (ADHD), major and persistent depressive disorders, and PTSD. Interviews also address experiences of abuse or neglect, which are common in these patients. Group and individual therapy is behaviorally oriented, with immediate rewards for progress toward objectively defined goals in treatment and punishments for lapses to previous behaviors. Patients may attend on-site schools with special education teachers, together with planned recreational activities. The patients' parental families, often very chaotic, can be engaged in modifications of structural family therapy or multisystemic therapy, both of which have strong empirical support. Participation in 12-step programs is often required. Treatment interventions may be coordinated closely with interventions by community social workers and probation officers. Medications are typically not prescribed for inhalant use disorders per se, but a psychiatrist often prescribes antidepressants for depression, disulfiram (i.e., Antabuse) for comorbid AUD, or atomoxetine (Strattera) or bupropion (Wellbutrin) for ADHD (stimulant medications are often avoided in youth with SUDs). Progress is monitored with urine and breath samples analyzed for alcohol and other drugs at intake and during treatment. However, frequent monitoring of breath for inhalants, or of urine for hippuric acid, is too expensive for most public treatment programs. Patients who fail to abstain in day treatment may be transferred to residential care. Treatment usually lasts 3 to 12 months. Termination is considered successful if the youth has practiced a plan to stay abstinent; is showing fewer antisocial behaviors; has a plan to continue any needed psychiatric treatment (e.g., treatment of comorbid depression); has a plan to live in a supportive, drug-free environment; is interacting with family in a more productive way; is working or attending school; and is associating with drug-free, nondelinquent peers. In many cases, of course, these goals are only partially accomplished. As with all other recommended treatments of inhalant use disorder, controlled studies of long-term outcomes are lacking.

Laboratory Examinations

Therapy for SUDs often uses repeated tests of biological samples to validate patients' reports of abstinence or use. However, with inhalants such tests may be difficult to interpret. First, these volatile compounds have a relatively brief sojourn in the body and may be detected in urine for only a few hours after use. Second, even if the compounds occur in urine, they may volatilize out of samples during transfer or storage. Third, although hippuric acid, a toluene metabolite, can be detected longer than toluene, hippuric acid also may be produced from foods, raising a question of false-positive findings. Fourth, inhalants may bind to, or pass through, the plastic of urine cups or breadth collection bags, reducing concentrations and making the compounds undetectable. Thus, the most careful monitoring of inhalant use would involve frequent urine samples (e.g., two or three per week) at random times, collected in tightly sealed glass containers with little or no air space, and carefully refrigerated until analysis. Analyses for inhalants usually require gas chromatography, followed by flame ionization, electron capture, or mass spectroscopic procedures. Simultaneous analyses of the ratio of hippuric acid to creatinine may reveal recent toluene use. However, even under ideal conditions, the short half-life of inhalants makes inhalant monitoring much less valuable clinically than monitoring for many other substances. Similarly, breath samples may be collected in specially designed glass traps or in Tedlar (DuPont) (not Mylar) bags. However, the half-life of inhalants in alveolar air apparently is a matter of minutes, and so breath samples may be useless for monitoring treatment progress in patients who show no current signs of intoxication. Moreover, the authors have examined costs in their region for monitoring tests, as a proportion of the per-patient, per-day reimbursement rates that must cover *all* costs of intensive day treatment. Each urine test for toluene uses up more than one day's reimbursement. Each urine test for hippuric acid uses up about 1.8 days' reimbursement. Each also requires about a week to complete. Thus, frequent biological monitoring for inhalants is not very practical.

INHALANT INTOXICATION

Diagnosis and Clinical Features

Inhalant intoxication is a clinically significant mental disorder that occurs during, or immediately following, the inhalation of volatile hydrocarbons. Inhalation of hydrocarbons may be intentional or unintentional (as in the case of occupational exposures). Inhalation of toxic gases from products such as glues and paints, which often contain toluene and/or admixtures of other volatile hydrocarbons, can lead to a constellation of problematic psychological and behavioral impairments including dizziness, incoordination, slurred speech, unsteady gait, lethargy, tremor, depressed reflexes, psychomotor retardation, generalized muscle weakness, diplopia,

stupor or coma, nystagmus, and subjective feelings of elation or euphoria. Two or more of these signs/symptoms occurring contemporaneously, or in close temporal contiguity, with hydrocarbon inhalation should occasion a DSM-5 diagnosis of inhalant intoxication. Other reported concomitants or sequelae of inhalant intoxication include vomiting, aspiration of vomitus, bronchospasm, chest pain, cardiac arrhythmias or arrest, accidental burns, seizures, illusions, auditory or visual hallucinations, delusions, perceptions of altered body size, and suffocation in a plastic bag placed over the mouth and nose. If known, the specific inhalant should be indicated in the DSM-5 diagnosis (e.g., butane intoxication instead of inhalant intoxication). However, many inhaled products are mixtures of psychoactive hydrocarbons and it is therefore often difficult to identify the specific agent(s) leading to intoxication.

DSM-5 and ICD-10-CM require that an assessment for inhalant use disorder accompany a diagnosis of inhalant intoxication. Individuals meeting inhalant intoxication diagnostic criteria with a comorbid mild or moderate/severe inhalant use disorder receive codes of F18.129 and F18.229, respectively. If no comorbid inhalant use disorder is identified in an individual meeting inhalant intoxication criteria, the code is F18.929.

Factors that should heighten suspicion for inhalant intoxication include the identification of chemical odors on the clothing and/or breadth of an intoxicated person or the discovery of products that can be inhaled for psychoactive effects in the possession of such individuals. General indicants of enhanced risk of inhalant intoxication include signs of intoxication in a youth between ages 10 and 17 and in persons who appear intoxicated but who test negative on commonly administered toxicological screens for substance use (which usually do not test for inhalant use).

No nationally representative estimates are available for the annual or lifetime prevalence of inhalant intoxication. However, if it is assumed that most inhalant users experience intoxication at least once in their inhalant-using experiences, then lifetime inhalant intoxication may be relatively common, especially among youth and in diverse high-risk populations (juvenile and criminal justice, rural, etc.).

Laboratory Examinations

Random, intermittent monitoring of biological samples has modest value for confirming self-reported abstinence in currently nonintoxicated patients with inhalant use disorder, because inhalants are so briefly detectable in the body. However, patients displaying behavioral signs of inhalant intoxication do, in most cases, have detectable concentrations of inhalants in urine and breath. Confirming their presence may have later clinical or forensic value. Urine samples should be collected in glass vessels with little or no air space, sealed tightly, and refrigerated until analyzed for inhalants themselves and for the ratio of hippuric acid (a toluene

metabolite) to creatinine. Breath samples may also be collected in specially designed glass traps or in Tedlar (not Mylar) bags, and should be analyzed for inhalants within a few hours of their collection.

Differential Diagnosis

Differentiation from other intoxications, particularly those due to sedative (e.g., ethanol, benzodiazepine, and barbiturate) use, is aided by a history of inhalant use, the presence of inhalant odors or paraphernalia (e.g., solvent-soaked rags), and residues on the skin or clothing, a characteristic perioral or perinasal rash from contact with organic solvents, and toxicological examination of body fluids. Polysubstance use is common among inhalant users, and concurrent intoxications with other drugs may be assessed by history and toxicological examinations. As noted above, intoxication in persons who test negative on conventional drug screens or who are members of groups at high risk of inhalant use (e.g., homeless or aboriginal youth) also raise suspicion for inhalant intoxication. Likewise, parental, familial, or peer reports of patient inhalant use and/or known associations with inhalant users increase the likelihood of an inhalant-related intoxication. Despite evidence of inhalant intoxication, in comatose patients other explanations (e.g., closed head injury) must be sought. Dextrose 50 percent for injection (50 g) and naloxone (Narcan) 2 mg intravenously help rule out coma of diabetic or opioid origin. If delirium develops in the course of intoxication with inhalants, the diagnosis is inhalant intoxication delirium, rather than inhalant intoxication. If a depressed mood, anxiety, or psychosis appears prominently during an inhalant-related intoxication and if those symptoms are severe enough to warrant independent clinical attention, the diagnosis should be inhalant-induced depressive disorder, inhalant-induced anxiety disorder, or inhalant-induced psychotic disorder, respectively.

Course and Prognosis

The onset of intoxication is almost instantaneous following inhalation of volatile hydrocarbons, given the rapid absorption of those inhalants across pulmonary membranes and their quick distribution into the brain and other lipids. Inhalants are rapidly metabolized and excreted, and inhalant intoxication usually lasts a few hours or less, unless administration is continuous or occurs intermittently. Unless trauma, hypoxia, cardiac arrest, burns, or other problems ensue, there probably are no lasting effects from one or a few intoxications, except that each use of these reinforcing drugs may increase the probability of further use. Prolonged, repeated use causes persisting effects.

Treatment

Inhalant intoxication, like alcohol intoxication, usually receives no medical attention and resolves spontaneously. However, effects of the intoxication, such as coma, bronchospasm, laryngospasm, cardiac arrhythmias, trauma, or burns, require treatment. Otherwise, care primarily involves reassurance, quiet support, and attention to vital signs and level of consciousness. Some evidence suggests that physical agitation during inhalant intoxication may precipitate cardiac arrhythmias or cardiac arrest, so the environment should be calming and reassuring. However, sedative drugs, including benzodiazepines, are contraindicated, since they may potentiate inhalant effects. Because of the short half-life of inhalants, inhalant intoxication usually improves considerably after about 30 minutes of abstinence, unless other drugs were also consumed. Following resolution of the intoxication a careful evaluation is needed with appropriate intervention or referral for inhalant use disorder, other SUDs, conduct disorder, antisocial personality disorder, and possible comorbid mood, anxiety, neurocognitive, and other psychiatric conditions.

INHALANT INTOXICATION DELIRIUM

DSM-5 provides a diagnostic category for inhalant intoxication delirium, which is a disturbance of consciousness and a change in cognition that is a direct physiological consequence of intoxication with inhalants. Inhalant intoxication delirium should be diagnosed instead of inhalant intoxication if disturbances in attention and other cognitive functions (e.g., memory, language, and perception) predominate in the clinical picture and are severe enough to warrant clinical attention. In DSM-5 and ICD-10-CM, all persons diagnosed with inhalant intoxication delirium must be assessed for a comorbid inhalant use disorder. Persons with inhalant intoxication delirium and mild or moderate/severe inhalant use disorder receive codes of F18.121 and F18.221, respectively. Individuals with inhalant intoxication delirium and no co-occurring inhalant use disorder receive a code of F18.921. The course and treatment of inhalant intoxication delirium are like those of inhalant intoxication, but the additional confusion requires special attention to patient safety. If the delirium results in severe behavioral disturbances, short-term treatment with a dopamine receptor antagonist (e.g., haloperidol [Haldol]) may be necessary. Benzodiazepines should be avoided because of the possibility of adding to the patient's respiratory depression. ICD-10 provides a comparable diagnostic category, "acute intoxication due to use of volatile solvents, with delirium."

INHALANT-INDUCED MAJOR AND MILD NEUROCOGNITIVE DISORDER

Diagnosis and Clinical Features

Studies of inhalant-related neurocognitive disorder have yielded equivocal findings and have been beset by numerous methodological problems,

including a focus on adolescent users with briefer lifetime exposures. Several studies of adjudicated adolescent inhalant and noninhalant-using polysubstance users identified greater executive function, processing speed, behavioral, language, and memory problems in the inhalant-using group. A series of Australian studies comparing adolescent inhalant users to cannabis/other drug users and to community controls found that the inhalant-using and cannabis/other drug-using groups displayed a variety of neurocognitive impairments relative to controls, but did not differ significantly among themselves across many measures. However, clinical and some research evidence suggests that some inhalant-using adults develop inhalant-induced neurocognitive disorder. For example, among toluene users (average age, 29 years) studied with MRI, the neuropsychological deficits correlated strongly with the severity of cerebral white matter abnormalities, and those abnormalities appear to be caused by inhalants.

The cardinal feature of major neurocognitive disorder is a significant decline in level of performance in one or more cognitive domains such that the affected individual experiences problems in the performance of instrumental activities of daily living. The cognitive impairments can include disturbed executive functioning (i.e., planning, organizing, sequencing, and abstracting), deficits in learning and memory, and difficulties with language, social cognition, and perceptual-motor abilities. The symptoms must represent a decrement from earlier functioning, not occur exclusively in the course of a delirium, and persist beyond the usual duration of inhalant intoxication. Most affected persons will meet criteria for a moderate/severe inhalant use disorder. The ICD-10 category of dementia in other specified diseases classified elsewhere includes inhalant-induced persisting dementia.

Inhalant-induced mild neurocognitive disorder differs from inhalant-induced major neurocognitive disorder primarily in that the impairment(s) identified represent(s) a “modest” rather than “significant” decline in level of performance of one or more cognitive capacities and the deficits identified do not interfere with complex instrumental activities of daily living. In both major and mild inhalant-induced neurocognitive disorder, it is expected that abstinence from inhalants (concomitantly with abstinence from other neurotoxic agents) is associated with a stabilization or improvement of neurocognitive impairments (i.e., not progressive decline). Both DSM-5 and ICD-10-CM require that any patient meeting criteria for a major or mild inhalant-induced neurocognitive disorder be assessed for a comorbid inhalant use disorder. Individuals with an inhalant-induced mild neurocognitive disorder and no, a mild, or a moderate/severe inhalant use disorder are assigned codes of F18.988, F18.188, or F18.288, respectively. Patients with an inhalant-induced major neurocognitive disorder and no, a mild, or a moderate/severe inhalant use disorder are assigned codes of

F18.97, F18.17, or F18.27, respectively.

Differential Diagnosis

Nearly all of these patients have a moderate/severe inhalant use disorder, and many are dependent on alcohol, which can also produce dementia. Moreover, histories of head injury are very common among such patients. Thus, despite clear evidence of prolonged inhalant use, this disorder requires a full evaluation for the multiple causes of neurocognitive dysfunction.

Course and Prognosis

Few of these patients have been studied prospectively. A recent 15-year follow-up of 77 inhalant users documented significant improvements in standard neurological, cognitive, and ocular-motor tests, except for a subset of 17 users who had previously developed lead encephalopathy. Inhalant users with a history of lead encephalopathy showed little improvement in cognitive function over the 15 years of follow-up. Long-term abstinence may allow for recovery of brain function in a currently unclear percentage of inhalant users. Current findings suggest, however, that some inhalant users develop irreversible neurological deficits following prolonged inhalant use. These impairments may cause inhalant users to have less cognitive capacity to avoid relapses, complicating treatment, and each relapse may add to the accumulating cerebral toxicity.

Treatment

There is no established treatment of inhalant-induced neurocognitive disorder. Low-key street outreach and extensive social service support have been offered to severely deteriorated inhalant-dependent homeless adults. Patients may require extensive support within their families or in foster or domiciliary care.

INHALANT-INDUCED PSYCHOTIC DISORDER

The essential features of inhalant-induced psychotic disorder are hallucinations or delusions judged by laboratory tests, history, or physical examination to be due to the direct physiological effect of inhalant substances. Such psychotic symptoms sometimes develop during intoxication with inhalants; therefore, this diagnosis applies to patients who meet criteria for inhalant intoxication but who also have psychotic symptoms in excess of those usually associated with inhalant intoxication. The psychotic symptoms must cause clinically significant functional impairment and/or distress and be severe enough to warrant independent clinical attention. This diagnosis is not made in the presence of inhalant intoxication delirium.

Consistent with DSM-5 and the ICD-10-CM, persons diagnosed with inhalant-induced psychotic disorder are assessed for a comorbid inhalant use disorder. Specific codes are available for the diagnosis of inhalant-induced psychotic disorder with no (F18.959), a mild (F19.159), or a moderate/severe (F18.259) co-occurring inhalant use disorder.

The course and treatment of inhalant-induced psychotic disorder are like those of inhalant intoxication. The disorder is brief, lasting a few hours to (at most) a very few weeks beyond the intoxication. Clinical evidence suggests that tetraethyl lead is a potent contributor to psychotic symptoms in countries still using leaded gasoline. Vigorous treatment of such life-threatening complications as respiratory or cardiac arrest, together with conservative management of the intoxication itself, is appropriate. Confusion, panic, and psychosis mandate special attention to patient safety. Severe agitation may require cautious control with haloperidol (5 mg per 70 kg intramuscularly). Sedative drugs, including benzodiazepines, may potentiate and worsen inhalant intoxications.

In a randomized clinical trial, 40 hospitalized psychotic patients with DSM-III-R diagnoses of inhalant dependence and inhalant-induced organic mental disorder volunteered for treatment with either haloperidol (up to 40 mg per day orally) or carbamazepine (Tegretol) (up to 1,600 mg per day orally) for up to 5 weeks. Both groups could receive intramuscular haloperidol for acute agitation. The antipsychotic response of the two groups was similar, but there were significantly fewer extrapyramidal reactions in the carbamazepine group. However, without a placebo-control group, these data cannot establish that either medication was better (or worse) than no medication. One case study of a 25-year-old inhalant dependent Caucasian man found that 0.5 mg risperidone daily for four weeks, following by a dosage increase to 1 mg daily, was effective in significantly reducing paranoia symptoms and craving for, and use of, inhalants. It was not clear, however, to what extent the psychotic symptoms the patient displayed were due to an independent disorder or inhalant-induced.

Despite similarities in name, DSM-5's inhalant-induced psychotic disorder differs from ICD-10's residual and late-onset psychotic disorder due to volatile solvent use. The former is a variant of inhalant intoxication, whereas the latter is conceptualized as persisting long after direct psychoactive substance effects abate. Controversy continues as to whether inhalants produce persisting psychotic states in the absence of delirium or neurocognitive disorder. In ICD-10, acute intoxication due to the use of volatile solvents includes patients with marked psychotic symptoms arising in the course of the intoxication.

INHALANT-INDUCED DEPRESSIVE DISORDER

The essential feature of inhalant-induced depressive disorder is a

prominent and persistent disturbance in mood characterized by depressed mood and/or anhedonia judged to be due to the direct physiological effect of inhalant substances. Such mood symptoms sometimes develop during intoxication with inhalants, so this diagnosis applies to patients who meet criteria for inhalant intoxication but who also have depressive symptoms in excess of those usually associated with inhalant intoxication. The depressive symptoms must be severe enough to warrant independent clinical attention and associated with distress, functional impairment, or both. This diagnosis is not made in the presence of inhalant intoxication delirium. In ICD-10, acute intoxication due to volatile solvent use includes patients with marked affective symptoms arising in the course of the intoxication.

In accordance with DSM-5 and ICD-10-CM, individuals presenting with inhalant-induced depressive disorder are also assessed for a comorbid inhalant use disorder. Persons with inhalant-induced depressive disorder and no, a mild, or a moderate/severe inhalant use disorder are assigned the following codes: F18.94, F18.14, and F18.24, respectively.

The course and treatment of inhalant-induced depressive disorder are like those of inhalant intoxication. Inhalant-induced depressive disorder is brief, lasting a few hours to (at most) a very few weeks beyond the intoxication. Although antidepressant drugs are seldom appropriate for these relatively brief disorders, the risk of suicide requires a careful history and may require psychosocial treatment. Suicide has been implicated in about 40 percent of both inhalant-related medical examiners' death reports and inhalant-related visits to hospital emergency departments.

INHALANT-INDUCED ANXIETY DISORDER

The essential features of inhalant-induced anxiety disorder are panic attacks or other prominent anxiety symptoms judged to be due to the direct physiological effect of inhalant substances. Such anxiety symptoms sometimes develop during intoxication with inhalants; thus, this diagnosis applies to patients who meet criteria for inhalant intoxication but who also have anxiety symptoms that predominate in the clinical presentation and that are in excess of those usually associated with inhalant intoxication. The anxiety symptoms must cause distress and/or functional impairments severe enough to warrant independent clinical attention. This diagnosis is not made in the presence of inhalant intoxication delirium. In ICD-10, acute intoxication due to the use of volatile solvents includes patients with marked anxiety symptoms arising in the course of the intoxication.

DSM-5 and ICD-10-CM diagnostic protocols require that all patients with inhalant-induced anxiety disorders be assessed for the presence and severity of comorbid inhalant use disorder. Persons with an inhalant-induced anxiety disorder and no comorbid inhalant use disorder received a code of F18.980, whereas those with a mild or moderate/severe co-

occurring inhalant use disorder receive codes of F18.180 and F18.208, respectively.

The course and treatment of inhalant-induced anxiety disorder are like those of inhalant intoxication. Sedative and antianxiety drugs are contraindicated, since they worsen inhalant intoxication, which precipitates inhalant-induced anxiety disorder.

UNSPECIFIED INHALANT-RELATED DISORDER

The diagnosis of unspecified inhalant-related disorder is reserved for inhalant-related clinical presentations associated with clinically significant distress and/or functional impairment that do not meet full diagnostic criteria for any specific inhalant-related disorder.

NITROUS OXIDE-RELATED DISORDERS

DSM-5 includes nitrous oxide-related disorders among other (or unknown) substance-related disorders because of differences between nitrous oxide and other inhalants in modes of action and associated problems. Still, many studies categorize nitrous oxide and nitrites as "inhalants," inconsistent with the DSM-5, and relatively few studies examine nitrous oxide and nitrites separately from other "inhalants." Epidemiological studies suggest that while toluene-containing products may frequently be utilized by adolescents, among adult past-year "inhalant" users, 68 percent had used nitrous oxide or whippets, making this the most commonly used "inhalant" in that group.

Nitrous oxide was introduced for clinical practice in 1844 and is still a widely used inorganic gas anesthetic. It also is sold as a propellant in whipped cream dispensers in "whippets" for whipping cream. Not surprisingly, nitrous oxide misuse seems to appear most commonly among the health care and food service workers who use these preparations. In 2009, the National Survey on Drug Use and Health estimated that between 16.3 percent and 31.6 percent of 12- to 17-year-old American past-year inhalant initiates had inhaled "nitrous oxide or 'whippets'" at least once annually between 2002 and 2007. In one recent and relatively large study of conduct-disordered youth, nearly 16 percent reported lifetime nitrous oxide use. Prior diagnosis with a psychiatric disorder and temperamental fearlessness were significant predictors of nitrous oxide use. Youth who had used volatile solvents and nitrous oxide had significantly greater psychiatric and SUDs comorbidity compared to users of volatile solvents alone, nitrous oxide alone, or neither nitrous oxide nor volatile solvents.

Nitrous oxide has a rapid onset and offset of action. It is mostly excreted in the breath with little or no biotransformation. Nitrous oxide reinforces self-administration behavior in animals. Although there is some evidence that nitrous oxide influences opioid systems, the subjective effects of the drug are not reduced by the opioid antagonist, naloxone.

In studies of persons without SUDs, some prefer inhaling nitrous oxide more than placebo, but others do not. In such studies the drug produces feelings described as "drunk," "dreamy," "coasting or spaced out," and "pleasant bodily sensations." Individuals reporting more of these feelings are more likely to choose nitrous oxide over placebo. When compared in the laboratory to oxygen, nitrous oxide produces significantly higher ratings of euphoria but no significant increases in subjective sleepiness, distinguishing it from the CNS depressant effects of most inhalants. Nitrous oxide-exposed subjects do as well as oxygen-exposed subjects on tasks of coordination and reaction time but perform significantly worse on some measures such as tasks of logical reasoning and digit symbol substitution, where subjects replace numbers with symbols according to a digit-symbol code.

Chronic use may produce diffuse polyneuropathy and myelopathy with extensive, although sometimes reversible, neurological symptoms mimicking those of vitamin B_{12}-related pernicious anemia. Active vitamin B_{12} requires reduced cobalt, but nitrous oxide irreversibly oxidizes cobalt, suppressing the activity of an important enzyme, methionine synthase.

A 35-year-old male dentist with no history of other substance use problems complained of problems with nitrous oxide abuse for 10 years. This use had begun as experimentation with what he had considered a harmless substance. However, his level of use increased over several years, eventually becoming almost daily for months at a time. He felt a craving for nitrous before sessions of use. Then, using the gas while alone in his office, he immediately felt numbness, a change in his temperature and heart rate, and an alleviation of depressed feelings. "Things would go through my mind. Time was erased." He sometimes fell asleep. Sessions might last a few minutes, or up to 8 hours; they ended when the craving and euphoria ended. He had often tried to stop or cut down, sometimes consulting professionals about the problem.

Cases in the literature strongly suggest that nitrous oxide intoxication, misuse, and dependent use do occur. Considering the drug's brief duration of action and users' intermittent inhalation patterns, nitrous oxide appears unlikely to produce a clinically significant withdrawal syndrome. It is not clear, however, whether nitrous oxide produces other substance-related disorders. There are no clinically practical urine or breath tests for the presence of nitrous oxide. Misusers of the drug who develop neurological symptoms may have low serum concentrations of vitamin B_{12}, and remission of neurological symptoms has followed combined administration of B_{12} and folate, together with abstinence from nitrous oxide. In the absence of controlled studies of treatment of nitrous oxide use disorders, general principles for treating other SUDs should guide the treatment of these patients.

AMYL, BUTYL, AND ISOBUTYL NITRITE-RELATED DISORDERS

DSM-5 includes amyl, butyl, and isobutyl nitrite-related disorders among

"Other (or Unknown) Substance-Related Disorders." In a national household survey in the United States, among adult past-year inhalant users, amyl nitrite/poppers/rush were the second most commonly used "inhalant" (32 percent) behind nitrous oxide; in comparison, nitrites were the seventh most commonly used "inhalant" among adolescents in that same survey. Although in the general population prevalence of nitrite use is relatively low, in some populations (i.e., urban HIV-positive gay and bisexual men) high rates of nitrite/popper use have been shown, and nitrite use has been associated with risky sexual behaviors.

The vapors of amyl nitrite, a volatile liquid, produce vasodilation and smooth muscle relaxation. Once supplied in easily "popped" glass vials, amyl nitrite found wide use as acute inhalation therapy for angina pectoris between 1867 and about 1980. Sublingual nitroglycerin tablets have now superseded that use. Amyl nitrite and its close relative butyl nitrite enjoyed a flurry of use as recreational drugs during the 1970s and 1980s. Reports from that time suggest that the use mainly was among adolescents, users of other drugs, and homosexual men. The latter group especially reported using the drugs to enhance orgasm, and these nitrites also were considered to produce a high, a feeling of wild abandon, or an altered state of consciousness. A recent study of antisocial youth found that 83 percent of nitrite users reported that they experienced intoxication when they inhaled nitrites. Nonmedical users called amyl nitrite vials poppers. Those vials' availability declined with their declining medical use. Butyl nitrite and related compounds were then increasingly sold as room deodorizers (but really for inhalation) under suggestive names such as Rush or Locker Room.

In one study of patients who had misused these drugs, most experienced dizziness, light-headedness, cardiac palpitations, blurred vision, and a feeling of warmth immediately after inhalation. Others complained of immediate headache, nasal burning, nausea, cough, dyspnea, or syncope. Nearly half found the experience not at all pleasant; the others rated it only fair to good. Some experts believe the drug does not have primary reinforcing effects mediated in the CNS, but that certain pleasurable sensations associated with peripheral vasodilation lead to its continued use.

A few case reports document severe, but nonfatal, methemoglobinemia after nitrite inhalation by persons genetically deficient in methemoglobin reductase. There also are reports of rapidly fatal methemoglobinemia after oral ingestion (not inhalation) of these compounds by genetically normal persons. Some epidemiological studies further suggested that among homosexual men nitrite use might increase the risk of HIV infection or of developing Kaposi sarcoma after HIV infection. However, those observations were heavily confounded, since nitrite users tended to have more partners and to pursue unsafe sexual practices more frequently.

Considered together, the lack of a compelling instantaneous high, the immediate unpleasant feelings, and the longer-term risks from amyl and butyl nitrite seem to contribute to a rather low prevalence of use. In 2000 to 2001, among Americans 12 to 17 years of age living in households, 1.5 percent said that they ever had used "amyl nitrite, 'poppers,' locker room odorizers, or 'Rush.'" Recent Monitoring the Future survey findings and the results of a large study of conduct-disordered youth also found that less than 2 percent of 12- to 17-year-olds had tried inhaling nitrites.

Some persons with sustained industrial exposures to nitrites experience withdrawal headaches on weekends or vacations, but it seems unlikely that intermittent, brief exposures to these short-acting drugs would produce clinically significant withdrawal in recreational misusers of nitrites. There are no studies for the treatment of amyl or butyl nitrite-related disorders, and so such treatment should follow general principles for the treatment of other substance-related disorders.

FUTURE DIRECTIONS

Although considerable progress has been made in better understanding the pharmacology of inhalants and the epidemiology of inhalant use and inhalant use disorders, there exist several areas that should be the focus of future investigation. First, there is limited empirical research to guide the treatment of adolescents or adults with inhalant-related disorders; such work is critical to clinicians working with such populations. Second, substance abuse treatment generally requires regular objective monitoring for substance use; the currently available methods for monitoring inhalant use are cost-prohibitive, are specific to individual inhalants (do not cover the broad spectrum of possible inhalants), generally can only detect very recent use (i.e., testing for specific inhalants), or may lack specificity (i.e., testing for inhalant metabolites such as hippuric acid). Development of sensitive and specific monitoring methods, which are cost effective, would represent a major contribution. Third, individuals with inhalant-use disorders are very likely to have other mental health problems. Little research thus far has attempted to better understand the reasons underlying that comorbidity (directionality of causality or a common shared risk). Given that inhalant use is common and that a single use may be fatally toxic, it is important to support continued efforts to develop empirically tested, effective inhalant prevention programs that may be broadly and cost-effectively implemented. Taxonomic research identifying subtypes of inhalant users could provide data that would facilitate matching of inhalant users to specific treatment interventions. Studies of the short- and long-term consequences of inhalant use and longitudinal investigations of the natural history of inhalant use disorders are needed. Ethnographic investigations of cultural and subcultural factors relevant to inhalant use could prove informative as to the reasons why and how

inhalants are misused. Psychometric research is needed to examine the reliability, validity, and latent structure of DSM-5 inhalant use and inhalant-induced disorders diagnostic criteria in different populations of inhalant users. Health services research is needed to better understand the health care utilization patterns of persons with inhalant-related disorders. Additional investigation of the phenomenology of inhalant intoxication across different inhalants and mixtures of inhalants would provide additional insight into the reinforcing properties of inhalants.

REFERENCES

Beckman NJ, Zacny JP, Walker DJ. Within-subject comparison of the subjective and psychomotor effects of a gaseous anesthetic and two volatile anesthetics in healthy volunteers. *Drug Alcohol Depend*. 2006;81:89.

*Bowen SE, Batis JC, Paez-Martinez N, Cruz SL. The last decade of solvent research in animal models of abuse: mechanistic and behavioral studies. *Neurotoxicol Teratol*. 2006;28:636.

Butland BK, Field-Smith M, Ramsey JD, Anderson HR. Twenty-five years of volatile substance abuse mortality: a national mortality surveillance programme. *Addict*. 2012;108:385–393.

Cairney S, O'Connor N, Dingwall KM, et al. A prospective study of neurocognitive changes 15 years after chronic inhalant abuse. *Addict*. 2013;108:1107–1114.

Compton WM, Conway KP, Stinson FS, Colliver JD, Grant BF. Prevalence, correlates, and comorbidity of DSM-IV antisocial personality syndromes and alcohol and specific drug use disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry*. 2005;66:677.

Filley CM, Halliday W, Kleinschmidt-DeMasters BK. The effects of toluene on the central nervous system. *J Neuropathol Exp Neurol*. 2004;63:1.

Garland EL, Howard MO. Phenomenology of adolescent inhalant intoxication. *Exp Clin Psychopharacol*. 2010;18:498–509.

Garland EL, Howard MO. Adverse consequences of acute inhalant intoxication. *Exp Clin Psychopharmacol*. 2011;19:134–144.

*Garland EL, Howard MO. Volatile substance misuse: clinical considerations, neuropsychopharmacology and potential role of pharmacotherapy in management. *CNS Drugs*. 2012;26:927–935.

Garland EL, Howard MO, Vaughn MG, Perron BE. Volatile substance misuse in the United States. *Subst Use Misuse*. 2011;46:8–20.

*Howard MO, Bowen SE, Garland EL, Perron BE, Vaughn MG. Inhalant use and inhalant use disorders in the United States. *Addict Sci Clin Pract*. 2011;6(1):18–31.

Howard MO, Garland EL. Volatile substance misuse: toward a research agenda. *Am J Drug Alc Abuse*. 2013;39:3–7.

Jain R, Verma A. Laboratory approach for diagnosis of toluene-based inhalant abuse in a clinical setting. *J Pharm Bioallied Sci*. 2016;8:18–22.

Kandasamy A, Jayaram N, Benegal V. Baclofen as an anti-craving agent for adolescent inhalant dependence. *Drug Alcohol Rev*. 2015;34:696–697.

Lee DE, Schiffer WK, Dewey SL. Gamma-vinyl GABA (Vigabatrin) blocks the expression of

toluene-induced conditioned place preference (CPP). *Synapse*. 2004;54:183.

*MacLean S, Cameron J, Harney A, Lee NK. Psychosocial therapeutic interventions for volatile substance use: a systematic review. *Addict*. 2012;107:278–288.

*Marsolek MR, White NC, Litovitz TL. Inhalant abuse: monitoring trends by using poison control data, 1993–2008. *Pediatrics*. 2010;125:906–913.

Nguyen J, O'Brien C, Schapp S. Adolescent inhalant use prevention, assessment, and treatment: a literature synthesis. *Int J Drug Policy*. 2016;31:15–24.

Ridenour TA, Bray BC, Cottler LB. Reliability of use, abuse, and dependence of four types of inhalants in adolescents and young adults. *Drug Alcohol Depend*. 2007;91:40.

Storck M, Black L, Liddell M. Inhalant abuse and dextromethorphan. *Child Adolesc Psychiatr Clin N Am*. 2016;25:497–508.

Substance Abuse and Mental Health Services Administration. *Drug Abuse Warning Network, 2003: Area Profiles of Drug-Related Mortality. DAWN Series D-27, DHHS Publication No. (SMA) 05–4023*. Rockville, MD: Office of Applied Studies; 2005.

Substance Abuse and Mental Health Services Administration. *Drug Abuse Warning Network, 2011: National Estimates of Drug-Related Emergency Department Visits. DAWN Series D-39, DHHS Publication No. (SMA) 07–4256*. Rockville, MD: Office of Applied Studies; 2013.

Substance Abuse and Mental Health Services Administration. *Treatment Episode Data Set (TEDS). 2002–2012. State Admissions to Substance Abuse Treatment Services, BHSIS Series: S-72, HHS Publication No. (SMA) 14–4889*. Rockville, MD: Office of Applied Studies; 2014.

Substance Abuse and Mental Health Services Administration. *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings. NSDUH Series H-48, HHS Publication No. SMA 14–4863*. Rockville, MD: Office of Applied Studies; 2014.

Swaim RC. The moderating effects of perceived emotional benefits on inhalant initiation among American Indian and white youth. *Am J Addict*. 2015;24:554–560.

Tormoehlen LM, Tekulve KJ, Nanagas KA: Hydrocarbon toxicity: a review. *Clin Toxicol*. 2014;52:479–489.

Wu LT, Howard MO. Psychiatric disorders in inhalant users: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Drug Alcohol Depend*. 2007;88:146.

Wu LT, Pilowsky DJ, Schlenger WE. Inhalant abuse and dependence among adolescents in the United States. *J Am Acad Child Adolesc Psychiatry*. 2004;43:1206.

Wu LT, Ringwalt CL. Inhalant use and disorders among adults in the United States. *Drug Alcohol Depend*. 2006;85:1.

Wu LT, Schlenger WE, Ringwalt CL. Use of nitrite inhalants ("poppers") among American youth. *J Adolesc Health*. 2005;37:52.

Yuncu Z, Zorlu N, Saatcioglu H, et al. Abnormal white matter integrity and impairment of cognitive abilities in adolescent inhalant abusers. *Neurotoxicol Teratol*. 2015;47:89–95.

▲ 11.8 Tobacco-Related Disorders

DOUGLAS M. ZIEDONIS, M.D., MAKENZIE E. TONELLI, B.A., AND SMITA DAS, M.D., PH.D.

Tobacco use disorder (TUD) is the most prevalent, deadly, and costly of SUDs. The prevalence of addiction to nicotine, the primary psychoactive substance in tobacco, surpasses all other SUDs. Tobacco currently remains unrivaled as the leading cause of premature death and preventable morbidity in the United States and globally with an incidence of 1 in 5 deaths in the United States being attributed to cigarette smoking. TUDs are the leading major public health issue and uniquely affect psychiatric patients.

The literature points to several key reasons why addressing tobacco use and TUDs must occur in mental health and addiction treatment settings, including the fact that 44 percent of all cigarettes in the United States are consumed by smokers with mental illness and among patients in mental health and addiction treatment settings, 50 to 95 percent are smokers. This is in stark contrast to the general trend of decreasing smoking rates in the United States at large (17.8 percent). The most common cause of mortality among individuals with a psychiatric or SUD is related to tobacco and tobacco accounts for the disturbing statistic that people with mental illness die on average 25 years earlier than the general population. In addition to the negative physical health consequences, there is a probable independent association between suicide and current smoking and longer lifetime smoking duration. (In a recent analysis, for example, longer lifetime smoking (≥41 vs. ≤10 years) was associated with higher odds of suicide (odds ratio [OR] = 2.26, 95% confidence interval [CI] = 1.30–3.93).

Unfortunately, TUD has also been one of the most often ignored SUDs in mental health and addiction treatment settings, particularly by psychiatrists; the need to address tobacco in these settings is clear and clinical practice guidelines delineate expectations to do so. Historically, mental health and addiction treatment settings not only neglected to address tobacco use and TUDs; there was a minimization and rationalization that supported use. In a survey by the American Association of Medical Colleges (2007) of more than 3,000 physicians, 47 percent of psychiatrists felt patients had more immediate problems to address and 22 percent reported that cessation heightens other symptoms. This survey marked a change of improved attitudes for the majority; however, these attitudes of rationalizing not addressing the topic isolated psychiatrists compared to all other physician specialties surveyed. Some psychiatrists continue to erroneously see TUD as a lesser priority as compared to other SUDs or worry that smoking cessation or limiting smoking may exacerbate other psychiatric symptoms.

These assumptions are not supported by the literature. While psychiatric patients may have a worse tobacco withdrawal experience, their long-term mental status is not worsened and likely improves with their physical health. In fact, accruing evidence demonstrates significant improvement in mental health and quality of life as result of tobacco

cessation. A recent systemic review and meta-analysis conducted by Taylor et al. (2014) assessed 26 smoking cessation studies and found a significant decrease in anxiety, depression, and stress as well as improvements in overall mood and quality of life among quitters compared to nonquitters. Tobacco cessation also enhances the likelihood of successful sobriety from other addictive substances among individuals with an SUD, as continued tobacco use is actually found to hinder recovery from other addictions. While they have worse withdrawal as a result of heavier dependence, people with psychiatric comorbidities are able to quit with overall stable mental health and much better physical health, quality of life, finances, and often lower dosages of psychiatric medication.

Psychiatrists have an opportunity to help the well-being and extend the lifespan of their patients by evaluating and treating TUD using evidence-based approaches in psychiatric treatment settings, including psychosocial treatments, pharmacotherapies, community resources, and carbon monoxide (CO) monitors. Seven U.S. Food and Drug Administration (FDA)-approved pharmacological agents are available to help individuals with TUD, including five nicotine replacement therapies (NRTs) and two non-NRT pills (bupropion [Zyban] and varenicline [Chantix]). All seven of these medications can be effectively used in psychiatric patients; three NRT are available over-the-counter (OTC) or by prescription (patch, gum, lozenge), two NRT are only by prescription (nicotine inhaler and nasal spray), while the two non-NRTs are by prescription only. The integration of these medications with TUD-specific psychosocial interventions, community resources, and providing guidance and appropriate monitoring can improve clinical outcomes by 50 percent. Treating TUD will not only help improve a patient's well-being, but these services can be reimbursed by completing "meaningful use" assessments of tobacco, performing CO assessments, and performing TUD treatments. Ultimately treating TUD in people with mental illness has multiple benefits at the individual, provider, and system levels.

HISTORY

Tobacco use in the New World dates back to at least 600 AD and was introduced into European culture in the 16th century. Early on, most tobacco use was via pipes, smokeless tobacco, or cigars. The cigarette became popular beginning in the early 1900s with the invention of the cigarette-making machine, the match, and the use of acidifying agents to permit nicotine to enter the lower respiratory tract, where it could be rapidly absorbed into the arterial circulation. Cigarette use grew dramatically in the first half of the 20th century, initially through distribution to military and next with targeting women as well. The first reports of the association of smoking and disease began in the 1950s and culminated with the 1964 U.S. *Surgeon General's Report on Smoking and*

Health. The tobacco industry sought to fight the clear risk of cigarettes by marketing "light" or "low tar" cigarettes, which in reality are no less harmful than other cigarettes and are often just smoked more efficiently. With the tobacco industry efforts in marketing, tobacco use in the United States did not decline significantly until the 1970s and 1980s owing to tobacco-control activities and taxation. In the 1990s and 2000s, tobacco use in the United States has declined little, and that decline may be largely due to environmental smoke laws after the Environmental Protection Agency declared secondhand smoke dangerous as well. Of note, the tobacco industry also tried to suppress these findings. Although tobacco use in many Western nations has declined, use in developing countries is substantially increasing in part due to relaxed policy enforcement and increased marketing practices. Recent attempts by the FDA and the World Health Organization (WHO) to regulate and discourage tobacco use are important, but have been only partially successful.

The tobacco industry targeted individuals with mental illness and/or addiction, a likely cause for the high rates of TUD in this population, as demonstrated by Prochaska and evidence documented in The Legacy Tobacco Documents Library (http://legacy.library.ucsf.edu). For example, the tobacco industry attempted to validate the "self-medication hypothesis," that people with schizophrenia needed to smoke to manage their psychiatric symptoms. R.J. Reynolds in a Legacy Library record even specifically studied what they considered the "downscale market profile," people who are homeless or mentally ill, and documents from them note the population is "more impressionable to marketing/advertising . . . they're more susceptible. They're less formed intellectually . . . more malleable." Research promoting smoking among those with mental illness was widely distributed in meetings, publications, and through lobbying/policy efforts. Advertising was directed at people with mental illness, implying that tobacco was part of psychiatric illness (i.e., advertisement headline reading, "Schizophrenic. For New Merit, having two sides is just normal behavior"). Also worrisome is the intimate history the tobacco industry had with psychiatric institutions, supplying them with free or low cost or tax-free cigarettes. While much of this activity took place from the 1950s to 1970s, even in the 1990s mental health centers were still communicating with companies like R.J. Reynolds for cigarette donations. Tobacco companies' also lobbied to allow for smoking in psychiatric facilities after Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) mandated smoke-free hospitals.

Beyond the health consequences of tobacco use, the 1988 U.S. Surgeon General's report, *The Health Consequences of Smoking—Nicotine Addiction*, comprehensively reviewed the evidence and concluded that smoking was a form of substance dependence. The most recent report in 2014, "The Health Consequences of Smoking—50 Years of Progress" clearly

outlines the dangerous history of tobacco and the strong effects of tobacco control policies in the United States. Nonetheless, tobacco remains the leading preventable cause of death in America.

COMPARATIVE NOSOLOGY AND DIAGNOSTIC FEATURES

Although tobacco dependence prevalence has been high over much of the last century, it was only in the 1980s that the term nicotine dependence was first accepted and included, (though not coded), in the *Diagnostic and Statistical Manual of Mental Disorders*, Third Edition (DSM-III-TR) and the ninth edition of the *International Classification of Diseases* (ICD-9) as an SUD. In the 1970s, tobacco corporations sought to influence the American Psychiatric Association (APA) task force revising the DSM; these efforts may have led to the narrow definition of tobacco dependence initially. The 1988 U.S. Surgeon General's report, *The Health Consequences of Smoking—Nicotine Addiction* comprehensively reviewed the evidence and concluded that smoking was a form of substance dependence. Also in DSM-III, *abstinence-induced nicotine withdrawal syndrome* was first classified. Nicotine *dependence* and *withdrawal* were the two defined nicotine-related disorders in the revised fourth edition of the DSM-IV-TR, which distinguished nicotine from other substances. The essential feature of any substance *dependence* disorder in DSM-IV-TR is that "the individual continues use of the substance despite significant substance-related problems." Because 50 percent of smokers die of a smoking-related illness, this definition clearly is applicable to nicotine use. The essential feature of *withdrawal* is "a substance-specific maladaptive behavioral change . . . that is due to the cessation of, or reduction in, heavy and prolonged substance use." Since nicotine withdrawal produces an observable, well-defined, time-limited syndrome in more than half of smokers, this definition also was appropriate.

Contrary to other substances, nicotine *abuse* was omitted as a diagnosis in DSM-IV due to the strong propensity for use to transition to what was defined as dependence (when tolerance and withdrawal criteria are met). Nicotine abuse was also confined to significant psychosocial problems but not physical problems, and the former is rare with nicotine use. (Physical problems, however, are present even with nondaily smokers.) The tenth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) included *harmful use*, a category that is similar to abuse but includes continued use that causes physical problems (which was warranted given the high rates of use despite numerous health consequences).

May 2013 marked the release of the current DSM-5, which now includes among the 10 substance-related and addictive disorders the classification of tobacco-related disorders. Throughout this chapter prior DSM-IV *nicotine dependence* terminology has been replaced with the current DSM-

5 classification of *TUD*. The new DSM-5 diagnostic criteria will be discussed in more detail in this chapter in the assessment section related to establishing a TUD diagnosis.

EPIDEMIOLOGY

In the United States, 18 percent of the adult population smokes cigarettes, with TUD being the most prevalent of all substance-related and addictive disorders. The prevalence of smoking in the United States has significantly decreased—it was just shy of 50 percent—over the past 5 decades. Unfortunately some groups still disproportionately use tobacco, such as 50 to 85 percent of those with mental illness/SUDs, those with lower income, and lower education. Women have slightly lower rates of smoking (15.3 percent in women vs. 20.5 percent in men). Lower smoking rates are observed in those older than 65 years of age (8.8 percent) likely as a result of premature death. Although smoking in adolescents (<18) has declined, smoking in young adults (18- to 24-year-olds) has not changed or even increased recently. Worldwide, 1 billion men and 250 million women use tobacco products, and consumption is increasing globally, though decreasing in some high-income and upper middle-income countries.

Ninety percent of smokers tried their first cigarette by age 18 and 99 percent by age 26. Sixty-nine percent of current smokers want to quit smoking completely and 52.4 percent of smokers in the United States have made a quit attempt for >1 day in the past year. While 6.2 percent have sustained cessation, less than half of smokers (48.3 percent) who saw a health professional in the past year reported receiving advice to quit. In those who try to quit, use of counseling and/or cessation medications was 31.7 percent.

In terms of the diagnosis of nicotine dependence per se (based on DSM-III-R, given the lack of newer studies with DSM-5 criteria to date), approximately 24 percent of people in the United States develop nicotine dependence at some point in their lives, making it the most prevalent psychiatric disorder. Approximately 50 percent of current adult daily smokers are nicotine dependent. Nicotine withdrawal occurs in approximately 50 percent of smokers who try to quit.

Prevalence of Other Traditional and Emerging Tobacco Products

As a reference to describe other products, a description of a cigarette is warranted. A cigarette is a cylinder of tobacco rolled in thin paper for smoking. Nicotine, a naturally occurring pesticide, is the primary psychoactive and addictive chemical in tobacco, and hence cigarettes. Other chemicals are added to the shredded tobacco mixtures such as humectants (i.e. propylene glycol) to keep the tobacco moist, flavoring products (e.g., menthol), enhancers, sugars, and chemicals, such as

ammonia, which boost the addictive properties of cigarettes, especially when burned. A cigarette contains 6 to 11 mg of nicotine, allowing for 1 to 2 mg of absorptions and smoking a pack per day equals 20 to 40 mg of nicotine absorbed per day.

Use of other forms of tobacco products, including older products (e.g., cigars), in addition to new emerging products, such as smokeless tobacco (e.g., snus, hookahs, and dissolvables) and other alternatives, such as electronic nicotine delivery systems (ENDS or e-cigarettes), are gaining popularity today. National estimates based on the 2012–2013 National Adult Tobacco Survey (NATS) demonstrate that among adults meeting criteria as established users of each of the following tobacco products, current daily use prevalence are as follows: cigars, cigarillos, or filtered little cigars, 12.8 percent; regular pipes, 12.6 percent; water pipes or hookahs, 1.2 percent; e-cigarettes, 17.9 percent; chewing tobacco, snuff or dip, 48.1 percent; dissolvable tobacco products, 16.8 percent; and snus, 11.3 percent. Unfortunately, these products are becoming more appealing and accessible among children and adolescents. There are also valid concerns around these products serving as gateways to other tobacco products, especially among children. McMillen et al. (2012) assessed use prevalence among former smokers and found that smokers abstinent for a shorter amount of time were more likely than more long-term quitters to try these products, raising a concern around these products as possible deterrents to successful quit attempts. Patients should be educated on the addictive nature of emerging forms of tobacco as well as possible misconceptions regarding safety of these products, which are not FDA approved.

Adverse health effects of cigars and cigarillos are comparable to those associated with cigarette use, including cancers (oral, esophageal, and lung) and tobacco-related illnesses that are significantly correlated with level of smoke exposure. Cigars are arguably more harmful than cigarettes as they contain higher concentrations of known carcinogens, toxins, and tar. Whether a cigar user inhales or not, they are still absorbing nicotine; when inhaled, nicotine is delivered more quickly through the lungs, whereas if not inhaled, nicotine is absorbed at a slower rate through the tissue in the mouth. Cigars are offered in a wide range of sizes, thus the amount of nicotine that one might contain can vary largely. Cigars typically contain around 100 to 200 mg of nicotine, but can contain as much as 444 mg of nicotine. It can be approximated that in proportion to the total nicotine amount and yield of cigarettes, that nicotine exposure from cigars can range anywhere from 9 to 68 mg. The smallest version of cigars is equivalent to the standard size of a cigarette, called a little cigar (containing about 1 g of tobacco—slightly more than the amount in a cigarette). A marginally larger and narrower version is called a cigarillo (containing about 3 g of tobacco).

Electronic cigarettes, also referred to as e-cigarettes or ENDS, have

increasingly been linked to marketing statements that promote them as safer alternatives to tobacco, though, there is a dearth of clinical research, including nicotine pharmacology studies, in this area to support this claim. Electronic cigarettes still contain, in addition to the addictive substance, nicotine, toxic chemicals such as propylene glycol (an irritant in antifreeze). As supported by recent clinical studies, the nicotine exposure provided by e-cigarettes, compared to traditional cigarettes, varies according to the experience level of the user, where blood levels of nicotine, and its metabolite, cotinine, tend to be similar to that delivered by cigarettes for current e-cigarette smokers and less concentrated for inexperienced users. The amount of nicotine in the e-liquid ranges in strengths of 6 to 24 mg/mL in a cartridge, though there is variation among different e-cigarette brands and brand models in overall nicotine content as well as in the amount that is actually delivered to the user. The flavored nature of the substance can add appeal to vulnerable populations (youth) and also enhances the addictive potential. In addition, the products are not subject to the same regulation as traditional cigarettes, lending to surges of advertising and widespread availability in convenience stores, vape shops, etc.

A *water pipe*, also referred to as hookah or shisha, is a device with origins in India. Water pipes date back to the 1700s, and are used for smoking tobacco by delivering cooled charcoal tobacco smoke that passes through water before inhaled. Levels of CO, nicotine, tars, and heavy metal exposure from this product can often be even higher than what traditional cigarettes produce, contrary to the common belief that water pipe smoking has less health consequences than cigarette smoking. Variation in nicotine exposure is influenced by a variety of determinants, including the temperature at which the tobacco is burned, type of tobacco and water pipe product, and individual patterns of smoking behaviors.

Snus, originally from Sweden, is a form of smokeless tobacco that is encapsulated in a porous pouch that the user stores in between the cheek and the gum; it does not require the need to spit, as is the case with chewing tobacco/dip. Though it has been suggested to be associated with less significant harmful effects than cigarettes if entirely replacing cigarette smoking, it can be a gateway to other tobacco products and may decrease the likelihood of successful smoking cessation attempts. In Sweden, snus use has exceeded smoking, and the country has seen a dramatic decline in lung cancer rates; however, many former smokers are now dependent on snus and there is evidence of an increased risk of pancreatic cancer.

Chewing tobacco and snuff are fairly similar smokeless products, with the former being a more finely ground form of tobacco and the latter provided in the form of loose leaves. Both are stored in the mouth and then spit out after use. Snuff is encapsulated in a porous pouch, whereas chew or finely ground tobacco is directly placed as a wad between the gum and the

cheek to allow nicotine to be absorbed by the mouth tissue. Typically, the nicotine content of these products can range from 3.6 to 4.5 mg. Compared to the seven seconds it takes for the brain to respond to nicotine from inhalation of cigarette smoke, the same effect can be delayed to up to 30 minutes for chew or snuff users, though is continuously delivered for up to 60 minutes after use (whereas nicotine delivery ceases immediately after the last puff of a cigarette). In terms of nicotine content, one dip or pouch of these forms of "spit tobacco" is equivalent to five cigarettes, whereas one can, containing 144 mg of nicotine, equates to smoking 4 packs of cigarettes. Smokeless tobacco users suffer the same systemic adverse effects as smokers due to tars and tobacco additives and have higher rates of oral, prostate, pancreatic, and cervical cancers compared with nonsmokers. Snuff users have a 50-fold increased risk of cancer of the cheek and gum. Smokeless tobacco is associated with numerous dental, periodontal, and oral soft tissue problems, including gingival recession, periodontal attachment loss, tooth staining, halitosis, and leukoplakia (a precursor to oral cancer).

Dissolvables are smokeless products that melt in the mouth, and similar to snus, are promoted as forms of tobacco not subject to prohibited use in smoke-free locations. They are provided in forms that mimic candy, such as lozenges, orbs, pellets, strips, or sticks, where the juices are swallowed and are often sweet or minty in flavor. While little is known about the safety of dissolvable products, they are still addictive in nature.

ETIOLOGY

Nicotine and acetylcholine bind to and activate nicotinic acetylcholine receptors. Originally, these receptors were thought to be confined to the ganglia and neuromuscular junction but have now been found in several areas of the CNS. The two types of ligand-gated receptors that appear to be involved in the reinforcing properties nicotine are the α4/β2 and α7 subtypes. These receptors are unusual in that they quickly desensitize. This phenomenon, plus the fact that repeated use of nicotine increases, not decreases, the number of receptors, suggests that nicotine may actually act as much as an antagonist as an agonist at this receptor. Emerging data indicate that distinct nicotinic receptors including those that contain the β4 subunit are involved in nicotine withdrawal symptoms.

The reinforcing effects of nicotine appear to be modulated by dopamine and glutamate release due to nicotinic receptors on dopamine cells in the ventral tegmental area and the core and shell of the nucleus accumbens; whereas withdrawal symptoms upon nicotine cessation are modulated by glutamate and corticotropin-releasing factor (CRF) release in the habenula–interpeduncular axis. The mesocorticolimbic system neural circuits are now seen as a common pathway for addictive drugs.

There are multiple contributors to tobacco dependence. Nicotine

dependence is partly due to its multiple pharmacological effects and the cigarette delivery device. Nicotine is best absorbed in the oral mucosa, nose, skin, and lung (more so than the stomach due to acidity). It also is able to enter the bloodstream without being processed by the liver (first pass effects). Nicotine via cigarettes is rapidly absorbed directly into the arterial circulation and reaches the CNS in less than 15 seconds. Peak behavioral and cardiovascular effects occur within a few minutes. Nicotine is metabolized via the liver and has a half-life of 2 hours. The rapid effects of nicotine and the rapid decline lead to withdrawal symptoms that can form rather quickly in smokers.

Binding of nicotine to receptors leads to the subjective gratifying effects (through modulation of dopamine, glutamate, GABA, norepinephrine, and other systems) in the mesocorticolimbic system. The release of these neurotransmitters (via nicotine receptors) leads to the experience of reinforcing effects of smoking such as pleasure, arousal, and subjective lessening of anxiety. Many people with dependence subjectively feel as though nicotine is required to cope with life demands. Over time, continued use may be to suppress withdrawal. Tolerance also develops (part of dependence) as shown in several studies, and is established quickly and increases over the course of a day. Compared to nonsmokers, smokers often report less of the pleasurable effects of nicotine, such as arousal, vigor, and a subjective experience often referred to as "buzz," as a result of this tolerance.

Ironically, nicotine's lack of a typical effect of drugs of abuse—that is, intoxication—may also contribute to nicotine dependence. Lack of intoxication allows the smoker to imbibe large amounts of nicotine on a daily basis without producing an obvious behavioral interference in daily activities, which, in turn, increases the chance of physical or psychological dependence. Legal, and relatively inexpensive compared to other drugs, cigarettes can be obtained with few social skills or resources. Being legal, punishment for smoking only occurs from social approbation.

The ideal reinforcer is one that occurs rapidly, frequently, and consistently, with minimal effort, and is unlikely to be punished. Nicotine via cigarettes has all these attributes. Smoking is the most rapid method to deliver a drug to the brain. It is also one of the most, if not the most, frequently reinforced behaviors. Taking 10 puffs per cigarette times 20 cigarettes per day times 365 days per year times 20 years results in millions of drug self-administrations. Because cigarettes are manufactured, they deliver precise doses of drug. Of course the tobacco industry has long been invested in making cigarettes more addictive via design changes, paper choice, droplet sizes, additives (such as propylene glycol to keep the tobacco moist, or menthol flavoring), sugars, and chemicals, such as ammonia, which boost the addictive properties of cigarettes, especially when burned. The combination of multiple drug effects, lack of

intoxication, and method of delivery makes nicotine users highly susceptible to developing dependence. In addition to the primary reinforcement described above, tobacco is highly subject to secondary reinforcement (classical conditioning) where conditioned stimuli (seeing a cigarette, smelling it, or being in situations where one usually smokes) promote the urge to smoke.

Although not widely known, initiation and cessation of tobacco use are as heritable as alcohol dependence or abuse. Some of the genetic effects are shared with alcohol, and some are specific to tobacco. Specific genes may affect responses of the body and the brain to nicotine (metabolism of nicotine, receptor sensitivity, and the levels of neurotransmitters). An example is variants in the *CYP2A6* gene that influence the rate at which a smoker clears nicotine; factors like this also contribute to an individual's susceptibility to tobacco addiction and response to treatments. Similarly, adolescent initiation of smoking can affect the trajectory to dependence (more than 80 percent of smokers begin as teenagers), social environment, peer modeling, and family environment.

Children who are more likely to start smoking are those who have a high need to conform, low academic performance, rebelliousness, depressive symptoms, and poor self-esteem. Peer and family influences are paramount. Attention-deficit, conduct, and alcohol or drug use disorders increase the risk of initiation and maintenance of smoking. Adolescent initiation of smoking can affect the trajectory to dependence (more than 90 percent of smokers begin as teenagers), social environment, peer modeling, and family environment.

DIAGNOSIS AND CLINICAL FEATURES

New classification of substance-related and addictive disorders in the DSM-5 now includes the following tobacco-related disorders (APA, 2013): (1) *TUD,* (2) *tobacco withdrawal*, (3) *other tobacco-induced disorders*, and (4) *unspecified tobacco-related disorders*. TUD is characterized by ongoing use of tobacco products resulting in harmful consequences exemplified by clinically significant impairment or distress. As with all SUDs, TUD criteria are a combination of the DSM-IV criteria for abuse (i.e., work, school and home responsibilities are not met as a result of use; continued use in dangerous contexts; persistent use regardless of substance-related—caused or worsened—interpersonal adverse effects) and dependence (i.e., - tolerance; withdrawal; increased amount of substance and length of time used; recurrent ineffective attempts to quit, cut down or control use; use over an extended period of time; repeated use in spite of known psychological and/or physical consequences; limited ability to fulfill social, recreational, and work-related obligations).

A new feature of the DSM-5 classification for all SUDs, including tobacco, is the addition of craving as a criterion, defined as a strong desire

or urge for the substance. To meet diagnostic criteria for SUD, at least two of these 11 conditions must occur within a 12-month period: (1) Tobacco use often persists at a higher volume and over a longer duration of time than intended; (2) Repeated desire or unsuccessful attempts to quit, cut down, or control tobacco use; (3) A significant amount of time is spent engaging in activities related to acquiring or using tobacco products; (4) Craving, or an intense desire, to use tobacco; (5) Inability to fulfill major role obligations at work, school, or home due to persistent use; (6) Continued use despite continued or recurrent problematic social or interpersonal consequences, occasioned or worsened by the effects of tobacco (e.g., arguments with others about tobacco use); (7) Important social, recreational, and work-related activities are given up or limited due to tobacco use; (8) Tobacco use persists in situations where physically unsafe (e.g., smoking in bed); (9) Tobacco use is continued despite knowledge of the probable worsening or contributing effects of tobacco use on persistent or recurrent physical or psychological problems; (10) Withdrawal: nicotine associated withdrawal symptoms occurs with stopping use; (11) Tolerance: An increasingly larger doses of nicotine is needed to obtain the desired effect.

Abstinence threshold is defined as presenting none of the TUD criteria, with the exception of craving, for 3 or more months, with a cutoff of 12-months for early remission, and a threshold of 12 or more months for sustained abstinence. Severity can be assessed by the ICD-10 codes, which are integrated throughout and used in conjunction with the DSM-5 tobacco-related disorder classification. Mild severity is represented by the presence of two to three symptoms, moderate is four to five symptoms, and severe is six or more symptoms. Of note, appropriate use of ICD-10 codes for TUD will vary according to severity in cases where TUD is comorbid with tobacco withdrawal or tobacco-induced sleep disorder.

Tobacco withdrawal follows use for at least many weeks and either cessation or reduction in use resulting in: anxiety; depressed mood; irritability, frustration, or anger; insomnia; difficulty concentrating; restlessness; and increased appetite. To meet withdrawal criteria at least four of the seven signs or symptoms listed above must apply within 24 hours of abrupt cessation or volume reduction. Contrary to DSM-IV classification of withdrawal, DSM-5 omitted reduced heart rate or weight gain symptoms. In support of this change, research suggests that these variables have no predictive value on smoking cessation outcomes. The insomnia appears to be specific to increased awakenings and intense dreaming. Craving is common, and increased coughing, constipation, mouth ulcers, and decreased performance on vigilance tasks can occur. The syndrome is typically most severe with abstinence from cigarettes, intermediate with abstinence from smokeless tobacco, and least severe with abstinence from nicotine-replacement products. Most withdrawal

symptoms peak at 1 to 3 days and can last up to 2 to 3 weeks. In addition, craving can persist for 6 months or more.

Tobacco use and TUD are two to three times more prevalent among patients with anxiety, attention-deficit, mood, substance use, and other psychiatric disorders. Conversely, these psychiatric disorders are two to three times more common among current smokers than among people who have never smoked or are ex-smokers. There are several possible reasons for this association; for example, shared genetic influences on smoking and psychiatric disorders, subjective improvement by nicotine of a psychiatric disorder, modeling other psychiatric patients, and boredom. Depressive symptoms, substance use, and attention deficit predict the likelihood of starting to smoke, the probability of becoming dependent, and the inability to stop smoking. Smokers with a current or past history of most psychiatric disorders are more dependent, have more withdrawal on cessation, and have a lower rate of smoking cessation.

DIFFERENTIAL DIAGNOSIS

Many of the symptoms of tobacco withdrawal can mimic, exacerbate, or mask the symptoms of psychiatric disorders or the adverse effects of psychiatric medications; for example, akathisia, anxiety, depression, irritability, insomnia, and weight gain. Individuals who do quit smoking may have increased psychiatric medication side effects or have caffeine intoxication due to the metabolic changes in the P450 1A2 isoenzyme with stopping smoking. Stopping smoking and maintaining the same caffeine intake will increase the availability of caffeine and its intoxication effects. A differential diagnosis for change in mental status or increased medication side effects should include investigating whether there was a change in smoking status. Nicotine intoxication (rare) causes abdominal pain, dizziness, headaches, nausea, pallor, palpitations, sweating, vomiting, and weakness; treatment is supportive. Children are at risk of nicotine intoxication if they are able to access and consume nicotine medications or other nicotine products, including e-cigarettes.

COURSE AND PROGNOSIS

Twenty percent of all mortality in the United States is due to smoking, making it the leading cause of death in the United States, with a third of those deaths related to cardiovascular/cerebrovascular causes, another third related to cancer, and about a fifth related to respiratory diseases. About 10 percent of smoking-related deaths are from secondhand smoke exposure in nonsmokers and approximately 45 percent of smokers die of a smoking-related disease. In total, there are more than 480,000 deaths per year in the United States, which includes 41,000 due to secondhand smoke. Smoking is a major risk factor for lung cancer and chronic obstructive pulmonary disease (COPD), accounting for more than 75 percent of all

deaths from these disorders. Other common smoking-related diseases include low-birth weight offspring, perinatal complications, other cancers (e.g., throat, bladder, and pancreas), and ulcers. Additional, less obvious health effects of tobacco demonstrated in research including worse wound healing, diabetes, impotence, and osteoporosis. Secondhand smoke increases the risk of cancer and heart disease in spouses and the incidence of respiratory and ear problems in children.

With reference to cancer, the 2014 Surgeon General's report, *The Health Consequences of Smoking—50 Years of Progress*, reinforced there is sufficient evidence to infer a causal relationship between smoking and cancers of the lung, larynx, oral cavity, pharynx, esophagus, pancreas, bladder, kidney, cervix, stomach, and acute myeloid leukemia. Cigarette smoke has carcinogens that can act as direct carcinogens, cocarcinogens, or tumor promoters/initiators; smoke, free radicals, and other components can cause irritation, inflammation, cell proliferation/hyperplasia, oxidative stress and damage, and decreased organ function. Lung cancer, the leading cause of cancer deaths in the United States, is predominantly attributable to cigarette smoking, and demonstrates a clear dose response where the risk is associated with the number of cigarettes smoked per day and years smoked—this dose response is evident in other cancers as well.

Aside from cancer, cigarette toxicity in the lungs leads to several other health effects: loss of cilia, hyperplasia, metaplasia, mucus plugging of small airways, destruction of alveoli, and a reduced number of small arteries. Other lung/respiratory sequelae of TUD include COPD (emphysema and chronic bronchitis; the third leading cause of death in the United States, with 90 percent of cases attributable to cigarette smoking). Unfortunately, COPD is less noticeable in early stages of the disease process since small airways are affected first and are less clinically prominent. This often leads to later notice of effects on the large airways (emphysema) and irreversible progression of the disease. Other respiratory effects include increased risk of respiratory infection, such as pneumonia, in part due to cilia damage and triggering of asthma attacks.

Heart disease is the leading cause of death in the United States, and 30 percent of heart disease mortality is caused by smoking where smokers have two to four times the risk of coronary heart disease, and women are more affected. The end effects of cardiovascular disease include myocardial infarctions, strokes, aortic aneurysm, and the impacts of peripheral vascular disease. Acutely, nicotine causes increases in systolic blood pressure, heart rate, and cardiac output while causing vasoconstriction; another acute effect of smoking is that CO reduces oxygen availability to the heart. Chronically, cigarette smoke increases inflammation, platelet activation/thrombosis, and endothelial dysfunction that can lead to atherosclerosis. Later in the inflammation process smoking destabilizes plaques and promotes oxidation of low-density lipoproteins; the result is

reduced myocardial blood flow, reduced oxygen, and artery occlusion. Nicotine is also arrhythmogenic (it increases serum catecholamines) and this, combined with carboxyhemoglobin lowering the threshold for ventricular fibrillation, can lead to sudden cardiac death, accounting for half of all deaths from coronary heart disease. Quitting smoking reduces the risk of sudden death immediately and in the long term. The risk of myocardial infarction is dose dependent, but as few as one to four cigarettes a day raises the risk of dying from ischemic heart disease. Also, when it comes to quitting smoking and heart disease, the risk is reduced by about half after just the first year of cessation and falls to that of "never-smokers" after 15 years of abstinence.

The course and prognosis of the health effects of TUD are worse with comorbid disorders and exposures. There is an additive effect of environmental carcinogens with smoking that worsens cancer outcomes, for example, with radiation and asbestos exposure. Similarly with oral, laryngeal, and esophageal cancers, alcohol use can worsen cancer outcomes, possibly by chemical interaction; for example, the risk of larynx cancer is 75 percent higher in people who use tobacco and alcohol versus those exposed to either substance alone. Finally, due to the deleterious effects on the cardiac and respiratory systems, smokers are less able to be physically active. This increases the risk of many other comorbidities.

Environmental tobacco smoke: Secondhand smoke impacts the prognosis of TUDs and the likelihood of quitting when living with other smokers. There are an estimated 7,357 chemical compounds in cigarette smoke, which include respiratory irritants (acrolein and acetaldehyde), cardiovascular damaging chemicals (cyanide, arsenic, and the cresols), and carcinogens (other metals, *N*-nitrosamines, and polycyclic aromatic hydrocarbons); over 70 of these chemicals are carcinogenic. Often smokers cohabitate with other smokers, especially in psychiatric group or extended living situations. However, even separating smokers from nonsmokers within the same air space, cleaning the air, opening windows, and ventilating buildings do not eliminate secondhand smoke exposure. Therefore, secondhand smoke is very relevant in psychiatry. Secondhand smoke exposure causes nearly 34,000 heart disease deaths annually among adult nonsmokers in the United States. Secondhand smoke is a cause of cardiovascular disease and lung cancer (increase in risk is 25 to 30 percent). There is also an association between exposure to secondhand smoke and risk of stroke.

TREATMENT

Given the widespread evidence base of the harmful effects of TUD, a great deal of research has been done on treatment and allows for the effective use of evidence-based treatments. Several websites and organizations offer information on TUD treatment, including free booklets. There are excellent

guidelines for assessing and treating TUDs published in 2008 by the U.S. Public Health Service (USPHS) and for the treatment of SUDs in 2006 by the APA. Also, most state health departments have telephone lines to assist smokers in quitting and have websites with information for providers and patients (Table 11.8–1).

All patients should be assessed for (1) smoking status and (2) motivation to quit (Table 11.8–2). A useful approach to TUD treatment is the 5As model: *Ask, Advise, Assess, Assist,* and *Arrange* follow-up or the minimum standard of AAR: *Ask, Advise,* and *Refer*. Given the overwhelming prevalence and morbidity/mortality of TUD in psychiatric settings, all patients should be *asked* about current and past tobacco use, type of tobacco used, and frequency of use; TUD should be listed in their problem list and treatment plan with strategies either for either increasing their motivation or on how to quit, perhaps including psychosocial treatment, education, community resource referral, and/or medication options and guidance. In addition to asking about tobacco use, assess for DSM-5 criteria. All patients using tobacco should be *advised* to quit with a personalized message, for example, including their specific health risks. With psychiatric patients, it is often important to state one's belief that the patient has the ability to stop. For some, the above advice has no immediate effect. However, the cumulative effect of repeating the advice at later visits and the receipt of advice from others can often motivate a quit attempt. Helping smokers who are not ready to quit to reduce the number of cigarettes smoked daily (aided by nicotine replacement medications) can motivate them to try to quit in the near future. The clinician should *assess* for desire to quit. Motivational interviewing/enhancement can be employed even if a patient is not ready to quit. Common motivators to quit are health concerns, effects of smoking on others, and social pressure. Common barriers to cessation are withdrawal, fear of failure, and fear of weight gain. Next, smokers need *assistance*, which can include, setting a quit date, preparing for the quit date, need for medication or telephone or in-person therapy, weight gain concerns, high-risk situations, making cigarettes unavailable, discussions with significant others (especially if they too are smokers), and so forth. Finally, the fifth A is for *arranging* follow-up since TUD is an addictive disorder and relapse is likely. Follow-up can include working with the psychiatric provider or using other community resources, such as state quit lines (1-800-QUIT-NOW) (Table 11.8–2). Because relapse is often rapid, the first follow-up phone call or visit should be 2 to 3 days after the quit date.

Psychosocial Therapies and Community Resources

Psychosocial interventions can be varied based on motivational level for quitting. Lower motivated to quit individuals can benefit from motivational interviewing/motivational enhancement therapy with personalized

feedback. Feedback that has more impact includes CO meter scores, cost of cigarettes per year for them, health problems that they already have that are worsened or caused by ongoing tobacco use, and effect on family and significant others. Educational programs such as learning about healthy living provide information that can help enhance motivation. If patients worry about weight gain with quitting, exercise and nutrition to help them begin to lose weight before quitting can be a helpful strategy.

Table 11.8–1.
5 A's Model of Tobacco Cessation Intervention

1. **Ask About Tobacco Use:** Screen (and document) current and past tobacco use

Open-ended → all-inclusive and specific questions: "Tell me about your tobacco use?" "Do you use any tobacco products, such as cigarettes, cigars, smokeless tobacco, or e-cigarettes?"

Yes: Understand type and quantify amount and frequency of use.

No: "Did you use tobacco products in the past?" "Have you used any tobacco product in the last 30 days, even a puff of a cigarette, for example?"

- ► More dependent users: Smoking within the first 30 minutes after awakening and/or over 20 cigarettes per day
- ► Assess for DSM-5 criteria, withdrawal symptoms, prior quit attempts, types of treatment, social support for quitting, medical conditions
- ► CO meter to assess: 15–50 ppm = smoker; 0–5 ppm in 2–3 days of quitting

2. **Advise to Quit** (or reinforce people who have quit) with a tailored message

- ► "Quitting smoking can be one of the most important things you can do for your health"
- ► "People with psychiatric diagnoses suffer from smoking related illnesses"
- ► Personalize with comorbidities (physical and psychiatric) or motivation for quitting

3. **Assess for Desire to Quit**

- ► "What are your thoughts about quitting" with a follow-up such as, "Would you have any interest in quitting in the next month?"

4. **Assist Patients**

Strategies for Lower Motivated Patients:

- ► For those in precontemplation or contemplation, listen to reasons for not being in preparation
- ► Try to increase motivation CO meter scores, $$ spent, other disorders, increasing freedom, options in relationships, more job opportunities, effect of tobacco smoke on their loved ones (family, friends, pets), longer life, and fewer sick days
- ► Do Pros/Cons or "Importance" or "Confidence" ruler
- ► Consider the 5 Rs **relevance, risks, rewards, roadblocks,** and **repetition** (revisit the conversation in the future).

Strategies for Higher Motivated Patients:

- ► Shared decision making
- ► Assess for triggers to plan for relapse prevention followed by, how can we plan for those difficult situations
- ► Reassure worst withdrawal occurs within the first 2 weeks
- ► Discuss the "quit plan" and help prepare for the "quit date"
- ► Provide prescription for one or combination of 7 FDA medications for TUD
- ► Encourage some form of counseling support (community resources, counseling, or as part of medication management).

5. **Arrange Follow-Up**

- ► Follow-up counseling to support cessation and prevent relapse
- ► Assess engagement with community resources/other providers

- ► Follow-up with you as psychiatric provider addressing TUD in plan
- ► Engagement with state's tobacco quit line (1-800-QUIT-NOW)

Reprinted with permission from FOCUS: Journal of Lifelong Learning in Psychiatry, (Copyright ©2015). American Psychiatric Association. All Rights Reserved.

Table 11.8–2.
Selected Community Resources for Tobacco Use Disorder Treatment

Internet Resources	
http://www.surgeongeneral.gov/tobacco	U.S. Public Health Service guidelines
http://www.cancer.org	American Cancer Society
http://www.cdc.gov/tobacco	Centers for Disease Control and Prevention
http://www.lungusa.org	American Lung Association
http://www.nci.nih.gov	National Cancer Institute
http://www.treatobacco.net	Society for Research on Nicotine and Tobacco
www.attud.org	Association for Treatment of Tobacco Use and Dependence
www.makesmokinghistory.org	Make Smoking History
www.quitnet.com	Quit Net
www.nicotine-anonymous.org	Nicotine Anonymous
Quit-Lines and Live-Chat Support	
1-800-QUIT-NOW	National Network of Tobacco Cessation Quit-lines
1-855-QUIT-VET	US Department of Veterans Affairs Quit-line
https://livehelp.cancer.gov/	National Cancer Institute LiveHelp Online Chat
Mobile Apps	
www.text2quit.com	Text2Quit
www.smokefree.gov/smokefreetxt/	Smokefree TXT
www.livestrong.com/quit-smoking-app/	My Quit Coach
www.becomeanex.com	Become An Ex
http://teen.smokefree.gov/sftapps.aspx	Smoke-Free Teen
http://women.smokefree.gov/sfwapps	Smoke-Free Women

Other Local Resources to Explore in Your Area

- ► Nicotine Anonymous (www.nicotine-anonymous.orga)
- ► Contact Employee Assistance Program (EAP). They may offer free help
- ► Contact insurance company to ask for help. Some assistance may be free or partially reimbursable, including counseling and NRT medications

Behavior therapy, including relapse prevention, is the most widely accepted and well-proven psychological therapy for treating TUDs. Behavior therapy consists of several techniques, three of which are supported by good evidence. *Skills training* and *relapse prevention*

identify high-risk situations and plan and practice behavioral or cognitive coping skills for these situations. *Stimulus control* refers to eliminating cues for smoking in the environment. This approach also requires preparation such as removing smoking triggers and working with loved ones to plan for support.

Mindfulness-based interventions are being utilized to help individuals. Mindfulness approaches help individuals stay present in the moment and recognize cravings as temporary thoughts and feelings that will dissipate over time and which they can notice and accept without needing to act on those thoughts or feelings. Mindfulness is also helpful for stress reduction in general.

Twelve-Step Facilitation is a standard evidence-based treatment in most addiction treatment; however, there are no studies of Nicotine Anonymous (NicA) in the literature. The strength of attending meetings can be in the social support and a plan for recovery. There are national phone meetings for easy access for anyone, as well as local meetings, but those are less numerous than AA.

There are an increasing number of other community resources. The most widely used are quitlines that provide individual coaching over the phone and sometimes with free NRT. 1-800-QUIT-NOW works in all states. There are online Internet websites/apps for those who prefer computer support (i.e., www.becomeanex.org and www.ffsonline.org). Other resources are clinic run smoking cessation groups, peer specialists, peer websites (www.njchoices.org); training module for peer specialists (www.Rxforchange.org); and information for club houses (www.genesisclubhouse.org). Other therapies, such as acupuncture, hospitalization, and hypnosis, have been studied, but overall the results are equivocal and yet to be proven effective.

Many of these resources are becoming increasingly more accessible and many are available at no-cost. Medicare, Medicaid, and other insurance companies are also now supporting reimbursement of many psychosocial treatments, with more and more options being considered.

Pharmacological Therapies

The USPHS and the APA guidelines recommend that all smokers be offered medication to aid in smoking cessation. The seven FDA-approved medications include five NRTs, bupropion (Zyban), and varenicline (Chantix). NRT patch, gum, and lozenge are available OTC, while NRT inhaler, NRT nasal spray, varenicline, and bupropion are by prescription only in the United States. All the medications approved by the FDA for smoking cessation double the quit rate (similar to results obtained with behavior therapy) and appear equally effective (Table 11.8–3). The one exception may be varenicline, which has been shown superior to some other medications. The medications produce few side effects; less than 5

percent of patients stop medication due to adverse effects (except for nicotine nasal spray). While some medications have prescribing considerations for pregnant women or people with heart conditions, it is certainly the case that continued tobacco use is very harmful for these groups.

History and patient input are essential to choosing medications: cost can be a consideration, previous attempts with a medication, medical comorbidities, severity of dependence/withdrawal, and breakthrough symptoms. Typically, combined patch and another NRT or monotherapy with varenicline are first tried. Combinations of medication have greater efficacy. Updated FDA guidelines on NRTs allow for use while smoking for risk reduction and promote using more than one medication (as shown in practice and research for many years).

Nicotine Replacement Therapies. Because nicotine undergoes first-pass metabolism by the liver, nongastrointestinal routes have had to be used to deliver nicotine. NRTs are effective, presumably because they reduce nicotine withdrawal, without the immediate reinforcing effect of rapidly absorbed nicotine in tobacco smoke. Often NRT is used in inpatient or other clean air settings. Most replacement therapies use a short period of maintenance (6 to 12 weeks), often followed by a gradual reduction period (6 to 24 weeks). However, longer-term use may be necessary for some. Most NRTs should be used with caution in cardiovascular patients (specifically those who are 2 weeks post myocardial infarction, have serious arrhythmias, or have unstable angina). Finally, to optimize oral absorption of nicotine in the gum, lozenge, or inhaler, minimize acidic beverage (coffee, juice, soda) consumption 15 minutes prior to and during use.

Table 11.8–3.
Pharmacotherapies for Tobacco Use Disorders

Seven FDA-Approved Medications for Tobacco Use Disorders (please consult the PDR for details): 5 Nicotine Replacement Therapies (Nicotine Agonists) and 2 Non-Nicotine Agonists (Bupropion and Varenicline)

Product	Initial Prescription	Patient Education Issues	Precautions	Side Effects	Approximate Costs
Nicotine patch (OTC)	Apply 21 mg patch for 4–6 weeks; taper to mid-dose (14 mg) for an additional 2–4 weeks and can	Apply each day to clean, dry, hairless skin, using tape if needed; rotate patch site daily to minimize skin irritation	Caution within 2–6 weeks of heart attack or with arrhythmias	Skin reaction/rash (50% of patients—usually mild, treat with 1% hydrocortisone cream or 0.05% triamcinolone	$45/2-week supply (generics available and effective); $3/day

	use lower dose (7 mg) for another 2 weeks; start at lower dose for lighter smokers			cream); vivid dreams or sleep disturbances can occur if worn overnight	
Nicotine gum (OTC)	2 mg; 4 mg dose if smoking ≥25 cpd); chew every 1–2 hours as needed; usual range is 9–16 pieces/day	"Chew and park"; avoid coffee and other acidic beverages 15 minutes before/during use	Dentures	Jaw pain, hiccups, mouth irritation; nausea if swallowing saliva	$50/2-week supply; generics available and effective; $3/day
Nicotine lozenge (OTC)	2 mg; 4 mg (if smoke within 30 minutes of waking); dissolve (do not chew) in mouth; use 9–20/day for 6 weeks, then taper over 6–12 weeks	Avoid food, coffee and acidic drinks 15 minutes before/during use	N/A	Hiccups, nausea, heartburn	$43 for 72 lozenges ($0.60/lozenge); approximately $6/day average

Reprinted with permission from FOCUS: Journal of Lifelong Learning in Psychiatry, (Copyright ©2015). American Psychiatric Association. All Rights Reserved.

Nicotine patches (OTC), applied daily to a clean and hairless location, provide long-acting nicotine administration. Typically patients use a 21 mg patch, although 7 and 14 mg patches can be used for lighter smokers (10 to 15 cigarettes per day). Side effects include local skin reactions (minimized with creams and by rotating patch sites) and insomnia/vivid dreams (minimized by removing the patch prior to bed).

Nicotine gum (OTC) may be used alone or in combination with other medications and is generally utilized hourly in initial phases of cessation. The 2 mg gum is recommended for those smoking less than 25 cigarettes per day, while the 4 mg is generally used for heavier smokers (but dose can vary by preference). Proper usage education is important for efficacy and to decrease adverse effects: patients should "chew" a few crunches until a peppery or tingly flavor emerges and then "park" between the cheek and

the gum until the tingle goes away. This will help maximize nicotine absorption (for about 30 minutes) and reduce transfer to the stomach, which can irritate the gastrointestinal tract. The nicotine lozenge (OTC) is easier to take consistently and will achieve 25 to 27 percent higher plasma concentrations over time compared to the gum (allow the lozenge to dissolve, do not chew it). Patients who smoke their first cigarette within 30 minutes of awakening should use the 4 mg lozenge, while others might consider the 2 mg dose, but again preference should be considered. Both the gum and lozenges provide relieve of immediate withdrawal especially in high-risk situations.

Prescription NRT in the United States includes the inhaler and spray. The inhaler is intended for oral mucosal absorption like the gum and lozenge and provides the smoker with the hand-to-mouth routine of smoking. The nicotine nasal spray provides the fastest absorption of the NRT products and has the highest potential for dependency, as well as more perceived complications of use such as rhinitis.

Non-Nicotine Prescription Medications. Bupropion (Zyban), which affects dopamine, norepinephrine and nicotinic–cholinergic receptors, can help decrease cravings and withdrawal symptoms. Dosages of 300 mg per day reliably double quit rates in smokers with and without a history of depression. Combining bupropion and a nicotine medication has greater efficacy. Contraindications to use of bupropion include a seizure disorder (the risk of seizures in those appropriately screened is less than 1 in 1,000), current/past bulimia or anorexia nervosa, recent/concurrent monoamine oxidase inhibitor (MAOI), or other bupropion use. More common side effects are insomnia and dry mouth. There is a black box warning for neuropsychiatric side effects. This medication can be started 1 to 2 weeks before a quit date and is often used for up to 6 months post quit.

Varenicline is a partial agonist at the α4β2 neuronal nicotinic acetylcholine receptor; it both relieves craving and withdrawal and, unlike other medications, reduces the reinforcing effects of nicotine by blocking dopaminergic stimulation responsible for smoking reinforcement/reward. While there were concerns about the FDA-issued black box warnings for neuropsychiatric adverse events (depressed mood, agitation, changes in behavior, suicidal ideation, and suicide), more recent research of varenicline in psychiatric patients has found that psychiatric symptoms were not worsened and there is no evidence of an increased suicide risk. Furthermore, a 2014 FDA review did not support the original warning. A second black box warning is for cardiovascular adverse events ("a small, increased risk of certain cardiovascular adverse events in people who have cardiovascular disease"). Clinicians should counsel patients on these two black box FDA warnings and monitor for symptoms related to mental status and cardiac status.

Combined Psychosocial and Pharmacological Therapy

Combining medication and behavior therapy increases quit rates over either therapy alone. Also, combination medications have higher efficacy than monotherapy. Combination examples include using nicotine patch and short-acting NRT, bupropion and NRT, and varenicline and bupropion or varenicline and NRT. The USPHS Clinical Practice Guidelines recommend, contingent on the absence of medical contraindication, the use of at least one of the seven FDA-approved front-line agents for smoking cessation in conjunction with counseling to increase the likelihood of a successful quit attempt. Patients should be encouraged to use both pharmacological and psychosocial treatment approaches; however, medications can be effective in the absence of psychosocial therapy; thus, the latter should not be a requirement for the former.

Ms. H is a 45-year-old patient with schizophrenia who smokes 35 cigarettes per day. She began her cigarette use at approximately 20 years of age during the prodromal stages of her first psychotic break. During the first 20 years of treatment, no psychiatrist or physician advised her to stop smoking, largely because they believed that she could not stop.

At 43 years of age, her primary care physician recommended smoking cessation, especially with her history of asthma and heart disease. The patient attempted to stop on her own but this lasted only 48 hours, partly because her housemates and friends smoked. During a routine medication check, her psychiatrist *asked* about her smoking history and *advised* that she stop smoking, and the patient described her prior attempts. The psychiatrist *assessed* for motivation, using a 1 to 10 scale of motivation to quit and discussing where Ms. H was. He also had her breathe into a CO meter to demonstrate decreased oxygen capacity of her lungs and how that affects her asthma and heart disease. The psychiatrist and the patient discussed ways to avoid smokers and had the patient announce her intent to quit and request her friends to try not to smoke around her and to offer encouragement for her attempt to quit. The psychiatrist also noted that she became irritable, slightly depressed, and restless and had insomnia during prior cessation attempts and thus recommended medications for *assistance*. The patient chose to use nicotine patch plus nicotine gum as needed.

The psychiatrist had the patient call 2 days after her quit attempt as had been *arranged* at the appointment. At this point, the patient stated that the patch and gum were helping. One week later, the patient returned after having relapsed back to smoking. The psychiatrist praised the patient for not smoking for 4 days. After some motivation enhancement Ms. H was still reluctant to try again immediately.

At the next routine visit 2 months later, the psychiatrist *asked* about smoking again as it was in the assessment and treatment plan. After *advice* and *assessment*, the psychiatrist and patient listed several activities that the patient could do to avoid being around friends who smoked, phoned the patient's boyfriend to ask him to assist the patient in stopping, asked the nurses on the inpatient ward to call the patient to encourage her, plus *arranged* for the patient to be in a support group for the next 4 weeks. This time the psychiatrist prescribed the medication varenicline (Chantix). The patient was followed with 15-minute visits for each of the first 3 weeks. She had two "slips" but did not go back to smoking and remained an ex-smoker.

Ms. H's psychiatrist was correct in using pragmatic plans to help the patient overcome specific problems, in following the patient with short visits or phone calls, and in recommending medication. The total amount of time spent with the patient on smoking was approximately 3 hours and there was reimbursement; more importantly the intervention was important to decreasing her morbidity and mortality and was unlikely to be given by other care providers.

Tobacco Use Disorder Treatment of Psychiatric Patients and Other Special Populations

Psychiatric patients who are trying to stop smoking should be encouraged to undertake psychosocial and pharmacological therapies because these patients experience greater withdrawal symptoms and have a poorer ability to stop as a result of more efficient smoking and higher levels of nicotine. In addition, these patients should be followed weekly to detect any possible exacerbation of their psychiatric problem; often the perceived exacerbation results from withdrawal from nicotine and can be managed with NRT. Newer research demonstrates that smokers with mental illness do not have worsening of clinical symptoms of unipolar depression, bipolar disorder, posttraumatic stress disorder, or schizophrenia with tobacco treatment. Patients on smoke-free psychiatric wards can have withdrawal symptoms that require treatment. Treating TUDs in psychiatric patients not only reduces illness and lengthens life but can also present other benefits such as more discretionary income, less discrimination, and more options for relationships and employment.

People with SUDs should also be encouraged to quit smoking. Current evidence suggests that, for the large majority (but not all) of patients, stopping smoking does not threaten sobriety. Some newer studies suggest that quitting smoking can increase likelihood of maintaining sobriety from other drugs.

Non-nicotine chemicals in tobacco smoke activate cytochrome P450 liver cytochrome (CYP1A2), thereby decreasing the levels of several medications. As a result, smoking cessation increases the concentrations of these medications. Many of these medications are psychiatric medications, and often the increase can be clinically significant in psychiatric patients; for example, haloperidol (Haldol), clozapine (Clozaril), and fluvoxamine (Luvox) concentrations increase 30 to 40 percent with abstinence. Other examples are more antipsychotics (olanzapine and fluphenazine), antidepressants (amitriptyline, nortriptyline, imipramine, clomipramine, and trazodone), benzodiazepines, and caffeine. Therefore, it is important to monitor psychiatric medication doses in smokers where quantity smoked is changing. With a relapse to smoking in a patient on clozapine, for example, clozapine blood levels can be reduced and psychiatric symptoms may re-emerge and can be life threatening. Similarly, smoking cessation in someone maintaining steady caffeine consumption can incur caffeine "intoxication," which will feel like severe nicotine withdrawal, including agitation and restlessness. It is also useful to motivate patients to quit smoking with the potential that they may decrease psychiatric medication doses.

Another special population includes the pregnant smoker: the health of a pregnancy and outcomes in children born to smokers or born to

environments with smoke exposure should be addressed. Smoking can even make it harder for a woman to become pregnant and once pregnant, can increase the risk of preterm (early) delivery, stillbirth (death of the baby before birth), low birth weight, sudden infant death syndrome (SIDS or crib death), ectopic pregnancy, and orofacial clefts in infants. There may also be an association that children born to smoking mothers can also have increased risks of mental retardation, learning problems, and ADHD. Fathers are not immune from risks: Smoking can affect sperm and sexual function, which can reduce fertility and also increase risks for birth defects and miscarriage. The use of medications for smoking cessation during pregnancy is debatable. Whether the harm from smoking during pregnancy is due to nicotine, CO, or other chemicals in cigarette smoke is unclear. The concentrations of nicotine from medications are much lower than those from smoking, plus no CO and tar exposure occurs. Often, medications are reserved for pregnant smokers who have failed at cessation without medication.

Nonabstinence Goals

Given the highly addictive nature of tobacco, some tobacco users may not be able to or want to try to fully quit tobacco. The idea of using less to have less harm or to use an e-cigarette or another nicotine product is raised by some tobacco users. Harm reduction (using less tobacco) has been a method of gradual reduction for lower motivated individuals; however, this strategy often comes with compensatory smoking (i.e., smoking each cigarette more intensely to obtain the necessary nicotine), which appears to minimize any potential benefit of these strategies. Although some have feared that offering or agreeing to the patient's plan for reduction of smoking as a treatment goal would undermine motivation to quit, reduction in tobacco use appears to increase the chance of later quitting and may help motivates some tobacco users initially not willing to quit to try later to quit. Thus, agreeing to monitor a reduction method may not improve health but may lead to cessation by continuing to engage and work with the patient on the ultimate goal of cessation.

PREVENTION AND POLICY INITIATIVES

Policy initiatives to reduce smoking prevalence have focused on tobacco taxation, smoking restrictions, restriction of tobacco advertising, counter advertising via mass media, attacks on the tobacco industry as immoral, restriction of sales to minors, and warnings on cigarette packages. Among these, the effect of taxes is especially large and replicable and influences initiation among youth. Whether public health approaches in the United States have prompted the less dependent, less psychiatrically ill smokers to quit and left a more hardcore group of smokers who will require treatment

in addition to social pressure is widely debated due to the paucity of evidence.

A major policy advancement in many psychiatric settings has to do with tobacco-free public buildings and spaces. In 1993, JCAHO banned smoking in hospitals. In addition to reducing access to smoking, thereby compelling a smoker into a quit attempt or at least reduction in cigarettes smoked, environmental tobacco smoke is reduced and tobacco related cues are eliminated as well.

Given that TUDs have their onset typically in adolescence, general school-based tobacco prevention programs have been a logical approach to consider. There may be opportunities to consider high-risk adolescent groups, including those with psychiatric disorders being treated in mental health settings. Tobacco use often precedes mental illness and/or other SUDs. Strategies in school-based settings have focused on preventing the initiation of tobacco or nicotine-related products (such as e-cigarettes) through methods such as education (e.g., about risks of smoking), denormalizing smoking (e.g., correcting overestimates of prevalence of smoking), skills training (e.g., refusal training), and advocacy (e.g., students demonstrate against tobacco industry). Among these, only denormalizing and skills training have good empirical support; however, even with these interventions, the long-term effect of school-based programs is not large.

References

Agaku IT, King BA, Husten CG, et al. *Tobacco Product Use Among Adults—United States, 2012–2011 [Internet]*. Atlanta, GA: Centers for Disease Control and Prevention. Morbidity and Mortality Weekly Report (MMWR). 2014:542. Available at http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6325a3.htm.

American Association of Medical Colleges. *Physician Behavior and Practice Patterns Related to Smoking Cessation, Summary Report*. Washington, DC: Association of American Medical Colleges; 2007.

American Cancer Society. *How are Cigars Different From Cigarettes? [Internet]*. Atlanta, GA: American Cancer Society; 2014 [cited 2015 February 17]. Available at http://www.cancer.org/cancer/cancercauses/tobaccocancer/cigarsmoking/cigar-smoking-cigars-vs-cigarettes.

Anthenelli RM, Benowitz NL, West R, et al. Neuropsychiatric safety and efficacy of varenicline, bupropion, and nicotine patch in smokers with and without psychiatric disorders (EAGLES): a double-blind, randomised, placebo-controlled clinical trial. *Lancet*. 2016;387:2507–2520.

Balbuena L, Tempier R. Independent association of chronic smoking and abstinence with suicide. *Psychiatr Serv*. 2015;66(2):186.

Benowitz NL. Neurobiology of nicotine addiction: implications for smoking cessation treatment. *Am J Med*. 2008;121:S3.

Brewer JA, Elwafi HM, Davis JH. Craving to quit: psychological models and neurobiological

mechanisms of mindfulness training as treatment for addictions. Psychology of Addictive Behavior. *J Soc Psychol Addict Behav*. 2013;27(2):366.

Centers for Disease Control and Prevention (CDC). Quitting smoking among adults—United States, 2001–2010. *MMWR Morb Mortal Wkly Rep*. 2011;60(44):1513.

Chelladurai Y, Singh S. Varenicline and cardiovascular adverse events: a perspective review. *Ther Adv Drug Safety*. 2014;5(4):167.

Fiore M. and 2008 PHS Guideline Update Panel, Liaisons, and Staff. Treating tobacco use and dependence: 2008 update U.S. Public Health Service Clinical Practice Guideline executive summary. *Respir Care*. 2008;53(9):1217.

Hitsman B, Moss TG, Montoya ID, George TP. Treatment of tobacco dependence in mental health and addictive disorders. *Can J Psychiatry*. 2009;54(6):368.

*Lasser K, Boyd JW, Woolhandler S, Himmelstein DU, Mccormick D, Bor DH. Smoking and mental illness: a population-based prevalence study. *JAMA*. 2000;284:22.

Lawn S, Pols R. Smoking bans in psychiatric inpatient settings? A review of the research. *Aust N Z J Psychiatry*. 2005;39(10):866.

Mcmillen R, Maduka J, Winickoff J. Use of emerging tobacco products in the United States. *J Environ Public Health*. 2012;2012:1–8.

Neuman MD, Bitton A, Glantz SA. Tobacco industry influence on the definition of tobacco related disorders by the American Psychiatric Association. *Tob Control*. 2005;14(5):328.

Prochaska JJ, Delucchi K, Hall SM. A meta-analysis of smoking cessation interventions with individuals in substance abuse treatment or recovery. *J Consult Clin Psychol*. 2004;72(6):1144.

*Prochaska JJ, Hall S M, Bero LA. Tobacco use among individuals with schizophrenia: what role has the tobacco industry played? *Schizophr Bull*. 2008;34:555.

Roberts E, Eden Evins A, McNeill A, Robson D. Efficacy and tolerability of pharmacotherapy for smoking cessation in adults with serious mental illness: a systematic review and network meta-analysis. *Addiction*. 2016;111:599–612.

Rodgman A, Perfetti TA. *The Chemical Components of Tobacco and Tobacco Smoke*. Boca Raton, FL: CRC Press, Taylor & Francis Group; 2009.

Rogers ES, Smelson DA, Gillespie CC, et al. Telephone smoking-cessation counseling for smokers in mental health clinics: a patient-randomized controlled trial. *Am J Prev Med*. 2016;50:518–527.

Schroeder MJ, Hoffman AC. Electronic cigarettes and nicotine clinical pharmacology. *Tobacco Control*. 2014;23:30.

Taylor G, McNeill A, Girling A, Farley A, Lindson-Hawley N, Aveyard P. Change in mental health after smoking cessation: systematic review and meta-analysis. *BMJ (clinical Research Ed.)*. 2014;348:1151.

Thurgood SL, McNeill A, Clark-Carter D, Brose LS. A systematic review of smoking cessation interventions for adults in substance abuse treatment or recovery. *Nicotine Tob Res*. 2016;18:993–1001.

U.S. Department of Health and Human Services (USDHHS). *The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General*. Atlanta: Georgia: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014.

*Williams JM, Ziedonis DM, Vreeland B, et al. A wellness approach to addressing tobacco in mental health settings: learning about healthy living. *Am J Psychiatr Rehabil*. 2009;12:352.

World Health Organization (WHO). *Western Pacific Regional Office (WPRO), 2014-last update, Tobacco Fact Sheet [Homepage of World Health Organization, Western Pacific Region]*, [Online]. 2014. Available at http://www.wpro.who.int/mediacentre/factsheets/fs_201203_tobacco/en/.

Young-Wolff KC, Kline-Simon AH, Das S, Mordecai DJ, Miller-Rosales C, Weisner C. Smoking trends among adults with behavioral health conditions in integrated health care: a retrospective cohort study. *Psychiatr Serv*. 2016;67(9):996–1003. doi: 10.1176/appi.ps.201500337.

Zevin S, Benowitz NL. Drug interactions with tobacco smoking. *Clin Pharmacokinet*. 1999;6:425.

Zhao-Shea R, Liu L, Pang X, Gardner PD, Tapper AR. Activation of GABAergic neurons in the interpeduncular nucleus triggers physical nicotine withdrawal symptoms. *Curr Biol*. 2013;23(23):2327.

*Ziedonis DM, Williams JM, Steinberg MB, Foulds J. Addressing tobacco addiction in office-based management of psychiatric disorders: practical considerations. *Primary Psychiatry*. 2006;13(2):51.

*Ziedonis D, Hitsman B, Beckham JC, et al. Tobacco use and cessation in psychiatric disorders: National Institute of Mental Health Report. *Nicotine Tob Res*. 2008;10(12):1691.

▲ 11.9 Opioid Use and Related Disorders: From Neuroscience to Treatment

SEAN X. LUO, M.D., PH.D., AND ADAM BISAGA, M.D.

The term **opioid** describes a class of psychoactive compounds, both naturally occurring and chemically synthesized, that are related to *opiates*: alkaloids compounds found as natural products in the opium poppy plant, *Papaver somniferum*. Opioids are among the oldest known drugs worldwide: opium poppy use predates recorded history. Opioids can be used medically for pain relief, but are also commonly misused by means of nonprescribed use of pharmacological opioids or illicitly obtained heroin. Opioid use disorders affect a wide range of individuals around the world, and the United Nations Office on Drugs and Crime (UNODC) estimates that 0.3 percent of the global population, or more than 15.6 million people, abused opioids in 2009. In the United States, opioid use disorder has dramatically increased in the past few decades, and now constitutes a large public health problem affecting the young and the old, the affluent and the poor, professionals, and the unemployed. At the same time, there have been significant advances in knowledge about opioids and their effect on the CNS. Practical and effective treatment strategies that target specific brain abnormalities have emerged. There is now a consensus in the medical community that opioid use disorder is a chronic, relapsing disorder involving multiple brain circuits, and is amenable to pharmacological and psychosocial interventions.

HISTORICAL BACKGROUND

The Greek word "opos," which means juice, suggests people may have first discovered the liquid from the unripe poppy seed capsule when it is notched. The written record of the early history of opium use and abuse is ambiguous. The preparation described by Homer in *The Odyssey,* given by Helen to Telemachus and his friends to help them forget their grief over Odysseus' absence was attributed to Homer's imagination by Theopharastus (300 BC). Sumerians cultivated poppies and isolated opium at the end of the third millennium bc Arab traders brought opium to India and China around the 8th century ad and between 10th and 13th centuries opium travelled from Asia Minor to Europe. Thomas Syndeham noted in 1680 that "Among the remedies which it has pleased Almighty God to give man to relieve his sufferings, none is so universal and efficacious as opium." However, written records of drug abuse and tolerance can be found in Turkey, Egypt, Germany, and England, and the problematic practice of smoking opium reached its apogee in China in the mid-17th

century after tobacco smoking was banned. Efforts by the Chinese government to criminalize opium use conflicted with British and French commercial imperatives, leading to the Opium Wars in the 19th century, which greatly influenced post-Colonial era geopolitical landscape.

Sertürner used advanced chemical methods in 1806 to isolate the active ingredient in opium, naming it morphine after Morpheus, the god of dreams. Codeine was isolated from opium a few years later. Over the next century, these purified compounds gradually replaced crude opium for medicinal purposes, although nonmedical use of opium, especially through smoking, still persisted in many parts of the world. The first semisynthetic opium derivative—diacetylmorphine or heroin—was introduced into medicine in 1898 and marketed as an effective cough suppressant, and at the time, thought to have very little abuse liability. The first purely synthetic morphine-like opioids, meperidine was serendipitously discovered in 1939 in a search for atropine replacement. Methadone was discovered in the 1940s and introduced to medical practice in the subsequent decades.

The recognition of opioid withdrawal and addiction by the medical community, especially in the United States, generally began in the 19th century. It was noted at the time that morphine addiction was common among White middle-class women who were given it frequently for numerous medical symptoms such as neurasthenia and cough, as well as in the postpartum period. A variety of remedies were used to treat opioid withdrawal and "cure" morphine addiction. However, these were often more dangerous than untreated withdrawal and were not effective to prevent relapse. A variety of other opioid-related issues became prominent in the early 20th century: prevalent opium smoking by Chinese immigrants working on railroads, intravenous (IV) use of opioids with the newly introduced hypodermic needle, an increase in illegal activities surrounding drug use, and the importation of opioids from other countries. A variety of medications on the market included opioids components surreptitiously and was mostly unregulated. As drug use was gradually linked to crime, the U.S. Federal Government first passed the Pure Food and Drug Act of 1906, which required accurate labeling of component contents. The subsequent Harrison Narcotics Tax Act of 1914 demarcated a starting point for federal regulation and taxation of the production, importation, and distribution of opiates and Coca products. The Harrison Act had a profound impact on persons exhibiting physical and psychological symptoms of tolerance and withdrawal because it was interpreted by the Treasury Department to exclude the provision of using opioids to treat these symptoms for medical purposes. Clinics that had been established specifically to provide therapeutic morphine to patients with opioid use problems were shut down by the early 1920s. Doctors were urged to avoid treating patients with these symptoms and some were prosecuted for doing so. Detoxification efforts

were plagued by frequent relapses and many sought alternative access to opioids. As medical prescription of opioids decreased, illicit traffic increased substantially, providing mostly morphine and heroin to addicted individuals. Users were frequently arrested for possession and sale of opioids and incarcerated.

In the 1930s, the U.S. Public Health Service established two federal hospitals—at Lexington, Kentucky, and Fort Worth, Texas—to deal with the growing number of federal prisoners who were also addicted to opioids. These hospitals provided long-term residential treatment of prisoners and some voluntary patients, and a significant amount of early addictions research was conducted there. Follow-up studies of these prisoners found that relapse rates were high even after long periods of abstinence and treatment. In light of the lack of effective treatment strategies and increasing criminal activities surrounding drug use, worldwide penalty for the illicit use of opioids increased after World War II. The Single Convention on Narcotic Drugs of 1961 established an international treaty to prohibit production and supply of specific drugs that was ratified by most member countries. The United Nations Commission on Narcotic Drugs, as part of World Health Organization, was empowered to add, remove, and transfer drugs among the treaty's four schedules of controlled substances. This treaty also required criminalization of cultivation, production, distribution, possession, and trafficking of scheduled opioids. Governmental regulation of opioids through criminal law became a worldwide phenomenon. In the United States, the Controlled Substance Act of 1970 served as the national implementing legislation for the Single Convention on Narcotic Drugs, and heroin was categorized as a Schedule I drug, the most stringent category of regulation (indicating no valid medical use). Other opioids, such as morphine, oxycodone, and methadone, were designated Schedule II drugs.

Since the 1960s, it has been increasingly recognized that legal regulation is an insufficient intervention both for individuals with problematic opioid use and for the community as a whole. Attention has been shifting toward the development of better treatment programs for opioid-dependent individuals. The Synanon organization was founded in California in 1958 and constituted an early model of therapeutic community. The 12-step programs, such as Narcotics Anonymous (NA), which grew out of AA, first emerged in Los Angeles in the early 1950s. Since the late 1960s federal and state agencies have increasingly focused on treatment research. Methadone maintenance began as a research program at Rockefeller University in 1964. This effective treatment was delivered in the setting of outpatient methadone maintenance clinics, a model that subsequently expanded and became widely adopted throughout the United States and many parts of the world. Problematic opioid use became a condition amenable to medical treatment by physicians and allied

healthcare professionals.

Naltrexone was synthesized in the 1960s and later used to treat opioid addiction, was approved by the FDA in 1984. Buprenorphine was initially formulated in the United Kingdom in 1978 as an injection to treat severe pain, and a sublingual formulation was released in 1982, providing a convenient treatment of opioid use disorders. As the availability of residential substance abuse treatment decreased significantly in the 1990s, the U.S. Congress passed the Drug Abuse Treatment Act of 2000, which permitted physicians with qualifying training to prescribe approved Schedule III, IV or V medications (such as buprenorphine), for the treatment of addiction or detoxification outside of a methadone program setting. Since this law was passed, office-based treatment of opioid use disorder with buprenorphine has expanded substantially as the number of individuals presenting for the treatment increased with the growing epidemic of misuse of prescription opioids in the United States.

COMPARATIVE NOSOLOGY

The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) divides opioid and related disorders into five main diagnoses: opioid use disorder, opioid intoxication, opioid withdrawal, other opioid-induced disorders, and unspecified opioid-related disorder (also see Clinical Features section below).

Opioid use disorder is diagnosed for users who meet at least two of the eleven criteria in the previous 12-month period (Table 11-9–1). This disorder primarily reflects a pattern of compulsive, prolonged self-administration of opioid substances that have no legitimate medical purpose or at doses greatly exceeding the amount needed for a medical condition, with continuing use of opioids despite significant substance-related problems. The opioid use disorder diagnosis is further specified by a set of additional modifiers: (1) the duration of remission: either *early remission*, where none of the criteria are met for more than 3 months but less than 12 months, or *sustained remission*, where none of the criteria have been met for more than 12 months; (2) whether the individual is *on maintenance therapy*: if the individual is maintained on a medication specifically for the treatment of opioid use disorder; and (3) whether the individual is in a *controlled environment*: where access to opioids is restricted. The diagnosis is also coded based on symptom severity. Mild opioid use disorder denotes the presence of two to three symptoms, moderate, four to five symptoms, and severe six or more symptoms.

Table 11.9–1.
Summarized DSM-5 Diagnostic Categories and Criteria for Opioid Use Disorder

Category	Criteria
Impaired control	► Opioids used in larger amounts or longer than intended ► Unsuccessful efforts or desire to cut back or control opioid use ► Excessive amount of time spent obtaining, using, or recovering from opioids ► Craving to use opioids
Social impairment	► Failure to fulfill major role obligations at work, school, or home as a result of recurrent opioid use ► Persistent or recurrent social or interpersonal problems that are exacerbated by opioids or continued use of opioids despite these problems ► Reduced or given up important social, occupational, or recreational activities because of opioid use
Risky use	► Opioid use in physically hazardous situations ► Continued opioid use despite knowledge of persistent physical or psychological problem that is likely caused by opioid use
Pharmacological properties	► Tolerance as demonstrated by: increased amounts of opioids needed to achieve desire effect; diminished effect with continued use of the same amount ► Withdrawal as demonstrated by: symptoms of opioid withdrawal syndrome; opioids taken to relieve or avoid withdrawal

Opioid intoxication is diagnosed when an individual who recently used an opioid exhibits clinically significant problematic behavioral or psychological changes, such as dysphoria, impaired judgment, or psychomotor agitation or retardation that developed during, or shortly after opioid use. These individuals also develop symptoms such as drowsiness or coma, slurred speech, and impaired attention and memory. *Opioid withdrawal* is diagnosed when an individual ceases or reduces opioid use, or receives an opioid antagonist after a period of opioid use. Three of the nine withdrawal symptoms must develop and cause clinically significant distress for this diagnosis to apply. In addition, DSM-5 also contains unspecified opioid-related disorder to denote clinically significant opioid-related disorder that do not meet any criteria as above, as well as substance/medication-induced depressive disorder, substance/medication-induced psychotic disorder, substance/medication-induced bipolar and related disorder, and substance/medication-induced anxiety disorder, which describe a set of psychotic, depressive, bipolar-related or anxiety-related symptoms that occur soon after exposure or withdrawal from opioids (or other substances).

The DSM-5, released in 2013, represents a significant revision and update of criteria from the DSM-IV-TR and ICD-10 (WHO: International Statistical Classification of Diseases and Related Health Problems 10th revision) definitions of opioid-related disorders. SUD in DSM-5 combines the DSM-IV categories of substance abuse and dependence into a single disorder measured on a continuum from mild to severe, as it has been increasingly observed that these conditions are not dichotomous entities. In DSM-5, each substance is now addressed as a separate use disorder but the overarching set of criteria is the same. Although DSM-IV only requires

one symptom to diagnose a substance abuse disorder, DSM-5 requires at least two symptoms from a list of 11. Drug craving is added to the list, and law enforcement is eliminated because legal contexts differ across cultures and DSM-5 aims to provide a set of criteria that can be applied internationally.

There is a high level of agreement between the DSM-5 and ICD-10, and both have as their central feature an emphasis on the compulsive aspect of drug-using behavior. However, ICD-10 continues to make the distinction between abuse (harmful use category) and dependence. In DSM-5, the code for opioid use disorder, mild now corresponds to the ICD-10 definition of harmful use of opioids (F11.10), while the DSM-5 criteria for opioid use disorder, moderate and severe correspond to the ICD-10 code for opioid dependence syndrome (F11.20). It is important to note that within both classification systems, physical dependence, as evidenced by tolerance or withdrawal, is a criterion that is neither necessary nor sufficient for a diagnosis of either opioid use disorder or opioid dependence. One of the rationales for the new DSM-5 terminology of opioid use disorder, rather than opioid dependence, is to avoid the confusion of symptoms of physical dependence versus the syndrome of drug dependence, as well as avoiding the potentially stigmatizing label of *addiction*, which is commonly used in the lay literature to describe more severe forms of an SUD, but also has many other meanings. A more neutral label of a "Use Disorder" was selected and the diagnosis of *addiction* does not appear in either DSM-5 or ICD-10, though it remains in the title of the DSM-5 "Substance-Related and Addictive Disorders" chapter.

ICD-10 differs from DSM-5 in that the concept of harmful use is limited to physical and mental health (e.g., hepatitis, overdose, and skin abscess) and specifically excludes social impairments with the reasoning: "Harmful patterns of use are often criticized by others and frequently associated with adverse social consequences of various kinds. The fact that a pattern of use or a particular substance is disapproved of by another person or by the culture, or may have led to socially negative consequences such as arrest or marital arguments, is not in itself evidence of harmful use."

EPIDEMIOLOGY

Extensive epidemiological surveys from both the United States and around the word track trends in opioid use and misuse. In the United States, there are several national and public surveillance systems that provide useful information on different aspects of opioid use and addiction over time: the annual NSDUH estimates prevalence of substance use, abuse, and dependence within community dwelling persons age 12 years and older; the Monitoring the Future (MTF) survey reports on the prevalence of substance use as well as attitudes about drug use among a national sample of 8th-, 10th-, and 12th-grade students, and the TEDS collects demographic

and drug use data from persons of all ages admitted to substance abuse treatment facilities receiving state or federal funding (with some exceptions). Internationally, the World Drug Report provides an annual review of the major developments in drug markets for the various drug categories, from production to consumption. These publically available data sources, combined with new methods in data mining and visualization, provide an easier way for researchers and public health workers to more rapidly respond to emerging trends in opioid misuse. In this section, heroin and other opioids are discussed separately, as large epidemiological differences have been observed for these two categories of substances.

Prevalence of Heroin Use and Dependence

In 2012, the NSDUH reported the lifetime prevalence of heroin use in the United States among persons aged 12 years or older to be approximately 4.5 million people or 1.8 percent of the population. In 2012, there were 669,000 persons aged 12 or older (0.3 percent of the population) who had used heroin for the first time within the past 12 months. This represents a significant increase from 2002, when the prevalence of use in the past year was 404,000 persons. Prevalence of heroin use in the past month has almost doubled (from 166,000 to 335,000) since 2002. The estimate of the number of people who have DSM-IV opioid abuse or dependence within this group has also significantly increased over the past decade and is currently estimated at 467,000 (0.2 percent of the population).

For adolescents (8th, 10th, and 12th graders), the MTF survey shows stabilization of lifetime prevalence of heroin use. The lifetime prevalence measure was approximately 1 percent in 2013 as compared to approximately 2 percent at peak prevalence in 1999. Annual prevalence of heroin use also continues to decline, and in 2013 is approximately 0.5 percent. The lower prevalence among school age children and the increased prevalence in adults present a conundrum, possibly due to a number of factors including the clustering of distribution and geographical concentration.

According to a TEDS (last available for 2007), heroin continues to constitute a large fraction of inpatient treatment admissions, about 14 percent, and is behind alcohol and marijuana. Heroin appears to affect all ethnic groups, with admissions in 2007 at 60.4 percent for Whites, 22.1 percent for Blacks, and 22.4 percent for Hispanics, and predominantly affects male patients (68.5 percent of admissions). Treatment of heroin abuse is common among inpatients aged 50 and older. In 1992, 8.9 percent of all substance abuse treatment admissions for persons aged 50 to 54 years old were for heroin use; this percentage increased to 18.4 percent by 2005, stabilizing at 13.7 percent in 2007. This cohort of baby boomers, aging into retirement, may have a high prevalence of heroin use relating to

more prevalent illicit drug use in their adolescence of the 1960s.

Prevalence of Prescription Opioid Misuse and Dependence in the United States

Illicit prescription opioid use has become an epidemic in the United States. The NSDUH defines prescription opioid misuse as use by a person who has not been prescribed the medication by a physician (regardless of the reason why the person took it) *or* use to achieve euphoria. The lifetime, past-year, and past-month prevalence of opioid misuse has increased significantly in the past decade. In 2002, it was estimated at 29.6 million lifetime, 11 million past-year, and 4.4 million past-month individuals who misused prescription opioids. In 2012, these numbers increased to 37.0 million lifetime (14.2 percent of the population), 12.5 million past-year (4.8 percent of the population), and 4.9 million past-month individuals (1.9 percent of the population) misusing these drugs. Past-year (8.5 percent) and past-month (3.3 percent) misuse was highest among 26- to 29-year-olds. In 2012, approximately 1.8 million people initiated misuse of prescription opioids within the previous 12 months, a number that is above all other illicit substances except marijuana. In 2012, an estimated 2.1 million people were opioid dependent, nearly double the 2006 figure, and almost a triple the 2002 figure. The ratios of number of persons with prescription opioid dependence divided by the number of persons reporting past 30-day use (47.8 percent) and lifetime use (5.7 percent) are similar to these ratios for heroin indicating that the risk of developing opioid dependence is similar in people who misuse heroin or prescription opioids.

Data from MTF also indicate an increasing prevalence of nonmedical prescription opioid use in school age children and adolescents. In the new "narcotics other than heroin" category, prevalence in 12th grade adolescents increased from 6.1 percent in 1992 to 12.2 percent in 2012. In the 2007 TEDS, opiates other than heroin account for 90,489 total admissions. The admitted patients are predominantly White (88.6 percent) with a slight preponderance of males (53.4 percent). Admissions are most common among individuals between the ages of 25 and 29.

Worldwide Trends

Abuse of opioids is now a leading global public health burden. According to the WHO World Drug Report, the global area of illicit opium cultivation in 2013 stood at close to 300,000 hectares and has been increasing since 1998, when estimates became available. Afghan heroin is increasingly reaching Oceania and South-East Asia and lucrative markets in Western and Central Europe. Global past-year use of all illicit opiates including both heroin and prescription opiates has slowly increased from 1999 to 2012 and currently stands at about 0.7 percent. Opioid use remains high in North America and Oceania, with prevalence rates of 4.3 percent and 3 percent,

respectively, with the main increase in the United States. While recent and reliable estimates are not available from Asia and Africa, many experts from countries in those regions also perceive an increase in opioid use (Fig. 11.9–1). The use of opiates (heroin and opium), however, remained stable globally, with 12.8 to 20.2 million past-year users.

ETIOLOGY AND PATHOGENESIS

Opioid use disorder is a complex phenomenon that involves multiple etiological factors at different levels of analysis: genetic and epigenetic influences, synaptic plasticity, alterations in neural circuits, deficits in coping skills, abnormal patterns of thoughts and behavior, and dysfunction in family and social surroundings. The abnormal behavior characteristic of addiction represents the manifestation of a common final pathway and a heterogeneous set of underlying causal factors. These factors influence each other in space and in time, in complex and nonlinear ways. For example, peer pressure (a social factor) may commonly precipitate the initiation of heroin use, but differences in neural circuits, likely generated, in part, by differences in genetic variations, eventually yield a difference in drug-related experiences, specifically the balance between positive effects (e.g., euphoric) versus negative effects (e.g., nausea and vomiting) that determine further drug-taking behavior. The interaction between the effects of repeatedly self-administered substance and the resulting experiences subsequently cause synaptic plasticity, either reinforcing/strengthening or extinguishing/weakening selected brain circuitries, which in turn affects individual's behavior—either continuing with a drug-using peer group or a withdrawal from the group.

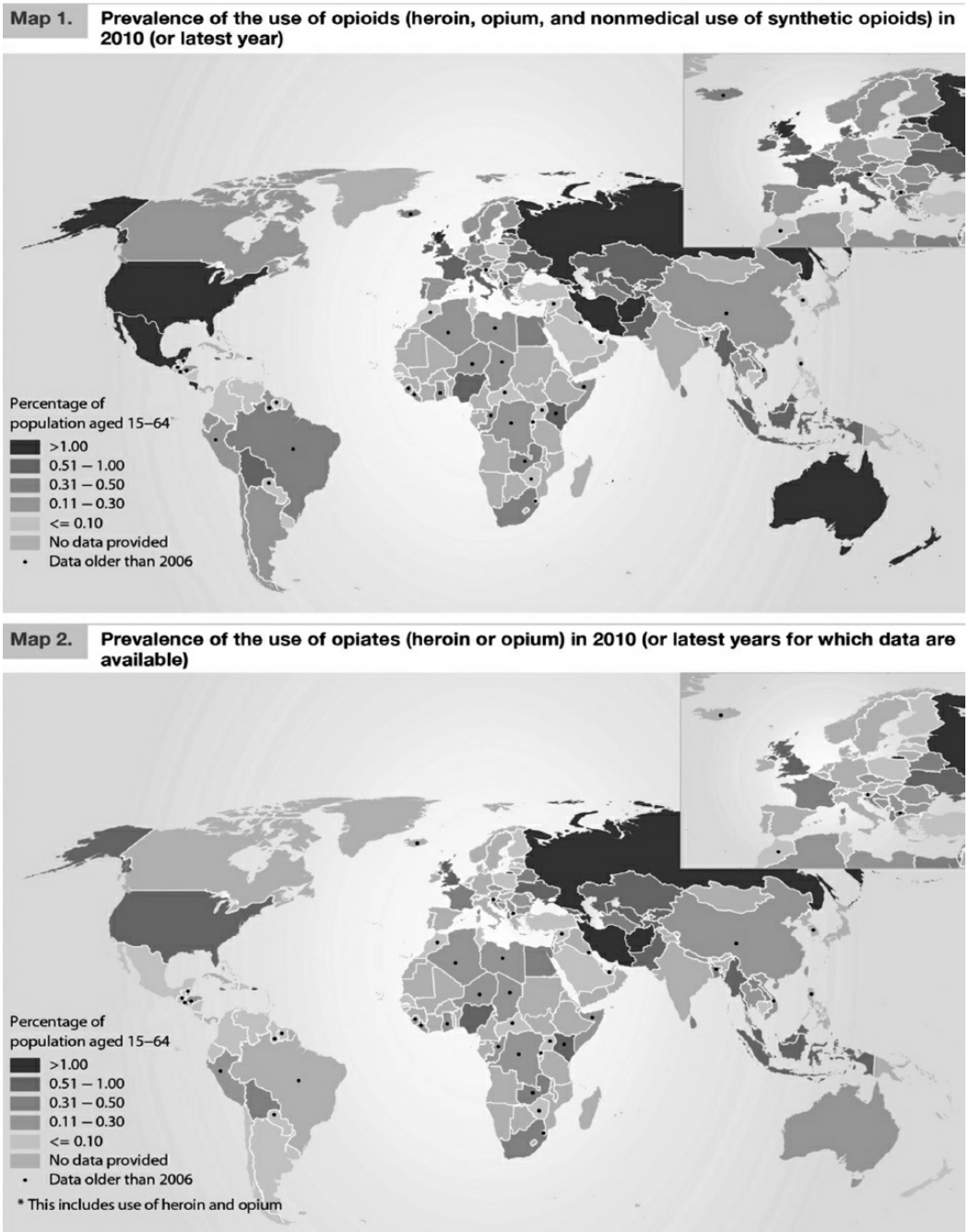

FIGURE 11.9–1. Worldwide distribution of opioid use. Map 1 shows the use of all opioid chemicals including medical opioids, showing a significant distribution in the United States as well as Mexico. Map 2 shows the prevalence of illicit opioid use (heroin and opium), showing a peak distribution Russia as well as in Central Asia, as well as wide prevalence in United States and Western Europe. (Adapted from UNODC, World Drug Report, Research and Trend Analysis Branch. Division for Policy Analysis and Public Affairs. United Nations Office on Drugs and Crime. Vienna, Austria.)

Initiation of opioid use is dependent on social and cultural context. When opioids are injected or inhaled, the rapid onset of action in the brain induces a *rush* or *flash*, a sudden, brief but intensely pleasurable sensation lasting 1 to 2 minutes. This euphoric sensation subsequently leads individuals to seek repetition of the drug-taking (reinforcement). Repeated

drug-taking and euphoria causes opioid tolerance; changes in neural circuitry that makes it harder to achieve the same intensity of euphoria with an equal amount of opioids. An abrupt cessation of use by an individual that has been using repeatedly over an extended time period produces opioid withdrawal, a set of physical symptoms such as nausea, vomiting, and intense craving signifying the presence of physiological dependence. The emergence of withdrawal on stopping use and euphoria with continuing use are the basis of a long-term learning process that assigns salience to external and internal cues and experiences associated with positive and negative drug effects (conditioning) and alters the behavioral repertoire of the addicted individual in the form of increased impulsivity or decreased response to social rewards.

Neurobiology of Opioid Use Disorders

Opioids and Opioid Receptors. A variety of opioid compounds exist (Fig. 11.9–2). *Morphine*, *codeine,* and *oxycodone* are commonly prescribed medications for pain relief, but can also be diverted for nonmedical use. *Heroin* or diacetylmorphine is the opioid that is most commonly abused by injection; common street names include dope, horse, smack, and tar. Opium extracted from the poppy plant, which is often smoked, is more commonly abused in the Middle East and Asia. Most commonly encountered opioids act at one of the three main opioid receptor systems (mu, kappa, and delta). The term *endorphin* refers to a subclass of opioid-like substances, consisting of peptides that are produced naturally by the human body, including dynorphins, enkephalins, endorphins, endomorphins, and nociceptin.

Opioid receptors are a group of G protein-coupled receptors (GPCRs) with opioids as ligands and respond to a family of endogenous opioids (endorphins). The opioid receptors are distributed widely in the brain, the spinal cord, and digestive tract. There are four currently known types of opioid receptors, delta (δ), kappa (κ), mu (μ), and nociceptin receptor (NOP). All four types of receptors have 7 transmembrane domains (TM1-7). Amino-acid residues at the same positions in TM3, TM6, and TM7 interact with ligands to produce downstream effects—the "message" region of the ligands. Another region—the "address" region—produces opioid selectivity and occupies one of two different areas of the binding pocket, depending on the type of opioid (Fig. 11.9–3).

The downstream activity of opioid receptors results in two known signal transduction pathways: classical and alternative. The classical signal transduction relates to opioid receptor by inducing G-proteins to interact with an inward rectifying potassium channel, Kir3. This process causes cellular hyperpolarization and inhibits neural activity. All four opioid receptors also cause a reduction in Ca^{2+} content both intracellularly and in synaptic vesicles and synaptosomes, with compensatory upregulation of

vesicular Ca^{2+} content during the development of opiate tolerance. In addition, μ-, δ-, and κ-opioid receptors inhibit adenylyl cyclase activity and inhibit the cAMP-dependent Ca^{2+} influx. More recently, activation of opioid receptors has been shown to induce β-arrestin recruitment and binding in the alternative transduction pathway. Arrestin molecules bind to phosphorylated opioid receptors to regulate their traffic and this interaction plays a role in the development of tolerance. Phosphorylated arrestin-bound opioid receptor complex also recruits alternate signal transduction cascades, including MAP Kinases, inducing numerous other complex downstream molecular events.

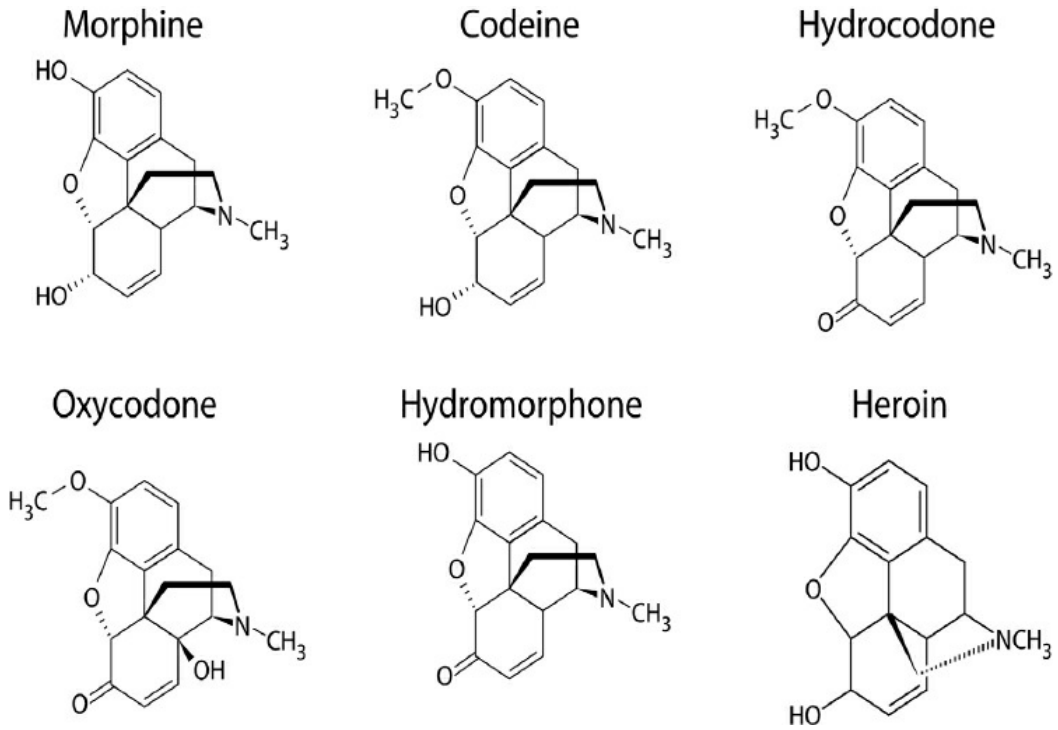

FIGURE 11.9–2. Chemical structures of common medicinal and illicit opioids. They share a three-ring phenanthrene structure as well as a methylamine group. All opioids are chemically related to morphine, the parent compound, and bind to the opioid receptors. Hydromorphone, hydrocodone, oxycodone, and heroin are semisynthetic. These compounds all have analgesic effects and can potentially be substances of abuse, though they have differing pharmacokinetic and pharmacodynamic properties.

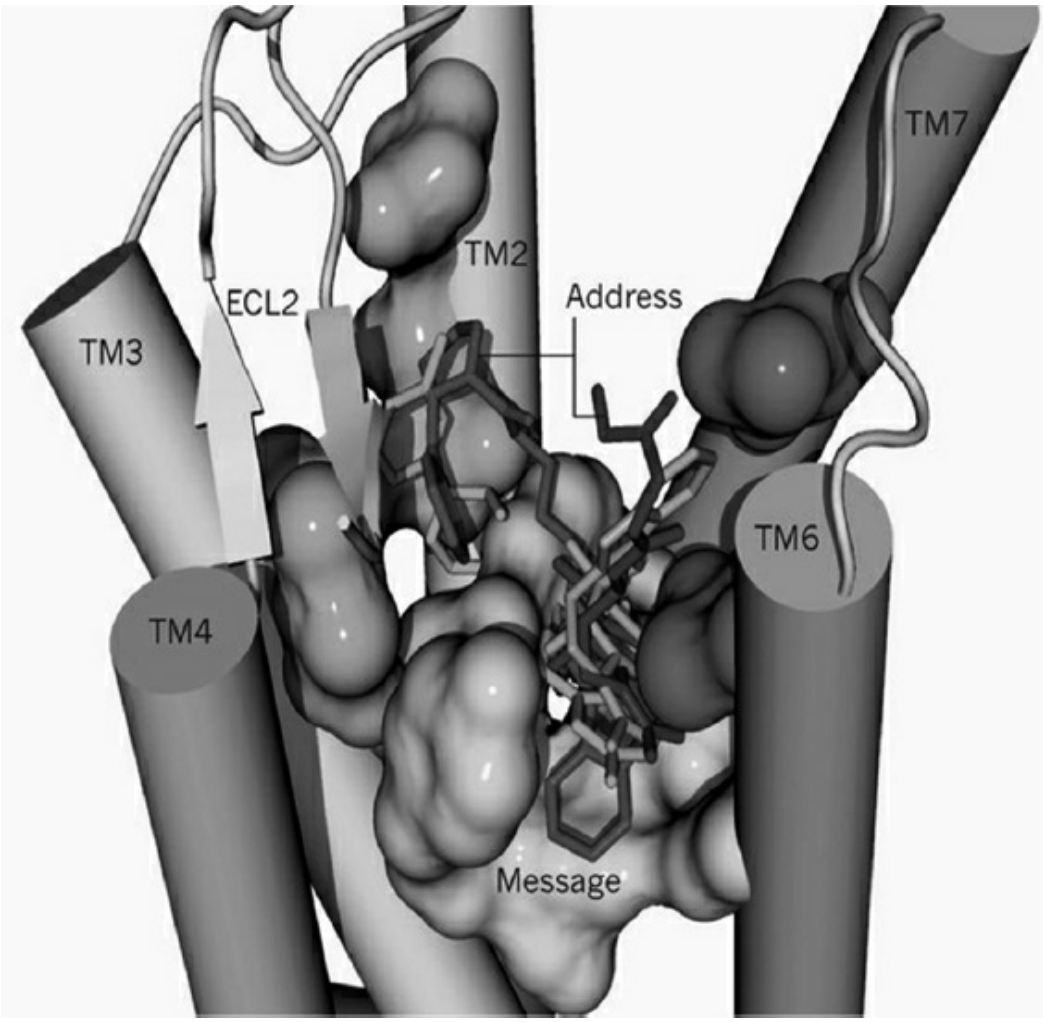

FIGURE 11.9–3. Common structural elements of the opioid receptors. Several features are shared by all four opioid receptor types: seven transmembrane helices (TM1–7), five shown (*gray cylinders*). ECL2 is a β-hairpin loop region; the *arrows* represent β-sheets. Ligands are depicted as stick representations inside the binding pocket. The cyan surface indicates the amino-acid residues from TM3, TM6, and TM7 that interact with the ligands "message" regions. The magenta surfaces indicate the residues from TM6 and/or TM7 that interact with the "address" region—responsible for opioid selectivity. (Adapted from Filizola M, Devi LA. Structural biology: how opioid drugs bind to receptors. *Nature*. 2012;485(7398):314–317.)

The molecular and cellular mechanisms mediating the development of tolerance to morphine continue to be debated. One hypothesis proposes that downregulation of opioid receptors following chronic agonist exposure induces tolerance. More recent in vivo studies show that downregulation does not occur consistently with each and every agonist. Newer evidence suggests that the receptors may be desensitized and uncoupled from downstream signaling pathways and that chronic exposure changes levels of second messengers, such as cAMP, and upregulation of adenylyl cyclase.

Opioid Receptor Pharmacology. Opioids can interact with opioid receptors in several ways (Fig. 11.9–4). An opioid *agonist* is a molecule that binds to an opioid receptor and induces downstream signal transduction

events. Many opioids produce actions similar to that of morphine (a prototypical μ-opioid receptor agonist), and some bind to various receptor subtypes in a pattern that is distinct from that of morphine, producing a large, diverse array of possible patterns of opioid agonism. Molecules that bind to opioid receptors but initiate no signal transduction action are opioid *antagonists* at those receptors. Naloxone or naltrexone acts as μ-opioid receptor antagonists. When both antagonist and agonist are present, they compete against each other at the binding site. *Partial agonists* such as buprenorphine bind to opioid receptors but only produce a limited effect in intensity or maximal effects compared to a full agonist.

Agonist effect at μ-opioid receptors is associated with opioid use disorders. The direct physiological action of μ-agonists includes analgesia, respiratory depression, miosis, euphoria, drowsiness, decreased ability to concentrate, changes in endocrine and other functions regulated by the hypothalamus, and increased tone of smooth muscle in the GI tract. Agonists at the κ-opioid receptor often produce dysphoria and analgesia but no significant pupillary change. Heroin, morphine, hydromorphone, and oxycodone are all μ-opioid receptor agonists, and they differ from each other another primarily in metabolism, pharmacokinetics, and relative affinities at other opioid receptors, which may produce considerable differences in their subjective effects. For instance, heroin (diacetylmorphine) is more potent and more lipid soluble than morphine, crossing the blood–brain barrier more rapidly and producing a more rapid onset of subjective effects. Patients are often able to discriminate heroin from oxycodone.

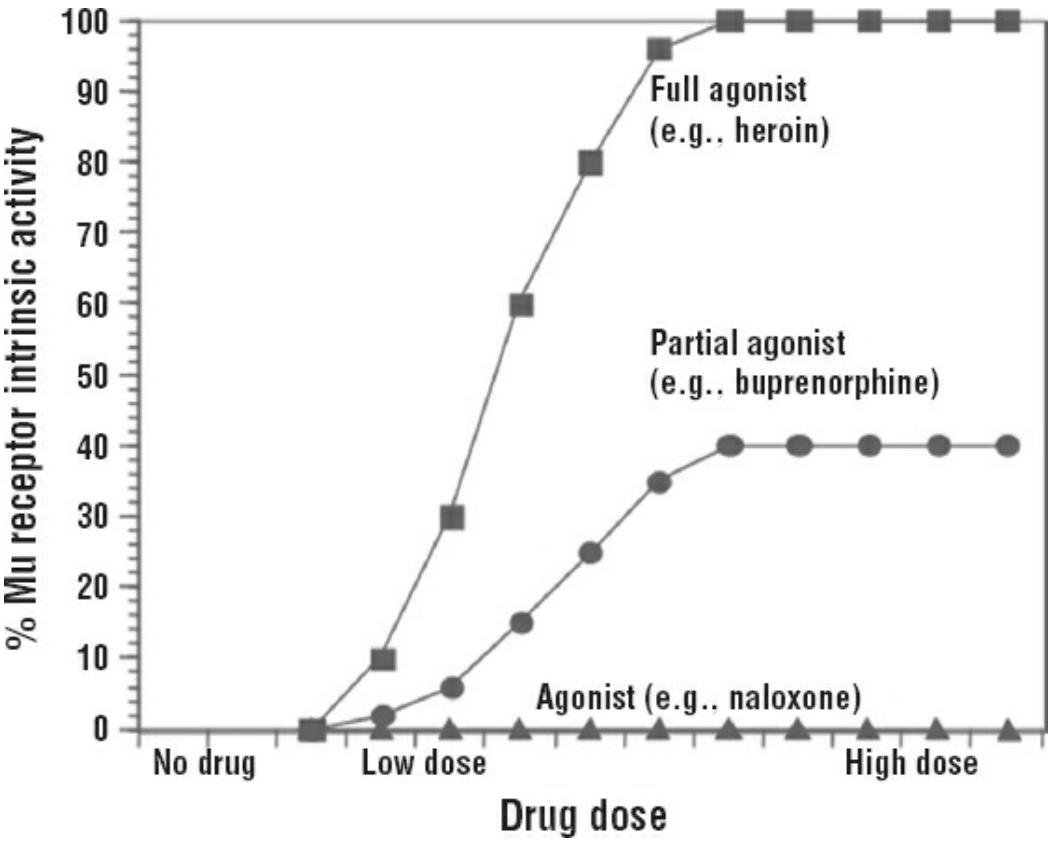

FIGURE 11.9–4. Schematic of pharmacological action of different classes agents affecting μ-opioid receptors. A full agonist, such as heroin, induces all of the capacity for intrinsic activity of the opioid receptor. A partial agonist, such as buprenorphine, only allows for half of the activity to proceed, blocking other full agonists from achieving full biological effect. Antagonist, such as naloxone and naltrexone, eliminates the potential for any other agent from binding to the receptor and activating the receptor.

Neural Circuits in Opioid Use Disorders. A variety of animal models have provided a pathway to dissect the neural circuits that mediate the effect of opioid molecules on the opioid receptors to affect animal behavior. The focus of experimental work has been converging on three aspects of opioid associated behavior: self-administration, remodeling occurring with the chronic exposure, and adaptation and relapse.

CIRCUITS IN OPIOID SELF-ADMINISTRATION. The μ-opioid receptor mediates the reinforcing effects of heroin and morphine and blockade of μ-opioid receptor but no other opioid receptors attenuate opiate self-administration. Knockout mice with no μ-opioid receptor expression in the brain do not readily associate contextual cues when these cues are paired with opiate administration (conditioning). Opioid drug conditioning is mediated by neuronal hyperpolarization and neuronal inhibition in multiple structures, including those that provide major afferent input to the ventral tegmental area (VTA); that is, nucleus accumbens (NAc), ventral palladium, prefrontal cortex (PFC), amygdala, mediodorsal thalamus, and the pedunculopontine tegmental nucleus. This interconnected network is also

known as the cortical–pallidal–striatal circuit, and is characterized by a high density of opiate receptors reciprocal connections. Site-specific blockade of μ-opioid receptor in the VTA or NAc attenuates heroin self-administration. Damage to the NAc can abolish opioid self-administration. These findings support the hypothesis that this circuitry mediates the rewarding effects of these agents.

Dopamine appears to play a central role in mediating the rewarding effect of opioids in this circuit as μ-opioid receptor agonists activate VTA dopaminergic neurons, and dopamine release in the NAc is enhanced following heroin self-administration. Electrophysiological studies support the idea that overlapping and heterogeneous subpopulations of NAc and medial PFC neurons are activated during opiate self-administration and mediate its reinforcing effects. There is also evidence that a diverse group of neurons releasing GABA and glutamate, as well as the cannabinoid systems, are involved in mediating the reinforcing effects of opiates.

CIRCUITS INVOLVED IN CHRONIC OPIOID EXPOSURE. Cessation of chronic opioid exposure alters the activity of mesolimbic dopaminergic neurons and the subsequent decreased dopamine neurotransmission may be related to anhedonia and dysphoria during the early phase of opioid withdrawal. The network level reconfigurations observed in chronic opioid users remain unclear, but dopamine receptor agonists can attenuate some signs of withdrawal in animal models. Chronic opioid use also alters the activity of NAc medium spiny neurons that are primary targets of dopamine neurons. These neurons are involved in a diverse number of transmitter release functions, including GABA and the opioid peptides dynorphin and enkephalin, as well as substance P, and project to the ventral palladium and VTA. Alterations in these output neurons may also contribute to the aversive consequences of opiate withdrawal.

Opioid receptor antagonists, such as naloxone, are hypothesized to function primarily in blocking the μ-opioid in NAc and other regions, including the amygdala, VTA, and locus coeruleus. Selective δ-opioid receptor antagonists can precipitate affective but not somatic components of withdrawal syndrome in animal models. Dynorphin is the endogenous ligand for the κ-opioid receptor and agonists at these receptors produce dysphoria and negative effective mediated in part by the activation of NAc opioid receptors and inhibition of dopamine release. Increased dynorphin expression after abstinence is related to opiate-induced upregulation of the cAMP pathway and the induction of the gene transcription regulator, CREB. Long-term abstinence from opiates in a previously dependent animal often produces a phenomenon of hypersensitivity in dopaminergic neurons—with more dopamine release subsequently with a same dose of drugs. This is possibly related to behavioral sensitization to various drugs of abuse as well as plays the role in the reinstatement of compulsive drug-

seeking behavior.

Chronic opioid use also induces circuit reconfiguration in basolateral amygdala, locus coeruleus, and the bed nucleus of the stria terminalis involving a variety of neurotransmitter such as glutamate and norepinephrine. Each of these adaptations may be involved in expression of the aversive effects of the opiate-withdrawal syndrome.

CIRCUITS MEDIATING RELAPSE. A large number of diffuse neural circuits appear to be engaged in reinstatement of opioid use in animal models of relapse. Factors that are involved in relapse in humans such as stress, renewed contact with the drug, or drug cues and increased drug availability reliably reinstate drug-seeking behavior in laboratory animals. The mesolimbic dopamine system plays a major role in mediating the priming effects of opiates in drug-seeking behaviors. Intra-NAc infusion of amphetamine, which stimulates dopamine release, can reinstate heroin seeking, while dopamine D2 receptor antagonists block heroin-induced reinstatement of drug seeking. Glutamatergic neurons may also be involved in reinstatement. Stress can reliably reinstate heroin seeking in rats, potentially mediated through CRF releasing neurons and may result from a distinct neural circuit outside of the mesolimbic dopamine system. Cue induced relapse has been relatively poorly understood. Dopamine plays a less central role in this phenomenon and PFC glutamate AMPA receptor plasticity is important in cue-induced relapse to heroin seeking.

Human Neuroimaging of Opioid Use Disorders. Key evidence supporting the hypothesis that drug addiction is a chronic disease of the brain comes from brain imaging studies examining a number of critical neural circuits. These studies provide a link between animal models of drug self-administration and human clinical studies. The existing evidence suggests that across all substances there is a general circuit that is engaged in drug self-administration, independent of the specific substance of abuse. There are then distinct substance-dependent circuits that mediate reinforcing effects of a specific substance. In opioids, the complex opioid receptor systems throughout the brain interact with the dopamine system generating and maintaining the complex clinical phenomenon of opioid use disorders. The main brain imaging techniques are structural MRI, functional MRI (fMRI), resting state MRI, MR spectroscopy (MRS), positron emission tomography (PET), and SPECT. Each of these methods can reveal different aspects of brain structure and function.

Converging brain imaging data suggest a distributed neural network of brain regions regulated by the μ-opioid receptor system to mediate pain, reward, motivation, and integration of emotional value. These regions include the anterior cingulate and PFC, the anterior and posterior insular cortex, orbitofrontal cortex, ventral striatum, ventral palladium, thalamus,

hypothalamus, amygdala, periaqueductal gray (PAG), and others. The κ-opioid receptor system is involved in suppressing the antinociceptive actions of opioids at the μ-opioid receptor sites and is the most widely expressed receptor system in the brain. Multiple studies report a push-and-pull antagonism between the μ- and κ-receptor systems. The μ-opioid receptor system is also associated with the positive reinforcing effect of other abused drugs including cocaine and alcohol, in mediating their reward-related behaviors, and increasing binding in μ-opioid receptor is correlated with self-reported craving in cocaine.

PET studies show that dopamine is implicated in opioid use disorders in a number of different ways. As the central neurotransmitter that mediates reward, dopamine D2/3 receptor availability is reduced in heroin-dependent subjects at baseline, which is an effect seen similarly in the cocaine-, alcohol-, tobacco-, and methamphetamine-dependent subjects. Presynaptic dopamine release is also blunted in heroin-dependent subjects (Fig. 11.9–5). There might also be a correlation between dopamine release and the subjective effects of heroin. These results suggest that the dopamine system may be involved in regulating a common underlying circuit in producing a general addiction phenomenon independent of the specific substance of abuse.

Task-dependent fMRI studies in opiate-dependent subjects suggest that a decreased activity of a cingulo–fronto—parietal–cerebellar neural circuit may be involved in opioid use, with the region dorsal anterior cingulate (dACC) in particular as being associated with performance deficits on tasks requiring behavioral control and error estimation. MRS suggests that opiate-dependent patients also have reduced concentrations of *N*-acetyl aspartate (NAA) and glutamate/glutamine within the dACC. Resting state fMRI in opioid-dependent subjects shows a significant decline in connectivity of NAc to a number of areas including laterobasal amygdala, hippocampus, parahippocampus, and the orbital frontal and mid-cingulate cortices.

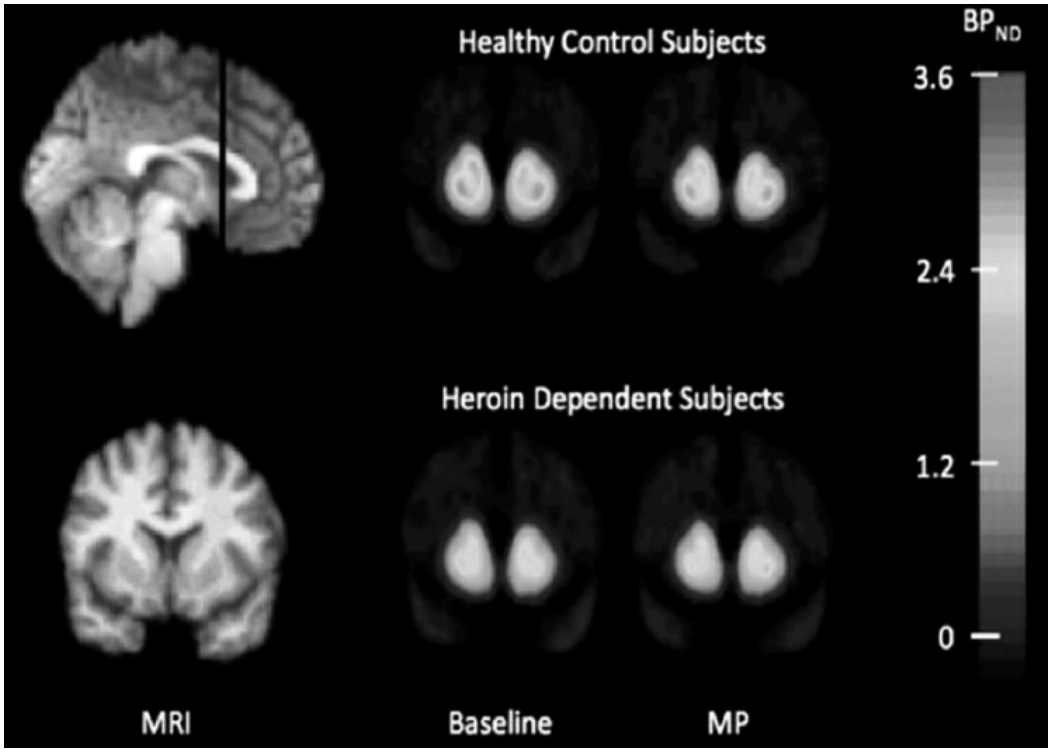

FIGURE 11.9–5. Deficit in dopamine neurotransmission in heroin-dependent subjects. [11C]raclopride is a dopamine radioligand and its binding potential (BPND) maps averaged across all subjects. Left column: representative magnetic resonance imaging (MRI) from one subject showing the level of the coronal positron emission tomography slices (bottom). The black line in the sagittal view (top) shows the level of the coronal slices. Center and right columns: healthy control subjects (top) and heroin-dependent subjects (bottom) at baseline (center) and after methylphenidate (MP) (right). (Adapted from Martinez D, Saccone PA, Liu F, et al. Deficits in dopamine D2 receptors and pre-synaptic dopamine in heroin dependence: commonalities and differences with other types of addiction. *Biol Psychiatry*. 2012;71(3):192–198. (Open access))

While current neuroimaging evidence strongly suggests that there is an underlying circuit level abnormality that mediates the initiation and maintenance of problematic opioid use, the precise mechanisms, both at the level of networks activities of cortical and striatal neurons and their transmitter release, continue to be unclear. There is also no current neuroimaging method for predicting and prognosticating outcomes for opioid use disorders. These important questions represent active areas of research.

Molecular Genetics and Genomics of Opioid Use

The genetic contribution to the development and maintenance of opioid dependence is significant. Twin studies estimate that around 50 to 60 percent of the liability for heroin addiction is genetic. There is also evidence that genetic vulnerability is made up of common factors that increase the likelihood of developing any drug addiction (16 percent) and factors that increase the likelihood of developing opioid addition specifically (38 percent).

Linkage study for opioid dependence shows clustered linkage on chromosome 17, as well as linkage on a region of chromosome 14 overlying the neurexin 3 gene (NRXN3). They also identified a male-specific linkage peak on chromosome 10q and potentially a signal chromosome 4q. Functional polymorphism in *OPRM1* gene (Asn40 Asp), which encodes the mu-opioid receptor, and *OPRD1* (encoding delta receptor) and *OPRK1* (encoding kappa receptor) has been extensively studied in haplotype-specific case-controlled designs but results have been inconsistent. A potentially robust positive association between *OPRD1* variants and opioid dependence has been reported, but requires replication.

The first genome-wide association study for opioid dependence disorder was completed in 2013, signaling the beginning of the application of genomics approaches to opioid use disorders. The study enrolled 5,697 subjects and identified significant associations from multiple loci; the most compelling results involved genes in the potassium signaling pathways as well as genes involved in calcium signaling and long-term potentiation. The functional and clinical significance of these implicated targets will likely become important areas of future research.

Psychosocial Factors

Genetic and neural circuits eventually contribute to addiction behavior through their interaction with complex environmental and psychosocial factors. Social attitudes, peer pressure, individual temperament, and drug availability are all determinants of the emergence of problematic opioid use. There is less social stigma associated with prescription opioids relative to heroin, and prescription opioids are believed to be safer in comparison to heroin and are more widely available. As prescription opioid users develop a number of functional impairments, their opioid use often transitions to heroin and the IV route of administration due to lower cost and increasing tolerance for slower acting agents. The presence of a significant number of opioid users in a neighborhood, town, or city also can create a subculture of opioid experimentation. In the United States, areas in which prevalence of opioid use disorder is high also often have associated high crime rates, unemployment rates, and poorly performing school systems. While correlation does not equal causation, these factors can all be conceptualized as contributing to reduced resistance to opioid use and may worsen the prognosis and treatment efficacy.

Epidemiological studies show that people who have opioid use disorders place more value on independence and less value on academic achievement. Patients with heroin use disorders have more novelty seeking, risk-taking, and impulsivity traits and show significant signs of delinquency before their first experimentation with opioids. A significant proportion of heroin users meet the criteria for antisocial personality disorder.

Most urban heroin users come from single-parent families, and poor parental functioning has been consistently reported as a risk factor for opioid addiction. Family risk factors involve disruption of family structure and disrupted parental relationships. Cross-generational alliances between the drug user and one parent against another parental figure are common. Dysfunctional family equilibrium can threaten the recovery process, and an assessment of family dynamics is an important component in the assessment and treatment of someone with opioid use disorder.

Integrative Approach

Across different levels of analyses, the etiology of opioid use disorder is now conceptualized as a complex interaction of genes, neural circuits, and environment. While the underlying mechanistic details remain subjects of active research, several consensus hypotheses have emerged from a large amount of converging experimental data. First, it appears that there is a large component of risk of developing opioid use disorder that derives from genetic influences, and these genetic variations change the pattern of biochemical signaling pathways involved in both opioid receptor and nonopioid receptor systems. These genetic factors of risk interact with environmental risk factors, which derive from sensitization during critical period, traumatic experiences, and adverse social and economic conditions. The combination of genetic and environmental factors interacting in complex ways eventually leads to a heterogeneous set of neural circuit level abnormalities involving multiple cortical circuits and transmitter systems, likely involving abnormal release and reinforcement of the dopamine neurotransmission. These systems subsequently influence both behavior and the rest of the body through its influence on the stress system (pituitary–adrenal system). The initiation of opioid use subsequently induces neuroadaptations in systems regulating reward, motivation and drive, memory, and behavioral control. Changes in other neural systems lead to the development of tolerance and physiological dependence. As a result, affected individuals develop a range of problematic behavior: not only maintaining drug use but also relapse following cessation of use. These behaviors lead to criminality, poverty, and deficit in social support, which in turn strengthen drug-seeking and -using behavior.

This complex, multifactorial etiological model stipulates that abnormalities in brain function are behind the etiology and maintenance of the opioid use disorder. It is important that patients and their families recognize the significance but also limitations of this model. Opioid use disorder is a chronic, relapsing-remitting behavioral brain disorder that is not a character flaw or moral weakness. However, the central role of neural circuit level abnormalities does not imply that patients have no control over choices they make, particularly choices related to their own recovery. The brain is an incredibly plastic organ, and neural circuits can be remodeled in

response to environmental changes, new behaviors administration of medications. Effective treatment of opioid addiction should therefore target all of these etiological factors in a multilevel fashion.

DIAGNOSIS AND CLINICAL FEATURES

Opioid Intoxication and Overdose

The DSM-5 criteria for opioid intoxication are broadly inclusive, and in clinical situations opioid intoxication symptoms can vary in severity and presentation. Usually initial euphoria is followed by apathy, dysphoria, psychomotor agitation or retardation, and impaired judgment. Pupillary constriction, drowsiness, and slurred speech are common. There may also be X-ray evidence of pulmonary edema and frothing at the mouth. Other physical signs include cyanosis, cold clammy skin, and decreased body temperature. Blood pressure is decreased but only falls dramatically with severe anoxia. Cardiac arrhythmias have also been reported and may be related to anoxia or to the presence of quinine as an adulterant in the opioid. A marked tolerance usually develops over time to subjective intoxicating effects of opioids.

On occasion intoxication is followed by an overdose, a state characterized by unresponsiveness to external stimulation (coma). Signs of opioid overdose include coma with pinpoint, but reactive pupils and respiratory depression or apnea. In severe cases, severely depressed respiration can develop, which can lead to anoxia and large, nonreactive pupils.

Opioid Withdrawal

Opioid withdrawal develops after the cessation of or reduction in opioid use that has been heavy and prolonged. The withdrawal syndrome also occurs when an opioid partial agonist such as buprenorphine or an opioid antagonist such as naloxone or naltrexone is administered to a patient who is regularly taking an opioid agonist. The withdrawal symptoms are usually the physiological opposite of the intoxication symptoms. Most patients undergoing withdrawal complain of anxiety and irritability, restlessness, “achy feeling” in the legs and back, and increased sensitivity to pain. Few objective signs may be present in milder withdrawal. In more severe withdrawal patients present with nausea or vomiting, abdominal cramps and diarrhea, severe muscle aches, lacrimation and rhinorrhea, yawning, pupillary dilation, and increased sweating and chills. Fever, severe insomnia, confusion, and delirium can occur in severe withdrawals.

Opioid withdrawal can be either *spontaneous* (stopping daily use) or *precipitated* (due to the ingestion of an antagonist or partial agonist). Spontaneous withdrawal starts 6 to 8 hours after last use of the short-acting opioid (heroin, morphine, oxycodone, hydrocodone) and 24 hours

after last use of a longer acting opioid such as methadone. Withdrawal severity peaks around 36 to 48 hours and resolves after 5 to 7 days for short-acting agents and after 2 or more weeks in case of methadone. Administration of an antagonist precipitates opioid withdrawal within minutes after parenteral administration (naloxone) and 2 to 3 hours after oral administration (naltrexone). Opioid withdrawal precipitated by naltrexone is usually more severe than spontaneous withdrawal but resolves faster, generally within 2 to 3 days. Usually a period of several weeks of continuous receptor occupation is required before spontaneous withdrawal is clinically significant, but a precipitated withdrawal can occur even after a few doses of an opioid.

There is considerable variability between individuals in presenting symptoms during opioid withdrawal. The severity of opioid withdrawal also varies across individuals and in time, as well as based on the pharmacological characteristics of the opioid, such as its potency, degree of opioid agonism (e.g., full versus partial agonist), dose, the duration of use, and the rate at which the opioid is removed from the receptors (e.g., a slow taper versus sudden abstinence). In general, the higher the potency, agonist efficacy, and dose, the longer the duration of opioid use; and the faster the removal of opioid from the receptor, the more intense the withdrawal syndrome. Standardized instruments have been developed to quantify the severity of opioid withdrawal symptoms in clinical practice. The Clinical Opiate Withdrawal Scale (COWS) is the most commonly used and utilizes resting pulse rate, sweating, restlessness, pupil size, pain, rhinorrhea, GI symptoms tremor, yawning, anxiety, and piloerection as anchor points (Table 11.9–2).

Opioid withdrawal symptoms can occur after abrupt discontinuation of maintenance treatment with buprenorphine, a partial opioid receptor agonist, but such symptoms are generally less severe than seen with cessation of a full agonist. Some patients report significant insomnia, irritability, and depression. These symptoms tend to resolve within 7 to 10 days. Drug effects from larger doses of buprenorphine (16 to 32 mg or more sublingually) can persist for several days due to its long half-life, and withdrawal symptoms can be delayed for 96 hours or more. This feature of buprenorphine has led to its use on an intermittent basis in some patients (e.g., at intervals of 48 hours).

Following the acute opioid withdrawal syndrome, there also appears to be a protracted period of physiological abnormality that lasts for months, called *protracted or postacute withdrawal syndrome*. It is characterized by depressed mood and anhedonia, irritability, labile mood, stress intolerance, insomnia, and recurrent urges to use opioids. This protracted syndrome is postulated to result from alterations and remodeling of the neural circuitry associated with withdrawal, such as changes in the serotonergic and dopaminergic systems. This syndrome, while not formally

included in the DSM-5, is likely a significant contributor to relapse to opioid use.

Table 11.9–2.
Clinical Opioid Withdrawal Scale (COWS)

Resting Pulse Rate: (record beats per minute)
Measured after patient is sitting or lying for 1 minute
0 pulse rate 80 or below
1 pulse rate 81–100
2 pulse rate 101–120
4 pulse rate greater than 120

Sweating: *Over past ½ hour not accounted for by room temperature or patient activity*
0 no report of chills or flushing
1 subjective report of chills or flushing
2 flushed or observable moistness on face
3 beads of sweat on brow or face
4 sweat streaming off face

Restlessness: *Observation during assessmen*t
0 able to sit still
1 reports difficulty sitting still, but is able to do so
3 frequent shifting or extraneous movements of legs/arms
5 Unable to sit still for more than a few seconds

Pupil size
0 pupils pinned or normal size for room light
1 pupils possibly larger than normal for room light
2 pupils moderately dilated
5 pupils so dilated that only the rim of the iris is visible

Bone or joint aches: *If patient was having pain previously, only the additional component attributed to opiates withdrawal is scored*
0 not present
1 mild diffuse discomfort
2 patient reports severe diffuse aching of joints/muscles
4 patient is rubbing joints or muscles and is unable to sit still because of discomfort

Runny nose or tearing: *Not accounted for by cold symptoms or allergies*
0 not present
1 nasal stuffiness or unusually moist eyes
2 nose running or tearing
4 nose constantly running or tears streaming down cheeks

GI Upset: *Over last ½ hour*
0 no GI symptoms
1 stomach cramps
2 nausea or loose stool
3 vomiting or diarrhea
5 Multiple episodes of diarrhea or vomiting

Tremor: *Observation of outstretched hands*
0 No tremor
1 tremor can be felt, but not observed
2 slight tremor observable
4 gross tremor or muscle twitching

Yawning: *Observation during assessment*

0 no yawning
1 yawning once or twice during assessment
2 yawning three or more times during assessment
4 yawning several times/minute

Anxiety or Irritability

0 none
1 patient reports increasing irritability or anxiousness
2 patient obviously irritable anxious
4 patient so irritable or anxious that participation in the assessment is difficult

Gooseflesh skin

0 skin is smooth
3 piloerection of skin can be felt or hairs standing up on arms
5 prominent piloerection

Score:
5–12 = **mild;**
13–24 = **moderate;**
25–36 = **moderately severe;**
more than 36 = **severe withdrawal**

From Wesson DR, Ling W. The Clinical Opiate Withdrawal Scale (COWS). *J Psychoactive Drugs*. 2003;35:253–259.

Opioid Use Disorder

DSM-5 opioid use disorder diagnosis spans a spectrum of problematic opioid use behaviors and combines the DSM-IV abuse and dependence categories. These criteria have physiological, behavioral, and cognitive components and have been organized into four categories: impaired control, social impairment, risky use, and pharmacological changes (Table 11.9–1). Two or more criteria are set as a threshold for the diagnosis and three levels of severity are specified (mild, moderate, and severe). The diagnostic criteria are primarily assessed during a detailed clinical interview, while repeated urine toxicology and toxicology analysis of hair can be used as corroborating evidence of chronic use. No single criterion is given special weight: the presence of tolerance and withdrawal are neither necessary nor sufficient for the diagnosis. The broad emphasis for these criteria is to specify the degree of impairment in social and occupational function due to a pattern of problematic opioid use, rather than on the timing, quantity, type, or route of use.

Tolerance to opioids' hedonic effects is a common reason for escalation of use, but even chronic heroin using patients continue to experience brief periods of euphoria 30 to 60 minutes each time that they receive injection. Ongoing opioid use is therefore not solely used to avoid withdrawal. Because of tolerance to physical symptoms of opioids such as respiratory depression and sedation, opioid effects may not be readily detected in chronic users by even the careful observer. While tolerance is advantageous for patients maintained on higher therapeutic doses of methadone or buprenorphine, it makes detection of opioid use difficult, especially in

certain vulnerable populations such as healthcare providers.

Opioid-Induced Mental Disorders. DSM-5 specifies a category of substance-induced disorders for a number of different substances that include a variety of opioid-induced mental disorders (psychotic, bipolar, depressive, anxiety, obsessive-compulsive, sleep, sexual, delirium, and neurocognitive disorders). Chronic opioid use can produce a number of psychiatric symptoms, and opioid-using patients can also have comorbid psychiatric diagnoses that preceded or are clearly independent of opioid use. The important differential considerations here therefore rely on a detailed history of chronicity and order of presence for these symptoms. The onset of opioid-induced psychiatric syndromes needs to be temporally connected with the substance use (within 1 month), as opposed to occurring prior to the initiation of substance use or during a prolonged period of abstinence, and opioids need to be capable of producing the specific syndrome. Clinically, it is common to see some degree of depressed mood during and for several weeks after opioid withdrawal. Opioid-induced mood disorder should not be diagnosed after opioid withdrawal unless the severity of mood disturbance exceeds what is normally encountered or persists for more than a few weeks and is of sufficient intensity to warrant independent clinical attention. Opioid-induced psychotic symptoms are rare, although have been occasionally observed in clinical practice, especially manifesting as derealization and occasionally hallucinations. Opioid-induced sleep disorder and opioid-induced sexual dysfunction are common, and both hypersomnia and insomnia are common complaints of patients who either have opioid use disorder or are maintained on opioids for therapeutic purposes. Common sexual dysfunctions for chronic opioid users include erectile dysfunction and orgasmic difficulties. While the presentation and consequences of opioid-induced mental disorders can be identical as in the case of the independent mental disorder, most opioid-induced disorders are likely to improve within days or weeks of abstinence.

Toxicology, Physical, and Laboratory Examination

There is currently no biomarker for the diagnosis of opioid use disorders. The most commonly used urine toxicology tests can detect recent opioid use. Urination is a major route for drug elimination and urine testing is a convenient alternative to testing blood. Drugs are first detectable in urine 2 to 6 hours after use. Morphine, heroin, codeine, and hydrocodone can be detected in urine for 1 to 3 days after use. Heroin metabolite (6-AM) can be detected in urine for up to 8 hours after use after which heroin can be detected as morphine (its metabolite). The ingestion of modest amounts of poppy seeds (of the type and amount used in baking) may also produce positive results on tests with a low threshold for detection. Fully synthetic

opioids such as methadone, fentanyl, buprenorphine, tramadol, and oxycodone require specific urine toxicology tests. Methadone can be detected in urine for up to 4 days after a single use episode and for longer with the repeated use. Analysis of hair clippings can provide information and offers an average measure of use over the preceding 2 to 3 months. Hair testing is a better indicator of chronic use as compared to urine but drug use within the previous week will not be detected in hair. Samples of oral fluids can detect recent opioid use with approximately the same sensitivity as urine testing but drugs can be detected in saliva immediately after smoking or injecting and within 1 to 2 hours after swallowing.

Patients with a history of IV opioid use should also be tested for hepatitis (B and C) and for HIV. Sharing straws for intranasal opioid use is a risk factor for hepatitis C. At the time of initial evaluation, it is also important to measure liver functions, complete blood count, and basic metabolic panel, as opioid using patients tend to have poor nutritional status, and acute opioid intoxication can cause metabolic abnormalities and leukocytosis. Chest X-rays may show evidence of pulmonary fibrosis if the person has been using injection materials contaminated with microcrystalline talc or cotton particulates. Methadone may prolong the QT interval on ECG in rare cases.

Physical exams are often normal though signs of intoxication or withdrawal may be present. Intranasal use of heroin may cause nasal irritation and necrosis. IV users often have needle marks throughout the body because they use injection sites that are concealed (e.g., soles of the feet). Infections (e.g., thrombophlebitis and skin abscess), lymphatic obstruction, and edemas are not uncommon. Subcutaneous and muscle tissue hardening can result from repeated intramuscular injections of meperidine (seen among health professionals). Skin lesions consisting of scars, skin atrophy, and hyperpigmentation suggest chronic use. Endocarditis may produce fever and heart murmurs. Signs of liver abnormalities such as stigmata and ascites should also be documented. Mental status examinations often show abnormalities in mood, affect, and cognition.

Differential Diagnosis

Diagnosing opioid use disorder is usually straightforward in clinical situations when an individual presents with a request for treatment and is open about the duration and frequency of use and the nature of associated issues. However, in earlier stages of the disease, collateral information from family and peers confirmed with toxicology results are necessary, as individuals have a tendency to under-recognize or diminish the existence and impact of drug-related problems. Patients with a long history of addiction who relapsed will try to hide the history of illness when trying to obtain opioids in the medical setting—engaging the family or reaching out

to previous medical providers is important for an accurate diagnosis and an adequate treatment plan.

Opioid use disorder can be misdiagnosed in patients with chronic pain or intermittent acute pain syndromes (e.g., sickle cell disease) who exhibit a range of problematic behavior such as arguing with medical staff, requesting higher doses and more potent opioids, and seeking medications from multiple doctors. Ongoing and poorly controlled pain can elicit these behaviors, a phenomenon called *pseudoaddiction*. Careful longitudinal observation, outside informants, and urine toxicology testing, alongside continued appropriate attempts at pain relief and treatment of co-occurring medical and psychiatric problems are required in generating the differential diagnosis. An empathic, compassionate, and nonargumentative approach to working with such patients is essential and physician–patient contract specifying the treatment plan and procedures to follow is helpful.

In most cases, opioid withdrawal is easy to diagnose, as its physiological signs are very typical. Diagnosing withdrawal in early phases when patients mostly report subjective symptoms of various aches, pains, and anxiety may be more difficult and further monitoring may be necessary. Poor response to the treatment of withdrawal with opioid agonists such as methadone suggests withdrawal from other substances; however, administering buprenorphine to an opioid-dependent individual who does not have clear signs of withdrawal can worsen the withdrawal. The presence of severe withdrawal (autonomic instability, confusion or seizures) without classic opioid-withdrawal signs raises the strong possibility of withdrawal from other substances such as alcohol or sedative-hypnotics. Toxicology evaluation is essential in all cases presenting with drug-withdrawal symptoms.

Similarly, opioid intoxication must be differentiated from mixed intoxications in which opioids only play a minor role. In general, a failure to respond to high doses of naloxone suggests that intoxication is caused by a nonopioid. Elderly patients may respond to opioids with dysphoria and confusion in *atypical opioid intoxications* rather than opioid-induced delirium or opioid-induced psychotic disorders, which are rare.

COURSE AND PROGNOSIS

The clinical course for opioid use disorders is highly variable. Some persons apparently can use opioids occasionally (e.g., several times a month) over periods of months or years without meeting the criteria for an opioid use disorder. This is the more common course for patients who use prescription opioids but is also seen in heroin users. Data from national surveys suggest that approximately one-third of recreational users develop moderate-to-severe opioid use disorder in their lifetimes, and the rates are lower in people who use prescription opioids recreationally. The NSDUH data show that 10 percent of individuals who report lifetime use of heroin

and 6 percent of individuals who report lifetime use of prescription opioids meet the DSM-IV criteria for opioid dependence in the past year. Even though the majority of recreational users do not go on to develop a use disorder, there is no reliable way to predict among first-time users on an individual level who will progress to opioid use disorder, although presence of a co-occurring psychiatric disorder and IV use of heroin are major risk factors. Patients who meet opioid use disorder at some point in their lives can achieve significant periods of abstinence, often without formal treatment. Most commonly, affected individuals will experience a chronic and recurring course of the illness.

Patterns of Use and Relapses and a Long-Term Prognosis

Opioid use disorder has a typical course, although in a particular individual the progress and remission through this course are highly variable. Opioid use typically begins with single doses used intermittently (e.g., weekend use) with longer periods of abstinence, either self-imposed or involuntary in the form of imprisonment or hospitalization, lasting from a few weeks to years. Gradually individuals transition to daily use and increase the dose frequency; multiple uses of short-acting opioids throughout the day are common, often with a first dose in the morning preceding all other daily activities in order to eliminate or avoid early withdrawal symptoms. Periods of self-detoxification and voluntarily imposed abstinence are typical but often followed by a relapse, especially under stress. Depression and life crises (especially arguments and interpersonal losses) are commonly associated with relapse. Some persons can maintain a lower-level of use early on following the relapse but more commonly the frequency of use and associated problems rapidly escalate, usually at a faster pace than during early periods of the disorder. The most common reasons for seeking abstinence are accumulating occupational, family, financial, and legal problems, feelings of fatigue and exhaustion with activities needed to maintain daily use, and inevitable periods of withdrawal. For others, emergence of severe depression or a legal mandate is the primary reason they want to stop using and seek out treatment.

Treatment allows individuals to initiate and maintain long-term abstinence. One of the major goals of treatment is to favorably modify the course of the disorder by minimizing the impact of risk factors on relapse (relapse prevention). In most cases, more than one treatment episode is necessary to achieve a long-term remission; however, among individuals who have a good treatment response, multiple treatment attempts can have a cumulative effect, and the duration of each drug-using episode tends to shorten. Many patients who have had positive experience in the treatment will be able to recognize high-risk situations and protect themselves from relapse or will be able to return to the treatment after the first few episodes of drug use (lapse) and minimize the risk that lapse would progress to a

full-blown relapse. Repeated relapse, although common, is not an inevitable course of opioid use disorder. Cohort studies show that 10 to 40 percent of interviewed opioid users are abstinent at long-term follow-up. Treatment research conducted throughout the past 4 decades shows that a proportion of opioid users enter into long-term remission. Sustaining ongoing treatment and monitoring, longer lengths of stay in treatment, and self-motivation have been identified as the most important reasons for long-term recovery. Other protective factors include religion, spirituality, employment, and family. Positive outcomes, such as greater involvement in prosocial activities (e.g., employment) and improved mental health, are both causes and effects of long-term recovery. Many patients who develop an opioid use disorder do not enter treatment and tend to have a less severe form and course of the disorder and less coexisting medical and psychiatric problems as compared to those seen in treatment programs.

For a large proportion of patients, opioid use disorder is a lifelong condition with severe health and social consequences. A longitudinal study of 581 patients with moderate-to-severe opioid use disorder who were placed under the California Civil Commitment Program in the early 1960s showed that during the 33-year follow-up period (Hser et al., 2001). Most of the deaths involved drug-related causes or violence. Among the survivors, many had serious medical and psychiatric problems and were involved in the criminal justice system. At the time of the follow-up interview 20 percent of were using heroin, 10 percent refused testing, 4 percent were incarcerated, and 56 percent were currently abstinent, but 40 percent had used heroin within the past year. More than 60 percent of subjects reported long periods of abstinence (>1 year). A review of long-term follow-up studies found abstinence rates between 10 percent and 20 percent for both drug-free and methadone maintenance treatment. In more recent studies, such as the ALIVE Cohort, conducted over the 12-year study period, four patterns were noted in patients with a pattern of injection drug use: 29 percent of participants remained persistent drug injectors, 20 percent ceased injection, 14 percent relapsed once, and 37 percent had multiple transitions.

Predictors of treatment outcome—legitimate work, low crime, less drug use, family relationships, and psychological adjustment—are favorable indicators for good long-term prognosis. Severity of comorbid psychiatric conditions at the beginning of treatment is usually a bad prognosis. No important differences in long-term outcome have been found for different treatment modalities. Over shorter follow-up periods, methadone maintenance, therapeutic communities, and drug-free programs were significantly superior to detoxification only programs.

Opioids and Crime

The relationship between drug use and crime is complex. Heroin use

correlates to increased levels of criminal activity, especially income-generating crime. Observational studies have shown that a reduction in crime is seen in heroin users in opioid maintenance treatment. Retention in treatment is consistently shown to be an important factor in mediating this reduction. Patients in continuous treatment had the greatest reductions in criminal activity, while those who drop out after a several treatment episodes have the highest rates of criminal convictions. Longer periods of methadone maintenance treatment in the community for individuals with a history of incarceration are associated with improved long-term outcomes. However, there is a smaller association between crime and illicit use of other opioids, and other causal factors of crime, such as poverty, lack of social services, and discrimination, can also contribute to illicit opioid use.

Morbidity and Mortality

Chronic use of pure opioids, including methadone, in nontoxic doses has relatively few known side effects as users develop tolerance to most of the acute effects. The cognitive impairment seen with long-term alcohol and sedative use is not generally found with long-term use of opioids. The only organ toxicity appears to be endocrinopathy with lower levels of testosterone and other sex hormones and resulting physiological changes (loss of libido, impotence, amenorrhea). Occasionally patients report chronic constipation and sleep disturbances.

Despite the relative safety of chronic opioids, the life expectancy of chronic opioid users, especially heroin users, is markedly reduced. Death in many opioid users occurs at a young age with an average of 44 years of potential life lost for each fatality. The rate of mortality in opioid users is 6.5 times the rate expected in the general population, with a higher rate for younger persons. In the United States, the overall standardized morality rate is about 1.0 to 1.5 percent per person per year, 0.8 to 1.2 percent in a cohort of people treated for opioid dependence in Australia, and 2 to 3 percent in British clinics providing clean injectable drugs to younger patients. Overdose is the most frequent cause of death followed by unintentional injures (accidents and violence) and suicide. Death related to liver disease, particularly viral hepatitis, is very common, while HIV-related mortality has been steadily decreasing thanks to HIV prevention efforts and harm reduction programs such as needle exchange.

Overdose. Even though overdose is the most common cause of death in individuals with opioid use disorder, most overdoses are nonlethal. Only 1 out of 25 to 50 of overdoses results in death; however, harm associated with nonlethal overdoses can be significant because even brief periods of complete anoxia can irreversibly damage CNS. Most cases of overdose occur in older, long-term users, many of them drug-injectors, and many of those individuals have one or more nonlethal overdoses before a lethal one.

It is not well understood why some individuals are at increased risk of overdose. A majority of overdoses involve multiple CNS depressants, most commonly alcohol and benzodiazepines but also other opioids such as methadone.

Medical Comorbidity. Most of the medical complications seen in chronic users of opioids, particularly users of heroin, are due not to heroin but to medical complications related to exposure to other substances used concurrently (e.g., impurities and adulterants in heroin and acetaminophen in prescription painkillers), harmful drug administration practices (e.g., nonsterile injecting), and unhealthy behaviors associated with lifestyle choices (e.g., poor nutrition and unprotected sex). Infections of skin and subcutaneous tissue (abscesses, cellulitis), thrombophlebitis, endocarditis, pneumonia, and hepatitis (B and C) are frequently seen medical complications. Acute infections are the major reason for inpatient admission of people who inject drugs.

Viral hepatitis, particularly hepatitis C virus (HCV), is the leading infectious cause of death in the United States and the leading cause of liver transplantation and liver cancer. In the United States, up to 80 percent of HCV cases are seen among persons who inject drugs, as are half of all new cases of infection. Abnormal liver function is seen in two-thirds of heroin users, and it may persist after the cessation of injections. High-risk sexual and drug-use behaviors are the most common mode of transmission of hepatitis and HIV in heroin users. Despite education efforts and other prevention measures, the frequency of these behaviors remains high in the population of IV drug users. All patients who are in treatment should be screened for hepatitis B and C and offered vaccination for hepatitis B.

Acquired immune deficiency syndrome (AIDS) remains an important factor in mortality among IV opioid using patients. HIV seroprevalence has a wide geographic variability and is lower among those who have continuously been in drug treatment. The incidence of tuberculosis is also higher among heroin users than in the general population and has been linked to the compromised immune system of these individuals regardless of whether or not they have HIV. IV drug users should have access to measures that reduce spread of hepatitis and HIV, such as clean injecting equipment, condoms, and access to medical treatment.

Other infections related to IV substance use and sharing of needles include malaria, tetanus, osteomyelitis, and syphilis. Endocarditis related to staphylococcus but also pseudomonas or candida may result in serious complications. Endocarditis can cause abscesses in the brain or spinal cord, meningitis, diffuse vasculitis, and subarachnoid hemorrhage from rupture of infected aneurysm. Various contaminants (e.g., talc) entering the bloodstream can cause pulmonary emboli, which can eventually result in angiothrombotic pulmonary hypertension and right ventricular failure.

Staphylococcal pneumonitis may also be related to septic emboli. Poor IV injecting technique can cause chronic edema of extremities. Subcutaneous or intracutaneous injection (known as skin popping) may cause widespread ulceration and disfigurement.

Patients surviving overdose may experience long-term sequelae of anoxic brain injury, such as transverse myelitis, amblyopia, plexitis, peripheral neuropathy, parkinsonian syndrome, movement disorder, intellectual impairment, and personality changes. Pathological changes in muscles and degeneration of peripheral nerves have also been seen. Synthetic opioids can affect cardiac conduction by lengthening the QT interval. Large increases in QT intervals can cause serious arrhythmias, such as torsade des points. LAAM (levo-α-acetylmethadol) and methadone appear to have the lowest margin of safety. Prolonged QT interval (≥500 msec) is seen in approximately 2 percent of patients treated with methadone and clinical use of LAAM has been discontinued, in part, because of ventricular rhythm problems.

Chronic pain is commonly seen within the opioid-dependent population. Pain is often cited as the reason for initiating the use of prescription opioids and heroin, and chronic, and usually post-traumatic pain is often reported by patients who use opioids.

Psychiatric Comorbidity. Comorbid psychiatric conditions are the rule rather than the exception in opioid use disorder. The type and severity of additional psychiatric diagnoses can influence the course of the disorder and treatment. The most recent US-based NESARC found that 69 percent of men and 74 percent women with opioid use disorder in their lifetimes met criteria for an additional mood, anxiety, or another SUD at some point in their lives. Some of the most common diagnoses included mood disorders, antisocial personality disorder, anxiety disorders, and AUD (Table 11.9–3). Findings from the Australian Treatment Outcome Study (ATOS) also show a high prevalence of mental disorders among opioid users, among those both who are in treatment and those who are not. In the ATOS the most prevalent comorbid psychiatric conditions were major depression, posttraumatic stress disorder (PTSD), and antisocial personality.

Among the NESARC cohort, women were twice as likely as men to have a lifetime mood or anxiety disorder. Women were twice as likely as men to have paranoid personality disorder, but half as likely as men to have antisocial personality disorder. Women were less likely than men to meet criteria for abuse on cocaine, marijuana and less likely than men to be dependent on alcohol. Women in the ATOS cohort also were more likely to have depression or PTSD.

The long-term outcome for opioid dependence among persons with additional major psychiatric disorders, such as schizophrenia or bipolar I

disorder, is not generally known. However, a study that specifically examined the efficacy of additional psychiatric care for dual diagnosis subjects in an outpatient drug-dependence clinic found that, although they initially had significantly greater percentages of positive urine toxicologies than patients without comorbid psychiatric conditions, they had higher retention rates in treatment and, after 6 months of treatment had comparable success rates to the single-diagnosis group.

TREATMENT

Clinical Pharmacology of Medications Used in the Treatment of Opioid Use Disorders

The primary pharmacological treatment strategy used in the management of opioid use disorder aims at engaging the μ-opioid receptor. The three classes of medications used in treatment are full agonists, partial agonists, and antagonists, each exerting a different pharmacological effect at the receptor (Fig. 11.9–4, Table 11.9–4). Opioid full agonists, which include methadone, heroin, and oxycodone, bind to the receptor and activate it to the maximum capacity, with the pharmacological and clinical effect that is directly correlated with the dose. Methadone is used for the purpose of detoxification and as a maintenance agent.

Table 11.9–3.
Substance and Non–substance-Related Psychiatric Disorders in Opioid Users

Mood, anxiety, and personality disorders among individuals with lifetime opioid use disorders in NESARC by gender

(*N* = 355)	Men (%) (*N* = 223)	Women (%) (*N* = 578)	Total (%)	Odds Ratio (95% Confidence Interval)
Heroin/other opioids				
Abuse	75.6	67.5	73	0.67 (0.43, 1.04)
Dependence	24.4	32.5	27	1.49 (0.96, 2.31)
Cocaine				
Abuse	37.6	17	30.9	0.34 (0.20, 0.58)***
Dependence	17.4	20.5	18.4	1.22 (0.73, 2.03)
Marijuana				
Abuse	56.2	39.3	50.7	0.51 (0.33, 0.78)**
Dependence	19.4	19.9	19.6	1.03 (0.62, 1.72)
Alcohol				
Abuse	22.6	20.4	21.8	0.88 (0.55, 1.41)
Dependence	67.1	54.5	63	0.59 (0.40, 0.86)**
Sedatives				
Abuse	37.7	23.3	33	0.50 (0.30, 0.85)*

Dependence	9.1	12.9	10.3	1.48 (0.77, 2.87)
Amphetamine				
Abuse	35.9	28	33.3	0.70 (0.45, 1.08)
Dependence	13.3	18.5	15	1.47 (0.80, 2.72)
Hallucinogen				
Abuse	37.4	27.5	34.2	0.63 (0.40, 1.01)
Dependence	9	9.5	9.1	1.06 (0.51, 2.20)
Tranquilizers				
Abuse	37.7	22.7	32.9	0.49 (0.30, 0.78)**
Dependence	8.7	12.4	9.9	1.49 (0.80, 2.81)
Inhalants				
Abuse	13	6.7	10.9	0.48 (0.24, 0.95)*
Dependence	1.7	2.1	1.9	1.22 (0.30, 4.99)
Any other substance				
Abuse	23.1	21.6	22.6	0.92 (0.60, 1.41)
Dependence	72.4	62.3	69.1	0.63 (0.43, 0.91)*
Any mood or anxiety disorder	68.5	73.8	70.3	1.30 (0.84, 2.00)
Any mood disorder	54.1	71.2	59.7	2.10 (1.35, 3.26)**
Major depression	45.7	63.5	51.5	2.07 (1.38, 3.12)**
Dysthymia	15.3	25.3	18.6	1.88 (1.20, 2.93)**
Manic disorder	16.7	27.1	20.1	1.86 (1.17, 2.94)*
Hypomanic disorder	9	10.9	9.6	1.24 (0.61, 2.50)
Any anxiety disorder	33.6	50.1	39	1.98 (1.35, 2.92)***
Panic disorder w/o agoraphobia	13.2	18.3	14.9	1.46 (0.88, 2.45)
Panic disorder w/agoraphobia	4.1	6.6	4.9	1.65 (0.63, 4.33)
Social phobia	9.3	19.8	12.7	2.42 (1.36, 4.30)**
Specific phobia	12.9	30.1	18.5	2.91 (1.69, 5.02)***
Generalized anxiety disorder	9.7	17.7	12.3	2.01 (1.13, 3.56)*
Any personality disorder	49.2	50.7	49.7	1.06 (0.74, 1.52)
Paranoid	12.9	22.2	15.9	1.93 (1.17, 3.21)*
Avoidant	9.5	12.4	10.4	1.35 (0.71, 2.57)
Dependent	3	2.3	2.8	0.77 (0.25, 2.37)
Obsessive compulsive	16	23	18.3	1.56 (0.95, 2.57)
Schizoid	11.5	12.8	11.9	1.13 (0.61, 2.11)
Histrionic	9.8	8.5	9.4	0.86 (0.42, 1.76)
Antisocial	34.1	22.4	30.3	0.56 (0.36, 0.87)*

Asterisks indicate significance level for p-values. ** indicates P ≤ 0.05; ***indicates P ≤ 0.01; ****indicates P ≤ 0.001.

Adapted from Grella CE, Karno MP, Wards US, Niv N, Moore AA. Gender and comorbidity among individuals with opioid use disorders in the NESARC study. *Addict Behav*. 2009;34(6–7):498–504. (Open access)

The opioid partial agonist buprenorphine binds to the receptor but activates it only partially; at most the effect is approximately 50 percent of the maximum effect produced by the agonist. The agonist effect of

buprenorphine is directly correlated with the dose at the lower range of doses but as the dose increases the pharmacological and clinical effect increases very little and it reaches a ceiling level beyond which no further increase in agonist effect is possible. Functionally, buprenorphine acts primarily as an agonist (receptor activation). However, buprenorphine has a high affinity (binding strength) to the receptor and a slow dissociation from the receptor, therefore it can function as an antagonist (blocker) preventing agonists with lower receptor affinity from activating the receptor. Buprenorphine can also act as a functional antagonist if given to individuals who are physically dependent with agonist occupying receptor sites. In such a situation buprenorphine will abruptly displace the agonist from the receptor, but since it will have only partial effect at the receptor, it will cause a sudden drop in the level of agonist effect resulting in a precipitated withdrawal. Buprenorphine is used to assist detoxification and as a maintenance agent.

Opioid antagonists (e.g., naloxone, nalmefene, and naltrexone) bind to the receptor but do not exert any activity. These medications function as antagonists—they compete for binding at the receptor with agonists and partial agonist, and if the dose of the antagonist is high enough, they may displace the agonist from the receptor and prevent the agonist from binding to the receptor and exerting any effect. In an individual physically dependent on opioids, antagonists given in sufficient doses will displace remaining agonist from the receptor and will precipitate an immediate opioid withdrawal. Naloxone is used in the treatment of overdose; because it is a short-acting agent it can be dosed very accurately without the risk of an unnecessarily prolonged effect. Naltrexone is a long-acting agent used following detoxification, as a strategy to prevent redependence and relapse.

Table 11.9–4.
Pharmacological Properties of Different Opioids

Opioid	Equivalent Dose to 10 mg IM Morphine	Half-Life (hours)	Peak Effect (hours)	Duration (hours)	Detection Period in Urine
Morphine	10 mg IM	2–3	0.5–1	3–4	48–72 hours
	20–30 mg PO	2–3	1–2	3–6	
Controlled release morphine	20–30 mg PO	2–3	NA	8–12	
Hydromorphone	1.5 mg/IV/SQ	2–3	0.5–1	3–4	2–4 days
	7.5 PO	2–3	1–2	3–6	
Oxycodone	20–30 PO	2–3	1–2	3–6	2–4 days
Hydrocodone	30 PO	2–4	1–2	3–6	2–4 days
Fentanyl	50–100 ug IV/SQ	7–12	<10 minutes	1–2	48–72 hours
Codeine	150–200 mg PO	2–3	0.5–1	4–6	48 hours

Methadone	VARIABLE See below	12–150	1–2	6–8	7 days
Heroin	2 mg IV	10 minutes	<1 minute	<30 minutes	48–72 hours as morphine
Buprenorphine	0.4 mg PO	20–70 hours	2–4 hours	Up to 72	72 hours

Oral Morphine-Equivalent Daily Dose (mg/day)	Methadone Dose Conversion Ratio
0–99	4:1
100–299	8:1
300–499	12:1
500–999	16:1
>1,000	20:1

Steps for conversion of opioid to methadone:

1. Convert to ORAL morphine equivalent (24-hour total dose)
2. Divide by ratio above
3. Divide by 50% to account for incomplete cross-tolerance
4. Divide by 3 for frequency (every 8-hour dosing)
5. Round down to the nearest tablet size—2.5-mg intervals

The last group of medication used in the treatment of opioid use disorder includes medications that target various signs and symptoms emerging during acute and protracted opioid withdrawal (e.g., clonidine, lofexidine, and prochlorperazine).

Treatment of Opioid Overdose

Opioid overdose or mixed opioid-sedatives overdose generally presents with unconsciousness and respiratory depression and constitutes a medical emergency. As most overdoses that come to medical attention are nonlethal, an immediate treatment is essential to minimize the impact of anoxia and prevent irreversible injury to the CNS and other organs. Immediate interventions include protection of airways and support of breathing in case of very low respiratory rate and use of supplemental oxygen attached to a bag-valve mask. Naloxone, an opioid antagonist, should be administered parenterally as soon as possible to reverse opioid overdose, preferably via the IV route.

Naloxone is given at the dose of 0.2 to 0.4 mg to a patient in coma, to aid in the maintenance of respiration. Higher initial doses (1 to 2 mg) should be given to patients with apnea or in cardiorespiratory arrest. Naloxone has a relatively short half-life (60 to 90 minutes) and must be repeatedly administered or given as an IV infusion in patients who have overdosed on an opioid with a longer half-life (e.g., extended-release oxycodone), otherwise the patient can fall back into coma after the effect of naloxone wears off. Naloxone can also be administered subcutaneously or intramuscularly if it is difficult to establish IV access. Signs of improvement —increase in respiratory rate and pupillary dilation—should occur within 2 to 3 minutes following IV administration. If no response is seen during this

time, a higher (1 to 2 mg) repeated doses should be given at 2 to 3 minute intervals while ventilation and oxygenation is maintained. Higher doses of naloxone may be required for buprenorphine, methadone, of fentanyl overdose.

The primary goal of naloxone administration is to restore adequate ventilation (above 12 respirations per minute) and oxygenation (above 92 percent) rather than full alertness. Administering higher doses of naloxone can precipitate opioid withdrawal and may cause difficulties with controlling agitation. An acute lung injury (with hypoxia, rales, and mouth frothing) can be seen in rapid overdose reversals in patients with marked hypoventilation. If there is no improvement after 10 mg of naloxone is given, alternative diagnoses need to be considered. Most overdoses include other sedatives such as alcohol and benzodiazepines and therefore strategies aimed at managing toxicity from other substances should also be considered if the response to naloxone in insufficient.

Following the reversal of an overdose, a strong attempt should be made to enroll such individuals in the treatment of their opioid use disorder using one of the pharmacological strategies; otherwise the risk or repeated overdose is very high. The most important strategy to reduce rates of overdose is the treatment of opioid use disorder. Most individuals who overdose on heroin are not in treatment at the time of overdose. Those who enter treatment significantly reduce use of opioids and other CNS depressants and have a significantly lower risk of overdose.

Treatment of Opioid Withdrawal (Detoxification)

Opioid withdrawal is usually treated in the setting of detoxification, which continues to be a major medical treatment modality used in working with opioid-dependent individuals. The primary goals of detoxification are (1) to provide a safe removal of the opioid and the physiological dependence from the organism and prevent life-threatening complications that might happen if opioid withdrawal is left untreated and (2) to stabilize the patient's physical and psychological health to allow them to initiate medication-assisted relapse-prevention treatment. Detoxification is necessary before starting relapse-prevention treatment with the opioid antagonist naltrexone but it is not needed before starting maintenance treatment with the opioid agonists methadone or buprenorphine. It is important to note that detoxification by itself is never a sufficient treatment modality. Detoxification that is not followed by a relapse-prevention treatment has a very high rate of early relapse, as much as 90 percent during the first few months following treatment, and the risk of opioid overdose is much greater in individuals who were detoxified as the detoxification also removes the physiological tolerance to opioids, which protects against overdose.

Improved detoxification methods have evolved over the past 150 years,

reflecting a more humanitarian view of addicted individuals and a better understanding of addiction as a medical disorder. Medication-assisted detoxification usually takes place in the medical facility, at an inpatient or an outpatient treatment setting. The inpatient setting is reserved for patients with a history of severe or complicated withdrawal, those expected to have severe withdrawal based on the baseline pattern of substance use (e.g., patients with mixed opioid and sedative-hypnotic withdrawal), and those presenting with coexisting medical and psychiatric problems. For others inpatient detoxification may not be a realistic option because of occupational or domestic responsibilities and lack of adequate medical insurance to pay for intensive hospital treatment, whereas outpatient detoxification and relapse prevention treatment initiation allows more individuals to benefit from treatment.

Three main detoxification strategies are currently in the clinical use: (1) symptomatic treatment only with nonopioid medication, (2) agonist-assisted detoxification, and (3) antagonist-assisted detoxification with symptomatic treatment. Each of these regimens is expected to produce a different dynamic of the onset, severity, and duration of withdrawal symptoms. The main difference is that in agonist-assisted detoxification withdrawal symptoms tend to be more severe toward the later part of the treatment and may continue after treatment completion while in symptomatic treatment only and antagonist-assisted detoxification, withdrawal symptoms are more severe early in the treatment and are usually minimal later in the treatment.

Symptomatic Treatment of Opioid Withdrawal. Opioid withdrawal is characterized by a wide range of signs and symptoms affecting multiple organ systems. The severity of withdrawal can be significantly alleviated with medications directed at each component of withdrawal. This so called "symptomatic treatment" or treatment with ancillary medications is frequently utilized, particularly in situations where opioids are not used in detoxification because of legal restrictions. Ancillary medications are also a mainstay of treatment during antagonist-assisted detoxification and can also be used in the course of agonist-assisted detoxification to treat residual withdrawal symptoms after the agonist is stopped. The following classes of medications are used to target specific syndromes emerging during opioid withdrawal:

1. Autonomic (sympathetic) arousal can be decreased using α-2 adrenergic agonists. Clonidine is the most widely used symptomatic-treatment medication, effective in relieving a wide range of autonomic withdrawal symptoms (sweating, restlessness, tremor, and rhinorrhea). Sedation is a particularly welcomed effect of clonidine; however, other aspects of withdrawal (such as muscle/bone pain and insomnia) are not helped by clonidine. Clonidine is given orally, starting at doses of 0.1 to 0.3

mg three to four times a day. In outpatient settings, doses tend to be lower due to the potential for a postural hypotension that is more difficult to monitor. Hypotensive risks of clonidine are most pertinent in the situation of fluid loss (vomiting/diarrhea). Aggressive oral hydration with electrolyte fluids is helpful to increase tolerability of clonidine. Lofexidine is another adrenergic agent that has effectiveness profile similar to that of clonidine; however, the risk of hypotension appears to be lower with lofexidine.

2. Anxiety and psychomotor agitation is a particularly troublesome aspect of opioid withdrawal that is often the reason why some patients prematurely withdraw from the treatment. Clonidine may only partially help with anxiety and therefore anxiolytic/sedating agents that work through a different mechanism may be very helpful. Benzodiazepines are the most commonly used anxiolytics because they have a wide margin of safety and are available in the form of several agents with varying duration of action, allowing great flexibility in use. Clonazepam (0.5 to 2.0 mg) or oxazepam (15 to 30 mg) can be given two to four times per day to prevent and alleviate anxiety/agitation. Benzodiazepines are especially suited for use in the inpatient setting as there is concern about the potential for benzodiazepines to be misused or unsafe (if taken in combination with an opioid) during outpatient detoxification. Other classes of sedative agents that can be used include antihistamines such as diphenhydramine or hydroxyzine.

3. Insomnia, which at times can be severe, is seen in most patients undergoing opioid withdrawal and targeting insomnia can greatly improve tolerability of detoxification though rarely can it be completely treated. Specific hypnotic agents used include zolpidem (5 to 10 mg) or zopiclone (7.5 to 15 mg), but sedating antidepressants (e.g., trazodone 50 to 150 mg) or atypical neuroleptics (e.g., quetiapine 50 to 200 mg) can also be used.

4. Other symptoms of withdrawal can also be treated symptomatically. Musculoskeletal pain can be alleviated using nonsteroidal anti-inflammatory drugs (ibuprofen or ketorolac) or antispasmodic agents (cyclobenzaprine or methocarbamol). GI distress responds well to antiemetic agents (prochlorperazine, ondansetron) and diarrhea can be treated with loperamide.

Agonist-Assisted Detoxification. Two agonist medications are approved in the United States for use in the treatment of opioid withdrawal: methadone and buprenorphine. Both methadone and buprenorphine are effective in treating signs and symptoms of opioid withdrawal and have a long elimination half-life that makes them good candidates to be used for the purpose of detoxification. Agonist-assisted detoxification involves the substitution of the short-acting agonist at the receptor site (usually heroin or oxycodone) with a long-acting agonist, followed by a gradual and controlled reduction in the dose of the agonist to minimize the severity of

associated withdrawal. The severity of associated withdrawal is directly proportional to the rapidity of dose reduction; however, it tends to be relatively greater in the final stages of agonist discontinuation and will persist after the last dose is given.

Methadone is a full opioid agonist and can be given immediately after the cessation of opioids to preclude withdrawal. However, buprenorphine is a partial opioid agonist and it will precipitate withdrawal if given to a physically dependent patient who does not show signs of withdrawal; thus, a wait period to allow for the spontaneous withdrawal to emerge is necessary before the first buprenorphine dose can be given. Therefore it may be more difficult for patients to initiate detoxification with buprenorphine as compared to methadone. However, the severity of terminal withdrawal after administration of the final dose is commonly lower in case of the buprenorphine-assisted detoxification than in the case of methadone-assisted detoxification.

METHADONE. For patients who have been using heroin or illicit prescription opioids the initial dosage of methadone is usually 20 to 30 mg orally. If objective signs of withdrawal persist after the first dose of methadone, the dose can be repeated after approximately 2 to 4 hours. As a general rule, initial stabilization does not require more than 40 mg of methadone during the first 24 hours but it can be higher in the inpatient setting if clear signs of moderate or severe withdrawal are present after the initial 6 to 10 hours of observation period. For example individuals who have easy access to large doses of high-potency opioids, such as healthcare providers, may require higher initial doses. After 24 to 48 hours of dose stabilization, the patient's methadone dose can be gradually reduced by 10 to 20 percent or by 5 mg per day. Inpatient detoxification with methadone is usually completed within 1 week.

In case of detoxification in an outpatient setting, the first-day dose should not exceed 40 mg. Over the next few days, the dose may be raised (generally by no more than 10-mg increments every 2 to 3 days) until patient has no discernible withdrawal symptoms. In most cases, the dose of 40 to 60 mg per day is sufficient to prevent any opioid withdrawal; however, this dose may not be adequate to eliminate craving. In case of continuing opioid use, the dose of methadone needs to be increased further and the attempt at detoxification needs to be postponed and the goal changed to stabilization—no withdrawal symptoms, craving, or opioid use. Generally, a period of dose stabilization, at minimum 4 weeks in duration, is necessary prior to a slow dose reduction schedule. Studies of outpatient methadone withdrawal show that gradual dose reductions (e.g., 3 percent per week) have a greater likelihood of success versus more rapid dose reductions (e.g., 10 percent per week). Dose reductions are generally well tolerated, but as the dose is reduced to below 20 to 30 mg per day, there is

increasing withdrawal distress with each dose change and increased risk of illicit opioid use. Another factor that interferes with the final discontinuation is the negative expectancy effect, where patient expects the increase in withdrawal as the medication is stopped and experiences a significant distress. Conducting a single-blind detoxification, where patient agrees to be blinded to the actual dose and rate of taper, can have a significant positive effect on the outcome. Outpatient detoxification using an opioid agonist can usually be accomplished in 1 to 2 months, but slower reductions can be more effective in patients who have difficulty tolerating the process.

BUPRENORPHINE. While methadone has been safely and effectively used for detoxification for almost 50 years, buprenorphine has become an alternative option in the past decade, following FDA approval of sublingual buprenorphine products indicated for the treatment of opioid withdrawal. The sublingual buprenorphine is marketed in two formulations: a mono therapy product containing buprenorphine alone and a combination therapy product that contains buprenorphine and naloxone in a buprenorphine to naloxone dose ratio of four to one. A combination therapy product is a preferred formulation, as it carries a lower risk of diversion and misuse, to be used for induction and maintenance. A mono product is generally used in pregnant and nursing mothers and in those patients that cannot tolerate the naloxone side effects, which is a rare occurrence.

Detoxification using buprenorphine can be safely accomplished in the inpatient setting under direct monitoring. The first dose of buprenorphine should be given when the patient experiences at least moderate severity of withdrawal, with several, preferably objective signs present. The first dose of buprenorphine is typically 2 mg, and subsequent 2 mg doses are given every 2 to 4 hours until the withdrawal symptoms are eliminated. Usually, the total daily dose needed to effectively suppress withdrawal on the first day is 8 to 12 mg. Buprenorphine taper is usually started on the second day, while the daily dose is decreased by 2 mg every day, with 1 mg decreases on the last 2 days of taper. The complete taper can be usually accomplished in 5 to 7 days and the long elimination half-life of buprenorphine (31 to 35 hours) is very useful in extending the medication effect after it is discontinued.

Buprenorphine-assisted detoxification can also be accomplished in the outpatient setting and the strategy is generally similar as with the methadone-assisted detoxification. Ideally, the first day's dosing should occur under medical supervision to minimize the risk of buprenorphine-precipitated withdrawal. Patients are instructed to arrive in withdrawal, which typically starts 12 hours after the last dose of a short-acting opioid or 24 hours after a long-acting opioid like methadone. The first dose of

buprenorphine is usually 2 mg, and if no precipitated withdrawal is observed, subsequent 2 mg doses are given at the clinic every 2 hours, with additional doses given to take at home, until the withdrawal symptoms are eliminated. If withdrawal worsens after the first buprenorphine dose, it is best to wait for 4 to 6 hours before giving the next buprenorphine dose while the precipitated withdrawal can be treated symptomatically with medications such as clonidine. Usually the total daily dose needed on the first day is 8 to 12 mg. Subsequent daily dose increases of 2 to 4 mg are recommended if needed to relieve opioid withdrawal and initiate opioid abstinence. Once a period of stabilization is established, slow dose reductions can begin, usually in 2 mg dose increments. Because of the prolonged elimination half-life of buprenorphine, dose reductions are best introduced at weekly intervals, as the patient may not experience the effect of dose reduction until several days later. As is the case with methadone, slower dosage reduction schedules are usually better tolerated and less likely to prompt relapse.

Antagonist-Assisted Detoxification. Although agonists-assisted and symptomatic treatment can effectively minimize distress of detoxification, these strategies do not shorten and may extend the duration of detoxification to 7 to 10 days and many individuals drop out of treatment before its completion. Several methods have been developed to shorten the duration of detoxification and to use it as a starting point for the antagonist-based treatment to prevent post-detoxification relapse (naltrexone induction). To accomplish it, opioid antagonist, such as naltrexone, is introduced during detoxification and the emerging withdrawal symptoms are treated symptomatically (e.g., clonidine and clonazepam). To minimize severity of precipitated withdrawal naltrexone is introduced 2 to 3 days after the last dose of a full agonist opioid in low dosage (e.g., 3 to 6 mg) and quickly titrated upwards over the ensuing days up to the full blocking dose of naltrexone (50 mg oral tablet or an injection of extended-release preparation). As a variant to this procedure, 1 day of buprenorphine (4 to 8 mg) can be used as a step-down from a full agonist before the initiation of naltrexone.

Precipitated withdrawal can be quite severe if naltrexone is given early and in a high initial dose and it could be dangerous in medically compromised patients (e.g., severe fluid loss, hypotension, and delirium), therefore it requires close monitoring on an inpatient basis and is best used in patients with low level of dependence and without medical problems that may worsen during withdrawal (e.g., asthma, diabetes, and heart disease). Detoxification under heavy sedation or a general anesthesia, so-called ultra-rapid procedures, has been advocated in an effort to initiate naltrexone immediately after cessation of opioids and to shorten duration of detoxification to 1 to 3 days. While these methods may increase the

likelihood of successful induction onto naltrexone comparing to more gradual detoxification, these methods require a highly experienced anesthesiology team, are expensive, and can produce life-threatening adverse events including aspiration pneumonia, pulmonary edema, renal failure, and ketoacidosis, complications that are never seen in the course of traditional detoxification methods. Therefore, ultra-rapid methods should not be suggested as an alternative to other methods of detoxification; however, in highly selected cases there might be a benefit to using these methods to initiate relapse-prevention treatment with long-acting opioid antagonists.

Relapse after Detoxification. An extensive amount of evidence shows that detoxification—treatment of opioid withdrawal after the cessation of opioid use—is by itself not adequate to treat opioid use disorder. In the absence of relapse-prevention treatment, most patients will relapse within the first 6 months and such relapse can very dangerous, as detoxified individuals do not have physiological tolerance to protect them from acute opioid toxicity. Therefore, detoxification can only be seen as a first step of engagement in treatment while the primary treatment goal is the maintenance of abstinence.

Relapse-prevention treatment strategies that can be implemented after the cessation of opioid use generally include a combination of pharmacological treatment, with antagonist or agonist agents, and psychosocial interventions. The outcomes of treatment that include only psychosocial approaches, even if delivered in the high-quality and optimized manner, are inferior to treatments that include pharmacological support. The exception might be a long-term residential treatment (e.g., therapeutic community) where individuals have no ready access to opioids while undergoing intensive treatment and rehabilitation and are monitored for protracted withdrawal symptoms (e.g., insomnia, low energy, anxiety, and drug craving), which emerge during early abstinence. Nevertheless, the strongest evidence points to the benefits of added pharmacological support to minimize the risk of relapse during early stages of abstinence from opioids.

Treatment of Opioid Use Disorder

The ultimate goals of opioid use disorder treatment vary not only for patients but also for providers, as there are several perspectives as to what constitutes the primary goal of the treatment of opioid use disorder. The goals range from a sustained, life-long abstinence from all psychoactive substances to the minimization of harm to the individual that continues using opioids. For the medically oriented providers, goals of treatment usually include the reduction in symptoms of the diagnosis, more specifically (1) cessation of illicit opioid use, (2) protection against the risk

of overdose and death, and (3) improvement in physical and psychological health. Often, complete abstinence early in the treatment is an unattainable goal and more realistic immediate goals are treatment engagement, reduction in use amount or frequency, and the elimination of particularly risky drug-use practices. For other treatment providers, the primary goal of treatment is the achievement of recovery, often defined as a living in abstinence from all psychoactive substances, including psychoactive medications, supported by changes in many aspects of life (social, recreational, vocational) and participation in a mutual help and support group (e.g., NA). Unfortunately, sometimes the goal of recovery is seen as incompatible with the use of relapse prevention medications, particularly opioid agonists, in support of abstinence. However, these two approaches do not have to be exclusionary and can be complementary on a time--continuum, with more immediate medically focused goals as a first step gradually transitioning to more recovery-oriented treatment that can be accomplished with or without the continuing support of pharmacotherapy.

Psychosocial Treatments for Opioid Dependence. A broad range of psychosocial interventions has been used and evaluated as treatment strategies for opioid dependence. Psychological interventions commonly include unstructured, supportive psychotherapy, and drug counseling but may also include various psychological techniques such as CBT or contingency management (CM). Social interventions include assistance with daily needs (food, clothing, and housing), employment counseling, assistance with access to healthcare, and engagement in social support network.

CBT is a particularly valuable approach to the treatment of opioid dependence with a strong evidence for its effectiveness. CBT is a structured and goal-oriented approach that is based on the assumption that drug-taking is a learned behavior. During treatment, patients are introduced to new coping skills and cognitive strategies to replace the dysfunctional behaviors and cognitions and through these changes support abstinence. Clinical trials evaluating psychosocial interventions have shown that regular counseling reduces illicit drug use and the quality of counseling and rapport between patient and counselor have a significant impact on outcome measures. In patients with higher levels of psychiatric symptoms or depression, CBT or supportive psychotherapy produced superior results as compared to drug counseling alone. CBT can also be effective in reducing HIV risk behaviors associated with IV drug use. A CM approach is often used as a part of treatment that focuses on promoting new behaviors that are competing to drug use. CM involves giving patients tangible rewards to reinforce positive behaviors such as abstinence, treatment attendance, and compliance with medication. Monitored urine collection and testing with immediate consequences are necessary for CM

effectiveness. Patients do not receive cash payments for abstinence but rather they receive vouchers to purchase goods and services consistent with their treatment goals. Toxicology also provides an objective index of treatment progress and test results can be discussed in therapy sessions in order to promote better understanding of situations that lead to continued drug use. Both CBT and CM can work in synergy with medication treatment as these can be focused on medication adherence.

NA is a self-help group of abstinent drug abusers modeled on the 12-step principles of AA. This approach emphasizes two major strategies to maintain abstinence: spirituality (faith and hope for recovery) and pragmatism (belief in doing "whatever works" to avoid using drugs again). Such groups now exist in most large cities and can provide useful group support. Although the scientific evidence supportive of the usefulness of such groups has been very limited, observational studies support its effectiveness and therefore patients receiving pharmacological treatment should be encouraged to participate in the self-help groups. Because 12-step groups emphasize abstinence from all psychoactive substances, patients treated with opioid agonists may find themselves being ostracized in traditional NA groups. An alternative approaches developed specifically for patients on agonist medications including Methadone Anonymous (MA), a variant of a 12-step program specifically for patients receiving methadone medication. Additional approaches, such as family and network therapy, can improve compliance with medication and can enhance outcomes in opioid addiction treatment.

Medications for the Treatment of Opioid Dependence. In the United States, four medications are currently approved by the FDA for the treatment of opioid dependence: methadone (used since the mid-1960s and approved in 1972), oral naltrexone (used since the 1970s and approved in 1984), buprenorphine (approved in 2002), and injectable, long-acting naltrexone (approved in 2010). LAAM is a long-acting agent similar to methadone that was in use since its approval in 1993, but it was discontinued in Europe and in the United States in early 2000s due to limited uptake (in the United States) and cardiac toxicity concerns (in Europe).

METHADONE. In the United States, there are approximately 250,000 patients in outpatient treatment programs receiving methadone. The original approach pioneered by Drs. Dole and Nyswander in 1964 emphasized using methadone dosage sufficient to suppress opioid drug craving and producing a cross-tolerance blockade of other opioids. Doses of methadone were in the 80 to 120 mg per day range and patients were encouraged to remain on methadone indefinitely. Subsequent treatment programs used lower doses of methadone (less than 50 mg per day), as the

methadone maintenance treatment was seen not as a goal in itself but rather a transitional stage before the eventual detoxification. Lower doses of methadone were often adequate for eliminating withdrawal but often these doses only partially suppressed craving and drug-seeking behavior and did not produce adequate cross-tolerance to large doses of heroin. Empirically, patients maintained on lower doses of methadone are far more likely to discontinue treatment, and those who do so are unlikely to maintain abstinence from illicit opioids.

Meta-analytical studies show methadone's clear superiority over treatments without medication on improving treatment retention and reducing heroin use. The majority of patients treated in methadone programs also show decreased nonopioid drug use, criminal behavior, symptoms of depression, and increased gainful employment. The effectiveness of opioid maintenance treatment is influenced by the dose and the quality of the additional services provided. The optimal dose is generally in the 80 to 120 mg per day range, with some patients requiring higher doses to achieve an optimal clinical response. Effective individual drug counseling can significantly augment treatment outcomes by decreasing illicit opioid and cocaine use, decreasing needle sharing and crime, and increasing psychological well-being.

Among patients who initiate methadone maintenance, there is an increase in mortality due to respiratory depression (due to ongoing illicit opioid use); however, after that time there is a reduction in mortality that remains as long as the patient is maintained on the medication. The risk of overdose is again increased after patients stop taking methadone. Methadone is relatively safe as a long-term maintenance agent. The main safety concern is the risk of excessive sedation with a rapid dose escalation and where other sedating agents, prescribed or illicit, are used in combination with methadone.

Among the most common side effects with methadone are constipation, excessive sweating, decreased libido, and sexual dysfunction. Opioids can reduce plasma levels of testosterone and follicle-stimulating hormone (FSH). Sleep abnormalities (insomnia and nightmares) and altered electroencephalogram (EEG) sleep patterns are frequently found during the first months of methadone treatment and return to baseline, but complaints of sleep abnormalities may persist.

Clinicians should be aware that a variety of medications could cause a change in plasma levels of methadone by inducing or inhibiting hepatic enzymes that metabolize methadone (CYP3A4). For example, adding rifampin, phenytoin, carbamazepine, darunavir, or efavirenz to the methadone regimen may cause a decrease in methadone levels with resulting opioid withdrawal, while adding fluconazole/ketoconazole, fluvoxamine, erythromycin, or delavirdine may produce an increase in methadone levels, causing excessive sedation. Methadone appears to

increase the QTc interval in some patients, an effect than can increase the risk of a cardiac conduction abnormality (torsade de pointes). This effect is dose related and is more commonly seen with very high doses of methadone. However, other risk factors for QTc prolongation may increase the risk of cardiac complications in patients treated with methadone including coadministration with other medications increasing QT interval such amifampridine, metronidazole, thioridazine, or ziprasidone.

In the United States, methadone can be dispensed only at an outpatient opioid treatment program (OTP) certified by SAMHSA and registered with the Drug Enforcement Administration (DEA) or to a hospitalized patient in an emergency. SAMHSA-certified OTP facilities provide daily doses of methadone under a direct supervision until the patient is deemed stable enough to receive take-home doses. Decision to dispense methadone from the OTP, and to be taken unsupervised at home, should balance the benefits of reduced frequency of clinic visits in stable patients with the risks of misuse (e.g., injection) and diversion of methadone to the illicit drug market.

BUPRENORPHINE. Buprenorphine is a high-affinity partial μ-opioid receptor agonist, κ-receptor antagonist, and ORL-1 agonist. At lower doses of up to 16 mg buprenorphine has a dose-dependent opiate-like effect but at higher doses there is progressively less increase in opiate-like effects up to a point where a maximum effect is reached. This ceiling to the μ-opioid agonist effects of buprenorphine limits adverse effects such as respiratory depression, which increases its safety profile. Buprenorphine has poor oral bioavailability and therefore it is usually administered sublingually as a tablet or film and as a buccal film. To decrease the potential for abuse of buprenorphine given on a chronic basis, pharmaceutical formulations contain buprenorphine mixed with naloxone in a four to one dose ratio. Because naloxone has poor sublingual bioavailability, it exerts minimal effect when a medication is taken sublingually as prescribed. However, if a buprenorphine–naloxone product is dissolved and injected by a person physically dependent on opioids such as heroin, naloxone exerts a maximal effect and it blocks some of the effects of buprenorphine and may even precipitate withdrawal.

Buprenorphine is highly lipid soluble and has a long elimination half-life (24 to 48 hours) allowing for a once-daily dosing and even less frequent dosing with higher doses. Agonist effects of buprenorphine prevent withdrawal symptoms from emerging. Because buprenorphine has a higher receptor affinity than most of other opioids, it effectively blocks effects of agonists by a direct occupation of the receptor sites unlike methadone that produces blocking effect primarily via the mechanism of cross-tolerance (maintaining high levels of tolerance). Buprenorphine has been shown to be effective in the treatment of opioid dependence in an office-based

setting and buprenorphine 12 mg is equally effective in promoting treatment retention and suppression of heroin use as 50 to 60 mg of methadone. There is new evidence that buprenorphine doses above 16 mg per day may be effective in patients with limited response to lower doses.

To avoid precipitation of opioid withdrawal, treatment with buprenorphine is best suited to individuals who are not physically dependent or present with at least mild severity withdrawal symptoms. Unlike with methadone, there is little concern about overdose during treatment initiation with buprenorphine. However, combining buprenorphine with benzodiazepines or other sedative-hypnotics can produce excessive sedation and there are case reports of fatalities associated with a parenteral misuse of buprenorphine along with a benzodiazepine. Common side effects of buprenorphine include constipation, nausea, headache, upset stomach, excessive sweating, somnolence, and decreased interest in sex. Most of these side effects are mild in severity and decrease over time.

Similarly to methadone, buprenorphine is primarily metabolized by the CYP3A4 and medications that block this enzyme, such as atazanavir/ritonavir may produce excessive sleepiness while enzyme inducers, such as rifampin may precipitate opiate withdrawal. The clinical impact of pharmacokinetic interactions with buprenorphine is generally less significant than with methadone; however, dose adjustment may be necessary to achieve optimal clinical response and minimize risks for adverse effects. There is no evidence that buprenorphine produces significant cognitive impairment or ECG QTc prolongation.

In the United States, physicians can treat opioid dependence using buprenorphine, with some restrictions, in a manner similar to treatment with other controlled substances. The law passed in the 2000 (the Drug Addiction Treatment Act) allowed office-based opioid treatment of opioid dependence (OBOT) with schedule III, IV, or V medications that were approved for an opioid dependence indication, and buprenorphine, which was approved by the FDA in 2002, meets these criteria. In order to prescribe buprenorphine, physicians must complete a training course and receive a waiver granted by the DEA. Waivered physicians are qualified to treat up to 30 patients in the first year of their waiver, after which they can request a permission to increase the government-mandated patient cap to 100. Buprenorphine can also be dispensed in a manner similar to methadone in the OTP, and in this case there is no restriction on the number of patients treated by a prescribing physician.

NALTREXONE. The use of opioid antagonists to treat opioid dependence is based on the notion that after detoxification an antagonist will prevent opioids from exerting euphoric effects and that in individuals who resume opioid use while on naltrexone the reinforcing effects of the opioid will

gradually extinguish. Such individuals will cease use after few attempts and will maintain abstinence. The effectiveness of this treatment is contingent on the full compliance with the antagonist. In addition to providing blockade of opioid effects in case of use, patients taking naltrexone experience less craving in the presence of opioid-related cues, which may assist in maintaining abstinence.

Naltrexone is a high-affinity and long-acting μ-opioid receptor antagonist that provides a dose-dependent blockade of heroin and synthetic opioids effects. Naltrexone is an attractive treatment option, as in contrast to methadone or buprenorphine it is not a controlled substance so it can be prescribed without a special license, it has no abuse potential, and there is no withdrawal distress on stopping treatment as patients treated with naltrexone are not physiologically dependent on opioids. Naltrexone can be useful in (1) patients who have not had treatment success with agonists, (2) patients with a high motivation for abstinence from all opioids such as those very active in a 12-step program, and (3) patients who have been successful in agonists therapy and wish to discontinue it but would like to be protected against relapse with medication. However, naltrexone remains a rarely used treatment option due to the need to be fully detoxified before treatment can be started (7 to 10 days of an opioid-free period), poor compliance in many patients, and continuing protracted withdrawal symptoms that may persist in early weeks of treatment with naltrexone.

Naltrexone is available as an oral 50 mg tablet that can be taken daily or three times a week (e.g., 100 mg on Mondays and Wednesdays, and 150 mg on Fridays) to provide stable blood levels to block substantial doses of heroin. Naltrexone is also available in extended-release preparations that can maintain sufficient blocking serum levels of naltrexone. One of such preparations approved by the FDA is a monthly intramuscular injection. There are also implantable preparations of naltrexone with a much longer duration of action (3 to 6 months) that are available in other countries; however, only one of these implants has full regulatory approval of the national drug monitoring agency in Russia.

The effectiveness of treatment with injectable extended release naltrexone has more than double the success rate of treatment with oral naltrexone. Oral treatment is most effective in the context of involving a family member or a professional to provide ongoing support and monitoring of medication administration, which is critical at the beginning of treatment. Alternatively, compliance with oral naltrexone may be improved with the incorporation of a CM behavioral approach into the treatment.

Because naltrexone is not metabolized by the P450 cytochromes interactions with other medications are not of clinical concern. Naltrexone is contraindicated in patients who are taking any opioids, including partial

and mixed agonists such as buprenorphine and nalbuphine, as it will precipitate a severe withdrawal in such patients.

OTHER MEDICAL APPROACHES. There is a considerable number of patients that are unable to benefit from treatment with methadone, buprenorphine, or naltrexone and alternative approaches to provide these individuals with effective treatment have been explored. This includes a maintenance treatment with slow-release oral morphine and with a medically prescribed pharmaceutical-grade heroin.

Morphine has been used in individuals who are unable to tolerate or do not find effects of methadone useful as being different from effects of heroin. In those patients, morphine can have positive effects on mood and sleep and eliminate craving for heroin in contrast to methadone. It may also minimize the use of other medications (e.g., sedatives) often seen when such patients are treated with methadone. Maintenance treatment with a once daily slow-release preparation of morphine is available in several European countries. However, because of a short duration of action many such patients continue to need methadone to prevent withdrawal symptoms overnight and the use of morphine in a treatment setting is complicated by the difficulties in assessing heroin use in such patients. At the present time there is not a sufficient number of experimental studies to draw generalizable conclusions.

Prescribing heroin as a treatment strategy in heroin dependence has been available in the United Kingdom since 1926 and in the past 30 years this approach has been piloted and researched in other European countries and in Canada. Heroin maintenance is usually offered as a "last resort" treatment and there is a broad consensus that it should be offered to only a very selective group of patients with chronic illness who were offered and failed other approaches including a high-dose methadone maintenance treatment, the most effective of all medication-assisted approaches. Recent studies focused on the effectiveness of supervised prescription injectable heroin in participants that failed treatment with methadone, and where prescribed heroin was used in combination with oral methadone. In such studies patients treated with heroin have better treatment retention and reduced use of illicit drugs as compared to patients treated with methadone, but heroin-treated patients had a greater risk of adverse events. At the present time, there is not a sufficient number of experimental studies to draw conclusions for the practice of heroin treatment in the community.

Treatment of Special Populations

There is a complex relationship between the criminal justice system, patients with substance abuse disorders, and drug-dependence treatment systems. Treatment of criminal justice patients includes efforts directed at

persons who are incarcerated, and interventions are designed for persons under legal restraints but still residing in the community. Specialized interventions for opioid-dependent criminal justice patients in the community have been developed. Drug courts, which now operate within each state in the United States, combine close criminal justice supervision with long-term drug abuse treatment that typically requires drug testing of parolees.

In general, illicit opioid use during pregnancy is associated with adverse effects for both the pregnant woman and the fetus, but is not associated with any specific teratogenic effects. However, there is a higher risk of infections, pre-eclampsia, miscarriage, premature rupture of membranes, and premature labor as well as stillbirths, premature infants, and infants with low birth weight and neonatal opioid withdrawal (also called the neonatal abstinence syndrome). Infants who survive may be infected with HIV and other diseases as a result of maternal high-risk behavior and are at risk of sudden infant death syndrome (SIDS). Most of these risks can be substantially reduced with methadone or buprenorphine maintenance treatment.

Healthcare professionals often come to treatment when their drug use is discovered by colleagues or by drug enforcement agencies. Under supervision, most (as much as 85 percent) of these patients recover and are able to return to practice. Returning to specialties with easy access to opioids can pose a serious hazard to continued abstinence. Treatment with extended-release naltrexone, which is easy to supervise, may be particularly useful in those patients.

References

Amass L, Pukeleviciene V, Subata E, et al. A prospective, randomized, multicenter acceptability and safety study of direct buprenorphine/naloxone induction in heroin-dependent individuals. *Addiction*. 2012;107(1):142–151.

American Psychiatric Association. Practice guideline for the treatment of patients with substance use disorders, second edition. *Am J Psychiatry*. 2007;164[Suppl]:4.

Bukten A, Skurtveit S, Gossop M, et al. Engagement with opioid maintenance treatment and reductions in crime: a longitudinal national cohort study. *Addiction*. 2012;107(2):393–399.

Cornish R, Strang J, Vickerman P, Hickman M. Risk of death during and after opiate substitution treatment in primary care: prospective observational study in UK General Practice Research Database. *BMJ*. 2010;341:c5475.

Darke S, Mills KL, Ross J, Teesson M. Rates and correlates of mortality amongst heroin users: findings from the Australian Treatment Outcome Study (ATOS), 2001–2009. *Drug Alcohol Depend*. 2011;115(3):190–195.

Ferri M, Davoli M, Perucci CA. Heroin maintenance for chronic heroin-dependent individuals. *Cochrane Database Syst Rev*. 2011;(12):CD003410.

Filizola M, Devi LA. Structural biology: how opioid drugs bind to receptors. *Nature*.

2012;485(7398):314–317.

Gelernter J, Kranzler HR, Sherva R, et al. Genome-wide association study of opioid dependence: multiple associations mapped to calcium and potassium pathways. *Biol Psychiatry*. 2014;76(1):66–74.

Goldstein RZ, Volkow ND. Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. *Nat Rev Neurosci*. 2011;12(11):652–669.

Gowing L, Farrell MF, Bornemann R, Sullivan LE, Ali R. Oral substitution treatment of injecting opioid users for prevention of HIV infection. *Cochrane Database Syst Rev*. 2011; (8):CD004145.

Grella CE, Karno MP, Warda US, Niv N, Moore AA. Gender and comorbidity among individuals with opioid use disorders in the NESARC study. *Addict Behav*. 2009;34(6-7):498–504.

Havnes I, Bukten A, Gossop M, Waal H, Stangeland P, Clausen T. Reductions in convictions for violent crime during opioid maintenance treatment: a longitudinal national cohort study. *Drug Alcohol Depend*. 2012;124(3):307–310.

Hser YI, Hoffman V, Grella CE, Anglin MD. A 33-year follow-up of narcotics addicts. *Arch Gen Psychiatry*. 2001;58(5):503–508.

Jones HE, Heil SH, Baewert A, et al. Buprenorphine treatment of opioid-dependent pregnant women: a comprehensive review. *Addiction*. 2012;107(Suppl 1):5–27.

Jones HE, Kaltenbach K, Heil SH, et al. Neonatal abstinence syndrome after methadone or buprenorphine exposure. *N Engl J Med*. 2010;363(24):2320–2331.

Kakko J, Svanborg KD, Kreek MJ, Heilig M. 1-year retention and social function after buprenorphine-assisted relapse prevention treatment for heroin dependence in Sweden: a randomised, placebo-controlled trial. *Lancet*. 2003;361(9358):662–668.

Kendler KS, Jacobson KC, Prescott CA, Neale MC. Specificity of genetic and environmental risk factors for use and abuse/dependence of cannabis, cocaine, hallucinogens, sedatives, stimulants, and opiates in male twins. *Am J Psychiatry*. 2003;160(4):687–695.

Koob GF, Volkow ND. Neurocircuitry of addiction. *Neuropsychopharmacology*. 2010;35(1):217–238.

Kreek MJ, Nielsen DA, Butelman ER, LaForge KS. Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction. *Nat Neurosci*. 2005;8(11):1450–1457.

Krupitsky E, Nunes EV, Ling W, Illeperuma A, Gastfriend DR, Silverman BL. Injectable extended-release naltrexone for opioid dependence: a double-blind, placebo-controlled, multicentre randomised trial. *Lancet*. 2011;377(9776):1506–1513.

Kunoe N, Lobmaier P, Vederhus JK, et al. Naltrexone implants after in-patient treatment for opioid dependence: randomised controlled trial. *Br J Psychiatry*. 2009;194(6):541–546.

MacArthur GJ, Minozzi S, Martin N, et al. Opiate substitution treatment and HIV transmission in people who inject drugs: systematic review and meta-analysis. *BMJ*. 2012;345:e5945.

Martinez D, Saccone PA, Liu F, et al. Deficits in dopamine D2 receptors and pre-synaptic dopamine in heroin dependence: commonalities and differences with other types of addiction. *Biol Psychiatry*. 2012;71:192–198

Mattick RP, Breen C, Kimber J, Davoli M. Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. *Cochrane Database Syst Rev*. 2008;

(2):CD002207.

McLellan AT, Arndt IO, Metzger DS, Woody GE, O'Brien CP. The effects of psychosocial services in substance abuse treatment. *JAMA*. 1993;269(15):1953–1959.

McNicholas L. *Clinical Guidelines for the Use of Buprenorphine in the Treatment of Opioid Addiction: A Treatment Improvement Protocol TIP 40*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2004.

SAMHSA. *Results from the 2012 National Survey on Drug Use and Health: Summary of National Findings*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2013.

Saxon AJ, Ling W, Hillhouse M, et al. Buprenorphine/Naloxone and methadone effects on laboratory indices of liver health: a randomized trial. *Drug Alcohol Depend*. 2013;128(1-2):71–76.

Sigmon SC, Bisaga A, Nunes EV,, Kosten T, Woody G. Opioid detoxification and naltrexone induction strategies: recommendations for clinical practice. *Am J Drug Alcohol Abuse*. 2012;38(3):187–199.

Strain EC, Bigelow GE, Liebson IA, Stitzer ML. Moderate- vs high-dose methadone in the treatment of opioid dependence: a randomized trial. *JAMA*. 1999;281(11):1000–1005.

Strain EC, Stitzer ML. *The Treatment of Opioid Dependence*. Baltimore, MD: Johns Hopkins University Press; 2006.

Teesson M, Mills K, Ross J, Darke S, Williamson A, Havard A. The impact of treatment on 3 years' outcome for heroin dependence: findings from the Australian Treatment Outcome Study (ATOS). *Addiction*. 2008;103(1):80–88.

UNODC. *United Nations Office on Drugs and Crime. World Drug Report 2014*. New York: United Nations; 2014.

Weiss RD, Potter JS, Fiellin DA, et al. Adjunctive counseling during brief and extended buprenorphine-naloxone treatment for prescription opioid dependence: a 2-phase randomized controlled trial. *Arch Gen Psychiatry*. 2011;68(12):1238–1246.

WHO. *World Health Organization Guidelines for the Psychosocially Assisted Pharmacological Treatment of Opioid Dependence*. Geneva: WHO Press; 2009.

Woody GE, Poole SA, Subramaniam G, et al. Extended vs short-term buprenorphine-naloxone for treatment of opioid-addicted youth: a randomized trial. *JAMA*. 2008;300(17):2003–2011.

▲ 11.10 Sedative-, Hypnotic-, or Anxiolytic-Related Disorders

JOHN J. MARIANI, M.D.

DEFINITION

In this chapter, the epidemiology, neurophysiology, clinical characteristics, and management of sedative, hypnotic, and anxiolytic-related disorders will be discussed. Sedative, hypnotic, and anxiolytic medications constitute a broad group of agents that are classified by their clinical effects, with most medications in this grouping having GABA agonist or GABAergic transmission facilitation effects. In this chapter, these medications will be collectively referred to as "sedative-hypnotics," with the understanding that within this class of medications, distinctions as to clinical effects are to some degree artificial. From both a public health and clinical management perspective, the benzodiazepines are the most important sedative-hypnotic agents because of their common use and wide availability, both by prescription and from nonprescription sources. Benzodiazepines and other sedative-hypnotic agents are widely prescribed because of their efficacy in the treatment of anxiety, insomnia, and other therapeutic indications. Like other prescribed controlled substances, most individuals will be able to use sedative-hypnotic agents without developing a pattern of problematic or compulsive use. However, for a minority of individuals exposed to sedative-hypnotic agents, a pattern of use meeting criteria for a substance use disorder may develop. These individuals need specialized assessment and care to avoid the potential behavioral and medical consequences of sedative-hypnotic use disorder.

Studies of sedative, hypnotic, or anxiolytic use disorders are limited by the operational definition of substance use disorders set forth in the American Psychiatric Associations' *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; DSM-5). Sedative-hypnotic agents, and related anxiolytic medications, have the potential to induce tolerance, and with regular use individuals may be at risk of a discontinuation syndrome, or frank withdrawal state. However, drug addiction is a sociobehavioral phenomenon, albeit with significant pharmacological origins and medical consequences. There is a tendency to conflate the neuroadaptive changes that develop in response to chronic sedative-hypnotic administration (physiological dependence) with the concept of a substance use disorder or addiction. Many individuals who have difficulty tapering or discontinuing sedative-hypnotic agents because of neuroadaptive changes may be considered to be "addicted," by themselves or clinicians, but do not meet full criteria for a substance use disorder unless other symptoms are present.

The addictive potential of a substance is usually inferred from data on illicit trafficking; population surveys of nonprescription use; audits of prescribing practices; postmarketing studies; monitoring of diversion from physician's offices, pharmacies, or manufacturing sites; emergency room treatment; and overdose deaths. Pharmacological characteristics also predict misuse liability: Drug-induced euphoria and the development of tolerance, an abstinence syndrome, or cross-tolerance with known addictive substances may all predict the risk of misuse or addictive potential. Finally, a number of laboratory methods for determining drug preferences in human and animal models have been developed to determine misuse liability. Sedative-hypnotic medications that have general GABAergic effects tend to have the potential for misuse and addiction. In general, because of their risk of misuse and addiction, sedative-hypnotic medications are classified as controlled substances. This risk is related to, but distinct from the tendency for GABAergic medications to induce pharmacological tolerance and be predisposed to developing withdrawal. An individual who developed tolerance and withdrawal while prescribed a sedative-hypnotic agent, and displayed no other symptoms, would not meet criteria for a substance use disorder.

Various terms are used to refer patterns for the use of controlled substances. *Misuse* refers to the use of a drug in a manner that is not consistent with generally accepted medical practice or social and legal custom (e.g., use without a valid prescription or use to deliberately produce intoxication, pleasure, or a high). *Recreational use* refers to the use of drugs solely for their hedonic value, but lacks specificity, and has an inappropriate harmless connotation. Substance use disorder as defined in the DSM-5 combines the DSM-IV-TR categories of substance abuse and substance dependence into a single disorder measured on a continuum from mild to severe. Although a diagnosis of substance abuse previously required only one symptom, mild substance use disorder in DSM-5 requires two to three symptoms from a list of 11. The WHO Tenth Revision of the International Classification of Diseases and Health Problems (ICD-10) retains the term "substance dependence" to indicate "addiction." The term *physiological dependence* is applied more narrowly in the psychopharmacological context "referring solely to the development of withdrawal symptoms on cessation of drug use."

The *abstinence or discontinuation syndrome* that develops with the discontinuation of sedative-hypnotic agents has specific characteristics referable to the autonomic nervous system (e.g., tremor, sweating, tachycardia, and startle response) and must be clearly distinguished from recurrence of the features of the underlying disease for which the drug was originally prescribed. It can be reversed or attenuated by readministering the discontinued drug or administering a drug with cross-tolerance. A *withdrawal syndrome* is associated with many sedative-hypnotic agents,

which in some cases can be associated with severe medical sequelae. In this chapter, abstinence syndrome will refer to subjective symptoms that cannot be objectively measured (e.g., anxiety, insomnia, and craving) and withdrawal syndrome will refer to a more severe neurophysiological disturbance where symptoms can be objectively measured (e.g., elevated pulse or blood pressure) or observed (e.g., tremor).

Substances

The drugs discussed in this section are referred to as *anxiolytic* or *sedative-hypnotic drugs*. The terminology for this group of drugs is not clearly established. Their sedative or calming effects are on a continuum with their hypnotic or sleep-inducing effects. Anxiolytic drugs cover a wide spectrum of pharmacological agents, and the terminology implies more specificity than actually exists. The broad group of drugs that historically has been included in that class exhibits considerable variation in clinical use, toxicity, risk for addiction, and potential for diversion and misuse. Sedative-hypnotic agents are classified in this section as follows: (1) benzodiazepines and functionally related drugs, (2) barbiturates, and (3) miscellaneous sedative-hypnotic drugs. In the practice of psychiatry, the drugs that are most important clinically are the benzodiazepines.

ETIOLOGY AND NEUROPHARMACOLOGY

Although most of the drugs discussed in this chapter have their primary effects on the γ-aminobutyric acid type A ($GABA_A$) receptor complex, there are subtle differences in their mechanisms of action at this site. Furthermore, neuroadaptive changes that occur with prolonged drug exposure affect not only the $GABA_A$ receptor but other neurotransmitters, the hypothalamic pituitary axis, and calcium modulated systems. GABA is the major inhibitory neurotransmitter in the brain.

GABA type A receptors are a family of ligand-gated ion channels responsible for inhibitory regulation of the CNS. Benzodiazepines bind directly to a specific modulatory site that is present on $GABA_A$ receptors, which enhances the effect of the inhibitory neurotransmitter, but does not open the chloride channel in the absence of GABA. Chronic exposure to benzodiazepines leads to alterations of GABAergic neurotransmission, which are manifested as the symptoms of tolerance, dependence, and withdrawal. The mechanism of these changes is unclear, and may involve alterations in the expression of individual $GABA_A$ receptor subtypes, but is known to be linked to the dose and duration of benzodiazepine use. While the direct effects of benzodiazepines are GABA-related, the withdrawal syndrome may be in part be mediated by calcium channel and glutamatergic mechanisms. In addition, neuroendocrine pathways, including regulation of the hypothalamic–pituitary–adrenocortical axis by

CRF, may be involved in the mechanism of benzodiazepine dependence.

Activation of $GABA_A$ receptors by agonists or positive modulators leads to an influx of chloride ions into neurons, which then produce hyperpolarization, thereby decreasing the influence of excitatory input. $GABA_A$ receptors exist as protein structures that consist of five subunits. In the brain, these structures commonly contain two α, two β, and one γ subunit. Each subunit constitutes a unique sequence of amino acids, which are differentiated into numbered subtypes based on minor variations in the amino acid sequence. For example, the α subunit is thought to have six subtypes, designated α1 to 6. Receptors containing α1, α2, α3, and α5 subunits are sensitive to the effects of benzodiazepines. In contrast, the presence of α4 or α6 subunits leads to a lack of sensitivity to benzodiazepines. Benzodiazepines bind at the interface of the γ2 and α subunits.

Agonists that interact with the benzodiazepine receptor produce alterations in the structure of the $GABA_A$ receptors, which can enhance the affinity of these receptors for GABA and direct receptor agonists, and are often referred to as positive modulators, to differentiate them from direct agonists, such as GABA and muscimol. The benzodiazepine-induced allosteric changes in the $GABA_A$ receptor complex result in an increase in the frequency of opening of the chloride channel located in the center of $GABA_A$ receptor complex.

Barbiturates share some similarity in their actions with those of benzodiazepines. At low concentrations, barbiturates act at a receptor site other than the GABA binding site on the $GABA_A$ receptor complex to positively modulate the actions of GABA via an allosteric mechanism. However, at higher concentrations barbiturates act as $GABA_A$ receptor agonists, directly increasing the activity of these receptors. $GABA_A$ receptors that contain the β2 or β3 subunit show greater sensitivity to the effects of pentobarbital than receptors with the β1 subunit. The type of α subunit may be a factor in determining the extent to which a barbiturate may act as a direct GABA agonist. Pentobarbital shows the greatest efficacy as a GABA receptor agonist in receptor complexes that contain the α6 subunit, which is a benzodiazepine insensitive subunit.

Barbiturates also may have actions on neurons that involve mechanisms other than inhibition of $GABA_A$ receptor activity. At least in the hippocampus, pentobarbital may inhibit the activity of Dl-α-amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA) receptors. These receptors are involved in the mediation of some of the effects of excitatory neurotransmitter glutamate. Barbiturates also may inhibit the release of neurotransmitters by blocking voltage-sensitive calcium channels.

Diazepam (Valium) and other classical benzodiazepines appear to bind with equal affinity to benzodiazepine-sensitive receptors with different

subunit composition. In contrast, so-called nonbenzodiazepine hypnotics, such as the imidazopyridine, zolpidem (Ambien), the pyrazolopyrimidine, zaleplon (Sonata), and the cyclopyrrolone, eszopiclone (Lunesta), act more selectively on $GABA_A$ receptors that contain the α1 subunit (often referred to as $GABA_{A1}$ receptors). These differences in selectivity may be of significance because the nature of the subunits may be related to the specific clinical effects of the agonist. $GABA_A$ receptors that contain α1 subunits appear to mediate the sedative-hypnotic effects of benzodiazepines. This subunit has also been implicated in playing a key role in the production of the ataxic effects of both benzodiazepines and zolpidem. The antianxiety effects of benzodiazepine may be mediated by receptors containing the α2, and possibly the α3 subunit. The α2 and α3 subunits also may be involved in the mediation of the muscle relaxant effects of benzodiazepines.

Tolerance develops to many of the actions of benzodiazepines. Tolerance first develops to the sedative effects of these drugs. This may then be followed by tolerance to benzodiazepine-induced amnesia and to the anticonvulsant effects of these drugs. In addition to tolerance, benzodiazepines may produce a withdrawal syndrome upon abrupt discontinuation.

The development of tolerance to the effects of benzodiazepines may involve mechanisms that include the uncoupling of the allosteric interaction that occurs between benzodiazepines and GABA binding sites and the internalization into neurons of the $GABA_A$ receptor complex. Tolerance to benzodiazepines also may involve alterations in the subunit composition of $GABA_A$ receptors. Chronic exposure to classical benzodiazepine compounds such as diazepam may lead to an alteration in the gene expression for the mRNAs that are responsible for the synthesis of specific $GABA_A$ subunits. These changes may include a decrease in the mRNAs for cortical and hippocampal α1 subunits. Within the hippocampus, binding to hippocampal α5-GABA receptors is decreased in animals that are tolerant of the sedative effects of diazepam. The tolerance to the motor activity depressant effects of diazepam was not observed in animals missing the gene for the α5 subunit.

Chronic exposure to the α1 selective drugs zolpidem and zaleplon results in less extensive changes in the expression of $GABA_A$ receptor units than is observed following prolonged treatment with classical benzodiazepines. The repeated administration of either zolpidem or zaleplon may not decrease mRNAs encoding for the α1 and other $GABA_A$ receptor subunits. Following short-term withdrawal from these agents, however, an elevation in α4 mRNA may occur, and changes in the expression of mRNA for other $GABA_A$ receptor subunits are similar to that seen with withdrawal from diazepam, including a decrease in the mRNA for

the α1 subunit.

It was once believed that the primary mechanisms underlying withdrawal symptoms following the discontinuation of chronic benzodiazepine treatment were related to changes in the GABA system: $GABA_A$ subunit expression, $GABA_A$ receptor down regulation, or uncoupling of benzodiazepine receptors. More recently, changes in the glutamatergic systems have been shown to be involved in withdrawal. In several brain regions, including the frontal cortex and hippocampus, withdrawal symptoms occur in association with increased expression of the AMPA receptor GluR1 subunit. Electrophysiological evidence also indicates that the activity of AMPA receptors is increased during withdrawal from benzodiazepines. In addition to changes in AMPA receptor function, chronic exposure to benzodiazepines may result in upregulation of L-type calcium channels. Such upregulation occurs in association with physical dependence on a variety of abused substances, including ethanol and nicotine.

Tolerance to barbiturates may involve both an increase in the rate of metabolism of these compounds and several neuroadaptive responses. Barbiturates such as pentobarbital (Nembutal) and phenobarbital (Bellatal) can increase the activity of cytochrome P450 enzymes, including the CYP2B and CYP3A enzymes. This effect can lead to enhanced barbiturate metabolism that in turn results in a reduction in barbiturate plasma concentrations and pharmacokinetic tolerance. The prolonged administration of barbiturates may produce some alterations in the expression of different $GABA_A$ receptor subunits. The expression of the α6 $GABA_A$ receptor subunit, for example, is increased in pentobarbital tolerant animals. Chronic phenobarbital administration may result in an increase in NMDA receptors in the cortex during withdrawal. This upregulation of NMDA receptors may explain the high level of CNS excitability, which is a feature of withdrawal from barbiturates.

MISUSE LIABILITY

Surveys of prescribing practices, patient-initiated dosage changes, and recreational or nonmedical benzodiazepine use along with reports from emergency rooms, medical examiners, and law enforcement agencies provide a complex picture of the public health implications of the benzodiazepines. The vast majority of medical and psychiatric patients use benzodiazepines appropriately, although rates of misuse by patients dependent on alcohol or other drugs may be higher than individuals with anxiety or insomnia without a history of substance use problems. Furthermore, long-term use is high among patients with psychiatric disorders and the elderly, with the latter group being sensitive to drug toxicity.

One approach to assessing misuse liability is by measuring the reinforcing properties of the drug. If its pharmacological effects increase behavior (i.e., self-administration of a dose or the work required to permit self-administration), then the drug is a positive reinforcer and has abuse liability. Animal and human studies have assessed the reinforcing properties of benzodiazepines.

$GABA_A$ Receptor Subtypes and Addiction Risk

Benzodiazepines are positive allosteric modulators of the γ-aminobutyric acid type A receptors ($GABA_A$Rs). Recent work on the neural basis of the addictive properties of benzodiazepines suggests that receptors that contain the $GABA_A$R α1 subunit are important for triggering drug-evoked synaptic plasticity at ventral tegmental area (VTA) synapses, which is an important first step underlying drug reinforcement. Benzodiazepines have a stronger direct effect on GABA neurons than on dopamine neurons. GABA neurons are more strongly hyperpolarized in the presence of benzodiazepines and no longer inhibit dopamine neurons, resulting in an increase in dopamine transmission. This disinhibition of VTA dopamine release may be a mechanism for the development of benzodiazepine addiction, as well as potential benzodiazepine facilitation of addiction to other substances.

Animal Studies

Animal studies have consistently demonstrated that benzodiazepines have minimal reinforcing effects. Although a small number of studies show that animals self-administer benzodiazepines more frequently than other drugs, they are significantly less potent than other drugs of abuse (e.g., cocaine and other CNS stimulants) or even many other sedative-hypnotic drugs, including most barbiturates. Studies that use the conditioned place preference paradigm have also demonstrated that benzodiazepines have minimal reinforcing effects. In some studies, diazepam has produced place preference among laboratory rats, but to a lesser extent than classic drugs of abuse, such as morphine or amphetamine.

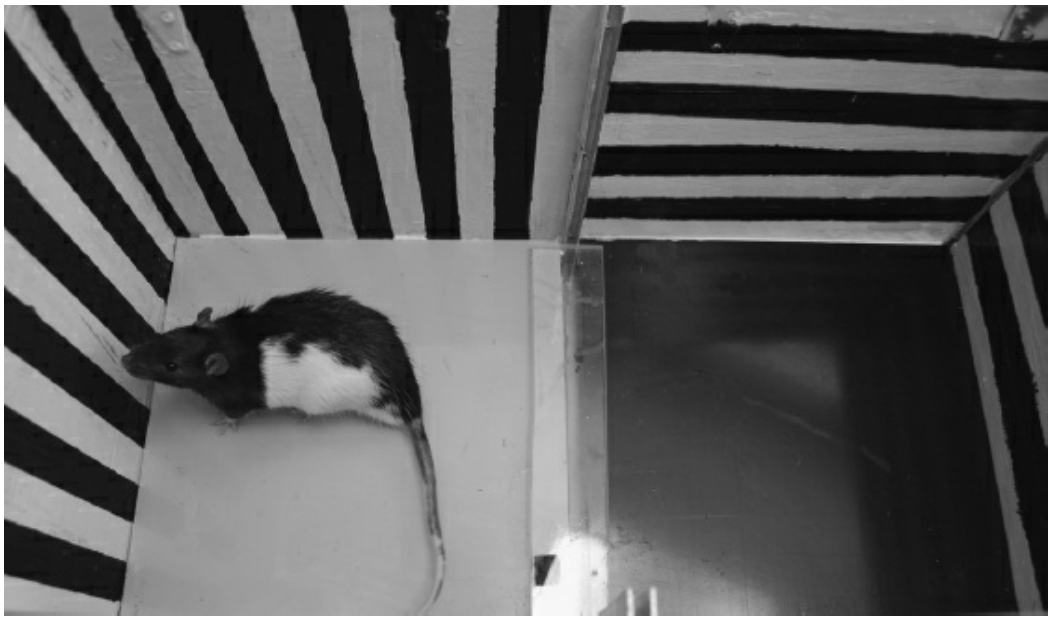

FIGURE 11.10–1. In this conditioned place preference experiment the animal spends more time in the chamber compartment in which a reinforcing drug had been repeatedly administered (*the white floor*) than in compartment he associates with saline administration (*the black floor*). (Courtesy of Conan Kornetsky, Boston University School of Medicine.)

Animal studies using schedule-induced polydipsia (the use of intermittent food administration to the experimental animal, which promotes ingestion of fluids, including drug solutions) indicate that midazolam (Versed) is ingested more often than water, but self-administration of chlordiazepoxide (Librium) and flurazepam (Dalmane) does not differ significantly from that of water.

Another animal model predicting abuse liability consists of drug infusion in response to activity by the animal, such as lever pressing (Fig. 11.10–1). Drug infusion may allow continuous or intermittent drug availability, and rates of administration of the drug are compared to those of the vehicle, other sedative-hypnotic drugs, and standard drugs of abuse, such as cocaine. Studies with continuous drug access suggest that triazolam (Halcion), diazepam, midazolam, clobazam (Frisium), and chlordiazepoxide have reinforcing effects; studies with intermittent drug availability show that diazepam, midazolam, triazolam, alprazolam (Xanax), lorazepam (Ativan), chlordiazepoxide, and bromazepam (Lectopam) maintain responses at a higher level than vehicle, indicating abuse liability. With the possible exception of triazolam and midazolam, in studies using this model, the benzodiazepines consistently demonstrate lower abuse liability than barbiturates other than phenobarbital, which also has relatively low abuse liability. The nonbenzodiazepine hypnotics that act on the $GABA_1$ receptor, such as zolpidem, zaleplon, and eszopiclone (Lunesta), are similar to the benzodiazepines in animal models of abuse. Partial benzodiazepine receptor agonists, such as abecarnil and bretazenil,

have low abuse liability in these models. Prior exposure to ethanol or sedative-hypnotic drugs may increase the reinforcing effects of benzodiazepines.

Animal studies have demonstrated that a withdrawal syndrome occurs on abrupt discontinuation of benzodiazepines and related sedative hypnotics. The animal models using the social interaction test and social conflict paradigms, along with isolation-induced ultrasonic vocalizations recordings, indicate that abrupt discontinuation of benzodiazepine tranquilizers may induce numerous anxiogenic responses, including decreased social interaction and increased aggressive behaviors. Studies that use a competitive benzodiazepine receptor antagonist, such as flumazenil (Romazicon), to precipitate withdrawal, have also demonstrated the development of a withdrawal syndrome using electrophysiological measures.

Human Studies

The assessment of misuse liability in human laboratory studies relies on variations in two predictive models: self-administration of benzodiazepines by experimental subjects who have a history of substance use problems and subjective responses that correlate with high misuse potential. In the first paradigm, one type of study allows subjects with a history of sedative-hypnotic use disorder to self-administer orally (i.e., sample) different color-coded drugs and placebo under double-blind conditions in a simulated social environment. During subsequent sessions, they are permitted to self-administer the drug that they like the best. Some studies require that the subjects perform a task (e.g., riding a stationary bicycle) to earn doses. The misuse liability of different drugs is assessed by how often they are chosen for self-administration or by how much work a subject does to earn a dose. Scales measuring drug effect are also administered. The paradigm has the advantage of directly observing the behavior involved with ingesting the drug, albeit under somewhat artificial circumstances. It works best for individuals with a history of a sedative-hypnotic use disorder in remission without a co-occurring psychiatric disorder. It is probably safe to infer that those subjects are administering the drug for its hedonic value. The reason for self-administration by other subjects may not be clear. Do patients with anxiety disorder self-administer the drugs because they are effective anxiolytics or because they induce euphoria? Similarly, if subjects are in withdrawal from sedative-hypnotic drugs or ethanol, do they choose the agent that alleviates their withdrawal symptoms most effectively or the drug that produces the greatest euphoria? An additional problem with this design is that it exposes subjects to multiple doses of potentially addictive drugs.

An alternative strategy is to administer only a single dose of the test drugs and then to compare subjective responses associated with misuse

liability. A set of scales, which differs slightly among laboratories, usually consists of the Morphine Benzedrine Group subscale of the Addiction Research Center Inventory (ARCI-MBG); self-reports of sedation, liking, and intensity of drug effect; monetary street value of the drug; and similarity to reference drugs of abuse. Studies demonstrate that estimates of misuse liability using this paradigm are consistent with assessments based on the self-administration model. One disadvantage of the single-dose paradigm is that it measures only subjective effects, not behavior. It also has some of the same limitations as the self-administration paradigm. Does the subject like the drug because it has a therapeutic effect or because it is intoxicating? Both designs attempt to circumvent the problem by including an adequate battery of scales that measure several different dimensions of drug effect. Dose equivalency and homogeneity of the subjects (especially with respect to a personal or family history of psychoactive substance abuse and psychiatric illness) are critical factors in the study design and in the interpretation of data from both models.

The predictive models in humans clearly establish that benzodiazepines occupy an intermediate position on the spectrum of misuse liability of sedative-hypnotic drugs. Methaqualone (Sopor) and barbiturates generally produce higher ARCI-MBG and drug-liking scores than do benzodiazepines. Buspirone (BuSpar), a nonbenzodiazepine anxiolytic agent (i.e., it is essentially devoid of benzodiazepine agonist activities), has little or no abuse liability in those models. With respect to differences among the benzodiazepines, the weight of the evidence suggests that diazepam has greater abuse liability than halazepam (Paxipam), oxazepam (Serax), chlordiazepoxide, or clorazepate (Tranxene). Diazepam, lorazepam, alprazolam, and triazolam exhibit similar profiles of abuse potential using those paradigms. The clinical relevance of the findings has not been adequately studied.

Flunitrazepam (Rohypnol), a benzodiazepine agonist, is not marketed in the United States but has entered the country through illegal sources in Central and South America. Reports of misuse and resulting health consequence have been reported. The drug has been implicated in sexual assaults, in which flunitrazepam is surreptitiously added to an alcoholic beverage consumed by the victim to reduce judgment, inhibition, and physical ability to resist sexual activity. Memory for the event may also be impaired by the combination of alcohol and flunitrazepam. The weight of the evidence supports the safety and efficacy of flunitrazepam in clinical populations; however, individuals with alcohol or drug dependence commonly misuse it in combination with alcohol, illicit drugs, or both.

Pharmacological characteristics, especially pharmacokinetics, of benzodiazepines and related compounds influence misuse liability. Some drugs (e.g., diazepam) may appear to have high abuse liability because of rapid absorption from the GI tract after oral doses. However,

pharmacodynamic differences among the drugs are not well studied. For example, little is known about differences in abuse liability between full and partial agonists or among drugs that have preferential effects on specific subtypes of the benzodiazepine receptor. Zolpidem, for example, is an imidazopyridine hypnotic agent with a rapid onset of action and a short elimination half-life. When the drug was first marketed in the 1980s, the manufacturer presented data indicating a low abuse liability for zolpidem, but current investigations may show otherwise. Although in vitro evidence indicates that zolpidem binds preferentially to the benzodiazepine receptor with a high affinity ratio for the α1 subunit, it is unlikely that this characteristic would provide lower potential for abuse; rather, selective binding more likely explains the lack of anticonvulsant and muscle relaxant effects of the drug. Also supporting the similarity between zolpidem and classic benzodiazepines is the fact that zolpidem and triazolam depress energy metabolism in the same areas of the brain. With respect to subjective, psychomotor, and memory effects, there are no significant differences between a single 10-mg dose of zolpidem and a single 0.25-mg dose of triazolam. High doses of triazolam (0.75 mg) are identified more often as similar in effect to barbiturates, benzodiazepine, or alcohol than are high doses of zolpidem (45 mg). In the United States, New Zealand, and some European countries, zolpidem and zopiclone (Imovane) (another nonbenzodiazepine hypnotic) have been shown to have reduced efficacy after a few weeks of treatment, requiring increases in doses, which may contribute to a higher risk of abuse and dependence. In some cases, zopiclone dosages have reached levels as high as 82.5 mg per day, after an initial prescribed dosage of 3.75 to 15.0 mg per day. Dosages of zolpidem have reached 800 mg per day after an initial prescribed dose of 5 to 20 mg per day. Abrupt discontinuation of zolpidem or zopiclone after months of use has resulted in severe abstinence syndromes, which may include seizures. Eszopiclone has been studied in 6-month trials of insomnia and did not show tolerance, but was associated with a mild withdrawal syndrome, consisting of rebound insomnia and anxiety in a very small percentage (2 percent or less) of patients. Eszopiclone at doses of 6 and 12 mg produced euphoric effects comparable to 20 mg of diazepam in previous sedative-hypnotic abusers. Studies examining the abuse liability of 25, 50, and 75 mg of zaleplon in subjects with a history of sedative-hypnotic abuse indicated abuse potential similar to benzodiazepines and benzodiazepine-like hypnotics. A withdrawal syndrome similar to benzodiazepines, including seizures, should be expected in patients who have had long-term, high-dose treatment with zaleplon. In general, clinical experience with zolpidem, zopiclone, eszopiclone, and zaleplon (nonbenzodiazepine hypnotics or z-drugs) has been favorable; however, they produce subjective effects associated with misuse in experienced drug users and caution should be exercised when they are prescribed in

addiction-prone individuals. Although the nonbenzodiazepine hypnotics (z-drugs) are similar to benzodiazepines in human misuse liability measures, the actual rates of misuse may be lower according to some authorities.

Clinical experience and research findings have identified patient characteristics that influence misuse liability. For example, when individuals with a history of alcohol use disorder and their adult children are given a single dose of alprazolam or diazepam, they have greater increases on misuse potential scales than do nonalcoholic controls without a family history of alcoholism. The differences in reinforcing effects are not attributable to pharmacokinetic differences as measured by peripheral plasma concentrations; however, pharmacodynamic differences have been noted using functional MRI and EEG measures. Studies have found that patients with prior alcohol or substance use disorders exhibit a clear preference for benzodiazepines in self-report models. Enhanced mood states after benzodiazepine challenges have also been found in moderate drinkers, although the confound of a family history of substance use problems has not been adequately controlled in all studies. Other findings indicate that moderate drinkers self-administer higher doses of diazepam than do light drinkers. Whether these results are drug specific or apply to the class of benzodiazepines as a whole is not known. Patients with a family history of substance abuse have reported a benzodiazepine preference in some studies. A survey of 5,000 male college students found that 23 percent of those with a positive family history of alcohol use problems reported using benzodiazepines (vs. 0.9 percent of those without such a family history).

Anxiety disorder patients with a history of alcohol or other substance use disorders have been commonly thought to be at a higher risk of misusing benzodiazepines. A study conducted by the Harvard/Brown Anxiety Disorders Research Program (HARP) examined total daily dose, as needed (PRN) use, and continued use of benzodiazepine in patients with and without a history of alcohol dependence. That study found that prior alcohol use disorder was not a strong predictor for benzodiazepine misuse.

Experimental paradigms do not provide evidence that individuals with anxiety disorders are at greater risk of benzodiazepine misuse when compared to nonanxious subjects. The fact that some anxious subjects choose diazepam over placebo in self-administration experiments underscores the limitations of the study design rather than indicating abuse potential. In such studies, it is impossible to determine if subjects choose the drug to relieve psychic distress or to experience its pleasurable effects. In clinical experience, the risk of benzodiazepine misuse in patients with anxiety disorders is low.

EPIDEMIOLOGY

The prevalence of benzodiazepine use disorder is increasing in the United States; the number of treatment admissions for benzodiazepines tripled between 1998 and 2008, while overall admissions increased only 11 percent, and 1 percent of US residents report past-year nonmedical use of sedative-hypnotic medications, a rate that has increased 20 percent from 2008 to 2010. The past year prevalence estimates of sedative hypnotic use disorders in the United States, which consist primarily of benzodiazepine users, range up to 977,000 individuals, which exceed the estimates for heroin or methamphetamine. The lifetime risk of sedative-hypnotic use disorders is approximately 1 percent. Benzodiazepine treatment admissions are more likely than other treatment admissions to receive detoxification and inpatient treatment. Risk factors for benzodiazepine use disorder include participation in substance use disorder self-help groups, younger age, longer duration of benzodiazepine use, higher dose of benzodiazepine use, lower level of education, non-native cultural origin, and outpatient treatment of substance use disorders.

According to the US SAMHSA's 2013 NSDUH, of the estimated 2.8 million persons aged 12 or older who used illicit drugs for the first time within the past 12 months, about 5.2 percent initiated use with tranquilizers and 0.2 percent with sedatives. The 2013 NSDUH reported that in 2013 approximately 522,000 individuals in the United States met criteria for a sedative-hypnotic use disorder. It is estimated that in 2013, 376,000 individuals sought treatment of problems with sedative-hypnotics. Averaged over 2012 and 2103, the NSDUH found that more than half of the nonmedical users of sedative-hypnotics aged 12 or older got the prescription drugs they most recently used "from a friend or relative for free." More than four in five of sedative-hypnotic users who obtained the medication for free indicated that their friend or relative had obtained the drugs from one doctor. An analysis of the NSDUH 2011 data found that white race, older age, delinquency, and lower family income are positively associated with adolescent benzodiazepine misuse (Fig. 11.10–2).

A large-scale population-based survey of benzodiazepine and z-hypnotics found that among 58,967 respondents in Norway, 13,774 (23 percent) received at least one prescription of benzodiazepines or z-hypnotics in the period 2004 to 2008, and benzodiazepine use was associated with a higher risk of severe anxiety, depression, and sleep outcomes.

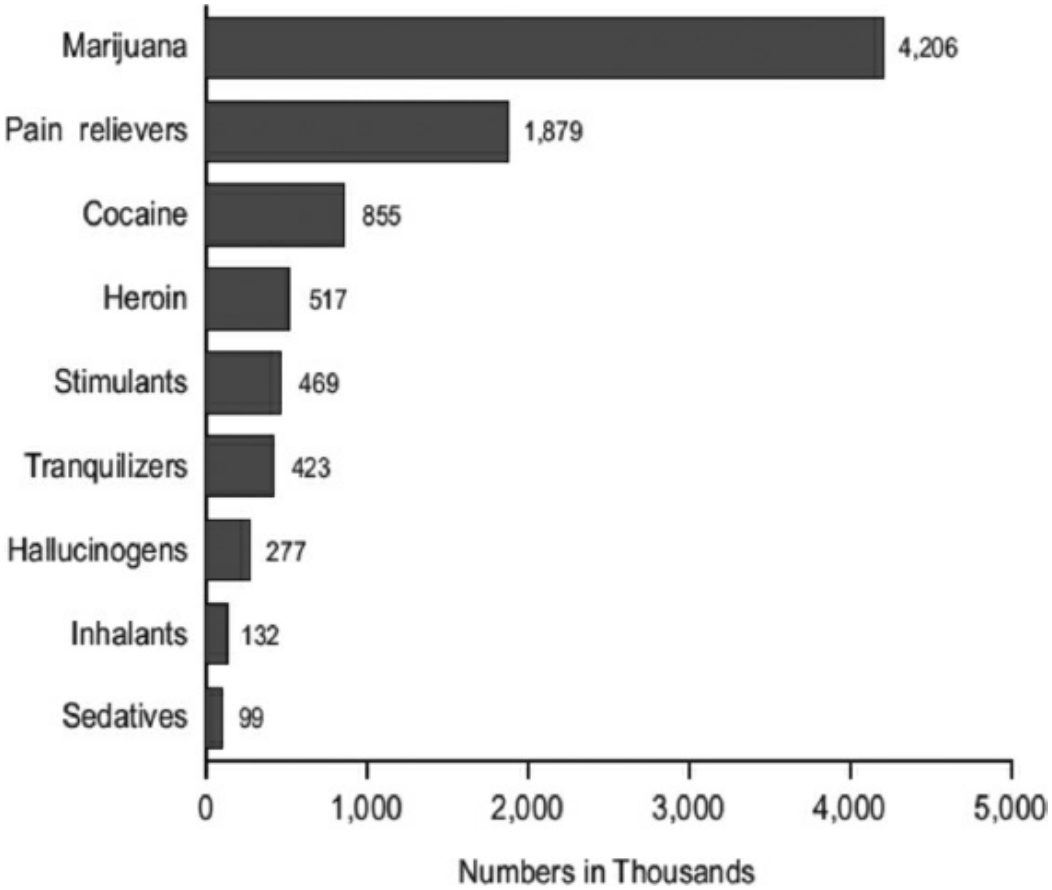

FIGURE 11.10–2. Specific illicit drug dependence or abuse in the past year among persons aged 12 or older: 2013. (From *Survey on Drug Use and Health: Summary of National Findings*, NSDUH Series H-48, HHS Publication No. (SMA) 14–4863. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014.)

MTF is a long-term study of American adolescents, college students, and adult high school graduates. According to the 2014 MTF survey, the prevalence of use of "tranquilizer" (anxiolytic) drugs had annual rates of 1.7 percent, 4.0 percent, and 4.7 percent in grades 8, 10, and 12, respectively. The 2014 MTF survey reports that 14 percent of 12th graders said that tranquilizers would be fairly or very easy to get if they wanted some, and identified alprazolam as the sedative-hypnotic most likely to be used by 12th graders. Other sources have estimated the prevalence of benzodiazepine use among United States high school seniors from 2007 to 2011, with the lifetime prevalence of medical use of prescription benzodiazepine anxiolytics being 4.9 percent, while the lifetime prevalence of nonmedical use was 7.5 percent. Among those high school seniors who were ever prescribed benzodiazepine anxiolytics, approximately 40.6 percent reported medical use only, the remainder using the medication in a nonprescribed manner.

A survey of college students at a US university found in 2013 that the lifetime prevalence of nonmedical use of "sleeping" and "sedative/anxiety" medication was 4.8 percent and 5.4 percent, respectively. The prevalence of

past-year nonmedical use of “sleeping” and “sedative/anxiety” medication was 2.3 percent and 3.0 percent, respectively. For medical users of these medications, 18.9 percent were approached to divert sleeping medication and 22.2 percent were approached to divert sedative/anxiety medication.

Individuals with benzodiazepine use disorder, as compared to therapeutic users of benzodiazepines, were found to differ in personality traits, including higher neuroticism and introversion, and a higher prevalence of emotional rather than task-based coping mechanisms. There were also significant correlations between the addiction and situational factors such as benzodiazepine—such as treatment circumstances and pretreatment adverse life events.

Surveys of Medical Use in Clinical Populations

A survey of unsupervised changes in dosage indicated that 12 percent of patients had decreased their anxiolytic dosage and 6 percent had increased it; 9 percent of patients taking hypnotics decreased their dosage and 8 percent increased it. To put these unauthorized increases in perspective, the same survey indicated that 13 percent of patients had unauthorized increases in antidepressant dosages, and those drugs have low abuse liability. A prospective study found that patients were more likely to use antidepressants than benzodiazepines on a long-term basis, suggesting that long-term use should not, in itself, be used to measure misuse liability. Concerns about long-term use of anxiolytics and hypnotics differ from chronic use of antidepressants because the former may be associated with mild cognitive impairment, falls, and automobile accidents, although there is not a consensus on the severity or causal relationship of these adverse effects to benzodiazepine use.

Surveys of psychiatric patients (except those diagnosed as having substance use disorders) have demonstrated high rates of benzodiazepine prescription but almost uniformly low rates of misuse. In one study of 2,719 outpatients, none of the 178 patients who had received benzodiazepines was diagnosed with a sedative-hypnotic use disorder. Studies of psychiatric inpatients suggest that the rate of benzodiazepine use disorders ranges from 0.4 to 13.0 percent of admissions. A study from the University of Munich found that 6.7 percent of 9,408 admissions had a diagnosis of benzodiazepine abuse or dependence, approximately half of whom were dependent on benzodiazepines alone. Lorazepam was the most commonly abused benzodiazepine, and oxazepam was the least commonly abused benzodiazepine, although the latter drug was the most commonly prescribed benzodiazepine in the country. Italian studies indicated that long-term prescription of benzodiazepine is common in psychiatric patients.

Surveys of Special Clinical Populations

Substantial evidence indicates that patients who abuse benzodiazepines frequently have a history of misusing other drugs. Results from the TEDS, which compiles national information about individuals admitted to publicly funded programs, indicate that benzodiazepines are rarely the primary drug of abuse. The number of substance abuse treatment admissions reporting both benzodiazepine and narcotic pain reliever misuse increased 569.7 percent from 5,032 admissions in 2000 to 33,701 admissions in 2010, while the number of all other admissions decreased by 9.6 percent during the same period. In the month prior to treatment admission, 57.1 percent of benzodiazepine and opioid pain reliever combination admissions reported daily use of narcotic pain relievers and 45.5 percent reported daily use of benzodiazepines. Nearly half (45.7 percent) of benzodiazepine and opioid pain reliever combination admissions reported a co-occurring psychiatric disorder compared with slightly more than one quarter (27.8 percent) of other admissions.

The concurrent use of opioids and benzodiazepines is ubiquitous around the world. However, the reasons for the concurrent use of these medications are not entirely clear. While in some cases, the possibility is that opioid users are using benzodiazepines therapeutically to self-medicate anxiety, mania or insomnia, a systematic review of the available literature from 1970 to 2012 suggests that benzodiazepine use is primarily for the purpose of achieving an intoxicating state. Benzodiazepine use is associated with elevated levels of harm among heroin users. However, baseline benzodiazepine status does not equate to poor long-term outcome among heroin users, although reductions in benzodiazepine use are associated with reductions in harm.

A retrospective analysis of buprenorphine and methadone cases from 2002 to 2010 reported to the American Association of Poison Control Centers' National Poison Data System (NPDS) found that there were 692 methadone–benzodiazepine cases and 72 buprenorphine–benzodiazepine cases reported. Clinical effects reported for methadone–benzodiazepine and buprenorphine-benzodiazepine interactions were lethargy (71.1 percent, 59.7 percent), respiratory depression (29.0 percent, 15.3 percent), coma (22.4 percent, 5.6 percent), respiratory arrest (4.5 percent, 0), hypotension (11.8 percent, 2.8 percent) and cardiac arrest (1.9 percent, 0), respectively. Patients in the methadone–benzodiazepine group were four-times more likely to receive naloxone (60.4 percent vs. 15.3 percent) or be intubated (16.3 percent vs. 4.2 percent) than in the buprenorphine–benzodiazepine group. There were no buprenorphine–benzodiazepine deaths, while exposure to methadone–benzodiazepine resulted in 16 deaths. These data suggest that nonmedical use of benzodiazepines with methadone is associated with higher morbidity and mortality as compared to nonmedical use of benzodiazepines with buprenorphine.

Benzodiazepine use disorder is a serious clinical problem in the

methadone maintenance treatment population. Estimates of the prevalence of benzodiazepine use disorder among methadone maintenance treatment patients range from 21 percent to 66 percent, with roughly half of patients starting benzodiazepine use after entering methadone maintenance treatment. Methadone maintenance treatment patients with a benzodiazepine use disorder are more likely to be using other drugs, including heroin, engage in high risk behaviors, and have increased rates of depression, anxiety, and global psychopathology. Women in methadone maintenance treatment who misuse benzodiazepines during pregnancy have babies of significantly lower birth weight than those women who do not. Methadone maintenance treatment patients with benzodiazepine use disorders have an eightfold greater likelihood of death compared to other methadone maintenance treatment patients and benzodiazepines are also likely to contribute to deaths from methadone toxicity in patients receiving methadone maintenance treatment. In summary, benzodiazepine use disorders in methadone maintenance treatment programs are widespread and are associated with a negative impact on treatment outcomes. Patients taking methadone presumably use benzodiazepines to boost the effects of methadone, but some studies have suggested that self-medication is also an important motivating factor. Individuals with opioid use disorder who use benzodiazepines are more likely to engage in other high-risk behaviors, such as needle sharing.

Cocaine users appear less likely to use benzodiazepines than patients in methadone clinics, with alcohol and opioids being the preferred secondary drugs of abuse. Benzodiazepines are used to attenuate the cocaine crash or to alleviate the anxiety associated with cocaine use. These patients generally use therapeutic doses of benzodiazepines on an intermittent basis.

Although the previously cited HARP study found that long-term therapeutic use of benzodiazepines did not lead to dose escalation, caution should be exercised when prescribing benzodiazepines to persons with a history of alcohol use disorder. The majority of alcohol use disorder patients often present for such treatment without revealing the extent of their alcohol use to the prescribing physician. The benzodiazepines do, however, offer added risks because of the development of tolerance to the psychomotor and sedative effects, dependence after an intermediate term of use, a withdrawal syndrome, and the potential for cognitive changes, especially among the elderly. When patients with a history of alcohol use problems or anxiety disorders present for treatment, medical professionals must exercise clinical wisdom without unnecessarily withholding potentially effective and relatively safe medication. Because benzodiazepines appear to be more rapidly effective than antidepressants for the treatment of anxiety disorders, they are often the preferred initial pharmacological treatment. There is substantial evidence that patients with

panic symptoms, especially those with a lifetime history of substance abuse, have suicide potential. With these patients, rapid symptom reduction is an important goal.

Between 16 and 25 percent of patients presenting for anxiety disorder concurrently drink alcohol in an unhealthy manner. Alcoholic women may have higher rates of concurrent use of benzodiazepines than alcoholic men. How many alcoholic patients misuse benzodiazepines is not entirely clear, but one study found that 15 percent of alcoholic outpatients who also used sedatives (mostly benzodiazepines) used them as prescribed, 24 percent used them appropriately and misused them, and 61 percent misused them. Another study reported that a little more than half of alcoholic outpatients with positive urine drug screens for benzodiazepines were misusing them. These data should be regarded cautiously, because treatment population surveys are biased to select only patients for whom previous therapy has failed.

It has been estimated that approximately 15 to 20 percent of all alcoholic persons presenting for treatment may be misusing benzodiazepines. Benzodiazepines are used to self-medicate withdrawal symptoms or anxiety disorders, to produce euphoria, and to enhance the effects of ethanol. Other groups may also be at high risk of benzodiazepine abuse. People with disabilities appear to misuse tranquilizers at as much as twice the rate of people without disabilities.

Another group that may be at higher risk of benzodiazepine toxicity is elderly patients. The overall prevalence of substance use disorders among geriatric psychiatry outpatients has been estimated to be approximately 20 percent. In the same study, 11 percent of patients were found to be dependent on benzodiazepines; a significantly higher number than other studies have previously reported, which may be attributed to methodological differences among the different studies. Although it is unknown whether long-term administration of benzodiazepines accelerates cognitive decline, results of some studies suggest that long-term use of benzodiazepines is a risk factor for increased cognitive decline and increased risk of falls among the elderly. A survey of patients between 61 and 95 years of age who were treated with benzodiazepines in a primary care setting found that these older patients underestimated adverse effects, resisted taper or discontinuation, and believed that there were no alternative therapies that would help them. A case-control study matched 1,796 individuals recently diagnosed with Alzheimer disease with 7,184 controls and found that any benzodiazepine use was associated with an increased risk of Alzheimer disease. Adjustment for anxiety, depression, and insomnia did not change the finding. No association was found for a cumulative prescribed daily doses, but the strength of association increased with exposure density. While possible that benzodiazepine use may be an early marker of a condition associated with an increased risk of dementia,

the stronger association observed for long-term exposures raises the suspicion of a possible direct association.

The Drug Abuse Warning Network (DAWN) was a public health surveillance system that monitored drug-related emergency department visits in the United States. An analysis of patient visits collected by DAWN between 2005 and 2011 (943,032 visits) revealed that combining benzodiazepines with opioid pain relievers or alcohol significantly increases the risk of a more serious emergency department visit outcome. Older patients experienced increased risk of a more serious outcome, even for visits involving benzodiazepines alone (Fig. 11.10–3).

DIAGNOSIS AND CLINICAL FEATURES

Diagnosis

DSM-5 lists a number of sedative-, hypnotic-, or anxiolytic-related disorders, but contains specific diagnostic criteria only for sedative, hypnotic, or anxiolytic intoxication and sedative, hypnotic, or anxiolytic withdrawal. Other sedative-, hypnotic-, or anxiolytic-related disorders have their diagnostic criteria outlined in the DSM-5 sections that are specific for the major symptom (e.g., sedative-, hypnotic-, or anxiolytic-induced psychotic disorder).

Substance Use Disorder. Sedative, hypnotic, or anxiolytic use disorders are diagnosed according to the general criteria in DSM-5 for substance use disorders.

Intoxication. DSM-5 contains a single set of diagnostic criteria for intoxication by any sedative, hypnotic, or anxiolytic substance. Although the intoxication syndromes induced by all those drugs are similar, subtle clinical differences are observable, especially with intoxications that involve low doses. The diagnosis of intoxication by one of that class of substances is best confirmed by obtaining a blood sample for substance screening.

Withdrawal. DSM-5 contains a single set of diagnostic criteria for withdrawal from any sedative, hypnotic, or anxiolytic substance. The clinician can specify *with perceptual disturbances* if illusions, altered perceptions, or hallucinations are present but are accompanied by intact reality testing. Benzodiazepines are associated with a withdrawal syndrome, and withdrawal from barbiturates can be life threatening. Withdrawal from benzodiazepines can also result in serious medical complications, such as seizures.

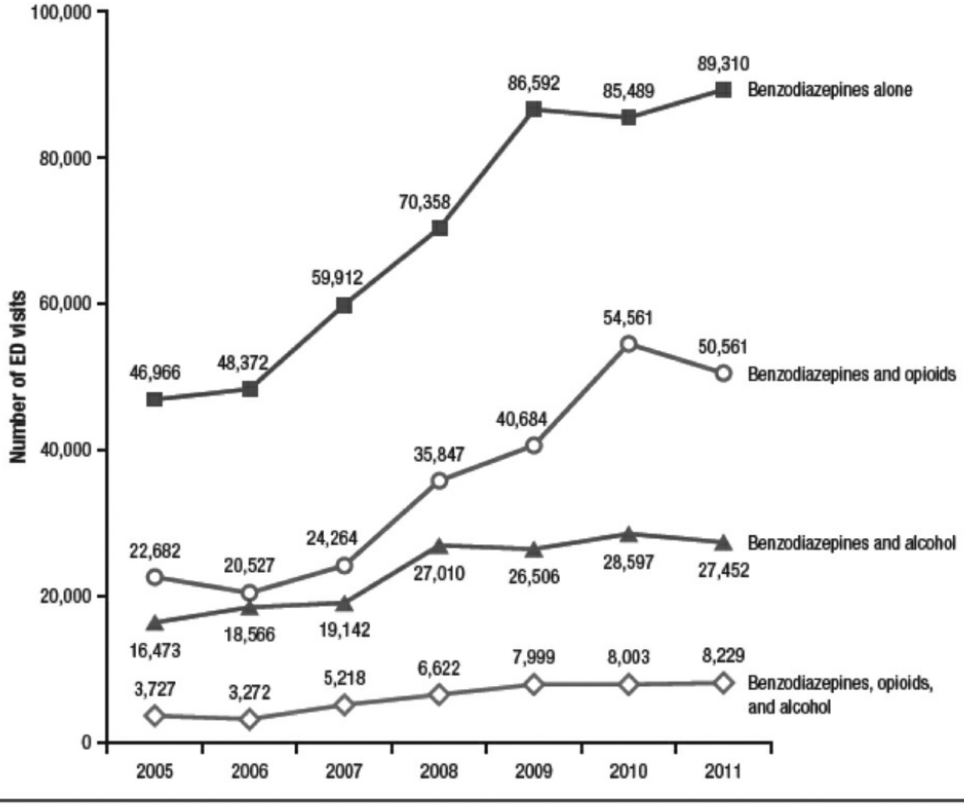

* No other drugs were involved.
Source: 2005 to 2011 SAMHSA Drug Abuse Warning Network (DAWN).

FIGURE 11.10–3. Estimated number of emergency department (ED) visits involving benzodiazepines alone or in combination with opioids or alcohol,* by year and drug combination (patients aged 12 and older): 2005 to 2011. (From Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. (December 18, 2014). The DAWN Report: Benzodiazepines in Combination with Opioid Pain Relievers or Alcohol: Greater Risk of More Serious ED Visit Outcomes.)

Delirium. DSM-5 allows diagnosis of sedative, hypnotic, or anxiolytic intoxication delirium and sedative, hypnotic, or anxiolytic withdrawal delirium. Delirium that is indistinguishable from delirium tremens associated with alcohol withdrawal is more commonly seen with barbiturate withdrawal than with benzodiazepine withdrawal. Delirium associated with intoxication can be seen with barbiturates or benzodiazepines if the dosages are high enough.

Neurocognitive Disorders. DSM-5 allows diagnosis of sedative-, hypnotic-, or anxiolytic-induced, neurocognitive disorders. The diagnosis of the disorder is complex because it is difficult to know whether a neurocognitive disorder is due to the substance use itself or to associated features of the substance use.

Psychotic Disorders. The psychotic symptoms of barbiturate withdrawal can be indistinguishable from those of alcohol withdrawal delirium. Agitation, delusions, and hallucinations are usually visual, but

sometimes tactile or auditory features develop after approximately 1 week of abstinence. Psychotic symptoms associated with intoxication or withdrawal are much more common with barbiturates than with benzodiazepines and are diagnosed as sedative-, hypnotic-, or anxiolytic-induced psychotic disorders. The clinician can further specify whether delusions or hallucinations are the predominant symptoms.

Other Disorders. Sedative, hypnotic, and anxiolytic use has also been associated with mood disorders, anxiety disorders, sleep disorders, and sexual dysfunctions. When none of those diagnostic categories is appropriate for a person with a sedative-, hypnotic-, or anxiolytic-use disorder, the appropriate diagnosis is sedative-, hypnotic-, or anxiolytic-related disorder not otherwise specified (NOS).

Clinical Features and Patterns of Abuse

Sedative, hypnotic, and anxiolytic drugs can nearly all be taken orally, occasionally to achieve a time-limited specific effect or regularly to obtain a constant, usually mild, intoxication. The physical dangers of injection include transmission of the HIV, cellulitis, vascular complications from accidental injection into an artery, infections, and allergic reactions to contaminants. IV use is associated with rapid and profound tolerance, dependence, and a severe withdrawal syndrome. One survey reported that 28 percent of heroin addicts injected benzodiazepines, and these patients had higher rates of polydrug use, HIV risk-taking behavior, criminal activity, psychological distress, health problems, and overdose than addicts who did not inject benzodiazepines.

Benzodiazepines

Since the introduction of chlordiazepoxide in 1960, benzodiazepines have become the primary drugs used to treat anxiety and insomnia, largely replacing barbiturates and other sedative-hypnotic agents. Zolpidem, an imidazopyridine hypnotic drug, is chemically distinct from the benzodiazepines but has similar clinical effects and acts at the $GABA_{A1}$ benzodiazepine receptor complex. At present it has the largest market share of hypnotic prescriptions in the United States; other hypnotics acting at the $GABA_{A1}$ receptor, such as zaleplon and eszopiclone, also have effects very much like the classical benzodiazepines. The benzodiazepines have lower abuse liability than most barbiturates, pose a much lower risk when taken in overdose, and have fewer interactions with other drugs. The advantages of the benzodiazepines must be weighed against the risk of abuse and physiological dependence.

Intoxication. Benzodiazepine intoxication can be associated with

behavioral disinhibition, potentially resulting in hostile or aggressive behavior. The effect is perhaps most common when benzodiazepines are taken in combination with alcohol. Benzodiazepine intoxication is associated with less respiratory depression than barbiturate intoxication.

Withdrawal Syndrome

SIGNS AND SYMPTOMS. Studies in the early 1960s by Leo Hollister established that abrupt discontinuation of high doses of chlordiazepoxide or diazepam could lead to a withdrawal syndrome. More recent studies show that therapeutic doses given for weeks to months may also be associated with withdrawal syndrome.

The American Psychiatric Association's Task Force Report on Benzodiazepine Dependence, Toxicity, and Abuse defined *withdrawal* as a true abstinence syndrome consisting of "new signs and symptoms and worsening of preexisting symptoms following drug discontinuance that were not part of the disorder for which the drugs were originally prescribed." Many authorities have taken issue with that definition of withdrawal and prefer to distinguish withdrawal only from recurrence, not from rebound symptoms, viewing the true abstinence syndrome as consisting of rebound symptoms plus new signs and symptoms.

The signs and symptoms of the benzodiazepine discontinuation syndrome (Table 11.10–1) have been classified as major or minor, like those of the alcohol withdrawal syndrome. According to that classification, minor symptoms include anxiety, insomnia, and nightmares. Major symptoms (which are extremely rare) include grand mal seizures, psychosis, hyperpyrexia, and death.

The discontinuance syndrome may also be divided into symptoms of rebound, recurrence, and withdrawal. *Rebound symptoms* are those for which the benzodiazepine was originally prescribed that return in a more severe form than they had before treatment. They have a rapid onset after termination of therapy and a brief duration. *Recurrence* refers to return of the original symptoms at or below their original intensity. The pattern and course of these symptoms reflect the anxiety disorder for which treatment was originally instituted.

Discontinuation symptoms are described with similar terms for benzodiazepines and SSRIs. Withdrawal reactions to SSRIs appear to be similar to those for benzodiazepines; referring to these reactions as part of a dependence syndrome in the case of benzodiazepines, but not SSRIs, may not be rational.

Table 11.10–1.
Signs and Symptoms of the Benzodiazepine Discontinuation Syndrome

The following signs and symptoms may be seen when benzodiazepine therapy is discontinued; they

reflect the return of the original anxiety symptoms (recurrence), worsening of the original anxiety symptoms (rebound), or emergence of new symptoms (true withdrawal):

- *Disturbances of mood and cognition*
  Anxiety, apprehension, dysphoria, pessimism, irritability, obsessive rumination, and paranoid ideation
- *Disturbances of sleep*
  Insomnia, altered sleep–wake cycle, and daytime drowsiness
- *Physical signs and symptoms*
  Tachycardia, elevated blood pressure, hyperreflexia, muscle tension, agitation/motor restlessness, tremor, myoclonus, muscle and joint pain, nausea, coryza, diaphoresis, ataxia, tinnitus, and grand mal seizures
- *Perceptual disturbances*
  Hyperacusis, depersonalization, blurred vision, illusions, and hallucinations

Withdrawal symptoms (Table 11.10–1) are loosely categorized into four types: (1) disturbances of mood and cognition, (2) disturbances of sleep, (3) physical signs and symptoms, and (4) perceptual disturbances. Mood and cognitive symptoms are anxiety, apprehension, dysphoria, irritability, obsessive ruminations, and paranoia. Sleep disturbances include insomnia, altered sleep–wake cycle, and daytime drowsiness. Somatic symptoms are agitation, tachycardia, palpitations, motor restlessness, muscle tension, tremor myoclonus, nausea, coryza, diaphoresis, lethargy, muscle and joint pain, hyperreflexia, ataxia, tinnitus, and seizures. Perceptual disturbances include hyperacusis, depersonalization, blurred vision, illusions, and hallucinations.

The temporal sequence of symptom development is not well established, but, on the abrupt cessation of benzodiazepines with short elimination half-lives, symptoms may appear within 24 hours and peak at 48 hours. Symptoms arising from abrupt discontinuance of benzodiazepines with long half-lives may not peak until 2 weeks later. Although some investigators suggest that a subgroup of patients have withdrawal syndromes that last for many months, no medical or scientific evidence validates the existence of such a syndrome. Prolonged symptoms are almost certainly attributable to the recurrence of the original anxiety or progression of the anxiety disorder itself.

RISK FACTORS. Factors influencing the development of the discontinuance or withdrawal syndrome are listed in Table 11.10–2. Studies on humans and animals have attempted to identify risk factors for the development of physiological dependence. Animal studies demonstrate that the severity of withdrawal is greater with higher doses and longer periods of drug administration, but that has not been as consistently demonstrated in clinical populations. When patients are stratified into high- and low-dose groups, the withdrawal syndrome is seen to be more severe in the high-dose group. However, two recent studies, one using a daily 21-mg dose of diazepam equivalent as the cutoff separating high- and low-dose groups

and the other using three groups based on diazepam daily dose (less than 6 mg, between 6 and 10 mg, and more than 10 mg), indicate no significant differences in withdrawal symptoms between groups. One possible explanation for the failure to demonstrate a consistent relation between dose and withdrawal symptoms is that subjects in clinical studies are being treated for an anxiety disorder, and, when the benzodiazepine is discontinued, the re-emergence of the original anxiety symptoms may be confused with withdrawal symptoms. Furthermore, the drug dosage is probably higher in patients with the most severe disease, complicating any relationship between dosage and the intensity of the discontinuance syndrome.

Table 11.10–2.
Factors Influencing the Development of the Benzodiazepine Discontinuation Syndrome

Dosage of benzodiazepine
Duration of benzodiazepine treatment
Rate of drug taper
Psychopathology

According to one study, withdrawal symptoms are related to treatment duration when the length of the treatment is less than 8 months but not when it is 1 year or longer. The findings of another group indicate that patients taking benzodiazepines for more than 5 years have more withdrawal symptoms than those taking them for less than 5 years.

Mild withdrawal symptoms may occur with abrupt discontinuation of therapeutic doses after 4 weeks of benzodiazepine treatment. There is a risk of rebound insomnia after a few days to 1 week of treatment with benzodiazepine hypnotic drugs with short elimination half-lives. Benzodiazepine hypnotics with longer elimination half-lives are less likely to induce rebound insomnia on abrupt discontinuation. The likelihood of a serious withdrawal syndrome increases as treatment continues, and many authorities see 4 months of treatment at therapeutic doses as a critical point in the development of clinically significant physiological dependence. That does not imply, however, that 4 months is an upper limit for treatment duration.

The risk of withdrawal is also influenced by the rate at which the drugs are discontinued. Gradual tapering of the benzodiazepines with short elimination half-lives is associated with fewer withdrawal symptoms than their abrupt discontinuation. Clinical research shows less distinct differences between gradual and abrupt discontinuation of benzodiazepines with long half-lives, because these drugs have a self-tapering action or because the period of tapering and observation in many studies is insufficient.

If benzodiazepines are abruptly discontinued, the withdrawal syndrome related to short half-life agents appears earlier and may be more intense than that with long half-life drugs. Differences in intensity have not been proved because most studies do not monitor withdrawal symptoms for 2 weeks, which may be the time of peak symptoms with some drugs with long half-lives. Furthermore, any difference in withdrawal severity between short and long half-life agents is not entirely supported by animal studies. There is no difference in symptom severity of the withdrawal syndrome after animals are treated with midazolam or chlordiazepoxide, for example, because essentially all benzodiazepines, regardless of their pharmacokinetic properties in humans, are eliminated rapidly in small animal species.

The development of physiological dependence differs depending on whether it results from benzodiazepine partial agonists or full agonists. In animal models, the partial agonists bretazenil and abecarnil are associated with fewer withdrawal symptoms than are classic agonists such as diazepam or clorazepate (the prodrug for desmethyldiazepam). Clinical studies in anxious patients, however, indicate that abrupt termination of abecarnil is associated with withdrawal symptoms in some patients.

Eszopiclone has demonstrated only mild withdrawal symptoms with discontinuation of therapeutic doses after 6 months of treatment. Supratherapeutic doses, however, lead to an abstinence syndrome that is indistinguishable from that produced by classical benzodiazepines. Clinical experience with eszopiclone, zolpidem, and zaleplon is limited compared to benzodiazepine hypnotics, although many clinicians feel that rebound insomnia, upon discontinuation, may be mild and of a shorter duration then with classical benzodiazepines, however, sufficient studies are lacking.

Some studies in animals indicate that periodic administration of the benzodiazepine antagonist flumazenil during the chronic administration of lorazepam, diazepam, triazolam, or clobazam may attenuate the withdrawal syndrome.

Personality traits may be a risk factor for the development of the benzodiazepine withdrawal syndrome. Withdrawal severity is greater in patients with higher scores on the dependence scale of the Minnesota Multiphasic Personality Inventory 2 (MMPI-2), high prewithdrawal levels of anxiety and depression, lower educational level, and passive-dependent personality disorder.

Overdose. The benzodiazepines, in contrast to the barbiturates and the barbiturate-like substances, have a large margin of safety when taken in overdoses, a feature that contributed significantly to their rapid acceptance. The ratio of lethal to effective dose is approximately 200 to 1 or higher because of the minimal respiratory depression associated with the benzodiazepines. Even when grossly excessive amounts (more than 2 g) are

taken in suicide attempts, the symptoms include only drowsiness, lethargy, ataxia, some confusion, and mild depression of the user's vital signs. A much more serious condition prevails when benzodiazepines are taken in overdose in combination with other sedative-hypnotic substances, such as alcohol. In such cases, small doses of benzodiazepines can cause death. The availability of flumazenil, a specific benzodiazepine antagonist, has reduced the lethality of the benzodiazepines, because flumazenil can be used in emergency rooms to reverse the effects of the benzodiazepines.

Barbiturates

Since the advent of the benzodiazepines, barbiturate use has been limited in modern medicine. Phenobarbital is still prescribed as an anticonvulsant and as a sedative, especially for children. It is also a common component of many combination products and reduces the stimulating effects of sympathomimetic agents. Butalbital is an intermediate-acting barbiturate found in a widely used combination product that also contains acetaminophen and caffeine (Fioricet) and is approved for the treatment of muscle contraction headaches. Given the limited availability of barbiturates, it is uncommon, at least in the United States, for addiction units to treat patients dependent on barbiturates other than butalbital and phenobarbital. The barbiturates marketed currently in the United States are listed in Table 11.10–3.

A major disadvantage of the use of the barbiturates is the development of pharmacokinetic and pharmacodynamic tolerance. *Tolerance* is defined as reduced drug response as a result of decreased drug concentration at the site of action, usually the result of increased drug metabolism (pharmacokinetic) or of cellular adaptive changes with unchanged or higher drug concentrations at the site of action (pharmacodynamic). Pharmacodynamic tolerance begins after acute doses and continues to develop over weeks to months. Tolerance to the mood-altering and sedative effects develops to a greater extent than does tolerance to the lethal effects, increasing the risk of accidental overdose.

Intoxication. When barbiturates and barbiturate-like substances are taken in relatively low doses, the clinical syndrome of intoxication is indistinguishable from that associated with alcohol intoxication. The symptoms include sluggishness, incoordination, difficulty in thinking, poor memory, slowness of speech and comprehension, faulty judgment, disinhibition of sexual and aggressive impulses, a narrowed range of attention, emotional lability, and exaggeration of basic personality traits. The sluggishness usually resolves after a few hours, but the impaired judgment, distorted mood, and impaired motor skills may remain for 12 to 24 hours, depending primarily on the half-life of the abused substance. Other potential symptoms are hostility, argumentativeness, moroseness,

and occasionally, paranoid, and suicidal ideation. The neurological effects include nystagmus, diplopia, strabismus, ataxic gait, positive Romberg sign, hypotonia, and decreased superficial reflexes.

Table 11.10–3.
Barbiturates and Other Sedative-Hypnotic Agents

Generic Name	Trade Name	Dose (mg)
Amobarbital	Amytal	100
Aprobarbital	Alurate	40
Butabarbital	Butisol	100
Butalbital	Many combination products (e.g., Fioricet)	100
Pentobarbital	Nembutal	100
Secobarbital	Seconal	100
Chloral hydrate	Aquachloral Supprettes	500
Ethchlorvynol	Placidyl	500
Glutethimide	Doriden	250
Meprobamate	Miltown	400

Note: The substitution technique for sedative-hypnotic withdrawal requires calculation of equivalent doses of phenobarbital to replace the sedative–hypnotic agent that the patient is taking. The above are doses of various sedative-hypnotic agents for which a 30-mg dose of phenobarbital should provide adequate coverage of a withdrawal syndrome. Daily doses of phenobarbital should rarely exceed 600 mg using this protocol.

Dependence and Withdrawal. Physiological dependence may develop after a daily dose of 400 mg of pentobarbital for 3 months; abrupt discontinuation results in paroxysmal abnormalities on the EEG in approximately 30 percent of patients. At a daily dose of 600 mg of pentobarbital for 1 to 2 months, a withdrawal syndrome characterized by anxiety, insomnia, anorexia, tremor, and EEG changes occurs in approximately half of patients, and 10 percent may have a single seizure. At higher dosages of 800 to 2,200 mg per day for several weeks to months, abrupt discontinuation leads to minor symptoms, which include apprehension and uneasiness, insomnia, muscular weakness, twitches, coarse tremors, myoclonic jerks, postural faintness and orthostatic hypotension, anorexia, vomiting, and EEG changes, within 24 hours of the last dose. Minor symptoms may persist as long as 2 weeks. At high doses, major symptoms develop. As many as 75 percent of patients may have grand mal seizures on the second or third day after withdrawal, and two-thirds have more than one seizure. The interictal EEG shows four spike-wave discharges per second. Two-thirds of these patients develop delirium between the third and eighth days of withdrawal, which is sometimes accompanied by hypothermia, which may be fatal. Disorientation, visual

hallucinations, and frightening dreams may precede the onset of full delirium. Delirium may be exceedingly difficult to reverse, even with large doses of a barbiturate; thus, clinicians should never wait for the appearance of withdrawal symptoms before instituting therapy. The duration of the withdrawal syndrome is between 3 and 14 days; most end by the day 8.

Overdose. Barbiturates are lethal when taken in overdose because they induce respiratory depression. In addition to intentional suicide attempts, accidental or unintentional overdoses are common. Barbiturates in home medicine cabinets are a common cause of fatal drug overdoses in children. As with benzodiazepines, the lethal effects of barbiturates are additive to those of other sedative-hypnotic drugs, including alcohol and benzodiazepines. Barbiturate overdose is characterized by induction of coma, respiratory arrest, cardiovascular failure, and death.

The lethal dose varies with the route of administration and the degree of tolerance for the substance after a history of long-term abuse. For the most commonly abused barbiturates, the ratio of lethal to effective dose ranges between 3 to 1 and 30 to 1. Dependent users often take an average daily dose of 1.5 g of a short-acting barbiturate, and some have been reported to take as much as 2.5 g per day for months. The lethal dose is not much greater for the long-term abuser than it is for the neophyte. Tolerance develops quickly to the point at which withdrawal in a hospital becomes necessary to prevent accidental death from overdose.

Miscellaneous Sedative-Hypnotic Drugs

The older sedative-hypnotic drugs are similar in effect to the barbiturates, and they lead to tolerance and physiological dependence. The withdrawal syndrome is similar to the barbiturate withdrawal syndrome, and the protocols described for barbiturates should be used for safe withdrawal of patients dependent on those drugs.

Meprobamate and Carisoprodol (Soma). Meprobamate, a carbamate derivative, has weak efficacy as an antianxiety agent. At clinical doses in humans, it has minimal muscle-relaxant effects, but it may have a mild analgesic effect in musculoskeletal pain and may potentiate analgesics. Typical daily doses are between 1,200 and 1,600 mg, with a maximum of 2,400 mg, in three or four divided daily doses. A discontinuance syndrome occurs after several weeks of treatment with a daily dose of 2,400 mg, and mild symptoms may be seen with long-term therapy at doses of 1,600 mg daily. Onset of the discontinuance syndrome occurs within 12 to 48 hours after abrupt discontinuation and lasts for an additional 12 to 48 hours. Seizures may be common in withdrawal from meprobamate, and reports published in the 1970s suggest that serious withdrawal symptoms are more common with meprobamate than with barbiturates. Fatalities from

overdose have been reported after doses as low as 12 g. The manufacturer offers the following guidelines for blood concentration: 5 to 20 μg/mL is the normal range with recommended doses; 30 to 100 μg/mL is associated with mild-to-moderate overdose, with patients in stupor or light coma; 100 to 200 μg/mL is associated with serious overdose, deeper coma, and fatalities; concentrations greater than 200 μg/mL are associated with more fatalities than survivals. Meprobamate induces microsomal enzymes and may exacerbate intermittent porphyria.

Carisoprodol is a commonly prescribed muscle relaxant that is metabolized to meprobamate, hydroxyl meprobamate, and hydroxyl carisoprodol and has been associated with abuse, dependence, and impairment of driving ability.

Chloral Hydrate. Chloral hydrate is used as a hypnotic in doses of .5 to 2.0 g. After oral administration, it is rapidly transformed to trichloroethanol, the metabolite responsible for its pharmacological activity. It is also effective in blocking experimentally induced seizures; however, anticonvulsive and hypnotic doses are similar in humans, making the benzodiazepines and barbiturates better choices for anticonvulsive medications.

At high doses, chloral hydrate may produce respiratory depression, hypotension, gastric necrosis, depressed cardiac contractility, and a shortened refractory period. Even at therapeutic doses, it is associated with gastric distress and flatulence. Somnambulism, disorientation, and paranoid ideation may occur. Tolerance and physiological dependence develop with long-term use, and the withdrawal syndrome is similar to the barbiturate abstinence syndrome. Several drug interactions occur with chloral hydrate, including displacement of oral anticoagulants and acidic drugs from protein binding sites by the metabolite trichloroacetic acid. When combined with furosemide (Lasix), flushing, tachycardia, hypotension or hypertension, and diaphoresis have been reported. It potentiates the effects of ethanol, and it should not be used in patients with intermittent porphyria. Fatalities from overdose may occur with as little as 4 g, although patients have survived after taking 30 g.

Ethchlorvynol. Ethchlorvynol is a rapidly acting sedative-hypnotic agent with anticonvulsant and muscle-relaxant properties. Recommended doses are 500 to 1,000 mg. Initial effects (which may include euphoria) occur 15 to 30 minutes after an oral dose, and plasma levels peak at 1 to 1.5 hours. The elimination half-life of the parent compound is 10 to 20 hours. The duration of the hypnotic effect is approximately 5 hours. Long-term use of ethchlorvynol has been associated with toxic amblyopia, scotoma, nystagmus, and peripheral neuropathy, which are usually reversible on drug discontinuation. IV self-administration may cause pulmonary edema.

Idiosyncratic reactions of excitement and stimulation have been described. Hypersensitivity reactions with urticaria and rare fatal thrombocytopenia may also arise. It should not be used in patients with intermittent porphyria. Tolerance, physiological dependence, and a withdrawal syndrome are seen with long-term administration. Lethal doses are typically between 10 and 25 g; one person was reported to have died after taking 2.5 g with ethanol, and one patient survived but was in a coma for 7 days after taking 50 g.

Glutethimide and Methyprylon. Glutethimide (Doriden) and methyprylon (Noludar) (no longer marketed in the United States) are piperidinedione sedative-hypnotic drugs that have high liability for abuse. Glutethimide resembles the barbiturates in most respects but differs from them in having significant anticholinergic activity. An overdose can cause ileus, bladder atony, mydriasis, hyperpyrexia, myoclonic jerks, and convulsions. A lethal dose is between 10 and 20 g; intoxication may be seen at a dose of 5 g. An unusual characteristic of glutethimide is that a withdrawal syndrome can occur in patients taking therapeutic doses (0.5 to 3.0 g daily), even when they have not stopped taking the drug. The abstinence syndrome includes tremulousness, nausea, tachycardia, fever, tonic muscle spasms, and grand mal seizures. Withdrawal from a combination of glutethimide and antihistamine has caused catatonia and dyskinesia. When glutethimide is taken with codeine (the street name is a *load*), a euphoria similar to that with heroin results. In these patients, detoxification may require opioid withdrawal and sedative withdrawal.

Methaqualone. Methaqualone is no longer available in the United States, and much of what is sold on the street as methaqualone is actually diazepam; however, it remains a major public health problem in South Africa and India. Methaqualone is a quinazoline with sedative, anticonvulsive, local anesthetic, and antispasmodic activity. It has antitussive effects comparable to those of codeine and weak antihistaminic activity. During the 1960s and 1970s, those in the drug culture considered it to provide the ultimate high of all sedative-hypnotic agents, producing euphoria comparable to that from heroin. It was usually taken orally, often with alcohol, which was referred to as "luuding out." Chronic users may take daily oral doses between 75 mg and 2 g, with an average daily dose of 775 mg. Onset of effects after oral dosing occurs after 30 minutes and lasts 2 to 4 hours. Although oral dosing was most common in the United States and Europe, methaqualone is smoked in South Africa and India, frequently used in a preparation of crushed methaqualone tablets, cannabis, and tobacco, referred to as "white pipe." Onset of euphoria is rapid by the smoked route and is followed by a relaxed state lasting up to 6 hours. Adverse effects include peripheral neuropathy, nightmares,

somnambulism, gastric discomfort, and urticaria. Death has been reported after 8 g, but one report attributed most deaths under the influence of methaqualone to accidents due to impairment produced by the drug. Tolerance and physiological dependence are seen with chronic dosing.

An overdose of methaqualone may result in restlessness, delirium, hypertonia, muscle spasms, convulsions, and (in high doses) death. Unlike barbiturates, methaqualone rarely causes severe cardiovascular or respiratory depression, and most fatalities result from combining methaqualone with alcohol.

Gamma Hydroxybutyrate. Gamma hydroxybutyrate (GHB) was used for decades in Europe with a limited number of reported adverse effects and abuse. However, beginning in the late 1980s, when clinical use of GHB as a dietary supplement, growth promoter, and mild sedative became more prevalent, public health officials began inquiring about its safety and efficacy. By the early 1990s, concern that GHB was developing into a popular drug of abuse and evidence that GHB was being used as a recreational club drug led to FDA removal of GHB from the market. The FDA has now approved GHB (sodium oxybate; Xyrem) as a treatment of excessive daytime sleepiness and cataplexy in patients with narcolepsy. Also, it is still possible to obtain products containing GHB and its precursors through the Internet and various health food stores.

GHB is a naturally occurring short-chained fatty acid naturally distributed throughout the CNS. It is synthesized in the brain from GABA via deamination to succinic semialdehyde (SSA) by GABA aminotransferase. Other precursors of GHB such as 1,4-butanediol and γ-butyrolactone may have important roles as well. GHB activity is terminated by synaptic reuptake and subsequent metabolism of SSA to succinate, which is further degraded to carbon dioxide and water. Although the half-life of GHB is reported as 30 to 50 minutes, its pharmacokinetics are nonlinear, so that high doses may take considerably longer to clear from the body.

The pharmacological actions of GHB are mediated through specific GHB receptors that are G-protein coupled. GHB may also bind to GABA type B ($GABA_B$) receptors, although some investigators believe that the affinity is so low that this may not be of physiological importance. Other data suggest that weak $GABA_B$ agonist activity has been demonstrated in studies that have shown that $GABA_B$ antagonists block GHB effects, and that the $GABA_B$ agonist baclofen (Lioresal) produces effects similar to GHB. Euphoric effects of GHB may be mediated by enhanced dopamine release in the striatum and mesolimbic system, although the relationship of dopaminergic effects to abuse liability is not well understood. Abrupt discontinuation of chronic high doses have been associated with anxiety,

restlessness, insomnia, nausea, lethargy, tachycardia, sweating, tremors, muscle cramps, and psychosis. However, 6-month open label clinical trials of clinical doses of sodium oxybate have not shown tolerance. The discontinuation effects of sodium oxybate have not been systematically studied in humans, although in rare cases abrupt discontinuation may be associated with anxiety and insomnia.

Determining the risk of GHB administration in humans is difficult due to insufficient laboratory testing of GHB intoxication, limited availability of data reporting blood and tissue levels of hospitalized patients, or coingestion of alcohol or other drugs in clinical cases. GHB in combination with other CNS depressants, especially ethanol, creates amplified CNS depression and has been found to result in acute intoxication, often requiring medical attention. Similar to benzodiazepines and barbiturates, when GHB and alcohol are combined, memory impairment, confusion, increased sleep, and unconsciousness are likely to occur. Other adverse effects include mild hypothermia, dizziness, nausea, vomiting, weakness, loss of peripheral vision, agitation, hallucinations, decreased respiratory effort, and coma.

TREATMENT

Withdrawal

Benzodiazepines. Guidelines for the treatment of benzodiazepine discontinuance syndrome and the clinical management of withdrawal are presented in Table 11.10–4. The most common clinical situations that require medically supervised withdrawal from the benzodiazepines involve patients with anxiety disorders or chronic insomnia who have been maintained on therapeutic doses of benzodiazepines for several months or years, patients taking supratherapeutic doses of benzodiazepines for treatment-resistant types of anxiety or because of inappropriate dosage escalation, and patients who abuse benzodiazepines as part of a mixed abuse pattern (e.g., those who use benzodiazepines to self-medicate alcohol withdrawal or to end a cocaine run).

In planning a rational withdrawal, clinicians should remember the factors that influence the development of physiological dependence. The risk of withdrawal is greatest in patients taking high doses over long periods. The onset of withdrawal and its peak severity are earlier with agents with short half-lives, and, if they are not tapered, withdrawal from them may be more severe than that from agents with long elimination half-lives, although the last point has not been definitively established. Concurrent use of other sedative-hypnotic agents (e.g., barbiturates or ethanol) may alter the time course and severity of the benzodiazepine withdrawal syndrome.

Table 11.10–4.
Guidelines for the Treatment of Benzodiazepine Discontinuance Syndrome

Evaluate and treat concomitant medical and psychiatric conditions
Obtain drug history and urine and blood samples for drug and ethanol assay
Determine required dose of benzodiazepine or barbiturate for stabilization, guided by history, clinical presentation, drug-ethanol assay, and (in some cases) challenge dose
Detoxification from supratherapeutic dosages:
Hospitalize if there are medical or psychiatric indications, poor social supports, or polysubstance dependence or if the patient is unreliable
Some clinicians recommend switching to a longer-acting benzodiazepine for withdrawal (e.g., diazepam [Valium], clonazepam [Klonopin]); others recommend stabilizing on the drug that the patient was taking or on phenobarbital
After stabilization, reduce dosage by 30% on the second or third day and evaluate the response, keeping in mind that symptoms that occur after decreases in benzodiazepines with short elimination half-lives (e.g., lorazepam [Ativan]) appear sooner than with those with longer elimination half-lives (e.g., diazepam)
Reduce dosage further by 10–25% every few days, if tolerated
Use adjunctive medications if necessary—carbamazepine (Tegretol), gabapentin (Neurontin), β-adrenergic receptor antagonists, divalproex (Depakote), clonidine (Catapres), and sedative antidepressants have been used, but their efficacy in the treatment of the benzodiazepine abstinence syndrome has not been established
Detoxification from therapeutic dosages:
Initiate 10–25% dose reduction and evaluate response
Dose, duration of therapy, and severity of anxiety influence the rate of taper and the need for adjunctive medications
Most patients taking therapeutic doses have uncomplicated discontinuation
Psychological interventions may assist patients in detoxification from benzodiazepines and in the long-term management of anxiety

TAPERING. The basic principle underlying safe withdrawal is gradual tapering of the benzodiazepine dosage. In cases of detoxification from therapeutic doses, the daily dose is known, and the tapering protocol is easily calculated. Anyone taking a benzodiazepine for 2 weeks or longer should be tapered from the drug. Initially, the dose is reduced by approximately 10 to 25 percent, and the patient is observed for any signs and symptoms of withdrawal. Withdrawal signs appear earlier for drugs with short half-lives than for drugs with long half-lives. Subsequent reductions depend on how the patient responds to initial changes and require individualization based on careful observation of the patient's condition after each dose change.

Lower doses of benzodiazepines and no previous withdrawal attempts predict successful discontinuation of benzodiazepines in patients with benzodiazepine use disorder. In individuals with benzodiazepine use disorder, clinically significant dose reductions are associated with improvements in quality of life measures. Treatment options for benzodiazepine use disorder are limited, with no approach clearly effective.

In some cases, patients taking high therapeutic doses of a benzodiazepine for a year or more may require several months or longer to stop the medication completely, not primarily for pharmacological reasons, but so they can learn alternative strategies for coping with anxiety. Most patients taking midrange therapeutic doses for a shorter time tolerate weekly reductions ranging from 10 to 25 percent. Most authorities agree that the last phases of drug discontinuation are the most difficult for the patient, and it may be necessary to slow the taper rate or to stop the taper entirely. The adjunctive use of nonbenzodiazepine medications may be useful in these circumstances.

Detoxification from supratherapeutic dosages of benzodiazepines may require a slightly different approach. Most clinicians hospitalize such patients because of the greater medical risks associated with supratherapeutic dosage withdrawal. Patients may be tapered using the benzodiazepine that they have been taking or switched to a drug with a long elimination half-life, such as diazepam or clonazepam, using the dose equivalencies listed in Table 11.10–5 to determine the initial daily dose. The estimated equivalent dose is then administered in divided doses on day 1 to ensure that an accurate history and appropriate equivalency are established. Some clinicians stabilize patients on that dose for 2 to 3 days; others prefer to reduce the dose by 30 percent on the second day, followed by 5 to 10 percent daily reductions.

Table 11.10–5.
Approximate Therapeutic Equivalent Doses of Benzodiazepines

Generic Name	Trade Name	Dose (mg)
Alprazolam	Xanax	1
Chlordiazepoxide	Librium	25
Clonazepam	Klonopin	0.5–1.0
Clorazepate	Tranxene	15
Diazepam	Valium	10
Estazolam	ProSom	1
Flurazepam	Dalmane	30
Lorazepam	Ativan	2
Oxazepam	Serax	30
Temazepam	Restoril	20
Triazolam	Halcion	.25
Quazepam	Doral	15
Zolpidem	Ambien	10
Zaleplon	Sonata	10

Research studying the management of discontinuation of therapeutic benzodiazepine treatment has generally supported the approach of a slow, gradual taper. For patients with benzodiazepine use disorder, a consensus exists that gradual taper is preferred to abrupt discontinuation. Small

clinical trials have tested propranolol, buspirone, and progesterone without any clear evidence of efficacy. Hydroxyzine can reduce anxiety symptoms associated with a benzodiazepine taper. Carbamazepine has shown promise as an adjunct to gradual taper of benzodiazepine. Trazodone and valproate were associated with improved abstinence rates, but had no effect on withdrawal symptoms. Cyamemazine, an anxiolytic antipsychotic agent, has shown promise as a treatment of the benzodiazepine withdrawal syndrome. The use of a slow infusion of the benzodiazepine antagonist flumazenil has been shown to be feasible, but carries a high risk of seizures. No clearly effective pharmacotherapy for the treatment of benzodiazepine use disorder has been identified. The use of carbamazepine (Tegretol) or divalproex (Depakote) in high-dosage withdrawal often permits even larger daily reductions in the benzodiazepine dose. The efficacy of other anticonvulsants such as levetiracetam (Keppra), gabapentin (Neurontin), and topiramate (Topamax) has not been established but are sometimes used by experienced clinicians. Pregabalin shows promise as a treatment of tapering long-term benzodiazepine use. Most high-dose detoxification can be completed in 2 weeks or less using this protocol. For some patients, the dose cannot be reduced that rapidly, however, and slower, longer tapers are sometimes necessary. The efficacy of anticonvulsants for the treatment of the benzodiazepine discontinuance syndrome has not been established, although clinical experience to date is encouraging.

Some authorities have proposed agonist substitution for high-dose benzodiazepine users using a model successful for treating opioid use disorder. While research has not yet thoroughly examined this approach, the concept of using a long-acting benzodiazepine as an agonist replacement strategy is intriguing.

When the physician knows that the patient has been taking supratherapeutic doses but cannot determine the exact dose, 20 mg of diazepam may be given to estimate tolerance. The dose should be repeated every 2 hours until mild sedation occurs. The total dose required to induce mild sedation is then considered the initial dose, and the detoxification proceeds as described previously.

If the primary drug that the patient is taking is not going to be used in the withdrawal protocol, clonazepam or lorazepam may be the best drug to use when detoxifying patients who have been taking supratherapeutic dosages of triazolobenzodiazepines (e.g., triazolam). It is not clear whether higher doses of diazepam would have alleviated the alprazolam withdrawal syndrome in the published case reports. Some authorities recommend the use of phenobarbital in such cases. Withdrawal from the nonbenzodiazepine hypnotics zolpidem, zaleplon, and eszopiclone is usually treated with benzodiazepines, although case reports describe gradual taper of the hypnotic as also being effective.

Despite the risk of benzodiazepine use problems among methadone

maintenance treatment patients, no clearly effective strategy for managing this clinical problem has been developed. As for the specific circumstance of managing methadone maintenance treatment patients with a benzodiazepine use disorder, limited research has been conducted. Potential treatment interventions that have been investigated, including benzodiazepine maintenance, contingency reinforcement, and slow outpatient detoxification have not been accepted into clinical practice. The optimal strategy for the treatment of benzodiazepine use disorders in methadone maintenance treatment patients remains unknown.

DETOXIFICATION FROM MULTIPLE DRUGS OF ABUSE. Polysubstance abuse may present special problems for safe withdrawal. The most common polysubstance abuse pattern influencing medical management of detoxification from benzodiazepines is the concurrent use of alcohol, opioids, and cocaine. Combination with alcohol alters the time course and increases the severity of the sedative withdrawal syndrome. Adequate dosages of a benzodiazepine usually suffice for safe withdrawal; however, in rare instances, even large dosages of diazepam can be ineffective, and switching to a barbiturate becomes necessary, with phenobarbital being the preferred agent.

When supratherapeutic doses of benzodiazepines are abused with large amounts of opioids, it is best to discontinue the former drug first and then stabilize the patient on methadone or buprenorphine. When the period of risk of major symptoms of sedative-hypnotic withdrawal passes, opioid withdrawal may begin. Addiction to low or moderate dosages of benzodiazepines or opioids often permits simultaneous withdrawal from both classes of drugs.

Benzodiazepines are often used to terminate a cocaine or other psychostimulant run or to decrease anxiety induced by cocaine. Benzodiazepines used intermittently in that manner usually do not require detoxification. The chronic use of high doses of benzodiazepines with cocaine is less common but does require medically supervised benzodiazepine tapering. The cocaine abstinence syndrome is characterized by depression, irritability, lethargy, lack of motivation, hypersomnolence, confusion, and drug craving. The clinical presentation of the sedative-hypnotic withdrawal syndrome may be altered by recent cocaine use and is best monitored by carefully measuring vital signs (e.g., increases in pulse, temperature, and blood pressure).

Some clinicians prefer to use phenobarbital for detoxification of patients with mixed-drug abuse. A dose of phenobarbital equivalent to the benzodiazepine and other sedatives the patient is taking is calculated (Tables 11.10–3 and 11.10–5) or, preferably, the dose is determined from a challenge test, and the patient is stabilized on that amount in divided daily doses for the period of peak risk of withdrawal. The maximum daily dose of

phenobarbital is 600 mg. After stabilization, phenobarbital dose is reduced by 30 to 60 mg every 2 or 3 days.

ADJUNCTIVE MEDICATIONS. Several adjunctive medications are used to treat the benzodiazepine withdrawal syndrome, but in no case has their efficacy been consistently supported in controlled trials nor is there evidence that any specific adjunctive medication is superior to another.

Carbamazepine, given in doses that achieve the therapeutic concentrations for treatment of seizures, may partially attenuate the withdrawal syndrome from benzodiazepines. Initial dosages are 200 mg twice a day, although some clinicians recommend a single dose of 400 mg at bedtime to take advantage of the sedative effect. If initial doses are well tolerated, carbamazepine is increased to 600 mg on the second or third day. Typical total daily doses range from 400 to 800 mg. After several days of treatment, benzodiazepine dosage is tapered. Despite reports that tapering can be completed within a week, even from supratherapeutic dosages, some patients require a slow taper that lasts several weeks. Carbamazepine can be tapered quickly 2 to 4 weeks after the last benzodiazepine dose. Carbamazepine may not be successful in withdrawing patients with panic disorder because it is ineffective in blocking panic attacks, but it has shown some success even in these patients, as long as a slow benzodiazepine taper is used. The efficacy of other anticonvulsants is undetermined, but clinical experience suggests that valproate (Depacon) may reduce withdrawal symptoms as well. Enteric-coated tablets of divalproex (Depakote) are used in starting dosages of 500 to 1,000 mg divided into two or three daily doses. Dosages are increased to achieve serum concentrations of 50 to 120 μg/mL. Some clinicians prefer to use daily loading doses of 20 mg/kg. For patients who abuse alcohol and benzodiazepines, valproate should be avoided when liver enzyme activities are three times their normal values. Gabapentin and pregabalin may also reduce benzodiazepine withdrawal symptoms, but experience is limited.

The α-adrenergic receptor antagonists may also attenuate withdrawal symptoms, although meta-analyses have challenged their efficacy. Propranolol (Inderal) in doses of 60 to 120 mg may reduce the symptoms of benzodiazepine withdrawal, although tachycardia and elevated blood pressure are the primary symptoms affected, whereas the subjective sense of discomfort and dysphoria persists. The α-adrenergic receptor antagonists do not have prophylactic effects against seizures and should not be used in monotherapy for withdrawal. Nevertheless, in some cases, propranolol or atenolol (Tenormin), 50 to 100 mg, may be helpful adjuncts to withdrawal treatment.

Clonidine (Catapres), .1 mg twice a day to .2 mg three times a day or as a patch, has been used in the treatment of low-dose benzodiazepine withdrawal with variable results. Clonidine does not provide protection

against withdrawal seizures. It may be effective only when used before withdrawal symptoms develop.

Although antidepressants have not been rigorously studied in the treatment of benzodiazepine withdrawal, many experienced clinicians prescribe them. Agents with well-established antipanic and sedative effects may be useful adjunctive medications for patients whose anxiety recurs during taper or after drug discontinuation. Imipramine (Tofranil), amitriptyline (Elavil), doxepin (Adapin, Sinequan), nefazodone (Serzone), trazodone (Desyrel), SSRIs, and MAOIs may all be useful in that role. Melatonin may be a useful treatment of insomnia during withdrawal, although clinical studies have been contradictory.

PSYCHOLOGICAL TREATMENTS. Specific psychological treatments for withdrawal from anxiolytics and hypnotics include brief advice, various types of cognitive behavioral therapy, and medically managed withdrawal. Cognitive behavioral strategies for anxiety management have been applied to patients withdrawing from benzodiazepines with mixed results. Patients who meet the criteria for generalized anxiety disorder are poor candidates for relaxation training. Cognitive behavioral therapy appears particularly effective for patients withdrawing from hypnotics, although one study found structured medical management as effective as self-help cognitive behavioral therapy (which incorporates education, planned sage reductions, engaging patients in their taper schedule, and medical support.

Many practitioners believe that successful withdrawal requires transmitting to the patient a sense of control over the withdrawal symptoms. In support of this belief, placebo substitution is associated with fewer withdrawal symptoms after drug discontinuation than is withdrawal without placebo. Apparently the belief that one is taking an effective medication is enough to reduce symptoms.

One way to foster a sense of control over withdrawal symptoms is to link the symptoms of anxiety to environmental intrapsychic stressors. Cognitive restructuring teaches patients about the withdrawal syndrome and helps them identify and relabel withdrawal symptoms as anxiety, thus permitting the implementation of adaptive coping strategies, which include systematic desensitization, in vivo graded exposure, and group problem solving. Using a diary to record mood states and their precipitants sometimes helps cognitive restructuring. A diary may also alert the clinician to maladaptive responses to benzodiazepine withdrawal, such as increased alcohol consumption.

The concept of *cognitive coping* should be introduced to patients, and examples of self-statements (talking to oneself) that are applicable to withdrawal and anxiety management should be discussed and provided as flash cards. They include such statements as “I feel nervous now, but it won’t last,” “I feel bad, but it’s not worse than when I had the flu,” or “I feel

uncomfortable, but nothing bad is going to happen to me." Self-statements should be short, simple, and relevant to the patient.

Patients are encouraged to be active participants in developing their withdrawal schedule. Clinicians should be flexible about the rate of taper and should encourage the patient to join in the withdrawal process, especially those who are withdrawing from therapeutic doses. It is often useful to request that patients choose which doses should be decreased or eliminated according to their perceived need. For some patients, setting a maximum permissible daily dose may be necessary.

LONG-TERM OUTCOME. Few adequately designed follow-up studies have been carried out on patients who terminated benzodiazepine treatment. Many studies suffer from sample selection bias, diagnostic heterogeneity of the sample, and failure to assess environmental factors or to control for concurrent psychotherapy or psychoactive medication; however, despite those shortcomings, several interacting findings have emerged.

In one study of patients who had discontinued benzodiazepines from 10 months to 3.5 years earlier, 70 percent were no longer taking benzodiazepines and had few or no symptoms, 22 percent were not taking benzodiazepines but required other psychotropic medications, and 8 percent had returned to benzodiazepine use. Another study assessed patients who were evaluated for a discontinuation program 3 to 6 years earlier and found that 45 percent were not using benzodiazepines, although some were using other psychoactive medication. Of those who successfully completed the tapering program, 73 percent were benzodiazepine free at follow-up, compared to 39 percent of patients who entered but did not complete the program, and 14 percent of patients who did not participate in the program. Another study of patients treated for benzodiazepine dependence found that 56 percent of benzodiazepine-free (no use during the preceding month) patients had moderate-to-severe anxiety, and 38 percent had moderate-to-severe depression at 1- to 5-year follow-up. A different study found that although 66 percent of patients were not taking benzodiazepines at follow-up, 75 percent had taken them at some time during the 5 years after entering a discontinuation program.

The follow-up studies suggest that most patients can successfully discontinue taking benzodiazepines. They also indicate that some patients experience recurrent anxiety that requires alternative psychoactive medication or reinstitution of benzodiazepine therapy. However, a subgroup of patients who discontinue benzodiazepines actually shows improvement in anxiety over pretapering levels, which suggests that those who continue to have significant symptoms of anxiety or depression while taking a benzodiazepine should probably undergo a drug-free trial period. The value of a structured program for discontinuation and the substitution of medications with low abuse liability are supported in several studies.

There appears to be no relationship between long-term outcome (i.e., not using benzodiazepines at follow-up) and the duration of treatment or the specific benzodiazepine that was discontinued. Some evidence suggests that low doses of benzodiazepines predict successful discontinuation, although other studies have found weak or no correlation between dose and return to benzodiazepine use.

Table 11.10–6.
Pentobarbital (Nembutal) Test Dose Procedure for Barbiturate Withdrawal

Symptoms after Test Dose of 200 mg of Oral pentobarbital	Estimated 24-Hr Oral pentobarbital Dose (mg)	Estimated 24-Hr Oral phenobarbital Dose (mg)
Level I: Asleep, but arousal is possible; withdrawal symptoms not likely	0	0
Level II: Mild sedation; patient may have slurred speech, ataxia, and nystagmus	500–600	150–200
Level III: Patient is comfortable; no evidence of sedation; may have nystagmus	800	250
Level IV: No drug effect	1,000–1,200	300–600

Adapted from Ciraulo DA, Shader RI, eds. Clinical Manual of Chemical Dependence. Washington, DC: American Psychiatric Press; 1991 and data from Ewing JA, Bakewell WE. Diagnosis and management of depressant drug dependence. *Am J Psychiatry*. 1967;123:909.

Barbiturates. There are three common methods of establishing the barbiturate dosage required for safe withdrawal: (1) administration of a test dose of pentobarbital to determine tolerance, (2) calculation of required doses based on estimated equivalencies, and (3) administration of phenobarbital loading doses.

In the first method, the initial step is to determine tolerance. Once intoxication has subsided but before withdrawal symptoms have developed, an oral dose of 200 mg of pentobarbital is administered on an empty stomach, and the effects are observed after 1 hour to determine the stabilization dose (Table 11.10–6). If no changes are observed, the test is repeated 3 hours later using 300 mg. If there is still no response, the total requirement may exceed 1,600 mg per day. The calculated daily dose is divided and given every 4 to 6 hours for a 2- to 3-day stabilization period. Daily dose reductions are usually 10 percent of the stabilization dose, but withdrawal regimens must be individualized. Phenobarbital may be substituted for pentobarbital for stabilization and withdrawal at one-third of the dose.

The second method of barbiturate detoxification requires calculating the equivalent hypnotic dose of phenobarbital on the basis of the dosage of barbiturate or other sedatives the patient reports taking, also taking into

consideration any alcohol consumed. A dose of 30 mg of phenobarbital is substituted for an equivalent hypnotic dose of the sedative. The patient is stabilized for the period of peak vulnerability to withdrawal, 2 to 3 days for sedative-hypnotic agents with short half-lives, and reductions of 30 to 60 mg of phenobarbital are made every 2 or 3 days. Clinicians using this method report that only rarely are doses of phenobarbital greater than 600 mg required. The procedure is not recommended, because published equivalencies are only approximations, and patients vary greatly in their metabolism and pharmacodynamic response to the same dose of the sedative. Furthermore, animal studies indicate that cross-tolerance among the various anxiolytic and sedative-hypnotic agents is not complete. Furthermore, the method relies on the patient's history of drug use, which is often unreliable. Plasma concentrations can establish which sedatives have been ingested and may indicate the degree of tolerance when clinical correlations are made. Estimation of tolerance (and thus determination of dosage requirements) is best made after a test dose or a loading dose is administered. Tables 11.10–3 and 11.10–5 list approximate equivalencies of selected barbiturates, benzodiazepines, and sedatives.

The final method for medical withdrawal from barbiturates involves using loading doses of phenobarbital. According to the original protocol, 120-mg doses are administered every 1 to 2 hours until three of five signs are present—nystagmus, drowsiness, ataxia, dysarthria, or emotional lability—or, if patients are in withdrawal, until abstinence symptoms abate. No additional drug is administered, and, because of its long half-life, the drug self-tapers. Patients are assessed before each dose. Some clinicians have modified the protocol to use doses of 60 mg every 1 to 2 hours and to use a gradual taper rather than abrupt discontinuation. The originators of the protocol reported a mean loading dose of 1,440 mg, with hourly doses sometimes required for 15 to 20 hours. They have also administered phenobarbital IV at .3 mg/kg per minute, with close medical supervision, for medically ill patients. A loading dose procedure for withdrawal from butalbital containing headache remedies has also been described. When patients are known to be taking less than 10 tablets (i.e., Fioricet) per day, reductions of 1 tablet every 2 to 3 days may be successful. However, since patients do not always reliably report drug intake, an alternative strategy is to use a loading dose of 120 mg of phenobarbital hourly guided by a rating scale to document effects on nystagmus, dysarthria, ataxia, and drowsiness. One study found that the mean number of phenobarbital doses required was 9.7 (range 7 to 14) and withdrawal was completed in less than 24 hours.

Overdose

The treatment of overdose of the general class of substances involves gastric lavage, activated charcoal, and careful monitoring of vital signs and

CNS activity. Overdose patients who come to attention while awake should be kept from slipping into unconsciousness. Activated charcoal should be administered to delay gastric absorption. If the patient is comatose, the clinician must establish an IV fluid line, monitor the patient's vital signs, insert an endotracheal tube to maintain a patent airway, and provide mechanical ventilation, if necessary. Hospitalization of a comatose patient in an intensive care unit is usually required during the early stages of recovery from such overdoses.

REFERENCES

Allison C, Pratt JA. Neuroadaptive processes in GABAergic and glutamatergic systems in benzodiazepine dependence. *Psychopharmacol Ther*. 2003;98:171.

Billioti de Gage S, Moride Y, Ducruet T, et al. Benzodiazepine use and risk of Alzheimer's disease: case-control study. *BMJ*. 2014;349:g5205.

Bruce SE, Vasile RG, Goisman RM, Salzman C, Spencer M. Are benzodiazepines still the medication of choice for patients with panic disorder with or without agoraphobia? *Am J Psychiatry*. 2003;160:1432.

Chen KW, Berger CC, Forde DP, D'adamo C, Weintraub E, Gandhi D. Benzodiazepine use and misuse among patients in a methadone program. *BMC Psychiatry*. 2011:11:90.

de las Cuevas C, Sanz E, de la Fuente J. Benzodiazepines: more "behavioural" addiction than dependence. *Psychopharmacology*. 2003;167:297.

*Denis C, Fatseas M, Lavie E, Auriacombe M. Pharmacological interventions for benzodiazepine mono-dependence management in outpatient settings. *Cochrane Database Syst Rev*. 2006;3:CD005194.

Druid H, Holmgren P, Ahlner J. Flunitrazepam and evaluation of use, abuse, and toxicity. *Forensic Sci Int*. 2001;122:136.

Griffiths AN, Jones DM, Richens A. Zopiclone produces effects on human performance similar to flurazepam, lormetazepam and triazolam. *Br J Clin Pharmacol*. 1986;21:647.

Hood S, O'Neil G, Hulse, G. The role of flumazenil in the treatment of benzodiazepine dependence: physiological and psychological profiles. *J Psychopharmacol*. 2009;23(4):401–409.

Huang B, Dawson DA, Stinson FS, et al. (2006). Prevalence, correlates, and comorbidity of nonmedical prescription drug use and drug use disorders in the United States: results of the National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry*. 2006;67(7):1062–1073.

Johnston LD, O'Malley PM, Miech RA, Bachman JG, Schulenberg JE. *Monitoring the Future National Survey Results on Drug Use: 1975–2014: Overview, Key Findings on Adolescent Drug Use*. Ann Arbor: Institute for Social Research, The University of Michigan; 2015.

*Lee SC, Klein-Schwartz W, Doyon S, Welsh C. Comparison of toxicity associated with nonmedical use of benzodiazepines with buprenorphine or methadone. *Drug Alcohol Depend*. 2014;138:118–123.

Martinez-Hernandez E, Sandoval A, Gonzalez-Ramirez, R, Zoidis G, Felix R. Inhibition of recombinant N-type and native high voltage-gated neuronal Ca2+ channels by AdGABA:

mechanism of action studies. *Toxicol Appl Pharmacol*. 2011;250(3):270–277.

*Mohler H. The rise of a new GABA pharmacology. *Neuropharmacology*. 2011;60(7-8):1042–1049.

Morin CM, Bastien C, Guay B, Radouco-Thomas M, Leblanc J, Vallieres A. Randomized clinical trial of supervised tapering and cognitive behavior therapy to facilitate benzodiazepine discontinuation in older adults with chronic insomnia. *Am J Psychiatry*. 2004:161(2):332–342.

Mumford GK, Evans SM, Fleishaker JC, Griffiths RR. Alprazolam absorption kinetics affects abuse liability. *Clin Pharmacol Ther*. 1995;57:356.

Nardi AE, Freire RC, Valenca AM, et al. Tapering clonazepam in patients with panic disorder after at least 3 years of treatment. *J Clin Psychopharmacol*. 2010;30(3):290–293.

Nordfjærn T. A population-based cohort study of anxiety, depression, sleep and alcohol outcomes among benzodiazepine and z-hypnotic users. *Addict Behav*. 2012;37(10):1151–1157.

*Otto MW, McHugh RK, Simon NM, Farach FJ, Worthington JJ, Pollack MH. Efficacy of CBT for benzodiazepine discontinuation in patients with panic disorder: further evaluation. *Behav Res Ther*. 2010;48(8):720–727.

Paterniti S, Dufouil C, Alperovitch A. Long-term benzodiazepine use and cognitive decline in the elderly: the epidemiology of vascular aging study. *J Clin Psychopharmacol*. 2002;22:285.

Posternak MA, Mueller TI. Assessing the risks and benefits of benzodiazepines for anxiety disorders in patients with a history of substance abuse or dependence. *Am J Addict*. 2001;10:48.

Song J, Shen G, Greenfield LJ Jr, Tietz EI. Benzodiazepine withdrawal-induced glutamatergic plasticity involves up-regulation of GluR1-containing alpha-amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptors in Hippocampal CA1 neurons. *J Pharmacol Exp Ther*. 2007;322(2):569–581.

Specka M, Bonnet U, Heilmann M, Schifano F, Scherbaum N. Longitudinal patterns of benzodiazepine consumption in a German cohort of methadone maintenance treatment patients. *Hum Psychopharmacol*. 2011;26:404–411.

Substance Abuse and Mental Health Services Administration, Office of Applied Studies. *Substance Abuse Treatment Admissions for Abuse of Benzodiazepines*. The TEDS Report. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2011.

Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. *The DAWN Report: Benzodiazepines in Combination with Opioid Pain Relievers or Alcohol: Greater Risk of More Serious ED Visit Outcomes*; 2014.

Substance Abuse and Mental Health Services Administration. *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings, NSDUH Series H-48, HHS Publication No. (SMA) 14–4863*. Rockville, MD: Substance Abuse and Mental Health ServicesAdministration; 2014.

*Vorma H, Naukkarinen H, Sarna S, Kuoppsalami K. Long-term outcome after benzodiazepine withdrawal treatment in subjects with complicated dependence. *Drug Alcohol Abuse*. 2003;70:309.

Vorma H, Naukkarinen H, Sarna S, Kuoppasalmi K. Symptom severity and quality of life after benzodiazepine withdrawal treatment in participants with complicated dependence. *Addict*

Behav. 2004;29(6):1059–1065.

Vorma H, Naukkarinen HH, Sarna SJ, Kuoppasalmi KI. Predictors of benzodiazepine discontinuation in subjects manifesting complicated dependence. *Subst Use Misuse*. 2005;40(4):499–510.

Wafford KA. $GABA_A$ receptor subtypes: any clues to the mechanism of benzodiazepine dependence? *Curr Opin Pharmacol*. 2005;5(1):47–52.

Xiang K, Tietz EI. Benzodiazepine-induced hippocampal CA1 neuron alpha-amino-3-hydroxy-5-methylisoxasole-4-propionic acid (AMPA) receptor plasticity linked to severity of withdrawal anxiety: differential role of voltage-gated calcium channels and N-methyl-D-aspartic acid receptors. *Behav Pharmacol*. 2007;18(5-6):447–460.

Zawertailo LA, Busto UE, Kaplan HL, Greenblatt DJ, Sellers EM. Comparative abuse liability and pharmacological effects of meprobamate, triazolam, and butabarbital. *J Clin Psychopharmacol*. 2003;23:269.

▲ 11.11 Anabolic-Androgenic Steroid-Related Disorders

HARRISON G. POPE, JR., M.D., AND GEN KANAYAMA, M.D., PH.D.

Table 11.11–1. Examples of Commonly Used Anabolic Steroids

Compounds usually administered orally
Fluoxymesterone (Halotestin, Android-F, Ultrandren)
Methandienone (formerly called methandrostenolone; Dianabol)
Methyltestosterone (Android, Testred, Virilon)
Mibolerone (Cheque Drops[a])
Oxandrolone (Anavar)
Oxymetholone (Anadrol, Hemogenin)
Mesterolone (Mestoranum, Proviron)
Stanozolol (Winstrol)
Compounds usually administered intramuscularly
Nandrolone decanoate (Deca-Durabolin)
Nandrolone phenylpropionate (Durabolin)
Methenolone enanthate (Primobolan Depot)
Boldenone undecylenate (Equipoise[a])
Stanozolol (Winstrol-V[a])
Testosterone esters blends (Sustanon, Sten)
Testosterone cypionate
Testosterone enanthate (Delatestryl)
Testosterone propionate (Testoviron, Androlan)
Testosterone undecanoate (Andriol, Restandol)
Trenbolone acetate (Finajet, Finaplix[a])
Trenbolone hexahydrobenzylcarbonate (Parabolan)

Note: Many of the brand names listed above are foreign but are included because of the widespread illicit use of foreign steroid preparations in the United States.

[a]Veterinary compound. With permission.

Anabolic–androgenic steroid (AAS) abuse has become a major form of substance abuse worldwide, with tens of millions of individuals who have now used these drugs to improve athletic performance or, more commonly, to simply enhance personal appearance. However, AAS abuse is a recent phenomenon, having become common in the general population only within the last two or three decades. Because widespread AAS abuse is so new, psychiatric research findings on the epidemiology, features, and treatment of this condition remain limited in comparison to the volume of research on other forms of substance abuse. In particular, there are still only a modest number of published clinical studies where AAS users have been recruited and evaluated in person. Thus, many findings in this field remain tentative, and erroneous beliefs about AAS use are still common among both lay individuals and mental health professionals.

The AAS are a family of hormones that includes the natural male hormone testosterone, together with numerous derivatives of testosterone that have been synthesized over the last 75 years (Table 11-11–1). All AAS produce both *anabolic* (muscle-building) and *androgenic* (masculinizing) effects. Importantly, AAS should be distinguished from other families of steroid hormones such as *corticosteroids,* which include the natural hormone cortisol, together with its synthetic derivatives such as prednisone. Corticosteroids have no anabolic properties and hence little abuse potential. Often, when the generic term "steroids" is used, individuals confuse these two families of hormones, and may think that the "steroids" administered by doctors for shoulder or back pain are the same "steroids" used illicitly by weightlifters to gain muscle. To avoid this confusion, the term "AAS," rather than "steroids" is used throughout this chapter to indicate that the authors are referring to testosterone and its synthetic relatives.

HISTORY

Some of the highlights in the history of AAS are summarized in Figure 11.11–1. Testosterone was first synthesized in 1935 in Nazi Germany, and over the next few years, numerous synthetic analogues of testosterone were developed, creating the family of AAS. Both the anabolic and androgenic actions of these drugs were quickly recognized; they were reportedly first used by Hitler's troops in World War II to increase aggressiveness, although this claim has never been verified. In the late 1930s and early 1940s, a number of psychiatric reports suggested that testosterone and other AAS might possess antidepressant properties. These hormones were believed to be especially effective for the "male climacteric" and were prescribed in Europe and in the United States for depressed middle-aged men. With the advent of ECT and standard antidepressant medications, however, the use of AAS for psychiatric purposes rapidly waned. Then starting in the 1990s, interest in the therapeutic effects of AAS in psychiatry

was renewed, partially as a result of the recognition that testosterone produced antidepressant effects in men with HIV infection. Also, by the 2000s, various transdermal testosterone gels were developed and marketed, eliminating the need to administer testosterone by cumbersome intramuscular injections. Thus, the medical use of testosterone for its potential antidepressant properties, or simply (and more controversially) as an "anti-aging" treatment, has increased markedly in the United States over the last two decades.

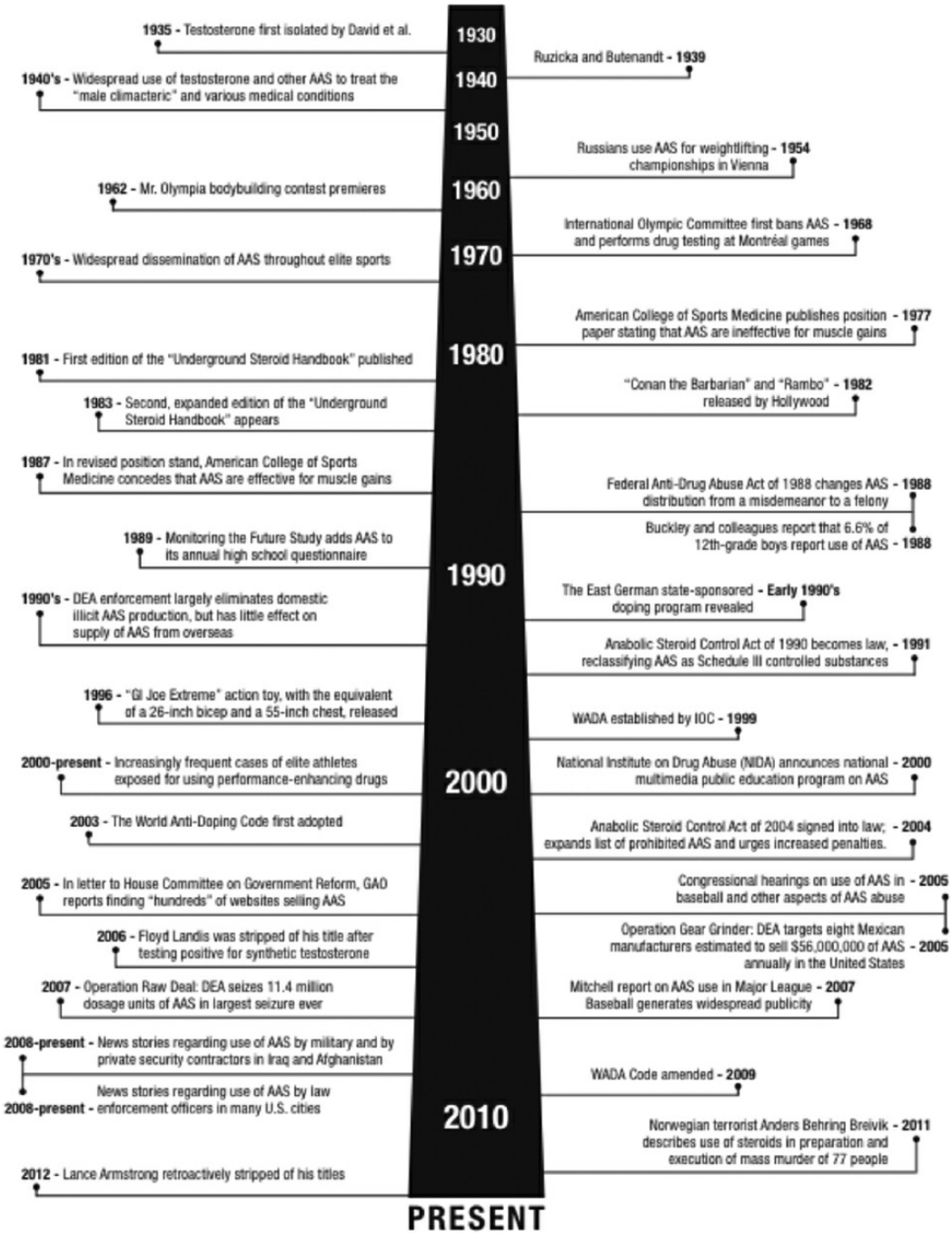

FIGURE 11.11–1. An historic timeline of the evolution of image- and performance-enhancing drug use.

In parallel with the evolution of medical applications for AAS, the use of these drugs for muscular development and athletic performance also began to evolve soon after World War II. Elite athletes quickly discovered that these drugs could create muscle gains greater than could be achieved by any natural individual, and by the 1950s, AAS use had already entered competitive sports, especially those requiring muscle mass and strength, such as bodybuilding, powerlifting, American football, and field events such as the shot-put. The first widely documented use of AAS in international competition was by the Russians at the world weightlifting championships in Vienna in 1954. By the 1960s and 1970s, AAS were used extensively in the Olympics and other major athletic events. Perhaps the most dramatic example of AAS use was in the German Democratic Republic, which developed an extensive state-sponsored doping program, which allowed a country with only 18 million citizens to win more than 150 Olympic gold medals between 1960 and 1990.

Although the power of AAS was widely recognized by athletes, the medical community was embarrassingly slow to acknowledge these effects. During the 1960s and 1970s, numerous placebo-controlled studies attempted to assess whether AAS were effective for gaining muscle mass. Many of these studies produced negative results, almost certainly because they employed dosages and durations of AAS administration far below those used by actual athletes in the field. On the basis of these often-flawed studies, many reputable textbooks and reviews in pharmacology and endocrinology concluded that AAS were ineffective for athletic purposes. Athletes who had actually used these drugs tended to hold the medical profession in contempt for this display of ignorance. Evidence suggests that even to this day, AAS users often distrust medical professionals on the subject, and consequently discount medical warnings about the possible physical and psychiatric dangers of these agents. For example, in one recent study of 43 AAS users, 56 percent reported that they had never disclosed their AAS use to any physician that they had ever seen, and 40 percent reported that they would trust information on AAS from their local drug dealer at least as much as information from any physician.

AAS use remained largely confined to athletes until the 1980s, but then began to break out of the athletic world and onto the street. Underground handbooks on how to use AAS began to appear, providing detailed information about the various AAS available on the illicit market, and instructions on how to use and self-inject them. Also during this era, Western culture showed a growing emphasis on male muscularity. Images of muscular male bodies increasingly appeared in Western magazines, advertising, television, and Hollywood movies. Even children's action toys, such as "G.I. Joe," grew steadily more muscular from the 1970s onward. These trends helped to propel a rapid rise in the prevalence of illicit AAS use among ordinary rank-and-file men lifting weights in gymnasiums, who

often had no athletic aspirations at all, but simply wanted to look bigger. Starting in the late 1980s, these trends began to be documented by epidemiologic studies of AAS use among students, using anonymous questionnaires. Several of these studies noted that a majority of students were using AAS not for athletic purposes, but only for personal appearance. More recent studies of adult AAS users have repeatedly reinforced this impression, finding that a great majority of AAS users have not used these drugs for any competitive athletic activity at all. Nevertheless, many people still erroneously believe that AAS use is primarily a problem of athletes.

With the growing illicit use of AAS in the 1980s, preliminary reports also began to document adverse psychiatric and medical effects associated with using high-dose AAS. These reports included cases of hypomanic and manic syndromes associated with AAS use and depressive syndromes associated with AAS withdrawal. These findings led to increasing concern about the public health consequences of using AAS, and in 1991 the United States Congress passed the Steroid Trafficking Act, adding AAS to the list of Schedule III controlled substances and placing jurisdiction over these substances in the hands of the Drug Enforcement Administration (DEA). In the wake of this legislation, it quickly became difficult for AAS abusers to find doctors willing to prescribe these drugs, and production of AAS by American pharmaceutical companies declined rapidly. Thus, on the illicit market, the supply of genuine AAS diminished, and a larger number of counterfeit AAS, with no pharmacological effects, began to appear. Underground guides to use AAS began to include sections with photographs of bottles and syringes to help users to distinguish genuine preparations from counterfeits. However, despite the existence of counterfeits, an abundant supply of genuine AAS continues to be available because, unlike drugs of abuse such as cocaine and marijuana, AAS may be purchased legally without a prescription in many countries, and can be sent illicitly by mail to the United States (e.g., to post office boxes). Thus, with the rise of the Internet, many AAS users can receive these drugs by mail from abroad, usually without interdiction.

Although the general public is quite familiar with many forms of illicit substance abuse, most lay persons and arguably even most physicians are surprisingly unfamiliar with the effects of AAS or the extent of their use. In particular, most people are unaware that, without massive doses of AAS, it is virtually impossible to achieve the combination of low body fat and high muscularity exhibited by many competition bodybuilders and professional athletes. Since athletes and their promoters have large personal and financial motivations to deny such drug use, the public has been hoodwinked for decades about the magnitude of the AAS problem. However, recent years have seen mounting scandals and revelations about AAS use by elite athletes in the Olympics and other competitions, as well as by prominent sports professionals, movie and television actors, and law

enforcement personnel. Public awareness has also increased in response to reports of premature deaths, including suicide, in athletes known or suspected to have taken AAS.

COMPARATIVE NOSOLOGY

Psychiatric syndromes associated with use of AAS are cited in the fifth edition of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) under the general category of "other substance-related use disorders." Within this category are subcategories of other substance use disorder, other substance intoxication, other substance withdrawal, other substance–induced disorders, and unspecified other substance–related disorder. Similarly, the 10th revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) classifies AAS-related disorders under the category of "mental and behavioural disorders due to multiple drug use and use of other psychoactive substances." Unfortunately, these various terms do not apply easily to AAS use disorders, because AAS produce little "intoxication" in the manner of classical drugs of abuse and often (but not always) lack major psychoactive effects. Therefore, in the text below, the authors will attempt to parallel the standard terminology of DSM-5 for other drugs of abuse, while recognizing that AAS deviate from this terminology in a number of important respects.

EPIDEMIOLOGY

The epidemiology of AAS use has long been associated with persistent erroneous beliefs, including particularly the misconceptions that (1) AAS are used primarily by athletes, (2) AAS users typically start use in adolescence, and (3) girls and women often use AAS.

It has been very difficult to estimate the lifetime prevalence of illicit AAS use in the United States, because most available data rely on anonymous surveys of students, particularly students of high-school age. There are two serious problems with the use of these data. First, the median age of onset of AAS use, both in the United States and other countries, is approximately 23 years—an older age of onset than for almost any other major category of illicit drugs. In pooled data from five American studies, collectively evaluating more than 2,000 AAS users, in which age of onset was ascertained, only 5.9 percent of users had initiated AAS prior to the age of 17, and only 39.3 percent prior to age 21. Thus, studies of high-school students greatly underestimate the number of individuals who will ultimately use AAS in their lifetimes, because the individuals being surveyed have only barely begun to enter the age of risk for initiating AAS use.

Second, anonymous surveys are vulnerable to serious errors in the opposite direction as a result of false-positive responses to the survey questions about "steroids." When students are asked a typical survey

question, such as whether they have used “steroid pills or shots without a doctor’s prescription,” they may falsely respond that they have used “steroids” when in fact they have taken only corticosteroids, or have purchased nutritional supplements over the counter that they mistakenly believed were “steroids.” Many students fail to realize that genuine AAS are available only on the black market in the United States and most other Western countries. Thus, they may often respond positively to a “steroid” question when in reality they have never used AAS at all. A recent analysis of anonymous survey data in the United States has suggested that virtually all high-school girls claiming “steroid” use on questionnaires represent false positives, and that the true use prevalence of AAS use in girls may be as low as 0.1 percent. This finding is hardly surprising, when it is considered that girls rarely desire to become extremely muscular, and are also vulnerable to the androgenic effects of AAS, such as beard growth, deepening of the voice, and masculinization of secondary sexual characteristics. Similarly, anonymous survey data for high-school boys also need to be corrected sharply downward.

After allowing for these two sources of error, and digesting all available American survey data, a recent analysis has suggested that the true prevalence of AAS use in high-school girls is close to 0 percent, and the prevalence of AAS use in high-school boys is about 1 percent. However, the same analysis suggested that by adulthood, a far larger number of American men have used AAS, with an estimate of 2.9 to 4.0 American men who have used AAS at some time in their lives. However, the corresponding prevalence for adult women is likely extremely low; in five available studies since the year 2000 that have recruited AAS users without regard to gender, the mean proportion of female users, calculated using a random effects model, was 1.8 percent—a male–female ratio of more than 50 to 1. Moreover, the authors are aware of only one study worldwide in the last 25 years in which a group of investigators recruited actual female AAS users in person, and even in this study, the investigators located only 25 such women over a 2-year interval in three major American cities. Thus, from a public health standpoint, the problems of AAS use are largely confined to males.

AAS users often use other *ergogenic* (performance-enhancing) drugs to gain muscle, lose fat, improve aerobic capacity, or to lose water for bodybuilding competitions. These drugs include thyroid hormones, such as liothyronine (Cytomel) and levothyroxine (Synthroid); human growth hormone and its daughter compound, somatomedin-C (insulin-like growth factor-1 [IGF-1]); insulin; sympathomimetic agents such as amphetamines, ephedrine and pseudoephedrine (Sudafed); β-adrenergic receptor agonists such as clenbuterol; the amino acid derivative GHB; and various laxatives and diuretics.

AAS users also frequently display other substance use disorders

involving classical drugs of abuse, such as cannabis, cocaine, and opioids. One interview study of 44 AAS users found that 44 percent had a history of abuse of or dependence upon at least one other drug or alcohol. Another study of 62 users found that 65 percent reported a history of abuse or dependence upon at least one drug other than alcohol. Similarly, among a large sample of 10,552 college students in 2001, a diagnosis of alcohol dependence in the past year was over seven times more likely among past-year AAS users than nonusers of AAS after adjusting for age, gender, race, intercollegiate athletic status, grade-point average, and geographical region. Several European studies have also documented an association between AAS use and abuse of multiple other classical drugs.

Opioid abuse and dependence appear to be particularly associated with AAS use in the United States. One study found that 21 (9 percent) of 227 sequential male heroin addicts at one facility had apparently been introduced to opioids through AAS; they reported that they had first learned about opioids from fellow AAS users in the gym and had subsequently first purchased opioids from the same individual who had sold them AAS. Another study of 88 men admitted for opioid dependence found that 22 (25 percent) had used AAS at some time in their lives. Interestingly, most of these men reported that AAS were the first drugs that they had learned to use by injection. Collectively, these data suggest that AAS use is often part of a pattern of polydrug use, and that AAS have increasingly become "street" drugs, rather than drugs used purely for athletic performance.

PHARMACOLOGY

Chemistry

All steroid drugs—including AAS, estrogens, and corticosteroids—are synthesized in vivo from cholesterol and resemble cholesterol in their chemical structure (Fig. 11.11–2). Testosterone has a four-ring chemical structure containing 19 carbon atoms (Fig. 11.11–2). Most synthetic derivatives of testosterone are produced by one or more of the following modifications: (1) forming an ester at the C-17-β-hydroxyl group, (2) adding an alkyl group to the C-17-α position, or (3) altering the structure at a different carbon site.

Aqueous testosterone suspension is essentially inactive when ingested by the oral route because of first-pass metabolism in the liver, and is only briefly active when injected because it is rapidly absorbed and metabolized. Synthetic AAS overcome these problems. Testosterone esters (e.g., testosterone cypionate [Depo-Testosterone] and testosterone enanthate [Delatestryl]) confer the advantage of slower absorption than testosterone when injected, and their effects last from 2 to 4 weeks. Testosterone esters are hydrolyzed to testosterone after they are absorbed, so these agents can

be monitored by blood concentrations. Testosterone is metabolized in part to dihydrotestosterone, which is up to 10 times more potent than testosterone as an androgen, and in part to the estrogen estradiol, which accounts for some of the adverse effects seen in men, such as gynecomastia. Testosterone is also metabolized to neurosteroids in the brain, especially 3-α-androstanediol, which may account for some of its psychoactive effects. The effects of testosterone depend both on the balance of metabolic enzymes and on the concentration of androgen receptors in a particular tissue. The C-17-alkylated synthetic AAS make oral administration possible, but once-daily use is generally required.

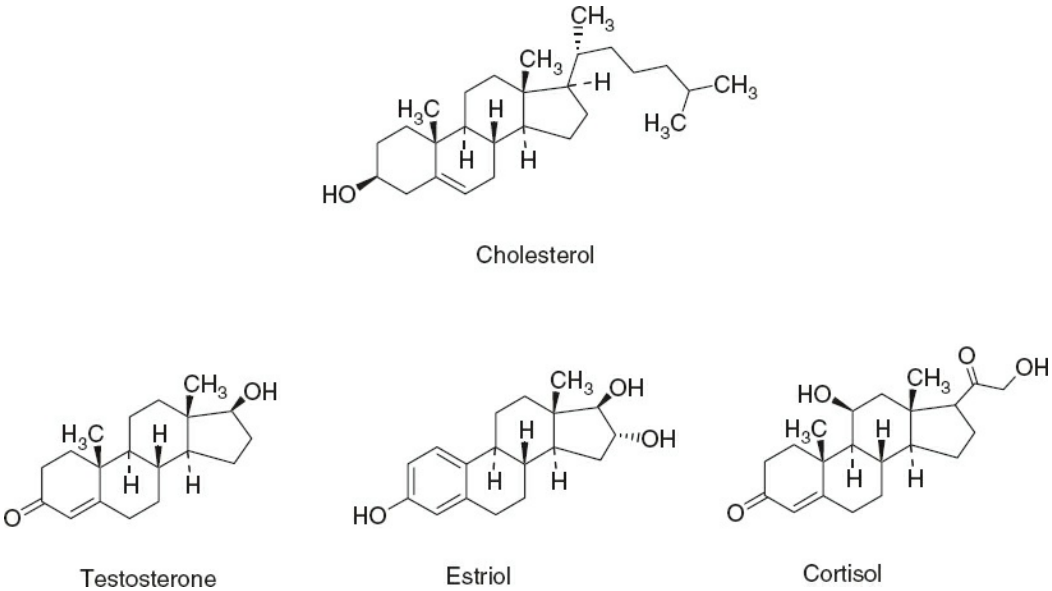

FIGURE 11.11–2. Chemical structure of cholesterol and three hormones derived from cholesterol: Testosterone (an anabolic–androgenic steroid), estriol (an estrogenic steroid), and cortisol (a corticosteroid).

Pharmacodynamics

AAS bind androgen receptors in the cytoplasm of cells. The androgen-receptor complex is then translocated to the cell nucleus where it augments gene transcription and ultimately new protein synthesis. The anabolic effects desired by illicit users result especially from new protein synthesis in muscle tissue, which likely results from a genomic action of AAS. The mechanism of action for psychoactive effects is poorly understood, but may involve both genomic and nongenomic actions of AAS as well as actions on both neurons and glial cells. Nongenomic actions may include binding to membrane-associated androgen receptors, including allosteric sites on GABA receptors. Androgen metabolites such as estradiol and neurosteroids can also have psychoactive effects. The neurosteroid metabolite of testosterone, 3-α-androstanediol, for example, is a positive allosteric modulator of GABA receptors, thereby potentiating the action of GABA.

A number of neural structures and neurotransmitters have been

implicated in the psychiatric effects of AAS. Most of this work comes from animal studies, demonstrating that parts of the limbic system (including the amygdala, hippocampus, and nucleus accumbens) and the hypothalamus are all target sites for the actions of AAS. In addition, androgen receptors are known to be expressed in the cerebral cortex. Animal studies also suggest that AAS-induced increases in vasopressin, substance P, and glutamate as well as AAS-induced decreases in brain serotonin and 5-HIAA may be associated with aggressive behavior. Forebrain GABAergic mechanisms may also be involved in AAS-induced aggression and anxiety. Finally, AAS may exert their reinforcing effects through their actions on GABA, opioid, and dopamine systems, as will be discussed below.

In humans, a recent experimental study reported significantly higher levels of 5-hydroxyindoleacetic acid (5-HIAA) in the cerebrospinal fluid of men treated with methyltestosterone as compared to men receiving placebo. Moreover, 5-HIAA levels correlated significantly with certain AAS-associated effects, including increased energy, diminished sleep, and sexual arousal.

Normal testosterone plasma concentrations for men range from 300 to 1,000 ng/dL. Generally, 200 mg of testosterone cypionate taken every 2 weeks will restore physiological testosterone concentrations in a hypogonadal male. A eugonadal male who initiates physiological dosages of testosterone will have no net gain in testosterone concentrations, because exogenously administered AAS will shut down endogenous testosterone production via feedback inhibition of the hypothalamic-pituitary-testicular (HPT) axis. Consequently, illicit users take dosages 10 to 100 times greater than therapeutic discs in order to achieve the desired effects. Doses in this range are most easily achieved by taking combinations of oral and injected AAS, which illicit AAS users often do. Transdermal testosterone gels or patches are generally inadequate to produce massively supraphysiologic AAS levels.

Effects on Specific Organs and Systems

The first effects of endogenous androgens are to induce male sexual differentiation in the fetus. Accumulating evidence suggests that fetal and perinatal androgens cause sexual dimorphism of the brain (at least in nonhuman mammals) as well as differentiation of the sexual organs. During puberty, testosterone levels increase in boys to produce male secondary sex characteristics. The phallus, prostate, and seminal vesicles all grow and develop, and testosterone and gonadotropins are involved in sperm production. Androgens cause the larynx to enlarge, the vocal cords to lengthen and thicken, and the voice to deepen. A male pattern of body hair develops, including hair on the face and chest. Sebaceous glands grow and increase their secretions, predisposing the adolescent to acne. Bones

and muscles also grow during puberty as a result of androgen stimulation. Exogenous administration of AAS to a prepubertal male will also induce all of these effects.

In adults, physiological levels of testosterone continue to be important for spermatogenesis, although abnormally high levels will suppress spermatogenesis by inhibiting release of gonadotropins, which are also required for spermatogenesis. Exogenous AAS increase muscle growth, promote mineralization of bones, and stimulate erythropoiesis by increasing erythropoietin levels and by direct action on the bone marrow.

Medical Uses

AAS are primarily indicated for testosterone deficiency (male hypogonadism), hereditary angioedema (a congenital skin disorder), and some uncommon forms of anemia caused by bone marrow or renal failure. In women, they are given, although not as first-choice agents, for metastatic breast cancer, osteoporosis, endometriosis, and adjunctive treatment of menopausal symptoms. In men, they have been used experimentally as a male contraceptive and for treating major depressive disorder and sexual disorders in eugonadal men. More recently, they have been used to treat wasting syndromes associated with AIDS where they may also have antidepressant effects.

Adverse Medical Effects

The most common acute adverse medical effects of AAS include acne (which occasionally may be severe), gynecomastia (caused by aromatization of AAS to estrogenic hormones), striae (caused by overly rapid muscle growth beneath the skin), and acceleration of male pattern baldness. These features, if observed in an unusually muscular individual, may alert the clinician to surreptitious AAS use.

Although the *acute* adverse medical effects of AAS are usually modest, the *long-term* dangers of these drugs are substantial, and are only beginning to be studied. One reason for the dearth of research is the fact that widespread illicit AAS use did not appear until the 1980s, as mentioned above. Thus, even most of the older AAS users—those who first tried AAS as youths in the 1980s—are only now reaching middle age and entering the age of risk for various adverse medical effects.

Perhaps the best-documented long-term adverse effects of AAS are on the cardiovascular system. A rapidly evolving literature has shown that AAS exposure can cause a dilated cardiomyopathy, characterized by both systolic dysfunction (as evidenced by a reduced ventricular ejection fraction) and by diastolic dysfunction (as manifested by reduced diastolic tissue velocity). Numerous studies have now shown that the effects of weightlifting alone cannot explain these phenomena, in that non-AAS-

using weightlifters display few abnormalities on these indices. Tentative data suggest that AAS-induced cardiomyopathy is at least partially reversible, with normalization of systolic and diastolic function occurring in a majority of AAS users within months after stopping the drugs. However, impaired diastolic tissue velocity may be at least partly attributable to fibrotic changes caused by AAS-induced cellular injury, and hence may be potentially irreversible.

Another serious cardiovascular effect of AAS is accelerated atherosclerosis. These atherosclerotic changes appear largely attributable to the fact that AAS increase low-density lipoprotein and decrease high-density lipoprotein cholesterol. Dyslipidemia of this type is particularly severe with use of 17-α-alkylated AAS that are administered orally. Atherosclerotic disease may cause premature death from myocardial or cerebral infarction among occasional AAS users even in their 20s or 30s; at least a dozen such case reports have been published in the last 20 years. One imaging study of 14 professional weightlifters with long-term AAS use found markedly higher coronary-artery calcium scores than would be expected for men of comparable age. Another study performed coronary computed tomography angiography in 57 AAS users age 35 to 55, all of whom reported at least 2 years of lifetime exposure to AAS, and 36 age-matched men reporting no AAS use. On measures of the number of diseased coronary artery segments, the worst degree of stenosis, and total plaque volume, AAS users showed greater pathology than nonusers, and the degree of pathology was strongly associated with total duration of lifetime AAS exposure. The potential clinical consequences of such changes are suggested by a study from Finland that examined the mortality of 62 male powerlifters who placed 1st to 5th in Finnish championships during 1977 to 1982 (all highly likely to have used AAS) and 1,094 population controls. Eight (12.9 percent) of the powerlifters but only 34 (3.1 percent) of controls had died—a ratio of 4.6—and 3 of the 8 deaths among powerlifters were ascribed to myocardial infarction.

AAS have also been implicated in causing hypertension, cardiac conduction abnormalities, and thrombotic events in both human and animal studies, although the frequency and severity of these phenomena in AAS users remain poorly understood. Chronic use of high doses of AAS may also frequently cause polycythemia, with some users developing a hematocrit of 55 to 60 percent, and sometimes higher than 60 percent. Again, however, the magnitude and frequency of clinical morbidity associated with AAS-induced polycythemia are uncertain.

AAS may also exert substantial neuroendocrine toxicity, because AAS suppress HPT function. Thus, male users will often become hypogonadal after stopping a course of AAS use. Although HPT function usually recovers spontaneously within a few weeks to a few months, hypogonadism may sometimes persist for more than a year after discontinuing AAS, and may

possibly be irreversible in some cases. There may be several different mechanisms causing AAS-induced hypogonadism in different men. Some appear to show hypogonadotropic hypogonadism, due to incomplete recovery of gonadotrope function. Several reports have described successful treatment of such cases with gonadotropin releasing hormone agonists such as triptorelin, selective estrogen receptor modulators such as clomiphene, or agents mimicking pituitary hormonal function, such as human chorionic gonadotropin (HCG). However, these approaches sometimes fail. For example, some patients may respond to clomiphene with normalization of pituitary hormone levels, but then fail to normalize testosterone levels, suggesting possibly irreversible damage to Leydig cells. Recent reports have also described former AAS users who showed persistent erectile dysfunction and loss of libido even when treated with physiologic levels of exogenous testosterone—possibly reflecting some type of end organ resistance.

One recent study compared 19 former long-term AAS users and 36 non-AAS-using weightlifters using measures of psychopathology, neuroendocrine function, and erectile function. Five of the former AAS users displayed markedly reduced testosterone levels of less than 200 ng/dL despite having discontinued AAS 3 to 26 months earlier. The former users also displayed markedly and significantly lower mean scores on ratings of sexual desire. Five former users had experienced episodes of major depressive disorder during AAS withdrawal.

Also of concern is the effect of prolonged AAS use on tissues sensitive to androgens, such as the prostate. AAS may contribute to prostatic hypertrophy, and two case reports have documented prostate cancer among AAS users who were still in their 40s. However, both of these reports appeared more than 25 years ago; the apparent lack of reports since 1991 argues against a major association of AAS use with prostatic neoplasia.

Among the musculoskeletal effects of AAS is an increased risk of tendon rupture. One recent study found that 15 (17 percent) of 88 long-term AAS users had ruptured an upper body tendon (biceps, triceps, or pectoralis) at some time in their lives as opposed to 0 cases among 50 age-matched non-AAS-using weightlifters. Notably, in this and other studies, most of the AAS users did not rupture tendons while doing heavy weightlifting, but instead as a result of lifting objects, falling, or other daily activities. The increased risk of tendon rupture in AAS users appears at least partially attributable to the fact that these men develop muscles that become too strong relative to their tendons. However, some animal data suggest that AAS may additionally interfere with the architecture of the tendons themselves by creating collagen dysplasia, causing tendons to become stiffer and less flexible, with an increased crimp angle and earlier liability to failure.

Rarely, AAS may cause hepatotoxicity, with the effects including peliosis

hepatis (an accumulation of blood-filled cysts in the liver), and the development of both benign and malignant hepatic tumors. These hepatotoxic effects occur almost exclusively with use of orally active 17-α-alkylated AAS. The true incidence of AAS-induced hepatotoxicity is often overestimated, because rhabdomyolysis from weightlifting can increase transaminase levels, and this finding is often incorrectly interpreted as evidence of hepatotoxicity. In reality, true AAS-induced serious hepatotoxicity is very rare.

Rhabdomyolysis from heavy weightlifting may also possibly lead to renal failure in occasional AAS users, and one study has described a series of cases of focal segmental glomerulonephritis in long-term AAS users. There is some possibility that AAS is very rarely associated with renal neoplasms.

Several recent studies have raised the alarming possibility that high doses of AAS may display neurotoxicity, thus possibly leading to cognitive deficits. For example, in several laboratory studies, supraphysiologic levels of AAS—within the range likely attainable by human AAS abusers—led to accelerated apoptotic effects on neuronal cells. Following up on these findings, one recent study has assessed cognitive function in long-term AAS users versus comparison of non-AAS-using weightlifters. The study found significant deficits on two tests of visuospatial memory in the AAS users, and the severity of these deficits was significantly associated with total lifetime AAS exposure. These ominous findings will require replication in further investigations using neuropsychological testing and neuroimaging approaches.

In the rare females who use AAS, these drugs may induce shrinkage of breast tissue, irregular menses (diminution or cessation), and masculinization (clitoral hypertrophy, hirsutism, and deepened voice). Masculinizing effects in women may be irreversible, although this issue has not been well studied. Abuse of AAS by children has led to concerns that AAS-induced premature closure of bony epiphyses could cause shortened stature. However, as discussed above, use of true AAS by young teenagers is rare, and the authors have never personally seen an actual case of short stature unequivocally caused by AAS use. Other uncommon adverse effects include edema of the extremities due to water retention, exacerbation of tic disorders, and sleep apnea.

Laboratory Findings

As mentioned above, the use of AAS can be associated with elevations in liver function tests such as bilirubin, lactate dehydrogenase (LDH), alkaline phosphatase, aspartate amino transferase (AST or SGOT), and alanine amino transferase (ALT or SGPT). However, elevations in LDH, ALT, and AST, as well as the muscle enzyme creatine phosphokinase (CPK), are common in weight-training athletes, even in the absence of AAS use, as a

result of muscle trauma from intensive weightlifting. The enzyme gamma-glutamyl transaminase (GGT), which is present exclusively in liver and not in muscle, may be more useful to distinguish frank liver toxicity from mere muscle trauma in users of AAS.

As also discussed above, an altered cholesterol profile (decreased levels of high-density lipoprotein cholesterol and increased levels of low-density lipoprotein cholesterol) is frequently seen, especially with oral administration of the C-17-alkylated AAS. Total cholesterol and triglyceride concentrations may also be elevated. Increases in the hematocrit and hemoglobin levels may also occur.

Endocrine Tests

AAS inhibit the release of luteinizing hormone (LH) and follicle-stimulating hormone (FSH), resulting in lower serum levels. Serum testosterone and estradiol levels vary with the particular agent and timing of the sample. For example, the use of supraphysiological doses of testosterone esters will increase both testosterone and estradiol levels, whereas the exclusive use of most other AAS will decrease these hormonal levels. Similarly, testosterone and estradiol levels are decreased during AAS withdrawal because of the lingering inhibition of the HPT axis, as discussed above. Semen analysis may reveal decreased sperm count and motility as well as abnormal sperm morphology.

ETIOLOGY

The major reason for taking illicit AAS is to enhance either athletic performance or physical appearance. The dramatic effects of AAS on muscle growth are illustrated in Figure 11.11–3, which compares a "natural" bodybuilder who has never used these drugs with a bodybuilder of identical height and body fat who has used AAS extensively. Although the anabolic properties of AAS represent the prime motivation for use, psychoactive effects may also contribute to the persistent and dependent use of AAS. Anecdotally, some AAS users report feelings of power, aggressiveness, and euphoria, which become associated with and can reinforce AAS-taking.

As mentioned earlier, AAS use occurs almost entirely in males, despite persistent misconceptions that females commonly use these drugs. AAS use is widespread in Western cultures, with Scandinavian countries showing perhaps the highest prevalence, followed by the United States, British Commonwealth countries, and Brazil. Moderate rates of AAS use have been documented in studies in continental Europe and the Middle East, but AAS use is rare in Eastern Asian countries such as China, Korea, and Japan. It has been speculated that in these latter societies, masculinity is defined in the Confucian tradition, as someone who has intellect, integrity, and force of character—but not hypertrophied muscles. For example, far Eastern

traditions have few equivalents to muscular Western statues depicting gods of Greek and Roman mythology, or to the depictions of the Norse gods or the Vikings. Thus, East Asian men may see little reward in attaining high levels of muscularity and hence are rarely attracted to AAS.

Two well-documented risk factors for AAS use among Western men are conduct disorder and body image disorders. One study examined 233 experienced American weightlifters age 18 to 40, comprising 102 AAS users and 131 nonusers. In this study, men reporting a history of adolescent conduct disorder were significantly more likely to have subsequently developed AAS use. In the same study, men reporting higher levels of adolescent concern about muscularity showed a strikingly higher risk of developing AAS use than those reporting lower levels. Several other studies have shown a strong association between muscle dysmorphia—a form of body dysmorphic disorder characterized by preoccupations that one is not muscular enough—and use of AAS. Interestingly, this literature suggests that muscle dysmorphia is a risk factor for initiation of AAS use, but paradoxically may worsen after initiating AAS use, perhaps because the individual becomes more involved in the AAS-using culture and becomes increasingly focused on body muscularity.

ANABOLIC STEROID DEPENDENCE

The possibility of AAS dependence was mentioned in the scientific literature as early as 1980. Three case reports appeared in the medical literature between 1988 and 1990, which described young men who initially took AAS to enhance their weightlifting and bodybuilding activity, and later used the drugs to combat depression when they tried to stop using them. The men reported feeling addicted and sought professional help to stop using illicitly obtained AAS because they were unable to accomplish discontinuation on their own.

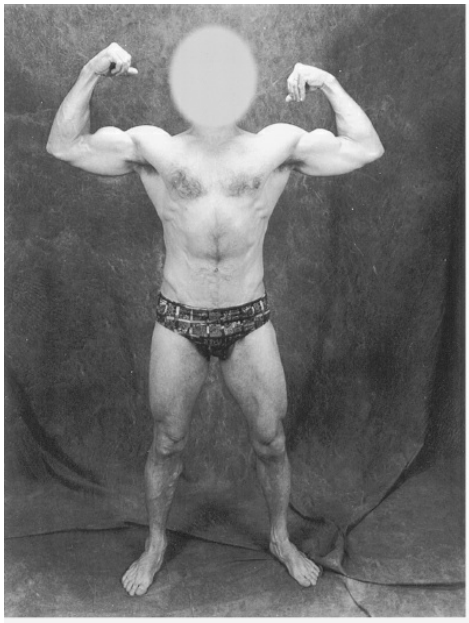
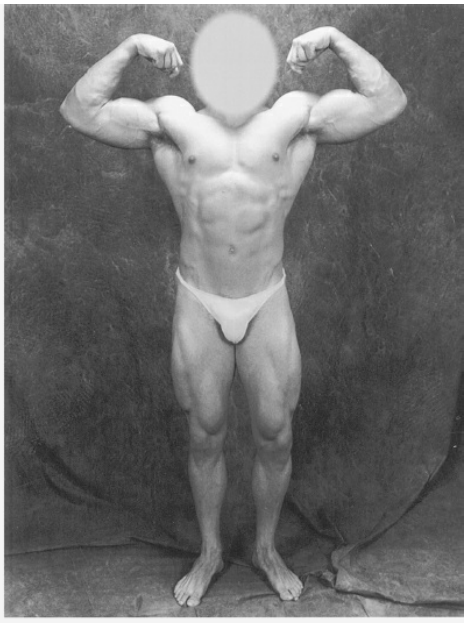
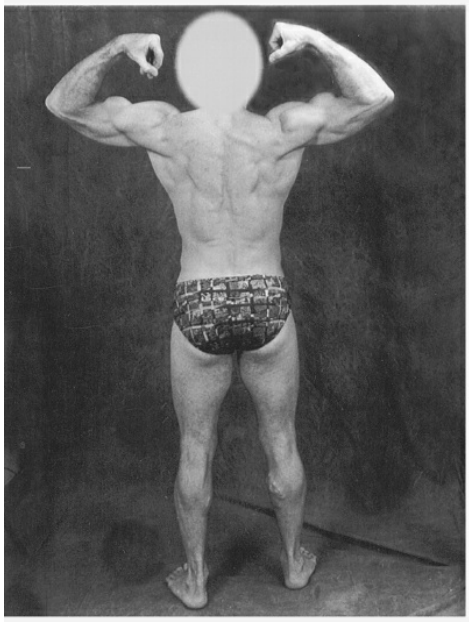
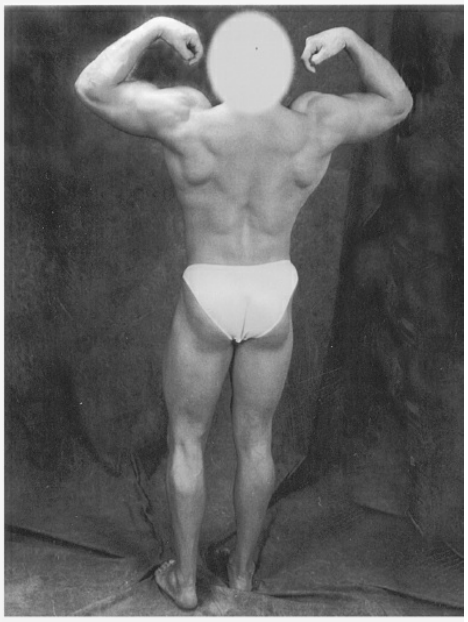
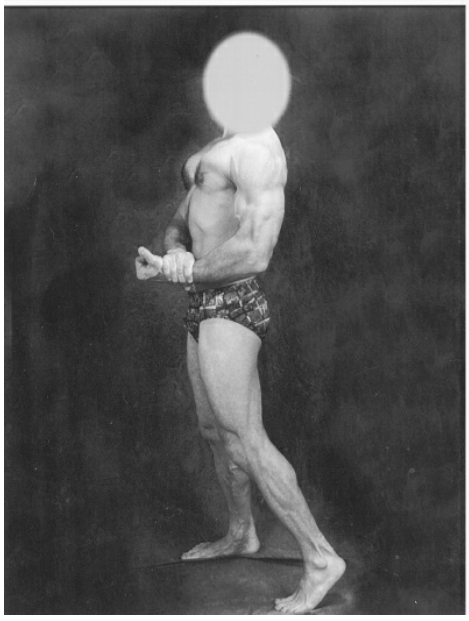
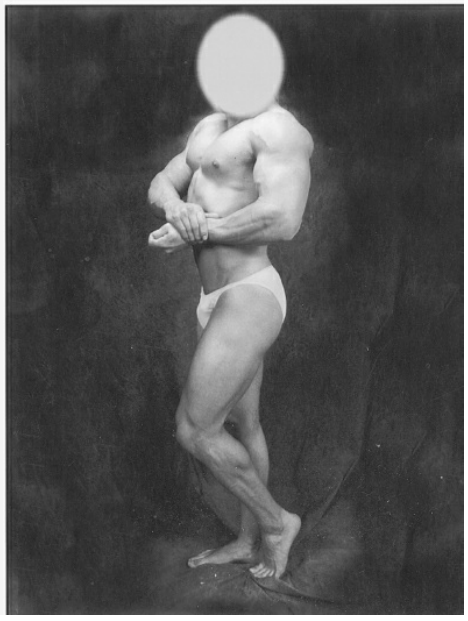

FIGURE 11.11–3. Physical effects of anabolic steroid use. These photographs compare a "natural" bodybuilder who has never used anabolic steroids (*left*) with a man who used large doses of anabolic steroids over several years (*right*). Both men are 67 inches tall and have 7 percent body fat. The man on the left weighs 170 lb and represents approximately the maximum degree of muscularity obtainable without drugs. His fat-free mass index is 25.4 kg/m^2 (by the formula of Kouri EM, Pope HG Jr, Katz DL, Oliva PS. Fat-free mass index in users and nonusers of anabolic–androgenic steroids. *Clin J Sport Med.* 1995;5:223). The man on the right weighs 213 lb and has a fat-free mass index of 31.7 kg/m^2. Note that the muscle hypertrophy from steroid use is particularly marked in the upper body in the pectoralis, deltoid, trapezius, and biceps muscles. Any man significantly more muscular than the man on the left has almost certainly abused anabolic steroids.

Over the last three decades, nine studies, to the authors' knowledge, have utilized criteria for substance dependence adapted from DSM-III-R or DSM-IV to diagnose AAS dependence in 363 of 1247 AAS users—a prevalence of 32.5 percent. Notably only 2 (0.6 percent) of the AAS-dependent individuals in these studies were women. One group has proposed formal diagnostic criteria for AAS dependence based on DSM-IV criteria for substance dependence. These criteria include symptoms of tolerance and withdrawal; prolonged use of AAS and difficulty in discontinuing AAS (e.g., because of anxiety about losing muscle mass); spending large amounts of time involved in AAS-related activities (weightlifting, attending to particular diets for gaining muscle mass, and associating with other AAS users); and continued use of AAS despite adverse effects, such as gynecomastia, sexual dysfunction, or evidence of cardiovascular toxicity.

Another group has suggested that it may be too simplistic to diagnose AAS dependence alone using traditional substance dependence criteria, because AAS use typically occurs as part of a larger complex of behaviors, including the use of other appearance- and performance-enhancing drugs, compulsive patterns of dieting and exercise, and behaviors associated with body image disturbance.

Etiology

AAS dependence may develop via at least three different etiologic pathways (Fig. 11.11–4). First, individuals with body-image disorders may become dependent upon AAS because they develop intense anxiety that they will lose muscularity if they stop taking these drugs. Such individuals may finish a course (a so-called "cycle") of AAS use, but then start a new course of AAS more quickly than planned because of fears that they are beginning to get smaller. Eventually, they may come to use AAS continuously, with no off-drug periods, as a result of this anxiety. Second, as discussed above, men will often develop hypogonadism after stopping a course of AAS, with attendant loss of libido and erectile function, and possible symptoms of major depression. In response to these dysphoric symptoms, they may be quick to resume AAS use in order to restore normal mood and sexual

function, again potentially leading to almost continuous AAS use as a result of this neuroendocrine "addiction."

Third, AAS have been shown to possess hedonic effects, as evidenced by animal studies. For example, rats and mice will exhibit conditioned place preference to AAS, preferring to go to an environment where they have previously received AAS. Also, male hamsters will self-administer testosterone even to the point of death, and this self-administration can be blocked by treatment with opioid antagonists. These data suggest that in addition to their genomic actions described above (binding to cognate intracellular androgen receptors), AAS also likely possess nongenomic effects mediated by binding to plasma membrane receptors. These nongenomic effects may perhaps increase dopaminergic activity in the nucleus accumbens, leading to hedonic and reinforcing properties. The degree of these hedonic effects in humans is difficult to quantitate, although there is no doubt that some AAS users anecdotally describe feelings of increased self-confidence, invincibility, and sometimes euphoria. Thus, it seems likely that hedonic effects combine together with the previous two pathways to contribute to the evolution of AAS dependence in some individuals.

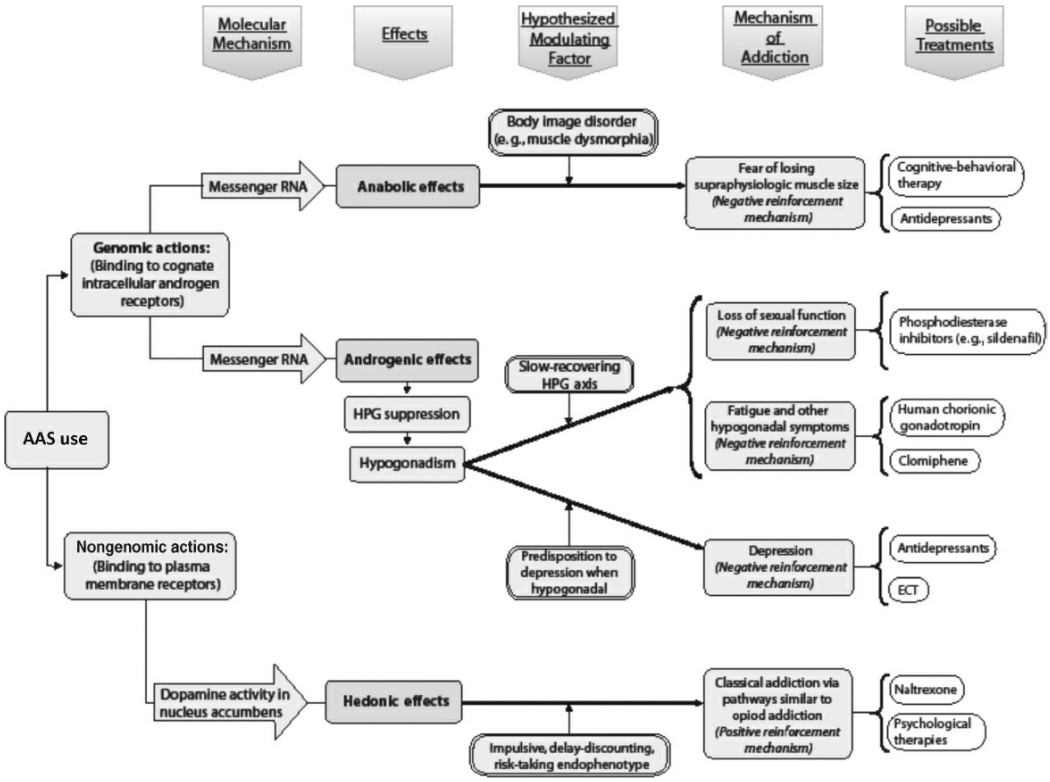

FIGURE 11.11–4. A model showing three hypothesized pathways by which anabolic steroid dependence may develop, together with potential treatment strategies to address each of these pathways. "HPG axis" indicates hypothalamic–pituitary–gonadal axis; "ECT" indicates electroconvulsive therapy.

In another recently proposed model of AAS dependence, the roles of

body-image disturbance and neuroendocrine factors are similarly acknowledged. This model postulates that AAS use and exercise initially produce desirable secondary gains (e.g., muscle gains and decreased body fat), but these behaviors put pressure on the homeostatic system as a result of increased metabolic/musculoskeletal allostatic load. This combines with increased hormonal allostatic load to contribute to physiological addiction, characterized in part by use of ancillary substances to reduce the side effects of allostatic load.

Course and Prognosis

No longitudinal studies of dependent AAS users have been published to date, in part because AAS-dependent individuals rarely see clinical treatment. One study estimated that AAS users accounted for less than 0.1 percent of patients entering treatment programs for substance abuse. The low number of documented treatment cases is striking when it is considered that among the 2.9 to 4.0 million American men who have used AAS, perhaps 1 million have experienced AAS dependence.

Treatment

Few descriptions and no controlled trials have been published of treatment for AAS-related disorders. Given the lack of published treatment studies, perhaps the best guidelines for approaching AAS treatment are those suggested in Figure 11.11–4, which summarizes each of the three proposed pathways to AAS dependence discussed above. First, body image concerns, such as muscle dysmorphia, may improve with cognitive behavioral therapies. Although the authors are aware of no formal trials of cognitive behavioral therapy in muscle dysmorphia per se, a growing literature suggests that cognitive behavioral approaches are effective in other forms of body dysmorphic disorder, and thus may also be helpful in muscle dysmorphia. Individuals able to address cognitive distortions about their muscularity may be more successful in managing anxiety about loss of muscle mass from discontinuing AAS.

Another modality of demonstrated efficacy for body dysmorphic disorder is treatment with serotonergic antidepressants. Although the authors are not aware of any studies with these drugs in muscle dysmorphia per se, it appears likely that they would show efficacy in such individuals, and might be of value in AAS users showing prominent and refractory obsessions about body image.

Second, it is imperative that patients be assessed and treated for AAS-induced hypogonadism. Men who experience loss of libido, erectile dysfunction, and possible major depression during AAS withdrawal will be strongly tempted to resume AAS use if the symptoms are not appropriately treated by an experienced endocrinologist. Typically, such patients will

require initial treatment with physiologic doses of exogenous testosterone after they stop using supraphysiologic doses of AAS. Then, to restore HPT function, treatment typically involves use of clomiphene, which stimulates the pituitary to resume production of LH and FSH. Simultaneously, patients can be treated with HCG to stimulate the testis to resume production of testosterone and spermatozoa. Then exogenous testosterone can gradually be tapered as the patient's own HPT axis begins to reawaken. Patients may also require treatment with phosphodiesterase inhibitors such as sildenafil to treat erectile dysfunction. Importantly, as noted earlier, hypogonadism may be prolonged and refractory in some patients, requiring persistent and aggressive treatment by an endocrinologist. In addition, some individuals will develop major depression during AAS withdrawal, and require treatment with antidepressant medications or even ECT (see below).

Finally, given the evidence for the hedonic effects of AAS, individuals may experience craving and other withdrawal effects comparable to those experienced by individuals attempting to achieve abstinence from classical drugs. It follows that dependent AAS users may benefit from psychological treatments of established efficacy in the treatment of disorders such as opioid dependence, including, for example, contingency management, behavioral couples therapy or behavioral family counseling, and supportive–expressive therapy. Couples therapy may be particularly indicated in cases where women are reporting distress about living with an AAS-dependent male partner.

ANABOLIC STEROID-INDUCED MOOD DISORDERS

Hypomanic or Manic Episodes

The effects of AAS on mood are highly variable, with a majority of individuals showing only modest mood changes, or no mood changes at all, even when taking markedly supraphysiologic doses of AAS. However, some individuals show substantial irritability, aggressiveness, hypomania, and frank mania associated with AAS use, and these psychiatric effects represent a significant public health issue associated with these drugs. Although AAS users have long recognized that these drugs could cause syndromes of anger and irritability, these syndromes were little recognized in the scientific literature until the late 1980s and 1990s. Since then, numerous observational field studies have described individuals who exhibited prominent hypomanic or even manic symptoms during AAS use. In a recent review of 12 such studies, 11 reported at least some such effects in association with AAS use, with an apparent association between the dosage of AAS used and the frequency of psychiatric effects. Specifically, in individuals using the equivalent of 300 mg of testosterone a week or less, psychiatric effects appear rare; in individuals taking intermediate dosages,

between 300 and 1,000 mg of testosterone equivalent per week, mood syndromes appear more frequent, albeit still uncommon; and in individuals taking the equivalent of more than 1,000 mg a week mood syndromes become quite common and are occasionally severe.

The association of AAS dosage with psychiatric effects is illustrated by examining individual studies. At the low end of the dosage range, for example, one study of 53 men examined individuals who were using a mean dose of only 318 mg of testosterone or equivalent a week. The men in this study exhibited essentially no psychiatric changes during use of AAS. By comparison, another study examined 71 men and 6 women who were using a mean dose estimated as somewhat more than 500 mg a week. The authors of this study noted that 1 (1.3 percent) of the subjects developed mania and 6 (7.8 percent) developed hypomania. One subject (1.3 percent) described major depressive disorder during AAS use, and 5 (6.5 percent) subjects reported major depressive disorder in association with AAS withdrawal. Three (3.9 percent) of the 77 users in this study reported a suicide attempt during a period of withdrawal from AAS use. At the high end of the dosage range, a third study examined 39 men and 2 women who had used a mean of 750 mg per week. In this study, five (12 percent) of the subjects reported a manic syndrome, and five (12 percent) reported psychotic symptoms at some time in association with AAS use. Interestingly, among the 8 subjects reporting manic or psychotic symptoms or both in this study, the average weekly dose was about 900 mg. A further illustration of the relationship between AAS dosage and the prevalence of psychiatric syndromes is provided in Table 11.11–2, which provides the frequency of hypomanic, manic, and major depressive syndromes reported among 88 users of AAS in one study where the users are grouped on the basis of the maximum dosage of AAS used. Psychiatric syndromes were almost absent at the lowest dosages, but affected nearly half of the individuals using dosages equivalent to 1,000 mg or more of testosterone a week.

Observational field studies, however, are subject to a number of methodological limitations common to retrospective studies of this nature. First, subjects may be vulnerable to recall bias because they are being asked to describe psychiatric symptoms that may have occurred well in the past. Second, subjects are generally taking drugs of unknown potency or authenticity obtained on the illicit market. Thus, assumptions regarding the nature and dosage of the drugs are necessarily crude at best. Third, selection bias may influence which subjects present for interviews in observational studies: subjects who have experienced prominent psychopathology in association with AAS use might be either more likely or less likely to participate. Fourth, expectational effects and influences from the subculture of the gymnasium may greatly influence subjects' reactions to AAS. For example, an individual lifting weights regularly in a "hard core"

gymnasium might be primed to expect increases in his irritability and aggression with AAS use, even if such changes do not actually occur.

Table 11.11–2.
Relationship of Weekly Steroid Dosage to Prevalence of Mood Disorders

No. (%) of Subjects Displaying Mood Disorders[a]

Weekly Steroid Dose[b] (No. of Users)	Manic Episode	Hypomanic Episode	Major Depressive Disorder	Total[c]
Low (*N* = 12)	0	1 (8)	0	1 (8)
Medium (*N* = 51)	0	5 (10)	3 (6)	7 (14)
High (*N* = 25)	4 (16)	3 (12)	7 (28)	11 (44)

[a]Significance of relationship between increasing doses of steroids and prevalence of mood disorders, $p = 0.003$; total, $p = 0.004$.

[b]Low represents less than 300 mg per week of testosterone or equivalent; medium, 300–1,000 mg per week; and high, more than 1,000 mg per week.

[c]Total is less than the sum of individual categories because some individuals displayed both a manic or hypomanic episode and a more depressive episode on association with steroids.

From Pope HG Jr, Katz DL. Psychiatric and medical effects of anabolic-androgenic steroid use: a controlled study of 160 athletes. *Arch Gen Psychiatry*. 1994;51:375, with permission. Copyright © 1997 American Medical Association. All rights reserved

Laboratory studies using AAS in normal volunteers avoid most of these limitations, but are subject to problems of their own. Specifically, most laboratory studies, including endocrinological investigations, physiological studies, and studies using AAS to treat various medical conditions, have used very modest doses of testosterone or other AAS, rarely exceeding the equivalent of 300 mg of testosterone a week. Furthermore, such studies have rarely attempted to measure psychiatric effects systematically. Not surprisingly, these studies have generally failed to note prominent psychiatric changes in their subjects. As a result, some such studies have erroneously concluded that AAS lack psychiatric effects, because they have inappropriately extrapolated from modest doses used under laboratory conditions, and fail to consider that illicit AAS users use far larger doses in the field. This error is similar to that of many earlier negative studies of the effects of AAS on muscle growth, again because the studies typically tested doses far lower than those actually employed by illicit AAS users.

We are aware of only four laboratory studies in which normal male volunteers were given dosages of at least 500 mg a week of testosterone or equivalent under placebo-controlled, double-blind conditions. These studies represent the only controlled investigations that have used doses of AAS more closely approximating those used by actual AAS abusers—although even these laboratory studies still do not approach the massive doses of AAS employed by some actual illicit users. Of 109 men receiving

such dosages in these four studies combined, 5 exhibited prominent hypomanic or manic syndromes during AAS exposure. These subjects displayed levels of symptomatology that might be expected to seriously impair normal social or occupational functioning; one study subject, while receiving methyltestosterone 240 mg per day under blinded conditions, even requested that he be placed in seclusion on the ward where the study was conducted. Notably, however, the great majority of the men in these four studies exhibited few behavioral changes with high dosages of AAS. Thus, it appears that most individuals taking AAS, even at fairly high dosages, will experience few psychiatric effects, whereas an occasional individual may develop a severe manic or hypomanic reaction. At present the mechanism of these reactions is unknown; they may represent an idiosyncratic phenomenon unrelated to premorbid psychiatric function.

Figure 11.11–5 shows changes in ratings of manic symptoms in three subjects who received testosterone cypionate in dosages up to 600 mg a week versus inert placebo, each for periods of 6 weeks in a double-blind, crossover design. Subjects varied widely in their response to testosterone. The subject on the left side of the figure exhibited virtually no change in manic ratings during either the testosterone or placebo periods; the subject in the middle displayed a modest rise in manic symptoms in association with testosterone, but not sufficient to significantly affect his social or occupational functioning. By contrast, the subject on the right displayed a marked manic reaction, approaching the range of patients hospitalized for the treatment of manic episodes, requiring the investigators to discontinue testosterone administration 1 week earlier than planned. The symptoms of mania resolved promptly within 2 weeks following discontinuation of testosterone.

It is difficult to extrapolate from these findings to estimate the frequency of prominent hypomanic or manic reactions among AAS users in the field. Illicit users may employ higher dosages than could ethically be administered in the laboratory; such users may take several different AAS together ("stacking"); they may use alcohol or other drugs while taking AAS; and of course they do not screen themselves with the same care as would be the case in a laboratory study. Finally, some users may become more sensitized to the psychiatric effects of AAS after using them for multiple courses over time; since most of the subjects in available laboratory studies were AAS-naive prior to the treatment, they may not be representative of actual AAS users in the field. For all these reasons, the crude rate of approximately 5 percent estimated above may represent a lower bound for the true prevalence of AAS-induced hypomanic and manic syndromes in the overall population of illicit users.

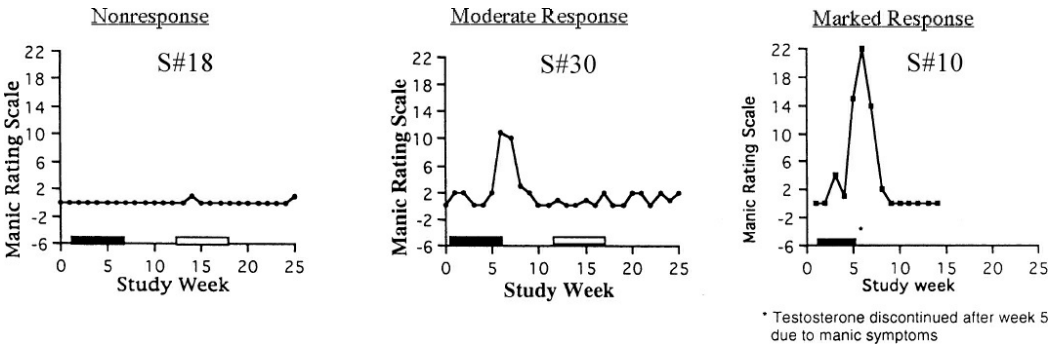

FIGURE 11.11–5. Examples of nonresponse, moderate response, and marked response to testosterone versus placebo on the Young Manic Rating Scale. The black bar represents the period of testosterone cypionate (Depo-Testosterone) administration in each subject. Testosterone was administered in gradually increasing doses (150 mg per week for the first 2 weeks, 300 mg per week for the next 2 weeks, and 600 mg per week for the final 2 weeks of the period). The white bar represents the period of placebo administration. A 6-week washout period was interposed between the two treatment periods to minimize carryover effects.

Major Depressive Episodes

AAS-induced depressive syndromes have generally not been reported in laboratory studies, but have been documented in field studies. Several reports of completed suicides associated with AAS withdrawal have also appeared. Usually it appears that a brief and self-limited syndrome of depression occurs upon AAS withdrawal, probably as a result of the depression of the HPT axis following exogenous AAS administration. Such syndromes might be expected to occur more often in the field than in laboratory settings because men typically take AAS for longer intervals than would occur in the laboratory, resulting in more pronounced neuroendocrine depression and shrinkage of the testes.

Not all depressive syndromes of this nature, however, are self-limited. Some reports have described episodes of depression that continued for months after discontinuing AAS, and which required treatment with fluoxetine (Prozac), tricyclic antidepressant agents, or ECT. Again, however, given the uncontrolled and retrospective nature of these observations, it is difficult to judge the true frequency of such syndromes in the field.

Mixed Episodes

Although the authors are not aware of specific published descriptions of mixed affective episodes associated with AAS use, the authors have witnessed many such cases in their anecdotal experience. Mixed episodes appear to be particularly frequent in individuals who have developed AAS dependence, and who are typically using AAS regularly with only brief drug-free intervals. Prolonged and relatively continuous AAS use of this type typically produces profound HPT suppression, so that the individual becomes markedly hypogonadal, and often seriously depressed, promptly

upon stopping AAS. Therefore, he will often quickly resume AAS to self-treat the depression, only to experience a recurrence of depression when he next decreases the dose or attempts to discontinue the drugs. In the authors' experience, such individuals sometimes develop "brittle" mood disorders, with a mixture of manic and depressive symptoms, or sharp swings between manic and depressive symptoms, often associated with prominent impulsivity, sometimes leading to suicide attempts. In some instances, such a suicide attempt may be the first event that brings the AAS user to the attention of clinicians.

Violence

A possible serious consequence of AAS-induced mood disorders, especially manic, hypomanic, or mixed episodes, may be violent or even homicidal behavior. Several published reports have anecdotally described individuals with no apparent history of psychiatric disorder, no criminal record, and no history of violence, who committed violent crimes, including murder, while under the influence of AAS. In a number of cases, AAS use has been cited in criminal trials as a possible mitigating factor in the defense of such individuals. One study of 133 sequential, newly incarcerated male felons found that 3 (2.3 percent) reported committing violent crimes while under the influence of AAS. Another recent paper mentions six cases of AAS-induced criminal behavior seen by the authors, including three homicides and three violent assaults. The authors reported that the men in all cases apparently experienced psychotic symptoms, and that the mental status of all six perpetrators cleared within 2 months after discontinuing AAS; all reported a clear memory of the violent act and of their delusional thinking at the time. Forensic researchers in Sweden have described several similar cases of AAS users, many with no prior history of criminality or violence, who committed violent offenses, including homicide, in association with AAS use.

Finally, another study suggests that violence associated with AAS is frequently directed at women. This study compared 23 male AAS users with 14 male nonusers, using the Conflict Tactics Scales to assess users' relationships with their girlfriends or spouses. Significant differences emerged on several of these scales between users on drugs and nonusers and between users on drugs and the same users when off drugs.

A Swedish forensic research group has recently used epidemiological techniques to assess possible links between AAS use and violence. For example, these investigators found that deceased AAS users were significantly more likely to have died from homicide or suicide than a comparison group of deceased amphetamine and/or heroin users. These investigators also found that deceased AAS users, during the last 3 years of their lives, were more likely to have been sentenced for violent crimes, but less likely to have been sentenced for crimes against property, than

deceased individuals who tested positive for other drugs, but negative for AAS. Another study from the same group showed that prisoners testing positive for AAS were significantly less likely to have been convicted of a crime against property, but significantly more likely to have been convicted of a weapons offense, than AAS-negative prisoners (using an analysis that excluded cases referred from substance abuse centers). Although these findings do not permit causal inferences in specific cases, they appear consistent with the findings described in previous paragraphs.

ANABOLIC STEROID-INDUCED PSYCHOTIC DISORDER

Psychotic symptoms are rare in association with AAS use, but have been described in a few cases, primarily in individuals who were using the equivalent of more than 1,000 mg of testosterone a week. Usually these symptoms have consisted of grandiose or paranoid delusions, generally occurring in the context of a manic episode, although occasionally occurring in the absence of a frank manic syndrome. In most cases, the reported psychotic symptoms have disappeared promptly (within a few weeks) after the discontinuation of the offending agent, although temporary treatment with antipsychotic agents was sometimes required. As with manic reactions to AAS, the mechanism for these seemingly idiosyncratic psychotic reactions is unknown. It is interesting to note, however, that the chemically related family of hormones, corticosteroids, also produces idiosyncratic manic and psychotic symptoms in occasional individuals, while creating few psychiatric effects in the majority of patients. The mechanism of action of these idiosyncratic effects with corticosteroids is also unknown.

ANABOLIC STEROID-RELATED DISORDER NOT OTHERWISE SPECIFIED

Anxiety Disorders

Occasional anecdotal reports have described decreases in symptoms of anxiety disorders, such as panic disorder and social phobia, during AAS use, and exacerbations of these disorders during AAS withdrawal. Conversely, some individuals report increases in anxiety symptoms while taking AAS. However, no systematic data are available on these entities.

Associated Substance Abuse or Dependence

As mentioned earlier, AAS use may be associated with use of other ergogenic drugs, especially among the minority of AAS users who are competitive athletes. Among the larger number of noncompetitive “street” AAS users, extensive use of illicit drugs is common, as described in detail above. Indeed, individuals with a history of AAS use may first come to clinical attention when they require treatment for dependence upon alcohol or illicit drugs such as opiates, cocaine, or cannabis. For example, in one

recent study of 223 men admitted to a substance abuse treatment unit, 29 were found to have a past history of AAS use, but this history was documented in physicians' admission evaluations in only four of the cases. Nineteen of the 29 men were interviewed in detail, and five were found to exhibit a prior history of AAS dependence. As this finding illustrates, AAS use is often undetected by clinicians, despite its potential importance in causing the various psychiatric syndromes described in the above sections.

Mr. A. is a 26-year-old, single white man. He is 69 inches tall and presently weighs 204 pounds with a body fat of 11 percent. He reports that he began lifting weights at age 17, at which time he weighed 155 pounds. About 2 years after beginning his weightlifting, he began taking AAS, which he obtained through a friend at his gymnasium. His first "cycle" (course) of AAS, lasting for 9 weeks, involved methandienone, 30 mg orally a day, and testosterone cypionate, 600 mg intramuscularly a week. During these 9 weeks, he gained 20 pounds of muscle mass. He was so pleased with these results that he took five further cycles of AAS over the course of the next 6 years. During his most ambitious cycle, approximately 1 year ago, he used testosterone cypionate, 600 mg per week; nandrolone decanoate (Deca-Durabolin), 400 mg a week; stanozolol (Winstrol), 50 mg a day; and oxandrolone (Anavar), 10 mg a day.

During each of the cycles Mr. A. has noted euphoria, irritability, and grandiose feelings. These symptoms were most prominent during his most recent cycle, when he felt "invincible." During this cycle, he also noted a decreased need for sleep, racing thoughts, and a tendency to spend excessive amounts of money. For example, he impulsively purchased a $2,700 stereo system when he realistically could not afford to spend more than $500. He also became uncharacteristically irritable with his girlfriend, and on one occasion put his fist through the side window of her car during an argument—an act inconsistent with his normally mild-mannered personality. After this cycle of AAS ended, he became mildly depressed for about 2 months.

Mr. A. smoked marijuana almost daily during his last 2 years of high school and continues to smoke at least twice a week. He has also used hallucinogens on about 50 occasions, and has used cocaine on about 200 occasions over the last 8 years. He has also used a number of drugs to lose weight in preparation for bodybuilding contests. These include ephedrine, amphetamine, triiodothyronine, and thyroxin. Recently, he has also begun to use the opioid agonist-antagonist nalbuphine (Nubain) intravenously to treat muscle aches from weightlifting. He also uses oral opioids, such as controlled-release oxycodone (OxyContin), at least once a week. He uses oral opioids sometimes to treat muscle aches, but often simply to get high. He reports that use of nalbuphine and other opioids is widespread among other AAS users of his acquaintance.

Mr. A. exhibits characteristic features of muscle dysmorphia. He checks his appearance dozens of times a day in mirrors, or when he sees his reflection in a store window or even in the back of a spoon. He becomes anxious if he misses even 1 day of working out at the gym, and acknowledges that his preoccupation with weightlifting has cost him both social and occupational opportunities. Although he has a 48-inch chest and 19-inch biceps, he has frequently declined invitations to go to the beach or a swimming pool for fear that he would look too small when seen in a bathing suit. He is anxious because he has lost some weight since the end of his previous cycle of AAS and is eager to resume another cycle of AAS in the near future.

Mr. B. is a 37-year-old personal trainer and body builder. He describes a normal childhood and adolescence, although he was periodically mildly depressed. By the age of 15, he began to drink alcohol regularly, and at age 18 used cocaine several days per week for about a year. He describes "mood swings" starting in his late adolescence, but these were not severe enough to compromise his social or academic functioning. He obtained a master's degree in physical therapy from a local community college at age 21.

Starting in his late teens, Mr. B. became "obsessed" with bodybuilding and with his

physique. At age 21, he began his first 16-week "cycle" of injectable testosterone cypionate up to 600 mg per week, plus oral stanozolol (Winstrol) up to 25 mg per day. After stopping the cycle, he became mildly depressed, with fatigue, decreased sex drive, and lack of enthusiasm for going to the gym. He was also distressed to notice that he regained some body fat and lost some muscle during the weeks after stopping the cycle. He also began to consume alcohol more regularly at this time, including sneaking drinks at work.

Four months later, he started a second cycle of AAS with a larger "stack" of drugs, including injectable testosterone cypionate, up to 600 mg per week; injectable nandrolone decanoate (Deca-Durabolin) 400 mg per week; and oral oxymetholone (Anadrol) 50 mg per day. Although Mr. B. had intended to stop the cycle after 16 weeks, he was so gratified by his muscle gains that he decided to extend the cycle to 30 weeks. Upon stopping, he again experienced depression, despite injecting illicit HCG in an attempt to prevent hypogonadism. He drank alcohol heavily and again began to use cocaine almost daily during this period.

Mr. B. continued to take cycles of AAS regularly for the next 5 years; he estimates that he was on the drugs "about two thirds of the time" from age 22 to 27. A. During his cycles he felt "high" and very sexually active; in between the cycles he felt increasingly depressed and eager to resume taking AAS. At the age of 27, Mr. B. got married, and his wife was eager to have a child. An endocrinologist found that Mr. B. was sterile as a result of the suppression of sperm cell production due to his AAS use. To please his wife, Mr. B. stopped AAS entirely and managed to remain off of AAS for more than 2 years. His sperm count finally rose to the normal range 7 months after stopping AAS; soon thereafter his wife became pregnant. The couple had a son, followed a year and a half later by a daughter. Mr. B. used virtually no illicit drugs, and only occasional alcohol, during this period.

Shortly after his wife became pregnant with the second child, Mr. B. resumed using AAS despite his wife's objections. Once back on AAS, he felt self-confident and proud of his muscle gains; he became sexually promiscuous. He reports that his promiscuity contributed to his divorce from his wife 2 years later. Each time upon stopping AAS, he became markedly depressed; he was started on fluoxetine during one such period of AAS withdrawal at the age of 31. Fluoxetine was initially quite helpful, but it did not deter him from going back onto a cycle of AAS within a few months.

Over the next 6 years, Mr. B. continued to take AAS about "80% of the time," by his account —but from age 35 to age 37, he used AAS continuously, remaining on a modest dose of testosterone cypionate even during the spaces between the cycles. He described increasingly "brittle" mood swings, with periods of euphoria, irritability, and increased self-confidence alternating with sudden dips into depression with suicidal ideation. At times, Mr. B described a mixture of affective symptoms where he would experience euphoria, irritability, and suicidal thoughts almost simultaneously. Over the course of these years, his physician tried several other antidepressants (citalopram, venlafaxine, and duloxetine), plus several concomitant medications for mood stabilization (lithium carbonate, valproate, and lamotrigine), with inconsistent benefit—partially because Mr. B. would often become noncompliant. Also, Mr. B. gradually relapsed into drinking alcohol almost daily, and began to use increasing amounts of controlled-release oxycodone (OxyContin) to self-treat his dysphoric moods. Mr. B. also became increasingly impulsive, and on one occasion took a substantial overdose of citalopram, necessitating a 3-day hospitalization on a psychiatric unit.

Following this hospitalization, Mr. B. became more consistently compliant with a regimen of duloxetine and lithium carbonate. He continued to be highly focused on weightlifting and his body appearance. However, he still used oxycodone frequently and continued to complain of sharp "dips" of severe depression. During one of these "dips," shortly after tapering down from high-dose AAS to his usual "maintenance" level of testosterone cypionate, he made a second, nearly successful suicide attempt with a combination of illicitly obtained oral opiates. He was again hospitalized, treated with ECT, and responded well. A cardiac evaluation during this hospitalization, using Doppler echocardiography, revealed evidence of dilated cardiomyopathy judged likely attributable to AAS.

Mr. B. has resolved to not resume taking AAS, but 6 months after hospital discharge, he is already contemplating going back on a low dose of AAS because he is distressed by his deteriorating physique.

Mr. B.'s case exemplifies several AAS-related disorders, including apparently AAS-induced bipolar disorder, AAS withdrawal, and AAS dependence. In addition to physiological

withdrawal, Mr. B. met several other diagnostic criteria for AAS dependence: taking more AAS than originally intended, resuming use after several attempts to stop completely, requiring increasing doses of AAS (tolerance), and continuing use despite knowledge of adverse consequences such as divorce. Another important aspect of this case was concurrent abuse of or dependence on alcohol, opioids, and cocaine. All four classes of substances share a final common pathway of dopamine augmentation; in addition, alcohol, opioids, and AAS have effects on the endogenous opioid system; and alcohol and AAS both also act as GABA receptor modulators. Taking these drugs in combination typifies the loss of control commonly seen with addiction, but also may reflect an attempt to enhance the rewarding effects of AAS (dopamine and opioid systems) and analgesic aspects of AAS (opioid system), while minimizing some of the hyperactivating effects of AAS (via the antianxiety/sedative effects of GABA and opioid systems).

References

*Bhasin S, Storer TW, Berman N, et al. The effects of supraphysiologic doses of testosterone on muscle size and strength in normal men. *New Engl J Med*. 1996;335:1–7.

Bjørnebekk A, Walhovd KB, Jørstad ML, Due-Tønnessen P, Hullstein IR, Fjell AM. Structural brain imaging of long-tem anabolic-androgenic steroid users and nonusing weightlifters. 2016. *Biological Psychiatry*. Epub before print. Available at http://dx.doi.org/10.1016/j.biopsych.2016.06.017

*Brower KJ. Anabolic steroid abuse and dependence in clinical practice. *Phys Sportsmed*. 2009;37:131–140.

Carrillo M, Ricci LA, Melloni RH. Glutamate-vasopressin interactions and the neurobiology of anabolic steroid-induced offensive aggression. *Neuroscience*. 2011;185:85–96.

*Daly RC, Su TP, Schmidt PJ, Pickar D, Murphy DL, Rubinow DR. Cerebrospinal fluid and behavioral changes after methyltestosterone administration: preliminary findings. *Arch Gen Psychiatry*. 2001;58:172–177.

Estrada M, Varshney A, Ehrlich BE. Elevated testosterone induces apoptosis in neuronal cells. *J Biol Chem*. 2006;281(35):25492–25501.

*Frye CA. Some rewarding effects of androgens may be mediated by actions of its 5alpha-reduced metabolite 3alpha-androstanediol. *Pharmacol Biochem Behav*. 2007;86:354–367.

Garevik N, Strahm E, Garle M, et al. Long term perturbation of endocrine parameters and cholesterol metabolism after discontinued abuse of anabolic androgenic steroids. *J Steroid Biochem Molec Biol*. 2011;127:295–300.

Givens ML, Deuster PA, Kupchak BR. CHAMP symposium on androgens, anabolic steroids, and related substances: what we know and what we need to know. *Mil Med*. 2016;181(7):680–686.

Graham MR, Ryan P, Baker JS, et al. Counterfeiting in performance- and image-enhancing drugs. *Drug Testing and Analysis*. 2009;1(3):135–142.

Gruber AJ, Pope HG Jr. Psychiatric and medical effects of anabolic-androgenic steroid use in women. *Psychother Psychosom*. 2000;69:19–26.

Hartgens F, Kuipers H. Effects of androgenic-anabolic steroids in athletes. *Sports Med (Auckland, NZ)*. 2004;34(8):513–554.

Hildebrandt T, Lai JK, Langenbucher JW, Schneider M, Yehuda R, Pfaff DW. The diagnostic dilemma of pathological appearance and performance enhancing drug use. *Drug Alcohol Depend*. 2011;114(1):1–11.

Ip EJ, Barnett MJ, Tenerowicz MJ, Perry PJ. The Anabolic 500 survey: characteristics of male

users versus nonusers of anabolic-androgenic steroids for strength training. *Pharmacotherapy*. 2011;31(8):757–766.

Kanayama G, Brower KJ, Wood RI, Hudson JI, Pope HG. Treatment of anabolic-androgenic steroid dependence: emerging evidence and its implications. *Drug Alcohol Depend*. 2010;109:6–13.

Kanayama G, DeLuca J, Meehan WP, et al. Ruptured tendons in anabolic-androgenic steroid users: a cross-sectional cohort study. *Am J Sports Med*. 2015;43(11):2638–2644.

Kanayama G, Hudson JI, Pope Jr HG. Culture, psychosomatics and substance abuse: the example of body image drugs. *Psychother Psychosom*. 2012;81(2):73–78.

Kanayama G, Kean J, Hudson JI, Pope HG. Cognitive deficits in long-term anabolic-androgenic steroid users. *Drug Alcohol Depend*. 2012;130:208–214.

Klotz F, Garle M, Granath F, Thiblin I. Criminality among individuals testing positive for the presence of anabolic androgenic steroids. *Arch Gen Psychiatry*. 2006;63:1274–1279.

Kouri EM, Pope HG Jr, Katz DL, Oliva PS. Fat-free mass index in users and non-users of anabolic-androgenic steroids. *Clin J Sport Med*. 1995;5:223–228.

Parssinen M, Seppala T. Steroid use and long-term health risks in former athletes. *Sports Med* 2002;32:83–94.

Pope H, Kanayama G, Hudson J. Risk factors for illicit anabolic-androgenic steroid use in male weightlifters: a cross-sectional cohort study. *Biol Psychiatry*. 2012;71:254–261.

*Pope HG Jr., Kouri EM, Hudson JI. The effects of supraphysiologic doses of testosterone on mood and aggression in normal men: a randomized controlled trial. *Arch Gen Psychiatry*. 2000;57:133–140.

*Pope HG, Wood R, Rogol A, Nyberg F, Bowers L, Bhasin S. Adverse health consequences of performance-enhancing drugs: an Endocrine Society scientific statement. *Endocrine Rev*. 2014;35:341–375.

Rasmussen JJ, Selmer C, Østergren PB, et al. Former abusers of anabolic-androgenic steroid exhibit decreased testosterone levels and hypogonadal symptoms years after cessation: a case-control study. *PLoS One*. 2016;11(8):e0161208.

Rubinow DR, Schmidt PJ. Androgens, brain, and behavior. *Am J Psychiatry*. 1996;153(8):974–984.

Skarberg K, Nyberg F, Engstrom I. Multisubstance use as a feature of addiction to anabolic-androgenic steroids. *Eur Addict Res*. 2009;15(2):99–106.

Su TT, Pagliaro M, Schmidt PJ, Pickar D, Wolkowitz OM, Rubinow DR. Neuropsychiatric effects of anabolic steroids in male normal volunteers. *JAMA*. 1993;269:2760–2764.

*Thiblin I, Petersson A. Pharmacoepidemiology of anabolic androgenic steroids: a review. *Fundamental Clinical Pharmacol*. 2005;19(1):27–44.

*Wood RI. Anabolic-androgenic steroid dependence? Insights from animals and humans. *Front Neuroendocrinology*. 2008;29(4):490–506.

12

Schizophrenia and Other Psychotic Disorders

▲ 12.1 Introduction and Overview

CAROL A. TAMMINGA, M.D.

Schizophrenia is arguably one of the most puzzling yet disabling of all brain diseases, with its severe and persistent psychotic manifestations accompanied by variable cognitive dysfunction and profound psychosocial impairment. Its disease onset, at least of the psychotic manifestations, occurs at an unusual time of life in late adolescence/early adulthood. Its symptoms are often undetectable in early life, are at their worst during child-bearing years, and often improve, to a degree, with aging. People with psychotic illnesses, in general, suffer greatly throughout life and medical need worldwide continues to be very high. Fortunately, in the pages of this section, the broad aspects of this condition are described and discussed, as far as they are known, by experts in the field.

Substantial progress in understanding and treating schizophrenia will depend on discovering the neural mechanisms of the elements of the condition, one of those being the mechanism for psychosis. This understanding is necessary for developing drug targets, novel therapeutics, and for biologically based disease definitions. It is hard to conceive of the exact neural mechanisms whose dysfunction could interfere so extensively with such broad cognitive functions as does psychosis. This section lays out current perspectives of schizophrenia presentation, treatment, and biology that include much of what is known about its disease manifestations today and brain dysfunctions in the illness. What provides clear hope of progress for the future is not only the quality of the tools that are available to study brain function but, in particular, the wealth of basic neurobiology that there is available and the extent to which this understanding of basic brain mechanisms translates to human diseases. In this section, the curious reader will find the facts of psychosis presentation and the sum of the manifestations and biomarkers that any model needs to take into account

in the formulation of a disease model. Of course, blatant heterogeneity exists in the different psychotic illnesses, with nearly as much heterogeneity within a diagnosis as between, along with a lack of distinguishing biological characteristics. Several scientists represented in this section speculate on the use of alternative strategies for classification, attempting to increase biological homogeneity within disease-like clusters to advantage mechanistic discovery.

It is always poignant to review the long history of psychotic illness, including its most severe variant, schizophrenia. Descriptions of schizophrenia can be found in works of literature from earliest times. Schizophrenia-like symptoms are described in individuals labeled as seers and prophets, as well as witches and devils. The perspective of schizophrenia being a human brain disease did not develop until the 19th century. Then, "humane" treatment was prescribed for "insanity," an approach that advocated care, protection, and human understanding for those afflicted. It was not until the middle of the 20th century that antipsychotic drug treatments became widely available. Now in the early 21st century, there have been two, maybe three, generations of antipsychotic medications developed, several known risk genes discovered, an evolving anatomy, and several regional human brain-based models for symptoms manifestations of cognitive dysfunction (prefrontal cortex) and psychosis (hippocampus). Meanwhile, the focus of inquiry into the illness has moved from conventional diagnostic criteria to dimensions of dysfunction, and there has developed a fundamental understanding of the neural basis of complex functional brain diseases. Therefore, it might be optimistically predicted that it will eventually be possible to describe the genes, cellular and molecular mechanisms, and rational treatments for psychosis manifestations over the next several decades.

The articulation of a modern biological hypothesis for schizophrenia began with dopamine. The pivotal observation that drugs that diminish dopamine-mediated neurotransmission in brain reduce psychotic symptoms was made only in the mid-20th century. The first antipsychotic drug (chlorpromazine) was discovered serendipitously but with a strong clinical rationale and astute observation. The experimental result that implicated the action of chlorpromazine connected to dopamine receptor blockade was a hypothesis that fueled decades of drug discovery and was the basis for a related etiological hypothesis that dopamine was indeed the core pathophysiology of schizophrenia. The speculative extension of this well-supported treatment hypothesis has been harder to prove and to establish as the basis for schizophrenia pathophysiology, although ideas are still being tested. Although data have consistently supported the dopamine mechanism of antipsychotic drug action, support for the dopamine hypothesis of pathophysiology has been inconsistent.

At the same time, the understanding of the pharmacology and basic

biology of human cognition and its dysfunctions has evolved into a very complex picture, one unlikely to be explained simply as overactivity of dopamine. A component of cognition like declarative memory, which have been intensively studied since Henry Moliason lost declarative memory function, is understood now to be served by alterations in synaptic strength, postsynaptic calcium interactions with receptor trafficking, and epigenetic remodeling in hippocampus and the memory neocortex. An alteration in this process of memory physiology could be at the basis of a model of psychosis. Many speculations have emerged that multiple manifestations of psychotic illness may be associated with distinct pathophysiologies, for example, one for psychosis, another for cognitive dysfunction. Anyone interested in the pathophysiological basis of psychiatric conditions will have to gain an understanding of normal brain processes as already being articulated within neuroscience laboratories around the world. Full attention will have to be directed to these discoveries in basic neuroscience to understand modern articulations of brain-based pathophysiologies for psychosis and cognitive dysfunction.

In clinical research in schizophrenia, there has been a willingness to abandon conventional diagnostic categories, experimentally and hypothetically, for the sake of examining alternate views of pathophysiology, etiology, and treatment. This has involved a focus on pathophysiology of *dimensions* of dysfunction for postulating mechanisms and exploring treatments. Not that the dimension will be the final explanatory element, but that it will allow a resorting among similar disease manifestations to gain biological homogeneity. Dimensions of normal cerebral function have been proposed, connected with known neural systems, a dysfunction or perversion of which would generate psychopathology. Specifically, the independence of symptom domains such as psychosis, cognitive dysfunction, affective disturbance, and negative symptoms is now being explored in a Research Domain Criteria (RDoC) framework. The implication of this formulation, which is now being explored, is that each of these composite symptom complexes will have its own etiology, mechanism, and selective treatment and can be treated independently and individually. Personalized medicine, as well as new disease formulations, is likely to spring from these various approaches.

The field of schizophrenia practice and research needs to gain understanding in some obvious areas. First, effective diagnostic and treatment algorithms for the early detection and treatment of psychotic young adults need to be established. Effective medications are necessary, even when optimized across each individual person, but they are by no means sufficient. Psychotic illness, especially at its inception or when it threatens to break through, needs broad, diverse, and comprehensive treatment. This kind of comprehensive treatment extends beyond antipsychotic medication to include cognitive remediation approaches,

psychosocial treatments, behavioral therapies, family training, and community involvement. The newest study in the treatment of early psychosis, the RAISE study, found that implementing all of these treatments early on in disease manifestations generates better outcomes sooner in young people. Treating a psychotic illness such as schizophrenia "piecemeal" generates only partial gains and falls short of the success that comes with a comprehensive plan allowing potentiated gains.

It would certainly be useful if more was known about the mechanism of action of the effective drugs that are used for psychosis, all of them being dopamine receptor antagonists. What actually do these antidopaminergic drugs do in brain that quiet the storms of psychotic symptoms? This would be an approach to discovering disease pathophysiology. One of the single issues that affect psychotic illnesses most adversely is the lack of disease understanding. What the field needs is a simple, direct, and concrete presentation of the mechanism(s) of psychosis to combat stigma, direct treatment development, and on which to base behavioral and cognitive treatment approaches.

What if there were biological formulations of psychotic diseases that were confirmed and demonstrated? How different would that be for the serious mental illnesses that psychiatrists treat? To answer this question, medical conditions that changed after full disease understanding could be examined. For example, there was the disease called "dropsy" in the late 19th century, which was a morbid edema that progressed rapidly and caused extreme morbidity and finally death. Its determinants and treatments have been progressively understood and more precisely treated by discoveries in medicine throughout the 20th and 21st centuries. Dropsy can be explained by a plethora of cardiac, pulmonary, renal, and metabolic conditions, some genetic, others, developmental, others metabolic, structural. Its association with specific organ failures focused attention on a component of the dropsy cases (all phenomenologically quite similar, if not identical) on dysfunctions of the particular organ systems. Some treatments are specific for the diseased organ, some medications specific for the edema. Now edema rarely becomes morbid anymore. And our increased understanding of *prophylaxis* directed toward cardiac, metabolic, and diseases has improved health and helped avoid dropsy-like conditions. These kinds of disease structures and their determinants need to be found in cognitive and affective brain diseases so that the same kind of understandings can be gained for psychosis brain health.

References

Carlsson A, Lindquist L. Effect of chlorpromazine or haloperidol on formation of 3-methoxytyramine and normetanephrine in mouse brain. *Acta Pharmacologica et Toxicologica*. 1963;20:140–145.

Clementz BA, Sweeney JA, Hamm JP, et al. Identification of distinct psychosis biotypes using

brain-based biomarkers. *Am J Psychiatry*. 2016;173:373–384.

Hyman SE. Revolution stalled. *Sci Transl Med*. 2012;4(155):155cm11.

Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010;167:748–751.

Lewis DA, Curley AA, Glausier JR, Volk DW. Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. *Trends Neurosci*. 2012;35:57–67.

Li W, Ghose S, Gleason K, et al. Synaptic proteins in the Hippocampus indicative of increased neuronal activity in CA3 in schizophrenia. *Am J Psychiatry*. 2015;172:373–382.

Tamminga CA. Development in somatic treatments: introduction. *Am J Psychiatry*. 1994;151:216–219.

▲ 12.2 Phenomenology of Schizophrenia

STEPHEN F. LEWIS, M.D., RODRIGO ESCALONA, M.D., AND SAMUEL J. KEITH, M.D.

Schizophrenia continues to be defined and diagnosed as a psychiatric disorder solely based on its clinical phenomenology. Efforts to include the use of biomarkers to make the diagnosis and predict treatment response or course of illness have not been successful thus far. The concept of schizophrenia as a distinct clinical entity is a recent historical development. However, rich descriptions of psychotic disorders can be found through centuries of written history, with early Greek physicians documenting cases that resemble schizophrenia in its current form. Psychosis emerged only as a medical condition worthy of scientific study and treatment during the 18th century in Europe. At that time, the concept of psychosis included a large undifferentiated group of disorders such as psychotic symptoms secondary to medical conditions such as general paresis and syphilitic insanity, as well as primary or "functional" psychoses such as affective psychosis and schizophrenia. It would take at least another half century for a more consistent and systematic classification of psychotic disorders to appear and set the basis for modern psychiatric diagnoses.

DIAGNOSTIC CRITERIA

History

Emil Kraepelin (1856 to 1926) pioneered the psychiatric nosology that would set the foundation for the development of modern psychiatric diagnostic systems. He first separated, from the broad group of psychosis, the organic psychoses such as syphilitic insanity and general paresis, and then most importantly, he described two major groups of primary psychotic disorders: The *manic depressive psychosis* and *dementia praecox* (dementia of the young, a term coined earlier by Morel). He used

the long-term prognosis and course of illness as principal features to define the two distinct categories. The assumption at that time was that all affective psychosis had a nondeteriorating course as opposed to dementia praecox, which was deemed as deteriorating and irreversible. He believed that dementia praecox was a loss of the inner unity of the activities of intellect, emotion, and volition. He included in the concept of *dementia praecox* a broad group of disorders known separately at that time, such as paranoia, catatonia, and hebephrenia, and defined them as subtypes of the same illness.

In 1911, the term *schizophrenia,* as the disease is known today, was introduced by Eugen Bleuler (1857 to 1939). He proposed the name to denote a "splitting" of psychic functions, which he considered to be the core feature of the illness. He was less concerned with the course of illness than Kraepelin had been and much more concerned with the understanding of the psychological mechanisms underlying the disease process. His description of the illness included primary and secondary symptoms. His four primary symptoms ("the Four As") were *abnormal associations, autistic behavior and thinking, abnormal affect,* and *ambivalence*. Bleuler considered the loss of harmonic coordination among thought, emotion, and behavior to be the hallmark of the illness. Bleuler considered hallucinations, delusions, social withdrawal, and diminished drive as secondary manifestations of the illness that depended on the adaptive capacity of the individual and environmental circumstances. Thus, according to Bleuler, the most obvious and striking manifestations of schizophrenia were only "accessory symptoms." In addition, he saw the symptoms of schizophrenia as a continuum with normal behavior, a concept that led to major diagnostic controversies among psychiatrists around the world for the next several decades, until the recent formalization of diagnostic criteria.

Kurt Schneider (1887 to 1967) in his classification of thought disorders attempted to make the diagnosis of schizophrenia more reliable by identifying a group of symptoms that were the most characteristic of the illness. His so-called first-rank symptoms included various forms of hallucinations, thought withdrawal, thought insertion, thought broadcasting, delusional perceptions, and experiencing feelings and actions as made or influenced by external agents (Table 12.2–1). The presence of one of these symptoms, in the absence of intoxication, brain injury, or clear affective illness, was sometimes taken as sufficient for making the diagnosis of schizophrenia.

The different conceptualizations of the illness and diagnostic criteria used in clinical practice and research led to wide disparities in the rates of diagnosis of schizophrenia among clinicians and researchers around the world. This problem ignited a search for agreements and conventions to facilitate communication and define more formal and reliable diagnostic

criteria giving birth to modern nosology.

Modern Diagnostic Criteria

In the United States, the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) of the American Psychiatric Association (APA), in its second edition (DSM-II), was heavily influenced by Bleularian criteria to make the diagnosis of schizophrenia. The predominance of psychological and psychosocial theories in the field at the time, such as the theories of Adolph Mayer, Harry Stack Sullivan, and the Freudian followers, was fertile ground for the adoption of a view of schizophrenia according to Bleuler. The consequences of adopting such concepts led to broad and arbitrarily applied criteria to diagnose schizophrenia, sometimes making the rate of diagnosis rather high. Many conditions now classified elsewhere were included in the diagnosis of schizophrenia according to the DSM-II. As a consequence, studies performed during the 1960s and 1970s in the United States and in Europe showed that the reliability of diagnosis locally or across countries was very low. In the United States, the more inclusive diagnostic criteria for schizophrenia led to rates that were much higher than in Europe, where many more patients were instead diagnosed with affective psychosis. Through the middle of the 20th century, the rate of schizophrenia as an admission diagnosis for psychiatric facilities in London stayed similar over a 40-year period, whereas the diagnosis of schizophrenia appeared to be increasing as an admission diagnosis in New York City, with no clear explanation for the differences. The United States–United Kingdom Cross National Project study of hospitals in London and New York showed that approximately 60 percent of the admissions in New York State hospitals were diagnosed with schizophrenia, whereas in London hospitals, admissions with the diagnosis of schizophrenia were about half that rate. When standardized diagnostic interviews were applied in these settings, the rates of schizophrenia in New York decreased to levels that matched the rates in London. These findings triggered the need to define more standardized diagnostic criteria on both sides of the Atlantic. A later groundbreaking study that propelled the adoption of explicit diagnostic criteria was the International Pilot Study of Schizophrenia (IPSS) sponsored by the World Health Organization (WHO). One thousand two hundred patients with psychotic disorders from nine countries were interviewed using a structured diagnostic interview, the Present State Examination (PSE). This study showed that schizophrenia manifested itself in similar ways across very diverse cultures, with similar rates of hallucinations, delusions, first-rank symptoms, and affective symptoms among people with schizophrenia in developing and developed countries. The findings from this study reaffirmed the belief that useful and uniform formal criteria could be developed and would contribute to better communication among clinicians and researchers alike.

Table 12.2–1.
First-Rank Symptoms of Kurt Schneider

Audible thoughts	Auditory hallucinations of a person's voice being spoken aloud
Voices arguing or discussing	Auditory hallucinations of two or more voices arguing or discussing, usually about the person experiencing the hallucination
Voices commenting on patient's actions	Auditory hallucinations commenting on a person's behaviors
Somatic passivity	Tactile or visceral hallucinations that are imposed by some external agent; can be combinations of different somatic hallucinations
Thought withdrawal	Sensation of thoughts being actively removed from a person's mind
Thought insertion	Thoughts inserted into a person's mind by some external agent
Thought broadcasting	The sense that a person's thoughts are experienced as real phenomena by others—the thoughts are made audible or may be experienced by others through telepathy
Made feelings	Feelings that are not a person's own are imposed on that person by an external agent
Made impulses or drives	An impulse for action is imposed on a person by some external agent
Made volitional acts	A person's actions are from and are controlled by an external agent; the person is a passive participant in the action
Delusional perception	A perception that has a unique and idiosyncratic meaning for a person, which leads to an immediate delusional interpretation

The changes between DSM-II to DSM-III in the diagnostic criteria of schizophrenia represent the biggest conceptual departure from previous diagnostic classification systems and set the foundation of current nomenclature. This was a major break with decades of American and European nosological tradition. The result was a much more narrow definition of the disorder. This major conceptual change was possible due to the efforts published in 1972 by the Washington University psychiatry research group in St. Louis, Missouri, that introduced a diagnostic classification system aiming at greater validity and reliability of diagnoses. These criteria, known as the Feighner criteria after the first author of the initial report, were rigorous in trying to define homogeneous diagnostic groups. The principles underlying this classification system were (1) clinical description, (2) laboratory study, (3) delimitation from other disorders, (4) follow-up studies, and (5) family studies. The next major step toward current diagnostic criteria was the introduction of the Research Diagnostic Criteria (RDC) in 1978. This effort borrowed many concepts from the Feighner criteria and the Schneiderian tradition. The RDC emphasizes the temporal length of observable psychopathology and differentiation based on the presence or absence of mood symptoms. The main difference between the Feighner and the RDC criteria was the reduction of the duration of active symptom phase from 1 month in the Feighner criteria to

2 weeks in the RDC. The duration of illness was kept at 6 months to reach diagnosis. In addition to the definition of formal criteria for schizophrenia, a structured diagnostic interview was developed (Schedule for Affective Disorders and Schizophrenia, or SADS). This kind of structured diagnostic interview became the precursor to several instruments currently in use in research.

The DSM-III, published in 1980, for the first time defined schizophrenia much more narrowly, placed more emphasis on Schneiderian concepts, and included a multiaxial diagnostic system. Dimensions such as secondary diagnosis, impact of functioning, and relevant medical illness and severity were now included in the formal diagnosis. In addition, clear exclusion criteria were described, and the maximum age of onset at 45 years was eliminated. The principles used to guide the development of DSM-III were also used for the development of DSM-IV and DSM-5. Among the many important features considered to make the diagnosis of schizophrenia, there was a new emphasis on using course of illness in addition to cross-sectional symptoms in an effort to improve reliability and validity of the diagnosis. A requirement for symptoms of at least 6 months' duration was added as a requirement for the diagnosis of schizophrenia.

Current Diagnostic Criteria

The current versions of the DSM and *International Classification of Diseases* (ICD) diagnostic systems are the DSM-5 and the ICD-10. In spite of some differences, both systems have a great deal of concordance in terms of the definition of schizophrenia. Both systems reflect a clear narrowing of the diagnosis of schizophrenia compared to previous diagnostic practices, especially in the United States. One of the major differences between them is the longitudinal course criterion. The DSM-5 requires a minimum duration of the illness of 6 months (with at least 1 month of active symptoms) as opposed to only 1-month total duration of illness in the ICD-10. Also, the DSM-5 requires deterioration in social and occupational function, in contrast to the ICD-10, which does not have such a requirement for the diagnosis of schizophrenia. In addition, the ICD-10 continues to include a category of simple schizophrenia that does not require the presence of psychotic symptoms for the diagnosis. All subtypes of schizophrenia were eliminated in DSM-5.

The DSM-5 criteria for schizophrenia were released in 2013 and included some important changes while retaining the core structure of DSM-IV. The symptom threshold for the diagnosis was raised from needing one core symptom in DSM-IV-TR to at least two symptoms. Core symptoms continue to be delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior, and negative symptoms. It also eliminated the special diagnostic importance for first-rank Schneiderian

symptoms such as running commentary hallucinations, two or more voices conversing with each other, and having bizarre delusions. Delusions are now treated equally toward the diagnosis of schizophrenia regardless of their content. A major departure from previous editions of the DSM was the elimination of all subtypes of schizophrenia due to lack of data supporting their stability over time and any clinical or research predictive value. In ICD-10, the symptom criterion is defined in more detail than in DSM-5, and there is a greater emphasis on Schneiderian first-rank symptoms. ICD-10 also makes a clear distinction between primary and secondary negative symptoms. Another important difference from DSM-5 is the strict requirement of 1-month duration of symptoms, even in the presence of successful treatment.

Diagnostic Criteria: Functioning. In addition to the cross-sectional criteria, DSM-5 requires the presence of significant deterioration in one or more major areas of functioning, such as work, interpersonal relations, or self-care. In contrast, ICD-10 does not require functional deterioration for a diagnosis of schizophrenia and in fact states that deterioration is not inevitable.

Diagnostic Criteria: Duration. The DSM-5 criteria require a total duration of 6 months. Within this 6-month period, there must be at least 1 month of active-phase symptoms (overt psychotic symptoms). A shorter duration of the active phase is allowed only if successful treatment is instituted. The rest of the 6-month period may include continuing psychotic symptoms, prodromal symptoms preceding clear-cut psychosis, or residual symptoms after the resolution of the psychotic symptoms. Residual symptoms are defined as attenuated forms of psychotic symptoms, such as odd beliefs, magical thinking, ideas of reference, odd perceptual experiences, peculiar or concrete thinking, vague speech, or odd behavior. Negative symptoms may also be included as residual symptoms. If the 6-month duration criterion is not met, but the other criteria are present, the diagnosis of schizophreniform disorder should be made instead.

In contrast, ICD-10 requires only a minimum duration of 1 month. This time does not include the prodromal phase and does not consider treatment as a variable to shorten the duration of the acute phase. If this duration requirement is not met, a diagnosis of acute schizophrenia-like disorder should be made instead.

Diagnostic Criteria: Exclusions. The DSM-5 criteria require that mood and schizoaffective disorder be ruled out to make the diagnosis of schizophrenia. The patient should not meet criteria for a manic or depressive episode during the psychotic phase, or if there is a concomitant

mood episode with active phase symptoms, the duration of the mood episodes should be brief relative to the duration of the psychotic illness. This includes residual symptoms. In contrast, according to ICD-10, a patient with a depressive or manic episode can still meet criteria for schizophrenia, if criteria for schizophrenia were met before the onset of the mood episode.

DSM-5 excludes a diagnosis of schizophrenia if the symptoms are the result of a direct physiological effect of a substance, a drug of abuse, or a medication, as well as when the symptoms are due to a neurological or medical condition. ICD-10 includes very similar exclusion criteria.

DIAGNOSTIC SUBTYPES

The diagnostic subtypes of schizophrenia were dropped in DSM-5 whereas the ICD-10 continues to reflect little change from Kraepelin's original classification system. Schizophrenia subtypes have been controversial constructs and their clinical utility has been consistently questioned. They have a weak relationship to biological variables, poor long-term stability, and poor predictive value. The paranoid subtype appears to have greater stability than the other subtypes and is used more often in clinical practice along with the undifferentiated type. Instead, in order to capture the rich and diverse clinical presentation of schizophrenia with heterogeneous symptoms and signs that vary over time, the DSM-5 included in its assessment measures section, the Clinician-Rated Dimensions of Psychosis Symptom Severity, an assessment measure that allows quantifying the five core dimensions that define schizophrenia spectrum disorder (hallucinations, delusions, disorganized speech, abnormal psychomotor behavior, and negative symptoms) plus impaired cognition, depression, and mania. Each dimension is rated on a 0 (not present) to 4 (present and severe) scale and can be used for treatment planning, outcomes, and longitudinal follow-up. Catatonia was designated as a separate category in DSM-5 that can now be linked to multiple disorders such as schizophrenia, schizoaffective disorder, and mood disorders.

THE SYMPTOMS OF SCHIZOPHRENIA

The descriptive power of Kraepelin's dementia praecox and the explanatory power of Bleuler's schizophrenia took hold in psychiatry during the first half of the 20th century. By the middle of the century, varying conceptual orientations among psychiatrists led to widely differing applications of the diagnosis of schizophrenia. Schizophrenia was broadly applied, with the diagnosis given to both chronically hospitalized patients with constant hallucinations and bizarre behaviors and independent working people with problems of histrionic behaviors and difficulty making empathic connections with their therapists. With the introduction of new medical therapies to treat the symptoms of depression and psychosis in the 1950s, it

became essential to define which symptoms were associated with which illnesses, so that the right treatments could be applied to the right illnesses and treatment responses could be determined.

To apply statistics to the diagnosis and treatment of psychiatric illnesses required that both symptoms and responses be quantified, and the tool for quantification arose with the generation of rating scales for psychiatric disorders. The first of the broadly used symptom scale was the Brief Psychiatric Rating Scale (BPRS) from Overall and Gorham, published in 1962. This scale was derived from earlier scales and from the evaluation of hundreds of patients, and it assessed aspects of psychopathology that were familiar to most clinicians. It was short (in its first form consisting of just 16 items), easy to administer, and reliable. It was initially intended to provide efficient assessment of treatment response for many diagnoses seen in inpatient settings, but its utility led to wide use in measuring treatment responses of severely ill patients across a broad range of clinical settings. The BPRS was the standard tool for measuring treatment response in psychotic disorders for decades, and it remains widely used in research and as an aid to clinicians treating chronically and severely ill patients.

At about the same time the BPRS was being developed, John Wing and his colleagues were trying to address the poor reliability of psychiatric diagnoses with creation of the PSE. The PSE is a much longer instrument than the BPRS that evaluates 140 signs and symptoms and requires experienced clinicians for reliable administration. Algorithms to facilitate computer-based interpretation of the large bodies of data generated with the use of the PSE were developed in parallel, with an aim of using the PSE to identify groups of patients united by similar symptom patterns. Despite the time involved in administering the PSE and the large bodies of data created when administered to groups, the PSE was used in both the New York–London project and in the IPSS, and its use helped define the parameters of schizophrenia. Like the BPRS, the PSE was intended to be an instrument applicable across a broad range of psychiatric disorders, but it has its greatest impact in studies of schizophrenia.

Reliable rating scales of symptoms and psychopathology, such as the PSE and BPRS, were essential to the progress made in understanding the treatment and course of schizophrenia and for finding the biologic and social correlates of the illness. The authors of the BPRS and the PSE, and the authors of the many rating scales that followed, were all motivated to address particular questions about illness and treatment, and none of them intended to constrain the clinical interview of patients suffering with schizophrenia. But the existence and reliable utility of these scales has nonetheless had the unintended effect of focusing researchers' and clinicians' attention on the questions asked by the scales and diverting some attention from the stories patients have to tell. The scales, like the brief and reliable contemporary diagnostic criteria psychiatrists apply, have

become the definition of illness in the mind of many mental health clinicians, and this view often limits inquiry about the individual experience of patients.

An early application of the PSE was in the WHO's IPSS. The results of the IPSS showed that populations of patients sampled in a range of rural and urban sites around the world all had similar profiles of symptom types, although at each site study participants displayed widely varying clinical presentations. Across sites, the most common symptoms were lack of insight and an inadequate description of problems, suspiciousness, unwillingness to cooperate, ideas of reference, flat affect, persecutory delusions, delusional mood, poor rapport, a range of auditory hallucinations including verbal hallucinations, thought alienation, and gloomy thoughts. Subsequent follow-up studies, other studies conducted across countries, and studies of different cultures within individual countries confirmed that when studied as groups there was similarity in the overall distribution of symptoms within study populations.

By the 1970s, the body of data produced through the use of the PSE in large samples led to new efforts to find correlates between the symptoms of the disease and functional outcomes. Strauss, Carpenter, and Bartko made an early attempt to find patterns in the expression of the symptoms of schizophrenia. In their studies, the PSE-rated symptoms of schizophrenia clustered into groups of positive and negative symptoms and into a group of symptoms associated with "disorders of relating." By the mid-1980s, factor analytic techniques were being more broadly applied to the assessment of the symptoms of schizophrenia, and separate investigators found that three dimensions or subsyndromes of schizophrenia could be derived from rating scales. These dimensions were psychotic symptoms (hallucinations and delusions), negative symptoms (e.g., social withdrawal, blunting of affect, interest, and motivation), and disorganization symptoms (e.g., disorganization of thought and behavior). A similar clustering of symptoms was found to exist in studies of populations from around the world. Later studies showed that two of these subsyndromes showed temporal stability, the negative syndrome especially and to a lesser extent the disorganization syndrome, and that each subsyndrome had its own distinct relationship to demographic factors, cognition, and heritability.

Within the past decade, it has become clear that the number of subsyndromes within schizophrenia is defined in part by the number of signs and symptoms measured, so that more thorough ratings of a larger number of symptoms lead to finding more subsyndromes. Scales or combinations of scales that include more diverse examples of psychopathology lead to the reliable derivation of psychotic, negative, disorganization, depression and anxiety, and agitation dimensions. There have been thousands of patients enrolled in studies using the Positive and Negative Syndrome Scale (PANSS) conducted across continents and

cultures, and there has been a remarkable consistency in the finding of these same five factors. The content of an individual symptom might vary across cultures but the structure of symptoms within the illness seems remarkably consistent. For the most commonly used rating scales, age, symptom severity, duration of illness, and short-term medication withdrawal do not seem to affect the finding of separate subsyndromes.

This dissection of the syndrome of schizophrenia into subsyndromes will multiply if more and more elaborate assessments of signs, symptoms, and history are completed. A study that detailed premorbid and prodrome histories along with routine assessments of clinical symptoms found separate factors for educational achievement and premorbid function with unique heritability. More detailed assessment of disorganization symptoms finds subfactors for motor and verbal disorganization, and when a larger number of mood and anxiety questions are asked, mood and anxiety may be split into separate factors. However, each subfactor is most highly correlated with the other symptoms in the same class. At this time, the addition of additional factors, or the division of the five PANSS factors into separate subfactors, is primarily important for investigators with a narrow focus on therapeutics or neurobiology for one group of symptoms—the five-factor model is a valid and useful way to organize the clinical symptoms that are assessed to make decisions about patient's care.

Psychotic Symptoms

The psychotic symptoms of schizophrenia define schizophrenia for the general public and for many clinicians. They are the obvious and dramatic symptoms of the illness, and for most people, they are what is fascinating and frightening about schizophrenia. Within psychiatry, the interest in delusions and hallucinations has overshadowed investigation and treatment of the other facets of the psychopathology of schizophrenia. This emphasis on psychotic symptoms was reinforced with the break from the impressionistic diagnostic scheme of DSM-II to a symptom-based diagnosis in DSM-III, when the diagnosis came to depend in large part on clinicians' ability to recognize the psychotic symptoms considered characteristic of schizophrenia.

It is now clear that there is no particular symptom, not even the most carefully specified kind of hallucination or delusion, which by itself generates or justifies a diagnosis of schizophrenia. It is also clear that a focus on treating positive symptoms to the exclusion of the other signs and symptoms of schizophrenia will ignore significant sources of morbidity and might exacerbate other forms of psychopathology and disability in clinically significant ways. Nonetheless, psychotic symptoms, because of their both unique nature and easy identification, aid assignment of diagnoses and serve as markers of treatment response, and these symptoms will remain at the core of concepts about the nature of schizophrenia.

The use of the term psychotic in relation to symptoms should be restricted to hallucinations and delusions. In the past, symptoms of inappropriate affect, purposeless activity, and disorganization of speech and behavior were also considered psychotic symptoms, but it is now clear that they are better considered in the separate disorganization dimension. Psychotic symptoms are also not synonymous with the term "positive symptoms," which has traditionally included symptoms, such as incoherence or loosening of associations, or motor restlessness, that are assigned to disorganization and agitation factors. Similarly, the symptoms in the original PANSS "Positive" subscale are a composite of psychotic, disorganization, and agitation factors. At this time, there is no reason to use the original positive subscale.

The responsiveness of psychotic symptoms to antipsychotic medications has led to some unforeseen costs in attempts to effectively treat schizophrenia. The focus on psychotic symptoms of schizophrenia that arose with the adoption of the diagnostic scheme of Kurt Schneider and its emphasis on "first-rank" symptoms of psychosis (Table 12.2–1) became prominent in the United States at about the same time the demonstrable benefits of antipsychotic medications became clear to most practicing psychiatrists. This evolved into a seemingly reasonable treatment approach that focused on maximizing antipsychotic doses to minimize or eliminate the psychotic symptoms that were judged to define the illness. The costs in this approach were not apparent to most clinicians, and only in the 1970s and 1980s were there prominent psychiatrists who suggested that high-dose medication treatments might impose burdens such as secondary negative symptoms, dyskinesias, and increased rates of medical complications that were not always offset by the benefits achieved with decreased psychosis.

Hallucinations. The earliest descriptions of schizophrenia did not share the more contemporary emphasis on psychotic symptoms, but there was certainly clear awareness of the burden hallucinations impose. Although for Bleuler hallucinations were accessory symptoms, in his *Textbook of Psychiatry* (1924), he evoked the disruption that hallucinations must impose on the inner life of those afflicted with schizophrenia.

> The patients hear blowing, roaring, humming, rattling, shooting, thundering, music, crying and laughing, but above all whispering, speaking, calling; they see individual things, landscapes, animals, people and all sorts of impossible figures; they smell and taste all sorts of pleasant and unpleasant things; they touch things, animals and people and are struck by rain drops, fire, and bullets, they feel all the tortures and perhaps all the pleasant things also that our bodily sensations can transmit.

Although an exclusive or even predominant focus on the psychotic symptoms of schizophrenia is not appropriate, hallucinations do hold a special place in the experience of people with schizophrenia. Perceptions

that can cause people to “feel all the tortures and perhaps all the pleasant things” will certainly influence behavior, and much that seems incomprehensible in schizophrenia may become understandable when a patient’s experiences with hallucinations are revealed. Hallucinations also drive some of the understanding of illness that patients, their families, and their physicians share or do not share. Families and health care providers may not have the overwhelming experience of distorted perception that people with schizophrenia often have, but families and providers have had dreams and vivid perceptions, and sometimes these experiences can help them imagine something of what happens with patients. Alternatively, families may chide patients for listening to people who are not there, and physicians will hurry through cursory encounters with patients who seem to pay more attention to voices of intangible people than they do to the doctor in front of them.

Many of Kurt Schneider’s first-rank symptoms, proposed to specify a diagnosis schizophrenia—hearing one’s thoughts spoken aloud, voices referring to oneself, a running commentary on the patient’s thoughts or behaviors, a sense of being the recipient of externally imposed sensations—are hallucinations and were initially considered to make a diagnosis of schizophrenia likely in the absence of identifiable brain pathology. In fact, hallucinations are common in other psychiatric illnesses, and the first-rank symptoms have been found to vary from a minority of patients to very high rates in both schizophrenia and in other psychiatric disorders, depending on the population being studied. Although there is no particular unique feature of hallucinations in schizophrenia, there is typically greater chronicity in schizophrenia than that in other psychiatric illnesses. First-rank symptoms had a special place in DSM-III and DSM-IV, where the symptom of auditory hallucinations consisting of a running commentary on a person’s behavior or thoughts, or voices conversing with each other, was sufficient by itself to meet the symptom criteria for schizophrenia. This special status is eliminated in DSM-5.

Hallucinations and other psychotic symptoms are not rare in the general population. A review suggested that 5 to 8 percent of the general population experience hallucinations or persecutory delusions, and other reviews cite figures of up to 15 percent of the general population as hearing voices. Symptoms of anxiety are particularly associated with hallucinations, and similar to schizophrenia, there is a particular increase in the frequency of auditory hallucinations. Not surprisingly, substance abuse is also associated with all kinds of hallucinations, but a particular association between substance abuse and auditory hallucinations is not found. These findings suggest that conceiving of hallucinations as a discrete manifestation of psychopathology that establishes the presence of mental illness might be a convenient assumption that provides diagnosticians some security but does not in fact reflect the nature of the experience. In a

Netherlands-based study, younger age, living in an urban environment, low income, less education, unemployment, female gender, and being single are all associated with increased rates of hallucinations. Many of these demographic factors seem similar to those of people who will develop schizophrenia. Also similar to studies of schizophrenia is the finding that in the general population, hallucinations predict the presence of delusions, and that acute stressors increase the likelihood of hallucinations. It has been clear for decades that hallucinations do not confirm or exclude schizophrenia versus any other psychiatric diagnosis, and the finding of frequent hallucinations in the general public suggests that the presence of hallucinations does not by itself confirm psychiatric illness.

Since the time of Karl Jaspers, there has been a view that hallucinations that seem to arise from outside the body are "true" hallucinations, while hallucinations that seem to arise inside the body are pseudohallucinations; that this is something of a distortion of Jaspers' actual discussion has not affected views. There is a similar view among some clinicians that hallucinations experienced by a person with schizophrenia who has good insight into their nature are not really hallucinations. Neither of these distinctions, nor the volume, number of voices heard, or content of the hallucinations, has diagnostic significance or demographic or functional correlates. Many patients will describe auditory hallucinations that arise both inside and outside their heads or will develop insight into the nature of their hallucinations only after decades of full belief in their external reality. When considering patients, these distinctions should not be used to make a diagnosis or generate a treatment plan.

There is more information about the nature of psychosis than any other symptoms of schizophrenia, but even in the study of hallucinations, there are large gaps in the current knowledge. Most investigations focus on the presence or absence of a symptom, and when there is more detailed inquiry, it is typically limited to questions about frequency. There is a literature dealing with questions of when hallucinations arise, what provokes them, how patients cope with them, or what other perceptions or thoughts are associated with hallucinations, but the answers to these questions arise from a surprisingly limited number of studies.

Studies that examine the prevalence of hallucinations across cultures come up with similar rank orders, though the exact prevalence of each type of hallucination in each setting varies. Auditory hallucinations are the most common type, followed by visual hallucinations, and tactile (or *haptic*), olfactory, and gustatory hallucinations are less common. Visceral and other deep tissue hallucinations (*cenesthetic* hallucinations) can provoke odd descriptions that are inconsistent with the internal sensory apparatus, with complaints of burning in the brain or movement of the thoracic organs to new positions. There is very little contemporary research regarding visceral hallucinations and their prevalence, nature, or associations with other

symptoms, but they are likely much less frequent than gustatory and olfactory hallucinations.

AUDITORY HALLUCINATIONS. In schizophrenia, auditory hallucinations are clearly the most common type of hallucination. This may not always have been the case; a century ago, visual hallucinations were more common than auditory hallucinations in London, and visual hallucinations have been decreasing in frequency and auditory hallucinations have been increasing in frequency over the past century in Vienna. Also, auditory hallucinations are relatively less common, and visual hallucinations are relatively more common in African, Near Eastern, and some Asian populations than in Western societies. Nonetheless, evidence from the WHO's IPSS and other sources confirms that the most prevalent type of hallucination and the most common symptom of psychosis in schizophrenia is auditory hallucinations.

In keeping with their prevalence, auditory hallucinations are the best characterized of the psychotic symptoms. The IPSS studies suggest that internationally, more than 70 percent of people with schizophrenia have auditory hallucinations, and that percentage is probably higher in industrialized societies. As with many symptoms, careful questioning in prospective studies is associated with a higher rate of reporting for auditory hallucinations, leading to a prevalence of 98 percent in one report.

Patients generally believe that their hallucinations are real manifestations of someone, somewhere, talking to them or transmitting a voice or voices. They frequently attempt to test the reality of their voices, but patients' tests often rely on attempts to confirm that the voices really know what the patient is thinking or doing or planning. Since the voices inevitably do manage to know such things, their reality is proved to the patient. Even when voices fail their experiments, for example, failing to accurately predict what will be served for supper, a delusional explanation will be generated that confirms the reality of an external cause. Ideas of reference and other delusions further reinforce these interpretations of the hallucinations. One powerful explanation for the belief in external sources of hallucinations by persons experiencing them is that no one would want to generate such an experience; the hallucination is usually so unpleasant, so intrusive that it makes no sense that it would arise from within the person who experiences it. Although there have been decades of speculation about why the content of auditory hallucinations is so often threatening or demeaning, there is at present no clear explanation for the generally negative content of the auditory hallucinations.

The content of verbal auditory hallucinations, both in words and in prosody, is generally derogatory or threatening and often quite profane. Patients will hear expletives, threats, demeaning personal comments, and accusations of vile thoughts or behaviors; this kind of negative content comes more commonly from male voices. Not surprisingly, patients are

often distressed and frightened, or angered by these experiences, and can have a complicated sense of guilt, depression, or belligerence in response. Patients typically feel less control over the experience of negative hallucinations than they do over benign or positive hallucinations, and severity of delusions and negative symptoms can further inhibit coping with hallucinations. Emotional experiences, often sadness or apprehension, often precede hallucinations, but patients also often experience the somatic symptoms of anxiety, and less frequently fear or anger, before the recurrence of hallucinations. Social stressors, physical illness, and chronic pain can all increase the frequency of hallucinations and the intrusiveness and dysphoria associated with the experience. Hallucinations can be unremitting and last from waking until bedtime or can come in irregular, infrequent periods of seconds or minutes a few times a month. Although it is not a rule, evenings and nights before retiring are often times when auditory hallucinations arise or escalate. It is common that people with schizophrenia have some period during their illness when their auditory hallucinations are present for the majority of the day.

Voices are the most common kind of hallucination and tend to be clear and understood, but unintelligible sounds of whispers or distant conversations are not rare. Patients can also hear footsteps, telephones ringing, the sounds of pebbles hitting a window, helicopters, buzzing, clicking, and musical phenomena that range from single notes present throughout the day to discrete complex passages of orchestral music. Bleuler described “extracampine hallucinations” in which the perception was received by the patient outside the sensory apparatus—for example, a sound is appreciated through the knee.

When it is voices that are heard, single words are probably the most common. Although there is an enormous range in the number of voices that are heard by different patients, the mean number is been estimated to be 3. The gender of the speaker can usually be identified. The person afflicted often hears voices that are interpreted as coming from a person who has more social status than the patient. Younger patients will tend to hear somewhat younger voices, while older patients hear somewhat older voices. The person afflicted by them often reliably assesses the location, loudness, and clarity of the voices, and in general, persistent auditory hallucinations will eventually lead the patient to settle on identities for their voices.

With the passage of time, most patients will experience a decrease in the frequency of their hallucinations or a change in the nature of their hallucinatory experience. Where hallucinations are initially very upsetting, provoking anger or panic, after the first stages of illness have subsided, patients may generate coping strategies through a process of trial and error over a period of years and can come to mitigate the effects of hallucinations on their lives. Simple strategies such as yelling at the voices can sometimes

decrease or stop the experience, though clinicians and families may see this as a distressing sign of deterioration. Less disturbing to people around the patient, singing or humming or listening to music can block the experience, and sometimes shifts in posture or talking to someone else can help. Unfortunately, for those afflicted, the only available regular contact with the voices of other people may be through the radio or the television, and these can reliably exacerbate hallucinations for some. After years or decades, patients often start to question the content of their hallucinations, and if their questions do not lead to insight, at least they sometimes provoke awareness that the threats or predictions rarely come true. Some patients can gain limited control of their hallucinations and can shift the prosody or content of the discussion, and there is a sizable minority of patients who can regularly start and stop their hallucinations at least some of the time. Religious themes are common in auditory hallucinations, and while they commonly start with a focus on the sins of the patient, with time there can be a conversion to a view of the patient having a special relationship with his or her god or gods, or that the individual might have some special message to spread or task to perform.

As a person gets used to the presence of these voices, the hallucinatory voice or voices can become companions, and given the frequent social isolation that is experienced in schizophrenia, this companionship can be welcome. The voices can take on the role of a guide through daily life for the person with schizophrenia, often in middle age or later, so that the hallucinations will tell a person what to buy at the grocery store, how to structure his or her day, sometimes even how to respond in social interactions. The voices might even take on a moral role for the patient, suggesting the emotional consequences of particular behaviors, or act as a conscience that spurs generosity or remorse in response to the patient's actions. Although some people with schizophrenia will with time develop good insight into the nature of their hallucinations, despite ongoing delusions or negative symptoms, other patients with predominately positive symptoms or residual schizophrenia will go through long periods with few obvious symptoms and no hallucination but retain a full delusional interpretation of their past hallucinatory experiences.

COMMAND AUDITORY VERBAL HALLUCINATIONS. Sensational violent crimes committed by people with severe mental illnesses have led to an increased popular and clinical focus on the phenomenon of command verbal auditory hallucinations. Command auditory verbal hallucinations are very common, present at some phase of the illness for at least one-third and maybe as many as three-quarters of patients with schizophrenia; there are estimates that the lifetime prevalence of command hallucinations approaches the prevalence of auditory hallucinations as a whole. The emphasis on command hallucinations in clinical assessments of risk is unfortunate,

because other clinical symptoms and features of the illness history that contribute more to risk assessments are sometimes neglected once a patient's experience, or lack of experience, with command hallucinations is elicited in the clinical interview. There is a consensus that patients are more likely to comply with command hallucinations if (1) the voice is believed to be a real communication from someone else, especially if there is an identity for that person, (2) voices have benevolent intentions toward the patient or the action described will help the patient, (3) the voice has some omnipotence or other power greater than the patient, (4) there is are adverse consequences to the patient for not complying, and (5) the command is for a nonviolent action. Trait anger and impulsivity also increase the risks of compliance. Likely more important than any of these specific attributes of the hallucination is the patient's delusional structure; if the hallucinatory experience is congruent with his or her delusional beliefs, the person is more likely to comply with the commands.

VISUAL HALLUCINATIONS. Visual hallucinations are less common than auditory hallucinations in schizophrenia, but they are not rare. Estimates of the prevalence of visual hallucinations do not generally exceed 55 percent, and a more widely accepted estimate would include around one-third of patients with schizophrenia having visual hallucinations at some point in their illness. As is the case with other psychiatric symptoms, careful investigation by clinicians can sometimes elicit a history of visual hallucinations when more casual questioning fails to detect their presence. Up to half of those with schizophrenia and auditory hallucinations will have visual hallucinations, but it is relatively uncommon for someone with schizophrenia to have visual hallucinations in the absence of auditory hallucinations. Visual hallucinations have been suggested to predict a more severe illness, and there are some data to support this contention. It may be that patients with predominately psychotic symptoms are less likely to have visual hallucinations than patients with significant primary negative symptoms or disorganizations symptoms.

There is a broad range of visual experiences that afflict people with schizophrenia. The most commonly reported visual hallucinations are formed images of animate objects, people, or parts of people (especially heads and faces), religious images, animals, and fantastic creatures that may be similar to images in books or movies or television. Inanimate objects are less common, but flashes of light, streaks of color, or grids or lines in the visual field may also be described. Images can be in color or black or white, and hallucinations in only shades of white have been described. Visual hallucinations are generally more discrete and more limited in duration than auditory hallucinations, but there are patients who will experience visual hallucinations throughout the course of a day. In vision, the continuum with illusions (distortions of objects seen in the

environment) is more obvious than it is with auditory perceptions. Shadows that seem like people or animals but reveal themselves as only shadows with closer inspection, objects that change size or shape, outlines of objects that seem to float just outside the object itself, and objects that have some transparency are all kinds of visual distortions that are reported. Some particularly unfortunate people will have fluctuating periods of dramatic distortions of the world itself; the ground beneath someone will seem to be at angles that are inconsistent with the patient's proprioceptive awareness, sidewalks will briefly appear to point into the air, and walls will appear or disappear in buildings. Lilliputian hallucinations—hallucinations of small people, frequently busy attending to some task—are generally considered a sign of an encephalopathic process but they are rarely present in schizophrenia.

The content of auditory and visual hallucinations is often dependent on the culture of the person experiencing hallucinations, but there seem to be some culturally appropriate visual hallucinations in non-Western societies that have only parallels to religious-themed hallucinations in the West. Even in cases of unequivocal schizophrenia, in non-Western societies there can be a social acceptance that the hallucinatory experience is an extension of a commonly shared reality. It seems plausible, though not demonstrated, that cultural acceptability might influence the sensory mode of hallucinations experienced (e.g., visual instead of auditory), as well as the content of the hallucinations.

OLFACTORY, GUSTATORY, AND TACTILE HALLUCINATIONS. Tactile hallucinations are present in a range of 15 to 25 percent of people with schizophrenia, with no clear pattern of variation across cultures. Particular tactile hallucinations, such as the sense of bugs crawling on or under the skin (*formication*), are found in a variety of mental illnesses and encephalopathies. Like auditory and visual hallucinations, the tactile hallucinations of schizophrenia can take on a broad variety of forms. Feelings of being touched, burned, and cut are common, and there are surprising numbers of patients who will feel as though they are subject to a kind of sexual assault, inflicted on their genitals or anus by unseen (and sometimes seen) entities. Less localized sensations of electric shocks or tearing or stabbing in the limbs and torso also occur. Patients with tactile hallucinations can also have a physical sense that somehow other people or magical beings enter into and exit their bodies, and while they can communicate that this sense is particularly distressing for them, articulating what it actually feels like may be very difficult.

Olfactory and gustatory hallucinations tend to be reported by a small minority of patients, with olfactory hallucinations the more common of the two types. As with other kinds of hallucinations, the experience tends to be unpleasant for most patients, with the smells of rotting meat, garbage, and

feces common, and the taste of blood or metal frequently described. Often, tastes and smells will go together in an unpleasant combination.

Delusions. Like hallucinations, delusions hold a special place in both the medical and popular perception of schizophrenia, with a view that delusions are a discrete category of thought that is generated and maintained through unique mechanisms and consisting of ideas that are inaccessible to normal reason. A standard definition of delusions is frequently reduced to calling them "fixed false beliefs." In the glossary for DSM-5, delusions are defined as "A false belief based on incorrect inference about external reality that is firmly held despite what almost everyone else believes and despite what constitutes incontrovertible and obvious proof or evidence to the contrary" and "not ordinarily accepted by other members of the person's culture or subculture." This emphasis on falsity and imperviousness to argument owes its origin in large part to the influence of the German psychiatrist and philosopher Karl Jaspers. In *General Psychopathology,* Jaspers wrote that delusions were "held with an extraordinary conviction, with an incomparable, subjective certainty," that they were "impervious to other experiences and counter-argument," and that their content is "impossible." This description makes a delusion sound like something entirely unlike the ideas generally held and understood, which clinicians like to believe are arrived at through considered investigation and reasoned thought. Jaspers further argued that the origin of "delusions proper" arose from a unique kind of knowing or experiencing the world—not from the exercise of faulty judgment about personal experience but from a novel sort of experience that intrudes itself into the mind, a "direct experience of meaning." This exists on no continuum with the experience of the general population, who derive meaning for their experiences from context, history, emotion, cognitive style, and state of mind. However, the variability in the type and character of delusions presented by patients has confronted the notion that delusions arise de novo, immutable, and impervious to reason or experience. To deal with this incompatibility between observation and theory, the awkward idea of "delusion-like-ideas" has been introduced.

Like hallucinations, delusions are not experiences that happen only to the mentally ill. When a representative sample of the German population was asked about their beliefs with a questionnaire, a quarter reported that they had special abilities or powers, and a third endorsed a belief in telepathy. Reflection on daily life shows that many people clinicians know outside of clinical settings, and clinicians themselves frequently hold ideas that have no grounding in a shared external reality.

Furthermore, the delusions that are held by persons with schizophrenia are frequently quite flexible, and while argumentative confrontation accomplishes little, life circumstances or shifting attention can change

delusions in dramatic ways. There are patients who will cling tenaciously to the idea that they have been persecuted by agents of a foreign government, or who have had chips implanted in their brains by aliens from space, but every day around the world other patients will check out fantastic ideas with family, friends, nursing staff, and doctors, and happily accept the assurance that their odd concerns are unfounded.

When delusions are examined as they exist in schizophrenia, there is a clear continuum with the overvalued ideas everyone is prone to. Delusions can vary from fleeting concerns about personal meanings taken from the way a television news reader intonated a particular word to a preoccupation with ideas about how giant corporations and governments are organized to persecute the individual with schizophrenia. The delusions found in schizophrenia frequently are different from the overinvested ideas that almost everyone has, but the differences do not seem to arise from a novel process that is found only in schizophrenia. Contrary to Jaspers' formulation, delusions in schizophrenia often develop from normal cognitions, and the path by which delusions are developed often seems quite clear to clinicians when the circumstances and state of mind of the patient are described. Delusions encountered in clinical settings can be presented as impressions that quickly abate, and it is a common experience for clinicians to inquire about a delusional concern previously reported by a patient, only to find that the patient has dismissed his or her old idea, as though it is of no consequence, and now has new preoccupations. Instead of arising from a new kind of perception, delusions often arise from social stressors, and threats to self-esteem seem particularly prone to provoke persecutory or grandiose delusions in response.

What does seem to distinguish delusions in schizophrenia is analogous to what distinguishes schizophrenia itself: the frequency of ideation, severity of the impact of the symptom on the patient's life, and the pattern of delusions in relation to other symptoms, instead of any unique or pathognomonic content or structure. The number of delusional beliefs, the strength of preoccupation with delusions, and the amount of distress they cause for patients all seem to distinguish delusions in schizophrenia. Although delusions in schizophrenia are not always "firmly sustained," there is in general a stronger conviction that the delusional belief is accurate despite contradictory views from others. Delusions are common in many psychiatric syndromes but tend to be held more strongly, preoccupy more time, and are more pervasive in their influence on patient cognition than is found in other psychiatric illnesses. Most people with schizophrenia will have delusions, and the percentage of people who will experience clinically significant delusions may exceed 70 percent. In Martin Harrow's long-term follow-up studies, 25 percent of patients had delusions at all assessments over a 20-year period. As is the case with hallucinations, many patients will moderate the effects of delusions on their lives, often by

titrating amounts of social contact. This increasing engagement can decrease their preoccupation, although when the delusions center on concerns of persecution this relief may not be available. For most patients, the presence of one delusion predicts the presence of other delusions.

CONTENT OF DELUSIONS. It is an unfortunate feature of delusions in schizophrenia that the focus of the delusion is so often based around fears of persecution or threats to a person's safety. People with schizophrenia are more likely than people in the general population to fear that people look at them in an unusual or odd way, to believe that they are being watched or being harmed, or that their feelings and actions are not under their control. In schizophrenia, most delusions will fall into some broad patterns of content. The most common delusions in schizophrenia are delusions of persecution, which are endorsed by up to 80 percent of inpatients. These persecutory fears can range from transient doubts about the intentions of strangers around them to profound convictions that charitable organizations or family members are plotting to have the patient tortured or killed. These may arise from misinterpretation of social interactions or social events or from misattribution of intent on the part of others.

A sizable literature has arisen around the formation of persecutory delusions. Patients with paranoia tend to have a worried thinking style, sleep disturbance, and anomalous internal experience. They are people with increased interpersonal sensitivity and negative beliefs about themselves and are influenced by life experiences and social circumstances to view the world as threatening. Living in urban environments and co-occurring substance abuse seem to exacerbate the problem.

Other highly prevalent delusions, such as delusions of bodily control and of having one's thoughts audible by others, are also delusions that can lead to persistent dysphoria. *Delusions of reference* refers to delusions in which neutral stimuli or events or people take on a particular significance in reference to the person who experiences the delusions; in the DSM-5 glossary, they are "a delusion in which events, objects, or other persons in one's immediate environment have a particular and unusual significance." There is usually a persecutory slant to these delusions, although there may also be self-aggrandizing or religious meaning derived from them. In DSM-5, they are considered distinct from *ideas of reference,* "feeling that causal incidents and external events have a particular and unusual meaning that is specific to the person." As already discussed, delusional ideas lie on a continuum with normal beliefs, so the distinction between delusions of reference and ideas of reference seems of little value in clinical settings and a source of unreliability in assessments for research purposes. As important, "delusions of reference" and "ideas of reference" are frequently treated as synonyms by clinicians. Other common delusions encountered in schizophrenia, whose names do not clearly define their meaning, include

thought broadcasting and *thought insertion*. These refer to a delusion that the afflicted person's thoughts are audible to others and a delusion that certain thoughts are not the person's own but are inserted into the deluded person's mind, respectively. Also common are grandiose ideas of personal power or significance and religious themes, which can be persecutory as often as grandiose in their content. There are a series of eponymous delusions, most commonly referring to delusions that focus on questions of the identity of the deluded person or those around him or her, or referring to a particular impossible state of existence. When first identified, each of these eponymous delusions was supposed to define a particular diagnostic association with schizophrenia, but since they can be found in almost any medical or psychiatric condition where delusions arise, there is little value of knowing each particular name.

A type of delusion that held special significance in diagnosing schizophrenia is the concept of "bizarre delusions." In DSM-III and DSM-IV, the presence of bizarre delusions was deemed important enough that its existence, in isolation from other characteristic symptoms, was sufficient to meet Criteria A for schizophrenia. Like the idea that auditory verbal hallucinations of voices conversing or commenting, the significance of these delusions arose from the tradition of Schneider's first-rank symptoms. Bizarre delusions are those that are considered implausible by people who are in the patient's culture, and this is generally taken to mean something that is judged physically impossible. There have been many criticisms of the concept of bizarre delusions, the most prominent being that when assessing presence or absence, the interrater reliability varies from fair to very low, even when trained raters are used. Bizarre delusions are not a universal finding and are not pathognomonic of any disorder, but they are more common in schizophrenia than in any other psychiatric illness, and when present, they do discriminate schizophrenia from other psychiatric illness with fair reliability. However, when used as diagnostic tool, they do not affect the reliability of the diagnosis of schizophrenia versus other illnesses with psychosis, and the special status of bizarre delusions is dropped in DSM-5.

Although the most common theme in the delusions found in schizophrenia is persecutory, it might be that the most common inciting event for the formation of a delusion is the need to explain the experience of auditory hallucinations. The validity of the hallucinatory experience seems self-evident to most patients, but there is frequently considerable difficulty generating any clearly communicated idea about how the sounds of someone else's voice is perceived by the patient and the patient alone. A common resolution is for the patient to invoke a fantastic explanation about science and technology and conspiracies; in less technologically focused societies, magic and witchcraft and religion are the most common explanations. In both settings, the explanations fall into culturally validated

explanations for otherwise incomprehensible phenomena. Complex and persistent hallucinations are associated with complex and systematized delusions, where contradictions and conflicts with the reality of a patient's situation have ready answers in the form of ever more elaborate and internally consistent delusions. In these circumstances, the power of a delusional idea can be impressive and sufficient to make patients discount widely shared knowledge possessed for most of their lives when it conflicts with their delusions. Systematized delusions are most common in middle age, and this may be a reflection of how much time it takes after the onset of schizophrenia to accommodate a new reality of the illness. Other hallucinatory experiences and other severe symptoms of schizophrenia can also provoke delusional interpretations from patients. Olfactory and gustatory hallucinations in particular may predict more severe delusions.

As noted earlier, careful application of detailed assessments can lead to hallucinations and delusions being considered separate subfactors in the general class of psychotic symptoms. Outside of research settings, the separation seems to have little utility. Hallucinations and delusions together remain the most treatable cluster of symptoms in schizophrenia, and the severity and change in one symptom class is highly correlated with change in the other.

Potentially less distressing themes, such as grandiose and religious content in delusions, are experienced by half or fewer people with schizophrenia. It seems that awareness of the frequent occurrence of grandiose delusions in schizophrenia is being lost in clinical settings; like the tendency in the past to diagnose anyone experiencing hallucinations with schizophrenia, it has become common to diagnose everyone with grandiose delusions with bipolar disorder, disregarding other diagnostic discriminators. It can be generally expected that delusions of persecution or religious content will have the most effect on a patient's behavior, although the response of each patient to each thought is dependent on that person and that situation.

DELUSIONS AND COGNITION. A review of investigations into the nature of delusions finds that in fact there does appear to be a route to delusions that can be understood outside the idiosyncratic experience of the patient who experiences a "direct experience of meaning." Although the ultimate cognitive deficits underlying delusions remain the subject of conjecture, it is clear that contrary to Jasper's position, there are reasoning errors that accompany delusion formation in schizophrenia. When tested, people with schizophrenia who have delusions will make decisions based on less information than people with schizophrenia and no delusions, or compared to normal populations. These judgments are held in greater confidence than the decisions of control groups despite being poorly supported. It is not the case that people with schizophrenia and delusions are incapable of

making reasoned judgment because when these patients are instructed to hold on their decision making until all information is presented, they tend to make the same decisions as controls. This "rush to judgment" is a trait characteristic that does not disappear when delusions remit, and in schizophrenia, it may for a common cause for the formation of delusions.

Negative Symptoms

Negative symptoms remain under recognized and underappreciated as a cause of disability in schizophrenia. It could be claimed that negative symptoms are the most important clinical symptoms in schizophrenia because the severity of negative symptoms predicts short-term and long-term disability better than the severity of psychotic or disorganization symptoms. Negative symptoms may also be the most significant predictor of social function; in a laboratory study, the severity of negative symptoms had greater predictive power than impairment of social skills on measures of work performance. Negative symptoms are more stable across time than are positive and disorganization symptoms. Negative symptoms "represent a loss or diminution of normal functions" and stand in contrast to positive symptoms where there are perceptions and cognitions and behaviors added to normal mental functions.

In the original conceptions of schizophrenia, negative symptoms held a more prominent place. In *Dementia Praecox and Paraphrenia,* Kraepelin remarked on "emotional dullness" and related it to "the loss of interest, the loss of inner sympathy, with the giving way of those emotional main-springs which move us to exert our mental powers, to accomplish our tasks, to follow trains of thought." For Bleuler "'emotional deterioration' stands at the forefront of the clinical picture." Kurt Schneider appreciated the emotional withdrawal and lack of empathy that clinicians frequently experienced in dealing with patients with schizophrenia but did not include these disturbances in his characteristic symptoms of schizophrenia. Following Schneider, the emphasis on reducing or hopefully even eliminating hallucinations and delusions as the primary goal in the treatment of schizophrenia led to the underrecognition and undertreatment of negative symptoms, and that lack of recognition and treatment persists to the present.

The publication of DSM-III did little to change the emphasis on positive symptoms. The only acknowledgment of negative symptoms came in a reference to blunted or flat affect, which could make a formal thought disorder constitute a symptom when it occurred in conjunction with the disturbance of affect. However, the accumulating research of John Strauss and William Carpenter, T. J. Crow, Nancy Andreasen, and others through the 1970s and 1980s led to an acknowledgment of the role of negative symptoms in schizophrenia, and flat affect was added to the diagnostic criteria for schizophrenia in DSM-III-R. In the DSM-5 diagnostic criteria,

negative symptoms are manifest in "diminished emotional expression or avolition" to constitute one of the characteristic symptoms necessary for fulfilling Criteria A for the diagnosis of schizophrenia. Adding negative symptoms to the diagnostic criteria in the DSM has not increased the incidence of schizophrenia, an indirect confirmation that these symptoms have been present but unrecognized throughout the period of modern diagnostic criteria.

The two-syndrome concept as formulated by T. J. Crow was especially important in spurring research into the nature of negative symptoms. Crow's formulation of two separate dimensions of psychopathology, with a treatment responsive set of positive symptoms and a less responsive and potentially irreversible set of negative symptoms tied to demonstrable neuropathology, led to the generation of testable hypotheses and research efforts that continue to be active to this day. These investigations were facilitated by the creation of reliable and valid rating scales that extended a focus on negative, as well as positive, symptoms, with the PANSS, the Scale for the Assessment of Negative Symptoms (SANS) (Table 12.2–2), and the Scale for the Assessment of Positive Symptoms (SAPS) (Table 12.2–3) of Nancy Andreasen as the most important of these new tools. It would be shown that the groupings of symptoms into negative and positive subsyndromes in the PANSS, and in the SANS and SAPS, did not reflect the reality of the five different subsyndromes—psychotic, disorganization, negative, mood and anxiety, and agitation—but this does not diminish the creative efforts that led to these scales or importance of these scales for research. In fact, it was only through careful analysis of the structure of symptoms in these scales that a more accurate characterization of the phenomenology of schizophrenia was possible.

Table 12.2–2.
Items in the Scale for the Assessment of Negative Symptoms

I. Affective flattening or blunting
   a. Unchanging facial expressions
   b. Decreased spontaneous movement
   c. Paucity of expressive gesture
   d. Poor eye contact
   e. Affective nonresponsivity
   f. Inappropriate affect
   g. Lack of vocal inflections
II. Alogia
   a. Poverty of speech
   b. Poverty of content of speech
   c. Blocking
   d. Increased latency of response
III. Avolition—apathy
   a. Grooming and hygiene

b. Impersistence at work or school
c. Physical anergia

IV. Anhedonia—asociality
a. Recreational interests and activities
b. Sexual interest and activities
c. Intimacy and closeness
d. Relationships with friends

V. Attention
a. Social inattentiveness
b. Inattentiveness during testing

Negative symptoms do not exist only in schizophrenia or only in psychiatric illness. A full range of negative symptoms are seen in people who have had brain injuries, and mild expression of these symptoms are seen in 5 to 10 percent of the population who do not meet criteria for a psychiatric illness. A significant negative syndrome does seem to be much more common in schizophrenia than in other psychiatric diagnoses, including in comparison with mood disorders, and in psychiatric illnesses the pattern of multiple enduring negative symptoms not due to medication or social effects appears most often in schizophrenia. In chronic populations about half the patients will have at least one overt negative symptom. In Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE), 40 percent of patients had significant negative symptoms, and of those, only one-half had significant positive symptoms. Although the diagnostic schemes are tied firmly to psychotic and disorganization symptoms, negative symptoms may better distinguish schizophrenia from other psychiatric syndromes, and it appears that negative symptoms are most reliably associated with outcome in schizophrenia.

There is a range of demographic, treatment, and symptom variables associated with the presence of significant negative symptoms. These include male gender, poor premorbid education and vocational function, earlier onset of illness, and longer duration of untreated illness. Negative symptoms in the prodrome, and especially at onset, predict persistence. There is a reciprocal relationship with affective symptoms, where more prominent affective symptoms are associated with a milder severity of negative symptoms. Although there is some contradictory evidence, it appears that negative symptoms are more heritable than the psychotic symptoms but less heritable than the disorganization symptoms.

Table 12.2–3.
Items in the Scale for the Assessment of Positive Symptoms

I. Hallucinations
a. Auditory hallucinations
b. Voices commenting

c. Voices conversing
d. Somatic or tactile hallucinations
e. Olfactory hallucinations
f. Visual hallucinations

II. Delusions

a. Persecutory delusions
b. Delusions of jealousy
c. Delusions of guilt or sin
d. Grandiose delusions
e. Religious delusions
f. Somatic delusions
g. Delusions of reference
h. Delusions of being controlled
i. Delusions of mind reading
j. Thought broadcasting
k. Thought insertion
l. Thought withdrawal

III. Bizarre behavior

a. Clothing and behavior
b. Social and sexual behavior
c. Aggressive behavior
d. Repetitive or stereotyped behavior

IV. Positive formal thought disorder

a. Derailment
b. Tangentiality
c. Incoherence
d. Illogicality
e. Circumstantiality
f. Pressure of speech
g. Distractible speech
h. Clanging

It is possible to conceive that any function of the mind might be lost, and it seems that the range of negative symptoms could be vast. In patient populations, the number of negative symptoms is fairly circumscribed. Many functions of the kind assessed with cognitive testing are impaired in schizophrenia, but except for deficits in the recognition of emotion, these functions are generally not included in the negative symptoms. There is a modest but consistent correlation between negative symptoms and cognition, but neither is well explained by the other, and they should be considered separate domains of psychopathology. More important, the kinds of function lost in negative symptoms in schizophrenia have been consistent in studies conducted at various times in various sites around the world.

There is also a weak but consistent relationship between levels of negative symptoms and levels of disorganization symptoms in populations with both kinds of symptoms. In both groups of symptoms, there is an increase in familial risk that is not seen for psychotic symptoms. There is a

greater concordance in monozygotic twins for negative versus psychotic symptoms, and there is an overall increase in negative symptoms in the relatives of patients with negative symptoms. Negative symptoms have also been found to correlate with measures of neuropathology, especially those measures of frontal lobe function or the volume of cerebral tissue.

Categories of Negative Symptoms. The MATRICS consensus conference on negative symptoms, convened under the auspices of the National Institutes of Mental Health (NIMH) in 2005, suggested that there are five general categories of negative symptoms: avolition, anhedonia, social withdrawal (asociality), affective blunting, and alogia. Subsequent analysis shows that these symptoms can be subdivided into two main groups—a group with loss of drive and pleasure—the avolition symptoms, also including anhedonia and asociality, and the affective blunting symptoms, including alogia. These two groups are the basis for DSM-5 including avolition and affective blunting as the representative negative symptoms in Criteria A. The first of these subdivisions—the avolition group—is assessed with interview, while affective blunting and alogia are signs observed by the clinician.

The most common negative symptoms are *avolition* and *anhedonia*. Avolition is the loss of will or drive (in neurology this is sometimes referred to as *abulia*). In psychiatry the loss of will or drive is sometimes considered to be a manifestation of anhedonia, but in fact, these refer to separate kinds of pathology. Avolition is similar to apathy, and these may be considered closely related, with avolition identifying a deficit in the ability to act and apathy a loss of concern for an idea or task. The definition of avolition in the DSM-5 glossary is "An inability to initiate and persist in goal-directed activities" and specifies subsequent deficits in work, intellectual pursuits, and self-care as resulting from avolition. Avolition in particular seems associated with deficits in grooming and hygiene, and it seriously impairs educational and vocational progress. Although often overlooked, this loss of will can be severely disabling for patients.

Anhedonia is a loss of the ability to find or derive pleasure from activities or relationships and may be the most persistent of the negative symptoms. In DSM-5, anhedonia is "Lack of enjoyment from, engagement in, or energy for life's experiences; deficits in the capacity to feel pleasure and take interest in things." Estimates of the prevalence of anhedonia in schizophrenia vary widely, but probably half of patients experience considerable anhedonia. Although present in depressive disorders, when anhedonia is present as part of a negative syndrome, it should not be considered a manifestation of depression. The concept of a loss of the ability to feel pleasure or reward can seem at odds with the description of their lives that patients provide. Tales of binging on junk food or chain-smoking throughout the day because "it's the only thing I like to do" seem

to suggest an overwhelming drive for reward that negates the suggestion of anhedonia in these patients. Recent investigation into the difference between the anticipation of pleasure and the ability to experience pleasure seems to reconcile the discordance between the idea of anhedonia and the experiences reported by patients. It has been found that people with schizophrenia and anhedonia can experience levels of pleasure similar to control populations while participating in pleasurable activities but experience much less anticipatory pleasure than controls. The loss of anticipatory pleasure correlates with measures of anhedonia in rating scales and with functional outcomes. A loss of anticipatory pleasure should decrease motivation, and this may be an explanation for the strong correlation between measures of avolition and anhedonia.

Social withdrawal is sometimes referred to as "passive or apathetic social withdrawal" and includes indifference to social relationships and decreases in the drive to socialize. This includes a decrease in the desire for relationships with family or friends, a decrease in sexual interests, constricted relationships with family and friends, and an inability to feel intimacy or closeness. Patients with social withdrawal have deficits in *theory of mind,* which refers to the ability to understand how another person might be thinking when there is common knowledge of their current circumstance.

Affective *blunting,* consisting of both an inability to understand or recognize displays of emotion from others and an inability to express emotion, is an important predictor of functional impairment in schizophrenia. The blunting of expression includes deficits in production, facial expression, gestures, and prosody, and understanding these social signals is similarly impaired. In the DSM-5 glossary, it is "blunted affect" and defined as a "reduction in the intensity of emotional expression," and "flat affect," the "absence or near absence of affective expression." It is more common in men, in people with an early onset illness, and in people with poor premorbid function, and it predicts lower scores on quality-of-life measures.

Alogia is a decrease in verbal communication, or "poverty of speech," and it is found in up to 25 percent of people with schizophrenia. Although alogia was previously considered both the loss of production and a paucity of content with a normal volume of words, it is only the loss of production of speech that is a negative symptom. The lack of speech production is considered to result from a decreased rate of verbal cognition. In this way, alogia is conceived of as a "negative thought disorder." This loss of production includes an increased latency to response, short verbal responses, and a paucity or complete lack of spontaneous production. In DSM-5, it is "impoverishment in thinking that is inferred from observing speech and language behavior. There may be brief and concrete replies to questions and restriction in the amount of spontaneous speech." The

paucity of meaningful content in the presence of a normal amount of speech that is sometimes included in alogia is actually a disorganization of thought and not a negative symptom and is properly included in the disorganization cluster of symptoms. Affective blunting and alogia are two measures of communication, and some authors have labeled these symptoms as "disorders of relating." Poor attention had been considered a negative symptom, but it is often more related to disorganization symptoms than to negative symptoms.

Course and Associated Features of Negative Symptoms

Although the prototype of a patient with negative symptoms is someone with years of illness, residing in a group home, and passive and indifferent to the events around him or her, negative symptoms can precede the onset of illness. Anhedonia and social isolation can predict schizophrenia in high-risk populations. These symptoms may be obscured by the drama around fear, agitation, and psychosis in first-break patients, but a broader range of negative symptoms can be found early in overt illness, with most patients having some negative symptoms. Anhedonia is most common, present in about half of patients at onset, and avolition and affective blunting are also common. If recognizable *akinesia* (an obvious difficulty in initiating motor activity) is present at the first admission, there is an increased risk of subsequent significant negative symptoms. All of the negative symptoms can progress during the first years of illness, and alogia, in particular, can increase from its relatively low rates at onset; up to 25 percent of patients will have significant negative symptoms within 2 years of onset of the illness. There is no course of illness that any particular patient can be expected to follow, and for many, there is a variable course for their negative symptoms, with significant expression at some points, and subsiding of symptoms at other times. But prominent negative symptoms at onset are better predictors of the future course of illness than are levels of psychotic or disorganization symptoms.

Negative symptoms arise from multiple sources, including iatrogenic causes. Long-term institutional care or subsistence-level existence on the streets and in shelters can remove normal, expectable pleasures and reinforcing activities and will narrow the number of relationships someone with schizophrenia can have. People with schizophrenia may lose the capacity to experience drives and social relationships and reward simply because circumstances eliminate them. A frequent and frequently unrecognized cause of negative symptoms is treatment with antipsychotic medications. Medications for the treatment of schizophrenia can cause a phenocopy of primary negative symptoms that can be distinguished only with a full history of illness or empiric trials of medication adjustment. Reports of excessive doses of antipsychotic medications causing blunting of affect, diminution of motor activity, impairments in grooming and hygiene,

and anhedonia have existed almost since the introduction of neuroleptics, but for many patients, overtreatment with medications remains the rule and not the exception. Extrapyramidal symptoms should not be a proxy measure for the presence of medication-induced negative symptoms, as extrapyramidal rating scales may remain stable even when dose changes sufficient to reduce negative symptoms have been made.

Negative symptoms can also arise from the presence of other psychiatric symptoms. Positive symptoms are a common cause of apparent apathy, social withdrawal, and alogia. Persecutory delusions can precipitate frantic attempts to protect safety but can also engender fear, leading to a very active social withdrawal, and prominent hallucinations can reinforce these fears. Frequent and intrusive hallucinations can lead directly to withdrawal and inaction, strictly through their effects on attention. Depression is common in schizophrenia, and with inattention or perfunctory clinical interviews, it can be overlooked, and the appearance of apathy and affective blunting can be attributed to the disease processes intrinsic to schizophrenia instead of a treatable mood disorder. Secondary causes of negative symptoms, with the exception of medication effects, will often remit within a year of their onset, and sometimes the passage of time can help distinguish primary from secondary negative symptoms; symptoms that do not diminish over the course of a year and do not remit with manipulation of the medication regimen should lead to consideration of primary negative symptoms.

Negative symptoms are frequently part of the clinical picture associated with exacerbations of symptoms. As is the case with depression in schizophrenia, the presence of negative symptoms may parallel an increase in positive symptoms, and these symptoms may remit when the positive symptoms remit. In these cases, use of antipsychotic medications would be expected to decrease the severity of negative symptoms. In fact, this is what is found in studies of medication treatment of recently admitted patients, where antipsychotic medication reliably reduces the severity of negative symptoms.

Primary Negative Symptoms. Primary negative symptoms are those negative symptoms that appear intrinsic to the disease process for the affected person. Primary negative symptoms are the least variable of the symptoms of schizophrenia, so that a person with prominent primary negative symptoms is likely to continue to have significant manifestation of these symptoms throughout the course of his or her illness. This persistence of negative symptoms predicts future disability. According to Bleuler, “It has been known since the early years of modern psychiatry that an ‘acute curable’ psychosis became ‘chronic’ when the affects began to disappear.” When there are multiple negative symptoms and a careful history suggests that these are enduring and not due to secondary causes,

the deficit syndrome (Table 12.2–4) is established.

People with schizophrenia and the deficit syndrome have a very poor prognosis regarding the possibility for a full functional recovery. They will tend to have a lower clinician-rated quality of life, more time in a hospital, an increased frequency of symptomatic periods, and more social and physical anhedonia. They will generally have lower intelligence quotients, less education, worse work histories, and they are less likely to be married and more likely to be men than people with schizophrenia who do not have the deficit syndrome. There is a more insidious onset, with a progression of negative symptoms, and more disorganization of behavior, and the negative symptoms in a person with the deficit syndrome are likely to respond poorly to medication treatment or social skills training. Not all symptoms and prognostic factors are worse in people with the deficit syndrome; they have no more delusions or hallucinations than nondeficit groups, they are less likely to have severe delusions around social themes and tend to have less severe suspiciousness, they are less likely to endorse prominent dysphoria, and they are less likely to have suicidal ideation or suicide attempts than nondeficit populations. Negative symptoms are associated with impairments in insight, and patients with significant negative symptoms rate themselves higher on subjective assessments of quality of life, despite lower scores on scales of function. There are profound implications for treatment and function for a patient with the diagnosis of the deficit syndrome, but reliable assessment has been found to depend on special training and the use of the Schedule for the Deficit Syndrome.

Table 12.2–4.
The Schedule for the Deficit Syndrome

The Schedule for the deficit syndrome requires the presence of two or more negative symptoms in a patient with schizophrenia. At least two symptoms should have been present in the preceding year, and the symptoms are not better attributed to other symptoms of schizophrenia, to another psychiatric illness, or to a medication effect.

The six negative symptoms listed it the schedule for the deficit syndrome are as follows:

1. Restricted affect
2. Diminished emotional range
3. Poverty of speech
4. Curbing of interests
5. Diminished sense of purpose
6. Diminished social drive

The deficit syndrome affects a minority of people with schizophrenia, but it is a sizable minority. The estimated rates of the deficit syndrome in schizophrenia vary widely, with rates between 10 and 30 percent; rates are lower early in course of illness and progress for the first 5 years. A meta-analysis of 23 studies calculated a rate of almost 30 percent. Demographic

associations for the deficit syndrome include an increase in summer births, reversing the usual pattern of risk for schizophrenia.

The deficit syndrome is clearly of great importance in clinical populations, but in usual treatment settings, time constraints and insufficient resources might make it impossible to differentiate long-standing and problematic negative symptom from true deficit symptoms. For many patients, a regular sequenced trial of medications and medication doses, attendance in psychosocial rehabilitation programs, or changes in living situations, which all might help tease out a secondary versus primary cause for negative symptoms, is not possible. So the distinction between deficit and nondeficit symptoms can be difficult or impossible for the practicing clinician. Robert Buchanan suggested the idea of "persistent negative symptoms" where possible secondary negative symptoms fail to respond to usual remediation methods. The persistent negative symptom criteria requires presence of a diagnosis of schizophrenia and at least two negative symptoms, for at least 12 months, present during periods of clinical stability and not due to psychiatric comorbidities or medication effects. The potential for extrapyramidal symptoms to cause negative symptoms is emphasized. If widely adopted, these simple-to-apply criteria might help focus increasing clinical resources on what is currently an underrecognized problem.

Disorganization

The disorganization syndrome has been part of the conception of schizophrenia since Kraepelin included hebephrenia in dementia praecox, but the behaviors, experiences, and associated pathology have been subject to less investigation than the psychotic and negative symptoms. Consequently, the boundaries of the disorganization syndrome are less clear than those of the psychosis and negative syndromes. However, the disorganization syndrome appears to be the most heritable of the subsyndromes of schizophrenia. Family and twin studies show an association between degrees of relatedness and the presence of prominent disorganization. This evidence parallels the finding that hebephrenia is the most heritable of the traditional subtypes of schizophrenia.

The disorganization symptoms certainly include the formal thought disorders, bizarre and catatonic behaviors, and inappropriate affect. Factor analytic studies consistently find that impairment of attention and poverty of content of speech fit best in the disorganization symptoms. Disorganization symptoms can affect almost any kind of behavior and disrupt almost any form of thought. The disruptions of behavior can be marked, with wildly inappropriate expressions of affect including odd or exaggerated gestures, singsong or childlike changes in prosody, silly or inappropriately bright affects, or alternatively, grimacing and wide-eyed expressions of surprise or anger. This disruption of affect can be socially

disabling and unfortunately seems to respond more poorly to treatment than other symptoms in the disorganization subsyndrome. During acute episodes, approximately 20 percent of patients exhibit an inappropriate affect. The motor manifestations of disorganization may change with time; recent studies of first-episode patients find that there is only a modest relationship between disorganized motor behavior and disorganization of speech, despite the strong correlations in chronic populations. Although previous versions of the DSM referred to "flat, blunted, or inappropriate affect" in the diagnostic criteria for schizophrenia, there is no reason to group inappropriate affect with the negative symptom of flat or blunted affects.

Disorganization can also include erratic behaviors such as broad movements of arms or the trunk as the patient ambulates, jerking the head or the neck, or posturing that can resemble aggressive stances. Hygiene is frequently poor, and as a group, those with disorganized schizophrenia are more indifferent to routine activities of daily living, or health care and health care maintenance, than are other patients with schizophrenia. Attire can be simply unusual or inappropriate (e.g., adults squeezing on adolescents' clothing, multiple layers of clothes during hot summer days, or light-weight attire in the middle of winter), or sometimes dramatically eccentric.

Although disorganization symptoms follow a variable course for most patients, there are a small number of patients who have multiple persistent disorganization symptoms throughout the course of their illness, and in some cases, these symptoms can be progressive. This group constitutes the traditional hebephrenic subtype of schizophrenia and is less prevalent than the group of patients with enduring primary negative symptoms or the deficit syndrome. These people are also markedly and persistently disabled and should be the subject of a greater research focus.

Motor Symptoms. Motor behaviors can include subtle repetitive hand movements or broad, complex, and purposeless movements that involve the limbs and the trunk. These movements can occur when speaking and seem to be the gestural equivalent of a neologism or can involve multiple movements of a limb or limbs, as though the person were performing some complex but undecipherable manual task. Patients may sometimes mimic the motor behaviors of others, which is referred to as *echopraxia*. Rocking when sitting down or standing is common, as is patients hugging themselves, wringing their hands, or toying with their clothing, or hair, or small objects around them. More complex motor behaviors, odd but apparently purposeful, have been described. There can be repetitive directed behaviors such as gestures and hugging others, or patients might have periods of appearing to read materials over and over (even if the text is upside down), or staring for protracted periods into the mirror. An early

onset of prominent motor symptoms has been associated with significant thought disorder and may predict greater disability and a deteriorating course.

Symptoms of catatonia are included in the disorganization symptom, although it is important to note that catatonia is more common in psychotic mood disorders and in cases with brain injury and encephalopathies than schizophrenia. This is made explicit in DSM-5, where catatonia can be a diagnostic specifier for other mental illnesses. The DSM-5 defines catatonia as "a marked decrease in reactivity to the environment" and its essential feature as a "psychomotor disturbance that may involve decreased motor activity, decreased engagement during interview or physical examination, or excessive and peculiar motor activity." In its appearance, it is as though central motor programs are engaged without direction from frontal areas engaged in higher-level planning. The diagnostic criteria require at least 3 of 12 motor signs; stupor, catalepsy, waxy flexibility, mutism, negativism, posturing, mannerism, stereotypy, agitation, grimacing, echolalia, and echopraxia.

There is a long history of description of catatonic symptoms, and many general classes have been identified. *Catatonic excitement,* with constant purposeless motor activity, often including both upper and lower limbs, might be the most common type seen in psychiatric settings in the Western world; in DSM-5, this would be "agitation." Catatonic *rigidity* is similar but refers to patients holding simple, fixed, rigid postures. DSM-5 clarifies some definitions—*negativism* refers to lack of response or automatic resistance instructions or stimuli, catatonic *posturing* refers to spontaneous and active maintenance of a posture, and *mannerisms* are an "odd, circumstantial caricature" of normal motor behaviors. Some people with *catatonic stupor* can be obviously awake but immobile without rigidity and will sit or lay in these postures until moved. *Catalepsy* is holding a posture held against gravity and is differentiated from waxy flexibility or *cerea flexibilitas,* which is in DSM-5 a tendency to pose modest resistance to repositioning of limbs. In the past, catalepsy and waxy flexibility were often used as synonyms. *Echopraxia* and *echolalia* retain their familiar meanings of unconscious mimicking action or speech, respectively.

Catatonia is less prevalent in the Western world than it was before the introduction of somatic therapies (convulsive therapies, antipsychotic medications). In old photographs or movies from state hospitals, odd postures and gestures, grimacing, or patients immobile and slumped in chairs are ubiquitous. In the first edition of this textbook, this description was given:

> Until the middle 1930s, mental hospitals were filled with stuporous catatonics, many of whom would lie motionless for weeks or months, unresponsive to almost every stimulus, their eyelids flickering, saliva drooling from their mouths, unresponsive to almost every stimulus. Since they

would not take any nourishment, they had to be fed by stomach tube twice a day as long as they remained stuporous.

Despite its decreasing prevalence in schizophrenia, some signs of catatonia are present in a large minority of patients during some phase of their illness. One estimate puts catatonia at 8 percent lifetime prevalence in schizophrenia, and DSM-5 suggests up to a 35 percent rate in inpatient settings. Unlike other markers of severity of illness in schizophrenia, the prevalence is essentially equal across genders. The risk may increase with increasing maternal age, and there is an increased risk of suicide, without an increase in the prevalence of mood disorders.

The experience of catatonia is often difficult to elicit from patients after an episode has resolved. Patients often have little recall, or will describe disjointed, impressionistic memories, like those of patients who have recovered from a prolonged delirium. Recollections of motivation or understanding the cognitive sequence that led to their behaviors are particularly hard to obtain. Other patients have recovered to report an awareness of activity outside of their control, without an ability to respond to others or organize any meaningful communication.

In the developing world, catatonia and other marked disturbances of behavior are more common in schizophrenia, and this may be particularly true in Africa. It may be that even the infrequent or inconsistent treatment with antipsychotic medications that patients are exposed to in industrialized societies has a prophylactic benefit for catatonia.

Thought Disorder. Thought disorder is the most studied form of the disorganization symptoms. It is referred to as "formal thought disorder," or "conceptual disorganization," or as the "disorganization factor" in various studies that examine cognition or subsyndromes in schizophrenia. It has been said that the thought disorder of schizophrenia expressed in speech is easy to recognize but hard to define, but that has not limited attempts to define it. As a primary care physician uses the fundoscopic examination as a way to visualize the vasculature of the central nervous system, in psychiatry speech is examined as a means to determine a patient's verbal cognition. "Thought disorder" here refers to disorganization of the form of thought and not content. An older use of the term "thought disorder" included the phenomena of delusions and sometimes hallucinations, but this is confusing and ignores the clear differences in the relationships between symptoms that have become apparent over the past 30 years. Delusions and hallucinations should be identified as psychotic symptoms, and thought disorder should be taken to mean formal thought disorders or a disorder of verbal cognition.

In its more mild manifestations, thought disorders do not preclude essential communication. Speech is frequently stilted or vague, and sentences may be incomplete. Speech may be *tangential,* so that the

associative chain moves obliquely off topic. With the progression of time, or an increased severity of illness, speech can further deteriorate, and inappropriate content can intrude. *Loosening of associations* refers to changes in topic without identifiable links between the subjects and indicates a lack of internal structure for thought expressed as speech. The DSM-5 glossary distinguishes *derailment* as a shift in topics between clauses, where *incoherence* is a shift within clauses. Patients with prominent thought disorder have more confabulation, more concrete speech, and engage in more symbolic speech. In its most severe form, patients can present with *neologisms* (novel words, and in schizophrenia, taken to indicate that the word has an idiosyncratic or nonsense meaning), *echolalia* (in DSM-5 "pathological, parrotlike, and apparently senseless repetition (echoing) of a word or phrase"), and incomprehensible speech without meaning or logical internal connections—*word salad*.

There is a recurrent or persistent identifiable thought disorder identified in 20 to 50 percent of subjects across a variety of studies, and this number goes up to approximately 75 percent of patients when they present with first episode or acute exacerbations of their illness. Of those subjects who do present with a thought disorder, many will have only mild or moderate disruption, and this can respond to treatment. The possibility that there is a secondary thought disorder, similar to secondary negative symptoms, due to distraction or anxiety that arises during psychotic exacerbations has been poorly explored but may account for some of the improvement seen following these episodes.

There are disturbances of thought or language found in many psychiatric disorders, but the thought disorder in schizophrenia is distinctive. When compared to subjects with bipolar disorders, siblings, and controls, those with schizophrenia are found to be more impaired in comprehension, attention, semantic organization, and fluency and complexity of speech. Prominent thought disorder at illness onset predicts a worse prognosis. It is associated with an earlier onset of illness, increased risk of hospitalization, decreased functional outcomes, higher rates of disability, and more inappropriate social behaviors. Verbal production tends to be less complex in patients with thought disorder. Thought disorder that does not respond well to treatment also predicts a worse course. In one study, only 20 percent of those subjects with a severe thought disorder at a posthospital follow-up regained a high level of function at a subsequent follow-up visit. There is a modest but consistent association with some measures of formal cognition, including deficits in abstracting ability, mental flexibility, planning, and both response initiation and inhibition. The thought disorders found in schizophrenia tend to be more stable in their severity than are hallucinations and delusions, and when prominent, they seem less likely to attenuate in middle age than positive symptoms.

Depression and Anxiety

Depression. Most people with schizophrenia will experience significant depression and anxiety during the course of their illness, and their frequency has provoked a discussion about whether these symptoms are comorbidities, or if they should be considered, like psychotic, disorganization, and negative symptoms, as intrinsic parts of schizophrenia. Awareness of their presence is not universal; clinicians and families may not notice depression and anxiety in their patients and relatives with schizophrenia or may be too distracted by patients' psychotic and disorganization symptoms to notice. Anxiety and depression contribute to the disability associated with schizophrenia, and anxiety and depression symptoms can worsen the other core symptoms of schizophrenia. Clinicians should attempt to rule in or rule out these conditions, so that if present, they can be properly treated.

Depression can be confused with negative symptoms and vice versa, though there is almost no concordance between negative symptoms and depression in schizophrenia. There is also an unfortunate tendency to consider that depression cannot occur in schizophrenia simply because DSM-III and DSM-IV excluded major depressive in the presence of schizophrenia and most other psychotic disorders. The diagnosis of major depression in DSM-5 is neither specifically excluded nor endorsed, though the reality of depression in schizophrenia is acknowledged in the presence of a depression dimension in the Clinician Rated Dimensions of Psychosis Symptom Severity. Despite the uncertain role for depression symptoms in diagnostic schemes for schizophrenia, when patients are asked about the experience of depression they are usually fully capable of acknowledging its presence. In fact, patients are more likely to endorse their past experiences of depression than their past psychotic experiences. Depression has been variably reported as a good prognostic feature and as a poor prognostic feature; it seems likely that depression in schizophrenia carries the same costs it does for other people, and it clearly adds to the burden of the illness and impairs functional recovery. However, the presence of significant depression might also be an indicator of the absence of the deficit syndrome and the long-term disability associated with primary negative symptoms.

Depression is a common symptom in the early course of illness, with one-half or more of all patients experiencing significantly depressed mood during the prodrome and first break. In cross-sectional studies, the prevalence of major depression is fairly low, with rates of around 10 percent. When symptoms are assessed, instead of a complete depressive syndrome, patients followed long-term after first hospitalization reveal rates of some depressive symptoms of up to 75 or 80 percent. In one group of patients followed from their first episode, follow-up visits showed

depression present in about one-quarter of visits when the subjects were also psychotic; in almost all cases, this depression resolved in parallel with resolution of the psychotic exacerbation. Estimates of the lifetime rates of depression vary widely and are often based on point prevalence studies. In studies with ascertainment periods of up to 15 years, find significant depression or a full depressive syndrome in from 25 to 75 percent of subjects. Some subjects in these studies never exhibit depression, and this group is similar in size (20 to 30 percent) to the percentage that would be expected to have a deficit syndrome. It is possible that many of these patients may not be able to generate enough affective experience to feel real depression.

The presentation of depression in schizophrenia may be masked by the frequent focus in clinical settings on the vegetative symptoms of depression. Disturbances of sleep, appetite, energy, and concentration can be so common in schizophrenia that they become useless in screening for depression in these patients. Also, antipsychotic medications can produce a syndrome that appears identical to major depression, and an increasing emphasis on mood symptoms, poor self-esteem, guilt, and nihilism may be necessary to differentiate a primary mood state from medication effects. The Calgary Depression Scale (Table 12.2–5) emphasizes the psychic symptoms of depression, has good reliability, and is useful for detecting changes in mood in schizophrenia. This scale should be the first choice for monitoring depression and treatment of depression in schizophrenia.

There has been a common focus on postpsychotic depression in studies of depression in schizophrenia. Postpsychotic depression has been attributed to demoralization and increasing insight following resolution of exacerbations of psychotic symptoms, with patients' increasing insight revealing to them the severity of their illness and the likelihood of a future of disability and relapse. Although periods of depression that follow acute exacerbations of symptoms are periods of increased risk of suicide, there is little else to distinguish them from depression that arises during periods of relative stability. At any time in the illness, the combination of depression symptoms and increased insight predicts an increasing sense of distress among patients.

Table 12.2–5.
Items in the Calgary Depression Scale for Schizophrenia

1. Depression
2. Hopelessness
3. Self-depreciation
4. Guilty ideas of reference
5. Pathological guilt
6. Morning depression
7. Early wakening
8. Suicide

Anxiety. Since the last edition of this text, there has been considerable progress toward establishing the impact of anxiety disorders in schizophrenia. Much of this progress has come in the form of systematic reviews and meta-analyses of existing studies, and these reports confirm a high frequency of anxiety disorders in schizophrenia. The presentation of frequent anxiety complaints by patients with schizophrenia is well known to most clinicians, but the frequency of full syndromes fulfilling diagnostic criteria for specific disorders is probably not as well known. Anxiety disorder diagnoses are not excluded in schizophrenia by DSM-5 criteria, though there is a suggestion that specific phobias are sometimes better explained by delusions, and if generalized worry occurs only during the course of schizophrenia, generalized anxiety disorder should not be diagnosed. While obsessive–compulsive spectrum disorders and traumatic stress disorders are not considered "Anxiety Disorders" in DSM-5, their impact is reviewed here because the literature has continued to group them as anxiety disorders.

Social phobia, panic attacks and panic disorder, and generalized anxiety disorder are common in schizophrenia. Studies of the prevalence of anxiety disorders in schizophrenia find social phobia to be the most common anxiety disorder, with a prevalence of 15 percent. Life histories of patients with schizophrenia frequently find premorbid social anxiety that continues through the prodrome and first stages of illness and adds considerably to the social, educational, and vocational deficits associated with the illness. There is evidence of increased suicidality, decreased functional recovery, and self-esteem in patients with social phobia, but there is a relative paucity of research about social phobia in schizophrenia as compared to studies of panic, traumatic stress disorders, and obsessive–compulsive spectrum disorders.

Panic attacks are more common than panic disorder, and in some patients, the frequency of panic attacks peaks at around the onset of psychotic symptoms. Although there is a wide range of estimates of the prevalence of panic attacks in schizophrenia, an estimate of 30 percent of patients experiencing panic at some point in their illness is reasonable. There is also a broad range of estimates of the frequency of panic disorder in schizophrenia, but a recent meta-analysis finds a rate of 12 percent, as compared to rates of 2 or 3 percent in the general population. In schizophrenia, panic seems especially related to psychotic symptoms, and there might be feed-forward cycles of persecutory delusions and related hallucinations precipitating panic, which then can worsen psychosis. Unaffected relatives of patients with schizophrenia and panic have increased risks of panic, raising the possibility of a neurobiological link between panic and schizophrenia.

There is a high prevalence of posttraumatic stress disorder (PTSD) in schizophrenia, but the relationship between the two diagnoses makes it difficult to tease out causation versus shared risks and increased vulnerability. Stress, including traumatic stress, during childhood and adolescence increases the risks of developing schizophrenia, but it is also true that the difficult life situations for many people with schizophrenia puts them at an increased risk of experiencing traumatic stressors. In fact, there is a growing body of evidence that shows psychotic experiences themselves can be traumatic stresses. Inpatients and outpatients in public mental health facilities in the Eastern United States were found to have a rate of about 30 percent of schizophrenia and related diagnoses in male and female patients; the rates of physical and sexual assault experienced by patients of both genders in this study were very high. A meta-analysis found a rate of PTSD in schizophrenia of 12 percent, which is higher than the prevalence in the general population. The comorbidities that attend PTSD are already increased in schizophrenia, and the combination of PTSD and schizophrenia further amplifies their severity and effect. The core symptoms of schizophrenia are more severe in patients with comorbid PTSD, depression is more common, suicidal ideation and suicide attempts are increased, patients require more care (including requiring more hospitalizations), and it is likely that substance abuse rates are higher in this population.

Probably because of the well-defined neural circuitry in obsessive-compulsive disorder (OCD) and the overlap in these circuits with some of those implicated in schizophrenia, there has been informed speculation about potential overlap in the underlying pathology of schizophrenia and OCD. It does appear that the prevalence of obsessive-compulsive symptoms (OCS) and OCD in schizophrenia exceeds the prevalence in the general population, with estimates of up to a third of people with schizophrenia having comorbid OCD. However, ascertainment of OCD can be difficult, with a special difficulty separating delusions and ruminative concern about these delusions from true obsessions. Systematic analyses suggest somewhat lower rates, with a review finding a rate of 25 percent for patients with OCD and only slightly more with OCS. A meta-analysis found an even lower rate of 12 percent, and an analysis of CATIE data found only 5 percent of patients with OCD. Although there have been some downward revisions in the estimate of OCD in schizophrenia, in each case, the rates remain considerably higher than the rate in the general population.

Comorbid OCD was once considered a good prognostic factor in schizophrenia, but this is clearly not the case. Obsessive–compulsive symptoms are associated with an earlier age of onset, and OCS or OCD often precedes the onset of psychosis. The presence of OCS or OCD is associated with more severe psychotic symptoms, more depression, more disability, increased hospitalizations, and an increased risk for suicide

attempts. As with panic, there is an increased risk for OCD in the relatives of patients with schizophrenia and OCD as compared to relatives with schizophrenia alone. A considerable body of literature has arisen over the last 30 years that suggests that for a subset of patients with schizophrenia a real "schizo-obsessive" form of the illness might exist.

Agitation and Hostility

The clinical features of agitation, irritability or frank hostility, uncooperativeness, and impulsivity are familiar to those who deal with clinical populations with schizophrenia, and these symptoms segregate together independent from the influence of culture. They are found more commonly in male patients and in those with substantial cognitive impairments. Comorbidities and treatment variables can increase the risk for agitation, with substance abuse particularly prominent. There are other treatment-specific risks, including poor medication adherence, exacerbations of psychosis, and extrapyramidal symptoms (including tardive dyskinesia) that increase the risk of agitation and hostility. Unfortunately, beyond this, little progress has been made in understanding agitation in schizophrenia. The link between agitation and mania has not been explored; it seems intuitively obvious that this might be a subsyndrome, such as psychosis, in which some of the genetic overlaps between schizophrenia and bipolar disorders are expressed. The evidence base that would clarify whether agitation and mania are variable manifestations of a similar underlying process or analogous manifestations of unrelated causes does not exist. Over the last decade, the predominant focus in research into agitation has concerned its links to violence.

Violence. The topic of violence in schizophrenia is contentious. In the mind of the general public schizophrenia is sometimes synonymous with unreasoned violence. For people with schizophrenia, and for their families and advocates, the real problem is that people with schizophrenia are far more likely to be the victims of violence than the perpetrators. Although this is true, and although there are studies with negative findings, it also remains true that a small minority of people with schizophrenia will commit a disproportionate amount of violence and even murder. Recent systematic reviews and meta-analyses have led to some downward revision in the estimated rates of violence among those with schizophrenia, especially in the absence of psychiatric comorbidities, but the finding of an increased rate remains. Large cohort studies and smaller diagnosis-specific studies, conducted in a variety of countries that have varying baseline rates of violence, show that the presence of a diagnosis of schizophrenia increases the risks of committing violence or being arrested.

No single variable associated with schizophrenia accounts for the increased rates of violence. A series of factors increase risks, though these

factors are so prevalent in the population that the predictive power of any one feature is very small. Some demographic and premorbid features are associated with violence, including gender, with men committing the majority of violent acts. Violent victimization, including abuse during childhood or adolescence, increases rates of violence in adults with schizophrenia, as does a history of poor impulse control. Some people with schizophrenia will exhibit a premorbid conduct disorder in childhood and will move into more serious antisocial behaviors as they age. This group is particularly likely to engage in severe violence and to calculate and plan their violent acts. How much of this violence is due to antisocial personality disorder and how much can be attributed to schizophrenia is uncertain, but clearly, this is not solely a problem due to a diagnosis of schizophrenia.

Some factors intrinsic to schizophrenia also increase the risks of violence. Poor impulse control and cognitive deficits can predate the onset of the illness but can also arise with the expression of the illness. Insight is variable in schizophrenia, but those with significant impairments have an increased risk of violence. Not surprisingly, threat delusions, especially those that provoke negative affects, increase risks of violence, and the deficits in interpreting social and emotional communication that are common in schizophrenia also increase risks. As reviewed in the discussion of auditory hallucinations, the congruence of the content of hallucinations with a delusional structure that focuses on threat and the benefits of compliance increases the likelihood of following the commands; command hallucinations without this delusional structure likely pose little threat. Delusions in the absence of hallucinations are often sufficient to provoke violence. For many people with schizophrenia, the persecutory delusions that make them believe their lives are at risk are significant contributors to their violent behaviors, and for those suffering from the delusions violence is an act of self-defense. The influence of paranoid delusions on an increased rate of violence persists even when the results are controlled for the effects of demographic factors and substance abuse. In particular, systematized persecutory delusions that suggest that the person is a particular target, with co-occurring delusions of control, may increase the rates of violence. In this group, there is a higher rate of violence by women, and in inpatient psychiatric units, the rates of assault by men and women may be equivalent. It has been suggested that confusion, frequently seen in emergency and inpatient settings, contributes to a significant percentage of assaults in these settings. Acute exacerbations of illness are usually times of a higher burden of all symptoms, which also increases the risk of violence.

A broad range of comorbidities also increases the risk of violence. The effects of substance abuse, in particular, the effects of alcohol abuse or dependence parallel the effects of substance abuse in the general population. A widely reported meta-analysis found an odds ratio for violence of almost 9 with comorbid schizophrenia and substance abuse

diagnoses, and an odds ratio of 2 with schizophrenia alone. Reporting at the time emphasized the dramatic increase in risk with substance abuse, but the study also clearly showed that there is increase without the presence of substance abuse, and it poses a clinical and societal problem. Along with substance abuse, and frequently co-occurring with substance abuse, nonadherence with medication treatment is associated with increased rates of violence in people with schizophrenia. Suicide threats and attempts, and neurologic comorbidities, each increases rates.

In chronically ill and disabled patients, there are often less severe acts of violence, frequently directed against family, friends, caregivers, and health care providers. This may be the most common type of assault in treatment settings and frequently occurs as medical staff members attempt to modify patient behavior. These incidents seem more likely related to the effects of impoverished and constricted lives, where there is difficulty for patients to gain experience modulating affect and impulses. Impulsivity and poorly moderated affect can lead to transient displays of aggression, and the effects of substance abuse frequently exacerbate these.

People with schizophrenia are also overrepresented in the numbers of people who commit homicide. One estimate puts the percentage of homicide offenders with schizophrenia at greater than 6 percent; a common estimate is that the risk of homicide is 10 to 20 times higher in schizophrenia. In any location, the risks reflect the community rate of homicide, so that areas with low homicide rates see smaller rates by people with schizophrenia, and the rates rise in areas with higher community rates. The MacArthur study found that violence was more likely to be directed at family members than people outside the family or strangers. Multiple other lines of evidence suggest that there is a considerable risk for stranger homicide by people with schizophrenia. The large majority of these individuals were men; many were in the early stages of their illness with a recent exacerbation of their illness and had a recent change in the quality of their delusions. Homelessness, a lack of psychiatric treatment, and conduct disorder or a history of antisocial behaviors are also associated with stranger homicide.

Suicide. Suicide accounts for a good part of the excess mortality that is usually found in schizophrenia. Twenty to 40 percent of people with schizophrenia will make a suicide attempt sometime during their illness. A commonly cited figure for suicide rate in schizophrenia is 10 percent, although there is little contemporary evidence to support this figure. Reanalysis of the evidence suggests that 5 percent of people with schizophrenia will commit suicide, with increased risk early in the course of illness, with frequent relapses, and at around the time of admission or immediately following discharge from a psychiatric facility. Also, periods of increased psychosis, especially when paranoid delusions are prominent,

preserved or emerging insight, and a fear of mental deterioration seem to increase risk in this population. Nonspecific factors such as depression, hopelessness, motor restlessness and agitation, impulsivity, and anxiety increase the risk for suicide in schizophrenia, as they do in other psychiatric illnesses. People with schizophrenia who attempt suicide tend to use more lethal means in their attempts, have a strong intent to die, and make multiple attempts. As noted previously, patients with significant and persistent primary negative symptoms seem to have a lower risk for suicide.

Insight

Lack of insight was the most common symptom of schizophrenia in the IPSS studies of symptoms and course of schizophrenia. Impaired insight has long been recognized as contributing to diminished functional outcome, lowered quality of life, decreased treatment adherence, increased levels of psychopathology, and recurrent exacerbations of illness, and it has been clear that better understanding of insight has the potential to generate treatment strategies that significantly improve outcomes—though there is evidence that better insight is associated with increased suicidality, depression, and self-reported distress. For decades, the primary investigators of insight in schizophrenia came from the philosophic tradition of phenomenology, where the aim is to understand as completely as possible the subjective experience of a person; in schizophrenia, the major focus has been on the psychotic symptoms and the nature of the perception and understanding in psychosis. The functional consequences of impaired insight or the neurobiology behind the impairments are less easily appreciated in this approach.

A lack of self-reflection is frequent in schizophrenia, and as is seen in delusions, there can be an undue certainty in judgments that are based on insufficient evidence. Splits in insight are often observed in patients, and it is common in chronically ill patients to find someone who accepts that treatment helps functioning and prevents worsening of symptoms but who deny illness and the idea that psychotic symptoms are mental phenomena. Poor insight is correlated with negative symptoms, inappropriate affect and thought disorder, and delusions, although the strength of the correlations is fairly weak. Over the last 20 years, there has been a broad surge of interest in insight in schizophrenia, continuing the phenomenological approach but also driven in part by broad emphasis on understanding the neuropsychological underpinnings of schizophrenia, in part because of the emphasis in outcomes over symptoms, and in part because of the interest of cognitive therapists.

Insight is not a unitary phenomenon, and impaired insight is not due to any single underlying process. There is a continuum of insight, from complete understanding of illness and effects of illness on function and

production of symptoms to understanding that auditory hallucinations are associated with illness but nonetheless feeling that the voices are by some external agent to total denial of illness and symptoms and dysfunction. Cognitive mechanisms certainly play a part, and there are relationships between neurocognition as measured with neuropsychological testing and the severity of impairment of insight, but results are inconsistent and taken together the relationships are modest. A separate approach addresses "clinical insight" and focuses on those aspects of insight that effect care or are observed in clinical settings. One useful scheme for conceptualizing insight in schizophrenia suggests four dimensions: awareness of illness, awareness of symptoms and attribution of symptoms to illness, awareness of the need for treatment, and awareness of the consequences of the illness. There is separately a recent focus driven by on "Cognitive Insight," which actually refers to the metacognitive processes that underlie the ability to reflect on conscious cognition and the certainty underlying cognitive processes. Presumably, there will soon be an elaboration of cognitive therapy techniques to address this aspect of insight.

Illustrative Case

A 37-year-old man presents for assessment. He has a chief complaint of "don't like my medicine," though a review of his medication profile shows only irregular requests for refills of the several different atypical antipsychotic agents and serotonin reuptake inhibitors he has been prescribed over the last year. There are notes by two other psychiatrists in the electronic record, but this record includes only a handful of notes and a pattern of infrequent visits and missed appointments over the last 4 years.

As the appointment proceeds, he explains that he feels as though his current medication "makes me too tired" and "makes me fat" but also acknowledges that he has not been taking the medication "for a while" and that current complaints of an altered sleep–wake cycle, daytime fatigue, increased appetite, and lack of self-initiated hobbies or social activities cannot be attributed to medication effects. He denies depressed mood, but blunted responses of affect and little investment in ideas or hopes for his future are observed. He discloses regular but brief verbal auditory hallucinations of a male and female voice that are familiar to him but that he does not know of the identity of either voice. He feels as though the voices are entities controlled by some unknown group "who have been watching me for a long time" with cameras, microphones, and airplanes. He feels "They used to want to get me" and do physical harm to him, but for an unknown period, they have been focused just on monitoring his whereabouts and "what I'm thinking." Mild perioral and lingual dyskinetic movements and occasional grimacing are evident, but he is also seen to shift to awkward postures and there are occasional complex repetitive movements of his fingers that are judged to be mannerisms. Hygiene is good, but he has not recently shaved, and his hair and clothing are unkempt. Speech is very terse and frequently lacking in pertinent detail, with occasional allusions to ideas that are not clear and cannot be explained. He is aware of the clinical setting, his living situation, and the month and the year but cannot recall the day of the month or the week. After some extensive prodding, he agrees to work with his uncle to get his antipsychotic medications refilled and to follow-up within a month.

On his return visit, there is a dramatic decrease in the rate of ambulation, few spontaneous directed movements, and he has a flat, slack-jawed appearance. There is a clear low-frequency tremor in his hands, more prominent in the right than in the left, and significant stiffness in the elbows bilaterally, but the oral dyskinesias are now absent. Speech production is generally reduced to one- or two-word answers, but he does say that his auditory hallucinations are

gone. His uncle accompanies him and confirms good adherence to medication treatment, but that although the patient is much slower, his temper seems worse. The patient says that he feels “real bad” and wants to discontinue his medications and wishes he could “go away,” which he indicates means to escape the experience of existence, but he does not want to die or kill himself. He nods assent when asked whether he is depressed. He is prescribed a lower dose of antipsychotic medication and as needed anticholinergic medications, with follow-up in 2 weeks.

In the interval, copies of his old paper-based chart are retrieved, and there is a more thorough review of his care. He was first seen in a child clinic for presumed depression at 16 years of age, with significant social withdrawal and declining academic performance, along with decreased participation in sports and music classes, disrupted sleep–wake cycle, and frequent irritability. He was treated with antidepressant medications and brief individual and family therapy. His school performance improved slightly, but he remained withdrawn from former friends and interests. After 15 months, he was discharged from care.

He was next seen in the hospital system at 20 years of age. He had graduated from high school but then held a series of entry-level jobs for very short periods and frequently fired after an inability to get to work on time or inadequate job performance. His parents suspected drug abuse but never found any evidence of illicit drug use. After a half year of getting, and getting fired from, a series of jobs, he stopped looking for jobs and stayed at home. He was occasionally incomprehensible when explaining his online interests to his parents and became more and more cryptic and more and more irritable with them while his daily rhythms and relationships fragmented. His return to the mental health system was precipitated by an arrest, and then transfer from jail, after he threatened his parents with some broken glass while arguing that they were “helping to poison me.”

During his first admission, he explained that his parents were trying to poison him so that they could steal his ideas; he knew this because people who were simultaneously plotting with his parents and trying to protect him were warning him through radio waves. He was very agitated, and there were also infrequent but odd motor movements in gait, posture, and gesture. He had a slow initial response to a high dose of an atypical antipsychotic agent, and the slow pace of his response prompted an extensive medical workup by his team. The only significant findings were of mild obesity, a modest elevation in fasting glucose, and a magnetic resonance image of the head that showed lateral ventricles judged to be “minimally enlarged for patient’s age.” After 18 days on the inpatient unit, he was discharged from the clinic.

Because of resistance on the part of the patient, follow-up was erratic through most of his twenties. He was seen in several of the other community hospitals in the same city and treated with a variety of antipsychotic, mood-stabilizing, and antidepressant medications. Several years after his first admission, he was transferred to an inpatient unit from an intensive care unit, where he had been brought treated after stabbing himself in the neck after a period of nonadherence and increasing psychosis. Through his twenties, there were two other suicide attempts, the onset of alcohol abuse and heavy cigarette smoking, an inability to find employment, and variable periods of treatment adherence and nonadherence. During one stretch of almost 2 years in his late twenties, he went to a day hospital program and then to a psychosocial rehabilitation program where he made friendly acquaintances with whom he would sometimes go out to eat or go to a movie. Despite these activities, he still spent most of his time at home, watching television, eating, and smoking cigarettes. For some reason he could not explain, he discontinued the medication he had been taking, got progressively more psychotic, complained of recurrent depressed mood, and had delusions of persecution reinforced by the content of his auditory hallucinations. He was again threatening to his parents and had his last psychiatric admission. After discharge from this hospital stay his parents were unwilling to take him back home, and he moved in his uncle. He has been treated in the same system since the last hospitalization, with tobacco use, obesity, and hyperlipidemia as comorbidities.

At subsequent follow-up visits, a low dose of antipsychotic is found that causes minimal stiffness and no other perceptible extrapyramidal symptoms. The patient and his uncle are educated on sleep hygiene, and his sleep–wake cycle stabilizes. The depressed mood and wish to not exist expressed at the second visit quickly diminish with parallel with the resolution of his extra-pyramidal symptoms. His auditory hallucinations returned, but with discussion, it becomes clear that he usually finds the content of the voices helpful, and that they provide

companionship that he likes; he is reluctant to consider trials of other medications because they might silence his voices. His dyskinetic movements return, but the odd postures and gestures are alleviated. Despite urging from family and his psychiatrist, his paucity of external activities and lack of exercise only slowly improve, though his better mood allows him to go out to eat or see movies with a variety of family members. Efforts to return him to a psychosocial rehabilitation program are ongoing.

RESEARCH IMPLICATIONS

The recognition and adoption of the five-factor model of symptoms in schizophrenia research has been comprehensive and rigorous in some areas and disappointingly weak in others. In investigations into novel therapeutics, including psychosocial therapies, there has been widespread acceptance of the model, and the new therapies are often designed to address deficits in treatment in a particular factor. The NIMH MATRICS Consensus Conference on negative symptoms and the endorsement of negative symptoms as a therapeutic target serve as a model for pursuing this kind of research. As noted earlier in the chapter, investigation into the nature of some symptoms, particularly in the disorganization and agitation factors, lags behind the effort in psychosis and negative symptoms. There are clearly opportunities to use phenomenological approaches to better define the intersections between schizophrenia and bipolar disorders, especially in psychosis and agitation symptoms, and to carry out parallel investigations into concordant (or discordant) nature of negative and disorganization symptoms in schizophrenia and the symptoms of autism spectrum illness and intellectual disabilities. In studies that attempt to relate neuroimaging findings to symptoms in schizophrenia, there has been very poor recognition of the benefit of using factor-derived symptom clusters, and there remains an entrenched pattern using SAPS, SANS, or PANSS Positive and PANSS Negative scales. This is particularly puzzling, because researchers no longer have to generate a factor analysis from their own data. Wallwork and colleagues generated a five-factor model for the PANSS built on previous models and on their own analysis that has provided a good fit across multiple separate studies conducted in North and South America, Europe, and Asia. This model allows investigators to plug their data into a set factor structure for analysis; there is no guarantee that its widespread use would strengthen the relationship between neuroimaging or genetic findings and symptoms, but not adopting empirically derived factors certainly decreases the likelihood of finding associations.

REFERENCES

Addington D, Addington J, Maticka-Tyndale E. Assessing depression in schizophrenia: the Calgary Depression Scale. *Br J Psychiatry Suppl.* 1993;22:39–44.

Andreasen NC. *The Scale for the Assessment of Negative Symptoms (SANS).* Iowa City, IA: University of Iowa; 1983.

Andreasen NC. *The Scale for the Assessment of Positive Symptoms (SAPS)*. Iowa City, IA: University of Iowa; 1984.

Barch DM, Bustillo J, Gaebel W, et al. Logic and justification for dimensional assessment of symptoms and related clinical phenomena in psychosis: relevance to DSM-5. *Schizophr Res*. 2013;150:15–20.

Berrios GE. Phenomenology, psychopathology and Jaspers: a conceptual history. *Hist Psychiatry*. 1992;3:303–327.

Bleuler E. *Textbook of Psychiatry*. New York: Macmillan; 1924.

*Bleuler E. *Dementia Praecox or the Group of Schizophrenias*. New York: International Universities Press; 1952.

Buchanan RW. Persistent negative symptoms in schizophrenia: an overview. *Schizophr Bull*. 2007;33:1013–1022.

Carpenter WT Jr, Strauss JS, Bartko JS. The diagnosis and understanding of schizophrenia. Part I. Use of signs and symptoms for the identification of schizophrenic patients. *Schizophr Bull*. 1974;11:37–49.

Crow TJ. The two syndrome concept: origins and current status. *Schizophr Bull*. 1985;11:471–486.

Fazel S, Gulati G, Linsell L, Geddes JR, Grann M. Schizophrenia and violence: systematic review and meta-analysis. *PLoS Med*. 2009;6:e1000120.

Feighner JP, Robins E, Guze SB, Woodruff RA, Winokur G, Munoz R. Diagnostic criteria for use in psychiatric research. *Arch Gen Psychiatry*. 1972;26:57–63.

Gard DE, Kring AM, Gard MG, Horan WP, Green MF. Anhedonia in schizophrenia: distinctions between anticipatory and consummatory pleasure. *Schizophr Res*. 2007;93:253–260.

Garrett M, Silva R. Auditory hallucinations, source monitoring, and the belief that "voices" are real. *Schizophr Bull*. 2003;29:445–457.

Harrow M, Jobe TH. How frequent is chronic multiyear delusional activity and recovery in schizophrenia: a 20-year multi–follow-up. *Schizophr Bull*. 2010;36:192–204.

Jaspers K. *General Psychopathology*. Hoenig J, Hamilton MW, trans-eds. Chicago, IL: University of Chicago Press; 1968.

Kay SR, Opler LA, Lindenmayer JP. Reliability and validity of the positive and negative syndrome scale for schizophrenics. *Psychiatry Res*. 1988;23:99–110.

*Kirkpatrick K, Buchanan RW, Ross DE, Carpenter WT Jr. A separate disease within the syndrome of schizophrenia. *Arch Gen Psychiatry*. 2001;58:165–171.

*Kraepelin E. *Dementia Praecox and Paraphrenia*. Edinburgh: Livingstone; 1919.

Liemburg E, Castelein S, Stewart R, van der Gaag M, Aleman A, Knegtering H. Two subdomains of negative symptoms in psychotic disorders: established and confirmed in two large cohorts. *J Psychiatr Res*. 2013;47:718–725.

Nayani TH, David AS. The auditory hallucination: a phenomenological survey. *Psychol Med*. 1996;26:177–189.

Overall JE, Gorham DR. The brief psychiatric rating scale. *Psychol Rep*. 1962;10:799.

Pompili M, Amador XA, Girardi P, et al. Suicide risk in schizophrenia: learning from the past to change the future. *Ann Gen Psychiatry*. 2007;6:1.

Rabinowitz J, Berardo CG, Bugarski-Kirola D, Marder S. Association of prominent positive and

prominent negative symptoms and functional health, well-being, healthcare-related quality of life and family burden: a CATIE analysis. *Schizophr Res*. 2013;150:339–342.

Schneider K. *Clinical Psychopathology*. New York: Grune & Stratton; 1959.

Shinn A, Heckers S, Öngur D. The special treatment of first rank auditory hallucinations and bizarre delusions in the diagnosis of schizophrenia. *Schizophr Res*. 2013;146:17–21.

Spitzer RL, Endicott J, Robins E. *Research Diagnostic Criteria*. New York: Biometrics Research Division, New York State Psychiatric Institute; 1975.

Stefanovics EA, Elkis H, Zhening L, Zhang XY, Rosenheck RA. A cross-national factor analytic comparison of three models of PANSS symptoms in schizophrenia. *Psychiatr Res*. 2014;219:283–289.

Strauss JS, Carpenter WT Jr, Bartko JJ. The diagnosis and understanding of schizophrenia. Part III. Speculations on the processes that underlie schizophrenic symptoms and signs. *Schizophr Bull*. 1974;1:61–69.

Wallwork RS, Fortgang R, Hashimoto R, Weinberger DR, Dickinson D. Searching for a consensus five-factor model of the Positive and Negative Syndrome Scale. *Schizophr Res*. 2012;137:246–250.

▲ 12.3 Worldwide Burden of Schizophrenia

ASSEN JABLENSKY, M.D.

In the 1990s, the WHO, in collaboration with the World Bank and the Harvard School of Public Health, initiated the Global Burden of Disease (GBD) study that produced estimates of morbidity and mortality from all causes by age, sex, and geographical region, introducing for the first time a new integrated unit of measurement, the disability-adjusted life year (DALY). One DALY corresponds to 1 year of healthy life lost in a given population due to the combined effects of disability and premature mortality resulting from disease. For any given disease, DALYs sum up the years of life lost to mortality and years of healthy life lost to disability. The surprising outcome of the GBD assessment for the year 2000 was that, worldwide, several psychiatric conditions, including unipolar depressive disorders, bipolar affective disorder, schizophrenia, and self-inflicted injuries were among the top 10 conditions accounting for the largest proportions of the total DALYs in age groups 15 to 44 years for both sexes among 135 diseases or health conditions. In terms of DALYs, schizophrenia ranked eighth, accounting for 2.6 percent of the total, and in terms of years lost to disability (YLD), it was third, accounting for 4.9 percent of the total. The findings on schizophrenia are likely to reflect broadly the burden of morbidity and disability that the disorder inflicts on individuals, families, and communities worldwide. Global DALYs have remained stable from 1990 to 2010. However, global disease burden has continued to shift away from communicable to noncommunicable diseases and from premature

death to years lived with disability. Important shifts have occurred in DALY composition, including a 34 percent increase in major depressive disorders and other psychiatric conditions, indicating a rising burden from mental and behavioral disorders.

This chapter reviews evidence from past and present research, bearing on the individual components that jointly make up the aggregate burden of disease and disability associated with schizophrenia—the variations in its incidence and prevalence, course and outcome, comorbidity and mortality, and socioeconomic costs. In evaluating the evidence, it is important to bear in mind that schizophrenia still remains a broad clinical syndrome (rather than a "disease"), and that the neurobiology and genetics underlying its manifestations are not yet fully understood.

However, research into the genetic underpinnings of schizophrenia recently reached a new milestone with the genome-wide association study (GWAS) of the largest sample to date of patients and healthy controls (36,989 cases and 113,075 controls), which identified 108 associated genomic loci significantly associated with schizophrenia risk. The study, by an international consortium of some 80 research groups, found robust associations of the disorder with the dopamine receptor DRD2, as well as with molecules involved in glutamatergic neurotransmission, suggesting plausible pathogenetic pathways. A surprising novel discovery was the strongest association of schizophrenia risk with the major histocompatibility complex on chromosome 6, pointing to an important, yet poorly understood, role of the immune system in the pathophysiology of the disorder. However, these findings still explain a relatively small fraction (~30 percent) of the estimated heritability of schizophrenia and can be seen as the prelude to the next-generation whole-genome sequencing studies. Such studies will have the potential to identify rare genomic variants (mutations present in less than 2 to 5 percent of the population) that may contribute to familial and individual risk of the disorder.

The clinical concept underlying the present diagnostic classifications of schizophrenia does not indicate whether one or several pathological processes are involved, or whether the causes of schizophrenia are the same in all cases. These gaps in the biological characterization of schizophrenia reflect its complexity, as well as the currently incomplete understanding of the basic neurobiology underpinning key psychological functions that are impaired in schizophrenia, such as memory, multimodal sensory integration, self-monitoring, and intentionality of behavior.

Throughout the chapter, the terms *developing countries* and *developed countries* are used to refer to a broad (and somewhat fluid) grouping of societies still characterized by traditional community structures, beliefs and values, and a relatively recent transition to industrialization and large-scale urbanization, and societies that have undergone the transition to

"modernity" much earlier, with all the consequences for community life, family support patterns, and systems of care for the sick and disabled. In the opinion of the author, recent alternative terms, such as *established market economies* or *emerging market economies,* fail to capture all of the nuances referred to previously and may suggest a closer relationship between the political economy and the burden of disease and disability than is warranted by the available evidence.

SCHIZOPHRENIA BEFORE AND AFTER KRAEPELIN

Since its early formulation in 1896 by Emil Kraepelin (1856 to 1926), the concept of *dementia praecox,* later renamed as schizophrenia by Eugen Bleuler (1857 to 1939), has been a hotly disputed diagnostic entity, raising, among other things, the question of its applicability to different populations and cultures. Although the Western medical literature before the 18th century contains unequivocal descriptions of depressive, manic, hysterical, and organic-dementing clinical pictures, it provides hardly any explicit reference to what later became known as schizophrenia. A plausible early description of schizophrenia can be attributed to Philippe Pinel (1745 to 1826), who referred in his *Nosographie Philosophique, ou la Méthode de l'Analyse Appliquée a la Médecine* (Vol III, Paris: Brosson, 1803), under the rubric of *démence ou succession d' idées incoherents,* to "frequent or continual appearance of unconnected ideas and shallow, dissociated emotions; disorganized movements and extravagant actions ... a kind of automatic existence." Similar case observations were reported by John Haslam (1764 to 1844), the apothecary to Bethlem Hospital in London, in his *Observations on Madness and Melancholy* (London: J. Callow, 1809). However, for most of the 19th century, such occasional descriptions of schizophrenic disorders were completely overshadowed by the all-embracing notions of "melancholia" and "mania." How was it possible that a condition which in the 20th century was found to have a prevalence of about 5 per 1,000, and accounted for more than 50 percent of all inmates in mental hospitals, could escape recognition? There are at least two possible explanations. One, less likely, is that until the second half of the 19th century schizophrenia indeed was a rare disease and that the significant increase of the share of schizophrenia in treated psychiatric morbidity did not occur until the last decades of the 19th century. Statistics on patients admitted to the University Psychiatric Clinic in Munich in 1908—some of them diagnosed personally by Kraepelin—indicate that as late as 10 years after the establishment of the diagnostic category of *dementia praecox,* the frequency of this diagnosis among all first admissions was as low as 9.1 percent in men and 7.3 percent in women. However, little was known at that time about the prevalence of untreated cases in the community. A second possible explanation is that schizophrenia did exist and was frequent, but for most of the 19th century, it was assimilated with other

syndromes, “mania” and “melancholia” in particular. There was no phenomenological description of schizophrenia as such because the concept did not exist, although all the symptoms and signs that were to be integrated later into the diagnostic category had been present all the time. Social factors must also have played a role. Diminishing tolerance in growing urban communities to bizarre behavior and deviance, and the commitment of increasing numbers of “insane” persons to the mental hospital where “scientific” diagnosis and treatment were believed to be practiced, may have facilitated the nosographic description of new clinical entities. According to statistics from Austria, the community census of “insane persons” rose slightly from 0.94 per to 1.35 per 1,000 population between 1873 and 1905, while the proportion of such persons institutionalized in hospitals during the same period increased from 32.5 percent to 61.9 percent. Such unprecedented concentration of psychotic persons in medical institutions undoubtedly made it easier for clinicians to identify syndromes, subtypes, and other varieties of mental disorders.

Comparative Psychiatry and Ethnopsychiatry

During the same period, there was a growing interest in the manifestations and frequency of mental disorders, primarily the psychoses, in non-Western cultures. British, Dutch, and French colonial administrations imported into their overseas dependencies the asylum model of care for the mentally ill and built mental hospitals for patients with intractable chronic illness, in particular, including “dangerous” psychotics. In 1903, Kraepelin travelled to Java and spent several weeks at the Buitenzorg (now Bogor) Hospital. He came to the conclusion that the basic forms of *dementia praecox* and *manic-depressive insanity* in the Javanese were essentially the same as those in Europe, though “racial characteristics, religion, and customs” modified their clinical manifestations. Kraepelin anticipated “rich rewards” from the capacity of cross-cultural studies to “throw light on the causes of mental disorder” and proposed *comparative psychiatry* as the systematic study of mental disorders and personality traits across cultures. However, for decades, the epidemiology of schizophrenia and other psychotic disorders in non-European cultures remained restricted to mental hospital statistics or rudimentary field surveys undertaken by colonial public health administrators. Typically, his conclusion was that “true” schizophrenia and depression were rare among indigenous peoples, and that chronically ill psychotic people were well tolerated in the community. The ideological framework in which such early *ethnopsychiatric* observations were embedded varied between benign paternalism and overt racism. In a 1953 monograph, John Carothers (1903 to 1989) claimed that the paucity of structured delusional contents and the lack of systematization of delusions in “the African” could be explained by a congenital underdevelopment of the frontal lobes of the brain. Similar

assumptions led other authors to conclude that depression was rare in sub-Saharan Africa or in Asia because of the lack of Judaeo-Christian cultural values that made the experience of guilt possible.

The sketchy and often distorted picture of the epidemiology of psychoses in the developing countries began to change in the postcolonial era when locally born psychiatrists, educated in the West, entered practice and teaching in their countries. Although trained in the colonial metropolis, they were keen to understand the nature of mental disorders in their own cultures and to introduce culturally appropriate alternatives to the colonial mental hospital, such as the model "psychiatric village" in Ibadan, Nigeria, founded by Thomas A. Lambo (1923 to 2004). In Asia, Pow Meng Yap (1921 to 1971) charted systematically the so-called *culture-bound syndromes,* highlighting their differentiation from schizophrenia. Epidemiological surveys, providing data on the incidence and prevalence of schizophrenia, were carried out in Mauritius, Nigeria, India, Taiwan, and mainland China. Since the 1960s, such studies expanded to include a variety of populations in Africa, Asia, and South America about which little had been known previously.

The comparative study of schizophrenia across populations and cultures gained in scope and momentum with the research program initiated by WHO in the 1960s. Two multicenter collaborative studies, the International Pilot Study of Schizophrenia (IPSS) and the Ten-Country Study (Determinants of Outcome of Severe Mental Disorders), generated cross-sectional and follow-up data on more than 3,000 cases of schizophrenia in 16 geographical areas in 12 countries in Africa, the Americas, Asia, and Europe. These studies introduced standardized assessment instruments, research procedures, and explicit diagnostic criteria in data collection at community and hospital facilities by local psychiatrists and mental health workers who had been trained to use such research tools and participated in their development. The WHO studies enabled for the first time direct comparisons of the population rates, psychopathology, and burden of disease associated with schizophrenia across various cultures.

DIAGNOSTIC COMPARABILITY OF THE SCHIZOPHRENIA PHENOTYPE ACROSS POPULATIONS

There is, as yet, no validated biomarker that could assist diagnosis. Cognitive testing and neuroimaging may help in ruling out alternatives, such as schizophrenia-like psychotic features secondary to brain or other medical disease, but clinical reasoning retains the last word. The paradigm-transformative DSM-III and subsequently ICD-10, DSM-IV, and DSM-5 are descendants of the Kraepelinian nosology that postulated close relationships between clinical symptoms, course and outcome, and ultimately, brain pathology. However, these diagnostic manuals are essentially classifications of diagnostic concepts and not of "natural kinds"

such as people or diseases. There is little evidence that most currently recognized mental disorders, including the psychoses, are separated by natural boundaries. There is a growing understanding that the syndromal spectrum of schizophrenia is the end-point for heterogeneous gene networks, pathophysiological pathways, and environmental modifiers. Increasing experimentation with research-oriented models and cross-disciplinary methodologies linked to the emerging RDoC may eventually reshape diagnostic classification, but this is a long-term project. Until its products reach clinical implementation, we shall have to live with DSM-5 and ICD-11, not forgetting Karl Jaspers' dictum that "a classification has only a provisional value. It is a fiction which will discharge its function of it proves to be the most apt for the time."

Despite the wide acceptance of the DSM and ICD criteria that promote reliable diagnostic identification, the question whether cases diagnosed as schizophrenia in different populations and cultures are phenotypically homologous is of critical importance, considering that the biological basis of the disorder is largely unknown and no objective diagnostic test is at present available. Phenotypic comparability implies that the constellation of symptoms is coherent and replicable, and that course, outcome, and response to treatment are broadly predictable.

In the past, the belief was widespread that schizophrenia was a "Western" disease and had no counterpart in indigenous populations untouched by modern technology and lifestyles. That this belief was mistaken is demonstrated by the fact that no human community has yet been found to be free of schizophrenia, diagnosable by commonly accepted criteria. Although no single symptom is *pathognomonic* of the disorder, the overall clinical presentation of schizophrenia is remarkably similar across cultures. Acutely ill patients in different cultural settings describe the same characteristic symptoms, such as hallucinatory voices commenting in third person on their thoughts and actions; thoughts being broadcast or taken away; or their surroundings being imbued with special meaning. Negative symptoms, such as psychomotor poverty, social withdrawal, and amotivation, are common, irrespective of the cultural setting. The conclusion that patients diagnosed with schizophrenia in different cultures suffer from the same disorder is further supported by the similar age- and sex-specific distribution of the onset of symptoms, which tends to peak in early adulthood. Considering the variety of social norms and beliefs about illness across cultures, the similar way in which the core symptoms of schizophrenia are experienced and described by people affected by the disorder is indeed striking. Nevertheless, variations exist that may affect the recognition and treatment of schizophrenia. Nigerian psychiatrists have described in patients with schizophrenia a characteristic symptom–complex consisting of anxiety, depression, vague hypochondriacal symptoms, bizarre magicomystical ideas, episodic twilight or confusional

states, atypical depersonalization, emotional liability, and retrospective falsification of memory based on hallucinations or delusions. Certain variants of the schizophrenia syndrome, such as an acute onset form and a catatonic subtype, appear to be more common in traditional rural communities. Research conducted within the WHO program of cross-cultural schizophrenia studies has confirmed the existence of population differences in the mode of presentation of the disorder. For instance, acute onset was found to characterize 40.3 percent, and marked catatonic symptoms 10.3 percent of cases assessed at research centers in India and Nigeria, compared to 10.9 percent and 1.2 percent at European and North American centers.

Differentiation from Other Disorders

A common clinical problem in developing countries is the differentiation of schizophrenia from psychoses due to infectious or parasitic diseases. African trypanosomiasis often results in a symptomatic psychosis that has a slow, insidious onset and may mimic schizophrenia. Since a variety of infectious, parasitic, and nutritional diseases are endemic in the developing world, it has been suggested that a large proportion of the cases of schizophrenia in those populations may in fact be symptomatic psychoses associated with tropical communicable diseases. Another brain pathology known to contribute to psychoses with schizophrenia-like features is temporal lobe epilepsy. The association between epilepsy and a chronic, interictal psychosis that is difficult to distinguish from schizophrenia has been well documented clinically. Studies in Europe and North America suggest a nearly 10-fold increase in the risk of schizophrenia-like psychosis among people suffering from epilepsy. No comparable data are available from developing countries, but considering that epilepsy is highly prevalent in many regions, it could be expected that undiagnosed and untreated epilepsy may account for a proportion of the cases of schizophrenia-like psychoses, especially in areas where epilepsy is endemic. It is unlikely, however, that either communicable diseases or organic brain disorders such as epilepsy account for more than a small fraction of the cases presenting as schizophrenia. Among some 5,000 individuals with psychotic illnesses who were screened in India and Nigeria for inclusion into the WHO Ten-Country Study, only 11.7 percent were excluded on grounds of having a physical disease that might explain their psychotic symptoms. Thus, problems of differential diagnosis in respect of organic brain diseases may arise in individual cases, but the majority of schizophrenic disorders in the Third World cannot be attributed to somatic disease or gross brain pathology. Considering all available evidence, schizophrenic disorders in non-Western populations can be reliably distinguished from acute transient psychoses, culture-bound syndromes, and probably from affective disorders, although the boundary with the latter has not been adequately

explored and symptomatic overlap may exist. Family morbidity data are still scarce but where such information is available, it suggests that genetic factors contribute to the transmission of schizophrenia in the same way as in the developed countries.

VARIATIONS IN THE INCIDENCE AND PREVALENCE OF SCHIZOPHRENIA

To date, no population or culture has been identified in which schizophrenia does not occur. The rates of its occurrence are broadly comparable, provided that the populations being compared are large enough to minimize the effects of small-area variation. This certainly does not imply that the incidence of the disorder is uniform across all populations, but rather that the range of variation is modest, in comparison with other complex or multifactorial diseases, such as non–insulin-dependent diabetes, ischemic heart disease, or multiple sclerosis, in which 10- to 30-fold differences among populations are not uncommon.

Prevalence

The prevalence of a disease is the estimated the number of cases per 1,000 persons at risk present in a population at a given time (point prevalence), or over a defined period (period prevalence). Since schizophrenia cases in prolonged remission are likely to be missed in point prevalence surveys, it will always be useful to estimate the lifetime prevalence by supplementing the assessment of the present mental state with data about past episodes of the disorder. An historical overview of selected prevalence studies of schizophrenia in different populations, spanning nearly seven decades of research, is presented in Table 12.3–1. These studies differ in many aspects of methodology but are generally characterized by a high intensity of case-finding. Several of the studies involve repeated surveys of the same population following intervals of 10 or more years. The majority have produced point prevalence estimates in the range 2.4 to 6.7 per 1,000 population at risk in developed countries and in the range 1.4 to 6.8 per 1,000 population at risk in developing countries. A systematic review of 188 studies in 46 countries, published between 1965 and 2002, estimated the median value for point prevalence at 4.6 per 1,000 persons and for lifetime prevalence at 7.2 per 1,000.

The recent 2010 Australian National Survey of psychosis covered a population of 1,464,923 people aged 18 to 64 years, that is, approximately 10 percent of all Australians in that age range. The study employed a two-phase design: a psychosis screen of all users of mental health services was followed by a random selection of 1,825 people who were interviewed using a standardized schedule. In addition, two cognitive tests were employed, and all participants provided blood samples for a screen of metabolic abnormalities. The study found the 1-month prevalence of schizophrenia to

be 3.10 cases per 1,000 population at risk and the estimated lifetime morbid risk at 3.45 per 1,000. The mean cognitive ability of the participants was 1.6 SD below the population average, but higher scores on the cognitive tests were significantly associated with better independent functioning. A striking finding was the high prevalence of the metabolic syndrome, affecting 60.8 percent of participants (65.9 percent were current smokers; 47.4 percent were obese; 32.4 percent were sedentary; 50.8 percent had a lifetime history of cannabis dependence, and 49.8 percent were chronic alcohol users). These alarming findings pointed to an urgent need for physical health interventions, cognitive remediation therapies, and workforce planning focusing on supported employment.

Incidence

The incidence rate is the estimated annual number of first-onset cases in a defined population per 1,000 persons at risk. Incidence is of greater interest for risk-factor epidemiology than prevalence since it is interpretable in terms of probability of disease occurrence in a given population and is temporally closer to the antecedents of the disease or its precipitating factors. However, its estimation depends critically on the capacity to determine reliably the point of the disorder—a requirement that is difficult to satisfy in schizophrenia due to its long prodromal period and the fuzzy boundary between premorbid state and onset of psychosis. Since objective biomarkers of the disease process are still lacking, onset is usually defined as the point in time when clinical manifestations become recognizable and can be diagnosed according to specified criteria. The first hospital admission, which has been used as a proxy for disease onset in many studies, is not a robust indicator because of the variable time lag that may exist in treatment facilities between the earliest appearance of symptoms and hospitalization (and many cases do not get hospitalized). A better approximation is provided by the time point of the person's *first contact* with any psychiatric, general medical, or alternative "helping" agency, for reasons likely related to emerging psychosis. Clearly, both first-admission and first-contact studies share the limitation of measuring "treated" incidence and risk-omitting cases that do not present for assessment or treatment. This limitation can be overcome by periodically repeated door-to-door surveys of the same population or by longitudinal cohort studies (though both are difficult to mount for reasons of cost and logistics).

Table 12.3–2 presents the essential features of 12 selected incidence studies of schizophrenia. Studies using a "broad" definition of schizophrenia (ICD-8, ICD-9, or unspecified) report approximately a threefold difference in the rates, for example, from 0.17 to 0.54 per 1,000 population per year for first admissions or first contacts. Studies using specified diagnostic criteria, such as the RDC, DSM-III, DSM-III-R, DSM-

IV, or ICD-10, produce considerably lower incidence rates. A systematic review of incidence data from some 160 studies from 33 countries, published between 1965 and 2001, yielded a median value of 0.15 and mean value of 0.24 per 1,000, with a fivefold range of the rates and a tendency for recent studies to report lower rates. A good deal of the variation highlighted by the systematic review can be attributed to methodological differences and confounders, including type of study sample and design, coverage of case finding (hospital-based, field surveys, case registers, birth cohorts), diagnostic practices, sample size, statistical power, and methods of data analysis. For example, birth cohort studies and case registers tend to produce higher rates than cross-sectional community surveys and hospital admission studies. Notwithstanding such limitations, real variation is undoubtedly present in schizophrenia (as in any human disease). The interesting question concerns the extent and sources of this variation, and whether its study adds any clues to the understanding of schizophrenia etiology.

Table 12.3–1.
Selected Prevalence Studies of Schizophrenia

Author	Country	Population	Method	Prevalence per 1,000 Population at Risk
Brugger (1931)	Germany	Area in Thuringia (*n* = 37,561); age 10+ years	Census; interview of sample	2.4
Strömgren (1938); Bøjholm and Strömgren (1989)	Denmark	Island population (*n* = 50,000)	Census interviews; repeat census	3.9 → 3.3
Böök (1953); Böök et al. (1978)	Sweden	Genetic isolate (*n* = 9,000); age 15–50 years	Census interviews; repeat census	9.5 → 17.0
Essen–Möller et al. (1956); Hagnell (1966)	Sweden	Community in southern Sweden	Census interviews; Repeat census	6.7 → 4.5
Rin and Lin (1962); Lin et al. (1989)	Taiwan	Population sample	Census interviews; repeat census	2.1 → 1.4
Crocetti et al. (1971)	Croatia	Sample of 9,201 households	Census based on hospital records and interviews	5.9
Dube and Kumar (1972)	India	4 areas in Agra (*n* = 29,468)	Census based on hospital and clinic records	2.6
Rotstein (1977)	Russia	Population sample (*n* = 35,590)	Census based on hospital and clinic records	3.8
Keith et al. (1991)	USA	Aggregated data across 5 ECA sites	Sample survey; interviews	7.0 (point) 15.0 (lifetime)
Jeffreys et al. (1997)	UK	London health district	Census; interview of	5.1

		(*n* = 112,127)	sample (*n* = 172)	
Kebede et al. (1999)	Ethiopia	25 districts of Addis Ababa (*n* = 2,228,490)	Screening by self-report questionnaire, interviews of sample (*n* = 2,042)	7.0 (point) 9.0 (lifetime)
Jablensky et al. (2000)	Australia	4 urban areas (*n* = 1,084,978)	Census, screen for psychosis; interviews of sample (*n* = 980)	3.1–5.9 (point)[a] 3.9–6.9 (period, 1 year)[b]
Waldo et al. (1999)	Micronesia	Island of Kosrae Genetic isolate	Screen of hospital records, interviews	6.8 (point)
Arajärvi et al. (2005)	Finland	Birth cohort (*n* = 14,817) Genetic isolate	Case register data; interviews of 55% of register cases	15.0 (lifetime) 19.0[c] (lifetime)
Wu et al. (2006)	USA (California)	Medicaid/Medicare health insurance data	20% random sample of insured subjects	5.1 (period, 1 year)
Perälä et al. (2007)	Finland	National sample (*n* = 8,028)	Screen for psychosis, interviews of sample; register and case note data also used	10.0 (lifetime) 22.9[d] (lifetime)

[a]All psychoses.

[b]Schizophrenia and other nonaffective psychotic disorders.

[c]Schizophrenia spectrum disorders.

[d]Nonaffective psychotic disorders.

→ Changes in prevalence found in repeat surveys of the same population.

Small Area Variations

Since schizophrenia is a low-incidence disorder, variations in its prevalence are more prominent in comparisons of rates obtained from small geographical areas and communities. In a study of an ethnically and socioeconomically homogeneous rural region in Ireland, the overall prevalence of schizophrenia at 3.9 per 1,000 was well within the modal range, but analysis by small district electoral divisions revealed significant variation in rates, ranging from 0.0 to 29.4 per 1,000. Such variation stands in contrast to the more similar rates typically found in studies of large urban areas or at national level. Local variation can be attributed to several factors, for example, spatial clustering of cases due to shared genetic vulnerability within extended pedigrees, differential mobility and mortality, and, possibly, to differential exposure to risk factors influencing intrauterine growth and early neurodevelopment. Such effects may give rise to “outlier” pockets of high or low incidence and prevalence that tend to cancel each other in larger population agglomerations.

Genetic Isolates

Isolated populations provide another source of incidence and prevalence variation. Isolates are characterized by a common descent from a small number of *founder* ancestors, a degree of endogamy or inbreeding, and a restricted admixture of immigrants, due to geographical or cultural seclusion over multiple generations, often ranging over hundreds of years. Such populations may vary considerably in size but are likely to be less heterogeneous with regard to genetic makeup and environmental exposures than the outbred (panmictic) populations constituting the world's majority, in which theoretically all individuals are potentially mating partners. The so-called young isolates comprise up to 20 to 30 generations and typically have arisen following drastic population size reductions (bottlenecks) due to wars, famine, epidemics, religious persecution, or other cataclysms. Subsequent population expansion results in a more uniform genetic background, including wider intervals of linkage disequilibrium, a more uniform environment and lifestyle, and significantly higher or lower prevalence of certain diseases, including psychiatric disorders. If coupled with availability of genealogical memory or records, such isolates present unique opportunities for genetic linkage and association studies of Mendelian (monogenic) diseases, and, hopefully, of complex traits such as schizophrenia, bipolar disorder, and other psychiatric syndromes.

Table 12.3–2.
Selected Incidence Studies of Schizophrenia

Author	Country	Population	Method	Rate per 1,000
Ödegaard (1946)	Norway	Total population	First admissions 1926–1935 (*n* = 14,231)	0.24 (hospital diagnoses)
Walsh (1969)	Ireland	City of Dublin (*n* = 720,000)	First admissions	0.57 (males, ICD-8) 0.46 (females, ICD-8)
Murphy and Raman (1971)	Mauritius	Total population (*n* = 257,000)	First admissions	0.24 (Africans) 0.14 (Indian Hindus) 0.09 (Indian Moslems)
Lieberman (1974)	Russia	Moscow district (*n* = 248,000)	Follow-back of prevalent cases	0.20 (males) 0.19 (females)
Helgason (1977)	Iceland	Total population	First admissions (case register)	0.27 (ICD-8)
Lin et al. (1989)	Taiwan	3 communities (*n* = 39,024)	Door-to-door survey	0.17 ("Bleulerian" criteria)
Castle et al. (1991)	UK	London (Camberwell)	First admissions (case register)	0.25 (ICD-9) 0.17 (RDC)

				0.08 (DSM-III)
Rajkumar et al. (1993)	India	Area in Madras (*n* = 43,097)	Door-to-door survey and key informants	0.41 (ICD-9)
Wig et al. (1993)	India	A rural area (*n* = 103,686) and an urban area (*n* = 348,609) in northern India	Case-to case finding and key informants	0.38 (urban, ICD-9) 0.09 (urban, Catego S+) 0.44 (rural, ICD-9) 0.12 (rural, Catego S+)
Brewin et al. (1997)	UK	Nottingham	2 cohorts of first contacts (1978–1980 and 1992–1994)	0.25 → 0.29 (all psychoses, ICD-10) 0.14 → 0.09 (ICD-10 schizophrenia)
Mahy et al. (1999)	Barbados	Total population (*n* = 262,000)	First contacts; PSE interviews; Catego	0.32 (ICD-9) 0.28 (Catego S+)
Bresnahan et al. (2000)	USA (California)	Birth cohort (*n* = 12,094)	Case register study; cumulative risk by age 38 years	0.93 (males, DSM-IV) 0.35 (females, DSM-IV)

A number of isolated populations in different parts of the world, including Finland, Iceland, and northern Sweden; the Pima Indians; the Bedouins; the people of the Central Valley of Costa Rica, several areas in Quebec, as well as religious communities, such as the Old Order Amish, the Hutterites, and the Mennonites, have been studied by epidemiologists and geneticists with a view to identifying large pedigrees, informative for diseases ranging from asthma and diabetes to schizophrenia and bipolar disorder. High rates of psychoses (two to three times the national or regional rate) have been reported for population isolates in northern Sweden and several areas in Finland. Although the whole population of Finland shares some features of an old isolate (approximately 2,000 years), the northern and eastern regions of Finland have been settled relatively recently (in the 16th and 17th centuries), and one particular subregion with a current population of 18,000 has been founded by 40 families at the end of the 17th century, that is, only 12 generations back. Genetic–epidemiological studies in this isolate estimate the lifetime risk of schizophrenia at 2.2 percent, compared to 1.2 percent for the whole of Finland. A case register–based study of a birth cohort (14,817 individuals) from this region established a lifetime prevalence of 1.5 percent for schizophrenia spectrum psychotic disorders.

Daghestan in the Northern Caucasus (Russian Federation) is a region that has been inhabited for more than 3000 years by some 26 small ethnic groups constituting together at least five genetically distinct populations, varying considerably in their morbidity patterns. Possibly, the highest ever

lifetime prevalence estimate for schizophrenia (4.9 percent) has been reported from one such highland subisolate (3,000 members).

The population of the Palau islands (Micronesia), currently 20,470 people, has been geographically and ethnically isolated from other Pacific populations for nearly 2,000 years. A genetic epidemiological study of treated cases estimated the lifetime prevalence of schizophrenia at 2.77 percent in males and 1.99 percent in females, that is, high in excess of the modal risk of about 1 percent usually reported for large outbred populations. All of the 160 Palau cases were concentrated in 59 families, each traceable to a single common founder, with 11 of them having 5 to 11 affected members each.

At the other extreme, the lowest known prevalence rate of schizophrenia in any population (and a very low rate of bipolar disorder) has been found among the Hutterites in South Dakota, a Protestant sect of European descent (founded in the 16th century) whose members live since the 1870s in closely knit, endogamous rural communities in Manitoba (Canada) and South Dakota (United States). According to preserved pedigree records, all of the present 45,000 Hutterites are descendants of fewer than 90 ancestors who lived in the 18th and early 19th centuries. Reduced genetic heterogeneity and communal lifestyle with minimum variation in environmental exposures make this population an attractive laboratory for a variety of disease studies, including psychiatric disorders. An early epidemiological study, in which the entire population of several Hutterite communities was screened, resulted in a schizophrenia lifetime prevalence of only 1.1 per 1,000. Subsequent reanalysis of the data using DSM-III-R criteria, and a repeat survey, replicated the original finding. Both genetic (low frequency of psychosis-predisposing alleles) and lifestyle factors (protective community support) have been proposed as an explanation for the unusually low rate of psychosis. Negative selection for individuals with schizoid traits who fail to adjust to the communal lifestyle and eventually outmigrate without leaving progeny has also been suggested but not definitively proven.

Table 12.3–3.
Incidence Rates per 1,000 Population, Age 15–54 Years, for a "Broad" and a "Narrow" Case Definition of Schizophrenia (WHO Ten-Country Study, 1992)

		"Broad" Definition (ICD-9)			"Narrow" Definition (CATEGO S+)		
Country	**Area**	**Male**	**Female**	**All**	**Male**	**Female**	**All**
Denmark	Aarhus	0.18	0.13	0.16	0.09	0.05	0.07
India	Chandigarh (rural area)	0.37	0.48	0.42	0.13	0.09	0.11
	Chandigarh (urban area)	0.34	0.35	0.35	0.08	0.11	0.09
Ireland	Dublin	0.23	0.21	0.22	0.10	0.08	0.09
Japan	Nagasaki	0.23	0.18	0.20	0.11	0.09	0.10

Russia	Moscow	0.25	0.31	0.28	0.03	0.03	0.02
UK	Nottingham	0.28	0.15	0.22	0.17	0.12	0.14
USA	Honolulu	0.18	0.14	0.16	0.10	0.08	0.09

Migrant Populations

A special case is the exceptionally high incidence rate of schizophrenia (about 6.0 per 1,000) that has been found in the African–Caribbean population in the United Kingdom. The excess morbidity is not restricted to recent immigrants and is actually higher in the British-born second generation of migrants. Similar findings of nearly fourfold excess over the general population rate have been reported for the Dutch Antillean and Surinamese immigrants in the Netherlands. The causes of the phenomenon remain obscure. Incidence studies in the Caribbean do not indicate any excess schizophrenia morbidity in the indigenous populations from which migrants are recruited. Explanations in terms of biological risk factors have found little support. A finding in need of replication is the significantly higher incidence of schizophrenia among the siblings of second-generation African–Caribbean patients with schizophrenia than the incidence of the disorder in the siblings of white patients. Such "horizontal" increase in the morbid risk could result from an environmental factor boosting the penetrance of predisposing genes carried by a proportion of pedigrees within the African–Caribbean population. Psychosocial hypotheses involving acculturation stress, demoralization due to racial discrimination, and blocked opportunities for upward social mobility are being explored but have not yet been definitively tested.

The WHO Ten-Country Study

Inferring a more general perspective on the comparative incidence or prevalence of schizophrenia in different populations from individual studies conducted at different times and with different methods, or from pooled data from multiple studies, is beset with difficulties. The WHO Ten-Country Study remains to date the only investigation that has applied a uniform design and common research tools to generate directly comparable incidence data for different populations. Incidence counts in the study were based on first-in-lifetime contacts with any "helping agency" within defined areas (including traditional healers in the developing countries), which were monitored over a 2-year period. Potential cases and key informants were interviewed by clinicians using standardized instruments, and the timing of onset was reliably ascertained for the majority of the patients. In 86 percent of the 1,022 patients, the onset of diagnostic symptoms of schizophrenia was within the year preceding their first contact, and, therefore, the first-contact incidence rate was adopted as a reasonable approximation to the "true" onset rate. Two definitions of

caseness, differing in the degree of sensitivity and specificity, were used: a "broad" clinical definition comprising ICD-9 schizophrenia and paranoid psychoses, and a more restrictive definition of "nuclear" schizophrenia with Schneiderian *first-rank symptoms* of "nuclear" schizophrenia. Intersite reliability of case classification was ensured by applying a computerized diagnostic algorithm. The rates for eight of the catchment areas in which a nearly complete case finding was achieved are shown in Table 12.3–3. The differences between the area rates for "broadly" defined schizophrenia (0.16 to 0.42 per 1,000) were significant (P <. 001) but those for "nuclear" schizophrenia were not, suggesting that the frequency of the latter subgroup varies less across different populations. No differences were found between the "broad" and "nuclear" cases with regard to age at onset, or 2-year course and outcome. Therefore, it is unlikely that "nuclear" and "broad" schizophrenia identify different clinical disorders. Several replications of the WHO design, including its research procedures and instruments, have been carried out with very similar findings in India, the Caribbean, and the United Kingdom.

VARIATION IN THE COURSE AND OUTCOME OF SCHIZOPHRENIA

Systematic investigations into the course and outcome of schizophrenia were initiated by Kraepelin, who believed that the natural history of the disorder could provide a provisional validation of the disease concept, until final verification by brain pathology and understanding of etiology. However, contrary to Kraepelin's postulate of "mental weakness" as the ultimate outcome of *dementia praecox,* it is precisely the long-term course of schizophrenia that reveals the greatest extent of variation across populations and cultures. The course and outcome of schizophrenia exhibit striking heterogeneity and variability, both across and within populations. Patients with similar clinical and diagnostic characteristics at baseline assessment develop a broad spectrum of outcomes ranging from stable clinical and social recovery after a single psychotic episode to chronic unremitting psychosis and severe impairment. Long-term follow-up studies lend credibility to the conclusion that a high proportion (more than 30 percent) of patients meeting the diagnostic criteria for schizophrenia may achieve relatively favorable outcomes.

Untreated Schizophrenia

Since the great majority of schizophrenic patients are today receiving pharmacological treatment, one could assume that current research does not reflect the "natural" course and outcome of untreated schizophrenia. However, several studies, in an urban area in Scotland, and in remote rural areas in India and China, have succeeded in detecting and assessing *never hospitalized* patients with schizophrenia. About half of the Scottish patients

had been prescribed neuroleptics by their general practitioners, while the Indian and Chinese patients were virtually untreated. In all three settings, the outcomes of these samples were heterogeneous but, except for a larger proportion of Chinese patients having marked psychotic symptoms, they did not differ much from the outcomes in the treated groups. In a historical study of 70 Swedish patients with first admissions in 1925, lifetime records were retrieved and rediagnosed in accordance with DSM-III. None of these patients had received neuroleptics. The final outcome was rated as "good" in 33 percent, as "profoundly deteriorated" in 43 percent, and as "intermediate" in 24 percent.

Secular Trends in the Outcome of Schizophrenia

A long-term perspective on the course of schizophrenia over successive generations is provided by a meta-analysis of 320 outcome studies on schizophrenia or *dementia praecox* published between 1895 and 1992, which comprise a total of 51,800 subjects. Overall, about 40 percent of the patients have been described as "improved" after an average length of follow-up 5.6 years. There was a significant increase in the rate of improvement during 1956 to 1985 compared to 1895 to 1955, clearly related to the introduction of neuroleptic treatment, but a secular trend toward better outcomes with every successive decade had been present for much longer. Coupled with the virtual disappearance of the most malignant or "catastrophic" forms of schizophrenia that tended to result in a profound "defect state" after a single psychotic episode, or in death (*lethal catatonia*), these observations suggest that a transition to a less deteriorating course of schizophrenia had been taking place prior to the advent of modern pharmacological treatment. Among the factors that might explain this shift are improvements in general care, progressive changes in attitudes and hospital regime that occurred in a number of institutions on both sides of the Atlantic in the 1930s and 1940s, and the expectation that psychosocial measures such as psychotherapy or rehabilitation could result in a cure or remission of the disorder.

Concurrently with the above trends in Europe and North America, reports from developing countries, usually based on small clinical samples, pointed to a less disabling course and a high rate of recovery from schizophrenic psychoses in traditional societies, such as Mauritius and Sri Lanka. Good outcome cases included patients who would be expected to have poor outcome, if "Western" prognostic criteria had been applied. However, selection bias could not be excluded, as these studies were based on hospital admissions only. In addition, clinical improvement could have been confounded with the social adjustment many patients achieve in a comparatively undemanding environment. Thus, room was left for doubts concerning the validity of the early observations of a better prognosis of schizophrenia in non-Western settings.

Many of the methodological issues attending such research were later addressed by the WHO multicenter studies that employed standardized assessment, recruitment of homogeneous, incidence-based samples, and more refined, dimensional measures of course and outcome. These studies included a total of 2,736 patients in 16 countries, diagnosed with schizophrenia according to strict criteria.

Course and Outcome in the WHO Studies

The earliest WHO investigation, the IPSS, found that higher proportions of patients in India, Colombia, and Nigeria had better outcomes on most dimensions than patients in the developed countries at 2- and 5-year follow-up assessment. However, the IPSS was not free of bias, since patients were recruited from hospitals, where admission policies could have led to overinclusion of chronic cases in the developed countries, and of recent onset, acute cases in the developing countries.

Such confounding factors were largely eliminated in the subsequent WHO Ten-Country Study, in which epidemiological cohorts of first-episode cases were uniformly assessed upon their first contact with community or hospital services. Diagnostic misclassification due to potential admixture of transient psychotic episodes of good prognosis, or of psychoses due to acute physical illness, was excluded, as such cases were carefully screened out. The follow-up of these incidence cohorts provided ample confirmation of the finding that the outcome of schizophrenia was generally better in developing than in developed countries (Table 12.3–4). The better course and outcome in the developing country areas could not be attributed to any particular clinical subtype of the disorder. The main outcome difference across the study areas was in the average length of symptom-free remissions. Analysis of the data revealed that the better overall pattern of course and the less disabling outcome in the study areas in developing countries were mainly attributable to a significantly greater percentage of patients remaining in a remission of symptoms over longer periods after an acute psychotic episode, rather than to milder or shorter psychotic episodes. Statistically, this pattern was significantly predicted by the setting (developing versus developed country); acute onset; being married or cohabiting with a partner; and having access to a supportive network (family or friends). The length of remission was unrelated to antipsychotic treatment, which generally was administered for much shorter periods of time to patients in the developing countries.

A subsequent long-term follow-up investigation, the International Study of Schizophrenia(ISoS), involving 18 research centers in 14 countries, traced 75 percent of cases assessed in the earlier WHO studies referred to previously, as well as additional cohorts from China, Hong Kong, and India. Follow-up data were collected on a total of 1,633 cases (surviving or dead), and 890 patients were reinterviewed at either 15- or 25-year follow-

up since their first assessment. Key findings from this study are presented in Table 12.3–5.

Several general conclusions can be drawn from the WHO studies. The frequencies of both relapses and remissions tend to increase over time. However, regardless of the increasing relapse rate, the cumulative proportion of follow-up time during which patients have psychotic symptoms (as a percentage of the total follow-up time) tends to remain stable or even decrease. At the end of the 5-year follow-up, 57 percent of the patients had experienced a total of less than 9 months of active psychosis; only 22 percent had been psychotic for 45 to 60 months. At 15-year and 25-year follow-up, 43 percent and 41 percent, respectively, have been free of active psychotic symptoms for the past 2 years. The levels of social impairment assessed at 2 years changed little during the subsequent years of follow-up. Overall, most of the observed change in the clinical state and social functioning of patients between the 2-year follow-up and the 5-year follow-up were toward improvement rather than deterioration. This was congruent with the findings at 15-year and 25-year follow-up. Independently of the WHO studies, similar findings of better outcomes of schizophrenia in developing countries have been reported by other investigators.

Table 12.3–4.
Two-Year Course and Outcome Features of 1,070 Patients with Schizophrenia in the WHO Ten-Country Study

Course and Outcome Descriptor	Percent Patients in Developing Countriesa (gn = 467)	Percent Patients in Developed Countries[b] (n = 603)
Remitting, complete remissions	62.7	36.8
Continuous or episodic, no complete remission	35.7	60.9
Psychotic <5% of the follow-up	18.4	18.7
Psychotic >75% of the follow-up	15.1	20.2
No complete remission during follow-up	24.1	57.2
Complete remission for >75% of the follow-up	38.3	22.3
On antipsychotic medication >75% of the follow-up	15.9	60.8
No antipsychotic medication during follow-up	5.9	2.5
Hospitalized for >75% of follow-up	0.3	2.3
Never hospitalized during follow-up	55.5	8.1
Impaired social functioning throughout follow-up	15.7	41.6
Unimpaired social functioning	42.9	31.6

>75% of follow-up

[a]Study areas in Colombia, India, and Nigeria.

[b]Study areas in the Czech Republic, Denmark, Ireland, Japan, Russia, the United Kingdom, and the United States.

Factors Influencing the Variations in Outcome

The differences in outcome between patients in developing countries and patients in developed countries remain insufficiently understood but are unlikely to be reducible to a single variable or to a small number of explanatory variables. The variation in course and outcome may result from additive or interactive effects of genetic differences between populations; from lower incidence in traditional societies of the type of chronic stress to which people with schizophrenia are particularly vulnerable; or from higher odds of an individual–environment fit in traditional communities that minimize social isolation. While genetic differences undoubtedly exist (e.g., in the allelic structure of genes coding for key neurotransmitter receptors), nothing more specific can at present be said as to their role in the course and outcome of schizophrenia. On the contrary, a strong modifying influence of the psychosocial environment on the course of schizophrenia is plausible, considering the differences between developing and developed countries with regard to social support systems, kinship networks, and beliefs about mental disease. As regards psychosocial stress, the WHO studies found that the index of *expressed emotion* (EE), a short-range predictor of psychotic relapse, was as effective in Indian families as in European and North American families, but that high-EE families were much rarer in India than in Denmark or the United Kingdom. If this finding could be replicated in other settings in developing countries, the relative rarity of at least one type of pathogenic stress in the daily environment of schizophrenic individuals in such communities could be demonstrated. However, EE may not be the only type of stress to which schizophrenic individuals respond with psychotic exacerbation. Ethnopsychiatrist Henry Murphy (1915 to 1987) proposed four criteria for schizophrenia-evoking stress: (i) a situation demanding action or decision; (ii) complexity or ambiguity of the information supplied to deal with the task; (iii) unless resolved, the situation demanding action or decision persists; (iv) the person has no “escape route” available. Each component of this model may occur at different frequencies in traditional and in industrialized societies, a proposition that should be testable epidemiologically or experimentally. For example, important differences exist between traditional cultures and Western societies regarding the perception of the sick role and the beliefs and practices related to mental illness. Shared magical–mystic explanations of mental illness and traditional healing practices, still prevalent in traditional rural

communities, certainly provide no cure for schizophrenia, but they may play a role in lowering the barriers to recovery and reintegration in the community. The findings of a better outcome for schizophrenia in traditional societies are compelling enough to set a research agenda that may lead to important insights concerning the management and treatment of schizophrenia in both developing and developed countries.

Table 12.3–5.
Long-Term (15- and 25-Year) Outcome in Patient Cohorts Assessed in the International Study of Schizophrenia (ISoS; 2007)

Outcome Variable	Incidence Cohorts (N = 1,171, Including 15-Year Follow-up of the WHO 10-Country Cohort)	Prevalence Cohorts (N = 462, Including 25-Year Follow-up of 373 Cases from the WHO IPSS)
Recovered at follow-up (Bleuler's criteria)	48.1	53.5
Not psychotic in the past 2 years	42.8	40.8
GAF-S >60	54.0	56.7
Working most of past 2 years	56.8	73.9
GAF-D >60	50.7	60.3
SMR (range)	0.00–5.71[a]	1.04–8.88[b]

GAF-S, Global Assessment of Functioning—Symptoms scale; GAF-D, Global Assessment of Functioning—Disability scale; SMR, standard mortality ratio.

[a]Rochester (0.00), Moscow (1.41), Chandigarh urban (1.88), Prague (2.53), Chandigarh rural (3.02), Honolulu (3.13), Nottingham (3.31), Dublin (4.10), and Nagasaki (5.71).

[b]Sofia (1.04), Cali (1.31), Madras (1.90), Agra (1.86), Beijing (2.97), Prague IPSS (3.84), Mannheim (5.55), Hong Kong (5.76), and Groningen (8.88).

To sum up, the most striking finding from long-term follow-up studies is the high proportion of patients who recover, completely or with mild residual abnormalities, after decades of severe illness. This contrasts with the ingrained image of schizophrenia as an intractable, deteriorating illness that many clinicians tend to adopt on the basis of a limited follow-up horizon and patient samples selected for unfavorable course and treatment response. It is unlikely that the high percentage of recoveries in the long-term studies could be explained by affective illness or brief transient psychoses being misdiagnosed as schizophrenia. Similarly unlikely is the attribution of all the good outcomes to antipsychotic treatment, since comparable proportions of improvement of recovery had been reported for patients who never received neuroleptics. A tentative conclusion is that schizophrenia is not an invariably chronic deteriorating disorder. It should

be noted, however, that some studies suggest a recent trend of worsening clinical and social outcomes in patients with schizophrenia in both developed countries and developing countries. People with schizophrenia are profoundly sensitive to their immediate social environment. Increasing social and economic stresses experienced by both rural and urban communities in many developing countries in the last two decades may be eroding the traditional support systems, thus resulting in worse outcomes.

THE BURDEN OF COMORBIDITY

Physical Disease

There is significant medical comorbidity in schizophrenia, comprising mainly common diseases that affect patients with schizophrenia more frequently than attributable to chance, as well as certain rare conditions or abnormalities that tend to co-occur with the disorder. Physical disease in patients with schizophrenia tends to be seriously undetected and underdiagnosed. Between 46 and 80 percent of inpatients with schizophrenia, and between 20 and 43 percent of outpatients, have been found in different surveys to have concurrent medical illnesses. Persons with schizophrenia, and especially those who are homeless or injection drug users, are at increased risk for potentially life-threatening communicable diseases, such as HIV/AIDS, hepatitis C, and tuberculosis.

However, it is the chronic noncommunicable diseases that in the past decades are taking an increasingly heavy toll on people with schizophrenia. Cardiovascular disease (CVD) and the associated premature mortality are significantly increased in schizophrenia. A recent comprehensive meta-analysis of mortality in schizophrenia reported a standardized mortality ratio (SMR) of 2.54. It has been estimated that the high prevalence of CVD impacts on the life expectancy of people with schizophrenia by an average reduction of 18.7 years for men and 16.3 for women, and that the gap between those affected with the disorder and the general population is increasing. Although suicide rates in schizophrenia are high and equal those for major depression, recent studies suggest that CVD is at present the main cause of premature death, already evident in early or middle adulthood. The contribution of “atypical” antipsychotic medication to weight gain and hence to metabolic imbalance and type II diabetes is likely, but its effects are collinear with other, modifiable lifestyle risk factors, including nutrition, smoking, and lack of physical activity. Obesity and the concomitant metabolic syndrome involving insulin resistance and type II diabetes are at present common problems in patients with schizophrenia.

Certain rare genetic or idiopathic disorders, such as metachromatic leucodystrophy, acute intermittent porphyria and coeliac disease, and dysmorphic features such as high-steepled palate, malformed ears, and other minor physical anomalies, also tend to co-occur with schizophrenia.

Substance Abuse

Substance abuse is at present by far the most common associated health problem among patients with schizophrenia and may involve any drug of abuse or a polydrug combination. The addictive use of cannabis, stimulants, and nicotine is disproportionately high among patients with schizophrenia and may be related to the underlying neurobiology of the disorder. In a population sample of patients with psychotic disorders in Australia, a diagnosis of comorbid drug abuse or dependence was made in 36.3 percent of males and 15.7 percent of females with schizophrenia (compared to 3.1 percent and 1.3 percent in the general population). In addition to poor prognosis of schizophrenia in patients with heavy cannabis use, a systematic review of published data on cannabis exposure and the onset of schizophrenia concluded that early use increased the risk of psychosis in a dose-related manner, especially in persons at high genetic risk of schizophrenia. A much debated issue is the role of illicit substance use in precipitating the onset of schizophrenia. While cannabis, amphetamines, cocaine, and polydrug abuse are at present almost ubiquitous among patients with schizophrenia worldwide, their potential of inducing irreversible functional alterations in brain neurotransmitter receptors remains uncertain. A recent study examined information on substance use during the year prior to illness onset in 1,642 participants in the Australian National 2010 Survey of Psychosis with a diagnosis of schizophrenia or affective psychosis. The findings confirmed previous reports and indicated that, controlling for covariates, history of polysubstance use and predominant cannabis abuse was highly significantly associated with an earlier age at onset of schizophrenia spectrum disorders but not with age at onset of affective psychoses.

VARIATIONS IN MORTALITY

In both developing and developed countries, schizophrenia is associated with excess mortality, well documented by epidemiological studies on large cohorts over extended periods. Population-based data from Norway, 1926 to 1941 and 1950 to 1974, indicate that, while the total mortality of psychiatric patients had been decreasing, the relative mortality of patients with schizophrenia remained unchanged at a level higher than twice that of the general population. Similar findings have been reported from other European countries and North America, with SMR of 2.6 or higher for patients with schizophrenia, which corresponds to about 20 percent reduction in life expectancy. In Taiwan, data collected over a 15-year period indicate that of all mental disorders followed up, schizophrenia was associated with the highest mortality, representing an 80 percent increase over the mortality of the general population. A meta-analysis of 18 studies estimated a crude mortality rate of 189 deaths per 10,000 population per

year and a 10-year survival rate of 81 percent. Mortality was significantly higher in males than in females, and the difference was primarily due to a male excess in suicides and accidents.

Data from successive patient cohorts in Denmark, the United Kingdom, and Australia suggest a trend of increasing mortality in the early stages of the disorder. In Denmark, the 5-year cumulated SMR increased from 5.30 (males) and 2.27 (females) for the period 1971–1973 to 7.79 (males) and 4.52 (females) for the period 1980–1982 (56). Particularly striking was the SMR of 16.4 for males in the first year after a diagnosis of schizophrenia.

Whereas in the past the excess mortality among individuals with schizophrenia was due mainly to communicable diseases such as tuberculosis, the leading causes of premature death among patients with schizophrenia at present are CVD, suicide, accidents, and common physical diseases. Suicide accounts for 28 percent of the excess mortality in schizophrenia, with aggregated SMR of 9.6 for males and 6.8 for females. In developed countries, the suicide risk associated with schizophrenia is now nearly as high as that associated with major depression. In China, the *relative risk* (RR) of suicide in individuals with schizophrenia, compared to those without the disorder, has been estimated at 23.8. Suicide in schizophrenia may occur at any stage of disorder, but the risk is particularly high in the first 6 months after the first psychotic episode, as well as following periods of frequent hospital admission and discharge. Several risk factors, relatively specific to schizophrenia, have been suggested: being young and male, experiencing multiple relapses and remissions, comorbid substance abuse, awareness of the deteriorating course of the condition, and a sense of hopelessness with loss of faith in treatment. In the Australian study, suicide risk was highest in the first 7 days after discharge from inpatient care. The trend seems to parallel the significant reduction in the number of psychiatric beds. Whether the increased suicide mortality is associated with the shift in the management of schizophrenia from hospital to community care remains to be established.

Unnatural causes of death apart, in both developed and developing countries, schizophrenia is associated with excess mortality from common diseases, caused or exacerbated by poor self-care, inadequate nutrition, heavy smoking, and medical neglect. At least part of this excess mortality is preventable.

CRIMINAL OFFENDING AND VICTIMIZATION

Much remains to be understood about the risk of offending by individuals with schizophrenia and related psychotic disorders. Inconsistent research findings have been attributed to the use of varied and sometimes flawed methodologies and biased samples, such as inpatient and prison populations. The risk of violent offending has been estimated to be much

higher in people with severe mental illness than that in the general population, partly attributable to the presence of substance abuse. In a large-scale, whole-of-population Western Australian study of a cohort of all people born (1955 to 1969), arrest history had been documented for the period 1985 to 1996 and cross-linked with the psychiatric case register. The prevalence of mental illness among offenders was 11.1 percent; 6.5 percent of offenders had substance use disorders, but only 1.7 percent had schizophrenia. Notably, for the majority of offenders with schizophrenia, first arrest preceded first contact with mental health services. Given a large proportion of schizophrenia offenders being arrested prior to their first contact with psychiatric services, there are important implications for mental health and criminal justice policy and practice. Very similar findings were reported from a meta-analysis of nine studies, indicating that a substantial proportion of first-episode psychosis individuals commit offences before presenting for treatment. Another aspect of the problem is victimization—violent assaults on people with severe mental illness. Among adults with psychotic disorders, especially women, the annual prevalence of any victimization was 38.6 percent, and of violent victimization 16.4 percent—nearly five times higher than the general population rate.

FOCUS ON HIGH CLINICAL RISK AND FIRST-EPISODE PSYCHOSIS

During the past two decades, the study of high clinical risk states and first-episode psychosis has become a rapidly expanding area of research, often coupled with intervention programs and dedicated services. While the concept of schizophrenia prodrome and the transition to psychosis is not new, the prospect of delaying, attenuating, or even preventing the onset of the disorder has a strong appeal to researchers and, in many instances, to funding agencies. Several comprehensive reviews and meta-analyses take stock of what has been learned and which questions remain unanswered in this burgeoning field. Among the issues highlighted in this literature are problems of heterogeneity in the definitions of “clinical high-risk” or “ultra-high-risk” (UHR), which tend to overemphasize positive psychotic symptoms at the expense of negative symptoms, cognition, and social functioning; the distinction between trait and state factors; small study samples; different entry criteria for clinical trials; and selection bias in reporting outcomes. The estimates of risk of transition to psychosis in individuals presenting with clinical high risk are consistent, indicating 18 percent developing overt psychotic symptoms at 6 months of follow-up, 22 percent at 1 year, and 36 percent at 3-year follow-up. The high proportion of false-positive prognostication using UHR criteria is indeed disquieting, but the field has demonstrated potential benefits in terms of better, integrated treatment and care for young people in prepsychotic distress.

SOCIAL AND ECONOMIC COSTS

The costs of schizophrenia are high. A recent Australian study estimated that in 2010, costs of hospitalization accounted for 42 percent of the total health sector costs, and that productivity losses due to mental morbidity were assessed at 53 percent of the total costs. On a wider horizon, the GBD analysis found that psychiatric disorders, including schizophrenia, accounted for 22.9 percent of all years lived with disability.

Notwithstanding the variations in course and outcome discussed previously, schizophrenia is associated with a greater burden of chronic disability than any other mental disorder. Both the positive and negative symptoms of the disease interfere seriously with a person's capacity to cope with the demands of daily living. Patients with schizophrenia experience particular difficulty in dealing with complex demands and environments, especially those that involve social interaction and decoding of social communication. Moreover, the onset is usually at a developmental stage of incomplete social maturation, educational attainment, and acquisition of occupational skills. The intrusion of psychosis at this stage results in a severely truncated repertoire of social skills and lifelong socioeconomic disadvantage. These factors are exacerbated by the societal reaction to individuals manifesting behaviors associated with "insanity," which generally involves stigma and social exclusion.

These adverse factors interact to cause the *social breakdown syndrome,* a term introduced by Ernest Gruenberg (1915 to 1991) to designate the cluster of secondary impairments in schizophrenia resulting in a loss of social support networks and a greatly diminished quality of life for a substantial proportion of patients. Many patients with schizophrenic end up on the streets or in the criminal justice system and are exposed to abuse, even in psychiatric hospitals. Such outcomes are not uncommon in either developed or developing countries, although in the latter settings traditional family and community structures are still capable of providing a protective environment, and probably fewer patients are marginalized by society.

The social and economic costs of schizophrenia are disproportionately high relative to its incidence and prevalence. As pointed out previously, estimates by WHO and the World Bank indicate that schizophrenia accounted for 2.6 percent of the burden of disease (DALYs) in 1990 in the developed countries and for 0.8 percent in the developing countries. In terms of DALYs, the demographic trends to the year 2020 predict more than a 50 percent increase in the disease burden attributable to schizophrenia in developing countries, a burden approaching that of malaria and nutritional deficiency.

The total material cost of schizophrenia is disproportionately high relative to the population point prevalence of the disease, or the lifetime morbid risk. In developed countries, the direct costs of schizophrenia, incurred by hospital or community-based treatment, supervised

accommodation, and related services, amount to 1.4 to 2.8 percent of the national health care expenditure and to up to one-fifth of the direct costs of all mental disorders. Although estimates of the indirect costs of schizophrenia vary greatly depending on the method of analysis and the underlying assumptions, these costs are likely to be comparable in scale to the direct costs, considering lost productivity and employment, the economically devastating long-term impact of the illness on the patient's family, other caregivers' opportunity costs, the increased mortality of people with schizophrenia, the costs to the criminal justice system, and other issues related to public concerns about safety. Estimates based on the Epidemiological Catchment Areas study in the United States put the direct costs of schizophrenia in 1990 at $17.3 billion and the indirect costs at $15.2 billion per annum.

An important aspect of the economics of schizophrenia is the so-called *funding imbalance effect:* studies indicate that 97 percent of the total lifetime costs of schizophrenia are incurred by fewer than 50 percent of the patients diagnosed with the disorder. While this finding points to a hardcore subset of cases with severe chronic illness, multiple disabilities, and excessive dependence on services and other support, it also suggests that in more than 50 percent of those with schizophrenia, the disorder is less disabling or treatment more effective.

Most of the economic evidence on schizophrenia comes from studies conducted in Western societies. However, mental health economics is a young discipline. Evidence on the costs of schizophrenia is at present limited even in developed countries, and such data are quite scarce for the majority of developing countries, although individual studies provide some insight into the likely economic impact of the disease. Since both the direct and indirect costs of schizophrenia are context-bound, extrapolations not only from the developed to the developing world but also across countries at comparable levels of *gross domestic product* (GDP) per capita must be made with caution, given the diversity of cultures, social structures, and health care systems.

Thus, although the generic cost-driving factors associated with schizophrenia are likely to be similar around the world (management of the chronic or relapsing symptoms and impairments, provision of inpatient and outpatient care and medication, mortality, lost productivity and unemployability, impact on the family, and impact on the community), their relative weights and, hence, the structure of the direct and indirect costs of the illness are likely to vary considerably. Hospital or other residential care, which accounts for more than 75 percent of the direct costs of schizophrenia in high-income countries, is likely to account for a smaller fraction in developing countries because of lower staffing and equipment costs. On the contrary, the proportion of direct costs attributable to the dispensing of antipsychotic medication, which is less than 5 percent of the

total direct costs in developed countries, is likely to be higher in developing countries.

It is difficult to estimate the economic impact of schizophrenia on families in developing countries. A study conducted in Nigeria revealed that caregiver *opportunity costs* (e.g., family members' lost productivity or income) were not of a different order from those in the developed world. Yet, the aggregate family costs in developing countries may, in fact, be substantially higher since a larger proportion of schizophrenic patients live with their families, as compared with patients in Western societies, and the family is likely to be the first line of treatment and management of psychotic episodes. The cost of purchasing prescribed maintenance medication is usually borne by the family. Additional direct costs may be incurred by the necessity to travel, often long distances, to the nearest hospital or clinic, as well as by payment for the services of traditional healers. In many traditional communities, the stigma associated with mental illness may affect the family as a whole and restrict, for instance, marital opportunities for younger family members. While lost educational opportunities are likely to be a problem, lost paid employment is more difficult to quantify and, therefore, less likely to appear prominently in estimates of the indirect illness burden in developing countries. On the contrary, reintegration of a family member who has suffered a psychotic episode into the domestic economy may be much easier to achieve than formal employment, and this may be one factor in the better and longer remissions of patients with schizophrenia observed in developing countries.

Another aspect of the social and economic costs of schizophrenia is the commonly perceived association between the disease and the criminal behavior. Offences committed by persons with schizophrenia tend to receive wide media coverage and to reinforce popular ideas about dangerousness associated with mental illness. It is important to dispel such prejudicial attitudes. Carefully designed studies in Europe and elsewhere indicate that a small proportion of patients with schizophrenia tend to be overrepresented among the perpetrators of violent offenses, including homicide. However, the *population-attributable fraction* of such offenses committed by persons with schizophrenia is negligible compared with the total number of offenses in the community. Moreover, the rate of arrest and incarceration is likely to be high among patients with schizophrenia because of their conspicuous behavior and appearance, rather than the seriousness of the offense. Issues of protecting patients' rights must, therefore, be an important part of antistigma campaigns in both developing and developed countries.

FUTURE DIRECTIONS

Important insights into the nature and causes schizophrenia have been

gained from population-based studies, although essential questions still remain unanswered. The clinical syndrome of schizophrenia appears to be robust and identifiable reliably in diverse populations and cultures, suggesting that a common pathophysiology and, possibly, a common genetic predisposition may underlie its manifestations. At the level or large population aggregates, no major differences in incidence and morbid risk have to date been detected, though small geographical area variation exist and appear to be related to a mix of locally operating risk factors whose effects may be attenuated in large, heterogeneous populations.

The study of "atypical" populations, such as genetic isolates or minority groups, may be capable of detecting unusual variations in the incidence of schizophrenia that could provide novel clues to the etiology and pathogenesis of the disorder. Notwithstanding the difficulties in the genetic dissection of complex disorders, emerging powerful methods of genomic analysis will eventually identify polymorphisms and haplotypes associated with schizophrenia risk. The majority are likely to be in genes of small effect, although one cannot rule out the possibility that rare mutations of moderate or major effect will also be found, especially in isolate populations or at the level of neurocognitive and neurophysiological abnormalities associated with the disorder. Establishing their population frequency and associations with a variety of phenotypes, including premorbid personality traits, will be a major task for comparative epidemiology.

At present, no single, or major, environmental risk factor influencing the incidence of schizophrenia or other psychoses has been conclusively demonstrated. Further studies using large samples are required to evaluate potential risk factors, antecedents, and predictors, for which the present evidence is inconclusive. The relationship between genotype and phenotype in schizophrenia is likely to be mediated by complex causal pathways involving gene–gene and gene–environment interactions, "programmable" neural substrate, and stochastic events.

Current epidemiological research is increasingly making use of existing large databases, such as cumulative case registers or birth cohorts to test hypotheses about risk factors in case–control designs. Methods of genetic epidemiology are being integrated within population-based studies. These trends predict an important future role for epidemiology in the unraveling of gene–environment interactions that are likely to be the key to the understanding of the etiology of schizophrenia. In this context, research into psychotic disorders in non-Western populations can provide valuable information on genetic heterogeneity, the impact of the environment, and the course and outcome of schizophrenia. Both traditional communities and societies undergoing transition in their social and economic organization can contribute to the better understanding of the relationships between culture and mental disorder and the variety of human experience

in coping with mental illness.

The vast majority of people whose lives are profoundly affected by schizophrenia reside in developing countries. As indicated previously, the incidence and prevalence of schizophrenia in developing countries are not different from what is being found in the developed world, or may even be higher. Although a greater proportion of people with schizophrenia in developing countries experience longer periods of symptomatic recovery than patients in the West, the burden of illness is severe and affects the productivity and quality of life of many families. The predicament of patients and families is likely to worsen as many of these countries are experiencing the pressures of economic restructuring, increasing income inequality, unemployment, and cuts in public spending on health care. The traditional economic and psychosocial resources of the family, which in the past have been capable of absorbing much of the impact of severe mental illness, may quickly become eroded.

A common denominator in the great variety of situations and systems of care for people afflicted with psychotic illnesses in the developing world is that the modest successes achieved in some of these countries are vulnerable to the local repercussions of economic globalization. In all societies, the stigma surrounding mental illness is likely to mean that people with severe mental disorders are among the first to be disadvantaged by shrinking government health expenditures. It is, therefore, imperative to initiate proactive measures on an international scale to forestall such developments.

References

Arajärvi R, Haukka J, Varilo T, et al. Clinical phenotype of schizophrenia in a Finnish isolate. *Schizophr Res*. 2004;67:195–205.

Bhugra D, Bhui K, eds. *Textbook of Cultural Psychiatry*. Cambridge: Cambridge University Press; 2007.

Brown S. Excess mortality of schizophrenia. A meta-analysis. *Br J Psychiatry*. 1997;171:502–508.

Bulayeva KB, Leal SM, Pavlova TA, et al. Mapping genes of complex psychiatric diseases in Daghestan genetic isolates. *Am J Med Genet B Neuropsychiatr Genet*. 2005;132B:76–84.

*Committee on Nervous System Disorders in Developing Countries, Institute of Medicine Board on Global Health. *Neurologic, Psychiatric, and Developmental Disorders. Meeting the Challenge in the Developing World*. Washington, DC: The National Academies Press; 2001.

Eaton JW, Weil RJ. *Culture and Mental Disorders*. Glencoe, IL: The Free Press; 1955.

Gelder MG, López-Ibor JJ Jr, Andreasen NC, eds. *New Oxford Textbook of Psychiatry*. Vol 1, Part 4, Section 4.4. Oxford: Oxford University Press; 2000.

Gruenberg EM. The epidemiology of schizophrenia. In: Arieti S, ed. *American Handbook of Psychiatry*. 2nd ed. Vol. 2. New York: Basic Books; 1974.

Harrison G, Glazebrook C, Brewin J, et al. Increased incidence of psychotic disorders in migrants

from the Caribbean to the United Kingdom. *Psychol Med*. 1997;27:799–806.

Hegarty JD, Baldesarini RJ, Tohen M, Waternaux C, Oepen G. One hundred years of schizophrenia: a meta-analysis of the outcome literature. *Am J Psychiatry*. 1994;151:1409–1416.

Hirsch SR, Shepherd M, eds. *Themes and Variations in European Psychiatry. An Anthology*. Bristol: John Wright & Sons; 1974.

*Hopper K, Harrison G, Janca A, Sartorius N, eds. *Recovery From Schizophrenia. An International Perspective. A Report from the WHO Collaborative Project, The International Study of Schizophrenia*. Oxford: Oxford University Press; 2007.

*Jablensky A, Sartorius N, Ernberg G, et al. Schizophrenia: manifestations, incidence and course in different cultures. A World Health Organization Ten-Country Study. Psychological Medicine. *Psychol Med Monogr Suppl*. 1992;20:1–97.

Keith SJ, Regier DA, Rae DS. Schizophrenic disorders. In: Robins LN, Regier DA, eds. *Psychiatric Disorders in America. The Epidemiologic Catchment Area Study*. New York: The Free Press; 1991.

Kraepelin E. *Dementia Praecox and Paraphrenia*. Edinburgh: Livingstone; 1919.

McGrath J, Saha S, Welham J, El Saadi O, MacCauley C, Chant D. A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC Med*. 2004;2:13.

Mortensen PB, Juel K. Mortality and causes of death in first admitted schizophrenic patients. *Br J Psychiatry*. 1993;163:183.

Murphy HBM. *Comparative Psychiatry. The International and Intercultural Distribution of Mental Illness*. Berlin: Springer; 1982.

*Murray RM, Jones PB, Susser E, van Os J, Cannon M, eds. *The Epidemiology of Schizophrenia*. Cambridge: Cambridge University Press; 2003.

Myles-Worsley M, Coon H, Tiobech J, et al. Genetic epidemiological study of schizophrenia in Palau, Micronesia: prevalence and familiality. *Am J Med Genet (Neuropsychiatr Genet)*. 1999;88:4–10.

Nimgaonkar VL, Fujiwara TM, Dutta M, et al. Low prevalence of psychoses among the Hutterites, an isolated religious community. *Am J Psychiatry*. 2000;157:1065–1070.

*Perälä J, Suvisaari J, Saarni SI, et al. Lifetime prevalence of psychotic and bipolar I disorders in a general population. *Arch Gen Psychiatry*. 2007;64:19–28.

Phillips MR, Yang G, Li S, Li Y. Suicide and the unique prevalence pattern of schizophrenia in mainland China: a retrospective observational study. *Lancet*. 2004;364:1062–1068.

Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia. *PLoS Med*. 2005;2(5):e411.

Scully PJ, Owens JM, Kinsella A, Waddington JL. Schizophrenia, schizoaffective and bipolar disorder within an epidemiologically complete, homogeneous population in rural Ireland: small area variation in rate. *Schizophr Res*. 2004;7:143–155.

Selten JP, Slaets J, Kahn RS. Schizophrenia in Surinamese and Dutch Antillean immigrants to the Netherlands: evidence of an increased incidence. *Psychol Med*. 1997;27:807–811.

Semple DM, McIntosh AM, Lawrie SM. Cannabis as a risk factor for psychosis: systematic review. *J Psychopharmacol*. 2005;19:187–194.

Varilo T, Peltonen L. Isolates and their potential use in complex gene mapping efforts. *Curr Opin Genet Dev*. 2004;14:316–323.

Wig NN, Menon DK, Bedi H, et al. Expressed emotion and schizophrenia in North India. I. Cross-cultural transfer of ratings of relatives' expressed emotion. *Br J Psychiatry*. 1987;151:156–160.

World Health Organization. *The World Health Report 2001. Mental Health: New Understanding, New Hope*. Geneva: World Health Organization; 2001.

▲ 12.4 Genetics of Schizophrenia

KIMBERLEY M. KENDALL, B.Sc. (HONS.), M.B.B.CH., M.R.C.PSYCH., GEORGE KIROV, M.D., PH.D., AND MICHAEL J. OWEN, M.D., PH.D.

Ever since Kraepelin delineated the disorder dementia praecox in 1899, psychiatrists could not fail to notice that this illness tended to run in families. Four years later, the first description of a human autosomal genetic disorder (brachydactyly type A1) brought attention to the burgeoning new science of genetics and efforts to establish the etiology of schizophrenia soon began.

FAMILY STUDIES

Any attempt to establish that a disorder has a genetic susceptibility must begin with the demonstration that relatives of probands have an increased risk of developing the disorder. The first systematic family study to show this was published in 1916 by Ernst Rüdin and, over the next half a century, several more family studies would follow. The combined data demonstrated convincingly the familial basis of schizophrenia. Figure 12.4–1 clearly shows that the risk of developing schizophrenia is increased in the relatives of probands. The risk of developing schizophrenia is higher in siblings and children of probands than in other classes of relatives, reaching a lifetime risk of 10 percent or more.

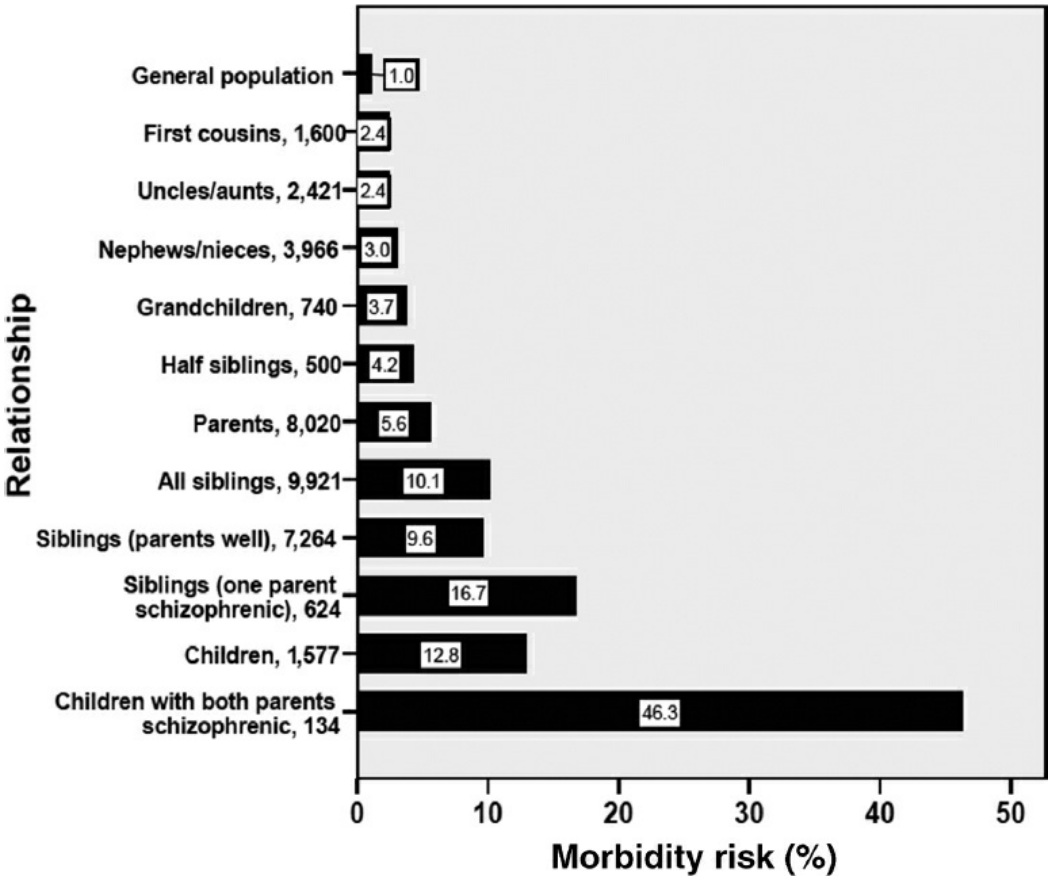

FIGURE 12.4–1. Morbidity risk for schizophrenia in the relatives of schizophrenic probands. The numbers are based on the study by Gottesman and Shields (1982) and used with permission by Irving Gottesman.

The validity of these early studies was later questioned due to the lack of control groups, the nonblind nature of the diagnoses, and the lack of operationalized diagnostic criteria. The possibility was raised that these factors had led to overestimates of the morbidity risk, so a new wave of family studies followed. Between 1989 and 1995, several further large family studies were published, this time using more rigorous diagnostic methodology and including control groups. These studies confirmed that schizophrenia runs in families but calculated smaller morbidity risks than earlier studies (1.4 to 6.5 percent in first-degree relatives of probands with schizophrenia versus 0 to 1.9 percent in controls). These differences may have been due to varying methodologies and diagnostic criteria.

The most recent family studies separate out data on the morbid risk among first-degree relatives. Of 725 parents studied, 17 had schizophrenia, giving a lifetime risk of 2.3 percent. Of 645 siblings studied, 48 had schizophrenia, giving a lifetime risk of 7.4 percent. Of 246 offspring studied, 37 developed schizophrenia, resulting in a lifetime risk of 15 percent. The newer studies confirmed the higher risk among siblings and offspring than that among parents and established that the morbidity risks were not vastly different from those calculated in older studies (Fig. 12.4–

2). The morbidity risks calculated in the older and newer studies were of similar magnitude for siblings (10.1 percent old versus 7.4 percent new) and offspring (12.9 percent old versus 15 percent new), but for parents, the difference was more substantial (5.6 percent old versus 2.3 percent new) and may have resulted from the use of more restrictive diagnostic criteria in recent studies that excluded a proportion of parents with milder forms of the illness. In addition, a large proportion of parents display features of the extended phenotype of schizophrenia (e.g., schizotypal personality disorder).

Overall, both the old and new family studies established the high familial loading of schizophrenia with siblings consistently showing around an 8- to 10-fold increased risk of developing schizophrenia compared to the rest of the population. This is greater than the increased risk reported for any environmental factor.

The estimated risk of schizophrenia among the offspring of matings between two affected parents is particularly high (Fig. 12.4–1), but this was based on a small number of observations. In 2010, a Danish national register-based cohort study examined the risk of severe mental disorder in offspring with two psychiatrically ill patients. The risk of schizophrenia in offspring of parents both affected by schizophrenia was 27.3 percent. This increased to 39.2 percent when schizophrenia-related disorders were included, largely confirming the previous findings.

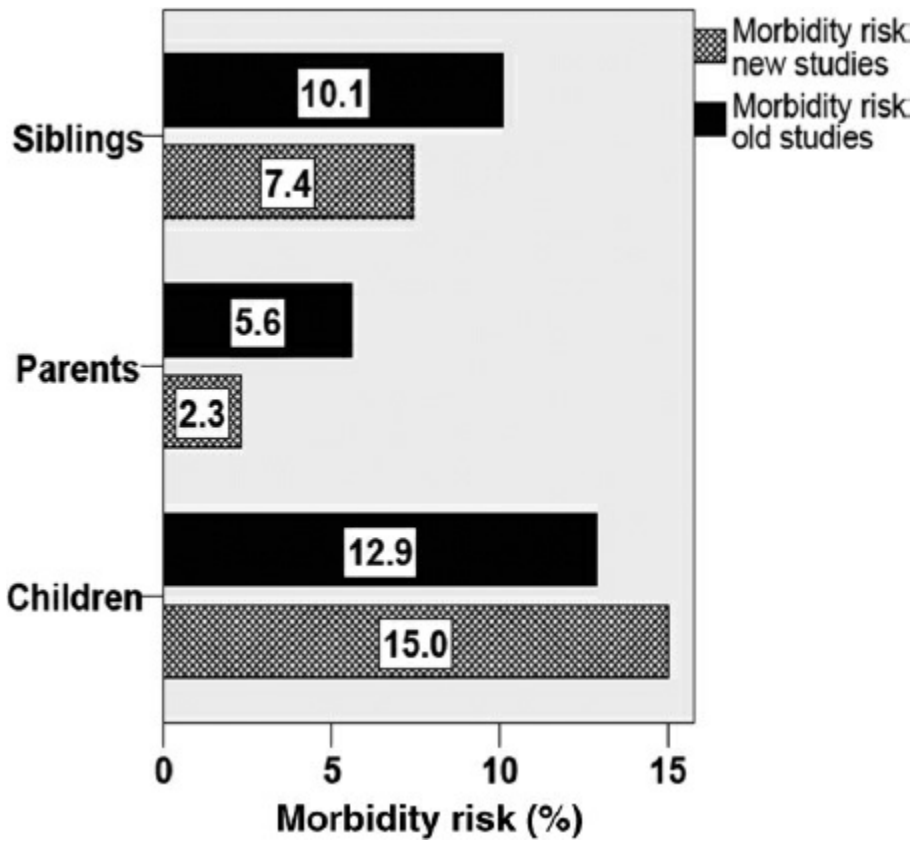

FIGURE 12.4–2. Morbidity risk for schizophrenia in the old and the newer family

studies.

TWIN STUDIES

Twins reared together share many aspects of their environment. Monozygotic (MZ) twins share all of their genes while dizygotic (DZ) twins share, on average, 50 percent. Twin studies examine the proportion of twin pairs where both twins are affected or the proportion of co-twins who are also affected by a particular trait (i.e., the concordance rate). If it is assumed that MZ and DZ twins share their environments to a roughly equal extent (known as the shared environment assumption), a higher concordance in MZ twins implies that the disorder is, at least partly, genetic. A concordance of less than 100 percent in MZ twins indicates that environmental factors are also at play. The two methods of estimating the concordance rate are the pairwise method and the probandwise method. In the pairwise method, concordance is simply the number of concordant pairs divided by the total number (e.g., if five pairs of twins are affected and in each of another five pairs, only one twin is affected, the pairwise concordance rate is 5/10 = 50 percent). In the probandwise method, if both members of a pair were ascertained independently (i.e., they are both probands), the pair is counted twice because each of these probands has an affected co-twin. The probandwise concordance rate is given by the number of affected co-twins divided by the number of probands. Each doubly ascertained pair will have two probands and two affected co-twins. In the example described previously, the probandwise concordance rate is $(2 \times 5)/(2 \times 5 + 5) = 10/15 = 66.7$ percent. The probandwise method indicates how many independently ascertained probands have an affected twin, or the risk of the twin of a person with schizophrenia developing the disorder him- or herself. On average, the probandwise method results in concordance rates that are higher than those obtained by the pairwise method by around a quarter. Probandwise concordance rates are preferred by geneticists as they are technically more correct and directly comparable to the population risks reported in family studies.

Table 12.4–1 summarizes the major reported twin studies in schizophrenia that used systematic ascertainment. The lack of raw data in some publications prevents the estimation of probandwise concordance rates. In some cases, these have been obtained via personal communication by previous reviewers or assumptions have been made. Authors have used different severities of illness (e.g., “strict schizophrenia” or “any psychotic disorder”) and different diagnostic classifications. In an attempt to preserve some uniformity, the authors of this chapter followed the interpretations from the reviews by Gottesman and Kendler. Heritability estimates are those reported originally or estimated in reviews. The concordance rates are remarkably consistent. Heritability estimates vary between 53 percent and 90 percent, with two meta-analyses providing

estimates of 81 percent and 88 percent.

Identical Twins Reared Apart

In theory, the shortcomings of the shared environment assumption can be avoided by studying MZ twins who have been reared apart. However, such twins are very rare. Gottesman and Shields described fourteen pairs of MZ twins reared apart, nine of whom were concordant for schizophrenia. This gives a pairwise concordance (64 percent) that is higher than that for twins reared together. However, these must have been very unusual cases of adoption.

The Offspring of Twins Discordant for Schizophrenia

There are two possible explanations for discordance of a heritable disorder in MZ twins. The first is that the affected twin suffers from an environmentally determined form of the disorder (termed a phenocopy). The second is that both twins inherited the same increased genetic liability but that this is expressed only in the affected twin. Two studies have attempted to distinguish between these two explanations by comparing the rates of illness in the offspring of discordant MZ pairs. In one study, 11 twins with schizophrenia had 47 offspring, six of whom developed schizophrenia (16.8 percent risk). Six unaffected co-twins had 24 offspring, four of whom developed schizophrenia (17.4 percent risk). These data are consistent with the explanation that both twins inherited the same increased genetic liability but that this was expressed only in the affected twin. In another study of 28 offspring of twins with schizophrenia, three developed the disorder (10.7 percent risk). Of 45 offspring of unaffected twins, only one developed the disorder (2.2 percent risk). The differences in the conclusions of these studies may be due to small sample sizes.

ADOPTION STUDIES

One shortcoming of the twin method is that there tend to be greater similarities in the environments shared by MZ twins compared to DZ twins. It is difficult to separate the contribution of factors in families because children share both genes and environment with their parents and siblings. One way to control for environmental factors is by conducting adoption studies. If there is a genetic component to the disorder being studied, the similarity between adopted children and their biological parents should be higher than the similarity between adopted children and their adoptive parents. It is assumed that adoption does not increase the risk of developing schizophrenia and that placement is random. Performing such studies is extremely difficult as it involves the tricky and ethically problematic process of tracing biological parents and adopted children and interviewing them.

Table 12.4–2 summarizes the results of adoption studies in schizophrenia. All of these adoption studies came to very similar conclusions. Children of parents with schizophrenia have at least a 10 percent risk of developing schizophrenia, schizoaffective disorder, or other narrow-spectrum schizophrenic disorder when adopted away very soon after birth. This risk is very similar to that reported among children of parents with schizophrenia examined in family studies (Fig. 12.4–1), suggesting that most of the familial transmission is genetic.

ADVANCES IN THE MOLECULAR GENETICS OF SCHIZOPHRENIA

Family, twin, and adoption studies have demonstrated that genetics play an important role in conferring risk of schizophrenia, and the estimated heritability of the disorder has been calculated at around 80 percent. This means that 80 percent of the variation in the trait is due to variation in genetic factors. However, identification of risk at the level of DNA variation initially proved challenging, slow to advance, and difficult to replicate. The advent of more ambitious and technologically advanced techniques, including GWA studies, copy number variant (CNV) analysis, and next-generation sequencing, has recently resulted in greater success and optimism for the future.

Table 12.4–1.
Concordance Rates in Twin Studies and Heritability Estimates

		MZ				Same Sex DZ				
		Pairwise		Probandwise		Pairwise		Probandwise		
Author/Year	Country	N con/ Total	%	N Co-Twins/ Proband	%	N Con/ Total	%	N Co-Twins/ Proband	%	Herit ability[a]
Luxenburger (1928)	Germany	11/19	58	14/22	64	0/13	0	0/13	0	—
Rosanoff (1934)	US, Canada	25/41	61	Not known	—	7/53	13	Not known	—	63–84
Essen–Möller (1941)	Sweden	7/11 2/8	64 25	7/11 2/8	64 25	4/27	15	Not known	—	87
Essen–Möller (1970)[b]										
Kallmann (1946)	US	120/174	69	191/245	78	34/296	11	59/318	19	90
Slater (1953)	UK	24/37	65	28/41	68	8/58	14	11/61	18	73
Inouye (1961, 1972)	Japan	34/58	59		—	3/20	15		—	66
Kringlen (1968)	Norway	21/55	38	31/69	45	8/90	9	14/96	15	61
Fisher (1973)	Denmark	10/21	48	14/25	56	8/41	20	12/45	27	41

Gottesman and Shields (1972)	UK	11/22	50	15/26	58	3/33	9	4/34	12	86
Tienari (1975)	Finland	3/20	15	6/21	29	3/42	7	6/45	13	53
Kendler and Robinette (1983)	US	30/164	18	60/194	31	9/268	3	18/277	6	71
Onstad (1991)	Norway	8/24	33	15/31	48	1/28	4	1/28	4	87
Kläning (1996)	Denmark	4/13	31	7/16	44	1/18	6	2/19	11	83
Cannon (1998)	Finland	20/67	30	40/87	46	9/186	5	18/195	9	82
Franzek and Beckmann (1998)	Germany	11/22	50	20/31	65	5/23	22	7/25	28	79
Cardno (1998)	UK	13/40	33	20/47	43	0/50	0	0/50	0	83
Leonhard (1976, 1981, 1982)[c]	Germany	19/32	59	30/44	68	6/36	17	7/34	21	63
Meta-analysis, Cardno (2000)										88
Meta-analysis, Sullivan (2003)										81

MZ, monozygotic; DZ, dizygotic; S, systematic ascertainment; NS, nonsystematic ascertainment; CA, complete ascertainment; RPA, random population ascertainment; "N con/total," number of concordant pairs/total number of pairs used to calculate the pairwise concordance rate; "N co-twins/probands," number of affected co-twins/number of probands, used to calculate the probandwise concordance rate.

[a]Heritability is based on population rates of SZ provided in the study or estimated in previous reviews of these studies.

[b]Essen–Möller's own strict diagnostic interpretation of the same MZ twins, after a 30-year long follow-up.

[c]Probandwise rates for Leonhard's twins are taken from his personal communication to Kendler (1982), pairwise rates are from his publications in 1976 and 1981.

Table 12.4–2. Adoption Studies and Schizophrenia Risk

Study	Design	Genetic Relatives of a Schizophrenic Person	Not Genetically Related to a Schizophrenic Person	Comment
Heston (1966), Oregon, US	Subjects born mostly in psych. hosp. to schizophrenic/psychotic	5 (10.6%) of 47 adopted away children became schizophrenic, 4 more	0% schizophrenia in 50 control	

	mothers, separated from birth	had low IQ <70	children adopted away from healthy mothers	
Higgins (1976), Denmark	Comparison of 25 children separated from schizophrenic mothers at age 0–36 months, with 23 children reared by their schizophrenic mothers	Adopted away children: 4 (16%) developed chronic schizophrenia, 2 suicided, another 4 had SA or acute psychotic episodes. Of 23 children reared by their schizophrenic mothers, 4 became schizophrenic, 4 had schizoid personality disorders	—	Based on th "Copenha sample
Rosenthal (1971), Denmark	Compared 52 adopted away children of schizophrenic parents most of whom became ill after the adoption, with 67 control adoptees	14 (26.9%) of 52 children of SZ parents developed a broad schizophrenia spectrum illness	12 (17.9%) of 67 control adoptees developed a broad schizophrenia spectrum illness	Parental diagnosis schizophr possible, reactive c borderlin schizophr "Copenha sample
Wender (1974), Denmark	Cross-fostering: 28 children of normal parents raised by schizophrenic parents and 69 adopted children of schizophrenic parents	13 (18.8%) of 69 children of schizophrenic parents developed "probable borderline or more severe" disorders (Overlap with Rosenthal)	3 (10.7%) of 28 children of normal parents, adopted by future schizophrenics developed "probable borderline or more severe" disorders	"Copenhage sample + small frac of the "Provinci sample"
Kety (1994), Denmark	Compared the biological and adoptive relatives of 47 schizophrenic patients who were adopted with the relatives of 47 control adoptees	14 (5%) of 279 biological relatives were chronic schizophrenic and 30 (10.8) had "latent" schizophrenia	0% schizophrenia and 2 (1.8%) "latent" schizophrenia out of 111 adoptive relatives. 1 (0.3%) schizophrenia and 6 (1.7%) "latent" schizophrenia out of 351 biological and adoptive relatives of 47	Combined d from "Copenha and "Provinci samples

			control adoptees	
Mirsky (1995), Israel	Kibbutz-reared children compared with children reared by their families	Among 25 children of schizophrenic parents reared in kibbutz: 3 developed schizophrenia (12%) and 3 schizophrenia spectrum disorders	25 control children reared in kibbutz: no case of schizophrenia	Children in kibbutz h contact w biological parents
Tienari (2000), Finland	Compared offspring of schizophrenic or paranoid mothers with offspring of healthy mothers	17 (10.4%) of 164 offspring of schizophrenic/paranoid mothers developed schizophrenia (11), schizoaffective (1), schizophreniform (1), or schizotypal disorder (4), that is, narrow spectrum schizophrenia	4 (2%) of 197 control adoptees developed schizophrenia	Nationwide register o SZ/paran psychotic mothers, prospecti
Lichtenstein et al. (2009), Sweden	Multigeneration population register of 9,009,202 individuals	Adopted away children of affected parents had a relative risk of schizophrenia of 13.7	No cases of schizophrenia in adoptive parents of affected adoptees	Nationwide register

GWA Studies

GWA studies examine common alleles throughout the genome for association with a particular trait. The HapMap Project annotation of millions of single nucleotide polymorphisms (SNPs) has demonstrated that many SNPs are highly correlated with neighboring SNPs. Therefore, it is not necessary to genotype all of them in order to gain a nearly complete coverage of common variation in the genome. Furthermore, array-based technology has made it possible to genotype millions of SNPs on a single array, making the examination of variation across the whole genome in large samples a much less daunting prospect.

The earliest GWA studies were not particularly fruitful, but they did provide some useful lessons, in particular the need to study extremely large samples, due to the need to correct for the enormous multiple testing involved and the small odds ratios for risk conferred by the common variants. It was estimated that such studies require in a target of P value of $<5 \times 10^{-8}$ to achieve significance at a genome-wide level. This requires sample sizes in the order of tens of thousands. These could only be reached via collaboration between multiple research groups.

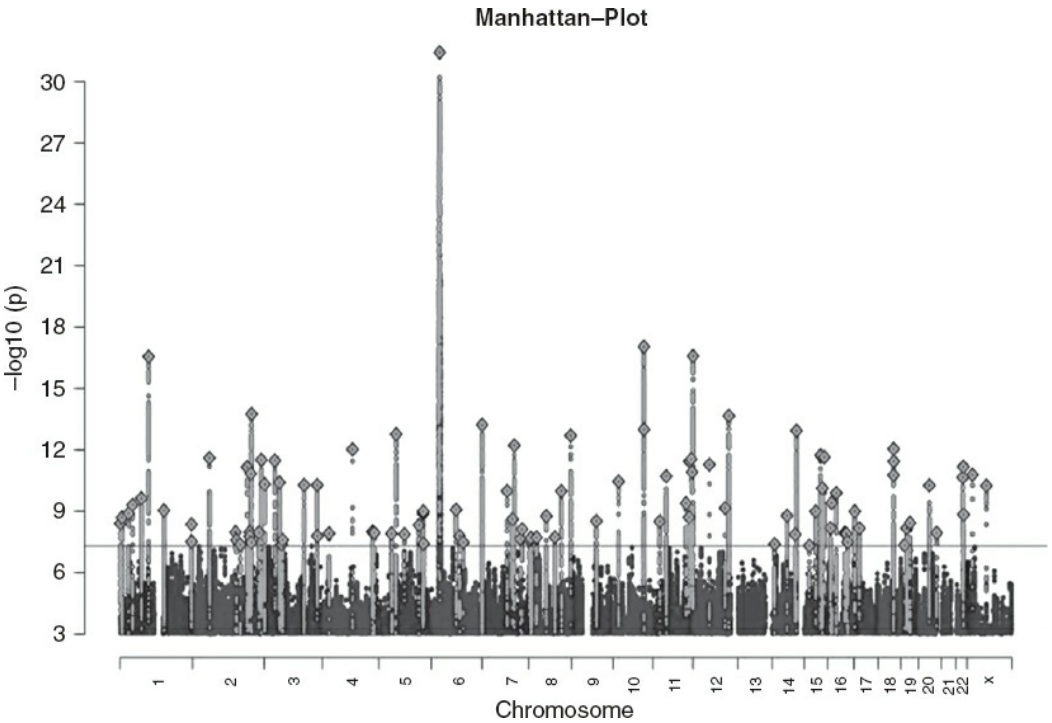

FIGURE 12.4–3. Manhattan plot of the discovery genome-wide association meta-analysis of 49 case–control samples (34,241 cases and 45,604 controls) and three family-based association studies (1,235 parent affected-offspring trios). Markers are shown according to their chromosome positions and level of association with schizophrenia, expressed as a negative log10 *P*-value (i.e., the numbers show the number of decimal points after the zero). (From Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511(7510):421–427.)

The largest and most successful schizophrenia GWA studies to date were conducted by the Psychiatric Genomics Consortium, a collaboration bringing together data from research groups around the world. This study combined most available schizophrenia samples with GWA genotypes into a single analysis and gives the most recent snapshot of progress in the area to date. The authors took genome-wide genotype data and constructed 49 ancestry matched nonoverlapping case–control samples (46 European ancestry, 3 East Asian ancestry; 34,241 cases and 45,604 controls) and three family-based samples of European ancestry (1,235 parent-affected offspring trios). After quality control and controlling for population stratification, 9.5 million genetic variants were considered. The results are depicted in Figure 12.4–3.

The authors then took independent SNPs, achieving a P value of $<1 \times 10^{-6}$ and obtained extra results from the deCODE genetics sample (1,513 cases and 66,236 controls of European ancestry). The deCODE data were combined with the primary sample to give a final data set of 36,989 cases and 113,075 controls. This analysis discovered 128 independent SNPs at 108 loci that exceeded genome-wide significance ($P < 5 \times 10^{-8}$) for association with schizophrenia. Seventy-five percent of these loci included

protein-coding genes and a further 8 percent were within 20 kilobase (kb) of a gene. Of note were associations with genes thought to be related to the etiology and treatment of schizophrenia (DRD2—dopamine receptor D2), glutamatergic neurotransmission and synaptic plasticity (GRM3, GRIN2A, SRR, GRIA1), and calcium signaling (CACNA1C, CACNB2, and CACNA1I). For each locus associated with schizophrenia in this study, the authors identified a set of SNPs that likely account for the association. However, signals could be potentially attributed to a known nonsynonymous exonic polymorphism in only 10 cases. This suggests that the majority of variants associated with schizophrenia that were identified by this study do not exert their influence by altering protein structure but rather that they alter gene expression. However, further studies, including sequencing of the associated regions, are required to identify the actual risk variants.

The common variants identified at genome-wide levels of statistical significance represent the tip of an iceberg of common allele effects. This has become apparent with the development of methods, such as the polygene score, to identify the en masse effects of common risk alleles including the many that are not individually supported at genome-wide levels of significance. These have estimated that relatively common small-effect alleles account for at least 25 percent of total liability to schizophrenia (about 33 percent of genetic liability). Such approaches now allow genetic risk to be estimated directly from genomic data rather than family history and will be of great utility to future population and clinical studies.

CNVs and Schizophrenia

CNVs are deletions or duplications of chromosomes of at least 1 kb in size. The earliest report of a rare CNV associated with an increased risk of schizophrenia was a deletion of about 2.5 Mb on chromosome 22q11.2 that has been known since about 1976 to cause a severe congenital disorder described variably as DiGeorge, Velo-Cardio-Facial, or Shprintzen syndrome. Carriers of this CNV were shown to have up to 30 percent risk of developing schizophrenia or other psychotic disorders. It took until 2008, after the development of high-resolution and high-throughput microarray technologies, for researchers to interrogate the whole genome and find other CNVs that increase risk for schizophrenia. All such CNVs turned out to be very rare, found in <1 percent of cases, and, therefore, very large sample sizes (of >10,000 cases and controls) were required to confirm their role. As of today, there are 12 CNVs in 11 loci that have reached the accepted genome-wide levels of statistical significance of *P* value of $<4.1 \times 10^{-4}$.

The first new CNV locus to be identified was *NRXN1*. It is different from the other CNVs, as it is the only one that involves a single gene. The gene is disrupted by CNVs that are of different sizes and affect different parts of

the gene. Rujescu et al showed in 2009 that it was the CNVs that intersected exons that increased the risk for schizophrenia. *NRXN1* proved to be an excellent candidate gene. It codes for a cell adhesion molecule, neurexin 1, which is specific to synapses in the brain, where it mediates interactions between pre- and postsynaptic structures. It plays a vital role in the formation, maintenance, and release of neurotransmitters at synapses.

All remaining CNVs shown in Table 12.4–3 share other features. They are flanked by large stretches of homologous DNA called low copy repeats that can misalign during meiosis or mitosis and cause the genetic material between them to be deleted or duplicated. Therefore, they are usually of the same size and their mutation rates are quite high compared to other CNVs. They all disrupt multiple genes (between three and more than 40), which probably explains their high pathogenicity. All the CNVs shown in Table 12.4–3 also increase the risk of developing intellectual deficit, developmental delay, autism spectrum disorders, and various congenital malformations and somatic diseases.

Table 12.4–3.
Schizophrenia-Associated Copy Number Variants

Locus	Chromosome Position	Combined Results from Meta-analysis			
		Cases	Controls	ORs (95% CI)	P-value
1q21.1 del	chr1:146,57–147,39	0.17% (33/19,056)	0.021% (17/81,821)	8.35 (4.65–14.99)	4.1×10^{-13}
1q21.1 dup	chr1:146,57–147,39	0.13% (21/16,247)	0.037% (24/64,046)	3.45 (1.92–6.20)	9.9×10^{-5}
NRXN1 exonic del	chr2:50,15–51,26	0.18% (33/18,762)	0.020% (10/51,161)	9.01 (4.44–18.29)	1.3×10^{-11}
3q29 del	chr3:195,73–197,34	0.082% (14/17,005)	0.0014% (1/69,965)	57.65 (7.58–438.44)	1.5×10^{-9}
WBS dup	chr7:72,74–74,14	0.066% (14/21,269)	0.0058% (2/34,455)	11.35 (2.58–49.93)	6.9×10^{-5}
15q11.2 del	chr15:22,80–23,09	0.59% (116/19,547)	0.28% (227/81,802)	2.15 (1.71–2.68)	2.5×10^{-10}
Angelman/ Prader–Willi dup	chr15:24,82–28,43	0.083% (12/14,464)	0.0063% (3/47,686)	13.20 (3.72–46.77)	5.6×10^{-6}
15q13.3 del	chr15:31,13–32,48	0.14% (26/18,571)	0.019% (15/80,422)	7.52 (3.98–14.19)	4.0×10^{-10}
16p13.11 dup	chr16:15,51–16,30	0.31% (37/12,029)	0.13% (93/69,289)	2.30 (1.57–3.36)	5.7×10^{-5}
16p12.1 del	chr16:21,95–22,43	0.16% (33/20,402)	0.05% (12/26,628)	3.3 (1.61–7.05)	3.4×10^{-4}
16p11.2 dup	chr16:29,64–30,20	0.35% (58/16,772)	0.030% (19/63,068)	11.52 (6.86–19.34)	2.9×10^{-24}
22q11.2 del	chr22:19,02–	0.29%	0.00%	NA (28.27 –	4.4 ×

20,26	(56/19,084)	(0/77,055)	∞)	10^{-40}

Numbers are according to the 2014 review (Rees et al., 2014), except for the 16p12.1 del which is reported in Rees et al., (2016).

The Penetrance of CNVs. The high risk of developing schizophrenia associated with these CNVs suggests that they could be informative for genetic counseling. Therefore, it is important to estimate their penetrance, that is, the risk of a CNV carrier developing disorder. Initial studies reported modest penetrance but they took into account only the risk of schizophrenia. Over the years, it became evident that each of the CNVs listed in Table 12.4–3 also increased the risk for developmental delay, autism spectrum disorders, and various congenital malformations. This was established in large studies on (usually young) patients referred for clinical genetic testing because of a variety of early-onset disorders. It became clear that the risk among carriers of these CNVs to develop one of these early-onset disorders was even greater than the risk of schizophrenia. Therefore, when considering the penetrance of these CNVs, one should take into account the risk of developing any serious disorder, not just schizophrenia. The authors recently estimated the penetrance of CNVs for both types of conditions and showed that the penetrance for the group of developmental delay/autism spectrum disorders and congenital malformations was much higher than that for schizophrenia. Taken together, the penetrance for any serious developmental disorder is substantial, averaging just more than 40 percent and is accounted for mostly by nonpsychotic disorders (Fig. 12.4–4).

CNVs Point to Biological Pathways. The rate of de novo CNVs in patients suffering from schizophrenia is about 5 percent. This is at least twice higher than among controls. Analysis of these de novo CNV loci suggests that genes within them tend to cluster in certain biological pathways. A systems-based statistical approach found significant enrichment at CNV loci for genes involved in the postsynaptic density. This enrichment was explained by genes encoding components of the *N*-methyl-D-aspartate (NMDA) receptor signaling complex and synaptic proteins that interact with the activity-regulated cytoskeleton-associated protein (ARC) complex. Arc is a member of the immediate-early gene (IEG) family and its mRNA is localized to activated synaptic sites in an NMDA receptor–dependent manner. It is believed to play a critical role in learning and memory-related molecular processes.

Sequencing Studies

When examining association with disease, GWA studies are used to detect common variation, particularly SNPs, and the genotyping platforms used

can also detect CNVs. However, these platforms will not detect rare mutations such as single nucleotide variants (SNVs) and small insertions and deletions (indels). Several teams have used exome sequencing in an attempt to detect small and rare variants, including de novo mutations. Whole exome sequencing involves the sequencing of all the exons in the genome and, due to recent technological advances, can be done on a high throughput basis.

In the largest study of its kind to date, researchers used massively parallel short-read sequencing to analyze the exomes of 2,536 individuals with schizophrenia and 2,543 controls. To correct for multiple testing across every nucleotide in the exome would result in the requirement for very high significance levels to support association with schizophrenia. Therefore, the authors studied sequences from a set of 2,500 genes with preexisting evidence (e.g., GWA studies or certain pathways) for involvement in schizophrenia. The variant achieving the highest level of significance occurred at a known schizophrenia risk locus; a missense allele in the *CCHCR1* gene in the major histocompatibility complex (MHC).

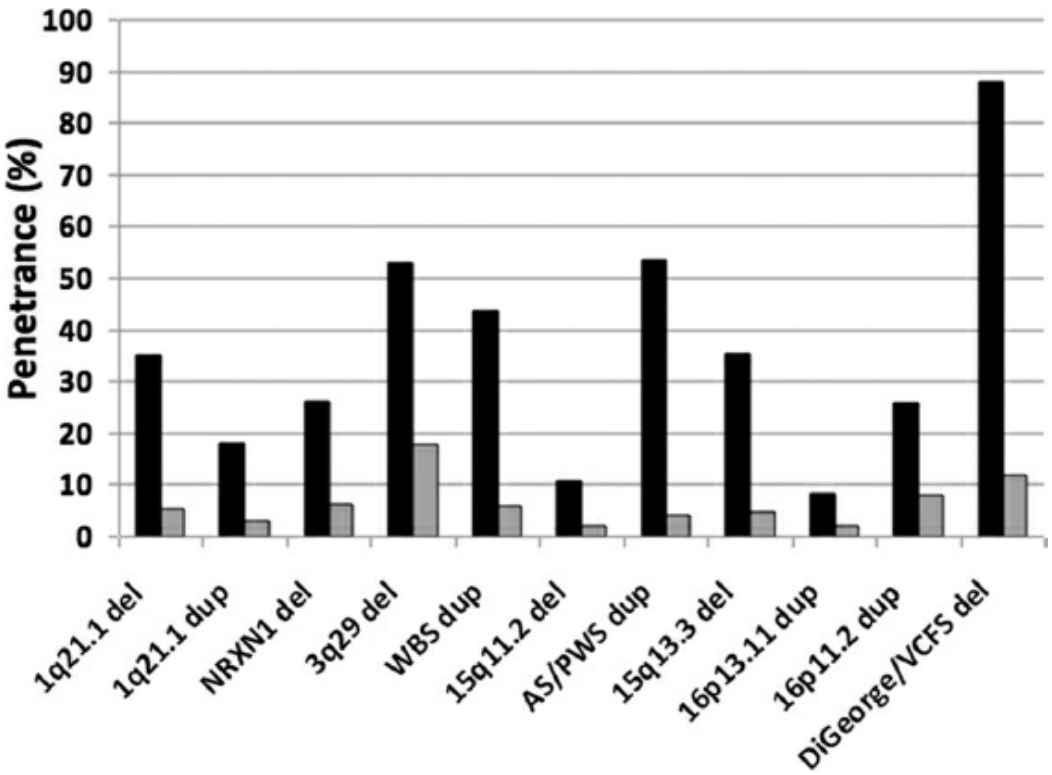

FIGURE 12.4–4. Penetrance of schizophrenia-associated CNVs. Numbers follow the estimates in the study by Kirov et al. (2014). Grey bars: penetrance of schizophrenia, black bars: penetrance of the other developmental disorders.

Of the 635,944 coding and splice site variants discovered in this study, 56 percent were singletons. Those most likely to affect protein function were mostly singleton observations (69 percent of nonsense variants, 58 percent of missense variants, and 51 percent of silent variants were

singletons). There was a significantly higher rate of rare disruptive mutations in cases, even when taking into account potential differences in exome-wide burden between cases and controls and correction for multiple testing. Overall, the data suggest a polygenic burden of rare variants relatively concentrated in genes found to be associated with schizophrenia in other genetic studies. Examination of sets of genes enriched for these rare variants yielded similar results to previous studies. Enriched gene sets included the voltage-gated calcium ion channel family (implicated by GWA studies), postsynaptic gene sets particularly the ARC complex (implicated in de novo CNVs), and targets of the fragile X mental retardation protein (FMRP) (implicated by GWA studies).

In a second study, published at the same time, exome sequencing was performed on the largest sample of parent–offspring trios studied to date in order to identify de novo SNVs and small indels in schizophrenia. This analysis also revealed a convergence on components of the NMDA receptor signaling complex and synaptic protein interactors of ARC, as well as an enrichment of mutations in the targets of the FMRP complex.

THE EMERGING GENETIC ARCHITECTURE OF SCHIZOPHRENIA

While much genetic risk is still unexplained at the level of individual risk variants, the findings to date support a general framework for the genetic architecture of schizophrenia. As expected from genetic epidemiology and population genetics, it appears that there is a spectrum of allelic risk. There are contributions from alleles that are common in the population. The effect sizes of these tend to be small due to the effects of natural selection, but as group, they probably account for about a third of genetic liability. There is also a contribution from rare alleles, some of which have a large effect on individual risk of disease pending their removal from the population by selection. The relationship between effect size and allele frequency is illustrated in Figure 12.4–5. The relative contribution of alleles of different effect sizes and frequencies is unclear, but this should be clarified as larger whole exome and whole genome sequencing studies are conducted. However, it seems likely that alleles across the risk spectrum will be found with their effect sizes constrained by natural selection.

Pleiotropy

An important implication of recent studies is that genetic risk does not recognize the current definitions of disease. Quite extensive pleiotropy with respect to diagnosis has been demonstrated for common alleles at the level of individual SNPs, genes, and the en masse effects of multiple common risk alleles, and for rare alleles including both CNVs and SNVs. Of course, the power of many studies to detect relatively nonspecific risk factors will be greater than it is to detect those associated with more specific outcomes,

and there is accumulating evidence for genetic effects with a degree of specificity at the phenotype level. However, an important question is whether current diagnostic categories are the best phenotypes for future genetic studies, for research aiming to understand relationships between genetic risk and clinical outcome, and to best understand pathophysiology.

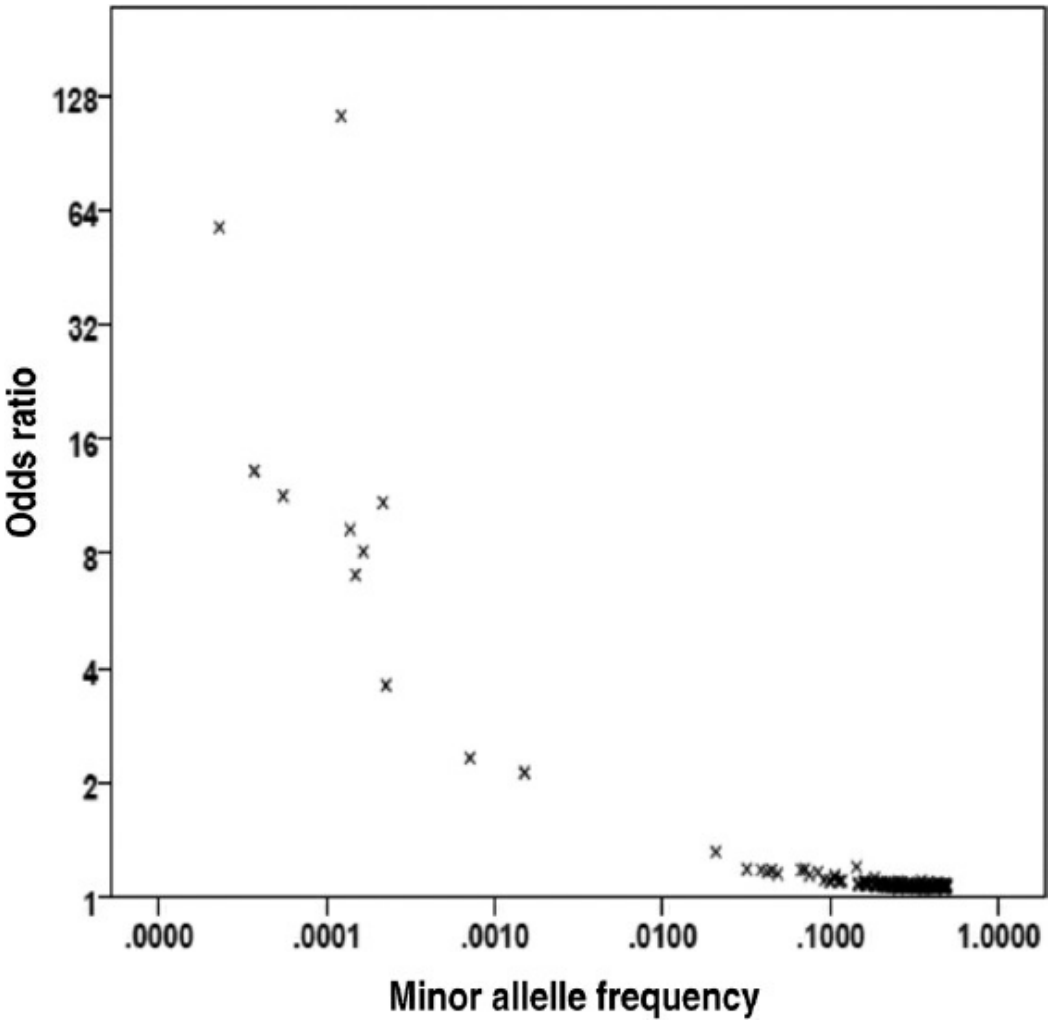

FIGURE 12.4–5. The relationship between frequency of alleles and their effect sizes. Data with OR >2 refer to CNVs reported in the study by Rees et al. (2014). Remaining data are SNPs reported in the Schizophrenia Working Group of the Psychiatric Genetics Consortium paper, 2014.

CONVERGENCE ONTO BIOLOGICAL MECHANISMS

As indicated, despite the undoubted complexity of the genetic picture that is emerging, work using different approaches is beginning to reveal some convergence onto a coherent set of biological processes. Results from GWA studies, CNV, and sequencing studies all point to a functionally related set of synaptic proteins involved in synaptic plasticity, learning, and memory. Among the gene sets implicated are the ARC and NMDAR complexes, targets of FMRP, and voltage-gated calcium channels. Although it seems highly likely that there are other systems and mechanisms are involved,

including the dopamine system, and that these can be expected to emerge with further and larger GWA studies and sequencing studies, the current findings offer numerous entry points for neuroscientists to probe the biological basis of schizophrenia.

RISK COUNSELING

It is understandable that individuals may wish to know how likely they, or their children, are to develop a psychotic illness, particularly if they have affected relatives. This task remains problematic. In the case of schizophrenia, one may inherit an increased susceptibility to develop the disorder but that this does not make it inevitable that the disorder will develop, as even MZ twins are concordant for the disorder in only about half the cases. As previously discussed, an increased genetic risk of developing schizophrenia may be accounted for by rare alleles of large effect and multiple common alleles of small effect. This complicated situation means that, at the current level of evidence, meaningful genetic counseling is unlikely to be appropriate for most patients. The exception is for carriers of a small number of specific CNVs that are found in only about 2.5 percent of patients. Fairly accurate estimates of risk are available for carriers of these CNVs and these can be used in counseling. It may be possible to use polygenic risk scoring to stratify individuals into high, medium, and low-risk groups, but the predictive power of such estimates is currently of no useful clinical value. As the genetics of schizophrenia is understood in better detail, additional information will likely allow the more appropriate use of genetic counseling.

REFERENCES

Bassett AS, Marshall CR, Lionel AC, Chow EW, Scherer SW. Copy number variations and risk for schizophrenia in 22q11.2 deletion syndrome. *Hum Mol Genet*. 2008;17(24):4045–4053.

Cardno AG, Gottesman II. Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics. *Am J Med Genet*. 2000;97:12–17.

Cardno AG, Rijsdisk FV, West RM, et al. A twin study of schizoaffective-mania, schizoaffecive-depression and other psychotic syndromes. *Am J Med Genet B Neuropsychiatr Genet*. 2012;159(B):172–182.

Fromer M, Pocklington AJ, Kavanagh DH, et al. De novo mutations in schizophrenia implicate synaptic networks. *Nature*. 2014;506:179–184.

Gottesman II. *Schizophrenia Genesis. The Origin of Madness*. New York: WH Freeman; 1991.

Gottesman II, Laursen TM, Bertelsen A, Mortensen PB. Severe mental disorders in offspring with 2 psychiatrically ill patients. *Arch Gen Psychiatry*. 2010;67(3):252–257.

Gottesman II, Shields J. *Schizophrenia. The Epigenetic Puzzle*. New York: Cambridge University Press; 1982.

Hall J, Trent S, Thomas KL, O'Donovan MC, Owen MJ. Genetic risk for schizophrenia: convergence on synaptic pathways involved in plasticity. *Biol Psychiatry*. 2015;77(1):52–58.

Ingraham LJ, Kety SS. Adoption studies of schizophrenia. *Am J Med Genet (Semin Med Genet)*. 2000;97:18–22.

Kendler KS. Overview: a current perspective on twin studies of schizophrenia. *Am J Psychiatry*. 1983;140:1413–1425.

Kendler KS, Gardner CO Jr. The risk for psychiatric disorders in relatives of schizophrenic and control probands: a comparison of three independent studies. *Psychol Med*. 1997;27:411–419.

Kirov G, Gumus D, Chen W, et al. Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet*. 2008;17(3):458–465.

Kirov G, Pocklington AJ, Holmans P, et al. De novo CNV analysis implicates specific abnormalities of postsynaptic signaling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry*. 2012;17:142–153.

Kirov G, Rees E, Walters JTR, et al. The penetrance of copy number variations for schizophrenia and developmental delay. *Biol Psychiatry*. 2014;75(5):378–385.

Lichtenstein P, Yip BH, Björk C, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet*. 2009;373(9659):234–239.

McGuffin P, Owen MJ, O'Donovan M, Thapar A, Gottesman II. *Seminars in Psychiatric Genetics*. London: Gaskell; 1994.

Owen MJ. New approaches to psychiatric diagnostic classification. *Neuron*. 2014;84(3):564–571.

Purcell SM, Moran JL, Fromer M, et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*. 2014;506(7487):185–190.

Rees E, Kendall K, Pardinas AF, et al. Analysis of intellectual disability copy number variants for association with schizophrenia. *JAMA Pschyiatry*. 2016;73(9):963–969.

Rees E, Moskvina V, Owen MJ, O'Donovan MC, Kirov G. De novo rates and selection of schizophrenia-associated copy number variants. *Biol Psychiatry*. 2011;70(12):1109–1114.

Rees E, Walters JT, Georgieva L, et al. Analysis of copy number variations at 15 schizophrenia-associated loci. *Br J Psychiatry*. 2014;204(2):108–114.

Rujescu D, Ingason A, Cichon S, et al. Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum Mol Genet*. 2009;18(5):988–996.

Schizophrenia Working Group of the Psychiatric Genetics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511:421–427.

Stefansson H, Meyer-Lindenberg A, Steinberg S, et al. CNVs conferring risk of autism of schizophrenia affect cognition in controls. *Nature*. 2014;505:361–366.

Südhof TC, Malenka RC. Understanding synapses: past, present, and future. *Neuron*. 2008;60(3):469–476.

Sullivan PF, Daly MJ, O'Donovan MC. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet*. 2012;13(8):537–551.

Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait. Evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry*. 2003;60:1187–1192.

Szatkiewicz JP, O'Dushlaine C, Chen G, et al. Copy number variation in schizophrenia in Sweden. *Mol Psychiatry*. 2014;19(7):762–773.

Tienari P, Wynne LC, Laksy K, Moring J, Nieminen P. Genetic boundaries of the schizophrenia spectrum: evidence from the Finnish Adoptive Family Study of Schizophrenia. *Am J*

Psychiatry. 2003;160:1587–1594.

▲ 12.5 The Clinical Epidemiology of Schizophrenia

JIM VAN OS, PH.D., AND ULRICH REININGHAUS, PH.D.

WHAT IS CLINICAL EPIDEMIOLOGY?

Clinical epidemiology is the basic science of preventive medicine. It examines the variation in susceptibility, onset, course, and outcome of disease and the reasons behind this variation, within defined populations. It is the science of making predictions about the emergence (onset) and persistence (prognosis) of disease, through the systematic identification and evaluation of inherited and acquired characteristics, attributes, behaviors, and exposures, known as risk factors (determinants). These risk factors increase an individual's likelihood of developing (1) a precursor of the clinical disorder (a condition or state preceding the recognized onset of a disease), (2) a clinical disease, or (3) a particular course and outcome trajectory for the disease process. Determinants rarely act in isolation to increase risk, rather their effects are cumulative. Such complementary actions can be additive (the linear sum of the individual effects) or synergistic (interaction of effects) when individual risk factor effects are modified by particular environmental conditions, biological stochastic events, or with other known risk factors. Furthermore, there is no reason to assume that the risk factors for the emergence of a clinical disorder and those influencing persistence are congruent. Evidence that the same risk factor is implicated in both onset and persistence of a disease (discrete major effect) could aid in the understanding of etiological mechanisms. However, lack of discrete effects suggests that the course of a disease is not predetermined and is potentially modifiable after onset.

Clinical epidemiology attempts to harness the heterogeneity observed in risk factor profiles, course, and outcome and uses this variation to generate hypotheses and develop empirically based conceptual frameworks (models) of disease. Identifying the risk factors operating prior to the onset of disease is fundamental to effective evidence-based primary prevention, while the risk factors predicting course and outcome help identify patients at higher risk of poor outcome, so guiding the development of secondary and tertiary preventative strategies and interventions tailored for specific individuals in clinical settings.

Heterogeneity is the hallmark of schizophrenia, with manifest variation in risk factors for onset, symptomatology, course, and outcome. A central tenet of modern risk factor epidemiology is that virtually all diseases are a consequence of the interplay between genetics and environmental characteristics; therefore, recent studies investigate how differences in

genetic endowment moderate the effect of environmental risk factors (gene–environment interactions).

RISK FACTORS FOR THE ONSET OF SCHIZOPHRENIA

Genetic Risk: Indirect Evidence from Epidemiological Studies

Family Studies. Family studies compare the morbidity risk (the proportion of affected to unaffected relatives, usually adjusted for variation in nonaffected participants' time at risk) for probands with schizophrenia to the morbidity risk in the general population or demographically matched control groups. Results from a systematic review of the studies published during the period 1980 to 2003 found the first-degree relatives of probands to have a higher morbidity risk for schizophrenia than the relatives of controls. Estimates ranged from 2 to 9 percent (in studies that adjusted for time at risk using Weinberg, Strömgren, or Kaplan–Meier methods) compared to a morbidity risk in the relatives of controls at about 0.5 percent. The effect varied according to the diagnostic classification employed in the study, with broader definitions of schizophrenia, such as Research Diagnostic Criteria, tending to show larger estimated effects than the narrower definitions employed in post–third editions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM). Similarly, studies using control groups unscreened for psychiatric disorder showed more conservative estimates. There was also evidence of clinical heterogeneity, with the relatives of early-onset cases having a higher morbidity risk than the relatives of late-onset cases. Further studies have examined familiar aggregation for categorical subtypes, symptom dimensions derived from factor analysis, and individual symptoms. The findings are inconsistent and show significant methodological heterogeneity; there is nevertheless some support for higher morbidity risk for schizophrenia associated with negative symptoms and the disorganization dimension/hebephrenic subtypes. Most family studies have investigated morbidity risk in first-degree relatives; the few studies of second- and third-degree relatives suggest an increase in risk with increasing relationship closeness, the structural form of which may not be linear.

Adoption Studies. Most children inherit their genetic characteristics as well as their early rearing environment from their biological parents. Adoption provides an opportunity to disentangle these genetic and rearing effects. Evidence from adoption studies have primarily come from a small group of studies. An adoptee's family study, using the Danish national linked data set, compared rates of schizophrenia in the adoptive and biological families of adoptees with schizophrenia. Results reported over the past 30 years (using different diagnostic criteria) consistently show the biological relatives to be at a higher risk of schizophrenia than the adoptive

relatives, who have a similar estimated risk to the rest of the general population. The second strategy compares the rates of schizophrenia in adopted children for parents with and without a diagnosis of schizophrenia. Early studies using this methodological approach consistently demonstrated higher rates of schizophrenia in the adopted offspring of affected mothers, compared to the rates in the offspring of the unaffected controls. A recent Finnish study has improved on these earlier methodologies by identifying larger samples, using standardized interviews with home assessments and DSM operationalized diagnostic criteria. Its results support the early findings, showing a morbidity risk in the adopted offspring of mothers with schizophrenia of 5.1 percent, as compared with 1.6 percent in the adopted offspring of control mothers, a finding that has been replicated in a recent, large Swedish registry study.

Twin Studies. There are some studies that compare the concordance rates for monozygotic (identical) twins (assumed to inherit all their genes in common) with those of dizygotic (nonidentical) twins (who on average share around 50 percent of their genes). Such studies further assume that twin pairs are equally exposed to environmental risk factors for schizophrenia. Statistical combination of the published studies that used DSM-III-R criteria to define caseness estimated pooled concordance rates of 50 percent in monozygotic twins and 4.1 percent in the dizygotic pairs; the pooled estimates of concordance for studies using different diagnostic classification systems are actually very similar. The concordance rate for monozygotic twins was higher in twins with an early rather than a late onset of schizophrenia. Twin studies investigating classical schizophrenia subtypes show higher monozygotic concordance rates for the hebephrenic subtype compared to the paranoid type. Combination twin and family studies find the risk for schizophrenia in the offspring of the identical and unaffected twin to be the same as that seen for the offspring of the affected twin.

SPECIFICITY OF HERITABILITY

Relationship to Subclinical-Psychoses Phenotypes

Clinicians have long recognized that some close relatives of patients with schizophrenia, although never frankly psychotic, display symptoms and signs similar to the pathological experiences seen in the clinical disorder. Some are attenuated forms of the positive/disorganized symptoms of schizophrenia (odd behavior, odd speech, muddled thinking, loose associations, ideas of reference, magical thinking, suspiciousness, and illusions), while others resemble the negative symptoms (social withdrawal, aloofness, guardedness, poor rapport). The DSM classification system has combined these quasi-psychotic (schizotypic) features under the diagnostic

category of schizotypal personality disorder. As schizotypal, schizoid, and paranoid personality disorders are found to co-occur, they have been grouped together under the rubric Cluster A personality disorder within the manual. Family studies measuring lifetime prevalence for DSM Cluster A personality disorders (some studies have also included avoidant personality disorder) in first-degree relatives of patients with schizophrenia (and other psychotic disorders including *psychotic* bipolar disorder) find rates in the relatives of the probands (range: lifetime prevalence of 2.8 to 8.2 percent) that are two to four times the rates found in the control families. The significant heterogeneity in estimated lifetime prevalence is primarily due to methodological differences. Adoption and twin studies confirm this familiar aggregation. The familiar coaggregation of pathological and attenuated schizotypal symptoms suggests that there may be a single continuum of liability across the schizotypic features and pathological psychotic symptoms.

Relationship to Other Psychotic Disorders and Bipolar Disorder

Family and adoption studies of schizophrenia have generally shown an excess of risk in biological relatives, not only for schizophrenia but also for a broad range of psychotic disorders, such as schizophreniform, schizoaffective, delusional, and atypical delusional disorders. Less accepted is the possible familiar coaggregation of bipolar disorder and schizophrenia in families, stemming primarily from the limited and conflicting data available from family studies of bipolar disorder, which have been interpreted to support the disorder breeding true. However, schizoaffective disorder occurs at similarly increased rates in the families of probands with schizophrenia and bipolar disorder, and family studies of schizoaffective disorder demonstrate increased risk for both schizophrenia and bipolar disorder, suggesting at least some genetic nonindependence. This is further supported by a large, recent registry study of Swedish families, reporting evidence that schizophrenia and bipolar disorder partly share common genetic determinants. However, these studies do not help to elucidate the nature of this shared liability. A possible explanation would be an underlying continuous common liability for all psychotic disorders. A twin study from Finland tested the hypothesis that affective disorders and other psychotic disorders represented an intermediate threshold on a liability continuum shared with schizophrenia. The results did not support a single continuum liability model, although the investigators did not compare alternative hybrid models. However, a further study investigated both common liability (shared genetic effects) and disorder specific liability models, using the Maudsley twin data set. The researchers relaxed the normal hierarchical diagnostic approach, allowing within person comorbidity of schizophrenia, schizoaffective and manic syndromes. Independent and common pathway models were used to compare the

associations of the syndrome pairs and found that for schizophrenia and mania, the model encompassing both shared (common) and syndrome-specific genetic effects fits the data best, while for schizoaffective disorder, the shared genetic effect model was sufficient to explain the associations. These results suggest that transmission may operate through a set of susceptibility genes specific to schizophrenia, and another set specific to bipolar disorder, while a third set of shared (common) effects acts across the three syndromes. This is complemented by emerging evidence on shared intermediate and clinical phenotypes of schizophrenia, schizoaffective and bipolar psychotic disorders. However, interpretation of these studies must be tentative, findings are preliminary and somewhat crude (based on additive and static models of genetic liability), and further replication is required. These codicils notwithstanding, the evidence points toward a model of liability, which incorporates a common (shared) genetic susceptibility for psychosis and concomitant separate schizophrenia-specific liability (see Section 12.4 on the molecular genetics of schizophrenia). There have been substantially fewer studies examining the risk of nonaffective and nonpsychotic Axis I disorders in the relatives of probands with schizophrenia; in a recent population-based country-wide study in Denmark, however, almost any psychiatric disorder among first-degree relatives increased the individual's risk of schizophrenia. The population attributable risk associated with psychiatric family history in general was nearly 30 percent, whereas family histories including schizophrenia accounted for only 6 percent.

GENETIC MODEL OF LIABILITY FOR SCHIZOPHRENIA

The family, adoption, and twin studies provide irrevocable evidence for the involvement of a major genetic component in the liability for schizophrenia. However, the familiar distributions are inconsistent with simple Mendelian modes of inheritance (single major locus) and are more compatible with a polygenetic mode of inheritance. One possible model is the *polygenetic threshold model,* which assumes that the risk of schizophrenia is due to additive effects of multiple small genes at different loci. However, the polygenetic threshold model does not adequately account for the exponential rise in risk of schizophrenia observed with increasing familiar closeness, which is more consistent with (at least some) epistatic (gene–gene) interaction. Nonetheless, the monozygotic/dizygotic twin correlations are significantly lower than would be expected with a full epistatic model. While the initial surge for the molecular genetic basis of schizophrenia was characterized by slow progress and methodological concerns, recent years have seen rapid advances through large-scale collaboration in GWAS, which have generated replicated findings on a number of common risk alleles. However, the common variants identified to date explain not all of the genetic risk of schizophrenia, and a number of

common risk alleles (with small effects) remain to be determined. Furthermore, heritability estimates of the overall contribution of common genetic variants based on molecular genetic data are considerably smaller (i.e., 23 to 33 percent) than heritability estimates from twin studies (i.e., 81 percent). Therefore, it is more likely that both genetic and environmental factors contribute to the development of schizophrenia. The mechanisms by which these factors interact to influence risk remain unclear but will probably involve the complex interplay of both additive and interacting genes, environmental factors, gene–environment interactions (genetic moderation of environmental factors and environmental moderation of gene expression), and neural, cognitive, and affective mechanisms underlying the interplay of genetic and environmental factors. These are likely to be continuously distributed within the population. This model of risk may vary within and across individuals, families, or clinically defined subgroups of the disease, depending on the specific risk factors active within the person at different stages of development. If this probabilistic model of risk transmission is accepted, then environmental risk factors appear to be necessary complementary components in the development of schizophrenia in many if not all cases.

Gene–Environment Interaction

Traditional epidemiological risk factor methodologies measure the linear association of environmental exposures to risk for onset and persistence of schizophrenia. Such studies demonstrate that people respond differently to the same environmental exposure. Some individuals seem to have a low risk of developing the disorder as a result of an environmental exposure, while others are more vulnerable; in some circumstances, persons with increased liability for a disorder seem to have a higher exposure frequency to noxious environmental agents than the general population. Variation in genotype may, in part, account for this heterogeneity. There are two most often studied gene–environment relationships. *Gene–environment correlation* occurs when a genotype could increase the likelihood of a person becoming exposed to an environmental risk factor. For example, biological parents not only determine a child's genotype but also provide his or her early rearing environment, or a child's temperament (at least partially genetically determined) influences how he or she selects and/or shapes his or her environment. *Gene–environment interaction* (*synergism*) (G × E) posits that instead of inheriting a disease state per se, a person actually inherits a set of genetic susceptibility factors that act to moderate the effects of (causal) environmental exposures. If this assumption is correct, integrated epidemiological approaches, which examine the joint influence of genes and environment on the risk for schizophrenia, will improve the accuracy and precision of the measured risk estimates and the predictive power of environmental risk factors.

While long ignored in molecular genetic research, and still an emerging field, the beginning of this century has seen a small number of G × E studies on candidate genes in schizophrenia. These studies have tested individual, a priori selected SNPs, with very few attempts at replication and limited evidence on the potential mechanisms underlying G × E in schizophrenia. Indeed, there remain a number of conceptual and methodological challenges in contemporary molecular G × E research, which are currently sought to be addressed in integrated, large-scale investigations such as the "The European Network of National Networks studying Gene-Environment Interactions in Schizophrenia" (EU-GEI).

Environmental Risk Factors for the Onset of Schizophrenia

A number of environmental factors are statistically associated with an increased risk for the onset of schizophrenia. Evidence of association, however, does not equate to causality. Decisions regarding causation require deductive judgments. In order to aid this logic, causal criteria have been developed (Table 12.5–1). This section uses the most frequently applied criteria to structure the review of the current evidence for the risk factors for the onset of schizophrenia. However, it is important to stress that these criteria should not be used in a mechanistic fashion, as all have exceptions and are useful at best as a framework for thought.

Table 12.5–1.
Causal Criteria for the Onset of Schizophrenia

1. **Strength of the association**
   The argument states that strong associations are more likely to be causal, because if they could be explained by another factor, the effect of the confounder would have to be even stronger than the observed association. While weaker associations are more likely due to noise and bias, it is important to remember that measures of association (e.g., relative risk for odds ratios) are a characteristic of the sample under study and depend on the relative prevalence of other causal factors.
2. **Dose–response relationship (biological gradient)**
   Defines a unidirectional trend in disease frequency with increasing levels of exposure; however, this is neither necessary nor sufficient for causality as the presence of an unadjusted confounder will also show a gradient. Furthermore, some causal relationships show single jumps (thresholds).
3. **Temporal association**
   This is a sine qua non for causality. Requires that for a factor to be considered causal, it must precede the disease. This is irrefutable; however, it does not provide evidence for or against causation when the cause is present prior to onset of the disorder.
4. **Consistency of findings**
   Is the repeated observation of an association in different populations under different circumstances? However, lack of consistency does not rule out a causal association because some risk factors only produce effects under certain situations in which a complement of causes is present (or have been present). Consistency can really be assessed only after the relevant details of the causal mechanism are understood.
5. **Specificity of effect**
   This is still emphasized in psychiatric epidemiology, perhaps due to the continual nosological

debates regarding classification of disorders. However, there are no good grounds to refute a causal interpretation of an exposure because it is related to multiple effects or disease outcomes.

6. **Biological plausibility**
   Plausibility is important; however, biologically plausible hypotheses can be constructed that are later found to be erroneous, and vice versa causal associations can be dismissed due to current knowledge or belief systems.
7. **Coherence**
   Related to biological plausibility and, therefore, similar problems, the association should not conflict with what is known of the natural history and biology of the disease.
8. **Experimental evidence**
   Commentators are unclear what is meant by experimental evidence, whether it was evidence coming from laboratory work or human experiments such as removal of an exposure. However, this is a test for causality rather than a criterion.
9. **Analogy**
   Analogy may provide a source for more elaborate hypotheses about the associations but does not falsify a hypothesis.

Critique of the causal criterion

With the exception of temporality, none of the aforementioned situations can be viewed as a necessary or sufficient criteria for determining whether an observed association is causal. Causal inference cannot be deduced with certainty.

Risk Factors Operating during Early Development

Paternal Age. Recently, several large population-based cohort studies from Israel, the United States, Denmark, and Sweden have provided strong evidence linking advanced paternal age to the risk of schizophrenia in offspring. The association is strong and independent and the findings are consistent. There is a higher risk of schizophrenia (around three to four times) in the offspring of fathers who are older than 50 years, at the time of conception, than that in the offspring of fathers in their early 20s. The small variation of estimated effect size across studies reflects methodological differences, with higher associations found in studies of patients with narrowly defined schizophrenia. Nearly all studies found a dose–response relationship, with the risk increasing across consecutive age bands. The functional form of this association has not been established, but there is some evidence for a slightly increased risk of schizophrenia in younger fathers (younger than 20 years) with J-shaped associations having been suggested. Temporal association exposure clearly predates the diagnosis of schizophrenia. All the cohort studies have estimated broadly similar strengths of association, and while adjustments for known risk factors have been made, there may be residual confounding by a factor currently unidentified, which directly increases the risk of schizophrenia and is also associated with paternal age. The specificity of effect association is stronger for narrowly defined schizophrenia than for nonaffective psychosis, supporting relative specificity for schizophrenia within the spectrum of psychosis. There is no evidence for an association between age of first admission and paternal age. In evidence of gene–environment interactions, there is a stronger association between paternal age and

schizophrenia in people *without* a family history. This observed pattern of effect lends support for two genetic theories. The first and most likely is that advancing paternal age results in accumulation of de novo mutations in the germ cells of older fathers, or second, advancing paternal age interferes with the deoxyribonucleic acid (DNA)-methylation process of gene expression. Nevertheless, this finding needs further investigation as a large Danish study suggested that the association with higher paternal age may be due simply to selection into late fatherhood accompanying a predisposition to schizophrenia.

Season of Birth. Winter birth in people who later develop schizophrenia is a robust epidemiological finding, at least in the northern hemisphere. It is likely to be a proxy indicator for some seasonally fluctuating environmental factor. The most popular hypotheses relate to seasonal variation in exposure to intrauterine viral infections around the time of birth, or variation in light, temperature/weather, or external toxins. There is a statistically significant but small effect associated with winter birth and later development of schizophrenia; a meta-analysis of northern hemisphere studies estimated a pooled odds ratio of 1.07 (95 percent confidence interval [CI], 1.05 to 1.08). For a dose–response relationship, the effect size increases with latitude (although there are fewer studies from the lower latitudes of the northern hemisphere in the pooled analyses) until about 50 degrees latitude, where it then starts to fall, suggesting a nonlinear inverted U-shaped association, which may in part be explained by methodological heterogeneity across studies. For temporal association, exposure clearly predates diagnosis of schizophrenia. The studies from northern hemisphere countries consistently show excess winter births in people with schizophrenia; however, the effect seems to be greater for those born in urban than for those in rural settings. A meta-analysis of studies carried out in the southern hemisphere did not find an effect for season of birth (pooled odds ratio, 1.04; 95 percent CI, 0.99 to 1.08). Studies examining season of birth association to other psychiatric conditions are small and their findings inconsistent. The largest and most robust study, carried out in Denmark, did not find an association with season of birth and development of bipolar disorder, suggesting some specificity of effect for schizophrenia (spectrum) disorder as compared with bipolar disorder. Few studies have looked at age of onset in relation to seasonality of birth in people with schizophrenia. The largest study to date found higher rates of winter births in patients who developed schizophrenia later in life; however, most studies are underpowered to investigate this issue. Little information is currently available on evidence of gene–environment interactions; however, evidence from small studies suggests that further research on interaction between winter birth and underlying genetic liability may be productive.

Pregnancy and Birth Complications. There are a large number of published studies demonstrating some relationship between pregnancy and birth complications and the development of schizophrenia. The literature suggests that pregnancy and birth complications can have a small effect on the risk of later development of schizophrenia. Meta-analytic studies have estimated a pooled odds ratio of 2.0 (95 percent CI, 1.6 to 2.4) for patients with schizophrenia having a history of pregnancy and birth complications compared with nonaffected controls. The term pregnancy and birth complications covers a broad range of pre- and perinatal events, but even a large systematic review of prospective population-based studies (period 1995 to 2001) was unable to identify one individual complication or unifying mechanism. Taken together, findings may implicate an increased risk of hypoxia.

Studies to date have mostly been underpowered to uncover variation in effect size (odds ratio) across groups ordered by degree of exposure to pregnancy and birth complications. One population-based cohort study from Helsinki used continuous measures and found an effect gradient of increasing risk with decreasing birth weight, length at birth, and placental weight (controlled for family history of schizophrenia), and the National Collaborative Perinatal Project found risk increased when multiple hypoxia-associated complications were present. However, it is difficult to interpret these gradients while the underlying mechanism remains elusive. Exposure clearly predates diagnosis of schizophrenia, although fetal growth retardation may be an early indicator of aberrant neurodevelopment. The associations for individual obstetric complications with later development of schizophrenia are mostly negative and inconsistent; however, statistically pooled estimations suggest a modest effect with the broad category of pregnancy and birth complications. There is methodological heterogeneity with a discrepancy in the estimated effects for case–control studies (that generally show stronger associations) and cohort studies. The most likely explanation for this is selection and information bias that may arise in case–control studies.

Pregnancy, birth, and neonatal complications do not act independently of one another; however, current studies are underpowered to detect interaction effects. There have been only a few small studies examining differential effect of pregnancy and birth complications and symptom dimensions or subtypes of schizophrenia. A recent systematic review of pregnancy and birth complications and bipolar disorder estimated the pooled odds ratio for exposure to pregnancy and birth complications and later development of bipolar disorder to be 1.01 (95 percent CI, 0.76 to 1.35) compared with controls, and 0.61 (95 percent CI, 0.39 to 0.95) compared with people who later developed schizophrenia. Furthermore, a Danish national register study did not find fetal growth indicators to be associated with the development of bipolar disorder. This suggests some

specificity for schizophrenia (spectrum) disorder among psychiatric conditions. Indirect evidence of gene–environment interactions comes from a population-based cohort study in Helsinki that found that a history of fetal hypoxia was associated with increased structural brain abnormalities (reduced gray matter and increased ventricular size) among patients with schizophrenia and their nonaffected siblings but not among controls at lower genetic risk for the disorder. Moreover, the effect size was greater for these patients than that for the unaffected siblings, suggesting that while they show some sensitivity to the schizophrenia-promoting effects of hypoxia, there is a quantitative difference in this sensitivity. Similarly, there is evidence from a single study reporting that serious obstetric complications interact with hypoxia-regulated/vascular expression genes to influence schizophrenia risk. Evidence of gene–environment correlation from a meta-analytic study of pregnancy and birth complications in children born to parents with schizophrenia suggests that individuals with schizophrenia have an increased risk of obstetric complications, low birth weight, and poor neonatal condition. The risk is greater for mothers with schizophrenia but is not confined to mothers with onset prior to delivery or to the births of the children who develop schizophrenia themselves later in life. These findings possibly reflect a gene–environment correlation whereby the genetic liability for schizophrenia in the parent increases the likelihood of social adversity that is associated with obstetric complications.

Other Putative Prenatal Risk Factors. Emerging evidence from ecological studies suggests that a number of other prenatal factors may have a weak association with the later development of schizophrenia. A series of studies have found higher rates of schizophrenia in people exposed to the 1957 influenza pandemic during the second trimester of their gestation. This is not a fully replicated finding and remains controversial. A study from New York City found a 10 to 20 percent increased risk for schizophrenia in people who had serologically confirmed prenatal rubella exposure, while further studies have implicated prenatal exposure to toxoplasmosis, poliovirus, and other common respiratory infections. It has been suggested that the effect may be in part due to cytokines and chemokines, which mediate host response to infection. There is some evidence that rhesus incompatibility increases the risk of developing schizophrenia; this may not be through the direct toxic effect of hyperbilirubinemia on the developing brain but rather via maternal–fetal genotype incompatibility effects. There may be higher rates of schizophrenia in the offspring of mothers who experienced significant levels of stress during pregnancy, such as the death of a spouse or living through a natural disaster or military invasions. Nutritional deficiency in pregnancy may also increase the risk of schizophrenia in offspring. A

twofold increase in rates of schizophrenia was found in the children of mothers who were exposed to the Dutch famine during early gestation. These putative environmental risk factors all require further replication.

RISK FACTORS OPERATING DURING CHILDHOOD AND ADOLESCENCE

Urban Birth and Upbringing

More than 20 studies have demonstrated an association between urban living and rates of schizophrenia and nonaffective psychosis. Generally, studies show a twofold increase in risk of schizophrenia in urban as compared to rural settings, as evidenced by a recent meta-analysis. Furthermore, the weight of evidence supports a risk gradient, with unidirectional increasing rates of schizophrenia with increasing levels of urbanicity. There is compelling evidence suggesting that elevated rates of schizophrenia in urban areas are not due to social drift during the prodrome (reverse causation). It is urban exposure prior to the onset of the disorder that seems to be more important rather than the level of exposure at the time of illness onset. A case register study from the Netherlands demonstrated an increased risk of schizophrenia for people born in urban areas, irrespective of where they were living at the time of presentation, while those living in urban areas were at increased risk only if they had also been born within an urban area. Furthermore, a Danish study measured the effect of cumulative urban exposure on risk of schizophrenia and found that the first 15 years of life were the most critical as well as evidence of a dose-response relationship. Analyses from the same cohort further suggest that people conceived (including period of gestational development) in urban areas but reared in less urban areas may be at higher risk of schizophrenia. This phenomenon of increased risk seen when there is a mismatch between in utero developmental environments and environment in adult life is recognized in other chronic complex diseases, such as cardiovascular disease. The findings are consistent with similar effect sizes seen across studies from different geographical areas. The effect is not confounded by other known risk such as pregnancy and birth complications and is present after controlling for sociodemographic variables. Most analyses have examined the association of urban living within the broad psychosis category rather than specific diagnoses. Studies that have compared across diagnostic categories have found the risk to be relatively specific for schizophrenia (spectrum) disorders compared to bipolar disorder. Studies investigating family history and urban exposure support a possible interaction with genetic liability. For example, in the Danish cohort, those *without* a family history of schizophrenia had a much smaller risk of schizophrenia (0.054 percent increased risk, per unit change of urbanicity) than those with a family history (0.22 percent increased risk, per unit change of urbanicity).

COULD GROUP LEVEL SOCIAL PROCESSES ACCOUNT FOR THE URBAN UPBRINGING EFFECT?

Growing up in an urban environment is a proxy environmental risk factor. Recent research has started to examine the social and economic structure of neighborhoods as a possible explanation for this effect. The social environment varies widely and systematically across dimensions of material deprivation, neighborhood organization (social fragmentation), and ethnic composition. Studies have consistently found a relationship (not necessarily linear) between material deprivation and risk of schizophrenia, although study design has generally been cross-sectional, precluding assessment of the direction of the relationship, and could be accounted for by people drifting into more deprived areas before presenting to clinical services. Ecological studies have consistently shown area-level social fragmentation to influence strongly the rates of schizophrenia. A small number of recent studies have examined levels of social cohesion and found neighborhoods richly endowed with stocks of social capital (interpersonal trust, norms of reciprocity, mutual aid) to have reduced rates of psychosis independent of material deprivation. Ethnic density is an area-level characteristic, defined as the percentage of composition (in a particular geographical area) of people with a given ethnicity, and it has been found to be protective. That is, the higher the area's ethnic density, the lower the rates of schizophrenia for the ethnic group under study, independent of area measures of deprivation, a finding that has been consistently reported in a number of studies. Taken together, these reports suggest a possible neighborhood-level social process influencing the risk of schizophrenia, which will require further elaboration and replication.

Migration

Many studies have reported increased rates of schizophrenia in migrants, especially among the second generation born in the new homeland. Increased risk is found among black Caribbean and black African immigrants to the United Kingdom; Surinamese, Dutch Antillean, and Moroccan immigrants to the Netherlands; Australian, African, and Greenland immigrants to Denmark; and Greek immigrants to Belgium. A meta-analysis of incidence studies (published between 1970 and 2003) found a higher risk of schizophrenia in almost all immigrant groups than that in the majority population. Although the pooled effect size suggested a threefold increase in risk, there was evidence of systematic heterogeneity across studies that could not be accounted for by variation in study design. Analyzing first- and second-generation migrants separately found a pooled relative risk of 2.7 (95 percent CI, 2.3 to 3.2) for the first generation and a higher pooled relative risk of 4.5 (95 percent CI, 1.5 to 13.1) for the second generation. There was also a higher risk of schizophrenia in migrants from

developing countries; however, the most conspicuous finding was the increased risk associated with being black, with a pooled relative risk of 4.8 (95 percent CI, 3.7 to 6.2). Migrant status is a binary exposure; one is either a migrant or one is not. Therefore, a dose–response association will be testable only after the underlying mechanism operating to increase the risk is established.

The selective migration of people at increased risk of schizophrenia is possible (reverse causation). However, this does not explain the increased risk apparent in groups who migrated en masse, such as the Surinamese mass migration to the Netherlands. Furthermore, it cannot explain the increased risk found in the second-generation immigrants. Increased risk of schizophrenia in migrant groups is a replicated and consistent finding. The validity of this finding has been intensely debated, however. Increased rates do not appear to be explained by psychiatrists diagnosing schizophrenia differently in migrant groups, differential service use, or underenumeration, which may have hampered some of the earlier studies. Poor obstetric care in migrant mothers has also been muted as an explanation, but there is no evidence to support this view. The few published studies examining migrant status across dimensions or subtypes of schizophrenia have not detected any differential associations between migrant group and the majority population. While there has been only a limited amount of research on candidate explanatory factors of the high rates of schizophrenia in migrant and ethnic minority groups, the evidence reported to date by investigations such as the Aetiology and Ethnicity in Schizophrenia and Other Psychoses (AESOP) study, a large epidemiological incidence and case–control study of first-episode psychosis conducted in the United Kingdom, suggests that it is a broad range of social adverse experiences in both childhood and adulthood that may combine and interact in complex ways to increase risk in some groups. Finally, studies mostly from the United Kingdom have found the risk of schizophrenia to vary across different ethnic groups. The black Caribbean and black African groups in England are at higher risk than the Asian group, as are the Surinamese and Moroccan groups compared to the Turkish group in the Netherlands. The association of effect is not specific to schizophrenia, with rates of nonaffective and bipolar disorder also higher in migrants. For instance, the AESOP study found a sevenfold increase in mania among the black Caribbean and black African group compared to the majority group and a twofold increase in mania among other migrant groups. There is no difference in the age of onset in migrant groups as compared to the majority population. Family studies of schizophrenia have found a differential morbidity risk for schizophrenia in the relatives of UK-born black Caribbean individuals (15.9 percent) compared with the rates in the relatives of white probands (1.8 percent), which suggests a possible gene–environment interaction.

Cannabis Use

There is little doubt that cannabis intoxication is associated with transient psychotic symptoms in some individuals. Furthermore, case–control studies have consistently demonstrated higher rates of cannabis use among patients with schizophrenia as compared with controls. However, controversy remains about whether cannabis is a risk factor for the development of schizophrenia or other psychotic disorders. There are only a few prospective studies examining prior cannabis use and subsequent development of schizophrenia (or other psychotic disorders), but more studies have examined broadly defined psychotic outcomes, such as positive or negative symptoms and their associated disability.

The best evidence for an association between cannabis use in adolescence and subsequent development of schizophrenia comes from a study of Swedish conscripts who were followed up over a 27-year period. Cannabis use prior to conscription was associated with higher rates of schizophrenia, an estimated odds ratio of 1.5 (95 percent CI, 1.1 to 2.0) after adjusting for potential confounders such as disturbed behavior, low intelligence quotient, urban upbringing, and poor social adjustment. Furthermore, a meta-analysis of prospective cohort studies found similarly high rates of broadly defined *psychosis outcome* (either psychotic disorder or psychotic symptoms and associated disability) in people who had at some time used cannabis compared with those who had never used, with a pooled odds ratio of 1.41 (95 percent CI, 1.20 to 1.65). There are a number of small studies examining association of cannabis use with symptom dimensions or subtypes of schizophrenia; most have found no association with negative symptoms and cannabis exposure. In the Swedish conscript study, the effect size increased with frequency of cannabis use, and this is a consistent finding in the studies that used the broader definition of psychosis outcome. It is possible that individuals start using cannabis in the prodrome of schizophrenia to reduce their distressing symptoms (reverse causation), but this cannot be fully discounted. However, investigators have generally screened and excluded people who may be experiencing early psychotic symptoms at the start of the prospective studies.

Studies examining cannabis use and the subsequent development of other psychiatric conditions are small, and their findings are inconsistent. The evidence to date may suggest some specificity of effect for schizophrenia and nonaffective psychosis; however, it is too early to tell. The association with prior cannabis use and subsequent psychosis is a consistent finding across studies. Adjusting for other drug abuse, known risk factors of schizophrenia and sociodemographic variables attenuate the effect, but they remain statistically significant. A family study of adolescents with acute psychosis found a 10-fold increase in the morbidity

risk for schizophrenia in relatives of probands who tested positive for cannabis use, compared to the relatives of probands who tested negative for cannabis. Furthermore, there is tentative evidence that specific polymorphisms may moderate the association between cannabis and psychosis outcomes, but more work is required, for example, using polygenic risk scores based on the Psychiatric Genomics Consortium case–control analyses.

STRESSFUL LIFE EVENTS AND EARLY CHILDHOOD TRAUMA

Many studies report an excess of stressful life events in the few weeks prior to the onset of psychotic and affective disorders. A recent systematic review and meta-analysis of studies published during the period 1968 to 2012 suggests a more than threefold increased risk of psychosis in those exposed to adult life events. However, the authors noted several methodological limitations of included studies such as small sample size and the use of checklist measures. Although these life events may precipitate the disorder, there is some evidence to suggest that they reflect a concatenation of stressful life events over a more sustained period before the onset of psychosis. However, it is difficult to establish whether increased stress is a cause or a consequence in the development of psychotic disorder. This is difficult to disentangle with traditional risk factor methodologies; gene–environmental interaction designs may be a more fruitful way of exploring the effect of stress.

Early childhood trauma studies describe a range of severe adverse experiences including sexual, physical, and emotional abuse, and neglect. There are now a modest number of population-based studies that were recently combined into a meta-analysis (covering the period from 1980 to 2011), suggesting that the risk of psychotic experiences and psychotic disorder is increased almost threefold in those exposed to early childhood trauma. Whether or not childhood trauma is a risk factor for schizophrenia remains an intensely debated topic; recent treatment studies nevertheless show that specific treatments targeting PTSD in schizophrenia are safe and effective.

PREMORBID VULNERABILITY INDICATORS/MARKERS

Premorbid Indicators: Early Developmental Abnormalities

As a group, individuals who develop schizophrenia as adults are more likely to manifest subtle motor/language, cognitive abnormalities during childhood and adolescence than individuals who do not develop schizophrenia. There does not appear to be any association between developmental abnormalities and prenatal insults, suggesting that they may be indicators of genetic liability. A series of large birth cohort studies have consistently found evidence for delays in attainment of developmental

milestones for children, who later develop schizophrenia. Abnormalities in motor milestones are most conspicuous before the age of 2 years, while receptive language abnormalities are most prominent between the ages of 5 and 15 years. The relationship between the age of acquiring motor/language milestones and risk of schizophrenia appears to be linear, making a possible developmental subgroup unlikely. Furthermore, severe motor/language abnormalities tend to predict an earlier age of onset for schizophrenia. A study from Philadelphia found increased rates of motor and language abnormalities, not only in the children who later developed schizophrenia but also in their siblings who remained unaffected, suggesting that manifest motor abnormalities may be markers of underlying genetic liability. The early emerging motor and language deficits appear to be relatively specific for schizophrenia (or spectrum disorder). As a group, children who later develop schizophrenia score lower on standardized measures of intelligence than their unaffected counterparts throughout childhood. However, the difference widens around the ages of about 4 to 7 years. Studies that compared affected and unaffected siblings found that the children, who later develop schizophrenia, had a lower intelligence quotient than their unaffected sibling. This deviance in cognitive impairment seems to be relatively specific for schizophrenia (spectrum) disorder and is not present in bipolar disorder. Twin and family studies have shown that the alterations in adult cognitive performance seen in patients and their first-degree relatives are markers for schizophrenia genetic liability. It is likely that the developmental cognitive alterations similarly reflect the influence of genetic factors associated with schizophrenia.

As a group, the children who later develop schizophrenia are more likely to show signs of subtle social maladjustment, for example, poor relationships with other children, social isolation, lone play, and social anxiety as compared to their peers. However, these subtle social markers are common among children who later develop a broad range of different psychiatric disorders. Although developmental abnormalities are common in children who later develop schizophrenia, many of them have normal developmental histories, which results in low sensitivity (the ability to predict children who will later develop psychosis correctly). Furthermore, they are relatively prevalent in the general population (estimated at around 15 percent), which consequently reduces their specificity (the ability to correctly predict children who will not develop psychosis) and in turn makes them poor predictors of future schizophrenia (positive predictive power estimates range from 3 to 5 percent). Developmental abnormalities may be most useful as a possible measurable indicator of genetic liability for schizophrenia.

Premorbid Indicators: Psychopathology/Vulnerable Mental States

Childhood Psychiatric Disorders. The Dunedin Multidisciplinary Health and Development study is a longitudinal follow-up of a representative birth cohort of 1,037 children born in Dunedin, New Zealand, during the year 1972 to 1973. Its strength lies in the fact that it has sequentially assessed not only childhood developmental indices but also serially collected self-reported psychopathology (ages 11, 13, 15, 18, and 21 years), with further psychiatric interviews carried out at ages 11 and 26 years. Investigators first examined diagnoses given to individuals across the follow-up period and found that more than half of those who had been given a major DSM-IV diagnosis at the age of 26 years had also been given a diagnosis at the age of 11 to 15 years, and three-quarters had been given a diagnosis before the age of 18 years. Generally, there was child-to-adulthood syndrome continuity found for anxiety, depression, and antisocial disorders. However, no such syndrome continuity was evident for adult schizophreniform disorder, which was preceded by a number of different childhood disorders, including anxiety, depression, conduct/oppositional disorder, and attention-deficit/hyperactivity disorder (ADHD). This group of juvenile diagnoses did not specifically predict adult psychotic disorder but were in fact the harbinger of many different adulthood psychiatric illnesses. Such findings clearly demonstrate significant psychopathology prior to adult-onset psychosis. However, the juvenile categorical diagnoses are demonstrably poor discriminators for what adult psychiatric disorder a child is likely to develop in the future.

Subclinical Psychotic Experiences. Subclinical psychotic experiences predict later development of psychotic disorders. Evidence that subclinical psychotic experiences may be vulnerability markers (or even precursors) for the later development of psychotic disorders comes from the Dunedin Multidisciplinary Health and Development study. In their most recent analyses, psychotic experiences at the age of 11 years were associated with increased rates of psychotic disorder by the age of 38 years. In further analyses, the investigators examined whether self-reported delusional beliefs and hallucinatory experiences at the age of 11 years would predict schizophreniform psychiatric disorder 15 years later (age 26 years). By the age of 11 years, nearly 15 percent of the total cohort reported some psychotic experience. Children were grouped according to the strength/frequency of their psychotic experiences (at the age of 11 years): The majority of children had no psychotic symptoms (control group, n = 654), an intermediate group had weak symptoms (n = 95), and a small proportion had symptoms described as strong (n = 12). At the age of 26 years, 2 percent of the controls had a diagnosis of schizophreniform disorder, compared with 9.5 percent in the weak and 25 percent in the strong symptom group. Early psychotic experiences specifically increased the risk for schizophreniform disorder moderately in the weak symptom

group (odds ratio, 5.1; 95 percent CI, 1.7 to 18.3) and strongly in the strong symptom group (odds ratio, 16.4; 95 percent CI, 3.9 to 67.8), compared to controls. Although 25 percent of the strong group had developed schizophreniform disorder, a further 70 percent (although not fulfilling the full diagnostic criteria for schizophreniform disorder) still reported some psychotic symptoms (persistence of preclinical psychotic experiences) at the age of 26 years and 90 percent had occupational and social difficulties. The strong symptom group also manifested impairments in motor, language, and cognitive ability, suggesting that the pervasive and persistent psychotic experiences may be indicators of an underlying ongoing psychotic process, reflecting increased genetic liability. The weak symptom group, on the contrary, did not have such prominent developmental abnormalities, although it had significant receptive language impairment, suggesting a weaker genetic liability. Replication of this transition over time from subclinical psychotic experiences to full-blown psychotic disorder comes from a longitudinal general population study in the Netherlands (the Netherlands Mental Health Survey and Incidence Study [NEMESIS]). To identify new cases of subclinical psychotic experiences, the cohort was followed for 1 year. Individuals who reported subclinical psychotic experiences at the 1-year assessment, but not at baseline, were considered *incident cases*. Next, a second-wave follow-up (over 2 years) allowed the risk of transition of subclinical to clinical psychotic disorder (both affective and nonaffective psychosis) in the incident case group to be compared with the risk in the group of individuals with no clinical or subclinical psychotic experiences (controls). The incident cases were at significantly higher risk of developing psychotic disorders than the controls (odds ratio, 65.1; 95 percent CI, 19.4 to 218.1), and the positive predictive value over 2 years was 8 percent. There was a dose–response relationship in the association between the number of subclinical psychotic experiences and the transition to clinical psychosis going from a risk of around 30 in individuals who had only one subclinical psychotic experience (odds ratio, 27.3; 95 percent CI, 5.2 to 143.6) to a much higher risk in individuals who had multiple subclinical experiences (odds ratio, 211.2; 95 percent CI, 51.6 to 864.1). In both the Dunedin cohort and the NEMESIS study, transition from subclinical to clinical psychotic disorder was higher in individuals characterized by stronger (persistent) subclinical psychotic experiences. In the NEMESIS study, transition was also partly dependent on the emotional context of the subclinical psychotic experience. Emerging evidence from general population studies shows that subclinical psychotic experiences are relatively common (prevalence of around 8 percent, based on a recent meta-analysis). However, most are brief, self-limiting developmental phenomena. The reason why approximately 15 percent of children have subclinical psychotic experiences during development and the others do not is not fully established, but it may in part reflect the shared genetic

liability for psychosis continuously distributed in the general population. In two large independent general population samples, NEMESIS and the German Early Developmental Stages of Psychopathology Study (EDSP), investigators examined the hypothesis that these relatively common, subclinical developmental psychotic experiences could become abnormally persistent when synergistically combined with known risk factors for onset of psychotic disorders (the study used cannabis, urban upbringing, and developmental trauma). Over a follow-up period of 3 years, the studies found the rates of subclinical psychotic experiences that *persisted* to be relatively low, 26 percent (NEMESIS) and 31 percent (EDSP). However, rates of persistence increased progressively with higher baseline doses of exposure to known risk factors. This suggests that environmental determinants of psychotic disorders may in part operate by driving up the rates of subclinical psychotic experiences, which become persistent in individuals with a genetic liability for psychosis. A dynamic model of psychosis evolution or proneness–persistence–impairment model may best conceptualize this trajectory from developmental expression of psychosis liability (transitory subclinical psychotic symptoms) to more intrusive persistent subclinical psychotic symptoms to significant impairment and dysfunction and the eventual onset of psychotic disorder. Such a model is supported by recent meta-analytic evidence and provides a framework for the further investigation of the ontogenesis of psychotic disorders. The psychosis–proneness–persistence–impairment model of psychosis offers a possible insight into the underlying mechanism for development of psychosis. However, a positive predictive power of around 40 percent for persistent subclinical symptoms to estimate onset of future psychotic disorders is still too low to be useful as a general population screening test.

THE COURSE AND OUTCOME OF SCHIZOPHRENIA

Prognostic Pessimism and Schizophrenia: Does the Evidence Stack Up?

Origins of Prognostic Pessimism. The pervading clinical and societal view of schizophrenia is one of a uniformly "debilitating" and "deteriorating" disorder of poor outcome, the basis of which probably stems from clinically biased observations and a degree of conceptual confusion. For example, a carefully conducted Finnish study showed that the prevalence of the group of related (affective and nonaffective) psychotic disorders is as high as 3.5 percent, of which less than a third had a diagnosis of schizophrenia. As the different psychotic disorders share etiological factors, symptoms, and treatments, schizophrenia may be considered the (relatively) poor outcome subgroup of a general broader psychotic syndrome. As schizophrenia is much more studied than the other psychotic syndromes, the relatively poor outcome of schizophrenia may have affected expectations of psychosis prognosis across the range of

psychosis expression.

The idea that schizophrenia is a deteriorating disorder originates directly from the work of Kraepelin, who systematically investigated the course of psychoses in patients admitted to a large custodial institution in Germany during the late 19th century. He combined a number of clinical syndromes into a single disease entity (dementia praecox) because they appeared to share a poor longitudinal course (estimated permanent cure: range, 2.6 to 16 percent). However, the asylum provided long-term residential care for people with the most severe, complex, and debilitating conditions (i.e., a biased sample, skewed toward the worst outcomes), suggesting that Kraepelin's seminal description may not fully reflect the complete range of outcomes for schizophrenia.

Although most clinicians no longer work in long-term custodial settings, clinical caseloads remain biased toward patients who require sustained contact with psychiatric service and have a continued need for care, rather than the representative sample of all persons ever diagnosed with schizophrenia, therefore perpetuating the clinical illusion of schizophrenia as an *invariably* chronic and deteriorating disorder. During the past century, many studies have attempted to examine the outcome of schizophrenia, yet, despite this effort, the data on longitudinal course remain weak, primarily due to the inherent and complex methodological difficulties associated with carrying out follow-up studies, which are especially sensitive to confounding and bias (Table 12.5–2). There is actually little information regarding the natural history of schizophrenia in the literature, as throughout history, psychosis has tended to result in custodial care or other medical, religious, or cultural interventions, which could potentially modify the longitudinal trajectory of the disorder. From the 1950s, the vast majority of individuals diagnosed with schizophrenia have received antipsychotic (neuroleptic) treatment, so it is unlikely that contemporary studies actually reflect natural history. A recent study from India was able to examine outcomes in individuals who had never received treatment of schizophrenia (a closer approximation to natural history, notwithstanding the potential influence of cultural context). The course for the untreated group was heterogeneous and rather similar to the course of those who had received treatment.

Descriptions of the longitudinal (symptom) pattern of schizophrenia have to date been derived from systematic clinical observation (rather than statistical modeling techniques). Clinical researchers, working during the early part of the 20th century, followed patients with schizophrenia over the long term and noted that the significant variability in the course of treatment could be usefully described by a number of course typologies (European researchers recognizing eight categories, while American observers described 12). These longitudinal patterns (slightly modified) have been incorporated into the current diagnostic classification systems as

additional descriptors. The World Health Organization has described operationalized (symptom) descriptors for course (Table 12.5–3) that are useful and reliable in clinical and research settings. Outcome is a multidimensional construct that at minimum requires description of domains for (1) clinical (symptoms and treatment) and (2) social (e.g., independent living, maintenance of social relationships, and employment). Such domains are likely, in part, to overlap. However, studies examining risk factors for the persistence of schizophrenia and disease–environment interactions should measure symptom course and social domains independently, as the different domains probably do not covary directly but are more likely to be the result of different influences, at least in part (Tables 12.5–4 and 12.5–5).

There is a small body of literature comparing the patterns and frequencies of outcome using different perspectives, but few studies have examined multiple perspectives in the same sample. One such study, carried out in South Verona's community care service, found a distinction between objective indicators and patient-reported outcomes. Patients were especially troubled by persistent unmet social and functional needs. This may have important implications for the clinical management of people with schizophrenia, particularly since psychiatric services currently emphasize the pharmacological and psychological treatments, while de-emphasizing social and functional interventions, yet social unmet needs seem to influence disability over the longitudinal course.

A systematic review of prospectively designed studies examining the outcome of first episode schizophrenia (and other psychosis) published during the period 1966 to 2003 found good outcomes in 42 percent, intermediate outcomes in 35 percent, and poor outcomes in 27 percent of the samples with a diagnosis of schizophrenia. In a recent 10-year follow-up of the AESOP first-episode psychosis cohort, 23 percent of cases did not experience a remission and 13 percent remitted and remained symptom free, with an intermediate group of 44 percent having had at least one remission of 6 months or longer and at least one episode of 6 months or longer. What is more, although recently life expectancy in the general population has risen, there was evidence that the excess mortality, which has been repeatedly reported in people with psychosis, remains high. The absence of a uniform definition for outcome across studies makes direct comparison difficult. This frequency distribution is comparable with the findings from an earlier meta-analysis of long-term follow-up studies (published between the turn of the 20th century until 1998, with follow-up period greater than 1 year). The evidence from recent systematic reviews is summarized in Table 12.5–6.

Table 12.5–2.

Methodological Heterogeneity: Limiting Comparability between Studies of Course and Outcome of Schizophrenia

1. ***Variation in the populations from which patients are selected.*** Many studies have relied on inpatient samples only. However, approximately 20% of patients with schizophrenia are never admitted to a hospital and they may differ from those admitted in important variables associated with outcome.
2. ***Variation in diagnostic criteria.*** Lack of reliability between different classification systems may result in different samples being identified. Also, if broad concepts of schizophrenia are employed, the results tend to be more favorable than studies using narrow definitions.
3. ***Variation in the length of illness before entry into the follow-up study.*** If this varies, then subjects with an admixture of follow-up epochs and levels of chronicity are being examined.
4. ***Variation in length of follow-up.***
5. ***Variability of attrition rates.*** In any follow-up study, a proportion of cases will be lost to follow-up due to death, migration, or refusal, or are untraceable for other reasons. Many published studies have high rates of attrition (greater than 20%); furthermore, if the analyses were done at a time where the researcher was unable to take account for this attrition (using imputation methods or other latent variable methods), loss to follow-up is likely to be related to particular patterns of course and outcome, which will bias the outcome.
6. ***Variation in methods used to assess course and outcome.*** There is no single measure of course and outcome for complex disorders, and qualitative summary terms (e.g., recovery, improvement, or deterioration) are likely to obfuscate intraindividual variability over time and the terms are used inconsistently in the literature. Outcome is multidimensional (at least three domains should be considered, viz., symptom severity, functional impairment, and disability, in social and occupational roles). However, these domains are likely to be related, and few studies have used data reduction techniques to deal with the problems of multiple analyses. Furthermore, few studies examine serial follow-up measures, and a single outcome measure will be less representative of a longitudinal course the longer the follow-up period is.
7. ***Variation in the characteristics of the general population.*** For example, social outcomes (e.g., unemployment) will be influenced by the general unemployment rates in the area.
8. ***Variation in statistical techniques and adjustment of confounding.*** Statistical methods, which deal with multiple comparisons made at different time points, were not available or not utilized at the time of publication of the study. Further studies have often not had the resources to collect adequate data on potential confounding variables when assessing potential predictor (or prognostic indicators).
9. ***Variation in long-term management of schizophrenia.*** Treatment may vary across time and place, which may differentially modify the natural history of the disorder.

Table 12.5–3.
World Health Organization Life Chart Description of Clinical Symptom Course

1. Single psychotic episode, followed by complete remission
2. Single psychotic episode, followed by incomplete remission
3. Two or more psychotic episodes, with complete remission between episodes
4. Two or more episodes, with incomplete remission between episodes
5. Continuous (unremitting) chronic psychosis

Risk Factor for the Persistence of Schizophrenia (5 Years' Follow-Up)

Although there have been considerable advances and concerted research efforts into the determinants of onset of schizophrenia, there has been a

paucity of studies examining the determinants of persistence over the past decade. The body of research evidence reflects this neglect. Many early studies of family risk and outcome of schizophrenia found an association between good outcome in schizophrenia and a family history of affective disorders. However, the findings were inconsistent and at times contradictory. There was significant methodological heterogeneity that could potentially bias the results, such as retrospective design, insufficient adjustment for potential confounders (family size, age, and sex of the relatives), the use of global unidimensional outcome measures, and mixed samples (recent onset and chronic disorders studied together).

There have been few recent studies examining the effect of family risk on the course of schizophrenia. The largest study to date examined outcome and family history in a sample of 253 DSM-III patients with schizophrenia and 723 first-degree relatives. There was a higher morbidity risk of schizophrenia (spectrum) disorder in the relatives of probands with poor outcome disorders, although this did not reach statistical significance. However, there was a (statistically significant) higher morbidity risk of affective disorder in the relatives of probands with good outcome disorder. A study from Camberwell examined the associations between 4-year outcome and familiar loading for a range of psychotic disorders in a sample of 150 consecutively admitted patients with recent onset disorder. Higher familiar loading for psychosis predicted higher rates of persistent negative symptoms (estimated odds ratio of 1.5), more time in hospital, and poorer social functioning at follow-up. The effect of familiar loading was greatest in the group of patients with a diagnosis of schizophrenia compared to other psychotic disorders. These findings suggest poor outcome disorders, with higher levels of negative symptoms that may “breed true” within families. There are relatively few studies examining the relationship of pregnancy and birth complications and schizophrenia. However, taken together, they suggest that patients with schizophrenia who have a history of pregnancy and birth complications have a less remitting course and poorer functional outcome. In addition, they may have a poor response to antipsychotic treatment.

There is strong evidence that childhood cognitive ability is associated with outcome; lower intelligence has been shown to predict unfavorable clinical and functional outcomes at follow-up. Furthermore, studies consistently find subtle social premorbid adjustment to be associated with less remitting course and poorer functional outcome. This group of patients is also more likely to have prominent negative symptoms. There are now a number of relatively small- to medium-term (2 to 5 years) follow-up studies comparing the course and outcome of patients with schizophrenia exposed to cannabis versus nonexposed patients. They are methodologically heterogeneous but some consistency across studies is evident. The cannabis-exposed patients have more positive symptoms (but not negative

symptoms) and the course tends to be relatively unremitting. Poorer outcome in the domains of work and education have been demonstrated, which are most likely reflecting factors that independently increase the likelihood of early cannabis use and later education and labor difficulties. By contrast, psychosis that follows stressful life events tends to have a better outcome. The risk factors for the persistence of schizophrenia have been less intensively and comprehensively studied than the risk factors for the onset of schizophrenia. Further research is required to draw any firm conclusions that risk factors for onset may also be determinants influencing outcome (discrete effect models) but is worthy of further investigation. Discrete effects can be broadly grouped into two groups: (1) high familiar morbidity risk for schizophrenia, obstetric complications, and early developmental abnormalities, which are more strongly associated with a clinical symptom profile characterized by prominent negative symptoms (and the associated neurocognitive markers), insidious onset, unremitting course, and functional deterioration, which are strongly linked to the neurodevelopment model of schizophrenia; (2) familiar morbidity risk of affective disorder, adverse life events, and migrant status, which have less specific association with the schizophrenia syndrome and whose clinical correlates suggest less prominent negative symptoms, acute onset, remitting course, and better functional outcomes, which may be linked to a mechanism of social stress sensitivity.

Table 12.5–4.
Variation in Proportion of Good Outcome due to Variation in Outcome

Study	*N*	Period	Event Rate	Upper Limit	Lower Limit	Definition of Good Symptom Outcome
Muller, 1951 (Germany)	200	30	0.33	0.26	0.40	Recovered or substantially improved at the time of discharge (residual symptoms)
Bland, 1978 (Alberta)	43	14	0.21	0.09	0.33	Symptomatic recovery with no social or intellectual deficit through follow-up
Stephens, 1978 (Baltimore)	349	>5	0.24	0.20	0.29	"Recovered" included patients with symptoms but no social impairment
Ciompi, 1980 (Switzerland)	289	37	0.27	0.22	0.32	Remitted completely or had only mild residual symptoms
Salokangas, 1983 (Finland)	161	8	0.55	0.47	0.62	Complete recovery or occasional mild psychotic symptoms
Rabiner, 1986 (New York)	36	1	0.56	0.39	0.72	Symptomatic remission for at least 3 months prior to assessment
Sartorius, 1986 (Multinational)	1,352	2	0.39	0.36	0.42	One episode with no or minimal symptoms at follow-up
Shepherd, 1989 (United Kingdom)	107	5	0.22	0.15	0.30	Had no relapse during follow-up period
McCreadie, 1989 (Scotland)	49	2	0.37	0.23	0.50	Asymptomatic and functioning adequately at the time of follow-up
Marneros, 1992 (Germany)	355	25	0.07	0.04	0.10	Full remission after first episode
Thara, 1994 (India)	76	10	0.17	0.09	0.26	Complete recovery without relapse during follow-up period
an der Heiden, 1995 (Germany)	56	14	0.25	0.14	0.36	Full recovery over follow-up period
Mason, 1996 (United Kingdom)	58	13	0.52	0.39	0.65	Full remission in 2 years prior to outcome assessment
Wieselgren, 1996 (Sweden)	101	5	0.30	0.21	0.39	Dichotomized variable "good outcome" versus "poor outcome"
Ganev, 1998 (Bulgaria)	55	16	0.38	0.25	0.51	Complete remission in 2 years prior to follow-up assessment
Wiersma, 1998 (Netherlands)	82	15	0.27	0.17	0.36	Complete remission never readmitted
Vazquez–Barquero, 1999 (Spain)	76	10	0.32	0.21	0.42	Symptomatic recovery may have residual symptoms
Stirling, 2003 (United Kingdom)	49	10	0.47	0.33	0.61	Categorical measure Global Assessment of Functioning (GAF) score = 4 at the time of follow-up

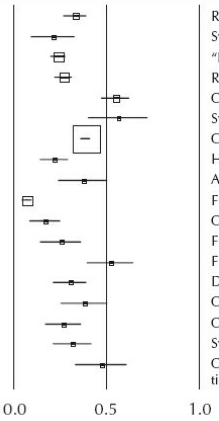

Table 12.5–5.
Variation in Proportion of Poor Outcome due to Variation in Outcome

Study	N	Period	Event Rate	Upper Limit	Lower Limit	Definition of Poor Symptom Outcome
Muller, 1951 (Germany)	200	30	0.41	0.34	0.48	Insidious chronic course, continual need for full-time care
Bland, 1978 (Alberta)	43	14	0.16	0.05	0.27	Severe chronic social and/or intellectual deficit
Stephens, 1978 (Baltimore)	349	>5	0.30	0.25	0.35	"Unimproved" evidence of chronic sustained psychotic symptoms
Ciompi, 1980 (Switzerland)	289	37	0.18	0.14	0.22	Severe chronic phase
Salokangas, 1983 (Finland)	161	8	0.24	0.18	0.31	Continous psychotic symptoms
Rabiner, 1986 (New York)	36	1	0.44	0.28	0.61	Relapsed or in-episode evidence of psychotic symptoms
Sartorius, 1986 (Multinational)	1,352	2	0.40	0.37	0.42	Unremitting psychotic symptoms
Shepherd, 1989 (United Kingdom)	107	5	0.43	0.34	0.52	Remained impaired throughout follow-up
McCreadie, 1989 (Scotland)	49	2	0.39	0.25	0.52	Psychotic symptoms present at 2-year follow-up
Marneros, 1992 (Germany)	355	25	0.17	0.13	0.21	Psychotic symptoms present at the time of follow-up
Thara, 1994 (India)	76	10	0.07	0.01	0.12	Continuous psychotic symptoms
an der Heiden, 1995 (Germany)	56	14	0.59	0.46	0.72	Moderate to severe symptoms at the time of follow-up
Mason, 1996 (United Kingdom)	58	13	0.34	0.22	0.47	Continuous symptoms in 2 years prior to folllow-up
Wieselgren, 1996 (Sweden)	101	5	0.14	0.07	0.21	"Poor outcome" evidence of ongoing persistant psychotic symptoms
Ganev, 1998 (Bulgaria)	55	16	0.45	0.32	0.59	Chronic continuous psychotic symptoms in 2 years prior to assessment
Wiersma, 1998 (Netherlands)	82	15	0.11	0.04	0.18	Chronic continuous psychotic symptoms over full follow-up time
Vazquez–Barquero, 1999 (Spain)	76	10	0.09	0.03	0.16	Chronic continuous psychotic symptoms
Stirling, 2003 (United Kingdom)	49	10	0.06	–0.01	0.13	Global Assessment of Functioning (GAF) score = 1 at follow-up

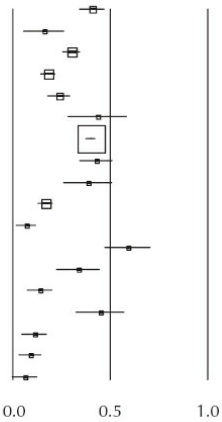

Table 12.5–6.
Summary of Findings from Follow-up Studies on Course and Outcome of Schizophrenia

1. The course of schizophrenia is highly variable both *within* and *between* patients, even after taking account of methodological heterogeneity across studies. There is a broad range of possible course patterns ranging from complete recovery to continuous unremitting psychopathology, cognitive performance, and social functioning. Between such extremes, a substantial number of patients present with multiple episodes of psychosis interspersed with partial remission.
2. Less than half of patients diagnosed with schizophrenia show substantial clinical improvement after follow-up time averaging 6 years. Although these estimates are more optimistic than the earlier reports based on observations in institutional settings, they importantly emphasize the fact that schizophrenia is more often associated with persistent chronic disability that profoundly affects a person's development and quality of life.
3. Examining the trends in course and outcome over the past century revealed substantial gains in favorable outcome from the 1920s until the 1970s, after which time the gradient seemed to reverse (slightly). This biphasic trend may reflect developments in treatment, changes in diagnostic criteria across the decades, and possible sampling bias. Moreover, a secular trend (benign metamorphosis) toward better outcome and the virtual disappearance of the most lethal forms of catatonia was evident prior to the introduction of antipsychotics. The reasons for this trend are unclear, but commentators have speculated that they reflect changes in care and expectations within health service settings, or the development of psychological, social, and rehabilitation approaches to treatment. Unfortunately, the current cannon of outcome studies (as published) does not lend itself to easily disentangling direct disease effects from environmental factors that may influence the persistence of the disorder.
4. Course descriptors vary as a function of the length of follow-up, with the greatest variability being in the first 5 years, after which time the course tends to plateau. However, cross-sectional outcome measures of psychopathology do not differ greatly according to study duration, suggesting that there is no clear pattern of deterioration.
5. On average, patients with a diagnosis of schizophrenia have the poorest outcome, with schizoaffective patients occupying an intermediate position between schizophrenia and affective psychosis; however, diagnostic categories of psychosis poorly predict outcome. Furthermore, course and outcome estimates vary depending on the diagnostic classification used. Patients diagnosed using broadly defined schizophrenia have generally better outcomes than those diagnosed using narrowly defined (post–DSM-III) schizophrenia.
6. There is no reliable set of predictors yet identified for course and outcome. Furthermore, the predictive value of putative prognostic indicators is generally poor and likely to depend on the

duration of illness (stage) and the projection time. Significant methodological differences in definition of predictor variables, adjustments for other known risk factors, differential attrition rates (loss to follow-up), and variation in outcome measurement reduce the comparability of studies on the prediction of prognosis and make exploratory subgroup analyses difficult, despite a large number of published studies. Best estimates to date suggest that all known prognostic factors or psychopathology explain only a small amount of the total variance in outcome (estimates at around 30%, depending on the outcome domain under study), which translates into low predictive power.

TOWARD AN EMPIRICALLY BASED MODEL OF SCHIZOPHRENIA: A WORK IN PROGRESS

Understanding cause is fundamental to the development of scientific knowledge and the development of empirically based models of disease. Modern medicine has developed a deductive system for evaluating causation based on Koch's postulates and Hill's criteria (and their elaborations). The integrated interpretation of epidemiological, clinical, and experimental research aided by such causal criteria has resulted in the successful identification of important risk factors for some leading causes of morbidity and mortality (e.g., cardiovascular disease), some forms of cancer, and infectious diseases. Why has this type of research, which has been reasonably successful for other diseases, been less successful for schizophrenia? At least three (interrelated) factors may be impeding progress: (1) The lack of a sufficiently validated operational definition for schizophrenia (caseness definition). (2) The disorder has a complex time course with a broad range of age of onset. The most common course pattern is episodic (single or multiple), with delusions and hallucinations being prominent symptoms; however, mounting evidence from population studies suggests that these symptoms are also relatively prevalent in the general population. Furthermore, risk factors for schizophrenia probably operate throughout the life course, perhaps even prior to birth. (3) Classical epidemiological strategies assume a static unidirectional cause–effect linear association, which may be of limited usefulness if psychosis is better conceptualized as a dynamic transitional process of small iterations of phenotypic change, mediated by the complex interplay and coevolution of environmental and biological risk factors that are continuously distributed in the general population. While the most widely accepted contemporary model of schizophrenia is based on a neurodevelopmental process, in recent years, there has been a substantial move toward an integrated model of schizophrenia, based on consistent and strong evidence that implicates socioenvironmental risk factors in the development of schizophrenia.

Neurodevelopmental Model of Schizophrenia

The broad neurodevelopmental hypothesis posits that schizophrenia is the behavioral outcome of subtle deviances in early brain development, the full effect of which is not manifest until adolescence or early childhood. There

is convincing evidence for the broad neurodevelopmental model, in agreement with the epidemiological findings for prenatal risk factors and developmental abnormalities reviewed in this section. Children who later develop schizophrenia also have higher rates of minor physical anomalies, indicative of subtle disruption of ectoderm development (e.g., craniofacial and dermatoglyphic anomalies and neurological soft signs). Further evidence is the absence of inflammatory reactions in the brains of people with a diagnosis of schizophrenia. Although earlier histological studies implicated abnormal neuronal migration during corticogenesis, this was not been borne out with further investigation. However, reduced neuronal size, dendritic arborization, and spine density in the cortex and hippocampus appear to be consistent findings across studies. Such findings are important and support a neurodevelopmental model when put in the context of a burgeoning realization that ultrastructural refinement at the synapse is shaped by patterns of environmental input that occur across the life span and seems to have functional significance. Such plasticity in the connectivity of neurons is thought to underlie new learning and the development of new schemata and mental skills. There is generally a long latency period between exposure (before adulthood) to known risk factors and onset to schizophrenia. Such a delayed effect is consistent with a causal model in which the environmental risk factor increases a person's liability, but the full manifestation of the disorder requires the addition or interaction of other complementary (gene or environmental) causal components.

Integrated Model of Schizophrenia

The strong and consistent evidence that implicates socioenvironmental risk factors in the etiology of schizophrenia reviewed in this section has recently led investigators to posit an integrated model of schizophrenia. In this model, the neurodevelopmental model is extended to, and combined with, sociodevelopmental, cognitive, neurobiological, G × E, and epigenetic hypotheses. Given the variation in incidence and expression of schizophrenia across different social contexts, this model introduces a level of explanation focusing on the impact of the environment, in interaction with genetic risk, on cognitive schemata as well as on brain neurobiological mechanisms across the life course. So, for example, exposure to trauma during childhood may lead to the development of negative schematic beliefs and, thereby, predispose individuals to a paranoid way of thinking, which, when coupled with a sensitized dopamine system, may increase liability to psychotic disorder. What is more, it has been hypothesized that susceptibility genes for schizophrenia may have specific developmental correlates that influence brain development across the life course, and it is possible that this variable expression that forms the basis for the widespread subtle developmental delays and subclinical psychotic

experiences can be seen years before the onset of schizophrenia and later emergence of the full-blown disorder. Moreover, it could potentially help explain the broad variation in age of onset of schizophrenia. Epigenetic hypothesis suggests that the onset of schizophrenia is influenced not only by the additive or interactive action of static gene and environmental factors but also via epigenetic dysregulation of gene activity. This model fits well with the familiar (non-Mendelian) features of schizophrenia, and it is likely that epigenetic strategies will develop the underlying mechanisms of the phenotypic differences in monozygotic twins discordant for schizophrenia. Unlike DNA sequencing that usually remains stable over time, epigenetics refers to the regulation of gene expressions that are controlled by heritable but potentially reversible changes in DNA methylation and chromatin structure. Epigenetic aberrations (epimutations) may originate from three sources acting individually or in combination: (1) inherited through the germline; (2) influence of environmental factors acting across different developmental stages; or (3) generated via biological stochastic events. Such epigenetic modification of DNA and chromosomal proteins may have a significant impact on regulation of gene expression via age- and hormone-dependent neurochemical changes rather than the structural changes to the brain. Recent studies suggest that this may be a useful model to better understand developmental aspects of schizophrenia. A case study of two pairs of monozygotic twins (one pair concordant for schizophrenia the other discordant) found a higher degree of DRD2 (dopamine D_2 receptor) methylation in the lymphocytes of the affected co-twin as compared to the unaffected twin. Furthermore, the affected twin in the discordant pair was epigenetically more similar to the unrelated affected twins than they were to their own unaffected co-twin. Such epigenetic studies may further the current understanding of how abnormally expressed (developmental) genes affect the development of the psychosis phenotype.

REFERENCES

Bhopal R. Cause and effect: the epidemiological approach. In: *Concepts of Epidemiology: An Integrated Introduction to the Ideas, Theories, Principles and Methods in Epidemiology*. New York: Oxford University Press; 2002:98–133.

Bromet EJ, Naz B, Fochtmann LJ, Carlson GA, Tanenberg-Karant M. Long-term diagnostic stability and outcome in recent first-episode cohort studies of schizophrenia. *Schizophr Bull.* 2005;31(3):639–649.

Brown AS. Prenatal infection as a risk factor for schizophrenia. *Schizophr Bull.* 2006;32(2):200–202.

Cannon M, Jones P, Murray RM. Obstetric complications and schizophrenia: historical and meta-analytic review. *Am J Psychiatry*. 2002;159:1080–1092.

*Cardo A, Murray RM. The “classical” genetic epidemiology of schizophrenia. In: Murray RM,

Jones PB, Susser E, van Os J, Cannon M, eds. *The Epidemiology of Schizophrenia*. Cambridge: Cambridge University Press; 2002:194–219.

Caspi A, Moffitt TE. Gene-environment interactions in psychiatry: joining forces with neuroscience. *Nature Rev Neorosci*. 2006;7:583–590.

Davies G, Welham J, Chant D, Torrey EF, McGrath J. A systematic review and meta-analysis of northern hemisphere season of birth studies in schizophrenia [review]. *Schizophr Bull*. 2003;29(3):587–593.

European Network of National Networks studying Gene-Environment Interactions in Schizophrenia. Identifying gene-environment interactions in schizophrenia: contemporary challenges for integrated, large-scale investigations. *Schizophr Bull*. 2014;40:729–736.

Hegarty JD, Baldessarini RJ, Tohen M, Waternaux C, Oepen G. One hundred years of schizophrenia: a meta-analysis of the outcome literature [see comment]. *Am J Psychiatry*. 1994;151(10):1409–1416.

Heinz A, Deserno L, Reininghaus U. Urbanicity, social adversity and psychosis [review]. *World Psychiatry*. 2013;12(3):187–197.

Hollander AC, Dal H, Lewis G, Magnusson C, Kirkbride JB, Dalman C. Refugee migration and risk of schizophrenia and other non-affective psychoses: cohort study of 1.3 million people in Sweden. *BMJ*. 2016;352:i1030.

Howes OD, Murray RM. Schizophrenia: an integrated sociodevelopmental-cognitive model. *Lancet*. 2014;383:1677–1687.

Krabbendam L, van Os J. Schizophrenia and urbanicity: a major environmental influence-conditional on genetic risk. *Schizophr Bull*. 2005;31(4):795–799.

Lichtenstein P, Yip BH, Bjork C, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet*. 2009;373:234–239.

Malla A, Iyer S, McGorry P, et al. From early intervention in psychosis to youth mental health reform: a review of the evolution and transformation of mental health services for young people. *Soc Psychiatry Psychiatr Epidemiol*. 2016;51(3):319–326.

Messias E, Kirkpatrick B, Bromet E, Ross D, Buchanan RW. Summer birth and deficit schizophrenia: a pooled analysis from 6 countries. *Arch Gen Psychiatry*. 2004;61(10):985–989.

*Moore TH, Zammit S, Lingford-Hughes A, et al. Cannabis use and risk of psychotic or affective mental health outcomes: a systematic review. *Lancet*. 2007;370(9584):319–328.

Morgan C, Gayer-Anderson C. Childhood adversities and psychosis: evidence, challenges, implications. *World Psychiatry*. 2016;15(2):93–102.

Ottman R. Theoretical epidemiology: gene-environment interaction: definitions and study designs. *Prev Med*. 1996;25:764–770.

*Pedersen CB, Mortensen PB. Are the cause(s) responsible for urban-rural differences in schizophrenia risk rooted in families or in individuals? *Am J Epidemiol*. 2006;163(11):971–978.

Rapoport JL, Addington AM, Frangou S. The neurodevelopmental model of schizophrenia: update 2005. *Mol Psychiatry*. 2005;10:434–449.

Reininghaus U, Dutta R, Dazzan P, et al. Mortality in schizophrenia and other psychoses: a 10-year follow-up of the ÆSOP first-episode cohort. *Schizophr Bull*. 2015;41(3):664–673.

Riley B, Kendler KS. Molecular genetic studies of schizophrenia. *Eur J Hum Genet*. 2006;14:669–680.

*Rothman KJ, Greenland S. Causation and causal inference. In: Rothman KJ, Greenland S, eds. *Modern Epidemiology*. 2nd ed. Philadelphia, PA: Lippincott-Raven Publishers; 1998:7.

Schunemann HJ, Guyatt GH. Clinical epidemiology. In: Ahrens W, Pigeot I, eds. *Handbook of Epidemiology*. Berlin: Springer; 2005:1169–1224.

Shih RA, Belmonte PL, Zandi PP. A review of the evidence from family, twin and adoption studies for a genetic contribution to adult psychiatric disorders. *Int Rev Psychiatry*. 2004;14(4):260–283.

Sipos A, Rasmussen F, Harrison G, Tynelus P, Lewis G. Paternal age and schizophrenia: a population based cohort study. *BMJ*. 2004;329(7474):1070.

van Os J, Kapur S. Schizophrenia. *Lancet*. 2009;374:635–645.

*van Os J, Reininghaus U. Psychosis as a transdiagnostic and extended phenotype in the general population. *World Psychiatry*. 2016;15(2):118–124.

Zammit S, Allebeck P, Andreasson S, Lundberg I, Lewis G. Self reported cannabis use as a risk factor for schizophrenia in Swedish conscripts of 1969: historical cohort study [see comment]. *BMJ*. 2002;325(7374):1199.

▲ 12.6 Cellular and Molecular Neuropathology of Schizophrenia

JESSICA MARIE PEREZ, M.S., AND SUBROTO GHOSE, M.D., PH.D.

Schizophrenia is a chronic brain disease for which the basic pathophysiology has remained elusive. The anatomical structure and function of the living human brain can be studied using indirect techniques while the direct microscopic and molecular characterization can be achieved using only human postmortem brain tissue. Early human postmortem studies conducted in the first half of the past century met with success for some neuropsychiatric illnesses such as Alzheimer disease, Pick disease, and Huntington chorea but not schizophrenia. It was not until the 1980s that there was a resurgence of interest in the neuropathological characteristics of the schizophrenic brain, as the tools and understanding in neuroscience itself became more sophisticated, sensitive, and reliable.

It is widely accepted that while no obvious, identifiable neuropathological lesion is seen in schizophrenia, a more subtle pathology does exist. A large number of modern postmortem studies of pathology in tissue of schizophrenic persons have now been published. However, due to incomplete guiding hypotheses, the postmortem literature of schizophrenia is very broad and at times fragmented. Nonetheless, changes that have most consistently been found suggest localization to specific brain regions such as the prefrontal cortical regions, thalamus, and limbic structures in the brain.

MORPHOMETRIC STUDIES AND CYTOARCHITECTURAL CHANGES

Whole Brain

The most consistent observation in schizophrenia is the normal macroscopic appearance of the postmortem brain. However, over the course of the past three decades, postmortem brain tissue studies have begun to provide clues into anatomical and histological correlates of schizophrenia pathology. It should be noted that their findings are disparate, sometimes discrepant, or even conflicting. This may be due to disease heterogeneity or methodological issues such as quality of postmortem tissue, sample size, and confounds of postmortem research such as tissue quality. Nonetheless, a coherent picture supported by converging data is beginning to emerge.

A meta-analysis looking at the postmortem brain weight found a small but significant reduction in schizophrenia that inversely correlates with the age of onset. Consistent with this finding, structural imaging studies demonstrate that brain volumes are smaller in the prodrome, at the time of first break psychosis, and in the chronically ill subject. A recent

longitudinal study reported reduced prefrontal and temporal lobe volumes in individuals with high genetic load for developing schizophrenia. Those individuals who did go on to develop schizophrenia showed even greater reductions in prefrontal volumes. Furthermore, there is evidence that gray matter volume reductions may be progressive in the initial stages of the illness. Regional postmortem studies have identified evidence of subtle changes in cellular morphology with regions most often implicated being the prefrontal cortex (PFC), thalamus, and medial temporal lobe (Table 12.6–1). An initial study in 1982 found that a large proportion of patients with schizophrenia showed evidence of gliosis. While several studies since then have not supported this finding, more recent studies suggest that the different types of glia may play roles in schizophrenia pathology. While the absence of other classical attributes of the neurodegenerative diseases, such as obvious cell loss or neurofibrillary tangles, has led to the belief that schizophrenia is not a degenerative disease in the traditional sense, the reported glia-related changes raise the possibility of an immunological or inflammatory component to disease pathophysiology. A landmark study in 1976 reported ventricular dilatation in subjects with chronic schizophrenia. Since then, numerous studies have replicated this observation, primarily an enlargement of the third and lateral ventricles in subjects with schizophrenia. This observation, however, lacks diagnostic specificity, as it can occur in numerous other brain disorders, such as Alzheimer disease and hydrocephalus. In addition, high-risk and family studies show that enlarged ventricular volume is also found in unaffected relatives who share genetic material (and, by extension, schizophrenia risk genes).

Table 12.6–1.
Summary of Golgi Studies Examining Dendritic Spine Density in Schizophrenia

Region	Reported Finding	Observations
Prefrontal cortex	Lower	In layer 3
Temporal cortex	Lower	In layer 3
Visual cortex	Unchanged	In layer 3
Hippocampus		
Subiculum	Lower	Apical dendrites
CA3	Higher	Apical dendrites

Prefrontal Cortex

The PFC has been hypothesized as a major site of dysfunction in schizophrenia due to the reduced task-related prefrontal activation in patients with the illness and due to the prominent deficits in certain cognitive functions that depend on the PFC. Structural imaging studies have provided evidence that the prefrontal lobe volume is reduced in

schizophrenia; stereological analysis of postmortem brain samples has refined this observation by ascribing the volume deficit primarily to the gray matter that has been found to be approximately 3 to 12 percent smaller in schizophrenia subjects than in their normal controls. Gray matter is composed of neurons, glial cells, and neuropil that include neuronal processes such as dendritic trees and axonal arbors. Initial studies using a two-dimensional stereological approach detected neuronal loss in the PFC with no evidence of change in glial numbers. More recent studies have employed a three-dimensional stereological approach that takes into account cell size and number per unit volume. These studies have not found changes in neuron number, suggesting that the volume decrease is not attributed to neuronal loss but rather reductions in neuropil. Indeed, quantitative analyses of neuron number and density in the PFC (Brodmann areas 9 and 46) demonstrated increases in neuronal density, smaller pyramidal somal cell volumes, and reduced neuropil in schizophrenia in layer 3 (Fig. 12.6–1) but not in layer 5 or 6. The reductions in neuropil suggest that schizophrenia is associated with a quantitative deficit in neuronal connections. Several studies have examined dendritic spine density in schizophrenia (Table 12.6–1) and consistently find reductions in spine density in layer 3 of the PFC. The layer 3 pyramidal neurons serve as a major site for integration of information from other intrinsic cortical neurons and input from the thalamus, implicating these neural pathways in schizophrenia pathophysiology.

The issue of influence of long-term antipsychotic treatment and its possible effect on these cellular measures has been examined. Exposure of nonhuman primates to chronic antipsychotic medication does not appear to alter cortical neuronal density, suggesting that the elevated neuronal density found in schizophrenia is not a medication effect. Interestingly, in these chronically treated animals, the numbers of glia were increased as much as 33 percent in certain cortical layers. The implication of this is not yet clear but may suggest a role for glia in terms of antipsychotic response.

Hippocampal Formation

Magnetic resonance imaging (MRI) studies find consistent reductions in hippocampal or amygdala-hippocampal volumes in schizophrenia that may be lateralized to the left side. Meta-analyses report hippocampal reductions of the magnitude of 4 percent in schizophrenia. Smaller hippocampal volumes have been shown to be present at the time of disease onset and in those with high genetic loading for the illness. Human postmortem studies suggest an anteroposterior gradient of the neuropathological findings, with the anterior part putatively more affected than the posterior part. If the volume decrease holds true, what causes it is still a matter of debate. The current consensus is that there are no significant differences in neuronal density in the hippocampus in schizophrenia, although it remains a

possibility that there are cell-type specific differences in subregions of the hippocampus. Studies of neuronal size have demonstrated somewhat more consistent results. As in the PFC, smaller pyramidal neurons have been reported in the hippocampus, but with normal-sized neurons in the visual and motor cortex in schizophrenia. A question that remains to be answered pertains to the time line of changes. Although reduced hippocampal volumes can be identified at the beginning of the illness, imaging data attest to a progressive course (i.e., hippocampal volumes continue to decrease with disease advancement), which may suggest a potential toxic effect of the ongoing disease or drug treatment, particularly first-generation antipsychotic medications. Two Golgi impregnation studies have examined spine density in hippocampal pathways. One reported reductions in spine density while another recent study reported a specific increase in spine density on CA3 pyramidal cell apical dendrites postulated to reflect an increase in CA3 basal activity, thereby raising the possibility of subregional changes within the hippocampus. In the entorhinal cortex, early studies showed neurons in aberrant and poorly formed clusters. Since the spatial arrangement of cells is largely determined during brain development in utero, a disruption in this arrangement could reflect a developmental abnormality. Well-designed follow-up studies taking into account the rostrocaudal heterogeneity of the entorhinal cortex have not replicated these findings, and the results remain controversial. Similarly, initial reports of reduced number and smaller size of neurons in the entorhinal cortex have not held up to replication.

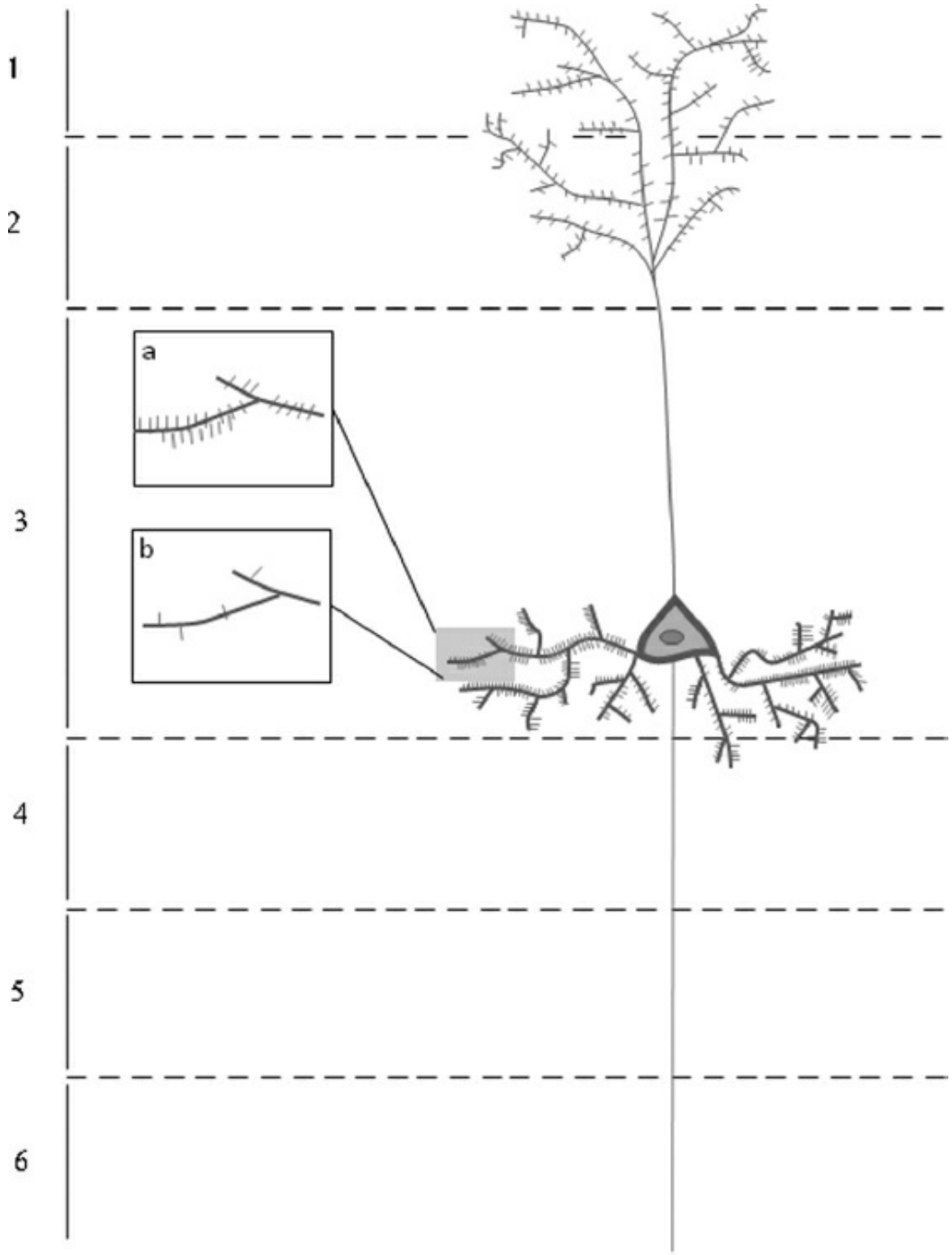

FIGURE 12.6–1. Schematic representation of dendritic spine density in layer 3 prefrontal cortex pyramidal neuron in controls (a) compared to schizophrenia (b).

Thalamus

The thalamus, because of its pivotal position in relationship to afferent sensory information and as a station in the cortical-subcortical circuits, is a brain region of interest. The volume, cell number, and cell densities have been investigated in several thalamic nuclei, including the mediodorsal (MDL) and anterior thalamic nuclei, where the main efferent connections to the PFC and limbic regions arise. In the mediodorsal nuclei, total

neuronal counts are decreased, although not in all studies. These reductions in neuron number are reported to be in the parvocellular and densocellular parts of the MDL, which project primarily to the dorsolateral PFC. More recent studies of larger sample size have not, however, replicated these findings of reduced neuronal number.

Basal Ganglia

Postmortem and in vivo imaging studies have reported volumetric increases of the caudate–putamen complex in schizophrenic patients. Although animal and human studies initially suggested that the volumetric increase could be explained by antipsychotic medications, electron microscopic studies suggest an increase in the number of synapses in the caudate and putamen, suggesting increased neuronal activity in the basal ganglia. Whether this is disease related or a medication effect is not yet conclusive.

Table 12.6–2 summarizes the morphological findings seen in schizophrenia.

WHITE MATTER CHANGES IN SCHIZOPHRENIA

The white matter comprises axons and their myelin sheaths and glial cells, together comprising about half of the brain by volume. Depending on age, gray to white matter ratios of 1.1 to 1.3 have been reported. The development of the human brain involves increases in both gray and white matter with the majority of myelination occurring within the first few years of life, followed by a protracted progression. About half the brain comprises myelinated axonal tracts that connect cortical regions with each other and subcortical regions and is referred to as the connectome. Schizophrenia has been increasingly viewed as a syndrome of impaired information management within neural circuits, a phenomenon that has been widely demonstrated by functional imaging studies leading to greater attention devoted to abnormalities of myelin and oligodendroglia, the myelin sheath–forming cells. Dysfunction in either of them could impair the electrical properties of the axons, leading to defective information transfer along the neural networks.

One of the most consistent alterations observed in postmortem gene expression studies using human postmortem brain tissue from donors with schizophrenia is a decrease in oligodendrocyte pathway–related genes. Replicated decreases of myelin associated genes, myelin-associated glycoprotein (MAG), myelin and lymphocyte protein (MAL), 2′,3′-cyclic nucleotide 3′-phosphodiesterase (CNP), and oligodendrocyte transcription factor 2 (OLIG2, which encodes the transcription factor central to oligodendrocyte development) have been demonstrated in the dorsolateral PFC in schizophrenia although negative studies have also been reported.

MAG, a gene expressed only by myelin-forming cells, enhances the survival of oligodendrocytes. It is interesting to note that the putative risk genes for schizophrenia, neuregulin (*NRG1*) and *DISC1*, are associated with oligodendrocyte and myelin function. *NRG* signaling is involved in nervous system development by modulating a wide array of processes, including glial differentiation, axonal–glial signaling, and myelination while *DISC1* plays a role in oligodendroglial differentiation during development.

Table 12.6–2.
Summary of Main Morphological Findings in Human Postmortem Studies in Schizophrenia Compared to Control Cases

Region	Major Macroscopic Finding	Neuronal Cell Size	Neuronal Cell Number	Observations
Ventricles	Enlargement			Present at onset
Prefrontal cortex	Slight cortical thinning	Smaller	Unchanged	Increased neuronal density in Brodmann areas 9 and 46 Reductions in dendritic spines
Anterior cingulate cortex	Unchanged	Unchanged	Unchanged	
Hippocampal formation	Decreased volume	Smaller	Unchanged	Some evidence of laterality (L < R)
Thalamus	Unchanged	Unchanged	Unchanged	Earlier studies showing reduced neuronal size and number in certain nuclei have not been replicated in recent larger studies
Basal ganglia (enlargement of caudate)	Mixed findings			Increased number of synapses. Possible medication effect

Subtle ultrastructural changes, such as abnormalities in myelin compaction and lamellar arrangement, have been described in postmortem brains of people with schizophrenia. There is also some evidence of a decrease in oligodendrocyte number Brodmann areas 9 and 24 along with abnormal oligodendrocyte morphology. Furthermore, it is of interest that white matter integrity and function are influenced by a number of factors, including glutamatergic signaling. Glutamate is released in an activity-dependent manner from white matter in brain; additionally, glutamate receptors and reuptake transporters are expressed in glial cells within white matter fiber tracts. In animals, altered glutamate homeostasis is associated with oligodendrocyte and axonal damage, suggesting that the impairment in oligodendrocyte integrity seen in schizophrenia could be related to its sensitivity to glutamate. In addition to oligodendrocytes, the other two types of glia are astrocytes and microglia. Although early reports of astrogliosis have not been replicated, the expression of astrocyte-related

genes has been reported in schizophrenia. Astrocytes play an important role in the synaptic metabolism of neurotransmitters such as glutamate, γ-aminobutyric acid (GABA), and monoamines, and could contribute to synaptic dysfunction in schizophrenia. There is some suggestion that microglia are altered in schizophrenia. There are reports of an increase in microglia in the PFC and white matter but not in all studies. A genome-wide association study recently reported that astrocyte and oligodendrocyte gene sets, but not microglia gene sets, are associated with an increased risk for schizophrenia. To clarify these observations, larger and more comprehensive studies are needed to clarify the involvement of astrocytes and microglia in schizophrenia.

NEUROTRANSMITTER SYSTEMS

Neurochemical studies began in schizophrenia at a time when scientists were anticipating a single neurobiological defect. Although no single defect has been found, investigators interpret studies in a broader systems context. Nonetheless, findings are often still organized by neurotransmitter system.

Dopamine

In the 1960s, it was proposed that hyperdopaminergic activity may be responsible for the positive symptoms in schizophrenia. This was based in part on the pharmacological observation that blockade of dopamine receptor subtype 2 (D_2) reduced psychotic symptoms. The other circumstantial observation was that high doses of amphetamine, known to act via the reversal of the dopamine transporter (DAT), elicit psychosis. These findings served as the compelling impetus to study biochemical measures of dopamine neurotransmission in the human postmortem brain in schizophrenia. A series of studies in the nonhuman primate showed that a reduction in dopamine transmission in the PFC was associated with cognitive deficits such as those found in schizophrenia. This led to a modification of the dopamine hypothesis to the presence of an imbalance of dopamine transmission in schizophrenia, with dopaminergic hyperactivity in subcortical regions, but hypoactivity in prefrontal cortical regions. Although the dopamine hypothesis has been the prevailing hypothesis for decades, evidence of dopamine dysfunction in human postmortem studies has not been forthcoming.

The measurement of dopamine and metabolites in lumbar cerebrospinal fluid (CSF) as a reflection of neurotransmitter activity in the brain has not contributed to the understanding of schizophrenia pathophysiology. Studies examining D_1 expression in the PFC and striatum have been largely negative. Increases in D_2 receptors are a consistent finding in a number of postmortem studies, but since chronic antipsychotic

treatment is known to upregulate D_2 receptors, these findings have been largely attributed to a medication effect. D_3 receptor expression was found to be increased in the ventral striatum of medication-free cases of schizophrenia but normal in subjects treated with antipsychotics. D_3 mRNA levels are, however, reported to be normal in the accumbens. There was considerable interest in the D_4 receptor, fueled in part by the high affinity of clozapine for D_4 receptors. Ligand binding studies were initially promising but not replicated in subsequent studies, including one that used a selective D_4 ligand. DAT expression in the striatum has been examined by several research groups, the majority finding that levels are unaltered in schizophrenia. It should be noted, however, that although human postmortem studies have not identified clear evidence of dopaminergic dysfunction in schizophrenia, there are recent imaging studies that suggest higher occupancy of D_2 receptors by dopamine and changes in dopamine release in schizophrenia.

Glutamate

Several lines of evidence implicate excitatory amino acids in the pathophysiology of schizophrenia. The glutamate hypothesis of schizophrenia is centered on the clinical observation that NMDA receptor antagonists, such as phencyclidine (PCP) and ketamine, produce a syndrome that is indistinguishable from schizophrenia.

Brain regions repeatedly implicated in the pathophysiology of schizophrenia include the PFC, hippocampus, and thalamus. The major connections between these brain regions are glutamatergic, making this neurotransmitter system central to the understanding of the abnormal connectivity in schizophrenia. Glutamate is the major excitatory amino acid neurotransmitter in the brain and is known to activate both ionotropic and metabotropic glutamate receptors (mGluRs). The ionotropic receptors, α-amino-3-hydroxyl-5-methyl-4-isoxazolepropionic acid (AMPA), kainate (KA), and NMDA receptors, have been extensively studied at the mRNA, protein, and ligand-binding level in the cortex, striatum, thalamus, and limbic regions in schizophrenia. The data have not been easy to integrate, but the most consistently replicated findings are a decrease in AMPA receptor mRNA and protein expression in the hippocampus and downregulation of the NMDA GluN1 subunit in the temporal cortex, hippocampus, and thalamus. Upregulation of the GluN2B subunit is seen in the temporal cortex and the hippocampus. Studies showing a hypoactive NMDAR are supported by further studies showing reduced expression of the protein complexes that mediate their downstream signaling, such as PSD-95, neurofilament light, and SAP102 in the PFC. Among the mGluRs in schizophrenia, mGluR2 and 3 have been implicated in schizophrenia in animal and human models of schizophrenia.

One of the glutamatergic systems' major postnatal neurodevelopmental processes is the functional switch from expressing NMDA receptors primarily containing the GluN2B subunit to primarily containing the GluN2A subunit, as a transition from juvenile to adult neural processing. There is recent evidence that an early environmental stressor, such as maternal deprivation, in animal models leads not only to schizophrenia-like behavior but also to deficiencies in the GluN2B to GluN2A developmental switch via epigenetic and transcriptional regulation of the *grin2b* gene. This can relate back to increases in GluN2B that are seen in the hippocampus and thalamus of postmortem tissue of patients with schizophrenia.

γ-Aminobutyric Acid

Evidence of GABAergic involvement is found in reduced expression of presynaptic markers in subpopulations of interneurons in cortical brain regions. Increased $GABA_A$ receptor binding has been reported in the prefrontal and cingulate cortices and the hippocampus. GABA is synthesized from glutamate by glutamic acid decarboxylase (GAD). There are two isoforms of this enzyme, GAD65 and GAD67. A consistent finding in schizophrenia is the reduction in GAD67 mRNA levels in the PFC in schizophrenia. GABAergic neurons can be defined by the presence of one of the calcium-binding proteins, namely, parvalbumin (PV), calretinin, and calbindin. The most characteristic morphological types of neurons that express PV are the large basket and chandelier cells. The reduction in GAD67 found in the PFC is associated with GABA transporter reductions in PV-containing inhibitory neurons, perhaps the chandelier cells. These changes suggest reduced synthesis and reuptake of GABA in these neurons and, thereby, altered GABAergic transmission at the chandelier cell–pyramidal cell synapses. It is interesting to note that these GABAergic abnormalities occur in specific GABAergic cells, speaking against a medication effect that would be expected to have a more general effect. In addition, primates treated chronically with antipsychotic medication do not show the GABA deficits found in the PFC in schizophrenia. The GABAergic abnormalities could potentially be related to dendritic spine changes found in schizophrenia. Reelin, a protein secreted by GABA neurons involved in dendritic spine formation, is also decreased in the PFC and hippocampus in schizophrenia. Since neuronal numbers are unchanged in the PFC, the decrease in GAD67 and reelin is likely due to decreased expression per neuron. GABA is modulated by a number of other genes implicated in schizophrenia such as *BDNF*, *DISC1*, and *NRG1*, which impact migration of PV+ interneurons to the cerebral cortex, GAD67 expression, and development of GABA circuitry.

Acetyl Choline

Clinically, it is well known that schizophrenic patients have a much higher incidence of cigarette smoking. The upregulation in nicotinic receptors seen in "normal" smokers is not seen in schizophrenia. On the contrary, decreased levels of nicotinic and muscarinic receptors are reported in the hippocampus, frontal cortex, thalamus, and striatum in schizophrenia. In addition, cholinergic neurotransmission is known to be integral to cognition and memory, functions disrupted in schizophrenia. Decreases in M_1 and M_4 receptors have been reported in the PFC and striatum in schizophrenia. There is also evidence of reduced expression of nicotinic receptor subunits α7 in the frontal, but not parietal, cortex and reduced expression of α4β2 subtypes in the hippocampus in schizophrenia. These findings suggest cholinergic dysfunction in schizophrenia.

In summary, molecular abnormalities are found in a number of anatomic regions and in several neurotransmitter systems in the neuropathology of schizophrenia. Identification of primary pathology from epiphenomenon is essential. For example, neurotransmitter systems are dynamic, and disruption of one system would lead to compensatory mechanisms in other relevant pathways. A sound strategy to adopt once an abnormality is discovered is to first confirm the finding by replication in a second cohort and then determine whether the molecular abnormality is part of the primary pathology. Converging data from in vivo human studies, postmortem human studies, and animal model studies would provide clues to the primary pathology.

IS SCHIZOPHRENIA A DEVELOPMENTAL DISORDER?

The development of the brain is a complex process that involves an intricate sequence of events marked by sensitive or critical time periods when neural circuits are their most plastic, allowing for essential developmental changes to occur. These critical time periods, however, also make neurodevelopment a highly vulnerable stage where disruptions to essential processes can cause permanent adaptations in brain circuitry and even behavior. Although the etiology of schizophrenia is unknown, genetic predisposition is clear and the role of environmental factors in the development of schizophrenia is evident from the fact that monozygotic twins have approximately 50 percent concordance rates of schizophrenia. A currently accepted construct to understand schizophrenia is one in which there is an interplay between genetic risk and exposure to environmental factors. The development of the brain is a protracted process during which time cortical growth and remodeling occurs, axonal tracts undergo myelination and the GABA, and glutamate and monoamine systems develop and mature from birth through youth and adolescence to stable adult levels, changing slowly into senescence.

Neuronal Migration and Neurogenesis

During the second trimester of fetal development, neurons migrate upward from the ventricular wall to their target cortical layer. Initial postmortem studies in cortical tissue from schizophrenic individuals found ectopic neurons and abnormal cytoarchitecture in the PFC and entorhinal cortex. These data were consistent with an impairment of neuronal migration of these particular cells into the cortex during their critical developmental period (second trimester) in schizophrenia and interpreted as evidence that schizophrenia was a developmental disorder. These intriguing findings led to a series of more precise and carefully designed studies that have not been able to replicate or confirm the initial findings. With recent advances in the field of adult neurogenesis, modern tools have become available to study the proliferation and differentiation of stem cells in the brain. The use of antibodies previously used in cancer research has recently been applied in histological studies of the postmortem human hippocampus, including Ki-67 and proliferating cell nuclear antigen. There are initial observations that suggest a reduction in cell division in the hippocampus in schizophrenia. Although provocative and interesting, replication studies are needed.

Spinogenesis and Pruning

Adult schizophrenia subjects have fewer synaptic connections in specific brain regions, and there is some evidence, although divergent and inconsistent, that synaptic proteins such as synaptophysin, GAP-43, complexin I and II, and SNAP-25 are reduced in these individuals. Early electron microscopy studies showed that cortical synaptic density reaches a maximum at 2 to 4 years of age when it is about double adult levels, with reduction in the number of synapses (synaptic pruning) toward adult levels occurring mainly during adolescence. This prompted the suggestion that schizophrenia is a defect of excessive pruning during adolescence. A more recent extensive study of newborns to those aged 91 years shows that synaptic elimination may be a more extended process continuing throughout the third decade of life before synaptic density stabilizes at adult levels. While the hypothesis of excessive pruning may still hold true, the new studies raise the possibility that pruning may occur over a broader time course. Another possible explanation of reduced spine density in certain brain regions in schizophrenia could be related to altered development and stabilization of dendritic spines.

Myelination

Myelination is critical for the proper axonal conductance and communication between brain regions. The first 2 years of life see the rapid myelination of axonal fibers that continue into adulthood following a region-specific course. In the PFC, for example, myelination occurs well

into the second decade of life. As mentioned earlier, abnormalities in myelin-related genes are observed in schizophrenia. The protracted development course provides opportunities for genetic and environmental exposures to interfere with normal myelin maturation in the brain.

Risk Genes

Schizophrenia is a highly heritable and polygenic illness. GWAS and CNV studies have provided important evidence, suggesting a role of a large number of common alleles of small effect size and rare CNVs of large effect size in schizophrenia. Expression profiles of some of the putative risk genes, including neuregulin (*NRG*), dystrobrevin-binding protein 1 (*DTNBP1*), disrupted in schizophrenia 1 (*DISC1* and *2*), and regulator of G protein-signaling-4 (*RGS 4*), have been characterized in human postmortem tissue. Besides examining expression patterns and levels of expression of the candidate gene transcripts, human postmortem tissue has been invaluable in providing a resource from which to identify alternate splice variants, determine developmental expression profiles, and to determine relationships between variants and other molecular species. As an example, the SNP rs1344706 in *ZNF804A*, a gene that encodes zinc finger protein 804A, is one of the best-supported risk variants for schizophrenia from GWAS, independent replications, and meta-analyses. A developmentally regulated truncated *ZNF804A* transcript was identified using human brain tissue, which is regulated by the SNP in the fetal brain. The full *ZNF804A* transcript is also expressed in the brain but not regulated by this SNP. These data suggest that the truncated transcript might be a critical isoform during an early neurodevelopmental stage of life important to schizophrenia pathophysiology.

Array Studies

Several transcriptome studies have been conducted using postmortem tissue since the first study over a decade ago. Gene expression changes are found in several brain regions including the PFC, hippocampus and temporal cortices, and in different cell types. Correlation networks are increasingly being used in bioinformatics applications. For example, weighted gene coexpression network analysis is a method used to describe the correlation patterns among genes across microarray samples. This approach, successfully used in other biological contexts such as cancer studies, can be used for finding clusters (modules) of highly correlated genes and for relating modules to one another. In the schizophrenia PFC, pathways identified include oxidative phosphorylation, myelination, synaptic transmission, and immune function while data from the hippocampus show deficits in immune response and inflammation, transcriptional regulation, and epigenetic control in PV-containing

interneurons.

Epigenetics

Epigenetic processes dynamically regulate gene expression independently of DNA sequence variation and can regulate key neurobiological processes. Modification of the chromatin structure of DNA (chromatin remodeling) is associated with reversible modifications of DNA rather than with inherited variability within the DNA sequences and is, therefore, called epigenetic modification. This important process is tightly regulated by histones that control how tightly DNA is bound and whether the copy machinery will have access to individual genes. Epigenetic regulation of chromatin can occur via several mechanisms including processes such as DNA methylation and posttranslational modifications of histones. There appear to be periods during development when epigenomic changes take place in the human brain, for example, DNA methylation patterns show a unique temporal pattern across the lifespan with greatest changes occurring during the prenatal period, slowing after birth, and gradual slowing with aging. DNA methyltransferases (DNMT1 and 3), which mediate DNA methylation, are highly expressed in brain and are overexpressed in GABAergic neurons in the PFC in schizophrenia. A differential DNA methylation profile has been reported in leucocytes and postmortem brain in schizophrenia. One genome-wide methylation study using human postmortem schizophrenia tissue identified 2,929 unique genes that were differentially methylated, many of which were also identified in prior studies using leucocytes. Environmental factors can also influence epigenetic changes and, of relevance to schizophrenia, maternal malnutrition, certain intrauterine conditions, and viral infections, can lead to hypermethylation of DNA. The epigenome might serve as a substrate that links environmental exposures, genetic variants, and psychopathology.

TOWARD A PATHOPHYSIOLOGY OF SCHIZOPHRENIA

There is neither a single brain region nor a single neurochemical alteration but several, which have been associated with schizophrenia. The PFC, hippocampus, and thalamus are the regions most often implicated, perhaps in part because these are the regions most studied. At the cellular level, reduced gray matter volumes, reduced size of neurons but without cell loss, and reduced dendritic arborization and spines are the main observations seen in schizophrenia. White matter changes also contribute to connectivity deficits between brain regions implicated in schizophrenia. The cause remains unknown, but the implication of these findings is one of reduced connectivity in the brain. Neurochemical deficits in neurotransmitter systems indicate abnormal synaptic communication between neurons. There may be specific neurotransmitter systems involved in certain cortical

regions with deficits in the GABAergic system being the most consistent. A plausible formulation is that schizophrenia is a disease of abnormal connectivity, one that may occur at the cellular level, synaptic level, or circuit level. The nature, location, and timing of pathology seem to be key factors in the development of schizophrenia and warrant further exploration. The advent of next-generation sequencing (NGS) allows for the rapid sequencing of entire genomes, transcriptomes, and epigenomes from specific brain regions, specific cell types, and specific time points in development. The collaborative use of modern translational research methodologies and tools will be critical to the successful pursuit of understanding the pathophysiology of schizophrenia.

REFERENCES

Akbarian S. Epigenetic mechanisms in schizophrenia. *Dial Clin Neurosci*. 2014;16(3):405–417.

Arnold SE. Cellular and molecular neuropathology of the parahippocampal region in schizophrenia. *Ann N Y Acad Sci*. 2000;911:275–292.

Bernstein H, Steiner J, Guest PC, Dobrowolny H, Bogerts B. Glial cells as key players in schizophrenia pathology: recent insights and concepts of therapy. *Schizophr Res*. 2015;161(1):4–18.

Catts VS, Fung SJ, Long LE, et al. Rethinking schizophrenia in the context of normal neurodevelopment. *Front Cell Neurosci*. 2013;7:60.

Catts VS, Lai YL, Weickert CS, Weickert TW, Catts SV. A quantitative review of the postmortem evidence for decreased cortical N-methyl-D-aspartate receptor expression levels in schizophrenia: how can we link molecular abnormalities to mismatch negativity deficits? *Biol Psychol*. 2016;116:57–67.

Coyle JT, Tsai G, Goff D. Converging evidence of NMDA receptor hypofunction in the pathophysiology of schizophrenia. *Ann N Y Acad Sci*. 2003;1003:318–327.

Davis KL, Kahn RS, Ko G, Davidson M. Dopamine in schizophrenia: a review and reconceptualization. *Am J Psychiatry*. 1991;148(11):1474–1486.

Eastwood SL. The synaptic pathology of schizophrenia: is aberrant neurodevelopment and plasticity to blame? *Int Rev Neurobiol*. 2004;59:47–72.

Gibbons A, Dean B. The cholinergic system: an emerging drug target for schizophrenia. *Curr Pharm Des*. 2016;22(14):2124–2133.

*Glausier JR, Lewis DA. Dendritic spine pathology in schizophrenia. *Neuroscience*. 2013;251:90–107.

*Harrison PJ. The hippocampus in schizophrenia: a review of the neuropathological evidence and its pathophysiological implications. *Psychopharmacology (Berl)*. 2004;174(1):151–162.

Heckers S, Konradi C. Hippocampal neurons in schizophrenia. *J Neural Transm*. 2002;109(5–6):891–905.

Kleinman JE, Law AJ, Lipska BK, et al. Genetic neuropathology of schizophrenia: new approaches to an old question and new uses for postmortem human brains. *Biol Psychiatry*. 2011;69(2):140–145.

Knable MB, Barci BM, Webster MJ, Meador-Woodruff J, Torrey EF. Molecular abnormalities of

the hippocampus in severe psychiatric illness: postmortem findings from the Stanley Neuropathology Consortium. *Mol Psychiatry*. 2004;9(6):609–620, 544.

Konopaske GT, Lange N, Coyle JT, Benes FM. Prefrontal cortical dendritic spine pathology in schizophrenia and bipolar disorder. *JAMA Psychiatry*. 2014;71:1323–1331.

*Lewis DA, Hashimoto T, Volk DW. Cortical inhibitory neurons and schizophrenia. *Nat Rev Neurosci*. 2005;6(4):312–324.

Meador-Woodruff JH, Healy DJ. Glutamate receptor expression in schizophrenic brain. *Brain Res Brain Res Rev*. 2000;31(2–3):288–294.

Pidsley R, Mill J. Epigenetic studies of psychosis: current findings, methodological approaches, and implications for postmortem research. *Biol Psychiatry*. 2011;69(2):146–156.

Schmidt MJ, Mirnics K. Neurodevelopment, GABA system dysfunction, and schizophrenia. *Neuropsychopharmacology*. 2015;40(1):190–206.

Selemon LD. Regionally diverse cortical pathology in schizophrenia: clues to the etiology of the disease. *Schizophr Bull*. 2001;27(3):349–377.

*Selemon LD, Goldman-Rakic PS. The reduced neuropil hypothesis: a circuit based model of schizophrenia. *Biol Psychiatry*. 1999;45(1):17–25.

*Sequeira PA, Martin MV, Vawter MP. The first decade and beyond of transcriptional profiling in schizophrenia. *Neurobiol Dis*. 2012;45(1):23–36.

Tao R, Cousijn H, Jaffe AE, et al. Expression of ZNF804A in human brain and alterations in schizophrenia, bipolar disorder, and major depressive disorder: a novel transcript fetally regulated by the psychosis risk variant rs1344706. *JAMA Psychiatry*. 2014;71(10):1112–1120.

Wockner LF, Noble EP, Lawford BR, et al. Genome-wide DNA methylation analysis of human brain tissue from schizophrenia patients. *Transl Psychiatry*. 2014;4:e339.

▲ 12.7 Structural Brain Imaging in Schizophrenia

AMANDA E. LYALL, PH.D., MAREK KUBICKI, M.D., PH.D., AND MARTHA E. SHENTON, PH.D.

THE BRAIN AND SCHIZOPHRENIA: HISTORICAL PERSPECTIVE

After more than 100 years of research, the neuropathology of schizophrenia is still unknown. This is despite the fact that both Eugen Bleuler and Emil Kraepelin, who first described "dementia praecox" and "the schizophrenias," believed that brain abnormalities would ultimately be linked to the etiology of schizophrenia. This conviction was further fueled by critically important inroads that were being made during this same time period into Huntington's chorea, some forms of epilepsy, Pick disease, tertiary syphilis, and Alzheimer disease. Alois Alzheimer, in fact, was one of the first to study the neuropathology of schizophrenia, although he went on to investigate what he thought were more tractable brain disorders.

Researchers at the end of the 19th and beginning of the 20th centuries were nevertheless hopeful that they would make similar inroads into the neuropathology of schizophrenia. Results from postmortem studies,

however, were quite disappointing and often led to conflicting findings, due in part to the crude measurement tools available, and in part due to the high expectations that researchers had that they would find large abnormalities when in fact such abnormalities have proven to be small and subtle. This state of affairs led to a waning of interest in the neuropathology of schizophrenia and, although a small number of researchers continued to investigate the brain and its role in schizophrenia, the general consensus, as clearly articulated by the eminent neurologist Fred Plum, was that "schizophrenia is a graveyard of neuropathologists." Consequently, research investigating brain abnormalities in schizophrenia lost favor.

Computed Tomographic Studies

It was not until 1976 that interest in investigating brain abnormalities in schizophrenia was renewed with the first computer-assisted tomographic study of schizophrenia by Johnstone and colleagues that showed enlarged lateral ventricles in patients when compared to healthy control subjects. This one study led to a proliferation of computed tomographic (CT) and postmortem studies, with the overwhelming majority showing enlarged lateral ventricles in schizophrenia, compared with healthy controls. This early study was the impetus for a renewed interest in structural brain changes as markers for schizophrenia pathology, an interest that has continued to the present day.

Nuclear Magnetic Resonance/Magnetic Resonance Imaging Studies

Following the introduction of CT, nuclear magnetic resonance (NMR) was developed in the 1970s on the basis of the seminal work of Peter Mansfield and Paul Lauterbur, who shared the Nobel Prize in Medicine in 2003 for their independent contributions to the theory that led to the use of MRI. Following its introduction, MRI soon became the most powerful tool for visualizing soft tissue contrast in the brain (e.g., gray and white matter). In contrast to CT, which uses x-rays, MRI has no known adverse effects, with only a few contraindications. It can thus be used as a research tool in at-risk populations, such as children and pregnant women, and it can be repeated as needed, thereby opening the door to longitudinal studies to investigate brain changes over time, or to monitor more precisely treatment efficacy.

MR scanners were first built and clinically used in the 1980s, and today they are used routinely to measure tissue properties in both the body and the brain. The use of MRI in clinical research made it possible to evaluate specific brain regions in vivo. R.C. Smith and colleagues conducted the first MRI study of the brain in schizophrenia in 1984. This study, and later studies conducted in the late 1980s, could not, however, image the entire brain, as the technology was not yet in place. Nonetheless, the promise of

what MRI offered the field led to a proliferation of studies in schizophrenia. Of further note, MRI studies conducted between 1988 and 2000 led to compelling evidence that brain abnormalities are present in schizophrenia. These studies also observed, not surprisingly, that most abnormalities are small and subtle, rather than large and dramatic, suggesting still further that more advanced and accurate measurements were needed to detect such brain abnormalities. As detailed later, a number of novel imaging technologies have since become available that make possible even more precise identification and characterization of subtle structural abnormalities in schizophrenia.

Diffusion Tensor Imaging Studies

Diffusion tensor imaging (DTI) is a relatively new imaging technology for evaluating brain structure. This technique, inspired by earlier work on diffusion MR from 1986, can detect the direction of water diffusion in living tissues. DTI was first introduced by Basser and colleagues in 1994 and then applied to human brain research by Carlo Pierpaoli and colleagues in 1996. The first study of schizophrenia employing DTI was by Monte S. Buchsbaum and coworkers in 1998. The number of studies since this time period has increased dramatically, as described later.

Prior to the advent of DTI, it was difficult to characterize white matter abnormalities in schizophrenia because white matter appears uniform and homogeneous in conventional MR scans. DTI, however, affords a unique opportunity to investigate and to quantify more accurately microstructural white matter abnormalities in the brain. Specifically, this technique provides an indirect measure of the molecular motion of water in brain tissue, which differs depending on tissue type. For example, in CSF and in gray matter, the diffusion of water is random and equal in all directions. This diffusion is said to be isotropic, or, geometrically spherical in shape. In contrast, in white matter, the motion of water is more restricted and is said to be anisotropic, or, geometrically ellipsoidal in shape, maximal along axons, and most restricted perpendicular to the main axonal axis. Furthermore, the tissue properties of white matter, including axonal fiber density, fiber diameter, thickness of myelin sheaths, and the direction or orientation of fibers within in each voxel (volume of a pixel), all affect the diffusion of water molecules. Anisotropic diffusion is interpreted as an indicator of white matter integrity, where higher anisotropy indicates greater integrity of fiber tracts and lower indicates less integrity.

To date, DTI studies in schizophrenia have focused on both large areas of the brain, such as whole brain white matter, and white matter in the prefrontal and temporal lobes, as well as smaller, more anatomically defined regions of interest, such as interhemispheric connections including the corpus callosum, and frontotemporal connections, including the uncinate fasciculus, cingulate fasciculus, and arcuate fasciculus.

Summary

The time period between 1988 and 2000 witnessed a burgeoning of neuroimaging studies in schizophrenia "that have led to more definitive findings of brain abnormalities than in any other time period in the history of schizophrenia research" (see review by Shenton and colleagues in 2001). Combined with the wider application of MRI methodologies, there were also unparalleled advances in MR image acquisition protocols. For example, image acquisition paradigms evolved from evaluating a single 1-cm slice in the brain on MR scanners with low field strength (0.5 T) to acquiring 1-mm slices throughout the entire brain on MR scanners with higher field strengths (1.5-T and 3-T, with 4- and 7-T magnets just beginning to be used). This change in technology made it possible to evaluate the entire brain, using thin, contiguous slices, thereby resulting in greater anatomical specificity. Furthermore, the subsequent development of cutting-edge image analysis tools, for MRI and DTI, has led to an unprecedented number of novel findings in schizophrenia. More specifically, the introduction of segmentation algorithms that automatically classify tissue into gray matter, white matter, and CSF, replacing more manual measurements, as well as newer automated atlas-based parcellation methods, makes it possible to conduct larger studies with increased accuracy, greater speed, and higher reliability when compared to previous studies that employed manual tracing methods (i.e., manual region of interest—ROI methods). Nonetheless, manual tracing methods still remain the gold standard in the field. Figure 12.7–1 shows a coronal slice through the brain where the temporal lobe, amygdala, parahippocampal gyrus, and superior temporal gyrus (STG) have been manually drawn. Figure 12.7–2 shows a lateral view of the brain where the surface shows the different lobes of the brain.

Recent advances in imaging technology also make it possible to associate gray matter and white matter abnormalities with genetic and environmental contributions, cognitive and behavioral testing, and neuropsychological measures. Significant efforts are also being made to link multimodal imaging techniques within the same patients, so as to explore more fully the entire brain and its role in schizophrenia. It has also become increasingly clear that inroads to understanding the biological underpinnings and developmental progression of this disorder hinge on more methodologically rigorous studies that introduce more innovative and sensitive measures for both brain structure and function. In summary, the 21st century has been, and will continue to be, an exciting time for brain imaging studies in schizophrenia, as advances in imaging technology over the past 20 years make new insights into the brain possible.

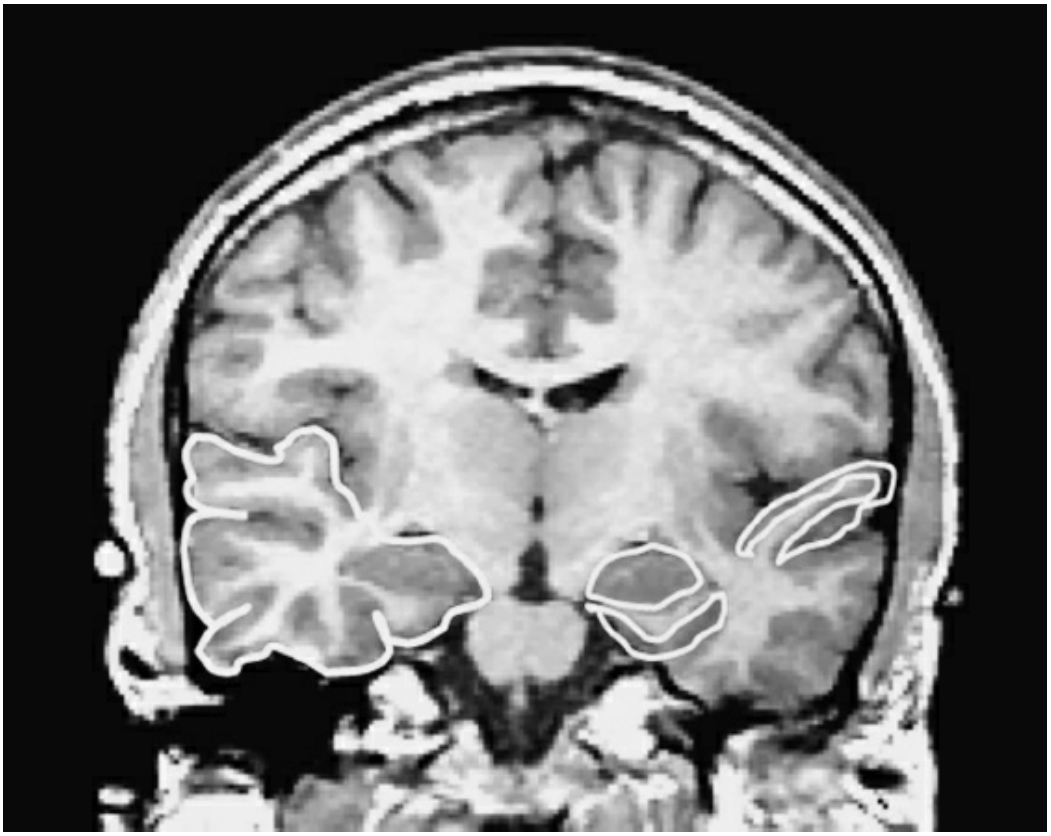

FIGURE 12.7–1. Coronal 1.5-mm slice that shows the medial temporal lobe region of the brain, with the temporal lobe delineated in white on the left side of the image (subject's right). Regions delineated in white on the right side of the brain (subject's left) are the superior temporal gyrus, which borders the Sylvian fissure, the amygdala (the almond-shaped object), and the parahippocampal gyrus, inferior to the amygdala. (From Shenton ME, Kikinis R, Jolesz FA, Pollak SD, LeMay M. Left temporal abnormalities in schizophrenia and thought disorder: a quantitative MRI study. *N Engl J Med*. 1992;327(9):604–612.)

OVERVIEW

As described previously, advances in imaging were critical for confirming structural brain abnormalities in schizophrenia. These early seminal studies inspired new studies to include (1) more homogeneous patient groups to explore relationships between structural abnormalities and clinical and symptoms; (2) more at-risk populations, based on increased clinical or genetic high risk; (3) the application of newer, more sensitive methods such as DTI to investigate whole-brain brain circuitry and white matter fiber tracts (see also last section in chapter); and (4) the use of multiple methods to investigate brain structure and function within the same patient (e.g., PET-MR, functional MRI [fMRI], DTI, or magnetic resonance spectroscopy [MRS], to name just a few). It also became clear that future investigations need to elucidate the timing of brain abnormalities and how they may change over time.

The goal of this chapter is to review progress in MR/DTI imaging techniques to further our current understanding of schizophrenia. The

graph in Figure 12.7–3 shows the number of structural MRI publications per year since 2000, and Figure 12.7–4 shows the number of DTI publications per year since 2000. Figures 12.7–5 and 12.7–6 document the shift in the type of clinical populations investigated, showing an increase in the number of studies on first-episode patients and on individuals with an increased clinical or genetic high risk. This greater focus on earlier stages, including prodromal stages, and the increased interest on the developmental progression of structural abnormalities are all clearly in line with recommendations from the authors' previous review.

In what follows, this chapter will provide a selective review of neuroimaging findings in schizophrenia. First, this chapter summarizes findings from 1988 to 2000, which are described in detail in the review by Shenton and colleagues, 2001. This is followed by a review of structural MR findings in chronic schizophrenia, first-episode schizophrenia, and at-risk populations from 2000 to 2014. Next, this chapter reviews DTI studies from 1998 to 2014, including chronic schizophrenia, first-episode schizophrenia, and at-risk populations. Finally, this chapter highlights current challenges in the field and discuss future directions for schizophrenia neuroimaging research.

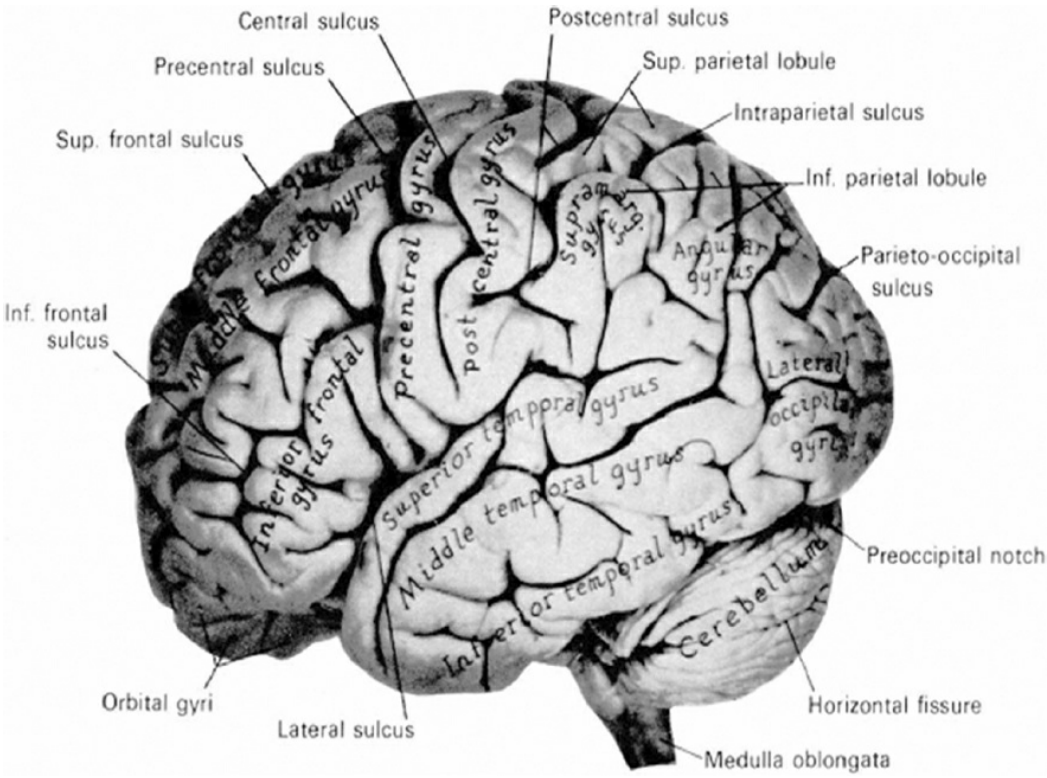

FIGURE 12.7–2. Lateral view of the human brain. (From Carpenter MB. *Human Neuroanatomy*. New York, NY: Williams & Wilkins; 1983.)

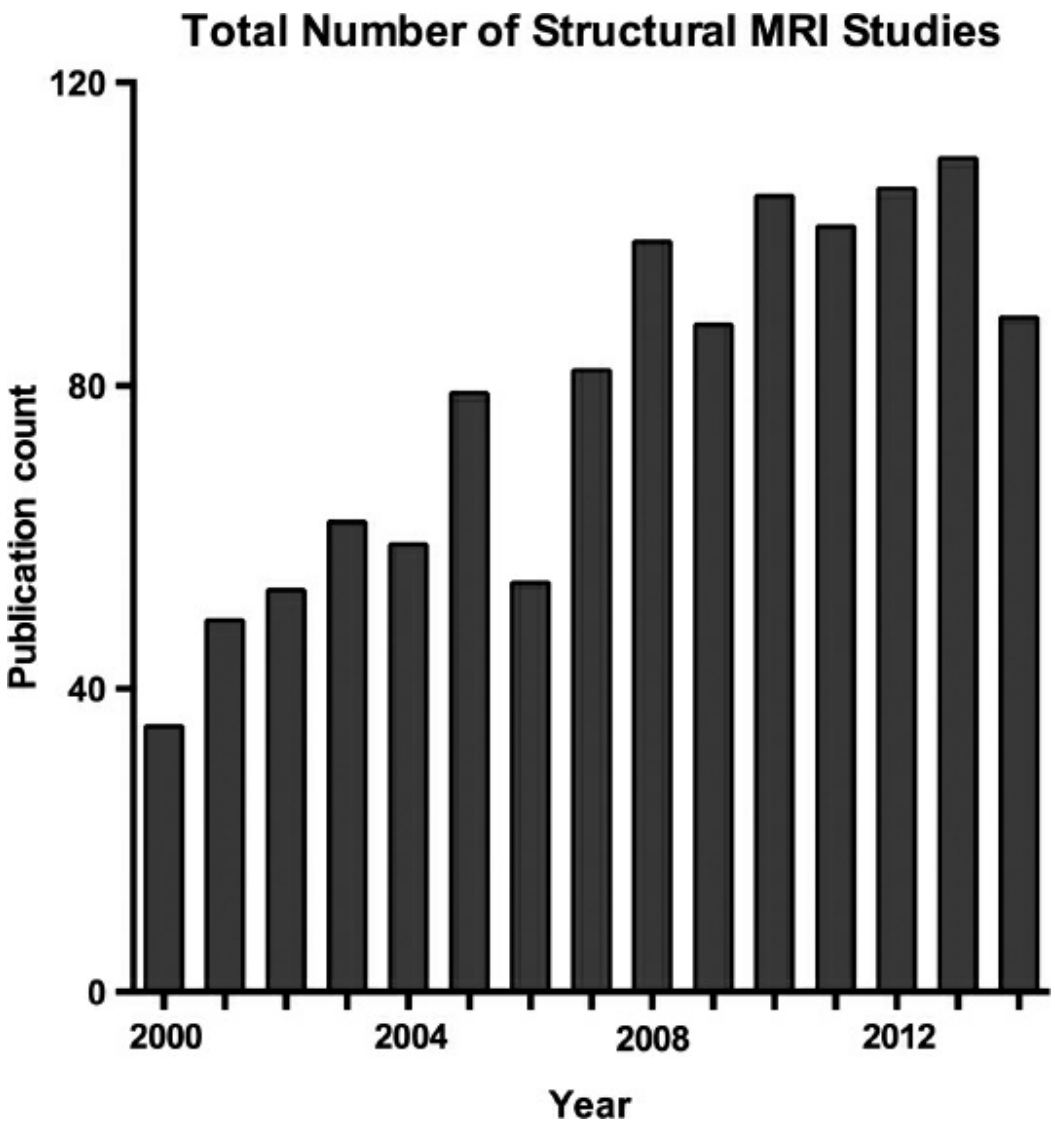

FIGURE 12.7–3. Total number of structural MRI studies on schizophrenia published from 2000 to 2014. Since 2000, of the 1,410 imaging studies identified (MRI and DTI), the authors identified 1,171 published structural MRI studies related to schizophrenia, excluding case studies, studies with no control group, studies with small sample sizes ($n < 10$), studies on schizotypal personality disorder or childhood-onset schizophrenia, reviews and meta-analyses, or studies from other modalities, such as PET, MRS, and fMRI. These studies are displayed by aforementioned years.

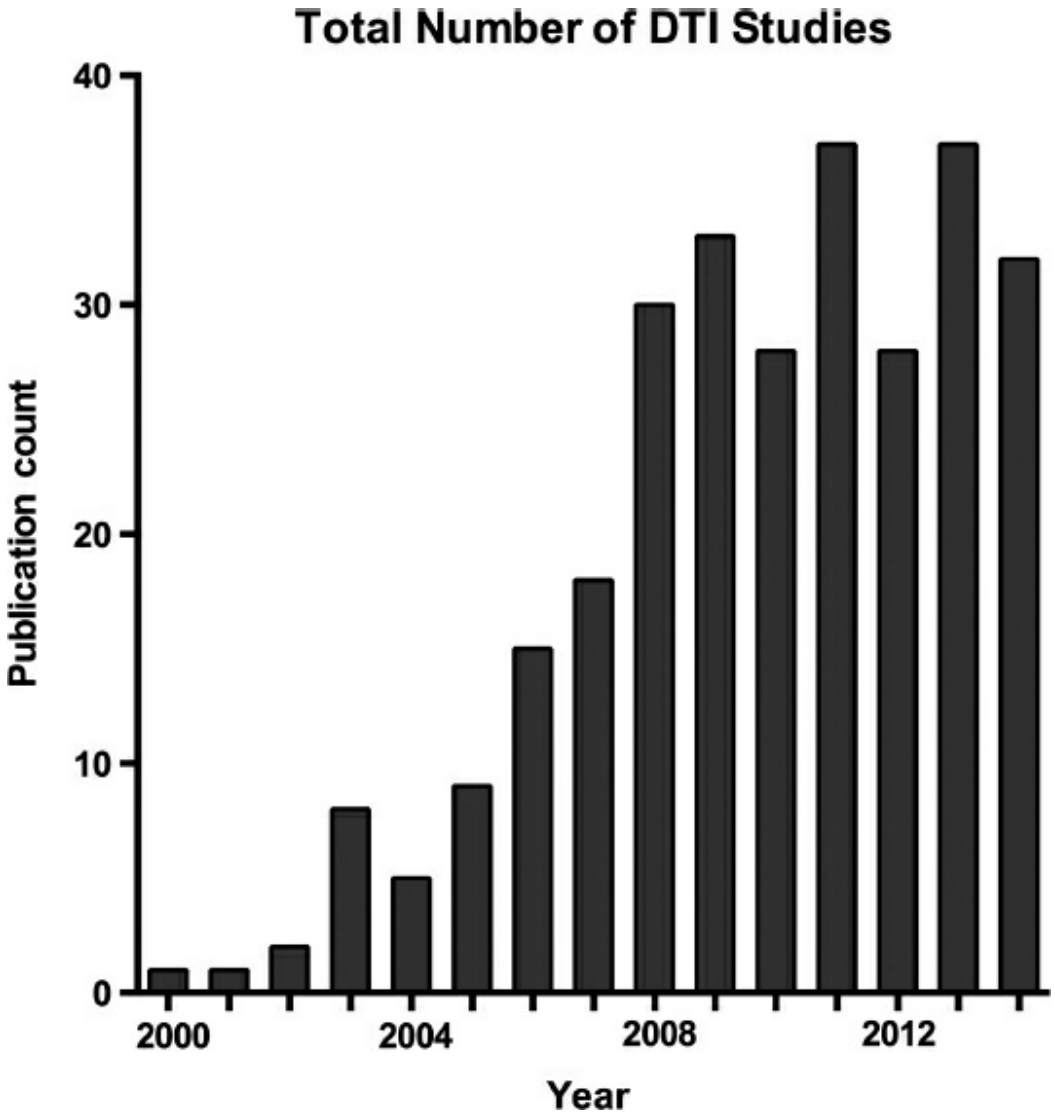

FIGURE 12.7–4. Total number of DTI studies on schizophrenia published from 2000 to 2014. Since 2000, of the 1,410 imaging studies identified (MRI and DTI), the authors identified 284 published DTI studies related to schizophrenia, excluding case studies, studies with no control group, studies with small sample sizes (n < 10), studies on schizotypal personality disorder or childhood-onset schizophrenia, reviews and meta-analyses, or studies from other modalities, such as PET, MRS, and fMRI. These studies are displayed by aforementioned years.

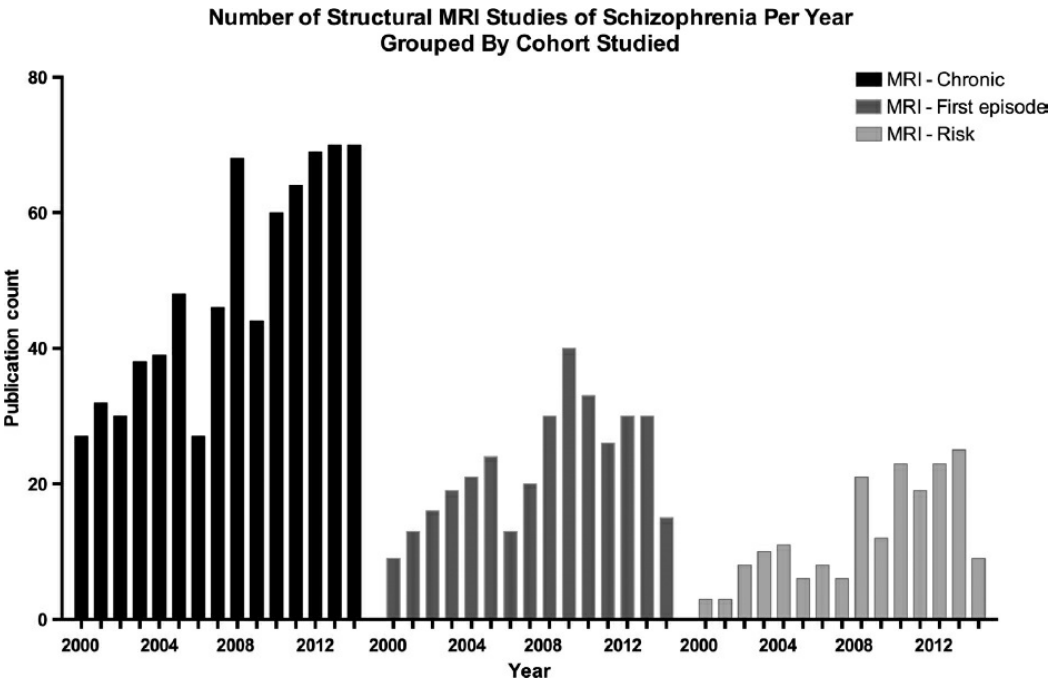

FIGURE 12.7–5. Total number of structural MRI studies in schizophrenia published from 2000 to 2014: Separated by cohort. In total, there were 740 chronic schizophrenia studies (*black*), 341 first-episode studies (*gray*), and 210 studies of at-risk for schizophrenia populations (*light gray*) that the authors identified after applying their exclusion criteria (see Figs. 12.7–3 and 12.7–4). These studies are displayed by aforementioned years of publication. It is important to note that for these yearly counts for cohorts, a study might be counted twice if the study utilized two populations, in the same study, from the aforementioned categories (e.g., a study of chronic schizophrenia and their unaffected relatives will be included in both MRI-chronic and MRI-risk categories). Therefore, the sum of all chronic, first-episode, and at-risk studies presented in this graph will be greater than the total number of MRI and studies presented in Figure 12.7–3.

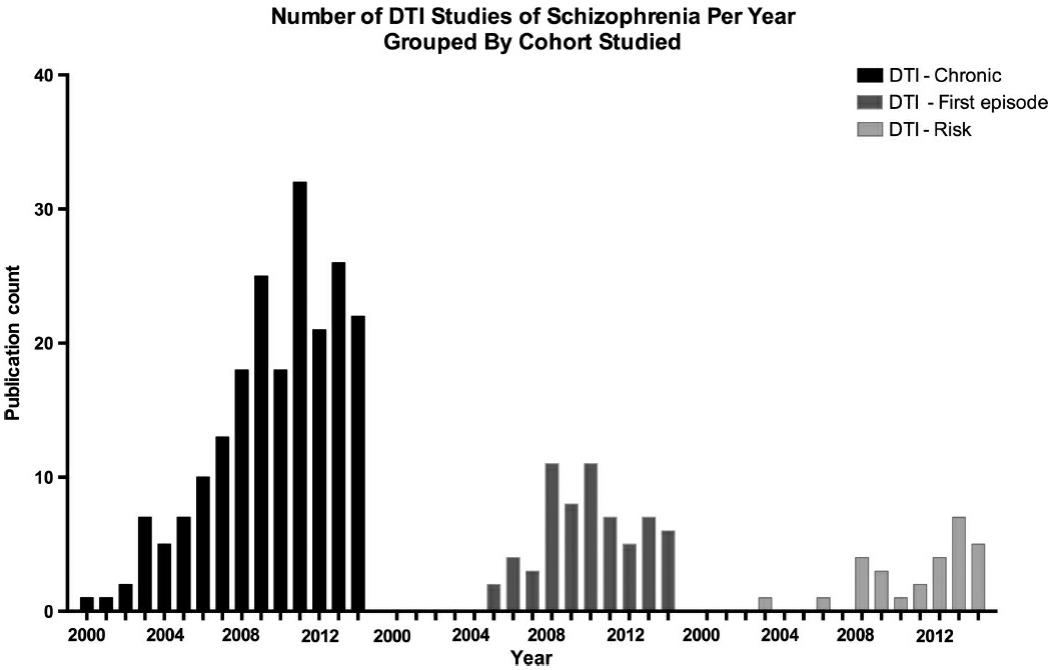

FIGURE 12.7–6. Total number of DTI studies in schizophrenia published from 2000 to 2014: Separated by cohort. In total, there were 211 chronic schizophrenia studies (*black*), 64 first-episode studies (*gray*), and 28 studies of at risk for schizophrenia populations (*light*

gray) that the authors identified after applying their exclusion criteria (see Figs. 12.7–3 and 12.7–4, previously). These studies are displayed by year of publication. It is important to note that for these yearly counts for cohorts, a study might be counted twice if the study utilized two populations, in the same study, from the aforementioned categories (e.g., a study of chronic schizophrenia and their unaffected relatives will be included in both DTI-chronic and DTI-risk categories). Therefore, the sum of all chronic, first-episode, and at-risk studies presented in this graph will be greater than the total number of DTI studies presented in Figure 12.7–4.

MR FINDINGS IN SCHIZOPHRENIA

Summary of MR Research Findings in Schizophrenia from 1988 to 2000

The first studies of schizophrenia, as noted earlier, used relatively crude neuroimaging techniques by today's standards. Nonetheless, the early studies were seminal in confirming that schizophrenia is in fact "a brain disease."

In 2001, Shenton and colleagues reviewed 193 peer-reviewed structural MRI studies of schizophrenia conducted between 1988 and 2000. Such a review would not be possible today, given the exponential increase in the number of studies since 2000 (see Fig. 12.7–3). This review confirmed a number of findings described earlier in CT and postmortem studies. For example, 80 percent of the studies reviewed reported enlarged lateral ventricles in patients with schizophrenia compared to controls. This was a common finding in early CT studies and has proven to be the most replicated finding in schizophrenia. The second most replicated finding in this review was deficits in the medial temporal lobe, where 74 percent of studies reported volumetric reductions in medial temporal lobe structures, most notably in the amygdala, hippocampus, parahippocampal gyrus, and proximal neocortical areas, such as the STG. Other abnormalities included cavum septi pellucidi (92 percent), third ventricle enlargement (73 percent of studies), parietal lobe volume reduction (67 percent), frontal lobe volume reduction (59 percent), occipital lobe volume reduction (44 percent), as well as thalamus (42 percent) and cerebellar volume reduction (31 percent). Basal ganglia showed abnormalities (68 percent; however, some increased volume, some decreased volume), as did the corpus callosum (63 percent showed abnormalities).

Work following this time period, reviewed later, has led to a further understanding of brain structural abnormalities in schizophrenia, through the incorporation of newer tools, more accurate measurements, and information gleaned from other imaging modalities. As will be made clear later, a further appreciation of the importance of brain circuitry abnormalities and neural systems, with respect to both the development of specific cognitive deficits and clinical symptoms of schizophrenia, has also become a topic of interest in current research.

MR Findings in Schizophrenia from 2000 to 2014

The vast majority of studies today are no longer investigating whether schizophrenia is "a brain disease" but instead are asking: "What sort of disease is it?" Researchers are more interested in posing questions such as: Are there disease-specific structural abnormalities in schizophrenia? At what point in brain development do brain abnormalities first appear? Do these brain abnormalities change over time? If brain abnormalities change over time, do they have a neurodevelopmental origin or are they the result of neurodegenerative processes, or both? Will early treatment be effective at mitigating or preventing the progression of the disorder? Could early treatment help attenuate cognitive decline? A focus on answering these questions will bring psychiatrists to a closer understanding of the nature and time course of schizophrenia.

As shown in Figures 12.7–3 and 12.7–4, there has been a substantial increase in the number of MRI/DTI publications in schizophrenia over the past 14 years. Since 2000, a PubMed Web search identifies 4,358 publications relevant to MRI/DTI studies in schizophrenia. This number is reduced to 1,410 when publications that do not meet the criteria are excluded. Specifically, case studies are excluded, studies with no control group, studies with small sample sizes ($n < 10$), studies investigating schizotypal personality disorder or childhood-onset schizophrenia, reviews and meta-analyses, or studies using other modalities, such as PET, MRS, and fMRI ($n = 2{,}948$ publications excluded). The sections that follow provide a selective review of studies conducted since 2000 and up to 2014. This chapter also highlights studies that the authors believe will provide critical insights for future neuroimaging investigations into the underlying pathophysiology of schizophrenia and populations at high risk for developing schizophrenia.

Chronic Schizophrenia

Chronic schizophrenia represents the most severe end of the clinical spectrum. These individuals have been living with a chronic illness and most have received medications over a long period of time, both factors that may confound findings. Nonetheless, between 2000 and 2014 the number of studies addressing abnormalities in chronic schizophrenia has grown exponentially (see Fig. 12.7–3, which documents 740 chronic schizophrenia studies published since 2000). While advances in image acquisitions, most notably the increased use of 3T magnets, have led to higher resolution and higher image quality, the most remarkable observation is the consistency of findings and the number that report brain abnormalities in chronic schizophrenia, including prefrontal, frontal, temporal, and parietal cortical involvement, as well as subcortical structures such as the amygdala and hippocampus.

An increase in data quality and the availability of more refined measurement tools have also led to the emergence of new, target regions

that were previously difficult to delineate, such as the temporal pole, insula, cingulate gyrus, and dorsolateral prefrontal cortex (DLPFC). Studies are also more focused on relating structural abnormalities with deficits in cognitive and clinical functions, such as impairments in face recognition (fusiform gyrus and amygdala), hallucinations (STG), deficit and nondeficit symptoms (prefrontal cortex), impairments in social cognition (anterior cingulate), language disturbances (medial temporal lobe measures), and formal thought disorder (STG, posterior temporal lobe), to name a few. Some of these specific findings are discussed in detail later.

More recent studies have attempted to disentangle the effects of chronicity, medication, substance use, and aging on findings in chronic schizophrenia. Such studies, including a study by Ho and colleagues in 2008, suggest that medication might play a significant role in volume decrease over time, especially in white matter. Furthermore, it has been suggested by Tang and colleagues in 2009 that volume deficits observed in chronic subjects might be a function of "accelerated aging," rather than necessarily resulting from progressive neurodegenerative processes. Others, including Egashira and coworkers in 2014, suggest that duration of illness may be a key contributor to the degree of cortical volume loss observed in chronic populations.

More recent studies in chronic patients also utilize cortical surface measures, that is, cortical thickness and surface area, to model the biological underpinnings of schizophrenia cortical volume loss. Cortical surface measures deconstruct global cortical volume into two primary components: cortical thickness and cortical surface area, which provides a more detailed picture of cortical changes. Also, because cortical thickness and surface area have been shown to be genetically independent morphological features, a greater understanding of which structural phenotype is affected in schizophrenia may narrow the window for investigating underlying aberrant neurodevelopmental mechanisms and related genes.

A further example of using cortical surface measures is demonstrated by Thompson and coworkers, who in 2001 reported cortical volume loss, mediated by a posterior to anterior wave of cortical thinning that begins in the parietal lobe and progresses forward to frontal and temporal areas. In a more recent study, Rimol and coworkers, in 2012, reported reductions in cortical surface area in the frontal, temporal, and occipital lobe, where cortical area reductions were observed most in areas where cortical thinning was less prominent. Interestingly, Rimol and colleagues also showed that cortical surface area reductions in chronic patients were a disorder-specific finding for schizophrenia compared to adult bipolar subjects.

The use of cortical surface models also makes possible the study of gyral patterns in chronic schizophrenia. A recent study by Nesväg and colleagues

in 2014, for example, investigated 207 chronic patients and 206 healthy control subjects, where lower local gyrification index (LGI) was observed compared with controls in the left precentral gyrus, right middle temporal gyrus, and right precuneus. The finding of decreased gyrification in chronic schizophrenia provides yet further evidence for neurodevelopmental aberrations in schizophrenia, as the principle folds of the cortex are established in late stages of prenatal and early perinatal development.

Finally, an exciting new development in the field of neuroimaging is the availability of much larger sample sizes for clinical populations. For instance, in 2014, Haijma and colleagues conducted a meta-analysis on more than 8,000 chronic schizophrenic patients with structural imaging data. This kind of analysis is made possible by collaborative efforts among multiple research groups worldwide. The availability of such great numbers of schizophrenic patients with imaging data will make possible the application of more cutting-edge and computationally intensive methods, such as structural equation modeling, network analysis, and machine learning algorithms, all of which generally require larger sample sizes in order to achieve statistical power. In addition, many of these contributing studies have collected multiple types of imaging data and the task now is to link other modalities, such as fMRI, DTI, MRS, and so forth, to the observed structural abnormalities in schizophrenia. This kind of approach will allow clinical researchers to begin to delineate further regional or network brain abnormalities that may be specific to schizophrenia compared to other psychiatric disorders, or to identify abnormalities that are associated with other contributing factors, such as medications or substance abuse.

In summary, and as highlighted previously, a number of significant structural abnormalities have been reported in chronic schizophrenia. These abnormalities are commonly observed in frontal, temporal, and parietal cortices, as well as in multiple subcortical structures and the cerebellum. These structural abnormalities have also been associated with functional and cognitive differences in chronic schizophrenia populations. However, despite the many benefits of studying chronic subjects, schizophrenia is hypothesized to be a "neurodevelopmental disorder," meaning that the structural abnormalities outlined previously are the result of an aberration that occurred in the earlier stages of neurodevelopment. Thus, to gain a greater understanding of schizophrenia, it is important to focus attention on earlier stages of the disorder.

First-Episode Patients

Since 2000, neuroimaging studies in schizophrenia have gradually shifted from the study of chronic, clinically stable, and medicated patients to investigating patients at their first psychotic episode (see Fig. 12.7–5, which documents 341 first-episode schizophrenia studies published since 2000).

A focus on first-episode patients, in fact, is critical to answering a number of questions, including the following: (1) Are brain abnormalities present at first onset of illness, where chronic illness and long-term effects of medication are not confounding factors? (2) If present, are these brain abnormalities static, or do they progress over the postonset period? If they progress, what is the time course of progression and what are the implications? (3) And, finally, if present, are the brain abnormalities observed at first episode specific to schizophrenia?

Early MRI studies of first-episode patients were primarily cross-sectional and used less sophisticated MR technology than what is available today. Not surprisingly, this led to more equivocal findings in terms of which brain regions were abnormal at the first episode of illness.

In 2005, Christos Pantelis and coworkers published an important review of MR findings from longitudinal first-episode studies and prodromal and high-risk individuals. Based on this review, the authors concluded that there is an accelerated gray matter reduction in prefrontal regions of the brain that they believe interacts with other factors, such as substance abuse and stress. This reduction, in turn, leads to further progressive changes in medial temporal and orbitofrontal brain regions, all of which are observed at first onset of illness. This view thus incorporates multiple pathological processes that occur at different neurodevelopmental time points. The first time point likely involves an early neurodevelopmental insult or lesion that "renders the brain vulnerable to later brain maturation processes" that take place during adolescence or early adulthood. These speculations are reminiscent of Sarnoff Mednick and Thomas S. McNeil's 1968 two-hit theory of schizophrenia and of I. Feinberg's 1982 theory that schizophrenia results from dysfunctional programmed synaptic elimination, or pruning, that occurs as part of normal development during adolescence, which goes awry in schizophrenia.

Interestingly, in Hulshoff Pol and Kahn's 2008 review, progressive gray matter reductions were also extant in first-episode schizophrenia subjects and were more pronounced in the first 20 years of the illness compared with age-matched controls and compared to changes that occur later in the course of the disease. Changes occurring in the first 20 years primarily manifested as cortical gray matter loss in the frontal and temporal lobes, paired with an increase in lateral ventricle volume. Of note, the hippocampus did not appear to show progressive brain changes over time in the majority of studies that investigated hippocampal volume reduction following a first episode of illness, although some studies did show an increase in third ventricle volume over time. Moreover, more pronounced brain changes appear to be associated with poorer outcome, more negative symptoms, and poorer performance on neurocognitive measures, although the latter are more equivocal. There is also some indication that

abnormalities in brain maturation are involved and have their strongest influence in the first 20 years, prior to the age of 45 years.

Despite the increasing number of publications focusing on first episode or early schizophrenia and following them longitudinally—the location, extent, as well as the progression of structural changes remain a matter of active debate. For example, in a 2006 study that followed first-episode patients over a 2- to 3-year period of time, Whitford and coworkers used voxel-based morphometry measures to evaluate the whole brain, and they reported widespread gray matter reduction in frontal, parietal, temporal lobe (including STG), cerebellum, and parts of the occipital lobe. However, in a more recent study from 2012, Vita and coworkers conducted a meta-analysis of current longitudinal studies on first-episode populations that had been scanned within 24 months of illness onset. In this analysis, they reported a loss in whole-brain gray matter volume as well as volume loss in frontal, temporal, and parietal lobe gray matter but did not report differences within the cerebellum and circumscribed regions of the occipital lobe.

The magnitude of structural changes reported in first-episode studies is, in fact, quite small compared to the changes observed in chronic patients. For example, in 2007, Nakamura and coworkers reported, at baseline, reduced neocortical gray matter, larger sulcal CSF, and increased lateral ventricles in first-episode patients compared with controls. When rescanned after 1.5 years, neocortical gray matter volume was reduced by 1.7 percent, and this was localized to frontal (2.4 percent) and temporal regions (2.6 percent). In addition, there was a 7 percent increase in sulcal CSF compared with controls and an increase of 10.4 percent for lateral ventricles over the 1.5-year follow-up period. Since the observed changes in first-episode subjects are inherently small, analyses require large data sets to achieve statistical power. Multisite neuroimaging consortia, such as CIDAR and MIND studies, focus on the recruitment of first-episode subjects to acquire large databases that are publicly available and include imaging data so that the schizophrenia research community can conduct more powerful analyses.

Nakamura and coworkers also found that poorer outcome was associated with postonset brain changes in patients diagnosed with a first episode of schizophrenia. In contrast, patients experiencing a first episode of an affective disorder, with psychotic features, showed increases in neocortical gray matter volume (3.6 percent) at follow-up 1.5 years later, which were associated with lithium and/or valproate medications. Since increases observed in the affective group were not associated with clinical symptoms, and also seemed to be specific to those receiving mood stabilizers, these investigators concluded that the changes observed in neocortical gray matter in the affective group were likely not intrinsic to the disorder, as in the first episode schizophrenia sample, but were more likely

associated with medication effects. It was further concluded that the progressive changes in neocortical gray matter, observed in first-episode schizophrenic patients, were consistent with neuropil loss reported in postmortem studies.

Taken together, these findings suggest that relatively small brain abnormalities are most likely present at first episode of illness, and that some brain regions may exhibit progressive changes over a relatively short period of time after illness onset (see also later studies in the Neurobiology of DTI Changes in Schizophrenia section, which focus on changes in DTI that also show both possibly neuroinflammatory changes at early onset of illness and more progressive changes in prefrontal regions). More studies, however, are needed that follow patients over longer periods of time in order to delineate further which regional brain abnormalities are present at illness onset and, ultimately, to determine which are specific to schizophrenia, which progress over time, and what the implications of progression versus nonprogression might be. In addition, given that many MR studies suggest brain changes in the early postonset period, this time period would appear to be critically important, and attention should be given to understanding processes taking place at this stage of the illness. Understanding these processes might also lead to earlier interventions in the course of illness and to prevention, amelioration, or, at the very least, to mitigating further progressive brain changes.

Of further note, most of the first-episode studies are naturalistic and include patients as they enter the hospital. Some of these patients have been receiving psychotropic medications prior to admission and are, therefore, not necessarily medication naive. Although it can be argued that these patients have not been receiving psychotropic medications for long periods of time, as is the case with chronic patients, medication is still an issue as there is evidence to suggest that even short-term effects of medications can affect brain structures such as the basal ganglia.

Another area that needs further attention is the study of cognitive impairments, clinical symptoms, and their association with progressive and nonprogressive brain abnormalities. Although some work has been done here, with studies reporting associations between brain regions and both clinical and cognitive impairments, evidence suggests that cognitive and clinical symptoms improve early in the course of illness, while structural brain abnormalities progress. It would thus be important to understand why these two seemingly incongruous events take place and to understand further their implications with respect to timing and progression, possible underlying mechanisms, and intervention and treatment strategies.

Finally, as more is learned about brain structures that are compromised in individuals as they transition into a first episode of schizophrenia, it is also important to turn to MR brain studies of individuals prior to the onset of illness. Focusing on the prodromal period, prior to illness onset, makes it

possible to determine whether structural brain abnormalities are present in a subset of individuals who are at risk for developing schizophrenia, who then go on to develop schizophrenia, versus another subset of individuals who are at risk for developing schizophrenia but who never become ill. Such a focus will also make it possible not only to learn more about the timing of brain abnormalities in schizophrenia but also to predict brain markers or brain signatures that predispose an individual to develop schizophrenia.

Another strategy is to evaluate family members so as to discern brain abnormalities that are associated with genetic variations in brain structure but are neither necessary nor sufficient for the development of psychosis (i.e., they appear in the nonpsychotic first-degree relatives). These latter topics are reviewed later, where there are 210 clinical high-risk and family genetic high-risk studies conducted between 2000 and 2014 (Fig. 12.7–5).

Clinical High-Risk Studies

An important question to answer in schizophrenia, as noted previously, is whether brain abnormalities are present prior to the onset of illness (i.e., before psychosis begins). To some extent, this question can be addressed by longitudinal studies that begin prior to illness onset, in what is called the *prodrome period*. A further question to be addressed, also noted previously, is whether the brain abnormalities observed are independent of psychosis, in that they are present in family members without psychotic symptoms, and are therefore necessary, but not sufficient, for psychosis to occur, but may, nonetheless, be important as an indicator of genetic vulnerability to schizophrenia. Furthermore, as reviewed by Pantelis and coworkers in their extensive review in 2005, brain changes observed at onset of psychosis may reflect an additional brain change that is neurodevelopmental in origin but does not take place until transition to active psychosis.

There are three large research studies that focus on those who are at risk for developing psychosis. The first is the Edinburgh High-Risk Study (EHRS), the second is the Melbourne Personal Assessment and Crisis Evaluation (PACE) study, and the third is the North American Prodrome Longitudinal Study (NAPLS), which is a multisite study focusing on the prodromal period where individuals exhibit some clinical symptoms but have not converted to psychosis.

The EHRS study defines individuals at risk on the basis of measures of cognitive impairment, derived from the Structural Interview for Schizotypy (SIS), where those who are above designated cutoff scores are identified as at risk. Findings from this study suggest that schizotypal features and abnormal right prefrontal cortical folding are predictive of those who go on to develop schizophrenia, compared with those at high risk who do not. This work was based on findings of cortical folding in the prefrontal lobe in

family members of patients with schizophrenia, who later developed schizophrenia. These investigators suggest that abnormalities in cortical folding, observed in individuals who are at risk for schizophrenia, and who go on to develop schizophrenia, reflect disordered connectivity in the right prefrontal lobe.

In the Melbourne study, Pantelis and colleagues define individuals at UHR for developing psychosis (i.e., schizophrenia as well as other psychoses) on the basis of number of trait and state risk factors. Of interest here, 35 percent of UHR went on to develop schizophrenia within 1 year. Moreover, individuals who went on to develop schizophrenia showed an increase in brain volumetric abnormalities in medial temporal and prefrontal (particularly orbitofrontal) brain regions, compared with UHR subjects who did not go on to develop schizophrenia. The authors further speculated that their findings may indicate a different brain trajectory for individuals who develop schizophrenia versus those who develop nonschizophrenic psychosis or affective psychosis, a notion that has received a great deal of attention, most notably in studies focusing on white matter, as discussed later.

In a recent publication from the NAPLS consortium in 2015, Cannon and coworkers investigated a cohort of 274 clinical high-risk (CHR) cases, where 35 converted to psychosis. Baseline scans and follow-ups were performed 12 months after baseline or to point of conversion to psychosis. No significant differences were found at baseline for cortical thickness and subcortical or ventricular volumes. However, CHR individuals who converted to psychosis exhibited a significantly faster rate of gray matter loss in right superior frontal, right middle frontal, and right medial frontal gyri as well as a significantly larger expansion in third ventricular volume compared to nonconverters and healthy controls. In addition, the authors found that converters with no exposure to antipsychotic medications during the interscan interval actually showed a greater reduction in right prefrontal cortical thickness and greater expansion of third ventricle volume than nonconverters (with and without antipsychotic exposure) and healthy controls. This finding suggests that the differential rate of gray matter loss between converters and nonconverters was not explained by exposure to antipsychotic medications, further suggesting that cortical thinning may be the result of the pathophysiology of onset of psychosis.

Based on these findings, these investigators speculate that the brain abnormalities observed likely reflect abnormal brain maturation in those who develop schizophrenia, which occurs, in concert, with other events such as substance abuse and stress. These investigators also note that there are likely early markers of brain abnormalities, including frontal lobe impairments that occur prior to illness onset, as well as later markers that occur at the time of active transition to psychosis, such as temporal lobe abnormalities, which may be more influenced by environmental factors and

show progression over time. The former were interpreted by Pantelis and colleagues as likely involving early lesions or early neurodevelopmental insults to the brain that render the brain vulnerable to later abnormal processes, including accelerated gray matter loss in frontotemporal regions and abnormal connectivity in prefrontal brain regions. In reviewing some of the family studies, however, it would appear that abnormalities in temporal lobe brain structures (particularly hippocampus) may reflect a genetic vulnerability to schizophrenia. Thus, the story may be slightly more complex than previously thought.

Finally, the relationship between genes and brain structure and function is an area important to investigate. Recently, several international consortia have been created to acquire large genetic and imaging data sets from schizophrenia and at-risk populations, where thousands of cases will be available for investigating gene-imaging associations (an example of which can be found in the 2016 study by Franke and coworkers). The availability of thousands of cases will allow researchers to make use of GWAS, which provide statistical power to search for gene-imaging associations. Data generated from these consortia will lead to further subtyping of high-risk subjects based on brain markers, genes, and outcome, which will be an important and informative direction for future studies (see the Implications and Future Directions for Research section).

Family Studies: Genetic High-Risk Studies

MRI studies of brain abnormalities in the unaffected family members of patients diagnosed with schizophrenia show abnormalities that can be characterized as intermediate between patients and healthy controls. These results are similar to findings reported in clinical high-risk individuals, most notably in the hippocampus, discussed previously. Studying unaffected family members is important for many reasons. First, and most importantly, any structural brain abnormalities are independent of psychosis. These individuals have not been hospitalized for a psychiatric disorder and, therefore, are not suffering from the possible putative or neurotoxic effects of psychosis. Second, unaffected families members are generally not medicated with neuroleptic medications, thus removing the potential confound neuroleptics have on structural brain findings. Information acquired through the study of unaffected family members thus likely reflects an increased genetic vulnerability to developing schizophrenia. Importantly, however, this notion needs to be tempered with the knowledge that genetic factors may interact with environmental factors, thus making it difficult to determine with certainty that brain abnormalities observed in family members of patients with schizophrenia are necessarily driven only by genetic factors.

In a 2007 meta-analysis, Boos and colleagues reviewed 25 MRI studies of unaffected first-degree relatives of patients with schizophrenia and

reported a reduction of left hippocampal volume and increased third ventricle volume. Of note, hippocampal findings were associated with verbal and declarative memory deficits, similar to findings in schizophrenia. The association of hippocampal abnormalities with unaffected family members may be a combination of genetic or environmental influences. For example, it is known that obstetrical complications can lead to hypoxia, which affects the hippocampus, and the hippocampus is also involved in the regulation of stress through the hypothalamic–pituitary–adrenal axis. Although obstetrical problems and stress regulation may be genetic, they may also be associated with environmental factors or a combination of the two. Accordingly, a further investigation is needed to evaluate more closely the cause of hippocampal volume reduction in the unaffected family members of patients with schizophrenia.

A recent review by Bois and colleagues in 2015 summarizes structural MRI findings from multiple familial and genetic high-risk studies and meta-analyses. Cross-sectional analyses found decreasing gray matter volume in genetic high-risk subjects in the DLPFC, Broca's area, Wernicke's area, bilateral anterior cingulate cortex, right STG, left precuneus, left medial frontal gyrus, right middle frontal gyrus, and bilateral hippocampal/parahippocampal regions. Similar results have been reported in studies employing cortical surface measures, with siblings and unaffected relatives showing decreased cortical thickness and surface area in the cingulate gyrus, STG, and prefrontal cortex when compared to healthy controls. Investigations into gyrification patterns in familial high-risk individuals have proven to be generally inconclusive, with some studies reporting significant differences in cortical gyrification in unaffected relatives and siblings, whereas others have not. As stated previously, abnormal gyrification patterns are posited to be indicative of aberrant pre- or perinatal neurodevelopment. Equivocal findings in gyrification in familial risk studies may indicate that the unaffected relatives of individuals with schizophrenia might not have experienced a potential early-stage insult, thereby contributing to their resilience against a transition to psychosis despite their genetic predisposition.

Longitudinal familial risk studies, which follow family members over time to determine whether familial high-risk individuals convert to psychosis, also provide a powerful neuroimaging paradigm to assess potential neurodevelopmental progression. Despite the generally low conversion rate seen in familial high-risk studies, comparisons by Bois and colleagues in 2015 between high-risk converters (HR-T) and nonconverters (HR-NT) have found that HR-T subjects exhibit different developmental trajectories from HR-NT and controls. HR-T subjects tend to show reductions in cortical thickness over time in frontal, temporal, and parietal areas, as well as localized reductions in cortical surface area, while HR-NTs

and controls exhibit cortical thickening or surface area expansion over time. These studies demonstrate that structural differences in schizophrenia may precede symptom-onset and, therefore, provide researchers and clinicians with an avenue for the identification of structural biomarkers of risk that may be used for targeted therapeutic interventions, which might mitigate or prevent the onset of psychosis.

Summary

In summary, MRI has proven to be an extremely powerful tool in identifying structural abnormalities in schizophrenia. Differences in cortical and subcortical structures are reported at all stages of the disorder, from onset to chronicity, as well as in individuals who are at increased risk for developing the disorder. Thus far, the majority of evidence points to schizophrenia being a neurodevelopmental disorder. Early motor and cognitive abnormalities, neurological soft signs prior to onset of illness, sulcal–gyral pattern abnormalities, facial dysmorphology, lack of evidence from postmortem findings, including lack of loss of neurons, and lack of gliosis, would all tend to suggest that schizophrenia is not a neurodegenerative disorder, at least as defined by standard criteria. Also, one must consider the fact that neurodevelopment does not stop at birth. Thus, a better question might be: Are there some patients who show progressive changes following the first episode of illness, while others do not, and if so, what are the factors that contribute to these differences (genes, severity of brain abnormalities at illness onset that lead to putative neurotoxic brain changes, predisposing prodromal brain abnormalities, and so forth), and what are the implications with respect to treatment and outcome? A more complete discussion of future neuroimaging studies addressing the latter questions will be conducted more extensively in the Implications and Future Directions for Research section.

DIFFUSION TENSOR IMAGING FINDINGS IN SCHIZOPHRENIA

DTI Clinical Applications—Overview

As noted previously, DTI is a relatively new and powerful imaging technique for evaluating white matter architecture in the brain. What is most appealing about this technique is that it can be used both to visualize and to quantify white matter fiber bundles in the brain. It is invaluable for investigating white matter integrity in multiple sclerosis (MS), stroke, normal aging, Alzheimer disease, schizophrenia, and other psychiatric disorders, as well as for characterizing subtle diffuse axonal injuries in mild cases of traumatic brain injury (TBI).

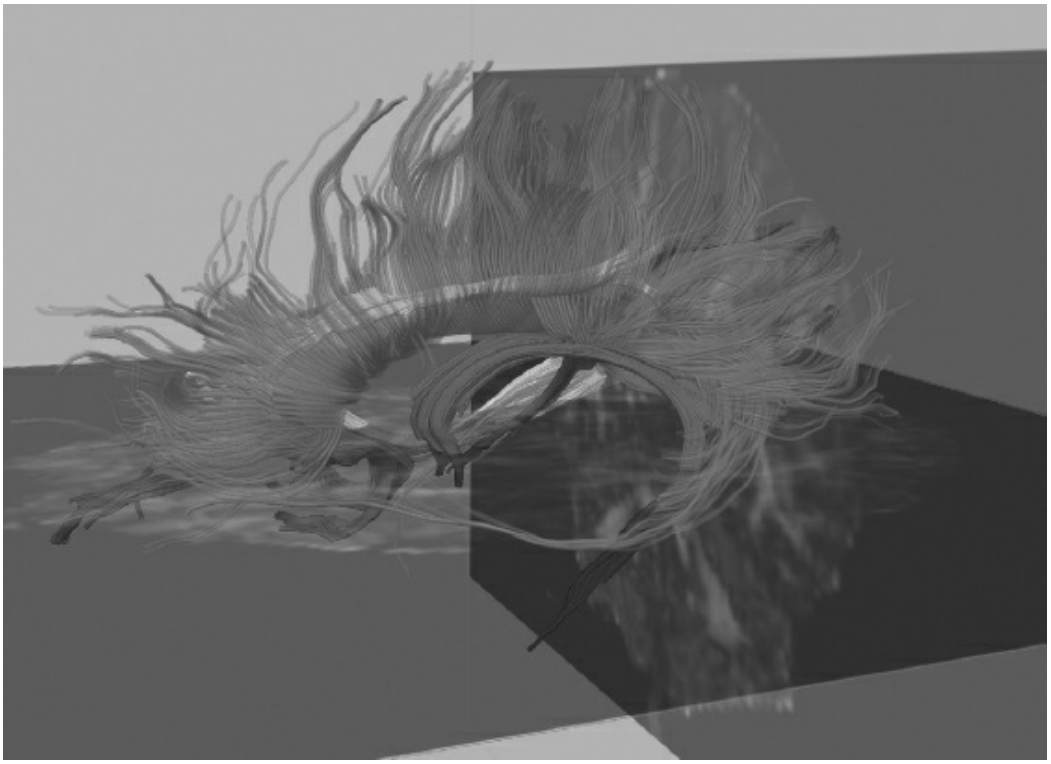

FIGURE 12.7–7. Three-dimensional image reconstructed based on diffusion data acquired on a 3-T General Electric scanner, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, which shows several major white matter fiber bundles identified through diffusion tensor imaging: Fornix (*magenta*), right cingulum (*green*), right inferior longitudinal fasciculus (*yellow*), right uncinate fasciculus (*blue*), and corpus callosum (*orange*). (See color atlas.)

DTI is also emerging as an important tool in presurgical planning, where not only edema associated with brain tumors can be evaluated but also white matter fiber tracts. Moreover, the involvement of white matter and its proximity to the brain tumor can be characterized and quantified so that surgical procedures can be planned that will minimize axonal damage. Furthermore, this technique, while relatively new, is rapidly changing, and although there are many limitations and many questions unanswered with respect to what some of the measures mean, DTI is nonetheless the most powerful tool available for noninvasive, in vivo exploration of white matter pathology in the human brain. Figure 12.7–7 shows a three-dimensional reconstruction of several white matter fiber tracts in the brain derived from DTI images acquired on a 3-T scanner.

The Application of DTI to Schizophrenia

As demonstrated previously, a great deal of progress, especially over the last three decades, has been made in schizophrenia research in characterizing gray matter abnormalities in the brain. This has not been true for white matter, however, until very recently. The main reason for this, as noted previously, is that white matter appears uniform and homogeneous on conventional MR scans. With DTI, it is now possible to

characterize white matter abnormalities in schizophrenia.

A further interest in evaluating white matter in schizophrenia comes from the pattern and number of gray matter abnormalities observed in schizophrenia that are consistent with a disturbance of connectivity between and within brain regions, most likely of developmental origin. Kraepelin early speculated that abnormalities in frontotemporal connections would ultimately be found in schizophrenia. More recently, Weinberger and coworkers suggest that these abnormalities reflect a "disconnection" in the brain that is likely neurodevelopmental in origin.

Additional evidence regarding white matter alterations in schizophrenia comes from neuropathological and genetic studies that suggest the specific involvement of oligodendrocytes, the neural cells that provide protection to axons and improve communication between brain areas, in the pathophysiology of schizophrenia. Thus, an important area of research is to evaluate white matter fiber tract abnormalities in schizophrenia, particularly in frontotemporal tracts as they might help explain the pattern and number of gray matter abnormalities observed in this disorder. Furthermore, focusing on white matter fiber tracts enables researchers to move from evaluating isolated gray matter regions in the brain toward evaluating neural systems and networks that are biological substrates of cognition, social cognition, emotion, and behavior.

As shown in Figure 12.7–4, there have been 284 DTI-related papers published since 2000 (239 with DTI only and 45 with DTI and MRI combined). DTI has been used to test anatomical/functional schizophrenia hypotheses, focusing on specific structures/locations/networks and their involvement in schizophrenia psychopathology, such as abnormalities within connections forming functional networks thought to be most affected in schizophrenia, including language processing, memory, executive attention, emotional processing, and abnormalities within connections related to specific clinical symptoms, that is, establishing a link between the arcuate fasciculus and hallucinations. Recently, the introduction of DTI-related measures, with increased specificity, has made possible the investigation of hypotheses focusing on the neurobiological nature of white matter changes in schizophrenia, including (1) axonal and/or myelin integrity disruptions caused by either disease itself, medications, or other environmental factors; (2) orientation/coherence disruptions likely associated with neurodevelopmental factors; (3) genetically programmed age-/development-related abnormalities affecting axonal growth and maturation; (4) cell/axonal density abnormalities originating from faulty pruning; and (5) neuroinflammation that would either be associated with or predate psychosis onset. Important studies published to date that applied DTI techniques to schizophrenia are discussed below. First discussed are the localization of white matter abnormalities and their association with illness onset, clinical symptoms,

age, medication, and other possible confounds. Then studies combining DTI with other imaging technologies are discussed, followed by future directions of research for the field.

Anatomical Location of White Matter Abnormalities: Functional Consequences

Chronic Schizophrenia. The first DTI studies in schizophrenia evaluated whole brain in relatively small samples (i.e., five chronic patients and six controls in the first study published in 1998 by Buchsbaum and coworkers). Other early studies were also relatively small, with 20 or fewer patients. Even with these small sample sizes, however, reductions in anisotropy were reported and, in fact, in a review by Kubicki and colleagues in 2007, it is clear that the majority of white matter connections are affected in chronic patients (including frontotemporal, fronto-occipital, temporo-occipital, thalamocortical, and interhemispheric connections). This early research not only established the presence of white matter pathology in chronic schizophrenia but also reinforced the observed relationship between white matter pathology and cognitive deficits (i.e., white matter pathology within the cingulum bundle and deficits in attention, or uncinate fasciculus abnormalities and deficits in episodic memory), as well as clinical symptoms (i.e., the relationship between white matter pathology within the arcuate fasciculus and auditory hallucinations, or overall white matter pathology and negative symptoms). Figures 12.7–4 and 12.7–6 show 284 DTI studies in schizophrenia, with 211 in chronic schizophrenia since 2000.

First-Episode Schizophrenia. While early DTI studies in schizophrenia investigated, almost exclusively, medicated schizophrenia patients, it soon became clear that antipsychotic medication, substance abuse, duration of psychosis, and other environmental factors may alter white matter volume, cell density, and oligodendrocyte count, thereby affecting DTI-derived measurements. Thus, in order to understand white matter pathology in schizophrenia, it is important to study patients who are not yet medicated. As seen in Figure 12.7–6, since 2000, there have been 64 DTI studies that have focused on first-episode schizophrenia. The majority of these studies confirm that even when medication, chronicity, and aging effects are taken into account, schizophrenia is still associated with white matter abnormalities, and these abnormalities are located in brain areas similar to those reported in chronic schizophrenia.

It has also been suggested, although replication is needed, that white matter changes might not be progressive in nature but, instead, that the longitudinal changes reported in some DTI studies might be due to medication. A small number of recently published papers also suggest that the location of white matter pathology may strongly depend upon the clinical profile of individual patients, which further emphasizes the need

for clinically homogenous samples or for using an anatomy atlas to characterize individual profile of abnormality for each patient (see later under the Implications and Future Directions for Research section).

Schizophrenia Risk: Clinical and Genetic High Risk. Thus far, only a small number of DTI studies (n = 28; see Fig. 12.7–6) have targeted at-risk populations to determine whether or not white matter abnormalities exist before schizophrenia onset, thereby constituting a vulnerability factor for developing psychosis. This hypothesis is being tested by evaluating white matter in subjects at either high genetic (familial) or clinical (prodromal) risk for developing schizophrenia. Since the rate at which these populations convert to schizophrenia is in both cases relatively low (10 to 15 percent), these populations contain both those who will and will not develop schizophrenia. Thus, detecting changes that are directly related to schizophrenia risk is quite challenging, and resulting findings are equivocal (as is clear from a 2015 review by Peters and colleagues). In general, however, these studies suggest that white matter pathology is likely present before schizophrenia onset. Carletti and coworkers in 2012, for example, suggest that some aspects of white matter pathology might constitute a risk for schizophrenia (i.e., neurodevelopmental pathology), while progressive changes might be a predictor of conversion to schizophrenia. To this end, in a recently published study from our laboratory in 2013 (http://pnl.bwh.harvard.edu), Kubicki and coworkers reported that individuals with genetic risk for schizophrenia exhibit different trajectories of white matter changes over time compared to a healthy, low-risk population. Further studies are needed to understand better the neurodevelopmental trajectories of white matter maturation for both healthy and clinical high-risk schizophrenia populations.

Neurobiology of DTI Changes in Schizophrenia

Early studies suggest that fractional anisotropy fractional anisotropy (FA), the most popular index of white matter integrity derived from DTI, should be considered a proxy for myelin integrity. This interpretation was promoted primarily because FA abnormalities were initially observed in demyelinating diseases such as MS. Later DTI studies, however, have led to questions regarding the specificity of FA to myelin disruptions. For example, it is now believed that FA, while sensitive to myelin integrity disruptions, may also be modulated by other factors, including tract geometry, axonal coherence, axonal size, and changes in the volumes of extracellular water surrounding axons. Several new DTI-derived measures have been proposed as specific to particular micropathologies. Understanding the neurobiology of white matter changes and their longitudinal changes from risk for schizophrenia, through first-episode schizophrenia, and into chronic schizophrenia, will be helpful in the

development of new diagnostic tools as well as for novel intervention strategies. Thus far, several measures have been proposed that partially address the issue of low FA specificity. For example, Seal and coworkers in 2008 investigated the separate contributions of myelin and axonal pathology to FA abnormalities observed in schizophrenia and suggested that myelin may be the principal contributing factor to FA deficits. Similar conclusions have been reached in a few newer studies that used DTI in combination with other measures of myelin integrity, all of which suggest that chronic schizophrenia might be associated with myelin, rather than axonal pathology. In patients with first episode, however, the source of white matter pathology might be different. In a recent 2012 publication by Pasternak and colleagues from our laboratory, a novel DTI-derived measure of free water was used to distinguish between cellular (measuring integrity of axons and oligodendrocytes) and extracellular (proxy of edema and neuroinflammation) water diffusion, rather than separating the cellular components of white matter pathology into contributions made by the axons or myelin. Results suggest that first-episode schizophrenia is likely predominantly associated with increased extracellular water, implying that neuroinflammation, rather than cell degeneration, may underlie FA changes at the very early stages of schizophrenia. These authors also suggest that cellular pathology (oligodendrocyte/axon-related) might be the principal pathology in chronic stages, where prolonged neuroinflammation leads to the neurodegeneration of myelin, a notion that was further confirmed in their subsequent investigation in 2014. This viewpoint is also shared in another recent study by Voineskos and coworkers in 2013, in which genetic data, brain imaging, and cognitive data were combined and used to highlight still further the impact of oligodendrocyte genes (*MAG*, *OLIG2*, and *CNP*) on white matter integrity and cognitive performance in patients with schizophrenia. The question of whether neuroinflammation plays a role in conversion to psychosis is an important question that only future studies can address.

Conclusions

While it will still understanding the time course and biological nature of white matter changes in schizophrenia, progress has been made in this regard within the last several years. Early DTI studies have provided evidence for the existence of white matter abnormalities in schizophrenia. However, only recent studies, which include large and homogenous schizophrenia populations, as well as more precise and more specific measurements, have shed new light into the dynamics and possible underlying neurobiology of these white matter abnormalities. Future studies of those at risk for developing schizophrenia (both genetic and clinical) need to include longitudinal, follow-up scans of those individuals who convert to schizophrenia versus those who do not, as this would help

psychiatrists to better understand the processes associated with psychosis onset. The impact of age and gender on healthy white matter architecture and microstructure, both in terms of localization and longitudinal changes, must also be better understood before any definitive conclusions about abnormal brain development can be made. Finally, studying larger and more homogenous populations will make it possible for the specificity of white matter changes to be understood as they relate to particular symptoms or clinical profiles. All this knowledge directly translates to the development of new treatment strategies. The more specific the understanding regarding both the neurobiology and the time course of schizophrenia-related white matter changes, the earlier, and more targeted the treatment, and the better the chances for good recovery.

IMPLICATIONS AND FUTURE DIRECTIONS FOR RESEARCH

With new methods being developed almost every year, interest in morphological abnormalities in schizophrenia has grown exponentially, as has current knowledge about the extent of brain abnormalities in schizophrenia. The field, however, is still far from understanding fully the neuropathology of schizophrenia. Questions that need to be addressed include the following: What are the neurodevelopmental and possible progressive features of brain changes observed following the first episode of schizophrenia? What brain changes are associated with cognitive and clinical symptoms, and how does the field understand improvement in cognitive and clinical measures and increased volume reduction in discrete brain regions following a first episode of schizophrenia? How does one reconcile increased functional connectivity, with decreased white matter tract integrity, often reported in schizophrenia, as well as before psychotic onset? Are some brain abnormalities stable over time, while others progress? Can brain and cognitive phenotype features be used to predict who will develop schizophrenia and who will not? What is the association between gray and white matter alterations in schizophrenia, and how can these findings be related to functional alterations? Are there points of entry where treatment would be most efficacious? How can brain changes following the first year of psychosis be understood, and can these changes be prevented or at the very least ameliorated?

As technology advances, and new, more refined technologies for in vitro postmortem and in vivo MRI methods emerge, there will be a better understanding of schizophrenia (see Fig. 12.7–8 for use of a two-tensor white matter fiber tracking algorithm). Hypotheses regarding the relationship between neuroanatomical connections and functions in schizophrenia can now be approached empirically, something not possible even 15 years ago. It has become increasingly clear that schizophrenia is a disorder characterized by subtle brain changes. As such, there is presently a shift in focus from the analysis of discrete brain regions towards the study

of the connections between regions. Future studies need to combine multiple, histologically validated imaging techniques, to understand fully neurobiological pathologies observed in schizophrenia, and merge them with genetic data, in order to understand fully imaging phenotypes associated with schizophrenia and those at-risk for schizophrenia. A present day example is of this groundbreaking work is The Human Connectome project, which utilizes high-resolution multi-modal imaging techniques to characterize and further understand the possible brain network abnormalities in patients with schizophrenia.

Furthermore, by studying schizophrenia over the life span, including chronic, first-episode, early- and late-onset, and prodromal patients, and also subgroups of patients who respond and do not respond to medication, an understanding can be developed about the timing of brain abnormalities, about which brain changes can be detected early in the course of the illness, or even before illness onset, and if, and how, these changes progress, and whether or not there are differences in the brain between different clinical subtypes of schizophrenia. By studying brains of patients with genetic syndromes related to schizophrenia, such as velocardiofacial or Klinefelter syndrome, there may be a further appreciation of the role that genes play in schizophrenia. Brain changes over time due to progression also need to be tracked and differentiated from aging and medication-related changes. Moreover, the relationships between genes, brain development, and illness onset and progression need to be investigated further, so as to begin to understand possible mechanisms underlying the neurobiology of schizophrenia.

FIGURE 12.7–8. Panel A shows a coronal view of white matter fiber tracts using a two-tensor model through the corpus callosum. Panel B shows a sagittal view of the corpus callosum shown. (From Shenton ME, Hamoda HM, Schneiderman JS, et al. A review of magnetic resonance imaging and diffusion tensor imaging findings in mild traumatic brain injury. *Brain Imaging Behav*. 2012;6(2):137–192. doi:10.1007/s11682-012-9156-5. Review. PubMed PMID: 22438191; PubMed Central PMCID: PMC3803157, with permission.)

The next decade will thus be an exciting one as more information will emerge that is relevant to what is known about brain structure and function in schizophrenia during the prodrome period as well as at first episode, and

there will be a greater appreciation and understanding of brain abnormalities, cognitive impairments and clinical symptoms, and the impact of what is likely the weak effect of multiple genes on the susceptibility to developing schizophrenia. There will likely also be more studies of treatment efficacy and more attention paid to the critical period following first episode, where progressive changes in the brain are most evident.

REFERENCES

Basser PJ, Mattiello J, LeBihan D. MR diffusion tensor spectroscopy and imaging. *Biophys J*. 1994;66(1):259–267.

Bois C, Whalley H, McIntosh A, Lawrie S. Structural magnetic resonance imaging markers of susceptibility and transition to schizophrenia: a review of familial and clinical high risk population studies. *J Psychopharmacol*. 2015;29(2):144–154.

Boos HB, Aleman A, Cahn W, Hulshoff Pol H, Kahn RS. Brain volumes in relatives of patients with schizophrenia: a meta-analysis. *Arch Gen Psychiatry*. 2007;64(3):297–304.

Buchsbaum MS, Tang CY, Peled S, et al. MRI white matter diffusion anisotropy and PET metabolic rate in schizophrenia. *Neuroreport*. 1998;9(3):425–430.

Cannon TD, Chung Y, He G, et al. Progressive reduction in cortical thickness as psychosis develops: a multisite longitudinal neuroimaging study of youth at elevated clinical risk. *Biol Psychiatry*. 2015;77(2):147–157.

Carletti F, Woolley JB, Bhattacharyya S, et al. Alterations in white matter evident before the onset of psychosis. *Schizophr Bull*. 2012;38(6):1170–1179.

Egashira K, Matsuo K, Mihara T, et al. Different and shared brain volume abnormalities in late- and early-onset schizophrenia. *Neuropsychobiology*. 2014;70(3):142–151.

Franke, B. Stein JL, Ripke S, et al. Genetic influences on schizophrenia and subcortical brain volumes: large-scale proof of concept. *Nat Neurosci*. 2016;19(3):420–431.

*Haijma SV, Van Haren N, Cahn W, Koolschijn PCMP, Hulshoff Pol HE, Kahn RS. Brain volumes in schizophrenia: a meta-analysis in over 18 000 subjects. *Schizophr Bull*. 2013;39(5):1129–1138.

Ho B-C, Andreasen NC, Ziebell S, Pierson R, Magnotta V. Long-term antipsychotic treatment and brain volumes: a longitudinal study of first-episode schizophrenia. *Arch Gen Psychiatry*. 2011;68(2):128–137.

Hulshoff Pol HE, Kahn RS. What happens after the first episode? A review of progressive brain changes in chronically ill patients with schizophrenia. *Schizophr Bull*. 2008;34(2):354–366.

Johnstone EC, Crow TJ, Frith CD, Husband J, Kreel L. Cerebral ventricular size and cognitive impairment in chronic schizophrenia. *Lancet*. 1976;2(7992):924–926.

Kubicki M, McCarley R, Westin CF, et al. A review of diffusion tensor imaging studies in schizophrenia. *J Psychiatr Res*. 2007;41(1-2):15–30.

Kubicki M, Shenton ME, Maciejewski PK, et al. Decreased axial diffusivity within language connections: a possible biomarker of schizophrenia risk. *Schizophr Res*. 2013;148(1-3):67–73.

Nakamura M, Salisbury DF, Hirayasu Y, et al. Neocortical gray matter volume in first-episode schizophrenia and first-episode affective psychosis: a cross-sectional and longitudinal MRI

study. *Biol Psychiatry*. 2007;62(7):773–783.

Nesvåg R, Schaer M, Haukvik UK, et al. Reduced brain cortical folding in schizophrenia revealed in two independent samples. *Schizophr Res*. 2014;152(2-3):333–338.

*Pantelis C, Yücel M, Wood SJ, et al. Structural brain imaging evidence for multiple pathological processes at different stages of brain development in schizophrenia. *Schizophr Bull*. 2005;31(3):672–696.

*Pasternak O, Westin C-F, Bouix S, et al. Excessive extracellular volume reveals a neurodegenerative pattern in schizophrenia onset. *J Neurosci*. 2012;32(48):17365–17372.

Pasternak O, Westin C-F, Dahlben B, Bouix S, Kubicki M. The extent of diffusion MRI markers of neuroinflammation and white matter deterioration in chronic schizophrenia. *Schizophr Res*. 2015;161:113–118.

Peters BD, Karlsgodt KH. White matter development in the early stages of psychosis. *Schizophrenia Res*. 2015;161:61–69.

Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G. Diffusion tensor MR imaging of the human brain. *Radiology*. 1996;201(3):637–648.

Rimol LM, Nesvåg R, Hagler DJ, et al. Cortical volume, surface area, and thickness in schizophrenia and bipolar disorder. *Biol Psychiatry*. 2012;71(6):552–560.

Seal ML, Yücel M, Fornito A, et al. Abnormal white matter microstructure in schizophrenia: a voxelwise analysis of axial and radial diffusivity. *Schizophr Res*. 2008;101(1–3):106–110.

*Shenton ME, Dickey CC, Frumin M, McCarley RW. A review of MRI findings in schizophrenia. *Schizophr Res*. 2001;49(1–2):1–52.

Smith RC, Calderon M, Ravichandran GK, et al. Nuclear magnetic resonance in schizophrenia: a preliminary study. *Psychiatry Res*. 1984;12(2):137–147.

Tang B, Chang W-L, Lanigan CM, Dean B, Sutcliffe JG, Thomas EA. Normal human aging and early-stage schizophrenia share common molecular profiles. *Aging Cell*. 2009;8(3):339–342.

Thompson PM, Vidal C, Giedd JN, et al. Mapping adolescent brain change reveals dynamic wave of accelerated gray matter loss in very early-onset schizophrenia. *Proc Natl Acad Sci USA*. 2001;98(20):11650–11655.

Vita A, De Peri L, Deste G, Sacchetti E. Progressive loss of cortical gray matter in schizophrenia: a meta-analysis and meta-regression of longitudinal MRI studies. *Transl Psychiatry*. 2012;2:e190.

*Voineskos AN, Felsky D, Kovacevic N, et al. Oligodendrocyte genes, white matter tract integrity, and cognition in schizophrenia. *Cereb Cortex*. 2013;23(9):2044–2057.

Whitford TJ, Grieve SM, Farrow TF, et al. Progressive grey matter atrophy over the first 2–3 years of illness in first-episode schizophrenia: a tensor-based morphometry study. *Neuroimage*. 2006;32(2):511–519.

▲ 12.8 Functional Brain Imaging in Schizophrenia

RAQUEL E. GUR, M.D., PH.D., AND RUBEN C. GUR, PH.D.

Functional neuroimaging has become a major translational tool for

advancing the understanding of the neurobiology of schizophrenia. Neuroimaging affords dissecting complex mental processes disrupted in schizophrenia to probe the nature and extent of aberrant activity in the underlying brain circuitry. The functional imaging methodology can thus help incorporate cognitive and affective neuroscience paradigms in efforts to understand brain regulation of behavior. The challenge is to harness the power of functional neuroimaging to clarify the neuropsychiatric manifestations of schizophrenia, integrating this knowledge with advances in genomics. Functional neuroimaging, applied in combination with pharmacologic probes, can augment the bridge between basic and clinical neuroscience. Such advances are helping elucidate mechanisms underlying schizophrenia with implications for treatment.

In the years since the previous edition of this textbook, the literature on functional neuroimaging in schizophrenia has continued to grow, especially in the application of functional magnetic resonance imaging (fMRI) methods applied for activation and for examining resting state connectivity. This chapter will highlight fMRI applications in schizophrenia and clinical risk populations, summarizing main areas of progress, common issues, and potential future directions.

MAGNETIC RESONANCE FUNCTIONAL IMAGING METHODS

The introduction of fMRI was an exciting development in functional imaging research. Functional MRI methods offer multiple advantages over positron emission tomography (PET) for imaging brain function, including higher spatial resolution, higher temporal resolution, noninvasiveness and lack of ionizing radiation, direct correlation with anatomical imaging, greater repeatability, and economy. Disadvantages of fMRI techniques include the presence of loud background noise generated by the gradients, difficulties in presenting stimuli and performing tasks in the magnet bore, claustrophobia, low signal-to-noise ratios for most methods, and lack of quantitation in physiologic units for most methods. Many of these disadvantages can be overcome through the use of specialized equipment compatible with the MRI environment (Fig. 12.8–1). Considering the costs versus benefits and availability of fMRI, most PET studies in schizophrenia now examine receptor function, whereas brain behavior research is conducted with fMRI. The most commonly applied methods are blood oxygenation level–dependent (BOLD) and arterial spin–labeled (ASL) imaging of cerebral blood flow rates in physiologic units (mL/100 g/min).

Blood Oxygenation Level–Dependent Imaging

The technique relies on magnetic susceptibility effects of deoxyhemoglobin, which cause regional signal decreases in imaging sequences that are sensitive to susceptibility (e.g., echoplanar or routine gradient echo

sequences). With regional brain activation studies, a net increase in signal intensity is observed in regions activated by the task. The increase in image intensity corresponds to a local decrease in deoxyhemoglobin. This effect is attributed to a greater increase in regional blood flow than in regional oxygen consumption, as indicated by PET measurements of blood flow and oxygen consumption associated with brain activation. A variety of pulse sequences can be used to obtain BOLD measures. Many simple activation paradigms have been tested, and activation has been observed with both fast and slow imaging. A typical response is a 1 to 25 percent change in regional image intensity that develops over 3 to 8 seconds following task initiation. Susceptibility effects of deoxyhemoglobin are field dependent. Thus, a scanner with 1.5 T field strength would typically record signal changes with functional activation of about 0.25 to 5 percent, while at 3 T and 7 T scanning changes even higher than 25 percent have been observed. The main advantage of novel ultrafast imaging, especially using multiband parallel acquisition, is that the time course of signal change can be observed and multislice imaging can be carried out in a reasonable time period.

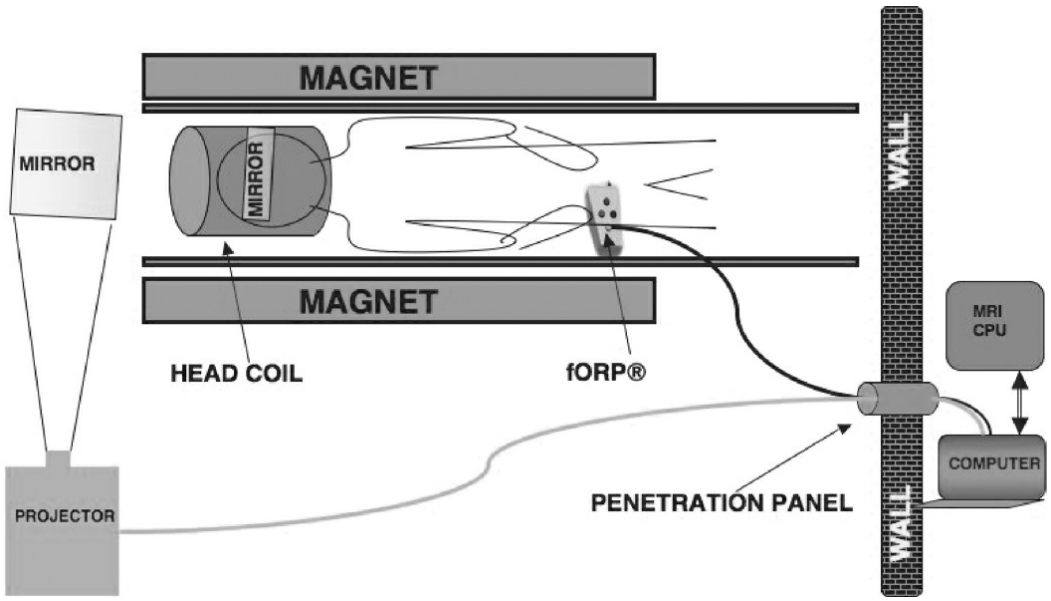

FIGURE 12.8–1. A schematic functional magnetic resonance image setup. fORP, fiber-optic response pad.

BOLD studies have initially relied primarily on blocked conditions where signal change was integrated over periods ranging from 20 seconds to minutes. In such designs, a sequence of stimuli is presented and either the type of stimulus changes in blocks or the task may change. Signal is contrasted between different blocks. For example, in a study investigating neural substrates of identifying facial expressions, a sequence of faces is presented, separated by a control stimulus. The face stimuli repeat throughout the study, but the task changes creating four blocks or

conditions: (1) Push button for happy; (2) Push button for sad; (3) Push button for anger; (4) Push button for fear. Comparing signals integrated across face trials to that of the control stimulus will identify brain regions recruited for facial affect identification. Contrasting each condition with baseline and with each other would highlight activation specific for each block of a specific target emotion (Fig. 12.8–2).

Event-Related fMRI

This method is a variant of the BOLD technology where, rather than aggregating the tasks into blocks, an estimate of the hemodynamic response is obtained by interspersing the appearance of stimuli and contrasting the signal following stimulus presentation to that following a control stimulus. These contrasts provide information on brain regions whose activation is time locked to the appearance of the specific stimulus class. For example, if the task is to push the button for a specific emotion such as fear, the question may be asked: Which brain regions activate every time a fearful face appears, compared to when the face is neutral? This is a different question from that posed by the blocked analysis. The condition is the same, searching for fearful expressions that are intermixed with other emotions and with neutral faces. However, the event-related analysis will identify those brain regions within the facial affect–processing network that are specifically engaged when a fearful face appears rather than a neutral face. Furthermore, an event-related analysis can identify brain regions whose activation is associated with correct responses and separate them from brain regions associated with errors. This feature permits a closer link of brain activation with performance. An fMRI study can be done with blocked and event-related paradigms, or, by carefully spacing stimulus classes, hybrid designs can be applied where blocked analysis can identify brain regions engaged in specific tasks, and event-related analysis can be used to establish time-locked components of the hemodynamic response related to specific stimulus classes. For example, in the facial affect identification task, the block design revealed a circuitry engaged in this task that included sensory, limbic, and evaluative components. However, the spacing of the presentation of stimuli within blocks enabled contrasting hemodynamic changes associated with the appearance of fearful faces to those associated with neutral faces, and this contrast showed greater limbic activation associated with correct identification of fearful and angry faces in healthy people (Fig. 12.8–3).

Arterial Spin Labeling

This technique provides a quantitative perfusion imaging method that uses magnetization tagging of endogenous arterial water to determine perfusion of brain tissue by comparing images obtained with and without labeling of

the arterial supply. The method is analogous to steady-state techniques used in PET since the regional signal intensity is dependent on the arterial blood flow, which delivers labeled spins, and the T1 relaxation rate, which causes the labeling to decay. This technique has the important advantage of providing quantitative cerebral blood flow parameters in physiological units (milliliter per 100 g of tissue per minute). Furthermore, perfusion is measured in brain parenchyma directly and is thus better localized than measurements obtained using an intravascular tracer, which is most sensitive to venous outflow effects. There may also be less motion sensitivity than BOLD (Fig. 12.8–4).

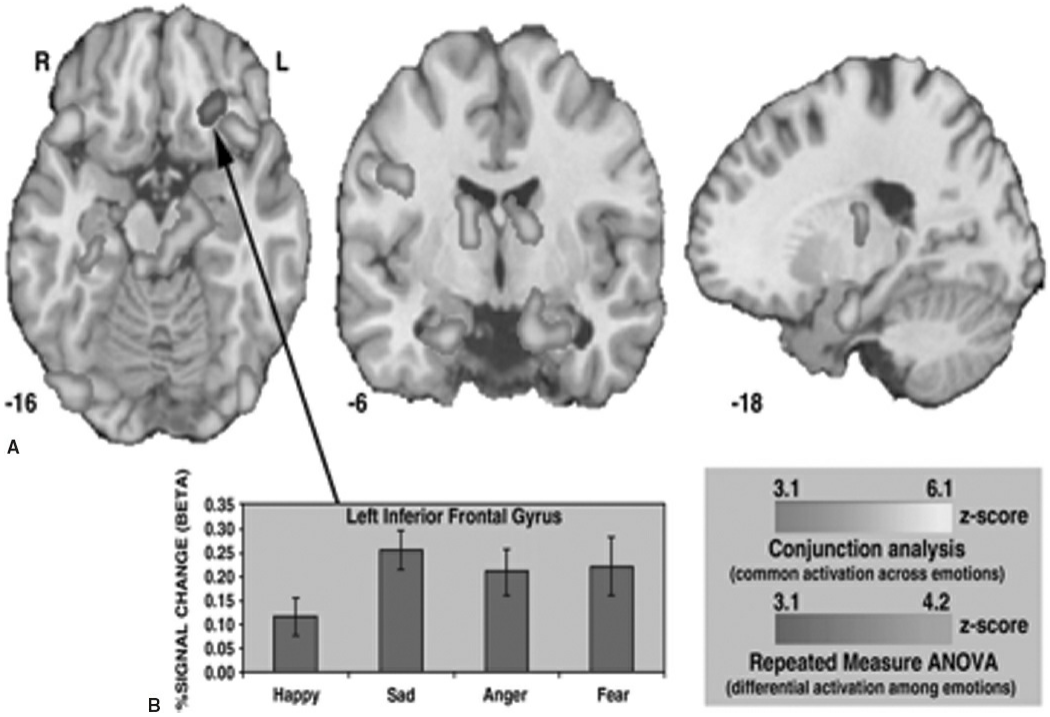

FIGURE 12.8–2. Block analysis results. **A:** Group analysis (*n* = 17) showing average brain activation for all target emotions and target emotion-sensitive activation in the inferior frontal gyrus. **B:** Mean percentage signal change for inferior frontal gyrus by target emotion. Images are displayed over a Talairach-normalized template in radiological convention (left hemisphere to viewer's right). (From Loughead J, Gur RC, Elliott M, Gur RE. Neural circuitry for accurate identification of facial emotions. *Brain Res.* 2008;1194:37–44.)

Analytic Approaches

The application of analytic methods in fMRI has been increasingly standardized and includes motion correction, high pass filtering, spatial smoothing, and scaling. Resulting translational motion parameters should be examined to ensure that there is no excessive motion. Non–brain areas should be removed and the functional images coregistered to the structural volume. This image is then usually transformed into a standard anatomical space, and transformation parameters are later applied to statistical images, after analysis is carried out in the subject's space. A more advanced

spatial normalization procedure for fMRI data involves application of the high-dimensional deformations determined from the anatomic images to the coregistered fMRI images. This provides higher registration accuracy across subjects and, therefore, potentially improves sensitivity in activation detection. This approach also allows correlating brain volumes with fMRI activity voxel-wise, as structural and functional images follow the same transformation to the standard space.

Subject-Level Time Series. This statistical analysis is carried out where condition events are modeled with a canonical hemodynamic response function and, when appropriate, its temporal derivative. Whole brain statistical analysis is completed for each individual in subject space, and resulting contrast maps, obtained through linear contrasts of event types, are spatially normalized.

Group-Level Random Effects. This analysis is performed including within- and between-group comparisons. Significance thresholds are based on spatial extent (k) and peak height (u). Often, a height threshold corresponding to an uncorrected P value of 0.005 is used, requiring a minimum cluster $P \leq 0.05$. This results in corrected probability of $P < 0.05$, based on the theory of Gaussian fields.

In addition to the whole brain analysis, a region-of-interest analysis can be performed to obtain the percentage signal change values for specific nodes in a hypothesized network. Predictor variables are a set of lagged indicators of stimulus type associated with each scan. Convolution with a priori hemodynamic response functions is not required. Resulting regression coefficients (β parameters) represent unbiased estimators of incremental hemodynamic responses at each time point, beginning at stimulus presentation and continuing for 16 seconds poststimulus. These data can be extracted for offline analysis.

Pattern Analysis and Classification. Voxel-by-voxel or region-of-interest-by-region-of-interest analysis of data is not able to capture correlations among structural and functional measurements at different brain locations. Therefore, spatially distributed and complex structural and functional patterns are better captured with high-dimensional pattern analysis and classification methods.

Connectivity Analysis. Although most fMRI studies have focused on examining brain regions activated by specific tasks, increased attention is being drawn to examining intercorrelations among brain regions. Such "connectivity analyses" can help identify the strength of association among nodes in the network and to establish in clinical populations whether deficits relate to reduced connectivity. Defined as the temporal correlations

in neural activity among brain regions, functional connectivity provides a conceptual and methodological framework for identifying spatially distributed patterns of brain activity and for quantifying interregional interactions. Several specific approaches have been implemented, and their relative merits are being investigated. The time series correlation method examines interregional correlations within individual subjects over the time course of an experiment and has been effectively applied to measure functional connectivity at a resting "default mode" state as well as across a wide range of cognitive and physiological states.

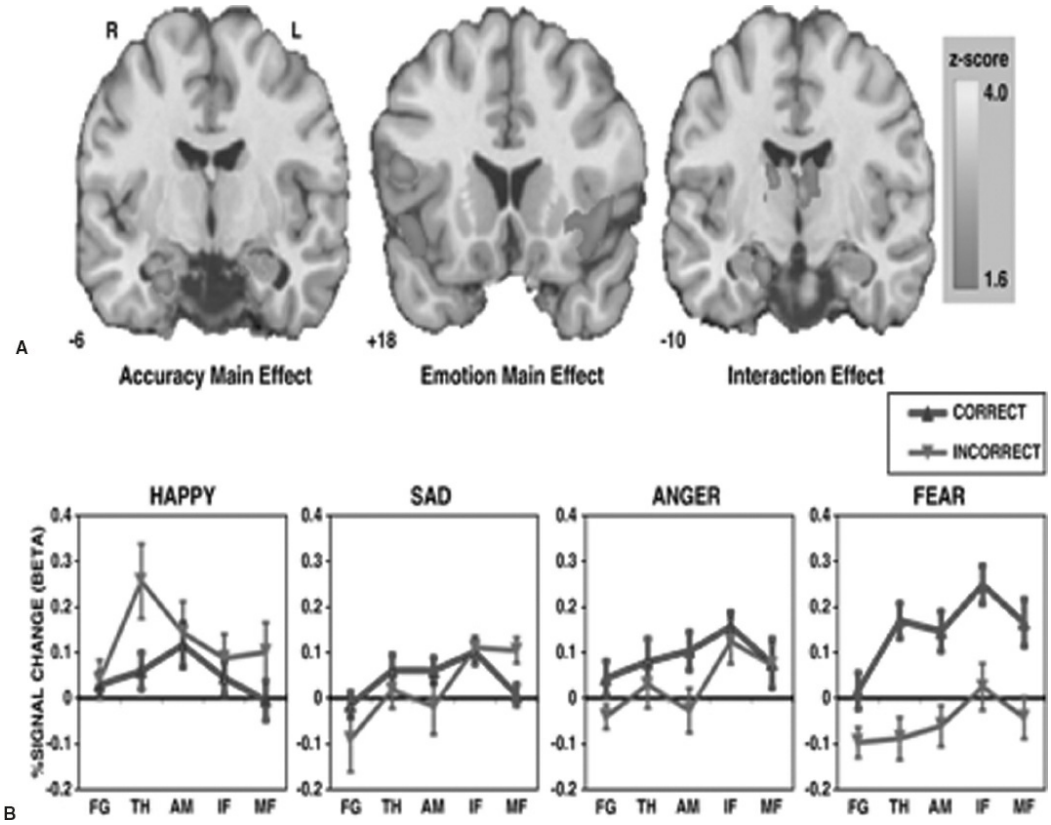

FIGURE 12.8–3. Event-related results. **A:** Emotion (happy, sad, anger, fear), accuracy (correct vs. incorrect), and interaction effects for whole brain repeated measure ANOVA highlighting mesial temporal lobe regions. **B:** Mean percentage signal change for regions showing main or interaction effects (error bars represent S.E.M. [standard error of the mean]). Images are displayed over a Talairach-normalized template in radiological convention (left hemisphere to viewer's right), masked by block activation and thresholded $P = 0.05$. (From Loughead J, Gur RC, Elliott M, Gur RE. Neural circuitry for accurate identification of facial emotions. *Brain Res.* 2008;1194:37–44.)

More recently, graph theory has been applied to the analysis of functional connectivity. This approach considers the brain as consisting of nodes ("vertices") that communicate with each other, with an index of the strength of connectivity represented by edges. The values determining significance of edges could be any index of relatedness such as coordinated fluctuations in activity measured in changes in the BOLD signal time series. Graph theory permits the identification of networks and quantitation of the

predominance and strength of within- and between-networks communications. For example, each node may show fluctuations that are synchronous not only with adjacent nodes but also with nodes that are in a homotopic region or nodes that belong to a different region entirely. Graph theory can calculate parameters of "community structure" such as modularity and cohesiveness. Graph theory has helped identify networks such as the default mode, salience, and cognitive control.

The methodological developments summarized previously for measuring regional brain activation have been embraced by investigators because of the diverse parameters of brain structure and function they offer. Furthermore, MR scanning facilities are available in academic settings around the world. Consequently, numerous studies have applied neurobehavioral paradigms in fMRI studies aimed at dissecting complex behavior in schizophrenia. This has resulted in a large body of literature, with converging results that are contributing potential biomarkers that can advance the search for dimensional underpinnings of schizophrenia along the continuum of psychosis.

FUNCTIONAL MAGNETIC RESONANCE IMAGING IN SCHIZOPHRENIA

The literature on fMRI studies in schizophrenia has continued to grow with efforts to examine samples early in the course of psychosis and across psychotic disorders. Such efforts have capitalized on functional brain abnormalities identified in schizophrenia including abnormal activity in response to motor tasks, working memory, attention, word fluency, emotion processing, and medication effects. These functional neuroimaging studies have suggested links between deficits in performance to abnormalities in activation patterns. Before summarizing such findings, it is important to outline some of the special challenges confronting the application of functional neuroimaging in schizophrenia. Although these challenges are not unique to functional imaging studies, they should be addressed systematically as the field matures.

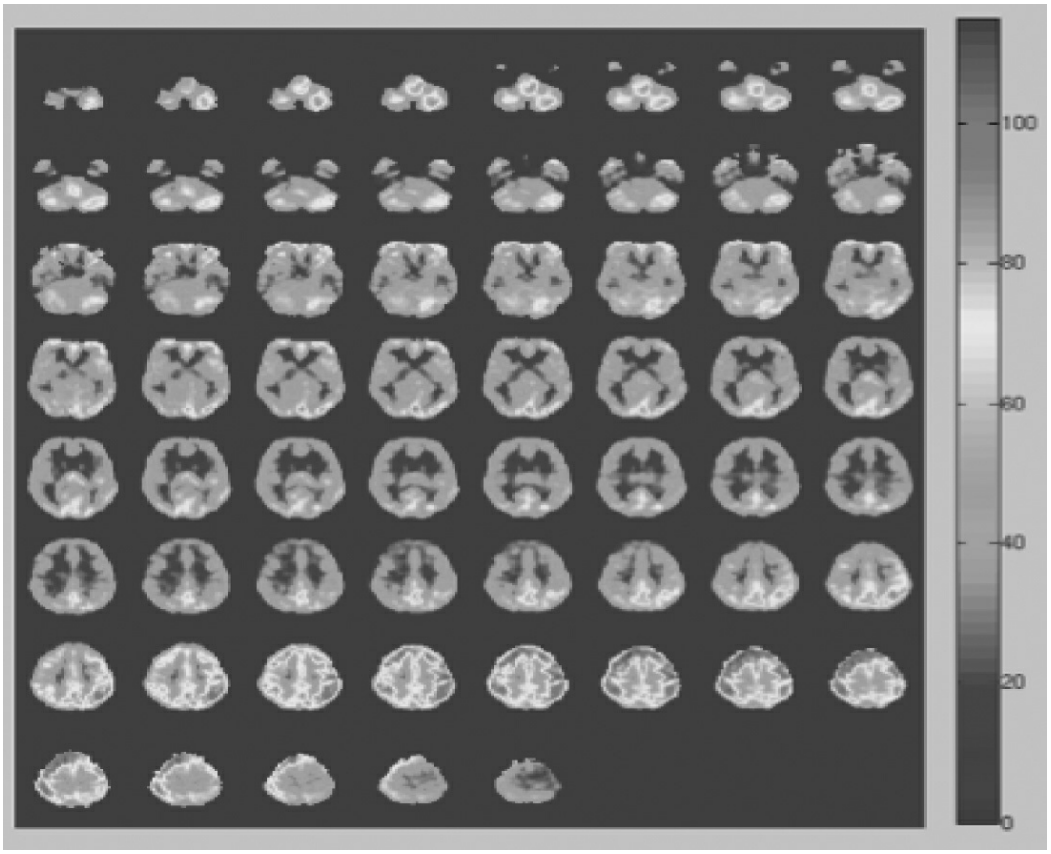

FIGURE 12.8–4. Map of quantitative blood flow obtained in a healthy individual with arterial spin labeling magnetic resonance imaging. (See color atlas.)

Special Considerations in fMRI Applications in Schizophrenia

In addition to methodological considerations that are common to all fMRI studies, several factors merit further consideration in the study of schizophrenia. These are likely to affect data interpretation and generalizability and relate broadly to several categories.

Clinical Variables

SAMPLE CHARACTERISTICS. The heterogeneity of schizophrenia can influence fMRI results. Samples vary in age, gender, and clinical status. Regarding sample size, most fMRI studies include relatively small samples that are not sufficiently powered to address pertinent clinical questions. Data on healthy people in fMRI studies are often obtained in university undergraduates, and small samples have sufficient power to detect task effects. However, in the study of schizophrenia, such samples could be marginally adequate to demonstrate differences between patients and healthy comparison participants but not informative for correlation with important clinical variables or for comparison with other clinical samples. The heterogeneity of schizophrenia may preclude adequate representation of some diagnostic subcategories, which can be further confounded by demographic characteristics. For example, patients are likely to be older

than college undergraduates, unless studied early in the course of illness or have an early age of onset. Age and gender are important moderating variables in examining the study of brain and behavior, and they are even more so in schizophrenia. Age of onset is earlier in men, and premorbid functioning and outcome are worse. These considerations entail recruitment of a comparison sample that is sociodemographically balanced to that of patients, and there must be a sufficiently large sample of patients to examine the possible effects of age and chronicity of illness on the measures obtained.

MEDICATIONS. Another relevant potential confound that needs consideration when applying functional neuroimaging in schizophrenia is the medication status of the participants. Unless studied at first presentation, when neuroleptic-naïve, patients are likely to be treated with psychoactive medications that affect brain function, and it is unclear how these effects will be reflected in BOLD signal changes. Short of controlled clinical trials, there are no simple and ethically acceptable ways to study or control potential medication effects. It is important to document medication status and examine, to the extent possible, whether they can account for the differences observed between patients and controls.

SYMPTOM SEVERITY. The clinical status of patients with schizophrenia is a factor that, in addition to influencing the feasibility of participation, may also compound the complexity of data interpretation. This factor is especially relevant when attempting to link imaging data to clinical measures of symptomatology. For example, men are likely to have more prominent negative symptoms than women, and, therefore, samples need to be adequately powered to clarify this potential confound. Obviously, the exclusion of acutely ill patients may limit the generalizability of data to patients who are clinically stable or who have more negative symptoms.

PERFORMANCE. A critical yet often overlooked issue in understanding results of functional neuroimaging studies that involve neurobehavioral probes is how to deal with performance. The typical goal of tasks in functional imaging studies with healthy people is to activate a specific neural circuitry, and performance is usually required to be perfect or is used as a “nuisance” variable. When the aim is to identify a set of regions that gets uniquely activated for a specific behavioral process or domain, making the task difficult will introduce individual differences that can diminish experimental effects and may engage extraneous regions related to frustration, anxiety, or motivation. This methodological requirement for tasks that do not produce variable performance becomes problematic in studies in which the goal is to examine neural substrates for behavioral deficits. Thus, when comparing patients with schizophrenia to healthy controls on tasks where patients are putatively impaired, should it still be

insisted that performance be perfect for patients? Such methodological considerations have compelled some investigators to seek ways for equating difficulty levels when contrasting activation between patients and controls by making an easier version of the task for patients through titrating performance. This practice, however, may result in comparing activation maps between groups who received different tasks. Other approaches include changing the level of prescanning training to maximize patients' performance. However, such practices could result in group differences due to habituation.

Arguably, different paradigms are appropriate at different stages of the research. To establish whether the neural system under investigation engages the same regions in patients with schizophrenia as in controls, it would make sense, even when investigating a domain in which impairment is expected in patients, to make the task so easy that patients and controls will perform near perfection. Failure of patients to activate a specific region under these conditions will indicate that they do not recruit the requisite circuitry for that domain, which may explain why they fail when the going gets tough. However, in the next phase of the research, it would be important to make the task harder so that individual differences in activation could be correlated with individual differences in performance both in patients and in control groups.

Event-related paradigms can help address this issue to some extent because they enable separation of correct from incorrect responses, thereby "leveling the playing field" with regard to performance. For example, on the emotion identification task described previously, a blocked analysis showed that patients with schizophrenia activated the same regions as controls but to a lesser degree (Fig. 12.8–5). An event-related analysis, however, showed a more dramatic difference when comparing activation time linked to correct and incorrect responses. Activation was similar in patients and controls for correct responses. Yet, for threat-related expressions (angry and fearful faces), unlike controls in which greater limbic activation portended correct identification, in patients greater activation in these regions was associated with incorrect responses (Fig. 12.8–6).

Major Findings in fMRI Studies in Schizophrenia

Cognition. Impaired cognition in schizophrenia has stimulated efforts to examine brain processes related to information processing cascades. Diffuse deficits have been documented, with relatively greater impairment in executive functions and in learning and memory. These deficits have been related to frontotemporal systems. Such impairments are core features of schizophrenia, important for elucidating underlying mechanisms, as buttressed by their prominence at initial illness presentation, limited amelioration with symptomatic relief, link to functional outcome, and utility as endophenotypes or biomarkers

(quantitative intermediate phenotypic measures).

Neurobehavioral probes with functional imaging provide a powerful method for exploring the neural circuitry underlying such observed deficits. Abnormal activations in ventromedial and superior temporal lobe, prefrontal cortices, and limbic structures were documented with memory and executive tasks. These complex processes could be compromised as downstream effects of sensory integration deficits.

Verbal learning deficits are well established in schizophrenia. They are prominent at early clinical presentation, persist despite therapeutic intervention and amelioration of symptoms, and relate to functional outcome. The relevance of learning and memory deficits in schizophrenia provides an impetus to integrate well-developed paradigms in cognitive neuroscience to advance the understanding of the neurobiology of these deficits.

The application of fMRI enables researchers to examine distinct memory processes such as encoding and retrieval. Studies in patients, compared to healthy participants, have consistently demonstrated that during the learning phase, there are frontotemporal abnormalities. Most studies in schizophrenia have reported decreased activation of the frontal cortex, especially the inferior prefrontal region. The data are less consistent with regard to the temporal lobe. Most studies have observed decreased activation in patients' hippocampus and parahippocampal gyrus, but other studies have noted increased activation. As emphasized earlier, performance may be a contributing factor as the studies differ in task difficulty and their approach to performance variance.

Examining frontal lobe function, event-related fMRI studies have shown deficits in working memory and cognitive control in schizophrenia. Studies that focus on the temporal lobe in schizophrenia have progressed from language-processing tasks to examination of the hippocampus and parahippocampal gyrus and probes evaluating long-term memory. Such reports noted that hippocampal dysfunction was often concomitant with aberrations in prefrontal function, suggesting that frontotemporal connectivity is disrupted in schizophrenia. As summarized previously, advances in image analysis methods enable researchers to measure the connectivity of distributed brain circuitry and test hypotheses implicating specific network models. Verbal episodic memory is differentially affected in schizophrenia, in part due to use of inefficient organizational strategies. Indeed, providing a semantic organizational strategy during word encoding normalizes subsequent recognition accuracy in patients. This finding suggests that deficits in verbal memory may result in part from aberrant frontotemporal interactions, a hypothesis that fits well with the prevailing view of schizophrenia as a disorder of disturbed neural and psychological integration.

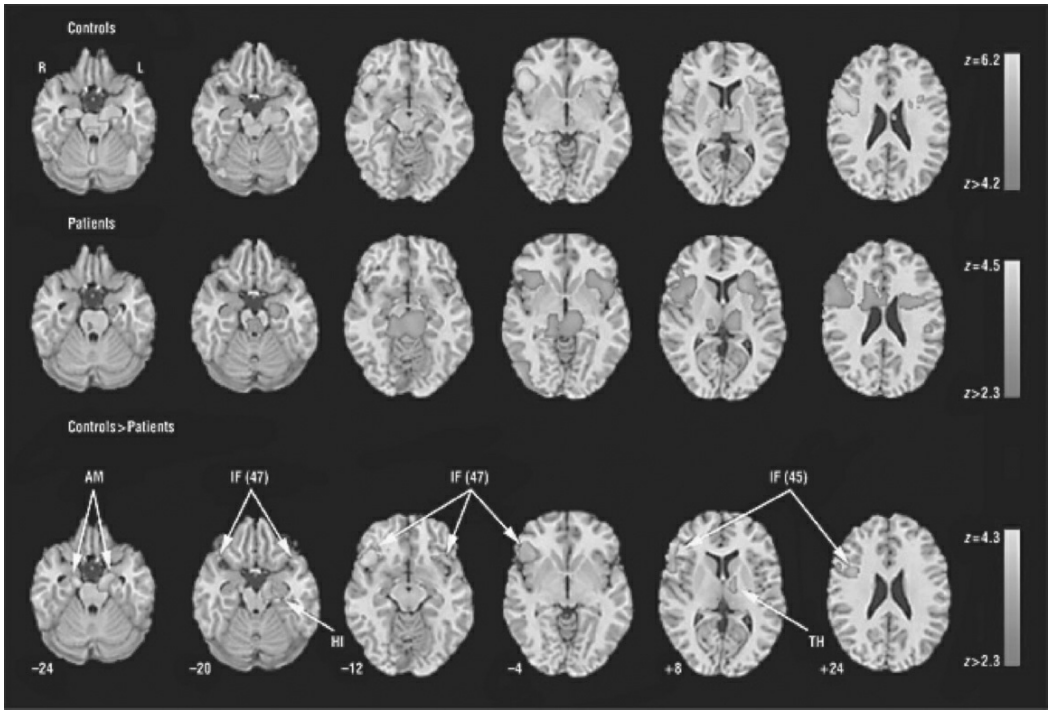

FIGURE 12.8–5. Regions activated for emotion identification task relative to baseline (block analysis) in controls (*upper row*), patients (*middle row*), and the controls–patients contrast (*bottom row*). No patients–controls contrast survived correction. Significance thresholds are based on spatial extent using a height of $z \geq 3.1$ and a cluster probability of $P \leq 0.05$. Images are displayed over a Talairach-normalized template in radiological convention (left hemisphere to viewer's right). The *z*-level coordinates are provided. AM, amygdala; IF (47), inferior frontal (Brodmann area 47); HI, hippocampus; IF (45), inferior frontal (Brodmann area 45); TH, thalamus. (From Gur RE, Loughead J, Kohler CG, Elliott MA, Lesko K. Limbic activation associated with misidentification of fearful faces and flat affect in schizophrenia. *Arch Gen Psychiatry.* 2007;64:1356–1366.)

Supporting this hypothesis, functional neuroimaging studies in schizophrenia have revealed abnormalities in both frontal and temporal activities. Earlier PET studies showed that healthy people, when performing verbal tasks, activated prefrontal regions and deactivated superior temporal gyrus. In contrast, patients with schizophrenia showed less superior temporal gyrus deactivation and a positive correlation between activity in the prefrontal and superior temporal gyrus regions. Further evidence of reciprocal changes in left frontal and left temporal regions was provided in fMRI studies of word fluency and in studies of word encoding and recognition.

Defined as the temporal correlations in neural activity among brain regions, functional connectivity provides a conceptual and methodological framework for identifying spatially distributed patterns of brain activity and for quantifying interregional interactions. Several specific approaches have been implemented as outlined previously, and their relative merits have not been established. The time series correlation method examines interregional correlations within individual subjects over the time course of an experiment and has been effectively applied to measure functional

connectivity across a wide range of cognitive and physiological states.

In a study examining within-subject correlations of frontal and temporal time series during verbal encoding, aberrant frontotemporal connectivity was noted in schizophrenia, confirming previous reports. The study also identified distinct alterations within dorsal and ventral prefrontal cortex. Relative to healthy controls, patients with schizophrenia had reduced connectivity between the dorsolateral prefrontal cortex and temporal lobe areas, including parahippocampus and superior temporal gyrus. Patients, however, showed *increased* connectivity between a region of ventrolateral prefrontal cortex and these same temporal lobe regions. Higher temporal–dorsolateral prefrontal cortex connectivity during encoding was associated with better subsequent recognition accuracy in healthy participants but not patients. Temporal–ventrolateral prefrontal cortex connectivity was uncorrelated with recognition accuracy in either group. The results suggest that reduced temporal–dorsolateral prefrontal cortex connectivity in schizophrenia could underlie encoding deficits, and increased temporal–ventrolateral prefrontal cortex connectivity may represent an ineffective compensatory effort (Fig. 12.8–7).

Notwithstanding an extensive literature applying fMRI in schizophrenia, data on early information processing are more limited. A visual stimulation study demonstrated normal activation of visual, motor, somatosensory, and supplementary motor regions to a flashing checkerboard paradigm with a motor response. However, abnormal activation was reported for more complex sensory conditions, especially when requiring integration. A block design study presented simultaneous visual and auditory stimuli to first-episode neuroleptic-naive patients. Reduced activation in parietal lobes, right thalamus, and prefrontal cortex, implicated in the dorsal visual processing pathway, was observed in patients. Diminished activation in patients in prefrontal regions implicated in regulating inhibition was reported in a study of the inhibitory P300 for a no-go condition. A three-stimulus auditory oddball task showed diffuse cortical and subcortical hypofunction during target detection and novelty processing in schizophrenia. Individuals with prodromal symptoms demonstrated smaller differential activation in frontal regions between relevant and irrelevant stimuli. In a visual oddball study, patients and healthy participants showed activation patterns specific to targets and novels, and activation of both neural systems was associated with faster performance. Specifically, reduced activation in regions involved in target and novelty processing in patients was accompanied by increased activation in circuits related to elaborated stimulus processing. For targets, abnormal activation was noted in regions related to ideational and visual association, and for novels patients overactivated sensory and frontal areas related to visual–spatial processing and working memory (Fig. 12.8–8). Thus, the attenuated electrophysiological response to targets seen in event-

related potential studies may relate to insufficient top-down activation of target circuitry, while the attenuated-evoked response to novel distractors reflects overprocessing of bottom-up events. Notably, abnormal activation in patients was associated with more severe symptoms.

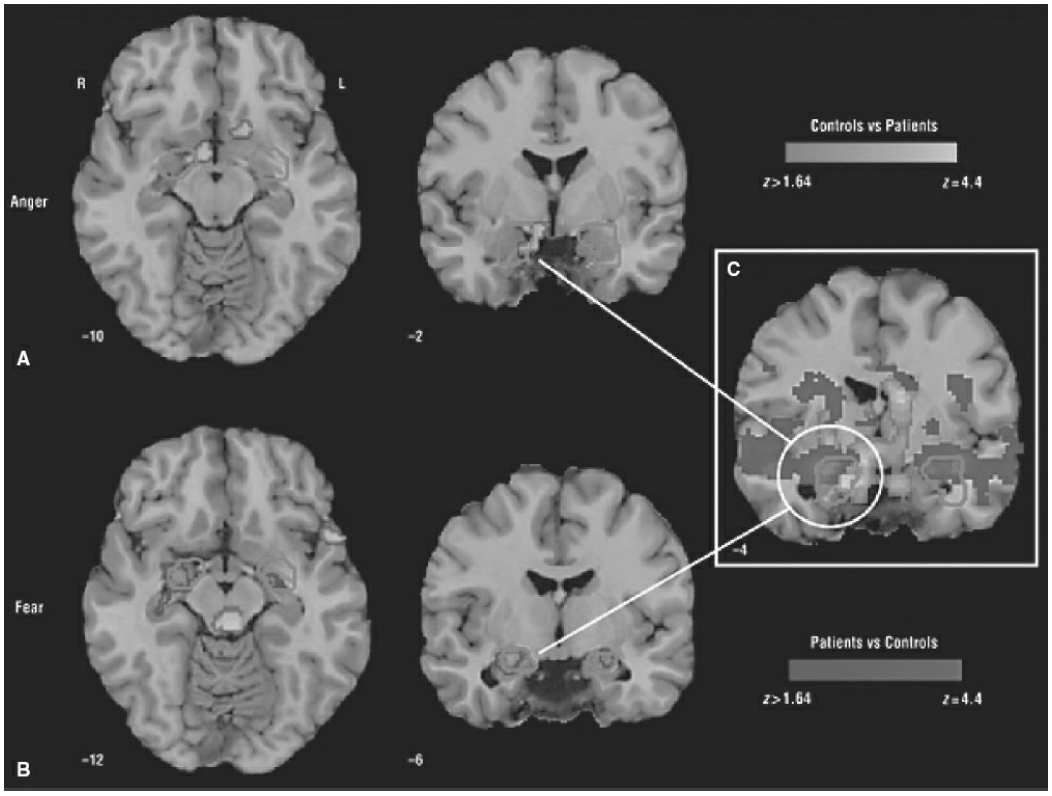

FIGURE 12.8–6. Activation maps showing peak amygdala response for anger (**A**) and fear (**B**) conditions for the event-related analysis. Images are displayed over a Talairach-normalized template in radiological convention and thresholded at $z \geq 3.1$, uncorrected (>25 continuous voxels). Outline shows extent of atlas-derived amygdala regions of interest (Wake Forest University pickatlas) used for percentage of signal change extraction. Inset (**C**) highlights "patients vs. controls" incorrect responses superimposed on "controls vs. patients" correct responses at $z \geq 1.64$, uncorrected. (From Gur RE, Loughead J, Kohler CG, Elliott MA, Lesko K. Limbic activation associated with misidentification of fearful faces and flat affect in schizophrenia. *Arch Gen Psychiatry.* 2007;64:1356–1366.)

Emotion. Social cognition, the cognitive processes involved in perception, interpretation, and processing of social information, is increasingly recognized as a core domain of deficit in schizophrenia. Neurobiological models of social cognition have confirmed that the processing of social information requires complex and synergistic interactions between several neural regions, which encompass the basic perception of social stimuli and the later evaluation of these stimuli.

The increased interest in affective neuroscience has likewise generated several tasks that probe emotion processing and social cognition. Such tasks may have relevance to schizophrenia, perhaps especially for understanding the intransigence of negative symptoms and poor social functioning. For example, functional imaging studies have demonstrated

the role of limbic response in identification of facial emotions in healthy people and diminished response in schizophrenia. As can be seen in the middle row of Figure 12.8–2, top-down (task related) limbic activation is diminished in schizophrenia. However, patients showed abnormally increased limbic activation time locked to the appearance of threat-related facial emotions of anger and fear. Furthermore, increased amygdala activation for fear was associated in patients both with failure to identify the emotion and with more severe flat affect (Fig. 12.8–9).

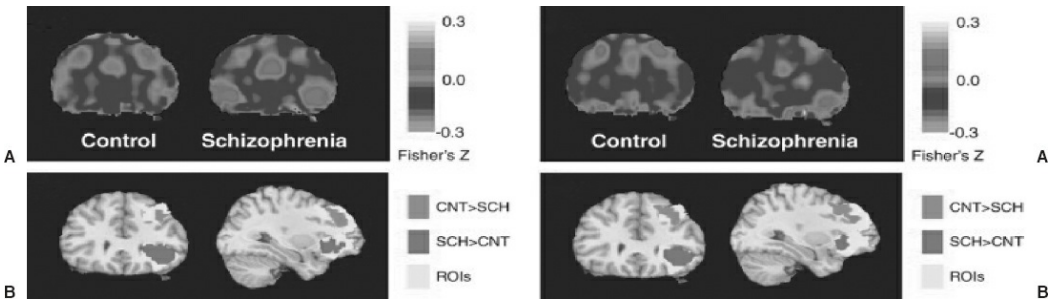

FIGURE 12.8–7. Group differences in left STG–DLPFC (superior temporal gyrus–dorsolateral prefrontal cortex) and STG–VLPFC (superior temporal gyrus–ventrolateral prefrontal cortex) connectivity (*left column*) and in left PHIP–DLPFC (parahippocampal gyrus–dorsolateral prefrontal cortex) and PHIP–VLPFC connectivity (*right column*). **A:** Left column, group-averaged STG time series correlation maps for control group (*left*) and schizophrenia group (*right*), from a representative coronal plane (y = 30). A region of DLPFC shows greater connectivity (measured as average Fisher's Z-transformed correlation coefficients) with STG in controls, while a region in VLPFC shows greater connectivity with STG in patients. **B:** Results of voxel-by-voxel t tests are overlaid on the two prefrontal regions of interests (ROIs) and a standard reference brain (Colin, MNI [Montréal Neurological Institute]), in the same coronal plane. The group differences seen within DLPFC and VLPFC are statistically significant. **A:** Right column, group-averaged PHIP time series correlation maps reveal a region within DLPFC, showing greater connectivity with PHIP in controls, while a region within VLPFC shows greater connectivity with PHIP in patients. **B:** Group differences seen within DLPFC and VLPFC are statistically significant. (From Wolf DH, Gur RC, Valdez JN, Loughead J, Elliott MA. Alterations of fronto-temporal connectivity during word encoding in schizophrenia. *Psychiatry Res.* 2007;154:221–232.)

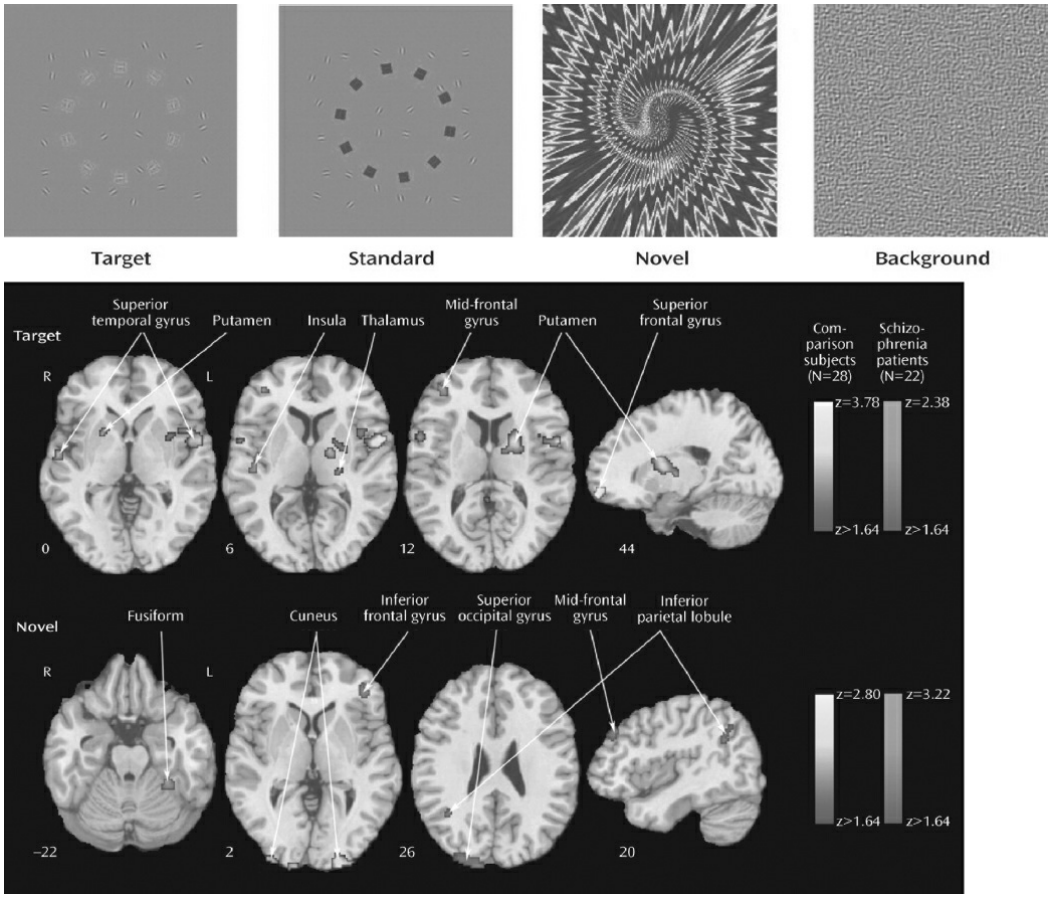

FIGURE 12.8–8. Examples of stimuli used in "oddball" studies of attentional processing (*top*) and contrast images of patients with schizophrenia and comparison subjects for target and novel stimuli. Greater activation in patients is depicted by the scale on the right, whereas greater activation in comparison subjects is shown by the scale on the left. Images are in radiological convention (left hemisphere to viewer's right). (See color atlas.) (From Gur RE, Turetsky BI, Loughead J, Snyder W, Kohler C. Visual attention circuitry in schizophrenia investigated with oddball event-related functional magnetic resonance imaging. *Am J Psychiatry.* 2007;164:442–449.)

Impaired emotional functioning is a prominent feature of schizophrenia. Although positive symptoms have traditionally attracted more attention and targeted treatment, negative symptoms, including flat affect, are increasingly recognized as more debilitating and resistant to intervention. Early studies reported impaired identification of facial affect in schizophrenia, but there was limited information on its relationship to clinical features and cognition.

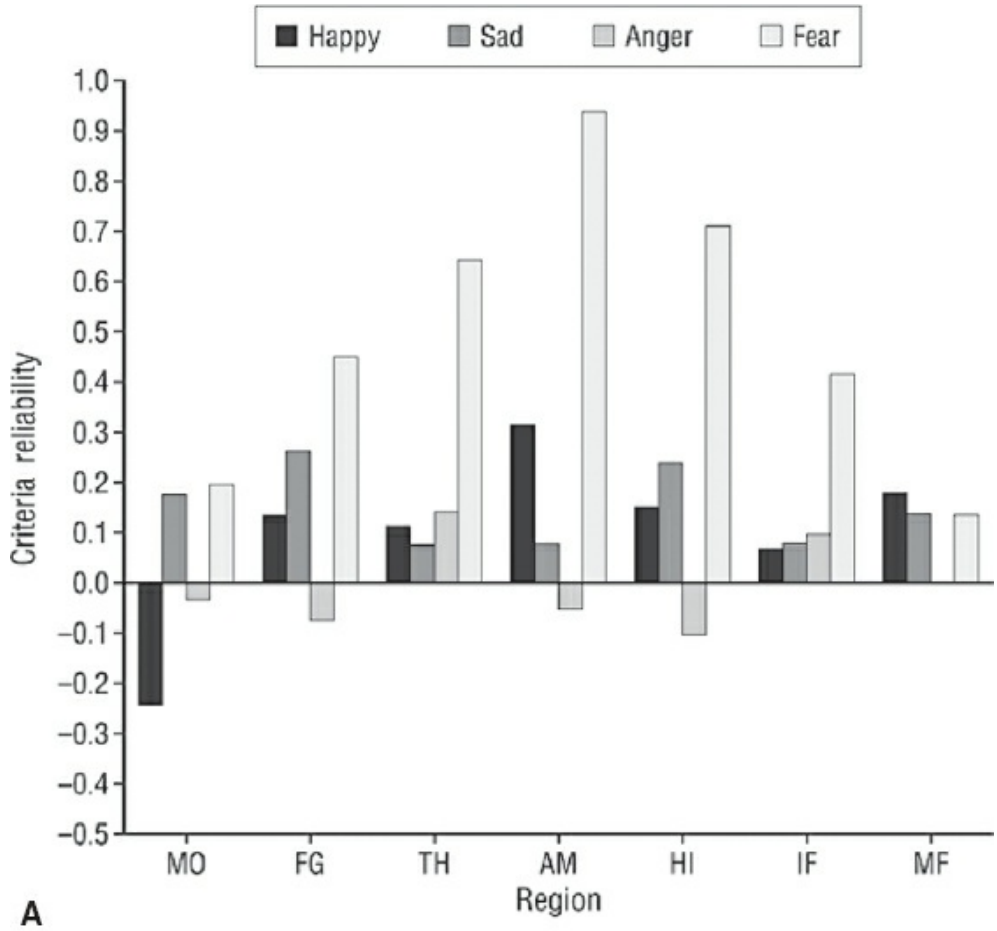
Happy
Sad
Anger
Fear
Criteria reliability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
MO
FG
TH
AM
HI
IF
MF
Region

A

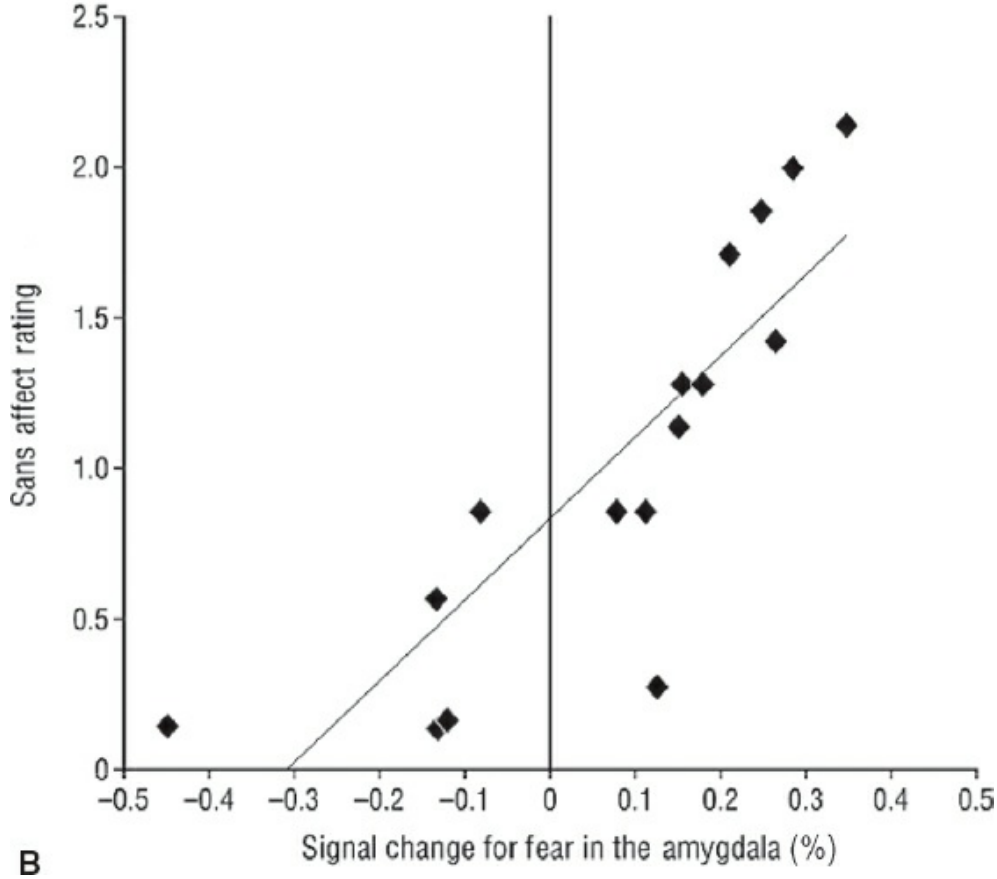
Sans affect rating
2.5
2.0
1.5
1.0
0.5
0
−0.5
−0.4
−0.3
−0.2
−0.1
0
0.1
0.2
0.3
0.4
0.5
Signal change for fear in the amygdala (%)

B

FIGURE 12.8–9. Association between brain activity and clinical measures. **A:** Correlations between event-related activation for the four emotional expressions in activated regions and severity of clinical ratings for flat affect. **B:** Scatterplot of the association between percentage of signal change for the appearance of fear expressions and severity of flat affect. MO, midoccipital; FG, fusiform gyrus; TH, thalamus; AM, amygdala; HI, hippocampus; IF, inferior frontal; MF, midfrontal. (From Gur RE, Loughead J, Kohler CG, Elliott MA, Lesko K. Limbic activation associated with misidentification of fearful faces and flat affect in schizophrenia. *Arch Gen Psychiatry.* 2007;64:1356–1366.)

Although there is a growing literature in healthy people elucidating the role of amygdala in affect processing, fewer functional neuroimaging studies have examined this circuitry in schizophrenia. Consistently in fMRI and PET studies, with emotionally valenced pictures, faces, and odors, patients showed diminished activation in limbic and paralimbic regions, especially the amygdala. Abnormal activation of related cortical regions has also been observed. Functional MRI studies with block design report lack of amygdala activation for sad mood induction, also evident in unaffected siblings. Decreased activation and accuracy in identifying expressions of fear, anger, and disgust, contrasted with responses to mild happiness, especially in nonparanoids. Presentation of fearful and neutral faces showed a disconnection between brain activity and arousal in paranoid patients, where increased arousal was associated with decreased amygdala/medial prefrontal activity. In a study using a block design alternating between emotion and age identification, patients showed decreased activation in left amygdala and bilateral hippocampus for the emotion condition. In a study examining intensity judgment, patients had an *exaggerated* amygdala response: Positive faces induced right amygdala activation in both groups, while negative faces activated only the right amygdala in controls and bilaterally in patients. Picture categorization (pleasant, unpleasant, neutral) produced less activity in amygdaloid–hippocampal and cortical–basal ganglia-thalamic circuitry in patients.

The growing fMRI literature complements PET studies that also showed abnormal cerebral blood flow in patients engaged in emotion-processing tasks. Pictures with aversive and nonaversive content activated right insula in patients and controls, but patients showed greater activation of medial prefrontal cortex for aversive stimuli. Furthermore, controls activated bilateral amygdaloid and orbitofrontal regions for nonaversive pictures, but patients activated only left orbitofrontal cortex. Activity in left amygdala correlated with positive symptoms. In a subsequent study, emotionally salient pictures produced abnormal activation in patients' ventral striatum and amygdala and increased tonic activation of the amygdala. Notably, greater abnormality was associated with less severe positive and negative symptoms. The amygdala did not activate in patients for unpleasant pictures during valence categorization. Areas of decreased cerebral blood flow included prefrontal cortex and subcortical regions such as thalamus

and cerebellum.

These studies of emotion processing in schizophrenia vary in methodology and design. Nonetheless, there seems to be considerable convergence of evidence that patients show abnormal activation in amygdala and associated regions. It also appears that the abnormalities are more pronounced for negatively valenced stimuli, although no further differentiation for specific emotions has been established. In most studies, patients' performance was carefully examined and tasks were often constructed so as to minimize performance difference to avoid confounding of physiologic measures. In all cases in which performance was evaluated, it did not explain the difference in activation patterns. However, reduced task-related (top-down) activation could reflect bottom-up interference from abnormally increased amygdala activation related to stimulus valence. Thus, if the task is to identify emotions, sensitivity of the amygdala for specific stimuli may disrupt cortical processes required for categorization and response. Increased activation associated with incorrect responding was also seen for faces expressing anger. Thus, it appears that threat-related faces induce greater activation in patients in a way that interferes with the top-down process, thereby producing errors. Whenever this response differed between patients and controls, it indicated more error-associated activation for fear or anger in patients.

Bottom-up interference in cognitive functioning likely operates by disrupting the coordinated activation of regions required for its support. Although there is considerable work on conditions associated with task activation, recent efforts have been aimed at studying interregional connectivity. These studies have used structural equation modeling that permits assessment of whether specific causal models linking activation of one region to another are not rejected by the data. To examine connectivity, it is useful to consider the pattern of correlations among activation parameters obtained from regions recruited for a specific task. These correlations would reflect the extent to which activation in one region is proportional to that in another. Controls have high specific connectivity for activation to the top-down task and for bottom-up activation associated with correct responding. Patients, by contrast, show abnormally high connectivity associated with activation to threat-related stimuli, leading to incorrect responding. Thus, disrupted operation of a top-down process in schizophrenia could be linked to excessive connectivity for bottom-up amygdala activation related to stimulus valence. Such disruptive activation could not only reduce connectivity among regions participating in the top-down process but also trigger recruitment of other regions, including cortical ones, that may further disrupt task-related processing. Therefore, abnormal amygdala activation will lead to dysfunctional activation in other regions involved in information processing and disrupt downstream processes involving complex behavior. This will manifest in cognitive

deficits.

Ultimately, the importance of abnormal activation depends on whether it relates to disease variables. This question has received less attention and results diverge. Studies show abnormally increased amygdala activation associated with more severe positive symptoms, both positive and negative symptoms, or less severe symptoms. It may be that the direction of such correlations depends on whether the abnormality is that of over- or underactivation and whether the activation potentiates symptoms or reflects a compensatory process. Although correlations between activation parameters and symptom severity are small or moderate, abnormally increased bottom-up amygdala activation for fearful faces was correlated highly with increased severity of affective flattening. Determining such correlations could help guide efforts to ameliorate symptoms with knowledge of regions where activation change can be palliative.

A likely specific consequence of abnormal bottom-up activation of amygdala is the disruption of memory processes. There is considerable evidence that the amygdala interacts closely with the hippocampus in the formation of episodic memory, not only when affective valence is prominent. This area needs further investigation.

The convergence of findings pointing to the amygdala and its connectivity as a likely source for the devastating effects of schizophrenia propels a search for factors that could modulate amygdala response. Fortunately, there is an expansion of rigorous studies with fMRI in healthy people that have attempted to identify such modulators in efforts to construct mechanistic models of affect regulation. Emerging findings from meta-analyses support the notion of a fear-specific response of the amygdala and partially support the laterality hypothesis with increased left amygdala activation across tasks. Potential contributions of modulating processes such as attention, habituation, arousal and anxiety, learning, and extinction have been examined. Overall, while the amygdala response to fearful expressions can be influenced by such factors, the response is quite robust and specific for fear and could be neither explained nor completely extinguished by such manipulations. There is evidence that fear-related stimuli are especially potent activators of the amygdala, although cognitive processing modulates stimulus effects. The amygdala's extensive connections with subcortical regions are likewise essential for its role in affect processing. It has been proposed that rapid visual input from a subcortical thalamic pathway provides an early alarm signal for threat detection, whereas slower input through a geniculostriate pathway enables more detailed evaluation of environmental input. This slower pathway could explain why amygdala activation seems to habituate rapidly in passive viewing but shows no signs of habituation when the task involves top-down processing where affective valence is relevant.

Studies attempting to elucidate mechanisms underlying fear response

have often used faces to link with animal evolutionary approaches. Affective information on the face was systematically manipulated by methods including spatial frequency, binocular suppression, chimeric faces, masked fearful eye whites, and gaze direction. These methods have suggested that amygdala response can be potentiated or suppressed by such factors. For example, amygdala response is stronger when information is presented in low rather than high spatial frequency, and the effect is especially pronounced for fear in the upper part of the face, presumably the eyes region. Notably, bilateral amygdala lesions specifically diminish a patient's ability to utilize fear-related information in the eyes region and directing the patient to note that the eyes could temporarily reverse this deficit. This finding is encouraging in that rehabilitation could possibly ameliorate deficits.

Gaze direction is another modulator of amygdala response. In healthy people, directed fear and averted anger expressions produced greater amygdala response than averted fear and directed anger. This effect was interpreted to suggest that the amygdala is especially sensitive to threat-related ambiguity. Since poor eye contact is a major feature of impaired social interactions in schizophrenia, gaze direction and evaluation of the patient's ability to use information in the eyes region seem the most appropriate target for the next phase of investigation. Although scan path studies reported a restricted range of visual scanning in schizophrenia for happy, sad, and neutral faces, these studies have not indicated a difference between the eyes and the mouth regions. However, fear and anger stimuli have not been used. Because amygdala damage leads to avoidance of the eyes region in fearful faces, and the eyes region is important for distinguishing anger from fear, the difficulty of patients in identifying fear could be related to impairment in recognizing changes in the eyes region, thereby diminishing amygdala response to fear in the eyes relative to the mouth region. As with other bottom-up activation abnormalities, it is expected that the severity of these abnormalities is associated with poorer eye contact and affective flattening.

Pharmacology. The literature on the possible effects of antipsychotic agents on BOLD signal change is limited. Few studies have examined neuroleptic-naive patients, let alone in a repeated measures design. Furthermore, small samples and methodologically limited designs have precluded systematic examination of the possible effects of therapeutics. Nonetheless, the recognition that impaired cognition is a hallmark of schizophrenia and that fMRI enables researchers to examine the brain systems underlying the aberrant functioning has extended studies of cognition to fMRI.

Examples of studies include a working memory evaluation with an n-back paradigm after patients switched from first-generation antipsychotics

to a second-generation agent. Increased dorsolateral prefrontal cortex and parietal cortex activity were reported in the patients with the new treatment. Similarly, normalization of brain activity was reported in patients treated with long-acting risperidone (Risperdal) compared to conventional depot medication while performing an n-back task. Normalization of prepulse inhibition, which is impaired in schizophrenia, was also noted in patients treated with olanzapine (Zyprexa) and risperidone compared to those on first-generation antipsychotics. Studies that examined motor control in relation to medications have also noted improved brain activity in patients treated with second-generation agents compared to first-generation antipsychotics.

The well-established literature on heritability of endophenotypic brain behavior measures in unaffected family members of individuals with schizophrenia can provide a path to investigate the potential effect of psychoactive medications on brain function in unmedicated family members. For example, an abnormally strong alprazolam-induced reduction in amygdala and hippocampus activation during emotion identification was observed in first-degree family members relative to healthy controls, compared to equal reduction in both groups for the emotion memory task. An aberrant response may be related to unmasking underlying amygdalar and GABAergic abnormalities. Such pharmacological fMRI paradigms could prove useful for developing drugs targeting specific neural circuits in schizophrenia.

Given the limited number and scope of the available literature, there is no conclusive evidence yet as double-blind studies are lacking and most reports consist of cross-sectional design. Thus, it is difficult to conclude that the reported "normalizing" effects of the medications are indeed attributable to specific medication effects and are not part of clinical stabilization or practice effects. Furthermore, as no absolute measures of brain activity are provided with BOLD, the relative change needs to be interpreted with caution. Pharmacologic fMRI is a potentially powerful methodology when integrated with well-designed and informative pharmacologic designs.

Genetics. The genomic architecture of schizophrenia has been investigated in large samples that have relied on the clinical phenotype with more recent efforts including endophenotypes. As detailed in other chapters, it has become increasingly clear that this heritable, heterogeneous disorder does not have shared genetic contributions of relatively few genes. Thus, in samples of more than 20,000 cases and controls per group, 13 risk alleles estimating that about 6,000 to 10,000 independent and largely common SNPs contribute to the heritability and etiology of schizophrenia. Further efforts by the Schizophrenia Working Group of the Psychiatric Genomics Consortium identified 108 loci with

small effects associated with the disorder and, more recently, 128 established and novel loci have been identified (Schizophrenia Working Group of the Psychiatric Genomics Consortium).

The application of fMRI in genetic paradigms is more limited relative to structural MRI. Initial studies have examined genetic effects indirectly by comparing activation patterns in probands with schizophrenia, unaffected siblings, and healthy comparison subjects with no family history of schizophrenia. Such studies have demonstrated, for example, that there were abnormalities in siblings that were less severe than those seen in affected individuals. Such findings support the application of fMRI as a quantitative phenotypic marker of schizophrenia.

The availability of sequencing of the human genome offers unprecedented opportunities for direct evaluation of the effects of genetic variability on brain activity. Early work exploiting this potential has demonstrated such effects in healthy people by comparing activation patterns between genotypically characterized groups. Studies applying genetic strategies used functional polymorphisms to group individuals for comparisons. For example, a common Val108/158Met substitution in the gene for catechol-O-methyltransferase (COMT) leads to decreased activity of this enzyme in dopamine catabolism and has been linked to decreased prefrontal cortical activity. Studies, therefore, examined val/met polymorphism and prefrontal cortex activation. Individuals homozygous for met allele had diminished prefrontal and hippocampal engagement while performing episodic memory encoding and retrieval compared to val/val subjects. Notably, however, a recent meta-analysis of fMRI working memory studies in people with schizophrenia and healthy controls did not find evidence for an impact of COMT Val158Met genotype on brain activation. Diverse analytic approaches and small sample sizes are likely to contribute to inconsistent findings.

Using a similar approach, a risk haplotype was examined in *GRM3,* a gene encoding a metabotropic glutamate receptor. The findings were of reduced neuronal function in prefrontal cortex and impaired activation in the hippocampus during performance of a verbal memory task. The risk allele in the promoter region was associated with decreased activation of prefrontal and temporal lobe cortex.

The incorporation of functional neuroimaging into genomic studies has potential for constructing mechanistic models for the pathophysiology of schizophrenia but poses additional challenges. Since the expected effects of any single polymorphism are small, it is essential to pay close attention to experimental design in neuroimaging genomics. It is important to use well-characterized neurobehavioral probes to evaluate and control confounding variables such as age, gender, basal abilities, performance, and clinical status. In addition, neuroimaging genomic studies require large samples that can optimally be obtained by multisite studies. Such efforts introduce

the additional variability related to scanner characteristic and image acquisition. Nonetheless, the field is positioned to make new discoveries that will illuminate neural pathways to specific features of schizophrenia in ways that will lead to novel interventions.

Functional Neuroimaging in Clinical At-Risk Samples

The investigation of schizophrenia as a brain disorder began with studies that evaluated chronic patients, establishing neurocognitive deficits associated with abnormalities in parameters of brain structure and function. With the increased availability of noninvasive neuroimaging methodology, studies of younger patients became feasible, and this work indicated that the brain abnormalities and neurocognitive deficits are evident already in first-episode individuals. This work has produced two shifts of emphasis in the study of psychosis that can advance the understanding of the underlying neurobiology. The first is a shift to study even earlier phases of psychosis—the clinical risk stage; the second, a shift from dichotomous symptom categorization to quantitative dimensional endophenotypes, which can be examined across psychosis. This development will benefit genomic investigations because such vulnerability indices are heritable brain and neurobehavioral parameters that are trait markers related to illness.

Functional MRI has been applied to individuals at risk for psychosis, commonly in small samples of people at genetic risk because of family history or help-seeking individuals with attenuated psychosis spectrum symptoms. Neurobehavioral probes applied were those that have shown differences between patients with schizophrenia and controls such as working memory using the n-back paradigm. Overall, psychosis risk groups show decreased activation in the BOLD response in dorsolateral and medial prefrontal regions. The pattern of activity is similar to that seen early in the course of schizophrenia, but less pronounced abnormalities are observed. Similar to data obtained from patients with schizophrenia, individuals at risk show *increased* amygdala activation for identification of threat-related expressions (Fig. 12.8–10).

To evaluate activation changes with disease progression, longitudinal designs are necessary. Such designs have been applied in several fMRI studies. This small literature suggests that individuals who transition to psychosis differ from those who do not, with the latter group showing normalization of brain activity. Thus, the application of fMRI holds promise as a tool that may facilitate elucidating the brain circuitry dysfunction underlying the psychotic process.

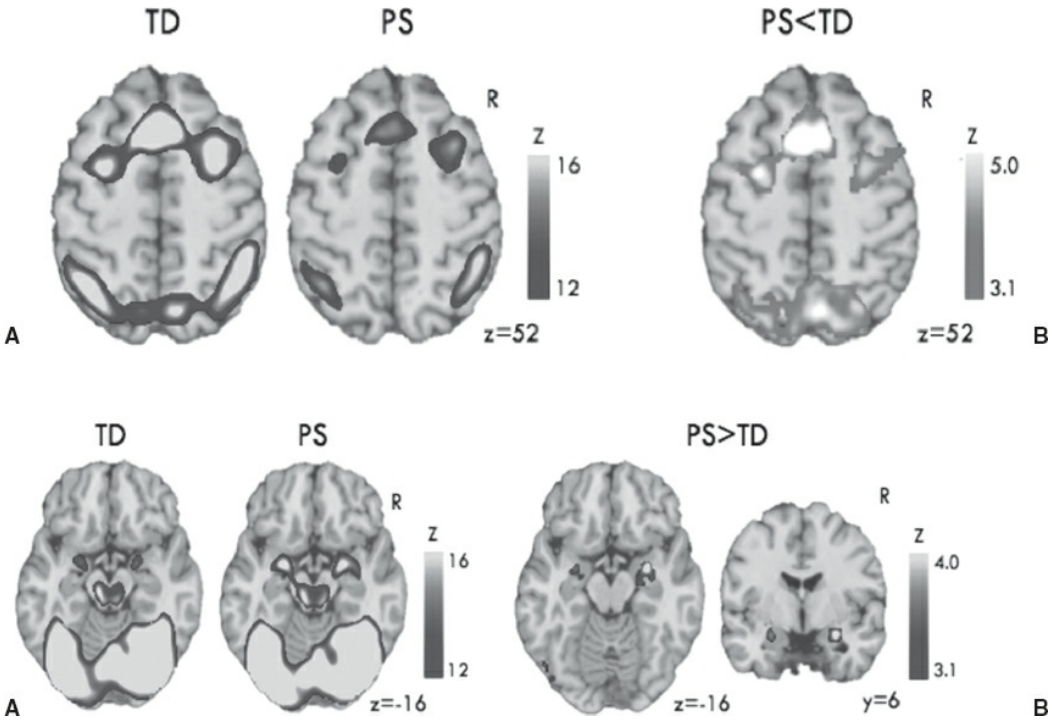

FIGURE 12.8–10. Aberrant activation in youth with psychosis spectrum symptoms. Top panel shows the working memory n-back task effect and bottom panel shows effects of emotion identification task. For top panel: **A:** The contrast evaluating the effect of working memory load (2-back greater than 0-back) robustly recruited executive function regions, including dorsal anterior cingulate and bilateral dorsolateral prefrontal cortex, with a qualitatively similar regional pattern in both the psychosis spectrum (PS) and typically developing (TD) groups. **B:** The PS group showed significantly reduced activation in these regions (significant clusters shown in light gray). R indicates right. For bottom panel: **A:** Facial emotion-processing activated amygdala and other regions implicated in emotion and face processing, with a qualitatively similar regional pattern in both the PS and TD groups. **B:** The PS group showed significantly greater activation in bilateral amygdala (significant clusters shown in black). R indicates right. (From Wolf DH, Satterthwaite TD, Calkins ME, et al. Functional neuroimaging abnormalities in youth with psychosis spectrum symptoms. *JAMA Psychiatry.* 2015;72:456–465.)

Furthermore, prior fMRI studies of psychosis risk have typically examined only categorical effects, lacking adequate power to identify correlates of illness dimensions. Therefore, it remains unclear to what extent the reported abnormalities relate to positive psychotic symptoms that are used to select the samples versus negative symptoms or cognitive deficits that are likely also present.

Dysconnectivity of resting BOLD has been reported in several brain networks in adults with schizophrenia, including the default mode and the cingulo-opercular network. More recently, youth at risk for psychosis showed multifocal dysconnectivity, implicating the bilateral anterior cingulate, frontal pole, medial temporal lobe, opercular cortex, and right orbitofrontal cortex. These results were driven by hyperconnectivity among default mode regions and diminished connectivity among cingulo-opercular regions, as well as diminished coupling between frontal and default mode regions. Thus, functional dysconnectivity in psychosis

spectrum youths corresponds to abnormalities reported in adults with established psychotic disorders.

Challenges for Functional Neuroimaging Research in Schizophrenia

Considerable advances with functional neuroimaging technologies have been made in efforts to understand the neurobiology of schizophrenia. fMRI has become the dominant method to examine brain systems because it provides noninvasive measures with good anatomic resolution and increasingly reliable quantitation. Abnormalities in schizophrenia have been documented at many levels of neurobehavioral and neurophysiological processing and in multiple neural systems. However, these diffuse abnormalities, which implicate neurodevelopmental aberrations, are not fully expressed as a disorder until after brain development and maturation is largely completed.

The recognition that the underlying pathological processes of schizophrenia are present throughout the central nervous system (CNS) suggests that there may be some core defects in neurotransmission whereby the ability of neurons to respond to and process stimuli is curtailed. However, within the context of global abnormalities, some neural systems are differentially affected. Brain regions with greater vulnerability, such as the hippocampus, prefrontal cortex, and olfactory system, are dynamic and maintain a high degree of plasticity during brain maturation and thereafter. Thus, neuronal plasticity, evident in molecular, chemical, morphological, and physiological modifiability with neural activity, is abnormal in schizophrenia. This may ultimately be manifested in the array of complex features that characterize schizophrenia. However, functions mediated through other neural systems, where greater plasticity is not required, show subtle abnormalities, and motor abilities, sensation, and autonomic functions are relatively spared.

These challenges are further compounded by the absence of suitable animal models for the disorder as phenotypically defined, buttressing the need to identify endophenotypes that can be studied in both humans and animals. For example, sensorimotor gating and encoding can be characterized in animals and in humans, as can fear conditioning and affiliative behavior. The identification of such quantitative phenotypes in schizophrenia that have their counterparts in rodents enables researchers to investigate the genetic, molecular, biochemical, anatomic, and physiological aspects of these behaviors in ways not possible in the human. Findings from studies in rodents will provide new information about these behaviors and elucidate the effects of genetic, pharmacological, and behavioral manipulations, which can then be pursued in humans. This top-down–bottom-up interchange holds great promise for advancing the understanding of the pathophysiology of schizophrenia, where functional neuroimaging provides a bridge between the molecular and the system

levels.

REFERENCES

Anticevic A, Repovs G, Barch DM. Working memory encoding and maintenance deficits in schizophrenia: neural evidence for activation and deactivation abnormalities. *Schizophr Bull*. 2013;39:168–178.

Carter CS, Barch DM. CNTRICS Executive Committee: imaging biomarkers for treatment development for impaired cognition: report of the sixth CNTRICS meeting: biomarkers recommended for further development. *Schizophr Bull*. 2012;38:26–33.

*Dickinson D, Straub RE, Trampush JW, et al. Differential effects of common variants in SCN2A on general cognitive ability, brain physiology, and messenger RNA expression in schizophrenia cases and control individuals. *JAMA Psychiatry*. 2014;71:647–656.

Ford JM, Palzes VA, Roach BJ, et al. Functional Imaging Biomedical Informatics Research Network, Mathalon DH: visual hallucinations are associated with hyperconnectivity between the amygdala and visual cortex in people with a diagnosis of schizophrenia. *Schizophr Bull*. 2015;41:223–232.

*Fornito A, Bullmore ET. Connectomics: a new paradigm for understanding brain disease. *Eur Neuropsychopharmacol*. 2015;25:733–748.

*Fusar-Poli P, Borgwardt S, Bechdolf A, et al. The psychosis high-risk state: a comprehensive state-of-the-art review. *JAMA Psychiatry*. 2013;70:107–120.

Grimm O, Heinz A, Walter H, et al. Striatal response to reward anticipation: evidence for a systems-level intermediate phenotype for schizophrenia. *JAMA Psychiatry*. 2014;71:531–539.

Gur RE, Gur RC. Functional magnetic resonance imaging in schizophrenia. *Dialogues Clin Neurosci*. 2010;12:333–343.

Hager BM, Keshavan MS. Neuroimaging biomarkers for psychosis. *Curr Behav Neurosci Rep*. 2015;2015:1–10.

*Libby LA, Ragland JD. FMRI as a measure of cognition related brain circuitry in schizophrenia. *Curr Top Behav Neurosci*. 2012;11:253–267.

Lui S, Yao L, Xiao Y, et al. Resting-state brain function in schizophrenia and psychotic bipolar probands and their first-degree relatives. *Psychol Med*. 2015;45:97–108.

*Nickl-Jockschat T, Janouschek H, Eickhoff SB, Eickhoff CR. Lack of meta-analytic evidence for an impact of COMT Val158Met genotype on brain activation during working memory tasks. *Biol Psychiatry*. 2015;78(11):e43–e46.

*Röder CH, Dieleman S, van der Veen FM, Linden D. Systematic review of the influence of antipsychotics on the blood oxygenation level-dependent signal of functional magnetic resonance imaging. *Curr Med Chem*. 2013;20:448–461.

Satterthwaite TD, Vandekar SN, Wolf DH, et al. Connectome-wide network analysis of youth with psychosis-spectrum symptoms. *Mol Psychiatry*. 2015;20(12):1508–1515.

*Taylor SF, Kang J, Brege IS, Tso IF, Hosanagar A, Johnson TD. Meta-analysis of functional neuroimaging studies of emotion perception and experience in schizophrenia. *Biol Psychiatry*. 2012;71:136–145.

Wolf DH, Satterthwaite TD, Calkins ME, et al. Functional neuroimaging abnormalities in youth with psychosis spectrum symptoms. *JAMA Psychiatry*. 2015;72:456–465.

Wolf DH, Satterthwaite TD, Loughead J, et al. Amygdala abnormalities in first-degree relatives of individuals with schizophrenia unmasked by benzodiazepine challenge. *Psychopharmacology (Berl)*. 2011;218:503–512.

*Yan C, Yang T, Yu QJ, et al. Rostral medial prefrontal dysfunctions and consummatory pleasure in schizophrenia: a meta-analysis of functional imaging studies. *Psychiatry Res*. 2015;231:187–196.

▲ 12.9 Molecular Brain Imaging in Schizophrenia

DEAN F. WONG, M.D., PH.D., JAMES ROBERT BRAŠIĆ, M.D., ANDREW HORTI, PH.D., AND GERHARD GRÜNDER, M.D.

Molecular *brain imaging*, the visualization of in vivo brain chemistry in humans and animals, has undergone remarkable growth since the first mapping of neuroreceptors in humans more than three decades ago in the early 1980s. Molecular brain imaging comprises radiological techniques that represent unique opportunities to study disease and therapeutic mechanisms in people with schizophrenia. In addition, molecular brain imaging represents a powerful tool to document the expression of animal models to simulate the symptoms and signs of schizophrenia as parallel studies can be performed in animals and humans. Furthermore, molecular brain imaging provides the means to determine the effects of therapeutic interventions for schizophrenia, such as receptor occupancy, in animal models, healthy adult human volunteers, and people with schizophrenia. Table 12.9–1 provides a glossary of terms encountered in molecular brain imaging.

MOLECULAR BRAIN IMAGING AND THEORIES OF SCHIZOPHRENIA

Arvid Carlsson's well-known *dopamine hypothesis of schizophrenia* (i.e., hyperactivity of the subcortical dopamine system contributes to the pathogenesis and pathophysiology of schizophrenia) is closely associated with the original recognition of the role of dopamine independent of its function as a precursor for norepinephrine biosynthesis some 50 years ago. The dopamine hypothesis originated from the finding that traditional neuroleptic drugs, such as chlorpromazine (Thorazine) and haloperidol (Haldol), antagonize dopamine activity by blocking postsynaptic D_2-like dopamine receptors. Subsequently, research in schizophrenia has provided the basis to propose additional hypotheses, including the serotonin hypothesis of schizophrenia, substantiated by studies to determine the mechanisms of action of atypical second-generation antipsychotics, and more recently the glutamate hypothesis of schizophrenia. Molecular imaging provides the means to test these hypotheses directly in living humans with schizophrenia. In other words, the original dopamine hypothesis that schizophrenia results from excessive dopaminergic activity likely must be expanded to reflect contributions of dysfunction of serotonergic, glutamatergic, and several other neurotransmitter systems.

The study of molecular imaging in schizophrenia has a historical root in the very basis of neuroreceptor imaging with PET (see Fig. 12.9–1) and single photon emission computed tomography (SPECT). Indeed, the first

successful imaging study in 1983 in living human brain was with the antipsychotic spiperone (Spiroperidol), first labeled with carbon-11 [^{11}C]methyl iodine 3-*N*-methylspiperone ([^{11}C]NMSP)and later with fluorine-18 [^{18}F]. This was followed later with the radiotracer [^{11}C]raclopride, an antipsychotic and D_2 and D_3 antagonist that had been studied in humans in clinical trials. Although both of these antipsychotics were not eventually used clinically in the United States (except for clinical investigational use), their radiolabeled PET analogs both led the way for studying of D_2-like dopamine receptors and indeed opened up the entire field for studying neuroreceptors as essentially a new subspecialty of neuroimaging with tremendous application in neuropsychiatry. Both of these radiotracers were quickly shown to be displaceable by unlabeled marketed antipsychotics such as haloperidol (Haldol), which led ultimately to occupancy studies and therapeutic drug dose levels now being visualized in living human brain.

Following these initial dopamine studies, there was increased interest in the serotonin (5-HT) system, which is also believed to be important in some aspects of schizophrenia symptomatology and potentially in treatment. Indeed, many atypical antipsychotics starting with clozapine (Clozaril) and subsequent later generation antipsychotics such as risperidone (Risperdal) and olanzapine (Zyprexa) have both serotonin- and dopamine-binding properties. Hence, the development of molecular imaging of the serotonin sites was a natural progression. It was discovered with molecular imaging technology that almost all second-generation "atypical" antipsychotics (with the exception of amisulpride) profoundly antagonize serotonin 5-HT_2 receptors at the lowest clinically applied doses.

Table 12.9–1.
Glossary of Terms, Abbreviations, and Definitions for Molecular Brain Imaging in Schizophrenia

Term	Abbreviation	Definition
α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid	AMPA	A glutamate agonist
α-methyl-para-tyrosine	AMPT	A pharmacological agent to deplete presynaptic dopamine by inhibition of tyrosine hydroxylase
β-carbomethoxy-3-β-(4-iodophenyl)-tropane	β-CIT	An agent to bind to dopamine transporters (DATs) and serotonin transporters (SERTs)
Binding affinity	$1/K_D$	Tightness of the attachment between a neurotransmitter and its neuroreceptor
Binding potential	BP	An estimate of the density of the indicated receptors in specific brain regions estimated by the ratio of B_{max} to K_D
[^{11}C]-3-amino-4-[2-(dimethylamino-methylphenylthio)]benzonitrile	[^{11}C]DASB	A radiotracer to bind to serotonin transporters (SERTs)

Dopa-decarboxylase	DDC	The rate-limiting enzyme in the presynaptic metabolism of dopamine
Dopamine receptor	D_2/D_3 R	D_2/D_3 R represents the receptors labeled by benzamide PET radiotracers, e.g., [^{11}C]raclopride.
Dopamine release	DAR	The increment in intrasynaptic dopamine following physiological stress or pharmacological stimulation by a dopaminergic agent, e.g., amphetamine and methylphenidate
Dopamine transporter	DAT	The reuptake site for dopamine from the synapse into the presynaptic neuron to be packaged again for rerelease into the synapse
(+)-enantiomer of 1,2,3,4a,5,6-hexahydro-9-hydroxy-4-n-propyl- 4H-naphth[1, 2-b][1, 4]-oxazine	(+)-PHNO	A potent dopamine D_2/D_3 agonist
Equilibrium dissociation constant	K_D	The inverse of the affinity of radioligand binding
Functional magnetic resonance imaging	fMRI	A brain imaging method detecting functional changes using magnetic waves
γ-aminobutyric acid	GABA	An inhibitory neurotransmitter
[^{123}I]iodobenzamide	[^{123}I]IBZM	An iodinated benzamide
Lysergic acid diethylamide	LSD	A hallucinogen
Molecular imaging	MI	The use of molecular-based probes to visualize physiological process
Magnetic resonance spectroscopy	MRS	The visualization of in vivo brain chemistry using magnetic waves
Maximal binding potential	B_{max}	The theoretical density of the indicated neuroreceptor in portions of the brain
N-methyl-D-aspartic acid	NMDA	An agonist of the NMDA glutamate receptor
3-*N*-methylspiperone	NMSP	The methylated derivative of spiperone, a potent agent to bind to dopamine $D_2/D_3/D_4$ and serotonin (5-HT_{2A}) receptors
Positron emission tomography	PET	A technique for the visualization of positrons in an organ after the administration of a radioactive compound
3-quinuclidinyl benzilate	QNB	A potent antagonist of muscarinic receptors
Serotonin	5-HT_{2A}	An excitatory neurotransmitter
Serotonin receptor	5-HT_{2A} R	Molecules on the postsynaptic site to activate the neuron after being stimulated by serotonin
Serotonin transporter	SERT	The reuptake site for serotonin from the synapse into the presynaptic neuron to be packaged again for rerelease into the synapse
Single photon emission computed tomography	SPECT	A technique for visualization of single photons in an organ after the administration of a radioactive compound
2-deoxy-2-[^{18}F]fluoro-D-glucose	[^{18}F]FDG	Radioactive sugar
H_2[^{15}O]	H_2[^{15}O]	Radioactive water

FIGURE 12.9–1. High-resolution positron emission tomography (PET) images are produced by this high-resolution research tomography (HRRT) scanner. (Photographed by Jeffrey Galecki.)

More recently, a "glutamate hypothesis" has developed in schizophrenia. This has been related to some extent to dopamine, where one of the prevailing theories is that a hypoglutamatergic state, especially in frontal cortical regions in schizophrenia, has led to the hyperdopaminergic state that can be observed especially in the striatal areas. Indeed, if this glutamate hypothesis is verified, this might help explain the evidence for the increase in psychostimulant (e.g., amphetamine)-induced dopamine release (DAR), which can be observed in schizophrenia. Similarly, NMDA antagonists, such as ketamine (Ketalar), which affect the 1-(1-phenylcyclohexyl) piperidine (PCP) also known as PCP site of the NMDA receptor (a calcium ion channel), provide experimental tools for pharmacological challenges to measure DAR changes in schizophrenia.

Thus, neuroreceptor imaging, while a technical achievement of neuropsychopharmacology and radioisotope methodology, originates from research to try to understand the pathophysiology and treatment of neuropsychiatric disorders such as schizophrenia. Indeed, once it is possible to study the dopamine, serotonin, glutamate, and other important systems in the same individuals and similar cohorts, it will be such neuroreceptor imaging that will continue to help push forward the frontiers of schizophrenia research.

Hence, there has been a developing synergy among molecular imaging

with radioisotopes labeled to drugs, including PET, SPECT, as well as magnetic resonance spectroscopy (MRS), and structural MRI (sMRI) and fMRI, and the current knowledge of schizophrenia, including its pathophysiology and treatment. The value of MRS is the ability to image amino acids (such as GABA and glutamate). The different frequencies of protons in lipid and water facilitate the differentiation of tissues in the nervous system by MRS. These measurements are complementary to the information obtained by PET and SPECT molecular imaging methods of neurotransmitter metabolism, transporters, and receptors that are observed in lower (nanomolar) concentrations in the brain. It is important to consider the contribution of changes in brain structure with sMRI in the interpretation of neurochemical and functional imaging data methods by means of fMRI. Thus, PET and SPECT and MR methods provide unique information, and the technique can be applied to address specific research questions relevant to pathophysiology and treatment mechanisms.

Molecular Imaging of Cerebral Blood Flow and Cerebral Metabolism

Prior to the first neuroimaging receptor studies in schizophrenia in the mid-1980s, there was over a decade of imaging involving cerebral blood flow measurements with xenon [^{133}Xe] and radioactive water ([^{15}O]$_2$H) and almost simultaneously studies of regional cerebral glucose metabolism using 2-deoxy-2-[^{18}F]-fluoro-D-glucose ([^{18}F]FDG) PET. These latter studies suggested the presence of "hypofrontality" in patients with schizophrenia. PET with radioactive water ([^{15}O]$_2$H) demonstrated reduced blood flow in the frontal cortex, anterior cingulate, cerebellum, and thalamus in the left hemisphere, with possibly compensatory increased blood flow in the right hemisphere. Nonetheless, the findings were difficult to relate to brain neurochemistry, something that can now be obtained from molecular imaging studies. The ability to image receptors, which began with the imaging of drugs primarily used as antipsychotics, has now extended to include other neurotransmitter systems as well as pre- and intrasynaptic sites. These molecular imaging methodologies (see Section 1.17 on radiotracer imaging with PET and SPECT) have indeed revolutionized the entire area of CNS neuroimaging since the same methodology has been applied to other neuropsychiatric disorders and more recently to drug development. Furthermore, this initial research primarily focused on schizophrenia has led the way for the study of such neuroreceptor systems in other organs.

Molecular Imaging of D_2-Like Receptors

Historical advances in both molecular imaging and schizophrenia research extend beyond investigations of postmortem autoradiography and tissue samples to include the in vivo human imaging of radiolabeled drugs such as

[^{11}C]chlorpromazine. However, it did not differentiate the D_2-like rich striatum from cortical imaging and so was not considered useful for further development. The first successful molecular imaging, a translation of in vitro neurochemistry, began with the imaging of [^{11}C]*N*-methylspiperone ([^{11}C]NMSP), a methylated derivative of spiperone, one of the highest affinity antipsychotics tested in clinical trials in the 1970s. The [^{11}C] methylation of spiperone provided the technical tool to image D_2-like dopamine receptors to quantify their density in the basal ganglia as well as cortical serotonin (5-HT_{2A}) receptors. Now more specific D_2 and D_3 ligands exist such as [^{11}C]raclopride and [^{18}F]fallypride.

TYPES OF MOLECULAR IMAGING

Molecular imaging techniques include PET, SPECT, and fMRI. These are all *functional brain imaging techniques* to visualize the metabolism and other physiological processes in the living brain.

Since the first studies of the distribution and the density of dopamine D_2 receptors in the human brain, there has been an explosion of neuroreceptor PET ligands available for the study of the pathophysiology of schizophrenia. Furthermore, in particular, these neuroreceptor compounds form the basis of near-biomarkers for measuring the mechanism of action and in particular rational drug dosing of drugs for antipsychotic action by virtue of the occupancy of receptor systems thought to be important in treatment. Table 12.9–2 documents a large number of these radiopharmaceuticals that are available for measuring neuroreceptor systems relevant to schizophrenia. These include but are not limited to dopamine D_1, dopamine D_3, extrastriatal dopamine D_2 (cortical), dopa-decarboxylase (DDC), nicotinic acetylcholinergic subtypes α4 β2, α7, cannabinoids CB1, phosphodiesterase subtypes 4 (PDE4) and 10A (PDE10A), serotonin 5-HT_{1A}, serotonin 5-HT_{2A}, glycine transporter subtype 1 ($GlyT_1$), opiate (mu, delta, and kappa) receptors, metabotropic glutamate receptors ($mGluR_2$, $mGluR_3$, and $mGluR_5$), serotonin transporters (SERT), NMDA ion channel site, and the translocator protein (TSPO).

The goal of functional brain imaging techniques is to identify and characterize the dynamic chemical and biological changes of the brain under varying conditions. Functional brain imaging contrasts with *structural brain imaging*, procedures to precisely demarcate anatomical regions of the brain. Structural brain imaging procedures include x-rays, MRI, CT, and ultrasound. Thus, molecular brain imaging is a subclass of functional brain imaging.

Indirect Molecular Imaging Estimates of Neurotransmitter Action

Ideally, concentrations of neurotransmitters would be measured in the living human brain. Procedures to directly measure them are not yet available. Therefore, molecular brain imaging requires indirect techniques to estimate the density and distributions of neurotransmitter in the living brain. For example, there exist no noninvasive procedures to directly measure glucose concentration in the living human brain. Instead, [^{18}F]FDG PET is a technique used to estimate the density and distribution of [^{18}F]FDG, a radioligand incorporating a derivative of glucose in the brain that measures glucose metabolism. Similarly, the concentrations of dopamine and serotonin cannot be directly measured in the brain. Instead, radiotracers exist to estimate the density and distribution of the presynaptic DAT and SERT and the postsynaptic dopamine (both D_1- and D_2-like) dopamine receptors (D_1R and D_2R) and serotonin receptor (5-HT_{2A}R). For example, molecular imaging demonstrates the alterations in the density of dopamine D_2 receptor with the menstrual cycle. In fact, there is PET evidence that the binding to the dopamine D_2 receptor is reduced in the follicular phase and increased in the periovulatory and luteal phases of the estrous cycle, and estrogen replacement has produced an increase in cortical 5-HT_{2A} receptors. In addition, the density of dopamine D_2 receptors is inversely proportional to age and the fall with age differs between sexes. Molecular imaging also has been used to confirm the dopamine hypothesis that people with schizophrenia in exacerbation have an excess of intrasynaptic dopamine in the brain. The dopamine hypothesis represents a cornerstone of schizophrenia pathophysiology. Drugs to block the binding of dopamine to dopamine D_2 receptors remain a fundamental tool of the pharmacology of schizophrenia.

Table 12.9–2.
Abbreviated and Chemical Names of Available Nuclear Neuroimaging Radiotracers and the Indicated Neurotransmitter Systems

Target	Abbreviated Name	Chemical Name
Acetylcholinesterase (AChE) activity	[^{11}C]MP4A	*N*-[^{11}C]methylpiperidin-4-yl acetate
Cannabinoid CB_1 receptors	[^{3}H]MePPEP	[^{3}H](3*R*,5*R*)-5-(3-methoxy-phenyl)-3-((*R*)-1-phenyl-ethylamino)-1-(4-trifluoromethyl-phenyl)-pyrrolidin-2-one
Cannabinoid CB_1 receptors	[^{11}C]OMAR	1-(2,4-dichlorophenyl)-4-cyano-[^{11}C]-5-(4methoxyphenyl)-*N*-(piperidin-1-yl)-1*H*-pyrazole-3-carboxamide
Cannabinoid CB_1 receptors	[^{18}F]MK-9470	*N*-[2-(3-cyano-phenyl)-3-(4-(2-[^{18}F]-fluoroethoxy)phenyl)-1-methylpropyl]-2-(5-methyl-2-pyridyloxy)-2-methylproponamide
Dopamine synthesis and	6-[^{18}F]DOPA	6-[^{18}F]dihydroxyphenylalanine

storage		
Dopamine D_1 receptors	(+)-[^{11}C]NNC112	(+)-8-chloro-5-(7-benzofuranyl)-7-hydroxy-3-[^{11}C]methy1-2,3,4,5-tetrahydro-1H-3-benzazepine
Dopamine D_1 receptors	[^{11}C]SCH23390	(*R*)-(+)-8-chloro-2,3,4,5-tetrahydro-3-[11C]methyl-5-phenyl-1*H*-3-benzazepin-7-ol
Dopamine D_2 receptors	[^{11}C]RAC	[^{11}C]raclopride
Dopamine D_2 receptors	[^{18}F]NMB	[^{18}F](*N*-methyl)benperidol
Dopamine D_2 receptors	[^{123}I]IBZM	(*S*)-*N*-[(1-ethyl 2-pyrrolidinyl)] methyl-2-hydroxy-3-[^{123}I]iodo-6-methoxybenzamide
Extrastriatal dopamine D_2 receptors	[^{11}C]FLB 457	(*S*)-*N*-((1-ethyl-2-pyrrolidinyl)methyl)-5-bromo-[^{11}C]2,3-dimethoxybenzamide
Dopamine D_2/D_3 receptors	[^{18}F]FAL	[^{18}F]fallypride
D_3 receptors	[^{11}C]-(+)-PHNO	[^{11}C]-(+)-4-propyl-3,4,4a,5,6,10b-hexahydro-2H-naphtho[1,2-b][1,4]oxazin-9-ol
Dopamine transporter	[^{123}I]β-CIT	β-CIT2-betacarbomethoxy-3 beta-(4-[^{123}I]iodophenyl)tropane
γ-aminobutyric acid (GABA)	[^{18}F]FMZ	[^{18}F]fluoroflumazenil
Glycine transporter type 1 ($GlyT_1$)	[^{11}C]GSK931145	N-((1-([$^{11}CH_3$]dimethylamino)cyclopentyl) (phenyl)methyl)-2,6-dimethylbenzamide
Glycine transporter type 1 ($GlyT_1$)	[^{11}C]R05013853	[5-[11C]methanesulfonyl-2-((*S*)-2,2,2-trifluoro-1-methyl-ethoxy)-phenyl]-[5-(tetrahydro-pyran-4-yl)-1,3-dihydro-isoindol-2-yl]-methanone
Muscarinic acetylcholine receptor (mAChR)	[^{11}C]NMPB	[^{11}C]*N*-methyl-4-piperidyl benzylate
Muscarinic acetylcholine receptor subtype 1 ($mAChR_1$)	[^{11}C]GSK1034702	[11C]4-fluoro-6-methyl-1-(1-(tetrahydro-2H-pyran-4-yl)piperidin-4-yl)-1,3-dihydro-2H-benzo[d]imidazol-2-one
Nicotinic acetylcholine receptor (nAChR) subtype α4β2	[^{18}F]AZAN	(−)-2-(6-[^{18}F]fluoro-2,3'-bipyridin-5'-yl)-7-methyl-7-aza-bicyclo[2.2.1]heptane
Nicotinic acetylcholine receptor (nAChR) subtype α4β2	2-[^{18}F]FA	(*S*)-3-(azetidin-2-ylmethoxy)-2-[^{18}F]fluoropyridine
Nicotinic acetylcholine receptor (nAChR) subtype α4β2	6-[^{18}F]FA	6-[^{18}F]fluoro-3-(2(*S*)-azetidinylmethoxy)pyridine
Nicotinic acetylcholine receptor (nAChR) subtype α4β2	5-[^{123}I]IA	(*S*)-5-[^{123}I]iodo-3-(2-azetidinylmethoxy)pyridine
Nicotinic acetylcholine receptor (nAChR) subtype α7	[^{11}C]CHIBA-1001	4-[^{11}C]methylphenyl 2,5-diazabicyclo[3.2.2]nonane-2-carboxylate
Nicotinic acetylcholine receptor (nAChR) subtype α7	[^{18}F]ASEM	[^{18}F] 3-(1,4-diazabicyclo[3.2.2]nonan-4-yl)-6-fluorodibenzo[b,d]thiophene 5,5-dioxide
Opioid receptors	[^{18}F]MK-0911	1-(2-[18F]fluoroethyl)-3-(3-(hydroxymethyl)-1-(spiro[2.5]octan-4-ylmethyl)piperidin-4-yl)-1,3-dihydro-2H-benzo[d]imidazol-2-one
Phosphodiesterase 10A	[^{18}F]JNJ42259152	2-[[4-[1-(2-[^{18}F]fluoroethyl)-4-(4-pyridinyl)-1*H*-

(PDE_{10A}) receptors		pyrazol-3-yl]phenoxy]methyl]-3,5-dimethyl-pyridine
Phosphodiesterase 10A (PDE_{10A}) receptors	[^{18}F]MNI-654	2-(2-(3-(1-(2-[^{18}F]fluoroethyl)-1*H*-indazol-6-yl)-7-methyl-4-oxo-3,4-dihydroquinazolin-2-yl)ethyl)-4-isopropoxyisoindoline-1,3-dione
Phosphodiesterase 10A (PDE_{10A}) receptors	[^{18}F]MNI-659	2-(2-(3-(4-(2-[^{18}F]fluoroethoxy)phenyl)-7-methyl-4-oxo-3,4-dihydroquinazolin-2-yl)ethyl)-4-isopropoxyisoindoline-1,3-dione
Serotonin 5-HT2a receptors	[^{11}C]MDL100907	(*R*)-(+)-[3-O-methyl-[11C]]-4-(1-hydroxy-1-(2,3-dimethoxyphenyl)methyl)-*N*-2-(4-fluoro-phenylethyl)piperidine
Serotonin 5-HT2a receptor	[^{18}F]altanserin	{3-(2-[4-(4-[^{18}F]fluorobenzoyl)piperidin-1-yl]-ethyl)-1,2-dihydro-2-thioxo-quinazolinone, *1*}
Serotonin (5-HT) transporter	[^{11}C]DASB	[11C]-3-amino-4-(2-dimethylaminomethylphenylsulfanyl)-benzonitrile
Translocator protein (TSPO)	[^{11}C]PK11195	*N*-butan-2-yl-1-(2-[^{11}C]chlorophenyl)-*N*-methylisoquinoline-3-carboxamide
Translocator protein (TSPO)	[^{18}F]FEPPA	N-(2-(2-[18F]fluoroethoxy)benzyl)-N-(4-phenoxypyridin-3-yl)acetamide
Vesicular acetylcholine transporter	[^{123}I]IBVM	[^{123}I]iodobenzovesamicol
Vesicular monoamine transporter type 2 (VMAT-2)	[^{11}C]DTBZ	[^{11}C]dihydrotetrabenazine

Pharmacological Challenges

Pharmacological challenges are experimental tools to test the effects of the administration of drugs hypothesized to alter the levels of neurotransmitters in regions of the brain. Examples of pharmacological challenges of the dopaminergic system include the administration of pharmacological agents to increase, for example, amphetamine (Dexedrine), or decrease, for example, α-methyl-para-tyrosine (AMPT), metyrosine (Demser), intrasynaptic dopamine with [^{11}C]raclopride presumably competing with dopamine for binding to D_2/D_3 receptors. These techniques demonstrate that intrasynaptic dopamine is increased following the administration of amphetamine. After AMPT administration, patients with schizophrenia apparently still have more dopamine available than normal, although both are decreased compared to subjects without AMPT. This is consistent with elevations of baseline dopamine D_2 receptors in schizophrenia.

Pharmacological challenges also represent tools to evaluate the mechanisms of actions of novel agents. For example, agents hypothesized to alter the intrasynaptic concentration of dopamine can be administered before scans to estimate dopamine binding. The results can then be compared and contrasted with the effects of agents known to increase or decrease intrasynaptic dopamine.

More recently despite the lower cortical density of D_2 receptors, amphetamine challenges with [^{11}C]FLB457 and [^{18}F]fallypride have shown cortical DAR.

Important new challenges are being attempted in other neurotransmitter systems and dopamine. There has been a long search for methods of inducing intrasynaptic serotonin release. Additional pharmacological challenges while potentially less relevant to schizophrenia have been changes in endorphin release measured with the μ-opiate receptor ligand [^{11}C]carfentanil. Much more relevant to schizophrenia, however, is the suggestion that intrasynaptic glutamate might be measured in cortex, following drugs that increase glutamate such as *N*-acetylcysteine. In addition, there have been recent studies involving ketamine infusions at the doses similar to those used for challenge studies to induce psychotic-like behavior in clinical research studies in patients with schizophrenia and controls. Both of these have employed the radiotracer [^{11}C]ABP688 (see section later).

MOLECULAR IMAGING IN PATHOPHYSIOLOGY

Several theories have been proposed to facilitate an understanding of the pathogenesis and the pathophysiology of schizophrenia. Promising conjectures to explain the abnormalities of schizophrenia include dopamine, glutamate, and inflammation hypotheses.

The dopamine hypothesis proposes that positive symptoms of schizophrenia result from an excess of intrasynaptic dopamine, and negative symptoms result from a deficit of dopamine.

The glutamatergic hypothesis proposes that positive symptoms of schizophrenia result from dysfunction of glutamate metabolism in the frontal cortex. Glutamate is the main excitatory neurotransmitter and GABA is the main inhibitory neurotransmitter in the human brain. Glutamatergic neurons from the prefrontal cortex project to GABAergic neurons in the ventral tegmental area (VTA). The activated excitatory glutamatergic neurons then stimulate GABAergic neurons. The GABAergic neurons in the VTA project to the striatum. The activated inhibitory GABAergic neurons inhibit dopaminergic neurons in the striatum. People with schizophrenia are hypothesized to exhibit dysfunction of the glutamatergic neurons in the prefrontal cortex, resulting in reduced activation of the GABAergic neurons in the VTA, leading to reduced inhibition of dopaminergic neurons in the striatum. The uninhibited dopamine activity in the striatum is then hypothesized to produce the positive symptoms of schizophrenia.

The inflammatory hypothesis of schizophrenia posits that oxidative stress and resulting cerebral inflammation underlie the pathophysiology of schizophrenia. TSPO, a putative marker of inflammation, is increased in the prefrontal cortex of autopsy specimens of adults with schizophrenia.

Radiotracers to bind to the TSPO, for example, [^{11}C]PK11195 and [18F]FEPPA (see Table 12.9–2) permit the assessment of inflammation in healthy people and in patients with schizophrenia and other neuropsychiatric disorders.

Molecular Imaging of Dopamine in Schizophrenia

The development of molecular imaging studies resulted from the successful radiolabeling of dopamine receptor antagonists, typical antipsychotic drugs, in the mid-1980s, consistent with the investigation of postsynaptic dopamine D_2-like receptors.

Estimation of the Binding of Dopamine D_2-Like Receptors. The study of the binding of D_2-like receptors is typically made by a single PET or SPECT scan at high specific activity (with low mass of the nonpharmacological moiety injected) to measure the so-called binding potential (BP), which is a ratio of the receptor density to the binding affinity.

Measurement of the BP has been accomplished using two main classes of radiotracers. The first class of radiotracers to visualize D_2 or D_3 receptors in the brain includes butyrophenones, such as [^{11}C]3-N methylspiperone, [^{11}C]NMSP, and [^{18}F]methyl spiperone. The second class of radiotracers to visualize receptors in the brain includes radiolabeled benzamides, such as [^{11}C]raclopride (Fig. 12.9–2), [^{18}F]fallypride ([^{18}F]FP) (Fig. 12.9–3), [^{11}C]FLB457, or [^{123}I]iodobenzamide ([^{123}I]IBZM).

Estimation of Absolute Dopamine D_2-Like Receptor Density. In contrast to the typical approach of measuring the BP of D_2-like receptors, techniques using two PET scans have the unique ability to measure absolute receptor density (B_{max}). The measurement of absolute B_{max} and affinity K_D can be carried out by obtaining two brain scans with the administration of [^{11}C]raclopride or [^{11}C]NMSP, with the latter after a known potent dose of a cold, nonradioactive D_2/D_3 receptor antagonist such as unlabeled raclopride itself or haloperidol, respectively, to reduce some of the binding. Following some additional mathematical modeling, these results then provide an estimate of the B_{max} and binding affinity ($1/K_D$). These results correspond to the results from autoradiography studies or in vitro homogenate studies.

Subsequently, simplified studies involving single high-specific activity (typically radiotracer mass injections under 20 μg) were carried out to estimate the BP, that is, the ratio of the B_{max} to K_D. Nonetheless, in order to separate these variables, B_{max} and K_D from binding potential (BP_{ND}), it is necessary to carry out receptor saturation studies with varying unlabeled mass injections such as done in vitro with “Scatchard” analyses of receptor

binding assays.

In general, the absolute B_{max} studies show elevations ranging from small to large as much as a threefold increase with [^{11}C]NMSP, a butyrophenone in drug-naïve schizophrenia. Although B_{max} elevations have been consistently shown with [^{11}C]NMSP, no changes have been demonstrated with B_{max} using benzamides such as [^{11}C]raclopride or [^{123}I]IBZM. Interestingly, one of the characteristics of these benzamides is that they are also quite susceptible to endogenous dopamine and recently have gained popularity for measuring the effect of pharmacological challenges or psychological stress that alters endogenous dopamine levels. Some have suggested that the differences in B_{max} between raclopride and NMSP could be attributed to the preferential affinity for a monomer versus a dimer form of the dopamine receptor in the case of the [^{11}C]NMSP and possibly receptor internalization and susceptibility to endogenous neurotransmitter for [^{11}C]raclopride. Regardless of the reasons for the differences, one meta-analysis has shown that there is a small 12 percent significant elevation of striatal D_2 receptors with an effect size of 0.51, and there is an association between increased D_2 receptor density and positive symptoms.

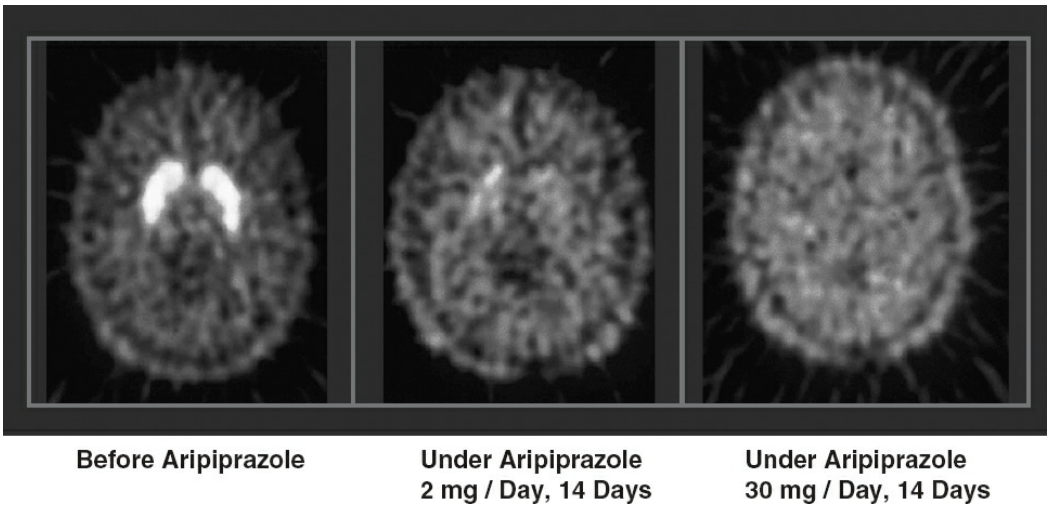

FIGURE 12.9–2. Representative transversal images from positron emission tomography (PET) with [^{11}C]raclopride, a radiotracer for striatal D_2/D_3 receptors in the brain, of a healthy adult volunteer without medication (*left*) and after treatment for 14 days with either 2-mg (*middle*) or 30-mg (*right*) aripiprazole (Abilify). The baseline study (*left*) demonstrates marked uptake of the radiotracer in the striatum (caudate nucleus and putamen), indicating a high density of striatal D_2/D_3 receptors. After treatment with 2-mg aripiprazole for 14 days, approximately 70 percent of the D_2/D_3 receptors in the striatum are occupied by aripiprazole, so the radiotracer uptake is markedly reduced (*middle*). After treatment with 30-mg aripiprazole for 14 days, almost all of the D_2/D_3 receptors in the striatum are occupied by aripiprazole, so the radiotracer uptake is almost completely blocked (*right*). These results confirm that aripiprazole occupies the D_2/D_3 receptors in the striata of healthy adults in a dose-dependent manner, and that clinical doses of the drug lead to almost complete saturation of striatal D_2/D_3 receptors. (Modified from Yokoi F,

Gründer G, Biziere K, et al. Dopamine D_2 and D_3 receptor occupancy in normal humans treated with the antipsychotic drug aripiprazole (OPC 14597): a study using positron emission tomography and [^{11}C]raclopride. *Neuropsychopharmacology*. 2002;27:248–259, with permission.)

Dopamine Precursor Estimates. Dopamine precursor measures using [^{18}F]fluorodopa or [^{11}C]dopa were among the first markers of dopamine neurotransmission to be studied. When properly quantified, these measurements reflect DDC rates of activity or represent some measure of dopamine presynaptic function. A number of these studies have demonstrated that the dopa uptake was higher in people with schizophrenia and other psychoses in contrast to healthy matched controls. Elevated [^{18}F]fluorodopamine ([^{18}F]Fdopa) turnover has been shown to coexist with decreased steady-state storage of [^{18}F]Fdopa and its decarboxylated metabolites in the caudate nucleus and amygdala of untreated people with schizophrenia. Interestingly, the decreased steady-state storage of [^{18}F]Fdopa in the amygdala correlated inversely with positive symptoms of schizophrenia, suggesting, furthermore, a role for this brain structure in the pathophysiology of the disease. Hence, the striatal findings are among the strongest evidence of a "biomarker" in schizophrenia and are consistent with the concept of decreased tonic dopamine and elevated phasic dopamine originally proposed by Anthony Grace.

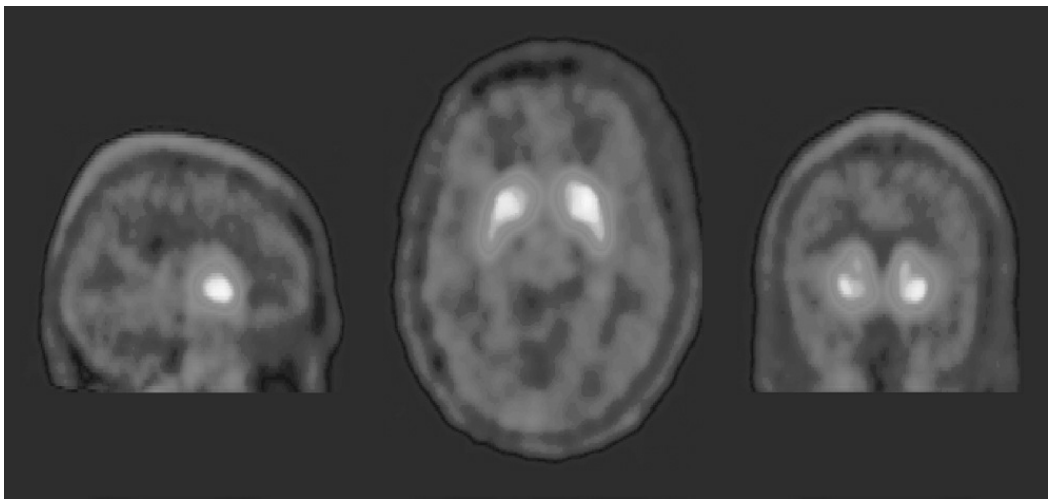

FIGURE 12.9–3. Representative sagittal (*left*), transversal (*middle*), and coronal (*right*) images from positron emission tomography (PET) with [^{18}F]fallypride ([^{18}F]FP), a radiotracer for extrastriatal D_2/D_3 receptors in the brain, of a 24-year-old man with his first episode of schizophrenia treated with aripiprazole until 78 hours before the tomography. The average striatal D_2/D_3 receptor occupancy was 82 percent. At the moment of tomography, the aripiprazole serum concentration was 106 ng/mL. (From Gründer G, Fellows C, Janouschek H, et al. Brain and plasma pharmacokinetics of aripiprazole in patients with schizophrenia: an [^{18}F]fallypride PET study. *Am J Psychiatry*. 2008;165:988–995, with permission.)

Other Dopamine Subtypes. There is considerable interest in the D_3 receptor as a potential therapeutic target for schizophrenia. But until relatively recently, no studies are explicitly able to image the D_3 receptor in vivo without also measuring the much more densely distributed D_2 receptors. A high-affinity selective D_3 radioligand is needed, and some recent studies, using the D_3 selective PET ligand [^{11}C]PHNO, have been carried out in schizophrenia, which show some selective D_3 binding.

Dopamine D_1 receptors may also play roles in the development and maintenance of the symptoms and signs of schizophrenia. A number of studies have shown no significant difference in D_1 receptors in untreated patients with schizophrenia. However, a twin study showed reduced striatal D_1 receptors in the affected twins who were on medications compared to the unaffected cotwins. Dopamine D_1 receptor density in the dorsolateral prefrontal cortex is associated with impairments in working memory performance and negative symptoms. The limitations of these studies of the D_1 receptors are that the radiotracers are not only selective for D_1 receptors but also bind to 5-HT_{2A} receptors. Thus, further radiotracer development is needed for the D_1 receptor site. Several years ago, D_4 was considered a serious candidate for antidepressant action, but initial PET ligands and small therapeutic trials for schizophrenia failed. It is still of interest, perhaps in other areas of neuropsychiatry as personality, and is a binding site for [^{11}C]NMSP.

Measuring Intrasynaptic Dopamine. In the past decade, an important property probably unique to radiolabeled benzamides has been found. Specifically, [^{123}I]IBZM and [^{11}C]raclopride (Fig. 12.9–2) and probably [^{18}F]fallypride (Fig. 12.9–3) can have their D_2/D_3 receptor binding reduced in the presence of a pharmacological challenge with a pharmacological agent that induces increased or decreased levels of intrasynaptic dopamine. In the case of increased intrasynaptic dopamine, challenges with dopaminergic agents, including methylphenidate (Ritalin) as well as amphetamines, have been employed. Investigators have shown that these benzamides can have a reduced BP in the presence of an intravenous amphetamine challenge (0.3 mg/kg body weight). Researchers showed significantly greater decreases in BPs in patients with schizophrenia in contrast to healthy normal control volunteers. There is still some controversy about the meaning of these decreases in BP, and what is probably most defensible is that there is an increase in occupancy of the D_2/D_3 sites. This reduction of occupancy could be due to the increased intrasynaptic dopamine released into the synapse through the use of psychostimulants. However, a change of the affinity state, which might increase with dopamine, might be occurring. These results, however, are

not without controversy. Instead of the radiotracer returning to baseline as one might expect after a bolus amphetamine challenge, it stays depressed for many hours, sometimes exceeding 6 to 8 hours, which could be consistent with other factors than simply intrasynaptic DAR. Furthermore, at least one study has not demonstrated significant elevations in DAR, which in itself is an interesting finding suggestive of a potential heterogeneity in dopamine response to challenges with stimulants. The heterogeneity may be related to heterogeneity in clinical symptoms, as at least one study has shown that DAR is greater in patients who have more positive symptoms at the time of scanning.

Although controversial, it is interesting to note that one other neuropsychiatric disorder, Tourette syndrome, has showed highly significant elevations of DAR after amphetamine challenge and bolus intravenous [^{11}C]raclopride PET studies. These DAR measures are greater with Tourette syndrome patients with comorbid obsessive–compulsive disorder. Hence, the elevated DAR appears not to be specific to schizophrenia.

Molecular Imaging after Dopamine Depletion. Another procedure used for changing intrasynaptic dopamine is dopamine depletion. This has been typically tried either with reserpine (Serpalan), a drug for animals, or with AMPT, a drug which is approved in humans for pheochromocytoma. Using such AMPT challenges has allowed the demonstration of elevation in BPs presumably reflecting the decrease of competition with intrasynaptic dopamine. Such measures have allowed an estimate of baseline or basal dopamine in the synapse. One such study demonstrated that the elevation or change in BP post-AMPT dopamine depletion is greater in people with schizophrenia than in healthy controls, and patients with greater elevations in D_2 receptor binding demonstrated greater improvement in positive symptoms with antipsychotic treatment. This is consistent not only with an increase in basal intrasynaptic dopamine but also with elevated basal densities of elevated D_2 receptors. Thus, these findings support the results of a meta-analysis that shows a small but definite elevation of D_2 receptors in schizophrenia.

Future Molecular Imaging of Agonists of Dopamine and Glutamate

In the future, dopamine agonists, such as the [^{11}C](+)-enantiomer of 1,2,3,4a,5,6-hexahydro-9-hydroxy-4-n-propyl-4H-naphth[1,2-b][1,4]-oxazine ([^{11}C](+)PHNO), or glutamate agonists, such as [^{11}C]α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate ([^{11}C]-AMPA), may help determine whether there is truly intrasynaptic DAR. Dopamine agonists are more reflective of the competition with the intrasynaptic dopamine itself. On the contrary, a recent study with [^{11}C](+)PHNO did not produce a

significant change in DAR.

Extrastriatal Dopamine D_2/D_3 Receptors

With the advent of higher-affinity radiotracers of the D_2/D_3 receptors, such as [^{18}F]fallypride (Fig. 12.9–3) and [^{11}C]FLB457, it has been possible to image, with both PET and SPECT, extrastriatal dopamine D_2/D_3 receptors in the cortex in schizophrenia. Several PET/SPECT studies have been carried out to date, all showing decreases or no change. Some studies have inconsistently demonstrated decreased binding in the thalamus. Also, decreases have been shown in the temporal cortex, amygdala, and cingulate cortex, although inconsistently. Some of the challenges are technical in that absolute receptor density has not been carried out, as in the case of striatal D_2 receptors.

Summary of All Dopamine Neuroreceptor Studies in Schizophrenia

Since the vast majority of studies have been with the dopamine system in schizophrenia, and since this is the area where technically most of the PET and SPECT radioligands have been developed and validated, it is useful to have a summary table of contemporary neurotransmitters (Table 12.9–3). There are increasing lines of evidence that the glutamatergic system is dysfunctional in schizophrenia, and this in turn may lead to especially positive symptoms in schizophrenia in which antipsychotic medications appear to have benefit. Anthony Grace hypothesized that, in contrast to healthy normal controls, people with schizophrenia demonstrate abnormal intrasynaptic levels of dopamine in both (1) the resting, basal, unstimulated, tonic state and (2) the phasic state of maximal arousal such as the one induced pharmacologically by amphetamine or another agent that produces maximal intrasynaptic levels of dopamine. As described by studies of Grace and others, one unifying theory may be described by two states of dopamine: A tonic state usually associated with extrasynaptic dopamine (homeostatic or basal level) and a phasic state, which is more episodic and perhaps more associated with symptomatology such as hallucinations usually associated with intrasynaptic dopamine. Pharmacological challenges confirm the proposal that the tonic or resting level of intrasynaptic dopamine is reduced in schizophrenia while the phasic or excited level of intrasynaptic dopamine is increased in schizophrenia. In addition, hallucinations, delusions, and other positive symptoms are increased during the phasic state resulting from the increased intrasynaptic dopamine. As shown in Table 12.9–3, the pre-, post-, and intrasynaptic measurements of dopamine would be quite compatible with this phasic tonic dopamine theory. In this theory, patients with schizophrenia have lower tonic dopamine and hence less inhabitation of their phasic dopamine (thereby alluding to the hyperdopaminergic

state). Indeed, as shown in Table 12.9–3, a lower tonic dopamine would be consistent with upregulation of the presynaptic enzymes of DDC as measured by [^{18}F]fluorodopa, elevated amphetamine-induced DAR, and elevated D_2-like receptor postsynaptic in response to the lower tonic dopamine. In the future, it may be possible to integrate other abnormalities such as serotonin, and once glutamate studies have been fully verified with multiple radioligands, this should also contribute to this model.

Table 12.9–3
Imaging Studies of Dopamine Neurotransmission in Schizophrenia: Observations and Possible Mechanisms

Site	Observation	Salient and Related Findings	Possible Mechanism
D_2 (striatum)	Elevations with [^{11}C]NMSP (butyrophenone)	Varied postmortem changes, internalization, monomer elevation	Decreased tonic DA
D_2 (striatum)	No change (benzamides) except following AMPT using [^{123}I]IBZM or [^{11}C]raclopride	Due to competition with intrasynaptic DA or dimer configuration, unmasking of D_2 due to intrasynaptic DA removal with AMPT	Increased phasic DA Decreased tonic DA
DDC (striatum)	Elevation with [^{18}F]fluorodopa	Majority of imaging studies	Decreased tonic DA
DA (striatum) release after AMP challenge	Elevation with [^{123}I]IBZM or [^{11}C]raclopride	Majority of imaging studies	Increased phasic DA
DA (striatum) release after ketamine challenge	Elevation with [^{11}C]raclopride Elevation with [^{11}C]FLB457	Glutamate antagonist stimulates hypoglutamatergic state	Increased phasic DA
Prefrontal D_1	Decrease with [^{11}C]SCH23390		Unknown
ACCX D_2	Decrease with [^{11}C]FLB457	Postmortem tyrosine hydroxylase immunoreactive changes	Unknown

[^{11}C]NMSP, (3-*N*-[^{11}C]methylspiperone; DDC, dopa decarboxylase; DA, dopamine; AMPT, α-methylparatyrosine; [123I]IBZM, (*S*)-*N*-[(1-ethyl 2-pyrrolidinyl)]methyl-2-hydroxy-3-[123I]iodo-6-methoxybenzamide; AMP, amphetamine; [^{11}C]FLB457, (*S*)-*N*-((1-ethyl-2-pyrrolidinyl)methyl)-5-bromo-2[^{11}C], 3-dimethoxybenzamide; [^{11}C]SCH23390, (*R*)-(+)-8-chloro-2,3,4,5-tetrahydro-3-[11C]methyl-5-phenyl-1*H*-3-benzazepin-7-ol; ACCX, anterior cingulate cortex.

Adapted from Wong DF. In vivo imaging of D_2 dopamine receptors in schizophrenia: the ups and downs of neuroimaging receptors [Erratum]. *Arch Gen Psychiatry*. 2002;59(5):440, with permission.

Molecular Imaging of Serotonin in Schizophrenia

The second major neurotransmitter system studied in schizophrenia is serotonin. This has most recently been driven by theories of the mechanisms of action of atypical antipsychotics. However, the role of

serotonin in schizophrenia can also be traced back to the psychotic symptomatology from the hallucinogenic effects of lysergic acid diethylamide, which is chemically similar to serotonin.

Estimation of the Density of the Serotonin 5-HT_{2A} Receptors. Initial brain imaging studies of serotonin examined the 5-HT_{2A} receptors. Indeed 5-HT_{2A} receptors were originally imaged with radiotracers such as [^{11}C]NMSP that bind to D_2-like alpha adrenergic and 5-HT_{2A} receptors. A number of postmortem studies have suggested decreased 5-HT_{2A} binding in the prefrontal cortex, but PET studies in drug-free or drug-naive patients with schizophrenia have found normal 5-HT_{2A} receptor binding. One study demonstrated significant decreases in prefrontal cortex in drug-naive schizophrenic patients, and another study has demonstrated decrease in cortical binding in drug-free patients. Estimation of the density and distribution of serotonin receptors in the living human brain is an area that merits further investigation, as more 5-HT_{2A} radioligands become available such as [^{18}F]altanserin.

Estimation of the Density of Serotonin 5-HT_{1A} Receptors. More recently, the ability to measure 5-HT_{1A} receptors has led to investigations showing increases in binding in multiple brain regions, including medial temporal cortex using PET imaging with [^{11}C]-WAY100635, a potent 5-HT_{1A} antagonist. Subsequent studies have shown decreased binding using the same radiotracer in the amygdala, while a third showed no differences. Only some of the imaging studies have confirmed the results of postmortem studies, which suggest increases in 5-HT_{1A} in cortical areas.

Estimation of the Density of the SERT. The second major serotonin imaging area is with the SERT. Initial studies were carried out with nonselective SPECT radiotracers such as [^{123}I]-β-carbomethoxy-3-β-(4-iodophenyl)-tropane ([^{123}I]-β-CIT) and [^{11}C]-3-amino-4-(2-dimethylaminomethylphenylsulfanyl)benzonitrile ([^{11}C]-DASB), which labeled SERT in the midbrain and where no differences were found. This is despite the fact that some postmortem studies have shown decreases in SERT in frontal cortex and in cingulate cortex.

Molecular Imaging of Cholinergic Receptors in Schizophrenia

Acetylcholine represents another key neurotransmitter system likely to play a role in the pathogenesis and pathophysiology of schizophrenia. The neuronal receptors for acetylcholine in the brain are subdivided into two primary classes: muscarinic and nicotinic. Both muscarinic and nicotinic acetylcholine receptors participate in the development and maintenance of

schizophrenia.

Estimation of the Density of Muscarinic Acetylcholine Receptors. Only one study of the muscarinic acetylcholine receptor (mAChR) system using SPECT and 3-quinuclidinyl-4-[123I]iodobenzilate ([123I]IQNB), a potent antagonist of muscarinic receptors, has shown a very marked decrease of mAChR in schizophrenia, consistent with some postmortem studies. However, the functional role of this finding is not fully understood. Using the same radiotracer, the investigators demonstrated greater mAChR occupancy by clozapine than by olanzapine, consistent with the observation of greater extrapyramidal side effects associated with the two drugs.

Estimation of the Density of Nicotinic Acetylcholine Receptors. Some preliminary studies are proceeding to determine whether the decreases in nicotinic acetylcholine receptors (nAChRs) observed in postmortem studies are also observed with PET imaging. This is quite relevant to schizophrenia, given the high prevalence of smoking among patients with schizophrenia.

Many people with schizophrenia smoke cigarettes and inhale large doses of nicotine per puff. Nicotine, the markedly habit-forming component of cigarette smoke, exhibits high affinity for compounds with β_2 nicotinic acetylcholine receptors (β_2*-nAChRs). SPECT with [^{123}I]-5-iodo-3-[2(*S*)-2-azetidinylmethoxy]pyridine ([^{123}I]5-IA) provides quantitative measurements of β_2*-nAChR availability. People with schizophrenia exhibit a reduction in the availability of β_2*-nAChRs in contrast to healthy people without schizophrenia. Smokers with schizophrenia exhibit increased availability of β_2*-nAChRs in contrast to nonsmokers with schizophrenia, indicating an upregulation of β_2*-nAChRs in smokers with schizophrenia. Reduced cortical availability of β_2*-nAChRs in smokers with schizophrenia is associated with increased negative symptoms and reduced executive function.

Molecular Imaging of Histamine H_1 Receptors in Schizophrenia

The histamine H_1 system has also been implicated as having a role in the pathophysiology of schizophrenia. Molecular imaging with [^{11}C]doxepin has shown reductions of BP values for H_1 receptors in the frontal and prefrontal cortices and the cingulate gyrus among medicated patients with schizophrenia.

Molecular Imaging Studies of the Glutamate System in Schizophrenia

Glutamate, the excitatory neurotransmitter, likely plays multiple roles in the pathogenesis and pathology in schizophrenia. Receptors for glutamate, the excitatory neurotransmitter, include the NMDA receptor. One approach

is to directly image one of the sites of the NMDA receptor channel, the PCP site. This was done with preliminary studies with *N*-(1-naphthyl)-*N′*-(3-[^{123}I]iodophenyl)-*N*-methylguanidine ([^{123}I]CNS1261), which binds to the PCP site in the thalamus and other brain regions. This radiotracer binding is reduced by ketamine, an antagonist of the NMDA receptor, which also binds to this site in part. Ketamine, an antagonist of the NMDA receptor, reduces binding of [^{123}I]CNS1261, a radiotracer binding to the NMDA receptor, measured by SPECT, in the thalamic and other regions of the brain in healthy human adults. Even in normal volunteers, ketamine-induced changes in [^{11}C]raclopride occurred, which was greater than placebo. These changes were not greater than that caused by the intravenous amphetamine, but the ketamine-induced changes were correlated with symptoms seen in schizophrenia. Although initial radiotracer binding studies did not show differences in patients with schizophrenia compared to healthy controls, treatment with clozapine did reduce the binding of [^{123}I]CNS1261, suggesting that clozapine exerts some of its therapeutic actions through glutamatergic mechanisms. At the same time, negative symptoms of schizophrenia are increased as measured by the Brief Psychiatric Rating Scale. This is consistent with the glutamate hypothesis of schizophrenia proposing that reductions of glutamate in the brain lead to negative symptoms of schizophrenia.

Since negative symptoms of schizophrenia were measured to be increased in people with schizophrenia, one possibility is that the glutamate hypothesis in schizophrenia may include reductions of glutamate that lead to negative symptoms. Furthermore, blockade of the NMDA receptor by glutamate antagonist may precipitate positive symptoms of schizophrenia, as has been shown by the administration of ketamine to humans. Even in healthy controls, ketamine induces hallucinations; hence the ketamine challenge is a human model of schizophrenia. Studies involving ketamine infusion have been shown to significantly reduce the binding of [^{11}C]FLB457 in the posterior cingulate/retrosplenial cortices, which suggests intrasynaptic DAR due to competition of endogenous dopamine with the radiotracer binding. This has further been supported by microdialysis studies in rodents. At the same time, ketamine-induced psychotic symptoms were associated with the changes in the radiotracer binding. These results suggested that cortical dopamine mechanisms may relate to ketamine-induced psychotic-like symptoms in humans. The second study showed that psychotomimetic doses of s-ketamine decreased the in vivo binding of [^{11}C]raclopride to striatal D_2-like receptors along with eliciting a psychotic-like symptom with alterations in mood and cognition and hallucinations. This decreased the raclopride BP by approximately 17 percent in the ventral striatum, 14 percent in the caudate, and 14 percent in the putamen, suggesting an increase in intrasynaptic dopamine (i.e., DAR). The change in raclopride BP in the ventral striatum also correlated with

heightened mood, including euphoria and grandiosity. This showed additional evidence that the glutamate NMDA receptor may contribute to psychotic symptoms. In summary then, there is early evidence of this relationship between indirect NMDA antagonists (indirect) and enhanced DAR in the striatum and elsewhere and extrastriatal areas. Such challenge studies with NMDA antagonists may simulate a hypoglutamatergic state, which is increasingly thought to be consistent with schizophrenia and the downstream effects on dopamine and may also underline positive symptoms.

Interestingly, studies with acute ketamine did not alter serotonin binding with [^{18}F]altanserin in healthy volunteers. This was considered by the authors to be more of a technical finding, however, in that this radiotracer, [^{18}F]altanserin, may be more stable in acute ketamine challenges but might be useful in studying chronic plastic effects of ketamine on 5-HT_{2A} receptors. It does not necessarily mean, however, that there is no effect on serotonin of NMDA antagonists, and this remains an unanswered question.

Thus, glutamate and dopamine likely interact in people with schizophrenia. Blockade of the NMDA receptors by glutamate antagonists may precipitate positive symptoms of schizophrenia and produce DAR in the lateral prefrontal and the anterior cingulate cortex.

An important recent finding in the relationship between glutamate and potentially with influence on dopamine is two recent studies. While not a large signal, there is suggestion with a drug *N*-acetylcysteine. There was a group difference that was significant, suggesting a potential displacement of [^{11}C]ABP688, a negative allosteric modulator of glutamate. A second study suggests a displacement in a number of normal human subjects using [^{11}C]ABP688 in the absence and presence of a pharmacologically active IV ketamine challenge. This ketamine challenge was typical of that used by this group to induce pharmacological effects of ketamine as a model for psychosis and hallucinations typically seen in schizophrenia (as a model), but there was significant individual variability. These two findings are important and a fact that could dramatically revolutionize challenge.

MOLECULAR IMAGING IN DRUG DEVELOPMENT

One of the most challenging aspects in psychiatric pharmacotherapy is the administration of clinically effective drug dose that at the same time causes no or minimal side effects. At a given dose, plasma concentrations of psychotropic drugs vary extremely because of large interindividual differences in absorption, metabolism, and excretion. PET and more recently SPECT have been used extensively to characterize the relationships between occupancy of target molecules in the brain (neurotransmitter receptors and transporters) and plasma concentrations of the respective drug. This has been attempted to correlate with clinical

efficacy and side effects. PET occupancy measures have now reached the status of almost surrogate markers that have been proven to be critical in decision making in the development of new psychotropic drugs, especially for choosing a dose for expanded clinical trials. Determination of the appropriate occupancy level for various classes of drugs is vital to streamlining drug development. These studies have also helped elucidate the mechanism of action of some psychotropic drugs, including some of the newer antipsychotics.

D_2 Receptor Occupancy and Beneficial and Adverse Effects of Antipsychotics

With regard to the relationship between drug kinetics in plasma and brain and clinical effects and side effects, antipsychotics belong to the best-explored drugs. Lars Farde and collaborators demonstrated that clinically effective doses of typical neuroleptics occupy D_2-like dopamine receptors in the range between 65 and 90 percent. The suggestion of a "therapeutic window" between 60 and 80 percent striatal D_2 receptor occupancy for sufficient treatment response and a "ceiling" of around 80 percent occupancy for extrapyramidal side effects (EPS) was later confirmed by the same and a number of other groups. Although the class of antipsychotics is characterized by marked pharmacological heterogeneity, the general rule that the incidence of EPS increases in a dose-dependent manner does also apply for most of the second-generation antipsychotics, including olanzapine and risperidone. They lose their "atypical" properties when striatal D_2 receptor occupancy is raised above a threshold of approximately 80 percent. The 5-HT_2 antagonism, which characterizes most of the second-generation antipsychotics, seems to protect from EPS only at moderate doses; this preservation is lost at striatal D_2 receptor occupancies above the 80 percent threshold. Thus, when the doses of these drugs are raised above a certain threshold (usually approximately 6 mg for risperidone, 30 mg for olanzapine, and 160 mg for ziprasidone [Geodon]), striatal D_2 occupancy increases to levels that are associated with a higher incidence of extrapyramidal symptoms (EPS). However, there are a few exceptions from this general rule. Some second-generation antipsychotics do not induce EPS, even when their doses (or plasma concentrations) are raised to unusually high values. In those cases, the upper threshold of the "therapeutic window" is defined by other side effects rather than EPS. This pharmacologically heterogeneous group of antipsychotics can be differentiated by the distinct characteristics of these drugs in PET studies. Clozapine and quetiapine (Seroquel) occupy a maximum of 60 to 70 percent of striatal D_2-like dopamine receptors even at extremely high doses or plasma levels. Twenty-four hours after the last administration of quetiapine, a significant occupancy of striatal D_2 receptors is not detectable

anymore. These particular characteristics of clozapine and quetiapine are most likely due to their low affinity rather than their rapid dissociation (k_{off}) from the D_2 receptor. Thus, the tolerability and safety of these drugs is limited by anticholinergic, antihistaminic, and α-adrenergic rather than EPS side effects. On the contrary, aripiprazole (Abilify) leads to complete saturation (occupancy) of striatal D_2 receptors at clinically used doses, although the incidence of EPS is very low. Again, tolerability is not restricted by the occurrence of EPS but has other side effects such as psychomotor activation (Fig. 12.9–4). The characterization of aripiprazole as a partial dopamine agonist explains this unique feature. Thus, the "therapeutic window" for treatment with most antipsychotic drugs is defined downward by a minimum D_2 receptor occupancy, which has to be achieved, and upward by the EPS threshold. This rule applies both to first- and second-generation antipsychotics, with the exceptions of aripiprazole, clozapine, and quetiapine.

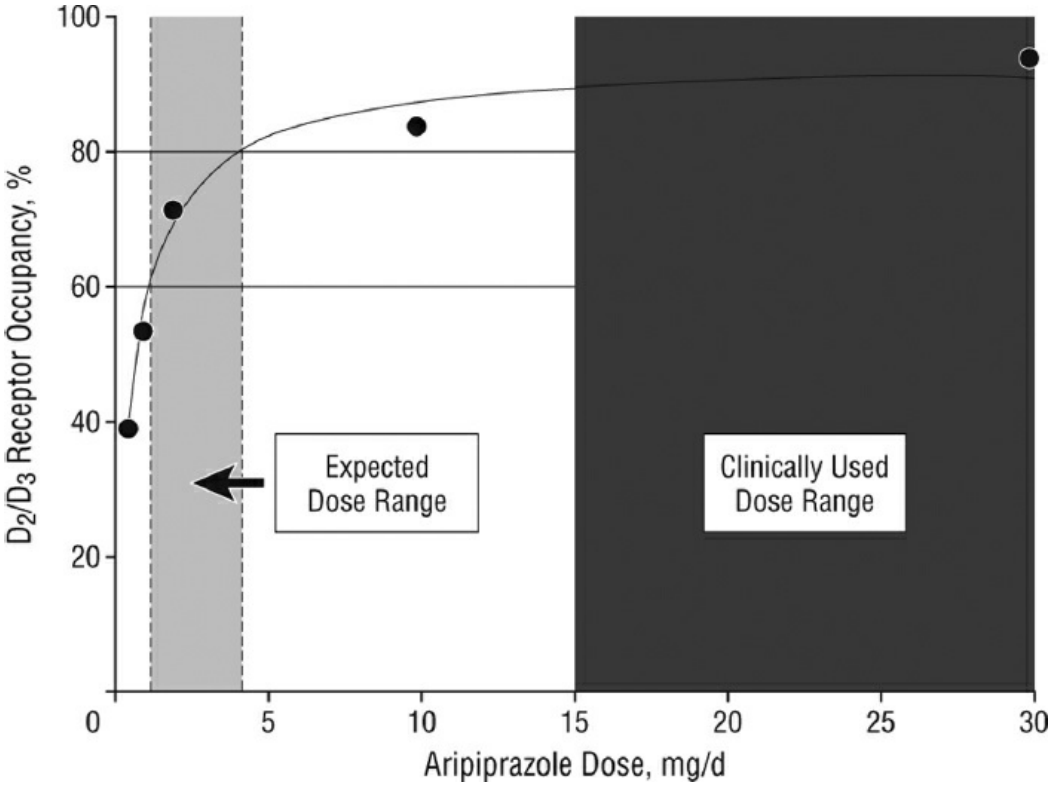

FIGURE 12.9–4. Relationship between aripiprazole (Abilify) daily dose, mean D_2 and D_3 receptor occupancy in the putamen, and clinically expected and used aripiprazole doses. *Horizontal lines* indicate the hypothesized therapeutic window between 60 and 80 percent striatal D_2 receptor occupancy. The *light-shaded area* designates the dose range that should lead to receptor occupancy within this therapeutic window, whereas the *dark-shaded area* indicates the dose range that has proved to have antipsychotic efficacy. (From Gründer G, Carlsson A, Wong DF. Mechanism of new antipsychotic medications: occupancy is not just antagonism. *Arch Gen Psychiatry*. 2003;60:974–977, with permission.)

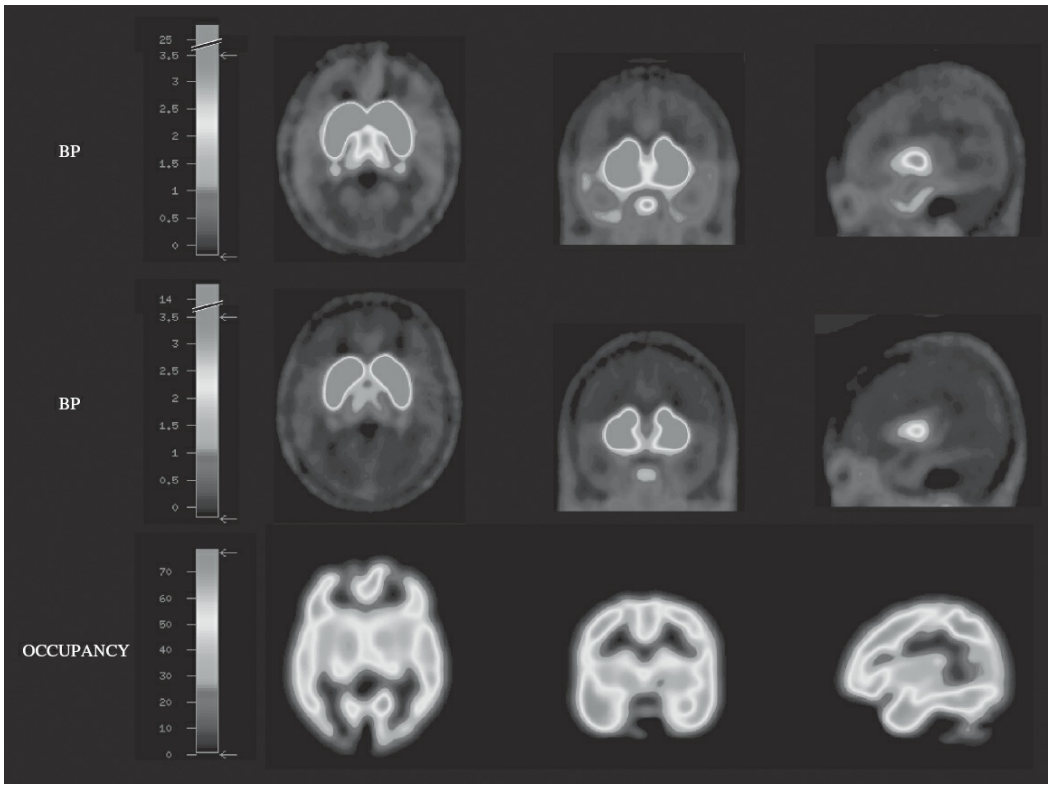

FIGURE 12.9–5. [^{18}F]fallypride positron emission tomography (PET) parametric mean binding potential (BP) images in normal volunteers (*n* = 7, *top row*) and in clozapine (Clozaril)-treated patients with schizophrenia (*n* = 15, *middle row*). The *bottom row* illustrates the D_2/D_3 receptor occupancy in each single voxel in the patient group, clearly demonstrating the significantly higher binding in cortical regions than that in the striatum. (From Gründer G, Landvogt C, Vernaleken I, et al. The striatal and extrastriatal D_2/D_3 receptor-binding profile of clozapine in patients with schizophrenia. *Neuropsychopharmacology*. 2006;31:1027–1035, with permission.)

However, all the aforementioned observations were made with reference to receptor occupancy in striatal structures only. With the advent of high-affinity radiotracers belonging to the class of substituted benzamides such as [^{18}F]fallypride, it became possible to visualize and quantify extrastriatal dopamine receptors (as cortical receptor). Although earlier PET studies with the moderate affinity D_2/D_3 antagonist [^{11}C]raclopride demonstrated that the second-generation antipsychotics, clozapine and quetiapine, occupy striatal D_2/D_3 receptors to a significantly lesser extent than do other antipsychotics, it was later consistently shown in studies with high-affinity ligands that these two compounds nonetheless occupy a significantly higher proportion of temporolimbic than striatal D_2/D_3 receptors. Investigators have demonstrated that the clinical antipsychotic efficacy of clozapine is likely more related to its binding in temporal cortex than to its striatal binding. This “preferential” extrastriatal binding of second-generation antipsychotics has since been shown for several other second-generation antipsychotics, at least to a modest extent (Fig. 12.9–5).

However, there are a number of important clinical and theoretical issues that remain to be resolved. First, the "therapeutic window" of 65 to 80 percent was established in patients with schizophrenia of mostly younger age. Lower occupancy values might be sufficient for relapse prevention in old age, but the studies suggesting this might have been biased by the lack of correction for the decrease in dopamine receptor density with age. Second, further studies with long-acting injectable antipsychotics are necessary to better inform clinicians about doses and injection intervals for long-term relapse prevention with these compounds. Third, molecular imaging might provide important insights into the nondopaminergic mechanisms of current and future antipsychotics. For example, a PET study has shown that the main metabolite of quetiapine, norquetiapine, occupies a substantial fraction of the noradrenaline transporter (NET). Whether the NET occupancy of 20 to 35 percent in the thalamus under doses between 150 and 300 mg daily is sufficient to explain quetiapine's antidepressant and mood-stabilizing effects remains a matter of debate.

Last but not least, while many antipsychotics are antagonists both at D_2 and D_3 (and possibly D_4) dopamine receptors, their effects on the D_3 receptor subtype in vivo in the human brain are quite surprising. Treatment with haloperidol, olanzapine, or risperidone leads to marked reduction of the binding of the D_3 preferring agonist radiotracer [^{11}C]PHNO in the striatum, where its binding represents binding to D_2 receptors. However, antipsychotic treatment leads to an increase in tracer binding rather than a decrease in brain structures where [^{11}C]PHNO binding represents binding to D_3 receptors, namely, the D_3-rich regions globus pallidus and substantia nigra. Thus, while (most) antipsychotics act as D_3 antagonists in vitro, they seem to lead to an upregulation of D_3 receptors, a finding that is yet completely unexplained. This increased D_3 binding can be completely blocked with a single dose of the D_3-preferring dopamine agonist pramipexole. Because of its preferential binding to D_3 receptors, [^{11}C]PHNO is now being used as a tool to explore the functional role of the D_3 receptor in neuropsychiatric disorders. The first occupancy studies with selective D_3 antagonists such as ABT-925 have been performed, and they demonstrate that [^{11}C]PHNO binding can be markedly displaced by 75 and 65 percent, respectively, in the substantia nigra and the globus pallidus by ABT-925, while the displacement is much lower in caudate and putamen.

Target Engagement of Compounds with Novel Mechanism of Action

Molecular imaging techniques have been more recently used for determination of target occupancy of some novel compounds with

nondopaminergic mode of action. The most important example is the group of glycine transporter type 1 (GlyT1) inhibitors. Glycine acts as a cotransmitter at the NMDA glutamate receptor, which is a calcium ion channel activated by NMDA in the presence of glycine. It has been hypothesized that inhibition of the GlyT1 by increasing the availability of glycine at the NMDA receptor attenuates NMDA receptor malfunction and represents a major neurobiological underpinning of schizophrenic psychopathology. Following initial smaller positive studies with the naturally occurring GlyT1 inhibitor sarcosine, several compounds with higher affinity and selectivity were synthesized and clinically tested. The occupancy of the glycine transporter by one of these compounds, RG1678 (bitopertin), was recently assessed with PET and [^{11}C]RO5013853, a radiotracer selective for the GlyT1, which was specifically developed for this purpose. Healthy male volunteers received bitopertin doses between 5 and 175 mg, and a clear nonlinear dose/occupancy relationship could be established, with the highest dose leading to complete saturation of the GlyT1. Interestingly, in the first phase II study in patients with schizophrenia, bitopertin improved negative symptoms at doses of 10 mg and 30 mg daily, while the highest applied dose (60 mg) was ineffective. Both studies together suggested that mild to moderate antagonism of the GlyT1 rather than complete blockade could be associated with beneficial effects of this class of drugs. Although bitopertin subsequently failed in phase III clinical trials for schizophrenia indications, studies in nonhuman primates confirmed that pharmacologically induced cognitive deficits could be reversed by the GlyT1 inhibitors bitopertin and Org 25935. This reversal of cognitive dysfunction induced by ketamine or scopolamine was achieved in the GlyT1 occupancy range of 30 to 70 percent, while higher occupancy was not effective. These studies taken together nicely demonstrate the translational potential of molecular imaging in conjunction with clinical studies in patients for the elucidation of the mechanism of action of novel compounds for the treatment of specific dimensions of schizophrenia psychopathology. While bitopertin failed for various reasons, the full potential of GlyT1 inhibition remains to be explored.

Other potentially important targets for future drug treatment of schizophrenia that are currently being intensely explored are both α4β2 and α7 nAChRs and phosphodiesterase type 10A. One new and promising PET tracer [^{18}F]ASEM is capable of demonstrating occupancy in schizophrenia with some α7-nAChR drugs.

MOLECULAR IMAGING AND GENETICS

Schizophrenia has a heritability of about 80 percent, but the detailed molecular genetic basis of the disease remains to be fully determined. Molecular brain imaging as applied to the study of schizophrenia can be a useful means to quantify mediators of this genetic vulnerability. Two

approaches have been applied. Molecular imaging studies have been performed in individuals at genetic risk for schizophrenia. Increased cortical D_1 and caudate D_2 densities have been observed in monozygotic compared to dizygotic siblings of patients with schizophrenia. First-degree relatives of patients with schizophrenia have displayed increased uptake of 6-[^{18}F]fluorodopa ([^{18}F]FDOPA) in caudate–putamen. These observations in at-risk samples are consistent with neuroreceptor-binding studies in patients with schizophrenia who have demonstrated increased D_1, D_2, and FDOPA uptake. These studies highlight the potential role of molecular brain imaging in identifying genetic markers for schizophrenia as well as individuals at risk for the disease who might benefit from pharmacological intervention.

Furthermore, the potential role of molecular brain imaging in identifying genetic markers for schizophrenia is confirmed by the findings of a decrease in the prefrontal and cortical binding of 5-HT_{2A} receptors using PET and [^{18}F]altanserin in subjects at enhanced risk of schizophrenia. Those subjects at risk who developed a first-episode psychosis showed an additional decrease of 5-HT_{2A} receptor BP in the caudate in comparison to the "nonconverters," suggesting that the progressive reduction of 5-HT_{2A} receptor binding in the caudate could provide an indicator of illness activity and help predict imminent conversion to schizophrenia. This study's findings, although interesting, may be diminished by the very low concentration of 5-HT_{2A} receptors in the caudate that makes it problematic to measure their binding.

The second approach is to characterize genetic polymorphisms of the same transporters or receptors that are being imaged.

The use of molecular imaging also seems helpful in clarifying the contribution that different genotypes associated with dopamine receptors play in defining the phenotypic expression of the receptors themselves in schizophrenia. Impaired DAT function has been thought to be involved in antipsychotic-induced EPS. Nonetheless, studies so far conducted have been negative to this regard. The frequency of the more common alleles of the variable number of tandem repeats (VNTR) polymorphism of DAT gene (*SLC6A3*) was not preferentially associated with patients who had EPS. Furthermore, no significant differences in DAT BP have been found among the three major VNTR genotype groups. In addition, no significant association has been found between VNTR polymorphism and amphetamine-induced DAR and DAT BP density in patients with schizophrenia. These studies seem to suggest a lesser role for DAT as compared to D_2 or D_1 receptors in the pathophysiology of schizophrenia. There is a need for research to determine whether common gene variants of the brain dopamine pathway explain the variability in nicotine-induced DAR. An overall decrease in ventral caudate/nucleus accumbens BP in

those who smoked compared with those who did not smoke was also found. These findings suggest that dopamine system genotype variabilities can explain a significant proportion of the interindividual variability in smoking-induced DAR and indicate that smoking-induced DAR has a genetic predisposition. COMT has been shown to be critical for prefrontal dopamine flux and prefrontal cortex-dependent cognition and activation. Several potentially functional variants in the gene have been identified, but considerable controversy exists regarding the contribution of individual alleles and haplotypes to risk for schizophrenia. A pattern of inefficient prefrontal cortex engagement during working memory studied with fMRI has been found in normal volunteers in association with the Val158Met polymorphism of the COMT gene. A polymorphism of the metabotropic glutamate receptor 3 or GRM3, a G protein-linked receptor modulating synaptic glutamate, has also been associated, even in normal subjects, with a pattern of inefficient prefrontal cortex engagement during working memory studied with fMRI analogous to that found with COMT. A recent study has observed an apparent epistatic interaction of these two genes on the engagement of prefrontal cortex during working memory. Specifically, the GRM3 genotype putatively associated with suboptimal glutamatergic signaling was significantly associated with inefficient prefrontal engagement and altered prefrontal–parietal coupling on the background of COMT Val-homozygous genotype. Conversely, COMT Met-homozygous background mediated against the effect of GRM3 genotype. These findings extend putative brain dopaminergic and glutamatergic relationships indexed by COMT and GRM3 to a systems-level interaction in human cortical circuits implicated in working memory dysfunction such as in schizophrenia.

MOLECULAR IMAGING IN CLINICAL PRACTICE

Molecular Imaging in the Treatment in Schizophrenia

Molecular imaging of psychotic disorders is still a research tool rather than a method to evaluate treatment effectiveness for patients. However, the findings that antipsychotics typically produce maximal therapeutic effects with minimal adverse effects when 65 to 80 percent of the (striatal) dopamine D_2 receptors are occupied (see previously and Figs. 12.9–2 through 12.9–5) have important implications for clinical practice. Once a plasma concentration range that leads into the "therapeutic occupancy window" has been established by means of PET, it can be tested whether this range is the most useful in large patient populations. This has indeed been demonstrated for various antipsychotic drugs. A well-documented and clinically important example is risperidone. PET studies show that at 6-mg risperidone daily, which is the highest dose administered under usual clinical conditions, a significant proportion of patients present with striatal

D_2-like occupancies sufficient to cause EPS, while a daily oral dose of risperidone of 3 to 4 mg produces a striatal D_2-like occupancy in the range of 70 to 80 percent in most patients. The available data suggest that the 80 percent threshold associated with a higher probability of EPS is reached with a plasma concentration of approximately 40 ng/mL. A plasma concentration of approximately 15 ng/mL is associated with 60 percent striatal D_2 occupancy. The same therapeutic window seems to apply for long-acting injectable risperidone. Clinical studies with long-acting injectable risperidone at various doses have demonstrated that a dose of 75 mg (used in clinical testing but is not approved) is associated with a relatively high EPS incidence and correspondingly high plasma concentrations above 40 ng/mL, while most patients treated within the approved dose range of 25 to 50 mg every 2 weeks present with plasma concentrations in the range of 15 to 40 ng/mL. Indeed, the 25-mg dose might be too low in some patients. This example illustrates that the knowledge of an individual's plasma level can provide meaningful information on psychotropic drug concentrations at brain target molecules.

FUTURE DIRECTIONS

The future of molecular imaging in schizophrenia is closely linked to the progress in the discovery and understanding of the molecular mechanisms related to the etiology and pathophysiology of the disease. The role of the glutamatergic system has become increasingly clear over the past decade. Both NMDA and non-NMDA glutamate receptors have been implicated in the pathophysiology of positive, negative, and cognitive symptoms of schizophrenia. A hypofunction of cortical glutamate has been postulated to interact with subcortical dopamine to cause increased release of presynaptic dopamine and elevated numbers of postsynaptic dopamine receptors. There exists a need to develop ligands for glutamatergic receptors such as metabotropic glutamatergic receptors type 2/3 (mGlu2/3). While pomaglumetad (LY404039) failed in clinical trials in schizophrenia, a PET dose-finding had not been performed due to the lack of a radiotracer. A specific radiotracer for this target would allow to further explore its functional relevance and therapeutic potential. Likewise, there exists a tremendous future role for ligands for (1) the glycine and serine modulatory sites of the NMDA receptor and (2) the enzyme racemase that regulates the production of d-serine in the brain. Furthermore, the study of nicotinic receptors in vivo in patients with schizophrenia has recently become possible with the availability of probes (especially the α-7 subtype) for these receptors that likely play a significant role in the mechanisms of schizophrenia and can aid in target engagement for cognitive enhancement therapy in schizophrenia. New probes are also needed to image those neuronal structures (i.e., cytoskeleton proteins) that are involved in

neuronal migration and synaptic formation and that form the basis of the neurodevelopment hypothesis of schizophrenia. The demonstration of microglial activation utilizing PET and (*R*)-[^{11}C]PK11195 in people with recent-onset schizophrenia suggests a promising avenue for future research, which can now be explored with better radiotracers such as [^{11}C]DPA-713 [^{11}C]PBR28, [^{11}C]AC-5216, or [^{18}F]FEPPA.

ACKNOWLEDGMENT

The authors thank Nicola Cascella, MD, for prior contributions.

REFERENCES

Anderson BA, Kuwabara H, Wong DF, et al. The role of dopamine in value-based attentional orienting. *Curr Biol*. 2016;26(4):550–555.

Boileau I, Payer D, Chugani B, et al. *In vivo* evidence for greater amphetamine-induced dopamine release in pathological gambling: a positron emission tomography study with [^{11}C]-(+)-PHNO. *Mol Psychiatry*. 2014;19:1305–1313.

Brašic´ JR, Cascella N, Kumar A, et al. Positron emission tomography (PET) experience with 2-[18F]fluoro-3-(2(*S*)-azetidinylmethoxy)pyridine (2-[18F]FA) in the living human brain of smokers with paranoid schizophrenia. *Synapse*. 2012;66:352–368.

Caravaggio F, Borlido C, Wilson A, Graff-Guerrero A. Examining endogenous dopamine in treated schizophrenia using [^{11}C]-(+)-PHNO positron emission tomography: a pilot study. *Clin Chim Acta*. 2015;449:60–62.

Carlson PJ, Diazgranados N, Nugent AC, et al. Neural correlates of rapid antidepressant response to ketamine in treatment-resistant unipolar depression: a preliminary positron emission tomography study. *Biol Psychiatry*. 2013;73:1213–1221.

Ceccarini J, De Hert M, Van Winkel R, et al. Increased ventral striatal CB_1 receptor binding is related to negative symptoms in drug-free patients with schizophrenia. *Neuroimage*. 2013;79:304–412.

Celen S, Koole M, Ooms M, et al. Preclinical evaluation of [^{18}F]JNJ42259152 as a PET tracer for PDE10A. *Neuroimage*. 2013;82:13–22.

Coughlin JM, Wang Y, Ambinder EB, et al. In vivo markers of inflammatory response in recent-onset schizophrenia: a combined study using [(11)C]DPA-713 PET and analysis of CSF and plasma. *Transl Psychiatry*. 2016;6:e777. doi:10.1038/tp.2016.40.

Davis RE, Vanover KE, Zhou Y, et al. ITI-007 demonstrates brain occupancy at serotonin 5-HT_{2A} and dopamine D_2 receptors and serotonin transporters using positron emission tomography in healthy volunteers. *Psychopharmacology (Berl)*. 2015;232:2863–2872.

DeLorenzo C, DellaGioia N, Bloch M, et al. In vivo ketamine-induced changes in [^{11}C]ABP688 binding to metabotropic glutamate receptor subtype 5. *Biol Psychiatry*. 2015;77:266–275.

Esterlis I, Ranganathan M, Bois F, et al. In vivo evidence for β2 nicotinic acetylcholine receptor subunit upregulation in smokers as compared with nonsmokers with schizophrenia. *Biol Psychiatry*. 2014;76:495–502.

Fusar-Poli P, Meyer-Lindenberg A. Striatal presynaptic dopamine in schizophrenia, part II: meta-analysis of [^{18}F/^{11}C]-DOPA PET studies. *Schizophr Bull*. 2013;39:33–42.

George N, Gean EG, Nandi A, et al. Advances in CNS imaging agents: focus on PET and SPECT tracers in experimental and clinical use. *CNS Drugs*. 2015;29:313–330.

Harada A, Suzuki K, Miura S, et al. Characterization of the binding properties of T-773 as a PET radioligand for phosphodiesterase 10A. *Nucl Med Biol*. 2015;42:146–154.

Hostetler ED, Sanabria-Bohórquez S, Eng W, et al. Evaluation of [^{18}F]MK-0911, a positron emission tomography (PET) tracer for opioid receptor-like 1 (ORL1), in rhesus monkey and human. *Neuroimage*. 2013;68:1–10.

Howes OD, Shotbolt P, Bloomfield M, et al. Dopaminergic function in the psychosis spectrum: an [^{18}F]-DOPA imaging study in healthy individuals with auditory hallucinations. *Schizophr Bull*. 2013;39:807–814.

Kang JI, Park H-J, Kim SJ, et al. Reduced binding potential of GABA-A/benzodiazepine receptors in individuals at ultra-high risk for psychosis: an [^{18}F]-fluoroflumazenil positron emission tomography study. *Schizophr Bull*. 2014;40:548–557.

Kenk M, Selvanathan T, Rao N, et al. Imaging neuroinflammation in gray and white matter in schizophrenia: an in-vivo PET study with [^{18}F]-FEPPA. *Schizophr Bull*. 2015;41:85–93.

Kim J, Horti AG, Mathews MB, et al. Quantitative multi-modal brain autoradiography of glutamatergic, dopaminergic, cannabinoid, and nicotinic receptors in mutant disrupted-in-schizophrenia-1 (DISC1) mice. *Mol Imaging Biol*. 2015;17:355–363.

Kim J-H, Son Y-D, Kim J-H, et al. Serotonin transporter availability in thalamic subregions in schizophrenia: a study using 7.0-T MRI with [^{11}C]DASB high-resolution PET. *Psychiatry Res*. 2015;231:50–57.

Kreisl WC, Jenko KJ, Hines CS, et al. A genetic polymorphism for translocator protein 18 kDa affects both in vitro and in vivo radioligand binding in human brain to this putative biomarker of neuroinflammation. *J Cereb Blood Flow Metab*. 2013;33:53–58.

Kuwabara H, Gao Y, Stabin M, et al. Imaging α4β2 nicotinic acetylcholine receptors (nAChRs) in baboons with [18F]XTRA, a radioligand with improved specific binding in extra-thalamic regions. *Mol Imaging Biol*. 2016. doi:10.1007/s11307-016-0999-9.

Leurquin-Sterk G, Postnov A, de Laat B, et al. Kinetic modeling and long-term test-retest reproducibililty of the mGluR5 PET tracer 18F-FPEB in human brain. *Synapse*. 2016;70(4):153–162.

Martin-Facklam M, Pizzagalli F, Zhou Y, et al. Glycine transporter type 1 occupancy by bitopertin: a positron emission tomography study in healthy volunteers. *Neuropsychopharmacology*. 2013;38:504–512.

Mizrahi R, Kenk M, Suridjan I, et al. Stress-induced dopamine response in subjects at clinical high risk for schizophrenia with and without concurrent cannabis use. *Neuropsychopharmacology*. 2014;39:1479–1489.

Nakajima S, Takeuchi H, Plitman E, et al. Neuroimaging findings in treatment-resistant schizophrenia: a systematic review: lack of neuroimaging correlates of treatment-resistant schizophrenia. *Schizophr Res*. 2015;164:164–175.

Oswald LM, Wand GS, Wong DF, et al. Risky decision-making and ventral striatal dopamine responses to amphetamine: a positron emission tomography [11C]raclopride study in healthy adults. *Neuroimage*. 2015;113:26–36.

Pasternak O, Kubicki M, Shenton ME. In vivo imaging of neuroinflammation in schizophrenia.

Schizophr Res. 2016;173(3):200–212.

Salavati B, Rajji TK, Price R, et al. Imaging-based neurochemistry in schizophrenia: a systematic review and implications for dysfunctional long-term potentiation. *Schizophr Bull*. 2015;41:44–56.

Selvaraj S, Arnone D, Cappai A, Howes O. Alterations in the serotonin system in schizophrenia: a systematic review and meta-analysis of postmortem and molecular imaging studies. *Neurosci Biobehav Rev*. 2014;45:233–245.

Umbricht D, Alberati D, Martin-Facklam M, et al. Effect of bitopertin, a glycine reuptake inhibitor, on negative symptoms of schizophrenia: a randomized, double-blind, proof-of-concept study. *JAMA Psychiatry*. 2014;71:637–646.

Vasconcelos GS, Ximenes NC, de Sousa CN, et al. Alpha-lipoic acid alone and combined with clozapine reverses schizophrenia-like symptoms induced by ketamine in mice: participation of antioxidant, nitrergic and neurotrophic mechanisms. *Schizophr Res*. 2015;165:163–170.

Winton-Brown TT, Fusar-Poli P, Ungless MA, Howes OD. Dopaminergic basis of salience dysregulation in psychosis. *Trends Neurosci*. 2014;37:85–94.

Wong DF, Kuwabara H, Pomper M, et al. Human brain imaging of alpha7 nAChR with [^{18}F]ASEM: a new PET radiotracer for neuropsychiatry and determination of drug occupancy. *Mol Imaging Biol*. 2014;16:730–738.

▲ 12.10 Neurocognition in Schizophrenia

PHILIP D. HARVEY, PH.D., RICHARD S.E. KEEFE, PH.D., AND CHARLES E. EESLEY, PH.D.

NEUROCOGNITIVE IMPAIRMENT IN SCHIZOPHRENIA

Neurocognition as a Core Feature of Schizophrenia

Neurocognitive impairment in schizophrenia is clinically relevant and profound. Patients with schizophrenia perform one to two standard deviations below healthy controls on various neurocognitive tests. The severity of this impairment is greatest in the domains of memory, attention, working memory, problem solving, processing speed, and social cognition. These deficits are present prior to the initiation of antipsychotic treatment and are not caused by psychotic symptoms in those patients who are able to complete testing, which includes the overwhelming majority. Many of the various neurocognitive deficits in schizophrenia have been shown to be associated with functional outcomes such as difficulty with community functioning, difficulty with instrumental and problem-solving skills, reduced success in psychosocial rehabilitation programs, and the inability to maintain successful employment. Neurocognitive deficits are better able to explain important functional outcomes such as work performance and independent living than positive or negative symptoms.

The emphasis on neurocognitive deficits in schizophrenia goes beyond their severity and relation to functional outcomes. Neurocognitive deficits

are present to an extent in some patients with schizophrenia prior to the onset of psychosis, are prominent during the prodrome, and are correlated with measurable brain dysfunction more than any other aspect of the illness. Perhaps most importantly, neurocognition is increasingly considered as a primary target for treatment.

Despite the importance of neurocognitive impairment in the biology, function, and treatment of schizophrenia, and despite the emphasis placed on neurocognitive impairment by Emil Kraepelin and Eugen Bleuler more than 100 years ago, it still is not included as a criterion for diagnosis in the Fifth Edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) or the eleventh edition of *International Statistical Classification of Diseases and Related Health Problems* (ICD-11) criteria. It is noteworthy, however, that this newest edition of the DSM has extensive discussions of cognition and its relation to disability. These discussions include the recommendation that assessment of cognition and everyday functioning should be a routine part of the assessment of schizophrenia and the assessment scales section of the DSM-5 (Section 3) includes assessment of the severity and types of cognitive deficits. Thus, recognition of the importance of cognition in schizophrenia is now truly "official" and viewed as centrally important to the assessment of the condition.

Profile and Magnitude of Neurocognitive Impairment Associated with Schizophrenia

The level of severely of impairment in performance on neurocognitive tests is strong evidence for the importance of neurocognitive deficits in schizophrenia. Neurocognitive deficits appear to be present in patients with schizophrenia early in the course of the illness. A number of studies have demonstrated that, on various neurocognitive test domains, patients with either first-episode or chronic schizophrenia perform more than a standard deviation below healthy controls, with some domains much more impaired (Fig. 12.10–1). Furthermore, those cases who are apparently in a prodromal state and convert to psychosis have more severe cognitive deficits at the time of the detection of their state than those cases who do not convert. The most striking aspect of the profile of neurocognitive deficits in patients with schizophrenia is that so few neurocognitive functions remain similar in level of performance compared to healthy controls. In fact, an estimate of the severity of neurocognitive impairment in patients with schizophrenia, based on an examination of the most methodologically sound studies completed, suggests that many important neurocognitive functions are in the impaired range (one to two standard deviations below the normal mean). Meta-analysis of more than 300 studies shows a consistent and stable difference between patients with schizophrenia and healthy controls in a wide range of domains of neurocognitive functioning (Fig. 12.10–2).

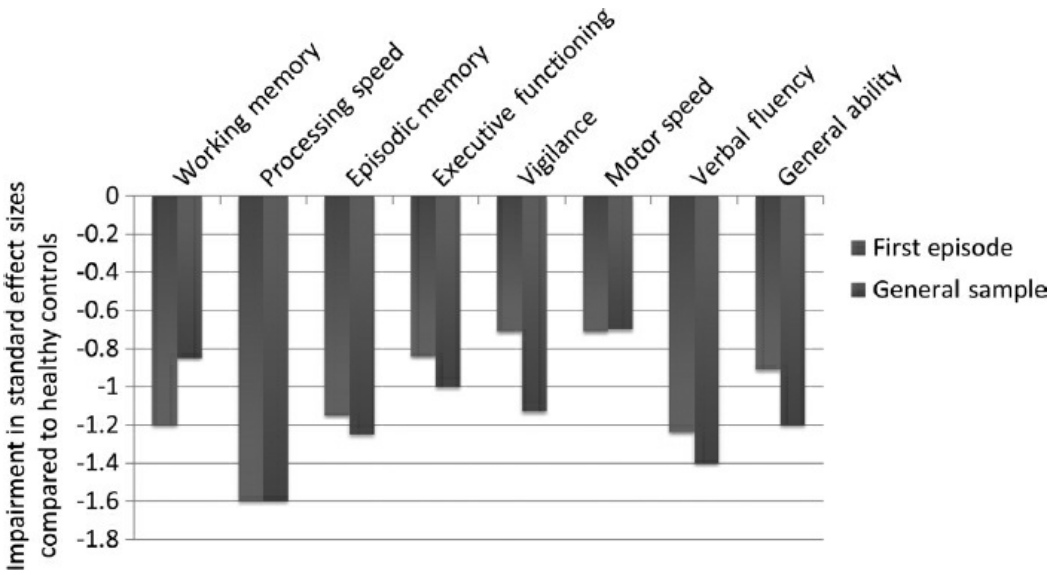

FIGURE 12.10–1. Impairment in major cognitive domains in first-episode patients compared to patients with schizophrenia in general. Based on two systematic meta-analyses: Mesolam-Gately et al. (2009) for first-episode patients and Dickinson et al. (2007) for patients in general. (Adapted from Mesholam-Gately RI, Guiliano AJ, Goff KP. Neurocognition in first episode schizophrenia: a meta-analytic review. *Neuropsychology*. 2009;23:315–336; Dickinson D, Ramsey ME, Gold JM. Overlooking the obvious: a meta-analytic comparison of digit symbol coding tasks and other cognitive measures in schizophrenia. *Arch Gen Psychiatry*. 2007;64:532–542.)

Neurocognitive tests often assess more than one domain of functioning, and many tests do not fit neatly into a single domain. Thus, descriptions of the profile of neurocognitive deficits in schizophrenia have varied across literature reviews and can be confusing to the nonpsychologist. The domains that are most consistently cited as being severely impaired in schizophrenia are verbal memory, executive functions, attention or vigilance, verbal fluency, and motor speed. Deficits in social cognition also appear to be severe, but there are less normative data available on healthy people to provide information about the severity of impairment. The opinion of a group of experts who served on the Neurocognition Subcommittee for the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Project is that the most important domains of neurocognitive deficit in schizophrenia are working memory, attention/vigilance, verbal learning and memory, visual learning and memory, reasoning and problem solving, speed of processing, and social cognition. These domains were chosen because they were viewed as substantially impaired, separable from one another, and to have demonstrated relationships to functional outcomes such as independent community living.

Attention/Vigilance. *Vigilance* refers to the ability to maintain attention over time. A standard vigilance test used in many studies is the Continuous Performance Test (CPT). The “Identical Pairs” version of the

CPT requires that patients attend to a series of two- to four-digit numbers presented sequentially on a computer screen at a rate of one per second. Patients respond with a button press on the computer mouse or keyboard each time a number is identical to the previous number. This type of CPT reveals moderately severe vigilance impairments in patients with schizophrenia. Many neurocognitive tests require intact vigilance for good performance, even if the test itself is not a measure of "pure" vigilance.

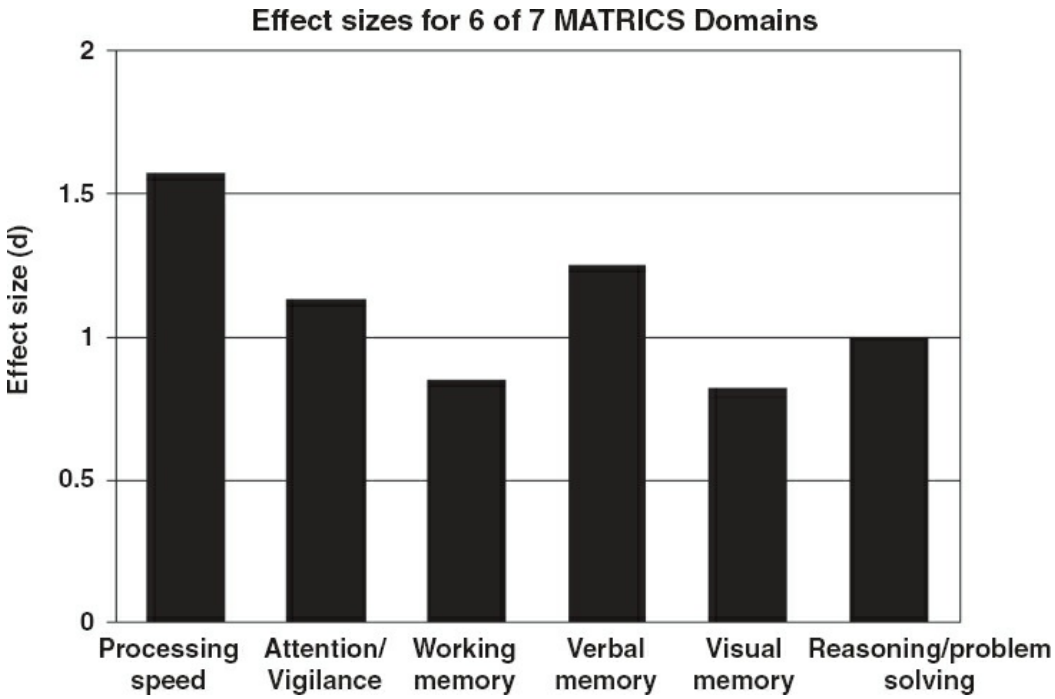

FIGURE 12.10–2. Effect sizes for six of the seven Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) domains. The following tests were chosen to represent each domain. There was not enough information for the social cognition domain, so it is not represented in this figure. Processing speed: Digit symbol coding test; Attention/Vigilance: AX CPT (the AX version of the Continuous Performance Test); Working memory: Letter–number sequencing test; Verbal memory: Word list learning test; Visual memory: Visual reproduction test; Reasoning/Problem solving: Wisconsin Card Sorting Test (WCST) categories.

Impairments in vigilance can result in difficulty following social conversations and an inability to follow important instructions regarding treatment, therapy, or work functions. Simple activities such as reading or watching television can become labored or impossible. Reviews of the literature have suggested that vigilance deficits in patients with schizophrenia are related to various aspects of outcome, including social deficits, community functioning, and skills acquisition.

Verbal Learning and Memory. *Verbal memory* functioning includes, but is not limited to, abilities associated with learning new information, retaining newly learned information over time, and recognizing previously presented material. In general, patients show larger deficits in learning

than in retention. The findings for recognition are more equivocal, with most studies suggesting relatively mild deficits, although large deficits have also been reported.

The tests used to measure learning typically involve the ability to learn lists of words or written passages. Verbal list learning tasks usually require the patient to listen to 12 to 16 words, then immediately recall as many of the words as possible. Using the California Verbal Learning Test, healthy controls can recall about 8 of 16 words after the first trial; patients with schizophrenia can recall only about five. After five consecutive trials of the same word list, most controls can recall at least 13 of the words, while patients with schizophrenia, on average, can recall only nine. Thus, patients with schizophrenia are not only impaired in their ability to immediately recall verbal material, they are also impaired compared to controls in their ability to learn over time. Patients are also impaired in recalling more engaging verbal material, such as stories. Much empirical evidence points to the connection of verbal memory impairment with impairments in everyday activities and occupational deficits in patients with schizophrenia. Success in many modern endeavors requires the ability to learn information that is presented in a verbal context. Interestingly, there is little evidence that people with schizophrenia forget more verbal material that they have learned than healthy controls, with rapid forgetting being much more common in cortical dementias such as Alzheimer disease.

Visual Learning and Memory. Since visual information is not as easily expressed as verbal information, fewer tests sensitive to the deficits of schizophrenia have been developed. Most tests expose subjects to one or more verbal figures and require the subjects to draw them from memory or to indicate which among an array of figures was previously presented.

Studies of the relationship between poor visual memory and functional outcome have yielded mixed findings. Visual memory has been found to correlate modestly with employment status, job tenure, psychosocial rehabilitation success, social functioning, quality-of-life ratings, and strongly with functional capacity. Other studies have reported no significant correlations. It may be that tests of visual memory are simply less sensitive to impairment or have other measurement problems.

Reasoning and Problem Solving. Although there are numerous tests of reasoning and problem solving, also described in some contexts as "executive functioning," the most well-known and frequently utilized in schizophrenia research is the Wisconsin Card Sorting Test (WCST). In this test, patients are given a deck of cards with various numbers of colored shapes on them and are asked to match their cards to four "key" cards that also have shapes on them that differ by color, form (i.e., shape), and number. The first principle to which the subject needs to learn to sort the

cards is *color*. After a patient demonstrates that he or she has learned that *color* is the correct sorting principle, the principle changes to *form* without warning. Repeated sorting attempts by the previously correct principle are referred to as *perseverations*, and patients with schizophrenia, like patients with frontal lobe damage, often make these errors. In fact, the very poor performance of patients with schizophrenia on the WCST and the reduced activity of the dorsolateral prefrontal cortex during performance of this test led to widespread pursuit of the hypothesis of frontal hypoactivation in schizophrenia. It is important to note, however, that the WCST measures a variety of neurocognitive functions including vigilance as well as visual memory and is not a "pure" measure of executive functions.

The rules of society and the workplace change regularly, and success in these arenas is often measured by one's ability to adapt to changes. Patients with schizophrenia who are impaired on measures of reasoning and problem-solving often have difficulty adapting to the rapidly changing world around them.

Speed of Processing. Many neurocognitive tests require subjects to process information rapidly and can be compromised by impairments in processing speed. A standard example of this type of task is the Wechsler Adult Intelligence Scale Digit Symbol Test. Each numeral (1 through 9) is associated with a different simple symbol. Subjects are required to copy as many of the symbols associated with the numerals as possible in 120 seconds. This nonspecific neurocognitive impairment has been found to correlate with a variety of clinically important features of schizophrenia, such as daily life activities, job tenure, and independent living status. It is also sensitive to medication side effects such as somnolence and extrapyramidal symptoms. It has been argued that processing speed tasks such as digit symbol and symbol coding may be the strongest neurocognitive discriminator between patients with schizophrenia and healthy controls. Neurocognitive data from 1,331 patients with schizophrenia assessed prior to treatment randomization for the NIMH CATIE schizophrenia trial suggested that a digit symbol task accounted for more variance in the overall composite score than any other measure. Thus, while seemingly a small component of the profile of neurocognitive deficits in schizophrenia, brief measures of processing speed have a very important role in characterizing relevant aspects of neurocognitive impairment.

The relevance of these deficits is clearly demonstrated by the effect of reduced processing speed on the ability of some patients with schizophrenia to keep in step with the task-oriented jobs that are frequently held by patients with schizophrenia. Psychomotor slowing has the potential to affect many different aspects of everyday life as well as hampering performance on many other cognitive testing measures.

Verbal Fluency. Most neurocognitive assessments in treatment studies of schizophrenia have included measures of verbal fluency as a separate domain of functioning. Most of tests of verbal fluency measure either phonological fluency (also referred to as letter fluency) or semantic fluency. *Phonological fluency* refers to a patient's ability to produce as many words as possible beginning with a particular letter (e.g., "F") within, for instance, 60 seconds. *Semantic fluency* refers to the ability to produce words within a particular meaning-based category, such as "animals." Not only do patients with schizophrenia produce fewer words than normal controls but they are more likely than healthy people to produce words that are not outside the required category. Impaired verbal fluency can damage functioning in social and vocational settings by making communication difficult and awkward.

Immediate/Working Memory. *Immediate memory* refers to the ability to maintain a limited amount of information "online" for a brief period of time (usually a few seconds). Repeating a string of digits (digits forward) is an example of immediate memory. The definition of working memory, on the contrary, is more complex and varies across studies. Some investigators consider working memory to be synonymous with immediate memory, whereas others describe that it should require some manipulation of the information being held online. For example, repeating a series of digits in the reverse order than they were presented (digits backward) requires maintenance of the initial string and an active manipulation because the information needs to be both held online and then subsequently reordered. Patients typically show deficits on both tasks. Visual working memory tasks involve keeping a visuospatial stimulus, such as a dot on a computer screen, or series of visual objects in mind while working on a related or unrelated task.

Working memory has been described by various authors as a core component of the neurocognitive impairment in schizophrenia and is related to functional outcomes such as employment status and job tenure. Much of the clinical relevance of working memory deficits in schizophrenia comes from strong correlations that working memory measures have with a variety of other neurocognitive domains impaired in schizophrenia, such as attention, planning, memory, and intelligence. In addition, the advanced understanding of the neuroanatomy of working memory functions in human and nonhuman primates adds to the current understanding of impairments in working memory in schizophrenia. This neuroanatomical work has suggested that neural circuitry that includes prefrontal cortical regions mediates aspects of working memory functions, and that this circuitry may be impaired in schizophrenia.

Social Cognition. Theory-of-mind skills and social and emotional

perception have been the general focus of the literature on social cognition in schizophrenia. *Theory-of-mind* is the ability to infer about another's intentions and/or to understand the mental states of others. Individuals with schizophrenia perform poorly on measures of theory-of-mind abilities. The evidence regarding whether impairments in theory-of-mind skills are independent of a general neurocognitive deficit is mixed, although the most recent data suggest minimal overlap. Facial affect recognition and social cue perception are the two general areas into which studies of social perception in schizophrenia can be broken down. Reviews of the literature on facial affect recognition suggest that individuals with schizophrenia have stable deficits on tests of facial affect perception compared with healthy controls and psychiatric controls, and that perception of negative emotions and fear may be particularly impaired. Tests of social cue perception use more dynamic stimuli that require multiple sensory modalities, such as watching videotapes of persons interacting. Patients with schizophrenia show consistent impairments on these tasks. In particular, they have more difficulty discerning other individuals' goals and intentions than what they are wearing or saying.

Although a variety of reports have described the strong relationship between general neurocognitive impairments and social deficits in schizophrenia, there is growing evidence that social cognition is more strongly related to social impairments in schizophrenia, even after controlling for performance on neurocognitive tasks. Path analysis has been used to suggest that the relationship between neurocognitive impairment and social deficits is almost entirely mediated by impairments in social cognition.

Prevalence of Neurocognitive Deficits in Schizophrenia

About 15 percent of patients with schizophrenia (and 85 percent of the general population) are rated as "unimpaired" by clinical neuropsychological assessment. However, these patients tend to have the highest levels of premorbid functioning and demonstrate neurocognitive functioning that is considerably below what would be expected of them based on their premorbid levels and the educational level of their parents. Data from 150 patients with schizophrenia suggest that up to 98 percent of patients perform more poorly on neurocognitive tests than would be predicted by their parents' level of education. In addition, comparisons of monozygotic twins discordant for a formal diagnosis of schizophrenia suggest that almost all affected twins perform worse than their unaffected twin on neurocognitive tests, although the unaffected twins are performing more poorly than expectations on the basis of previous educational attainment. Therefore, it is likely that almost all patients with schizophrenia are functioning below the level that would be expected in the absence of the illness.

Experts in neurocognition and schizophrenia have come to a clear consensus that neurocognitive impairment is a core feature of the illness. The profile of deficits is broad, severe, and is likely present in most, if not all, patients. Neurocognitive impairment has clear clinical relevance, as neurocognitive impairment interferes with the everyday lives of patients in many important ways, from limiting social relationships to reducing the likelihood of employment. Furthermore, cognitive impairments in schizophrenia are similar across different countries, educational systems, and racial and ethnic groups, suggesting minimal influence of environmental and social variables.

NATURAL HISTORY OF NEUROCOGNITIVE IMPAIRMENT

Although the development of neurocognitive deficits in patients with schizophrenia follows a predictable pattern over time, it of course can vary greatly in any individual. In most patients, detectable deficits may be present in childhood, followed by a decline in neurocognitive function that occurs sometime prior to the first episode. However, many individuals with schizophrenia demonstrate no obvious neurocognitive deficits prior to the onset of psychosis. In most patients, the severity of neurocognitive impairments becomes greater around the time that psychosis develops. The long-term stability of neurocognitive impairment over time in the majority of patients is clear. It is also clear that progress of deficits over the illness is generally absent, even in very poor outcome cases, until later life and then is limited to a subset of cases (Fig. 12.10–3).

The Genetics of Neurocognition in Schizophrenia

Neurocognitive impairment in patients with schizophrenia has a clear heritable component. First-degree relatives of individuals with schizophrenia are impaired on a variety of neurocognitive measures, with effect sizes ranging from small to medium. The relationship between the heritability of schizophrenia and related neurocognitive deficits has been estimated, with overall intelligence quotient (IQ) and working memory demonstrating phenotypic correlations with the presence of schizophrenia that appear slightly higher than other aspects of neurocognition. The heritability of neurocognitive impairments in schizophrenia families appears to be moderate (h^2 = 0.25 to 0.40) and larger than their environmental correlations. A number of aspects of neurocognition have been tentatively linked to specific genes or single nucleotide polymorphisms. Some of these findings have been replicated and most genes associated with cognitive deficits to date implicate the functioning of the glutamatergic system. Searches for the genomics of cognition in schizophrenia may lead to clues about the origin of disability, the etiology of schizophrenia, and similarities and differences between schizophrenia

and other severe mental illnesses such as bipolar disorder.

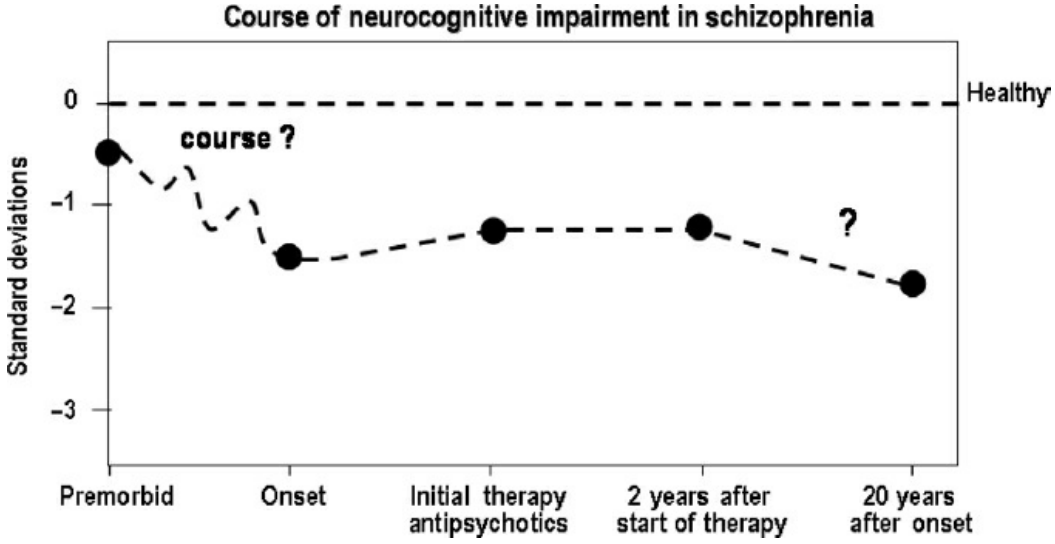

FIGURE 12.10–3. Course of neurocognitive impairment in schizophrenia. The y axis is the number of standard deviations below the mean of healthy controls throughout the course of the illness. The course of cognitive decline between the premorbid phase and onset of illness is not well known. It is controversial whether patients with chronic schizophrenia have cognitive decline over time.

Deficits in Children at Risk for Schizophrenia and Follow-Back Studies

Individuals who are genetically vulnerable to schizophrenia have notable neurocognitive impairments, while individuals who are examined with neurocognitive assessments before they develop schizophrenia are found to have impairments in a variety of areas. High-risk studies of children with one or two biological parents with schizophrenia have suggested that attention deficits can predict which children will develop schizophrenia in the future.

Several studies have utilized the "follow-back" method of identifying adult patients with schizophrenia and then linking these patients with their records of neurocognitive assessments performed while they were children or adolescents. In the UK National Survey of Health and Development Study, data from 5,362 people indicated that children who went on to develop schizophrenia as adults differed significantly from the general population in a variety of neurocognitive and behavioral domains. Low educational test scores in verbal, nonverbal, and mathematics at all ages assessed were significant risk factors. Similar findings were generated from a population-based study that investigated the risk of schizophrenia in a sample of 50,000 18-year-old males conscripted into the Swedish army between 1969 and 1970. In the United States, scores from grades 4, 8, and 11 on the Iowa Tests for 70 children who later developed schizophrenia show no significant differences from controls at grades 4 and 8. However, for children who developed schizophrenia, test scores dropped significantly

between grades 8 and 11, corresponding to the onset of puberty. Yet, limitations of this study, including small sample size and a nonrandom sample, restrict the generalizability of these results.

Several studies have utilized a link between the Israeli Draft Board Registry and the National Psychiatric Hospitalization Case Registry. Israeli law requires that all adolescents between the ages of 16 and 17 years undergo preinduction assessment to determine their intellectual, medical, and psychiatric eligibility for military service. This assessment is compulsory and is administered to the entire, unselected population of Israeli adolescents. It includes individuals who will be eligible for military service, as well as those who will be excused from service on the basis of medical, psychiatric, or social reasons.

The results suggest that neurocognitive functions are significantly impaired in adolescents who are later hospitalized for schizophrenia. These deficits precede the onset of psychosis in young people destined to develop schizophrenia. Along with social isolation and organizational ability, neurocognitive deficits are significant predictors of which young people will eventually develop a psychotic disorder. One study from this series examined the neurocognitive performance of 44 patients with a first episode of schizophrenia who had previously undergone neurocognitive assessment as part of their registration with the Israeli Draft Board. The stability of the deficits in these patients suggests that most of the neurocognitive impairment seen occurs prior to the first psychotic episode. These follow-back studies consistently demonstrate that neurocognitive impairment precedes the onset of illness in schizophrenia and schizoaffective illness. However, the impairments are small on average and to date have not proven useful for identifying cases at higher risk for developing schizophrenia.

Prodrome

Neurocognitive deficits are also found in individuals who are identified as being at “ultrahigh” risk for schizophrenia by virtue of their family history of schizophrenia in a first-degree relative or the manifestation of mild signs and symptoms consistent with the prodromal symptoms of schizophrenia. Preliminary data also suggest that olfactory identification and verbal memory deficits may help predict which individuals at ultrahigh risk will develop schizophrenia. These findings are particularly interesting given that neurogenesis continues later into adulthood in the brain regions that contribute to these functions. However, at present, the ability of these neurocognitive deficits to predict which young individuals will eventually develop schizophrenia is weak. Since schizophrenia is a rare illness (about 1 percent of the population) and the percentage of individuals at ultrahigh risk for schizophrenia who eventually develop the illness ranges from 10 to 40 percent, a predictive factor must be highly specific to be useful in

clinical practice; otherwise, an unacceptably large number of young people will be falsely identified as vulnerable to schizophrenia. The results from large-scale studies have shown that in cases with identifiable prodromal symptoms, cognitive impairments are present at the time of the detection of prodrome, and those cases with the greatest impairments are at the highest risk to develop psychosis.

First Episode of Psychosis Studies

Once psychosis develops, neurocognitive deficits are severe. Patients with a first episode of schizophrenia who have never taken antipsychotic medication already exhibit neurocognitive impairment of a magnitude similar to cases with an established illness. Motor functions and language functions may be more mildly impaired prior to treatment with antipsychotic medication. Initial treatment with antipsychotic medications does little to benefit these neurocognitive deficits.

Neurocognitive Changes in Patients Recovering from Acute Exacerbation

Several studies have investigated the longitudinal course of neurocognitive deficits in schizophrenia by repeatedly assessing neurocognition in samples that combine first-episode patients and those who have had an acute exacerbation of their symptoms. Patients recovering from an acute exacerbation of illness do not appear to demonstrate improvements in the severity of their neurocognitive impairment, despite clear improvements in symptoms with treatment.

Longitudinal Change in Neurocognitive Function

There are many age-related changes in neurocognitive functioning in the healthy population, especially in motor speed and memory functions. The magnitude of change in long-term memory and ability to access and utilize previously learned material is more modest. Reading and vocabulary are tests of “old learning” or “crystallized intelligence.” Baseline functioning can thus be estimated from performance on these tests that show minimal aging-related changes. The prospect of detecting neurocognitive decline with aging is diminished because of floor effects on numerous tests. The types of neurocognitive functions that change with normal aging, such as new learning and recall, attention, and processing speed, are related to the deficits most commonly observed in patients with schizophrenia early in life. Thus, aging-related changes may affect those functions that are already severely impaired in adults with schizophrenia.

The notion of neurodegeneration in schizophrenia has been controversial. Evidence has been presented that suggests that schizophrenia is a neurodegenerative process, and some have concluded that schizophrenia is progressive. However, with a few notable exceptions,

cross-sectional studies generally have not found evidence of increased neurocognitive impairment in association with duration of illness, and older nonelderly patients do not manifest greater impairment than younger patients. Definitive answers to questions regarding possible progression of neurocognitive impairment in schizophrenia must come from longitudinal studies.

Recent studies have shown a complex relationship between relapse, retreatment with antipsychotic medications, changes in cortical structure, and changes in cognition. Several studies have shown that there is loss of white and gray matter and ventricular enlargement seen following relapses due to medication discontinuation after initial successful treatment. One study has shown that these progressive changes in brain volume are correlated with progression of impairments in executive functioning. The mechanism through which brain volume is reduced following relapse is unclear and why only some studies show cognitive decline in early course patients is unknown, but attention to this topic may help understand both the course of cognitive functioning in relapsing patients and the initial origin of these cognitive deficits.

Neurocognitive Functioning in Elderly Patients

Neurocognitive impairment in schizophrenia may worsen over time in at least a subgroup of elderly patients with schizophrenia. In older schizophrenic patients with a lifetime of poor functional outcome, prominent neurocognitive impairments resembling dementia have been reported.

Based upon cross-sectional studies, elderly patients with schizophrenia appear to show some decline in neurocognitive function toward the end of life. However, this decline may be restricted to those patients who had an early onset of illness, a long period of psychosis prior to antipsychotic treatment, or a lifetime of poor functioning. Some of the inconsistency of these results may derive from the subject selection processes in these studies. The patients who are available for these studies tend to be residents at tertiary care facilities that house only the patients with the greatest deterioration in functional ability. Finally, neurocognitive decline in some elderly patients with schizophrenia has been found to be associated with tardive dyskinesia and neurological dysfunction and could potentially be an outcome of decades of antipsychotic treatments and unenriched environments.

Summary

Several studies have demonstrated that neurocognitive impairment is present in a milder form prior to the onset of psychosis in young people destined to develop schizophrenia. Neurocognitive impairment is severe in

patients who have experienced their first psychotic episode, even before antipsychotic treatment is initiated. Although some early-phase patients may demonstrate a slight improvement in neurocognitive impairment with treatment, many patients' neurocognitive abilities do not improve at all. Following this early phase, neurocognitive impairment appears to be remarkably consistent compared to functioning early in life, even in the presence of positive and negative symptom change, although few longitudinal studies have had follow-up periods of longer than 5 years. The course of neurocognitive impairment in elderly patients is uncertain, but some studies suggest that patients with the most chronic courses of illness may manifest a further, albeit gradual, neurocognitive decline in the latest years of life.

RELATIONSHIP OF NEUROCOGNITIVE IMPAIRMENT TO SCHIZOPHRENIA SYMPTOMS

This section addresses the relationship between neurocognitive impairment and symptoms. If neurocognitive deficits were the result of the symptoms of the illness, then the deficits would disappear simultaneously to symptom reduction. However, this does not occur in patients with schizophrenia. In contrast to some patients with psychotic bipolar illness, whose performance on neurocognitive tests may improve somewhat when their psychosis remits, the performance of the patients with schizophrenia does not change when psychotic symptoms are minimized.

Cross-Sectional Studies of the Relation to Symptoms of Schizophrenia

Neurocognitive deficits are largely separate from both the positive and negative symptoms of the disorder. Cross-sectional relationships between neurocognition and positive symptoms, particularly for hallucinations and delusions, are usually quite weak. Generally, the variance shared by these variables is considerably less than 10 percent. Relationships tend to be slightly stronger between neurocognition and negative or "deficit" symptoms.

Positive Symptoms

It has been repeatedly reported that neurocognitive ability is not strongly correlated with severity of psychotic symptoms in patients with schizophrenia. During periods of acute exacerbation, the severity of positive symptoms has been found to be significantly correlated with *better* performance on select domains of neurocognitive function. Although some exceptions exist, such as isolated reports of significant correlations of positive symptoms with working memory, source monitoring, and auditory distractibility, the overall trend is for general neurocognitive impairment not to be correlated with positive symptoms. The correlations between

positive symptoms and neurocognitive function in five different domains and an overall composite score collected from 1,331 patients with schizophrenia during the baseline period of the NIMH CATIE trial were near zero, with *P* values greater than 0.05 despite the tremendous statistical power in this study. Given that more than 91 percent of patients were able to complete neurocognitive testing in this all-comer trial, it is clear that the absence of a relationship between neurocognition and positive symptoms is generalizable. This low correlation is also found in various patient samples, including first-episode, chronic, and elderly samples, supporting the idea that positive symptoms are not the sole cause of the neurocognitive impairment found in any schizophrenia patient population.

Various factors may mediate the absence of a relationship between neurocognitive deficits and positive symptoms. The low validity and temporal stability of positive symptom assessment may be a crucial factor. The subjective report of a psychotic patient may not reflect the true level of the patient's psychosis, and these symptoms change considerably over time compared to neurocognitive deficits, which are temporally stable. Another factor is that the exclusion of patients who are too psychotic to be tested may weaken any potential correlation between the severity of psychosis and neurocognitive impairment. However, the 91 percent success rate in CATIE suggests that this limitation may be overstated. Finally, those patients who have more intact neurocognitive abilities may be better able to recall and express their internal state including detailed delusions and hallucinations. These patients would thus receive higher scores on positive symptom rating scales, resulting in a reduction in any potential correlation between positive symptoms and neurocognitive impairment. There are clear limitations to the assessment of positive symptoms, and the absence of cross-sectional correlations between positive symptoms severity and neurocognitive impairment may be more complex than can be determined in typical research designs.

Negative Symptoms

Neurocognitive dysfunction is correlated with various types of negative symptoms. The greater variance shared between neurocognition and negative symptoms compared with positive symptoms may result from measurement overlap. For instance, tests of verbal fluency and the negative symptom of "poverty of speech" both measure the level at which a patient generates speech. A patient who generates speech at a slow rate will do so during a test of verbal fluency as well as in an interview during which he or she is being rated for poverty of speech.

Empirical studies suggest that motor functions are strongly correlated with negative symptoms, job success, and outcome. Deficient motor skills are represented in both the negative symptom and the neurocognitive

dysfunction domain, since symptoms such as blunted affect and motor retardation are observational measures of motor functioning. Thus, impaired motor skills lie at the core of negative symptoms in schizophrenia in many ways.

The association between reduced motivation and poor performance of patients with schizophrenia on neurocognitive tests is complex. Although monetary reinforcement has been shown to improve performance on effortful neurocognitive tests such as the WCST in some studies, others have not. Increases in pupil size are associated with increased neurocognitive processing demands. Therefore, pupillary response can measure engagement in a task. If the neurocognitive deficits of schizophrenic patients were due to lack of interest or motivation, their pupillary response would be low throughout the period of neurocognitive assessment. Only during high-processing conditions do patients have abnormal pupillary responses, suggesting that they put forth a normal amount of effort when being given neurocognitive tests of modest difficulty, yet their decreased processing capacity leads them to be unable to engage in difficult tasks. Other studies in this have similarly found that patients do not increase their effort as task difficulty increases, even in the face of financial rewards that differentially reward more effort. These findings do not suggest that performance on neurocognitive tests is affected in general by poor motivation but that motivation does need to be considered as a factor.

On the contrary, neurocognitive deficits may cause reduced motivation or the appearance of reduced motivation. First, patients who are cognitively capable to engage in a task or a social interaction may *appear* not to be motivated to succeed in that endeavor. Furthermore, individuals with neurocognitive deficits are often less likely to be motivated to have goals and pursue them. Patients with neurocognitive impairment are likely to be met with failure if they attempt to pursue employment, social, and even recreational avenues that require suitable neurocognitive skill. Repeated failures are likely to cause discouragement and reduced motivation in anyone, including people with schizophrenia.

In sum, a cross-sectional correlation of negative symptoms and neurocognitive impairment has been consistently reported, but the magnitude of this relation is modest. The directionality of the relation between these aspects of schizophrenia is complex and not well understood.

Formal Thought Disorder

Deficits in semantic memory may lie at the heart of the neurocognition–thought disorder relation. This argument is supported by empirical data suggesting that the difference between semantic fluency (the number of names of animals produced in 60 seconds) and phonological fluency (the

number of words beginning with "F" produced in 60 seconds), an indication of the severity of the impairment of the "semantic network" in schizophrenic patients, predicted the severity of their formal thought disorder. Thus, a patient's inability to have verbal information available (referred to as *semantic priming*) may be the most important neurocognitive factor in formal thought disorder. Other hypotheses regarding the cognitive origins of formal thought disorder include deficits in working memory (e.g., being confused about what has already been said in a conversation) and attentional focus.

Affective Symptoms

Since many patients with schizophrenia also report depressed mood if not a full depressive disorder, and since depression is associated with some neurocognitive impairment, the role of depressed mood in neurocognitive impairment is important. Very few studies have examined this relationship directly. Furthermore, the distinction between depressive symptoms and negative symptoms is sometimes difficult for raters and clinicians to make, although they appear to be separate dimensions of schizophrenia based on factor analyses. Higher depression scores are significantly correlated with poorer verbal memory, which remains even when psychomotor retardation and processing speed performance are controlled statistically. Thus, depression may influence the association between negative symptoms and neurocognitive impairments, and it may also have a direct deleterious effect on some domains of neurocognitive impairment.

The cross-sectional correlations between neurocognitive impairment and symptoms are weaker than might be expected and vary depending on symptom domain. The consistency of this finding strongly supports the idea that neurocognitive impairment is not caused by psychosis. The episodic nature of symptoms in schizophrenia in the context of the high temporal stability of neurocognitive impairment contributes to their disassociation.

Longitudinal Course of Neurocognition and Symptoms

The relative independence of neurocognitive impairment and symptoms in cross-sectional studies appears to be supported by longitudinal studies. When patients with schizophrenia are successfully treated with antipsychotic medications, the severity of their symptoms is substantially reduced. However, if neurocognitive function improves with first-generation antipsychotic treatment, the effects are small.

The stability of neurocognitive impairment occurs in the context of frequent variability of the symptoms of schizophrenia, especially the positive symptoms. When patients are tested during a psychotic exacerbation and then again when symptoms had improved considerably, 1

to 2 years later, neurocognitive function is very stable over time. Over a 1-year interval, neurocognitive deficits and negative symptoms are significantly correlated at each of the assessments, but there is little predictive relationship between these variables over time. If negative symptoms were somehow causing poor performance on neurocognitive tests (or vice versa), then a longitudinal relationship between these variables would be expected. Second, when the differential relationship between neurocognitive and negative symptoms and adaptive functioning deficits has been examined, neurocognitive impairments have been found to be more strongly correlated with functional deficits, particularly in everyday activities and vocational outcomes, than negative symptoms. Third, in those patients whose functional status worsens over time, changes in neurocognitive impairment predict the level of change, while changes in negative symptoms do not. These data provide yet another line of support for the relative independence of these domains and the importance of neurocognitive impairment in the mediation of functional difficulties.

Another way to investigate the longitudinal relationship between neurocognitive impairment and symptoms is to calculate the correlation between change in neurocognitive impairment and symptom change. This approach addresses the issue of *pseudospecificity*, that is, changes in neurocognition caused by other clinical changes. Several studies have performed such analyses and though results are mixed, the findings have generally suggested that symptom change does not contribute substantially to neurocognitive change.

The literature consistently demonstrates that neurocognitive deficits are not caused by the symptoms of schizophrenia. Most impressive are the negligible cross-sectional correlations between various domains of neurocognitive function and the positive symptoms of the illness. Although other aspects of schizophrenia, such as negative symptoms, appear to correlate with neurocognitive impairment, there is as much or more evidence supporting the idea that neurocognitive impairment causes negative symptoms, just as there is evidence that negative symptoms cause neurocognitive impairment. Although there is limited evidence of a relationship between changes in neurocognitive impairment and changes in symptoms over time, there is certainly no clear evidence that changes in neurocognitive impairment are caused by symptom changes.

CLINICAL IMPORTANCE OF NEUROCOGNITIVE IMPAIRMENT

Cross-Sectional Correlations with Function

Although the literature on neurocognitive impairment suggests that it is not strongly correlated with symptoms, it has been consistently proven to be related to a variety of other important aspects of the illness, such as social functioning, functional impairments, unemployment, quality of life,

relapse prevention, medical status, and economic cost. Neurocognitive impairment also has considerable power to predict functional status years later.

Relationship of Neurocognitive Impairment to Functioning

Functional outcome in schizophrenia is difficult to define and measure. The three types of functional outcome that most studies of neurocognitive deficits have examined are community (social and occupational) outcome, the ability to solve simulations of interpersonal interactions, and success in psychosocial rehabilitation programs. Strong support for associations between key areas of neurocognition and functional outcome are offered by several reviews of the literature.

The connection between actual community functioning and neurocognition has been weaker than the strong link between laboratory measures of community functioning or skills acquisition and neurocognition. In a 1-year follow-up study of patients with first-episode psychosis, the size of the correlations between the number and quality of actual social relations and various neurocognitive measures were found to be in the 0.25 to 0.35 range.

All of the key neurocognitive domains reviewed previously have demonstrated significant relationships to functional outcome, with effect sizes in the medium range. Neurocognitive impairments are also correlated with deficits in the performance of specific skills critical for independent living.

Neurocognitive Impairment and Unemployment

Several studies of neurocognitive performance and vocational functioning have been conducted in rehabilitation settings. Lower ratings of work behavior and performance are related to poorer baseline scores on verbal memory tests and the WCST. In addition, improvement in patient work performance in a 6-month work rehabilitation program was predicted by better baseline performance on various neurocognitive tests.

Studies examining correlates of competitive employment have consistently emphasized the role of neurocognition. Patients enrolled in school full time or holding competitive employment show superior performance across measures of working memory, sustained attention, problem solving, and episodic memory when compared to unemployed patients, with the scores of part-time workers falling between the other two groups. In addition, vocational functioning is significantly associated with performance on speed of processing tasks such as Trailmaking A and B.

The determinants of employment in patients with schizophrenia are complex and involve a variety of socioeconomic variables that are outside a clinician's control, such as the availability of disability payments. However,

beyond these societal factors, neurocognitive performance plays a more important role than clinical symptoms in the ability of patients with schizophrenia to work and in the ability to benefit from work support programs.

Neurocognitive Impairment and Quality of Life

Quality of life is often defined by the quality of a person's social, occupational, and interpersonal aspects of life, all of which are related to neurocognitive functioning. Reductions in quality of life are more strongly associated with neurocognitive deficits than other symptomatic features of the illness. Specifically, the relationship between subjective experience and social functioning has been shown to be mediated by executive functioning. Patients have less realistic impressions of their social functioning if they also have more severe executive deficits. One study of patients with schizophrenia that examined the relationship between coping abilities and neurocognitive dysfunction suggested that more severe executive and memory deficits are related to decreased use of coping mechanisms. Neurocognitive deficits have also been found to predict long-term quality of life in a few studies. However, persistent symptoms also have an impact on quality of life, and clinicians need to balance the impact of treatment on these two aspects of the disorder separately.

Neurocognitive Impairment and Relapse Prevention

Neurocognitive functions have been shown to be the strongest predictors of patients' ability to manage medications. Neurocognitive deficits contribute to patterns of medication mismanagement that are associated with poor adherence and risk of relapse. In one study, memory impairment was the strongest predictor of partial compliance. Patients performing poorly in medication management tests have also demonstrated poor global scores on a cognitive assessment battery. Decreased medication compliance in schizophrenic patients has been shown to be related to poor performance on tests of attention and visual memory. Since antipsychotic medication has little impact on neurocognitive deficits, there can be a fair amount of confidence that the direction of causation is running from neurocognitive impairment to decreased medication compliance.

Neurocognitive Impairment and Medical Comorbidity

Neurocognitive impairment is also related to medical comorbidities in schizophrenia. Neurocognitive deficits related to organization (such as reasoning and problem solving) directly affect patients' ability to seek treatment of medical problems. In elderly patients with schizophrenia, neurocognitive and functional impairments predict the later incidence of new-onset medical problems, while medical problems do not predict the

subsequent worsening of neurocognitive and self-care deficits. Inability of patients with schizophrenia to reduce damaging habits such as smoking has been correlated with deficits in memory and attention. Thus, neurocognitive impairments directly affect new-onset medical problems in older patients.

The Costs of Neurocognitive Impairment

Neurocognitive impairment is also a major factor in the direct and indirect costs associated with schizophrenia. Factors leading to the increased cost include loss of ability for self-care, level of inpatient and outpatient care needed, and loss of productivity (for both patient and caretaker). Although neurocognitive impairment by itself is rarely a reason for acute hospitalization of a patient with schizophrenia, it may contribute to the overall length of hospitalization. It may also lead to early admission to nursing homes or long-term care facilities in the case of elderly patients and other patients who are consistently unable to live in the community.

Neurocognitive impairment is the single greatest contributor to the functional disability of schizophrenia. Patients whose symptoms have remitted continue to experience profound neurocognitive deficits that interrupt their ability to function independently, benefit from rehabilitation programs, find and maintain employment, and experience healthy levels of quality of life.

NEUROCOGNITION AND FUNCTIONAL CAPACITY

In the last years, research has focused on the relationships between neurocognitive performance and the ability to perform everyday functional skills. These skills, including financial skills, shopping, travel and transit, and self-maintenance, have been shown to relate to real-world deficits in these areas. Functional capacity assessment instruments, such as the University of California–San Diego Performance-based Skills Assessment, require patients to complete a series of challenging and relevant everyday tasks such as making a doctor's appointment, paying a bill, or planning for a leisure activity. One possible mechanism of the impact of neurocognitive deficits on everyday disability is by inducing impairments in the ability to perform the skilled acts that are required for everyday functioning. Research studies have shown very high and consistent correlations between neurocognitive performance and functional capacity measures, and these same studies have shown that functional capacity measure may be more closely related to real-world functional outcomes than neurocognitive deficits are.

TREATMENT OF NEUROCOGNITIVE IMPAIRMENT

As described in the preceding sections, neurocognitive impairment in

patients with schizophrenia has a devastating impact on various aspects of their illness and their lives. The search for treatments that can improve neurocognition has accelerated in recent years. Although antipsychotic medications have a pronounced beneficial impact on psychotic symptoms, there is currently little pharmacological relief that can be provided for the neurocognitive impairment found in patients with schizophrenia.

Impact of Antipsychotic Treatment on Neurocognition

For decades, patients with schizophrenia were treated with excessively large doses of conventional, first-generation, "typical" antipsychotics, which may have had direct adverse neurocognitive effects. Larger doses of conventional antipsychotics cause lethargy, somnolence, and extrapyramidal symptoms, all of which impair neurocognition. In addition, the anticholinergic medications that are used to control side effects cause additional neurocognitive impairment. Furthermore, lowering dosages of conventional antipsychotic medications has some modest neurocognitive benefit, supporting the idea that the negative impact of first-generation antipsychotics may be dose dependent. As lower doses did not make neurocognitive impairment less severe than that seen in never-medicated patients, the impact of these older antipsychotic drugs at lower doses is less clear.

Although many studies and meta-analyses have suggested that second-generation, "atypical" antipsychotic treatments provide greater neurocognitive benefit to patients with schizophrenia than first-generation antipsychotics, many of these studies have had substantial weaknesses, such as small sample sizes, short duration of treatment, no comparator or a comparator of relatively high doses of first-generation antipsychotic treatment, and inattention to important clinical factors such as the relationship of neurocognitive improvement with symptom change, anticholinergic treatment, and change in extrapyramidal symptoms. The impact of these various design weaknesses on study results is controversial.

The NIMH CATIE schizophrenia trial provided the opportunity to examine the neurocognitive effects of antipsychotics without these methodological weaknesses. This study assessed the neurocognitive effects of four second-generation medications, olanzapine (Zyprexa), quetiapine (Seroquel), risperidone (Risperdal), and ziprasidone (Geodon), and a representative first-generation medication, perphenazine (Trilafon), using randomized, controlled, double-blind methods in 817 patients. The neurocognitive measures in this trial improved in each treatment group after 2 months of treatment, and none of the groups were statistically distinguishable from the others, including patients receiving the older medication, perphenazine. However, the amount of improvement in these patients was interpreted to be consistent with practice effects and expectation biases, suggesting that the effect of antipsychotic medications

on these patients was negligible.

Less is known about the effects of antipsychotic treatments on patients with first-episode psychosis. Two industry-sponsored studies have suggested that olanzapine and risperidone improve neurocognition more than even low doses of haloperidol (Haldol) in first-episode patients, but the lack of improvement in patients using haloperidol appears to be associated with other clinical changes and side effects in these patients. The amount of neurocognitive improvement in these studies is also small and may be consistent with practice and expectation effects.

The Search for Pharmacological Treatments for Neurocognition

Although neurocognition is increasingly viewed as a central feature of the disorder, the inability of antipsychotic treatment to improve neurocognition has led to broad efforts to identify pharmacological treatments that may provide neurocognitive benefit to patients with schizophrenia. Currently, many large- and medium-sized pharmaceutical companies have compounds in their development pipelines that are being considered as possible treatments of the neurocognitive deficits of schizophrenia. There have been some failures and false starts in this area, and at present there are no U.S. Food and Drug Administration (FDA)–approved treatments of neurocognition in schizophrenia.

Cognitive Remediation

A variety of cognitive remediation programs have been developed to address neurocognitive deficits in schizophrenia, with recent refinements of these programs going beyond older drill and practice exercises and moving toward teaching strategies to improve neurocognitive functioning, compensatory strategies to reduce the effects of persistent neurocognitive impairments, and group exercises aimed at deployment of the trained skills. In contrast to earlier strategies that converted neurocognitive rehabilitation programs designed for individuals with brain injury following healthy premorbid histories and were found to be too rote and tedious for patients with schizophrenia, these newer strategies are engaging and perceived as entertaining by patients. The potential for the success of these programs to improve neurocognition and functional outcomes in patients with schizophrenia has been supported by a number of recent studies.

A meta-analysis of 26 randomized clinical trials involving a total of 1,151 patients concluded that neurocognitive remediation produces moderate improvements in neurocognitive performance and, when combined with psychiatric rehabilitation, also improves functional outcomes. Unfortunately, the largest of the trials had only 67 patients, and each study was conducted at a single or a small number of settings with highly trained

research personnel. Since that time, there have been several larger studies that supported the benefits of cognitive remediation, particularly in concert with psychosocial interventions, for improving both cognitive test performance and improving everyday outcomes. Furthermore, as in the pharmaceutical industry, the tendency of the developers of neurocognitive remediation programs to evaluate the efficacy of their own programs without bias has been questioned. Thus, the generalizability of these findings is limited and the real-world applicability of neurocognitive rehabilitation is unknown.

One relevant controversy in this research area is that pharmacological interventions to improve neurocognition may be minimally effective only if patients with schizophrenia are not exposed to rich neurocognitive environments or receive targeted psychosocial interventions. A new compound that has the potential to have a beneficial neurocognitive effect by altering the neurochemistry of neurocognition in the human brain may be prevented from demonstrating benefit if patients do not have sufficient daily neurocognitive stimulation, which is frequently the case in many patients with schizophrenia. It is likely that one of the mechanisms of neurocognitive improvement in schizophrenia will be an enhanced neuroplasticity, which is often experience dependent. Thus, neurocognitive enhancement trials in schizophrenia may require neurocognitive remediation in all subjects to test the effectiveness of new compounds adequately.

Summary

There are currently no clearly effective treatments of neurocognitive impairment in schizophrenia. Larger doses of first-generation antipsychotics, perhaps especially the higher potency medications such as haloperidol, may have a deleterious effect. Despite initial hopes to the contrary, it appears as though even the second-generation antipsychotic medications do not provide more than minimal benefit for neurocognitive impairment. A broad effort is currently under way to find new medications that may serve as "cotreatments" to antipsychotics to benefit neurocognition specifically. A few limited studies have suggested that neurocognitive remediation programs may have modest beneficial effects on neurocognition in patients with schizophrenia, and large-scale efforts are needed.

Table 12.10–1.
Advantages and Disadvantages of Selected Neurocognitive Batteries

Battery	Pros	Cons
WAIS-III and WMS-III	Long history of use allows comparison with previous studies, norms available	Lengthy battery, not designed

		specifically for research in schizophrenia
MCCB	Designed by panel of experts for research in schizophrenia, allows for domain scores with minimal testing norms available	Many domain scores based on performance on one test
SCoRS	High face validity, easily administered by clinicians, minimal time requirements, demonstrated correlation with other measures of cognitive and functional outcomes	The right informant/observer is required to obtain valid information
UPSA	High face validity, proxy test of real-world functioning, minimal time requirements	Domain-level analysis not possible
RBANS	Minimal time requirements, small practice effects, performance correlated with that on WAIS-III and WMS-III	Lacks measures of important cognitive domains in schizophrenia, significant ceiling effects
BACS	Minimal time requirements, designed for use in schizophrenia research, high correlation with composite scores from more extensive batteries, minimal practice effects, available in nine languages, norms available	Domain-level analysis is minimal
BCA	Extreme brevity of administration, high correlation with extensive cognitive battery and functional outcome measures	Domain-level analysis minimal; minimal cover of many areas
Computerized batteries	Automatic administration reduces rater error	Validity of many tests has not been examined; norms are often questionable
Psychophysiological tasks	Much of underlying neurobiological circuitry is known, afford quick assessment of schizophrenia-related endophenotypes	Relationship between performance of these tasks and functional outcome unknown

WAIS, Wechsler Adult Intelligence Scale; WMS, Wechsler Memory Scale; MCCB, MATRICS (Measurement and Treatment Research to Improve Cognition in Schizophrenia) Consensus Cognitive Battery; SCoRS, Scale for Cognition in Schizophrenia; UPSA, University of California Performance-Based Skills Assessment; RBANS, Repeatable Battery for Neuropsychological Status; BACS, Brief Assessment of Cognition in Schizophrenia; BCA, Brief Cognitive Assessment. (From Kraus MS, Keefe RS. Cognition as an outcome measure in schizophrenia. *Br J Psychiatry Suppl*. 2007;50:s46–s51.)

METHODS OF ASSESSMENT

A large number of tests have been developed to assess neurocognition, and various batteries comprising these tasks have been employed in clinical

practice and research with patients with schizophrenia (Table 12.10–1). The most useful method of assessment will depend on the clinical or research questions being addressed, the study settings, and the available resources. A battery comprising many different tests that may take several hours to complete has the advantage of enhanced reliability of neurocognitive estimates and will allow the assessment of various domains of neurocognition. However, these long test batteries have the disadvantage of patient, tester, and resource burden. In addition, missing data and attrition rates may increase, which may be higher in most impaired subjects, thus skewing results. If the research or clinical question involves the efficacy of a treatment intervention on neurocognition, a composite score from a small battery may be sufficient and allow for a larger number of patients to be assessed while requiring less staff resources. If, however, the question involves relative strengths and weaknesses of various neurocognitive domains, a thorough battery composed of multiple tasks in each domain may be required. If the question is primarily focused on one neurocognitive domain, a brief general battery in combination with multiple tests of the domain of primary interest is likely to be appropriate.

A recent development in neurocognitive assessment for clinical trials is the availability of computerized test batteries that allow direct data transfer to study databases. These methods offer the hope of minimizing rater error and reducing the costs required for human quality assurance. However, many of these methods have not been fully validated and, therefore, results obtained using them must be evaluated carefully and scientifically. Many of these tests have been made commercially available without nearly as much information on their validity as would be hoped for.

Complex experimental procedures derived from cognitive neuroscience have also been used to characterize detailed aspects of the neurocognitive impairments in schizophrenia. These procedures have the potential to be useful since the underlying neurobiological circuitry that mediates these functions has been uncovered in animal and human studies. However, care must be taken when inferring an etiological basis or treatment strategy for schizophrenia based on the performance of patients on these tasks. The outstanding question remains whether interventions that improve the performance of patients on these tasks would have any effect on functionally relevant cognitive ability or on functional outcomes, as correlations with everyday outcomes have sometimes been found to be low. These cognitive neuroscience tasks have been particularly appealing for use in genetic studies as tools to quickly assess an endophenotype that might reflect a specific genotypic vulnerability to schizophrenia.

Finally, neurocognitive deficits in schizophrenia and the impairments in everyday functioning that they mediate have begun to be measured with methods that may more closely track changes in functional outcomes. Interview-based assessments of neurocognition such as the Scale for

Cognition in Schizophrenia (SCoRS) and the Cognitive Assessment Interview (CAI) have been developed for use in clinical practice and clinical trials. These measures involve addressing a series of questions about the severity of neurocognitive impairment and how these impairments reduce functioning in daily life. Their strength is that they can be completed in a short time period (15 to 30 minutes) and yield a reliable estimate. However, information obtained directly from patients with schizophrenia tends to have reduced reliability and validity, and a caregiver or other informant is usually needed to provide information from an objective perspective. Performance on these tests is highly correlated with levels of functioning in everyday life, and it is believed that improvements on these tasks herald the potential for real-world change.

References

Bilder RM, Goldman RS, Robinson D, et al. Neuropsychology of first-episode schizophrenia: initial characterization and clinical correlates. *Am J Psychiatry*. 2000;157(4):549–559.

*Bowie CR, McGurk SM, Mausbach BT, Patterson TL, Harvey PD. Combined cognitive remediation and functional skills training for schizophrenia: effects on cognition, functional competence, and real-world behavior. *Am J Psychiatry*. 2012;169:710–718.

Bowie CR, Reichenberg A, Patterson TL, Heaton RK, Harvey PD. Determinants of real-world functional performance in schizophrenia subjects: correlations with cognition, functional capacity, and symptoms. *Am J Psychiatry*. 2006;163(3):418–425.

Buchanan RW, Keefe RS, Umbricht D, Green MF, Laughren T, Marder SR. The FDA-NIMH-MATRICS guidelines for clinical trial design of cognitive-enhancing drugs: what do we know 5 years later? *Schizophr Bull*. 2011;37:1209–1217.

Dickinson D, Ramsey ME, Gold JM. Overlooking the obvious: a meta-analytic comparison of digit symbol coding tasks and other cognitive measures in schizophrenia. *Arch Gen Psychiatry*. 2007;64:532–542.

*Greenwood TA, Lazzeroni LC, Murray SS, et al. Analysis of 94 candidate genes and twelve endophenotypes for schizophrenia from the Consortium on the Genetics of Schizophrenia. *Am J Psychiatry*. 2011;168:930–946.

Harvey PD. Pharmacological cognitive enhancement in schizophrenia. *Neuropsychol Rev*. 2009;19:324–335.

Harvey PD, Keefe RSE. Studies of cognitive change in patients with schizophrenia following novel antipsychotic treatment. *Am J Psychiatry*. 2001;158(2):176–184.

Heinrichs RW, Zakzanis KK. Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. *Neuropsychology*. 1998;12:426–445.

Keefe RS, Bilder RM, Davis SM, et al; CATIE Investigators; Neurocognitive Working Group. Neurocognitive effects of antipsychotic medications in patients with chronic schizophrenia in the CATIE trial. *Arch Gen Psychiatry*. 2007;64(6):633–647.

Keefe RSE, Buchanan RW, Marder SR, et al. Clinical trials of potential cognitive-enhancing drugs in schizophrenia: what have we learned so far? *Schizophr Bull*. 2013;39:417–435.

Keshavan MS, Vinogradov S, Rumsey J, Sherrill J, Wagner A. Cognitive training in mental

disorders: update and future directions. *Am J Psychiatry*. 2014;171:510–522.

Mesholam-Gately RI, Guiliano AJ, Goff KP, Faraone SV, Seidman LJ. Neurocognition in first episode schizophrenia: a meta-analytic review. *Neuropsychology*. 2009;23:315–336.

Seidman L, Giuliano AJ, Meyer EC, et al. Neuropsychology of the prodrome to psychosis in the NAPLS consortium: relationship to family history and conversion to psychosis. *Arch Gen Psychiatry*. 2010;67:578–588.

Vinogradov S, Fisher M, Holland C, Shelly W, Wolkowitz O, Mellon SH. Is serum brain-derived neurotrophic factor a biomarker for cognitive enhancement in schizophrenia? *Biol Psychiatry*. 2009;66:549–553.

Vinogradov S, Fisher M, Warm H. The cognitive cost of anticholinergic burden: decreased response to cognitive training in schizophrenia. *Am J Psychiatry*. 2009;166:1055–1062.

*Wykes T, Huddy V, Cellard C. A meta-analysis of cognitive remediation for schizophrenia: methodology and effect sizes. *Am J Psychiatry*. 2011;168:472–485.

▲ 12.11 Phenotypes of Psychosis

BRETT A. CLEMENTZ, PH.D., MATCHERI S. KESHAVAN, M.D., GODFREY D. PEARLSON, M.D., JOHN A. SWEENEY, PH.D., AND CAROL A. TAMMINGA, M.D.

Thaker's (2012) contribution to this textbook offered a wide-ranging overview of schizophrenia's phenotypic manifestations. He provided a brief but erudite historical summary of the schizophrenia construct, highlighted its diagnostically complex and superficially bewildering constellation of clinical features, and touched on epidemiology (both population and genetic), illness course, and the vagueness of determining presence/absence of overt psychotic illness in "boundary" cases. Dr. Thaker surveyed what was known about the extra-clinical (laboratory-based) phenotypic characteristics of schizophrenia, covering disturbances in cognition, psychophysiological manifestations of early and late information processing deviations, abnormalities in neural communication/synchronization over short and long cortical distances, and abnormalities of ocular motor control. Much of this information formed the background for a large-scale study of the psychosis continuum called the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) that had begun just a few years before, and of which Dr. Thaker was a founding Principal Investigator.

DO CLINICAL PHENOMENOLOGICAL PSYCHOSIS DEFINITIONS CAPTURE BIOLOGICAL REALITY?

In the ensuing 4 years a new Diagnostic and Statistical Manual of Mental Disorders (DSM-5) has been published, with the next International Classification of Diseases (ICD-11) not far behind. These manuals continue (or will continue) to use clinical phenomenological characteristics alone to define different psychotic illnesses, including schizophrenia, schizoaffective disorder, and bipolar disorder with psychosis. Operational diagnostic criteria for psychotic disorders, the backbone of these manuals, unquestionably facilitate diagnostic reliability for bipolar disorder with psychosis and schizophrenia, although are less successful for schizoaffective variants. The validity of these diagnostic distinctions, however, remains questionable for multiple reasons. Clinical phenomenology does not uniquely differentiate psychotic disorders. Poor outcome patients (traditionally considered a hallmark of schizophrenia psychosis) often have prominent affective symptoms and bipolar disorder family histories. Considerable overlap exists across psychosis diagnoses on neurocognitive features although schizophrenia cases tend to have more severe cognitive dysfunction and negative symptoms. Family, twin, and

adoption studies indicate the importance of constitutional/genetic factors in risk for both bipolar disorder with psychosis and schizophrenia. Curiously, however, clinically defined psychotic disorders do not "breed true." a fact problematic for the classical Kraepelinian dichotomy theory (a formulation that persists in the modern psychiatric diagnostic consciousness despite Kraepelin's, 1920, deep concerns with its veracity). All antipsychotics have largely overlapping dopaminergic mechanisms of action, are only partially effective, and treat psychosis generally, not specific clinical psychosis diagnoses. Apparently, to quote Dr. Steve Hyman, "DSM [and related classifications] was a poor mirror of nature."

One could reasonably wonder whether DSM-like psychosis diagnoses have justified their pertinence. As theoretical constructs (hypotheses to be tested rather than entities to be reified), they have been only modestly successful. There are certainly statistical tendencies for schizophrenia and bipolar disorder with psychosis to differ in predictable directions on a range of measures. Striking heterogeneity within and overlap between these theoretical constructs on a range of molecular, genetic, brain imaging, perceptual, neurocognitive, symptomatic, treatment response, and outcome variables, however, raise concerns with how accurately DSM psychosis definitions capture neurobiological reality. As a result, just 4 years after Thaker's (2012) contribution, it seems contrary to contemporary understanding to organize a presentation of psychosis phenotypes around current DSM psychosis constructs.

In other branches of medicine, data do not support the assumption that clinical phenomenological disease definitions alone are congruent with biologically based disease entities. The National Academies of Sciences recently proposed revising the classification of disease across all medicine based on molecular/biomarker data rather than basing them solely on clinical phenomenology, a charge not obviously embraced by framers of the current psychiatric diagnostic manuals. In fairness, the DSM framers had few (if any) sturdy biomarker/molecular/neurobiological hooks on which to hang their diagnostic hats.

After decades of research on the neurobiology of psychotic disorders, why is this case? A standard, and somewhat nihilistic, answer is that psychiatric brain disorders are hopelessly complex, heterogeneous by their nature, and beyond the reach of simple biological explanations; they are qualitatively different from diseases in other branches of medicine. Perhaps. An alternative explanation is that continued reliance on DSM-type diagnoses, which are notoriously pleomorphic and only vaguely tied to biological reality, but are nevertheless allowed to surmount neurobiological observations, impedes the scientific development of psychiatry enjoyed by other medical disciplines. The B-SNIP consortium was formed to address this dilemma.

STUDYING NEUROBIOLOGICAL DISTINCTIVENESS OF DSM PSYCHOSIS DIAGNOSES

Discovery of neurobiological/etiological mechanisms should be an important goal of disease research in psychiatry, with such knowledge informing disease definitions and eventually supporting treatment developments. So far attempts to develop a systematic classification of psychotic (or any other psychiatric) syndromes, informed by neurobiological mechanisms, have been unsuccessful. A continuing challenge for modern psychiatry is to discover the (broad sense) distinctive biological substructures of its diseases. This was an organizing goal of the B-SNIP project from its inception. Other consortia, including the Consortium on the Genetics of Schizophrenia (COGS) and the North American Prodrome Longitudinal Study (NAPLS) group have complimentary goals.

B-SNIP working hypotheses borrowed elements from both dimensional and categorical (which are complimentary rather than competitive) approaches to understanding DSM-type psychosis manifestations. Simplistically, a dimensional approach to description/diagnosis of diseases can be understood as a continuum of severity, with this continuum shading from disease state into subsyndromal and normal variants on the construct(s) of interest (e.g., frontal lobe gray matter volume, cortical thickness, neuropsychological measures of cognitive control, oddities of perceptual experience). Categorical approaches can be understood as capturing qualitative, not just continuous quantitative (dimensional), differences between diseases. In their purest form, such diseases are qualitatively distinct from normalcy. It is unlikely, however, that any purely psychiatric brain disease is similar to a categorical entity like Huntington's disease (an autosomal dominant mutation of the Huntingtin gene). Even the characters of Mendel's peas (e.g., yellow versus green color) were quantitative in nature (although "yellow" and "green" can be considered categorical descriptors). The nuances of various approaches to nosology in psychiatry have been discussed with considerable sophistication in other locations.

B-SNIP was prominently influenced by the proposals of Craddock, Owen, and colleagues on the diagnostic reformulation of the psychoses. Initially, the authors had two main organizing hypotheses for understanding psychosis neurobiology and its relationship to clinical manifestations. First, there was reasonable evidence for continua of severity on both neurobiological and symptomatic features of relevance to the larger psychosis category, with schizoaffective variants bordered by psychotic bipolar disorder on the less, and schizophrenia on the more, pathologically extreme ends. Craddock, O'Donovan, and Owen made similar proposals, although schizoaffective variants do not always conform to a simple dimensional scheme. Second, it seemed plausible that there are

etiological factors independent of the psychosis neurobiological continua. These factors, superimposed on the psychosis-relevant continua, would be responsible for cases tending toward particular DSM-type manifestations. The psychosis severity background upon which these etiological "hits" occur could cause the sometimes confusing and seemingly arbitrary nature of clinical manifestations observed across psychosis cases. In 2010 Craddock and Owen presented a comprehensive model with multiple levels of analysis (ranging from genes through intermediate neurobiological and behavioral features to clinical phenotypic psychosis manifestations) that captures the essence of such a proposition.

From the authors' perspective, there were three keys to the success of B-SNIP aims. First, it was crucial to collect psychosis samples (regardless of DSM diagnosis) of sufficient size to capture representative clinical and neurobiological heterogeneity for two reasons: (i) peculiarity of small and moderately sized samples across diverse clinical and research settings could account for inconsistent results across psychosis studies—this was an avoidable obstacle to knowledge acquisition; and (ii) larger sample sizes would support more sophisticated approaches to multivariate statistical analyses.

Second, the psychosis dimension(s) and any independent etiological factors should be quantified at a level removed from clinical symptomatology. The authors focused on biomarkers (traditional so-called endophenotypes) at the perceptual/cognitive level of analysis. These laboratory measures were selected given their promise for identifying neurobiological deviations and understanding genetic risk for psychosis. Starting with a particular class of these measures, at the perceptual/cognitive level of analysis, had a number of useful features: (i) in the late 2000s (2008) there was considerable information available on these measures; indeed, they were the most researched endophenotypes for psychosis; (ii) findings would be easily replicable and portable across diverse clinical and laboratory settings; and (iii) starting at an intermediate level of neurobiological deviation (between the more basic molecular and structural/functional anatomical and more superficial psychosocial functioning and clinical descriptive) offered the opportunity to shorten the causal distance between genes and clinical phenes. Theoretically meaningful associations could be tested in both downward and upward directions in this extensive and complex causal chain, strengthening the theoretical network.

Third, to define biomarkers leading to biosignatures of various psychosis manifestations, it seemed likely multiple, statistically integrated biomarkers, rather than biomarkers considered individually, would be required to capture the essence of psychosis disease factors/constructs. Epigenetics, etiologic heterogeneity, pleiotropy, and variable expressivity influence phenotypic manifestations, making it difficult for more molar

level laboratory measures to be precise indicators of more molecular causative agents. Just as one would never use a single item from the WAIS Information subscale to estimate full scale IQ, it seemed overly optimistic to assume a single measure could adequately assess any brain functioning-related construct (e.g., context updating in working memory, behavioral inhibition; behavioral response speed). Nevertheless, single laboratory measures are often used in experimental psychopathology and psychiatry research to test hypotheses of interest. This was another source of possible error in biomarkers research on the psychoses that could be easily avoided. B-SNIP relied on statistically integrating multiple individual measures to assess constructs of interest (e.g., antisaccade response latency and prosaccade latency both index speed of visual orienting; the auditory ERP N100 response during the paired-stimuli task is assessing a highly similar neural response as the N100 response during the oddball task), thus enhancing signal-to-noise of the authors' biomarker variables and maximizing meaningful variance in estimating important brain functions of relevance to psychosis.

B-SNIP RESULTS USING DSM PSYCHOSIS DIAGNOSES AS THE GOLD STANDARD

The focus of B-SNIP was psychosis, a clinical phenotype with reliable and quantifiable symptom manifestations. Although fuzzy at its border with normal variations, a clinical psychosis diagnosis can be made with high inter-rater reliability. Psychosis also subsumes multiple diagnostic groups (schizophrenia, schizoaffective disorder, bipolar disorder with psychosis) that share clinical features, has neuropharmacological treatment mechanisms that cross diagnostic boundaries, and is clearly a group of brain diseases. The initial B-SNIP goal was to identify neurobiological features via the authors' dense biomarker panel that were (i) shared between and (ii) distinctive to the DSM psychosis diagnosis categories. B-SNIP recruited over 2500 volunteers (psychosis probands, a subset of their first-degree biological relatives, and a community comparison sample) and performed dense biomarker phenotyping across multiple levels of analysis (symptoms, social functioning, cognitive, psychophysiological, brain imaging, and molecular measures). The psychosis probands were clinically stable outpatients who met consensus diagnoses of DSM-IV-TR schizophrenia, schizoaffective disorder, or bipolar disorder with psychosis.

Initially the authors investigated whether symptoms might distinguish the three psychosis groups. Linear combinations of multiple psychosis-related clinical symptoms were created that best differentiated the three DSM psychosis diagnoses. One practical outcome of this statistical enterprise would be to help clinicians know what clinical features are most closely associated with a particular psychosis diagnosis.

The main outcome of this analysis is illustrated in Figure 12.11–1 (upper graph). There were two statistically significant dimensions that

differentiated the diagnostic groups. One dimension the authors labeled "general psychosis" given its primary symptom correlates, which included hallucinatory behavior, delusions, blunted affect, unusual thought content, difficulty in abstract thinking, lack of judgment and insight, suspiciousness/persecution, social withdrawal, and lack of spontaneity. Schizophrenia cases scored highest and bipolar disorder with psychosis cases scored lowest on this symptom dimension. The other dimension the authors labeled "affective instability" given its symptom correlates, which included depression, sadness, somatic concern, lassitude, inner tension, reduced sleep, suicidal thought, concentration difficulty, irritability, reduced appetite, and anxiety. This dimension also had prominent correlates with the general psychosis symptoms of hallucinatory behavior, delusions, and suspiciousness/persecution. The affective instability dimension differentiated the schizoaffective disorder cases from the other two diagnostic groups.

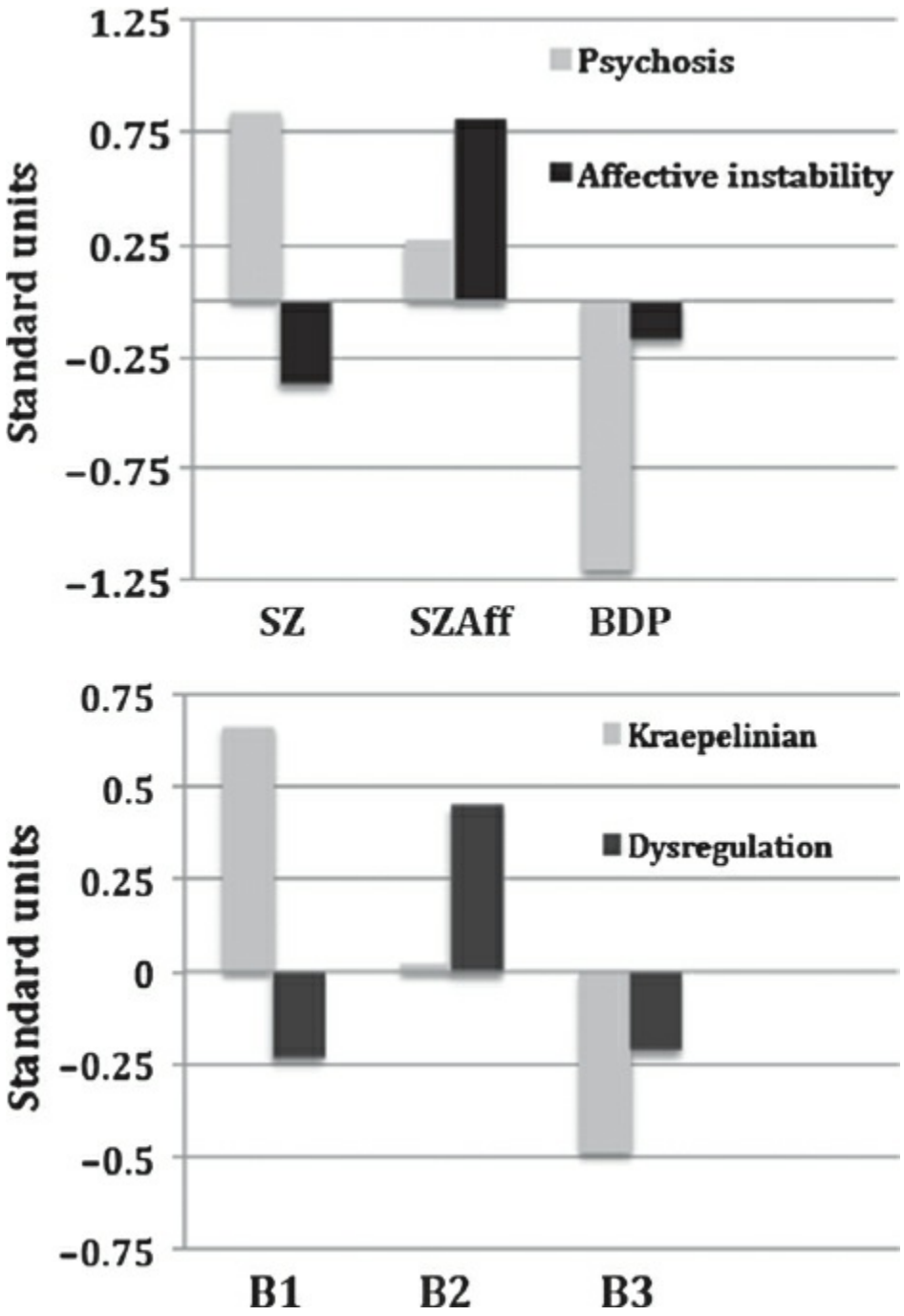

FIGURE 12.11–1. Symptom dimensions of psychosis. The ordinate shows standard score units on psychosis dimensions that best characterized DSM diagnoses (upper plot) and psychosis biotypes (lower plot). For DSM diagnoses, symptomatic differences were captured by "General Psychosis" and "Affective Instability" dimensions. For psychosis biotypes, symptomatic differences were captured by "Kraepelinian" and "Neuro-Cognitive Dysregulation" dimensions. SZ, schizophrenia; SZAff, schizoaffective disorder; BDP, bipolar disorder with psychosis; B1, biotype-1; B2, biotype-2; B3, biotype-3.

The symptom correlates of the DSM psychosis diagnoses were interesting for four main reasons. First, schizophrenia was not characterized by any special set of clinical features but rather by more extreme scores on psychosis symptom ratings. Second, schizoaffective disorder, not bipolar disorder with psychosis, was characterized by

affective instability. Third, affective instability had prominent associations with general psychosis-related symptoms, perhaps indicating multiple pathways to indistinguishable clinical presentations. Fourth, bipolar disorder with psychosis was associated with low scores on the general psychosis dimension. Affective instability played no role in bipolar disorder's differentiation from the other psychosis groups.

The symptomatic differentiations of the DSM psychosis groups were revealing. There was clear indication of a psychosis severity continuum, and some indication that schizoaffective disorder was unique in its symptomatic presentation from schizophrenia and bipolar disorder with psychosis. The next task was to investigate the neurobiological uniqueness of the DSM psychosis diagnoses using measures in the authors' biomarker battery. The main biomarkers used in the first B-SNIP iteration were as follows:

Brief Assessment of Cognition in Schizophrenia (BACS). This battery is a widely used test of global neuropsychological function. It covers multiple cognitive domains (Verbal Memory, Processing Speed, Reasoning and Problem Solving, Working Memory), although a global neuropsychological functioning composite score integrating over these domains yielded the best measure of psychosis-related cognitive deviation in B-SNIP. Dysfunction of cognition is a prominent feature of psychotic disorders, so it was critically important to have measures reasonably assessing these domains in the B-SNIP biomarkers battery.

Pro- and Anti-saccade Tasks. Prosaccade tasks require subjects to look as quickly and accurately as possible to a peripheral visual cue, and assess speed of visual orienting (psychosis cases show variable slowed or speeded response times). Antisaccade tasks require subjects to look as quickly and accurately as possible to the mirror image location (same amplitude, opposite direction) of a peripheral visual cue, and they assess inhibitory control under perceptual conflict because the visual stimulus and required response location are incompatible (psychosis cases have delayed reaction times and increased error rates). Participants performed pro- and anti-saccade tasks under identical conditions, and eye movement recordings were analyzed using established methods. Saccade latency (time from peripheral cue onset to saccade onset) and percentage of error responses (glances to the incorrect spatial location) were recorded for each condition.

Stop Signal Task. Stop signal performance measures efficiency and adequacy of cognitive control when response activation and generation to a single stimulus location are placed in conflict (there are delayed response times and increased errors in psychosis). In a typical stop signal paradigm, trials begin with a central fixation cross after which subjects are shown a "Go" cue to the left or right. On a minority of trials a "Stop" signal is presented at central fixation (with variable delays of, e.g., 50–250 ms after

Go cue onset). Participants are instructed to respond to the Go cue as quickly as possible unless they encounter the Stop signal. Reaction time measures and proportion of Stop signal errors are often quantified, as they were in the B-SNIP project.

Auditory Paired-Stimuli and Oddball Evoked Brain Responses. Evoked brain responses to repetitive auditory stimuli (paired-stimuli task, used to measure what is sometimes called "P50 gating") and predetermined auditory targets randomly interspersed with nontarget (or standard) auditory events (oddball task) are deviant in psychosis. These paradigms assess the neural dynamics of preparation for and recovery from auditory sensory activations, neural responses to stimulus salience, and neural differentiation of relevant from irrelevant auditory stimulus events. For the typical paired stimuli task, subjects passively listen to multiple (e.g., 120–150) binaural auditory click pairs with a short (500 ms) interval between the two clicks and a long (8–10 sec) interval between the click pairs. For a typical auditory oddball task, subjects listen to hundreds (e.g., 500–600) of auditory stimuli occurring every 1 to 1.5 sec. Most of the stimuli are the same (e.g., 1000-Hz tones), and are called "standards," with 15% to 25% of the stimuli, randomly interspersed with the standards, being different (e.g., 1500-Hz tones); those are called "targets." Targets elicit a different brain response from the standards, with the most prominent the so-called P300 (a positive voltage waveform in the event-related brain potential most prominent over central parietal lobe sensors occurring 300–400 ms after stimulus onset). Electroencephalography data can be scored in multiple ways. A goal of B-SNIP was to maximize use of available information, so auditory evoked brain responses were scored in both the temporal and frequency domains. Rather than using a single or few sensors (the most typical approaches in the psychosis literature), the authors integrated brain activity over 64 sensors distributed around the whole head to estimate patterns of neural response associated with auditory stimulus processing. Scoring data in the temporal domain yields information on strength of brain signals in voltage units at particular time points (e.g., 50, 100, 150, or 300 ms after stimulus onset). Scoring in the frequency domain yields information on magnitude of brain signals in particular frequency ranges (e.g., delta, theta, alpha, beta, gamma); this frequency information also can be quantified as a function of time (e.g., gamma activity in the first 100 ms after stimulus onset). Extracting temporal voltage and frequency magnitude information provided more complete information on complex brain responses occurring during auditory processing than could be obtained using either quantification approach alone.

After data integration within paradigm set (BACS, stop signal, saccade tasks, electroencephalography paradigms), there were nine biomarker composite variables that differentiated psychosis probands from healthy persons. Those variables and their associated effect size separation from

healthy persons by DSM diagnosis are shown in Figure 12.11–2 for both probands (upper chart) and relatives (lower chart). Even though these biomarker composites have familiar names (e.g., P300 ERP), it is important to highlight that every biomarker variable listed below the abscissa is actually a composite of multiple measures (e.g., the P300 ERP has primarily six temporal and frequency domain measures determining the score).

There are three noteworthy features of Figure 12.11–2. First, as with most measures in psychosis research, the cases have deficient values (values below the mean of healthy persons) on most of the biomarker composites. The exceptions are for Paired S2 ERP (where schizophrenia cases have modestly "accentuated" activity), and Saccade Latency (with bipolar disorder with psychosis and schizophrenia cases having modestly faster reaction times). Although the Paired S2 ERP effect seems to be consistent with the traditional view that schizophrenia cases have difficulties with stimulus filtering, a more nuanced explanation of the factors determining that difference are available in the authors' published papers. Second, for most of the measures (BACS, Stop Signal, antisaccade, N100 ERP, and P300 ERP) there is a continuum of severity, with bipolar disorder with psychosis cases being least impaired, schizophrenia cases being most impaired, and schizoaffective disorder cases being intermediate. There were two variables (Intrinsic EEG, P200 ERP) where the three psychosis groups were indistinguishable. Third, the relatives show the same general pattern of effects as the probands to whom they are related, although the magnitudes of the effects are attenuated. Only two measures (BACS, antisaccade) showed effect sizes and overall effects among relatives suggesting they would be useful for family/genetic investigations. Of greatest interest, the clinically affected relatives (those with psychosis-related clinical diagnoses) of both bipolar disorder with psychosis and schizophrenia probands were similarly and significantly worse than normal on BACS scores. Among the clinically healthy relatives, however, only those related to schizophrenia probands had BACS scores worse than normal. On antisaccade performance, the relatives of the three probands groups did not differ from each other, although all three relative groups were worse than normal.

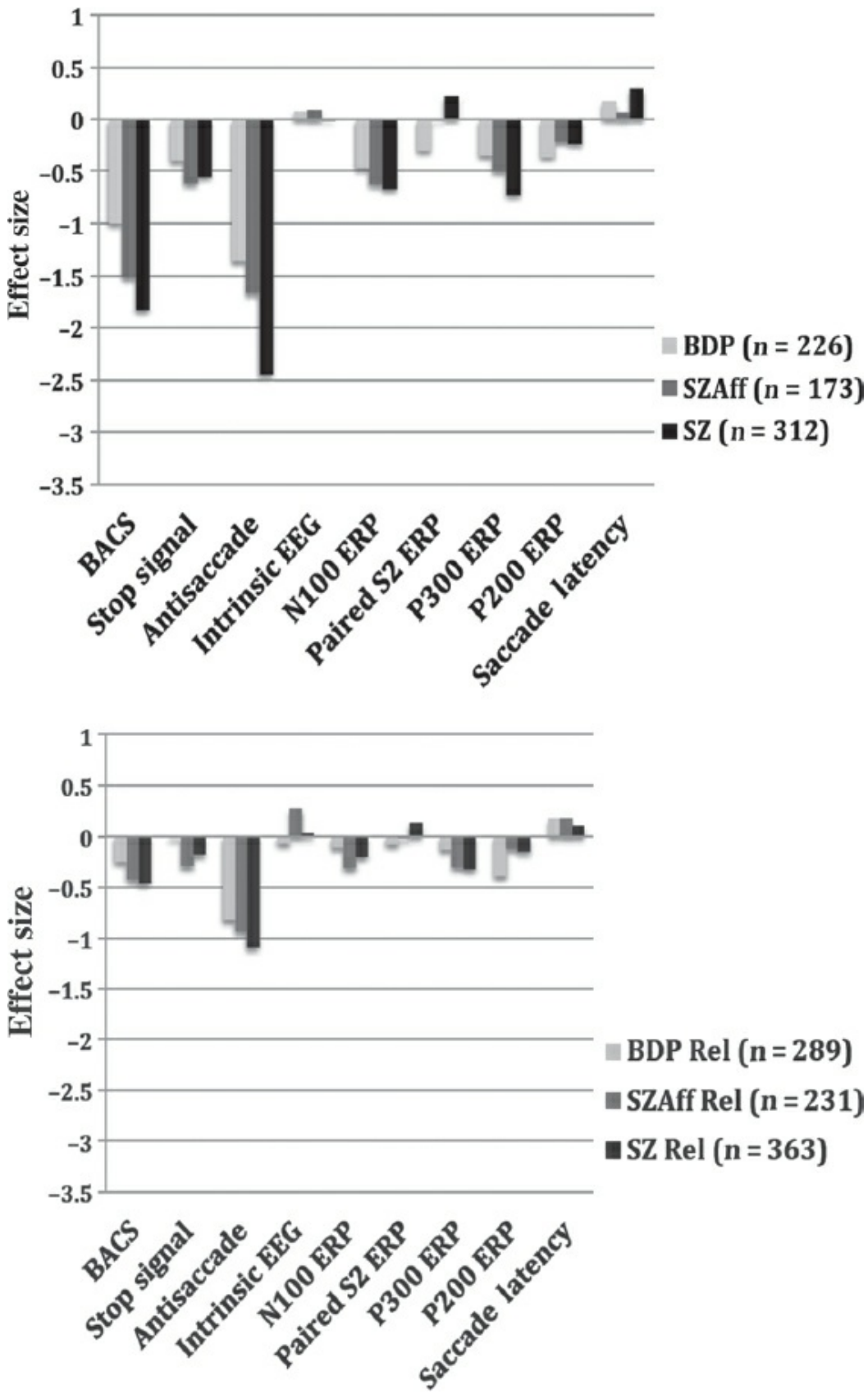

FIGURE 12.11–2. Biomarker scores by DSM diagnoses. Psychosis group differences from healthy persons in effect size units on the nine biomarker composite variables. Probands are shown in the top plot and relatives are shown in the bottom plot. Values less than zero indicate lower scores than healthy persons. BDP, bipolar disorder with psychosis;

SZAff, schizoaffective disorder; SZ, schizophrenia.

A prominent outcome of these initial B-SNIP analyses was, similar to analyses of the symptom data, demonstration of a continuum of psychosis severity on multiple biomarker measures. Only the BACS and antisaccade composites had effect size separations (greater than one standard deviation) suggesting their usefulness for identification of psychosis cases. There was also considerable overlap of bipolar disorder with psychosis, schizoaffective disorder, and schizophrenia groups on all of the biomarker composites. Indeed, the most extreme differences between diagnostic groups occurred on the antisaccade composite, where there was a one standard deviation separation between the most extreme scoring psychosis groups. This overlap on neurobiological features is ubiquitous in the psychosis literature, in this case on measures ranging from basic sensory processing through sensori-motor performance, inhibitory control, target detection, memory functions, and general executive control abilities. In addition, DSM diagnosis groups showed the same level of overlap on structural and functional brain imaging measures. There was only one measure that showed the hint of a distinctive, DSM diagnosis-related response pattern (Paired S2 ERP composite), but the difference between bipolar disorder with psychosis and schizophrenia groups on that measure was only 0.5 standard deviation units (a modest effect at best).

B-SNIP provides the largest and most comprehensive evaluation of biomarker-endophenotypes by DSM diagnostic status to date. The main outcomes do not support the thesis that DSM psychosis diagnoses capture neurobiologically distinct brain disorders. Only 45% of the probands were correctly classified when using the biomarker composite variables to predict diagnosis. Differences in degree, rather than differences in kind, of deviation described the overwhelming majority of DSM diagnosis-related variance. Perhaps B-SNIP used the wrong laboratory measures to address this issue (although measures were selected for their demonstrated success); this possibility can be tested (see below). In addition, despite many of the biomarker composite variables being reasonably familial (having shared score variance within a family), evidence that these measures could be helpful for understanding the genetics of DSM psychosis diagnoses was meager. The individual biomarker composites seemed less useful in this regard than their level of familial similarity would indicate.

ALTERNATIVES TO CURRENT DSM-TYPE PSYCHOSIS DIAGNOSTIC SYSTEMS

There were a number of factors leading B-SNIP collaborators to question whether DSM diagnoses-based analyses effectively captured the neurobiological uniqueness of psychotic disorders. Prominent among these factors was the considerable variability both within and between DSM psychosis groups on the laboratory measures. The authors had excellent

and expertly quantified data on a range of psychosis-relevant biomarkers. Rather than doubting the veracity of the laboratory data and their quantification, the authors questioned whether DSM-type diagnoses could be capturing neurobiological uniqueness.

Our primary alternative hypothesis was that biologically defined disease characteristics of serious psychiatric illness lie orthogonal to clinical diagnostic disease definitions. Research in other medical disciplines indicates that clinical phenomenological diagnoses fail to capture etiologically meaningful disease categories. Biological differentiation using biomarker/molecular/genetic data has resulted in classification of diseases with remarkably similar clinical presentations and pathology into distinct disorders. There is no reason to believe the psychoses should be different in this regard. As Fischer and Carpenter proposed in 2009: "Psychotic experience is to the diagnosis of mental illness as fever is to the diagnosis of infection—important but non-decisive in the differential diagnosis." If true, the correct approach would be to ignore DSM-type clinical diagnoses (except for the presence of psychosis) in favor of alternative methods of classification. One solution the authors explored was to develop biomarker-based classifications of psychoses informed by clinical neuroscience.

The next step in analysis of the B-SNIP data, therefore, was to rely on variance in the biomarker panel to capture neurobiologically distinctive psychosis subgroups independent of DSM definitions. This approach considered [big] psychosis as the main clinical syndrome of interest. The question, then, was how to meaningfully parse psychosis-relevant biomarker variance. The authors' answer was to use *k*-means clustering of the nine biomarker composite variables to form psychosis subgroups. This analytical method forms clusters that are defined by their centroids (or center of mass in the nine variable space of the biomarkers), with all cases in the cluster being as close to their respective centroid as possible. Only psychosis cases were used in this analysis because the biomarkers differentiated psychotic and healthy persons, so the problem was meaningfully parsing variance among psychosis cases.

It was established that three subgroups best captured neurobiological variance among the biomarker composite variables. Based on the *k*-means analysis, subjects were assigned to one of the three subgroups given their scores on these variables. Given the apparent neurobiological distinctiveness of the subgroups, the authors called them psychosis *biotypes* (biologically distinctive types). The associated effect size separations from healthy persons by psychosis biotype membership are shown in Figure 12.11–3 for both probands (upper chart) and relatives (lower chart). Figure 12.11–3 is analogous to Figure 12.11–2 that was organized by DSM diagnoses, using exactly the same subjects, but now grouped by biotype membership.

There are three noteworthy features of Figure 12.11–3. First, the effect

size separations of the biotypes from normal are considerably larger than they were for DSM diagnosis groups. Second, there are four different patterns of abnormalities between biotype groups: (i) on BACS, Stop Signal, antisaccade, and saccade latency biotype-1 probands have the lowest scores, biotype-3 probands have the highest scores, and biotype-2 probands are intermediate; (ii) on N100, P300, and P200 biotype-1 probands have the lowest scores, but biotype-3 probands are now the intermediate group; (iii) on Intrinsic EEG (a measure of accentuated background or nonspecific brain activity) biotype-1 probands are low while biotype-2 probands show accentuated activity; and (iv) on Paired S2 ERP biotype-1 probands have an accentuated response while biotype-2 probands have lower amplitude responses than normal. There was not a single variable on which the three biotypes were indistinguishable. Third, the relatives show the same general pattern of effects as the probands to whom they are related, although, as with the DSM data, the magnitudes of the effects are attenuated. Even when removing the clinically affected relatives, the same patterns of statistically significant effects remained, suggesting that these variables are indexing liability for disease among relatives when they are sorted by psychosis biotype.

It is impossible to illustrate the distribution of individual psychosis cases in the nine variable space of the biomarker composites. To simplify this task, the authors statistically formed two linear combinations (functions) of the biomarker composite variables that maximally differentiated the biotypes. One function had highest correlations with the BACS, stop signal, and antisaccade tasks, and the authors labeled it "cognitive control." The other function had highest correlations with the remaining biomarker composites, and the authors labeled it "sensorimotor reactivity." Each subject could be assigned a score on these functions, and Figure 12.11–4 shows the distributions of those scores by biotype (upper left) and DSM diagnosis (lower right). In the biotypes plot, in addition to displaying the scatter plots by biotype, the location of the healthy subjects in this bivariate space is also displayed. These bivariate scatter plots are a simplification for illustrative purposes, but they highlight some important points. First, biotype-3 probands, even though psychotic, largely overlap the distribution of healthy persons. Second, biotype-1 probands were profoundly low on both cognitive control and sensorimotor reactivity. Third, biotype-2 probands are moderately impaired on cognitive control but have accentuated sensorimotor reactivity. Fourth, when the same data were sorted by DSM diagnoses, there were obvious tendencies for schizophrenia to be more deviant than bipolar disorder with psychosis cases on both functions, schizoaffective disorder was indistinguishable from schizophrenia cases, and there was tremendous heterogeneity within and overlap between DSM psychosis diagnostic groups.

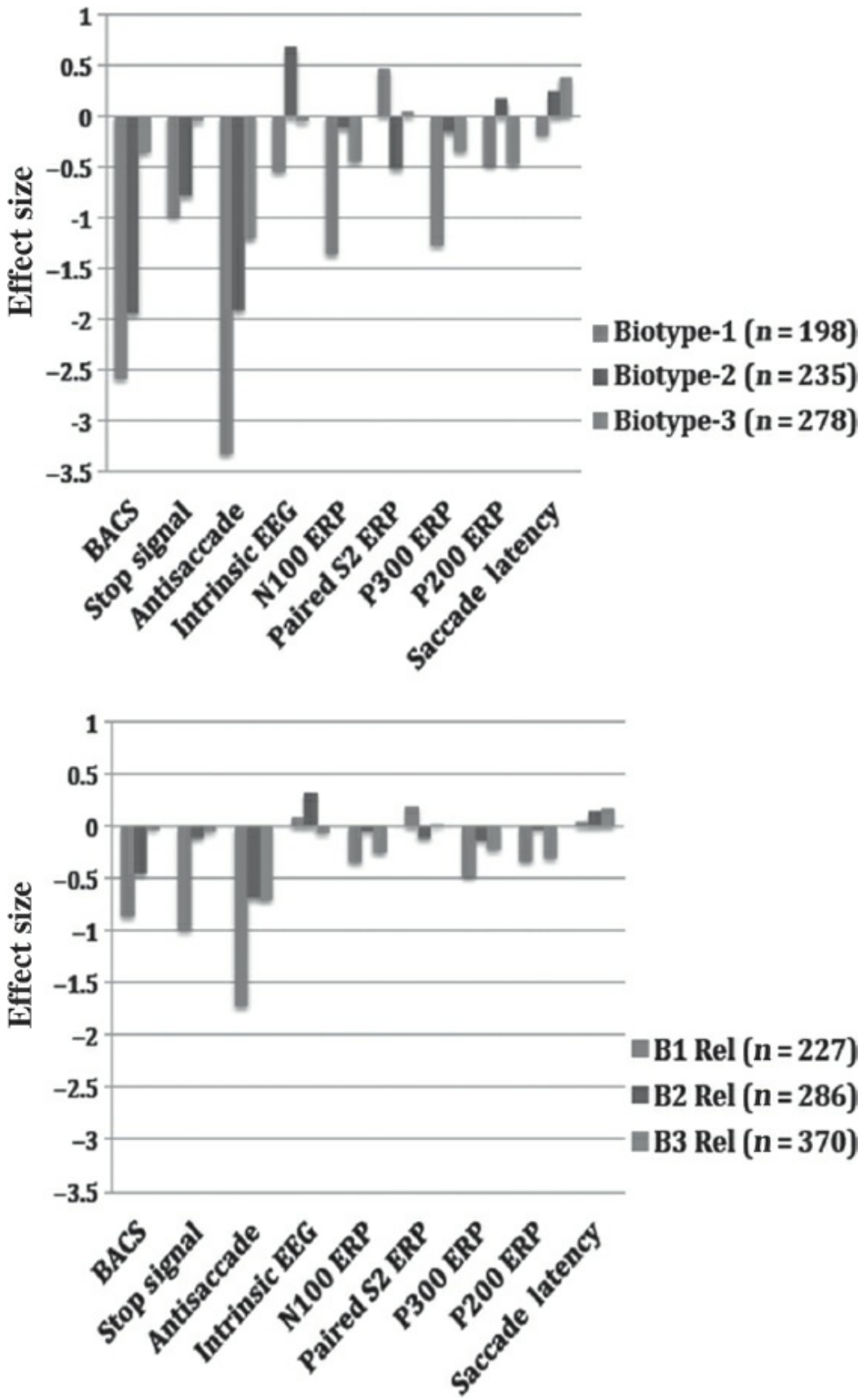

FIGURE 12.11–3. Biomarker scores by psychosis biotype. Psychosis group differences from healthy persons in effect size units on the nine biomarker composite variables. Probands are shown in the top plot and relatives are shown in the bottom plot. Values less than zero indicate lower scores than healthy persons. B1, biotype-1; B2, biotype-2; B3,

biotype-3; Rel, relative groups.

In addition to the first-degree relative data above, the biotype categories were tested for uniqueness using measures not included in biotype class formation. These external validators included social functioning, measures of brain volume from structural magnetic resonance images, and clinical diagnoses of the first-degree relatives. A summary of the main defining characteristics of the biotypes is provided below, including group comparisons on the external validating measures:

Biotype-1: [59% schizophrenia, 20% bipolar disorder with psychosis – 28% of psychosis sample]. These probands were severely impaired on both cognitive control and sensorimotor reactivity. They had the highest proportion of inhibition errors, the lowest amplitude brain responses to auditory stimuli but comparatively accentuated neural responding to repeated auditory stimuli, had sluggish responses (slow reaction times) to sensory inputs, and the poorest target (critical stimulus) detection. On the external validating measures, they had the worst social functioning among the biotypes, and the most extensive and severe reduction in structural brain volumes that were in predominantly frontal, cingulate, temporal, and parietal cortices as well as basal ganglia and thalamus.

Biotype-2: [46% schizophrenia, 27% bipolar disorder with psychosis – 33% of psychosis sample]. These probands were moderately impaired on cognitive control, but had accentuated sensorimotor reactivity, including normal to higher amplitude neural responding to auditory inputs, the highest intrinsic (not specifically stimulus-related) neural activities, but intact neurophysiological evidence of target detection. A pattern of moderately impaired cognitive control and high sensorimotor reactivity is inconsistent with a continuum-of-severity characterization of the biotypes. On the external validating measures, they had social functioning intermediate between the biotype-1 and biotype-3 probands, and structural brain volume reductions that were spatially similar to but less extreme than those observed among biotype-1 probands.

Biotype-3 [32% schizophrenia, 44% bipolar disorder – 39% of psychosis sample]. These probands did not differ from healthy persons on cognitive control (despite having a psychosis diagnosis) but were modestly deviant on Sensorimotor Reactivity. They had modestly more inhibition errors than was normal, modestly lower neural responding to auditory inputs, modestly impaired target detection, and the fastest visual orienting times of the three biotypes (significantly faster than healthy persons). On the external validating measures, they had the highest social functioning of the biotype groups, although still significantly worse than normal, and structural brain volume reductions that were modest and primarily located in anterior limbic regions.

When relatives were categorized by their relationship to biotype probands they showed the same, although less extreme, performance patterns on cognitive control and sensorimotor reactivity as the proband groups. The relatives also had the same patterns of compromised social functioning and structural brain deviations in less extreme form as those observed among the biotype probands. Finally, rates of psychosis-related illnesses among biotype-1 (26%) and biotype-2 (24%) relatives were significantly higher than those observed among biotype-3 relatives (17%). It is important to note than none of these systematic biological relative effects were observed when groups were categorized by DSM diagnoses.

As the authors did for DSM diagnoses, they also created linear combinations of the psychosis-related clinical symptoms that best differentiated the three biotypes. The main outcome of this analysis is illustrated in Figure 12.11–1 (lower graph). There were two statistically significant dimensions that differentiated the biotypes. One dimension the authors labeled "Kraepelinian" given its primary symptom correlates, which included difficulty in abstract thinking, blunted affect, conceptual disorganization, poor attention, disorientation, hallucinatory behavior, and lack of spontaneity. Biotype-1 cases, who had the most profound neurocognitive deficits, were highest on this dimension and biotype-3 cases, who were most normal across biomarkers, were lowest. The other dimension the authors labeled "neurocognitive dysregulation" given its symptom correlates, which included decreased need for sleep, unusual thought content, stereotyped thinking, thought disorder, mannerisms and posturing, concentration difficulty, and anxiety. This dimension also had a prominent correlate with the Kraepelinian symptom of disorientation. The Dysregulation dimension differentiated the biotype-2 cases from the other two biotypes.

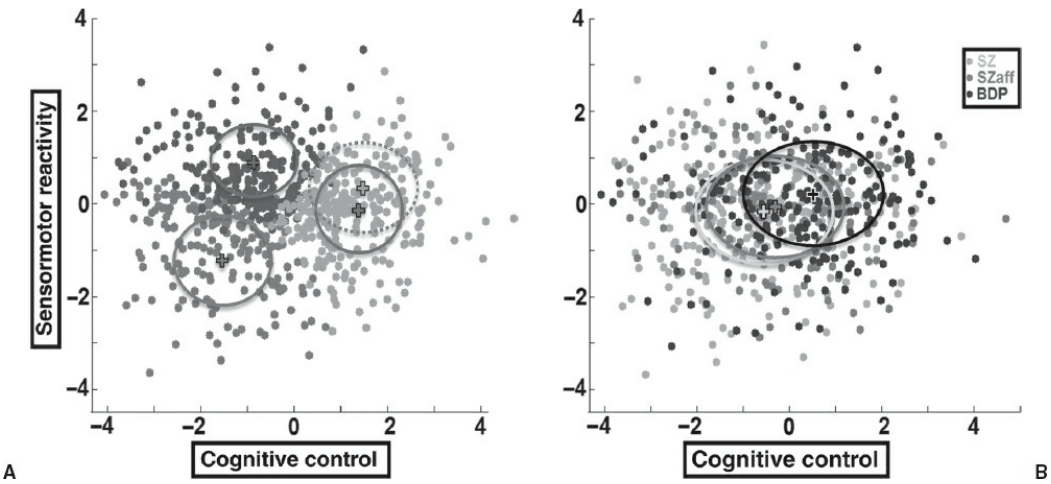

FIGURE 12.11–4. **A:** Sensorimotor reactivity by cognitive control as a function of biotype. Each dot indicates the location of an individual in this bivariate biomarker space (red, biotype-1; blue, biotype-2; green, biotype 3). The color-coded crosses and ellipsoids show the centroids and ±1 SD for each biotype. The gray cross and ellipse show the same

values for the healthy group. **B:** Sensorimotor reactivity by cognitive control as a function of DSM diagnosis. Each dot indicates the location of an individual in this bivariate biomarker space (light gray, schizophrenia [SZ]; medium gray, schizoaffective disorder [SZAff]; black, bipolar disorder with psychosis [BDP]). The gray scale coded crosses and ellipsoids show the centroids and ±1 SD for each DSM diagnosis. (See color atlas.)

The symptom correlates of the biotypes were interesting by virtue of their differences from the DSM categories. Although the Kraepelinian dimension contains a psychotic symptom (hallucinatory behavior), psychotic symptoms are not the main features differentiating the biotypes. This fact is consistent with the position that psychotic symptoms are not decisive for differentiating neurobiologically distinct types. Psychosis symptoms may be akin to the symptoms of fever, dropsy, or congestive heart failure, and, like most symptoms, not typically useful for differential diagnosis. The clinical characteristics differentiating the biotypes, however would seem to provide useful clues to neurobiological features (i.e., biotype-2, the cases with accentuated sensorimotor reactivity, also had differentiating symptom features concordant with a neural dysregulation syndrome). It may be possible, therefore, to identify symptomatic features characteristic of neurobiological types working from the "bottom up" rather than by working from the "top down" (the approach taken by framers of DSM-type psychosis definitions). A related approach has shown promise in diseases relevant to psychiatry, for instance, 22q11 deletion syndrome.

PSYCHOTIC DISORDERS: DIFFERENCES IN KIND?

The answer to this question for DSM-type clinical psychosis diagnoses has been known for some time. Using clinical symptoms alone to define diseases fails to capture biological reality. A reason psychiatry has shown modest progress as a medical discipline may be related to a discomfort with abandoning a primarily clinical/descriptive approach to diagnosis. Prominent scientists, among them Drs. Hyman, Insel, and Jablensky, have questioned psychiatry's reliance on this approach and have elaborated on how such persistence may impede the field's ability to discover etiological mechanisms undergirding discrete disease entities and the development of successful treatments such discoveries would support.

An impediment to the development of classifications with stronger biological foundations is the lack of data supporting such distinctions between diseases. This is a legitimate concern, but one with a certain level of circularity. Historically, the gold standards for determining the usefulness of neurobiological measures of the psychoses are DSM-type psychosis definitions. To the extent that such definitions are orthogonal to neurobiological reality, the ability to identify etiological distinctiveness of psychosis subtypes has been severely limited. The B-SNIP project offers proof of concept that by eschewing DSM-type clinical psychosis diagnoses it is possible to more accurately define neurobiologically distinctive

psychosis subtypes. Although the authors' approach has limitations (e.g., the identification of subtypes is determined by the biomarkers used to dissect those types, the longitudinal characteristics of the biotypes is uncertain, there are certainly more than three etiologically distinctive psychosis classes), the verisimilitude of biotypes was supported by multiple external validating measures.

The specific characteristics of the psychosis biotypes must be evaluated and replicated. The biomarker panel must be refined and expanded in theoretically important ways. More varied and powerful statistical analysis must be used to support the parsing of neurobiological uniqueness as manifest among the psychosis-relevant biological variables. Molecular foundations must be demonstrated for the biotypes. Nevertheless, members of the B-SNIP consortium are excited for the future and the possibility that the authors' work, and the work of similar research teams, will support a transformative reconceptualization of the psychoses.

References

Clementz BA, Sweeney JA, Hamm JP, et al. Identification of distinct psychosis biotypes using brain-based biomarkers. *Am J Psychiatry*. 2016;173(4):373–384.

Clementz BA, Wang J, Keil A. Normal electrocortical facilitation but abnormal target identification during visual sustained attention in schizophrenia. *J Neurosci*. 2008;28:13411–13418.

Craddock N, O'Donovan MC, Owen MJ. Psychosis genetics: modeling the relationship between schizophrenia, bipolar disorder, and mixed (or "schizoaffective") psychoses. *Schizophr Bull*. 2009;35:482–490.

Erickson MA, Ruffle A, Gold JM. A meta-analysis of mismatch negativity in schizophrenia: from clinical risk to disease specificity and progression. *Biol Psychiatry*. 2016;79(12):980–987.

Ethridge LE, Hamm JP, Pearlson GD, et al. Event-related potential and time-frequency endophenotypes for schizophrenia and psychotic bipolar disorder. *Biol Psychiatry* 2015;77:127–136.

Ethridge LE, Soilleux M, Nakonezny PA, et al. Behavioral response inhibition in psychotic disorders: diagnostic specificity, familiality, and relation to generalized cognitive deficit. *Schizophr Res*. 2014;159:491–498.

Gatt JM, Burton KL, Williams LM, Schofield PR. Specific and common genes implicated across major mental disorders: a review of meta-analyses. *J Psychatr Res*. 2015;60:1–13.

Gershon ES, Alliey-Rodriguez N, Liu C. After GWAS: searching for genetic risk for schizophrenia and bipolar disorder. *Am J Psychiatry*. 2011;168:253–256.

Gottesman II, McGue M. Endophenotype. In: Cautin RL, Lilienfeld SO, eds. *The Encyclopedia of Clinical Psychology*. 1st ed. John Wiley & Sons, Inc; 2015.

Hamm JP, Ethridge LE, Boutros NN, et al. Diagnostic specificity and familiality of early versus late evoked potentials to auditory paired stimuli across the schizophrenia-bipolar psychosis spectrum. *Psychophysiol*. 2014;51:348–357.

Hamm JP, Ethridge LE, Shapiro JR, et al. Family history of psychosis moderates early auditory

cortical response abnormalities in non-psychotic bipolar disorder. *Bipolar Disord*. 2013;15:774–786.

Hill SK, Reilly JL, Keefe RS, et al. Neuropsychological impairments in schizophrenia and psychotic bipolar disorder: findings from the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) study. *Am J Psychiatry*. 2013;170:1275–1284.

Hyman SE. The diagnosis of mental disorders: the problem of reification. *Annu Rev Clin Psychol*. 2010;6:155–179.

Insel TR, Cuthbert BN. Brain disorders? Precisely. *Science*. 2015;348:499–500.

Ivleva EI, Bidesi AS, Keshavan MS, et al. Gray matter volume as an intermediate phenotype for psychosis: Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *Am J Psychiatry*. 2013;170:1285–1296.

Ivleva EI, Morris DW, Moates AF, Suppes T, Thaker GK, Tamminga CA. Genetics and intermediate phenotypes of the schizophrenia-bipolar disorder boundary. *Neurosci Biobehav Rev*. 2010;196:38–44.

Jablensky A. The diagnostic concept of schizophrenia: its history, evolution, and future prospects. *Dialogues Clin Neurosci*. 2010;12:271–287.

Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry*. 2012;17:1174–1179.

Kendler KS, Parnas J. (Eds). *Philosophical Issues in Psychiatry ii: Nosology*. Oxford: Oxford University Press; 2012.

Keshavan MS, Clementz BA, Pearlson GD, Sweeney JA, Tamminga CA. Reimagining psychoses: an agnostic approach to diagnosis. *Schizophr Res*. 2013;146:10–16.

Keshavan MS, Morris DW, Sweeney JA, et al. A dimensional approach to the psychosis spectrum between bipolar disorder and schizophrenia: the schizo-bipolar scale. *Schizophr Res*. 2011;133:250–254.

Khadka A, Meda SA, Stevens MC, et al. Is aberrant functional connectivity a psychosis endophenotype? A resting state functional magnetic resonance imaging study. *Biol Psychiatry*. 2013;74:458–466.

Lichtenstein P, Yip BH, Björk C, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet*. 2009;373:234–239.

Meda SA, Ruaño G, Windemuth A, O'Neil K, Berwise C, Dunn SM, et al. Multivariate analysis reveals genetic associations of the resting default mode network in psychotic bipolar disorder and schizophrenia. *Proc Natl Acad Sci USA*. 2014;111:E2066–2075.

Narayanan B, Ethridge LE, O'Neil K, et al. Genetic sources of subcomponents of event-related potential in the dimension of psychosis analyzed from the B-SNIP study. *Am J Psychiatry*. 2015;172(5):466–478.

Reilly JL, Frankovisch K, Hill S, et al. Elevated antisaccade error rate as an intermediate phenotype for psychosis across diagnostic categories. *Schizophr Bull*. 2014;40:1011–1021.

Rzhetsky A, Wajngurt D, Park N, Zheng T. Probing genetic overlap among complex human phenotypes. *Proc Natl Acad Sci USA*. 2007;104:11694–11699.

Schneider M, Debbane M, Bassett AS, et al. Psychiatric disorders from childhood to adulthood in 22q11.2 deletion syndrome: Results from the International Consortium on Brain and Behavior in 22q11.2 Deletion Syndrome. *Am J Psychiatry*. 2014;171:627–639.

Skudlarski P, Schretlen DJ, Thaker GK, et al. Diffusion tensor imaging white matter endophenotypes in patients with schizophrenia or psychotic bipolar disorder and their relatives. *Am J Psychiatry*. 2013;170:886–898.

Tamminga CA, Ivleva EI, Keshavan MS, et al. Clinical phenotypes of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *Am J Psychiatry*. 2013;170:1263–1274.

▲ 12.12 Schizophrenia: Pharmacological Treatment

JOHN M. KANE, M.D., AND CHRISTOPH U. CORRELL, M.D.

Nearly every patient with schizophrenia will benefit from pharmacological treatment. Antipsychotic medications—the mainstay of pharmacological treatment—are effective for reducing the impact of psychotic symptoms such as hallucinations, delusions, and disorganized thoughts and behavior, as well as related symptoms of agitation and aggression. In many patients, these symptoms can be completely eliminated. Once these symptoms are minimized, medications can decrease the likelihood that symptoms will recur. Nevertheless, negative symptoms and cognitive dysfunction are only very inadequately treated with currently available antipsychotics.

Moreover, controlling positive psychotic symptoms is often insufficient for improving the community functioning and the quality of life of patients with schizophrenia. There is also an increased awareness of the serious health problems of patients with schizophrenia and the possible role of some antipsychotics in worsening these problems. These observations have led to principles of drug treatment of schizophrenia that emphasize the necessary, but not always sufficient, role of antipsychotic treatment in helping patients achieve their personal goals.

HISTORY

The history of somatic therapies in schizophrenia can be divided into two eras, with the discovery of chlorpromazine (Thorazine)—the first clearly effective antipsychotic drug—as the dividing line. Prior to the introduction of antipsychotics in the early 1950s, several treatments had been administered to individuals with psychotic illness, with results that are difficult to interpret because careful research methods in psychiatry had not been developed. During the late 19th and early 20th centuries, schizophrenia was viewed as an illness that was associated with an inevitable deterioration into dementia. As a result, patients were frequently hospitalized for long periods of time. Somatic treatments were used to help control the most severe symptoms of the disorder and to make hospitals safer for both patients and staff. Sedating agents such as bromides and barbiturates were used to control agitation, and physical treatments such as

hydrotherapy and wet sheet packs were also used for their calming effects. In the early 1920s, sleep treatment with barbiturates was introduced. This treatment was based on the observation that patients tended to improve following an overdose of barbiturates. The method involved maintaining patients in a highly sedated state for days during which they would awaken only for necessary activities, such as eating and personal hygiene.

Insulin coma treatment was introduced during the 1930s. Patients were administered gradually increasing doses of insulin until a coma was introduced. After an hour of monitoring, glucose was administered, terminating the coma. Patients were commonly administered as many as 20 comas. Insulin coma was widely used in the treatment of psychosis, suggesting that it may have been somewhat effective. Unfortunately, it was never exposed to adequate research trials, and it remains unclear if the treatment was effective. It was abandoned when antipsychotics were introduced.

Prefrontal lobotomy was proposed as a treatment for serious mental illnesses by Moniz in 1935. The support for this treatment came from animal studies in which frontal lobe extirpations in monkeys resulted in an animal that appeared less easily frustrated. The use of frontal lobotomy was common prior to the introduction of effective antipsychotics, although there is a remarkable lack of controlled studies comparing psychosurgery to other treatments. Although reports suggest that lobotomy may have been effective in reducing severe psychotic symptoms, they also resulted in deteriorations in other areas. Following lobotomies, patients frequently demonstrated personality deterioration with impulsive and psychopathic behaviors, as well as impairments in concept formation and the ability to plan. Psychosurgery was abandoned as a treatment for schizophrenia after the introduction of effective antipsychotic medications.

Convulsive therapies were developed after it was observed that some patients improved after a seizure. Drugs such as camphor and pentylenetetrazol (Metrazol) were used initially to induce seizures, but were abandoned after Cerletti and Bini proposed the use of electrically induced convulsions. In its early days, electroconvulsive therapy (ECT) was administered without anesthetics or muscle relaxants. The lack of anesthetics inspired fear in many patients, and the lack of muscle relaxants led to injuries from forceful muscle contractions. ECT continues to have a role in certain patients with schizophrenia and is discussed in a later section.

The first effective antipsychotic medications were probably derived from extracts of the rauwolfia plant. Publications from the 1930s and 1940s suggested that these agents were effective for both hypertension and psychosis. Reserpine, the most potent of the rauwolfia alkaloids, was introduced in the early 1950s and was widely prescribed in the United States and elsewhere for schizophrenia and other psychotic illnesses.

Studies comparing reserpine to dopamine receptor antagonists suggested that their efficacy was similar. However, reserpine's side effects, particularly depression, led most clinicians to prefer the dopamine receptor antagonists. Thus, reserpine is only rarely used for managing psychosis.

The discovery of chlorpromazine in the early 1950s may be the most important single contribution to the treatment of a psychiatric illness. Laborit, a surgeon in Paris, noticed that administering chlorpromazine to patients prior to surgery resulted in an unusual state in which they seemed less anxious regarding the procedure. In 1952 he convinced Delay and Deniker and other psychiatrists to administer chlorpromazine to psychotic and excited patients. The effects were extraordinary. Chlorpromazine was effective at reducing hallucinations and delusions as well as excitement. It was also noted that it caused side effects, including rigidity, tremor, and bradykinesia, that appeared similar to parkinsonism. The use of chlorpromazine spread rapidly through the psychiatric hospitals in Paris and eventually to the rest of the world. Because chlorpromazine was effective and relatively easy to use, it and other antipsychotic medications were partially responsible for a substantial reduction in the number of patients in psychiatric hospitals.

Thioridazine (Mellaril) and fluphenazine (Prolixin), as well as newer antipsychotic classes, such as the butyrophenones (e.g., haloperidol [Haldol]) and the thioxanthenes (e.g., thiothixene [Navane]), were developed after the introduction of chlorpromazine. Although these newer agents differed in their potency and their side-effect profiles, all were similar in their effectiveness. Clozapine (Clozaril), the first effective antipsychotic with negligible extrapyramidal side effects (EPS), was discovered in 1958 and first studied during the 1960s. However, in 1976 it was noted that clozapine was associated with a substantial risk of agranulocytosis. This property resulted in delays in the introduction of clozapine. In 1990 clozapine finally became available in the United States, but its use was restricted to patients who responded poorly to other agents. The introduction of risperidone (Risperdal) in 1994, olanzapine (Zyprexa) in 1996, quetiapine (Seroquel) in 1997, ziprasidone (Geodon) in 2001, aripiprazole (Abilify) in 2002, paliperidone (Invega) in 2007, asenapine (Saphris) in 2009, iloperidone (Fanapt) in 2009, lurasidone (Latuda) in 2010, and brexpiprazole (Rexulti) and cariprazine (Vraylar) in 2015 have given clinicians new alternatives for treating a large number of patients with schizophrenia. Table 12.12–1 summarizes pharmacologic, formulation, and dosing information for SGAs and selected FGAs.

PHASES OF TREATMENT IN SCHIZOPHRENIA

Pharmacological treatment differs depending on the phase of a patient's illness. The acute stage is usually characterized by positive psychotic symptoms and related agitation/aggression that require immediate clinical

attention. These symptoms may represent a first psychotic episode or, more commonly, a relapse in an individual who has experienced multiple prior episodes. Treatment during this phase focuses on alleviating the most severe psychotic symptoms. Following the acute phase, which usually lasts from 4 to 8 weeks, patients will generally enter a stabilization phase in which acute symptoms have been controlled, but patients remain at risk for relapse if treatment is interrupted or if the patients are exposed to stress. During this phase, treatment focuses on consolidating therapeutic gains, with similar treatments as those used in the acute stage. This phase may last as long as 6 months following recovery from acute symptoms. The third stage is the stable or maintenance phase when the illness is either in a relative stage of remission or symptomatically stable. The goals during this phase are to prevent psychotic relapse or exacerbation and to assist patients in improving their level of functioning. Although the control of positive psychotic symptoms is a major goal across all three treatment stages, minimization of negative symptoms and cognitive dysfunction is also highly relevant since negative symptoms and cognitive disturbances have been associated even more with functional impairment than positive symptoms, leading to considerable gaps in achieving recovery.

FIRST EPISODE SCHIZOPHRENIA

The treatment of patients in early phase illness differs from that of patients with chronic schizophrenia. First episode patients are generally more treatment responsive, require lower antipsychotic doses (approximately 50%), and are generally more sensitive to side effects. Patients with first episode schizophrenia often have a difficult time accepting their illness, are at high risk for nonadherence, relapse, psychosocial deterioration, and suicidality. Although, as in chronic-stage illness, antipsychotic treatment is the cornerstone of management, multidisciplinary interventions, focusing on engagement, treatment continuation, relapse prevention, physical health, and functional recovery are paramount.

EFFECTIVENESS OF ANTIPSYCHOTIC MEDICATIONS

A large body of evidence supports the effectiveness of antipsychotics for schizophrenia. Many of these studies were carried out in the 1960s, when there was skepticism that these agents were truly antipsychotic rather than more effective tranquilizers. An evaluation of these studies by the 1995 Schizophrenia Patient Outcomes Research Team (PORT) found that about 70 percent of patients treated with an antipsychotic achieved relative remission or substantial improvement. In contrast, only about 25 percent of patients treated with placebo had similar degrees of response. Most studies compared one or more antipsychotic with either a placebo or an agent such as phenobarbital that served as a control. Antipsychotics were found to be more effective than either placebo or tranquillizers. Recent

meta-analyses of high-quality randomized controlled, parallel design studies confirm that antipsychotics are significantly superior to placebo for total symptomatology and positive symptoms.

Table 12.12–1.
Pharmacologic, Formulation and Dosing Information for Second-Generation Antipsychotics and Selected First-Generation Antipsychotics

Antipsychotic	Principal Liver Enzyme Target	Protein Binding	Bio-Availability	Time to Peak Level	Half-Life	CPZ Dose Equivalent, mg[a]	Typica ing Do
Second-Generation Antipsychotics							
Partial Dopamine D2 Agonists							
Aripiprazole	2D6 >3A4	>99%	87%	po: 3–5 hours; im short: 1–3 hours; im long: 5–7 hours	po: 75 hours; im long: 30–47 days	7.5	po: 10- long (glu 400 wee arip 10–:
Brexpiprazole	2D6, 3A4	>99%	95%	4 hours	91 hours	N/A	0.5–1
Cariprazine	3A4>2D6	91–97%	52% (1 mg)	3–6 hours	2-5 days (didesmethyl-cariprazine: 1–3 weeks)	N/A	1.5
Dopamine D2—Serotonin 2A Anatagonists							
Asenapine	1A2>3A4	98	35% (≤2% if swallowed)	0.5–1.5 hours	First: 6 hours, terminal: 24 hours	7.5	5 bid
Clozapine	1A2 (30%) >2C19 (24%) >3A4 (22%) >2C9 (12%)	97%	50–60%	1.5–2.5 hours	12 hours	50	12.5

	>2D6 (6%)						
Iloperidone	2D6> 3A4>1A2	95%	96%	2–4 hours	18 hours	5	1 mg b day, 1 da incr mg/ to ther dose
Lurasidone	3A4	99%	9–19%	1–3 hours	18 hours	25	40–80
Olanzapine	1A2, 2D6, 3A4	93%	60%	po: 6 hours; im short: 15–45 minutes; im long: 7 days	po: 30 hours; im short: 30 hours; im long: 30 days	5	po: 5–: long (glut 210, or 4 wk (10), wee oral 20)
Paliperidone	<10% first-pass hepatic clearance	74%	28%	po: 24 hours; im long: 13 days	po: 23 hours; im long: 25–49 days	3	po: 6; i 234 into day 156 8, th 234 delt glut ever wee
Quetiapine	3A4	83%	100%	IR: 1.5 hours; XR: 6 hours	6–7 hours	75	IR: 25– XR: 300
Risperidone	2D6>3A4	90%	70%	3 hours	3 hours	2	po: 2; i (del glut with wee risp

Ziprasidone	Aldehyde oxidase (2/3) >3A4 (1/3)	>99%	60%	po: 6–8 hours; im short: ≤60 minutes	po: 7 hours; im short: 2–5 hours	60	20–40
First-Generation Antipsychotics (Dopamine D2 Antagonists)							
Chlorpromazine	2D6	>90	20%	2–4 hours	30 hours	100	po: 25- mg; shor follc 25–! as n afte hou
Fluphenazine	1A2	>90	<50%	po: 2 hours; im long: 8–10 hours	po: 14–16 hours; im long: 14 days	2	po: 2.5 2–3 dose
Haloperidol	3A4	92%	60%–70%	po: 2–6 hours; im short: 10–20 minutes; im long: 6–7 days	po: 18 hours; im short: 10–20 hours; im long: 3 weeks	2	po: 1–2 shor mg 4–8 as n im l (del glut 10–2 time oral ever wee (≤10
Perphenazine	2D6	>90%	40%	1–3 hours	9–12 hours	10	Inpatie 16 d in 2- dose Out 4–8 in 3

CPZ, chlorpromazine; Diss, dissolvable tablet; ER, extended release; im, intramuscular; IR, immediate release; iv, intravenous; N/A, not available; q, every; po, oral; Supp, suppository; XR, extended release.

[a]Chlorpromazine dose equivalents (i.e., dose given in the table is equivalent to 100 mg of chlorpromazine).

[b]Doses need to be individualized based on efficacy and tolerability.

Source. Package insert information for each medication.

TREATMENT OF ACUTE EPISODES

Indications for Pharmacological Treatment

Nearly all schizophrenia patients with acute psychotic symptoms will benefit from an antipsychotic medication, albeit to different degrees. Aside from relieving the patient from symptoms, there is some evidence indicating that lengthy delays in initiating antipsychotic treatment may worsen the long-term course of schizophrenia. Some other studies—but not all—suggest that a longer time between the first onset of psychosis and the initiation of treatment is related to a worse outcome. Many of the studies suggesting that delayed treatment worsens outcomes have important limitations, such as lack of randomization and comparing individuals treated during different decades. However, a definitive study will never be carried out to determine if withholding treatment worsens the long-term course of schizophrenia. As a result, it is probably prudent for clinicians to consider the possibility that untreated psychosis can result in a poorer prognosis. Moreover, recent evidence also suggests that more frequent relapses and longer time spent in relapse are associated with worse treatment outcome and, possibly, cortical volume reductions.

These data do not mean that all patients need to be treated immediately. There are circumstances in which the management of a patient may improve if drug treatment is delayed several days. A brief delay may permit clinicians to conduct a more thorough diagnostic evaluation and rule out causes of abnormal behavior, such as substance abuse, extreme stress, medical illnesses, and other psychiatric disorders.

Assessment

Whenever possible, patients should receive a physical, neurological, and mental status examination, and a laboratory evaluation before medications are started. Blood tests for complete blood cell count (CBC), electrolytes, fasting glucose, hemoglobin A1C, lipid profile, liver, renal, and thyroid function should be ordered. Other evaluations that should be considered are pregnancy tests in women, as well as human immunodeficiency virus (HIV) and syphilis tests when relevant. The presence of movement disorders, particularly preexisting tardive dyskinesia (TD), should be assessed as this may influence the choice of antipsychotic.

Since antipsychotics are relatively safe drugs, treatment can usually

begin before the results of laboratory tests are known. An exception is clozapine treatment, which should only begin after the patient is confirmed as having a normal white cell and granulocyte cell count. Under emergent conditions, for example, in potentially dangerous situations when patients refuse to cooperate with an evaluation, antipsychotics can be administered prior to a medical evaluation.

Selection of an Antipsychotic Drug

Although the usefulness of the distinction has been questioned, antipsychotics can be categorized into two main groups: the older conventional antipsychotics, which have also been called first-generation antipsychotics (FGAs) or dopamine receptor antagonists, and newer drugs, which have been called second-generation antipsychotics (SGAs) or serotonin-dopamine antagonists. The FGAs are further categorized as being low, mid, or high potency, with the higher potency drugs having more specificity and greater affinity for D_2 receptors and a greater tendency to cause EPS. Lower potency drugs are less likely to cause EPS, but tend to have more effects on other neurotransmitter receptors and are thus more likely to cause postural hypotension, sedation, weight gain, and anticholinergic effects. The SGAs are further categorized as D2 partial agonists and D2-serotonin 2a antagonists (Table 12.12–1).

Comparisons of Antipsychotics

Numerous studies have examined the comparative effectiveness and tolerability of various antipsychotic drugs. In general, all available antipsychotics appear to be similarly effective, at least at a group level, but with substantial differences in adverse effects. The exception is clozapine, which has been consistently shown to be more effective than other medications for individuals who have symptoms that persist despite adequate treatment with other antipsychotics, at least when clozapine is adequately dosed.

The makers and marketers of antipsychotics have sponsored many comparative studies, but research has shown that the sponsor's drug is frequently found to be superior. Independently conducted meta-analyses have found few effectiveness differences between drugs. An exception is that amisulpride (not available in the United States), olanzapine, and risperidone have been shown in some meta-analyses to produce greater symptom response than antipsychotics other than clozapine.

Because of the perception of bias in industry-sponsored studies, government-funded agencies in the United Kingdom and the US sponsored studies intended to produce independent research findings regarding the longer term effects of antipsychotic drugs. These studies, the CATIE schizophrenia trial in the United States and the Cost Utility of the Latest

Antipsychotic Drugs in Schizophrenia Study (CUtLASS) in the United Kingdom, found that clozapine seemed advantageous for individuals with refractory symptoms, but no clearly important advantages in effectiveness for any other drug or drug class for most patients. Both studies found that people tended to continue taking olanzapine longer than other antipsychotic medications, but neither study found greater reductions in symptoms or improvements in quality of life for olanzapine. Neither CATIE nor CUtLASS found substantial advantages in overall tolerability or acceptability for SGAs over FGAs. The CATIE study found that some antipsychotics, in particular clozapine, olanzapine, and quetiapine, were associated with weight gain and serum lipid abnormalities that are risk factors for cardiovascular disease.

Route of Administration

The decision regarding route of administration in the acute phase is usually straightforward. Under most conditions, patients should be treated with an oral antipsychotic. Short-acting intramuscular drugs are useful when the patient refuses oral dosing and when a very rapid onset is desired. Intramuscular administration of most antipsychotics results in peak plasma levels in about 30 minutes with clinical effects emerging within 15 to 30 minutes (Table 12.12–1). Most orally administered antipsychotics result in a peak plasma level in 1 to 6 hours following administration (Table 12.12–1).

Long-Acting Injectable Antipsychotics

Since nonadherence to antipsychotics is common and one of the most relevant and preventable risk factors for insufficient response and relapse, long-acting injectable formulations of antipsychotics have been developed (Table 12.12–1). These drugs differ from short-acting compounds in that they are released slowly over a period of several weeks and as a result can take days or weeks to achieve therapeutic serum concentration and weeks to months to achieve steady state. As a result, long-acting injectable antipsychotics are helpful for continuation and maintenance treatment, but not necessarily for rapid acute treatment. In the United States, the two FGAs, fluphenazine and haloperidol, have been formulated as oil-based esters formed between the alcohol group of the drug and a long chain fatty acid. After injection the drug is slowly released from the injection site by enzymatic hydrolysis of the ester and diffusion of the free drug. Currently, four SGAs, risperidone, olanzapine, paliperidone, and aripiprazole, are available as long-acting injectable antipsychotics in the United States. All four agents are water-based formulations that require biweekly (risperidone, certain doses of olanzapine), monthly (aripiprazole, certain doses of paliperidone, certain doses of olanzapine), 6-weekly (882 mg of

aripiprazole lauroxil) or 3-monthly (certain doses of paliperidone) injections. Oral supplementation is necessary for risperidone (3 weeks) and aripiprazole (2 weeks for the monohydrate and 3 weeks for the lauroxil formulation) while optimum plasma levels of their long-acting injectable formulation are being achieved. Long-acting forms of other SGAs are currently in development. In view of data that relapses reduce chances of response, at least in a subgroup of patients, the merits of preventing relapses earlier in the illness course and not only after multiple relapses have led to a chronic and debilitating illness course needs to be considered.

Starting Antipsychotics

Prior to starting an antipsychotic, clinicians should describe to patients the medication that is being prescribed, its target symptoms, and its possible side effects, particularly those that are common and unpleasant when a drug is first started (e.g., muscle stiffness, sedation, and akathisia). Patients who are severely disturbed may be unable to participate meaningfully in this discussion. However, most will benefit from information about the goals of treatment and important risks associated with antipsychotic medication. Since patients with schizophrenia may be suspicious, it is particularly important to emphasize that patients can participate as collaborators in interpreting medication effects. Since psychotic individuals may be dependent on the help and support of their families, it is frequently helpful to involve one or more family members in decision making about drug treatment.

In some settings and locations, it is necessary for patients to give written or verbal consent prior to receiving an antipsychotic medication. This can be a dilemma for patients who are conceptually disorganized and find it difficult to understand the risks and benefits of drug treatment. Under these circumstances, clinicians should adjust the complexity of the discussion to the patient's state of mind. Thus, it may be appropriate to provide a limited amount of information that focuses on the most common acute side effects of the medication when the patient is most seriously impaired. As the patient improves, clinicians may then elaborate on the costs and benefits of medication. For example, detailed discussions about TD or weight gain and metabolic side effects associated with chronic treatment may be deferred until the patient has improved and long-term maintenance is being considered.

Thus, it is important for psychiatrists to evaluate whether acutely disturbed patients are able to participate meaningfully in decisions about their medication. Clinicians should become familiar with local and state laws that affect a patient's right to refuse or accept drug treatment. The most difficult situation is when a patient who desperately needs medication refuses it. Under some conditions, family members who have been educated about schizophrenia may be helpful in convincing patients to

accept medication. Every locality has provisions for treating patients against their will under emergency conditions. Some permit involuntary treatment when certain conditions are met. Many localities now have laws allowing for compulsory outpatient treatment, particularly for patients who have a history of violent or aggressive behavior.

Dosage Selection

Finding the best dose for an antipsychotic is both difficult and important. The difficulty exists because the physician is unable to titrate dose against clinical effects due to the delay between a clinical intervention and the patient's clinical response. In some individuals, there is a delay of days or even weeks between the time when a medication is started and when the patient eventually responds. The notion of delayed response was initially supported by findings indicating that the neurochemical response to antipsychotic is complex and includes an initial blockade of central dopamine receptors followed by delayed decrease in dopamine turnover. However, more recent studies and meta-analyses have found that in a 4- to 6-week treatment course, the greatest reduction in symptoms occurs within the first 1 to 2 weeks. It has also been found that little or no response during the first week or two (assuming a valid diagnosis, appropriate adherence, and therapeutic dose level) is a powerful predictor of subsequent poor response.

The recommended dosage range for FGAs is in the range of 300 to 1,000 mg daily of chlorpromazine or the equivalent of other antipsychotics. A number of dosage comparison studies have failed to support the routine use of higher doses. That is, when groups of patients are assigned to higher doses such as more than 2,000 mg of chlorpromazine or 40 mg of haloperidol, the rate of improvement and the amount of improvement are no greater than for those assigned to more moderate doses. Many clinicians prescribe doses of SGAs—particularly olanzapine and quetiapine—that are substantially higher than those that were studied in controlled trials. Clinicians are sometimes impressed by individuals who require these higher doses, suggesting that there is a small group of patients who should be treated with high doses. However, most patients who are receiving these high doses are only partial responders to an antipsychotic who have endured dosage increases that were not associated with improvement. The use of drug blood levels and assessment of cytochrome-P450 polymorphism might be appropriate in this context (Table 12.12–1). Nevertheless, studies investigating higher doses in patients selected for not responding to the maximum tested doses are few are far between.

There are also only limited data from controlled trials to assist clinicians in finding the best dose of clozapine. The mean dose of clozapine prescribed differs between Europe and the United States, with Europeans commonly treating with less than 300 mg of clozapine daily and clinicians

in the United States often prescribing 500 mg or more. These experiences support the practice of treating most clozapine patients with doses in the range of 300 to 500 mg daily. However, side effects, particularly sedation and orthostatic hypotension, are often limiting factors that prevent clinicians from reaching a targeted dose. Although some patients have an optimal response to doses between 600 and 900 mg, the risk of seizures increases substantially in this dose range. Studies suggest that patients are more likely to respond to clozapine at plasma levels of 350 ng/mL or higher, suggesting that measuring plasma levels may be useful for poor responders. Therapeutic dose levels of other antipsychotics are under researched.

Large multicenter trials indicate that risperidone is most effective at 4 to 8 mg daily (Table 12.12–1). Higher doses may lead to EPS without an advantage in increased effectiveness. In the United States, the average dose of risperidone prescribed for schizophrenia is slightly more than 4 mg daily. This suggests that a reasonable practice would be to manage patients with schizophrenia with 4 mg of risperidone and increase the dose if they fail to respond after 4 to 6 weeks. Olanzapine is usually effective in the range of 10 to 20 mg daily (Table 12.12–1), although a number of case reports describe individuals who demonstrated optimal responses at doses of 25 mg and higher. Case reports also identify patients who have demonstrated substantial improvements when doses of quetiapine were raised well above 800 mg. At this time, there is no evidence that prescribing doses of ziprasidone greater than 160 mg daily increases its effectiveness (ziprasidone should be taken with at least a 500 kcal meal) (Table 12.12–1). Similarly, there is no evidence that doses of aripiprazole greater than 30 mg daily are more effective than the recommended range of 10 to 30 mg daily (Table 12.12–1). Paliperidone's optimal dose range appears to be 6 to 12 mg daily, with higher doses associated with more adverse effects, but not better effectiveness (Table 12.12–1). Asenapine's optimal dose seems to be between 5 mg bid and 10 mg bid, taken sublingually (avoid eating or drinking for 10 minutes after administration) (Table 12.12–1). Iloperidone's effective dose range is between 6 and 12 mg bid, but titration is needed to avoid orthostatic hypotension (i.e., 1 mg bid on day 1, 2 mg bid on day 2, then increase by 2 mg/day until therapeutic dose is reached) (Table 12.12–1). Finally, lurasidone is effective between 40 and 160 mg (taken with at least a 350 kcal meal), brexpiprazole between 2 and 4 mg, and cariprazine between 1.5 and 6 mg, with higher doses resulting in greater adverse effects but not better efficacy (Table 12.12–1).

A number of recent findings suggest a reasonable strategy for treating acute schizophrenia. The dose of an antipsychotic that is likely to be effective is the dose that occupies approximately 60 to 70 percent of D2 receptors measured in the striatum. The therapeutic response might depend to some extent on processes that occur after these receptors have

been occupied for a period of time. This observation is supported by findings from both positron emission tomography (PET) scanning and the measurement of plasma homovanillic acid, which suggest that clinical improvement is not associated with the immediate effects of the drug on dopamine receptors, but on processes that occur later. At the same time, improvement in psychotic symptoms can begin to occur within the first 24 hours of treatment, particularly with the use of short-acting intramuscular antipsychotics.

Therefore, the goal of the first days of treatment is to administer a drug dose that occupies an adequate proportion of dopamine receptors and to keep the patient comfortable until the drug is increasingly effective. This may involve time-limited cotreatment with a benzodiazepine. If a patient does not adequately respond in the first week or two, this does not necessarily indicate that the current treatment is inadequate. Since most of the improvement on an antipsychotic occurs during the first 2 to 4 weeks, patients should be observed for this interval before changing a drug. Also, the strategy of using medications as needed as a guide to finding the optimal dose makes little sense since the immediate and delayed responses are very different. However, if there is no symptom improvement whatsoever after 2 weeks, despite adherence to a therapeutic dose level, later response is unlikely to occur, and a switch or dose increase within the therapeutic range, should be considered.

Managing Agitation in Acute Psychosis

Agitation in acute schizophrenia can result from disturbing psychotic symptoms, such as frightening delusions or suspiciousness, or from other causes, including drug abuse or withdrawal, or akathisia. Patients with akathisia can appear agitated when they experience a subjective feeling of motor restlessness. Differentiating akathisia from psychotic agitation can be difficult, particularly when patients are incapable of describing their internal experience. In case of agitation a dose increase of the antipsychotic should be beneficial, whereas in the case of akathisia, the symptoms would worsen. Benzodiazepine treatment can improve both agitation and akathisia.

Clinicians have a number of options for managing agitation that results from psychosis. Antipsychotics and benzodiazepines can result in relatively rapid calming when psychotic patients are agitated. An advantage of an antipsychotic is that a single intramuscular injection of haloperidol, fluphenazine, olanzapine, aripiprazole, or ziprasidone will often result in calming without an excess of sedation. Low potency antipsychotics are often associated with sedation and postural hypotension, particularly when they are administered intramuscularly (such as chlorpromazine). Intramuscular ziprasidone, aripiprazole, and olanzapine are similar to their oral counterparts in not causing substantial EPS during acute treatment.

This can be an important advantage over haloperidol or fluphenazine, which can cause frightening dystonia or akathisia in some patients, although coadministration of the anticholinergic benztropine (Cogentin) or the antihistamine promethazine (Phenergan) can reduce the risk of dystonia substantially. Rapidly dissolving oral formulation of olanzapine, risperidone, or aripiprazole may also be helpful as an alternative to an intramuscular injection, at least in order to ensure that the patient has ingested the medication fully.

Benzodiazepines are also effective for agitation during acute psychosis. Lorazepam (Ativan) has the advantage of reliable absorption when it is administered either orally or intramuscularly. The combination of lorazepam with an antipsychotic has been found to be safer and more effective than large doses of antipsychotics in controlling excitement and motor agitation. Moreover, the use of benzodiazepines may reduce the amount of antipsychotic that is needed to control psychotic patients. Due to the sedating properties of short-acting injectable olanzapine that may depress respiratory drive, it should not be combined with benzodiazepines within a 2-hour window prior or after its administration.

MANAGING SIDE EFFECTS

Patients will frequently experience side effects of an antipsychotic before they experience clinical improvement. Whereas a substantial clinical response may be delayed for days or weeks after drugs are started, side effects will often begin almost immediately. This early onset of side effects is important because a patient's interpretation of a drug's effectiveness is often associated with how that drug makes them feel. Moreover, one of the challenges of treating acutely psychotic individuals is maintaining the trust of those who may misinterpret experiences and become suspicious. Warning patients about the potential side effects of medication can lead to prompt management and will often improve the trust between patient and clinician. Moreover, minimizing adverse effects can have long-lasting effects since one of the powerful predictors of drug reluctance or drug refusal is an earlier experience of side effects.

Extrapyramidal Side Effects

One of the most widely accepted benefits of SGAs is their reduced propensity to cause EPS. Some debate has centered around the comparator drugs and dosages used in establishing these benefits, but even relatively low doses of high-potency conventional antipsychotics (e.g., haloperidol 4 mg per day) are associated with more EPS than recommended doses of SGAs. Lower potency FGAs have a lower propensity to cause EPS than higher potency drugs. However, because EPS can occur with any currently available antipsychotic, including SGAs, awareness of EPS remains critical.

A common form of EPS is akathisia, a side effect consisting of a subjective feeling of restlessness with or without restless movements, usually in the legs or feet. Patients who experience severe akathisia will often pace continuously or move their feet restlessly while they are sitting. Some complain that they are unable to feel comfortable, regardless of what they do. Severe akathisia can cause patients to feel anxious or irritable, and some reports suggest that severe akathisia can result in aggressive or suicidal acts. Researchers have estimated that 25 to 75 percent of patients treated with a high-potency FGAs will experience akathisia. This side effect can be difficult to assess and is frequently misdiagnosed as anxiety or agitation. The partial D2 agonists aripiprazole and cariprazine can lead to early restlessness that may not be true akathisia, but rather due to early stimulatory effects that are minimized or controlled by switching slowly from a full D2 blocker or a sedating or anticholinergic antipsychotic to aripiprazole, using a lower starting dose or cotreating temporarily with a benzodiazepine or antihistamine. Often, when the initial restlessness subsides, dose escalation can be tolerated without development of true akathisia. Of note, akathisia rates were significantly reduced with lurasidone despite use of higher doses when the administration was moved from the morning to the evening, suggesting that movement-related side effects that may be peak concentration related can be minimized when administering the antipsychotic in the evening, having the peak level occur during sleep.

Since patients may experience akathisia as irritability or agitation, asking patients whether they are restless or if they have difficulty sitting still can be helpful in early stages of treatment. At this point, a dosage adjustment, a β-blocker, or an anticholinergic antiparkinson drug may provide considerable relief. Notably, coadministration of a benzodiazepine can treat both akathisia and psychotic agitation effectively. Also, patients who have a history of developing severe akathisia that responds poorly to these treatments are likely to do better if they are treated with an SGA.

Dystonias are probably the most frightening type of EPS. They are intermittent or sustained muscular spasms and abnormal postures affecting mainly the musculature of the head and neck, but sometimes the trunk and lower extremities. Common forms of dystonia include abnormal positioning of the neck, impaired swallowing (dysphagia), hypertonic or enlarged tongue, and deviations of the eyes (oculogyric crisis). These reactions usually appear within the first few days of therapy. Dystonias are more likely to occur in younger patients, particularly young males, and respond acutely to intramuscular biperiden.

Antipsychotic-induced parkinsonism consists of tremor, muscular rigidity, and a decrease in spontaneous movements. All of these features resemble the movement disorder in idiopathic parkinsonism. Examination will usually reveal a positive glabella tap. This motor disturbance affects

about 30 percent of patients who are chronically treated with FGAs. The first evidence of drug-induced parkinsonism may be a diminished arm swing or decreased facial expressiveness. Akinesia can be subtle and can be mistaken for or exacerbate negative symptoms. Risk factors for antipsychotic-induced parkinsonism include increasing age and dose, FGA use, a history of parkinsonism, and underlying basal ganglia damage.

When patients develop EPS, clinicians have a number of alternatives. These include reducing the dose of the antipsychotic (which is most commonly an FGA), adding an antiparkinson medication, or changing the patient to an SGA that is less likely to cause EPS. The most effective antiparkinson medications are the anticholinergic antiparkinson drugs (which are described in another section). Although these medications are frequently effective, they also cause their own side effects, including dry mouth, constipation, blurred vision, and often cognitive disturbance. Also, these medications are often only partially effective, leaving patients with substantial amounts of lingering EPS. Conversely, chronic anticholinergic use can often be discontinued in the maintenance treatment phase in patients without enduring EPS, as D2 receptor adaptation may make chronic anticholinergic use unnecessary. Centrally acting β-blockers such as propranolol are frequently effective for treating akathisia. Most patients respond to doses between 30 and 90 mg daily.

If FGAs are being prescribed, clinicians may consider prescribing prophylactic antiparkinson medications for patients who are likely to experience disturbing EPS. These include patients who have a history of EPS sensitivity or those who are being treated with high-potency drugs. Prophylactic antiparkinson medications may also be indicated when high-potency drugs are prescribed for young men who tend to have an increased vulnerability for developing dystonias. These patients may be good candidates for SGAs. However, the need for chronic anticholinergic treatment should be reevaluated as indicated above.

Some individuals are highly sensitive to EPS at dosages necessary to control their psychosis. For many of these patients, medication side effects may seem worse than the illness itself. These patients should routinely be treated with drugs associated with lower incidence of EPS, which means that in general SGAs would be preferred to FGAs. However, highly sensitive individuals may experience EPS on any antipsychotic. Among the drugs with lower risk, clozapine and quetiapine seem to cause the least EPS, while all other antipsychotics may cause dose-related EPS.

Tardive Dyskinesia and Other Tardive Syndromes

As with EPS, TD is less common with the SGAs than with conventional drugs, although there is probably a range of risk within both groups of drugs. Prospective studies lasting 6 months or more are consistent in demonstrating a significantly lower risk of TD with SGAs (i.e., clozapine,

risperidone, olanzapine, quetiapine) than haloperidol, the most common comparator FGA. Fewer data are currently available with the newer SGAs, but early experience certainly suggests a low risk with these drugs as well. These results were somewhat expected in that early occurring EPS are a significant risk factor for TD. However, the risk of TD is not absent with SGAs and it is important for clinicians to be aware of the identification and management of TD even when patients are treated with SGAs.

TDs commonly consist of abnormal, involuntary movements of the mouth, face and tongue, trunk, and extremities. The oral-facial movements occur in about three-fourths of TD patients and can include lip smacking, sucking, and puckering as well as facial grimacing. Other movements might include irregular movements of the limbs, particularly choreoathetoid-like movements of the fingers and toes, and slow, writhing movements of the trunk. Younger patients with TD tend to develop slower athetoid movements of the trunk, extremities, and neck.

The abnormal movements of TD are usually reduced by voluntary movements of the affected areas and are increased by voluntary movements of unaffected areas. The abnormal movements of TD are usually increased with emotional arousal and absent when the individual is asleep. According to the diagnostic criteria in the fifth edition text revision *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), the abnormal movements should be present for at least 4 weeks, and patients should have been exposed to an antipsychotic for at least 3 months (at least one month in the elderly). The onset of the abnormal movements should occur either while the patient is receiving an antipsychotic or within 4 weeks of discontinuing an oral or 8 weeks after the withdrawal of a depot antipsychotic.

Prevalence surveys indicate that 20 to 30 percent of patients who are chronically treated with an FGA will exhibit symptoms of TD. Three to 5 percent of young patients receiving an FGA develop TD each year. The risk in elderly patients is much higher. Although seriously disabling dyskinesia is uncommon, in a small proportion of cases walking, breathing, eating, and talking are affected. Individuals who are more sensitive to acute EPS appear to be more vulnerable to developing TD. Patients with organic mental illness and affective disorders may also be more vulnerable to TD than those with schizophrenia.

A task force on TD of the American Psychiatric Association issued a report in which they made a number of recommendations for preventing and managing TD. These include: (1) establishing objective evidence that antipsychotic medications are effective for an individual; (2) utilizing the lowest effective dose of antipsychotic; (3) prescribing cautiously with children, elderly patients, and patients with mood disorders; (4) examining patients on a regular basis for evidence of TD; (5) when TD is diagnosed, consider alternatives to the antipsychotic being used, obtain informed

consent, and also consider dosage reduction; (6) if the TD worsens, consider a number of options, including discontinuing the antipsychotic or switching to a different drug. Clozapine has been shown to be effective in reducing severe TD or tardive dystonia.

Regular monitoring for TD should be a component of management strategies with antipsychotics. Usually, the initial symptoms are mild. The monitoring should be particularly careful for those patients with an increased risk for TD, including elderly patients, patients who are sensitive to EPS, and individuals with affective illness. Routine monitoring should include examination every 3 to 6 months, and monitoring for high-risk groups should be carried out every 3 months.

Other Side Effects

Sedation and postural hypotension can be important side effects for patients who are being treated with low-potency FGAs, such as chlorpromazine or thioridazine. Sedation can be a problem with clozapine, olanzapine and quetiapine, and postural hypotension with clozapine or iloperidone treatment. These effects are often most severe during the initial dosing with these medications. As a result, patients treated with these medications, particularly clozapine, may require weeks until they reach a therapeutic dose. Although most patients develop tolerance to sedation and postural hypotension, sedation may continue to be a problem. In these patients, daytime drowsiness may interfere with a patient's attempts to return to community life.

All FGAs as well as risperidone and paliperidone elevate prolactin levels, which can result in galactorrhea and irregular menses. There is also concern that long-term elevations in prolactin and the resultant suppression in gonadotropin-releasing hormone can cause clinically important suppression in gonadal hormones. These, in turn, may have effects on libido and sexual functioning. In addition, via the same mechanism, elevated prolactin may cause decreases in bone density and lead to osteoporosis. The concerns about hyperprolactinemia and sexual functioning and bone density are based on experiences with prolactin elevations related to tumors and other causes. It is unclear to what degree these risks are also associated with the lower elevations that occur with prolactin-elevating drugs.

Prior to the introduction of SGAs, elevated prolactin was a consequence of treatment with all antipsychotics. Clozapine and quetiapine do not appear to elevate prolactin above normal levels, while the partial D2 agonists, aripiprazole, brexpiprazole and cariprazine may reduce prolactin levels. As a result, when patients demonstrate side effects related to prolactin, such as galactorrhea, menstrual disturbances or sexual side effects, changing patients to a prolactin-sparing agent may be effective.

Side Effects of Clozapine

Clozapine has a number of side effects that make it a difficult drug to administer. The most serious is a risk of agranulocytosis. This potentially fatal condition occurs in approximately 0.3 percent of patients treated with clozapine during the first year of exposure. Subsequently, the risk is substantially lower. As a result, patients who receive clozapine in the United States are required to be in a program of weekly blood monitoring for the first 6 months, biweekly for the next 6 months, and monthly thereafter.

Clozapine is also associated with a higher risk of seizures than other antipsychotics. The risk reaches nearly 5 percent at doses over 600 mg. Patients who develop seizures with clozapine can usually be managed by reducing the dose and/or adding an anticonvulsant, usually valproate (Depacon). Myocarditis has been reported to occur in approximately 5 patients per 100,000 patient years. Other side effects with clozapine include hypersalivation, sedation, tachycardia, weight gain, fever, and postural hypotension, which generally can be managed.

Metabolic Side Effects of Antipsychotics

Patients with schizophrenia are more likely than the population at large to suffer from a number of illnesses including coronary heart disease, diabetes, and hypertension. The high prevalence of these illnesses may explain why individuals with schizophrenia have a 15 to 20 years shorter life expectancy than the population at large. The risk of these illnesses is partly explained by unhealthy lifestyles of many patients, which may include smoking, poor dietary habits, and lack of exercise.

In addition, there is increasing evidence that antipsychotics can contribute to these problems. Across a large body of evidence, the risk for weight gain and abnormalities in glucose and lipid metabolism is greatest with clozapine, olanzapine, and chlorpromazine. The lowest risk is observed with aripiprazole, brexpiprazole, cariprazine, lurasidone, and ziprasidone, although problems can even occur with these agents, especially in first episode patients and those with minimal antipsychotic exposure. All other antipsychotics have intermediate risk for weight gain and glucose and lipid metabolism with the exception of a higher risk for lipid abnormalities with quetiapine.

Health Monitoring in Patients Receiving Antipsychotics

Routine cardiometabolic monitoring is recommended, at a minimum the assessment of body weight, blood pressure and fasting glucose, and lipids with or without fasting or random hemoglobin A_{1C} at baseline, 3 months and annually after antipsychotic initiation. More frequent assessments are recommended if significant weight gain occurs. When abnormalities are

observed, healthy lifestyle counseling or interventions or switching to a lower risk agent should be considered. Findings from CATIE and other studies indicate that when patients gain weight on an antipsychotic, they may lose substantial weight when they are changed to other antipsychotics that are not associated with severe weight gain. If healthy lifestyle interventions or switching treatments fail or are not feasible, addition of medications that target significant weight gain or abnormal blood pressure, glucose or lipid metabolism may be warranted, and consultation with medical specialists may be needed.

NEGATIVE, MOOD, AND COGNITIVE SYMPTOMS

Negative symptoms and cognitive impairments are associated with a substantial amount of the social and vocational disability in schizophrenia. This observation has resulted in a reappraisal of the goals of treatment, placing a greater emphasis on treatment strategies for decreasing the severity of these impairments. Negative and cognitive symptoms can be classified as being either primary, which is related to the illness, or secondary negative symptoms, which is that they are related to non-illness related factors. Reasons for secondary negative symptoms that can be addressed directly include depression, psychotic symptoms of suspiciousness or paranoia, social anxiety, EPS, sedation, sleep apnea, chronic pain or environmental deprivation. The management of secondary negative symptoms begins with the management of the condition that caused these symptoms. For depression, this may include the addition of an antidepressant medication; for EPS this may involve the addition of an antiparkinson medication, a dose reduction, or a change to an antipsychotic, usually an SGA, associated with less EPS.

If the previously mentioned causes of secondary negative symptoms have been ruled out, the patient is likely to be demonstrating a type of enduring primary negative symptom. There is no strong evidence that any antipsychotic is better at reducing primary negative symptoms than any other. Although some studies suggested that the SGAs are more effective in reducing negative symptoms than FGAs, the advantage may be only in reducing secondary negative symptoms that are due to EPS. Meta-analyses have shown no significant advantages of SGAs over haloperidol in reduction of negative symptoms. However, until this issue is decided by adequate controlled studies, it is probably reasonable for clinicians to consider changing to an SGA for patients with substantial negative symptoms in an attempt to minimize any that are secondary to EPS. Several meta-analyses have indicated that antidepressant augmentation of antipsychotics may improve negative symptoms, but whether the improvement pertains predominantly to primary or secondary negative symptoms remains unclear.

Patients with schizophrenia frequently suffer from impairments in

memory, attention, and information processing. These cognitive impairments can also interfere with the social and vocational rehabilitation of patients, even when their psychotic symptoms have been well controlled. As with negative symptoms, cognitive impairments can also be secondary to other causes, including depression, psychotic symptoms of thought disorganization or hallucinations, substance abuse, side effects of antipsychotics, such as EPS, sedation, side effects of anticholinergic antiparkinson medications used to treat EPS, or sleep apnea. Addressing these factors, including decreasing the use of anticholinergic medication by changing to drugs that do not require antiparkinson medications, particularly SGAs, may be helpful. However, cognition was carefully monitored in the CATIE study and did not differ meaningfully among the FGA, perphenazine, and the SGAs.

SUICIDAL BEHAVIOR

Patients with schizophrenia and schizoaffective disorder are at considerable risk for suicidal behavior. Approximately 20 to 40 percent of patients make suicide attempts and 5 to 10 percent succeed. Suicidal behavior appears to be an independent domain from psychosis; however, depression and comorbid substance abuse increase the risk. Based on a large-scale study of clozapine (vs. olanzapine) in patients at risk for suicide, clozapine received an indication for the prevention of suicidal behavior. This suggests that people with schizophrenia who are experiencing unremitting suicidal thoughts should be considered as candidates for clozapine.

STRATEGIES FOR POOR RESPONDERS

When patients with acute schizophrenia are administered an antipsychotic medication, approximately 50 percent will improve to the extent that they will achieve a complete remission of positive symptoms or experience only mild symptoms. The remaining 50 percent of patients will improve, but will still demonstrate variable levels of positive symptoms that are resistant to the medications. Rather than categorizing patients into responders and nonresponders, it is more accurate to consider the degree to which the illness is improved by medication. Some patients are so severely ill that they require chronic institutionalization. Others will respond to an antipsychotic with substantial suppression of their psychotic symptoms, but will demonstrate persistent positive symptoms, such as hallucinations or delusions.

Before deciding that a particular antipsychotic is not working adequately for a patient, it is important to ensure that the diagnosis is correct, to rule out substance related or medical reasons, and ensure that the patient has been adherent to the medication regimen and that they received an adequate trial of the medication. A 6-week trial on an adequate

dose of an antipsychotic represents a reasonable trial for most patients. If patients demonstrate even a mild amount of improvement during this period, it may be reasonable to wait before changing a medication, since data from groups of patients indicate that patients may improve at a steady rate for 3 to 6 months. However, early lack of at least minimal improvement (i.e., within the first 2 weeks) seems to identify a subgroup of patients who achieve a lower level of response even after 3 months of subsequent treatment. It may also be helpful to confirm that the patient is receiving an adequate amount of the drug by monitoring the plasma concentration (if the patient is receiving a drug where there are adequate data to define a therapeutic level). This information is available for a number of antipsychotics including haloperidol, clozapine, fluphenazine, trifluoperazine (Stelazine), and perphenazine. A very low plasma concentration may indicate that the patient has been noncompliant or, more commonly, only partially compliant. It may also suggest that the patient is a rapid metabolizer of the antipsychotic or that the drug is not being adequately absorbed. Tests are also available to assess polymorphisms of the cytochrome-P450 enzyme system to identify atypical metabolizer profiles. Under these conditions raising the dose may be helpful. If the level is relatively high, clinicians should consider whether side effects may be interfering with therapeutic response.

If the patient is responding poorly, many clinicians will consider raising the dose above the usual therapeutic level. The use of high doses in poor medication responders has been studied under a number of circumstances. Nearly all studies found that higher doses were not associated with greater improvement than conventional doses. This suggests that changing to another drug is more likely to be helpful than changing to a high dose since many adverse effects may be dose related.

If a patient has responded poorly to an FGA, it is unlikely that this individual will do well on another FGA. Studies suggest that a poor response to one FGA is likely to be followed by a poor response to another. This has led many clinicians to change to an SGA in the hope that a somewhat different profile of receptor interactions may lead to better response.

There is substantial evidence indicating that clozapine is effective for patients who respond poorly to other antipsychotics. The most definitive evidence of clozapine's advantages in this population comes from a multicenter trial reported by Kane et al. in 1988 in which clozapine was compared with chlorpromazine. Patients in this study were a severely psychotic group of individuals who had failed in trials with at least three antipsychotics. Clozapine was significantly more effective than chlorpromazine in nearly every dimension of psychopathology, including positive and negative symptoms. This study found that 30 percent of patients treated with clozapine met improvement criteria by the end of the

6-week trial compared to only 4 percent of patients on chlorpromazine. Studies with a longer duration indicate that as many as 60 percent of patients are likely to meet these same improvement criteria when patients are maintained on clozapine for 6 months.

There is also evidence suggesting that risperidone and olanzapine may be helpful when an FGA is only partially effective. At least one controlled trial with each of these drugs has shown superiority to FGAs in poor or partial responders, or "noninferiority" to clozapine. However, other studies suggest clozapine to be superior. The CATIE and CUtLASS trials also confirmed clozapine's superiority when FGAs or SGAs have failed. A number of meta-analyses have been conducted supporting clozapine's superiority in patients with refractory symptoms, especially when clozapine was adequately dosed (>400 mg).Given clozapine's side-effects profile, a case can be made for the practice of first trying patients on risperidone or olanzapine when they have responded poorly to an FGA. However, if patients fail to respond adequately, a trial of clozapine is clearly warranted. Whether patients failing SGAs should have at least one trial of an FGA prior to clozapine is unclear, but since partial and covert nonadherence is a common cause of treatment "resistance," a trial with a long-acting injectable antipsychotic may be considered prior to switching a patient to clozapine.

In recent years, there has been an increase in the use of antipsychotic polypharmacy. Although there are a variety of factors contributing to such practice, poor initial response to monotherapy is often cited as a reason, but patients also end up on two or more antipsychotics if they start improving during a switch that is then aborted. Unfortunately, there is no high-quality evidence supporting antipsychotic polypharmacy for treatment-resistant patients, except possibly for augmenting dopamine antagonists with a partial D2 agonist to improve negative symptoms. However, whether this finding reflects improvement of primary or secondary negative symptom is unclear.

Other "augmentation" strategies for the treatment of positive symptoms with benzodiazepines, lithium, antiepileptics, and β-blockers have been studied to some extent and controlled trials and meta-analyses are generally negative. These findings do not rule out the possibility of individual benefit for some patients; however, the data supporting the potential value of clozapine are far more consistent. The use of adjunctive medications to treat symptoms of affective disturbance, anxiety, agitation, and so forth may be more appropriate.

ECT can also be considered as a treatment of last resort for refractory individuals, even in those not adequately responding to clozapine.

MAINTENANCE THERAPY

During the stable or maintenance phase, patients are often in a relative

state of remission with only minimal positive psychotic symptoms or at a stable plateau with moderate symptoms. The goals during this stage are to prevent patients from suffering psychotic relapse or exacerbation and to assist patients in improving their level of functioning. Pharmacotherapy plays an important part in both of these goals. Medications are effective in preventing or delaying psychotic relapse and they may also be an important adjunct in managing functional impairments that may interfere with psychosocial rehabilitation. The art of maintenance treatment results from the unfortunate problem that medication side effects or lack of patient's insight, cognitive or organizational abilities, and motivation to continue long-term treatment can sometimes interfere with these goals. In addition, titrating dosage in maintenance treatment can be very difficult, because the relapse or exacerbation that can result from too low a dose is not likely to occur for weeks to months.

Drug and Route of Administration for Maintenance Therapy

Stable patients who are maintained on an antipsychotic have a much lower relapse rate than patients who have their medications discontinued. Although studies differ, pooling large amounts of data suggests that 15 to 25 percent in a year will experience a relapse while receiving medications and 50 to 75 percent will relapse without medications. Clinicians are often tempted to discontinue medications in patients who have been well and stable for prolonged periods of time. Unfortunately, these patients also have high relapse rates when their medications are discontinued. Other evidence indicates that patients who experience relapses while they are receiving an antipsychotic have milder episodes than patients who relapse on no medication. It has also been suggested that patients who had their medications discontinued are more likely to show dangerous behavior and to be admitted involuntarily. Moreover, a subgroup of up to one in six patients may not respond to the next antipsychotic trial as well anymore as they did before the relapse.

As previously discussed, TD has been a major concern in establishing the benefit-to-risk ratio of long-term treatment. Even with FGAs, experts agreed that the benefits of maintenance treatment outweighed the risks. As SGAs have been introduced with a substantively reduced risk of TD, one of the major concerns about long-term treatment has been diminished.

The evaluation of benefit-to-risk has been particularly challenging in those patients who have experienced only one psychotic episode and have responded well to treatment. It is important to recognize that although many such patients achieve remission, recovery (which includes sustained, relatively normal social and vocational adjustment) is far less common.

It has been well established that even patients who have had only one episode have a four in five chance of relapsing at least once over the next 5 years, and that stopping medication is the most significant risk factor. At

the same time, it is often difficult for late adolescents or young adults to accept the nature of their illness and the need for ongoing pharmacotherapy. This represents a psychotherapeutic and psychoeducational challenge to clinicians and underscores the importance of combining psychosocial treatments, family therapy, vocational rehabilitation, and pharmacological management.

Although published guidelines do not make definitive recommendations about the duration of maintenance treatment following the first episode, recent data suggest that 1 or 2 years might not be adequate. This is a particular concern when patients have achieved good employment status or are involved in educational programs, as they have an enormous amount to lose if they experience another psychotic episode.

It is generally recommended that multiepisode patients receive maintenance treatment for at least 5 years. However, this recommendation is based on controlled data until this time and the implication that discontinuing medication at that point is not without substantial risk is a mistake. Therefore, many experts would recommend pharmacotherapy on an “indefinite” or “for the foreseeable future” basis.

The first 3 to 6 months following an acute episode or relapse is a period of particular vulnerability. With short lengths of hospital stay, adequate linkages with ambulatory programs are critical to ensure continuity of care. After stabilization for 6 months, gradual dosage reduction may be attempted, but its need is unclear if there are no limiting adverse effects that are dose related. Moreover, there are very few studies that help define minimum maintenance dose for SGAs. Given reduced concern regarding TD and the lack of a strong dose–response curve for weight gain, one could argue that there is less incentive to define the lowest effective dose for maintenance treatment with SGAs.

There are some data, including a large meta-analysis, suggesting that the SGAs are more effective in preventing relapse than FGAs. The authors’ impression at present is that this advantage is not due to enhanced compliance. Although there are data suggesting some improvement in adherence with SGAs, these differences are modest at best.

Nonadherence rates with long-term antipsychotic treatment are very high. Average estimates suggest that more than 40 to 50 percent of patients become at least partially noncompliant within 1 or 2 years.

Given the high rates of relapse following discontinuation of medication and the potentially severe consequences (loss of job, interference with school, family burden, suicidality, homelessness, aggressive or violent behavior), attempts to enhance adherence are critical. Although psychosocial treatments focusing on compliance can be helpful, the use of long-acting injectable antipsychotics should be considered as a preventive measure and not just reserved for patients who have repeatedly experienced noncompliance and consequent relapse.

Given these data, long-acting injectable antipsychotics are a valuable treatment option that currently appears to be underutilized. Although in meta-analyses of randomized controlled trials, it has been difficult to demonstrate consistent superiority over oral antipsychotic treatment, this is likely due to the fact that highly selected, relatively compliant patients are included and that the careful monitoring entailed is not representative of routine clinical care. In contrast, almost all mirror image studies have shown that in the patients eligible for long-acting injectable treatment, these are associated with significantly less hospitalizations than when the patient was prescribed oral antipsychotics.

There are a number of potential advantages of long-acting injectable medication. First, clinicians know immediately when nonadherence occurs and clinicians have some time to initiate appropriate interventions before the blood levels become nontherapeutic. Second, there is less day-to-day variability in blood levels and peak concentrations are lower, potentially making it easier to establish minimum effective doses. Third, many patients who have had experience with such treatment often prefer it.

INTEGRATING PHARMACOTHERAPY AND PSYCHOSOCIAL TREATMENT

Most patients with schizophrenia will benefit from a combination of pharmacotherapy and psychosocial treatments. Recent improvements in both domains suggest that the overall outcome of this disorder can be improved if patients receive the optimal forms of both treatments at the appropriate stage of their illness. Both studies and clinical experience suggest that psychosocial treatments are probably most effective when patients have recovered from severe psychotic episodes. During the acute psychotic phase, clinical management should emphasize maintaining cooperativeness and trust. This focus is particularly important when there is overt suspiciousness or a tendency to misinterpret the intentions of the treatment team. A successful strategy is likely to include clear explanations of the rationale for treatment and possible antipsychotic side effects. Since family members may be important allies in ensuring cooperativeness, family psychoeducation programs have been demonstrated to be helpful during this phase.

It is difficult to generalize about the interactions of antipsychotics and psychosocial treatments for stable patients since psychosocial treatments can be quite different in their content and their goals. Nevertheless, a number of important treatment principles can be drawn from the literature on combining treatments. The first is that psychosocial treatments are most likely to be effective when patients have been adequately stabilized on antipsychotics. Early studies indicated that psychosocial treatments could actually lead to a worse outcome when outpatients with schizophrenia were treated with a placebo. Other studies indicate that patients are most likely to respond to psychosocial treatments when their condition is stable. For

example, a study with social skills training found that patients who received a type of pharmacotherapy that minimized the proportion of time that they were in a psychotic state also demonstrated the greatest improvements in social adjustment.

Psychosocial treatments may also improve the response to pharmacotherapy by improving medication adherence. This effect was suggested in a study in which patients received a form of family treatment that also encouraged medication adherence. In addition, specific adherence-focused group sessions have been shown to be helpful. Other studies have indicated that psychosocial treatments, particularly family treatment, may decrease the amount of stress that the patient experiences within the family, and that this, in turn, decreases the amount of antipsychotic medication required by the patient.

DRUG DEVELOPMENT FOR SCHIZOPHRENIA

Studies, such as CATIE and CUtLASS, indicate that both FGAs and SGAs have limits to their effectiveness. Although these agents are effective for reducing the severity of positive symptoms in schizophrenia, their effects on cognitive and negative symptoms are often inadequate. These domains are important since the long-term functional outcome—or the ability of patients to function in work, school, and social relationships—is related to the severity of these symptoms. This shortcoming has led to a search for medications that have specific effects on negative symptoms and cognitive impairments.

Considerable attention has focused on cognitive impairments as a target for drug development in schizophrenia. Research on the nature of these impairments and the neurobiology of cognition suggested that this was a promising target. However, there were several obstacles that discouraged development in this area. These included the lack of a consensus on how cognition should be measured in clinical trials, concerns as to whether the FDA and agencies in other countries would approve drugs for this indication, a lack of guidance as to how clinical trials of cognition enhancing drugs should be designed, and a lack of a consensus regarding the most promising molecular targets for drug development. All these issues were addressed by an NIMH initiative known as MATRICS (Measurement and Treatment Research to Improve Cognition in Schizophrenia). Through a series of consensus development activities a battery known as the MATRICS Consensus Cognitive Battery has been developed and is currently in widespread use in clinical trials. The FDA has clarified that the agency will approve drugs for treating the cognitive impairments associated with schizophrenia if the studies meet guidelines developed in an NIMH-FDA consensus meeting, including the requirement for a functional coprimary outcome measure. The same is true for agencies in other parts of the world. Although it has remained problematic to find

effective treatments for cognitive dysfunction in schizophrenia, a number of drugs addressing varied molecular targets are currently in different stages of development.

Other activity has focused on drugs that may improve negative symptoms in schizophrenia. Here, again, consensus on the targets, designs, and measurement tools has prepared the ground for large development programs. However, like in the area of cognitive dysfunction, despite many attempts and promise in the area of proglutamatergic mechanisms, to date, no drugs have been approved for the treatment of negative symptoms in schizophrenia. A number of agents are currently in clinical trials.

If studies of specific drugs for negative symptoms and cognitive impairments are successful, treatment of schizophrenia in the future may involve managing patients with an antipsychotic drug and comedications for these other symptom dimensions. It is hoped that this will significantly improve functional outcomes. Whether any of these or other drug mechanisms will also enhance efficacy for positive symptoms remains to be seen, but additional mechanisms of action for the treatment of all domains of schizophrenia are highly desirable.

References

Buchanan RW, Kreyenbuhl J, Kelly DL, et al. Schizophrenia Patient Outcomes Research Team (PORT). The 2009 schizophrenia PORT psychopharmacological treatment recommendations and summary statements. *Schizophr Bull*. 2010;36(1):71–93.

Citrome L. New second-generation long-acting injectable antipsychotics for the treatment of schizophrenia. *Expert Rev Neurother*. 2013;13(7):767–783.

Correll CU, Citrome L, Haddad PM, et al. The use of long-acting injectable antipsychotics in schizophrenia: evaluating the evidence. *J Clin Psychiatry*. 2016;77(suppl 3):1–24.

Correll CU, Leucht S, Kane JM. Lower risk for tardive dyskinesia associated with second-generation antipsychotics: a systematic review of one-year studies. *Am J Psychiatry*. 2004;161(3):414–425.

De Hert M, Detraux J, van Winkel R, Yu W, Correll CU. Metabolic and cardiovascular adverse effects associated with antipsychotic drugs. *Nat Rev Endocrinol*. 2011;8(2):114–126.

Dixon LB, Dickerson F, Bellack AS, et al. Schizophrenia Patient Outcomes Research Team (PORT). The 2009 schizophrenia PORT psychosocial treatment recommendations and summary statements. *Schizophr Bull*. 2010;36(1):48–70.

Dold M, Fugger G, Aigner M, Lanzenberger R, Kasper S. Dose escalation of antipsychotic drugs in schizophrenia: a meta-analysis of randomized controlled trials. *Schizophr Res*. 2015;166(1-3):187–193.

Essock SM, Covell NH, Davis SM, Stroup TS, Rosenheck RA, Lieberman JA. Effectiveness of switching antipsychotic medications. *Am J Psychiatry*. 2006;163(12):2090–2095.

Fusar-Poli P, Papanastasiou E, Stahl D, et al. Treatments of negative symptoms in schizophrenia: meta-analysis of 168 randomized placebo-controlled Trials. *Schizophr Bull*. 2015;41(4):892–899.

Galling B, Roldán A, Hagi K, et al. Antipsychotic augmentation vs. monotherapy in schizophrenia: systematic review, meta-analysis and meta-regression analysis. *World Psychiatry*. (in press)

Garay RP, Citrome L, Samalin L, et al. Therapeutic improvements expected in the near future for schizophrenia and schizoaffective disorder: an appraisal of phase III clinical trials of schizophrenia-targeted therapies as found in US and EU clinical trial registries. *Expert Opin Pharmacother*. 2016;17(7):921–936.

Helfer B, Samara MT, Huhn M, et al. Efficacy and safety of antidepressants added to antipsychotics for schizophrenia: a systematic review and meta-a. *Am J Psychiatry*. 20161;173(9):876–886.

Jones PB, Barnes TR, Davies L, et al. Randomized controlled trial of the effect on Quality of Life of second- vs first-generation antipsychotic drugs in schizophrenia: Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1). *Arch Gen Psychiatry*. 2006;63(10):1079–1087.

Kahn RS, Fleischhacker WW, Boter H, et al. Effectiveness of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: an open randomized clinical trial. *Lancet*. 2008;371:1085–1097.

Kahn RS, Sommer IE, Murray RM, et al. Schizophrenia. *Nat Revi Dis Primers*. 2015;1:15067.

Kane J, Kishimoto K, Correll CU. Evaluation and management of treatment-resistant schizophrenia. http://www.uptodate.com/contents/evaluation-and-management-of-treatment-resistant-schizophrenia. Accessed January 29, 2015.

Kane JM, Correll CU. Past and present progress in the pharmacologic treatment of schizophrenia. *J Clin Psychiatry*. 2010;79(9):1115–1124.

Kane J, Honigfeld G, Singer J, Meltzer H. Clozapine for the treatment-resistant schizophrenic. A double-blind comparison with chlorpromazine. *Arch Gen Psychiatry*. 1988;45(9):789–796.

Kane JM, Robinson DG, Schooler NR, et al. Comprehensive versus usual community care for first-episode psychosis: 2-Year outcomes from the NIMH RAISE early treatment program. *Am J Psychiatry*. 2016;173(4):362–372.

Kishimoto T, Agarwal V, Kishi T, Leucht S, Kane JM, Correll CU. Relapse prevention in schizophrenia: a systematic review and meta-analysis of second-generation antipsychotics versus first-generation antipsychotics. *Mol Psychiatry*. 2013;18(1):53–66.

Leucht S, Cipriani A, Spineli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*. 2013;382(9896):951–962.

Leucht S, Tardy M, Komossa K, et al. Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. *Lancet*. 2012;379(9831):2063–2071.

Lieberman JA, Stroup TS, McEvoy JP, et al. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med*. 2005;353(12):1209–1223.

Lopez LV, Kane JM. Plasma levels of second-generation antipsychotics and clinical response in acute psychosis: a review of the literature. *Schizophr Res*. 2013;147(2–3):368–374.

McEvoy JP, Lieberman JA, Stroup TS, et al. CATIE Investigators. Effectiveness of clozapine versus olanzapine, quetiapine, and risperidone in patients with chronic schizophrenia who did not respond to prior atypical antipsychotic treatment. *Am J Psychiatry*. 2006;163(4):600–610.

Millan MJ, Agid Y, Brüne M, et al. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. *Nat Rev Drug Discov*. 2012;11(2):141–168.

Millan MJ, Fone K, Steckler T, Horan WP. Negative symptoms of schizophrenia: clinical characteristics, pathophysiological substrates, experimental models and prospects for improved treatment. *Eur Neuropsychopharmacol*. 2014;24(5):645–692.

Mizuno Y, Suzuki T, Nakagawa A, et al. Pharmacological strategies to counteract antipsychotic-induced weight gain and metabolic adverse effects in schizophrenia: a systematic review and meta-analysis. *Schizophr Bull*. 2014;40(6):1385–1403.

Nielsen J, Correll CU, Manu P, Kane JM. Termination of clozapine treatment due to medical reasons: when is it warranted and how can it be avoided? *J Clin Psychiatry*. 2013;74(6):603–613.

Penttilä M, Jääskeläinen E, Hirvonen N, Isohanni M, Miettunen J. Duration of untreated psychosis as predictor of long-term outcome in schizophrenia: systematic review and meta-analysis. *Br J Psychiatry*. 2014;205(2):88–94.

Pharoah F, Mari J, Rathbone J, Wong W. Family intervention for schizophrenia. Cochrane *Database Syst Rev*. 2010;(12):CD000088.

Samara MT, Dold M, Gianatsi M, et al. Efficacy, acceptability, and tolerability of antipsychotics in treatment-resistant schizophrenia: a network meta-analysis. *JAMA Psychiatry*. 2016;73(3):199–210.

Samara MT, Leucht C, Leeflang MM, et al. Early improvement as a predictor of later response to antipsychotics in schizophrenia: a diagnostic test review. *Am J Psychiatry*. 2015;172(7):617–629.

▲ 12.13 Psychiatric Rehabilitation

ROBERT E. DRAKE, M.D., PH.D., GARY R. BOND, PH.D., AND DROR BEN-ZEEV, PH.D.

INTRODUCTION

Approximately 5 percent of adults in the United States experience disability due to psychiatric illness. Psychiatric rehabilitation denotes a wide range of interventions to help people with disabilities due to mental illness improve their functioning, quality of life, and recovery by enhancing the opportunities, skills, and supports they need to succeed in regular adult roles and in the environments of their choice. Regular adult roles include living independently, working, relating to family, attending school, having friends, pursuing leisure activities, and having intimate relationships. Psychiatric rehabilitation emphasizes independence, self-management, self-reliance, and community integration by avoiding dependence on professionals and segregated settings.

Psychiatric rehabilitation has become a fundamental component of the treatment for schizophrenia, consistent with current knowledge regarding the etiology, pathophysiology, manifestations, treatment, and course of the disorder. Current evidence, reviewed in other chapters, indicates that

schizophrenia is almost certainly a heterogeneous set of disorders, causally related to genetic factors, prenatal and perinatal insults, and postnatal stress. Although the definitive neurobiological substrate of the disorder remains to be defined, many patients manifest mild neurodevelopmental abnormalities related to attention, cognition, social function, and motor function in childhood, long before the onset of overt psychotic illness. Research on schizophrenia shows that many but not all patients exhibit abnormalities of neural structure, circuitry, and function. Clinically, the disorder is characterized by positive symptoms of psychosis, such as hallucinations and delusions; negative symptoms, such as lack of motivation and poverty of speech; cognitive deficits, such as problems with memory and executive functioning; and psychosocial difficulties, such as poor social role performance, chronic unemployment, high rates of substance abuse, and increased risk for homelessness, victimization, and other problems. Long-term follow-up studies have shown over many years that about 10 to 15 percent of schizophrenia patients recover completely, most others improve (Bleuler termed the usual course "healing with scarring"), but many continue to be plagued by symptoms, cognitive difficulties, and psychosocial problems for decades.

Antipsychotic medications often reduce positive symptoms of psychosis, but they have limited impact on negative symptoms, cognitive impairment, and psychosocial functioning. For the great majority of patients, medications do not restore premorbid levels of functioning and do not produce normative role performance. Some 20 percent or more of schizophrenia patients have psychotic symptoms that do not respond to antipsychotic medications, and many experience only a partial symptomatic response. Further, medications do not reverse the consequences of stigma, impaired learning, failure to master adult developmental tasks, and social withdrawal. Therefore, maximizing functional recovery through psychiatric rehabilitation is critical for most patients.

Current interventions for schizophrenia are multifaceted. Historically, rehabilitation specialists distinguished rehabilitation from treatment by its focus on improving functional performance rather than controlling symptoms or illness. Rehabilitation specialists also argued that they focused on the patient's strengths rather than deficits. In practice, these distinctions are not valid. Many cognitive behavioral, skills training, and technology interventions aim to improve management of symptoms as well as functional status, and most interventions take strengths and deficits into account. In this book, several interventions for schizophrenia could be classified as treatment or rehabilitation, and the authors have divided up a selection of current interventions for the sake of illustration. Readers should consult both chapters.

This chapter illustrates several approaches to psychiatric rehabilitation

with specific models and examples. Presenting the entire field is beyond the scope of this chapter, but interested readers should consult textbooks listed at the end of the chapter.

HISTORY

Current approaches to psychiatric rehabilitation represent the confluence of several historical traditions. In the 19th century, the reform movement called moral treatment emphasized putting people with mental illness in benign environments and normal roles, though usually in institutions. Moral treatment recognized that participation in education, work, social activities, and other normal roles could have a healthful effect.

In the United States, government programs for employment of people with disabilities began after World War I for persons with physical disabilities and were subsequently extended to those with psychiatric disabilities in 1943. These programs transferred ideas from physical rehabilitation to psychiatric rehabilitation and emphasized employment.

Deinstitutionalization and the development of community mental health centers in the 1960s and 1970s not only transferred the locus of care for people with schizophrenia to the community, but also promoted interest in making services accessible, comprehensive, and focused on helping people to function in regular environments and roles. With leadership from the National Institute of Mental Health and the Center for Mental Health Services of the Substance Abuse and Mental Health Services Administration, the community mental health ideology evolved into the community support system model, which included approaches such as case management, assertive outreach, coordination, integration, continuity, advocacy, self-help, supported employment, supported education, supported housing, self-management, and an emphasis on personal goals, termed recovery. The community support system model enlarged the role of community mental health and rehabilitation beyond reversing the disabling effects of institutionalization to include prevention of disability among young adults with severe psychiatric illness who had minimal experience in hospitals.

In the early years of deinstitutionalization, patients and nonprofessionals came together to form psychosocial clubs, such as Fountain House in New York and Thresholds in Chicago, to promote mutual aid and self-help. The early self-help clubs transformed into a range of psychosocial rehabilitation centers around the country. Despite great diversity, these centers consistently emphasized mutual support, strategies for coping, health promotion rather than a focus on illness, normative roles such as worker, and belief in the potential for productivity. Current ideas regarding psychiatric rehabilitation emerged directly from these activities.

Developments in the fields of skills training and environmental modification have also stimulated modern approaches to rehabilitation.

Skills training methods derived from social learning principles have been applied to numerous areas of everyday functioning: for example, daily living tasks, social activities, employment, education, housing, relating to the health care system, and managing one's illnesses. Approaches to making environments more accommodating are considerably more diverse. They range from collaborative work with families, landlords, and employers, to a wide range of tactics used in the community to remove barriers faced by people with mental illness related to societal stigma, legal constraints, economic disincentives, and rigidities in the health care system.

The consumer-initiated movement called "recovery" became an influential ideology for advocates and mental health care systems during the 1990s. Recovery has dozens of definitions, but the ideology includes hope, self-management, pursuing personal goals, finding meaning in life, and participating fully in the community. Since rehabilitation has always been patient-centered, recovery goals for most patients are isomorphic with rehabilitation goals.

Most recently, health information technology has been expanding rapidly to enhance self-management and rehabilitation goals. Web-based and mobile technology tools give people with mental disorders and their families immediate access to the latest information, local resources, research, and rehabilitation tools. Increasingly, these tools are shifting control of decision-making and management from professionals to patients and families.

OVERVIEW OF TECHNIQUES

Psychiatric rehabilitation uses three basic approaches: creating opportunities, providing supports, and increasing skills. In practice, these three approaches are combined in nearly all models of psychiatric rehabilitation.

People with mental disorders have always been discriminated against, stigmatized, and treated unfairly in nearly every society, even compared to people with other disabilities. In the past, long-term institutionalization was the primary mechanism; today segregation and lack of opportunity persist in the community. Sensationalism in the media constantly reinforces stigma. Psychiatrists and other professionals often convey stigmatizing attitudes and beliefs. Psychiatric rehabilitation aims to increase opportunities for normative housing, education, employment, socialization, leisure, and other activities that represent full citizenship under the rubric of community integration. To increase opportunities, rehabilitation specialists use legal mechanisms, regulations, community action, social networks, family supports, and personal relationships.

The second basic approach to psychiatric rehabilitation is to provide intensive supports. People with mental illness often need extra supports to

succeed initially in functional roles. Once achieving success in a specific setting, such as school or work, supports are typically withdrawn substantially or completely, as the experience itself increases self-confidence, functional skills, and natural supports.

Increasing personal skills for illness-management and social functioning is the third basic technique of psychiatric rehabilitation. Due to cognitive, experiential, and neurological difficulties, people with schizophrenia often need to enhance their abilities to manage their internal experiences, social environments, and roles they are pursuing in the community. Skills training has been a core feature of rehabilitation for decades. As illustrated below, more of the learning is now shifting from professionals to technology tools.

We illustrate how opportunity, support, and skills combine by describing three current approaches: Housing First, Individual Placement and Support, and technology tools, such as FOCUS. These specific examples cover only a small part of psychiatric rehabilitation but do demonstrate the enormous potential to improve life for people disabled by schizophrenia.

HOUSING FIRST

Homelessness and Housing

Mental health researchers have often described homelessness as a psychiatric problem, suggesting that people become homeless, repeatedly hospitalized, and incarcerated because of severe mental illness, poor treatment, or a combination of both. By contrast, homelessness researchers from other backgrounds, such as anthropology, epidemiology, and housing, consider homelessness to be primarily due to a lack of affordable and appropriate housing. Research shows that people with the most severe mental disorders and social disadvantage, for example, schizophrenia complicated by poverty and drug addiction, need safe housing, supports, and treatment.

The more difficult questions relate to type of housing and timing: Do some people need to be in supervised, congregate housing or is that a prejudice of mental health professionals in the deinstitutionalization era? If many people are eligible for independent housing, do they need to be stable and engaged in treatment to be housed? Alternatively, can they succeed in independent housing before they are engaged and committed to treatment? Policy makers, researchers, clinicians, patients, and advocates have debated these issues for at least four decades.

Recent research on Housing First, a well-defined housing model for homeless people with mental illness and/or substance use disorders, has had a remarkable impact on the psychiatric rehabilitation field.

Housing Models

Traditional approaches to housing people with schizophrenia and other serious mental illness used segregation and/or linkage with required clinical services. Segregated housing included nursing homes, group homes, family care, boarding homes, and housing units linked by proximity and closely overseen by mental health workers. All of these models assumed that people with serious mental disorders needed constant supervision, which was most easily provided by grouping them together, and were not competent to live on their own. Sometimes people were placed in independent apartments, often owned by the mental health system, because they agreed to participate completely in mental health treatment.

In the early 1980s, research clinicians developed assertive community treatment for the express purpose of helping people to leave the hospital and maintain tenure in the community. Assertive community treatment initially involved moving a multidisciplinary team out of the hospital with the patients to provide comprehensive services 24 hours a day, 7 days a week. The team helped people to find housing, helped them learn independent living skills, and provided intensive supports during crises. For example, team members would help people learn to shop in their neighborhoods, use public transportation, prepare meals in their apartments, take their medications, and find activities and friends. But the care was still centered on mental health treatment.

Encouraged by the success of assertive community treatment and as a reaction against the paternalistic psychiatric practices common in that era, advocates for community integration developed a set of housing principles that they labeled "supported housing." Core principles included normalized permanent housing (regular tenancy agreements, assurance of privacy, location in integrated communities), client choice in type and location of residence, assistance in finding affordable housing, individualized and flexible support, separation of housing and services. As supported housing proponents noted, most people preferred living in their own homes, not in supervised congregate housing.

Housing First has emerged as the most prominent model of supported housing. A systematic approach to reversing homelessness, it combines two primary interventions: offering homeless people affordable housing of their choice as rapidly as possible and providing supportive services to help them succeed in housing. To ensure that residents can afford housing, some Housing First programs provide housing vouchers or subsidies. The approach does not require that people participate in treatment or any mental health program as a precondition for obtaining housing. Instead, Housing First assumes that people need housing as a foundation from which they can pursue other goals.

Typically, Housing First integrates housing assistance with assertive community treatment, intensive case management, or some other team-based approach to comprehensive services. The same team assists people in finding affordable housing and also helps them learn to live independently and access other services as they express additional needs. Services are not mandatory, coercive, or chosen by professionals: tenants learn about services and make informed decisions, using a shared decision-making model.

James was a 50-year-old man who had been in and out of hospitals, homeless shelters, jails, and mental health centers for over 20 years. He often lived on the streets. Delusions, auditory hallucinations, and paranoia, all responding minimally to medications, accompanied him constantly. Regular marijuana use exacerbated his symptoms and got him entangled with drug dealers and the criminal justice system. James dressed eccentrically, wore a patch over one eye, and preached his religious beliefs in public settings, which occasioned rejections, confrontations with police, and trips to the emergency room. He seemed not to care about housing, failing to stay in several nursing homes, group homes, apartments, and shelters for more than a few weeks.

After a lengthy forensic hospitalization, James came to a Housing First team. Surprised that he was offered an apartment without preconditions, he readily agreed to move into his own place. The team helped him with furnishings, his landlord, shopping, and learning the new neighborhood.

James experienced many challenges, including disagreements with his landlord, restaurant managers, other tenants, drug dealers, and the police. But the team was available to intercede 24 hours a day, 7 days a week, and helped him to resolve many difficult situations. Listening respectfully to delusions and helping with practical matters was the central strategy. For example, James needed assistance to keep other drug users from moving into his apartment. He understood that to keep the apartment, which he liked, he needed to allow the team and his landlord to set limits with these acquaintances. Over time, James found delusional rationales for giving up marijuana, restricting his preaching, having a girlfriend, attending church, and taking a smaller dose of medications. Several team members admired his courage and integrity, and he often called on them for help. He did not return to homelessness, jail, or the hospital over several years.

Research on Supported Housing

The absence of operationally defined program models limited most of the early research on supported housing and other residential models. The inconsistent use of different labels, including “supported” and “supportive” housing, confounded the problem. Researchers have not always agreed on the core components of supported housing, and some principles are value statements not easily translated into objective criteria.

Housing First defined more clearly the core model components, and consequently research has yielded a coherent set of replicated findings. Large-scale pragmatic trials in the United States and Canada have demonstrated that most people with serious mental disorders can maintain independent housing over the long term when they have safe, affordable housing and adequate psychosocial supports. Housing First studies have also shown that offering housing to people with substance use disorders, again without preconditions, does not lead to the dire consequences that

paternalistic approaches assumed. Several studies have found that clients receiving supported housing assistance are more satisfied with their housing and have reduced use of psychiatric hospitals and emergency rooms. Yet research has not demonstrated that improvement in housing outcomes generalizes to other areas of functioning or to clinical outcomes, such as substance use and psychiatric symptoms.

INDIVIDUAL PLACEMENT AND SUPPORT

History

Impairment of vocational role performance complicates adjustment for patients with schizophrenia and other serious mental disorders. Across the United States, less than 15 percent of patients are employed, even though 50 to 75 percent state that competitive employment is a primary goal. Due to patient interests and historical factors discussed earlier, vocational rehabilitation has always been a centerpiece of psychiatric rehabilitation.

Clinicians, rehabilitation specialists, researchers, and patients themselves have developed many approaches to help people with serious mental illness achieve employment over the last half-century. Most have involved train-and-place models, with extensive pre-employment assessments and preparatory experiences, such as day centers, sheltered workshops, transitional employment, and skills training. These aimed to evaluate and improve the patient's attitudes, skills, appearance, readiness, and other abilities presumed necessary to obtain and sustain competitive employment. Evaluations of these step-wise approaches to vocational rehabilitation found that they did not improve the rate of competitive employment; most patients became discouraged, disengaged, or stalled during preparatory experiences.

Supported Employment

In the late 1980s, the emphasis of vocational rehabilitation began to shift to place-and-train models, termed supported employment, which were borrowed from the field of developmental disabilities. According to the federal definition, supported employment involves "competitive work in integrated work settings … consistent with the strengths, resources, priorities, concerns, abilities, capabilities, interests, and informed choice of the individuals, for individuals with the most significant disabilities for whom competitive employment has not traditionally occurred; or for whom competitive employment has been interrupted or intermittent as a result of a significant disability." In contrast to other vocational approaches, supported employment programs (a) did not screen people for work readiness, but help anyone who says they want to work; (b) did not provide intermediate work experiences, such as prevocational work units, transitional employment, or sheltered workshops; (c) actively facilitated

job acquisition, often accompanying clients on interviews; and (d) provided ongoing support once the person is employed.

In the 1990s a model of supported employment adapted specifically for people with severe mental illness appeared. Individual Placement and Support (IPS), also known as evidence-based supported employment, emphasizes the patient's personal preferences for work, a rapid job search (without extensive pre-employment assessment or training), and individualized supports as needed. Employment specialists provide practical assistance at every stage of the employment process. Jobs are typically part-time, beginning at just a few hours per week and expanding over time. IPS emphasizes regular jobs in the competitive labor market, owned by the employee rather than the rehabilitation program, paid at or above minimal wage, and supervised by the employer.

Several features of IPS supported employment correlate with better employment outcomes and thus are considered evidence-based principles.

(1) Zero Exclusion. Supported employment services are available for anyone who wants to work, regardless of symptoms, cognitive impairments, work history, substance use, or other personal characteristics. Research shows that professionals cannot reliably identify who will succeed or fail in employment.

(2) Competitive Employment. The goal of supported employment is competitive employment in work settings integrated in a community's economy. The great majority of patients who seek work want competitive employment, particularly if they are encouraged to believe that they can work. Few patients want sheltered or volunteer jobs.

(3) Rapid Job Search. Clients seek employment directly, rather than following lengthy pre-employment assessment, training, or practice. Rapid job search typically results in finding a first job within 4 months.

(4) Systematic Job Development. The employment specialist conducts targeted job development to help find jobs that match people's skills and preferences. Job development often involves the entire team identifying or pursuing relationships with employers.

(5) Integration of Rehabilitation and Mental Health Treatment. Supported employment should be an integral component of mental health treatment, not a separate service. Employment specialists are integrated into case management teams, assertive community treatment teams, or other multidisciplinary teams. Employment specialists meet regularly with their teams and collaborate on helping patients to succeed in jobs.

(6) Patient Preferences. Services are based on patients' preferences and choices. Patients report higher satisfaction and longer job tenure when the jobs are consistent with their preferences.

(7) On-going Support. Follow-along supports align with the individual patient's needs and continue until the employee is independent.

The multidisciplinary team provides a range of supports over time, with greater intensity as needed during times of crisis. For example, the psychiatrist may alter medications to help with the initial anxiety of working; the case manager may meet with the family to help them be supportive; and the employment specialist may provide benefits counseling, off-site supports, and skills training.

(8) Individualized Benefits Counseling. Many patients who receive disability, welfare, or other benefits fear they could lose these benefits if they work. Individualized, accurate information about benefits often corrects misconceptions and helps patients make good decisions.

Employment specialists typically are college graduates supervised by Master's level vocational specialists. Employment specialists provide vocational assessment, placement, and support, but the multidisciplinary team is involved throughout. The case manager and medical providers need to encourage patients to try employment. The entire team may help in finding a job and providing supports. For example, the case manager may help solve problems related to interpersonal difficulties on the job, the social worker may meet with the family to understand their concerns about employment, and the psychiatrist may provide extra medications to help with initial anxiety on the job or decrease medications to reduce side effects that interfere with working.

The key to success is finding the right match—a job that aligns with the patient's interests, abilities, ways of coping with illness, needs regarding benefits, family situation, and so on. Different jobs often lead to an optimal job match.

Antonio was a 45-year-old man who has been a client of a mental health agency for over 10 years. He attended a rehabilitative day treatment program until it was converted to a supported employment program. His case manager encouraged him to think about the possibility of working part-time. Antonio told his case manager that he couldn't work because of schizophrenia and because he was helping to raise his two kids and needed to be home at 3 PM when they arrived home from school every day. The case manager explained to Antonio that going to work doesn't necessarily mean a 40-hour a week job and that lots of people in the agency's supported employment program work in part-time jobs, even jobs that are only a few hours a week.

Antonio agreed to meet one of the employment specialists to discuss the possibility of work. Over the next couple of weeks, the employment specialist met with Antonio several times, read his clinical record, and talked with his case manager and psychiatrist. The employment specialist learned that Antonio loved to drive his car, that he had attendance problems in past jobs, and that he was a kind, friendly, and likeable person.

Antonio told the employment specialist that he was willing to do any job. He didn't have one specific job in mind. The employment specialist suggested a job at Meals on Wheels, as a driver for the lunch delivery. Antonio was hired and loved it right from the start. Absenteeism was never a problem, because he knew that people were counting on him to help with meal delivery. The hours were perfect (10 AM to 2 PM) so he could be at home when his kids returned from school. He became good friends with the other people who worked there. He told his case manager that it was wonderful to be bringing home a paycheck again. And best of all, he said, was that his kids saw him going to work just like their friends' dads.

Research on IPS Supported Employment

Research has produced several consistent findings. Interventions that do not target job placement directly, such as housing, counseling, and medications, have very little impact on employment outcomes. In order to help patients achieve competitive employment, the intervention needs to be specifically directed at this goal.

Interventions that focus on pre-employment skills training do not produce better vocational outcomes of any kind, including type, quality, amount, length, or satisfaction related to employment. Supported employment interventions that focus on direct and individualized placements with follow-along supports produce much better outcomes.

The evidence strongly favors IPS supported employment over other approaches to vocational rehabilitation. More than 20 randomized controlled trials show that IPS supported employment produces better employment outcomes, including higher rates of employment, greater earnings, more hours of work, and longer tenure in jobs. The magnitude of these effects is substantial. For example, rates of employment are three times higher in IPS supported employment programs than in the comparison conditions in the United States. Most of these trials are pragmatic trials, conducted with typical patients, providers, and programs. Figure 12.13–1 summarizes competitive employment rates for IPS and control participants in 22 randomized trials.

All types of patients do better in IPS supported employment than in traditional vocational programs: for example, patients with poor vocational histories, those from minority backgrounds, specifically African Americans and Latino patients, those with co-occurring disorders, and inner-city residents who have recently been homeless. Research also shows that patients themselves express high satisfaction with IPS supported employment jobs and that they typically stay in first jobs for 5 to 6 months, just like the general population starting first jobs. Several long-term follow-up studies indicate that patients maintain initial high employment rates over 5 to 10 years.

Although some clinicians still fear that high expectations will destabilize patients with schizophrenia or other serious mental disorders, IPS supported employment services do not produce clinical instability, higher rates of relapse, or other untoward outcomes. On the contrary, patients who are currently in competitive employment appear to benefit in numerous ways and report higher satisfaction with services, better self-esteem, and greater quality of life.

TECHNOLOGY TOOLS

Technology Use among People with Schizophrenia

People with psychiatric illnesses, including those with schizophrenia, use

contemporary technologies in a manner similar to the general population. Survey studies suggest that 80 to 94 percent of people with schizophrenia access the internet regularly, use email, and make good use of social networking websites, online forums, chat rooms, and blogs. Moreover, people with serious mental illnesses such as schizophrenia are already using a host of digital resources for mental health reasons: obtaining information about mental disorders and treatments, searching for mental health services, connecting with peers with similar conditions, and engaging with online communities and supports. For example, people with psychiatric conditions are posting video testimonials in which they openly disclose their illness and provide first-person accounts of their symptoms and coping strategies on YouTube, a popular video-sharing website. Their posts receive numerous comments from viewers who describe their own experiences with illness, provide positive feedback, share symptom management suggestions, and thank the individual for sharing their video. In this manner, a virtual community can support people who may not have access to local supports, and help erode the stigma that often prevents people from disclosing their mental health struggles.

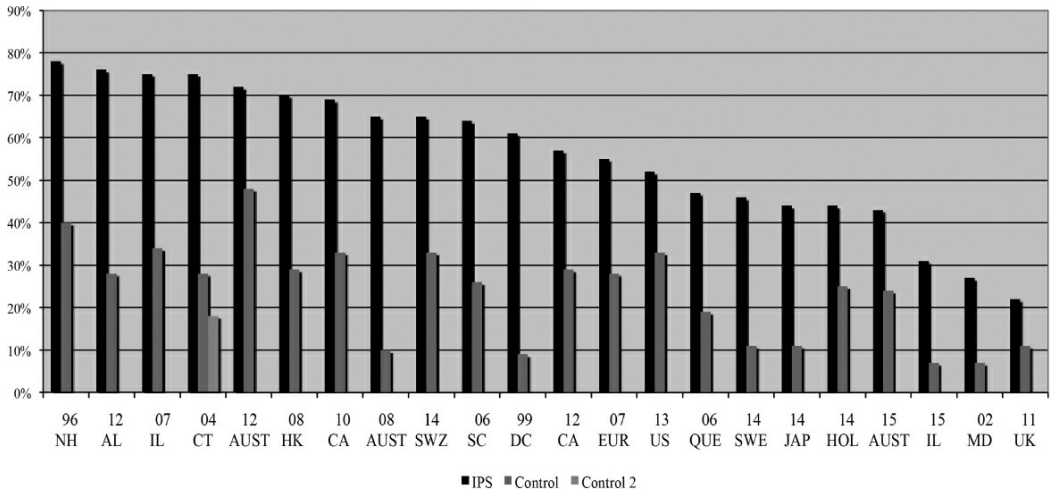

FIGURE 12.13–1. Competitive employment rates in 22 randomized controlled trials of individual placement and support.

Mobile-cellular technology is pervasive—one of the fastest adopted technologies in human history. The number of mobile phone subscriptions has gone from 1 billion in 2001 to 7 billion in 2014, with over three-quarters of the world's mobile devices now being used in low- and middle-income countries. To contextualize the potential healthcare implications of such a ubiquitous resource, more people in the world now have access to mobile phones than to working toilets. When mobile phones first became commercially available at the end of the 20th century, they were inaccessible for people with significant disabilities and limited resources. But rapid advancements in development of mobile hardware coupled with

vast expansion in global mobile-cellular infrastructure have contributed to significant drops in the costs of devices, as well as access to data plans. Consequently, the majority of the population can make calls and send text messages using their personal mobile device with relative ease. In the United States, several federal and state programs now subsidize basic mobile phones and monthly call/text minutes to make them accessible to individuals below a certain income level. Thus, many more people with schizophrenia have access to a wealth of information and resources. Survey studies in the United States have shown that 72 to 97 percent of people with serious mental health conditions own and use mobile technologies including cellular phones, text messaging devices for the hearing impaired, and smartphones (i.e., mobile phones with significant computational and data storage capacities). Similar access occurs in several European countries, the UK, and Australia.

Both community-based outpatients and acutely ill inpatients with schizophrenia can use mobile devices successfully. Over two decades of clinical research using mobile devices, including beepers, personal digital assistants (PDAs), and smartphones to record behaviors, symptoms, and functioning, suggest that people with schizophrenia are quite open to using mobile self-tracking tools, and that they experience little to no iatrogenic effects from doing so. Even highly symptomatic patients with paranoid schizophrenia have successfully engaged in Ecological Momentary Assessment (EMA) studies that require them to self-report on distressing thoughts and even illegal substance abuse using mobile devices. People with schizophrenia consistently report their interest in using technological tools to support illness management and functional recovery. Research groups in different countries have developed innovative technological approaches that aim to accomplish this goal.

Technological Approaches for Assessment and Treatment

Web-Based Interventions. People with schizophrenia, their family members, and their supporters can access a range of resources on the internet, for example, illness management programs, chat forums, support groups, and informational webpages. Acceptability, usability, and helpfulness of web-resources depend upon whether they were developed with close consideration of the needs of people with severe psychiatric illnesses (i.e., possible cognitive impairments, low literacy, dexterity problems).

Web-resources with appropriately designed user interfaces can be highly engaging and effective. For example, *Coping with Voices,* a self-directed web-based Cognitive Behavioral Intervention, addresses the functional impact and distress caused by auditory hallucinations. Users progress through an interactive 10-lesson program at their own pace and revisit content form previous lessons. Video tutorials educate people about

psychosis and dysfunctional cognitive styles using interactive games, quizzes, and practice opportunities. Participants who complete *Coping with Voices* experienced reductions in the severity and intensity of auditory hallucinations as well as other psychotic symptoms and general psychopathology.

Schizophrenia Online Access to Resources (SOAR) is a multifamily intervention for people with schizophrenia and their supports. The key elements of the *SOAR* intervention include empathic engagement of users, education about illness and treatments, and use of coping strategies. *SOAR* offers three therapy forums: one for people with schizophrenia, one for support persons, and one for both groups. Users can post questions and receive answers from trained clinicians within 24 to 48 hours. *SOAR* also contains a record of responses to previously posted questions, a library of relevant reading materials, and a list of community events, activities, and resources that may be of interest. People with schizophrenia who engage in SOAR experience a significant reduction in positive symptoms of psychosis. Support persons also showed a significant and sustained increase in knowledge about the course and prognosis of schizophrenia.

The *HORYZONS* system is an online platform developed for long-term recovery of people experiencing a first episode of psychosis. The system delivers several evidence-based and interactive psychosocial interventions and facilitates a moderated online peer-to-peer social networking environment. Users respond to a set of standardized questions that guide the tailored modules that target risk factors for psychotic relapse and other areas that are salient to people in early recovery, including psychoeducation, early warning signs, depression, social anxiety, and stress management. First-episode patients who used the *HORYZONS* system reduced depression and found the approach to be engaging, empowering, and safe.

Mobile Health. The immense capabilities and functionality of mobile technology have led to the development of the interdisciplinary field of Mobile Health (mHealth). Mobile devices can support the delivery of healthcare in a variety of formats, enabling patients and providers to send and receive calls and text messages, access websites, provide clinical decision support, and capture and transmit data collected in real-time and in users' own environment. Mobile devices with computational capacity can also house health-focused applications (apps) designed to facilitate patient illness self-management.

FOCUS, the first smartphone intervention specifically designed for people with schizophrenia, was developed iteratively and in close consultation with patients and clinicians. The *FOCUS* system can be activated to prompt users multiple times a day with questions about their symptoms, social functioning, mood, medication adherence, and sleep, which appear on the devices touchscreen. Depending on users' responses,

the system launches brief interactive interventions in the form of written text, images, and cartoon screens, designed to both give the individual skills they can use in the moment (e.g., relaxation strategies, medication reminders, social skills tips), as well as train the user in techniques that can be generalizable to a range of scenarios (e.g., cognitive restructuring, hypothesis testing, behavioral activation, distress tolerance). In addition to prompted interactions, all *FOCUS* interventions are available on-demand as part of a toolkit that can be accessed at any time. All user responses are recorded on the device and the system can send the data to a remote server whenever the phone has connectivity. Authorized individuals (e.g., clinicians, researchers) can access a web-based dashboard summarizing all user responses shortly after they engaged the system. People with schizophrenia find the *FOCUS* intervention highly acceptable and useable. Patients who used *FOCUS* experienced significant reduction in symptoms of psychosis, general psychopathology, and depression.

Lakesha is a 38-year-old woman with schizophrenia who receives community-based mental health services. Lakesha suffers from persecutory ideation and auditory hallucinations that warn her about making eye contact with strangers who are "out to humiliate her." Despite having access to a host of treatment options, she makes only sporadic use of the available resources because she is apprehensive about using public transportation to get to the clinic, and is concerned about "sitting in the waiting area with a bunch of crazies." Consequently, she is becoming increasingly isolative and is experiencing greater distress and impairment linked with her symptoms. Lakesha has an elderly mother with whom she is quite close, but whom she has been visiting less and less because she is convinced that the local bus drivers are intending to kidnap her.

Lakesha's psychiatrist suggested she enroll in a treatment study in which participants are given the *FOCUS* smartphone system for 3 months. Initially, Lakesha thought she would not be able to use the intervention because she was unfamiliar with smartphones. However, after a 45-minute training session with research staff she used the device successfully and learned how to operate the touchscreen to respond to *FOCUS* questions and suggestions. Lakesha started using FOCUS that evening, and continued to respond to clinical prompts daily. She found the guided breathing techniques to be particularly helpful as she could use them flexibly, whenever she wanted, including on the bus ride to see her mother.

After eight weeks of smartphone intervention use, Lakesha reported leaving her house daily to go to the local store, community center, or to visit a neighbor. She reported feeling less anxious about others' intentions. While she still experienced auditory hallucinations regularly, she described being less concerned about them. She successfully integrated several *FOCUS* strategies into her illness management repertoire, and is now able to engage in relaxation exercises and basic thought challenging of dysfunctional beliefs (e.g., "what is the evidence that I'm in danger when I leave the house?") independently.

Short Message Service (SMS) or text messaging or approaches have also been used to support psychiatric rehabilitation with patients with schizophrenia. For example, *Mobile Assessment and Treatment for Schizophrenia (MATS)* is a mobile phone intervention employed to administer questions and algorithm-driven follow-up suggestions to patients in the areas of medication adherence, social functioning, and management of auditory hallucinations. *MATS* initiates a sequence of brief

automated back-and-forth exchanges, including illness-management suggestions drawn from cognitive behavioral treatments for psychosis.

In addition to enabling completely novel treatments, mobile technologies can supplement traditional models of care. Community-based research has shown that patients with psychosis and co-occurring substance use find texting to be a highly engaging method for connecting with their case manager. When a *Mobile Interventionist* used daily texting to check in and provide daily support, patients responded to 87 percent of the messages sent and rated highly the therapeutic alliance with the *Mobile Interventionist*.

Virtual Reality. Virtual reality (VR) refers to multimedia techniques in which individuals engage in computer-generated scenarios on a computer screen or head-mounted stereoscopic display that creates a highly immersive 3D environment. VR has been used primarily to facilitate exposure therapy for people with anxiety disorders, but several studies have demonstrated its potential utility in supporting rehabilitation for people with psychosis. These include help identifying contextual triggers for persecutory ideation, learning skills, and practicing job interviews. Preliminary evidence suggests VR may be an effective tool, including benefits from continuous feedback from an on-screen virtual coach.

SUMMARY

Psychiatric rehabilitation comprises a range of techniques and interventions that help people with disabilities related to psychiatric illness to manage their illness, succeed in regular social roles in the community, pursue their personal goals, and "recover" in the sense of moving beyond feeling dominated by illness to create a meaningful life. Rehabilitation requires creating normative opportunities, such as for education, employment, and housing; providing supports as needed to ensure success; and helping people acquire the skills they want to pursue life goals.

REFERENCES

Alvarez-Jimenez M, Alcazar-Corcoles MA, González-Blanch C, Bendall S, McGorry PD, Gleeson JF. Online, social media and mobile technologies for psychosis treatment: a systematic review on novel user-led interventions. *Schizophr Res*. 2014;156(1):96–106.

Aubry T, Tsemberis S, Adair CE, et al. One-year outcomes of a randomized controlled trial of Housing First with ACT in five Canadian cities. *Psychiatric Serv*. 2015;66:463–469.

Baumel A, Correll CU, Hauser M, et al. Health technology intervention after hospitalization in schizophrenia: service utilization and user satisfaction. *Psychiatric Serv*. 2016;67(9):1035–1038.

Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. *Schizophr Bul*. 2014;40(6):1244–1253.

Ben-Zeev D, Kaiser SM, Krzos I. Remote “hovering” with individuals with psychotic disorders and substance abuse: feasibility, patient engagement, and therapeutic alliance with a text-messaging mobile interventionist. *J Dual Diagn*. 2014;10(4):197–203.

Burns T, Catty J, Becker T, et al. The effectiveness of supported employment for people with severe mental illness: a randomised controlled trial. *Lancet*. 2007;370:1146–1152.

Corrigan PW. *Principles and Practice of Psychiatric Rehabilitation: An Empirical Approach* (2nd ed.). New York: Guilford Press; 2016.

Drake RE, Bond GR, Becker DR. *Individual Placement and Support: An Evidence-Based Approach to Supported Employment*. New York: Oxford University Press; 2012.

Drake RE, Bond GR, Goldman HH, Hogan MF, Karakus M. Individual placement and support services boost employment for people with serious mental illness, but funding is lacking. *Health Affairs*. 2016;35:1098–1105.

Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile phone ownership and endorsement of “mHealth” among people with psychosis: a meta-analysis of cross-sectional studies. *Schizophr Bull*. 2016;42(2):448–455.

Hoffmann H, Jäckel D, Glauser S, Mueser KT, Kupper Z. Long-term effectiveness of supported employment: five-year follow-up of a randomized controlled trial. *Am J Psychiatry*. 2014;171(11):1183–1190.

Liberman RP. *Recovery From Disability: Manual of Psychiatric Rehabilitation*. Washington, DC: American Psychiatric Press; 2009.

Marshall T, Marshall T, Goldberg RW, et al. Supported employment: assessing the evidence. *Psychiatr Serv*. 2014;65:16–23.

Rog DJ, Marshall T, Dougherty RH, et al. Permanent supportive housing: assessing the evidence. *Psychiatr Serv*. 2014;65:287–294.

Rotondi A, Anderson C, Haas G, et al. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: one-year outcomes. *Psychiatr Serv*. 2010;61(11):1099–1105.

Swanson SJ, Becker DR. *IPS Supported Employment: A Practical Guide*. Lebanon, NH: Dartmouth Psychiatric Research Center; 2013.

Tabol C, Drebing C, Rosenheck RA. Studies of “supported” and “supportive” housing: A comprehensive review of model descriptions and measurement. *Eval Prog Plann*. 2010;33:446–456.

Thornicroft G, Szmukler G, Mueser KT, Drake RE, eds. *Oxford Textbook of Community Mental Health*. New York: Oxford Press; 2011.

Tsemberis S, Gulcur L, Nakae M. Housing First, consumer choice, and harm reduction for homeless individuals with a dual diagnosis. *Am J Public Health*. 2004;94:651–656.

▲ 12.14 Medical Health in Schizophrenia

JOHN W. NEWCOMER, M.D., MARTIN T. STRASSNIG, M.D., PETER A. FAHNESTOCK, M.D.*, AND MICHELE L. PERGADIA, PH.D.

INTRODUCTION

Consistent evidence spanning decades, countries, and clinical settings indicates that individuals with schizophrenia suffer reduced life expectancy relative to the general population. In 1841, William Farr published the first study explicitly examining death rates in psychiatric patients, titled "Report on the Mortality of Lunatics." The report was based on a handful of English asylums, detailing increased mortality overall as well as discrepancies between male and female patients and between economic groups. In 1936, Ornulv Odegard published an assessment of mortality rates in Norwegian mental hospitals. He observed that patients with severe mental illnesses had higher rates of mortality than those in the general population, mainly due to tuberculosis, pneumonia, and other infectious diseases, with relatively low rates for heart disease and cancer. In retrospect, the low rates of the latter conditions are likely related to premature death from the former conditions, as well as lower levels of smoking and obesity compared to contemporary rates. Even in these early studies, mortality rates were four to five times higher in psychiatric patients compared to the general population. Odegard believed that this excess mortality was partly a function of environmental factors, such as overcrowding and poorly prepared food, and partly a function of specific symptoms of mental illness, such as apathy resulting in an unwillingness to complain or be examined by a physician. In 1959, a period when tobacco use was at its highest level, Abraham published "Cardiovascular Disease in Psychotics," which showed an increased prevalence of atherosclerotic and hypertensive CVD in hospitalized patients with schizophrenia. He rejected the popular view of the time that "the protective life of people with schizophrenia seems to guard them from early onset of degenerative diseases of the heart." By the late 1970s, accumulating mortality data from psychiatric patients showed a consistent pattern of increased mortality rates relative to the general population. However, despite such reports there were only a few controlled studies in that period that aimed to understand the causes and contributions to this excess mortality.

Today, a growing body of literature addresses the public health problem of excess mortality among the mentally ill, including especially patients with schizophrenia. Leucht and colleagues (2007) searched MEDLINE database using the term "schizophrenia" in combination with 23 physical disease categories. The search yielded 44,202 abstracts, with "eighty-six

percent of the studies … from industrialized countries." Information on mortality for patients with schizophrenia in developing countries remains limited, where reduced access to health care is expected to be associated with further reductions in life expectancy.

ALL-CAUSE MORTALITY IN SCHIZOPHRENIA

The impact of medical morbidity and mortality in patients with schizophrenia has received increasing attention in recent years. The average life expectancy for individuals with schizophrenia has been estimated to be 10 to 30 years shorter than that of the general population. Numerous studies indicate that, compared with the general population, schizophrenia patients are at higher risk for development of various medical illnesses, are less likely to receive and benefit from appropriate medical treatment, are more likely to experience unfavorable clinical outcomes, and are more likely to die prematurely. In 2006, Colton and Manderscheid compiled data from the public mental health systems of eight states in the United States and compared the life expectancy of those patients with a major mental illness with general population norms. Focusing on the states with available outpatient as well as inpatient data sources, this study indicated that individuals with a major mental illness have a mean life span that is 25 to 30 years shorter than that observed in the general population over the same years in the same states (Table 12.14–1). In this study, "major mental illness" includes affective disorders like major depression and bipolar disorder, ADHD, schizophrenia, and schizoaffective disorders. Importantly, these results also indicated that the leading cause of early deaths is coronary heart disease (CHD). Also represented in the tabulation was death due to stroke or cerebrovascular disease. With the term CVD referring to the combination of CHD, cerebrovascular disease, and peripheral vascular disease, this study illustrated that CVD is responsible for more than 35 percent of deaths in this sample. Suicide, by contrast, was responsible for fewer than 5 percent of deaths. Such observations have led to growing clinical interest in the risk factors that contribute to morbidity and mortality in patients with schizophrenia.

Table 12.14–1.
Mortality Associated with Major Mental Disorders (Mean Years of Potential Life Lost Compared with the General Population; Persons with Major Mental Illness Lose 25 to 30 Years of Normal Life Expectancy)

Year	AZ	MO	OK	RI	TX	UT
1997	—	26.3	25.1	—	28.5	—
1998	—	27.3	25.1	—	28.8	29.3
1999	32.2	26.8	26.3	—	29.3	26.9

2000	31.8	27.9	—	24.9	—	—

Data from Colton CW, Manderscheid RW: Congruencies in increased mortality rates, years of potential life lost, and causes of death among public mental health clients in eight states. *Prev Chronic Dis.* 2006;3(2):A42.

In some of the prior literature in this area, SMRs have been used to quantify the additional mortality risk associated with schizophrenia relative to the general population, although caution is needed when interpreting this metric. SMRs are calculated as the ratio of observed to expected deaths within a specified sample or given age range (sometimes times 100). The *expected* mortality rate is the rate seen in the general population, whereas the *observed* mortality rate is that seen in the population in question (e.g., those with schizophrenia). Therefore, an SMR for myocardial infarction (MI) among schizophrenia patients describes the likelihood of a person with schizophrenia dying of an MI, expressed as a function of the likelihood that a member of the general population will die of an MI. It is therefore a measure of relative and not absolute risk. SMRs can be useful for identifying factors that contribute to the relative risk of death for psychiatric patients in comparison with the general population. In 2007, Saha and colleagues studied the distribution of all-cause SMRs in schizophrenia patients, comparing data from 37 different studies. This analysis yielded a median all-cause SMR of 2.58 compared to the general population. Increased SMRs were found for most major causes of death, and notably, the analysis showed that overall SMRs have been increasing over the past three decades, indicating that the gap between mortality rates observed in schizophrenia patients and rates in the general population is growing wider despite increasing awareness of the problem.

Register-based studies comparing mortality in patients with schizophrenia versus the general population over a specified period of years have contributed greatly to the current understanding of the clinical importance of medical comorbidities to long-term health outcomes, and are currently a major source of morbidity and mortality data in patients with schizophrenia. The age-adjusted risk for all-cause mortality is consistently high in schizophrenia patients among registry studies. A recent study by Brown and colleagues of 370 schizophrenia patients over 25 years confirmed that most patients died from natural causes rather than unnatural causes (such as suicide or accidents); in this study, 81 percent of patients died as a result of medical diseases. The all-cause mortality SMR was 2.89 in the overall study population, with disease-specific SMRs of 1.49 for cancer, 2.25 for CVD, 2.65 for lung cancer, 2.58 for circulatory diseases, 3.08 for cerebrovascular disease, 4.27 for diseases of the nervous system, 6.14 for diabetes mellitus, 8.01 for endocrine diseases, 4.99 for respiratory diseases, and 8.35 for pneumonia.

Evidence from population-based cohort studies consistently identifies a

significant increase in overall all-cause mortality ranging from two to four times higher in patients with schizophrenia compared to the general population, with somewhat smaller increases sometimes observed in exclusively inpatient samples. In a population-based sample of 7,784 inpatients with schizophrenia analyzed by Osby and colleagues, male and female subjects had an SMR of 2.8 and 2.4, respectively, for death due to any cause. All natural causes of death, except cancer in men and nervous system diseases in women, were significantly elevated. Overall, patients were more than twice as likely to die from cardiovascular, cerebrovascular, or respiratory disease. Addressing the sample size limitations that have previously limited conclusions about underlying causes of excess mortality in this population, a recent large study by Olfson and colleagues reported on SMRs for over 1 million US Medicaid patients followed for almost 5 million person-years of follow-up, with over 74,000 deaths. Patients with schizophrenia were 3.7 times more likely to die in the follow-up period compared to the general population. CVD had an SMR of 3.6 and the highest mortality rate (403.2 per 100,000 person-years) of any individual cause. Cancer overall had the next highest mortality rate, with lung cancer contributing the most with an SMR of 2.4 and a mortality rate of 74.8 per 100,000 person-years. Elevated SMRs were also observed for chronic obstructive pulmonary disease (COPD) (SMR = 9.9) and influenza and pneumonia (SMR = 7.0), suggesting the effects of high smoking rates in this population, as well as for diabetes mellitus (SMR = 4.2, mortality rate = 61.8 per 100,000 person-years). Nonsuicide-related substance-induced death, mostly from alcohol and/or other drugs, was also an important cause of death (95.2 per 100,000 person-years). Finally, accidental deaths (119.7 per 100,000 person-years) accounted for more than twice the number of deaths attributable to suicide (52.0 per 100,000 person-years). While much attention understandably goes to the effects of suicide in this population, this large and definitive study from the public sector underscores the important effects of medical morbidity and mortality in this population.

MEDICAL CAUSES OF EXCESS MORTALITY IN SCHIZOPHRENIA

A growing number of studies indicate that comorbid medical conditions, rather than unnatural causes of death such as suicide, confer most of the excess mortality seen in persons with schizophrenia. Medical comorbidities often remain underdiagnosed and undertreated in these patients, despite the well-established literature of registry studies, prospective trials, and case reports that indicate increased risk.

Cardiovascular Disease

In 2011, heart disease was listed as the top cause of death, accounting for

23.7 percent of all cases, or 1 of every 4.2 deaths in the United States, with cerebrovascular disease accounting for an additional 5.1 percent, numbers similar to those seen in other developed countries. During the past several decades, CVD mortality in the general population has markedly declined in the United States, from more than 50 percent to approximately 30 percent, but this improvement in public health outcome has not extended to patients with schizophrenia. CVD is also the leading cause of mortality in individuals with schizophrenia, who are even more likely to experience premature cardiovascular mortality than individuals in the general population. The prevalence of CVD in people with schizophrenia is approximately two- to threefold increased, particularly in younger individuals. In a retrospective cohort study conducted by Curkendall and colleagues, 3,022 individuals with schizophrenia were compared to a general population cohort. The overall prevalence of CVD was increased in the patients with schizophrenia (10.6 vs. 8.7 percent), as was with the incidence of ventricular arrhythmia (odds ratio [OR] = 2.3; 95 percent confidence interval [CI], 1.2 to 4.3), stroke (OR = 1.5; 95 percent CI, 1.2 to 2.0), diabetes (OR = 1.8; 95 percent CI, 1.2 to 2.6), and heart failure (OR = 1.6; 95 percent CI, 1.2 to 2.0). Cardiovascular mortality in schizophrenia has also been evaluated in large population-based samples using long periods of observation. Patients with schizophrenia were evaluated for mortality risk over a period of 19 years in a Swedish registry study, analyzed by Osby and colleagues. Between 1976 and 1991, death rates due to CVD increased 4.7-fold in men and 2.7-fold in women.

The Framingham Heart Study (a long-term cohort study funded by the National Heart, Lung, and Blood Institute) and other large population-based samples have identified modifiable risk factors for CVD and how these risk factors interact. The knowledge gained from these studies has important implications for decreasing cardiovascular-related morbidity and mortality, including in higher-risk populations such as people with schizophrenia. It is important to note here that persons with schizophrenia have an increased prevalence of all key modifiable cardiovascular and metabolic risk factors, including obesity, smoking, hypertension, dyslipidemia, and hyperglycemia.

A study conducted by McCreadie and colleagues, used the Framingham assessment methodology to determine the risk of CHD in 8,217 patients with schizophrenia in the United Kingdom over 10 years. The impact of diet, weight, smoking, and exercise habits on CHD risk was evaluated. Men with schizophrenia had a significantly higher risk of CHD (10.5 vs. 6.4 percent) and stroke (4.2 vs. 2.3 percent) compared with the general population. Risk among women in this study was not statistically higher than that of women in the general population. However, this result contrasts with cardiometabolic risk data from schizophrenia patients entering the US-based CATIE study. Joseph P. McEvoy and colleagues

compared the baseline presence of cardiometabolic risk factors in the CATIE patient sample with an age-, gender- and race/ethnicity-matched general population sample from the third National Health and Nutrition Examination Survey (NHANES III). Not only did both sexes in the CATIE population have an elevated prevalence of cardiometabolic risk factors compared to NHANES III controls, women tended to have even higher risk than men. In another study comparing CATIE patients with schizophrenia to an age-matched sample from NHANES III, Goff and colleagues reported that schizophrenia patients had a significantly higher 10-year risk of CHD than matched controls: 9.4 versus 7.0 percent in men, and 6.3 versus 4.2 percent in women, respectively. Lawrence and colleagues examined excess mortality due to CHD for psychiatric patients in comparison to the general population and found that men with schizophrenia were only 60 percent as likely to be admitted to a hospital for CHD, but were almost two times more likely to die from CHD. Noted above in the largest cohort study to date, from Olfson and colleagues, CVD made the highest individual contribution to overall mortality, with a rate of 403.2 deaths per 100,000 person-years.

Diabetes Mellitus

Diabetes-related morbidity and mortality has a major effect on the general population, in the context of increasing overweight and obesity rates as observed in developed countries. As of this writing, diabetes is the seventh leading cause of death in the United States. The prevalence of diabetes is expected to double or triple by 2050, according to projections from the US Centers for Disease Control and Prevention (CDC).

The prevalence of type 2 diabetes mellitus is substantially higher in individuals with schizophrenia than it is in the general population, with prevalence estimates ranging from two to four times higher. Multiple studies indicate that roughly 15 to 18 percent of patients with schizophrenia have type 2 diabetes mellitus, compared with an overall prevalence of approximately 9 percent in the general population. The incidence of diabetes in the general population increases progressively with a person's age, while schizophrenia is associated with an increase in the incidence of diabetes in earlier adult years, with prevalence then maintained at that elevated level over the life span. The largest cohort study in patients with schizophrenia to date reported an elevated SMR and an elevated absolute mortality rate from diabetes mellitus (SMR = 4.2; mortality rate = 61.8 per 100,000 person-years).

The interaction between schizophrenia and diabetes mellitus is complex and remains incompletely understood, but parsimony suggests that the well-documented prevalence of a number of risk factors in this population may be sufficient to explain the increased disease prevalence. Multiple risk factors underlie the increased rate of diabetes among those with schizophrenia, including poverty, urbanization, crowding, psychological

stress, smoking, and the effects of treatments like antipsychotic medications on rates of obesity, insulin resistance and dyslipidemia, as well as genetic factors. There are limited data that might suggest that increased activation of the hypothalamic–pituitary–adrenal (HPA) axis and the sympathetic nervous system may contribute to at least acute hyperglycemia drug-naive schizophrenia patients. However, the finding of increased adiposity or glucoregulatory impairments in unmedicated patients is not consistently observed, and HPA activation is only variably associated with schizophrenia and is generally reduced by antipsychotic treatment. Familial, possibly genetic, factors may play a role, as one report found that 18 to 19 percent of individuals with schizophrenia had a family history of type 2 diabetes mellitus, and, in general, studies of probands with diabetes can be subject to an ascertainment bias that would tend to increase the prevalence of diabetes observed in family members, independent of any effect of schizophrenia. It remains to be seen to what extent the increased prevalence of diabetes in this population cannot be fully explained by increases in the prevalence of overweight and obesity, reductions in the overall level of physical fitness, and other secondary detrimental effects related to lifestyle.

Underrecognition of schizophrenia as a marker for increased risk for diabetes may contribute to the impact of diabetes-related risk on the overall elevated mortality, in the absence of extra efforts in primary or secondary prevention. Of note, the Canadian Diabetes Association and the American Association of Clinical Endocrinologists have added schizophrenia to the list of risk factors for diabetes in their screening guidelines. Ongoing evidence documents generally low rates of screening for abnormalities in plasma glucose or lipids in this population.

Infectious Disease

In developing countries, infectious disease remains a substantial cause of mortality for persons with and without schizophrenia. Conditions in some parts of the world today remain similar to those described in the industrialized world of the early 20th century by Odegard. In developed countries, rates of certain infectious diseases among patients with schizophrenia still remain higher than in the general population. For example, the rate of human immunodeficiency virus (HIV) infection has been estimated at 3 percent in patients with a severe mental illness, an eightfold increase over the general population rate, and data have shown that knowledge of exposure risks is particularly low among these patients. Rates of hepatitis B (23.4 percent) and C (19.6 percent) infections are also significantly elevated among those with a severe mental illness, approximately 5 and 11 times the overall estimated population rates for these infections. Studies also consistently show a higher incidence of tuberculosis among patients with schizophrenia, compared to the general

population.

Cancer

For 2014 the American Cancer Society estimated over 1.6 million new cases of cancer and over half a million cancer-related deaths in the United States. Over the past several decades, the association between schizophrenia and cancer risk has been an area of increasing interest and debate. Although some studies have shown an increased risk of cancer associated with schizophrenia, others have found the opposite, showing reductions in the age-adjusted rate of cancer versus the general population. The latter findings have led to speculation that schizophrenia may confer some sort of protective effect, possibly genetic, that makes development of some cancers less likely, suggested to be similar to the protective effect conferred by sickle trait on malaria infection. This hypothesis is supported by analyses that suggest reduced incidence of multiple types of cancer in the first-degree relatives of schizophrenia patients. However, it is crucial to note that virtually all research reports concerning causes of death for schizophrenia patients use age-standardized or age-corrected analyses, rather than calculating age-specific death rates, that is death rates by cause within separate, progressive age groups (e.g., 21 to 25, 26 to 30, 31 to 35). Given the increased overall mortality in persons with schizophrenia, with a mean age of death typically in the 50s and largely related to CVD, the probability that schizophrenia patients will live long enough to clinically express many cancers is significantly reduced. Data comparing cause of death in patients and the general population by age-specific group would be very useful to clarify this issue, along with the broader use of a registry approach to better track mortality in this population. In the largest cohort study to date, from Olfson and colleagues, overall mortality related to cancer was second only to CVD, with lung cancer as the single largest specific contributor, with an SMR of 2.4 and a mortality rate of 74.8 per 100,000 person-years. Colon (SMR = 1.7; mortality rate = 14.1 per 100,000 person-years) and breast cancer (SMR = 1.6; mortality rate = 20.7 per 100,000 person-years) had the next highest rates, suggesting the hypothesis that smoking and obesity may be important risk factors.

Obesity- and smoking-related cancers have previously been reported to be more prevalent in schizophrenia patients. One registry study by Grinshpoon and colleagues evaluated the incidence of various types of cancer in 33,372 individuals with schizophrenia in Israel. The overall standard incidence ratio (SIR, defined as the ratio of incidence among schizophrenia patients to incidence in the general population, similar to SMR) of cancer was lower (0.86 for men and 0.91 for women) in people with schizophrenia. Three key areas of significantly increased risk were identified in the schizophrenia sample: The SIR for lung cancer was 1.38 in men, the SIR for breast cancer was 1.11 in women, and the SIR for cancer of

the corpus uteri was 1.64 in women. When the sample was segmented based on place of birth, some groups were found to have substantially increased risk. Women who were born in either Africa or Asia had an SIR of 1.37 for developing breast cancer, while their male counterparts had an SIR of 1.58 for lung cancer. Israeli-born women had a further elevated risk of cancer of the corpus uteri, with an SIR of 2.75. The results suggest that obesity- and smoking-related cancers may be more common in patients with schizophrenia and readily detected despite premature mortality from CVD in this population.

An elevated rate of breast cancer among women with schizophrenia has previously been reported. In a very large retrospective cohort study comparing outcomes in 52,819 women exposed to antipsychotics versus 55,289 controls, Wang and colleagues found that women treated with antipsychotics had a 16 percent higher risk of developing breast cancer than women in the control group. An increased risk of breast cancer may relate to the prolactin-elevating effects of commonly used antipsychotic medications, but may also relate to an increased incidence of insulin resistance, type 2 diabetes mellitus, and related metabolic abnormalities. In a study of 512 women with early stage breast cancer, who had not been diagnosed with diabetes, Goodwin and colleagues studied the association between fasting insulin levels and cancer outcomes. Insulin levels in the sample ranged from 8 to 340 pmol/L. They found that women with insulin levels in the uppermost quartile had double the risk of distant cancer recurrence and triple the risk of death compared with those who had fasting insulin levels in the lowest quartile. Given the elevated rates of obesity, insulin resistance, and type 2 diabetes associated with increased risk of breast and other obesity- and insulin resistance–associated cancers (e.g., colon, endometrial, gallbladder) in the mentally healthy, it is reasonable to consider that this association also contributes to the risk for developing these cancers in patients with schizophrenia. Notably, another study by Dalton and colleagues showed no difference in rates of breast cancer between those taking neuroleptic medications and the general population. The methodological issues discussed above, with a lack of age-specific comparisons, make these findings difficult to interpret. In any case, the study by Wang and colleagues mentioned above further indicated that the 52,819 women taking antipsychotics were significantly less likely than the controls to undergo appropriate screening for breast cancer, including mammograms and outpatient medical visits.

Pulmonary Disease

Chronic Lower Respiratory Disease, including COPD, is the third leading cause of death in the United States. In a sample population of outpatients with severe mental illness, including schizophrenia, Himelhoch and colleagues found that the overall prevalence of COPD, including chronic

bronchitis or emphysema, was 22.6 percent. This can be compared to a COPD prevalence of roughly 5 percent in a group of NHANES III general population controls matched for age, race, and gender to the study sample. The patients with severe mental illness had a 19.5 percent rate of chronic bronchitis as compared to 6.1 percent in the general population sample and a 7.9 percent rate of emphysema versus 1.5 percent in the general population sample. People with schizophrenia are also reported to be more than twice as likely as members of the general population to develop asthma. In addition to the increased mortality from pulmonary cancer mentioned above, increased mortality due to noncancer-related pulmonary disease was observed in a large Medicaid cohort of over 1 million schizophrenia patients followed for almost 5 million person-years. Elevated SMRs and absolute mortality were noted for COPD (SMR = 9.9; mortality rate = 89.5 per 100,000) and influenza and pneumonia (SMR = 7.0; 33.3 per 100,000).

Increased abdominal adiposity, seen in many patients with schizophrenia, can contribute to the risk of COPD, just like in the general population, where increased abdominal mass restricts diaphragmatic excursion among other effects. However, the strongest independent risk factor for COPD is smoking. People with serious mental illness (SMI), including individuals with schizophrenia, are much more likely to smoke than the general population, according to data from the National Association of State Mental Health Program Directors (NASMHPD). In a recent large study that included over 5,000 individuals with schizophrenia, Hartz and colleagues estimated the rate of lifetime daily smoking at 72 percent, which was more than double that of the general population. The mentally ill are also more likely to be exposed to the effects of secondhand smoke.

ORAL HEALTH IN SCHIZOPHRENIA

People with schizophrenia have poor dental health, among the poorest reported in the literature. Reports suggest that patients with schizophrenia brush their teeth irregularly, leading to functional difficulties such as difficulty with chewing and eating, in addition to increased risk for decay, infection, and inflammation.

INCREASED PREVALENCE OF MODIFIABLE RISK FACTORS IN PATIENTS WITH SCHIZOPHRENIA

Key risk factors for major causes of morbidity and mortality that form the basis of public health efforts in the general population including overweight or obesity, smoking, hypertension, dyslipidemia, and hyperglycemia, are all more prevalent in schizophrenia. The success of efforts aimed at addressing the excess burden observed in persons with schizophrenia is therefore highly likely to depend on the success of primary and secondary prevention

efforts targeting modifiable risk factors.

Overweight and Obesity

Increased adiposity is associated with increased morbidity and mortality from hypertension, dyslipidemia, type 2 diabetes mellitus, coronary artery disease, stroke, and certain types of cancers, including colon, breast, endometrial, and gallbladder cancer. Both overweight and obesity can be defined by body mass index (BMI), calculated as weight in kilograms divided by the square of height in meters (the latter is intended to be proportional to body surface area). Persons with a BMI greater than 25 kg/m^2 are considered overweight, and a BMI greater than 30 kg/m^2 generally defines obesity. At present, approximately 35 percent of adults in the United States are obese, and the prevalence of obesity is increasing both nationally and worldwide. Obesity is roughly twice as prevalent in people with schizophrenia, in comparison to the general population. Evidence from numerous studies characterizes the problem. Substantial evidence suggests that treatment with psychotropic medications, particularly certain antipsychotic medications, leads to increases in body weight and adiposity. Medications that include certain antipsychotics with clinically significant antagonism for histamine (H_1) receptors, and to a lesser extent, α_1-adrenoceptors, can increase risk of weight gain. In general, pharmacotherapies that decrease energy expenditure, via sedation or reduced motor activity, or increases caloric intake, via increased appetite or reduced satiety, can increase risk for body weight gain. Rates of significant short-term weight gain, defined as a >7 percent increase in body weight over baseline, are available from pooled registration trial data reported in the US packaging insert for available medications. For example, the reported rates of incident weight gain during registration trials using this definition are 29 percent for olanzapine (vs. 3 percent for placebo), 23 percent for quetiapine (vs. 6 percent for placebo), 18 percent for risperidone (vs. 9 percent for placebo), 10 percent for ziprasidone (vs. 4 percent for placebo), 8 percent for aripiprazole (vs. 3 percent for placebo), and 2.4 percent for lurasidone (vs. 0.7 percent for placebo).

Longer-term data show a similar pattern, again with more potent antihistaminic medications associated with more substantial weight gain. Over 1 year of treatment, again using pooled registration clinical trial data, olanzapine treatment is associated with a mean 12 kg of weight gain at doses in the range of 12.5 to 17.5 mg per day, while quetiapine produces a mean 3.2 kg increase, risperidone 2.2 kg, and ziprasidone and aripiprazole approximately 1 kg. It should be noted that the 1-year data for quetiapine may not be directly comparable to data for the other drugs, as available quetiapine data are based on “observed case” analyses where all patients completed the 1-year exposure, in contrast to a commonly reported “last

observation carried forward" method of analysis for the other medications (with later generally underestimating treatment-related weight change). The totality of available data suggests a rank order of risk of weight gain with second-generation antipsychotics, with clozapine and olanzapine offering the greatest risk, followed by quetiapine, risperidone, and asenapine offering much lower but ordinally "intermediate risk." Aripiprazole, ziprasidone, and based on available data, lurasidone have the most limited effects on weight. It should be noted that treatment with individual typical antipsychotics and other psychotropic medications including mood stabilizers and antidepressants, can also be associated with clinically significant weight gain. Complex polypharmacy, including common use of antipsychotics in combination with agents like valproic acid, can substantially increase the risk of weight gain.

Hypertension

Schizophrenia patients are about twice as likely to have hypertension as are members of the general population. Estimates of the rate of hypertension in schizophrenia patients range from 19 to roughly 47 percent. By comparison, approximately 15 percent of the general population has hypertension. Goff and colleagues reported that 27 percent of CATIE patients with schizophrenia had hypertension at baseline, compared with 17 percent in age- and gender-matched controls from NHANES III.

Increased adiposity and insulin resistance are associated with increased sympathetic nervous system activity and sodium retention, both of which can increase the risk of hypertension. The same risk factors that contribute to obesity and insulin resistance in schizophrenia patients can therefore increase the risk for hypertension, including treatment with medications that produce substantial weight gain. Although measureable hypertension can be attenuated by the α_1-adrenoceptor antagonism associated with some medications, antipsychotic treatment-induced hypertension has been observed in context of substantial weight gain. Cigarette smoking may also be associated with increased blood pressure both acutely due to the stimulant effects of nicotine, and secondarily due to chronic smoking increasing risk of various forms of CVD.

Dyslipidemia

Another important risk factor for CVD, dyslipidemia, has increased prevalence in schizophrenia. In the 1980s, the Multiple Risk Factor Intervention Trial established that the relationship between elevated serum cholesterol and death from CHD is continuous, incremental, and strong. Heiskanen and colleagues evaluated patients with schizophrenia to determine the frequency of the metabolic syndrome, reporting that 31 percent had serum triglyceride levels greater than 149 mg/dL. In addition,

58 percent of men were at risk with low serum high-density lipoprotein (HDL) cholesterol levels, and 25 percent of women had low HDL cholesterol levels. Among patients entering the CATIE study, 50.7 percent had elevated triglyceride levels, and 48.9 percent had low HDL cholesterol levels.

Some antipsychotic medications are associated with clinically significant increases in serum lipid concentrations, generally, but not exclusively, in association with their propensity to produce weight gain. In phase I of the CATIE study, olanzapine was associated with the greatest worsening of lipid parameters from baseline. Patients switched to olanzapine experienced an exposure-adjusted mean triglyceride increase of 40.5 mg/dL, and a mean increase in total cholesterol of 9.4 mg/dL, in contrast, for example, to exposure-adjusted mean decreases of 16.5 mg/dL and 8.2 mg/dL, respectively, in patients switched to ziprasidone. Notably, 23 percent of patients randomized to olanzapine in phase 1 had already been taking olanzapine prior to the study, and therefore they may have already experienced some of the drug-induced effect on plasma lipids, attenuating observed change during the study. In CATIE phase II, where the study methodology prevented assignment to a treatment arm that the patient was already taking in phase I, initial drug effects on lipids were more apparent. For example, patients in phase II who were assigned to olanzapine showed even greater increases in exposure-adjusted mean triglycerides (94.1 mg/dL) and total cholesterol (17.5 mg/dL) than seen in phase I. Pooled studies suggest an overall prevalence of decreased HDL (M < 40 mg/dL; F, 50 mg/dL) and triglyceride increases >150 mg/dL of approximately up to 40 percent in schizophrenia.

Insulin Resistance and Hyperglycemia

The prevalence of insulin resistance, variously defined, in patients with schizophrenia has been estimated at 1.5 to 2 times the prevalence in the general population. Hyperglycemia in schizophrenia (glucose >100 mg/dL) is common, affecting up to 20 percent of patients in pooled samples. The progression of insulin resistance to pancreatic β-cell failure and related hyperglycemia can progress to incident type 2 diabetes mellitus, which increases risk for morbidity and mortality from microvascular (i.e., retinopathy, neuropathy, and nephropathy) and macrovascular disease (i.e., CHD, cerebrovascular disease, peripheral vascular disease) complications. Antipsychotics increase risk of insulin resistance and hyperglycemia in patients with schizophrenia, with substantial evidence for the contribution of medication effects to risk and limited evidence suggesting risk for hyperglycemia in this population may be related to the disease state. Lifestyle factors can certainly affect outcomes, with poor diet, inactivity, smoking, and lack of exercise commonly associated with overweight and obesity and insulin resistance in untreated nonpsychiatric

samples as well as in treated patients.

Different antipsychotic medications are associated with different levels of risk for insulin resistance and hyperglycemia, generally proportional to their association with drug-induced weight gain, although additional adiposity-independent effects may contribute to risk with some medications. In phase I of the NIH-funded CATIE study, plasma glucose levels increased by a mean 15.0 ± 2.8 mg/dL in patients randomized to olanzapine, in comparison, for example, to 2.3 ± 3.9 mg/dL in patients randomized to ziprasidone. Neither change was statistically significant in this study, but the direction and magnitude of effects is largely consistent across multiple studies. In general, the ability to detect drug-induced risk for hyperglycemia during the routine duration of most clinical trials is limited by the natural course of a condition like type 2 diabetes. The development of clinically significant insulin resistance is initially compensated for by increased insulin production by pancreatic β-cells and reduced clearance of insulin, with resulting compensatory hyperinsulinemia often capable of buffering the onset of clinically detectable hyperglycemia for several years. Only when the β-cell secretory capacity fails sufficiently will clinical hyperglycemia occur, and the time course and likelihood of this will depend on patient factors that include genetic risk, pre-existing level of insulin resistance and β-cell function prior to the period of observation. Although antipsychotic medications may confer most of their risk for worsening insulin resistance through increases in body weight and adiposity, evidence from controlled studies suggests the existence of some additional, adiposity-independent treatment effects on insulin sensitivity and blood glucose (e.g., drug effects on glucose transporter function). For example, antipsychotic agents such as clozapine and olanzapine achieve relatively high intracellular concentrations where they can interact with intracellular glucose transporter proteins and reduce glucose transport from extracellular to intracellular space, potentially contributing to insulin resistance (e.g., reducing insulin-stimulated glucose uptake at muscle) and thereby risk for hyperglycemia, independent of any drug effects on adiposity.

Metabolic Syndrome

The metabolic syndrome is not a separate entity from the other modifiable risk factors discussed in this section. It can be described as a constellation of several risk factors, interrelated by insulin resistance, associated with risk for CVD and diabetes-related morbidity and mortality, or cardiometabolic risk. Using a widely accepted definition of the metabolic syndrome, from the third adult treatment panel (ATP III) of the National Cholesterol Education Program (NCEP), three or more of the following are required for diagnosis: Abdominal obesity (waist circumference greater than 102 cm in men or greater than 88 cm in women), elevated

triglycerides (150 mg/dL or more), decreased HDL cholesterol (less than 40 mg/dL in men or less than 50 mg/dL in women), hypertension (blood pressure 130/85 mm Hg or more), or hyperglycemia (fasting plasma glucose [FPG] 100 mg/dL or more).

The components of the metabolic syndrome, as well as the composite syndrome, are major risk factors for type 2 diabetes mellitus and CVD. As discussed above, obesity, dyslipidemia, hypertension, and type 2 diabetes mellitus are all recognized as common conditions in people with schizophrenia. The current general population prevalence of the metabolic syndrome is approximately 23 percent in US adults, similar to other developed countries. By contrast, studies indicate that the prevalence of metabolic syndrome in patients with schizophrenia ranges from 30 to over 40 percent. One meta-analysis estimated the overall rate of the metabolic syndrome in schizophrenia is approximately 32.5 percent. Older age and illness duration increase the rate of metabolic syndrome. In the CATIE study, McEvoy and colleagues found the rate of metabolic syndrome in patients at study baseline to be approximately 41 percent. They also studied the prevalence of individual metabolic syndrome criteria at baseline in CATIE, comparing the results with an age-matched general population sample from NHANES III, and found that all criteria for metabolic syndrome were present at elevated rates in CATIE patients of both sexes, with the sole exception of hyperglycemia in male patients (Table 12.14–2), the latter likely related to the relatively young age of the sample where greater pancreatic β-cell function can be observed despite insulin resistance.

Table 12.14–2.
Comparison of Metabolic Syndrome and Individual Metabolic Syndrome Criterion Prevalence in Fasting CATIE Subjects and Age-, Gender-, and Race/Ethnicity-Matched NHANES III Subjects

	Males			Females		
	Fasting CATIE Subjects	**Matched NHANES Subjects**	***P***	**Fasting CATIE Subjects**	**Matched NHANES Subjects**	***P***
	N = 509	*N* = 509	—	*N* = 180	*N* = 180	—
Metabolic Syndrome Prevalence	36.0%	19.7%	0.0001	51.6%	25.1%	0.0001
Waist Circumference Criterion	35.5%	24.8%	0.0001	76.3%	57.0%	0.0001
Triglyceride Criterion	50.7%	32.1%	0.0001	42.3%	19.6%	0.0001
HDL Criterion	48.9%	31.9%	0.0001	63.3%	36.3%	0.0001
BP Criterion	47.2%	31.1%	0.0001	49.6%	26.8%	0.0001

Glucose Criterion	14.1%	14.2%	0.9635	21.7%	11.2%	0.0075

CATIE, Clinical Antipsychotic Trials in Intervention Effectiveness; NHANES, National Health and Nutrition Examination Survey; HDL, high-density lipids; BP, blood pressure.

From McEvoy JP, Meyer JM, Goff DC, et al. Prevalence of the metabolic syndrome in patients with schizophrenia: Baseline results from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia trial and comparison with national estimates from NHANES III. *Schizophr Res.* 2005;80(1):19–32.

Antipsychotic medications are associated with risk of metabolic syndrome to differing degrees, again generally in proportion to their risk for weight gain. For example, pooled data from two 26-week randomized double-blinded studies, one comparing aripiprazole with placebo and the other comparing aripiprazole with olanzapine, showed incidences of metabolic syndrome of 19.2 ± 4.0 percent for olanzapine, 12.8 ± 4.5 percent for placebo, and 7.6 ± 2.3 percent for aripiprazole.

Smoking and Substance Abuse

Cigarette smoking is a well-known risk factor associated with all the major disease classes described in this chapter. The recent US Surgeon General's report details the number of premature deaths associated with smoking and second-hand smoke exposure from 1964 to 2014. Since the first report on smoking 50 years ago, there have been close to 21 million smoking-attributable deaths in the United States. Forty-eight percent of these deaths were due to CVD and metabolic diseases, 33 percent from smoking-related cancers, and 18 percent from pulmonary diseases. While the overall rate of current smoking in the United States declined during this time from 42 to 18 percent, the authors are unaware of reports of the comparable trend in cigarette use for individuals with schizophrenia, but evidence suggest that these individuals continue to smoke at a much greater level than the general population, and thus remain at elevated risk. Studies, such as that by Hartz and colleagues, indicate that over 70 percent of individuals with schizophrenia have a lifetime history of daily smoking, as opposed to 29 percent of the general population. According to a study by Brown and colleagues that examined the factors associated with 25-year mortality in 370 patients with schizophrenia, smokers were twice as likely to die during this time frame and 70 percent of deaths by natural causes were from diseases known to be related to smoking.

The co-occurrence of substance use, including nicotine, alcohol, marijuana, and other drugs is highly prevalent among individuals with schizophrenia. While Hartz and colleagues found that lifetime daily smoking was the most prevalent substance reported in individuals with schizophrenia, rates of lifetime heavy drinking (28 percent), marijuana use (43 percent), and other illicit drug use (35 percent) were also more than

twice that of the general population (10, 16, and 10 percent, respectively). Heavy alcohol exposure may lead to alcoholic hepatitis and eventually cirrhosis, with secondary cognitive impairment via hepatic encephalopathy in advanced cirrhosis, caused by the accumulation of nitrogenous compounds that are inadequately metabolized by the liver. In the CNS, prolonged alcohol abuse leads to cerebellar degeneration in about 1 percent of the chronically alcohol dependent, presenting as unsteady gait and mild nystagmus. These symptoms are irreversible, and they are often accompanied by global cognitive decline from the direct or indirect effects of chronic alcohol consumption including head injury, nutritional (e.g., thiamine) deficiency, or direct neurotoxicity. Peripheral neuropathy is seen in approximately 5 to 15 percent of chronic alcoholics due to nutritional deficiency and the direct toxic effects of alcohol on neuronal axons. Cocaine and amphetamines have parallel health effects based on their similar pharmacological activities, which result in increased synaptic levels of dopamine and other monoamines. This results in sympathomimetic and vasoconstrictive properties, which are responsible for many of the acute and chronic medical effects of ingestion, including MI, cardiac arrhythmias, cerebrovascular accidents, vasospasm, and hypertension. Injection drug use also increases risk for infectious disease, particularly hepatitis C and HIV.

In patients with schizophrenia, substance-use disorders exacerbate the risk of medical comorbidities. Patients with schizophrenia and concomitant substance-use disorders have higher adjusted odds for heart disease, asthma, gastrointestinal (GI) disorders, skin infections, and respiratory disorders than do patients with schizophrenia without substance-use disorders.

PRIMARY AND SECONDARY PREVENTION

CVD, which includes CHD, stroke, and peripheral vascular disease, is poised to remain the leading cause of death in the United States and most developed Western countries for the foreseeable future. Together heart and cerebrovascular diseases were the underlying cause of death in 28.8 percent (725,270) of all 2,512,873 deaths in the United States in 2011, or 1 of every 3.5 deaths. However, there have been remarkable declines in CVD mortality in the US general population over the past several decades, from over 50 to approximately 30 percent, as the underlying cause of death, largely due to improved diagnosis and treatment of CVD as well as success in the area of primary prevention (i.e., reduced rates of smoking in the general population). By contrast, patients with schizophrenia lose up to 30 years of life expectancy, with the majority of the excess premature deaths due to CVD. To date, the benefits of improved prevention seen in the general population have not been extended to patients with schizophrenia.

In the general US population during the past decade, most of the

observed reductions in CVD mortality are attributable to improvements in the treatment of acute cardiovascular events and in long-term secondary prevention. For example, the decrease in the fatality rate for hospitalized MI has been attributed to the acute utilization of aspirin, thrombolytics, β-blockers, and angiotensin-converting enzyme (ACE) inhibitors. Longer-term post-MI use of aspirin, β-blockers, ACE inhibitors, and statins, as well as therapeutic lifestyle changes, have made additional important contributions to reductions in MI-related mortality. Unfortunately, patients with schizophrenia who experience an acute MI are significantly less likely than the general population to receive drug therapies of proven benefit, including thrombolytics, aspirin, β-blockers, and ACE inhibitors. Patients with schizophrenia are also significantly less likely than members of the general population to undergo cardiac catheterizations and receive emergency angioplasties or coronary artery bypass grafts. Druss and colleagues performed a study of over 88,000 Medicare patients hospitalized for MI. Their data showed that mortality in the follow-up period was increased by 19 percent in the presence of any mental disorder and by 34 percent in persons with schizophrenia, with these increases in mortality attributable to reductions in the quality of care.

In the general US population, cigarette smoking has been the leading avoidable cause of all premature deaths. However, smoking rates in the general population have declined considerably over the past several decades, from over 50 percent in the 1950s to around 18 percent at present. Using a general population-based sample, Grant and colleagues estimated that individuals with a psychiatric disorder consume 34 to 46 percent of all cigarettes in the United States. As noted above, over 70 percent of schizophrenia patients have been estimated to be smokers in some samples. Although patients with major mental illness are overrepresented in state programs such as Medicaid, only some states cover all the smoking cessation treatments recommended in the U.S. Public Health Service Task Force guidelines. Further, there are often barriers to receiving these smoking cessation treatments, including requiring prior authorization and copayments and limiting the number of courses of treatment, all which create additional hurdles for disabled patients with fewer resources and fixed incomes.

The evidence base for the smoking cessation treatments recommended in the clinical practice guidelines is derived from clinical trials of smokers generally selected to have no history of severe psychiatric illness. Recent reviews, including a recent Cochrane review by Tsoi and colleagues, provide support that some of the recommended smoking cessation treatments also show efficacy in smokers with stabilized schizophrenia. They pooled results from seven trials that compared smoking cessation effects of bupropion relative to placebo in smokers with schizophrenia. Results suggested that bupropion increased rates of quitting smoking by

over three times that of placebo, both at the end of treatment and at 6-month follow-up, without obvious worsening of psychiatric symptoms. Pooling results from two studies that compared varenicline to placebo found varenicline to increase rates of quitting by over four times that of placebo at the end of treatment, but the evidence for long-term efficacy and a tolerable safety profile for varenicline in smokers with schizophrenia is still mounting. Surprisingly, comparable systematic studies on the effects of nicotine replacement and behavioral therapy on cessation in the population of smokers with schizophrenia appeared to be lacking.

Persons with schizophrenia have elevated rates of modifiable risk factors for CVD, including approximately 1.5 to 2 times the general population prevalence of diabetes, dyslipidemia, hypertension, smoking, and obesity. Some portion of the high prevalence of modifiable CVD risk factors can be explained by underdiagnosis and undertreatment. Nasrallah and colleagues studied the rates of treatment for existing modifiable risk factors among schizophrenia patients entering the CATIE study. Among the approximately 1,500 patients entering the CATIE study, which was conducted at 57 US sites spanning a range of academic and public sector treatment settings, 88 percent of those patients with dyslipidemia were receiving no lipid-lowering pharmacotherapy, 30 percent of those with diabetes mellitus were receiving no antidiabetes medications, and 62 percent with hypertension were receiving no antihypertensive medication (Fig. 12.14–1).

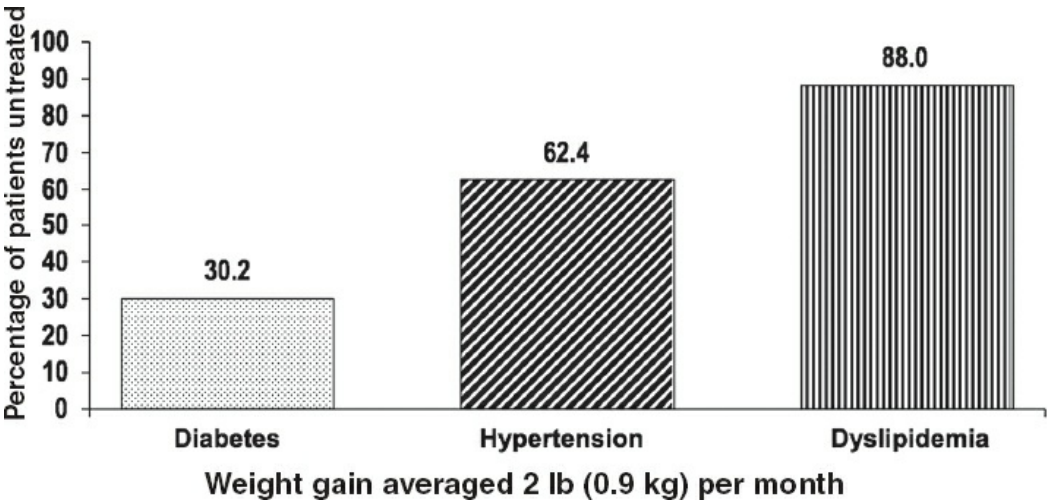

FIGURE 12.14–1. Prevention opportunities missed: Low rates of treatment for metabolic disorders in schizophrenia in the Clinical Antipsychotic Trials in Intervention Effectiveness (CATIE). Percentage of patients with diabetes mellitus, hypertension, or dyslipidemia receiving no treatment for those conditions at the baseline assessment time, print in the CATIE schizophrenia trial (*N* = 1,460). (From McEvoy JP, Meyer JM, Goff DC, et al. Prevalence of the metabolic syndrome in patients with schizophrenia: Baseline results from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia trial and comparison with national estimates from NHANES III. *Schizophr Res.* 2005;80(1):19–32.)

- Secondary safety and tolerability outcomes were evaluated at months 1, 3, 6, 9, 12, 15, and 18. This included changes in weight and laboratory analysis.
- The results for weight change are based on 185 ziprasidone patients, 341 risperidone patients, 337 quetiapine patients, 336 olanzapine patients, and 261 perphenazine patients.
- Patients in the olanzapine group demonstrated more weight gain than patients in any other group.
- Weight gain averaged 2 lb (0.9 kg) per month.

Screening for metabolic risk occurs at very low rates in patients with schizophrenia. This phenomenon is observed even in the context of treatment with antipsychotic medications associated with risk for disturbances in glucose and lipid metabolism, despite recommendations by the American Psychiatric Association, the American Diabetes Association and other groups, and FDA warnings. The frequency of baseline glucose and lipid monitoring varies from 8 percent to less than 30 percent, and follow-up assessments done in only 8.8 percent of patients. A large cohort study by Morrato and colleagues, involving Medicaid claims data for 55,436 enrollees with mental illness from several US states, showed that in the 4 months before and after a new antipsychotic prescription, less than one-third of patients overall received any plasma glucose measurement and less than 10 percent received any plasma lipid measurement. When patients with an existing diagnosis of diabetes were excluded from the analysis, glucose testing dropped to 10 to 15 percent of patients, and lipid measurements to only 5 percent of patients, suggesting that prevention-oriented screening was particularly uncommon. Such low levels of screening may contribute to the observed low levels of diagnosis and treatment for modifiable CVD risk factors in patients with schizophrenia. NCEP guidelines recommend that patients with diabetes mellitus be treated as aggressively as patients with prior MI or stroke who do not have diabetes. However, patients with diabetes and schizophrenia are less likely than nonmentally ill diabetes patients to receive standard-of-care treatments.

Based on a consensus development statement from the American Diabetes Association, patients initiating antipsychotic therapy, as well as those changing antipsychotic treatment, are appropriate targets for a baseline assessment of family and personal medical history, as well as an assessment of metabolic risk indicators like weight (BMI) or waist circumference, blood pressure, FPG, and fasting lipid measurements (total, low-density lipoprotein [LDL] and HDL cholesterol, and triglycerides), and possibly glycosylated hemoglobin (HbA_{1C}), with repeat measurements of weight at every visit thereafter. Regular measurements of weight can provide important information on treatment-related risk, requiring only an office scale. Laboratory tests should be repeated approximately every 3 months into initial treatment and at least annually thereafter, or more often in the setting of increased risk. All patients with schizophrenia should

have waist circumference, BMI, and blood pressure recorded at least every 6 months, with more frequent screening for higher-risk patients. For patients with hypertension, for example, the National Heart, Lung, and Blood Institute (NHLBI) recommends monthly monitoring of blood pressure following the initiation of antihypertensive therapy, with follow-up every 3 to 6 months once blood pressure has been stabilized.

Ideally, medical comorbidities that might arise during treatment should be prevented rather than treated after the fact. Clinicians should strongly consider taking cardiometabolic risk into account when selecting psychotropic agents. It should be noted that in October 2007, the FDA prescribing information for olanzapine was altered to include warnings about higher risk of weight gain, hyperglycemia, and hyperlipidemia, with an ongoing class warning that includes all atypical antipsychotic medications.

When elevated cardiometabolic risk is discovered, secondary causes of risk should be addressed if possible, according to the NCEP. This might be achieved, for example, by switching a patient from an agent with a higher to a lower risk of weight gain and dyslipidemia. In addition, behavioral therapy can be an effective method for overweight or obese patients with schizophrenia to manage weight gain. If these interventions fail, treatment of risk factors such as dyslipidemia with a medication such as an HMG-CoA (3-hydroxy-3-methyl-glutaryl-CoA) reductase inhibitor or other lipid-lowering therapy should be initiated.

Given the differential weight gain associated with different antipsychotic medications, it could be predicted that reductions in weight might be associated with a switch from an agent with high weight gain risk to one with lower risk, but not with switches between two treatments of similar weight gain liability. Indeed, these predictions have been confirmed in clinical trials involving switches to aripiprazole and to ziprasidone. In a recent example of a longer-term study, patients were switched to ziprasidone from risperidone, olanzapine, or a high-potency first-generation antipsychotic. After 52 weeks of treatment with ziprasidone, patients switched from olanzapine had a mean 9.8 kg weight loss versus 6.9 kg in those switched from risperidone and little change in the switch from agents like haloperidol, providing an important proof of concept regarding the prediction of weight change during therapeutic substitutions.

Investigations are under way into disparate strategies to address obesity in this population or to reverse iatrogenic weight gain associated with antipsychotic medications. A 2007 Cochrane database review by Faulkner and colleagues studied the effects of pharmacological and nonpharmacological intervention strategies. Notably, the strategy of switching antipsychotic medications was not included in the review but recent evidence from two well-powered randomized clinical trials supports the safety and effectiveness of switching antipsychotic medications under

certain circumstances to achieve weight loss. Adjunctive pharmacology with various agents was found to yield generally underimpressive results, based on limited data, typically from one or two studies for each agent with small sample sizes. No single agent was found to yield consistently superior results. The various adjunctive agents studied were amantadine (Symmetrel), D-fenfluramine, dextroamphetamine sulfate, nizatidine, phenylpropanolamine, sibutramine, topiramate, and fluoxetine. Among nonpharmacological interventions, cognitive-behavioral therapy yielded encouraging results in two studies, with significantly greater weight reduction than standard care.

Increasing evidence points to excessively low physical fitness in this population, representing another modifiable CVD risk factor and an intervention opportunity. Various fitness interventions, successful at managing metabolic comorbidities in the general population, such as the Diabetes Prevention Program, or LOOK AHEAD, clearly indicate the value of cardiovascular exercise, but have systematically excluded patients with schizophrenia. Small exercise interventions among patients with schizophrenia show feasibility in this population, and larger trials will more accurately define the value of physical exercise, currently clearly underutilized, in the management of metabolic comorbidities in schizophrenia.

At this time, switching antipsychotic medications may be one of the most effective strategies for dealing with antipsychotic-associated weight gain. Weight loss has been found to extend at least over a 1-year time frame, with concordant improvements in indices of cardiovascular risk such as fasting lipid profiles. An early study by Newcomer et al. established proof of concept for the safety and weight-loss-efficacy of switching overweight or obese patients on olanzapine to a low weight gain-risk antipsychotic like aripiprazole. In addition, the CATIE investigators' Comparison of Antipsychotics for Metabolic Problems (CAMP) study, randomly assigned overweight patients (BMI 27 or greater) with elevated non-HDL cholesterol to either stay on their existing antipsychotic medication (olanzapine, risperidone, or quetiapine) or switch to aripiprazole. The prespecified primary analysis included 89 switchers and 98 stayers who had at least one postbaseline non-HDL cholesterol measurement. Non-HDL cholesterol decreased more for the switch group than for the stay group (−20.2 mg/dL and −10.8 mg/dL, respectively). Switching was associated with larger weight reductions (least squares mean = 2.9 kg) and a net reduction of serum triglycerides of 32.7 mg/dL. Twenty-two switchers (20.6 percent) and 18 stayers (17.0 percent) experienced protocol-defined efficacy failure. Forty-seven switchers (43.9 percent) and 26 stayers (24.5 percent) discontinued the assigned antipsychotic medication before 24 weeks, indicating that while switching may be an effective way to reduce metabolic comorbidity, treatment noncompliance

must also be taken into account.

Appropriate monitoring and intervention can reduce long-term medical costs, improve general medical outcomes, and improve quality of life among patients with schizophrenia. In 2000 Lisa Dixon and colleagues showed that more patients with schizophrenia who were being treated for type 2 diabetes mellitus were satisfied with their lives than were schizophrenia patients with the same comorbid status who were not receiving treatment for diabetes. Concern for the physical health of schizophrenia patients has also been hypothesized to improve the therapeutic alliance and potentially increase the effects of, and compliance with, psychiatric treatments.

Improvements in primary and secondary prevention offer the largest potential impact on medical mortality in schizophrenia patients, but to achieve such improvements, clinicians will need to be better educated regarding this area of risk and have better coordination between primary health care providers, endocrinologists, cardiologists, and psychiatrists. Careful attention, and future research, should be directed toward the appropriate role of cardiometabolic risk assessments during psychiatric treatment, as well as the role of adjunctive pharmacotherapies and behavioral and physical fitness interventions that might target excess weight, dyslipidemia, hypertension, hyperglycemia, and smoking in this population. Without a greater degree of collaboration in the future between primary health care providers, endocrinologists, cardiologists, and psychiatrists, the large burden of avoidable premature mortality from CVD in patients with severe mental illness is likely to continue and increase in severity.

RECOMMENDATIONS

Schizophrenia patients have elevated rates of medical comorbidities such as heart disease, cancer, and diabetes, in addition to other prevalent concomitant complications such as substance use. Addressing prevention, screening, referral, and treatment of medical comorbidities is essential to improving health outcomes among patients in this population. There is a need for strategies that can provide clear guidance to administrators and medical directors on how to incorporate changes into already existing medical and psychiatric systems.

In the current literature, mortality data are collected, analyzed, and presented in diverse, often noncomparable ways, and there should be a continuation of efforts to standardize these methods and to allow comparison of data across different populations (e.g., developed and developing countries) and practice settings (e.g., public sector and private). As increased mortality and morbidity from medical causes continue as major concerns in the outcome of schizophrenia, there is a need to establish large, prospective population-based data sets from different

countries in order to enable comparisons of risks and outcome from various geographical areas and to assess time trends in outcome.

Care providers have a particular responsibility to address those components of medical risk that are iatrogenic in origin (e.g., medication side effects). The Institute of Medicine has suggested that all psychiatric and medical systems involved in the care of patients with schizophrenia, including mental health, substance abuse treatment, general health care, and other services, need to effectively collaborate in an effort to coordinate care of their patients and eliminate gaps and redundancies in necessary services. This collaboration should include the establishment of structured and routine comorbidity risk assessments, in addition to scheduled monitoring of antipsychotic treatment for side effects and adherence issues, and medical comorbidities of schizophrenia should receive the same established interventions that are provided to the general population. Implementation of such screening before the initiation of antipsychotic treatment, and during treatment, will likely improve outcomes among individuals with schizophrenia.

Although there is likely no one perfect model for health care delivery to the seriously mentally ill, one important model is the integrated clinic in which medical care providers are on site, colocated with psychiatrists and other team members. Early programs suggest that such an approach can improve service delivery and medical outcomes without increasing total costs. This can include programs designed to modify the behavior of both patients and health care providers in order to increase access to effective medical care. This approach can also involve increases in primary prevention, as well as diagnostic and treatment programs specifically tailored to this population.

Regardless of the structure of health care delivery, the psychiatrist will need to play a central role in a schizophrenia patient's overall care. Such a role may involve coordination of care with a patient's primary care physician. During the first visit between a treating psychiatrist and a patient, it is important for the psychiatrist to review results of the most recent physical and laboratory examination, if available. If such records do not exist, a medical screening should be arranged, ideally with the assistance of a case manager.

The American Psychiatric Association recently gave its Gold Achievement Award to the State of Missouri for successful implementation of enhanced medical screening and improvement of physical health outcomes within a HealthCare Home model. In 2011, Missouri (MO) became the first state in the United States to establish HealthCare Homes, as defined under the Affordable Care Act, for individuals with psychiatric conditions. The HealthCare Home Initiative is a CMS-approved collaborative effort between MO HealthNet, DMH, the MO Coalition for Behavioral Healthcare, and the MO Primary Care Association. The program

was developed to provide coordinated psychiatric and primary medical care to individuals with an SMI and ≥1 chronic medical condition (diabetes, asthma/COPD, CVD, BMI >25, tobacco use, developmental disability) receiving community-based psychiatric rehabilitation services in one of over 100 designated sites throughout MO. Enrollment in the program resulted in measureable improvements in BP, glucose, and cholesterol for individuals with diabetes, with >60 percent of the population with diabetes having normal HbA1c and BP values.

There is a crucial need for new paradigms like this in the prevention and treatment of medical illness in patients with schizophrenia, including closer attention to choice of psychotropic drug treatment regimens and more aggressive use of monitoring and interventions to identify and reduce risk. In the late 1990s, the US surgeon general concluded that mental health is "inextricably intertwined" with general medical health, and high-quality mental health systems must therefore integrate medical care with psychiatric treatment. Similarly, an Institute of Medicine report concerning the need to improve overall health care for patients with mental illness underscored the importance of integrating or colocalizing psychiatric and medical services. However, efforts to improve coordination of services face short- and long-term challenges, ranging from fiscal concerns to lack of awareness among health care providers who might play key roles.

The primary care physicians and specialists who provide most of schizophrenia patients' cardiovascular health care are likely to be less familiar than psychiatrists and other mental health care providers with the existing situation, or the potential solutions. Improvements in secondary prevention will require better education and involvement of primary health care providers, endocrinologists, and cardiologists to reduce disparities in the level and quality of treatment services received by patients with schizophrenia needing treatment for CVD, CVD risk-equivalent conditions like diabetes, and other CVD risk factors. In the short term, the existing mental health care system will have to put forth a significant effort, reallocating existing resources to coordinate screening, interventions, including needed referrals, and follow-up monitoring. Mental health care providers and systems will need to improve working relationships with primary care and specialty collaborators, including proactive efforts to facilitate evaluation and follow-up for patients. Future collaborations between primary health care providers, medical specialists, and psychiatrists, will be required to reduce the large burden of avoidable premature mortality in patients with schizophrenia, and to reduce health care disparities between schizophrenia patients and the general population.

References

Brown S, Kim M, Mitchell C, Inskip H. Twenty-five year mortality of a community cohort with schizophrenia. *Br J Psychiatry*. 2010;196(2):116–121.

*Charlson FJ, Baxter AJ, Dua T, Degenhardt L, Whiteford HA, Vos T. Excess mortality from mental, neurological, and substance use disorders in the Global Burden of Disease Study 2010. In: Patel V, Chisholm D, Dua T, Laxminarayan R, Medina-Mora ME, eds. *Mental, Neurological, and Substance Use Disorders: Disease Control Priorities*. 3rd ed (Volume 4). Washington (DC): The International Bank for Reconstruction and Development/The World Bank; 2016. Chapter 3.

Colton CW, Manderscheid RW. Congruencies in increased mortality rates, years of potential life lost, and causes of death among public mental health clients in eight states. *Prev Chronic Dis*. 2006;3:A42.

Dalton SO, Johansen C, Poulsen AH, et al. Cancer risk among users of neuroleptic medication: A population-based cohort study. *Br J Cancer*. 2006;95(7):934–939.

Goff DC, Sullivan LM, McEvoy JP, et al. A comparison of ten-year cardiac risk estimates in schizophrenia patients from the CATIE study and matched controls. *Schizophr Res*. 2005;80(1):45–53.

Grant BF, Hasin DS, Chou SP, Stinson FS, Dawson DA. Nicotine dependence and psychiatric disorders in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Arch Gen Psychiatry*. 2004;61(11):1107–1115.

Hartz SM, Pato CN, Medeiros H, et al. Genomic Psychiatry Cohort Consortium. Comorbidity of severe psychotic disorders with measures of substance use. *JAMA Psychiatry*. 2014;71(3):248–254.

Haupt DW, Fahnestock PA, Flavin KA, et al. Adiposity and insulin sensitivity derived from intravenous glucose tolerance tests in antipsychotic-treated patients. *Neuropsychopharmacology*. 2007;32:2561.

Leucht S, Burkard T, Henderson J, Maj M, Sartorius N. Physical illness and schizophrenia: a review of the literature. *Acta Psychiatr Scand*. 2007;116:317.

*Leucht S, Winter-van Rossum I, Heres S, et al. The optimization of treatment and management of schizophrenia in Europe (OPTiMiSE) trial: rationale for its methodology and a review of the effectiveness of switching antipsychotics. *Schizophr Bull*. 2015;41(3):549–558.

McEvoy JP, Meyer JM, Goff DC, et al. Prevalence of the metabolic syndrome in patients with schizophrenia: baseline results from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia trial and comparison with national estimates from NHANES III. *Schizophr Res*. 2005;80(1):19–32.

Mitchell AJ, Vankampfort D, Sweers K, van Winkel R, Yu W, De Hert M. Prevalence of metabolic syndrome and metabolic abnormalities in schizophrenia and related disorders – A systematic review and meta-analysis. *Schizopr Bull*. 2013;39(2):306.

Morrato EH, Newcomer JW, Allen RR, Valuck RJ. Prevalence of baseline serum glucose and lipid testing in users of second-generation antipsychotic drugs: a retrospective, population-based study of Medicaid claims data. *J Clin Psychiatry*. 2008;69:316.

Nasrallah HA, Meyer JM, Goff DC, McEvoy JP, Davis SM. Low rates of treatment for hypertension, dyslipidemia and diabetes in schizophrenia: data from the CATIE schizophrenia trial sample at baseline. *Schizophr Res*. 2006;86:15.

Newcomer JW. Second-generation (atypical) antipsychotics and metabolic effects: a comprehensive literature review. *CNS Drugs*. 2005;19(Suppl 1):1.

Newcomer JW, Hennekens CH. Severe mental illness and risk of cardiovascular disease. *JAMA*. 2007;298:1794.

Newcomer JW, Campos JA, Marcus RN, et al. A multicenter, randomized, double-blind study of the effects of aripiprazole in overweight subjects with schizophrenia or schizoaffective disorder switched from olanzapine. *J Clin Psychiatry*. 2008;69:1046.

Newcomer JW, Ratner RE, Eriksson JW, et al. A 24-week, multicenter, open-label, randomized study to compare changes in glucose metabolism in patients with schizophrenia receiving treatment with olanzapine, quetiapine and risperidone. *J Clin Psychiatry*. 2009;70(4):487–499.

Newcomer JW, Weiden PJ, Buchanan RW. Switching antipsychotic medications to reduce adverse event burden in schizophrenia: Establishing evidence-based practice. *J Clin Psychiatry*. 2013;74(11):1108–1120.

Olfson M, Gerhard T, Huang C, Crystal S, Stroup TS. Premature mortality among adults with schizophrenia in the United States. *JAMA Psychiatry*. 2015;72(12):1172–1181.

Osborn DP, Levy G, Nazareth I, Petersen I, Islam A. Relative risk of cardiovascular and cancer mortality in people with severe mental illness from the United Kingdom's general practice research database. *Arch Gen Psychiatry*. 2007;64:242.

*Patel V, Chisholm D, Parikh R, et al; DCP MNS Author Group. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. *Lancet*. 2016;387(10028):1672–1685.

Roerig JL, Steffen KJ, Mitchell JE. Atypical antipsychotic-induced weight gain: insights into mechanisms of action. *CNS Drugs*. 2011;25(12):1035–1059.

Saha S, Chant D, McGrath J. A systematic review of mortality in schizophrenia. Is the differential mortality gap worsening over time? *Arch Gen Psych*. 2007;64:1123.

Singleterry J, Jump Z, Lancet E, Babb S, MacNeil A, Zhang L. Centers for Disease Control and Prevention (CDC). State Medicaid coverage for tobacco cessation treatments and barriers to coverage—United States, 2008–2014. *MMWR Morb Mortal Wkly Rep*. 2014;63(12):264–269.

*Singleterry J, Jump Z, DiGiulio A, et al. State medicaid coverage for tobacco cessation treatments and barriers to coverage—United States, 2014–2015. *Morb Mortal Wkly Rep*. 2015;64(42):1194–1199.

Strassnig M, Clarke J, Mann S, Remington G, Ganguli R. Body composition, pre-diabetes and cardiovascular disease risk in early schizophrenia. *Early Interv Psychiatry*. 2015. doi: 10.1111/eip.12225. [Epub ahead of print]

Strassnig M, Signorile J, Gonzalez C, Harvey PD. Physical performance and disability in schizophrenia. *Schizophr Res Cogn*. 2014;1(2):112–121.

Tsoi DT, Porwal M, Webster AC. Interventions for smoking cessation and reduction in individuals with schizophrenia. *Cochrane Database Syst Rev*. 2013;2:CD007253.

Walker ER, McGee RE, Druss BG. Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. *JAMA Psychiatry*. 2015;72(4):334–341.

*The contents of this publication do not represent the views of the Department of Veterans Affairs or the United States Government.

▲ 12.15 Recovery in Schizophrenia

DINA HOOSHYAR, M.D., MARK RAGINS, M.D., AND FREDERICK J. FRESE, PH.D.

HISTORY OF RECOVERY

Early Background

Since biblical times people with serious mental illnesses were believed to be possessed by demons and treated accordingly. But in 1798 Philippe Pinel, often seen as the father of scientific psychiatry, began to characterize "mental derangements" as diseases in need of medical attention. Prior to the first comprehensive description of the disorder now called schizophrenia, Pinel described aspects of this syndrome.

Concerning the concept of recovery from this and similar conditions, Pinel held that some types of madness could have long and even permanent remission. Pinel's views and his humane *traitement moral* approach to caring for patients greatly influenced the development of treatment for the insane during the 19th Century, including the efforts of the Tukes in England and the reforming activities of Dorothea Dix in the United States.

During the 19th century, with such respectful approaches to treatment, success rates, or recovery from serious mental illnesses, were reported to range from 45 percent to as high as 90 percent, depending in large part on how soon the patients were admitted to hospitals after onset of their condition.

Return of the "Dark Ages"

Around the turn of the 20th century, hope for recovery began to significantly diminish for a certain category of persons with serious mental illnesses. The renowned German psychiatrist and nosologist, Emil Kraepelin, wrapped certain categories of patients whose illness he viewed as having a deteriorating course into a single classification and referred to them with the term, *dementia praecox* (early dementia). Even this name itself, connoted that this condition was one of inevitable deterioration. Kraepelin particularly emphasized the downward course of the disease and strongly suggested that those with this condition had little hope for recovery. Both Kraepelin and later Eugen Bleuler, the first to use the term *schizophrenia*, acknowledged that there might be occasional signs of improvement in these patients, but they tended to see these as unusual and temporary remissions.

While allowing for some slight possibility for improvement, these, and other early writers, overwhelmingly emphasized the high likelihood of deterioration for persons with schizophrenia. As a result, this pessimistic view of the condition held sway over mental health professionals' perceptions of this disease until fairly recently.

The Kraepelin/Bleuler perspective concerning the prospect of recovery from schizophrenia ushered in a particularly difficult era for persons given this diagnosis. On the one hand, some considered the lack of effective treatments for these patients to be a challenge and proceeded to devise a number of desperate attempts at cures for these hopeless individuals. These attempts included such innovations as having patients injected with malaria and being the recipients of frontal lobotomies, the developers of both these techniques having received Nobel Prizes for these innovations. Other problematic treatments during this time included insulin and metrazol shock therapies, and the application of "surgical bacteriology" to the patients at Trenton State Hospital, a technique that reportedly killed upwards of 30 percent of the recipients.

Also during this time, with the growing influence of the eugenics movement, a number of intellectuals began to argue that the evolution of the human race would be better served if society eliminated defective genetic material from the human gene pool. In the Western industrial societies, attempts at such elimination included the introduction of forced sterilization in the majority of US states and in many European countries. In Germany, the government initiated the infamous T-4 program, under which, officials engaged in the systematic murder of those considered to be the "deranged useless eaters" who were the patients in the psychiatric hospitals.

Emergence of Hope and the Rehabilitation Approach

After this dark half century, persons with schizophrenia began to see some hope with the introduction of anti-psychotic medications in the 1950s. Indeed these medications inspired sufficient optimism in policy-makers that a process of deinstitutionalization began, which stimulated the discharge to the community of the vast majority of the over half million persons who were in psychiatric hospitals in the 1950s. In reality, at the current time (2015) in the United States over 95 percent of those who were, or would have been, in government psychiatric hospitals in the 1950s are now living "in the community." However, as this move-to-the-community saga progressed, difficulties emerged in this massive reintegration attempt. Many of the discharged former patients could not find employment and were increasingly becoming homeless or being incarcerated as criminals.

In reaction, mental health professionals began to see a need for ex-patients to improve their functional abilities, particularly their social interaction skills, so that they could better fit into society. The term "psychosocial rehabilitation (PSR)" started to be used to describe an approach, which stressed "rehabilitation," opposed to medical "treatment" for the symptoms of these disorders. Importantly, stated principles of PSR included the concepts of "client choice," "strengths," and "empowerment of consumers."

Before long, it became apparent that the "psychosocial" aspects of these disorders could not be divorced from traditional medical considerations. Increasingly, the importance of wellness and a "holistic approach" became to be recognized and aspects of medical management were enveloped into the PSR concept. As attention to both medical and psycho-social factors began to merge, the term "psychiatric rehabilitation" was coined to reflect this integration of psycho-social and medical perspectives.

This psychiatric rehabilitation approach continued to focus on developing such skills as self-maintenance activities, vocational skills, use of public transportation, and social interaction skills. Importantly, the concepts of consumer choice and empowerment were increasingly seen as critical aspects of one's ability to function.

From Rehabilitation to the Recovery Approach

As the concept of mental health care was expanding to include input from the ex-patients, some of these persons in recovery began to organize. Local groups, many using the term "liberation" in their names, began in various parts of the United States in the early 1970s. In 1973, the first national meeting of such rights groups was held in Detroit. This gathering was referred to as the Conference on Human Rights and Psychiatric Oppression (CHRPO), which then met every year under that rubric through 1985 and continues to meet annually under the name, "Alternatives." From the beginning, members of these groups produced writings about their personal experiences of recovery. Increasingly these voices began to demand that their own perspectives and goals should take on more importance than just being add-on elements of treatment. Indeed many of the more strident voices characterized the treatment they had received as oppression, often viewing professionals as part of the problem. They also worked to have their views become a stronger force in society's perceptions of persons with mental illnesses. Judi Chamberlin, Sally Zinman, Howie The Harp, and others began to be heard and in time a national consumer movement developed.

As consumer voices became stronger, their calls for changes in society began to have an impact. Politically, consumers began to demand that society's discrimination and stigma toward mentally ill persons had to change. Many consumers saw their cause as being similar to the civil rights movements of other traditionally marginalized persons, especially that of the disability rights movement.

In response to these cries, in 1991 William Anthony, who had been a pioneer in championing a psychiatric rehabilitation model, expanded his view of this concept. Anthony's new vision included significantly more focus on the concerns of recovering persons. He issued a call for Recovery to become the "guiding vision" as to how to go about structuring the approach to caring for those with psychiatric disabilities. This vision has

evolved into what many observers characterize as the Recovery model. It is important to realize that by emphasizing the importance of consumer views, the emerging Recovery Movement necessarily embodied calls for the societal (political) changes, as well as calls for improved medical and rehabilitative care.

The cries from Recovering persons to have more say in their care have included calling for the inclusion of recovered persons on mental health policy making bodies, the creation of jobs for peer support specialists, the beginning of Recovery-oriented practice, the use of advanced directives, and numerous other innovations regarding the way persons with mental illnesses are treated and viewed by society.

DEFINING RECOVERY AND ITS EVOLUTION

By the 1980s and 1990s, severe problems were emerging in the public community mental health system and the dream of deinstitutionalization was fading. Failures were often attributed to lack of funding and political will, but suspicions also began to arise that the "medical model" itself was inadequate to meet the needs of community based mental health services. Fueling the dissatisfaction were outcome studies that showed that, contrary to common knowledge and expectations, people with schizophrenia "Recovered" to live reasonably full lives, free from seriously disabling symptoms about two-thirds of the time. Why were people settling for stabilization and isolation if Recovery is not just possible, but probable? Maybe the diagnosis of schizophrenia should not impose a destiny of life-long treatment, medication, disability, and dependence.

Alternative programs and approaches began to be developed to pursue better outcomes and lives:

- Beginning with Fountain House in New York, clubhouses and psychosocial rehabilitation programs addressed people's loss of usefulness and meaningful, productive roles.
- Psychiatric rehabilitation championed at UCLA by R.P. Liberman and at Boston University by Bill Anthony used behavioral and medical rehabilitation techniques to build a variety of skills including educational, employment, housing, and social skills.
- "Housing first" programs and "supported" programs with a variety of quality of life goals—e.g., housing, employment, education, parenting, physical health care—emerged that focused on people learning by doing instead of being given treatments to prepare them to do things.
- Integrated dual diagnoses approaches for people with both mental illnesses and substance abuse ranged from Fred Osher's early efforts at National Institute of Mental Health (NIMH) to Ken Minkoff's integrated systems of care, to "harm reduction," and Miller and Rolnick's "Motivational Interviewing."

- Trauma informed approaches were developed to reduce retraumatizing people while serving them and have been used to radically decrease seclusion and restraints in inpatient settings.
- Client choice, empowerment, and shared decision-making strategies were developed and used to enhance engagement with and participation in services.

Taken individually, each of these techniques adds to the "tool kit" of effective treatment approaches. More strikingly, however, was how they organically coalesced the Recovery Movement by the mid-1990s, no longer isolated "outsiders" but now part of a strong nationwide movement.

Previously within psychiatry, the concept of recovery had been used primarily within the context of substance abuse recovery. Now "Recovery" was being repurposed for serious mental illnesses.

The word "recovery" itself brings along with it a variety of traditions, histories, and values ranging from 12-step recovery, to recovery from physical illness, to trauma recovery, to spiritual recoveries. This broad spectrum of grass roots, "bottom up" influences gives the Recovery Movement an impressive diversity and practicality and widespread appeal. It also makes it unwieldy. Even defining "Recovery" has been difficult.

From a medical model point of view, people are recovering from the illness and its symptoms. For acute illnesses, this is usually accomplished by curing the illness directly. By contrast, recovery from chronic illnesses like schizophrenia depends more upon people learning to live well with their illnesses. For example, people can recover from a heart attack even though their heart muscle turns irreparably into scar tissue, if they take steps to protect their heart from future damage and if they are able to go on with their lives and do activities, like work and have sex, that are important to them.

Recovery from chronic medical conditions usually results from:

- Being able to maintain self-image and have hope
- Being able to maintain wellness and responsibility for self-care
- Being able to do things that make life meaningful
- Being able to replace professional supports with natural supports.

Two "consensus" approaches to defining Recovery have been put forward by Substance Abuse and Mental Health Services Administration (SAMHSA). The first, released in 2006, emphasized the "individual journey" aspect of Recovery: "Mental Health Recovery is a journey of healing and transformation enabling a person with a mental health problem to live a meaningful life in a community of his or her choice while striving to achieve his or her full potential." This approach emphasizes the individual's struggles and triumphs in overcoming mental illnesses and conceptualizes service providers as guides, coaches, and fellow travelers

promoting Recovery journeys.

The second approach is to define Recovery as achieving a quality of life despite the burdens of mental illness. SAMHSA's 2011 "working definition" of Recovery was: "A process of change through which individuals improve their health and wellness, live a self-directed life, and strive to reach their full potential." It goes on to define four domains to support a life in Recovery: *Health:* overcoming or managing one's disease(s) as well as living in a physically and emotionally healthy way; *Home:* a stable and safe place to live; *Purpose:* meaningful daily activities, such as a job, school, volunteerism, family caretaking, or creative endeavors, and the independence, income and resources to participate in society; and *Community:* relationships and social networks that provide support, friendship, love, and hope.

These two approaches to defining Recovery can be combined—Recovery is both the "journey" of overcoming serious mental illnesses and the "destination" of having a better life. A person can be "in Recovery" and/or "Recovered."

The Medical Model conceptualizes the process of recovery as:

- ► The fundamental problem is the symptoms of their mental illness.
- ► Diagnosis, assessments, treatment, monitoring, and achieving success are focused on symptom reduction.
- ► Once the illness is treated, other services may be needed to help people deal with residual symptoms, to keep symptoms away, and to rebuild lost functions or quality of life.
- ► Some fortunate people respond to treatment and other services to achieve recovery.

In contrast, the Recovery Model conceptualizes the process of Recovery as:

- ► Mental illnesses are a highly individualized, complex personal reaction to a variety of biological, emotional, developmental, traumatic, and social influences.
- ► Each person's outcome is more dependent upon how they personally react, adapt, and cope than it is dependent upon the type or severity of the symptoms themselves.
- ► Promoting positive personal reactions (e.g., maintaining hope, self-belief, managing risks, engaging and collaborating with helpers, taking self-responsibility, being motivated to build skills and supports, finding personal and social meanings, and avoiding isolation) throughout the course of each person's illness is crucial.
- ► Some fortunate people find enough inner strengths and resilience to make personal changes to achieve Recovery.

These differences in conceptualization translate into three major

transformations in service delivery: (1) from illness-centered to person-centered: *Recovery is something a person does, not what an illness does.* (2) From professionally driven to client-driven: *Recovery isn't something a professional can do to a person. It is a process they must engage in.* And (3) from deficit-based to strength-based: *Recovery doesn't happen when the person never has any more symptoms, even serious ones, but when they have the strength and resilience to handle whatever symptoms occur.*

RECOVERY OUTCOMES

"Schizophrenia is a chronic disabling illness affecting approximately 1 percent of the US population" is a familiar introduction to schizophrenia in the medical literature. The pessimistic tone of this introduction does not take into account the heterogeneity of outcomes associated with Recovery in this population. Existing challenges to objectively measuring the longitudinal progress of people with schizophrenia include: (1) evolving, nonuniform definitions of schizophrenia across the world; (2) lacking universally accepted criteria for measuring Recovery and rather measuring the nonexistence of symptoms associated with schizophrenia; and (3) having self-reinforcing conclusion (i.e., bias of researcher/clinician that worst outcomes are typical based on the historical, mainstream medical society's perception).

Although Emil Kraepelin's intent was to advance the understanding of psychiatric conditions through the categorization of symptoms into manic depression and dementia praecox, a result was dementia praecox (i.e., schizophrenia) became associated with poor outcome while manic depression with good outcome. However, 17 percent and 8 to 13 percent of dementia praecox cases demonstrated clinical long-term improvement and long-term recovery, respectively. Eugen Bleuler, who coined the word "schizophrenia," also propagated this sentiment when he revised his initial more optimistic assessment to one never associated with a "full restitutio ad integrum."

In the 1970s, published studies contested the dogma that schizophrenia equals poor outcomes. The 1987 findings of the Vermont Longitudinal Research Project, which had a more rigorous study design as compared to some of the 1970s studies, supported the findings of these earlier studies. One-half to two-thirds of the Vermont study participants attained considerable improvement or recovery 20 to 30 years post completion of a comprehensive rehabilitation program and planned deinstitutionalization (Table 12.15–1). Vermont Longitudinal Research Project's cohort consisted of clients residing in Vermont State Hospital (n = 269) between January 1955 to December 1960 who were considered the most chronically mentally ill among the Vermont State Hospital's client population. They had not responded to chlorpromazine (Thorazine) treatment and had been referred to a comprehensive rehabilitation and community placement program

sponsored by the Vermont State Hospital and Vocational Rehabilitation Division of the Vermont State Department of Education.

Table 12.15–1.
Results from the Strauss–Carpenter Levels of Function Scale for the 168 Subjects of the Vermont Study Who Were Alive and Interviewed

Area of Functioning	N	%
Not in hospital in past year	140	83
Met with friends every week or two	111	66
Had one or more moderately to very close friends	128	76
Employed in past year[a]	79	47
Displayed slight or no symptoms	121	72
Able to meet basic needs	133	79
Led moderate to very full life	128	76
Slight or no impairment in overall function	92	55

[a]Quality of work could not be rated; issues of confidentiality prevented visits to subjects' work sites.

Data from Harding CM, Brooks GW, Ashikage T, Strauss JS, Breier A. The Vermont Longitudinal Study of Persons with Severe Mental Illness, I: Methodology, Study Sample, and Overall Status 32 Years Later. *Am J Psychiatry*. 1987;144:718–726.

Despite the evidence that Recovery occurs in schizophrenia, the concepts that schizophrenia is a progressive disease and associated psychosis may be deleterious to the brain assisted in program development for minimizing the duration of untreated psychosis among young people experiencing a first episode of psychosis. PACE (Personal Assessment and Crisis Evaluation) Clinic was established in 1994 as the first clinical and research clinic in the world designed to prevent or delay the onset of psychosis among persons aged 14 to 25 years old with prodromal symptoms, brief self-resolving psychotic episodes, and significant decrease in functioning in the past year. Levels of remission (defined as mild positive symptoms with no greater than moderate negative symptoms) are high at 70 to 83 percent within first year of antipsychotic treatment. Among those who adhere to their prescribed antipsychotic treatment, their maintenance of remission rate is also favorable at around 80 percent; however, remission of symptoms does not necessarily equate to functional Recovery. "Poor outcome" rate was also observed to be stable at approximately 25 percent of the study population. Currently, a clear association between duration of untreated psychosis and magnitude of structural brain differences does not exist.

World Health Organization's International Study of Schizophrenia found the best predictor of symptom and disability scores at 15-year follow-up was the percentage of time spent psychotic during the first two follow-up years post first episode of psychosis. Existing evidence on maintenance

antipsychotic treatment (10 or more years) is mixed. Not all people with schizophrenia require continuous antipsychotic treatment for the remainder of their lives. Their characteristics include: better pre-morbid developmental condition, less vulnerability to anxiety and psychosis, and better neurocognitive skills. Another predictor of outcome is country of residence. In another World Health Organization study, participants living in East Asia experienced the highest clinical remission rate at 84.4 percent; however, participants living in North America experienced the highest functional remission rate at 35.0 percent. Environmental factors associated with Recovery include repeated opportunities for rehabilitation, positive expectations and relationships, family psychoeducation, and access to good-quality mental health care (i.e., therapeutic alliance with clinician, satisfaction with antipsychotics, general medical condition care, assistance in managing deficits in resources, and responsiveness to needs and services).

The principle that schizophrenia is not uniformly synonymous as a chronic disabling illness was not a prevalent view during the *Diagnostic and Statistical Manual of Mental Disorders* (DSM)-III era. DSM-III described schizophrenia as "a complete return to pre-morbid level of functioning in individuals is so rare as to cast doubt upon the accuracy of the diagnosis." However, DSM-5 reads: "course and outcome may not be reliably predicted. The course appears to be favorable in about 20% of those with schizophrenia, and a small number of individuals are reported to recover completely. However, most individuals with schizophrenia still require formal or informal daily living supports, and many remain chronically ill, with exacerbations and remissions of active symptoms, while others have a course of progressive deterioration." As appreciation of Recovery being possible grows, a question naturally arises regarding what are the meaningful outcomes that should be evaluated to establish a practice as being a *recovery evidence-based practice*, especially given the majority of existing peer-reviewed literature, analyzes symptom rather than functional, everyday life task improvement.

EVIDENCE-BASED PRACTICES

The presented practices are not an exhaustive list of Recovery evidence-based practices. For example, US Department of Health and Human Services' Substance Abuse and Mental Health Services Administration alone lists 400 interventions. Selected practices based on the recommendations of peer-reviewed review articles are presented in detail.

Table 12.15–2.
Significant Outcomes for Assertive Community Treatment in 25 Randomized Controlled Trials[a]

Parameter	Effectiveness of ACT Compared with Control Conditions (No. of Trials [%])		
	Better	No Different	Worse
Psychiatric hospital use	17 (74%)	6 (26%)	0
Housing stability	8 (67%)	3 (25%)	1 (8%)
Symptoms	7 (44%)	9 (56%)	0
Quality of life	7 (58%)	5 (42%)	0
Social adjustment	3 (23%)	10 (77%)	0
Jail/arrests	2 (20%)	7 (70%)	1 (10%)
Substance use	2 (33%)	4 (67%)	0
Medication compliance	2 (50%)	2 (50%)	0
Vocational functioning	3 (37%)	5 (63%)	0
Patient satisfaction with services	7 (88%)	1 (12%)	0
Family members satisfaction with services	2 (67%)	1 (33%)	0

[a]Patient numbers in these studies ranged from 28 to 873, with a median of 130 and mean of 171.2 patients. The follow-up period ranged from 3 to 36 months, with a median of 18 and a mean of 17.7 months.

Pharmacotherapy

Pharmacotherapy is the mainstay of evidence-based treatments for schizophrenia. First-generation antipsychotics are associated with a 25 to 70 percent reduction in annual relapse rates of positive symptoms in efficacy trials. Comparable but not better rates exist for risperidone, olanzapine, quetiapine, ziprasidone, and aripiprazole. Neither first- or second-generation antipsychotics effectively treat primary negative symptoms or cognitive impairments; second-generation antipsychotics are associated with minimal cognitive gains that are not associated with functional improvements. Antipsychotics improve functional impairment as a consequence of reducing positive symptoms; however, the main causes of functional impairment are negative symptoms and cognitive impairment. Clinical Antipsychotic Trials of Intervention Effectiveness found that before 18 months of antipsychotic treatment 74 percent of the study's population discontinued their antipsychotic due to such factors as inefficacy or intolerable side effects. This finding stresses the importance of the inclusion of non-pharmacological therapies in tandem with pharmacotherapy in the Recovery process for people with schizophrenia.

Assertive Community Treatment (ACT)

ACT is a comprehensive, individualized approach to delivering care for people whose functional impairments traditionally from chronic psychosis prevents them from effectively navigating their lives and health care systems in less intensive care models. ACT's fundamental principles include multidisciplinary staffing, integration of services, team approach

(i.e., shared caseloads and daily team meetings), low client–staff ratios (i.e., 10:1 ratio which varies depending on the severity of clients' functional impairment), point-of-contact care delivery in the community, medication management, focus on everyday concerns, rapid access, assertive outreach (i.e., being persistent in offering services to clients who are not engaging in their care), individualized services, and time-unlimited services. ACT is one of the most researched treatment models and is associated with numerous beneficial outcomes (Table 12.15–2). However, critics view ACT as not being a Recovery-orientated practice due to its "assertive outreach" model of repeatedly contacting clients who have repeatedly declined this service, its medication management component, some of its fidelity standards are clinician rather than client preference based, and utilization of behavioral contracts, outpatient commitments, and representative payeeships. ACT's "time-unlimited services" principle is also contrary to Recovery's hopeful theme of eventual functional improvement.

Supported Employment (SE)

SE's core principles include obtaining competitive employment rather than sheltered workshop or transitional work participation, being available to all clients living with psychosis who are interested in competitive employment, designing job searches based on client employment preferences, engaging in rapid job search, integrating mental health and employment services, and providing individualized job supports and benefits counseling. Study findings favor SE with a large effect size in job acquisition, total weeks worked, and job tenure as compared to non-SE vocational programs. This employment has been associated with reduction in outpatient psychiatric treatment and improvement in self-esteem. Employment is inconsistently associated with positive outcomes in symptom severity, psychiatric hospitalization, life satisfaction, and global wellbeing though employment is consistently unrelated to worsening outcomes.

Cognitive-Behavioral Therapy (CBT)

CBT, among the different psychotherapeutic approaches for schizophrenia, has the most evidence-based support for its therapeutic utility. It has significant effect size for decreasing positive symptoms. Evidence supports that these benefits occur among chronic, non-recent onset schizophrenia population and are sustained post-completion of treatment. Some effective CBT techniques to target psychotic symptoms include belief modification, focusing/reattribution, and normalizing psychotic experiences in context that client identifies specific symptoms that s/he would like to address.

Cognitive Remediation

Cognitive Remediation, also known as cognitive training and cognitive

rehabilitation, is "a behavioral training based intervention that aims to improve cognitive processes (attention, memory, executive function, social cognition, or metacognition) with the goal of durability and generalization," per Cognitive Remediation Experts Workshop (Florence, Italy, April 2010). A 2007 meta-analysis found that cognitive remediation's effect size is in the medium range for improving overall cognitive functioning and for improving six of seven cognitive domains across an array of program and client characteristics.

Family Psychosocial Intervention

Family Psychosocial Intervention to reduce high levels of criticism, hostility, or over involvement is promoted given people with schizophrenia from families with high expressive emotions have more frequent symptom relapses. Other potentially beneficial family interventions include illness education and associated coping skills training, crisis intervention, and emotional support for at least 6 to 9 months for clients who have ongoing contact with their families. Benefits also include reduction in hospitalizations or relapses over 2 year period of 25 to 50 percent.

Housing First (HF)

HF's foundation is that permanent housing is a basic human right to be provided in tandem with a psychiatric rehabilitation approach that values client choice. "Continuum of Care," the traditional model, requires clients to complete a progressive housing program model starting from outreach to transitional housing to end with permanent housing. Given its progressive nature, if clients are unable to meet the "housing readiness" expectations for each step, such as maintaining sobriety or adhering to prescribed medications, then clients do not advance to permanent housing and/or potentially return to homelessness. Studies have shown that clients of HF as compared to Continuum of Care are more likely to retain permanent housing (i.e., 79 percent compared to 27 percent at end of 6 months), obtain housing earlier, report higher perceived choice, and demonstrate no difference in substance use or psychiatric symptoms although utilization of substance abuse treatment is significantly higher among Continuum of Care clients.

Illness Self-Management Training

Illness Self-Management Training spans the spectrum from psychoeducation to skills training. Psychoeducation entails education about the nature of schizophrenia including its associated symptoms, the principles of the stress–vulnerability model, and treatment modalities along with their associated risks and benefits. Controlled studies have found that psychoeducation increases the participants' knowledge but is

not consistently associated with behavior change. Skills training activities are aimed at improving social interactions, independent living, and community reintegration. Although the content of these trainings differ, they share many key features, such as behaviorally based instruction, role modeling, rehearsal, corrective feedback, positive reinforcement, and practice in everyday, non-office based environment. While social skills training does not have much of an effect on relapse and hospitalization rates, it is associated with improvement in social functioning.

Wellness Recovery Action Planning (WRAP)

WRAP is a peer-led, self-management intervention program for chronic mental health conditions. Unlike traditional medical models, WRAP does not depend on clients obtaining formal mental health care. WRAP also fosters an environment which empowers clients to build a life worth living, rather than simply managing symptoms, through emphasizing holistic health, wellness, strengths, and social support. WRAP intervention has been associated with psychiatric symptom reduction and improvement in physical health, quality of life measures, and Recovery-orientated perspectives such as hopefulness.

Integrated Schizophrenia and Substance Use Disorder Treatment Models

Integrated Schizophrenia and Substance Use Disorder Treatment Models started mainly in the mid-1980s when it became evident that approximately 50 percent of people with schizophrenia had a substance use disorder, were having difficulty navigating the disconnected mental health and substance use disorder services, and were at times being denied substance use disorder care. The founding principle of integrated care is a multidisciplinary team that contains both mental health and substance use specialists who provide coordinated mental health and substance use care for all aspects of the client's life in a system which has the same administrative structure and funding source. A consistent finding is reduction in substance use disorder behaviors leading to stable remission; less consistent findings are associated with hospitalization rates, mental health symptoms, and quality of life measures.

Motivational Interviewing (MI)

MI's aim is to resolve ambivalence as the means to stimulate behavioral change. Its elements include rapport building, reflective listening, assessing client's motivation for change through use of open-ended questions, highlighting client's change-related statements and efforts, demonstrating gap between baseline and desired behaviors, soliciting permission prior to providing assessment or recommendations, utilizing resistance as prompt to alter approach, bolstering client's self-efficacy for change, and creating

an action plan that client verbalizes willingness to follow. Meta-analysis has found utilization of MI associated with statistically significant improvement in clinical outcomes across a range of behavioral related health conditions.

Electroconvulsive Therapy (ECT)

ECT is indicated for the treatment of schizophrenia when antipsychotic therapy has not been associated with a beneficial response.

Low-Frequency Repetitive Transcranial Magnetic Stimulation (rTMS)

rTMS is also recommended for acute treatment of auditory hallucinations when antipsychotic therapy has not been associated with a beneficial response.

Crisis Intervention Team Model and Mental Health Courts

Crisis Intervention Team Model and Mental Health Courts are considered "Best Practices" because they do not have the needed body of literature for determination of being Evidence-Based Practices. They are being presented given the number of people with serious mental illness in prisons and jails surpass the number in state psychiatric hospitals by tenfold. Crisis Intervention Team (CIT)'s goal is to divert people with serious mental illness from the criminal justice system to services, if appropriate, through a collaborative approach between law enforcement and mental health systems. The main components of CIT include (1) 40 hours of mental health related training for CIT officers; (2) training for call dispatchers to recognize calls potentially involving persons with serious mental illness and assigning them to CIT officers; (3) central designated psychiatric emergency drop-off site that accepts all people CIT officers bring for assessment. Initial findings suggest that CIT programs are associated with reduced arrests, increased diversion to mental health services, and increased safety. Regarding safety outcomes, compared to non-CIT officers, CIT officers were found to use less force as subject resistance increased. When CIT officers used force in the 15 percent of encounters rated as high violence risk, they generally relied on low-lethality methods. A recent example of the necessity of CIT occurred on October 18, 2016 when a New York City police sergeant killed Deborah Danner, a person with schizophrenia, in the context of "not following the training or protocols for dealing with those with mental illness," per New York City Major's assessment.

Mental health courts (MHC) were created to address the over-representation and disproportionate jail stays of inmates with mental health illness; US jail system incarcerates approximately two million adults with serious mental illness annually and retains them 2.5 to 8 times longer

than adults without mental health illness. MHC's goal is to divert people with mental illness from the criminal justice system to community treatment while maintaining public safety. Participation in MHCs is voluntary given this decision is usually accompanied by a mandatory guilty plea; however, some courts expunge guilty pleas upon successful completion. Generally, clients are required to sign a written contract pledging their commitment to MHC's mandates, such as clients taking their prescribed medications. Penalties for non-adherence to this contract include increased supervision and re-incarceration. Participation in MHCs has been associated with lower annualized re-arrest rate, fewer arrests, and fewer incarceration days compared to treatment as usual during an 18 month observation period following MHC treatment. Criticism of MHCs include (1) insufficient data regarding clients' appreciation of MHC's voluntary nature and consequences of participation; (2) its occasional utilization with clients who have impairments in legal competence; (3) twice as long processing time between arrest to outcome compared to conventional court system for same inmate population (70 vs. 37 days, respectively); (4)adding to the stigma associated with mental illness; (5) prioritizing access to limited community mental health services to those in MHCs.

Among expert clinicians, the trend is to define practices as being the "gold standards" if they are "evidenced-based." Some clients do not share this belief given this approach limits their ability to choose from treatment modalities and has the propensity to diminish the validity of their first-hand experience with treatment modalities. Actions that might bridge this gap include interventions that evaluate outcomes generated by clients, active recruitment of people with mental health conditions into mental health fields given this level of training is required to appreciate the nuances of evidence-based practice creation and evaluation, and self-disclosure by current mental health practitioners of their mental health conditions as a means to demonstrate that mental health policy and research decisions are not being created in isolation from clients.

JOURNEY OF RECOVERY

From an illness-centered point of view, the treatment of schizophrenia is based upon the course of the illness: It may begin with a prodromal phase before a "first break." There is often a period of untreated psychosis, before multiple attempts to acutely stabilize the illness undermined by relapses. Over time the illness may stabilize and go into a residual phase or even remission, though acute relapses can always recur.

From a person-centered point of view, the service needs depend less upon the diagnosis or stage of the illness and more upon how the illness is impacting the person and how they are relating to it—for example, out of control and at high risk, ambivalently engaging with helpers, taking self-

responsibility and building coping skills, incorporating the illness into a rebuilt self-identity, or adapting to reconnect to meaningful community roles.

There are perceptive autobiographies written by people who have worked hard to recover from schizophrenia that everyone can all learn from, ranging from Joanne Greenberg's fictionalized classic "I Never Promised You a Rose Garden" describing her relationship with her psychiatrist Frieda Fromm-Reichmann at Chestnut Lodge and her return to the community in the premedication, psychoanalytic era, to the more contemporary "The Center Cannot Hold" describing Elyn Saks' ongoing struggles with psychosis even as she has become a law school professor and associate dean at USC.

Patricia Deegan built her exemplary career promoting Recovery practice upon her own experiences and her taped interviews of people describing their own Recoveries. Based on these "lived experiences" of Recovery, she created her "conspiracy of hope" promoting self-help "Coping with Voices" techniques and her "Common Ground" program promoting "personal medicine" and shared decision making with psychiatrists. The approach of building strategies on personal stories and experiences has expanded to families in NAMI (National Alliance on Mental Illness) with their "In Our Own Voice" programs. Sharing personal experiences has proved a powerful tool for bridging differences and reducing antagonism to work effectively together.

Conceptually, it's important to begin with, "What is this person recovering from?" From a person-centered point of view, there are many other issues beyond illness symptoms to recover from including hopelessness, isolation, lost functions, roles, and income, self-esteem, damaged families, internal and external stigma, faith, etc.

Mark Ragins has described four common stages in a book *A Road to Recovery:* Hope, Empowerment, Self-Responsibility, and Having Meaningful Roles in Life. Specific tools have been developed to help people along the way.

Stage 1: Hope

To be motivating, hope has to be more than just an ideal. It has to take form as an actual image of how things could be if they were to improve. Tools to promote hope include:

- ► Positive goal visualizing and goal setting incorporating what people are most passionate about and likely to persist with
- ► Exposure to new possibilities
- ► Working with peer staff who have achieved considerable Recovery as role models
- ► Sharing pain and suffering, willing to start by "meeting people where

they are" avoiding empty reassurance and encouragement
- Hopeful staff.

Stage 2: Empowerment

To move forward, people need to have a sense of their own capability, their own power. Their hope needs to be focused on things they can do rather than new cures or fixes someone else will discover or give to them. It is often needed for someone else to believe in them before they're strong enough to believe in themselves and to start focusing on their strengths instead of their losses. It also often takes some actual experience of success to really believe one can be successful. Waiting until someone is ready to move on can often be stagnating and disempowering because "readiness" often occurs only in retrospect after something has been done successfully. Tools to promote empowerment include:

- Client-driven, self-directed treatment plans
- Offering meaningful education and choices in a process of shared decision making instead of just informed consent
- Heavy consumer inclusion—"nothing about us without us"
- Skill building services
- Internalizing successes as self-pride and self-confidence.

Stage 3: Self-Responsibility

At some point, most people who recover realize that no one else can do it for them, that they have to take charge of their Recoveries. People can be supported, but they can't be caretaken or protected into Recovery. Taking one's own risks, setting one's own goals and path, and learning one's own lessons are essential parts of Recovery. Tools to promote self-responsibility include:

- Motivational interviewing and enhancement
- Avoiding overprotection from "natural consequences" and learning from failures
- Creating opportunities for self-responsibility and learning from successes
- Self-help tools like WRAP
- Growth oriented risk taking.

Stage 4: Having Meaningful Roles in Life

Ultimately to recover one must achieve some meaningful role apart from the schizophrenia. Connectedness to other people, belonging, and feeling accepted, which may have for a while only been possible with others who have mental illnesses, within families, or with compassionate helpers

becomes possible in a variety of contexts. The isolation is increasingly broken and life re-entered. Tools to promote meaningful roles in life include:

- "Supported services," e.g., education, housing, employment
- Family connection and reunification services
- Peer support and "giving back"
- Community development, e.g., finding and building community welcoming, inclusion, and meaningful niches.

It is also important to have a conception of where the Recovery journey is going. People will naturally be more likely to be motivated to pursue their own goals rather than staff or community goals. Commonly, Recovery goals will begin practically with functions, external things, and roles. As people progress in their Recovery, their goals progress as well. Their view of what Recovery means to them and where this journey is taking them develops and deepens along the way.

The journey of Recovery should take place primarily in the community, rather than in isolated islands of tolerance and asylum. The overarching social culture, as well as many local community cultures, often makes Recovery difficult. Illness-centered, psychoeducational, anti-stigma efforts are likely less effective than relationship based, storytelling efforts. The overall goals are to increase community acceptance, inclusion, compassion, relationships, and belonging for people with schizophrenia. Since the middle ages when a spiritual healing center for people with mental illnesses was established there, the small town of Gheel, Belgium has maintained a pervasive counter-culture integrating people with severe mental illnesses within their families and community, proving it can be done.

ROLE OF PSYCHIATRIST

Psychiatrists are likely to be the most important staff in a person's Recovery (all too often they are the only staff) despite the profound shortages of psychiatrists and limited time available. This situation is due to the crucial importance of medications and hospitalizations within the current system, both usually under direct psychiatrist control. There are important tools psychiatrists can incorporate into their practice.

Person-Centered Formulations instead of Illness-Centered Diagnoses

DSM diagnoses are very broad symptom clusters without connection to childhood, stressors, strengths, or causes. A person-centered formulation goes beyond the "what's wrong" of diagnoses, to include a best guess as to "why" their life unfolded as it did. It includes thinking about deficits and strengths, genetics and environment, biology and trauma, family and culture, and risk and protective factors interwoven into a narrative that

makes sense to them and they can recognize as their story. This is the shared story that forms the basis for treatment collaboration.

Relationship-Based Services

Most people with schizophrenia are likely to have considerable barriers to forming trusting relationships. It takes skill to assess these barriers (e.g., childhood trauma, paranoia, warnings from voices, fear of being locked up, poor prior experiences with psychiatrists, substance abuse, secondary gain, internal stigma, and cultural barriers) and to overcome them to create and sustain a trusting treatment relationship.

Trauma Informed Care

Many people with schizophrenia have been severely traumatized, sometimes as an important trigger for their psychosis. Trauma perspectives redefine "what's wrong with you?" to "what happened to you" and "symptoms" to "coping techniques you used to survive that make problems now." Recovery is a normal expectation overcoming trauma. Many times there's nothing "post" about their trauma—they're unsafe and being actively traumatized right now.

Goal-Driven Medications and Treatment

Medications are "used" to make it easier to achieve personal goals, rather than "taken" because they're sick and need them and "the doctor told me to." Initially, medications can be for "short term" effects until a "customer relationship" is built. Taking medication to improve symptoms doesn't have to precede rebuilding lives. Medications enable self-help coping techniques, rather than competing with them. Getting off medications happens not when symptoms are relieved, but when they're no longer needed to attain and maintain goals.

Shared Decision Making

Most people (including doctors and patients) are more motivated to do what they want to do than what someone else wants them to do. Informed consent is almost always educating people about alternatives that are staff choices, while shared decision making combines the person's expertise about themselves and their own motivations with staff's expertise about a variety of things (based on "evidence," professional knowledge, or experience) to come to a decision that hopefully is better than either would have arrived at alone and that they're both motivated to work on. Shared decision making is the foundation for collaboration as opposed to compliance or adherence.

Activated Patients Promoting Their Own Recoveries

Being "compliant" and leaving it to the doctor to prescribe medications to cure schizophrenia, or even to put it into remission, hardly ever works. Patients must take a key role in promoting an array of activities to promote their own Recovery and in facilitating access to a broad array of support for those activities. A "prescription pad" of "What else are you going to do? And how are you going to get help to do it?" can include:

- ► Talk to other people instead of isolating.
- ► Actively build security in your life—money to survive on, safe housing, and connections with family, friends, and spirituality.
- ► Actually feel feelings and emotions instead of deadening them, medicating them, avoiding them, or getting high.
- ► Learn some emotional coping skills.
- ► Learn to "use" medications instead of just "taking" medications.
- ► Engage (or re-engage) in activities that make you more fun and interesting.
- ► Take responsibility for your own life and make some changes in yourself.
- ► Go to work even when you're not feeling well.
- ► Do things outside of being a "mental patient" and outside the mental health system.
- ► Improve physical health and wellness.
- ► Love other people—family, partners, kids, and friends.
- ► Work on acceptance and forgiveness instead of blaming and vengeance.
- ► Give back by helping others.
- ► Find meaning and blessings in suffering and reconnect with God and spirituality.

Taking Strengths Seriously and Building Resilience

- ► Building Social Determinants of Health: Often the outcome of treatment depends more upon the social characteristics of the person than upon their illness itself. Integrating poverty services with mental health services without saying "that's not my job" will likely improve outcomes. Political and community activism supporting social policies that promote improved social determinants of health will also likely improve outcomes.
- ► Building protective factors: Protective factors have been shown to prevent risk factors from becoming predictive. For people experiencing repeated mental health crises, it may be more effective to build protective factors to prevent further mental health crises than to treat each crisis after it occurs.
- ► Building self-efficacy: The more someone believes and has experienced that what they do makes a difference in their life, the more likely they are to take self-responsible action to recover. Examples of self-efficacy

include: Increasing the ability to affect inner states directly, increasing one's role in their own life instead of relying on others, developing specific goal driven skills to achieve personal growth, making developmental progress through life's stages, building interpersonal and community based efficacy, and increasing spirituality.

- Building resilience by finding strengths in struggles: Sometimes the process of struggling helps people discover and develop strengths inside of themselves they didn't know they had, receive gifts from their wounds, and figure out what's really important in life. Examples of these strengths include: internal qualities, external resources, meaning making, and enhanced connectedness and belonging.

This example illustrates how the psychiatrist-patient relationship keeps evolving to meet the needs of each stage of the Recovery journey:

Martha had suffered a great deal by the time I met her. She grew up in a small town in Missouri, the fourth of eight children. She had only a ninth grade education and could barely read and write. When she was 15 years old, Martha's father left for St Louis to find work and her mother struggled to take care of the family even with his financial help. Martha started spending more time reading her Bible and began believing that God was telling her what to do. Unfortunately, the instructions included spending her day trying to hail down cars on the highway to teach people about the Bible and taking off all of her clothes in church to prove she was pure. After trying to deal with her for several years, Martha's family took her to a psychiatric hospital when she was standing over her mother with a knife one night. She was repeatedly put into seclusion and sedated for disrobing in the hospital. One day, she reported being raped by another patient, but the staff didn't believe Martha. After two weeks, she was discharged back home with a prescription she threw away because it drugged her. After five more years of psychosis, fighting with her family, never going to the clinic in the next town, police being called repeatedly on her, and several hospitalizations and jailings, she ran away to Los Angeles where she arrived without any funds or friends. She was huddled in the corner mumbling to herself, disheveled, holding tight to her Bible.

Help Them Trust Us

In a Recovery-based treatment, clients do not trust their psychiatrist just because of their ability to treat them. They also trust their psychiatrist to personally care enough to create a relationship where they can find strengths in themselves to use their treatments and rebuild their lives.

Martha was wary of psychiatrists because she was afraid of being locked up so staff arranged for me to meet her a few times informally eating lunch at our client run Deli. We talked about my memories of St Louis from medical school and about how she had helped raise her younger siblings. We shared family pictures from our wallets. I told Martha about a couple other young women we'd met who came to Los Angeles all alone and how, even though it is hard, I thought she had the strengths to make it. The third time, I helped Martha with forms to get welfare payments and a discount bus pass and offered to help her get social security if she came in for a full assessment. We let Martha use our showers and got her some clean clothes from our donations.

Help Them Regain Control of their Lives

In a recovery-based treatment, the focus isn't on the psychiatrist taking care of them and helping them feel better. It is on them taking control of their own life so they can take care of themselves and function even if they're not always feeling well.

Martha enrolled in our program. We asked her what her goals were, and she said she wanted a place to stay of her own, someone to talk to and feel normal with, and to go to church. We introduced Martha to another member who was a little older and further on in her Recovery, and they became friends even though Martha is hard to understand sometimes and still talked to voices. We loaned Martha money against her pending back check form SSI to get a cheap weekly hotel room. In return, she had to come to the Village and do some light office work for us one hour-a-day. Martha's favorite job was helping to decorate the office for Christmas, but it made her miss her family. We helped Martha call them and tell them she was safe, and she sent off a box of presents for them. I showed her the award I'd gotten from the local Baptist church for community service and told her that although it isn't my job to help find business for churches, I could give her a list of local churches to try. She got dressed up to try the church her friend went to that caters to homeless people. As she was feeling stronger, she asked to meet our teacher to learn to read better. She became attached to him as a surrogate father figure, but mostly she was shocked and pleased with herself that she could learn to read when no one was yelling at her and hitting her when she made a mistake.

Help Them Rebuild their Life

In a recovery-based treatment, the goal is not just to treat a mental illness. The goal is to help someone with a mental illness have a better life.

My report helped Martha get SSI, and we gave her the money for a deposit for her own apartment. Several staff and clients went to Martha's housewarming party. I brought diet soda and chocolate. She said she was having trouble sleeping and wondered if I had any medications that might help. She noticed that other clients weren't being drugged up here at the Village and had grown to trust me. I asked Martha about what was keeping her awake, and we ended up talking about how she was tormented by demons which she hated as well as hearing God which was a blessing. We talked about how the demons took advantage of her when she felt weak and vulnerable and maybe medications could strengthen her to keep them away. We decided together that she would try taking low dosage antipsychotic medications at bedtime, and I encouraged more nightly prayers to prepare to sleep safely. She started coming to regular appointments with me in my office. As she felt stronger, she wanted to earn some money to donate to the church, and we helped Martha get a part-time training job at our Deli. She enjoyed preparing my daily salad the way I like it, without any carrots.

Help Them Heal

In a recovery-based treatment, healing isn't just from the illness. Healing is also from the destruction and the rejection.

She was proud to share her success with her mother and younger brother, but she also began to realize how much she'd hurt them over the last few years. We began talking about her regrets more in our sessions than her medications, and she began to accept what she'd done enough to write a letter asking her mother for forgiveness. She never sent it, but she felt a weight lifted nonetheless. When our teacher left to take another job, she was devastated. She started isolating and stopped taking her medications, and the demons returned with a vengeance.

Take the Long View

In a recovery-based treatment, psychiatrists don't just assess and treat illnesses as a detached professional. Psychiatrist and patient walk together, sharing the joys and sufferings of Recovery.

> We resisted the impulse to hospitalize Martha and get her back on her medications even though it was heartbreaking to see Martha deteriorate again because we didn't want to push her away. She lost her job at the Deli. She wouldn't respond to repeated phone calls so I went with Martha's case worker to Martha's apartment to reach out to her. She opened the door and broke into tears. As we talked, we realized together that the loss of our school teacher had reactivated the feelings of abandonment and betrayal she'd felt when Martha's father had left them. We talked about how she was older and stronger now and could handle this departure. She agreed to take the bottle of pills I'd brought in my pocket.

Help Them Move On

Even though many serious mental disorders persist indefinitely, a Recovery-based treatment can prepare them to deal with whatever happens and keep going with their lives without needing to be in treatment forever.

> After Martha rebuilt her life again, she enrolled in a WRAP group to develop better tools besides giving up and isolation to deal with life's inevitable losses in the future. And she'd have to do it without me coming to her home to help rescue her. I started spending less and less time helping her, and she spent more time mentoring new clients. The next Christmas she saved money to take a bus back to Missouri, and this time did ask her mother for forgiveness which her mother gave her. After she returned to Los Angeles, she realized that Missouri really was her home, and we made arrangements for her to return home including a phone call and letter from me so she could get medication from the clinic there. At our farewell party, she gave me a poster she'd made with her friend, as she was leaving, for my office wall that said, "The Village opens a door for us to have a better life. It's up to us to walk through it."

PEER SUPPORT SPECIALIST

Peer support is emotional support offered by a person with mental health condition(s) to another who has similar mental health condition(s) so as to effect beneficial change. Peer Support Specialists (PSS) are people who are active in their Mental Health Recovery and potentially have obtained a PSS certificate. In 2005, over ten thousand PSS were working in the United States alone. Although fluctuating in its prominence in the history of psychiatry, the role of PSS can trace its roots to at least the "moral treatment era," 18th century concept whose aim was to treat people with mental health disorders humanely. In 1935, Alcoholics Anonymous formed and utilized the peer support concept of a "sponsor." During the 1970s, Mental Health Recovery Movement advocated for more inclusion of people with mental health disorders in their mental health care, and as such the concept of incorporating people with mental health disorders into mental health treatment teams further emerged. In the early 1990s, the current version of PSS developed; however, their roles and responsibilities were

limited as compared to today's PSS.

Early roles for PSS generally included less clinical activities or activities that non-PSS clinicians were too busy to perform, such as driving clients to their appointments, filing, and answering phone calls. PSS initially did not participate in treatment planning given their marginalized, unclear clinical role. With time, PSS' roles have become more clinical in nature to include assisting in identifying goals, teaching new skills, advocating, job coaching, providing counseling, and leading groups along with documenting these clinical interactions in the clients' medical records. One of the main roles of PSS is to disclose their mental health conditions to clients as a means to serve as a role model. This disclosure also potentially makes them more approachable to clients due to this acknowledged shared similar mental health related experiences. These shared experiences can lay the foundation for engagement with clients whom non-PSS clinicians have been unable to reach. Another important role of PSS is to educate non-PSS clinicians about what methods might assist clients based on the PSS' past experience with similar experiences.

SAMHSA considers the role of PSS as evidence-based practice in the context that "there is a need to identify specific attitudes, knowledge, and skills that are essential for peer support and effective operation of these services and programs." The bulk of existing research includes ability to train and hire PSS and comparison between PSS and non-PSS clinicians' outcomes. These studies have shown that PSS are effective in engaging people in care, reducing emergency visits and hospital admissions, and reducing substance use. If PSS provide positive self-disclosure and role modeling to clients, these actions increase the clients' self-care ability, decrease clients' depression and psychosis, and increase clients' sense of hope, control, ability to effect changes, community belonging, and life satisfaction.

Despite these benefits associated with PSS, hiring officials at times are hesitant about hiring PSS for multiple reasons. One reason is concern that job related stressors will disproportionately negatively impact PSS compared to non-PSS clinicians. As a result, "stability" eligibility criteria are sometimes utilized in hiring practices, such as a time requirement since last mental health hospitalization or length of sobriety. These eligibility restrictions are not based on any evidence-based predictive findings and potentially constitute discrimination based on disability rights legislation. Another concern is PSS might break clients' confidentiality; however, like other clinicians, PSS have national practice guidelines that emphasize confidentiality and privacy and should be provided with on the job training and supervision which also emphasize these topics. Yet another concern is that PSS will make the supervisor's work harder. Initially, PSS might benefit from more supervision time to acclimate to working in their new environment than non-PSS clinicians; however, long-term benefits of

improved client engagement and associated outcomes generally outweigh this time investment.

PSS do encounter unique on-the-job challenges that non-PSS clinicians do not encounter. The identification of therapeutic boundaries might be more challenging, given their past experiences, and their role of being the more relatable people in the treatment team. For example, in the past, PSS could have been hospitalized with a client and as such has a pre-existing relationship with this client as a client. Another unique challenge is clients often view PSS as reaching the pinnacle of success and as such are viewed as role models who they overtly idolize. Simultaneously, some clients may be concerned about a possible relapse for the PSS, impugning the reputation of recovering persons as being able to perform in a competent manner. These challenges can usually be effectively dealt with in supervision with PSS' clinical supervisor either in individual or group meetings with other PSS and through structured training. In addition to challenges, the role of PSS often brings personal growth to both PSS and the other members of the treatment team. The PSS benefit from the experience of having fulfilling employment. The non-PSS clinicians are able to interact with people with openly disclosed mental health conditions who are contributing to their team's mission, challenging misconceptions about the abilities of people living with mental health conditions.

INVOLUNTARY TREATMENT

From 1880 through 1950, the population of the nation's public psychiatric hospitals increased by more than fivefold on a per capita basis. Unfortunately, during this time of increased hospital population, resources for the hospitals were greatly restricted, particularly during the time of the World Wars and the Great Depression. Hospitals became outrageously overcrowded and conditions for the patients became horrendous.

However, during World War II, many of those who became classified as conscious objectors (COs) were given jobs replacing orderlies and other staff in the public psychiatric hospitals. The COs, particularly the Mennonites assigned to work in the Cleveland State Hospital in Ohio, became outraged by the conditions they found in that institution. They observed human beings being treated worse than animals and managed to bring what they found to the attention of reporters from the local *Cleveland Press*. The publication of stories about the terrible conditions in the hospital became the crucial spark that initiated a virtual whirlwind of national media attention on the plight of hospitalized psychiatric patients in the next few years. *Readers Digest, Life,* and numerous other popular magazines published major expose's. Books such as Morton Deutch's *The Shame of the States* and Mary Jane Ward's *The Snake Pit* were widely read, and the public began to demand changes.

Until this time, psychiatrists had enjoyed great deference from society

and had almost undisputed authority concerning who could be committed to these facilities and to the type of treatment, or lack thereof, they received once they became inpatients. However, with the adverse media campaign focusing on the callousness and abuse existing in the psychiatric hospitals, the absolute authority enjoyed by hospital psychiatrists began to be eroded. Until the 1950s, having a person declared insane and committing them psychiatrically was usually dependent on the clinical judgment of psychiatrists. Their decisions concerning "treatment" delivered to hospitalized patients, including those concerning the use of restraints, seclusion, medications, various forms of shock therapies, etc., were rarely challenged. The Community Mental Health Centers Act of 1963, which initiated the deinstitutionalization movement, had as a consequence a diminution in the authority of psychiatrists. The Lanterman-Petris-Short Act (1969) in California and Lessard decision in Wisconsin (1972) were particularly important in shifting the primary responsibility for involuntary treatment from the medical professionals to the courts. Increasingly, before a person could be involuntarily treated, it became necessary to provide proof before a judge that a mentally ill person was "immediately dangerous" to themselves or others. By 1980, every state had changed or reinterpreted its laws to make it more difficult to provide treatment to mentally ill persons who could not be shown to be imminently dangerous.

Importantly during this time, many who had been released from psychiatric hospitals began to organize and to make demands for changes. Early in these organizing activities, recovering persons found valuable allies in the legal profession, who were interested in changing laws so that psychiatric patients could be given maximum freedom, in spite of the opinions of mental health professionals about their need for treatment.

The call for "no forced treatment" by consumer spokespersons became the rallying cry for the rapidly emerging consumer advocacy groups and their allies in the civil rights movement, and increasingly in the mental health bureaucracies, particularly during the last 40 years.

However, many mental health advocates and activists increasingly found that "no forced treatment" by the psychiatric professionals meant that many of the recently freed mentally ill would engage in behaviors that caused them to be placed in the nation's jails and prisons. Indeed as of 2014, it was well documented that there were over ten times as many seriously mentally ill persons criminally incarcerated in jails and prisons than there were in public psychiatric hospitals and that a large majority of those left in the public psychiatric hospitals had criminal charges filed against them.

Steps to turn around the adverse consequences of criminalizing the mentally ill began to be taken. Working closely in conjunction with NAMI, police officers began to receive more formalized training in their handling calls involving mentally ill persons, which was about 10% of the total police

calls. CIT training programs began in the late 1980s in Memphis and thousands of police departments have now received CIT training, in almost every state, and several foreign countries.

Additionally specialized mental health courts began to be established where mentally ill persons accused of crimes could avoid incarceration and criminal histories by enrolling as participants in these courts. Similarly themed specialty dockets for veterans, persons with addictions, etc., have also become widespread.

One of the most controversial of the initiatives to decriminalize the mentally ill involved attempts to reverse the plethora of new laws that were driving this criminalization process. To the effect, NAMI established an advocacy group, which has subsequently become independent from the parent organization. The purpose of this group, named the Treatment Advocacy Center (TAC), has been to roll back laws throughout the nation so that mentally ill persons in crisis, particularly those who have demonstrated tendencies to become dangerous, can be given treatment, on an outpatient basis, even without their consent. TAC uses the term, assisted outpatient treatment (AOT), to refer to this process of coerced treatment. Since TAC's founding in 1998, 22 states have made significant changes to their civil commitment standards and treatment laws, in large part due to TAC's advocacy activities.

But these efforts, focused on allowing the delivery of mandatory treatment, are hugely controversial. Advocacy groups, such as Mental Health America, the National Disability Network, the National Coalition for Mental Health Recovery, and the esteemed Bazelon Center for Mental Health Law, all strongly oppose any further infringement of the rights of any person with a mental illness. Although under the various AOT laws patients cannot be forced to take medication, these advocacy groups see such initiatives as a "slippery slope" leading to a return to the horrors of the old psychiatric institutions.

There are former patients, including the current co-author (FJF), who are deeply saddened by the criminalization of persons with mental illness, but are also wary of returning to the world extreme psychiatric authoritarianism. It is frequently asserted that absolutely no person, who does not need it, should ever be given forced treatment. But unfortunately the nature of serious mental illness is that from time to time, some of people should be given treatment over their psychosis-driven objections.

However, the question remains, who should determine whether forced treatment should be delivered. Until about 50 years ago, psychiatrists virtually alone had this authority. More recently, this authority has increasingly been transferred to judges.

Consistent with the mantra, "Nothing about us without us," many people see initiatives, such as advanced directives and "shared decision making," as positive steps in that these approaches provide increasing

input concerning what treatments are received. As the efforts to decriminalize those with serious mental illness proceed, many are hopeful that, at some level, persons in Recovery or their family members, can have some input as to how decisions are made concerning being forced to receive unwanted treatments.

STIGMA/ADVOCACY

Most people with schizophrenia who have been released into the community in the wake of deinstitutionalization have encountered various forms of barriers to their being reintegrated into society. One class of barrier is that which is imbedded in linguistic and cultural heritage. The other, more easily changed, class of barriers consists of laws, rules, and regulations that obstruct people with schizophrenia from being afforded opportunities to integrate.

Linguistically and culturally, due at least in part to society's history of viewing people with schizophrenia as being possessed by demons and later their systematic isolation from others with the use of large institutions, society still tends to employ demeaning, highly pejorative terms, and other characterizations, when referring to people with schizophrenia:

"Get the dangerous 'crazies' off the streets."

"Only a 'nut job' would commit such an act."

"Schizophrenia is a shameful disease."

The above comments are just a sample of similar statements the author (FJF) hears frequently in the mass media and even during public lectures by mental health professionals.

Words like "crazy," "nuts," and even "schizophrenia" when used in the manner above, are fairly obvious examples of pejorative terms, similar to those used in the past to perpetuate exclusion of ethnic and religious minorities, gays and lesbians, and other traditionally marginalized groups.

People with mental illness generally are very aware that such pejorative terms are frequently used to refer to people with schizophrenia. However in the general parlance, words such as these are also used very frequently when there is no explicit intention by the speaker to refer to mental illness.

When one says something like: "That's a crazy idea!" or "You must be nuts to think such a thing!" the speaker usually means that what is being talked about is very unusual or that one is thinking "outside the box." Used in this manner, the purpose of speaker is usually not to demean the mentally ill.

Nevertheless, the words, "crazy," "nuts," "berserk," etc. are the same as the terms that are used in highly demeaning ways toward people with schizophrenia. Normal persons often think that it's okay to use these terms because they are not thinking of mentally illness. Indeed, in common parlance, if one wishes to dismiss an idea or proposition as unworthy of consideration, very often the idea in question is labeled as "crazy" or

"insane." In most people's minds, these terms immediately connote something that is beyond the pale for consideration by reasonable people. That which is crazy is that which can immediately and safely be dismissed without any further consideration. However, many persons in Recovery from serious mental illnesses tend to be very sensitive when hearing such words used in a pejorative, exclusionary manner.

In other words, even if those with mental illness are able to totally recover, they face a world of traditional attitudes and customs in the use of language that form a major linguistic/social barrier to being accepted into the broader society.

But the question remains, what is the politically correct term to use when referring to persons who experience schizophrenia and other serious mental illnesses. From the advocates' perspective an appropriate answer to this question becomes problematic.

In advocacy circles, those in Recovery split over this issue in 1985 at a gathering in Pottstown, PA. One group wanted to be called "consumers"; the other group adopted the term, "survivor." In the consumer literature, the term "C/S/X" (consumer/survivor/ex-patient) is often the preferred appellation.

Around this same time, NAMI was suggesting that when people with schizophrenia were in the hospital they should be called "patients," and after their release, they became "consumers." Accordingly, the NAMI "Client Council" changed its name to the "Consumer Council" around that time. Also about then the mental health bureaucracies, both state and federal, began to strongly encourage the use of "person first" language.

This taxonomy gets very complex. Europeans use "service user" or more frequently just "user." And, at some point, both New York and Michigan officially adopted the term, "recipient." Generally, the words, client, consumer, and patient or ex-patient, tend to be more generally accepted in advocacy circles in this country than some of the more loaded words.

The authors of this chapter are not going to suggest that they have a totally adequate solution to this labeling problem but might suggest that as a general rule the use of "person first" language is usually appropriate, provided one is not interrupting the flow of the language.

Cultural barriers take many forms in addition to the use of these pejorative words. Continued portrayal of the mentally ill as dangerous, weird, incapable, etc., in the media and elsewhere perpetuate stereotypes that those in Recovery from mental illness see as impeding their acceptance by society.

In this vein, at the end of the twentieth century, The American Film Institute selected the 50 most villainous cinema characters for the century. The sometimes psychotic killers, Hannibal Lecter (*The Silence of the Lambs*) and Norman Bates (*Psycho*) held the top two spots. Other persons with mental illness, their family members, and mental health professionals

(stigma by association) held a vastly disproportionate number of places on the list.

The second category of barriers, the more concrete rules and regulations, include social security and other disability programs, which act to provide subsistence-level resources. Unfortunately, these also often have the effect of disincentivizing people with schizophrenia from seeking employment.

In the United States, the Americans with Disabilities Act (ADA), which became law in 1990, prohibited discrimination against persons with psychiatric disabilities and ensured that reasonable accommodation would be given to disabled employees provided this would not be an "undue hardship" on the employer. While this law prohibits the discriminatory activities of employers, many fields have continued to effectively screen out disabled persons, particularly those with mental illnesses, by limiting their access to required education programs and restricting opportunities to become properly licensed.

Given these barriers to inclusion, how can an individual consumer work to diminish stigma? The author suggests the following.

(1) According to the psychiatrist, Howard Goldman, there is research revealing that the most powerful way to reduce discrimination and prejudice is exposure to persons who are open about being in Recovery from serious mental illness. Therefore, obviously one step that can be taken by a Recovering person is to become open about having such a condition. A major reason for not being open is shame, but recovering persons should be encouraged to not to be ashamed. Persons in Recovery from schizophrenia have met and overcome what is often described as the most devastating of all mental illnesses. Persons meeting this challenge should stand tall and be proud of what they are accomplishing. Confidentiality is the keeping of secrets. It allows persons in Recovery to keep others from finding out about their "shameful" condition. Remember that the right to confidentiality is the right of the patient, not the professional. So the patient has the right to be open and honest about their condition if they so desire. Many who have self-disclosed feel good about not having to hide anymore. In this regard, many of those in the advocacy movement are mental health professionals who themselves have recovered from schizophrenia or other serious mental illnesses. However, there are many other professionals who have psychiatric histories who are not yet ready to become open about these conditions. Because there is so much stigma and discrimination, being open about one's psychiatric history can be problematic as well as possibly threatening to one's career and social standing. Because of this negative climate, working professionals need to be wary about being too transparent about their mental health history. However, those who are more secure in their careers, as in

being close to or past retirement age, or academics with tenure, would perform a significant service by becoming open about their experiences with these disorders.

(2) Whether or not you are ready to become open and honest about your psychiatric history, you can still join with others who have similar conditions. There are numerous advocacy groups that meet frequently. NAMI has regular meetings of consumers, family members, and other advocates that meet all over the country. You do not have to necessarily identify yourself as a person with a psychiatric history to participate. Among many such groups, one often hears the mantra, "Advocacy is the best therapy."

(3) Regarding fighting the use of the demeaning words and phrases so often heard with reference to the mentally ill, significantly changing this situation in the general public will probably take at a least generation or two. But where can one start? The author recommends that one starts with him/herself. You probably use terms like "crazy," "nuts," and "loony," just like the rest of the population does. First, discipline yourself to stop speaking in this manner. Once you have accomplished this self-control, then confront others in the mental health community who use these terms, in a polite manner of course. The author recommends that you start these confrontations first with those in the mental health community because mental health advocates, professionals, family members and other consumers have the most vested interest in fighting discrimination and making things change.

REFERENCES

Anthony WA. Recovery from mental illness: the guiding vision of the mental health service system in the 1990s. *Psychosoc Rehab J*. 1993;16:11.

Australian Health Ministers' Advisory Council "A National Framework for Recovery-Oriented Mental Health Services: Guide for Practitioners and Providers" 2013. Available at https://www.health.gov.au/internet/main/publishing.nsf/Content/67D17065514CF8E8CA257

Baldwin ML. *Beyond Schizophrenia: Living and Working with a Mental Illness*. New York: Rowman & Littlefield; 2016.

Bond GR, Drake RE, Mueser KT, Latimer E. Assertive community treatment for people with severe mental illness critical ingredients and impact on patients. *Dis Manag Health Out*. 2001;9(3):141.

Ciompi L, Harding CM, Lehtinen K. Deep concern. *Schizophrenia Bull*. 2010;36:437.

Cook JA, Copeland ME, Jonikas JA, et al. Results of a randomized controlled trial of mental illness self-management using wellness recovery action planning. *Schizophrenia Bull*. 2012;38(4):881.

Copeland ME. *Wellness Recovery Action Plan*. West Dummerston, VT: Peach Press; 2002.

Danner D. Living with Schizophrenia. Available at https://assets.documentcloud.org/documents/3146953/Living-With-Schizophrenia-by-

Deborah-Danner.pdf

*Davidson L, Rakfeldt J, Strauss J. *The Roots of the Recovery Movement in Psychiatry: Lessons Learned*. Hoboken, NJ: Wiley-Blackwell; 2010.

*Dickerson FB, Lehman AF. Evidence-based psychotherapy for schizophrenia 2011 update. *J Nerv Ment Dis*. 2011;199:520.

Dixon L, Adams C, Lucksted A. Update on family psychoeducation for schizophrenia. *Schizophrenia Bull*. 2000;26(l):5.

Dixon LB, Dickerson F, Bellack AS, et al. The 2009 schizophrenia PORT psychosocial treatment recommendations and summary statements. *Schizophrenia Bull*. 2010;36(1):48.

*Drake RE, Mueser KT, Torrey WC, et al. Evidence-based treatment of schizophrenia. *Curr Psychiat Rep*. 2000;2:393.

Frese FJ. Advocacy, stigma, and self-disclosure: a personal perspective. In: Bromet EJ, ed. *Long-Term Outcomes in Psychopathology Research: Rethinking the Scientific Agenda*. New York: Oxford University Press; 2015.

*Frese FJ, Stanley J, Kress K, Vogel-Scibilia S. Integrating evidence-based practices and the recovery model. *Psychiatr Serv*. 2001;52:1462.

Harding CM, Brooks GW, Ashikage T, Strauss JS, Breier A. The Vermont longitudinal study of persons with severe mental illness, I: methodology, study sample, and overall status 32 years later. *Am J Psychiat*. 1987;144:718.

Haro JM, Novick D, Bertsch J, Karagianis J, Dossenbach M, Jones PB. Cross-national clinical and functional remission rates: worldwide schizophrenia outcome health outcomes (W-SOHO) study. *Brit J Psychiat*. 2011;199:194.

Harrow M, Jobe TH. Does long-term treatment of schizophrenia with antipsychotic medications facilitate recovery? *Schizophrenia Bull*. 2013;39(5):962.

Hasan A, Falkai P, Wobrock T, et al; on behalf of the WFSBP Task Force on Treatment Guidelines for Schizophrenia. World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Schizophrenia. Part 3: Update 2015 Management of special circumstances: depression, suicidality, substance use disorders and pregnancy and lactation. *The World Journal of Biological Psychiatry*. 2015;16:142–170.

Jacobson N. *In Recovery: The Making of Mental Health Policy*. Nashville: Vanderbilt University Press; 2004.

*Kreyenbuhl J, Buchanan RW, Dickerson FB, Dixon LB. The schizophrenia patient outcomes research team (PORT): updated treatment recommendations 2009. *Schizophrenia Bull*. 2010;36(1):94.

Lehman AF, Buchanan RW, Dickerson FB, et al. Evidence-based treatment for schizophrenia. *Psychiat Clin N Am*. 2003;26:939.

Lehman AF, Lieberman JA, Dixon LB, et al. American Psychiatric Association Steering Committee on Practice Guidelines: Practice guideline for the treatment of patients with schizophrenia, second edition. *Am J Psychiat*. 2004;161(2S):1.

Lieberman JA, Stroup S, McEvoy JP, et al., for the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Investigators. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med*. 2005;353(12):1209.

Luciano A, Bond GR, Drake RE. Does employment alter the course and outcome of schizophrenia

and other severe mental illnesses? A systematic review of longitudinal research. *Schizophr Res*. 2014;159:312.

Lundahl B, Moleni T, Burke BL, et al. Motivational interviewing in medical care settings: a systematic review and meta-analysis of randomized controlled trials. *Patient Educ Couns*. 2013;93:157.

McGuire AB, Kukla M, Green M, Gilbride D, Mueser KT, Salyers MP. Illness management and recovery: a review of the literature. *Psychiatric Services*. 2014;65(2):171–179.

Menezes NM, Arenovich T, Zipursky RB. A systematic review of longitudinal outcome studies of first-episode psychosis. *Psychol Med*. 2006;36:1349.

Pilon D, Ragins M. Milestones of Recovery scale. Available at http://milestonesofrecoveryscale.com/. 2015.

Ragins M. A road to Recovery. Available at http://mhavillage.squarespace.com/storage/08ARoadtoRecovery.pdf. 2015.

Shorter E. *A History of Psychiatry*. New York: John Wiley & Sons; 1997.

Silver J, Nemec PB. The role of the peer specialists: unanswered questions. *Psychiatric Rehabilitation Journal* 2016;39(3):289–291.

Slate RN, Johnson WW. *Criminalization of Mental Illness*. Durham, NC: Carolina Academic Press; 2008.

Steadman HJ, Redlich A, Callahan L, Robbins PC, Vesselinov R. Effect of mental health courts on arrests and jail days. *Arch Gen Psychiat*. 2011;68(2):167.

The New York Times' Editorial Board. The Death of Deborah Danner. October 20, 2016. Available at http://www.nytimes.com/2016/10/21/opinion/the-death-of-deborah-danner.html?ref=opinion&_r=1

Torrey EF. *American Psychosis: How the Federal Government Destroyed the Mental Illness Treatment System*. New York: Oxford University Press; 2013.

Tsemberis SJ, Moran L, Shinn M, Asmussen SM, Shern DL. Consumer preference programs for individuals who are homeless and have psychiatric disabilities: a drop-in center and a supported housing program. *Am J Community Psychol*. 2003;32:305.

U.S. Department of Health and Human Services' Substance Abuse and Mental Health Services Administration (SAMHSA)'s National Registry of Evidence-Based Programs and Practices (NREPP). Available at http://www.nrepp.samhsa.gov/Search.aspx?ret=r. 2015.

Waldemar AK, Arnfred SM, Petersen L, Korsbek L. Recover-oriented practice in mental health inpatient settings: a literature review. *Psychiatric Services*. 2016;67:596–602.

Watson AC, Fulambarker AJ. The crisis intervention team model of police response to mental health crises: A Primer for Mental Health Practitioners. *Best Pract Ment Health*. 2012;8(2):71.

Wykes T, Huddy V, Cellard C, McGurk SR, Czobor P. A meta-analysis of cognitive remediation for schizophrenia: methodology and effect sizes. *Am J Psych*. 2011;168:472.

Zipursky RB, Reilly TJ, Murray RM. The myth of schizophrenia as a progressive brain disease. *Schizophrenia Bull*. 2013;39(6):1363.

▲ 12.16 Psychosis as a Defining Dimension in Schizophrenia

DSM-5: AN UPDATED DEFINITION OF PSYCHOSIS

In the fifth edition of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), the psychosis phenotype is categorized into a broad range of clinical diagnoses including schizophrenia, schizoaffective disorder, other functional psychoses (e.g., delusional disorder, psychosis due to substance use or general medical condition), as well as schizotypal personality disorder. The key symptom domains comprising "psychosis" include delusions, hallucinations, disorganized thought process (formal thought disorder), and disorganized or abnormal motor behavior (including catatonia). Despite the absence of unique pathognomonic psychosis features for each diagnosis, that is, any type of psychotic symptoms can occur in any diagnosis, DSM-5 emphasizes that some aspects of psychosis manifestations are characteristic for various psychosis diagnoses. In schizophrenia spectrum disorders (e.g., schizoaffective disorder, schizophreniform disorder), psychotic symptoms are typically comprised of prominent hallucinations and/or delusions as well as formal thought disorder leading to rather severe behavioral manifestations (e.g., disorganized, bizarre, or catatonic behavior). In psychotic disorders due to general medical conditions as well as in substance-induced psychotic disorders, psychosis refers to delusions and/or hallucinations temporally overlapping with the presumably causative event (medical illness, substance use) which do not occur exclusively during the course of delirium. In such rare psychotic diagnoses as delusional disorder and shared psychotic disorder, psychosis is essentially equivalent to isolated persistent delusions. Importantly, only disorders grouped under the "Schizophrenia spectrum and other psychotic disorders" section of DSM-5 are conventionally characterized by psychotic symptoms as the key defining feature comprising the clinical "core" of the diagnosis. Nevertheless, multiple other psychiatric disorders, such as dementias, neurodevelopmental disorders, affective disorders, some personality disorders (besides schizotypal type) may also manifest psychotic symptoms as associated "secondary" diagnostic features.

In schizophrenia, psychosis presents in its most severe form, frequently with bizarre delusions and hallucinations leading to dramatic behavioral manifestations. Although any of five senses may be affected by hallucinatory experiences, the most common hallucinations in schizophrenia are auditory, with multiple voices conversing among themselves, commenting on the patient's life or behavior, or talking directly to the patient. Visual hallucinations are rather common, whereas tactile, olfactory, and gustatory hallucinations are less typical and should alert the clinician to look for underlying neurological or medical conditions that mimic psychotic disorder. Delusions in schizophrenia vary broadly in

content with most classic examples including paranoid/persecutory, religious, or somatic forms. Although not entirely specific for schizophrenia, such delusions as being influenced or controlled by an outside entity or force (delusions of passivity), phenomena of thought insertion, withdrawal or thought broadcasting, and delusional preoccupation with esoteric, symbolic, abstract, or philosophical ideas are more typical for patients with schizophrenia than for other psychotic illnesses. Importantly, Schneiderian psychotic features (i.e., presence of bizarre delusions or hallucinations consisting of a voice that keeps a running commentary on the person's behavior or thoughts or multiple voices conversing with one another) have been de-emphasized in DSM-5 leaving criterion A to a broader range of symptom clusters including (any) delusions, (any) hallucinations, disorganized speech, grossly disorganized or catatonic behavior, and negative symptoms.

Although delusions and hallucinations are commonly perceived as core symptoms of schizophrenic psychosis, they are not an obligate component of the formal diagnosis of schizophrenia. According to DSM-5, since at least two Criterion A symptoms are required to fulfill the diagnosis, patients without either hallucinations or delusions could receive a diagnosis of schizophrenia; however, in these cases other psychotic symptoms are required such as thought disorder (e.g., looseness of associations, thought blocking, derailment, tangentiality, circumstantiality, neologisms, verbigeration, echolalia, global incoherence), grossly disorganized behavior or catatonic symptoms. Although the phenomena of a gross impairment in reality testing and a loss of ego boundaries are not included into formal psychosis definitions, they are commonly seen in patients with schizophrenia, especially during acute psychotic states. These clinical entities describe the lack of a clear sense of where the patient's own mind or body end and where those of other people or inanimate objects begin. In the most severe manifestations, the patients may have experiences that they mentally or physically disintegrated or fused with an outside object or with the entire universe. These components of psychosis may be poorly verbalized by schizophrenia patients due to the complexity of the experiences; however, they can manifest in peculiar, impulsive, or dangerous behavior, and therefore, these symptoms should be carefully evaluated.

Schizophrenic psychosis is usually characterized by a chronic course with either continuous psychotic symptoms or intermittent periods of disease exacerbations and remissions. Course types with a single (or few) episodes followed by a long-term stable remission or clinical recovery are uncommon. However, in other psychotic disorders the lifetime course of psychosis can be intermittent or fleeting, with relatively short episodes of hallucinations and delusions manifesting at the peak of severe depression or mania, decompensated medical conditions, or induced by substance

exposure or medication. In the course of schizophrenia, psychotic symptoms tend to become less severe with time, whereas a decline in cognitive functioning and negative symptoms become more apparent. A further deterioration in the patient's baseline functioning can follow each relapse of schizophrenic psychosis. This failure to return to baseline functioning after each relapse is commonly taken as the major distinction between schizophrenia and psychotic mood disorders, although recent studies report that about 20 to 25 percent of patients with affective disorders show considerable long-term impairment in psycho-social and basic daily functioning. Without treatment, psychosis in schizophrenia and related disorders is frequently incapacitating and contributes, along with negative symptoms and cognitive decline, to an overall poor prognosis of the illness.

CATEGORICAL AND DIMENSIONAL DIAGNOSTIC APPROACHES IN SCHIZOPHRENIA

The symptoms comprising DSM-5 criteria for schizophrenia do not define a diagnosis-specific syndrome and reflect substantial clinical heterogeneity of the illness. No single clinical sign is pathognomonic for schizophrenia; every symptom observed in schizophrenia can occur in different psychiatric disorders. The current categorical concept of schizophrenia (and other psychiatric diagnoses) is based on having a number of different combinations of symptoms, such that many variations of the defining and associated symptoms and symptoms domains are possible. The phenomenology-driven diagnostic systems, while serving the purpose of clinical utility, have critical long-term limitations resulting in classifications that are heterogeneous, lack clear boundaries, and, most importantly, neither map on emerging experimental biomarker constructs nor reflect causal biological disease mechanisms. Despite considerable progress in understanding neurobiological bases of psychiatric disorders, at this point, it is not possible to propose a biologically based diagnostic system due to a lack of defining knowledge about the neurophysiologic and molecular mechanisms of psychiatric disorders. One of the alternative diagnostic approaches espoused to reduce the clinical heterogeneity of psychiatric diagnoses is the identification of psychopathological illness dimensions (such as psychosis) with a research and clinical focus on underlying disease neurobiology. Such symptom dimensions are comprised of groups of symptoms that resemble one another, occur together, and are, hypothetically, based on interrelated pathophysiological and molecular mechanisms.

Three major symptom dimensions are often described in schizophrenia: Psychosis, disorganization, and negative symptoms. The psychosis dimension includes distortions or exaggerations of inferential thinking (delusions) and perceptual disturbances (hallucinations). The

disorganization domain manifests as illogical disorganized thought process and/or grossly disorganized behavior. Although schizophrenic disorganization is commonly perceived as one of the accompanying features of acute psychotic state which tend to resolve as severity of psychosis diminishes, disorganized features are phenomenologically unique, and are specifically associated with poor prognosis, especially in cases of chronic presentation. Therefore, disorganization has been recognized as a distinct clinical domain of schizophrenia, not just a subtype of psychotic symptoms. Negative symptoms dimension includes restrictions in the range and intensity of emotional expression (affective flattening), in the fluency and productivity of thoughts and speech (alogia), and in the initiation of goal-directed behavior (avolition). Recently, a broader dimensional model of schizophrenia has been proposed, including cognitive dysfunction (i.e., deficits in general intelligence level, various aspects of memory, executive function, global processing speed, and attention) as well as affect dysregulation and mood symptoms (Fig. 12.16–1). Although neither is yet a part of the current DSM-5 diagnostic definition, cognitive dysfunction is known to be one of the defining features of the illness which frequently manifests long before first-break psychosis and is highly predictive of disease course and overall prognosis.

Comparative studies contrasting dimensional composites with traditional diagnostic categories have shown that the dimensional approach may be superior in predicting course, demand for psychiatric care, treatment outcome, social adaptation, and global prognosis of illness. Originally, the dimensional models were developed to aid diagnostic classification. However, more recently these dimensional constructs have been targeted in various research paradigms, including genetic and disease biomarker studies, as well as in identifying clinical targets for novel drug development. Dimensional components within the schizophrenia diagnosis may be characterized by unique underlying pathogenetic mechanisms and appear to be, at least partially, pharmacologically independent. To date, effective treatments are available for only a few of these symptom dimensions. Psychotic symptoms may be satisfactorily controlled with the arsenal of first- and second-generation antipsychotic medications, although treatment-refractory cases are not uncommon. Affective symptoms are typically alleviated by antidepressants, mood stabilizers, and some antipsychotic agents. However, no effective treatments exist either for cognitive deficits or negative symptoms, although treatment development is very active in these areas. Clinical experience, as well as well-controlled treatment trials, suggests that psychosis across many diagnostic categories, including such diagnoses as schizophrenia, bipolar disorder, psychotic depression, functional and substance-induced psychoses, and various organic psychoses, show a similar pharmacological treatment response. One of the current foci in drug development for schizophrenia is treatment

for cognitive dysfunction, with hypothetical treatment targets focused on glutamatergic, cholinergic, and serotoninergic neurotransmitter systems.

PSYCHOSIS SPECTRUM DISORDERS: DOES PSYCHOSIS CUT ACROSS THE FORMAL DIAGNOSTIC CATEGORIES?

Among the clinical manifestations of psychiatric illnesses, psychosis represents perhaps the most dramatic and clinically captivating phenomenon. Historically, in early German and French schools of psychiatry in the 18th and early 19th centuries, psychosis was described as a single clinical entity and often was the hallmark of mental illness. Since 1896, when Emil Kraepelin demarcated "dementia praecox" from "manic depressive insanity," the nosological (categorical) mode of thinking and classification has been predominant in psychiatry worldwide. This is despite the fact that Kraepelin's concept underwent a complex evolution that led him to a considerable personal uncertainty about making clear distinctions among psychosis spectrum disorders, such as dementia praecox, manic depressive psychosis, intermediate conditions, and paranoid disorders. In the current psychiatric classification, psychosis is acknowledged as a clinical phenotype within numerous psychiatric disorders, including schizophrenia, mood disorders, substance use disorders, cognitive and neurodevelopmental disorders, and disorders related to general medical conditions, although it is not an obligate component of their formal diagnoses (Fig. 12.16–2).

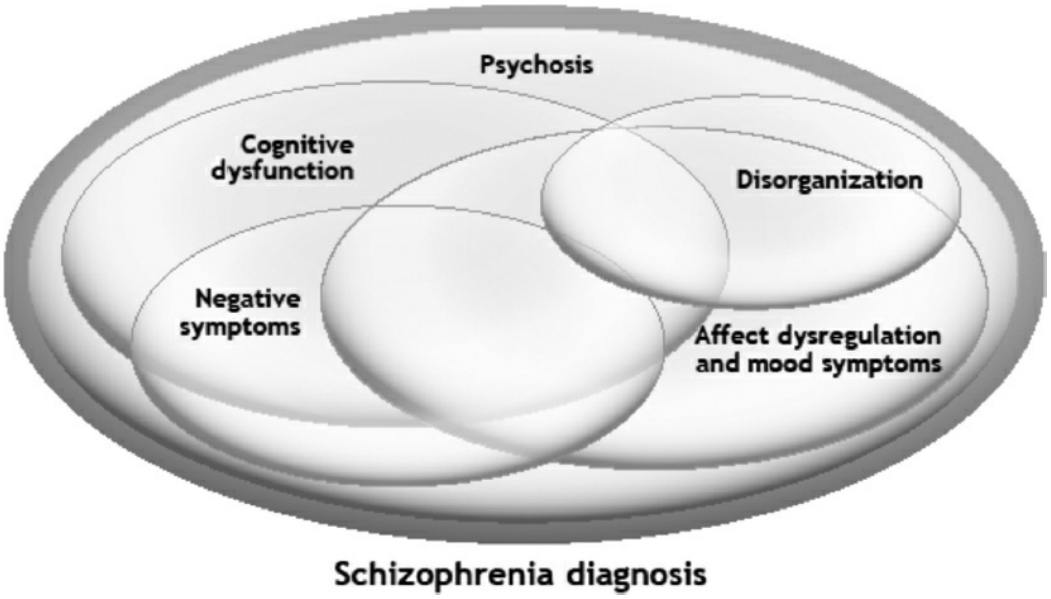

FIGURE 12.16–1. Symptom dimensions of schizophrenia.

Large-scale research studies that utilized multivariate factor analytic approaches applied to samples spanning various categorical diagnoses suggest that psychosis represents a distinct clinical domain overlapping a

broad spectrum of psychiatric disorders. The strategy of breaking down phenomena into distinct dimensions has been broadly used in different medical sciences and is commonly used in psychology. But, at the present time, there is not a full consensus on how to subclassify various psychosis spectrum disorders into meaningful, biologically based dimensions. In recent studies of patients with psychotic disorders (schizophrenia, schizoaffective disorder, psychotic mood disorders, and other functional psychosis) and their biological relatives, five distinct symptom domains spanning these diagnoses were described, including reality distortion (psychosis), disorganization, mania, depression, and negative symptoms, commonly referred to as the "pentagonal model" of psychotic disorders. Different diagnostic categories showed distinct dimensional profiles, with schizophrenic persons scoring highest in the reality distortion, negative, and disorganized symptoms; and affective diagnoses scoring highest in manic and depressive dimensions and lower in the psychosis and negative symptoms. Interestingly, the mania factor is considered to be the best discriminator between schizophrenia and mood disorder diagnoses, while psychosis turned out to be a domain of overlap between these diagnostic categories. In all dimensional constructs, schizophrenia patients typically show higher variability compared to patients with mood disorders, suggesting that bipolar disorder diagnosis may signify a better defined phenomenological construct than schizophrenia, which appears to be highly heterogeneous. These studies suggest quantitative variations in symptom dimensions across diagnostic categories rather than qualitative differences. However, multiple questions remain, including whether or not the phenomenological structure of psychosis is the same across diagnoses, what is the extent of similarities and uniqueness of underlying genetic and pathophysiological mechanisms of psychosis is, and whether various psychosis spectrum disorders show the same molecular basis, treatment response and functional outcomes.

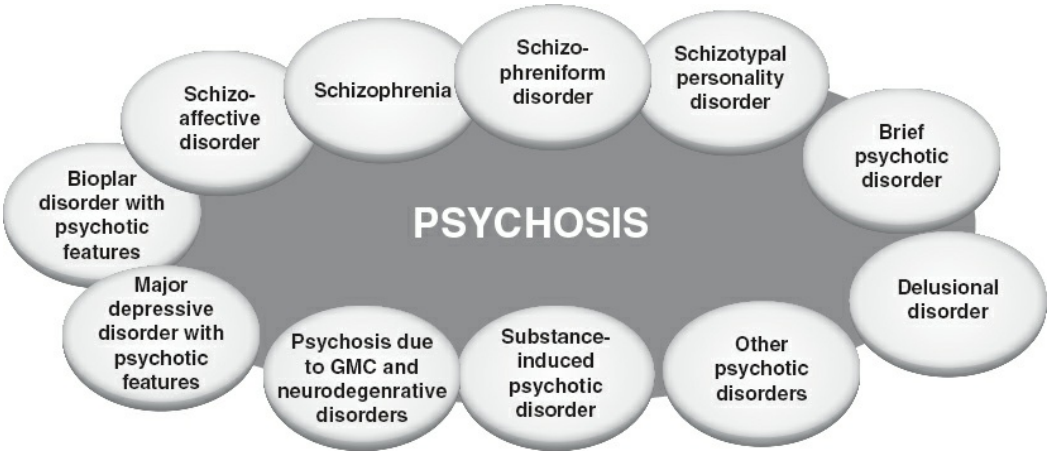

FIGURE 12.16–2. DSM-5 diagnoses overlapping the psychosis dimension. GMC, general medical conditions.

BIOLOGICAL FACTORS IN PSYCHOSIS: EVIDENCE FROM TRANSLATIONAL NEUROSCIENCE INVESTIGATIONS

Genetic Studies

Psychotic illnesses are broadly heterogeneous disorders, not caused by a single etiological factor, but rather due to a complex interplay of different endogenous and environmental pathogenic influences. There is consistent evidence that a genetic component contributes to the etiology of psychosis; however little is known for certain about particular genomic regions or individual risk genes. Recent findings from large-scale whole genome studies provide evidence for contribution of numerous genes, each of a very small effect, to ethiopathogenesis of schizophrenia. Moreover, a substantial overlap in genetic susceptibility across a broad spectrum of psychiatric disorders, for example, schizophrenia, bipolar disorder, autism, has been observed. Some of the identified candidate genes show strong associations with dimensional constructs, such as psychosis or mood symptoms, that are not projected directly onto categorical diagnoses. Hypothetically (not yet well established) genes or sets of genes, interacting with environmental factors, may predetermine vulnerability to psychosis and other clinical disease manifestations. Depending on genetic influence and other endogenous (e.g., neurodevelopmental) and/or environmental interactions, psychosis may coexist with different clinical phenotypes (e.g., mood symptoms or cognitive dysfunction), composing epiphenomenal categorical diagnoses. This conceptualization of psychosis draws preliminary support from epidemiological and molecular genetic studies.

Family Studies. Family studies of schizophrenia and affective psychoses have established that psychosis strongly aggregates in families. Having a relative with schizophrenia or bipolar disorder is a single most powerful risk factor for developing psychosis. The life-time risk for developing schizophrenia increases approximately 8 to 12 folds in first-degree biological relatives of schizophrenic probands. Although the results of genetic studies in affective psychoses are less consistent, the familial aggregation of bipolar disorder and major depressive disorder has been observed. First-degree relatives of individuals with bipolar I disorder have elevated rates of bipolar I disorder (4 to 24 percent), bipolar II disorder (1 to 5 percent), and major depressive disorder (4 to 24 percent). While it was originally thought that schizophrenia and affective psychoses are inevitably segregated, recent reports challenge this. Family studies show that schizophrenia and affective psychoses can co-occur in the same families, suggesting shared familial risk. For example, a large epidemiologic study, using a genetically homogenous population isolate in north-eastern Finland, reported that 16 percent of siblings of schizophrenic probands carried a lifetime diagnosis of psychotic disorder, including schizophrenia

(0.5 percent), schizoaffective disorder (3.3 percent), bipolar I disorder (1.1 percent), major depressive disorder with psychotic features (2.2 percent), alcohol-induced (2.2 percent), and other substance-induced (1.1 percent) psychoses, delusional disorder (1.6 percent) and psychotic disorder, not otherwise specified (4.4 percent). Over half of the siblings (54 percent) in this study had a lifetime diagnosis of any mental disorder, with major depressive disorder (15.3 percent) and anxiety disorders (12 percent) leading among non-psychotic diagnoses. Numerous reports confirmed that the increased risk for psychotic illness in relatives of individuals with schizophrenia does not appear to be confined to schizophrenia alone. On the other hand, bipolar illness has been associated with increased risk of schizophrenia in relatives. In one of the recent family studies it was reported that relatives of women with early onset of bipolar disorder had the highest morbid risks for both bipolar illness and schizophrenia. The presence of more than one patient with bipolar disorder in a family increased the risk for schizophrenia nearly fourfold. In a different study it was shown that affective disorders were more frequently inherited from the same parental side of the family as schizophrenia psychosis, suggesting that the same genetic factors could contribute to susceptibility to both schizophrenic and affective psychoses. Several studies showed that, in fact, parents of schizophrenic probands suffer from affective disorders (especially, major depressive disorder) more often than from schizophrenia. Schizoaffective disorder occurs at similarly increased rates both in families of probands with schizophrenia and bipolar disorder. Likewise, schizophrenia and bipolar disorder have been shown to occur at increased rates in relatives of probands with schizoaffective disorder.

Further, there is growing evidence that certain clinical phenotypes and symptom dimensions may predict familial risk of psychosis across categorical diagnosis. For example, in a large family study from the Netherlands a syndrome, characterized by bizarre behavior, inappropriate affect, catatonia and poor rapport, predicted psychosis in biological relatives of schizophrenic probands. The syndrome characterized by affective blunting and insidious early onset of illness predicted psychosis in the first-degree relatives of probands with various types of functional psychoses. A lifetime history of mania specifically predicted affective psychosis in the relatives of probands with various psychotic disorders. Similarly, in a recent study, conducted in a sample of psychotic inpatients and their 2,987 first-degree relatives in Spain, severe primary negative symptoms in probands were related to familial liability for schizophrenia, whereas familial morbid risk of mood disorders was predicted by mania, depression and catatonia syndromes in probands. Importantly, this familial association pattern was completely independent of the probands' diagnosis of schizophrenia or major mood disorder. These studies provide evidence that familial liability for psychosis clearly overlaps DSM categories of

schizophrenia and mood disorder. From a dimensional perspective, psychosis may represent a shared phenotype with unique genetic etiologies, running through family generations.

Twin Studies. Twin studies show that the concordance rate for schizophrenia is higher in monozygotic twins (47 to 56 percent) than in dizygotic twins (12 to 16 percent), suggesting a strong heritability component for the illness. It has been reported that the more severe the schizophrenia, the more likely it is for the twins to be concordant for the disorder. Some European studies have reported a concordance rate for monozygotic twins of over 80 percent in cases of severe schizophrenia with typical core symptoms. These observations suggest that the schizophrenia diagnosis includes clinically and etiologically heterogeneous subgroups with potentially distinct underlying genetic backgrounds. Further, twin studies suggest that a schizophrenia diagnosis in one twin equally increases the risk for schizophrenia and affective disorders in the co-twin. A clear overlap in genetic risk for schizophrenia, schizoaffective, and manic phenotypes was reported in a sample of 77 monozygotic and 89 same-sex dizygotic twin pairs, ascertained from the Maudsley Twin Register in London. In this study, if one member of a monozygotic twin pair had schizophrenia, there was about an 8 percent chance of schizoaffective disorder diagnosis in the co-twin and an 8 percent risk of mania. Another report, based on the Maudsley twin series, showed that the maximum monozygotic/dizygotic concordance ratio was produced by a combination of schizophrenia, affective disorder with mood-incongruent psychotic features, schizotypal personality disorder, and atypical psychosis, suggesting that these psychosis spectrum disorders share a common genetic background.

Genetic Linkage Studies. Genetic linkage studies have identified several chromosomal regions showing linkage to both schizophrenia and affective psychosis, including 1q32, 10p11-15, 13q32, 18p11.2, and 22q11-13. In one study significant linkage to psychotic bipolar disorder was obtained on chromosomes 9q31 and 8p21. Several other chromosomal sites were supportive of linkage with psychosis in bipolar pedigrees, including 5q, 6q, 8p, 13q, 15q, 17p, and 18q, which are also implicated in schizophrenia. A recent genome-wide linkage scan in schizoaffective disorder, using British and Irish pedigrees with at least one member diagnosed with schizoaffective disorder, bipolar type, confirmed the existence of loci that influence susceptibility across the functional psychosis spectrum. This study demonstrated genome-wide significant linkage at chromosome 1q42 and suggestive linkages at 22q11 and 19p13. The evidence for these linkages was equally contributed by schizophrenia pedigrees (where schizophrenia was predominant among relatives in addition to schizoaffective disorder)

and bipolar disorder pedigrees (with higher proportion of affective psychosis). Noteworthy, two candidate genes, *DISC1* (disrupted in schizophrenia 1) and *COMT* (catechol-O-methyltransferase), which have been consistently implicated in schizophrenia and, more recently, in bipolar disorder, map to 1q42 and 22q11, respectively.

Studies of Individual Genes and Genome-Wide Association Studies. Studies of individual genes also support the evidence of overlap between schizophrenia and different psychotic disorders. In most studies, the candidate genes were first tested in schizophrenia, but a growing number of reports suggest that some of these genes may also be involved in etiologies of affective psychoses and even broader psychopathology, for example, autism spectrum disorders and epilepsy. In particular, recent genetic studies have shown increasing evidence for an overlap in genetic susceptibility across the traditional classification categories, including association findings at *DISC1*, *COMT*, *DAOA* (G72)/G30), (D-amino acid oxidase activator (G72)/G30), dysbindin (dystrobrevin-binding protein 1), *BDNF* (brain derived neurotrophic factor), *NRG1* (neuregulin 1), and *RGS4* (regulator of G protein signaling 4). For instance, *NRG1* has been consistently implicated in schizophrenia; more recently, its relationship to bipolar disorder as a candidate susceptibility gene has also been reported. A recent study found the association of *NRG1* with a clinical phenotype of bipolar disorder with mood-incongruent psychotic symptoms, as well as with schizophrenia with lifetime manic episodes, suggesting that *NRG1* may confer susceptibility to a specific clinical phenotype with combined features of psychosis and mania. A number of independent genetic linkage and association studies in diverse populations support that variation in *DISC1* gene influences susceptibility to disorders of psychosis spectrum, including schizophrenia, schizoaffective disorder, bipolar disorder, and major depression. Although this gene has been named "disrupted in schizophrenia," in fact, the family in which the *DISC1* translocation was originally observed included cases of both psychosis and mood disorder. Recently several independent groups have implicated association of *DISC1* with bipolar disorder, and in one study the strongest association was observed with schizoaffective disorder, suggesting that *DISC1* influences susceptibility to a mixed psychosis-mood phenotype that cuts across the traditional schizophrenia/bipolar divide. *Dysbindin* is another example of a candidate gene that was initially implicated in schizophrenia. Nevertheless, recent reports have also demonstrated an association between a set of polymorphisms in this gene and a clinical subtype of bipolar disorder with recurrent psychotic episodes, specifically psychotic mania. Overall, these findings support variations in *NRG1*, *DISC1*, and dysbindin as potential genetic effects contributing (along with many other genetic and environmental factors) to susceptibility for psychosis or a mixed phenotype

with features of both psychosis and mood symptoms rather than to the DSM "schizophrenia" phenotype, although these genetic markers were indeed originally tested as schizophrenia candidate risk genes.

Polymorphisms in *G72 (DAOA)/G30* have been broadly linked to both schizophrenia and bipolar disorder. Furthermore, a recent study from the United Kingdom, conducted with a large sample of schizophrenia and bipolar disorder cases, implicated several genetic variations in *G72 (DAOA)/G30* as specific risk markers for major mood episodes across the bipolar and schizophrenia diagnoses. Thus, although this locus was originally described as a schizophrenia risk gene, this large-scale report suggests that G72 (DAOA)/G30 appears to be more strongly associated with the mood symptoms domain, not psychosis, within the schizophrenia-bipolar diagnostic continuum.

Epidemiological and molecular genetic studies support the hypothesis that psychosis is a clinical phenotype with multifactorial etiologies, among which a genetic component is crucial. Family studies of schizophrenia and affective psychoses have established that psychosis strongly aggregates in families. Twin studies have suggested high heritability estimates of psychosis with a complex mode of transmission. Recent whole genome linkage studies have suggested a number of chromosomal loci that influence susceptibility to psychosis, independently of diagnostic categories. Detailed studies of linked genomic regions have identified several putative candidate genes that appear to be involved in genetic etiologies of schizophrenia and affective psychoses. Since schizophrenia is a relatively rare illness with presumably multiple genes of small to very small effects involved in its etiology, the majority of available to date genetic studies are still statistically underpowered. Geneticists suggest that tens of thousands of cases and controls may be needed to find firm associations. For example, the Schizophrenia Psychiatric Genome-Wide Association Consortium recently conducted a multi-stage genome-wide association study (GWAS) including 36,989 schizophrenia cases and 113,075 controls identified 108 loci associated at genome-wide significance with the schizophrenia phenotype. Examination of specific candidate genes at these loci suggested the involvement of dopaminergic regulation (*DRD2* gene), multiple genes involved in glutamatergic neurotransmission and synaptic plasticity (e.g., *GRM3, GRIN2A, SRR, GRIA1*), as well as neuronal calcium signaling (e.g., *CACNA1C, CACNB2*, *CACNA11*). Independent of genes expressed in brain, associations were enriched among genes expressed in tissue that play important role in immunity (in particular, relevant to B-lymphocyte lineages involved in acquired immunity), providing support for the hypothesized link between the immunological effects and schizophrenia. In addition to polygenic, complex inheritance observed in schizophrenia and other psychotic disorders, there is also strong indication that in some patients a deletion or insertion of a larger chromosomal

region, so-called copy number variation (CNV), might play a crucial role in the disease pathogenesis. This could be specifically important in sporadic cases of schizophrenia, since a higher frequency of *de novo* mutations has been associated with these CNVs. Noteworthy, CNV markers linked to schizophrenia show substantial overlap with neurodevelopmental disorders, such as intellectual disability and autism spectrum disorders, providing support for partially overlapping etiologies underlying these disorders.

Current clinical targets available for genetic studies, for example, "schizophrenia" or "bipolar disorder," are highly heterogeneous and do not clearly map onto molecular or neurophysiologic disease models. Understanding the biological effect of genetic risk markers is undoubtedly complex. Although several such candidate markers have been implicated, the associating variations are likely different in different disease subtypes and symptom clusters, and it is therefore difficult to determine the specific biological effect of each risk gene characteristic for specific disease components. Various epigenetic factors, although poorly explored so far, add to the complexity of the picture. In addition, environmental factors, perhaps interacting with risk genes, are likely to contribute and further complicate endogenous and environmental factors interplay. Although originally the candidate risk genes were implicated in schizophrenia, recent findings provide evidence that they show strongest associations with symptom dimensions, such as psychosis (e.g., *NRG1*, *DISC1*, and *dysbindin*) or mood symptoms (e.g., *G72/G30, BDNF*), across the schizophrenia–mood disorder continuum. Furthermore, recent GWAS studies support much broader genetic overlaps, from psychotic disorders, to autism spectrum disorders, to epilepsy and neurodegenerative disorders, to medical conditions (e.g., autoimmune disorders). Although intriguing, these observations are still preliminary and should be treated with caution, given the difficulties in establishing defining evidence for genetic etiologies in complex, polygenetic disorders. Future genetic studies derived from large samples, targeting more homogenous disease components (e.g., symptom dimensions) and/or biomarker-derived constructs underlying clinical phenomenology, are needed.

Emerging Candidate Biomarkers for Psychosis

Despite considerable progress in epidemiological and molecular genetics of psychiatric syndromes made in the past decade, clinical phenotypes, such as schizophrenia or bipolar disorder, when directly applied in linkage analyses or studies of individual genes, generally fail to result in consistent findings. Although multiple genome-wide "hotspots" and associated candidate genes have been implicated, the causative "psychosis genes" have not been identified. Even less is known about specific pathways and brain circuitries through which these genetic effects may contribute to symptom

manifestations of psychotic disorders. The inconsistent genetic findings may be, in part, explained by the presumed complexity of the pathophysiologic mechanisms of psychiatric disorders, resulting from a multicomponent interplay of genetic/epigenetic, endogenous and environmental factors. In addition, the phenomenological heterogeneity of psychotic disorders as well as lack of biologically based definitions and clear boundaries in the existing diagnostic categories contribute to the limited success in identifying causative genetic etiologies. Based on these kinds of considerations, the focus of genetic and translational studies has recently shifted toward disease intermediate phenotypes, or biomarkers. Intermediate phenotypes (often called biomarkers) are specific measurable alterations in brain function or structure, or abnormalities in peripheral systems (e.g., hyperactive inflammatory cascades, endocrine effects) reflecting components of disease pathogenesis which are thought to be "intermediate" between the aberrant genes and the overt clinical manifestations of disease. These alterations are frequently referred to as "disease biomarkers" and can be assessed by objective neurophysiological, neuroanatomical, or behavioral measurements. An ideal intermediate phenotype for an illness would be a measure that is associated with the clinical syndrome, heritable and state-independent. In addition, it co-segregates in families while being expressed in unaffected family members and may serve as an informative predictor or risk marker observed long before full disease manifestations. Intermediate phenotypes are a promising area of study in the search for genetic factors in psychosis which may provide more direct clues to the genetic underpinnings than does the clinical syndrome itself. On the other end, dimensions of disease, such as psychosis, may be particularly useful in identifying the underlying intermediate phenotypes, as dimensions clearly provide less heterogeneous clinical constructs.

Biomarker-focused studies in psychiatry started in the field of schizophrenia research and up-to-date led to the recognition of several putative disease biomarkers. These include oculomotor abnormalities (smooth pursuit eye movement alterations and saccadic eye movement disinhibition); deficits in sensorimotor gating, measured with electroencephalography-based event-related potential (ERP) and prepulse inhibition (PPI) of startle response paradigms; abnormal neurocognitive performance; as well as structural and functional brain abnormalities captured with various imaging methodologies. In contrast to the considerable amount of data in schizophrenia, much less is known about intermediate phenotypes of other psychotic disorders, although selected candidate biomarkers have recently begun to be explored in disorders of neurodevelopmental and affective spectrum. Similar to overlapping clinical symptom manifestations, the three most common psychotic disorders, schizophrenia, schizoaffective disorder and bipolar disorder, appear to

share common etiologic determinants and biomarker signatures (Fig. 12.16–3). Within the array of neurophysiological, anatomic, and cognitive measures, there are specific elements that can be targeted as candidate psychosis biomarkers.

Oculo-Motor Alterations in Psychotic Disorders. Similar abnormalities in smooth pursuit eye movement (i.e., inability to closely follow a moving target) and saccadic eye movement disinhibition have been observed in schizophrenic and psychotic bipolar disorder patients, as well as in their first-degree relatives and persons with mild psychotic symptoms, such as those with schizotypal personality disorder. These oculomotor phenotypes are thought to be mediated by frontal cortex and have been linked to working visual spatial memory function as well as deficits of prefrontal inhibitory control. These oculo-motor deficits are stable, trait-like characteristics, observed in both acute (psychosis, acute mania) and remission disease phases, and seem to be minimally influenced by medications. Recent genetic studies of oculomotor biomarkers suggested linkage of smooth pursuit eye movement alterations to chromosome 6p21 in schizophrenia families. Other studies, using fMRI simultaneously with eye tracking during the imaging sessions, have reported that schizophrenic patients tend to activate different brain regions in order to maintain smooth visual tracking compared to healthy subjects. Patients showed reduced activation in frontal and, specifically, supplemental eye fields regions, medial superior temporal and anterior cingulate cortices. Interestingly, activation in medial occipito-temporal cortex was increased in patients compare to healthy individuals, perhaps, reflecting recruitment of different brain regions in order to compensate for a primary prefrontal deficits underlying smooth pursuit eye movement system dysfunction in schizophrenia.

Sensory and Sensory-Motor Gating Deficits in Psychotic Disorders. Sensory/sensory-motor gating is a neurophysiological function of a normal brain that allows to “filter out” irrelevant and/or excessive environmental stimuli. It has long been observed that patients with evolving psychosis and mania display hypersensitivity to environmental stimuli as well as aberrant stimuli attributions, thought to be mediated by altered sensory and sensory-motor gating mechanisms. These phenomena are typically measured by various ERP paradigms using electroencephalography (EEG) with such probes as P50 (positive EEG wave observed within 50 ms of a target auditory stimulus) and P300 (positive EEG wave observed within 300 ms of an “odd-ball” auditory stimulus), as well as electromyography (EMG) reflecting sensory-motor gating function (e.g., blink response in the prepulse inhibition [PPI] paradigms).

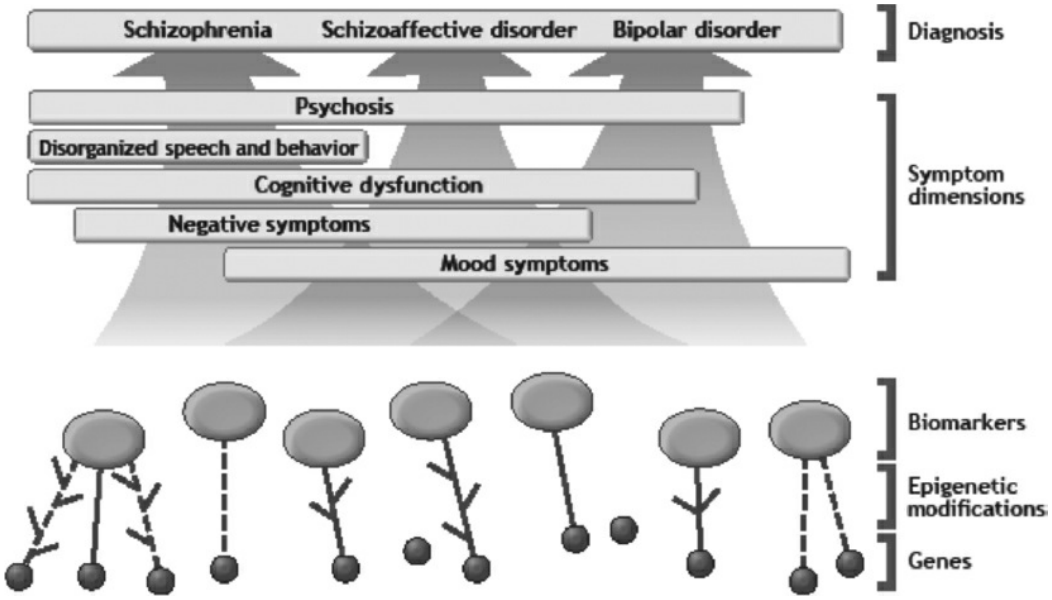

FIGURE 12.16–3. Clinical, biomarker, and genetic components in the schizophrenia–bipolar disorder continuum. Schizophrenia, schizoaffective disorder, and bipolar disorder are the three main psychiatric diagnoses overlapping the psychosis dimension. In addition, other symptom dimensions, such as cognitive dysfunction, disorganization, negative and affective symptoms, are variably presented in these diagnoses. These symptom manifestations are thought to result from a complex interplay of environmental and biological factors, expressed in disease biomarkers such as altered brain anatomy, cognitive or neurophysiologic abnormalities, captured with electroencephalography, brain imaging and other objective measurements. Some, though not all, of these biomarkers may be heritable and subserved by genetic and epigenetic effects.

ELECTROENCEPHALOGRAPHY-BASED BIOMARKERS IN PSYCHOTIC DISORDERS. **Auditory Paired Stimuli ERPs**. An overall pattern of gating deficits characteristic for psychosis is a reduced sensory and sensory-motor gating function thought to underlay environmental stimuli hypersensitivity phenomena. In the standard auditory "paired stimuli" paradigm, so called "P50" response amplitude (i.e., the largest EEG wave within 50 ms after the stimulus presentation) is measured in response to each of two auditory "clicks." In healthy individuals, the P50 wave following the second stimulus is lower than that to the first stimulus, reflecting normal sensory gating function. In contrast, both schizophrenia and bipolar patients (specifically, patients with acute mania) show diminished suppression of P50 ERP associated with insufficient sensory gating. These sensory gating deficits are thought to be stable "trait" measures largely independent of active severity of psychotic or mood symptoms or treatment effects. Some studies report more prominent P50 abnormalities in patients with disorganized schizophrenia, although an association between disrupted P50 and negative symptoms has also been suggested. An intriguing observation, suggesting that P50 sensory gating deficit may be a unique psychosis intermediate phenotype, has been recently reported in patients with bipolar disorder. Persons with bipolar I and lifetime history of psychosis and those with schizoaffective disorder showed P50 deficits similar to those

observed in schizophrenia, whereas bipolar patients with no history of psychosis exhibited normal P50 suppression. Sensory gating deficits are thought to be heritable and have been linked to various psychotic disorders aggregating in families, including schizophrenia. "Schizophrenia-like" P50 suppression deficits are also observed in people with schizotypal personality disorder as well as in individuals at high risk for developing psychosis, such as subjects originating from high genetic density families (i.e., those with several psychotic family members) and individuals with schizophrenia prodrome, typically preceding frank psychosis manifestations. Evidence for a linkage between disrupted P50 gating and chromosome 15q14 has been reported in several studies, specifically to the promoter region of the α-7 nicotinic cholinergic receptor subunit gene, in both individuals with schizophrenia and their relatives.

Odd Ball Paradigm and P300 ERP. P300 is typically measured in auditory or visual "odd ball" paradigms where frequently presented stimuli ("targets") are interleaved with infrequent and unexpected atypical ("odd ball") stimuli and refers to a positive EEG wave occurring within 300ms following an "odd ball" stimulus. Several large meta-analyses reported reduced P300 ERP amplitude and delayed P300 latency in probands with schizophrenia and their biological relatives, as well as in "mild psychosis phenotypes" such as schizophrenia prodrome and schizotypy. Recently studies of P300 ERP have been extended to bipolar disorder, and suggested similar "schizophrenia-like" alterations in P300 amplitude and latency in individuals with psychotic bipolar disorder and their relatives, albeit less consistently across different studies. In several studies a frontal subcomponent of P300 ERP was reported to correlate with severity of hallucinations in patients with schizophrenia, perhaps due to shifting of attentional resources from external to internal (psychotic) stimuli. Similarly, patients with psychotic bipolar disorder showed reductions in frontal lobe P300 voltage, suggesting a component of overlapping neural substrate for dysfunction in both schizophrenia and affective psychosis. Diminished P300 amplitude and increased latency have been also linked to severity of negative symptoms, illness duration and early age of psychosis onset, as well as several cognitive dysfunctions, that is, deficits in attention. Several studies, conducted on healthy twin pairs suggested moderate heritability estimates for P300 ERP biomarker. Interestingly, the P300 heritability in females appeared to be lower (45 percent compared to 56 percent in males), suggesting that endogenous or environmental factors may have a stronger modifying influence on the P300 phenotype in females. Genetic association studies linked P300 alterations to the *COMT* genotype (i.e., Met homozygous variant) as well as *NRG1* polymorphisms in patients with schizophrenia.

Sensorimotor Gating Deficits in Psychosis: Prepulse Inhibition. PPI paradigms measure inhibition of the startle reflex. In

healthy individuals, a weaker "prepulse" stimulus leads to an inhibition of a response magnitude to a startle ("pulse") stimulus, used as an index of sensorimotor gating. In individuals with psychotic disorders insufficient PPI reflecting an exaggerated startle response is commonly observed. Diminished PPI has been reported in patients with schizophrenia, schizotypal personality disorder, and both affected and unaffected biological relatives of schizophrenia probands. Similar findings have been noted in persons with bipolar disorder, especially during acute manic phase, although the studies in bipolar are few and less consistent. Several studies reported associations between PPI deficits and florid psychotic symptoms, such as disorganized thought process and behavior, greater distractibility, and bizarre psychotic behavior. However, at least a few studies in schizophrenia observed disrupted PPI even in remission phases, supporting this candidate biomarker as a stable "trait" measure. Likewise, individuals with acute psychotic mania show significantly reduced PPI and less startle habituation, along with a significant reverse correlation between PPI sensorimotor gating and severity of psychotic symptoms. In contrast, euthymic bipolar patients failed to demonstrate abnormalities in PPI. These observations of similar PPI deficits in schizophrenia and psychotic mania, abnormal sensorimotor gating may represent one of the putative biomarkers subserving psychosis as a unique clinical phenotype.

Cognitive Dysfunction as a Candidate Psychosis Biomarker. Cognitive dysfunction has recently drawn interest as a distinctive phenotype across psychotic disorders, as well as a treatment target. The presence of specific cognitive deficits is not yet part of the diagnostic DSM-5 criteria for schizophrenia; nonetheless it is recognized as a common and consistent feature of the disorder. Extensive cognition studies have shown stable, trait-like deficits in the domains of working memory, verbal episodic memory, attention, processing speed, language comprehension, problem solving, and executive function in individuals with schizophrenia. Patients with bipolar disorder manifest deficits in overlapping cognitive domains, with attention, working memory, and verbal memory difficulties being among the most prominent, albeit generally less severe that those found in schizophrenia. A considerable clinical heterogeneity of the bipolar phenotype correlates with variable presentations of cognitive deficits, with some patients manifesting profound "schizophrenia-like" abnormalities, and others showing little or none alterations in cognitive function. It appears that the greater the burden of illness associated with the disorder, as manifest in the lifetime history of psychosis, number of past manic episodes, an overall duration of illness and number of past hospitalizations, the greater the neurocognitive deficits typically found in bipolar patients. Moreover, bipolar patients with a lifetime history of psychosis appear to have a slightly different profile of neurocognitive deficits than those without a history of psychosis, manifesting more severe impairment on

measures of executive function and spatial working memory. More severe bipolar patients, especially those with psychosis, begin to resemble schizophrenia patients with regard to the level and profile of neurocognitive deficits. One of the commonly reported cognitive features that may help to disambiguate schizophrenia and bipolar disorder is a general intelligence level, especially, premorbid IQ, with lower scores typically observed in patients with schizophrenia but normal IQ levels in bipolar disorder. An extensive research work has been done in attempt to discover genetic associations for cognitive dysfunction in schizophrenia and related psychotic disorders, with most consistent findings supporting links for *COMT* and executive function, *DISC1* and verbal declarative and working memory deficits in schizophrenia and, less consistently, bipolar probands, and their biological relatives.

Anatomical Brain Alterations in Psychotic Disorders. Anatomical brain changes have been examined as candidate psychosis phenotypes across schizophrenia and related psychotic disorders. Studies in schizophrenia have consistently reported a modest but measurable reduction in whole brain volume, particularly in gray matter (averaging 2 to 8 percent for various brain structures, based on meta-analyses), as well as an increase in ventricular size. Several definite areas of the brain have been consistently reported to demonstrate volumetric changes, such as decreased volume in several frontal (e.g., dorso-lateral, orbito-frontal), temporal (medial temporal lobe [MTL], particularly, the hippocampal formation; STG), and parietal cortical regions, as well as thalamus. Several meta-analyses of volume change in schizophrenia reported whole brain, and, particularly, fronto-temporal gray matter volume reductions in individuals at the time of their first psychotic break, preceding any treatments.

The structural findings in psychotic mood disorders are less consistent, ranging from decreased, to normal, to increased gray and white matter volumes in various brain regions. Recent large-scale psychosis-focused studies report regionally overlapping but more focused and lesser in magnitude gray matter volume reductions in psychotic bipolar disorder compared to schizophrenia. The most reproducible regional volume reductions in psychotic bipolar disorder are typically localized to anterior limbic regions, that is, orbito-frontal, anterior cingulate, insular cortex, as well as amygdala and anterior temporal cortex. At the same time, volume increases in amygdala have been reported. Some studies, comparing individuals across the psychosis dimension, find no abnormalities in gray matter volume, while volume deficits of white matter were evident throughout the neocortex and brainstem, with similar regions of white matter loss found in schizophrenia and bipolar patients.

The question of whether volume loss is progressive during the lifetime course of psychosis has been explored. Measurable gray matter volume and

density changes are typically observed in individuals near psychosis onset, progress with illness duration, more aggressively during the initial years of psychosis in some cases, and ultimately result in characteristic gray matter alterations found in chronic psychosis samples. Data from relatives with psychosis spectrum disorders (e.g., schizotypal personality disorder) demonstrate milder but similar gray matter volume reductions suggesting a heritable link between psychosis manifestations and brain structure phenotypes. Moreover, recent large-scale reports demonstrate overlapping gray matter volume reductions in the psychosis probands (including schizophrenia, schizoaffective disorder, and psychotic bipolar disorder cases) and their first-degree relatives with psychosis spectrum disorders, but normal gray matter characteristics in relatives unaffected by lifetime psychosis, even those with clear non-psychotic diagnoses (e.g., mood disorders, anxiety disorders). These gray matter alterations may reflect "psychosis burden," from extensive reductions in probands where psychosis is fully manifested, to similarly distributed but milder alterations in relatives with mild psychosis spectrum phenotypes, to normal gray matter in relatives without lifetime psychosis. Moreover, associations between lifetime duration of psychosis and gray matter volume, with more loss observed as illness progresses, have been reported in schizophrenia and schizoaffective probands in frontal, temporal cortex, basal ganglia and thalamus, the regions long implicated in psychosis in both first episode and chronic samples.

The exact mechanisms underlying these brain structure changes are unknown, but can be interpreted in the context of available post-mortem findings of cellular neocortical pathology in psychotic disorders. Histological examination of schizophrenia cases indicates no neocortical neuronal loss, yet altered neuronal cell packing, presumably due to reduced interstitial neuropil, resulting in higher neuronal density and overall decreased cortical thickness. In contrast, analyses from bipolar cases show decreased neuronal and glial density but normal overall cortical thickness. Dendritic spine loss, as well as reduced dendritic length in dorsolateral prefrontal cortex, has been observed in both schizophrenia and bipolar post-mortem tissue samples. These post-mortem findings suggest both overlapping and unique anatomical underpinnings for schizophrenia and bipolar disorder, providing plausible cellular correlates for the structural brain changes captured with imaging techniques.

Furthermore, longitudinal studies provide support for a measurable effect of various psychotropic medications on brain structure. Long-term treatment with first-generation antipsychotics has been linked to increased gray matter volume and density in basal ganglia. Both first- and second-generation antipsychotics are associated with decreased gray and white matter volume in the fronto-temporal regions. Lithium has been linked to higher gray matter volume/density in diffuse neocortical regions as well as

larger volume of specific subcortical structures, that is, amygdala, in psychosis and mood disorder samples. Given that the majority of psychotic and mood disorder patients undergo years of treatment with various psychotropic agents, it is possible that the gray matter deficits consistently reported in patients with schizophrenia and schizoaffective disorder are, at least in part, accounted for by effect of a lifetime antipsychotics use, whereas a relative gray matter volume preservation in bipolar patients could be secondary to a medication effect (i.e., chronic treatment with lithium) and that, even if a "primary" disease-associated loss of neocortical volume exists in these patients, it could be obscured by a volume-enhancing effect of chronic lithium use. Disentangling a "primary" disease and medication effects on brain structure presents considerable difficulty and requires carefully designed longitudinal research protocols.

Some, though not all studies, have captured associations between brain anatomy changes and clinical manifestations in schizophrenia and related disorders. In particular, deficits in the fronto-temporal volume have been linked to severity of positive symptoms in schizophrenia and schizoaffective disorder; reductions in various structures within the MTL have been shown to correlate with the severity of delusions. Additionally, volume reductions in the MTL have been linked to declarative memory dysfunction in first-degree relatives of schizophrenic probands. Subsequent data suggest that the various subgroups within bipolar disorder may have different structural brain characteristics, depending on whether they have psychosis. The majority of studies report similar brain structure changes in persons with schizophrenia and psychotic bipolar disorder, perhaps as one of the overlapping intermediate phenotypes indicative of psychosis liability.

Intermediate phenotypes represent smaller, homogenous, measurable units of illness more proximal to pathogenesis of illness and genetic etiologies than the global diagnosis constructs. Intermediate phenotypes may be conceptualized somewhat similar to symptom dimensions: While the dimensions of illness represent more discreet and homogenous clinical complexes, intermediate phenotypes represent smaller, homogenous components of illness with strong biologic ties. From this perspective, intermediate phenotypes may be a promising target in the search for genetic factors in psychosis, although even this has been difficult to demonstrate. They may also provide more direct clues for understanding of psychosis pathogenesis because they are thought to be more intimately linked to causal biologic pathways of the illness than the clinical syndromes. At first, the intermediate phenotype/biomarker-focused approach was used in schizophrenia research and resulted in recognition of several neurophysiological, neuroanatomical, and cognitive markers. More recently, these same candidate biomarkers have been used to explore different psychotic illnesses, specifically psychotic bipolar disorder, in order to develop valid, objectively defined biological definitions for the

psychosis dimension. Among the variety of candidate biomarkers, there are specific elements that appear to be promising as indicators of psychosis liability. Future studies focused on intermediate psychosis phenotypes will explore their genetic linkage, as well as improve overall conceptualization of clinical and biological boundaries of psychosis spectrum disorders.

PSYCHOSIS AS A CLINICAL CONTINUUM: FROM ATTENUATED PSYCHOSIS PHENOTYPES TO FRANK PSYCHOTIC DISORDERS

Psychosis manifestations include a broad array of symptom presentations of various severities which appear to lay on a continuum from mild subthreshold psychotic phenomena to frank and florid psychosis. Mild sporadic psychotic symptoms, such as illusions, mild hallucinations (e.g., undifferentiated noises or calling a person's name), ideas of reference, magical thinking, unusual beliefs and perceptual experiences, and the like are fairly common in otherwise healthy individuals who don't carry any diagnosis of psychiatric illness. Although these symptoms are psychotic by nature, they are below the threshold for a formal psychosis diagnosis. It appears that there may be no discrete demarcations between frank psychosis and minor subclinical psychotic symptoms, other than severity, breadth and persistence; rather, psychosis represents a clinical continuum with multiple subtypes and grades of symptom manifestations.

Large longitudinal studies evaluated the frequency of mild psychotic symptoms in the general population and found that subclinical psychotic experiences are not uncommon. For example, the large community sample study using the British National Psychiatric Morbidity Survey data showed that 4.4 percent of the general population reported incident psychotic symptoms over an 18-month period. The vast majority of individuals who reported such psychotic symptoms did not meet full diagnostic criteria for any psychiatric disorder. These observations may indicate the presence of latent continuum-like psychopathology in the general population. Moreover, they are consistent with a multifactorial etiological model of schizophrenia where many different genes of small effect interact with one another and with environmental factors to result in the illness. Depending on a complex interplay of genetic and environmental influences, the manifestations of psychosis can vary from mild episodic experiences (e.g., déjà vu, infrequent thought blocking, or visual illusions), to moderate psychotic-like symptoms (e.g., magical thinking, ideas of reference, bodily illusion, or out-of-body experiences), or to severe and persistent frank psychosis, fulfilling DSM criteria (e.g., delusions or hallucinations). From the psychosis continuum perspective, subthreshold psychotic symptoms are considered indicators of proneness to psychosis, with clinical psychosis emerging when a critical genetic threshold is reached, accompanied by modifying or triggering epigenetic and environmental factors (Fig. 12.16–4).

Clinically, the continuum nature of psychosis is well illustrated by premorbid and prodromal symptoms in individuals with schizophrenia and affective psychosis, psychopathology in biological relatives of psychotic probands, as well as by certain types of personality disorders, displaying a wide range of psychotic symptoms, varying from very mild psychotic-like phenomena to frank psychosis during decompensation phases.

Premorbid and Prodromal Signs and Symptoms in Schizophrenia and Affective Psychoses

The course of schizophrenia is generally conceptualized as starting with premorbid (pre-prodromal) signs and symptoms, and then a prodromal phase of illness, leading to typical clinical manifestations as the illness evolves. However, it is not always possible to clearly demarcate these periods due to presence of schizophrenia spectrum personality disorders and traits characteristic for a proportion of individuals who subsequently develop schizophrenia, as well as minor psychotic, affective, and behavioral symptoms that sometimes appear long before the debut of schizophrenia. In the typical, although not invariable, premorbid history of schizophrenia, patients have had schizoid or schizotypal personality features and are characterized as quiet, passive, introverted, emotionally aloof, and reclusive. Other characteristics, more typical for schizoid personality, include a low need for emotional bonds: In childhood these individuals have few friends, by the adolescent period they may show no interest in dating, and avoid school sports and competitions. Instead, they prefer solitary activities and seem content by themselves. When schizotypal features are predominant, thinking, perception, and behavior appear to be odd or peculiar. Various psychotic-like symptoms, such as magical thinking, peculiar beliefs, ideas of reference, perceptual illusions, odd fantasies, or derealization–depersonalization symptoms, are commonly present. Especially during the adolescent period, peers may exclude them from the group activities or tease them, as schizotypal individuals have poor communication skills and distorted social judgment. This often leads to increased anxiety, feelings of “not fitting in,” and further withdrawal from the social group. Although different types of personalities (including paranoid, avoidant, obsessive-compulsive, and, less commonly, borderline) are also seen in the premorbid period in individuals with schizophrenia, schizoid, and schizotypal personality traits are believed to be the most common. Still, the presence of these premorbid and prodromal symptoms does not inevitably result in conversion to schizophrenia or other psychotic disorder.

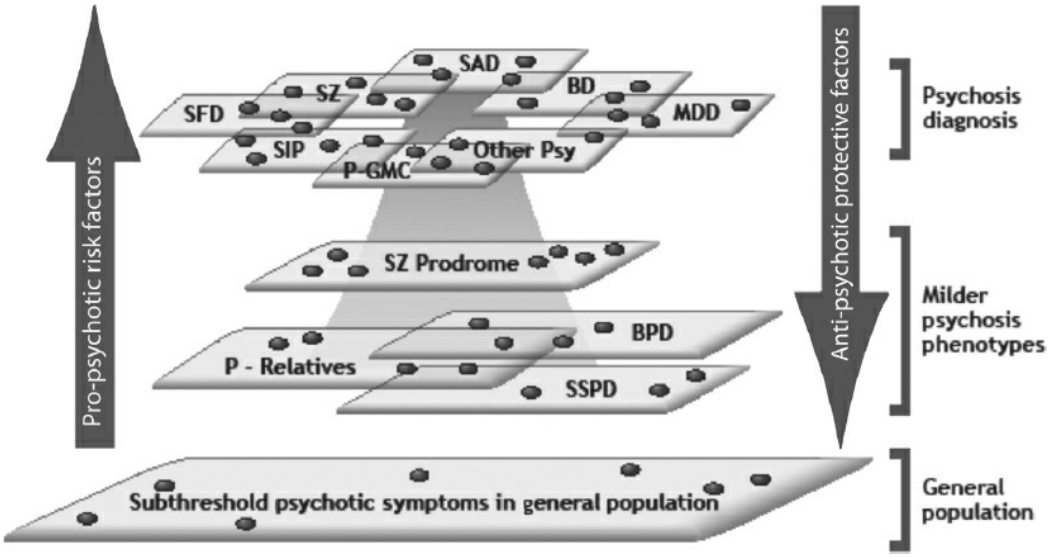

FIGURE 12.16–4. Threshold model of psychosis. In the general population, considerable evidence indicates the presence of fleeting psychotic symptoms without any functional impact. When more severe or persistent, psychotic symptoms appear at subthreshold of formal psychosis diagnosis and manifest in entities like schizophrenia prodrome and schizophrenia spectrum personality disorders, as well as are frequently seen in biological relatives of probands with psychotic illnesses. In a proportion of these individuals psychosis may reach the level of florid clinical manifestations severe enough to fulfill criteria for DSM diagnosis. Genetic, epigenetic, other biological and environmental factors can protect against or advance psychosis. Examples of pro-psychotic risk factors include obstetrical complications, drug abuse, specifically, adolescent cannabis use, and psychosocial stressors. Less is known about anti-psychotic resilience factors, though high premorbid functioning, structured environment and cognitive remediation administered early in the course of evolving psychosis appear to have protective effects. •, candidate psychosis genes; P-Relatives, psychotic relatives of probands with psychotic disorder; SSPD, schizophrenia spectrum personality disorders (cluster A: Schizoid, paranoid, schizotypal); BPD, borderline personality disorders; SZ, schizophrenia; SAD, schizoaffective disorder; BD, bipolar disorder with psychotic features; MDD, major depressive disorder with psychotic features; SFD, schizophreniform disorder; SIP, substance-induced psychosis; P-GMC, psychosis due to general medical conditions; Other Psy, other psychotic disorders.

Although the first hospitalization is often taken to be the beginning of the illness, in many cases premorbid and prodromal signs and symptoms may have been present for months or years prior. The majority of schizophrenia cases are characterized by insidious, gradual onset of psychosis with no clear date of its first manifestation. The prodromal symptoms may start with somatic complaints, such as headaches, digestive problems, autonomic symptoms, or fatigue and generalized weakness. Family and friends commonly notice that the person has changed, became withdrawn and isolated, and is no longer functioning well in occupational, personal, and social settings. Deterioration in personal hygiene and grooming, decreased appetite, and peculiar eating habits are also common. During this stage, a patient may begin to develop new, unusual interests in abstract ideas, philosophy, occult, or religious topics. Gradually, although acute onset may also occur, subthreshold psychotic symptoms, such as unusual speech, strange perceptual experiences and bodily illusions,

bizarre ideas, and markedly peculiar behavior, build up and subsequently progress to the level of frank psychosis in the form of hallucinations, delusions, typical thought and speech disorders, and/or grossly disorganized behavior.

The validity of the prodromal syndrome is uncertain, since symptoms are frequently recognized only after the diagnosis of schizophrenia has been made. During the prodromal phase, patients usually lack insight regarding their developing symptoms, although some verbalize feeling as if they have changed. Once schizophrenia is diagnosed, the retrospective remembrance and the patient's subjective report of early signs and symptoms may be affected, therefore obtaining the information about the early (prodromal) period of illness from the patient's family may be of significant help in building a diagnosis. The presence of prodromal symptoms in affective psychoses is unclear. Although clinical observations suggests that such symptoms as mood lability, impulsivity, destructibility, and physical hyperactivity are often present long before the development of affective psychosis, phenomenology of the bipolar prodrome is not well characterized. One of the recent studies conducted with early onset bipolar I disorder individuals showed that there may be unique prodromal characteristics in patients with and without subsequent bipolar phenotype that manifests as psychotic mania. In this study, attenuated psychotic symptoms emerging late in the prodrome and increased energy and goal-directed activity were more common in individuals with later psychotic mania. Studies comparing prodromal symptoms in schizophrenia and affective psychosis revealed considerable overlap; for example, such symptoms as suspiciousness, sub-threshold hallucinatory experiences, anxiety, insomnia, impaired school, or work functioning were frequently present during the prodromal period in both psychoses. On the other hand, depressed mood, suicidality, mood lability, difficulty thinking/communicating clearly, tiredness/lack of energy, obsessions/compulsions, and physical agitation were more specific for the bipolar than schizophrenia prodrome.

In research, one of the approaches that is being utilized to characterize clinical profiles of the prodromal state of psychotic illness and to identify predictive factors for conversion to fully manifested psychosis is observing large samples of individuals who are at increased risk for developing schizophrenia and related psychotic disorders (e.g., offsprings of schizophrenic parents and individuals who have several schizophrenic family members). Recent such studies focusing on prodrome and other "at risk samples" have shown that premorbid and prodromal symptoms, clinically well accepted as a phase in a course of schizophrenia, are not necessarily pathognomonic of schizophrenia psychosis. For example, in a study by the Melbourne group, high-risk individuals with either a family history of psychotic disorder, schizotypal personality disorder,

subthreshold psychotic symptoms, or brief transient psychotic symptoms were followed for a 12-month period. In that group 40.8 percent of the high-risk individuals subsequently developed psychosis; however, diagnoses ranged between schizophrenia, schizoaffective disorder, bipolar disorder, and major depression with psychosis. Other large-scale studies in this field suggest that approximately one-third of "at risk" individuals eventually convert to psychosis, one-third will continue to experience stable chronic mild psychotic phenomena without full conversion to DSM-level psychotic illness, and one-third will improve overtime, some to the level of full recovery. These results indicate that although early signs and symptoms, including prodromal psychotic-like phenomena, can be predictive of conversion to a spectrum of psychotic disorder, they do not indicate the exact type of psychosis, and the long-term outcomes are highly variable ranging from fully manifested psychotic disorder to full recovery. Given these research outcomes, in DSM-5, a new diagnostic category, "Attenuated psychosis syndrome," has been introduced. It refers to a broadly defined syndrome characterized by psychotic-like phenomena that are below a threshold for full psychosis, for example, the symptoms that are less severe and more transient, and frequently accompanied by a relatively preserved insight. This diagnostic category is appropriate for some (but not all) individuals with prodromal or subthreshold symptoms and calls for a clinician's judgment on where to draw the line for mild psychotic-like presentations that do not deserve DSM diagnosis and where it is appropriate to introduce this diagnosis.

Psychosis Manifestations as a Component of Personality Disorders

The occurrence of brief psychotic symptoms is common in individuals with personality disorders. Typically, psychotic experiences in these people are triggered by psychological (e.g., change in employment, loss of a relationship) or physiological (e.g., medical illness) stressors, or substance use, when compensatory mechanisms weaken. Certain types of personality disorders are especially predisposed to psychosis, including cluster A and borderline personality disorders. In DSM-5, cluster A personality disorders include paranoid, schizoid, and schizotypal types. These personality disorders are also called *schizophrenia spectrum personality disorders* because their core features resemble symptoms of schizophrenia, albeit in a milder manifestation. In addition, evidence demonstrates the familial aggregation of these personality disorders in biological relatives of probands with schizophrenia and other psychotic disorders.

Schizotypal Personality Disorder. Schizotypal personality disorder is now described under DSM-5 category of "Schizophrenia Spectrum and Other Psychotic Disorders" emphasizing long-acknowledged biological link between this particular Cluster A personality disorder and schizophrenia.

Schizotypal personality disorder is characterized by cognitive and perceptual distortions, eccentric behavior, as well as a persistent pattern of social and interpersonal deficits. Mild psychotic symptoms, such as magical thinking, perceptual alterations, bizarre rituals, ideas of reference, paranoid ideation, derealization/depersonalization, idiosyncratic speech, odd or eccentric behavior, are core features of schizotypal personality disorder. In response to stress, schizotypal individuals may decompensate and display frank psychotic symptoms (e.g., hallucinations, delusions, disorganization of thought process, and behavior); however, these episodes are generally brief and fragmentary. Usually, these psychotic symptoms are transitory and insufficient in duration and severity to draw an additional diagnosis of a psychotic disorder. Nevertheless, in some cases clinically significant psychotic symptoms may develop that ultimately meet criteria for brief psychotic disorder, schizophreniform disorder, delusional disorder, or even schizophrenia. In these cases schizotypal personality disorder may be indicated in the diagnosis as a premorbid personality disorder, in parallel with a diagnosis of psychotic illness.

Sufficient research evidence has shown that among cluster A personality disorders schizotypy is most intimately related to schizophrenia and other psychoses. There is a higher density of schizotypal personality disorder cases in the biological relatives of schizophrenia probands than in controls. Although schizotypal personality disorder occurs in 0.3 to 3 percent of the general population, studies have reported rates of 0.8 to 4.7 percent in the biological relatives of psychotic patients. Recent linkage studies from genome-wide scans in schizophrenia and schizotypal personality disorder, conducted with the large sample of high-density schizophrenia families, have suggested that at least a subset of candidate schizophrenia susceptibility genes is also involved in etiology of schizotypy in nonpsychotic relatives of schizophrenic probands.

In the large epidemiologic study called the Roscommon Family Study, which collected 534 probands with psychotic and affective illnesses and 2,043 of their biological relatives, schizotypal personality disorder showed a strong familial relationship not only with schizophrenia, but also with the other major psychotic disorders. Further factor analysis from this research sample showed that schizotypy is a complex, multidimensional clinical construct, whose various dimensions differ widely in the specificity with which they reflect the familial liability to psychosis. In this study seven schizotypal dimensions were described, among which sub-psychotic thought disorder, negative schizotypal signs, such as poor rapport and odd behavior, deficient occupational functioning, and social isolation/avoidance best disambiguated relatives of schizophrenic probands from relatives of matched controls. Different studies reported that the disorganization dimension of schizotypy was common to relatives of both schizophrenia and psychotic bipolar disorders, but not in the

relatives of nonpsychotic bipolar probands. Interestingly, schizotypal disorganization symptoms were suggested as a unique phenotype associated with an increased familial risk for both schizophrenia and affective psychosis. From this perspective, schizotypy represents a subthreshold psychosis phenotype that cuts across the categorical diagnosis of schizophrenia and bipolar disorder.

Paranoid Personality Disorder. The essential feature of paranoid personality disorder is a pattern of long-standing suspiciousness and mistrust of people in general such that their motives are interpreted as malevolent. Subthreshold psychotic symptoms are not uncommon in persons with paranoid personality disorder, of which ideas of reference and associated verbal and visual illusions are seen most often. Under stressful conditions, specifically in a situation naturally provoking paranoid feelings (e.g., when these patients are being criticized or disagreed with, relocating, facing new professional environment), they may decompensate and experience brief psychotic episodes, lasting minutes to hours. Among frank psychotic symptoms, paranoid and persecutory delusions are the most common, although brief visual and auditory hallucinations may also occur.

There is evidence of a genetic association between paranoid personality disorder and psychosis spectrum illnesses. For example, in some instances, paranoid personality disorder may appear as the premorbid antecedent of schizophrenia or delusional disorder. An increased incidence of major depressive disorder with psychotic features has been also reported in patients with paranoid personality disorder. There have been few controlled studies of familial risk of paranoid personality disorder based on direct interviews of relatives, blind to proband's diagnosis. However, there is substantial evidence for a familial relationship between paranoid personality disorder and psychotic disorders with reports of higher prevalence of paranoid personality disorder among first-degree relatives of schizophrenic, schizoaffective and psychotic bipolar probands ranging from 1.7 to 8.6 percent in different studies compared to 0.5 to 2.5 percent in the general population. In contrast to schizotypal personality disorder, paranoid personality disorder is less common in relatives of probands with affective psychosis, although one study reported increased familial risk of paranoid personality disorder for unipolar depression (2.9 percent), which was higher than for relatives of schizophrenia patients (1.7 percent).

Schizoid Personality Disorder. Individuals with schizoid personality disorder are characterized by a pervasive pattern of detachment from social relationships and a restricted range of emotional expressions. This behavioral pattern begins in childhood and is usually clinically obvious by early adulthood, appearing in a variety of contexts. Although subclinical psychotic symptoms are not uncommon in persons with different Cluster A

personality disorders, individuals with schizoid personality disorder do not show many unusual perceptual experiences, distorted thinking processes, or bizarre behavioral patterns. Although, they appear self-absorbed, detached, and emotionally aloof, they have a normal capacity to recognize reality. On rare occasions, particularly in response to social or interpersonal stressful circumstances, individuals with this disorder may experience very brief psychotic episodes (lasting minutes to hours). The content of psychotic symptoms may vary but is usually limited to delusions of reference, paranoid or persecutory delusions, and brief visual and auditory hallucinations relevant to the social or interpersonal stressors.

Although a genetic association between schizoid personality disorder and functional psychoses has been suggested, the data to support this are contradictory. Some studies (e.g., the Roscommon Family Study, Bipolar-Schizophrenia Network on Intermediate Phenotypes) have reported that schizoid personality disorder has significantly increased prevalence in the relatives of individuals with schizophrenia (1.2 to 3.4 percent), although this association is modest compared to schizotypal personality disorder. In contrast, several studies showed that the rate of schizoid personality disorder among biological relatives of schizophrenic probands is not any higher than in the general population (reports vary between 0.3 to 0.7 percent). Similarly, reports of schizoid personality disorder co-segregated in families with affective psychoses are uncommon. One of the difficulties is that people with isolated schizoid personality disorder are not commonly seen in clinical settings, and the prevalence of this personality disorder in the general population is not clearly established. Schizoid personality disorder may appear as the premorbid precursor of schizophrenia; however, the proportion of patients who eventually convert to schizophrenia is unknown. Schizoid personality disorder most often co-occurs with schizotypal and paranoid personality disorders, as well as with avoidant personality disorders, which has been also reported to be associated with schizophrenia.

Borderline Personality Disorder. The crucial feature of borderline personality disorder is a pervasive pattern of instability of self-image, interpersonal relationships, and affective reactions, as well as marked impulsivity and self-damaging behavior. Borderline personality disorder is clinically perceived as a condition “on the border” of psychotic functioning; therefore, historically, in psychoanalytic literature it has been called “ambulatory schizophrenia” (a term introduced by Helene Deutsch) and “psychotic character disorder” (described by John Frosch). Individuals with borderline personality disorder are prone to transient psychotic episodes that occur most frequently during periods of stress, specifically in response to a real or imagined abandonment. In fact, psychotic-like symptoms are included as diagnostic criteria for borderline personality disorder (criterion

9: Transient, stress-related paranoid ideation or severe dissociative symptoms). Other broadly defined psychotic symptoms are also often reported, including ideas of reference, hypnagogic phenomena, and body-image distortions. More severe psychotic symptoms, including frank hallucinations and delusions, occur less frequently. Typically, the symptoms tend to be transient, lasting minutes to hours. However, one study by a Cornell group reported that 27 percent of the patients with borderline personality disorder had psychotic episodes, lasting for weeks, which were not attributed to comorbid affective or substance abuse disorders.

It is noteworthy that in some cases psychotic symptoms in persons with borderline personality disorder may be of a factitious nature (e.g., Canadian studies report up to 13 percent of factitious psychotic symptoms in inpatients with borderline personality disorder), thus a thorough psychiatric history, observation of patient's behavior, and objective information obtained from the family members are crucial. In some cases, the psychotic symptoms in borderline personality disorder respond well to external structuring and situational changes (e.g., the real or perceived return of subjectively valuable person or caregiver may result in remission of the transient psychosis). However, in more severe cases antipsychotic medications are warranted.

Borderline personality disorder shows a strong familial pattern. For example, the risk to acquire borderline personality disorder for first-degree biological relatives of those with the disorder is about five times greater than in the general population. There is also an increased familial association of borderline personality disorder with mood disorders, antisocial personality disorder, substance-related disorders, and eating disorders. Although considerable co-occurrence of borderline personality disorder with schizotypal personality disorder has been reported, as well as occurrence of premorbid borderline personality features in persons with schizophrenia, the majority of studies do not support a familial link between borderline personality disorder and schizophrenia. The large-scale epidemiologic studies of schizophrenia-related personality disorders in relatives (e.g., the Roscommon Family Study) reported that borderline personality disorder was rare among first-degree relatives of schizophrenic probands, but showed a modest clustering of cases in relatives of patients with psychotic and nonpsychotic affective disorders. Recent findings from a large multisite consortium in biomarkers of psychosis, Bipolar-Schizophrenia Network on Intermediate Phenotypes, demonstrated that 1.3 percent of first-degree relatives of schizophrenia patients met rigorous DSM-IV-TR criteria for borderline personality disorder, whereas 4.1 and 3.2 percent of relatives of patients with schizoaffective disorder and psychotic bipolar I disorder, respectably, met such criteria. Family studies of borderline personality disorder have also indicated increased rates of

mood disorders in the relatives of borderline probands compared with healthy controls. Recent studies have suggested that borderline personality disorder and bipolar disorder may exist on a spectrum and, perhaps, share partially overlapping pathogenesis and may be rooted in the same genetic etiology, similar to cluster A personality disorders and schizophrenia. However, the question of whether a borderline personality disorder is related specifically to psychotic bipolar disorder and to psychosis in general remains open. Pilot studies suggest that persons who have borderline personality disorder with a history of psychotic-like symptoms show distinct eye tracking abnormalities, similar to those in patients with schizophrenia. This may suggest that borderline personality disorder with psychotic symptoms may fall into the psychosis spectrum disorders continuum and share some endophenotypic and genetic markers, underlying psychosis. Psychotic-like symptoms and narrowly defined psychosis appear to be a promising target dimension for future research in borderline personality disorder and mood disorders.

Clinical and research evidence suggest that psychosis is best conceived of as a clinical continuum with multiple subtypes and grades of symptoms, similar to hypertension or congestive heart failure. In fact, there appear to be no clear demarcation between frank, clinically significant, psychosis and mild psychotic-like symptoms, other than severity and duration. Large epidemiologic studies suggest that fleeting psychotic symptoms are common in the general population. The majority of individuals who have these psychotic experiences show no functional impairment and are considered to be free of any formal psychiatric diagnosis. These observations suggest that psychosis can manifest in mild forms long before its manifestation to a clinically significant level. Moreover, many of the individuals with such subthreshold symptoms will remain stable for years, without further progression into a full psychosis phenotype. Only a modest proportion of them will convert to a psychotic illness in the future. These “at risk” individuals, while rarely evaluated in psychiatrists’ offices, may be a valuable source for studying proneness to psychosis, as well as for exploring pro-psychotic risk factors and individual resilience against psychosis.

In more severe and persistent presentations, psychotic symptoms appear at the border of formal psychosis diagnosis and manifest in entities like schizophrenia prodrome and cluster A personality disorders, frequently seen in biological relatives of probands with psychotic illnesses. Although cluster A personality disorders (in particular, schizotypy) are generally thought to be biologically related to schizophrenia, recent family and whole genome linkage studies have shown that schizotypal personality disorder may present a unique phenotype associated with an increased familial risk for psychosis in general, independent of the categorical diagnosis. Similarly, prodromal symptoms, which are classically

conceptualized as a phase in a course of schizophrenia, appear to be not unique to schizophrenia psychosis. Instead, they signify risk for a spectrum of functional psychoses but not the exact diagnosis that will develop.

From the disease continuum perspective, the final diagnosis of a psychotic disorder is the end point of a longitudinal pathway from asymptomatic, to subthreshold psychotic experiences, to frank but infrequent psychotic symptoms, to development of persistent psychosis. It is possible that this pathway is based on a dynamic matrix of genetic, gene expression and epigenetic influences which are in a complex interplay with environmental factors, with frank psychosis emerging when a critical genetic effects threshold is reached, accompanied by modifying or triggering epigenetic, individual biological and/or environmental factors. Although subthreshold manifestations of psychosis are considered indicators of proneness to schizophrenia and other functional psychoses, they do not signify an inevitable conversion to psychosis. Whether or not an individual at risk will traverse this pathway may depend on interaction between pro-psychotic factors (e.g., individual biological vulnerability, obstetrical complications, maternal/in-utero immune stimulation, adolescent cannabis use, childhood trauma) and protective resilience factors (e.g., high premorbid functioning, high IQ, stable social environment).

ACKNOWLEDGMENTS

We would like to thank Four Seasons Decorations, Dallas, TX, for their help with graphic design and illustrations for this chapter.

REFERENCES

*Allardyce J, Gaebel W, Zielasek J, van Os J. Deconstructing psychosis conference February 2006: the validity of schizophrenia and alternative approaches to the classification of psychosis. *Schizophr Bull*. 2007;33:863.

Arajärvi R, Ukkola J, Haukka J, Suvisaari J, Hintikka J. Psychosis among “healthy” siblings of schizophrenia patients. *BMC Psychiatry*. 2006;6:6.

Balaratnasingam S, Janca A. Normal personality, personality disorder and psychosis: current views and future perspectives. *Curr Opin Psychiatry*. 2015;28(1):30–34.

Bramon E, McDonald C, Croft RJ, Landau S, Filbey F. Is the P300 wave an endophenotype for schizophrenia? A meta-analysis and a family study. *Neuroimage*. 2005;27:960.

*Dikeos DG, Wickham H, McDonald C, Walshe M, Sigmundsson T. Distribution of symptom dimensions across Kraepelinian divisions. *Br J Psychiatry*. 2006;189:346.

Fanous AH, Neale MC, Gardner CO, Webb BT, Straub RE. Significant correlation in linkage signals from genome-wide scans of schizophrenia and schizotypy. *Mol Psychiatry*. 2007;12:958.

Hamshere ML, Bennett P, Williams N, Segurado R, Cardno A. Genomewide linkage scan in schizoaffective disorder: significant evidence for linkage at 1q42 close to DISC1, and suggestive

evidence at 22q11 and 19p13. *Arch Gen Psychiatry*. 2005;62:1081.

Hong LE, Tagamets M, Avila M, Wonodi I, Holcomb H. Specific motion processing pathway deficit during eye tracking in schizophrenia: a performance-matched functional magnetic resonance imaging study. *Biol Psychiatry*. 2005;57:726.

Ivleva EI, Bidesi AS, Keshavan MS, et al. Using Brain Gray Matter Volume as an Intermediate Phenotype of Psychosis: Bipolar and Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *Am J Psychiatry*. 2013;170(11):1285–1296.

Kato T. Whole genome/exome sequencing in mood and psychotic disorders. *Psychiatry Clin Neurosci*. 2015;69(2):65–76.

*Kendler KS, McGuire M, Gruenberg AM, O'Hare A, Spellman M. The Roscommon Family Study: III. Schizophrenia-related personality disorders in relatives. *Arch Gen Psychiatry*. 1993;50:781.

Kwapil TR, Barrantes-Vidal N. Schizotypy: looking back and moving forward. *Schizophr Bull*. 2014;41(Suppl 2):S366–S373.

Links PS, Steiner M, Mitton J. Characteristics of psychosis in borderline personality disorder. *Psychopathology*. 1989;22:188.

Narayanan B, Ethridge LE, O'Neil K, et al. Genetic sources of subcomponents of event-related potential in the dimension of psychosis analyzed from the B-SNIP study. *Am J Psychiatry*. 2015;172(5):466–478.

Olincy A, Martin L. Diminished suppression of the P50 auditory evoked potential in bipolar disorder subjects with a history of psychosis. *Am J Psychiatry*. 2005;162:43.

Owen MJ, Craddock N, Jablensky A. The genetic deconstruction of psychosis. *Schizophr Bull*. 2007;33:905.

Park N, Juo S, Cheng R, Liu J, Loth J. Linkage analysis of psychosis in bipolar pedigrees suggests novel putative loci for bipolar disorder and shared susceptibility with schizophrenia. *Mol Psychiatry*. 2004;9:1091.

Pearlson G, Folley B. Endophenotypes, dimensions, risks: is psychosis analogous to common inherited medical illnesses? *Clin EEG Neurosci*. 2008;39(2):73.

Peralta V, Cuesta MJ. The relationship between syndromes of the psychotic illness and familial liability to schizophrenia and major mood disorders. *Schizophr Res*. 2007;91:200.

*Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511(7510):421–427.

Tamminga CA, Ivleva EI, Keshavan MS, et al. Clinical phenotypes of psychosis in the Bipolar and Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *Am J Psychiatry*. 2013;170(11):1263–1274.

*Tamminga CA, Ivleva EI. Toward a dimensional understanding of psychosis and it's treatment. Treatment of psychiatric disorders. In: Gabbard GO, ed. *Gabbard's Treatments of Psychiatric Disorders*. 5th ed. 2014.

Tamminga CA, Pearlson G, Keshavan M, Sweeney J, Clementz B, Thaker G. Bipolar and schizophrenia network for intermediate phenotypes: outcomes across the psychosis continuum. *Schizophr Bull*. 2014;40(Suppl 2):S131–S137.

Thaker G. Neurophysiological endophenotypes across bipolar and schizophrenia psychosis. *Schizophr Bull*. 2008;34(4):760–773.

Valles V, Van Os J, Guillamat R, Gutiérrez B, Campillo M. Increased morbid risk for schizophrenia in families of in-patients with bipolar illness. *Schizophr Res*. 2000;42:83.

van Os J, Linscott RJ, Myin-Germeys I, Delespaul P, Krabbendam L. A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. *Psychol Med*. 2008;8:1–17.

Webb CT, Levinson DF. Schizotypal and paranoid personality disorder in the relatives of patients with schizophrenia and affective disorders: A review. *Schizophr Res*. 1993;11:81.

Wilhelmsen KC. The feasibility of genetic dissection of endophenotypes. *Psychophysiology*. 2014;51(12):1337–1338.

Wiles NJ, Zammit S, Bebbington P, Singleton N, Meltzer H. Self-reported psychotic symptoms in the general population: Results from the longitudinal study of the British National Psychiatric Morbidity Survey. *Br J Psychiatry*. 2006;188:519.

Williams NM, Green EK, Macgregor S, Dwyer S, Norton N. Variation at the DAOA/G30 locus influences susceptibility to major mood episodes but not psychosis in schizophrenia and bipolar disorder. *Arch Gen Psychiatry*. 2006;63:366.

▲ 12.17 Other Psychotic Disorders

RAMIN MOJTABAI, M.D., PH.D., LAURA J. FOCHTMANN, M.D., AND EVELYN J. BROMET, PH.D.

This section describes disorders included in the section on other psychotic disorders in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). It also describes the tenth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) category of acute and transient psychoses. The disorders discussed in this section occur less frequently and are more poorly understood than schizophrenia or mood disorders with psychotic features. Nevertheless, they can have profound psychosocial consequences and are important to identify and treat.

ACUTE AND TRANSIENT PSYCHOTIC DISORDERS AND BRIEF PSYCHOTIC DISORDER

History

Descriptions of psychotic disorders with an acute onset and transient course have a long history in psychiatry. Along with the classic syndromes of schizophrenia and affective psychoses, Emil Kraepelin (1856 to 1926) mentioned amentia, a psychotic disorder with a remitting course and favorable outcome, originally described by Theodor Meynert (1833 to 1898). Similar psychotic syndromes were described by other 19th- and 20th-century European psychiatrists under various terms such as "cycloid psychosis," "bouffée délirante," "reactive psychosis," and others, as briefly described in Table 12.17–1. At the same time, investigators reported a high prevalence of acute benign psychotic episodes in Africa, the Caribbean islands, and India.

The descriptions of these psychotic syndromes are rather similar, a remarkable observation since these descriptions grew out of different psychiatric traditions separated by cultural and linguistic boundaries. Although all portray a psychotic condition with acute onset and remitting course, the multiplicity of descriptions and labels has been a major barrier to better understanding and classifying these psychoses. In recent years, attempts have been made to surpass this barrier by defining a syndrome based on the common features of the traditional descriptions. In this context, the introduction of the acute and transient psychotic disorders in ICD-10 was an important step. Many of these cases would be classified under the category of brief psychotic disorder in DSM-5. Some would be categorized under schizophreniform disorder and other specified schizophrenia spectrum and other psychotic disorders in the DSM-5.

The historical roots of the DSM-5 diagnosis of brief psychotic disorder,

similar to ICD-10 acute and transient psychotic disorders, can be found in European psychiatry. The Scandinavian concept of reactive or psychogenic psychosis was especially influential in the formulation of the DSM-III brief reactive psychosis diagnosis (the forerunner of the DSM-5 brief psychotic disorder) as a psychotic condition lasting less than 2 weeks that followed a significant psychosocial stressor and that involved emotional turmoil and one of the symptoms of loosening of associations, delusions, hallucinations, or disorganized or catatonic behavior. This diagnosis was maintained in the DSM-III-R, but the allowable duration of symptoms was extended to 1 month. In DSM-IV and DSM-IV-TR, the diagnosis of brief reactive psychosis was replaced by brief psychotic disorder, eliminating an identifiable psychosocial stressor as a diagnostic criterion. The change was partly motivated by the observation that many cases of brief psychosis are not precipitated by a marked stressor and hence cannot be labeled as "reactive." Furthermore, the DSM-III and DSM-III-R criterion of emotional turmoil was removed from the DSM-IV and DSM-IV-TR criteria for brief psychotic disorder. The aim of these revisions in later editions of the DSM was to broaden the scope of the diagnosis of brief psychotic disorder. Nonetheless, DSM-IV and DSM-IV-TR distinguished between brief psychotic disorder with and without marked stressors and indicated that the former group is equivalent to the brief reactive psychosis. There were few changes in the definition and characterization of brief psychotic disorder between the DSM-IV-TR and the DSM-5. The DSM-5 requires that one of the criteria be the presence of delusions, hallucinations, or disorganized speech. The DSM-5 also includes a "with catatonia" specifier and ratings of current severity of symptoms on a scale ranging from not present to present and severe.

Comparative Nosology

Acute and transient psychotic disorders are defined by ICD-10 as psychotic conditions with onset within 2 weeks and full remission within 1 to 3 months. A more detailed description along with criteria for specific conditions under this general rubric are presented in Table 12.17–2.

Brief psychotic disorder is defined by DSM-5 as a psychotic condition that involves the sudden onset of psychotic symptoms and lasts 1 day or more but less than 1 month. Remission is full and patients return to their premorbid level of functioning. Thus, most individuals diagnosed with brief psychotic disorder under DSM-5 would be classified as having acute and transient psychotic disorders under ICD-10. DSM-5 criteria for brief psychotic disorder are presented in Table 12.17–3.

Table 12.17–1.
Partial Listing of the Classic Descriptions of Acute and Transient Psychotic

Disorders

Diagnosis	Publications	Description
Amentia	Meynert T. Amentia, die Verwirrtheit. *Jahrb Psychiatr Neurol*. 1889;9:1.	A psychotic illness with acute onset characterized by confusion and perplexity, agitation, rapidly changing vivid hallucinations and delusions, misidentification phenomena, anxiety and apprehension, first introduced by Theodore Meynert (1833–1898). An association with physical illness and exhaustion was noted in some patients. Full recovery occurs in a few weeks or months. Although somewhat influential at its time, amentia never produced much interest in later years.
Cycloid psychosis	Kleist K. über zykloide, paranoide und epileptoide Psychosen und über die Frage der Degenerations-psychosen. *Schweizer Arch Neurol Psychiat*. 1928;23:1. Leonhard K. Cycloid psychoses-endogenous psychoses which are neither schizophrenic nor manic depressive. *J Ment Sci*. 1961;107:632. Perris C, Brockington IF. Cycloid psychoses and their relation to the major psychoses. In: Perris C, Struwe G, Jansson B, eds. *Biological Psychiatry.* Amsterdam: Elsevier; 1981:447.	A psychotic disorder with acute onset and good prognosis but frequent recurrences, characterized by confusion, mood incongruent delusions, hallucinations, overwhelming anxiety, deep feelings of happiness or ecstasy, motility disturbances of akinetic or hyperkinetic type, a particular concern with death, mood swings, rapid change in symptoms within an episode. Two variants of the disorder were first described by Karl Kleist (1879–1960): Confusional insanity characterized by contrasting phases of confused excitement and stupor and motility psychosis characterized by contrasting phases of hyperkinesis and akineses. A third variant, anxiety-elation psychosis was introduced by Karl Leonhard (1904–1988). The diagnosis of cycloid psychosis is still used by German, Scandinavian, and other European psychiatrists and was influential for the formulation of acute and transient psychotic disorders in ICD-10. Carlo Perris and Ian Brockington introduced a set of explicit criteria for the diagnosis of cycloid psychosis in 1981.
Bouffée délirante	Magnan V, Legrain M. *Les Dégénérés (Etat Mental et Syndromes Episodiques).* Paris: Rueff et Cie; 1895. Pull CB, Pull MC, Pichot P. Nosological position of schizoaffective psychoses in France. *Psychiatr Clin*. 1983;16:141.	A psychotic disorder with acute onset in a person without previous psychiatric history. The episode remits completely with no residual symptoms. The episodes are characterized by delusions, hallucinations, depersonalization and derealization, confusion, mood change, and changing symptoms during the course of episode. Episodes are not due to organic or substance use. The syndrome was introduced by Valentin Magnan (1835–1916) and Paul-Maurice Legrain (1860–1939) in 1895 and was said to be caused by degeneration, a 19th-century view that attributed mental illness to the spread of

		modern civilization and urban life. Although this etiological formulation is no longer accepted, the diagnosis is still used by French speaking clinicians in Europe, West Africa, and the Caribbean. The concept of bouffée délirante was influential in formulating the ICD-10 acute and transient psychotic disorders. Due to its common occurrence in Africa and the Caribbean, it is also categorized as a culture-bound syndrome in the DSM-IV-TR. Charles Pull and colleagues proposed a set of explicit diagnostic criteria for bouffée délirante in 1983.
Psychogenic or reactive psychosis	Wimmer A. *Psykogene Sindssygdomsformer*. Copenhagen: St. Hans Hospitals Jubilæumsskrift; 1916.	A psychotic disorder with acute onset following external stress. Compared to schizophrenia, the onset is more likely to be acute and to be later in life. Premorbid adjustment tends to be better than in schizophrenia. There are more affective and confusional symptoms and less bizarre symptoms, and there is less family history of schizophrenia. Prognosis is better than schizophrenia. Diagnoses of psychogenic or reactive psychoses were popular among Scandinavian psychiatrists in the early part of the 20th century and continued to be used throughout the past century but found limited use outside of these countries.
Schizophreniform psychosis or disorder	Langfeldt G. *The Schizophreniform States*. Copenhagen: Ejnar Munksgaard; 1939. American Psychiatric Association. *Diagnostic and Statistical Manual*. 3rd ed. Washington, DC: American Psychiatric Press; 1980.	The Norwegian psychiatrist Gabriel Langfeldt (1937–1966) introduced the concept of schizophreniform psychosis as a condition with a good outcome that had a sudden onset following an identifiable precipitating factor in an individual with well-adjusted premorbid personality. Patients often present with disturbance of mood and clouding of consciousness. The term, but not the concept, was adopted by the DSM-III as a nonaffective psychotic syndrome with schizophrenic symptoms, which is distinguished from schizophrenia by a duration of less than 6 months. The concept of schizophreniform disorder in the latter sense was perpetuated in the later editions of the DSM.
Oneirophrenia	Meduna LJ. *Oneirophrenia: The Confusional State*. Urbana: University of Illinois Press; 1950.	Ladislas von Meduna (1896–1964) described oneirophrenia in 1939 as a syndrome characterized by acute onset of confusion, nightmare, or dream-like quality of all perceptions (hence the term “oneirophrenia”), extreme fear and anxiety, delusions, and visual hallucinations. The prognosis is generally good with full recovery. Meduna proposed an endocrinological explanation for the syndrome.
Hysterical psychosis	Hollander MH, Hirsch SJ. Hysterical psychosis. *Am J Psychiatry*. 1964;120:1066.	Although the term hysterical psychoses had been in use since the early 20th century, Marc Hollander and Steven Hirsch presented the first formal description of the syndrome in 1964. These authors described a psychotic episode with sudden and dramatic onset related to a profoundly upsetting event in the context of a “hysterical” personality. Symptoms include hallucinations, delusions, depersonalization,

and disorganized behavior. The episode seldom lasts longer than 1–3 weeks.

Table 12.17–2.
ICD-10 Diagnostic Criteria for Research for Acute and Transient Psychotic Disorders

F23 Acute and transient psychotic disorders

G1 There is acute onset of delusions, hallucinations, incomprehensible or incoherent speech, or any combination of these. The time interval between the first appearance of any psychotic symptoms and the presentation of the fully developed disorder should not exceed 2 weeks.

G2 If transient states of perplexity, misidentification, or impairment of attention and concentration are present, they do not fulfill the criteria for organically caused clouding of consciousness as specified for F05.-, criterion A.

G3 The disorder does not meet the symptomatic criteria for manic episode (F30.-), depressive episode (F32.-), or recurrent depressive disorder (F33.-).

G4 There is insufficient evidence of recent psychoactive substance use to fulfill the criteria for intoxication (F1x.0), harmful use (F1x.1), dependence (F1x.2), or withdrawal states (F1x.3 and F1x.4). The continued moderate and largely unchanged use of alcohol or drugs in amounts or with the frequency to which the individual is accustomed does not necessarily rule out the use of F23; that must be decided by clinical judgment and the requirements of the research project in question.

G5 Most commonly used exclusion clause. There must be no organic mental disorder (F00–F09) or serious metabolic disturbances affecting the central nervous system (not including childbirth). A fifth character should be used to specify whether the acute onset of the disorder is associated with acute stress (occurring 2 weeks or less before evidence of first psychotic symptoms): F23.x0 Without associated acute stress, F23.x1 With associated acute stress. For research purposes, it is recommended that change of the disorder from a nonpsychotic to a clearly psychotic state is further specified as either abrupt (onset within 48 hours) or acute (onset in more than 48 hours but less than 2 weeks).

F23.0 Acute polymorphic psychotic disorder without symptoms of schizophrenia

A. The general criteria for acute and transient psychotic disorders (F23) must be met.
B. Symptoms change rapidly in both type and intensity from day to day or within the same day.
C. Any type of either hallucinations or delusions occurs, for at least several hours, at any time from the onset of the disorder.
D. Symptoms from at least two of the following categories occur at the same time:
    1. Emotional turmoil, characterized by intense feelings of happiness or ecstasy, or overwhelming anxiety or marked irritability;
    2. Perplexity, or misidentification of people or places;
    3. Increased or decreased motility, to a marked degree.
E. If any of the symptoms listed for schizophrenia (F20.0–F20.3), criteria G(1) and (2), are present, they are present only for a minority of the time from the onset (i.e., criterion B of F23.1 is not fulfilled).
F. The total duration of the disorder does not exceed 3 months.

F23.1 Acute polymorphic psychotic disorder with symptoms of schizophrenia

A. Criteria A, B, C, and D of acute polymorphic psychotic disorder (F23.0) must be met.
B. Some of the symptoms specified for schizophrenia (F20.0–F20.3) must have been present for the majority of the time since the onset of the disorder, but not necessarily meeting these criteria completely, i.e., at least any one of the symptoms in F20, G1.1a to G1.2g.
C. The symptoms of schizophrenia in criterion B above do not persist for more than 1 month.

F23.2 Acute schizophrenia-like psychotic disorder

A. The general criteria for acute and transient psychotic disorders (F23) must be met.
B. The criteria for schizophrenia (F20.0–F20.3) are met, with the exception of the criterion for duration.
C. The disorder does not meet criteria B, C, and D for acute polymorphic psychotic disorder (F23.0).
D. The total duration of the disorder does not exceed 1 month.

F23.3 Other acute predominantly delusional psychotic disorders

A. The general criteria for acute and transient psychotic disorders (F23) must be met.
B. Relatively stable delusions or hallucinations are present but do not fulfill the symptomatic criteria for schizophrenia (F20.0–F20.3).
C. The disorder does not meet the criteria for acute polymorphic psychotic disorder (F23.0).
D. The total duration of the disorder does not exceed 3 months.

F23.8 Other acute and transient psychotic disorders

Any other acute psychotic disorders that are not classifiable under any other category in F23 (e.g., acute psychotic states in which definite delusions or hallucinations occur but persist for only small proportions of the time) should be coded here. States of undifferentiated excitement should also be coded here if more detailed information about the patient's mental state is not available, provided that there is no evidence of an organic cause.

F23.9 Acute and transient psychotic disorder, unspecified

From World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. 10th revision. Geneva: World Health Organization; 1992, with permission.

There is no specific section for acute and transient psychotic disorders in the DSM-5. As a result, patients who meet these ICD-10 criteria are classified under a number of different diagnostic categories in the DSM-5. The large majority would meet the DSM-5 criteria for brief psychotic disorder, schizophreniform disorder or other specified schizophrenia spectrum, and other psychotic disorder, although the actual proportion receiving each diagnosis varies considerably across studies. In the DSM-IV field trial for schizophrenia and other psychotic disorders, of all individuals meeting the ICD-10 criteria for acute and transient psychotic disorders, 42 percent also met the DSM-III-R criteria for schizophreniform disorder, 21 percent for psychotic disorder not otherwise specified (NOS), and only 13 percent for brief reactive psychosis. In contrast, in a study from Germany, 62 percent of patients with acute and transient psychotic disorders also met the DSM-IV criteria for brief psychotic disorder, 31 percent for schizophreniform disorder, and 5 percent for psychotic disorder NOS. Part of the difference in the results of the two studies may be due to differences in the criteria of DSM-III-R brief reactive psychosis and DSM-IV brief psychotic disorder.

Notably, the mode of onset is not a diagnostic criterion for any of the above DSM-5 diagnoses. Sudden onset (i.e., within 2 weeks) is described in DSM-5 as an essential feature of brief psychotic disorder and onset of prominent psychotic symptoms within 4 weeks is included as a specifier for favorable prognosis in the criteria for schizophreniform disorder. Nonetheless, some patients with remitting psychoses also meet the criteria

for these diagnoses despite having a nonacute onset of symptoms.

Table 12.17–3.
DSM-5 Diagnostic Criteria for Schizophreniform Disorder

Diagnostic criteria

295.40 (F20.81)

A. Two (or more) of the following, each present for a significant portion of time during a 1-month period (or less if successfully treated). At least one of these must be (1), (2), or (3):
   1. Delusions.
   2. Hallucinations.
   3. Disorganized speech (e.g., frequent derailment or incoherence).
   4. Grossly disorganized or catatonic behavior.
   5. Negative symptoms (i.e., diminished emotional expression or avolition).
B. An episode of the disorder lasts at least 1 month but less than 6 months. When the diagnosis must be made without waiting for recovery, it should be qualified as "provisional."
C. Schizoaffective disorder and depressive or bipolar disorder with psychotic features have been ruled out because either (1) no major depressive or manic episodes have occurred concurrently with the active-phase symptoms, or (2) if mood episodes have occurred during active-phase symptoms, they have been present for a minority of the total duration of the active and residual periods of the illness.
D. The disturbance is not attributable to the physiological effects of a substance (e.g., a drug of abuse and a medication) or another medical condition.

Specify if:

With good prognostic features: This specifier requires the presence of at least two of the following features: onset of prominent psychotic symptoms within 4 weeks of the first noticeable change in usual behavior or functioning, confusion or perplexity, good premorbid social and occupational functioning, and absence of blunted or flat affect.

Without good prognostic features: This specifier is applied if two or more of the above features have not been present.

Specify if:

With catatonia (refer to the criteria for catatonia associated with another mental disorder, pp. 119–120, for definition).

Coding note: Use additional code 293.89 (F06.1) catatonia associated with schizophreniform disorder to indicate the presence of the comorbid catatonia.

Specify current severity:

Severity is rated by a quantitative assessment of the primary symptoms of psychosis, including delusions, hallucinations, disorganized speech, abnormal psychomotor behavior, and negative symptoms. Each of these symptoms may be rated for its current severity (most severe in the last 7 days) on a 5-point scale ranging from 0 (not present) to 4 (present and severe). (See Clinician-Rated Dimensions of Psychosis Symptom Severity in the chapter "Assessment Measures.")

Note: Diagnosis of schizophreniform disorder can be made without using this severity specifier.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Epidemiology

Acute and transient psychotic disorders are rare, especially in industrialized settings. As a result, data on their incidence or prevalence

are limited with a particular paucity of research on their occurrence in developing countries. Much of the current knowledge of these disorders has come from clinical samples, but their brief nature means that many individuals with acute and transient psychoses will never come to the attention of mental health services. Additional information about the epidemiology of these conditions is the result of international studies in developing and industrialized settings in the 1980s and 1990s. For example, a first-episode psychosis study from Nottingham, England, estimated the incidence rate of cases that remained diagnostically stable for 3 years to be 1.36 per 100,000 (0.74 per 100,000 for men and 1.99 for women). Another epidemiological study reported a substantially higher incidence in developing countries compared with industrialized countries. These psychoses comprised 35 percent of all first contact nonaffective psychoses in developing countries compared to only 6 percent in industrialized countries. Age of onset was also younger (mid-20s) in the developing than industrialized (mid-20s to 30s) countries.

The most common specific disorder in the group of ICD-10 acute and transient psychotic disorders is likely polymorphic psychotic disorder without symptoms of schizophrenia, comprising between a third to a half of all cases of acute and transient psychotic disorders, followed in frequency of occurrence by polymorphic psychotic disorder with symptoms of schizophrenia.

Brief psychotic disorder is perhaps even less common than acute and transient psychotic disorders, although the two disorders overlap considerably. In a sample of over 1,000 patients with nonorganic psychotic or major affective disorder hospitalized in in Germany, individuals with brief psychotic disorder comprised only 2.5 percent of the cases. Similarly, in the Suffolk County, New York, first-admission cohort, a primary diagnosis of brief psychotic disorder was found in only 2 percent of the cohort sample. The incidence may be higher in first onset samples compared to first admission samples, because some of these cases may remit before clinical attention is sought. A study of U.S. Air Force recruits during basic training reported a 6-week incidence rate of 1.43 cases of DSM-III-R brief reactive psychosis per 100,000 recruits. However, this number may be an overestimate of general population rates because basic training is a particularly stressful period.

Similar to acute and transient psychotic disorders, brief psychotic disorder is more common among women than men. Also similar to acute and transient psychotic disorders, age of onset of brief psychotic disorder in industrialized settings appears to be older than in developing countries.

Etiology

Little is known about the etiology of acute and transient psychotic disorders and even less is known about brief psychotic disorder. The

available evidence for acute and transient psychotic disorders points to both biological and sociocultural factors as underlying risk factors. In one study from India, for example, a history of antecedent fever was reported in 47 percent of individuals with an acute and transient psychotic disorder. In all patients, the fever had resolved before the onset of psychotic symptoms, making it unlikely that the psychosis was the direct effect of an acute infection. Similar reports from other developing country settings and classic European psychiatric literature also note high rates of febrile illness prior to the development of acute and transient psychotic disorders, which may in part explain the higher incidence of these disorders in developing country settings where infectious diseases are more common than in industrialized settings.

Information on the biochemical, physiological, and anatomical correlates of these disorders is limited, and the available data do not reveal any distinctive abnormalities that would suggest a specific etiology. Researchers from Japan have noted hypothalamic–pituitary axis abnormalities in subsets of female patients with recurrent atypical psychoses who are more likely to experience episodes at the time of menstruation and parturition; however, the relevance of these findings to other patients with acute and transient psychotic disorders is unclear. Most available data on biological correlates of these conditions are based on work involving patients with cycloid psychosis, which overlaps with ICD-10 acute transient psychotic disorders and DSM-5 brief psychotic disorder. Auditory evoked potential studies in patients with cycloid psychosis recorded P300 peaks over the left hemisphere, similar to normal controls, but patients with cycloid psychosis had a higher P300 amplitude than normal controls, suggesting a higher level of arousal. This finding is consistent with a regional cerebral blood flow study of cycloid psychosis that recorded increased hemispheric blood flow during the psychotic episode that was directly related to symptom severity and that returned to a normal level after remission. However, results of computed tomography (CT) scan studies and electrophysiological studies of patients with cycloid psychosis show little or no differences from findings in normal control groups.

No genetic studies and few family studies of ICD-10 acute and transient psychotic disorders or DSM-5 brief psychotic disorder have been conducted. A study from India reported a higher risk of acute transient psychotic disorders and a lower risk of schizophrenia and mood disorders in the first-degree relatives of probands compared with schizophrenia probands. Similar results were obtained in family studies of cycloid psychosis and reactive psychosis in other settings.

The high incidence rate of acute and transient psychotic disorders in developing country settings has also led to speculations about the role of sociocultural factors in the etiology of these syndromes. It has been

suggested that rapid cultural change and modernization in the developing countries expose individuals to loss of status and stress arising from role confusion, making individuals vulnerable to psychotic reactions.

Acute stress (within about 2 weeks of onset of psychosis) has been noted by ICD-10 as a feature of illness in some individuals with acute and transient psychotic disorders. However, the actual proportion of patients who experience acute stress in this time frame varies considerably across studies, ranging from 10 to 69 percent. Assessment of acute stress is also complicated by lack of a clear definition of such stress in ICD-10 and the possible impact of recall bias. Furthermore, the stressors likely vary according to social and cultural context. Thus, for example, among women in developing countries, departure from or return to their parental village was a major stressor in one study, whereas among men in these settings, job-related problems were the main stressor.

The specifiers "with" or "without marked stressor(s)" are noted in the DSM-5 criteria of brief psychotic disorder. DSM-5 also notes that these events "would be markedly stressful to almost anyone in similar circumstances in the individual's culture." However, there are few reliable data on the proportion of cases of brief psychotic disorder that are preceded by such stressors.

Clinical lore suggests an association between pre-existing personality disorders and psychoticism traits and both acute and transient psychotic disorders as well as brief psychotic disorder. However, reliable assessment of premorbid personality disorders in patients with psychotic conditions is often difficult and may be affected by recall bias or residual psychotic symptoms. This point was highlighted in a Danish study of acute and transient psychotic disorders in which the majority of patients were diagnosed with a pre-existing personality disorder shortly after recovery from psychosis, whereas a year later only a minority received such a diagnosis. Similarly, a study from Germany did not find significant differences with regard to premorbid adjustment or the "big five" personality dimensions (neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness) measured by self-report between patients who had recovered from an acute transient psychotic episode and normal controls.

A psychodynamic formulation of brief reactive psychosis links this condition to deficient ego strength. According to this formulation, psychosis occurs in response to an intense emotionally stressful event that, in the absence of support in the environment, overwhelms the individual's ego defense mechanisms.

Diagnosis and Clinical Features

Acute and Transient Psychotic Disorders. Recognition of acute and transient psychotic disorders early in their course is difficult. In the first

few weeks of the illness, there are often few specific indicators that would distinguish acute and transient psychotic disorders from other nonaffective psychotic disorders with acute onset. Furthermore, even in retrospect, it is often difficult to assess the duration of a psychotic episode or distinguish spontaneous remission from remission due to antipsychotic medication treatment. Therefore, a definitive diagnosis of acute and transient psychotic disorder is often only possible retrospectively and when extensive information on the patient's past history is available.

The ICD-10 provides a sequence of three "key features" to be used for diagnosing acute and transient psychotic disorders. In order of priority, these features are (a) an acute onset (within 2 weeks); (b) presence of typical syndromes, which are the basis for the subcategorization into specific disorders; and (c) the presence of associated acute stress (within about 2 weeks of onset). Acute onset is defined as "a change from a state without psychotic features to a clearly abnormal psychotic state, within a period of 2 weeks or less." Time to onset should not be confused with the time between symptom onset and peak illness severity, as the maximum severity of symptoms may occur weeks after symptom onset. Similarly "prodromal periods of anxiety, depression, social withdrawal, or mildly abnormal behaviour" should not be counted in calculating time to onset. Thus, the criterion of acute onset will still be met even if the individual reports weeks or months of these prodromal symptoms before the onset of illness. ICD-10 also suggests that a distinction be made between acute onset, as described above, and abrupt onset (i.e., onset within 48 hours), noting that the latter may be associated with even better outcome.

With regard to the presence of a typical syndrome, ICD-10 describes two such syndromes: the polymorphic syndrome and the typical schizophrenic syndrome. The polymorphic syndrome consists of marked hallucinations, delusions, and perceptual disturbances that change from day to day or even from hour to hour. This state is frequently associated with symptoms of emotional turmoil, "with intense transient feelings of happiness and ecstasy or anxieties and irritability." Although mood symptoms might be present, the criteria for a manic or depressive episode are not met.

The schizophrenic syndrome requires that the ICD-10 symptom criteria for schizophrenia be met for the majority of the time since the onset, but less than 1 month. The schizophrenic syndrome might be associated with the polymorphic syndrome as in "acute polymorphic psychotic disorder with symptoms of schizophrenia" or stand alone as in "acute schizophrenia-like psychotic disorder."

The ICD-10 section on acute and transient psychotic disorders defines *acute stress* as "events that would be regarded as stressful to most people in similar circumstances, within the culture of the person concerned." ICD-10 mentions "bereavement, unexpected loss of partner or job, marriage, or

the psychological trauma of combat, terrorism and torture" as typical examples of acute stress.

A diagnosis of acute and transient psychotic disorder requires ruling out typical manic or depressive episodes, alcohol or drug intoxication, delirium, dementia, and mental disorders due to medical etiologies.

DSM-5 Brief Psychotic Disorder. The diagnosis of brief psychotic disorder in DSM-5 is based on the presence of one or more of the following psychotic symptoms for a period of at least 1 day but less than 1 month: (1) delusions, (2) hallucinations, (3) thought disorder as evidenced by disorganized speech, and (4) disorganized or catatonic behavior. DSM-5 specifies that one of the first three symptoms must be present for the diagnosis to be made. Although sudden onset is discussed as an essential feature of brief psychotic disorder in the text of DSM-5, it is not included in the diagnostic criteria. Similar to the ICD-10 diagnostic criteria for acute and transient psychotic disorders, the DSM-5 criteria for brief psychotic disorder allow for specifiers indicating presence or absence of a marked stressor or stressors. These are defined as events that would be "markedly stressful to almost anyone in similar circumstances in the individual's culture." The manual does not specify stressors that would qualify as marked stressors, but examples reported in the literature include serious surgery or medical illness, immigration, war and other mass violence, torture, and even intensive training programs such as military training.

Although in some cases the temporal order of stressor and psychotic episode is clear, it is not always easy to determine whether a specific stressor was a precipitant or consequence of the illness. Other sources of information, such as information from spouse or friends about the level of functioning before the onset of episode, as well as history of response to similar stressors in the past, would be helpful in determining the order of events. The DSM-5 criteria for brief psychotic disorder also allow for a "with postpartum onset" specifier for episodes starting within 4 weeks of giving birth. Comorbid catatonia can be identified by a "with catatonia" specifier and the additional diagnostic codes 293.89 (F06.1). Similar to other psychotic disorders in the DSM-5, the current severity of psychotic symptoms can be optionally specified using a 5-point rating scale.

"Emotional turmoil" and "confusion" are not among the diagnostic criteria for brief psychotic disorder in DSM-5. However, they are noted as typical experiences in the text of the manual. The text also notes rapid shifts from one intense affect to another in some individuals. A study of brief reactive psychosis from Swaziland, South Africa, also notes that the clinical presentation in that setting is characterized more by volatile affect and behavior, altered states of consciousness, and dissociative experiences than by clear-cut psychotic symptoms.

Differential Diagnosis

A definitive diagnosis of acute and transient psychotic disorders or brief psychotic disorder early in the first episode of illness is difficult if not impossible. Unless the psychosis has fully remitted, the duration of an episode of illness cannot be determined. Therefore, during the first few days or weeks of an episode, the diagnosis is usually provisional.

Many aspects of the differential diagnosis of ICD-10 acute and transient psychotic disorders relate to symptom duration. A provisional diagnosis of acute polymorphic psychotic disorder without symptoms of schizophrenia would be changed to a diagnosis of ICD-10 persistent delusional disorder or other nonorganic psychotic disorder if the illness lasts more than 3 months. Similarly, a diagnosis of acute polymorphic psychotic disorder with symptoms of schizophrenia would be changed to ICD-10 schizophrenia if the episode lasts more than 1 month, which is the minimum duration criterion for schizophrenia in ICD-10 unlike the 6 month minimum duration of symptoms required for a diagnosis of schizophrenia by DSM-5.

Similar considerations apply to DSM-5 brief psychotic disorder, in which the duration of an episode is defined as less than 1 month. If the duration of illness (from prodrome through recovery) exceeds 1 month, the diagnosis would need to change depending on other symptoms and the duration of the illness. Remission of psychotic symptoms within 1 month as a result of successful treatment also makes it difficult to distinguish brief psychotic disorder from other longer lasting disorders. In such instances, the presence of continued residual symptoms or impairment may be helpful in ruling out a diagnosis of brief psychotic disorder. In many cases, however, only the long-term course of illness can resolve diagnostic uncertainties.

Perhaps the most critical differential diagnostic issue for both acute and transient psychotic disorders and brief psychotic disorder is to rule out delirium, psychotic disorder due to another medical condition, substance/medication-induced psychotic disorder and drug or alcohol intoxication or withdrawal. Careful physical and laboratory examination is often indicated. Toxicological screens and imaging studies may also be necessary. A positive drug or alcohol screen does not necessarily rule out a diagnosis of acute and transient psychotic disorder or brief psychotic disorder. Some patients with acute and transient psychotic disorder or brief psychotic disorder may use substances in an attempt to ameliorate their symptoms. The temporal relationship between onset of psychosis and the pattern of substance abuse may be helpful in determining if a psychotic disorder is substance/medication-induced. Symptoms that persist for many days after all traces of the substance are eliminated from blood and urine would support a diagnosis of acute and transient psychotic disorder or brief

psychotic disorder.

Both ICD-10 acute and transient psychotic disorder and DSM-5 brief psychotic disorder may be associated with confusion and perplexity, making it difficult to distinguish these disorders from delirium or medical causes of psychosis. Diurnal variation in consciousness, which is a characteristic of delirium, may also be confused with variable "polymorphic" psychotic symptoms seen in the acute and transient psychotic disorders. However, compared with delirium, changes in consciousness in acute and transient psychotic disorders are generally less severe and less persistent. Furthermore, a medical cause of delirium can often be identified through history, physical examination, laboratory tests, or imaging studies.

Acute psychotic episodes have been reported in a number of medical conditions including head trauma, cerebral anoxia, epilepsy, and endocrine disorders such as hyper- or hypothyroidism. Many of these episodes are associated with disturbances of consciousness. Differential diagnosis is often based on a thorough medical history and physical examination as well as laboratory investigations.

Good prognosis of brief psychotic disorder and of acute and transient psychotic disorders, variability of symptoms, and the presence of mood symptoms sometimes make it difficult to differentiate these diagnoses from mood disorders. Presence of a full mood syndrome that meets the criteria for a manic or a depressive episode excludes diagnosis of brief psychotic disorder in DSM-5 and excludes diagnosis of acute and transient psychotic disorders in ICD-10.

Patients with personality disorders, especially borderline personality disorder, sometimes experience psychotic symptoms under stressful life circumstances. These symptoms are generally transient and do not require a separate diagnosis. If the symptoms last more than 1 day, an additional diagnosis of brief psychotic disorder may be warranted.

Finally, a number of "culture-bound" syndromes with acute onset of psychotic and dissociative symptoms have been described in various cultural settings. Many of these syndromes are transient and resolve spontaneously without antipsychotic medication treatment. The nosological relationship of these conditions to acute and transient psychotic disorders or brief psychotic disorder remains unresolved.

Course and Prognosis

Acute and transient psychotic disorders have a favorable early course by definition. Full remission within 1 to 3 months is a diagnostic criterion for these disorders in the ICD-10. Although there are few first-admission studies that allow for reliable assessment of course and outcome, the available evidence indicates that most of the recovered patients remain well in ensuing years, and even among those who experience further episodes,

these episodes are transient, and the outcome is generally favorable. Recurrence rates of 39 to 47 percent have been reported in first-admission samples followed for 3 to 4 years, but full recovery is common. Only about 6 to 18 percent of patients have any residual symptoms at a later point of time, although some studies have shown that the duration of nonaffective psychoses with acute onset and remitting course may extend up to 6 months. A 12-year follow-up study from India also reported favorable long-term outcome. The long-term course and outcome of acute and transient psychotic disorders in industrialized settings have not been extensively studied. In a prospective study from Latvia, 40 percent of a group of first-admission patients with a diagnosis of acute and transient psychotic disorder were readmitted in a 26-month follow-up.

A number of studies from various countries have assessed the diagnostic stability of the ICD-10 acute and transient psychotic disorders. In studies from India, Denmark, and Latvia, between 27 and 48 percent of patients had a different diagnosis 1 to 3 years after their index episodes. Some studies have found that diagnostic stability was associated with the patients' clinical features. Thus, a study from Germany found differences between the acute schizophrenia-like psychotic disorder subtype, of whom only 30 percent maintained their original diagnosis at follow-up, and other subtypes of acute and transient psychotic disorders, of whom 67 percent maintained their original diagnosis. Similarly, in a retrospective long-term study from Japan, 63 percent of patients with acute polymorphic psychotic disorder without symptoms of schizophrenia maintained their original diagnosis for an average follow-up period of 20 years. In a prospective study from Latvia, presence of thought disorder was the strongest predictor of change of diagnosis to schizophrenia at follow-up. The diagnostic stability of an acute and transient psychotic disorder may also vary according to the patient's gender. In a British study, the diagnoses of acute and transient psychotic disorders remained unchanged in 73 percent of women but only 14 percent of men at 3-year follow-up. Most changes were into the schizophrenia or mood disorder categories.

Brief psychotic disorder also has a benign early course by definition. Full remission within 1 month is a DSM-5 criterion for this diagnosis. Some individuals, however, experience relapses. In a study of consecutive admissions for psychosis in Germany, 83 percent of the patients initially diagnosed with brief psychotic disorder maintained this diagnosis when re-evaluated 2 years later, and 70 percent experienced relapses during the 2-year follow-up. Because changes in diagnosis are more likely to occur in the early course of psychotic conditions and the German study used cross-sectional diagnosis at the time of follow-up, the rates of diagnostic change and of relapse would likely be higher than in a first-admission or first-episode sample. Nevertheless, the social functioning of these individuals was much better than individuals with schizophrenia; over 90 percent were

living independently and over half were employed.

Several other studies have also examined the stability of the brief psychotic disorder diagnosis. In a prospective epidemiologic study of a first-admission cohort from Suffolk County, New York, only 27 percent of those receiving a consensus research diagnosis of brief psychotic disorder at 6-month follow-up maintained this diagnosis when re-evaluated at the 24-month follow-up, with 9 percent rediagnosed with schizophreniform disorder and 18 percent with psychotic disorder NOS, a residual diagnostic category in DSM-IV. The remaining individuals were rediagnosed with mood disorders (27 percent), schizophrenia (9 percent), or other conditions (9 percent). A Brazilian study of first-episode patients admitted to an emergency psychiatric unit also found that only about 23 percent of patients with a clinical discharge diagnosis of brief psychotic disorder maintained this diagnosis 1 year later. These findings suggest that assessment of the prognosis of brief psychotic disorder is confounded by the diagnostic instability of the disorder.

SCHIZOPHRENIFORM DISORDER

History

The concept of schizophreniform disorder was introduced in 1939 by the Norwegian psychiatrist Gabriel Langfeldt. He divided a group of patients who had initial symptom profiles consistent with schizophrenia into two groups based on their subsequent course. One group had a poor course and was referred to as "typical schizophrenia," whereas the "schizophreniform" group had a sudden onset, associated mood symptoms, and clouding of consciousness with a more benign course.

Forty years later, schizophreniform disorder was officially included in the DSM-III in the section on psychotic disorders not elsewhere classified. However, the term was used to describe patients with a different set of characteristics than those identified by Langfeldt. In fact, when the charts of 55 of Langfeldt's original patients diagnosed with schizophreniform psychosis were re-evaluated, most patients met the DSM-III-R criteria for a mood disorder. In contrast, DSM-III described a condition that was identical to schizophrenia except with regard to its duration, with the prodromal, active, and residual phases of schizophreniform disorder lasting more than 2 weeks but less than 6 months.

DSM-III-R added the requirement that a patient given a diagnosis of schizophreniform disorder should not meet the criteria for brief reactive psychosis (i.e., the duration of symptoms should not be less than a month). In addition, the presence or absence of a series of good prognostic features was added as a specifier in order to more closely approximate Langfeldt's original conception of the disorder. DSM-III-R also dropped the criterion B of schizophrenia (impairment in functioning) from the diagnostic criteria

of schizophreniform disorder and noted that the diagnosis should be made provisionally if recovery had not yet occurred. Additional changes in the diagnostic criteria for schizophreniform disorder reflect changes to the underlying diagnostic criteria for schizophrenia between DSM-III and DSM-III-R (i.e., differences in the specific psychotic symptoms needed to establish a diagnosis, removal of the requirement for age of onset of less than 45 years). In DSM-IV and DSM-IV-TR, the diagnostic criteria for schizophreniform disorder were maintained with little change. DSM-5 introduced a few changes consistent with changes in the A criterion for schizophrenia and specifiers added to virtually all psychotic disorder diagnoses. In addition, in keeping with the other psychotic disorder diagnoses, the DSM-5 diagnosis for schizophreniform disorder allows for a "with catatonia" specifier and ratings of current severity of symptoms.

The A criterion for schizophrenia in DSM-5 removed the higher weight given in prior editions of DSM to bizarre delusions, auditory hallucinations keeping up a running commentary on the person's behavior or thoughts and auditory hallucinations of two or more voices conversing with each other. In DSM-III-R, DSM-IV, and DSM-IV-TR, presence of one of these symptoms was enough to meet the A criterion requirement; whereas, at least 2 symptoms were required if none of these highly weighted symptoms were present. In contrast, the DSM-5 gives equal weight to symptoms included under the A criterion. However, it requires that one of the symptoms be a delusion, hallucination, or disorganized speech.

Comparative Nosology

The nosology of schizophreniform disorder and, in particular, its relationship to the schizophrenia diagnosis has changed little in successive editions of the DSM. Similar to previous editions of the DSM, in DSM-5, schizophreniform disorder is conceptualized as a variant of schizophrenia (Table 12.17–3). Thus, the first criterion specifically requires that a patient meet the A criterion for schizophrenia. The criteria further specify the minimum and maximum duration at least 1 month but less than 6 months, that the patient does not have schizoaffective or mood disorder with psychotic features and that the disorder is not due to the physiological effects of a substance or another medical condition. With a similar constellation of symptoms but a duration of less than 1 month, a diagnosis of brief psychotic disorder would be made. If the duration of illness extended beyond 6 months and there is associated functional impairment, a diagnosis of schizophrenia would be warranted.

Under the broad group of acute and transient psychotic disorders, the ICD-10 contains a conceptually similar category to schizophreniform disorder, namely, acute schizophrenia-like psychotic disorder in which patients fulfill the ICD-10 criteria for schizophrenia, but the illness is of a shorter duration (Table 12.17–2). However, with acute schizophrenia-like

psychotic disorder, the duration of psychotic symptoms is less than 1 month and complete recovery usually occurs within a few days or weeks, which is shorter than the illness duration in DSM-5 schizophreniform disorder. Another distinction from schizophreniform disorder is that each of the acute and transient psychotic disorders must have an acute onset of symptoms, with 2 weeks or less from a nonpsychotic to a clearly psychotic state. This is in contrast to schizophreniform disorder in DSM-5 in which the acute onset of symptoms is one of the good prognostic features rather than a requirement for establishing the diagnosis. Many cases of DSM-5 schizophreniform disorder would be categorized as schizophrenia in ICD-10 because the minimum duration of psychosis for a diagnosis of schizophrenia in ICD-10 is only 1 month.

Epidemiology

Data on the prevalence of schizophreniform disorder are limited. One study of DSM-IV schizophreniform disorder in the Finnish general population found a lifetime prevalence of 0.07 percent overall with a fivefold greater rate in men than in women. Somewhat higher estimates were noted in a systematic review of the literature, which found a 1-year prevalence of 0.09 percent and a lifetime prevalence of 0.11 percent, although a great deal of heterogeneity in the data was observed. In one study from Suffolk County, New York, approximately 5 percent of patients with a first admission psychotic disorder met the diagnostic criteria of schizophreniform disorder on admission.

Etiology

The etiology of schizophreniform disorder remains unclear. Studies of risk factors for psychotic illness have typically combined results from individuals with schizophreniform disorder with those from individuals with schizophrenia, making it difficult to delineate risk factors for the development of schizophreniform disorder separate from schizophrenia. The heterogeneity in the course and outcome of the schizophreniform disorder further complicates etiological research. As currently conceptualized, this diagnostic category includes both individuals with schizophrenia in the early course of illness who are given the diagnosis of schizophreniform disorder provisionally and whose diagnosis would switch to schizophrenia in follow-up, as well as individuals with a remitting psychotic illness who would not meet the criteria for schizophrenia in future.

Nevertheless, available research supports the role of genetic factors in the etiology of schizophreniform disorder. One meta-analysis found that rates of schizophrenia and of mood disorder in relatives of schizophreniform probands did not differ from those rates in the relatives

of probands with schizophrenia or mood disorders, respectively. A subsequent study showed an excess risk of schizophrenia spectrum illness in relatives of schizophreniform probands but no increase in the risk of mood disorder in family members. However, an additional study of patients with schizophreniform disorders with "good prognostic features" showed an increased likelihood of mood disorder in first-degree relatives of probands. Despite this variability in the findings of family studies, there does seem to be an increase in the rates of other psychiatric illnesses among family members of those with schizophreniform disorder, suggesting that genetic factors may be playing a role in some patients.

Studies using brain imaging to look for abnormalities in individuals with schizophreniform disorder have had mixed results. One large study that examined hippocampal and amygdalar volumes found no difference between patients with schizophreniform disorder and controls. Studies of ventricular enlargement in patients with schizophreniform disorder have produced mixed results. In one study, left ventricular enlargement was correlated with a progression of illness to schizophrenia or schizoaffective disorder. Although two studies showed no progressive increases in ventricular size over 2 to 7 years of follow-up for the schizophreniform patient group as a whole, outcomes were poorer in patients with initially enlarged ventricles than in those without ventricular enlargement.

Diagnostic and Clinical Features

The mode of onset of schizophreniform disorder is similar to that of schizophrenia. Although many patients with schizophreniform disorder may experience functional impairment during an episode, they are unlikely to report a progressive decline in social and occupational functioning. The initial symptom profile is the same as that of schizophrenia in that two or more psychotic symptoms (hallucinations, delusions, disorganized speech, grossly disorganized or catatonic behavior, and negative symptoms) must be present. There is an increased likelihood of emotional turmoil and confusion or perplexity, the presence of which may indicate a good prognosis. Although negative symptoms may be present, they are relatively uncommon in schizophreniform disorder and are considered poor prognostic features. By definition, patients with schizophreniform disorder return to their baseline state within 6 months. Otherwise, the provisional diagnosis of schizophreniform disorder needs to be changed to schizophrenia. In some individuals, schizophreniform disorder is episodic with more than one episode occurring after long periods of full remission.

Differential Diagnosis

It is important to first differentiate schizophreniform disorder from psychoses that can arise from medical conditions or use of substances. This

is accomplished by taking a detailed history and physical examination, and where indicated, performing laboratory tests or imaging studies. Although it is not always possible to distinguish substance-induced psychosis from other psychotic disorders cross sectionally, a rapid onset of psychotic symptoms in a patient with a significant substance history should raise the suspicion of a substance/medication-induced psychotic disorder. Because individuals with schizophreniform disorder also may have an acute onset of psychotic symptoms, longitudinal assessment is often crucial in distinguishing a substance/medication-induced psychotic disorder from schizophreniform disorder with incidental substance use. Detailed substance-use history and toxicological screen are also important for treatment planning in an individual with a new onset of psychosis. A detailed history of medication use, including over-the-counter medications and herbal products, is essential because such therapies may also produce an acute psychosis.

The duration of psychotic symptoms is one of the features that distinguishes schizophreniform disorder from other syndromes. Schizophrenia is diagnosed if the duration of illness lasts more than 6 months, whereas symptoms that occur for less than 1 month indicate a brief psychotic disorder. One of the difficulties in making a diagnosis of schizophreniform disorder involves dating the onset of a psychotic episode that evolves insidiously. Without a distinct symptom onset, it becomes difficult to decide whether the illness has actually lasted less than 6 months. Another difficulty in establishing the diagnosis of schizophreniform disorder is that it requires knowledge of the patient's clinical course over at least 6 months after onset of psychosis. In the event that the diagnosis is made before 6 months have elapsed, the diagnosis of schizophreniform disorder is considered provisional. For some patients, prolonged observation will be needed before the duration of symptomatology can be established with certainty, particularly if the patient has received antipsychotic treatment.

It is sometimes difficult to distinguish mood disorders with psychotic features from schizophreniform disorder. The acute onset of symptoms may resemble the onset of mood disorder episodes. Additionally, mood symptoms such as loss of interest and pleasure may be difficult to distinguish from negative symptoms, avolition, and anhedonia. A thorough longitudinal history is important in elucidating the diagnosis since the presence of psychotic symptoms exclusively during periods of mood disturbance is an indication of a primary mood disorder.

Course and Prognosis

The course and prognosis of schizophreniform disorder are variable and depend largely on whether the diagnosis remains schizophreniform disorder or whether the provisional diagnosis of schizophreniform disorder

changes over the course of longitudinal follow-up. There are conflicting data on the diagnostic stability of schizophreniform disorder. One small study found that all patients in a first episode cohort who were initially diagnosed with schizophreniform disorder were subsequently diagnosed with schizophrenia or schizoaffective disorder at 2-year follow-up. However, most studies have found that 45 to 75 percent of patients who receive a diagnosis of schizophreniform disorder will have a subsequent change in diagnosis to schizophrenia or schizoaffective disorder with longitudinal assessment. When predictive factors have been assessed, progression to schizophrenia seems to be associated with poor premorbid functioning, the presence of blunted affect, or the designation of schizophreniform disorder without good prognostic features. Evolution of schizophreniform disorder to a mood disorder was less common, occurring in about 5 to 15 percent of patients in most studies with a maximal rate of 35 percent in a single study. Based on data from one study, greater rates of progression to a mood disorder may be seen in individuals with schizophreniform disorder who have good prognostic features. Only about 20 to 40 percent of patients maintained the diagnosis of schizophreniform disorder over follow-up periods of several years. With longer periods of follow-up, the proportion of individuals who retain the diagnosis of schizophreniform disorder is even smaller. In the Suffolk County, New York, first-admission sample, less than 5 percent of those who initially received a diagnosis of schizophreniform disorder retained that diagnosis at the time of the 10-year follow-up assessment. These findings underscore the importance of regarding the diagnosis of schizophreniform disorder as provisional when it is made cross sectionally.

For individuals whose diagnosis remains schizophreniform disorder after longitudinal follow-up, the illness course is usually benign. In general, these patients will have met criteria for “good prognostic features” upon first presentation.

Individuals classified as “without good prognostic features” often have a course that resembles that of patients initially diagnosed with schizophrenia. For many of these patients, their diagnosis indeed changes to schizophrenia with longitudinal observation. For patients with schizophreniform as a whole, outcomes tend to be less favorable than for individuals with mood disorders, with a more favorable course and a lesser degree of negative symptoms than in individuals with schizophrenia.

DELUSIONAL DISORDER

History

Delusional disorders, which were once referred to as paranoid disorders or paranoia, have been recognized since the time of the ancient Greeks. In the early 19th century Johann Heinroth and later Jean Esquirol described a

disorder that influenced the intellect in which irrational ideas and actions appeared to develop through plausible and logical reasoning. In 1863 Karl Kahlbaum, who referred to the syndrome as "diastrephia," made the important observation that personality is preserved during the course of the illness. Late in the 19th century Valentin Magnan described a chronic systematized delusional disorder in which delusional ideas were well circumscribed and highly structured.

Kraepelin described paranoia as a third form of psychosis, distinct from both schizophrenia and mood disorder and consisting of chronic, nonbizarre, well-systematized delusions without the typical deterioration of dementia praecox. He noted that, "One invariably observes the insidious development of a permanent unshakable delusional system from inner causes in which clarity and order of thinking, willing and action are completely preserved." In the sixth edition of his textbook, Kraepelin also introduced the concept of paranoia in the elderly, which would later be termed paraphrenia.

Table 12.17–4.
DSM-5 Diagnostic Criteria for Delusional Disorder

Diagnostic criteria
297.1 (F22)

1. The presence of one (or more) delusions with a duration of 1 month or longer.
2. Criterion A for schizophrenia has never been met.
   Note: Hallucinations, if present, are not prominent and are related to the delusional theme (e.g., the sensation of being infested with insects associated with delusions of infestation).
3. Apart from the impact of the delusion(s) or its ramifications, functioning is not markedly impaired, and behavior is not obviously bizarre or odd.
4. If manic or major depressive episodes have occurred, these have been brief relative to the duration of the delusional periods.
5. The disturbance is not attributable to the physiological effects of a substance or another medical condition and is not better explained by another mental disorder, such as body dysmorphic disorder or obsessive–compulsive disorder.

Specify whether:

Erotomanic type: This subtype applies when the central theme of the delusion is that another person is in love with the individual.

Grandiose type: This subtype applies when the central theme of the delusion is the conviction of having some great (but unrecognized) talent or insight or having made some important discovery.

Jealous type: This subtype applies when the central theme of the individual's delusion is that his or her spouse or lover is unfaithful.

Persecutory type: This subtype applies when the central theme of the delusion involves the individual's belief that he or she is being conspired against, cheated, spied on, followed, poisoned or drugged, maliciously maligned, harassed, or obstructed in the pursuit of long-term goals.

Somatic type: This subtype applies when the central theme of the delusion involves bodily functions or sensations.

Mixed type: This subtype applies when no one delusional theme predominates.

Unspecified type: This subtype applies when the dominant delusional belief cannot be clearly determined or is not described in the specific types (e.g., referential delusions without a prominent persecutory or grandiose component).

Specify if:

With bizarre content: Delusions are deemed bizarre if they are clearly implausible, not understandable, and not derived from ordinary life experiences (e.g., an individual's belief that a stranger has removed his or her internal organs and replaced them with someone else's organs without leaving any wounds or scars).

Specify if:

The following course specifiers are only to be used after a 1-year duration of the disorder:

First episode, currently in acute episode: First manifestation of the disorder meeting the defining diagnostic symptom and time criteria. An acute episode is a time period in which the symptom criteria are fulfilled.

First episode, currently in partial remission: Partial remission is a time period during which an improvement after a previous episode is maintained and in which the defining criteria of the disorder are only partially fulfilled.

First episode, currently in full remission: Full remission is a period of time after a previous episode during which no disorder-specific symptoms are present.

Multiple episodes, currently in acute episode

Multiple episodes, currently in partial remission

Multiple episodes, currently in full remission

Continuous: Symptoms fulfilling the diagnostic symptom criteria of the disorder are remaining for the majority of the illness course, with subthreshold symptom periods being very brief relative to the overall course.

Unspecified

Specify current severity:

Severity is rated by a quantitative assessment of the primary symptoms of psychosis, including delusions, hallucinations, disorganized speech, abnormal psychomotor behavior, and negative symptoms. Each of these symptoms may be rated for its current severity (most severe in the last 7 days) on a 5-point scale ranging from 0 (not present) to 4 (present and severe). (See Clinician-Rated Dimensions of Psychosis Symptom Severity in the chapter "Assessment Measures.")

Note: Diagnosis of delusional disorder can be made without using this severity specifier.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

The distinction between paranoid disorder and paranoid schizophrenia was maintained by Eugen Bleuler. However, Kurt Kolle reassessed Kraepelin's patients at a later time and showed that their pattern of deterioration was similar to that of patients with dementia praecox. This view was expressed in DSM-II in which paranoid states were regarded as possible variants of schizophrenia.

Beginning with the third edition, the DSM adopted the Kraepelinian concept of paranoia as consisting of nonbizarre delusions without prominent hallucinations. However, DSM-III-R referred to paranoia as delusional disorder for clarity and to encompass other types of delusion in addition to delusions of jealousy or persecution. The DSM-IV and DSM-IV-TR also acknowledged that delusions may show varying levels of conviction

and that delusional beliefs may be partially responsive to evidence or reason. The major change between DSM-5 and prior editions of the DSM in the definition of delusional disorder is the removal of the requirement that the delusions be nonbizarre. Furthermore, in DSM-5 a specifier for delusions "with bizarre content" was added if the delusions are clearly implausible, not understandable, and not derived from ordinary life experiences. The example of bizarre delusion provided in the manual is the delusional belief that the person's internal organs have been replaced with someone else's organs without leaving any wounds or scars. The DSM-5 further added course specifiers and current severity specifiers to the criteria (Table 12.17–4).

Comparative Nosology

In the DSM-5 criteria for delusional disorder, the emphasis is on the presence of one or more delusions lasting for more than a month, in the absence of other criterion A symptoms of schizophrenia. Hallucinations if present are not prominent and are related to the content of delusions.

The ICD-10 description of delusional disorder, similar to the prior editions of DSM requires that the delusions be "other than those listed as typically schizophrenic" (i.e., other than completely impossible, culturally inappropriate or Schneiderian first-rank symptoms) (Table 12.17–5).

Table 12.17–5.
ICD-10 Diagnostic Criteria for Delusional Disorders

A. A delusion or a set of related delusions, other than those listed as typically schizophrenic in criterion G1(1)b or d for paranoid, hebephrenic, or catatonic schizophrenia (i.e., other than completely impossible or culturally inappropriate), must be present. The commonest examples are persecutory, grandiose, hypochondriacal, jealous (zelotypic), or erotic delusions.
B. The delusion(s) in criterion A must be present for at least 3 months.
C. The general criteria for schizophrenia are not fulfilled.
D. There must be no persistent hallucinations in any modality (but there may be transitory or occasional auditory hallucinations that are not in the third person or giving a running commentary).
E. Depressive symptoms (or even a depressive episode) may be present intermittently, provided that the delusions persist at times when there is no disturbance of mood.
F. There must be no evidence of primary or secondary organic mental disorder as listed under organic, including symptomatic, mental disorders, or of a psychotic disorder due to psychoactive substance use.

Specification for possible subtypes

The following types may be specified if desired: Persecutory; litigious; self-referential; grandiose; hypochondriacal (somatic); jealous; erotomanic.

Other persistent delusional disorders

This is a residual category for persistent delusional disorders that do not meet the criteria for delusional disorder. Disorders in which delusions are accompanied by persistent hallucinatory voices or by schizophrenic symptoms that are insufficient to meet criteria for schizophrenia should be coded here. Delusional disorders that have lasted for less than 3 months should, however, be

coded, at least temporarily, under acute and transient psychotic disorders.
Persistent delusional disorder, unspecified

From World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization; 1993, with permission.

Furthermore, ICD-10 criteria for delusional disorder require that the duration of the delusion be greater than 3 months. Delusional syndromes that last less than 3 months would be classified under acute and transient psychotic disorders as other acute predominantly delusional psychotic disorders (Table 12.17–2).

Notwithstanding these differences, delusional disorder is typically chronic in nature, and both systems of classification will most likely include similar patients. Similar to DSM-5, the ICD-10 allow for the presence of transient and nonprominent hallucinations.

Epidemiology

Studies of the epidemiology of delusional disorders have been difficult to conduct because the condition occurs infrequently. However, in a general population study from Finland, the lifetime prevalence of DSM-IV delusional disorder was 0.18 percent. In clinical studies, individuals with delusional disorder represent from 1 to 4 percent of psychiatric admissions. The persecutory type is the most common type of delusional disorder. Other than the jealous type that appears to be more common in men, delusional disorders appear to be equally common in the two genders.

Etiology

The etiology of delusional disorder is not known and its infrequent occurrence limits the ability to examine genetic or other risk factors. However, small family studies have failed to identify strong genetic links between delusional disorder and schizophrenia.

Psychodynamic theories hypothesize that persecutory delusions, such as those found in delusional disorder, are a protective psychological response to stresses or conflicts that represent a profound threat to the self (e.g., fear of the unknown, homosexual panic, and adjustment to immigration). In responding to the threat, there is an affective withdrawal from intimate ties and an intense effort to retain the appearance of normalcy through defensive mechanisms such as projection. At its extreme, the paranoid defense fails to relieve anxiety and escalates to a distinct maladaptive state, which is postulated to result in the development of a delusional disorder.

Another theory of the etiology of delusional disorder is based on research by cognitive and experimental psychologists and suggests that persons with persecutory delusions selectively attend to threatening information, jump to conclusions on the basis of insufficient information,

attribute negative events to external personal causes, and have difficulty in envisaging others' intention and motivations. Preferential recall of threatening episodes has also been postulated as a mechanism for reinforcing the delusional belief.

Diagnostic and Clinical Features

The DSM-5 lists five specific subtypes of delusional disorder based on their content: persecutory, jealous, somatic, grandiose, and erotomanic. It also allows for mixed and unspecified types.

Persecutory Type. The persecutory type is the best recognized type of delusional disorder in which the individual is convinced that he or she (or someone to whom the person is close) is being malevolently treated or harmed in some way. These persecutory beliefs are often associated with querulousness, irritability, anger, and at times aggressive or even homicidal behavior. The individual may also become preoccupied with formal litigation. In contrast to the persecutory delusions of schizophrenia, the delusions in delusional disorder are systematized and coherent. With the exception of impairments related to the delusion itself, social functioning is not compromised. In some of these individuals, the demarcation between normality, pathological behavior, overvalued ideas, and delusion may not be very clear.

Jealous Type. Feelings of jealousy can fall on a continuum from normal jealousy to pathological jealousy and to jealousy that is delusional in nature, with the thresholds for abnormal jealousy being, in part, socially determined.

The most common delusion of jealousy is the belief that one's spouse or partner is unfaithful. In some instances, some degree of infidelity may actually be occurring, yet the magnitude of the jealous response and the "evidence" accumulated to support accusations of unfaithfulness may take on a delusional quality. Jealousy is dangerous for it not only evokes anger but also arms the jealous individuals with a sense of righteous indignation to justify their acts of aggression. Jealousy focuses primarily on the partner and can sometimes lead the partner or others to be harmed or even murdered.

Erotomanic Type. In the erotomanic type of delusional disorder, which has also been referred to as De Clerambault syndrome or "psychose passionelle," the patient has the delusional conviction that another person, usually of higher status, is in love with him or her. Patients will generally describe the love object as having fallen in love with them first and as having made the first advances. Patients also describe elaborate rationalizations when the love object ignores or rejects their contacts and

communications. Patients who develop erotomania often tend to be solitary, withdrawn, dependent, and sexually inhibited, as well as having poor levels of social or occupational functioning. Symptoms typically have an acute onset and then become chronic.

Erotomanic symptoms may also occur in the context of other psychiatric disorders. In fact, the majority of patients with erotomanic symptoms will have a diagnosis other than delusional disorder.

Stalking may be observed in individuals with the erotomanic type of delusional disorder, although most stalkers do not meet the criteria for erotomania. Stalkers most often persecute their targets by unwanted communications that consist of frequent telephone calls, letters, e-mail, text messages, graffiti, notes, or packages. Typically, stalkers will also follow their target and make unwelcome or intrusive and unexpected appearances in his or her public domain. Along with this pattern of intrusive behavior is an implicit or explicit threat, which leads the stalking target to experience considerable fear. In some circumstances, the violence may escalate until the stalker actually murders the target or individuals associated with the target.

Somatic Type. The somatic type of delusional disorder is the only type of the 5 specific types that was not included in Kraepelin's original description of paranoia. It is diagnosed when the central themes of the delusional beliefs are of a hypochondriacal or somatic nature. As a result, in the older literature, this type of delusional disorder is sometimes referred to as monosymptomatic hypochondriasis. As with other subtypes of delusional disorder, the patient must have a clear sensorium and the symptoms may not result from an underlying general medical condition or another psychiatric disorder.

In the somatic type of delusional disorder, the onset of symptoms may be gradual or sudden. In most patients the illness is unremitting, although the severity of the delusion may fluctuate. Anxiety and hyperalertness also characterize patients with this subtype. Some themes recur, such as concerns about infestation in delusional parasitosis, preoccupation with body features with dysmorphic delusions, and delusional concerns about body odor, which are sometimes referred to as "bromosis." In delusional parasitosis, tactile sensory phenomena are often linked to the delusional beliefs. In Europe, delusional parasitosis is generally thought to represent a primary hallucination with an accompanying secondary delusion that provides a rationalization for the sensory perception, whereas in the United States, the sensory phenomena associated with delusional parasitosis are viewed as components of a systematized and encapsulated delusion.

Patients with the somatic type of delusional disorder rarely present for psychiatric evaluation, and when they do, it is usually in the context of a psychiatric consultation. Instead, these patients generally present to a

specific medical specialist for evaluation. Thus, these individuals are more often encountered by dermatologists, plastic surgeons, urologists, and other medical specialists, including dentists.

Grandiose Type. The grandiose type of delusional disorder was included in Kraepelin's original description of paranoia. However, it is relatively rare to encounter a patient who has grandiose delusions alone without associated symptoms of mania that would suggest a diagnosis of bipolar disorder with psychotic features, or other psychotic symptoms that would indicate a diagnosis of schizophrenia. Some patients with delusional disorder may have elements of grandiosity in their thoughts, but these are rarely elaborated into a full-blown delusion.

Mixed Type. The mixed type of delusional disorder includes individuals who exhibit more than one of the preceding types simultaneously. These delusions may be bizarre in nature. Of note, under DSM-IV-TR, the presence of a bizarre delusion would have satisfied the criterion A for schizophrenia or schizophreniform disorder and ruled out a diagnosis of delusional disorder. However, under DSM-5, at least two of the criterion A symptoms need to be met for a diagnosis of schizophrenia or schizophreniform disorder and delusional disorder can be diagnosed even if the delusions are bizarre.

Unspecified Type. Although the preceding themes are the most common themes observed among individuals with delusional disorder, other delusions may occur as well. Examples include the misidentification syndromes such as the Fregoli syndrome, in which strangers are misidentified as a single persecutor wearing different disguises and the Capgras syndrome, in which the individual believes that his or her family members or other familiar people have been replaced by identical looking imposters. Another example is the Cotard syndrome, in which individuals report that they have lost all of their possessions and status as well as their entire being and even their internal organs. Individuals experiencing misidentification delusions may act violently toward the presumed imposter, believing that the imposter is doing harm to the individual or a loved one.

Differential Diagnosis

The first step in the clinical assessment is establishing whether a delusion exists. Some statements that initially seem to be delusional may, in fact, be true. In contrast, reports of circumstances that initially seem believable may only be clearly identified as delusions as the symptoms worsen, the delusion becomes less encapsulated, or more information comes to light. At times having multiple sources of information is helpful in deciding the

delusional nature of a belief.

Certain behavioral features such as hypersensitivity, guardedness, evasiveness, secretiveness, litigiousness, overly detailed descriptions, hostility, and humorlessness are characteristic of the delusional patient yet are not pathognomonic. The patient may also be secretive, vague, and not voluntarily provide the clinician with necessary details. At other times, the patient may see no need to convey information, believing that the interviewer is part of the "conspiracy" and already knowledgeable about its specifics. For example, the patient may view the interview as a staged game where both parties know the hidden meanings behind the overt text. As a result, certain phrases are commonly heard from individuals with persecutory ideas in the context of delusional disorder (e.g., "Don't play games with me," "You know what 'they' have been doing").

Characteristic observations or reports about the patient's behavior (e.g., checking for microphones and checking the spouse's clothes) may also lend support to the diagnosis. Even with such observations it is sometimes difficult to distinguish between a true situation, an overvalued idea, or a delusion. Confabulation may at time be difficult to distinguish from delusions, although delusions are usually more systematized. Often, the inappropriateness of the patient's behavior or emotional response, rather than the simple truth or falsity of the belief, will indicate its delusional nature. Consequently, when evaluating individuals with possible delusional disorder, the most sensible approach is to gather as much data as possible from many sources and then utilize clinical judgment in determining whether a threshold indicating psychopathologic disturbance has been passed.

In making a diagnosis of delusional disorder, the first step is to eliminate general medical conditions as a potential cause of delusions. Many general medical conditions may be associated with the development of delusions, and delusion may at times accompany a delirious state.

Metabolic conditions and disorders affecting the limbic system and basal ganglia are most often associated with the emergence of delusional beliefs. Complex delusions occur more frequently in patients with subcortical pathology. In Huntington disease and in individuals with idiopathic basal ganglia calcifications, for example, more than 50 percent of patients demonstrate delusions at some point in their illness. Anosognosia (i.e., nonrecognition of disability) and reduplicative paramnesia (i.e., individuals believing they are in different places at the same time) that accompany right-side cerebral infarctions may seem like delusional beliefs. Capgras syndrome has been observed in a number of medical disorders including CNS lesions, vitamin B_{12} deficiency, hepatic encephalopathy, diabetes, and hypothyroidism. Focal lesions associated with such misidentification delusions have more often involved the right than the left hemisphere. Delusions of infestation, lycanthropy (i.e., the false belief that

the patient is an animal, often a wolf or werewolf), and erotomania have been reported in small numbers of patients with epilepsy, CNS lesions, or toxic-metabolic disorders. In such cases, a diagnosis of psychotic disorder due to another medical condition is appropriate.

In dementia syndromes of all types, 15 to 50 percent of individuals experience delusions. Most of these delusions are simple, although a minority of patients may have complex, well-systematized delusions. Persecutory delusions are particularly common in early dementia and may be the first manifestation of a dementing illness in some individuals.

Once general medical conditions, delirium, and dementia have been eliminated as possible diagnoses, substance use and medications must be explored as possible causes. Among the substances of abuse, amphetamines and cocaine are most likely to result in delusional beliefs, which are typically persecutory in nature. Prescribed substances, over-the-counter medications, and herbal products such as ephedra may also be associated with the development of delusions. Delusional beliefs clearly linked to a substance of abuse or medication are categorized as substance/medication-induced psychotic disorder.

In differentiating delusional disorder from other psychotic disorders, a number of factors need to be taken into consideration. Individuals with schizophrenia will often have hallucinations or disorganized thoughts or behaviors that are not present in delusional disorder. In addition, prominent negative symptoms or deterioration in function from previous premorbid levels would suggest a diagnosis of schizophrenia. Functional impairment in delusional disorder is a consequence of the delusions themselves (e.g., being fired from work after harassing the boss in the context of an erotomanic delusion).

Mood disorders are also important in the differential diagnosis of delusional disorder. Mood symptoms are highly prevalent in delusional disorders. Symptoms of depression may seem to be an understandable response to the perceived delusional experiences (e.g., feeling overwhelmed and demoralized over “harassment” by neighbors or coworkers). In addition, given the prevalence of mood disorders, many individuals who present with a delusional disorder will be found to have had a history of depression. Somatic, grandiose, or persecutory delusions that are similar to those observed in patients with delusional disorder may also be seen in individuals with psychotic mood disorders. In distinguishing mood disorders with psychotic features from delusional disorder, it is therefore important to determine whether a full mood syndrome (manic, depressed, or mixed episode) is present or whether milder mood symptoms that do not meet criteria for a mood disorder are present. It is also important to consider the temporal pattern of symptom onset and progression, recognizing that the boundaries between mood disorder and delusional disorder will generally become clear over time. If mood disorders have

occurred concurrently with the delusions, their duration should have been brief relative to the duration of delusions.

Perhaps most difficult in terms of differential diagnosis is differentiating among an overvalued idea, an obsession with lack of insight, and a delusion. For example, the somatic type of delusional disorder may be easily confused with illness anxiety disorder (hypochondriasis). In illness anxiety disorder, there is a preoccupation with fears or ideas of having a serious disease, and these concerns are based on the person's misinterpretation of bodily symptoms. In addition, the somatic preoccupation persists despite appropriate medical evaluation and reassurance. Although the patient may be briefly relieved and begin to doubt his or her ideas of having an illness for a short while after each evaluation, the fear generally emerges again. Although many individuals with illness anxiety disorder have some realization of the excessive nature of their fears, some individuals may lack insight, which makes differential diagnosis difficult. In most individuals with illness anxiety disorder, however, there is a longer pattern of being preoccupied with physical symptoms, and the preoccupations usually involve more than one symptom or organ system. Furthermore, somatic delusions are generally more bizarre than the health concerns in illness anxiety disorder.

In body dysmorphic disorder, there is a preoccupation with a slight or imagined defect in appearance. These individuals may recognize that their view of their appearance is distorted, yet a significant proportion of individuals with body dysmorphic disorder will hold their beliefs with delusional intensity. In these cases, a diagnosis of body dysmorphic disorder, with "absent insight/delusional beliefs" specifier is appropriate.

Similarly, reality testing may be poor in some individuals with OCD. The DSM-5 diagnosis of OCD requires that the individual must recognize that the obsessions or compulsions are excessive or unreasonable at some time. Nonetheless, patients show varying degrees of insight into the validity of their beliefs, and some individuals do not recognize that their symptoms are unreasonable. In DSM-IV-TR, these individuals would have been diagnosed as having OCD "with poor insight." If the obsession reaches "delusional proportions," DSM-5 suggests that a diagnosis of OCD with the "absent insight/delusional beliefs" specifier is the appropriate diagnosis.

When persecutory ideas are present and patients are guarded about sharing full details of their delusional beliefs, it may be particularly difficult to differentiate delusional disorder from a paranoid personality disorder. In these cases, the longitudinal history will be important in establishing a diagnosis of a personality disorder and also in determining whether delusions are now superimposed on the individual's premorbid personality. If a patient met the criteria for a paranoid personality disorder prior to the onset of psychosis and after the psychosis has subsided, then such individuals would be given a diagnosis of delusional disorder as well as a

diagnosis of paranoid personality disorder.

Course and Prognosis

The course of the disorder has been difficult to study for a number of reasons including the relatively low prevalence, the unstable nature of the diagnosis, and the tendency for patients with delusional disorder to avoid treatment by medical professionals in general and by psychiatrists in particular. The age of onset of delusional disorder is generally in the mid- to late-30s, with the vast majority of individuals having an insidious onset of symptoms. In addition, delusional disorder is generally associated with less deterioration in functioning over time than schizophrenia or schizoaffective disorder. Although the specific delusions may persist and remain fixed, they are generally stable. Functioning tends to be preserved, and negative outcomes from the disorder typically occur when individuals are unable to refrain from acting on their delusional beliefs. The impact of age of onset on course is unclear, with some studies showing more favorable prognosis with earlier age of onset and others showing no effect of age of onset on course. In some studies, the presence of precipitating factors has been associated with better outcome, along with female sex and being married. The presence of co-occurring diagnoses, particularly depression, may be an additional factor that negatively influences the course of the disorder.

SCHIZOAFFECTIVE DISORDER

History

Kraepelin divided the psychoses into dementia praecox characterized by progressive deterioration and manic-depressive illness characterized by a remitting course; however, his rich clinical examples included many patients with features of both disorders. In the early 20th century, Lange, Bowman, and Dunton also described intermittent forms of dementia praecox that presented with periods of excitement or depression along with gradual deterioration in functioning. Bleuler believed that the existence of mood and schizophrenic symptoms during the same episode did not alter the diagnosis of schizophrenia and pointed out the need for prolonged observation to correctly differentiate schizophrenia from manic-depressive psychosis.

Jacob Kasanin was the first to coin the term schizoaffective psychosis to identify patient with characteristics of both schizophrenia and mood disorders. He based his observation on a distinct subgroup of nine young patients with good premorbid functioning, a sudden onset of both schizophrenic and mood symptoms, and good outcome. However, a review of his detailed case histories shows that all of these individuals could have been diagnosed as having a mood disorder with psychotic features under

the DSM classification.

DSM-I described a schizoaffective type of schizophrenia reaction diagnosed in individuals "showing significant admixtures of schizophrenic and affective reactions." The manual also noted that "The mental content may be predominantly schizophrenic, with pronounced elation or depression. Cases may show predominantly affective changes with schizophrenic-like thinking or bizarre behavior." However, the long-term course of these individuals was said to be "basically schizophrenic in nature." DSM-II maintained the diagnosis of schizophrenia, schizoaffective type and further divided the category into excited and depressed subtypes.

Under DSM-III, disorders involving full affective syndromes with schizophrenia-like symptoms were classified as mood disorders with mood-incongruent psychotic symptoms. As long as the schizophrenic symptoms occurred at the same time as the mood disturbance, the patient was considered to have a mood disorder. The DSM-III relegated schizoaffective disorder to the psychotic disorders not elsewhere classified, reserved for cases that did not clearly meet the diagnostic criteria for schizophrenia or mood disorders. In the DSM-III-R, schizoaffective disorder became more clearly defined with the specification that "there have been delusions or hallucinations for at least 2 weeks, but no prominent mood symptoms" and that "the duration of all episodes of a mood syndrome has not been brief relative to the total duration of the psychotic disturbance." DSM-IV and DSM-IV-TR incorporated the stricter time frame of 1 month duration of schizophrenia symptoms (as part of criterion A). DSM-5 further elaborated on the duration of mood symptoms by requiring that the symptoms for a major mood episode should be present for the majority of the total duration of the active and residual portions of the illness.

Comparative Nosology

Both the DSM-5 and ICD-10 criteria for schizoaffective disorder require that there be a period of concurrent psychosis and disturbance of mood (Tables 12.17–6 and 12.17–7).

DSM-5 specifically requires "An uninterrupted period of illness during which there is a major mood episode (major depressive or manic) concurrent with Criterion A of Schizophrenia."

ICD-10 requires that the individual "meets the criteria for one of the affective disorders of moderate or severe degree," and that during the same episode of the disorder (and concurrently for at least part of the episode) psychotic symptoms similar to those of schizophrenia must also be present. The primary distinction between DSM-5 and ICD-10 is that DSM-5 requires a period of at least 2 weeks during the lifetime duration of illness in which delusions or hallucinations occur in the absence of a major mood episode. In addition, DSM-5 requires that "Symptoms that meet criteria for a mood episode are present for the majority of the total duration of the

active and residual portions of the illness." In contrast, ICD-10 requires that mood and psychotic symptoms both be a prominent part of the clinical picture with a balance between the number, severity, and duration of the schizophrenic and mood symptoms. ICD-10 also does not include typical mood-congruent psychotic symptoms, such as persecutory or grandiose delusions, among the psychotic symptoms that can fulfill the diagnostic criteria.

In both DSM-5 and ICD-10, schizoaffective disorder is divided into two subtypes based on whether or not manic episodes are present with or without major depressive episodes (bipolar type) or the mood episodes are limited to the major depressive episodes (depressive type). Both criteria also exclude individuals whose symptoms are the result of general medical conditions or substance-induced disorders.

Epidemiology

Few studies have examined the prevalence of schizoaffective disorder in general population samples. A general population study in Finland using DSM-IV criteria found a lifetime rate of schizoaffective disorder of 0.32 percent.

Gender differences in the rates of schizoaffective disorder in clinical samples generally parallel gender differences seen in mood disorders, with approximately equal numbers of men and women who have the bipolar subtype and a more than twofold female to male predominance among individuals with the depressed subtype of schizoaffective disorder. Individuals with schizoaffective disorder are more likely to be married or employed than individuals with schizophrenia but less likely than those with bipolar disorder. Other sociodemographic risk factors have not been well studied.

Etiology

In etiologic research, analyses have typically combined individuals with schizoaffective disorder and with schizophrenia into a single category. As a result, little is known about the distinctive etiologic factors associated with schizoaffective disorder. However, genetic factors may play a role because relatives of probands with schizoaffective disorder have higher rates of mood disorder than relatives of probands with schizophrenia, and higher rates of schizophrenia than relatives of probands with mood disorder. Recent studies of the disrupted in schizophrenia 1 (*DISC1*) gene, located on chromosome 1q42, suggest its possible involvement in schizoaffective disorder as well as schizophrenia and bipolar disorder.

Table 12.17–6.
DSM-5 Diagnostic Criteria for Schizoaffective Disorder

Diagnostic criteria

1. An uninterrupted period of illness during which there is a major mood episode (major depressive or manic) concurrent with Criterion A of schizophrenia.
   Note: The major depressive episode must include Criterion A1: Depressed mood.
2. Delusions or hallucinations for 2 or more weeks in the absence of a major mood episode (depressive or manic) during the lifetime duration of the illness.
3. Symptoms that meet criteria for a major mood episode are present for the majority of the total duration of the active and residual portions of the illness.
4. The disturbance is not attributable to the effects of a substance (e.g., a drug of abuse and a medication) or another medical condition.

Specify whether:

295.70 (F25.0) **Bipolar type:** This subtype applies if a manic episode is part of the presentation. Major depressive episodes may also occur.

295.70 (F25.1) **Depressive type:** This subtype applies if only major depressive episodes are part of the presentation.

Specify if:

With catatonia (refer to the criteria for catatonia associated with another mental disorder, pp. 119–120, for definition).

Coding note: Use additional code 293.89 (F06.1) catatonia associated with schizoaffective disorder to indicate the presence of the comorbid catatonia.

Specify if:

The following course specifiers are only to be used after a 1-year duration of the disorder and if they are not in contradiction to the diagnostic course criteria.

First episode, currently in acute episode: First manifestation of the disorder meeting the defining diagnostic symptom and time criteria. An acute episode is a time period in which the symptom criteria are fulfilled.

First episode, currently in partial remission: Partial remission is a time period during which an improvement after a previous episode is maintained and in which the defining criteria of the disorder are only partially fulfilled.

First episode, currently in full remission: Full remission is a period of time after a previous episode during which no disorder-specific symptoms are present.

Multiple episodes, currently in acute episode: Multiple episodes may be determined after a minimum of two episodes (i.e., after a first episode, a remission and a minimum of one relapse).

Multiple episodes, currently in partial remission

Multiple episodes, currently in full remission

Continuous: Symptoms fulfilling the diagnostic symptom criteria of the disorder are remaining for the majority of the illness course, with subthreshold symptom periods being very brief relative to the overall course.

Unspecified

Specify current severity:

Severity is rated by a quantitative assessment of the primary symptoms of psychosis, including delusions, hallucinations, disorganized speech, abnormal psychomotor behavior, and negative symptoms. Each of these symptoms may be rated for its current severity (most severe in the last 7 days) on a 5-point scale ranging from 0 (not present) to 4 (present and severe). (See Clinician-Rated Dimensions of Psychosis Symptom Severity in the chapter "Assessment Measures.")

Note: Diagnosis of schizoaffective disorder can be made without using this severity specifier.

Note: For additional information on Development

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Diagnostic and Clinical Features

According to the DSM-5, a diagnosis of schizoaffective disorder is based on the presence of both criteria A for schizophrenia (i.e., delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior, and negative symptoms) and presence of concurrent symptoms meeting the criteria for a depressive or a manic episode. To distinguish depressive symptoms from negative symptoms of schizophrenia, the DSM-5 criteria for schizoaffective disorder, depressed type emphasize that the A1 criterion of a major depressive episode must be fulfilled (i.e., the presence of "depressed mood" and not simply anhedonia and/or loss of interest).

Determining whether an individual fulfills the DSM-5 criteria for schizoaffective disorder also requires that the clinician take a detailed temporal history of the onset and cessation of mood and psychotic symptoms. To accomplish this, the patient must be able to clearly distinguish when mood and psychotic symptoms started and stopped. Because retrospective reporting is inherently unreliable and both types of symptoms can begin or end insidiously, this is not always a simple task for either the clinician or the patient. Without sufficient information about the length of psychosis and mood symptoms, it is difficult to know whether criterion A (psychotic symptoms of at least 1 month's duration), criterion B (psychotic symptoms in the absence of the mood syndrome), and criterion C (mood symptoms occurring for of the majority of the total duration of the active and residual portions of the illness) are fully met. Such estimates are often difficult to make in practice. In a series of approximately 700 first-admission patients with psychotic features, 6 percent received a diagnosis of schizoaffective disorder at hospital discharge, but only half of these (3 percent of the total) met DSM-III-R criteria for schizoaffective disorder upon review by the research psychiatrists. Two years later, only one third of those initially diagnosed with schizoaffective disorder were again diagnosed with this disorder, one third were diagnosed with schizophrenia, and another third diagnosed with an affective psychosis. Recent phenomenological research has failed to support the validity of the 2-week cutoff for duration of psychotic symptoms in the absence of mood episode for distinguishing schizoaffective disorder from other conditions with concurrent mood and psychotic symptoms.

Table 12.17–7.
ICD-10 Diagnostic Criteria for Schizoaffective Disorders

Note. This diagnosis depends upon an approximate "balance" between the number, severity, and duration of the schizophrenic and affective symptoms.

G1. The disorder meets the criteria for one of the affective disorders of moderate or severe degree, as specified for each category.

G2. Symptoms from at least one of the groups listed below must be clearly present for most of the time during a period of at least 2 weeks (these groups are almost the same as for schizophrenia):

1. Thought echo, thought insertion or withdrawal, thought broadcasting (criterion G1(1)a for paranoid, hebephrenic, or catatonic schizophrenia);
2. Delusions of control, influence, or passivity, clearly referred to body or limb movements or specific thoughts, actions, or sensations (criterion G1(1)b for paranoid, hebephrenic, or catatonic schizophrenia);
3. Hallucinatory voices giving a running commentary on the patient's behavior or discussing the patient among themselves, or other types of hallucinatory voices coming from some part of the body (criterion G1(1)c for paranoid, hebephrenic, or catatonic schizophrenia);
4. Persistent delusions of other kinds that are culturally inappropriate and completely impossible, but not merely grandiose or persecutory (criterion G1(1)d for paranoid, hebephrenic, or catatonic schizophrenia), e.g., has visited other worlds; can control the clouds by breathing in and out; can communicate with plants or animals without speaking;
5. Grossly irrelevant or incoherent speech, or frequent use of neologisms (a marked form of criterion G1(2)b for paranoid, hebephrenic, or catatonic schizophrenia);
6. Intermittent but frequent appearance of some forms of catatonic behavior, such as posturing, waxy flexibility, and negativism (criterion G1(2)c for paranoid, hebephrenic, or catatonic schizophrenia).

G3. Criteria G1 and G2 above must be met within the same episode of the disorder, and concurrently for at least part of the episode. Symptoms from both G1 and G2 must be prominent in the clinical picture.

G4. The disorder is not attributable to organic mental disorder, or to psychoactive substance-related intoxication, dependence, or withdrawal.

Schizoaffective disorder, manic type

A. The general criteria for schizoaffective disorder must be met.
B. Criteria for a manic disorder must be met.

Other schizoaffective disorders

Schizoaffective disorder, unspecified

Comments

If desired, further subtypes of schizoaffective disorder may be specified, according to the longitudinal development of the disorder, as follows:
Concurrent affective and schizophrenic symptoms only
Symptoms as defined in criterion G2 for schizoaffective disorders
Concurrent affective and schizophrenic symptoms beyond the duration of affective symptoms.

From World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization; 1993, with permission.

In clinical practice, the diagnosis of schizoaffective disorder is more commonly made than justified by strict application of DSM-5 criteria. Individuals with a chronic longitudinal course of illness and elements of psychosis and mood disturbance are sometimes erroneously classified as cases of schizoaffective disorder, as are some individuals who have psychosis in the absence of mood symptoms and prominent mood episodes but who do not fulfill criteria for having mood episodes during a significant portion of the illness. Schizoaffective disorder has also been mistakenly viewed as a compromise between the diagnoses of schizophrenia and mood disorder when the exact diagnosis is unclear.

Differential Diagnosis

A critical step in differential diagnosis is determining whether a substance-use disorder or a general medical condition could account for the patient's symptoms. A substance/medication-induced mood disorder or a psychotic disorder or a mood or psychotic disorder due to another medical condition is diagnosed when there is evidence from the history, physical examination, or laboratory tests indicating that the symptoms are the direct physiological consequence of a specific substance or a general medical condition. A wide range of medical etiologies including neurological disorders, infectious diseases, metabolic abnormalities, and endocrinopathies may result in a combination of psychosis and mood change that may resemble schizoaffective disorder. Similarly, prescribed and over-the-counter medications and substances of abuse may produce mood change and psychosis either with chronic use, acute intoxication, or withdrawal. For example, acute or chronic use of alcohol can produce hallucinosis and can also predispose the patient to a depressive state. The chronic use of cannabis may precipitate psychosis and can also lead to an anergic state, which shares many features with depression (amotivational syndrome). Chronic cocaine use may similarly be associated with psychosis as well as changes in mood. Thus, careful screening for general medical conditions and substance-use disorders is an essential aspect in the differential diagnosis of schizoaffective disorder.

In distinguishing between schizoaffective disorder and mood disorder with psychotic features, several elements should be taken into consideration. Many clinicians fail to assess the relative persistence of psychotic and mood symptoms between acute episodes, and others expand the definition of schizoaffective disorder to include chronic psychosis with a few prominent mood symptoms as well as mood disorders with an acute, severely psychotic presentation. Patients often view depression as more understandable or socially acceptable than psychosis and may retrospectively emphasize the duration of depressive symptoms in their illness. It is also difficult to disentangle negative symptoms from depressive symptoms by history alone. Without a clear history of a euphoric mood, retrospectively assessed manic episodes may be viewed as episodes of acute psychosis associated with agitation or irritability. Thus, it is important to take a thorough history using information from as many informants as possible to delineate whether criteria were met for both a mood episode and schizophrenia. Such a temporal history can also be useful in distinguishing a mood disorder, typically depression that is superimposed on schizophrenia. Even with a detailed history, the diagnosis of schizoaffective disorder may only become clear over time.

Course and Prognosis

Although the course and prognosis of schizoaffective disorder may vary, the outcome is generally poorer than for patients with mood disorder but

somewhat better than for patients with schizophrenia. In the Suffolk County, New York, first-admission series, 75 percent of patients with this disorder did not remit fully over the first 2 years of follow-up, compared to 87 percent of patients with schizophrenia and 22 and 40 percent of those with bipolar disorder with psychotic features and major depressive disorder with psychotic features, respectively. Individuals with schizoaffective disorder may also have levels of neurocognitive impairment and deficits in social and occupational functioning that are greater than those of individuals with bipolar disorder. In terms of mortality, one European study showed that individuals with schizoaffective disorder have comparable rates of death from unnatural causes to individuals with bipolar disorder and greater rates of death from unnatural causes than individuals with schizophrenia. Studies have also shown that schizoaffective disorder has low diagnostic stability over time as compared to schizophrenia or mood disorders that impact comparisons of prognosis and outcomes among disorders.

SUBSTANCE/MEDICATION-INDUCED PSYCHOTIC DISORDER

History

Psychotic symptoms in alcoholics and individuals who used opioids or hashish were described by classic psychiatrists. In the DSM-I, these disorders were introduced under the rubrics of acute brain syndrome associated with alcohol intoxication, acute brain syndrome associated with drug or poison intoxication, chronic brain syndrome with psychotic reaction associated with alcohol intoxication, and chronic brains syndrome with psychotic reaction associated with drug or poison intoxication. However, the DSM-I did not distinguish between psychotic symptoms that occur in the context of delirium and those that occur in clear sensorium. These syndromes continued to be classified under organic brain syndromes in DSM-II. However, the DSM-II also introduced the categories of other alcoholic hallucinosis characterized by accusatory or threatening auditory hallucinations in a state of relatively clear consciousness and alcohol paranoid state (alcoholic paranoia) that develops in chronic alcoholics and is characterized by excessive jealousy and delusions of infidelity by the spouse. The DSM-IV and DSM-IV-TR combined the organic hallucinosis and delusional syndromes into a single category of substance-induced psychotic disorders in the section on schizophrenia and other psychotic disorders and introduced specifiers for onset during intoxication and onset during withdrawal. This approach to classification of these disorders was maintained in the DSM-5, although the name of the disorder was changed to substance/medication-induced psychotic disorder.

Diagnostic and Clinical Features

According to the DSM-5, substance/medication-induced psychotic disorder is characterized by the presence of delusions and/or hallucinations that have developed during intoxication or withdrawal from a medication or substance or soon after exposure to it. The substance or medication involved is one that is known to produce psychotic symptoms. Furthermore, the symptoms are not better explained by an independent psychotic disorder, such as schizophrenia. The manual presents some features of the disorder that could be helpful in differential diagnosis of an independent psychotic disorder, including psychotic symptoms that preceded substance/medication use, recurrent episodes of the disorder or symptoms that last for a long period of time (1 month or longer) after withdrawal or intoxication symptoms are resolved, all suggesting an independent psychotic disorder.

Symptoms that occur exclusively during the course of delirium may not be diagnosed as a substance/medication-induced psychotic disorder. Also, when the psychotic symptoms are associated with symptoms of intoxication or withdrawal, a diagnosis of substance/medication-induced psychotic disorder can only be made when the psychotic symptoms predominate the clinical picture and are sufficiently severe to need clinical attention. Specifiers "with onset during intoxication" and "with onset during withdrawal" can be used to signal association with intoxication or withdrawal, respectively, and to indicate that the diagnostic criteria for intoxication or withdrawal are met.

Comparative Nosology

Corresponding diagnoses are included in the ICD-10 in the section of "behavioural disorders due to psychoactive substance use" (Table 12.17–8). A group of psychotic syndromes were described in this section that occurred during or immediately after psychoactive substance use and were characterized by "vivid hallucinations (typically auditory, but often in more than one sensory modality), misidentifications, delusions, and/or ideas of reference." The sensorium was noted to be usually clear. ICD-10 also identified specifiers to further characterize the symptomatic picture of the disorder, such as "schizophrenia-like," "predominantly hallucinatory," and "predominantly polymorphic."

Epidemiology

The prevalence of substance/medication-induced psychotic disorder likely varies according to the distribution of substances and medications in the population. In the Lundby study that examined prevalence of psychiatric disorders in a sample of over 3,500 participants in rural communities in Sweden who were followed over a period of 50 years, the prevalence of substance-induced psychotic disorders was 0.59 percent over the lifetime

of an individual and about six times more common among males than females. In a 2005 study of 400 patients with psychosis in conjunction with substance use presenting to New York City hospital emergency departments, 44 percent were diagnosed as cases of substance-induced psychotic disorder. When these patients were re-evaluated a year later, 25 percent of cases initially diagnosed as substance-induced were rediagnosed as cases of psychotic disorders not related to substances.

The substances involved are likely different in different age groups. Substances of abuse may be a more common cause of substance/medication-induced psychotic disorder in the younger age group; whereas, medications in the context of polypharmacy for multiple health conditions are a more common cause in the older age group.

Etiology

A broad range of substances have been implicated in cases of substance/medication-induced psychotic disorder. These include alcohol, cannabis, hallucinogens, phencyclidine, inhalants, sedative-hypnotics, anxiolytics, cocaine, and other stimulants. Withdrawal from alcohol, sedative-hypnotics, and anxiolytic drugs has also been associated with psychotic symptoms.

Table 12.17–8.
ICD-10 Diagnostic Criteria for Psychotic Disorder due to Psychoactive Substance Use

Mental and behavioral disorders due to psychoactive substance use (F10-F19)

F1*x*.5 Psychotic disorder

A cluster of psychotic phenomena that occur during or immediately after psychoactive substance use and are characterized by vivid hallucinations (typically auditory, but often in more than one sensory modality), misidentifications, delusions and/or ideas of reference (often of a paranoid or persecutory nature), psychomotor disturbances (excitement or stupor), and an abnormal affect, which may range from intense fear to ecstasy. The sensorium is usually clear but some degree of clouding of consciousness, though not severe confusion, may be present. The disorder typically resolves at least partially within 1 month and fully within 6 months.

Diagnostic guidelines

A psychotic disorder occurring during or immediately after drug use (usually within 48 hours) should be recorded here provided that it is not a manifestation of drug withdrawal state with delirium (see F1*x*.4) or of late onset. Late-onset psychotic disorders (with onset more than 2 weeks after substance use) may occur, but should be coded as F1*x*.75. Psychoactive substance-induced psychotic disorders may present with varying patterns of symptoms. These variations will be influenced by the type of substance involved and the personality of the user. For stimulant drugs such as cocaine and amphetamines, drug-induced psychotic disorders are generally closely related to high dose levels and/or prolonged use of the substance.

A diagnosis of a psychotic disorder should not be made merely on the basis of perceptual distortions or hallucinatory experiences when substances having primary hallucinogenic effects (e.g., lysergide [LSD], mescaline, and cannabis at high doses) have been taken. In such cases, and also for confusional states, a possible diagnosis of acute intoxication (F1*x*.0) should be considered. Particular care should also be taken to avoid mistakenly diagnosing a more serious condition (e.g.,

schizophrenia) when a diagnosis of psychoactive substance-induced psychosis is appropriate. Many psychoactive substance-induced psychotic states are of short duration provided that no further amounts of the drug are taken (as in the case of amphetamine and cocaine psychoses). False diagnosis in such cases may have distressing and costly implications for the patient and for the health services.

Includes: Alcoholic hallucinosis

Alcoholic jealousy

Alcoholic paranoia

Alcoholic psychosis NOS

Differential diagnosis. Consider the possibility of another mental disorder being aggravated or precipitated by psychoactive substance use (e.g., schizophrenia (F20.-); mood [affective] disorder (F30-F39); paranoid or schizoid personality disorder (F60.0, F60.1)). In such cases, a diagnosis of psychoactive substance-induced psychotic state may be inappropriate.

The diagnosis of psychotic state may be further specified by the following five-character codes:

F1*x*.50 Schizophrenia-like

F1*x*.51 Predominantly delusional

F1*x*.52 Predominantly hallucinatory

(includes alcoholic hallucinosis)

F1*x*.53 Predominantly polymorphic

F1*x*.54 Predominantly depressive symptoms

F1*x*.55 Predominantly manic symptoms

F1*x*.56 Mixed

From World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization; 1993, with permission.

A number of common medication groups are associated with psychotic symptoms including nonsteroidal anti-inflammatory medications, other over-the-counter medications (e.g., phenylephrine and pseudoephedrine), corticosteroids, gastrointestinal medications, muscle relaxants antihistamines, antihypertensive and cardiovascular medications, antimicrobial medications, and chemotherapeutic agents (e.g., cyclosporine and procarbazine). The list includes medications commonly used in psychiatry and neurology, anticonvulsants, antiparkinsonian medications, and disulfiram. Additionally, some toxins commonly used in agriculture or industry have been reported to induce psychotic symptoms including organophosphate insecticides, carbon monoxide, carbon dioxide, and volatile substances such as fuel or paint.

Diagnostic and Clinical Features

Prominent delusions and/or hallucinations that are judged to be due to the physiological effects of a substance/medication are the core features of the diagnosis of substance/medication-induced psychotic disorder. The DSM-5 notes that "Hallucinations that the individual realizes are substance/medication-induced are not included here and instead would be diagnosed as substance intoxication or substance withdrawal with the accompanying specifier 'with perceptual disturbances.'" The manual also

notes that the psychotic symptoms are required to have developed during or soon after substance intoxication or withdrawal, as identified by specifiers "with onset during intoxication" and "with onset during withdrawal." The diagnosis is recorded along with the name of offending substance, and, in the case of substances of abuse, presence of a substance disorder if appropriate (e.g., cocaine-induced psychotic disorder with onset during intoxication, with severe cocaine use disorders). The DSM-5 also notes some differences in coding of diagnoses according to ICD-9-CM and ICD-10-CM. The reader is referred to the text of the criteria for a detailed description of these differences.

Differential Diagnosis

A major differential diagnosis is between substance/medication-induced psychosis and an independent psychotic disorder that is not substance induced. The DSM-5 provides a few guidelines for differential diagnosis of nonsubstance/medication-induced psychotic disorders, including temporal order of symptom onset, course of psychotic symptoms and past history. If the psychotic symptoms preceded the onset of the substance/medication use or the symptoms persist for a substantial period of time (e.g., about 1 month) after the acute withdrawal or severe intoxication, then the psychotic symptoms are more likely to be related to a primary psychotic disorder. A history of a recurrent primary psychotic disorder also argues against substance/medication-induced psychosis. Nevertheless, information about the temporal order of symptoms and substance use and past history are not always clear. Furthermore, a past history of psychotic disorder unrelated to substances does not rule out a current diagnosis of substance/medication-induced psychotic disorder.

Specific features of the disorder, such as atypical age of onset (e.g., first onset of delusions in a person older than 35 years without a known history of a primary psychotic disorder) or the nature of symptoms, may provide some guidance in differentiating substance/medication-induced psychotic disorder from primary psychotic disorders. Certain types of psychotic symptoms and accompanying symptoms are more common with certain substances. For example, the tactile hallucination of bugs crawling under the skin (formication) is more common in cocaine-induced psychosis; whereas, cannabis-induced psychotic disorder that typically develops shortly after high-dose cannabis use, usually involves persecutory delusions, anxiety, emotional lability, and depersonalization experiences. Alcohol-induced psychotic disorder with hallucinations often presents with auditory hallucinations after prolonged and heavy use of alcohol. Evidence of substance use from the history and physical examination, or laboratory tests for substances that can be detected using such tests may be helpful in solidifying a diagnosis.

Assessment of cognitive and mood symptoms is vital for differential

diagnosis from delirium and mood disorders. Hallucinations are relatively common in the context of delirium associated with substance intoxication or withdrawal and do not justify a separate diagnosis of substance/medication-induced psychotic disorder. Similarly, delusions in the context of dementia or mild neurocognitive disorder should be coded as major or mild neurocognitive disorder, with behavioral disturbance.

The DSM-5 clearly states that if the individual recognizes that the hallucinations are the result of substance or medication use and does not act on them, then these experiences should not count toward a diagnosis of substance/medication-induced psychosis. Some individuals who are intoxicated with stimulants or hallucinogens, meperidine or phencyclidine, or withdrawing from alcohol or sedatives may experience altered perceptions that they recognize as such. In these cases, the appropriate diagnoses are substance intoxication or substance withdrawal with perceptual disturbances.

Prolonged use of hallucinogens is sometimes associated with recurrent hallucinations long after stopping the use of these drugs. Reality testing remains intact in these cases and the person is aware that these experiences are related to drug effects. Such "flashbacks" should be diagnosed as hallucinogen persisting perception disorder.

A difficult differential diagnosis is to separate the effects of medications from the effects of underlying medical or mental disorder for which the medication is prescribed. The clinical history, course of illness, and the nature and course of treatments may provide clues to this differential diagnosis. At times medication substitution or discontinuation may be the only way to definitively establish a causal link between medication use and psychosis. If the disturbance is deemed to be related to both a medical condition and substance/medication use, both diagnoses should be made (e.g., psychotic disorder due to systemic lupus erythematosus and steroid-induced psychotic disorder).

Course and Prognosis

The course and prognosis of substance/medication-induced psychotic disorder is generally favorable and psychosis typically remits after removal of the offending agent. However, there have been reports of continuing psychotic symptoms for weeks or months after stopping stimulants, cocaine or phencyclidine. Treatment with antipsychotic medications may be needed in cases of severe substance/medication-induced psychotic disorder associated with agitation, or suicidal/homicidal ideation or other life-threatening behavior. Hospitalization may be needed for safety of the patient or others in some of these cases.

PSYCHOTIC DISORDER DUE TO ANOTHER MEDICAL CONDITION

History

An association between mental and physical illness has been noted since antiquity. The concept of "sympathetic insanity" that became popular in the 19th century hypothesized an association in normal and abnormal functioning between all body organs that leads to severe mental illness among individuals with sicknesses in other organs. In DSM-I, these disorders were introduced under the rubric of "acute brain disorders" that included acute brain disorders associated with infections, trauma, circulatory disturbances, metabolic disturbances, convulsive disorder, and neoplasms. However, the DSM-I did not distinguish between psychotic symptoms that occur in the context of delirium and those that occur in conjunction with clear sensorium. The DSM-III introduced two categories of "organic delusional disorder" and "organic hallucinosis" that presented with psychotic symptoms in the absence of clouding of consciousness, as in delirium. The DSM-IV and DSM-IV-TR combined the two categories into one category of "psychotic disorder due to…[the general medical condition] and introduced specifiers: "with delusions" and "with hallucinations" to identify the prominent symptoms. This approach to classification of these disorders was maintained in the DSM-5.

Comparative Nosology

Corresponding diagnoses are included in the ICD-10 under the overall category of "other mental disorders due to brain damage and dysfunction and to physical disease." This overall category includes "organic hallucinosis," characterized by persistent or recurrent hallucinations that occur in clear consciousness and may or may not be recognized by the subject as such, and "organic delusional [schizophrenia-like] disorder," characterized by recurrent delusions in clear consciousness (Table 12.17–9).

Epidemiology

There are few studies of the population prevalence of "psychotic disorder due to another medical condition." The prevalence of this disorder likely varies across different communities and over time corresponding to the prevalence of the common underlying disorders. In the Lundby study that examined the prevalence of psychiatric disorders in a sample of over 3500 participants in rural communities in Sweden who were followed over a period of 50 years, the prevalence of psychotic disorders related to a general medical condition was 0.76 percent over the lifetime and did not vary across genders. The prevalence is likely higher in the old age group, reflecting the prevalence of the underlying disorders. The nature of medical disorders giving rise to psychotic symptoms may also change across the lifespan, with epilepsy, head trauma and autoimmune diseases being more

common causes earlier in life and stroke, anoxic events, and cardiovascular diseases more common later in life.

Table 12.17–9.
ICD-10 Diagnostic Criteria for Organic Hallucinosis and Organic Delusional Disorder

F06.0 Organic hallucinosis

A disorder of persistent or recurrent hallucinations, usually visual or auditory, that occur in clear consciousness and may or may not be recognized by the subject as such. Delusional elaboration of the hallucinations may occur, but insight is not infrequently preserved.

Diagnostic guidelines

In addition to the general criteria in the introduction to F06 above, there should be evidence of persistent or recurrent hallucinations in any modality; no clouding of consciousness; no significant intellectual decline; no predominant disturbance of mood; and no predominance of delusions.

Includes: Dermatozoenwahn

Organic hallucinatory state (nonalcoholic)

Excludes: Alcoholic hallucinosis (F10.52)

Schizophrenia (F20.-)

F06.2 Organic delusional [schizophrenia-like] disorder

A disorder in which persistent or recurrent delusions dominate the clinical picture. The delusions may be accompanied by hallucinations but are not confined to their content. Features suggestive of schizophrenia, such as bizarre delusions, hallucinations, or thought disorder, may also be present.

Diagnostic guidelines

The general criteria for assuming an organic etiology, laid down in the introduction to F06, must be met. In addition, there should be delusions (persecutory, of bodily change, jealousy, disease, or death of the subject or another person). Hallucinations, thought disorder, or isolated catatonic phenomena may be present. Consciousness and memory must not be affected. This diagnosis should not be made if the presumed evidence of organic causation is nonspecific or limited to findings such as enlarged cerebral ventricles (visualized on computerized axial tomography) or "soft" neurological signs.

Includes: Paranoid and paranoid-hallucinatory organic states

Schizophrenia-like psychosis in epilepsy

Excludes: Acute and transient psychotic disorders (F23.-)

Drug-induced psychotic disorders (F1*x*.5)

Persistent delusional disorder (F22.-)

Schizophrenia (F20.-)

ICD-10, *International Statistical Classification of Diseases and Related Health Problems,* 10th revision.

From World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization; 1993, with permission.

Etiology

The list of medical conditions associated with psychotic symptoms is quite long, including neurological conditions (e.g., head trauma, neoplasms, cerebrovascular disease, Huntington disease, multiple sclerosis, epilepsy, visual nerve injury, deafness, migraine, and CNS infections), endocrine disorders (e.g., hyper- and hypothyroidism, hyper- and

hypoparathyroidism, hyper- and hypoadrenocorticism), metabolic conditions (e.g., hypoxia, hypercarbia, and hypoglycemia), fluid or electrolyte imbalances, hepatic insufficiency, kidney failure, and autoimmune disorders with CNS involvement (e.g., systemic lupus erythematosus and anti-NMDA receptor antibody encephalitis). Among these conditions, endocrine, metabolic, and autoimmune disorders are the most common contributors to psychotic disorder due to another medical condition. Psychotic symptoms are also not infrequent in epilepsy observed in 2 to 7 percent of these patients. Psychotic symptoms are specifically associated with postictal phase of temporal lobe epilepsy. Table 12.17–10 provides a partial listing of medical conditions associated with psychotic symptoms.

Diagnostic and Clinical Features

The essential feature of psychotic disorder due to another medical condition is the existence of prominent delusions or hallucinations that are judged to be directly caused by another medical condition evidenced by clinical history, physical examination findings, or laboratory data including neuropsychological testing and imaging studies. The psychotic symptoms are not better explained by another mental disorder, such as a mood disorder with psychotic features or schizophrenia and do not occur only in the course of a delirium. The DSM-5 criteria also require that the psychotic symptoms be associated with significant distress or impairment in functioning. The DSM-5 allows the diagnostician to identify the predominant psychotic symptoms by using specifiers "with delusions" or "with hallucinations" and to specify the severity of psychotic symptoms.

Although hallucinations in psychotic disorder due to another medical condition can occur in any sensory modality, certain types of hallucinations are suggestive of specific etiologies. For example, olfactory hallucinations are suggestive of temporal lobe epilepsy. Furthermore, visual hallucinations and hallucinations in more than one sensory modality may be more common in psychotic disorder due to another medical condition than in schizophrenia. Psychotic disorder due to another medical condition is generally not diagnosed if the individual maintains reality testing for the hallucinations and recognizes that they result from the medical condition. Delusions are most commonly persecutory. The association between the content of delusions and particular medical conditions appears to be less specific than is the case for hallucinations.

Psychotic disorder due to another medical condition is coded in DSM-5 by recording the name of the medical condition (e.g., psychotic disorder due to malignant lung neoplasm, with delusions).

Differential Diagnosis

Psychotic disorder due to another medical condition must be distinguished from delirium, which is common in many acute medical conditions and may be associated with hallucinations and delusions. A separate diagnosis of psychotic disorder due of another medical condition is not justified if the psychotic symptoms only present in the context of delirium. Delusions in the context of a neurocognitive disorder (e.g., dementia) would be diagnosed as neurocognitive disorder, with behavioral disturbance.

Table 12.17–10.
A Partial List of Medical Disorders That May Cause a Psychotic Disorder due to Another Medical Condition

Neurodegenerative disorders	Huntington disease, basal ganglia calcification, multiple sclerosis, metachromatic leukodystrophy
Other CNS disorders	Brain tumors, especially temporal lobe and deep hemispheric tumors; epilepsy, especially complex partial seizure disorder; head trauma, anoxic brain injury, migraine
Vascular disease	Atherosclerotic vascular disease, especially when associated with diffuse, temporoparietal, or subcortical lesions; hypertensive encephalopathy; subarachnoid hemorrhage; subdural hematoma; fat embolism
Infectious disease	HIV/AIDS, encephalitis lethargica, Creutzfeldt–Jakob disease, syphilis, malaria, acute viral encephalitis
Metabolic disorder	Hypercalcemia, hyponatremia, hypoglycemia, uremia, hepatic encephalopathy, pancreatic encephalopathy, porphyria
Endocrinopathies	Addison disease, Cushing disease, hyper- or hypothyroidism, hyper or hypoparathyroidism, panhypopituitarism
Vitamin deficiencies	Vitamin B_{12} deficiency, folate deficiency, thiamine deficiency, niacin deficiency
Autoimmune diseases	Systemic lupus erythematosus, temporal arteritis, anti-NMDA receptor antibody encephalitis

CNS, central nervous system; HIV, human immunodeficiency virus; AIDS, acquired immunodeficiency syndrome; NMDA, *N*-methyl-D-aspartate.

Severe medical conditions, especially if they lead to serious disability or prolonged hospitalization, can be extremely stressful and may be associated with a psychologically mediated brief psychotic disorder with marked stressor. Distinguishing between these diagnoses is especially difficult in the absence of other information regarding past history and the course of medical illness and psychosis.

Many patients with medical conditions also use medications. If there is evidence of recent (within 4 weeks) use of medications or substances of abuse, a substance/medication-induced psychotic disorder should be considered as part of differential diagnosis. If the disturbance is judged to be due to both a medical condition and substance use, both diagnoses (i.e., psychotic disorder due to another medical condition and substance/medication-induced psychotic disorder) can be given.

Psychotic disorder due to another medical condition must be

distinguished from other psychotic disorders including schizophrenia, delusional disorder, schizoaffective disorder, and/or mood disorders with psychotic features. The presence of features such as later age at onset or presence of visual or olfactory hallucinations may suggest psychotic disorder due to another medical condition or a substance/medication-induced psychotic disorder.

Course and Outcome

The course and outcome of psychotic disorder due to another medical condition depend on the course and outcome of the underlying medical illness. With appropriate management of the underlying medical disorder, the psychotic symptoms generally remit. Most cases are transient and very rarely do the psychotic symptoms continue after remission of the underlying medical disease.

CATATONIA

History

Catatonia was originally described by the German psychiatrist Karl Kahlbaum in 1870s as a distinct mental disorder with psychiatric, motor, and behavioral symptoms. Although, originally, Kahlbaum considered catatonia to have an overall benign course, reports of chronic catatonia led later psychiatrists, including Kraepelin, to reconsider and view catatonia as a chronic disorder similar to hebephrenia. This view of catatonia contributed to Kraepelin's proposal that catatonia be viewed as a subtype of dementia praecox. While Kraepelin recognized that catatonic symptoms could occur in mood disease and organic disorders, he distinguished between catatonic dementia praecox and catatonic symptoms occurring in other disorders.

This view of catatonia as a distinct subtype of schizophrenia dominated modern classification systems including the DSM and ICD. Multiple editions of DSM through DSM-IV-TR and ICD-10 discussed catatonia as a subtype of schizophrenia. DSM-III-R introduced the catatonic features specifier for mood disorders when the clinical picture of a mood disorder episode is dominated by two or more catatonic symptoms. The category of catatonic disorder due to a general medical condition was introduced in DSM-IV. Furthermore, the ICD-10 recognized a category of organic catatonic disorder in addition to catatonic schizophrenia.

In a clear break from the Kraepelinian conception of catatonia as subtype of schizophrenia, DSM-5 introduced a separate rubric for catatonia. This change was motivated by evidence regarding the occurrence of catatonia in a number of other psychiatric and medical disorders besides schizophrenia and mood disorders and evidence of under-recognition of catatonia in current psychiatric practice. The manual recognized three

types of catatonia: (1) catatonia associated with another mental disorder, (2) catatonic disorder due to another medical condition, and (3) unspecified catatonia. Although DSM-5 did not recognize catatonia as an "independent class," many scholars view catatonia as a discrete entity that is distinct from the psychotic disorders.

Epidemiology

Epidemiologic research on prevalence, incidence, and correlates of catatonia has been hampered by differences in definition and method of assessment of catatonia. A review of 20th century studies found that catatonia was reported in 7 to 17 percent of patients hospitalized with acute psychotic disorders and 13 to 31 percent of patients with mood disorders. There is some evidence that the prevalence of catatonic type of schizophrenia may have declined over the 20th century and that catatonia may be more common in developing countries than industrialized countries. Catatonia may be seen in individuals of all ages, including children and adolescents. Gender and age distributions of patients with catatonia likely reflect the epidemiology of the underlying disorder.

Etiology

The etiology of catatonia is unknown. The long and growing list of medical conditions associated with catatonia underscores the potential heterogeneity in the etiology of catatonia. The list includes major neurologic disorders such as epilepsy, encephalitis and trauma, endocrine disorders including thyroid disease, metabolic disturbances such as uremia and hypercalcemia, paraneoplastic disorders, folate deficiency, and autoimmune diseases. Catatonia can also occur in the context of a number of mental disorders, including psychotic and mood disorders and neurodevelopmental disorders including autism spectrum disorder.

It is plausible that these different conditions are the distal causes that work through a common mechanism to produce catatonic symptoms. The favorable response of the disorder to GABAergic treatments (e.g., benzodiazepines) points to a role of GABA system in catatonia. There is also a growing interest in the role of the NMDA receptors in catatonia based on numerous case reports associating anti-NMDA receptor antibody encephalitis with catatonic symptoms.

Diagnostic and Clinical Features

The essential feature of catatonia is a psychomotor disturbance that may involve decreased engagement with the interviewer during clinical interview or physical examination, decreased motor activity, or excessive and peculiar motor activity. The psychomotor disturbance may range from marked unresponsiveness to environmental stimuli (stupor) to marked

agitation. In some cases, the individual's symptoms may quickly switch between decreased motor activity and excessive motor activity, making the diagnosis difficult. Some of the peculiar motor and verbal symptoms such as catalepsy (holding a posture induced by the examiner), posturing (active adoption of often awkward postures by the patient), waxy flexibility (slight resistance when the examiner tries to position the patient's body parts), mannerism (odd or caricature-like mimicking of normal actions), grimacing, echolalia, and echopraxia (mimicking another person's speech or behavior) are more pathognomonic and helpful for the diagnosis of a catatonic syndrome. During severe stages of immobility or increased activity, the individual may need careful supervision to prevent dehydration, malnutrition, exhaustion, hyperpyrexia, aspiration, embolic events, decubiti, or contractures of extremities. Self-inflicted injuries and harm to others can also occur, particularly during periods of increased activity.

The DSM-5 diagnosis of catatonia requires that the clinical picture be dominated by at least 3 of the 12 characteristic symptoms including stupor, catalepsy, waxy flexibility, mutism, negativism, posturing, mannerism, stereotypy, agitation, grimacing, echolalia, and echopraxia. The diagnosis of catatonia associated with another medical condition also requires evidence from clinical history, physical examination, or laboratory examinations findings (including imaging studies) that the catatonic symptoms are caused by another medical condition, that they are not better explained by another mental disorder and do not occur exclusively during the course of delirium, and that the symptoms cause significant distress or impairment in social or occupational functioning.

In coding catatonia associated with another mental disorders or medical disorder, the name of the associated mental or medical disorder is recorded first, followed by the specifiers "catatonia associated with…" in the case of mental disorders (e.g., "major depressive disorder; catatonia associated with major depressive disorder") or "catatonic disorder due to…" in the case of medical conditions (e.g., "hypothyroidism; catatonic disorder due to hypothyroidism").

The residual category of unspecified catatonia is reserved for situations in which the catatonic symptoms are present and cause clinically significant distress or impairment. However, either the number of symptoms does not reach the minimum required of three symptoms, the underlying medical or mental disorder is not clear or there is insufficient information regarding the underlying cause (e.g., in the emergency department settings).

Differential Diagnosis

Movement disturbances are common during the course of a delirium or neuroleptic malignant syndrome and a separate diagnosis of catatonic

disorder due to another medical condition is not justified in these cases. Catatonia-like symptoms are not uncommon among patients treated with antipsychotic medication (e.g., neuroleptic-induced acute dystonia). History of recent exposure to medications, especially first-generation antipsychotic medications, may help in distinguishing these diagnostic possibilities. Generalized rigidity is a cardinal feature of neuroleptic malignant syndrome, in which significant increases in creatine kinase levels are common. Autonomic instability and increases in temperature can be seen in individuals with catatonia, but these findings are typically more severe in neuroleptic malignant syndrome and can be potentially life-threatening.

Course and Prognosis

Catatonic symptoms sometimes remit with the management of the underlying psychiatric or medical condition. However, many individuals will also require specific treatment of the catatonic syndrome. Although a rare type of remitting and relapsing catatonia, known as "periodic catatonia" has been described in classic European literature, the majority of cases of catatonia do not have a recurring course. Catatonic symptoms are especially responsive to treatment with high doses of benzodiazepines and ECT, which may be indicated as the main or adjunctive treatment in the acute management of psychiatric patients presenting with severe catatonic symptoms. Use of antipsychotic medications, particularly high-potency first-generation antipsychotic medications, in individuals with catatonia may worsen outcomes.

OTHER SPECIFIED SCHIZOPHRENIA SPECTRUM AND OTHER PSYCHOTIC DISORDER AND UNSPECIFIED SCHIZOPHRENIA SPECTRUM AND OTHER PSYCHOTIC DISORDER

History

The concept of "unspecified psychosis" was introduced in DSM-II but was specifically noted to be "for librarians and statisticians to use in coding incomplete diagnoses" rather than for use by clinicians. However, in DSM-III, the diagnosis of "atypical psychosis" was introduced as "a residual category for cases in which there are psychotic symptoms ... that do not meet the criteria for any specific mental disorder." DSM-III-R changed the name of the diagnosis to psychotic disorder not otherwise specified (atypical psychosis) and eliminated mention of the diagnosis as a "residual category." Instead, it emphasized the use of the diagnosis for "Disorders in which there are psychotic symptoms (delusions, hallucinations, incoherence, marked loosening of associations, catatonic excitement or stupor, or grossly disorganized behavior) that do not meet the criteria for any other nonorganic psychotic disorder." It also noted that the diagnosis

of psychotic disorder not otherwise specified was preferable to deferring a diagnosis in individuals for whom there was inadequate information to make a specific diagnosis.

The concept of psychotic disorder NOS persisted with the development of DSM-IV, although the references to atypical psychosis and nonorganic psychotic disorders were dropped. As in prior editions of DSM, there were no specific diagnostic criteria for this category. However, it was now defined as including psychotic symptoms "about which there is inadequate information to make a specific diagnosis or about which there is contradictory information, or disorders with psychotic symptoms that do not meet the criteria for any specific Psychotic Disorder."

The DSM-5 divided the DSM-IV category of psychotic disorder NOS into two distinct categories of "other specified schizophrenia spectrum and other psychotic disorder " and "unspecified schizophrenia spectrum and other psychotic disorder" to underscore the reason why the disorder could not be classified as one of the specific schizophrenia spectrum disorders. While in the other specified schizophrenia spectrum and other psychotic disorder the characteristics of the condition clearly do not meet the criteria for a specific disorder, in the unspecified schizophrenia spectrum and other psychotic disorder there is inadequate information to make a more specific diagnosis. The latter diagnosis may thus be appropriate for acute care and emergency department settings where there is insufficient information available.

DSM-5 provided four examples of other specified schizophrenia spectrum and other psychotic disorder: (1) persistent auditory hallucinations, (2) delusions with significant overlapping mood episodes, (3) attenuated psychosis syndrome, and (4) delusional symptoms in partner of individual with delusional disorder. The last example, variously called "shared delusional disorder," "shared psychotic disorder," and "induced psychosis," and "folie à deux," was discussed as a specific psychotic disorder in the previous edition of DSM and in ICD-10. The new category of attenuated psychosis syndrome is a new category in the DSM and its introduction was associated with some controversy. These two categories are discussed in some detail later in the chapter.

Comparative Nosology

ICD-10 includes a somewhat related diagnostic category: "F 23.8 Other Acute and Transient Psychotic Disorders." This category is used for "Any other acute psychotic disorders that are not classifiable under any other category" in the section of acute and transient psychotic disorder (see Table 12.17–2). However, like other acute and transient psychotic disorders, these psychotic conditions are characterized by acute onset and typically brief duration and in this regard differ from the DSM-5 categories of "other specified" and "unspecified schizophrenia spectrum and other psychotic

disorder," which have no associated mode of onset or time-related requirements.

Epidemiology

There is minimal information about the prevalence of other specified schizophrenia spectrum and other psychotic disorder and unspecified schizophrenia spectrum and other psychotic disorder in the community. In clinical practice, the psychotic disorder not otherwise specified (NOS) diagnosis has been often used in first admission or emergency psychiatric patients when there is inadequate information to make a diagnosis.

Etiology

Given the heterogeneity of the other specified and unspecified schizophrenia spectrum and other psychotic disorder diagnoses, it is not surprising that there has been minimal investigation into etiologies of these disorders. However, one family study has examined the occurrence of other psychotic illnesses in relatives of probands with psychotic disorder NOS according to prior editions of the DSM and found a substantially elevated risk of schizophrenia and schizophrenia spectrum disorders in these family members but not mood disorders.

Diagnostic and Clinical Features

The diagnostic category of other specified schizophrenia spectrum and other psychotic disorder includes cases with psychotic symptoms and clinically significant distress or impairment that do not conform to diagnostic criteria for specific schizophrenia spectrum and other psychotic disorders. The specific presentation can be recorded as well. For example, for a patient who present to the hospital with persistent auditory hallucinations that are not accompanied by mood disturbance and that are not consistent with symptom criteria for schizophrenia or any other DSM-5 category a diagnosis of "other specified schizophrenia spectrum and other psychotic disorder, persistent auditory hallucinations" may be recorded. Other examples include psychotic symptoms that are below the threshold for a full psychotic episode (attenuated psychosis syndrome), delusional beliefs in partners of individuals with a primary delusional disorder, and persistent delusions with periods of overlapping mood symptoms that have been present for a substantial portion of the delusional disturbance.

A diagnosis of unspecified schizophrenia spectrum and other psychotic disorder may be appropriate for a patient presenting with psychotic symptoms causing clinically significant distress or impairment but where information regarding symptoms and history is not sufficient to make a diagnosis of a specific schizophrenia spectrum and other psychotic disorder. For example, the clinician may not be able to determine whether

the psychosis is primary, due to a general medical condition, or the result of a substance or medication use. Such distinctions can be difficult to make, especially when adequate clinical history is not available (e.g., in the emergency department settings).

Differential Diagnosis

Because other specified and unspecified schizophrenia spectrum and other psychotic disorder are diagnoses of exclusion, the differential diagnosis includes other possible diagnoses with psychotic symptoms. Co-occurring substance use is one of the most frequent confounding factors in establishing a specific diagnosis, since substance use can be found in more than half of hospitalized patients with psychotic symptoms. The possibility that a general medical condition may be causing the psychotic symptoms must always be kept in mind.

If mood symptoms are present, the differential diagnosis will include major depressive disorder with psychotic features or bipolar disorder with psychotic features, depending on the type of mood change that is present. Schizoaffective disorder is also possible if the temporal requirements for the psychotic symptoms and mood changes are met.

If no mood symptoms are present and substance, medication, and medical disease can be ruled out as causes of psychotic symptoms, the differential diagnosis will include schizophrenia, schizophreniform disorder, brief psychotic disorder, and delusional disorder depending on the duration and type of psychotic symptoms exhibited.

Course and Prognosis

The course and prognosis are highly variable as other specified and unspecified schizophrenia spectrum and other psychotic disorder are heterogeneous. Furthermore, the diagnosis of unspecified schizophrenia spectrum and other psychotic disorder is often a provisional diagnosis by design that would be changed upon gathering more information. In a first-admission study in Suffolk County, New York, 12 percent of individuals were diagnosed by the research team with psychotic disorder NOS at the time of first admission, but only 4.8 percent received this diagnosis by the 6-month follow-up, and by 24-month follow-up, the rate was 3.9 percent. Of 25 individuals with a diagnosis of psychotic disorder NOS at 6-month follow-up, only 11 maintained this diagnosis by 24-month follow-up, a stability rate of only 44 percent. In addition, 8 of the 25 subjects with psychotic disorder NOS had met the criteria for schizophrenia or schizoaffective disorder.

ATTENUATED PSYCHOSIS SYNDROME

The introduction of the new category of attenuated psychosis syndrome in

the DSM-5 was subject to some controversy. While prodromal symptoms of schizophrenia have long been recognized by psychiatric investigators, earlier classifications did not include a specific category for these symptoms. Introduction of these symptoms in the diagnostic system was motivated by efforts over the past two decades to improve early detection of schizophrenia and to implement early treatment with the aim of reducing chronicity and long-term disability often associated with schizophrenia. In recognition of this aim, the category was initially labeled the "psychosis risk syndrome." The name change to attenuated psychosis syndrome was intended to highlight the importance of current symptoms and the associated distress and disability as a focus of clinical attention rather than the future risk of a full-blown psychotic disorder.

The DSM-5 provides few guidelines and no discrete criteria for diagnosis of attenuated psychosis syndrome. The definition provided specifies "psychotic-like symptoms that are below a threshold for full psychosis." The manual mentions symptoms that are of lower severity are more transient or are associated with relatively intact insight as examples of attenuated psychosis syndrome. However, no cutoffs for severity or duration of symptoms are provided in the definition of attenuated psychosis syndrome. Like other conditions included under other specified schizophrenia spectrum and other psychotic disorders, attenuated psychosis syndrome is associated with "clinically significant distress or impairment in social, occupational, or other important areas of functioning."

The prevalence of attenuated psychosis syndrome in the general population is unknown. Psychotic-like symptoms are reported by between 4 and 8 percent of the general population. However, in many cases these experiences are not associated with distress or disability, as required by the attenuated psychosis syndrome definition. In a study of adolescents aged 11 to 13 years old, approximately 6.9 percent reported psychotic-like symptoms associated with some degree of distress. These symptoms were more common in males than females.

Etiology of attenuated psychosis syndrome remains unknown. Genetic risk, such as history of psychotic disorders in first-degree relatives, is a criterion in some alternative definitions of attenuated psychosis risk. Structural imaging studies of attenuated psychosis syndrome cases have produced mixed results. Some functional imaging studies have suggested abnormalities in dopamine and glutamate systems.

The future risk of a full-blown psychotic disorder among individuals with attenuated psychosis syndrome has been a major focus of research. The risk of transition varies in different studies and has declined in more recent studies. In one meta-analysis of short-term studies, less than 40 percent switched to a full-blown psychotic illness in a 3-year period. However, the risk of transition remained elevated beyond 3 years and was

as high as 65 to 79 percent in a long-term study. Individuals with attenuated psychosis syndrome may also be at increased risk of other adverse psychological and social outcomes, including suicidal behavior.

There are few guidelines regarding the management of attenuated psychosis. Given that the majority of individuals with attenuated psychosis syndrome would not experience a full-blown psychotic illness in near future, treatment with antipsychotic medication may not be indicated for self-limited symptoms. In many cases, careful assessment and frequent monitoring may be appropriate.

DELUSIONAL SYMPTOMS IN PARTNER OF INDIVIDUAL WITH DELUSIONAL DISORDER

The category of "delusional symptoms in partner of individuals with delusional disorder" was first described in 1860 by Jules Baillarger, who called the syndrome "folie à communiquée," and again described in 1877 by Charles Lasègue and Jean Pierre Falret who gave the syndrome the commonly used name, "folie à deux." Over the years, many other names including shared psychotic disorder, shared delusional disorder, and induced psychotic disorder have been used to describe this syndrome in which two individuals with a close and generally long-term relationship share the same delusional belief. Shared paranoid disorder was introduced in DSM-III in the section on paranoid disorders and the persecutory nature of the shared delusions was emphasized. In DSM-III-R and later, DSM-IV and DSM-IV-TR, the content of the shared delusion was broadened to include all types of delusional beliefs. DSM-5 designated this condition to the category of other specified schizophrenia spectrum and other psychotic disorder. Due to the rarity of cases that present at clinical settings, the prevalence of delusional symptoms in partner of individual with delusional disorder is unknown.

Etiological theories regarding delusional symptoms in partner of individual with delusional disorder are primarily based on family and interpersonal dynamics. In general, individuals who develop a shared delusion will tend to have had a submissive role in the relationship with the partner who was first to develop a delusional disorder. The person who develops the shared delusional symptoms may be particularly impressionable due to personality factors or intellectual limitations. Social isolation, poverty, language difficulties, and geographical isolation may also contribute to the development of the condition. In some cases, the shared delusion may involve more than two individuals and even the entire family unit. Persecutory delusional beliefs are most commonly seen in shared psychotic disorder, comprising about 70 percent of patients in one study; however, religious, grandiose, and somatic delusions may also be observed. Because of the relative rarity of shared delusional disorder, its clinical course and prognosis are unclear, although some reports of successful

treatment are present in the literature.

References

Addington J, Chaves A, Addington D. Diagnostic stability over one year in first-episode psychosis. *Schizophr Res*. 2006;86:71.

Bogren M, Mattisson C, Isberg PE, Nettelbladt P. How common are psychotic and bipolar disorders? A 50-year follow-up of the Lundby population. *Nord J Psychiatry*. 2009;63:336.

Bromet EJ, Kotov R, Fochtmann LJ, et al. Diagnostic shifts during the decade following first admission for psychosis. *Am J Psychiatry*. 2011;168:1186.

Caton CL, Drake RE, Hasin DS, et al. Differences between early-phase primary psychotic disorders with concurrent substance use and substance-induced psychoses. *Arch Gen Psychiatry*. 2005;62:137.

*Cheniaux E, Landeira-Fernandez J, Telles LL, Lessa JL, Dias A. Does schizoaffective disorder really exist? A systematic review of the studies that compared schizoaffective disorder with schizophrenia or mood disorders. *J Affect Disord*. 2008;106:209.

Coryell W. Schizoaffective and schizophreniform disorders. In: Fatemi SH, Clayton PJ, eds. *The Medical Basis of Psychiatry*. New York, NY: Springer; 2008:109–123.

Fennig S, Craig TJ, Bromet EJ. The consistency of DSM-III-R delusional disorder in a first-admission sample. *Psychopathology*. 1996;29:315.

Fink M. Rediscovering catatonia: the biography of a treatable syndrome. *Acta Psychiatr Scand*. 2013;127(Suppl 441):1.

Fiorentini A, Volonteri LS, Dragogna F, et al. Substance-induced psychoses: a critical review of the literature. *Curr Drug Abuse Rev*. 2011;4:228.

Frances A. Catatonia: diagnosis, classification, and treatment. *Curr Psychiatry Rep*. 2010;12:180.

*Fusar-Poli P, Borgwardt S, Bechdolf A, et al. The psychosis high-risk state: a comprehensive state-of-the-art review. *JAMA Psychiatry*. 2013;70:107.

Fusar-Poli P, Cappucciati M, Bonoldi I, et al. Prognosis of brief psychotic episodes: a meta-analysis. *JAMA Psychiatry*. 2016;73:211–220.

Fusar-Poli P, Cappucciati M, Borgwardt S, et al. Heterogeneity of psychosis risk within individuals at clinical high risk: a meta-analytical stratification. *JAMA Psychiatry*. 2016;73:113–120.

Hodgkinson CA, Goldman D, Jaeger J, Persaud S, Kane JM. Disrupted in schizophrenia 1 (DISC1): association with schizophrenia, schizoaffective disorder, and bipolar disorder. *Am J Hum Genet*. 2004;75:862.

*Kendler KS. Demography of paranoid psychosis (delusional disorder). *Arch Gen Psychiatry*. 1982;39:890.

Kotov R, Leong SH, Mojtabai R, et al. Boundaries of schizoaffective disorder: revisiting Kraepelin. *JAMA Psychiatry*. 2013;70:1276.

Laursen TM, Labouriau R, Licht RW, Bertelsen A, Munk-Olsen T. Family history of psychiatric illness as a risk factor for schizoaffective disorder: a Danish register-based cohort study. *Arch Gen Psychiatry*. 2005;62:841.

Malhi GS, Green M, Fagiolini A, Peselow ED, Kumari V. Schizoaffective disorder: diagnostic issues and future recommendations. *Bipolar Disord*. 2008;10:215.

*Marneros A, Pillman F. *Acute and Transient Psychoses*. Cambridge: Cambridge University Press; 2004.

Naz B, Bromet EJ, Mojtabai R. Distinguishing between first-admission schizophreniform disorder and schizophrenia. *Schizophr Res*. 2003;62:51.

Nugent KL, Paksarian D, Mojtabai R. Nonaffective acute psychoses: uncertainties on the way to DSM-V and ICD-11. *Curr Psychiatry Rep*. 2011;13:203.

Nykiel SA, Baldessarini RJ, Bower MC, Goodwin J, Salvatore P. Psychosis NOS: search for diagnostic clarity. *Harv Rev Psychiatry*. 2008;16:55.

Pillmann F, Haring A, Balzuweit S, Marneros A. A comparison of DSM-IV brief psychotic disorder with "positive" schizophrenia and healthy controls. *Comprehen Psychiatry*. 2002;43:385.

Santelmann H, Franklin J, Bußhoff J, Baethge C. Interrater reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depression - a systematic review and meta-analysis. *Schizophr Res*. 2016;176:357–363.

Schimmelmann BG, Conus P, Edwards J, McGorry PD, Lambert M. Diagnostic stability 18 months after treatment initiation for first-episode psychosis. *J Clin Psychiatry*. 2005;66:1239.

*Schwartz JE, Fennig S, Tanenberg-Karant M, Carlson G, Craig T. Congruence of diagnoses 2 years after a first-admission diagnosis of psychosis. *Arch Gen Psychiatry*. 2000;57:593.

Strakowski SM. Diagnostic validity of schizophreniform disorder. *Am J Psychiatry*. 1994;151:815.

Susser E, Wanderling J. Epidemiology of nonaffective acute remitting psychosis vs schizophrenia: sex and sociocultural setting. *Arch Gen Psychiatry*. 1994;51:294.

Tondo L, Vázquez GH, Baethge C, et al. Comparison of psychotic bipolar disorder, schizoaffective disorder, and schizophrenia: an international, multisite study. *Acta Psychiatr Scand*. 2016;133:34–43.

Wilson JE, Niu K, Nicolson SE, Levine SZ, Heckers S. The diagnostic criteria and structure of catatonia. *Schizophr Res*. 2015;164(1-3):256.

13 Mood Disorders

▲ 13.1 Mood Disorders: Historical Introduction and Conceptual Overview

HAGOP S. AKISKAL, M.D.

CLINICAL AND PUBLIC HEALTH SCOPE OF MOOD DISORDERS

Prevalence

Depressive disorders afflict one of five women and one of ten men at some time during their lives. Depressive episodes alternating with mania or hypomania represent the domain of bipolar disorders. Increasingly, the conventional figure of 1 percent for bipolar disorders in the general population is being challenged, and there are now convincing data that this group of disorders may account for 5 percent of the population and up to 50 percent of all depressions. The enlargement of the boundaries for bipolar disorder is largely due to better detection of the bipolar II subtype (depression plus hypomania rather than mania). The current evidence for clinical, therapeutic, and public health implications of such a broadened bipolar concept has been summarized in a World Psychiatric Association monograph.

Despite the availability of effective treatments, many persons with mood disorders are disabled, and rates of suicide (which occurs in approximately 15 percent of depressive patients) are high in young and, particularly, elderly men. An alarming increase of suicide rates among middle-aged women has been recently reported. Although depressive disorders are more common in women, the traditional view that more men than women die of suicide may therefore need revision in light of these data. High-profile cases of infanticide publicized on television have brought into the public's awareness the role of the reproductive cycle in severe postpartum psychosis and, more generally, the high burden of all forms of depression in women. It is therefore relevant that psychiatric journals targeting clinicians are increasingly devoting larger space to reproductive-related mood disorders in women. Obviously, more than hormones and

physiology are involved in the morbidity and mortality of women, and broader developmental and life cycle–related traumatic, social, and economic factors should be addressed in the research and clinical literature.

The suboptimal outcome of mood disorders documented in recent research reports cannot be ascribed to underdiagnosis and undertreatment alone, for several reasons. First, Gerald Klerman and colleagues suggested that the incidence of mood disorders may be increasing in younger age groups, especially in cohorts born in the 1960s, and may be associated with rising rates of alcohol and substance abuse. Second, mood disorders, once believed to be essentially adult disorders, are increasingly being diagnosed in children and adolescents. Third, clinical studies suggest higher rates of chronicity, recurrence, and refractoriness than previously believed. For instance, chronicity, reported by Emil Kraepelin to occur in no more than 5 percent of this population in the early 20th century in Germany, is now seen in varying degrees in one of three affectively ill patients. Nonetheless, outcome studies coming from university centers tend to overestimate the proportion of cases with less favorable prognosis, and, undeniably, many patients seen in private practice experience a favorable outcome. In addition, not unexpectedly, current data indicate that depressed patients treated by psychiatrists in private settings receive much better care than those in other settings.

Concepts of Mood Disorders

In the European tradition, the broader rubric of *affective disorder* (which subsumes mood and anxiety disorders) has been conceptualized along two influential schools. Aubrey Lewis and his followers from the Maudsley school have promoted a continuum model—from anxiety disorders to mild neurotic depressions to severe endogenous and psychotic depressions. The Newcastle school, led by Martin Roth, has sharply demarcated those conditions from one another. Although vestiges of both approaches are still influential in clinical and basic research, their significance is now overshadowed by European studies in German-speaking countries that subdivide mood disorders on the basis of polarity: unipolar (depressive episodes only) and bipolar (depressive episodes plus manic, hypomanic, or mixed episodes). That subdivision, in part supported by studies in the United States, has served as the basis for much recent research into the biology, treatment, and classification of mood disorders, and is reflected in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) and the tenth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10). Despite such official sanction, many authorities continue to see considerable continuity between recurrent depressive and bipolar disorders. This has led to widespread discussion and debate about the bipolar spectrum (BPS),

which incorporates classic bipolar disorder, bipolar II, and recurrent depressions.

Childhood bipolarity, too, is receiving increasing clinical attention, thanks to the seminal work of Elizabeth Weller and colleagues, originally conducted at Ohio State University. In addition, clinical observations at the University of Tennessee on the juvenile offspring of adult patients with bipolar disorder have led to a greater appreciation of the bipolar nature of complex clinical presentations of affective illness in the juvenile offspring and kin of adult bipolar probands. More recent work by Biederman's group at Harvard suggests intriguing links between pediatric bipolar disorder and attention-deficit/hyperactivity disorder (ADHD).

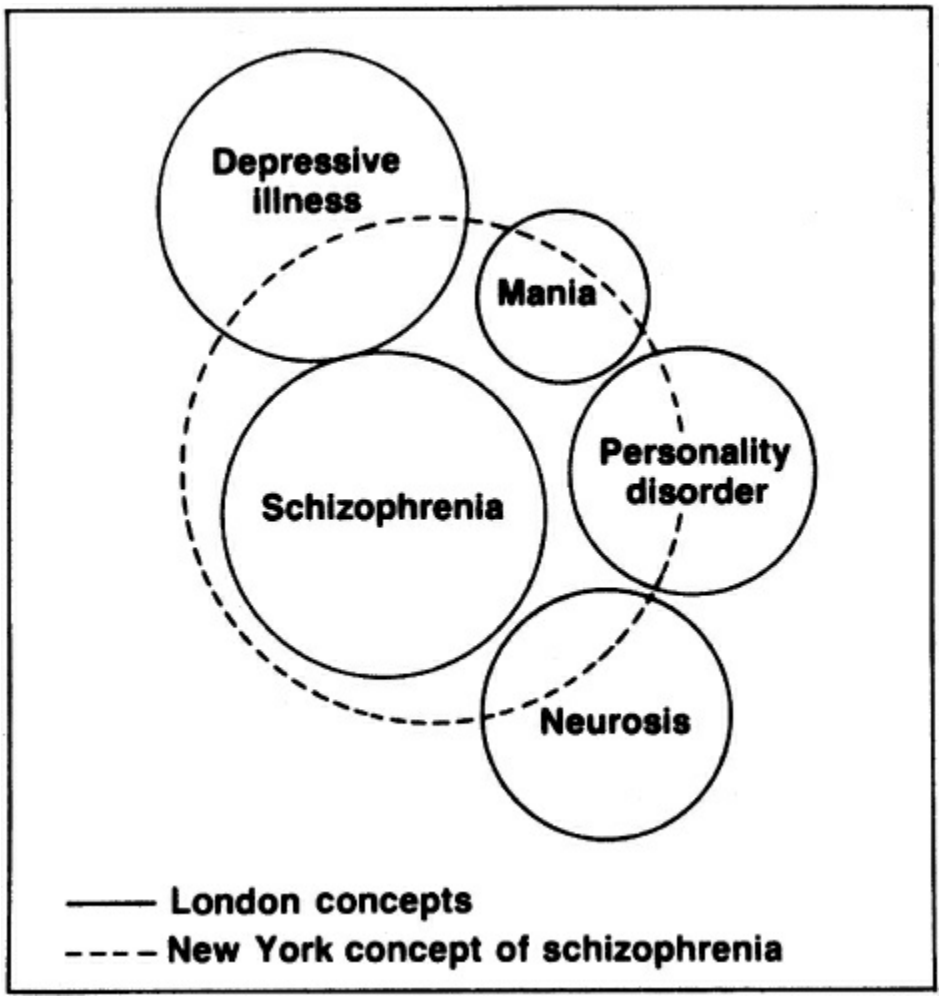

FIGURE 13.1–1. Comparison of British (London) and U.S. (New York) concepts of schizophrenia. (Adapted from Cooper JE, Kendell RE, Garland BJ, Sharpe L, Copeland JRM. *Psychiatric Diagnosis in New York and London*. London: Oxford University Press; 1972.)

Current concepts of mood disorders in the United States embrace a wide spectrum, including many conditions previously diagnosed as schizophrenia, personality disorder, or neurosis. The diagnostic shift occurred in part as a result of the U.S.–U.K. Diagnostic Project, which

demonstrated that schizophrenia was being diagnosed at the expense of mood disorders (Fig. 13.1–1). Conceptual boundaries were further broadened by the availability of new and effective treatments and by the unacceptable risk for tardive dyskinesia and suicide in persons with misdiagnosed mood disorders. More generally, present research interest in mood disorders in the United States emanated from a landmark 1969 NIMH conference on the psychobiology of affective illnesses. The NIMH Collaborative Depression Study—a long-term prospective project deriving directly from recommendations made at the conference—has legitimatized the broader perspective.

Morbidity and Mortality

Unfortunately, findings published by Martin Keller and colleagues in the 1980s documenting gross undertreatment of mood disorders continue to describe the current treatment landscape worldwide. Whatever changes have occurred in diagnostic practice do not appear to have significantly affected the morbidity and mortality of mood disorders. This is all the more scandalous because the 1990s saw new classes of user-friendly antidepressant and mood-stabilizing agents, as well as depression-specific psychotherapies. Thus state of affairs results, in part, from the fact that clinical exposure to mood disorders in both specialized (psychiatric) and primary care (general medical) training is suboptimal. Mood disorders, as highly prevalent and lethal disorders, must command a greater share in the clinical curriculum of both psychiatrists and general medical practitioners. Because most mood disorders are chronically relapsing conditions, long-term exposure to patients with these disorders in mood or bipolar clinics should be obligatory training for young doctors. Unfortunately, few academic centers have such clinics, and those that exist are largely devoted to research. The primary goal of these clinics is the execution of research protocols rather than the acquisition of clinical experience in caring for such patients.

Nearly half of all cases of depression, just like those with adult-onset diabetes, remain undetected for years or inadequately controlled—both of which seem to lag behind hypertension, in which early detection and treatment have significantly reduced complications such as stroke. Efforts by patient-advocacy organizations—often in concert with national psychiatric organizations and governmental mental health agencies—appear to be increasing public and government awareness of mood disorders. Ultimately, however, the challenge is to provide all primary care physicians with the requisite hands-on experience in this prevalent group of disorders. User-friendly tools to detect suicidal patients in the general medical sector would further enhance preventive efforts. These would require significant changes in the structure of health care, including, but not limited to, the greater participation of nurses and social workers as

liaisons in primary mental health and continuity of care for depressed or suicidal patients.

Because mood disorders underlie 50 to 70 percent of all suicides, effective treatment of these disorders on a national level should, in principle, drastically reduce this major complication of mood disorders. That elderly depressed patients, often with medical comorbidity, constitute the highest-risk group for suicide yet escape clinical detection and treatment is particularly problematic for public health. A small-scale prospective Swedish study undertaken by Rihmer and Rutz, although not specifically targeted to the elderly, yielded promising results in this regard. In addition, clinical findings in recurrent mood disorders have clearly shown the value of lithium prophylaxis in the prevention of suicide and overall mortality. Emerging data suggest that such benefits may accrue from all efficacious treatments for mood disorders. Suicide prevention at the national level, attributed to the widespread use of selective serotonin reuptake inhibitors (SSRIs) and primary care physician education, has been achieved in many but not all European countries. Such prevention does not have a tradition in the United States and, as mentioned earlier, suicide rates have actually increased in the United States in the last few years, from 10.5 percent to 11 percent, a net increase of 5 percent.

DEFINITIONS

Mood disorders encompass a large group of psychiatric disorders in which pathological moods and related vegetative and psychomotor disturbances dominate the clinical picture. Known in previous editions of DSM as *affective disorders*, the term *mood disorders* is preferred today because it refers to sustained emotional states, not merely to the external (affective) expression of the present emotional state. Mood disorders are best considered as syndromes (rather than discrete diseases) consisting of a cluster of signs and symptoms, sustained over a period of weeks to months, that represent a marked departure from a person's habitual functioning and tend to recur, often in periodic or cyclical fashion.

Major Depressive Disorder and Bipolar Disorder

Major depressive disorder (MDD) (unipolar depression) is reported to be the most common mood disorder. It may manifest as a single episode or as recurrent episodes. The course may be somewhat protracted—up to 2 years or longer—in those with the single-episode form. Whereas the prognosis for recovery from an acute episode is good for most patients with major depressive disorder three of four patients experience recurrences throughout life, with varying degrees of residual symptoms between episodes. *Bipolar disorders* (previously called *manic-depressive psychosis*) consist of at least one hypomanic, manic, or mixed episode. *Mixed episodes*

represent a simultaneous mixture of depressive and manic or hypomanic manifestations. Although a minority of patients experience only manic episodes, most bipolar disorder patients experience episodes of both polarity. Manias predominate in men, depression and mixed states in women. The bipolar disorders were classically described as psychotic mood disorders with both manic and major depressive episodes (MDEs) (now termed *bipolar I disorder*), but recent clinical studies have shown the existence of a spectrum of ambulatory depressive states that alternate with milder, short-lived periods of hypomania rather than full-blown mania (bipolar II disorder). Bipolar II disorder, which is not always easily discernible from recurrent major depressive disorder, illustrates the need for more research to elucidate the relation between bipolar disorder and major depressive disorder.

The border with schizophrenia continues to be fuzzy. The Kraepelinian dichotomy of dementia praecox and manic-depressive illness has been challenged in favor of a spectrum concept that goes beyond mood-incongruent mood episodes and might spill over into "good-prognosis" schizophrenia.

Dysthymia and Cyclothymia

Clinically, major depressive episodes often arise from a low-grade, intermittent, and protracted depressive substrate known as *dysthymic disorder* (called persistent depressive disorder in DSM-5, but the terms are interchangeable). Likewise, many instances of bipolar disorders, especially ambulatory forms, represent episodes of mood disorder superimposed on a cyclothymic background, which is a biphasic alternating pattern of numerous brief periods of hypomania and numerous brief periods of depression. Dysthymic and cyclothymic disorders represent the two prevalent subthreshold mood conditions roughly corresponding to the basic temperamental dysregulations described by Emil Kraepelin and Ernst Kretschmer as predisposing to affective illness.

It is not always easy to demarcate full-blown syndromal episodes of depression and mania from their subthreshold counterparts commonly observed during the interepisodic periods. The subthreshold conditions appear to be fertile terrain for interpersonal conflicts and postaffective pathological character developments that may ravage the lives of patients and their families. In North America and some Western European countries many such patients end up being labeled with *borderline personality disorder*, which, unfortunately, often tends to obscure the affective origin of the presenting psychopathology.

Cyclothymic and dysthymic conditions also exist in the community without progression to full-blown mood episodes. As such, they are best considered, respectively, as trait bipolar and trait depressive conditions. Understanding the factors that mediate transition from trait to clinical

state is important for preventing manic and major depressive episodes.

A fascinating development related to these temperaments pertains to the "positive" attributes of the depressive and related temperaments in subserving sensitivity to human suffering, thereby playing a role in altruistic behavior. The more tempestuous cyclothymic, based on recent data from different centers, appears to be the temperament involved in poetry, the arts, architecture, and other forms of creative human endeavor. In other words, the "dilute" forms of affective illness may have evolved as traits of fundamental importance for human culture.

Other Subthreshold Mood States

Epidemiological studies in both Europe and North America have also revealed other subsyndromal conditions with depressive and hypomanic manifestations with few symptoms (oligosymptomatic mood states) and of short duration (brief episodes). Variously referred to as *minor*, *subsyndromal*, *brief*, or *intermittent*, these descriptions do not merely represent arbitrary lowering of diagnostic thresholds, but they also herald increasing realization of their importance in early detection of at-risk individuals, as has happened in other medical fields (e.g., diabetes mellitus and essential hypertension). If disabling mood disorders afflict 5 to 8 percent of the general population (Epidemiologic Catchment Area [ECA] study), milder but still clinically significant mood disorders would raise lifetime rates to 17 percent (National Comorbidity Study [NCS]); if subclinical mood states are added, that figure doubles to involve one-third of the general population (as reported, e.g., by Kenneth Kendler and colleagues). New evidence from both Europe and the United States has shown that BPS conditions (bipolar I, bipolar II, and unspecified bipolar disorder in formal diagnostic manuals such as the DSM-5 and the ICD-10) may account for at least 50 percent of all mood disorders in the community and in psychiatric practice.

Comorbidity in mood disorders involves considerable overlap with anxiety disorders. As summarized in an NIMH monograph, anxiety disorders can occur during an episode of depression, may be a precursor to the depressive episode, and, less commonly, may occur during the future course of a mood disorder. Those findings suggest that at least some depressive disorders share a common diathesis with certain anxiety disorders. More recent clinical experience suggests intriguing comorbidity patterns between bipolar II disorder on one hand and panic, obsessive-compulsive, and social phobic states on the other. Furthermore, bipolar I and II disorders are particularly likely to be complicated by the use of alcohol, stimulants, or both. In many cases, the alcohol or substance abuse represents attempts at "self-treatment" of the depression and associated anxiety or insomnia (or both) and, in the case of mania and hypomania, attempts to maintain or enhance the positive moods and energy. Finally,

physical illness—both systemic and cerebral—occurs in association with depressive disorders with a greater frequency than expected by chance alone. Unless properly treated, such depression negatively affects the prognosis of the physical disorder. More provocatively, there is current reawakening to the contribution of cerebral and cardiovascular factors to the origin of late-onset psychotic depressions (previously classified as *involutional melancholia*).

An integrated framework of pathogenesis is necessary for understanding psychopharmacological, somatic, and psychotherapeutic approaches in the clinical management of patients with mood disorders. A historical perspective on current developments is also a valuable lesson in the study of mood disorders.

ANCIENT GREEK AND ROMAN DESCRIPTIONS

Much of what is known today about mood disorders was described by the ancient Greeks and Romans, who coined the terms *melancholia* and *mania* and noted their relation. The ancients also hypothesized a temperamental origin for those disorders. Much of modern thinking about mood disorders (e.g., the work of French and German schools in the middle and latter parts of the 19th century, which influenced current British and U.S. concepts) can be traced back to these ancient concepts.

Melancholia

Hippocrates (460 to 357 Before the Common Era [bce]) described melancholia ("black bile") as a state of "aversion to food, despondency, sleeplessness, irritability, and restlessness." Thus, in choosing the name of the condition, Greek physicians (who may have borrowed the concept from ancient Egyptians) postulated the earliest biochemical formulation of any mental disorder. They believed that the illness often arose from the substrate of the somber melancholic temperament, which, under the influence of the planet Saturn, made the spleen secrete black bile, ultimately leading to mood darkening through its influence on the brain.

One Hippocratic aphorism recognized the close link between anxiety and depressive states: "Patients with fear of long-standing are subject to melancholia." Hippocrates, who described the first historical case of melancholia, may have also been the first to describe a depressive mixed state, an activated form of depression:

> A woman of Thasos became morose because of a justifiable grief, and although she did not take to her bed, she suffered from insomnia, loss of appetite . . . she complained of fears and *talked much*; she showed despondency and . . . talked at random and used foul language . . . many intense and continuous pains . . . she leapt up and could not be restrained. [Emphasis added]

According to Galen (131 to 201), melancholia manifested in "fear and

depression, discontent with life, and hatred of all persons." A few hundred years later, another Roman, Aurelianus, citing the now-lost works of Soranus of Ephesus, amplified the role of aggression in melancholia (and its link to suicide) and described how the illness assumed delusional coloring: "Animosity toward members of the household, sometimes a desire to live and at other times a longing for death, suspicion on the part of the patient that a plot is being hatched against him."

In addition to natural melancholia, which, presumably, arose from an innate predisposition to overproduce the dark humor and led to a more severe form of the malady, ancient medicine recognized such environmental contributions to melancholia as immoderate consumption of wine, perturbations of the soul due to the passions (e.g., love), and disturbed sleep cycles. Autumn was considered the season most disposing to melancholy.

Mania

A state of raving madness with exalted mood was noted by the ancient Greeks, although it referred to a somewhat broader group of excited psychoses than that in modern nosology. Its relation to melancholia was probably noted as early as the first century bce, but, according to Aurelianus, Soranus discounted it. Nonetheless, Soranus had observed the coexistence of manic and melancholic features during the same episode, consisting of continual wakefulness and fluctuating states of anger and merriment and, sometimes, of sadness and futility. Thus, Soranus seemed to have described what today are called *mixed episodes*. Natural melancholy was generally considered a chronic disorder, but Soranus noted the tendency for attacks to alternate with periods of remission.

Although others before him hinted at it, Aretaeus of Cappadocia (ca. 150) is generally credited with making the connection between the two major mood states: "It appears to me that melancholy is the commencement and a part of mania." He described the cardinal manifestations of mania as it is known today:

> There are infinite forms of mania, but the disease is one of them. If mania is associated with joy, the patient may laugh, play, dance night and day, and go to the market crowned as if a victor in some contest of skill. The ideas the patients have are infinite. They believe they are experts in astronomy, philosophy, or poetry.

Aretaeus described the extreme psychotic excitement that could complicate the foregoing clinical picture of mania:

> The patient may become excitable, suspicious, and irritable; hearing may become sharp . . . [and they might] get noises and buzzing in the ears; or may have visual hallucinations; bad dreams and his sexual desires may get uncontrollable; aroused to anger, he may become wholly mad and run unrestrainedly, roar aloud; kill his keepers, and lay violent hands upon himself.

Noting the fluctuating nature of symptoms in the affectively ill, Aretaeus commented: "They are prone to change their mind readily; to become base, mean-spirited, illiberal, and in a little time extravagant, munificent, not from any virtue of the soul, but from the changeableness of the disease." Aretaeus was thus keenly aware of the characterological distortions so commonly manifested during the different phases of cyclical mood disorders.

Finally, consolidating the knowledge of several centuries, Aretaeus described mania as a disease of adolescent and young men given intermittently to "active habits, drunkenness, lechery" and an immoderate lifestyle (what today might be called *cyclothymic disorder*). Exacerbations were most likely to occur in the spring.

Affective Temperaments

The concept of health and disease in ancient medicine was based on harmony and balance of the four humors, of which sanguine humor was deemed the healthiest. Even a desirable humor, however, such as blood, which made persons habitually active, amiable, and prone to jest, could, in excess, lead to the pathological state of mania. The melancholic temperament, dominated by black bile and predisposed to pathological melancholia, was described as lethargic, sullen, and given to brooding or contemplation; its modern counterpart might be persistent depressive disorder as defined in DSM-5. A long tradition dating back to Aristotle (384 to 322 bce) attributed creative qualities to the otherwise tortured melancholic temperament in such fields as philosophy, the arts, poetry, and politics. The remaining two temperaments—choleric and phlegmatic—were less desirable, in that yellow bile made persons choleric (irritable, hostile, and given to rage) and phlegm made them phlegmatic (indolent, irresolute, and timid). The choleric and phlegmatic temperaments would probably be recognized today as borderline personality disorder and avoidant or schizoid personality disorder, respectively.

Many of the original Greek texts on melancholia were transmitted to posterity through medieval Arabic texts such as those of Ishaq Ibn Imran and Avicenna (and their Latin renditions by Constantinus Africanus). In describing different affective states, Avicenna developed the theory of the temperaments to its fullest. He speculated that a special form of melancholia supervened "if black bile be mixed with phlegm," when the illness was "coupled with inertia, lack of movement, and quiet." Furthermore, mania was not necessarily linked to the sanguine (what today is termed *hyperthymic*) temperament, in that many forms of excited madness were believed to represent a mixture of black and yellow bile.

Avicenna further observed that the mixture of anger and restlessness in melancholia indicated that the disease was manic in nature and that the appearance of such signs and symptoms along with violence heralded the

transition from melancholia to mania. Those elaborations on Galen's temperamental types might be considered the forerunners of current personality dimensions, deriving mood states from various mixtures of neuroticism and introversion-extroversion. (What both the ICD-10 and the DSM-5 describe as *cyclothymic disorder* represents the intense mood lability of high neuroticism coupled with cyclic alternation between extroversion and introversion.) Speculation on how diverse depressive phenomena could be understood as a mix of humors anticipated modern multiple-transmitter hypotheses of depression. Ishaq Ibn Imran summarized the existing knowledge of melancholia by considering the interaction of genetic factors ("injured prenatally as the result of the father's sperm having been damaged") with a special temperament given to "mental overexertion"—although not necessarily physical overactivity—that, in turn, was associated with "disruption of the correct rhythms of sleeping and waking." Those views, too, have a very modern ring to them.

MODERN ERA

The first English text (Fig. 13.1–2) entirely devoted to affective illness was Robert Burton's *Anatomy of Melancholy*, published in 1621. A scholarly review of medical and philosophical wisdom accumulated in past centuries, it also anticipated many modern developments. The concept of affective disorder endorsed by Burton was rather broad (as it always has been in the United Kingdom), embracing mood disorders and many disorders that are today considered somatoform disorders, including hypochondriasis. Although he described "causeless" melancholias, Burton also categorized the various forms of love melancholy and grief. Particularly impressive was his catalogue of causes, culminating in a grand conceptualization:

> Such as have Saturn misaffected in their genitures such as are born of melancholy parents as offend in those six non-natural things, are of a high sanguine complexion, are solitary by nature, great students, given to much contemplation, lead a life out of action, are most subject to melancholy. Of sexes both, but men more often. Of seasons of the year, autumn is most melancholy. Jobertus excepts neither young nor old.

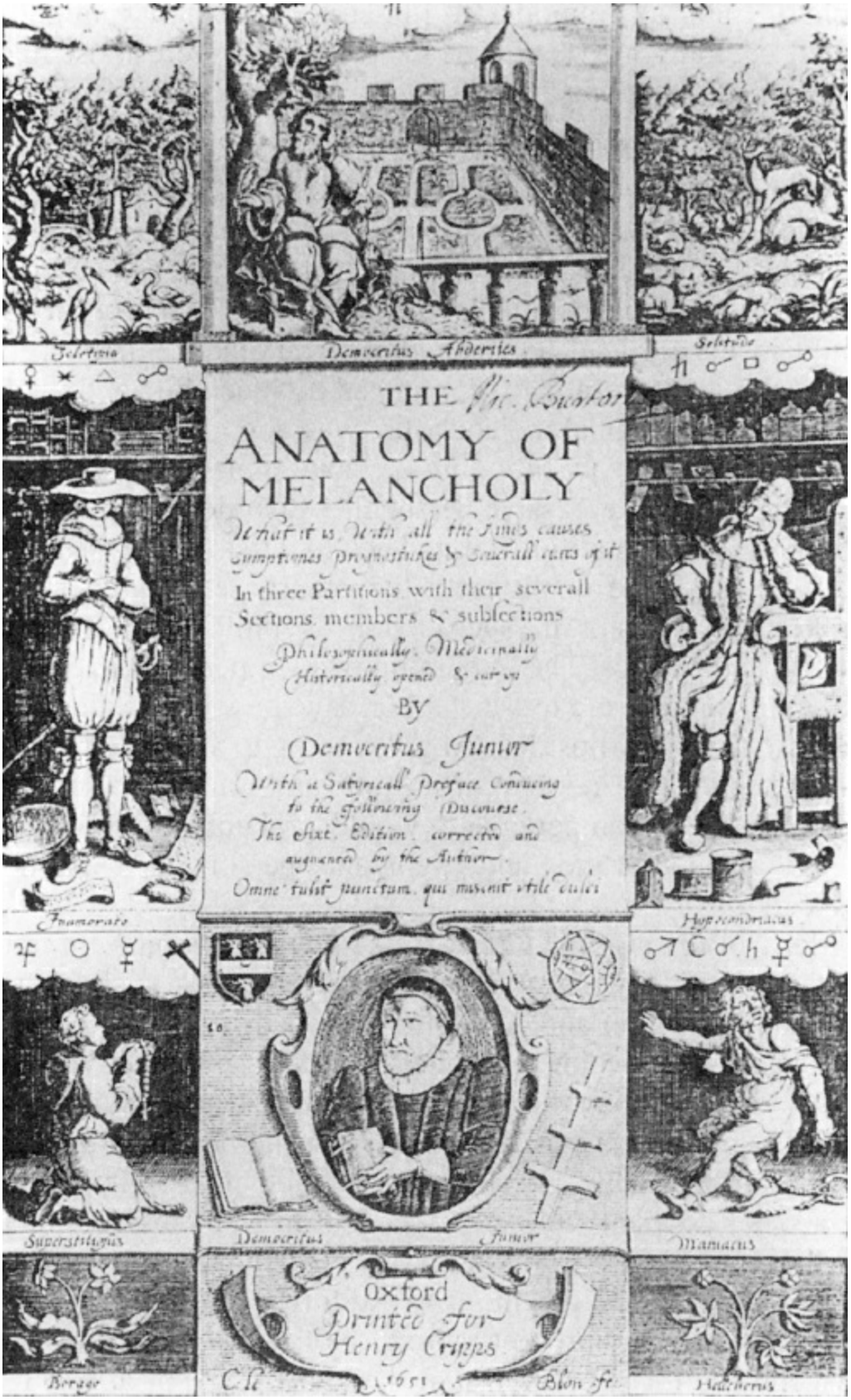

FIGURE 13.1–2. Frontispiece of Robert Burton's *Anatomy of Melancholy* (1621).

Burton's six nonnatural things referred to such environmental factors as

diet, alcohol, biological rhythms, and perturbations induced by passions such as intense love. Burton did not definitively indicate age prevalences. Like nearly all of his predecessors, he favored male (rather than the currently reported female) preponderance.

Finally, Burton considered both the melancholic (contemplative) and the sanguine (hot-blooded) temperaments to be substrates of melancholia. Burton's work thus linked certain forms of depression with the softer expressions of the manic disposition, or bipolar II disorder, from which he appears to have suffered.

The 18th and 19th centuries introduced humane hospital care of the mentally ill, thereby permitting systematic clinical observation of the psychopathology and outcome of mood disorders.

Concept of Affective Disorder

Although Celsus (ca. 30) had described "forms of madness that go no further than sadness," the French alienist Jean-Philippe Esquirol (1840) may have been the first psychiatrist in modern times to suggest that a primary disturbance of mood might underlie many forms of depression and related paranoid psychoses. Until Esquirol's work, melancholia had been categorized as a form of insanity (i.e., ascribed to deranged reasoning or thought disturbance). Esquirol's observations on melancholic patients led him to postulate that their insanity was partial (dominated by one delusion, a monomania) and that "the symptoms were the expression of the disorder of the affections. The source of the evil is in the passions." He coined the term *lypemania* (from the Greek, "sorrowful insanity") to give nosological status to a subgroup of melancholic disorders that were affectively based. Esquirol cited Benjamin Rush (1745 to 1813), the father of American psychiatry, who had earlier described *tristimania*, a form of melancholia in which sadness predominated.

Esquirol's influence led other European psychiatrists to propose milder states of melancholia without delusions, which were eventually categorized as simple melancholias and, ultimately, as primary depressions. Such descriptions culminated in the Anglo-Saxon psychiatric term *affective disorder*, coined by Henry Maudsley (1835 to 1918), the renowned British psychiatrist after whom a London hospital is named.

Manic-Depressive Illness and the Question of Psychogenic Depressions

Although the connection between mania and depression had been rediscovered sporadically since it was first described 2,000 years ago, the clinical work that finally established *circular insanity* (Jean-Pierre Falret's term) and *folie à double forme* (Jules Baillarger's term) as discrete nosological entities with both depressive and manic poles was undertaken by these two Esquirol disciples in the 1850s. That accomplishment built on

Philippe Pinel's reforms, which championed humane treatment of the mentally ill in Paris around the turn of the 18th century and emphasized systematic clinical observations of patients, which were detailed in case records. French alienists made longitudinal observations on the same patient from one psychotic attack into another. Furthermore, Esquirol had introduced the practice of chronicling events in statistical tables. Thus, the Hippocratic approach to defining a particular case by its onset, circumstances, course, and outcome was applied by French alienists in studying the affectively ill. The humanitarian reforms introduced in the 19th century ensured that standards of general health and nutrition would improve the outlook for the mentally ill—especially those with potentially reversible disorders such as affective disorders, who could now be discharged from the asylums. The French school, by segregating the nondeteriorating mood disorders from other types of insanity, then paved the way for the Kraepelinian system.

Kraepelin's (1856 to 1926) unique contribution was not so much his grouping together of all the forms of melancholia and mania, but his methodology and painstaking longitudinal observations, which established manic-depressive illness as a nosological entity and (he hoped) a disease entity. His rationale was based on the following: (1) the various forms had a common heredity measured as a function of familial aggregation of manic and depressive cases, (2) frequent transitions from one to the other occurred during longitudinal follow-up of patients, (3) a recurrent course with illness-free intervals characterized most cases, (4) the superimposed episodes were commonly opposite to the patient's habitual temperament—that is, mania could be superimposed on a depressive temperament and depression on a hypomanic temperament—and (5) both depressive and manic features could occur during the same episode (mixed states). Kraepelin's synthesis was developed as early as the sixth (1899) edition of his *Lehrbuch der Psychiatrie* and most explicitly stated in the opening passages of the section on manic-depressive psychosis in the eighth edition (published in four volumes, 1909 to 1915):

> Manic-depressive insanity includes on the one hand the whole domain of so-called periodic and circular insanity, on the other hand simple mania, the greater part of the morbid states termed melancholia and also a not inconsiderable number of cases of confusional insanity. Lastly, we include here certain slight and slightest colorings of mood, some of them periodic, some of them continuously morbid, which on the one hand are to be regarded as the rudiment of more severe disorders, on the other hand, pass over without boundary into the domain of personal predisposition.

For Kraepelin, the core pathology of clinical depression consisted of lowered mood and slowed (retarded) physical and mental processes. In mania, by contrast, the mood was elated, and physical and mental activity were accelerated. His earlier observations on what he termed *involutional melancholia* (referring to 40- to 65-year-old patients with extreme anxiety,

irritability, agitation, and delusions) had led him to separate that entity from the broader manic-depressive rubric. In the eighth edition of *Lehrbuch der Psychiatrie,* however, he united melancholia with the manic-depressive group, with the justification that it was a special form of mixed state and that follow-up conducted by his pupil Dreyfus had demonstrated unmistakable excited phases.

The classification of mood disorders is still evolving. Karl Leonhard in 1957, Jules Angst in 1966, Carlo Perris in 1966, and George Winokur, Paula Clayton, and Theodore Reich in 1969, working independently in four different countries, proposed that depressive disorders without manic or hypomanic episodes (unipolar depressive disorder) that appear in middle age and later are distinct from depressive episodes that begin at earlier ages and alternate with manic or hypomanic episodes (bipolar disorder). The main difference between the two affective subtypes is the greater familial loading for mood disorder—especially for bipolar disorder—among bipolar disorder probands.

Kraepelin had conceded the occurrence of psychogenic states of depression occasioned by situational misfortune. However, he believed that manic-depressive illness was hereditary, although he could not document postmortem anatomopathological findings in the brains of manic-depressive patients. Therefore, manic-depression had to be considered a functional mental disorder in which brain disturbances were presumed to lie in altered physiological functions. Such biological factors were deemed absent in the psychogenic depressions. Thus, Kraepelin's classification of mood disorders is both dualistic and unitary. It is dualistic to the extent that he designated them as either psychologically occasioned or somatically caused. It is unitary with respect to disorders in the latter group, which have been termed *endogenous affective disorders* (i.e., due to internal biological causes). In other words, Kraepelin restricted the concept of clinical depression to what the DSM-5 terms *major depressive disorder with melancholic features*. Moreover, he postulated a continuum between that condition and what the DSM-5 and the ICD-10 now term *bipolar disorders*.

Cartesian thinking in 17th-century France conceptually separated mind from body, thereby providing physicians autonomy over the somatic sphere, free from interference by the Church. The dichotomous paradigm ensured that the study of the two aspects of the human organism would not be confounded by the complexities of mind–body interactions. That is one reason Kraepelin's descriptive observations have proved valuable to subsequent generations of clinicians. Furthermore, his approach exemplifies the best tradition of scientific humanism in medicine: Description and diagnostic categorization of an individual patient are necessary for the physician to apply the knowledge gained from past observation of similarly described and diagnosed patients. One limitation

to the Kraepelinian approach is that because of its biological reductionism, it is not sufficiently articulate to account for mind–body interactions in the genesis of mental disorders.

Depressions as Psychobiological Affective Reaction Types

It was the ambition of Swiss-born Adolf Meyer (1866 to 1950) to bridge the divide between psyche and soma. Meyer, who dominated psychiatry from his chair at Johns Hopkins University during the first half of the 20th century, coined the term *psychobiology* to emphasize that both psychological and biological factors could enter into the causation of depressive and other mental disorders. Because of the nascent state of brain science during Meyer's time, he was more adept at biography than biology and therefore paid greater attention to psychosocial causation. He preferred the term *depression* (pressed down) to *melancholia* because of its lack of biological connotation. He conceived of depressive states in terms of unspecified constitutional or biological factors interacting with a series of life situations beginning at birth or even at conception. From that viewpoint arose the unique importance accorded personal history in depressive reactions to life events.

Meyer's terminological revision left a somewhat confusing legacy, in that the term *depression* is now applied to a broad range of affective phenomena ranging from sadness and adjustment disorders to clinical depression and bipolar disorders. Nosological nuances to which Meyerians paid little attention, such as the difference between melancholic depression and more mundane depressions, are not just a matter of semantics. To the extent that those two forms of depression are seen in different clinical settings, hypotheses based on one population may not apply to the other. For instance, uncontrollable traumatic events may have taught study participants to feel helpless or to view the world in a negative light, but that does not equate with clinical depression; nor does the process appear to be specific to depression. Failure to make such nosological distinctions further clouds interpretations of the results of trials comparing psychotherapy and pharmacotherapy for depressive disorders.

On the other hand, the Meyerian emphasis on biographical factors for the patient represented a more practical approach to depth psychology. Recent sociological interpretations of depression can also be traced to Meyer's work. In final analysis, however, the Meyerian concern for the uniqueness of the individual has proved heuristically sterile. It de-emphasizes what is diagnostically common to different individuals, thereby obscuring the relevance of accrued clinical wisdom for the index patient. For that reason, the Meyerian approach, after enjoying clinical popularity for several decades in North America, has given way to neo-Kraepelinian rigor. However, the psychobiological vision of bridging biology and psychology, one of the major preoccupations of psychiatric thought and

research today, owes much to Meyer's legacy.

CONTEMPORARY MODELS OF DEPRESSION

Aggression-Turned-Inward Model

Sigmund Freud was initially interested in a psychoneural project for all mental phenomena. Limitations of the brain sciences of the day led him to adopt instead a model that relied on a concept of mental function borrowed from physics. The notion that depressed affect is derived from retroflexion of aggressive impulses directed against an ambivalently loved internalized object was actually formulated by his Berlin disciple Karl Abraham and later elaborated by Freud. Abraham and Freud hypothesized that turned-inward anger was intended as punishment for the love object that had thwarted the depressed patient's need for dependency and love. Because, in an attempt to prevent the traumatic loss, the object had already been internalized, the patient then became the target of his or her own thanatotic impulses. A central element in those psychic operations was the depressed patient's ambivalence toward the object, which was perceived as a frustrating parent. Aggression directed at a loved object (parent) was therefore attended by considerable guilt. In the extreme, such ambivalence, guilt, and retroflexed anger could lead to suicidal behavior.

According to this model, depression is an epiphenomenon of the transduction of thanatotic energy, a reaction that takes place in the closed hydraulic space of the mind. Freud's earlier writings had similarly portrayed anxiety as being derived from the transformation of dammed-up sexual libido. Although Freud envisioned that psychoanalytic constructs would one day be localized neuroanatomically, the hydraulic mind is a metaphor that does not refer to actual physiochemical space in the brain.

The concept of emotional behavior as an arena of incompatible forces confined to a psyche that is relatively impervious to current influences outside of the organism is the major liability of the aggression-turned-inward model and, perhaps, of orthodox psychoanalysis itself. Although the sexual-energy-transduction hypothesis of anxiety has been discarded in modern psychoanalytic thought, in modified versions, the aggression-turned-inward model continues to be used in clinical conceptualization. The lingering popularity of the model may be due, in part, to its compatibility with the clinical observation that many depressed patients suffer from lack of assertion and outwardly directed aggressiveness. However, a substantial number of hostile, depressed patients are also encountered in clinical practice (indeed, "depression with anger attacks" has been recently described), and clinical improvement in most patients typically leads to decreased, not increased, hostility. Such observations shed doubt on the aggression-turned-inward mechanism as a universal explanation for depressive behavior. Finally, little evidence exists to

support the contention that outward expression of anger has therapeutic value in clinical depression.

Outwardly directed hostility in depression is not a new clinical observation; in fact, ancient Greek and Roman physicians had noted it. Hostility is best considered a manifestation rather than a cause of depressive disorder, especially when the disorder is attended by mixed bipolar features. The hostility of the depressed patient can also be understood as an exaggerated reaction to frustrating love objects, as secondary to self-referential attribution, or simply as nonspecific irritability of an ego in affective turmoil; this could, in part, be a function of a concurrent personality disorder from the erratic cluster. Such commonsense explanations that do not invoke unobservable hydraulic transmutations have greater appeal from heuristic and clinical perspectives.

Object Loss and Depression

Object loss refers to traumatic separation from significant objects of attachment. Ego-psychological reformulations of the Abraham–Freud conceptualization of depression have paid greater attention to the effect of such losses on the ego, de-emphasizing the id-libidinal and related hydraulic aspects. The depressant effect of separation events often resides in their symbolic meaning for a person rather than in any arbitrary objective weight that the event may have for clinical raters. However, love loss, bereavement, and other exits from the social scene, as defined by the London psychiatrist Eugene Paykel, are presently the concepts most commonly used in practice and research.

Although love melancholy had been described since antiquity, the two affective states were systematically compared for the first time in Freud's 1917 paper on mourning and melancholia. According to current data, the transition from grief to pathological depression occurs in no more than 10 percent of adults and 20 percent of children. These figures suggest that such transition occurs largely in persons predisposed to mood disorders.

John Bowlby of the Tavistock Clinic, London, did a comprehensive clinical investigation of the attachment that the child establishes with the mother or mother substitutes during development, a bond considered the prototype for all subsequent bonds with other objects. Like many psychoanalytic explanations of adult symptom formation, the object loss model is formulated as a two-step hypothesis, consisting of early breaks in affectional bonds, which provide the behavioral predisposition to depression, and adult losses, which are said to revive the traumatic childhood loss and so precipitate depressive episodes. However, the role of proximate separations in provoking depressive reactions rests on more solid clinical evidence than the hypothesized sensitization resulting from developmental object loss. That realization led Bowlby to regard childhood

sensitization resulting from early deprivation as a generic characterological vulnerability to a host of adult psychopathological conditions.

Compared with aggression turned inward, object loss is more directly relevant to clinical depression, yet it is still pertinent to question whether it is an etiological factor. Studies at the Wisconsin Primate Center indicated that optimal homeostasis with the environment is most readily achieved when the individual is securely attached to significant others, and the dissolution of such ties appears to be relevant to the emergence of a broad range of psychopathological disturbances rather than depression per se. A related methodological question is whether object loss operates independently of other etiological factors. For instance, a history of early breaks in attachment may reflect the fact that one or both of the patient's parents had mood disorder, with resultant separation, divorce, suicide, and so forth.

On balance, the ego-psychological object loss model is conceptually superior to its id-psychological counterpart. In postulating an open system of exchange between a person and the environment, the model permits consideration of etiological factors other than separation, such as heredity, character structure, and adequacy of social support—all of which might modulate the depressant effect of adult separation events. Conceptualizing the origin of depression along those lines is in the mainstream of current ideas of adaptation, homeostasis, and disease. An important treatment implication is the value of social support in preventing relapse and mitigating chronicity of depression. That is, indeed, an ingredient in the interpersonal psychotherapy of depression, which can be conceptualized as a form of brief, focused, and practical psychodynamic therapy.

Loss of Self-Esteem and Depression

Reformulation of the dynamics of depression in terms of the ego suffering a collapse of self-esteem represents a further conceptual break with the original id-psychological formulation; depression is said to originate from the ego's inability to give up unattainable goals and ideals. The model further posits that the narcissistic injury that crushes the depressed patient's self-esteem is imposed by the internalized values of the ego rather than the hydraulic pressure of retroflected thanatotic energy deriving from the id. Because the construct of the ego is rooted in social and cultural reality, loss of self-esteem may result from symbolic losses involving power, status, roles, identity, values, and purpose for existence. Thus, the existential and sociocultural implications of depression conceived as a derivative ego state provide the clinician with a far more flexible and pragmatic tool for understanding depressed persons than the archaic hydraulic metaphors related to libidinal vicissitudes. That model represents one of the first attempts to formulate depression in terms that subsequent psychological theory and research could operationalize in more testable

form.

Self-esteem is part of the habitual core of the individual and, hence, is integral to the personality structure. Indeed, low self-esteem conceived as a trait is a major defining attribute of the depressive (melancholic) personality. Although it is understandable that such individuals can easily sink into melancholia in the face of environmental adversity, it is not obvious why persons with apparently high self-esteem (e.g., those with hypomanic and narcissistic personalities) also succumb to melancholy with relative ease. To explain such cases, one must invoke an underlying instability in the system of self-esteem that renders it vulnerable to depression. The opposite is also known to occur; that is, manic episodes may develop from a baseline of low self-esteem, as in the case of bipolar disorder patients with antecedent traits of shyness, insecurity, and dysthymia.

The foregoing considerations suggest that the vicissitudes of self-esteem deemed central to the model of depression as loss of self-esteem are manifestations of a more fundamental mood dysregulation. In classic psychoanalysis, such dysregulation is considered to be of constitutional origin. In general, attempts by psychoanalytic writers to account for bipolar oscillations have not progressed beyond metapsychological jargon, with the notable exception of denial of painful affects as a mechanism in the phenomenology of mania.

Cognitive Model

The cognitive model, developed by Aaron Beck at the University of Pennsylvania, hypothesizes that thinking along negative lines (e.g., thinking that one is helpless, unworthy, or useless) is the hallmark of clinical depression. In effect, depression is redefined in terms of a cognitive triad, according to which patients think of themselves as helpless, interpret most events unfavorably vis-à-vis the self, and believe the future to be hopeless. In more recent formulations in academic psychology, these cognitions are said to be characterized by a negative attributional style that is global, internal, and stable and that exists in the form of latent mental schemata that generate biased interpretations of life events.

Because the cognitive model is based on retrospective observations of already depressed persons, it is virtually impossible to prove that causal attributions such as negative mental schemata precede and, therefore, predispose to clinical depression; they can just as readily be regarded as subclinical manifestations of depression. The theoretical importance of the cognitive model lies in the conceptual bridge it provides between ego-psychological and behavioral models of depression. It has also led to a new and widely accepted system of psychotherapy that attempts to alter the negative attributional style, to alleviate the depressive state, and, ultimately, to fortify the patient against future lapses into negative

thinking, despair, and depression.

The cognitive model, therefore, has the cardinal virtue of focusing on key reversible clinical dimensions of depressive illness, such as helplessness, hopelessness, and suicidal ideation, while providing a testable and practical psychotherapeutic approach. That approach, however, is less likely to succeed in patients with the full-blown melancholic manifestations of a depressive disorder. It is doubtful that negative cognitions alone could account for the profound disturbances in sleep, appetite, and autonomic and psychomotor functions encountered in melancholic depressions. Furthermore, conceptualizing a multifaceted malady such as depression largely or solely as a function of distorted cognitive processes is reminiscent of pre-Esquirolian notions that emphasized impaired reasoning in the development of depression. Finally, recent extensions or modifications, or both, of cognitive therapy in association with behavioral therapy (cognitive–behavioral therapy) for all emotional disorders (and even for schizophrenia) are reminiscent of earlier global claims of the psychodynamic perspective.

Learned Helplessness Model

The learned helplessness model is, in some ways, an experimental analogue of the cognitive model. The model proposes that the depressive posture is learned from past situations in which the person was unable to terminate undesirable contingencies. The model is based on experiments in dogs that were prevented from taking adaptive action to avoid unpleasant electrical shock and subsequently showed no motivation to escape such aversive stimuli, even when escape avenues were readily available. Armed with evidence from many such experiments, University of Pennsylvania psychologist Martin Seligman postulated a trait of learned helplessness (a belief that it is futile to initiate personal action to reverse aversive circumstances) formed from the accumulation of past episodes of uncontrollable helplessness.

The learned helplessness paradigm is a general one and refers to a broader mental disposition than depression. Thus, it is potentially useful in understanding such diverse conditions as social powerlessness, defeat in sporting events, and posttraumatic stress disorder (PTSD). In addition, past events might shape a characterological cluster—consisting of passivity, lack of hostility, and self-blame—relevant to certain depressive phenomena. The low hostility observed in some patients during clinical depression could, for instance, be ascribed to the operation of such factors. Learned helplessness could thereby provide plausible links between aspects of personal biography and clinical phenomenology in depressive disorders. Therapeutic predictions for alleviating depression and related psychopathological states capitalize on new cognitive strategies geared to modifying expectations of uncontrollability and the negative attributional

style. This illustrates how insights gained from experimental paradigms can be combined fruitfully to address clinical disorders.

Nonetheless, the clinician should be wary of unwarranted clinical extrapolations. For example, some therapists have argued that the depressed patient's passivity is "manipulative," serving to obtain interpersonal rewards. It has also been suggested that such factors have a formative influence on the development of the depressive character. That interpretation appears more relevant to selected aspects of depression than to the totality of the disorder. Depressive behavior and verbalizations clearly have a powerful interpersonal impact, but casting depression as merely a masochistic lifestyle developed to secure interpersonal advantages represents a mechanistic circular argument that could be viewed as disrespectful of the clinical agony of patients with mood disorders. Finally, although most formulations focusing on helplessness have emphasized acquisition through learning, recent experimental research in animals tends to implicate genetic factors in the vulnerability of learning to behave helplessly. The value of the helplessness paradigm may reside in its utility in predicting a variety of subthreshold affective disturbances generic to civilian reactions to adversity and trauma.

Depression and Reinforcement

Other behavioral investigators, including, notably, psychologist Peter Lewinsohn, have developed clinical formulations of depression that hinge on certain deficits in reinforcement mechanisms. According to the reinforcement model, depressive behavior is associated with lack of appropriate rewards and, more specifically, with receipt of noncontingent rewards. The model identifies several contributory mechanisms. Some environments may consistently deprive persons of rewarding opportunities, thereby placing them in a chronic state of boredom, pleasurelessness, and, ultimately, despair. That reasoning, however, may offer more insight into social misery than clinical depression. A more plausible postulated mechanism is the provision of rewards that are not in response to the recipient's actions; in other words, the gratis provision of what a person considers undeserved rewards may lead to lowering of self-esteem. Predisposition to depression is formulated in terms of inadequate social skills, which are hypothesized to decrease a person's chances of responding to potentially rewarding contingencies in the environment. Indeed, recent research on the relationship between personality and depression suggests that such deficits might underlie certain depressive states. Therefore, psychotherapeutic approaches designed to enlarge a patient's repertoire of social skills may prove valuable in preventing some types of depression.

The concepts of depression that have been derived from behavioral methodology and developed in the last several decades are scientifically

articulate and, therefore, testable approaches to clinical depression. However, the important distinction between depression on self-report inventories and clinical depression tends to be overlooked in investigations testing the reinforcement paradigm. Furthermore, the behavioral model does not address the distinct possibility that reinforcement deficits may, in part, represent the psychomotor deficits of depressive illness. Nevertheless, by focusing on reward mechanisms, the behavioral model provides a conceptual bridge between purely psychological and emerging biological conceptualizations of depression.

Biogenic Amine Imbalance

The formulation of sophisticated biological explanations of mood disorders had to await the development of neurobiological techniques that could probe the parts of the brain involved in emotions. Although the complex physiology of the limbic-diencephalic centers of emotional behavior generally cannot be directly observed in humans, much has been learned from animal work. The limbic cortex is linked with both the neocortex, which subserves higher symbolic functions, and the midbrain and lower brain centers, which are involved in autonomic control, hormonal production, and sleep and wakefulness. Norepinephrine-containing neurons are involved in many functions that are profoundly disturbed in melancholia, including mood, arousal, appetite, reward, and drives. Other biogenic amine neurotransmitters that mediate such functions are the catecholamine dopamine—especially important for drive, pleasure, sex, and psychomotor activity—and the indoleamine serotonin, which is involved in the regulatory control of affects, aggression, sleep, and appetite. Cholinergic neurons, secreting acetylcholine at their dendritic terminals, are generally antagonistic in function to catecholaminergic neurons.

Although the opioid system might, on experimental and theoretical grounds, also serve as one of the neurochemical substrates for mood regulation, no cogent model of mood disorders involving that system has appeared. Likewise, biochemical formulations of mood disorders have paid relatively little attention to the major excitatory brain neurotransmitter glutamate and the inhibitory neurotransmitter γ-aminobutyric acid (GABA).

Biogenic Amine Hypotheses. Joseph Schildkraut at Harvard University and William Bunney and John Davis at NIMH published the first formal hypothesis connecting depletion or imbalance of biogenic amines (specifically norepinephrine) and clinical depression. The serotonin counterpart of the model was emphasized in the models proposed by Alec Coppen in England and I. P. Lapin and G. F. Oxenkrug in Russia. Both catecholamine and indoleamine hypotheses were essentially based on two sets of pharmacological observations. First, reserpine (Serpasil), which

decreases blood pressure by depleting biogenic amine stores, precipitates clinical depression in some patients. Second, antidepressant medications, which alleviate clinical depression, raise the functional capacity of the biogenic amines in the brain. This style of thinking is known as the *pharmacological bridge*, extrapolating from evidence on the mechanism of drug action to the neurotransmitter pathologies presumed to underlie a given psychiatric disorder. Such pharmacological strategies have been of heuristic value in developing research methods for the investigation of mood disorders and schizophrenia. Indeed, the research methodology developed by the relatively few investigators working in this area during the last half century is among the most elegant in the history of psychiatry.

Variations of the biogenic amine model assign somewhat different relative weights to the biogenic amines norepinephrine and serotonin in the development of pathological mood states. Arthur Prange and colleagues at the University of North Carolina formulated a permissive biogenic amine hypothesis in which serotonin deficits permit the expression of catecholamine-mediated depressive or manic states. That hypothesis was supported by subsequent animal research showing that an intact serotonin system is necessary for optimal functioning of noradrenergic neurons. Omission of tryptophan from the diet of antidepressant-responsive depressed patients may annul the efficacy of the antidepressant; among healthy volunteers, that special diet also induces sleep electroencephalographic characteristics of clinical depression. Although such findings are provocative, the precursor-loading strategy to increase the brain stores of serotonin (e.g., with l-tryptophan) has not been unequivocally successful in reversing clinical depression. Dietary loading with catecholamine precursors has fared even worse than serotonin-precursor loading in the treatment of depression.

The cholinergic–noradrenergic imbalance hypothesis proposed by David Janowsky and colleagues represents yet another attempt to elucidate the roles of biogenic amines. This hypothesis, along with the related cholinergic supersensitivity hypothesis developed by J. Christian Gillin, has been tested extensively at the University of California at San Diego. Subsequent formulations by Larry Siever and Kenneth Davis at the Mount Sinai Hospital in New York have refocused on noradrenergic dysregulation. The model assumes oscillation from one output mode to the other at different phases of depressive illness. In a provocative extrapolation from that model, bipolar depression would have low noradrenergic output, but many instances of major depressive disorder as with some anxiety disorders, could be biochemically conceptualized as high-output conditions.

Despite more than four decades of extensive research and indirect evidence, however, no deficiency or excess of biogenic amines in specific brain structures has been shown to be necessary or sufficient for the

occurrence of mood disorders. It has not been possible either to confirm the putative role of central norepinephrine in depression or to discard it altogether. The role of dopamine as formulated, among others, by the Italian pharmacologist Gian Luigi Cessa, although studied less extensively than that of norepinephrine, deserves greater recognition, in that it might have relevance to atypical and bipolar depression, as well as to mania.

Preliminary data from a small brain imaging study has shown blunted serotonin responsivity in prefrontal and temporoparietal areas in unmedicated patients with major depressive disorder. Such data, considered in the context of the overall serotonin literature in depression, are provocative but not conclusive and serves to illustrate the fact that the case for serotonergic disturbance in depression continues to be based on indirect evidence. Moreover, the putative permissive role of serotonin is better documented for aggressive suicide attempts. Serotonergic dysfunction might subserve other conditions characterized by lack of inhibitory control, among them obsessive-compulsive disorder (OCD), panic disorders, bulimia nervosa, certain forms of insomnia, alcoholism (alcohol abuse or dependence), and a host of impulse-ridden personality disorders. Such considerations have led Dutch psychiatrist Herman van Praag and colleagues to postulate a dimensional neurochemical disturbance generic to a large group of disorders within the traditional nosology. This hypothesis might be variously regarded as a challenge to psychiatric nosology or as a statement of the need to supplement clinical classification with biochemical parameters. Both interpretations are in line with clinical observations during the last two decades testifying to the high prevalence of comorbidity in depressive, other emotional disorders, and certain impulse-control disorders.

It is implied that the foregoing postulated biochemical faults are genetically determined. Although biogenic amine models of mood disorders were developed retrospectively from the pharmacological action of antidepressant and thymoleptic agents, they have stimulated the development of new classes of antidepressants with more selective action on specific neurotransmitter receptors. Their introduction has virtually revolutionized the treatment of depression. Yet the fundamental biochemistry of mood disorders is still far from being understood. Curiously, although selective in action, the new compounds working on the serotonin system have broad effectiveness in a variety of mood-related conditions, such as dysthymic disorder, PTSD, OCD, panic disorder, social phobia, bulimia nervosa, and borderline personality disorder. Such data indirectly favor the hypothesis of an underlying biological commonality to several of these disorders. The foregoing considerations have, in turn, led to a provocative formulation of an increasingly prevalent “social syndrome” in populations experiencing social disruption, immigration, and abuse and characterized by anxiety, depression, violence-proneness, impulsivity, and

suicidality—reflecting a perturbed serotonin system, the oldest, most basic brain structure involved in human socialization and territoriality, coping with stress, danger, and survival.

New antidepressants with dual action on both serotonergic and noradrenergic receptors and emerging data on their possible greater efficacy in melancholic depressions suggest that the biochemistry of mood disorders involves more complex dysregulation than is implied in single-neurotransmitter hypotheses. The work of George Henninger and colleagues at Yale University further suggests that monoamines better explain how antidepressants facilitate recovery from depression than their being the fundamental causes of depression. Moreover, emerging biochemical paradigms are moving away from distal biochemical lesions to focus on molecular perturbations closest to the putative genetic underpinnings of mood disorders. Originally tied to the mechanism of action of mood stabilizers in bipolar disorder, such work is exploring second messenger systems, phosphorylation G proteins, signal transduction, DNA transcription, and messenger RNA translation. Again, such search for molecular mechanisms represents "backward logic" from the putative mechanism of action of selected thymoleptic agents. The same can be said about Frederick Petty's GABAergic and Shih-Jen Tsai's brain-derived neurotrophic factor (BDNF) hypothesis on the origin of bipolar disorder.

Neuroendocrine Links. Functionally inadequate mobilization of neurotransmitters in the face of continued or repeated stress, as indirectly reflected in pathological modification of noradrenergic and serotonergic receptor function, could represent neurochemical final common pathways of homeostatic failure. Such mechanisms could also provide links with psychoendocrine dysfunction; the hypothesized neurotransmitter deficits may underlie the disinhibition of the hypothalamic–pituitary–adrenal axis, characterized by steroidal overproduction, the most widely studied endocrine disturbance in depressive illness. When challenged with dexamethasone (Decadron), the altered axis resists suppression, which offered Bernard Carroll's team (then at the University of Michigan) the possibility of developing the dexamethasone-suppression "test" (DST) for melancholia. This procedure is of uncertain specificity for depressive illness and, thus, is unsuitable to serve as a diagnostic test. However, that line of research has been useful in pathogenetic understanding. For instance, it led to the demonstration by Emory University's Charles Nemeroff of increased concentrations of corticotropin-releasing factor (CRF) in the cerebrospinal fluid (CSF) of patients with major depressive disorder. CRF also appears relevant to the pathophysiology of anxiety disorders, such as panic disorder, and PTSD. The research of Florian Holsboer's group at Munich's Max Planck Institute has shown impaired glucocorticoid and

mineralocorticoid receptor function in these disorders, with relevant pathophysiological and therapeutic implications.

Another neuroendocrine index of noradrenergic dysregulation, blunted growth hormone response to the α_2-adrenergic receptor agonist clonidine (Catapres), likewise points to limbic-diencephalic disturbance. However, studies performed in the United States suggest that it is positive in both endogenous depression and severe anxiety disorder (panic disorder). Thyroid-stimulating hormone (TSH) blunting on thyrotropin stimulation, another common neuroendocrine disturbance in depression, also shows limited specificity.

What is remarkable, however, is that the DST, clonidine, and thyrotropin challenge data, in aggregate, identify most persons with clinical depression. Such evidence of midbrain disturbance argues for considering clinical depression to be a legitimate disease. The disease concept of depression is further buttressed by computed tomography (CT) scans showing enlarged pituitary and adrenal glands, a state marker of depressive illness.

Stress and Depression. The concept of a pharmacological bridge implies two-way traffic. The hypothesized chemical aberrations may be primary or genetically based. Provision should also be made, however, for the likelihood that psychological events that precipitate clinical depression might initiate or exacerbate neurochemical imbalance in vulnerable subjects. That suggestion is supported by studies in animals in which early separation from peers and inescapable frustration effect profound alterations in the turnover of biogenic amines and in postsynaptic receptor sensitivity. Thus, in genetically predisposed persons, environmental stressors might more easily lead to perturbations of limbic-diencephalic neurotransmitter balance. Finally, in vulnerable individuals, especially during the formative years of childhood, psychological mechanisms might more easily perturb midbrain neurochemistry. Traumatic experiences appear particularly potent in this regard. The hippocampus has been the subject of intense recent research as the possible neuroanatomical substrate linking such loss and trauma to adult depression. Ongoing ingenious experimental paradigms in primates and rodents continue to explore the role of early experience and stress in subsequent depressive-like behaviors in these animals. In humans, a new provocative finding indicates that a polymorphism of the serotonin transporter gene would identify who among traumatized children would develop adult depression. Likewise, a polymorphism of the monoamine oxidase (MAO), a gene plays a significant role in determining who among battered children will grow into an adult sociopath. Animal models of mania are sparse and problematic.

Neurophysiological Approaches

Neuronal Hyperexcitability. Lithium is known to replace intracellular sodium and hyperpolarize the neuronal membrane, thereby decreasing neuronal excitability. Abnormalities in neuronal electrolyte balance (an excess of residual sodium, defined by radioisotope techniques) and hypothesized secondary neurophysiological disturbances were the focus of British investigations by Alec Coppen and colleagues in the early 1960s. The data appear compatible with the hypothesized movement of excess sodium into the neuron during an episode of mood disorder and redistribution toward the preillness electrolyte balance across the neuronal membrane during recovery. Intraneuronal sodium leakage is postulated in both depressive and manic disorders but is deemed more extreme in the latter. Because the harmonious activity of the neuronal cell and, by implication, that of a group of neurons depends on the electrical gradient maintained across its membrane by differential distribution of sodium, abnormalities in sodium concentrations and transport are hypothetically relevant to the production of an unstable state of neurophysiological hyperexcitability. In formulating their thesis of neurophysiological arousal in melancholic states, Joseph Mendels and Peter Whybrow (both of whom have worked at the University of Pennsylvania) have capitalized on the foregoing electrolyte disturbances. The view that mania represents a more extreme electrophysiological dysfunction in the same direction as depression violates the commonsense notion of symptomatological "opposition" between the two kinds of disorder, yet it may, in part, account for the existence of mixed states in which symptoms of depression and mania coexist. The NIMH team led by Frederick Goodwin first showed that a substantial minority of depressed patients with a bipolar substrate respond to lithium salts, which further supports the concept of a neurophysiological common denominator to mania and depression. Perturbations of calcium metabolism also appear relevant to bipolar patients. Therapeutic implications of this observation (e.g., the use of calcium channel inhibitors in bipolar I disorder) have not yielded consistent results. Finally, rubidium, another alkali metal, has been explored in the depressive phase of bipolar disorders, again with inconclusive results.

Rhythmopathy. European studies have shown that depressed patients are phase advanced in many biological rhythms, including the latency to the first rapid eye movement (REM) in sleep. Shortened REM latency, which has been extensively studied by David Kupfer and colleagues at the University of Pittsburgh, has been proposed as another laboratory "test" for depressive disorder. Shortened REM latency may serve as a trait marker for depression because it has been found in dysthymia and so-called borderline

personality, as well as among the clinically "well" offspring of adults with major depression.

Formulations of circadian rhythms by Thomas Wehr and Norman Rosenthal, working at NIMH, have focused on abnormalities on brain regulation of temperature, activity, and sleep cycles. Others have investigated the role of the pineal hormone melatonin in mood disorders, without achieving consistent results. The application of circadian rhythm research concepts to women with mood disorders has also led to imaginative methods, but, again, without definitive characterization of the neurophysiologic faults.

At a basic level, instrumental REM sleep deprivation in neonates has been recently shown to lead to adult "depression-like" behaviors in rats. In human studies, sleep deprivation and exposure to bright white light has been shown to correct phase disturbances and thereby terminate depressive episodes, especially in subjects with periodic and seasonal depressions. It has even been shown that the average person is light deprived, and that phototherapy can benefit even those without clear-cut seasonal patterns. Unfortunately, except for their use in mild seasonal depressions and suppression of sleep deprivation to prevent hypomania in bipolar disorder, the foregoing circadian studies have not had a palpable impact on practice. Their application in the large segment of mood-disordered patients remains cumbersome, if not elusive. Their effect can be better assessed at the level of theory. Although the specificity and efficacy of these neurophysiological indices and manipulations for clinical depression and bipolar disorder require more extensive research, cumulatively, they point to midbrain dysregulation as the likely common neurophysiological substrate of affective disorders. The foregoing considerations further suggest that the ancient Greeks, who ascribed melancholia to malignant geophysical influences, did not indulge in mere poetic metaphor. The ancients had observed the disturbed circadian patterns and advocated their readjustment to restore euthymia.

Affective Dysregulation. A major challenge for research in mood disorders is to characterize the basic molecular mechanisms that underlie the neurophysiological rhythmopathies, which, in turn, might account for the recurrent nature of the affective pathology as envisioned by Kraepelin. This means that in the most typical recurrent forms of the disorders, the constitutional foundations (manifested as cyclothymic and dysthymic traits and/or a broad range of emotional disequilibrium covered by the rubric of "neuroticism") are so unstable that the illness may run its entire course more or less autonomously, with the environment largely serving to turn on and off the more florid phases (episodes). Parisian psychiatrist Jean Delay, a pioneer in psychopharmacology in the 1950s, also emphasized affective dysregulation as the fundamental pathology in the spectrum of mood

disorders. Robert Post, at NIMH, hypothesized that the electrophysiological substrates could be so kindled that an oligoepisodic disorder initially triggered by environmental stressors could assume an autonomous and polyepisodic course. He hypothesized that this phenomenon might occur because neuronal perturbations brought about by stressors in the early course of mood disorders get incorporated into the DNA. This fascinating kindling hypothesis, however, does not seem to pertain to garden-variety mood disorders but to those with extreme cyclicity. The monograph on manic-depressive illness by Fred Goodwin and Kay Jamison presents in-depth arguments for this cyclical paradigm of thymopathy.

The amygdala, through its connection with the serotonin transporter, appears to play a crucial role in the dysregulation of the emotional life of affectively ill patients. Data indicate that the dysregulation is both in function and anatomic structure.

THEORETICAL SYNTHESIS

Pathophysiological Understanding

Modern psychobiology attempts to link experience and behavior to the central nervous system (CNS). Building conceptual bridges between the psychological and biological approaches to mood disorders requires sophisticated strategies that go beyond the Cartesian notion of limited mind–body interactions through the pineal gland and the generalizations of the Meyerian school.

William McKinney and Hagop Akiskal in 1973 developed a conceptual framework that considers the affective syndromes as the final common pathway of various psychological and biological processes. The overarching hypothesis is that psychological and biological etiological factors converge in reversible deficits in the diencephalic substrates of pleasure and reward. Those areas of the brain subserve the functions that are disturbed in melancholia and mania. The integrative model links the central chemistry and physiology of reward mechanisms with the object loss and behavioral models of depression, both of which give singular importance to the depressant role of loss of rewarding interpersonal bonds. A key element of the model is the circadian disturbances observed since ancient times in both depressive and manic syndromes. Both syndromes are conceptualized as clinical manifestations of a disordered limbic system with its subcortical and prefrontal extensions. Brain imaging studies in melancholic patients by Wayne Drevets, originally at Washington University, have visualized limbic disturbances extending into subcortical structures and occurring primarily in patients with a familial diathesis for depression; the amygdala appears to be the focal limbic structure in the latter studies. Clinical experience and research data suggest that multiple factors described in the following

subsections converge to produce or exacerbate dysregulation in these brain regions, leading to the final common pathway of clinical depression. The data on mania are more tentative and are mentioned when relevant.

Heredity. Evidence indicates a significant genetic role in the causation of bipolar and recurrent major depressive disorders. Although it is not known exactly what is inherited and biological endophenotypes have not been delineated, temperamental dysregulation might be hypothesized to fulfill the role of a behavioral endophenotype. The depressive inheritance might translate into impaired coping under stress (neuroticism), and bipolar inheritance might translate into affective dysregulation (cyclothymia), involving over- and underreactivity to life situations, circadian events, and biological stressors. Whatever the precise nature of the inherited fault or excess, current research suggests that heritability involves a broad spectrum of disorders, including milder affective states, as well as temperamental inclinations. Recent findings, from both clinical and genetic investigations, have emphasized the importance of broad affective phenotypes that incorporate panic and anxiety reactivity within both traditional unipolar and bipolar disorders. For instance, the affective dysregulation underlying bipolar disorder can manifest in euphoria, irritability, depression, panic attacks, and social anxiety. Such tendencies are observed among patients and their first-degree relatives. Genetic heterogeneity is likely and may involve inheritance of a single dominant gene with variable penetrance in some families or specific subtypes, or oligogenic inheritance in the majority of cases. Different genetic mechanisms will, in all likelihood, involve more than one disorder (e.g., depression and generalized anxiety; bipolar I disorder, psychosis, stimulant abuse, and dipsomania; bipolar II disorder, panic disorder, and bulimia nervosa). Another distinct possibility is that some forms of schizophrenia, bipolar I and II disorders, and recurrent depressions lie on an oligogenic BPS. A polymorphism involving the short alleles of the serotonin transporter gene appears relevant to depressive dispositions. The genetic mechanism involving mood reactivity and panic appears to be subserved by chromosome 18q. The partial overlap of manic and schizophrenic phenomenology might be related to a polymorphism related to a variety of genes.

It is relevant to point out at this juncture that at least five genetic loci have been identified with shared liability for both bipolar and schizophrenic disorders. These are the loci for BDNF, catechol-*O*-methyl transferase (COMT), neuroregulin (NRGl), d-amino acid oxidate activator (DAOA; also known as G72), and Disrupted in Schizophrenia 1 (DISC1). It is unlikely that these findings coming from laboratories worldwide represent systematic diagnostic error. One heuristic favored theoretical interpretation of these data is that bipolar and schizophrenic disorders

represent oligogenic inheritance, requiring multiple additive genes. A competing hypothesis is that there exist distinct mood disorder phenotypic spectra, each with their own genotypes, only some of which cross the border into psychosis; others cross the border with anxiety and panic; still others may involve "comorbidity" with migraine.

Developmental Predisposition. Parents with mood disorders are often in conflict, which may lead to separation, divorce, and suicide. It can be said that heredity often determines the type of environment into which the child predisposed to mood disorder is born. Developmental object loss, although not specifically involved in causing mood disorder, might modify the expression of the illness, possibly by leading to earlier onset, more severe episodes, and an increased likelihood of personality disorder and suicide attempts. The serotonin transporter polymorphism mentioned earlier appears to mediate the relationship between early trauma and depression. Likewise, this polymorphism appears relevant to neuroticism and suicide attempts.

Temperament. Since ancient times, persons prone to mania and melancholia have been described as possessing certain temperamental attributes, representing variations on the theme of what today is subsumed under cyclothymic, dysthymic, and anxious-inhibited temperaments, as well as the traits of high neuroticism describing emotionality. Many monozygotic twins discordant for full-blown mood disorders, studied by Aksel Bertelsen's Danish research team, exhibited affective instability with temperamental moodiness, which strongly suggests that such attributes are genetically determined. Research conducted by Kendler's team at the Medical College of Virginia further suggests that several of the temperamental attributes might be transmitted as part of the genetic liability to mood disorders. Research has also identified such temperaments in the prepubertal offspring of parents with bipolar I disorders, suggesting that they precede by years to decades the overt onset of major mood disorder episodes. The atmosphere of high expressed emotion and the negative critical remarks by relatives and affectively unstable patients documented in the recent psychological literature on mood disorders often reflect the interpersonal clashes between patients and their temperamentally intense relatives. Thus, temperaments appear to be intimately involved in generating much interpersonal friction, emotional arousal, and sleep loss (just to cite common perturbations), thereby eliciting many of the life stressors that precipitate affective episodes. The use of stimulant drugs either to self-treat lethargy or to enhance hypomanic traits could further contribute to episode precipitation. As for the depressive disposition, the work of Maria Kovacs at the University of Pittsburgh has shown that dysthymia in children evolves into major

depressive episodes postpubertally, of which a proportion switch to bipolar states. These data cohere with the work conducted at the University of Tennessee, Memphis, showing shortened REM latency in early-onset dysthymic subjects. The familial bipolar diathesis revealed in the Tennessee work, along with its tendency to switch to hypomania, suggests commonalities between depressive and bipolar II disorders.

Life Events. Most individuals do not develop clinical depression when exposed to environmental adversity. Such adversity seems to play a pathogenic role primarily in those with an affective diathesis. In fact, the work of Kendler at the Medical College of Virginia indicates that genetic factors might underlie the depressive disorder patients' susceptibility to life events. Furthermore, current data suggest that social stressors in the onset of depression are more relevant to the first few episodes of the illness. The evidence linking such events to mania is less convincing. At any rate, stressful events often appear to be triggered by the temperamental instability that precedes clinical episodes. Interpersonal losses are common events in the lives of individuals with intense temperaments. The arousal and sleep loss associated with such events can precipitate both depressive and manic states.

A recent study by Peter McGuffin's team at the Institute of Psychiatry, London, raised the possibility that one mechanism by which heredity produces depression is the creation of environmental adversities in the lives of individuals predisposed to this illness. This work has been replicated by independent groups of investigators. Whatever the origin of environmental adversity, it is common clinical experience that loss represents an important, perhaps even central, theme in clinical depression. Variables that seem to modulate the effect of adult losses include concurrent life events, resultant changes in lifestyle, lack of interpersonal support, deficient social skills, and the symbolic meaning of the putative loss. The research program of George Brown and his followers in London capitalizes on the foregoing considerations, particularly the importance of early and proximate losses in socioeconomically disadvantaged women who lack supportive relationships. However, that conceptualization downplays the degree to which the social context of the depression reflects the dysthymic temperamental liabilities of those depressed women. Recent research indicates that even social support is determined to a considerable degree by the genetic mechanisms that underlie mood disorders. Indeed, the short alleles of the serotonin transporter gene are now implicated in mediating between adverse events and clinical depression.

Biological Stressors. Many physical diseases and pharmacological agents are known to precede the onset of both depressive and manic

episodes. Like psychosocial stressors, however, they do not generally seem to cause de novo episodes, but mobilize them in persons with a personal and family history of mood disorders. Thyroid disturbances have a role in practice because they are associated with rapid cycling in bipolar patients, especially women; lithium is often contributory to such disturbances occurring in the depressions in which bipolar women are not uncommonly "stuck."

Sex. Clinical and epidemiological studies concur in suggesting that women are at higher risk for mood disorders, with the risk highest for depression. This now appears to be, in part, a function of anxious-depressive traits represented by neuroticism. These traits have strong genetic determinants. Women have higher concentrations of MOA (the enzyme that breaks down monoamine transmitters) in the brain and more precarious thyroid status. In addition, low estrogen and high progesterone concentrations have been postulated as possible mediating factors in postpartum depressions, premenstrual accentuation of affective instability, and women's vulnerability to the depressant effect of steroidal contraceptives. Finally, recent data point to the role of estradiol in depressions occurring during the transition to menopause. Personality factors might also be relevant to the sex differences in depression. In recent collaborative work with University of Pisa psychiatrist Giulio Perugi, Akiskal proposed the hypothesis that female sex might favor greater expression of dysthymic attributes, whereas hyperthymic traits appear to be favored by male sex. Those considerations tend to parallel, respectively, the *ruminative* and *active* cognitive response styles reported by Susan Nolen-Hoeksema, originally at Stanford University, to distinguish the sexes. What specific sex-related biographical factors might interact with sex-related biological factors to produce such trait differences is largely unknown. An intriguing possibility is that women, because of temperamental inclination to depressive cognitions, might react more intensely to childhood adversities, as well as be more specifically vulnerable to adult stressors related to bonding with men and child rearing. Research by Mark George and colleagues has raised the provocative possibility that women overrespond to sad circumstances over a lifetime, thereby permanently altering anterior limbic and prefrontal brain function in a "depressive" direction. The high prevalence of anxious-depressive conditions in women—most pronounced at the clinical level—might be linked, hypothetically, to an evolutionary adaptive advantage conferred by traits of fear, inhibition, and avoidance to women who bear the responsibility of pregnancy and child rearing. The higher prevalence of minor depressions in women, rediscovered in a recent Danish study, is in line with the foregoing hypothesis.

The integrative model presented here (Fig. 13.1–3) goes beyond the

general provisions of the unified approach developed three decades ago. It is submitted that, at least in the highly recurrent forms of the malady, affective temperaments represent the intermediary stage between remote (hereditary) and proximate (stressful) factors and that limbic-diencephalic dysfunction is best characterized as the biological concomitant of the clinical manifestations of the affective syndromes. Like the temperamental dysregulations, these biological disturbances represent a putative stage in the pathogenetic chain. They emerge as temperamental instabilities that react to, provoke, or invite life events, substance use, and alterations in circadian rhythms—which, in turn, appear to usher in the behavioral, emotional, and cognitive manifestations of the illness. It is finally relevant to point out that biological stressors such as hormonal disturbances and traumatic brain injury appear to compromise limbic-diencephalic function as their depressant mechanism.

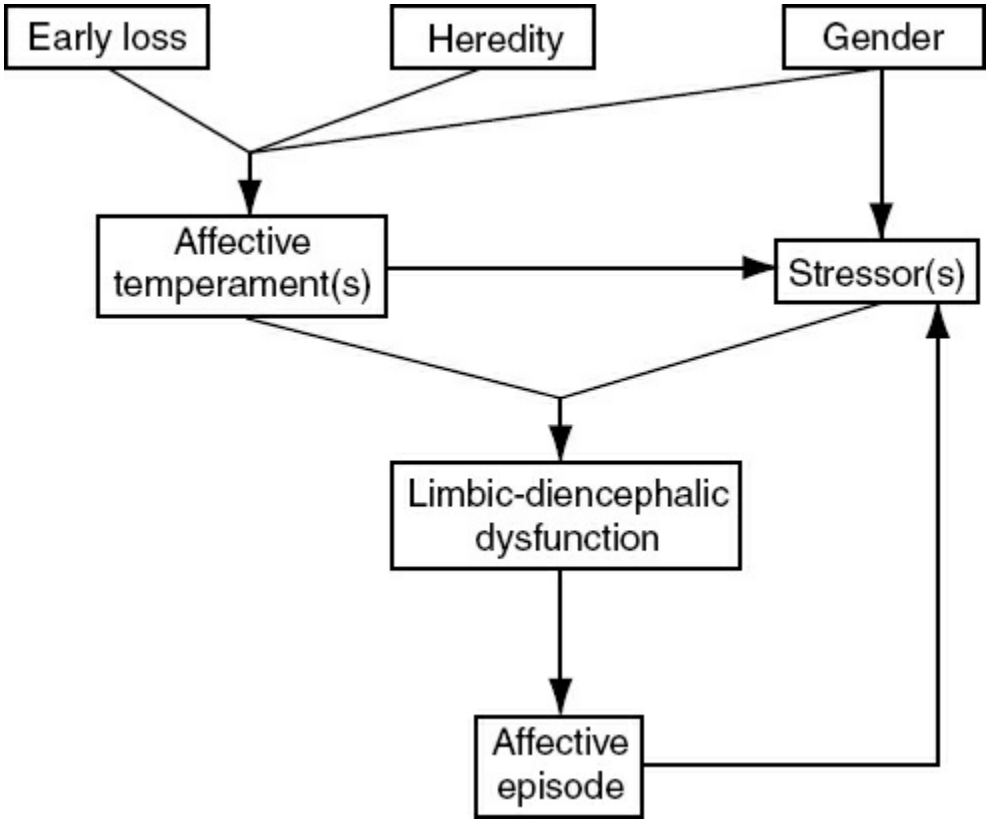

FIGURE 13.1–3. An integrative pathogenetic model of mood disorders.

Therapeutic Perspectives

The foregoing integrative model mandates the joint use of somatic-pharmacological and psychosocial interventions. Although the milder forms of mood disorders can be managed with psychotherapy, somatic treatments are usually required to reverse the biological disturbances in melancholia before the patient can respond to interpersonal feedback. Depressive disorders with psychotic features often necessitate more

definitive somatic interventions, such as electroconvulsive therapy (ECT). Continued psychopharmacological treatment is also effective in decreasing rates of relapse and future recurrence in most. Bipolar disorder and most forms of depression are considered lifelong illnesses needing indefinite maintenance pharmacotherapy.

Psychosocial therapy by skilled clinicians can provide support, combat demoralization, change maladaptive self-attributions, and improve conjugal and vocational functioning. Recent fascinating data support the notion that psychosocial interventions such as cognitive–behavioral therapy modulate cortical-limbic function described earlier as the final common pathway of depressive illness. Whether such therapy can also modify personality traits to fortify the patient against new episodes is a future research challenge. It might prove more profitable to attempt to help patients to explore professional and object choices that match their temperamental proclivities and assets, which, in turn, might provide them greater harmony and adaptation in life. Although much needs to be learned about the indications for medication and psychotherapy in different subtypes of mood disorders, research not only does not support a negative interaction between the two forms of treatment, but, on selected parameters, suggests additive and even synergistic interaction. There is a great need for patients, their families, and clinicians to understand how a biologically driven illness like depression should be approached from a pragmatic psychotherapeutic perspective.

The challenge for psychiatric research in the decade ahead is to elucidate the basic mechanisms by which the predisposing, precipitating, and mediating variables reviewed here and others yet to be identified interact to produce the final common path of decompensation in melancholia. Because of the heterogeneity of depressive conditions presenting as a psychobiological final common clinical syndrome and because antidepressant agents, irrespective of specificity to one or another biogenic amine, are approximately equally effective in two-thirds of those with depressive disorders, the antidepressant agents may be acting not on the primary lesions of these disorders but on a neurochemical substrate distal to the underlying biological faults. Choice of antidepressants is still highly determined by the side effect profile least objectionable to a given patient's physical status, temperament, and lifestyle. That so many different classes of antidepressants—with different mechanisms of action—have been marketed since the 1990s represent indirect evidence for heterogeneity of putative biochemical lesions. The investigation of central neurotransmitter receptor function continues to occupy much current effort to delineate the mechanism of antidepressant action and side effects of classic agents, as well as the new compounds that have made the treatment of depression "clinician and patient friendly." Whether study of specific receptors will unravel the molecular mystery of depression remains

to be seen. Since the 1990s, studies have begun on antidepressant and mood-stabilizing effects on molecular mechanisms believed to be closer to the "genetic underpinnings" of mood disorders. Herein is the promise of the future—a new generation of psychiatrists conversant with both clinical phenomenology and molecular biology. Data suggest that the biological specificity of genetic factors in mood disorders might be translated into distinct temperamental dysregulations, which, in turn, might predispose to different affective subtypes.

Returning to the therapeutic arena, mounting clinical evidence indicates that, in a special subgroup of depressed patients with bipolar disorder, antidepressants might provoke mixed episodes, hypomanic episodes, or both, and possibly increase later cycling. The kindling-sensitization model suggests the utility of anticonvulsant medication on episode escalation and might represent yet another example of pathophysiological intervention. Whatever the merit of this model, the 1990s witnessed intense clinical and research interest and U.S. Food and Drug Administration (FDA) approval of the clinical introduction of divalproex (Depakote) and lamotrigine (Lamictal) for bipolar disorder, and many other promising anticonvulsants are being developed for that disorder. Anticonvulsant mood stabilizers appear to possess a broad spectrum of activity on bipolar disorders, including mixed dysphoric and rapid-cycling forms. Lithium, by contrast, seems more specific to euphoric or "classic" mania. On the other hand, the introduction of atypical antipsychotics—olanzapine (Zyprexa), risperidone (Risperdal), quetiapine (Seroquel), and aripiprazole (Abilify)—for mania raises intriguing questions about a common neural substrate for schizophrenic and manic disorders. Psychoeducational interventions geared to disturbed rhythms of the disorder represent another example of rational therapeutics. Mood clinics should help patients and their significant others to dampen stimulation so that it is kept at an optimal level for depressed patients with cyclothymic traits. All offending drugs (e.g., cocaine, caffeine, and sedative-hypnotic agents) should be gradually eliminated and circadian disruptions and sleep loss minimized. The greater challenge is to learn how to curb the ill-advised actions of patients with cyclical depressions. Unfortunately, current data indicate neurocognitive deficits, particularly in executive function, that underlie and/or complicate these disorders. Psychoeducation and psychotherapy have the task of ameliorating the resulting social problems, and despite much progress in this realm, these deficits remain stable over time. Compliance with mood-stabilizer regimens that, for many, would attenuate episodes and prevent such sequelae is difficult to achieve. Further research on treatment or medication-adherence techniques is needed for promoting more efficient use of mood stabilizers.

It is tempting to suggest that biogenic amines, the "humors" of modern psychobiology, play the same heuristic role as the ancient humors did for

many centuries. The black humor, appropriately evoked in the construct of melancholia in DSM-5, may not have the same claim for etiological relevance to depressive disorders as norepinephrine and serotonin, but at least it has a classic heritage. Dopamine, by contrast, may represent the sanguine humor that drives hypomanic temperaments and manic behavior. When genetic factors contributing to clinical depression and mania are discovered, in all likelihood, they will be more linked to temperamental dispositions than to full-blown affective disease phenotypes. The clinician will still need to interpret the myriad influences that effect such inclinations to produce disease in an individual patient—that is, fundamental scientific advances in mood disorders, rather than diminishing the role of practitioners, will actually increase it. It is to be regretted that, despite destigmatization efforts and a rich armamentarium of therapeutic developments, the clinical care of affectively ill patients at the severe end of the spectrum continues to be grossly inadequate. More research will not improve this dismal situation unless the human and social dimensions of severe mood disorders are addressed with the requisite clinical and public health policies. Caring for patients with mental illness is a dimension distinct from evidence-based treatments along psychopharmacological and psychotherapeutic lines: It requires the allocation of human resources and integrated mental health structures geared to the total patient.

In any discipline, scientific truth is a function of its technology, but understanding the phenomena under consideration is a matter of philosophical temperament that seeks integration and the hope for a unified vision. Research into the causes and treatment of mood disorders has generated abundant recent data suitable for integration into theory and practice, and conceptualizing the origin and treatment of mood disorders can no longer be justified on the grounds of ideological preference alone.

References

Adams F, ed. *Aretaeus of Cappadocia: The Extant Works of Aretaeus, the Cappadocian*. London: Sydenham Society; 1856.

Ahrens B, Muller-Oerlinghausen M, Schou M, Wolf T, Alda E. Excess cardiovascular and suicide mortality of affective disorders may be reduced by lithium prophylaxis. *J Affect Disord*. 1995;33:67.

Akiskal HS. Dysthymia and cyclothymia in psychiatric practice a century after Kraepelin. *J Affect Disord*. 2001;62:17.

Akiskal HS, Akiskal KK. In search of Aristotle: Temperament, human nature, melancholia, creativity and eminence. *J Affect Disord*. 2008;100:1.

*Akiskal HS, McKinney WT. Depressive disorders: Toward a unified hypothesis. *Science*. 1973;182:20.

Akiskal HS, Yerevanian BI, Davis GC, King D, Lemmi H. The nosologic status of borderline

personality: clinical and polysomnographic study. *Am J Psychiatry*. 1985;142:192.

Berretta S, Pantazopoulos H, Lange N. Neuron numbers and volume of the amygdala in subjects diagnosed with bipolar disorder or schizophrenia. *Biol Psychiatry*. 2007;62:884.

*Bertelsen A, Harvald B, Hauge M. A Danish twin study of manic-depressive disorders. *Br J Psychiatry*. 1997;130:330.

Biederman J, Mick E, Faraone SV, Spencer T, Wilens TE. Pediatric mania: a developmental subtype of bipolar disorder? *Biol Psychiatry*. 2000;48:458.

Brandon AR, Shivakumar G, Freeman MP. Depression: covering mood and vasomotor symptoms. *Curr Psychiatry*. 2008;7:39.

*Caspi A, Sugden K, Moffit TE, Taylor A, Craig IW. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*. 2003;301:386.

Cytryn L, McKnew D. *Growing Up Sad: Childhood Depression and Its Treatment*. New York: WW Norton; 1996.

Davidson RJ, Scherer KR, Goldsmith HH, eds. *Handbook of Affective Sciences*. Oxford: Oxford University Press; 2003.

Delay J. *Les Dèréglements de l'Humeur*. Paris: Presses Universitairesde France; 1946.

Depression Guideline Panel. *Depression in Primary Care (Clinical Practice Guideline, Number 5)*. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; 1993.

Frasure-Smith N, Lesperance F. Depression—A cardiac risk factor in search of a treatment. *JAMA*. 2003;289:3171.

Freud S. Mourning and melancholia. In: *Standard Edition of the Complete Psychological Works of Sigmund Freud*. Vol. 4. London: Hogarth Press; 1975.

Gershon ES, Hamovit J, Guroff JJ, Dribble E, Leckman JF. A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. *Arch Gen Psychiatry*. 1982;39:1157.

*Goodwin FK, Jamison KR. *Manic-Depressive Illness*. 2nd ed. New York: Oxford University Press; 2007.

Hariri AR, Mattay VS, Tessitore A, Kolachana B, Fera F. Serotonin transporter genetic variation and the response of the human amygdale. *Science*. 2002;19:400.

Jorge RE, Robinson RG, Moser D, Tateno A, Crespo-Facorro B. Major depression following traumatic brain injury. *Arch Gen Psychiatry*. 2004;61:42.

Kendler KS. Social support: a genetic-epidemiological analysis. *Am J Psychiatry*. 1997;154:1398.

Kendler KS, Karkowski-Shuman L. Stressful life events and generic liability to major depression: genetic control of exposure to the environment? *Psychol Med*. 1997;27:359.

Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. A longitudinal twin study of personality and major depression in women. *Arch Gen Psychiatry*. 1993;50:853.

Klerman GL, Lavori PW, Rice J, Reich T, Endicott J. Birth-cohort trends in rates of major depressive disorder among relatives of patients with affective disorder. *Arch Gen Psychiatry*. 1985;42:689.

Kovacs M, Akiskal HS, Gatsonis C, Parrone PL. Childhood-onset dysthymic disorder: clinical features and prospective naturalistic outcome. *Arch Gen Psychiatry*. 1994;51:365.

*Kraepelin E. *Manic-Depressive Insanity and Paranoia*. Edinburgh: Livingstone; 1921.

Krishman KR. Pituitary size in depression. *J Clin Endocrinol Metab*. 1991;72:256.

Kupfer DJ. REM latency: a psychobiologic marker for primary depressive disease. *Biol Psychiatry*. 1976;11:159.

Kuswanto C, Chin R, Sum MY, et al. *The Psychoses Dissected: The Kraepelinian Dichotomy, a Continuum or One Disorder? Program No. S4.* Annual Meeting of the American College of Neuropsychopharmacology, Boca Raton, FL; 2007.

Kuswanto C, Chin R, Sum MY, et al. Shared and divergent neurocognitive impairments in adult patients with schizophrenia and bipolar disorder: whither the evidence? *Neurosci Biobehav Rev*. 2016;61:66–89.

Lewinsohn PM, Youngren MA, Grosscup SJ. Reinforcement and depression. In: Depue RE, ed. *The Psychobiology of Depressive Disorders; Implications for the Effects of Stress*. New York: Academic Press; 1979.

Lewis A. States of depression: their clinical and aetiological differentiation. *Br Med*. 1938;J2:875.

Lozano AM, Mayberg HS, Giacobbe P, Hamani C, Craddock RC. Subcallosal cingulated gyrus deep brain stimulation for treatment-resistant depression. *Biol Psychiatry*. 2008;64:461.

MacQueen GM, Hajek T, Alda M. The phenotypes of bipolar disorder: relevance for genetic investigations. *Mol Psychiatry*. 2005;10:811.

Malhi GS, Byrow Y, Fritz K, et al. Mood disorders: neurocognitive models. *Bipolar Disord*. 2015;17(Suppl 2):3–20.

McMahon FJ, Simpson SG, McInnis MG, Badner JA, MacKinnon DF. Linkage of bipolar disorder to chromosome 18q and the validity of bipolar II disorder. *Arch Gen Psychiatry*. 2001;58:1025.

Merikangas KR, Akiskal HS, Angst J, Greenberg PE, Hirschfeld RM. Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2007;64:543.

Mur M, Portella MJ, Martinez-Aran A, Pifarre J, Vieta E. Long-term stability of cognitive impairment in bipolar disorder: a 2-year follow-up study of lithium-treated euthymic bipolar patients. *J Clin Psychiatry*. 2008;69:712.

Murray CJL, Lopez AD, eds. *The Global Burden of Disease*. Geneva: World Health Organization; 1996.

Oedegaard KJ, Fasmer OB. Is migraine in unipolar depressed patients a bipolar spectrum trait? *J Affect Disord*. 2005;84:233.

Perugi G, Musetti L, Simonini E, Piagentini F, Cassano GB. Gender mediated clinical features of depressive illness: the importance of temperamental differences. *Br J Psychiatry*. 1990;157:835.

Petty F. GABA and mood disorders: a brief review and hypothesis. *J Affect Disord*. 1995;34:275.

Post RM. Transduction of psychosocial stress into the neurobiology of recurrent affective disorder. *Am J Psychiatry*. 1992;149:999.

Prange AJ Jr, Wilson IC, Lynn CW, Alltop LB, Stikeleather RA. l-Tryptophan in mania: contribution to a permissive hypothesis of affective disorders. *Arch Gen Psychiatry*. 1974;30:56.

Puig-Antich J. Affective disorders in children and adolescents. In: Meltzer HY, ed. *Psychopharmacology: The Third Generation of Progress*. New York: Raven; 1987.

Recupero PR, Harms SE, Noble JM. Googling suicide: surfing for suicide information on the

Internet. *J Clin Psychiatry*. 2008;69:878.

Rihmer Z, Akiskal H. Do antidepressants t(h)reat(en) depressives? Toward a clinically judicious formulation of the antidepressant-suicidality FDA advisory in light of declining national suicide statistics from many countries. *J Affect Disord*. 2006;94:3.

Seligman MD. *Helplessness: On Depression, Development and Death*. San Francisco: Freeman; 1975.

Torres F, Barbosa M, Maciel P. Recurrent copy number variations as risk factors for neurodevelopmental disorders: critical overview and analysis of clinical implications. *J Med Genet*. 2016;53(2):73–90.

*Widlocher DJ. Psychomotor retardation: clinical, theoretical, and psychometric aspects. *Psychiatr Clin North Am*. 1983;6:27.

Zaninotto L, Solmi M, Toffanin T, Veronese N, Cloninger CR, Correll CU. A meta-analysis of temperament and character dimensions in patients with mood disorders: comparison to healthy controls and unaffected siblings. *J Affect Disord*. 2016;194:84–97.

▲ 13.2 Mood Disorders: Epidemiology

KATHLEEN RIES MERIKANGAS, PH.D., AND ZOLTáN RIHMER, M.D., PH.D., D.SC.

Application of the tools of epidemiology to psychiatry over the past several decades have led to numerous advances in the understanding of the magnitude, correlates, patterns of comorbidity, course, impact, and service patterns of mood disorders. Advances in epidemiology were comprised of methodologic developments including the introduction of structured and semistructured diagnostic interviews, statistical methods for estimating prevalence and correlates of mental disorders, and the shift from clinical to population-based samples to obtain estimates of the magnitude and correlates of mental disorders unbiased by treatment seeking. The results of recent epidemiological studies that have shown the high prevalence of subthreshold conditions and pervasive comorbidity between disorders have illustrated the need for further development of the psychiatric diagnostic system. There is now widespread agreement regarding the need to expand the current categorical diagnostic system because of the arbitrary thresholds that distinguish disorders, and the lack of distinct boundaries between different classes of disorders.

During the past decade the descriptive epidemiology of mood disorders has come to maturity. In recent years there has been a proliferation of epidemiologic research, both in the United States and abroad, that has strengthened the evidence based on the magnitude, correlates and consequences of mood disorders including bipolar disorder and major depressive disorder. This work has highlighted the dramatic personal and societal impact of bipolar disorders. This section will provide an update on recent large scale studies of the prevalence of mood disorder subtypes in

adults and children, and summarize the correlates, patterns of comorbidity, impact, and service patterns for mood disorders from population-based studies. Findings show substantial concordance with those reported in earlier volumes of this textbook, and highlight the need to shift from cross-sectional descriptive epidemiology to analytic epidemiology in order to identify risk factors, correlates and consequences of these disabling conditions.

DESCRIPTIVE EPIDEMIOLOGY

Bipolar Disorder

The largest series of international studies that provide prevalence rates of mood disorders is the World Health Organization (WHO) World Mental Health (WMH) surveys. Data on bipolar disorder collected in a subset of 11 population-based surveys were carried out in the Americas (São Paulo metropolitan area, Brazil; Colombia; Mexico; and United States), Europe (Bulgaria and Romania), Asia (Shenzhen, People's Republic of China; Pondicherry region, India; and 9 metropolitan areas in Japan), Lebanon, and New Zealand. With the exception of Japan (unclustered 2-stage probability sample), all surveys were based on stratified multistage clustered area probability samples. In the past few decades, the lifetime prevalence of Bipolar I (BPI) disorder has generally estimated at about 1 percent. The most recent estimates from the largest population-based survey, the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) have reported substantially greater lifetime prevalence estimates of bipolar I disorder of 3.3, Bipolar II (BPII) of 1.1 with corresponding 12 month rates of 2.0 and 0.8, respectively. Recent studies have begun to include estimates for bipolar II disorder and BPS disorders as well as bipolar I disorder. As would be expected, prevalence tends to increase with successively more inclusive disorder definitions. For example, in the WMH Surveys, the cross-national lifetime prevalence of BPS disorders was 2.4 percent; of which 0.6 percent met criteria for bipolar I disorder, 0.4 percent for BPII, and 1.4 percent for subthreshold bipolar disorder. In many (but not all) studies assessing bipolar II disorder, its prevalence is actually lower than that of bipolar I disorder; this may indicate that the majority of people who experience hypomanic episodes also experience manic episodes, or may be explained by the major depressive episode requirement for diagnosis of bipolar II disorder.

An increasing number of studies have also begun to estimate the occurrence of bipolar disorder in youth. Lifetime prevalence rates of bipolar disorder in youth range from 0.2 percent in the Great Smoky Mountains Study to 2.9 percent for bipolar I disorder or bipolar II disorder in the National Comorbidity Survey-Adolescent supplement (NCS-A). Young adults aged 19 to 24 in the Canadian Community Health Survey had

a lifetime prevalence of 3.8 percent. A few studies have also estimated 12-month prevalence rates of bipolar disorder ranging from 1.3 to 2.5 percent. Prevalence rates of mania range from 0.4 (12-month) to 2.0 percent (12-month); and rates of hypomania range from 0.1 (lifetime) to 0.9 percent (6-month). The results of longitudinal studies converge in estimating the prevalence of bipolar disorder at between 1.4 and 2.1 percent, which approximates cross-sectional prevalence rates in adult samples. This was confirmed by a recent meta-analysis of epidemiologic studies of bipolar disorder in children and adolescents, which reported a mean prevalence of 1.8 percent.

Major Depression

The WMH surveys also provided the largest aggregate data on the prevalence of major depressive disorder. As shown in Figure 13.2–1, lifetime and 12-month prevalence rates were estimated in 18 countries, divided according to high and low-middle income. The lifetime prevalence estimates average 11.1 (range 8.0 to 18.4) in low and 14.6 (range 6.6 to 21.0) in high-income countries, whereas the 12-month prevalence rates average 5.5 in high (range 2.2 to 8.3) and 5.9 (range 3.8 to 10.4) in low-income countries. More recent prevalence estimates from the NESARC study were 13.2 for lifetime and 5.3 for 12-month major depressive disorder. These aggregate findings suggest that the descriptive epidemiology of mood disorders has come to maturity. Although there is a wide range of estimates, the average rates of both lifetime and 12-month depression are fairly consistent across studies that employ comparable methodology.

Several cross-sectional and prospective studies have also reported rates of major depressive disorder in youth. The lifetime rates of major depressive disorder in childhood range from about 0.6 to 4.8 percent with a median of 2.2 percent. Results of the NCS-A study of adolescents in the United States yielded lifetime and 12-month prevalence of major depressive disorder of 11.0 and 7.5 percent, respectively. Both the TRacking Adolescents' Individual Lives Survey (TRAILS) study in the Netherlands also characterized major depressive disorder with severity based on impairment. As expected the lifetime rates in these studies were substantially lower than those of nonsevere major depressive disorder with lifetime and 12-month rates of 3.0 and 2.3 percent, respectively. In the NCS-A, the prevalence of major depressive disorder increased significantly across adolescence, with markedly greater increases among females than among males. Most cases of major depressive disorder were associated with psychiatric comorbidity and severe role impairment, and a substantial minority reported suicidality. Treatment in any form was received by the majority of adolescents with 12-month DSM-IV major depressive disorder (60.4 percent), but only a minority received treatment that was disorder-

specific or from the mental health sector.

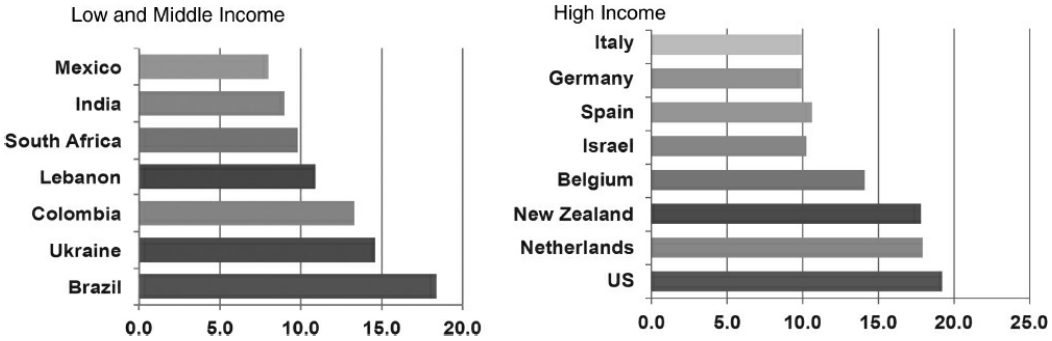

FIGURE 13.2–1. Lifetime prevalence of major depression in the world mental health.

Numerous studies have begun to expand the diagnostic criteria for mental disorders to collect information on the spectra of expression of particular conditions. With respect to mood disorders, the spectrum concept of bipolar disorder long advocated by clinicians has now been validated in community samples of adults and children. Application of the spectrum concept of bipolar disorder has yielded rates of bipolarity of up to 5 percent and depression of up to 20 percent. Likewise, many individuals in the community may show some (a few or more) depressive symptoms that do not reach the severity or duration threshold for specific mood disorders in the DSM-5 system but, nevertheless, also have substantial morbidity and dysfunction. Although these subthreshold disorders may be less severe forms of major depressive or bipolar disorders, they may also cause substantial suffering as well as disability. Subtypes of subthreshold depression identified in the Zurich Cohort Study and replicated in numerous subsequent studies include recurrent brief depressive disorder and minor depressive disorder (as well as recurrent brief hypomania).

CORRELATES OF MOOD DISORDERS

Gender

The most consistent finding across all of the studies on the prevalence and incidence of unipolar major depression is that it is approximately twofold more common among women than men. This gender difference begins in early adulthood, is most pronounced in people between the ages of 30 and 45, and also persists in the elderly. Because there are no data (aside from biological–hormonal differences) that show that female gender per se means increased vulnerability for depression, increased stress sensitivity, maladaptive coping strategies, and multiple social roles (all of which are frequently seen in women) and substance use disorders that can mask depressive symptoms (more frequently seen in men) have been suggested for the explanation of the gender difference. In addition to these psychosocial theories, recent studies show that because prior anxiety

disorders are more common in women, preceding anxiety disorder may also be a significant factor contributing to the higher depressive morbidity in women. Thus, gender difference in unipolar major depression seems to result from the complex interaction of the mentioned biological and psychosocial variables. Minor depressive disorder and recurrent brief depressive disorder are also more common among women, but the difference is not so marked as that in major depression.

In contrast to unipolar depression, the gender ratio in bipolar disorders (all subtypes combined) is approximately 1:1. However, among bipolar II disorder patients and in special subpopulations (mixed/dysphoric mania, mixed depressive episode, winter depression, bipolar depression with atypical clinical features, rapid cycling bipolar disorder), women are overrepresented. Looking at the depression–mania continuum as a whole spectrum, there is a clear trend: the higher the depressive component, the higher the proportion of women. Consequently, in the rare cases of unipolar mania (manic episode without any major or minor depression), men are markedly overrepresented. The gender differences in lifetime prevalence rates are more marked than in 1-year and current prevalence rates, which may be attributable—at least in part—to the stronger male tendency to forget previous episodes and deny emotionally negative events. Recent population-based epidemiological surveys showed that the lifetime prevalence and 1-year prevalence of major depression, dysthymia, and bipolar (I + II) disorders were much higher among people with same-sex sexual behavior, (including bisexuality) particularly in the case of men.

Age

Depressive disorders show much higher lifetime prevalence among people younger than 45 years, but the age of onset differs between unipolar and bipolar and familial and nonfamilial cases. The average age of onset of recurrent unipolar major depressive episode falls between the ages of 30 and 35 years, whereas single-episode major depression that often lacks family history of mood disorders usually begins some years later. The genetic predisposition for depression decreases with age. However, social stressors appear to place younger individuals at a greater risk for depression than elderly ones. On the other hand, isolation, loss of interpersonal contacts, medical disorders, and disability play a more important role in the development of depression in later life. Early-onset depression is associated with a higher female to male ratio than late-onset depression. Because major depressive disorder is a frequent and highly recurrent illness, the probability of recurrence does not decrease with age. The incidence of major depressive disorder in old age is lower in both sexes, but first incidence and prevalence of minor depressive disorder show the opposite trend. The incidence of unipolar major depressive disorder in the postpartum period is just slightly increased, but this is not the case in

the postmenopausal years. Pure (primary) dysthymic disorder typically starts in late adolescence or early adulthood, but, in the absence of appropriate treatment, almost all cases develop later into major depression. The first incidence of dysthymia is also common in old age, but this is typically the consequence of the adverse psychosocial and biological conditions frequently seen among the elderly.

The age of onset of bipolar disorder most commonly around 20 years of age, is substantially (about 10 years) lower than that of unipolar depression. Bipolar men appear to have 4 to 5 years earlier onset than bipolar women. In more than half of the cases, onset is before the age of 20, frequently in late adolescence. In contrast to this, first-onset mania is very rare among elderly people. The ages of onset of bipolar I disorder and bipolar II disorders are similar, but there is a slight tendency for higher age of onset in bipolar II disorder patients. The incidence of the depressive phase of bipolar disorder after childbirth is relatively high, and the majority of patients with "postpartum depression" have bipolar disorder. Like unipolar depression, bipolar patients with positive family history of mood disorders are significantly younger at the beginning of the illness and need less stressors to precipitate the illness than those lacking such history.

Race and Ethnicity

It is well known that cultural differences may influence the clinical presentation of depression and mania, leading to unreliable figures in prevalence studies performed by standard methodology. However, there is now consistent evidence that rates of mood disorders are lower in blacks and Hispanic than in whites. Because of their increased exposure to psychosocial stress and other risk factors for mood disorders, the lower rates in these ethnic subgroups is paradoxical.

SOCIAL CORRELATES

Marital Status

The relationship between marital status and mood disorders is quite complex. For example, being single, divorced, or separated can be either a risk factor for depression or the result of the adverse life events generated by depressive or manic psychopathology, or both. Major depressive disorder and bipolar illness are most frequent among divorced, separated, or widowed individuals. Single women have lower rates of depression than married women do, but the opposite is true for men. However, being single as a result of having never married, as a result of the dissolution of a difficult marriage, and as a result of widowhood represent three very different conditions. The risk of an major depressive episode is very high among recently widowed individuals of all ages, but is particularly high in the elderly. Patients with mood disorder are overrepresented among the

divorced, and the rate of family breakdown (separation, divorce) is elevated slightly in dysthymic patients, substantially in major depressive patients, and markedly in bipolar I disorder and bipolar II disorder patients. The presence of mania, hypomania, or major depression is a strong predictor for future separation or divorce, which can cause serious distress for the patients and for their spouses and may generate negative life events for their children. These early negative life events (e.g., parental loss before adolescence) are well-known predisposing factors for adult mood disorders, particularly in the case of family loading (case of subjects with positive family history of mood disorders).

Socioeconomic Factors

Although the relationship between depressive symptoms and low social class is well documented, most studies found only a weak (but consistent) correlation between major depressive disorder or bipolar I disorder and lower socioeconomic status. Individuals with lower socioeconomic status have a lower level of educational, lower income, and poorer living conditions, as well as a higher rate of unemployment and, ultimately, homelessness. As has been demonstrated in the NCS, the proportion of major depressive episode was approximately three times higher among individuals without a workplace than among those with one. However, as in the case of marital status and mood disorder, cause and effect may be reversed here, too. Mood disorder can easily lead to unemployment, divorce, or low income, resulting in regression on the social hierarchy scale. Because hypomania is not as disruptive as mania, in terms of academic and social carrier, the educational level of bipolar II disorder patients is above the average and, in contrast to unipolar major depression and bipolar I disorder, bipolar II disorder patients tend to belong to higher social classes and are relatively overrepresented among socially active, creative people.

Residence

As urban communities are more stressful than rural ones, it is not surprising that most studies performed in Western societies concluded that major depression was more frequent in urban residents than in their rural counterparts. In the ECA study, this rural–urban difference also persisted after controlling for marital status, race, and socioeconomic status. The results of the NCS show that respondents living in rural areas had approximately 40 percent lower odds of 1-year comorbidity of three or more mental disorders than did those living in urban areas. The urban–rural distinction provides little infomation about the real living and social conditions in general, but it can be a good marker for the density of the population and for other important sociological variables.

Seasonal Factors

Despite the fact that more than two-thirds of patients with recurrent major mood disorders show irregular seasonal patterns individually, statistically, spring and fall are the peak times for depression, just as summer is for mania. Because seasonal affective disorders (fall–winter depression and spring–summer depression) occur in approximately 20 to 25 percent of the patients with recurrent major mood disorders, it is possible that the seasonal pattern observed in unipolar depression and bipolar disorder in general is the consequence—at least in part—of the characteristic annual rhythms of these specific seasonal subtypes. The seasonal profiles of committed and attempted suicides, the prescription of antidepressants and ECT, and the availability of L-tryptophan (the main precursor of serotonin) are very similar to the seasonal onset of major depression. However, acute and long-term pharmacotherapy of mood disorders (more precisely, its cessation) can change the seasonal pattern of depression and mania, which should be taken into account either in planning studies or in the interpretation of the results. On the other hand, the decreasing seasonal variation in suicides that has been observed in many countries in the last two decades may reflect, among others, the reduced rate of depression-related suicides in the given population.

Geographic Trends

There is a general, but weak, trend for lower prevalence of depression and higher rate of mania in regions located closer to the Equator. Consistently, at least in the Northern Hemisphere, winter depression (which affects between 1 and 6 percent of the community) seems to be more frequent in countries situated farther from it. A significant positive correlation was found between prevalence of winter depression and latitude in North America (where its prevalence is twice as high as in Europe), whereas a similar, but only slight, tendency was observed in Europe. On the other hand, the distribution of summer depression across the latitudes shows the opposite tendency. However, the relationship between these two forms of seasonal affective disorder and latitude (e.g., the length of daily photoperiod) is weak. Other climatic and genetic influences, as well as social and cultural factors—which can also be interrelated with the daily photoperiod—may play a role, too. In the NCS-R, major depressive disorder (seasonal and nonseasonal forms combined) was largely unrelated to geography. Regarding the West–East dimension, the lower prevalence rates of major depressive and bipolar I disorders reported from Far Eastern countries may be related primarily to cultural differences and methodological shortcomings, but there is some recent evidence indicating that the prevalence of major mood disorders in these countries may, in fact, be lower.

PSYCHOSOCIAL FACTORS

Social Stressors

Social stressors, in general, have been well recognized as risk factors for mood disorders. However, different kinds of social stressors (i.e., childhood vs. adulthood events, acute vs. chronic stressors, positive vs. negative life events) can play different roles in the predisposition and precipitation of depressive or manic disorders. In the case of major depression and bipolar disorder, the association of acute stressors and the onset of illness become progressively weaker with the increasing number of previous episodes, and patients at high genetic risk for mood disorders commonly experience depressive or manic episodes without any negative life event. In the development of major mood disorders, chronic stressors (e.g., unemployment, difficult marriage) play a more important role than specific, acute stressors. However, accumulation of stressful negative life events is the strongest predisposing factor. The higher level and, probably, the different nature of social stressors in individuals living in urban communities may be among the main sources accounting for their higher psychiatric morbidity. The fact that subjective perception of life events is more important than the event itself makes the estimation of the causative role of life events more difficult. On the other hand, it has also been demonstrated that acute, positive life events (that are quite rare in Western communities) can also precipitate either major depression or mania in vulnerable individuals. Negative life events should not be considered as only predisposing or precipitating factors, as they are frequently the result of the behavior of patients with major depressive disorder and, particularly, patients with bipolar disorder.

Social Support

Social support can improve coping and modify (e.g., reduce or eliminate) the occurrence of psychosocial stressors or the adverse consequences of them. The consistent finding in the literature that living alone, having low socioeconomic status, and being unemployed are significant risk factors for mood disorders. In other words, it means that weak or lacking social support (including social network, social interaction, and instrumental support) can also be considered a major risk factor. However, regarding social interactions, the frequency of the interactions is more important than the amount. Poor social support is related to onset, relapse, and recurrence of depression, but there is no evidence for the excess of depression among women caused by reduced social support.

COMORBIDITY

Individuals with major mood disorders are at an increased risk of having one or more additional comorbid Axis I disorders. The most frequent

disorders are alcohol abuse or dependence, panic disorder, OCD, and social anxiety disorder. Conversely, individuals with substance use disorders and anxiety disorders also have an elevated risk of lifetime or current comorbid mood disorder. In both unipolar and bipolar disorder, men more frequently present with substance use disorders, whereas women more frequently present with comorbid anxiety and eating disorders. In general, bipolar patients more frequently show comorbidity of substance use and anxiety disorders than do patients with unipolar major depression. Recent clinical and epidemiological studies show that, if bipolar II disorder is also considered, comorbidity of substance use disorders and anxiety disorders is the highest in the bipolar II disorder subgroup. Because comorbid substance use disorders and anxiety disorders worsen the prognosis of the illness and markedly increase the risk of suicide among unipolar major depressive and bipolar patients, the data presented above are consistent with previous findings showing that bipolar patients (particularly bipolar II disorder patients) are at the highest risk of suicide. These findings also indicate that Axis I comorbidity might be among the most important contributing factor to that risk. However, there is a substantial overlap between mood and substance use disorders. Data from the NESARC study revealed that, 41 percent of those with alcohol use disorders met criteria for primary depression, 17 percent for concurrent depression, and 42 percent for secondary depression. This finding supports and extends the results of previous studies conducted among alcoholics in treatment and has strong diagnostic and treatment implications.

Untreated major mood disorders (and particularly bipolar disorder) carry extremely high risk of both completed and attempted suicide. This is partly due to the fact that they use more violent/more lethal methods. If patients with major mood disorders attempt or commit suicide, they do it mostly during their **severe** major (mostly mixed) depressive episode and less frequently in dysphoric mania but very rarely during euphoric mania and euthymia suggesting that suicidal behavior in patients with major mood disorder is a "state-dependent" and severity-dependent. Consequently, the early recognition and appropriate acute and long-term treatment of mood disorders is crucial for suicide prevention. The careful and systematic exploration of clinically detectable suicide risk factors helps clinicians to identify suicidal patients, but unfortunately, at the stage of present knowledge, all suicides cannot be predicted and prevented.

SERVICE PATTERNS

Adults

Despite the great progress in the diagnosis and treatment of depressions in the last two decades, they still remain underdiagnosed and undertreated. North American and European surveys show that approximately half of

those who develop mood disorders seek treatment for them, but only a small fraction (about one-third) of recognized depressives receive appropriate treatment. The results of the NESARC study revealed that only 36.8 percent of those with a current mood disorder had sought disorder-specific treatment for a mood disorder. Interestingly, people with manic episodes were less likely than those with major depressive episodes to seek disorder-specific treatment. An increasing rate of health service utilization is related to increasing severity of depression and to comorbid psychiatric and medical disorders.

Despite the fact that many patients with depressive disorders seek help in primary care, general practitioners still have difficulties recognizing and treating depression. The current prevalence of major depression in primary care is approximately 10 to 15 percent and that of dysthymic disorder is approximately 6 to 8 percent. Paradoxically, because severe major depression is much more common among those with comorbid chronic medical disorders, depressions with significant somatic comorbidity, in particular, remain unrecognized in primary care. Population-based studies also show that in individuals with two or more chronic physical illnesses, the 1-year prevalence of severe major depression is fourfold higher than that of individuals without such conditions. Women more often seek treatment for their depression and are more compliant with the treatment than are men, but, unfortunately, "male depression" is less frequently recognized. Because there is no gender difference in the response to antidepressant pharmacotherapy, these facts can help to explain the so-called suicide paradox—that is, despite the fact that major depression is much more common in women, it is men who are markedly overrepresented among suicide victims. Besides general practitioners, other professionals (i.e., internists, cardiologists, neurologists) also frequently see depressed patients. The point prevalence of major depressive disorder in acute (medical–surgical hospital) care is also higher than 10 percent, a figure well in excess of the point prevalence rate for the general population. Concomitant depression increases the morbidity and mortality from concurrent medical illness, and patients with simultaneous medical disorder and depression are less compliant with treatment and take longer to recover than nondepressed medical patients.

Depressed patients, particularly those who go unrecognized, frequently seek general health care, but depression, if correctly identified, can be effectively treated even in primary care. On the other hand, despite the fact that more than half of suicide victims contact different levels of health care 4 weeks before committing suicide, the rate of adequate (antidepressive) pharmacotherapy among depressed suicide victims is disturbingly low. However, epidemiological and clinical studies, as well as the steadily increasing use of antidepressants (the latter is accompanied by a constant decline in national suicide rates in most well-developed countries,

particularly in those with traditionally high suicide rates) indicate that, in the last 25 to 30 years, the referral, recognition, and successful treatment of depression increased, suggesting that the prognosis for depressive illness, in general, is improving. Given the high prevalence of depressive disorders in the population, and particularly in most clinical settings, as well as the strong relationship between untreated depression and suicidal behavior indicates that better and more widespread treatment of depression is an important part of suicide prevention. The repeated postgraduate education of GPs and other healthcare workers in combination with gatekeeper training and public education is particularly useful.

Youth

Despite the magnitude and serious consequences of depression in youth, only about one-quarter to less than half of those with mental disorders receive mental health services. Factors associated with service utilization include ethnicity, high global impairment, comorbidity, prior history of depression, suicide attempt, and impact of the child's problem on the family. School services are the most common point of entry for children seeking services, although those who enter through the education sector are least likely to transition to specialty mental health services. The actual diagnostic process and services provided differ dramatically by the context of entry to service.

There has been accumulating controlled research on the effects of antidepressants in youth, but few studies of either the comparative efficacy of various agents that are commonly used in clinical practice or of individual and family therapies either in conjunction with or independent from drug treatments. One study that demonstrated the efficacy of combined pharmacologic and psychotherapeutic treatment has provided a model for future studies that can guide policy in health services for adolescents.

IMPACT OF MOOD DISORDERS

The importance of role disability has become increasingly recognized as a major source of indirect costs of illness because of its high economic impact on ill workers, their employers, and society. The introduction of the concept of Disability Adjusted Life Years (DALYs), which estimate the disease-specific reduction in life expectancy attributable to disability and increased mortality has highlighted the dramatic public health impact of mental disorders. By the year 2020, it is estimated that psychiatric and neurologic disorders will account for 15 percent of the total burden of all diseases Major depression is the leading cause of disability among those age 5 and over, and the second leading source of disease burden surpassing cardiovascular diseases, dementia, lung cancer, and diabetes.

In the most recent estimates of the impact of chronic diseases by the WHO, 2.5 percent of the total proportion of DALYs was attributable to major depressive disorder and 0.5 percent to bipolar disorder. In terms of DALYs for Mental, Neurological, and Substance Use Disorders, 24.5 percent was attributable to major depressive disorder and 5 percent to bipolar disorder. Comparative studies of role disability reveal that the effects of mood disorder outweigh those of most other mental and neurological disorders with a few exceptions. The dramatic impact of mood disorders on lifetime disability highlights the importance of epidemiology in surveillance, understanding and control of the major mental disorders.

The global burden of mental disorders in children and adolescents up to 24 years of age has also been examined. Similar to studies within the United States, there is a broad range to estimates of the rates of mental disorders. A recent review reported estimates that ranged from 8 (in the Netherlands) to 57 percent (for young people receiving services in five sectors of care in San Diego, California, USA). Results from The Australian National Survey of Mental Health and Well Being showed that at least 14 percent of adolescents younger than 18 years had a diagnosable mental or substance use disorder within the previous 12 months, and this figure rose to 27 percent in the 18- to 24-year age. Meta-analysis of the childhood and adolescent data has shown that at least one out of every four to five young people in the general population will suffer from at least one mental disorder in any given year. However, there is much less information on the burden of mental disorders in developing countries and substantial cross-cultural variations are likely.

Mortality and Mood Disorders

There are now several prospective studies that have investigated mortality among those with a range of mental disorders. The increased mortality associated with mood disorders has been well-documented. Data from large population registries in Sweden and Denmark have shown that bipolar disorder is associated with substantial elevation in risk of both mortality in general, particularly death by suicide. The mortality risk associated with mood disorders tends to decrease with age; that is, the high risk of death by suicide associated with affective disorders in the Danish Registry was much greater during the first 12 months following the initial admission compared to later years. No differences in excess mortality have been shown for different subgroups of mood disorders in either of these registries.

HISTORICAL TRENDS

Whereas early community surveys suggested a strong cohort effect for mood disorders, both retrospective and prospective studies show that there

has been no direct evidence for a marked increase in the incidence and prevalence of depression over the past decades. The observed changes are mainly due, in particular, to artifacts of memory and, to a lesser degree, to higher subjective awareness, better definitions and methods of case identification, redistribution by age and sex (with higher rates among younger women being of recent origin), and increase in childhood-onset mood disorders. Similar findings have emerged for depression in youth with comprehensive meta-analyses showing that there has been no increase in rates of depression across the last three decades.

REFERENCES

Akiskal HS. Validating the bipolar spectrum. *J Affect Disord*. 2003;73:1–2.

Alonso J, Petukhova M, Vilagut G, et al. Days out of role due to common physical and mental conditions: results from the WHO World Mental Health surveys. *Mol Psychiatry*. 2001;16(12):1234–1246.

Angst J, Cui L, Swendsen J, et al. Major depressive disorder with subthreshold bipolarity in the National Comorbidity Survey Replication. *Am J Psychiatry*. 2010;167(10):1194–1201.

Avenevoli S, Swendsen J, He JP, Burstein M, Merikangas KR. Major depression in the national comorbidity survey-adolescent supplement: prevalence, correlates, and treatment. *J Am Acad Child Adolesc Psychiatry*. 2015;54(1):37–44.e2.

Bauer M, Pfennig A. Epidemiology of bipolar disorders. *Epilepsia*. 2005;46:8–13.

Boyd RC, Joe S, Michalopoulos L, Davis E, Jackson JS. Prevalence of mood disorders and service use among US mothers by race and ethnicity: results from the National Survey of American Life. *J Clin Psychiatry*. 2011;72(11):1538–1545.

Bromet E, Andrade LH, Hwang I, et al. Cross-national epidemiology of DSM-IV major depressive episode. *BMC Med*. 2011;9:90.

*Costello EJ, Foley DL, Angold A. 10-year research update review: the epidemiology of child and adolescent psychiatric disorders: II. Developmental epidemiology. *J Am Acad Child Adolesc Psychiatry*. 2006;45(1):8–25.

*Faedda GL, Marangoni C, Serra G, et al. Precursors of bipolar disorders: a systematic literature review of prospective studies. *J Clin Psychiatry*. 2015;76(5):614–624.

Frías Á, Palma C, Farriols N. Comorbidity in pediatric bipolar disorder: prevalence, clinical impact, etiology and treatment. *J Affect Disord*. 2015;174:378–389.

*Gore FM, Bloem PJ, Patton GC, et al. Global burden of disease in young people aged 10–24 years: a systematic analysis. *Lancet*. 2011;377(9783):2093–2102.

*Hasin DS, Grant BF. The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) Waves 1 and 2: review and summary of findings. *Soc Psychiatry Psychiatr Epidemiol*. 2015;50(11):1609–1640.

Laursen TM, Munk-Olsen T, Nordentoft M, Mortensen PB. Increased mortality among patients admitted with major psychiatric disorders: a register-based study comparing mortality in unipolar depressive disorder, bipolar affective disorder, schizoaffective disorder, and schizophrenia. *J Clin Psychiatry*. 2007;68(6):899–907.

Lewinsohn PM, Seeley JR, Buckley ME, Klein DN. Bipolar disorder in adolescence and young

adulthood. *Child Adolescent Psychiatry Clin N Am*. 2002;11(3):461–475.

McIntyre RS, Konarski JZ, Soczynska JK, et al. Medical comorbidity in bipolar disorder: implications for functional outcomes and health service utilization. *Psychiatr Serv*. 2005;57(8):1140–1144.

Merikangas KR, Akiskal HS, Angst J, et al. Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication. *Arch Gen Psychiatry*. 2007;64(5):543–552.

Merikangas KR, Cui LH, Kattan G, Carlson GA, Youngstrom EA, Angst J. Mania with and without depression in a community sample of US adolescents. *Arch Gen Psychiatry*. 2012;69(9):943–951.

Merikangas KR, Herrell R, Swendsen J, Rossler W, Ajdacic-Gross V, Augst J. Specificity of bipolar spectrum conditions in the comorbidity of mood and substance use disorders. *Arch Gen Psychiatry*. 2008;65(1):47–52.

*Merikangas KR, Jin R, He JP, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch Gen Psychiatry*. 2011;68(3):241–251.

Merikangas KR, Tohen M. Epidemiology of bipolar disorder in adults and children. In: Tsuang MT, Tohen M, Jones P, eds. *Textbook of Psychiatric Epidemiology*. 3rd ed. West Sussex, UK: Wiley; 2011:329–342.

Mortensen PB, Pedersen CB, Melbye M, Mors O, Ewald H. Individual and familial risk factors for bipolar affective disorders in Denmark. *Arch Gen Psychiatry*. 2003;60(12):1209–1215.

Murphy JM, Gilman SE, Lesage A, et al. Time trends in mortality associated with depression: findings from the Stirling County study. *Can J Psychiatry*. 2010;55(12):776–783.

Ormel J, Oldehinkel AJ, Sijtsema J, et al. The Tracking Adolescents' Individual Lives Survey (TRAILS): design, current status, and selected findings. *J Am Acad Child Adolesc Psychiatry*. 2012;51(10):1020–1036.

*Patel V, Chisholm D, Parikh R, et al; *DCP MNS Author Group*. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd ed. *Lancet*. 2015;387(10028):1672–1685.

Ramsey CM, Spira AP, Mojtabai R, Eaton WW, Roth K, Lee HB. Lifetime manic spectrum episodes and all-cause mortality: 26-year follow-up of the NIMH Epidemiologic Catchment Area Study. *J Affect Disord*. 2013;151(1):337–342.

Rihmer Z. Prediction and prevention of suicide in bipolar disorders. *Clin Neuropsychiat*. 2005;2:48.

Rihmer Z, Szádóczky E, Füredi J, Kiss K, Papp ZS. Anxiety disorders comorbidity in bipolar I, bipolar II and unipolar major depression: results from a population-based study in Hungary. *J Affect Disord*. 200l;67:175–179.

Sala R, Goldstein BI, Morcillo C, Liu SM, Castellanos, M, Blanco C. Course of comorbid anxiety disorders among adults with bipolar disorder in the U.S. population. *J Psychiatr Res*. 2012;46(7):865–872.

Tsuchiya KJ, Byrne M, Mortensen PB. Risk factors in relation to an emergence of bipolar disorder: a systematic review. *Bipolar Disord*. 2003;5(4):231–242.

*Van Meter AR, Moreira ALR, Youngstrom EA. Meta-analysis of epidemiologic studies of pediatric bipolar disorder. *J Clin Psychiatry*. 2011;72(9):1250–1256.

Waraich P, Goldner EM, Somers JM, Hsu L. Prevalence and incidence studies of mood disorders: a systematic review of the literature. *Can J Psychiatry*. 2004;49(2):124–138.

▲ 13.3 Mood Disorders: Genetics

JOHN R. KELSOE, M.D., AND TIFFANY A. GREENWOOD, PH.D.

INTRODUCTION

The substantial role of genes in the susceptibility to mood disorders has long been supported by family, twin and adoption studies. That mood disorders run in families is a common observation of patients and clinicians. However, genes clearly only contribute a predisposition that must interact with environmental factors in order to cause disease. Studies of the inheritance of different forms of mood disorder and other psychiatric disorders suggest complex genetic relationships. Different forms of illness such as bipolar II may share some, but not all, genes with bipolar I. Studies of a variety of phenotypes related to mood disorder suggest a similar pattern of partial genetic overlap. Alternatively, some aspects of mood disorder seem more quantitative in nature. Early attempts to identify the specific genes involved through linkage analysis of large families and association analyses of specific candidate genes were met with mixed success and few reproducible results. Very high throughput microarray-based methods more recently made possible large-scale genome-wide association studies (GWAS) that promised to identify additional genes. However, these studies also suffered initially from poor replicability, which has been improved through the use of increasingly larger samples. Yet, the identified risk genes presently explain only a small portion of the genetic variance in mood disorders. The identification of specific genes has a number of important clinical ramifications. Most importantly, a better understanding of basic pathophysiology may lead to novel treatments and drug targets. Better knowledge of pathophysiology may also inform future diagnostic systems based more on underlying biology, rather than symptom profiles. Additionally, pharmacogenetics, or the study of how genes influence medication response, is making progress and promises tools to aid physicians in drug selection.

GENETIC EPIDEMIOLOGY

Genetic epidemiological studies can provide a great deal of information about the familial and genetic nature of a disorder by examining the rates of illness among relatives. Such studies also address how different forms of mood disorder may differ in their genetic transmission and how they may relate to each other. However, these studies must be considered in the light of their various methodological limitations. First, many studies were

conducted before the distinction between bipolar disorder and major depression; thus, the statistics from these studies pool both illnesses. Similarly, many studies preceded the introduction of operationalized diagnostic criteria, making comparisons across studies difficult. Additionally, if the results of a study are to be meaningfully generalized to the population, it is critical that the subjects to be studied are selected in a systematic fashion, free of ascertainment bias. For example, if probands (the first ill subject identified in a family) are selected based on their strong family history, then the results of a family study may inaccurately indicate a strong familial rate of illness in the population. Finally, for reasons of feasibility, many of these studies used the family history method, in which the rates of illness in family members are determined by questioning the proband about his or her family. Although there are several excellent standardized instruments available, this method is inherently less accurate than the direct interview method, in which each family member is interviewed to make a diagnosis.

Family studies address the question of whether a disorder is familial and compare the rates of illness among first-degree relatives of a proband with the disorder to either the rates in the general population or the rates in first-degree relatives of control subjects. Rates of illness are typically adjusted for age to indicate the morbid risk. Compared to a rate of 1 percent in the general population, family studies indicate a sevenfold higher morbid risk to first-degree relatives of bipolar probands, suggesting a strong familial risk. Similarly, family studies of probands with major depression reveal morbid risks for depression among first-degree relatives that are elevated two- to threefold over the general population. These data strongly support the familial nature of mood disorders. Furthermore, major depression is generally found at an elevated rate in the families of bipolar probands, although the rate of bipolar disorder is only slightly elevated in the families of probands with major depression. This familial overlap suggests some degree of common genetic underpinnings between these two forms of mood disorder.

Family studies clearly indicate that mood disorders are familial but are unable to distinguish whether it is genetic or environmental factors that mediate the familial transmission. Families might share a variety of different environmental factors that could transmit the illness, including shared exposure to infectious agents, toxins, and other brain insults. Twin studies provide the most powerful approach to separating genetic from environmental factors, or “nature” from “nurture.” The most common strategy is one in which both monozygotic (MZ) and same-sex dizygotic (DZ) twin pairs are identified with one twin having a mood disorder. The cotwins are then examined to determine the proportion of twin pairs in which both twins are affected. This is termed the concordance rate. As the twin pairs have been raised together and share the same environmental

factors, a difference in concordance rate between MZ and DZ pairs reflects the role of heritable genetic factors. Twin studies generally find a two to fourfold increase in concordance rate for mood disorder in MZ twins compared to DZ twins, providing the most compelling data for the role of genetic factors in mood disorders. It is equally noteworthy that the concordance rate for MZ twins is only around 70 to 80 percent, not 100 percent. This is clear evidence that nonheritable environmental factors also play a significant role in mood disorders. In studies that distinguish bipolar disorder from major depression, the MZ to DZ concordance ratio for bipolar–bipolar pairs is higher than that for depression–depression pairs, indicating a greater role for genetics in bipolar disorder than in depression. Furthermore, the rate of depression is elevated in the MZ cotwins of bipolar probands, and, to a lesser extent, the rate of bipolar disorder is elevated in the cotwins of probands with depression. This is consistent with the family data and argues for a genetic overlap between bipolar disorder and major depression.

A variety of limitations of twin studies have been raised, including the argument that parents treat MZ twins and DZ twins differently, so environment is not equally shared. Adoption studies provide an alternative approach to separating genetic and environmental factors in familial transmission. The most common experimental design is one in which probands are identified who have a mood disorder and were adopted at birth. The rates of psychiatric illness are then determined in both the biological and adoptive parents. Although only a limited number of such studies have been reported due to the difficulty in obtaining subjects, these results of such studies are supportive of the role of genetics and are generally consistent with the twin data.

In summary, family and twin data collectively suggest that genes explain approximately 75 percent of the etiology of bipolar disorder and 37 percent of major depression. More recently, methods have been developed to estimate heritability based on the genome-wide genotyping commonly done for GWAS as described in detail below. In this method, allele sharing is used as a measure of shared genetic vulnerability. The current understanding of the heritability of mood disorders has also been greatly enhanced by the recent availability of large population-based samples. For example, one study examined the extent of allele sharing among relatives in the Swedish national registry of more than nine million individuals from more than two million families. This study suggested that the heritability of bipolar disorder was 59 percent. Another study used population summary statistics to estimate heritability in a Danish sample of more than 2.6 million individuals. This study revealed heritability estimates of 62 and 32 percent for bipolar disorder and major depression, respectively. The heritability of bipolar disorder using population-based methods is quite a bit lower than that obtained from twin samples. This may be due to biases

in how the twin samples were collected, as discussed above, which would inflate estimates of the genetic contribution. However, heritability estimates for major depression are reasonably consistent, whether derived from twin or population-based samples. Regardless of origin, all heritability estimates demonstrate that environment or other nonheritable factors must explain a substantial portion of risk for mood disorders, particularly major depression, since genes appear to contribute much less than 100 percent of the risk. Recent molecular methodologies, as discussed below, are helping to increase the current understanding of the nature of this heritability.

GENETIC FEATURES OF MOOD DISORDERS

For mood disorders, there is no 1-to-1 relationship between the genes (genotype) and the expressed trait (phenotype) that is transmitted in a simple and predictable fashion as observed for Mendelian traits. Therefore, mood disorders are said to be complex genetic disorders. What are the factors that contribute to this complexity?

Studies reporting heritability estimates of bipolar disorder that are less than 100 percent clearly demonstrate that it is a predisposition or susceptibility to disease that is inherited. The probability that someone will manifest a trait given that they have a certain genotype is termed the penetrance of the gene. Mood disorders are said to have reduced penetrance (less than 100 percent). Furthermore, the penetrance of mood disorder genes increases with age, from a low risk for illness in childhood to a maximum in adulthood. Thus, in families of people with mood disorders, there are likely individuals who have the genes for mood disorder but do not develop the disease. These are termed nonpenetrant carriers. The converse of this situation is individuals who have mood disorder but do not have the genes. Such individuals with purely environmentally caused disease are termed phenocopies. These factors conspire to produce an indirect relationship between genes and disease.

Variable expressivity refers to the phenomenon of the same gene or group of genes resulting in a variety of different forms of illness. The twin data clearly demonstrate this for mood disorders. Identical twins with identical genomes are observed in which one twin has bipolar disorder and the other has depression. The sharing of genetic vulnerability between different forms of illness or disorders is also a manifestation of variable expressivity. Nonheritable factors must play a role in the specific manifestation of the predisposition to mood disorder. Variable expressivity may also result from modifier genes that influence the expression of the susceptibility gene. Such variability in expression is not unique to psychiatric disorders. For example, in neurofibromatosis, ill individuals range in manifestations and severity from those with only pigmented retinal lesions to others with multiple large tumors. Such a range in

expression results from the same disease gene.

Another of the various factors complicating efforts to map genes for mood disorders is genetic heterogeneity. Heterogeneity refers to the likely role of multiple genes in the etiology of illness. The data for genetic heterogeneity in mood disorders at this point are compelling. There are several critical questions regarding the nature of heterogeneity in mood disorders: How many genes are involved? How large an effect does each gene have? In what ways do the genes interact to produce illness? These questions define a variety of models of heterogeneity, which are described below. Such models can be broadly grouped into those in which disease results from a few genes of major effect (major loci) and those in which disease derives from the combined action of many genes of small effect (polygenic). The answers to these questions are not currently known and are a topic of great debate currently framed as the portions of the missing heritability transmitted by common versus rare variants.

MODE OF TRANSMISSION OF COMPLEX GENETIC DISORDERS

Early attempts to understanding the mode of genetic transmission for mood disorders involved an examination of the pattern of transmission in families. This approach, termed segregation analysis, had relatively poor statistical power to distinguish forms of transmission but provided some support for autosomal dominant transmission. These studies also provided support for the polygenic multithreshold model described below.

Molecular tools and much larger sample sizes have led to a more powerful approach to these same questions. It is clear that mood disorders are complex traits, but in what ways might genes be transmitted so as to create the observed patterns of transmission and fit the complex nature of the phenotype? Current data suggest a mixture of genes of both large and small effect, which are transmitted in a variety of ways. The basics of genes and genetic transmission are covered elsewhere in this text. Here, different possible models of transmission will be reviewed in the context of their implications for mood disorders.

In single major locus transmission, one gene explains the majority of the genetic transmission in families. It results in the common Mendelian transmission patterns of autosomal dominant, autosomal recessive, and X-linked, as shown in the top pane of Figure 13.3–1. In autosomal dominant transmission, only one copy of the disease allele can cause illness, while in recessive transmission, two copies are necessary. X-linked describes the transmission of trait loci on the X chromosome. As described above, single major locus transmission of the same gene across many families seems highly unlikely, given existing data.

Alternatively, some phenotypes are better represented as numbers or quantitative traits, with blood pressure or height being classic examples. Such phenotypes are usually modeled as polygenic traits. Here, each gene

contributes a small amount toward the overall genetic effect. The more polygenes an individual has, or the higher the genetic loading, the larger the value of the trait. In the simplest model, such traits are normally distributed in the population based on the distribution of individuals in the population with different numbers of polygenes. In Figure 13.3–1, this is illustrated in the middle pane. The question is whether such a model fits the observed data for mood disorders. Is the mood disorder phenotype best described as a continuous variable? One approach is to argue that the quantitative trait cannot be observed directly, but rather represents some latent trait that predisposes to mood disorder, like affective temperament. When an individual has a trait value above a certain threshold, they are more likely to develop the disease. These concepts are illustrated in Figure 13.3–2, according to a model in which affective temperaments mediate the underlying genetic disposition to bipolar disorder. This quantitative trait model is also consistent with the observation of milder forms of mood disorder spectrum traits in family members and the range in severity in the population. However, it is not clear exactly how to order different forms of illness as being "more" mood disordered. Is severe major depression less of a mood disorder than mild bipolar disorder? Though aspects of this model are attractive, the transmission of mood disorders is clearly more complex.

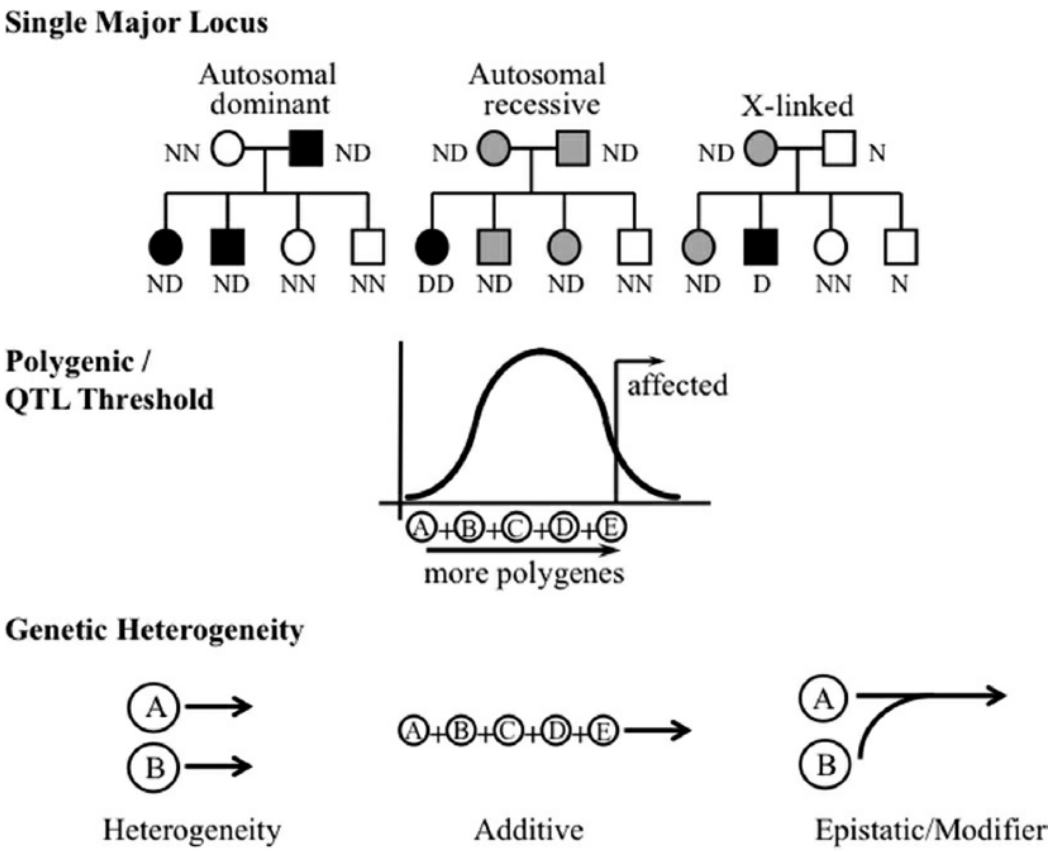

FIGURE 13.3–1. Models of genetic transmission. Many different models have been considered and tested to see if they explain the transmission of mood disorders. This is a selection of some of the more prominent models. In Mendelian or single major locus

transmission (**top pane**), one gene transmits the illness. In a polygenic quantitative trait (QTL) model (**middle pane**), multiple genes add together to contribute to a quantitative trait. In this figure, the x-axis represents the number of polygenes that a given individual is carrying, as well as, the value of the resulting quantitative trait. The frequency of that trait value in the population is represented on the y-axis. In a QTL threshold model, illness results when an individual's trait value, and therefore genetic loading, exceeds a critical threshold. In the **bottom pane**, some possible models of genetic heterogeneity are illustrated, including multiple single major loci, many loci of small additive effects, and gene interactions.

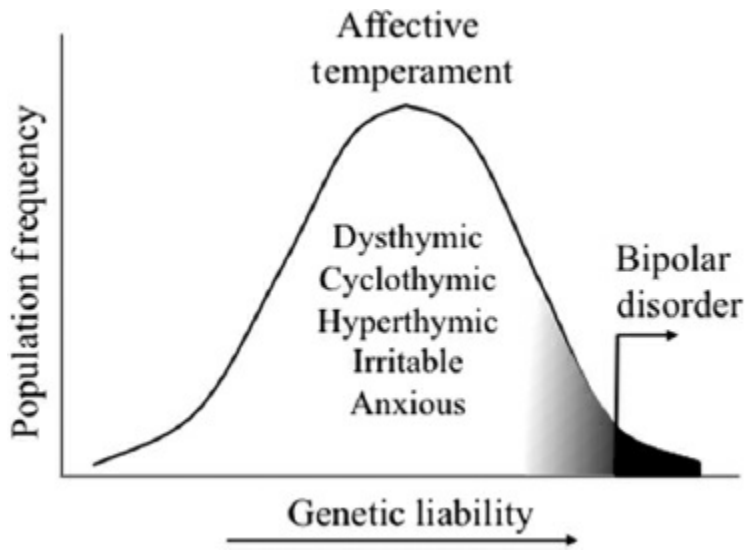

FIGURE 13.3–2. Affective temperament as a quantitative trait for bipolar disorder. This figure illustrates a polygenic, quantitative trait threshold model of bipolar disorder. If bipolar disorder is to some extent a polygenic trait, then it is likely that the risk variants occur and are transmitted primarily in mildly affected individuals who may have affective temperaments, but not illness. The affective temperaments would be normally distributed in the population, with more extreme temperaments conferring a greater risk for illness. This model is consistent with the observation of increased temperament traits in family members or bipolar patients, who are at a higher genetic liability for illness. At some critical threshold, the high genetic liability results in bipolar disorder.

The bottom pane in Figure 13.3–1 illustrates some models of genetic heterogeneity. Clearly, multiple genes are involved, but in what way do they combine or interact to produce illness? The first is a heterogeneity model involving multiple single major loci. In this model, multiple genes for mood disorder exist in the population, each able to alone cause mood disorder. As stated above, it is clear at this point that there are many genes for mood disorder. However, it remains very plausible that there could be many genes each operating as a single major locus in a subset of families. If the number of genes in the model is quite high, then mapping such genes could be very difficult. The second is an additive model, in which many genes may add together in order to produce a cumulative predisposition to illness. This is simply the polygenic quantitative trait model described above. A third model involves gene interactions. In this model, the probability of one gene manifesting the phenotype is modulated by the interacting effect of another gene. If an individual has one gene or the

other, the risk for illness might be small, but if the individual has both disease alleles, then the risk increases. This is termed epistasis and is distinguished from an additive model in that the increase in risk from having two genes is greater than the sum of the risks of the individual genes. In a modifier model, the modifying gene has little or no effect on predisposition alone, but rather modifies the effect of another gene.

These three models of heterogeneity have very different implications for mapping genes for mood disorder. The heterogeneity model calls for possibly a large number of rare alleles, each of strong effect. In each individual and their family, the illness would result primarily from one gene, but that variant would be quite rare. In the polygenic model, each polygene contributes only a small amount to the overall genetic vulnerability. Under this model, each of the polygenes would be relatively common in the population, but the effect size of each would be small, such that many polygenes would add together to produce the observed population prevalence of mood disorders. This implies that the alleles for mood disorder would primarily reside in individuals without mood disorder who have too few polygenes to develop the illness. Some of these individuals with a few risk alleles might instead have a spectrum trait, such as cyclothymic temperament. In the gene interaction model, each gene alone might have only a small effect. As it is particular combinations of genes that would convey larger genetic effects, the genes would best be identified by considering multiple genes simultaneously. These different models call for different approaches to mapping or identifying the disease genes.

GENETICS OF THE MOOD DISORDER SPECTRUM

The genetic relationship between the various forms of mood disorder is a topic of much recent interest. Among mood disorders, bipolar disorder and major depression are widely felt to have some sort of common genetic underpinning, although its exact nature is unclear. Twin and family data argue for major depression occurring in the twins or other relatives of bipolar probands at a greater rate than that expected by chance. However, it is less clear that bipolar disorder occurs in the relatives of probands with depression at an elevated rate. Twin studies indicate that "polarity" is usually consistent in MZ twins—in other words, bipolar–bipolar or depression–depression pairs are much more common than bipolar–depression pairs. Yet bipolar–depression pairs do occur. These data suggest that bipolar disorder and depression are neither completely identical nor completely distinct genetically. Rather, there is a partial genetic similarity.

A partial commonality of susceptibility genes would also predict that a portion of those with depression carry genes that may also predispose to bipolar disorder. Such patients have been said to have "bipolar III"

disorder and have been the subject of much discussion and investigation. They are presumably identified by a family history of bipolar disorder or a history of developing hypomania or mania only in response to antidepressant treatment. Similarly, a hypomanic-like personality style termed hyperthymic temperament has been described. Patients with depression and a hyperthymic temperament are more likely to have a family history of bipolar disorder and to develop mania spontaneously. These patients with depression and a bipolar genetic diathesis may also be more likely to respond to lithium augmentation of antidepressant treatment. These data suggest that bipolar disorder and depression may be better understood as existing on a spectrum, consistent with a quantitative genetic trait.

Other forms of mood disorders have also been postulated to be, at least to some extent, genetically distinct. Although the risk for bipolar I is similar in the relatives of bipolar I or bipolar II probands, several studies have reported that the risk for bipolar II is greater than that for bipolar I in the relatives of bipolar II probands. This suggests that bipolar II, to some extent, breeds true and that a subset of the genes for bipolar disorder predisposes preferentially to bipolar II.

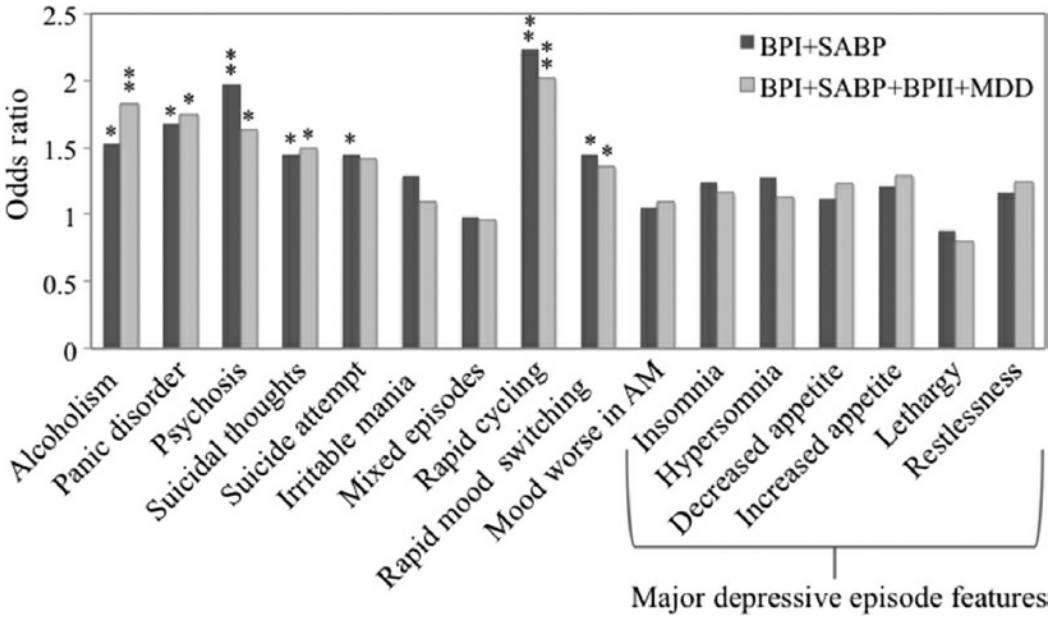

FIGURE 13.3–3. Familiality of clinical features in mood disorder. This figure illustrates how a family-based strategy can be used to evaluate different features of mood disorder that may be more or less genetically determined. In this sample of 589 bipolar families, relevant clinical features were evaluated for cluster within sibling groups. Familiality was measured as an odds ratio, which refers to the odds in favor of expressing the clinical feature. The strength of this association is indicated by a *p* value <0.05 (*) or <0.001 (**). Several clinical features displayed familiality under a narrow model that included only siblings with bipolar I disorder (BPI) and schizoaffective disorder bipolar type. These included features of bipolar disorder, such as psychosis, rapid cycling, and rapid mood switching, as well as comorbid conditions, such as alcoholism and panic disorder. These clinical features were more or less significant under a broad model that added bipolar II disorder (BPII) and major depression (MDD).

This family-based strategy has been used for a number of different features of mood disorder. Such evaluations have focused on clinical features that cluster within bipolar families and display familiality, as shown in Figure 13.3–3. Bipolar disorder with psychotic features is more commonly found in the families of probands with psychotic symptoms, suggesting that bipolar disorder with psychotic features may be a genetically distinct form of illness. Comorbid panic disorder also tends to cooccur in families and be associated with more rapid mood switching. The most strongly familial of bipolar disorder subforms is rapid cycling illness, which is twice as likely to occur in a relative if the proband has rapid cycling illness. Suicidality has also been shown to cluster in families in a way that suggests that it may represent a genetically distinct from of illness. This might result from a subset of mood disorder genes that are particularly virulent in terms of suicide or a modifier gene that influences suicidal risk in those vulnerable to depression. Early age at onset and comorbid alcoholism are other traits that have been reported to some extent to "breed true" in families with both bipolar disorder and depression.

Some traits have been associated with a higher heritability and hence a more pronounced role of genes. For example, major depression in general has a lower heritability than bipolar disorder. However, when families or twins are selected through a proband with an early age at onset and recurrency, a much higher heritability is observed. Furthermore, these traits of early onset and recurrent illness run in the families. This suggests that these traits are associated with a stronger genetic effect and may result from a subset of the broader set of mood disorder genes. Such traits are important to identify, as they may enable the selection of a set of subjects with a more genetic form of illness that may be easier to map.

Several other kinds of traits have been hypothesized to share genes with mood disorders. It has been repeatedly observed that bipolar patients who respond well to lithium have a stronger family history of bipolar disorder. Lithium response has also been reported to cluster in families, suggesting that lithium responsive bipolar disorder is a genetically distinct form of illness. Some somatic illnesses have also been genetically associated with mood disorder. Migraine headache occurs at a higher frequency in the families of bipolar probands. Migraine in particular may be associated with bipolar II illness. Such somatic associations should not be surprising. At least half of genes expressed in brain are also expressed in other tissues. Therefore, mutations that predispose to mood disorder may also affect similar systems in different organs.

Aspects of temperament and personality may also be genetically related to mood disorders. Clinically, it is not unusual to observe relatives of mood disorder patients who have milder manifestations of mood disorder, such as dysthymia or cyclothymia. Family studies have supported this and

shown that measures of cyclothymic, dysthymic, and anxious temperaments are elevated in relatives of bipolar patients. Similarly, measures of harm avoidance, reward dependence, novelty seeking, and self-directedness may be personality traits that are stable, genetically influenced, and associated with risk for depression. These observations raise the intriguing question of whether the alleles that contribute to risk for major mood disorders also play a role in affective temperament or personality in the general population. This is consistent with a polygenic quantitative genetic model as discussed above.

Why Does Bipolar Disorder Exist? Do Mood Disorder Genes Convey a Selective Advantage?

A possible selective advantage for mood disorders has long been speculated, particularly for bipolar disorder. Individuals with bipolar disorder may go through periods of hypomania or mania where their productivity is better than usual, or better than the general population. Bipolar disorder has long been associated with creativity, as illustrated by the mania-inspired artistic genius of Tchaikovsky, Lord Byron, Ernest Hemingway, and van Gogh, among many others. Many population-based studies support the familial connection between bipolar disorder and creativity. Milder spectrum forms of bipolar disorder, such as cyclothymic disorder, or bipolar spectrum traits, such as cyclothymic or hyperthymic temperaments may also convey greater productivity or creativity on a more common, less dramatic, and more consistent basis. It is easy to see how a need for less sleep, a greater drive to work, faster thoughts and speech, and charisma would be advantageous in aspects of today's society. Hypersexuality is adequate alone from a genetic standpoint for such individuals to have an advantage in propagating their genes. Yet, bipolar disorder also results in great dysfunction and suffering.

Such a combination of positive and negative selective factors is not unusual in genetics. In the context of certain environmental conditions, selective advantages will tend to increase disease allele frequencies, while selective disadvantages will decrease them. The result is what is termed balanced selection. The frequencies of disease alleles in the population gradually change until the advantageous effects on selection are balanced by the deleterious effects. In this way, selection by the environment determines the population allele frequency and the frequency of the disease. In the context of bipolar disorder, creativity, affective temperament, and certain personality traits, such as openness to experience, extraversion, and positive schizotypal traits, provide attractive targets for a balancing selection model. Such traits would be expected to increase with mild to moderate loading for genes causing risk for bipolar disorder. As genetic loading for illness increases, trait values would begin to decrease according to an inverted-U–shaped curve along with the

increasing impairment associated with illness. This model is summarized in Figure 13.3–4 and suggests the influence of common variants distributed across the entire population, which is consistent both with the polygenic model of bipolar risk and with the observance of a stable worldwide prevalence rate. According to this model, first-degree relatives of patients and other allele carriers serve to perpetuate bipolar risk genes in the general population and benefit from the increased fitness that they confer. Such an idea has been difficult to test empirically. However, current methods that examine the correlation of the combined effects of all common genetic variants in the genome have begun to enable the direct testing of this idea, and an initial study demonstrated that bipolar disorder and creativity do indeed share genetic underpinning. Further progress in this area awaits the availability of potentially advantageous phenotypes for a sufficiently large cohort.

IDENTIFYING GENES FOR MOOD DISORDERS

Strategies for Gene Identification

The search for genes for mood disorders has, to date, focused primarily on bipolar disorder because of the stronger evidence for its genetic basis. In its early years, this search had a series of ups and downs of reported linkage findings and subsequent nonreplications likened to the highs and lows of the illness itself. However, the recent development of new methods for gene mapping has made this a very exciting time for psychiatric genetics. These methods will be addressed here specifically in relation to mood disorders. Two general strategies to disease gene identification have been employed: genome-wide mapping and candidate gene studies.

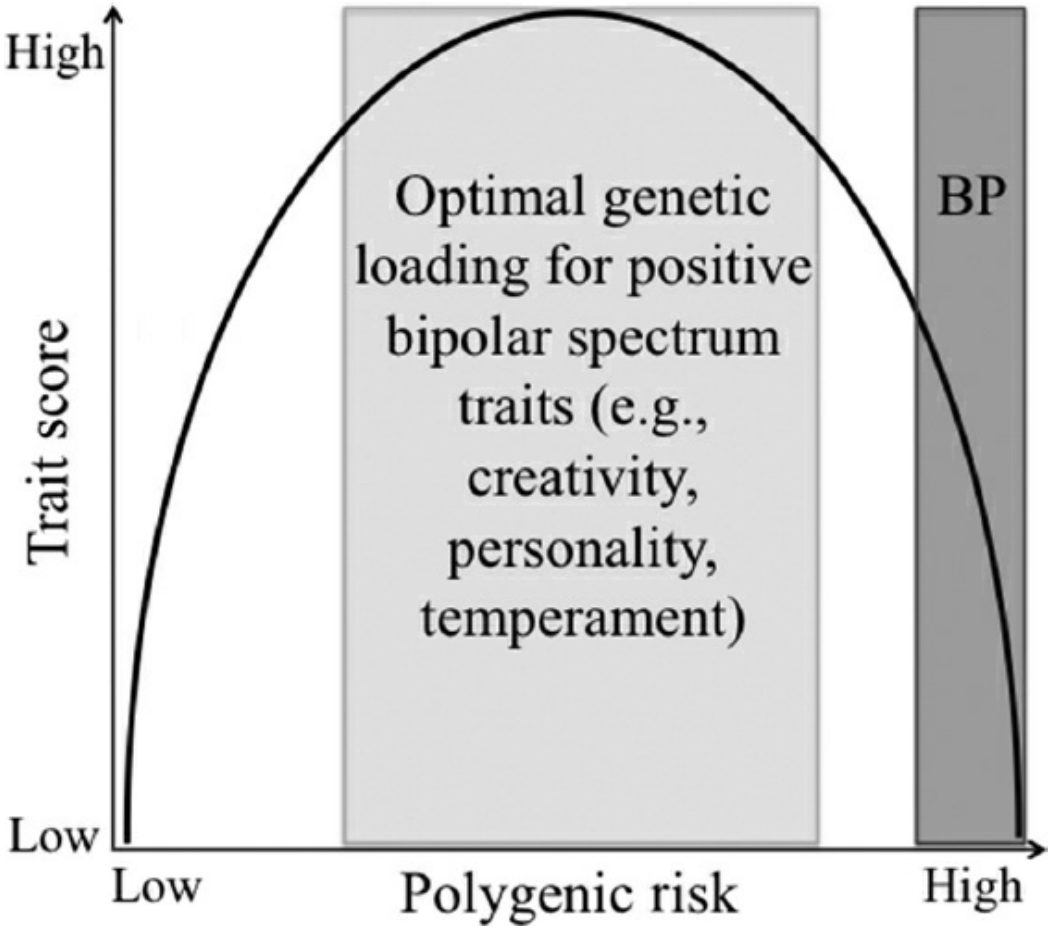

FIGURE 13.3–4. Balancing selection model of the bipolar spectrum. This figure provides an evolutionary and population genetics perspective on the bipolar spectrum. Shown is the proposed relationship of positive bipolar spectrum traits, such as creativity and affective temperaments, and the underlying shared genetic vulnerability. According to this model, positive traits increase with genetic loading for bipolar disorder (BP) up to a threshold, beyond which they begin to diminish according to an inverted-U–shaped curve along with the increasing impairment associated with illness. Polygenic risk reflects the combined influence of common genetic variants. This common variation is maintained in the population by clinically unaffected individuals, including first-degree relatives of patients and other allele carriers, who benefit from the increased fitness conferred by the associated positive, adaptive traits. The selective advantage in these individuals may greatly outweigh any selective disadvantage in severely ill individuals who are far fewer in number.

In genome-wide studies, markers are selected so as to cover the entire genome and are each tested for their genetic involvement in disease. The two primary advantages of this approach are that it is comprehensive, since all chromosomal regions are tested, and it is agnostic with no hypotheses regarding disease mechanism required. Genes are simply identified based on their position in the genome relative to the tested markers. This enables the discovery of new genes and mechanisms not previously suspected. However, this approach has some significant disadvantages. It may be quite expensive, requires very large sample sizes, and once a region is identified it may be difficult to identify the specific gene and risk variant.

The other major approach to disease gene identification is the candidate

gene approach. In this approach, a specific gene is chosen based on its known role in systems relevant to disease. DNA markers are then genotyped and tested for genetic association. This approach has the advantage of being simpler and more focused. However, since it is not systematic or genome-wide and also requires a hypothesis, it is less likely to result in novel discoveries.

Initially, genome-wide studies were conducted using genetic linkage. In this strategy, several hundred or several thousand markers that cover the genome are examined in families segregating the disease. Markers that are consistently inherited along with disease indicate chromosomal regions that may harbor disease genes. However, such regions may be quite large (10 to 20 Mb) and include hundreds of genes, such that identifying the specific gene may be difficult. Similarly, many linkage peaks may include multiple susceptibility genes, which further complicates disease gene mapping. While linkage has had some success in finding reproducible chromosomal regions, only a few specific genes have emerged from this work. It is felt by many experts that though linkage may work well for simple Mendelian traits, it may not be very powerful for complex or polygenic disorders, such as mood disorders. It may be the wrong tool for illnesses that are highly heterogeneous or have small gene effect sizes.

Over the past decade, GWAS have increasingly replaced linkage to become standard practice. It had long been known that genetic association was more powerful in detecting small gene effects than linkage. In genetic association, the frequencies of the alleles for a marker are compared between cases with the illness and controls. If an allele is more common in cases than controls then it may be involved in the illness. Association may be detected if the DNA variant used as a marker is the actual functional mutation itself or if it is simply very close to the actual mutation. The most common kind of DNA variant is a single nucleotide polymorphism (SNP), which consists of a single base substitution in the DNA sequence. The abundance and ease of genotyping SNPs has led to them becoming the most commonly used marker for association studies. As association requires the marker to be very close to the mutation, on the order of 1 to 10 kilobases, a very high density of markers is required to cover the genome. Several developments made this possible. First, the second phase of the Human Genome Project, called the HapMap Project, identified millions of SNPs and determined their frequency in different populations. Secondly, microarray technologies were developed that enabled the genotyping of hundreds of thousands of SNPs simultaneously on a chip for very low cost. Thus, it is now possible to genotype one million SNP markers spanning the genome.

Mood disorders likely result from a mixture of effects, ranging from rare variants of large effect to many common variants of small effect. Interactions among risk variants, as well as interactions of risk variants

with the environment, are also likely to play a role in susceptibility. Although the mode of transmission in mood disorders appears very complex, genome-wide investigations through GWAS and sequencing may help resolve questions as to which variants contribute risk and how they are transmitted. While GWAS may be useful in identifying common risk variants, sequencing of many individuals will be needed to identify rare variants contributing to risk.

Linkage Studies

Numerous linkage studies of bipolar disorder have been conducted in a wide range of family samples that vary in the number of families, size of families, genetic strategy (e.g., sibling pair), and geographical and ethnic origin. These studies have implicated many different chromosomal regions meeting statistical significance, with some having been replicated with statistical significance. Across studies, linkage has been demonstrated for 4p, 6q, 8q, 16p, 18p, 18q, 21q, and Xq. The 22q region was first implicated in psychiatric illness by the observation of psychotic and mood symptoms in adolescents with a 22q11 deletion syndrome called velocardiofacial syndrome. The accumulation of multiple complete linkage scans has also enabled the application of meta-analysis methods to examine the results of multiple studies and calculate statistics that test the strength of a finding across multiple data sets. One such study examining 11 published genome scans of bipolar disorder identified 13q and 22q as the two regions with strongest evidence for linkage and as those most clearly replicated. Temperament traits have also been examined for linkage, as they have been proposed to represent the primary expression of the genes underlying illness. It has been argued that such traits are a more powerful phenotype for linkage and other gene mapping studies because they are closer to the genetic effect, and as quantitative traits, they convey more information. One study of bipolar families identified linkage for hyperthymic temperament on 1q, 2p, 6q, and 14q; for dysthymic temperament on 3p and 13q; and for irritable temperament on 6q.

While fewer linkage studies have examined major depression, several chromosomal regions, including 2q, 3q, 12q, 15q, and 18q, have been reported for a recurrent, early-onset form of illness that appears to be more heritable. Neuroticism is a personality trait, characterized by anxiety and dysphoria, and related to risk for major depression. Linkage of neuroticism has been reported for several chromosomal regions, including 1q, 4q, 6p, and 11p. Another personality trait related to depression, Harm Avoidance, has been mapped to 8p.

Candidate Gene Association Studies

Numerous candidate genes have also been examined for their possible role

in the susceptibility to mood disorders. These candidates are selected largely because of their role in signal transduction, catecholamine or serotonin neurotransmission in particular. Other genes have been examined in association studies because they reside regions linked to illness, or they have shown association in other disorders, like schizophrenia.

The serotonin transporter (SLC6A4) is probably the most studied gene in psychiatric genetics. It is of great interest because it mediates the reuptake of serotonin and modulates synaptic serotonin levels; it is also the site of action of SSRI antidepressants. The promoter of the gene contains a repeat region that varies in the number of repeats between two major alleles. The short form of the repeat has been shown to express less serotonin transporter mRNA and protein than the long form and is associated with response to SSRI antidepressants. Alone it shows a relatively weak association with major depression and bipolar disorder. However, a prospective study measuring life events and the incidence of depression demonstrated that negative life events, such as death of a spouse or parent, strongly correlated with risk for major depression in those who carried the short form of the gene. This was one of the first demonstrations of an interaction between environmental stresses and a specific gene. It also emphasizes the importance of considering environment when examining gene effects.

Another candidate gene of high interest is brain-derived neurotrophic growth factor (BDNF), which is involved in the development and maintenance of catecholaminergic neurons. BDNF has been associated with bipolar and unipolar mood disorders and schizophrenia. A Val/Met substitution has been identified in the gene that affects processing of the peptide. This gene is particularly intriguing because it is a growth factor involved in brain development and neuronal maintenance. Antidepressants and lithium have been shown to increase its expression, and it likely plays a role in their neuroprotective functions.

Many candidate genes for mood disorder have been associated with multiple phenotypes. DRD4 and SLC6A3 (DAT1) have been associated with both mood disorder and ADHD. Most prominent is the overlap between genes for bipolar disorder and schizophrenia, as shown in Table 13.3–1. DAOA, NRG1, and DISC1 were each identified first through positional cloning studies of schizophrenia. Subsequently, they were shown to also be associated with bipolar disorder. These results are consistent with the family data and linkage data, where a large number of linkage peaks identified for bipolar disorder have also been reported for schizophrenia. They suggest that bipolar disorder and schizophrenia are more related in terms of genetics and mechanism than had been previously thought, as discussed below.

Table 13.3–1.
Chromosomal Regions and Genes Associated with Bipolar Disorder or Shared between Bipolar Disorder and Schizophrenia

	Bipolar Disorder Only	Bipolar Disorder and Schizophrenia
Clinical features	Mania Mood cycling	Psychosis Cognitive deficits Depression
Linkage regions	4p16 8q24 9q21 12q24 16p 18q22 21q21 Xq26	6q21–25 10p14 13q32 18p11 22q11–13
Candidate genes	SLC6A3 (5p15.33) HTR4 (5q32) DRD4 (11p15.5) DRD2 (11q32.2) HTR2A (13q14.2) SLC6A4 (17q11.2) MAOA (Xp11.3)	DISC1 (1q42.2) DTNBP1 (6p22.3) NRG1 (8p12) BDNF (11p14.1) DAOA (13q33.2) COMT (22q11.21)
GWAS significant genes	VRK2 (2p16.1) TRANK1 (3p22.2) NDST3 (4q26) ANK3 (10q21.2) CACNA1C (12p13.33) SLC25A17 (22q13.2)	LMAN2L (2q11.2) ITIH1/3/4 (3p21.1) FSTL5 (4q32.2) ADCY2 (5p15.31) TFAP2B (6p12.3) RPS6KA2 (6q27) MPP6 (7p15.3) IMMP2L (7q31.1) ADD3 (10q25.1) FADS2 (11q12.2) TENM4/ODZ4 (11q41.1) ZNF592 (15q25.3) HLF (17q22) NCAN (19p13.11) MACROD2 (20p12.1)

Genome-wide Association Studies

The advent of GWAS has completely revolutionized the field of psychiatric genetics. GWAS gained primacy over linkage several years ago but have only recently begun to provide replicable results. As described above, this technology enables the genotyping of hundreds of thousands of markers simultaneously, thereby allowing association studies to systematically cover the genome. Many of the early GWAS did not yield strong evidence for association for any single gene, likely due to the use of sample sizes that were underpowered to detect small genetic effects. However, several genes began to show trends toward association across multiple studies, lending confidence to the results. A calcium channel, CACNA1C showed such

correspondence between multiple studies, as did the gene encoding ankyrin G (ANK3). Other reported genes included the epidermal growth factor gene and diacylglycerol kinase eta (DGKH).

In 2007 the Psychiatric Genomics Consortium (PGC) was formed to combine GWAS datasets across the major psychiatric disorders in the hopes of identifying robust associations to genes and pathways underlying risk. By using a dataset consisting of more than 16,000 bipolar patients and controls, the Bipolar Working group of the PGC was able to confirm association to CACNA1C and ANK3 and identify a novel association to teneurin transmembrane protein 4 (TENM4). More recent analyses in an even larger sample of more than 52,000 individuals implicated 19 genes and regions, most of which reflect novel associations, as shown in Table 13.3–1. While the Depression Workgroup was initially unable to identify significant associations in a sample of over 53,000 individuals, ongoing efforts to increase the sample's size have finally begun to bear fruit, with nine significant associations now identified in a sample totaling over 219,000 patients and controls. The large difference in the number of subjects needed to detect associations reflects the higher phenotypic complexity and lower heritability of major depression compared to bipolar disorder, as well as gender differences and a greater role of gene–environment interactions However, it is anticipated that with increased sample sizes for mood disorders, additional genes will be identified that collectively explain a higher portion of the variation in disease risk. This phenomenon was successfully demonstrated by the Schizophrenia Workgroup, whereby quadrupling the sample to more than 150,000 patients and controls resulted in the identification of more than 100 additional significant associations and more than 30 percent of the risk for schizophrenia now explained by common genetic variants. It is expected that mood disorders will follow a similar pattern.

Several studies have shown that the specificity of the phenotype is much more important than large sample sizes in genetic analyses of complex disorders. The thought is that by using a more specific phenotype, the grouping of subjects in the analysis is more clinically similar and therefore more likely to share the same genetic influences. Thus, a better understanding and better utilization of subclinical phenotypes and comorbid states may lead to more powerful analyses for gene discovery. For example, a GWAS of temperament in a large bipolar sample revealed genome-wide significant associations to three distinct regions for the hyperthymic and irritable temperaments, where analysis of bipolar disorder as a phenotype failed to demonstrate significance to any region. This study demonstrates the potential power of quantitative phenotypes that may better index the underlying genetic and biological processes. Similarly, a recent study of comorbid eating disorders in the same bipolar disorder sample implicated the FXR1/SOX2-OT region, albeit at a level

slightly below genome-wide significance. However, this region demonstrated genome-wide significance in two independent and very large GWAS evaluating eating disorders and schizophrenia, respectively. Thus, an analysis of a small sample with a specific phenotype was able to replicate a prior association observed in related samples. Finally, another study examined the presence of comorbid externalizing disorders as a subclinical phenotype of bipolar disorder. Externalizing disorders included alcohol and substance abuse/dependence, pathological gambling, antisocial personality disorder, ADHD, and conduct disorder and were present in over 60 percent of the bipolar sample. The presence of such disorders significantly predicted a more severe course of illness characterized by an early age of onset of symptoms, more rapid mood switching, rapid cycling, more episodes of both depression and mania, and increased suicidality. This phenotype thus holds potential promise as a more severe bipolar phenotype to be evaluated in larger samples.

It is noteworthy that very little overlap is seen between the chromosomal regions implicated by linkage and GWAS of bipolar disorder. This is likely a reflection of the types of risk variants captured by each method. GWAS can only identify risk variants that are relatively common in the population, generally those with frequencies of 1 to 5 percent. While linkage can, in theory, identify both common and rare variants, the lack of convergence with GWAS data may be an indication that prior linkage studies have primarily detected rare variants in families. However, this remains to be verified through sequencing, as discussed below.

Studies of Rare Variation

While GWAS has begun to define some of the common variation contributing to risk for mood disorders, recent data suggest that many psychiatric disorders, including mood disorders, may result from multiple rare mutations of stronger effects. It is important to keep in mind that common SNPs currently explain only about 30 percent of the heritability of bipolar disorder, with the remaining genetic variance still to be discovered. This has been termed the "missing heritability," and rare variation has been proposed as the primary constituent. In this model, genetic variants are rare because they are the result of relatively recent mutations, and enough generations have not passed for them to be positively selected and increase in frequency or negatively selected against as deleterious and decrease in frequency. It is thought that rare variants may include more variants of large effect, such that in each individual, disease results from only one or a few variants, yet each variant is very rare in the population or even unique to a family.

Several different types of rare variants exist, ranging from large structural variants, such as copy number variants (CNVs), to single nucleotide variants (SNVs). CNVs have been proposed to explain a

significant portion of the genomic variation in humans. CNVs typically comprise a region of DNA that is duplicated many times, and the number of duplications is variable from individual to individual. Some are relatively stable and inherited, while others are spontaneous mutations in which one or more copies are added or deleted. The mechanism for these insertions and deletions is unclear but likely results from an error in DNA replication. These insertions and deletions range from a few thousand base pairs (kilobases) to many millions of base pairs (megabases). Typically, parts of one or more genes are deleted or interrupted such that gene function is eliminated or substantially altered. The human population contains a large number of very rare CNVs distributed across the genome.

While CNVs have been shown to play a significant casual role for some psychiatric disorders, particularly schizophrenia and autism, the role of CNVs in mood disorders is likely less prominent. Studies to date suggest that inherited stable CNVs do not contribute significantly to the cause of bipolar disorder, however, spontaneous CNV mutations do occur more often in bipolar families. One of the challenges in studying CNVs is their rarity; though it is possible to demonstrate an increased rate of CNVs in general, identifying specific causal CNVs is difficult. CNVs were originally studied using the same DNA chips used for GWAS. Differences in the signal intensity indicated differing numbers of copies of a region. However, this technology made it difficult to detect CNVs below a certain size range, typically 10,000 to 100,000 base pairs. Though there is a great deal of copy number variation below this size range, its contribution to bipolar disorder remains unclear.

In order to access variation below this size range, including SNVs, DNA sequencing must be performed. The last 15 years have brought a dramatic decrease in the cost of DNA sequencing from approximately $3 billion for the first genome to around a thousand dollars. Though still expensive, it has begun to make possible a complete examination of the genome. Studies of this nature come in two main varieties: exome sequencing and whole genome sequencing. By using methods that select for exonic DNA, all of the protein coding sequence in an individual, their exome, can be selectively sequenced. However, this enables only an examination of the 2 percent of the genome that codes for proteins and is unable to examine regulatory regions in the genome that control when genes are turned on or off. In order to fully examine the genome, including regulatory sequences, whole genome sequencing must be performed. For many years, it was felt that the 98 percent of the genome that does not code for proteins was largely without function, or “junk DNA.” However, more recently, this estimate has been reversed to indicate that almost all DNA is functional and has a role in regulation of gene expression.

To date, only a limited number of exome or whole genome sequencing studies have been performed. As most variants are rare, these analyses

typically compare the number of rare variants or functional variants between cases and controls on a gene-by-gene level. One study examining families using whole genome sequencing identified an excess of functional variants in genes in pathways related to voltage-gated calcium channel function and GABA receptors. These results fit well with those from GWAS, suggesting that susceptibility genes are likely affected by both common and rare variants. These studies also suggest that most of the variants involved in bipolar disorder are in regulatory regions of DNA and that most families are segregating several variants. Whole genome sequencing is the tool that delivers almost complete genomic information. The heritable causes of bipolar disorder will reside somewhere in these data. However, this era of human genetics is just beginning. Much larger sample sizes will be necessary to detect the rare variants that explain some portion of the missing heritability.

Molecular Pathways

Having discovered SNPs associated with bipolar disorder, the next challenge is to translate this into actionable biological knowledge. What abnormalities in neuronal function result from the causative SNPs, and how does this result in the symptoms of bipolar disorder? One approach to this important problem is to test gene pathways for enrichment of evidence of association. A large number of pathways are tested for enrichment, and the function of those pathways that are enriched may inform the cause of disease. This approach has highlighted a number of pathways of relevance, including CRF signaling, beta-adrenergic signaling, phospholipase signaling, glutamate signaling, and histone methylation. Pathway analysis, though informative, has a number of limitations, including the current availability of a very incomplete collection of pathways. As much of the pathway information comes from studies of other tissues and diseases, their relevance to the CNS is unclear. Furthermore, the pathways vary in quality, with some derived from direct protein–protein interaction experiments and others simply a correlation in levels of expression. As a result, many are large pathway categories comprising many genes, which, at this time, provide relatively limited information.

Functional Validation of Key Genes

Several significant SNPs and genes identified by GWAS have been shown to have an effect on biological functions relevant to psychiatric illness. These “post-GWAS” studies provide invaluable validation of the genetic results and begin to inform the biology of mood disorders. One of the most important findings to emerge from GWAS of bipolar disorder is the involvement of the voltage-gated calcium channel complex. Many subunits of this complex have shown association to bipolar disorder, and its role in

triggering neurotransmitter release in response to action potentials places it in a key regulatory position. A functional effect of the primary subunit involved, CACNA1C, has been suggested by several fMRI studies of bipolar disorder. CACNA1C genotype has been associated with reward responsiveness in healthy controls, anterior cingulate and hippocampal activation in episodic recall in controls, and amygdalar activation to emotional stimuli. Another prominent finding from GWAS is the gene ANK3, or giant-ankyrin. Ankyrin plays a role in anchoring components on the cell surface to the underlying cytoskeleton. It has recently been shown to also play a role in synapse formation and forming subdomains around AMPA receptor complexes. Knockdown of gene expression of ANK3 in mice resulted in lower anxiety and more activity, both of which were reversed by lithium. Though it is still early days for these functional studies, these are critical and the direction for the future, if lists of associated genes are to be transformed to a useful model of bipolar disorder pathways.

GENETIC OVERLAP BETWEEN MOOD AND OTHER DISORDERS

The relationship between bipolar disorder and schizophrenia has been debated since the distinction was made by Emil Kraeplin in 1899. Are they distinct disease entities, or do they exist on a spectrum? The definition of a disease entity should be based primarily on disease mechanism. Genetics is only one aspect of disease mechanism, but it does inform the discussion. This question has been explored in family studies by asking whether bipolar and schizophrenia "breed true." A seminal study of families of bipolar probands by Tsuang and colleagues revealed an increased risk for bipolar disorder, as well as a smaller but significant increase in the risk for schizophrenia. Similarly, in families of probands with schizophrenia, there is a clear increase in the risk for schizophrenia, as well as a smaller increase in the risk for bipolar disorder. This is consistent with the model of partial genetic overlap, wherein a portion of the genes for one illness also predisposes to the other. The Maudsley twin study supports this same conclusion. Patients were shown over time to vary in presentation, such that at some points they were diagnosed with bipolar disorder and at others with schizophrenia. Also, a substantial portion of the genetic variance was explained by genetic factors common to the two disorders.

This genetic commonality between schizophrenia and bipolar disorder bears on the nature of schizoaffective disorder. Though commonly observed clinically, schizoaffective disorder has been variously defined through the years. It is consistent with a spectrum of illness between bipolar disorder and schizophrenia, where schizoaffective disorder lies in the middle. The genetic relationship of schizoaffective disorder to mood disorders also involves the role of psychosis in the genetics of mood disorders. Studies examining familial risks in relatives of schizoaffective

patients have been undertaken in order to attempt to clarify this issue. However, they have led to largely inconsistent results, some finding an increased risk of schizophrenia and some an increased risk of bipolar disorder. A possible explanation is that schizoaffective disorder represents a mixture of patients, some with bipolar and some with schizophrenia diatheses. This is supported by data that find an increased rate of bipolar disorder among the relatives of probands with the bipolar subtype of schizoaffective disorder. Similarly, schizophrenia has been reported to be increased among the relatives of probands with the depressive type of schizoaffective disorder. However, at least from a genetic perspective, these results suggest that the Kraepelinian distinction between the disorders is only partially valid and that both schizophrenia and mood disorders lie at the extremes of a spectrum of a common genetic liability.

The partial genetic overlap between bipolar disorder and schizophrenia has been supported by the identification of chromosomal regions and genes involved in both disorders. Some of this genetic overlap is shown in Table 13.3–1. One study investigated the contribution of common genetic variation to several psychiatric disorders, as well as the degree of overlap between them. Such studies use GWAS data to estimate the contribution of common variants, or the SNP heritability. This study revealed that common variants contribute approximately 25 percent of the risk for bipolar disorder and that 68 percent of this risk is shared with schizophrenia, as shown in Figure 13.3–5. These estimates therefore reflect a total genetic overlap, or coheritability, of about 17 percent due to common variants. Furthermore, many of the genes from the most recent GWAS conducted by the PGC Bipolar Working Group reflect shared associations with schizophrenia, as shown in Table 13.3–1, and likely contribute to this observed coheritability. Both bipolar disorder and schizophrenia also show significant overlap with major depression, with more than 40 percent of the common variation being shared between these disorders. These numbers are likely to increase with increasingly larger sample sizes for these types of analyses. It is also important to note that these heritability and coheritability estimates do not account for the potential effects of rare risk variants, some of which may also be shared between disorders. This may account for the significantly lower heritability estimates observed using GWAS data compared to those obtained from twin and population-based studies.

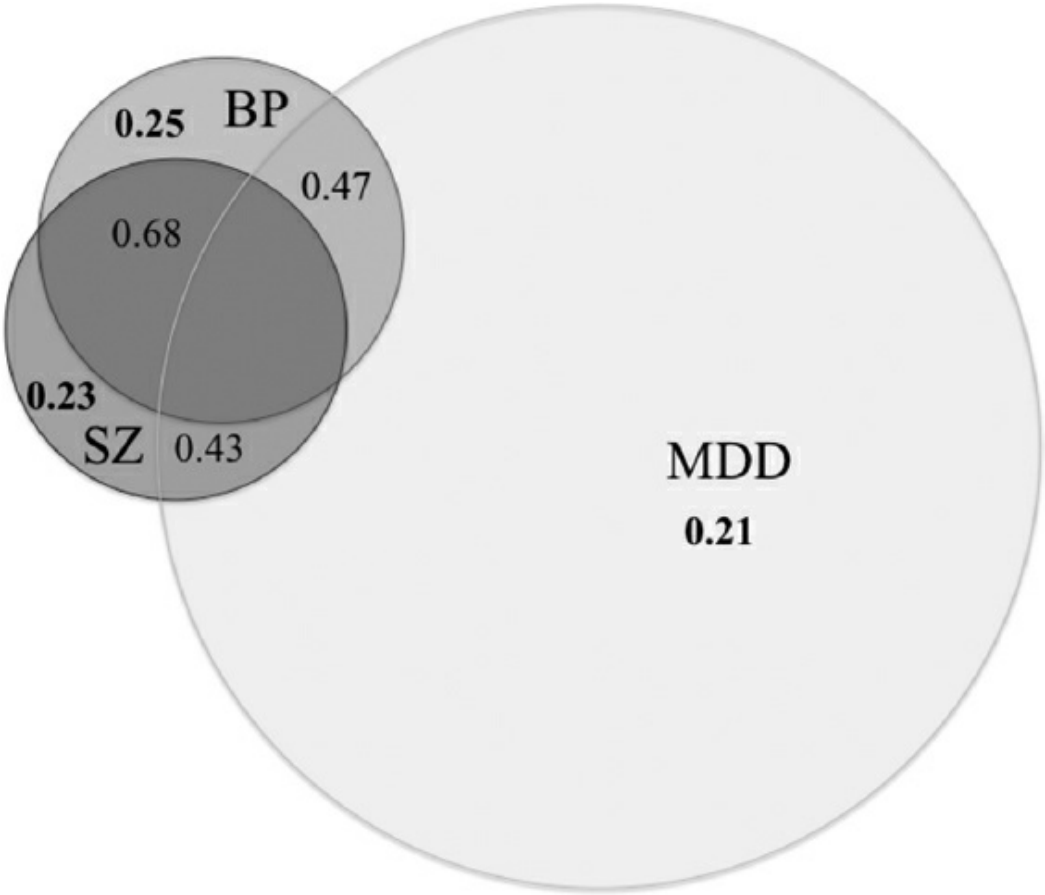

FIGURE 13.3–5. Genetic relationships between psychiatric disorders. This figure illustrates the genetic relationship between bipolar disorder (BP), major depression (MDD), and schizophrenia (SZ), each of which are indicated with respect to their prevalence rate. GWAS data were to evaluate the proportion of the genetic variance in each disorder that is due solely to common variation, or the SNP heritability. A common genetic vulnerability of >20% was shown for all three disorders, as indicated in bold text. These SNP heritability values are much lower than the heritability estimates derived from twin and population-based samples, which include all types of genetic variation and indicate a heritability of about 60% (vs. 25%) for bipolar disorder, 37% (vs. 21%) for major depression, and >60% (vs. 23%) for schizophrenia. Some genes are less specific and contribute to risk across multiple phenotypes. This is reflected in the shared common genetic variance, or coheritability, of 68% for bipolar disorder and schizophrenia and >40% for major depression with both disorders.

Genetic overlap with mood disorders has also been hypothesized for other psychiatric disorders. ADHD, for example, may share some genetic risk factors with bipolar disorder. Risk for ADHD has been shown to be elevated in the children of bipolar probands, and the risk for bipolar disorder is elevated in the parents of ADHD probands. Distinguishing these two disorders is a common diagnostic dilemma that may result in part from shared genetic contributions and mechanisms. However, studies of SNP coheritability have not revealed a significant genetic overlap as might be expected from the family data. The reason for this is unclear, but it may involve samples that are too small to detect an effect or perhaps a more prominent role for rare variants.

PHARMACOGENETICS

One of the first clinical applications of genetics in psychiatry may come from the newly emerging field of pharmacogenetics. Medication responses in mood disorders frequently have robust individual differences. Though response has been related to certain clinical factors, genetic differences likely play a major role. Much of the focus of this field to date has been on drug metabolism. Genetic variation in the cytochrome P450 system influences the degradation of a number of drugs. This can result in wide variation in serum level in different patients on the same dose. Monitoring such effects may help explain the lack of efficacy of drugs in some patients or predict possible toxicity due to high serum drug levels in poor metabolizers. Much, if not most, of the variation in response is likely due not to hepatic metabolism, but rather variation in genes that are drug targets or involved in the pathways modulated by drugs. One of the most intriguing aspects of pharmacogenetics is the implication for disease mechanism. A drug may work in one patient with depression and not another because the two patients have illnesses that are biochemically distinct. In other words, a drug may be effective only in a subset of patients whose illness involves dysfunction of the pathway that is modulated by the drug. A gene variant may predict response because it is associated with a form of mood disorder that involves that pathway. In this way, drug response may be a powerful phenotype for dissecting mechanistically different forms of illness. Alternatively, the gene variant may impact the ability of the drug to modulate its target pathway and have little effect on disease vulnerability.

Several other genes have been reported to predict drug response in mood disorder. A G protein subunit gene (GNB3) has been reported to predict response to antidepressants and lithium. BDNF and the gene for its receptor, NTRK2, have been associated with lithium response. In addition to influencing SSRI response, the serotonin transporter has also been reported to predict antidepressant-induced mania and antidepressant response to sleep deprivation. The risk for antidepressant-induced mania has also been shown to be associated with this gene. Recently, GWAS results have been reported for lithium response, including a large collaborative sample of over 2,500 patients retrospectively assessed for lithium response. This study identified a long noncoding RNA (lncRNA) significantly associated with lithium response, which was replicated in a smaller prospective sample. The intriguing aspect of this finding is that the lncRNA likely regulates the expression of a group of genes that may provide further information regarding genes involved in lithium response.

Other studies have focused on the clinical effectiveness of a panel of SNPs associated with drug response. These panels largely comprised SNP variants in CYP 450 enzymes along with some pharmacodynamics SNPs. When doctors and patients were randomized as to whether the doctor

received the results of such tests, the group with the pharmacogenetics data had superior outcome. Much of this effect resulted from avoiding medications that the patient could not metabolize and would thus cause intolerable side effects. Although research in this area is still in its infancy, it promises a powerful tool for future clinicians who one day may use a DNA test panel to optimize the selection of medications for individual patients, thereby reducing the trial and error involved in medication treatment.

GENETIC COUNSELING

Based on the observation of psychiatric illness in their own families or the increasing public awareness of psychiatric genetics, patients frequently ask clinicians two questions: Are mood disorders genetic? What is the risk to my children or grandchildren? The answer to the first question is easy: Yes, mood disorders are genetic, as is evidenced by the large body of epidemiological data summarized in this chapter. However, it is important that patients understand that mood disorders are only in part genetic. The twin studies argue strongly that only 70 to 80 percent of the etiology of mood disorders is genetic. Therefore, it is a predisposition to illness that is inherited and interacts with other nonheritable factors.

The risk to children and grandchildren is the more difficult question and the one that deserves the greatest consideration. The family data indicate that if one parent has a mood disorder, a child will have a risk for mood disorder of between 10 and 25 percent. If both parents are affected, this risk roughly doubles. It is important to take a careful family history to more accurately predict risk for a specific family. Several factors from the family history should be considered. The more members of the family that are affected, the greater the risk to a child. The risk is also greater if the affected family members are first-degree relatives, rather than more distant relatives. A family history of bipolar disorder conveys a greater risk for mood disorders in general and, specifically, a much greater risk for bipolar disorder. The presence of more severe illness in the family also likely conveys a greater risk. These factors should be considered together in forming an estimate of risk for the concerned parent.

Equally important to providing the estimate of risk is providing guidance in interpreting and responding to that information. Patient's reactions to risk information will vary greatly depending on his or her own personal experience with the illness. Some will be relieved that it is so low, whereas others will be fearful that it is so high. It is important to emphasize that their child carries a risk or predisposition to illness, rather than a certainty of illness. It is also useful to emphasize the range of illness, from mild to severe, that could result, as well as the availability and efficacy of treatment. Ultimately, the use of such information in family planning is a highly personal decision. Some patients may choose to not have children.

For existing children, it is important to educate parents about the typical age of onset, presenting symptoms, and the importance of early recognition and treatment. However, this must be balanced with the goal of not labeling the child or being overly protective.

FUTURE DIRECTIONS AND IMPLICATIONS OF IDENTIFYING SUSCEPTIBILITY GENES

Though much is understood about the familiality and heritability of mood disorders, the identification of specific genes has been challenging. Studies to date have reproducibly identified a number of genes; though, these genes together explain only a small portion of the genetic variance. It remains unclear how many genes are involved and how the illness is transmitted.

New technologies are now being used that may yield answers to some of these questions. GWAS have shown that common variants contribute significantly to heritability. Yet, mood disorders may not be primarily polygenic and the result of common alleles. It appears likely that they are highly heterogeneous and that cases may result from either many common variants or very rare mutations of strong effect. To identify these very rare variants, each individual would need to be fully sequenced. As discussed above, these studies are just beginning but already have begun to provide information about the relationship between common and rare variants in disease causation.

The availability of large datasets generated by DNA sequencing thousands of subjects and millions of SNPs will also enable the testing for a variety of interactions. It is very likely that much of the genetic variance is explained by gene–gene interactions. Genes may interact to produce a risk for illness greater than the sum of their individual effects. It is presently, unclear how important such interactions are for mood disorders. Detection of interaction effects is only possible with large datasets. Ultimately, a more comprehensive understanding of the genetics of mood disorder may entail complex models of the interaction of numerous genes. As a better understanding of environmental effects becomes available, these will be incorporated into such models. Although some environmental effects may be nonspecific, it seems likely that at least some will reflect specific gene–environment interactions. Such interactions may be difficult to detect and will add substantial complexity to the overall model of mood disorder genetics. At this point, it seems promising that most genes for mood disorders can be identified through such large, though tractable, projects.

What will be the implications of this knowledge? The most immediate impact may well be on public opinion. Psychiatrically ill persons have long experienced the stigma and discrimination that portions of the public impose out of fear and ignorance. The acceptance of major mental illnesses as brain disorders has been slow. The definitive identification of genetic causes may have a highly beneficial impact on public understanding and

acceptance.

This knowledge should also have a dramatic impact on the understanding of pathophysiology and approaches to treatment for mood disorders. Most current theories of pathophysiology are based on the mode of action of therapeutic agents, which were, for the most part, discovered serendipitously. The site of action of therapeutic drugs is not necessarily either the site of the genetic defect or the primary site of the pathophysiology. The identification of disease genes may point to entirely new systems involved in the pathophysiology or components of currently implicated systems that were not previously known. It is hoped that such an understanding may lead to the rational drug design long sought by patients, clinicians, and pharmaceutical companies. The result may be new drugs that act via completely novel mechanisms, possibly with greater efficacy and specificity.

The identification of disease genes will likely have a major impact on diagnosis and nosology. Just as the diagnosis of jaundice has given way to a classification scheme based on pathophysiology, it is likely that, in the future, the diagnosis of major mood disorders may be specific to disease mechanism. Mechanism-specific diagnoses may dictate the use of different, more specific, and more effective treatments. Finally, such knowledge will carry with it danger and responsibility. Premorbid DNA testing that would indicate the degree of genetic vulnerability to major mood disorders could become available. Those with psychiatric disorders will then face the same issues of genetic testing currently faced by families with Huntington disease or breast cancer: Will at-risk family members want such information? How will they use it? How will they cope with it? How can psychiatrists assist them in these decisions? Finally, how can patients be protected from discrimination based on such information?

In summary, genetic studies promise a new era of understanding and treatment of mood disorders. Identification of genes alone will not elucidate pathophysiology but will merely point the way for the application of modern neuroscience methods in the equally large task of understanding mechanisms. Recent results suggest that such guidance may not be far away.

References

Akiskal HS. The prevalent clinical spectrum of bipolar disorders: beyond DSM-IV. *J Clin Psychopharmacol*. 1996;16:4S–14S.

*Ament SA, Szelinger S, Glusman G, et al. Rare variants in neuronal excitability genes influence risk for bipolar disorder. *Proc Natl Acad Sci U S A*. 2015;112:3576–3581.

Badner JA, Gershon ES. Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. *Mol Psychiatry*. 2002;7:405–411.

Caspi A, Sugden K, Moffitt TE, et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*. 2003;301:386–389.

Craddock N, O'Donovan MC, Owen MJ. Genes for schizophrenia and bipolar disorder? Implications for psychiatric nosology. *Schizophr Bull*. 2006;32:9–16.

*Cross-Disorder Group of the Psychiatric Genomics Consortium. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet*. 2013;45:984–994.

Evans L, Akiskal HS, Keck PE Jr, et al. Familiality of temperament in bipolar disorder: support for a genetic spectrum. *J Affect Disord*. 2005;85:153–168.

Gershon ES, Hamovit J, Guroff JJ, et al. A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. *Arch Gen Psychiatry*. 1982;39:1157–1167.

Greenwood TA, Akiskal HS, Akiskal KK, Study BG, Kelsoe JR. Genome-wide association study of temperament in bipolar disorder reveals significant associations with three novel Loci. *Biol Psychiatry*. 2012;72:303–310.

Greenwood TA, Badner JA, Byerley W, et al. Heritability and genome-wide SNP linkage analysis of temperament in bipolar disorder. *J Affect Disord*. 2013;150:1031–1040.

Heun R, Maier W. The distinction of bipolar II disorder from bipolar I and recurrent unipolar depression: results of a controlled family study. *Acta Psychiatr Scand*. 1993;87:279–284.

Kelsoe JR. Arguments for the genetic basis of the bipolar spectrum. *J Affect Disord*. 2003;73:183–197.

*Lichtenstein P, Yip BH, Bjork C, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet*. 2009;373:234–239.

Liu X, Kelsoe JR, Greenwood TA, The Bipolar Genome Study. A genome-wide association study of bipolar disorder with comorbid eating disorder replicates the SOX2-OT region. *J Affect Disord*. 2016;189:141–149.

McQueen MB, Devlin B, Faraone SV, et al. Combined analysis from eleven linkage studies of bipolar disorder provides strong evidence of susceptibility loci on chromosomes 6q and 8q. *Am J Hum Genet*. 2005;77:582–595.

Mondimore FM, Zandi PP, Mackinnon DF, et al. Familial aggregation of illness chronicity in recurrent, early-onset major depression pedigrees. *Am J Psychiatry*. 2006;163:1554–1560.

Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci*. 2015;18:199–209.

Nurnberger JI Jr, Koller DL, Jung J, et al; Psychiatric Genomics Consortium Bipolar Group. Identification of pathways for bipolar disorder: a meta-analysis. *JAMA Psychiatry*. 2014;71:657–664.

Oedegaard KJ, Fasmer OB. Is migraine in unipolar depressed patients a bipolar spectrum trait? *J Affect Disord*. 2005;84:233–242.

PGC Bipolar Disorder Working Group. *GWAS of 18,000 Cases Identifies 19 Loci for Bipolar Disorder*. Toronto: World Congress on Psychiatric Genetics; 2015.

*Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet*. 2011;43:977–983.

Sadovnick AD, Remick RA, Lam R, et al. Mood Disorder Service Genetic Database: morbidity risks for mood disorders in 3,942 first-degree relatives of 671 index cases with single depression, recurrent depression, bipolar I, or bipolar II. *Am J Med Genet*. 1994;54:132–140.

*Saunders EH, Scott LJ, McInnis MG, Burmeister M. Familiality and diagnostic patterns of subphenotypes in the National Institutes of Mental Health bipolar sample. *Am J Med Genet B Neuropsychiatr Genet*. 2008;147B:18–26.

Segurado R, Detera-Wadleigh SD, Levinson DF, et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: bipolar disorder. *Am J Hum Genet*. 2003;73:49–62.

*Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet*. 2012;13:537–551.

Swaminathan S, Koller DL, Foroud T, et al. Characteristics of bipolar I patients grouped by externalizing disorders. *J Affect Disord*. 2015;178:206–214.

Tsuang MT, Faraone SV. *The Genetics of Mood Disorders*. Baltimore: Johns Hopkins University Press; 1990.

Wray NR, Gottesman II. Using summary data from the Danish national registers to estimate heritabilities for schizophrenia, bipolar disorder, and major depressive disorder. *Front Genet*. 2012;3:118.

Zubenko GS, Maher B, Hughes HB 3rd, et al. Genome-wide linkage survey for genetic loci that influence the development of depressive disorders in families with recurrent, early-onset, major depression. *Am J Med Genet*. 2003;123B:1–18.

▲ 13.4 Mood Disorders: Clinical Features

HAGOP S. AKISKAL, M.D.

TERMINOLOGY

Mood disorders are characterized by pervasive dysregulation of mood and psychomotor activity and by related biorhythmic and cognitive disturbances. The rubric of *affective disorder,* which, in some European classifications, also subsumes morbid anxiety states, is increasingly being replaced by the nosologically more delimited concept of *mood disorder*. Thus, *mood disorder* is now the preferred term in the WHO's tenth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) and the American Psychiatric Association's (APA's) fifth edition of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). Official mood disorder categories in current use include bipolar disorders (with manic or hypomanic, depressive, or mixed episodes) and major depressive disorders and their respective attenuated variants known as *cyclothymic* and *dysthymic disorders* (called persistent depressive disorder in DSM-5). The term *unipolar* is avoided in DSM-5 because major depressive disorder, especially when early in age at onset, is an unstable diagnosis and is liable to switch to bipolar. Conditions that, in earlier editions of these manuals, were categorized as *endogenous depression, involutional melancholia,* and *psychotic depressive reaction* have been incorporated into major depressive disorder. Although the neurotic–endogenous distinction has been officially deleted, the term *melancholic features* is now used to qualify major depressive disorders in which biological concomitants predominate. The United States and international classifications recognize the common occurrence of mixed anxiety–depressions, which underscores the value of the broader *affective disorder* rubric. It is uncertain whether classic neurasthenic conditions, which have recently reemerged and overlap to some extent with the so-called chronic fatigue syndrome, should be classified under this broad rubric.

RECENT DIAGNOSTIC TRENDS

The reshuffling and reclassification of various affective conditions into the mood disorders chapter of the APA classification in DSM-5 have, on balance, considerably broadened their boundaries. This change reflects, in part, developments in pharmacotherapy that have resulted in considerable alleviation of suffering for persons whose illnesses fall short of and sometimes beyond the boundaries of classic mood disorders. As a result, many individuals with recurrent mood disorders who would have been

disabled can now expect to lead a life relatively free of major episodes and major personal and occupational disruptions. Such gratifying results have, in turn, helped to destigmatize this group of disorders. Destigmatization has been further facilitated by published self-revelations of famous persons with depressive and bipolar disorders. Unfortunately, access to competent mental health care is not available to large segments of the population. Moreover, destigmatization, which has led to global efforts to detect and to treat mood disorders—particularly depression—in public and private sectors, may have had unintended sequelae. Despite the obvious public health advantages of such destigmatization, some authorities nonetheless now contend that a global *depression-centric* perspective has emerged in psychiatric and general medical sectors at the expense of bipolar disorders and may have led to overtreatment with antidepressants and unmasking of latent bipolar disorders. Others believe that the opposite to be true—that it is bipolar disorder that is overdiagnosed and mood stabilizers that are overused. The data, however, indicate that, on balance, it is the continuing underdiagnosis of bipolar disorder that presents the greatest threat to public health. Indeed, recent epidemiologic data suggest considerable disability in untreated or inadequately treated bipolar disorder.

SPECTRUM OF MOOD DISORDERS

The foregoing considerations notwithstanding, the increasing clinical diagnosis of major depressive and bipolar disorders should not be dismissed as mere therapeutic fad. External validating strategies, such as familial-genetic studies and prospective follow-up, can now be used to buttress the broadened concept of mood disorders, including that of bipolar disorder. New research comparing monozygotic and dizygotic twins has demonstrated that the genetic propensity to mood disorders embraces entities that extend beyond endogenous depression (melancholia) to subsume a larger variety of depressions, including some encountered in persons in the community who have never received psychiatric treatment. Although such data might seem counterintuitive to those who would restrict depression to a core primary "biological" disease, they suggest that the constitutional predisposition for affective dysregulation occurs in as many as one of every three persons. That ratio is similar to the proportion of those who progress to a full depressive syndrome after bereavement, rhesus monkeys developing depressive-like behavior after a separation paradigm, and dogs that develop learned helplessness after inescapable shock. The fact that these rates are considerably higher than those observed in clinical populations suggests that many people possess protective factors against major depressive episodes; alternatively, the data suggest that other factors determine which person with emotional distress becomes a clinical case. A great deal might therefore be revealed about the nature of pathological affective processes through the study of self-limiting affective

conditions on the border of mood disorders.

Evidence from long-term prospective studies indicates that the course of mood disorders consists of various gradations of affective oscillations from the subsyndromal level to the syndromic level. Major episodes represent an operational convention to define a clinical threshold. Subchronic course, dominated by subthreshold symptoms, occurs in 50 to 60 percent. This is equally true for major depressive disorder and bipolar disorders. Detecting the illness in its subthreshold stage and treating interepisodic subthreshold phases are important clinical considerations.

The suffering and dysfunction resulting from mood disorders are among the most common reasons for consulting psychiatrists and other physicians. In fully developed cases of depression, all activity stops—including creative powers—and life is grim and in total disarray, as portrayed in Albrecht Dürer's engraving (shown in Fig. 13.4–1) of interest, the ladder and the angel-like portrayal of the melancholic subject in this engraving promise that the illness that can bring descent into the hell of depression can also permit ascent into the heavens and creativity. Such creativity, however, is largely limited to a few talented individuals with the milder forms of bipolarity, such as bipolar II and at the temperamental level. Many with the manic extremes of the illness have poor judgment and insight, refuse treatment, and suffer a deteriorative course.

All great physicians of the past, beginning with Hippocrates, have devoted considerable space in their general medical texts to the clinical characterization of melancholic and manic states, as well as their alternations in the same patient. The relationship of anxious and melancholic states was also known. A broad spectrum of affective disturbances—ranging from the relatively mild temperamental variants (represented in the official US nosology by dysthymic and cyclothymic disorders) to their severest forms (including what today is considered mood disorder with mood-congruent and mood-incongruent psychotic features)—has been described in classical medical and psychiatric treatises. Finally, classical authors noted that melancholia and certain physical diseases shared seasonal exacerbation, and they described the common occurrence of alcohol indulgence, especially in individuals prone to mania. These boundary problems continue to pose challenges today, with the addition of substance use disorders.

FIGURE 13.4–1. *Melancholia* by Albrecht Dürer (1471 to 1528).

AFFECTS, MOODS, TEMPERAMENTS, AND MORBID MOOD STATES

Ethological Considerations

Affects and *moods* refer to different aspects of emotion. Affect is communicated through facial expression, vocal inflection, gestures, and posture and (according to current ethological research) is intended to move human beings and other primates to appraise whether an individual is satisfied, distressed, disgusted, or in danger. Thus, joy, sadness, anger, and fear are basic affects that serve a communicative function in primates as well as many in other mammalian species.

Affects tend to be short-lived expressions reflecting momentary emotional contingencies. Moods convey sustained emotions; their more enduring nature means that they are experienced long enough to be felt inwardly. Moods are also manifested in subtle ways, and their accurate assessment often requires empathic understanding by the interviewer. The words that persons use to describe their inner emotions may coincide with the technical terms used by researchers or clinicians and often vary from one culture to another. Furthermore, the inward emotion and the prevailing affective tone may be discordant. This conflict could be due to deliberate simulation (i.e., the person does not wish to reveal his or her inner emotion), or it could result from a pathological lesion or process that has altered the emotions and their neural substrates. Thus, evaluating moods and affective expression requires considerable clinical experience. Such evaluation is obviously of great importance in differential diagnosis. Finally, it is of great usefulness in the therapeutic process; the admonition to "go after the affect" in psychotherapy helps to find where it hurts, thereby making treatment possible.

Sadness and Joy

The normal emotions of sadness and joy are part of everyday life and should be differentiated from major depressive disorder and mania. Sadness, or normal depression, is a universal human response to defeat, disappointment, or other adversities. The response may be adaptive, in an evolutionary sense, by permitting withdrawal to conserve inner resources, or it might signal the need for support from significant others. Transient depressive periods also occur as reactions to certain holidays or anniversaries, as well as during the premenstrual phase and the first week postpartum. Termed, respectively, *holiday blues, anniversary reactions,* premenstrual tension, and *maternity blues,* they are not psychopathological per se, but individuals predisposed to mood disorder may develop clinical depression during such times.

Grief

Normal bereavement or grief, considered the prototype of reactive depression, occurs in response to significant separations and losses, such as death, divorce, romantic disappointment, leaving familiar environments, forced emigration, or civilian catastrophes DSM-5 tends to limit the concept of normal grief to loss due to the death of a loved one—a condition that it considers as an exclusionary criterion for major depression. However, the work of Elie Karam and colleagues showed that bereavement and other losses associated with the civil war in Lebanon served as potent forces in depression formation. The same was true for the losses due to the earthquake in Armenia. The boundary behavior and clinical depression are

blurred in reactions to such complex losses. In addition to depressed affect that is appropriate to the loss, bereavement reactions are characterized by the prominence of sympathetic arousal and restlessness believed to represent, from an evolutionary perspective, physiological and behavioral mechanisms to facilitate the search for the lost object. Like other adversities, bereavement and loss do not generally seem to cause depressive disorder, except in individuals predisposed to mood disorder. Reactions to major catastrophes might represent a partial exception.

Elation

The positive emotion of elation is popularly linked to success and achievement. However, paradoxical depressions may also follow such positive events, possibly because of the increased responsibilities that often have to be faced alone. Elation is conceptualized psychodynamically as a defense against depression or as a denial of the pain of loss, as exemplified by *maniacal grief,* a rare form of bereavement reaction in which elated hyperactivity may replace the expected grief. The character of the "merry widow" in opera seems to be of similar derivation.

Other pseudomanic states include the brief energetic and unusually lucid periods encountered in dying patients or in those who need to take superhuman action in the face of unusual duress, both of which have been conceptualized as a *flight into health*. In predisposed persons, such reactions might be the prelude to a genuine manic episode. Sleep deprivation, which commonly accompanies major stressors, might represent one of the intermediary mechanisms between stressor and adverse clinical outcome.

Affective Temperaments

Another mediating factor between normal and pathological moods is temperament. Most persons have a characteristic pattern of basal affective oscillations that defines their temperament. For instance, some are easily moved to tears by sad or happy circumstances, whereas others tend to remain placid. Normally, oscillations in affective tone are relatively minor, tend to resonate with day-to-day events, and do not interfere with functioning. Some persons exhibit greater variability of emotional responses whereby, with no obvious provocation, the person alternates between normal mood and sadness or elation. Temperament traits tend to cluster into basic types. A worrying temperament associated with generalized anxiety disorder is often complicated by depressive episode: It overlaps considerably with the depressive temperament; such inclination to melancholy makes the person more easily sink into weary and sad moods and occurs in 3 to 6 percent of the general population. The hyperthymic temperament, in which the person is naturally inclined toward cheerful

moods, has been reported in 2 to 5 percent; the cyclothymic temperament, swinging between cheerful and sad moods, characterizes 4 to 6 percent of young adults. All four types have an early insidious onset and tend to persist throughout adult life. Of interest, these temperaments may be the prelude to episodes of depressive or hypomanic and manic polarity, which underscores the inherent instability of temperamental inclinations. Marked irritable-explosive traits occur in 2 to 3 percent of young persons and tend to attenuate by middle age. Current data suggest that such traits often coexist with the mood–labile cyclothymic type, representing the dark "borderline" side of this temperament.

An examination of the traits associated with these temperaments can provide the rationale for the hypothesis about the social and evolutionary functions that they subserve (Table 13.4–1). Thus, the person with a depressive temperament is hard working, dependable, sensitive to the suffering and needs of others, and suitable for jobs that require long periods of devotion to meticulous detail; such persons are said to shoulder the burdens of existence without experiencing its pleasures. A person with the hyperthymic temperament, endowed with high levels of energy, extroversion, and humor, assumes leadership positions in society or excels as a performer in media or entertainment; such persons are also successful from a Darwinian perspective in that they are adept in amorous advances and engender a large number of offspring. In talented persons, the cyclothymic temperament, which alternates between sadness and elation, could provide the inspiration for love and for the emotional intensity needed for composing music, writing poetry, and painting. A person with the irritable temperament might be best suited for a military career or even revolutionary action. The danger with persons with extreme temperaments is that they could swing too far in one or the other direction or in both directions (i.e., major depressive, manic, or mixed episodes). Use of such substances as alcohol, caffeine, and other stimulants might further destabilize affective regulation in persons with those attributes. Some adolescent girls with irritable cyclothymia might develop the extreme emotional disequilibrium that, in contemporary psychiatry, is considered borderline personality disorder. Temperament concepts can enrich understanding of the boundary between normal moods and emotional disorders and can supplement the DSM-5 descriptors of personality disorders with valuable information about individual vulnerability and assets. A new instrument—the Temperament Evaluation of Memphis, Pisa, Paris and San Diego—has been developed in an autoquestionnaire version (TEMPS-A) and validated in at least 12 languages.

Table 13.4–1.
Attributes, Assets, and Liabilities of Depressive and Hyperthymic Types

Derived from Classic Concepts of Temperaments

Depressive	Hyperthymic
Gloomy, incapable of fun, complaining	Cheerful and exuberant
Humorless	Articulate and jocular
Pessimistic and given to brooding	Overoptimistic and carefree
Guilt-prone, low self-esteem, and preoccupied with inadequacy or failure	Overconfident, self-assured, boastful megalomaniac
Introverted, with restricted social life	Extroverted and people seeking
Sluggish, living a life out of action	High energy level, full of plans
Few but constant interests	Versatile with broad interests
Passive and "sensitive"	Overinvolved and meddlesome
Reliable, dependable, and devoted	Uninhibited and stimulus seeking

Morbid Mood States

Mood disorders represent abnormal or extreme variations in mood and associated manifestations and are characterized by the following features: pathological mood change, endoreactive moods, recurrence, and impairment.

Pathological Mood Change. Pathological moods are distinguished from their normal counterparts by being out of proportion to any concurrent stressor or situation; being unresponsive to reassurance; being sustained for weeks, months, and, sometimes, years; and having a pervasive effect on the person, such that judgment is seriously compromised by the mood.

Endoreactive Moods. Depression and mania are diagnosed, respectively, when sadness or elation is overly intense and continues beyond the expected impact of a stressful life event. Indeed, the morbid mood might arise without apparent or significant life stress. The pathological process in mood disorders is thus partly defined by the ease with which an intense emotional state is released and, especially, by its tendency to persist autonomously even when the offending stressor is no longer operative. Rather than being endogenous (i.e., occurring in the absence of precipitants), mood disorders are best conceptualized as endoreactive (i.e., once released, they tend to persist autonomously). The homeostatic dyscontrol of mood, which is part of a more pervasive mood dysregulation, resists reversal to the habitual baseline affective tone.

Recurrence. In a more descriptive vein, what sets mood disorders apart from their normal emotional counterparts is the clustering of signs and symptoms into discrete syndromes that typically recur on an episodic basis or pursue an intermittent, subthreshold course over the span of many years, if not a lifetime. Cyclic course and, in some cases, regular periodic recurrence are other signs of mood dysregulation that are particularly relevant to bipolar disorder.

Impairment. Normative reactions to adversity and stress, including biological stress, typically consist of transient admixtures of anxiety and dysphoria that are best captured under the DSM-5 rubric of adjustment disorder with mixed anxiety and depressed mood. That is, self-limiting reactions are best qualified broadly as normal affective states that produce little, if any, impairment in the main areas of functioning.

Although anxiety, irritability, and anger do occur in various types of mood disorders, pathologically sustained mood states of depression and elation characterize those disorders. Morbid mood states (mood disorders) then consist of protracted emotional reactions that deepen or escalate, respectively, into clinical depression or mania, with a tendency to recur or to evolve into unremitting chronicity in 15 to 20 percent of cases. The contribution of temperamental peculiarities to such outcomes should be apparent. The impaired functioning characteristic of mood disorders is thus based on a combination of factors, including severity, autonomy, recurrence, and chronicity of the clinical features.

To recapitulate, dysregulation in mood disorders can take different forms. It could be expressed as a single severe episode that persists autonomously for many months and sometimes years, or it might recur with episodes of varying severity, years apart or in rapid succession, with or without interepisodic remission. In general, the earlier the age at onset, the more likely are recurrences, especially those of bipolar nature. Thus, depending on the course of the illness, impairment could be state dependent and could occur during an episode, or it could extend into the interepisodic period. According to National Institute of Mental Health (NIMH) estimates, on average, a woman with bipolar disorder spends 12 years in florid episodes (often hospitalized), loses 14 years from a productive career and motherhood, and has her life curtailed by 9 years. More recent weekly prospective observations over up to two decades in the NIMH Collaborative Depression Study have shown that patients with bipolar disorder I are symptomatic 47 percent of the time, and bipolar disorder II patients are symptomatic 53 percent of the time, much of it spent in subthreshold depression. In this study, unipolar major depressive disorder was even more pervasive, with subacute chronicity 59 percent of the time.

Recent observations have also revealed another pattern of impairment. In dysthymic and cyclothymic disorders, which represent an intensification of temperamental instability, impairment is not due to the severity of the mood disturbance per se but to the cumulative impact of the dysregulation beginning in the juvenile or early adult years and continuing unabated or intermittently over long periods; hence, the frequent confusion with character pathology. Here, the impairment is more subtle but nonetheless is pervasive. Persons with cyclothymic disorder tend to be dilettantes, whereas those with dysthymic disorder often lead morose and colorless

lives. The fundamental causes of mood disorders must be sought in the preclinical expressions in the offspring of adults with these disorders.

PSYCHOPATHOLOGY

Depressive Syndrome

Like other illnesses, depressive disorder clusters into signs and symptoms that constitute what DSM-5 and ICD-10 term *major depressive episode* (Table 13.4–2). These criteria attempt to set an operational threshold for depressive disorder based on a specified number of items and their temporal patterns. The diagnosis of clinical depression cannot be accomplished by a checklist: The DSM-5 diagnostic criteria for major depressive disorder provide a general guide. Only after an in-depth phenomenological approach can a clinician ascertain the diagnosis of a depressive disorder. Disturbances in all four spheres (mood, psychomotor activity, cognitive, and vegetative) should be ordinarily present for a definitive diagnosis of major depressive disorder.

Mood Disturbances. Mood change, usually considered the sine qua non of morbid depression, manifests in a variety of disturbances, including (1) painful arousal, (2) hypersensitivity to unpleasant events, (3) insensitivity to pleasant events, (4) insensitivity to unpleasant events, (5) reduced anticipatory pleasure, (6) anhedonia or reduced consummatory pleasure, (7) affective blunting, and (8) apathy. The phenomenology and psychometric properties of this broad range of mood disturbances are under investigation at the Salpêtrière Hospital in Paris. The focus in the description that follows is primarily on painfully aroused mood (depression) and diminished capacity for pleasure (anhedonia), two mood disturbances given selective weight in DSM-5 and ICD-10.

DEPRESSED MOOD. The term *depressed mood* refers to negative affective arousal, variously described as *depressed, anguished, mournful, irritable,* or *anxious*. These descriptions tend to trivialize a morbidly painful emotion, typically experienced as worse than the severest physical pain. Thus, depressed mood has a somatic quality that, in the extreme, is indescribably painful. Even when not so severe, depressive suffering is qualitatively distinct from its "neurotic" counterparts, taking the form of groundless apprehensions with severe inner turmoil and torment. This description is particularly apt for middle-aged and elderly persons, who were once considered to be experiencing *involutional melancholia*. The sustained nature of the mood permits no respite, although it tends to lift somewhat in the evening. Suicide may represent an attempt to find deliverance from such unrelenting psychic torment; death can be conceived as comforting.

Patients with a milder form of the malady typically seen in primary care

settings might deny experiencing mournful moods and instead complain of physical agony from headache epigastric pain, precordial distress, and so on, in the absence of any evidence of diagnosable physical illness. Such conditions have been described as *depressio sine depressione* or *masked depression*. In such cases, which are commonly observed in older patients, the physician should corroborate the presence of mood disturbance by the depressed affect in the patient's facial expression, vocal inflection, and overall appearance.

ANHEDONIA AND LOSS OF INTEREST. Paradoxically, the heightened perception of pain in many persons with depressive disorder is accompanied by an inability to experience normal emotions. Patients exhibiting the disturbance may lose the capacity to cry, a deficit that is reversed as the depression is lifting.

In evaluating anhedonia, inquiring whether the patient has lost the sense of pleasure is not enough; the clinician must document that the patient has actually given up previously enjoyed pastimes. When mild, anhedonia evidences with a decreased interest in life. Later, patients complain that they have lost all interest in things. This is best illustrated in Hamlet's disgust: "How weary, stale, flat, and unprofitable seem to me all the uses of the world" (William Shakespeare, *Hamlet,* Act I, Scene II). In the extreme, patients lose their feelings for their children or spouses, who once were a source of joy. Thus, the hedonic deficit in clinical depression might represent a special instance of a more pervasive inability to experience emotions.

Patients with severe depression may complain of being emotionally cut off from others and experience depersonalization and a world that seems strange to them (derealization). The impact of the loss of emotional resonance can be so pervasive that patients may denounce values and beliefs that had previously given meaning to their lives. For instance, members of the clergy might present with the complaint that they no longer believe in the Church and that they have lost God. The inability of the person with depressive disorder to experience normal emotions (commonly observed among young depressed patients) differs from the flat affect of patient with schizophrenia, in that the loss of emotions is itself experienced as painful; that is, the patient suffers immensely from the inability to experience emotions.

Table 13.4–2.
DSM-5 Criteria for Major Depressive Disorder

A. Five (or more) of the following symptoms have been present during the same 2-week period and represent a change from previous functioning: at least one of the symptoms is either (1) depressed mood or (2) loss of interest or pleasure.
Note: Do not include symptoms that are clearly attributable to another medical condition.

1. Depressed mood most of the day, nearly every day, as indicated by either subjective report (e.g., feels sad, empty, and hopeless) or observation made by others (e.g., appears tearful). (**Note:** In children and adolescents, can be irritable mood.)
2. Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation).
3. Significant weight loss when not dieting or weight gain (e.g., a change or more than 5% of body weight in a month), or decrease or increase in appetite nearly every day. (**Note:** In children, consider failure to make expected weight gain.)
4. Insomnia or hypersomnia nearly every day.
5. Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings or restlessness or being slowed down).
6. Fatigue or loss of energy nearly every day.
7. Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).
8. Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others).
9. Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.

B. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.

C. The episode is not attributable to the physiological effects of a substance or to another medical condition.

Note: Criteria A–C represent a major depressive episode (MDE).

Note: Responses to a significant loss (e.g., bereavement, financial ruin, losses from a natural disaster, a serious medical illness, or disability) may include the feelings of intense sadness, rumination about the loss, insomnia, poor appetite, and weight loss noted in Criterion A, which may resemble a depressive episode. Although such symptoms may be understandable or considered appropriate to the loss, the presence of an MDE in addition to the normal response to a significant loss should also be considered. This decision inevitably requires the exercise of clinical judgment based on the individual's history and the cultural norms for the expression of distress in the context of loss.[1]

D. The occurrence of the MDE is not better explained by schizoaffective disorder, schizophrenia, schizophreniform disorder, delusional disorder, or other specified and unspecified schizophrenia spectrum and other psychotic disorders.

E. There has never been a manic episode or a hypomanic episode.

Note: This exclusion does not apply if all of the manic-like or hypomanic-like episodes are substance-induced or are attributable to the physiological effects of another medical condition.

[1]In distinguishing grief from an MDE, it is useful to consider that in grief the predominant affect is feelings of emptiness and loss, while in MDE it is persistent depressed mood and the inability to anticipate happiness or pleasure. The dysphoria in grief is likely to decrease an intensity over days to weeks and occurs in waves, the so-called pangs of grief. These waves tend to be associated with thoughts or reminders of the deceased. The depressed mood of MDE is more persistent and not tied to specific thoughts or preoccupations. The pain of grief may be accompanied by positive emotions and humor that are uncharacteristic of the pervasive unhappiness and misery characteristic of MDE. The thought content associated with grief generally features a preoccupation with thoughts and memories of the deceased, rather than the self-critical or pessimistic ruminations seen in MDE. In grief, self-esteem is generally preserved, whereas in MDE feelings of worthlessness and self-loathing are common. If self-derogatory ideation is present in grief, it typically involves perceived failings vis-à-vis the deceased (e.g., not visiting frequently enough, not telling the deceased how much he or she was loved). If a bereaved individual thinks about death and dying, such thoughts are general focused on the deceased and possibly "joining" the deceased, whereas in MDE, such thoughts are focused on ending one's own life because of feeling worthless, undeserving of life, or unable to cope with the pain of depression.

Coding and Recording Procedures

The diagnostic code for major depressive disorder is based on whether this is a single or recurrent episode, current severity, presence of psychotic features, and remission status. Current severity and psychotic features are only indicated if full criteria are not currently met for a MDE. Remission specifiers are only indicated if the full criteria are not currently met for a MDE. Codes are as follows:

Severity/Course Specifier	Single Episode	Recurrent Episode[a]
Mild (p. 188)	296.21 (F32.0)	296.31 (F33.0)
Moderate (p. 188)	296.22 (F32.1)	296.32 (F33.1)
Severe (p. 188)	296.23 (F32.2)	296.33 (F33.2)
With psychotic features[b] (p. 186)	296.24 (F32.3)	296.34 (F33.3)
In partial remission (p. 188)	296.25 (F32.4)	296.35 (F33.41)
In full remission (p. 188)	296.26 (F32.5)	296.36 (F33.42)
Unspecified	296.20 (F32.9)	296.30 (F33.9)

[a]For an episode to be considered recurrent, there must be an interval of at least 2 consecutive months between separate episodes in which criteria are not met for a major depressive episode. The definitions of specifiers are found on the indicated pages.

[b]If psychotic features are present, code the "with psychotic features" specifier irrespective of episode severity.

In recording the name of a diagnosis, terms should be listed in the following order: major depressive disorder, single or recurrent episode, severity/psychotic/remission specifiers, followed by as many of the following specifiers without codes that apply to the current episode.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Psychomotor Disturbances. In depression, psychomotor changes consist of abnormalities in the motor expression of mental and emotional activity. In severe cases, these changes manifest in specific facial features (Fig. 13.4–2).

PSYCHOMOTOR AGITATION. Although agitation (pressured speech, restlessness, hand wringing, and hair pulling) is the more readily observed abnormality, it appears to be less specific to the illness than retardation (slowing of psychomotor activity). Psychophysiological studies have documented that such slowing often coexists with agitation.

PSYCHOMOTOR RETARDATION. Underlying many of the deficits seen in clinical depression, some authorities believe psychomotor retardation to be the core, or primary, pathology in mood disorders. Morbid depression—what patients describe as being "down"—can be understood in terms of moderate-to-extreme psychomotor slowing. The patient experiences inertia, being unable to act physically and mentally. Recent brain imaging research that has revealed subcortical (extrapyramidal system) disturbances in mood disorders tends to support the centrality of psychomotor dysfunction in these disorders.

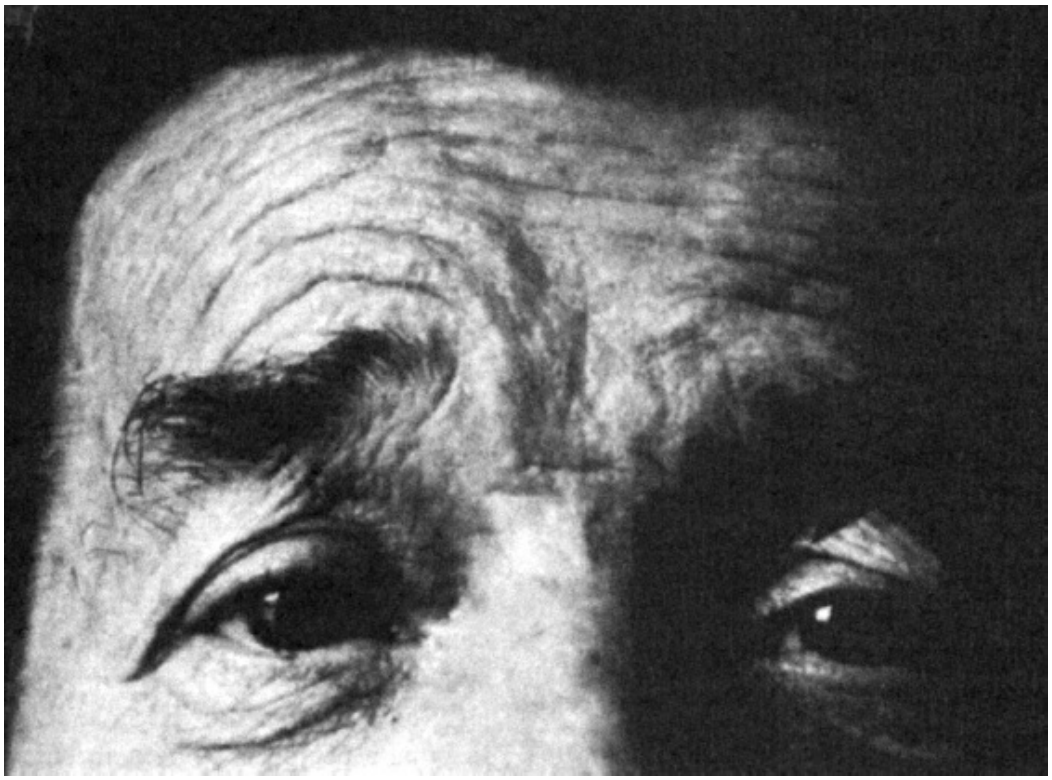

FIGURE 13.4–2. The Swiss neuropsychiatrist Otto Veraguth described a peculiar triangle-shaped fold in the nasal corner of the upper eyelid. The fold, often associated with depression, is referred to as *Veraguth's fold.* The photograph illustrates this physiognomic feature in a 50-year-old man during a major depressive episode. Veraguth's fold may also be seen in persons who are not clinically depressed, usually while they are harboring a mild depressive affect. Distinct changes in the tone of the corrugator and zygomatic facial muscles accompany depression, as shown in electromyograms. (Courtesy of Heinz E. Lehmann, M.D.)

Long neglected in psychopathological research, psychomotor retardation can be measured with precision. The Salpêtrière Retardation Scale developed by Daniel Widlocher and colleagues places special emphasis on the following disturbances: (1) paucity of spontaneous movements; (2) slumped posture with downcast gaze (Fig. 13.4–3); (3) overwhelming fatigue (patients complain that everything is an effort); (4) reduced flow and amplitude of speech and increased latency of responses, often giving rise to monosyllabic speech; (5) a subjective feeling that time is passing slowly or has stopped; (6) poor concentration and forgetfulness; (7) painful rumination or thinking that dwells on a few (usually unpleasant) topics; and (8) indecisiveness or an inability to make simple decisions.

DSM-5 places greater emphasis on the more easily observable objective or physical aspects of retardation. For the patient, however, the subjective sense of slowing is as pervasive and disabling. This more psychological dimension of retardation is most reliably elicited from depressed persons with good verbal skills.

Ms. A, a 34-year-old literature professor, presented to a mood clinic with the following

complaint: "I am in a daze, confused, disoriented, staring. My thoughts do not flow, my mind is arrested. . . . I seem to lack any sense of direction, purpose. . . . I have such an inertia, I cannot assert myself. I cannot fight; I have no will."

Less linguistically sophisticated patients would simply complain of an inability to perform household chores or difficulty in concentrating on their studies. Such psychomotor deficits, in turn, underlie depressed patients' diminished efficiency or their inability to work.

PSEUDODEMENTIA. The slowing of mental functions can be so pronounced in elderly persons that they experience memory difficulties, disorientation, and confusion.

STUPOR. Psychomotor slowing in young persons is sometimes so extreme that patients might slide into a stupor, unable to participate even in basic biological functions, such as feeding themselves. Such an episode is often the precursor of bipolar disorder, which later declares itself in a manic episode. Today, depressive disorder is diagnosed in its earlier stages, and subtle stupor is much more likely to be encountered clinically.

A B

Figure 13.4–3. A 38-year-old woman during a state of deep retarded depression **(A)** and 2 months later, after recovery **(B).** Note the turned-down corners of her mouth, her stooped posture, her drab clothing, and her hairdo during the depressed episode. (Courtesy of Heinz E. Lehmann, M.D.)

A 20-year-old male college student seen in the emergency room spoke of "being stuck—as if I have fallen into a black hole and can't get out." Further evaluation revealed that the patient was metaphorically describing his total loss of initiative and drive and was engulfed by the disease process. A clinician without the requisite phenomenological training might consider such a patient bizarre and perhaps even psychotic. Yet, the patient responded dramatically to

fluoxetine (Prozac) and, in 2 weeks, was back in school.

Cognitive Disturbances. The cognitive view of depression considers negative evaluations of the self, the world, and the future (the negative triad) central to understanding depressed mood and behavior, but it is equally likely that the depressed mood colors perceptions of the self and others or that disturbed psychomotor activity leads to negative self-evaluations. Therefore, instead of being considered causal, the cognitive triad in depression is best approached empirically as a psychopathological manifestation of depression. Faulty thinking patterns are clinically expressed as (1) ideas of deprivation and loss; (2) low self-esteem and self-confidence; (3) self-reproach and pathological guilt; (4) helplessness, hopelessness, and pessimism; and (5) recurrent thoughts of death and suicide.

The essential characteristic of depressive thinking is that the patient views everything in an extremely negative light. The self-accusations are typically unjustified or are blown out of proportion, as in the case of a middle-aged woman who was tormented by guilt because, as a child, she had not repaid a nickel that she had borrowed from a classmate. Some of the thoughts may verge on the delusional. For instance, an internationally renowned scientist complained that he was "nothing." Self-evaluations that indicate an extremely low image of one's self might, nonetheless, reflect an accurate perception of one's impairment from psychomotor retardation.

MOOD-CONGRUENT PSYCHOTIC FEATURES. In depressive disorders with psychotic features, negative thinking acquires grossly delusional proportions and is maintained with such conviction that the thoughts are not amenable to change by evidence to the contrary. According to Kurt Schneider, delusional thinking in depression derives from humankind's four basic insecurities—those regarding health, financial status, moral worth, and relationship to others. Thus, severely depressed patients may have delusions of worthlessness and sinfulness, reference, and persecution: They believe that they are being singled out for their past mistakes and that everyone is aware of their incompetence. Persecutory ideation in depression is often *prosecutory,* in that it derives from the belief that one deserves punishment for such transgressions. A severely depressed man may feel so incompetent in all areas of functioning, including the sexual sphere, that he may suspect his wife of having an affair (delusion of infidelity).

Other depressed persons believe that they have mismanaged their finances in such a way that their children will starve (delusions of poverty); that they are harboring an occult illness, such as cancer or the acquired immune deficiency syndrome (AIDS) (delusions of ill health); or that parts of their bodies are missing (nihilistic delusions). In more severe illness, the

patient might feel that the world has changed and that calamity and destruction await everyone. In rare tragic instances, a parent with such delusions might kill his or her young children to save them from moral or physical decay and then commit suicide. In women with psychotic depression, infanticide is most likely to occur in the postpartum period, often leading to all types of inappropriate interpretations in the media. Finally, a minority of depressed persons have fleeting auditory or visual hallucinations with extremely unpleasant content along the lines of their delusions (e.g., hearing accusatory voices or seeing themselves in coffins or graveyards). All of these psychotic experiences are genuine affective delusions or hallucinations. They are mood congruent in the sense that they are phenomenologically understandable in light of the prevailing pathological mood.

MOOD-INCONGRUENT PSYCHOTIC FEATURES. Sometimes so-called first-rank or schneiderian-type symptoms can arise in the setting of a major depressive episode.

A 42-year-old civil servant said that she was so paralyzed by depression that she felt that she had no personal initiative and volition left; she believed that some malignant force had taken over her actions and that it was commenting on every action that she was undertaking. The patient recovered fully with thymoleptic medication. There is no reason to believe that, in this patient, the feelings of somatic passivity and running commentary indicated a schizophrenic process.

Thus, with proper phenomenological probing, certain classes of apparently mood-incongruent psychotic experiences can be understood as arising from the pathological mood and the profound changes in psychomotor activity that accompany them. (In other instances, the clinician must seek a history of alcohol or substance use disorder or withdrawal as a putative explanation for mood incongruence in psychotic depression.) In brief, incidental schneiderian first-rank symptoms should not distract from the diagnosis of an affective psychosis if otherwise typical signs and symptoms of mood disorder are present.

HOPELESSNESS AND SUICIDE. Given that most, if not all, clinically depressed patients find themselves locked in the private hell of their negative thoughts, it is not surprising that many untreated or inadequately treated patients give up hope of ever recovering and kill themselves. The suicide attempt is not, however, undertaken in the depth of melancholia. When asked if she had any suicide plans, a severely depressed patient replied, “Doctor, I don’t exist—I am already dead.”

Thus, the risk of suicide is less pronounced during acute severe depression. Emil Kraepelin observed that it is when psychomotor activity is improving, yet mood and thinking are still dark, that the patient is most likely to muster the requisite energy to commit the suicidal act. Aaron

Beck's work has shown that hopelessness on mental status evaluation in a patient recovering from depression should alert the clinician to the possibility of such an outcome. The discovery of a therapeutic modality that could achieve improvement in psychomotor, mood, and cognitive components in tandem will constitute a major clinical advance.

There is no basis for the common belief that inquiring about suicide provokes such behavior. On the contrary, patients are often relieved that the physician appreciates the magnitude of their suffering. Suicidal ideation is commonly expressed indirectly (e.g., in a wish not to wake up or to die from a malignant disease). Some depressed persons are tormented with suicidal obsessions and are constantly resisting unwanted urges or impulses to destroy themselves. Others might yield to such urges passively (e.g., by careless driving or by walking into high-speed traffic). A third group harbors elaborate plans, carefully preparing a will and taking out insurance. Deliberate planning indicates a high suicidal risk. The foregoing examples are not exhaustive; they are meant to remind clinicians with depressed patients to be alert always to the possibility of suicide.

Vegetative Disturbances. The Greeks considered depression a somatic illness and ascribed it to black bile; hence, the term *melancholia*. The mood change in depressive disorder is accompanied by measurable alterations of biorhythms that implicate midbrain dysfunction. Once the changes occur, they tend to be independent of the environment throughout much of the episode, and, as a consequence, they do not respond to interpersonal feedback of a pleasant or upbeat nature. The biological concomitants of melancholia include profound reductions in appetite, sleep, and sexual functioning, as well as alterations in other circadian rhythms, especially matinal worsening of mood and psychomotor performance. These disturbances are central to the concept of melancholia, a form of depression in which such biological concomitants predominate. An equally prominent subgroup of depressed persons exhibits a reversal of the vegetative and circadian functions, with increases in appetite and sleep—and sometimes in sexual functioning—and an evening worsening of mood; in this atypical pattern, patients characteristically exhibit mood reactivity and sensitivity to rejection. Marked retardation might herald a manic switch, whereas atypicality should raise the suspicion of bipolar II disorder.

ANOREXIA AND WEIGHT LOSS. The most reliable somatic indicators of depressive disorder include anorexia and weight loss. In addition to being a hypothalamic-based disturbance in depression, anorexia might be secondary to blunted olfactory or taste sensations or a decreased enjoyment of food, or (rarely) it might result from a delusional belief that the food has been poisoned.

Inanition, especially in elderly persons, can lead to malnutrition and

electrolyte disturbances that represent medical emergencies in their own right. If weight loss is severe, especially after 40 years of age, the psychiatrist should first use appropriate medical consultation to rule out the likelihood of an occult malignancy.

WEIGHT GAIN. Overeating, decreased activity, or both may result in weight gain. In middle-aged patients, it may aggravate pre-existing diabetes mellitus, hypertension, or coronary artery disease. In younger patients, especially women, weight problems may conform to a bulimic pattern that is often the expression of the depressive phase of a bipolar disorder with infrequent hypomanic periods (bipolar II disorder).

INSOMNIA. Sleep disturbance, a cardinal sign of depression, often is characterized by multiple awakenings, especially in the early hours of the morning, rather than by difficulty falling asleep. The light sleep of a depressed person, in part a reflection of the painful arousal of the disorder, tends to prolong the depressive agony over 24 hours. Thus, deep stages of sleep (III and IV) are decreased or deficient. The attempt to overcome the problem by drinking alcohol may initially succeed but ultimately aggravates the sleep patterns and insomnia. This is also true for sedative–hypnotic agents, which are often prescribed by the busy general practitioner who has not spent enough time diagnosing the depressive condition. Although sedatives (including alcohol) effectively reduce the number of awakenings in the short term, they are not effective in the long run because they further diminish stage III and stage IV sleep. They are not antidepressants, and they tend to prolong the depression.

HYPERSOMNIA. Young depressed patients, especially those with bipolar tendencies, often exhibit excessive sleep and have difficulty getting up in the morning.

Kevin, a 15-year-old boy, was referred to a sleep center to rule out narcolepsy. His main complaints were fatigue, boredom, and a need to sleep all the time. Although he had always started the day somewhat slowly, he now could not get out of bed to go to school. That alarmed his mother, prompting sleep consultation. Formerly a B student, he had been failing most of his courses in the 6 months before referral. Psychological counseling, predicated on the premise that his family's recent move from another city had led to Kevin's isolation, had not been beneficial. Extensive neurological and general medical workup had also proven negative. He slept 12 to 15 hours per day but denied cataplexy, sleep paralysis, and hypnagogic hallucinations. During psychiatric interview, he denied being depressed but admitted that he had lost interest in everything except his dog. He had no drive, participated in no activities, and had gained 30 pounds in 6 months. He believed that he was "brain damaged" and wondered whether it was worth living like that. The question of suicide disturbed him, as it was contrary to his religious beliefs. These findings led to the prescription of desipramine (Norpramin) in a dose that was gradually increased to 200 mg per day over 3 weeks. Not only did desipramine reverse the presenting complaints, but it also pushed him to the brink of a manic episode.

The affective nature of the disorder in such patients is often

unrecognized, and their behavior is attributed to "laziness." The vignette also illustrates the emergence of manic behavior during antidepressant treatment. Such shifts in polarity are common in major depressive disorder and necessitate revising the diagnosis to a bipolar disorder. Hypersomnia associated with lethargy should actually raise strong suspicion of bipolar disorder in young depressive patients. In the elderly, organic or brain pathology should be suspected.

CIRCADIAN DYSREGULATION. Many circadian functions, such as temperature regulation and cortisol rhythms, are disrupted in major depressive disorder. Disturbances of sleep rhythms, however, have received the greatest research focus. These include deficits in stage IV or delta sleep, as well as more intense REM activity in the first one-third of the night. More specific to depressive disorders—whether suffering from insomnia or hypersomnia—nearly two-thirds of patients exhibit a marked shortening of REM latency, the period from the onset of sleep to the first REM period. This abnormality is observed throughout the depressive episode and may also be seen during relatively euthymic periods in persons with recurrent depression. The occurrence of short REM latency in the younger, clinically well relatives of the affectively ill suggests that neurophysiological abnormalities might precede the overt psychopathological manifestations of the illness; on closer scrutiny, these "well" relatives are often found to meet criteria for subthreshold mood conditions, such as dysthymic disorder, intermittent depression, or labile temperament.

Few data exist on the consistency of sleep electroencephalographic (EEG) abnormalities in patients from episode to episode. However, clinical experience suggests that a patient observed over time (even during the same episode) may exhibit insomnia and morning worsening of mood and activity during one period of the disorder and hypersomnia extending to late morning hours during another period. In either case, persons with depressive disorder are characteristically tired in the morning, which means that even prolonged sleep is not refreshing for them. The propensity to exhibit such divergent patterns of sleep disturbance is more likely in bipolar disorders. Patients with major depressive disorder tend to exhibit insomnia more stereotypically episode after episode; despite extreme fatigue, they rarely oversleep. Such fatigue coexisting with negative affective arousal is even more exhausting.

SEASONALITY. Another classic biorhythmic disturbance in mood disorders is seasonal (especially autumn–winter) accentuation or precipitation of depression. Most of those patients experience increased energy and activation, if not frank hypomania, in the spring. In the fall and winter, they complain of fatigue, tend to crave sugars, and overeat and oversleep. The hypersomnia in some of these patients is associated with delayed (rather

than short) REM latencies. These data suggest dysregulation of circadian rhythms in depressive disorders rather than mere phase advance. Given the biphasic nature of the clinical phenomenology, some autumn–winter depressions belong to the bipolar spectrum (type II). Although autumn–winter depression has received the greatest attention, there also exist summer depressions; the former appear to be related to reduction in daylight (photoperiods), and the latter appear to be related to increased temperature.

SEXUAL DYSFUNCTION. Decreased sexual desire is seen in depressed men and women. In addition, some women experience temporary interruption of their menses. Depressed women are often unresponsive to lovemaking or are disinclined to participate in it, a situation that could lead to marital conflict. Psychotherapists might mistakenly ascribe the depression to the marital conflict and might devote unnecessarily zealous psychotherapeutic attention to conjugal issues. Decreased or lost libido in men often results in erectile failure, which may prompt endocrinological or urological consultation. Again, depression may be ascribed to the sexual dysfunction rather than the reverse, and definitive treatment may be delayed by the focus on the sexual complaint. In the past, some men with depressive disorder were subjected to permanent penile implants before receiving more definitive treatment of their depression. This is less likely to occur in the sildenafil (Viagra) era, but even treatment with such agents would not necessarily resolve erectile disorder in clinically depressed patients without competent treatment of the mood disorder. SSRIs typically aggravate the sexual dysfunction, whereas bupropion (Wellbutrin) tends to improve it.

A small subgroup of persons with depressive disorder may exhibit increased sexual drive or activity of a compulsive nature. These patients tend to have other atypical features as well; hence, the increased sexual drive can be considered the *fifth reverse vegetative sign* (after evening worsening of mood, initial insomnia, hypersomnia, and weight gain). Several patients have derived temporary reversal of their depression after intense sexual encounters. In these depressed persons, increased sexual drive, intensity, or both may indicate a mixed episode of bipolar disorder (type II). Further scrutiny in such cases often reveals a premorbid cyclothymic or hyperthymic temperament. Current data suggest that, depending on the breadth of the criteria used, 30 to 70 percent of depressions with atypical features belong to the bipolar spectrum.

Manic Syndrome

As with clinical depression, the psychopathology of mania (Table 13.4–3) can be conveniently discussed under mood, psychomotor, circadian, and cognitive disturbances. The clinical features of mania are generally the opposite of those of depression. Thus, instead of lowered mood, thinking,

activity, and self-esteem, there is elevated mood, a rush of ideas, psychomotor acceleration, and grandiosity. Despite those contrasts, the two disorders share such symptoms as irritability, anger, insomnia, and agitation. Actually, an excess of the latter symptoms of escalating intensity suggests a mixed phase or mixed episode of mania and depression occurring simultaneously. Manic and mixed episodes represent the hallmark of what was once termed *manic-depressive psychosis* and is currently termed *bipolar I disorder*.

Table 13.4–3.
Manic Episode

A. A distinct period of abnormally and persistently elevated, expansive, or irritable mood and abnormally and persistently increased goal-directed activity or energy, lasting at least 1 week and present most of the day, nearly every day (or any duration if hospitalization is necessary).
B. During the period mood disturbance and increased energy or activity, three (or more) of the following symptoms (four if the mood is only irritable) are present to a significant degree and represent a noticeable change from usual behavior:
   1. Inflated self-esteem or grandiosity.
   2. Decreased need for sleep (e.g., feels rested after only 3 hours of sleep).
   3. More talkative than usual or pressure to keep talking.
   4. Flight of ideas of subjective experience that thoughts are racing.
   5. Distractibility (i.e., attention to easily drawn to unimportant or irrelevant external stimuli), as reported or observed.
   6. Increase in goal-directed activity (either socially, at work or school, or sexually) or psychomotor agitation (i.e., purposeless non-goal-directed activity.)
   7. Excessive involvement in activities that have a high potential for painful consequences (e.g., engaging in unrestrained buying sprees sexual indiscretions, or foolish business investments).
   8. The mood disturbance is sufficiently severe to cause marked impairment in social or occupational functioning or to necessitate hospitalization to prevent harm to self or others, or there are psychotic features.
   9. The episode is not attributable to the physiological effects of a substance (e.g., a drug of abuse, a medication, other treatment) or to another medical condition.

Note: A full manic episode that emerges during antidepressant treatment (e.g., medication and electroconvulsive therapy) but persists at a fully syndromal level beyond the physiological effect of that treatment is sufficient for a manic episode and therefore, a bipolar I diagnosis.
Note: Criteria A–D constitute a manic episode. At least one lifetime manic episode is required for the diagnosis of bipolar I disorder.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Although milder or hypomanic features (Table 13.4–4) can contribute to success in business, leadership roles, and the arts, recurrences of even mild manic symptomatology are typically disruptive. The elated mood tends to produce overoptimism concerning one's abilities, which, coupled with the impulsivity characteristic of mania, often leads to disaster. Thus, accurate and early diagnosis is paramount.

Classic mania, as formulated in the DSM-5, operationalism of manic

episode is relatively easy to recognize. Misdiagnosis was once rampant in North American practice as clinicians confused severe mania with schizophrenia and confused its milder variants with normality or with narcissistic and sociopathic personality disorders. Like the misdiagnosis of depressive conditions, such errors of clinical judgment are due to a lack of familiarity with the phenomenology of the classic illness. Again, DSM-5 criteria provide only a guideline. The actual diagnosis requires careful history and phenomenological understanding by an empathic observer. The manic patient lifts the observer's mood and makes the examiner smile and even laugh but can also be irritating. The patient's speech is fast and may even appear to the novice psychiatry student as "loose," but it also can often be witty. Finally, the behavior is typically dramatic, expansive, and jesting. Current research indicates that social disinhibition and pathological overfamiliarity with strangers represent cardinal features of mania and can be contrasted with the withdrawal of most schizophrenic patients. For the experienced clinician, the overall Gestalt experienced in the presence of manic patients is emotionally and qualitatively distinct from that of persons with schizophrenia or frontal lobe diseases; the latter conditions tend to leave the examiner cold. These considerations become clearer when the clinical observer systematically examines the psychopathology of mania in the areas of mood, behavior, and thinking.

Table 13.4–4.
Hypomanic Episode

A. A distinct period of abnormally and persistently elevated, expansive, or irritable mood and abnormally and persistently increased activity or energy, lasting at least 4 consecutive days and present most of the day, nearly every day.
B. During the period of mood disturbance and increased energy and activity, three (or more) of the following symptoms (four if the mood is only irritable) have persisted, represent a noticeable change from usual behavior, and have been present to a significant degree:
   1. Inflated self-esteem or grandiosity.
   2. Decreased need for sleep (e.g., feels rested after only 3 hours of sleep).
   3. More talkative than usual or pressure to keep talking.
   4. Flight of ideas or subjective experience that thoughts are racing.
   5. Distractibility (i.e., attention to easily drawn to unimportant or irrelevant external stimuli), as reported or observed.
   6. Increase in goal-directed activity (either socially, at work or school, or sexually) or psychomotor agitation.
   7. Excessive involvement in activities that have a high potential for painful consequences (e.g., engaging in unrestrained buying sprees, sexual indiscretions, or foolish business investments).
C. The episode is associated with an unequivocal change in functioning that is uncharacteristic of the individual when not symptomatic.
D. The disturbance in mood the change in functioning is observable by other.
E. The episode is not severe enough to cause marked impairment in social or occupational functioning or to necessitate hospitalization. If there are psychotic features, the episode is, by definition, manic.
F. The episode is not attributable to the physiological effects of a substance (e.g., a drug of abuse, a medication, other treatment).

Note: A fully hypomanic episode that emerges during antidepressant treatment (e.g., medication and electroconvulsive therapy) but persists at a fully syndromal level beyond the physiological effect of the treatment is sufficient evidence for hypomanic episode diagnosis. However, caution is indicated so that one or two symptoms (particularly increased irritability, edginess, or agitation following antidepressant use) are not taken as sufficient for diagnosis of a hypomanic episode, nor necessarily indicative of a bipolar diathesis.

Note: Criteria A–F constitute a hypomanic episode. Hypomanic episodes are common in bipolar I disorder but are not required for the diagnosis of bipolar I disorder.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Mood Disturbance. Mood disturbance in mania represents a contrast to that observed in depression, but not entirely.

MOOD ELEVATION. The mood in mania is classically one of elation, euphoria, and jubilation, typically associated with laughing, punning, and gesturing.

LABILITY AND IRRITABILITY. The prevailing positive mood in mania is not stable, and momentary crying or bursting into tears is common. In addition, the high is so excessive that many patients experience it as intense nervousness. When crossed, patients can become extremely irritable and hostile. Thus, lability and irritable hostility are as much features of the manic mood as is elation. In mixed manic states, they dominate the clinical picture, giving rise to what is now termed *dysphoric mania* (and what Kraepelin characterized as *anxious-depressive mania*).

Psychomotor Acceleration. Accelerated psychomotor activity, the hallmark of mania, is characterized by overabundant energy and activity and rapid, pressured speech. Subjectively, the patient experiences an unusual sense of physical well-being (eutonia).

FLIGHT OF IDEAS. Thinking processes are accelerated, subjectively experienced as flight of ideas, and thinking and perception are unusually sharp. The patient may speak with such pressure that associations are difficult to follow; such "clang" associations are often based on rhyming or chance perceptions and can be lightning fast. The pressure to speak may continue despite the development of hoarseness.

IMPULSIVE BEHAVIOR. Manic patients are typically impulsive, disinhibited, and meddlesome. Pathological familiarity with total strangers is also a feature not specifically listed in the DSM-5 schema for mania, yet it is one of its cardinal signs. They are intrusive in their increased involvement with others, leading to friction with family members, friends, and colleagues. They are distractible and move quickly, not only from one thought to another but also from one person to another, showing heightened interest

in every new activity that strikes their fancy. They are indefatigable and engage in various activities in which they usually display poor social judgment. Examples include preaching or dancing in the street; abusing the use of long-distance calling; buying new cars, hundreds of records, expensive jewelry, or other unnecessary items; paying the bills of total strangers in bars; giving away furniture; marrying impulsively; engaging in risky business ventures; gambling; and taking sudden trips. Such pursuits can lead to personal and financial ruin.

DELIRIOUS MANIA. Although not listed in DSM-5, an extremely severe expression of mania (once known as *Bell's mania*) called delirious mania involves frenzied physical activity that continues unabated and leads to delirium and disorientation—a life-threatening medical emergency. This complication, the manic counterpart of stupor, is rare today.

WITH CATATONIC FEATURES. As a symptom, catatonia can be present in several mental disorders. The presence of catatonic features in patients with mood disorders may have prognostic and treatment significance. The hallmark of symptoms of catatonia—stuporousness, blunted affect, extreme withdrawal, negativism, and marked psychomotor retardation—can be seen in both catatonic and noncatatonic schizophrenia, major depressive disorder (often with psychotic features), and medical and neurological disorders. Clinicians often do not associate catatonic symptoms with bipolar I disorder because of the marked contrast between these symptoms of stuporous catatonia and the classic symptoms of mania.

VEGETATIVE DISTURBANCES. Vegetative disturbances are more difficult to evaluate in mania than in depression.

Hyposomnia. Hyposomnia is the cardinal sign of decreased need for sleep—the patient sleeps only a few hours but feels energetic on awakening. Some patients may actually go sleepless for several days. This could lead to dangerous escalation of manic activity, which might continue despite signs of physical exhaustion.

Inattention to Nutrition. There does not seem to be a clinically significant level of appetite disturbance as such, but weight loss may occur because of increased activity and neglect of nutritional needs.

Sexual Excesses. Hypersexuality is a cardinal sign of mania. The sexual appetite is typically increased and may lead to sexual indiscretion. Married women with previously unblemished sexual lives may associate with men below their social status. Men typically overindulge in alcohol, frequent bars, and may squander their savings on prostitutes. The sexual misadventures of manic patients result in marital disasters—hence the multiple separations or divorces that are almost pathognomonic of the

disorder. The sexual excesses of bipolar patients are even more problematic today, in view of the specter of AIDS.

Cognitive Distortions. Manic thinking is overly positive, optimistic, and expansive.

GRANDIOSITY, LACK OF INSIGHT, AND POOR JUDGMENT. The patient exhibits inflated self-esteem and a grandiose sense of confidence and achievements. Behind that facade, however, may be a vague and painful recognition that the positive self-concepts do not represent reality. However, such insight (if present at all) is transient, and manic patients are notoriously refractory to self-examination and insight. Denial and lack of insight are cardinal psychological derangements of mania. It is this lack of insight—coupled with poor judgment—that leads manic patients to engage in activities that harm themselves and their loved ones. It also explains, in part, their nonadherence with medication regimens during the manic phase.

DELUSION FORMATION. Manic patients often harbor delusional beliefs, including delusions of exceptional mental and physical fitness and talent; delusions of wealth, aristocratic ancestry, or other grandiose identity; delusions of assistance (i.e., well-placed persons or supernatural powers are assisting their endeavors); or delusions of reference and persecution, based on the belief that enemies are observing or following them out of jealousy at their special abilities. At the height of mania, patients may even see visions or hear voices congruent with their euphoric mood and grandiose self-image (e.g., they might see images of heaven or hear cherubs chanting songs to praise them). The denial characteristic of mania—and the frequently psychotic nature of episodes—means that clinicians must routinely obtain diagnostic information about past episodes from significant others. Lack of insight also unfortunately means that hospitalization must often be arranged on an involuntary basis.

MOOD-INCONGRUENT PSYCHOSIS. Psychosis in the setting of mania and mixed manic episodes is typically mood congruent. The sense of physical well-being and mental alacrity is so extraordinary that it is understandable why manic patients believe that they possess superior powers or perhaps are great scientists or famous reformers. Moreover, their senses are so vivid that reality appears richer and more exotic and can be easily transformed into a vision.

Likewise, their thoughts are so rapid and vibrant that they feel that they can hear them. Thus, certain first-rank schneiderian-type symptoms that have been traditionally considered mood incongruent can be understood phenomenologically to arise from the powerful mental experiences of mania.

A 37-year-old engineer had experienced three manic episodes for which he had been hospitalized; all three episodes were preceded by several weeks of moderate psychomotor retardation. Although he had responded to lithium (Eskalith) each time, once outside the hospital, he had been reluctant to take it and eventually refused to do so. Now that he was "euthymic," after his third and most disruptive episode during which he had badly beaten his wife, he could more accurately explain how he felt when manic. He experienced mania as "God implanted in him," so he could serve as "testimony to man's communication with God." He elaborated as follows: "Ordinary mortals will never, never understand the supreme manic state which I'm privileged to experience every few years. It is so vivid, so intense, so compelling. When I feel that way, there can be no other explanation: To be manic is, ultimately, to be God. God himself must be supermanic: I can feel it when mania enters through my left brain like laser beams, transforming my sluggish thoughts, recharging them, galvanizing them. My thoughts acquire such momentum, they rush out of my head, to disseminate knowledge about the true nature of mania to psychiatrists and all other ordinary mortals. That's why I will never accept lithium again—to do so is to obstruct the divinity in me." Although he was on the brink of divorce, he would not yield to his wife's plea to go back on lithium.

The vignette illustrates the possibility that even some of the most psychotic manifestations of mania represent explanatory delusions, the patient's attempt to make sense of the experience of mania. Many manic patients abuse alcohol and stimulants to enhance their mental state; mood incongruence can sometimes be explained on that basis

HYPOMANIA VERSUS MANIA. Nonpsychotic and nondisruptive variants of mania are much more common and are recognized by DSM-5 as hypomanic episodes. They are the historical clinical marker for bipolar II. Diagnostically, history of hypomania is preferably obtained from significant others who have observed the patient; the experience is often pleasant, and the subject may be unaware of it or may tend to deny it. Others experience irritable or dysphoric hypomanic episodes, which may be difficult to recognize clinically because of a stereotype that emphasizes the positive aspects of hypomania. Some patients feel that they benefit from the energy and confidence of hypomania. This is true if hypomania is mild and "sunny." But, many patients, especially those with "darker" irritable activation, experience impairment in occupational and interpersonal functioning over time. DSM-5 stipulates a minimum duration of 4 days for hypomania; however, current evidence indicates that bipolar II disorders with long (4 or more days) and short (2 to 3 days) hypomania are indistinguishable on the basis of bipolar family history. It is, therefore, recommended that the threshold for detecting the duration of hypomania be set at 2 days. Ultimately, the duration of hypomanic experiences might be less important than the fact that they recur. In other words, the occurrence of brief recurrent hypomanias, even if their duration is 1 day, interspaced with major depressive episodes can be taken as presumptive evidence for bipolar II. Finally, although DSM-5 states that treatment-emergent hypomania in a depressed patient does not count toward a diagnosis of bipolarity, prospective observations show that nearly all such episodes are followed eventually by spontaneous hypomania (or mania);

moreover, family history for bipolar disorder is comparable in patients with spontaneous and antidepressant-associated hypomania. In brief, antidepressant-associated hypomanias are best considered as a genetically less penetrant variant of bipolarity (which, in the literature, are often referred to as *bipolar III disorders*).

DIAGNOSTIC CLASSIFICATION

The classification of mood disorders in DSM-5 subsumes a large variety of patients seen in private, public, ambulatory, and inpatient settings. The main demarcation in that large clinical terrain is between bipolar and depressive disorders. Thus, bipolar disorders range from the classic manic and depressive episodes, often of psychotic intensity (bipolar I disorder), to recurrent major depressive episodes alternating with hypomanic episodes (bipolar II disorder) and cyclothymic mood swings. Likewise, depressive disorders include those with psychotic severity, melancholia, atypical features, and dysthymic variants.

Major and specific attenuated subtypes are distinguished on the basis of severity and duration. In dysthymic and cyclothymic disorders, a partial mood syndrome consisting of such subthreshold features as subdepressive and hypomanic periods is maintained, intermittently or continuously, for at least 2 years. Subdepressive periods dominate in dysthymia; in cyclothymia, they alternate with brief hypomania. The onset is typically in adolescence or childhood, and most persons with these diagnoses first seen clinically in young adulthood have had low-grade mood symptoms for 5 to 10 years. Major mood disorders, which generally begin much later in life, require the presence of a full manic episode or a full depressive episode—sustained for at least 1 and 2 weeks, respectively—and an episodic course, typically permitting recovery or remission from episodes. As many as one-third of persons with major depressive disorders fail to achieve full symptomatic recovery and should thus be qualified as chronic or in partial remission.

Dichotomy or Continuum?

Although, in the extreme, bipolar and depressive disorders can be discriminated clinically and therapeutically (Table 13.4–5), clinical observations testify to a vast overlap between those extremes. Thus, the distinctions among the various affective subtypes are not as hard and fast as DSM-5 attempts to portray. For instance, full-blown bipolar disorder can be superimposed on cyclothymic disorder that tends to persist after the resolution of manic or major depressive episodes. Even more common is major depressive disorder complicating cyclothymic disorder, which should be reclassified as an important course variant of bipolar II disorder. Likewise, recent evidence indicates that dysthymic disorder may precede

major depressive disorder in as many as one-third of cases. Moreover, as much as 50 percent of persons with major depressive disorder during long-term prospective follow-up develop hypomanic or manic episodes and should be reclassified as having bipolar disorder. In some, if not many, instances, apparent switching of polarity might simply be due to earlier misclassification of bipolar disorder as major depressive disorder. Finally, unexpected crossing from dysthymic disorder to hypomanic or manic episodes has also been described, suggesting that some forms of dysthymic disorder are subaffective precursors of bipolar disorder. The concept of *sub-bipolar dysthymia* (not a formal category in DSM-5) can bring clinical attention to this subgroup of patients. Such observations are in line with Kraepelin's historic attempt to bring all mood disorders under one rubric. Epidemiological studies in the community have also shown much fluidity among various subthreshold and major mood disorders.

Table 13.4–5.
Differentiating Characteristics of Bipolar and Unipolar Depressions

	Bipolar	Unipolar
History of mania or hypomania (definitional)	Yes	No
Temperament and personality	Cyclothymic and extroverted	Dysthymic and introverted
Sex ratio	Equal	More women than men
Age of onset	Teens, 20s, and 30s	30s, 40s, and 50s
Postpartum episodes	More common	Less common
Onset of episode	Often abrupt	More insidious
Number of episodes	Numerous	Fewer
Duration of episode	3–6 months	3–12 months
Psychomotor activity	Retardation > agitation	Agitation > retardation
Sleep	Hypersomnia > insomnia	Insomnia > hypersomnia
Family history		
Bipolar disorder	Yes	±
Unipolar disorder	Yes	Yes
Alcoholism	Yes	Yes
Pharmacological response		
Most antidepressants	Induce hypomania–mania	±
Lithium carbonate	Prophylaxis	±

Heterogeneity undoubtedly exists among mood disorders. However, the classic unipolar–bipolar distinction might not be the best way to capture it. The foregoing observations suggest that much of the recurrent depressive terrain might be *pseudo-unipolar* (i.e., soft bipolar). The clinical significance of these considerations lies in the fact that many DSM-5 subtypes of mood disorders are not pure entities, and considerable overlap and switches in polarity take place. They also provide some rationale, for instance, for why lithium or other mood stabilizer augmentation may be

effective in some apparently unipolar depressions; such patients do not necessarily experience brief spontaneous hypomanic episodes but instead often exhibit a high baseline level of hyperthymic traits. Finally, several studies have shown that bipolar patients with cyclothymic premorbid adjustment and interepisodic adjustment are at considerable risk of *antidepressant-induced rapid cycling,* defined as a rapid succession of major episodes with few or no intervals of freedom.

As Kraepelin illustrated in his monograph, course is best captured graphically (Fig. 13.4–4). Kraepelin, after diagramming 18 illustrative patterns for the entire spectrum of manic-depressive illness, declared that the illness pursued an indefinite number of courses. Although not represented in the official US and international classifications, some of these course patterns are of considerable interest today. For instance, a biphasic course (a sequence in which episodes of opposite polarity succeed each other and then are followed by a free interval) is relevant to treatment response. Thus, depression followed by mania or hypomania—compared to mania or hypomania followed by depression—appears, on the basis of replicated studies, to be less responsive to lithium.

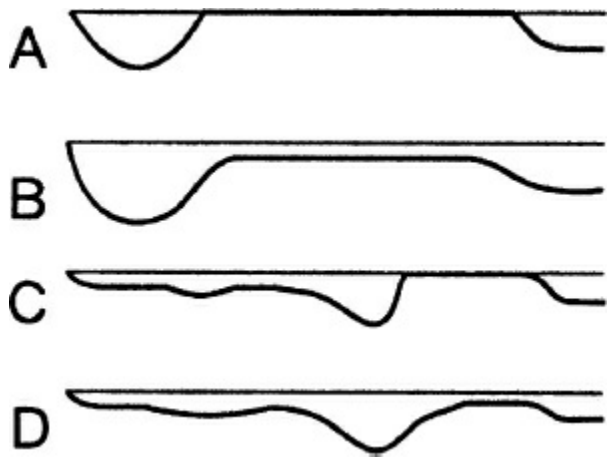

FIGURE 13.4–4. Graphs depicting prototypical courses. **A:** Course of major depressive disorder, recurrent, with no antecedent dysthymic disorder and a period of full remission between the episodes. This pattern predicts the best future prognosis. **B:** Course of major depressive disorder, recurrent, with no antecedent dysthymic disorder but with prominent symptoms persisting between the two most recent episodes (i.e., partial remission is attained). **C:** Rare pattern (present in fewer than 3 percent of persons with major depressive disorder) of major depressive disorder, recurrent with antecedent dysthymic disorder but with full interepisode recovery between the two most recent episodes. **D:** Course of major depressive disorder, recurrent, with antecedent dysthymic disorder and no period of full remission between the two most recent episodes. This pattern, commonly referred to as *double depression,* is seen in approximately 20 to 25 percent of persons with major depressive disorder. (From American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Text rev. Washington, DC: American Psychiatric Association; 2000, with permission.)

DEPRESSIVE DISORDERS

The broad category of depressive disorders includes major depressive

disorder, dysthymic disorder, and depressive disorder not otherwise specified.

Major Depressive Disorder

Episodes usually begin over a prodromal period of weeks to months. The DSM-5 diagnosis of major depressive disorder requires one of the following: (1) dysphoric mood or (2) decreased interest in usual activities. The full range of signs and symptoms are listed in Table 13.4–2. Such symptoms must be sustained for at least 2 weeks, and cannot be explained by another process known to cause depressive symptoms, such as normal bereavement, certain physical conditions commonly associated with depression, or another mental disorder. It can be a single episode or, commonly, recurrent, or both.

Comorbid Physical Disease. The foregoing considerations raise the question of whether major depressive disorder should be limited to depressions of unknown etiology (i.e., those without documented physical causes). The DSM-5 approach has basically been that, when the cause is known, the condition should be diagnosed as mood disorder due to a general medical condition that must be specified or substance-induced mood disorder The problem with this approach is that many common medical factors historically associated with depression do not seem to be causative in the etiological sense but rather are triggering agents in otherwise predisposed persons. This is analogous to the situation with life events, which no longer are used in making distinctions between reactive and endogenous subtypes of depression. A more troubling implication is that major depressive disorders without demonstrable physical disease are not medical or otherwise biological. More important, there appears to be no reliable or valid way for a clinician to decide that a depressive condition is due to a specified medical condition. In brief, the designation *due to a general medical condition* is cumbersome and redundant. Major depressive disorder should be considered to represent the final common pathway of multifactorial interacting factors—physical and psychological—a syndrome that should be diagnosed irrespective of presumed cause.

Diagnostic Threshold. Another question concerning the definition of major depressive disorder relates to the threshold at which a constellation of depressive features becomes a condition distinct from the ordinary blues. Within the current definition, if a person responds to a setback with lowered spirits and self-doubt, difficulty in sleeping and concentration, and decreased sexual interest for 14 days, he or she would qualify for the diagnosis of a major depressive disorder of mild intensity. Many clinicians would consider such a condition a relatively minor departure from normality, probably no more than an adjustment disorder. Obviously,

criteria other than signs, symptoms, and duration are necessary to differentiate a major depressive disorder from adjustment reactions to life situations. The presence of the following characteristics might assist in such a differentiation:

- By definition, a major depressive disorder should be incapacitating. Previously, much attention was paid to the interpersonal consequences of depression. Recent evidence indicates that measurable deficits in work performance are often early manifestations. Afflicted persons also do not benefit from taking leisure time, and, hence, prescribing vacations is futile.
- Major depressive disorder is usually perceived as a break from a person's usual or premorbid self, which can be so striking that patients may feel as though they are losing their minds. The important point is that the patient and significant others can usually relate the onset of the illness to a given month or quarter of a year, which is not true, for instance, for dysthymic disorder.
- Major depressive disorder is often experienced by the patient as qualitatively distinct from grief or other understandable reactions to loss or adversity. William James described it as follows:

> There is a pitch of unhappiness so great that the goods of nature may be entirely forgotten, and all sentiment of their existence vanish from the mental field. For this extremity of passion to be reached, something more is needed than adversity; the individual must in his own person become the prey of pathological melancholy. Such sensitiveness and susceptibility of mental pain is a rare occurrence where the nervous constitution is entirely normal: one seldom finds it in a healthy subject even where he is the victim of the most atrocious cruelties of outward fortune; it is an active anguish, a sort of psychical neuralgia wholly unknown to healthy life.

Two additional features, when present, would further validate the diagnosis of major depressive disorder.

- History of past episodes.
- Consecutive-generation family history of mood disorder—especially when a large number of family members are afflicted with depression or mood disorder—is characteristic of clinical depression. For instance, one study that prospectively followed persons with minor or neurotic depression found that such pedigrees predicted the development of future major episodes. In clinical practice, these factors would strongly weigh whether depression is taken seriously.

Single-Episode and Recurrent Subtypes. A significant minority—perhaps one-third—of all major depressive episodes do not recur. Such patients tend to be older and less likely to have a positive family history for mood disorders and have a more protracted (1 to 2 years) course of the disorder. Patients with single-episode major depressive disorder should be distinguished from those experiencing their first episodes of recurrent

major depressive disorder. The latter group tends to be younger, and the disorder is more likely to have been preceded by a depressive temperament or dysthymic disorder. Those who switch to bipolar disorder are more likely to have experienced recurrent depressions (5 or more episodes).

Research has established that recurrent major depressive disorders are more familial than their single-episode counterparts. The average length of episodes is 6 months, whereas the mean interval between episodes tends to vary (typically years). The mean number of major episodes over a lifetime, according to retrospective and prospective studies, is five to six, in contrast to an average of eight to nine major episodes in bipolar disorder. These figures are probably underestimates, in that they are typically ascertained on the basis of clinical referral or hospitalization, or both.

Melancholic Features. *Melancholic features* is a qualifying phrase for major depressive disorders in which anhedonia, guilt, and psychomotor–vegetative disturbances dominate the clinical picture. In addition, severe suicidal ideation may be present.

Atypical Features. Reverse vegetative signs with rejection sensitivity, often contrasted to melancholia on phenomenological and pharmacological grounds, represent a major depressive disorder qualifier occurring in as many as one-third of all major depressive disorders. Atypical features are so common in bipolar disorder, especially bipolar II disorder, that some consider them to be clinical markers for soft bipolar disorders. Although there is no consensus on this question, before diagnosing major depressive disorder in such cases, it is clinically wise to exclude bipolar II disorder, especially where hostile–labile features predominate.

Psychotic Features. From 10 to 15 percent of major depressive disorders, usually from the rank of those with melancholic features, develop into delusional depressions. In young persons, they tend to be retarded, even stuporous, and are best considered initial episodes of a bipolar disorder until proven otherwise. More typically, psychotic depression that develops for the first time after 50 years of age often presents with severe agitation, delusional guilt, hypochondriacal preoccupations, early-morning awakening, and weight loss. The premorbid adjustment of such patients is classically characterized as *obsessoid*. Their mournful-anxious mood and agitation are autonomous, being refractory to psychological interventions, and they endure great psychic suffering. Except for the fact that generally one to two episodes occur in late-onset (so-called involutional) depressions, they represent a severe variant of DSM-5 melancholia. Kraepelin's postulation of a cerebrovascular basis for such cases makes the ventricular enlargement and white matter opacities reported in psychotic depressions of some interest. Their etiological

specificity for persons with late-onset psychotic depression has been controversial, however, because younger (more bipolar) persons with psychotic depression may exhibit similar findings. Brain imaging findings tend to be correlated with the neurocognitive deficits observed in psychotic depressions. Those features do not seem to define a distinct depressive subtype but one of greater severity that some authorities now classify as *vascular depressions*. Finally, despite attempts to suggest a neurochemical uniqueness based largely on the need for antipsychotic treatment in the acute phase of many of those patients, familial and other external validations have failed to support psychotic depression as a separate entity; Emerging data, nonetheless, might eventually force a change in this convention. For instance, William Coryell and collaborators in the NIMH Collaborative Depression Study showed that psychotic depression was the most consistent unipolar subtype across episodes. Alan Schatzberg's work, originally conducted at Harvard, likewise underscored the uniqueness of psychotic depression based on neuroendocrine and putative neurochemical considerations. Finally, consideration should be given to Athanasios Koukopoulos' clinical formulation that many agitated psychotic depressions might represent mixed states (i.e., activated depressions that belong to the bipolar spectrum).

Chronic Depression. The symptom profile in chronic depressions usually displays low-grade intensity rather than severe syndromic chronicity. Severe depressive disorder, in its psychotic forms, is so agonizing that the patient is at risk of committing suicide before the disorder has a chance to become chronic. More commonly, the psychotic symptoms respond to medication or to ECT, but residual depressive symptoms may linger for a long time. In other persons with chronic depressions, the chronicity arises from more mundane (nonpsychotic) major depressive episodes, representing depressive residua following one or several clinical episodes that fail to remit fully. Instead of the customary remission within 1 year, the patients are ill for years. The level of depression varies, fluctuating between syndromic illness and milder symptoms. Recent landmark analyses from the NIMH Collaborative Depression Study by Lewis Judd and colleagues actually showed that as many as 60 percent of patients with major depressive disorder have a subthreshold fluctuating chronic course.

Rather than exhibiting a frankly depressive mood, many persons with chronic depression experience deficits in their ability to enjoy leisure and display an attitude of irritable moroseness. They also show a sense of resignation, generalized fear of an inability to cope, adherence to rigid routines, and inhibited communication. Such deficits, along with the irritable humor, tend to poison their conjugal lives: Their marriages are typically in a state of chronic deadlock, leading neither to divorce nor to

reconciliation. In other patients, the residual phase is dominated by somatic features, such as sleep and other vegetative or autonomic irregularities. That these interpersonal, conjugal, and autonomic manifestations represent unresolved depression is shown by persistent sleep EEG (especially REM and delta phase) abnormalities that are indistinguishable from their acute counterparts. Unfortunately, self-treatment with ethanol or iatrogenic benzodiazepine dependence rather than definitive treatment of the ongoing low-grade depression is common.

Failure to recover from major depressive disorder is associated with increased familial loading for depression, disabled spouses, deaths of immediate family members, concurrent disabling medical disease, use of depressant pharmacological agents, and excessive use of alcohol and sedative–hypnotic agents. Social support is often eroded in persons with residual depression through the death or illness of significant others. Therefore, a thorough medical evaluation and socially supportive interventions should be essential ingredients of the overall approach to these patients.

Interpersonal disturbances in such patients are usually secondary to the distortions produced by longstanding depression. Therefore, observed pathological characterological changes—clinging or hostile dependence, demanding tendencies, touchiness, pessimism, and low self-esteem—are best considered as *postdepressive personality* changes. A dangerous stereotypical thinking holds that, because a patient has not responded adequately to standard treatments (the illness has become chronic), the disorder must have a characterological substrate. The long duration of the disorder often leads the patient to identify with the failing functions of depression, producing the self-image of being a depressed person. This self-image itself represents a malignant cognitive manifestation of the depressive disorder and dictates vigorous treatment targeted at the mood disorder.

Dysthymic Disorder (Persistent Depressive Disorder in DSM-5)

Dysthymia, called Persistent Depressive Disorder in DSM-5 (Table 13-4–6), is distinguished from chronic depressive disorder by the fact that it is not a sequel to well-defined major depressive episodes. Instead, in the most typical cases, patients complain that they have always been depressed. Thus, most cases are of early onset, beginning in childhood or adolescence and certainly by the time patients reach their 20s. A late-onset subtype, much less prevalent and not well characterized clinically, has been identified among middle-aged and geriatric populations, largely through epidemiological studies in the community.

Table 13.4–6.

DSM-5 Diagnostic Criteria for Dysthymia

This disorder represents a consolidation of DSM-IV-defined chronic major depressive disorder and dysthymic disorder.

A. Depressed mood for most of the day, for more days than not, as indicated by either subjective account or observation by others, for at least 2 years.
   Note: In children and adolescents, mood can be irritable, and duration must be at least 1 year.
B. Presence, while depressed, of two (or more) of the following:
   1. Poor appetite or overeating.
   2. Insomnia or hypersomnia.
   3. Low energy or fatigue.
   4. Low self-esteem.
   5. Poor concentration or difficulty making decisions.
   6. Feelings of hopelessness.
C. During the 2-year period (1 year for children or adolescents) of the disturbance, the individual has never been without the symptoms in Criteria A and B for more than 2 months at a time.
D. Criteria for a major depressive disorder may be continuously present for 2 years.
E. There has never been a manic episode or hypomanic episode, and criteria have never been met for cyclothymic disorder.
F. The disturbance is not better explained by a persistent schizoaffective disorder, schizophrenia, delusional disorder, or other specified or unspecified schizophrenia spectrum and other psychotic disorder.
G. The symptoms are not attributable to the physiological effects of a substance (e.g., a drug of abuse, a medication) or another medical condition (e.g., hypothyroidism).
H. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
   Note: Because the criteria for a major depressive episode include four symptoms that are absent from the symptom list for persistent depressive disorder (dysthymia), a very limited number of individuals will have depressive symptoms that have persisted longer than 2 years but will not meet criteria for persistent depressive disorder. If full criteria for a major depressive episode have been met at some point during the current episode of illness, they should be given a diagnosis of major depressive disorder. Otherwise, a diagnosis of other specified depressive disorder or unspecified depressive disorder is warranted.

Specify if:

With anxious distress (p. 184)
With mixed features (pp. 184–185)
With melancholic features (p. 185)
With atypical features (pp. 185–186)
With mood-congruent psychotic features (p. 186)
With mood-incongruent psychotic features (p. 186)
With peripartum onset (pp. 186–187)

Specify if:

In partial remission (p. 188)
In full remission (p. 188)

Specify if:

Early onset: If onset is before 21 years of age.
Late onset: If onset is at age 21 years or older.

Specify if (for most recent 2 years of persistent depressive disorder):

With pure dysthymic syndrome: Full criteria for a major depressive episode have not been met in at least the preceding 2 years.
With persistent major depressive episode: Full criteria for a major depressive episode have been

met throughout the preceding 2-year period.

With intermittent major depressive episodes, with current episode: Full criteria for a major depressive episode are currently met, but there have been periods of at least 8 weeks in at least the preceding 2 years with symptoms below the threshold for a full major depressive episode.

With intermittent major depressive episodes, without current episode: Full criteria for a major depressive episode are not currently met, but there has been one or more major depressive episodes in at least the preceding 2 years.

Specify current severity:

Mild (p. 188)
Moderate (p. 188)
Severe (p. 188)

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyri ght © 2013). American Psychiatric Association. All Rights Reserved.

Although the dysthymic disorder category in DSM-5 can occur as a secondary complication of other psychiatric disorders, the core concept of dysthymic disorder refers to a subaffective disorder with (1) low-grade chronicity for at least 2 years, (2) insidious onset with origin often in childhood or adolescence, and (3) persistent or intermittent course. Although not part of the formal definition of dysthymic disorder, the family history is typically replete with depressive and bipolar disorders, which is one of the more robust findings supporting its link to primary mood disorder.

Social Adjustment. Dysthymic disorder is typically an ambulatory disorder compatible with relatively stable social functioning. However, the stability is precarious; recent data document that many patients invest whatever energy they have in work, leaving none for leisure and family or social activities, which results in marital friction. These empirical findings on the work orientation of persons with dysthymic disorder echo earlier formulations in the German and Japanese literature. For instance, Kraepelin described such persons as follows: "Life with its activity is a burden which they habitually bear with dutiful self-denial without being compensated by the pleasures of existence."

The dedication of persons with dysthymic disorder to work has been suggested to be an overcompensation and a defense against their battle with depressive disorganization and inertia. Nevertheless, Ernst Kretschmer suggested that such persons are the "backbone of society," dedicating their lives to jobs that require dependability and great attention to detail. Epidemiological studies have demonstrated that some persons with protracted dysthymic complaints, extending over many years, have never experienced clear-cut depressive episodes. Some of them may seek outpatient counseling and psychotherapy for what some clinicians might consider *existential depression,* with feelings of being empty and lacking any joy in life outside of their work. Such persons have been described as

leading "monocategorical existences." Others present clinically because their low-grade dysphoria has intensified into a major depression disorder.

Course. An insidious onset of depression dating back to late childhood or the teens, preceding any superimposed major depressive episodes by years or even decades, represents the most typical developmental background of dysthymic disorder. A return to the low-grade depressive pattern is the rule after recovery from superimposed major depressive episodes, if any; hence, the designation *double depression* as a prominent course pattern. This pattern, commonly seen in clinical practice, consists in the baseline dysthymic disorder fluctuating in and out of depressive episodes. The more prototypical patients with dysthymic disorder often complain of having been depressed since birth or of feeling depressed all the time. They seem, in the apt words of Kurt Schneider, to view themselves as belonging to an "aristocracy of suffering." Such descriptions of chronic gloominess in the absence of more objective signs of depression earn these patients the label of *characterological depression*. The description is further reinforced by the fluctuating depressive picture that merges imperceptibly with the patient's habitual personality makeup. This conceptual uncertainty notwithstanding, given the confluence of data on the efficacy of many classes of antidepressants, patients with dysthymia should not be denied the potential benefit of antidepressants.

Clinical Picture. The profile of dysthymic disorder overlaps with that of major depressive disorder but differs from it in that symptoms tend to outnumber signs (more subjective than objective depression). This means that marked disturbances in appetite and libido are uncharacteristic, and psychomotor agitation or retardation is not observed. This all translates into a depression with attenuated symptomatology. However, subtle endogenous features are not uncommonly observed: inertia and anhedonia that are characteristically worse in the morning.

Although dysthymic disorder represents a more restricted concept than its parent, neurotic depression, it is still quite heterogeneous. Anxiety is not a necessary part of its clinical picture, yet dysthymic disorder is often diagnosed in patients with anxiety and neurotic disorders. That clinical situation is perhaps to be regarded as a secondary or *anxious dysthymia* or, in the framework of Peter Tyrer, as part of a *general neurotic syndrome*. For greater operational clarity, it is best to restrict dysthymic disorder to a primary disorder, one that cannot be explained by a nonmood disorder. The essential features of such primary dysthymic disorder include habitual gloom, brooding, lack of joy in life, and preoccupation with inadequacy. Dysthymic disorder then is best characterized as longstanding, fluctuating, low-grade depression experienced as part of the habitual self and representing an accentuation of traits observed in the depressive

temperament. Sleep EEG data indicate that many persons with dysthymic disorder at baseline exhibit the sleep patterns of those with acute major depressive disorder, providing support for the constitutional nature of the disorder. Further evidence for that position comes from studies demonstrating high rates of familial affective disorder in dysthymic disorder or depressive temperament, or both.

The clinical picture of dysthymic disorder that emerges from the foregoing description is quite varied, with some patients proceeding to major depression, whereas others manifest the pathology largely at the personality level. The foregoing considerations suggest that a clinically satisfactory operationalism of dysthymia must include symptomatic, cognitive, and trait characteristics.

A 27-year-old male grade-school teacher presented with the chief complaint that life was a painful duty that had always lacked luster for him. He said that he felt "enveloped by a sense of gloom" that was nearly always with him. Although he was respected by his peers, he felt "like a grotesque failure, a self-concept I have had since childhood." He stated that he merely performed his responsibilities as a teacher and that he had never derived any pleasure from anything he had done in life. He said that he had never had any romantic feelings; sexual activity, in which he had engaged with two different women, had involved pleasureless orgasm. He said that he felt empty, going through life without any sense of direction, ambition, or passion, a realization that itself was tormenting. He had bought a pistol to put an end to what he called his "useless existence" but did not carry out suicide, believing that it would hurt his students and the small community in which he lived.

Dysthymic Variants. Dysthymia is not uncommon in patients with chronically disabling physical disorders, particularly among elderly adults. Dysthymia-like clinically significant subthreshold depression lasting 6 or more months has also been described in neurological conditions, including stroke. According to a recent WHO conference that generated a book, this condition aggravates the prognosis of the underlying neurological disease and therefore deserves pharmacotherapy. Ongoing studies should provide more explicit clinical recommendations on this topic.

Prospective studies on children have revealed an *episodic course* of dysthymia with remissions, exacerbations, and eventual complications by major depressive episodes, 15 to 20 percent of which might even progress to hypomanic, manic, or mixed episodes postpuberty. Persons with dysthymic disorder presenting clinically as adults tend to pursue a chronic unipolar course that may be complicated by major depression. They rarely develop spontaneous hypomania or mania. However, when treated with antidepressants, some of them may develop brief hypomanic switches that typically disappear when the antidepressant dose is decreased. In this special subgroup of persons with dysthymic disorder, the family histories are often positive for bipolar disorder. These patients represent a clinical bridge between major depressive disorder and bipolar II disorder.

Other Specified Depressive Disorder and Unspecified Depressive Disorder

Minor Depressive Disorder. In so-called minor depression, observed in primary care settings, the depression is subthreshold, milder than major depressive disorder, and yet not protracted enough to be considered dysthymic. These varied manifestations of depression argue for a continuum model (Fig. 13.4–5) as originally envisaged by Kraepelin. Lewis Judd and collaborators at the University of California at San Diego have suggested that subthreshold depressive symptoms—without necessarily meeting the criterion for mood change—might actually represent the most common expressions of a depressive diathesis. From such a subsyndromal symptomatic depressive base, individuals predisposed to depressive illness are said to fluctuate in and out of the various DSM-5 and subthreshold subtypes of depressive disorders. This viewpoint is presently most cogent for subsyndromal symptomatic depression that follows major depressive disorder, a strong predictor of subsequent frequent relapse or chronic course. There is an important message for the clinician here: Treat subsyndromal symptomatic depression residual to major depressive disorder.

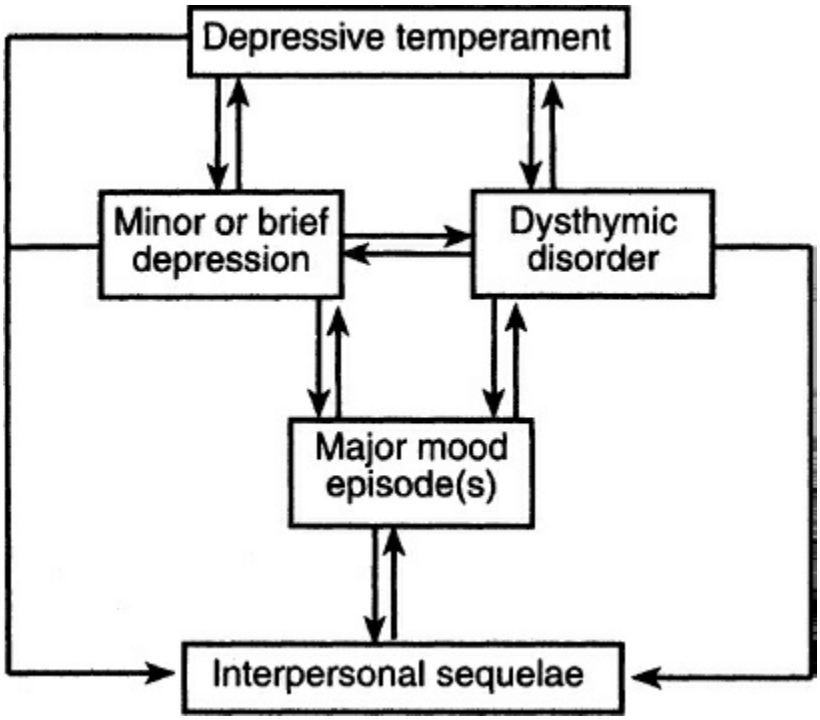

FIGURE 13.4–5. Relation of various depressive conditions supporting a spectrum concept. (From Akiskal HS. Dysthymia: clinical and external validity. *Acta Psychiatr Scand*. 1994;89[Suppl]:19, with permission.)

Premenstrual Dysphoric Disorder. In view of the higher prevalence of depressive disorders in women, premenstrual affective changes—dysphoria, tension, irritability, hostility, and labile mood—have received clinical and, more recently, research attention. The attempt to establish a

specific premenstrual dysphoric disorder beyond this more normative premenstrual tension has neglected the occurrence of premenstrual eutonia, increased energy, and sexual drive. The not uncommon occurrence of these positive emotions, along with the labile mixed affective manifestations, tends to point toward a bipolar phenomenon. Although women with severe premenstrual complaints appear to have higher rates of lifetime major mood disorders, a recent twin study found that genetic and environmental factors contributing to premenstrual depression and major depressive disorders are largely distinct. Furthermore, events such as migraine, epileptic attacks, and panic states may, in some instances, be associated with the premenstrual phase. The foregoing considerations suggest the hypothesis that premenstrual psychobiological changes exacerbate different neuropsychiatric disorders to which women are otherwise predisposed. Whether the exaggerated premenstrual variability in emotional equilibrium constitutes a specific mood disorder or a more familiar affective disorder variant (e.g., mixed panic–depressive, bipolar) must await more-definitive studies. Premenstrual accentuation or precipitation of bipolar episodes, especially of a mixed dysphoric nature, is described in the clinical literature.

Recurrent Brief Depressive Disorder. Recurrent brief depressive disorder derives from British work on young adults with frequent suicide attempts and epidemiological studies conducted in a young adult cohort in Zurich. It is described as short-lived depressions that usually recur on a monthly basis but are *not* menstrually related. They could coexist with major depressive disorder or dysthymic disorder and with hypomania. Again, the existence of hypomanic-prone recurrent brief depressive disorder argues for a bridge between depressive and bipolar II disorders. Such patients are believed to be more prevalent in primary care than in psychiatric settings. Those seen in psychiatric settings are likely to be given Axis II diagnoses, such as borderline personality disorder. The current nosological status of recurrent brief depressive disorder is uncertain, but it testifies to Kraepelin's observation that many transitional forms link the depressive temperament to affective episodes:

> A permanent gloomy stress in all the experiences of life usually perceptible already in youth, and may persist without essential change throughout the whole of life (or) there is actually an uninterrupted series of transitions to periodic melancholia in which the course is quite indefinite with irregular fluctuations and remissions.

Reactive Depression. Classically, *reactive depression* is defined as resulting from a specific life event. In an ideal case, the depression would not have occurred without the event (e.g., love loss) to which it is a reaction. It continues as long as the event is present, and it terminates with the reversal of the event (e.g., return of the lover). Depressions exhibiting

all of those features are almost never seen in clinical practice. With interpersonal support, most people can face life's reverses, which explains why reactive depression tends to be self-limiting. Hence, adjustment disorder is the more appropriate diagnosis for many cases of reactive depression.

Chronic Demoralization. Conceptually, however, one can envision chronically unsatisfactory life situations that might lead to chronic helplessness. However, such a condition, which could warrant the designation of chronic reactive depression, is a contradiction in terms. The question often raised is why a person would continue to stay in the situation. Sometimes psychodynamic authors invoke the concept of masochism to explain why certain persons cannot rid themselves of painful life situations, implying that they somehow contribute to their maintenance. Current thinking is that some of those presumed self-defeating traits are more situation specific than previously believed and might resolve with the elimination of the situation. So-called self-defeating features then are best considered psychodynamic mechanisms rather than indicators of a specific personality. At the present stage of knowledge, they do not deserve to be raised to the level of a nosological entity (hence, their non-inclusion in DSM-5). Chronic adjustment disorder might describe the chronic demoralization observed among some individuals stuck in chronically unsatisfactory life situations. Others might fulfill the criteria for dysthymia. Finally, if the chronically unsatisfactory life conditions involve conjugal violence and trauma, PTSD may be the appropriate diagnosis.

Neurasthenia. A century-old term developed by the American neuropsychiatrist George Beard, *neurasthenia* refers to a more chronic stage of anxious–depressive symptomatology. The anxiety generated by overstimulation is so excessive that it is replaced by a chronic disposition to irritability, fatigue (especially mental fatigue), lethargy, and exhaustion. It is as if the patient's mind refuses to take on new stresses. The clinical picture described by Beard suggests that anxious manifestations were pre-eminent in his time. They included headache, scalp tenderness, backache, heavy limbs, vague neuralgias, yawning, dyspepsia, palpitations, sweating hands and feet, chills, flushing, sensitivity to weather changes, insomnia, nightmares, pantophobia, asthenopia, and tinnitus.

Although the diagnosis of neurasthenia is now used more in China than in the rest of the world, the recent worldwide popularity of the concept of chronic fatigue syndrome attests to the clinical acumen of classic physicians. Despite much energy invested in finding a viral or immunological cause, current descriptions tend to suggest an anxiety or mood disorder basis for many, but certainly not all, of those with the syndrome. Some patients with neurasthenia would meet the criteria for

major depressive disorder with atypical features and history of panic and phobic disorders. What circumstances would lead anxiety or depression to manifest primarily in fatigue is as elusive as it was 100 years ago. Like many other patients presenting to primary care settings with somatic complaints, those with chronic fatigue tend to denounce psychiatric diagnoses as inadequate explanations for their ills.

Postpsychotic Depressive Disorder of Schizophrenia. A postpsychotic depression can occur during the residual phase of schizophrenia. It is characterized by the persistence of negative symptoms or positive symptoms that are in an attenuated form (e.g., odd beliefs, unusual perceptual experiences). The depression is characterized by loss of interest or pleasure and sad or depressed mood. Most typically, the Major Depressive Episode follows immediately after remission of the active phase of the psychotic episode. Sometimes it may follow after a short or extended interval during which there are no psychotic symptoms. Mood symptoms due to the direct physiological effects of a drug of abuse, a medication, or a general medical condition are not counted toward postpsychotic depressive disorder of schizophrenia.

In all postpsychotic depressions, one must first exclude a missed bipolar diagnosis. Negative symptoms due to classic antipsychotics—especially depot phenothiazines and those due to the residuum of schizophrenia once positive symptoms are brought under control—should be distinguished from the depressive episodes that complicate the course of schizophrenia in young, intelligent patients. This phenomenon is so common (at least 30 percent of patients with schizophrenia) that it can be considered as part of the natural course of schizophrenia rather than a separate nosological entity.

BIPOLAR DISORDERS

There are several types of bipolar disorder categories included in DSM-5. Some are categorized on a descriptive level such as bipolar I disorder, bipolar II disorder, cyclothymic disorder, and unspecified bipolar disorder; some are classified according to etiology such as substance or medication induced bipolar disorder and bipolar disorder due to another medical condition. The clinical picture for each is similar regardless of etiology.

Tables 13.4–3 through 13.4–5 describe the syndromes of depression, mania and hypomania that make bipolar disorder I or bipolar II.

Bipolar I Disorder

Typically beginning in the teenage years, the 20s, or the 30s, the first episode of bipolar I disorder could be manic, depressive, or mixed. One common mode of onset is mild retarded depression or hypersomnia for a few weeks or months, which then switches into a manic episode. Others

begin with a severely psychotic manic episode with schizophreniform features; only when a more classic manic episode occurs is the affective nature of the disorder clarified. In a third group, several depressive episodes occur before the first manic episode. A careful history taken from significant others often reveals hyperthymic or cyclothymic traits that antedated the frank onset of manic episodes by several years, if not longer.

Single manic episode (describes patients having a first episode of mania (most such patients eventually develop depressive episodes). The remaining subcategorization is used to specify the nature of the current or most recent episode in patients who have had recurrent mood episodes. For clinicians and researchers alike, it is more meaningful to chart a patient's course in color over time—for example, using red for manic, blue for depressive, and violet for mixed episodes, with hypomanic, dysthymic, and cyclothymic periods drawn in the appropriate colors on a smaller scale between the major episodes. Life events, biological stressors, and treatment can be indicated by arrows on the time axis. This approach, originally championed by Kraepelin, is routinely used in mood or bipolar clinics to chart retrospectively the past course of illness in mood disorders, especially bipolar disorders. Robert Post at the NIMH has developed this approach into systematic clinical science. Its adoption for clinical use was recommended by an international conference that took place in Barcelona. Recent data from the NIMH Collaborative Depression Study—using a prospective method of life charting—of bipolar I disorder clinical cohorts followed in five major US medical centers show more depression, especially its subthreshold variants, during an average of 12 years of follow-up. These patients were actually symptomatic nearly 50 percent of the weeks through the entire follow-up, suggesting a more pervasive course of illness than previously appreciated. Although such data indicate some degree of tertiary care bias in favor of patients willing to be retained in a long-term study, they underscore a clinical reality faced by many clinicians in their daily practice (see Fig. 13.4–6).

On average, manic episodes predominate in youth, and depressive episodes predominate in later years. Although the overall sex ratio is approximately 1:1, men, on average, undergo more manic episodes, and women experience more mixed and depressive episodes. Bipolar I disorder in children is not as rare as previously thought; however, most reported cases are in boys, and mixed manic (dysphoric–explosive) and rapid-cycling presentations are the mode. Childhood-onset depression must also be considered a major risk for ultimate bipolar transformation. This is based on the following characteristics: (1) early age at onset; (2) even sex ratio; (3) prominence of irritability, labile moods, and explosive anger, suggesting mixed episodes; (4) questionable response to antidepressants, hypomanic switches, or both; (5) high recurrence rate after depression; and (6) familial affective loading. Mania can also first appear after 65 years of

age, although a diligent search often reveals a past mild, forgotten, or untreated depressive or excited episode in earlier years.

Acute Mania. Mania typically escalates over a period of 1 to 2 weeks; more-sudden onsets have also been described. The DSM-5 criteria stipulate (1) a distinct period that represents a break from premorbid functioning, (2) a duration of at least 1 week, (3) an elevated or irritable mood, (4) at least three or more classic manic signs and symptoms, and (5) the absence of any physical factors that could account for the clinical picture. The irritable mood in mania can deteriorate to cantankerous behavior, especially when the person is rebuffed. Such patients are among the most aggressive seen in the emergency room. Florid grandiose psychosis with paranoid features, a common presentation of mania, further contributes to the aggression. Alcohol use, observed in at least 50 percent of bipolar I patients (often during the manic phase), further disinhibits the patient and might lead to a dangerous frenzy. Such patients may attack loved ones and hurt them physically. So-called crimes of passion have been committed by patients harboring delusions of infidelity on the part of spouses or lovers, usually when the patient is under the influence of alcohol.

The genesis of delusional, hallucinatory, even first-rank, psychotic experiences in mania has already been described. Recent research has documented that most types of formal thought disorders are common to schizophrenic and mood psychoses; only poverty of speech content (vagueness) emerges as significantly more common in schizophrenia. In addition, posturing and negativism occur may occur during manic periods. Finally, even pseudodemented presentations, can occur. Mania is most commonly expressed as a phase of bipolar I disorder, which has strong genetic determinants. The evidence does not permit separating recurrent mania without depressive episodes from that type as a distinct nosological entity.

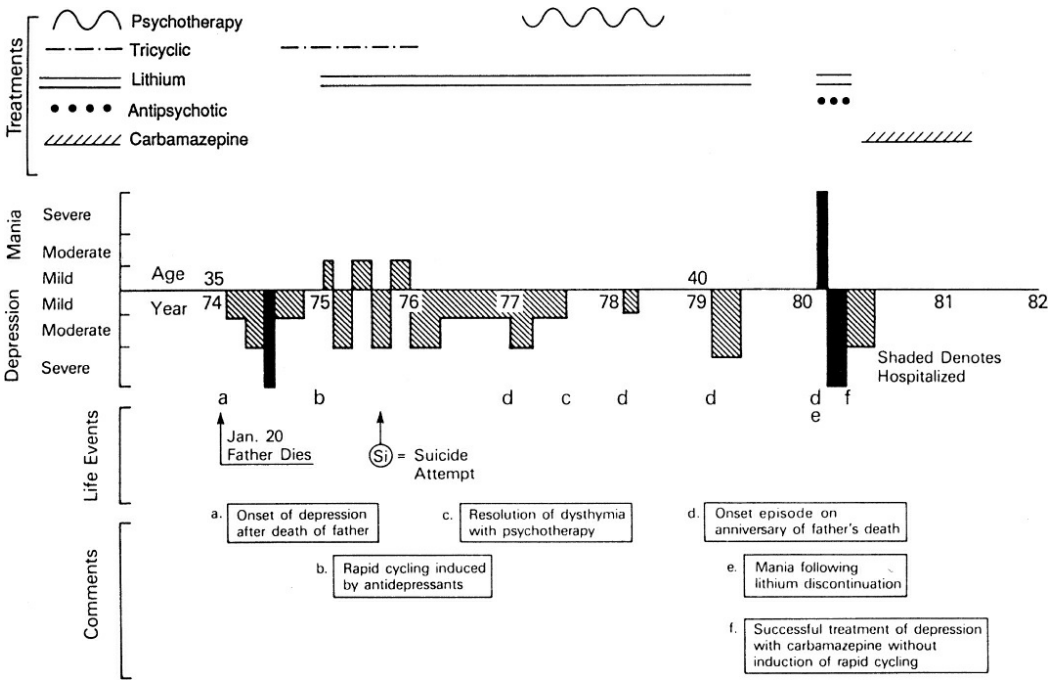

FIGURE 13.4–6. Graphing the course of a mood disorder. Prototype of a life chart. (Courtesy of Robert M. Post, M.D.)

Secondary Mania. Although there is some suggestion that postpartum mania without depression is distinct from familial bipolar I disorder (in which depressive, manic, and especially mixed manic episodes occur in the postpartum period), the evidence for a distinct puerperal mania is not compelling. Mania without prior bipolarity can arise in the setting of such somatic illnesses as thyrotoxicosis, systemic lupus erythematosus or its treatment with steroids, rheumatic chorea, multiple sclerosis (MS), Huntington disease, cerebrovascular disorder, diencephalic and third ventricular tumors, head trauma, complex partial seizures, syphilis, and (most recently) AIDS. The family history is reportedly low in such cases, suggesting a relatively low genetic predisposition and, thus, a lower risk of recurrence. These patients do not easily fit into the DSM-5 category of mood disorder due to a general medical condition) because most of the conditions appear to be cerebral. Such factors must always be diligently sought in manias of late life.

Less well-defined forms of mania are the so-called reactive manias. Personal loss and bereavement are hypothesized to be triggering factors, and the reaction is conceptualized psychodynamically as a denial of loss. Although such explanations may be plausible in individual cases, no systematic data suggest that these patients differ in family history from persons with other manias. The same is generally true for depressed patients who switch to hypomania or mania after the abuse of stimulant drugs, treatment with antidepressants, or sleep deprivation. In all of these situations, a bipolar diathesis is usually manifest in a family history of mania or in spontaneous excited episodes during prospective observation.

First-onset manic episodes can also occur in persons who abruptly abstain from alcohol after one or more decades of chronic use and then develop classic bipolar I disorder.

Chronic Mania. About 5 percent of bipolar I disorder patients have a chronic manic course. These cases commonly represent deterioration of course dominated by recurrent manic episodes grafted onto a hyperthymic baseline. Noncompliance with pharmacological treatment is the rule. Recurrent excitement is personally reinforcing, subjective distress is minimal, and insight is seriously impaired. Thus, the patient sees no reason to adhere to treatment. Episodic or chronic alcohol abuse, prevalent in such patients, has been suggested as a contributory cause of the chronicity. Some authorities further consider comorbid cerebral pathology to be responsible for nonrecovery (and increased mortality) from manic excitements occurring in late life.

Grandiose delusions (e.g., delusions of inventive genius or aristocratic birth) are not uncommon in chronic mania and may lead to the mistaken diagnosis of paranoid schizophrenia. Because of their social deterioration, Kraepelin subsumed such patients under the category *manic dementia*. Organic factors, such as head trauma and chronic alcohol abuse, may contribute to the deterioration. Nonschizoid premorbid adjustment, a family history of bipolar I disorder, and the absence of flagrant formal thought disorder can be marshaled in establishing the affective basis of these poor-prognosis manic states.

Bipolar Mixed Phase. Momentary tearfulness and even depressed mood are commonly observed at the height of mania or during the transition from mania to retarded depression. These transient labile periods, which occur in most bipolar I disorder patients, must be contrasted with mixed episodes proper.

The latter, variously referred to as *mixed mania* or *dysphoric mania,* are characterized by dysphorically excited moods, irritability, anger, panic attacks, pressured speech, agitation, suicidal ideation, severe insomnia, grandiosity, and hypersexuality, as well as persecutory delusions and confusion. Severely psychotic mixed states that involve hallucinations and schneiderian symptoms risk being labeled *schizoaffective*. A correct diagnosis helps to avoid conventional antipsychotic drugs known to exacerbate the depressive component; in such patients, failure to use mood stabilizers can prolong the patient's misery.

New research data from mood centers worldwide on mixed mania suggest that dysphoric mania—mania and full-blown depression occurring simultaneously—is relatively uncommon. Two to four depressive symptoms from the list of depressed mood, helplessness, hopelessness, fatigue, anhedonia, guilt, and suicidal ideation, impulses, or both in the setting of a

manic syndrome appear to suffice for the diagnosis of mixed manic states, which occur in 50 percent of patients with bipolar disorder sometime during their lives. Mixed states occur predominantly in women in whom mania is superimposed on a depressive temperament or a dysthymic baseline. The emerging conceptualization of mixed mania is a manic state intruding on long-term depressive traits. Other mixed manic states, especially in men, often arise from the interaction of substance or alcohol use in bipolar I disorder.

Depressive Phase. Psychomotor retardation, with or without hypersomnia, marks the uncomplicated depressive phase of bipolar I disorder. Onset and offset are often abrupt, although onset can also occur gradually over several weeks. Patients may recover into a free interval or may switch directly into mania. Switching into an excited phase is particularly likely when antidepressants have been used. However, not all patients develop mania after antidepressant treatment of bipolar depression. Some develop a mixed agitated depression; indeed, patients may be stuck for many months in a severe depressive phase with some manic admixtures, such as racing thoughts and sexual arousal. Delusional and hallucinatory experiences are less common in the depressive phase of bipolar I disorder than in the manic and mixed manic phases. Stupor is the more common psychotic presentation of bipolar depression, particularly in adolescents and young adults. Pseudodemented organic presentations appear to be the counterpart of stupor in elderly adults.

Cyclothymic Disorder

An attenuated bipolar disorder that typically begins insidiously before 21 years of age, cyclothymic disorder is characterized by frequent short cycles of subsyndromal depression and hypomania. The course of cyclothymia is continuous or intermittent, with infrequent periods of euthymia. Shifts in mood often lack adequate precipitants (e.g., sudden profound dejection with social withdrawal for a few days switching into cheerful, gregarious behavior). Circadian factors may account for some of the extremes of emotional lability, such as the person's going to sleep in good spirits and waking up early with suicidal urges. The mood changes of cyclothymia are best described as *endoreactive* in the sense that endogenous over-reactivity seems to determine the sudden shifts in mood and behavior (e.g., falling in love with a person one has just met and as quickly falling out of love).

Mood swings in these ambulatory patients are overshadowed by the chaos that the swings produce in their personal lives. Repeated romantic breakups or marital failures are common because of interpersonal friction and episodic promiscuous behavior. Uneven performance at school and work is also common. Persons with cyclothymic disorder are dilettantes; they show great promise in many areas but rarely bring any of their efforts

to fruition. As a result, their lives are often a string of improvident activities. Geographical instability is a characteristic feature; easily attracted to a new locale, job, or love partner, they soon lose interest and leave in dissatisfaction. Polysubstance abuse, which occurs in as many as 50 percent of such persons, is often an attempt at self-treatment. Suicidality has emerged as a public health hazard in adolescents. Many such youth, often labeled "borderline," may actually suffer from undiagnosed temperamental expression of bipolar disorder.

Bipolar II Disorder

Research conducted during the last three decades has shown that, between the extremes of manic-depressive illness defined by at least one acute manic episode, which could be mood congruent or incongruent (bipolar I disorder), and strictly defined major depressive disorder without any personal or family history of mania (pure unipolar disorder), there exists an overlapping group of intermediary forms characterized by recurrent major depressive episodes and hypomania. Current data worldwide indicate that bipolar II disorder is actually more prevalent than bipolar I disorder. This certainly appears to be true in the outpatient setting, in which an average of 50 percent (and in one French national study up to 65 percent) of persons with major depressive disorder have been reported to conform to the bipolar II disorder pattern. The depressive phase of bipolar disorder, especially as expressed in bipolar II and its "soft" expressions, has emerged as a major public health problem.

The following self-description provided by a 34-year-old poet illustrates the pattern:

> I have known melancholy periods, lasting months at a time, when I would be literally paralyzed: All mental activity comes to a screeching halt, and I cannot even utter one word. I become so dysfunctional that I was once hospitalized. Although the paralysis creeps into me insidiously—often lasting months—it typically reverses within hours. I am suddenly alive and vibrant, I cannot turn off my brain neither during the day nor at night; I usually go on celebrating like this for many days, needing no more than few hours of slumber each day.

This vignette is nearly identical to the autobiographical description provided by the British poet William Cowper three centuries earlier:

> I have known many a lifeless and unhallowed hour ... long intervals of darkness interrupted by short returns of peace and joy. . . . For many succeeding weeks to rejoice day and night was all my employment. Too happy to sleep much, I thought it was lost time that was spent on slumber.

Although some hypomanias last for weeks, the hypomania at the end of depressive episodes in most bipolar II disorder patients does not last long; it is usually measured in days. As discussed earlier, hypomanias with short duration (less than 4 days) are familially similar to those with longer duration (more than 4 days). Another common form of bipolar II disorder is major depressive disorder superimposed on cyclothymic disorder, in

which hypomania precedes and follows major depressive disorder, with the entire interepisodic period characterized by cyclothymic mood instability. Such patients tend to have an unstable course, stormy interpersonal relationships, and more irritable and hostile hypomanic episodes—hence the confusion with borderline personality disorder. Current data indicate that cyclothymic and noncyclothymic bipolar II disorder have similar rates of familial bipolarity, which suggests that the borderline characterological features represent a complication of the unstable course that cyclothymia imparts to bipolar II disorder. Such individuals with unstable course have a high risk of suicide—indeed, based on research in Budapest, it appears that the high burden of suicide in mood disorders appears to be due to bipolar II disorder. Descriptively, *hypomania in bipolar II disorder* can be defined as mini manic episodes occurring spontaneously. Bipolar II disorder—especially when major depressions are superimposed on cyclothymia—is thus best characterized as cyclical or *cyclothymic depression*.

The depressive episodes of patients with bipolar disorder often have admixtures (e.g., flight of ideas and increased drives and impulsivity in sexual and other domains). These are best regarded as depressive mixed states, which are basically ignored in the DSM-IV-TR schema. Depressive mixed states are not as severe as dysphoric mixed states but are refractory to antidepressants nonetheless, and, according to current data, are overrepresented in suicide. Current Italian research conducted by Franco Benazzi and the author of this chapter indicates that such depressive mixed states occur in as much as 60 percent of cases of bipolar II disorder and as much as 30 percent of cases of "unipolar" major depressive disorder. Both groups have similar rates of familial bipolarity, suggesting that one of three patients with "unipolar" depression belongs to the bipolar II spectrum.

Hypomania. The common denominator of the soft spectrum of bipolar disorders is the occurrence of hypomania as episode or trait. *Hypomania* refers to a distinct period of at least a few days of mild elevation of mood, sharpened and positive thinking, and increased energy and activity levels, typically without the impairment characteristic of manic episodes. It is not merely a milder form of mania. Hypomania occurring as part of bipolar II disorder rarely progresses to manic psychosis and insight is relatively preserved. Hypomania is distinguished from mere happiness in that it tends to recur (happiness does not!) and can sometimes be mobilized by antidepressants. In cyclothymic disorder, it alternates with mini depressions; in hyperthymic temperament, it constitutes the person's habitual baseline. These definitions, developed by empirical research, recognize three patterns of hypomania: *Episodes* heralding the termination of a retarded depressive state (bipolar II disorder), cyclic alternation with mini depressions (cyclothymic disorder), and an elevated baseline of high mood, activity, and cognition (hyperthymic or chronic hypomanic *traits*).

Because hypomania is experienced as a rebound relief from depression or as a pleasant, relatively short-lived, ego-syntonic mood state, persons with bipolar II disorder rarely report it spontaneously. Skillful questioning is thus required to make the diagnosis of soft bipolar conditions; as in mania, collateral information from family members is crucial. In interviewing the patient, it is customary to use questions along these lines: "Have you had a distinct sustained high period when your mood was so intense that you felt nervous, you were endowed with such energy that others could not keep up with you, and when your thinking and perceptions were unusually vivid or rapid?" Many bipolar patients react to inquiries about "high moods" or "high periods" with a negative: "I am not a manic-depressive!" Benazzi and the author have recently shown that it is best first to inquire about signs and symptoms referring to behavioral activation and to items related to mood. Once energized periods are remembered, the patient is more likely to recall the mood change associated with these periods.

Clinical and epidemiological studies in the United States and Europe have revealed a rich range of hypomanic manifestations including an increase in cheerfulness and jocularity; gregariousness and people seeking; greater interest in sex; talkativeness, self-confidence, and optimism; and decreased inhibitions and sleep need. The clinician must ascertain that those experiences are not due to stimulant or alcohol withdrawal. Depressive and hypomanic periods are often not easily discerned because chronic caffeine use, stimulant abuse, or both may complicate the depression. In such instances, diagnosis should be deferred for 1 month after detoxification. Cocaine and stimulant use is also not uncommon in hypomania, especially in patients with hyperthymic traits, and constitutes an attempt to augment the baseline energy and mood rather than self-treatment of low moods. When in doubt, direct clinical observation of hypomania—sometimes elicited by antidepressant pharmacotherapy—provides definitive evidence for the bipolar nature of the disorder.

Seasonal Patterns. Seasonality is observed in many cyclic depressions, often with autumn or winter anergic depression and energetic periods in the spring. This natural propensity explains why phototherapy may provoke mild hypomanic switches. Seasonal depressions conform, in large measure, to the bipolar II disorder or bipolar III disorder pattern. Furthermore, preliminary evidence suggests that treatment with classic antidepressants disrupts the baseline seasonality, with the depressive phase appearing in the spring and summer. The changes that antidepressants induce in seasonal depressions probably represent a special variant of the rapid-cycling phenomenon.

Temperament and Polarity of Episodes. New systematic clinical

observations have revealed that bipolar II disorder (characterized predominantly by depressive attacks) arises more often from a hyperthymic or cyclothymic baseline, whereas bipolar I disorder (defined by manic attacks) not uncommonly arises from the substrate of a depressive temperament. In line with these observations, a prospective 11-year NIMH study of major depressive disorder patients who switched to bipolar II disorder showed that *mood–labile* (cyclothymic) and *energetic–active* (hyperthymic) temperament traits were highly specific (86 percent) and reasonably sensitive (42 percent) predictors of such an outcome.

Bipolarity is conventionally defined by the alternation of manic (or hypomanic) and depressive episodes. The foregoing data on temperaments suggest that a more fundamental characteristic of bipolarity is the reversal of temperament into its "opposite" episode (in the case of the bipolar II disorder spectrum, from cyclothymia and hyperthymia to major depression). Such findings suggest that the intrusion of cyclothymic and hyperthymic traits into a depressive episode may underlie the instability of the bipolar II disorder subtype and could partly explain why bipolar II depression often has mixed features. These considerations may have important implications for preventing recurrence. For instance, a prospective study of the onset of bipolar disorder in the offspring or siblings of adults with the disorder found that children with depressive onsets as their first episode (which were usually treated with antidepressants) had significantly higher rates of recurrence than those with manic or mixed onsets (treated with lithium) during a 3-year prospective observation. It appears that temperamental instability in the depressive group might have predisposed them to the cycling effect of antidepressants.

Alcohol and Substance Abuse. New evidence supports the high prevalence of alcohol and substance abuse in mood disorder subtypes, especially those with interepisodic cyclothymia and hyperthymia. The relation appears to be particularly strong in the teenage and early adult years, when the use of such substances often represents self-medication for the mood instability. It is not just self-treatment of selected symptoms associated with the down or up phases (e.g., alcohol to alleviate the insomnia and nervousness characteristic of both phases), it also augments certain desired ends (e.g., stimulants to enhance high-energy performance and sexual behavior associated with hypomania). Alcohol may be sought both for its disinhibiting effects and for sedation to calm the nervousness of the high periods. How many patients display alcohol and substance abuse secondary to an underlying bipolar diathesis remains to be determined. However, in view of findings suggesting a link between polysubstance abuse and suicide in adolescents with bipolar familial backgrounds, the use of mood stabilizers in these adolescents should be strongly considered.

Although alcohol and stimulant use continues into adult years in a considerable number of bipolar disorder patients, such use is often unrelated to familial alcoholism and frequently tends to dwindle during long-term follow-up, which supports the self-medication hypothesis.

Rapid-Cycling Bipolar Disorder. *Rapid cycling* is defined as the occurrence of at least four episodes of retarded depression and hypomania (or mania) per year. Thus, rapid cyclers are rarely free of affective symptoms and experience serious vocational and interpersonal incapacitation. Lithium is only modestly helpful to those patients, as are traditional antipsychotic agents; most antidepressants readily induce excited episodes and thus aggravate the rapid-cycling pattern. A balance among mood stabilizers, atypical antipsychotic drugs, and antidepressants may be difficult to achieve. Antidepressants should be avoided because every provoked hypomania is followed by depression, and the roller coaster is further prolonged. Many such patients require frequent hospitalization because they develop explosive excitement and precipitous descent into severe psychomotor inhibition. The disorder is a roller coaster nightmare for the patient, significant others, and the treating physician. Treating these patients is an art. Lamotrigine (Lamictal), a mood stabilizer that stabilizes "from below," has been reported to be efficacious in this group of patients.

As expected, rapid cycling commonly arises from a cyclothymic substrate, which means that most rapid cyclers have bipolar II disorder. Factors favoring its occurrence include (1) female gender, (2) borderline hypothyroidism, (3) menopause, (4) temporal lobe dysrhythmias, (5) alcohol, minor tranquilizer, stimulant, or caffeine abuse, and (6) long-term, aggressive use of antidepressant medications. Most clinically identified patients are bipolar II women in middle age or upper socioeconomic status. Rapid cycling is uncommon from a bipolar I disorder base.

Leadership and Creativity. Persons with hyperthymic temperament and soft bipolar conditions in general possess assets that permit them to assume leadership roles in business, the professions, civic life, and politics. Increased energy, sharp thinking, self-confidence, and eloquence represent the virtues of an otherwise stormy life.

Creative achievement is relatively uncommon among those with the manic forms of the disorder, which are too severe and disorganizing to permit the necessary concentration and application. Notable artistic achievements are found among those with soft bipolar disorders, especially cyclothymic disorders. Psychosis, including severe bipolar swings, is generally incompatible with creativity. That conclusion tends to refute the romantic tendency to idolize insanity as central to the creative process. Because talent is the necessary ingredient of creativity, how might soft bipolarity contribute? The simplest hypothesis is that depression might

provide insights into the human condition, and the activation associated with hypomania helps in producing the artistic work. A more profound interpretation suggests that the repeated self-doubt that comes with recurrent depression might be an important ingredient of creativity because original artistic or scientific expression is often initially rejected and the self-confidence that accompanies repeated bouts of hypomania can help in rehearsing such ideas or expressions until they are perfected. Finally, the tempestuous object relations associated with bipolarity in the parent's or the patient's life often create the unique biographical landmarks that might be immortalized in an artistic medium.

Ultrarapid Cycling. In these patients, hypomanic or manic and depressive episodes alternate rapidly over a few days. In many, the alternation occurs in a regular fashion. The episodes in such cycling patients have greater amplitudes than in cyclothymia, leading to marked impairment. They can be considered as intermediary between cyclothymia and full-blown rapid cycling. Of considerable theoretical and methodological interest, they are difficult to manage clinically. Some individuals with this condition learn to change their work days such that they coincide with the hypomanic periods.

Recurrent Brief Hypomania. This hypomanic counterpart of recurrent brief depressive disorder is best conceptualized as a variant of soft bipolarity described previously. It tends to coincide with cyclothymic depression, in which hypomanias are short (less than 4 days).

Hysteroid Dysphoria. Hysteroid dysphoria combines reverse vegetative signs with the following characteristics: (1) giddy responses to romantic opportunities and an avalanche of dysphoria (angry-depressive, even suicidal, responses) to romantic disappointment; (2) impaired anticipatory pleasure, yet the capability to respond with pleasure when such is provided by others (i.e., preservation of consummatory reward); and (3) craving for chocolate and sweets, which contain phenylethylamine compounds and sugars believed to facilitate cellular and neuronal intake of the amino acid L-tryptophan, hypothetically leading to synthesis of endogenous antidepressants in the brain. The use of the epithet *hysteroid* was used to convey that the apparent character pathology was secondary to a biological disturbance in the substrates governing affect, drives, and reward. The intense, unstable life of the patient with hysteroid dysphoria suggests links to cyclothymic disorder or bipolar II disorder. This suggestion is further supported by the Columbia group's tendency to subsume those patients under atypical depressions (some of which, as indicated, have bipolar affinities). Like patients with bipolar depression, they respond preferentially to monoamine oxidase inhibitors (MAOIs). In brief,

hysteroid dysphoria appears to be a variant of bipolar II disorder with cyclothymic-irritable traits. Such patients are often relegated to the realm of borderline personality. Recent formulations deriving from the Pisa–San Diego Collaboration suggest that cyclothymic traits represent the underlying diathesis of atypical depression, borderline personality, and bipolar II.

UNSPECIFIED MOOD DISORDERS

After all diagnostic information has been obtained, some patients with affective symptoms do not meet the specific criteria for the mood conditions described thus far, nor do they fit the examples listed in the not otherwise specified categories for depressive and bipolar disorders. One example is that of individuals with acute suicidal ideation of an ego-dystonic nature, not accompanied by other affective symptoms. Family or close follow-up, or both, often suggests a more discrete mental condition. In general, such cases should be considered as undiagnosed mood disorders. However, two hybrid conditions that have not been specifically addressed elsewhere in this chapter can be conveniently discussed here.

Mixed Anxiety–Depressive Disorder

Both anxiety and depression can occur together when confronted with a major aversive life situation. The admixture implies that the psychopathology progresses from anxiety to depression, that the patient's mental state is still in flux, and that the ongoing dynamics partly explains the subacute or chronic nature of the disorder. Anxious depression serves to point to the common presence of anxiety in depressive states, especially its greater visibility when the depression is less prominent. Patients with the latter presentation are reportedly most prevalent in general medical settings. This should not come as a surprise, because depressive symptoms that motivate medical consultation commonly complicate generalized anxiety states with a subthreshold level of symptomatology. Some authorities argue that neurotic depressions arise as maladaptive responses to anxiety and, on that basis, suggest retaining the *neurotic depression* rubric. Recent preliminary genetic data indirectly support the contention that certain (unipolar) depressive and (generalized) anxiety states are related. As currently defined, anxious depressions are heterogeneous. In patients refractory to anxiolytic or antidepressant treatment, practitioners must entertain the diagnosis of a complex bipolar II disorder with mixed features. Indeed, recent familial-genetic investigations suggest that bipolar II disorder with panic attacks might represent a special form of bipolar disorder.

Atypical Depression

Originally developed in England and under investigation at Columbia University in New York City, *atypical depression* refers to fatigue superimposed on a history of somatic anxiety and phobias, together with reverse vegetative signs (mood worse in the evening, insomnia, tendency to oversleep and to overeat). Experience suggests that another reverse vegetative sign is increased sexual interest and/or desire, although it remains undescribed in this literature. Sleep is disturbed in the first half of the night in many persons with atypical depressive disorder, and so irritability, hypersomnolence, and daytime fatigue would be expected. The temperaments of these patients are characterized by sensitive traits. The MAOIs and serotonergic antidepressants seem to show some specificity for such patients, which is the main reason that atypical depression is taken seriously.

Other research suggests that reverse vegetative signs can be classified as (1) the anxious type just described or (2) a subtle bipolar subtype with protracted hyperphagic–hypersomnic–retarded dysthymic disorder with occasional brief, extroverted, hypomanic-type behavior, often elicited by antidepressants. Increasing evidence indicates considerable affinity between atypical depression and bipolar II and III disorders. Furthermore, many patients with dysthymic disorder exhibit atypical features at various times. Recent Italian research suggests that many patients with atypical depression meet criteria for brief hypomania or cyclothymic disorder. Such patients also resemble those with hysteroid dysphoria, the precursor of the construct of atypical depression.

DIFFERENTIAL DIAGNOSIS

A missed mood disorder diagnosis means that the disorder does not receive specific treatment, which has serious consequences. Many such persons drop out of school or college, lose their jobs, get divorced, or may commit suicide. Those with unexplained somatic symptoms are frequent users of the general health system. Others are unwell, despite interminable psychotherapy. Some, treated with conventional neuroleptics, develop tardive dyskinesia unnecessarily. As with other medical disorders for which specific treatments are available, accurate diagnosis and early treatment are within the purview of all physicians and mental health professionals. Psychiatrists, in particular, should develop the competence to detect the entire spectrum of mood disorders. Despite massive educational efforts, underdiagnosis and undertreatment of mood disorders remain serious problems worldwide. The most pernicious of these is downplaying the mood component because a personality disorder is given undue prominence.

Although much enthusiasm was generated three decades ago about the potential use of certain biological markers (e.g., REM latency, DST, and thyrotropin-releasing hormone [TRH] test) to corroborate the

differentiation of mood disorder from adjacent disorders, no definitive progress justifies their routine use in clinical practice. The same is true, at least for now, for brain imaging, neuropsychological testing, and genetic screening. Faced with unusual or confusing presentations, a systematic clinical approach is still the best method in differential diagnosis (1) to detail all clinical features of the current episode, (2) to elicit a history of more typical major mood episodes in the past, (3) to assess whether the presenting complaints recur periodically or cyclically, (4) to substantiate adequate social functioning between periods of illness, (5) to obtain a positive family history for classic mood disorder and to construct a family pedigree, and (6) to document a history of unequivocal therapeutic response to thymoleptic medication or ECT in the patient or the family.

Normal Bereavement

Bereaved persons exhibit many depressive symptoms during the first 1 to 2 years after their loss, so how can the 5 percent of bereaved persons who have progressed to a depressive disorder be identified?

- ► Grieving persons and their relatives perceive bereavement as a normal reaction, whereas those with depressive disorder often view themselves as sick and may actually believe they are losing their minds.
- ► Unlike the melancholic person, the grieving person reacts to the environment and tends to show a range of positive effects.
- ► Marked psychomotor retardation is not observed in normal grief.
- ► Although bereaved persons often feel guilty about not having done certain things that they believe might have saved the life of the deceased loved one (guilt of omission), they typically do not experience guilt of commission.
- ► Delusions of worthlessness or sin and psychotic experiences in general point toward mood disorder.
- ► Active suicidal ideation is rare in grief but common in major depressive disorder.
- ► "Mummification" (i.e., keeping the belongings of the deceased person exactly as they were before his or her death) indicates serious psychopathology.
- ► Severe anniversary reactions should alert the clinician to the possibility of psychopathology.

In another form of bereavement depression, the patient simply pines away, unable to live without the departed person, usually a spouse. Although not necessarily pathological by the foregoing criteria, such persons do have a serious medical condition. Their immune function is often depressed, and their cardiovascular status is precarious. Death can ensue within a few months of that of a spouse, especially among elderly

men. Such considerations (highlighted in the work of Sidney Zisook and his colleagues at the University of California at San Diego) suggest that it would be clinically unwise to withhold antidepressants from many persons experiencing an intensely mournful form of grief. A case vignette illustrates this point:

> A 75-year-old widow was brought to the treatment by her daughter because of severe insomnia and total loss of interest in daily routines after her husband's death 1 year before. She had been agitated for the first 2 to 3 months and thereafter "sank into total inactivity—not wanting to get out of bed, not wanting to do anything, not wanting to go out." According to her daughter, she was married at 21 years of age, had four children, and had been a housewife until her husband's death from a heart attack. Past psychiatric history was negative; premorbid adjustment had been characterized by compulsive traits. During the interview, she was dressed in black, appeared moderately slowed, and sobbed intermittently, saying "I search everywhere for him ... I don't find him." When asked about life, she said "Everything I see is black." Although she expressed no interest in food, she did not seem to have lost an appreciable amount of weight. Her DST result was 18 mg/dL. The patient declined psychiatric care, stating that she "preferred to join her husband rather than get well." She was too religious to commit suicide, but, by refusing treatment, she felt that she would "pine away ... find relief in death and reunion."

Anxiety Disorders

Anxiety symptoms, including panic attacks, morbid fears, and obsessions, are common during depressive disorders, and depression is a common complication of anxiety states. Systematic British studies have shown that early-morning awakening, psychomotor retardation, self-reproach, hopelessness, and suicidal ideation are the strongest clinical markers of depression in that differential diagnosis. On follow-up of depressed patients, the manifestations tend to remit, whereas those with anxiety states continue to exhibit marked tension, phobias, panic attacks, vasomotor instability, feelings of unreality, and perceptual distortions, as well as hypochondriacal ideas. A predominance of such anxiety features antedating the present illness suggests the diagnosis of an anxiety disorder. Because anxiety disorders rarely first appear after 40 years of age, a late appearance of marked anxiety features strongly favors the diagnosis of melancholia. The clinical picture is often one of morbid groundless anxiety with somatization, hypochondriasis, and agitation. The depressive nature of the illness is further supported by a superior response to ECT. Periodic monosymptomatic phobic and obsessional states exist that can be regarded as affective equivalents on the basis of a family history of mood disorders and their response to thymoleptic agents. Recent data from a large clinical series suggest that 15 percent of cases have hypomanic symptoms; these patients are best considered to have bipolar II disorder and are treated with lithium salts. Social phobias often usher in adolescent depression, even a bipolar II disorder; the latter is particularly likely when the phobia is paired with alcohol abuse.

The psychopathological differentiation of anxiety and depressive states has not been entirely resolved. Cognitive factors may differentiate them best (Table 13.4–7). Although recurrent (especially retarded) major depressive disorder is a distinct disorder from anxiety states, at least some forms of depression may share a common diathesis with anxiety disorders, particularly generalized anxiety disorders. Before assigning patients to such a putative mixed anxiety–depressive group (not yet an official nosological entity), the clinician must note that anxiety that arises primarily during depressive episodes is best considered as epiphenomenal to depressive disorder. The same is generally true for anxiety symptoms that occur in a person with depressive disorder who is using alcohol or sedative–hypnotic or stimulant drugs. Finally, anxiety symptoms could be prominent features of mixed bipolar states, as well as complex partial seizures.

Table 13.4–7.
Unique Cross-Sectional Profiles of Clinical Anxiety and Depression

Anxiety	Depression
Hypervigilance	Psychomotor retardation
Severe tension and panic	Severe sadness
Perceived danger	Perceived loss
Phobic avoidance	Loss of interest—anhedonia
Doubt and uncertainty	Hopelessness—suicidal
Insecurity	Self-deprecation
Performance anxiety	Loss of libido
	Early-morning awakening
	Weight loss

Reprinted with permission from Akiskal HS. Toward a clinical understanding of the relationship of anxiety and depressive disorders. In: Maser JP, Cloninger CR, eds. *Comorbidity of Mood and Anxiety Disorders*. Washington, DC: American Psychiatric Press; 1990.

Major Depressive Disorder versus Bipolar Disorder

This "internal" boundary question of whether a patient has major depressive disorder versus bipolar disorder has been discussed extensively throughout this work. With the increasing availability of mood-stabilizing agents, it has emerged as a major challenge in clinical practice. Numerous studies have shown that bipolar disorder is confused not only with personality, substance use, and schizophrenic disorders but also with depressive and anxiety disorders. Prospective and other systematic studies have shown that each of the following features—especially in combination—is predictive of bipolar disorder:

- Early age at onset
- Psychotic depression before 25 years of age
- Postpartum depression, especially one with psychotic features

- Rapid onset and offset of depressive episodes of short duration (less than 3 months)
- Recurrent depression (more than five episodes)
- Depression with marked psychomotor retardation
- Atypical features (reverse vegetative signs)
- Seasonality
- Bipolar family history
- High-density three-generation pedigrees
- Trait mood lability (cyclothymia)
- Hyperthymic temperament
- Hypomania associated with antidepressants
- Repeated (at least three times) loss of efficacy of antidepressants after initial response
- Depressive mixed state (with psychomotor excitement, irritable hostility, racing thoughts, and sexual arousal *during* major depression)

More broad indicators of bipolarity include the following conditions, none of which, by itself, clinches a bipolar diagnosis but should raise clinical suspicion in that direction: Agitated depression, cyclical depression, episodic sleep dysregulation, or a combination of these; refractory depression (failed antidepressants from three different classes); depression in someone with an extroverted profession; periodic impulsivity, such as gambling, sexual misconduct, and wanderlust, or periodic irritability, suicidal crises, or both; and depression with erratic personality disorders.

Personality Disorders

The state dependency of most personality measures is well documented. In general, clinicians should refrain from using personality disorder labels in describing patients with active affective illness and should focus instead on competent treatment of the mood disorder. Unfortunately, such advice is not always heeded. Even in patients with chronic or subthreshold mood disorders, personality maladjustment is best considered postaffective, arising from the distortions and conflicts that affective disturbances produce in the life of the patient. The most problematic of the personality labels used in patients with mood disorders is borderline personality disorder, usually applied to teenage and young adult women. Table 13.4–8 shows that the overlap between borderline personality and mood disorders is extensive, so that giving a *borderline* diagnosis to a person with mood disorder is redundant. Use of personality disorder diagnoses may lead to a neglect of the mood disorder or, perhaps, half-hearted treatment of the mood disorder; failure to respond would then be blamed on the patient's "self-defeating character" or "resistance to getting well," thus exculpating the clinician.

Although more systematic research is needed on the complex interface of personality and mood disorders, clinically, they are often inseparable. As with alcohol and substance use disorders, it is generally preferable to diagnose mood disorders at the expense of personality disorders, which should not be difficult to justify in most cases that satisfy the validating strategies outlined previously. When features of personality and mood disorders coexist, it is good practice to defer personality disorder diagnoses and to embark on competent treatment of the concurrent mood disorder. Although not all personality disturbances recede with the competent treatment of mood disorders, so many experienced clinicians have seen such disturbances melt away with the successful resolution of the mood disorder that erring in favor of mood disorders is justified.

A 19-year-old single woman presented with the chief complaint that "all men are bastards." Since her early teens, with the onset of her menses, she had complained of extreme variability in her moods on a nearly daily basis; irritability with hostile outbursts was her main affect, although more-protracted hypersomnic depressions with multiple overdoses and wrist slashings had led to at least three hospitalizations. She also had migrainous headaches that, according to her mother, had motivated at least one of those overdoses. Despite her tempestuous and suicidal moods that led to these hospitalizations, she complained of "inner emptiness and a bottomless void." She had used heroin, alcohol, and stimulants to overcome this troubling symptom. She also gave history of ice cream craving and frequent purging. She was talented in English and wrote much-acclaimed papers on the American confessional poet Anne Sexton. She said that she was mentally disturbed because of a series of stepfathers who had all forced "oral rape" on her when she was between 11 and 15 years of age. She subsequently gave herself sexually to any man that she met in bars, no longer knowing whether she was a "prostitute" or a "nice little girl." On two occasions, she had inflicted cigarette burns inside her vagina "to feel something." She had also engaged in a "brief lesbian relationship" that ultimately left her "emptier" and guilt ridden; nonetheless, she now believed that she should burn in hell because she could not get rid of "obsessing" about the excitement of mutual cunnilingus with her much older female partner. The patient's mother, who owned an art gallery, had been married five times and gave a history of unmistakable hypomanic episodes; a maternal uncle had died from alcohol-induced cirrhosis. The patient's father, a prominent lawyer known for his "temper and wit," had committed suicide. The patient was given phenelzine (Nardil), eventually raised to 75 mg per day, at which point the mother described her as "the sweet daughter she was before age 13." At her next premenstrual phase, the patient developed insomnia, ran away from home at night, started "dancing like a go-go girl, met an incredibly handsome man" of 45 years of age (a pornography shop owner), and had a clandestine marriage to him. After many dose adjustments, she is now maintained on a combination of lithium (900 mg per day) and divalproex (Depakote; 750 mg per day). The patient now attends college and has completed four semesters in art history. In addition to control of her irritable and suicidal moods, bulimic and migraine attacks have abated considerably. Her marriage has been annulled on the basis that she was not mentally competent at the time of the wedding. She is no longer promiscuous and now expresses fear of intimacy with men that she is attracted to. She is receiving individual psychotherapy for this problem.

Table 13.4–8.
Overlap of Borderline Personality Disorder and Mood Disorders

Familial
High rates of mood disorder

Phenomenology
Dysthymic disorder
Cyclothymic disorder
Bipolar II disorder
Mixed state
Pharmacological response
Worsening on most antidepressants
Stabilization on anticonvulsants
Prospective course
Major mood episodes
Suicide

Adapted from Akiskal HS, Chen SE, Davis GC, Puzantian VR, Kashgarian M. Borderline: an adjective in search of a noun. *Clin Psychiatry.* 1985;46:41.

This case illustrates the intimate relationship among atypical depression, borderline personality, and bipolar II disorder. These three conditions may share an underlying psychobiological or genetic diathesis. The complaint is often heard that, even when a mood disorder is diagnosed in a "borderline" patient, response to antidepressants is disappointing. The problem is that affective disorders in these patients usually conform to bipolar II disorder—often complicated by ultrarapid cycling—and many clinicians trained in an earlier era, including some with biological orientation, may lack sufficient experience in the art of pharmacologically managing patients who markedly deviate from classic bipolar I disorder. Recently, lamotrigine has shown promise for such patients.

Alcohol and Substance Use Disorders

The high comorbidity of alcohol and substance use disorders with mood disorders cannot be explained as merely the chance occurrence of two prevalent disorders. Self-medication for mood disorders is insufficiently appreciated by psychiatrists and other professionals who deal with addiction. Given the clinical dangers of missing an otherwise treatable disorder, mood disorder should be seriously considered as the primary diagnosis if marked affective manifestations persist or escalate after detoxification (e.g., 1 month). This consideration also pertains to cyclothymic disorder and dysthymic disorder, which appear particularly likely to invite self-medication. The clinical validating strategies listed previously can further buttress a mood disorder diagnosis.

The DSM category of substance-induced mood disorder is difficult to validate clinically because, in the absence of an affective diathesis, detoxification should, in principle, rapidly clear affective disturbances in persons whose primary problem is that of substance abuse. A dual diagnosis of a mood disorder and a substance use disorder is a more realistic clinical approach to this group of patients. Bipolarity, particularly bipolar II disorder, should be sought in the interface of mood and

substance use disorders.

A 27-year-old married businessman employed in an international family venture owned by his father presented with a court-ordered request for psychiatric treatment. He had been found "bringing" cocaine across the United States–Mexican border and was briefly jailed. He had used stimulants since his late teens to enhance his already high level of energy. His family was rich, and he had no difficulty affording cocaine. During the previous year, he had needed more cocaine because of greater moodiness and fleeting suicidal ideation, which he linked to increasing tensions between him and his father: "My father was never satisfied with me and demanded greater and greater performance from me." His arrest by police was a major embarrassment for him and his family and motivated his compliance with psychiatric hospitalization to detoxify him. He had not had cocaine for 10 days, exhibited marked lability of mood, and gradually sank into a severe hypersomnic-retarded depression of stuporous proportions. He was treated in an inpatient unit with tranylcypromine (Parnate) 20 mg twice a day, and, within 10 days, he switched into hypomania, his mind "exploding with creativity and confidence," marked jocularity and witticisms that entertained other patients, and marked seductiveness toward the nurses. His wife recalled that the patient previously had several such periods naturally (i.e., "off cocaine"), which had strained their marriage due to "brief sexual liaisons." Reducing the tranylcypromine dose by 50 percent did not eliminate the hypomanic behavior, and lithium, 900 mg per day, was added. He has since been maintained on a combination of tranylcypromine and lithium for 4 years; he has not relapsed into cocaine use, and, following a few psychoeducational sessions involving father and spouse, relationships with family and spouse have been less tempestuous. (Since then, consultation was sought by the patient's 60-year-old mother, an attractive, sophisticated woman who confessed that for years she had been engaged in "love relationships" with young artists, with, apparently, her husband's "tacit consent"; since at least her mid-20s, she, by history, would meet the criteria for bipolar II disorder, only treated "on the couch," and her sister and brother had received treatment of "alcohol excesses.") The patient stated that pharmacotherapy, which did require adjustment now and then, helped in balancing the "rough edges" of his "high-nervous temperament" and his "periodic lapse into paralyzing fatigue states that occurred at stressful times." If it had been assumed that these mood states were merely due to cocaine withdrawal, the patient's bipolar II disorder would never have been treated. There is emerging interest in treating dually diagnosed patients with anticonvulsant mood stabilizers and their judicious combination. Among the antidepressants, bupropion (Wellbutrin) is a suitable agent to "replace" stimulant craving; serotonin reuptake inhibitors can be useful for concurrent social phobia and anxious components. The intention is to attenuate any withdrawal phenomena from substances of abuse and to reduce craving while treating any underlying or emerging bipolar diathesis.

Physical Disease

Somatic complaints are common in depressive disorders. Some, such as vegetative disturbances, represent the hypothalamic pathology that is believed to underlie a depressive disorder. Autonomic arousal, commonly associated with depression, could explain such symptoms as palpitations, sweating, and headache. In some instances, the physical symptoms might reflect delusional experiences. The clinician must be vigilant about the likelihood that somatic complaints in depression can also reflect an underlying physical illness. Table 13.4–9 lists the most common medical conditions that have been associated with depression. When depressive symptoms occur in the setting of physical illness, it is not always easy to determine whether they constitute a genuine depressive disorder. Before diagnosing depression, psychiatrists must ensure that they are not dealing

with pseudodepression: (1) functional loss due to physical illness; (2) vegetative signs, such as anorexia and weight loss, as manifestations of such an illness; (3) stress and demoralization secondary to the hospitalization; (4) pain and discomfort associated with the physical illness; and (5) medication adverse effects. However, nonpsychiatric physicians who manage such patients must consider the diagnosis of depression in the presence of persistent anhedonia; observed depressed mood with frequent crying; observed psychomotor retardation or agitation; indecisiveness; convictions of failure, worthlessness, or guilt; and suicidal ideation. The physician should also suspect clinical depression in all patients who refuse to participate in medical care.

Table 13.4–9.
Pharmacological Factors and Physical Diseases Associated with Onset of Depression

Pharmacological
Steroidal contraceptives
Reserpine (Serpalan), á-methyldopa
Anticholinesterase insecticides
Amphetamine or cocaine withdrawal
Alcohol or sedative–hypnotic withdrawal
Cimetidine (Tagamet), indomethacin (Indocin)
Phenothiazine antipsychotic drugs
Thallium, mercury
Cycloserine (Seromycin)
Vincristine (Oncovin), vinblastine (Velban)
Interferon
Endocrine-metabolic[a]
Hypothyroidism and hyperthyroidism
Hyperparathyroidism
Hypopituitarism
Addison disease
Cushing syndrome
Diabetes mellitus
Infectious
General paresis (tertiary syphilis)
Toxoplasmosis
Influenza, viral pneumonia
Viral hepatitis
Infectious mononucleosis
Acquired immune deficiency syndrome
Collagen
Rheumatoid arthritis
Lupus erythematosus
Nutritional
Pellagra
Pernicious anemia
Neurological
Multiple sclerosis

Parkinson disease	
Head trauma	
Complex partial seizures	
Sleep apnea	
Cerebral tumors	
Cerebrovascular infarction (and disease)	
Neoplastic	
Abdominal malignancies	
Disseminated carcinomatosis	

[a]Cholesterol is not mentioned because low levels as a factor in depression have been inconsistently reported.

Diagnosing depression in medically ill elderly patients can be particularly difficult. This task should be undertaken diligently because it was recently reported that (especially in those with cardiovascular disease) mortality is accelerated by depression. Depressed elderly adults often deny being depressed but complain of anxiety, fatigue, and worsening memory. Hypochondriacal symptoms and pain are common. Patients may exhibit extreme negativism and querulousness when invited to participate in medical procedures; others develop poor fluid and food intake out of proportion to their physical conditions.

Another important diagnostic problem at the interface of mood disorder and physical disease is the rare development of malignancy in patients with an established mood disorder. Patients who had responded well to a given antidepressant during previous episodes now have an unsatisfactory response to the same medication. Even a small dose may cause such alarming symptoms as agitation, dizziness, depersonalization, and illusions, which might indicate an occult malignancy, perhaps in the abdomen or the brain. The psychiatrist should always be vigilant about the development of life-threatening physical diseases in patients with pre-established depressive disorder.

A 55-year-old woman had four previous episodes of severe depression that had responded to ECT, amitriptyline (Elavil), or a combination in her native city in the Middle East. She immigrated to the United States at 43 years of age and encountered several major stresses, including an unfamiliarity with English, her daughter dating a man of Chinese extraction, and a complicated series of operations for uterine prolapse. Then, her husband confessed that he had had an affair with a much younger woman. For months, the patient had complained of intermittent fatigue and expressed hostility toward her husband. He had slapped her on her face on two occasions. Her ensuing fifth depressive episode (with pain localized to her face) appeared fully "understandable," but her physician's prescription of 25 mg of amitriptyline resulted in dizzy spells that culminated in syncope. Citalopram (Celexa), 10 mg, did not fare any better. An extensive medical workup revealed a retroperitoneal lymphoma. She died 6 months later.

Stupor. Although less common today, stupor still raises a diagnostic problem in differentiating between a mood disorder and somatic disease, as well as other psychiatric disorders. Depressive stupor is relatively easy to

distinguish from so-called hysterical mutism; in the latter, behavior is meaningfully directed to significant others in the patient's environment. The rubric of catatonic stupor is best reserved for a phase of schizophrenia; in such patients, the schizophrenic origin of the catatonia might be apparent from the patient's history. Otherwise, most acute-onset stupors are probably of affective origin. The main differential diagnosis here is from organic stupor (due to drugs or acute intracranial events); the physical and neurological examination is not always decisive in such cases, and diagnosis depends on a high index of suspicion of possible somatic factors.

Depressive Pseudodementia. The geriatric equivalent of semistupor in younger persons with depressive disorder, depressive pseudodementia, is distinguished from primary degenerative dementia by its acute onset without prior cognitive disturbance, a personal or family history of past affective episodes, marked psychomotor retardation with reduced social interaction, self-reproach, diurnal cognitive dysfunction (worse in the morning), subjective memory dysfunction in excess of objective findings, circumscribed memory deficits that can be reversed with proper coaching, and a tendency to improve with sleep deprivation. Current clinical considerations have revealed "maniform" pseudodementia as well, some of which might be in the early stages of a dementing disease and manifesting with such manifestations as mood lability, aggressivity, and sexual disinhibition in the setting of memory decline and confusion. Family history of bipolar disorder or past history of mild manic-like expressions helps in distinguishing these patients from those with depressive pseudodementia. The distinction is important because antidepressants and anti-Alzheimer agents often worsen the maniform conditions, which instead often respond favorably to anticonvulsant mood stabilizers. The definitive diagnosis in all such cases will be revealed during prospective follow-up. The main clinical message here is not to miss treatable and/or potentially reversible depressive and bipolar states in elderly patients with dementiform manifestations or early and even mid-stage dementia.

Chronic Fatigue Syndrome. Chronic fatigue syndrome is a complex differential diagnostic problem in view of the subtle immunological disturbances presumably associated with it. The following self-report by such a patient illustrates many of the uncertainties marking the present knowledge of the interface between the syndrome and mood disorders.

> I am a 39-year-old, never-married woman, trained as a social worker but currently on disability. I have experienced extreme lethargy and fatigue for many years. I have always felt foggy headed and had trouble thinking and concentrating. My complaint is of fatigue not of depression. My body feels like lead and aches all over. My brain feels achy and sore. I feel much worse in the morning, and I cannot get out of bed; I feel better at night. I feel bad every day. I ache all over, as though someone had beaten me up. Exercise has been prescribed to me, but it makes me worse. Also, I am very sensitive to hot and cold. My sexual drive is low. I have a

general feeling of anhedonia. As far back as I remember—in junior high school—I was always exhausted. I always complained about fatigue, not depression, because that has been the overwhelming problem. I feel the depression is secondary to the fatigue. In high school, I was a compulsive overeater, and I was bulimic for a few years, but it was never severe, and I was only about 10 pounds overweight. In those days, I would sleep 10 or 12 hours a night on the weekend and still feel exhausted; I could not get up for school on Monday. As an adolescent, I felt inferior. I could not make decisions, I did not want to go to camp or leave home for long periods of time—I felt so insecure. Recently, I had a sleep study done, which showed a short latency to stage REM sleep (49 minutes). I was diagnosed as having dysthymic disorder and began taking antidepressants. When I took tranylcypromine, it was the first time in my life that I felt like a normal person. I could play sports, I had a sex drive, I had energy, and I was able to think clearly, but the benefits lasted for barely 2 months. My response was equally short-lived to phenelzine, imipramine (Tofranil), selegiline (Eldepryl), and bupropion. I have not responded to SSRIs at all. I also wish to point out that I had never experienced high periods before I took antidepressants. My main problem has always been one of exhaustion. When I responded to medications, they worked very quickly (within a few days), and I felt great, but they all stopped working after a short time. The dose would be raised, and, again, I would feel better. Eventually, when I got to high doses, I either could not tolerate the high dose or the drug would no longer help. I have taken different combinations of drugs for 10 years, and I have not been able to feel well for more than 6 weeks at a time. Recently, I went to an immunologist. He said I have an abnormality in regulating antibody production and recommended γ-globulin shots. They did not help. When I first started working, I always felt tired and foggy headed, so it was difficult to be sharp while at work. At times, I would close the door to my office and put my head down. Working has become increasingly difficult for me. I had two great jobs, which I blew. As of last year, I had to go on disability. I am desperate for relief, as my condition has drastically affected my life. Disability has been hard for me. I am single and have no other financial resources. I am very despondent, as I feel that my life is passing by without the hope of my ever really improving.

The foregoing clinical picture is compatible with a pseudounipolar or bipolar III disorder as described previously. If so, mood stabilizer augmentation of the tranylcypromine would have been a therapeutic choice to pursue. Some biologists and immunologists, as well as some psychiatrists, believe that abnormal substances circulate in the bloodstream, supplying the brains of such patients. Industrial toxins have also been suggested. Although awaiting more definitive research on the etiology of chronic fatigue syndrome, the psychiatrist can cautiously consider certain patients for thymoleptic trials. That decision can be bolstered by the following considerations: The patient wakes up with fatigue and dread of facing the day; fatigue is part of a more generalized psychomotor inertia or lack of initiative; fatigue is associated with anhedonia, including sexual anhedonia; and fatigue coexists with anxious and pessimistic ruminations. Although none of the foregoing considerations alone is pathognomonic for depression, in aggregate, they point in that direction. The occurrence of hypomanic-like periods (as in the previous vignette) further supports the link between some cases of chronic fatigue and mood disorder. Several recent neuroendocrine challenge studies suggest that some chronic fatigue patients might have a strong anxiety substrate and could be managed accordingly. This is not to say that chronic fatigue is largely a matter of missed affective diagnoses with bipolar

II and panic disorder features; yet, it would be a pity to miss potentially treatable diagnoses. A family or past personal history of classic affective illness or episodes should strongly weigh in this direction. Obviously definitive data are lacking on the essential nature of chronic fatigue, and practitioners should be guided by their clinical experience with therapeutic trials of nonsedating thymoleptic agents while they await new research developments.

Schizophrenia

Despite some genetic overlap between schizophrenic and bipolar disorders, clinically, the two groups of disorders are in the main separate. Young patients with bipolar disorder might seem psychotic and disorganized and, thus, schizophrenic. Their thought processes are so rapid that they may seem loose, but, unlike patients with schizophrenia, they display an expansive and elated affect, which is often contagious. By contrast, the severely retarded bipolar depressive person, whose affect may superficially seem flat, almost never exhibits major fragmentation of thought. The clinician, therefore, should place greater emphasis on the pattern of symptoms than on individual symptoms in the differential diagnosis of mood and schizophrenic psychoses. No pathognomonic differentiating signs and symptoms exist. Differential diagnosis should be based on the overall clinical picture, phenomenology, family history, course, and associated features. Because the two groups of disorders entail different combinations of pharmacological treatments on a long-term basis (mood stabilizers are usually needed as the mainstay in bipolar disorder), the differential diagnosis is of major clinical importance. Table 13.4–10 summarizes clinical experience in the area and lists the most common pitfalls in diagnosis. In the past, many bipolar patients, especially those with prominent manic features at onset, were labeled as having *acute schizophrenia* or *schizoaffective schizophrenia*. In the past, such misdiagnoses (which typically led to long-term treatment with conventional neuroleptic agents) were costly in terms of tardive dyskinesia, vocational and social decline, and even suicide. For instance, some patients with postpsychotic depressive disorder of schizophrenia have postmanic depressions that were treated with classic neuroleptic monotherapy without the benefit of more definitive thymoleptic agents.

Table 13.4–10.
Common Causes of Misdiagnosis of Mood Disorder as Schizophrenia

Reliance on cross-sectional rather than longitudinal picture
Incomplete interepisodic recovery equated with schizophrenic defect
Equation of bizarreness with schizophrenic thought disorder
Ascribing irritable and cantankerous mood to paranoid delusions

Mistaking depressive anhedonia and depersonalization for schizophrenic emotional blunting
Flight of ideas perceived as loose associations
Lack of familiarity with the phenomenological approach in assessing affective delusions and hallucinations
Heavy weight given to incidental schneiderian symptoms

Adapted from Akiskal HS, Puzantian VR. Psychotic forms of depression and mania. *Psychiatr Clin North Am.* 1979;2:419.

Modern treatments—even with atypical antipsychotics—that tend to keep many persons with schizophrenia out of the hospital do not always seem to prevent an overall downhill course in many cases. By contrast, the intermorbid periods in bipolar illness are relatively normal or even supernormal, and yet, over time, some social impairment may result from the accumulation of divorces, financial catastrophes, and ruined careers. (Although rapid-cycling disorders, which have sharply risen during the past two decades, cause considerable social impairment, mood symptoms are so prominent that differentiation from schizophrenia is generally not difficult; in addition, such patients usually display more classic bipolar phases before the rapid cycling.)

Postpsychotic depressions in persons with established schizophrenia are sometimes due to inadequate control of schizophrenic symptomology. In other patients, especially more intelligent young patients with schizophrenia, they reflect the experience of losing one's ego and sanity. It would be more meaningful to give such patients a diagnosis of schizophrenia and a superimposed major depressive diagnosis and to treat the patient accordingly.

Schizoaffective Disorder

As the previously mentioned considerations suggest, depression in the setting of a schizophrenic disorder does not necessarily constitute a distinct nosological entity. The concept of schizoaffective (or cycloid) psychosis should be restricted to recurrent psychoses, with full affective and schizophrenic symptoms occurring nearly simultaneously during each episode. This diagnosis should *not* be considered for a mood psychosis in which mood-incongruent psychotic features (e.g., schneiderian and bleulerian symptoms) can be explained on the basis of one of the following: (1) affective psychosis superimposed on mental retardation, giving rise to extremely hyperactive and bizarre manic behavior; (2) affective psychosis complicated by concurrent brain disease, substance abuse, or substance withdrawal, known to give rise to numerous schneiderian symptoms; or (3) mixed episodes of bipolar disorder (which are notorious for signs and symptoms of psychotic disorganization). Official diagnostic systems, such as DSM-5, use the category of schizoaffective disorder broadly. Thus, patients with clear-cut manic episodes receive a schizoaffective diagnosis if

delusions or hallucinations occur in the interepisodic period in the absence of prominent affective symptoms. Many psychotic symptoms in mood disorders are often explanatory (albeit delusional), whereby the patient tries to make sense of the core experiences of the affective illness. In patients with recurrent episodes, delusional thinking can be carried over into the interepisodic period. Such patients are thus delusional in the absence of prominent mood symptoms and technically might be considered schizoaffective.

However, affective illness is typically a lifelong process, and limiting its features to discrete episodes is artificial. In the past, neuroleptic agents were prescribed on an as-needed basis to reduce the strong affective charge of those interepisodic delusions, but they did not effectively eliminate the affect-laden experiences. Atypical antipsychotic agents introduced during the 1990s are more likely to be beneficial in this regard. Augmenting with continued thymoleptic treatment (resorting to ECT, if necessary) and an empathic psychotherapeutic approach may prove more rewarding in the long run.

A 29-year-old female college graduate, mother of two children and wife of a bank president, had experienced several manic and retarded depressive episodes that had responded to lithium carbonate. She was referred to me because she had developed the delusion that she had been involved in an international plot. Careful probing revealed that the delusion represented further elaboration, in a rather fantastic fashion, of a grandiose delusion that she had experienced during her last postpartum manic episode. She believed that she had played an important role in uncovering the plot, thereby becoming a national hero. Nobody knew about it, she contended, as the circumstances of the plot were top secret. She further believed that she had saved her country from the international scheme and suspected that she was singled out for persecution by the perpetrators of the plot. At one point, she had even entertained the idea that the plotters sent special radio communications to intercept and to interrupt her thoughts. As is typical in such cases, she was on a heavy dose of a lithium–antipsychotic combination. The consultation was requested because the primary mood symptoms were under control, and yet she had not given up her grandiose delusion. She flippantly remarked that "I must be crazy to believe in my involvement in an international plot," but she could not help but believe in it. Over several months, seen typically in 60-minute sessions weekly, the patient had developed sufficient trust that her beliefs could be gently challenged.

She was, in effect, told that her self-professed role in the international scheme was highly implausible and that someone with her superior education and high social standing could not entertain a belief, to use her own words, "as crazy as that." She eventually broke into tears, saying that everyone in her family was so accomplished and famous that to keep up with them she had to be involved in something grand; in effect, the international scheme, she said, was her only claim to fame: "Nobody ever gives me credit for raising two kids, and throwing parties for my husband's business colleagues: My mother is a dean, my older brother holds high political office; my sister is a medical researcher with five discoveries to her credit [all true] and who am I? Nothing. Now, do you understand why I need to be a national hero?" As she alternated, over subsequent months, between such momentary flashes of insight and delusional denial, antipsychotic medication was gradually discontinued. Maintained on lithium, she now only makes passing reference to the grand scheme. She was encouraged to pursue her career goal toward a master's degree in library science.

The vignette illustrates how phenomenological understanding, rational pharmacotherapy, and practical psychotherapeutic or vocational guidance

can be fruitfully combined in the approach to patients with psychotic mood disorders. It also indicates that problems with insight are not necessarily state dependent in mood disorders. Cognitive disturbances have recently been described in all mood disorders, even when they are nonpsychotic. At a more fundamental level, the vignette suggests that clinical diagnoses cannot be based entirely on operational criteria; one's opinion of a patient's illness is not infrequently changed by its response to treatment.

REFERENCES

Akiskal HS. Bipolar depression: focus on phenomenology. *J Affect Disord*. 2005;84(2–3):105–106.

Akiskal HS, Akiskal KK. TEMPS: temperament evaluation of Memphis, Pisa, Paris and San Diego. *J Affect Disord*. 2005;85(1–2):3–16.

Akiskal HS, Akiskal KK. In search of Aristotle: temperament, human nature, melancholia, creativity and eminence. *J Affect Disord*. 2007;100:1–6.

Akiskal HS, Akiskal KK, Lancrenon S, Hantouche EG, Fraud J-P. Validating the bipolar spectrum in the French National EPIDEP Study: overview of the phenomenology and relative prevalence of its clinical prototypes. *J Affect Disord*. 2006;96:197–205.

Akiskal HS, Azorin JF, Hantouche EG. A proposed multidimensional structure of mania: beyond the euphoric–dysphoric dichotomy. *J Affect Disord*. 2003;73:7–18.

Akiskal HS, Benazzi F. Family history validation of the bipolar nature of depressive mixed states. *J Affect Disord*. 2003;73:113–122.

Akiskal HS, Benazzi F. Optimizing the detection of bipolar II in outpatient private practice: toward a systematization of clinical diagnostic wisdom. *J Clin Psychiatry*. 2005;66:914–921.

Akiskal HS, Benazzi F, Perugi G, Rihmer Z. Agitated "unipolar" depression re-conceptualized as a depressive mixed state: implications for the antidepressant-suicide controversy. *J Affect Disord*. 2005;85:245–258.

*Akiskal HS, Bourgeois ML, Angst J, Post R, Moller HJ. Re-evaluating the prevalence of and diagnostic composition within the broad clinical spectrum of bipolar disorders. *J Affect Disord*. 2000;59(Suppl 1):S5–S30.

Akiskal HS, Cassano GB, eds. *Dysthymia and the Spectrum of Chronic Depressions*. New York: Guilford Press; 1997.

Akiskal HS, Hantouche EG, Lancrenon S. Bipolar II with and without cyclothymic temperament: "Dark" and "sunny" expressions of soft bipolarity. *J Affect Disord*. 2003;73:49–57.

Akiskal HS, Mallya G. Criteria for the "soft" bipolar spectrum: treatment implications. *Psychopharmacol Bull*. 1987;23:68–73.

Akiskal HS, Maser JD, Zeller P, Endicott J, Coryell W. Switching from "unipolar" to bipolar II: an 11-year prospective study of clinical and temperamental predictors in 559 patients. *Arch Gen Psychiatry*. 1995;52:114–123.

Akiskal HS, Walker PW, Puzantian VR, King D, Rosenthal TL. Bipolar outcome in the course of depressive illness: phenomenologic, familial, and pharmacologic predictors. *J Affect Disord*. 1983;5:115–128.

Alhajji L, Nemeroff CB. Personalized medicine and mood disorders. *Psychiatr Clin North Am*.

2015;38(3):395–403.

Angst J, Ajdacic-Gross V, Rossler W. Classification of mood disorders. *Psychiatr Pol*. 2015;49(4):663–671.

Balázs J, Benazzi F, Rihmer Z, Rihmer A, Akiskal KK. The close link between suicide attempts and mixed (bipolar) depression: implications for suicide prevention. *J Affect Disord*. 2006;91:133–138.

Clayton PJ. The sequelae and nonsequelae of conjugal bereavement. *Am J Psychiatry*. 1979;136:1530–1534.

Coplan JD, Aaronson CJ, Panthangi V, Kim Y. Treating comorbid anxiety and depression: psychosocial and pharmacological approaches. *World J Psychiatry*. 2015;5(4):366–378.

Davidson JR, Miller RD, Turnbull CD, Sullivan JL. Atypical depression. *Arch Gen Psychiatry*. 1982;39:527–534.

*Judd LL, Akiskal HS, Maser JD, et al. A prospective 12-year study of subsyndromal and syndromic depressive symptoms in unipolar major depressive disorders. *Arch Gen Psychiatry*. 1998;55:694–700.

Kuswanto C, Chin R, Sum MY, et al. Shared and divergent neurocognitive impairments in adult patients with schizophrenia and bipolar disorder: whither the evidence? *Neurosci Biobehav Rev*. 2016;61:66–89.

Malhi GS, Byrow Y, Fritz K, et al. Mood disorders: neurocognitive models. *Bipolar Disord*. 2015;17(Suppl 2):3–20.

Ng B, Camacho A, Lara DR, Brunstein MG, Pinto OC, Akiskal HS. A case series on the hypothesized connection between dementia and bipolar spectrum disorders: bipolar type VI? *J Affect Disord*. 2008;107:307–315.

Post RM, Altshuler LL, Kupka R, et al. More illness in offspring of bipolar patients from the U.S. compared to Europe. *J Affect Disord*. 2016;191:180–186.

Reynolds CF, Hoch CC, Kupfer DJ, Buysse DI, Hofick PR. Bedside differentiation of depressive pseudodementia from dementia. *Am J Psychiatry*. 1988;145:1099–1103.

Rosenthal NE. *Winter Blues*. New York: Guilford Press; 1993.

Rumble ME, White KH, Benca RM. Sleep disturbances in mood disorders. *Psychiatr Clin North Am*. 2015;38(4):743–759.

Singh MK, Garrett AS, Chang KD. Using neuroimaging to evaluate and guide pharmacological and psychotherapeutic treatments for mood disorders in children. *CNS Spectr*. 2015;20(4):359–368.

Torres F, Barbosa M, Maciel P. Recurrent copy number variations as risk factors for neurodevelopmental disorders: critical overview and analysis of clinical implications. *J Med Genet*. 2016;53(2):73–90.

*Widlocher DJ. Psychomotor retardation: clinical, theoretical, and psychometric aspects. *Psychiatr Clin North Am*. 1983;6:27–40.

*Winokur G, Turvey C, Akiskal HS, et al. Alcoholism and drug abuse in three groups—bipolar I, unipolars, and their acquaintances. *J Affect Disord*. 1998;50:81–89.

Yarlagadda A, Acharya G, Kasaraneni J, Hampe CS, Clayton AH. Placental barrier and autism spectrum disorders: the role of prolactin and dopamine on the developing fetal brain. *Innov Clin Neurosci*. 2015;12(9–10):14–17.

Young AS, Fristad MA. Family-based interventions for childhood mood disorders. *Child Adolesc Psy*chiatr *Clin North Am*. 2015;24(3):517–534.

Zaninotto L, Solmi M, Toffanin T, Veronese N, Cloninger CR, Correll CU. A meta-analysis of temperament and character dimensions in patients with mood disorders: comparison to healthy controls and unaffected siblings. *J Affect Disord*. 2016;194:84–97.

▲ 13.5 Mood Disorders: Intrapsychic and Interpersonal Aspects

JOHN C. MARKOWITZ, M.D., AND BARBARA L. MILROD, M.D.

The symptom criteria listed in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) reliably define mood disorders, but they do not completely describe these syndromes or explain how patients experience them. Depressed patients suffer emotionally, cognitively, and physically. There are three key theoretical approaches that have been applied to mood disorders: psychoanalytic/psychodynamic, cognitive, and interpersonal. They vary in their frequency of use in clinical practice and in the amount of research that they have received.

Theory has uses and limitations. Theory can help to organize clinical and research thinking about depression and can pinpoint aspects of the syndrome. A theory can provide a narrative thread or focus to anchor a clinical formulation and to give coherence to a therapy. Indeed, some theoretical approach is necessary for clinicians to impose order on the overwhelming amount of clinical data with which patients present for the treatment. Theories may also allow predictions to be made about the treatment mechanisms and outcomes. The authors submit that an organizing theoretical framework is as essential to structure psychotherapy as grammar is to speech. Hence, understanding the theoretical backgrounds of psychotherapies is crucial for the psychotherapeutic clinician.

Theories should be considered approximations of the truth, needful of testing and subject to disconfirmation. Thomas Huxley in 1870 famously described "The great tragedy of Science—the slaying of a beautiful hypothesis by an ugly fact." Dogma and ideology have done psychiatrists and other mental health professionals more harm than good, for example, in separating psychotherapy and pharmacotherapy into opposing camps—a rift from the 1950s that is only now healing. Inasmuch as the etiology of mood disorders is complex and no single theory comprehensively explains it, rigid adherence to any theory is unwarranted and potentially problematic. Clinicians should understand and use multiple theories, flexibly choosing the optimal theoretical perspective to apply to a given patient's illness and then persevering in using that model for an appropriate course of treatment. They should then proceed following that theoretical approach rather than devolving into eclecticism. (The authors of this chapter recommend formal training in any treatment modality the clinician plans to undertake. Theoretical understanding does not suffice.) However, to develop clinical expertise in any therapy, it is essential to be able to understand patients' problems through the lens of the theory

underlying it.

Psychodynamic theory has the longest historical tradition. Both cognitive theory and psychodynamic theory largely focus on intrapsychic phenomena, whereas interpersonal theory, the newest and least formally developed, focuses more on interpersonal, extrapsychic reality. It is widely accepted that aspects of both cognitive and interpersonal theories derive from psychoanalysis. Each approach provides a potentially useful explanation of the plight of the depressed patient. Although the utility of therapies derived from cognitive theory (cognitive behavioral therapy [CBT]) and interpersonal theory (interpersonal psychotherapy [IPT]) has been better tested for depression than psychodynamic therapy to date, no theories of depression have been formally tested in all of their assumptions. They are best considered working clinical models.

PSYCHODYNAMIC ASPECTS OF MOOD DISORDERS

A contemporary psychoanalytic understanding of mood disorders includes a comprehensive focus on biological underpinnings, cognitive function, and interpersonal situation and style. What is unique to psychoanalysis is its attention to intrapsychic, unconscious pressures in its consideration of psychological symptoms, including mood disorders. In order to describe intrapsychic aspects of mood disorders, several basic psychoanalytic ideas about mental life first must be defined:

1. From a psychodynamic view, mental life exists on two levels: Both within the realm of consciousness and also within a less accessible realm, described as "the unconscious." Psychic or emotional symptoms arise from aspects of mental life that are at least in part unconscious. This is true of mood disorders.
2. Psychoanalysts have found it useful to conceptualize the mind as comprising three basic theoretical psychic structures, *the id*, *the ego*, and *the superego*. In brief, *the id* is considered the aspect of the mind that subsumes the drive derivatives and desires. *The ego* serves as intermediary between the id and the external world. It contains many intrapsychic functions, including motor action, perception, self-esteem, the relationship to reality, and the ability to modulate the drives and to modulate anxiety. Defense mechanisms, which are unconscious intrapsychic mechanisms employed to modulate both the drives and anxiety, are ego functions. *The superego* is conceptualized to comprise the person's value system. The superego can both punish and reward the person, depending upon whether or not his or her actions are consonant with his or her moral values.
3. Moods are pervasive ego states that color the entire ego with the same affect state. Unlike simple affective responses to events, they are not focused but general, either because the affect is too strong or because the

ego is too weak to contain a focused response. Because moods tend to be generalized states, they typically involve some degree of denial of the opposite feelings. From a psychoanalytic vantage, moods carry an unconscious significance, notwithstanding their presumed biological and neurotransmitter underpinnings.

4. Since Sigmund Freud's discovery of the importance of the unconscious in everyday mental life and of individuals' capacity to shut off unwanted, painful, emotionally laden experiences using defense mechanisms, the current understanding of moods and symptoms has been enhanced by the recognition that these psychic phenomena represent a breakthrough of *unconscious fantasy* into consciousness. Persistent unconscious fantasies often underlie people's psychological symptoms, dreams, personalities, and important life choices.
5. Freud initially described another central principle in the organization of mental life: That people unconsciously avoid "unpleasure" and that ideas that produce unpleasure are screened from consciousness by "repression," or processes now called *defenses*. Clinically, the degree to which unpleasure is avoided varies from person to person, and Freud later modified his original theory about people avoiding unpleasure. In *Beyond the Pleasure Principle*, he attempted to describe intrapsychic mechanisms that could account for phenomena in which the pursuit of pleasure is given up for the more fundamental need to discharge intense emotional tensions to protect the individual. Thus, major depression is a condition in which the person is living beyond the pleasure principle.

PSYCHOANALYTIC DESCRIPTIONS OF MAJOR DEPRESSION

Psychoanalytic formulations highlight the phenomenology of aspects of major depression, helping to make the emotional backdrop of the illness more understandable. Although these theories provide a clinically rich set of ideas that can be useful in treating such patients, it should be emphasized that no theoretical positions are supported by systematic research. The field of psychoanalysis has been fairly resistant to research. Nor are these psychoanalytic formulations, despite their differences, mutually exclusive. Some of these formulations shall be outlined, and illustrative clinical examples shall be provided.

Response to Loss/Anger Turned Inward

The classic psychoanalytic understandings of depression were stated by Karl Abraham, Freud, and Sandor Rado and emphasized depressed patients' reactions to object loss, in reality or in fantasy. In these formulations, the profound response to loss is believed to occur in part because the current loss invokes an earlier, childhood loss, also either of a fantasy or a reality nature. These authors noted depressed patients'

ambivalent or hostile object relations, along with object attachments characterized by excessive dependency, laced with an emphasis on need gratification in emotionally charged relationships. Major depression occurs only after the tie to the object is shattered. In *Mourning and Melancholia*, Freud highlighted the way in which depressed patients irrationally attack themselves. In his formulation, this occurs because aspects of an ambivalent object become internalized, or incorporated, into the patient's sense of self, and hostility directed toward the object is instead directed at the self. This state of affairs serves to preserve the relationship with the other person (the object) in reality.

Stuart Asch described a variation in these dynamics in which some patients become depressed not because of object loss but due to maintaining a submissive tie to the object. This tie is maintained because the patient perceives separation to be an aggressive act with destructive consequences. Rage at having to play this part is directed toward the self, as the patient feels devalued because of the submission.

Joseph Sandler and Walter Joffe, also focusing on the phenomenon of loss leading to depression, called the Hampstead Index, a comprehensive clinical registry of childhood responses to abandonment and loss, for cases of childhood depression. They inferred that depression is a basic affective response to loss. They emphasized that more than just the other person is lost and that the child feels that he or she has also lost a sense of self in relationship to the lost object, or a previous set of feelings about the self. Thus, as Freud and Abraham also noted, Sandler and Joffe emphasize a symbiotic or narcissistic tie to the object. They saw individuals predisposed to depression as struggling against feelings of helplessness and injured self-esteem in childhood.

Ms. A, a 26-year-old woman, presented with severe major depression in the context of her boyfriend B's expressing increasing anger at her and his saying that he wanted to move out of their home. An anxious person at baseline, Ms. A took to her bed for several weeks, closing the blinds so that no light could enter her room. She stopped eating, lost weight, and slept most of the time. She dropped out of school and took a leave of absence from work. She thought almost constantly of suicide, although she had not picked a method and her deep religious commitment made this an unacceptable choice. In her first session, she spoke about how worthless she felt and how she blamed herself for her boyfriend wanting to break up with her. She said she thought that she was too "pointless," "anxious," and "boring" for him and that her presence stifled him. She expressed no anger toward B at the time of her presentation, saying that she "understood" why he was being like this. She also said she hoped that this was just a "bad phase" that he would get over, although he told her repeatedly that he was dating another woman. In psychotherapy it gradually emerged, however, that over the past month B had been breaking dishes and furniture in their home. He yelled at Ms. A almost constantly. This patient's depression remitted with psychodynamic psychotherapy only as she became aware of her own murderous rage toward B, at which time she began to view herself in a less negative and denigrated manner.

Guilt

Melanie Klein postulated that depressed patients fear that they cannot protect an idealized, or good, internalized "other" from destructive, rageful impulses. Although emphasizing a different facet of major depression, this view coincides with Freud's focus on the destruction of the object tie in major depression. As a result, the depressed patient's characteristic guilt, inhibitions, and punitive superego develop. However, not all depressions are characterized by excessive guilt, and Klein's description applies only to this subset of patients. Klein also highlighted the danger that depressed patients foresee in "triumphing" over parents or siblings via any life success: Success is experienced as aggressively humiliating to loved ones or as damaging to others. Klein theorized that idealization and devaluation are "manic defenses" against the guilt and sense of loss experienced in the resulting depression.

Ms. C, a 23-year-old woman, became acutely depressed when she was accepted to a prestigious graduate school. Ms. C had been working diligently toward this acceptance for the past 4 years. She reported being "briefly happy, for about 20 minutes" when she learned the good news but rapidly slipped into a hopeless state in which she recurrently pondered the pointlessness of her aspirations, cried constantly, and had to physically stop herself from taking a lethal overdose of her roommate's insulin. In treatment, she focused on her older brother, who had regularly insulted her throughout the course of her life, and how "he's not doing well." She found herself very worried about him. She mentioned that she was not used to being the "successful" one of the two of them. In connection with her depression, it emerged that Ms. C's brother had had a severe, life-threatening, and disfiguring pediatric illness that had required much family time and attention throughout their childhood. Ms. C had become "used to" his insulting manner toward her. In fact, it seemed that she required her brother's abuse of her in order not to feel overwhelmed by survivor guilt about being the "healthy, normal" child. "He might insult me, but I look up to him. I adore him. Any attention he pays to me is like a drug," she said. Ms. C's acceptance to graduate school had challenged her defensive and essential compensatory image of herself as being less successful, or damaged, in comparison with her brother, thereby overwhelming her with guilt. Her depression remitted in psychodynamic psychotherapy as she better understood her identification with and fantasy submission to her brother.

Impairment in Self-Esteem Regulation

A general trait of patients with major depression is the loss of self-regard. Yet loss of self-regard can occur in the absence of depression. Edward Bibring disagreed with Klein's formulation that emphasized the importance of a punitive superego and argued that conflicts about aggression and object loss were secondary determinants in depression. He viewed depression instead as resulting from a sense of helplessness, impaired self-esteem, and self-directed anger triggered by failures to live up to the narcissistic aspirations of any developmental phase.

Charles Brenner also de-emphasized the classic psychoanalytic focus on object loss, seeing a depressive propensity as equally likely to connect with organizing fantasies of narcissistic injury, and particularly of castration. Brenner states that these fantasies are accompanied by reactive aggression against those blamed for the painful affects, with consequent guilt.

Mr. D, a 39-year-old successful and aggressive lawyer, had "never been depressed a day in my life" before developing a kidney stone that required a 5-day hospital stay full of medical interventions. He was well on discharge but developed major depression, with lethargy, decreased energy, hopelessness, and weight loss. He cried all the time. "I feel like a big, fat failure," he announced. Handsome but insecure, Mr. D prided himself on his physical prowess and appearance and on his ability to attract the amorous attentions of "beautiful women." "I'm not supposed to be sick," he said, "Especially down there. It makes me feel old, ugly, and unattractive. Disgusting, in fact." In describing his experiences in the hospital, he reported that the humiliation, or narcissistic blow to his self-esteem resulting from the illness and hospitalization, had so overwhelmed him ("I was hooked up to these disgusting tubes—even in my penis!" he wailed) that he had uncharacteristically not even complained about being placed into a four-bed hospital room with three homeless people who hadn't showered. "I figured they were in control, so I didn't even think about it," he said. His symptoms remitted in psychoanalysis with therapeutic exploration of the way in which the hospitalization had in fantasy felt like a castration.

Many contemporary psychoanalysts amplify upon these models in their understanding of depression, while acknowledging the prominence of impaired self-esteem regulation. Edith Jacobson emphasized the development of self and object representations in depressed patients. She noted depressed patients' disappointment with parental figures, resulting in devaluation and degradation of their images and of the self-representation, especially when a mature separation had not been achieved.

Inadequacy of Early Caregivers

Psychoanalysts have given a personal, intrapsychic face to the well-known epidemiological observation of the connection between parental (particularly maternal) depression and subsequent depression in children. Hans Kohut described depression as connected to experiences of profound emptiness in patients whose parents were unable to empathize with their early affective experiences. Such is the case, as many parents of depressed patients are themselves depressed. These patients crave compensatory relationships ("selfobject" relationships, mirroring experiences, and idealizing relationships), leaving them vulnerable to disappointment, as real relationships cannot live up to these compensatory fantasies.

Leo Stone highlighted depressed patients' refusal to accept the separateness and autonomy of the object. Again, this view emphasizes the symbiotic tie to the object to compensate for an incomplete sense of self. Stone suggested that depressed patients unconsciously coerce objects; they are disappointed in them and prone to envy and rage because of an early history of "oral frustration." Aggressive fantasies about disappointing and hurting loved ones give rise to the severe guilt with which these patients struggle.

Ms. E, a 21-year-old college student, presented with major depression and panic disorder since early adolescence. She reported hating herself, crying constantly, and feeling profoundly

hopeless in part because of the chronicity of her illness. Even at the time of presentation, she noted her sensitivity to her mother's moods. "My mother's just always depressed, and it makes me so miserable, I just don't know what to do," she said. "I always want something from her, I don't even know what, but I never get it. She always says the wrong thing, talks about how disturbed I am, stuff like that, makes me feel bad about myself." In one session, E poignantly described her childhood: "I spent a lot of time with my mother, but she was always too tired, she never wanted to do anything or play with me. I remember building a house with blankets over the coffee table and peeking out, spying on her. She was always depressed and negative, like a negative sink in the room, making it empty and sad. I could never get her to do anything." This patient experienced extreme guilt in her psychotherapy when she began to talk about her mother's depression. "I feel so bad," she sobbed. "It's like I'm saying bad things about her. And I love her so much, and I know she loves me. I feel it's so disloyal of me." Her depression remitted in psychodynamic psychotherapy as she became more aware of and better able to tolerate her feelings of rage and disappointment with her mother.

Sidney Blatt, a psychoanalytic researcher, contrasts *anaclitic* depressed patients, anxiously attached individuals who struggle with excessive dependence on others, with *introjective* depressed patients, who are compulsively self-reliant. Blatt contends that anaclitic patients suffer more from feelings of loneliness, helplessness, and weakness; introjectives from a sense of worthlessness, self-criticism, and guilt.

PSYCHOANALYTIC FORMULATIONS OF DYSTHYMIA

Psychoanalysts have largely written about major depression, but some have also addressed dysthymia. Asch noted underlying masochistic pathology in dysthymic patients, a view that has been de-emphasized by contemporary dysthymia researchers, who have underlined that chronic depression, can appear phenomenologically like masochistic character structure, simply because of its chronicity. (In other words, masochism is a redundant concept in chronic depression.) Milton Horowitz emphasizes the "pleasure of revenge" in the patient's defeating of all around him or her through failure, hopelessness, and negativity, and he regards this as more important than the experience of personal suffering for many such patients. David Milrod describes both the rewarding and the punitive aspects of superego function in response to narcissistic injury in patients with chronic dysthymia and self-pity.

Ms. F, a 20-year-old college student, came to psychotherapy complaining of chronic depression. She described herself as "dull, but difficult too." While she enjoyed school, she had trouble finding things outside of school that she enjoyed doing. She was bored, lethargic, and unhappy most of the time. Although she had friends and a boyfriend—albeit a man much less desirable than she might otherwise have been expected to find—she resented their demands on her time, often because she found herself upset when she was with them. Her relationships with her boyfriend, some of her closer friends, and her mother were characterized by intermittent "crises," events when Ms. F felt ignored, criticized, and abused by them, and she wound up locking herself in her room to cry, sometimes for days. As psychotherapy unfolded, it became clear that her boyfriend was very demanding of her time, yet rarely actually spoke to her. For example, he preferred to have her "around" while he worked on his computer for hours on end, but when he went out "to hang out" he preferred to spend time with his friends, without her. Ms. F handled this situation with a combination of resignation and resentment:

She stayed with him while he worked, then often ignored him at parties that they both attended, in order to "get back at him." In psychodynamic psychotherapy, her dysthymia and passive masochism resolved simultaneously, enabling her to stop feeling worthless in social situations and with her boyfriend and to become more assertive in all of her relationships.

SYNTHESIS: DYNAMICS OF DEPRESSION

Features common to many psychoanalytic theories of depression include feelings of exquisite narcissistic vulnerability stemming from a variety of sources, including early loss or experiences with parents perceived as traumatically unempathic, frustrating, or rejecting. A sense of helplessness or inadequacy in relation to these experiences, with accompanying fantasies of damage or castration, can contribute to this vulnerability. The resulting impairment in self-esteem regulation is common to all depressed patients, who are prone to a self-image of being unlovable, damaged, or inadequate.

Depressed patients perceive that they have either failed to live up to their ambitions or to their moral values in the ego ideal, the intrapsychic mechanism that triggers guilt in depression. Many psychoanalysts hypothesize that the resulting aggression toward a frustrating parent, or toward the self as damaged, contributes decisively to the propensity toward depression. In depressed patients, aggression is largely self-directed. Guilt (conscious or unconscious) or shame theoretically results from the patient's perceived sense of failure, with a diminished sense of self. Difficulties in self-esteem regulation contribute to a self-representation of being "bad" or shamefully out of control, aggravating the original problem in a vicious cycle.

The sine qua non of depression in this conceptualization is aggression directed toward the self-representation, which proves uncontainable and spreads to a mood state. This may arise from a withdrawal of positive, or "libidinal," supplies to the self-representation and replacement with aggression in the following ways:

1. By not living up to personal aspirations (giving rise to shame, rather than guilt);
2. By not living up to the ego ideal (precipitating guilt);
3. In an interpersonal depression, as described by Freud, in which a symbiotic bond to an ambivalent object tie is shattered. From a psychoanalytic perspective, this type of depression is associated with more primitive pathology, and these patients are more likely to be in a borderline or psychotic spectrum, because of the nature of the primitive, symbiotic interpersonal bond.

These different pathways trigger a common self-assessment of being worthless, bad, inadequate, or unlovable. The resulting hostility directed toward the self is a core feature of depression.

Retroflected hostility has been pinpointed in the historical

psychoanalytic literature as also contributing to borderline and obsessional pathology. Indeed, many patients with these underlying character structures also experience major depression. From a psychoanalytic perspective, the essential feature of all depressions is the self-directed aggression, which becomes a mood state. Difficulty with self-esteem regulation, which is not an essential component of other syndromes, is always prominent in depression, coupled with aggression turned inward that exacerbates guilt and shame.

PSYCHOANALYTIC FORMULATIONS OF MANIA

From a psychoanalytic perspective, the clinical presentation of mania arises as a result of a global, massive regression that affects all three psychic structures: The id, the ego, and the superego. The regression leads to a primitive mental state in which the pleasure principle is reinstated. In *Group Psychology and Analysis of the Ego*, Freud described mania as a fusion of the ego with its superego. Less cryptically, psychoanalysts have highlighted a common organizing fantasy in manic patients of a fantasy incorporation, or mystical union, of the patient with someone of great power, often an aristocrat, or God, as in the story of St. Theresa's mystical union with God. Such organizing fantasies, couched in both sexual and "oral" terms, magically impart a sense of omnipotence and specialness to the patient, highlighting one aspect to the common phenomenology of mania. Bertam Lewin noted that in mania early relationships with both parents that had become desexualized during the process of superego formation during latency become resexualized during the manic episode and that in fantasy the manic patient is identified with both partners (male and female) during the sexual act. The states of mania, and of hypomania, involve massive utilization of the defense mechanism of denial, in which sad and negative aspects of reality are entirely ignored. One author hypothesized that the manic episode was precipitated by the need to control the memory and experience of intolerable pain. In this way, mania could be considered a defensive reaction, different from depression, in response to pain and distress.

Ms. G, a 42-year-old housewife and mother of a 4-year-old boy, developed symptoms of hypomania and later of frank mania without psychosis, when her only son was diagnosed with acute lymphocytic leukemia. A profoundly religious woman who had experienced 10 years of difficulty with conception, Ms. G was a devoted mother. She reported that she was usually rather "down." Prior to her son's illness, she used to joke that she had become pregnant with him by divine intervention. When her son was diagnosed and subsequently hospitalized, he required painful medical tests and emergency chemotherapy, which made him very ill. The doctors regularly barraged Ms. G with bad news about his prognosis during the first few weeks of his illness.

Ms. G was ever present with her son at the hospital, never sleeping, always caring for him, yet the pediatricians noted that as the child became more debilitated and the prognosis more grim, she seemed to bubble over with renewed cheerfulness, good humor, and high spirits. She could not seem to stop herself from cracking jokes to the hospital staff during her son's painful

procedures, and as the jokes became louder and more inappropriate, the staff grew more concerned. During her subsequent psychiatric consultation (requested by the pediatric staff), Ms. G reported that her current "happiness and optimism" were justified by her sense of "oneness" with Mary, the mother of God. "We are together now, she and I, and she has become a part of me. We have a special relationship," she winked. Despite these statements, Ms. G was not psychotic, and said that she was "speaking metaphorically, of course, only as a good Catholic would." Her mania resolved when her son achieved remission and was discharged from the hospital.

COGNITIVE THEORY OF DEPRESSION

Learning theory long has been a branch of behavioral psychology. Aaron Beck, finding that psychoanalytic theory did not sufficiently explain dreams of depressed patients, developed a theory of depression based on educating the patient about his or her negative thinking, or cognitions. Beck and colleagues then successfully tested CBT, a treatment built on this theory, in clinical trials. The cognitive model is based on the recognition that people are not objective; rather, an individual's idiosyncratic perception of events affects his or her emotions and behaviors. Depressed individuals perceive reality in subjectively depressed ways. Elaborate discussions of cognitive theory exist, and cognitive explanations have been extended from their original depressive origins to a range of psychopathology. Brad Alford and Beck argue that cognitive theory provides a comprehensive and coherent paradigm for psychopathology.

Beck's initial observations about major depression have a salience and simplicity worth repeating. He noted that depressed patients tend to have characteristically skewed and negative thoughts about (1) themselves, (2) their environment, and (3) the future, a cluster he termed the *cognitive triad*. Indeed, depressed individuals frequently report negative thinking about themselves: "I'm a loser," "Everything I do goes wrong," "I'm weak and defective." The environment appears hostile and overwhelming: "Even if I felt capable—which I don't—there's no way I could cope with what I have to do"; "My friends will react badly if I try to speak up"; "She will reject me." Finally, not only do things look grim in the present, but there is no prospect of relief in the future: "It's never going to get any better." These three aspects of negative perspective converge to provide a convincingly, despairingly bleak and mood congruent view of the world. This outlook helps to explain why depressed patients see no way out of their misery and contemplate suicide.

Cognitive theory has explored the form as well as the content of thinking characteristic of depressed patients. Not only are cognitions skewed to the negative and pessimistic, but particular types of distortions occur. Depressed individuals tend to engage in "all or nothing," dichotomous thinking: If things aren't entirely one way, then they must be the opposite. Depressed individuals make arbitrary (negative) inferences about events, selectively abstract negative details out of context,

overgeneralize (concluding negative rules from single instances), magnify (the negative) and minimize (the positive), and take personally events that may not be directly about them. Depressed patients "catastrophize," leaping from one imagined worst case scenario to the next in disastrous cascade:

> He will take what I said as an insult, and will never speak to me again, and will tell all our friends. Then no one will like me, I'll be all alone, and my social life will be ruined forever.

They selectively focus on negative evidence, failures, and setbacks that confirm their theories of defectiveness, while ignoring or discounting the successes that they have as "flukes."

> J, a 19-year-old college sophomore, made two mistakes on a test. Although his overall grade was high and the mistakes themselves appeared minor to the therapist, J took them as proof that he was defective and deficient. His perfect score on the next test, by contrast, he discounted as "luck ... the teacher was being easy on me," but not as any evidence of his personal competence. If, while depressed, you consider yourself a loser, then you pay most attention to outcomes that fit that schema.

Beck made the important point that the depressed outlook is not objective but irrational. "Automatic thoughts" of characteristically self-deprecatory and hopeless nature pop involuntarily into the patient's head. Because these thoughts are mood congruent, depressed individuals find them believable and tend not to question them. The negative thinking characteristic of depressed patients is damaging in two respects. First, it is painful and depressing: Many patients are barraged with negative ideas about themselves and their situation, in effect a stream of insults. Second, these thoughts tend to inhibit action: If you know you are incompetent and going to fail, then you will also have thoughts like, "Why bother?" Depressive thinking patterns not only hurt, they paralyze individuals into inaction, which then leaves them more time to ruminate on their inactivity ("I'm not doing anything with my life") and to suffer self-criticism. The deterioration of productive behavior due to negative thinking will then lead to more things going wrong in the patient's life, providing more examples of the patient's incompetence and reinforcing the negative thinking.

The case examples above provide good examples of self-critical thinking, which in cognitive theory would be understood as the irrational, automatic thoughts of a depressed patient: Ms. A felt "pointless" and "boring," Mr. D "like a big, fat failure," "old, ugly, and unattractive ... disgusting"; Ms. F as "dull, but difficult too."

Immediate and specific negative cognitions ("He doesn't like me") fit into larger, more basic and stable patterns of self-conception called "schemas" or core beliefs. For depressed patients, these tend to be global, negative, all-or-nothing assumptions based on early childhood experience: "If I do something imperfectly, it's (and I'm) worthless"; "If I'm not in total

control, I'm completely helpless." Early cognitive theory presented such negative thinking as cause rather than effect of depression. Although cognitions are clearly an important aspect of depression, it seems naive at this point to expect that this is the single etiology of a complex and multidetermined syndrome that has both genetic and environmental components. Hence, over time cognitive theory has tended to back away from that etiological stance.

A general cognitive explanation of an major depressive episode is that a vulnerable individual, perhaps predisposed both by biology and by negative schemas based on early childhood, experiences a current life situation that evokes "automatic" thoughts. For example, a relatively minor setback at work might activate underlying schemas by eliciting the thoughts, "I've done a bad job," and thereby "I'm a terrible worker," "I've let the company and my co-workers down," "My boss hates me," and therefore "I'll never succeed," "I'm a failure." Or a disagreement in a relationship might rouse ideas such as, "K doesn't like me," evoking and seemingly confirming the more general core beliefs: "I'm a terrible person," "I'm a lousy parent," and furthermore "No one could possibly love me." The onset of these thoughts and accompanying depressive symptoms further compromises functioning, reinforcing the negative outlook. As mood darkens and the negative thoughts become ever more credible and pervasive, the patient gives up activities ("Why bother?") and feels still more helpless and hopeless.

Cognitive therapy, the treatment that follows from this approach, includes a Socratic discussion and evaluation of the patient's thoughts, weighing the evidence supporting and contradicting such thoughts. The patient actively tests hypotheses based on automatic thoughts ("I'll fail at whatever I do") by attempting various selected behaviors as homework. As the patient learns to recognize the irrational nature of depressive thinking, he or she can challenge rather than simply believe it and can begin to extinguish such thinking, replacing automatic irrational thoughts with rational responses. Outcome research repeatedly has shown that this approach is efficacious in treating mood disorders and other psychiatric syndromes.

Dysthymic and Bipolar Disorders

Cognitive theory has not been particularly altered for dysthymic disorder. Rather, chronicity of depression would simply seem to ingrain negative core beliefs more deeply. Cognitive Behavioral Analysis System of Psychotherapy (CBASP), an eclectic treatment that James McCullough developed specifically to treat chronic depression, postulates that chronically depressed patients fail to behave appropriately in interpersonal encounters. CBASP combines elements of cognitive, behavioral, psychodynamic, and interpersonal therapies in focusing on the patient's behaviors and thinking in interpersonal encounters, including those with

the therapist.

Cognitive behavioral treatments have been developed not only for unipolar depression but also as an adjunct to medication for bipolar disorder. Mania may be viewed as the mirror of depression, wherein overvalued positive ideas ("I'm special" or "I'm the greatest") automatically arise, again distorting reality, but this time in the opposing direction. Thinking itself accelerates in mania. By and large, cognitive theory has had less to say about manic than depressed states, and adaptations of cognitive therapies for mania are more practical than theoretical. They focus on psychoeducation about the illness, behavioral stabilization such as maintaining a regular sleep schedule, and emphasize the importance of adherence to pharmacotherapy. Following these principles, Francisco Colom and Eduard Vieta have developed an empirically tested psychoeducational treatment of bipolar patients.

Aspects of cognitive theory have been tested for unipolar depression, with sometimes contradictory results. CBT, which directly addresses depressive negative cognitions and helps patients to weigh their evidence rather than simply believe them, has repeatedly demonstrated efficacy as a treatment of major depression and does improve negative cognitions. Yet other antidepressant treatments relieve depression and improve cognitions without specifically addressing cognitions, and behavioral activation—a behavioral approach without the cognitive component—has been shown to be at least as efficacious as CBT. Efficacious treatments of depression may spread from a particular theorized target to have generalized benefits.

Comparison to Psychodynamic Theory

It will be seen that cognitive theory does not contradict and to some degree overlaps with psychoanalytic theory. The treatments emerging from these theories differ greatly, however, in large part due to their differing definitions of the role of the therapist, handling of the therapeutic relationship, and approach to the meaning of that relationship: in psychoanalytic terms, the transference. Both theories focus on intrapsychic processes and on self-critical thinking, but cognitive theory emphasizes the distinction between rational and irrational thinking and their connection to mood rather than to unconscious fantasy.

INTERPERSONAL THEORY OF DEPRESSION

Interpersonal theory dates back to the era after World War II, when it arose as a heretical response to the more intrapsychic emphasis of psychoanalysis. Psychoanalytic theory emphasized the importance of early life experience, and many therapists at that time saw the patient's psychic structure as essentially formed by the end of adolescence. Psychiatrists such as Adolf Meyer, Harry Stack Sullivan, Erich Fromm, and Frieda

Fromm-Reichmann challenged then current theory by emphasizing the influence of the real impact of current life events on their patients' psychopathology, focusing on environmental and interpersonal encounters rather than underlying intrapsychic drives and structures.

Sullivan coined the term "interpersonal" as a rubric for considering current life experience. He scrutinized communications in the social field, a more "external" outlook than traditional psychoanalysis. The interpersonal theorists worked mainly with inpatients diagnosed with schizophrenia in a prepharmacological era. Although their work stirred great controversy at the time and a Sullivanian school distinct from the psychoanalysis of drives and ego psychology still exists, Sullivan was trained in psychoanalysis and did not entirely disagree with his forebears. Over time, the rift between Sullivanians and other psychoanalysts has narrowed.

The consideration of current interpersonal factors gained currency over succeeding decades, and it is now mainstream clinical thinking that current life events and interpersonal functioning affect and are affected by psychopathology. A school of interpersonal psychoanalysis—not particularly focused on mood disorders—arose. Psychoanalytically trained therapists like Silvano Arieti and Jules Bemporad emphasized interpersonal factors in the treatment of depressed patients. Jack Anchin and Donald Kiesler have reviewed other interpersonally based psychotherapies. None of these theories has received empirical testing.

Researchers did develop a host of related data about interpersonal issues associated with depression. For example, research showed that interpersonal support protects an individual against depression: Having a confidant to talk to reduces the risk of developing a depressive episode. Major life stressors, including the death of a significant other, struggles in important relationships, and upheavals such as a change in marital status, housing, job status, or physical health have been shown to increase the risk of depressive episodes in vulnerable individuals. Moreover, the onset of depressive episodes leads to deterioration in relationships and social functioning.

John Bowlby postulated that people have an evolutionarily determined, instinctual drive to form emotional attachments. Animal evidence now supports this theory. This basic component of human nature ensures infant survival: Children need to have parents nearby or available for feeding and protection. As children develop, they begin to explore their environment, gradually moving out from the "secure base" of their attachment figure. Disruptions in this early caregiving connection may lead to vulnerability of attachment style. For example, loss of one's mother in the first decade of life has been shown to be a risk factor for subsequent depression. Children with insecure childhood attachments may not learn to ask for help from others. When such vulnerable individuals face stressors or feel an absence or inadequacy of interpersonal support during times of stress, they may be

helpless to respond effectively and prone to developing symptoms. Furthermore, individuals with insecure attachment styles may have difficulty in developing comfortable relationships on which they can rely for support in times of need.

In the 1970s, when Gerald L. Klerman, Myrna M. Weissman, and their colleagues were conducting a randomized controlled trial of outpatients with major depressive episodes, they recognized that many such patients received psychotherapy in community treatment. They sought accordingly to add a psychotherapy to their trial but realized that it was unclear then (as now) of what such community psychotherapy consisted. They then developed IPT as a manualized, time-limited treatment of outpatients with major depressive disorder based on interpersonal and attachment theories as well as empirical evidence of the psychosocial nature of depression. IPT is thus one specific therapeutic application of more general interpersonal theory. IPT demonstrated efficacy for major depression in this and in subsequent outcome trials.

In simplest terms, interpersonal theory as applied to IPT can be understood as a link between mood and events. For all people, upsetting external events evoke a sad or demoralized mood: In lay terms, they are "depressing." For biologically or environmentally predisposed individuals, however, a sufficiently disturbing life event can trigger an episode of major depression. Examples of such life events are the death of a significant other (complicated bereavement), a problematic relationship (role dispute), or other major life change (role transition). Once a depressive episode starts, its symptoms compromise functioning, producing more negative life events in a vicious downward cycle. This formulation seems straightforward, even commonsensical, but depressed patients have a peculiar amnesia for external events and tend to blame themselves for how they feel or to see the depressed state as who they are. It can be helpful clinically to remind them that they are ill, not defective, and that outside events may have contributed to their distress.

IPT therapists do not propose this as an etiological theory of depression, but as a pragmatic one: The depressive mood episode can be linked either to a precipitating life event or to consequent life events that become the focus for the treatment. The IPT therapist defines major depression as a medical illness—a treatable medical problem that is not the patient's fault—and links it to an interpersonal focus such as a role dispute. The therapeutic contract for the patient is to solve the interpersonal focus within a time-limited period. Solving the interpersonal problem is at once a realistic relief (e.g., improving a marriage), relieves depressive symptoms, and builds interpersonal skills that may hopefully protect against future interpersonal triggers and depressive episodes. Typical areas of interpersonal skill building are issues such as self-assertion, confrontation, effective expression of anger, and the taking of social risks.

Clinicians armed with differing theories approach the same material in different ways. IPT seeks a life event or interpersonal situation as a plausible and pragmatic fiction on which to focus a time-limited treatment in which the patient can work on interpersonal skills. For example, the case of Ms. A might be conceptualized as major depression arising in the context of a role dispute with her boyfriend, B. Had she been treated in IPT, the focus might have been on recognition and appropriate expression of her own anger as part of learning to renegotiate that relationship. Alternatively, or additionally, clinical judgment might have defined the breakup of that relationship as a role transition that Ms. A needed to mourn and accept in order to move on to better relationships or activities. Ms. C's difficulty in tolerating her acceptance to graduate school might similarly be considered a role transition, as might Mr. D's bout of renal stones. Medical illnesses, even if transient, frequently provoke role transitions by shifting a patient's perception of his life trajectory; Mr. D may have interpreted his hospitalization as evidence of his fading potency, aging, and mortality. Ms. E's situation might conceivably be defined either as a role transition—adjusting to college life away from her mother—or as a role dispute with her mother.

Preliminary evidence supports the theory underlying IPT as a treatment. Patients in two IPT treatment trials were asked to report the degree to which they had resolved the interpersonal crisis (e.g., role transition) on which their treatment had focused. Subjects' report of making changes in this interpersonal focal area correlated with symptomatic improvement during the trial.

Dysthymic Disorder

An adaptation of IPT has been developed for dysthymic disorder, but differences in the treatment are largely those of technique. Patients with chronic depression tend to be socially isolated (Ms. F above is slightly unusual in having a boyfriend, although her choice of boyfriend was in keeping with the disorder) and invariably have interpersonal difficulties. If the syndrome began early in life, as is often the case, then patients may never have developed important communication skills. They tend to be passive and unassertive, to experience self-assertion as "selfish" rather than a healthy and sometimes necessary negotiation of one's own goals in life, and to view anger as a "bad" feeling rather than a useful interpersonal signal that someone is bothering them. Even if they have had such capabilities in the past, their interpersonal functioning will have eroded under the pounding discouragement of chronic depression. Chronicity breeds resignation and entrenches maladaptive interpersonal patterns.

Bipolar Disorder

In all of the theoretical approaches under discussion, mania has received less attention than unipolar depression. There is no real adaptation of interpersonal theory for mania. Clinically, it is apparent that manic patients lose social judgment and become overinvolved and labile in relationships with others. Ellen Frank and colleagues have adapted IPT for bipolar patients as an adjunctive treatment to pharmacotherapy by combining it with a behavioral strategy. Interpersonal social rhythms therapy (IPSRT) is a hybrid of IPT and social zeitgeber theory, which stresses the importance of maintaining diurnal structure through regular patterning of circadian events, or social rhythms. In particular, this involves ensuring a regular bedtime to prevent the loss of sleep that often triggers a manic episode. As with CBT for bipolar disorder, this represents more of a clinically reasonable, pragmatic extension of an already established treatment than a theoretical evolution. This approach, described in a manual, has shown benefit as an adjunct to pharmacotherapy in a randomized controlled trial for bipolar patients.

Future Directions

Psychoanalytic, cognitive, and interpersonal theories all provide insights into mood disorders. None is received truth, none can be considered fully tested in empirical studies, and all are surely incomplete. Nor has this section covered all theories of mood disorder: Biological, behavioral, feminist, and many other theories have been proposed for depression. Psychoanalytic, cognitive, and interpersonal theories each capture important aspects of complex syndromes. In the hands of skilled clinicians, therapists and patients may find each approach useful in conceptualizing the patient's status and organizing his or her treatment.

Future research may elaborate more comprehensive applications of these theories to bipolar disorder. And theory alone does not suffice to train a clinician: clinical experience and supervision, as well as personal therapy, help to expand clinical understanding and ability.

References

Abraham K. Notes on the psycho-analytical investigation and treatment of manic-depressive insanity and allied conditions (1911). In: Bryan D, Jones E, eds. *Selected Papers on Psychoanalysis*. London: Hogarth Press; 1927:137.

Anchin JC, Kiesler DJ, eds. *Handbook of Interpersonal Psychotherapy*. New York: Pergamon Press; 1982.

Arieti S, Bemporad J. *Severe and Mild Depressions*. New York: Basic Books; 1978.

Basco M, Rush AJ. *Cognitive-Behavioral Therapy for Bipolar Disorder*. New York: Guilford Press; 1996.

Beach SRH, Kogan SM, Brody GH, Chen YF, Lei MK, Murry VM. Change in caregiver depression as a function of the Strong African American Families Program. *J Fam Psychol*.

2008;22(2):241.

*Beck AT, Rush AJ, Shaw BF, Emery G. *Cognitive Therapy of Depression*. New York: Guilford Press; 1979.

Bowlby J. *Attachment and Loss*. New York: Basic Books; 1973.

Colom F, Vieta E. *Psychoeducational Manual for Bipolar Disorder*. New York: Cambridge University Press; 2006.

Copeland LA, Zeber JE, Salloum IM, Pincus HA, Fine MJ, Kilbourne AM. Treatment adherence and illness insight in veterans with bipolar disorder. *J Nerv Ment Dis*. 2008;196:16.

DeRubeis RJ, Hollson SD, Amsterdam JD, Shelton RC, Young RC. Cognitive therapy vs. medications in the treatment of moderate to severe depression. *Arch Gen Psychiatry*. 2005;62:409.

Eells TD, ed. *Handbook of Psychotherapy Case Formulation*. 2nd ed. New York: Guilford Press; 2007.

Frank E. *Treating Bipolar Disorder: A Clinician's Guide to Interpersonal and Social Rhythm Therapy*. New York: Guilford Press; 2005.

*Freud S. Mourning and Melancholia. In: Strachey A, eds. *Standard Edition of the Complete Psychological Works of Sigmund Freud*. Vol. 14. 1915:237–258.

Freud S. Beyond the Pleasure Principle. In: Strachey A, eds. *Standard Edition of the Complete Psychological Works of Sigmund Freud*. Vol. 18. 1920:3–67.

Freud S. The Ego and The Id. In: Strachey A, eds. *Standard Edition of the Complete Psychological Works of Sigmund Freud*. Vol. 19. 1923:3–66.

*Jacobson E. *Depression: Comparative Studies of Normal, Neurotic, and Psychotic Conditions*. Madison, CT: International Universities Press; 1971.

Klein M. Mourning and its relation to manic-depressive states. *Int J Psychoanal*. 1940;21:125.

*Klerman GL, Weissman MM, Rounsaville BJ, Chevron ES. *Interpersonal Psychotherapy of Depression*. New York: Basic Books; 1984.

Lewin BD. Comments on hypomanic and related states (1941). In: Arlow JA, ed. *Selected Works of Bertram D. Lewin*. Richmond, VA: William Byrd Press; 1973:78.

*Lipsitz JD, Markowitz JC. Mechanisms of change in interpersonal psychotherapy. *Clin Psychol Rev*. 2013;33:1134–1147.

Lorenzo-Luaces L, German RE, DeRubeis RJ. It's complicated: the relation between cognitive change procedures, cognitive change, and symptom change in cognitive therapy for depression. *Clin Psychol Rev*. 2014;41:3–15.

Markowitz JC. *Interpersonal Psychotherapy for Dysthymic Disorder*. Washington, DC: American Psychiatric Press; 1998.

McCullough JP Jr. *Treatment for Chronic Depression: Cognitive Behavioral Analysis System of Psychotherapy*. New York: Guilford Press; 2000.

*Milrod D. Self-pity, self-comforting, and the superego. *Psychoanal Study Child*. 1972;27:505.

Rado S. The problem of melancholia. *Int J Psychoanal*. 1928;9:420.

Rush AJ, Beck AT, Kovacs M, Weissenburger J, Hollon SD. Comparison of the effects of cognitive therapy and pharmacotherapy on hopelessness and self-concept. *Am J Psychiatry*. 1982;139:862.

Sandler J, Joffe WG. Notes on childhood depression. *Int J Psychoanal*. 1965;46:88.

Sullivan HS. *The Interpersonal Theory of Psychiatry*. New York: WW Norton; 1953.

Weissman MM, Markowitz JC, Klerman GL. *Clinicians' Quick Guide to Interpersonal Psychotherapy*. New York: Oxford University Press; 2007.

Zeber JE, Copeland LA, Good CB, Fine MJ, Bauer MS, Kilbourne AM. Therapeutic alliance perceptions and medication adherence in patients with bipolar disorder. *J Affect Disord*. 2008;107(1–3):53.

▲ 13.6 Mood Disorders: Suicidal Behavior

ZOLTáN RIHMER, M.D., PH.D., D.SC., AND MAURIZIO POMPILI, M.D., PH.D.

Suicidal behavior is one of the most alarming signs not only in psychiatry but also in everyday life, and it is among the most frequent sources of contact with medical services, particularly with psychiatric services. Suicidal behavior (completed suicide and suicide attempt) as well as suicidal ideation or planning are very rare in the absence of current major mental disorders; they are not the linear or direct consequences of them. Although suicidal behavior is very complex, multicausal phenomenon involving several medical–biologic, psychosocial, demographic, and cultural components, a history of untreated major mood disorders (particularly in the presence of previous suicide attempt) constitutes the most important risk factor. However, because the majority of patients with mood disorders never complete (and more than half of them never attempt) suicide, other familial–genetic disorders, personality, psychosocial, and demographic risk factors also play a significant contributory role. The exploration of the role of these factors help to provide a better understanding of what makes the minority of patients with mood disorders suicidal, while the majority of patients are not suicidal. In addition to suicidal behavior, major mood disorders (both unipolar major depression and bipolar disorder) are also associated with a considerable amount of illness-related health and economic problems, such as loss of productivity, secondary substance-use disorders, family breakdown, and increased somatic morbidity and mortality. Given the high lifetime prevalence of unipolar major depressive and bipolar disorders, they are among the most frequent as well the most potentially life-threatening psychiatric illnesses. Despite the great clinical and public health significance of major mood disorders, people with these disorders are still under-referred, under-diagnosed and under-treated. This is particularly true for those who die by suicide and for those who attempt suicide and receive medical attention. As the act of killing oneself is considered to be a major cause of years of life lost, its prevention is receiving more and more attention nowadays.

RELATIONSHIP BETWEEN MOOD DISORDERS AND SUICIDAL BEHAVIOR

Psychological autopsy studies from several different countries consistently show that around 90 percent of consecutive suicide victims have at least one but frequently more (mostly untreated) major psychiatric disorders at the time of their death, and the rate of major mood disorders is between one-half and two-thirds. Comorbid anxiety and personality disorders are also frequently present, but they are not the main diagnosis, as these patients experience suicidal behavior mainly in the case of a coexisting major depressive episode. Although the ratio of attempted to completed suicide in the general population is about 15–40:1, it is much lower (5–10:1) among patients with major mood disorders, suggesting that these patients use more lethal (or more violent) suicide methods. Patients with mood disorders, particularly bipolar patients, often present a low ratio of suicide attempts to completed suicides, which points to the relatively high lethality of suicide attempts in bipolar disorder. Such findings indicate that bipolar patients have a particularly high level of intent to die and a high lethality of the methods used. Among all individuals hospitalized due to severe suicide attempts in Finland from 1997 to 2003 (N = 23,321), mood disorders (N = 5,164) raised the risk of completed suicide by 72 percent and of suicide attempt by 59 percent during the mean follow-up time of 3.6 years. As in other studies, the risk was very high just immediately after discharge from hospital.

Completed suicide and suicide attempts are two different but greatly overlapping phenomena: around half of those who complete suicide have attempted suicide at least once previously, and the first attempt (even if the method used is nonviolent or nonlethal) significantly increases the risk of future completed suicide. This is partly because of the fact that those who repeatedly attempt suicide frequently switch their method from nonviolent to violent or from nonlethal to lethal. Between 10 and 18 percent of adults in the general population worldwide report lifetime suicidal ideation, and 3 to 5 percent have made at least one suicide attempt in their lifetime. Suicidal ideation, suicide attempt, and completed suicide are three different but greatly overlapping features. Prior suicide attempt and current major depression are the two strongest predictors of future suicide, and the vast majority of suicide attempters or completers come from a population of people with current suicidal ideation, particularly in the presence of an untreated major depressive episode.

The standardized mortality ratio of suicide death in (unipolar and bipolar) patients with major mood disorders compared to that in the general population has been reported to be 10- to 30-fold. While suicide is around the 10th most common cause of mortality in the general population in well-developed countries, it is among the most common leading causes of death among patients with unipolar major depression and bipolar disorder, and the life expectancy of these patients is decreased by about 10

to 12 years. Studies published many decades ago estimated that 15 to 19 percent of severe (mostly hospitalized) patients with unipolar and bipolar major depressive episodes would die by suicide. However, most recent studies—maybe reflecting the more widespread and effective treatments—have found that this figure is less than 10 percent, which is, of course, still very high. A study from the United States found that there was a severity-dependent hierarchy in suicide risk among patients with major depressive disorders. Whereas the lifetime risk of completed suicide in the general U.S. population was 0.5 percent, the same figures in nonsuicidal major depressive outpatients, major depressive inpatients hospitalized without specification of suicidality, and severe major depressives hospitalized for suicidality were 2.2, 4.0, and 8.6 percent, respectively. In a national long-term follow-up study on suicide risk in Denmark on all subjects born between 1955 and 1991 who had been treated in a psychiatric hospital in Denmark up to the year 2006, or as outpatients after 1995, it was found that the absolute lifetime risks of completed suicide in patients with diagnosis of bipolar disorders were 7.8 percent for males and 4.8 percent for females, while the same figures for patients with diagnosis of unipolar depression were 6.7 and 3.8 percent, respectively. It should be noted, however, that because the diagnostic switch between unipolar depression and bipolar disorder was not taken into account during this long-term follow-up study, the real lifetime risk of completed suicide should be higher in bipolar disorder and lower in unipolar disorder.

Other long-term follow-up studies also show that the rate of completed suicide is higher in bipolar disorder (types I and II combined) than in unipolar depression, and similarly, population-based epidemiological studies also found a substantially higher rate of suicide attempts in bipolar I and II disorders than in patients with unipolar major depressive disorder: more than half of bipolar patients and around one third of unipolar depressives report a lifetime history of one or more suicide attempts, and suicidal ideation, a major precursor of suicidal behavior in patients with mood disorders, is even more common. The majority of the studies show that out of the three different clinical manifestations of major mood disorders (unipolar depression, bipolar I disorder, and bipolar II disorder), bipolar patients in general, and bipolar II subjects in particular, carry the highest risk of both attempted and committed suicide. Considering only 10 studies, published between 1976 and 2004, in which unipolar, bipolar I, and bipolar II patients were analyzed separately, and summarizing the data, it was found that the rate of previous suicide attempts was the lowest in unipolar major depression patients (13 percent), highest in bipolar II patients (33 percent), and intermediate in bipolar I patients (28 percent). Most recent studies found that completed and attempted suicide is more frequent in bipolar II patients than in bipolar I patients, but meta-analyses found this difference statistically insignificant. However, in everyday

clinical practice, bipolar II patients are frequently overlooked. Clinicians often acknowledge the fact that the risk of suicide is higher among bipolar II patients as they often have a cyclothymic affective temperament, which predisposes them to suicide risk, and they often have rapid mood shifts, which are found in the clinical history of such patients.

The risk of completed and attempted suicide in bipolar disorder is more pronounced at younger ages, especially in the first few years after onset. However, bipolar patients with past suicide attempts have earlier age at onset than those without prior suicidality. Since the correct diagnosis of bipolar disorders usually comes at an average of 8 to 10 years after the onset of the illness, and delays in the correct diagnosis as well as delays in the initiation of mood stabilizers at illness onset confer an increased risk of suicidal behavior, these findings strongly suggest that early recognition of the bipolar nature of depressive disorders and adequate short-term and long-term treatment is crucial for preventing suicide.

The first suicide attempt significantly increases the risk of future suicide attempts or completed suicide, and about 20 percent of those who attempt suicide eventually will die by their own hand. However, up to half of suicide victims have made at least one prior suicide attempt, indicating that about 50 percent of suicide victims die by their first suicidal act. Therefore, to detect the risk of suicide in patients with mood disorders and to plan prevention strategies, as early as possible, even before the first suicidal act, is crucial in suicide prevention. This chapter summarizes the most relevant clinically detectable suicide risk and protective factors in mood disorders and briefly highlights the most effective preventive strategies. It will focus on completed suicide (an act of deliberately taking one's life) and suicide attempt (a self-injurious act with intent to end one's life), both of which are commonly referred to as "suicidal behavior," but it will not cover deliberate self-harm (defined as an intentional injuring of one's body without apparent suicidal intent).

CLINICALLY DETECTABLE SUICIDE RISK AND PROTECTIVE FACTORS IN PATIENTS WITH MOOD DISORDERS

Because suicide is a rare event in the community, its precise prediction in individual cases is very difficult but not impossible. However, suicidal behavior (completed suicide and suicide attempt) is quite frequent among patients with mood disorders who contact different levels of the health care some weeks or months before their death or attempt, which emphasizes the potentially key role of health care workers in suicide prevention. The scientific literature identifies several clinical, biological, psychosocial, and demographic risk and protective factors that have been found to be statistically significant with suicidal behavior.

Suicide risk factors are classified as short-term versus long-term, proximal versus distal, changeable versus unchangeable, psychiatric–

medical versus sociodemographic, and clinical versus biological. As for biological risk factors, a decreased (dysregulated) central serotonergic system in suicide victims and suicide attempters (particularly in those with violent method), regardless of the psychiatric diagnosis, is a very strong correlate of suicidal behavior and is one of the most consistent neurobiological findings in the field of psychiatry. Some studies, however, suggest that alterations in other neurotransmitter (norepinephrine, dopamine, GABA, and *N*-methyl-D-aspartate [NMDA]) systems and in cholesterol as well as omega-6 and omega-3 metabolism might also be important in this respect. Studies consistently show that hyperactivity of the hypothalamic–pituitary–adrenal axis (as reflected in the abnormal dexamethasone suppression test), is a robust predictor of future completed suicide, primarily in patients with major depressive episodes. As decreased central serotonin metabolism is more strongly related to impulsive–aggressive personality characteristics than to any specific psychiatric disorder, these trait-related features increase the likelihood of suicidal behavior, particularly in the simultaneous presence of an acute episode of major psychiatric disorder and adverse psychosocial situations, including lack of social support. Despite the fact that current studies on the biology of suicidal behavior are very promising, in the everyday clinical practice (particularly in the case of urgent need of action), they have just very limited practical value. Therefore, clinically explorable risk and protective factors—as will be discussed below—are the most important aids in this complicated and highly responsible task.

Specific Clinical Characteristics of the Current Mood Episode

If patients with major mood disorders complete or attempt suicide, in the majority of the cases, they so during major depressive or mixed episodes but practically never during euphoric mania or euthymia (clinical recovery), indicating that suicidal behavior in patients with major mood disorders is both state and severity dependent. Severe major depressive episodes and the simultaneous presence of major depressive and manic syndrome (DSM-IV mixed affective episode and DSM-5 major depressive episode or manic episode “with mixed features”) are the most important time-related risk periods of suicidal behavior in patients with mood disorders. Both the DSM-IV mixed affective episode and DSM-5 mania or major depressive episode “with mixed features” are characterized not only by increased suicide risk but also by earlier age of onset, higher frequency of manic and depressive episodes, higher rate of rapid cycling course, as well as substance use and anxiety disorder comorbidity, all of which increase the risk of suicidal behavior.

According to both clinical and research observations, depression of suicide victims differs qualitatively from that of living depressed controls.

Suicidal depressives often experience more severe forms of depression, with features highly related to suicide risk, such as insomnia, hopelessness, anxiety, agitation, weight or appetite loss, feelings of worthlessness or inappropriate guilt, and thoughts of death or suicidal ideation. Comorbidity with substance (including alcohol) use disorder(s) is also often observed in depressed suicide victims. Patients with mood disorders who exhibit volatile and erratic moods associated with dysphoria and agitation or who present the classic mixed states have a particularly higher risk of suicide. Predisposing factors interact with precipitating stressors, but deaths, separations, and other major losses, scandals, or imprisonment rarely precipitate suicide in the absence of a psychiatric disorder. Current mood states are critical determinants of suicidal risk in patients with mood disorders, particularly depressive and dysphoric-irritable mixed states (DSM-5 major depressive or manic episode with mixed features) that together are present in at least 75 percent of suicides in bipolar patients, whereas suicide is extremely rare in pure mania, hypomania, and after clinical recovery.

Subthreshold lifetime or current (hypo)manic symptoms during DSM-IV unipolar major depressive disorder (subthreshold bipolar depression and mixed depression that shows a great overlap with the clinical category of agitated depression) or the DSM-5 category of major depressive or manic episode with mixed features also increase the risk of suicidal behavior. Analyzing the US National Comorbidity Survey-Replication database also showed that the history of lifetime suicide attempts was the highest in bipolar I disorder (66 percent), lower in bipolar II disorder (50 percent), and lowest in unipolar major depression without subthreshold bipolarity (30 percent). However, in patients with unipolar major depression with lifetime or current (intradepressive) subthreshold (hypo)manic symptoms (major depressive episode "with mixed features" in DSM-5), the rate of patients with prior a suicide attempt was intermediate between pure unipolar major depression and bipolar II disorder (41 percent). Studies also show that the psychotic nature of major mood episodes also increases the risk of suicide, particularly in patients with bipolar depression. In contrast to a major depressive episode, minor depression and pure dysthymic disorder (dysthymia without "comorbid" major depression) do not markedly increase the risk of suicide, but many patients with these milder forms of mood disorders develop major depressive episodes, and consequently an increased risk of suicide, during the long-term course.

Comorbid Psychiatric, Personality, and Medical Disorders

Patients with mood disorders with comorbid anxiety disorders, substance use disorders, cigarette smoking, personality disorders (mainly borderline-type), and serious, disabling, painful and fatal medical illness are also at increased risk of attempted and committed suicide. Young persons with

bipolar disorder and ADHD have an increased likelihood of suicidal behavior compared with young patients who have only bipolar disorder. Besides the well-known role of chronic alcoholism, acute alcohol intake is also an important risk factor for attempted and completed suicide even in nonalcoholic depressives, as it could trigger the actual suicidal behavior and increases the lethality of suicidal acts. Strong evidence suggests that in addition to alcohol and substance use disorders, cigarette smoking is also positively related to suicidal behavior in the general population and—according to most studies—also in samples of patients with major mood disorders. Major depression, bipolar disorder, and schizophrenia are associated with higher rates of cigarette smoking than that of the general population, but smoking remains a significant suicide risk factor even after controlling for psychiatric morbidity. Smokers, in general, are more impulsive, and current smokers with major depressive disorder show a greater number of depressive symptoms and higher levels of suicidality (suicide attempts; suicidal ideations) during their depressive episode than those with major depressive disorder who have never smoked. Patients suffering from disabling, painful, and/or life-threating medical disorders, including cancer, are also at elevated risk of suicide, particularly when depression is also present. The actual risk of suicide in patients with cancer varies with anatomical location, stage of the cancer, and with the result of cancer therapy. Despite the recent progress in cancer treatments, this diagnosis is still a major stressor for many patients, especially those with other psychiatric or psychosocial suicide risk factors.

Prior Course of the Mood Disorder

Clinically detectable suicide risk factors related to the prior course of the illness in patients with mood disorders are early onset of the unipolar or bipolar disorder, relatively early stage of the illness, first episode depression, predominantly depressed polarity during the prior course of the illness (which is typical for bipolar II disorder), rapid cycling bipolar disorder, and, of course, suicidal ideation or suicide attempt in the past: all of these have been repeatedly demonstrated to increase the risk of attempted or completed suicide in patients with mood disorders. Suicide risk appears to be particularly high primarily early in the course of bipolar illness, and this also underlines the importance of precise diagnosis and effective treatment of specific mood disorders as early as possible.

Specific Personality Features

Impulsivity, aggressivity, pessimism, and few reasons for living, as personality traits, markedly increase the risk of all forms of suicidal behavior when combined with depression. There is a strong association between impulsiveness and bipolar disorder, and impulsivity becomes

more pronounced during depressive or (hypo)manic episodes. Even modest manic symptoms during bipolar depressive episodes are associated with greater levels of impulsivity and higher rates of suicide attempts. Current findings also show that in contrast to hyperthymic temperament and cyclothymic, irritable or depressive affective temperaments, the first two are also characteristic for bipolar (mainly for bipolar II) disorder, which also increases the risk of suicidal behavior in patients with both bipolar and unipolar major depressive episodes. In addition, the role of cyclothymic temperament in suicidal behavior is supported by the finding that anger and hostility are most common in remitted bipolar and major depressive patients with cyclothymic affective temperament. The role of unstable affective temperament in suicidal behavior is also demonstrated by findings showing that cyclothymic temperament significantly predicts future bipolar transformation and suicide attempts in adult and juvenile depressives, and bipolar I, bipolar II, and unipolar major depressive disorder patients with cyclothymic temperament have a significantly higher rate of prior suicide attempts and lifetime or current suicidal ideation than noncyclotyhmic depressives. The persistent, frequent, and marked instability of mood, thinking, and behavior are the most characteristic features of cyclothymia, which is the attenuated manifestation of major bipolar mood disorders. However, marked affective temperaments, which can be detected in about 15 to 20 percent of the general population, become suicide risk factors only during major depressive episodes, as persons with marked affective temperament have a much higher chance of developing major mood disorders. Patients with cyclothymic and irritable affective temperaments who attempt suicide significantly more frequently report childhood physical and/or sexual abuse, suggesting that besides impulsivity, cyclothymic or irritable temperaments are further mediating variables between these early negative life events and adult suicidal behavior. Pessimistic personality features, cognitive rigidity, ruminative thinking, maladaptive coping strategies, and disturbed neurocognitive functions like impaired decision making, as could be detected also by specific neurocognitive tests, are also characteristics of most suicidal depressives. The interaction between personality features and illness characteristics in the suicidal process is best formulated in the "stress-diathesis model," which proposes that suicidal behavior is determined not only by the stressor (acute major psychiatric illness or acute major personal crisis) but also by a diathesis or predisposition (impulsive, aggressive, pessimistic personality traits).

Previous Suicide Attempt

Previous suicide attempts, particularly in the case of violent or more lethal methods, are the most powerful single predictor of future attempts and fatal suicide both in patients with major mood disorders or with other

psychiatric diagnoses. One third of unipolar and about half of bipolar depressives attempt suicide at least once during their lifetime, and suicidal behavior shows high consistency across major depressive episodes. In the majority of cases, suicidal depressives become suicidal again in the next depressive episode. Bipolar patients, particularly bipolar II patients, use more violent and more lethal suicide methods than unipolar depressives, respectively, and is more common in males. Nonviolent or low-lethality suicide attempts do not mean low suicide risk, as those who repeatedly attmept suicide commonly change their method from nonviolent to violent, but the opposite pattern is rare. Bipolar depressives with a history of suicide attempts have a more severe symptomatology in general; they report more hopelessness, self-blame, guilt, current suicidal ideation; and more frequently show aggressive–impulsive personality traits than nonsuicidal bipolar depressives do. Therefore, higher rates of suicidal behavior (mainly completed suicide) in patients with bipolar disorder than in unipolar major depression may be partly due to a specific effect of bipolar disorder, resulting in more dangerous or more lethal suicidal behaviors.

Family History of Suicidal Behavior

Family, twin, and adoption studies show that a family history of suicidal behavior and/or major mood disorders in first- and second-degree relatives are among the strongest risk factors for both attempted and completed suicide in psychiatric patients in general, and in patients with major mood disorder in particular. The heritability index of completed suicide is approximately 40 percent, and it is even higher among depression-related suicides. However, the familial component of suicidal behavior seems to be partly independent of specific psychiatric disorders, as relatives of suicidal persons are more than 10 times more likely than relatives of comparison subjects to attempt or complete suicide after controlling for psychopathology. Similarly, a recent longitudinal study found that the children of parents with mood disorders who attempted suicide were five times more likely to attempt, even after controlling some other offspring risk factors for suicidal behavior, such as previous suicide attempt and history of mood disorder (both assessed at baseline) or mood disorder at the time point before the suicide attempt. Bipolar and unipolar major depressive patients who have a family history of suicidal behavior and who were exposed to childhood physical and/or sexual abuse are at greater risk of suicidal behavior. Impulsivity and irritable temperament seem to be the link between childhood abuse and suicidal behavior, particularly in the case of pure or mixed major depressive episodes. Research data, including twin and adoption studies, show that the familial aggregation of suicidal behavior has several components: genetic background of major psychiatric disorders, genetic–familial nature of impulsive–aggressive personality

traits irrespective from psychiatric disorder, and in some cases copycat mechanisms. However, it should be noted that the "suggestive" effect of a suicidal family member becomes active mainly when a major depressive or other severe mental disorder develops.

Negative Life Events and Adverse Life Situations

Although negative life events do not lead to suicidal acts in the general population, they could trigger suicidal behavior in vulnerable persons, particularly in high-risk groups, such as those with major depression. Adverse, unwanted, or stressful life events play an important role in the suicidal process as predisposing (childhood events, including physical and sexual abuse) and precipitating (adulthood events) factors. Although negative life events do not inevitably lead to suicidal behavior in healthy persons or even in the high-risk groups such as psychiatric (particularly mood disorder) patients, they may actually trigger the suicidal process, particularly if other risk factors or adverse situations are also present. One third of patients with unipolar major depression or bipolar disorder report a history of childhood physical and/or sexual abuse or neglect, and depressed patients with such a history have an earlier age of onset of bipolar illness, greater psychiatric comorbidity, and increased rates of suicide attempts. These adverse childhood events may lead to developmental and neuroendocrinological scars that render the subject vulnerable for suicidal behavior in a crisis situation during adulthood, including severe major depression or bipolar disorder.

Affect mastery, or the capacity to regulate affect, is gradually acquired after birth—it is not present early on. Only with time and with reasonable parenting do children attain the capacity to regulate themselves and to keep themselves in reasonable emotional composure. The various developmental stages bring different emotional stresses to children. They must learn to deal with separations, frustrations, or other emotionally challenging circumstances, hence the importance of good early attachment experiences is unambiguous. The parent who is frightened, insecure, or vindictive, prevents the child from making a "secure attachment." Insecure attachment with affective dysregulation, which is now identified in various central nervous system activities as well as in the hypothalamic–pituitary–adrenal (HPA)-axis activity, is considered a suicide risk as well.

About half of all completed suicides in both bipolar and unipolar mood disorders are associated with recent severe, acute negative life events, or adverse life situations, such the death of a relative or close friend, isolation, living alone, separation or divorce, and being unmarried. However, the stressors are commonly dependent on the victims' own behavior, particularly in the case of bipolar I or bipolar II disorder. Hypomanic episodes, but mainly manic episodes, can easily lead to financial extravagance, aggressive–impulsive behavior, or episodic promiscuity, thus

generating several interpersonal conflicts, marital breakdown, and new negative life events, all of which have a negative impact on the further course of the illness and on the family atmosphere of the patients as well, which may ultimately trigger a new depressive episode and suicidal behavior. Other relevant stressful events identified as suicidal risk factors include interpersonal or occupational difficulties; personal or economic losses; retirement; bereavement; social isolation; feeling stigmatized, as also commonly seen in patients with mood disorders; and limited access to support or health care or social services. Permanent adverse life situations (e.g., unemployment and social isolation) as well as acute psychosocial stressors (e.g., loss events and financial breakdowns) in adult patients with unipolar or bipolar disorders are useful indicators of suicidality in the clinical practice, primarily if other risk factors are also present. It should be noted, however, that unemployment is frequently the consequence of a major mood disorder, particularly in cases of comorbid substance use disorder. The risk of suicide is very high in the first few days of hospitalization and a few days and weeks after discharge from an inpatient psychiatric departments particularly in the case of unplanned discharge, in patients with a short hospital stay, with a high number of previous hospitalizations, and, of course, in the case of lack of follow-up care. Doctors should be alert when a patient discharged from an inpatient or outpatient psychiatric department seeks help for any psychological or even medical problems.

Demographic Factors

Although Caucasians, males, older persons, and urban residents, as well as minority groups (e.g., immigrants, ethnic minorities, specific professions, prisoners, lesbian, gay, bisexual, and transgender persons) are more and less overrepresented among completed suicides, females and young persons more commonly attempt suicide. Suicide mortality is elevated among those in a certain profession—especially veterans, health care professionals, and agricultural workers. However, the role of these demographic factors in suicidal behavior is less pronounced among patients with mood disorders; therefore, these do not have a clinically significant powerful predictive value in the case of patients with unipolar major depression or bipolar disorder, and they have just weak importance in the prediction of suicidal behavior in individual cases. However, due to the interaction of genetic biological and psychosocial risk factors, close relatives of suicide victims are at a particularly elevated risk of suicidal behavior, and these persons always need some kind of professional help ideally immediately after the suicide of their relatives. Suicidal behavior shows a typical seasonal variation: spring and early summer is the peak, and winter is the low of completed suicides on both hemispheres, and this seasonal pattern is more pronounced among depression-related suicides,

among males, and among those who use a violent suicide method.

Clinical Utility of Suicide Risk Factors

As suicidal behavior is very rare in the absence of a current major psychiatric disorder, a mental state examination is an integral part of the assessment. The suicide risk factors in mood disorders are shown in Table 13.6–1. This table organizes the risk factors from the point of view of the clinical examination when clinicians collect information about the present clinical manifestations as well as about the past history of patients with mood disorders. However, these risk factors have different clinical relevance and diverse predictive utility; illness-related factors are more alarming and more powerful than personality-related risk factors. Suicide risk factors have an additive effect; the more risk factors present, the higher the suicide risk. For proper prevention, clinical assessment of suicide risk should pay attention to warning signs. The American Association of Suicidology proposed to remember the acronym "Is Path Warm":

- ► I Ideation—threatened or communicated
- ► S Substance abuse—excessive or increased
- ► P Purposeless—no reasons for living; anhedonia
- ► A Anxiety, agitation and insomnia
- ► T Trapped—feeling no way out; perceived burdensomeness
- ► H Hopelessness
- ► W Withdrawal—from friends, family, society
- ► A Anger (uncontrolled)—rage, seeking revenge
- ► R Recklessness—risky acts, unthinking
- ► M Mood changes (dramatic)

Table 13.6–1.
Clinically Explorable Suicide Risk Factors in Mood Disorders

A. Illness-related risk factors
   1. Risk factors related to current mood episodes
   a. Severe major depressive episode
      - Current suicide attempt, plan, ideation,
      - Hopelessness, guilt, pessimism, few reasons for living, unusual behavior (e.g., making a will or testament)
      - Major depressive episode with mixed features (mixed depression), agitated depression
      - Past mania or hypomania (bipolar I or II diagnosis)
      - Insomnia, anxiety
      - Appetite and/or weight loss
      - Comorbid anxiety disorders or anxiety disorders, substance use disorder, acute alcohol intake, cigarette smoking
      - Comorbid personality disorder(s)
      - Comorbid disabling medical disorder(s)
      - Lacking medical care and lacking family and social support

- Noncompliance, treatment-resistance
- First few days or weeks of the treatment (particularly if appropriate care and co-medication is lacking) and first few weeks after hospital discharge

b. Major depressive or manic episode with mixed features

2. Risk factors related to prior course of the mood disorder
   - Previous suicide attempts or suicidal ideation (particularly in the case of violent and highly lethal methods)
   - Early onset or early stage of the illness
   - Predominantly depressive course or high proportion spent in major depressive or mixed mood episode, first episode depression
   - Rapid cycling course of bipolar I or II disorder

B. Personality-related and historical risk factors

1. Risk factors related to personality features
   - Aggressive–impulsive–pessimistic personality traits
   - Cyclothymic, irritable, depressive, anxious affective temperament
2. Risk factors related to personal history and/or family history
   - Early childhood adverse events (parental loss, separation; emotional, physical, and sexual abuse)
   - Permanent adverse life situations (unemployment, financial problems, isolation, living alone, chronic and disabling medical disorders)
   - Acute psychosocial stressors (loss events, acute financial catastrophe)
   - Family history of depressive or bipolar disorders (first- and second-degree relatives)
   - Family history of suicide and/or suicide attempt (first- and second-degree relatives)

Clinicians should be aware of communications about suicide; very often, patients talk about suicide, death, and/or having no reason to live. Most suicidal individuals give definite warnings of their suicidal intentions, but significant others are either unaware of the significance of these warnings or do not know how to respond to them. However, suicidal intent is sometimes denied or minimized. Clinicians should always remember to ask about suicide; questioning patients about suicide thoughts and plans is always indicated and crucial for proper prevention. Talking about suicide does not increase the risk. During suicidal crisis, patients may withdraw from friends and/or social activities, and they may also lose interest in hobbies, work, school, and so on. Suicidal individuals may also prepare for death by writing out a will (unexpectedly) and making final arrangements, giving away prized possessions, clearing up business matters pertaining to survivors, or accumulating drugs or poisons for taking his or her life, and so on. Because these warning signs, which are highly indicative for current suicidal danger, are reported mostly by the relatives of patients, keeping contact with family members or with significant others is also important. Several suicidal intent scales have been developed that could be useful in special circumstances, but they are routinely used mostly in clinical research.

Suicide Protective Factors in Patients with Mood Disorders

In contrast to the numerous suicide risk factors, only a few circumstances

are known to have a protective effect against suicidality, and suicidal behavior is always the result of the complex interplay between risk and protective factors. Good family and social support, pregnancy and postpartum period, having a large number of children, holding strong religious beliefs (regardless of the nature of religion), regular physical activity, and restricting lethal suicide methods whenever possible, (e.g., to reduce domestic and car exhaust gas toxicity, barriers at train stations and bridges), as well as introducing stricter laws on drug and gun control seem to have some protective effect. Recent studies have found that optimistic personality features as well as hyperthymic temperament may be also a protective factor and that there is a strong relationship between psychological resilience and hyperthymic temperament both in depressive and healthy individuals. However, one of the most extensively studied and changeable suicide protective factors in major mood disorders is acute and long-term treatment, both pharmacological and nonpharmacological.

Suicidal behavior, as a final act, is always the result of a dynamic interplay between risk and protective factors. However, in individual cases, suicide risk factors have a more practical relevance, and the presence of one or more suicide risk factors is more important than the lack of protective ones. In the majority of cases, more than one suicide risk factors are present simultaneously, and their effect is cumulative: the higher the number of risk factors, the higher the suicide risk. They are usually closely related to each other. For example, elderly major depressives frequently have multiple medical disorders, and they commonly experience loss events and isolation; major depression easily leads to secondary alcoholism, loss of job, family breakdown, financial problems, and so on. Considering the clinically relevant suicide risk factors in patients with mood disorders, in the majority of cases, suicidal behavior is predictable, but, of course, not guaranteed. Severe major depression with agitation, anxiety, insomnia, or hopelessness, particularly in combination with prior suicide attempt and, of course, with current suicidal ideation, are the most important signs for identifying acute or hyperacute suicide risk that requires prompt action.

ANTIDEPRESSANT-ASSOCIATED SUICIDAL BEHAVIOR: THE ROLE OF UNDERLYING BIPOLARITY

Given that suicidality is the most visible outcome of *untreated or inadequately treated* depressive illness, the U.S. Food and Drug Administration (FDA) warning of emergent suicidality in children and adolescents based on the antidepressant arm of placebo-controlled randomized trials has created understandable concern in clinical practice. The issues involved are of broader public health importance for all patients, not just for youngster patients. In 2004, the FDA advised that antidepressants could sometimes be related to suicidal behavior, which in

their database examined mainly pertains to children and adolescents. Newer meta-analyses have confirmed the FDA's position for a slightly increased risk of suicidality (suicidal ideation and attempts but not completed suicide) in juvenile patients in randomized controlled antidepressant trials on unipolar major depression. As in other branches of medicine, psychiatrists must always be vigilant of the rare risk of iatrogenesis when prescribing potent agents like antidepressants to patients with depressive disorders where the risk of suicidality is inherent and quite common.

The most challenging paradox of contemporary psychiatry is that antidepressants prevent suicidal behavior among severely ill, frequently suicidal "real-life" unipolar major depressives, but they can induce such behavior rarely in less severe, actually nonsuicidal unipolar major depressives in randomized controlled clinical trials, particularly in the youngest age cohorts. The main cause of the many disbeliefs about antidepressants and suicide is that some people do not realize that these two opposite effects are not in the same magnitude. As *antidepressants prevent depression-related suicidal behavior in the vast majority of cases*, they could *increase the risk only in a small, vulnerable subpopulation*; therefore, in the everyday clinical practice, the suicide preventive effect of antidepressants highly overshadows this unwanted iatrogeny. There is no doubt that suicide mortality is much higher among untreated depressives than among antidepressant-treated depressives. The rarely occurring suicidal behavior in patients taking antidepressants is mostly the consequence of the lack of antidepressant effect (antidepressant-associated suicidality) but is rarely the result of suicide-inducing potential of antidepressants (antidepressant-induced suicidality), as antidepressants can sometimes worsen depression. Unrecognized bipolar depressives (mainly bipolar II depressives in the daily clinical practice) and mixed "unipolar" depressive patients with intradepressive (hypo)manic symptoms (DSM-5 major depressive episode "with mixed features") in the randomized controlled trials (RCTs) on antidepressant monotherapy in unipolar major depression are considered "unipolar" major depressives, which means that these patients receive antidepressants and—if needed—co-administered anxiolytics, but not mood stabilizers. Antidepressant monotherapy without the concomitant use of mood stabilizers or atypical antipsychotics in threshold and subthreshold bipolar depressives—who have typically early illness onset and young current age—can result in high rates of antidepressant resistance and can sometimes worsen the cross-sectional picture of depression, particularly in adolescents and young adults, not only by causing (hypo)manic switch but also by inducing or aggravating depressive mixed state or agitation, also called "activation syndrome," which is the major substrate of suicidal behavior. However, 30 to 40 percent of diagnosed unipolar patients with major depressive

disorders have clinically significant lifetime or current subthreshold hypomanic symptoms (subthreshold bipolar depression and mixed depression or major depressive episode "with mixed features," respectively) as well as other basic clinical features, such as a family history of bipolar disorder, early onset, and bipolar conversion, which are characteristic and external validators for bipolar disorder. This means that more than one third of DSM-5 diagnosed unipolar major depressives are in fact subthreshold or mixed bipolar depressives, and in addition to the antidepressant resistance, primarily these patients are the subjects of bipolar conversion as well as ultimately antidepressant-associated worsening of depression and related suicidality. In contrast to "officially" diagnosed bipolar I and bipolar II depressives as well as actively suicidal or psychotic depressives, subthreshold bipolar and unipolar major depressives with mixed features are not excluded from the RCTs on "unipolar" major depression. The slightly elevated (but in absolute sense very low) risk of suicidal behavior (attempts but not completed suicides) among younger patients taking antidepressant drugs compared to those taking a placebo in randomized controlled antidepressant trials in DSM-III and DSM-I diagnosed unipolar major depressive disorders might be the consequence of the depression-worsening effect of antidepressant monotherapy in subthreshold and mixed bipolar depressed patients who were included into these trials and falsely diagnosed as suffering from unipolar depression and receiving antidepressants but not mood stabilizers. The depression-worsening effect of antidepressant monotherapy in threshold and subthreshold bipolar depression is more pronounced in the everyday clinical practice where not only most major depressives with subthreshold bipolarity but also many bipolar II patients are diagnosed and treated as unipolar depressives. Among antidepressant-treated patients, the rate of prior suicide attempts is much higher in the case of unrecognized versus recognized bipolarity, and antidepressant-induced "activation syndrome" or agitation is also more common in bipolar than in unipolar depression. The recognition of the role of "pseudounipolar" depressive mixed states in suicidal behavior has clear implications for suicide prevention. The correct identification of the "covert" bipolar nature of the given depressive episode, as reflected in mixed or agitated clinical presentation, is crucial for selecting the most appropriate treatment.

As a consequence of the FDA Black Box Warning in 2004, the decreased use of antidepressants in children and adolescents seen in some countries has been accompanied by a concurrent increase in suicide rates in those age groups, whereas in middle-aged and older persons, the utilization of antidepressants increased continuously, and suicide rates further decreased. These findings indicate that the Black Box Warning, contrary to its original intention, resulted in an increasing number of young depressives not receiving antidepressants and an increase in the suicide

mortality of this age group. This increase, which occurred among young depressed persons without antidepressant treatment, also shows that in everyday clinical practice, the suicide preventive effect of antidepressants highly overshadows the rare and unwanted suicide-inducing effect. The precise and early diagnosis of bipolar disorder, even in first-episode depressives, and the formal recognition of bipolar depressive mixed states in the official diagnostic system (mixed features specifier across all mood syndromes in the DSM-5) will help to red flag those pseudounipolar mixed depressives for whom antidepressant monotherapy is contraindicated and for whom mood stabilizers and/or atypical antipsychotics are indicated on clinical grounds.

PREVENTION OF SUICIDE IN PATIENTS WITH MOOD DISORDERS

Prevention of suicide in patients with depressive disorders is not only a great challenge for everyone who cares for or has contact with them but could also be among the most reliable indicators of the efficacy as well. As much of the putative psychosocial and demographic suicide risk factors are not modifiable in the frame of individual health care, consequently, suicide preventive efforts of psychiatrists and other medical professionals should focus on the treatable and amendable contributory psychiatric disorders involved in such behavior. The role of the health care system in suicide prevention varies as a function of the setting (e.g., psychiatric settings vs. primary care practice), with specialized suicide prevention centers (where they exist) playing a most important role. Because suicidal behavior is a very complex phenomenon, even in patients with mood disorders, its prevention should be also complex, and the majority of suicide attempts and completed suicides in patients with mood disorders are preventable. As suicidal behavior in patients with mood disorders is very rare in the absence of a major depressive or mixed affective episode, reducing the time spent in these high-risk states is essential for their suicide prevention. While successful acute treatment can only prevent suicide connected with the given mood episode, given the highly recurrent nature of major mood disorders, it is adequate long-term therapy that can provide effective protection for a longer period of time.

In the case of acute suicide danger, the patient needs close observation and urgent hospitalization, even against of his or her own wish. Crisis intervention, whenever needed, is also helpful. After an open discussion with the patient and relatives, involuntary admission is rarely needed. It should be explained that hospitalization will be as short as possible and that it is best for the patient. In cases when acute hospitalization is not indicated, close observation by family members and removing possible means of suicide (i.e., guns, drugs, pesticides, and car keys) is advised, as is consultation with a local outpatient psychiatrist in the GP and other non-psychiatric practice. As anxiety, psychomotor agitation, and insomnia

markedly increase acute suicide risk, prompt pharmacological treatment of these symptoms is always needed, and these treatments provide immediate results. In the most severe suicidal (sometimes stuporous or catatonic) depressives, electroconvulsive therapy (ECT) is the most useful treatment strategy.

Suicide Prevention in the Frame of Psychiatry

Suicide and suicide attempts are not only among the most tragic events of human life, but they are also the most hard end points or most visible treatment outcomes in patients with psychiatric and particularly mood disorders. In addition, unfortunately many psychiatrists will lose a patient to suicide, which can cause severe distress for the doctors as well as the entire treating team. Although the elimination of acute suicide danger by all possible means is a common task for everyone, it is especially important that health care professionals diagnose those who are at risk so that thay can be managed effectively.

Suicide prevention in patients with mood disorders is much more than detecting and managing the actual suicide risk. Targeting only those persons who are at acute suicide risk is frequently ineffective because it is often too late to make a successful intervention at this late stage. As suicidal behavior does not usually occur in the very early stages of depression and other major mental illnesses, this offers some time for making a correct diagnosis and effective acute and long-term treatment even before the first suicidal act.

Large-scale, retrospective, and prospective open-naturalistic and controlled long-term clinical studies, including severely ill, comorbid, frequently suicidal depressives (usually inpatients), have shown that appropriate acute and long-term pharmacotherapy (antidepressants in unipolar depression and mood stabilizers in bipolar disorder) markedly reduces suicide morbidity and mortality—and probable other forms of aggressive, violent behaviors—even in this high-risk population. Bipolar patients on mood stabilizers with or without antidepressants or antipsychotics, and unipolar depressives on antidepressants, show a 70 to 80 percent (range 56 to 93 percent) risk reduction of completed and attempted suicide compared to those without pharmacological treatment. Although long-term treatment with lithium and antiepileptic mood stabilizers results in a similar reduction in suicide mortality, lithium may have some superiority in preventing suicide. Register-based observational cohort studies also show that former inpatients with unipolar major depression who continued treatment with antidepressants, and former inpatients with bipolar disorder who continuously took mood stabilizers (lithium or antiepileptic mood stabilizers), had a markedly decreased rate of completed suicide compared with those who stopped taking antidepressants or mood stabilizers.

This marked antisuicidal potential of lithium seems to be more than the simple result of its **episode-prophylactic effect,** as it has been shown that during the long-term lithium prophylaxis of recurrent bipolar or unipolar affective disorder patients with at least one prior suicide attempt, a significant reduction in the number of suicide attempts was found not only in the excellent responders but also in moderate responders and in nonresponders. The clinical importance of this finding is that in the case of lithium nonresponse, when the patient has one or more suicide risk factor, instead of switching lithium to another mood stabilizer, the clinician should retain lithium (even on a lower dose) and combine it with another mood stabilizer. The well-documented strong antisuicidal effect of lithium in patients with bipolar and unipolar major mood disorder has also been supported by results derived from the general population: investigating the lithium levels in tap water and drinking water in Texas, Japan, Austria, and Greece, lithium levels were significantly and negatively associated with suicide rates. The evidence that suicide rates among depressed patients have progressively and significantly lowered through the first half "pretreatment era" (1900 to 1939), "ECT era" (1940 to 1959), and "antidepressant era" (1960 to 1992) also supports the suicide-preventive effect of antidepressants in depressed patients. As comorbid anxiety, agitation, and insomnia increase the short-term suicide risk for depressed patients, anxiolytics, atypical antipsychotics, and sleep-promoting drugs—mostly for short-term use—are also part of the therapeutic arsenal. Most recent findings show that ketamine, a glutamate antagonist, is a promising drug for rapidly reducing current suicide risk. Overall, the data reviewed here justify the claim that the increased use of antidepressants (and mood stabilizers) in the last two to three decades was an important contributor in *reduced* suicidal mortality in Canada, Australia, and most European countries.

Pharmacotherapy, however, is a necessary but not sufficient method of reducing suicidal behavior in depressed patients, and psychosocial interventions are always needed. Psychoeducation and supportive psychotherapy are also always needed. There is some evidence that concurrent depression-focused psychotherapies, in combination with pharmacotherapy, also improve the compliance of patients and increase the effectiveness of pharmacotherapy and may therefore contribute to suicide prevention for patients with severe unipolar or bipolar disorders. The interaction between pharmacotherapy and psychosocial interventions is quite complex. It has been reported, for example, that successful episode-preventive medication with mood stabilizers in bipolar patients counteracted dysfunctional cognitions (including lowered self-esteem), and adjunctive cognitive therapy could help to optimize the long-term course of bipolar illness.

However, psychotherapy for depression *alone* is not enough for

preventing depression-related suicide. A systematic review and meta-analysis show that there is little evidence for the assumption that suicidality in depressed patients can be reduced with psychotherapy for depression despite decreasing depressive symptomatology, indicating that in suicidal depressives psychotherapy and psychosocial interventions should be always combined with appropriate acute (and if necessary long-term) pharmacotherapy. Concurrent depression-focused psychotherapies, in combination with pharmacotherapy, also improve the compliance of patients and increase the effectiveness of pharmacotherapy, and may therefore contribute to suicide prevention for patients with severe recurrent unipolar or bipolar disorders. Both hopelessness–pessimism and aggressiveness–impulsiveness may be amenable to cognitive-behavioral therapy and pharmacotherapy (e.g., selective serotonin reuptake inhibitor (SSRI) antidepressants, lithium, and other mood stabilizers). As a significant part of depressives stop their medication in the first few weeks or months, aftercare of recently discharged depressive patients is essential for improving compliance and maintaining efficacy. The role of psychosocial interventions in suicide prevention is also supported by the World Health Organization (WHO) organized study of 1867 suicide attempters (many of whom should have current major depression) seen at emergency departments. A brief individual psychoeducational and psychosocial intervention before discharge and regular contact thereafter (eight telephone, postcard, or personal contacts during the 18-month follow-up) resulted in significantly fewer completed suicides (0.2 percent) in the intervention group (n = 872) than in the treatment as usual group (n = 827, 2.2 percent).

Suicide Prevention in the Primary Care

Despite the fact that more than two thirds of suicide victims contact different levels of health care (mostly GPs and psychiatrists) during last few weeks or months before their death, the rate of recognition of depression and adequate pharmacotherapy among depressed suicide victims is around 20 to 25 percent, which is disturbingly low. While the current prevalence of major mood disorders in the primary care practice is around 8 to 12 percent, many depressed patients are not recognized by their GPs. However, more recent studies reported much higher rates of recognition and treatment of depression in primary care practice (62 to 85 percent), and 33 to 50 percent of them were treated with antidepressants. Physician factors related to poor recognition of depression include lack of experience, insufficient or suboptimal knowledge about symptoms, treatment and good prognosis in treated depression, prejudices about mental illness, lack of postgraduate psychiatric training, insufficient interview skills, lack of cooperation with psychiatrists, and a low level of empathy. Although suicidal behavior is a relatively rare event in primary care, given that

depression is quite frequent in GP practice and is very common among completed suicides, major depression, particularly in combination with past or current suicidal behavior, it should be taken very seriously even in primary care. Compared to nonsuicidal patients, suicide victims visit their GPs much more frequently in the last 4 weeks of their life. Many physicians and even mental health care professionals avoid discussing suicide directly and frankly with patients for fear of provoking suicidal behavior or, more likely, because of personal discomfort. However, asking questions about suicidal ideation and past suicide attempts does not trigger suicide. This is particularly true if such a discussion is accompanied by some sentences explaining that depressive disorders can be successfully treated and that suicidal intent will vanish after (or even before) the recovery from depression. Asking simple questions (e.g., "What do you think about the future?" "Do you feel that life is not worth living?") can easily facilitate further, more deep and honest discussion on the topic of suicide. This is beneficial, as many patients think they are alone or unique in their suicidal ideas. Leaflets, posters, and fliers left in the waiting room as well as providing appropriate websites indicating the symptoms, dangers, and treatable nature of depression and preventable nature of suicide are also beneficial.

Postgraduate training on better recognition and management of depressive disorders among psychiatrists and particularly in primary care appears to prevent suicide. The first example for the significant role of GPs in suicide prevention comes from the Swedish Gotland Study. The study demonstrated that after a short intensive postgraduate training for GPs on the island of Gotland on the diagnosis and treatment of depression, suicide rates and the number of depression-related hospital admissions declined significantly and antidepressant prescription increased markedly a few years after the training. The rate of depressive suicides among all suicides decreased significantly after the training, indicating that the decline in suicide mortality after the education resulted directly from a robust decrease in depressive suicides, and suggesting that this result might not be caused by random fluctuation or by other unknown factors. However, the decline in depressive suicides after the GP training had mostly been the result of a decline in female depressive suicides. Few suicidal males were known to the local medical services, although many of them were known to the police and social welfare services. The main reasons for this are that male depression is frequently masked by abusive, aggressive behavior, and males seek professional help less frequently than females. The favorable effects of the GP training faded after a few years, and repeated education again led to another decrease in suicides, again, mainly in females. Indeed, recent international data have supported the findings of the pioneering Gotland Study, also showing the importance of GPs as the front line in diagnosing and treating depression and in reducing suicide mortality. The

community-intervention suicide prevention project in Hungary between 2000 and 2005 also showed that education of GPs and other health-care professionals as well as the public is an effective method of reducing suicide mortality. Similarly, a randomized-controlled study from Australia demonstrated that suicidal behavior during the 2-year follow-up period was lower for older persons treated by educated GPs as compared to control GPs. However, primary care education in isolation does not have any significant long-term effect on suicide mortality, and primarily complex, permanent, and multilevel educational and organizational interventions are useful. The multilevel community program should be based on the involvement and the strong collaboration with psychiatrists, GPs, gatekeepers, and media. The effect of these multilevel community suicide prevention programs has been demonstrated recently by the German Nuremberg Alliance Against Depression Study and the International European Alliance Against Depression Project, a multifaceted community-based action program against depression and suicide. Better management of depression requires not only improved recognition and treatment skills from the doctors but also good compliance from the patients, as nonadherence to antidepressant therapy is one of the most common causes of treatment failure. GPs should work in close and permanent collaboration with the local mental health services. Outpatient psychiatric consultation is also helpful in the cases of differential diagnostic problems, treatment resistance, and comorbid substance use disorder regardless of whether the patient is suicidal or not. If long-term prophylactic pharmacotherapy is needed (bipolar disorder, recurrent unipolar major depression), then the GP may direct the patient to a psychiatrist for optimizing the therapy. Although depressive disorders represent an important target for suicide prevention in old age, there should be increased attention also for psychosocial and somatic-medical risk factors in this age range.

Regular aftercare with fixed appointments is highly recommended, particularly for those patients with previous suicide attempts. Psychological support should be available for each suicidal depressed patient. This is important because the actual clinical picture immediately after suicide attempt is often misleading, due to the cathartic effect of self-aggression, resulting in a short-lived but sometimes marked improvement in the depression. This can also serve as one of the explanations why some health care workers misinterpret suicide attempts as manipulative acts.

Health care professionals, of course, are unable to prevent all suicides, including those occurring in the context of patients with mood disorders. Nevertheless, the current theoretical knowledge and the available treatment and preventive strategies are sufficient to prevent many, probably most of them. The most important methods of suicide prevention in patients with mood disorders are shown in Table 13.6–2. However, as

health and social care workers can only help those who contact them, public education media campaigns on symptoms, dangers, treatability, and referral pathways are vitally important. As psychosocial and community factors also play an important role in suicide, it is not only health care workers that are responsible for its prevention. All members of the society have their own task, major or minor, and entailing more or less responsibility, in addressing this issue. However, prevention of mood disorder–related suicidal behavior is primarily the duty of psychiatric and primary health care. On an aggregate level, however, organization and societal changes (e.g., the improvement in a primary health care system, empowerment strategies in the workplace, and community or the alleviation of unpredictability and stressful transitions in a society) seem to reduce increased suicide figures. To reduce suicide mortality the following must be done: improve the well-being of people in general (including decreasing unemployment and providing more support for health and social services); restrict lethal suicide methods whenever possible (e.g., reduce domestic and car exhaust gas toxicity and introduce stricter laws on gun control); and finally, initiate more restrictive alcohol and drug policies, which also reduces suicide mortality. These are the responsibilities of the health care system and all of society.

Table 13.6–2.
Suicide Prevention Strategies in Patients Mood Disorders

A. Eliminating acute suicide danger (physical inhibition, emergency hospitalization, sedation, anxiolysis, crisis-intervention)
B. Improving the early diagnosis and treatment of mood disorders with particular regard to the soft (subthreshold) clinical manifestations of bipolar disorder
    1. Education of health care workers, patients, relatives and gatekeepers (clergymen, policemen, teachers, social workers, peer-helpers, etc.)
    2. Adequate acute and long-term treatment aftercare (pharmacotherapy, nonpharmacological interventions such as psychoeducation, psychotherapy, cognitive therapy, family counseling and family therapy, regular long-term care, regular contact, etc.)
C. Improving the patients' compliance (psychoeducation, psychotherapy, cognitive therapy, etc.)
D. Educating the public via printed and electronic media, including the Internet
    1. Educating about the symptoms, dangers, and treatable nature of mood disorders and the preventable nature of suicidal behavior
    2. Reducing the stigma against mood disorders and suicide
    3. Providing information on how and where to get help in the case of mood disorder and suicide crisis

References

Akiskal HS. Targeting suicide prevention to modifiable risk factors: has bipolar II been overlooked? *Acta Psychiatr Scand*. 2007;116:395.

*Akiskal HS, Benazzi F. Psychopathologic correlates of suicidal ideation in major depressive outpatients: is it all due to unrecognized (bipolar) depressive mixed states? *Psychopathology*.

2005;38:273.

Angst J, Cui L, Swendsen J, et al. Major depressive disorder with subthreshold bipolarity in the National Comorbidity Survey Replication. *Am J Psychiatry*. 2010;167:1194.

Bostwick JM, Pankratz VS. Affective disorders and suicide risk: a reexamination. *Am J Psychiatry*. 2000;157:1925.

Brent DA, Melhem NM, Oquendo M, et al. Familial pathways to early-onset suicide attempt: a 5.6-year prospective study. *JAMA Psychiatry*. 2015;72:160.

Cuijpers P, de Beurs DP, van Spijker BA, Berking M, Andersson G, Kerkhof AJ. The effects of psychotherapy for adult depression on suicidality and hopelessness: a systematic review and meta-analysis. *J Affect Disord*. 2013;144:183.

Fawcett J. *Depressive disorders*. In: The American Psychiatric Publishing Textbook of Suicide Assessment and Management, Second Edition, Simon RI, Hales RE, eds. Washington, DC: American Psychiatric Publishing; 2012;109–123.

Fleischmann A, Bertolote JM, Wasserman D, et al. Effectiveness of a brief intervention and contact for suicide attempters: a randomized controlled trial in five countries. *Bull World Health Organ*. 2008;86:703.

Fountoulakis KN, Gonda X, Siamouli M, Rihmer Z. Psychotherapeutic intervention and suicide risk reduction: a review of the evidence. *J Affect Disord*. 2009;113:21.

*Goodwin FK, Jamison KR. *Manic Depressive Illness. Bipolar Disorders and Recurrent Depression*. New York: Oxford University Press; 2007.

Haukka J, Suominen K, Partonen T, Lönqvist J. Determinant and outcomes of serious attempted suicide: a nationwide study in Finland; 1996–2003. *Amer J Epidemiology*. 2008;167:1155.

Hawton K, van Heeringen K. Suicide. *Lancet*. 2007;373:1372.

Isacsson G, Ahlner J. Antidepressants and the risk of suicide in young person—prescription trends and toxicological analyses. *Acta Psychiatrica Scandinavica*. 2014;129:296.

*Isometsa E. Suicidal behaviour in mood disorders—who, when, and why? *Canad J Psychiatry*. 2014;59:120.

Luoma JB, Martin CE, Pearson JL. Contact with mental health and primary care providers before suicide: a review of the evidence. *Am J Psychiatry*. 2002;159:909.

*Mann JJ, Apter A, Bartolote J, et al. Suicide prevention strategies: a systematic review. *JAMA*. 2005;294:2064.

Nordentoft M, Mortensen PB, Pedersen C. Absolute risk of suicide after first hospital contact in mental disorder. *Arch Gen Psychiatry*. 2011;68:1058.

O'Connor R, Nock MN. The psychology of suicidal behaviour. *Lancet Psychiatry*. 2014;1:73.

Pompili M, Gonda X, Serafini G, et al. Epidemiology of suicide in bipolar disorders: a systematic review of the literature. *Bipolar Disord*. 2013;15:457.

Pompili M, Innamorati M, Rihmer Z, et al. Cyclothymic-depressive-anxious temperament pattern is related to suicide risk in 346 patients with major mood disorders. *J Affect Disord*. 2012;136:405.

Pompili M, Rihmer Z, Innamorati M, Lester D, Girardi P, Tatarelli R. Assessment and treatment of suicide risk in bipolar disorders. *Expert Rev Neurother*. 2009;9:109.

Rihmer Z. Depression and suicide and attempted suicide. In: O'Connor R, Platt S, Gordon J, eds. *International Handbook of Suicide Prevention: Research, Policy and Practice*. Oxford: Wiley

Blackwell; 2011.

*Rihmer Z, Gonda X. Pharmacological prevention of suicide in patients with major mood disorders. *Neurosci Biobehav Rev*. 2013;37:2398.

Rihmer Z, Rutz W. Treatment of attempted suicide and suicidal patients in primary care. In: Wasserman D, Wasserman C, eds. *Oxford Textbook of Suicidology and Suicide Prevention*. New York, Oxford University Press; 2009.

Schaffer A, Isometsa E, Tondo L, et al. International society for bipolar disorders task force on suicide: meta-analyses and meta-regression of correlates of suicide attempts and suicide deaths in bipolar disorder. *Bipolar Disord*. 2015;17:1.

Sudak HD. Suicide. In: Sadock BJ, Sadock VA, Ruiz P, eds. *Kaplan and Sadock's Comprehensive Textbook of Psychiatry*. 9th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2009.

Takeshima M, Oka T. Association between the so-called "activation syndrome" and bipolar II disorder, a related disorder, and bipolar suggestive features in outpatients with depression. *J Affect Disord*. 2013;151:196.

Van Heeringen K, Mann JJ. The neurobiology of suicide. *Lancet Psychiatry*. 2014;1:63.

Vita A, De Peri L, Sacchetti E. Lithium in drinking water and suicide prevention: a review of the evidence. *Int Clin Psychopharmacol*. 2015;30:1.

Wasserman D, Rihmer Z, Rajescu D, et al. The European Psychiatric Association (EPA) guidance on suicide treatment and prevention. *Eur Psychiatry*. 2012;27:129.

▲ 13.7 Mood Disorders: Pharmacological Treatment of Depression and Bipolar Disorders

LAKSHMI N. YATHAM, M.B.B.S., M.R.C.PSYCH. (UK), AND SIDNEY H. KENNEDY, M.D.

INTRODUCTION

Mood disorders comprise two separate chapters in the DSM-5: "Depressive disorders" and "Bipolar and Related Disorders." Among the major changes, mixed episode as a diagnostic entity was eliminated and a course specifier "with mixed features" has been added for both major depressive episodes and manic/hypomanic episodes. Interestingly, this allows an individual with an major depressive episode with three or more (hypo)manic symptoms and no prior manic/hypomanic episodes to remain within the major depressive disorder domain, rather than be classified as a forme fruste of bipolar disorder. Other changes within the depressive disorders chapter are the inclusion of disruptive mood dysregulation disorder (DMDD) and premenstrual dysphoric disorder as new diagnoses, and the replacement of chronic depression and dysthymia with a single diagnostic entity—persistent depressive disorder (PDD). The essential feature of PDD is depressed mood and at least two depressed symptoms present for 2 years or more (1 year for children or adolescents). There is still the possibility that major depressive episodes may occur intermittently (*double*

depression).

Abnormally and persistently increased goal directed activity or energy has been added to the stem criterion for a manic/hypomanic episode. Further, manic/hypomanic episodes emerging during antidepressant treatment are considered to be indicative of bipolar disorder, provided that the symptoms persisted beyond the physiological effects of treatment. The "not otherwise specified" (NOS) category has been replaced with "other specified depressive disorder," "other specified bipolar and related disorder," and "unspecified bipolar and related disorder."

Unfortunately, there are limited systematic treatment data for new disorders such as disruptive mood dysregulation disorder and other specified depressive or bipolar disorder or unspecified bipolar disorder. Therefore, the main focus of this chapter will be on the treatment of major depressive disorder and bipolar disorder.

TREATMENT OF MAJOR DEPRESSIVE DISORDER

Major depressive disorder is typically a disorder of recurrent depressive episodes rather than a single episode. Several treatment options are available for managing patients with major depressive disorder and the list of treatments continues to expand, although the progress in identifying search for novel strategies has been slow. This section will provide an outline of foundations of the management of major depressive disorder and provide guidance to clinicians on selection of first-line antidepressants, and the next steps if first-line treatments fail. The reader is referred to other chapters in this book for the management of major depressive disorder during pregnancy and postpartum period, and the management of major depressive disorder in children and adolescents and elderly.

FOUNDATIONS OF MANAGEMENT

Address Acute and Maintenance Treatment

Effective management of the major depressive disorder requires a focus on both acute and maintenance aspects of the disorder from the outset (see Table 13.7–1). While some authors differentiate continuation and maintenance phases after the acute treatment phase, there is limited evidence to support this differentiation. The primary goal in the acute phase is to achieve remission—a state that comes close to being symptom free (typically remission represents a score of ≤7 on the Hamilton Rating Scale for Depression (HAM-D) or ≤10 on the Montgomery-Asberg Depression Rating Scale [MADRS]). Since failure to achieve remission is associated with a higher risk of recurrence, this has replaced the previously accepted treatment aim of response, defined as a reduction of ≥50 percent on one of the above scales.

Table 13.7–1.
Phases of Antidepressant Treatment

Treatment Phase	Duration	Goals	Activities
Acute and continuation	8–12 weeks	► Achieve symptomatic remission ► Monitor side effects ► Restore function	► Establish therapeutic alliance ► Provide psychoeducation ► Select optimal antidepressant treatment(s) ► Supportive and measurement-based care ► Monitor progress
Maintenance	6–24 months or longer	► Return to full function and quality of life ► Prevention of recurrence	► Continue psychoeducation ► Rehabilitate ► Manage comorbidities ► Monitor for recurrence

Modified from Lam RW, McIntosh D, Wang J, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 1. Disease Burden and Principles of Care. *Can J Psychiatry*. 2016;61(9):510–523.

Included in the psychiatric assessment, it is important to assess anxiety symptoms, substance use, suicidality, psychotic features, bipolarity, and other psychiatric comorbidities, as well as conduct a careful assessment of physical health, psychosocial, functional factors, and concurrent prescription medications. Details of prior treatment response are also valuable in guiding future treatment decisions.

During the acute phase of treatment, it is also important to monitor comorbid conditions and tailor the treatment appropriately to individual needs. Reports on psychiatric comorbidity indicate that anxiety disorders (60 percent), substance use disorders (25 percent), and impulse control disorders (30 percent) are highly prevalent among patients with lifetime major depressive disorder diagnosis. There is also a bidirectional relationship between major depressive disorder and physical health. Individuals with prior depressive episodes have high risk of medical disorders including heart disease and diabetes, while patients with medical disorders, including cardiovascular and stroke disease, chronic pain, obesity, diabetes, epilepsy and other central nervous system (CNS) disorders, and cancer have two to fourfold increases in risk of major depressive disorder.

Although symptomatic improvement/remission has been the traditional goal for the treatment of depression, it is increasingly recognized that improvement in function (e.g., return to work or other meaningful activities) and overall quality of life are most valued by patients. It is important to listen to the apprehensions and expectations of patients and provide education about depression and treatment options. Many patients will also have searched the Internet for information and will welcome a

chance to discuss the materials they have downloaded. These discussions are likely to recur during ongoing supportive care and it is important to establish good rapport at the initial visit.

Since many depressed patients have medical comorbidities, requiring a range of pharmacotherapies, and frequently receive more than one antidepressant medication, it is essential to consider the risk of drug–drug interactions. Most antidepressant-related interactions are associated with cytochrome (CYP) P450 or P-glycoprotein effects. Examples include the potent inhibitory effects of fluoxetine and paroxetine on CYP2D6 that can elevate serum levels of coadministered drugs including beta-blockers. Paroxetine and sertraline are also potent inhibitors of P-glycoprotein that may elevate levels of digoxin, calcium channel blockers, and some chemotherapy drugs. Web-based drug interaction checkers are readily available.

Regular follow-up appointments are important, particularly in the first 2 to 3 months. The first follow-up visit should generally be 2 weeks after initiating treatment, although high-risk patients with suicide intent or psychotic features should be seen sooner and more frequently. Each visit should include an assessment of symptom change, side effects, and function, as well as an evaluation of psychosocial stressors and supports. There is evidence to suggest that patients who do not obtain at least a 20 percent reduction on a standard symptom rating scale, such as the HAM-D by 2 weeks, have a worse outcome compared to those who do show early symptom reduction. In practical terms, this supports a dose adjustment or confirmation of treatment adherence after 2 to 4 weeks visit.

Although some patients may achieve remission early in the course of treatment (2 weeks or less), the majority of those who do remit will do so after 6 to 8 weeks. Typical rates of remission are between 30 to 40 percent with most antidepressants and response rates are generally in the 50 to 65 percent range, for patients who are not considered treatment resistant.

Consider Measurement-Based Care

Traditionally, the use of clinician-rated scales to evaluate outcomes in the treatment of major depressive disorder has been restricted to clinical trials and other research protocols. More recently, brief clinician-administered and patient self-report measures have been introduced into routine clinical practice (measurement-based care). The advantages of administering a rating scale include longitudinal monitoring and the opportunity to refer to previous evaluations in clinical discussions about areas of improvement or deterioration. However, some clinicians are still reluctant to quantify symptom severity during their time with a patient and feel this detracts from the therapeutic relationship.

Several clinical instruments are readily available and completion times range from 5 to 20 minutes depending on the scale length. Abbreviated

versions such as the HAMD-7 item scale are also available. The Patient Health Questionnaire (PHQ-9) is a self-report scale containing the nine items listed in DSM-5, rated according to their frequency of occurrence. An alternative is the Quick Inventory for Depression Symptomatology (QIDS), which is available in both observer-rated and self-report forms. Two easy to administer scales that evaluate functional outcomes are the Sheehan Disability Scale (SDS) and the Lam Employment Absence and Productivity Scale (LEAPS). An emerging trend involves the use of smart phones or other personal devices to provide remote data capture of objective measures such as sleep–wake patterns, body temperature, or physical activity.

Evaluate Adherence

Adherence is a pivotal component of any therapy. Unfortunately, many patients neglect to take their medication on a frequent basis, missing doses for days or weeks, or discontinuing altogether. This increases the risk of relapse or recurrence. "Forgetting" to take medication as prescribed is more likely to occur as patients begin to feel better and about 50 percent admit to decreased adherence by 3 months. On the positive side, there is evidence that adherence improves in the weeks leading up to and just after a follow-up appointment. Although not currently a routine practice, plasma monitoring of antidepressant blood levels provides an additional method to assess adherence. There is also evidence that a brief telephone contact to monitor symptoms and side effects early in the course of treatment improves adherence and ultimately outcome.

ANTIDEPRESSANT TREATMENT STRATEGIES

Recognize First-, Second-, and Third-Line Approaches

Selecting an antidepressant should reflect a balance between physician judgment and patient preference, taking into account symptom profile, past history of treatment response and tolerability, potential for drug-drug interactions, and in many instances, cost. *First-line* antidepressants include SSRIs, SNRIs, other second generation, and novel antidepressants. These agents have comparable efficacy but superior tolerability to the first-generation MAOI and tricyclic antidepressant (TCA) agents, which are considered *third-line* treatments for major depressive disorder. *Second-line* options include novel antidepressants and adjunctive treatments. Although there are very few direct comparisons of effectiveness/efficacy and tolerability among antidepressants, two major studies contribute most information on treatment outcomes. Combining knowledge from these trials with course specifiers for individual patient episodes of depression can help to refine treatment selection.

Comparative Antidepressant Studies. The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) is the largest published effectiveness trial to evaluate sequential monotherapies and adjunctive antidepressant treatment strategies. It includes medications from each of the above drug classes over four sequential stages in approximately 3,000 patients starting with citalopram in Phase I, with remission as the primary outcome measure. Based on the QIDS scale, 33 percent of those in Phase I achieved remission and 47 percent achieved response. Those who did not achieve remission were eligible to enter the second phase of treatment. In this phase, one arm involved a comparison of switching to another SSRI (sertraline); an SNRI (venlafaxine); bupropion, or CBT. The second arm of Phase II evaluated adjunctive treatment with bupropion, buspirone, or CBT. Despite the large initial sample size, the number of subjects evaluated in the third and fourth phases was substantially smaller and it is difficult to draw conclusions about the relative effectiveness of TCA, MAOI, and other adjunctive strategies. Overall, STAR*D highlighted the relatively low rate of remission (approximately 30 percent) following an adequate trial of citalopram. In addition, the study did not support conventional thinking that switching to a different drug class (e.g., from SSRI to SNRI) following failure with an SSRI was superior to switching within the same class. There was also no difference in outcomes between switching and adjunctive treatment strategies. While the time to achieve remission was longer in those who received CBT, the overall rates of remission did not differ significantly from pharmacotherapy strategies. STAR*D also highlighted the marked reduction in remission rates after two consecutive trials, and the high rates of dropout after each phase.

Among meta-analyses, the most influential study involved a multiple comparisons network meta-analysis of 12 second-generation antidepressants from 117 published trials that collectively included almost 26,000 participants. To incorporate all trials, the authors accepted different severity rating measures (HAM-D, MADRS, and Clinical Global Impression) with "response" as the primary efficacy outcome. The percentage of patients dropping out was a proxy for intolerability. Escitalopram, mirtazapine, sertraline, and venlafaxine were ranked as the top four antidepressants for efficacy, although escitalopram and sertraline were preferred when tolerability and efficacy rankings were combined. While sample size was a major strength in this study, methodological differences across trials may have influenced conclusions. Other meta-analyses suggest comparable efficacy between bupropion and SSRIs, as well as superiority for escitalopram compared to other SSRIs. However, many have argued that differences among first-line antidepressants are minimal.

Course Specifiers for Depressive Episodes. Course specifiers for

depressive episodes may also provide some guidance for treatment selection. In the case of *anxious distress* (tense, restless, fearful, etc.), it would be preferable to select an antidepressant with established efficacy in the treatment of generalized anxiety disorder (GAD). There is also evidence that patients who display *melancholic features* (specifically anhedonia and lack of reactivity) may have a different response to current first-line antidepressants (e.g., SNRI and SSRI) compared to those with nonmelancholic features. Patients with *atypical features* (preservation of mood reactivity, hypersomnia, hyperphagia, weight gain, and a heightened sensitivity to interpersonal rejection) were historically thought to respond better to MAOIs, but it appears that current first-line antidepressants are all equally effective for patients with atypical features. Two specifiers that are most likely to influence treatment selection are *with seasonal pattern*, where light therapy is recommended at least as an adjunctive treatment, and *with mood congruent or mood incongruent delusions*, where either an antidepressant/antipsychotic combination therapy or ECT is recommended.

An additional specifier in DSM-5 is major depressive disorder *with mixed features*. This requires at least two of the hypo(mania) criteria (e.g., grandiosity, expansive mood, and pressured speech) but without meeting full criteria for (hypo)mania. Currently, it is unclear whether patients within this subtype will remain as patients with major depressive disorder or cross over to meet criteria for bipolar disorder. Lurasidone is an appropriate first-line therapy currently for such patients given the evidence from a recent clinical trial that showed improvement in both depressive and hypomanic symptoms in patients with major depressive disorder with mixed features.

Select a First-Line Antidepressant. Early in the course of treatment, most clinicians would select a first-line agent from SSRI, SNRI, other second-generation or novel drug classes. While overall efficacy is likely similar among these agents, differences in their receptor pharmacology and the adverse event profile may aid the clinician in treatment selection. For instance, the SSRI antidepressants share blockade of the serotonin (5HT) transporter as their common mechanism of action, but they differ in their actions on additional receptors. Differences among SNRI antidepressants are even more distinct. Venlafaxine, duloxetine, and desvenlafaxine preferentially block the 5HT transporter, while levomilnacipran predominantly blocks the norepinephrine (NE) transporter. These differences in transporter and receptor affinities make some antidepressants more activating (e.g., bupropion) and others more sedating (e.g., mirtazapine). These differences may allow clinicians to tailor the selection of antidepressants for individual patients to increase acceptability, regardless of the fact that the efficacy may be similar.

Another major consideration when choosing an antidepressant is the side effect profile of the drug and patient sensitivity to side effects. Gastrointestinal and sexual side effects are particularly common with SSRIs, while patients with pre-existing hypertension may be more prone to further elevation in blood pressure with an SNRI such as venlafaxine. Prolongation of QTc interval has also been noted at higher doses with various antidepressants, including citalopram. The relative frequencies of side effects associated with SSRIs and SNRIs are listed in Table 13.7–2. Clinicians need to be vigilant for the emergence of hypo(manic) symptoms during treatment with antidepressants, particularly in patients with a family history of bipolar disorder and in those who display reverse vegetative symptoms as a part of a major depressive episode. It is prudent to offer information about hypo(manic) symptoms to such patients and advise them to make an early appointment for assessment if they experience these symptoms. The risk of antidepressant-induced mood elevations in patients who meet criteria for a major depressive episode with mixed features in the context of major depressive disorder is likely higher, and hence such patients also need to be educated and monitored closely.

Table 13.7–2.
Prevalence of Side Effects among Selective Serotonin-Reuptake Inhibitors and Serotonin- and Norepinephrine-Reuptake Inhibitors

Drug Class	Recommended Dose Range (mg)	Side Effect Frequency	
		10–30%	>30%
SSRI			
Citalopram	20–40	Nausea, dry mouth, sweating	None
Escitalopram	10–20	Male sexual dysfunction and nausea	None
Fluoxetine	20–60	Nausea, dry mouth, somnolence, nervousness, anxiety, insomnia, tremor, anorexia	None
Fluvoxamine	100–300	Dry mouth, headaches, somnolence, agitation, insomnia, sweating, tremor, anorexia, dizziness, constipation	Nausea
Paroxetine	20–60	Nausea, diarrhea, dry mouth, headaches, somnolence, insomnia, sweating, asthenia, male sexual dysfunction, dizziness	None
Sertraline	50–200	Nausea, diarrhea, dry mouth, headaches, somnolence, insomnia, fatigue, tremor, male sexual dysfunction, dizziness	None
SNRI			
Venlafaxine	75–375	Headaches, somnolence, dry mouth, dizziness, nervousness, insomnia, sweating, male sexual dysfunction	Nausea
Desvenlafaxine	50–100	Dry mouth, dizziness, nausea, sweating	None
Duloxetine	30–120	Nausea, dry mouth, constipation, insomnia, male sexual dysfunction	None
Levomilnacipran	20–80	Nausea, dry mouth, headaches, male sexual	None

		dysfunction	
Other Second-Generation and Novel Antidepressants			
Agomelatine*	25–50	None	None
Bupropion	150–450	Insomnia, dry mouth, nausea	Headaches
Mirtazapine	15–60	Dry mouth, constipation, increased appetite, weight gain	Somnolence
Moclobemide*	300–600	None	None
Vilazodone	10–40	Diarrhea, nausea, headaches	None
Vortioxetine	10–20	Nausea	None

Data from Kennedy SH, Lam RW, McIntyre RS, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 3. Pharmacological Treatments. *Can J Psychiatry*. 2016;61(9):540–560; Kennedy SH, Rizvi SJ. Chapter 246: SSRIs and Related Compounds. In: Stolerman I, ed. *Encyclopedia of Pharmacology*. New York: Springer; 2010.

*Not routinely available in the United States.

An increase in reported suicidal ideation in adolescents and young adults during antidepressant clinical trials resulted in "black box warnings" about prescribing SSRIs and other antidepressants to this age group. However, the relationship between increased suicidal ideation and completed suicide remains controversial (see Table 13.6–2). Other rare but serious adverse effects of antidepressants include the 5HT syndrome, increased upper gastrointestinal bleeding (particularly when antidepressants and nonsteroidal anti-inflammatory agents are coprescribed); hyponatremia has also been of concern, especially in elderly depressed patients, and there is a reduction in seizure threshold with most antidepressants, particularly bupropion.

Some second-generation antidepressants with unique mechanisms of action also have noradrenergic effects. Bupropion is both an antidepressant (Wellbutrin) and a medication for smoking cessation (Zyban) with predominantly noradrenergic and dopaminergic effects. It is an activating antidepressant; hence, it is frequently prescribed in combination with a more sedating 5HT enhancing antidepressant. Mirtazapine promotes release of both 5HT and NE through blockade of α2-noradrenergic receptors. It has a sedating profile and is often prescribed to depressed patients with severe insomnia. Mirtazapine is also associated with weight gain. Trazodone is another antidepressant with dual effects on 5HT and NE receptors. Because of its highly sedating properties, it is more often prescribed at subtherapeutic antidepressant doses (50 to 100 mg) as a hypnotic. Nefazodone has a similar mode of action to trazodone with additional NE transporter blocking effects. Despite a favorable efficacy profile, reports of life-threatening hepatotoxicity have resulted in the virtual discontinuation of its use. Moclobemide is a reversible MAO-AI that may be prescribed to treat major depressive disorder without dietary precautions, although it is not available in the United States. For a list of

additional side effects, see Table 13.7–2.

Novel Antidepressants. In addition to the SSRI, SNRI and other "second generation" antidepressants, a number of new antidepressants with additional modes of action are available in many countries to treat depression. Vilazodone is an SSRI with additional 5HT-1A partial agonist effects. This mimics the effect of combining a 5HT-1A partial agonist, such as buspirone, with an SSRI. Favorable aspects of the drug include low rates of sexual side effects, lack of cardiac arrhythmias, and minimal weight gain. Vortioxetine is also a 5HT transporter blocker, with additional receptor effects including 5HT3 and 5HT7 antagonism, as well as 5HT1A and 5HT1B agonism. During the clinical trial development program, measures of cognition including executive function, verbal learning, memory, and processing speed were evaluated and found to be significantly improved. Agomelatine is available in Europe, South America, and many other countries as an antidepressant, but is not available in North America. The mode of action of this antidepressant involves agonism of melatonin-1 and melatonin-2 receptors, combined with 5HT-2C antagonism. There are no direct effects on the 5HT system, although 5HT-2C antagonism facilitates release of dopamine in the prefrontal cortex and may account for its favorable effect on anhedonia and minimal sexual side effects.

Inadequate Response to First-Line Treatments: Next Steps

When a patient has received at least 4 to 6 weeks of treatment with a first-line antidepressant and shown no improvement despite a dose increase, the following options should be considered: (i) further dose increase, particularly when the medication appears to be well tolerated; (ii) switching to a different antidepressant within the same or a different class; and (iii) adjunctive therapy with an established agent such as an atypical antipsychotic, lithium, triiodothyronine or a second antidepressant with a different neurotransmitter action (e.g., bupropion added to an SSRI).

Consider Optimizing before Switching. Usually the first step is dose optimization, particularly when the drug is well tolerated at a relatively low therapeutic dose or if past treatment history suggests tolerance for high doses of other medications. Extending the initial trial beyond 4 to 6 weeks is indicated if there is an incremental but slow improvement during the first 4 weeks.

Potential advantages of switching from one monotherapy to another are cost, single side effect profile, and simpler dosing regimen. These factors are likely to be associated with improved adherence, when compared to adjunctive treatment regimens. Two clinician-monitored methods for switching monotherapies are typically used. The first requires a complete tapering of the initial drug before starting the second drug. This is known

as the "V-method." The second "X-method" involves a gradual tapering of the initial drug, while the second drug is introduced, resulting in a cross over process. This method minimizes discontinuation-emergent symptoms, and is preferred when there are similar pharmacodynamic drug profiles (e.g., switching from one SSRI to another or from an SSRI to an SNRI); however, it should not be used when the potential for a hazardous drug–drug interaction exists (e.g., switching from an MAOI drug to a TCA or SSRI).

Apply Adjunctive Strategies. In clinical practice, the majority of adjunctive treatments are administered sequentially, with the second agent being added only after the first has been ineffective or partially effective. This second-line approach was used at several stages in the STAR*D trial. In a few studies, concurrent combined treatments have been initiated from the outset with inconclusive results. Until there are specific indications for this approach, sequential adjunctive treatments are generally preferred.

Several agents, not primarily indicated for the treatment of depression, including lithium, triiodothyronine (T3), and atypical antipsychotics, are frequently used adjunctively with existing antidepressants. Most studies of lithium involve augmentation of TCAs and the daily dose varied widely (250 to 1,200 mg); nevertheless, two recent meta-analyses support the efficacy of lithium. There is also limited support for the use of T3 at a dose of 25 to 50 μg/day as an augmentation strategy. At the third level of STAR*D, lithium and T3 both produced modest increases in remission. Buspirone, a partial postsynaptic 5-HT1A agonist, was also evaluated in the STAR*D trial and was comparable in effectiveness to bupropion in some but not all outcome measures in the augmentation of citalopram.

There is a large body of evidence to support the adjunctive use of atypical antipsychotics in the treatment of major depressive disorder and also in treatment-resistant depression (TRD). Although they share a common dopamine D2 antagonist action, atypical antipsychotics differ in their affinities for D2 and other receptors. This supports switching from one atypical agent to another in the absence of a favorable response or if adverse effects of the first atypical are intolerable. Based on a series of studies, the combination of olanzapine and fluoxetine demonstrated superiority over either agent alone for TRD. The only SSRI involved in these trials was fluoxetine, but it is likely that olanzapine (5 to 10 mg) would also be effective when combined with other antidepressants. Quetiapine (150 to 300 mg) is the only atypical agent to demonstrate antidepressant efficacy as monotherapy in the treatment of major depressive disorder and is also effective when combined with a range of antidepressants in previously treatment-resistant patients. Augmentation with risperidone was efficacious in two double-blind, placebo-controlled trials. Aripiprazole (5 to 15 mg) has been shown to be an effective

adjunctive therapy in major depressive disorder, where response and remission rates were double the rates in those who received "add-on" placebo. Interestingly, patients who either worsened or failed to improve with a first antidepressant treatment still showed significant benefit with adjunctive aripiprazole. Adjunctive treatment with ziprasidone in a TRD trial was not significantly better than continuation of the SSRI (sertraline), supporting the view that atypical antipsychotic adjunct efficacy is not a class effect but depends on the individual receptor profiles of each agent. Although atypical antipsychotics provide a valuable additional treatment option for TRD, their potential to cause serious side effects, including weight gain, dyslipidemia, hyperglycemia, extrapyramidal symptoms and QTc interval prolongation, is of concern, particularly when long-term treatment is anticipated. An unanswered question is how long to continue any adjunctive agent particularly when the additional therapy carries a hazardous side effect burden (e.g., weight gain or hyperlipidemia).

Other strategies to consider when patients have failed to respond to the above strategies include the addition of psychotherapy, with most evidence supporting CBT or modified versions such as cognitive behavioral analysis systems psychotherapy. ECT is very effective in antidepressant treatment refractory patients with response rates as high as 70 percent. ECT also remains an established treatment of certain groups of depressed patients, particularly those with acute suicidal ideation or with mood congruent or incongruent delusions. Other neurostimulation approaches including repetitive transcranial magnetic stimulation (rTMS) and vagal nerve stimulation may also be considered as adjunctive treatments.

Consider Third-Line Antidepressants

Both tricyclic and MAOI antidepressants are restricted in their use today because of tolerability and safety concerns. Amitriptyline, the most evaluated TCA, has mixed reuptake inhibition effects on both 5HT and NE transporters with additional cholinergic, histaminergic, serotonergic, and noradrenergic receptor actions. These multiple effects are responsible for the wide range of side effects including sedation, weight gain, urinary retention, and cardiac conduction defects. Nortriptyline is a metabolite of amitriptyline that is often used to treat late-life depression and has a less extreme side effect profile. Imipramine, desipramine, and clomipramine are also TCAs that are typically reserved for third-line use, although clomipramine is indicated for the treatment of OCD and not major depressive disorder in the United States. In all cases, plasma monitoring may be helpful in determining appropriate doses.

Tranylcypromine, phenelzine, and isocarboxazid are irreversible inhibitors of both MAO-A and MAO-B. While this mode of action provides a broad-spectrum inhibition of monoamine reuptake and a favorable antidepressant effect, it also increases the risk of orthostatic hypotension,

and in the presence of tyramine-rich foods may produce a life-threatening hypertensive crisis. This necessitates a low tyramine diet during MAOI treatment. Hence, these agents are generally reserved for otherwise treatment-resistant patients. Selegiline is an irreversible inhibitor of MAO-B that has been used as an adjunct therapy for Parkinson disease and is available in both oral and transdermal patch delivery systems to treat major depressive disorder. Its use is generally restricted to TRD patients.

Recognize Experimental Antidepressant Options

Evidence has emerged that ketamine, given intravenously, provides a rapid but transient antidepressant response, even in formerly treatment-resistant depressed patients. Similar preliminary results have been reported for nitrous oxide. Both novel treatments support the growing interest in glutamate and NMDA receptors in the pathophysiology of depression. Similarly, evidence that proinflammatory cytokines such as interleukin 1 (IL-1) and IL-6 are frequently elevated in depression has stimulated interest in the potential role of anti-inflammatory agents as antidepressants. Celecoxib, minocycline, and infliximab are examples of anti-inflammatory agents with potential antidepressant properties. So far these remain experimental treatments, as does deep brain stimulation (DBS) to targeted brain areas.

MAINTAINING REMISSION

Address Risk Factors for Relapse and Recurrence

The primary aim of maintenance treatment is to prevent the occurrence of a new episode (recurrence). It is appropriate for patients with major depressive disorder who have experienced recurrent episodes or PDD but not following a single-episode major depressive disorder. Maintenance medication is more effective than placebo in virtually all studies to date. Given the high rate of relapse for individuals who have had three or more episodes, they should receive prolonged maintenance therapy.

Patients who consistently relapse following antidepressant discontinuation, those with recurrent psychotic episodes, persistent residual symptoms, or a strong family history of mood disorders should be considered as candidates for prolonged maintenance therapy. In any case, clinicians and patients need to decide collaboratively whether to initiate maintenance treatment or to maintain contact and rapidly reinstate treatment if there are emergent symptoms. If a new episode develops when the patient is free of treatment, then early intervention shortens the length of the new episode. There are virtually no data to suggest when to discontinue maintenance medication, although a study examining the impact of discontinuing venlafaxine after either 1 year or 2 years of

maintenance treatment supported the benefits of longer-term treatment. When the decision to discontinue treatment is made, careful monitoring is required, particularly for the first 6 months as this is a high-risk period for recurrence.

TREATMENT OF BIPOLAR DISORDERS

The discovery of the utility of atypical antipsychotics and some anticonvulsants such as valproate and lamotrigine for bipolar disorder in the 1990s heralded a new era in the pharmacological treatment of bipolar disorder. The availability of several newer treatments for the management of bipolar disorder has presented clinicians with options to tailor the treatment to suit individual patient needs based on the efficacy, adverse events, and patient preferences. The past two decades have also witnessed significant progress in understanding the adjunctive role of psychological treatments in managing bipolar disorder.

This section will address the foundations of the management of bipolar disorder and specific treatment options for the management of mania, bipolar depression, and prophylaxis of bipolar disorder. Clinical recommendations are provided for the management of each phase of bipolar disorder with guidance on first-, second-, and third-line choices, but it is important to acknowledge that these recommendations are mainly applicable for bipolar I disorder. This section also covers briefly the management of bipolar II disorder as well as the management of comorbidity in bipolar patients. The reader is advised to refer to other sections in the book for managing bipolar disorder during pregnancy and postpartum period as well as in children and adolescents and older adults.

FOUNDATIONS OF MANAGEMENT

Diagnose and Intervene Early

Although mania is the defining feature of bipolar I disorder, depressive episodes and subsyndromal depressive symptoms are more common in bipolar I disorder. Indeed, prospective cohort studies suggest that patients with bipolar I disorder spend more time (a ratio of 3:1) experiencing depressive episodes/symptoms than manic episodes/symptoms. Depressive symptoms are even more problematic for bipolar II patients as they spend 37 times more days experiencing depressive symptoms than hypomanic symptoms.

Given this, patients with bipolar disorder are more likely to seek help when they are depressed and often do not volunteer information about previous manic/hypomanic episodes. Hence, clinicians must be vigilant to the possibility of bipolar disorder in every patient presenting with depressive symptoms. Studies suggest that about two-thirds of patients with bipolar disorder are misdiagnosed initially, and major depressive

disorder is the most common misdiagnosis. While screening instruments such as the Mood Disorders Questionnaire are useful, they do not replace a thorough clinical assessment that systematically probes for previous history of manic/hypomanic symptoms. Clinicians should have a high index of suspicion for a diagnosis of bipolar disorder in patients presenting with a history of substance abuse (e.g., cocaine or alcohol), family history of bipolar disorder, or reverse vegetative symptoms (e.g., anergia, hypersomnia, and hyperphagia) and/or psychotic features as a part of depressive episode.

Early intervention has the potential to improve outcomes in bipolar disorder. Indeed, studies in first episode mania patients suggest that those that remain episode free after their first manic episode show improvements (reversal of cognitive deficits) in cognitive functioning and minimal decline in age-related gray matter volumes.

Conduct Systems Review, Baseline Investigations, and Implement Safety Monitoring for Psychotropic Medication

In addition to obtaining a comprehensive psychiatric history to diagnose bipolar disorder, clinicians are advised to conduct a systems review of medical history with particular attention to personal and family history of cardiovascular, endocrine, and metabolic disorders as the incidence of medical comorbidities is higher in bipolar patients compared with general population. Many psychotropic medications have a tendency to cause weight gain and some cause metabolic syndrome and therefore, it is important to conduct some baseline laboratory investigations and obtain weight and waist circumference either before when possible or within few days of the initiation of specific pharmacotherapy (see Fig. 13.7–1 for details) for bipolar disorder. Patients that require treatment with some psychotropic medications such as lithium or valproate need additional specific lab investigations prior to initiating treatment. During the follow-up visits, clinicians are advised to routinely monitor for medication-induced adverse events at each scheduled clinical follow-up and institute appropriate strategies to address them. Patients should have weight, appropriate laboratory investigations, and serum levels of lithium or valproate measured at regular intervals as outlined in Figure 13.7–1.

Treat the Disorder and Not Just Acute Episodes

Pharmacotherapy is the cornerstone of the management of bipolar disorder. When initiating treatment of acute episodes, while stabilization of acute episodes is the immediate goal, it is essential for clinicians to bear in mind that bipolar disorder is a recurrent illness, and hence prevention of further mood episodes is an equally important consideration in treatment selection. Therefore, when choosing a medication to treat an acute episode,

those that have proven efficacy in treating not only the index episode but also in preventing the recurrence of both manic and depressive episodes should be considered first before selecting the medications that have efficacy for only acute episodes. There is growing consensus that mood episode recurrences are "toxic" to the brain and are associated with cumulative reductions in brain volumes, lower brain-derived neurotrophic factor (BDNF) levels, cognitive impairment, and poorer response to pharmacotherapy, and thus prevention of recurrence of mood episodes is paramount in the management of bipolar disorder. Consensus is growing that prophylaxis should be offered for all patients with bipolar I disorder including for those after the first manic episode as prospective naturalistic cohort studies suggest that 60 to 75 percent of first episode manic patients experience a recurrence within 4 years. Adjunctive psychological treatments such as psychoeducation promote treatment adherence and reduce the risk of relapse of mood episodes in bipolar patients, and hence should be initiated as soon as the patient is amenable to such strategies.

Treat Acute Episodes to Full Remission

Acute manic, hypomanic, and depressive episodes must be treated to full and sustained remission. This is because several studies now suggest that residual subsyndromal mood symptoms not only significantly affect quality of life and functioning but also strongly predict relapse of subsequent full-blown mood episodes. Patients who recovered from an index mood episode with residual mood symptoms experience a new episode 3.4 times faster than those that recovered from an index episode with no residual symptoms. Median time to recurrence of a mood episode is five times faster in those with residual symptoms compared with those with full recovery from an index episode.

Assess for and Treat Comorbidities

Psychiatric and physical comorbidities are common in bipolar disorder. While the estimates of the prevalence vary, at least 75 percent of patients with bipolar disorder have at least one lifetime comorbid psychiatric disorder. Alcohol or substance use disorders are present in at least 40 percent of patients, and some type of anxiety disorder is present in 50 to 75 percent of patients. There is evidence that patients with alcohol and substance use disorders or anxiety disorder comorbidities take longer time to recover from an index episode and are more likely to relapse compared with those without these comorbidities. Hence, the optimal management of bipolar disorder should include not only treatment of mood symptoms/episodes associated with this condition but also assessing for comorbidities and instituting appropriate pharmacological/psychological management strategies to address comorbidities in order to improve

outcomes. During the maintenance treatment, clinicians should routinely monitor for onset of new psychiatric, substance or alcohol use comorbidities, and manage them as clinically indicated.

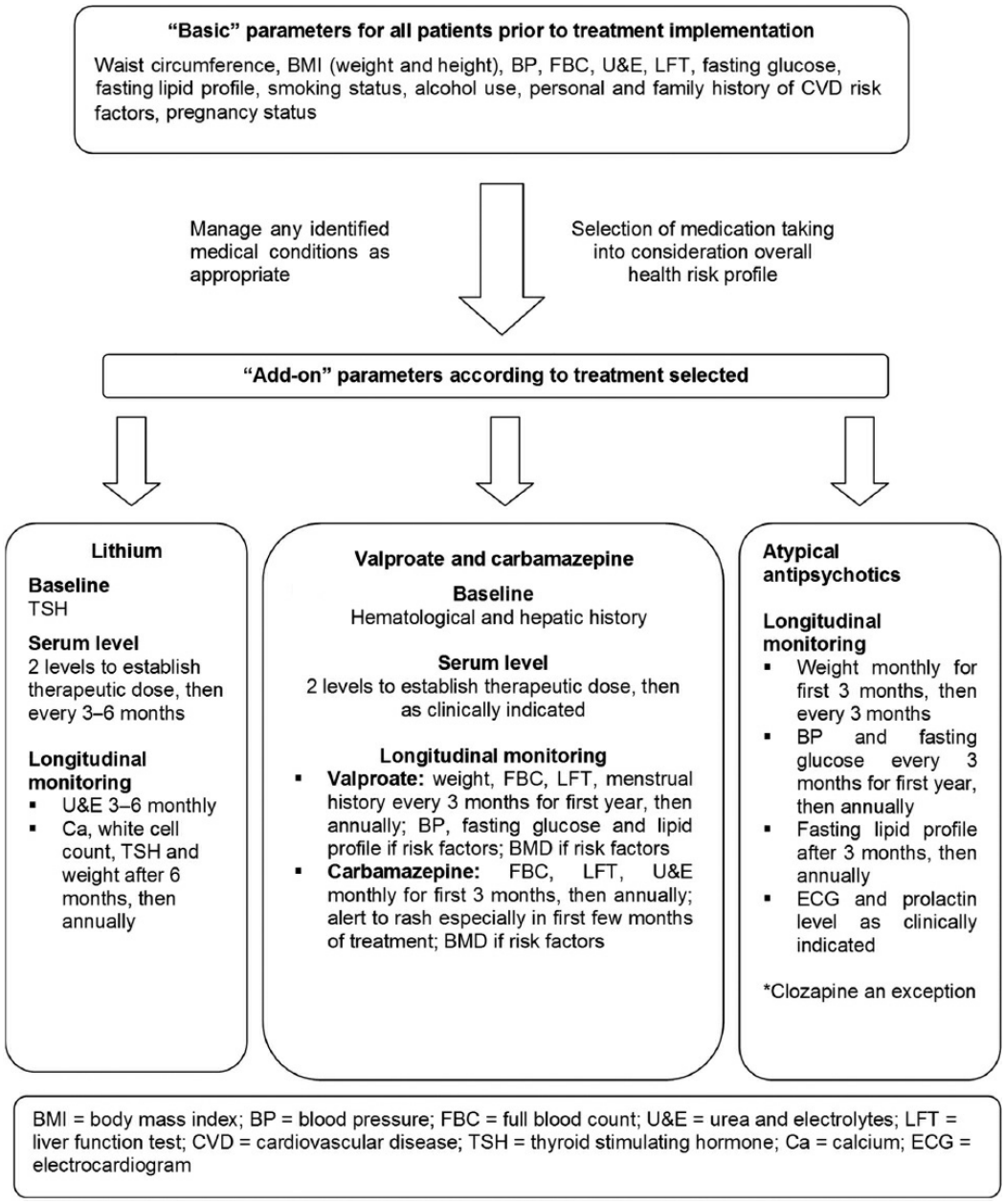

FIGURE 13.7–1. Recommendations for treatment monitoring in bipolar disorder. (Adapted from ISBD safety monitoring guidelines.)

Address Cognitive Impairment

Patients with bipolar disorder often complain of cognitive difficulties. A thorough clinical assessment should be conducted in such patients to address contributory factors such as comorbid substance use, psychotropic medications-induced confusion or sedation, lithium-induced hypothyroidism, or valproate-induced hyperammonemia, etc. for cognitive difficulties. Evidence has accumulated over the past 15 years to suggest that

cognitive dysfunction can be an inherent part of bipolar disorder for a substantial proportion of patients with bipolar disorder. Impairments are seen in several domains including in processing speed, verbal learning, working memory, and executive function and such impairments are present not only during acute mood episodes but also even during euthymic periods. The effect sizes for magnitude of impairment during euthymia are in the range of 0.6 to 0.8. Cognitive impairments particularly in the domains of verbal learning and executive function correlate with functional impairments. There is increasing interest in examining the efficacy of pharmacological treatments and cognitive remediation to address cognitive dysfunction. While evidence for efficacy is limited, there are no risks to attempting cognitive and functional remediation in such patients.

Monitor for Treatment Adherence

While the rates of treatment nonadherence vary depending upon how it is defined, there is evidence that up to 60 percent of patients with bipolar disorder are nonadherent or poorly adherent to pharmacotherapy. The consequences of poor adherence to pharmacotherapy include illness destabilization with increased relapse/recurrence of mood episodes, high suicide risk, increased hospitalization, disability, and poor functioning. Several factors such as poor insight/denial of illness, concerns about or experience of side effects of medications, inadequate efficacy of medications, younger age, low level of education, cognitive impairment all contribute to poor treatment adherence. It is important for clinicians to be cognizant of demographic variables and explore illness-related factors, attitudes toward illness and treatment-related factors that predict poor adherence and address them routinely at clinical follow-up visits. There is evidence that psychoeducation improves treatment adherence. More recently, customized treatment adherence programs have been developed that include modules on psychoeducation, substance abuse, medication routines, and communication with providers, and these programs can be tailored to address the specific reasons for the treatment of nonadherence in individual patients.

Monitor for the Emergence of Mood Symptoms and Suicide Risk

There is evidence that bipolar patients with subsyndromal mood symptoms are more likely to relapse into a full-blown mood episode compared with those without such symptoms. Therefore, it is prudent for clinicians to routinely screen for emergent subsyndromal symptoms at each clinical visit. Clinicians may find Young Mania Rating Scale (YMRS) and MADRS particularly helpful in systematically assessing and quantifying subsyndromal symptoms. The reasons for the emergence of such symptoms

such as poor treatment adherence, comorbidity, or partial efficacy of ongoing treatments should be identified when possible and addressed with psychoeducation or changes to pharmacotherapy as appropriate in order to rapidly resolve such symptoms. Subsyndromal depressive symptoms can also be treated with psychological strategies such as with interpersonal and social rhythm therapy (IPSRT) or CBT.

Clinicians should also routinely monitor for the emergence of suicidal ideation not only when treating acute episodes but also during the maintenance treatment. This is because suicidal ideation is common in patients with bipolar disorder and that between 25 to 50 percent of patients with bipolar disorder make a suicide attempt at least once during their lifetime. The rates of lifetime risk of suicide deaths in bipolar patients range from 8 to 19 percent with more recent studies in less severely ill populations suggesting lower rates. Regardless, the standardized mortality ratios for suicide in bipolar patients range from 10 to 30, which means that bipolar patients are 10 to 30 times more likely to die of suicide compared with the general population. A recent International Society for Bipolar Disorders (ISBD) Task Force identified several risk factors for suicide attempts and suicide deaths (see Table 13.7–3) in bipolar disorder.

Clinicians should be more vigilant for suicidal ideation in those with increasing number of risk factors. The risk of suicide attempt and suicide death is higher early in the course of the illness, during acute mood episodes or hospitalization, and within 4 to 12 weeks of discharge from the hospital. When suicidal ideation is endorsed, clinicians must take appropriate steps immediately to ensure the safety of the patient and these may include hospitalization, preventing/reducing access to methods of suicide (e.g., medication overdose), establishing therapeutic alliance, and involving family members/friends when possible in care, and addressing the illness related and psychosocial factors that led to the emergence of suicidal ideation.

PHARMACOLOGICAL TREATMENTS

This section briefly reviews the evidence for the treatment options for various phases of bipolar disorder and provides clinical recommendations.

ACUTE MANIA: TREATMENT OPTIONS

Lithium

There is robust evidence for efficacy of lithium in treating acute mania. When adequately dosed, therapeutic effect of lithium becomes apparent within 7 days as the separation from placebo was seen on as early as day 7 in clinical trials. While the response rates to lithium varied between the trials, on average, about 50 percent of patients met criteria for response to lithium in these 3 week trials.

Lithium is typically started at 900 mg/day and serum levels measured after 5 days of consecutive therapy, 12 hours after the last dose. Serum levels between 0.8 and 1.2 mEq/L are likely required to treat acute mania effectively, although lower levels may be appropriate in elderly. Lithium levels of ≥1.5 mEq/L can lead to symptoms of lithium toxicity and this should be treated as a medical emergency. Patients should be educated and clinically monitored for signs of lithium toxicity that may include gastrointestinal (vomiting and diarrhea), neurological (coarse tremor, agitation, dysarthria, drowsiness, lethargy, ataxia) renal (polyuria and polydipsia), and cardiovascular (syncope, dizziness, arrhythmias) symptoms.

Anticonvulsants

Carbamazepine. Carbamazepine is as effective as lithium and neuroleptics in treating acute mania. Carbamazepine extended release (ER) is also effective with response rates between 40 and 60 percent that are similar to the aggregate response rates of 52 percent observed with carbamazepine in older studies.

Carbamazepine is typically started at 100 to 200 mg bid and the dose gradually titrated upward to achieve therapeutic effect. Most patients will need between 600 and 800 mg/day. Therapeutic effect is observed typically within 1 week to 10 days. There is no evidence to suggest a relationship between serum carbamazepine levels and acute therapeutic efficacy in mania.

Oxcarbazepine. While oxcarbazepine may be better tolerated than carbamazepine, the data for its efficacy in treating acute mania is conflicting. While four small controlled studies in adult manic patients support the efficacy, a recent larger clinical trial in children and adolescents with mania that compared oxcarbazepine with placebo found no differences in improvement in manic symptoms between the two groups. Given the lack of larger placebo-controlled trials in adults with modern methodology and a negative trial in children and adolescents with acute mania, further studies are clearly needed to determine the efficacy of oxcarbazepine for acute mania. This statement is consistent with a recent Cochrane review which concluded that there is currently insufficient evidence for efficacy of oxcarbazepine in treating acute mood episodes in bipolar disorder.

Divalproex/Valproate. Divalproex consists of sodium valproate and valproic acid in a 1:1 molar ratio. The landmark trial by Bowden and colleagues (1994) that compared the efficacy of divalproex with placebo and lithium led to the approval of divalproex for the treatment of acute mania. Divalproex is as effective as lithium in adult manic patients. Two out of

three studies that compared divalproex with olanzapine showed similar efficacy at 3 weeks. Divalproex has similar efficacy to quetiapine in treating acute mania in adolescents.

Table 13.7–3.
Factors Associated with Suicide Attempts and Suicide Deaths in Bipolar Disorder

Variables	Increased Likelihood of Suicide Attempts in Bipolar Disorder	Increased Likelihood of Suicide in Bipolar Disorder
Sociodemographics		
Sex	Female	Male Females have higher SMR
Age	Younger Older age—higher lethality attempts	No clear effect Older age—higher ratio of deaths/attempts
Race	Minorities—specific to youth only	
Marital status	Single, divorced, single parents	
Religious affiliation	Discordant results	
Characteristics of Bipolar Disorder		
Age of illness onset	Younger age	
Duration of illness	Discordant results	
Polarity of 1st episode	Depression Mixed symptoms Mania—more violent attempts	
Predominant polarity	Depressive	
Polarity of current episode	Depressive episodes Mixed states	Depressive episodes Mixed states Manic states with psychotic features
Other mood episode characteristics	Mixed features Greater number and severity of episodes/rapid cycling Anxiety Atypical features Suicidal ideation	Hopelessness Psychomotor agitation
Psychosis at any time	Discordant results	No effect
Bipolar disorder subtype	No clear effect	No clear effect
Comorbidities		
Psychiatric comorbidity	Substance use disorder Cigarette smoking Coffee intake Anxiety disorder Obesity or high BMI Eating disorder	Substance use disorder—no clear effect Anxiety disorder
Personality disorders and characteristics	Personality disorders—including borderline or cluster bipolar disorder in particular Aggression or irritability Impulsivity (discordant results)	

Other Clinical Variables		
First-degree family history	Mood disorders Bipolar disorder Suicide	Mood disorders Bipolar disorder Suicide
Prior suicide attempts	Presence	Presence
Early life trauma	Childhood abuse Higher early life stress	
Psychosocial precipitants	Interpersonal problems Occupational problems Bereavement	Present within a week prior to death
Sexual dysfunction	Presence	

SMR, standardized mortality ratio.

From Schaffer A, Isometsä ET, Azorin J, et al. A Review of Factors Associated with Greater Likelihood of Suicide Attempts and Suicide Deaths in Bipolar Disorder: Part II of a Report of the International Society for Bipolar Disorders Task Force on Suicide in Bipolar Disorder. *Aust N Z J Psychiatry*. 2015;49(11):1006–1020.

Therapeutic effect of divalproex is apparent by day 7, although studies with divalproex oral loading (20 mg/kg/day or 30 mg/kg/day on days 1 and 2 followed by 20 mg/kg/day) suggest more rapid onset of action with effect apparent by day 5 or sooner. Posthoc analyses of data from these trials suggest that divalproex has a broader spectrum of efficacy with effect in manic patients with and without concurrent depressive symptoms, psychotic symptoms, rapid cycling, and regardless of number of prior mood episodes.

There is a linear relationship between serum valproate levels and therapeutic efficacy for acute mania with higher serum levels yielding greater efficacy. The evidence suggests that optimal efficacy may not be obtained when serum levels are below 94 μg/mL. Hence, clinicians should titrate the dose of valproate to achieve serum valproate levels to this target range. If tolerability is not a concern, the doses could be further increased if necessary as some patients may require higher levels for optimal efficacy.

Other Anticonvulsants. Lamotrigine, gabapentin, topiramate, and licarbazepine are not effective in treating acute mania as none separated from placebo in controlled clinical trials in improving manic symptoms. Lamotrigine or topiramate adjunctive therapies are also ineffective in treating acute mania.

Other anticonvulsants such as tiagabine, pregabalin, levetiracetam, felbamate, and zonisamide have not been assessed in controlled trials for acute mania.

Conventional/First-Generation Antipsychotics

Haloperidol is the most widely investigated first-generation antipsychotic for acute mania. Several modern placebo-controlled trials of atypical

antipsychotics such as risperidone, olanzapine, quetiapine, ziprasidone, and aripiprazole used haloperidol as an active comparator. In all these studies, haloperidol was more effective than placebo confirming the efficacy of haloperidol in treating acute mania.

The efficacy of haloperidol was apparent at early as day 2 in some trials. The doses of haloperidol used in these trials ranged from 2 mg to 15 mg per day, although in the ziprasidone trial up to 30 mg/day of haloperidol was allowed. These studies demonstrated that haloperidol was as effective as risperidone, olanzapine, and aripiprazole and more effective than quetiapine and ziprasidone. The extrapyramidal symptoms are common with haloperidol, and patients with bipolar disorder are more prone to movement disorders such as tardive dyskinesia. Furthermore, more patients treated with haloperidol experienced depression in some of these trials compared with placebo or atypical antipsychotics.

It is believed that the efficacy of haloperidol for acute mania is related to its ability to reduce dopamine transmission through the blockade of dopamine D2 receptors. Since all conventional antipsychotics reduce dopamine transmission through their effects on D2 receptors, they all are likely effective in treating acute mania. Chlorpromazine has been compared with lithium and found to have similar efficacy overall, although it appears to be more effective in "highly active" manic patients, which is likely related to its sedating effects while lithium has better efficacy in treating core manic symptoms. Loxapine-inhaled formulation was recently assessed for efficacy in treating agitation in acute manic patients and found to be effective, but its utility in treating core manic symptoms has not been assessed in placebo-controlled trials.

Haloperidol in combination with lithium or valproate is more effective than lithium or valproate monotherapy. Similarly, contrary to what clinicians might believe, valproate add-on to haloperidol/perphenazine is more effective than monotherapy with haloperidol/perphenazine.

Atypical Antipsychotics

Monotherapy. Several atypical antipsychotics are currently in use for the treatment of psychosis. Of these, risperidone, olanzapine, quetiapine, ziprasidone, aripiprazole, paliperidone, and asenapine have proven efficacy and hence approved for the treatment of acute mania in several countries. Table 13.7–4 provides a summary of the efficacy of atypical antipsychotics and the dose ranges commonly used in clinical practice. Clozapine and lurasidone are likely effective but no placebo-controlled trials have assessed their efficacy for acute mania to date. Cariprazine, a new atypical antipsychotic candidate with D2 and D3 partial agonist properties, is effective and was recently approved by the FDA for the treatment of mania.

Table 13.7–4.
Atypical Antipsychotics for Bipolar Disorder: Efficacy Summary and Dose Ranges

		Efficacy					
				Prophylaxis			
Drug	Dose Range (mg)	Acute Mania	Acute Bipolar Depression	Mood Episodes	Mania	Depression	Comments
Olanzapine	5–20	Yes	?Yes[1] (see comments)	Yes	Yes	Yes[2]	[1]Improvem was seer mostly ir sleep, ap and inne tension b in core depressiv sympton [2]Magnitude benefit f depressi than for
Risperidone	1–6	Yes	No data	[1]Yes	[1]Yes	No	[1]Prophylact efficacy v demonst with Risp Consta b studies v oral risperido
Quetiapine	300–800[1]	Yes	Yes	Yes	Yes	Yes	[1]For bipolar depressi 300 mg/ as effect 600 mg/ (see text guidance Appears to equal eff in prever both ma and depr
Ziprasidone	80–160	Yes	No	[1]Yes	[1]Yes	No	[1]Prophylact demonst for adjur therapy l data for monothe
Aripiprazole	15–30	Yes	No	Yes	Yes	No	
Paliperidone	6–12	Yes	No data	[1]Yes	Yes	No	[1]Less effect than olanzapi preventi

							mood ep
Asenapine	10–20	Yes	No data	Yes	?Yes	?Yes	Numerically fewer pa in the asenapin group ha relapse c manic ar depressiv episodes the diffe were not significar
Lurasidone	40–120	No data	Yes	Studies underway	Studies underway	Studies underway	Effective in depresse patients mixed fe
Cariprazine	3–12[1]	Yes	[1]Yes	No data	No data	No data	[1]For bipolar depressi 1.5–3 mg is recomm while for mania, u 12 mg/da appropri

The only study that compared two atypical antipsychotics head-to-head found no difference in efficacy between risperidone and olanzapine for acute mania. While other atypical antipsychotics have not been compared directly in double blind, controlled trials, there is no reason to believe that there are any clinically meaningful significant differences in efficacy between various atypical antipsychotics, given the broad similarity in effect sizes and placebo-corrected response rates.

If a medication is effective rapidly, it can shorten the duration of the manic episode and the hospital length of stay. When atypical antipsychotics are dosed adequately, some improvement in manic symptoms is apparent typically within the first week with all atypical antipsychotics. While there is a perception that sedating atypical antipsychotic drugs may have more rapid onset of action, this is not borne out in clinical trials. For instance, although aripiprazole and ziprasidone are less sedating, their therapeutic effects in mania were apparent at first post-baseline measurements of day 4 and day 2, respectively in clinical trials.

Depressive symptoms are common in acute manic patients and several trials with atypical antipsychotics have included some patients with DSM-IV mixed episodes. Posthoc analyses from these trials suggest that olanzapine, asenapine, aripiprazole, and ziprasidone are effective in treating manic symptoms in patients with mixed episodes/dysphoric

mania. Further, a meta-analysis suggests that atypical antipsychotics as a group are effective in treating acute mixed episodes and the effect size for their efficacy in treating mixed episodes is in the medium range that is similar to their magnitude of effect in manic patients. As to comorbid depressive symptoms, they improved with atypical antipsychotics in most studies but significant separation from placebo occurred only in few studies in favor of atypical antipsychotics (e.g., risperidone and quetiapine). Based on this, it is reasonable to conclude that atypical antipsychotics unlike conventional first-generation antipsychotics at the very least do not worsen or induce depressive symptoms in manic patients. DSM-5 has eliminated the mixed episodes category and introduced mixed features as a course specifier for manic as well as depressive episodes. While no clinical trials of atypical antipsychotic have been conducted to date in manic patients with mixed features as a course specifier, many patients who meet these criteria would have been included in previous trials. Indeed, a recent posthoc analysis of asenapine trials using DSM-5 mixed features specifier criteria supports the efficacy of asenapine in this population. It is likely that other atypical antipsychotics are also effective.

Combination/Add-On Therapy. There is evidence that as many as 80 percent of manic patients in clinical settings are treated with a combination of a mood stabilizer and an antipsychotic. In some cases, both medications are started at the same time (combination therapy) while in other cases, one medication is started and the other is added after a few days (add-on therapy).

In clinical trials of add-on therapy, although patients met entry criteria for mania severity, refractoriness to existing therapy was not established prospectively in many studies. Hence, it is unknown, if the add-on studies truly assessed the efficacy in refractory manic patients.

Regardless, there is evidence that response rates are greater when manic patients are treated with a combination/add-on of risperidone, olanzapine, quetiapine, aripiprazole, or asenapine with lithium or valproate compared with lithium or valproate monotherapy. On average, 20 percent more manic patients will respond with combination/add-on therapy compared with lithium or valproate monotherapy. While different combinations treatments are not directly compared, based on differences in response rates observed between monotherapy and combination therapies in various studies, there does not appear to be any clinically meaningful differences between different combinations of atypical antipsychotic and lithium or valproate. The only exception is ziprasidone adjunctive therapy that was not superior to lithium or valproate monotherapy in the only clinical trial that assessed this strategy.

There is preliminary evidence that adding an atypical antipsychotic to carbamazepine does not confer additional efficacy. For instance,

risperidone or olanzapine combined with carbamazepine is no more effective than carbamazepine monotherapy. This is likely because carbamazepine through induction of hepatic microsomal enzymes reduces serum levels of antipsychotics and thus affects their efficacy.

Some clinicians believe that antipsychotic monotherapy is as effective as the combination strategy. There is preliminary evidence to suggest that this is clearly not the case. For instance, a large double-blind trial suggested that lithium given in combination with quetiapine XR was more effective than quetiapine XR monotherapy. Similarly, valproate in combination with haloperidol/perphenazine was more effective than monotherapy with an antipsychotic.

It is, however, important to note that greater efficacy of combination therapy comes with a trade-off in terms of adverse events as combination therapy is often associated with more adverse events than monotherapy.

Miscellaneous Agents

Tamoxifen, a drug used in the treatment of breast cancer, is a protein kinase c (PKC) inhibitor. There is evidence that treatment with lithium and valproate alters PKC signaling pathway and more specifically reduces specific isoforms of this enzyme. There is evidence from four small double-blind, placebo-controlled trials (2 monotherapy and 2 adjunctive therapies) that tamoxifen is effective in treating acute mania. However, clinicians must be aware that tamoxifen use in patients with breast cancer is associated with two to sevenfold increase in endometrial carcinoma. Further, there is no longer term safety data for tamoxifen in bipolar patients.

Verapamil was reported to be effective in two small earlier controlled trials and some open trials but a larger double-blind, placebo-controlled trials found no evidence for efficacy of verapamil in treating acute mania. Two very small trials suggest that nimodipine may be useful in treating manic symptoms, but larger trials are needed before any conclusions about its efficacy can be drawn. Calcitonin was reported to be effective in treating acute manic symptoms in a trial conducted in 80s, but a more recent double-blind trial that compared calcitonin nasal spray versus saline nasal spray in refractory manic patients found no difference between the two groups.

CLINICAL RECOMMENDATIONS FOR THE MANAGEMENT OF ACUTE MANIA

General Principles of Management

Manic episodes can be a part of bipolar I disorder or secondary to an underlying medical condition, neurological condition, head injury, substance use disorder (e.g., cocaine or amphetamines), or induced by drugs such as steroids. When possible, verbal de-escalation techniques

should be tried to calm the manic patient and obtain a thorough psychiatric and medical history. Manic patients are often uncooperative and lack insight; collateral information is vital in such cases. Patients should be assessed for the risk of aggressive behavior, severity of mania, presence of psychotic features, and insight, all of which inform the clinician about the treatment setting for further management. Patients with severe mania and aggressive behavior with little insight will require hospitalization, while those with mild-to-moderate mania may be treated in an out-patient setting if adequate family support is available and if the patient has some insight.

A thorough review of medication history should be conducted. If a patient had been taking antidepressant medication, this should be discontinued. Similarly, other medications that might be potentially contributing to manic symptoms such as stimulants or steroids should be discontinued. A physical examination and appropriate baseline medical investigations should be conducted before definitive pharmacological treatment is instituted, although this may not be possible in some cases. If the patient has a history of substance use, this should be addressed simultaneously while mania is being treated. Psychoeducation should commence as soon as the patient is behaviorally stabilized and is amenable.

Management in an Emergency Setting

Patients presenting with a manic episode in an emergency setting are often agitated, irritable, uncooperative, and lack insight. In some cases, they may be verbally or physically aggressive and may escalate to becoming assaultive and violent. Hence, ensuring the safety of the acutely manic patient and those around is a high priority. It is important to move the patient to a low stimulus and comfortable environment as soon as possible.

The severity of agitation and aggressive behavior and patient's willingness to accept treatment dictate whether medications can be given orally or in an injectable form. Oral medications are less likely to disrupt the therapeutic alliance and hence are preferable but have the disadvantage of slower onset of action. Further, although oral administration of benzodiazepines such as lorazepam, first-generation antipsychotics such as haloperidol and loxapine, and second-generation antipsychotics (risperidone, olanzapine, quetiapine, aripiprazole, ziprasidone, asenapine) is a common clinical practice in the emergency setting to control agitation and aggressive behavior, they are not widely studied in controlled clinical trials for efficacy in treating agitation and aggressive behavior in mania. Nevertheless, clinicians often employ this strategy in many parts of the world. While haloperidol and several second-generation antipsychotics have proven efficacy in treating acute mania, there is little evidence to suggest that lorazepam is effective in treating core manic symptoms although this may be useful in calming the patient. Hence lorazepam

monotherapy for controlling agitation in mania is not recommended but this can be given in conjunction with an antipsychotic and in particular a first-generation antipsychotic such as haloperidol to reduce the risk of extrapyramidal symptoms and to enhance the calming effect.

While lithium and carbamazepine are effective in treating acute mania, they have not been studied for their efficacy in treating acute agitation and aggressive behavior in acute mania, and hence are not recommended as monotherapy for treating agitation in an emergency setting. Valproate oral loading dose has a similar faster onset of action as the antipsychotics in treating acute mania but has not been studied for controlling agitation. For patients who have a tendency to "cheek" the medication, orally disintegrating tablets (ODT) or liquid form of antipsychotics is preferable. However, there is no convincing evidence to support that these forms of administration have a faster onset of action in comparison to conventional tablet forms.

Two novel forms of oral administration of antipsychotics have shown rapid efficacy in controlling agitation. In a recent clinical trial, sublingual asenapine was shown to rapidly control agitation within 2 hours in psychiatric patients that included manic patients with a number needed to treat (NNT) of 3. Sublingual asenapine bypasses the first-pass metabolism provided that the patient does not drink or eat for 10 minutes after the administration and this may explain the rapid effect. Similarly, loxapine aerosol delivered through inhalation with staccato system is also rapidly effective in controlling agitation in manic patients with an NNT of 3 achieved within 30 minutes. Loxapine inhalation (ADASUVE, 10 mg) is approved by the FDA in the United States as well as by the European Medicines Agency (EMA) for the treatment of agitation in mania. Since ADASUVE can cause bronchospasm that can potentially lead to respiratory distress and arrest, it is approved under the Risk Evaluation and Mitigation Strategy (REMS) Program to mitigate the risk of bronchospasm. This basically means that ADASUVE can be administered only in an enrolled health-care facility. Given the risk of bronchospasm, it is contraindicated in patients with a history of respiratory conditions such as asthma, chronic obstructive pulmonary disease (COPD), or other lung diseases associated with bronchospasm.

When a rapid control of agitation is required to ensure safety and if a patient is unwilling to take the medication orally, intramuscular (IM) administration of an antipsychotic should be considered. Haloperidol and second-generation antipsychotic medications such as olanzapine, aripiprazole, and ziprasidone IM formulations are rapidly effective within 2 hours in controlling agitation with an NNT of 3. Clinicians, however, must recognize that IM administration of medication against patients' wishes can be a very negative experience for the patient and may interfere with establishment of therapeutic alliance. Hence, if IM formulations are used to

manage agitation and aggressive behavior, patients should be switched to oral medications as soon as the patient is willing to take the medication.

Treatment Selection for Acute Mania

If a manic patient was given an atypical antipsychotic with proven efficacy for mania to control agitation in the emergency setting and if this was well tolerated and improved agitation, this should be continued and the doses adjusted as clinically appropriate to treat acute mania. However, if a first-generation antipsychotic such as haloperidol, or loxapine or chlorpromazine was used to control agitation, patients should be switched to a first-line antimanic agent as soon as possible. Although the first-generation antipsychotic are likely as effective in treating acute mania, they are not recommended as first-line antimanic agents by most experts because first-generation antipsychotics such as haloperidol have a tendency to induce depressive symptoms and increase the risk of depressive episodes in bipolar patients.

First-Line Treatments

Lithium, valproate, and several atypical antipsychotics are all considered first-line agents for treating acute mania (see Table 13.7–5) as they all are effective and that the response rates are broadly similar. Although multiple treatment meta-analyses suggest that risperidone may be more effective, it is likely that this was mostly driven by a risperidone clinical trial conducted in India that enrolled more severely ill manic patients, hence the greater effect size. Thus, it is unlikely, that there are any clinically meaningful significant differences in efficacy between these agents, and hence, monotherapy with any of these agents would be considered an appropriate first-line option for acute mania.

Table 13.7–5.
Recommendations for Pharmacological Treatment of Acute Mania

First line	Monotherapy: Lithium, divalproex, divalproex ER, olanzapine, risperidone, quetiapine, quetiapine XR, aripiprazole, ziprasidone, asenapine, paliperidone ER, cariprazine Adjunctive therapy with lithium or divalproex: Risperidone, quetiapine, olanzapine, aripiprazole, asenapine
Second line	Monotherapy: Carbamazepine, carbamazepine ER, ECT, haloperidol Combination therapy: lithium + divalproex
Third line	Monotherapy: chlorpromazine, clozapine, tamoxifen Combination therapy: lithium or divalproex + haloperidol, lithium + carbamazepine, adjunctive tamoxifen
Not recommended	Monotherapy: gabapentin, topiramate, lamotrigine, verapamil, tiagabine Combination therapy: risperidone + carbamazepine, olanzapine + carbamazepine

ECT, electroconvulsive therapy; XR or ER, extended release.

Modified from CANMAT/ISBD Guidelines.

Some guidelines such as Canadian Network for Mood and Anxiety Treatments (CANMAT) that were published in collaboration with the ISBD recommend combination of an atypical antipsychotic with lithium or valproate also as a first-line option. This is because combination therapy in general results in approximately 15 to 20 percent more patients responding compared with monotherapy. As with monotherapy, there does not appear to be any clinically meaningful significant differences in efficacy between various proven combinations for acute mania. It is, however, important for clinicians to remember that combination therapy is likely associated with more adverse events.

Which First-Line Treatment? Given that several monotherapy and combination therapy first-line options are available for treating acute mania, clinicians have to make a choice as to which monotherapy or combination therapy should be used to treat a given patient with acute mania.

Since there are no efficacy differences between various first-line monotherapies and between different first-line combination therapies, physicians should consider clinical presentation, previous history of response/nonresponse, adverse events, and patient preference in deciding which monotherapy or combination to use to treat mania in a given patient. Patients are more likely to respond to combination therapy compared with monotherapy but this comes with additional adverse event liability. Therefore, clinicians must carefully weigh the risks and benefits of monotherapy versus combination therapy for each patient before making a decision on the most appropriate strategy.

If a decision is made to use monotherapy, patients presenting with acute agitation and aggressive behavior are commonly started on an atypical antipsychotic given their effectiveness in behavioral control and more rapid onset of action compared with mood stabilizers such as lithium. Lithium is the preferred agent in patients with “classical euphoric grandiose” type of mania whereas those with dysphoric/irritable mania, mixed features, comorbid alcohol or substance abuse or brain trauma may respond better to valproate. Similarly, patients with multiple previous mood episodes tend to respond poorly to lithium while this does not predict poor response to valproate. Atypical antipsychotics, and in particular, olanzapine, asenapine, aripiprazole and ziprasidone appear to be equally effective for patients with classical mania as well as those with mixed episodes/mixed features. While clinicians tend to believe that presence of psychotic features may require treatment with an antipsychotic, the evidence from clinical trials suggest lithium, valproate and atypical antipsychotics are equally effective in treating mania in

patients with and without psychosis. However, clinical trial data notwithstanding, if a manic patient who is on lithium or valproate continues to exhibit psychotic symptoms, clinicians should consider adding an antipsychotic. Given that there are clear differences in adverse event profile between various antimanic agents, this is also a very important consideration in treatment selection. For instance, olanzapine is not appropriate for manic patients who are overweight or obese or have a personal or family history of diabetes. Similarly, for patients with kidney disease with compromised kidney function, lithium is not appropriate.

Since there are no efficacy differences between combinations of different atypical antipsychotics (except ziparsidone) with lithium or valproate, the decision with regard to which combination to use for a given patient is also determined by taking into consideration previous history of response, adverse event profile of medications, clinical presentation and preference of the patient.

What If a Patient Does Not Respond to a First-Line Treatment? Since several first-line monotherapy and combination treatment options are available, if one first-line treatment is ineffective, it is appropriate to switch the patient to another first-line treatment. Most manic patients who respond to therapy will show some improvement in manic symptoms within the first 2 weeks. If no improvement is observed during this time period, a switch to a different first-line treatment should be considered. For instance, if a patient is not responding to an atypical antipsychotic monotherapy a switch to lithium or valproate or vice versa might be appropriate. A switch to a different monotherapy or combination could also be considered for tolerability reasons. In real-world clinical practice, it is unusual to encounter a patient who does not respond to one of the several first-line treatment options that are currently available to treat acute mania.

Second-Line Treatments

These should be considered mainly for those patients who have not responded to several first-line therapies.

Haloperidol is an effective treatment of acute mania. However, haloperidol has a tendency to induce depressive symptoms/switch patients into depression. Therefore, despite excellent efficacy for acute mania, haloperidol is recommended only as a second-line option.

Carbamazepine is also as effective in treating acute mania as other first-line antimanic agents but tolerability is a significant concern with carbamazepine. Further, carbamazepine induces hepatic microsomal enzymes, which results in induction of its own metabolism as well as the metabolism of other drugs, and this complicates the management of bipolar disorder.

The efficacy of lithium and valproate combination for acute mania is not examined in large double-blind, placebo-controlled trials. Nonetheless, given that both lithium and valproate are effective in monotherapy, it is likely that combination may provide some benefit for patients who are refractory to monotherapy with either agent or a combination of either agent with an atypical antipsychotic.

Third-Line Treatments

If a patient is a nonresponder to all first-line and second-line treatments, third-line options should be considered. These include first-generation antipsychotics such as chlorpromazine, or an atypical antipsychotic clozapine or tamoxifen.

Treatments Not Recommended

Given the efficacy of valproate and carbamazepine for acute mania, clinicians tend to believe that other anticonvulsants are also effective in treating acute mania. However, there is evidence that gabapentin, topiramate, and lamotrigine are ineffective for acute mania and hence should not be used. Similarly, it is not advisable to combine atypical antipsychotics such as risperidone or olanzapine with carbamazepine as no additional benefit in efficacy can be achieved with this strategy.

ECT IN ACUTE MANIA

Although there is a paucity of well-designed studies assessing the efficacy of ECT for acute mania, it is believed that ECT is highly effective for acute mania based on the data from retrospective and prospective studies. ECT may be used as a first-line option in first trimester of pregnancy or in patients with delirious mania where rapid response is required. In other situations, given the cognitive adverse event profile, ECT should be used primarily in only those manic patients who have failed several first and possibly second-line treatments.

ACUTE BIPOLAR DEPRESSION: TREATMENT OPTIONS

There is a significant unmet need for the treatments of acute bipolar depression as few drugs have proven efficacy.

Lithium

The evidence for efficacy of lithium monotherapy in treating acute bipolar depression is sparse. Earlier small trials that used cross-over designs with short duration of lithium and placebo periods reported that response rates were significantly higher during lithium periods compared with placebo periods. The mean response rate to lithium in these studies was 76 percent. About half of the patients experienced relapse of depression when lithium

was substituted with placebo.

However, the efficacy of lithium was not confirmed in a more recent large double-blind, placebo-controlled parallel design study as lithium was not different from placebo in improving depressive symptoms in bipolar I and bipolar II patients. The mean serum lithium levels in this study were 0.61 mEq/L, and therefore, it is possible that higher levels might have led to greater efficacy. In one study that compared lithium versus lithium plus paroxetine showed that the latter was more effective than lithium alone in patients with low serum lithium levels but in those that had serum lithium levels ≥0.8 mEq/L, there was no difference in efficacy, lending some support to the notion that higher serum lithium levels are required for lithium to be effective for treating acute bipolar depression.

Studies that compared antidepressant/other agent adjunctive therapy to mood stabilizer with mood stabilizer monotherapy (that included patients on lithium) in general found no differences in efficacy between the two groups. Based on this, it has been suggested that mood stabilizer (lithium) monotherapy is effective. Lithium has been reported to have antisuicidal effect and furthermore, lithium is effective in preventing relapse of depressive episodes.

Anticonvulsants

Divalproex/Valproate. Similar to lithium, the data for efficacy of divalproex in acute bipolar depression is also sparse. Divalproex was superior to placebo in 3 out of 4 small placebo-controlled trials. A meta-analysis of these trials suggests that divalproex is effective and that approximately 40 percent of patients will respond and meet criteria for remission. It is currently unknown if there is a relationship between serum valproate levels and therapeutic efficacy for improving depressive symptoms.

Lamotrigine. Five double-blind, placebo-controlled studies with reasonable sample sizes of 7 to 10 weeks duration failed to demonstrate the efficacy of lamotrigine for treating acute bipolar depression. Since the lamotrigine dose needs to be titrated up gradually due to concerns about risk of serious skin rash with rapid up-titration, it is possible that 7 to 10 week study durations were not sufficient to demonstrate the efficacy given that the target dose of 200 mg was not reached until the beginning of week 5 or 6 in these clinical trials.

A pooled analysis of response rates from these trials suggests that lamotrigine is effective in treating acute bipolar depression but the magnitude of effect is small. However, the therapeutic effect was larger in those that had more severe depression.

There is evidence that lamotrigine addition provides therapeutic benefit in bipolar depressed patients who fail to respond to lithium monotherapy.

Other Anticonvulsants. There is only limited data for carbamazepine as its efficacy was assessed in only 3 small placebo-controlled trials that had a total of 40 bipolar depressed patients. The mean response rates to carbamazepine were 68 percent; relapse occurred in 50 percent of these patients with placebo substitution.

Gabapentin and levetiracetam are found to be ineffective in small controlled trials. No placebo-controlled data exist for other anticonvulsants.

Atypical Antipsychotics

While all atypical antipsychotics are effective in treating mania, only few of these have proven efficacy in treating bipolar depression. Among these, quetiapine has the best evidence for efficacy. Clinical trials, however, assessed only two fixed doses (300 mg/day and 600 mg/day) and there is no significant difference in efficacy between these two doses. Thus, for vast majority of bipolar depressed patients, 300 mg/day of quetiapine is sufficient to improve depressive symptoms. Although the efficacy of lower doses of quetiapine has not been studied in acute bipolar depression in clinical trials, experience from clinical practice suggests that some bipolar depressed patients do respond to lower doses.

Olanzapine was more effective than placebo in clinical trials of bipolar depression but the magnitude of benefit is small. Olanzapine plus fluoxetine combination is clearly effective, and is approved by the FDA for the treatment of acute bipolar depression. Although not tested, it is likely that other SSRIs given in conjunction with olanzapine are also effective.

Lurasidone monotherapy and adjunctive therapy to lithium or valproate are also effective in treating acute bipolar depression and were recently approved by the FDA for this indication. Cariprazine, a D2 and D3 receptor partial agonist, at 1.5 mg/day was superior to placebo in one of the two double-blind trials in improving depression in acute bipolar depression. Drop-out rates were higher in the 3 mg/day group that may explain why 3mg dose was not significantly more effective than placebo.

Aripiprazole is not effective in treating acute bipolar depression as it was not superior to placebo in improving depressive symptoms in clinical trials. Although the studies allowed flexible doses between 5 to 30 mg/day, the mean doses in these studies were around 16 mg/day, which may have been too high. More patients in the aripiprazole group dropped out from these clinical trials. Thus, it is conceivable that tolerability related to higher doses may have impacted the efficacy. Indeed, in studies of patients with refractory unipolar depression, aripiprazole given in smaller doses has been effective in augmenting the efficacy of antidepressant medications. Ziprasidone monotherapy and adjunctive therapy are not effective in treating acute bipolar depression.

Antidepressants

There is continuing controversy about the role of antidepressants in treating acute bipolar depression. Most experts agree that monotherapy with antidepressant medications is not appropriate for bipolar I depression because of the concerns of manic/hypomanic switch and possible destabilization of the course of bipolar disorder. This advice is consistent with the recent data from a Swedish Registry study which demonstrated that manic switch is about 2.8 times greater in bipolar patients treated with antidepressant monotherapy compared with those taking antidepressants in conjunction with mood stabilizers. There is consensus that the risk of manic/hypomanic switch is higher with TCAs, older MAOIs, and possibly with SNRI antidepressants such as venlafaxine.

While adjunctive antidepressant therapy is widely used by clinicians to treat acute bipolar depression, the evidence for such strategy from clinical trials is conflicting. Fluoxetine adjunctive therapy to olanzapine is effective but paroxetine, citalopram, bupropion and agomelatine adjunctive therapy failed to provide benefit.

However, two out of three meta-analyses support the efficacy of antidepressants. Further, another recent meta-analysis that calculated the NNT for the various treatments of bipolar depression reported an NNT of 5.75 for antidepressants that is in the same range as the NNT of 5.62 for quetiapine and 5.24 for lurasidone. However, these meta-analyses included studies of both antidepressant monotherapy and combination therapy. This limits the applicability of findings from these meta-analyses to real-world practice as most experts do not recommend antidepressant monotherapy for treating acute bipolar I depression. A more recent meta-analysis that specifically examined the efficacy of modern non-SNRI antidepressant adjunctive therapy supports efficacy but the magnitude of benefit is small. There is little evidence to suggest that non-SNRI antidepressant adjunct therapy when used for the short-term treatment of bipolar depression increases the risk of manic/hypomanic switch or cycling in bipolar disorder.

Adjunctive Therapy with Dopamine Agonists/Psychostimulants

There is conflicting evidence for the efficacy of modafinil/armodafinil adjunctive therapy. While a modafinil and a phase 2 study of armodafinil were positive, only one out of three phase 3 studies of armodafinil adjunctive therapy supported its efficacy for bipolar depression. Armodafinil was well tolerated in these trials with no increase in manic switches.

Pramipexole adjunctive therapy has been shown to be useful in two small trials of patients with refractory acute bipolar depression, although only one of these had included patients with bipolar I depression.

Omega-3-Fatty Acid Adjunctive Therapy

Although there is evidence for the efficacy of adjunctive omega-3-fatty acid therapy in major depression, few studies assessed its efficacy in acute bipolar depression. The results of these studies are conflicting, but a meta-analysis of the data from these trials suggests efficacy. The evidence suggests that potential therapeutic benefits are likely to be derived from eicosapentaenoic acid (EPA) preparation. The omega-3-fatty acid preparations are in general well tolerated; a few patients report gastrointestinal symptoms such as diarrhea and bloating.

Nutritional Supplements as Adjuncts

N-acetylcysteine (NAC) is a glutathione precursor and hence is expected to have protective effects against oxidative stress. There is preliminary evidence that NAC, given at 2 g/day, may provide some benefit in treating depressive symptoms as one out of two studies that assessed the efficacy of adjunctive NAC versus placebo reported improvement in depressive symptoms in patients with bipolar disorder that included patients with bipolar I, bipolar II, and bipolar NOS. There were no significant tolerability issues with NAC.

Glutamatergic Modulators

A few small studies, some of which included treatment-resistant bipolar depressed patients, support the efficacy of the noncompetitive NMDA antagonist ketamine. When infused in subanesthetic doses of 0.5 mg/kg over 40 minutes as an adjunctive therapy, ketamine has been reported to rapidly improve depressive symptoms in patients with TRD. A recent systematic review concluded that a single ketamine infusion was associated with higher rates of response and remission at days 1, 3, and 7. The efficacy appears to be greater in those with unipolar depression than in those with bipolar depression. Ketamine may have antisuicidal effects. There is preliminary evidence that the benefits with ketamine could be maintained with further infusions or IM or oral or sublingual or intranasal administration of ketamine. However, administration of riluzole, another glutamatergic modulator, was not more effective than placebo in maintaining the antidepressant effects of ketamine. Ketamine infusion is associated with transient psychotomimetic effects as well as sedation, dizziness, lightheadedness, incoordination, blurry vision, and feeling "strange or unreal."

Limited open-label data suggest that riluzole at doses of 50 to 200 mg may be useful in treating bipolar depressive symptoms.

Somatic Treatments

Electroconvulsive Therapy. There is consensus among the experts that ECT is one of the most effective treatments for acute bipolar depression. Response rates to ECT are in the range of 50 to 70 percent in refractory major depressive disorder and bipolar depression. A recent meta-analysis concluded that there is no difference in efficacy of ECT between major depressive disorder and bipolar depression. There is some evidence that patients with bipolar depression respond faster and require fewer treatments to improve compared with those with unipolar depression. It is not surprising that a recent study found that ECT is more effective than algorithm-based pharmacological treatments in patients with bipolar depression who have failed at least two previous pharmacological treatments. Most patients with bipolar depression require six to eight treatments to respond.

Maintenance pharmacotherapy is necessary for the vast majority of patients to prevent relapse of depressive and manic episodes. This should be commenced, when possible, prior to completion of the ECT course. While ECT is very effective, cognitive side effects can be often problematic. Therefore, clinicians need to pay particular attention to various ECT stimulation parameters such as electrode placement, pulse width, and electrical dose in order to minimize cognitive side effects.

Magnetic Seizure Therapy. Magnetic seizure therapy (MST) is relatively new technique that involves inducing seizure through magnetic stimulation of brain. There is preliminary uncontrolled evidence that MST is effective in treating depression. However, few patients with bipolar depression were included in these reports. Further, while MST appears to be effective, head-to-head comparison with ECT suggests that response and remission rates are better with ECT compared with MST. It is conceivable that further refinements in stimulation parameters of MST might lead to better efficacy. As expected, MST is associated with much safer cognitive side effect profile as memory, language, and practice seem to be relatively unaffected by MST.

Transcranial Magnetic Stimulation. The data with regard to efficacy of rTMS in acute bipolar depression is limited, although there is extensive literature supporting its efficacy in treating major depressive disorder. A recent meta-analysis which included data on 181 patients with different sub-types of bipolar disorder and different stimulation parameters concluded that overall, there is support for efficacy of rTMS in treating bipolar depression with a NNT of 6. However, given the small number of studies and differences in methodology, but clearly further controlled trials are needed to fully understand its role in treating bipolar depression.

CLINICAL RECOMMENDATIONS FOR THE MANAGEMENT OF ACUTE BIPOLAR DEPRESSION

General Principles of Management

Treatment goals in a patient with bipolar depression include treating depression to full remission without destabilizing the course of bipolar disorder while ensuring the safety of the patient. The risk of suicide is higher during acute depressive episodes and also in depressed patients with mixed features. Therefore, clinicians must carefully assess for suicide risk and appropriate steps must be taken to ensure the safety of the patient. The severity of depression, presence of mixed features or psychotic features, substance abuse or other psychiatric comorbidities, and the severity of suicidal ideation should guide the clinician with regard to the most appropriate treatment setting for managing depression in each patient.

Adjunctive psychological treatments such as CBT, IPSRT, and family-focused therapy (FFT) are beneficial in reducing time to recovery, improving treatment adherence and recovery rates, and reducing subsyndromal symptoms in bipolar depressed patients and, therefore, form an important part of the optimal management of bipolar depression.

Treatment Selection for Acute Bipolar Depression

Many patients with bipolar disorder experience breakthrough depressive episodes. This may be because patients had been taking inadequate doses of maintenance medication, or partially or nonadherent to medications or the maintenance medications were ineffective in preventing the depressive recurrence or relapse. A thorough clinical assessment that gathers information about the course of bipolar disorder (e.g., number of previous manic and depressive episodes, their duration, and severity), the treatment history (current and previous medications, what worked and what did not work, tolerability issues), and adherence to treatment should help the clinician in determining the most appropriate treatment of a given patient.

First-Line Treatments

The evidence for the efficacy of many commonly used medications for treating depression in bipolar disorder is less robust and often conflicting. Therefore, it is not surprising that there is no general consensus among experts about the most appropriate first-line treatment options for bipolar depression. The CANMAT/ISBD group that published the most recent treatment guidelines took into account several factors such as efficacy from controlled trials, meta-analyses, clinical experience, safety of medication, and long-term efficacy in arriving at the first-line treatment recommendations for acute bipolar depression. A modified list of CANMAT recommendations for bipolar depression is provided in Table 13.7–6.

Which First-Line Treatment? If a patient had been taking one of the first-line treatments, optimizing the current treatment is the most appropriate first step before other interventions are considered. Similarly, a first-line treatment should be commenced in patients who are drug free or not taking first-line medications. While there is no clear research evidence to guide clinicians between various first-line monotherapy options, previous treatment history, illness course, symptom profile of depression, side effect profile of medications, and patient preferences are all important considerations in determining treatment choices among the first-line therapies. Lithium is more appropriate for patients with a history of classical manic episodes, an illness course that has mania, depression, and euthymia pattern, and a family history of bipolar disorder. Bipolar I patients with frequent depressions and infrequent mild manias are suitable for lamotrigine monotherapy while quetiapine is a better option for those with frequent depressions as well as manic episodes. Those with reverse vegetative symptoms such as hypersomnia and hyperphagia may be reluctant to take quetiapine because of its sedating and appetite stimulating properties, although there is no reason to believe that it is less effective for this population. Lamotrigine dose needs to be escalated gradually due to concerns of skin rash with rapid dose escalation, and hence although it is effective in severely depressed patients, it may not be the most appropriate option given that it takes a few weeks to reach effective therapeutic dose. ECT or olanzapine plus an SSRI combination may be a better option for patients with depression with psychotic features. Lurasidone is relatively weight neutral and is preferred by patients who are reluctant to take medications that can cause weight gain. While the pharmacological profile of this agent suggests that it may prevent mood episodes, given the lack of data, it may be prudent to consider combining lurasidone with lithium or valproate.

Table 13.7–6.
Recommendations for Pharmacological Treatment of Acute Bipolar I Depression

First line	Monotherapy: Lithium, lamotrigine, quetiapine, quetiapine XR, Lurasidone Combination therapy: lithium or divalproex + lurasidone, lithium or divalproex + SSRI, lithium or divalproex + bupropion, olanzapine + SSRI, lithium + lamotrigine
Second line	Monotherapy: Divalproex, Cariprazine, Carbamazepine, Olanzapine, ECT[a] Combination therapy: adjunctive SSRI to an atypical antipsychotic, adjunctive modafinil, adjunctive pramipexole, adjunctive EPA
Third line	Combination therapy: lithium + carbamazepine, adjunctive NAC, adjunctive rTMS
Not recommended	Monotherapy: Gabapentin, aripiprazole, ziprasidone Combination therapy: Adjunctive ziprasidone, adjunctive levetiracetam

[a]Can be a first-line treatment of severe depression with suicidal ideation or in the first trimester of pregnancy.

SSRI, selective serotonin reuptake inhibitor; MAOI, monoamine oxidase inhibitor; ECT, electroconvulsive therapy; AAP, atypical antipsychotic; NAC, N-acetylcysteine; rTMS, repetitive transcranial magnetic stimulation.

Modified from CANMAT/ISBD 2013 Guidelines Recommendations.

What If a Patient Does Not Respond to a First-Line Treatment? If optimization of monotherapy with a first-line is agent is ineffective, a switch to another first-line monotherapy therapy (e.g., a switch from lithium to quetiapine or vice versa) or combination therapy (e.g., lamotrigine plus lithium or lurasidone plus lithium) is appropriate. Addition of an SSRI or bupropion could be considered for bipolar depressed patients who were unresponsive to optimization of lithium or valproate monotherapy, especially if there is a previous history of response to adjunctive antidepressant therapy. However, antidepressants are not appropriate for patients with a history of rapid cycling or those with mixed features. If antidepressants are used for treating acute bipolar depression, an attempt should be made to taper and discontinue them within 8 weeks of remission of depression. Given the scarcity of evidence for efficacy of second-line treatments, all first-line treatments should be tried before considering second-line treatment options.

Second-Line Treatments

ECT should be considered for patients who have failed several first-line treatments. ECT may be considered earlier for bipolar depressed patients with psychotic features, those at high risk of suicide, elderly, and those with rapidly deteriorating physical status due to poor oral intake because of depression.

Valproate monotherapy is appropriate for bipolar depressed patients who had tolerability issues with other first-line monotherapy options. While no data exists, it may also be useful for bipolar depressed with mixed features. Similarly, cariprazine or carbamazepine monotherapy is appropriate for patients who were unresponsive or had tolerability issues to lithium, valproate, and other first-line treatments. Adjunctive pramipexole or modafinil may be beneficial in refractory patients who have significant anergia, anhedonia, and hypersomnia. Antidepressant adjunctive therapy to other atypical antipsychotics such as aripiprazole or quetiapine may be helpful for some patients. Adjunctive EPA between 1 and 4 g/day is appropriate for refractory patients or those with tolerability issues.

Third-Line Treatments

These options are only appropriate for patients who failed all first- and second-line options because some third-line treatments such as adjunctive venlafaxine and MAOIs carry significant risks of potentially destabilizing

the course of bipolar disorder. If these strategies are employed, clinicians need to educate patients and their family members about the potential risks of manic switch and illness destabilization.

Treatments Not Recommended

Monotherapy with aripiprazole and ziprasidone is not recommended. However, these agents are effective in treating and preventing mania. Hence, adding a first-line agent or SSRI to these medications is an acceptable option.

MAINTENANCE THERAPY/PROPHYLAXIS: TREATMENT OPTIONS

Lithium

Lithium is still considered the "gold standard" by many experts for the maintenance treatment of bipolar disorder. Lithium is effective in preventing mood episodes not only in bipolar I patients who had responded to lithium during acute episodes (enriched for lithium response) but also in those who had been treated with a different psychotropic medication for acute mood episodes (not enriched for lithium response). Lithium is effective in preventing both manic and depressive episodes, although the magnitude of benefit may be slightly lower for depression than it is for preventing mania. While earlier studies suggested that around 70 percent of patients with bipolar disorder remained episode free with lithium, more recent studies suggest that the response rates are probably below 50 percent. This may be due to several reasons including the cohort effect, and the fact that these studies recruited patients who were responders to other psychotropic medication during acute episodes.

There is evidence that lithium plus valproate combination is more effective than valproate alone in preventing relapse of mood episodes. In bipolar patients who are refractory to therapy with lithium or carbamazepine, combination appears to be helpful in reducing the risk of relapse of mood episodes.

It is recommended that lithium is given as a single daily dose as there is some evidence that multiple daily doses of lithium are more commonly associated with greater incidence of renal side effects and pathological changes in kidneys. Further, treatment adherence is likely better with single daily dose.

There is evidence from posthoc analyses of recent studies that serum lithium levels below 0.6 mEq/L are less effective. Therefore, for prophylaxis, serum lithium levels should be maintained between 0.6 and 1.2 mEq/L. There is inconsistent evidence with regard to whether serum levels above 0.6 mEq/L lead to greater efficacy. Higher lithium levels, however, are associated with greater incidence of adverse events. Lithium levels should be measured after the target dose has been achieved, in the

steady state (approximately 5 days after the last dose change), about 12 hours after the last dose. Serum levels should be repeated until two consecutive levels within the therapeutic range are obtained for the same dose.

Patients receiving maintenance lithium therapy should be monitored as per modified ISBD safety monitoring guidelines (see Fig. 13.7–1) or more frequently if clinically indicated.

Patients on maintenance treatment of lithium should be educated about symptoms of lithium toxicity and the preventive measures. Risk factors for lithium toxicity include renal diseases, dehydration, drug interactions, elderly and those with organic brain syndromes. Lithium toxicity is a medical emergency, and if not treated can result in irreversible organ damage and death.

If lithium is to be discontinued for any clinical reason in bipolar patients whose serum lithium levels had been in therapeutic range, it should be tapered and discontinued over a period of a several weeks. This is because there is evidence that time to 50 percent risk of recurrence of mood episodes is four times shorter with abrupt discontinuation (1 to 14 days) compared with gradual (15 to 30 days) discontinuation of lithium.

Anticonvulsants

Carbamazepine. There is limited data for the efficacy of carbamazepine from placebo-controlled trials. However, most, although not all, controlled studies that compared carbamazepine with lithium showed similar efficacy. While no study showed that carbamazepine was more effective than lithium, there is some evidence that carbamazepine may be more effective than lithium in patients with “non-classical bipolar disorder” (i.e., patients with bipolar II disorder/bipolar NOS or with mood-incongruent delusions or with anxiety disorders or substance use comorbidity or mixed states).

The clinical trials have not specifically examined the relationship between serum carbamazepine levels and therapeutic efficacy in maintenance treatment. Hence, it is currently unknown what levels might offer optimal prophylactic efficacy. In clinical practice, serum carbamazepine levels are typically maintained between 4 and 12 μg/L, the same range used for treating patients with epilepsy.

Carbamazepine induces hepatic microsomal enzymes that metabolize carbamazepine. As a result, within the first few weeks of commencing carbamazepine, the clearance of carbamazepine increases by approximately twofold; hence, dosage adjustments are necessary in order to keep serum carbamazepine levels within the clinically accepted range. Carbamazepine causes leukopenia in 10 to 20 percent of patients, and it occurs commonly within the first 3 months of commencing treatment and resolves with discontinuation of carbamazepine. Agranulocytosis and aplastic anemia are rare but have a more rapid onset and occur in an unpredictable pattern.

Therefore, periodic clinical monitoring is not likely to be of much help. Clinicians need to be vigilant to this possibility when hematological symptoms emerge clinically.

There are no controlled data to support the efficacy of oxcarbazepine in the maintenance treatment of bipolar disorder.

Valproate. In contrast to lithium that has robust data for its efficacy in maintenance treatment, only a single large double-blind, controlled trial assessed the efficacy of valproate in maintenance treatment of bipolar disorder. In this trial, valproate was not superior to placebo on the primary measure of time to any mood episode. This study was considered a failed trial for a variety of reasons including the fact that lithium was also not superior to placebo in preventing relapse of mood episodes. Valproate, interestingly, was superior to placebo on some secondary measures including time to discontinuation for a recurrent mood episode or depressive episode. In the subgroup of patients who were treated with valproate during the open-label phase, valproate was superior to placebo in time to any mood episode. Further, patients randomized to valproate were significantly less likely to drop out for a mood episode and less likely to be prematurely terminated for any reason. Thus, in aggregate, these data suggest that valproate may be beneficial for preventing mood episodes in patients who respond to valproate during acute episodes.

A randomized controlled study found no difference in efficacy between valproate and lithium in patients with rapid cycling bipolar disorder who were stabilized on a combination of valproate plus lithium. Consistent with this, the results of a recent Cochrane review indicated that valproate was more effective than placebo in preventing study withdrawal due to any mood episode and that there was no difference in efficacy between lithium and valproate. There is evidence that fewer patients drop out of treatment with valproate compared with placebo or lithium. This may explain why valproate is widely used for maintenance treatment.

Although there is a linear relationship between serum levels of valproate and improvement in acute manic symptoms, there is no data to guide the clinicians with regard to serum valproate levels for the maintenance treatment of bipolar disorder. Therefore, it is a common practice to maintain the same valproate dose that was effective in treating mania for maintenance treatment, provided serum levels are within the range used for epilepsy (45 to 125 μg/mL).

Most patients prefer to take valproate as a single bedtime dose. Valproate can cause nausea and vomiting, particularly in the first few days of commencement of therapy. Some patients experience weight gain and hair loss, women taking valproate should be monitored for polycystic ovary syndrome with questions about menstrual irregularities, fertility issues, hirsutism, and galactorrhea.

Lamotrigine. Lamotrigine is approved for prophylaxis of mood episodes in bipolar I disorder by the FDA. Lamotrigine has robust efficacy in preventing depressive relapses/recurrences, but it only has marginal efficacy in preventing manic episodes. There is no data to suggest a relationship between serum levels of lamotrigine and prophylactic efficacy. Lamotrigine add-on to aripiprazole reduces the risk of depressive relapses in bipolar patients who are on maintenance therapy with aripiprazole.

A large randomized open study that compared lamotrigine with lithium found no statistically significant differences in efficacy but as expected, the hazard ratios for lamotrigine relative to lithium were higher (1.91) for mania and lower (0.69) for depression. In rapid cycling bipolar disorder, lamotrigine was no more effective than placebo on the primary measure of time to additional pharmacotherapy, but lamotrigine patients stayed significantly longer in the study compared with those that received placebo.

Lamotrigine is titrated upward slowly due to concerns about skin rash with rapid titration. Other risk factors for skin rash include younger age, a prior history of rash with another drug, and taking lamotrigine in conjunction with valproate. Rash more commonly occurs within the first few weeks of therapy. A recent review concluded that the incidence of skin rash was about 10 percent in prospective studies. The incidence of serious skin rash (i.e., Steven Johnson syndrome or toxic epidermal necrolysis) is rare (about 1 to 10 per 10,000) with slow dose titration. A recent review reported that about 50 percent of patients with serious skin rash had received lamotrigine in conjunction with valproate.

If a patient on lamotrigine therapy develops skin rash, this needs to be evaluated carefully to determine if it is a benign or serious rash. When in doubt, if immediate dermatological consultation is not available, lamotrigine should be discontinued and the patient should be carefully monitored. If the rash resolves and if it was considered benign, a rechallenge with lamotrigine can be undertaken 4 weeks later if lamotrigine therapy is essential for the management of bipolar disorder. Those with serious skin rash require urgent medical evaluation and treatment.

Apart from cutaneous reactions described above, lamotrigine does not cause weight gain and is in general well tolerated.

Other Anticonvulsants. Gabapentin is not effective for treating acute mania or bipolar depression. In a small controlled trial, surprisingly, gabapentin adjunctive therapy was beneficial in maintaining mood stability. There are no controlled maintenance data for topiramate, levetiracetam, felbamate, or other anticonvulsants.

Antipsychotics

Typical (Conventional) Antipsychotics. Conventional antipsychotics are likely effective in preventing manic episodes. However, there is evidence

that conventional antipsychotic therapy is associated with increased risk of depressive symptoms and depressive episodes. Further, patients with bipolar disorder are at higher risk of developing extrapyramidal symptoms including tardive dyskinesia with conventional antipsychotics. Therefore, therapy with conventional antipsychotics including depot preparations should be mainly reserved for those patients who have failed treatment with traditional mood stabilizers and atypical antipsychotics.

Atypical Antipsychotics. Atypical antipsychotics are widely used for treating acute mania and all currently available agents in North America with the exception of lurasidone and clozapine have proven efficacy from double-blind, controlled trials. If a manic patient responds to an atypical antipsychotic agent, whether continuation of that atypical antipsychotic is effective in preventing relapse of mood episodes is an important clinical question. There is evidence that continuation of aripiprazole monotherapy in such patients is effective in preventing mood episodes and manic episodes but not depressive episodes. Similarly, aripiprazole adjunctive therapy to lithium or valproate is also beneficial in preventing manic episodes. Aripiprazole dose range is between 10 and 30 mg/day with a mean dose of about 17 mg/day when it is used as an adjunctive therapy and about 24 mg/day when used as a monotherapy for prophylaxis.

Continuation of olanzapine in manic/mixed episode patients who respond to olanzapine is effective in preventing manic, mixed, and depressive episodes, but the magnitude of benefit appears to be greater for preventing mania and mixed episodes relative to depression. Olanzapine appears to be as effective as lithium and valproate in preventing mood episodes. Olanzapine adjunctive therapy to lithium or divalproex appears to be useful in preventing symptomatic relapse of mood episodes.

While oral risperidone has not been assessed for its efficacy, manic patients who responded to oral risperidone and switched to risperidone long-acting injectable (Risperdal Consta) are less likely to experience relapse of mood episodes. Like aripiprazole, Risperdal Consta is effective in preventing manic episodes but not depressive episodes. Most patients require about 25 mg of Risperdal Consta every 2 weeks for preventing manic episodes. Risperdal Consta adjunctive therapy is also effective in preventing relapse of mood episodes and manic episodes but not depressive episodes in frequently relapsing bipolar patients.

Paliperidone ER monotherapy (mean dose of 6 mg/day) continuation in manic patients who responded to paliperidone is effective in prolonging time to and reducing the risk of relapse of mood episodes and manic episodes but not depressive episodes. In schizoaffective patients, paliperidone palmitate once monthly (between 78 and 234 mg monthly) given as either monotherapy or adjunctive therapy is also beneficial in reducing the risk of relapse of manic, depressive, and psychotic symptoms.

Quetiapine is effective in preventing mood episodes, manic/mixed episodes, and depressive episodes in bipolar patients who responded to quetiapine during acute mania or bipolar depression. Quetiapine appears to be equally effective in preventing both mania and depression. The mean doses of quetiapine in the maintenance studies ranged from 527 to 562 mg/day. Quetiapine adjunctive therapy to lithium or divalproex is also effective in preventing both manic and depressive episodes, and the magnitude of benefit of adjunctive therapy for prevention of mania and depression is similar with hazard ratios around 0.3. This means that there is approximately a 70 percent reduction in risk of manic and depressive episodes with quetiapine adjunctive therapy compared with lithium or valproate monotherapy.

While ziprasidone monotherapy is effective in treating acute mania, there is no data to guide clinicians with regard to whether continuation of ziprasidone in such patients is useful in preventing mood episodes. However, ziprasidone adjunctive therapy to lithium or valproate is effective in preventing mood episodes and manic episodes but not depressive episodes.

A recent and as yet unpublished double blind placebo controlled trial showed that asenapine monotherapy was more effective than placebo in preventing relapse of mood episodes in recently stabilized manic/mixed episode patients. Numerically fewer patients in the asenapine group had relapse of manic, mixed and depressive mood episodes. The studies assessing the efficacy of lurasidone in preventing mood episodes are ongoing. Clozapine is likely effective at least in preventing mania but no controlled studies have been conducted to date. In refractory bipolar/schizoaffective patients, clozapine add-on has been reported to be beneficial in reducing manic and psychotic but not depressive symptom scores on rating scales.

Antidepressants. There is evidence that TCAs increase risk of manic switch and are not recommended for treating acute bipolar depression. The maintenance studies that compared imipramine with lithium or lithium plus imipramine confirmed increased risk of manic switch with both imipramine mono and adjunctive therapy, and hence TCAs are clearly not appropriate for maintenance therapy of bipolar disorder.

If modern antidepressants are used as adjuncts to treat acute bipolar depression, is continuation of antidepressant therapy beneficial? This issue was addressed in a Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) trial that randomized patients who responded to adjunctive antidepressant therapy to continuation for 1 to 3 years versus discontinuation within 2 weeks. The results showed that continuation of antidepressant therapy was associated with a trend toward less depressive symptom burden without increasing manic/hypomanic symptoms.

Further, the time to depressive episode was significantly delayed in the antidepressant continuation group. There was no evidence of increase in manic episode risk. Interestingly, a rapid cycling course predicted a threefold increase in depressive episodes in patients who continued antidepressants versus those who discontinued. A recent meta-analysis of modern antidepressant adjunctive therapy showed an increased risk of manic switch with long-term therapy with a NNH of 19. Therefore, clinicians are advised to carefully assess the risks and benefits of the need for long term antidepressant adjunctive therapy for each patient and tailor the treatment accordingly.

CLINICAL RECOMMENDATIONS FOR MAINTENANCE TREATMENT/PROPHYLAXIS OF BIPOLAR I DISORDER

General Principles of Management

Most experts agree the need for long-term prophylaxis in all patients with bipolar I disorder and that prophylaxis should begin after the first manic episode. This is because bipolar disorder is a highly recurrent illness, frequently characterized by residual symptoms and chronicity. Over 90 percent of patients have recurrence of a mood episode within 5 years. A recent meta-analysis suggests that recurrence rates are lower in first manic episode patients with about 60 percent experiencing a mood episode within 4 years. Bipolar I patients experience syndromal or subsyndromal symptoms 50 percent of the time with depressive symptom periods predominating by a ratio of 3:1 to manic/hypomanic symptoms. There is evidence that subsyndromal symptoms increase the risk of full-blown mood episodes. Multiple mood episodes are more likely to be associated with cognitive impairment, structural brain changes, and functional impairment. Frequent episodes alter response to treatment such that those that have had multiple mood episodes are less likely to respond to treatment during acute mood episodes as well as to prophylactic pharmacotherapy and adjunct psychotherapy.

Therefore, the goals of the long-term treatment include prevention of mood episodes, prevention of/improvement in subsyndromal symptoms, prevention of suicidal behavior, and improvement in treatment adherence, quality of life, cognitive, social, and work function.

While pharmacotherapy is the mainstay of the long-term prophylactic treatment of bipolar disorder, there is evidence that adjunct psychological treatments are helpful in reducing risk of relapse in bipolar patients. Group or individual psychoeducation is effective in reducing relapse rates by 25 percent to 30 percent. Patients should be educated about illness, adverse effects of medication, early warning symptoms of mood episodes, triggers for mood episodes such as substance use and psychosocial stressors, as well as coping strategies. IPSRT and FFT are also effective. Bipolar patients who

recovered from mood episodes tend to benefit from adjunct individual CBT that reduces relapse rates and prolongs time to mood episode. The STEP-BD study reported that recovery rates at 1 year were significantly greater in acute bipolar depressed patients who received CBT, IPSRT, or FFT compared with those that received collaborative care. Thus, adjunct psychological treatments are essential components of the long-term treatment of bipolar disorder to improve treatment adherence and mood symptoms and reduce relapse rates.

Treatment Selection for Maintenance Treatment

Several factors are important in aiding the treatment selection for the maintenance treatment. These include information about course of the illness in the patient, response (lack of response), and tolerability to various treatments during acute episodes and maintenance treatment, predictors of response to various medications, predominant polarity of the illness, comorbidity, and whether the most recent episode was manic or depressive and what treatment/s were effective in remitting this episode. Family history of bipolar disorder if present and treatment response in family members are also helpful in some cases.

First-Line Treatments

Lithium, valproate, lamotrigine, and some atypical antipsychotics are effective and are recommended as first-line agents for the maintenance treatment of bipolar disorder (see Table 13.7–7). However, it is important to note that the benefits of these agents in reducing relapse rates have been demonstrated in bipolar patients who were responders to acute therapy with the same agents. The only exception to this is lithium that has been shown to be effective in the maintenance treatment regardless of whether patients were treated for an acute episode with lithium or not.

Which First-Line Treatment? If a patient was recently treated for an acute manic episode with monotherapy with one of the first-line agents listed in Table 13.7–7, it is advisable to continue the same medication for the maintenance treatment. Risperidone and ziprasidone are first-line agents for mania but not for the maintenance therapy because of lack of data. However, given that the other atypical agents that are effective in mania are also effective in maintenance treatment, clinicians may choose to continue these agents for the maintenance treatment after discussing the potential benefits and risks with such strategy. Although paliperidone ER is effective in maintenance treatment, it was less effective than olanzapine, and hence, it is listed as a second-line treatment for maintenance, but in manic patients who were responders to paliperidone ER, clinicians may also choose to continue this medication for prophylaxis.

Table 13.7–7.
Recommendations for Maintenance Pharmacotherapy of Bipolar Disorder

First line	Monotherapy: Lithium, lamotrigine (limited efficacy in preventing mania), divalproex, olanzapine,[a] quetiapine, risperidone LAI (mainly for prevention of mania), aripiprazole (mainly for preventing mania), asenapine Adjunctive therapy with lithium or divalproex: Quetiapine, risperidone LAI (mainly for prevention of mania), aripiprazole (mainly for preventing mania), ziprasidone (mainly for preventing mania)
Second line	Monotherapy: Carbamazepine, paliperidone ER, paliperidone palmitate Combination therapy: Lithium + divalproex, lithium + carbamazepine, lithium or divalproex + olanzapine, lithium + risperidone, lithium + lamotrigine
Third line	Monotherapy: Ziprasidone,[b] risperidone[b] Adjunctive therapy: Clozapine, ECT, omega-3-fatty acids, oxcarbazepine, gabapentin
Not recommended	Monotherapy: Gabapentin, topiramate, or antidepressants Adjunctive therapy: Flupenthixol

[a]Given the metabolic side effects, use should be carefully monitored.

[b]Manic patients who responded to these, continuation is appropriate as a first line.

LAI, long-acting injection; ECT, electroconvulsive therapy; ER, extended release.

Modified from CANMAT/ISBD Guidelines.

It is important to note that most atypical antipsychotic medications with the exception of quetiapine and possibly olanzapine are effective mainly in preventing mania. Hence, if a patient's illness course is characterized by predominant depressive polarity, clinicians are advised to consider adding lamotrigine or lithium to an atypical antipsychotic agent or replacing the atypical antipsychotic agent with either lithium or quetiapine for prophylaxis.

Manic patients who were treated with a combination of lithium or valproate with an atypical antipsychotic, there is evidence that continuation of combination therapy is more effective than lithium or valproate monotherapy for preventing relapse of mood episodes. A recent study that compared different durations of atypical antipsychotic adjunctive therapy with risperidone or olanzapine for the maintenance treatment reported that adjunctive therapy is beneficial for the first 6 months after remission of mania, but continuation beyond this period was associated with significant weight gain liability with no apparent additional benefit in efficacy. Hence, clinicians are advised to carefully assess the need for the continuation of atypical antipsychotic adjunctive therapy beyond 6 months after mania remission. For instance, a longer duration of adjunctive therapy may be appropriate in bipolar patients with a previous history of insufficient prophylaxis with lithium or valproate monotherapy.

All first-line monotherapy agents for acute bipolar depression with the exception of lurasidone are also effective in maintenance treatment of bipolar disorder, and hence should be continued for prophylaxis. In a

bipolar depressed patient treated with lurasidone monotherapy, clinicians may choose to continue this medication for prophylaxis despite lack of maintenance data given its D2 blockade (likely confers acute and prophylactic antimanic efficacy) properties. Lamotrigine monotherapy is effective mainly in preventing mania, and hence in bipolar patients with previous severe manic episodes, clinicians are well advised to add a prophylactic antimanic agent such as lithium or an atypical antipsychotic for optimal prophylaxis.

If a bipolar depressed patient responded to adjunctive modern antidepressant therapy, is continuation of adjunctive therapy with antidepressant beneficial? There is preliminary evidence that continuation in nonrapid cycling bipolar patients is beneficial in reducing the risk of depressive relapses without destabilizing mood. The CANMAT/ISBD guidelines recommend tapering and discontinuing antidepressants about 8 weeks after remission of depression, based on the assumption that antidepressants are primarily needed to cover the natural duration of the depressive episode, and beyond this, there is a risk of potential destabilization.

In making a decision with regard to the maintenance medication for bipolar patients who are currently not on any treatment, clinicians need to consider several factors outlined in treatment selection section above. If lithium and quetiapine have not been tried previously, given the robust evidence for efficacy in preventing both mania and depression, they are the most appropriate first-line agents for such patients. Patients with a history of alcohol or substance abuse or mixed features may be more appropriate for valproate therapy. For those with predominant manic polarity, any first-line agent with the exception of lamotrigine would be appropriate.

What If a First-Line Agent Is Ineffective? If a patient experiences breakthrough subsyndromal symptoms or relapse while on monotherapy with one of the first-line agents, a switch to an alternative first-line agent or a combination therapy is appropriate. For instance, if a patient on lithium experiences depressive symptoms, consider adding lamotrigine or quetiapine. If manic symptoms emerge on lithium, consider adding an atypical antipsychotic or valproate or if there are tolerability issues, consider switch to an alternate first-line agent. It is recommended that monotherapy or combination therapy with several first-line agents be tried before embarking on second-line strategies for the maintenance therapy.

Second-Line Treatments

Given that the evidence for prophylactic efficacy of carbamazepine and paliperidone is less robust, these options are mainly considered for bipolar patients who failed or had tolerability issues to monotherapy with first-line agents. Carbamazepine does not cause weight gain and is effective in

preventing relapses, particularly in patients with atypical forms of bipolar disorder.

However, many patients who failed several first-line agents are unlikely to respond to monotherapy. Hence, combination therapy strategies outlined in Table 13.7–7 are most appropriate for such patients.

Third-Line Treatments

Clinical experience and some controlled data suggest that clozapine adjunctive therapy is very effective in refractory bipolar disorder. However, in addition to the risk of agranulocytosis, clozapine is associated with significant weight gain and metabolic side effects, and hence patients need to be monitored closely for these adverse events and treated appropriately. For some refractory patients, maintenance ECT is an option. Adjunctive therapy with omega-3-fatty acids and other agents can be considered for patients who have tolerability issues or suboptimally controlled with other combination strategies.

TREATMENT OF BIPOLAR II DISORDER

Although bipolar II disorder is common, associated with significant morbidity and mortality, few clinical trials investigated the efficacy of treatments for various phases of bipolar II disorder. The limited data that is currently available is derived from studies of bipolar I disorder or major depressive disorder that had included a few patients with bipolar II disorder. Therefore, in the absence of controlled clinical trial data, by necessity, extrapolation is made from the treatment studies in bipolar I disorder, to aid clinical decisions in managing bipolar II disorder.

Treatment of Hypomania

It is generally believed that medications that are effective in treating acute mania are also effective in treating hypomania. Indeed, small placebo-controlled trials or open-label studies demonstrated that valproate, quetiapine, and risperidone are effective in treating hypomanic symptoms. These findings are consistent with clinical experience of experts in the field.

Thus, bipolar II patients with persistent hypomanic symptoms could be treated with an antimanic agent to rapidly alleviate hypomanic symptoms. Prior history of treatment response, adverse effects of medications, and patient preferences are important considerations in treatment selection. If a patient had been taking a prophylactic antimanic agent, an increase in the dose of that medication may be sufficient in some cases. Medications that could potentially be contributing to hypomanic symptoms such as antidepressants should be discontinued.

Treatment of Bipolar II Depression

The data for efficacy of various treatments is even more limited for treating acute bipolar II depression relative to bipolar I depression. Quetiapine is the only medication with some evidence for efficacy from placebo-controlled trials with adequate sample sizes in treating acute bipolar II depression. In three out of five studies, quetiapine/(XR) was more effective than placebo in improving depression with improvement apparent as early as week 1. Hence, quetiapine is an appropriate first-line option for bipolar II depressed patients who are drug naïve or have never tried quetiapine.

Most experts recommend lithium, lamotrigine, valproate, and lurasidone as appropriate monotherapy agents for some patients with bipolar II depression despite limited and conflicting data.

Modern antidepressants are widely used to treat bipolar II depression, although the evidence is limited and conflicting. Given the limited effective options for treating bipolar II depression, modern antidepressant adjunctive therapy is appropriate for treating some patients with bipolar II depression. A meta-analysis suggests that the risk of antidepressant mood elevation is lower in bipolar II disorder compared with bipolar I disorder (7 percent in acute and 14 percent in maintenance trials versus 14 percent and 23 percent, respectively for bipolar I). Therefore, there is less disagreement among experts about the use of modern antidepressant adjunctive therapy for bipolar II depression. This advice is consistent with the recommendations of a recent ISBD Task Force that endorses its use in bipolar II depression except in patients with rapid cycling bipolar disorder and in those with mixed features. Although a series of small studies suggest that antidepressant monotherapy is effective and is not associated with increased risk of switch, such strategy is not recommended for most bipolar II depressed patients.

ECT is an effective treatment of bipolar II depression. Although the evidence is very limited, adjunctive therapy with pramipexole, modafinil, omega-3 fatty acids or n-acetyl cysteine could be considered for treating bipolar II depressed patients who are refractory to above strategies.

Maintenance Treatment of Bipolar II Disorder

Syndromal or subsyndromal depressive symptoms comprise the vast majority of symptomatic periods in patients with bipolar II disorder. Therefore, the major challenge in the management of bipolar II disorder is not only to treat depressive episodes to full remission but also to prevent relapse/recurrence of depressive episodes without increasing the risk of hypomanic switch.

Clinicians typically continue the medication that worked for the acute episode for the maintenance treatment. While such strategy has not been widely studied in bipolar II disorder, there is some evidence that bipolar II depressed patients who respond to quetiapine or lamotrigine benefit from continuation of the same medication for the maintenance treatment.

Lithium is beneficial for maintenance regardless of whether it was used to treat acute episodes or not. If a patient has responded to adjunctive modern antidepressant therapy, continuation of the combination may be beneficial for some patients. Some studies suggest that patients who responded to fluoxetine monotherapy may benefit from continuation of this medication for prophylaxis. Although these studies did not report increased hypomanic switch rates, clinicians are well advised to warn patients about the potential risk of hypomanic/manic switch if such strategy was chosen. If the above options are ineffective, open-label data and clinical experience support the use of combinations of mood stabilizers or a mood stabilizer and an atypical antipsychotic with or without a modern antidepressant or pramipexole or other agents such as NAC or omega-3 fatty acids.

TREATMENT OF BIPOLAR DISORDER WITH COMORBIDITY

Both psychiatric and medical comorbidity is highly prevalent in patients with bipolar disorder and impacts clinical outcome. In general, bipolar patients with comorbidities tend to have poorer outcomes as they are less likely to respond or take longer time to respond to treatments and are more likely to relapse. Therefore, the optimal management of such patients includes treating not only symptoms associated with bipolar disorder but also addressing comorbidities without destabilizing bipolar disorder. However, clinical trials conducted for regulatory approval of treatments for bipolar disorder have systematically excluded patients with comorbidities. Therefore, in the absence of evidence-based strategies for the management of patients with bipolar disorder with psychiatric comorbidity, the treatment decisions for the management of comorbidity would have to be made by necessity, based on the secondary analysis of data on comorbid symptoms in trials of bipolar disorder and the treatment efficacy data from clinical trials of primary condition. In choosing treatments for managing comorbidity, clinicians must carefully weigh the risks of adjunctive treatments such as their propensity to destabilize bipolar disorder and adverse events versus benefits in treating comorbidity and the consequential positive impact on the outcome of bipolar disorder.

Anxiety Disorders

Since several psychological treatments such as CBT have been shown to be effective in treating primary anxiety disorders, their use is preferable in treating the comorbid anxiety disorders in bipolar patients as they pose little risk of illness destabilization.

Generalized Anxiety Disorder. Generalized anxiety disorder/anxiety symptoms are common in bipolar disorder, impact clinical outcomes, and hence must be treated. In addition to psychological treatments, if a patient

has been taking an anticonvulsant such as lamotrigine or valproate or an atypical antipsychotic such as quetiapine or olanzapine plus fluoxetine combination for bipolar disorder, these should be optimized as there is some evidence from secondary analysis of clinical trials or open-label data to support their efficacy. If optimization of ongoing pharmacotherapy is ineffective in resolving symptoms, medications that have been shown to be effective in treating primary anxiety disorder should be considered. Among these, pregabalin and quetiapine are preferable as they are not associated with the risk of destabilization of bipolar disorder. Gabapentin is also useful as anecdotal evidence supports its efficacy in treating comorbid anxiety symptoms in bipolar patients. Some SSRI antidepressants (escitalopram, paroxetine, sertraline) are effective and widely used in clinical practice. Similarly, some newer antidepressants such as agomelatine and vortioxetine are also effective in treating GAD; the risk of illness destabilization appears to be lower with agomelatine while no data are available as yet for vortioxetine. However, it must be remembered that since anxiety disorders/symptoms pose chronic symptom burden in bipolar disorder, the long-term use of SSRIs or other antidepressants in bipolar patients may pose the considerable risk of illness destabilization. Hence, other strategies outlined above should be tried first before considering the use of SSRIs. If SSRIs are used on a longer term basis, patients need to be on adequate cover with antimanic mood stabilizing medications (i.e., a combination of two mood stabilizers or a mood stabilizer and an atypical antipsychotic). Although SNRIs such as duloxetine and venlafaxine are effective in primary anxiety disorder, their use in bipolar patients is not recommended due to their potential higher risk of mood destabilization.

Social Anxiety Disorder. While the estimates of social anxiety disorder (SAD) in bipolar disorder vary, the National Comorbidity Survey reported that 39 percent of patients with any type of bipolar disorder have SAD. There is virtually no controlled efficacy data in this comorbid population. CBT is very effective in treating primary SAD and likely useful in comorbid patients as well. Among the agents shown to be effective in treating primary SAD, pregabalin and gabapentin are preferred options as they have no propensity to destabilize bipolar disorder. In nonresponders, SSRIs could be considered provided patients with bipolar disorder are on adequate mood stabilization. Beta-blockers such as propranolol and atenolol are not recommended as they do not appear to be beneficial in primary SAD.

Panic Disorder. Between 14 to 27 percent of patients with bipolar disorder have panic disorder and up to 35 percent of patients with bipolar disorder report experiencing panic attacks. There is limited open-label data to support the use of valproate in bipolar disorder with comorbid panic

disorder. There is posthoc evidence that gabapentin is useful in severe PD. If optimization of anticonvulsant therapy is ineffective, addition of an atypical antipsychotic such as quetiapine or risperidone may be appropriate. As with other anxiety disorders, SSRIs are effective but these should be given preferably in conjunction with an atypical antipsychotic and possibly a mood stabilizer in order to minimize the risk of manic/hypomanic switch. Clonazepam, alprazolam, and lorazepam are also effective and they may be appropriate for the short-term use to control panic attacks.

Post-Traumatic Stress Disorder. Lifetime post-traumatic stress disorder (PTSD) rates range from 16 to 39 percent in patients with bipolar disorder. Childhood sexual abuse, adult sexual assault, and adult survival of the suicide, homicide, or accidental death of a close friend or relative are commonly associated with comorbid PTSD in bipolar disorder.

There is virtually no controlled data in bipolar disorder with PTSD comorbidity. Anticonvulsants are likely beneficial and should be preferred particularly in light of their efficacy in reducing cravings for alcohol, as comorbid alcohol abuse is common in patients with PTSD. Atypical antipsychotic mono or adjunctive therapy has been reported to be beneficial for primary PTSD particularly in treating symptoms of intrusion, and hence are appropriate options in anticonvulsant nonresponders. As with other anxiety disorders, SSRIs are effective but should be used preferably in conjunction with an atypical antipsychotic. Benzodiazepines are not effective and there is anecdotal evidence that they may worsen symptoms.

Obsessive–Compulsive Disorder and Related Disorders. About 10 to 25 percent of patients with bipolar disorder have a lifetime history of OCD. The symptoms of OCD are particularly exaggerated during depressive episodes. Sexual and religious obsessions tend to be more common in bipolar disorder and their severity tends to wax and wane over lifetime. Psychological treatments such as CBT should be offered first. There is no controlled data for pharmacotherapy in comorbid patient population, but there is some evidence that adjunctive treatment with lamotrigine, topiramate, or atypical antipsychotics such as risperidone, aripiprazole, olanzapine or quetiapine is effective in treating primary refractory OCD. Since these medications with the exception of topiramate are effective in treating bipolar disorder, they should be optimized first. In those that continue to manifest OCD symptoms, therapy with topiramate or a combination of an atypical antipsychotic with an SSRI (i.e., escitalopram, fluoxetine, fluvoxamine, paroxetine, and sertraline) may be considered. The need for ongoing SSRI treatment should be evaluated periodically, and the risk versus benefit analysis should determine the duration of SSRI

treatment. While clomipramine is effective in primary OCD, it is not recommended in bipolar patients due to higher risk of manic switch.

Alcohol and Substance Use Disorders

About 50 percent of patients with bipolar disorder have alcohol or substance use comorbidity. There is evidence that valproate is effective in reducing various indices of alcohol use in bipolar patients with comorbid alcohol use, and hence is a preferred option in this population. Lamotrigine is also beneficial for some patients with alcohol comorbidity. Other anticonvulsants such as gabapentin and topiramate have shown efficacy in primary alcoholism and could be used as adjuncts to primary treatments for bipolar disorder in some patients to address alcohol comorbidity. While naltrexone and disulfiram are effective in primary alcohol use disorder, disulfiram is not recommend in bipolar patients given its inhibitory effects on dopamine β-hydroxylase and the potential for worsening of mania. There is some evidence for naltrexone add-on therapy in bipolar patients with alcohol comorbidity. While quetiapine is effective in primary bipolar disorder, it does not appear to have any utility in reducing alcohol-related behaviors.

Valproate, lamotrigine, quetiapine, and risperidone mono or add-on therapy appears to provide some benefit in patients with cocaine abuse. Add-on NAC or modafinil may be useful in some patients.

Attention-Deficit Hyperactivity Disorder

The estimates for the prevalence of ADHD in patients with bipolar disorder range from 9 to 30 percent, and therefore, all patients with bipolar disorder should be routinely screened for comorbid ADHD. Bipolar patients with ADHD tend to have an earlier age at onset of mood symptoms, more frequent mood episodes, more suicide attempts, and other psychiatric comorbidities. ADHD symptoms persist during euthymia and contribute to impairment in function. The diagnosis of comorbid ADHD in bipolar patients can be challenging, given the overlapping symptoms. Therefore, clinicians need to be aware of the possibility of both under-diagnosis and over-diagnosis of ADHD in bipolar disorder. Persistent impulsivity and impaired ability for sustained attention despite treatment with mood stabilizing medications and the absence of other symptoms of mood episodes should raise the index of suspicion for ADHD.

If ADHD symptoms and functional impairment persist despite psychoeducation and behavioral interventions in bipolar patients optimized on mood stabilizing medications, institution of specific pharmacotherapy targeting ADHD symptoms is appropriate. However, clinicians must be aware that ADHD treatments can destabilize the course of bipolar disorder. Adjunctive bupropion is often the first choice,

particularly in bipolar patients with comorbid ADHD and frequent depressive episodes/depressive symptoms, as it is less likely to induce manic switch compared with stimulants. In nonresponders, adjunctive modafinil or mixed amphetamine salts or methylphenidate could be considered. If a patient responds, clinicians are advised to periodically evaluate the need for ongoing treatment with stimulants and offer psychoeducation to patients on monitoring for the emergence of manic symptoms.

FUTURE DIRECTIONS IN THE TREATMENT OF MOOD DISORDERS

The ultimate goal is a stratified approach to the personalized treatment of major depressive disorder, bipolar disorders, and other mood disorders. However, this is limited by the failure of clinical phenotypes to inform selection of psychotropic medication. While the search for validated molecular, neuroimaging, and other biomarkers to subgroups within the major depressive disorder and bipolar disorders is underway, this has not yet influenced treatment selection or patient outcomes. A dimensional approach to diagnosis proposed by the National Institute of Mental Health, known as Research Domain Criteria, should facilitate a dimensional approach to evaluating behavioral characteristics associated with different psychiatric disorders at cellular, molecular, neuroanatomical, and behavioral levels.

REFERENCES

Arnow BA, Blasey C, Williams LM, et al. Depression subtypes in predicting antidepressant response: a report from the iSPOT-D trial. *Am J Psychiatry*. 2015;172(8):743–750.

Cipriani A, Barbui C, Salanti G, et al. Comparative efficacy and acceptability of antimanic drugs in acute mania: a multiple-treatments meta-analysis. *Lancet*. 2011;378(9799):1306–1315.

Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet*. 2009;373(9665):746–758.

Cipriani A, Reid K, Young AH, Macritchie K, Geddes J. Valproic acid, valproate and divalproex in the maintenance treatment of bipolar disorder. *Cochrane Database Syst Rev*. 2013;10:CD003196.

Durgam S, Earley W, Lipschitz A, Guo H, Laszlovszky I, Németh G, Vieta E, Calabrese JR, Yatham LN: A Randomized, Double-blind, Placebo-Controlled Evaluation of the Safety and Efficacy of Cariprazine in Patients with Bipolar Depression. *American Journal of Psychiatry*, 2016 Mar 1;173(3):271–281.

Gaynes BN, Warden D, Trivedi MH, Wisniewski SR, Fava M, Rush AJ. What did STAR*D teach us? Results from a large-scale, practical, clinical trial for patients with depression. *Psychiatr Serv*. 2009;60(11):1439–1445.

Geddes JR, Calabrese JR, Goodwin GM. Lamotrigine for treatment of bipolar depression: independent meta-analysis and meta-regression of individual patient data from five

randomised trials. *Br J Psychiatry*. 2009;194(1):4–9.

Ghaemi SN, Ostacher MM, El-Mallakh RS, et al. Antidepressant discontinuation in bipolar depression: a Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) randomized clinical trial of long-term effectiveness and safety. *J Clin Psychiatry*. 2010;71(4):372–378

Gignac A, McGirr A, Lam RW, Yatham LN. Recovery and recurrence following a first-episode of mania: a systematic review and meta-analysis of prospectively characterized cohorts. *J Clin Psychiatry*. 2015;76(9):1241–1248.

Ginsberg LD. Impact of drug tolerability on the selection of antidepressant treatment in patients with major depressive disorder. *CNS Spectr*. 2009;14(12):8–14.

Kennedy SH. A review of antidepressant therapy in primary care: current practices and future directions. *Prim Care Companion CNS Disord*. 2013;15(2):pii: PCC.12r01420.

Kennedy SH, Young AH, Blier P. Strategies to achieve clinical effectiveness: refining existing therapies and pursuing emerging targets. *J Affect Disord*. 2011;132:S21–S28.

Kennedy SH, Lam RW, McIntyre RS, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 3. Pharmacological Treatments. *Can J Psychiatry*. 2016;61(9):540–560.

Lam RW, McIntosh D, Wang J, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 1. Disease Burden and Principles of Care. *Can J Psychiatry*. 2016;61(9):510–523.

Lam RW, Parikh SV, Michalak EE, Dewa CS, Kennedy SH. Canadian Network for Mood and Anxiety Treatments (CANMAT) consensus recommendations for functional outcomes in major depressive disorder. *Ann Clin Psychiatry*. 2015;27(2):142–149.

McIntyre RS, Rosenbluth M, Ramasubbu R, et al; Canadian Network for Mood and Anxiety Treatments (CANMAT) Task Force. Managing medical and psychiatric comorbidity in individuals with major depressive disorder and bipolar disorder. *Ann Clin Psychiatry*. 2012;24(2):163–169.

Ng F, Mammen OK, Wilting I, et al. The International Society for Bipolar Disorders (ISBD) consensus guidelines for the safety monitoring of bipolar disorder treatments. *Bipolar Disorders*. 2009;11(6):559–595.

Pacchiarotti I, Bond DJ, Baldessarini RJ, et al. The International Society for Bipolar Disorders (ISBD) Task Force report on antidepressant use in bipolar disorders. *Am J Psychiatry*. 2013;170(11):1249–1262.

Parikh SV, Huniewicz P. E-health: an overview of the uses of the Internet, social media, apps, and websites for mood disorders. *Curr Opin Psychiatry*. 2015;28(1):13–17.

Raison CL, Rutherford RE, Woolwine BJ, et al. A randomized controlled trial of the tumor necrosis factor antagonist infliximab for treatment-resistant depression: the role of baseline inflammatory biomarkers. *JAMA Psychiatry*. 2013;70(1):31–41.

Rizvi SJ, Kennedy SH. Emerging drugs for major depressive disorder: an update. *Expert Opin Emerg Drugs*. 2012;17(3):285–294. (PMID: 22530790)

Schaffer A, Isometsä ET, Azorin J, et al. A review of factors associated with greater likelihood of suicide attempts and suicide deaths in bipolar disorder: Part II of a Report of the International

Society for Bipolar Disorders Task Force on Suicide in Bipolar Disorder. *Aust N Z J Psychiatry*. 2015;49(11):1006–1020.

Schoeyen HK, Kessler U, Andreassen OA, et al. Treatment-resistant bipolar depression: a randomized controlled trial of electroconvulsive therapy versus algorithm-based pharmacological treatment. *Am J Psychiatry*. 2015;172(1):41–51.

Sim K, Lau WK, Sim J, Sum MY, Baldessarini RJ. Prevention of Relapse and Recurrence in Adults with Major Depressive Disorder: Systematic Review and Meta-Analyses of Controlled Trials. *Int J Neuropsychopharmacol*. 2015;19(2). pii: pyv076.

Swartz MA, Thase ME. Pharmacotherapy for the treatment of acute bipolar II depression: current evidence. *J Clin Psychiatry*. 2011;72(3):356–366.

Yatham LN, Beaulieu S, Schaffer A, et al. Optimal duration of risperidone or olanzapine adjunctive therapy to mood stabilizer following remission of a manic episode: A CANMAT Randomized double blind placebo controlled trial. *Mol Psychiatry*. 2015; doi: 10.1038/mp.2015.158.

Yatham LN, Kennedy SH, Parikh SV, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2013. *Bipolar Disorders*. 2013;15(1):1–44.

Zhou X, Ravindran AV, Qin B, et al. Comparative efficacy, acceptability, and tolerability of augmentation agents in treatment-resistant depression: systematic review and network meta-analysis. *J Clin Psychiatry*. 2015;76(4):e487–e498.

▲ 13.8 Mood Disorders: Psychotherapy

JOHN R. MCQUAID, PH.D., AND MICHAEL E. THASE, M.D.

Psychotherapy, in various forms, has been one of the cornerstones of therapeutics for mood disorders for decades. Even before controlled studies were available to document the utility of these approaches, there was substantial clinical support or face validity for treating depression psychotherapeutically; depression frequently occurs during times of significant psychosocial stress and often is associated with relevant intra- and interpersonal vulnerabilities. As the field has become more evidence-based and awareness of issues such as cost-effectiveness and disseminability has grown, greater emphasis has been placed on models of therapy that are procedurally specified and empirically tested. These trends were first evident in the introduction of shorter-term or time-limited therapies, particularly those that emphasize more pragmatic behavioral, cognitive behavioral or interpersonal formulations, and interventions. More recently, time-limited and manual-based models of psychodynamic psychotherapy have been introduced for the treatment of mood disorders. All of these therapies aim to improve upon the considerable beneficial effects of supportive clinical management, with the potential for theoretically driven, mode-specific benefits in addition to relief of depressive symptoms. This chapter will review the current status of these therapies.

HISTORY OF PSYCHOTHERAPY FOR DEPRESSION

Psychotherapy for depression arose from psychoanalytic and subsequent psychodynamic models of intervention. Early theories conceptualized depression as an intrapsychic phenomenon stemming from internalized anger that arose from unconscious conflicts. Therefore, treatment was long-term, focused on developing awareness of these internal conflicts through identification and interpretation of historical experiences that reflected such conflicts. Intervention emphasized developing insight into these conflicts through the use of free association and the therapist's interpretations. This treatment also used the patient's projections onto the therapist (the transference neurosis) as a means of studying early object relationships and conflicts.

In contrast to more traditional approaches, behavioral theories of depression treatment emphasized observable factors, such as decreased participation in potentially rewarding or pleasurable activities, and suggested that the negative emotional symptoms of depression could be understood as a consequence of withdrawal of reinforcement. Treatment

interventions initially were guided by learning theories, but by the 1970s had broadened to include cognitive components (e.g., to increase self-efficacy or decrease negative thoughts or beliefs about self, world, and future). Some models emphasized social skills or problem-solving strategies, whereas others focused more on increasing positive activities or remediating cognitive symptoms and vulnerabilities. Whatever the differences in specific models of CBT, all emphasized the role of the therapist as a teacher of specific skills that are intended to help the depressed person reduce depressive symptoms and/or improve coping. Moreover, virtually all models of CBT emphasize the importance of homework assignments to enhance mastery and generalization of these skills in "real world" environments.

Arguably the most impactful of the models of CBT to emerge from the 1970s was the approach developed by Beck and colleagues. More commonly known as cognitive therapy (CT), this form of therapy focused more on identifying the negatively valenced automatic thoughts associated with depressed moods and using strategies to both test the accuracy of the negative thoughts and consider more rational (and less depressogenic) alternatives. As many people in the midst of a depressive state evince various kinds of cognitive distortions (e.g., thinking in a more overly general and absolute way or selective recall of negative mood relevant memories or associations), a skilled CT therapist can help most nonpsychotically depressed people see examples of "cause and effect" relationships between cognition, mood, and behavior within the first hour of treatment.

Whereas early models of behavior therapy either minimized the value of the therapeutic relationship or were silent about its importance, Beck and colleagues emphasized a collaborative approach between the therapist and the patient, referred to as collaborative empiricism, and specifically structured therapy sessions in discrete 10 to 20 minute segments to ensure that the patient had the opportunity to give feedback (e.g., whether the particular topic was personally relevant or made sense) and express reservations. A collaborative process using Socratic questioning was similarly incorporated to guide and facilitate mastery of techniques for testing and remediating negative thoughts and beliefs, and the patient practices these techniques.

Beyond CT's initial focus on automatic negative thoughts ("surface" cognitions), Beck's model of therapy also emphasized the vulnerability to depression was linked to "deeper" forms of cognition, including dysfunctional attitudes and schema. Schemas are underlying cognitive structures that help the brain organize information in memory, guide awareness, and predispose action. The complex thoughts and actions involved in tying a bow are an example of a simple, nonaffectively relevant schema. Schemas that are associated with depression are often reflected by

automatic negative thoughts or abiding beliefs about competence ("I'm just not good enough") or intimacy ("No one will ever really love me"), which may be activated by thematically relevant life events. Thus, another distinguishing feature of CT is a specific stress-diathesis model of vulnerability. As such schemas are "silent" (i.e., they must be inferred from the predominant themes or beliefs revealed in automatic negative thoughts), and Beck's model of CT did deviate significantly from the zeitgeist of the behavior therapy movement. Nevertheless, the fully developed model of therapy developed by Beck and colleagues incorporated more conventional behavioral strategies such as activity scheduling and graded task assignments.

Interpersonal psychotherapy (IPT), developed by Klerman, Weissman and colleagues, emerged in the 1980s as the leading alternative to CBT. Drawing upon psychodynamically influenced interpersonal models (Meyer and Sullivan) and attachment theory (Bowlby), IPT incorporated many of the more pragmatic methods used by social workers in case management to develop a time-limited therapy that addresses the common problematic patterns in relationships that plague the lives of people with depression, including unresolved grief, role disputes, role transitions, and interpersonal deficits. Beyond eliciting an interpersonal inventory and identifying the area or areas of interpersonal difficulty of greatest relevance to a particular patient, the interventions used in IPT are relatively eclectic and can include psychoeducation, nondirective questioning and empathic support, and role playing and social problem solving. IPT thus may be particularly well-suited for more traditionally trained psychotherapists or counselors who find it difficult to implement the structure, activity level, and the more directive posture that characterize most models of CBT. Significantly, IPT was the first psychotherapy outside of CBT to be evaluated in comparison to antidepressant medications using RCT method.

Several adaptations of IPT have particular utility of patients with more complex depressive disorders. Frank and colleagues incorporated the importance of social rhythm disruption and the adaptive value of social rhythm stability (i.e., the consistency of regular activities such as bedtime and meal times) in their adaptation of IPT, which was primarily targeted for recurrent forms of affective illness, including bipolar disorder. Known as interpersonal and social rhythm therapy (IPSRT), this approach adds an assessment of social rhythm stability and activity scheduling tailored to strengthen the role of social *zeitgebers* (time givers) in helping patients prevent recurrent affective episodes and promote recovery. Shear and colleagues adapted IPT to focus on complicated bereavement, including clinical presentations in which the thoughts, feelings, and behaviors pertaining to bereavement overlap with the signs/symptoms of PTSD.

More recent innovations in psychotherapy have incorporated techniques from a long-standing meditation tradition, and done so using

an evidence-based approach. Mindfulness meditation, developed from Buddhist techniques, emphasizes two skills: observing one's perceived sensations and accepting and experiencing those sensations nonjudgmentally. In the application of this technique to depression, sensations include cognition and emotional states, as well as input through the five senses. Segal, Teasdale, and colleagues propose that these techniques are relevant for three reasons. First, the training in observation facilitates gathering evidence, which is a central component in CBT. Second, learning to view thoughts—even the most negative or damning ones—as discrete, limited experiences helps patients dampen affective arousal in response to them, which in turn improves their ability considering alternate perspectives. Third, the emphasis in meditation on nonjudgmental observing moves the therapy toward a focus on acceptance rather than change per se. For more chronic depressive disorders, a focus on acceptance—rather than change—can offer a novel therapeutic component, particularly when ability to tolerate distress has interfered with coping behaviors.

Several models of treatment now incorporate mindfulness techniques and are often referred to as "third wave" techniques, reflecting the path from behavioral to cognitive to now mindfulness and acceptance-based interventions within the CBT tradition. Mindfulness-based cognitive therapy (MBCT) is a treatment intervention specifically developed to reduce risk of relapse in individuals who have a history of recurrent depression. Acceptance and commitment therapy (ACT), developed by Steven Hayes and colleagues, emphasizes identifying core values and committing to acting on those values rather than focusing on avoiding suffering. This model was particularly designed for addressing chronic depression. Another closely related therapy, dialectical behavior therapy (DBT), has recently been adapted for work with patients with more treatment-resistant depressions. Initially developed by Linehan and colleagues, DBT is a behaviorally oriented system of therapy that also places a strong emphasis on mindfulness and acceptance and has well-replicated effects on reduction in suicidal ideations and behaviors.

Family support is an important asset in life and is generally a favorable prognostic indicator across mood disorders. Conversely, there is an increased risk of relapse/recurrence for patients whose families show high levels of criticism and emotional expression. Drawing upon these observations, Miklowitz and colleagues developed FFT, which has primarily been adapted for the treatment of adolescents and young adults with bipolar disorder. This treatment focuses on educating both patients and family members on the disorder and helping them gain a shared perspective of the disorder. From this shared perspective, the therapist guides the family in developing a relapse prevention plan, including ways that the family can identify symptom exacerbations, an agreement about

how they will share this observation with the patient, and an understanding of how the patient will respond to this information. Options for managing symptoms are generated (e.g., visiting the doctor for medication adjustment or increasing structure in the home). Family members are also trained in communicating effectively in ways with the patient and reducing the level of hostility and criticism in the environment.

REVIEW OF RESEARCH FINDINGS

Unipolar Disorders

Individual Psychotherapy. A number of forms of psychotherapy (see Table 13.8–1) have been shown to be either efficacious or probably efficacious for outpatients with major depressive disorder. These treatments have been shown to be effective in outpatients treated in a variety of settings (e.g., psychiatric clinics, psychology clinics, college counseling centers, community mental health centers, and primary care clinics) and meta-analyses suggest that, across modes of therapy, differences tend to be modest to nonexistent. By contrast, no form of psychotherapy is indicated as the primary treatment of bipolar disorder, although a growing literature supports the value of several different forms of psychotherapy as adjuncts to ongoing therapy with mood stabilizers.

CBT has been evaluated in numerous RCTs, including placebo-controlled studies that included first- and second-generation antidepressants as active comparators. Although the findings of one very influential multicenter trial, the National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Research Program (TDCRP), suggested that CT may be less effective than pharmacotherapy in patients with more severe depressive symptoms, the findings of a number of other RCTs do not support such a relationship and, overall, outpatients with higher symptom severity scores tend to respond slower or less favorably to many different interventions. In general, depressed patients treated with CBT are as likely to respond/remit across 8 to 16 weeks of treatment as are patients treated with antidepressant medications.

Table 13.8–1.
Efficacious Treatments for Major Depression

Treatment	Conceptualization of Disorder Etiology	Sample Interventions
Behavioral therapy	Deficit of reinforcers, including pleasant activities and positive interpersonal contacts	Increase activity level Structured goal setting Interpersonal skills training
Cognitive-behavioral therapy	Interaction of beliefs with matching stressor	Identify and challenge automatic thoughts Engage in activities that provide evidence disproving dysfunctional

		beliefs Modify core beliefs by reviewing evidence
Interpersonal psychotherapy	Interpersonal vulnerabilities arising from early attachment and learned relationship patterns	Develop awareness of patterns in primary relationships and the therapeutic relationship Interpersonal skills training Communication analysis
Behavioral marital therapy	Marital distress increases stress while impairing support resources	Assertive communication training Active listening exercises Problem-solving skills Increasing reinforcing behaviors toward spouse

A number of studies have likewise demonstrated that IPT also is an effective treatment of depressed outpatients. Although results of the NIMH TDCRP study revealed trends favoring IPT over CT among the more severely symptomatic subset of patients, the findings of several other subsequent comparisons have not revealed any meaningful differences. Although there are numerically more studies of CBT than IPT for most depressive disorders, there is a particularly strong literature supporting the use of IPT for the treatment of depressions during and after pregnancy.

The results of several studies suggest that simpler behavioral interventions may be as effective for depressed outpatients as more elaborate models of CBT that include cognitive restructuring. The performance of a “pure” behavioral activation (BA) intervention was particularly strong for patients with more severe levels of depression in one study. Whether BA and CT might be sequenced for patients at increased risk of relapse or recurrence is an interesting topic for future research, particularly because it is now possible to provide a lead a portion of the curriculum covered in CBT over the internet, with additional benefit resulting from telephone support.

Although RCTs of time-limited variations in psychodynamic psychotherapy were sparse in recent years, there is now a sufficient body of research to conclude that these are efficacious for the treatment of depressed outpatients (as compared to waiting list control conditions) and have outcomes across 8 to 16 weeks that are comparable to CBT or IPT. To date, no reliable indicators of preferential responsiveness to psychodynamic psychotherapy versus other forms of psychotherapy have emerged.

Across studies and types of psychotherapies, there is evidence that a response to psychotherapy may be more durable or persistent for weeks or months after therapy is stopped than is the case following withdrawal of antidepressant medications. As a result, differences in the cost of delivering care that might favor pharmacotherapy in the short-run may dissipate or actually shift in favor of psychotherapy when the need for prophylactic

therapy is taken into account. Ongoing psychotherapy may, however, be useful when a typically adequate course of individual therapy (8 to 16 weeks) only leads to an incomplete remission. In one study of incompletely remitted patients judged to be at high risk for relapse, the provision of 8 months of continuation phase CBT had a relapse prevention effect that was comparable to pharmacotherapy. Whereas psychotherapies may not warrant ongoing care for relapse prevention, it is less certain that an effective course of time limited psychotherapy actually reduces the longer term risks of recurrent depression. In the classic study of maintenance phase therapies conducted by Frank and colleagues, patients with highly recurrent forms of major depressive disorder who recovered with the combination of IPT and imipramine were significantly more like to remain well when their maintenance treatment included the active imipramine than when they received monthly sessions of "maintenance" IPT. Nevertheless, other studies have indicated that the addition of CBT or MBCT during an ongoing course of continuation or maintenance phase therapy may improve residual symptoms and reduce the risk of recurrent illness after discontinuation of antidepressant medication.

Group and Couples Therapy. A number of studies indicate that cognitive and behavioral therapies are well-suited for delivery in groups. Group delivery of CBT may not only be a cost-effective alternative to individual therapy, but may have particular value for groups with subsyndromal symptoms, for whom individual treatment might be judged not to be necessary. Depressed individuals with marital discord similarly have been shown to benefit from behavioral marital therapy (BMT). In controlled studies, BMT appears to have an impact on depressive symptoms that is comparable to individual CBT, with the added value of greater improvements in marital functioning than CBT

Combined Psychotherapy and Pharmacotherapy. Many psychiatrists favor combining psychotherapy and pharmacotherapy, particularly for the treatment of outpatients with more chronic, severe, or treatment-resistant forms of depression. Although the effect sizes typically favoring combined treatment over competently administered monotherapies have tended to be modest, in a pooled analysis of participants from a series of studies conducted at the University of Pittsburgh the advantage of combined treatment was particularly large for patients with more severe, recurrent depressive disorders. Results of several larger scale, multicenter studies likewise showed advantages for patients with more severe or chronic depressions. Similarly, several studies that have adapted IPT and CBT to inpatient settings have shown that the addition of focused therapy improves outcomes in comparison to treatment as usual conditions that include antidepressant medication and milieu therapies. As noted earlier,

sequential studies of pharmacotherapy followed by focused psychotherapy have likewise shown advantages for promoting recovery or facilitating discontinuation of pharmacotherapy.

Bipolar Disorders

In addition to the models of psychotherapy described earlier, group psychoeducation has emerged as a cost-effective form of adjunctive intervention for people with bipolar disorder. As developed by Colom and colleagues in Barcelona, this practical treatment can be conducted in groups as large as 30 to 50 participants and their families and focuses on risk factors, warning signs, and coping strategies. Although the acute efficacy of psychoeducation has never been studied for bipolar depression, it has been shown in several studies to have a significantly positive effect on treatment adherence and relapse prevention.

With respect to the treatment of bipolar disorder, no form of psychotherapy has been shown to be effective as a monotherapy for the treatment of mania or mixed states. Like group psychoeducation, most formal therapies have been studied as adjuncts to mood stabilizers with the goal of prevention of relapses or recurrences. Looking across interventions, the evidence is fairly convincing that adjunctive therapies, including psychoeducation, IPSRT, FFT, and CBT, significantly increase well time and reduce the risk of relapses and recurrence. Although one should be cautious comparing across studies, it appears that the magnitude of benefit for adjunctive psychotherapy, as compared to no adjunctive treatment or treatment as usual alone, may be comparable to the prophylactic effects of mood stabilizers such as lamotrigine, compared to placebo in contemporary studies (see Table 13.8–2).

It is less certain if focused psychotherapies are effective treatments of acute bipolar depression, although the same caveat could be stated about pharmacotherapy with antidepressants or mood stabilizers. In one component study of the large-scale US study known by the acronym STEP-BD, adjunctive psychotherapy with FFT, IPST, or CBT had significant effects on both depressive symptoms and social functioning. The nature of this study did not permit valid comparison among the psychotherapies. Given the hazards of treatment-emergent affective switches and the potential for inducing rapid cycling associated with antidepressants in bipolar disorder, the potential for focused psychotherapies for adjunctive treatment of bipolar I depressive episodes warrants further exploration. Likewise, the use of focused psychotherapy alone for selected patients with bipolar II depressive episodes deserves further study.

Table 13.8–2.
Efficacious Treatments for Bipolar Disorder

Treatment	Conceptualization of Disorder Etiology	Sample Interventions
Cognitive-behavioral therapy	Biological vulnerability interacting with stress Skills deficits limit ability to manage symptoms	Identify and challenge automatic thoughts that interfere with treatment adherence Engage in rewarding activities that provide increased routine and stability Practice communication skills with providers
Interpersonal and social rhythm therapy	Interpersonal vulnerabilities arising from early attachment and learned relationship patterns, plus disruption of social rhythms	Develop awareness of patterns in primary relationships and the therapeutic relationship Track and stabilize social rhythms Interpersonal skills training Communication analysis
Family focused therapy	Biological vulnerability exacerbated by negative expressed emotion in family environment.	Education regarding the disorder, including precipitants, risk factors, and effective treatment Establish relapse prevention plan, agreed upon by all involved family members Communication skills training Problem-solving skills training

APPLYING THE RESEARCH TO CLINICAL PRACTICE

There have been important advances in the understanding the roles of psychotherapy in the treatment of mood disorders, and several good psychotherapy options for both unipolar and bipolar disorders now exist. Nonetheless, clinicians are often faced with providing treatment to patients who may not match the typical participant in research trials and need to extrapolate from the available data to develop a treatment plan for their patient. Of course, several important critical questions remain, including which treatment is best for which patient, what are optimal doses (e.g., number and length of sessions), and how best to address comorbid disorders. It is also not clear how critical-specific techniques associated with these interventions are for the successful treatment of mood disorders. A study of the therapeutic process of IPT (i.e., the actual in-session behaviors of the therapist and patient) found that the process was more consistent with ideal CBT than ideal IPT. A meta-analysis examining comparisons between specific psychotherapies and placebo conditions found that the difference between treatments was related to the adequacy of the placebo design. Studies with placebo comparisons deemed adequate (e.g., using nonspecific therapy techniques for the placebo condition and matching for level of contact with therapists) showed smaller differences between active interventions and placebo than those studies with placebo conditions that were inadequate (e.g., less therapist contact or constraint of topics to issues not related to the patient's disorder). These findings suggest that differences between psychotherapy techniques may be less critical

than some of the shared components (e.g., positive therapeutic relationship, shared goals, and facilitated behavior change). There are also no definitive answers regarding the underlying models of depression. Although BA, CBT, and IPT are based on empirical models of psychopathology that all have supportive evidence, the development of these treatments and the conceptualization of depression are extrapolations beyond the available data.

Clinicians then are faced with how to use this knowledge in treating patients. One key approach, regardless of model, is the development of a case formulation to guide treatment.

TREATMENT OF MAJOR DEPRESSIVE DISORDER WITH CBT

CBT is based on the assumption that dysfunctional beliefs, learned in response to life experiences, lead to dysfunctional behavioral and emotional responses. Treatment therefore modifies these beliefs to reduce problematic responses and to increase functional responses. Patients learn to modify beliefs by first learning to observe their thinking, affect, and behavior and then evaluating how thoughts influence mood and actions. As the patient becomes aware of these processes, the therapist teaches skills in questioning and testing the validity of beliefs and developing alternative, more accurate, functional beliefs. The long-term goal of treatment is for patients to learn the skills to such an extent that they are able to be their own therapist.

Assessment and measurement, by the therapist and the patient, are central to effective CBT. The therapist teaches the patient to quantify the issues of concern. Patients learn to observe and define their own assumptions that lead to depression and then to engage in a series of experiments to improve their mood. Patients learn to measure the effects of these experiments, so that they test the previous assumptions and can develop new, more constructive beliefs and actions that reduce depression and other problematic moods and behaviors.

A 23-year-old, unmarried, African-American woman presented to a community mental health clinic. The patient was in the midst of her first semester at law school. Her primary physician had referred her to the clinic after the patient reported frequent crying and feeling hopeless. Symptoms included depressed mood nearly every day; crying spells “out of the blue”; anhedonia; early, middle, and late insomnia; decreased appetite; weight loss; difficulty concentrating; feeling worthless; thoughts of death; and suicidal ideation.

CBT, like most other psychotherapies, uses a case conceptualization approach to determining treatment of each patient. In CBT, the conceptualization is developed based on an understanding of how different aspects of cognition are related. Aaron Beck and, later, Judith Beck proposed that cognition could be divided into three levels: automatic thoughts, intermediate beliefs, and cognitive schemata. The most

accessible, labeled, automatic thoughts represented the conscious response to stimuli. For example, the previously mentioned patient, when studying, often thought, "I'll never understand this."

The second level, intermediate beliefs, is assumptions about the self, the world, and the future that led to the automatic thought occurring in response to a particular stimulus. Using the previous example, the patient believed that she was not competent and that she was not good enough; she perceived the world (in this case, professors and classmates) as attacking and critical. She anticipated that the future would be filled with failure. At times when things went well, she was not unduly distressed. However, in the face of failure or criticism, she quickly experienced self-critical thoughts that were consistent with these beliefs.

The content (e.g., the beliefs) and the organization of that content define the third level, cognitive schema. An individual's schema determines which stimuli are most likely noticed and encoded in memory, which stimuli are ignored or discounted, how encoded information is linked or associated in memory, and which memories are most easily recalled. For the previously mentioned patient, she may suffer from a schema that could be conceptualized as, "I'm not good enough." She tends to notice information consistent with this belief and tends to ignore or to discount other information.

As the therapist completes an initial assessment and establishes a treatment plan, a central goal is to elaborate a case conceptualization that will facilitate treatment. By understanding the unique meaning of experiences to the patient, the therapist can be more effective in identifying interventions that truly test the beliefs and produce meaningful improvement.

Several authors have described models for developing case formulations in CBT, and they generally are based on gathering evidence to propose an underlying schema. Once the automatic thoughts, beliefs, and schema are identified, interventions are chosen based on these hypothesized structures. The therapist continues to gather evidence throughout treatment, refining the formulation over time.

Initial Assessment

Specificity and accurate data gathering are a key focus of CBT. In particular, assessment of diagnosis and symptoms is necessary. Primary assessment preferably includes standardized measures, such as the Structured Clinical Interview for DSM-5 (SCID), the Hamilton Rating Scale for Depression, and the Beck Depression Index (BDI). The assessment also needs to include the dimensions of functioning: interpersonal relationships, work and achievement, health, and recreation. As treatment progresses, a formulation is developed regarding the patient's case conceptualization, as well as particular cognitive and behavioral coping

strategies that the patient uses.

The previously presented patient experienced impairments in several areas of functioning. The patient had not contacted family or friends for several weeks, which was unusual for her. She had no current close friends at her school and had not had a dating relationship for several years. She reported that she was skipping classes and that she believed she was failing her program. The patient made statements such as, "I'm just a total loser" and "I don't belong." Despite excellent grades in college, she believed she was only accepted to law school due to her minority status. She primarily coped with stress by ruminating and had stopped many rewarding activities (e.g., exercise).

She had several relevant experiences from her early childhood. The patient had been adopted shortly after her birth by a white couple and was their only child. She grew up in an almost exclusively white neighborhood. She reported that, although she had some friends, she had no African-American friends and that she often felt discriminated against by whites and the few African-Americans with whom she had contact, who criticized her for "trying to be white." The patient described that, despite having devoted parents, she had a sense of not being good enough and being out of place since approximately 6 years of age, when she first started school.

Course of Therapy

In the initial session, the therapist has several goals, including (1) orienting the patient to the treatment approach, (2) establishing rapport, (3) identifying problems and treatment targets, (4) assessing current symptom and problem severity, (5) providing some initial symptom relief, and (6) introducing an initial homework assignment. In reality, these components tend to overlap. For example, it is often helpful to illustrate the CBT model using a life experience of the patient and, in doing so, to provide a chance at some symptom relief. An example could be processing a situation such as being late to a first therapy appointment:

Therapist: You mentioned that you were upset because you were late. Can you tell me what you were thinking when you realized you were late?
Patient: I figured you'd be angry.
Therapist: Okay, and when you thought I'd be angry, how did you feel?
Patient: Pretty nervous.
Therapist: So that is what we will focus on in here, how your thinking in different situations influences how you feel. In this case, you were late, thought "he'll be angry," and felt nervous. Our goal is to test out the beliefs, like "he'll be angry," and change them when they aren't healthy or accurate. The great thing is that you sort of implicitly tested out the belief that I'll be angry by showing up. Was I angry?
Patient: You didn't seem angry.
Therapist: How anxious did you feel once we started talking?
Patient: Well, I'm feeling more comfortable.
Therapist: What thoughts do you have now about me being angry?
Patient: I don't think you're angry.
Therapist: So, this is an example of what we're going to be doing in therapy—I'll be helping you to identify thoughts that make you feel bad. We're then going to work together to come up with ways to check them out and change them if they're not true or accurate. You checked out whether I was angry by observing me, and you changed your thought, and, now, I get the impression that you feel better.

In this example, the model of CBT is described for the patient and illustrated using her own thoughts, behaviors, and feelings. In particular, the effect of behavioral testing of beliefs (i.e., coming into session late and finding out if the therapist actually was angry) is pointed out. To increase her sense of efficacy, the therapist points out that her current actions can lead to mood change (feeling more relaxed after coming in late) and that she is already demonstrating some of the necessary skills to benefit from CBT.

To facilitate skills, an agenda is explicitly set forth and agreed on by the therapist and the patient every session. In the early sessions, the patient is oriented and trained in thought tracking and activity tracking. The decision of which thoughts or behaviors are targets for intervention is based on the specific needs of the patient, which the therapist assesses using the case formulation. In the case described, the patient reported feeling "miserable" when she arrived at class. The therapist used the situation to train the patient in using a thought record, as shown in Figure 13.8–1. The specific situation is identified, and the emotional response and the thought or thoughts are recorded as well as the patient's belief in the thought.

Patients learn techniques for modifying unhelpful thinking, including identifying factual inaccuracies (e.g., "I don't understand anything" is all-or-nothing thinking) and biased information processing (ignoring evidence that she does know some things) and generating more accurate thoughts by the use of logic and testing beliefs.

Therapist: Okay, in this situation, you are feeling miserable. Now let's check this thought. What problems do you see with the thought, "I don't understand anything?"
Patient: Well, it's pretty all-or-nothing, but it feels like it's true.
Therapist: Can you think of any times when you felt something was true and it wasn't?
Patient: Sure, I felt sure I wouldn't get into law school or college for that matter.
Therapist: What does that tell you about this thought?
Patient: Well, it might not be true either.
Therapist: Okay, what would be other evidence that this thought isn't completely true?
Patient: Well, I do know some things.
Therapist: What is your evidence? Have you taken any tests or completed any assignments?
Patient: Yeah—and I got a B on the first test.
Therapist: So, what can we write in the rational responses column?
Patient: It isn't all-or-nothing; I do know some stuff.
Therapist: How miserable do you feel if you tell yourself that?
Patient: Well, a little less.

Situation	Emotions (strength, 0–100)	Automatic Thoughts (strength of belief, 0–100)	Rational Responses	New Emotion (0–100)
ENTERING CLASS	MISERABLE (70–80)	I DON'T UNDERSTAND ANYTHING GOING ON IN THIS CLASS! (100)	THAT'S ALL-OR - NOTHING AND EMOTIONAL REASONING IN THE PAST, SOME THINGS I FELT WERE TRUE DIDN'T TURN OUT TO BE TRUE I DO KNOW SOMETHING, OR I WOULD NOT HAVE GOTTEN A B.	MISERABLE (50)

FIGURE 13.8–1. Thought record example.

In this case, the patient experiences several types of distortion, including all-or-nothing thinking and emotional reasoning (i.e., "I feel it is true; therefore, it must be true"). The patient is asked to generate evidence that challenges the negative belief. The evidence reduces but does not eliminate the negative feeling. This is often the case, particularly when working with retrospective examples and patients experiencing significant depression. One approach to increase the efficacy of thought challenging and evidence gathering in CBT is by developing behavioral experiments to test the belief. In this case, the patient later conducted an experiment to test the "I don't know anything" belief. She wrote down her own answers to questions asked in the class as well as the answers that other students made. She found that her knowledge was on par with her classmates, and she was able to answer most of the questions. This experiment also addressed her "belonging," because she got evidence of her equivalence to her peers.

As treatment progresses, the therapist points out recurring themes and helps the patient to identify and challenge core beliefs. In this example, the patient repeatedly experienced thoughts about not being good enough. She said that she felt this way all her life and attributed this to her sense of being different from her parents and her European-American friends and her experience of persistent racism.

One approach to changing schema is to generate alternative core beliefs and to help the patient to identify behaviors and cognitions that are consistent with the new beliefs. In this case, a belief of "I belong just like everyone belongs; others may not believe it or act that way, but I do" was consistent with the patients' values and overriding experience. The therapist would therefore challenge thoughts by asking "is that the 'old self'

or 'new self' talking?" This provided a shorthand for identifying dysfunctional thoughts and generating new ones consistent with the alternative schema.

In the final sessions, the tasks shift to reviewing progress, identifying potential future challenges, and learning to administer interventions independently. Returning to the example, the patient showed significant improvement in her mood. Interventions that had been beneficial were attending the university gym on a regular schedule (activity scheduling), joining a study group (activity scheduling, assertiveness training, and social interaction), and gathering evidence regarding beliefs about being good enough and belonging (thought challenging and experiments). In addition, another student had invited the patient on several dates. At the end of therapy, the patient had a depression score on the BDI in the nondepressed range, her school attendance and grades had improved, and she had daily contact with friends and family.

In preparation for termination, it is important to identify the improvements that the patient has made, the skills that have been learned, and potential challenges that may arise in the future. This client identified her risk factors for increased depression as a rejection if she started dating or failure at school. She also noted that, although she now did not fear attending class and she was doing well, she still was not talking in class unless she was called on.

Therapist: Is not answering questions in class a problem?
Patient: Yeah.
Therapist: Okay, you've been doing this a while, what do you think would be a good homework for you?
Patient: Well, I guess I could just plan to answer one or two questions in class, and see what happens.
Therapist: When you say "see what happens," what could you track?
Patient: Maybe how many I get right and my thoughts and feelings after each question.
Therapist: Sounds good. What are some things you can do to make that easy in the class setting?

Because CBT is a skills-based model, it is critical that the patient demonstrates an ability to use skills independently of the therapist. In this case, she is able to take on the role of therapist and to assign herself a task that can help her to modify her beliefs and behavior.

ADJUNCTIVE PSYCHOTHERAPY FOR BIPOLAR DISORDER

Psychotherapy for bipolar disorder is an adjunct to pharmacotherapy for the management of symptoms. As such, interventions including CBT, FFT, and IPSRT emphasize education about the disorder, learning to track symptoms and warning signs of relapse, and improving adherence to pharmacotherapy. In CBT, additional interventions include identifying thoughts and beliefs that can either (a) exacerbate symptoms or (b)

increase the risk of relapse. In CBT and IPSRT, there is an emphasis on establishing stable routines to reduce triggers for mania. Finally, FFT emphasizes the role of educating the entire family system about the disorder and enlisting family members to aid patients with the management of their disorder. The current vignette describes a case from a cognitive-behavioral conceptualization that incorporates techniques from other approaches.

A 47-year-old, divorced Caucasian man was referred for psychotherapy by his treating psychiatrist. The patient was currently unemployed, although he volunteered 5 hours per week at his church. Although he had a long history of symptoms, this was the first time that he had been referred for psychotherapy. At the time of referral, the patient reported mixed symptoms of depression and mania. Depressive symptoms included depressed and sad mood, feelings of guilt, and suicidal ideation. Currently, the patient reported manic symptoms of irritability, racing thoughts, decreased need for sleep, and impulsive spending. In addition, the patient was preoccupied with religious ideation.

In considering integrative interventions, it is key to have an organizing structure with which to work. In this current case, the psychotherapist starts from a cognitive-behavioral framework but then incorporates techniques associated with other approaches based on the formulation.

Initial Assessment

As with a unipolar disorder, accurate assessment of diagnosis and symptoms is necessary for effective treatment. In addition, it is particularly important in the treatment of bipolar disorder to assess level of activity, stability of routine, and supportive social networks. This is also an important point to identify patient beliefs regarding bipolar disorder. These could include beliefs about benefits of mania, fears that treatment would eliminate these benefits, concerns about medication side effects, or questioning the diagnosis. These issues will be key targets of treatment in any intervention.

The current patient presented for the treatment secondary to several significant life events. He reported that he had several arguments with other volunteers at his church. This had led to a warning from his pastor that he would not be able to continue volunteering if these conflicts continued. The patient was also under severe financial stress. Although he received disability income, he had spent the entirety of his most recent check on religious memorabilia and charitable donations and was at risk of eviction from his subsidized housing.

The patient had a long history of disorder. He reported that he was raised in an intact family by his mother and father and had two younger sisters. He reported that his childhood was uneventful and his parents loving and supportive. He had done well in high school and had joined the army. He served 4 years in the military as a communications specialist. While in the military he married a German woman that he met while stationed in Germany.

His first symptoms occurred while in the military. He reported a manic episode, leading to hospitalization, characterized by religious delusions, elated mood, goal-directed behavior (attempting to convert other soldiers to his beliefs), excessive spending, alcohol and drug use, and sexual acting out. This first episode led to his medical discharge from the military and divorce from his wife. After discharge from the hospital, the patient reported ongoing

depressive symptoms extending several years, punctuated by periods of mixed depression and mania. He reported three additional hospitalizations, one following an additional manic episode and two following suicide attempts secondary to severe depression.

Course of Therapy

Psychotherapy for bipolar disorder is an adjunctive treatment to pharmacotherapy, and initial sessions reflect this. As with unipolar disorders, treatment included education about the treatment model, establishing rapport, and identifying problems and goals and assessing symptoms. In addition, the therapist both assesses knowledge about bipolar disorder and provides education to both expand on that understanding and correct any misconceptions. An example of such education is below.

Therapist: What do you know about your diagnosis of bipolar one?
Patient: Well, it means I have ups and downs, and they get pretty bad.
Therapist: What else do you know about it?
Patient: Well, my doctor says I have to take my meds or else it will get worse.
Therapist: How do you feel about taking your medication?
Patient: I really don't like to.
Therapist: How come?
Patient: I'm just not me when I'm taking them. I feel like I am slowed down and I just can't think. I just don't know if they help. I'm taking them now, and I don't feel much better. And they don't fix the real problems I've got.
Therapist: Which are?
Patient: I have no friends, I have no girlfriend, I have no job, I have no money, no one helps me, my family's disowned me. And when I'm taking my medication, I just become a lump, and none of that gets better.
Therapist: That sounds pretty depressing. I am going to write down those things you listed as items for your problem list, including job, money, dating, family, friends, and medications.
Patient: OK.
Therapist: And one thing that stands out to me is that you aren't happy with taking medications. That is an important thing for us to deal with as well.

In this example, the question about the patient's understanding of his disorder was to actually get information for the problem list. In addition, it elicits information about concerns regarding medication. The therapist here now has the opportunity to both demonstrate the development of the problem list and over time educate the patient about medication use.

One issue that complicates treatment when patients are receiving combined pharmacotherapy and psychotherapy is whether the psychotherapist is also the prescribing provider. In the case where there is one provider, coordination of pharmacotherapy and psychotherapy is obviously easier, but frequently the prescriber does not have time and/or expertise to provide psychotherapy. Therefore, frequently separate providers do psychotherapy and pharmacotherapy, and therefore it is critical for these providers to be in frequent contact and coordination.

From a CBT perspective, the key aspects of psychotherapy for bipolar

disorder will include tracking symptoms and functioning, identifying and modifying cognitions that may exacerbate symptoms or impair functioning, and behavioral interventions designed to increase rewarding activities without producing manic activation. To do this, interventions can include those from other models, including the education and involvement of family or other support resources, and stabilization of social rhythms through the establishment of consistent routine as in IPSRT. The example below reviews identifying strategies to reduce risk of relapse following recent symptom exacerbation.

Therapist: What do you think led to you feeling worse?
Patient: I think that I felt like I had to do all the work for the church to prepare for the holidays. I started doing more and more, and I got really stressed.
Therapist: All right, so taking on too much is a problem. One thing we're going to do today is work on a list of warning signs that your symptoms are getting worse. Then we have to figure out how to help you notice when they come up and do something about them. Do you have anyone around who you trust who could help you notice when this is happening?
Patient: No. You know that.
Therapist: I was thinking about co-workers, like one of the other volunteers or your pastor. You've mentioned that he has been really supportive, and he knows about your symptoms.
Patient: Yeah, I know he gets worried about me.
Therapist: One thing we can do is practice asking him to give you feedback about your actions and to let you know when he notices warning signs that your symptoms are getting worse.
Patient: I just worry that I'm asking too much.
Therapist: Well, let's practice in here how you would ask. Also, let's generate some other ways to notice warning signs....

In this situation, identifying warning signs can help the patient apply skills to prevent full symptom exacerbation. However, patients often don't notice as symptoms are worsening, and therefore it is useful to enlist family members or other support to provide feedback. For this to be effective, it is helpful to (a) meet with the support person and educate them about the disorder and the goal of tracking warning signs and (b) to work with both the patient and the support person to agree on how the support person will inform the patient about concerns and how the patient will respond. In the current case, the patient's pastor eventually attended a session, and the patient and he developed a plan for tracking symptoms and responding when symptoms were exacerbated. This both provided pragmatic support and also challenged the patient's perception that no one cared. In this case, the fact that the provider was the pastor was also beneficial given that the patient had symptoms of hyperreligiosity when becoming manic. By educating the pastor and patient regarding these symptoms, they could quickly notice when there was a change and intervene through thought challenging, working with structure, attending to medication adherence, and working with the prescribing provider to adjust medications.

During the course of treatment of bipolar disorder, the role of the therapist is to facilitate the patient identifying warning signs, implementing

skills practice, and providing support and encouragement (Table 13.8–3).

As with unipolar treatment, later sessions emphasize not only continued practice of learned skills but challenging of core beliefs that may contribute to the risk of symptom relapse. For the current patient, the belief "I'm damaged and will never be accepted by anyone" was a key theme that came up. He perceived his bipolar disorder as a flaw that meant that he could never have a "normal" life and that others would always reject him. Assumptions that arose from this belief were that he could never have a romantic relationship, that he could never reconnect with his family, and that he could never have paid employment. He also was ambivalent about his treatment (particularly medication), doubting that he could make any improvement. Therefore, the later sessions focused on both challenging unhelpful beliefs and developing alternatives to unhelpful core beliefs.

At the end of treatment, the goals are for the patient to be able to identify symptom changes, have a range of skills with which to respond to those changes, and have improved communication with their pharmacotherapist so that medication adherence is consistent and appropriate changes in pharmacotherapy can be made as needed. In addition, if family members are available, the goal is that the patient and family have developed improved understanding of the disorder and communication skills so as to facilitate stable mood.

Table 13.8–3.
Examples of Warning Signs

Warning Signs	Depressed	Warning Sign	Normal	Warning Sign	Manic
Sleep	Sleeping >10 hours	Trouble waking up, sleeping >9 hours	8 hours	>7 hours	No sleep and energized
Energy	Cannot get out of bed	Tired in the evening	Up at alarm clock, able to work through day	Waking up early and energetic	"Superman mode"
Spending	Isolating—not buying necessities	It is an effort to get food	Normal	Start buying treats for friends	Giving away all my money to charity
Mood	Crying easily	Pessimistic	"OK"	Optimistic	Euphoric
Guilt	Thoughts of failure	Sensitive to criticism	"OK"	Flippant, not concerned about others	"I am special in God's eyes"
Religion	Abandoned by God	Feeling cynical and doubting	Comfortable with faith	Wanting to share the good news	"I am God's messenger"—seeking out people to convert
Suicide	Start planning	Thoughts of	None	None	None ("I will live

to kill self being dead forever")

FUTURE DIRECTIONS

Some basic guidelines can be derived in treatment. In general, it is incumbent on the clinician to apply interventions for which there is empirical support. In the case of mood disorders, there are several psychotherapeutic options that are now available. There are some patient characteristics that could recommend treatment, including the presence of significant marital distress, amenability to a particular validated treatment modality, and chronic depression.

An important component of good psychotherapy, regardless of model, is that the therapist has a treatment rationale that guides decisions regarding interventions. This is a shared theme of all of the current validated treatments for mood disorders. However, such a model is not likely sufficient for effective treatment. Psychotherapies generally are built around mediational models of change, with modification of the target symptom (e.g., depressed mood) being achieved through intervention on a more accessible construct (e.g., an irrational belief, a rewarding activity, increase in medication adherence, or a transference reaction in therapy). What distinguishes excellent clinicians, regardless of orientation, is their ability to hold a conceptual framework in mind and to apply that framework to the unique challenges of the patient.

REFERENCES

*Barth J, Munder T, Gerger H, et al. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS Med*. 2013;10(5):e1001454.

Beck AT, Rush AJ, Shaw BF, Emery G. *Cognitive Therapy of Depression*. New York: Guilford Press; 1979.

*Beck JS. *Cognitive Therapy: Basics and Beyond*. New York: Guilford Press; 1995.

Biesheuvel-Leliefeld KE, Kok GD, Bockting CL, et al. Effectiveness of psychological interventions in preventing recurrence of depressive disorder: meta-analysis and meta-regression. *J Affect Disord*. 2015;174:400.

Colom F. The evolution of psychoeducation for bipolar disorder: from lithium clinics to integrative psychoeducation. *World Psychiatry*. 2014;13:90.

Corruble E, Swartz HA, Bottai T, et al. Telephone-administered psychotherapy in combination with antidepressant medication for the acute treatment of major depressive disorder. *J Affect Disord*. 2016;190:6.

Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, Dobson KS. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Can J Psychiatry*. 2013;58:376.

*Cuijpers P, Geraedts AS, van Oppen P, Andersson G, Markowitz JC, van Straten A. Interpersonal psychotherapy for depression: a meta-analysis. *Am J Psychiatry*. 2011;168:581.

Dimidjian S, Hollon SD, Dobson KS, Schmaling KB, Kohlenberg RJ. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *J Consult Clin Psychol*. 2006;74:658.

Dobson KS, Hollon SD, Dimidjian S, Schmaling KB, Kohlenberg RJ. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. *J Consult Clin Psychol*. 2008;76:468.

Driessen E, Hegelmaier LM, Abbass AA, et al. The efficacy of short-term psychodynamic psychotherapy for depression: a meta-analysis update. *Clin Psychol Rev*. 2015;42:1.

Driessen E, Van HL, Don FJ, et al. The efficacy of cognitive-behavioral therapy and psychodynamic therapy in the outpatient treatment of major depression: a randomized clinical trial. *Am J Psychiatry*. 2013;170:1041.

Ekers D, Webster L, Van Straten A, Cuijpers P, Richards D, Gilbody S. Behavioural activation for depression; an update of meta-analysis of effectiveness and sub group analysis. *PLoS One*. 2014;9(6):e100100.

Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM. NIMH Treatment of Depression Collaborative Research Program: I. General effectiveness of treatments. *Arch Gen Psychiatry*. 1989;47:971.

Frank E. *Treating Bipolar Disorder: A Clinician's Guide to Interpersonal and Social Rhythm Therapy*. New York: Guilford Press; 2005.

Frank E, Kupfer DJ, Perel JM, et al. Three year outcomes for maintenance therapies in recurrent depression. *Arch Gen Psychiatry*. 1990;47:1093.

Frank E, Kupfer DJ, Thase ME, Mallinger AG, Swartz HA. Two-year outcomes for interpersonal and social rhythm therapy in individuals with bipolar I disorder. *Arch Gen Psychiatry*. 2005;62:996.

Hollon SD, DeRubeis RJ, Fawcett J, et al. Effect of cognitive therapy with antidepressant medications vs antidepressants alone on the rate of recovery in major depressive disorder: a randomized clinical trial. *JAMA Psychiatry*. 2014;71:1157.

Hollon SD, DeRubeis RJ, Shelton RC, Amsterdam JD, Salomon RM. Prevention of relapse following cognitive therapy vs. medications in moderate to severe depression. *Arch Gen Psychiatry*. 2005;62:417.

Jarrett RB, Minhajuddin A, Gershenfeld H, Friedman ES, Thase ME. Preventing depressive relapse and recurrence in higher-risk cognitive therapy responders: a randomized trial of continuation phase cognitive therapy, fluoxetine, or matched pill placebo. *JAMA Psychiatry*. 2013;70:1152.

Keller MB, McCullough JP, Klein DN, Arnow B, Dunner DL. A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *N Engl J Med*. 2000;342:1462.

*Klerman GL, Weissman MM, Rounsaville BJ, Chevron ES. *Interpersonal Psychotherapy of Depression*. New York: Basic Books; 1984.

Kocsis JH, Leon AC, Markowitz JC, et al. Patient preference as a moderator of outcome for chronic forms of major depressive disorder treated with nefazodone, cognitive behavioral analysis system of psychotherapy, or their combination. *J Clin Psychiatry*. 2009;70:354–361.

Linehan MM, Korslund KE, Harned MS, et al. Dialectical behavior therapy for high suicide risk in

individuals with borderline personality disorder: a randomized clinical trial and component analysis. *JAMA Psychiatry*. 2015;72:475.

McCullough JP. *Treatment for Chronic Depression: Cognitive Behavioral Analysis System of Psychotherapy*. New York: Guilford Press; 2000.

*Miklowitz DJ, Goldstein MJ. *Bipolar Disorder: A Family Focused Treatment Approach*. New York: Guilford Press; 1997.

Miklowitz DJ, Otto MW, Frank E, Reilly-Harrington, NA, Wisniewski SR. Psychosocial treatments for bipolar depression: a 1-year randomized trial from the systematic treatment enhancement program. *Arch Gen Psychiatry*. 2007;64:419.

Mohr DC, Ho J, Duffecy J, et al. Effect of telephone-administered vs face-to-face cognitive behavioral therapy on adherence to therapy and depression outcomes among primary care patients: a randomized trial. *JAMA*. 2012;307:2278.

Reinares M, Bonnín CM, Hidalgo-Mazzei D, Sánchez-Moreno J, Colom F, Vieta E. The role of family interventions in bipolar disorder: a systematic review. *Clin Psychol Rev*. 2015;43:47.

Scott J, Colom F, Vieta E. A meta-analysis of relapse rates with adjunctive psychological therapies compared to usual psychiatric treatment for bipolar disorders. *Int J Neuropsychopharmacol*. 2007;10:123.

*Segal ZV, Williams JMG, Teasdale JD. *Mindfulness-Based Cognitive Therapy for Depression*. New York: Guilford Press; 2002.

Shear K, Frank E, Houck PR, Reynolds CF 3rd. Treatment of complicated grief: a randomized controlled trial. *JAMA*. 2005;293:2601.

Swartz HA, Levenson JC, Frank E. Psychotherapy for bipolar II disorder: the role of interpersonal and social rhythm therapy. *Prof Psychol Res Pr*. 2012;43:145.

Thase ME, Greenhouse JB, Frank E, et al. Treatment of major depression with psychotherapy or psychotherapy-pharmacotherapy combinations. *Arch Gen Psychiatry*. 1997;54:1009.

*Weitz ES, Hollon SD, Twisk J, et al. Baseline depression severity as moderator of depression outcomes between cognitive behavioral therapy vs pharmacotherapy: an individual patient data meta-analysis. *JAMA Psychiatry*. 2015;72:1102.

▲ 13.9 Mood Disorders: Neurobiology

MICHAEL E. THASE, M.D.

It has been thought for centuries that at least some forms of depression are either caused by or maintained by disturbances of brain function, and since the 1960s, it has been possible to study selected neurobiological processes that are linked to the etiology and pathogenesis of mood disorders. Several key sets of observations have informed research in this area. First, the heritability of mood disorders has suggested that the neurobiological underpinnings of depression may be related to specific genes. Second, more detailed understanding of the neurobiology of stress response has informed interactive stress-diathesis models of vulnerability. Third, the discovery of the first generation of "somatic" treatments (i.e., ECT and the

TCA and MAOI antidepressants) in the 1940s and 1950s pointed to potentially reversible neurobiologic targets for intervention. As the methodology for studying the neurobiology of mood disorders has grown more sophisticated, investigations utilizing indirect indicators of brain function, such the levels of monoamine metabolites or cortisol urine, plasma, or CSF, have largely been supplanted by translationally guided studies of gene transcripts and proteomics. Likewise, coarse measures of regional brain function, such as recordings of evoked potentials or electroencephalographic (EEG) activity patterns during waking and sleeping, largely have given way to neuroimaging strategies that permit the activity of regions or specific neural circuits to be examined at rest and during provocative challenges. This chapter will survey the neurobiology of mood disorders, emphasizing clinically relevant associations with phenomenology, course of illness, and treatment outcomes.

CLINICAL PHENOMENOLOGY

The signs, symptoms, and subjective experiences associated with depression have long suggested dysfunction of basic central nervous system (CNS) processes. With respect to cortical function, depression involves multiple disturbances of information processing. Most depressed people automatically interpret experiences from a negative perspective, and their access to memory is negatively biased. In more severe depressive states, cognition and problem-solving skills are further compromised by poor concentration and decreased ability to use abstract thought. A virtual monologue of negative thoughts and images seems to run on autopilot, and, unlike normal states of sadness, ventilation to a confidante has little beneficial effect. In more extreme cases, delusions or hallucinations, or both, grossly distort reality testing. These neurocognitive changes point to dysfunction involving the hippocampus, prefrontal cortex (PFC), and the amygdala and other limbic structures.

Another biologically based characteristic of depression involves decreased interest and a loss of mood reactivity: Spontaneous, goal-directed activities decrease, and events that *should* improve mood have little or no effect. One correlate of loss of interest is decreased reinforcer salience. Even basic appetitive functions, such as appetite and libido, are diminished in severe depression. Anhedonia and decreased appetitive behavior point to dysfunction of the neural circuits involved in the anticipation and consummation of rewards, which involve the thalamus, hypothalamus, nucleus accumbens, anterior cingulate, and PFC.

Alterations in the initiation, expression, and spontaneity of movement are common in severe depression. A background state of decreased animation referred to as psychomotor retardation is sometimes accompanied by a superimposed state of agitation. In addition to the pacing and frequent postural shifts, stereotypical signs or behaviors, such

as a furrowed brow, hair pulling, biting at the lips or nail beds, and compulsive scratching, are observed in severe depressive states. Anhedonia and psychomotor disturbances are more pronounced among older depressed patients and in recurrent depressive episodes, with agitation typically limited to the most severe depressive syndromes. These observable signs point to dysfunction of subcortical circuits connecting the thalamus, basal ganglia, and striatum.

Almost all depressed people report low energy or fatigue and some difficulty involving sleep. Although insomnia is much more common (prevalence rates of 2/3rds to 4/5ths in most clinical samples), hypersomnolence is not uncommon, particularly earlier in life. Clinicians have long noted an association of hypersomnia with other "reversed" neurovegetative signs and symptoms, such as increased appetite and weight gain. The classical sleep disturbances of depression, such as difficulty maintaining sleep and early morning awakening, also tend to be more pronounced and severe in older individuals with a history of recurrent illness. Such terminal insomnia is, in turn, linked to diurnal mood variation, in which morning is the worst time of the day. Disturbances of these normally well-entrained circadian rhythms are indicative of dysregulation of nuclei in the thalamus, hypothalamus, and brain stem.

For most of the past century, clinical researchers have viewed the tendency for these signs and symptoms to co-aggregate or cluster together to reflect the neurobiological basis for the classical subtypes of depression. For example, anhedonia, loss of mood reactivity, and psychomotor disturbance are more commonly associated with weight loss, terminal insomnia, or diurnal mood disturbance, or a combination of these. This constellation forms the basis of the syndrome of *melancholia* (also known over the past five decades as *autonomous*, *biological*, or *endogenous* forms of depression). As these depressions are often recurrent and discrete episodes can begin without the provocation of a significant stressor, they were proposed to have an *internal* (i.e., neurobiological) etiology. The fact that such depressions were slow to remit spontaneously yet were quite responsive to ECT further reinforced perceptions of underlying biological dysfunction. The stable nature of this syndrome over the centuries has been recognized by the contemporary use of the ancient Greek term *melancholia*, even though black bile is no longer implicated in the pathophysiology of severe depressive states.

In contrast to *melancholia*, the depressions more commonly seen in young adult life often seemed inextricably tied to personality problems, interpersonal problems, and maladaptive responses to psychosocial difficulties. The terms *nonendogenous* and *reactive* were used to represent the observation that depression in young adults is almost invariably associated with stressful life events. As suggested by the names, these

depressions were presumed to be mediated by nonbiological factors related to personality and resilience to stress. The association of an early age of onset with co-occurring anxiety (both symptoms and discrete disorders)—coupled with a low incidence of psychotic symptoms (i.e., delusions or hallucinations)—further solidified impressions that younger depressed patients had a *neurotic* disorder. Another hallmark of a putatively neurotic disorder, persistence or chronicity, also characterized many early-onset depressions. The presumed primacy of psychosocial factors in neurotic or nonendogenous depression was further suggested by the too frequently poor responses to ECT and TCAs. That the onset of depression in early life may, of course, distort personality development, limit one's repertoire of coping skills and diminish resilience, change brain structure and function, and alter drug responsivity (i.e., pathoplasty) was not yet appreciated. It was also not yet known that early-onset depressions are associated with a relatively high familial loading in genetic studies (i.e., a hallmark of a biologically mediated disorder). Likewise, neither the heritability of neuroticism nor the neural basis of the characteristically increased reactivity to emotional stimuli had yet been established.

One historically important subgroup of patients with nonendogenous or neurotic depression was defined in part by *reverse* neurovegetative symptoms (i.e., increased appetite, weight gain, and hypersomnolence). As the patients treated by the hospital-based psychopathologists of the 1940s and 1950s seldom had such symptoms, this presentation was described as *atypical depression*. The notion of atypicality was reinforced by a relative poor response to ECT and TCAs and, by contrast, relatively preferential response to MAOIs. Importantly, as clinical research shifted to ambulatory settings and greater numbers of younger and less severely symptomatic people began to seek treatment of mood disorders, it was recognized that about 15 to 25 percent of depressed people have reversed vegetative features, which means that—despite the name—this presentation was not so atypical.

RISK FACTORS

Genetic Influences

Mood disorders and suicide clearly run in families, yet the contributions of nature and nurture have provoked intense debates for decades. It is now well established that bipolar affective disorder is more heritable than other mood disorders, heritability increases in proportion to the amount of shared genetic material, and that an early age of onset is associated with greater heritability. Thus, identical twins have greater heritable risk than fraternal twins, who are at approximately the same risk of other first-degree relatives (i.e., siblings and parents). First-degree relatives, in turn, have greater shared risk than half-siblings, grandparents, or cousins.

Research comparing same-sex fraternal and identical twins and studies using the adopted-away paradigm further demonstrated that heritable risk transcends environmental influences. Research using modern molecular genetic methods will eventually lead to identification of the gene products that transmit this risk. Already it is clear that a number of alleles are associated with a small but replicable increase in the risk of depression; yet others may be associated with relative resilience. Among the numerous genes associated with an increased risk of depression are the alleles that code for the promoter region of the serotonin transporter (5-HTT), which has been found by a number of researchers to moderate the association between stressful life events and the risk of depression (Fig. 13.9–1). Other genes, while showing no relation to risk of depression, are relevant to response to specific antidepressants. Whether batteries of these genes can improve the process of matching individual patients remains to be seen.

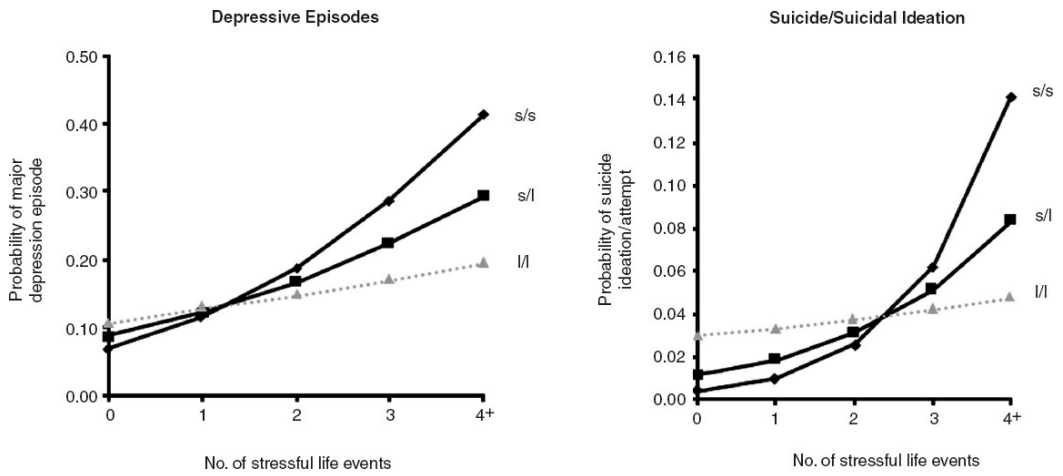

FIGURE 13.9–1. Interaction of serotonin transporter (5-HTT) gene polymorphism and life stress in depression risk. Results of multiple regression analysis estimating the association between number of stressful life events (between ages 21 and 26 years) and depression outcomes at age 26 as a function of 5-HTT genotype. s/s, short/short allele; s/l, short/long allele; l/l, long/long allele. (Data from Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW. Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science*. 2003;301(5631):386.)

Although the significance of genetics cannot be disputed, considerable variability exists within and across groups with comparable heritable risks. For example, identical twins do not have 100 percent concordance of bipolar affective disorder. Early-onset depressions are more heritable, and, conversely, late-onset mood disorders (i.e., after 60 years of age) have the lowest rates of heritability, presumably because of the greater impact of acquired risk factors such as cerebrovascular disease. In some pedigrees, men are at greater risk of alcoholism and sociopathy, and women are at greater risk of depression. There is some evidence that different environmental experiences may influence whether a woman develops depression or a generalized anxiety disorder. Conversely, for men, alcohol

or drug use disorders may constitute an alternate illness trajectory to early onset depressive disorders. Genetics thus constitute just one pathway of vulnerability.

Predispositions to various emotional states (i.e., mood set points or temperament) also appear to be partly heritable. Of particular importance is a temperament characterized by excessive emotional reactivity and behavioral inhibition. Modern theories of temperament and personality converge around several related dimensions, including high levels of harm avoidance (neuroticism) and low levels of novelty seeking (introversion).

These temperaments, which can be recognized in infancy, are associated with increased vulnerability to depression and anxiety, presumably via increased limbic reactivity and heightened sensitivity to the impact of stressful life events. Childhood onset of dysthymia similarly presages extremely high rates of depression and bipolar disorder in adulthood.

One reason for inconsistency in genetic research is the process known as epigenetics. This term refers to functionally relevant modification of genes—typically induced by DNA methylation or histone modification—in response to environmental stimulation that do not actually change the DNA sequence of the gene. Epigenetic changes in gene expression, both activating and inhibiting, are typically irreversible (i.e., will last for the life of the cell) and may persist across multiple generations. Intensive ongoing research indicates that some of the changes in brain function attributable to the impact of chronic stress may be mediated by epigenetic changes in key regions of the brain.

ETIOLOGY OF EMOTIONS

Some consider depression to be an extreme expression of sadness, a normal mood, with elation to be the healthier counterpart of the euphoria of mania. Such continuity between normal and pathological mood states is illustrated by bereavement. On the one hand, grief is a universal experience, yet, on the other hand, it can segue into a severe and disabling depressive state. Although there are less obvious parallels to mania, the intensity of new romantic love or other peak emotional experiences is associated with changes in perception, cognition, behavior, and judgment that border on manic excitement. The euphorigenic effects of cocaine and other psychostimulants amply document that the hard wiring for manic states exists in many if exposed to the right neurochemical milieu.

There is now considerable agreement that several basic emotions are observed across human cultures. Sadness and crying, for example, may be a universally recognized interpersonal distress signal. Moreover, the capacity to recognize the emotional states of others develops at such an early age that innate processes are undoubtedly implicated.

At the risk of anthropomorphizing excessively, basic emotional states are also observable across mammalian species. Behavioral predispositions

toward aggressivity (or passivity) and social dominance (or subordination) appear to extend across vertebrate species. However, by virtue of a large neocortex, humans are distinguished by a greater capacity to modify expression of basic emotional states via abstraction, integration, and synthesis of complex symbolic representations, communication of experiences to others in direct or elaborated forms, and development of self-concepts that guide behavior, both in relation to others and an anticipated future. The domains of competence are similarly broader and diverse. Whereas a highly competent primate may quickly gain a position of dominance if placed in a new troupe, only the human can intentionally misrepresent his or her competence, conceal motivations, or be thwarted by archival documentation of past behaviors.

When compared to other mammals, the relatively greater importance and intensity of attachment bonds facilitate the protracted task of human child rearing and enhance the advantages of kinship. Indeed, for hundreds of thousands of years, the ancestors of man lived in a world in which the survival into adulthood was uncommon. Indeed, it is only in the past few centuries that average life expectancy has exceeded 40 years. Young humans are arguably the most vulnerable of mammals and certainly remain dependent on their caregivers for the longest period of time. Certain necessary divisions of labor no doubt increased the chances of viable offspring and may have slowly shaped gender differences in affectivity and affiliative behavior that underpin gender differences in the risk of depressive disorders. Ethological considerations also may help to explain why rates of depression appear to be increasing. The profound changes in social systems and lifestyles that have taken place in the past several centuries have forced adaptations that are too rapid for natural selection. In relative terms, the breathtaking sociocultural and technological changes of the past century have taken place in less than 0.00001th of hominid experience! The stressors faced in modern life thus present complexities that, despite such a large neocortex and complex social systems, sometimes overwhelm basic adaptive capacities.

EMOTIONAL PROCESSING AND THE BRAIN

Modern affective neuroscience focuses on the importance of four brain regions in the regulation of normal emotions: the PFC, the anterior cingulate, the hippocampus, and the amygdala (Fig. 13.9–2). The PFC is viewed as the structure that holds representations of goals and appropriate responses to obtain these goals. Such activities are particularly important when multiple, conflicting behavioral responses are possible or when it is necessary to override affective arousal. The dorsolateral region of the PFC is particularly active in suppression of negative affects and is a key constituent in the cognitive control network (CNN). The medial PFC, along with the posterior cingulate and hippocampus (see below), are important

hubs in the default mode network (DMN), an interconnected circuit of regions that are active during resting wakefulness, self-referential thinking (i.e., autobiographical memory), or daydreaming. The DMN appears to show hyperconnectivity in depressive disorders, which predisposes to negative recall and rumination. There also is evidence of some hemispherical specialization in PFC function. For example, left-sided activation of regions of the PFC is more involved in goal-directed or appetitive behaviors, whereas regions of the right PFC are implicated in avoidance behaviors and inhibition of appetitive pursuits.

The anterior cingulate cortex (ACC) is thought to serve as the point of integration of attentional and emotional inputs. Two subdivisions have been identified: an affective subdivision in the rostral and ventral regions of the ACC and a cognitive subdivision involving the dorsal ACC. The former subdivision shares extensive connections with other limbic regions, and the latter interacts more with the PFC and other cortical regions. It is proposed that activation of the ACC as part of a CNN facilitates effortful control of emotional arousal, particularly when goal attainment has been thwarted or when novel problems have been encountered. There is evidence that connectivity among the components of the CCN is diminished in depressive states and, if persistent, such decreased connectivity may contribute to relapse/recurrence risk.

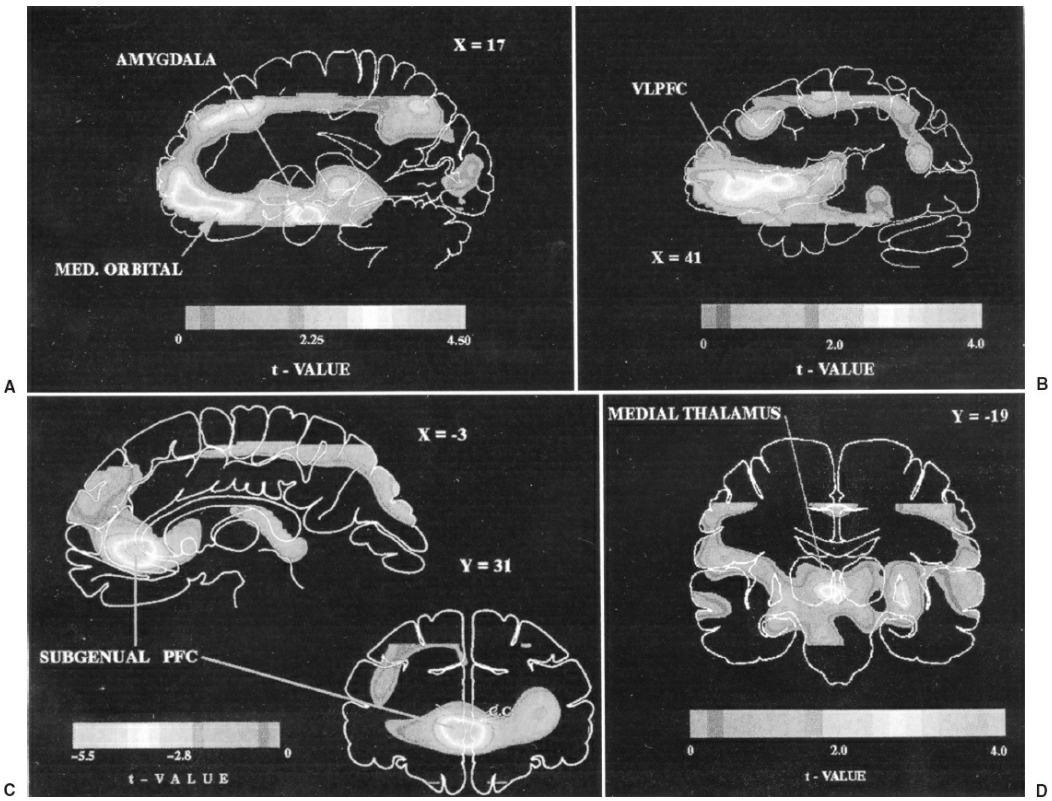

FIGURE 13.9–2. Composite coronal and sagittal sections of positron emission tomography scans show areas in which cerebral glucose metabolism is decreased in depressed patients relative to matched healthy controls. PFC, prefrontal cortex; VLPFC, ventrolateral prefrontal cortex. (See color atlas.) (From Wayne Drevets, M.D., and Annu Rev

The hippocampus is most clearly involved in various forms of learning and memory, including fear conditioning, as well as inhibitory regulation of the HPA-axis activity. Importantly, the hippocampus is also one of the few sites of the brain capable of neurogenesis (i.e., the process of differentiation of stem cells into potentially functional neurons). Emotional or contextual learning appears to involve a direct connection between the hippocampus and the amygdala.

The amygdala appears to be a crucial way station for processing novel stimuli of emotional significance and coordinating or organizing cortical responses. Located just above the hippocampi bilaterally, the amygdala has long been viewed as the functional center of the limbic system. Although most research has focused on the role of the amygdala in responding to fearful or painful stimuli, novelty and ambiguity can also bring the amygdala "on line."

STRESS AND ANIMAL MODELS OF DEPRESSION

Studies of rodents, dogs, cats, and nonhuman primates have demonstrated that acute stress responses involve activation of central and peripheral components of two interactive psychoneuroendocrine systems: the cortical HPA axis and the sympathomedullary (SM) axis. Acute responses to stress are phenomenologically and neurobiologically more akin to fear and anxiety than depression. However, whereas stress initially signals threat, it is the anticipation of loss and the meaning of the loss—as well as the loss of hope—that can provoke the transition from anxiety to sadness and despair.

Acute stress activates cell bodies of norepinephrine (NE) neurons in the locus ceruleus, whose ascending axons signal noradrenergically mediated increases in cortical arousal and whose descending axons signal increased adrenergic (predominantly epinephrine) output by the sympathetic nervous system and adrenal medulla. Behaviorally, the perception of acute stress elicits perceptual vigilance, preparedness for response, and inhibition of appetitive activities, such as foraging and pursuit of sexual activity.

Stress simultaneously triggers and releases corticotropin-releasing hormone (CRH) from neurons in the hypothalamus and cerebral cortex. Hypothalamic CRH activates synthesis and release of adrenocorticotrophic hormone (ACTH) from the anterior pituitary, which, in turn, triggers release of cortisol and other glucocorticoids (from the adrenal cortex). CRH also synergistically increases locus ceruleus activity and, directly or indirectly, increases synthesis of other stress-reactive gene products and anti-inflammatory responses and decreases synthesis of key neuropeptides such as BDNF. The excitatory amino acid glutamate may synergistically heighten CNS arousal in response to stress. The acute response to stress is

counterbalanced by homeostatic or adaptive mechanisms. These include feedback inhibition by glucocorticoid receptors in the hippocampus and pituitary, downregulation of postsynaptic noradrenergic receptors, and inhibitory auto- and heteroreceptors on presynaptic NE neurons. As is discussed subsequently, parallel input from serotoninergic, GABA, and glycinergic neurons exerts dampening or inhibitory effects.

Exposure to prolonged, inescapable stress is associated with numerous adaptations in neurobehavioral responses. Although CRH levels in the brain and corticosteroid levels in the periphery may remain elevated, levels of NE, 5-HT, dopamine (DA), and GABA in the brain stem and forebrain eventually decrease. Animals trapped in such a state eventually cease trying to resist or to escape and show decreased grooming and appetitive behavior. The behavior persists outside of the experimental situation. There are significant individual differences in the development of such learned helplessness, as well as differences across pedigrees, strains, and species. Like other forms of chronic stress, the negative behavioral effects of learned helplessness may be sustained by epigenetic changes in key regions of the brain. Importantly, antidepressant drugs have been shown to prevent, attenuate, or reverse learned helplessness across species.

Studies of the experiences of primates in the wild, of course, are more relevant to the stresses faced by people than a rat's response to inescapable, painful, electric shock. These naturalistic studies demonstrate not only that central 5-HT activity is partly under genetic control but also that loss of a dominant role within a primate social hierarchy results in decreased 5-HT neurotransmission and increased HPA. In the wild, animals with low levels of the serotoninergic metabolite 5-hydroxyindoleacetic acid (5-HIAA) are likely to be more aggressive, more socially subordinate, and less sexually active. When social status is manipulated by creating a new troupe from a cohort of subordinate primates, a new dominance hierarchy emerges, and the more dominant animals experience a corresponding increase in 5-HT function. Conversely, during times of adversity, such as drought or famine, it is the socially dominant primates that experience the largest increases in HPA activity.

MALTREATMENT AND EARLY ADVERSITY

Physical, verbal, and sexual abuse and parental neglect have an indelible effect on one's life trajectory. Although it has been known for decades that such maltreatment plays a key role in the pathogenesis of PTSD and borderline personality disorder, it is now clear that such a history results in a two- to threefold increase in the risk of depression and is associated with important phenomenological and biological differences in depressive states.

Studies of animal models confirm that lasting alterations in neuroendocrine and behavioral responses can result from severe early

stress. Animal studies further indicate that even transient periods of maternal deprivation can alter subsequent responses to stress. This vulnerability has been shown to be the result of enduring changes in gene expression. For example, protracted stress thus can induce changes in the functional status of neurons via inducing epigenetic changes and may suppress neurogenesis in the hippocampus. Recent studies in depressed humans indicate that a history of early trauma is associated with long-lasting changes in responsivity of the HPA axis, including accompanied by structural changes in the hippocampus and perhaps other focal regions within the cerebral cortex.

MONOAMINE SYSTEMS: FUNCTION AND DYSFUNCTION

An area of scientific inquiry is chosen, in part, because of the knowledge base and experimental paradigms available to investigators. In the early 1960s, Nobel prize-winning research had just elucidated the importance of monoamines in neurotransmission. Moreover, it was possible to measure the metabolites of the catecholamine NE in various body fluids and 5-HIAA, the principal metabolite of 5-HT, in CSF. Visualization of the functional brain, by contrast, was essentially limited by the low sensitivity of recording neuronal activity from the surface electrodes used for EEGs. Specifically, whereas the paroxysmal spikes in electrical activity caused by epileptogenic foci or the diffuse slowing associated with delirium could be documented by EEGs, the more subtle changes in brain activity or regional cerebral metabolism associated with affective states could not.

In addition, there were multiple lines of evidence from pharmacological studies implicating perturbations of monoamine systems in the therapeutic and iatrogenic effects of various medications. The relevant pharmacotherapies of the time, the TCAs and MAOIs, were known to increase the amount of neurotransmitter available at monoamine synapses. Conversely, the monoamine-depleting compound reserpine was known to cause depression as a side effect during the treatment of hypertension. Intoxication with potent noradrenergic agonists, such as amphetamine and cocaine, also was known to induce states of sleepless euphoria and activated psychosis, followed by a crashing, depression-like withdrawal.

The early monoamine hypotheses formulated by Joseph J. Schildkraut, Alexander H. Glassman, Arthur J. Prange, Jr., John Davis, Herman van Praag, Jonathan Cole, and others have undergone much revision, and it is now known that these ancient neuromodulatory systems comprise only a small fraction of the brain's neurons. Nevertheless, NE and 5-HT neurotransmission is still thought to play an important role in the pathophysiology of mood disorders and most antidepressant medications initiate their actions through these neurons. A third monoamine, the catecholamine DA, was thought to play a greater role in the pathophysiology and treatment of schizophrenia and thus received less

emphasis. Nevertheless, evidence of DA dysfunction in at least some forms of depression has emerged. Knowledge about monoamine receptor families, second messengers, neurokinins, excitatory amino acids, and gene transcription factors has added layers of complexity to the original monoamine hypotheses. As a result, there has been a progressive shift away from focusing on disturbances of single neurotransmitter systems in favor of studying neurobehavioral systems, neural circuits, and more intricate neuroregulatory mechanisms. The monoaminergic systems are thus now viewed primarily as broader, neuromodulatory systems, and disturbances are as likely to be secondary or epiphenomenal effects as they are directly or causally related to etiology and pathogenesis.

NORADRENERGIC SYSTEMS

The cell bodies of almost all noradrenergic neurons in the brain are located in the locus ceruleus of the brain stem and project rostrally to the hypothalamus, basal ganglia, limbic system, and cerebral cortex. This diffuse distribution belies NE's role in initiating and maintaining limbic and cortical arousal, as well as modulation of other neuronal systems. Noradrenergic projections to the amygdala and hippocampus are implicated in emotional memory and behavioral sensitization to stress, and prolonged activation of the locus ceruleus contributes to the state of learned helplessness. The medial forebrain bundle (MFB) is the key ascending NE pathway to anterior cortical structures. Stimulation of the MFB elicits increased levels of goal-directed and reward-seeking behavior. Sustained stress eventually results in decreased MFB neurotransmission, which may account for anergia, anhedonia, and diminished libido in depression. The locus ceruleus also is the origin of neurons that project rostrally to interact with the cell bodies of the sympathetic nervous system and, hence, the adrenal medulla. The locus ceruleus is phasically active. Increased activity is provoked by the perception of novel stimuli; decreased activity accompanies eating and sleeping. As noted previously, there is an additive interaction (i.e., positive feedback loop) between noradrenergic neurons in the locus ceruleus and CRH from the hypothalamus, as well as cortical regions. Other negative feedback loops must exert inhibitory control over this alerting synergism to prevent sustained pathological activation. For example, increased noradrenergic output stimulates α_2 heteroreceptors on co-localized inhibitory 5-HT neurons, and a sustained increase in CRH drives results in decreased hypothalamic release of ACTH, which blunts HPA activity. Changes in cognitive processing of sensory input can further dampen corticolimbic and sympathoadrenal responses to sustained threatening or painful stimuli.

SEROTONINERGIC SYSTEMS

Serotoninergic neurons project from the brain stem dorsal raphe nuclei to the cerebral cortex, hypothalamus, thalamus, basal ganglia, septum, and hippocampus. The 5-HT pathways have inhibitory and facilitatory functions in the brain. For example, much evidence suggests that 5-HT is an important regulator of sleep, appetite, body temperature, metabolism, and libido. Moreover, 5-HT inhibits aggressive behavior across mammalian and reptilian species. Serotoninergic neurons projecting to the suprachiasmatic nucleus (SCN) of the hypothalamus help regulate circadian rhythms (e.g., sleep–wake cycles, body temperature, and HPA-axis function). The 5-HT also permits or facilitates goal-directed motor and appetitive behaviors in conjunction with NE and DA.

As noted previously, there is some evidence that 5-HT neurotransmission is partly under genetic control, and a functional polymorphism of the 5-HTT has been shown to moderate risk of depression within the context of life stress. The 5-HT "tone" also shows a seasonal fluctuation and shows a characteristic temporal response to stress. Specifically, acute stress increases 5-HT release transiently, whereas chronic stress eventually causes decreased 5-HT activity and depletion of 5-HT stores. Chronic stress may also increase synthesis of 5-HT type 1A (5-HT_{1A}) autoreceptors in the dorsal raphe nucleus, which further decrease 5-HT transmission. Elevated glucocorticoid levels tend to enhance 5-HT functioning and thus may have significant compensatory effects in chronic stress.

DOPAMINERGIC SYSTEMS

There are four relatively discrete DA pathways in the brain. The tuberoinfundibular system projects from cell bodies in the hypothalamus to the pituitary stalk, exerting inhibitory control over prolactin secretion. The nigrostriatal system originates from cell bodies in the substantia nigra and projects to the basal ganglia, regulating involuntary motor activity. The cell bodies of the mesolimbic pathway are located in the ventral tegmentum and project diffusely to the nucleus accumbens, amygdala, hippocampus, medial dorsal nucleus of the thalamus, and cingulate gyrus. This pathway modulates emotional expression, learning and reinforcement, and hedonic capacity. The fourth DA pathway, also originating in the ventral tegmentum, is the mesocortical pathway, which projects to the orbitofrontal and the prefrontal cortical regions. The mesocortical pathway helps regulate motivation, concentration, and initiation of goal-directed and complex executive cognitive tasks. Decreased mesocortical and mesolimbic DA activity has obvious implications in the cognitive, motor, and hedonic disturbances that are associated with depression and is a target for action for some investigational compounds. Across pathways, increased DA activity is potentiated by nicotinic inputs and glucocorticoids, and DA levels are inversely correlated with indices of brain 5-HT activity.

SUMMARY OF RESEARCH ON MONOAMINES IN DEPRESSION

After nearly 30 years of research, it can be concluded that subsets of depressed people manifest one or more abnormalities of monoamine neurotransmission. Decreased central NE activity can be inferred, in part, from decreased urinary excretion of the metabolite 3-methoxy-4-hydroxyphenylglycol (MHPG). However, a partly overlapping subgroup of patients manifest elevated circulating levels of NE and its metabolites. This suggests a dissociation of NE activity in the MFB and SM systems. Functionally increased noradrenergic activity is also reflected by blunted α_1, β, and β-coupled second messenger (i.e., cyclic adenosine monophosphate [cAMP]) responses.

Chronic treatment with NE reuptake inhibitors (NRIs) (such as the TCA desipramine or the more selective drug reboxetine) causes a time-dependent decrease in firing of NE neurons and downregulation of postsynaptic β receptors. The antidepressant effects of NRIs are reversed by transiently decreasing the availability of NE in the brain by administration of the NE synthesis inhibitor α-methyl-p-tyrosine (AMPT).

5-HT dysfunction has been documented in overlapping subgroups of patients using a variety of methods, ranging from low CSF levels of 5-HIAA to decreased regional cerebral metabolism in response to a 5-HT agonist, such as D-fenfluramine. 5-HT dysfunction also is reflected by blunted responses to specific (e.g., the specific 5-HT_{1A} compound ipsapirone) and nonselective (e.g., L-tryptophan or D-fenfluramine) agonists, decreased neuroendocrine responses, and decreased 5-HT uptake sites on blood platelets. As discussed subsequently, decreased 5-HT neurotransmission also can be inferred from the findings of HPA and EEG sleep studies. Dietary depletion of L-tryptophan induces 5-HT mediated abnormalities in a subset of vulnerable individuals. Likewise, L-tryptophan depletion reverses responses to MAOIs and SSRIs but does not diminish the effect of interventions that target other systems (e.g., the antidepressant bupropion, cognitive therapy, or transcranial magnetic stimulation).

Low levels of the DA metabolite homovanillic acid have been implicated in psychomotor retardation. Some evidence suggests that increased DA activity, perhaps mediated by elevated glucocorticoid levels or lower levels of the enzyme DA β-hydroxylate, may contribute to the development of delusions and hallucinations.

OTHER NEUROTRANSMITTER DISTURBANCES

Acetylcholine (ACh) is found in neurons that are distributed diffusely throughout the cerebral cortex. Cholinergic neurons have reciprocal or interactive relationships with all three monoamine systems. Abnormal levels of choline, which is a precursor to ACh, have been found at autopsy in the brains of some depressed patients, perhaps reflecting abnormalities

in cell phospholipid composition. In an animal model of depression, strains of mice that are super- or subsensitive to cholinergic agonists have been found to be prone to or more resistant to developing learned helplessness. Cholinergic agonists can induce changes in HPA activity and sleep that mimic those associated with severe depression. Indeed, some remitted patients with mood disorders, as well as their never-ill first-degree relatives, have a trait-like increase in sensitivity to cholinergic agonists. Cholinergic agonist and antagonist drugs have differential clinical effects on depression and mania. Agonists can produce lethargy, anergia, and psychomotor retardation in normal subjects, can exacerbate symptoms in depression, and can reduce symptoms in mania. These effects generally are not sufficiently robust to have clinical applications, although some recent research has suggested that parenteral scopolamine can have rapid antidepressant effects. More conventional antidepressants that target serotonergic or adrenergic neurons also indirectly decrease cholinergic tone.

GABA has an inhibitory effect on ascending monoamine pathways, particularly the mesocortical and mesolimbic systems. Reductions in GABA have been observed in plasma, CSF, and brain GABA levels in depression. Animal studies have also found that chronic stress can reduce and eventually deplete GABA levels. By contrast, GABA receptors are upregulated by antidepressants, and some GABAergic medications have weak antidepressant effects.

The amino acids, glutamate and glycine, are the major excitatory and inhibitory neurotransmitters in the CNS. Glutamate and glycine bind to sites associated with the *N*-methyl-D-aspartate (NMDA) receptor, and an excess of glutamatergic stimulation can cause neurotoxic effects. Importantly, there is a high concentration of NMDA receptors in the hippocampus. Glutamate may thus work in conjunction with hypercortisolemia to mediate the deleterious neurocognitive effects of severe recurrent depression and may be proximally implicated in atrophic changes documented in some regions of the cerebral cortex, perhaps by first reducing the connectivity of neurons and volume of supporting glia. The potential significance NMDA receptors in the pathophysiology of depressive states is likewise suggested by growing evidence that a subanesthetic dose of ketamine (Ketalar), which noncompetitively antagonizes NMDA receptors, can exert large and rapid antidepressant effects that persist for 3 to 7 days. However, the effects require that ketamine is administered parenterally (e.g., intravenous [IV] infusion) and, because of well-characterized psychotomimetic and dissociative effects at somewhat higher doses, ketamine is considered a controlled substance (Schedule III) by the Drug Enforcement Administration. Development of drugs that produce antidepressant effects similar to ketamine that are easier to administer and/or have lower abuse potential is an area of intense

research activity.

SECOND MESSENGERS AND INTRACELLULAR CASCADES

The binding of a neurotransmitter and a postsynaptic receptor triggers a cascade of membrane-bound and intracellular processes mediated by second messenger systems. Receptors on cell membranes interact with the intracellular environment via guanine nucleotide-binding proteins (G proteins). The G proteins, in turn, connect to various intracellular enzymes (e.g., adenylate cyclase, phospholipase C, and phosphodiesterase) that regulate utilization of energy and formation of second messengers, such as cyclic nucleotide (e.g., cAMP and cyclic guanosine monophosphate [cGMP]), as well as phosphatidylinositols (e.g., inositol triphosphate and diacylglycerol) and calcium–calmodulin.

Second messengers regulate the function of neuronal membrane ion channels, neurotransmitter synthesis and release, and protein kinase activity. Protein kinase, for example, catalyzes phosphorylation, an energy-liberating process involved in the synthesis and degradation of receptors, ion channels, G proteins, and gene transcription and translation factors. Some studies have reported abnormalities in platelet adenylate cyclase activity, phosphoinositide hydrolysis, intracellular calcium metabolism, and G-protein function in depressive disorders. Moreover, the synaptic effects of antidepressants initiate a series of intracellular reactions that ultimately alter gene activity, resulting in downregulation of selected monoamine receptors and increased synthesis of BDNF. There is increasing evidence that mood stabilizing drugs similarly act on G proteins or other second messengers.

ALTERATIONS OF HORMONAL REGULATION

HPA Activity

As noted previously, elevated HPA activity is a hallmark of mammalian stress responses and one of the clearest links between depression and the biology of chronic stress. Hypercortisolemia in depression suggests one or more of the following central disturbances: decreased inhibitory 5-HT tone; increased drive from NE, ACh, or CRH; or decreased feedback inhibition from the hippocampus.

Evidence of increased HPA activity is apparent in 20 to 40 percent of depressed outpatients and 40 to 60 percent of depressed inpatients. Older patients, particularly those with highly recurrent or psychotic depressive disorders, are the most likely to manifest increased HPA activity. Although hypercortisolism is one of the best-replicated biological correlates of melancholia or endogenous depression, it is hardly a specific abnormality. For example, a short period of starvation or several weeks of partial sleep deprivation can induce hypercortisolism in otherwise healthy people.

Elevated HPA activity in depression has been documented via excretion of urinary free cortisol (UFC), 24-hour (or shorter time segments) IV collections of plasma cortisol levels, salivary cortisol levels, and tests of the integrity of feedback inhibition. A disturbance of feedback inhibition is tested by administration of dexamethasone (Decadron) (.5 to 2.0 mg), a potent synthetic glucocorticoid, which normally suppresses HPA-axis activity for 24 hours. Nonsuppression of cortisol secretion at 8:00 AM the following morning or subsequent escape from suppression at 4:00 PM or 11:00 PM is indicative of impaired feedback inhibition. Hypersecretion of cortisol and dexamethasone nonsuppression are imperfectly correlated (approximately 60 percent concordance), suggesting that they may measure overlapping but different aspects of HPA dysfunction. A more recent development to improve the sensitivity of the test utilizes an infusion of CRH to stimulate cortisol secretion after dexamethasone suppression.

Neither the sensitivity nor the specificity of these tests of feedback inhibition is sufficient for use as a diagnostic test, and adrenocortical hyperactivity (albeit usually less prevalent) is observed in mania, schizophrenia, dementia, and other psychiatric disorders. Nonsuppression may implicate a loss of inhibitory hippocampal glucocorticoid receptors more so than increased CRH drive, which also may account for the strong age dependence of cortisol nonsuppression. As noted previously, dexamethasone nonsuppression in adulthood is also associated with a history of early trauma, which may result from an enduring reduction in synthesis of glucocorticoid receptors, hippocampal cell death, and/or chronic hypersecretion of CRH.

Elevated HPA activity in depression is typically not associated with the physical stigmata of Cushing syndrome but is sufficient to be implicated in the genesis of neurocognitive and neuroimmunological disturbances. Patients with increased HPA activity are typically less responsive to attention placebo interventions and psychotherapy. Hypercortisolemic depressed patients thus may have a relatively greater need for active pharmacotherapy or ECT. Hypercortisolism does, however, usually resolve with effective treatment. When persistent despite effective treatment, increased HPA activity is associated with a high risk of relapse. This is presumably a consequence of incomplete resolution of the pathophysiology of the depressive episode. One implication of these relationships is that interventions that suppress HPA activity, including dexamethasone and the cortisol synthesis inhibitor ketoconazole (Nizoral), are sometimes used to treat patients with more refractory depressive disorders. Attempts to develop compounds that directly antagonize central CRH receptors as antidepressants are ongoing.

Thyroid Axis Activity

Approximately 5 to 10 percent of people evaluated for depression have previously undetected hypothyroidism, as reflected by low levels of circulating thyroid hormone, an elevated basal TSH level, or an increased TSH response to a 500-mg infusion of the hypothalamic neuropeptide thyroid-releasing hormone (TRH). Such abnormalities are often associated with elevated antithyroid antibody levels and, unless corrected with hormone replacement therapy, may compromise response to antidepressant medication.

An even larger subgroup of depressed patients (20 to 30 percent) shows a blunted TSH response to TRH challenge. This type of response would usually suggest hyperthyroidism, yet few depressed patients have clinically significant elevations of thyroid hormones. A blunted TSH response in a euthyroid person thus may result from pituitary downregulation consequent to increased TRH "drive." As neurons containing TRH have been identified in a variety of cortical regions, this abnormality may have a suprahypothalamic origin. Increased central TRH secretion, in turn, could result from a homeostatic response to decreased noradrenergic neurotransmission. The therapeutic benefits of adjunctive therapy with 1-triiodothyronine (T_3) or other thyroid hormones may be mediated by a dampening of this failed homeostatic response. This abnormality may be most common in individuals who have reduced ability to convert thyroxine to T_3. The major therapeutic implication of a blunted TSH response is evidence of an increased risk of relapse despite preventive antidepressant therapy. Of note, unlike the dexamethasone suppression test (DST), blunted TSH response to TRH often does not normalize with effective treatment.

GROWTH HORMONE

Growth hormone (GH) is secreted from the anterior pituitary after stimulation by NE and DA. Secretion is inhibited by somatostatin, a hypothalamic neuropeptide, and CRH. Pulsatile GH secretion follows a 24-hour circadian rhythm, with a characteristic secretory surge during the first few hours of sleep. The most consistent finding in depression is a blunted GH response to clonidine, an α_2-receptor agonist. Secretory responses after the onset of sleep or the administration of nonselective adrenergic agonists, such as desipramine, are also blunted in depression.

Although the hypothalamus has the highest concentrations of somatostatin, significant concentrations are also found in the amygdala, hippocampus, nucleus accumbens, PFC, and locus ceruleus. In addition to inhibition of GH release, somatostatin inhibits release or antagonizes the effects of CRH, GABA, and TSH. Decreased CSF somatostatin levels have been reported in depression, and increased levels have been observed in mania.

Prolactin

Prolactin is released from the pituitary by 5-HT stimulation and inhibited by DA. Most studies have not found significant abnormalities of basal or circadian prolactin secretion in depression, although a blunted prolactin response to various 5-HT agonists has been described. This response is uncommon among premenopausal women, suggesting that estrogen has a moderating effect.

ALTERATIONS OF SLEEP NEUROPHYSIOLOGY

Sleep is regulated by complex, temporally, and environmentally cued interactions between the monoamine neuromodulatory systems, neuroendocrine secretion, and neuropeptides. At the broadest levels, there is a seasonal tendency for longer and deeper sleep in winter and a circadian propensity for the onset of sleep (i.e., after dark and after midday). Onset of sleep is promoted by a surge in secretion of the pineal hormone melatonin (after the onset of darkness). Sleep onset is followed by the nocturnal surge in GH secretion, which occurs within the first 90 minutes of sleep onset. A reduction in core body temperature and diurnally low levels of cortisol secretion further promote the maintenance of sleep. Orchestrated across each 24-hour period is an oscillating, 90-minute, infradian cycle defined by rapid eye movement (REM) sleep. Within each cycle, there is a characteristic progression from light to deeper levels of sleep (defined by different types of EEG activity), culminating in the paradoxical central activation of REM sleep, during which most of the night's dreaming occurs. An 8-hour night of sleep thus typically includes four or five cycles consisting of nonrapid eye movement (NREM) sleep and REM sleep. The propensity for deep sleep is greatest within the first 3 hours after sleep onset. REM sleep periods, by contrast, tend to become longer and more intense as the night of sleep progresses.

Prefrontal cortical metabolism is normally decreased during NREM sleep, a time of physical and metabolic rest. The frontal cortex is essentially off line during deep sleep, and the characteristic rhythm of brain activity consists of slow (delta) desynchronized waves of thalamocortical origin. REM sleep, by contrast, is characterized by fast, low-amplitude electrical activity and increased glucose metabolism in the limbic system. REMs are under the direct control of cholinergic neurons in the pons, which are tonically inhibited by the reticular activating system (predominately by histaminergic and noradrenergic neurotransmission) during wakefulness. During sleep, inhibitory 5-HT projections from the dorsal raphe nuclei phasically suppress REM. Pharmacological manipulations that increase central cholinergic activity lighten sleep and increase phasic REM activity. Dietary depletion of 5-HT and exogenous administration of glucocorticoids similarly can increase phasic REM indices. Injections of CRH and ingestion

of potent noradrenergic agonists decrease total sleep time, reduce slow wave sleep, and suppress REM sleep.

Depression is associated with a premature loss of deep (slow wave) sleep and an increase in nocturnal arousal. The latter is reflected by four types of disturbance: (1) an increase in nocturnal awakenings, (2) a reduction in total sleep time, (3) increased phasic REM sleep, and (4) increased core body temperature. The combination of increased REM "drive" and decreased slow wave sleep results in a significant reduction in the first period of NREM sleep, a phenomenon referred to as reduced REM latency. Results of family and twin studies suggest that these related abnormalities are partly heritable. Consistent with the expected behavior of a heritable trait, reduced REM latency and deficits of slow wave sleep typically persist after recovery of a depressive episode. Blunted secretion of GH after sleep onset is associated with decreased slow wave sleep and shows similar state-independent or trait-like behavior. Difficulties maintaining sleep and increased phasic REM sleep are associated with hypercortisolism and increased limbic blood flow and glucose metabolism; although some forms of pharmacotherapy can distort sleep profiles, recovery from depression usually is accompanied by at least a partial reversal of these abnormalities. The combination of reduced REM latency, increased REM density, and decreased sleep maintenance identifies approximately 40 percent of depressed outpatients and 80 percent of depressed inpatients. False-negative studies are commonly seen in younger, hypersomnolent patients, who may actually experience an increase in slow wave sleep during episodes of depression. Approximately 10 percent of otherwise healthy individuals have abnormal sleep profiles, and, like dexamethasone nonsuppression, profiles associated with depression are not uncommonly seen in other psychiatric disorders.

Despite clear limitations as a diagnostic test, EEG sleep recordings continue to be an important research tool. For example, patients manifesting a characteristically abnormal sleep profile have been found to be less responsive to psychotherapy and to have a greater risk of relapse or recurrence and may benefit preferentially from pharmacotherapy. Most antidepressants suppress REM sleep, probably by directly or indirectly activating postsynaptic 5-HT_{1A} receptors. The efficacy of antidepressants that do not suppress REM sleep, such as bupropion, nefazodone (Serzone), and agomelatine (Valdoxan), in patients with pathologically increased REM sleep warrants further study. The effects of antidepressants on sleep maintenance and slow wave sleep are more variable and are largely mediated by antihistaminic effects and antagonism of 5-HT_2 and, possibly, melatonin receptors.

STRUCTURAL AND FUNCTIONAL BRAIN IMAGING

Modern magnetic resonance imaging (MRI) scans permit visualization of the living brain, with high resolution of cortical structures such as the hippocampus and anterior cingulate, as well as white matter lesions. One abnormality commonly observed in the depressive disorders is increased frequency of hyperintensities in subcortical regions such as periventricular regions, the basal ganglia, and the thalamus. More common in bipolar I disorder and among the elderly, these hyperintensities appear to reflect atherosclerotic change and, as such, illustrate the deleterious effects of the interaction of aging and recurrent affective episodes. Ventricular enlargement, cortical atrophy, and sulcal widening have been observed in some studies of patients with more severe recurrent affective illness. More recent studies utilizing higher resolution MRI scans have documented reduced volume in a number of relevant regions, including the hippocampus, medial orbital cortex, and anterior cingulate. Diffuse and focal areas of atrophy have been associated with increased illness severity and increased cortisol levels. In several studies, the hippocampal volume reduction has been associated with measures of illness chronicity, such as lifetime days of untreated depression. Volume reduction is thought to partly result from reduction in the glial matter that supports neurons; suppression of BDNF-mediated connectivity and, ultimately, neurogenesis also may be implicated. Nevertheless, reduced hippocampal volume has been documented in a study of individuals suffering from a first depressive episode and has been shown to be partly under heritable control in primates.

Positron emission tomography (PET) and functional MRI (fMRI) scanning are now widely used for research aimed at visualizing brain activity during rest and various states of activation. Normal sadness is associated with an increase in blood flow and neuronal activity in the thalamus and medial PFC, which may be thought of as a nonspecific change associated with diverse emotional responses. More specific activation is seen in the left amygdala, hippocampal formation, and parahippocampal gyrus. Sadness generated by one's own thoughts (as opposed to a video scenario) and anticipatory anxiety are associated with a relative increase in blood flow to the anterior insular cortex. The most widely replicated fMRI findings associated with depression pertain to resting state connectivity, which is abnormally increased in the DMN and decreased connectivity within the CNN.

The most widely replicated PET finding in depression is decreased anterior brain metabolism, which is generally more pronounced on the left side. Thus, depression may be characterized by a relative increase in nondominant hemispheric activity. Decreased cerebral blood and reduced glucose metabolism in regions such as the dorsolateral PFC appears to be a severity-dependent process and is correlated with difficulties inhibiting affective expression with higher cognitive processes such as abstraction and

executive problem-solving skills. Abnormalities of regional cerebral metabolism have been observed in unipolar and bipolar depressions and appear to be somewhat state dependent. Furthermore, there is a reversal of hypofrontality after shifts from depression into hypomania, such that there are greater left hemisphere reductions in depression compared to greater right hemisphere reductions in mania. Other studies have observed more specific reductions in reduced cerebral blood flow or metabolism, or both, in the dopaminergically innervated tracts of the mesocortical and mesolimbic systems in depression. Again, there is evidence that antidepressants at least partially normalize these changes.

In addition to a global reduction in anterior cerebral metabolism, increased glucose metabolism has been observed in several limbic regions, particularly among patients with relatively severe recurrent depression and a family history of mood disorder (Fig. 13.9–3). During episodes of depression, increased glucose metabolism is correlated with intrusive ruminations. Of note, increased metabolic activity in the amygdalae and decreased activity in the dorsolateral PFC appear to be independent dysfunctions. Increased paralimbic metabolism appears to be reversible with effective pharmacotherapy, but, in one study, the abnormality re-emerged when recently remitted patients were restudied off medication. State-dependent or trait-like paralimbic hypermetabolism has been proposed to be an "emotional amplifier" that helps to distort the signal of various stressors in those at high risk of depression. Consistent with this, studies using fMRI have established that increased amygdalar activity in response to emotionally relevant stimuli is associated with the less functional short or *s* allele of the gene coding for the promoter region of the 5-HT transporter.

IMMUNOLOGICAL DISTURBANCE

Depressive disorders are associated with several immunological abnormalities, including decreased lymphocyte proliferation in response to mitogens and other forms of impaired cellular immunity. These lymphocytes produce neuromodulators, such as CRF, and cytokines, peptides known as *interleukins*. There appears to be an association with clinical severity, hypercortisolism, and immune dysfunction, and the cytokine interleukin-1 may induce gene activity for glucocorticoid synthesis. There is increasing evidence that depression is a proinflammatory illness, which may partly explain the high incidence of comorbid disorders such as obesity, diabetes, and atherosclerosis. Research is now underway to determine if strategies that dampen neuroinflammation can have antidepressant effects.

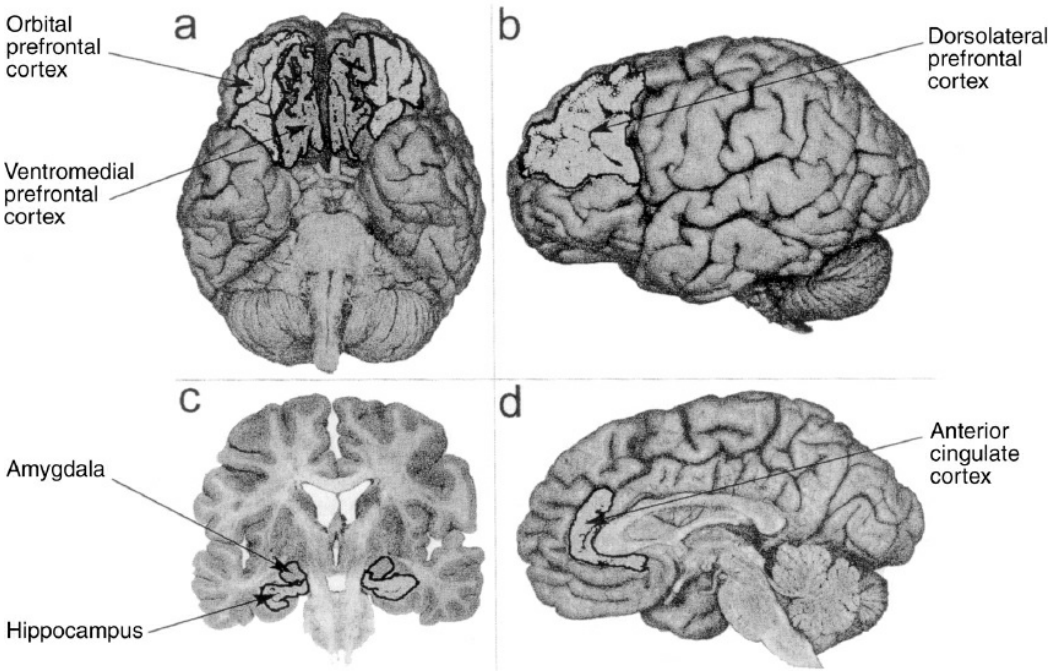

FIGURE 13.9–3. Key brain regions involved in affect and mood disorders. **A:** Orbital prefrontal cortex and the ventromedial prefrontal cortex. **B:** Dorsolateral prefrontal cortex. **C:** Hippocampus and amygdala. **D:** Anterior cingulated cortex.

SUMMARY

More severe and persistent depressive states, including those classified as major depressive disorder, are associated with a wide-range neurobiological disturbances, which in turn are linked to at least some of the differences observed in clinical presentations and response to specific treatments. Some of the disturbances are better understood as *traits*, which may be either inherited or acquired, whereas others are clearly state dependent and are reversible by treatment or spontaneous remission. Some of the state-dependent abnormalities associated with major depressive disorder, which occur more commonly in older patients with more severe symptomatology, include increased phasic REM sleep, poor sleep maintenance, hypercortisolism, impaired cellular immunity, reductions in anterior cerebral blood flow and glucose metabolism, and increased glucose metabolism in the amygdala. Together, these changes appear to reflect the progressive effects of otherwise adaptive short-term responses to sustained stress. Once manifest in this form, such severe or melancholic depressive episodes tend to be longer, more disabling, more prone to relapse, and more likely to benefit from pharmacotherapy or ECT (vis-à-vis nonspecific or psychotherapeutic interventions).

More trait-like abnormalities, which may include decreased slow wave sleep, reduced REM latency, blunted nocturnal GH secretion, and increased connectivity within the DMN. These abnormalities are associated

with an early age of onset and increased vulnerability to recurrent illness. The heritability of these abnormalities, which at least partly reflects decreased 5-HT inhibition of responses to stress or threat, is both inferred from family studies and other at-risk paradigms and has been linked to inheritance of the *s* allele of the 5-HT transporter.

Examples of more persistent, but acquired abnormalities may include global and focal changes, such as volume reduction in the hippocampus and anterior cingulate, hypertrophy of the adrenal cortex, periventricular hyperintensities, and alterations in CRH synthesis. Blunted TSH response (to a TRH infusion) and dexamethasone nonsuppression may represent *hybrids*, in that these abnormalities can be slow to normalize and, when persistent after remission, convey a high risk of relapse. Aging and the accumulating consequences of recurrent depressive episodes are inextricably connected. However, the diseases of aging that ravage brain function definitely increase the risk of depression and decrease responsivity to conventional forms of treatment. The late-onset form of depression associated with cerebral atherosclerosis illustrates the subtle interplay between brain function and mood disorder.

There is no doubt that vulnerability to mood disorders is partly heritable. This type of heritability is most likely polygenetic and, in all likelihood, will be best understood through models that include gene–environment interactions. Nevertheless, increased heritability is associated with an earlier age of onset, greater comorbidity, increased risk of recurrent illness, and an increased likelihood of hypomanic or manic episodes. Is it ironic that two of the more heritable forms of mood disorder, early onset chronic depressions and bipolar depression, are relatively less likely to be associated with state-dependent neurobiological disturbances? Different genetic vulnerabilities, different clinical presentations, and age-dependent differences in the effects of depression on brain function depression are likely to explain this apparent paradox.

Ultimately, depression remains a fundamentally human experience, partly because people can all relate to sadness, grief, and the heartbreak of lost love and partly because of the contrived nature of experiments that subject animals to prolonged, inescapable stress. Strong and sustaining affective bonds and an enduring sense of self-worth and competence are simply too important, and assaults on these fundamental aspects of well-being are too frequent for some humans to withstand without becoming depressed. That some are more vulnerable to depression than others is without question, as is the association of stress, depression, and numerous reproducible changes in brain function. Understanding the mechanisms of successful and unsuccessful adaptation and elucidating the alterations in brain function that predispose to and maintain depressive disorders represent the best hope to prevent and relieve the misery and suffering of tens of millions of people.

REFERENCES

Abdallah CG, Sanacora G, Duman RS, Krystal JH. Ketamine and rapid-acting antidepressants: a window into a new neurobiology for mood disorder therapeutics. *Annu Rev Med.* 2015;66:509–523.

Amstadter AB, Maes HH, Sheerin CM, Myers JM, Kendler KS. The relationship between genetic and environmental influences on resilience and on common internalizing and externalizing psychiatric disorders. *Soc Psychiatry Psychiatr Epidemiol.* 2015;51(5):669–678.

Bauer M, London ED, Silverman DH, Rasgon N, Kirchheiner J. Thyroid, brain and mood modulation in affective disorder: insights from molecular research and functional brain imaging. *Pharmacopsychiatry*. 2003;36(Suppl 3):S215.

Belmaker RH, Agam G. Major depressive disorder. *N Engl J Med.* 2008;358(1):55.

Bogdan R, Pagliaccio D, Baranger DA, Hariri AR. Genetic moderation of stress effects on corticolimbic circuitry. *Neuropsychopharmacology*. 2016;41:275–296.

Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science.* 2003;301(5631):386.

Coplan JD, Fathy HM, Jackowski AP, et al. Early life stress and macaque amygdala hypertrophy: preliminary evidence for a role for the serotonin transporter gene. *Front Behav Neurosci.* 2014;8:342.

Drevets WC, Thase ME, Moses-Kolko EL, Price J, Frank E. Serotonin-1A receptor imaging in recurrent depression: replication and literature review. *Nucl Med Biol.* 2007;34:865.

Dunlop BW, Mayberg HS. Neuroimaging-based biomarkers for treatment selection in major depressive disorder. *Dialogues Clin Neurosci.* 2014;16(4):479–490.

Furey ML, Drevets WC, Szczepanik J, Khanna A, Nugent A, Zarate CA Jr. Pretreatment Differences in BOLD Response to Emotional Faces Correlate with Antidepressant Response to Scopolamine. *Int J Neuropsychopharmacol.* 2015;18(8):pii: pyv028.

Gudayol-Ferré E, Peró-Cebollero M, González-Garrido AA, Guàrdia-Olmos J. Changes in brain connectivity related to the treatment of depression measured through fMRI: a systematic review. *Front Hum Neurosci.* 2015;9:582.

Homan P, Neumeister A, Nugent AC, Charney DS, Drevets WC, Hasler G. Serotonin versus catecholamine deficiency: behavioral and neural effects of experimental depletion in remitted depression. *Transl Psychiatry*. 2015;5:e532.

Kautzky A, Baldinger P, Souery D, et al. The combined effect of genetic polymorphisms and clinical parameters on treatment outcome in treatment-resistant depression. *Eur Neuropsychopharmacol.* 2015;25:441–453.

Kendler KS, Gardner CO, Prescott CA. Toward a comprehensive development model for major depression in women. *Am J Psychiatry*. 2002;159:1133.

Liu Z, Zhu F, Wang G, Xiao Z, Wang H. Association of corticotropin-releasing hormone receptor1 gene SNP and haplotype with major depression. *Neurosci Lett.* 2006;404:358.

Lyons DM, Yang C, Sawyer-Glover AM, Moseley ME, Schatzberg AF. Early life stress and inherited variation in monkey hippocampal volumes. *Arch Gen Psychiatry*. 2001;58:1145.

Miller AH, Raison CL. The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nat Rev Immunol.* 2015;16(1):22–34.

Nestler EJ. Role of the Brain's Reward Circuitry in Depression: Transcriptional Mechanisms. *Int Rev Neurobiol*. 2015;124:151–170.

Newport DJ, Carpenter LL, McDonald WM, Potash JB, Tohen M, Nemeroff CB; APA Council of Research Task Force on Novel Biomarkers and Treatments. Ketamine and other NMDA antagonists: early clinical trials and possible mechanisms in depression. *Am J Psychiatry*. 2015;172:950–966.

Phillips ML, Chase HW, Sheline YI, et al. Identifying predictors, moderators, and mediators of antidepressant response in major depressive disorder: neuroimaging approaches. *Am J Psychiatry*. 2015;172:124–138.

Posner J, Cha J, Wang Z, et al. Increased default mode network connectivity in individuals at high familial risk for depression. *Neuropsychopharmacology*. 2015;41(7):1759–1767.

Rajkowska G, Miguel-Hidalgo JJ. Gliogenesis and glial pathology in depression. *CNS Neurol Disord Drug Targets*. 2007;6:219.

Redei EE, Andrus BM, Kwasny MJ, et al. Blood transcriptomic biomarkers in adult primary care patients with major depressive disorder undergoing cognitive behavioral therapy. *Transl Psychiatry*. 2014;4:e442.

Rive MM, van Rooijen G, Veltman DJ, Phillips ML, Schene AH, Ruhé HG. Neural correlates of dysfunctional emotion regulation in major depressive disorder. A systematic review of neuroimaging studies. *Neurosci Biobehav Rev*. 2013;37(10 Pt 2):2529–2553.

Sanacora G, Banasr M. From pathophysiology to novel antidepressant drugs: glial contributions to the pathology and treatment of mood disorders. *Biol Psychiatry*. 2013;73(12):1172–1179.

Schmaal L, Veltman DJ, van Erp TG, et al. Subcortical brain alterations in major depressive disorder: findings from the ENIGM A Major Depressive Disorder working group. *Mol Psychiatry*. 2015;21(6):806–812.

Sheline YI, Gado MH, Kraemer HC. Untreated depression and hippocampal volume loss. *Am J Psychiatry*. 2003;160:1516.

Siegle GJ, Thompson WK, Collier A, et al. Toward clinically useful neuroimaging in depression treatment: prognostic utility of subgenual cingulate activity for determining depression outcome in cognitive therapy across studies, scanners, and patient characteristics. *Arch Gen Psychiatry*. 2012;69:913–924.

Smart C, Strathdee G, Watson S, Murgatroyd C, McAllister-Williams RH. Early life trauma, depression and the glucocorticoid receptor gene–an epigenetic perspective. *Psychol Med*. 2015;45:3393–3410.

Tafet GE, Nemeroff CB. The links between stress and depression: psychoneuroendocrinological, genetic, and environmental interactions. *J Neuropsychiatry Clin Neurosci*. 2015;28(2):77–88.

Thase ME. Depression and sleep: pathophysiology and treatment. *Dialogues Clin Neurosci*. 2006;8:217.

Thase ME. Using biomarkers to predict treatment response in major depressive disorder: evidence from past and present studies. *Dialogues Clin Neurosci*. 2014;16(4):539–544.

Thase ME, Berton O, Hahn CY. Biological aspects of depression. In: Gotlib IH, Hammen CL, eds. *Handbook of Depression*. New York: Guilford; 2014.

van Rossum EF, Binder EB, Majer M, Koper JW, Ising M. Polymorphisms of the glucocorticoid receptor gene and major depression. *Biol Psychiatry*. 2006;59:681.

van Winkel M, Peeters F, van Winkel R, et al. Impact of variation in the BDNF gene on social stress sensitivity and the buffering impact of positive emotions: replication and extension of a gene-environment interaction. *Eur Neuropsychopharmacol*. 2014;24(6):930–938.

14 Anxiety Disorders

▲ 14.1 Anxiety Disorders: Introduction and Overview

KATHLEEN RIES MERIKANGAS, PH.D.

INTRODUCTION

Anxiety disorders are the most prevalent psychiatric syndromes in the US population. Nearly one-fifth (17 percent) of adults report a lifetime history of one of the major anxiety disorders, and 1 in 10 suffer from a current anxiety disorder. Anxiety disorders are associated with intense subjective distress and social impairment, and their childhood and adolescent onset can interfere with accomplishment of social, educational, and occupational functioning. The most recent estimates of the Global Burden of Anxiety reveal that anxiety disorders explain 10 percent of the disability-adjusted life years for all mental, neurologic, and substance use disorders, second only to major depression. Despite their high burden, anxiety disorders have the lowest proportion of professional treatment of all the classes of mental disorders. The sections of this chapter document the phenomenology and classification of anxiety disorders in the DSM-5, epidemiology, risk factors and impact, biologic correlates including biochemical, neuroimaging, and genetics factors, and somatic and behavioral treatments.

BACKGROUND ON FEAR AND ANXIETY

Fear and anxiety can be conceptualized as two key core negative emotions. As such, definitions of "fear" and of "anxiety" must be based in broader definitions for emotions in general. For the current set of sections, the term "emotion" refers to the brain state associated with the perception of a motivationally salient stimulus, a stimulus that creates a need for the organism to act. Motivational stimuli can in turn be divided into "rewards" and "punishments," where a "reward" refers to a stimulus for which an organism will expend effort to approach, and a "punishment" refers to a stimulus for which an organism will expend effort to avoid.

"Fear" refers to the specific set of emotions or brain states that are

elicited in an organism when it confronts danger. Many current approaches refer to “fears” in the plural because “fears” are viewed as a collection of closely related but distinguishable brain states, as opposed to one, single state. This view reflects the fact that different forms of danger elicit different neural responses and associated differences in information processing and behavior: The response to a conspecific or some other form of innately dangerous stimulus can be distinguished, and both of these can in turn be distinguished from the response to objects that are viewed as dangerous through the effects of learning from experience. The term “danger,” in turn, refers to any stimulus or situation that is capable of producing harm to the organisms. The act of encountering a specifically dangerous object, such as a predator, can also be conceptualized as a “threat.” Moreover, because organisms tend to exert effort to virtually avoid all threats or dangerous scenarios, “threats” and “dangerous scenarios” can also be conceptualized as “punishments.”

The term “fear” can be distinguished from the term “anxiety” in a few respects. “Fear” is typically taken to refer to brain states associated with presentations of overtly dangerous stimuli. For example, fear is the emotion aroused when an organism immediately confronts a predator or a dangerous conspecific appearing in close proximity directly in front of the organism. As such, “fear” represents an acute, immediate reaction, manifest on abrupt encounters with highly salient threats. In research with rodents, this definition has at times been used to model “normal” responses in humans to dangers. For example, this type of “fear” response in the rodent is often modeled with the classic conditioning paradigm in which a rodent is placed in a cage and exposed to a cue, such as a tone or a light that predicts the immediate occurrence of an unconditioned aversive stimulus, such as a shock (UCS). In this fear conditioning paradigm the cue is conceptualized as a conditioned stimulus (CS+), and the organism’s response to the cue can be viewed as an analog of “fear.”

Unlike “fear,” from this perspective “anxiety” refers to brain states elicited by signals that predict impending, but not immediately presenting, danger. Thus, unlike “fear,” “anxiety” involves a more sustained change in the brain, manifest when a threat is still relatively removed from the organism in a spatial or temporal context. For example, “anxiety” might be conceptualized as the emotion aroused when an organism enters a deserted, isolated context where encounter with a predator appears likely, despite the isolated nature of the context and the absence of any overt cues indicating that the predator is lurking. In the fear conditioning experiment, “anxiety” is considered an analogue of pathologic reactions to danger in humans, as occurs in the anxiety disorders. This view reflects the relatively pervasive, chronic state of negative affect in clinical syndromes. In rodent experimental studies using fear conditioning, “anxiety” can be contrasted with “fear,” which manifests on immediate confrontation of the CS+;

"anxiety" typically is considered as manifesting to the general context of the fear-conditioning experiment. Thus, anxiety might be elicited when the organism is placed in the cage where the fear-conditioning experiment was previously conducted.

Thus, "fear" and "anxiety" have been distinguished on both temporal and spatial grounds. From the temporal perspective, fear refers to the acute, immediate response to suddenly appearing, imminent danger, whereas anxiety refers to the sustained, insidious response to danger, as might manifest when the degree to which a threat is present remains ambiguous. From the spatial perspective, fear refers to the acute response to proximal threats, whereas anxiety refers to the response to distal threats. As such, "fear" in rodents can be considered an analogue of "normal" human emotions elicited by thoughts about danger. "Anxiety" can be considered an analogue of "abnormal" human emotions in similar scenarios. Parenthetically, when an acute, proximal threat is particularly dangerous, however, the emotional state elicited in the organism might better be characterized as "panic" as opposed to "fear."

Conceptualizations of "normal" and "abnormal" fears and anxieties emerging from research directly among humans differ somewhat from conceptualizations in rodent-based or other basic research. Definitions in humans rest heavily on the presence of impairment, a disruption in normal functioning, or the presence of "clinically significant" distress. Basing definitions of abnormal anxiety on impairment or clinical judgments about distress does create some problems for attempts to conduct cross-species and neuroscience-focused research. The degree to which impairment or distress might manifest can be heavily influenced by the environmental circumstances to which the individual is exposed. Thus, two individuals manifesting identical neural responses to a dangerous stimulus might be classified differently in terms of the degree to which their associated anxiety can be considered "normal." If one individual faces hurdles not encountered by the other due to the nature of their environment and associated resources, these two individuals, who function identically from a neuroscience perspective, will be classified differently from a clinical perspective.

Thus, as delineated in the chapters pertaining to etiology, particularly those on neurophysiology, neurochemistry, and brain circuitry, the field is entering a period in which ongoing research in basic brain function and its cross-talk with the environmental context and stimuli will enhance the current understanding of the pathophysiology of disturbances in fear and anxiety. This knowledge has major future implications in its future potential for treatment and prevention.

DSM-5 MODIFICATIONS IN ANXIETY DISORDERS

The major subtypes of anxiety disorders in the DSM-5 include panic

disorder (with or without agoraphobia), agoraphobia (without a history of panic disorder), specific phobia, social phobia, and generalized anxiety disorder (GAD). Revisions to the classification of anxiety disorders in the DSM-5 involve removing obsessive-compulsive disorder and posttraumatic disorder have been subsumed under newly created "obsessive-compulsive and related disorders" and "trauma- and stressor-related disorders" categories, respectively. Therefore, both obsessive–compulsive disorder and posttraumatic stress disorder are not considered in this chapter. Other modifications to the proposed DSM-5 anxiety disorders category include the addition of separation anxiety disorder (contained under Disorders Usually First Diagnosed in Infancy, Childhood, or Adolescence in the DSM-IV), the identification of agoraphobia as a distinct and codable disorder (diagnosed only with reference to panic disorder in the DSM-IV), minor revisions to criterion language to enhance clarity, objectivity, and consistency across the anxiety disorders, and the relabeling of social phobia as social anxiety disorder (SAD). As such, the term "social phobia" will now be replaced with "social anxiety disorder."

ASSESSMENT OF ANXIETY

Anxiety may be assessed at the symptom or disorder level. Among adults, the most commonly used symptom measures include self-report checklists of both state and trait anxiety, fears, phobias, and avoidance, as well as clinician-administered symptom checklists. Anxiety disorders among adults are assessed by structured and semistructured diagnostic interviews designed for administration by either lay interviewers (e.g., the Composite International Diagnostic Interview [CIDI], the Diagnostic Interview Schedule for DSM-IV [DIS-IV]), or experienced clinicians (e.g., the Structured Clinical Interview for DSM Disorders [SCID], the Anxiety Disorder Interview Schedule-Revised [ADIS-R]). The ADIS-R is the most comprehensive diagnostic interview for the assessment of the diagnostic criteria of anxiety and affective disorders.

Assessment of symptoms and disorders of anxiety among children can involve obtaining information from the child alone or, more routinely, from individuals who observe the child in multiple contexts (i.e., caregivers and/or teachers), in addition to the child him- or herself. There are also several structured and semistructured diagnostic interviews of child anxiety disorders including the Diagnostic Interview Schedule for Children (NIMH-DISC-IV), the Composite International Diagnostic Interview modified for adolescents (CIDI), the Schedule for Affective Disorders and Schizophrenia for School-Age Children (KSADS-PL), and the Anxiety Disorders Interview Schedule for Children (ADIS-IV-C and P).

Several psychophysiologic indicators of anxiety have also been used to examine both human and animal anxiety. Experimental models that induce stress and measure autonomic output to test the human "fight or flight"

response to threat have been used to study the range of triggers, correlates and responses to fear-provoking situations. Behavioral tasks such as giving a speech or receiving the threat of shock have been used to experimentally induce anxiety states in unaffected individuals as well as those with ongoing anxiety disorders. Measures of changes in pulse, skin conductance, heart rate, and temperature regulation, as well as self-reported and observations of facial expression, blushing, and other overt signs of anxiety are presumed to provide a more accurate depiction of anxiety than self-reports or interviews about typical response patterns to stress.

OVERVIEW OF ANXIETY DISORDERS SECTION

The Anxiety Disorders section of this textbook covers the topics of phenomenology, epidemiology, genetics, neuroimaging, neurochemical aspects, and treatment of anxiety disorders.

In Section 14.2, Ryan J. Kimmel and Peter Roy-Byrne delineate the key clinical features of the anxiety disorders. The authors keep in line with the changes in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). The DSM-5 removed Post-Traumatic Stress Disorder and Obsessive Compulsive Disorder but continues to link panic, generalized anxiety, and phobic anxiety together. This section summarizes the key symptoms and discusses differential diagnosis as it pertains to delineating presentations that appear to occur in multiple disorders. The authors provide the historical background and possible implications of the changes (or the lack thereof) to each disorder in the DSM-5.

In Section 14.3, Kathleen R. Merikangas and John D. Eun describe the proliferation of epidemiologic research on anxiety disorders, both in the United States and abroad, that has strengthened the evidence base on the magnitude, correlates and consequences of bipolar disorder. Recent epidemiologic studies in youth have supported the findings in adults that anxiety disorders are the most prevalent mental disorders in the population. Prospective cohort studies have also demonstrated the stable course of anxiety disorders over lifetime that contribute to disability and increased risk of mortality. Epidemiologic research has highlighted the dramatic personal and societal impact of bipolar disorders. In the World Health Organization World Mental Health surveys, bipolar disorder was the second ranking cause of disability, measured by days spent our of role per year, among a range of physical and mental health conditions. Recent estimates of DALYs attributable to anxiety disorders show that they have the greatest attributable risk of all classes of mental disorders because of their early age of onset, and lifelong course.

Human research on the genetics of adult anxiety disorders is presented by John M. Hettema and Takeshi Otowa in HIN. The section provides a summary of family and twin studies that have shown that all of the anxiety disorder subtypes are familial with a two- to fourfold increased risk to

relatives and heritable, with a range from about 0.3 to 0.5. These studies established a foundation for employing approaches designed to identify the specific genetic factors involved in anxiety states including linkage, case-control candidate gene and genome-wide association studies. Despite few promising findings, the lack of consistency demonstrates that anxiety disorders are likely to be highly polygenic and etiologically and phenotypically heterogeneous. The two key limitations of extant research are the relatively small sample sizes, and the lack of inclusion of environmental factors and epigenetic mechanisms that may underlie their etiology.

Adriana Feder, James W. Murrough, Brian Iacoviello, and Dennis S. Charney summarize current knowledge of neurobiology and neurochemical basis of fear and anxiety in Section 14.5. They review the knowledge about anxiety disorders that arose from various animal models of stress and fear conditioning and the challenges in translating this work to humans. They conclude that panic disorder is caused not only by dysfunction to the norepinephrine and serotonin systems, but also by neurosteroids, opioid peptides, NPY, and other neurochemicals involved in the modulation of anxiety and, in some cases, response to treatment. Not surprisingly, other compounds, such as substance P, oxytocin, and orexin that have been shown in animal models to play a role in anxiety are now being investigated in human anxiety. They conclude that the use of techniques such as genetic analyses, neuroimaging combined with ligand binding, pharmacologic challenges, and behavioral and learning paradigms should provide a greater understanding of the diversity of these illnesses and aid in the development of novel therapeutic agents for these highly common and debilitating mental disorders.

In Section 14.6, Dylan Gee and BJ Casey review how neuroimaging approaches have informed the current models of neural circuitry underlying anxiety, fear, and stress-induced disorders. They provide an excellent summary of the prevailing models of anxiety disorders that emphasize interactions within limbic circuitry, especially involving prefrontal and sensory cortical regions, the amygdala, hippocampus, hypothalamus, bed nucleus of stria terminalis (BNST), and periaqueductal gray (PAG). They review current models of fear and stress-induced disorders, principally focusing on human neuroimaging studies. They demonstrate that the application of neuroimaging technology to investigations of the neurobiologic basis of anxiety disorders is yielding seminal information about the functional anatomical and neurochemical correlates of both normal and pathologic anxiety and fear states. Neuroimaging data from studies of panic disorder, specific phobias, SAD, and GAD support disease models that depend on altered interactions among the amygdala, hippocampus, and medial and orbital PFC. Future directions include gaining understanding of the neurodevelopmental

processes associated with anxiety disorders in childhood and adolescence and characterizing the heterogeneity of anxiety states through approaches that integrate imaging data with genetic and neurochemical information.

A review of the current state of cognitive-behavioral therapy is presented by Jonathan D. Huppert and Edna B. Foa (cognitive behavioral therapy [CBT]) in Section 14.7. There is evidence for the efficacy of CBT programs in the treatment of all anxiety disorders across children, adolescents, adults, and geriatric populations. Exposure in vivo to feared but otherwise safe stimuli and imaginal exposure to unwanted and upsetting thoughts are essential components of most CBT programs for anxiety. The cessation of safety behaviors or other subtle avoidance behaviors is essential for the anxiety disorders in order to disconfirm the patients' erroneous perceptions. The addition of interoceptive exposure appears to make a significant contribution to the treatment of panic disorder. In addition, exposure is typically conducted in the presence of cognitive challenging or discussions, helping the patient to attend to the corrective information. Anxiety management training and cognitive therapy have also been shown to be effective in the treatment of GAD. Cognitive therapy that includes behavioral experiments, which inherently involves in vivo exposure, is also an effective treatment for panic disorder and SAD. Overall, it seems that the integration of exposure therapy with discussions that increase attention of patients to the disconfirmation of their expectations of negative outcomes to exposures to the feared stimuli is the best method for optimizing treatment outcome.

In Section 14.8, Lakshmi N. Ravindran and Murray B. Stein review somatic treatments for each of the anxiety disorders. Several classes of agents have been evaluated in the treatment of anxiety and related disorders. These include antidepressants, benzodiazepines, anticonvulsants, and antipsychotics, among others. There are significant differences in efficacy and tolerability between classes as well as among agents within the same class. As well, there is significant variability in the symptoms they alleviate and the time points in the course of illness at which they produce optimal benefit. Although individually both pharmacotherapy and psychotherapy (e.g., CBT) show efficacy in the treatment of anxiety disorders, their combination has been little studied. However, available data shows inconsistent benefit for the combination. Some unresolved issues regarding somatic therapies are the optimal duration of pharmacotherapy, differential response by specific anxiety disorder subtypes, and the role of comorbidity in treatment response.

In summary, these sections provide an update on the current state of knowledge on the magnitude, clinical presentations, impact, etiologic factors, and treatments of anxiety disorders. There has been substantial progress in each of these domains since the last edition of this textbook. However, the challenges of etiologic and phenotypic heterogeneity, lack of

validity of diagnostic thresholds that distinguish disorder from basic human emotions of fear and anxiety, indistinct boundaries within specific anxiety subtypes and between anxiety disorders and mood and other emotional and behavior disorders continue to complicate the current understanding of their causal mechanisms.

References

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC; 2013.

Dunsmoor JE, Paz R. Fear generalization and anxiety: Behavioral and neural mechanisms. *Biol Psychiatry*. 2015;78(5):336–343.

Ganella DE, Kim JH. Developmental rodent models of fear and anxiety: from neurobiology to pharmacology. *Br J Pharmacol*. 2014;171(20):4556–4574.

Gilmartin MR, Balderston NL, Helmstetter FJ. Prefrontal cortical regulation of fear learning. *Trends Neurosci*. 2014;37(8):455–464.

Morrison FG, Ressler KJ. From the neurobiology of extinction to improved clinical treatments. *Depress Anxiety*. 2014;31(4):279–290.

Patel V, Chisholm D, Parikh R, et al; DCP MNS Author Group. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. *Lancet*. 2015;387(10028):1672–1685.

Tallot L, Doyère V, Sullivan RM. Developmental emergence of fear/threat learning: neurobiology, associations and timing. *Genes Brain Behav*. 2015;15(1):144–154.

▲ 14.2 Clinical Features of the Anxiety Disorders

RYAN J. KIMMEL, M.D., AND PETER ROY-BYRNE, M.D.

INTRODUCTION

Early clinicians delineated three different spheres of anxiety: phobias (context-dependent fears), generalized anxiety (context-independent fears), and panic. Antidepressant research that began in the 1960s contributed to the fundamental conception of the anxiety spectrum by demonstrating that various agents had efficacy for obsessive–compulsive, phobic, nonspecific anxiety, and panic symptom subsets. The American Psychiatric Association's interest in postcombat stress reactions rose in the 1970s in the wake of the Vietnam War. Based on a long and varied accumulation of evidence, then, during the 1980s DSM-III organized an "Anxiety Disorders" section that included specific diagnostic criteria for the component diagnoses of the Phobic Disorders, Generalized Anxiety Disorder (GAD), Panic Disorder, Obsessive–Compulsive Disorder, and Posttraumatic Stress Disorder (PTSD).

Though some diagnoses were added and the diagnostic criteria for the

original anxiety diagnoses were modified, the overarching organization of DSM-IV's "Anxiety Disorders" section was not a huge departure from DSM-III. The DSM-5, on the other hand, divides the anxiety spectrum into three pieces by separating out PTSD and OCD, but continuing to link panic, generalized anxiety, and phobic anxiety together. The DSM-5 also codifies numerous new diagnoses and asserts a close relationship between several of the classical anxiety disorders and a number of diagnoses that had been in other, nonanxiety classifications. Thus, the DSM-5's "Anxiety Disorders" section now contains Separation Anxiety Disorder, Selective Mutism, Specific Phobia, Social Anxiety Disorder, Panic Disorder, Agoraphobia, Generalized Anxiety Disorder, Substance/Medication-Induced Anxiety Disorder, Anxiety Disorder due to Another Medical Condition, Other Specified Anxiety Disorder, and Unspecified Anxiety Disorder.

The DSM-5 subsumes PTSD, Reactive Attachment Disorder, Disinhibited Social Engagement Disorder, Acute Stress Disorder, and Adjustment Disorder into a new section entitled "Trauma- and Stressor--Related Disorders." The final grouping is codified within the DSM-5 as the "Obsessive-Compulsive and Related Disorders" and includes Obsessive-Compulsive Disorder (OCD), Body Dysmorphic Disorder, Hoarding Disorder, Trichotillomania, and Excoriation Disorder.

With PTSD and OCD removed from the DSM-5 Anxiety Disorders section, modern nomenclature has made an interesting circuit. Emil Kraeplin, Anna Freud, John Bowlby, and other anxiety theorists from the prior century might recognize the logic behind the DSM-5's decision to recircumscribe its Anxiety Disorders, gathering together panic, generalized anxiety, phobias, and separation anxiety. Every generation, however, has researched, argued, and reshuffled the anxiety spectrum to see if drawing new parallels and new divisions would yield improved insight and enhanced treatment efficacy.

Keeping in line with the current nosology, the following material describes the contents of DSM-5's Anxiety Disorders section, whereas the Trauma-Related and Obsessive–Compulsive spectra are detailed in other chapters. Etiology and treatment are briefly referenced, but are more extensively covered in other sections of this book. Differential diagnosis is discussed, especially as it pertains to delineating between presentations that appear to straddle multiple disorders in the DSM-5's Anxiety Disorder grouping. However, given that anxiety-like symptoms appear in most of the diagnostic clusters in Section II of the DSM-5, there is not enough room in this section to address an exhaustive differential. The references included at the end of this section are a good source of catalogued confounders, as is the DSM-5 itself. Moreover, the authors of this section would recommend the reader to have a copy of the DSM-5 available in order to cross-reference the specific, diagnostic criteria that could not be repeated verbatim in the text below.

PANIC DISORDER AND THE PANIC ATTACK SPECIFIER

Definition, Diagnosis, and Clinical Features

The DSM-5 defines a panic attack as a rapidly escalating, but circumscribed period, of four of a cluster of cardiorespiratory, gastrointestinal, otoneurologic, autonomic, or cognitive symptoms (the specific symptoms are listed in Criterion A of Table 14.2–1). Certain culture-specific symptoms may be characteristic of panic attacks experienced by some patients, but these do not supplant the need for four symptoms from the DSM-5's list of panic attack criteria.

Table 14.2–1.
Criteria for Panic Disorder

A. Recurrent unexpected panic attacks. A panic attack is an abrupt surge of intense fear or intense discomfort that reaches a peak within minutes, and during which time four (or more) of the following symptoms occur:
Note: The abrupt surge can occur from a calm state or an anxious state.

1. Palpitations, pounding heart, or accelerated heart rate.
2. Sweating.
3. Trembling or shaking.
4. Sensations of shortness of breath or smothering.
5. Feelings of choking.
6. Chest pain or discomfort.
7. Nausea or abdominal distress.
8. Feeling dizzy, unsteady, light-headed, or faint.
9. Chills or heat sensations.
10. Paresthesias (numbness or tingling sensations).
11. Derealization (feelings of unreality) or depersonalization (being detached from oneself).
12. Fear of losing control or "going crazy."
13. Fear of dying.

Note: Culture-specific symptoms (e.g., tinnitus, neck soreness, headache, uncontrollable screaming or crying) may be seen. Such symptoms should not count as one of the four required symptoms.

B. At least one of the attacks has been followed by 1 month (or more) of one or both of the following:

1. Persistent concern or worry about additional panic attacks or their consequences (e.g., losing control, having a heart attack, "going crazy").
2. A significant maladaptive change in behavior related to the attacks (e.g., behaviors designed to avoid having panic attacks, such as avoidance of exercise or unfamiliar situations).

C. The disturbance is not attributable to the physiologic effects of a substance (e.g., a drug of abuse, a medication) or another medical condition (e.g., hyperthyroidism, cardiopulmonary disorders).

D. The disturbance is not better explained by another mental disorder (e.g., the panic attacks do not occur only in response to feared social situations, as in social anxiety disorder; in response to circumscribed phobic objects or situations, as in specific phobia; in response to obsessions, as in obsessive–compulsive disorder; in response to reminders of traumatic events, as in posttraumatic stress disorder; or in response to separation from attachment figures, as in separation anxiety disorder).

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

A single panic attack is not necessarily indicative of a psychiatric disorder. Moreover, panic attacks can be a feature of a wide variety of psychiatric and medical diagnoses. Thus, DSM-5's Panic Attack Specifier may be added to other diagnoses in an effort to further describe the experience and impairing condition of any single patient. For example, a patient who experiences panic attacks only during an exacerbation of severe major depressive episodes may have the Panic Attack Specifier coded alongside their Major Depressive Disorder.

The experience of a patient with Panic Disorder, on the other hand, involves the unrelenting fear of future panic attacks in the setting of a history of recurrent episodes of panic (Table 14.2–1). By definition, this fear results in behavioral changes designed to avoid panic triggers or avoid situations where the impact of a panic attack would be especially disconcerting, or to gain reassurance that nothing is wrong medically (by making ER visits or having repetitive medical testing). Panic attacks may "come out of the blue" without an obvious trigger. Panic attacks may also be the result of a specific trigger or may be associated with situations where a specific trigger is anticipated.

The DSM-5's definition of Panic Disorder maintains the DSM-IV's requirement that at least some of the panic attacks be seemingly random and unexpected. This helps separate the definition of Panic Disorder from, for example, the definition of a Specific Phobia of heights with a Panic Specifier. While an exposure to heights may predictably trigger a panic attack, unless there are context-independent panic attacks, the patient would not meet the criteria for Panic Disorder.

History and Comparative Nosology

Though the term "panic attack" is a 20th century construct, descriptions of what is now referred to as "Panic Disorder" date back to prior centuries. The disorder was identifiable enough that, several decades before DSM-III, psychopharmacology researchers were able to clearly identify patients with recurrent panic attacks and enable early studies that demonstrated the efficacy of antidepressants.

The biggest debate over the last 20 years regarding the Panic Disorder diagnosis has been in its relation to agoraphobia. The DSM-IV contained Panic Disorder with Agoraphobia and Panic Disorder without Agoraphobia, framing agoraphobia as a fear of public situations where a panic attack would be especially destructive. However, the DSM-5 identifies agoraphobia as its own diagnosis, rather than as a specifier of Panic Disorder. The newly-independent agoraphobia diagnosis will be discussed later in this chapter.

Epidemiology

Successive iterations of the DSM maintain that experiencing a single panic attack does not meet criteria for Panic Disorder. This is evidenced by data suggesting that only 1 in 25 people will have Panic Disorder in their lifetime, but at least 1 in 7 (with some measurements closer to 1 in 3) will have a panic attack. Panic Disorder has a high rate of comorbidity with affective disorders, substance use disorders, and other anxiety and trauma-related disorders. The age of onset of Panic Disorder peaks in the late teens and early twenties. Twice as many women as men are affected.

Differential Diagnosis, Etiology, Course, and Treatment

As mentioned previously, panic attacks may be a feature of other psychiatric conditions such that a more accurate description of a patient's condition may require the addition of the Panic Attack Specifier to another DSM-5 diagnosis. However, there are also a variety of medical conditions that can yield experiences that may be misconstrued as panic with a psychiatric etiology. For example, certain central nervous symptom pathologies, such as brain cancer or psychomotor epilepsy, can mimic panic attacks. There are many medications that can induce panic symptoms, with classic examples being corticosteroids, stimulants, certain hormones, and some asthma medications. Illicit stimulants and hallucinogens can induce panic symptoms. Certain endocrinopathies, whether acting centrally or peripherally, can yield panic-like symptoms. Examples would be thyroid disorders or parathyroid disorders. The episodic, but dramatic, symptoms of pheochromocytoma can also be initially misconstrued as panic. Cardiac abnormalities (e.g., causes of chest pain), as well as pulmonary conditions that result in acute respiratory distress (e.g., asthma, pulmonary embolus), may also masquerade as panic symptoms.

Though there is evidence that Panic Disorder has a heritable component, there is insufficient information about panic-related genetic factors to direct diagnosis or treatment. Often, stressful life events are noted to have clustered around the time of the onset of Panic Disorder. Certain "anxious temperaments," as well as childhood trauma, have been linked to higher rates of Panic Disorder in adulthood.

Once full criteria have been met, the disorder tends to be chronic, though the course is often fluctuating. Even after treatment to the point of remission, the rate of relapse is high. For example, naturalistic studies often demonstrate a >50 percent relapse rate within 12 months of discontinuing an effective antidepressant. Given the impact of Panic Disorder on psychosocial functioning, as well as the increased rate of suicide attempts in these patients, ongoing monitoring is warranted if the patient chooses to disengage from active treatment.

Serotonergic antidepressants are the mainstay of pharmacologic treatment. Several benzodiazepines also have FDA approval for Panic

Disorder and their efficacy and associated risks are discussed elsewhere in this book. There is also significant data to support treatment with CBT.

AGORAPHOBIA

Definition, Diagnosis, and Clinical Features

The DSM-5 defines agoraphobia as characterized by fear (imminent) or anxiety (anticipated) in multiple public situations, be they amongst crowds or simply being out alone. More specifically, patients with agoraphobia fear these milieus because they sense that, should they develop a condition that is frightening or humiliating, they would be unable to get help or quickly escape the setting. The fear or anxiety is, as per DSM-5, "out of proportion" with the real threat. The behavioral manifestation of this fear is avoidance. The emotions (fear or anxiety) or behavior (avoidance) persists for longer than 6 months.

Patients with agoraphobia may develop "panic-like" symptoms, as per DSM-5, without meeting full criteria for the Panic Attack Specifier. Patients also may have other psychosomatic symptoms, often gastrointestinal or autonomic, associated with fear or anxiety. This combination of fear and psychosomatic symptoms leads to major life dysfunction. Eventually, patients may avoid all employment and in-person social interactions.

History and Comparative Nosology

In DSM-IV, agoraphobia was subsumed under Panic Disorder as a required specifier. Clinicians thus had the option of coding for Panic Disorder With Agoraphobia or Panic Disorder Without Agoraphobia. This is understandable from a historical perspective, given clinical theorists previously placed agoraphobia within the broader subset of anxieties and behaviors that might be born from a fear of a panic attack. However, not every patient with agoraphobia will have symptoms that rise to meet DSM-5's full criteria for a Panic Attack. Thus, agoraphobia merits its own diagnosis in the DSM-5.

Epidemiology

Similar to Panic Disorder, more women than men have agoraphobia and the age of onset peaks in the late teens to early twenties. Agoraphobia in the absence of Panic Disorder is considered to be rarer than agoraphobia with comorbid Panic Disorder. However, there is some variability in the prevalence data. The measured prevalence of agoraphobia in specific clinical settings may evolve as the DSM-5's recognition of agoraphobia without Panic Disorder will spur clinicians to screen and consider the disorder more frequently, even in patients who do not present with panic attacks. Other anxiety disorders are seen alongside agoraphobia in comorbidity rates that often exceed 50 percent. Comorbid depressive

disorders are seen in 33 to 52 percent of cases, with some suggestion that the presence of comorbid panic attacks increases the risk of comorbid depressive episodes.

Differential Diagnosis, Etiology, Course, and Treatment

Although Specific Phobia could be confused with agoraphobia, Specific Phobia is limited to one stimulus, whereas agoraphobia's diagnostic criteria require at least two situations. Moreover, the difference can sometimes be parsed out by determining if the patient is, for example, loathe to walk across a bridge because they are afraid of heights (Specific Phobia) or because a bridge is an exposed, open area with limited exits should some frightening event occur (agoraphobia). Similarly, patients with Separation Anxiety Disorder may refuse to leave home because the fear of permanent separation from an attachment figure, patients with PTSD may avoid specific stimuli that trigger traumatic memories in certain community situations, patients with Social Anxiety Disorder may prefer home to situations where they might be judged and embarrassed, and patients with OCD may be effectively homebound due to specific obsessions about contamination when outside. However, these disorders can be usually be distinguished from agoraphobia by the breadth of the avoidance and the motivation for the behavior.

Agoraphobia seems to be, in part, heritable, with certain personality factors predicting the onset, as well. Agoraphobia tends to persist if not treated. It is presumed, based on the improvement in agoraphobia symptoms in DSM-IV era studies of Panic Disorder with agoraphobia, that the same pharmacotherapies that are efficacious for Panic Disorder would also be efficacious for agoraphobia. DSM-5's designation of agoraphobia as an independent diagnosis will help spawn prospective trials and retrospective, naturalistic studies that will test this assumption. In contrast to the pharmacotherapy literature, a number of studies already document the efficacy of behavioral treatment for agoraphobia independent of panic disorder.

SOCIAL ANXIETY DISORDER

Definition, Diagnosis, and Clinical Features

Social Anxiety Disorder is a fear of one or more social situations wherein the patient is concerned about being harshly judged and evaluated for their performance or interpersonal interactions. There is an out-of-proportion fear that they will be shamed by the experience. The symptoms last longer than 6 months and the specific social situation(s) are avoided altogether or tolerated only with significant anxiety. Many people describe increased anxiety in social situations, but a key to the Social Anxiety diagnosis, per DSM-5, is a "significant impairment" in an important facet of life

functioning.

Disabling performance anxiety when speaking in front of a crowd is a common example of Social Anxiety Disorder, as is severe anxiety at having to meet and engage with new people. However, patients can also fear having to do simple, everyday tasks in any setting that could be loosely described as "social." That is, unlike speaking in front of a crowd, the feared situations in Social Anxiety Disorder need not be ones where it would be common and reasonable to expect some level of judgment. An example would be a fear of writing one's signature in front of others out of concern that the manner of writing, no matter how ordinary, will be a source of derision. This fear may be compounded with anxiety about appearing anxious.

History and Comparative Nosology

The formulation of Social Anxiety Disorder has continued to evolve since DSM-III and the transition from DSM-IV to DSM-5 is no exception. The principal alteration in the most recent DSM is a specifier change. In DSM-IV, Social Anxiety Disorder was subdivided into "specific" and "generalized" forms. However, public speaking was far and away the most common "specific" form. Most patients with Social Anxiety disorder fear and avoid multiple social situations and thus the "specific" moniker, apart from performance anxiety, was seldom used. The DSM-5 thus conceives the generalized pattern to be the norm and the only specifier is "performance only."

The other nosologic debate that has impacted this diagnosis over the last several decades is whether the old, "generalized," subtype of Social Anxiety Disorder could be distinguished from Avoidant Personality Disorder. More specifically, Avoidant Personality Disorder has, at times, been conceived as a severe form of Social Anxiety Disorder. While there is an unusually high rate of comorbidity between the two diagnoses, as currently defined, it still appears that less than half of patients with Avoidant Personality Disorder have Social Anxiety Disorder. Moreover, some theories align Avoidant Personality Disorder closer to Schizophrenia than Social Anxiety. DSM-5 thus characterizes Avoidant Personality as phenomenologically distinct from the anxiety spectrum, though this issue has been, and will be, a topic of significant research and debate.

Epidemiology

Social Anxiety Disorder is more common than some of the other disorders described in this section, with a cross-national prevalence rate of 2 to 3 percent. It is more common in women than in men, has peak onset in the early teenage years (often associated with the move to junior high school or high school, or some other increase in social context complexity), and is

associated with an increased risk of depressive episodes.

Differential Diagnosis, Etiology, Course, and Treatment

As mentioned above, Avoidant Personality Disorder (described in more detail elsewhere in this book) has been the subject of debate over whether it is distinct from Social Anxiety Disorder, with the DSM-5 separating the two. Social Anxiety Disorder has a high comorbidity with other anxiety and affective disorders, as well. The use of illicit anxiolytics and sedatives leads to the relatively high rate of comorbid substance use disorders. The comorbidity between Social Anxiety Disorder and Selective Mutism is discussed elsewhere in this chapter. Social Anxiety Disorder is a common comorbidity in children with an Autism Spectrum diagnosis and also appears at rates above population baseline in patients with Schizophrenia.

Risk factors for Social Anxiety Disorder include female gender, family history, and childhood signs of behavioral inhibition. There is insufficient data on specific genetic factors mediating the increased familial risk, but parenting styles may also contribute to this familiality. The Mini-Social Phobia Inventory (Mini-SPIN) is an appropriate screening tool for adults.

One could see how agoraphobia could be confused with Social Anxiety Disorder or PTSD, though the context-dependent fears in the same situation would be different. For example, a patient with Social Anxiety Disorder might dislike a crowded party because they feel that everybody is looking at them and judging them. A patient with agoraphobia might avoid the same crowded party out of concern that, were they to develop anxiety, it would be difficult to sprint for the exit. A patient with PTSD might find that their desire for hypervigilance is overwhelmed by the multiple stimuli in a crowded room.

There is a markedly increased rate of suicide attempt amongst patients with Social Anxiety Disorder. This, along with the significant functional impairment associated with this disorder, should motivate clinicians to pursue aggressive treatment for these patients. CBT has demonstrated efficacy for Social Anxiety Disorder. The first-line pharmacologic treatments are serotonergic agents, but other, PRN medications, especially for the performance subtype (beta blockers), have been used effectively. Social Anxiety Disorder is a chronic condition, with a high rate of symptom re-emergence after symptom remission is achieved via selective serotonin reuptake inhibitors (SSRIs), for example. However, durable remission, even after cessation of CBT and/or antidepressants, has been seen in a small proportion of patients.

SPECIFIC PHOBIA

Definition, Diagnosis, and Clinical Features

Specific Phobias are fears of specific objects or situations, such as spiders

or blood, that go beyond the true threat of the stimulus and cause avoidance and functional impairment (Table 14.2–2). The patient is fearful nearly every time they are confronted with the phobic stimulus. The phobia, by rule, lasts longer than 6 months. Many patients have multiple phobic stimuli, in which case each would be coded separately. ICD-10 has specific codes for phobias of animals, natural environment (heights, storms, water), fear of blood, fear of injections, fear of transfusions, fear of other medical care, fear of injury, fear of a specific situation (elevators, airplanes, enclosed spaces, etc.), and fear of other (anything that does not fit into one the above).

History and Comparative Nosology

Circumscribed phobias have long been recognized. They are amenable to nonpharmacologic interventions and, as such, have been central to the delineation of developmental theories and psychotherapy strategies for over a century. A significant development over the course of the DSM iterations has been the separation and diagnostic refinement of Social Anxiety Disorder (formerly Social Phobia), agoraphobia (formerly subsumed under Panic Disorder), and Specific Phobia (formerly Simple Phobia). The number of subcategories of Specific Phobia, as well as the value of trying to group the various phobic stimuli, continues to evolve in successive DSM iterations. Determining how much distress and impairment is necessary to separate the range of normal from the range of pathologic remains a nosologic challenge.

Table 14.2–2.
Criteria for Specific Phobia

A. Marked fear or anxiety about a specific object or situation (e.g., flying, heights, animals, receiving an injection, seeing blood).
   Note: In children, the fear or anxiety may be expressed by crying, tantrums, freezing, or clinging.
B. The phobic object or situation almost always provokes immediate fear or anxiety.
C. The phobic object or situation is actively avoided or endured with intense fear or anxiety.
D. The fear or anxiety is out of proportion to the actual danger posed by the specific object or situation and to the sociocultural context.
E. The fear, anxiety, or avoidance is persistent, typically lasting for 6 months or more.
F. The fear, anxiety, or avoidance causes clinically significant distress or impairment in social, occupational, or other important areas of functioning.
G. The disturbance is not better explained by the symptoms of another mental disorder, including fear, anxiety, and avoidance of situations associated with panic-like symptoms or other incapacitating symptoms (as in agoraphobia); objects or situations related to obsessions (as in obsessive–compulsive disorder); reminders of traumatic events (as in posttraumatic stress disorder); separation from home or attachment figures (as in separation anxiety disorder); or social situations (as in social anxiety disorder).

Specify if:

Animal (e.g., spiders, insects, dogs).
Natural environment (e.g., heights, storms, water).

Blood-injection-injury (e.g., needles, invasive medical procedures)
Situational (e.g., airplanes, elevators, enclosed places).
Others (e.g., situations that may lead to choking or vomiting; in children, e.g., loud sounds or costumed characters).

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Epidemiology

The lifetime prevalence of Specific Phobia is 6 percent. The highest prevalence occurs in the teenage years. Women have a higher prevalence than men, though the issue of quantifying and interpreting gender-specific prevalence discrepancies, as in many psychiatric disorders, is complicated by the interaction of neurobiologic, cultural, regional, behavioral, psychodynamic, and psychosocial factors, both in the diagnosed and in those doing the diagnosing.

Differential Diagnosis, Etiology, Course, and Treatment

As was articulated previously, delineating between the context-dependent phobias can be a challenge when getting a history from any specific patient. Though some patients develop a Specific Phobia only after a traumatic event involving a threatening phobic stimulus, these patients, by definition, do not develop the breadth of symptoms that meet the criteria for PTSD. Similarly, patients can develop a Specific Phobia to an event following a panic attack, but Panic Disorder requires that at least some panic attacks that are unexpected, rather than just coming from a known phobic stimulus. As articulated previously, agoraphobia and Specific Phobia may be difficult to distinguish if the phobic stimulus is in a specific, public location. However, patients with agoraphobia ought to fear multiple public situations. Moreover, patients with agoraphobia specifically fear the panic or shame that can develop in public situations.

Some Specific Phobias, such as animal phobias, have a unique age of onset, suggesting a nontraumatic etiology perhaps more based in evolutionary preparedness. Other Specific Phobias, such as blood-injury-illness phobia, have a unique physiology (i.e., bradycardia and hypotension as opposed to the usual tachycardia and hypertensive reaction) and a greater rate of familiality that suggest still another distinct etiology.

There is a genetic component, but heritability does not seem specific for the type of Specific Phobia, but rather for the general risk of developing some kind of Specific Phobia. For many phobias, environmental factors, such as trauma, may be more important. As with other anxiety disorders, cognitive factors and conditioning may also play a role.

While the course of specific phobia may be especially chronic, there are effective treatments. Patients often recognize that their anxiety exceeds the true threat of the stimulus and this insight can often motivate a patient to

pursue therapy. The treatment of choice for Specific Phobia is a behavioral therapy series that includes exposure-based, systematic desensitization. There are small, placebo-controlled studies involving serotonergic antidepressants or atypical antipsychotics, but the data is very limited. There is ongoing work on augmenting exposure-based psychotherapy with d-cycloserine, though the data requires ongoing interpretation and study replication to avoid harm via administration of this medication in the wrong context. There is very little long-term data on the durability of remission in pharmacologic studies.

GENERALIZED ANXIETY DISORDER

Definition, Diagnosis, and Clinical Features

GAD is characterized by an uncontrollable worry, more days than not, lasting greater than 6 months, and causing significant impairment. The worry must be associated with at least three of the following six symptoms: "restlessness or feeling keyed up or on edge; being easily fatigued; difficulty concentrating or mind going blank; irritability; muscle tension; or sleep disturbance" (Table 14.2–3). All three symptoms need not be present simultaneously every day. The worry usually involves a broad swath of everyday life, such as simple daily activities, timeliness, finances, or health. These are issues that may concern the average patient on a daily or weekly basis. However, patients with GAD tend to exaggerate the likelihood and severity of the stressor to the point where catastrophe seems possible, likely, and imminent. Another key to the GAD diagnosis is that these concerns cannot be prioritized as less important in the face of the random, more pressing matters that pop up on a day-to-day basis. This inappropriate prioritization of certain anxieties contributes to the pathologic effect that GAD has on psychosocial functioning and to its chronicity.

History and Comparative Nosology

The definition of GAD has been beset by alterations in diagnostic criteria since its inclusion in DSM-III. These alterations have ranged from the broadly theoretical, such as whether GAD is a diagnosis of exclusion for patients whose anxiety does not seem to meet criteria for another anxiety disorder, to the empirically detailed, such as how many concurrent symptoms are necessary over what duration of time. Moreover, the tremendously high rate of comorbidity makes it difficult to isolate GAD for the purposes of evaluating the fidelity of the criteria, impairs the ability to track the durability of the diagnosis, and complicates evaluation of treatment efficacy. Interestingly, the pace of the nosologic evolution of GAD has slowed dramatically as evidenced by the diagnostic agreement between DSM-IV and DSM-5. This may reflect a lack of consensus on

proposed changes to the diagnostic criteria and/or indicate an awareness of the value of maintaining nosologic stability for the purposes of treatment efficacy research.

Table 14.2–3.
Criteria for Generalized Anxiety Disorder

A. Excessive anxiety and worry (apprehensive expectation), occurring more days than not for at least 6 months, about a number of events or activities (such as work or school performance).
B. The individual finds it difficult to control the worry.
C. The anxiety and worry are associated with three (or more) of the following six symptoms (with at least some symptoms having been present for more days than not for the past 6 months):
Note: Only one item is required in children.

   1. Restlessness or feeling keyed up or on edge.
   2. Being easily fatigued.
   3. Difficulty concentrating or mind going blank.
   4. Irritability.
   5. Muscle tension.
   6. Sleep disturbance (difficulty falling or staying asleep, or restless, unsatisfying sleep).

D. The anxiety, worry, or physical symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
E. The disturbance is not attributable to the physiologic effects of a substance (e.g., a drug of abuse, a medication) or another medical condition (e.g., hyperthyroidism).
F. The disturbance is not better explained by another mental disorder (e.g., anxiety or worry about having panic attacks in panic disorder, negative evaluation in social anxiety disorder [social phobia], contamination or other obsessions in obsessive–compulsive disorder, separation from attachment figures in separation anxiety disorder, reminders of traumatic events in posttraumatic stress disorder, gaining weight in anorexia nervosa, physical complaints in somatic symptom disorder, perceived appearance flaws in body dysmorphic disorder, having a serious illness in illness anxiety disorder, or the content of delusional beliefs in schizophrenia or delusional disorder).

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Epidemiology

GAD has a lifetime prevalence rate of 5 percent, often begins in the late teens, and is more common in women as men. The high lifetime psychiatric comorbidity has led some to view GAD as a prodromal or residual phase of a major depression, though there is insufficient empirical support for this view. Other common comorbidities include other anxiety spectrum disorders and substance use disorders.

Differential Diagnosis, Etiology, Course, and Treatment

For some patients, GAD distinguishes itself from other anxiety diagnoses with its relatively high, constant rate of symptoms, rather than the episodic course seen in the more context-dependent anxiety disorders. However, GAD may still have a fluctuating course. The biggest challenge, given the comorbidity between GAD and Major Depressive Disorder, can be

distinguishing whether a depressed patient's high anxiety is a feature of their depressive episode or represents an independent, comorbid disorder. The easiest way to determine this would be to evaluate the anxiety during periods of euthymia, though this often requires a long period of observation.

A number of genetic associations have been noted in GAD, though these are not specific enough to guide diagnosis or treatment. As with many other anxiety disorders, childhood sexual abuse is associated with an increased incidence of GAD. The same, wide variety of neurologic, medical, medication, and chronic illicit substance etiologies listed in the differential for Panic Disorder can also present with GAD-like symptoms. Beyond maintaining a high index of suspicion, presentations that include a marked change in intensity of GAD symptoms, a marked decline in functioning, or an atypical age of symptom onset may prompt a medical work-up.

The evolving criteria associated with successive DSM iterations have resulted in different classes of medications undergoing placebo-controlled trials under somewhat different diagnostic and inclusion criteria. For example, buspirone-garnered FDA approval under a pre-DSM-IV set of diagnostic criteria for how we would define GAD now. Placebo-controlled trials of benzodiazepines, similarly, were studied chiefly in the DSM-III and pre-DSM-III eras. It would be difficult to replicate the depth and breadth of these trials with every new DSM. One of the original motivations to delineate generalized (context-independent) fears and panic was early data suggesting a difference in the efficacy of antidepressants between the two. However, over the last 30 years, a variety of tricyclics, SSRIs, and SNRIs have demonstrated efficacy for both Panic Disorder and GAD.

Multiyear, naturalistic studies of GAD show that relapse is common. Thus, studies that follow patients for only a few months after cessation of an effective treatment modality may give clinicians a false sense of security. Patients with GAD are at an increased risk for suicide attempt. Clinicians thus need to regularly monitor for recurrence of GAD symptoms, similar to how one would routinely screen a patient who had a history of multiple major depressive episodes. The GAD-7 screening tool is a quick, valid, and useful instrument for tracking the course of this disorder. Screening for suicidal ideation is also recommended.

Long-term studies of GAD symptom relapse rates following discontinuation of antidepressants after treatment response have led to recommendations that pharmacologic treatment continue for longer than 6 months, though longer follow-up studies are needed. Given that context-independent anxiety disorders have long been formulated as separate from panic and context-dependent phobias, there is a significant body of literature on psychotherapy for the GAD spectrum, including CBT, relaxation, imagery exposure, and various meditation modalities. These modalities are discussed in more detail in other sections of this book.

SELECTIVE MUTISM

Definition, Diagnosis, and Clinical Features

Selective Mutism is a disorder of children wherein they are unable to speak in certain social situations. These children are capable of speaking and do so when at home. Though the lack of speech at school is what often leads to an evaluation with a mental health provider, the disorder can be evident during other interactions between the child and less-familiar adults. The mutism must impact social or educational functioning for greater than a month in order to meet DSM-5's diagnostic criteria. The social situation wherein the symptoms manifest must not be one where a relative mutism might be expected, such as just after starting school or while integrating into a setting with a new language.

History and Comparative Nosology

DSM-5's inclusion of this diagnosis in the Anxiety Disorders section stems from evidence that anxiety is often present as a symptom of Selective Mutism. However, there is heterogeneity in terms of comorbid social anxiety symptoms, developmental disorders, communication disorders, and many other symptom subsets. In the process of reframing the discussion of the disorder toward themes of anxiety and phobia, the name Elective Mutism was switched to Selective Mutism in order to remove the implication that the symptoms were born of defiance.

Epidemiology

Fewer than 1 percent of children have Selective Mutism. Selective Mutism has at times been formulated as part of the Social Anxiety Disorder spectrum. Though Social Anxiety Disorder is likely the most common comorbidity, Social Anxiety Disorder has an age of onset in the teens whereas Selective Mutism usually presents by age 5 years. The rarity of Selective Mutism makes it a challenge to quantify exactly, but there seems to be a higher prevalence in girls.

Differential Diagnosis, Etiology, Course, and Treatment

Communication disorders, autism spectrum, and developmental delay would be important etiologies to rule out when making the diagnosis of Selective Mutism. If there is a broader, autism spectrum diagnosis with prominent communication symptoms that better explains the symptoms, the Selective Mutism diagnosis is not applied. Selective Mutism should prompt further consideration and screening for comorbid anxiety disorders. A wide variety of genetic, temperament, environmental, and neurodevelopmental contributors have been suggested to contribute to the etiology of Selective Mutism.

CBT and behavioral therapies have shown efficacy in various studies. CBT programs that include parental and school-based interventions may have the most promise. These studies also demonstrate that the course of this disorder need not be chronic. Some SSRIs have demonstrated partial efficacy, though the studies are small and there is a low rate of full remission. The relative rarity of Selective Mutism, the high rate of psychiatric comorbidity, and the heterogeneity of developmental and genetic history in the children with this disorder make developing and interpreting the efficacy of treatment programs a challenge. However, the field has made strides with regards to this diagnosis in the last few years and one would expect that, with the inclusion of Selective Mutism in DSM-5's Anxiety Disorders section, research will continue to accelerate.

SEPARATION ANXIETY DISORDER

Patients with Separation Anxiety Disorder suffer greatly when physically apart from figures of significant attachment, or when anticipating periods of separation. These patients ruminate on the possibility of any harm that might befall them or their loved ones. This worry leads to the avoidance of activities apart from the presence of the attached figure. Patients can fear sleep, manifesting a kind of vigilance to reduce the possibility of a separation event. Even when they do manage to sleep, patients with Separation Anxiety Disorder often have nightmares regarding separation. Finally, the excessive worry can be associated with somatic symptoms of anxiety. Symptoms in children need only persist for 4 weeks to meet criteria, but persist longer than 6 months with adult-onset Separation Anxiety Disorder.

History and Comparative Nosology

The epidemiologic evidence documenting a high rate of onset occurring after 18 years old has led to a removal of the age-of-onset restriction in the DSM-5. The inclusion of the adult-onset formulation has moved Separation Anxiety Disorder from DSM-IV's "Disorders Usually First Diagnosed in Infancy, Childhood or Adolescence" to DSM-5's "Anxiety Disorders." As noted with Selective Mutism, the prominent symptom of worry ties Separation Anxiety Disorder to the other, classical diagnoses in the Anxiety Disorders section.

Epidemiology

In a recent worldwide study, the lifetime prevalence was 4.8 percent, with 43 percent of these patients experiencing onset after age 18 years. These data support DSM-5's removal of the age-of-onset restriction. The lifetime prevalence of Separation Anxiety Disorder is more common in women than in men, though this ratio is closer to unity when considering only the adult-

onset variant. More than half of patients show a remission of their disorder within the first 10 years after onset. Over the course of the lifetime, Separation Anxiety Disorder is associated with a broad range of comorbid mood and anxiety disorders, including Major Depressive Disorder, Bipolar Disorder, Panic Disorder, Social Anxiety Disorder, and Specific Phobia. The data on the relative rates of the various comorbidities will likely continue to evolve as DSM-5's recognition of the adult-onset variant spurs more recognition and research.

Differential Diagnosis, Etiology, Course, and Treatment

Patients with Separation Anxiety Disorder have emotional experiences and behaviors when out alone in public that overlap with agoraphobia, though the fear in Separation Anxiety Disorder would be the delay in returning to a loved one and the fear in agoraphobia would be secondary to the unavailability of timely help or escape should a frightening or embarrassing event occur. Similarly, patients with Separation Anxiety Disorder might experience panic-like symptoms, or even symptoms meeting full criteria for the Panic Attack Specifier. However, Panic Disorder requires that some of the attacks be context-independent and "out of the blue," whereas panic symptoms in Separation Anxiety Disorder would have their origins in a very specific theme of separation. While fear about the well-being of a loved one could be a theme in GAD, the worries in GAD involve a broader spectrum of themes and situations, rather than the more circumscribed worry about separation in Separation Anxiety Disorder. There are adult-onset courses of Separation Anxiety Disorder that seem, at least initially, to be features of a complicated bereavement process, suggesting an overlap with either PTSD or Complicated Grief.

Since the adult-onset variant is newly codified, there is limited data on etiology, genetic correlates, biomarkers, or course. The lifetime prevalence of traumatic experience is significantly associated with the lifetime prevalence of Separation Anxiety Disorder. There are no controlled intervention studies to guide treatment in adult-onset Separation Anxiety. Given the efficacy of CBT in childhood-onset Separation Anxiety Disorder, CBT is usually the first intervention in the adult-onset variant. However, it is unknown whether these interventions, or which facets of these interventions, developed for the childhood-onset variant would be efficacious for the adult-onset variant.

Studies show benefit for CBT, as well as for programs that include a component of education for parents and a focus on social skills in childhood-onset Separation Anxiety Disorder. CBT strategies often have to be modified based on the age of the child and there are few studies that include patients under the age of 7 years, for example. The data on pharmacotherapy for children with Separation Anxiety is mixed, with SSRIs sometimes employed for CBT nonresponders. While there are

children who appear to be refractory to current treatment programs, the format and targets of CBT interventions for Separation Anxiety Disorder continue to evolve and demonstrate efficacy.

SUBSTANCE/MEDICATION-INDUCED ANXIETY DISORDER AND ANXIETY DISORDER DUE TO ANOTHER MEDICAL CONDITION

Substance/Medication-Induced Anxiety Disorder can manifest as either generalized or panic anxiety. There needs to be a temporal association between the symptoms and either intoxication or withdrawal from the substance or medication. Additionally, the substance or medication in question must be known to be associated with the production of these symptoms. The DSM-5 contains documentation of specific ICD-10 codes for individual substances and medications, including alcohol, caffeine, cannabis, phencyclidine, other hallucinogens, inhalant, opioid, sedative/hypnotic/anxiolytic, amphetamine (or other stimulant), cocaine, and other (or unknown) substances. The range of "other" substances known to produce anxiety symptoms includes a wide range of hormonal, steroid, other endocrine, and neuromodulatory agents. ICD-10 mandates a comment about whether there is a comorbid substance use disorder, as well as a comment on the severity of the substance use disorder. There are also coding specifiers regarding whether the symptoms were the result of intoxication, withdrawal, or medication initiation. Ideally, the anxiety symptoms resolve with the removal of the offending agent or the resolution of the withdrawal. If the anxiety disorder persists for many months after the inciting incident, one may need to consider whether the patient would, at that point, merit the diagnosis of primary anxiety disorder. This point is interesting and part of a longer lasting debate about whether, for example, there is a prolonged withdrawal syndrome from benzodiazepines that can last for months to years. There is no good evidence for toxic or withdrawal effects of any substance lasting for durations longer than the physiologic duration of the withdrawal and, perhaps, the duration required to cognitively and psychologically "re-adjust."

Anxiety Disorder Due to Another Medical Condition can similarly present with anxiety or panic, yet the symptoms are temporally related to the onset or exacerbation of a diagnosed medical condition whose pathophysiology is known to yield anxiety symptoms. As was noted in the Panic Disorder section, there are numerous endocrine, metabolic, oncologic, cardiopulmonary, infectious, and neurologic etiologies that can manifest anxiety or panic symptoms.

FUTURE DIRECTIONS

The development of the nosology for anxiety disorders began with organized, direct observation. Differential responses to various psychotherapy and psychopharmacology interventions have contributed in

part to the current delineation of the disorders, though there is substantial overlap between psychopharmacologic efficacy among the disorders and even the specific CBT packages for distinct disorders is being challenged with transdiagnostic (unified) CBT protocols that have efficacy across disorders. Epidemiologic studies that help identify course, comorbidity, cultural differences, and biases have also helped shape the modern nomenclature. Biomarkers, genetic studies, gene–gene interactions, epigenetics, pharmacogenetics, neuroimaging, and the other developing tools of neuroscience have yet to be incorporated into successive iterations of the DSM but will hopefully contribute in the future to the increasing detailed nosology of the anxiety disorders. It is interesting that, with the culling of the Trauma-Associated and Obsessive–Compulsive spectra, DSM-5's Anxiety Disorders diagnostic grouping appears more neatly bounded. However, the extraordinarily high rate of comorbidity between the component diagnoses of the Anxiety Disorders, Trauma-Associated Disorders, and Obsessive–Compulsive Disorders argues that Anxiety Disorder etiology and treatment should not be viewed in isolation. Moreover, anxiety symptoms, or at least end-symptom presentations that look strikingly similar to facets of the Anxiety Disorders, appear in nearly every diagnostic cluster of Section II of the DSM-5. Thus, while there is diagnostic and treatment research value in the demarcation of the Anxiety Disorders, there is also value in seeing the forest through the trees.

References

American Psychiatric Association. *Desk Reference to the Diagnostic Criteria from DSM-5*. Washington, DC: American Psychiatric Publishing; 2013.

Bergman RL, Piacentini J, McCracken JT. Prevalence and description of selective mutism in a school-based sample. *J Am Acad Child Adolesc Psychiatry*. 2002;41(8):938–946.

Bogels SM, Alden L, Beidel DC, et al. Social anxiety disorder: questions and answers for the DSM-V. *Depress Anxiety*. 2010;27(2):168–189.

Bogels SM, Knappe S, Clark LA. Adult separation anxiety disorder in DSM-5. *Clin Psychol Rev*. 2013;33(5):663–674.

Bruce SE, Yonkers KA, Otto MW, et al. Influence of psychiatric comorbidity on recovery and recurrence in generalized anxiety disorder, social phobia, and panic disorder: a 12-year prospective study. *Am J Psychiatry*. 2005;162(6):1179–1187.

Chartrand H, Sareen J, Toews M, Bolton JM. Suicide attempts versus nonsuicidal self-injury among individuals with anxiety disorders in a nationally representative sample. *Depress Anxiety*. 2012;29(3):172–179.

Chou KL. Childhood sexual abuse and psychiatric disorders in middle-aged and older adults: evidence from the 2007 Adult Psychiatric Morbidity Survey. *J Clin Psychiatry*. 2012;73(11):e1365–e1371.

Donovan MR, Glue P, Kolluri S, Emir B. Comparative efficacy of antidepressants in preventing relapse in anxiety disorders—a meta-analysis. *J Affect Disord*. 2010;123(1–3):9–16.

Dummit ES 3rd, Klein RG, Tancer NK, Asche B, Martin J, Fairbanks JA. Systematic assessment of 50 children with selective mutism. *J Am Acad Child Adolesc Psychiatry*. 1997;36(5):653–660.

Eaton WW, Kessler RC, Wittchen HU, Magee WJ. Panic and panic disorder in the United States. *Am J Psychiatry*. 1994;151(3):413–420.

Ehrenreich JT, Santucci LC, Weiner CL. Separation anxiety disorder in youth: phenomenology, assessment, and treatment. *Psicol Conductual*. 2008;16(3):389–412.

Goodwin RD, Roy-Byrne P. Panic and suicidal ideation and suicide attempts: results from the National Comorbidity Survey. *Depress Anxiety*. 2006;23(3):124–132.

Hudson JL, Rapee RM, Lyneham HJ, McLellan LF, Wuthrich VM, Schniering CA. Comparing outcomes for children with different anxiety disorders following cognitive behavioural therapy. *Behav Res Ther*. 2015;72:30–37.

Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):593–602.

Kessler RC, Chiu WT, Jin R, Ruscio AM, Shear K, Walters EE. The epidemiology of panic attacks, panic disorder, and agoraphobia in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2006;63(4):415–424.

Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry*. 1994;51(1):8–19.

Klauke B, Deckert J, Reif A, Pauli P, Domschke K. Life events in panic disorder-an update on candidate stressors. *Depress Anxiety*. 2010;27(8):716–730.

Kristensen H. Selective mutism and comorbidity with developmental disorder/delay, anxiety disorder, and elimination disorder. *J Am Acad Child Adolesc Psychiatry*. 2000;39(2):249–256.

Mohatt J, Bennett SM, Walkup JT. Treatment of separation, generalized, and social anxiety disorders in youths. *Am J Psychiatry*. 2014;171(7):741–748.

Muris P, Ollendick TH. Children who are anxious in silence: a review on selective mutism, the new anxiety disorder in DSM-5. *Clin Child Fam Psychol Rev*. 2015;18(2):151–169.

Oerbeck B, Stein MB, Pripp AH, Kristensen H. Selective mutism: follow-up study 1 year after end of treatment. *Eur Child Adolesc Psychiatry*. 2015;24(7):757–766.

Remschmidt H, Poller M, Herpertz-Dahlmann B, Hennighausen K, Gutenbrunner C. A follow-up study of 45 patients with elective mutism. *Eur Arch Psychiatry ClinNeurosci*. 2001;251(6):284–296.

Rochester J, Baldwin DS. Adult separation anxiety disorder: accepted but little understood. *Hum Psychopharmacol*. 2015;30(1):1–3.

Roy-Byrne PP, Cowley DS. Course and outcome in panic disorder: a review of recent follow-up studies. *Anxiety*. 1994;1(4):151–160.

Seeley-Wait E, Abbott MJ, Rapee RM. Psychometric properties of the mini-social phobia inventory. Prim Care Companion. *J Clin Psychiatry*. 2009;11(5):231–236.

Silove D, Alonso J, Bromet E, et al. Pediatric-onset and adult-onset separation anxiety disorder across countries in the World Mental Health Survey. *Am J Psychiatry*. 2015;172(7):647–656.

Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092–1097.

Steinhausen HC, Juzi C. Elective mutism: an analysis of 100 cases. *J Am Acad Child Adolesc Psychiatry*. 1996;35(5):606–614.

Wittchen HU, Gloster AT, Beesdo-Baum K, Fava GA, Craske MG. Agoraphobia: a review of the diagnostic classificatory position and criteria. *Depress Anxiety*. 2010;27(2):113–133.

Wittchen HU, Zhao S, Kessler RC, Eaton WW. DSM-III-R generalized anxiety disorder in the National Comorbidity Survey. *Arch Gen Psychiatry*. 1994;51(5):355–364.

▲ 14.3 Epidemiology of Anxiety Disorders

KATHLEEN RIES MERIKANGAS, PH.D., AND JOHN DAVID EUN, B.A.

Population-based studies across the world have shown that anxiety disorders are the most common class of mental disorders. Epidemiology has made a valuable contribution to the current understanding of anxiety disorders by documenting the prevalence, correlates, and natural history of anxiety disorders, and highlighting the gap between untreated anxiety disorders in the general population compared to cases seen in clinical settings. During the past few years, more valid estimates of the global burden of anxiety disorders have also become available.

The aims of this chapter are: (1) to describe the general methods of epidemiology that have contributed to the current understanding of anxiety disorders; (2) to summarize the prevalence and correlates of DSM-IV anxiety disorders in adults and children from community samples; and (3) to describe the individual and societal impact of anxiety disorders.

BACKGROUND: PSYCHIATRIC EPIDEMIOLOGY

Epidemiology is the study of the distribution and determinants of diseases in human populations. Epidemiologic studies are concerned with the extent and types of illnesses in groups of people and with the factors that influence their distribution. Epidemiologists investigate the interactions that may occur among the host, agent, and environment to produce a disease state. Although the tools of epidemiology were developed to identify causes of infectious diseases, their application to chronic diseases such as cancer, heart disease, and mental disorders has also enhanced the current understanding of their multifactorial etiology.

The ultimate goal of epidemiologic studies is to identify the *etiology* of a disease in order to prevent or intervene in the progression of the disorder. To achieve this goal, epidemiologic studies generally proceed from studies that specify the prevalence and distribution of a disease within a population by person, place, and time (*descriptive* epidemiology) to more focused studies of the determinants of disease in specific groups (*analytic*

epidemiology). Descriptive epidemiologic studies are important in specifying the rates and distribution of disorders in the general population. Identification of differential distribution by sex, age, ethnicity, geographic site or by exposure to particular risk factors provides clues that may be tested systematically. Case-control studies generally proceed from retrospective designs and define cases and controls by the presence or absence of a disease in order to identify the disease's potential associations with a particular risk factor or set of risk factors. Cohort studies identify exposure to a putative risk factor in a group of people at baseline and then follow them retrospectively or prospectively to examine differential incidence of the disease over time.

Community study data can be examined to identify biases that may exist in treated populations and to construct case-registries from which persons may serve as probands for analytic epidemiologic studies. Such attention to sampling issues is a major contribution of the epidemiologic approach, as individuals identified in clinical settings often constitute the tip of the iceberg of the disease and may not be representative of the general population of similarly affected individuals with respect to demographic, social, or clinical characteristics.

Application of Epidemiology to Psychiatry

The traditional contributions of the application of the tools of epidemiology to psychiatry include: (1) methodologic developments including the introduction of structured and semistructured diagnostic interviews; (2) estimation of the prevalence and correlates of mental disorders across the world; and (3) identification of risk factors for the development of mental disorders.

Community studies of mental disorders have also highlighted the need to expand the categorical diagnostic criteria for mental disorders to collect information on the spectra of expression of their underlying domains. For example, findings from the National Comorbidity Survey-Replication (NCS-R) demonstrated the validity of the spectrum concept of bipolar disorder that had been proposed by clinicians for more than a decade. Prospective follow-up of community samples of youth has also demonstrated the heterogeneity of different patterns of developmental manifestations of anxiety subtypes and their underlying dimensions of fear and anxiety. Therefore, the imposition of arbitrary thresholds in the contemporary categorical diagnostic system may lead to inaccurate classification of these conditions. Likewise, the tendency for individuals to meet criteria for more than one condition both at the same time and across their lifetimes, or comorbidity, has been consistently demonstrated in community samples, highlighting the lack of distinct boundaries between classes of mental disorders, particularly anxiety and mood disorders.

One of the most important contributions of epidemiology to the current

understanding of mental disorders is the identification of the large number of individuals who suffer from one or more of these conditions but have not received clinical intervention yet. There is a growing gap between cases seen in specialty mental health settings and people with anxiety in the general community. As expected, clinical samples have a disproportionate representation of severe cases, with multiple comorbid conditions, and substantially greater impairment than those who do not receive mental health treatment. Below, updated estimates of the prevalence of anxiety disorders are provided.

MAGNITUDE OF ANXIETY DISORDERS

Adults

Anxiety disorders have emerged as the most prevalent mental disorders in both adult and child populations. The accumulating literature suggests that the prevalence of anxiety disorders in youth is similar, if not higher, to that reported in adults, thereby indicating the importance of a life course approach to the study of anxiety. During the past decade, the results of several large-scale studies focusing on cross-national and cross-cultural samples have emerged. The largest global collaboration to date is the aforementioned WHO World Mental Health 2000 Initiative (WMH2000), which encompasses a collection of nationally or regionally representative surveys on the prevalence, risk factors, correlates, and service patterns for mental disorders from over 25 countries, representing all regions of the world, with a total sample size exceeding 130,000 individuals. There has also been progress in descriptive epidemiology in countries outside of the United States and Europe that have been traditionally the sole source of population-based data on mental disorders. Below, an update of the prevalence of both the aggregate rates of anxiety as well as those of specific anxiety disorders is presented.

The new *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria removed Posttraumatic Stress Disorder and Obsessive–Compulsive Disorder from the Anxiety Disorder Section, so the review in this chapter is limited to the major subgroups of anxiety disorders including in DSM-5* Anxiety Disorders. Separation anxiety disorder and selective mutism are now included as anxiety disorders, and the former's diagnostic criteria no longer specify the age of onset. Other major changes include: the separate diagnoses of panic disorder and agoraphobia, removal of the requirement that individuals over age 18 years must recognize that their anxiety is excessive or unreasonable, and the extension of the 6-month duration criterion. However, because the *DSM-5* was introduced during the past year, prevalence rates based on the new criteria are not yet available. Several studies suggest that modifications in the anxiety disorders section of the *DSM-5* will not have substantial impact

on the prevalence rates of specific subtypes of anxiety.

Table 14.3–1 presents an overview of recent national and international community studies of *DSM-IV* anxiety disorders in adults.

Generalized Anxiety Disorder. GAD has also been extensively examined in samples from several continents. The 12-month prevalence rate of GAD falls between 2.1 and 3.1 percent in two large-scale studies in the United States, the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) and the NCS-R, respectively. The rates in other areas of the world ranged from 0 percent in Nigeria to 2.6 percent in Germany. The lifetime prevalence rates of GAD were even more variable, with the lowest rate reported in Nigeria (0.1 percent) and the highest in Italy (6.9 percent), with a median of approximately 2.3 percent.

Panic Disorder. The prevalence of *DSM-IV* panic disorder has been examined across the globe, and the rates remain fairly consistent across studies. Most studies yield 12-month prevalence rates between 0.2 and 1.1 percent. The highest rate was reported in the NCS-R Study in the United States (2.7 percent), whereas the lowest 12-month prevalence was reported in Nigeria (0.1 percent). The lifetime prevalence rates of panic disorder range from 0.2 percent in Nigeria to a high of 5.1 percent in the United States, with most studies falling between 1 and 2 percent.

Separation Anxiety Disorder. The *DSM-5* criteria no longer include an age-at-onset restriction because a substantial number of adults report onsets after age 18 years. However, there is limited data on separation anxiety from population surveys of adults. Recent investigation of adult versus childhood onset of separation anxiety disorders was completed in the World Mental Health Surveys. The 12-month prevalence rate averaged 1 percent across the 18 countries in the WMH Survey with the lowest prevalence in Nigeria and Italy, where no cases were identified, and the highest in Colombia that reported a 12-month rate of 2.7 percent. The lifetime prevalence rate averaged 4.8 percent, with 43.1 percent of lifetime onsets occurring after age 18 years, and a range of rates from 0.2 (Nigeria) to 9.8 percent (Colombia). However, the variability of estimates was far greater than that in the other WMH surveys, presumably reflecting methodologic differences in the surveys.

Table 14.3–1.
Lifetime and 12-Month Prevalence Rates of DSM-IV Anxiety Disorders in International Community Studies of Adults

	GAD	Panic Disorder	Separation Anxiety	Agoraphobia	Social Phobia	Specific Phobia
	12	12			12	12

Study	mo	LT	mo	LT	12 mo	LT	12 mo	LT	mo	LT	mo	LT	
U.S. (NESARC)[a]	2.1	4.1	2.1	5.1	—	—	0.1	0.2	2.8	5.0	7.1	9.4	1
Iran[b]	—	1.3	—	1.3	—	—	—	0.7	—	0.8	—	—	
Australia (NSMHWB)[c]	2.6	—	1.1	—	—	—	0.5	—	1.3	—	—	—	
Germany (GHS-MHS)[d]	1.5	—	2.3	—	—	—	—[4]	—	—[4]	—	—[4]	—	1
Italy, Sesto Fiorentino[e]	—	6.9	—	6.9	—	—	—	0.4	—	3.7	—	1.5	
Korea (KECA-R)[f]	0.8	1.6	0.1	1.6	—	—	0.2	0.2	0.4	0.5	3.4	3.8	
WHO WMH Surveys[g]													
US	3.1	5.7	2.7	4.7	1.9	9.2	0.8	1.4	6.8	12.1	8.7	12.5	1
Mexico	—	0.9	—	1.0	0.9	4.5	—	1.0	—	2.9	—	7.0	
Belgium	—	1.9	—	16	0.1	1.4	—	—	—	2.0	—	6.8	
Italy	0.5	1.9	0.6	16	0.0	1.5	0.4	1.2	1.0	2.1	2.7	5.7	
Ukraine	1.2	1.9	1.3	1.9	—	—	0.2	0.3	1.5	2.6	—	—	
Lebanon	1.3	—	0.2	—	1.9	6.9	0.3	—	1.1	—	8.2	—	1
Nigeria	0.0	0.1	0.1	0.2	0.0	0.2	0.2	0.4	0.3	0.3	3.5	5.4	
New Zealand	—	6.0	—	2.7	—	—	—	1.2	—	9.4	—	10.8	
Japan	1.2	—	0.5	—	—	—	0.3	—	0.8	—	2.7	—	
China	0.8	—	0.2	0.4	0.2	1.3	0.0	0.0	0.2	0.5	1.9	2.6	
South Africa	1.4	—	0.8	—	—	—	4.8	—	1.9	—	—	—	

[a]From Hasin DS, Grant BF. The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) Waves 1 and 2: review and summary of findings. *Soc Psychiatry Psychiatr Epidemiol*. 2015;50(11):1609–1640.

[b]Mohammadi MR, Davidian H, Noorbala AA, et al. An epidemiological survey of psychiatric disorders in Iran. *Clin Pract Epidemol Ment Health*. 2005;26;1:16.

[c]Oakley-Browne et al., 2006.

[d]Jacobi F, Wittchen HU, Holting C, et al. Prevalence, co-morbidity and correlates of mental disorders in the general population: results from the German Health Interview and Examination Survey (GHS). *Psychol Med*. 2004;34(4):597–611.

[e]Faravelli C, Abrardi L, Bartolozzi D, et al. The Sesto Fiorentino study: background, methods and preliminary results. Lifetime prevalence of psychiatric disorders in an Italian community sample using clinical interviewers. *Psychother Psychosom*. 2004;73(4):216–225.

[f]Cho MJ, Chang SM, Lee YM, et al. Prevalence of DSM-IV major mental disorders among Korean adults: A 2006 National Epidemiologic Survey (KECA-R). *Asian J Psychiatr*. 2010;3(1):26–30.

[g]The WHO World Mental Health Survey Consortium. Prevalence, Severity, and Unmet Need for Treatment of Mental Disorders in the World Health Organization World Mental Health Surveys. *JAMA*. 2004;291:2581–2590.

[h]Not including PTSD.

Agoraphobia. Apart from an elevated 12-month prevalence of

agoraphobia reported in South Africa (4.8 percent), the rates of agoraphobia were remarkably consistent across regions. Nine separate international studies reported estimates ranging from 0 (China) to 0.8 percent (United States). The lifetime prevalence estimates were similarly consistent, with the lifetime prevalence of *DSM-IV* agoraphobia ranging anywhere from 0 (China) to 1.2 percent of the population (Italy and New Zealand), and a median of 0.9 percent.

Social Anxiety Disorder. The 12-month prevalence of social phobia varied across regions, with the highest rates reported in the United States in the NCS-R study (6.8 percent) and the lowest appearing in China (0.2 percent), and a median of 4 to 5 percent. However, it cannot be determined whether differences in these rates are the result of methodologic or translation inconsistencies or true cross-cultural differences in defining social fears. Even within the United States, the 12-month prevalence dramatically differed, with the NCS-R study reporting 6.8 percent in comparison to the 2.8 percent prevalence reported from the NESARC study. The estimates of the lifetime prevalence of *DSM-IV* social phobia were also inconsistent across studies, with a range spanning from 0.5 percent in East Asian countries to a high of 9.4 percent in New Zealand.

Specific Phobia. Similar to the estimates reported for SAD, the 12-month and lifetime prevalence rates of *DSM-IV* specific phobia varied across studies. The 12-month prevalence estimates ranged from 1.9 (China) to 8.7 percent (US). Estimates for lifetime specific phobia were lowest in Italy (1.5 percent) and highest in New Zealand (10.8 percent), with a median of 6 percent. As with other disorders, reasons for these inconsistencies need to be examined further in order to distinguish between true differences in prevalence and those attributable to methodologic differences between studies.

Children and Adolescents

During the past decade, the first results of the prevalence of anxiety disorders in nationally representative samples of the United States have become available. The flagship study of the WMH initiative, the NCS-R, was extended to adolescents aged 13 to 18 years using the same diagnostic methodology. Table 14.3–2 compares the 12-month and lifetime prevalence rates in adults and adolescents in these studies. The high aggregate rates of anxiety in adults (28.8 percent) and youth (31.9 percent) demonstrate that anxiety disorder is the most common mental disorder in both adults and adolescents compared to mood disorder that affects approximately 15 to 20 percent of both adults and youth. The finding that the prevalence rates in youth approximate those of adults suggests that the onset of most of the anxiety disorders occurs in childhood and adolescence. In fact, the rates of

aggregate anxiety were actually higher in adolescents than in adults in the NCS-R for both 12-month (24.9 vs. 18.1 percent) and lifetime (31.9 vs. 28.8 percent) prevalence.

Table 14.3–2.
12-Month and Lifetime Rates of DSM-IV Anxiety Disorders in the National Comorbidity Survey Replication (Adults) and Supplement (Adolescents)

			NCS-A (Adolescents)						
	NCS-R (Adults)		Lifetime by Sex (%)		Lifetime by Age (%)				
Anxiety Disorder Subtype	12-mo (%)	Lifetime (%)	Female	Male	13–14 y	15–16 y	17–18 y	12-mo (%)	Lifetime (%)
Agoraphobia	0.8	1.4	3.4	1.4	2.5	2.5	2.0	1.8	2.4
Generalized anxiety disorder	3.1	5.7	3.0	1.5	1.0	2.8	3.0	1.1	2.2
Panic disorder	2.7	4.7	2.6	2.0	1.8	2.3	3.3	1.9	2.3
Separation anxiety disorder	6.6	5.2	9.0	6.3	7.8	8.0	6.7	1.6	7.6
Social anxiety disorder	6.8	12.1	11.2	7.0	7.7	9.7	10.1	8.2	9.1
Specific phobia	8.7	12.5	22.1	16.7	21.6	18.3	17.7	15.8	19.3
Any anxiety disorder	18.1	28.8	38.0	26.1	31.4	32.1	32.3	24.9	31.9

This is primarily attributable to adolescents reporting substantially higher rates of phobic disorders and separation anxiety than the adults. Prospective research described below has shown that many anxiety disorders in childhood subside as they progress to adulthood. By contrast, the rates of GAD (5.7 percent), panic disorder (4.7 percent), and SAD (12.1 percent) in adults were greater than those in adolescents, with lifetime rates of 2.2, 2.3, and 9.1 percent, respectively. Similar trends were generally observed for 12-month rates of these subtypes of anxiety disorders.

Prevalence rates of these conditions in adolescents with severe impairment were also evaluated in the NCS-A. As expected, lifetime prevalence rates of disorders with severe impairment were substantially lower than lifetime DSM-IV disorders: GAD – 0.9 percent; social phobia – 1.3 percent, and 0.6 percent of those with separation anxiety and specific phobias. This dramatic reduction in prevalence of youth with severe impairment was also found in the Tracking Adolescents' Individual Lives Survey (TRAILS) study that yielded a prevalence rate of aggregate anxiety disorders of 28.0 percent versus 5.2 percent for those with severe impairment.

Estimates of prevalence rates of anxiety disorders in children have been much more variable than those of adults because of substantial differences

in study methods and developmental manifestations of anxiety and fear that may influence prevalence estimates. The chief differences include: the age and sex composition of the samples, methods for assessing anxiety disorders (i.e., self-report vs. diagnostic interview); informant source (i.e., parent, child, teacher); the number and type of anxiety disorders included in the study; and the definition of diagnostic criteria for each disorder. For example, besides the NCS-A study, the only other study of mental health prevalence in a nationally representative sample of youth in the United States, the National Health and Nutrition Examine Survey (NHANES), only collected the diagnostic modules for Separation Anxiety Disorder and Panic Disorder. Moreover, a much more highly structured interview, the Diagnostic Interview Survey for Children (DISC) than the Composite International Diagnostic Interview (CIDI) employed in the NCS-A, was used in the NHANES. Thus, the prevalence rates of these conditions were substantially lower (i.e., 0.4 percent for panic disorder and 0.3 percent for GAD) than the 12-month rates of these found in the NCS-A for these disorders (i.e., 1.9 and 1.1 percent, respectively). However, the ranges of estimates across all cross-sectional studies that applied the *DSM-IV* criteria for the same disorders with similar interviews are fairly comparable, particularly for current and 12-month disorders.

During the past decade, prevalence rates of anxiety disorders in youth are also available from studies that have expanded beyond Europe, the United States, and Australia to Asia and the Middle East. Aggregate lifetime rates of anxiety disorders from cross-sectional studies of youth average about 15 to 20 percent, with phobic disorders as the most common subgroup, averaging about 15 percent, and progressively lower rates of social anxiety of 10 to 12 percent, generalized anxiety of 5 to 8 percent, agoraphobia of 4 percent, and panic of 2 to 3 percent.

There are now data from several prospective community studies that have followed community samples of children and young adults for more than a decade. These studies provide information on the cumulative incidence and stability of anxiety disorders in early and middle adulthood. Table 14.3–3 shows the prevalence rates of anxiety disorders from three key studies that follow different age groups across the life span. The Zurich Cohort Study of a community sample of 19- to 20-year-old adults from Zurich, Switzerland provides unique information on the prevalence and continuity of anxiety disorders across 30 years that tracks a single age cohort from early to middle adulthood. The cumulative prevalence rate of anxiety disorders across the lifetime was approximately 30 percent. Similar to cross-sectional studies, specific phobia was the most common disorder (26.9 percent), followed by GAD (20.8 percent), social anxiety (12.6 percent), agoraphobia (6.8 percent), and panic disorder (6.1 percent). Classification of the spectrum of disorders in the Zurich Cohort Study has further shown that there are a large proportion of people from the

community whose symptoms fall just below the diagnostic thresholds for mood and anxiety disorders, but nevertheless exhibit impairment and/or distress. As expected, inclusion of subthreshold manifestations of anxiety disorders leads to further elevation of the rates of these conditions in the general population.

The Early Developmental Stages of Psychopathology Study (EDSP), a prospective longitudinal study of a community sample of youth aged 14 to 24 years followed over 10 years with three interviews. They derived the following estimates of the cumulative incidence rates of anxiety disorders to age 33 years: aggregate anxiety disorders – 31.3 percent; specific phobia – 7 percent; social phobia – 6 percent; GAD – 4.3 percent; and panic – 3.3 percent. Projected rates to age 33 years were almost double those from the initial interview, demonstrating that the peak period for onset of anxiety disorders is late adolescence to early adulthood.

Table 14.3–3.
Lifetime Rates of DSM-IV Anxiety Disorders in Prospective Community Studies of Youth

Study	Age (y)	GAD (%)	Panic (%)	Separation Anxiety (%)	Agoraphobia (%)	Social Anxiety (%)	Specific Phobia (%)	Any (%)
Zurich Cohort Study[a]	18–50	20.8	6.1		6.8	12.6	26.9	30.0
Early Dev Study of Psychopathology[b]	14–33	4.3	3.3		6.0	7.0	14.7	31.3
The Great Smoky Mountain Study[c]	9–26	9.7	4.8	5.0	6.1	4.2	2.2	22.7

[a]Angst J, Paksarian D, Cui L, et al. The epidemiology of common mental disorders from age 20 to 50: results from the prospective Zurich cohort Study. *Epidemiol Psychiatr Sci*. 2015;24:1–9.

[b]Beesdo-Baum K, Knappe S, Asselmann E, et al. The Early Developmental Stages of Psychopathology (EDSP) study: a 20-year review of methods and findings. *Soc Psychiatry Psychiatr Epidemiol*. 2015;50(6):851–866.

[c]Copeland WE, Angold A, Shanahan L, Costello EJ. Longitudinal patterns of anxiety from childhood to adulthood: the Great Smoky Mountains Study. *J Am Acad Child Adolesc Psychiatry*. 2014;53(1):21–33.

The Great Smoky Mountain Study (GSMS) assessed 1,420 participants in the southeastern United States up to 11 times between ages 9 and 26 years. The cumulative incidence rates of anxiety across the study were 22.7 percent for all anxiety, and 9.7 percent for GAD, 6.1 percent for agoraphobia, 5 percent for separation anxiety disorder, 4.8 percent for panic disorder, 4.2 percent for social phobia, and 2.2 percent for specific phobia.

Although there is some comparability in the findings across the three

studies, variability in the samples and methods that make it difficult to aggregate findings on the prevalence of anxiety disorders. Nevertheless, the aggregate rates of anxiety disorders of approximately 30 percent confirm cross-sectional studies of adults that show that anxiety disorders are the most common class of mental disorders. Likewise, the relative rank ordering of prevalence rates of specific disorders is generally consistent across studies. However, all three studies show that there is substantial fluctuation across anxiety disorder subtypes, and across diagnostic thresholds over time. In general, phobic disorders are far more stable than panic and GAD that tend to resemble mood disorders in their waxing and waning across the life span. The correlates and risk factors for anxiety disorders that have been implicated in community studies of anxiety and other mental disorders are described below.

CORRELATES AND RISK FACTORS

Sex Differences

The results of community studies in adults reveal that compared to men, women have greater rates of almost all of the anxiety disorders. Women have an approximately two-fold elevation in lifetime rates of panic disorder, GAD, agoraphobia, and specific phobia than males in nearly all of the studies. By contrast, the sex ratio for social anxiety is approximately equal. Studies of youth also report that girls tend to have more of all subtypes of anxiety disorders, irrespective of the age composition of the sample. Epidemiologic surveys of adults reveal that the female preponderance of anxiety disorders is present across all stages of life but is most pronounced throughout early and mid-adulthood.

Age of Onset

Retrospective reports of adults and adolescents with anxiety disorders suggest that their onset generally occurs in childhood or adolescence. For example, results of the NCS-R reveal that anxiety disorders have the earliest age of onset of all of the major classes of mental and behavioral disorders, with a median onset by the age of 12 years. The NCS-A found an even earlier median onset at age 6 years. The onset is far earlier than the onset of mood disorders or substance use disorders, and comparable to that of impulse control disorders. The prospective GSMS found a U-shaped age prevalence curve for any anxiety disorder in which a sharp decrease in mid-childhood is followed by increasing levels from early adolescence to young adulthood.

Although there is substantial variation across studies, the results of prospective community-based research reveal differential peak periods of onset of specific subtypes of anxiety: separation anxiety and specific phobias in middle childhood; SAD in middle adolescence; agoraphobia and

panic disorder from late adolescence to young adulthood; and GAD in young adulthood.

Several follow-up studies of children and adolescents have shown that there is stability of both SAD and simple phobia, whereas GAD and panic symptoms and disorders are more variable, and tend to overlap with episodes of depression.

Social Class and Ethnicity

Community studies have consistently found that rates of anxiety disorders in general are greater among those at lower levels of socioeconomic status and education levels, but the associations might be more complex due to the significant interactions among the sociodemographic characteristics. For example, anxiety disorders were negatively associated with income and education levels in the NCS. There was almost a two-fold difference between rates of anxiety disorders in individuals in the highest income bracket over the lowest, and those who completed over 16 years of school versus less than 11. However, the association with education was not found in the NCS-R, and the association with income was not found in the NCS-A.

The research on how race/ethnicity and social class impacts the prevalence of anxiety disorders is inconclusive. While several community studies have yielded greater rates of anxiety disorders, particularly phobic disorders, among African Americans, others have shown the opposite. For example, non-Hispanic Blacks and Hispanics had lower rates of GAD and SAD, and non-Hispanic Blacks had lower rates of panic disorder than their non-Hispanic white counterparts. In contrast, non-Hispanic black adolescents had higher rates of anxiety disorders than non-Hispanic white adolescents in the NCS-A. The reasons for ethnic and social class differences have not yet been evaluated systematically; however, both methodologic factors and differences in exposure to stressors have been advanced as possible explanations. As described below, anxiety disorders may actually lead to lower educational attainment, so this may also be associated with ethnic differences in rates of anxiety disorders.

Patterns of Comorbidity

Subtypes of anxiety disorders tend to co-occur both cross-sectionally and across the lifetime. Anxiety disorders are associated with all the other major classes of disorders including mood disorders, disruptive behaviors, eating disorders, and substance use disorders. The co-occurrence of anxiety disorders and mood disorders is so common that there is emerging evidence that anxiety disorders may be part of the developmental sequence in which anxiety is expressed early in life followed by depression in adulthood.

Both family and twin studies have been used to examine sources of

overlap within the anxiety disorders, and between the anxiety disorders and other syndromes including depression, eating disorders, and substance abuse. Family studies have demonstrated the independence of familial aggregation of panic and phobias. With respect to comorbidity, panic disorder, GAD, and depression have been shown to share common familial and genetic liability, whereas there is substantial evidence for the independent etiology of anxiety disorders and substance use disorders. Similar results have emerged from studies of symptoms of anxiety and depression in both youth and adults in which anxiety and depression were found to result from a common genetic diathesis.

Familial and Genetic Factors

The familial aggregation of all the major subtypes of anxiety disorders has been well established, and the substantial role of genetic factors underlying familial transmission has been well documented, as described in Chapter 14.4 of this section. However, the relatively moderate magnitude of heritability also strongly implicates environmental etiologic factors. Components of anxiety that may be inherited include somatic manifestations of anxiety such as physiologic responses, which include pulse, respiration rate, and galvanic skin response. There is also substantial evidence for shared genetic risk factors for mood disorders and anxiety disorders that may manifest differentially across developmental stages.

Increased rates of anxiety symptoms and disorders have also been found among offspring of parents with anxiety disorders which suggest that there may be underlying psychological or biologic vulnerability factors for anxiety disorders in general which may already manifest in children prior to puberty. Children at risk for anxiety disorders throughout life are characterized by behavioral inhibition, autonomic reactivity, somatic symptoms, social fears, enhanced startle reflex, and respiratory sensitivity. These findings suggest that anxiety and fear states are highly heterogeneous and that future studies need to investigate the extent to which the components of anxiety result from common versus unique genetic factors and the role of environmental factors, either biologic or social, in either potentiating or suppressing their expression.

Temperament/Personality

The personality traits of neuroticism and negative affectivity have been associated with anxiety disorders in adults in both cross-sectional and prospective studies. However, these traits are also associated with mood disorders so there is a lack of specificity with respect to anxiety disorders. One of the earliest indicators of vulnerability to the development of anxiety is behavioral inhibition, characterized by increased physiologic reactivity or behavioral withdrawal in the face of novel stimuli or challenging situations.

Behavioral inhibition may be a manifestation of a biologic predisposition characterized by both overt behavioral and physiologic measures. In adolescents and adults with behavioral inhibition, studies have found alterations in neural circuits governing novelty detection, fear processing, anticipation, reward processing, and attention as well as alternations in amygdala, nucleus accumbens, caudate, and prefrontal cortex.

Anxiety sensitivity is another potential sensitive and specific trait marker for the development of anxiety disorders. Anxiety sensitivity is characterized by beliefs that anxiety sensations are indicative of harmful physiologic, psychological, or social consequences (e.g., fainting or an impending heart attack). Anxiety sensitivity has been shown to precede the development of anxiety symptoms and disorders and to have specificity in predicting the development of anxiety as opposed to depression.

Medical Symptoms/Disorders

Community samples have strongly confirmed the high magnitude of comorbidity of anxiety disorders with several medical conditions that had been described in clinical samples. There is a growing body of evidence for specific patterns of associations between anxiety disorders and a range of medical disorders including respiratory conditions, cardiovascular diseases, gastrointestinal disorders, metabolic diseases, and musculoskeletal disorders. For example, in the Australian National Survey of Health and Wellbeing, there was a two-fold greater rates of physical conditions among those with anxiety disorders compared to those without anxiety. Similar to other mental disorders, comorbid anxiety with physical disorders are associated with substantially greater disability than anxiety disorders alone.

Investigation of comorbidity of physical disorders and anxiety is complicated by the role of physical symptoms as a core feature of panic disorder and GAD. There is now evidence that anxiety disorders may represent the initial manifestations of several physical conditions, and there are also numerous physical conditions that lead to anxiety-like symptoms. In adults, anxiety disorders, particularly GAD and panic, are associated with cardiovascular risk factors and diseases. For example, analyses of data from electronic medical records and direct interviews regarding medical and mental disorders in the large Philadelphia Neurodevelopmental Cohort (PNC) revealed that asthma was the only medical condition specifically associated with anxiety disorders in childhood and adolescence. This link has been confirmed in a population-based birth cohort study in Australia that found that youth with more severe and persistent asthma at age 5 years were more likely to have anxiety problems from ages 5 to 17 years. Therefore, some of the other comorbid physical disorders associated with anxiety may not emerge until later in life. Similar to adults, rates of anxiety disorder in the PNC increased

with the severity of aggregate physical conditions, suggesting that physical–mental comorbidity is associated with greater impairment of both conditions. Prospective documentation of the evolution of anxiety disorders and physical conditions will be critical to the understanding of the explanations for comorbidity.

Exposure

The role of life experiences in the etiology of anxiety states, particularly phobias and panic disorder, has been widely studied. Life events have often been designated a causal role in the onset of phobias, which are linked inherently to particular events or objects. More broadly, life experiences that are likely to trigger specific future problems with safety and security are perceived to predict the onset of anxiety disorders. While it is likely that exposure may trigger the fear response, there is compelling evidence for the evolutionary tendency toward fear and avoidance of potential predators. Therefore, phobic states may first become apparent without exposure. The ability to extinguish such fears may be protective against the development of clinically significant phobias. Recent research suggests such life experiences that also trigger immediate distress (e.g., parental divorce or unemployment) might also more broadly precipitate the onset of comorbid depression and anxiety. In evaluating the evidence on the causal role of life experiences, it is critical to consider the subtypes of anxiety disorders separately. Moreover, stressful life events may interact with other risk factors such as family history of depression in precipitating episodes of panic.

IMPACT OF ANXIETY DISORDERS

Although the individual impact of anxiety disorders has been well recognized, early estimates of the global burden of disease tended to neglect anxiety disorders as a significance source of lifetime disability. Earlier estimates were based on a limited number of anxiety disorders, particularly obsessive–compulsive disorder and PTSD, that are no longer included in the class of anxiety disorders in the *DSM-5*. Recent studies of the global burden of disease have employed the concept of Disability Adjusted Life Years (DALYs) as an index of the impact of particular diseases on mortality and disability. Estimates of both person years lived with disability (YLDs) and years of lost life (YLL) as a measure of the societal impact of anxiety disorders. The most recent global estimate of YLDs is 390/100,000 population for all anxiety disorders, and the total DALY for all anxiety disorders is about 700 for females and 400 for males. The peak disability occurs in the young adult age group from 15 to 49 years. Anxiety disorders are the sixth leading cause of YLDs after low back pain, major depressive disorder, other musculoskeletal diseases, neck pain and

falls, with similar patterns across country-specific income levels. These findings highlight the significant public health impact of anxiety disorders, particularly when coupled with other comorbid conditions.

There is also new evidence regarding the increased risk of mortality associated specifically with anxiety disorders. Meta-analytic studies have shown that the population-attributable risk percent of mortality associated with anxiety disorders is only slightly lower than that attributable to depression, with 2.74 million deaths per year associated with anxiety disorders.

Evidence for the impact of anxiety on the life course of individuals who suffer from one or more of these conditions has emerged from prospective follow-up studies of children with anxiety into adulthood. The most important consequences are comorbid mental and physical conditions, substance use disorders, and, in a small proportion, suicide. The 15-year follow-up of the GSMS revealed that youth with anxiety disorders during childhood had substantial impairment in one or more of the domains assessed including the health, financial, and interpersonal functioning in adulthood. Youth with GAD tended to have poor functioning across all domains, whereas those with social phobia only displayed impairment in the interpersonal domain, and those with separation anxiety had more health problems than other youth. These studies illustrate that the burden of anxiety disorders is attributable to both their childhood and adolescent onset and to their continuity across the life span. In consideration of their association with the subsequent development of substance use disorders, behavior and mood disorders, these findings indicate the critical need for intervention at early stages of these conditions.

FUTURE DIRECTIONS

This chapter provides an update on the magnitude of anxiety disorders in adults and children in community surveys across the world. Data are expanding beyond North America and Europe to Asia, South America, Africa, and the Middle East that provide a more representative international depiction of the magnitude of anxiety. The most striking finding is the high prevalence of anxiety disorders and the consistency in correlates of anxiety disorders across geographic settings and methods of assessment. Recent estimates of global disability associated with anxiety disorders documents the substantial individual and societal impact of anxiety disorders. The impact of anxiety disorders on lower educational and occupational attainment, missed days in role, and comorbid associations with other mental and physical disorders, as well as their significant contribution to increased mortality, suggests that anxiety disorders should be a major public health priority. The strikingly low rates of mental health treatment, in light of the highly efficacious treatments available for anxiety disorders, should also elevate anxiety as a priority for

mental health outreach. The growing evidence for the emergence of anxiety in childhood based on both retrospective studies of adults and children and prospective studies of youth also suggests that preventive efforts should focus on youth to avoid progression and consequences of these conditions. Prospective studies of anxiety disorders across the life span have also documented the substantial comorbidity with both other mental and physical conditions, suggesting the need for a more comprehensive cross-specialty approach to clinical treatment and evaluation. Comorbidity as an important source of disability also continues to be poorly recognized in clinical samples.

This review also demonstrates that descriptive epidemiology has reached its maturity, and that future studies should focus on identifying the specific components underlying anxiety and their stability and overlap with other classes of mental disorders. Prospective studies of both clinical and community samples will be essential in the shift from descriptive analytic epidemiology designed to identify proximal and distal risk factors for anxiety disorders. The progress in neuroscience and treatment described in Chapters 14.3 through 14.6 should be better integrated into both community and clinical research. Better comprehension of the phenomenology of the specific anxiety disorders and their overlap should guide the development of the next phase of diagnostic categories of anxiety. In addition, as neuroscience and genetics inform knowledge regarding neural processes underlying anxiety disorders and the role of genetic and environmental factors in their evolution, studies of treatment and prevention strategies will assume increasing importance in reducing the magnitude and burden of this major source of mental disorders.

References

Angst J, Paksarian D, Cui L, et al. The epidemiology of common mental disorders from age 20 to 50: results from the prospective Zurich cohort study. *Epidemiol Psychiatr Sci*. 2015;24:1–9.

Baxter AJ, Vos T, Scott KM, Ferrari AJ, Whiteford HA. The global burden of anxiety disorders in 2010. *Psychol Med*. 2014;44(11):2363–2374.

*Beesdo-Baum K, Knappe S. Developmental epidemiology of anxiety disorders. *Child Adolesc Psychiatr Clin N Am*. 2012;21(3):457–478.

*Beesdo-Baum K, Knappe S, Asselmann E, et al. The early developmental stages of psychopathology (EDSP) study: a 20-year review of methods and findings. *Soc Psychiatry Psychiatr Epidemiol*. 2015;50(6):851–866.

Breslau J, Aguilar-Gaxiola S, Kendler KS, Su M, Williams D, Kessler RC. Specifying race-ethnic differences in risk for psychiatric disorder in a USA national sample. *Psychol Med*. 2006;36(1):57–68.

Cho MJ, Chang SM, Lee YM, et al. Prevalence of DSM-IV major mental disorders among Korean adults: a 2006 National Epidemiologic Survey (KECA-R). *Asian J Psychiatr*. 2010;3(1):26–30.

Clauss JA, Avery SN, Blackford JU. The nature of individual differences in inhibited

temperament and risk for psychiatric disease: a review and meta-analysis. *Prog Neurobiol*. 2015;127–128:23–45.

*Copeland WE, Angold A, Shanahan L, Costello EJ. Longitudinal patterns of anxiety from childhood to adulthood: the Great Smoky Mountains Study. *J Am Acad Child Adolesc Psychiatry*. 2014;53(1):21–33.

*Cosci F, Fava GA, Sonino N. Mood and anxiety disorders as early manifestations of medical illness: a systematic review. *Psychother Psychosom*. 2015;84(1):22–29.

Crome E, Grove R, Baillie AJ, Sunderland M, Teesson M, Slade T. DSM-IV and DSM-5 social anxiety disorder in the Australian community. *Austr N Z J Psychiatry*. 2015;49(3):227–235.

Faravelli C, Abrardi L, Bartolozzi D, et al. The Sesto Fiorentino study: background, methods and preliminary results. Lifetime prevalence of psychiatric disorders in an Italian community sample using clinical interviewers. *Psychother Psychosom*. 2004;73(4):216–225.

Goodwin RD, Scheckner B, Pena L, Feldman JM, Taha F, Lipsitz JD. A 10-year prospective study of respiratory disease and depression and anxiety in adulthood. *Ann Allergy Asthma Immunol*. 2014;113(5):565–570.

Hasin DS, Grant BF. The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) Waves 1 and 2: review and summary of findings. *Soc Psychiatry Psychiatr Epidemiol*. 2015;50(11):1609–1640.

Jacobi F, Wittchen HU, Holting C, et al. Prevalence, co-morbidity and correlates of mental disorders in the general population: results from the German Health Interview and Examination Survey (GHS). *Psychol Med*. 2004;34(4):597–611.

Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):593–602.

Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen HU. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int J Methods Psychiatr Res*. 2012;21(3):169–184.

Lewinsohn PM, Shankman SA, Gau JM, Klein DN. The prevalence and co-morbidity of subthreshold psychiatric conditions. *Psychol Med*. 2004;34(4):613–622.

Merikangas K, Avenevoli S, Costello J, Koretz D, Kessler RC. National comorbidity survey replication adolescent supplement (NCS-A): I. Background and measures. *J Am Acad Child Adolesc Psychiatry*. 2009;48(4):367–369.

Merikangas KR, Calkins ME, Burstein M, et al. Comorbidity of physical and mental disorders in the neurodevelopmental genomics cohort study. *Pediatrics*. 2015;135(4):e927–e938.

Merikangas KR, He JP, Brody D, Fisher PW, Bourdon K, Koretz DS. Prevalence and treatment of mental disorders among US children in the 2001–2004 NHANES. *Pediatrics*. 2010;125(1):75–81.

*Merikangas KR, He JP, Burstein M, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *J Am Acad Child Adolesc Psychiatry*. 2010;49(10):980–989.

Mohammadi MR, Davidian H, Noorbala AA, et al. An epidemiological survey of psychiatric disorders in Iran. *Clin Pract Epidemol Ment Health*. 2005;1:16.

Ormel J, Raven D, van Oort F, et al. Mental health in Dutch adolescents: a TRAILS report on

prevalence, severity, age of onset, continuity and co-morbidity of DSM disorders. *Psychol Med*. 2015;45(2):345–360.

Sareen J, Cox BJ, Clara I, Asmundson GJ. The relationship between anxiety disorders and physical disorders in the U.S. National Comorbidity Survey. *Depress Anxiety*. 2005;21(4):193–202.

Silove D, Alonso J, Bromet E, et al. Pediatric-onset and adult-onset separation anxiety disorder across countries in the World Mental Health Survey. *Am J Psychiatry*. 2015;172(7):647–656.

Slade T, McEvoy PM, Chapman C, Grove R, Teesson M. Onset and temporal sequencing of lifetime anxiety, mood and substance use disorders in the general population. *Epidemiol Psychiatr Sci*. 2015;24(1):45–53.

Teesson M, Mitchell PB, Deady M, Memedovic S, Slade T, Baillie A. Affective and anxiety disorders and their relationship with chronic physical conditions in Australia: findings of the 2007 National Survey of Mental Health and Wellbeing. *Aust N Z J Psychiatry*. 2011;45(11):939–946.

Tully PJ, Cosh SM, Baumeister H. The anxious heart in whose mind? A systematic review and meta-regression of factors associated with anxiety disorder diagnosis, treatment and morbidity risk in coronary heart disease. *J Psychosom Res*. 2014;77(6):439–448.

*Walker ER, McGee RE, Druss BG. Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. *JAMA Psychiatry*. 2015;72(4):334–341.

The WHO World Mental Health Survey Consortium. Prevalence, Severity, and Unmet Need for Treatment of Mental Disorders in the World Health Organization World Mental Health Surveys. *JAMA*. 2004;291:2581–2590.

▲ 14.4 Genetics of Anxiety Disorders

JOHN M. HETTEMA, M.D., PH.D., AND TAKESHI OTOWA, M.D., PH.D.

INTRODUCTION

Anxiety and fear are normal emotional responses to potential or actual threat. In anxiety disorders (ADs), these responses are dysregulated, that is, exaggerated or prolonged, in a manner that is disturbing to one's life. The major ADs include GAD, panic disorder (PD), and phobias (social phobia, agoraphobia, and specific phobia). ADs are among the most common mental disorders with a 12-month prevalence estimated around 20 percent in the United States. Because OCD and PTSD have purported distinct etiologic and neurobiologic characteristics from ADs and have not been included in diagnostic category of ADs in the Fifth Edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), this review will not examine these disorders.

Like most medical conditions, ADs are complex disorders, the etiology of which derives from a combination of genetic and environmental factors. The genetic basis of ADs has been indirectly supported by well-established animal models of fear behaviors. Although animal models are considered a

vital tool for advancing basic understanding of the genetic and neurobiologic basis of many human disorders, this chapter will mainly focus on human research in this review. Similarly, research continues into the genetic basis of normal fear, anxiety symptoms, and anxious temperament, but this review will primarily address findings regarding the adult clinical syndromes as defined in DSM. Most child genetics studies have focused on symptoms or temperament rather than ADs. This review is divided into two main divisions of genetic research: (1) genetic epidemiology and (2) molecular genetics.

GENETIC EPIDEMIOLOGY

Genetic epidemiology seeks to understand how diseases and their risk factors are distributed in families. First, family studies compare rates of illness in relatives of those who have the condition (case probands) with rates in relatives of healthy controls. Higher rates in the former group of relatives, as parameterized by a relative risk (RR) or odds ratio (OR) greater than 1.0, suggest familial aggregation. Next, one relies on either adoption studies (not available for ADs) or twin studies to differentiate genetic from within-family environment as sources of aggregation. Twin studies compare resemblance for a condition between members of a twin pair using the fact that identical (monozygotic [MZ]) twins share 100 percent of their genetic sequence while nonidentical (dizygotic [DZ]) twins share, on average, only 50 percent, like any two siblings. One commonly used measure of twin resemblance is the proband-wise concordance, that is, the proportion of co-twins of affected index twins who are also affected. If average concordance for MZ pairs is greater than that for DZ pairs, this is evidence for a genetic component to family resemblance. With larger twin samples, one may also estimate the proportion of individual differences due to the effects of genetic factors (heritability). For conditions with substantial heritability, gene finding (linkage or association) studies are undertaken to identify which specific genes contribute to risk.

A 2001 meta-analysis summarized findings across extant family and twin studies for various adult ADs. Those results suggest an overall moderate level of familial aggregation (OR = 4 to 6) and heritability (30 to 50 percent) across the ADs. This means that, on average, first-degree relatives (FDRs) of someone affected by an AD are four to six times more likely to develop one than a random individual. Fifty percent heritability means that genetic and nongenetic factors are about equally responsible for the etiology of ADs. Few family studies have been published since then, as it is now well established that all ADs aggregate in families. More twin studies have been conducted, however, with emphasis on the etiology of comorbidity or developmental risk. The results of that meta-analysis will be referred to and augmented with data from more recent studies, where available.

Family and Twin Studies

Panic Disorder. Data from 19 controlled family studies overall support the familial aggregation of PD, with risk to FDRs ranging from 3 to 17. Five of these meeting strict inclusion criteria were analyzed in the aforementioned meta-analysis, the results of which showed a highly significant association between PD in the proband and PD in FDR. Summary OR across the five studies was 5.0 (95 percent CI: 3.0 to 8.2) and the unadjusted aggregate risk based on 1,356 total FDR of PD probands was 10.0 percent, compared to 2.1 percent in 1,187 control relatives. Additionally, one study suggests higher familial risk associated with early-onset PD in the proband.

Two of the largest sources of adult twin data for the ADs, including PD, are the population-based Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD), and the Vietnam Era Twin (VET) Registry. The former consists of approximately 9,000 twins from male and female same-sex and opposite sex pairs born in Virginia, while the latter is of comparable size but contains only US male twins who served during the Vietnam War. The size of these samples permits the use of structural equation modeling to assess the relative contributions of genetics, common family (shared) environment, and individual-specific (nonshared) environment to the liability of PD. Both samples reported higher MZ than DZ concordance, suggesting a genetic component to PD, consistent with some prior smaller twin studies. Both VATSPSUD and VET studies estimated heritability of PD to lie between 30 and 40 percent, with a combined family-twin meta-analytic estimate of 43 percent. The remaining source of individual differences derived from individual-specific environment not shared between twins. In addition, the VATSPSUD found no evidence that genetic risk factors for PD significantly differ between men and women.

Generalized Anxiety Disorder. Two published family studies of GAD were included in the aforementioned meta-analysis. Both studies, derived from clinical probands, supported the familial aggregation of GAD, and together they show a significant association between GAD in the proband and in their FDR, with a summary OR of 6.08 (95 percent CI: 2.5–14.9). Both the VATSPSUD and the VET Registry examined broadened GAD syndromes via twin analyses, with an overall heritability of 32 percent (95 percent CI: 24 to 39 percent) when data from both samples were combined via meta-analysis.

Phobias. A group at Columbia University performed a series of analyses examining specific phobia, social phobia, and agoraphobia and their relationship with each other and PD. They reported higher rates of specific phobia in relatives of specific phobia probands compared with

control relatives (31 percent vs. 9 percent), higher rates of social phobia in relatives of social phobia probands (16 percent vs. 5 percent), and higher rates of agoraphobia in relatives of agoraphobia probands (10 percent vs. 3 percent). Another family study reported higher rates of agoraphobia in relatives of probands with agoraphobia compared to control relatives (11.6 percent vs. 4.2 percent, OR = 3.0). Two studies of social phobia found that familial aggregation is primarily due to the generalized subtype. Meta-analysis across family studies found a highly significant association between phobias in the proband and in FDR, with a summary OR of 4.07 (95 percent CI: 2.7 to 6.1).

Only two large, adult twin samples have comprehensively examined the genetics of phobias in adults. A series of analyses in the VATSPSUD examining phobic fears and disorders found that twin resemblance was largely due to genetic factors across agoraphobia, social phobia, and specific phobia, estimating that genetic factors explain 30 to 60 percent of their individual differences. Similar heritability estimates were reported from a study of about 1,400 female Norwegian twins.

Meta-analyses of social anxiety were conducted in 13 cohorts (42,585 subjects in total) of European ancestry. A meta-analysis of five studies that included diagnoses of SAD suggested that genetic influence accounted for 27 percent of variance, while the rest was mostly explained by nonshared environmental factors. Another meta-analysis was conducted using data from 5 twin studies on fears and 10 twin studies on specific phobias. The highest mean heritability among fear subtypes was animal fear (45 percent), and among specific phobias, the highest mean heritability was blood-injury-injection phobia (33 percent). In both studies, variance for the phenotypes could be mainly explained by genetic and nonshared environmental effects.

MOLECULAR GENETICS

Linkage and Candidate Gene Association Studies

Two main approaches have been applied to identify susceptibility genes in human studies: linkage and association studies. Linkage studies are performed in family pedigrees with several affected individuals to identify chromosomal loci likely to harbor a gene influencing a biologic trait or condition. While effective for identifying highly penetrant genes of large effect seen in classic Mendelian disorders, linkage has, with few exceptions, not been successful for most complex phenotypes encountered in medicine, including ADs. Several linkage studies have been performed for ADs, with few consistent findings between studies. Suggestive evidence for linkage to PD has been reported for regions on chromosomes 4q, 9q, 13q, 14q, and 22q. A more recent PD genome-wide linkage scan conducted in the isolated population of the Faroe Islands found an association of the amiloride-

sensitive cation channel 1 (ACCN1) gene, previously supported by animal studies. However, this association was not replicate in an independent sample. A reanalysis of several PD linkage studies supported linkage on chromosomes 4q and 7p. A Yale University group conducted linkage analyses of phobic disorder in a set of pedigrees ascertained for PD, reporting linkage on chromosome 3q for agoraphobia, 14q for specific phobia, and 16q for social phobia. There are no published linkage scans for GAD. A meta-analysis of 162 families with panic or other ADs found suggestive evidence for linkage on chromosomes 1, 5, 15, and 16. Of note, these AD linkage results moderately but significantly correlated with those for neuroticism, also examined in that study.

Association studies allow one to test specific genetic markers for their contribution to a phenotype. They include case-control comparisons in unrelated individuals or family-based transmission tests. To date, most association studies of ADs have focused on candidate genes, which have to be chosen using a priori knowledge, either from their position under a linkage peak or by their purported biologic relevance for the pathophysiology of the disease. Similar to depression, the most widely studied candidate genes for ADs are genes involved in neurotransmitter systems (e.g., serotonin, norepinephrine, GABA, glutamate, and dopamine) or the stress response system. A review summarizing more than 350 candidate gene findings for PD concluded that most of these results were inconsistent, negative, or not clearly replicated. (This has been the case for most candidate gene studies in psychiatry.) Failure to identify susceptibility loci may be due to differences in phenotypic assessment, heterogeneity at the genetic level, underpowered studies with small sample sizes, and poor candidate gene selection. A systematic review reported candidate gene studies for ADs or anxiety symptom measures. When restricted to studies with a sample size of at least 200 cases, 65 reports were found. PD had been the most frequently studied disorder (n = 30), compared to GAD (n = 4), social phobia (n = 1), and specific phobia (n = 1). The three most commonly studied genes were catechol-O-methyltransferase (COMT), serotonin transporter (SLC6A4, also known as 5-HTT), and brain-derived neurotrophic factor (BDNF). Meta-analyses showed significant association of the COMT 158 Val allele and the promoter region variable number tandem repeat region of the gene encoding monoamine oxidase A (MAOA) with PD in Caucasian female samples. However, even the largest studies that examined SLC6A4 (1,161 cases vs. 1,051 controls) found no consistent evidence of association with the often studied 5-HTT promoter variant. Given that the probability for a specific candidate gene to be causally related to the disorder is low due to limited a priori knowledge of the pathogenesis of ADs, an association with a candidate gene in a single study is most likely to be a false-positive finding (type I error). Furthermore, most complex disorders are believed to be highly polygenic, and common

variants underlying these disorders are expected to have small effect sizes. Therefore, much larger sample sizes than those included in previous studies are required for reliable results.

Genome-Wide Association Studies

Unlike candidate gene studies, which require a priori knowledge for gene selection, genome-wide association studies (GWAS) provide an unbiased survey of common genetic variation across the entire genome. The first AD GWAS was conducted in 200 PD cases and 200 controls in the Japanese population by a group at Tokyo University, reporting several potential novel loci for susceptibility to PD. However, a follow-up study in a larger sample failed to show significant association of these genes with PD. A combined meta-analysis of the two Japanese PD data followed by a replication produced no genome-wide significant findings. That group performed biologic pathway analyses using these Japanese PD GWAS data, finding support for immune-related involvement in PD. Another GWAS of PD was conducted in three German samples (in total, 909 cases vs. 915 controls) and reported two associated single nucleotide polymorphisms (SNPs) located in TMEM132D on 12q24.3. Association of this gene was subsequently replicated using a larger sample (1,670 cases vs. 2,266 controls). Risk genotypes for PD were associated with higher TMEM132D mRNA expression levels in the frontal cortex and with anxiety-related behavior in mice. Only one genome-wide copy number variation (CNV) study of PD has been conducted, with significant associations in a Japanese sample for common duplications in the 16p11.2 region; these are located near a large, rare CNV previously associated with autism and other neurodevelopmental disorders. Of note, none of these analyses found overlap with previously reported candidate genes, making these results tentative until replicated in independent, adequately powered samples.

Other AD phenotypes besides PD have been less extensively studied with GWAS. One study investigated candidate gene associations in a sample characterized for shared risk between ADs and depression, in which the candidate genes were selected from prior GWAS in mice and humans. Association was reported with the PPARGC1A gene, a transcriptional activator involved in energy expenditure and in metabolic conditions. A GWAS of anxiety-related behaviors (negative cognition, negative affect, fear, social anxiety, and anxiety composite) was conducted in 2,810 children followed up by replication in another sample of 4,804 children. Although no genome-wide significant signals were found, a significant heritability (10 percent) was captured by common variants across the genome. A GWAS of phobic anxiety symptoms performed in 11,127 individuals of European descent found no genome-wide significant signals.

At least two major considerations have thus far limited the search for AD susceptibility variants: (1) small sample sizes and (2) phenotypic

heterogeneity. Such limitations have thus far plagued many psychiatric GWAS endeavors. As mentioned above, the typical sample sizes examined in genetic studies of ADs have been underpowered by today's standards. Using the effect size (OR = 1.2) observed in recent GWAS studies in other psychiatric disorders, minor allele frequency of 0.5, and a genome-wide significant threshold of $p = 5 \times 10^{-8}$, power analyses indicate that a minimum of 5,000 cases and 5,000 controls would be necessary to obtain 80 percent power to detect significant association above the noise. Regarding phenotypic heterogeneity, the boundaries among the current definitions of ADs are not clear. Furthermore, standard case-control genetic association studies focus on single categorical phenotypes comparing allele frequencies for individuals who meet and do not meet criteria for one or another specific disorder. Twin studies suggest that ADs exhibit strong lifetime comorbidity, partially due to shared genetic risk factors between them; that is, genetic risk factors do not respect diagnostic boundaries delimited by clinical nosology. Therefore, it might be more efficient to focus on clusters of disorders with shared genetic risk factors rather than on a single individual AD.

To increase sample sizes, several international research collaborations have formed to investigate the genetics of ADs, similar to efforts by the Psychiatric Genetics Consortium for other psychiatric disorders. One such group, the Anxiety NeuroGenetics STudy (ANGST) consortium, has led a large GWAS meta-analysis effort using data from nearly 20,000 subjects in nine independent samples from the United States, Europe, and Australia. The approach uses both case-control comparisons and quantitative anxiety factor scores to detect variants playing a role in common genetic risk shared across ADs. That study has tentatively identified genome-wide association for two chromosomal regions containing potentially novel AD risk genes: a noncoding RNA locus LOC152225 on 3q12.3 and the CAMKMT gene on 2p21, with several other regions showing suggestive association. Replication in larger studies will determine whether these or other novel genes play a critical role in AD susceptibility.

Gene–Environmental Interactions

Although several large meta-analyses of GWAS of internalizing disorders such as major depressive disorder (MDD) have been conducted, it has been difficult to identify main effects of common variants in these disorders. Given the substantial role for nongenetic risk factors indicated by twin studies, additional approaches such as a gene-environment (G×E) interaction analyses might provide increased insight into the pathogenesis of these disorders. G×E interaction occurs when the effect of genetic risk factors is modified by life experiences or other environmental exposures. For example, individuals with certain genetic risk profiles might be the most susceptible to the anxiety-provoking effects of stress.

Candidate gene studies of G×E interaction have begun to appear, but this remains an underutilized strategy for genetic studies of most ADs. Similar to candidate gene studies reviewed above, G×E studies have mainly focused on the serotonin, hypothalamic–pituitary–adrenal (HPA) axis (e.g., CRHR1), and neurotrophin systems (e.g., BDNF) in predicting ADs. The most extensively investigated gene in relationship to environment is the serotonin transporter (SLC6A4, specifically its functional promoter length polymorphism, 5-HTTLPR). Some studies reported that the short variant modified the effect of current daily stress or previous childhood maltreatment on anxiety sensitivity, which may be a potential precursor or marker for ADs. In contrast, other studies reported that long-allele homozygotes with childhood maltreatment or family adversity displayed higher anxiety sensitivity and AD rates, respectively. A study examined interactions between the BDNF gene and early-life stress (ELS) exposure for depression and anxiety in 374 healthy European subjects. The authors found that the BDNF Met carriers predicted higher depression through reduced gray matter in hippocampus, while the BDNF Val homozygotes contributed specifically to elevated anxiety. Candidate genes in the HPA axis are also promising for the investigation of G×E, given the system's role in stress response. However, most of the previous research has examined the interaction between HPA function and ELS in depression and PTSD, with little focus on G×E effects for other ADs. Similar to the study of the main effects of candidate genes, G×E studies are limited in power and replicability.

So far, G×E studies have focused on a small number of putative functional markers in a limited number of candidate genes. Therefore, a systematic search for G×E interactions across the genome is essential. However, despite the availability of GWAS data, no genome-wide search for G×E in common psychiatric disorders including ADs has been conducted. The two major considerations are (1) sample sizes and (2) assessment of environment. Since the statistical power for detecting G×E interactions is even lower than that for detecting direct disorder-genotype association, very large samples will be needed. The second consideration is the quality of assessment of environmental variables. In most cases, life events are usually assessed as environmental variables, but the definition of stressors is inconsistent. There is converging evidence for a significant impact of cumulative as well as specific life events (e.g., threat, interpersonal, and health-related events in adulthood, and abuse or loss/separation experiences in childhood) on the pathogenesis of ADs. Given that meta-analyses from multiple cohorts will be required to meet power requirements, a more precise conceptualization and operationalization of "environment" or "candidate stressors" will be important for the design of future genome-wide G×E studies.

Epigenetics

Epigenetics is a rapidly emerging genetic approach that augments genetic association studies. The term "epigenetics" is now generally understood to refer to potentially heritable and functionally relevant modifications of gene expression and chromatin structure, accompanying no changes in genetic sequences. DNA methylation is one of the major forms of epigenetic modifications and occurs at five carbon positions of cytosine residues. Methylated cytosine residues usually suppress gene expression, interfering with the binding of transcription factors or recruiting methylated DNA binding factors. Such mechanisms have been considered to be possible mediators of responses to environmental factors, and therefore may play a certain role in pathogenesis of psychiatric disorders including ADs.

Similar to G×E studies, previous studies on methylation alteration in ADs have focused primarily on candidate genes putatively involved in stress response, neurotransmission, and neuroplasticity. One study reported hypomethylation of the oxytocin receptor (OXTR) gene in social phobia patients. The oxytocin neuropeptide system plays a role in social behavior and cognition. Methylation alteration in another candidate gene, glutamate-decarboxylase 1 (GAD1), was reported in PD patients. The GAD1 encodes the glutamate decarboxylase GAD67 isoform which is the rate-limiting enzyme in GABA synthesis from glutamate in the brain. CpG sites in the promoter and intron 2 region in GAD1 were found to be hypomethylated in PD patients compared with control subjects. Interestingly, GAD1 has shown genetic association to ADs in several prior studies. Another preliminary study has also shown that CpG sites in the promoter and exon/intron 1 of MAOA are significantly hypomethylated in female patients with PD, and negative life events were associated with the lower methylation level. A prior meta-analysis also supports a role of MAOA in PD. Methylation of the serotonin transporter gene was examined in children with ADs before and after cognitive behavior therapy. Increase in SLC6A4 methylation was related to response to the psychological therapy.

Stressful events, especially early-life adversities, are known to increase the risk of adult-onset ADs. Stress activates a number of neuronal circuits, in particular the HPA axis. Several genes associated with the HPA axis, such as corticotrophin-releasing factor (CRF), arginine vasopressin (AVP), proopiomelanocortin (POMC), glucocorticoid receptor (GR) (NR3C1) and FK506 binding protein 5 (FKBP5), have been targeted based on the hypothesis that methylation alterations in these genes mediate the biologic effects of stress. Previous animal studies have shown that changes in methylation of these genes upregulate the HPA axis and increase anxiety. One of these genes, NR3C1, has also been studied in humans, with hypermethylation reportedly associated with early-life adversity. Another

study reported a similar finding for hypomethylation in intron 7 of FKBP5. These studies suggest that dysregulation of the HPA axis resulting from excess stress can potentially increase anxiety through alterations in DNA methylation. Neurotransmission and neuroplasticity are also considered to be involved in anxiety-inducing stress responses. In particular, BDNF and SLC6A4 genes have been examined in several studies. Hypermethylation of promoter/exon IV of BDNF, which is thought to be related to neuroplasticity and play a role in the regulation of stress response, has been reported in MDD patients with histories of suicidal behavior associated with anxiety. Increased methylation of SLC6A4 has also been associated with early-life adversity.

Recently, genome-wide approaches have enabled the examination of DNA methylation patterns beyond candidate genes. Genome-wide methylation studies have identified genes such as BCL11A, JAG1, and MBP as novel candidate genes associated with anxiety. Global DNA methylation alterations have also been reported in individuals with anxiety phenotype. One study has reported that anxious participants had significantly higher levels of global DNA methylation than controls. Although further studies are needed to identify mechanisms underlying AD development, DNA methylation changes play a potentially important role, being a mediator of environmental factors like stress that have also been shown to increase risk.

FUTURE DIRECTIONS

Similar to most medical and behavioral/psychiatric phenotypes, the genetics of ADs has gone through several generations of research efforts and paradigms. Family and twin studies have unequivocally established that genetic factors play a key, but somewhat moderate, role in the risk for ADs, balanced by a similarly important contribution from nongenetic risk factors like stressful life events. Much subsequent effort was expended through linkage and association studies to identity the specific genes involved, with little success. The limitations have come from both the genetic side (genetic architecture with many genes of small effects, underpowered samples, poor choices of candidates to test) and the "disease" (phenotype) side (diagnostic definitions, boundaries, comorbidity, heterogeneity). Very few genes, thus far, have robust (i.e., consistently replicated) evidence for involvement in AD susceptibility. Appropriately, researchers seek to include the effects of environmental factors in AD genetic models through G×E and epigenetic study designs, but these are also limited by complexity of such effects and sample sizes needed to detect them.

Albeit somewhat behind genetics research for psychiatric disorders considered to be "serious mental illnesses" (schizophrenia, bipolar disorder, autism), AD genetics continues to advance through these

limitations. On the genetics side, researchers better appreciate the complexity of the genetic architecture of these phenotypes, pushing the boundaries of study design with the use of GWAS and very large sample sizes achieved through collaboration and meta-analysis. AD GWAS sample sizes are currently at the cusp of being able to detect the small effects of risk loci, making this an exciting period. Further sample size increases will continue to be needed.

The phenotypic limitations will likely require more time, effort, and novel approaches to overcome. Except in rare instances, psychiatric genetics research has not benefited from the "medical model," in which a clinical syndrome as described by psychiatric nosology is reified as the exclusive focus of study. As mentioned earlier, ADs do not seem to represent distinct genetic syndromes themselves, but rather, are likely multiple representations of dysregulated fear systems in the brain. Diagnostic syndromes are poor representations of the (unmeasured) biologic expression of underlying genetic susceptibility. One alternative approach is to construct phenotypic measures out of combinations of AD clinical features that share genetic risk factors, as currently applied by the ANGST consortium. With sufficient sample sizes and careful study methodology, this approach has the potential to discover many novel anxiety susceptibility genes but has little potential to provide insights into the mechanisms by which these genes eventually affect the biologic underpinnings of ADs. The other major approach, as advocated in the National Institute Of Mental Health's Research Domain Criteria (RDoC) initiative, is to identify more basic, neuroscience-based measures associated with one or more clinical syndromes that can be studied in both patients and unaffected individuals. Sometimes called endophenotypes, these are ideally laboratory-based, dimensional (rather than categorical) measures found to be abnormal in patients and their biologic relatives. Often, they have cross-species validity, making them feasible for study in model organisms and nonhuman primates. Several candidate endophenotypes have already been under study for ADs: behavioral response to novel stimuli, fear conditioning and extinction, sensitivity to chemical probes that elicit panic (e.g., carbon dioxide, caffeine), and cognitive, emotional, and physiologic responses to threat, including attentional biases, startle reactivity, and amygdala activation and its modulation by connections to the prefrontal cortex. There are burgeoning genetic studies of such phenotypes using twin and candidate gene approaches. However, the complexity of these phenotypes will likely make them impractical for the type of large sample studies necessary to detect novel genetic associations. Their major value will more likely lie in the dissection and elucidation of the effects of genes discovered through GWAS on the mechanistic pathways intermediate between them and clinical expression of ADs.

REFERENCES

*Binder EB. The genetic basis of mood and anxiety disorders – changing paradigms. *Biol Mood Anxiety Disord*. 2012;2:17.

Czajkowski N, Kendler KS, Tambs K, Roysamb E, Reichborn-Kjennerud T. The structure of genetic and environmental risk factors for phobias in women. *Psychol Med*. 2011;41:1987–1995.

Daskalakis NP, Yehuda R. Site-specific methylation changes in the glucocorticoid receptor exon 1F promoter in relation to life adversity: systematic review of contributing factors. *Front Neurosci*. 2014;8:369.

Domschke K, Tidow N, Kuithan H, et al. Monoamine oxidase A gene DNA hypomethylation – a risk factor for panic disorder? *Int J Neuropsychopharmacol*. 2012;15:1217–1228.

Domschke K, Tidow N, Schrempf M, et al. Epigenetic signature of panic disorder: a role of glutamate decarboxylase 1 (GAD1) DNA hypomethylation? *Prog Neuropsychopharmacol Biol Psychiatry*. 2013;46:189–196.

Erhardt A, Czibere L, Roeske D, et al. TMEM132D, a new candidate for anxiety phenotypes: evidence from human and mouse studies. *Mol Psychiatry*. 2011;16:647–663.

*Hettema JM, Neale MC, Kendler KS. A review and meta-analysis of the genetic epidemiology of anxiety disorders. *Am J Psychiatry*. 2001;158:1568–1578.

Hettema JM, Prescott CA, Myers JM, Neale MC, Kendler KS. The structure of genetic and environmental risk factors for anxiety disorders in men and women. *Arch Gen Psychiatry*. 2005;62(2):182–189.

Hettema JM, Webb BT, Guo AY, et al. Prioritization and association analysis of murine-derived candidate genes in anxiety-spectrum disorders. *Biol Psychiatry*. 2011;70:888–896.

Hovatta I, Barlow C. Molecular genetics of anxiety in mice and men. *Ann Med*. 2008;40:92–109.

Howe AS, Buttenschon HN, Bani-Fatemi A, et al. (2016). Candidate genes in panic disorder: meta-analyses of 23 common variants in major anxiogenic pathways. *Mol Psychiatry*. 2016;21:665–679.

Klauke B, Deckert J, Reif A, Pauli P, Domschke K. Life events in panic disorder—an update on "candidate stressors." *Depress Anxiety*. 2010;27:716–730.

*Maron E, Hettema JM, Shlik J. Advances in molecular genetics of panic disorder. *Mol Psychiatry*. 2010;15:681–701.

*McGrath LM, Weill S, Robinson EB, Macrae R, Smoller JW. Bringing a developmental perspective to anxiety genetics. *Dev Psychopathol*. 2012;24:1179–1193.

Murphy TM, O'Donovan A, Mullins N, O'Farrelly C, McCann A, Malone K. Anxiety is associated with higher levels of global DNA methylation and altered expression of epigenetic and interleukin-6 genes. *Psychiatr Genet*. 2015;25(2):71–78.

Otowa T, Kawamura Y, Nishida N, et al. Meta-analysis of genome-wide association studies for panic disorder in the Japanese population. *Transl Psychiatry*. 2012;2:e186.

Otowa T, Hek K, Lee M, et al. Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry*. 2016;21(10):1391–1399.

Roberts S, Lester KJ, Hudson JL, et al. Serotonin transporter methylation and response to cognitive behaviour therapy in children with anxiety disorders. *Transl Psychiatry*.

2014;4:e444.

Scaini S, Belotti R, Ogliari A. Genetic and environmental contributions to social anxiety across different ages: a meta-analytic approach to twin data. *J Anxiety Disord*. 2014;28:650–656.

Schraut KG, Jakob SB, Weidner MT, et al. Prenatal stress-induced programming of genome-wide promoter DNA methylation in 5-HTT-deficient mice. *Transl Psychiatry*. 2014;4:e473.

Schumacher J, Kristensen AS, Wendland JR, Nothen MM, Mors O, McMahon FJ. The genetics of panic disorder. *J Med Genet*. 2011;48:361–368.

*Shimada-Sugimoto M, Otowa T, Hettema JM. Genetics of anxiety disorders: genetic epidemiological and molecular studies in humans. *Psychiatry Clin Neurosci*. 2015;69(7):388–401.

Smoller JW, Faraone SV. Genetics of anxiety disorders: complexities and opportunities. *Am J Med Genet C Semin Med Genet*. 2008;148c:85–88.

Uher R. Gene-environment interactions in severe mental illness. *Front Psychiatry*. 2014;5:48.

Van Houtem CM, Laine ML, Boomsma DI, Ligthart L, van Wijk AJ, De Jongh A. A review and meta-analysis of the heritability of specific phobia subtypes and corresponding fears. *J Anxiety Disord*. 2013;27:379–388.

Webb BT, Guo AY, Maher BS, et al. Meta-analyses of genome-wide linkage scans of anxiety-related phenotypes. *Eur J Hum Genet*. 2012;20:1078–1084.

Ziegler C, Dannlowski U, Bräuer D, et al. Oxytocin receptor gene methylation: Converging multilevel evidence for a role in social anxiety. *Neuropsychopharmacology*. 2015;40(6):1528–1538.

▲ 14.5 Anxiety Disorders: Neurobiology and Neuroscience

ADRIANA FEDER, M.D., SARA COSTI, M.D., BRIAN M. IACOVIELLO, PH.D., JAMES W. MURROUGH, M.D., AND DENNIS S. CHARNEY, M.D.

There continues to be an expansion in the knowledge of the neurobiology of fear and anxiety. Preclinical research has implicated a range of neurochemical systems in the etiology of fear and anxiety responses. Dysfunction of these systems may also play a role in the development of anxiety disorders through their activity on certain important cortical and subcortical brain regions. As was recently emphasized by J. E. LeDoux, threat detection and defense responses in animals constitute nonconscious mechanisms, which in turn lead to responses in the brain and body that eventually result in conscious fear. While threat responses have been elegantly studied in animals and their study will continue to inform the understanding of anxiety disorders in humans, it is important to remember this distinction between automatic responses to threat and subjectively experienced fear and anxiety, as the latter can only be studied in humans. Thus, continued close collaboration between basic and clinical researchers is required to advance the current understanding of the neurobiology of anxiety disorders. Advances in functional neuroimaging have also allowed

for greater visualization of key brain regions in humans with and without anxiety disorders, and during emotion, learning, or anxiety-producing tasks. In addition, progress in molecular genetics allows for the identification of genetic polymorphisms that may confer vulnerability to anxiety disorders or a poorer treatment response. The rapidly growing field of imaging genomics, combined with new imaging techniques, will further help identify genes of key relevance for anxiety disorders.

NEUROCHEMICAL BASIS OF FEAR AND ANXIETY

The human body responds to perceived threats and fear-inducing stimuli by activating the release of specific neurotransmitters and neuropeptides located in key brain regions. The "fight-or-flight" reaction can be adaptive, in that increased vigilance, recall of potentially dangerous elements, and rapid physiologic response (evidenced by increased pupillary dilation, increased heart rate, and rapid breathing) prepare the organism for impending danger. On the other hand, overgeneralization of fear-learning to harmless stimuli is thought to characterize anxiety disorders.

Neurochemicals released during stress or periods of prolonged anxiety can result in chronic alterations to a range of neurobiologic systems. Severe stress can also have maladaptive consequences when it causes biologic systems to be chronically activated. Based on preclinical and clinical models of anxiety, there is evidence that exposure to early-life stressors or trauma may result in predisposition to the development of anxiety disorders later in life, by alterations in neuroendocrine systems or perhaps by direct neurotoxic effects. Why certain individuals are more susceptible to developing anxiety after such stressors while others are resilient represents an area of active investigation. This chapter reviews current knowledge about a range of neurotransmitters and neuropeptides, discusses how they interact in producing fear responses, and summarizes findings pertinent to anxiety disorders from animal models of anxiety and from human studies (Table 14.5–1).

NORADRENERGIC SYSTEM

Norepinephrine (NE) is a catecholamine released in the central nervous system (CNS) primarily from the locus coeruleus (LC), which is located in the pons. NE is a product of hydroxylation of dopamine, which is a metabolite of tyrosine. The NE system receives input from and sends projections to a variety of brain regions, modulating neurons that release a range of other neurotransmitters, including dopamine, serotonin, glutamate, and γ-amino butyric acid (GABA). Neurons in the LC and innervations from the LC to cortical and limbic regions are involved in the response to fearful stimuli and adaptation to acute and chronic stress. NE is also released by sympathetic nervous system ganglia and via synaptic vesicles in the adrenal medulla. It is responsible for regulating the body's

behavioral and physiologic responses to stress through activity of the adrenergic receptors or adrenoreceptors. These receptors are G-protein–coupled and are responsible for controlling NE release, in part by negative feedback. α-1 adrenergic receptors are postsynaptic, whereas α-2 adrenergic receptors are presynaptic. Animal studies of these receptors have led to a greater understanding of the role of NE in anxiety disorders.

Agonists of the α-2 adrenergic receptors, such as clonidine and guanfacine, are antiadrenergic, whereas α-2 antagonists, such as idazoxan and yohimbine, result in increased NE release. In addition, certain environmental stressors in animal studies, such as acute cold restraint and inescapable stress, cause a reduction in α-2 adrenergic receptors density in the hippocampus and amygdala. Mice lacking the α-2 receptor exhibit heightened autonomic activation and concomitant anxiety behaviors after administration of D-amphetamine.

Studies have led to the discovery of a functional polymorphism of the α-2C receptor in humans. This polymorphism, α-2C-Del322-325, is a deletion of codon repeats found to cause an increased release of NE caused by dysfunction in the negative feedback system. Found in higher rates in African Americans, it is associated with greater risk of cardiovascular diseases such as congestive heart failure due to increased circulating NE. During a yohimbine challenge, homozygotes of the α-2C-Del322-325 polymorphism were also more susceptible to increases in NE, heart rate, blood pressure, and anxiety than noncarriers. In a study of healthy males, involving the α-2B adrenergic autoreceptor, the short allele of the serotonin transporter polymorphism (5-HTTLPR) was found to influence attentional bias for negative information in opposite directions depending on the specific α-2B adrenergic autoreceptor genotype. These findings might have potential implications for the development of anxiety disorders.

Dysfunction of the NE system has been demonstrated in panic disorder (PD) and phobias. There has been longstanding clinical evidence that propranolol, a nonselective β-adrenergic receptor-blocking agent, is useful in the management of SAD, particularly stage fright and performance anxiety, although it has no direct impact on anticipatory anxiety. Beta blockers reduce peripheral autonomic symptoms such as hand tremors and tachycardia often associated with social anxiety. Further, pro-adrenergic agents such as yohimbine have been reported to be panicogenic in animal and human models of anxiety. Yohimbine is known to precipitate panic attacks in patients with PDs, but not in healthy subjects. In a study of children with various anxiety disorders, yohimbine caused greater anxiety responses than in the healthy control group. A more selective α-2 receptor antagonist than yohimbine, idazoxan, has been shown to provoke panic attacks in healthy subjects. Patients with PD also have an abnormal response to clonidine (Catapres), with greater decreases in blood pressure and plasma 3-Methoxy-4-hydroxyphenylglycol (MHPG), an NE metabolite,

compared to controls.

Studies of phobic individuals have demonstrated higher baseline rates of NE and catecholamines, and increased release of NE in response to anxiety-inducing triggers. This is not surprising, given the role that NE plays in fear conditioning and how repeated exposure to a stressful stimulus causes a potentiated release of NE, a process called sensitization. Patients with GAD also show higher plasma levels of NE and MHPG than healthy controls. Their response to yohimbine, however, is generally similar to that of healthy controls. This might be due to downregulation of α-2 receptors in patients with GAD, reported in platelets and thus also assumed in the CNS. Guanfacine, an α-2 receptor agonist approved in extended release form for attention-deficit hyperactivity disorder (ADHD), has been primarily tested for PTSD (formerly classified as an anxiety disorder) but showed no difference compared to placebo in clinical trials. A trial of guanfacine is ongoing in children and adolescents with anxiety disorders including GAD, SAD, and separation anxiety disorder.

Table 14.5–1.
Neurochemical Systems Involved in Anxiety Disorders and Related Treatment Approaches

Neurochemical	Brain Regions	Association with Anxiety Disorders	Treatment Approaches (Current and Future)
Noradrenergic system	Locus coeruleus, amygdala, hippocampus, hypothalamus, prefrontal cortex	Unrestrained, excessive system activation	SNRIs first-line therapy for anxiety disorders (along with SSRIs) Propranolol for performance anxiety Guanfacine did not show benefits for PTSD (no longer classified as anxiety disorder)
Hypothalamic–pituitary–adrenal (HPA) axis	Hippocampus, amygdala, hypothalamus, prefrontal cortex	Dysregulated HPA axis function (excessive cortisol release, abnormal feedback) in some studies	Cortisol administration under study for SAD and spider phobia Mifepristone under study for GAD and PD
Corticotropin-releasing hormone (CRH)	Prefrontal and cingulate cortices, amygdala, hippocampus, hypothalamus, bed nucleus of stria terminalis, nucleus accumbens, periaqueductal gray matter, locus coeruleus, dorsal raphe nuclei	Persistently increased CRH concentration	To date, CRH-1 receptor antagonists have failed to demonstrate efficacy in clinical trials
Neurosteroids	Hippocampus, amygdala, cortex	Abnormal peripheral levels	Paroxetine found to increase peripheral

		of neurosteroids in PD; possibly also in GAD and SAD (inconsistent findings)	levels of ALLO in one study of PD patients, but not in a second study Synthetic analogs of neurosteroids under development
Arginine vasopressin (AVP)	Paraventricular nucleus of hypothalamus, septum, hippocampus, cortex	Single nucleotide polymorphism in AVP V1b receptor gene linked to PD	AVP V1b receptor antagonist (SSR149415) failed to demonstrate efficacy for GAD in a clinical trial New AVP V1b receptor antagonists currently under study
Dopaminergic system	Amygdala, nucleus accumbens, prefrontal cortex	Excessive mesocortical dopamine release, persistently high levels of dopamine in prefrontal cortex	Bupropion is an NDRI used primarily to treat depression, sometimes as an adjunct for anxiety disorders
Serotonergic system	Dorsal raphe nuclei, amygdala, hippocampus, prefrontal cortex	Low activity of postsynaptic 5-HT1A receptors in PD and SAD	SSRIs and SNRIs first-line therapy for anxiety disorders
γ-amino butyric acid (GABA)	Substantia nigra, globus pallidus, hypothalamus, periaqueductal gray matter, hippocampus, amygdala, anterior cingulate	Reduced GABA-A and benzodiazepine binding in PD Reduced GABA levels in PD, possible imbalance between tonic GABAergic inhibition and glutamate-mediated excitation	Tiagabine, an SGRI (selective GABA reuptake inhibitor) and vigabatrin (inhibitor of GABA transaminase) are potential treatments Tiagabine equivocal findings in GAD Topiramate (blocks voltage-sensitive sodium channels, potentiates GABA) mixed findings in PD Compounds selective for specific GABA-A receptor subtypes under development **Note:** Gabapentin and pregabalin, although structurally related to GABA, do not act on GABA receptors
Glutamate	Amygdala, hippocampus, frontal and cingulate cortices	Possible imbalance between tonic GABAergic	Efficacy of DCS as adjunct to exposure therapy for

		inhibition and glutamate-mediated excitation in PD	acrophobia, SAD and PD Glycine transporter inhibitors under study Metabotropic receptor modulators under study NMDA receptor antagonism as potential anxiolytic—preliminary efficacy of riluzole for GAD
Neuropeptide Y (NPY)	Amygdala, hippocampus, brainstem, nucleus accumbens, locus coeruleus, hypothalamus	Low NPY levels in PTSD, less well studied in anxiety disorders	Intranasal NPY administration may reduce anxiety, under investigation Y1 and Y2 receptors potential targets for treatment
Galanin	Prefrontal cortex, amygdala, hippocampus, locus coeruleus	Very few studies in patients with anxiety disorders Galanin gene polymorphism associated with PD, only in women	Potential of galanin modulators as future approach; no known studies
Cholecystokinin (CCK)	Cerebral cortex, hippocampus, amygdala, caudate, putamen, thalamus, hypothalamus	Lower CSF levels of CCK in PD	To date, CCK-B receptor antagonists have failed to demonstrate efficacy for GAD or PD
Oxytocin	Hypothalamus, ventral tegmental area, amygdala	Oxytocin gene receptor polymorphism associated with increased risk for anxiety in individuals with early life stress	Intranasal oxytocin administration beneficial for SAD, under study; potential benefit for GAD
Endocannabinoid system	Prefrontal cortex, hypothalamus, amygdala, hippocampus	Dysregulation of endocannabinoid signaling	Cannabidiol reduced anxiety in patients with SAD during public speaking; under study Studies of FAAH inhibitors as potential treatment under way for PTSD (no longer classified as anxiety disorder); needed in anxiety

GAD, generalized anxiety disorder; PD, panic disorder; SAD, social anxiety disorder; NMDA, *N*-methyl-D-aspartate; NDRI, norepinephrine-dopamine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor; SNRI, serotonin–norepinephrine reuptake inhibitor; DCS, D-cycloserine.

To summarize, patients with all types of anxiety disorders exhibit symptoms consistent with an exaggerated noradrenergic system output, including excessive worry, avoidance, and heightened arousal and startle, along with increased autonomic and sympathetic activation. Drugs that act to attenuate this effect by decreasing the firing of NE neurons, including substances such as alcohol, opioids, and benzodiazepines, are commonly used, and abused, by patients with anxiety disorders as a form of self-medication. Drugs that increase NE signaling, such as SNRIs and tricyclic antidepressants, are thought to treat anxiety via eventual adaptive changes in NE (and serotonin) receptors occurring weeks later.

HYPOTHALAMIC–PITUITARY–ADRENAL AXIS

Cortisol is a glucocorticoid synthesized in the adrenal glands. It contributes to increased arousal, vigilance, focused attention and memory formation, inhibition of the growth and reproductive systems, and containment of the immune response. It is also known to regulate, marshal, and replenish energy stores. There is ample support for the notion that many types of psychological stress increase the synthesis and release of cortisol. Cortisol is a principal contributor in regulating hippocampus, amygdala, and prefrontal cortex (PFC) activity. It has been shown to enhance amygdala activity, increase corticotropin-releasing hormone (CRH) messenger ribonucleic acid (mRNA) concentrations in the central nucleus of the amygdala (CEA), increase the effects of CRH on conditioned fear, and facilitate the encoding of emotion-related memory. Of interest, this glucocorticoid-mediated facilitation of emotional memory encoding has been found to be contingent on endogenous arousal-induced activation of NE in the amygdala. In addition, glucocorticoids such as cortisol have biphasic effects on cognitive function, hippocampal excitability, and memory. Such effects may account, at least in part, for the adaptive role of cortisol in modulating behavior in response to acute stress.

The stress-induced increase in cortisol release is regulated through a complex negative feedback system mediated via moderate-to-low-affinity glucocorticoid receptors (GRs) and high-affinity mineralocorticoid receptors (MRs). The hippocampus has a very high concentration of GR receptors and possibly also MR receptors, and plays an important role in the negative feedback of the HPA axis. If stress-induced cortisol secretion is not contained, persistently elevated cortisol levels can cause diverse harmful effects both peripherally and in the CNS. Peripheral adverse effects include hypertension, osteoporosis, immunosuppression, insulin

resistance, dyslipidemia, dyscoagulation, and ultimately atherosclerosis and cardiovascular disease. The hippocampus is particularly sensitive to many types of insults. High glucocorticoid levels impair cell survival, alter cell metabolism, and induce changes in cell morphology. The stress-related impairment in hippocampal-dependent cognitive and memory function may be a consequence of this cellular damage. Furthermore, there is accumulating evidence that the amygdala may play an important role in mediating the negative effects of stress on the hippocampus.

Tests used to probe HPA axis function in patients with anxiety disorders include examination of cortisol suppression after administration of low-dose dexamethasone (DEX), the response of adrenocorticotropic hormone (ATCH) after CRH administration, and the DEX/CRH challenge test. Blunted ACTH responses to CRH have been reported in PD patients in some but not all studies. The DEX-CRH challenge test is a sensitive and specific technique for assessing HPA axis function. Responsiveness of the HPA system to this test in PD patients was found to be higher than in healthy controls but lower than in depressed patients. Urinary free cortisol and ACTH results in PD have been inconsistent. Plasma cortisol levels were elevated in some studies and normal in others. In a 24-hour study of ACTH and cortisol secretion in PD, patients had elevated overnight cortisol secretion and greater amplitude of ultradian secretory episodes. Cortisol levels may also be correlated with severity of PD. Cognitive intervention in PD has been shown to reduce significantly the elevation of cortisol and ACTH by pentagastrin challenge, although it had no effect on panic symptoms per se. Thus, HPA axis hypersensitivity in PD may be activated by environmental contextual cues such as novelty, whereas baseline activity might remain normal.

Despite some discrepancies in the recent literature, most evidence suggests that GAD is also associated with HPA axis abnormalities. The DEX suppression test showed a reduction in negative feedback sensitivity of the HPA axis in these patients. GAD has also been associated with higher levels of basal cortisol in plasma and saliva. Interestingly, cortisol levels, as well as severity of anxiety symptoms, were demonstrated to decrease after treatment, independently from the kind of therapy provided, benzodiazepine, SSRI, or psychotherapy.

Although several studies reported no abnormalities in basal HPA axis function in adult subjects suffering from SAD, a recent study found lower basal plasma cortisol levels during morning hours in patients with SAD compared to healthy controls. Data from studies in children and adolescents with SAD confronted with a psychosocial stressor show conflicting results, with some reporting similar salivary cortisol levels compared to controls and others increased salivary cortisol levels in children and adolescents with SAD.

There has been ongoing research on the use of glucocorticoids in the

treatment of anxiety disorders. In preclinical and clinical studies, glucocorticoids have been shown to block the recall of traumatic memories. Cortisol was also found to reduce the activation of fear responses in patients with SAD and spider phobia. In a recent double-blind placebo-controlled study, oral cortisol administration was shown to enhance therapeutic response to in vivo exposure-based group therapy in patients with specific phobia (SP) of spiders. In preclinical studies, RU-486 (mifepristone), a GR antagonist, impaired the recall of stressful memories in rats. In a recent pilot study conducted in older adults with anxiety disorder (GAD, PD, or anxiety disorder not otherwise specified) and co-occurring cognitive dysfunction, administration of mifepristone was associated with improvements in memory, executive functions, and severity of worry in patients with elevated cortisol levels at baseline. Further studies of glucocorticoids in humans are underway.

CORTICOTROPIN-RELEASING HORMONE

CRH, often referred to as corticotropin-releasing factor (CRF), is also a principal component of the response to stress, coordinating the adaptive behavioral and physiologic responses that occur after stress exposure. Driven by stress, hypothalamic levels of CRH are increased, and the HPA axis becomes active or perhaps even hyperactive, resulting in increased levels of cortisol and dehydroepiandrosterone (DHEA). CRH also exerts extrahypothalamic effects, which may be no less important than those mediated by the HPA axis. CRH-containing neurons are broadly distributed within the brain and can be found in the prefrontal and cingulate cortices, the CEA, the bed nucleus of the stria terminalis (BNST), the nucleus accumbens, the periaqueductal gray (PAG) matter, and brainstem nuclei such as the LC and dorsal and median raphe. The activation of amygdala CRH neurons triggers fear-related behaviors, and the engagement of cortical CRH neurons may lessen reward expectation. CRH is also involved in the regulation of diverse neurovegetative functions, such as food intake, sexual activity, and growth and reproduction. In most cases, the administration of CRH inhibits these behaviors, whereas the administration of CRH antagonists increases them.

It has been shown that early-life stress may bring about long-term alterations in brain CRH concentration. Furthermore, both CRH secretion and the individual response to heightened CRH levels apparently depend on numerous environmental factors such as current and/or past social environment, current and/or past trauma history, and behavioral dominance. Continuous high levels of both hypothalamic and extrahypothalamic CRH strongly contribute to psychobiological allostatic load. Preclinical studies have demonstrated that basal cerebrospinal fluid (CSF) CRH is higher in primates that have experienced early-life stress. The association between CRH and behavior occurs in a context-dependent

manner. Exposure to excessive levels of CRH in early life is associated with hippocampal damage later in life. Of interest, intracerebroventricular injection of CRH to the hippocampus produces an increase in MR levels, suggesting that CRH is an important regulator of HPA axis function. A recent study found that a SNP in the CRH gene was associated with behavioral inhibition, a childhood trait that predicts adult anxiety. Still, the exact functional impact of the SNP on the manifestation of the CRH gene is yet to be determined.

It has been suggested that CRH-1 and CRH-2 receptors have a major role in the regulation of anxiety. Both types of receptors are found in the pituitary and throughout the neocortex (especially in prefrontal, cingulate, striatal, and insular cortices), amygdala, and hippocampal formation in primate brain. However, other regions of the brain display predominantly one type of CRH receptor: CRH-1 receptors can be found within the LC, nucleus of the solitary tract, thalamus, and striatum, whereas CRH-2 receptors are chiefly located in the choroid plexus, certain hypothalamic nuclei, the nucleus prepositus, and the BNST. The varied distribution of these receptors suggests that each has a distinct role within the primate brain.

In rodents, CRH-1–deficient mice exhibit less anxiety-like behavior and a dampened response to stress. Conversely, CRH-2–deficient mice show more anxiety-like behavior and an exaggerated response to stress. Thus, the activation of CRH-1 receptors may contribute to increased anxiety-like responses, whereas the stimulation of CRH-2 may be pivotal in mediating anxiolytic behavior. Maintaining the appropriate functional equilibrium in the contribution of each of the two CRH receptor subtypes to brain CRH pathways might be essential to maintaining the protective and adaptive psychological and physiologic response to stressors. Although the study of CRH receptor function in humans is limited, genetic investigations have reported an association between nucleotide polymorphisms of the CRH-1 receptor gene and PD. There are also ongoing efforts to develop CRH receptor ligands for single-photon emission tomography (SPECT) and PET imaging. In light of the anxiogenic effects of CRH-1–receptor stimulation described previously, CRH-1 antagonists have been developed. To date, however, no CRH-1 receptor antagonist has demonstrated safety and efficacy in phase III trials for a stress-related psychiatric disorder. A large multicenter clinical trial failed to demonstrate superiority of the selective CRH-1 receptor antagonist pexacerfont over placebo in patients with GAD. Clinical trials of CRH-1 receptor antagonists verucerfont and emicerfont for SAD have been completed but results have not been published.

DOPAMINERGIC SYSTEM

Dopamine, a catecholamine, is released in response to acute stress in a number of specific brain areas important in the regulation of affective

behavior, including the basolateral nucleus of the amygdala, the nucleus accumbens, and the medial prefrontal cortex (mPFC). The dopamine innervation of the mPFC appears to be particularly vulnerable to stress. Thus, low-intensity stress (such as that associated with conditioned fear) or brief exposure to stress increases dopamine release and metabolism in this region even without any notable changes in other mesotelencephalic dopamine regions. This sensitivity of dopamine receptors in the mPFC is supported by animal studies showing that low-intensity electric foot shock increases in vivo tyrosine hydroxylase and dopamine turnover in the mPFC but not in the nucleus accumbens or striatum. In other areas receiving dopamine innervation, stress can enhance dopamine release and metabolism, provided that greater-intensity or longer-duration stress is applied. Thus, sensitivity and response of different brain regions to dopamine activation differ, with mPFC dopamine innervation activated prior to mesolimbic and nigrostriatal systems, and mesolimbic dopamine innervation more sensitive to stress than striatal dopamine innervation.

There is preclinical evidence that the extent of activation of the mesocortical dopamine system to stress may be, in part, genetically determined. It has been further suggested that extreme mesocortical dopamine release by stressful events may represent a vulnerability to depression and may lead to helpless reactions through an inhibition of subcortical dopamine transmission. Lesions of the amygdala before and after training in a conditioned stress model block stress-induced mPFC dopamine metabolic activation, suggesting amygdala control of stress-induced dopamine activation and a role in integrating the behavioral and neuroendocrine components of the stress response. On the other hand, lesions of mPFC dopamine neurons delay extinction (no effect on acquisition) of the conditioned fear response, indicating that prefrontal dopamine neurons are involved in facilitating the extinction of the fear response. One way to reconcile these two seemingly contradictory sets of data is to suggest that there is an optimal range for stress-induced increases in mPFC dopamine release to facilitate adaptive behavior responses. Too much mPFC dopamine release produces cognitive impairment, and an inhibition in nucleus accumbens dopamine activity results in abnormalities in motivation and reward mechanisms. Insufficient prefrontal cortical dopamine release delays the extinction of conditioned fear. In humans, preliminary studies have reported a reduced density of the dopamine transporter and D_2 receptors in patients with SAD. Notably, increased anxiety is among the symptoms usually associated with dopamine agonist withdrawal in patients treated for conditions like Parkinson's disease or restless legs syndrome.

The dopamine receptor gene (DRD2) and the dopamine transporter gene (DAT) have been previously linked to substance abuse, ADHD, and Tourette's syndrome. There has also been research showing a higher

association with phobic disorders of a polymorphism, Val158Met, of the enzyme catechol-*O*-methyltransferase (COMT), which catalyzes the degradation of the catecholamines NE, epinephrine, and dopamine. Although no overall association was found of COMT Val158Met in two meta-analyses of PD studies, there was an association of the 158Val allele reported in whites. In Asian populations, there was a trend toward association of the 158Met allele with PD. The clinical significance of these findings remains to be determined.

Only little clinical research on dopamine function in anxiety disorders has been performed. One research group found a higher concentration of the dopamine metabolite HVA in plasma in PD patients with high levels of anxiety and frequent panic attacks. PD patients were also shown to have a larger growth hormone response to the DAT agonist apomorphine than patients with depression. However, no direct alteration in CSF HVA level has been demonstrated in PD patients, and no correlations between anxiety severity and panic attacks have been found.

SEROTONERGIC SYSTEM

Serotonin (5-HT), the release of which is regulated by the dorsal raphe nuclei, is secreted in response to stress. Serotonin has been shown to have a dual role in anxiety. The serotonin hypothesis of anxiety states that 5-HT may be anxiogenic through its action on the PFC and amygdala, causing a heightened awareness to threats, and anxiolytic by its action on the dorsal periaqueductal gray (dlPAG), thereby inhibiting fight-or-flight behaviors. This model helps to explain the role of serotonin in producing panic and anticipatory anxiety, and excessive worry associated with generalized anxiety.

Various serotonin receptor subtypes have been shown to be involved in anxiety, in particular 5-HT_{1A} and 5-HT_{2A}. The 5-HT_{1A} receptor density is high in the cortex, hippocampus, amygdala, dlPAG, and raphe nuclei (presynaptic). Animal models initially suggested a role of 5-HT_{1A} receptors in the development of chronic anxiety and helped to generate animal models of anxiety-related disorders. However, it is not clear whether these models constitute relevant models for PD in humans. The development of highly selective 5-HT_{1A} receptor radioligands such as [^{18}F]*trans*-4-fluoro-*N*-2-[4-(2-methoxyphenyl)piperazin-1-yl]ethyl]-*N*-(2-pyridyl)cyclohexanecarboxamide (FCWAY) made it possible to assess central 5-HT_{1A} receptor binding in PD using PET. In PD patients, including those who also had depression, receptors were reduced by an average of nearly one-third in the anterior cingulate in the front middle part of the brain, the posterior cingulate, the rear middle part of the brain, and the raphe, as well as in the midbrain.

Preclinical studies of the genetic expression of the 5-HT_{1A} receptor have

shown that rodents lacking the receptor exhibit higher levels of anxiety. The opposite has also been shown in mice that have greater expression of 5-HT_{1A}. Of interest, animal studies using models of early-life stress, such as maternal separation, have shown that alterations to the 5-HT_{1A} receptor that occur early in life, such as the prenatal period, can cause persisting changes to the anxious phenotype, even after the restoration of the receptors at a later time. There is no such long-term effect if the receptors are reduced, and then reinstated later, in adult animals. These findings may be due to the effects of increased CRH and cortisol caused by the early stress, which thereby downregulate the number of 5-HT receptors. There may also be a genetic component. Studies in humans are using 5-HT_{1A}–binding ligands with neuroimaging to determine receptor density; they have reported reduced receptor binding in patients with PD and SAD. Interestingly, antidepressants may exert their anxiolytic effects through different receptor subtypes depending on the anxiety disorder involved. Chronic treatment with SSRIs in PD patients is thought to suppress panic attacks by enhancing the responsiveness of 5-HT_{1A} and 5-HT_{2A} receptors in the dorsal PAG matter, while in patients with generalized anxiety, SSRIs might act through the desensitization of 5-HT_{2C} receptors and, possibly, through stimulation of 5-HT_{1A} receptors in the forebrain.

One method of ascertaining the role of 5-HT in anxiety disorders has been to use pharmacologic challenges such as tryptophan depletion (TD), in which an amino acid mixture deficient in l-tryptophan, a precursor of serotonin, is administered, thereby resulting in an acute decrease in tryptophan and thus a decrease in brain serotonin levels. TD has also been shown to trigger depressive symptoms even in individuals with only a family history of depression. While more consistent in individuals with a history of depression, the results of TD challenges in individuals with anxiety disorders have been mixed. In studies of PD patients, TD did reduce tryptophan levels and increased anxiety, and was associated with a heightened panicogenic effect of CO_2 challenge when compared to controls. There are, however, studies of patients with PD in which TD did not produce anxiety, including a study using a cholecystokinin-4 (CCK-4) panic challenge. Another pharmacologic challenge of the serotonin system involves the 5-HT agonist *meta*-chlorophenylpiperazine (mCPP), used as a probe of 5-HT_2 receptor. Some but not all studies of oral and intravenous mCPP have shown increased anxiety in PD after the administration of the drug. Patients with SAD on SSRIs were also reported to have greater sensitivity to TD challenges, whereas another study found that mCPP did not induce a greater rate of panic attacks than in healthy controls.

Serotonergic agents, such as SSRIs, serotonin–norepinephrine reuptake inhibitors (SNRIs; venlafaxine [Effexor] and duloxetine [Cymbalta]), mirtazapine (Remeron), and buspirone (Buspar) remain effective

treatments for PD, social phobia, and GAD. The primary focus of investigation is on determining the underlying differences in treatment responses across patients, using assays of receptor binding and genetic differences. The serotonin transporter (5-HTT) gene and its polymorphism (5-HTTLPR) have been widely studied. Preclinical studies and studies in healthy volunteers have reported that the presence of two short alleles of the 5-HTTLPR is correlated with greater shyness and anxiety-like behaviors. Studies in patients with anxiety disorders have been equivocal. A PET study reported greater severity of anxiety symptoms in patients with SAD and increased right amygdala blood flow during public speaking compared to healthy controls. In PD, there are several studies showing no significant link between the 5-HTTLPR and development of the disorder or greater disease severity. Of interest, the 5-HTTLPR polymorphism has been associated with the efficacy of antidepressants. For example, one group reported that the efficacy of paroxetine (Paxil) in a sample of PD patients was dependent on allelic variability but that this association was found only in women. The clinical response to SSRIs also appears to be influenced by a polymorphism in the gene coding for RGS2, a protein involved in neuronal G-protein–coupled receptor responses to neurotransmitters (including 5-HT). It has been reported that reduced expression of RGS2 is associated with a poor response to sertraline treatment. Whether there are other genetic or neurochemical factors related to 5-HT that contribute to the therapeutic effect of anxiolytic medications will require more extensive investigation.

γ-AMINO BUTYRIC ACID AND THE BENZODIAZEPINE SYSTEM

GABA is the primary inhibitory neurotransmitter in the brain and a metabolic product of glutamate, the excitatory neurotransmitter. There are two subtypes of GABA receptors: GABA-A and GABA-B. GABA-B receptors are slow-acting channels and are agonized by agents such as baclofen (a drug used as a muscle relaxant). GABA-A receptors are fast-acting, ligand-gated channels made up of five protein subunits and consisting of multiple families of receptors. There is a wide range of drugs that act as agonists of the GABA-A receptor, including barbiturates, benzodiazepines, alcohol, anesthetics, neuroactive steroids, and several anticonvulsant medications. Not surprisingly, all of these agents have been shown to have anxiolytic effects. In particular, benzodiazepines have long been used in the treatment of anxiety disorders, in particular PD, phobias, and GAD.

Central benzodiazepine receptors have also been found in the human brain. They form a complex with GABA-A receptors and have different binding sites but positive allosteric effects on one another. The presence of these benzodiazepine receptors led to reports of endogenous benzodiazepine receptor ligands, also called endozepines, in the brain, and speculations that they might function in the modulation of fear and

learning. The exact role of these compounds in anxiety, however, is yet to be fully understood.

The role of GABA and benzodiazepine receptors in anxiety is well documented in animal studies. It has been reported that benzodiazepine partial agonists reduce anxiety-like behaviors in rats, whereas GABA inverse agonists produce anxiety in certain learning and behavioral tasks. There is also ample supporting evidence in humans. Studies using benzodiazepine antagonists have shown these agents to be anxiogenic. Flumazenil (Romazicon), an antagonist used in emergency settings to treat benzodiazepine overdose, has been shown to be a panicogenic agent in patients with PD but not in healthy controls. Clinical studies of GABA levels in the blood of patients with anxiety disorders have been limited. Studies of patients with PD compared to healthy controls have failed to show a change in plasma GABA at baseline and after a CCK-4 challenge.

Neuroimaging studies have enabled the visualization of GABA and benzodiazepine receptors in the human brain. Several studies of PD patients have shown reduced GABA-A and benzodiazepine binding, using PET and SPECT imaging, in areas such as the cortex, hippocampus, and insula when compared to controls. A study of PD patients taking medications (antidepressants and/or benzodiazepines), using proton magnetic resonance spectroscopy, showed reduced GABA levels in the anterior cingulate and basal ganglia. Although benzodiazepines continue to be among the most widely prescribed medications worldwide, they are no longer considered first-line treatment for PD, SAD, or GAD, having been replaced with SSRIs, cognitive–behavioral therapy, and other types of therapy. Part of this change in prescribing practices is due to concerns about the risk of tolerance, the potential for abuse or dependence, sedative properties, and potential long-term effects on cognition and memory. In spite of these concerns, there has been consistent evidence that benzodiazepines are safe and effective in the treatment of anxiety when prescribed appropriately.

Additionally, there has been a shift toward testing newer agents that enhance GABA in anxiety disorders. These drugs include several anticonvulsants that work in part as GABA agonists, including tiagabine (Gabitril), a selective GABA reuptake inhibitor (SGRI) and vigabatrin, an inhibitor of GABA transaminase (the latter is not approved for use in the United States). Controlled studies of tiagabine reported equivocal findings in GAD patients. Results from clinical trials of other anticonvulsants, such as topiramate (Topamax), in PD have been mixed. Further studies, using larger sample sizes, are indicated before the efficacy of these agents as primary treatments for anxiety disorders can be fully ascertained. Gabapentin and pregabalin, although structurally related to GABA, do not act on GABA receptors but instead bind to a subunit of voltage-sensitive calcium channels.

Neurosteroids, such as ALLO and pregnanolone (see separate section below), have been demonstrated to act as modulators of GABAergic neurotransmission mainly through a stereoselective interaction with GABA-A receptors. In animal models, studies of a synthetic analog of allopregnanolone (ALLO) reported an anxiolytic effect similar to benzodiazepines in the absence of tolerance phenomena. Targeting GABA-A receptors, and in particular the development of compounds selective for GABA-A receptor subtypes, appears to be an intriguing pharmacologic approach for the treatment of anxiety disorders. This approach, also taking advantage of the rapid turnover of GABA-A receptors, could yield a potential drug with high selectivity and reduced adverse effects. To date, results from animal trials appear promising, but further investigations in patients with anxiety disorders are required.

GLUTAMATE SYSTEM

Glutamate is the primary excitatory neurotransmitter of the CNS, is the precursor compound of GABA, and modulates pathways that mediate panic and defensive responses. Results from human and animal studies suggest that panic might arise from an imbalance between tonic GABAergic inhibition and glutamate-mediated excitation, with predominance of the latter. For example, reduced GABAergic tone has been reported in patients with PD. Further, glutamate is known to regulate other neurotransmitters involved in PD, including serotonin and NE.

Glutamate binds to two families of receptors: metabotropic and ionotropic receptors. Metabotropic receptors (mGluRs), which are G-protein–coupled membrane receptors, are made up of eight subtypes and are believed to serve as allosteric modulators. Ionotropic receptors, which are ligand-gated channels, include *N*-methyl-D-aspartate (NMDA), α-amino-3-hydroxy-5-methylisoxasole-4-propionic acid (AMPA), and kainate receptors. These receptors have become a focus of investigation in the area of anxiety disorders, primarily PD.

Preclinical studies have shown anxiolytic properties of group-2 and group-3 mGlu-receptor agonists, and more recently of selective group-2 receptor type 2 (mGluR2) agonists. Systemic infusion of these glutamate receptor agonists prevented panic-like responses after infusion of sodium lactate in panic-vulnerable rats, showing efficacy against panic similar to that of comparison drug alprazolam in this animal model of panic. With regard to human studies, a small randomized, double-blind clinical trial of an mGluR2/3 agonist, LY354740, in patients with PD, failed to show treatment effects compared to placebo. A more recent randomized, double-blind, placebo-controlled trial of the same compound in patients with GAD showed greater improvement in anxiety scores compared to placebo, but the study was discontinued due to findings of convulsions in animal studies of the compound. This is an understandable risk, given that glutamate-

enhancing drugs have neuroexcitatory properties and that glutamate inhibitors are commonly used as anticonvulsants. Similarly, there are studies of mGlu1, mGlu5, and mGluR7 antagonists, and AMPA/kainate–receptor blockers, showing antianxiety properties in animals. Further studies in humans are needed before the safety and tolerability of these compounds can be better understood.

NMDA receptors are prominently involved in the conditioning of fear learning, via entry of calcium through the receptors and voltage-gated calcium channels. Blockade of the NMDA receptor with antagonists such as (2R)-amino-5-phosphonovaleric acid (APV) can block fear acquisition, and some studies show that it may even block expression. In panic-prone rats, injection of NMDA antagonist MK-801 into the dorsomedial/perifornical hypothalamus blocked panic-like behavior induced by sodium lactate in a dose-dependent fashion. The NMDA receptor antagonist ketamine, which at subanesthetic doses has demonstrated efficacy in patients with treatment-resistant depression, has shown preliminary efficacy for PTSD in a recent study, and mixed results in patients with obsessive–compulsive disorder, both conditions formerly classified as anxiety disorders. Ketamine has not yet been studied in patients with conditions currently classified as anxiety disorders.

NMDA receptors are also critical for fear extinction, with preclinical studies showing that blockade of this pathway inhibits extinction. These findings were supported by animal research showing that a partial agonist at the glycine modulatory site on the NMDA receptor, D-cycloserine (DCS), a drug previously used for the treatment of tuberculosis, could facilitate extinction learning. Healthy volunteers receiving DCS during a stressful learning task had less stress and better performance than those receiving placebo. In addition, a small study using functional magnetic resonance imaging showed that healthy volunteers receiving DCS had a more blunted amygdala response to fearful faces compared to those receiving placebo. Several studies have shown the efficacy DCS as an adjunct to exposure therapy in patients with acrophobia (fear of heights), SAD, and PD. Negative studies have also been reported, including a study of DCS in spider phobia. Another potential strategy consists of inhibiting the glycine transporter, thereby increasing the availability of glycine at the NMDA receptor complex glycine modulatory site. Glycine transporter inhibitor Org 25935 was evaluated in a multicenter clinical trial as augmentation to cognitive behavioral therapy in patients with PD; panic symptoms decreased in every treatment group, but there was no statistically significant difference between Org 25935 and placebo.

Other glutamate-modulating agents, such as pregabalin, tiagabine, and lamotrigine (Lamictal), have been studied for GAD and SAD. Although pregabalin demonstrated efficacy in several randomized controlled trials in patients with GAD, it has not received formal FDA approval for the

treatment of this condition. Pregabalin, however, is recommended as a first line of treatment for GAD comorbid with epilepsy. Riluzole (Rilutek), a drug approved for amyotrophic lateral sclerosis, was studied in an 8-week, open-label trial in patients with GAD and found to be effective in lowering anxiety scores. The study of glutamate in larger-scale human studies should continue to provide novel therapeutic agents for anxiety disorders.

NEUROSTEROIDS

Endogenous steroids acting on neuronal function are called neuroactive steroids, or neurosteroids, and regulate neuronal excitability. They are highly selective and positive allosteric modulators of the GABA-A receptor, thereby making them potential anxiolytic agents, a finding that has been supported in animal models. Neurosteroids are thought to act on the amygdala by affecting intracellular receptors in a process involving NMDA receptors. They additionally regulate CRH gene expression, thereby reversing the anxiogenic effects of CRH. The progesterone metabolites believed to be involved in fear and anxiety include 3α,5α-tetrahydroprogesterone (THP) (ALLO) and 3α,5β-THP (pregnanolone).

Preclinical studies have shown that progesterone is anxiolytic, likely due to its effect of increasing ALLO. Progesterone is increased during pregnancy, and it has been shown in animal studies that this is associated with a reduction in anxiety behaviors and that this anxiolytic effect is reversed by finasteride, an inhibitor of the 5α-reductase enzyme that is involved in steroid biosynthesis. One study in humans reported that pregnant women had higher levels of progesterone, ALLO, and tetrahydrodeoxycorticosterone (THDOC), which were believed to be protective against maternal anxiety. A study of pregnant women with pre-existing PD found a reduction in anxiety symptoms during the pregnancy and rebound anxiety during the postpartum period, which is consistent with a rapid drop in hormone levels. The effect of hormone levels on anxiety has been studied during the menstrual cycle as well. One study reported that women with PD had elevated peripheral concentrations of ALLO both during the follicular and premenstrual phase, and suggested that neurosteroids may be a protective factor against spontaneous panic attacks.

Patients with PD have been found to have increased ALLO levels and decreased isopregnanolone levels. In a study of panic attacks induced by CCK-4 and lactate, patients with PD showed a decrease in ALLO and pregnanolone, suggesting decreased GABA tone not found in healthy controls. It is possible that alterations of neurosteroid levels might be related to the pathophysiology of panic attacks in PD and not simply reflect state anxiety. No data are available concerning alteration of neurosteroid levels during spontaneous panic attacks. Studies investigating the impact of pharmacotherapy on ALLO levels in PD patients revealed conflicting

results, with one trial showing no effect of paroxetine on ALLO levels or concentrations of other neuroactive steroids, but another one showing significant increase in ALLO levels after 2 months of treatment with paroxetine.

Studies of neurosteroid levels in patients with other anxiety disorders have yielded inconsistent findings. Two small-scale investigations reported low pregnanolone sulfate levels in both SAD and GAD, in both cases when compared to healthy controls. One group, however, demonstrated no difference in levels of DHEA, DHEA sulfate (DHEAS), or ALLO in patients with SAD compared to healthy volunteers. There have been no investigations of neurosteroids as potential therapeutic agents for anxiety disorder. Studies of synthetic analogs of neurosteroids are under way, with no published studies in anxiety disorders to date. Future investigations will have to consider possible confounding factors such as age, sex, or other factors in the evaluation of blood and CSF levels of these various neurosteroids.

ARGININE VASOPRESSIN

Arginine vasopressin (AVP) is released by the parvocellular neurons found in the paraventricular nucleus (PVN). AVP is best known for its effects on the control of body fluid levels, blood pressure, and blood sugar. Vasopressin is used as a treatment for nocturnal enuresis as well. It has also been linked to pain sensitivity, circadian rhythms, regulation of sleep, and learning and cognition. AVP exerts its effect through two receptors, V1a and V1b, located throughout the brain, including the septum, cortex, and hippocampus.

Although CRH is believed to be the critical factor in controlling the release of ACTH, AVP serves as an important mediator. The release of ACTH during periods of acute stress, as in rat studies using the forced swim test, can be suppressed by simultaneous administration of a CRH receptor 1 antagonist and an AVP V1b antagonist. Mice with the V1a receptor knock-out showed reduced anxiety on similar stress tasks. Similarly, mice with overexpression of the V1a receptor exhibited heightened anxiety-like behaviors. A human genetic study found an association between an SNP within the AVP V1b receptor gene and PD, compared with healthy controls. With regards to treatment, preclinical studies have yielded mixed results of AVP V1b antagonists on anxiety-like behaviors. Clinical studies investigating a V1b antagonist in patients with GAD failed to demonstrate clinical efficacy. Studies of new AVP V1b receptor antagonists are currently under way.

In addition, there are limited reports of AVP administration in anxiety disorders. A trial in healthy male and female volunteers found that the administration of intranasal AVP reduced skin conductance and heart rate responses during an angry-faces paradigm, and that there was a difference

between men and women in the social recognition paradigm, supporting the preclinical findings of sex differences in AVP.

OXYTOCIN

Oxytocin, a nine–amino acid peptide hormone synthesized in the paraventricular and supraoptic nuclei of the hypothalamus, and released from neurons within the brain and also into the bloodstream via the posterior pituitary, has been shown to interact with central reward processing regions, modulate social emotional functions, and attenuate stress responses via interaction with the HPA axis. The anxiolytic, stress-attenuating, and prosocial effects of oxytocin were first discovered in rodents.

In human studies, oxytocin administered intranasally has been associated with increased prosocial behavior, and improvement in emotion recognition and social memory. Its anxiolytic and stress-attenuating effects have stimulated significant interest in its potential therapeutic effect for anxiety disorders. A study using intranasal oxytocin in healthy male volunteers reported reduced amygdala activation to fearful faces, lower anxiety scores, and improved social cognition compared to intranasal placebo. In another study of men with early parental separation, a risk factor for anxiety and depressive disorders in adults, intranasal oxytocin administration attenuated decreases in salivary cortisol compared to placebo, thus dampening the HPA axis alteration associated with early separation. Further, a recent study reported associations between polymorphisms in the oxytocin receptor gene and increased risk for anxiety, stress, and depression in individuals with early-life stress exposure.

A series of studies in healthy and anxious volunteers has shown decreased anxiety during the Trier Social Stress Test (TSST) after intranasal oxytocin administration. In one study of patients with GAD, daily oxytocin administration over 3 weeks was beneficial, especially in male patients. Oxytocin facilitates extinction of social fear. In patients with SAD, intranasal oxytocin alone or as augmentation of exposure therapy was associated with more positive self-evaluation after public performance. In a recent neuroimaging study in patients with SAD, oxytocin administered intranasally was associated with greater enhancement of amygdala-frontal resting state connectivity in patients with higher social anxiety. Oxytocin also enhanced functional connectivity between the amygdala and other brain regions (insula, cingulate gyrus) in patients with SAD but not in healthy controls, in response to fearful faces. A recent study in patients with SAD and healthy controls found that decreased methylation of the oxytocin receptor gene was associated with SAD, increased cortisol response to the TSST, and increased amygdala reactivity during processing of words relevant to social anxiety. Further studies of oxytocin

administration in patients with SAD are necessary to develop this intervention for potential use in clinical practice.

NEUROPEPTIDE Y

Neuropeptide Y (NPY), a 36–amino acid peptide, is one of the most abundant and highly evolutionarily conserved polypeptides found in the brain. Its highest concentrations are in the LC, hypothalamus, septum, and periaqueductal gray, with moderate levels in the hippocampus, amygdala, and brainstem. NPY is linked with various roles in humans, such as feeding and sleep regulation. There is also ample preclinical evidence that NPY has anxiolytic effects. Of the five NPY receptor subtypes found in mammals (Y1 to Y5), the NPY-Y1 receptor has been most closely studied in anxiety. Studies of Y-1–receptor knockout mice and administration of Y-1 antagonists increased anxiety-like behaviors in the animals. In addition, NPY has been shown to be involved in fear consolidation, with preclinical studies showing that the administration of NPY impairs the retention of traumatic memories and reduces anxiety during stressful tasks. NPY also mediates the response to chronic stress by increasing expression of amygdala NPY mRNA. A study of rats exposed to maternal separation reported that treatment with acupuncture increased NPY expression in the basolateral amygdala and reduced anxious behaviors. These data suggest that NPY may play a role in the psychobiology of stress responses and resilience.

Neuropeptide Y has yet to be studied as a therapeutic agent in anxiety disorders and has only been investigated in small trials in healthy volunteers. Administering NPY via an intranasal route may be the most clinically relevant procedure. In a recent preclinical study, rats administered intranasal NPY before a single prolonged stress showed lower anxiety- and depressive-like behaviors than those administered placebo. Further studies in humans are needed to determine whether this route of administration can result in clinically significant levels in the CNS and whether NPY will be useful as a potential treatment for anxiety disorders.

GALANIN

Galanin, a 30–amino acid peptide, has a wide variety of functions in humans, including modulation of pain, learning, feeding and gastrointestinal motility, and cardiovascular control. Galanin, which is coexpressed with NE neurons in the LC, and thus projects to structures involved in the fear pathway (including the amygdala, hippocampus, and PFC), has been linked to regulation of mood and fear response. It also acts as an autoreceptor to reduce firing of the LC. Galanin acts on three subtypes of receptors: Gal-R_1, Gal-R_2, and Gal-R_3. Although preclinical data are inconsistent, the three receptors have been shown to have differing

effects on anxiety-like behaviors. Mice lacking Gal-R_1 receptor exhibited anxiety-like behaviors, whereas the injection of receptor agonists produced a reduction of such behaviors. Similarly, studies of Gal-R_3 antagonists in rodents showed them to have antidepressant and anxiolytic properties in several behavioral tasks. Overall, increases in central galanin, either by overexpression or via intracerebral administration, have been demonstrated to regulate anxiety behaviors in rats. More recent studies suggest that galanin might have anxiolytic effects under high stress/high noradrenergic activation. A trial in mice using galnon, a nonselective galanin-receptor agonist, showed that the drug reduced anxiety during certain stressful tasks, such as the forced swim test and the elevated zero maze, and that this effect was reversed by administration of a galanin-receptor antagonist. This finding has also been reported in another animal study using forced swim, a model of depression. The reports suggest that galanin's effects on mood and anxiety may be through its inhibition of hippocampal serotonin receptors.

There are very few human studies of galanin in patients with anxiety disorders. A genetic study of men and women with PD found an association between two of six haplotypes of a galanin gene polymorphism and severity of illness, but only in women; this finding was later replicated in a larger sample. Despite the therapeutic potential of galanin, there are no known published trials of galanin modulators as potential therapeutic agents in anxiety disorders.

CHOLECYSTOKININ

CCK is a neuropeptide derived from a 112–amino acid precursor compound and is classified by its various fragment peptides, of which CCK octapeptide (CCK-8) is the most abundant. CCK acts via two types of receptors, CCK-A and CCK-B (also referred to as CCK_1 and CCK_2, respectively). CCK-8 acts on both receptor subtypes, whereas CCK tetrapeptide (CCK-4) is a primary agonist of CCK-B receptors. CCK has its highest levels in the cerebral cortex, hippocampus, amygdala, caudate, and putamen, with intermediate levels in the thalamus and hypothalamus. It is also present in the gastrointestinal tract. CCK receptors are found in greatest density in the cortex, hypothalamus, substantia nigra, and PAG.

It has been shown in various preclinical studies that CCK-4 is anxiogenic and that this effect is blocked by CCK-B antagonists. CCK-8 has shown a similar effect on inducing anxiety in rats, but only after intracranial administration, suggesting that the drug does not cross the blood–brain barrier as effectively as CCK-4. Studies in humans have also shown similar results. Studies of both healthy subjects and patients with PD have shown the panicogenic effects of CCK-4 and of CCK-5, or pentagastrin. These effects could be blocked by the administration of CCK

antagonists or by treatment with GABA agonists such as vigabatrin or tiagabine. In addition, patients with PD have been shown to have lower CSF levels of CCK versus healthy controls, although it is unclear whether this is due to higher CCK turnover or greater receptor sensitivity. In addition, a recent study of healthy volunteers undergoing a CCK-4 challenge reported that cardiovascular changes may not be correlated with subjective panic, which suggests that this may not be an ideal model for studying the effect of anxiolytic drugs on panic unless strict criteria are used.

The aforementioned studies supported the role of CCK in anxiety and prompted the investigation of CCK-B–receptor antagonists as potential antianxiety drugs. In a randomized, double-blind, placebo-controlled study, however, the effect of CI-988, a CCK-B antagonist, on anxiety scores in GAD patients did not differ from placebo. Another study in GAD patients found that CI-988 failed to abate the anxiogenic effects of mCPP, and a similar study showed that it failed to treat the anxiety symptoms induced by CCK-4 in patients with PD. Trials of CCK-B antagonists in PD showed that they were not significantly better than placebo. This has led to the suggestion that the preclinical data were not as clear-cut in defining the significance of CCK-B in anxiety modulation. The role of CCK as a possible target of novel therapeutics remains an area of investigation.

ENDOCANNABINOID SYSTEM

Preclinical studies have stimulated a growing interest in the endocannabinoid system, involved in modulating sleep, appetite, pain, emotional memory, and fear responses, as a potential therapeutic target in patients with anxiety disorders. This system comprises polyunsaturated fatty acid endogenous ligands, anandamide (AEA), and 2-arachidonoylglycerol (2-AG), and two G-coupled receptors, cannabinoid receptors 1 (CB1) and 2 (CB2). CB1 and CB2 receptors are widely distributed in the frontal cortex and subcortical structures, including amygdala, basal ganglia, hippocampus, and other structures.

Findings from preclinical studies suggest that basal endocannabinoid levels are low and increase during exposure to threatening stimuli, counteracting fear responses under conditions of threat. In humans, new studies are focusing on targeting this system in patients with PTSD, no longer classified as an anxiety disorder. Cannabidiol (CBD), a Cannabis sativa constituent, binds to the CB1 receptor and several others, including the serotonin 5-HT_{1A} receptor. CBD has been studied in animal models of general anxiety and stress-induced anxiety. In human studies, CBD significantly reduced anxiety during simulated public speaking in healthy volunteers and in patients with SAD. Conversely, rimonabant, a CB1 antagonist/inverse agonist, increased anxiety in anticipation and during simulated public speaking in healthy volunteers. A recent study in healthy

volunteers implicated polymorphisms of the CB1 receptor gene and failure to extinguish fear.

While direct CB receptor modulators might potentially worsen anxiety or depression, inhibitors of fatty acid amide hydrolase (FAAH), the principal catabolic enzyme for AEA, show promise. Reduced FAAH expression has been associated with enhanced extinction learning in animal and human neuroimaging studies. Knock-in mice with a polymorphism associated with reduced expression and activity of FAAH exhibited enhanced fronto-amygala connectivity, increased fear extinction learning, and decreased anxiety-like behaviors. Further, administration of FAAH inhibitors to rodents was associated with decreased anxiety-like behaviors during experimental tasks designed to increase anxiety, such as the elevated-plus maze and the light-dark box. Clinical trials of these novel compounds are needed in social anxiety and other anxiety disorders.

Endocannabinoids implication with social anxiety is in accordance with dense expression of CB1 receptors in brain regions related to anxiety, fear, and aversion, including the mPFC, hippocampus, amygdala, and PAG.

Preclinical studies showed that anxiogenic-like effects of CB1 antagonists tend to be more evident when animals are subjected to high levels of aversion. Anandamide-hydrolysis inhibitors are more efficacious as anxiolytic drugs, when tested in a highly aversive environment. The basal levels of endocannabinoid synthesis and release tend to be low; however, the activity of this system is enhanced in response to neural activation when experimental animals are exposed to threatening stimuli, when endocannabinoids would work to counteract. Endocannabinoids implication with social anxiety is in accordance with dense expression of CB1 receptors in brain regions related to anxiety, fear, and aversion, including the mPFC, hippocampus, amygdala, and PAG.

FUTURE DIRECTIONS

Knowledge about anxiety disorders arose from various animal models of stress and fear, using primarily Pavlovian fear conditioning as a basis of understanding. Although such studies have provided invaluable information about the mechanisms of the fear response, translating this work to humans remains challenging. Anxiety disorders are highly unique in their etiology, susceptibility, presentation, treatment, and prognosis. Preclinical research will continue to refine the current understanding of the role and integration of neurochemical systems both in producing fear responses and in modulating anxiety responses in animal models. Additional progress will come from optimizing current animal models of anxiety. Further, animal studies are employing novel techniques including in vivo chemogenetics and optogenetics, coupled with retrograde tracing to probe neurobiologic systems involved in anxiety. New insights from basic research can then be applied to human studies in the search for novel

biomarkers of anxiety disorders and new treatment approaches.

Future research in humans will benefit from continued fine-tuning of intermediate phenotypes for anxiety disorders, yielding a more refined clinical categorization of disorders and facilitating translational research. The last decade has seen the emergence of a new investigational approach using target-selective agents to enhance the benefits of exposure therapy. Mechanistically novel compounds targeting glutamate, neuropeptide, and endocannabinoid systems are already under development and will continue to be tested in clinical trials. In addition to the neurochemical systems reviewed in this chapter, potentially promising avenues for future research include the neurotrophin, renin–angiotensin, and acetylcholine systems, among others. Finally, a new wave of studies in humans that promises to expand the current understanding of anxiety disorders combines neural circuit analyses with burgeoning work in psychiatric genomics, with the goal of yielding novel and more specific biomarkers of anxiety disorders, and ultimately more tailored treatment interventions for these highly common and debilitating mental disorders.

References

Blaya C, Salum GA, Lima MS, Leistner-Segal S, Manfro GG. Lack of association between the serotonin transporter promoter polymorphism (5-HTTLPR) and panic disorder: a systematic review and meta-analysis. *Behav Brain Funct*. 2007;3:41.

Bukalo O, Pinard CR, Holmes A. Mechanisms to medicines: elucidating neural and molecular substrates of fear extinction to identify novel treatments for anxiety disorders. *Br J Pharmacol*. 2014;171:4690.

Cameron OG, Huang GC, Nichols T, Koeppe RA, Minoshima S. Reduced gamma-aminobutyric acid(A)-benzodiazepine binding sites in insular cortex of individuals with panic disorder. *Arch Gen Psychiatry*. 2007;64:793.

de Kloet CS, Vermetten E, Geuze E, Wiegant VM, Westenberg HG. Elevated plasma arginine vasopressin levels in veterans with posttraumatic stress disorder. *J Psychiatr Res*. 2008;42:192.

Dunsmoor JE, Paz R. Fear generalization and anxiety: behavioral and neural mechanisms. *Biol Psychiatry*. 2015;78:336.

Farb DH, Ratner MH. Targeting the modulation of neural circuitry for the treatment of anxiety disorders. *Pharmacol Rev*. 2014;66:1002.

Gatt JM, Burton KL, Williams LM, Schofield PR. Specific and common genes implicated across major mental disorders: a review of meta-analysis studies. *J Psychiatr Res*. 2015;60:1.

Graeff FG, Zangrossi H Jr. The dual role of serotonin in defense and the mode of action of antidepressants on generalized anxiety and panic disorders. *Cent Nerv Syst Agents Med Chem*. 2010;10:207.

Griebel G, Holmes A. 50 years of hurdles and hope in anxiolytic drug discovery. *Nature Rev*. 2013;12:667.

Guastella AJ, Richardson R, Lovibond PF, et al. A randomized controlled trial of D-cycloserine

enhancement of exposure therapy for social anxiety disorder. *Biol Psychiatry*. 2008;63:544.

Harro J. CCK and NPY as anti-anxiety treatment targets: promises, pitfalls, and strategies. *Amino Acids*. 2006;31:215.

Harvey BH, Shahid M. Metabotropic and ionotropic glutamate receptors as neurobiological targets in anxiety and stress-related disorders: focus on pharmacology and preclinical translational models. *Pharmacol Biochem Behav*. 2012;100:775.

Hasler G, Nugent AC, Carlson PJ, Carson RE, Geraci M, Drevets WC. Altered cerebral gamma-aminobutyric acid type A-benzodiazepine receptor binding in panic disorder determined by [11C]flumazenil positron emission tomography. *Arch Gen Psychiatry*. 2008;65:1166.

Johnson PL, Fitz SD, Engleman ED, Svensson KA, Schkeryantz JM, Shekhar A. Group II metabotropic glutamate receptor type 2 allosteric potentiators prevent sodium lactate-induced panic-like response in panic-vulnerable rats. *J Psychopharm*. 2013;27:152.

Kalk NJ, Nutt DJ, Lingford-Hughes AR. The role of central noradrenergic dysregulation in anxiety disorders: evidence from clinical studies. *J Psychopharmacol*. 2011;25:3.

Kormos V, Gaszner B. Role of neuropeptides in anxiety, stress, and depression: from animals to humans. *Neuropeptides*. 2013;47:401.

LeDoux JE. Coming to terms with fear. *Proc Natl Acad Sci USA*. 2014;111:2871.

Linares IM, Chags MH, Machado-de-Sousa JP, Crippa JA, Hallak JE. Neuroimaging correlates of pharmacological and psychological treatments for specific phobia. *CNS Neurol Disord Drug Targets*. 2014;13:1021.

Long Z, Medlock C, Dzemidzic M, Shin YW, Goddard AW, Dydak U. Decreased GABA levels in anterior cingulate cortex/medial prefrontal cortex in panic disorder. *Prog Neuropsychopharmacol Biol Psychiatry*. 2013;44:131.

Morrison FG, Ressler KJ. From the neurobiology of extinction to improved clinical treatments. *Depress Anxiety*. 2014;31:279.

Murrough JW, Yaqubi S, Sayed S, Charney DS. Emerging drugs in the treatment of anxiety. *Expert Opin Emerg Drugs*. 2015;20:393.

Nash JR, Sargent PA, Rabiner EA, et al. Serotonin 5-HT1A receptor binding in people with panic disorder: positron emission tomography study. *Br J Psychiatry*. 2008;193:229.

Neumann ID, Slattery D. Oxytocin in general anxiety and social fear: a translational approach. *Biol Psychiatry*. 2016;79(3):213–221.

Perna G, Schruers K, Alciati A, Caldirola D. Novel investigational therapeutics for panic disorder. *Expert Opin Investig Drugs*. 2015;24:491.

Preter M, Klein DF. Panic, suffocation false alarms, separation anxiety and endogenous opioids. *Prog Neuropsychopharmacol Biol Psychiatry*. 2008;32:603.

Santos M, D'Amico D, Dierssen M. From neural to genetic substrates of panic disorder: Insights from human and mouse studies. *Eur J Pharmacol*. 2015;759:127.

Schüle C, Nothdurfter C, Rupprecht R. The role of allopregnanolone in depression and anxiety. *Prog Neurobiol*. 2014;113:79.

Singewald N, Schmuckermair C, Whittle N, Holmes A, Ressler KJ. Pharmacology of cognitive enhancers for exposure-based therapy of fear, anxiety and trauma-related disorders. *Pharmacol Ther*. 2015;149:150.

Unschuld PG, Ising M, Erhardt A, et al. Polymorphisms in the galanin gene are associated with

symptom-severity in female patients suffering from panic disorder. *J Affect Disord*. 2008;105:177.

Van't Veer A, Carlezon WA Jr. Role of kappa-opioid receptors in stress and anxiety-related behavior. *Psychopharmacology (Berl)*. 2013;229:435.

▲ 14.6 Neuroimaging and the Neuroanatomical Circuits Implicated in Anxiety, Fear, and Stress-Related Disorders

DYLAN G. GEE, PH.D., AND B.J. CASEY, PH.D.

Knowledge about the neuroanatomical, cellular, and molecular correlates of fear and anxiety has increased dramatically with the development of neurophysiological, lesion, and molecular techniques for investigating emotional behavior in experimental animals and of neuropharmacological and neuroimaging approaches for elucidating the biological correlates of fear and anxiety in humans. The convergent information obtained through these approaches has illuminated the neural circuitry underlying the mediation and modulation of adaptive fear and anxiety. A growing literature has begun to elucidate how dysfunction within these circuits may result in pathological anxiety syndromes such as panic and phobic disorders, as well as the neurodevelopmental processes underlying risk for pediatric anxiety.

The prevailing models of anxiety disorders emphasize interactions within limbic circuitry, especially involving prefrontal and sensory cortical regions, the amygdala, hippocampus, hypothalamus, BNST, and PAG. The bases for such models have progressed from theoretical considerations, extrapolated from fear-conditioning research in animals, to in vivo human imaging data. This section reviews current models of fear and stress-induced disorders, principally focusing on human neuroimaging studies. In this context, it is shown how data acquired using neuroimaging approaches inform the development and disruption of neural circuitry implicated in these disorders.

PHENOMENOLOGY OF FEAR AND STRESS-RELATED DISORDERS

Anxiety and fear are emotional states that, as in the case of a broader array of stress responses, normally serve adaptive functions. In contrast, anxiety disorders represent a category of psychiatric syndromes that are characterized by maladaptive fear and anxiety symptoms that cause distress and impair function. Panic disorder, specific phobia, SAD, GAD, and PTSD are considered to be fear and stress-related disorders. Panic disorder, specific phobia, SAD, and GAD are included as anxiety disorders in the fifth edition of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), whereas PTSD is a Trauma- and Stressor-Related Disorder. This shift acknowledges that a variety of clinical phenotypes are consistent with PTSD criteria and acknowledges the common precipitant of a traumatic stressor shared with other disorders in this category. Separation anxiety disorder and selective

mutism, which were previously categorized as Disorders Usually First Diagnosed in Infancy, Childhood, or Adolescence, are now included in the DSM-5 chapter on anxiety disorders in recognition of their role in adulthood as well. OCD, which was previously included as an anxiety disorder, was moved to Obsessive-Compulsive and Related Disorders in DSM-5 due to the shared features of obsessive preoccupation and repetitive behaviors that are present in OCD and disorders such as body dysmorphic disorder, trichotillomania, and Tourette's disorder.

Panic disorder is characterized by recurrent panic episodes, typically occurring spontaneously and without overt precipitants. Panic attacks involve a rapid escalation to extreme anxiety, accompanied by physical symptoms such as dyspnea, tachypnea, palpitations, sweating, dizziness, chest pain/tightness, and paresthesias, as well as emotional and cognitive symptoms, especially the feeling of dread that something catastrophic is about to happen or that immediate escape is necessary. Individuals with panic disorder often develop avoidance of contexts (i.e., places or situations) in which panic attacks previously occurred, are expected to be more likely, or from which escape may prove difficult (e.g., crowded or confined places). Panic attacks also commonly occur within the context of mood disorders or other anxiety disorders, and rarely occur in individuals with no psychiatric disorder. Consequently, the emphasis in modeling the pathogenesis of this disorder is directed toward the progression of panic attacks to panic disorder, as opposed to the underpinnings of panic attacks per se. Notably, temperament may relate to the vulnerability for developing panic disorder, given that heightened "anxiety sensitivity" (e.g., exaggerated awareness of interoceptive stimuli such as one's heartbeat) is a common characteristic of individuals who suffer from panic disorder.

Phobias are characterized by exaggerated anxiety responses that occur reliably during exposure to relatively innocuous stimuli or situations. Phobic responses include both the experience of fear and the peripheral manifestations of anxiety (e.g., changes in heart rate), and result in avoidance of the feared event. *Specific phobias* entail phobic responses to particular stimuli or situations (e.g., small animals, enclosed spaces). *Social anxiety disorder* (i.e., *social phobia*) entails phobic responses specific to social situations. In its most extreme form, avoidant behavior manifests as *Agoraphobia*, in which individuals experience anxiety about situations from which escape might be difficult (e.g., traveling in a bus or train, being in a crowd, being on a bridge). Though previously included as a qualifier for Panic Disorder (i.e., with or without agoraphobia), agoraphobia is now a separate diagnosis in DSM-5, recognizing that it can exist without panic attacks.

Generalized anxiety disorder is characterized by excessive worry about a variety of topics (e.g., job, health, finances, etc.), such as expecting the worst or worrying when nothing is wrong or in a manner that is

disproportionate to actual risk. Individuals with GAD find it difficult to control their worries and experience accompanying somatic or cognitive symptoms, including restlessness, fatigue, difficulty concentrating, irritability, difficulty sleeping, and muscle tension. Theories of GAD have focused on the function of worry, which is viewed as a form of cognitive avoidance. Consistent with this model, evidence suggests that worry may provide distraction from more emotional topics and reflect avoidance of future aversive events. In GAD, worry is used to cope with anticipated dangers or threats, and individuals often hold beliefs that worry is beneficial (e.g., "worrying keeps me safe").

Bridging from Phenomenology to Neurocircuitry

Neural circuitry underlying anxiety disorders has been characterized by: (1) identifying the functional domains that appear abnormal based on the phenomenology of the disorder; (2) establishing the mediating anatomy of the associated normal functions; (3) testing whether abnormalities exist at nodes within those circuits in individuals with the disorder; and, finally, (4) demonstrating associations between such abnormalities and the clinical features of the disorder. The following functional domains illustrate these concepts: (1) persistent or indelible learning about threat-related stimuli; (2) exaggerated fear responses; (3) impaired extinction learning or retention; (4) deficient ability to suppress attention and emotional/behavioral response to disorder-relevant stimuli; (5) overgeneralization of potential threat; (6) perceptual hypersensitivity to threat-related stimuli; and (7) exaggerated interoceptive/anxiety sensitivity. The neural circuits supporting these functional domains are discussed later.

In general, two lines of research have been most influential in understanding the neurocircuitry underlying these disorders. First, both animal studies using classic fear-conditioning paradigms and human neuroimaging studies in healthy humans have elucidated the neuroanatomical circuits that mediate fear, anxiety, and stress. In parallel, neuroimaging studies of individuals with anxiety disorders have begun to provide convergent evidence regarding how these neural substrates are altered in pathological states.

Preclinical Studies of Fear Learning

The design and interpretation of functional imaging studies of human fear and anxiety have relied in large part on animal studies. Given the high degree of cross-species conservation in the neural circuitry underlying fear learning, knowledge about anxiety and fear in humans has been informed by animal studies. Most of these preclinical studies investigate the anatomical correlates of fear processing within the context of Pavlovian

fear conditioning, fear-potentiated startle (FPS), and other putative anxiety-related paradigms in rodents (e.g., open field). The present discussion focuses on classic fear-conditioning paradigms as influential and pertinent examples. In these paradigms, an animal is exposed to a series of inherently innocuous sensory stimuli (e.g., auditory tones) in a conditioning procedure, such that one of these stimuli is paired with an inherently aversive stimulus (e.g., a shock, called the unconditioned stimulus [UCS]), whereas others are presented without shock and come to be associated with relative safety. These conditioned stimuli (CS) are termed the CS+ (predictive of the aversive UCS) and the CS− (predictive of no UCS). Following conditioning, exposure to the CS+ alone induces a constellation of physiological responses that can be conceptualized as defensive or preparatory for an anticipated aversive event and that bear resemblance to fear and anxiety, such as elevations in autonomic tone, increased startle, and freezing (in rodents).

Fear conditioning may be followed by extinction learning. During the process of extinction, a series of CS+ stimuli are presented without the delivery of an accompanying UCS, indicating that the CS+ no longer signals shock. Following a delay (e.g., the next day), the response to the CS+ is tested as an index of extinction retention. Extinction does not erase memory traces of the CS–UCS association but rather leads to the formation and strengthening of a competing safety memory which suppresses the original fear memory. Thus, previously extinguished fears can return because new learning is competing with old memory traces. The return of the fear response can also occur through reinstatement (unsignaled UCS presentations) or spontaneous recovery (following the passage of time or presentation of CS+ after extinction).

Unlike cued fear learning, paradigms that examine contextual fear learning are relevant to the phenomenology of anxiety disorders related to situation-specific anxiety. If the initial fear conditioning occurs in one context (e.g., a chamber with identifiable features; Cx+) and the extinction training occurs in a different context (e.g., a different chamber; Cx−), testing the following day reveals differential extinction retention versus renewal that occurs in the original context. Specifically, extinction learning is recalled in Cx− (the extinction/safe context), whereas the conditioned fear response is typically greater in Cx+ (the conditioned/threatening context). The context-gated return of the fear response following extinction is called “renewal.”

Recent studies have identified memory reconsolidation as a method for attenuation of fear memories. This concept is based on the notion that memories are dynamic; that is, every time a memory is retrieved, it returns to a fragile modifiable state and must restabilize again. Memory reconsolidation involves recalling an old memory such that it becomes labile followed by updating it with new information. The temporal window

during which a memory undergoes reconsolidation begins approximately 10 minutes after retrieval and lasts for at least 1 hour. Whereas a new memory must compete with the old fear memory in classic extinction, memory reconsolidation updates the old memory with nonfearful information (i.e., extinction is conducted during the reconsolidation window) such that fear responses are no longer expressed.

These preclinical studies of fear learning have important implications for the nature and treatment of anxiety disorders in humans. Individuals with anxiety may show exaggerated learned-fear responses, failure to extinguish learned-fear associations, failure to retain extinction learning, or failure to appreciate safe contexts. Exposure-based behavioral therapies for anxiety disorders build on the very principles of fear extinction. Moreover, the capacity for extinction learning, retention, and generalization across contexts may underlie the effectiveness of behavioral therapies and interventions.

ANATOMICAL CORRELATES OF FEAR AND ANXIETY

The application of fear learning and extinction paradigms in preclinical research has permitted the mapping of neural circuits that support core functions that are altered in fear and anxiety disorders (Fig. 14.6–1). Emotional processing in general can be divided into evaluative, expressive, experiential, and modulatory components. *Evaluation* of the emotional salience of a stimulus involves appraisal of its valence (e.g., appetitive vs. aversive) and its relationship to previous experiences and environmental contexts in which the stimulus arises. Emotional *expression* conveys the range of behavioral, endocrine, and autonomic manifestations of the emotional response, whereas emotional experience describes the subjective feeling accompanying the response. In contrast, the neural correlates of the subjective *experience* of fear and anxiety have been most directly addressed in neuroimaging studies of humans, who have the capability of verbally describing their experiences.

Finally, to optimize their capacity for guiding behavior, each of these aspects of emotional processing is *modulated* by complex neurobiological systems that prevent fear or anxiety responses from becoming persistent, excessive, or maladaptive. For example, fear conditioning and extinction paradigms have been used to investigate the “top-down” modulation of amygdala function by cortical structures, as well as the “bottom-up” influence of various hormones and neurotransmitters.

Fear Circuitry

Within the context of fear learning, the amygdala plays critical roles both in forming conditioned fear associations and in organizing fear responses. The amygdala responds to both UCS, which are inherently emotionally

salient and require no prior learning, and to CS that predict these emotionally salient events. The anatomical systems underlying fear are organized to permit both rapid responses to simple perceptual elements of potentially threatening stimuli and longer latency responses to highly processed information about more complex sensory stimuli and environmental contexts.

The amygdala is comprised of many subnuclei, called the amygdaloid complex, each of which serves a different function. The lateral nucleus of the amygdala (LA) serves as the primary sensory interface of the amygdala, receiving sensory information from the thalamus and cortex. Single neurons within the LA respond to auditory, visual, and somatic stimuli, and information within these domains about both the CS and the UCS converge in the LA. In contrast, olfactory input enters the cortical and medial nuclei of the amygdala via direct projections from the olfactory bulb. The projections from sensory thalamus to the LA are believed to support rapid conditioning to simple auditory or visual stimuli (i.e., bypassing cortical processing), enabling fear responses below the level of conscious awareness. Thus, lesioning the auditory cortex prior to conditioning does not prevent conditioning to single auditory tones. In contrast, projections to the LA from the primary sensory and sensory association cortices appear to be essential for some aspects of conditioned responding to more complex sensory stimuli. These relationships are modality-specific, such that disruption of the projections from the auditory thalamus and auditory cortex to LA specifically prevents the acquisition of fear conditioning to auditory stimuli and fear-conditioned responses to previously conditioned auditory stimuli.

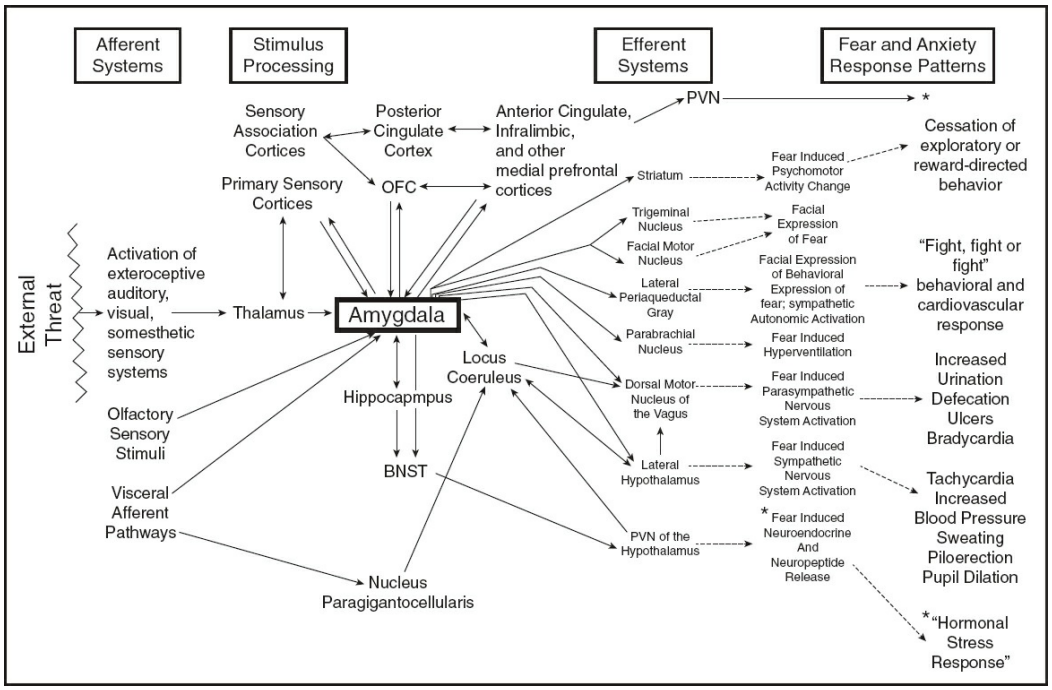

FIGURE 14.6–1. Neuroanatomical circuitry underlying fear and anxiety, organized by systems that evaluate the salience of potentially threatening stimuli (to the left) and systems that organize and mediate the expression of fear and anxiety behaviors (to the right). BNST, bed nucleus of the stria terminalis; OFC, orbitofrontal cortex; PVN, paraventricular nucleus. (From Charney DS, Drevets WC. The neurobiological basis of anxiety disorders. In: Davis K, Charney DS, Coyle J, Nemeroff CB, eds. *Psychopharmacology. The Fifth Generation of Progress*. New York: Lippincott Williams & Wilkins; 2002:901.)

After sensory input enters the LA, the neural representation of the stimulus is distributed in parallel to various amygdaloid nuclei, where it is modulated by diverse functional systems, such as those conveying mnemonic information about past experiences or interoceptive input related to ongoing homeostatic states. The most extensive amygdalar projections of the LA consist of reciprocal projections to the basal (BA) and accessory basal nuclei and the central nucleus of the amygdala (CE). Lesions of either the LA or CE—but not of other amygdala nuclei—disrupt fear conditioning to a tone CS, indicating that the direct projection from LA to CE is sufficient to mediate conditioning to simple sensory stimuli.

The projections from LA to the BA nuclei also participate in forming long-lasting memory traces for fear conditioning. Functional inactivation of the LA and BA nuclei prior to Pavlovian fear conditioning interferes with the acquisition of learning, whereas inactivation immediately after conditioning has no effect on memory consolidation. The BA nuclei have widespread intranuclear connections and also project to other amygdalar nuclei, including the CE and LA. The BA nuclei also share extensive, reciprocal projections with the orbital and medial PFC, through which they appear to modulate neuronal responses of both the PFC and the LA and CE.

Plasticity-related fear learning occurs in the amygdala and cortical areas. Plasticity within the amygdala involves long-term potentiation. Cortical regions are thought to establish explicit or declarative memories about the fear-related event through interactions with the medial temporal lobe memory system. The influence of the amygdala on cortically based memories has been most clearly characterized with respect to late plastic components of the auditory cortex neuronal responses to a CS. Single-unit recordings during fear conditioning indicate that some auditory cortex neurons that prior to conditioning did not respond to the CS tone develop late-conditioned responses (i.e., 500 to 1,500 ms after CS onset) that anticipate the UCS and show extinction-resistant memory storage. These late-conditioned auditory cortical neuronal responses take more trials to learn and respond more slowly than LA neurons within trials, and their late development is prevented by amygdala lesions. Thus, whereas the rapid conditioning of fear responses to potentially dangerous stimuli depends on plasticity in the amygdala, learning involving higher cognitive (i.e., mnemonic and attentional) processing of fear experiences may depend on

plasticity involving cortical neurons that is influenced by neural transmission from the amygdala.

Other auditory cortex neurons show an early (less than 50 ms after stimulus onset) plastic component during fear conditioning, in which the pre-existing electrophysiological responses of auditory cortex neurons to the CS become enhanced by conditioning. This short-latency plasticity within the auditory cortex appears to depend on input from the auditory thalamus and is unaffected by amygdala lesions. Nevertheless, such short-latency responses are extinguished more quickly (i.e., during repeated exposure to the CS alone) in animals with amygdala lesions, implying that the amygdala is involved in preventing the extinction of these responses.

Neural Circuitry Underlying Fear Expression

The amygdaloid output nuclei, especially the CE, receive convergent input from multiple amygdala regions and organize behavioral responses that reflect the sum of neuronal activity produced by multiple amygdala nuclei. The CE comprises the interface between the amygdala and the motor, autonomic, and neuroendocrine systems involved in expressing fear behavior. The CE projects to nuclei in the hypothalamus, midbrain, and medulla that mediate the neuroendocrine, autonomic, and behavioral responses associated with fear. For example, the amygdala facilitates stress-related corticotropin-releasing hormone (CRH) release via both intrinsic CRH-containing neurons within the amygdala and bisynaptic (double GABAergic [GABA, γ-aminobutyric acid]) anatomical projections to the paraventricular nucleus (PVN) of the hypothalamus (the nucleus containing the majority of CRH-synthesizing neurons in the hypothalamus). Electrical stimulation of the CE produces responses similar to those elicited by fear-conditioned stimuli, and lesions of the CE prevent the expression of fear responses of various types. In contrast, lesioning of specific structures efferent to the CE, such as the lateral hypothalamus or PAG, produces selective deficits in cardiovascular or somatomotor behavioral fear responses, respectively. The amygdala also sends projections to the thalamus, nucleus accumbens, ventromedial caudate, and ventral putamen, which participate in organizing motor responses to threatening stimuli. The amygdala may also influence motor behavior via projections through the hypothalamus and PAG.

The hypothalamic and brainstem structures that mediate the expression of emotional behavior can be activated directly via either the CE or the BNST. Anxiety-like responses elicited either by exposure to a threatening environment for several minutes or by intraventricular administration of CRH are mediated via the BNST rather than the CE, and the BNST has been implicated in sustained states of anxiety. This system is thus hypothesized to play a role in mediating anxiety during exposure to less explicit or less well-defined sensory cues and/or to contexts that occur over

a longer duration.

Fear Regulation Circuitry

Multiple areas of the medial and orbital PFC participate in modulating anxiety and other emotional behaviors. These structures are putatively involved in interpreting the higher-order significance of experiential stimuli, modifying behavioral responses based on reinforcement contingencies, and predicting social outcomes of behavioral responses to emotional events. These areas share extensive reciprocal projections with the amygdala through which the amygdala modulates PFC neuronal activity and the PFC can modulate amygdala-mediated responses to emotionally salient stimuli. The mPFC regulates the amygdala through input to the BA and the intercalated cells, a group of GABAergic neurons that regulate input from the BA to the CE and thus inhibit amygdala output.

The mPFC areas implicated in anxiety and fear-related behavior in humans and experimental animals include the infralimbic cortex, prelimbic cortex, anterior cingulate cortex (ACC), and a more anterior mPFC region extending from the rostral ACC (BA 24, 32) toward the frontal pole. The reciprocal projections between the amygdala and mPFC are hypothesized to play critical roles in attenuating fear responses and extinguishing behavioral responses to fear-conditioned stimuli that are no longer reinforced. Lesions of the ACC in rats enhance freezing to a fear-conditioned tone, suggesting that this mPFC region is involved in fear reduction. In addition, neurons in the rat prelimbic cortex reduce their spontaneous firing activity in the presence of a conditioned, aversive tone to an extent that is inversely proportional to the magnitude of fear. This suppression of prelimbic cortex neuronal firing activity is inversely correlated with increases in amygdala neuronal activity. Prelimbic cortex (which may correspond to the dorsal ACC in humans) projections to the amygdala also have been shown to enhance the magnitude of fear responses to aversively CS (Fig. 14.6–2). Finally, lesions of the infralimbic cortex (which is thought to correspond to the ventromedial PFC in humans) specifically interfere with the recall of extinction processes after long delays between the acquisition of extinction learning and re-exposure to the initial CS.

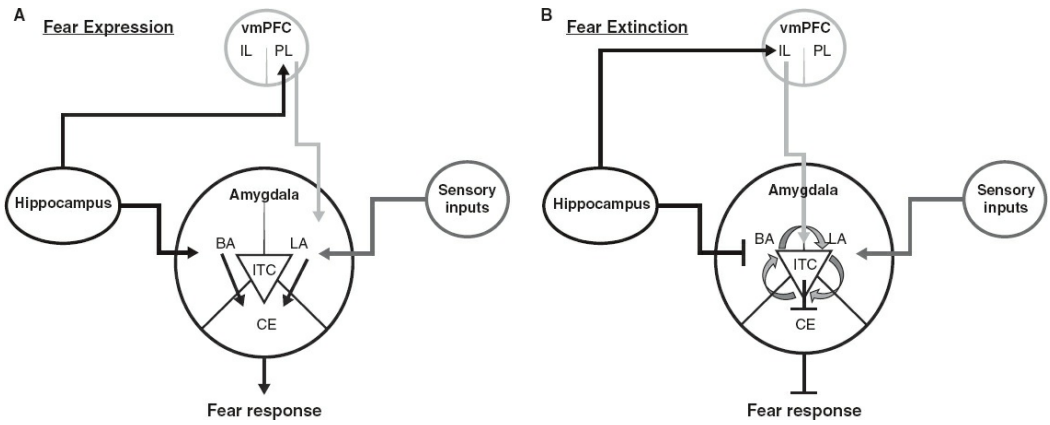

FIGURE 14.6–2. Neural circuitry of fear expression and extinction. **A**: During acquisition and expression of conditioned fear, projections from PL and thalamic nuclei (mediating converging sensory information) excited LA neurons, while hippocampal projections (mediating contextual inputs) lead to extinction of BA neurons directly or indirectly via connections with PL. LA and BA neurons activate CE output neurons, which project to downstream brainstem and hypothalamic nuclei responsible for mediating physiologic responses, resulting in fear expression. **B**: During extinction of conditioned fear, hippocampal projections (mediating contextual inputs) lead to divergent excitation of IL neurons and inhibition of BA neurons. IL projections directly activate GABAergic ITC cells within the amygdala. Integration of ITC, BA, and LA inputs during extinction results in a suppression of CE output neurons, resulting in a lack of physiologic response and suppression of fear expression. Arrowheads delineate pathway excitation; straight ends delineate pathway inhibition. *For simplicity, connection arrows are delineated as being unidirectional, although bidirectional projections exist. (From Pattwell SS, Lee FS, Casey BJ. Fear learning and memory across adolescent development: Hormones and Behavior Special Issue: Puberty and Adolescence. *Horm Behav.* 2013;64(2):380–389.)

In humans the pregenual ACC (pgACC; that is, the ACC situated anterior to the genu of the corpus callosum) shows areas of elevated hemodynamic activity during a variety of anxiety states elicited in healthy participants or participants with anxiety disorders. Electrical stimulation of this region elicits fear, panic, or a sense of foreboding in humans and vocalization in experimental animals. The subgenual ACC (sgACC) has been implicated in healthy sadness, MDD, bipolar disorder, and PTSD. In participants with PTSD, cerebral blood flow (CBF) increases in this region during internally generated imagery of past trauma.

Both the sgACC and pgACC share reciprocal anatomical connections with areas implicated in emotional behavior, such as the posterior orbitofrontal cortex (OFC), amygdala, hypothalamus, nucleus accumbens, PAG, VTA, raphe, locus coeruleus (LC), and nucleus tractus solitarius. Humans with mPFC lesions that include the pgACC and sgACC show abnormal autonomic responses to emotionally provocative stimuli, an inability to experience emotion related to concepts, and an inability to use information regarding the probability of aversive social consequences versus reward in guiding social behavior. In rats, lesions of the vmPFC comprised of infralimbic, prelimbic, and ACC cortices alter corticosterone secretion, sympathetic autonomic responses, and gastric stress pathology

during restraint stress or exposure to fear-conditioned stimuli. Finally, the ACC contains GC receptors that, when stimulated, inhibit stress-induced corticosterone release in rats.

Physiological activity also increases in more-anterior mPFC areas in healthy humans as they perform tasks that elicit emotional responses or require emotional evaluations. During anxious anticipation of an electrical shock, CBF increases in the rostral mPFC (vicinity of anterior BA 24, BA 32, and rostral BA 9), and the magnitude of the blood flow change correlates inversely with changes in anxiety ratings and heart rate. In rats, lesions of the rostral mPFC result in exaggerated heart rate responses to fear-conditioned stimuli, and stimulation of these sites attenuates defensive behavior and cardiovascular responses evoked by amygdala stimulation. In primates, whereas BA 24 and 32 have extensive reciprocal connections with the amygdala, the BA 9 cortex has only sparse projections to the amygdala. Nevertheless, all three areas send extensive efferent projections to the PAG and the hypothalamus through which cardiovascular responses associated with emotional behavior may be modulated.

Role of the Hippocampus and Other Structures in Contextual Fear

The hippocampus, perirhinal cortex, ventrolateral PFC (vlPFC), and anterior insula interact with the amygdala in the development of emotional responses to the environment. While much of the literature on anxiety has focused on frontoamygdala connections, the hippocampus plays a critical role in fear learning and extinction. Projections from the hippocampal formation to the amygdala via the fornix are implicated in spatial contextual conditioning. For example, lesioning these projections specifically prevents fear conditioning to the chamber or the position within a maze where aversive stimulation previously occurred. The hippocampus modulates frontoamygdala function by supplying contextual information about the degree of threat or safety in the environment that influences whether the extinction or original fear memory is expressed, which is accomplished via projections from the ventral hippocampus to the BA. Whether experienced as safe or threatening, these hippocampal–BA connections can modulate subsequent fear responses via projections to the CE. Connections between the ventral hippocampus and vmPFC (IL) modulate extinction learning by detecting contextual cues in the surrounding environment. Inactivation of the hippocampus before extinction training leads to impaired retrieval of cued extinction memory. In these ways, hippocampally mediated control of frontoamygdala connections increases an organism's flexibility to respond appropriately to danger cues across different environments.

The perirhinal cortex shares reciprocal anatomical connections with the amygdala and is believed to play a role in conveying information about

complex visual stimuli to the amygdala during the presentation of fear-conditioned visual stimuli. Lesions of the anterior perirhinal cortex, the basolateral nucleus of the amygdala, or the CE each can eliminate completely the FPS during exposure to some conditioned visual stimuli. In contrast, removal of the entire visual cortex, insular cortex, medial PFC, and posterior perirhinal cortex produces no significant effect on the magnitude of FPS, and lesions of the frontal lobe only partly attenuate FPS. The perirhinal cortex receives input regarding conditioned visual stimuli from the lateral geniculate nucleus, and lesions of this structure also can block FPS. Finally, the anterior perirhinal cortex receives afferent projections from the visual cortices and the ACC, infralimbic cortex, and parietal cortex, structures implicated in modulating behavioral responses to fear-conditioned stimuli. Lesions placed in the anterior perirhinal cortex after fear conditioning interfere with the expression of conditioned fear responses elicited by visual and auditory stimuli when these stimuli are presented in contexts that differ from the initial conditioning context. In contrast, lesions of the vlPFC and anterior insula reduce fear reactivity to contextual stimuli but do not affect CS acquisition or response extinction.

The temporopolar cortex has been implicated in modulating autonomic aspects of emotional responses and in processing emotionally provocative visual stimuli. Electrical stimulation of various sites within the temporopolar cortex can alter a variety of autonomic nervous system functions. In humans neurophysiological activity increases in the anterior temporopolar cortex during exposure to emotionally provocative visual stimuli, whether the stimuli convey sad, disgusting, or happy content, relative both to conditions involving exposure to emotionally neutral visual stimuli and to conditions in which corresponding emotional states are elicited via recall of autobiographical information. Portions of the temporopolar cortex may thus function as sensory association areas that participate in evaluating the emotional salience of actual or anticipated stimuli and in modulating autonomic responses to such stimuli.

Emotion Processing and Other Related Circuitry

Other areas implicated in studies of fear or anxiety in human and nonhuman primates include the posterior and lateral OFC, anterior insula, and vlPFC. Physiological activity increases in these areas during experimentally induced anxiety states in healthy participants and in participants with specific phobia or panic disorder. A complex relationship exists between physiological activity and anxiety symptoms in the OFC and vlPFC. Posterior OFC flow increases in participants with animal phobias during exposure to phobic stimuli and in healthy participants during induced sadness, but this change in CBF correlates inversely with changes in anxiety.

These data appear to be consistent with electrophysiological and lesion

analysis data showing that the OFC participates in modulating behavioral and visceral responses associated with fearful, defensive, and reward-directed behavior as reinforcement contingencies change. Nearly one-half of the OFC pyramidal neurons alter their firing rates during the delay period between stimulus and response, and this firing activity relates to the presence or absence of reinforcement. These cells are believed to play roles in extinguishing unreinforced responses to aversive or appetitive stimuli. The posterior and lateral OFC and the amygdala send projections to each other and to overlapping portions of the striatum, hypothalamus, and PAG through which these structures modulate each other's neural transmission. For example, the defensive behaviors and cardiovascular responses evoked by electrical stimulation of the amygdala are attenuated or ablated by concomitant stimulation of OFC sites, which, when stimulated alone, exert no autonomic effects.

Humans with OFC lesions show impaired performance on tasks requiring the application of information related to punishment or reward, perseverate in behavioral strategies that are unreinforced, and exhibit difficulty in shifting intellectual strategies in response to changing task demands. Similarly, monkeys with surgical lesions of the lateral OFC/vlPFC demonstrate "perseverative interference," which is characterized by difficulty in learning to withhold prepotent responses to nonreinforced stimuli as reinforcement contingencies change. The activation of the OFC during anxiety or obsessional states thus may reflect endogenous attempts to attenuate emotional expression or to interrupt unreinforced aversive thought and emotion. Conversely, OFC dysfunction may contribute to pathological anxiety and obsessional states by impairing the ability to inhibit nonreinforced or maladaptive emotional, cognitive, and behavioral responses to social interactions and sensory or visceral stimuli.

Many functional imaging studies report that exposure to aversive stimuli of various types increases physiological activity in the retrosplenial cortex and other portions of the posterior cingulate cortex. The posterior cingulate cortex is believed to serve as a sensory association cortex that processes the affective salience of experiential stimuli. The posterior cingulate cortex sends a major anatomical projection to the ACC through which it relays information into the limbic circuitry.

Changes in regional CBF associated with anxiety in other structures have been less consistently replicated. Several studies report anxiety-associated hemodynamic changes in the lateral temporal and inferior parietal cortical areas or the ventral striatum. These cortical areas comprise sensory association cortices, and physiological activity in these regions during anxiety may reflect sensory processing of the anxiety-provoking stimuli or feedback to these regions by the amygdala or other subcortical regions.

Finally, medial cerebellar activity consistently increases during experimentally induced anxiety or sadness in healthy participants or participants with anxiety disorders. One meta-analysis of PET studies investigating the anatomical correlates of multiple types of anxiety found two regions in which CBF increased across all normal and pathological anxiety states studied, namely the medial cerebellum and the midbrain (in the vicinity of the PAG). The medial cerebellum contains a set of nuclei that serve as functional complements to the hypothalamus in the modulation of autonomic responses. In addition, most or all areas of the cerebral cortex form closed, multisynaptic circuits through the cerebellum, although the functional significance of these anatomical loops is unclear.

Role of the Amygdala in Biologically Relevant Learning and Emotional Memory

In addition to its role in fear conditioning, evidence suggests the amygdala functions more broadly to facilitate biologically relevant learning, including responding to aversive and appetitive CSs, uncertainty, associability, violations of expectations, innate fear, processing affective elements of social interactions, and in nonspecific attention and arousal that facilitates learning. With regard to social dominance hierarchies, monkeys with amygdala lesions generally show reduced aggression and fear and become more submissive to dominant animals. Nevertheless, amygdala-lesioned monkeys appear to be oblivious to socially significant signals conveyed by the behavior or facial expressions of other monkeys (e.g., attempting to engage in sexual behavior with nonreceptive animals or without regard to dominance hierarchy), and presumably as a consequence, they are generally killed by other monkeys in wild (i.e., natural setting) primate colonies. The involvement of the amygdala in evaluating social behavior with respect to the emotional information conveyed by human facial expression or voice intonation has been demonstrated in humans. Humans with amygdala lesions are impaired in their ability to recognize fear in facial expression and fear and anger expressed in spoken language. In PET and fMRI studies, activity increases in the amygdala as humans view faces expressing various types of emotional expression. The automatic nature of this response is evident by its occurrence even when the face stimuli are shown so briefly (using backward masking techniques) that participants are not explicitly aware of having seen them.

Human neuroimaging, electrophysiological, and lesion analysis studies also demonstrate that the amygdala is involved in the recall of emotional or arousing memories. In humans, bursts of electroencephalographic (EEG) activity have been recorded in the amygdala during the recollection of specific emotional events. Moreover, electrical stimulation of the amygdala in humans can evoke the experience of fear or anxiety and the recollection of emotionally arousing life events from remote memory. Activation of the amygdala enhances the strength of long-term memory for emotional

stimuli or contexts represented in other cortical memory circuits as well. These circuits putatively involve the medial temporal lobe memory system, which has extensive anatomical connections with the amygdala through which interactions between storage and explicit recall of emotionally salient memories are supported. For at least some types of emotional learning, memory consolidation is critically dependent on noradrenergic stimulation of β- and α-1-adrenoreceptors in the basolateral nucleus of the amygdala. Acquisition of fear-conditioned responses thus requires an intact central noradrenergic system. The activation of norepinephrine release during emotional learning is facilitated by glucocorticoid (GC) hormone effects on noradrenergic neurons.

Structures Mediating Endocrine and Autonomic Responses of Fear or Stress

The peripheral hormonal and autonomic responses to threat mediated by the hypothalamic–pituitary–adrenal (HPA) axis and the sympathetic and parasympathetic autonomic nervous systems facilitate the generation of adaptive responses to threat or stress. Stimulation of the lateral nucleus of the hypothalamus via afferent projections from the CE, BNST, or ventral striatum activates the sympathetic nervous system, producing increases in blood pressure and heart rate, sweating, piloerection, and pupillary dilation. During stress, the secretion of CRH from the PVN of the hypothalamus increases peripheral adrenocorticotropic hormone (ACTH) levels, which stimulates the adrenal glands to secrete cortisol. The ACC, anterior insula, and posterior OFC send anatomical projections to the hypothalamus, which modulate cardiovascular and endocrine responses to threat and stress.

Regional differences in the regulation of CRH function by GC receptor stimulation and stress appear to play major roles in the mediation of fear and anxiety. Feedback inhibition of CRH release by GC-receptor stimulation (i.e., to suppress HPA axis activity) occurs at the level of the PVN and anterior pituitary, where systemically administered GC hormones reduce CRH release and CRH-receptor expression, respectively. In contrast, GC hormone administration can increase CRH mRNA expression within the CE and the BNST. The positive feedback of GH hormones on extrahypothalamic CRH function in the amygdala and the BNST may serve to amplify stress responses.

The vagus and splanchnic nerves constitute the major efferent projections of the parasympathetic nervous system to the viscera. The vagal nuclei receive afferent projections from the lateral hypothalamus, PVN, LC, amygdala, and infralimbic cortex. The splanchnic nerves receive afferent connections from the LC. The innervation of the parasympathetic nervous system from these limbic structures is believed to mediate visceral symptoms associated with anxiety, including alterations in cardiovascular, gastrointestinal, and genitourinary function.

Areas within the orbital and medial PFC also participate in modulating peripheral responses to stress, including heart rate, blood pressure, and GC hormone secretion. The neuronal activity within these PFC areas in turn is modulated by a variety of neurotransmitter systems that activate in response to stressors and threats. For example, the noradrenergic, dopaminergic, and serotonergic systems play roles in enhancing vigilance, modulating goal-directed behavior, and facilitating decision making about probabilities of aversive versus rewarding outcomes by modulating neuronal activity in the PFC.

NEUROCHEMICAL SYSTEMS INVOLVED IN FEAR AND ANXIETY

The chemical neurotransmitter systems that modulate the anatomical circuits supporting fear and anxiety processing include peptidergic neurotransmitters (e.g., CRH, neuropeptide Y, substance P, opioids), monoaminergic transmitters (norepinephrine, serotonin, dopamine), amino acid transmitters (GABA and glutamate), and adrenal steroids (e.g., cortisol). These neurochemical systems subserve important adaptive functions in preparing the organism for responding to threat or stress by increasing vigilance, modulating memory, mobilizing energy stores, and elevating cardiovascular function, but they can become maladaptive if they are chronically or inappropriately activated.

Central Noradrenergic System

Among the neurotransmitter systems that have been best studied in experimental animals in association with responses to stress or threat is the central noradrenergic system. Exposure to fear-conditioned stimuli, immobilization stress, and foot shock or tail pinch each increases norepinephrine turnover in the LC, hypothalamus, hippocampus, amygdala, and cerebral cortex. The firing activity of LC neurons also increases during exposure to fear-conditioned stimuli and other stressors or threats. Repeated exposure to severe, inescapable stressors, in contrast, results in reduced norepinephrine release, which is associated with the behavioral pattern termed “learned helplessness.” The responsiveness of LC neurons to novel stressors can be enhanced by previous chronic exposure to some stressors. In rats, the amount of norepinephrine synthesized and released in the hippocampus and the mPFC in response to a novel stressor or to local depolarization is exaggerated following repeated exposure to chronic cold stress. This effect of chronic stress on noradrenergic responses to subsequent novel stressors may constitute a form of “behavioral sensitization” in which single or repeated exposures to aversive stimuli or pharmacological agents can increase the behavioral sensitivity to subsequent stressors. Patients with PTSD and panic disorder show evidence of heightened peripheral sympathetic nervous system

arousal, which, because of the correlation between peripheral sympathetic activity and central noradrenergic function, is compatible with the hypothesis of increased central norepinephrine activity in these disorders.

Norepinephrine, cortisol, and CRH thus are tightly linked as a functional system that affords a homeostatic mechanism for responding to stress adaptively. GC-mediated inhibition of norepinephrine-induced CRH stimulation may be evident primarily during stress rather than under resting conditions as an adaptive response that restrains stress-induced neuroendocrine effects mediated by the PVN. In addition, alterations of both brain catecholamine and GC hormone levels affect the consolidation and retrieval of emotional memories in experimental animals. GCs influence memory storage via activation of GC receptors in the hippocampus, whereas norepinephrine effects are mediated in part through β-adrenoreceptor stimulation in the amygdala. In humans, adrenocortical suppression blocks the memory-enhancing effects of amphetamine and epinephrine, and propranolol impairs memory for an emotionally provocative story but not for an emotionally "neutral" story. The acute release of GC hormones and norepinephrine in response to trauma would thus modulate the encoding of traumatic memories.

Central Benzodiazepine–f-Aminobutyric Acid Receptor System

Several lines of preclinical and clinical evidence established that benzodiazepine-receptor agonists exert anxiolytic effects and suggested that benzodiazepine receptor function is altered in anxiety disorders. Benzodiazepine receptors are expressed throughout the brain but are most densely concentrated in the cerebral cortex. The benzodiazepine and $GABA_A$ receptors form parts of the same macromolecular complex, and although they constitute distinct binding sites, they are functionally coupled and regulate each other in an allosteric manner. Central benzodiazepine receptor agonists potentiate and prolong the synaptic actions of the inhibitory neurotransmitter GABA by increasing the frequency of GABA-mediated chloride-channel openings. Microinjection of benzodiazepine receptor agonists in limbic and brainstem regions such as the amygdala and PAG exert antianxiety effects in animal models of anxiety and fear. Conversely, the administration of benzodiazepine receptor inverse agonists produces behavioral responses and increases in heart rate, blood pressure, cortisol secretion, and catecholamine release resembling those seen in anxiety and stress, and these effects are blocked by the administration of benzodiazepine receptor agonists.

In experimental animals, stress downregulates benzodiazepine-receptor binding in the frontal cortex and hippocampus. In rats exposed to inescapable stress, the benzodiazepine-receptor binding decreases markedly in the frontal cortex, shows less consistent reductions in the

hippocampus and hypothalamus, and remains unchanged in the occipital cortex, striatum, midbrain, thalamus, and cerebellum. When such stressors are administered repeatedly or chronically, the benzodiazepine-receptor binding decreases more consistently in the hippocampus (i.e., in addition to frontal cortex) and also may decrease in the cerebellum, midbrain, and striatum.

Central Serotonergic System

In experimental animals, acute exposure to a variety of severe stressors, including restraint stress, tail shock, tail pinch, and high-level foot shock, results in increased serotonin (5-HT) turnover in the mPFC, nucleus accumbens, amygdala, and lateral hypothalamus. During exposure to fear-conditioned stimuli, the 5-HT turnover in the mPFC appears to be particularly sensitive to the severity of stress, increasing as the aversiveness of the UCS and the magnitude of the fear-behavioral response increase. In contrast, repeated exposure to electric shocks sufficient to produce learned helplessness is associated with reduced in vivo release of 5-HT in the frontal cortex. Preadministration of benzodiazepine receptor agonists or antidepressant drugs prevents this stress-induced reduction in 5-HT release and interferes with the acquisition of learned helplessness, whereas infusion of 5-HT into the frontal cortex after stress exposure reverses learned-helplessness behavior. The effect of stress in activating 5-HT turnover may stimulate both anxiogenic and anxiolytic pathways within the forebrain, depending on the region involved and the 5-HT receptor subtype that is predominantly stimulated.

Notably, 5-HT_{1A}–receptor knock-out mice show a marked elevation of anxiety and fear behaviors, which can be reversed by administration of 5-HT_{1A}–receptor agonists when these agents are administered prior to a critical developmental period. These findings are notable given that repeated or chronic stress decreases the genetic expression of 5-HT_{1A} receptors. Conversely, 5-HT_{2A}–receptor expression is upregulated during chronic stress and corticosterone administration and downregulated in response to adrenalectomy. In view of evidence that chronic administration of 5-HT_{1A}–receptor partial agonists and 5-HT_{2A}–receptor antagonists exerts anxiolytic effects, it is conceivable that these stress/corticosteroid-mediated effects on 5-HT_{1A}– and 5-HT_{2A}–receptor expression are relevant to the pathophysiology of anxiety.

Dopaminergic System

Acute stress increases dopamine release and turnover in multiple brain areas. The dopaminergic projections to the mPFC appear to be particularly sensitive to stress because brief or low-intensity stressors (e.g., exposure to

fear-conditioned stimuli) increase dopamine release and turnover in the mPFC in the absence of corresponding changes in other regions. For example, in rats, low-intensity electrical foot shock increases tyrosine hydroxylase activity and dopamine turnover in the mPFC but not in the nucleus accumbens or caudate putamen. In contrast, stress of greater intensity or longer duration additionally enhances dopamine release and metabolism in other areas as well. The regional sensitivity to stress thus follows a pattern in which dopaminergic projections to the mPFC are more sensitive to stress than the mesoaccumbens and nigrostriatal projections, and the mesoaccumbens dopaminergic projections are more sensitive to stress than the nigrostriatal projections. In healthy humans, the amount of dopamine released in the ventral striatum correlated negatively with anxiety and positively with euphoria.

NEUROIMAGING STUDIES OF HUMAN FEAR AND ANXIETY

The data obtained in experimental animals pertaining to the neural circuitry underlying the evaluation of emotionally salient stimuli, fear and extinction learning, emotional expression, and emotion modulation have been extended to humans using neuroimaging approaches. These lines of research were conducted initially in healthy humans to investigate the normal mediating anatomy of anxiety expression and modulation and to develop probes of functional response in these regions. In tandem, studies have been performed in individuals with anxiety disorders.

Overview of Neuroimaging Technology

Structural Neuroimaging Techniques. Volumetric magnetic resonance imaging (vMRI) involves automated or semiautomated segmentation of brain structures of interest, enabling investigators to measure indices of size or features of shape from high-resolution structural MRI data. Methods applied in this field include techniques for parcellating cortical subterritories, automated procedures for measuring cortical thickness, and voxel-based morphometry (VBM), which is an automated method for identifying voxels where structural differences can be inferred based on signal intensity differences on structural MRI scans. Diffusion tensor imaging (DTI) provides an index of regularity and myelination of white matter tracts by measuring the degree of anisotropic diffusion (fractional anisotropy) (FA). These methods have been used to examine regional size, shape, and tissue composition and how they relate to clinical symptoms in patients with anxiety disorders.

Functional Neuroimaging Techniques and Paradigms. A variety of experimental paradigms have been developed in conjunction with functional neuroimaging modalities to delineate the neurophysiological

correlates of normal brain functions, as well as the pathophysiological correlates of psychiatric disorders. Commonly applied methods include PET with tracers that measure regional cerebral blood flow (rCBF) or glucose metabolism, and fMRI to measure blood oxygenation level–dependent (BOLD) signal changes. Such brain activity profiles are sensitive to the cognitive behavioral state of the individual at the time of tracer distribution or image acquisition. Hence, functional imaging paradigms can be categorized based on the type of state manipulations employed, and the principal statistical comparisons applied.

In *neutral state* paradigms, participants are studied during a nominal "resting" state or while performing a nonspecific continuous task. Thus, hypotheses regarding between-group differences in brain activity are tested without particular consideration of momentary state variables.

In *symptom provocation* paradigms, participants are scanned during both a symptomatic state (after having their symptoms intentionally induced or exacerbated) and during a control condition. Within-group comparisons can be performed to test hypotheses regarding the mediating anatomy of the symptomatic state, and group-by-condition interactions can be assessed to distinguish responses in patient versus control samples. Either behavioral or pharmacological challenges can be used to induce symptoms. In some cases, when symptomatic states occur spontaneously, experiments can be designed to capture these events without the need for additional provocation or induction.

In *cognitive/behavioral activation* paradigms, participants are studied while performing specially designed cognitive/behavioral tasks. This approach aims to increase sensitivity and specificity by using tasks that activate particular brain systems of interest. In healthy participants, within-group analyses are conducted to establish the normal mediating anatomy for functional domains of interest. Subsequently, group-by-condition interactions are performed to compare the functional responsivity or integrity of the target brain systems in patient versus control samples.

Any of these paradigm types can be extended to longitudinal studies. In *developmental* studies, longitudinal designs can be used to characterize changes in functional or structural brain measures across childhood and adolescence in individuals with anxiety disorders or at risk for them. In *treatment* studies, participants are scanned both before and after a treatment trial. Within-group comparisons are then performed to test hypotheses regarding changes in brain activity profiles associated with symptom improvement, which may inform the mechanisms underlying effective treatments. Analyses also can be performed to identify pretreatment brain activity characteristics that predict subsequent treatment response.

Neuroreceptor/Neurochemical Imaging Tech niques. Neuroreceptor imaging studies use PET or single photon emission computed tomographic (SPECT) methods in conjunction with radiolabeled ligands that bind selectively to specific receptors, enzymes, or neural elements. In this way, the regional binding potential (modeled as density × affinity) has been characterized in vivo for some receptor species in anxiety disorders. The neuroreceptor imaging literature pertaining to stress and fear disorders is limited, however, by the paucity of radioligands available for in vivo investigation of the aforementioned neurochemical systems in humans. For example, radioligands for noradrenergic, GC, or CRF receptors are unavailable for human imaging, and thus limited information exists on these receptor systems in human anxiety disorders.

Other neurochemical imaging approaches include the use of *magnetic resonance spectroscopy* (MRS) to measure the regional relative concentration of select "MRS-visible" compounds. For example, MRS can be used to measure concentrations of the compound *N*-acetylaspartate (NAA), a purported marker of healthy neuronal density, or of neurotransmitters such as glutamate or GABA.

Human Neuroimaging Studies of Fear Conditioning and Extinction

Neuroimaging studies have been conducted to investigate the functional anatomy associated with conditioned fear acquisition, extinction, and extinction retention. Consistent with the findings from research conducted in experimental animals, PET and fMRI studies of fear conditioning in healthy humans have implicated the amygdala, vmPFC, and hippocampus. These studies have often used peripheral physiological endpoints (e.g., skin conductance) to examine conditioned fear and its association with functional activation and connectivity.

Many studies have shown amygdala activation during fear conditioning to the CS+ compared to the CS−, with correlations between amygdala activation and the conditioned response during acquisition. Activation of the vmPFC (specifically in the pgACC) and its functional connectivity with the amygdala have been observed, especially during extinction. During the recall of extinction after a 24-hour delay, activation of the vmPFC predicted extinction success and correlated with amygdala activation, consistent with animal models indicating that the infralimbic cortex acts to inhibit the amygdala response during the recall of extinction. A separate study showed that cortical thickness at a specific locus within medial OFC (mOFC) was directly correlated with the magnitude of extinction recall. Notably, studies of memory reconsolidation suggest that it is primarily mediated by the amygdala and does not rely on prefrontal engagement.

The role of the hippocampus in contextual learning has been supported by human imaging and lesion studies. Neuroimaging studies show hippocampal activation during contextual conditioning and in the context-

dependence of extinction retention. Lesions of the hippocampus in humans are associated with impairment in contextually-mediated reinstatement of an extinguished fear, consistent with the animal literature.

Taken together, these data implicate the amygdala, vmPFC, and hippocampus in human fear-conditioning/extinction processes and provide tools for investigating the neural substrates of these functions across the anxiety disorders.

Cognitive and Emotional Activation Probes

Paradigms comprising cognitive and emotional components have been used to study emotion processing and anxiety, as well as to probe the neural circuitry hypothesized to be involved in anxiety. These tasks have most commonly involved viewing pictures of emotionally expressive human faces or of other emotionally valenced visual stimuli. Humans have a vast amount of experience with facial expressions of emotion, which convey important information about the environment. Given this learning and reinforcement history, facial expressions of emotion can be conceptualized as CS.

Functional imaging studies have shown that passive viewing of human face stimuli produces reliable amygdala activation. The amygdala responds to fearful, happy, neutral, and angry faces; however, its responses appear to be greater to fearful faces than to neutral, happy, or angry faces. Some studies suggest that one reason the amygdala responds "preferentially" to fearful faces is that they predict uncertainty (i.e., increased probability of threat without informing its nature or location, as compared to angry faces which embody a direct threat). Consistent with the amygdala's responsivity to uncertainty, the amygdala robustly activates to surprised faces, which are ambiguous in nature (i.e., they could predict either a positive or negative event). Fearful and surprised faces both indicate an important but unknown event, and they share common features such as eye-widening.

Moreover, the processing of face stimuli recruits a broad array of other brain areas in addition to the amygdala (e.g., visuotemporal cortices, PFC), allowing assessment of the extended functional anatomical networks involved in the evaluation of emotional stimuli. Moreover, the development of paradigms that present faces using backward-masking techniques (so that the subject is not explicitly aware of having seen the masked face stimulus) provide an additional tool for exploring automatic amygdala responses relatively dissociated from top–down influences related to cognitive responses to the stimuli. Furthermore, the repeated presentation of face stimuli can be used to demonstrate habituation of amygdala responses, providing potential tools for characterizing habituation profiles in the amygdala. Other paradigms that involve viewing pictures (such as the International Affective Picture System) have indicated robust and reliable amygdala activation in response to emotionally valenced versus

“neutral” visual stimuli.

Paradigms that measure emotional interference and attention bias have commonly been used to study anxiety disorders because the abilities to detect emotional interference and to suppress attention/responses to a particular information stream (e.g., disorder-relevant information) in favor of another stream (e.g., nondisorder-relevant information) are particularly relevant to anxiety. The ACC plays a critical role in signaling the presence of conflict, and dorsal and ventral subregions of the ACC tend to signal cognitive interference (e.g., during the Counting Stroop task) and emotional interference (e.g., during the Emotional Counting Stroop task), respectively. The lateral PFC is more directly involved in augmenting cognitive control and mediating attentional bias. Patients with anxiety disorders show negative attention bias, such that they have difficulty suppressing attention to and redirecting attention away from threat-related stimuli.

The relationship between brain activation during an emotional event and the enhancement of long-term memory for that event also has been studied using functional neuroimaging. As healthy humans read stories of differing emotional valence, the magnitude of physiological activation in the amygdala correlated both with the negative emotional intensity and the subsequent recall performance of the narrative’s content. Physiological activity in the amygdala and hippocampus measured during memory encoding correlated with enhanced episodic memory for pleasant, as well as aversive, visual stimuli. The amygdalar role in modulating emotional memory thus may depend more generally on the degree of arousal or the behavioral salience associated with target information.

Interoceptive function has been explored with functional neuroimaging, with direct relevance to anxiety disorders in which interoceptive sensitivity can be heightened (e.g., in panic disorder). Human imaging studies have implicated the insular cortex in interoceptive function, using tasks such as the perception of one’s heartbeat, viewing faces expressing disgust, viewing disgusting pictures, or generating internal images of disgust.

NEUROIMAGING STUDIES OF ANXIETY DISORDERS

Theorists have extrapolated from the aforementioned data obtained in experimental animals to pose hypotheses regarding dysfunction in the homologous structures in humans with anxiety disorders. Specifically, it has been hypothesized that individuals with anxiety disorders exhibit intrinsically exaggerated amygdala hyperresponsivity and/or deficient top–down modulation of the amygdala response due to impairments in connectivity with the mPFC and/or the hippocampus. Behavioral data consistent with these hypotheses exist for some anxiety disorders within the context of fear-conditioning and extinction studies. For example, participants with PTSD were shown to acquire de novo conditioned

responses more readily and extinguish them more slowly than control samples. The neuroimaging data also appear to be compatible with such hypotheses, showing exaggerated amygdala responses and impaired function and/or abnormal structure in the mPFC and hippocampus.

Panic Disorder

The neurocircuitry underlying panic disorder encompasses a wide range of disparate elements, including dysregulated ascending noradrenergic and/or serotonergic systems, aberrant responsivity to panicogenic substances (e.g., CO_2) at the level of the brainstem, global cerebral abnormalities in lactate metabolism, and abnormal interactions among hippocampus, parahippocampal cortex, and amygdala. Although the behavioral sequelae of panic attacks (e.g., agoraphobic avoidance) putatively can be explained by learning theory and fear conditioning, a satisfactory model of panic disorder must also explain the occurrence of spontaneous panic episodes.

One hypothesis holds that a physiological (i.e., normal) anxiety response mediated by the anxiety/fear circuitry is recruited spontaneously as an aberrant event due to homeostatic deficits. This possibility is exemplified by the "false suffocation alarm" model and by theories regarding monoaminergic dysregulation. Another possibility is that panic attacks evolve in the context of what would be minor anxiety episodes, but because of failures in the systems responsible for limiting such normal responses, the anxiety response becomes excessive to an extent that is maladaptive. Similar to models of PTSD, hippocampal or other temporal cortical deficits might underlie such a mechanism in panic disorder. Finally, it is conceivable that panic attacks described as spontaneous (i.e., without identifiable precipitants) reflect fear responses to stimuli that are not processed at the conscious (i.e., explicit) level but, instead, recruit the fear circuitry without awareness (i.e., implicitly). According to this model, panic disorder might be characterized by amygdala hyperresponsivity to subtle environmental cues, triggering full-scale fear responses in the absence of conscious awareness. Furthermore, the absence of explicit awareness may be attributable to dysfunction involving hippocampal or parahippocampal regions, which may permit abnormal escalation in response to contextual stimuli that are not recognized as inherently threatening.

Functional Imaging Findings in Panic Disorder. The baseline or "neutral" state in panic disorder is characterized by mild to moderate levels of chronic anxiety (termed "anticipatory anxiety"). Thus, resting-state studies, during which participants are asked to simply relax in the absence of a task, have been used to characterize brain function at rest. Both PET and fMRI studies have identified altered resting-state rCBF and functional

connectivity in patients with panic disorder. Specifically, patients show altered function in regions such as the parahippocampal gyrus, hippocampus, superior temporal gyrus, and increased ACC–precuneus functional connectivity at rest. These findings suggest distinct patterns of brain function even in the absence of symptom provocation; however, it remains unclear to what extent these alterations cause chronic anxiety versus reflect the patients' anxious state.

Neuroimaging studies performed during symptom provocation in panic disorder have employed pharmacological challenges and cognitive/emotional challenge. SPECT and PET studies using pharmacological challenges have been mixed in their results. Global reductions in CBF, reduced rCBF in frontal cortex, and rCBF increases putamen and ACC have been shown among patients with panic disorder who experienced panic attacks. Findings for the insula have been inconsistent, with some evidence for reduced rCBF and other evidence for increased rCBF during panic attacks. Finally, several participants have had a spontaneous panic attack during imaging. Results have been mixed, showing a range of responses including decreased rCBF in right OFC, subgenual PFC, ACC, and anterior temporal cortex in one case; increased amygdala, insula, and PFC activation in another case; and increased PFC in another case with a less intense panic attack.

Panic-related stimuli, such as script-driven imagery, agoraphobic-specific pictures, and panic-related visual stimuli have been used to probe neural correlates of anxiety in patients with panic disorder. Patients have shown exaggerated activation in inferior frontal cortex, cingulate cortex, OFC, hippocampus, insula, and striatum. Neuroimaging studies have also presented emotional stimuli that are not disorder-specific, such as facial expressions of emotion, to probe for brain-related changes in panic disorder. Patients generally showed increased amygdala and insula activation, and less vmPFC and ACC activation to facial expressions, but findings have been mixed for some regions such as the amygdala and ACC. Similar patterns of increased insula activation have been observed to unpredictable, but not predictable, aversiveness, consistent with a general theme of intolerance of uncertainty that is observed in anxiety disorders.

Studies of CBT and pharmacological treatment have begun to identify potential mechanisms and predictors of treatment response in panic disorder. Medicated patients have shown distinct patterns of brain function and structure, including increased FA in the uncinate fasciculus following symptom remission, compared with nonmedicated patients. Symptom remission following CBT has been associated with less IFG activation and greater functional connectivity of the IFG with the amygdala, hippocampus, mPFC, and lateral PFC. Beyond identifying potential mechanisms underlying effective treatment, other efforts have focused on identifying predictors of treatment response, such as stronger negative functional

connectivity between the ACC and amygdala. Together, these studies suggest that treatments for panic disorder may act by altering frontolimbic connections and highlight the potential for neuroimaging to contribute to personalized medicine.

Complementary Neuroimaging Findings: Structure and Chemistry. Volumetric MRI studies of panic disorder have reported smaller OFC, putamen, and overall temporal lobe volume, and lower gray matter density in parahippocampal cortex in participants with panic disorder compared to controls. Increased volume has been reported in the insula and brainstem. There is some evidence for smaller amygdala volume in this population being negatively associated with anxiety severity, but other studies have not found abnormal amygdala volumes.

With respect to neurochemical abnormalities in panic disorder, a series of MRS studies demonstrated that in comparison with healthy controls, participants with panic disorder showed a significantly greater rise in brain lactate in response to the same level of hyperventilation. These studies also showed that the mean brain lactate level during lactate infusion was significantly greater in participants with panic disorder than controls. These data suggested reduced clearance, rather than higher production, of lactate in panic disorder. The broad distribution of this phenomenon further suggested that a global or widespread abnormality in cerebrovascular function exists in panic disorder.

Although receptor imaging studies of panic disorder focusing on benzodiazepine binding have yielded inconsistent results, they have repeatedly implicated prefrontal and temporal cortical regions of abnormality. In a SPECT-iomazenil study, participants with panic disorder exhibited a greater left/right ratio in benzodiazepine receptor uptake than controls, especially in the PFC. Another SPECT-iomazenil study of panic disorder found reduced distribution volume in the left hippocampus and precuneus in participants with panic disorder versus controls; participants who experienced a panic attack during tracer uptake exhibited reduced distribution in PFC. In a PET study of [C-11]flumazenil binding, participants with panic disorder exhibited a global reduction in benzodiazepine binding relative to controls, which was most pronounced in right OFC and right insular cortex. Whereas these initial experiments may have been confounded by prior treatment, a fourth study using SPECT-iomazenil found that medication-naive participants with panic disorder exhibited significantly *elevated* benzodiazepine-receptor binding in the right supraorbital frontal cortex compared to controls.

The imaging data showing differences in central benzodiazepine binding in panic disorder are noteworthy in light of evidence that the behavioral sensitivity to pharmacological challenges involving this system are altered in panic disorder. Either oral or intravenous administration of

the benzodiazepine receptor antagonist flumazenil produced panic attacks and increased anticipatory anxiety in a large proportion of participants with panic disorder but not in healthy controls. In addition, the effects of diazepam on saccadic eye movement velocity were abnormally reduced in panic disorder, implying that the functional sensitivity of the $GABA_A$/benzodiazepine complex was diminished in brainstem regions controlling saccadic eye movements. Similarly, participants with panic disorder showed abnormally reduced sensitivity to the suppressant effects of diazepam on plasma norepinephrine, epinephrine, and heart rate.

A PET study that used [F-18]FCWAY to assess 5-HT_{1A}–receptor binding in panic disorder showed that, relative to healthy controls, participants with panic disorder exhibited diminished 5-HT_{1A}–receptor binding in the ACC, posterior cingulate cortex, and raphe. This abnormality extended to participants with panic disorder both with and without comorbid depression and to participants with and without prior exposure to antidepressant drugs.

Studies of the sensitivity of specific serotonin receptor subtypes also implicated 5-HT_{1A}–receptor function in the pathophysiology of panic disorder. The elevation of plasma ACTH and cortisol and the hypothermic responses to the 5-HT_{1A}–receptor partial agonist ipsapirone were blunted in participants with panic disorder relative to healthy controls.

Summary. Taken together, the neuroimaging data suggest that panic disorder is associated with abnormalities in hippocampal/parahippocampal activity at rest, and that during acute panic, patients exhibit activation of the insular and striatal regions, as well as reduced activity in widespread cortical regions, including the PFC. Morphometric MRI studies also have reported abnormalities of gray matter volume involving the parahippocampal/temporal cortex. MRS studies of brain lactate implicate a global phenomenon consistent with an exaggerated hemodynamic response to hypocapnea, underscoring the potential limitations with using symptom-provoking paradigms (such as fear conditioning) to study panic disorder with fMRI or CBF-based methods. Similarly, receptor-binding studies suggest widespread abnormalities in the GABA/benzodiazepine system, which also appear to be most pronounced in prefrontal, paralimbic, and temporal cortices. In addition, 5-HT_{1A}–receptor binding abnormalities have been demonstrated within the raphe and cingulate cortex. Consistent with prevailing neurobiological models of panic disorder, fundamental abnormalities in monoaminergic neurotransmitter systems originating within the brainstem may underlie the abnormalities of metabolism, hemodynamics, and chemistry found in widespread territories of cortex. Moreover, these abnormalities conceivably may mediate the increased risk for developing

initial panic attacks. Furthermore, the regional abnormalities within the temporal lobes provide support for theories regarding dysfunctional influence over the amygdala via the temporal cortex and/or the hippocampus in panic disorder. These abnormalities may relate to the liability for developing panic disorder following initial panic attacks.

Specific Phobias

Phobias may represent the product of dysregulated systems specific to assessing potentially threatening, anxiety-provoking, or disgusting stimuli or situations. Neuroimaging data from humans are beginning to provide an empirical basis for neurocircuitry models of specific phobias. Specifically, although the amygdala had been the theoretical focus because of its critical role in mediating conditioned and innate fear, the human imaging data more directly have implicated anterior paralimbic regions and sensory cortical regions known to interact with the amygdala.

Functional Imaging Findings. Meta-analyses of brain activation in patients with specific phobias show consistent findings of increased insula and amygdala activation to phobic stimuli, as well as some evidence for increased activation in the ACC, OFC, thalamus, and cerebellum. Studies of specific phobias principally have used fMRI and PET symptom-provocation paradigms. For example, patients with dental phobia exposed to cross-modal symptom provocation (drilling sounds and dental images) showed increased ACC, insula, OFC, and thalamus activation to the auditory but not visual phobic stimuli. Several creative fMRI paradigms have been used to probe the nature of phobic-related brain responses in spider phobia. In one study, patients were presented with pictures of spiders or neutral pictures under predictable conditions (phasic) or unpredictable conditions (sustained). Amygdala activation was heightened during phasic fear to phobic stimuli, whereas heightened BNST and ACC activation was associated with sustained fear to phobic stimuli. Patients with spider phobia also showed stronger BNST–amygdala functional connectivity. These findings highlight the dissociation between the role of the amygdala in acute fear and the role of the BNST in sustained anxiety. Based on studies using nonphobic stimuli (e.g., facial expressions of emotion), amygdala hyperactivity in participants with specific phobias appears to be unique to phobic stimuli.

These findings suggest hypersensitivity to fear-related cues, which may be due in part to disruptions in cortical control regions or their connectivity with limbic and paralimbic regions. In a study of cognitive reappraisal in spider phobia, participants were asked to upregulate, downregulate, or simply look at pictures of spiders and aversive pictures not related to spiders. Patients with spider phobia showed less vmPFC activation during automatic regulation (looking) of spider versus aversive pictures, as well as

less dACC/dmPFC activation during effortful regulation (downregulate versus look) to spider versus aversive pictures.

Several studies have attempted to identify common versus distinct neural substrates of specific phobias by subtype (e.g., small animal versus blood-injection-injury). Initial results suggest that blood-injection-injury (e.g., dental phobia) may be more related to alterations in the PFC, thalamus, and occipitotemporoparietal regions, whereas spider phobia may be more related to dACC and insula disturbances. Both groups showed quicker time-to-peak in amygdala reactivity than healthy controls. There is also evidence that increased activation in sensory association cortices corresponds to the sensory modality associated with the symptomatic state in specific phobias. In this context, it is noteworthy that the amygdala sends anatomical projections to sensory cortices, and stimulation of these projections "tunes" or facilitates the responsivity in those areas during amygdala activity (e.g., in response to threatening stimuli).

Neuroimaging findings have begun to elucidate neural mechanisms associated with effective treatment of specific phobias. A meta-analysis suggested that exposure therapy for specific phobias was associated with reduced activation in the frontal cortex, limbic regions, and cerebellum. A longitudinal study of patients with spider phobia who underwent a single 2-hour exposure therapy session found that patients showed increased amygdala, insula, and ACC activation to phobic versus neutral stimuli prior to treatment. Effective treatment was associated with reduced amygdala activation and increased prefrontal engagement; however, 6 months later the participants showed persistent dampened amygdala activation in the absence of prefrontal engagement. In addition, lower activation in the visual cortex shortly after the exposure session predicted better treatment outcome 6 months later. The findings suggest that exposure therapy leads to lasting reorganization of brain responses to previously feared stimuli, such that limbic activation is normalized without the need for active prefrontal down regulation.

Complementary Neuroimaging Findings: Structure and Chemistry. Structural MRI studies have begun to identify volumetric differences associated with specific phobias and how they might differ by phobic stimulus. Participants with animal phobias show increased cortical thickness in pgACC, insular cortex, and posterior cingulate cortex. Participants with dental phobia show increased gray matter volume in sgACC, insula, OFC, and PFC. Participants with spider phobia show cortical thinning in the ACC, as well as reduced amygdala volume that correlates negatively with symptom severity. Each of these studies shows specificity which may be related to the complex neurocircuitry underlying representation of these complex categories of stimuli.

Summary. Imaging findings in specific phobias suggest activation of anterior paralimbic regions (i.e., anterior insula) and sensory association cortices corresponding to the sensory modality associated with the symptomatic state. Such results are consistent with a hypersensitive system for evaluation of, or response to, specific threat-related cues—a system in which the amygdala plays a major role. The fMRI data suggest that the amygdala hyperresponsivity seen in specific phobias occurs selectively to phobia-relevant stimuli and may be related to disrupted interactions with cortical regions that typically play a role in emotion regulation.

Social Anxiety Disorder

As in specific phobias, models of SAD consider increased sensitivity to the potential threat value of a specific class of stimuli—in this case, socially relevant stimuli. Although possibilities in this regard include enhanced sensitivity to social threat or scrutiny, it is also hypothesized that heightened responses to novel social stimuli and a decreased reward value of positive social stimuli play roles in SAD. Research paradigms in this area thus have focused on primary responses to social stimuli (e.g., faces), varied for valence, arousal, and novelty, as well as self- versus other-focused attention.

Functional Imaging Findings. PET and fMRI studies using a variety of symptom provocation paradigms, including public speaking challenge and contemptuous faces, have consistently shown exaggerated amygdala activity in patients with SAD compared with healthy controls. Findings have suggested that these differences in amygdala function persist across various conditions including facial expressions of both low- and high-level intensity, suggesting that the amygdala may be activated at a lower threshold of social stimulus in SAD. Moreover, several of the studies have found that amygdala activity in patients correlates with the severity of their social anxiety. Although many studies of SAD have focused on the amygdala, neuroimaging studies highlight alteration in a more distributed network of regions. Several studies have reported heightened insula activation to emotional stimuli and less striatal activation (e.g., during implicit learning) in SAD, and some evidence suggests altered function of the OFC, ACC, mPFC, and parahippocampal gyrus. Functional connectivity analyses suggest that exaggerated emotional reactivity (e.g., in the amygdala, insula) in SAD may result from weaker frontolimbic connections.

The degree to which abnormal function is specific to social versus nonsocial stimuli may inform the pathophysiology of SAD. Several studies suggest that patients with SAD show altered brain responses to social, but not nonsocial, stimuli. For example, an fMRI study that compared participants with SAD and controls during exposure to human face stimuli

or aversive odors found that the hyperresponsivity of the amygdala was selective for face stimuli in SAD. Patients also display increased amygdala and mPFC activation to self-referential critical comments, but not to critical comments of others or for neutral or positive comments about the self or others. Thus, it may be that processing of self-focused criticism or judgment is especially altered.

fMRI studies of pre- and posttreatment brain function point to possible mechanisms underlying effective treatments for SAD. In a randomized controlled trial of CBT for SAD, patients showed greater reduction in negative emotion and corresponding greater dlPFC and dmPFC activation and stronger negative dmPFC–amygdala connectivity during cognitive reappraisal following CBT, compared with a waitlist control condition. In addition, patients who received CBT showed earlier onset of dmPFC activation. Several studies have also examined fMRI predictors of treatment response in SAD. Prefrontal and amygdala activation, as well as amygdala–ACC functional connectivity, have predicted response to CBT, such that patients with greater regulatory capacity over emotional reactivity to faces may be more likely to subsequently benefit from treatment. Brain-based measures have also been shown to enhance prediction of treatment response in SAD over and above clinical measures alone, highlighting the potential contributions of neuroimaging for matching patients with treatments.

Complementary Neuroimaging Findings: Structure and Chemistry. Investigations of structural connectivity provide evidence of lower FA in the uncinate fasciculus, consistent with weaker amygdala–OFC functional connectivity observed in fMRI studies. In addition, patients with SAD show increased cortical thickness in distributed regions including the ACC, insula, dlPFC, and parietal cortex. Other studies have found lower gray matter volume in the amygdala, hippocampus, thalamus, and precuneus, and higher gray matter volume in regions including the lateral OFC, parahippocampus, and cerebellum.

Summary. Functional neuroimaging studies of SAD indicate that medial temporal lobe structures show exaggerated responsivity specifically to social stimuli. In particular, amygdala hyperresponsivity was demonstrated in response to social threat (e.g., harsh facial expressions, perceived evaluation during public speaking). These data are consistent with a hypersensitive system for the evaluation of faces with respect to social threat or salience as a neural substrate for the underpinnings of social anxiety in SAD, which may stem in part from weakened cortical control.

Generalized Anxiety Disorder

Neurocircuitry models for GAD have been less explored and emerged more recently than models for PTSD, panic disorder, and phobias. As compared with the focus on fear conditioning in PTSD and phobias, a satisfactory model of GAD must also account for the overgeneralization of fear that is a core feature of the disorder. Neuroimaging efforts to elucidate neural underpinnings have thus focused on core features of emotion dysregulation and excessive worry, as well as the overgeneralization of conditioned fear. High rates of comorbidity with major depressive disorder (MDD) and other anxiety disorders have presented challenges to testing neurocircuitry models of GAD. Though many patients with GAD have other comorbidities, the studies presented here have attempted to recruit "pure" GAD samples or to control for comorbidities by comparing patients with only GAD to those with other comorbidities.

Functional Imaging Findings. Findings on amygdala function in GAD have been mixed and point to a more complex picture than the classic pattern of amygdala hyperactivity to negative emotional stimuli seen in SAD, and phobias and stress-related disorders such as PTSD. While some fMRI studies have found that patients with GAD show no differences in amygdala reactivity to fearful faces, or even less reactivity than healthy controls in one study, evidence suggests that GAD may be characterized by nonspecific amygdala hyperactivity. For example, patients with GAD show elevated amygdala activation to both negative and neutral stimuli, and to both incongruent and congruent trials during emotional interference.

Given the role of the BNST in sustained anxiety, it has also been hypothesized that patients with GAD may display altered BNST function in addition to amygdala changes. During a gambling task designed to induce high uncertainty and anxiety about monetary loss, participants with GAD showed lower amygdala activation but higher BNST activation and higher sustained arousal throughout the task. The authors proposed that patients with GAD may disengage the amygdala and its response to acute stress earlier, while exhibiting increased BNST activation that maintains a more sustained state of anxiety. Further research on the temporal dynamics of these regions will help to clarify the nature of anxiety in GAD.

Consistent with a core deficit of emotion regulation in GAD, prefrontal function appears to be impaired in patients with GAD. One study used a worry induction procedure designed to target the excessive worry that characterizes GAD to test patients' activation of dorsal ACC and PFC regions, which are part of the "aversive amplification circuit." Participants with GAD and healthy controls both activated dACC and dmPFC during worry, but the participants with GAD failed to normalize this activation afterward. Instead, they showed persistent dACC/dmPFC activation even during resting-state scans that followed the worry induction. In this study, the dACC/dmPFC activation also correlated with patients' self-reported

worry symptoms. These findings may help to explain the sustained anxiety and worry that are classic features of GAD. Another line of work has attempted to understand prefrontal alterations in GAD that may relate to the overgeneralization of conditioned fear. In fear generalization paradigms, participants develop a conditioned response to a conditioned stimulus (e.g., a large rectangle) that is presented with an UCS (e.g., a shock). Participants then view target stimuli that become increasingly dissimilar from the CS+ (e.g., increasingly smaller rectangles). Patients with GAD show more overgeneralization (i.e., greater fear to the generalized stimuli, which are themselves nonthreatening), as well as less recruitment of the vmPFC that is associated with poorer discrimination between a CS+ and a generalized stimulus. These findings suggest neural substrates of fear overgeneralization and the failure to discriminate between threat and safety in GAD.

Investigations of connectivity are critical to understanding the neural bases of emotion dysregulation in GAD and how disrupted interactions between regions such as the amygdala and PFC might underlie patients' symptoms. Several independent studies provide evidence of altered corticolimbic functional connectivity in patients with GAD. A novel emotional conflict task involving facial affect identification was designed to test implicit emotion regulation in GAD. Participants viewed target facial expressions (fearful or happy) with overlying emotional words ("fear" or "happy") that either matched (congruent) or conflicted (incongruent) with the facial expression. Participants responded by identifying the facial expression, providing a measure of adaptation to emotional conflict through regulatory processes. In healthy participants, the task instantiated a strong negative coupling between the pgACC and the amygdala, consistent with downregulation of emotional reactivity despite the absence of explicit instructions to regulate emotion. In contrast, patients with GAD failed to activate the pgACC and did not show negative cingulate–amygdala functional connectivity. Consistent with altered amygdala connectivity, patients also show disrupted functional connectivity of the amygdala subregions at rest. These findings suggest that patients have impaired interactions between regions such as the ACC, that are important for exerting control, and subcortical limbic structures such as the amygdala, which may underlie deficits in emotion regulation.

Several functional neuroimaging studies have begun to examine possible predictors of treatment response or mechanisms of effective treatment in GAD. Results suggest that patients with higher anticipatory ACC activation, and with higher rostral ACC and lower amygdala activation to fearful faces, will respond better to venlafaxine. Patients who received citalopram showed lower activation in prefrontal and limbic regions to worry-related stimuli. Patients who received a course of cognitive behavioral therapy showed less amygdala and subgenual ACC activation to

fearful and angry faces, and increased insula activation to happy faces, suggesting a possible mechanism through which treatment improves symptoms.

Complementary Neuroimaging Findings: Structure and Chemistry. Consistent with altered corticolimbic functional connectivity, a study examining structural connectivity using DTI observed reduced FA in the bilateral uncinate fasciculus in patients with GAD. Moreover, lower FA in the uncinate fasciculus was associated with weaker negative functional connectivity between the pregenual cingulate and the amygdala during the anticipation of aversive stimuli. Structural and functional connectivity changes have also been linked with poorer discrimination between threat and safety in GAD. Thus, these structural and functional alterations in connectivity provide evidence of neural bases for emotion dysregulation and fear overgeneralization that are central to GAD.

Summary. Neuroimaging studies of GAD highlight the role of disrupted functional (ACC-amygdala) and structural (uncinate fasciculus) connectivity. As opposed to a clear finding of amygdala hyperactivity as seen in other anxiety disorders, this work highlights the importance of the BNST for sustained states of anxiety that are characteristic of GAD. Findings are consistent with core features of emotion dysregulation, which may be associated with weaker frontoamygdala connectivity, and fear overgeneralization, which appears to relate to alterations in the vmPFC and its connections. Initial evidence from treatment studies suggests that cognitive behavioral therapy and pharmacological interventions may reduce symptoms by targeting frontoamygdala circuitry.

Commonalities and Specificity across Anxiety Disorders

In general, the imaging data suggest some commonalities among anxiety disorders, including panic disorder and GAD, that appear to share amygdala hyperresponsivity to both disorder-specific and general threat-related stimuli. In contrast, in specific phobias and SAD, amygdala hyperresponsivity appears to be limited to disorder-specific stimuli. Of interest, exaggerated anterior insular activation has been observed across all of the disorders.

With respect to hippocampal involvement, the data are compelling for panic disorder in which temporal lobe volumetric abnormalities are reported in extrahippocampal regions, and most resting-activity findings suggest a rightward shift in laterality of function. The hippocampus has been less implicated in specific phobias, SAD, or GAD. In PTSD, the hippocampus is abnormally small with respect to volume and shows increased resting activity together with deficient capacity for activation in the context of explicit learning paradigms.

Finally, with respect to the PFC, panic disorder is associated with deficient recruitment of pgACC, as previously found for PTSD. In contrast, in specific phobias, the pgACC cortex thickness is reportedly increased (as opposed to the decreases in volume reported for PTSD), and evidence of PFC dysfunction during emotional processing is lacking. For SAD, the literature is in disagreement, with some but not other studies reporting diminished neurophysiological responses in the pgACC. A recent meta-analysis of SAD, specific phobias, and PTSD suggested that PTSD was characterized by hypoactivation of the dorsal and rostral ACC, as well as vmPFC. The vmPFC also appears to be implicated in overgeneralization of fear in GAD. The differences in findings for PTSD may be consistent with its separation from anxiety disorders in the DSM-5.

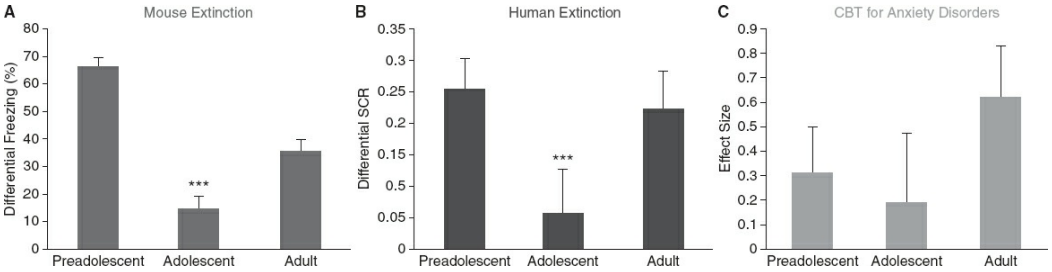

FIGURE 14.6–3. Fear extinction learning and improvement of anxiety by age. **A:** Adolescent (P29) mice show diminished extinction learning relative to adults (P70) and preadolescents (P23). **B:** Human adolescents (mean age: 13.9 ± 1.47 years, *n* = 25) show diminished fear extinction learning as indexed by skin conductance responses (SCR) compared with adults (mean age: 22.8 ± 2.57 years, *n* = 28) and children (mean age: 8.8 ± 1.78 years, *n* = 30). **C:** A similar pattern emerges when the effect size of cognitive behavioral therapy (CBT) relative to placebo is calculated separately for children (CBT: children, mean age: 9.46 ± 1.36 years, *n* = 90; adolescents, mean age: 14.34 ± 1.74 years, *n* = 49; placebo: children, mean age: 9.20 ± 1.31 years, *n* = 50; adolescents, mean age: 14.55 ± 1.64 years, *n* = 26) and adults. All results presented as mean ± SEM. *$p < 0.001$ compared with other groups. (Adapted from Drysdale AT, Hartley CA, Pattwell SS, et al. Fear and anxiety from principle to practice: implications for when to treat youth with anxiety disorders. *Biol Psychiatry*. 2014;75(11)e19–e20.)

These profiles of similarities and differences among anxiety subtypes including panic disorder, specific phobias, SAD, and GAD, as well as stress-related disorders (PTSD), concur with the nosological categorization of these disorders as being distinct from one another while sharing sufficient similarities in terms of the relevant neural circuitry to merit inclusion as anxiety or stress-related disorders. There has been little neuroimaging work conducted on separation anxiety disorder and selective mutism, thus the extent to which these disorders have shared versus unique neural substrates remains to be seen.

The Developmental Perspective and Markers of Risk, Resilience, and Treatment

Anxiety disorders peak during adolescence, and approximately 10 percent of youth have anxiety or stress-related disorders. Understanding

neurodevelopmental changes and the roles that they play in both the emergence of illness and in treatment efficacy is critical to reducing the burden of anxiety. Developmental neuroimaging studies can thus have a major impact on understanding risk, resilience, and treatment for pediatric anxiety disorders.

Evidence from animal and human studies shows dynamic changes in the neural circuitry supporting fear learning and extinction across development. The amygdala seems to undergo rapid structural development early in life and shows early functionality that is followed by continued change across development. The hippocampus also appears to develop earlier in life whereas the PFC shows protracted development and does not reach a mature functional architecture until early adulthood. Connections between the vmPFC and amygdala emerge during adolescence in rodents and show continued changes during this important developmental window in humans. Thus, although the basic structure of the amygdala is present at birth, the fine-tuning of its connections with PFC and hippocampus continues throughout childhood and adolescence. The earlier maturation of subcortical regions supporting fear reactivity, coupled with protracted maturation of PFC, may result in a period of altered fear regulation during adolescence. Less is known about the developmental trajectory of hippocampal connections with the amygdala and PFC.

Consistent with dynamic changes in brain function and connectivity across development, the processes of fear learning mediated by these structures undergo changes throughout childhood and adolescence. Several studies have examined fear acquisition in youth, with some evidence that simple discrimination between a CS+ and a CS− continues to improve linearly with age. The ability to express and extinguish fear undergoes dynamic, nonlinear changes across normative development. Translational studies across mice and humans have demonstrated a period of diminished extinction of cued fear during adolescence, relative to childhood and adulthood, which may be due to protracted maturational changes in the PFC and its connections with the amygdala (Fig. 14.6–3). These findings have raised the question of whether adolescents may benefit from treatments that are optimized to bypass prefrontally-mediated extinction processes. Because memory reconsolidation does not rely on the PFC, a recent study tested its utility as a novel way to reduce fear for adolescents. Consistent with prior work in adults, adolescents showed attenuated fear memory following memory reconsolidation (Fig. 14.6–4). Several studies have examined fear conditioning in youth with anxiety disorders, but the extent to which development in anxiety disorders is marked by similar changes in fear learning that have been found in typically developing youth remains a critical gap in the literature.

Neurodevelopmental and behavioral changes across development have

important implications for how stress and anxiety influence the brain and the efficacy of treatments targeting anxiety disorders at different developmental time points. Sensitive periods refer to times in development when heightened neuroplasticity renders the brain especially amenable to environmental influences. During these periods, environmental input can lead to a series of developmental cascades that ultimately have significant influences on behavior, of a positive *or* negative nature. Especially relevant to stress-related disorders, a sensitive period may render the brain more capable of responding to stress in adaptive ways. It could also magnify consequences of stressful life events in maladaptive ways. By contrast, stress that occurs during windows of reduced plasticity (e.g., after the closing of a sensitive period) may yield a brain that is less capable of remodeling itself. Thus, sensitive periods in neurodevelopment may render the developing brain more vulnerable to the effects of later stress, but they could also serve as windows of opportunity, during which there is increased potential for positive adaptation or effective intervention. Dynamic changes in brain and behavioral development likely influence how stress at unique developmental time periods alters the brain and how children and adolescents cope with stressors. Exacerbations of transient imbalances in brain circuitry, such as the effects of acute or chronic stress, may lead to altered stress reactivity and ultimately increase the risk for anxiety disorders. Neuroimaging studies across childhood and adolescence can inform the timing and nature of sensitive periods with direct relevance for risk and treatment of anxiety.

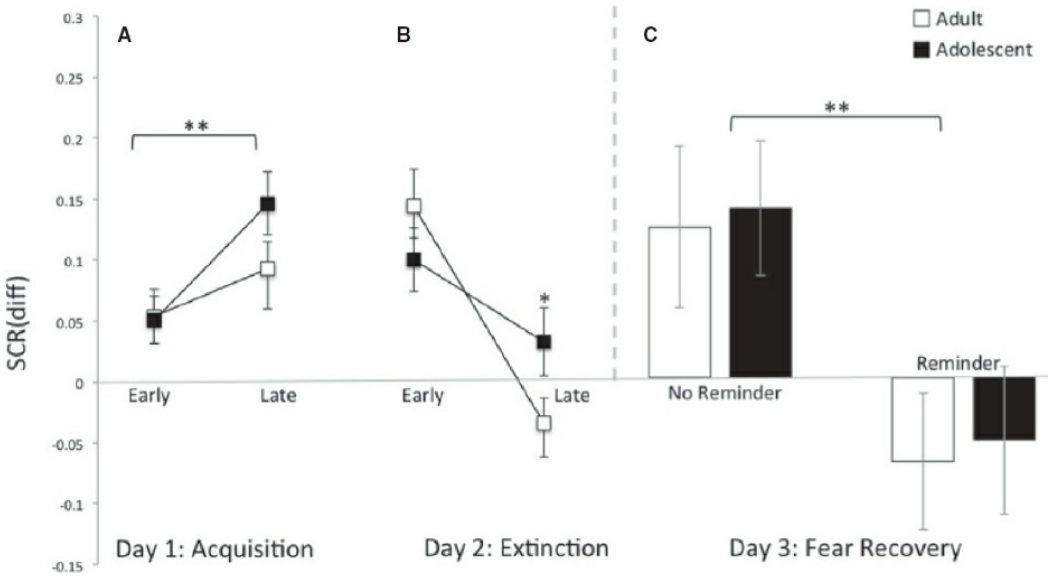

FIGURE 14.6–4. Acquisition, extinction, and recovery of fear memory by age group. **A:** There were no differences in differential skin conductance response (SCR) of the CS+ and CS − by age group during acquisition [F(1,70) = 0.979, p = 0.33] and a main effect of time [F(1,70) = 4.564, p = 0.036] in the mean SCR difference (CS+ − CS−) score from early to late trials. **B:** There was an interaction of age group × time in extinction learning as indexed by differential SCR of the CS+ and CS− [F(1,70) = 3.913, p = 0.05]. Post hoc t-tests showed significant within-session extinction learning for adults (t = 4.34, p < 0.0002) but not for

adolescents (t = 1.78, p = 0.08). All results are presented as a mean ± SEM, two-tailed t-test. **C:** Diminished fear memory with reconsolidation update. Participants who were reminded of the conditioned stimulus 10 minutes prior to extinction showed no recovery of fear 24 hours later, as indexed by SCR responses to the first CS+ trial of re-extinction (experimental day 3). There was a main effect of experimental condition [F(1,70) = 6.263, p = 0.015] and no age group × experimental condition interaction [F(1,70) = 0.002, p = 0.966]. All results are presented as a mean ± SEM, *p = 0.05, **p < 0.05. (From Johnson DC, Casey BJ. Extinction during memory reconsolidation blocks recovery of fear in adolescents. *Sci Rep.* 2015;5:8863.)

Developmental neuroimaging studies hold great promise for informing the nature of risk factors for anxiety disorders, and there is strong continuity between pediatric and adult anxiety. Individual differences, even in adulthood, often stem from early-life influences of environments (e.g., parental modeling of anxiety, stressful or traumatic events), genetic factors, or temperament. As an example, behavioral inhibition (BI), a behavioral phenotype evident in childhood, is characterized by withdrawal/avoidance and reactivity to novel stimuli. This temperament confers increased risk for pediatric anxiety disorders, which in turn confer increased risk for adult anxiety disorders. Because developmental studies permit the investigation of new-onset disease in a context that minimizes confounds of effects of illness chronicity and treatment, they hold the potential to identify the neural underpinnings of risk and predictors of disease onset. The elucidation of markers of risk is critical for developing strategies of prevention and early intervention. However, the distinction of markers of risk from markers of disease per se remains challenging. The example of BI to the unfamiliar illustrates this point. BI has been associated with neuroimaging findings of exaggerated amygdala responses to novel versus neutral faces. This finding underscores the potential liability in cross-sectional studies of adults with, as compared to those without, a given disorder with respect to interpreting between-group differences; although investigators may be inclined to attribute such differences to the pathophysiology of disease, such experiments cannot rule out the possibility that observed differences instead represent markers of risk. To address such confounding influences, investigators must study comparison groups of participants who have the risk factor but not the full disease. Moreover, pursuing the developmental perspective leads to the consideration of normal versus pathological trajectories over time with respect to brain structure and function. Brain-based criteria for disease must account for age, sex, medication effects, and other pertinent temporal variables (e.g., menstrual phase, diurnal variation).

Neuroimaging studies of pediatric anxiety have provided critical insight into the neural substrates of anxiety in children and adolescents, with initial evidence of how developmental processes may relate to risk for anxiety. Developmental neuroimaging involves unique challenges, such as dissociating between developmental processes and disease effects, and

adapting methods to be developmentally appropriate. Fear-conditioning paradigms used with children and adolescents have traditionally adapted UCSs to be developmentally appropriate, such as aversive noises, aversive air puffs, or aversive images (e.g., of negative faces that morph to scream, or of animals such as snakes). To date, these studies have primarily focused on SAD and GAD, whereas panic disorder, specific phobias, selective mutism, and separation anxiety disorder remain less explored. Future work will be necessary to elucidate the neural substrates of these disorders during development and to expand developmental neuroimaging to understand predictors of treatment response, which remain largely unexplored among children and adolescents.

Social Anxiety Disorder in Children and Adolescents. The incidence of SAD peaks during adolescence, when social stimuli become increasingly salient even in normative development. Several creative fMRI paradigms have been developed to examine social interaction during adolescence, and findings suggest important changes in brain responses to social stimuli in adolescents with SAD. During an "online" chatroom task, adolescents with SAD showed increased amygdala activation when anticipating evaluation from peers they rated as undesirable for social interaction. In addition, they displayed increased positive coupling between the amygdala and vlFPC, which was associated with symptom severity. A follow-up study showed that socially anxious adolescents also showed less activation in the nucleus accumbens during anticipation of feedback from peers they rated as desirable for social interaction. In addition, anxiety interacted with age, such that age was positively associated with rostral ACC activation for rejected peers and negatively associated for selected peers among participants with SAD, but the opposite was true for nonanxious participants. The authors suggested that increased regulatory resources may be devoted to potentially adverse social outcomes as individuals with SAD mature. An fMRI study in which participants received feedback from peers also showed age-related brain changes with anxiety. Specifically, adolescents with SAD showed increased striatal activation and negative frontostriatal functional connectivity during unexpected positive feedback from highly valued peers, whereas this pattern was not evident for adults with SAD or for healthy adolescents and adults. One longitudinal study found that N400 amplitudes and amygdala reactivity to facial expressions at age 8 to 9 years in healthy participants predicted symptoms of SAD at age 14 to 15 years, suggesting that increased limbic reactivity might predate or index risk for SAD. Much remains unknown about how developmental processes interact with anxiety, but these findings highlight the importance of how the underlying mechanisms of consequences of anxiety may differ across development.

Generalized Anxiety Disorder in Children and Adolescents. Consistent with studies of GAD in adults, studies of children and adolescents with GAD have demonstrated mixed findings with regard to amygdala hyperactivity. In one study, participants rated facial expressions of emotion, and the investigators manipulated attention. Youth with GAD showed greater amygdala, ventral PFC, and ACC activation to fearful than happy faces when attending to their own subjective fear, whereas comparison youth did not. In another study, youth with GAD did not show differences in amygdala reactivity to angry faces compared with controls. However, this may be a function of the design, as a subsequent study showed amygdala hyperactivity in youth with GAD to masked angry faces (i.e., when the faces were presented very briefly). Patients' amygdala reactivity was positively correlated with the severity of their anxiety symptoms, and they also displayed weaker functional connectivity between the ventrolateral PFC and amygdala. Youth with GAD have also shown elevated ventrolateral PFC activation, which was associated with fewer symptoms, suggesting that such activity may be compensatory in youth with GAD. Though less is known about resting-state functional connectivity in pediatric GAD, evidence suggests that adolescents with GAD display altered functional connectivity between the amygdala and mPFC at rest, consistent with studies of GAD in adults.

Though neuroimaging studies of GAD have primarily focused on brain function in children and adolescents, initial evidence has provided clues to structural changes and neural correlates of treatment response in pediatric GAD. Existing studies suggest that youth with GAD have larger amygdala volumes and increased cortical thickness across regions including the vmPFC, inferior and middle temporal cortex, and lateral occipital cortex. Investigations of treatment suggest that higher amygdala reactivity pretreatment predicts greater symptom improvement in cognitive behavioral therapy or fluoxetine treatment for youth with GAD and that these treatments were associated with higher vlPFC activation following intervention. Future studies will be necessary to evaluate the extent to which these developmental findings are similar to or distinct from adults with GAD.

Summary of Neuroimaging Findings in Pediatric Populations. The expanding literature on pediatric anxiety has begun to identify common neural mechanisms that are known to underlie anxiety disorders in adults, such as alterations in the amygdala and PFC. However, few studies have examined longitudinal developmental changes across childhood and adolescence in youth with anxiety disorders. Investigations that incorporate children, adolescents, and adults will contribute to the understanding of neurodevelopmental trajectories and changes related to anxiety. Identifying sensitive periods and factors that confer risk or

resilience will ultimately be essential for early identification and intervention, and for optimizing treatments for youth, one of the populations that is most vulnerable to anxiety and stress-related disorders.

FUTURE DIRECTIONS

The application of neuroimaging technology to investigations of the neurobiological basis of anxiety disorders is yielding seminal information about the functional anatomical and neurochemical correlates of both normal and pathological anxiety and fear states. The convergence of the systems neuroscience data obtained from human imaging studies with the physiological and anatomical data obtained from basic studies in experimental animals has facilitated the development of neurocircuitry models of anxiety disorders. The neuroimaging data from studies of panic disorder, specific phobias, SAD, and GAD support disease models that depend on altered interactions among the amygdala, hippocampus, and medial and orbital PFC.

Understanding the neurodevelopmental processes associated with anxiety disorders in childhood and adolescence will be critical to early identification of risk and targeting treatments for youth anxiety based on the biological state of the developing brain. Innovative study designs, which include developmental studies, also are needed to disentangle neural substrates of risk, compensatory processes, and epiphenomena from core substrates of pathophysiology. Moreover, some inconsistencies in the results across studies presumably reflect the neurobiological heterogeneity encompassed within criteria-based, psychiatric diagnoses. By characterizing this heterogeneity, research that integrates imaging data with genetic and neurochemical information ultimately may lead to the refinement of anxiety disorder phenotypes, the elucidation of their pathogenesis, and the development of targeted therapies that move toward personalized medicine to tailor treatments based on individual differences and developmental stages.

REFERENCES

Baur V, Bruhl AB, Herwig U, et al. Evidence of frontotemporal structural hypoconnectivity in social anxiety disorder: a quantitative fiber tractography study. *Hum Brain Mapp*. 2013;34(2):437–446.

Blair K, Geraci M, Devido J, et al. Neural responses to self- and other referential praise and criticism in generalized social phobia. *Arch Gen Psychiatry*. 2008;65(10):1176–1184.

Charney DS, Drevets WC. The neurobiological basis of anxiety disorders. In: Davis K, Charney DS, Coyle J, Nemeroff CB, eds. *Psychopharmacology. The Fifth Generation of Progress*. New York: Lippincott Williams & Wilkins; 2002:901.

Davis M, Shi C. The extended amygdala: are the central nucleus of the amygdala and the bed nucleus of the stria terminalis differentially involved in fear versus anxiety? *Ann N Y Acad Sci*.

1999;877:281.

*Drysdale AT, Hartley CA, Pattwell SS, et al. Fear and anxiety from principle to practice: implications for when to treat youth with anxiety disorders. *Biol Psychiatry*. 2014;75(11):e19–e20.

Etkin A, Prater KE, Hoeft F, Menon V, Schatzberg AF. Failure of anterior cingulate activation and connectivity with the amygdala during implicit regulation of emotional processing in generalized anxiety disorder. *Am J Psychiatry*. 2010;167(5):545–554.

Etkin A, Wager TD. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am J Psychiatry*. 2007;164(10):1476–1488.

Fredrikson M, Furmark T. Amygdaloid regional cerebral blood flow and subjective fear during symptom provocation in anxiety disorders. *Ann N Y Acad Sci*. 2003;985:341.

Goldin PR, Ziv M, Jazaieri H, Hahn K, Heimberg R, Gross JJ. Impact of cognitive behavioral therapy for social anxiety disorder on the neural dynamics of cognitive reappraisal of negative self-beliefs: randomized clinical trial. *JAMA Psychiatry*. 2013;70(10):1048–1056.

Guyer AE, Lau JY, McClure-Tone EB, et al. Amygdala and ventrolateral prefrontal cortex function during anticipated peer evaluation in pediatric social anxiety. *Arch Gen Psychiatry*. 2008;65(11):1303–1312.

Hahn T, Kircher T, Straube B, et al. Predicting treatment response to cognitive behavioral therapy in panic disorder with agoraphobia by integrating local neural information. *JAMA Psychiatry*. 2015;72(1):68–74.

Hauner KK, Mineka S, Voss JL, Paller KA. Exposure therapy triggers lasting reorganization of neural fear processing. *Proc Natl Acad Sci USA*. 2012;109(23):9203–9208.

Johnson DC, Casey BJ. Extinction during memory reconsolidation blocks recovery of fear in adolescents. *Sci Rep*. 2015;5:8863.

Klumpp H, Keutmann MK, Fitzgerald DA, Shankman SA, Phan KL. Resting-state amygdala-prefrontal connectivity predicts symptom change after cognitive behavioral therapy in generalized social anxiety disorder. *Biol Mood Anxiety Disord*. 2014;4(1):14.

LaBar KS, Gatenby C, Gore JC, LeDoux JE, Phelps EA. Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron*. 1998;20:937.

*LeDoux JE. *The Emotional Brain*. New York: Simon and Schuster; 1996.

Lissek S. Toward an account of clinical anxiety predicated on basic, neurally mapped mechanisms of Pavlovian fear-learning: the case for conditioned overgeneralization. *Depress Anxiety*. 2012;29(4):257–263.

McClure EB, Monk CS, Nelson EE, et al. Abnormal attention modulation of fear circuit function in pediatric generalized anxiety disorder. *Arch Gen Psychiatry*. 2007;64(1):97–106.

Monk CS, Telzer EH, Mogg K, et al. Amygdala and ventrolateral prefrontal cortex activation to masked angry faces in children and adolescents with generalized anxiety disorder. *Arch Gen Psychiatry*. 2008;65(5):568–576.

Munsterkotter AL, Notzon S, Redlich R, et al. Spider or no spider? Neural correlates of sustained and phasic fear in spider phobia. *Depress Anxiety*. 2015;32(9):656–663.

Myers KM, Davis M. Behavioral and neural analysis of extinction. *Neuron*. 2002;36(4):567.

Pattwell SS, Lee FS, Casey BJ. Fear learning and memory across adolescent development:

Hormones and Behavior Special Issue: Puberty and Adolescence. *Horm Behav*. 2013;64(2):380–389.

Phelps EA, Delgado MR, Nearing KI, LeDoux JE. Extinction learning in humans: role of the amygdala and vmPFC. *Neuron*. 2004;43(6):897.

*Pine DS. Research review: a neuroscience framework for pediatric anxiety disorders. *J Child Psychol Psychiatry*. 2007;48:631–648.

*Rauch SL, Drevets WC. Neuroimaging and the neuroanatomy of stress-induced and fear circuitry disorders: the agenda for future research. In: Andrews G, Charney DS, Sirovatka PJ, Regier DA, eds. *Stress-Induced and Fear Circuitry Disorders: Refining the Research Agenda for DSM-V*. Washington, DC: American Psychiatric Association; 2008:235.

Shin LM, Liberzon I. The neurocircuitry of fear, stress, and anxiety disorders. *Neuropsychopharmacology*. 2010;35(1):169–191.

Tromp DP, Grupe DW, Oathes DJ, et al. Reduced structural connectivity of a major frontolimbic pathway in generalized anxiety disorder. *Arch Gen Psychiatry*. 2012;69(9):925–934.

Vidal-Gonzalez I, Vidal-Gonzalez B, Rauch SL, Quirk GJ. Microstimulation reveals opposing influences of prelimbic and infralimbic cortex on the expression of conditioned fear. *Learn Mem*. 2006;13:728.

*Whalen PJ, Phelps EA. *The Human Amygdala*. New York: The Guilford Press; 2009.

Yassa MA, Hazlett RL, Stark CE, Hoehn-Saric R. Functional MRI of the amygdala and bed nucleus of the stria terminalis during conditions of uncertainty in generalized anxiety disorder. *J Psychiatr Res*. 2012;46(8):1045–1052.

▲ 14.7 Anxiety Disorders: Cognitive-Behavioral Therapy

JONATHAN D. HUPPERT, PH.D., AND EDNA B. FOA, PH.D.

In this section we will focus on cognitive-behavioral therapy (CBT) for *pathological anxiety* that leads to significant impairment in functioning or substantial distress. The fifth edition of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) specifies five primary anxiety disorders: phobias, panic disorder, agoraphobia, SAD, and GAD. Common to the anxiety disorders are intense, unrealistic anxiety, related to exaggerated and/or unrealistic cognitions about future threat of harm (either physical or psychological), physiological responses when confronted with anxiety-relevant stimuli, and behavioral tendencies to escape from or avoid anxiety triggers and to prevent anticipated harm. While the DSM-5 includes OCD and PTSD in separate categories from the anxiety disorders, most of the principles that will be discussed here also apply to both OCD and PTSD.

DESCRIPTION OF THE ANXIETY DISORDERS

Each anxiety disorder is distinguished by specific symptoms and specific negative cognitions. Panic attacks, for example, are discrete episodes of

intense anxiety associated with at least four physiological symptoms (e.g., rapid heart rate, chest pain, cold chills, or hot flashes) and/or cognitive symptoms (e.g., thoughts that one is going to die, go crazy, or lose control) that may occur in up to 30 percent of the general population in a given year, yet only 2.7 percent of the population has a diagnosable panic disorder during the same period. In addition to recurrent unexpected panic attacks, the diagnosis of panic disorder requires enduring anxiety about having another panic attack or concerns about the consequences of such attacks for one's well-being (see cognitive symptoms). The central feature of panic disorder, therefore, is the fear of one's own anxiety reactions because they are perceived as harmful (also known as "fear of fear"). Agoraphobic avoidance typically develops in response to an individual's attempt to prevent experiencing a panic attack, though it can also occur due to fears of other experiences, such as fear of vomiting or losing control of ones bowels. The avoidance in agoraphobia is related to the expectation that anxiety response or another trigger will lead to loss of physiological or behavioral control.

Individuals with other psychiatric disorders can have panic attacks, but they are usually not afraid of the consequences of such attacks per se. For example, individuals with a fear of flying may experience a panic attack while flying, and individuals with anxiety about public speaking may have a panic attack while giving a speech. However, in such situations, the individual is typically more afraid of the direct consequences of the situation and not of the fear itself. That is, the individual with a fear of flying is afraid that the plane will crash and the speech anxious individual is afraid of being rejected due to delivery of a bad speech. Similar conceptualizations can be made for SAD, in which the focus is on the fear of embarrassment and rejection; obsessive-compulsive disorder, in which the fear is of an idiosyncratic concern (contamination, fire, being an evil or harmful person, etc.); generalized anxiety, in which the worry is about specific future negative outcomes; posttraumatic stress disorder, in which the fear is about the consequences of thinking about the trauma or approaching situations that are similar to the trauma with the expectation of another traumatic event occurring in these situations, or inability to tolerate distress while thinking about the trauma; borderline personality disorder, in which the experience of intense emotion is the focus (and not its consequences); and specific phobia, in which the fear is circumscribed to a specific situation or object. With the exception of panic disorder, in all anxiety disorders the focus is on the harm related to the specific source of fear instead of on the panic itself.

Specific phobias are the most common of the anxiety disorders, and include fears of animals, insects, situations (e.g., heights), enclosed places (claustrophobia, which is phenomenologically close to panic disorder), and blood/injection phobias. Many other objects or stimuli can also be the

focus of concern (e.g., fear of vomiting, fear of the dark, fear of clowns, etc.). For specific phobias to meet criteria for a disorder, they must cause pronounced distress or interference in one's life. Therefore, if someone has a strong fear of snakes but lives in the city and has no fears of encountering snakes there (and has no plans for vacations in the wild), he or she will not meet criteria for a specific phobia.

SAD is characterized by a fear of embarrassment or criticism from others. Typically, the types of situations that elicit social anxiety can be divided into either performance situations or interpersonal interactions. While DSM-IV used a specifier of specific versus general social anxiety, with the latter defined as having four or more situations that are included in one's social fears, DSM-5 has eliminated this distinction because of a lack of reliability and specificity of the distinction. Indeed, there are some individuals who have one or two fears such as public speaking (the most common social concern) or using public bathrooms (also called shy bladder), but some have three, and some have four or more. The more situations one is concerned about, the more severe the impairment tends to be (almost by definition as it tends to affect more interactions, though even a specific fear can be very debilitating). Recent research has suggested that individuals who suffer from SAD have a fear of positive evaluation as well as a fear of negative evaluation. They also tend to have diminished positive affect, particularly due to a lack of pride (more than joy), which has been theorized to reflect some underlying low self-esteem.

GAD is the least focused of the anxiety disorders, and also has the greatest overlap with depressive disorders. GAD is one of the only psychiatric disorders with a specific cognitive criterion in the DSM-5: excessive, uncontrollable worry. Individuals with GAD tend to worry about minor matters (sometimes a distinguishing feature of GAD) as well as major ones (health, children's future, etc.). In addition to worry, individuals with GAD tend to have physiological responses that are less autonomic and more related to stress (e.g., sleep disturbance, muscle tension, irritability). GAD has the latest onset of the anxiety disorders.

Anxiety disorders are common (up to 29 percent of people experience at least one in their lifetime), persistent, and frequently co-occur with other disorders, making them a significant mental health problem. Fortunately, effective psychological and pharmacological treatments have been developed for the anxiety disorders. In fact, CBT has been endorsed as a first-line treatment for anxiety disorders by the United Kingdom's National Institute of Health and Clinical Excellence, the American Psychiatric Association Treatment Guidelines, and a number of other health care and government agencies around the world. These guidelines have been influenced by research suggesting that CBT for anxiety disorders is more cost-effective than medication or other treatments in the long term, and patient preferences tend to be 3:1 for psychosocial interventions over

medical ones. Now, the theory and practice of CBT for anxiety disorders is discussed and evidence of its efficacy is reviewed.

EMOTIONAL PROCESSING THEORY: A PSYCHOLOGICAL MODEL OF ANXIETY DISORDERS AND THEIR TREATMENT

Before describing the components of CBT and its efficacy in the treatment of the various anxiety disorders, a general psychological theory of anxiety disorders and the hypothesized conditions necessary for successful treatment are presented. Emotional processing theory, developed by Edna Foa and Michael Kozak, is based on the proposition that there is a cognitive *fear structure* that includes representations of stimulus, response, and meaning, and serves as a blueprint for avoiding or escaping from danger. The theory integrates information processing concepts and recent learning theories of acquisition and extinction. It posits that environmental stimuli activate matching information in the fear structure, resulting in the spreading of activation throughout the rest of the structure via associative connections, thereby activating the fear structure. A fear structure is adaptive if the stimuli represented in it are objectively harmful and the responses represented in it lead to effectively avoiding, escaping, or coping with the threat. A fear structure is pathological, as is the case in anxiety disorders, when the associations among stimulus, response, and meaning representations do not accurately reflect reality, such that harmless stimuli or responses are erroneously interpreted as being dangerous. For example, the fear structure for an individual with a dog phobia would include associations among representations of dogs; meaning representations of dogs as being dangerous and uncontrollable; and responses such as physiological arousal and various defensive behaviors (e.g., calling for help, running away). The fear structure in panic disorder is characterized by fear of physical sensations associated with panic symptoms (e.g., rapid breathing, heart palpitations) and interpretation of these sensations as indications of danger, such as a heart attack or going crazy. As a result of the erroneous meaning associated with these sensations, people with panic disorder avoid situations that they anticipate to give rise to panic or cause similar sensations, such as intense physical activity. The fear structure of SAD represents erroneous association between social interactions and negative consequences such as being perceived as boring or stupid. The fear structure in GAD is much more diffuse, as there is often not a central or "core fear."

Emotional Processing

Emotional processing refers to the changes in a pathological element of a fear structure (e.g., most dogs are not dangerous) that result in long-term fear reduction and resolution of the anxiety disorder. According to

emotional processing theory, two conditions are necessary for emotional processing to take place. First, the fear structure must be activated either through a match with stimuli in the environment (e.g., confronting the feared stimulus) or through symbolic means (e.g., imagining or thinking about the feared stimuli). Second, new information that is *incompatible* with the pathological aspects of the fear structure must be available and incorporated into the structure, thereby altering the structure and creating a more realistic or "nonfear" structure via the new, competing associations. These two conditions have been supported via research in animal and human studies, the latter of which include both basic research and treatment process research.

Emotional processing theory was explicitly designed to explain the phenomenology and mechanisms of anxiety disorders and the efficacy of exposure therapy in the treatment of pathological anxiety. Exposure therapy involves helping patients to systematically confront feared but otherwise safe stimuli in a manner that promotes eventual fear reduction. The efficacy of exposure therapy in reducing anxiety has been widely demonstrated, and exposure procedures are a basic component of nearly every CBT program that has been shown to be effective in the treatment of the various anxiety disorders. The relevance of emotional processing theory to other treatments for anxiety, specifically anxiety management training and cognitive therapy, is addressed later.

Recent elaborations of emotional processing theory specify two indicators of emotional processing that can be measured separately from treatment outcome. Following from the proposal that activation of the fear structure is necessary for change in the structure to occur, one indicator of emotional processing is sufficient emotional activation of the fear structure (emotional engagement), as reflected in assessment of fear and distress such as facial reactions, self-reports of fear and anxiety, or physiological reactions such as increased skin conductance or heart rate when the person confronts or thinks about the feared stimulus or response. The second indicator of emotional processing is reduction in the levels of fear or distress of across repeated sessions, or *between-session learning*. Fear reduction obtained in one session that does not at least partially transfer to a subsequent session (i.e., the lack of between-session learning) indicates that the fear structure was not altered and suggests that any previously observed fear reduction may have been brought about by some kind of compulsion, safety behavior, or other form of defensive tactic. By contrast, fear reduction that does transfer across sessions strongly suggests that the exposure experiences have led to some enduring change in the fear structure. Other indicators of between-session emotional processing may include reductions in fear-related cognitions or changes in biased cognitive processes, such as reductions in the investment of attentional resources to threat stimuli or decreased negative and/or increased positive

interpretations of ambiguous situations.

Earlier formulations of emotional processing theory included a third indicator of emotional processing: gradual reduction in these indexes of emotional reactivity to the feared situation over the course of an exposure exercise. In the context of exposure therapy, this is called *within-session habituation*. However, acute fear reduction can occur for reasons other than habituation, such as distraction or engaging in compulsions or other safety behaviors. Such measures limit emotional engagement and prevent the encoding of disconfirming information, thereby preventing therapeutic changes from taking place in the fear structure. For example, if a person with panic disorder who fears dying from a panic attack experiences an attack, thus activating the fear structure, but then immediately sits down and drinks water to prevent dying, they will not learn that the panic attack is not dangerous. Given the potential unreliability of this indicator and many null findings, revisions of emotional processing theory suggest that within-session habituation, although a possible indicator, is neither necessary nor sufficient.

Several studies conducted across a range of disorders, including specific phobias, SAD, panic disorder, OCD, and PTSD, investigated the relationship between indicators of emotional processing and treatment outcome of exposure therapy. Many of these studies showed a relationship between initial activation and long-term outcome and also between changes in either cognition or fear between sessions and outcome. The relationship between within-session changes in fear or cognition and outcome were less consistent.

Sources of Corrective Information and Conditions That Interfere with Emotional Processing

It is relatively easy to see how exposure therapy promotes emotional engagement as patients are encouraged to think and talk about feared thoughts, images, and memories; to confront feared objects and situations; and to engage in feared activities. It may not be as easy to see what sources of corrective or fear-disconfirming information is contained in an exposure exercise.

One important source of corrective information is within-session fear reduction, which provides new response and meaning information embedded in the experience of the *reduction* of physiological responding in the presence of feared stimuli that is incompatible with prior response information and the incorporation of nonthreat meaning related to this response. Take, for example, a patient with panic disorder; the experience of habituation disconfirms their belief that anxiety will persist unless the person escapes the situation, and it corrects misperceptions that physiological responses associated with fear and anxiety lead to heart attack or to going crazy. In-session activation also frequently teaches the

patient that the magnitude of the anxiety is not as great as initially expected. Experiencing the anxiety also frequently teaches the patient that they can tolerate the anxiety and that it does not last indefinitely.

Repeated and prolonged exposure may also provide corrective information regarding the actual likelihood or significance of feared consequences. Anxious people frequently overestimate the probability of feared events and the significance of those consequences, should they happen. For example, a person with social anxiety may make an unrealistically high estimate of the probability of being rejected in social interactions. Repeated exposure allows the person to test whether the feared consequence happens. Moreover, even if it does happen—for example, the person says something embarrassing—the person has the opportunity to learn that these consequences, although unpleasant, are not catastrophic.

Exposure to intrusive and unwanted thoughts can help a person to distinguish between unpleasant thoughts that occur in his or her mind and reality, or to separate thoughts from overt actions. Exposure therapy allows these individuals to learn that they can think upsetting thoughts without acting on them and that thinking about upsetting consequences do not make them happen.

Explicit provision of corrective information can also be extremely important and facilitate new associations within a fear structure. For example, psychoeducation about the causes and consequences of panic attacks can substantially reduce anxiety and panic attacks in approximately 30 percent of patients with panic disorder. This education includes corrective information regarding the likelihood of fainting as a result of panic attacks (which is close to zero in individuals other than those with blood/injection phobias) and distinguishing heart attacks and strokes from panic attacks, and a description of the biopsychosocial model of panic. On the other hand, many individuals require more-experiential exercises to incorporate such corrective information. For example, individuals with SAD can benefit significantly from provision of video and confederate feedback after an exposure to make explicit some of the corrective information that would otherwise not be observed. At times, within-session habituation does not occur in these circumstances, but between-session habituation does, making the latter a more reliable indicator of emotional processing than the former.

Failures of Emotional Processing

If repeated exposure to feared stimuli contains corrective or fear-disconfirming information, then why do most anxiety disorders have a chronic course? An important feature of all anxiety disorders, in addition to the experience of intense distress on exposure in the presence of the feared stimuli, is a strong tendency to escape from feared situations or avoid them

altogether. To the extent that such escape and avoidance strategies are effective at reducing anxiety, they are strengthened through the principle of negative reinforcement, making it even more likely that the person will engage in escape or avoidance in the future. Moreover, successful escape or avoidance limits the person's exposure to fear-disconfirming information. For example, a person who is afraid of dogs and successfully avoids even friendly dogs will not have the opportunity to learn that dogs can not only be safe but also fun to play with, and thus he or she will maintain the belief that avoiding dogs has prevented harm from occurring to them.

Two procedures studied in the context of exposure therapy that have been found to promote acute reductions in distress while in the presence of feared stimuli but that interfere with between-session habituation are distraction and the use of benzodiazepine medications. Several studies investigated the effects of distraction on exposure. These studies differed substantially in how distraction was manipulated, the population that was studied, and the duration of the exposure sessions, which may account for some of the variability in results across studies. Overall, a recent meta-analysis showed that, under some circumstances, distraction can produce acute reductions in anxiety that is equivalent to focused exposures. Results suggest that these effects are found more when multiple sessions are studied and distraction is more interpersonal/interactive in nature. It is possible that engagement of others during exposure recruits interpersonal emotion regulatory systems or that implicit learning via modeling or subtle verbal or nonverbal cues facilitates processing. Alternatively, such interactions may help facilitate engagement of approach/reward systems and deactivation of threat/fear systems. At times, uninstructed exposure seems to be superior to either focused or distracted exposure, suggesting that distraction may not be effective in some circumstances. It should be noted that the degree of distraction (e.g., complete engagement of cognitive resources on another task versus divided attention, splitting one's resources between the exposure and another task) may have different outcomes, and that the large majority of the studies conducted to date used analog participants or patients with specific phobia. Therefore, more research is need with more complex anxiety and related disorders.

When taken shortly before exposure to a feared stimulus, benzodiazepines can significantly reduce anxiety. As with distraction, however, the fear reduction that occurs with the medication does not transfer to subsequent tests in a nonmedicated condition and, indeed, appears to interfere with the normal process of between-session habituation. Researchers demonstrated this among a group of patients undergoing treatment for the fear of flying that involved taking two flights scheduled 1 week apart. Ninety minutes before taking the first flight, half of the patients took the high-potency benzodiazepine alprazolam (Xanax), and the remaining patients took a placebo. Patients taking the active

medication reported significantly less anxiety during the first trip than patients who took the placebo. Before the second flight, all patients were given a placebo pill. Patients who had taken the placebo before the first flight showed a reduction in their anxiety on the second flight, whereas those who had taken the active medication showed an increase in anxiety on the second flight, showing lack of extinction without fear activation. Similar findings were found in animal studies of extinction.

Other Approaches to Understanding the Mechanisms of Exposure and CBT for Anxiety Disorders

It is beyond the scope of this chapter to describe in detail all of the theories attempting to explain the mechanisms of exposure and CBT for anxiety disorders. However, two additional approaches should be noted given their prominence in the literature. First, the belief disconfirmation approach was developed first by Aaron T. Beck and colleagues. The main aspect of this model is that thoughts and beliefs lead to emotions and behaviors. This approach was initially called cognitive therapy, to contrast it with behavior therapy and exposure approaches. The belief disconfirmation model suggests that explicit discussions, contrived "experiments," and experiential exercises are used to dispel maladaptive beliefs the patient holds that are not realistic. Typically, such models do not attempt to explain implicit learning or concepts such as conditioning, habituation, or extinction. Ultimately, cognitive theories have become integrated with behavioral theories, including emotional processing theory, and their theoretical and practical distinctions are at most subtle. Indeed, in practice most researchers and clinicians today call their treatments CBT rather than pure behavioral or pure cognitive therapy.

A second approach which has been used recently by Michelle Craske and colleagues to contrast with emotional processing theory is the inhibitory learning approach. The basis of inhibitory learning is that exposure leads to fear reduction due to inhibition of fear associations. This model is based predominantly on modern, neoclassical learning theory. Accordingly, fear is thought to occur due to pairing of neutral stimuli with an aversive stimulus (classical conditioning). Fear is then "unlearned" through both habituation and extinction, leading to an inhibition of the association. The primary issue suggested here is that associations are not replaced, but only inhibited. Thus, they can be reactivated when the inhibitory processes weaken or fail, leading to relapse. There are competing models with the idea of inhibitory processes such as retrieval competition (suggesting that new and old memories both exist and compete to be expressed) and memory reconsolidation (suggesting that memory is constantly being reconstructed and not static). Some suggested implications of the inhibitory model are that it is essential to design exposures to maximize expectancy violation (common in many CBT models

today including cognitive therapy and emotional processing), remove safety signals, vary exposures, and change contexts of exposures. Most of these implications are not unique to inhibitory theory and are included in most current CBT models including those of belief disconfirmation and emotional processing as well as models that are not readily integrated into those models. Thus, many of the models proposed to explain the mechanisms of exposure and CBT for anxiety use different language to ultimately describe very similar processes. It is likely that the best explanatory model integrates all of these models, thereby incorporating their large overlap and adopting the subtle unique differences.

COGNITIVE-BEHAVIORAL THERAPY PROCEDURES FOR ANXIETY DISORDERS

Most CBT programs for anxiety disorders begin with an assessment, patient education, and specific treatment planning. The actual treatment procedures involve at least one of four components: (1) exposure to thoughts, objects, situations, and physiological sensations that are not dangerous but are nonetheless feared, avoided, or endured with great distress; (2) training in general anxiety or stress management techniques (including acceptance or mindfulness); (3) use of cognitive therapy techniques; and (4) training in specific skills, such as dating skills, assertiveness, and so forth.

Assessment, Psychoeducation, and Treatment Planning

CBT begins with a thorough assessment of the patient's presenting complaints, using empirically validated assessment procedures. Self-report and interviewer rating measures have been developed for the diagnosis and assessment of severity for each anxiety disorder. The assessment may also include measures of related psychopathology, such as depression and general anxiety, and specific theoretically relevant variables associated with the disorder of interest, such as worry-related metacognitions in GAD and anxiety sensitivity in panic disorder. Typically, assessments are repeated over the course and at the end of treatment for the purpose of establishing a diagnosis, selecting treatment procedures, monitoring the process of treatment, and evaluating treatment outcome.

The next step of CBT is psychoeducation. Patients are typically provided with a cognitive-behavioral formulation of their specific diagnosis and a related treatment rationale that serves as a guiding framework for future treatment sessions. In addition, alternative treatments shown to have similar efficacy are discussed to allow the patient to make an informed choice about treatment.

As part of the initial assessment and patient education, the therapist and patient work together to identify specific targets for treatment and to work out the details of the treatment plan. An example of this is the

development of a hierarchy for conducting in vivo exposure exercises or determining core beliefs and core fears of the patient in order to determine the proper exposures and their course. In addition, expectations are set that the patient will engage in homework between sessions to increase mastery of the therapy skills and self-monitoring procedures for symptoms.

Exposure Therapy

Exposure therapy involves intentionally confronting feared but otherwise not dangerous objects, situations, thoughts, memories, and physical sensations for the purpose of reducing fear reactions and avoidance associated with the same or similar stimuli. Systematic desensitization, where gradual imaginal and in vivo exposure was combined with relaxation, was the first exposure therapy technique to undergo scientific investigation. Although an effective treatment for some anxiety disorders, it has generally fallen out of use among researchers and cognitive-behavioral therapists due to the finding that concurrent relaxation was deemed unnecessary and possibly contraindicated. In addition, it does not seem necessary to progress systematically up the hierarchy (though it can be helpful at times). Finally, systematic desensitization did not appear helpful for OCD or agoraphobia while more intense in vivo exposures were helpful for these as well as specific phobias. The contemporary use of exposure therapy may be usefully divided into three classes of procedures: In vivo exposure, imaginal exposure, and interoceptive exposure.

In Vivo Exposure. In vivo exposure involves helping patients directly confront feared objects, activities, and situations. It is usually conducted in a graduated fashion according to a mutually agreed-on (between patient and therapist) hierarchy. For example, a hierarchy for a specific animal phobia, such as snakes or spiders, may begin with looking at pictures and other representations of the feared animal, followed by looking at the actual animal kept in a cage, first at a distance and then gradually moving closer. Depending on the animal, subsequent steps may involve touching and handling the animal, perhaps first while wearing a glove and then without the glove. These steps may be repeated across several different examples of the animal, differing in such dimensions as size and activity level.

In the case of SAD, individuals are asked to engage in conversations and other interpersonal situations. Recent evidence suggests that such exposures maximize the incorporation of corrective information by ensuring that patients with social anxiety make explicit, specific predictions prior to their exposures, focus their attention outward and drop safety behaviors during the exposure, and receive feedback from those they interacted with, and by carefully planned post exposure review of videotapes of the exposure. In the case of panic disorder, individuals

expose themselves to situations in which they fear having panic attacks such as riding in elevators, waiting in line at a store, and so on. GAD does not have standard in vivo exposure as part of its treatment because there is not usually a specific source of fears.

Imaginal Exposure. Imaginal exposure typically involves having the patient close their eyes and imagine feared stimuli as vividly as possible. One common form of imaginal exposure is imaginal flooding, in which the patient imagines the worst possible scenario repeatedly in order to help them imagine the full stream of occurrences that would need to happen in order for the feared outcomes to occur. Such imaginal exposure help the patient process the fear, which is often avoided both behaviorally (by avoiding contact with feared objects) and cognitively and emotionally (i.e., by avoiding thinking thoroughly about the feared consequences or avoiding how they would feel in the feared situations). Imaginal exposure has two general usages. The primary use is to help patients confront feared thoughts, images, and memories in order to process them. For example, individuals with social anxiety may experience extreme distress over thoughts of being rejected. Imaginal exposure to feared thoughts and consequences is used to promote reduction of emotional reactivity to the scenarios, to decrease the belief in the likelihood that the event will occur, to learn that the consequences are less severe than imagined, and to help the patient distinguish between thoughts and reality. In the case of phobias where there is a clear traumatic cause (e.g., fear of driving after a car accident), imaginal exposure is used to help the patient confront his or her memory of the traumatic event. Here again, the procedure promotes habituation to the memory and helps the patient to distinguish between the actual traumatic event, which was dangerous, and the current memory of the event, which, although distressing, is not harmful. With patients who have PTSD, imaginal exposure to traumatic memories is also used to help the patient get new perspective on the trauma and how they behaved during the trauma. For example, a patient who was raped and blamed herself for not fighting during the rape realizes during imaginal exposure that she actually fought vigorously during the rape; a soldier who blamed himself of not saving his friend during the battle realizes during imaginal exposure that he was too far away from his friend to be able to save him. In these cases, imaginal exposure also creates or organizes a coherent narrative, which may help the patient to cope with his or her fears.

A second use of imaginal exposure is in lieu of in vivo exposure when arranging direct contact with the feared situation is not safe or feasible, or as a preparatory exercise to facilitate subsequent in vivo exposures. In these cases, the patient imagines confronting the feared object or engaging in the feared task. When the patient's level of anticipatory anxiety is reduced, in vivo exposure may be conducted with the actual feared object

or task, assuming that it is both safe and available.

Interoceptive Exposure. Interoceptive exposure is the most recent form of exposure therapy to be introduced. This procedure is designed to induce feared physiological sensations under controlled circumstances. Interoceptive exposure exercises are most commonly used in the treatment of panic disorder and certain specific phobias, for example, the fear of vomiting, in which internal cues such as those associated with physiological arousal (e.g., rapid heart rate, dizziness, tingling in the fingertips) or gastric upset (e.g., after eating a spicy meal) elicit fear, anxiety, and further arousal.

A number of specific exercises have been developed by David Barlow and colleagues to induce specific panic-like sensations. For example, hyperventilating produces shortness of breath, dizziness, and tingling in the fingers and around the mouth and can produce a sense of unreality. Breathing through a thin straw produces the sensation of not getting enough air. Spinning in a chair or spinning in place produces dizziness and, in some people, mild nausea. Other possible activities include aerobic exercise to induce rapid heart rate and shortness of breath, consuming mild stimulants (e.g., eating chocolate-covered espresso beans), or eating spicy foods. The goal is to find an activity that produces similar sensations to the ones that patients find unduly distressing and attempt to avoid (e.g., the core panic symptoms in panic disorder). As with other forms of exposure, the purposes for inducing these sensations are to promote habituation of fear in response to them (i.e., the fear of fear or anxiety sensitivity) and to distinguish these sensations, which are normal responses to the various exercises, from physiological sensations that may actually signal a significant health problem. By intentionally inducing these feared sensations, the patients are able to learn to dissociate the sensations from danger per se by learning that these sensations do not cause harm and also gain a sense of control and mastery over their anxiety by accepting it. More recently, there have been suggestions and research supporting the use of interoceptive exposure as an exercise that assists patients to tolerate negative activating emotions (arousal), even outside the context of panic disorder. This is consistent with the notion that exposures (all types) in the context of acceptance of negative emotions (as opposed to attempts to suppress them) facilitate improvement during exposure.

Anxiety Management and Stress Inoculation Training

Anxiety management training is a general approach to treating anxiety, of which Meichenbaum's stress inoculation training is one specific example. Emotions in general, and stress and anxiety in particular, reflect the activity of three loosely coupled and interacting "channels," or modes, of responding: cognitive–phenomenological, physiological, and behavioral.

Stress inoculation training involves training patients in the application of specific skills and techniques designed to address each of these response modes in a flexible manner that is tailored to the individual's needs. The assumption underlying anxiety management training is that some people lack the skill of managing their anxiety and need to learn these skills. For example, spiraling negative cognitions that often occur during times of stress may be interrupted by the use of "thought stopping," in which the patient thinks the word "stop" in his or her mind while vividly imagining the word "stop" and then replaces it with positive or coping self-statements. The physiological manifestations of anxiety and stress (e.g., hyperventilation and its consequences, muscle tension, and gastric symptoms) may be addressed through training in diaphragmatic breathing and progressive muscle relaxation (e.g., Jacobson's "tense and relax" method). Changes in overt behavior are promoted through the use of role playing and covert rehearsal, in which patients practice new behaviors while executing other coping skills either with another person or in their mind. Exposure to feared situations either in vivo or in the imagination is frequently a part of stress inoculation training but is usually for the purpose of practicing new skills and is accompanied by the use of other coping skills. From the perspective of emotional processing theory, stress inoculation training brings about successful coping with fear and anxiety which provides new meaning information related to personal competence that counters typical fear-related perception of oneself as incompetent.

Cognitive Therapy

The basic assumption of cognitive therapy is that people's beliefs and appraisals of events and situations are at least as important in determining their emotional reactions as the actual events and situations. Problematic emotions such as anxiety are frequently the result of unrealistic and unhelpful beliefs about the world, the self, and others. The goal of cognitive therapy is to help patients to identify unhelpful cognitions and cognitive distortions (e.g., probability overestimation, catastrophic thinking, all-or-nothing thinking, overgeneralization, only considering evidence that is consistent with existing beliefs) and to modify them. Three of the most common traditional cognitive therapy procedures are (1) the use of Socratic dialogue, in which the therapist uses a series of questions to help the patient to identify and challenge unhelpful beliefs or to uncover evidence disconfirming fear-related beliefs; (2) the downward-arrow technique, in which the therapist helps to uncover deeper beliefs and meanings by repeatedly asking for greater clarification; and (3) the use of thought records, in which patients record automatic beliefs, list evidence for and against those beliefs, review potential cognitive errors that may be reflected in those beliefs, and then generate more realistic and helpful beliefs. All of these procedures can be conducted formally as traditionally was intended,

or in a more integrated fashion with exposures, as is more typically done in contemporary CBT. Indeed, behavioral experiments epitomize this integration. The goal of a behavioral experiment is to test beliefs about the nature and probability of feared consequences under conditions that will optimize exposure to, and the encoding of, disconfirming information. Accordingly, before conducting a behavioral experiment, the patient makes detailed predictions about what he or she believes will happen under a particular set of circumstances and then conducts the relevant experiment to see what happens. For example, consider a person with panic disorder who fears they will go crazy if they feel dizzy. A behavioral experiment to test predictions about the likelihood of going crazy might be to spin in a chair slowly at first, and then work up to spinning quickly, and seeing what the consequences of feeling intense dizziness are. Each time patient feels dizzy without going crazy (operationalized idiosyncratically for the given patient), the patient is confronted with evidence that disconfirms his or her beliefs about the probability of going crazy. The conceptual emphasis in cognitive therapy on changing beliefs through exposure to disconfirming evidence is consistent with the tenets of emotional processing theory. The practical and empirical issues involve determining the best ways to accomplish this goal.

Acceptance and Mindfulness-Based Training

Over the past decade, there has been an increase in interest in combining exposure and other forms of procedures to that aim at coping with anxiety via mindfulness, meditation, and a related concept frequently associated with mindfulness called “acceptance.” The interest is due to a number of CBT treatments that have their roots in Buddhism and have been integrated with humanism and CBT. Three of the most popular treatments are Acceptance and Commitment Therapy (ACT), Mindfulness-Based Stress Reduction (MBSR) and Mindfulness-Based Cognitive Therapy. All of these and other versions have in common the idea that fighting one’s anxiety (in the case of anxiety disorders) leads to a vicious cycle of attempting to suppress or otherwise force the downregulation of anxiety, which typically backfires. Therefore, acceptance-based techniques emphasize the need to first acknowledge that the anxiety exists and to try to objectively, nonjudgmentally observe it. Importantly, much of this concept was part of earlier versions of exposure therapies, but the aspect of radical acceptance was not emphasized. While some argue that acceptance-based techniques are in direct contrast with habituation models, anxiety reduction models, and cognitive therapy, the authors believe that all of these techniques likely facilitate one another, and that such techniques are easily integrated into exposure. Indeed, acceptance techniques and their integration with the other procedures may facilitate emotional processing.

Interpersonal Skills Training

Research has suggested that some individuals with SAD, as well as those with other anxiety disorders, have interpersonal deficits that contribute to the maintenance of their anxiety and potentially negatively affect treatment. Thus, some people with anxiety lack important skills needed to interact effectively with others. Consequently, they experience fewer interpersonal rewards and more punishments, leading them to avoid social interactions when possible, thereby further limiting their ability to acquire effective social skills and decreasing their social support. Accordingly, comprehensive treatment for anxiety provides explicit training to help patients acquire and use appropriate social skills and thereby reverse the negative cycle when necessary. These skills include assertiveness training and training in initiating, maintaining, and ending conversations. At times, it may be that training in such skills helps individuals have more confidence in their own skills and stop engaging in avoidance or safety behaviors. In addition, some research groups have suggested adding more interpersonal components to CBT to further help patients learn that goals of interpersonal interactions are not to pass a test, but rather to seek intimacy or simply to be in a relationship.

Efficacy of Cognitive-Behavior Therapy

Phobias and Agoraphobia. The key features of phobias are the presence of persistent and intense fear and avoidance of specific objects or situations that are excessive or unreasonable. If avoidance is not possible, the feared stimulus is endured with great distress. To qualify as a phobia, the fear and avoidance must significantly interfere with the person's social or occupational functioning or the person must experience marked distress about the fears. Although the range of phobic situations is potentially quite broad, the most common are animal phobias (e.g., dogs, insects), natural environment phobias (e.g., heights, storms, water), blood or injury phobias (including receiving or viewing injections), and situational phobias (e.g., bridges, flying, enclosed spaces). *Agoraphobia* is a pattern of pervasive avoidance of places and situations. If such places and situations cannot be avoided, then they are endured with great distress—the person with agoraphobia is afraid that he or she may experience a full-blown or symptom-limited panic attack and often perceives that help is unavailable. Most of the time, agoraphobia is associated with panic disorder; therefore, its treatment will be addressed later, along with that of panic disorder. Agoraphobia without panic typically involves fears of diarrhea or vomiting in public places due to anxiety. Common agoraphobic situations include being in various crowded public places (e.g., malls, restaurants, and theaters) and enclosed spaces (e.g., closets, elevators, and cars). Given the abundant evidence for the distinction of SAD from specific phobias, the

former is considered in a separate section.

A large body of research beginning in the mid-1960s has clearly demonstrated the efficacy of several exposure therapy protocols in reducing simple phobias and fear of public speaking. This research establishes exposure therapy as the treatment of choice for these disorders. The earliest studies focused on systematic desensitization. In the first controlled study, systematic desensitization was compared with a wait-list control condition in the treatment of students with a fear of snakes. In the first phase of the study, participants in the desensitization group separately practiced relaxation and created a hierarchy. In the second phase, participants in the desensitization group paired the feared images from the hierarchy with relaxation. Neither group showed any changes during the first phase, but the desensitization group showed significant improvement across the second phase, whereas the control group showed no change.

In a follow-up study, desensitization was found to be superior to a pseudotherapy that, like desensitization, included training in relaxation and hierarchy construction. After the initial preparation, participants in the desensitization group paired the feared images with relaxation, whereas the pseudotherapy group engaged in positive imagery during relaxation. Another study compared the efficacy of five sessions of systematic desensitization for the public speaking fears and other forms of social anxiety with wait-list and two alternative treatment conditions (insight-oriented psychotherapy and an attention-placebo condition in which subjects were given a placebo pill but told that it was a drug that reduced anxiety in stressful situations and then conducted a task that was said to be stressful but actually induced drowsiness). Results revealed that all three treatments were superior to wait-list and that systematic desensitization was superior to the alternative treatments, which did not differ from one another. This pattern of outcome was maintained for up to 2 years after the completion of treatment.

Research in the 1970s began to investigate the “flooding” approach to imaginal and in vivo procedures, which involved repeated and prolonged exposure to fear cues of high intensity without relaxation. Results from the earliest of these studies provided evidence for the efficacy of all three approaches (systematic desensitization, imaginal exposure, and in vivo exposure) but the results were mixed with regard to the relative efficacy of the treatments. One group found that both systematic desensitization and imaginal exposure were more effective than a nonspecific treatment control condition among a mixed sample of patients with specific phobias and patients with agoraphobia. However, they did not replicate the finding of superiority of imaginal exposure over desensitization for agoraphobia.

A more consistent pattern of outcome in favor of a particular form of in vivo exposure, called *participant modeling* or *guided mastery*, over other forms of exposure therapy in the treatment of phobias and agoraphobia has

been reported in series of studies by Albert Bandura and colleagues. In participant modeling, the therapist actively models handling the feared object for the patient and may use various "response induction aids" to help the patient to successfully handle the feared object. For example, in treating a patient with a spider phobia, the therapist may first demonstrate how to touch, handle, and control a tarantula. The therapist would then coach the patient in performing the same behaviors, first while wearing protective gloves and then without gloves. This active approach to in vivo exposure may be contrasted with more-passive approaches in which the patient may be instructed to sit and observe a snake in a cage and attempt contact with the object when their fear at the earlier step has decreased.

Systematic desensitization, prolonged imaginal exposure, and prolonged in vivo exposure have all been found to be effective in the treatment of phobias. In addition, systematic desensitization and imaginal exposure may be of some limited effectiveness in treating agoraphobia. Prolonged in vivo exposure is at least as effective as, and often more effective than, desensitization and prolonged imaginal exposure. The participant modeling variation of in vivo exposure has been found to be particularly effective in treating a range of phobias, including small-animal phobias, acrophobia, and agoraphobia. Some recent studies have suggested that use in vivo exposure plus a mildly distracting social interaction is more effective and more tolerable than full focus on the feared stimulus. Studies have also found virtual reality–based exposure therapy to be effective.

Panic Disorder. A panic attack is a sudden, intense rush of fear, anxiety, or impending doom that reaches a peak very quickly and is associated with at least 4 of 13 physical and cognitive symptoms (e.g., shortness of breath, dizziness, heart palpitations, fear of dying, fear of going crazy or losing control). Panic attacks may be caused by a specific situation, as when someone who is afraid of snakes encounters one, or they may be unexpected and perceived by the individual as coming out of the blue. For the diagnosis of panic disorder, a person must experience repeated unexpected panic attacks that result in persisting (at least 1 month) concerns about having additional attacks or worries about the physical or psychological consequences of the attack (e.g., having a heart attack, going crazy).

According to the cognitive-behavioral model of panic disorder, a consequence of experiencing unexpected panic attacks for some individuals is the development of a vicious cycle. The cycle begins with hypervigilance for somatic cues indicative of a panic attack that, if detected, are interpreted negatively (e.g., as an indicator of having a heart attack or going crazy). The negative interpretation of somatic sensations results in increased anxiety and arousal, spiraling into a panic attack. Agoraphobic avoidance may also develop as an attempt to avoid experiencing panic

attacks in situations in which help may not be available. Cognitive-behavioral treatments based on this model contain several components: (1) psychoeducation regarding the cognitive-behavioral model of panic disorder; (2) cognitive restructuring to challenge catastrophic cognitions about the likelihood and significance of panic attacks; (3) interoceptive exposure to feared somatic sensations to reduce the "fear of fear"; and (4) in vivo exposure to avoided cues, particularly in cases in which agoraphobic avoidance has developed. Oftentimes, later phases of treatment include combinations of phases 3 and 4 to help the patient learn to deal with physical sensations in the context of formerly avoided situations. Relaxation and breathing retraining were early components of CBT protocols, but research has suggested that they are at best unnecessary and possibly lead to greater dropout or interference with treatment gains.

The efficacy of CBT programs based on this model has been demonstrated in a series of well-controlled studies. Between 41 and 87 percent of patients receiving CBT were panic-free after treatment, compared to 36 percent receiving relaxation only, 50 percent receiving alprazolam, 20 percent receiving imipramine (Tofranil), 13 to 36 percent receiving placebo, and 30 to 33 percent undergoing wait-list. In the one study that combined imipramine with CBT, the combination treatment was found to be slightly more effective than CBT alone but not CBT combined with placebo. In this study, 6 months after treatment discontinuation, CBT alone and CBT plus placebo were superior to CBT plus imipramine, suggesting that medication may interfere with the maintenance of CBT gains after medication discontinuation. Several other research groups investigated variations of CBT, with similar results, replicating the superiority of CBT over relaxation, supportive counseling, and wait-list. More recently, another large, multicenter study examined CBT as a first line treatment for panic with agoraphobia. Individuals who did not respond to 10–12 sessions of CBT for panic and agoraphobia were randomized to either more sessions of CBT or to paroxetine. Individuals who received paroxetine responded more quickly to treatment than continued CBT, though individuals who received CBT had caught up by the 9-month follow-up. Individuals who responded to treatment were randomized to maintenance CBT or no maintenance treatment. Maintenance sessions decreased relapse significantly (5%) in comparison to those not randomized to maintenance treatment (18%).

Several studies attempted to identify which components add to treatment outcome. Two studies found that adding cognitive restructuring to in vivo exposure substantially improved outcome, although other researchers have not been able to replicate this. Another study compared CBT consisting of education, cognitive restructuring, and interoceptive exposure with the same program plus relaxation and found that adding relaxation did not improve outcome. A 1997 study compared two

treatments that incorporated cognitive restructuring and in vivo exposure. In addition, one included interoceptive exposure, and the other group substituted breathing retraining. Both treatments were associated with improvement, but the group receiving interoceptive exposure had superior outcome on panic frequency at posttreatment and follow-up assessments. In a study using a crossover design, patients were assigned to receive four sessions of cognitive therapy followed by four sessions of interoceptive exposure or the reverse. Both treatments were associated with improvement, and more improvement was observed for the first intervention administered, but there was no difference in the efficacy of the two components. An additional study examined compared cognitive therapy without interoceptive exposure to interoceptive exposure without cognitive therapy and found similar positive outcomes in both groups that were equivalent to other studies examining combined packages. Another study compared education, cognitive restructuring, and in vivo exposure with or without the inclusion of breathing retraining and found no reduction in treatment efficacy. Three variations of self-conducted exposure therapy—in vivo exposure, interoceptive exposure, and the combination of in vivo plus interoceptive exposure—were compared with a wait-list control in another study. All three treatments were superior to the wait-list condition, but there were no significant differences among treatment conditions.

Treatment for panic disorder has largely consisted of multicomponent packages consisting of education about the cognitive-behavioral model of panic along with one or more of the following: diaphragmatic breathing, relaxation, cognitive therapy, in vivo exposure, and interoceptive exposure. Dismantling studies have been partially successful in identifying the active ingredients, which seem to be cognitive restructuring combined with in vivo or interoceptive exposure. Indeed, the addition of interoceptive exposure may be particularly helpful. In a meta-analysis of 43 controlled studies of treatment for panic disorder, the largest mean effect size was obtained for the combination of cognitive therapy plus interoceptive exposure (0.88), followed by CBT without interoceptive exposure (0.68), CBT plus medication (0.56), and medication alone (0.47). Moreover, dropout rates from CBT were lower than dropout rates from medication conditions (alone or in combination with therapy). Another review by the Cochrane Collaboration found 21 trials comparing CBT to medications and their combination. Results suggested that combination treatment is superior to medication alone. Combination treatment also had superior symptom reduction to CBT alone in the acute phase, but also had more dropouts. After treatment discontinuation, no difference between CBT and combined treatment was found. Some trials showed that CBT for panic can also be helpful for treating medication-resistant patients, as well as for helping patients discontinue medications, including benzodiazepines and

SSRIs. Recent trials also showed the ability to disseminate treatment for panic in community mental health clinics in addition to primary care. Other studies showed that CBT is effective for nocturnal panic, and that focused CBT on panic also significantly reduces comorbid conditions. In addition, a number of trials suggested the efficacy of group CBT treatment for panic disorder. Finally, some studies showed promising results for early intervention with CBT in individuals at risk for developing panic disorder.

Social Anxiety Disorder

Individuals with SAD experience marked and persistent fear and avoidance of one or more social or performance situations. The person may fear doing or saying something embarrassing or fear manifesting the signs of anxiety (e.g., blushing, sweating, trembling) in such situations. Thus, SAD can be conceptualized as a fear of embarrassment or fear of social rejection. Most individuals with SADs fear and avoid multiple situations, falling within the generalized subtype of SAD.

Treatments for social phobia have generally consisted of exposure therapy, cognitive therapy, the combination of exposure plus cognitive therapy, and social skills training. A network meta-analysis summarized the outcome of 101 studies of 41 different treatments or control conditions for social anxiety. Most active treatments were superior to wait-list. However, individual CBT had the largest effect size (greater than group CBT and other treatments), and was significantly more effective than psychodynamic therapy, supportive therapy, or mindfulness. The results of individual dismantling studies yield an ambiguous picture. In the first such study, exposure alone was compared with exposure plus cognitive therapy. Although both groups improved, outcome was somewhat better for the combined treatment, particularly at the 3-month follow-up assessment. In a second study by the same group, the efficacy of each component alone, as well as the combination of the two, was compared with a wait-list control group. Compared to the wait-list control group, all three treatments were found to be effective in reducing social anxiety. Again, however, outcome was best among subjects who received the combined treatment. By contrast, another study did not find evidence for superiority of combined treatment over exposure alone. On some measures, patients who received exposure only had numerically better outcome than those who received the combined program.

One treatment program that has received considerable research attention is Richard Heimberg's cognitive-behavioral group therapy (CBGT), consisting of in-session exposure to feared social situations and cognitive restructuring and corresponding homework. The group format not only serves to make treatment more affordable, it is believed that group treatment provides a great resource for conducting exposures, obtaining feedback, and learning vicariously. One study found CBGT to be superior to

wait-list, and three studies found CBGT to be superior to a group attention control condition, in which patients received education about social anxiety and nondirective group therapy. Heimberg and colleagues also compared CBGT with phenelzine (Nardil) and placebo. Although there were no differences between CBGT and phenelzine immediately after treatment, there were more instances of relapse for patients receiving phenelzine during the 6-month follow-up. Another study also compared CBGT with the same treatment components administered individually and found the two methods to be comparable. However, more recent studies have suggested that individual treatment is superior to group treatment. One of the largest randomized trials found that group CBT including social skills training yielded similar outcomes to fluoxetine and the combination of fluoxetine plus CBT, and all of these treatments were superior to placebo.

The most promising results have come from a version of CBT developed by David M. Clark and colleagues. A number of randomized studies have been published, showing large effects compared to fluoxetine, exposure alone, psychodynamic therapy, and group CBT (although not CBGT). Recent data support the efficacy of this treatment in a fee-for-service clinic. Some differences between the newer program and other CBT for social anxiety include a stronger emphasis on identifying and eliminating safety behaviors and shifting attention outward from the self. In addition, exposures are integrated carefully with cognitive challenging in the context of making specific predictions and evaluating these predictions post hoc with feedback from multiple sources, including videotape and those with whom the patient interacted during the exposure. Much of this is consistent with modern theories of CBT discussed above. One of the latest advances in this treatment is an online, internet-based administration of much of the program, using videos, chats, and other interactive activities. Results seem to suggest excellent outcomes with less time spent by the therapist (most of which is online), though considerable patient investment in time on their own.

Generalized Anxiety Disorder

The key feature of GAD is the presence of excessive worry accompanied by anxiety and physiological symptoms of tension, irritability, and sleep disturbance. The content of the worry must encompass a range of domains (e.g., school or work performance, personal finances) and not be limited to features of another specific disorder, such as worrying about having a panic attack (as in panic disorder). The worry, anxiety, and arousal symptoms must be present more days than not over a minimum period of 6 months. Pathological worry, as occurs in GAD, is distinguished from normal worry, in that the frequency, intensity, and duration of the worry in addition to the attendant anxiety greatly exceed the actual probability or impact of the feared event and that the worry is experienced as difficult to control.

CBT packages for GAD have typically included training in relaxation and cognitive therapy. In addition, some packages have included some form of exposure therapy, exposing the patient to the worry content, and problem solving. Relaxation alone, cognitive therapy alone, and their combination (with and without the further addition of exposure) have all been found to be effective in reducing worry and anxiety compared to wait-list. In addition, CBT has been shown to be more effective than supportive control conditions. A recent meta-analysis found that CBT and relaxation did not yield different results for GAD. A number of meta-analyses and reviews have found CBT to be superior to wait-list and placebo conditions. One meta-analysis found that CBT effects were somewhat smaller than those for medications. The general results of these trials have been positive, but there is a general consensus that improvements in treatment efficacy are necessary, perhaps more than in CBT for the other anxiety disorders. A number of attempts have been made to improve treatment outcomes for GAD. One study added more treatment sessions for individuals predicted to have poorer outcome, with nonsignificant improvement in outcome compared to standard treatment. Others attempted to modify or add to existing models of CBT. Refinement of the CBT model includes a greater focus on meta-worry (worry about worry) or a greater focus on the intolerance of uncertainty and the separation of worry into worry about current problems (addressed by problem solving) and worry about future unknown events (addressed by increasing skills in tolerating uncertainty, including imaginal exposure). Additional areas and techniques to integrate with current CBT models include interpersonal treatment strategies, emotion regulation strategies, mindfulness and experiential avoidance strategies, additional focus on achieving positive well-being, and schema-focused treatments. Most data to date have not shown that these techniques provide improved outcome compared to standard CBT packages.

Future Directions

The foregoing literature review demonstrates the efficacy of CBT programs in the treatment of all anxiety disorders. These treatments appear to be effective in children, adolescents, adults, and geriatric populations, and are typically more effective than treatment as usual in the short and long term. Exposure in vivo to feared but otherwise safe stimuli and imaginal exposure to unwanted and upsetting thoughts are essential components of most CBT programs for anxiety. The cessation of safety behaviors or other subtle avoidance behaviors is essential for anxiety disorders in order to disconfirm patients' erroneous perceptions. The addition of interoceptive exposure appears to make a significant contribution to the treatment of panic disorder. In addition, exposure is typically conducted in the presence of cognitive challenging or discussions, helping the patient to attend to the

corrective information. Anxiety management training and cognitive therapy have also been shown to be effective in the treatment of GAD. Cognitive therapy that includes behavioral experiments, which inherently involves in vivo exposure, is also an effective treatment for panic disorder and SAD.

Several researchers have demonstrated the efficacy of multicomponent CBT packages that include components of anxiety management, cognitive therapy, and exposure. However, few of the relevant dismantling studies showed significantly improved outcome for these combined treatments as compared to the individual components. This may, to some extent, reflect low statistical power to detect the additive effects of treatments that individually are effective. Consistent with this hypothesis, meta-analyses have sometimes detected evidence for the superiority of combined treatments (typically of cognitive therapy plus exposure) that were not as evident in individual studies. Overall, it seems that the integration of exposure therapy with discussions that increase attention of patients to the disconfirmation of their expectations of negative outcomes to exposures to the feared stimuli is the best method for optimizing treatment outcome. Perhaps the addition of cognitive therapy to exposure enhances outcome because it serves to enhance attention to disconfirming information. It is important to note that, as was hypothesized by emotional processing theory, exposure should match the fear of the patient. It follows that exposures that maximize opportunities for correction of erroneous associations by matching the patient's fears carefully and providing corrective information via these experiences is likely to lead to the best outcomes. The need for relaxation or similar coping or management strategies seems questionable.

Regarding the execution of these specific techniques, it is important to note that some studies suggested that outcomes can differ across therapists using the same techniques, and some studies showed that expert sites are more effective in delivering the same techniques than non-expert sites (although this literature also includes data to suggest that some non-expert sites can obtain similar effects for some treatments). When therapist or site differences occur, it is due to two aspects of treatment delivery that likely interact. First, execution of these techniques typically requires experience and competence in formulating and articulating the treatment rationale (i.e., creating a flexible hierarchy that targets core fears, determining the proper focus of imaginal exposures, etc.), addressing the many overt and covert avoidance behaviors effectively, and knowing how to push exposures to their optimal level. In addition, although formation of the therapeutic alliance, warmth, empathy, and other nonspecific therapeutic skills are not the central focus of CBT, they are still important components that likely interact with the focused treatment techniques, facilitating compliance and thus better outcome. However, having nonspecific characteristics without

the expertise in technique is insufficient for provision of optimal treatment. Therefore, when experts are not available, it is essential to monitor patient progress carefully (also true when experts are available), and if there is not progress after ten sessions of treatment, referral to an expert out of one's locale may be warranted. Alternatively, there are increasing data to suggest that therapist-supported internet-based programs can be an effective treatment, which yields promising results that often are similar to face-to-face treatments.

Despite the success of CBT in the treatment of anxiety disorders, it is important to acknowledge two limitations. First, not every patient is responsive to treatment, and many people judged to have responded to treatment continue to experience significant residual symptoms. Thus, one avenue for future research is to identify ways to enhance existing treatments or to develop new treatments that are more effective. As noted, however, attempts to augment one psychological component with another treatment have not generally been very successful in significantly enhancing treatment outcome. Some of the recent improvements in outcome appear to be due to techniques that are developed based on specific maintenance mechanisms identified in studies on experimental psychopathology. Hopefully, continued study and improvement in understanding the psychological mechanisms involved in each of the anxiety disorders will help to improve outcome.

Another approach to potentially improving outcome is to combine psychotherapy with standard types of pharmacotherapy. This section did not provide coverage of the many studies that have investigated the efficacy of combining medication and CBT. The data tend to show that combining medication for anxiety disorders such as SRIs with CBT either provide a small advantage or do not enhance outcome beyond those achieved by CBT alone, whereas adding CBT to medication enhances the outcome of medication. Most studies of combined treatments, whether combining two forms of CBT or combining CBT with medication, implemented both treatments simultaneously. This strategy may actually serve to minimize the ability to detect the effects of combined treatments when the individual treatments are generally effective. An alternative strategy would be to use an augmentation strategy, in which partial responders to one active treatment are then randomized to continue with the original treatment alone or to add a second treatment, as was done in the panic study described above.

The second limitation of CBT is its relative inaccessibility to people who do not live in large cities or near a university-based medical school or graduate program in clinical psychology. Greater research is needed to determine the factors that limit the use of CBT among community-based clinicians and the development of models for the dissemination of these treatments to make them more widely available. Some advancements in

this area include using self-help and stepped care approaches to treatment, and using internet-based interventions.

References

Antony MM, Stein MB. *Oxford Handbook of Anxiety and Related Disorders*. New York: Oxford University Press; 2008.

Bandelow B, Reitt M, Röver C, Michaelis S, Görlich Y, Wedekind D. Efficacy of treatments for anxiety disorders: a meta-analysis. *Int Clin Psychopharmacol*. 2015;30(4):183–192.

*Bandura A. *Principles of Behavior Modification*. New York: Holt, Rinehart & Winston; 1969.

Barlow DH. *Anxiety and Its Disorders: The Nature and Treatment of Anxiety and Panic*. 2nd ed. New York: Guilford Press; 2002.

Barlow DH. *Clinical Handbook of Psychological Disorders*. 5th ed. New York: Guilford Press; 2014.

*Beck AT. *Cognitive Therapy and the Emotional Disorders*. New York: International University Press; 1976.

Chambless DL, Baker MJ, Baucom DH, Beutler LE, Calhoun KS. Update on empirically validated therapies II. *Clin Psychol*. 1998;51:3.

Clark DA, Beck AT. *Cognitive Therapy of Anxiety Disorders: Science and Practice*. New York: Guilford Press; 2011.

Clark DM, Ehlers A, McManus F, Hackmann A, Fennell M. Cognitive therapy versus fluoxetine in generalized social phobia: a randomized placebo-controlled trial. *J Consult Clin Psychol*. 2003;72:1058.

Craske MG, Treanor M, Conway CC, Zbozinek T, Vervliet B. Maximizing exposure therapy: an inhibitory learning approach. *Beh Res Ther*. 2014;58:10–23.

Cuijpers P, Sijbrandij M, Koole SL, Andersson G, Beekman AT, Reynolds CF. Adding psychotherapy to antidepressant medication in depression and anxiety disorders: a meta-analysis. *World Psychia*. 2014;13(1):56–67.

Cuijpers P, Sijbrandij M, Koole S, Huibers M, Berking M, Andersson G. Psychological treatment of generalized anxiety disorder: a meta-analysis. *Clin Psychol Rev*. 2104;34(2):130–140.

Davidson JRT, Foa EB, Huppert JD, Keefe FJ, Franklin ME. Fluoxetine, comprehensive cognitive behavioral therapy (CCBT) and placebo in generalized social phobia. *Arch Gen Psychiatry*. 2004;61:1005.

Foa EB, Huppert JD, Cahill SP. Update on emotional processing theory. In: Rothbaum BO, ed. *The Nature and Treatment of Pathological Anxiety*. New York: Guilford Press; 2006:3.

*Foa EB, Kozak MJ. Emotional processing of fear: exposure to corrective information. *Psychol Bull*. 1986;99:20.

Heimberg RG, Becker RE. *Cognitive–Behavioral Group Therapy for Social Phobia: Basic Mechanisms and Clinical Strategies*. New York: Guilford Press; 2002.

Heimberg RG, Turk CL, Mennin DS. *Generalized Anxiety Disorder: Advances in Research and Practice*. New York: Guilford Press; 2004.

*Hofmann SG, Smits JA. Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. *J Clin Psychiatry*. 2008;69:621.

Lang PJ. A bio-informational theory of emotional imagery. *Psychophysiology*. 1979;16:495.

Lang PJ, Lazovik AD. Experimental desensitization of a phobia. *J Abnorm Soc Psychol*. 1963;66:519.

Mayo-Wilson E, Dias S, Mavranezouli I, et al. Psychological and pharmacological interventions for social anxiety disorder in adults: a systematic review and network meta-analysis. *Lancet Psychiatry*. 2014;1(5):368–376.

McGlynn FD, Mealiea WL, Landau DL. The current status of systematic desensitization. *Clin Psychol Rev*. 1981;1:149.

Meichenbaum D. *Cognitive Behavior Modification*. Morristown, NJ: General Learning Press; 1974.

Mitte K. A meta-analysis of the efficacy of psycho- and pharmacotherapy in panic disorder with and without agoraphobia. *J Affect Disord*. 2005;88:27.

Podinaˇ IR, Koster EH, Philippot P, Dethier V, David DO. Optimal attentional focus during exposure in specific phobia: a meta-analysis. *Clin psychol rev*. 2013;33(8):1172–1183.

Watts E, Turnell A, Kladnitski N, Newby JM, Andrews G. Treatment-as-usual (TAU) is anything but usual: a meta-analysis of CBT versus TAU for anxiety and depression. *J Affective Disord*. 2015;175:152–167.

Wolitzky-Tylor KB, Horowitz JD, Powers MB, Telch MJ. Psychological approaches in the treatment of specific phobias: a meta-analysis. *Clin Psychol Rev*. 2008;28:1021.

*Wolpe J. *Psychotherapy by Reciprocal Inhibition*. Stanford, CA: Stanford University Press; 1958.

▲ 14.8 Somatic Therapies for Anxiety Disorders

LAKSHMI RAVINDRAN, M.D., AND MURRAY B. STEIN, M.D.

INTRODUCTION

Anxiety disorders are among the most widely diagnosed psychological conditions, with a lifetime prevalence of over 20 percent. They are often perceived as less disabling than depressive disorders, but patients with anxiety conditions also experience significant personal and societal burden, and a substantial proportion experience a chronic, fluctuating, long-term course of illness.

It must be noted that some patients with anxiety symptoms may not require treatment. For a proportion of anxious patients, symptoms can be transient and subsyndromal, and may result from stressful life events. There have been concerns that “medicalization” of anxiety symptoms such as normal worries and shyness may lead to inappropriate pharmacological treatments. However, studies repeatedly confirm that the occurrence of “overdiagnosis” is relatively low and that anxiety disorders remain under-recognized and undertreated. This has been attributed to factors including lack of awareness by patients of symptom significance and poor confidence by clinicians in their knowledge base for managing these conditions. In

addition, only a minority of patients seek help, and those that do tend to present with physical rather than psychological symptoms. Reluctance of patients to report emotional symptoms and of physicians to raise psychological issues are additional obstacles to diagnosis. Perhaps unsurprisingly, detection rates of anxiety and depression in primary care is less than 50 percent—the percentage of patients receiving appropriate evidence-based care for anxiety disorders in primary care has been estimated at 20 to 27 percent—but these rates appear to increase substantially upon repeated presentation of patients and associated expansion in reported symptoms and symptom severity. It has been suggested that recognition of anxiety disorders, and appropriate care, can be enhanced through systematic screening using questionnaires. Education of physicians, in particular those in primary care, about subgroups that may be more vulnerable to anxiety disorders, such as those with significant psychosocial difficulties and chronic physical illness, would enhance early recognition and intervention.

Table 14.8–1.
Principles of Management of Anxiety Disorders

- Careful diagnostic assessment with attention to medical and psychiatric comorbidities, contribution from life stressors, and genetic vulnerability.
- Appropriate physical examination and laboratory assessments, as indicated.
- Psychoeducation with a focus on adherence and self-management techniques.
- Systematic monitoring of treatment progress using structured instruments and scales.

Effective medication treatment of anxiety disorders is founded on observance of standardized principles of pharmacological management (see Table 14.8–1). However, patients' uncertainty about the value of pharmacotherapy and concerns about treatment costs can contribute significantly to nonadherence. Compliance with treatment can also be encouraged by dispelling the perception that antianxiety agents merely treat symptoms and not the condition. Psychoeducation for patients on the side effects of psychotropic treatment is also particularly relevant, as patients suffering from anxiety disorders are prone to sensitivity to such adverse events.

PRINCIPLES OF MANAGEMENT

As in the case with any psychological condition, the management of anxiety disorders should follow recognized principles (see Table 14.8–1) previously described. Of particular importance is the assessment of comorbid medical and psychiatric conditions, prescribed and over-the-counter medications, and excessive use of caffeine as contributing factors.

In general, the appropriate management of anxiety disorders initially involves pharmacotherapy, empirically supported psychotherapy (such as

cognitive–behavioral therapy [CBT]), or a combination thereof. The specific choices are informed by several key factors (see Table 14.8–2). Ideally, decisions about which treatment to use will be based on a good understanding of the patient's personal characteristics and personal and family history of illness, in conjunction with a strong awareness of the evidence base for the use of particular therapies for specific anxiety disorders. Both medication and psychological therapies have similar efficacy on their own in acute treatment but their comparative efficacy over long-term has not been well evaluated. Although it would seem intuitive that the combination would be superior to either treatment, this too has been poorly evaluated.

PHARMACOLOGICAL OPTIONS FOR TREATING ANXIETY DISORDERS

Several classes of agents have been evaluated in the treatment of anxiety and related disorders. These include antidepressants, benzodiazepines, anticonvulsants, and antipsychotics, among others. There are significant differences in efficacy and tolerability between classes as well as among agents within the same class. There is also significant variability in the symptoms they alleviate and the time points in the course of illness at which they produce optimal benefit.

Table 14.8–2.
Patient-Related Factors Influencing Choice of Treatment

- Severity and chronicity of illness
- Presence of medical and/or psychiatric comorbidity
- Personal or family history of response to a proposed intervention
- Concomitant medications and possibility of drug interactions
- Situation-specific versus generalized nature of symptoms (e.g., performance anxiety–related panic attacks)
- Patient preference
- Insight and likelihood of adherence
- Psychological-mindedness and motivation to change thoughts and behaviors
- Accessibility and cost of treatments

Antidepressants

Antidepressants include SSRIs, SNRIs, tricyclic antidepressants (TCAs), monoamine oxidase inhibitors (MAOIs), and reversible inhibitors of monoamine oxidase A (RIMA). Among these, SSRIs and SNRIs are often the first line agents prescribed.

Selective Serotonin Reuptake Inhibitors (SSRIs) and Serotonin Norepinephrine Reuptake Inhibitors (SNRIs). SSRIs are proven to be effective in a broad spectrum of anxiety and related disorders. They include fluoxetine (Prozac), sertraline (Zoloft), paroxetine (Paxil), fluvoxamine

(Luvox), citalopram (Celexa), and most recently, escitalopram (Lexapro or Cipralex). There are several SNRIs available, including venlafaxine (Effexor), desvenlafaxine (Pristiq), and duloxetine (Cymbalta). Most guidelines recommend SSRIs and venlafaxine as the initial antidepressant treatment for anxiety disorders. In practice, it is customary to start with lower initial doses to minimize an initial anxiogenic effect (which is almost always short lived) and a slow titration to an optimum dose. Although an initial response may be seen as early as 3 to 4 weeks, a trial up to 8 to 12 weeks at a therapeutic dosage is necessary to assess efficacy. The aim of treatment should be to achieve remission and not to settle for partial response, as partially treated patients are at greater risk for relapse and persistent psychosocial dysfunction.

The significant advantages for SSRIs over the older TCAs and MAOIs are their tolerability and safety. Although there are some differences between individual SSRIs, they are generally well-tolerated. Potential troublesome adverse effects include initial nervousness, insomnia, drowsiness, lightheadedness, nausea, and diarrhea; many of which improve with continued use. Treatment emergent sexual dysfunction across various domains (desire, arousal, erectile, ejaculatory, and orgasmic dysfunction are also commonly reported side effects of SSRIs, though their frequency varies to some extent within the SSRI class (slightly less with SNRIs) and to a significant extent between antidepressant classes (see Table 14.8–3). Since there is usually little or no tolerance for SSRI-induced sexual dysfunction, significant dysfunction often necessitates the switch to another class of agents such as agomelatine, bupropion, mirtazapine or moclobemide, or adjunctive sildenafil (Viagra) (for anti-depressant induced erectile dysfunction). However, adjunctive strategies provide limited benefit; dose reduction may be an alternative. Serious but uncommon adverse effects associated with SSRIs include serotonin syndrome and neuroleptic malignant syndrome. Hyponatremia, agranulocytosis, osteoporosis, and fractures in the elderly have been reported in a small subgroup of patients. Increased risk of GI bleeding specifically when prescribed with NSAIDs is another well documented serious adverse event.

Table 14.8–3.
Frequency of Sexual Dysfunction among Most Commonly Prescribed Agents for Anxiety Disorders

		Frequency of Sexual Dysfunction		
Class	**Drug**	**<10%**	**10–30%**	**>30%**
SSRI	Citalopram		X	
	Escitalopram		X	
	Fluoxetine			X
	Fluvoxamine			X
	Paroxetine			X

	Sertraline			X
SNRI	Duloxetine		X	
	Milnacipran		X	
	Reboxetine	X		
	Venlafaxine		X	
	Desvenlafaxine	X		
RIMA	Moclobemide	X		
Other Psychotropics	Agomelatine	X		
	Buspirone	X		
	Mirtazapine	X		
	Bupropion	X		
Atypical Antipsychotics	Quetiapine		X	
Anticonvulsant	Pregabalin		X	

With a similar side effect profile to SSRIs, venlafaxine extended release is generally well tolerated, although some patients additionally complain of dry mouth, increased sweating, and constipation. At the upper end of the therapeutic range (e.g., 225 mg per day or higher), routine blood pressure monitoring is also warranted because some patients (probably fewer than 5 percent) experience dose-related hypertension. The newer SNRI, duloxetine, has a similar efficacy of tolerability profile to venlafaxine, including discontinuation syndrome, though it may not be appropriate for use with individuals at risk of or with known liver dysfunction. Most recently, desvenlafaxine, the major metabolite of venlafaxine extended release, was released as a standalone compound. It has primarily been investigated as a treatment for major depression although often used off label for other conditions including anxiety disorders. It is thought to offer the advantage of dual uptake inhibition even at lower doses, unlike with venlafaxine where it occurs only in the upper dose range. To date, reports of tolerability relative to its parent compound are quite favorable.

Both with SSRIs and SNRIs, abrupt stopping can lead to discontinuation syndrome. Commonly reported symptoms of this condition, which tend to occur 2 to 4 days after medication cessation, include increased anxiety, irritability, tearfulness, dizziness or lightheadedness, malaise, sleep disturbance, and concentration difficulties. Although discontinuation symptoms can occur with all of the SSRIs and SNRIs, the most commonly noted are in agents with shorter half-lives (e.g., paroxetine and venlafaxine). Patients should be advised to discontinue SSRIs or SNRIs only under the supervision of their physician, with a gradual taper.

Ongoing controversy concerning antidepressant use and increased suicidality worse in the initial few weeks of treatment prompted the FDA to issue a black box warning on all antidepressant medication, cautioning their use in children and young adults. This led to more systematic evaluations of the risk of such occurrence, which found no clear evidence to

suggest SSRIs and other newer antidepressants are associated with increased risk of suicide in the adult population. In fact, they may confer a degree of protection in the adult group. With child and young adult samples, independent and more recent meta-analyses show a 1.5- to 2-fold risk of increasing suicidal thoughts and behaviors with newer antidepressants, but no completed suicide in the database, with only a very small increased absolute risk. Specifically, such risk may be higher with venlafaxine and lowest with fluoxetine and citalopram. Given the low absolute risk, and the significant impairment caused in children and youth by anxiety disorders, consideration of the use of pharmacotherapy in these populations is warranted.

Tricyclic and Heterocyclic Antidepressants (TCAs and HCAs). Until the advent of the SSRIs in the mid-1980s, TCAs and heterocyclic compounds (e.g., doxepin), along with benzodiazepines, were the mainstay of treatment for most anxiety disorders. TCAs and HCAs work by inhibiting the reuptake of several monoamines, including norepinephrine and serotonin, in the synaptic cleft. They also block cholinergic, histaminic, and alpha-1 adrenergic receptors, which contributes to their burdensome side effect profile. These include hypotension, dry mouth, blurred vision, and constipation, among others. Subgroups of patients often report an initial "jitteriness" when treatment is initiated with TCAs, which can often be minimized by beginning with very low doses and then titrating upward very slowly to a therapeutic dose. However, their effects on seizure threshold and potentially harmful cardiac effects, particularly in overdose, have further caused these agents to fall out of favor, despite equivalent efficacy to newer SSRIs and SNRIs. They are also associated with increased suicidality and should be avoided in at-risk patients. In addition, drug–drug interactions and discontinuation syndrome are still significant issues with this class of antidepressants.

Monoamine Oxidase Inhibitors (MAOIs). MAOIs are effective antidepressants whose use has been superseded by equally effective but significantly safer novel agents, relegating them to second-line use. The two most commonly used MAOIs in North America are phenelzine (Nardil) and tranylcypromine (Parnate). The benefit of phenelzine is well documented in panic disorder (PD) and SAD. MAOIs work by irreversibly inactivating both isoforms of the enzyme monoamine oxidase (MAO-A and MAO-B), leading to a relative increase in monoamines serotonin, norepinephrine, and dopamine in the synaptic cleft. The irreversibility of the inhibition requires patients to have a minimum 2-week "washout" period when switching from an MAOI to any other antidepressant class or even from one MAOI to another. The enzymatic inhibition is generic for monoamines and reduces the ability of the body to break down tyramine-rich foods, leading to the

risk of a hypertensive crisis. Though this reaction is rare, patients on MAOIs therefore need to follow a low-tyramine diet during treatment and for an additional 2 weeks after discontinuation. Other common side effects include postural hypotension, insomnia, and weight gain. Similar to TCAs, overdose is a potentially fatal event.

Reversible Inhibitor of Monoamine Oxidase Type A. Reversible inhibitors of monoamine oxidase type A (RIMAs) are unavailable in the United States, although one, moclobemide (Manerix), is available in Canada and several other countries globally. Randomized controlled trials (RCTs) have shown that moclobemide is effective in reducing symptoms of depression and several anxiety disorders, but anecdotal reports suggest that it may not be as efficacious as the older generation of MAOIs. Clinicians have also noted the gradual development of tolerance in a proportion of patients, leading to the need for incremental increase in dose. The main advantage for the use of RIMAs over MAOIs is the lack of dietary requirements, but caution is still advised with dietary restrictions recommended for higher doses. When side effects do occur, they consist mainly of sleep problems, dizziness, and nervousness; most are transient and go away with time. In generalized SAD in particular, RIMAs, if available, represent an alternative to SSRIs or SNRIs, especially for patients who have experienced problematic side effects with those agents.

Novel Antidepressants. Mirtazapine (Remeron) is a noradrenergic and specific serotonergic antidepressant (NaSSA). It has not been extensively studied in anxiety disorders but is claimed to be a useful agent for treating elderly patients with anxiety symptoms. Relatively free of sexual side effects with a propensity to increase appetite and weight gain, it is a useful agent in subgroups of patients. Its favorable effect on sleep often makes it a useful adjunctive agent.

Bupropion (Wellbutrin) is a dual norepinephrine dopamine reuptake inhibitor which has little benefit as the primary treatment of anxiety disorders. However, it may be used for adjunctive purposes to address residual symptoms.

Vortioxetine (Trintellix) is a new antidepressant in the North American market with a multimodal mechanism of action. It acts as an agonist and partial agonist at 5HT1A and 5HT1B receptors, respectively, but as an antagonist at 5HT3, 5HT7, and 5HT1B receptors. To date, the limited clinical trials of its use in anxiety have been primarily focused on GAD, where results were mixed. However, anxiety symptom reduction in major depressive disorder (MDD) patients would suggest benefit, as would its neurochemical profile.

Agomelatine (Valdoxan) is currently not available in North America. Its benefit is thought to be mediated through agonism of melatonin receptors

MT1 and MT2 and antagonism of 5HT2C receptors. The evidence for its benefit in anxiety disorders is still preliminary but with good evidence of efficacy in the acute treatment and prophylaxis of GAD. It has the advantage of low likelihood of sexual dysfunction, weight gain or discontinuation symptoms. Mild sedation can be an advantage for some patients. Regular monitoring of liver function is need in the early period of treatment.

Vilazodone (Viibryd), currently only available in the US, works by combining selective serotonin reuptake inhibition with partial agonism of the 5HT1A receptor. Currently, there are no published trials of its use in anxiety disorders, although a number of studies are currently underway.

Finally, milnacipran (Ixel) is an SNRI available in several countries worldwide but is not currently approved in North America. Two small open trials have suggested potential utility in GAD and PD, but controlled trials have yet to confirm this.

Benzodiazepines (BZDs)

For many years, BZDs were a mainstay of treatment for the anxiety disorders. As a class, they enhance the effects of GABA, the main inhibitory neurotransmitter in the brain. In addition to anxiolytic effects, they also possess sedative, anticonvulsant, and muscle relaxant properties. Most BZDs (e.g., clonazepam) are metabolized by the liver, primarily through CYP-mediated oxidation, but others, e.g., lorazepam, oxazepam, and temazepam, are metabolized via a conjugation process, and thus preferred choice for individuals with impaired liver function. The most common adverse effects of BZDs are mild dizziness and sedation, both of which are generally attenuated by time or dose reduction. Patients should be cautioned about using heavy or dangerous machinery and driving, especially in initial therapy and after dose changes. BZDs should be avoided in combination with alcohol as they can intensify its effects. Although potential for abuse and development of tolerance to therapeutic effects are of some concern, when used appropriately under medical supervision BZDs are efficacious, have a rapid onset of action, and are generally well tolerated. They are particularly useful for as-needed (PRN) use in individuals with acute episodes and exacerbations of anxiety, including performance anxiety and panic attacks.

In general, BZDs are classified by their half-life as short-, intermediate-, and long-acting. Agents with a relatively short or intermediate half-life (e.g., alprazolam [Xanax] or lorazepam [Ativan]) are most commonly prescribed for PRN use, but are also prescribed for regular use. A brief course of BZDs (3 to 4 weeks) is frequently used as an adjunct to an antidepressant (often initiated with the antidepressant), both to improve delayed or partial response or to help relieve initial adverse effects from the antidepressant (e.g., insomnia, jitteriness). At other times, BZD

monotherapy can be justified when other treatments have failed or are not tolerated, and when anxiety is persistent, severe and disabling. It is suggested that patients who have been stabilized on BZDs for many years with continued therapeutic response and no evidence of abuse are best maintained on it. There is increasing consensus that switching medications to take them off BZDs may not be of clinical benefit. It is of particular note that BZD monotherapy is not recommended in patients with a comorbid depressive disorder. They are also best avoided in individuals with histories of alcohol or substance abuse, unless there are compelling clinical reasons (e.g., non-response to other agents); their prescription in this context should be undertaken with due caution. There have been concerns expressed in the medical literature about long-term benzodiazepine use possibly increasing risk for dementia. Studies to date are suggestive, but inconclusive, particularly in implicating benzodiazepines, per se, as opposed to the indications for their use. But it will be important to follow this literature as it evolves in order to establish a risk-benefit profile for the use of benzodiazepines.

Anticonvulsants

Antiepileptic drugs have long been used in the treatment of mood disorders, but their use in anxiety disorders is still preliminary. An important development is the positive evidence for the benefit of pregabalin (Lyrica) in these disorders. Pregabalin, a structural analogue of GABA, has evidence from several good-quality RCTs supporting its use as a monotherapy as well as adjunctive therapy in several anxiety disorders, both in acute treatment and prophylaxis. It has been noted to improve sleep and may also reduce depressive symptoms. Dizziness and drowsiness are common side effects and weight gain occurs in about 20 percent of patients. It can be used safely in patients with hepatic dysfunction but is best avoided in patients with renal impairment. Other anticonvulsant medications, often used as adjunctive therapies to antidepressants, can be considered on a case-by-case basis in the treatment of patients with refractory anxiety disorders. These medications include gabapentin (Neurontin), levetiracetam (Keppra), topiramate (Topamax), tiagabine (Gabitril), and lamotrigine (Lamictal); however, evidence of efficacy from RCTs for most of these agents is either lacking or preliminary.

Antipsychotics

The use of antipsychotics for anxiety disorders has been explored in several clinical trials. For many years, the only antipsychotics available were the “typical” variety whose widespread use was limited by their adverse event profiles, including tardive dyskinesia (TD) and extrapyramidal symptoms (EPS). The newer “atypical” antipsychotics (AAPs), including olanzapine

(Zyprexa), risperidone (Risperdal), quetiapine (Seroquel), ziprasidone (Geodon), aripiprazole (Abilify), and lurasidone (Latuda), were introduced starting in the 1990s. AAPs are distinguished from the older generation of antipsychotics in part by their ability to affect not only dopamine blockade, but also blockade of the 5-HT2A serotonin receptor, and in part by a lower associated risk of developing the adverse events commonly found with the typical antipsychotics. AAP augmentation of antidepressants is widespread for treatment of depression. Investigations in anxiety disorders are less common; however, several studies—including randomized controlled trials—have found quetiapine to be a useful agent, either as monotherapy or as an augmentation agent for both acute and maintenance treatment of GAD. With this exception, evidence for use of AAPs in any of the anxiety disorders—either adjunctively or as monotherapy for refractory presentations—is equivocal or lacking. Even where the evidence of efficacy is solid, as in the case of quetiapine, the risks (e.g., weight gain and other adverse metabolic effects) may outweigh the benefits and other classes of antianxiety agents are preferred.

Other Antianxiety Medications

β-Adrenergic antagonists, also called β-Blockers, are used primarily on a PRN basis for performance anxiety. They reduce tachycardia (racing heart) and tremor (shaking), physical symptoms of anxiety on which patients with performance anxiety tend to focus. β-Blockers used for this purpose include propranolol (Inderal) and atenolol (Tenormin). Dosing regimens vary, with some patients started on fixed doses (20 to 40 mg per dose for propranolol and 25 to 50 mg per dose for atenolol) and others receiving dose titrations to yield an approximate 10-beat-per-minute reduction in heart rate compared to pretreatment; patient feedback on impact on performance anxiety can help to establish the appropriate dose. Of note, however, is the difference between the high efficacy rates described in the literature and anecdotal reports of benefit in less than 50 percent of cases; PRN short-acting benzodiazepines may provide a reasonable alternative for such patients. β-Blockers used PRN are generally well tolerated, but occasional side effects include dizziness, lightheadedness, or fatigue. They should not be prescribed to patients with cardiac or pulmonary diseases, angle-closure glaucoma, or diabetes.

Buspirone (BuSpar), an azapirone, is effective in the treatment of GAD. Buspirone is usually prescribed in three divided doses, with total daily doses ranging from 30 to 60 mg per day. The time course of response to buspirone is similar to that of antidepressants, and it usually takes at least several weeks, and sometimes several months, to reach full therapeutic efficacy. There is some evidence, most of it anecdotal, to suggest that buspirone may be useful as an adjunct to antidepressants in the treatment of anxiety disorders other than GAD.

Hydroxyzine (Atarax) is an antihistamine with preliminary evidence of benefit in the treatment of GAD. It may be a good short-term alternative to BZDs for some patients, but the long-term efficacy of hydroxyzine for anxiety disorders is unknown.

Novel Therapeutics for Anxiety

Despite the wealth of pharmacotherapeutic options for the treatment of anxiety disorders, clinicians continue to struggle with a subset of patients that remains resistant or only partially responsive to conventional therapies. In response, researchers have become increasingly interested in the use of novel therapies that may represent a better solution to this issue. To date, there has been a paucity of studies focused on strategies for the management of treatment-resistant anxiety disorders. Repetitive transcranial magnetic stimulation (rTMS) has shown promise as a treatment for depression or depression comorbid with anxiety disorder; however, evidence for rTMS applied for pure anxiety disorders is sparser. Case reports have suggested utility for rTMS in SAD and PD but a double-blind trial was unable to show a difference between real and sham rTMS treatments for PD. More promisingly, a small trial of fMRI guided rTMS for GAD showed positive results following acute treatment with gains generally maintained at 6-month follow-up; however, more controlled studies would be helpful to validate these findings.

Increasingly, consumers, particularly those with mood and anxiety disorders, are turning to complementary and alternative therapies, to either self-treat or augment their more conventional medications. These therapies include: herbal medications, acupuncture, biofeedback, and yoga, among others. However, many herbal preparations have the potential to cause drug–drug interactions, and so it is imperative that clinicians routinely ask after patients' use of these substances. Two recent meta-analyses noted that kava, one of the most common herbs taken for anxiety, does demonstrate benefit for mild-to-moderate anxiety, but its utility in specific anxiety disorders is still unclear and there is also possible risk of hepatotoxicity and documented interactions with medications metabolized through the cytochrome P450 enzyme system. Other herbal remedies with purported antianxiety properties but insufficient clinical data to support their use include valerian root, passiflora (passion flower), inositol, St. John's wort, gotu kola (Centella asiatica), California poppy, hawthorn, Ginkgo biloba, lemon balm, and sour date seed. Among other alternative therapies, exercise has sparse data for utility in anxiety conditions, but preliminary evidence suggests benefits of yoga for performance anxiety, and amino acid lysine and magnesium–calcium combination for mild anxiety.

TREATMENT RESPONSE AND COURSE OF ILLNESS ACROSS ANXIETY DISORDERS

WITH PHARMACOTHERAPY

In general, all first-line agents are thought to be equally effective, but in practice, SSRIs are often chosen as the first-line options. Starting with lower doses and increasing them gradually, as well as adding concomitant benzodiazepines, may improve tolerability. Partial or non-response to the initial trial of a first-line antidepressant may necessitate a switch to another first-line agent. Augmentation with benzodiazepines, e.g., alprazolam, clonazepam, etc., also has significant evidence for benefit. Guidelines also recommend combining pharmacotherapy with psychological treatment for this population.

Although individually both pharmacotherapy and psychotherapy (e.g., CBT) show efficacy in the treatment of anxiety disorders, their combination has been little studied. Available data shows inconsistent benefit for the combination; this may be because either treatment modality alone is sufficiently good that showing an additive effect is difficult without enormous sample sizes, or because superiority of the combination to individual therapies in acute treatment fails to be sustained at follow-up. Some authorities recommend that pharmacotherapy and CBT be started simultaneously. However, for patients with partial or non-response to pharmacotherapy, sequential provision of CBT is a good option.

Empirical data on the optimal duration of pharmacotherapy are lacking for most anxiety disorders. On average, 8 to 12 weeks of medications may be needed to evaluate efficacy in acute treatment. A number of controlled studies do show, however, that 20 to 50 percent of patients with chronic anxiety disorders relapse within several months of discontinuing antidepressant pharmacotherapy. It is thus recommended that such patients be maintained on treatment for at least 1 to 2 years before consideration is given to dose reduction or discontinuation, and that any discontinuation be done by gradual decrements of 10 to 25 percent every 1 to 2 months while observing for symptom return or exacerbation.

Response rates to pharmacotherapy differ somewhat among the anxiety disorders. PD is among the most treatment-responsive anxiety disorders, with response rates of 60 to 70 percent with effective interventions. Patients with comorbid agoraphobia may be somewhat more difficult to treat, in some cases requiring higher doses of medication or, more often, benefiting from concomitant CBT. Response rates to effective pharmacotherapy in GAD and SAD are slightly lower, with 50 to 60 percent of patients having a good clinical outcome.

INDIVIDUAL ANXIETY CONDITIONS AND THEIR MANAGEMENT

The following sections review the literature for the treatment of individual anxiety conditions and make medication recommendations based on the strength of evidence available for different classes of drugs. See Table 14.8–

4 for a summary of recommendations and Table 14.8–5 for general pointers.

Table 14.8–4.
Evidence-Based Recommendations for Monotherapy of Anxiety Disorders

	Anxiety Disorders		
Classes and Agents	**GAD**	**PD**	**SAD**
SSRIs	First line	First line	First line
SNRIs	First line	First line	First line
TCAs	Second Line	Second line	Not recommended
MAOIs	Insufficient Evidence	Second Line	Second line
RIMA	Insufficient Evidence	Insufficient evidence	Second Line
Other Psychotropics—Agomelatine, Buspirone, Mirtazapine	Second line	Second line	Second line
Benzodiazepines	Second line	Second line	Second line
Atypical Antipsychotics—Quetiapine	Second Line	Insufficient Evidence	Not Recommended
Anticonvulsant—Pregabalin	Second Line	Insufficient Evidence	Second line

Note: Beta-adrenergic blocking agents are recommended only for performance anxiety.

Table 14.8–5.
Pharmacotherapy of Anxiety Disorders: Key Pointers

- ► SSRIs as the first-line option.
- ► Start low and go slow.
- ► Routine increase to higher doses not recommended, but a subgroup might benefit.
- ► Concomitant benzodiazepines as initial therapy in the short-term may be useful.
- ► 8–12 weeks of pharmacotherapy at optimum doses may be needed to assess efficacy.
- ► Good evidence for the benefit of maintenance treatment at least up to 6 months.

Generalized Anxiety Disorders

GAD has a lifetime prevalence of 6 percent. Its key feature is excessive and difficult to control anxiety and worry about multiple events and activities, often accompanied by psychic and somatic symptoms of anxiety. It often presents with physical symptoms, and comorbidities, particularly of depression and other anxiety disorders, are common. Contrary to popular belief that GAD is a mild and acute condition, in many patients it is quite disabling and often runs a chronic course.

Large RCTs as well as systematic reviews and meta-analyses confirm the benefit of SSRIs (fluoxetine, citalopram, escitalopram, paroxetine, sertraline) and SNRIs (venlafaxine, duloxetine) as first-line choices. Agomelatine, pregabalin, buspirone and quetiapine are suggested as

alternative options. Imipramine, a TCA, has significant evidence for benefit but is often used as a second-line option due to adverse effects. Another highly prescribed class of agents for GAD is BZDs, which include alprazolam, lorazepam, diazepam, and bromazepam. Evidence for their benefit is strong but they are often only used short term, due to risk of tolerance and dependence. While quetiapine has significant efficacy in GAD, guidelines recommend its use as a second-line intervention due to metabolic adverse effects. Buspirone is often used similarly due to lack of comparative data with antidepressants. Other AAPs (olanzapine, risperidone and aripiprazole) and bupropion, which have lower evidence of efficacy, may also be useful as adjunctive therapies. However, propranolol, tiagabine, and ziprasidone have negative evidence in GAD.

For relapse prevention, there is good evidence to support the use of SSRIs and SNRIs as maintenance therapy. A similar benefit in prophylaxis has also been demonstrated with pregabalin and quetiapine. The benefit of continued maintenance medications may last at least 6 months and that CBT may be equally effective in prophylaxis.

Panic Disorder

Panic disorder has a lifetime prevalence of 5 percent and is characterized by the presence of recurring panic attacks in the absence of triggers, with significant physical and psychological symptoms. As with other anxiety disorders, comorbidities, particularly with depression and other anxiety disorders, are common. While periods between panic attacks are relatively symptom-free, persistent concern about future attacks and maladaptive changes in behavior to avoid attacks are present. These symptoms often overlap with agoraphobia, with marked fear and avoidance behavior relating to specific situations. Patients with panic disorder are often under-recognized and frequently present in the emergency and specialized medical clinics. Panic disorder is also commonly associated with medical conditions, in particular, thyroid disease, chronic pain, and malignant conditions, as well as cardiac and respiratory disorders.

Meta-analyses and RCTs have confirmed the benefit of several classes of antidepressants. SSRIs (citalopram, fluoxetine, paroxetine, sertraline, fluoxetine and escitalopram) and the SNRI venlafaxine, are often recommended as initial therapy. TCAs (in particular, clomipramine and imipramine), as well as MAOIs (phenelzine and tranylcypromine) have equal efficacy but are often used as second-line because of their adverse event profiles. Other second-line agents include mirtazapine, bupropion and reboxetine. Several BZDs (alprazolam, lorazepam, diazepam, and clonazepam) have also been proven effective in reducing panic symptoms. Although second-line interventions due to risk of dependence, they can be useful as concomitant short-term treatment to reduce or prevent panic attacks and reduce anticipatory anxiety. AAPs have only preliminary

evidence of benefit and are restricted to adjunctive third-line options for cases of partial or non-response to first-line agents. Studies with propranolol, tiagabine, gabapentin, and carbamazepine have reported negative findings, while there are no reports evaluating pregabalin.

Maintenance studies in panic disorder have reported that many treatment responders show further improvement over time and continued antidepressant therapy is effective in preventing relapse, with beneficial effects lasting up to 1 to 3 years. First-line agents show similar efficacy in maintenance therapy, and notably, evidence suggests that BZDs (alprazolam and clonazepam) may provide effective prophylaxis without development of tolerance. Discontinuing maintenance treatment should be done very gradually, particularly with BZDs.

Agoraphobia

Most presentations of agoraphobia are accompanied by panic attacks or PD, and literature confirms that agoraphobic symptoms improve in parallel with panic symptoms upon appropriate pharmacotherapy for PD. In other cases, agoraphobic symptoms overlap with those of social phobia, GAD, and avoidant personality disorder, and usually, treatment of the comorbid condition is accompanied by a reduction in agoraphobia. The efficacy of pharmacotherapy in "pure" agoraphobia has not been well evaluated; early studies with BZDs (alprazolam and diazepam), TCA imipramine, and MAOI phenelzine had negative findings but little research has been done since. However, CBT with a focus on exposure therapy appears to be an effective intervention. It has also been suggested that patients with "pure" agoraphobia are often housebound and avoid seeking help, leading to an underestimation of its true prevalence and limiting research on its treatment.

Social Anxiety Disorder

With an estimated lifetime prevalence of 8 to 12 percent, SAD is a very common anxiety disorder. The key symptom is an excessive or unreasonable discomfort in social situations, often leading to avoidant behavior. Two subtypes, a generalized form and a form limited to performance situations, are recognized, though in the DSM-5, the "generalized" subtype specifier has been removed with solely the "performance only" specifier remaining. As with other anxiety disorders, comorbidity is common, in particular with major depressive disorder and other anxiety-related conditions. SAD also often overlaps with avoidant personality disorder.

Evidence from meta-analyses, systematic reviews, and large RCTS suggest that all SSRIs, venlafaxine, pregabalin, and clonazepam have strong evidence for benefit and are recommended as first-line treatments.

Phenelzine is also efficacious for SAD but because of its poor side effect profile, it is used as a second-line agent for refractory illness. Other secondary alternatives include mirtazapine and moclobemide, gabapentin, as well as the BZDs bromazepam and alprazolam. Due to negative findings, TCAs, buspirone, and quetiapine, are not recommended. While propranolol and atenolol can be beneficial in performance anxiety, they do not appear to help generalized SAD.

While routine use of higher doses of pharmacotherapy is not recommended, subgroups may need them to attain and maintain remission. However, lack of remission with even optimal treatment is not uncommon. Unfortunately, systematic evaluations of strategies to treat refractory patients are few. Switching to an alternate first-line agent or phenelzine, augmentation with pregabalin or clonazepam, and combination with CBT are among the common strategies employed.

Longer-term studies have noted that following acute treatment, a significant proportion of patients show incremental continued response, increasing the overall response rate. Relapse prevention RCTs confirm the benefit of several agents, including SSRIs, pregabalin and clonazepam over a 6-month period. As with acute treatment, psychological interventions have been documented to help relapse prevention.

Specific Phobia

The key feature of specific phobia is an intense fear of a specific object or situation, with significant avoidance behavior; certain presentations of specific phobia may be difficult to distinguish from panic disorder. Subsyndromal conditions are common and lifetime prevalence is estimated at 10 to 13 percent, with many affected reporting multiple fears. Clinical presentations are rarely for the phobia itself, but more often for comorbid depression and other anxiety disorders. Psychological interventions, particularly exposure-based therapies are considered to have the most evidence for efficacy. The benefit of pharmacotherapy has not been well studied, but preliminary evidence suggests that SSRIs (escitalopram and paroxetine) may be useful as second-line treatments if patients do not respond to psychotherapy. So far, BZDs have not been found to be beneficial even as adjunctive to exposure therapy, and may even reduce the efficacy of psychological treatments. Thus, in general, there is little evidence to support the benefit of pharmacotherapy in the management of specific phobias. Of note, presence of comorbid depression and other anxiety disorders, which is frequently found, would necessitate treatment with appropriate first-line pharmacological agents.

TREATMENT OF ANXIETY DISORDERS IN CHILDREN AND ADOLESCENTS

Several RCTs have provided conclusive evidence that SSRIs are useful in

the treatment of certain anxiety disorders in children and adolescents, particularly GAD, SAD, and separation anxiety disorder (SepAD) in youth. Preliminary data also suggests that the combination of medication and psychotherapy is better than either CBT or medication alone for SepAD in children. Similarly, while psychosocial interventions are preferentially used in treatment of selective mutism, early case reports and pilot studies suggest that SSRIs (particularly fluoxetine and fluvoxamine) may be helpful especially in more severe cases. The black box warnings issued by the FDA regarding increased risks of suicidality with SSRI use in children and adolescents means these medications are now generally reserved for young people who do not respond to psychological interventions. Nevertheless, the findings of meta-analyses that the increase in absolute risk of suicide is very small, and recognition of the serious functional morbidity associated with chronic anxiety and avoidance in children and adolescents, would suggest that pharmacotherapy should not be unduly withheld in these situations.

FUTURE DIRECTIONS

Neurobiological advances in anxiety disorders continue to advance in parallel with progress in nosology. Currently in development are several new classes of agents with novel mechanisms of action, including corticotrophin-releasing factor type 1 (CRF-1)–receptor antagonists, metabotropic glutamate receptor agonists, and agents promoting neurogenesis. Newer targeted neurostimulation interventions may further enhance the ability to treat more refractory anxiety conditions. It is further hoped that the rapidly growing fields of functional neuroimaging and molecular genetics and their integration will also provide significant advances in the pathophysiology and the treatment of these conditions that affect the wellbeing of a significant proportion of the population.

REFERENCES

*Baldwin DS, Anderson IM, Nutt DJ, et al. Evidence-based pharmacological treatment of anxiety disorders, post-traumatic stress disorder and obsessive-compulsive disorder: a revision of the 2005 guidelines from the British Association for Psychopharmacology. *J Psychopharmacol*. 2014;28:403.

*Billioti de Gage S, Pariente A, Begaud E. Is there really a link between benzodiazepine use and the risk of dementia? *Expert Opin Drug Saf*. 2015;14:733–747.

*Craske MG, Stein MB. Anxiety. *Lancet*. 2016. (in press). doi: 10.1016/S0140-6736(16)30381-6.

Hua A, Major N. Selective mutism. *Curr Opin Pediatr*. 2016;28:114–120.

Kang EH, Kim B, Choe AY, Lee JY, Choi TK, Lee SH. Panic disorder and health-related quality of life: the predictive roles of anxiety sensitivity and trait anxiety. *Psychiatry Res*. 2015;225:157–163.

*Katzman MA, Bleau P, Blier P, et al. Canadian clinical practice guidelines for the management of

anxiety, posttraumatic stress and obsessive-compulsive disorders. *BMC Psychiatry*. 2014;14 Suppl(1):S1.

Montejo AL, Montejo L, Navarro-Cremades F. Sexual side-effects of antidepressant and antipsychotic drugs. *Curr Opin Psychiatry*. 2015;28:418–423.

*Olmer A, Iancu I, Strous RD. Exposure to antidepressant medications and suicide attempts in adult depressed inpatients. *J Nerv Ment Dis*. 2012;200:531.

Purohit MP, Zafonte RD, Sherman LM, et al. Neuropsychiatric symptoms and expenditure on complementary and alternative medicine. *J Clin Psychiatry*. 2015;76:e870.

Roest AM, deJonge P, Williams CD, de Vries YA, Schoevers RA, Turner EH. Reporting bias in clinical trials investigating the efficacy of second-generation antidepressants in the treatment of anxiety disorders: a report of 2 meta-analyses. *JAMA Psychiatry*. 2015;72:500–510.

Stein MB, Sareen J. CLINICAL PRACTICE. Generalized anxiety disorder. *N Engl J Med*. 2015;373:2059–2068.

*Strawn JR, Welge JA, Wehry AM, Keeshin B, Rynn MA. Efficacy and tolerability of antidepressants in pediatric anxiety disorders: a systematic review and meta-analysis. *Depress Anxiety*. 2015;32:149.

Umetsu R, Abe J, Ueda N, et al. Association between selective serotonin reuptake inhibitor therapy and suicidality: analysis of U.S. Food and Drug Administration Adverse Event Reporting System Data. *Biol Pharm Bull*. 2015;38:1689–1699.

15

Obsessive-Compulsive and Related Disorders

DAN J. STEIN, M.D., PH.D., AND CHRISTINE LOCHNER, PH.D.

Obsessive-compulsive and related disorders (OCRDs) constitute a new chapter in both the DSM-5 and the proposed ICD-11. This grouping of disorders includes obsessive-compulsive disorder, body dysmorphic disorder (BDD), hoarding disorder (HD), trichotillomania (TTM) (hair-pulling disorder), and excoriation (skin-picking) disorder. The DSM-5 revision process emphasized the importance of diagnostic validity and clinical utility, while the ICD-11 revision process emphasized the importance of clinical utility and global applicability. Both nomenclatures have included the OCRDs in a separate chapter in view of evidence that the disorders share certain phenomenological and psychobiological features, and with the aim of improving the recognition and management of these underdiagnosed and undertreated conditions. Here, each of the DSM-5 OCRDs will be covered in turn.

OBSESSIVE-COMPULSIVE DISORDER

Introduction

Obsessive-compulsive disorder (OCD) is the paradigmatic OCRD, and the one on which most research has been done. Like many of the other OCRDs, it is characterized by repetitive intrusive thoughts, and repeated mental acts or behaviors in response to those thoughts. OCD was once central to psychodynamic thinking, but it is now an exemplar par excellence of a neuropsychiatric disorder insofar as its underlying neurocircuitry is to some extent understood. Furthermore, whereas OCD was once considered rare it is now known to be quite prevalent—patients are often embarrassed by and ashamed of their symptoms, and so careful screening is needed to ensure case-finding. Finally, whereas OCD was once considered very refractory to treatment, the introduction of specific pharmacotherapeutic and psychotherapeutic interventions has led to more favorable outcomes for many.

Definition

OCD is characterized by the presence of obsessions and/or compulsions. Obsessions are intrusive and unwanted repetitive thoughts, urges, or impulses that often lead to a marked increase in anxiety or distress. Compulsions are repeated behaviors or mental acts that are done in response to obsessions, or in a rigid rule-bound way. Some patients may have only obsessions or only compulsions, but in most cases both obsessions and compulsions are present.

History

Although there are earlier descriptions of OCD symptoms in the medical literature, the English term derives from Westphal's description of "zwangsvorstellung" in the late 19th century, which was translated as "obsession" or "compulsion." Janet subsequently made an important contribution to the literature, describing the role of "incompleteness" in OCD patients, as well as symptoms such as tics. Freud's theories on the role of the unconscious in obsessive-compulsive symptoms were particularly influential on clinical practice in the early part of the 20th century, as was his description of a spectrum of symptoms ranging from obsessive-compulsive personality traits, through obsessive-compulsive neurosis, and on to obsessive-compulsive psychosis.

In the middle of the 20th century, case reports indicated that the serotonergic tricyclic antidepressant, clomipramine, was useful in OCD. Randomized controlled trials confirmed the efficacy of clomipramine, and gave impetus to a serotonin hypothesis of OCD. The introduction of behavior therapy and of selective serotonin reuptake inhibitors (SSRIs) for the treatment of OCD were additional major steps forward. The finding that OCD was accompanied by unique patterns of activation on brain imaging, and that both of these interventions helped normalize the functional neuroanatomy of OCD gave significant impetus to further research on this condition. Over time, interventions, including dopaminergic agents and neurosurgical interventions were also developed for treatment-refractory OCD.

The finding that clomipramine was more effective than desipramine, a noradrenergic tricyclic, in OCD as well as in a range of other conditions characterized by intrusive thoughts or repetitive behaviors, helped consolidate the notion of an obsessive-compulsive spectrum of disorders. SSRIs proved to be useful for a number of these conditions, but not all. Overlapping, but somewhat different behavioral approaches, have also proven useful for these conditions. At the start of the 21st century, cutting-edge research methods, including neurogenetics, have been used to help delineate the relationship between the OCRDs, particularly OCD and tic disorders.

Comparative Nosology

DSM-III provided operational criteria for OCD, defining obsessions and compulsions along the lines of the definition above. DSM-IV introduced the notion that compulsions can be mental acts. DSM-5 made relatively subtle adjustments to the DSM-IV criteria with the aim of further clarifying the symptoms, for example, replacing the term "impulse" with "urge" (to avoid confusion with impulse-control disorders) and noting that obsessions cause anxiety in most individuals (given that this is not the case in all individuals). However in DSM-5, OCD was shifted into a new chapter on OCRDs, a more substantial change. Another significant change from DSM-IV was the extension of the insight specifier to include absent insight; this emphasizes to clinicians that some patients with OCD (around 4 percent) may have delusional thoughts, and yet they should be diagnosed with OCD rather than with a psychotic disorder.

The ICD-10 description of OCD does not include the notion that obsessions give rise to compulsions. There is certainly some support in the literature for the idea that compulsions may not precede obsessions, and indeed some have even suggested that the disorder be termed compulsive-obsessive disorder. At the same time, from a treatment planning perspective, it may be useful for clinicians to understand the cycle of obsessions and compulsions, and the proposal for ICD-11 is to outline the way in which, in many patients, obsessions lead to anxiety or distress, and compulsions are undertaken to reduce these negative emotions. Like DSM-5, ICD-11 will for the first time have a chapter on OCRDs, and it will emphasize that some patients with OCD have absent insight.

Proposals for ICD-11 indicate that the OCRD chapter may include a few more disorders than the DSM-5 one; in particular it may include tic disorders, olfactory reference disorder, and hypochondriasis. Tic disorders are classified in both DSM-5 and ICD-11 as neurodevelopmental disorders, but ICD-11 may allow multiple-coding, so that these conditions may also be referenced in OCRDs, reflecting factors such as their close familial relationship to OCD. ICD-11 proposals emphasize that while patients with BDD are preoccupied with flawed appearance, those with olfactory reference disorder are preoccupied with smelling bad; with both groups displaying repetitive mental acts and behaviors in response to these concerns. DSM-5 includes illness anxiety disorder in the somatoform disorders, while ICD-11 proposals emphasize that this condition is characterized by intrusive repetitive thoughts and repeated behaviors in response, and so should also be in the chapter on OCRDs.

Epidemiology

Individuals with OCD are characteristically aware that their compulsions are excessive, and in the past they may not only have been too embarrassed to seek professional help, but also unlikely to have received efficacious intervention—factors that may have led to early authors describing OCD as

a rare condition. The Epidemiological Catchment Area study found that OCD was in fact one of the most prevalent of all psychiatric disorders, and follow-up cross-national surveys confirmed its presence across a number of different countries. The National Comorbidity Survey Replication (NCS-R) was able to delineate clinical from subclinical obsessive-compulsive symptoms in a rigorous way, and findings were that OCD had a lifetime prevalence of 2.3 percent, but that subclinical OCD was much more prevalent.

A number of additional findings with public health relevance emerge from community and clinical studies. First, although exact prevalence of OCD differs from study to study, OCD is a relatively common disorder in most countries surveyed, and subclinical OCD is even more prevalent. Second, OCD is accompanied by significant comorbidity; with particularly high rates of OCRDs, of mood, and of anxiety disorders. Third, OCD is accompanied by significant morbidity, including impairment in social, academic, and occupational functioning as well as a reduction in quality of life that can be comparable to that seen in other serious medical and mental disorders. Early Global Burden of Disease data indicated that OCD was one of the most disabling of all medical disorders, with more recent iterations emphasizing the important contributions of anxiety disorders (including OCD) as a whole. Fourth, even in high-income countries, there is significant underdiagnosis and undertreatment of OCD. Fifth, partly because of this underdiagnosis and undertreatment, OCD is associated with significant economic costs, both for the individual and for society as a whole.

OCD has a median age of onset of 19 in the NCSR. Approximately one-quarter of cases have onset by age 10, with early onset more common in males. In other cases OCD may have a later onset; in females onset is often during adolescence, with some developing OCD during pregnancy or the puerperium. Onset after 30 is rare. OCD has surprisingly similar features in children and in adults, but compulsions are typically more central to the presentation earlier in life, with some children either not experiencing obsessions, or being unable to articulate them. Community studies indicate a higher prevalence of OCD in females, but in clinical contexts there are roughly equal numbers of male and female patients. Males and females have similar symptoms, but males are more likely than females to have comorbid tics. Individuals with OCD have an increased family history of OCD, as well as of OCRDs including tic disorders.

Etiology

OCD is increasingly viewed as a neuropsychiatric disorder that is mediated by specific neurocircuits. From a cognitive-affective neuroscience perspective, OCD is characterized by a number of specific neuropsychological impairments, including cognitive inflexibility and

alterations in the habit system. Early clinical work established that OCD could be precipitated by damage to striatal circuitry, as was seen in the global influenza epidemic in the early years of the 20th century. More recent advances in structural and functional brain imaging have delineated the role of these and related circuits in mediating OCD, and translational work in animal models has allowed further delineation of the neuroanatomy of repetitive behavior. A range of laboratory and clinical work in recent decades has focused on the neurotransmitters and other molecules that are found in these circuits, and which may be involved in the pathogenesis of OCD.

Cognitive and neuropsychological studies of OCD have found a number of alterations, including a range of impairments in executive function, and with evidence for cognitive inflexibility, motor impulsivity, and excessive habit formation. Such work is consistent with the hypothesis that OCD represents impaired control of automated, habitual behaviors. A second group of studies has emphasized alterations in affective and reward processing in OCD, with perhaps the most specific findings those related to perceptions of disgust. Patients with OCD appear to show increased recognition of disgust, compared to other emotions. This finding is consistent with the view that OCD differs in important ways from anxiety disorders, which are characterized by alterations in fear processing. These cognitive-affective findings, as well as work on temperamental risk for OCD (such as increased behavioral inhibition) are consistent with a view which emphasizes that OCD symptoms, particularly in subclinical form, can be understood through the lens of evolutionary theory; they have an adaptive role in protecting individuals from harmful stimuli, whether by grooming to ensure the removal of harmful organisms or via the avoidance of disgust cues.

A number of brain regions, including anterior cingulate, orbitofrontal cortex, and striatum have consistently been implicated in OCD, although there is also evidence that a wide range of other regions play a role. One hypothesis has been that cortico-striatal-thalamic-cortical (CSTC) circuitry plays a particularly important role in mediating the cognitive-affective impairments seen in OCD, with activation or inhibition of different components of this circuitry driving the compulsive and impulsive features of OCRDs. Laboratory research has indicated that these circuits are involved in the control of automated behaviors such as grooming and habits. Clinical research has indicated that there are gray matter volumetric changes in components of this circuitry, as well as alterations in frontostriatal connectivity and a range of white matter tracts in OCD. A growing database of functional imaging studies have found increased activity in these circuits both at rest, and on exposure to feared stimuli. Early imaging work was seminal in showing that both pharmacotherapy and cognitive-behavioral therapy (CBT) were able to normalize functional

alterations in CSTC circuitry in OCD, and the relevant evidence-base has continued to grow.

Magnetic resonance spectroscopy and molecular imaging studies have begun to delineate the specific neurochemical alterations found in CSTC and other circuitry implicated in OCD. Such work has been complemented by work on blood and cerebrospinal fluid (CSF) metabolites, as well as by studies of pharmacological "challenges" in OCD, of pharmacotherapeutic dissection, and of neurogenetics. This work has indicated that a range of neurotransmitter systems contribute to OCD, including serotonergic, dopaminergic, glutamatergic, and GABAergic systems. A serotonin hypothesis of OCD has been influential; a range of evidence points to the involvement of this system in OCD, and patients respond to treatment with serotonin reuptake inhibitors (SRIs) but not to treatment with noradrenergic reuptake inhibitors. At the same time, there is relatively little evidence that serotonin plays a causal role in OCD. Furthermore, a range of evidence points to the involvement of the dopaminergic system in OCD, and it is notable that augmentation of serotonin reuptake blockers with dopamine receptor blockers is considered a first-line intervention for treatment-refractory OCD. Most recently, there has been interest in additional systems, including the glutamatergic system in OCD, with evidence that some glutamatergic agents may be useful in this condition.

Future work on the cognitive-affective neuroscience of OCD may be increasingly influenced by genetic findings. Twin and family studies indicate that genetic susceptibility plays a role in OCD, and more so in childhood-onset OCD. To date, OCD has been associated with serotonin-related polymorphisms, as well as with variants involved in catecholamine modulation; taken together meta-analysis of candidate gene studies suggest that OCD is associated with multiple genes of relatively small effect. There are growing data from genome wide association studies; and is notable that the most significant findings have included genes that contribute to glutamatergic function. It is notable that gene variants in OCD may be associated with some specific brain imaging findings, and the future hope is that a better understanding of the relevant intermediate phenotypes involved in OCD, and of their underlying molecular biology will be useful in developing a personalized medicine approach.

Diagnosis and Clinical Features

OCD can readily be diagnosed on the basis of a thorough psychiatric history and examination. While patients often attempt to neutralize obsessions with compulsions, these behaviors or mental acts are not connected in a realistic way with what they are designed to prevent, or are clearly excessive. The DSM-5 diagnostic criteria for OCD address the presence of obsessions and/or compulsions, but also emphasize that these symptoms are time-consuming or distressing/impairing, are not due to

substances/medications or to other medical conditions, and are not due to other psychiatric disorders (Table 15–1). Repetitive thoughts and behaviors can be seen in a range of other mental disorders, as is discussed in more detail below. When assessing the negative impact of OCD symptoms, it is important also to assess avoidance of triggers of obsessions and compulsions. In addition, key features to ascertain during the history and examination of an individual with OCD include the age of onset and course of the symptoms and any possible precipitants, the type and severity of symptoms, the extent of insight and of impairment, and the presence of comorbid disorders and symptoms.

Table 15–1.
DSM-5 Diagnostic Criteria for OCD

A. Presence of obsessions, compulsions, or both:
   Obsessions are defined by (1) and (2):
   1. Recurrent and persistent thoughts, urges, or images that are experienced, at some point during the disturbance, as intrusive and unwanted, and that in most individuals cause marked anxiety or distress.
   2. The individual attempts to ignore or suppress such thoughts, urges, or images, or to neutralize them with some other thought or action (i.e. by performing a compulsion).

   Compulsions as defined by (1) and (2):
   1. Repetitive behaviors (e.g., handwashing, ordering, checking) or mental acts (e.g., praying, counting, repeating words silently) that the person feels driven to perform in response to an obsession or according to rules that must be applied rigidly.
   2. The behaviors or mental acts are aimed at preventing or reducing distress, or preventing some dreaded event or situation; however, these behaviors or mental acts either are not connected in a realistic way with what they are designed to neutralize or prevent, or are clearly excessive.

   Note: Young children may not be able to articulate the aims of these behaviors or mental acts.
B. The obsessions or compulsions are time-consuming (e.g., take more than 1 hour per day) or cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
C. The obsessive-compulsive symptoms are not attributable to the direct physiological effects of a substance (e.g., a drug of abuse, a medication) or another medical condition. Etiological substances/medications include: intoxication with cocaine, amphetamines or other stimulants and exposure to heavy metals. Etiological medical conditions include: Sydenham chorea, medical conditions leading to striatal damage, such as cerebral infarction.
D. The disturbance is not better explained by the symptoms of another mental disorder (e.g., excessive worries, as in generalized anxiety disorder; preoccupation with appearance, as in body dysmorphic disorder; difficulty discarding or parting with possessions, as in hoarding disorder; hair pulling, as in trichotillomania [hair-pulling disorder]; skin picking, as in excoriation [skin-picking] disorder; stereotypies, as in stereotypic movement disorder; ritualized eating behavior, as in an eating disorder; preoccupation with having an illness, as in illness anxiety disorder; sexual urges or fantasies, as in a paraphilic disorder; impulses, as in disruptive behavior disorders; guilty ruminations, as in major depressive disorder; thought insertion or delusional preoccupations, as in a psychotic disorder; or repetitive patterns of behavior, as in autism spectrum disorder.)

***Specify* if:**

With good or fair insight: The individual recognizes that obsessive-compulsive disorder beliefs are definitely or probably not true or that they may not be true.

With poor insight: The individual thinks obsessive-compulsive disorder beliefs are probably true.

With absent insight/delusional: The individual is completely convinced that obsessive-compulsive disorder beliefs are true.

***Specify* if:**
Tic-related: The individual has a current or past history of a tic disorder.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

As noted above OCD may have an early age of onset; while there is often no specific precipitant, OCD onset has been linked to various infections, to certain substances, and to traumatic or stressful life events. In factor analytic studies of OCD, symptoms fall on a number of limited dimensions, including contamination/cleansing, forbidden thoughts/checking, symmetry/ordering, and hoarding symptoms. While patients with OCD typically have good insight into the fact that their OCD beliefs are not true, some patients have poor insight, and a small percentage have absent insight—their thoughts are in fact delusional. As such patients may be misdiagnosed with a psychotic disorder and treated with inappropriately high doses of antipsychotic medication, DSM-5 extends the insight specifier to include absent insight/delusional OCD.

DSM-5 also includes a tic-related specifier for OCD, as many patients with OCD have current or past history of tics. OCD with tics is more commonly seen in males, has an earlier age of onset, is more likely to be accompanied by certain OCD symptom dimensions (e.g., symmetry/ordering), and may be less likely to respond to SSRIs. Both OCD symptoms and tics can be preceded by sensory phenomena or premonitory urges, and in the case of complex tics, it may be quite difficult to differentiate tics from compulsions. Given their comorbidity, and the close familial relationship between OCD and tic disorders, there is a strong argument that Tourette disorder (characterized by motor and vocal tics) is the obsessive-compulsive related disorder par excellence. Proposals for DSM-5 argued, however, that TD was best classified as a neurodevelopmental disorder. Proposals for ICD-11 have suggested that TD be multiply-coded; classified in the sections on neurological disorders, neurodevelopmental disorders, and OCRDs.

Individuals with OCD are much more likely to suffer from other OCRDs than those without OCD, and they are also much more likely to provide a family history of such conditions. The most common comorbid disorder in OCD patients presenting to the clinic is, however, depression. From a clinical perspective, it is crucially important to assess for the presence of suicidal symptoms, which are quite prevalent in OCD. Comorbid anxiety disorders are also common in OCD. It is notable that substance use disorders are more common in community studies than in clinical studies of OCD. Individuals with OCD may also suffer from a number of personality disorders, including schizotypal personality disorder and

obsessive-compulsive personality disorder (OCPD). There is an argument that OCPD—although not characterized by obsessions and compulsions, but rather by a pattern of excessive perfectionism and control—is best conceptualized as an OCD spectrum condition; OCPD may share neuropsychological impairments with OCD, and may perhaps respond to treatment with SSRIs.

There is a growing literature on symptom measures and diagnostic interviews for screening for OCD, for assessing severity, and for ensuring accurate differential diagnosis. DSM-5 includes a number of screening questions for OCRDs, as well as a self-rating measure for OCRDs. DSM-5 screening questions for OCRDs ask about "unpleasant thoughts, urges, or images that repeatedly enter your mind," and about "feeling driven to perform certain behaviors or mental acts over and over again." The DSM-5 self-rating questions assess the time occupied by symptoms, distress caused by the symptoms, difficulty in controlling the symptoms, avoidance caused by the symptoms, and interference associated with the symptoms. The Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) is a clinician-rated instrument that assesses similar items, and that is commonly used as a primary outcome measure in clinical trials. Recent adaptations of this scale include a version that provides a score for each major OCD symptom dimension. The Structured Interview for OCRDs provides questions that target DSM criteria for each of the OCRDs, as well as key issues important in different diagnosis (e.g., are symptoms exacerbated by substance use), so allowing a comprehensive assessment of each of these conditions.

Diagnosis and assessment should also cover a number of other key issues. First, it is important to ascertain the patient's explanatory model of OCD. Many patients are quite comfortable with a biomedical explanation of OCD. However, others may have entirely different models, and it is then particularly important to provide psychoeducation and to negotiate a joint approach to management. In occasional cases, obsessions and compulsions may themselves interfere with management (e.g., patients may view medications as dangerous). Second, it is important to ascertain the extent to which family members are impacted by or accommodate symptoms. This is particularly important when planning CBT, which demands increased exposure to feared stimuli and refraining from compulsions. Third, it is important to take a very careful medical history. Recent work has suggested that OCD may be precipitated by Streptococcal infection, or by a range of other insults, and so a comprehensive medical history is warranted.

In children with OCD, the aim of compulsions may not be easily articulated. In this age group it is also relevant to differentiate normal developmental rituals from OCD. In children and adolescents with OCD it is particularly important to assess the impact of OCD on developmental outcomes, to ask about a family history of OCD, and to assess how symptoms impact on or are accommodated by the family. Males and

females have similar presentations, although males are more likely to have earlier-onset, certain OCD symptoms, and comorbid tics, and females may have onset during pregnancy and the puerperium, and perhaps exacerbation of symptoms during the premenstrual period. OCD symptoms appear largely independent of culture and geography, although the specific content of symptoms may be influenced by the local context.

Pathology and Laboratory Examination

At the current time, there are no biomarkers for OCD. Brain imaging may ultimately be clinically useful in predicting response to treatment, but at the present time it remains a research tool. Similarly, while neurogenetic investigations have been valuable in the research context, they do not have sufficient clinical utility for more widespread use.

Differential Diagnosis

The psychiatric disorder hierarchy criterion in DSM-5 lists a large set of psychiatric conditions that should be differentiated from OCD: generalized anxiety disorder, BDD, HD, TTM, excoriation disorder, stereotypic movement disorder, eating disorders, substance-related and addictive disorders, illness anxiety disorder, paraphilic disorders, disruptive, impulsive-control and conduct disorders, major depressive disorder, schizophrenia spectrum and other psychotic disorders, and autism spectrum disorders. Worries in generalized anxiety disorder are about real-life concerns and they tend to be less irrational and egodystonic than in OCD and similarly the ruminations of depression and the preoccupations of mania are typically mood-congruent and egosyntonic; in neither case are symptoms neutralized by compulsions. In BDD, illness anxiety disorder, TTM, excoriation disorder, stereotypic movement disorder, anxiety disorders, and eating disorders, the focus of preoccupations and repetitive behaviors is distinct from that typically seen in OCD.

Hoarding behaviors can be seen in OCD; for example, patients may hoard newspaper articles in order to prevent harm should they forget an important matter. However, in most cases of HD, there is no specific obsession. Patients with OCD may experience a “just right” feeling on completion of compulsions, but in paraphilic disorders, substance-related and addictive disorders such as gambling disorder, and disruptive, impulse-control and conduct disorders repetitive behaviors are typically gratifying. Thus although terms such as compulsive gambling or compulsive stealing are used, these conditions should not be conflated with OCD. While some patients with OCD have absent insight into their symptoms, the delusional beliefs are related to OCD and they do not have other features of schizophrenia spectrum disorders. That said, a subgroup of OCD patients suffers from comorbid schizophrenia spectrum disorders

including schizotypal personality disorder. Similarly, while patients with autistic disorder often have repetitive symptoms, they also have a broad array of other characteristic symptoms.

Course and Prognosis

The course of OCD is generally chronic, and in the absence of intervention symptoms may persist for decades. Nevertheless, recent data emphasize that long-term outcomes can be positive. Thus patients should be encouraged to remain hopeful, and to try a range of different evidence-based interventions over time.

Treatment

Psychoeducation is a crucial component of the treatment of OCD. Exploring the patient's explanatory model of OCD is a useful first step in key elements of treatment such understanding the patient's suffering, building a collaborative approach to care, and instilling hope. Many patients and families are relieved to learn basic facts about OCD, including its high prevalence, the fact that many others are embarrassed by/ashamed of symptoms, that it is not due to commonly voiced factors (e.g., bad mothering), and that treatments are efficacious for many. Consumer advocacy organizations, such as the Obsessive-Compulsive Foundation, may play a particularly important role in conveying such information.

Pharmacotherapy

Seminal early work reported that patients with OCD responded to clomipramine, a relatively serotonergic tricyclic antidepressant. Subsequent randomized controlled trials showed that clomipramine was efficacious in both children/adolescents and adults with OCD, and that clomipramine was more effective than the relatively noradrenergic tricyclic antidepressant, desipramine, in OCD. Clomipramine became the first registered agent for the treatment of OCD, and a "serotonin hypothesis" of OCD influenced a good deal of subsequent research on both OCD and OCRDs.

The subsequent introduction of the SSRIs was important in that these provided a set of agents that was not only efficacious but also well tolerated. In meta-analyses of serotonin reuptake inhibitors (SRIs) in OCD, clomipramine has the largest effect size, but this may reflect the fact that these were simply the earliest trials. In head-to-head comparisons of clomipramine and SSRIs, there is equal efficacy, but SSRIs are better tolerated. Several of the SSRIs are Food and Drug Administration (FDA) or European Medicines Agency (EMA) approved for the indication of OCD, including childhood OCD, and these agents are therefore considered first-line for OCD pharmacotherapy.

Key clinical questions include the optimal dose and duration of SRIs in the pharmacotherapy of OCD. Although there is not a very steep dose–response curve, meta-analysis indicates that higher doses of SRIs are associated with better outcomes in OCD. Clinical guidelines therefore suggest increasing the dosage of SRIs over time in OCD to quite high ranges (e.g., escitalopram 30 mg, fluoxetine 80 mg, sertraline 200 mg), with some patients responding to even higher doses. Exceeding the recommended doses is not, however, advisable for clomipramine or citalopram, due to safety concerns. There is also some evidence to the effect that response to SRIs is slower in OCD than in depression; clinicians are therefore encouraged to prescribe an SRI for at least 12 weeks before deciding to switch agents or to augment with another psychotropic. A general rule of thumb is that the dose required to obtain response should be continued for at least 1 to 2 years. When discontinuing medication, there is a real risk of relapse, so this should be done very gradually indeed, with small adjustments every few months.

Patients who fail to respond to one SRI may respond to another. A range of medications has also been used to augment SRIs, with the largest evidence base comprising studies of low-dose dopamine receptor blockers in adult OCD. Some studies have indicated that these are particularly useful in patients with tics, but not all data are consistent. From a clinical perspective it is notable that when response to these agents is seen, it typically occurs within a relatively short time (i.e., 4 weeks). Despite the relatively low doses used in this setting, it is important to monitor adverse events. There is some indication that agents with increased D2 and D3 binding (e.g., aripiprazole) are more effective in treatment-refractory OCD, but this hypothesis requires further investigation.

For other potential SSRI augmentation agents, outcomes have been negative, inconsistent, or modest. Perhaps most promising are small studies which have indicated that certain glutamatergic agents, such as memantine, riluzole, ketamine, lamotrigine, and topiramate, may be useful in this condition; however, these remain proof-of-principal investigations. Sometimes low doses of clomipramine are used to augment SSRIs, but careful electrocardiographic monitoring is required. D-cycloserine is a partial glutamatergic agonist that has been used to augment psychotherapy in OCD and anxiety disorders, but efficacy data in OCD are currently inconsistent. Choice of medication in refractory patients is guided by a range of clinical considerations, including the presence of comorbid disorders. Monotherapy with irreversible monoamine oxide inhibitors remains a potentially useful option in refractory cases.

Psychotherapy

There is a long history of understanding obsessive-compulsive symptoms from a psychodynamic perspective. However, there is no rigorous evidence

that a psychodynamic approach is useful in the treatment of OCD. Instead, empirical research on the psychotherapy of OCD has emerged from behavioral and cognitive perspectives. Exposure and response prevention has proven to be a particularly effective component of treatment, with more recent work also focusing on the potential role for cognitive and mindfulness approaches.

To some extent the principles of exposure and response prevention can be readily applied by most clinicians, including primary care physicians involved in the diagnosis and treatment of OCD. The essential principle is to minimize avoidance of distressing cues, and instead to encourage in vivo exposure to feared situations and imaginal exposure to feared consequences. After such exposures, it is important not to engage in the usual compulsive response. A hierarchy of gradual exposure to more distressing cues is then developed and followed. Self-help volumes and internet programs have been developed to assist with the process. There may, however, be significant resistance to undertaking exposure and responsive prevention, in which case a skilled cognitive-behavioral therapist is needed. As in the case of pharmacotherapy, it is important to ensure that the dose and duration of exposure and response prevention treatment is sufficient.

More cognitive approaches to the psychotherapy of OCD have also been developed. These focus on thought distortions; for example, those related to inflated responsibility. Therapist and patient collaborate to test such distortions, and to replace them with more adaptive thinking patterns. In recent years, such approaches have also been supplemented by the use of techniques from mindfulness, although the evidence-base for these sorts of interventions is currently not as well developed. Family sessions may also be an important component of psychotherapy, particularly in younger patients; families are likely to benefit from a better understanding of OCD, and of the principles of exposure and response prevention.

Combined Treatment

Clinical guidelines indicate that either pharmacotherapy or psychotherapy is a reasonable first-line intervention for the management of OCD. In younger patients, it seems reasonable to begin CBT first. In patients with more severe symptoms, or comorbid depression, it seems reasonable to begin pharmacotherapy first. There is some evidence from rigorous comparisons of CBT and pharmacotherapy and their combination that combined behavior therapy and SRIs are particularly useful, including in children and adolescents. Further, one careful study in adults indicated that ERP augmentation of SSRIs was more efficacious than risperidone augmentation. In very refractory cases, partial hospitalization for combined treatment may also be useful. In everyday clinical practice judicious sequencing of interventions maybe used; for example, patients on

pharmacotherapy may be encouraged to decrease avoidance and increase exposure while gradually tapering dosage.

Somatic Therapies

Targeted neurosurgical lesions to CSTC tracts have been shown efficacious in a proportion of highly treatment-refractory patients. Initial work focused on stereotactic neurosurgical lesions, with subsequent work using radiosurgical techniques without the need for craniotomy. The more recent focus of neurosurgical work has been on deep brain stimulation, with implantation of electrodes to stimulate specific brain regions. A small number of centers around the world have the skills to do such work, and the FDA has approved targeting of the ventral striatum/ventral capsule. Transcranial magnetic stimulation has also been studied for OCD, although data are not yet sufficiently strong to motivate for routine clinical use. Electroconvulsive therapy is not thought to be efficacious for OCD.

BODY DYSMORPHIC DISORDER

Introduction

Although long described in the medical literature, BDD was rarely recognized by clinicians or studied by researchers, was only introduced into the official nomenclature in DSM-III-R, and will only be recognized for the first time as an independent disorder in the ICD system in ICD-11. Since its introduction to the nomenclature, there has been a substantial increase in the literature on this underrecognized condition. Recent decades have witnessed growing awareness of its prevalence and morbidity—embarrassment and shame about symptoms means that careful screening is needed to ensure case-finding. Fortunately, approaches to evaluation and treatment developed for OCD and appropriately modified for BDD, have led to substantial advances in outcomes.

Definition

BDD is characterized by persistent preoccupations about one or more perceived defects or flaws in one's appearance. The defects or flaws appear slight or are not observable to others. In response to their appearance concerns, individuals with BDD demonstrate a range of mental acts or behaviors, including comparing themselves with others, checking in the mirror, or camouflaging their perceived flaws.

History

The term dysmorphophobia was introduced more than a century ago by an Italian psychiatrist, Morselli. It derived from the word "dysmorphia," meaning misshapenness or ugliness. Subsequently both Janet and Freud

described cases with BDD symptoms, conceptualizing them as obsessive-compulsive conditions. The introduction of BDD into the official nomenclature in 1987 helped to increase clinical awareness of this condition as an independent entity, and the question of how similar BDD was to OCD—which had been found to respond to selective pharmacotherapies and psychotherapies—gave impetus to ongoing research.

Comparative Nosology

DSM-5 made a number of changes from DSM-IV. Wording was altered to be more clear and perhaps less pejorative to patients. An additional diagnostic criterion was added to address the repetitive mental acts or behaviors seen in BDD. Importantly, an insight specifier was added, including absent insight, so that patients with delusional BDD are now classified as having an OCRD rather than a psychotic disorder. In ICD-10, BDD is collapsed together with hypochondriacal disorder. The proposal for ICD-11 is to follow DSM-5 in recognizing BDD as a separate disorder in the chapter on OCRDs.

Epidemiology

BDD has rarely been included in major nation-wide community surveys of the prevalence of mental disorders. However, point prevalence estimates in the United States and elsewhere have ranged from 1.7 to 2.4 percent; indicating that this is one of the more common OCRDs. Furthermore, clinical studies have found high rates of BDD in general adult psychiatric inpatients, in cosmetic surgery clinics, and in dermatology clinics.

Community and clinical studies confirm that BDD is associated with significant morbidity and comorbidity. Patients are often highly impaired, with very poor quality of life. Compared with OCD, there are high rates of depression and suicidality, further contributing to the burden of disease associated with this condition. Mean age of onset is in adolescence, and as in the case of OCD, community studies indicate a higher prevalence of females with BDD, while clinical reports indicate a roughly even male:female ratio.

Etiology

There is some evidence that BDD is characterized by specific cognitive-affective impairments; for example, neuropsychological studies have suggested deficits in executive functioning and in visual processing. Brain imaging and neurogenetic studies have provided some insights into the potential mechanisms that underlie such deficits, and into potential overlaps with other OCRDs. At the same time, the current evidence base remains relatively thin, and the etiology of BDD is still far from clear.

The few structural and functional brain imaging studies in BDD have tentatively suggested some involvement of the CSTC circuits that appear to play a role in OCD. At the same time, other neurocircuitry may also be relevant; for example, that involved in visual processing. Family and twin data provide preliminary evidence for genetic susceptibility, as well as a relationship to OCD. There are however a paucity of data on specific genetic variants underlying susceptibility to BDD. While childhood adversity, including history of teasing, has been associated with BDD, it is not clear this is association is a specific one.

Diagnosis and Clinical Features

As in the case of OCD, the diagnosis of BDD is made on the basis of a rigorous psychiatric history and examination. The DSM-5 diagnostic criteria for BDD address the presence of appearance preoccupations, repetitive behaviors or mental acts in response to these preoccupations, the consequent distress and impairment, and the points that such symptoms are not due to other medical conditions or other psychiatric disorders. It is important for clinicians to also assess avoidance due to BDD symptoms; this can range from minor social avoidance to being housebound. Again, key features to ascertain during the history and examination include the age of onset and course of the symptoms, the type and severity of symptoms, the extent of insight and impairment, and the presence of comorbid disorders and symptoms.

Patients incorrectly perceive themselves as looking ugly or even grotesque, when in fact they look normal or attractive. Symptoms may focus on any area of the body, but they most commonly focus on the face and head, with concerns most often involving the skin, nose shape and size, and hair. Individuals may be concerned with one feature or with many; over the course of their illness patients are preoccupied with about five to seven different body areas. More than a quarter of patients are concerned with the symmetry of their appearance.

A majority of individuals with BDD have ideas of reference, falsely believing that people are staring at or talking about their deformities. BDD by proxy has rarely been described in the literature, that is, preoccupation with the appearance of others, such as a family member.

Most patients attempt to camouflage their perceived flaws. Common repetitive behaviors in BDD include comparing body parts with the same areas in others, checking body parts in the mirror or other reflective surfaces, grooming (including applying make-up, hair-pulling and skin-picking), seeking reassurance, touching body parts to check them, changing clothes, dieting, tanning, exercising, lifting weights, and undergoing cosmetic procedures. Such symptoms can involve serious harm to self (e.g., skin-picking can lead to significant medical complications) and others (e.g., in one survey, 40 percent of cosmetic surgeons reported that a dissatisfied

BDD patient had threatened them legally and/or physically, and occasional homicide of such surgeons has been reported).

Patients with BDD have less insight into the truth of their beliefs than those with OCD, and a first DSM-5 specifier for BDD addresses level of insight. Only a quarter of BDD patients have fair or good insight, and around one-third have absent insight. Patients with delusional BDD have greater symptom severity and greater morbidity. Individuals with BDD sometimes focus on whether their body build is too small or not sufficiently muscular; such so-called muscle dysmorphia is a second DSM-5 BDD specifier. Patients with muscle dysmorphia mostly look normal, though some may be unusually muscular as a result of body-building. They are more likely to be male, and are more likely to have abnormal eating behavior (e.g., eating high-protein meals) and to abuse substances (e.g., anabolic steroids) related to their appearance concerns. Compared to other patients with BDD, they may also be more suicidal.

Of the various OCRDs, it appears that BDD is the one most associated with comorbid depression, with a lifetime rate of 75 percent, as well as with suicidality, with 80 percent having experienced suicidal ideation. About one-third of BDD patients have a lifetime history of OCD, and around 30 percent of BDD patients experience panic attacks triggered by appearance concerns. BDD is also associated with high levels of rejection-sensitivity and low self-esteem. Many patients with BDD use substances to self-medicate symptoms of social anxiety or emotional pain. Several measures are now available to assess BDD symptoms and their severity (see www.bodyimageprogram.com). The Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) has been adapted for use in BDD, and the Brown Assessment of Beliefs Scale can be used to assess insight. Useful screening question for BDD include "Are you very worried about your appearance in any way?" and "Are you unhappy with how you look?"

A majority of BDD cases have onset before 18; in children and adolescents it is important to assess the impact of BDD symptoms on developmental trajectories and family functioning. Males and females have similar presentations, although males are more likely to be preoccupied with thinning hair, small penis size, and small body build, while females are more likely to be preoccupied with weight, breasts, hips, legs, and excessive body hair. Women are more likely to camouflage their body, to check mirrors, and to pick their skin. BDD symptoms appear universal, although the precise range of appearance concerns may be influenced by local contexts. In the East, a variant of taijin kyofusho, termed shubo-kyofu, is characterized by a fear that one's ugliness will offend others.

Pathology and Laboratory Features

At the current time, there are no biomarkers for BDD. Brain imaging may ultimately be clinically useful in predicting response to treatment, but at

the present time it remains a research tool. Similarly, while neurogenetic investigations have been valuable in the research context, they do not have sufficient clinical utility for more widespread use.

Differential Diagnosis

The specific focus of BDD on appearance concerns serves to differentiate it from other conditions characterized by repetitive behaviors (e.g., OCD, TTM, excoriation disorder), anxiety or social avoidance (e.g., generalized anxiety disorder, panic disorder, social anxiety disorder, agoraphobia), depression (e.g., major depressive disorder), or delusional beliefs (e.g., psychotic disorder). For example, some patients with BDD have hair-pulling and many have skin-picking; but in BDD such behaviors are in response to concerns about body flaws, while in TTM and excoration disorder, this is not the case.

Differentiating BDD from eating disorder may require care; if the patient's appearance concerns focus only on body fat or weight, and symptoms meet criteria for an eating disorder, then an eating disorder should be diagnosed. That said, BDD and eating disorders are often comorbid. BDD should not be diagnosed when an individual is only concerned with his or her genitals and/or secondary sex characteristics, and when symptoms meet criteria for gender dysphoria.

Olfactory reference syndrome is characterized by repetitive concerns about having a foul odor, a concern which leads to a range of repetitive behaviors such as washing or changing one's clothes. In comparison to OCD and BDD, those with ORS are even more likely to demonstrate poor insight or no insight. Nevertheless, it may be useful to classify ORS as an OCRD, given that some patients with ORS appear to respond to treatment with SSRIs.

Illness anxiety disorder (in DSM-5) and hypochondriasis (in ICD-11) are characterized by repetitive thoughts about being ill, rather than about body flaws. As in the case of BDD, however, there are compulsive responses to the somatic preoccupations (e.g., checking, visiting physicians), and treatment approaches overlap to some extent with those used for OCD.

Course and Prognosis

The course of BDD is generally chronic, and in the absence of intervention symptoms may persist for decades. Earlier age of onset and more severe symptoms at intake may predict a worse course.

Treatment

Psychoeducation is a key component of the treatment of BDD. Exploring the patient's explanatory model is a useful first step in progressing key elements of treatment such as understanding the patient's suffering,

building a collaborative approach to care, and instilling hope. Basic facts, such as the frequent response of symptoms to treatment, may bring some relief. While patients cannot simply be talked out of their symptoms, key facts, such as that cosmetic procedures may exacerbate symptoms, may need to be presented. Engaging patients at the start of treatment is particularly important and difficult when there are low levels of insight. Given high rates of suicidality in BDD, evaluating and monitoring such symptoms is key throughout treatment.

Pharmacotherapy

Clomipramine is more effective than desipramine in BDD, and limited but rigorous data from randomized controlled trials indicate that SSRIs are efficacious in the treatment of BDD, including delusional BDD. There is also some limited evidence for the efficacy of these agents in children and adolescents with BDD. Unfortunately the pharmaceutical industry has not undertaken registration trials.

There is some evidence that as in the case of OCD doses higher than those ordinarily used in the treatment of depression are beneficial in BDD, and that longer duration of treatment is needed in order to see the full effects of treatment. Again, duration of treatment should be longer rather than shorter, with gradual tapering of dosage and close monitoring of emergent symptoms.

Patients who do not respond to one SRI may respond to another. There is anecdotal evidence that buspirone augmentation of SRIs is useful in partial-responders. The very limited controlled data on augmentation of SRIs with low-dose dopamine receptor blockers are not positive, but in clinical practice, strategies used for treatment-resistant OCD, including addition of antipsychotics such as aripiprazole, and monotherapy with irreversible monoamine oxidase inhibitors are employed.

Psychotherapy

CBT for OCD has been modified for BDD, and studies have indicated efficacy, including for individuals with delusional BDD. Core elements of such treatments include exposure, response prevention, and cognitive restructuring. As in the case of OCD, dose and duration of treatment should be sufficient; some have suggested weekly CBT sessions for at least 6 months, with daily homework.

Modifications of OCD manuals are necessary to address the substantial social avoidance often present in BDD (e.g., use of behavioral experiments to test accuracy of beliefs about others), and to target visual processing abnormalities (e.g., use of perceptual retraining to develop a more holistic view of appearance). Further, given that BDD is accompanied by ideas of reference, and often by absent insight, motivational interviewing may be

needed when ambivalence about treatment is expressed, and cognitive interventions may need to be more complex and more intensive than in OCD. Habit reversal may be needed for BDD-related skin-picking, hair-pulling, and body touching.

Combined Treatment

There are no data from rigorous trials comparing pharmacotherapy and psychotherapy in BDD. Nevertheless, in the clinical setting it would seem sensible to make judicious use of both. Thus, for example, pharmacotherapy may be useful at the start of treatment, for example, if patients are too depressed to undergo CBT for BDD, and psychotherapy may be useful while preparations are made to very gradually taper medication dosage over the long term.

HOARDING DISORDER

Introduction

Hoarding symptoms have long been described in the medical and psychiatric literature, but HD is a recent entry into the official nomenclature, appearing for the first time in DSM-5. This will hopefully give further impetus to research on this condition, which like so many other OCRDs is prevalent and impairing, but underdiagnosed and undertreated.

Definition

HD is characterized by the persistent and profound difficulty discarding with or parting from one's possessions. This in turn leads to significant congestion and clutter, and to substantive distress or impairment, so differentiating normal collecting from the clinical disorder.

History

Freud described hoarding symptoms as a component of OCPD. In the DSM era, clinicians interested in OCD typically viewed hoarding as a subtype of OCD. During the DSM-5 process, a review of the growing literature on hoarding symptoms emphasized that many patients with clinical hoarding did not meet diagnostic criteria for either OCPD or OCD, and noted important differences between OCD and hoarding in phenomenology and psychobiology, and so concluded that HD should be included in the nomenclature as a separate diagnostic entity. Hoarding symptoms have been shown to be a predictor of worse response to treatment for OCD, and one hope is that clearer diagnostic categories will help lead to better outcomes for patients.

Comparative Nosology

Hoarding is mentioned in DSM-IV in the section on OCPD, but is not described in the equivalent ICD-10 category of anankastic personality disorder. DSM-IV also noted that when differentiating OCD from OCPD, a diagnosis of OCD should be considered when hoarding is extreme. DSM-5 for the first time provides diagnostic criteria for HD, and includes this condition in the OCRD section. Similarly, the proposals for ICD-11 are to include HD as a separate entity. There are some distinctions between DSM-5 and the ICD-11 proposal, for example, ICD-11 emphasizes that excessive acquisition is seen in most patients with this condition.

Epidemiology

DSM-5 diagnostic criteria have not been widely employed. However, preliminary data indicate that HD is remarkably prevalent in Western countries, with one study that employed DSM-5 criteria yielding a point prevalence of 1.5 percent. Epidemiological data indicate that HD is associated with significant comorbidity and morbidity. In a study using DSM-5 criteria, male:female ratio was about 1. In contrast, in clinics, females are more common than males, perhaps indicating greater insight. Unlike the picture seen in most OCRDs, prevalence of hoarding increases with age, and in clinical settings the average age of first treatment is around 50. As in the case of most OCRDs, much further work is needed in low- and middle-income countries.

Etiology

Patients with HD may demonstrate impairment in various neuropsychological domains including spatial planning, working memory, response inhibition, and set-shifting. Structural and functional brain imaging studies have indicated that the relevant neurocircuitry overlaps only partially with that involved in OCD. Research on patients with neurological disorders has suggested that ventromedial prefrontal/anterior cingulate cortices as well as medial temporal regions are involved in hoarding.

Animal studies have pointed to a possible role for the dopaminergic system in hoarding. Twin studies indicate genetic susceptibility for developing pathological hoarding. While there is a relative dearth of neurogenetic studies, GWAS data from OCD studies provide a preliminary indication that specific areas of the genome are involved in mediating hoarding symptoms.

Diagnosis and Clinical Features

The diagnosis of HD is made on the basis of a rigorous psychiatric history and examination. DSM-5 diagnostic criteria for HD emphasize that

individuals have difficulty discarding or parting with possessions, regardless of their actual value. This difficulty is due to a perceived need to save the items and to distress associated with discarding them, and results in congestion and clutter so that the use of living areas for their intended use is severely compromised. If living areas are uncluttered it is only because others step in to assist, for example, friends or family members may agree to store items in their own homes. Hoarding is associated with clinically significant distress or impairment, and is not attributable to another medical condition or another mental disorder.

Key features to ascertain during the history and examination include the age of onset and course of the symptoms, the type and severity of symptoms, the extent of insight and impairment, and the presence of comorbid disorders and symptoms. A range of objects may be hoarded, but they are usually everyday items (e.g., newspapers, books, clothing), and the resulting clutter is usually highly disorganized. There is a literature on animal hoarding, which some conceptualize as a subtype of HD. Patients cite various reasons for hoarding including utilitarian value, sentimental attachment, concerns about losing information, and not being wasteful. Exposure to traumatic or stressful events may precede onset or exacerbation of hoarding in some cases.

Hoarding may be associated with a broad range of sequelae including health and safety problems (e.g., patients may not have sufficient space to cook or keep clean, and so may be at increased risk for a broad range of illnesses). “Clutter avalanches” may lead to falls and even mortality. In some individuals, conditions become entirely unsanitary; this is termed “domestic squalor.” Family relationships can be negatively impacted, and in some cases dependents may need to be removed. Threats of eviction or forced clearing of possessions may lead to substantial distress and even suicide.

DSM-5 provides two specifiers for HD. Excessive acquisition is specified if difficulty discarding possessions is accompanied by excessive acquisition of items that are not needed or for which there is no available space. As in OCD and BDD, insight can be specified as good or fair, poor, or absent/delusional. However, the definition is framed somewhat differently from that in OCD and BDD; in absent/delusional insight, the individual is completely convinced that hoarding-related beliefs and behaviors are not problematic despite evidence to the contrary. History from other informants may be key for accurate evaluation of symptoms when insight is absent.

Commonly associated features of HD include indecisiveness and procrastination, difficulty planning and organizing tasks, and distractibility and avoidance. Comorbid OCD is elevated in HD compared with the general population. Prevalent comorbid disorders in HD include generalized anxiety disorder and major depressive disorder. Anecdotal

experience suggests that other OCRDs may also be more commonly comorbid than expected by chance. In elderly patients with HD, there are higher rates of medical complications of HD compared with age-matched control subjects. A number of different measures, including the Saving Inventory-Revised and the Hoarding Rating Scale, can be used to assess the severity of hoarding symptoms. Home visits, or home photographs, are helpful for accurate assessment of symptoms.

Pathology and Laboratory Examination

There are no biomarkers for HD.

Differential Diagnosis

Differentiating normal collecting from HD is not typically problematic. In normal collecting, accumulation is confined to a narrow range of objects, acquired in a structured way, and then well organized. Furthermore, collecting is a pleasurable activity, that does not lead to impairment. In contrast, pathological hoarding is characterized by accumulation of a range of objects, unstructured or impulsive acquisition, and haphazard clutter, as well as consequent distress and impairment.

One key differential diagnosis to consider is that of OCD. In OCD hoarding is undertaken in response to an unwanted obsession; most commonly as a result of symmetry obsessions, feelings of incompleteness, or fears of harm or contamination (e.g., an object may be kept in case discarding that object leads to harm). In contrast in HD, there is no associated obsession; rather the individual wishes to retain saved items and is distressed by the thought of discarding them. In some cases both OCD and comorbid HD may be diagnosed.

Although the DSM-5 proposal was for OCPD to remove the hoarding criterion, no changes to the DSM-IV diagnostic criteria for personality disorders were made. Hoarding symptoms may also be seen in a range of other psychiatric disorders, including autism spectrum disorders (where collections may reflect a special interest) and psychotic disorders. HD should only be diagnosed if the hoarding symptoms are thought to be independent; that is, when they have a different onset and course.

A number of neurodevelopment and neurodegenerative conditions are associated with hoarding, for example, Prader–Willi Syndrome (where individuals are characterized by hoarding of food, obesity, and skin-picking) and Alzheimer's. Damage to particular brain regions, anterior ventromedial prefrontal and cingulate cortices has been associated with excessive and indiscriminate hoarding activity. The saving of "bizarre" items (e.g., feces, rotten food), or of severe squalor, may point to a diagnosis other than HD.

Course and Prognosis

HD has a somewhat later onset than most OCRDs. In the absence of intervention, its course appears to be often chronic and progressive. Hoarding symptoms often begin in childhood and adolescence, with diagnostic criteria met only in the 30s, and worsening of symptoms in each subsequent decade.

Treatment

Pharmacotherapy. The literature on pharmacotherapy of OCD has indicated that hoarding symptoms are a predictor of less favorable outcomes to treatment with SSRIs. Nevertheless, these agents remain the first-line intervention for HD in clinical practice, and uncontrolled data indicate that these agents as well as venlafaxine may be efficacious for this indication. The number of studies on the pharmacotherapy of HD is relatively small, but it is interesting that some case reports have noted a response to glutamate modulators.

Psychotherapy. CBT has been adapted for HD, and research to date indicates that this manualized treatment is superior to wait-list. Components include psychoeducation and goal setting, organization skills training, practice with resisting acquisition, and cognitive restructuring. Given that many patients have limited insight, motivational interview techniques are needed to engage patients in treatment. Involving family members and social services may be key, particularly when insight is absent.

Combined Treatment. There are no rigorous studies of combined treatment in HD, and clinical judgment is therefore used to determine the optimal sequencing of different modalities.

TRICHOTILLOMANIA (HAIR-PULLING DISORDER)

Introduction

The term TTM was coined by a French dermatologist, Hallopeau, in the 19th century. In the early part of the 20th century, a large case series of patients was described by two surgeons, who operated on individuals with trichobezoar (due to swallowing of hair after hair-pulling). Attention was directed to TTM when researchers discovered that like OCD, TTM responded selectively to clomipramine. Interest in the disorder further increased when consumer advocacy organizations rose in response to the demand for appropriate recognition and efficacious interventions for this underrecognized and undertreated condition.

Definition

TTM is characterized by pulling out of one's hair. This is not done for cosmetic reasons, but rather individuals often describe an irresistible urge to pull out their hair.

History

Renting of one's hair has long been described in the medical literature as a symptom of psychopathology, with descriptions dating at least as far back as Hippocrates. Hallopeau's term derives from the Greek for hair (trich) and pull (till). However, TTM is a relatively late entrant to the nomenclature; it was included for the first time in DSM-III-R. In DSM-5, TTM was for the first time characterized as an OCRD in view of some of its overlapping features with OCD. In ICD-11, TTM and excoriation (skin-picking) disorder may be classified as body focused repetitive behavior disorders (BFRBDs), giving appropriate recognition to this overlooked set of conditions.

Comparative Nosology. In DSM-IV, TTM was classified as an impulse control disorder not otherwise classified. Accordingly, diagnostic criteria include a description of rising tension prior to hair-pulling, followed by gratification during hair-pulling. *DSM-5 reviewers* noted that many patients with TTM did not meet diagnostic criteria for these two phenomena, and that patients who did and did not meet these diagnostic criteria did not demonstrate clinically important differences, and so replaced this with a criterion referring to the wish to stop hair-pulling.

Epidemiology. Studies of select populations, such as college students, indicate that noncosmetic hair-pulling has a prevalence of 10 percent or higher, with 1 to 3 percent meeting DSM-5 diagnostic criteria for TTM. Age of onset is typically at menarche, and female:male ratio is about 9:1. Pediatric samples include more girls, but with a less skewed gender ratio. There is also growing evidence of the comorbidity and morbidity associated with TTM; this includes comorbid psychiatric disorders, various medical sequelae (e.g., trichobezoars), and decreased quality of life.

Etiology. From a cognitive-affective neuroscience perspective, neuropsychological investigations of TTM have demonstrated some evidence of impairments, including deficits in working memory and visual-spatial learning. Structural and functional brain imaging studies have suggested some involvement of CSTC circuits relevant to habit learning, although the database of studies is sparse. There is also evidence of involvement of other regions such as those involved in reward processing and affect regulation, processes which may well be relevant to TTM.

Animal models of TTM include "barbering," and have investigated the role of various molecular systems in this behavior. Family and twin studies

provide evidence for genetic susceptibility, as well as a relationship to OCD and other BFRBDs. A number of interesting candidate genes have preliminarily been associated with TTM, including SAPAP3 and SLITRK1. There is evidence of increased childhood trauma in TTM, but this association is clearly not specific to this particular OCRD.

Diagnosis and Clinical Features

The diagnosis of TTM is made on the basis of a rigorous psychiatric history and examination. DSM-5 criteria refer to hair-pulling resulting in noticeable hair lose, repeated attempts to decrease or stop hair-pulling, clinically significant distress or impairment, with the hair-pulling or loss not attributable to another medical condition, nor better explained by the symptoms of another medical disorder. Key features to ascertain during the history and examination include the age of onset and course of the symptoms, the type of symptoms, the extent of impairment, and the presence of comorbid disorders and symptoms.

Hair-pulling may be from any area of the body, but is most commonly from scalp, eyebrows, and eyelashes. It is worth inquiring about pubic hair-pulling as this may be more embarrassing for patients to volunteer. Patients often pull hair daily, either more continuously throughout the day, or in fewer isolated episodes. Most patients pull with their fingers, but some use tweezers or other instruments. Patients may have a range of habits associated with the hair-pulling, for example, stroking or playing with hair, choosing a particular kind of hair to pull, and mouthing the hair once it is pulled. It is also important to inquire about trichophagy, partly because of associated shame, and partly because it may have serious medical sequelae. Some individuals pull hair from other dolls or other objects, from pets, or from other people.

Hair-pulling may be triggered by states of both hyperarousal (e.g., stress) and hypoarousal (e.g., boredom). It has been suggested that there are two major kinds of hair-pulling; focused and automatic. In focused hair-pulling, the person is aware of an urge to pull, resists it, and then gives in. In automatic hair-pulling, the person has less awareness of the hair-pulling, and before they realize it, have pulled a number of hairs. Most people with TTM report both kinds of hair-pulling, but individuals differ in the extent to which they employ focused or automatic hair-pulling. Hair-pulling may occasionally be seen in infants and toddlers; this is not well studied, but may be a self-limiting condition.

TTM is associated with substantial morbidity and comorbidity. First, TTM is associated with significant emotional distress. Repeated failures to control hair-pulling may lead to shame, frustration, and decreased self-esteem. Time-consuming camouflage and avoidance of situations or activities where hair loss may be discovered may further contribute to impairment. Comorbid conditions in TTM include other BFRBDs, with

excoriation (skin-picking) disorder the most common. OCD is also more prevalent in TTM than in the general population. More than half of treatment-seeking individuals with TTM have a comorbid psychiatric disorder, with mood and anxiety disorders the most common. A range of measures are now available to assess hair-pulling severity and type, including the Massachusetts General Hospital Hairpulling Scale, and the Milwaukee Inventory for Subtypes of Trichotillomania. Photographs of pulled areas may also be useful to assess progress during treatment.

Hair-pulling symptoms are somewhat similar across the life course although younger patients may be less likely to report urges or arousal prior to pulling, or gratification and relief after pulling. Females and males also have similar symptoms, but there is some evidence that males have slight differences in hair-pulling patterns (e.g., pulling more commonly from face and chest) and are more likely to have comorbid tics. While there are limited data on TTM from different countries and cultures, that which exists suggest that symptoms seem fairly similar across the world.

Pathology and Laboratory Examination. The majority of patients with TTM are able to describe their symptoms. In cases of alopecia, where hair-pulling is denied and the diagnosis is unclear, trichoscopy may be useful.

Differential Diagnosis. TTM should be differentiated from hair-pulling for cosmetic or cultural reasons, and from other repetitive behaviors that involve the hair but that do not lead to pulling. In TTM, there is no associated intrusive obsession or appearance concern, so differentiating it from hair-pulling which is due to OCD (with symmetry concerns) or BDD (with body appearance concerns). Similarly, on the rare occasion that hair-pulling is in response to a psychotic symptom or to a substance, TTM should not be diagnosed. Alopecia may occur for a range of reasons other than TTM.

Course and Prognosis. There are relatively few data on the long-term course of TTM. In the absence of intervention, its course is likely often chronic. However, some data indicate that in a proportion of cases remission does occur.

Treatment. Psychoeducation is a key component of the treatment of TTM. As in other OCRDs, exploring the patient's explanatory model is a useful first step in progressing key elements of treatment such understanding the patient's suffering, building a collaborative approach to care, and instilling hope. Many patients and families are relieved to learn basic facts about TTM, including its high prevalence, the fact that many others are embarrassed by/ashamed of symptoms, that it is not due to commonly voiced factors (e.g., bad mothering), and that treatments are

efficacious for many. Consumer advocacy organizations, such as the Trichotillomania Learning Centre (TLC) Foundation for Body-Focused Repetitive Behaviors, may play a particularly important role in conveying such information.

Pharmacotherapy. Early data indicated that TTM, like OCD and BDD, responded more robustly to clomipramine than to desipramine. Unfortunately subsequent trials of SSRIs and of venlafaxine in TTM have not consistently demonstrated efficacy. The largest positive clinical trial to date is with N-acetylcysteine (NAC) (1,200 to 2,400 mg/day), a nutraceutical that acts on the glutamatergic system. However, a trial of NAC in pediatric TTM was negative. There are small controlled trials indicating that dopamine receptor blockers may also be useful in these conditions. Given its favorable side-effect profile, it may be reasonable to use NAC as a first-line agent in adults with TTM, to use SSRIs when patients have comorbid conditions for which these agents are efficacious, and to consider low dose dopamine receptor blockers in more treatment-refractory cases.

Psychotherapy. Habit reversal therapy (HRT) is a set of cognitive-behavioral techniques that has been shown efficacious in the treatment of childhood and adult TTM. Components of HRT include awareness training, competing response training, and social support. Awareness training includes elements such as self-monitoring and collecting pulled hair. In competing response training, the patient substitutes hair-pulling with an incompatible action (e.g., first-clenching). Social support involves the use of another individual to reinforce the competing response, and may be particularly useful in pediatric TTM. In addition, stimulus control (SC) is used to change the patient's environment so that hair-pulling is more effortful and/or less reinforcing. There is also preliminary evidence that web-based administration of HRT/SC may be efficacious. Finally, controlled trials of augmentation of HRC/SC with acceptance and commitment therapy, dialectical behavior therapy, or cognitive therapy, which focus on strategies such as affect regulation, enhanced motivation, relapse prevention, or cognitive restructuring, have also suggested efficacy in TTM.

Combined Treatment. There is very preliminary evidence that in some individuals SSRIs in combination with psychotherapy are more beneficial than either treatment alone. From a clinical perspective, judicious sequencing of different treatment modalities may be used.

EXCORIATION (SKIN-PICKING) DISORDER

Introduction

Cases of "neurotic excoriation" or "psychogenic excoriation" have long been described in the medical literature. However, excoriation (skin-picking) disorder is new to DSM-5. This debut reflects a gradually accumulating research literature, and hopefully the inclusion of skin-picking disorder (SPD) in the nomenclature will give impetus to even more investigation, as well bring this often overlooked condition to the attention of clinicians. In ICD-11 the proposal is for TTM and SPD to be listed together under the umbrella of BFRBDs, given their many similarities.

Definition

SPD is characterized by the recurrent skin-picking, resulting in skin lesions. Individuals with skin-pulling attempt to decrease or stop skin-picking, but are unable to do so. While some picking of scabs is likely universal, in the case of SPD, there is associated clinically significant distress or impairment.

History

This disorder was first described in detail by a French dermatologist, Brocq, who noted the development of skin-picking in young women with acne ("acne excoriee"). The field of psychodermatology was influenced by psychodynamic constructs, as evidenced by the subsequent term "neurotic excoriation." In 1978, Griesemer, who trained in both dermatology and psychiatry, reported that 2 percent of dermatology patients had SPD. Early work with the SSRIs suggested that these agents were useful for SPD, and productive questions arose about the relationship between OCD, TTM, and SPD. More recently, consumer advocacy associations have further raised awareness of TTM, SPD, and their frequent cooccurrence.

Comparative Nosology. In DSM-IV, SPD is mentioned briefly as an "impulse control disorder not otherwise specified" in the chapter on "impulse control disorders not elsewhere classified." SPD is classified in the nomenclature for the first time in DSM-5. This inclusion recognized the increasing scientific evidence for the diagnostic validity and clinical utility of this entity, as well as the voice of patients and patient advocates. Diagnostic criteria in DSM-5 closely reflect those for TTM, consistent with the many phenomenological similarities between these conditions. The proposal for ICD-11 is very similar, but in that system both TTM and SPD may fall under the umbrella of BFRBDs. The most common of these conditions are TTM and SPD, but they include a range of other conditions with repetitive behaviors that damage the integument; including, for example, severe nail-biting, lip-chewing, and cheek chewing.

Epidemiology. Community-based prevalence studies of SPD have

found a point prevalence of 1.4 to 5.4 percent. Age of onset is somewhat more variable than TTM, although mean age is 12 years, typically coinciding with the onset of puberty. Although many more females than males have SPD, the ratio is more even than in the case of TTM. There is also growing evidence of the morbidity and comorbidity associated with SPD; this includes comorbid psychiatric disorders, various medical sequelae (e.g., infection), and decreased quality of life.

Etiology. From a cognitive-affective neuroscience perspective, neuropsychological investigations of SPD have demonstrated some overlap with OCD in that there is evidence of motor impulsivity. Brain imaging studies have similarly suggested some involvement of CSTC circuitry, although the database of investigations is sparse, and other regions also appear to be implicated.

Animal data, the finding that skin-picking is precipitated by dopaminergic stimulants, and response to dopamine receptor blockers suggests that the dopaminergic system plays a role in SPD. Family and twin studies provide evidence for genetic susceptibility, as well as a relationship with OCD and other BFRBDs. A number of interesting candidate genes have preliminarily been associated with SPD, including SAPAP3.

Diagnosis and Clinical Features. The diagnosis of SPD is made on the basis of a rigorous psychiatric history and examination. DSM-5 diagnostic criteria for SPD refer to skin-picking resulting in skin lesions, repeated attempts to decrease or stop skin-picking, clinically significant distress or impairment, with the skin-picking not attributable to another medical condition, nor better explained by the symptoms of another medical disorder. Key features to ascertain during the history and examination include the age of onset and course of the symptoms, the type of symptoms, the extent of impairment, and the presence of comorbid disorders and symptoms.

Skin-picking may be from any area of the body, but is most commonly from the face, followed by the hands, fingers, arms, and legs. Patients may pick from multiple sites; one study found that patients picked from an average of 4.5 sites. Patients may have a range of habits associated with the skin-picking, for example, finding a particular scab, picking at it in a specific way, and mouthing or swallowing the scab once it is picked. Most patients pick with their fingernails, but some use tweezers or other instruments. Patients often skin pick daily, either more continuously throughout the day, or in fewer isolated episodes. Patients may have a range of habits associated with the skin-picking, for example, stroking or playing with the skin, choosing a particular scab to pick, and mouthing the scab once it is pulled. Some individuals pick skin from other people.

Patients may report that skin-picking began after a dermatological

condition such as acne, but continued even with the resolution of that condition. Skin-picking may be triggered by states of both hyperarousal (e.g., stress) and hypoarousal (e.g., boredom). It has been suggested that there are two major kinds of skin-picking; focused and automatic. In focused skin-picking, the person is aware of an urge to pull, resists, it and then gives in. In automatic skin-picking, the person has less awareness of the hair-pulling, and before they realize it, has pulled a number of hairs. Most people with SPD report both kinds of skin-picking, but individuals differ in the extent to which they employ focused or automatic hair-pulling.

SPD is associated with substantial morbidity and comorbidity. First, SPD is associated with significant emotional distress. Repeated failures to control skin-picking may lead to shame, depression, and self-medication. Time-consuming camouflage and avoidance of situations or activities where skin lesions may be discovered may further contribute to impairment. Comorbid conditions in SPD include other BFRBDs, with TTM the most common. OCD and BDD are also more prevalent in SPD than in the general population. Other common comorbid psychiatric disorders in SPD include mood and anxiety disorders. A range of measures are available to assess hair-pulling severity and type, including the Skin Picking Impact Scale. A thorough physical examination may be useful to assess the extent and severity of skin-picking, and photographs of skin lesions may be helpful in assessing progress.

Skin-picking symptoms are somewhat similar across the life course, although further work is needed in this area. In one study, males had a later age of onset, more severe symptoms, and greater impairment than females with SPD. While there are limited data on SPD from different countries and cultures, that which exists suggest that symptoms seem fairly similar across the world.

Pathology and Laboratory Examination. The majority of patients with SPD are able to describe their symptoms, and dermatological examination reveals classical features. Dermatopathology of SPD also shows unique features.

Differential Diagnosis. Skin-picking is quite common in BDD, but in these individuals it is done in response to concerns about body appearance, and SPD should not be diagnosed. Similarly, skin-picking may also occur in response to obsessions about contamination, delusions or tactile hallucinations related to infestation (parasitosis), or may be induced by dopaminergic psychostimulants (e.g., cocaine). Skin picking may also be due to a factitious disorder (dermatitis artefacta). Finally, skin-picking may be seen in response to a broad range of dermatological conditions associated with itching (such as scabies), as well as in a number of medical conditions (such as Prader–Willi syndrome).

Course and Prognosis. Those data which exist on the long-term course of SPD suggest that in the absence of intervention, this is often chronic, albeit with fluctuations in severity over time. As is the case in other OCRDs, individuals often report that they have not sought medical attention because they are embarrassed by their symptoms, because they feel they should stop the symptoms by themselves, or because they are unaware that the symptoms may be part of a known condition or that there are efficacious treatments.

Treatment. Psychoeducation is a key component of the treatment of SPD. As in other OCRDs, exploring the patient's explanatory model is a useful first step in progressing key elements of treatment such understanding the patient's suffering, building a collaborative approach to care, and instilling hope. Many patients and families are relieved to learn basic facts about SPD, including its high prevalence, the fact that many others are embarrassed by/ashamed of symptoms, that it is not due to commonly voiced factors (e.g., bad mothering), and that treatments are efficacious for many. Consumer advocacy organizations, such as the TLC Foundation for Body-Focused Repetitive Behaviors, may play a particularly important role in conveying such information.

Pharmacotherapy. There is some evidence from randomized controlled trials that SPD responds to SSRIs, although studies have been few and with small samples. There are data that NAC, a nutraceutical that acts on the glutamatergic system, and *low dose dopamine receptor blockers* may also be useful in SPD. Data on the efficacy of lamotrigine in SPD are inconsistent.

Psychotherapy. There is a good deal of overlap in psychotherapy for TTM and SPD. HRT has been shown efficacious in the treatment of childhood and adult SPD in comparison to wait-list controls. Components of HRT for SPD have included awareness training, competing response training, and social support. Awareness training includes elements such as self-monitoring. In competing response training, the patient substitutes skin-picking with an incompatible action (e.g., first-clenching). Social support involves the use of another individual to help reinforce the competing response. Another manualized CBT treatment for SPD has also included cognitive restructuring and relapse prevention. There is interest in acceptance-enhanced augmentation of HRT for SPD, but this has not yet been rigorously studied.

Combined Treatment. There are no few rigorous studies comparing pharmacotherapy and psychotherapy in SPD. From a clinical perspective, judicious sequencing of modalities may be used.

OCRD DUE TO ANOTHER MEDICAL DISORDER

This is a new category in DSM-5, reflecting the point that in a new chapter on OCRDs it is relevant to exclude conditions that are due to another medical disorder. To determine whether OCRD symptoms are etiologically linked to another medical disorder, consideration should be given to the presence of a clear temporal relationship between such symptoms and the medical disorder. As noted earlier, obsessive-compulsive symptoms have long been described in the context of various neurological lesions, most notable in the aftermath of the influenza epidemic in the early 20th century. At postmortem, patients had striatal lesions. More recent work has documented the occurrence of OCD due to a number of other striatal pathologies, including stroke and infection. For example, it is notable that a significant proportion of cases of Sydenham's chorea have obsessive-compulsive symptoms.

Pediatric neuropsychiatric disorders associated with Streptococcus (PANDAS) is a condition thought similar to Sydenham's chorea, in that a Streptococcal infection is hypothesized to trigger an autoimmune reaction, resulting in sudden onset of obsessive-compulsive and other symptoms. Rigorous research has demonstrated that cases of PANDAS respond to immunotherapeutic interventions. At the same time, this is a relatively new area of research, and some have argued that sudden-onset obsessive-compulsive symptoms can be seen after a broad range of insults.

SUBSTANCE/MEDICATION-INDUCED OCRD

There is some evidence that particular substances or medications can trigger or exacerbate obsessive-compulsive symptoms. Thus, for example, skin-picking may be seen in patients with cocaine intoxication. Such work is theoretically interesting insofar as it points to the potential role for the dopaminergic system in OCRDs. There is less evidence, however, that chronic symptoms of OCRDs are due to substances or medications.

OTHER SPECIFIED AND UNSPECIFIED OCRDS

DSM-5 notes a range of other specified and unspecified OCRDs. Other specified OCRDs include obsessional jealousy and cultural variants such as shubo-kyofu, which is characterized by repeated concerns about body appearance, and koro, which is characterized by concerns that the penis will recede into the body. Other specified OCRDs are also diagnosed when the presentation differs in important ways from those outlined in the chapter; for example, BDD symptoms in patients who have observable body flaws, or who do not have repetitive behaviors.

CONCLUSION

The new DSM-5 chapter on OCRDs aims to improve the assessment and

management of these underrecognized and undertreated conditions, which are highly prevalent and associated with a great deal of comorbidity and morbidity. ICD-11 proposes a similar chapter, but including a somewhat broader spectrum of conditions. Both DSM-5 and ICD-11 list OCRDs adjacent to the anxiety disorders, in order to emphasize the overlap between these groups of conditions. Any particular approach to the meta-structure of the nomenclature arguably has pros and cons; on the one hand the OCRD chapter highlights important phenomenological and psychobiological similarities across these conditions, but there is the risk that naïve clinicians may overlook important differences required in their evaluation and treatment. Certainly, when assessing an individual with any one OCRD it may well be useful to ask about the presence of other OCRDs in the individual and in family members, and there are some similarities in the approach to pharmacotherapy and psychotherapy across these conditions. At the same time, as in other DSM-5 and ICD-11 chapters, it is important emphasize that there are also important distinctions that must be made between these conditions, and consequent differences in current treatment approaches.

References

Angelakis I, Gooding P, Tarrier N, Panagioti M. Suicidality in obsessive compulsive disorder (OCD): a systematic review and meta-analysis. *Clin Psychol Rev*. 2015;39:1–15.

Bienvenu OJ, Samuels JF, Wuyek LA, et al. Is obsessive-compulsive disorder an anxiety disorder, and what, if any, are spectrum conditions? A family study perspective. *Psychol Med*. 2012;42:1–13.

Ducasse D, Boyer L, Michel P, et al. D2 and D3 dopamine receptor affinity predicts effectiveness of antipsychotic drugs in obsessive-compulsive disorders: a metaregression analysis. *Psychopharmacology*. 2014;231:3765–3770.

de Wit SJ, Alonso P, Schweren L, et al. Multicenter voxel-based morphometry mega-analysis of structural brain scans in obsessive-compulsive disorder. *Am J Psychiatry*. 2014;171:340–349.

Fineberg NA, Chamberlain SR, Goudriaan AE, et al. New developments in human neurocognition: clinical, genetic, and brain imaging correlates of impulsivity and compulsivity. *CNS spectrums*. 2014;19:69–89.

Fineberg NA, Reghunandanan S, Kolli S, Atmaca M. Obsessive-compulsive (anankastic) personality disorder: toward the ICD-11 classification. *Revista brasileira de psiquiatria*. 2014;36 Suppl 1:40–50.

Fontenelle LF, Mendlowicz MV, Versiani M. The descriptive epidemiology of obsessive-compulsive disorder. *Prog Neuropsychopharmacol Biol Psychiatry*. 2006;30:327–337.

Grant JE, Odlaug BL, Chamberlain SR, Keuthen NJ, Lochner C, Stein DJ. Skin picking disorder. *Am J Psychiatry*. 2012;169:1143–1149.

Grant JE, Stein DJ. Body-focused repetitive behavior disorders in ICD-11. *Revista brasileira de psiquiatria*. 2014;36 Suppl 1:59–64.

Heuvel OA, Veale D, Stein DJ. Hypochondriasis: considerations for ICD-11. *Revista brasileira de*

psiquiatria. 2014;36 Suppl 1:21–27.

Ipser JC, Carey P, Dhansay Y, Fakier N, Seedat S, Stein DJ. Pharmacotherapy augmentation strategies in treatment-resistant anxiety disorders. *The Cochrane database of systematic reviews*. 2006:CD005473.

Kariuki-Nyuthe C, Gomez-Mancilla B, Stein DJ. Obsessive compulsive disorder and the glutamatergic system. *Curr Opin Psychiatry*. 2014;27:32–37.

Leckman JF, Denys D, Simpson HB, et al. Obsessive-compulsive disorder: a review of the diagnostic criteria and possible subtypes and dimensional specifiers for DSM-V. *Depress Anxiety*. 2010;27:507–527.

Lewis-Fernandez R, Hinton DE, Laria AJ, et al. Culture and the anxiety disorders: recommendations for DSM-V. *Depress Anxiety*. 2010;27:212–229.

Lochner C, Fineberg NA, Zohar J, et al. Comorbidity in obsessive-compulsive disorder (OCD): a report from the International College of Obsessive-Compulsive Spectrum Disorders (ICOCS). *Compr Psychiatry*. 2014;55:1513–1519.

Macy AS, Theo JN, Kaufmann SC, et al. Quality of life in obsessive compulsive disorder. *CNS spectrums*. 2013;18:21–33.

Mataix-Cols D, Frost RO, Pertusa A, et al. Hoarding disorder: a new diagnosis for DSM-V? *Depress Anxiety*. 2010;27:556–572.

Phillips KA, Stein DJ, Rauch SL, et al. Should an obsessive-compulsive spectrum grouping of disorders be included in DSM-V? *Depress Anxiety*. 2010;27:528–555.

Phillips KA, Wilhelm S, Koran LM, et al. Body dysmorphic disorder: some key issues for DSM-V. *Depress Anxiety*. 2010;27:573–591.

Radua J, Grau M, van den Heuvel OA, et al. Multimodal Voxel-Based Meta-Analysis of White Matter Abnormalities in Obsessive-Compulsive Disorder. *Neuropsychopharmacology*. 2014;39(7):1547–1557.

Romanelli RJ, Wu FM, Gamba R, Mojtabai R, Segal JB. Behavioral therapy and serotonin reuptake inhibitor pharmacotherapy in the treatment of obsessive-compulsive disorder: a systematic review and meta-analysis of head-to-head randomized controlled trials. *Depress Anxiety*. 2014;31:641–652.

Ruscio AM, Stein DJ, Chiu WT, Kessler RC. The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular psychiatry*. 2010;15:53–63.

Russell EJ, Fawcett JM, Mazmanian D. Risk of obsessive-compulsive disorder in pregnant and postpartum women: a meta-analysis. *J Clin Psychiatry*. 2013;74:377–385.

Sharma E, Thennarasu K, Reddy YC. Long-term outcome of obsessive-compulsive disorder in adults: a meta-analysis. *J Clin Psychiatry*. 2014;75:1019–1027.

Singer HS, Gilbert DL, Wolf DS, Mink JW, Kurlan R. Moving from PANDAS to CANS. *The J Pediatr*. 2012;160(5):725–731.

Stein DJ, Fontenelle LF, Reed GM. Obsessive-compulsive and related disorders in ICD-11. *Revista brasileira de psiquiatria*. 2014;36 Suppl 1:1–2.

Stein DJ, Grant JE, Franklin ME, et al. Trichotillomania (hair pulling disorder), skin picking disorder, and stereotypic movement disorder: toward DSM-V. *Depress Anxiety*. 2010;27:611–626.

Strauss C, Hale L, Stobie B. A meta-analytic review of the relationship between family

accommodation and OCD symptom severity. *J Anxiety Disord*. 2015;33:95–102.

Taylor S. Molecular genetics of obsessive-compulsive disorder: a comprehensive meta-analysis of genetic association studies. *Mol Psychiatry*. 2013;18:799–805.

Veale D, Matsunaga H. Body dysmorphic disorder and olfactory reference disorder: proposals for ICD-11. *Revista brasileira de psiquiatria*. 2014;36 Suppl 1:14–20.

16

Pathological and Problem Gambling: Gambling Disorder

HARVEY ROY GREENBERG, M.D.

Pathological gambling has always been a characteristic of humans, a scourge in the shadows. Because of an unprecedented worldwide increase in gambling opportunities over the past 25 years, the illness can no longer be condemned nor ignored. It is a universal public health problem, often as crippling as schizophrenia, major affective illness, and the anxiety disorders.

Mental health professionals have become more adept at diagnosing pathological gambling. Unfortunately efforts at increasing the public awareness of its ravages lag significantly behind the attention paid to other psychiatric problems which do not bear the unwarranted attainder of sinfulness.

One cannot precisely gauge the take from legitimate gambling in the United States. A reasonable guesstimate is approximately 150 billion dollars. The annual profit from American licit and illicit gambling enterprises is at least several hundred billion dollars, as much as one per cent of the US gross national product. Worldwide gambling profit is estimated at one-quarter to one-half trillion dollars.

The so-called brick-and-mortar gambling locales, casinos, bingo and videogame halls, etc., have spread from Biloxi to Buffalo, London to Las Vegas, Moscow to Montreal.

Macao, Singapore, and Hong Kong are the Xanadus of today. Their pleasure domes dwarf Las Vegas gaudy palaces.

A “whale” is casino slang for a multizillionaire gambler. Casinos make much of their money from whales, many of whom come from abroad. (Most high-end casinos have a private mini-casino for a “whale”.) At sumptuous casinos-within-casinos multibillionaire whales bet a developing nation’s wealth on the turn of a card. Workaday employees are more enthralled by the weekly football pool than their jobs. The weekly church bingo night seems infinitely more exciting to a housewife than her daily round. Welfare mothers wager their last few dollars on ubiquitous state lotteries.

Recreational gamblers are more eager than ever to toss away dollars or

euros, pounds or yen. Casino resort vacationers routinely figure acceptable losses into their holidays. But as surely as the house percentage destroys an ordinary roulette player's stake, the lives of a predictable percentage of gamblers will be ground down; careers ruined, family lives devastated, as they sink ever deeper into the idiosyncratic world of the pathological gambler.

HISTORICAL BACKGROUND

David Schwartz "Roll The Bones": The Casino Edition comprises an engrossing account of the archeology, anthropology, and history of gambling. Schwartz describes the ubiquity of gambling in virtually every place and era from West to East, primeval to postmodern. One envisions Cro-Magnons wagering flints on how far Og or Magog could chuck his spear, or whether Og would beat Magog in a prehistoric wrestling match.

Some forms of gambling originally had a religious basis, evolving from divination pitched at sussing out the future intentions of the Gods. One of the oldest of these is sortilege interpreting the patterns in a random throw of sticks, stones, or bones. It was only a matter of time before a punter of yore conceived it would be entertaining and profitable to bet on the immediate outcome of rolling the bones, rather than the outcome of a war next year.

Astralagali, the knucklebones of goats and sheep, often figured in the divination of antiquity. Shaved into cubes, marked with numbers or letters, they became forerunners of dice. Early examples have been exhumed from the tombs of Mesopotamia Ur and the palaces of Babylon.

The ancient Greeks and Romans were devotees of divining as well as extravagant gambling. Rolling the bones once used to predict the future now determined who went home broke at the end of the evening. Romans of every class emperors and plebes were ardent dicers. (Lest torture and execution follow, tyrants like Nero and Caligula were allowed to win large night after night.) Rabid fans also handicapped gladiatorial combats.

In the ancient East punters wagered on combat between men or beasts, birds, and insects. Cockfighting was popular in India several millennia ago. In the gambling houses of dynastic Chinese towns enthusiasts bet on matches between quails, fish, and crickets.

Board and tile games were popular in antiquity, played for pleasure or serious coin. Chinese gambling houses featured "mah jong-like games". Phaeronic Egyptians enjoyed early versions of checkers and backgammon. Modern chess derives partly from the ancient Indian game of chataranga, in which moves of the pieces were dictated by the throw of a die.

Gambling flourished in Europe throughout the medieval period and Renaissance. Crusaders played at hazard to pass tedious hours while laying siege to Arab castles, or staked fortune, limb, and life in savage jousting tournaments. Dicing was the most common gambling activity for the gentry

and commoner of the early renaissance.

"Contests between humans, animals, and both were denounced by authority, but nevertheless drew crowds." Bear baiting, forbidden in Elizabethan London, flourished just outside the city proper. It pleasured people of the better or lesser sort as much as the dramas staged in nearby theaters like Shakespeare's Globe. Audiences of bloody tragedies like "Hamlet" could smell genuine blood in the air from the nearby bear pits.

In the first half of the 15th century, a new gambling craze swept across Asia and Europe card playing. With more variations possible, card games quickly became dicing's chief rival. Cards also figured prominently in divination—and still are. The Tarot deck used for fortune telling derives from the deck employed several hundred years ago in Italian card play.

Gambling historians believe that playing cards originated in China and Korea, then were brought to Europe by travelers, traders, sailors, and mountebanks. Florence became a major center for the manufacture and export of cards. The deck of the common player was crude and cheap. The nobility commissioned exquisitely crafted decks at extravagant cost.

Venice figured significantly in the history of Renaissance wagering. The first legitimate European casinos were founded there: a stiff percentage of their earnings went to the Doge and his cronies, one of the first examples in early modernity of a government-gambling connection.

Lotteries were even more likely to be joined to the political fortunes of the Italian city states that sponsored them. Venetians, Genoese, and Romans were notable devotees (the Papacy sanctioned lotteries in 1732). The fiscal debacles that helped precipitate the French revolution were in part stoked by the diversion of state wealth into the pockets of lottery profiteers (Voltaire, included).

Probability theory was arguably first applied to gaming by Giralomo Cerdano, a 15th-century physician and pathological gambler. Before Cerdano, wins or losses were chalked up to the whims of Fortune. With a mathematical grasp of odds, shrewder bets could be laid. Claudius' elaborate wager in the duel between Laertes and Hamlet is based on the assumption that Laertes is the better swordsman. "Hence have we odds", says the wily king, even though he has already plotted a lethal payoff for his nephew.

Blaise Pascal, the 17th-century philosopher, mathematician, and theologian also studied probability theory. Pascal's wager famously proved that belief in God was the safest bet in the Heavenly House. Pascal is also credited with crafting a precursor of the modern roulette wheel, supposedly while trying to build a perpetual motion machine.

Roman soldiers allegedly had already invented a crude version of roulette by upending a chariot and wagering on the spin of its wheel. Although the attribution may be apocryphal, soldiers have always been ardent gamblers, a function of the perennial boredom of hurry up and wait

military life.

During the age of enlightenment gamblers throughout Europe continued to wager heavily against all reason. English gentlemen of the Enlightenment and Regency periods bet hugely on cards and dice as well as horse races, dog/rat fights, and savage bare-knuckle boxing contests.

Many a Lord lost his estate at the tables, then looked to marrying an heiress to mend his fortunes. Groundlings mostly wagered at dice. They also subscribed farthings to lotteries sponsored by the Crown to establish hospitals and charitable institutions. These honorable and occasionally rigged enterprises were the antecedents of the numbers racket and legitimate lottery play of today.

At least since the Renaissance gambling condemned by authority morphed into respectable and powerful businesses. Once more Venice, hub of international trade, became a singularly auspicious locale for gambling on or against the arrival of rich cargo ships from distant climes.

The venerable Lloyds of London was a coffeehouse-cum-bourse before it become a respectable insurance firm. Like the first casinos, early stock markets stood on equivocal moral and legal footing enormously lucrative for owners and the state, yet at the same time harshly condemned from the pulpit, by the press, and the courts for promoting gambling, as it were, by others means.

During the 18th and 19th centuries, elegant gambling halls were erected in spas frequented by the idle, where the pursuit of health served as a convenient pretext for wooing Lady Luck. The watering holes of Baden-Baden, Monte Carlo, Nice, and Bath were notorious for high-stakes play. Roulette was the most popular table game across the social spectrum; faro and variants of baccarat and blackjack were also played in the public halls. Separate private rooms were set aside for high stake card games of whist and piquet.

The fortunes of the great spas eventually waned, although their vestiges still can be found in Nice and Monte Carlo. European public gambling severely declined after World War II because of postwar social and economic devastation. However, renewed prosperity quickly led to the return and prosperity of gambling. As always the authorities took their cut from the proceeds, even as they intermittently denounced wagering. With the rise of powerful international gaming corporations, legitimate gambling has achieved undreamed of profit in Europe as well as Asia (*vide infra*).

Gambling in America

Long before Europeans set foot on the American continent, indigenous tribes were eagerly betting on games and sports. Roll the Bones contains an excellent discussion of early Indian gambling. Despite the Puritan ethic of the early New World settlers, gambling quickly became an omnipresent and predictably ambivalent presence in precolonial America. In 1665,

Newmarket Course, the first American racetrack, opened in Hempstead, Long Island. A merchant of the day recorded in his diary:

> . . . this day did go to Nu.Mark . . . For the race matches. Had news of a strong steed, Ceasar, from Ned Wells, the draper . . . did bet and see Ceasar be outrun. Drank pear cider with Ned . . . Swear to Goody (the diarist's wife) will wager no moor.

The pervasive double standard perennially applied to gambling is piquantly illustrated by the gaming behavior of the Founding Fathers. George Washington played cards almost nightly during the darkest years of the American Revolution. He also issued an edict strictly forbidding wagering among his officers and troops, proclaiming that:

> . . . at this time of public distress, men must find enough to do in service to their God and their country without abandoning themselves to vice and immorality . . .

Luckily for the nation, Washington was a better general than gambler. His meticulous expense accounts reveal he was consistently down a few pounds every year, predominantly at whist.

During the 18th century, American gentry imitated the gambling habits of their English counterparts, wagering on card games, horse racing, boxing, and wrestling matches. Everyday Americans diced as busily as their English brethren. Bearbaiting and similar blood sports were not as popular as abroad, except for cockfighting.

American variations of European rummy- and bridge-like games emerged during the first half of the 19th century. The game known today as poker derived from European models. Craps, which had once been the rage of well healed Parisians after the French revolution, became hugely popular among average Americans. By the Civil War, it was the favorite dice game among soldiers on both sides of the conflict.

As the borders of the nation expanded, the gambler followed close upon the heels of the prospector, the homesteader, and the entrepreneur. Gambling flourished everywhere—in the marginal milieus of mining camps and cattle towns; on Mississippi riverboats; in bustling New Orleans (where Bourbon Street was rechristened the Rue De Craps), in the slums and salons of major cities; in posh summer resorts where political and financial makers and shakers congregated.

American gambling legends spread. During a draw poker session, a known cheater declares he has beaten Wild Bill Hickock with a flush. But Bill calls a full house, turns over a pair of sixes and one ace. He draws his gun, says "Here's my six . . .", then plunges his knife into the table—"and here's my spike! (gambling lingo of that time for an ace)." The cheater turns pale and folds.

The professional gambler's persona acquired an ineluctable shimmer compounded out of a yeasty, penny-dreadful blend of romance and danger. Gambling idioms entered the vernacular, for example, passing the buck,

from the practice of denoting the dealer in a poker game by sticking a buckhorn knife in the table before him.

By mid-19th century, gambling in the United States figured prominently in coalitions between criminal elements and political machines. In the 20th century, widespread corruption of police, judicial, and other government officials accompanied the consolidation of powerful empires built around gambling, prostitution, loan-sharking, and bootlegging after the Volstead act was passed. The mob/government symbiosis especially flourished in Chicago, New York, Atlantic City, New Orleans, Little Rock, and Phoenix, and Las Vegas.

Vegas was the brainchild of the notorious Bugsy Siegal, gunned down in the 50s at the order of Mafia capos unhappy about his managerial skills. As the city grew, mob muscle power controlled casinos and as well as illicit gambling, loan-sharking, and related criminal enterprises throughout the 1950s and early 1960s.

Eventually the mob's iron control yielded to legitimate business interests. With the bad old Ratpack days gone, casinos on the Las Vegas Strip and elsewhere reinvented themselves into family entertainment centers. Middle- and working-class patrons with relatively modest budgets became prime targets of opportunity. Conventioneers replaced the high rolling "whales" of the past. Major money began to be reaped from slot machines rather than table games.

At this writing, family fun still exists in Las Vegas. But one also encounters a sanitized version of the mobbed-up 50s, catering to a younger clientele with money to burn. In garish, theme-based resort/gambling complexes like the Venetian and Caesar's Palace, tourists of every stripe enjoy carefully marketed wretched excess; ever more elaborate slot-gaming technology, celebrity/Cirque de Soleil mega- shows and dancing clubs. Five star restaurants rub elbows with cheap noodle bars and Burger Kings. The heady promise of unzipped sexuality pervades the scene.

The Las Vegas model has spread to casino/resort complexes elsewhere in America, Asia, and throughout the world. The multinational gaming corporations which run them have branched out into running film and record companies.

Since the 1990s brick-and-mortar casinos have been challenged by the Internet's virtual casino, with its ever expanding cornucopia of gaming websites. Internet poker garnered the greatest attention in America until it was banned several years ago. It is now being slowly relegalized, probably because of its potential for generating tax income for states in straightened circumstances. Meanwhile, Internet sport and poker gambling continues thrives worldwide, available at the tap of a computer key.

Nosological Considerations

Pathological gambling was designated as a separate diagnostic entity in the

third edition of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III). Parameters of the disorder were altered in the revision of the DSM-III, reformulated yet again for the fourth edition (DSM-IV), and were retained essentially unchanged in the revision of that edition (DSM-IV-TR). The illness continued to be subsumed in the category of Impulse Control Disorders: Not Otherwise Specified (which also included pyromania, kleptomania, trichotillomania, etc.).

The Problematic DSM-5

Substantive laboratory and clinical research over the past decade has definitively shown that pathological gambling should be considered an addiction, rather than an impulse control disorder or obsessional variant. Accordingly, the DSM-5 removes it from the group of Impulse Control Disorders, Not Otherwise Specified, and situates it in a new category of Non–Substance-Related Disorders.

Unfortunately, the DSM-5 also has replaced pathological/problem gambling with Gambling Disorder (DSM-5-312.21). The new diagnosis has a tame, homogenized ring. It does not intimate on first glance the devastating nature of the illness, as it should for the unknowledgeable reader. Pathological Gambling Disorder or Problem/Pathological Gambling Disorder would have been more appropriate if there must be a disorder at all. Problem gambling is also not mentioned as such in the DSM-5's Gambling Disorder.

Most of the criteria for Gambling Disorder essentially replicate earlier criteria for Pathological Gambling with various emendations. Some of these seem rather notional, and there are a few significant omissions.

To meet the diagnosis of any subcategory of Gambling Disorder, four of the following nine criteria must exist for 12 months. (One has paraphrased some of the DSM-5's clumsier wording for greater clarity.)

1. Wagering increasing amounts of money in order to achieve the peculiar excitement of the illness.
2. Restlessness and irritability accompanying efforts to decrease wagering, or abstain from it altogether.
3. Repeated unsuccessful efforts to control, diminish, or stop gambling.
4. An overwhelming preoccupation with past and future actual gambling activities and gambling-related endeavors such as a handicapping a horse race.
5. Gambling in the context of dysphoric affect anxiety, depression, etc. (One submits that the causes of dysphoria are not cited at sufficient length, not is it sufficiently emphasized that dysphoria may be a product of pathological gambling itself.)
6. Attempts to get even by repeatedly returning to gamble (chasing losses).

7. Lying to significant others in order to conceal the extent of gambling activity.
8. Gambling-related jeopardizing or loss of significant family and social relationships, jobs, educational/career opportunities.
9. Escalation of the search for sources of money to pay off gambling debts. Loan-sharking is not specifically mentioned.

The DSM-5 indicates that a diagnosis of Gambling Disorder cannot be made if excessive wagering is better explained by a manic episode. (Other disorders and conditions which can precipitate pathological gambling are not cited, e.g., delusional schizophrenia, or the administration of dopamine agonists in the treatment of Parkinson disease.)

The DSM-5 distinguishes between episodic and persistent forms of Gambling Disorder. (The factors which precipitate relapse are not cited.) The time necessary for achieving early or sustained remission is notionally set at 3 and 12 months, respectively. The distinction between mild, moderate, and severe types of Gambling Disorder is based upon how many of the nine essential criteria obtain at any given time in a patient's course.

Reflecting one's position on appropriate diagnosis, Pathological Gambling is used below instead of Gambling Disorder in discussing the illness severest presentation. (The majority of clinical studies at this time still employ Pathological Gambling [PG].) The abbreviation is frequently used in clinical practise and research. Problem Gambling is also employed rather than the mild subset of the DSM-5 Gambling Disorder which is characterized by presenting fewer of the criteria.

EPIDEMIOLOGY

Evidence-based research across the world consistently indicates that approximately 2 to 3 percent of any population are problem gamblers who do not meet diagnostic criteria for full-fledged pathological gambling; 0.5 to 1 percent of any population do meet those criteria. Some problem gamblers remain stable although compromised throughout their lives; others shift back and forth between problem and pathological gambling; or slide permanently into pathological gambling.

A definite connection has been established between new casino construction and the rise of problem/pathological gambling in the surrounding area, which did not exist before. Over 3 to 5 years' postconstruction, the prevalence rate then increases to the usual ca. 3 percent.

The escalation of casino and Internet gambling over the past several decades has been paralleled by increase in nonpathological and problem/pathological gambling. In this context, it is inevitable that many individuals with no previous history of problem/pathological gambling will have their latent disorder activated. (Quite possibly these subjects were

involved in some other form of addictive activity before gambling became available.)

The typical patient in earlier studies was a middle-aged white male from a comfortable economic background. However, reliable comprehensive surveys of treatment as well as nontreatment populations indicate that problem/pathological gambling now cuts across every ethnic, class, age, and occupational divide. (According to anonymous sources within the casino world, physicians are among their most consistent heavy players and losers.)

The prevalence of pathological gambling has especially increased in the following groups over the past two decades:

1. The poor: Particularly poor minorities dwelling in urban ghettos.
2. Adolescents and young adults: This group notably includes teen-age Internet poker players, who can gain access to the virtual tables without identification. In general, peer pressure to 'get with it and join the game' plays a major role in adolescent gambling of any sort. Male teenage gamblers predominate significantly over females. Adolescent gamblers are particularly likely to have parents, grandparents and other relatives with serious gambling pathology. Many preadolescent gamblers also bet on Internet poker, Magic Cards, fantasy sports contests, and videogames without a cash payoff. These young gamers are thought to be at risk for developing gambling difficulties later on.
3. The elderly: Particularly retirees with an inadequat social network, who live alone, have limited recreational or educational interests, or are homebound because of health problems. Problem drinking or frank alcoholism frequently accompanies gambling problems in this cohort, often with no prior history of alcohol abuse or gambling difficulties. "Day Trip" gambling by seniors in assisted living and nursing home facilities is often encouraged by nearby casinos and slot parlors. Inevitably, some day trippers become pathological gamblers.
4. Women: Several decades ago, male PGs were said to outnumber females by 5 to 1. Recent studies indicate a shift toward gender equality Although evidence-based study is limited, it is postulated that more women gamble today because of their increased presence in the workplace, with more ready cash at hand.

   Women formerly favored bingo, lotto, and slot machine play, but now participate in every wagering activity. Most women begin intemperate wagering later than men, but are quicker to become full-fledged PGs. Gambling-related suicidality is higher in this group.
5. Prisoners: Studies indicate that approximately one third of American prison inmates have formidable gambling issues. Many bring a previous gambling problem to incarceration. But some gambling begins in prison as an endemic solution to the tedium of prison life. Offenders with no

previous difficulties often become pathologically involved with card games like poker, hearts, backgammon, or chess. Cigarettes, drugs, privileges and cash comprise the common currency of prison gambling. Loan-sharking is common. Failure to pay debts can exact brutal punishment, even death.

Prisoners with gambling problems frequently suffer from cross addiction to alcohol and substance abuse, as well as other serious Axis I and II disorders. Unfortunately, prison systems rarely have treatment programs for PGs, compared with programs for substance and alcohol addiction.

One or another ethnic group is often declared particularly prone to recreational and pathological gambling. Chinese, Filipinos, English, Irish, Australians, and Americans have been nominated for pride of place. Absent reliable hard evidence, the author deems it most likely that gambling worldwide is inflected by national character and custom but no single group actually predominates.

FAMILY HISTORY

The family histories of pathological gamblers reveal that one or more parents, grandparents, or significant relatives are heavy or frankly pathological gamblers. Family members also have increased rates of affective illness—notably major depressive and bipolar disorder; substance abuse particularly of alcohol, cocaine, and nicotine; and also experienced verbal and sexual abuse in childhood. Both ADD and ADHD are supposedly found more often in PG family members, but to date there is no hard research on this score.

One conceives pathological gambling as the dark underbelly of the American dream for the PG's family of origin typically prizes competition; is highly materialistic, with little inclination to save or budget. Intense admiration for wealth and its signatures is common. Frequently a parent or grandparent has become well-to-do through hard work. Sadly, the same individual may have gambled away his business.

Clinical Features

Lesieur's description of the pathological gambler's signs, symptoms and life trajectory in 1977 is as relevant today. His typical subject was a middle-class married male in his 30s or 40s. This type is still quite common, but the clinician will encounter many similarities, differences, and variations of the model as a function of age, gender, ethnicity, and social class.

The PGs intense preoccupation with betting usually begins in adolescence, but can emerge even in the preteen years. In his youth, he acquires considerable gambling skills, often in a number of activities. A PG often experiences exhilarating big wins early in his career, which augments

an ineffable sense of specialness.

Because of his desire to make it big, he commonly drops out of high school or college and moves away from home. Frequent gambling continues into adulthood, but is reasonably well integrated into his life. Most PGs marry, raise children, work hard, often achieve considerable success and live large. Contrary to a popular image of the PG as a callous, entitled narcissist, he is usually deeply committed to family life; takes great pride in being a good provider and parent.

At some point the PG becomes embroiled in an ever escalating, ever tightening spiral of gambling involvements and options. He may start losing in the context of life stress; by wagering in new locales such as casinos and the Internet; or for no discernible reason whatsoever. Instead of stopping and cutting losses which he has often been able to do before he begins to chase, calling on more and more options to recoup mounting losses.

The PGs taps into sources of cash for gambling such as family checking and savings accounts, piggy bank and rainy day stashes; his personal or workplace business; legal and illegal lending institutions; credit cards; and fellow gamblers.

As ever more options are used, they are also used up. The spiral tightens, until that fewer and fewer options remain. The PG becomes trapped in a formidably powerful, self-enclosed system which ensures his inevitable downfall. He bets ever larger sums; becomes embroiled in more types of gambling for longer periods of time. Greater losses occur because of the grinding down of natural and house odds, as well as uncharacteristic impulsive, foolish play.

Most pathological gamblers are highly knowledgeable. But their expertise steadily crumbles as loss accrues. Indebtedness is compounded by the exorbitant interest rates (vigorish or the vig) on loans floated by sharks and credit card advances. When every option is exhausted, and every source is dried up, the PG reaches the disastrous moment of closure.

Crimes such as embezzlement, passing bad checks, and other types of fraud may be undertaken to get even. Female PGs may become prostitutes. Such offenses were once believed to be hallmarks of end-stage pathological gambling. Recent studies indicate that petty criminality often begins earlier, and continues throughout the patient's career. However, pathological gamblers without genuine antisocial or sociopathic traits rarely become involved in major criminal enterprises.

The revelation of the PG's transgressions is excruciatingly humiliating, and consequences are often catastrophic. Exposure may precipitate imprisonment, professional disgrace, divorce, escape to another state or country, or the ultimate deliverance—suicide.

But it is more likely for exposure to be followed by a favorable outcome with restitution of debts undertaken by family or friends; remission of

debts by a bookmaker, bank, or loan shark; an employer's forgiveness; a kind judge who orders probation rather than incarceration. Typically, the pardoner is deeply moved by one of the innumerable "sad tales" every PG always has at his or her beck and call.

The craving for action is central to the PGs driven modus vivendi. A PG frequently rationalizes this overwhelming desire by convincing himself that one simply must be wherever the action takes in order to catch up and stop chasing. But it is the pull of action itself which draws the pathological gambler ever deeper into the gambling vortex. It matters little whether he is losing or catches up. The ineluctable lure of action is captured by Dostoyevsky's mot "le jeu pour le jeur" ("The game for its' own sake").

Trippett observes that:

> Supposedly . . . action) speaks to the process of betting. I suspect it speaks more deeply to what happens within the gambler. He places the bet, juices flow, he really feels . . . alive. Action: when the bet is on, his existence is confirmed . . .

The idiosyncratic action rush is inflected by the peculiar ambience of each gambling locale the racetrack's steamy milieu, polluted by rumor and false report; the clockless eternal present of the casino, with its idiosyncratic symphony of sounds—the click of a roulette wheel; the mellow warbling of slot machines, the joyful uproar around a craps table when a winner on a roll makes his number again.

The action high can be obtained in so-called action systems a pool hall; bowling alley; the backroom of a candy store; a diner—where bookmakers, pros, recreational and pathological gamblers congregate. Here every kind of play is found or can be arranged.

Some PGs stay wedded to the game they have always played as they become enmeshed. However, a chasing PG usually finds that one kind of gambling is insufficient to satisfy his craving. So he branches out into other gambling activities. A lack of expertise in a new game often compounds losses, accelerating the chase.

As the spiral deepens, the trajectory of action inevitably extends beyond the time consumed in betting to embrace longer time before and after wagering—planning action; setting up a poker game; handicapping races at several tracks during a single day; getting to action as in leaving work prematurely to get to a racetrack in time for the daily double; dealing with the consequences of action—collecting winnings or, more often, dashing madly about to cover losses; making interminable excuses to a bookie for coming up short; to an employer for lateness and absences; to his wife for coming home at 4 AM only to leave a few hours later.

The pathological gambler is a master at manipulating cash flow—moving money. His previous business skills are now directed at the business of enabling play. He shuffles loans, bank accounts, cash, from one source to another in dazzling displays of financial legerdemain which would

do a Ponzi proud. His hectic strategies in moving money actually become an integral ingredient of the ominous, raw excitement of the action-seeking lifestyle.

The PG in serious chase mode is like a juggler, desperately fielding an ever increasing number of plates, oranges, bowling clubs tossed at him from offstage. Like his psychological twin, the adolescent, he is supremely engaged in an eternal present. A central cause of mounting losses is his utter incapacity to envision a future more than a few days away, a time when he wont be able to secure the cash to forestall closure and exposure.

The PGs trajectory over years is consistently marked by repeated cycles of spiraling, closure, exposure, bail-out, and relapse. Relapse is often precipitated by the prompting of fellow PGs who cannot abide the presence of an abstainer; by the intense boredom accompanying abstinence; or by the very fact that the problems precipitated by closure have been resolved. As the horror of indebtedness fades from memory the memory for gambling disasters is notably short the pursuit of action is all too easily resumed.

Pathological gambling occurs in companionable social circumstances at a blackjack or craps or poker table; at a horse or dog track or football game. Most PGs are outwardly gregarious, with an apparently wide network of relationships including other gamblers. These often have a curious ephemeral quality. For several hours of poker play, the man beside you is your best friend before he gets up and vanishes forever into the night. The decision to gamble with fellow PGs at the expense of the sufferer's family and social life is an especially significant choice in the course of the illness.

A lone wolf PG is rare. In aid of rationalizing schizoid tendencies, he will often assert that he shuns fellow gamblers to avoid bad advice or loaning hard won cash. Most Internet gamblers are not lone wolves. Computer play commonly alternates with social gambling. However, when Internet waging creates the pathological spiral in one so predisposed, outside wagering activities along with other social outlets fall away.

The induction of a gambling spiral has been ascribed *inter alia* to marital problems, death of a loved one, the birth of a child, business reverses or successes. However, on closer examination, these may really not be the actual causes of chasing but *post hoc propter hoc* events. Marital problems may be the consequence, not the precipitant of pathological gambling.

Whatever stressor precipitates the chase, the PG becomes an extraordinary source of stress for others, at home and at work. For many, the hallmark of the illness' increasing severity is not simply the time devoted to gambling, or the preoccupation with gambling, or mounting indebtedness—all wrenching enough. It is rather the devastating impact of pathological gambling upon every relationship outside gambling.

As the chase escalates, the PG's quality of life progressively deteriorates.

For years, spouse and family may have been totally unaware about the extent and severity of gambling, until indebtedness drives the patient to invade family funds. Exposure is especially calamitous for the PG's spouse shaking confidence in the PGs successful imagine, evoking angry feelings of betrayal.

Exposure regularly is followed by impassioned promises to cease gambling, leading to tearful reconciliation. But as the PG continues repeated chasing cycles and deceitful tactics to avoid the wife's "bitching and moaning", an atmosphere of pervasive mistrust comes to pervade the marriage. It degenerates into squalid ploys and counterploys revolving around the PG's concealment strategies, and the spouse's maneuvers to protect herself and her children.

Despite her suffering, the PG's wife does not easily oppose her husband and in gambling lingo go on the warpath. Once forged, the dysfunctional marital bind may continue for years before the wife issues a final ultimatum to cease or face divorce. Research does not support the common belief that PG wives are inherently masochistic. Nor do they subtly encourage the PGs illness.

It is rather that a wife, husband, or lover becomes locked into a uniquely compelling interpersonal system which eventually justifies its own existence. The wife or husband genuinely cares for the PG spouse, tends to deny or excuse gambling behavior, especially with the PG's convincing expressions of guilt. Even when a wife winds up loathing her PG husband, she may be unable to break free of the system because she fears being unable to support herself, or realistically cannot do so (one notes the similarity to the battered woman's plight).

The pathological gambler's children are tragic victims of the illness. They bear the burden of his neglectfulness, of parental disharmony, and of his actual abuse. Predictably they have a high rate of psychiatric symptoms and behavioral problems. By that poignant process through which one identifies with a traumatizing parent, they also stand at considerable risk of becoming PGs themselves.

The crumbling of the PG's value system is as traumatic to himself as others. He has always treasured his honor; loathes himself for abrogating it. Like addicts in the throes of their illness, he becomes both demoralized and de-moralized exhibits ego-alien, indeed abhorrent behavior, which would never exist were he not caught up in the chase. In end-stage gambling, the pathological gambler's entire existence revolves around play. At this point he may paradoxically derive little real pleasure from gambling. Winning has long since ceased to be an end, rather the means of securing further action.

Some PGs cease serious gambling in later life. Cessation does not appear related to treatment of any sort. Akin to the burn out of drug addicts, the craving for action seems to simply wither away for obscure

reasons. Mutatis mutandis, a PG may escalate chasing in old age, as gambling becomes his life's raison d'être.

PERSONALITY CHARACTERISTICS

Pathological gamblers display a strikingly similar cluster of personality traits. Depending on age, gender, and social class, some may be exaggerated or absent.

In pure culture pathological gamblers are competitive, going back as far back as latency or early adolescence. They are intelligent, although usually not deeply intellectual.

They consider themselves rugged individualists; are highly independent, overconfident to the point of abrasiveness. A keen interest in sports participation and viewing parallels their interest in gambling. They are perennial optimists, deniers; possess a distinctive gallows humor:

> Joe: Which fighter are you betting on tonight?
> Jim: Both.
> Joe: But you'll only come out even.
> Jim: I need the money!!!

PGs are by turns touchingly loyal, or gaggingly insensitive, notably when chasing. Despite their apparent sociability, they are often beset by profound feelings of loneliness. They do not easily express feelings, indeed may be so unable to get in touch with their emotions as to be alixithymic.

PGs as well as successful professional gamblers often project an air of grandiose elitism. They believe they always know the score at gambling and life in general, for which they hold gambling a Hobbesian dog-eat-dog metaphor. The pride of pathological gamblers in their idiosyncratic vocation, even at its disastrous end stage, is poignant.

COMORBIDITY

Extensive research indicates that impressive comorbidity exists between gambling and affective disorders, notably major depression and bipolarity and addictive disorders, notably to alcohol and cocaine, nicotine, and caffeine.

There is significant comorbidity between pathological gambling and attention deficit disorder, with or without hyperactivity, in adolescent as well as adult problem/pathological gamblers. One has the sense that many PGs have a forme fruste of full blown ADD/ADHD.

Various personality disorders have been described in PGs, especially narcissistic, borderline, and antisocial disorders. Although many pathological gamblers have obsessive I ve personality traits, full blown OCD is rare.

Mutatis mutandis, a higher rate of pathological gambling has been discovered in patients with a primary diagnosis of affective illness

(particularly bipolar) and schizophrenia. One's clinical impression is that patients with generalized anxiety disorder, panic disorder; narcissistic, borderline, and antisocial personality disorders are more likely to have gambling problems. However, the relationship has yet to be proven.

The pathological gambler's course is often worsened by serious comorbid conditions, particularly alcohol and drug addiction. The easy accessibility of these substances in or near gambling milieus further disinhibits the PG, resulting in wilder play and greater losses.

Exacerbation of an affective disorder may be a primary precipitant of a gambling binge. Heavy indebtedness may precipitate a dysthymic state or frank major depression, which then triggers off more betting binges.

Pathological gamblers paradoxically are often both hypochondriacal and dreadfully neglectful of their physical well-being. PGs eat poorly and excessively; are chronically sleep deprived and underexercised. They are more vulnerable to hypercholesterolemia, arteriosclerotic cardiovascular disease, diabetes, peptic/gastric ulcer, hepatic and pulmonary disorders (often a function of nicotine abuse).

DIFFERENTIAL DIAGNOSIS

Making a diagnosis of pathological gambling is often problematic because PGs so often minimize, rationalize, or absolutely deny their difficulties. Should doubts arise, the clinical interview may be supplemented by reasonably reliable tests such as the South Oaks Gambling Screen (SOGS) and the National Opinion Research Center DSM Screen for Gambling Problems (NODS). SOGS is better known, and frequently used in PG research. NODS is a population-based telephone interview which identifies gambling problems according to DSM criteria. Although initially useful, an extensive clinical interview obviously provides the most accurate information for the diagnosis to be made.

A problem gambler is most likely to be seen first by a clergyperson or general physician. Referral to a psychiatrist or other mental health practitioner usually occurs further down the line. Because of the patient's intense denial, both laypeople and practitioners need to recognize the symptoms and behavior which will differentiate problem/pathological gambling from the recreational betting practised by at least 85 percent of Americans.

Ordinary recreational wagering is based on a mix of conscious and unconscious factors unique to each unproblematic bettor, viz:

1. A desire for relaxation in a companionable social setting, away from the everyday grind. Recreational gamblers usually play on designated occasions such as a weekly poker game. They have a firm grip on acceptable losses.
2. The profit motive. One often hears players proclaim they don't care if

they win or lose, it is the fun that counts. If pressed, however, most recreational gamblers will admit that they would rather win than lose.

3. The inherent pleasure of exercising a variety of skills/and ego functions depending on the nature of game or sport.
4. The satisfaction of nonpathological competitive drives.
5. The attraction of risk taking in a betting situation where negative outcome produces minimal if any psychic trauma. In a quixotic world where mortality is the only certainty, yet the time of its coming ever in doubt, gambling allows one to indulge in the illusions of omnipotent control over destiny.
6. The impression of control is fostered by games in which the chance element can be reduced by studying the parameters of play and doping out the odds, for example, racetrack handicapping. When loss does occur, one is gently reacquainted with the limitation of one's powers.
7. The gratification of other obscure, possibly conflict-based drives and motives stemming from early stages of psychological development, etc.

For the professional gambler, wagering is a business, plain, and simple. Contrary to the myth of the glamorous high-roller, most pros avoid the limelight if only to avoid taxes. Their expertise is impressive. They are summarily prepared to take requisite risks; coolly acknowledge the losing streaks which come with the territory, and are able to weather losses on any given day without losing control. However, given the inherent nature of the work and personality factors, a pro is much more likely to slip into PG chase mode than the ordinary gambler.

Acute substance abuse accompanied by temporary disinhibition may precipitate a gambling bout without other symptoms of pathological gambling. Alcohol and cocaine are particular offenders in this regard.

Gambling binges are occasionally seen in the setting of an acute or chronic schizophrenic delusional system, for example, a patient who believed God had ordered him to stake his life savings at the dog track so as to hasten the end of days foretold in Revelations.

Excessive wagering may occur during the manic stage of a bipolar disorder in patients who have been unproblematic gamblers before, or have never gambled at all. The correlation between gambling episodes and mood swings in pathological gamblers with significant comorbid affective illness is a vexed question—obvious in some PGs, either nonexistent or obscure in others.

Bipolar PGs with mania in remission nevertheless can appear as wired to the universe as a patient in a full blown manic state; crackling with tension as they pursue their dizzy dance over the abyss. However, in the former case the PG quickly returns to a more temperate mood once play has ceased. He may be reactively and temporarily elated after a win, or just as appropriately depressed after a loss. In any case, a mood disordered

problem/pathological gambler will generally wager excessively whether clinically manic, depressed, or euthymic.

Individuals with antisocial personality disorder frequently gamble, but the nature of their behavior reflects their essential criminality and utter lack of empathy. They will do anything possible to win and are chronic cheaters. They are often poor players, blame others for their losses rather than recognize their own inadequate play. An antisocial individual may retaliate violently after a loss; indeed may subsequently rob an individual to whom he lost, or hold up a game where he lost.

Unlike antisocial gamblers, the majority of pathological gamblers have a better work record, a more stable family life, and a reasonably scrupulous moral set before their illness catches up with them. They are ashamed when driven to crime—usually petty—out of profound desperation, to cover losses and keep gambling.

Although gambling may seem to have an obsessive element, it is rare for an individual with obsessive-compulsive illness to become a pathological gambler. Unlike the tormenting ego-alien symptoms of an OCD patient, the pathological gamblers overwhelming preoccupation with play is usually quite egosyntonic, even persisting when the patient has reached a terminal, end-stage of gambling.

Bouts of pathological gambling are rarely due to compromise of cerebral function by injury, infection, or neoplasm. Intemperate wagering in patients with Parkinson disease treated with dopamine agonists is further discussed below.

ETIOLOGY

For centuries, pathological gambling was chiefly viewed as a moral stain rather than a medical problem. Gamblers were said to suffer from an engrained weakness of character that rendered them susceptible to temptation from the dark side. The evils of gambling were associated with the evils of drink, sexual depravity, and low companions.

Pseudopsychological theories surfaced in the 19th century, ridden with sanctimonious moralizing. The advent of psychoanalysis spurred awareness that pathological gambling might have a scientific, rather than a "weak-brain" or theological basis. During the last two decades, exploration of the PGs psyche has been joined by impressive research into the illness' genetics and neuroscience.

Psychoanalytic Theories

The literature on pathological gambling was sparse at the advent of psychoanalysis. Theories chiefly addressed the gambler's childhood trauma and psychic complexes.

Early analytic investigators drew analogies between pathological

gambling and sexual excitement. It was commonly supposed that compulsive wagering comprised some sort of neurotic substitute for mature erotic fulfillment.

However players state, and researchers have affirmed that the unique thrill of gambling action is utterly divorced from Eros. Nell Kimball, a madam of the frontier days whose establishment provided prostitutes and gambling house declared that "if you give a man a choice between gambling and sex, he will take gambling any time." The majority of pathological gamblers are fully capable of enjoying sex. All in all, the displacement of libido from bedroom to bookie must be rated a very vexed question.

The first psychoanalytic theories now appear simplistic or simply wrongheaded, but the findings of depth psychology over a century should not be dismissed out of hand. While analytically oriented therapy without other adjuvants will not benefit most PGs, it does yield helpful insights into the gambler's inner life and motivation, his vertiginous fluctuations between exhilaration and despair.

Simmel published the first account of a pathological gambler's analysis in 1920. He thought that pregenital and oedipal conflicts were implicated in the illness. Stekel described the gambler's frequent mood swings, preoccupation with superstition and rituals, and fantasies of omnipotence. He classified the PG as a compulsive neurotic with latent homosexual tendencies.

In his famous 1928 monograph, "Dostoyevsky and Patricide", Freud likened pathological gambling to other obsessive-compulsive behaviors. He argued that the great writer's gambling addiction stemmed from a displacement of childhood masturbatory urges. The frenetic drivenness, escalating tension, and use of the hands in roulette and other games were autoerotic equivalents.

Fenischel astutely compared the gambler's cycles of triumphant wins and desolating losses with manic-depressive cycling between giddy elation and devastating despair. He speculated that pathological gamblers believe they are entitled to special protections and gratifications, based on grandiose infantile expectations of parental bounty. Gambling is the magical means by which fate is compelled to act in the patient's behalf. In a variation of the *Schicksalneurosis* the destiny neurosis the pathological gambler surrenders autonomy and responsibility for his actions, perceiving himself either as destiny's favorite or victim, depending upon whether he is winning or losing.

Kris theorized that the pleasure of the PG's wagering articulates with guilt-ridden Oedipal sexual tensions, so that the heady joy of action becomes ineluctably fused with dread. The patient's ego is unable to control what it has initiated. He becomes trapped in a vicious cycle of anxiety: the need for reassurance of his survival through winning; more play; mounting losses; and the ominous escalation of tension. Gambling evolves from an

innocent pastime into a literal matter of life and death.

Edmund Bergler's theories about pathological gambling were immensely influential for decades. His signal assumption was that the pathological gambler is driven by a desire to lose. Bergler asserts that the PG bears profound grudges, first against parents, then later authority figures who are resented for imposing the reality principle on his infantile narcissism and omnipotent fantasies.

According to Bergler, aggressive impulses unconsciously directed against the misperceived deprivers of nursery days cause inordinate guilt and an intense need for punishment. Gambling recapitulates the cycle of childhood rage, aggression, and remorse. Pleasure in wagering reflects the exhilaration of grandiose rebellion; pain arises from the morbid certainty of punishment.

As the pathological gamblers anguish about losses becomes erotized, he develops psychic masochism: savoring his victimhood, relishing injustice collecting, reveling in self-pity and spurious righteous indignation. Lady Luck at base is his Bad Mother. He hopes his suffering will compel her to restore the blissful infantile nurturing she snatched from him.

While Bergler's ideas are now viewed as at least questionable, they nevertheless have provided grounds for more sophisticated psychoanalytic study. For instance, Boland and Boyd observe that:

> Gambling is (a) forbidden and an unconscious transgression because of the . . . intrapsychic satisfaction of multiple and libidinal determinants . . . (this gratification) is the origin of the abundant, pervasive guilt feelings which plague the gambler, as well as the origin of anticipated desired punishment in the form of gambling loss. Gambling is a unique situation with inherent infallible mathematical machinery . . . which insures guilt-alleviating loss, such that the unconscious books are balanced and restored to psychic equilibrium . . . (pathological gambling) activates guilt and contains its own system of probablistic loss and guilt cancellation . . .

Whatever the pathological gambler's other key conflicts, one speculates that the PGs devotion to wagering constitutes an uncanny spilt religion. His plethora of superstitions and rituals seem pitched at winning Fate over to his side, in aid of dealing with a devastating fear of death which has largely gone unappreciated in the literature. One wonders if the PG may court the erratic favors of Lady Luck in aid of making the ultimate score—a triumph over death itself.

MATERIALIST RESEARCH AND HYPOTHESES

Imaging Studies

More questions have been raised than answered about the neurobiology of the pathological gambler. Nevertheless, researchers are beginning to make progress in mapping cerebral centers and pathways which mediate judgment, mood regulation, decision making, and impulse control. These are compromised in pathological gambling, as they so often are in substance addictions.

Patients with cerebral trauma, tumors, or infectious processes which damage the ventromedial prefrontal cortex and amygdala *inter alia* may exhibit denial or unawareness of situations which require common sense judgments. Actions with immediate rewards are favored, while long-term negative consequences are ignored. Experts speculate that PGs may have some sort of analogous cerebral compromise which facilitates immoderate betting activity.

Many imaging studies have focused upon the ventral and ventrolateral striatal prefrontal cortex and connections with other brain loci mediating prudent judgment. Experts speculate that suppression of activity within these pathways may cause a blunted neurophysiological response to rewards and losses, in turn eroding the ability to make appropriate decisions. PGs, like substance abusers, regularly seek short-term gratification without regard for later consequence. This *belle indifference* may stem from deficient warning signals due to a dysfunctional VM/VR regulatory system.

The role of visual cueing in pathological gambling has been studied with various imaging techniques. PGs viewing videos of gaming scenarios demonstrate heightened activity in pathways between *inter alia* the right dorsolateral prefrontal cortex, right parahippocampal gyrus, and the left occipital cortex. Corollary conscious cravings to gamble were experienced by subjects after exposure to the gambling videos.

The impact of risk taking upon the PG brain has been demonstrated in PET research, in which subjects playing for real rather than simulated money showed an increased metabolic rate in the primary visual cortex, cingulate gyrus, putamen, and various prefrontal areas.

These findings are intriguing, but only scratch the surface. Any episode of recreational as well as pathological gambling involves many other activities than betting. A bout includes arousal of the urge to gamble; planning the time, place, and choice of gambling activity; experiencing the ebb and flow of action, for example, at the race track or card table; dealing with the outcome of winning or losing (e.g., moving money to cover losses). Investigators are still far from being able to monitor what precisely the PG brain is doing, which cerebral loci and pathways drive the doing, over the entire trajectory of a gambling episode. One summons up a “Star Trek” fantasy of ultratech halter monitoring.

Neurotransmitter Research

Many recent studies indicate that PGs suffer from complex neurotransmitter dysregulation similar to the abnormalities found in substance abusers and patients with other behavioral/impulsive problems. Imbalances across the entire neurotransmitter spectrum—serotonergic, noradrenergic, dopaminergic, GABA etc., have been variously postulated as significant factors in the dysfunction of cerebral sites and pathways

mediating behavioral arousal, inhibition/disinhibition, reward/reinforcement, and psychophysiological stress.

The particular focus on the role of increased dopaminergic activity in pathological gambling arose from the serendipitous discovery that approximately one-third of patients being treated for Parkinson disease with dopamine agonists develop the illness without any previous history. The gambling excesses of these patients usually disappear upon withdrawal of the offending drug.

Pathological gambling in the Parkinsonian group treated with dopamine agonists has been ascribed *inter alia* to disproportionate stimulation of dopamine receptors localized in mesolimbic pathways, which are thought to be involved in reward seeking.

However, one notes that two-thirds of the Parkinsonian patients studied do not develop pathological gambling, indicating that other factors must play a role in the afflicted third: pre-existing cerebral damage due to Parkinson disease itself, genetic, family, and sociocultural influences, etc.

Other Predisposing Materialist Factors

Adolescents, particularly male teenagers, have a higher incidence of impulsivity and stimulus seeking than other age groups, as well as a higher incidence of recreational gambling, arguably attributable to the cerebral and hormonal changes of that period. While it has not been established that average teenagers are more predisposed to pathological gambling than adults, adolescents with ADD and ADHD are significantly comorbid for pathological gambling, and vice versa.

A small study of adult nongambling subjects exposed to high- and low-risk gambling experiments, showed that those preferring stimulus-rich environments consistently sought long shots, wagering more to win more. One wonders if such subjects have some sort of inborn predisposition toward stimulus seeking which, given other factors, might lead them to become more attracted to social, and even problematic gambling.

Twin studies suggest that PGs, beyond identifying with a gambling relative, may have an inherited genetic potential to develop the illness.

Variations in several alleles have been discovered in PGs, but more research is needed to establish whether a predisposition for problem/pathological gambling exists at a chromosomal level.

Behavioral/cognitive research focuses on pathological gambling as a learned, disastrous habit. Based on the available research as well as one's own clinical work, it appears more likely that the PG's habit is driven by neurobiological factors, rather than simple introjection of a gambling relative's addiction. Hence the discussion of behavioral/cognitive factors has been placed in the category of materialist hypotheses.

Many pathological gamblers exhibit sensitivity to a gamut of reinforcers directly or indirectly related to wagering. An example of direct or primary

reinforcement is the slot-machines lure of high return for a repetitious low-response small investment.

An example of indirect or secondary reinforcement is the "comp" offered by casinos, depending on hours of play, player's wealth, etc. The comp can include a special room rate or no charge at all, free food, drinks, etc.

Contemporary casino design, technology, and management practices demonstrate an elegantly sinister deployment of virtually every behavioral stimulus for a vulnerable problem or pathological gambler to bet over his head, not with it.

Pathological gamblers are exceptionally prone to betting-related cognitive biases and distortions. These include: discounting the large sum of money owed as a result of making small bets toward achieving the big score; magnifying one's own skills while minimizing the real talents of other players; losing sight of the amount of gambling time needed to recover losses; selectively remembering memories of wins over losses; the abiding conviction that a chain of losses must herald an impending win, especially after a near miss (the so-called gamblers fallacy, but actually one of many).

A DARWINIAN SIDEBAR

At first glance, gambling would seem to have little if any evolutionary value. If the impulse to gamble is inherently unhelpful for survival today, why has it not simply mutated away?

There is obvious survival value in prudently choosing to stay inside the cave while outside the winds rage and the beasts howl. *Arguendo* the chance of group survival was improved when some Og or Magog proposed a better, if more dangerous plan to kill a mammoth in aid of providing better sustenance.

But what if the daring caveman subsequently began to wonder if a life better than bonechilling subsistence lay over the next mountain, or beyond? It is not unreasonable to suppose that he, his children, or grandchildren would eventually decide to defy the prevailing belief that monstrous creatures or molten pits of fire lay over that mountain, and seek a land where the weather was mild, and food plentiful.

One conjectures that certain individuals have always been more genetically/neurobiologically hardwired, as well as characterologically predisposed to appropriate risk taking, whether in geographical exploration, warfare, scientific research, finance, athletics, etc. Not only would these risk takers enjoy the prospect of their perilous venture's success, they might also savor the peril *sui generis* to an appropriate extent.

Now entertain the possibility that viable risk taking turns pathological for a small subset of the daring ones. In such cases, the risk inherent in a

dangerous enterprise may become the sole end, rather than merely one aspect of attaining whatever achievement is being sought. Throughout history perilous enterprises have suffered or even failed because a pathological risk taker courted greater danger than warranted for its own sweet sake. (One would expect to find a wide range of behavior between appropriate and absolutely pathological risk taking.)

Whyte acutely speculates that the appropriate risk taker is able to see a useful, albeit perilous pattern of opportunity in the environment, and takes prudent advantage of it. But an inappropriate risk taker may play down the pattern's inherent danger, indeed misperceive that one risk or another should never be taken, no matter how auspicious the intended outcome.

From this perspective, one conceptualizes gambling of any sort, recreational to pathological, as a spin-off of the Cro-Magnon's hard-wired risk-taking behavior. With no mammoth to slaughter, with no journeying into frightening terra incognita, wagering at best provides a pleasurable anticipation in the moment, joined to hope of a successful outcome; at the very worst the pathological gambler savors pleasure even when he knows the odds are stacked against him. Research cited above indicates that the PG's pleasure might actually be heightened by the heady promise of ultimate success, after heavy initial losses.

TREATMENT

Virtually every modality has been used to treat the pathological gambler. For decades results were ungratifying, and relapse rates high. Better understanding of the disorder's psychology and neurobiology have improved outcome somewhat—this in the context that only 10 to 20 percent of PGs actually seek or are compelled to seek professional help. Success hinges upon carefully individualizing treatment, case by case. Combined approaches are frequently required. It is not unusual for a patient to need psychotherapy; cognitive/behavioral work; medication; couple and family therapy, and membership in a Gamblers Anonymous (GA) group.

PGs make extremely difficult patients. They rarely want help, persistently deny flagrant problems, and can charm the pants off an unwary therapist. The average pathological gambler is often forced into treatment only after disclosure precipitates an ultimatum to get help or suffer divorce, prosecution, etc.

The PG's enormous pride, formidable denial, alixithymic lack of introspection, chronic impatience, and incurable optimism all militate against forming a viable working alliance particularly in psychoanalytically oriented therapy. These patients usually present a host of dire symptoms and problems related to the chase- depersonalization, hypochondriasis, panic attacks, substance abuse (notably of alcohol, cocaine, and tobacco) which do not readily yield to talking approaches. The patient is also likely

to suffer from other physical illnesses which have gone untreated or poorly attended to because of his own neglect- obesity, peptic ulcer, etc.

Pathological gamblers expect magical interventions; frequently quit treatment when miracles are not forthcoming. They often go through the motions of patienthood, while remaining inwardly convinced that a turn in luck and improved cash flow are all that is needed to resolve their problem. They are usually scornful of someone who doesn't gamble. However, the therapist who is a gambling tyro can put his ignorance to work by having a client teach him the ropes, much as an adolescent therapist will ask a stony silent youngster to instruct him about rap music.

Psychoanalytically oriented therapy with a pathological gambler is stormy indeed, akin to working with a resistant adolescent. One routinely expects missed sessions; relapses; and financial crises which cause disruption or cessation of treatment due to inability to pay.

Firm but kind limit setting is crucial. In most cases, the patient is asked to pay for each session on the spot, regardless of the cost. Greater responsibility in handling fees is often a sign of improvement. One has found that carrying a PG seriously undermines therapy, as does lowering a fee because of gambling losses unless serious physical illness supervenes or the patient loses a job for reasons not related to gambling.

The therapist should be vigilant vis-à-vis the veracity of patient's explanation about problematic treatment issues, as well as serious family and work difficulties. A PG in chase mode will minimize troubles or justify them with a "sad tale". Reality can be treated like toffee stretched here or there according to whatever will work. As in the treatment of seriously disturbed children and adolescents, it may be necessary to contact significant others to get a clearer picture of what is truly going on. The patient should be informed of this possibility at the beginning of treatment, with the understanding that the therapist would rather maintain confidentiality if the patient will hold up his end of the contract.

PGs require a tolerant, noncritical attitude regardless of the treatment modality employed. With their testy pride and razor-keen ability to sense disapproval, they will quickly reject help if they suss out even the smallest tincture of moralizing or criticizing.

The attractiveness of group techniques for pathological gamblers is not surprising, given their extraversion. Group, family, and couple therapy can all be helpful. The distressing impact upon significant others often registers more compellingly in family sessions than in individual therapy alone.

Psychopharmacological treatment, once largely unsuccessful, now frequently plays a significant role in management. Success has been reported with antidepressants, notably the SSRIs and Wellbutrin; mood stabilizers including lithium, lamictal; atypical antipsychotics, and opioid agonists such as naltrexone. (The use of benzodiazepine drugs should be avoided or carefully monitored because of the potential for triggering, or

increasing comorbid substance addiction.)

In many patients it is difficult to determine whether an antidepressant or mood stabilizer is alleviating gambling cravings directly, or through treating the comorbid condition—especially affective/bipolar disorders. Many short-term outcome studies have demonstrated improvement. Unfortunately, there are much fewer studies of long-term effectiveness. The same problem exists with other treatment strategies. Pathological gambling is not a monolithic entity. Different PGs may respond robustly to a medication which somehow addresses their specific neurophysiology.

GA is an international self-help organization administered by abstinent pathological gamblers, with branches in many large cities and small towns, here and abroad. GA groups are found throughout the first world, but at this writing are hard to come by in developing countries.

GA is run along traditional 12-step lines, analogous to Alcoholics Anonymous, with similar groups for relatives. These can be very worthwhile when the pathological gambler himself refuses help. Cross-addicted pathological gamblers often attend 12-step groups related to other forms of substance abuse (principally alcohol and cocaine) and indebtedness.

In a GA meeting, clients are encouraged to confront their problems in the company of others thoroughly acquainted with every turn and twist of chasing. Sponsors can be provided. Many new members eventually become sponsors themselves. Organizing financial restitution is a crucial aspect of GA work. Unfortunately, GA dropout rate is high, especially for clients facing exposure for the first time, and when GA is the only source of help. Many PGs quit GA groups after a bail-out: under the rubric of “no debts, no problem”.

A surprising number of pathological gamblers, especially adolescents, are unknowledgeable about their illness. Education about substance abuse is common in elementary and high school, but education about problem/pathological gambling remains sparse. Several studies show that PGs willing to learn more about their disorder from an instruction manual, structured telephone interviews and GA interventions have a better likelihood of ceasing, or at least reducing their wagering.

Cognitive and behavioral strategies appeal to the pathological gambler’s inherent pragmatism. These approaches aim at decreasing or redirecting the need for action, as well as demonstrating specifically how one or another gambling milieu incites or facilitates excessive gambling. Patients are instructed in relaxation exercises to decrease tension; taught to identify gambling triggers; substitute meaningful rewards for the imagined gain of gambling, etc.

Workbooks enable gamblers to impartially assess the time and cash spent gambling, as well as the time consumed thinking about gambling. Other helpful CBT methods include stopping techniques, structured

motivational sessions which encourage the PG to think ahead and explore future goals, strategies which help the PG recognize the irrationality of his particular gambling fallacies.

Evidence-based research suggests that a cohort of pathological gamblers will be more likely to accept therapy of some sort, or stay in therapy longer, when a therapist does not insist on total abstinence, at least early in the game. Indeed, some gamblers become abstinent once they get a clear understanding of their dysfunctional behavior.

It may be necessary to halt a debilitating gambling cycle by hospitalization, taking the PG out of his terribly stressful milieu. With remission of more serious symptoms, an effective outpatient treatment plan can be organized. At this time there are more OPD programs for pathological gamblers, but still relatively few inpatient units devoted to work with pathological gamblers.

A bingeing or end-stage PG will usually be placed in a standard medical hospital setting, or shoe-horned into inpatient drug and alcohol programs. Intriguingly, a hospitalized pathological gambler may suffer initially from physiological as well as emotional distress analogous to a hospitalized drug addict undergoing detoxification.

Given the PGs poor health habits, many clients require a complete physical and neurological examination, appropriate laboratory studies, and thorough evaluation of nutritional and exercise status beyond evaluating and treating their chief physical disorder(s).

Gambling and the Law

Federal, state, and local legal systems have traditionally been oriented toward the punishment of pathological gamblers, rather than their rehabilitation. According to one PG's description of a court appearance:

> . . . (the judge) asked me why I do it, and I told him that I have an uncontrollable urge to gamble. And he said, Ah I don't believe that. That's really a sickening thing. If you're a junkie with needle marks, you'll get away with it, or an alcoholic staggering in front of him. They don't treat degenerate gambling like any kind of sickness or anything. I don't think they ever will.

Approximately 60 percent of pathological gamblers commit illegal acts over the course of their illness. As noted, the majority of PG crimes are nonviolent—bouncing checks, financial scams, prostitution. Earlier in the course of his illness, a PG under arrest is unlikely to be incarcerated, but bailed out and quickly returned to the street where he is likely to relapse again. Incarceration itself can paradoxically become a disguised bail-out, forestalling restitution or eliminating it altogether.

Experts in the field have repeatedly told legal authorities that rather than jailing most PGs—particularly the majority who do not demonstrate antisocial pathology—these offenders should be offered the option to discontinue gambling, make restitution, and undergo some form of

treatment as a condition of strict probation. Several projects modeled after drug-related diversion programs have been set up across the United States, with promising results. Funding is a recurrent problem.

At this writing, most states have set up agencies and programs to deal with pathological gambling. The sum set aside for counseling derived from lotteries, taxing race-tracks and casinos is still relatively small, compared to cash committed to drug and alcohol treatment programs.

Casinos which once refused to admit that pathological wagering ever existed have begun to fund research into the prevention and treatment of the illness. Some casinos allow compulsive gamblers to set limits on their betting in advance, or even to bar entrance to a PG when so requested. Casino employees are also being taught to identify out-of-control gamblers and give information about sources of help. Unfortunately, some of these efforts turn out to be mere window dressing. It is exceedingly difficult to create programs to convince the clientele to bet with your head, not over it, while developing ever more creative methods of separating punters from their money.

Attempts to outlaw gambling by local, state, and federal authorities have proven no more successful than the prohibition against selling alcohol during the Roaring Twenties. Protests from antigambling groups often receive only token attention. Across America, straightened economic conditions and formidable lobbying by a flourishing gaming industry are chief causes of nullification of antigambling legislation as more casinos rise.

Internet poker was outlawed several years ago in the United States after cheating scandals and embezzlement schemes were uncovered. No such prohibitions have taken place elsewhere in the world. At this writing, online poker and sports wagering have been reestablished in Nevada, New Jersey, and Delaware. Over the next several years, most states are expected to follow in reestablishing online casino and sports wagering.

THE FUTURE

The Future for the Gambling Industry

The business of gambling has always been fraught with risk *sui generis*. Sponsors of illegal gambling enterprises, from floating crap games to unlicensed casinos, traditionally face the possibility of prosecution by authority as well as criminal harassment for protection. Legal gambling business is exquisitely sensitive to the fickle winds of fashion and the fashionable; the pressure of antigambling forces; and the vicissitudes of socioeconomic circumstances. War is no friend to gambling, except among the warriors.

“Roll The Bones” tells a fascinating and often dispiriting tale of the rise and fall of gambling meccas over the past three centuries. During the past

70 years, Las Vegas casinos have undergone several boom and bust cycles. As this is written, Atlantic City, which once posed a major threat to Las Vegas, is floundering.

The proliferation of brick-and-mortar casinos, notably tribal enterprises, poses a challenge to the gaming industry. Although there are more gamblers than ever, they have more places to gamble, diminishing the chances of small as well as large casinos to stay winners.

Today the greatest gambling industry profit is scored in Asia, particularly in Hong Kong, Singapore, and Macau. The latter's mega-casinos have generated higher yearly returns than Las Vegas since 2008. The Asian boom is linked to rising middle-class prosperity across the Pacific Rim, as well as the preference of regional whales to wager colossally closer to home. Japan's pachinko machine parlors continue to generate cascades of yen each year (the government, deriving much coin from the parlors, has gotten around its own gambling ban by declaring that pachinko is a painless pastime).

Punters will always seek out real gambling locales, but one predicts that the greatest future profit for the industry will be reaped from the Internet virtual casino, especially with the return of the American poker and sports network to the scene.

The Future for Problem/Pathological Gamblers

Ominously, millions of preteens and children visit Internet sites which—intentionally or otherwise teach them the rudiments of gambling in games where wagering wins prizes instead of cash. One wonders how many of these youngsters will become serious adult recreational gamblers, and for some tomorrow's problem/pathological gamblers.

Even were gambling opportunities to escalate only modestly over the next decade (unlikely), and the world population increases modestly (also unlikely), one forecasts that the silent public health crisis caused by problem and pathological gambling will grow even more dire, as more of the afflicted make their debut. PGs have a higher mortality rate than ordinary citizens, but they won't perish fast enough to mitigate disaster.

The avalanche of future PGs certainly could be reduced by a combination of truly useful legal measures, for example, mandating more diversion and treatment programs for gamblers charged with minor criminal offenses; increased public and physician awareness; and the development of more effective treatments.

Educating the public and health professionals about pathological gambling has increased substantively over the last 20 years. Film, broadcast, print, and Internet media devote more time and space to the problem. One lauds the work of groups like the National Council of Problem Gambling (see next section). Nevertheless, it is disheartening to discover that many people, lay and professional, still denounce the sick

gambler as a reprobate.

Research into the cause and treatment of gambling problems has increased dramatically, and will continue. It is possible that neurobiological study will yield more effective medications. Perhaps future science will engineer a better fit between the patient and medication.

One predicts there will be refinement of existing cognitive/behavioral strategies, as well as the development of new techniques. Marital, family, and group therapies including GA hopefully will flourish with heightened public awareness.

Brief talking therapies will continue, and perhaps new strategies will be developed. Prolonged psychoanalytically oriented individual treatment will unfortunately diminish, if only because most insurance schemes do not support it in the United States and worldwide.

The chronic obstacle for treatment success, now and in the future, remains the afflicted gambler's formidable resistance to getting help. The number of PGs entering any form of therapy is still scant. A prospective patient frequently presents in the advanced stage of the disorder which is extremely difficult to treat, a situation which one surmises won't change much in the near future. Early identification of fledgling PGs more likely to accept treatment would seem to be vital, but it is uncertain whether significant programs toward that end can be developed in this country and across the world, or whether the average PG-to-be would accept them.

The National Council on Problem Gambling

The National Council on Problem Gambling (NCPG) is a nationwide advocacy organization for pathological gamblers and their families. Membership includes former PGs, mental health practitioners, researchers, lawyers, and legislators. NCPG is an invaluable source of cutting-edge information about every aspect of gambling. It also maintains a hotline for PGs, relatives, and friends which offers immediate help and provides information about gambling treatment across the United States. Address: 730 1st Avenue NW, Washington, DC 20001. www.ncpg.org. Hotline: 1.800.522.4700.

CLINICAL VIGNETTES

Internet Gambling. Jim, a 19-year-old freshman, was introduced to Internet poker by his roommate. Initially he played in free games. He quickly graduated to low-stake cash games. After winning several hundred dollars, he thought he was ready to take on big-league high-stakes play.

By midterm, Jim had lost several thousand dollars, and was spending most days on his computer, playing several games simultaneously. He neglected studies, slept through examinations; ate irregularly and poorly.

He quickly exhausted the money his parents sent every month; maxed out several credit cards: and began borrowing heavily from friends. Eventually he was placed on academic probation and contemplated suicide. Luckily, a friend succeeded in getting him to go to the student health center, and he was able to get straight through individual therapy and GA membership.

Gambling in the Elderly. Fred, a 79-year-old Air Force sergeant, moved to Las Vegas after retiring in his mid-40s. He and his wife were easily able to afford a small house 10 miles out of town. For the next 20 years he worked as a casino security officer, then retired.

Fred lived comfortably for another decade, until his wife died from breast cancer. He was inconsolable, hid his grief beneath a stoic facade and became increasingly depressed. He had never gambled much, and had only visited the Strip for the occasional dinner or show.

To alleviate grief and boredom, he began working again as a part-time guard at the casino. Its glitter and excitement made him feel better; he found the warbling of slot machines was peculiarly soothing. He began playing a few coins before returning home. Within a month, he was spending every evening after work, and during the weekend at the slots. He enjoyed the company of other devotees, swapping stories about hot machines and big cashouts.

At first Fred hardly noticed that his hobby was cutting into his finances. He told himself that he was only betting quarters, and that his wife would want him to enjoy himself. But within 6 months he had poured his entire savings into the machines. He could not meet bills cut back on food and other essentials.

He realized how sick his gambling had become when he began thinking he might sell the home his wife loved so much, to support his habit. He joined GA, and subsequently exchanged his job for work at a supermarket near his house. “I never believed this could happen to me”, he says, “just as sure as I am it could happen tomorrow.”

A Classic Pathological Gambler. Gerry is a 35-year-old former auto dealership owner. Two of his uncles were compulsive gamblers, and his paternal grandfather had been hospitalized for bipolar illness. He began playing poker and going to the racetrack when he was 15. He dropped out of college after a few months and became a car salesman. Soon he was promoted to showroom manager, then went out on his own. By 32, he was a multimillionaire owner of a dealership chain, happily married with two children.

Throughout his life Gerry continued to gamble frequently. He was a successful weekend sports bettor; as well as a consistent winner in weekly

gin rummy, poker games, and occasional jaunts to Las Vegas and Atlantic City.

Gerry had always wanted a third child. His wife developed eclampsia, the baby was premature, and died of respiratory failure. Gerry became depressed for the first time in his life. He visited casinos more often, gradually increased his bets at blackjack and craps. His card games became boring—the action was zilch. He began frequenting a mobbed-backed local poker parlor which featured mega-stake action.

Over several years, Gerry slipped into a typical gambling spiral. He accumulated several million dollars in debts invaded his personal bank accounts, including his children's college money. He maxed out all his credit cards, began borrowing from loan sharks at exorbitant rates. He thought of killing himself in a car crash so that insurance would take care of his family.

Gerry's illness was exposed when his Porsche was repossessed one Sunday morning. His wife threated to divorce him when she learned the extent of his losses. But her wealthy father intervened and bailed him out. He swore never to gamble again, entered GA and within two months resumed chasing. Over the next 10 years, Gerry endured four more episodes of recovery and relapse. His wife left him; he lost his dealerships; and was forced into bankruptcy.

He finally signed himself into a medical center with a pilot dual-diagnostic recovery program, where he was diagnosed with atypical bipolar disorder. His treatment includes GA meetings; individual and family counseling; and pharmacotherapy with wellbutrin and lamictal.

Gerry eventually reconciled with his wife and family. He started selling cars again, now lives modestly, still attends meetings regularly, but thinks of himself as a degenerate gambler who could relapse on any given day.

A Pathological Lottery Gambler. Nelly, a 35-year-old housewife, had been playing the State Lotto for a dollar every Saturday for many years. She rotated numbers based on her children's birthdays, her anniversary, home address, etc. Occasionally she won a few dollars, which she always reinvested in more tickets.

Then Nelly missed a five million dollar jackpot because she inadvertently had reversed the last two numbers of her license plate. She felt inexplicably cheated. Thoughts about her near-triumph increasingly filled her mind, and she became obsessed with "finding my way back to the right combination—it's gotta happen."

Nelly began playing Lotto every day, and soon was buying 20 tickets. She also went to church bingo games several times a week, then over weekends. She gradually distanced herself from her family and friends, and became profoundly depressed.

In order to finance her wagering, Nelly spent a thousand dollars in pin money she had carefully saved over a decade; then surreptitiously sold jewelry inherited from her mother. Overwhelmed with guilt, she finally revealed her problem to her husband as well as a sympathetic priest. She reluctantly agreed to enter treatment.

She responded well to family support, medication, individual and family therapy, but shuns GA groups because "she doesn't want to hang around with a bunch of bums every day slurping coffee. Everyone thinks the lottery is harmless, but it nearly destroyed my life."

An Incarcerated Pathological Gambler. Lonnie, an 18-year-old young man with ADHD received a 3-year prison sentence for low-level drug dealing. On the outside he dealt with the chronic boredom characteristic of that disorder by snorting cocaine and playing hearts. He considered himself a master at the game.

He dealt with the tedium of prison life by playing hearts at every chance. He knew, yet denied, that he was up against career criminals who had honed their skills over years of play. He began borrowing heavily from the loan shark of a feared gang. When he fell behind he was repeatedly beaten and raped by gang members. Traumatized both physically and psychologically, he hung himself.

REFERENCES

*Achab S, Khazzal Y. Psychopharmacological treatment in pathological gambling: a critical review. *Curr Pharm Des*. 2011;17(14):1389–1395.

Aragay N, Roca A, Garcia B, et al. Pathological gamblers in a psychiatric sample. *Compr Psychiatry*. 2012;53;(1):9–14.

Balodis IM, Kobeer H, Worhunsky PD, Stevens MC, Pearlson GD, Potenza MN. Diminished fronto-striatal activity during processing of monetary rewards and losses in pathological gambling. *Biol Psychiatry*. 2012;71(8):749–755.

*Bellegard JD, Potenza M, eds. Neurobiology of compulsive gambling. In: Ross D, Spurret D, Collins P. *What Is Addiction*. Cambridge: Bradford Books; 2010:27–52.

Bergler E. *The Psychology of Gambling*. New York: International University Press; 1957.

Billiux J, Van Der Linden M, Khazaal Y, Zullino D, Clark L. Trait gambling cognitions predicting near-miss experiences persistence in laboratory slot-machine gambling. *Br J Psychol*. 2012;103(3):412–427.

Blaszynski A. *Overcoming Compulsive Gambling: A Self-Help Guide Using Cognitive Behavioral Techniques*. London: Robinson; 1998.

Bolen DS, Boyd WH. Gambling and the gambler: a review and preliminary findings. *Arch Gen Psychiat*. 1968;18(5):617–630.

Brasfield H, Febres J, Shorey R, et al. Male batterers' alcohol use and gambling behavior. *J Gambl Stud*. 2012;28(1):77–88.

Brewer, JA, Potenza MN, Desair RA. Differential associations between problem and pathological gambling and psychiatric disorders in individuals with and without alcohol abuse or dependence. *CNS Spectrums*. 2010;15(1):33–44.

Clarke L. From impulsivity to addiction: gambling disorder and beyond. *Psychiatric Times*. 2015. Available at: http://www.psychiatrictimes.com/special-reports/impulsivity-addiction-gambling-disorder-and-beyond

Crockford DN, Goodyear B, Edwards J, Quickfall J, el-Guebaly N. Cue-induced brain activity in pathological gamblers. *Biol Psychiatry*. 2005;58(10):767–795.

Cowlishaw S, Merkouris S, Dowling N, et al. Psychological therapies for pathological and problem gambling. *Cochrane Database Syst Rev*. 2012;11:CD008937.

Dodd ML, Kios KJ, Bower JH, Geda YE, Josephs KA, Ahiskog JE. Pathological gambling caused by drugs used to treat Parkinson's disease. *Arch Neurol*. 2005;62(9):1377–1381.

Fenichel O. *Psychoanalytic Theory of Neurosis*. New York: WW Norton; 1945.

Freud S. Dostoyevsky and parricide. In: Strachey J, ed. *Standard Edition of the Complete Psychological Works of Sigmund Freud. 21:177*. London: Hogarth Press; 1961.

Grant JE, Donahue CB, Odlang Bl, Kim SW. A 6 month follow up of imaginal desensitzation plus motivation interview in the treatment of pathological gambling. *Ann Clin Psychiatry*. 2011;23(1):3–10.

Gyollai A, Griffiths MD, Barta C, et al. The genetics of problem gambling: a systematic review. *Curr Pharm Des*. 2014;20(2):3993–3999.

*Hodgins DC, Stea JN, GrantJ F. Gambling disorder. Available at www.thelancet.com, Vol. 378, 11/26/12.

Joutsa SH, Johansson J, Niemelä S, et al. Mesolimbic dopamine release is linked to symptoms severity in pathological gambling. *Neuroimage*. 2012;60(4):1992–1999.

*Kennedy SH, Sidney H, Walsh BR, et al. Frequency and correlates of gambling problems in outpatients with major depressive and bipolar disorder. *Can J Psychiatry*. 2010;55(9):568–576.

Kozlowski C. Demand for stimulation and probability preference in gambling decision. *Pol Psychol*. 1977;77(8).

*LeSieur HR. *The Chase: Career of the Compulsive Gambler*. New York: Anchor Books; 1977.

Limbrick-Oldfield EH, van Holst RJ, Clark L. Fronto-striatal dysregulation in drug addiction and pathological gambling: consistent inconsistencies? *Neuroimage Clin*. 2013;2:385–393.

Lobo JSS, Kennedy JL. Genetic aspects of pathological gambling: with shared genetic vulnerabilities. *Addiction*. 2009;104(9):1454–1465.

Meng Y, Deng W, Wang H, et al. The prefrontal dysfunction in individuals with Internet gaming disorder: a meta-analysis of functional magnetic resonance imaging studies. *Addict Biol*. 2015;20:799–808.

*Potenza MN. Neurobiology of gambling behaviors. *Curr Opin Neurobiol*. 2013;23(4):660–667.

Schreiber I, Odlaug BL, Kim SW, Grant JE. Characteristics of pathological gamblers with a problem drinking parents. *American J Addict*. 2009;18(6):462–469.

*Schwartz D. *Roll the Bones: The History of Gambling: Casino Edition*. Las Vegas: Winchester Books; 2013.

Simmel E. On psychoanalysis of the gambler. *Int J Psychoan*. 1920;6:397.

Stekel, W: The Gambler, in Disorders of the Instincts and the Emotions; transl. James S. Teslar, Vol. 2, Ch. 13, London: Williams and Norgate, 1929.

Studer B, Apergis-Schoute AM, Robbins TW, Clark L. What are the odds? The neural correlates of active choice during gambling. *Front Neurosci.* 2012;6:46.

Toneatto T. Cognitive pathology of problem gambling. *Subst Use Misuse.* 1999;34(11):1593–1604.

Tripett F. *The Suckers. Look Magazine,* 5/19/1934.

Weintraub D, Koester J, Potenza MN, et al. Impulse control disorders in Parkinson disease: a cross-sectional study of 3090 patients. *Arch Neurol.* 2010;67:589–595.

17

Posttraumatic Stress Disorder

ARIEH Y. SHALEV, M.D., AND CHARLES R. MARMAR, M.D.

INTRODUCTION

Posttraumatic stress disorder (PTSD) is classified in the DSM-5 as a Trauma- and Stressor-Related Disorder. Like other DSM-5 disorders, PTSD is defined by its clinical expression and, specifically, by four groups of symptoms that develop following a person's exposure to a traumatic event. PTSD diagnostic criteria allow clinicians to reliably identify and diagnose affected individuals and provide a framework for research. However, they do not directly reflect the underlying neuropathological processes. Nonetheless, PTSD is best understood as a behavioral expression of neurobehavioral systems' pathology, shaped by individual, environmental, and cultural factors.

Individuals who suffer from PTSD re-live distressing instances of the traumatic event, with vivid emotional proximity and high, imperative intensity. They organize their lives trying to contain and mitigate the persistent effects of the traumatic experience. For those traumatized in a warzone, "*the war never ends*." Victims of rape, assault, or torture describe difficulties engaging and trusting other humans. Constantly re-living the trauma in the present, PTSD patients' lives become a series of effortful attempts to avoid reminders of the traumatic event. They scan the environment for threat signals, which they fearfully expect, and remain on guard, tense, restless, and exhausted.

PTSD has been diagnosed across continents and cultures. It is a common pathological outcome of a wide variety of traumatic events, from wars and disasters to individual events such as road traffic or work accidents. The specific content of PTSD symptoms (e.g., avoided situations, recurrent themes in nightmares) reflects aspects of the triggering events and thus differs from one event to another. Their core configuration of PTSD symptoms, however, is similar across traumatic events, nationalities, races, social classes and cultures, suggesting a common underlying neurobehavioral foundation.

Symptom persistence despite termination of threat, and PTSD patients' inability to regain a sense of safety are core features of the disorder and determine its "posttraumatic" signature. Another core feature is the

involuntary, uncontrollable and intense nature of symptoms. PTSD patients live under the shadow of past trauma. They are unable to distance themselves emotionally from their traumatic experiences, which are ever-present, powerful, and dominant. These features differentiate PTSD from an immediate reaction to traumatic events, and from remote memories of trauma exposure. The latter are not "re-lived" but rather remembered. They are not ever-present, but often occasional, such as anniversaries or on memorial days. They do not dominate behavior. Most PTSD patients have frequent, intrusive and distressing "flashbulb" recollection of the traumatic events during the daytime or in nightmares. They respond to remote reminders of the event by unnecessary, involuntary and uncontrollable alarm, and defensively avoid external and internal stimuli reminiscent of the trauma. They are tense, expecting harm and at the same time tired, emotionally numb and unable to enjoy previously pleasurable activities. Many express these symptoms for years and even decades.

Understanding the bio-psycho-social mechanisms leading from trauma to PTSD is central to predicting and preventing the disorder. Chronologically, these mechanisms comprise pretrauma vulnerability, trauma exposure impact and postexposure "symptom maintaining" factors. Vertically, each of these categories comprises molecular, brain circuit, neurobehavioral, and cognitive, emotional, interpersonal, and social dimensions. The etiology and the pathogenesis of PTSD are complex and multifactorial, triggered but not fully explained by the traumatic event.

DEFINITION

PTSD is a trauma and stress-related disorder, defined by the co-occurrence of reexperiencing, avoidance, negative beliefs and hyperarousal symptoms, in survivors of extreme adversity. PTSD is a common outcome of all types of traumatic events, from most horrifying and protracted (e.g., captivity and torture) to shorter events or incidents (e.g., accidents). The construct of PTSD has, in fact, replaced several event-related "syndromes" that pre-existed its definition (e.g., *"concentration camp syndrome," "war neurosis,"* or *"rape victim syndrome"*). Unlike most other mental disorders, the diagnosis of PTSD relies on associating current symptoms with a triggering traumatic life event (e.g., depression can be diagnosed regardless of its eventual onset following a life event). The association can be chronological (symptoms starting shortly after the event) or content-related (intrusive recollections of the traumatic event or avoidance of reminders).

The essential features of DSM-5 PTSD is the development of characteristic symptoms following exposure, witnessing or indirect exposure to one or more traumatic events. The clinical presentation of PTSD varies, however. Fear-based reexperiencing, emotional, and behavioral symptoms may predominate in one individual while others may

present anhedonia, dysphoria, negative cognitions, hyperarousal, irritability, or insomnia.

DSM-5 parses PTSD symptoms into four subcategories ("diagnostic criteria"): reexperiencing, avoidance, negative cognitions, and hyperarousal. To qualify for diagnosis of PTSD, a trauma-exposed individual must express a predefined minimal number of symptoms in each category. The numerous combinations of qualifying symptoms allowed by DSM-5 can accommodate significant heterogeneity in individual presentations of PTSD. Although this flexibility has been criticized as being too permissive and allowing different kinds of patients under the same diagnosis, it is actually in line with the multiple ways in which a dysfunction of specific brain circuits is expressed by individuals who have different neurocognitive, emotional, and interpersonal resources, different life history, environmental pressures, beliefs, and understanding of causality. PTSD symptoms are best understood as varying behavioral indicators of common underlying neuropathology.

The *directly experienced traumatic* events (PTSD criterion A) include, but are not limited to, exposure to war as a combatant or civilian, threatened or actual physical assault, threatened or actual sexual violence, being kidnapped, being taken hostage, terrorist attack, torture, incarceration as a prisoner of war, natural or human-made disasters, and severe motor vehicle accidents. Medical incidents that qualify as traumatic events involve sudden, catastrophic events (Table 17–1).

Table 17–1.
DSM-5 Diagnostic Criteria for Posttraumatic Stress Disorder

Posttraumatic Stress Disorder DSM-5 Diagnostic Criteria

A. Exposure to actual or threatened death, serious injury, or sexual violence in one (or more) of the following ways:
   1. Directly experiencing the traumatic event(s).
   2. Witnessing, in person, the event(s) as it occurred to others.
   3. Learning that the traumatic event(s) occurred to a close family member or close friend. In cases of actual or threatened death of a family member or friend, the event(s) must have been violent or accidental.
   4. Experiencing repeated or extreme exposure to aversive details of the traumatic event(s) (e.g., first responders collecting human remains: police officers repeatedly exposed to details of child abuse).

   Note: Criterion A4 does not apply to exposure through electronic media, television, movies, or pictures, unless this exposure is work related.

B. Presence of one (or more) of the following intrusion symptoms associated with the traumatic event(s), beginning after the traumatic event(s) occurred:
   1. Recurrent, involuntary, and intrusive distressing memories of the traumatic event(s). **Note:** In children older than 6 years, repetitive play may occur in which themes or aspects of the traumatic event(s) are expressed.
   2. Recurrent distressing dreams in which the content and/or affect of the dream are related to the traumatic event(s).
   3. Dissociative reactions (e.g., flashbacks) in which the individual feels or acts as if the traumatic

event(s) were recurring. (Such reactions may occur on a continuum, with the most extreme expression being a complete loss of awareness of present surroundings.)
4. Intense or prolonged psychological distress at exposure to internal or external cues that symbolize or resemble an aspect of the traumatic event(s).
5. Marked physiological reactions to internal or external cues that symbolize or resemble an aspect of the traumatic event(s).

C. Persistent avoidance of stimuli associated with the traumatic event(s), beginning after the traumatic event(s) occurred, as evidenced by one or both of the following:
1. Avoidance of or efforts to avoid distressing memories, thoughts, or feelings about or closely associated with the traumatic event(s).
2. Avoidance of or efforts to avoid external reminders (people, places, conversations, activities, objects, situations) that arouse distressing memories, thoughts, or feelings about or closely associated with the traumatic event(s).

D. Negative alterations in cognitions and mood associated with the traumatic event(s), beginning or worsening after the traumatic event(s) occurred, as evidenced by two (or more) of the following:
1. Inability to remember an important aspect of the traumatic event(s) (typically due to dissociative amnesia and not to other factors such as head injury, alcohol, or drugs).
2. Persistent and exaggerated negative beliefs or expectations about oneself, others, or the world (e.g., “I am bad,” “No one can be trusted,” ‘The world is completely dangerous,” “My whole nervous system is permanently ruined”).
3. Persistent, distorted cognitions about the cause or consequences of the traumatic event(s) that lead the individual to blame him-/herself or others.
4. Persistent negative emotional state (e.g., fear, horror, anger, guilt, or shame).
5. Markedly diminished interest or participation in significant activities.
6. Feelings of detachment or estrangement from others.
7. Persistent inability to experience positive emotions (e.g., inability to experience happiness, satisfaction, or loving feelings).

E. Marked alterations in arousal and reactivity associated with the traumatic event(s), beginning or worsening after the traumatic event(s) occurred, as evidenced by two (or more) of the following:
1. Irritable behavior and angry outbursts (with little or no provocation) typically expressed as verbal or physical aggression toward people or objects.
2. Reckless or self-destructive behavior.
3. Hypervigilance.
4. Exaggerated startle response.
5. Problems with concentration.
6. Sleep disturbance (e.g., difficulty falling or staying asleep or restless sleep).

F. Duration of the disturbance (criteria B, C, D, and E) is more than 1 month.

G. The disturbance causes clinically significant distress or impairment in social, occupational, or other important areas of functioning.

H. The disturbance is not attributable to the physiological effects of a substance (e.g., medication, alcohol) or another medical condition.

Specify whether:

With dissociative symptoms: The individual’s symptoms meet the criteria for posttraumatic stress disorder, and in addition, in response to the stressor, the individual experiences persistent or recurrent symptoms of either of the following:

1. Depersonalization: Persistent or recurrent experiences of feeling detached from, and as if one were an outside observer of, one’s mental processes or body (e.g., feeling as though one were in a dream; feeling a sense of unreality of self or body or of time moving slowly).
2. Derealization: Persistent or recurrent experiences of unreality of surroundings (e.g., the world around the individual is experienced as unreal, dreamlike, distant, or distorted).

Note: To use this subtype, the dissociative symptoms must not be attributable to the physiological effects of a substance (e.g., blackouts, behavior during alcohol intoxication) or another medical condition (e.g., complex partial seizures).

Specify if:

With delayed expression: If the full diagnostic criteria are not met until at least 6 months after the event (although the onset and expression of some symptoms may be immediate).

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Witnessed events include, but are not limited to, observing threatened or serious injury, unnatural death, physical or sexual abuse of another person due to violent assault, domestic violence, accident, war or disaster, or a medical catastrophe in one's child.

Indirect exposure through learning about an event is limited to experiences affecting close relatives or friends and experiences that are violent or accidental (e.g., violent personal assault, suicide, serious accident, and serious injury).

The characteristic symptoms resulting from the exposure to the extreme trauma include persistent intrusion symptoms associated with the traumatic event (criterion B), persistent avoidance of stimuli associated with the traumatic event (criterion C), negative alterations in cognitions or mood associated with the event (criterion D) and marked alterations in arousal and reactivity associated with the traumatic event(s) (criterion E). The full symptom picture must be present for more than 1 month (criterion F), and cause clinically significant distress or impairment in social, occupational, or other important areas of functioning (criterion G) and should not be attributable to the physiological effects of a substance or another medical condition.

Symptoms of intrusion (five symptoms, one needed to meet PTSD diagnostic criterion) reflect the unrelenting, unavoidable, and uncontrollable presence of the traumatic event in the person's life. Symptoms must be recurrent, intense, and marked. They evoke significant distress and, often, physiological alarm reactions (heart pounding, cold sweats). Most importantly, they are experienced and lived "here and now" and not as retrospective, distant recall of past events. PTSD intrusion symptoms differ, therefore, from other autobiographic or episodic memories in that they do not have time and space "tags"—that is—they are not perceived as having occurred "there" and "then" but rather as recurring and current. Their emotional "here and now" presence coexists with the patient's clear understanding that the traumatic event occurred in the past. At their extreme expression (criterion B3; dissociative reactions), however, they literally transport the patient back to the traumatic situation, overshadowing the sense of current reality.

Symptoms of avoidance (two symptoms, one needed) similarly reflect a direct linkage with the traumatic event. Criterion C2 (avoidance of reminders) reflects threat avoidance acquired during the traumatic event and reinforced thereafter, and is in line with the fear conditioning theory of

PTSD. Criterion C1 is even more intriguing in that it reflects an effort to avoid inner (mental) representations of the traumatic event, as if those representations are equivalents of direct exposure. In a sense, PTSD patients carry with them internalized representations of the trauma that, at the same time, they must avoid. Just being removed from a traumatic situation is not enough to quiet PTSD patients' avoidance: they avoid thoughts and memories. This is not unlike the obsession component of obsessive-compulsive disorder (OCD) where thoughts may become as distressing as actual encounters with feared objects. Studies have shown that the C2 criterion (avoidance of places and situations) tend to expand with time, such that new objects and situations are avoided (e.g., avoidance of combat situations may extend to avoiding weapons, uniforms, crowds, mail envelops reminding those used to call people to reserve duties, etc.). This "inappropriate generalization" can end by confining severely affected PTSD patients to restricted environments with "no surprises." Patients expressing C1, in contrast, have been described as constantly trying to push thoughts and images out of their consciousness and thereby paradoxically reinforcing their presence and never letting them fade away. Individuals who recover from traumatic experiences progressively forget them, rather than deliberately trying to suppress their recall.

The *negative alterations in cognitions and mood associated with the traumatic event* criteria (seven symptoms, two needed) are a new DSM-5 cluster, which did not exist in DSM-IV. It expands the scope of PTSD symptoms beyond reactions to the traumatic event into general dispositions toward life, other human beings and society. The "diminished interest" (D5), "detachment and estrangement" (D6), and "inability to experience positive emotions" (D6) criteria are remnants of one of PTSD theoretical sources, namely the loss and grief analogy, in which, during acute grief, unrelated aspects of life lose their appeal. Freudian theory suggested that the bereaved are not free to engage emotionally, until they mourn their loss. Those symptoms mark a profound dysphoria and anhedonia, and overlap with depression symptoms. The "inability to recall" criterion (D1) follows a long historical thread starting with the construct of "repression" (of traumatic memories), but also supported by a more modern construct of "failure to encode" traumatic experiences in memory at the height of their occurrence. Inability to recall is one of the most difficult symptoms to ascertain clinically. Other negative alterations criteria reflect hypothetical changes in cognitive schemata and belief (e.g., sustained belief that the world is dangerous, other people untrustworthy and oneself worthless). Indeed, what characterizes PTSD negative cognition is its permanence: the "world is *always* dangerous" (not sometimes); "I am irreversibly worthless;" unable to cope with another trauma (a construct known as "low self-efficacy"). Symptoms under criterion D are remarkably heterogeneous and not event-specific. They

nonetheless reflect the extent to which PTSD is associated with widespread modification of personal dispositions, mood, and beliefs—far beyond the direct effect of the traumatic event.

Finally the *alteration in arousal (E criterion, five symptoms, two needed)* combines an unrelenting scanning of the environment for threat signals (hypervigilance) with irritability, insomnia, loss of concentration exaggerated startle and reckless behavior. As with the Negative Alterations criterion, symptoms may begin with the traumatic event or worsen after it, but do not have direct link with the content of event or refer to aspects of the event. Frequently observed in PTSD patients, the resulting irritability, outburst of anger, restlessness, and eventual violence are often directed at family members and close relationship and further enhance the patient's isolation and loneliness.

Clinicians using DSM-5 diagnostic criteria for PTSD should be aware of the syndrome's inherent heterogeneity, and that PTSD is not the only response to traumatic circumstances. Clinicians must rely on symptoms with a direct link to the traumatic event when inferring a trauma-related disorder. Research shows that despite their heterogeneity, PTSD symptoms tend to change as one block, across criteria (e.g., after successful interventions or during spontaneous fluctuations). Relative to DSM-IV, DSM-5 heightened the bar for identifying PTSD and lowered the prevalence of the disorder in studies published to date. Furthermore, it identifies slightly different subsets of survivors as having PTSD. Like DSM-IV it offers a rough approximation of underlying brain processes. A formal diagnosis of PTSD clearly identifies subjects in need for help. However, it should not be used to exclude other, milder, of different distressing responses to traumatic events from receiving assistance, help, and treatment.

HISTORY

Traumatic syndromes resembling PTSD have been described and sharply debated during the 19th and 20th centuries. "Stragglers" in the American Civil War were described as "*trembling, staring into the middle distance and jumping at any loud noise.*" Rescuers of survivors of ship explosions in Toulon in 1907 and 1911 exhibited *"recapitulation of the scene, terrifying dreams, diffuse anxiety, fatigue, and various minor phobias."* World War I trench warfare survivors were labeled with "shell shock" and "war neurosis," the first suggesting an organic cause (traumatic brain injury) and the second a psychological origin. The term "combat fatigue," used by the US army during World War II implied a normal reaction to extreme demands—rather than a mental disorder. Resembling current PTSD symptoms, the manifestations of "operational fatigue," included "*Irritability, fatigue, difficulties falling asleep, startle reaction, depression, tremor and evidences of sympathetic over-reactivity,*

difficulties in concentration and mental confusion, preoccupation with combat experiences, nightmares and battle dreams, phobias, personality changes, and increased alcoholism." Signs of "traumatic neurosis" similarly included "*Fixation on the trauma, typical dream life, contraction of general level of functioning, irritability and proclivity to explosive aggressive reactions."*

Despite clinical observations during World War I and World War II, traumatic stress disorders were not included in the first diagnostic and statistical manuals of mental disorders, DSM-I (1952–1968) and DSM-II (1968–1980). DSM-I included a category of "Gross Stress Reaction" of a maximum 2 months duration, thereby implying that, unless there is other psychopathology, reactions to major stressors are transient. Those who remain ill after a traumatic exposure were seen as having another disorder (e.g., neurosis, depressive illness). DSM-II eliminated the "Gross Stress Reaction" and did not have a diagnostic category for reactions to major stressors. Paradoxically, American psychiatry did not have a diagnostic category to accommodate war-related disorders during one of the harshest military engagements of the United States, the Vietnam War.

The introduction of PTSD into DSM-III (1980) followed a delayed realization of psychopathological sequelae of the Vietnam War. The American Psychiatric Association (APA) Task Force on Nomenclature (the DSM-III Task Force) reviewed over 700 clinical observations of psychologically disabled veterans and rape victims. It coined the term Posttraumatic Stress Disorder, organized its symptoms following Kardiner's description of "traumatic neuroses" and UCSF researcher Mardi Horowitz's delineation of "Stress Response Syndrome." Like other DSM-III disorders, PTSD diagnostic criteria was an optimized template, based on tradition and preliminary studies, and meant to be validated by empirical research. Indeed, the APA *Committee on Reactive Disorders* included PTSD in DSM-III as a *"consensus"* rather than *"empirically validated"* disorder.

Empirical findings during the first decades after the publication of DSM-III refined PTSD diagnostic criteria, and these refinements were included in subsequent editions of the diagnostic and statistical manual: DSM-III-R (1987) and DSM-IV (1994). For example, survivor guilt, initially included in DSM-III was found to be rare and excluded from PTSD diagnostic criteria. Physiological reactivity to reminders was moved from the arousal to the reexperiencing criterion. The most radical change, however, involved the definition of a "traumatic event" (PTSD criterion A). The disorder was originally seen as resulting from extreme events that were *"out of the realm of normal life experience."* Based on evidence that PTSD also occurred in survivors of less dramatic occurrences (e.g., road traffic accidents, natural disasters) the array of traumatic events leading to PTSD was extended and the universally devastating qualification was replaced by

the individually shattering one. Traumatic events in DSM-IV also included instances of witnessing trauma to others (such as witnessing an execution or body handling). To palliate for the permissive inclusion of events, DSM-IV added emotional reaction of *intense fear, helplessness, or horror* to its Traumatic Event criterion. DSM-IV also introduced a related syndrome: Acute Stress Disorder (ASD) occurring during the first days and weeks following trauma exposure and characterized by symptoms of dissociation.

Reliability and Validity of the Syndrome

Originally introduced as a disorder awaiting empirical validation, the nature and persistence of PTSD have been repeatedly validated. Early studies evaluating the construct validity of PTSD found it to be very good, and to generally match DSM-III original symptom clusters (reexperiencing, avoidance, and hyperarousal). More recent work suggested that avoidance symptoms comprise two distinct subcategories: specific avoidance of trauma memories and reminders and broader restriction or alteration of social engagement and mood. DSM-5 criterion D reflects these new findings.

Originally thought to be a normal reaction to abnormal circumstances, PTSD is currently recognized as a psychopathological consequence of trauma exposure, occurring to some individuals and sparing others. The traumatic event is seen as triggering rather than causing PTSD (a necessary but not sufficient condition). The etiology of PTSD is multicausal, involving vulnerability (risk) factors, exposure characteristic, and postexposure recovery environment at biological, neurobehavioral, and psychosocial levels. DSM-5 PTSD diagnostic criteria are symptomatic expressions of underlying neuropathology, which current research is progressively evaluating. PTSD symptoms and their communication to clinician and researchers are influenced by self-awareness, social norms, belief systems, desirability and acceptance of disclosure, and help seeking. Contrary to a prevalent late 20th century criticism, PTSD is *not* a psychosocial construct. It is a major disturbance of the central nervous system's threat sensitivity, threat detection, emotion regulation, and processing of reward, with varying clinical expressions.

COMPARATIVE NOSOLOGY

Previously classified among the anxiety disorder (DSM-III through DSM-IV-TR), research has shown that many PTSD patients do not show prominent anxiety symptoms, but rather anhedonia, dysphoria, aggression, or dissociation. These finding justified the inclusion of PTSD in a new Trauma and Stress-Related Disorders category. The latter is defined by a linkage between current symptoms and an exposure to external stressors. Other disorders in that category include ASD, Adjustment Disorder,

Reactive Attachment Disorder, and Disinhibited Engagement disorder.

Contrasting with PTSD patients, trauma survivors who do not develop PTSD may have initial symptoms, and be occasionally reminded of past trauma (e.g., at "anniversary" dates) but their lives are not dominated by the imperative presence of the traumatic experience. PTSD, therefore, is a distinct pathological consequence of trauma exposure and subsequent disorder maintaining factors such as lack of social support and repeated exposure to stress and the traumatic event is best conceived as a *necessary but not sufficient* cause of PTSD.

Most trauma survivors do not develop PTSD. Moreover, traumatic events can accelerate or exacerbate other DSM-5 disorders, such as major depression or substance misuse. When trauma survivors meet both PTSD and Depression diagnostic criteria, both disorders can be diagnosed as comorbid, and clinicians have to prioritize their responses to address the most prominent, or life-threatening, condition.

EPIDEMIOLOGY

Prevalence of PTSD in the Community

Trauma exposure is common, with estimates of lifetime trauma exposure in the United States ranging from 50 to 89 percent. The most prevalent types of traumatic events include physical or sexual assault (52 percent) being involved in an accident or fire (50 percent). Lifetime exposure to multiple events is common. Despite the high prevalence of trauma exposure, only minority develops PTSD. The conditional prevalence of developing PTSD varies widely by trauma types: from 65 percent for men; 46 percent for women following rape/sexual to 2 percent for men and 22 percent for women following physical assault. The conditional prevalence of PTSD following accidents is similar for the two genders. Approximately 17 to 33 percent of people with acute PTSD go on to develop chronic PTSD, with survivors or interpersonal violence taking longer to remit. A minority of trauma-exposed individuals experience delayed onset of PTSD. Most individuals with PTSD (>50 percent) do not seek treatment. Furthermore, only 58 percent of individuals with PTSD in treatment are seen by a psychiatrist, clinical psychologist, or other mental health professional. Importantly, more than a third of individuals with PTSD have been found to have persistent PTSD (even for many years), regardless of treatment received.

The National Comorbidity Survey Replication (NCS-R) surveyed a nationally representative sample of 9,282 Americans aged 18 years and older. PTSD was assessed among 5,692 participants, using DSM-IV criteria. The NCS-R estimated the lifetime prevalence of PTSD among adult Americans to be 6.8 percent. Current past year PTSD prevalence was estimated at 3.5 percent. The lifetime prevalence of PTSD among men was

3.6 percent and among women was 9.7 percent. The 12-month prevalence was 1.8 percent among men and 5.2 percent among women.

No population-based epidemiological study has examined the prevalence of PTSD among children. The prevalence of PTSD among high-risk children who have experienced specific traumatic events varies greatly and suggests that children have a higher prevalence of PTSD than adults in the general population. In contrast with children the prevalence of PTSD has been estimated in a representative sample of adolescents. Researchers in 2003 surveyed a household probability sample of 4,023 adolescents between the ages of 12 and 17. Using DSM-IV criteria for PTSD, the 6-month prevalence was estimated to be 3.7 percent for boys and 6.3 percent for girls.

Prevalence of PTSD in Veterans

The National Vietnam Veterans Readjustment Study (NVVRS), conducted between 1986 and 1988, interviewed 3,016 American Vietnam Veterans and civilian controls selected to provide a representative sample of those who served in the armed forces during the Vietnam era. The estimated lifetime prevalence of PTSD among these veterans was 30.9 percent for men and 26.9 percent for women. Among Vietnam theater veterans, 15.2 percent of males and 8.1 percent of females were currently diagnosed with PTSD at the time the study was conducted. A 25-year follow-up of the NVVRS, the National Vietnam Veteran Longitudinal Study (NVVLS) was conducted between 2011 and 2013, assessing the course of PTSD 40 years after the war. The estimate for current DSM-5 warzone PTSD was 4.5 percent, 10.8 percent for current plus subthreshold warzone PTSD, and 17.0 percent for lifetime warzone PTSD. Comorbid major depression occurred in a third of veterans with current and subthreshold warzone PTSD. Over the past 25 years, 16.0 percent of theater veterans reported a greater than 20-point increase and 7.6 percent reported a greater than 20-point decrease in PTSD symptoms. The Health of Vietnam-Era Women's Study estimated that 15.9 percent of female veterans have current PTSD 40 years after the Vietnam War. In 2008, the RAND Corporation estimated the prevalence of PTSD to be 13.8 percent in a household telephone survey of a representative sample of Iraq and Afghanistan veterans. These findings underscore the need for mental health services over many decades for Vietnam veterans with PTSD symptoms and the urgent need to provide care for a new generation of warfighters.

ETIOLOGY

Risk Factors for PTSD

In a meta-analysis of risk factors for PTSD in the community, it was reported that female gender, age at trauma, race, lower education,

childhood abuse, greater severity of trauma exposure, lack of social support, and additional life stress increased risk for PTSD. A second meta-analysis identified seven predictors: prior trauma, prior psychological adjustment, family history of psychopathology, greater perceived life threat during the trauma, lower posttrauma social support, greater emotional distress during exposure, and greater dissociation during exposure. A meta-analysis of risk factors for combat-related PTSD identified lower education, nonofficer ranks, army service, combat specialization, higher numbers of deployments, longer cumulative length of deployments, more adverse life events, prior trauma exposure, and prior psychological problems. Various aspects of the trauma period also constituted risk factors including higher levels of combat exposure, discharging a weapon, witnessing someone being wounded or killed, severe trauma, and deployment-related stressors. Another study reported that killing enemy combatants, prisoners of war, and civilians in the warzone increased the risk of PTSD.

Genetics and Risk for PTSD

It is estimated that 30 percent of the variance in the risk for PTSD is determined by genes alone. A study of Vietnam veteran twin pairs reported that monozygotic twins of veterans with combat-related PTSD had greater mood disorder symptoms than monozygotic co-twins of combat controls or dizygotic co-twins of veterans with PTSD. The findings suggest a shared genetic vulnerability to PTSD and mood disorders. Researchers who studied 222 monozygotic and 184 dizygotic twin pairs reported PTSD symptoms to be moderately heritable with the remaining variability accounted for by unique environmental experiences. A number of common genetic variants have been associated with PTSD including polymorphisms in FKBP5, PACAP1, COMT, DRD2, GABA alpha-2 receptor, G protein signaling 2 (RSG2), an SNP in an intergenic region of the fourth chromosome and an estrogen response element on ADCYAP R1. In addition, the s/s genotype of the serotonin transporter gene interacts with childhood adversity to increase PTSD risk.

DIAGNOSIS AND CLINICAL FEATURES

As mentioned previously, the diagnosis of PTSD relies on identifying a traumatic event and confirming the presence of symptoms belonging to four diagnostic criteria: intrusion, avoidance, negative alteration in cognition, and hyperarousal.

To a degree PTSD symptoms, despite their apparent heterogeneity, reflect an inner logic and theoretically meaningful functional interrelations. Traumatic events remain lively and ever distressing, they are experienced as personal defeat, surrender to overwhelming stress, or failure to react

properly, hence their endlessly distressing rumination and lack of emotional resolution. Because they remain extremely painful, utterly uncontrollable and persistent, individuals with PTSD avoid stimuli that can provoke this re-living of them, including inner states of mind, thoughts, emotion, and bodily experiences. Fending threat, either external or internal, becomes an imperative priority and leads to always remaining on guard, ever more vigilant, attending to and seeking signs of potential threat, and unable to allocate attentional resources to dealing with real-life challenges. Ultimately, PTSD patients are defeated by the intensity, the uncontrollability and the timelessness of their posttraumatic distress and their Sisyphean attempts to escape their predicament, or regulate their emotions, and thus likely pessimistic, moody, and distant.

Importantly, posttraumatic psychopathology is not limited to PTSD, and milder and transient reactions are more frequent than lingering PTSD. Even within PTSD, the disorder can be expressed in a myriad of ways and with different configurations of its qualifying symptoms. PTSD, therefore, should be thought of as a prototypical reaction, a psychobiological trap in which what often appears as attempts to elaborate and overcome traumatic experiences (e.g., intrusive thinking, recall and reappraisal of traumatic experiences, self-protection and adequate attention to existing stressors) becomes senseless, timeless, intolerable, and inescapable.

Some researchers have suggested that PTSD is best construed dimensionally rather than categorically. According to this view, PTSD merely represents the endpoint of a stress response continuum, whereas partial PTSD would represent a subsyndromic degree of severity somewhere between no or mild response and full-blown PTSD. Recent work, however, suggests that impairment correlates do not smoothly vary as a function of increasing symptoms; those victims who meet full criteria for PTSD are more impaired than those who merely qualify for partial PTSD, a finding more consistent with a categorical rather than a dimensional construal of the illness.

Controversies Concerning PTSD's Validity and Social Origin

PTSD was born to an age of increased sensitivity to social justice, human rights, protective networks, and duties of care. From its inception, therefore, the disorder was associated with advocacy and a sense of mission among many of its proponents. PTSD has also been used extensively to identify and infer psychological damage and claim compensation. Surpassing its intended role as a descriptor of a mental disorder, PTSD has become a social, cultural, and economic phenomenon.

Additionally, because of its captivating simplicity, PTSD was quickly reified. Rather than a descriptor of an underlying condition, PTSD became the condition itself. Referring to PTSD as a real and distinct entity overlooks the difference between a set of symptoms and an underlying

condition. This distinction is too frequently forgotten.

By another DSM rule, complex clinical presentations that follow trauma are subsumed as comorbidities, that is, coexisting but distinct disorders. PTSD and Major Depression, for example, overlap in over 40 percent of the cases at any given time, and cooccur in 95 percent of the cases in a lifetime perspective. PTSD and Major Depression are nonetheless seen as distinct entities, particularly when it comes to defining research groups and/or licensed therapies. The amount of information lost by anchoring research and practices to categorical entities is hard to estimate.

Similarly, in a typical posthoc attribution error ("after this, therefore because of this"), the association between PTSD and the traumatic event is quickly perceived as causal. Part of the popularity of PTSD is due to its symbolizing such highly desirable, simplified, intuitive causal attribution.

Finally, advances in the clinical management of PTSD have not followed the accelerated path of neurobiological discovery. PTSD remains chronic, disabling, and frequent. The prevalence of PTSD among US veterans of recent conflicts, for example, matches seen after the Vietnam War. Not unrelated to the latter, the US Institute of Medicine's (IOM) 2008 report used uncompromising language to describe this state of affairs. "The studies conducted over the nearly three decades since Diagnostic and Statistical Manual of Mental Disorders (DSM) adoption of the PTSD definition do not form a cohesive body of evidence about what works and what does not."

PATHOLOGY AND LABORATORY EXAMINATION

PTSD epitomizes *a mental disorder occurring at the junction between external reality and a responsive brain:* it is a maladaptive outcome of a host-environment interaction gone wrong. Several clinical features of PTSD inform the underlying central nervous system (CNS) mechanisms: The disorder emerges after exposure to traumatic events, and therefore must centrally involve *new learning and CNS plasticity*. Its core symptoms of reexperiencing and avoidance directly involve reminders of the traumatic event, which implies *acquisition, consolidation, and conservation of aversive memories*. The presence of hypervigilance, hyperarousal, and emotional alterations extend the range of symptoms beyond responses to reminders, and suggest an *overgeneralization of aversive learning and dysregulation of emotions*. Acute PTSD symptoms often emerge shortly after trauma exposure and do not remit in survivors who develop PTSD, making PTSD a likely "disorder of recovery" and evoking *alterations of fear extinction and inability to process safety signals*. PTSD patients remain exceedingly reactive to ongoing life stressors and thus show *altered perception of threat signals and heightened threat sensitivity*.

Not surprisingly, studies of the neurobiology of PTSD have been extremely productive. To make sense of that abundant literature this

chapter addresses three layers of findings: (a) finding showing how the CNS mediates known PTSD symptoms and related behavior, (b) finding of alteration in neuroregulatory biological systems (e.g., neuroendocrine, neuropeptide, inflammatory) that moderate CNS responses, and (c) finding of basic neurobehavioral alterations, genetic endowment, and epigenetic regulation, often present before the traumatic event, and increasing CNS' vulnerability to develop PTSD. It then dedicates a special section to brain-imaging studies of PTSD.

In reading that literature, one should bear in mind that while several biological perturbations have been associated with PTSD, none has proven to be characteristic of the disorder in all affected individuals or account for the large proportion of the disorder's causation. Such a gap between group-based trends and finding and individual causation characterizes complex and multicausal disorders, in which there are many pathways to an outcome and individual-specific convergences of risk factors. To fully translate the complex biology of PTSD into practical prediction, diagnostics, or therapeutics, one must resort to advanced bioinformatics models, which are currently being developed. This section covers the existing evidence, and the few newly discovered links between layers of biological alterations and environmental pressures.

CNS Mediation of PTSD Clinical Manifestations

Several brain circuits have been implied in the mediation of PTSD clinical manifestations. Those include *fear learning, emotion regulation, executive control,* and *contextual learning*. Experimental methods used to explore these circuits encompass psychophysiology, neurocognitive assessments, and brain imaging.

Fear Conditioning Theory of PTSD. This model equates the acquisition and the maintenance of PTSD symptoms with the consolidation and extinction of conditioned fear responses. Within this model, the traumatic event is the conditioning stimulus (CS), the immediate responses to trauma are unconditioned responses (UCRs), and stimuli (sights, sounds) present during the traumatic event acquire the ability to elicit conditioned responses (CRs) resembling the CR in the absence of real threat. A series of psychophysiological studies has buttressed the fear-conditioning model of PTSD.

CNS Circuits Involved in Fear Conditioning. *Fear conditioning* and *extinction* are adaptive learning processes. The neural circuitry that mediates these processes is essential to survival and conserved across species. It commands the body's hard-wired, involuntary, instantaneous defensive responses to imminent threat (a "species-specific defense response"). This highly conserved stress response system includes

sympathetic and parasympathetic system activation, cardiovascular and respiratory reactions, activation of the hypothalamic–pituitary–adrenal (HPA) axis, defensive fight, flight, and freezing behaviors, and changes in information processing. It is orchestrated by the central nucleus of the amygdala, a brain structure that simultaneously connects to critical brain areas responsible for response execution. The activity of the central nucleus is modulated by layers of midbrain nuclei capable of filtering and evaluating threat stimuli (such as the basolateral amygdala [BLA]) and further by cortical and subcortical structures, such as the hippocampus, insula, and prefrontal structures. Each of these modulatory layers is highly plastic and therefore subject to new learning throughout life. PTSD is thought to involve new learning in these modulatory systems as follows.

During exposure to extreme threat, previously neutral stimuli overlapping in time with threat stimuli pass through the thalamus to the basolateral nucleus of the amygdala, which activates the central nucleus of the amygdala to enable defensive responding. The association between threat or "unconditioned" stimuli (US) and the previously neutral but currently conditioned stimuli (CS) is made in the basolateral nucleus of the amygdala. Once formed this association enables an exposure to the CS to trigger a defensive response. The magnitude of the defense response is modulated by brainstem serotonergic, noradrenergic, and dopaminergic neurons in the ventral tegmental area, dorsal and median raphe nuclei, and locus coeruleus. These monoamines affect both the magnitude of UCRs and the consolidation of US–CS associations. The balance of neurotransmitters and neuromodulators activated during unconditioned response also influences the extent to which contextual stimuli, more loosely associated in time with the US, may contribute to later overgeneralization of threat stimuli.

Activity within the amygdala is under inhibitory control of the prefrontal cortex (PFC). Specifically, glutamatergic neurons from the PFC activate inhibitory interneurons within the BLA, which in turn suppress BLA outputs to the CE. The monoaminergic effects are further modulated by neuroactive steroids and neuropeptides released by the adrenal gland during stress that further magnify, mitigate, or otherwise shape both fear acquisition extinction and recall.

Consolidation, Extinction, Extinction Recall, and Reconsolidation. Associations encoded in the amygdala during traumatic events tend to be robust and resistant to change. Their behavioral expression, however, is subject to extinction, which typically occurs when the conditioned stimulus is no longer associated with threat in a safe environment. Neuronal mechanisms underlying extinction include PFC inhibitory input to the BLA. Extinction involves active learning and as such is subject to recall and forgetfulness. Prolonged stress after the traumatic event can compromise

the PFC ability to efficiently extinguish fear memories, whereas soothing interpersonal contacts and a safe environment may enhance it. In the aftermath of traumatic events, therefore continuous, stress, and absence of soothing contacts may decrease the likelihood of efficient extinction and contribute to risk of developing PTSD.

Psychophysiological Studies

Bearing different names (e.g., Soldiers Heart; Shell Shock) the historical equivalents of PTSD have been associated with "physical" or "physiological" symptoms. Studies performed during World War I (1917–1918) recorded exaggerated heart rate responses to gunfire, sulfuric flames, epinephrine injection, and CO_2 inhalation among war veterans with combat stress reactions. Abraham Kardiner, whose description of War Neurosis inspired DSM-III definition of PTSD, defined these conditions as "physio-neuroses," involving bodily alarm responses to stressful combat reminders. Shortly after the inclusion in DSM-III, PTSD was equated with a "conditioned emotional response," similar to those observed in animal experiments of fear conditioning. Studies performed established the script-driven imagery protocol as standardized and reliable procedure to elicit physiological (heart rate, skin conductance, muscle tension) responses to individual trauma memories, and showed that these responses were exaggerated in trauma survivors with PTSD, but not in those who recovered. Research done in Jerusalem in 1992 showed abnormal auditory startle responses in PTSD.

Psychophysiological measurements added an objective dimension to PTSD. As such, it has helped deflect some of the initial skepticism regarding the disorder's validity. They have also opened ways to experimentally elicit PTSD symptoms, which was later used in functional brain-imaging "symptom provocation" experiments (refer to *Functional Brain Imaging*). Psychophysiological measurements have also shown that PTSD patients fail to extinguish conditioned fear responses, which is in line with their inability to recover from the early responses to traumatic events. Studies have also shown that PTSD patients fail to retain extinction, once they have acquired it, and return to reacting to adversity despite having safely recovered from it. Additionally, studies indicate an association between elevated heart rate shortly after traumatic events (e.g., during emergency department admission after injury) and subsequent PTSD, an association that reflects the intensity of the UCRs during trauma exposure. Finally, decreased heart rate variability (HRV) has been robustly associated with PTSD and linked to emotional cognition and emotion regulation.

Threat Detection and Emotion Regulation

Other specific circuits implicated in PTSD development mediate *threat*

information processing, mediated by the amygdala, dorsal Anterior Cingulate (dACC) and insula cortex, and *regulatory control mechanisms* that involve the hippocampus and medial and lateral prefrontal cortex regions (dMPFC and dLPFC). These regions undergo plastic modifications upon exposure to potentially traumatic events, which in turn modify future processing of stressful encounters. The amygdala assigns rapid "threat-potential" tagging of environmental cues and its high reactivity may be a risk factor for PTSD. The amygdala-associated PFC regions and insula have been implicated in detecting stimuli saliency and directing attention biases to threat. *Emotional regulation* circuits select appropriate response to threat—including emotional response—and involve the Medial Prefrontal cortex (MPFC) and Dorsolateral prefrontal cortex (DLPFC). Evidence from neuronal network studies demonstrates segregation between a *salience network* (SN) involving regions of the insula, ACC and amygdala an *attention network* (AN) involving the DLPFC and VLPFC, and a *default mode network* (DMN) regions like MPFC and hippocampus.

Neuroimaging findings in PTSD patients found abnormalities in threat detection and emotion regulation regions with increased reactivity of the extended amygdala, insula and dorsal ACC regions, and decreased regulatory activity in subregions of the prefrontal lobe. PTSD and MDD have shown excessive reactivity to potential threat stimuli in the amygdala and insula and hypoactivation of the prefrontal cortex. Prefrontal deficits involving ventromedial prefrontal cortex (vmPFC) have been implicated in emotion regulation during changing environmental demands in PTSD, MDD, and generalized anxiety disorder. Resting state connectivity studies in PTSD similarly show reduced network coupling within the DMN, greater coupling within the SN, and decreased segregation of the DMN and the SN, suggesting an overengagement of saliency detection in PTSD. Recent studies further suggest that trauma experience alters DMN connectivity shortly after its occurrence.

Cognitive Control, Memory, and Attention

Salient PTSD symptoms include inability to concentrate, memory disturbances and difficulties planning and monitoring one's activity. Activity in dACC and frontoparietal attentional networks has been implicated in cognitive control, including performance monitoring, response inhibition, task shifting, working memory, attention, and memory. This type of regulation also disengages medial prefrontal–medial parietal DMN regions involved in internally oriented, self-directed mentation. DMN regions are overactive or underdynamic in PTSD, resulting in poorer ability to disengage from threat responses and discriminate truly threatening stimuli from remote reminders of traumatic experiences.

Modulatory Processes Underlying PTSD

Neuroendocrine Modulation. Several "stress" hormones have been implied in the etiology, pathogenesis, and psychopathology of PTSD but, none are disorder-specific and together they exercise complex and sometimes opposing neuromodulatory effects. The HPA axis and sympathetic nervous system are important stress-reactive systems and thus central to understanding the disorder. They contribute to brain information processing, reactivity, and behavior in response to threat. Briefly stated, stress pathways involving cortisol signaling via the HPA axis are both blunted and hypersensitive in PTSD, with consequent influences on the HPA axis, sympathetic nervous system, brainstem, amygdala, and potentially accounting for co-occurring diagnostic patterns (PTSD, depression, chronic pain, refractory substance dependence).

Neurosteroids. Cortisol is centrally involved in stress response, and its regulation in PTSD has been extensively studied. Schematically, a surge in cortisol secretion by the adrenals follows the initial adrenergic response to stress. This surge, than, downregulates the brain's initial adrenergic response, via negative feedback mechanism, thereby attenuating the acute response to stress and the related US–CS pairing. Early studies revealed a pattern of low cortisol output and increased sensitivity to glucocorticoid negative feedback in PTSD, but the finding has not been consistent across samples of PTSD patients, and during the acute aftermath of traumatic events. Sources of the variability in these studies include different experimental designs, population characteristics, trauma characteristics (e.g., single occurrence vs. prolonged exposure [PE] to stressors) and variation in individual genetic endowment. Differences between findings may, therefore, reflect *differences between individuals' HPA axis sensitivity,* which accounts of whole groups are unable to discern. In other words, altered HPA axis response to stress may affect some individuals (e.g., those with history of child abuse, variation in genetic endowment, or different epigenetic regulations history) but not others. It is also important to evaluate the roles of cortisol in the pathophysiology of PTSD in the context of other relevant neurohormonal responses as follows:

Neuropeptide Y (NPY) has anxiolytic and anticonflict effects in numerous animal models and human studies. It stimulates neurogenesis in the hippocampus and as such may support recovery from stress. For example, higher plasma NPY levels during distressing mock captivity training were associated with stress resilience. NPY has also several bodily trophic effects including efficient restoration of homeostasis, improved sleep quality, lower heart rate response to adrenergic stimulation and better energy balance, all of which may affect the recovery from early PTSD symptoms.

Allopregnanolone/Pregnanolone are neuroactive steroids that selectively enhance the effect of the brain principal inhibitory neurotransmitter GABA. They also have neuroprotective effects such as increased myelination, decreased apoptosis (programmed cell death) and increased neurogenesis. Studies have shown low levels of these hormones in PTSD and depression and negative correlation with symptom severity.

Dehydroepiandrosterone (DHEA) is the immediate precursor for androgen. It is secreted from the adrenal gland along with cortisol in response to ACTH. DHEA antagonizes $GABA_A$ receptors and facilitates *N*-methyl-D-aspartate (NMDA) receptor function. In humans, an increasing DHEA level under stress appears to confer psychological and neurocognitive resilience and may protect against negative health outcomes.

Catecholamines. Exaggerated noradrenergic signaling is related to deficits in frontal lobe executive function and inhibition of amygdala reactivity as well as sleep fragmentation and nightmares, which are common in PTSD. An exaggerated adrenergic response during trauma exposure and in hours following exposure may contribute to the formation of more durable emotional memories. Evidence concerning the presence of sympathetic system hyperreactivity in PTSD has been generally consistent. Noradrenergic hyperreactivity may be a pretrauma vulnerability trait in some individuals, or develop as a result of trauma exposure in others. Factors that may contribute to the increased release of norepinephrine (NE) in response to traumatic stress or conditioned reminders of the trauma include decreased expression, affinity, or function of alpha 2-adrenergic autoreceptors, or genetic or stress-induced decrements in neuropeptide Y (NPY). Factors that may enhance amygdalar reactivity may include a downregulation of intra-amygdalar synaptic GABA receptors or reductions in GABAergic neuroactive steroids such as allopregnanolone.

Illustrating the role of catecholamines in maintaining PTSD symptoms, the alpha 2-antagonist yohimbine induces extreme hyperarousal, flashbacks intrusive memories, panic, increases in heart rate and systolic blood pressure in male combat veterans with PTSD and suppressed frontal lobe metabolism to a greater degree among PTSD patients. Resulting from these observations alpha 1-adrenergic postsynaptic receptor antagonists, such as prazosin have been used for the treatment of nightmares in PTSD and evaluated as an early prevention of PTSD (see “Treatment” section).

Dopamine release in the amygdala promotes the expression of unconditioned and conditioned stress responding. Additionally, dopamine may inhibit PFC projections into the BLA and thus diminish the inhibitory control over learned fear responses and lead to increasing hypervigilance and sensitivity to environmental cues. Dopamine is also highly relevant to reward signaling in the nucleus accumbens, a process that appears to be

downregulated in PTSD (see also PTSD criterion D symptoms).

Neuropharmacological treatment and genetic epidemiological studies suggest that the serotonin (5-HT) system impacts both PTSD risk and symptom severity. Administration of meta-chlorophenylpiperazine (mCPP), which interacts with multiple 5-HT receptor subtypes, and with the 5-HT transporter to release 5-HT, provokes anxiety, panic attacks, and PTSD symptoms, including flashbacks, and cognitive changes that are reversible by mixed 5-HT1c/5-HT2 antagonists. Selective serotonin reuptake inhibitors however, do not reduce the likelihood of developing PTSD when administered shortly after trauma exposure.

Genetic Contributions to the Risk of Developing PTSD

The genetic contribution to PTSD is multifactorial, person-specific, and significantly affected by early life experiences. The inheritance of PTSD is polygenic, with no single gene variant explaining more than a small percentage of the variance. Nonetheless, a large twin study has shown significant (>30 percent) inheritance of PTSD and PTSD symptoms, but re-analyses of the same database reported substantial genetic overlap between MDD and PTSD. As in other areas of psychiatry, the translation of heritability to specific gene variants is a work in progress and the distance between genetic endowment and clinical expression is currently insurmountable. Studies to date have, therefore, focused on gene variants that affect neuronal and neuromodulatory processes leading to PTSD.

Gene variants in the *NPY and noradrenergic signaling* pathways have been found to impact physiologic processes of relevance to PTSD, such as heart rate, blood pressure, NE release, and negative emotional reactions in response to low-dose yohimbine. These sympathetic system stress reactions would be expected to increase risk for PTSD, but whether any of these latter gene variants influence risk for PTSD is yet to be examined. Similarly, to the extent that the BDNF gene affects synaptic plasticity, neurogenesis and cognition, it might indirectly affect the likelihood of developing PTSD (or remaining resilient) at the aftermath of traumatic events. Numerous genes have polymorphisms that affect either ACTH or cortisol responses to stress (e.g., polymorphisms of genes for the mu-opioid receptor, catechol-O-methyl-transferase [COMT]) and angiotensin I–-converting enzyme (ACE-I), the glucocorticoid receptor, the ACTH receptor, and corticotropin-releasing factor (CRF) and the CRF receptor.

Gene–environment interactions that affect PTSD risk have been associated with the short allele of the serotonin transporter gene promoter (SLC6A4). In a group of Rawandan genocide refugees, the SLC6A4 promoter polymorphism was found to increase the risk for PTSD development in the context of relatively low trauma load. In 582 European American and 670 African American individuals, possessing either one or two alleles of the SLC6A4 promoter, polymorphism interacted with adult

trauma and childhood adversity to increase the risk for PTSD, especially for those with high rates of both types of trauma exposure.

Clearly the list of genetic polymorphisms with potential influence on the risk for PTSD is incomplete. Nor is it likely that possession of any one of these risk factors is a determinant. It may take a combination of such factors in the absence of countering protective factors, and in the context of facilitating environmental influences, to increase the risk for development of PTSD. This point is illustrated by work showing that both childhood trauma and decreased expression of the FKBP5 are required in order to increase the risk for PTSD.

Epigenetic Changes Associated with PTSD

The epigenetic regulation of gene transcription may provide new opportunities for understanding PTSD. Epigenetic effects consist of coordinated modification of gene promoters and chromatin such that access to the gene by regulatory elements is either enhanced or limited, including lower methylation or the glucocorticoid receptor gene F1 promoter in peripheral blood cells in veterans. Of note, research has shown differences between groups of persons with and without PTSD in methylation patterns of DNA in white blood cells. A cluster of unmethylated immune system function genes also characterized PTSD, which otherwise showed a greater number of methylated genes in general. In addition, there were group differences in methylation of the dinucleotide methyl transferase (DNMT) type 3 beta and 3L genes, which are involved in *de novo* rather than maintenance methylation, suggesting that specific adaptive epigenetic capacities themselves may be altered in PTSD.

The extent to which epigenetic mechanisms, operating after trauma exposure, underlie some of the persistent effects of exposure is an exciting new perspective. Research showed that immobilization stress in rodents *reduced* expression of BDNF gene transcripts I and IV in the hippocampus in association with decreases in the level of histone acetylation of the P1 and P4 promoters. Conversely, research has shown that extinction of conditioned fear was accompanied by an *increase* in the levels of histone acetylation of the gene promoters for BDNF transcripts I and IV. Together this work suggests that activation of specific PFC circuits *in combination* with agents that facilitate epigenetic processes that enhance synaptic plasticity and stabilization of new neuronal circuits may contribute to the persistence of PTSD and recovery from PTSD.

Structural and Functional Neuroanatomy

Structural Brain Imaging of PTSD. The first reported structural abnormality in PTSD was a reduced left hippocampal volume in Vietnam veterans with PTSD. This finding had been interpreted as reflecting the

hippocampus' vulnerability to excessive stress hormones levels due to prolonged stress. Numerous studies subsequently reported left, right, or bilateral hippocampal volume reduction in PTSD. Other studies, particularly in younger veterans, have failed to replicate this finding but meta-analyses generally support the existence of reduced hippocampal volume in PTSD. Findings of smaller hippocampus volume in PTSD, however, lack disorder specificity (similar findings have been shown in schizophrenia and affective disorders). There is also a controversy regarding their relationship with PTSD. Originally thought to be a consequence of developing PTSD due to a neurotoxic effect of protracted or "allostatic" stress, a longitudinal study of hippocampal volume in early PTSD has not shown a hippocampus shrinking effect and an elegant twin study of Vietnam veterans found smaller hippocampi in unexposed identical twins of war veterans with chronic PTSD, thereby positioning smaller hippocampi as vulnerability factor, rather than a consequence of PTSD.

More recent studies reported white matter volume reductions in the *corpus callosum* and *gray matter reduction in the amygdale and the* insula. The evidence concerning *structural alterations of the amygdala* is quite weak. Two meta-analyses reached somewhat different conclusions. One found that there is overall evidence for smaller left amygdala volumes in adults with PTSD, but not in pediatric PTSD, whereas the other did not find significant volumetric alteration in either population (Fig. 17–1).

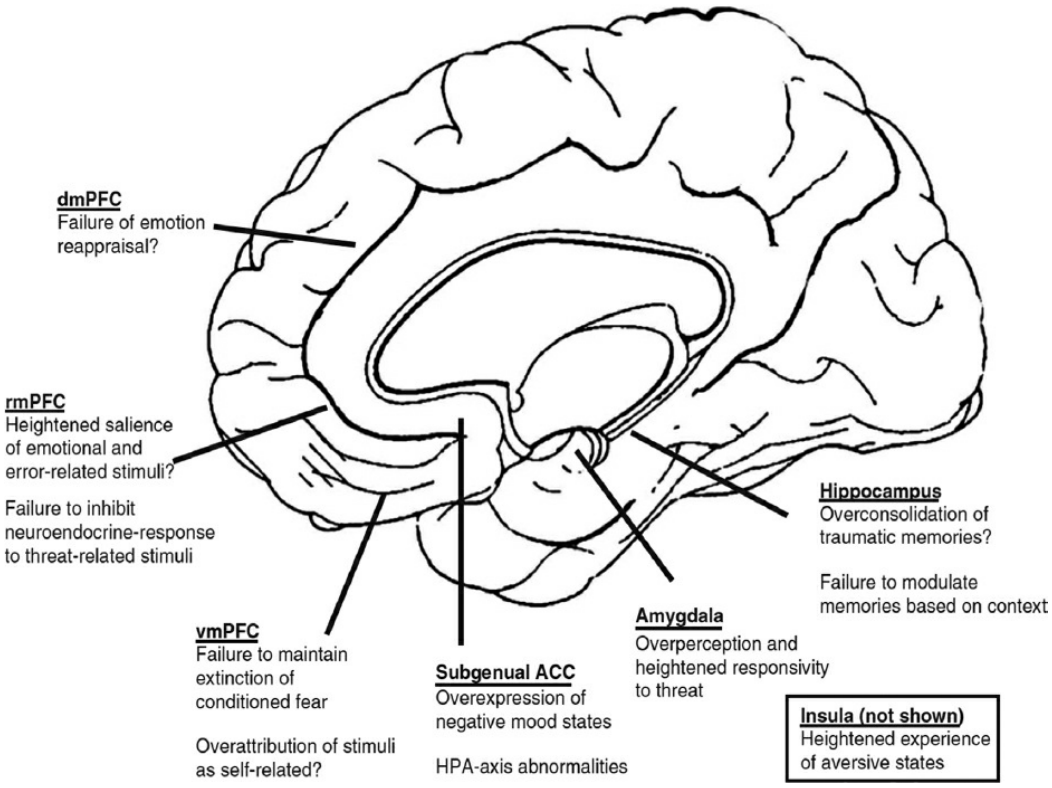

FIGURE 17–1. Neural regions implicated in PTSD. dmPFC, dorsomedial prefrontal

cortex; rmPFC, rostral medial prefrontal cortex; vmPFC, ventromedial prefrontal cortex; ACC, anterior cingulate cortex.

Several studies have pointed to abnormal structural characteristics in the *anterior cingulate cortex (ACC), including* reduced ACC gray matter volumes or white matter abnormalities of the cingulum. These reduced volumes may be particularly characteristic of the rostral and subcallosal portions of the ACC, rather than the dorsal ACC. Rostral ACC volume has been found to predict response to cognitive-behavioral interventions such that larger volumes were associated with greater symptom reduction. A study of pregenual ACC volumes in monozygotic twins discordant for PTSD suggested that reduced volumes might reflect an acquired trait of PTSD rather than a vulnerability factor.

Functional Brain Imaging. Functional neuroimaging studies of PTSD include both resting and activation studies. The latter include symptom provocation, and processing of both emotional and neutral stimuli. Within the limbic core, differences between PTSD and controls have been reported for the amygdala, and less so for the hippocampus. Outside the limbic core the vast majority of reported differences in activation between PTSD and controls occur in the posterior, dorsal and anterior cingulate, occipital and medial prefrontal cortices, and dorsolateral prefrontal cortex. The same overall pattern can be detected for neutral tasks: most abnormal activations appear to occur in the dorsal trend. There are, nonetheless several differences: the rostral ACC and subcallosal ACC do not show unique patterns of activity in neutral conditions and only the dorsal ACC appears to be recruited differently in PTSD compared with controls. There is also less consistency across studies, which may reflect the more varied tasks and smaller number of published studies.

Sleep

Up to 90 percent of people with PTSD report sleep disturbances such as nightmares and insomnia. In studies examining the physiological basis for sleep disturbances in PTSD, both *macrolevel* sleep architecture (i.e., time spent in various sleep stages overnight) and *microlevel physiological measures of* autonomic system activity during sleep (i.e., heart-rate variability, electrodermal activity) have been implicated. The frequent nightmares that are often the hallmark of PTSD, are thought to occur in rapid eye movement (REM) sleep. Additionally, REM has also been shown to play a role in the maintenance of fear memories. Sleep abnormalities also characterize the early phases of PTSD: While it is unclear why some survivors develop PTSD and others do not, persistent sleep abnormalities, which are present in upwards of 66 percent of those who develop the disorder has been suggested as a contributing factor. More specifically,

recent studies have shown that sleep disturbances, measured prior to exposure to traumatic stress, predict the development of PTSD.

Summary of the Biology of PTSD

Arguably the most intriguing feature of PTSD is its persistence, that is, its nonextinction despite the absence of further stress. Longitudinal studies have confirmed a frequent pattern of immediate expression of PTSD symptoms followed by recovery in most of those exposed, and persistence of symptoms in those who "develop" chronic PTSD. The conditioned fear analogy similarly predicts extinction in the absence of reinforcement, which does not occur in PTSD. A neurobiological account of PTSD, therefore, must also consider the disorder's maintaining factors. Several theoretical models account for the persistence of PTSD, including a *kindling hypothesis,* according to which patterns of pathological neuronal coactivation become "entrenched" by repeated use; an *allostatic stress hypothesis* presuming a "wear and tear" of CNS emotional-control systems (e.g., the hippocampus) by chronic stress and what might be referred to as a *subsystem clash* or *truncated response* view, according to which one brain subsystem blocks another from achieving the full processing of its inherent tasks.

DIFFERENTIAL DIAGNOSIS

PTSD cannot be diagnosed in the absence of exposure to a traumatic stressor. Indeed, it is the memory of trauma that generates and knits together the otherwise disparate symptoms of PTSD, many of which overlap with other mood and anxiety disorders. Hence, identifying the etiologically significant trauma is the crucial starting point for differential diagnosis. The symptoms must persist for at least 1 month following exposure to the stressor, thereby distinguishing PTSD from acute stress disorder, whose symptoms, by definition, cannot persist for longer than 1 month. PTSD arises in response to a traumatic stressor, whereas the diagnosis of adjustment disorder is suitable for people who experience anxiety, mood, or other symptoms following exposure to a nontraumatic life events, such as loss, separation, or illness.

PTSD comprises two kinds of symptoms: Event-related and non–event-related responses. Intrusive thoughts, nightmares, flashbacks, and avoidance of reminders are event related and thereby facilitate differential diagnosis by distinguishing PTSD from other syndromes that develop or are exacerbated after exposure to traumatic stressors. For example, OCD intrusive thoughts concern current and future threats, such as contamination. Likewise, generalized anxiety disorder concerns envisioning and worrying about threats that lie in the future.

PTSD non–event-related responses lack direct reference to the

traumatic event. These include loss of interest in activities, a sense of foreshortened future, difficulty in experiencing loving feelings toward others, exaggerated startle response, difficulty sleeping, difficulty concentrating, difficulty sleeping, and irritability. These symptoms present two challenges during differential diagnosis. First, the assessor must determine that these symptoms followed in the wake of trauma rather than predating exposure to trauma. Many people have sleep problems, are irritable, and so forth, and one cannot automatically assume that the presence of these problems signifies PTSD. Indeed, a person exposed to a trauma who exhibits irritability, difficulty concentrating, and insomnia must not be assumed to have "PTSD symptoms" unless these difficulties postdate the trauma or are exacerbated by the trauma, and are accompanied by trauma-related symptoms. Automatically interpreting insomnia in a trauma victim as a "symptom of PTSD" makes no more sense than calling coughing a "symptom of bacterial pneumonia" in the absence of a confirming culture.

The other challenge arises when assessors must determine whether apparent PTSD symptoms are direct consequences of memory for trauma or they signify the presence of another disorder. For example, sleep disturbance, irritability, difficulty envisioning one's future, loss of loving feelings, and loss of interest in activities are common symptoms of major depressive disorder—a condition often comorbid with PTSD. Likewise, other symptoms, such as hypervigilance and startle, may occur in other anxiety disorders as well. As a rule, therefore, PTSD requires a fuller presence of typical symptoms and a substantiated link in both content and timing to trauma exposure.

PTSD seldom occurs by itself, and the presence of other DSM-5 mental disorder does not preclude PTSD. For example, researchers studying Vietnam veterans found that 84 percent of outpatients as well as inpatients with PTSD met criteria for comorbid conditions, often alcoholism and antisocial personality. In another study of help-seeking veterans, researchers diagnosed current major depression, panic disorder, bipolar disorder, and social phobia more often in among those with PTSD than among those without PTSD. Comorbidity is also extremely common in epidemiological samples. In the NVVRS, 98.8 percent of the men who had a lifetime diagnosis of PTSD also qualified for at least one other mental disorder, whereas only 40.6 percent of those without PTSD did so. The most common comorbid disorders were alcohol abuse, depression, and generalized anxiety disorder. Comorbidity is common in epidemiology studies of civilian PTSD as well. In one study conducted in the Detroit area, about 80 percent of subjects with PTSD had at least one other disorder at some point during their lives. In a representative survey of the US population, the lifetime comorbidity rates were 88.3 percent in men and 79 percent in women with PTSD. Major depression and alcohol dependence

(AD) were common comorbid conditions. Retrospectively reported ages of onset suggested that other anxiety disorders often predated PTSD, whereas alcohol and mood disorders usually developed after PTSD.

In cases of comorbidity, clinicians are often challenged by identifying the primary disorder, or, in better terms, the disorder, or the configuration of symptoms that requires most urgent attention. Studies of PTSD suggest that treating "secondary" substance misuse is critical for addressing PTSD in treatment, ideally prior to initiating PTSD treatment. When not possible, simultaneously treating PTSD and substance abuse is feasible; thus, emphasizing harm reduction when abstinence cannot be achieved. Similarly, PTSD patients with prominent depression, suicidal ideation, severe impulse dyscontrol, or paranoid features should first be protected from the immediate consequences of these symptom configurations.

COURSE AND PROGNOSIS

Immediately following exposure to a traumatic event, many people experience shock, nightmares, intrusive thoughts about the trauma, and so forth. Indeed, it would be surprising if severely trauma exposed individuals had no emotional reaction whatsoever to horrific events. Accordingly, a diagnosis of PTSD requires symptoms that persist for at least 1 month following the trauma.

For most people exposed to trauma, symptoms do wane. For example, in a study of help-seeking rape victims, researchers found that 95 percent met PTSD symptom criteria within 2 weeks of the rape. Yet the proportion of victims who still met symptom criteria at 1, 3, and 6 months after the rape declined to 63.3, 45.9, and 41.7 percent, respectively. Among victims of nonsexual assault, 64.7 percent met PTSD symptom criteria 1 week after the trauma, whereas the proportion still meeting criteria at 1, 3, 6, and 9 months after the assault declined to 36.7, 14.6, 11.5, and 0 percent, respectively.

Epidemiological data from the National Comorbidity Study (NCS) indicate that the median time for PTSD to remit is 36 months for individuals who sought help for any mental health problem (not necessarily for PTSD) and about 64 months for individuals who never sought help for a mental health problem. Approximately one-third of those who qualify for PTSD had a chronic course.

A recent review of studies concerning delayed-onset PTSD confirmed that *symptom* onset is almost never delayed. This is a striking contrast to the original claim in the text of DSM-III that symptoms often only surface months or years after the trauma. In some rare cases, a person may develop immediate symptoms but fall just short of meeting full criteria for PTSD. A worsening of symptoms several months later may result in the person finally qualifying for the diagnosis. At first glance, this would appear to constitute a delayed onset of the PTSD *syndrome,* but it is not as if the

person has been wholly asymptomatic during the period following the trauma and prior to meeting full criteria. For example, in one study of Israeli combat veterans, most cases of apparent delayed-onset PTSD involved either exacerbation of chronic, subsyndromic symptoms or delayed help seeking in those with chronic PTSD. For example, some veterans finally met full criteria for PTSD after receiving notice that they were about to be reactivated for military service.

TREATMENT

Psychological Treatment of Chronic PTSD in Adults

A review of 70 studies of psychotherapies involving a total of 4,761 participants showed that many of the studies were rated as being at risk of bias, and sample sizes were generally small with limited follow-up data. There was evidence that individual Trauma-Focused Cognitive Behavior Therapy (TFCBT) including PE which focuses on reexperiencing the traumatic event through repeatedly engaging with the memories (imaginal exposure) and everyday reminders (in vivo exposure) rather than avoiding triggers, Eye Movement Desensitization and Reprocessing Therapy (EMDR), which includes repeatedly recalling distressing images while receiving sensory inputs, and non-TFCBT including Present Centered Therapy (PCT) which focuses on current relationship and work challenges rather than the trauma, are equally effective immediately posttreatment, with some evidence that TFCBT and EMDR are superior to non-TFCBT between 1 and 4 months following treatment. Individual TFCBT, EMDR, and non-TFCBT were found to be more effective than other therapies. There was evidence of greater dropout in active treatment groups. There is also evidence that Cognitive Processing Therapy (CPT) which emphasizes correcting faulty attributions including posttraumatic overgeneralizations of the world as dangerous, uncontrollable and unpredictable is effective for adults with chronic PTSD. A recent study in adults with PTSD reported that IPT, focusing on role expectations in dyadic relationships, was comparable to PE for treatment completers, with fewer dropouts in the IPT group for participants with comorbid depression. This study raises the provocative question “is exposure necessary?” for successfully treating PTSD.

Psychotherapy for Military-Related PTSD

A review of 37 controlled trials found that military-related PTSD is complex and difficult to treat. Trauma-focused treatments and CPT are associated with improvement in symptoms in approximately 60 percent of veterans, but are limited by relatively high dropout rates and treated patients often remain symptomatic with more than two-thirds retaining their diagnosis. Non–trauma-focused treatments, which are less demanding for patients and professionals, provide a reasonable option over trauma-focused

therapy, particularly if this leads to continued willingness to remain in care.

Psychological Therapies for the Treatment of PTSD in Children and Adolescents

A review of 14 controlled studies, including 758 participants exposed to sexual abuse, civil violence, natural disaster, domestic violence, and motor vehicle accidents, found no study was rated at high risk for selection or detection bias but a minority were rated at a high risk for attrition, reporting, and other bias. Across all psychological therapies, improvement was significantly better within a month of completing psychological therapy compared to a control group. The psychological therapy for which there was the best evidence of effectiveness was CBT, with greater improvement lasting up to a year.

Pharmacotherapy for Adults with Chronic PTSD

A meta-analysis was conducted to compare the efficacy of different medications for treating PTSD. The largest body of evidence for short- and long-term efficacy of medication currently exists for SSRIs, the strongest evidence is for sertraline and paroxetine which are approved by the FDA for adults with PTSD. There are promising initial findings for the noradrenergic and serotonergic reuptake inhibitor venlafaxine and the atypical antipsychotic risperidone. Evidence for the effectiveness of benzodiazepines is lacking, despite their continued use in clinical practice. Finally, the alpha 1-adrenergic antagonist prazosin and the atypical antipsychotics show some efficacy in treatment-resistant PTSD, and prazosin shows preliminary efficacy for treating nightmares. Insomnia is prevalent and disabling in chronic PTSD, with low-dose trazadone preferable to benzodiazepines. The recently developed orexin antagonist suvorexant is promising for managing trauma-related insomnia. Medications in the pipeline are targeting endocannabinoids including fatty acid amide hydrolase (FAAH) inhibitors, nabilone and cannabidiol, and other novel targets including glutamate, brain-derived neurotrophic factor (BDNF) and the oxytocin receptor.

Pharmacological Treatment of Comorbid PTSD and Substance Use Disorder

Researchers have reviewed advances in pharmacological treatment of comorbid PTSD and SUD. A randomized clinical trial of sertraline did not show overall efficacy for comorbid PTSD and AD, although it may have efficacy among light drinkers. Another clinical trial demonstrated the efficacy of both disulfiram and naltrexone for the treatment of AD in individuals with PTSD. A recent clinical trial suggested that norepinephrine reuptake inhibitors may also have efficacy for the treatment of comorbid PTSD and AD. Noradrenergic medications that are promising for comorbid PTSD and SUD include prazosin, guanfacine, and atomoxetine. Promising

glutamate/GABA medications include topiramate, memantine, acamprosate, N-acetylcysteine (NAC), and ketamine. The safety and efficacy of these medications for the treatment of PTSD and SUD need to be tested in controlled clinical trials.

Combined Pharmacotherapy and Psychological Treatments for PTSD

Researchers have assessed whether the combination of psychological therapy and pharmacotherapy is more effective for treating PTSD than either of these interventions alone. Patients of any age or gender, with chronic or recent onset PTSD arising from any type of event relevant to the diagnostic criteria were included. Four trials were eligible for inclusion, one of these trials (n = 24) was on children and adolescents. All used an SSRI and PE or a cognitive behavioral intervention. There was no strong evidence for superiority of combined treatment over psychological therapy or pharmacotherapy used as monotherapy. The authors conclude that there is not enough evidence to support or refute the effectiveness of combined pharmacotherapy and psychological compared to either of these interventions alone and that large randomized controlled trials are urgently required.

Nontraditional, Alternative, and Complimentary Treatments

There is preliminary evidence to support the efficacy of acupuncture, breathing and muscle relaxation, mindfulness meditation and yoga. For body-based therapies there is limited evidence, including no published studies for movement-based and energy therapies.

Early and Preventive Intervention

Unlike other mental disorders, PTSD often follows a distinct triggering event and has a clear onset point. Moreover, numerous trauma-exposed individuals are brought to the immediate attention of emergency care services and helpers. This creates a unique opportunity to identify individuals at risk and initiate preventive interventions.

Individuals' reactions to traumatic events follow characteristic trajectories that include persistent lack of symptoms, initial disarray followed by recovery, and nonremitting symptoms. A less frequent symptom trajectory consists of early resilience followed by delayed onset of symptoms. These trajectories define two primary goals for early interventions: (a) mitigate the development of early symptoms or (b) increase the likelihood of remission in survivors with initial symptoms and (c) prevent the delayed emergence of PTSD. Studies addressing the first goal include "stress-management," "preparation and training" or "need-based assistance (food, shelter, soothing human contact)," as well as pharmacological and psychological prevention provided immediately after

trauma exposure. Studies addressing the second goal must identify survivors at risk and use specific intervention protocols.

Psychological or Behavioral Interventions

Debriefing. Psychological debriefing, a widely used method in the 1980s–1990s, aimed at preventing long-term posttraumatic symptoms by promoting emotional processing of traumatic events. Debriefing was offered unselectively to anyone exposed to a potentially traumatic event. The method typically involved a single session within several days of the trauma exposure, either in a group or individual, and included general education about the possible effects of trauma exposure, as well as chronological retellings of the recent traumatic event. The method has face validity and is still well known, and therefore may be expected by laypeople when confronted with traumatic events. However, well-conducted studies showed no evidence of beneficial effects and even suggested that debriefing may have a negative effect on recovery. After a negative Cochrane review first published in 1997, most treatment guidelines have been updated to recommend against providing psychological debriefing on a routine basis for adults after trauma.

Early Cognitive Behavioral Therapy. Cognitive-behavioral therapy (CBT) can involve different strategies with distinct aims. *Exposure-based CBT,* exemplified by the PE protocol, aims to achieve and maintain fear extinction through repeated exposure to trauma-related stimuli in a safe context, thereby providing a sense of control over reactions and reducing avoidance. *Cognitive-based CBT* approaches aim to change the way patients continue to react to trauma-related reminders and to behavioral restrictions and rules derived from the traumatic experiences, and to negative appraisal of self and others by challenging the patient's beliefs about the meaning and current implication of the trauma. CBT is offered individually or in small group, to people who report symptoms (else there are no "intervention targets" and typically involves several weekly sessions, homework, and in vivo training exercises. The treatment may continue for over 3 months and require significant skills from the therapists.

Exposure-based CBT efficiently reduces PTSD symptoms in selected samples. A study was conducted using modified PE in rape, assault and motor vehicle accidents survivors around 12 hours after trauma, and found lower PTSD symptoms in the intervention group at 4 and 12 weeks after trauma, mainly for sexual assault victims. Another study found 5 weeks of exposure-based CBT to be effective in reducing PTSD in participants who meet ASD diagnostic criteria. Research has also found a reduction of PTSD symptoms at 13 months—but not 3 months after the traumatic events. In the largest study to date, researchers showed that both exposure and non–exposure-based CBT efficiently reduced the prevalence and the intensity of

PTSD symptoms 5 months after trauma exposure and that the outcome of delayed CBT (starting 5 months after the event), 9 months after trauma exposure did not differ from that of treatment administered within a month of the traumatic event. However, a small study with 3 weeks of PE did not find significant symptom improvement in the PE group compared to supportive counseling. Cognitive-based CBT has also shown effectiveness in some but not all studies. Researchers compared CBT to waitlist control subjects with acute PTSD, and found that CBT accelerated recovery, but makes no long-term difference.

Modifications of CBT-based intervention showed varied results. A telephone-based CBT study in patients with cardio defibrillator installed reported significant improvements in PTSD symptoms in the CBT group. Researchers developed a self-guided Internet-based intervention (Trauma TIPS) based on CBT to prevent the onset of PTSD symptoms. The result did not support the efficacy of Trauma TIPS.

Cognitive behavioral therapy is currently the mainstay of early prevention of PTSD. Several considerations, however, make the systematic implementation of CBT a major challenge. First, a meta-analysis of early interventions has indicated that CBT is only efficient in participants with diagnosable PTSD at treatment onset. Other findings have similarly demonstrated that survivors with subthreshold PTSD symptoms recover as well with or without CBT. Furthermore research has shown the known efficiency of CBT is limited to survivors of sexual assault, showing no effect among accidents victims. Moreover, attempts to deliver diluted versions of CBT (such as telephone or web-based communication) yielded negative results. Interestingly, studies do not show a reduced effect of CBT when administered later in the development of PTSD (e.g., 6 months after trauma exposure instead of 1 month). CBT, therefore is best positioned by current research as the treatment for ascertained clinical cases, optimally provided at some distance from the traumatic event. On the positive side, the early outcome of CBT is maintained over time and thus this intervention is virtually the unique means to stop the progression of early PTSD into chronicity. In its best performance, however, early CBT leaves substantial number of survivors unaffected and thus should be supplemented by “second-step” interventions or by interventions specifically affecting participants in the nonremission trajectory.

Pharmacological Interventions for Preventing PTSD

Various pharmacological agents have been examined in the prevention of posttraumatic symptoms. A recent review article concluded that in general, there is moderate quality evidence for the efficacy of hydrocortisone, and no evidence for propranolol, escitalopram, temazepam, and gabapentin. This field is rapidly developing as the neurobiological process underlying start to be clarified by more studies.

Hydrocortisone. Hydrocortisone has been shown to be effective especially in patients who have never been treated for psychiatric disorders. The underlying mechanism is not entirely understood. One hypothesis is that hydrocortisone can facilitate extinction learning through both nongenomic and genomic effects. Some also believe that high-dose exogenous hydrocortisone administered shortly after trauma may promote recovery through enhancing synaptic plasticity and connectivity. An animal model showed significantly increased dendritic growth and spine density, with increased levels of BDNF and obtunded postsynaptic density-95 (PSD-95) levels in steroid-treated stressed rats. Increasing cortisol levels also counter adrenergic activation, thereby reducing fear conditioning.

Propranolol. Propranolol is a beta-adrenergic antagonist that crosses the blood–brain barrier and therefore capable of reducing the CNS adrenergic drive associated with defensive threat responses. Experimental studies of propranolol in healthy subjects have shown that its administration prior to exposure to potentially traumatic narratives reduced the recollection of stressful elements of the narrative without affecting the general recall. It was thereby positioned as a prime candidate to affect traumatic recall in PTSD. Early treatment with Propranolol aims to overconsolidation of traumatic memories by blocking the memory-enhancing influence of stress hormones. Theoretically, therefore, it should be initiated while memories of the trauma are encoded and consolidated, preferably within hours of the traumatic event. However, controlled studies in the first hours after exposure have failed to show a preventive effect of propranolol on PTSD symptoms, although they showed a reduction in physiological responses to reminders. The documented gap between propranolol's effect of physiological responses to trauma reminders and its lack of effect of PTSD symptoms could be interpreted as suggesting that the pathogenesis of PTSD is not limited to amygdala-mediated threat conditioning and involves other brain areas and memory systems.

Benzodiazepines. Benzodiazepines are a group of GABA agonist, and thereby enhance the inhibitory transmission in wide areas of the brain. They are used as tranquilizers and sleep inducers but also interfere with long-term potentiation, learning, and memory. As such they were positioned as capable of reducing excessive "learning" occurring during or following trauma exposure. However, three human studies have shown that participants receiving benzodiazepines fared worse than those not treated at all. An animal "predator stress" study found that administering diazepam shortly after predator odor exposure enhanced the acquisition of long-term fear responses. While the exact mechanism of benzodiazepines PTSD enhancing effect is unknown, it is possible that these compounds interfere with postevent learning of extinction. Intriguingly, current data

on benzodiazepines rely on small studies while these compounds are widely used to mitigate distress following various stressors. Nonetheless, benzodiazepines are not recommended for use in the aftermath of traumatic events.

Morphine. Animal studies suggest morphine can produce retrograde amnesia for contextual conditioned fear, possibly through decreasing cAMP or activating NMDA receptors in hippocampus. Observational studies of hospital patients also suggested a possible beneficial effect of morphine administration within 48 hours after trauma exposure to survivors who have pain and similar result was reported, retrospectively, in 696 military personnel with severe combat injury. Given the retrospective nature of most studies, more research is needed to separate a specific effect of morphine from a generic "analgesic" effect. Pain after trauma exposure is a potent predictor of PTSD. It is unclear, therefore, whether morphine has any preventive value in trauma survivors without physical pain.

Other Approaches

Oxytocin is involved in emotion stress regulation, social engagement, and attachment and preliminary studies suggest that it may buffer the development of PTSD when administered shortly after trauma exposure. Neuropeptide Y (NPY) is another neuroendocrine candidate for early intervention. An animal study showed that delivery of NPY to rats' brain has a pronounced effect on reducing the development of PTSD-like symptoms, possibly through modifying stress-triggered dysregulation of HPA axis and central noradrenergic activity. Besides hormonal intervention, neurobehavioral retraining is also being tested for its capacity to reduce negative emotional processing and enhance executive control. Emerging evidence of impaired emotion regulation and executive functions in PTSD is likely to generate other early neurocognitive approaches.

Placing Prevention in Context

Most studies above summarize the "clinical" implementation of preventive interventions, that is, specialized care provided to survivors screened or formally diagnosed within a healthcare setting. They are therefore best construed as secondary and tertiary prevention. Studies to date, however, have not followed a primary prevention model. Challenging the restricted "clinical" model Zatzick et al. evaluated a stepped collaborative care model, which introduced care managers to address each patient's unique needs and coordinated different components of interventions (e.g., CBT, Motivation Interview, pharmacotherapy). The care team repeatedly measured patient's symptoms and adjusted levels of care accordingly. Results showed the feasibility and effectiveness of the method with reduced

PTSD symptoms in the intervention group.

IMPLICATIONS AND FUTURE DIRECTIONS

Studies on PTSD prevention demonstrated a large variety in study methodology. Differences in methodology need to be carefully examined before one can infer from study results and determine the most effective intervention for a certain population. Current evidence from research also represents a great extent of heterogeneity in sampling. Different studies recruited people with different features (e.g., trauma types, injury severity, and peritrauma reactions). Among all traditional prevention methods, TFCBT has been mostly reported as an effective approach, but only in people with acute stress symptoms. Taking into consideration the heterogeneity of trauma survivors, people with different age, gender, trauma type, and potentially different genetic features, childhood experience, and recovery environment may need different optimal strategies. For example, TFCBT was reported to be more effective in victims of traffic accidents, and exposure therapy was shown to be more beneficial to sexual assault victims or people with high genetic risk. Since no intervention is effective for all trauma survivors after trauma to improve outcome.

Most importantly, PTSD is likely multicausal and as such individuals with differing vulnerabilities and different exposure and postexposure circumstances may come to express the PTSD symptom complex through individual-specific pathways—and be responsive to individual-specific interventions. One way to advance the prevention of PTSD is to better map the variety of paths leading to this condition, and map those paths into subsets of trauma-exposed individuals. Once such knowledge becomes available, target-specific early intervention might replace generic treatment protocols applied to aggregates. Several steps should be taken in future studies. Psychiatrists need to step beyond diagnosis-based screenings and develop more complex accurate methods to predict individual level risks of chronic symptoms and impairment after traumatic events. By enhancing prediction models, intervention studies can take the important step to select the most relevant sample, for a more rigorous study design and the best clinical interest. At the same time, research need to continue exploring and confirming the underlying mechanisms of posttraumatic pathogenesis. Increasing knowledge to understand the mechanisms of posttraumatic psychopathology will open opportunities to explore more targeted intervention outside the range of traditional treatment-derived methods. These targeted intervention methods can allow clinicians to focus on both specific population and specific pathological process.

REFERENCES

*Amos T, Stein DJ, Ipser JC. Pharmacological interventions for preventing post-traumatic stress

disorder (PTSD). *The Cochrane Database of Sys Rev*. 2014;(7):CD006239.

*Bisson JI, Roberts NP, Andrew M, Cooper R, Lewis C. Psychological therapies for chronic post-traumatic stress disorder (PTSD) in adults. *Cochrane Database Sys Rev*. 2013;12:CD003388.

Brewin CR, Andrews B, Valentine JD. Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. *J Consult Clin Psychol*. 2000;68(5):748.

*Bryant RA. Early Intervention After Trauma. *Evidence Based Treatments for Trauma-Related Psychological Disorders*. Springer; 2015:125–142.

Dohrenwend BP, Turner JB, Turse NA, Adams BG, Koenen KC, Marshal R. The psychological risks of Vietnam for U.S. veterans: a revisit with new data and methods. *Science*. 2006;313(5789):979–982.

Forneris CA, Gartlehner G, Brownley KA, et al. Interventions to prevent post-traumatic stress disorder: a systematic review. *Am J Prev Med*. 2013;44(6):635–650.

Gabbay V, Oatis MD, Silva RR, Hirsch G. Epidemiological aspects of PTSD in children and adolescents. In: Silva RP, Ed. *Posttraumatic Stress Disorder in Children and Adolescents*. New York: Norton; 2004.

Galatzer-Levy IR, Ankri Y, Freedman S, et al. Early PTSD symptom trajectories: persistence, recovery, and response to treatment: results from the Jerusalem Trauma Outreach and Prevention Study (J-TOPS). *PLoS One*. 2013;8(8):e70084.

Gilbertson MW, Shenton ME, Ciszewski A, et al. Smaller hippocampal volume predicts pathologic vulnerability to psychological trauma. *Nat Neurosci*. 2002;5:1242–1247.

Hetrick SE, Purcell R, Garner B, Parslow R. Combined pharmacotherapy and psychological therapies for post traumatic stress disorder (PTSD). *Cochrane Database Syst Rev*. 2010;7:CD007316.

Holbrook TL, Galarneau MR, Dye JL, Quinn K, Dougherty AL. Morphine use after combat injury in Iraq and post-traumatic stress disorder. *N Engl J Med*. 2010;362(2):110–117.

*Institute of Medicine. *Treatment of Posttraumatic Stress Disorder: An Assessment of the Evidence*. Washington, DC: National Academies Press; 2007.

Ipser JC, Stein DJ. Evidence-based pharmacotherapy of post-traumatic stress disorder (PTSD). *Int J Neuropsychopharmacol*. 2012;15(6):825–840.

Kearns MC, Ressler KJ, Zatzick D, Rothbaum BO. Early interventions for PTSD: a review. *Depress Anxiety*. 2012;29(10):833–842.

*Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):617–627.

Kliem S, Kröger C. Prevention of chronic PTSD with early cognitive behavioral therapy. A meta-analysis using mixed-effects modeling. *Behav Res Ther*. 2013;51(11):753–761.

Kulka RA, Schlenger WA, Fairbanks JA, et al. *Trauma and the Vietnam War Generation: Report of Findings from the National Vietnam Veterans Readjustment Study*. New York: Brunner/Mazel; 1990.

Marmar CR, Schlenger W, Henn-Haase C, et al. Course of posttraumatic stress disorder 40 years after the Vietnam war: findings from the National Vietnam Veterans Longitudinal Study. *JAMA Psychiatry*. 2015;72(9):875–881.

Ozer EJ, Best SR, Lipsey TL, Weiss DS. Predictors of posttraumatic stress disorder and

symptoms in adults: a meta-analysis. *Psychol Bull*. 2003;129(1):52–73.

Pitman RK, Rasmusson AM, Koenen KC, et al. Biological studies of post-traumatic stress disorder. *Nature Rev Neurosci*. 2012;13(11):769–787.

Roberts NP, Kitchiner NJ, Kenardy J, Bisson JI. Early psychological interventions to treat acute traumatic stress symptoms. *Cochrane Database Syst Rev*. 2010;(3):CD007944.

Rose S, Bisson J, Churchill R, Wessely S. Psychological debriefing for preventing post traumatic stress disorder (PTSD). *Cochrane Database Syst Rev*. 2002;(2):CD000560.

Shalev AY, Ankri Y, Israeli-Shalev Y, Peleg T, Adessky R, Freedman S. Prevention of posttraumatic stress disorder by early treatment: results from the Jerusalem Trauma Outreach And Prevention study. *Arch Gen Psychiatry*. 2012;69(2):166–176.

Steenkamp MM, Litz BT, Hoge CW, Marmar CR. Psychotherapy for military-related PTSD: a review of randomized clinical trials. *JAMA*. 2015;314(5):489–500.

Tanielian T, Jaycox L, eds. *Invisible Wounds of War: Psychological and Cognitive Injuries, Their Consequences, and Services to Assist Recovery*. Santa Monica, CA: RAND; 2008.

Van Emmerik AA, Kamphuis JH, Hulsbosch AM, Emmelkamp PM. Single session debriefing after psychological trauma: a meta-analysis. *Lancet*. 2002;360(9335):766–771.

Xue C, Ge Y, Tang B, et al. A meta-analysis of risk factors for combat-related PTSD among military personnel and veterans. *PLoS One*. 2015;10(3):e0120270.

Yehuda R. Post-traumatic stress disorder. *N Engl J Med*. 2002;346(2):108–114.

Yehuda R, McFarlane AC. Conflict between current knowledge about posttraumatic stress disorder and its original conceptual basis. *Am J Psychiatry*. 1995;152(12):1705–1713.

Zatzick D, Roy-Byrne P, Russo J, et al. A randomized effectiveness trial of stepped collaborative care for acutely injured trauma survivors. *Arch Gen Psychiatry*. 2004;61(5):498–506.

18

Somatic Symptom and Related Disorders

JAVIER I. ESCOBAR, M.D., AND JOEL E. DIMSDALE, M.D.

INTRODUCTION

There appears to be a universal tendency to experience and communicate psychological distress in the form of physical symptoms and to seek medical attention for these symptoms. Often, these physical symptoms remain poorly explained and are associated with increased medical visits, unnecessary medical tests, and the performance of procedures that may result in iatrogenic complications. For many decades, the term "somatization symptoms" was used to describe these phenomena. This was gradually replaced by more descriptive terms such as "medically unexplained symptoms," "unexplained symptoms," and "functional somatic symptoms." The word "idiopathic" (meaning "arising spontaneously or from an obscure or unknown cause") is probably the most technically correct way to label the physical symptoms displayed by many patients with psychiatric disorders that remain poorly explained after detailed medical scrutiny and laboratory/radiological assessments. It also fits current thinking in other fields such as neurology where distinctions between "neuropathic," "inflammatory," "nociceptive," and "idiopathic" pain are common and useful. The evidence for the "psychogenesis" of many of these somatic presentations was originally derived from theoretical postulates such as Wilhelm Stekel's notion of somatization and has been nurtured by the frequent association of these symptoms with psychological distress and psychopathology. However, many of these somatic manifestations can arise spontaneously and their causes can remain obscure or unknown. The most distinctive characteristic of many patients with somatic symptom disorders is not the symptoms they complain about but the way in which they present and interpret them. Despite the relevance that these idiopathic physical symptoms have to medicine and psychiatry, their definition is difficult and their classification remains controversial. This has led to frequent revisions of nomenclatures that constantly alter the criteria used to designate "cases" and complicate the clinical recognition and management of these syndromes, recently and notably in DSM-5.

Everyone experiences somatic symptoms with certain frequency and

most are able to cope with them effectively. However, some people's lives are overwhelmed by their somatic concerns. Sometimes, the somatic concerns stem from well-established major medical illnesses; sometimes the origins of the concerns are never quite clear. What is common in both situations is that patients' lives are shadowed by disproportionate and excessive thoughts, feelings, and behaviors that center on the perceived somatic sensations. DSM-5 incorporated this perspective, because the older perspective (simply counting or cataloguing symptoms labeled as medically unexplained) embodied in the previous somatoform disorders' category in DSM-III and IV, was unreliable and such perspective often put the doctor and the patient at odds over the question of the legitimacy and "reality" of the patient's symptoms and personal suffering. In fact, research has shown that regardless of medical explanations, multiple physical symptoms predict disability and psychopathology (Fig. 18–1).

DEFINITION

According to the fifth edition of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), the key characteristic of somatic symptom disorders is the presence of one or more persistent somatic symptoms that are associated with excessive thoughts, feelings, and behaviors related to the symptoms. The stipulation that these symptoms be medically unexplained has been removed, because of the recognition that such a distinction in itself is unreliable and because psychiatrists commonly treat patients with medically established diagnoses who are troubled by or preoccupied with their physical symptoms, disproportionately. Furthermore, DSM-5 acknowledges that frequently, somatic symptom disorders co-occur with other psychiatric disorders. The diagnosis of a somatic symptom disorder is only made when there is significant distress or impairment associated with these chronic somatic symptoms.

In DSM-IV and earlier editions of the DSM, disorders such as factitious disorder and psychological factors affecting other medical conditions were grouped elsewhere. This led to some confusion in differential diagnosis. As a result, such disorders are now included within the rubric of the somatic symptom disorders chapter. Due to its strong linkage and similarity to obsessive-compulsive disorders, body dysmorphic disorder has now been moved out of the somatic symptom disorders and is now included in the obsessive-compulsive disorders group. Because pain symptoms are quite common and the requirement that symptoms be unexplained has been removed, pain disorder as a psychiatric diagnosis has been deleted in DSM-5, and these syndromes are now subsumed under somatic symptom disorder. Also, largely because of its pejorative connotation, the classical term "hypochondriasis" has been removed and the disorder renamed as Illness Anxiety Disorder in DSM-5 (Fig. 18–2).

Thus, the somatic symptom disorders' group in DSM-5 includes the following disorders: "somatic symptom disorder," "illness anxiety disorder," "conversion disorder" (functional neurological symptom disorder), "psychological factors affecting another medical condition" and "factitious disorder." Somatic syndromes that are below-threshold are classified under two residual categories, "other specified" and "unspecified somatic symptom and related disorders."

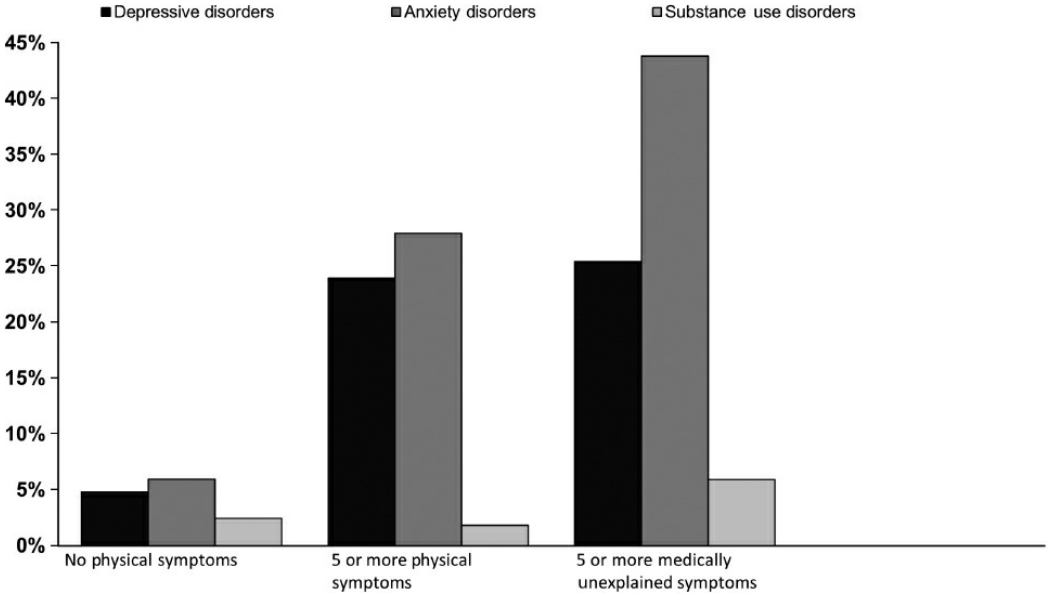

FIGURE 18–1. Whether medically explained or not, high levels of somatic symptoms augur psychopathology in general populations. (Adapted from Escobar JI, Cook B, Chen CN, et al. Whether medically explained or not, three or more concurrent somatic symptoms predict psychopathology and service use in community populations. *J Psychosom Res.* 2010;69:1–8.)

In the DSM-5 field trials, Somatic Symptom Disorder showed high levels of reliability, more robust even than those for more traditional diagnoses (Fig. 18–3).

HISTORY

Somatic presentations that have a presumed psychological origin have always mystified the medical establishment and have taken new forms and shapes as cultures evolve and medical paradigms shift. Throughout the ages, multiple folk concepts and terms have been used to designate the phenomena, such as "hysteria," "hypochondria," "spleen," "English malady," "surmenage," "humoral disorders," "neurovegetative dysfunction," "neurasthenia," "neurovegetative dystonias," "psychosomatic disorders," and many others. The idea that psychological processes are involved in these somatic experiences appears to have ancient origins but waxed and waned periodically. In the 17th century, Thomas Sydenham attributed symptoms and signs such as "pain," "convulsions," "diarrhea,"

and “dropsy” to disturbances of the mind.

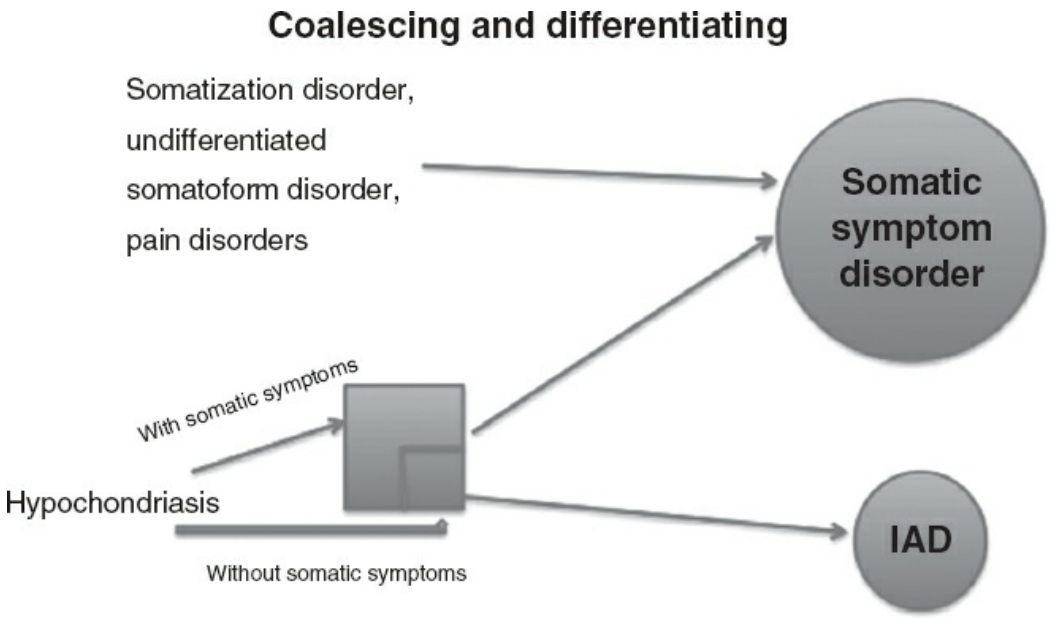

FIGURE 18–2. Main changes from DSM-IV to DSM-5. (From Dimsdale JE, Creed F, Escobar JI, et al. Somatic symptom disorder: an important change in DSM. *J Psychosom Res.* 2013;75(3):223–228.)

Hysteria and Hypochondria

The more dramatic somatic manifestations—“hysteria” and “hypochondria”—were key syndromes in traditional psychopathology and remained influential in medical discourse at the turn of the 20th century. The long survival of these labels (e.g., “hysteria,” “hypochondria” in the United States and “neurasthenia” in China) despite efforts to eliminate them, attests to their historical and practical value, but also to the fragility and inadequacy of current classifications. Michel Foucault, in his classical book, included “hysteria” and “hypochondria” among the four classic syndromes in psychopathology together with “mania” and “melancholia.” He indicated that from ancient times, somatic presentations have not fitted well the schemes of traditional psychopathology and have been viewed as outcasts in medicine and psychopathology.

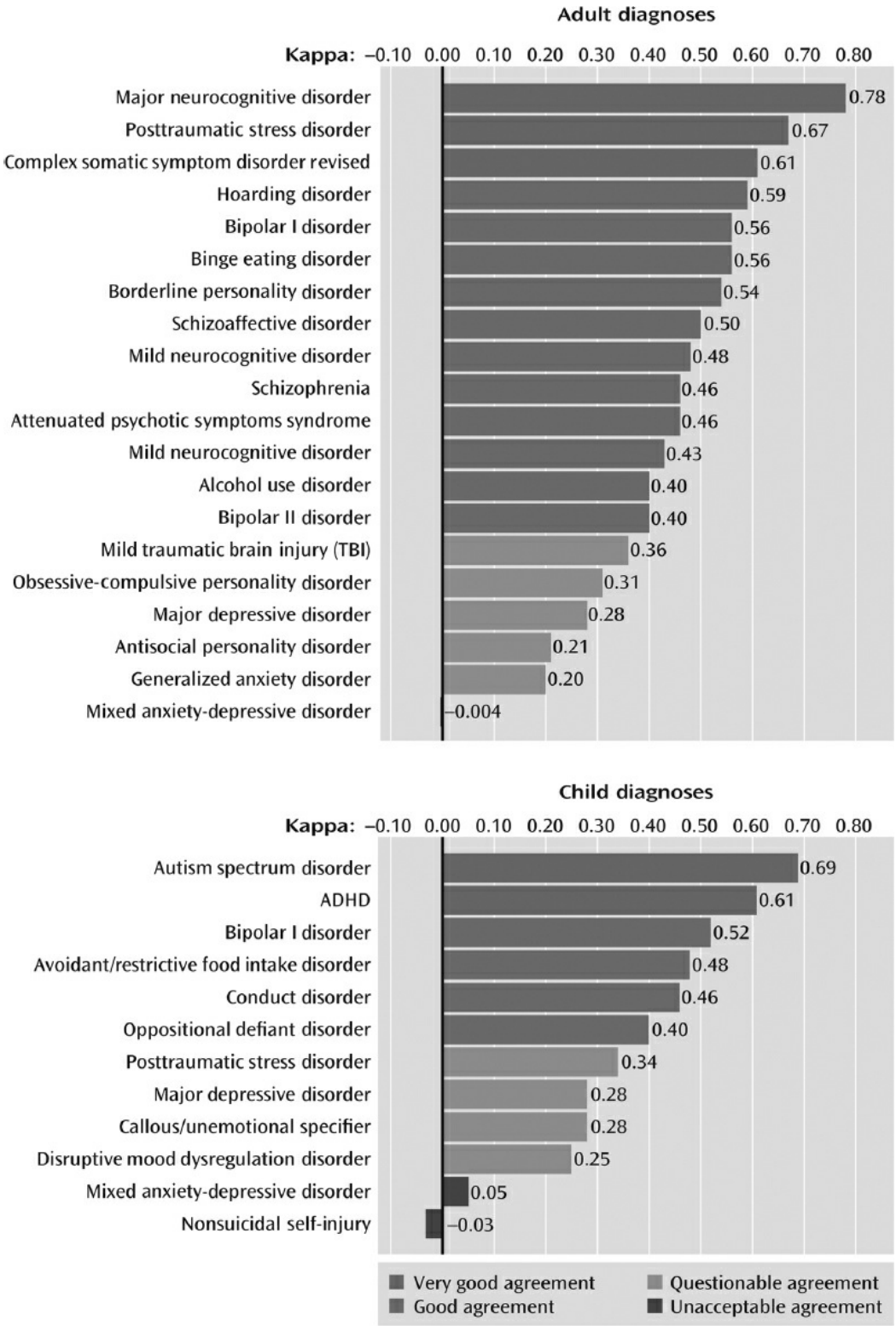

FIGURE 18–3. Reliability of somatic symptom disorders relative to other DSM-5 disorders. (From Freedman R, Lewis DA, Michels R, et al. The initial field trials of DSM-5: new blooms and old thorns. *Am J Psychiatry.* 2013;170(1):1–5 (Fig. 18–1).)

Somatic Presentations in Classical Psychopathology

In his "History of Mental Symptoms" the psychiatric historian German Berrios, a Peruvian-born psychiatrist practicing in England, maintained

that whereas initially physical and psychological phenomena were dealt with in a unitary fashion, after the 19th century, somatic symptoms were discarded as "nonspecific." Thus, according to Berrios, the leading psychopathological entities that came as a legacy from the 19th century excluded hysteria and hypochondria and included only "melancholia," "mania," "phrenitis," "delirium," "paranoia," "lethargy," "carus" (a term similar to "idiocy," "foolishness," stupidity," terms unfortunately utilized in the past to denote intellectual or cognitive deficits) and "dementia." Berrios also mentioned a disorder called "dysmorphophobia" and referred to the French term *cenestopathie* as precursor of "neurovegetative dystonias" and "psychosomatic syndromes."

Kurt Schneider viewed somatic presentations not as a separate group but as key components of other psychopathological syndromes. For example, he considered "anguish" as a corporal sensation localized in the chest. However, in Schneider's psychopathology, somatoform phenomena appear to have the best fit among his "psychopathic personalities," specifically the group he labeled "asthenic psychopaths." Schneider characterized them as "somatically labile" or "somatopaths"—individuals who completely focused their attention on their bodies, fatigued easily, and suffered from insomnia, headaches, and heart, bladder, and mental disturbances. Schneider also trivialized somatic presentations by asserting, for example, "occupation often dissipates all these problems."

Karl Jaspers mentioned "hypochondria," "hysteria," and the "somatopsychic," emphasizing the polarity that existed at that time regarding physical and psychological approaches to these problems. He affirmed that a healthy person lives in his body and does not think much of it and added that the great majority of physical suffering is due to psychological reflection and not to manifested physical disease. Thus, in Jaspers' opinion, many of the somatic presentations tend to become fixed and repetitive and can be understood as part of personal experience, situations, and conflicts.

The Swiss psychopathologist Eugen Bleuler dedicated very little attention to somatic phenomena, considering them to be outside the "endogenous" realm to which he devoted the most attention in his psychopathology. While Bleuler alluded to "hysterical syndromes" and "neurasthenic syndrome," he did not mention hypochondria or psychosomatic phenomena.

Henry Ey, an influential French psychiatrist and author of a well-known textbook of psychiatry in France, reflected the influence of psychoanalytic theories on European psychiatry during the first half of the 20th century. In his textbook, he devoted a full chapter to hysteria, which he defined as "the somatic hyperexpressivity of unconscious ideas, images, and affects." Ey did not mention hypochondriasis but included a long list of psychosomatic disorders that he labeled "organ neuroses" similar to the approaches of

Otto Fenichel and other psychoanalysts who are well known in North America.

The Spanish psychiatrist Juan José Lopez-Ibor, strongly influenced by German psychiatry, was perhaps the first to highlight the somatic ("endogenous" or "biological") component of some forms of anxiety and depression, proposing antidepressant treatments for them as early as the 1960s and thus anticipating by a number of years the development of the panic disorder concept in North American psychiatry. As to phenomena relevant to somatic presentations, Lopez Ibor defined anxiety as "vital anxiety or anguish," in his view, a corporeal sensation, coining the term *timopatia ansiosa* (anxious tymopathy) (not referring here to the thymus gland but to the Greek root *thymos,* meaning soul, spirit, mind, temperament). Lopez Ibor also placed "hysteria," "hypochondria," and "psychosomatic disorders" among his "disorders of mood."

Briquet Syndrome

The description in 1859 of a syndrome with multiple motor and sensory symptoms by the French neurologist Pierre Briquet allowed an initial separation of somatization from conversion phenomena. This approach initiated a trend to emphasize the descriptive aspects of these syndromes, gradually detaching them from psychological causality and placing a particular focus on the process of eliciting individual physical symptoms. This would eventually lead to the "symptom counting process" that would characterize somatization disorder under the Feighner, RDC, and DSM (DSM-III and IV) criteria.

Somatization

The psychoanalyst Wilhelm Stekel used this term for the first time in 1911, defining it as the bodily expression of a deep somatic neurosis or the process of how neurotic conflicts are presented as physical symptoms. Although the term "somatization" is credited to Stekel, its formal entry into medical discourse is due to the work and influence of Zbigniew J. Lipowski, a Polish-born psychiatrist practicing in Canada. In the 1960s, Lipowski provided a broader view of somatization, defining it as the tendency to experience, conceptualize, and/or communicate psychological states or contents as body sensations, functional changes, or somatic metaphors. By the time "somatization disorder" was enthroned as a full category in official nosology, an extensive body of research, principally conducted by investigators at Washington University/Barnes Hospital in St. Louis, had produced empirical evidence for the utility of its forbearer "Hysteria/Briquet syndrome" in tertiary care settings in North America.

The St. Louis Criteria

In 1951, James J. Purtell, Eli Robins, and Mandel Cohen, investigators from Washington University in St. Louis, compared the incidence of specific physical and mental symptoms among women with and without a diagnosis of hysteria and highlighted the polysymptomatic and chronic nature of this disorder. Next, Michael J. Perley and Samuel B. Guze in 1962 formulated for the first time a numerical definition of the symptoms needed to diagnose hysteria, which included the presence of at least 15 individual symptoms distributed over at least nine of ten groups of possible symptoms. These criteria were a modification of those proposed by E. Robins and P. O'Neal in a 1953 study of children that required at least 18 symptoms from at least seven symptom groups. The St. Louis group began adding other physical symptoms, psychiatric symptoms, and personality characteristics on top of the "pure" neurological symptoms that were included in the original Briquet syndrome. This grouping together of somatic and psychiatric symptoms was again used, with a few modifications, in the criteria proposed by John P. Feighner and collaborators for use in psychiatric research. In the Feighner criteria, the threshold for a hysteria diagnosis was elevated to at least 25 symptoms distributed in nine of ten possible groups. These criteria also required that the disorder should start prior to the age of 30 years.

COMPARATIVE NOSOLOGY

The Concept of "Somatoform"

Psychiatric nomenclatures (DSM-III and IV), relegated somatic phenomena to a secondary level and classified somatic syndromes of presumed psychological origin under the heading of "somatoform disorders," a category that is hierarchically inferior to other categories included in these nosologies (e.g., mood disorders). In the 1970s, the American nosologist Robert Spitzer coined the term "somatoform" (meaning somatic-like) to group these disorders among his research diagnostic criteria (RDC), which made its formal entry in North American psychiatry with the publication of the DSM-III in the 1970s. Although the creation of the somatoform disorders category was lauded as a nosological advance by some investigators, in the real world, patients presenting with multiple unexplained physical symptoms did not fit neatly into the taxonomies provided by the DSM and ICD systems. Even when somatoform criteria are met, there is also significant overlap with several other psychiatric syndromes, a reminder that nosologies are still quite imperfect and in need of continuous distillation. The restrictive quality of categorical somatoform diagnoses left out many of the "real cases" and this led investigators to propose dimensional classifications that included broader, less restrictive categories as an alternative for research, such as "abridged somatization," "multisomatoform disorder," and others. The use

of dimensional constructs rather than categories may have particular utility in primary care. In reviewing the literature on physical symptoms, it appears that there are two main dimensions that may have practical value. One is characterized by high levels of physical symptoms (the former somatization) and the other, by amplification of somatic presentations (hypochondriasis or health anxiety). This dimensional approach has been taken into account in the current nosological revision (DSM-5) and is included in Section III of the manual (Table 18–1).

Table 18–1.
DSM-5 Diagnostic Criteria for Somatic Symptom Disorder

300.82 (F45.1) Somatic Symptom Disorder

A. One or more somatic symptoms that are distressing or result in significant disruption of daily life
B. Excessive thoughts, feelings, or behaviors related to the somatic symptoms or associated health concerns as manifested by at least one of the following:
   1. Disproportionate and persistent thoughts about the seriousness of one's symptoms
   2. Persistently high level of anxiety about health or symptoms
   3. Excessive time and energy devoted to these symptoms or health concerns
C. Although anyone somatic symptom may not be continuously present, the state of being symptomatic is persistent (typically more than 6 months).

Specify if: With predominant pain
Specify if: Persistent
Specify current severity: Mild, moderate, severe

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Differences between DSM and ICD Classification of Somatic Symptom Disorders

As shown in Table 18–2, there were several differences between DSM-IV and ICD-10 in the way they classified these disorders. DSM-IV listed all these disorders under the Somatoform Disorders category. Chapter F45 of the ICD-10 included "Somatization Disorder" (this had a different symptom threshold from that in DSM-5), "Undifferentiated Somatoform Disorder," "Hypochondriacal Disorder," Other Somatomorph Disorders, and Somatomorph Disorder Not Otherwise Specified (note the use of "somatomorph" instead of "somatoform") and two disorders not listed in DSM-IV, "Somatomorph Autonomic Dysfunction" and "Persistent Somatomorph Pain Disorder." In ICD-10, conversion and dissociative disorders were placed in a separate category under the heading of "dissociative disorders [conversion]" while two disorders not in DSM-IV, "neurasthenia" and "Briquet's disorder" were listed in Chapter F48 under "other neurotic disorders."

Table 18–2.
Somatic Symptom Disorders in DSM-5 and ICD-10

DSM-5	ICD-10
Somatic symptom disorder	Somatization disorder, undifferentiated somatoform disorder, persistent somatoform pain disorder
Conversion disorder (functional neurological symptom disorder)	Dissociative (conversion)
Illness anxiety disorder	Hypochondriacal disorder
Factitious disorder	
Somatoform disorder NEC	Somatoform autonomic dysfunction, other somatoform disorders, somatoform disorder unspecified, neurasthenia (neurotic disorders category)
Psychological factors affecting another medical condition	

Data from DSM-5, American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association.

World Health Organization. *International Statistical Classification of Diseases and Related Health Problems*. 10th revision. Geneva: World Health Organization; 1992.

In terms of equivalences between DSM-5 and ICD-11, the expectation is that these two systems will be better synchronized than was the case with previous nomenclatures, but this cannot be confirmed until ICD-11 is finally released.

Functional Somatic Syndromes

Table 18–3 lists a conglomerate of syndromes frequently diagnosed in medical practice. These consist of clusters of "idiopathic" physical symptoms. Despite the fact that these functional diagnoses are "legitimately" used in many specialties, their association with psychopathology and psychological distress suggests that many of them would fall within the realm of the somatic symptom disorders. Excellent reviews of these syndromes have appeared in leading journals on both sides of the Atlantic (e.g., *Lancet, JAMA*), highlighting their common ancestry, as well as their epidemiologic, clinical, and psychopathological aspects. Besides the high levels of idiopathic symptoms, these syndromes also share such common elements as the following:

- The absence of a gold standard against which a specific diagnosis can be confirmed or ruled out.
- No consistent explanation emanating from physical and laboratory assessments.
- The absence of clearly articulated pathophysiology, despite a panoply of putative mechanisms (e.g., symptom amplification, muscle contraction, catecholamine release, persistent neurobiological dysfunction, neurological hyper reactivity, elevated cortisol, etc.).
- The simultaneous presence of multiple unexplained physical symptoms

originating from several different organ systems.

- High levels of psychiatric symptoms and "comorbidities."
- Comparable responses to certain psychological (e.g., cognitive-behavioral therapy [CBT]) and pharmacologic (e.g., antidepressants) interventions.

Table 18–3.
Functional Somatic Syndromes across Medical Specialties

Specialty Area	Functional Syndrome(s)
Allergy	Food allergies
Cardiology	Atypical chest pain, noncardiac pain, mitral valve prolapse
Dentistry	Temporomandibular joint syndrome, atypical facial pain
Ear, nose, and throat	Tinnitus, dizziness, globus syndrome
Gastroenterology	Irritable bowel, nonulcer dyspepsia
Internal medicine	Chronic fatigue syndrome, chronic Lyme disease, hypoglycemia, chronic candidiasis
Military medicine	Gulf war syndrome, chronic multisymptom illness
Neurology	Tension headache, pseudoseizures
Obstetrics/gynecology	Premenstrual syndrome, chronic pelvic pain
Occupational medicine	Multiple chemical sensitivity, sick-building syndrome, repetitive strain injury
Orthopedics	Carpal tunnel syndrome, chronic low back pain, chronic whiplash syndrome
Plastic surgery	Silicone-associated connective tissue disease
Pulmonary medicine	Dyspnea, habit cough, laryngeal dysfunction, hyperventilation
Rehabilitation medicine	Fibrositis, fibromyalgia
Urology	Interstitial cystitis

Modified from Henningsen P, Zipfel S, Herzog W. Management of functional somatic syndromes. *Lancet.* 2007;369:946.

EPIDEMIOLOGY

When available, specific epidemiological data will be provided under each individual disorder.

In general, somatic presentations are very common in all countries and cultures. A WHO study done in nine countries of the world, showed that only one-third of patients presenting to primary care had a documentable physical illness. At least one additional third had a mental disorder. Physical symptoms can be considered as the most common presentation for mental disorders worldwide, and their frequency seems to be on the increase. Canadian historian Edward Shorter pointed out that the shift in somatic symptoms commonly presented in medical and psychiatric settings seems to be related to prevailing medical models. He noted that according to US household surveys, there was a 150 percent increase in "episodes of illness" reported by the US population between the 1930s and the 1980s, despite the enormous improvements in health care. According to the data,

more than one-half of all patients presenting to primary care may present with "idiopathic physical symptoms." Yet, despite their predominantly somatic presentations, such patients do not appear to have a diagnosable physical illness. In fact, the proportion of patients in general medical settings with "idiopathic" physical symptoms for which no organic cause can be found has been estimated to be in the 20 to 80 percent range. It is quite common for these "idiopathic" physical symptoms to co-occur with psychiatric disorders—specifically with anxiety and mood disorders. With regard to recent data from countries outside the United States, more than 20 percent of the general population in Germany reported at least one "idiopathic" physical symptom leading to severe impairment. In Switzerland, about 80 percent of young adults reported somatic problems such as sleep disorders, backache, headache, and stomach or bowel complaints. In addition, research in primary care has shown that regardless of the presence or absence of medical explanations, as the number of individual somatic symptoms increases, so do the odds that the person will also meet criteria for common psychiatric disorders in particular depression and anxiety disorders.

ETIOLOGY

Somatic presentations of mental disorders should be conceptualized using a biopsychosocial model that includes genetic and biological vulnerability, environmental influences, as well as psychological and behavioral elements. According to this model, a number of contributing factors converge into the final common pathway of "somatic presentations." This appears to be a complex process in which genetic and biological vulnerability (increased sensitivity to pain, proprioceptive acuity), early traumatic experiences (violence, abuse, deprivations), learning factors (attention obtained from having an illness, lack of reinforcement of nonsomatic expressions of distress) and psychological elements, lead to a characterological style and a behavioral repertoire that focus on the somatic and conveys a diminished ability to express emotionally laden feelings. These traits may be more likely displayed in certain cultural/social frames that devalue or stigmatize psychological suffering in contrast to "genuine" medical disorders. This is reinforced or stimulated by psychiatric comorbidities, and in some instances, secondary gain (such as financial compensation).

Environmental and Cultural Factors

"Idiopathic" physical symptoms have been reported in most cultural groups, although the types of symptoms and their prevalence appear to differ across cultures. The phenomenon of "somatization" has also been well documented in international studies. Differences in medical care

across cultures obviously affect the presentation, recognition, and management of these somatic presentations. In general, "idiopathic" physical symptoms appear to be more common among individuals from lower socioeconomic strata, patients from developing countries, and, in the United States, those from certain ethnic groups. Often, clinical presentations dominated by idiopathic physical symptoms may represent "idioms of distress." Lay explanations embedded in culture concerning the origin of behavioral symptoms (e.g., *ataque de nervios* among Puerto Ricans) may play a role. Variations in symptom presentation are likely the result of the interaction of multiple factors within cultural contexts that affect how individuals identify and classify bodily sensations, perceive illness, and seek medical attention for them. Thus, somatic presentations can be viewed as expressions of personal suffering inserted in a cultural and social context. L. J. Kirmayer outlined a useful process for framing somatic symptoms that incorporates disease/disorder, symbolic expression of intrapsychic conflict, psychopathology, metaphorical expression of distress, positioning in the world, and formulation of commentary or social protest. According to Kirmayer, "somatization" and "psychologization" are not mutually exclusive ways in which to express distress. He advocated the adoption by clinical investigators of a sociocultural perspective to unravel the meaning of these somatic symptoms and to design effective therapeutic strategies.

Biological Factors

There is little information on the biology of somatoform disorders. Neurophysiological studies had suggested that somatoform and functional somatic syndromes might be related to proprioceptive acuity, abnormality of autonomic and proprioceptive responses, problems with the pituitary–hypothalamic axis, and other nonspecific findings. More recent studies using brain imaging techniques point to interesting differences in patients with conversion disorder, and functional somatic syndromes such as irritable bowel syndrome compared to controls or patients with demonstrated organic pathology. A study by Jon Stone and coworkers, using fMRI, showed a different type of brain activation in patients with conversion disorder, which suggested a more complex level of mental activity in conversion than in control patients. Another study, by Bruce D. Naliboff and collaborators, examined cerebral activation in patients with irritable bowel syndrome and control subjects during recto sigmoid stimulation. The results showed differences in activity of the right hemisphere and hypoperfusion of the nondominant hemisphere in the patients with functional gastrointestinal syndromes. A study on a large adolescent cohort (TRAIL study) suggested specific associations between cortisol responses and certain clusters of functional somatic symptoms. Regarding genetic factors, there is no clear evidence for familial

aggregation of somatic symptom disorders. However, epidemiological studies have suggested that there is a familial predisposition in the case of fibromyalgia. Studies in the 1950s and 1960s had also suggested that conduct disorder during childhood predisposed toward hysteria (somatization disorder) during adulthood and that antisocial personality in males may be associated with hysteria in female relatives.

Psychological Factors

Psychological factors play a significant role in somatic symptom presentations. Such traits as suggestibility, dramatic demeanor, flair, and flamboyance were all related to the classic notion of hysteria. Some of these traits survived in more recent nomenclatures and were included in the personality disorders category ("hysterical personality" and more recently "histrionic personality"). The concept of "alexithymia," used to describe an inability to express and normally process emotions, has also been associated with a tendency to emphasize somatic rather than psychological symptom presentations.

Attitudinal/Behavioral Factors

Learned behaviors and the attainment of "secondary" or "psychological" gain and "advantages" derived from the assumption of the sick role are all factors that may also play a role in the initiation and maintenance of somatoform symptoms.

DIAGNOSIS AND CLINICAL FEATURES

Two separate case reports will be used here to illustrate typical presentations and degrees of severity of somatic symptom disorders. These case reports will highlight diagnostic, prognostic, and management issues for these syndromes.

MILD SOMATIC SYMPTOM DISORDER

Mrs. M, a 40-year-old immigrant from Latin America presented to the primary care clinic with health anxiety and a number of specific somatic complaints. She had no previous psychiatric or medical history. She consumed alcohol occasionally but denied any other substance abuse. A systematic assessment disclosed the presence of several physical symptoms in the last few years, mostly related to pain, that she worried constantly about and that had resulted in either medical consultation or significant distress. These included low back pain, stomach pain, pain in arms, headaches and "dizzy" spells. Upon careful probing, these symptoms did not appear to have a clear medical explanation. Some of these symptoms had first occurred early in life (before the age of 15 years) and had continued intermittently. The Primary Care Evaluation of Mental Disorders (PRIME-MD) elicited diagnoses of Minor Depression and Anxiety Disorder Not Otherwise Specified. Depressive symptoms included diminished interest or pleasure, decreased energy/fatigue, and feelings of worthlessness. Anxiety symptoms included a subjective sense of worry, agitation, muscle pain, concentration difficulties, and irritability. These symptoms did not occur nearly every day, but only a few days per week. Although she reported occasional thoughts that life was

not worth living, she denied suicidal thoughts. The patient was managed in the primary care setting with a modified cognitive intervention that included relaxation techniques (diaphragmatic breathing, recommendation for brief daily naps, etc.) and resulted in a relatively fast improvement of somatic, mood, and anxiety symptoms. She did not require psychotropic medications. Treatment was ended after the 10th weekly session and the improvement in physical symptoms was maintained during a 1-year follow-up period.

SEVERE, PERSISTENT, SOMATIC SYMPTOM DISORDER

Mr. K, a white man in his mid-50s, consulted a general medicine clinic complaining of gastrointestinal problems. Major presenting symptoms were a long list of physical symptoms and concerns mostly related to the gastrointestinal system. These included abdominal pain, left lower quadrant cramps, bloating, persistent sense of fullness in stomach hours after eating, intolerance to foods, constipation, decrease in physical stamina, heart palpitations, and feelings that "skin is getting yellow" and "not getting enough oxygen." A review of systems disclosed disturbances from virtually every organ system, including tired eyes with blurred vision, sore throat and "lump" in throat, heart palpitations, irregular heartbeat, dizziness, trouble breathing, and general weakness. The patient reported that symptoms started prior to the age of 30 years. For more than a decade, psychiatrists, general practitioners, and all kinds of medical specialists, including surgeons, had seen him. He used the Internet constantly and traveled extensively in search of expert evaluations, seeking new procedures and diagnostic assessments. He had undergone repeated colonoscopies, sigmoidoscopies, and computed tomographic (CT) scans, MRI studies, and ultrasound examinations of the abdomen that had revealed Barrett esophagus but no other pathology. He was on disability and had been unable to work for more than 2 years due to his condition. About 3 years before his visit to the medicine clinic, his abdominal complaints and his fixed belief that he had an intestinal obstruction led to an exploratory surgical intervention for the first time, apparently with negative findings. However, according to the patient, the surgery "got things even worse," and since then he had been operated on at least five other occasions. During these surgeries he has undergone subtotal colectomies and ileostomies due to possible "adhesions" to rule out "mechanical" obstruction. However, available records from some of the surgeries do not disclose any specific pathology other than "intractable constipation." Pathological specimens were also inconclusive. A complete physical and neurological examination was normal except for examination of the abdomen, which revealed multiple abdominal scars. Right ileostomy was present, with soft stool in the bag and active bowel sounds. There was no point tenderness and no abdominal distension. During the examination, the patient kept pointing to an area of "hardness" in the left lower quadrant that he thought was a "tight muscle strangling his bowels." However, the examination did not disclose any palpable mass. Skin and extremities were all within normal limits, and all joints had full range of motion and no swelling. Musculature was well developed. Neurological examination was within normal limits. The patient was scheduled for brief monthly visits by the primary care physician, during which the doctor performed brief physicals, reassured the patient, and allowed the patient to talk about "stressors." The physician avoided invasive tests or diagnostic procedures, did not prescribe any medications, and avoided telling the patient that the symptoms were mental or "all in his head." The primary care physician then referred the patient back to psychiatry. The psychiatrist confirmed a long list of physical symptoms that started before the age of 30 years, most of which remained medically unexplained. The psychiatric examination revealed some anxiety symptoms, including apprehension, tension, uneasiness, and somatic components such as blushing and palpitations that seemed particularly prominent in front of social situations. Possible symptoms of depression included mild dysphoria, low energy, and sleep disturbance, all of which the patient blamed on his "medical" problems. The mental status examination showed that Mr. K's mood was rather somber and pessimistic, although he denied feeling sad or depressed. Affect was irritable. He was somatically focused and had little if any psychological insight. The examination revealed the presence of a few life stressors (unemployment, financial problems, and family issues) that the patient quickly discounted as unimportant. The psychiatrist diagnosed DSM-5 Somatic Symptom Disorder, severe type. Although the patient continued to deny having any psychiatric problems or any need for psychiatric intervention or treatment, he agreed to a few regular visits to continue

to assess his situation. He refused to have anyone from his family involved in this process. Efforts to engage the patient with formal therapy such as CBT or a medication trial were all futile, so he was seen only for "supportive psychotherapy," with the hope of developing rapport and preventing additional iatrogenic complications. During the follow-up period, the patient was operated on at least one more time and continued to complain of abdominal bloating and constipation and to rely on laxatives. The belief that there was a mechanical obstruction of the intestines continued to be firmly held by the patient and bordered on the delusional. However, he continued to refuse pharmacological treatment. The only medication he accepted was a low-dose benzodiazepine for anxiety. He continued to monitor his intestinal function 24 hours per day and to seek evaluation by prominent specialists, traveling to high-profile specialty centers far from home in search of solutions.

Comorbidity of Somatic Symptom Disorders

Patients with somatic symptom disorders present complicated social histories, including instances of abuse and neglect. The presenting symptoms can be quite broad, including features of mood, anxiety, personality disorders and, to some extent, substance use disorders (at least the abuse or misuse of prescription medications). In primary care studies, greater than 90 percent of patients presenting with high levels of physical symptoms qualify for a depression or anxiety diagnosis. The comorbidity with other somatoform disorders is also quite elevated, indicating significant overlap across the various categories. Thus, greater than 60 percent of patients with dissociative disorder also meet criteria for somatization disorder.

DIFFERENTIAL DIAGNOSIS

Because somatic symptom disorders are placed at the crossroad between physical and mental disorders, their differential diagnosis tends to be quite inclusive and elaborate. However, there are several features of these disorders that can help the differential diagnosis. Research in primary care has shown that as the number of physical symptoms increases, so does the likelihood that those individuals will also meet criteria for a psychiatric disorder, not a medical disorder.

The following features can help in deciding whether idiopathic physical symptoms may have a psychiatric etiology:

- The symptoms coexist with major psychiatric disorders such as depression or panic.
- The symptoms closely follow traumatic events.
- The symptoms lead to psychological "gratification" or "secondary gain."
- The symptoms represent predictable personality traits (coping mechanisms).
- The symptoms become persistent, join a conglomerate of other symptoms, and convey such attitudes as overuse of medical services and dissatisfaction with medical care.

The more of the foregoing features that are present, the more likely it is that symptoms may be classified as somatic symptom disorder. Differential diagnosis from other psychiatric disorders is also difficult because many of the somatic symptoms may be related to a psychiatric disorder such as depression (e.g., pain symptoms), anxiety (e.g., cardiorespiratory and gastrointestinal symptoms), or even psychotic disorders (somatic delusions). However, the presence of a high level of idiopathic physical symptoms has to be taken into account even if they appear in the context of what is viewed as another primary disorder such as depression or anxiety because they are likely to affect symptom severity, treatment response, and level of disability.

COURSE AND PROGNOSIS

The course of most somatic symptom disorders tends to be chronic, which is to be expected in the case of syndromes so closely associated with personality and cognitive styles. Epidemiological studies have shown that type and number of physical symptoms often change during follow-up periods, suggesting a certain inconsistency in symptom presentation. However, according to international studies, although idiopathic symptoms change, the number of symptoms present appears to remain high. The presence of high levels of idiopathic physical symptoms appears to be very disabling, leading to higher levels of disability than those reported for most medical and psychiatric disorders. Research in primary care has shown that psychiatric disorders such as depression or anxiety become more severe and disabling if they are associated with high levels of medically unexplained physical symptoms.

TREATMENT

Establishing an Alliance

Patients with severe and protracted somatic syndromes, as in the case presented earlier, approach medical encounters with a mixture of magical, unrealistic expectations, pessimism, and distrust of the profession. These attitudes are based on their previous experiences with physicians conveying lack of interest or disbelief in the patient's complaints and suffering. Building a trusting alliance in this context must begin with respect for the patient's symptoms and an acknowledgement of their validity. Active, receptive listening, tolerance for repetition, and a "neutral" approach (avoiding being dismissive, confrontational, or overly reassuring) are essential skills. It takes time and persistence to develop a trusting, therapeutic partnership.

Taking the History

Many of these patients usually bring with them "thick" charts. These

include descriptions of many clinical encounters and multiple tests and procedures that are often redundant and ordered without clear rationale. The prospect of reviewing these medical records is challenging and often leads to a negative attitude from the physician. The physician should keep an open mind, despite forewarnings in the medical records, and perform an independent assessment of the patient. An emphasis on psychological questioning and interpretations should be avoided at this stage. Premature reassurance, although seemingly appropriate from the physician's perspective, may be perceived by the patient as disinterest or dismissiveness.

Reassurance

The humanistic principles of the medical profession are well summarized in the maxim attributed to Oliver Wendell Holmes that the role of the physician is "to comfort always." Beyond the specificity of diagnoses and the effectiveness of treatments, the traditional practice of medicine gave high importance to bedside manner and interpersonal issues. However, 21st century medicine, with its reliance on technology and the changes in practice patterns (e.g., managed care and complex electronic medical records), leaves little time for face-to-face interactions, which is gradually eroding the doctor–patient relationship. It would appear that the profession has drifted away from the role "to comfort," meaning "to give strength and hope; to ease the grief or trouble; to cheer," to "to reassure," simply meaning "to assure anew; or restore to confidence." In this new context, "reassurance" simply means letting patients know that their symptoms do not appear to be caused by physical disease. Unfortunately, this does not appear to work well in the case of many patients presenting with "idiopathic" physical symptoms. The timing and degree of reassurance must be based on adequacy of data and the trust and security of the relationship.

Physical versus Psychological Focus

Patients with somatic symptom disorders appear to feel most comfortable providing details of symptoms in physical or somatic terms. However, they may also bring a repertoire of significant events (e.g., losses, trauma, disappointments, and deprivations) that may be related to the onset or persistence of symptoms and need to be explored at some point. However, many patients do not readily acknowledge or recognize "emotional" issues and will feel more comfortable dealing with questions related to their physical symptoms than questions related to psychological issues. Taking a history of multiple physical complaints can be carried over into subsequent appointments. A thorough physical examination should follow history taking at the initial visit, and briefer physical assessments should be

performed in subsequent visits. An inclusive drug history including prescription and nonprescription remedies also should be completed. As the history taking moves along, attitudes, beliefs, and attributions should become clearer and patterns of interaction and illness behavior discernible. Distorted beliefs, contradictory ideas, and fears can be addressed at several points during the process of gathering historical data.

General Therapeutic Approach

Because of the refractoriness of idiopathic physical symptoms to traditional interventions, physicians and therapists tend to be more successful when adopting a "caring" rather than a "curing" approach with these patients. Restraint in the use of medication is advised, unless these are clearly indicated for the relief of comorbid affective and anxiety symptoms. Clear assignment of appointment times at regular but relatively infrequent intervals (about once per month) is preferable to random appointments based on whim or fluctuation in symptoms. A clearly spelled out treatment plan is highly desirable. As the rapport and trust increase, the patient may be more willing to discuss his or her social world, family relationships, and the "stressors" that may be contributing to the symptoms. Although true insight may never fully develop, some awareness of emotional components could make the patient more receptive to referral for specialized behavioral intervention. Even if referral is successful, the patient should continue to see both the primary care physician and the psychiatrist. An ongoing working collaboration between the psychiatrist and the primary care physician should lead to better outcomes.

SPECIFIC SOMATIC SYMPTOM AND RELATED DISORDERS

Somatic Symptom Disorder

Introduction. In DSM-III, "somatization disorder" was the direct descendant of the hysteria/Briquet syndrome diagnosis described in the pioneer studies of the St. Louis group as one of the few valid diagnoses in psychiatry. This separated the concept of multiple idiopathic physical symptoms from the concepts of conversion and dissociative hysteria, which became separate diagnostic entities in the new nosologies. Once incorporated in the DSM process, "somatoform disorder" underwent various changes during the subsequent revisions of the nomenclature up to DSM-5. In clinical practice, somatization disorder appeared to be quite rare and there was disillusionment about its relevance in primary care settings and, in general, pervasive pessimism about its treatment and prognosis. "Full" somatization disorder (in DSM-III and IV) represented the severity extreme in dimensional models of somatization, and there was ample debate in the field about its survival in future editions of the DSM. Indeed, DSM-5 eliminated this diagnosis, merging it into the broader category of

somatic symptom disorder.

As discussed earlier, hysteria was the precursor of the concept of somatization and also of the DSM diagnosis of somatization disorder. In the golden age of descriptive psychopathology, the formal description of a syndrome with multiple neurological symptoms/complaints without organic findings (still called hysteria) was made by Pierre Briquet in France in 1859, and this began the separation of this polysymptomatic disorder from conversion hysteria. Toward the middle of the 20th century, North American nosologists from Washington University and Barnes Hospital in St. Louis started using a systematic approach to validate psychiatric diagnoses. One of their first targets was the diagnosis of hysteria. These investigators proposed a "revised" Briquet syndrome that included many other physical symptoms besides the neurological ones, as well as psychiatric and behavioral components. In the DSM classification, before DSM-III established the syndrome as Somatization Disorder, it had been called "conversion reaction" in DSM-I (1951) and "hysterical neurosis, conversion type" in DSM-II (1968).

Comparative Nosology

EVOLUTION OF THE DIAGNOSIS OF SOMATIZATION DISORDER. The confusion that arose in differentiating hysteria from hysterical conversion reactions in DSM-III was at least temporarily resolved by retaining Briquet's hysteria as Somatization Disorder and hysterical conversion as Conversion Disorder. Briquet's virtual encyclopedia of purely neurological symptoms was reduced in DSM-III to 14 symptoms for women and 12 for men out of a possible 37 symptoms attributable to several organ systems (cardiorespiratory, musculoskeletal, genitourinary, neurological, and gastrointestinal). Rather arbitrarily, DSM-III-R revised the criteria to require 13 symptoms for both males and females. Finally, in DSM-IV the criteria were changed once again. This time the revisions were stimulated by a need to simplify the criteria and by the reliance on epidemiologic data sets from outside of the United States. The same data set of adopted Swedish community residents for whom detailed medical records were available and which had led to the distinction of subtypes of somatization in the 1980s (high-frequency vs. low-frequency somatizers) was used again for this purpose. A method based on discriminant analysis was derived to detect "somatizers," and this led to four discrete sets, including bodily pains, gastrointestinal symptoms, pseudo neurological symptoms, and sexual or menstrual symptoms, that were the basis for the new diagnostic formulation. Thus, DSM-IV continued the deflationary trend and modified the criteria to having at least eight physical complaints referable to four pain sites or functions (e.g., back, chest, urination), two nonpain gastrointestinal symptoms (e.g., nausea, bloating), one nonpain sexual or

reproductive system symptom (e.g., menstrual irregularity, loss of libido), and one pseudo neurological symptom (e.g., urinary retention, aphonic, blindness). To qualify as symptoms of somatoform disorder, somatic symptoms should remain medically unexplained, although "exaggerations" of ordinarily expected symptoms of coexisting physical disease can be also considered in the symptom count. Somatoform disorder symptoms generally occur over a period of several years and are accompanied by significant impairment of social and occupational function and high utilization of medical resources, usually resulting in either ineffective and/or unnecessary medical/surgical treatments as illustrated in the third case presented earlier.

As mentioned earlier, DSM-5 took a new tack, completely redesigning this category, renaming it "somatic symptom and related disorders." There was concern about the very limited use of somatoform disorder diagnoses both in psychiatric practice and primary care. Somatization disorder in DSM-III and IV was a restrictive category, based on simple counts of medically unexplained symptoms, thus neglecting important attitudinal and behavioral elements.

The reliance on absence of medical explanations to catalog symptoms, was not precise or practical. Therefore, DSM-5 added key attitudinal and behavioral requirements (B criterion) and no longer requires that symptoms be medically unexplained.

In DSM-IV, hypochondriasis overlapped with somatization disorder and the label itself was stigmatizing. Pain disorders were quite nonspecific and these diagnoses were predominantly used in settings devoted to pain management. The most common somatoform diagnoses used in practice were undifferentiated somatization disorder and somatoform disorder NOS, both quite imprecise and rather meaningless labels. Thus, DSM-5 eliminated a few categories, renamed most and merged others (see Table 18–1 for criteria for DSM-5's somatic symptom and related diagnoses).

Epidemiology. The prevalence of somatic symptom disorder is not yet known. Because "medical explanation" for the symptom is no longer an exclusion and the somatic symptom threshold is much lower than that of DSM-IV, it has been argued that the prevalence of the DSM-5 disorder could be much higher than that of "full" somatization disorder in DSM-IV. However, the requirement in DSM-5 for "disproportionate and excessive thoughts, feelings, and behaviors" in addition to the reported somatic symptoms, makes the criteria more restrictive, and this may affect its overall prevalence in the opposite direction, that is, lowering the prevalence. However, there first needs to be well-designed, systematic studies in communities and clinical settings. These studies may require the use of more sophisticated instruments such as the SCAN-PSE that can elicit both the symptoms and the psychological/behavioral components of this

disorder.

Worldwide, a large majority of the comparative surveys that examined the old diagnosis of somatization disorder consisted of simple counts of physical symptoms with some effort made to rule out medical explanations. In many studies, the presence of high levels of physical symptoms was derived from physical symptom inventories such as those included in the Symptom Checklist 90 (SCL-90), the General Health Questionnaire (GHQ), and particularly the Diagnostic Interview Schedule (DIS) and the CIDI. The somatic syndromes were articulated as either "full" or "abridged" somatization disorders. Problems with these studies include variations in key instruments, ways of survey administration, and format of answers and the reliance on lifetime rather than current symptoms. In general, somatoform disorder and abridged somatoform disorder have been reported to be more common among females, individuals from lower socioeconomic strata, and, in the United States at least, among people from certain ethnic groups, such as Latinos. In addition, the comorbidity of high levels of these symptoms with psychiatric syndromes such as depression and anxiety has been well documented for most cultures and ethnic groups, although there appears to be cross-cultural variation in frequency and severity of the associated physical symptoms.

GENERAL POPULATIONS. Studies in the United States, Puerto Rico, Germany, and Italy found lifetime prevalence rates of "full" somatization disorder ranging from 0.1 percent in the United States to 0.8 percent in Germany, whereas lifetime rates of "abridged" somatization disorder ranged from 5.6 percent in Germany to 19 percent in Puerto Rico. Much greater prevalence rates have been reported for more broadly defined somatic symptom clusters. For example, 22 percent of the general population in Germany reported at least one unexplained physical symptom leading to severe impairment. Unfortunately, some of the most recent large-scale epidemiological studies in the United States and elsewhere, such as the National Comorbidity Survey (NCS), the National Epidemiologic Survey of Alcohol and Related Conditions (NESARC), and the World Mental Health Surveys, did not systematically assess unexplained physical symptoms, possibly due to the time and effort it takes to go through long and tedious physical symptom inventories.

Etiology

GENETIC AND BIOLOGICAL FACTORS. Some of the early family studies of the St. Louis group had suggested an association between hysteria/Briquet syndrome in females and antisocial personality in their male first-degree relatives. In addition, studies of adopted-away children highlighted the presence of alcoholism and violence in biological fathers of women with somatization disorder. However, no family studies have been performed

since the new diagnostic criteria were developed.

Treatment

GENERAL CONSIDERATIONS. Historically, the treatment of patients with severe somatic symptom disorders has been considered difficult and frustrating. Encounters of these patients with their physicians generally evoke frustration on both sides. When Freud began his practice as a clinical neurologist in Vienna, many of his patients (clearly many of those on whom the original theories of psychoanalysis were based), were "difficult" patients with dramatic physical presentations who were referred by other physicians who had failed to help them. Although the form of symptoms has been reshaped by cultural change and shifts in medical paradigms, this situation is not that different from the realities faced by many physicians today. Such a context represents a profound challenge to the most well-intentioned, dedicated physicians. Besides the numerous physical symptoms, somatoform disorder embodies traits that complicate therapeutic encounters and may render patients refractory to treatment. These traits include denial of psychological components; chronicity; lack of response to traditional types of reassurance; unshakeable belief about the presence of physical disease; high/inappropriate utilization of medical resources; risk of "occult" medical comorbidities; sensitivity to rejection; and frequent dysfunctional patient–physician relationships. Therapeutic goals should be modest at first and be limited to small, attainable gains such as a decrease in medical visits, a commitment to a single primary care physician, and the avoidance of unnecessary tests and procedures. The physician must be prepared for a long-term commitment to the patient. Specialist consultations should be limited and restricted to cases with evidence of comorbid physical conditions. Some of the elements of CBT such as diary keeping should be incorporated into the primary care physician's treatment strategy. Recommendations of such activities as exercise, yoga, relaxation, meditation, and massage may be useful and are generally better accepted by these patients than psychological recommendations. The major challenge, therefore, continues to be finding practical and effective ways to help primary care physicians to recognize, understand, and manage these patients in the context of a supportive doctor–patient relationship at the primary care site. Current emphasis on learned behaviors leads to the choice of cognitive-behavioral interventions, whereas other conceptualizations may encourage more relational, interpersonal, or psychodynamic approaches. Psychopharmacological approaches may be used independently or be combined with other interventions. Until recently, interventions with somatoform disorder patients generally had a pessimistic outlook and there was a paucity of well-controlled studies. Other issues that delayed effective interventions included the fact that

somatoform disorder patients were more likely to be seen in primary care settings where therapists may be poorly prepared to work with demanding, frustrating patients and comorbid physical disease and psychiatric disease are frequent confounders.

PHARMACOTHERAPY. Medications should be generally avoided in the management of the patient with somatoform disorder except in the presence of clearly delineated anxious, depressive, or psychotic symptoms. Despite the development of new pharmacological agents such as dual (norepinephrine and serotonin) reuptake inhibitors and a preliminary report on the possible efficacy of gabapentin, there is no evidence that a purely pharmacologic approach targeting the treatment of somatoform disorder will bear much fruit.

SMITH'S CONSULTATION LETTER. A groundbreaking controlled study by Richard Smith and colleagues at the University of Arkansas for Medical Sciences demonstrated the effectiveness of a rather simple intervention. This consisted of a brief "consultation letter" intended for primary care physicians providing them with "do's and don'ts" regarding their encounters with patients with multiple medically unexplained physical symptoms and briefly instructing them on the use of a few key management techniques. The letter urged the physicians to see these patients during regularly scheduled appointments, perform brief physical examinations focusing on the area of discomfort at each visit, avoid unnecessary diagnostic procedures, invasive treatments, and hospitalizations, avoid using statements such as "symptoms are all in your head," and briefly allow/encourage patients to talk about "stressors." Although patients did not show a very significant change in terms of somatic symptom relief per se, their functional capacity improved significantly and there was a decrease in their utilization of health resources that generated significant cost savings.

REASSURANCE AND SUPPORTIVE THERAPY. It is generally accepted that patients with somatoform disorder respond poorly to the simple reassurance that there are no positive findings in the physical and laboratory workup. Although no adequately controlled studies of psychotherapeutic intervention for "full" somatoform disorder existed prior to the study by Smith and colleagues, a number of studies of undifferentiated forms of somatoform disorder had suggested therapeutic benefits from the presence of an accepting attitude in the therapist, a shift in the patient's attention from somatic to emotional features, and the use of group approaches focusing on explanation, support, relaxation, and cognitive-behavioral approaches emphasizing adaptation to chronic somatic distress.

PSYCHOTHERAPY. There are reports indicating that brief psychodynamic

therapy of unexplained somatic symptoms may be effective, and at least one of these studies was a randomized, clinical study comparing brief psychodynamic therapy to regular medical treatment alone. However, most of the recent evidence in terms of randomized clinical trials involves CBT.

Cognitive-Behavioral Therapy. A body of research, particularly within the last decade, has reported on the efficacy of CBT for the treatment of full somatization disorder and related syndromes. CBT-type interventions appear to help patients by modifying thoughts and behaviors associated with somatization. In the typical CBT program, patients are systematically exposed to a number of behavioral techniques, which include relaxation training and graded increases in activities. From a cognitive perspective, CBT helps these patients to identify thoughts that contribute to increased stress, inactivity, and health concern. Often, patients with somatoform disorder syndromes may have a tendency to think catastrophically about their physical symptoms. Such thoughts lead them to conclude that they are sick and must limit physical activity, creating a cycle that perpetuates the somatic process. There are several large, well-designed studies that support the use of CBT-type interventions for patients suffering from somatization disorder and related syndromes. In primary care settings, studies between 1995 and 2000 strongly supported the efficacy of CBT for functional somatic syndromes.

In Germany, the group led by Professor Winfried Rief did a study in 140 primary care patients randomized to either *one* session of CBT or "standard medical care" and reported excellent improvement in both groups (and very good acceptance of CBT). It must be noted that in Germany, there exist psychosomatic units in hospitals that specialize in managing these patients (and have the time and the needed resources to do so thoroughly). It is generally acknowledged that to effectively treat patients with somatoform disorder, primary physicians must be recruited as participants. In the United States, Robert Smith and colleagues recently reported on a randomized clinical trial of patients with "high service utilization" who had medically unexplained symptoms treated with cognitive-behavioral and pharmacologic interventions by primary care clinicians (nurses). This intervention led to significant clinical improvement in mental and physical functioning.

Pivotal Study. In a 2006 study performed at a mental health care system in New Jersey by L. A. Allen and coworkers, 84 patients meeting full criteria for somatization disorder were randomly assigned to either standard medical care augmented by a psychiatric consultation letter or to a ten-session, manualized, individually administered CBT added to the psychiatric consultation letter. The primary outcome measure was the severity scale of the Clinical Global Impression Scale for Somatization

Disorder (CGI-SD). Secondary outcome measures were responder status as determined by clinical ratings, self-report measures of physical functioning and somatic symptomatology, and health care utilization assessed via medical records. Results of this study showed that 15 months after baseline, somatization symptoms were significantly ($p < 0.001$) less severe in the group treated with CBT. Patients treated with CBT also were significantly more likely to be rated as either "very much improved" or "much improved" than patients treated with only augmented standard medical care (39.5 percent vs. 4.9 percent). CBT was associated with greater improvements in self-reported functioning and symptomatology and led to a greater decrease in health care costs. The authors concluded that for patients diagnosed with somatoform disorder, CBT might produce clinical benefits beyond those that result from current state-of-the art treatments.

Illness Anxiety Disorder

Introduction. Together with conversion hysteria, the old notion of hypochondria has had a long and eventful history in psychiatry. Given the stigma attached to the original term, the name changed in DSM-5. DSM-5 recognizes that most cases of hypochondriasis can be subsumed under the rubric of Somatic Symptom Disorder. However, a minority of patients with hypochondriasis (~20 percent) are now given a diagnosis of Illness Anxiety Disorder. The distinction lies in the severity of the somatic symptoms. Where somatic symptoms are relatively minor and patients instead focus on concerns that they will get sick or have an undiagnosed illness, the diagnosis is Illness Anxiety Disorder. When there is preoccupation with health concerns superimposed by substantial somatic symptoms, the diagnosis is somatic symptom disorder. Phenomenologically, health anxiety can be broken down into several components, including idiopathic physical symptoms, such as amplification of bodily sensations; certain characterological features, such as low self-esteem, guilt, passivity, and personal suffering; and learned behaviors, such as the assumption of the sick role.

Definition. In DSM-5, key features of health anxiety are preoccupation with physical symptoms and the belief that they may be due to a serious physical illness that persists despite medical reassurance or objective evidence on the contrary. Some experts subscribe to a dimensional concept of health anxiety, according to which, these symptoms exist on a continuum from heightened awareness of bodily function to extreme delusional bodily preoccupation. Indeed, the fear of disease can become so intense that it has been given a phobic connotation, as is the case with historical terms such as "syphilophobia," "thanatophobia," "AIDS-phobia," and several others. These fears are particularly prevalent at times when certain diseases are thought to be incurable or terminal. The preoccupation

with health issues can also become fixed and reach delusional (psychotic) proportions.

History. Hippocrates was perhaps the first to point out that this disorder had its anatomical origin in organs located under the hypochondrium—literally below the cartilage or ribcage—and this notion was expanded by other classical Greeks such as Plato and Diocles. In the 16th and 17th centuries, Robert Burton, Thomas Willis, and Thomas Sydenham attributed the origin of hypochondria to "disturbances of the mind," and it was viewed by some as the male equivalent of hysteria. Since those early times, clinical descriptions of hypochondria have seemed very consistent with those formalized in modern diagnostic systems. In the mid-1600s the French dramatist Molière in his satire "Le Malade imaginaire" depicted a hypochondriac who feared the ministrations of physicians. Ironically, Molière, who suffered from pulmonary tuberculosis, collapsed on stage while playing the lead role and died later that evening.

Comparative Nosology. In DSM-5, illness anxiety is defined as the preoccupation with fears of having, or the idea that one has, a serious disease based on the person's misinterpretation of bodily symptoms, with a duration of at least 6 months. The preoccupation persists despite appropriate medical evaluation and reassurance; the belief is not of delusional intensity and is not restricted to a circumscribed concern about appearance. The preoccupation causes clinically significant distress or impairment in social, occupational, or other important areas of functioning; and the preoccupation is not better accounted for by another disorder such as anxiety disorder or depressive disorder. The ICD-10 outlines similar criteria, although it requires that the preoccupation with disease extends to a minimum of two disease entities, leads the patient to seek medical help or investigations, and does not involve phobic fears. In addition, in ICD-10, presumed body disfigurement or deformity is included in the diagnostic definition. Hypochondriasis has also been classified as "primary" or "transient" or "secondary" if symptoms are observed temporarily to follow a major physical event (e.g., cancer surgery) or if they are judged to be due to another mental disorder such as depression.

Epidemiology. The prevalence of health anxiety can vary according to how it is defined, how the criteria are elicited, and the type of settings studied. It has been estimated that it may affect as many as 4 to 6 percent of all patients. Health fears seem quite common in the general population. A national study in Germany found that 10 percent of the population reported "health anxiety" or the conviction of having a serious disease in the absence of medical explanations. A study in Italy reported a 1-year prevalence of health anxiety of about 5 percent. In primary care, studies

using various methodologies have shown prevalence rates of health anxiety in the 3 to 7 percent range. Psychiatric comorbidities are quite common among patients with health anxiety including depression (40 percent), panic disorder (10 to 20 percent), OCD and generalized anxiety disorder (5 to 10 percent). Longitudinal studies have also shown that the course of health anxiety can become chronic and incapacitating, with a majority (greater than 60 percent) of the cases still showing symptoms after several years of follow-up.

Treatment

REASSURANCE. The role of reassurance in health anxiety disorder has been a controversial one. Although the criteria includes inability to reassure the patient as a criterion for diagnosis, virtually all therapeutic interventions include reassurance as one of their main features. The ease with which patients with health anxiety can be successfully reassured may be related to patient factors such as chronicity, severity of symptoms, and personality characteristics and also to attitudes and treatment style of the therapist. Recommendations for approaching patients with health anxiety are generally similar to those already given for approaching patients with somatic symptom disorder. However, in general, patients with health anxiety seem less demanding than those with somatic symptom disorder, at least the severe types and some may even respond to properly conveyed reassurances. Once patients with health anxiety are systematically engaged in a therapeutic relationship, opportunities for insight development may arise.

PHARMACOLOGICAL TREATMENT. Little information is available on the pharmacological treatment of hypochondriasis. Most reports have been anecdotal, single blind studies. Such agents as tricyclics, selective serotonin reuptake inhibitors (SSRIs), and nefazodone (Serzone) have been reported to be clinically useful particularly in health anxiety associated with depressive symptoms. In secondary syndromes, symptoms of health anxiety have been reported to improve in parallel with anxiety or depressive symptoms. Antipsychotic agents such as risperidone have been recommended for delusional syndromes. No study has established the superiority of one drug over another in managing health anxiety. The only double blind, placebo-controlled clinical trial in patients with health anxiety (hypochondriasis) was that of Brian Fallon and colleagues. Following a placebo run-in, the investigators randomly assigned patients to receive fluoxetine (Prozac) in dosages ranging between 20 and 80 mg per day or a placebo. Fifty-seven patients were enrolled. Twelve discontinued during the placebo run-in and 45 were randomized to either fluoxetine or placebo for 12 weeks. Responders to acute treatment (much/very much improved in CGI) entered a 12-week maintenance phase to week 24.

Sustained responders at week 24 entered a double-masked discontinuation phase. More patients improved on fluoxetine compared to placebo starting at week 8 (50 percent vs. 19 percent, $p = 0.03$). The acute response was maintained to week 24, with more responders. The authors conclude that fluoxetine is a moderately effective and well-tolerated treatment for hypochondriasis.

PSYCHOTHERAPIES. Various group and individual therapies, including psychodynamic types, have been proposed over the years to manage patients with health anxiety. Unfortunately, there is a paucity of controlled studies. Robert Kellner, one of the pioneers in the treatment of somatic symptom syndromes was one of the first to report positive results of behavioral interventions in health anxiety by focusing on patients' misinformation and cognitive distortions.

Cognitive-Behavioral Therapy. Several recent controlled studies have confirmed the earlier suggestions of the efficacy of CBT interventions. This substantive evidence places CBT as the prototype, first-line treatment for this condition. For example, a review of several controlled studies showed that CBT-treated patients improved significantly more than control subjects in a great majority of the studies, with a "trend" toward improvement observed in the remaining trials.

Pivotal Study. In a 2004 study by Barsky and Ahern, patients from primary care practices and volunteers responding to public announcements, all with high hypochondriasis scores, were assigned to one of two treatment groups, either a six-session, individual CBT intervention or a "medical care as usual" group. The CBT was accompanied by a consultation letter sent to the patient's primary care physician. Main outcome measures included hypochondriacal beliefs, health fears, somatic symptoms, role function, and impairment. One hundred and two individuals were assigned to CBT and 85 were assigned to "medical care as usual." At the 12-month follow-up, CBT patients had significantly lower levels of hypochondriacal symptoms, beliefs, and attitudes ($p < 0.001$) and health-related anxiety ($p = 0.009$) than those in "medical care as usual." They also had significantly less impairment of social role functioning ($p = 0.05$) and intermediate activities of daily living ($p < 0.001$). Hypochondriacal somatic symptoms were not improved significantly by treatment. Barsky and Ahern concluded that this brief, individual CBT intervention, developed specifically to alter hypochondriacal thinking and restructure hypochondriacal beliefs, appears to have significant beneficial long-term effects on the symptoms of hypochondriasis.

Conversion Disorder (Functional Neurological Symptom Disorder)

Introduction. In DSM-IV, conversion disorder is the closest remnant of the old concept of hysteria. In its classic form (a motor or sensory deficit), it now appears to be a rare disorder, at least in Western societies. Although it was a key component of the classic syndrome of hysteria, conversion was separated from hysteria starting with the Briquet and the St. Louis criteria and passed as a separate category into the DSM system. In DSM-IV, the diagnosis of conversion disorder continued to invoke etiological assumptions (psychological factors) against the otherwise "atheoretical" background of the system. DSM-5 made two crucial modifications in diagnosing this condition. First, it appended a parenthetical diagnosis ("functional neurological symptom disorder") in recognition that this term is more familiar neurologists and arguably more acceptable to patients. Second, it clarifies that the diagnosis is made when the symptom is incompatible with known neurological disorders.

Definition. A conversion reaction is a rather acute and temporary loss or alteration in motor or sensory function that is incompatible with known neurological disorders. Because the onset frequently coincides with psychological issues (conflict), early theorists speculated that such issues were "converted" to neurological symptoms. The problem is that stress is omnipresent in life, and many patients present with conversion when no stressor has been demonstrated. The classic syndromes resemble neurological syndromes. Conversion motor symptoms mimic syndromes such as paralysis, ataxia, dysphagia, or seizure disorder (pseudo seizures), and the sensory ones mimic neurological deficits such as blindness, deafness, or anesthesia. There also can be disturbances of consciousness (amnesia, fainting spells). Nonneurological syndromes such as pseudocyesis (false pregnancy) or psychogenic vomiting have also been placed under the conversion disorder category. However, many clinicians continue to reserve the term conversion reaction for syndromes mimicking a neurological disease. In ICD-10 but not DSM-5, dissociation and dissociative disorders are attached to the somatoform category, closely linked to conversion syndromes. The demarcation between conversion disorder and somatic symptom disorder is not that clear, and conversion symptoms may form part of the constellation of symptoms seen in somatization disorder. Patients with conversion disorder usually present with symptoms suggestive of neurological disease such as muscle weakness, gait disturbance, blindness, aphonia, deafness, convulsions, or tremors. About one-third of patients diagnosed with conversion disorder presenting with motor symptoms also meet criteria for other Axis I psychiatric diagnoses, and 50 percent meet criteria for Axis II diagnoses.

History. The concept of *hysteria,* derived from the Greek word for womb or uterus, implied an unwanted migration of the organ to higher

sites. It was apparently known even to ancient Egyptians. In the middle Ages, hysterical symptoms were attributed to demonic influences, and their being placed at the moral level retarded the medical debate. During the Renaissance, hysteria returned to medicine, being considered a somatic disorder by physicians, who implied a connection or pathway between the uterus and the brain. By the middle of the 19th century, Briquet provided a detailed clinical description of a somatic syndrome affecting young women to which he gave the name hysteria and whose origin he situated in the brain. Charcot described the classic form as "la grande crise hystérique," which included phases such as the "prodromal," "epileptoid," "trance," and "terminal or verbal" phases. He also proposed the term "functional lesion" in an effort to resolve the absence of physical findings. His followers, Babinsky, Janet, and Freud, continued to place the emphasis on psychological factors. By the time Freud proposed his psychodynamic theory, Paracelsus, Thomas Willis, Sydenham, Mesmer, and Pinel had already suggested that hysteria had a psychogenic origin. Freud and Breuer jointly reported the first case of hysterical conversion, the case of Anna O. They theorized that symptoms of hysteria represented unwanted emotional distress or conflict that was suppressed and kept unconscious by the individual, only to appear in the form of medically unexplainable bodily symptoms. Freud named this process "somatic compliance" or "conversion," and thus, with the case report of Anna O, both "conversion hysteria" and the "talking cure" were born. Since then, however, conversion hysteria has gradually lost its central role in psychopathology, being more loosely described as hysteria, hysterical conversion, or simply conversion, and the term is often used in situations without clear evidence of psychological determination.

Comparative Nosology. In DSM-I (1951), hysteria became conversion reaction; in DSM-II (1968), hysterical neurosis, conversion type; and finally, in DSM-III (1980), conversion disorder, the term that remains to this date. In DSM-IV, conversion disorder is defined as the presence of one or more symptoms or deficits affecting voluntary motor or sensory function, suggesting a neurological or other medical condition; judged to be associated with psychological conflict or other stressors, not intentionally produced or feigned; and not fully explained by other medical conditions, use of substances, or cultural nuances. In DSM-IV, conversion symptoms are not limited to pain or sexual dysfunction, should not be a component of somatization disorder, and should not be considered due to or caused by another mental disorder. Of interest, although DSM-IV purports to be atheoretical, it continues to pay homage to psychodynamic tradition in this instance, based on the clinical presentation that originally set the psychoanalytic movement in motion. DSM-5 emphasizes that the neurological symptoms are incompatible with recognized neurological or

medical disorders and provides a specifier to denote whether or not a psychological stressor has been identified.

Epidemiology. The epidemiological information on conversion disorder is limited. Estimates vary broadly: Less than 1 percent in the general population, 5 to 14 percent among general hospital medical/surgical referrals to psychiatry consultation services, and 5 to 25 percent in treated psychiatric outpatients. The disorder appears to be more frequent in females and can be seen in children as young as 7 or 8 years old. It is rare after the age of 35 years.

Etiology

PERSONALITY FACTORS. Some of the traditional features of hysteria included a detached, unemotional, calmed attitude in front of what appears to be a severe and turbulent illness, a trait known as "la belle indifference." "Histrionic personality" is currently the term that summarizes the drama, flair, and flamboyance and exhibitionism attributed to these patients. However, none of these is given any relevance for the DSM-5 diagnosis of conversion disorder.

BIOLOGICAL FACTORS. The neurophysiologic conceptualization speaks of an inherent defect in certain brain functions, especially those in the dominant hemisphere that may interfere with verbal associations. In addition, functional MRI studies have shown differences in brain activation between patients with conversion disorder and controls.

PSYCHOLOGICAL FACTORS. The behavioral theory attributes conversion disorder to faulty childhood learning, with the nonadaptive behavioral responses used for secondary gain and control of interpersonal relationships. The psychoanalytic theory, on the other hand, describes symptoms as compromise formations with primary gain of conflict resolution through partial expression of the conflict without conscious awareness of its significance. Some have suggested a strong relationship between childhood traumatization by sexual or physical abuse and a later propensity for conversion disorder. Other studies, however, do not confirm such an association.

Treatment. Although a variety of explanations and treatments have been reported in the literature, there is little systematic research on the treatment of conversion disorder. Improved guidelines for treatment would be an important contribution, but these need to wait on methodologically controlled studies and long-term follow-up. Although psychological etiology is a requisite of DSM-IV diagnosis, it is often difficult to identify it early in treatment and may not reveal itself until extensive additional

history is obtained. Screening methods for assessing trauma, dissociative experience, and posttraumatic stress disorder (PTSD) should be helpful in this process. Any assessment of conversion symptoms must take into account idiosyncratic cultural or religious beliefs that may require specific culture-bound interventions for resolution of conflict and symptom amelioration. DSM-5 requires that careful examination document that the neurological symptom is incompatible with known neurological disorders.

APPROACH TO THE PATIENT. The usual caveats on history taking described for other disorders also apply here. There should be special attention given to history of trauma, sexual and physical abuse, and family history of conversion symptoms. Physical examination must pay particular attention to ruling out neurological diseases such as multiple sclerosis and other peripheral and central nervous system disorders. Routine laboratory studies are indicated, as well as electroencephalograms (to distinguish between epilepsy and pseudo seizures) and other special studies (e.g., MRI, x-rays, spinal tap, etc.) to rule out other possible organic etiologies.

REASSURANCE. Many conversion syndromes have an acute, benign course and may remit spontaneously with understanding and support. Early intervention can forestall potential chronicity and the progression into a well-entrenched somatization disorder.

HYPNOTHERAPY. This is perhaps the oldest treatment used for conversion disorder. However, there are very few randomized, controlled trials, and these seem to be inconclusive because hypnotizability was not predictive of treatment outcome.

SODIUM AMYTAL. The use of sodium amytal for diagnosing conversion disorder, uncovering traumatic events, and providing therapeutic relief was for some time a common approach to the patient with conversion hysteria. However, it is now rarely used because a number of observations questioned its clinical value.

PSYCHOTHERAPY. Once chronicity has developed, intensive treatment may use all treatment modalities, including hospitalization, individual or group therapy, insight-oriented therapies, behavioral techniques, hypnosis, sodium amytal interview, physical therapy, biofeedback, relaxation training, and medication (primarily for comorbid anxiety, depression, or other somatoform disorders). Psychological interpretations or explanations do not work well early in the process, but reassuring patients that critical tests are normal and that symptoms will eventually improve may be helpful. Any implication to the patient that he or she is malingering is very counterproductive. Behavioral interventions should focus on improving self-esteem, the capacity for emotional expression and assertiveness, and

the ability to communicate comfortably with others.

PHARMACOLOGICAL TREATMENT. Accompanying comorbid depression, anxiety, and behavior problems may respond to pharmacologic interventions. The use of hypnotic or narcoleptic techniques, if these are being considered, must be tentatively offered to patients whose fear of passivity or loss of control may induce overwhelming anxiety.

PHYSICAL THERAPY. With chronic conversion, muscle contractures can occur and physical therapy is necessary. Even in the absence of such contractures, however, many conversion patients find that physical therapy can be helpful for muscular symptoms or balance problems. The process of slowly progressive exercises and activity can help restore functioning.

Definition. Somatic Symptom Disorder Patients presenting predominantly with pain symptoms:

These patients complain of severe, persistent pain in one or more physical areas that are of sufficient severity to warrant clinical attention and cause significant impairment in function. Like the other presentations of Somatic Symptom Disorder, there is no assumption that the symptom is psychogenic or that there can be a clear-cut distinction between "physical" pain and "psychological" pain. The diagnosis of Somatic Symptom Disorder with pain as a specifier is made when the symptoms of pain are accompanied by extensive cognitive, affective, and behavioral repertoires. Differential diagnosis of "psychologically" versus "physiologically" induced pain is difficult. This task is made more complicated by incidental findings such as arthritic and degenerative changes in joints, vertebral disc problems, muscle spasms, and so on. In the presence of other psychiatric disorders such as depression or anxiety, it is often impossible to ascertain whether the pain is primary or secondary (due to the other disorders).

Theories of psychogenic pain formation invoke neural pathways, but they can be best understood if they are framed within a more comprehensive biopsychosocial model. The type and quality of pain response have been attributed to genetic and physiological vulnerabilities by some authors, whereas others emphasize psychosocial factors as the major determinants of the symptoms, including behavioral mechanisms. The notion of a "pain-prone personality" has waned in recent years, but it was part of a psychodynamic concept of chronic pain, with pain viewed as a psychological defense against guilt and anxiety possibly representing self-punishment. Patients with chronic pain often meet criteria for DSM personality disorders, but no specific personality profile has emerged as more relevant than others. Pain in the context of other depressive symptoms and in the absence of dysphoria or sadness was often called "depressive equivalent" or "masked depression," but this has never been officially endorsed. For a very long time, psychogenic pain complaints were

considered to be closely related to hysterical phenomena and described as part of classical hysterical syndromes. In previous classifications, pain of unclear origin had been called "idiopathic," a term that continues to have relevance in recent classifications of pain in specialties such as neurology. In the United States, pain was separated from conversion disorders, and the category of "pain-prone disorder" was first proposed in the RCD, possibly stimulated by a proliferation of pain treatment programs in the 1970s and 1980s. This led to the DSM category. DSM-III defined it as "psychogenic pain," DSM-III-R as "somatoform pain disorder," DSM-IV as simply "pain disorder," and DSM-5 merging it into Somatic Symptom Disorder.

Those who subscribe to the pathophysiological relevance of brain neurotransmitters to pain syndromes are likely to focus more on the pharmacological management of pain disorders. For decades, tricyclic antidepressants, particularly the more serotonergic ones (e.g., amitriptyline [Elavil]), have been used in open trials, and the evidence suggests that they are effective for pain syndromes at doses lower than the ones traditionally prescribed for depression (e.g., amitriptyline 75 to 100 mg per day).

There have been several randomized, controlled clinical studies on various pain syndromes. A placebo-controlled study of low-dose imipramine (Tofranil) in patients with unexplained chest pain ("abnormal cardiac pain perception") showed significantly better improvement in pain scores among patients receiving imipramine. In a placebo-controlled study of phenelzine (Nardil), imipramine, and placebo in patients with mixed pain/depression syndromes, phenelzine was found to be more effective than imipramine and placebo. A recent meta-analysis of controlled studies on various functional pain syndromes showed a robust effect for antidepressant medications, particularly tricyclics, on pain and functional symptoms. It has long been noted that pain is often associated with depression. Yet, Smith found only 13 studies from 1975 to 1990 that explored this relationship. Of these studies, only four addressed the treatment of depression and coexisting pain. All were open-trial designs, with and without antidepressants (amitriptyline or doxepin [Sinequan]) and comprehensive "pain programs." All showed impressive rates of response of depression scores when pain was focus of the treatment in these comorbid states. The development of dual antidepressants such as duloxetine has reawakened interest in examining somatic components of depression such as unexplained physical pain. Although this may have begun as a marketing strategy, it has served to reinforce the observation of clinicians of a significant association of depression with somatic syndromes such as pain.

Cognitive-Behavioral Therapy. CBT in noncardiac chest pain was studied in a randomized, controlled trial by a British group. Thirty-seven

subjects agreed to participate in the study after being reassured by a cardiologist that they did not have heart disease. The study found superiority of CBT over the control group in improving functional limitation and decreased tendency to worry about physical symptoms. A meta-analysis of 25 trials of treatment with CBT in various functional syndromes including pain showed that CBT was effective for a variety of chronic pain conditions. Additional studies comparing different therapeutic modalities may help to sharpen their specificity. Although psychodynamic treatments have fallen somewhat into disfavor, the assessment of pain programs that have this orientation can help shed light on the quality of patient–therapist interactions that predict good outcomes. Finally, collaborative approaches between mental health professionals and primary care physicians should contribute to more-integrated treatment approaches that may ease the burden of the therapists and maximize the therapeutic gains.

Psychological Factors Affecting Other Medical Conditions (PFAOMC)

According to DSM-5, the A criterion for this condition is the presence of a medical symptom or condition while the B criterion, the essential feature, is the presence of one or more clinically significant psychological or behavioral factors that adversely affect a medical condition by increasing the risk for suffering, death, or disability. Severity ranges from "mild" to "extreme."

Psychological or behavioral factors include psychological distress, patterns of interpersonal interaction, coping styles, and maladaptive health behaviors such as denial of symptoms or poor adherence to medical recommendations.

Common clinical examples include denial and refusal of treatment for an acute condition (such as myocardial infarct or abdominal emergencies) by individuals with certain personality styles (e.g., domineering or controlling), the exacerbation of asthma or irritable bowel attacks by anxiety, and the manipulation of insulin by an individual with diabetes, or diuretics in the case of hypertensive patients, in efforts to lose weight.

The diagnosis should be only made when the effect of the psychological factor on the medical condition is unambiguous and leads to documentable effects on the course and outcome of the medical condition.

The reverse situation, the psychiatric or psychological consequences of having a medical condition, is more properly classified as adjustment disorder. Other ambiguous situations where psychological and physical symptoms coexist (the so-called "comorbidities") present particular challenges, as it is often the case with somatic, anxiety, and mood disorders that not only overlap with one another, but are often associated with chronic medical conditions (diabetes, hypertension, hypercholesterolemia) or "functional" syndromes (irritable bowel, migraine headaches, and

several others).

The prevalence of PFAOMC is unknown. In mental health systems, the diagnosis is most frequently made in consultation-liaison psychiatry settings.

Differential diagnosis can be complex and should include other somatic symptom disorders (somatic symptom disorder, conversion disorder, factitious disorder), as well as personality disorders (e.g., borderline personality disorder).

Treatment. Treatment frequently involves communication with the patient's primary medical team as well as family. The psychoeducational intervention clarifies the role that emotional and behavioral factors play in aggravating the underlying medical condition. Medication to treat another underlying psychiatric disorder may be necessary.

Mr. A, a 55-year-old man was hospitalized on the intensive care unit following a cardiac arrest. He had been experiencing severe substernal chest pain but "ignored it." He became diaphoretic and reported reluctantly to a nearby emergency room where he had a cardiac arrest while awaiting evaluation. He was successfully resuscitated and transferred to the ICU. He emphatically rejected having had a heart attack, took his EKG leads off, and was preparing to leave the unit against medical advice. The psychiatric consultant obtained a history that the patient's father and brother had died of coronary disease. At the time of initial assessment, the patient was irritable and somewhat anxious. He agreed to stay in the ICU overnight after his wife insisted that he stay and the psychiatrist explained the tests further. He was given a low dose of a benzodiazepine and on the next day was agreeable to remain in the hospital for further treatment.

Factitious Disorder

This category was listed as a separate category in the previous DSM editions.

DSM-5 added it to the somatic symptom and related disorders category because the symptoms these patients present with are most frequently somatic and a differential diagnosis with other somatic symptom disorders such as conversion disorder (functional neurological disorder) is essential.

These patients feign, misrepresent, simulate or cause signs or symptoms or physical illness on themselves. A subtype of factitious disorder is Factitious Disorder imposed on another (previously known as Factitious Disorder by proxy).

DSM-5 distinguishes factitious disorder from malingering. According to DSM-5, malingering is differentiated from factitious disorder by the intentional reporting of symptoms for personal gain such as financial compensation or getting time off work. Thus, a distinguishing feature of factitious disorder in DSM-5 is the falsification of signs or symptoms or the induction of disease such as skin lacerations, adding blood to urine, inducing GI bleeding, laxative-induced diarrhea, induced fever and many

others in the absence of obvious external rewards.

Conversion disorder is an important differential diagnosis in patients presenting with neurological symptoms (fits, convulsions, paralysis, anesthesia, and so on), as unequivocal evidence of voluntary and deceptive fabrication of symptoms may be difficult to obtain.

Historically, simulated physical symptoms have been recorded since the time of the Greeks (Galen's report of induced physical symptoms such as vomiting or rectal bleeding). In recently released Vatican archives, self-induced stigmata, discovered, as part of inquisition trials is an interesting observation. In more recent times, factitious disorder was known as "Munchausen's syndrome" following the 1951 publication by Asher in the *Lancet*.

Epidemiology of FD is not known. Etiology is generally not clear, except in the case of religious instances or those cases involving people familiar with the health field. A common denominator nowadays is that these patients tend to be avid medical service seekers, curiously not a criterion included in defining the disorder.

Treatment is basically conservative, aimed at preventing further complications.

Other Specified Somatic Symptom and Related Disorders

Included here are below threshold somatic syndromes including brief somatic symptom disorder, brief illness anxiety disorder (both with less than 6 months duration of symptoms), illness anxiety disorder without excessive health-related behaviors (criterion D not met) and pseudocyesis.

Unspecified Somatic Symptom and Related Disorders and Other Below-Threshold Somatic Symptom Syndromes

Introduction. Several investigators responding to the expressed need of clinicians and researchers, particularly in primary care, have proposed below-threshold subtypes that may capture many of the patients often seen in primary care who do not meet the full criteria for somatization disorder and fall into the nonspecific categories of undifferentiated somatoform disorder or somatoform disorder not otherwise specified.

Definition. Syndromes that fall below the threshold for a somatic symptom disorder may include the old notions of "somatization syndrome," "abridged somatization disorder," and "polymorphous," "multisomatoform," or "polysomatoform" disorder.

History. Because of the difficulty in applying the unwieldy diagnostic criteria of full DSM-IV somatization somatoform disorder in primary care, based in a clinical study, Escobar proposed an abridged construct of

somatization that used a threshold of four idiopathic physical symptoms for males and six for females. This construct was shown to have similar discriminating power as the full somatization disorder. The abridged construct was systematized in 1987, incorporating the physical symptoms elicited with the DIS. Using these criteria, Kirmayer and Robbins in 1991 outlined three nonoverlapping syndromes in patients seen in a primary care service. They reported that the majority of patients presenting with medically unexplained somatic symptoms fitted one of the three definitions, each implying a different pathogenic process. In 1993, T. K. Craig and collaborators in the South London Somatization Study also used the abridged criteria for somatization and proposed four syndromes in primary care: "psychologizers," "somatizers," and "mixed" and "physical" cases. The notions of "psychologizers" and "somatizers" were also used in Spain by A. Lobo and collaborators. Other authors, such as K. Kroenke in the United States and W. Rief in Germany, also proposed abridged constructs of somatization.

Epidemiology. Overall, it has been estimated that in up to two-thirds of patients presenting to primary care, no allopathic disease entity can be differentially diagnosed and unequivocally established as the principal determinant of the presenting somatic complaints. In the United States, the proportion of patients with full DSM-IV somatization disorder and abridged somatoform disorder in primary care settings has been estimated to be 3 and 20 percent, respectively. Of the 20 percent or so primary care patients with high levels of idiopathic physical symptoms, most of them (75 percent or so) meet criteria for a major psychiatric disorder, lending credence to the statement that idiopathic physical symptoms may be the most typical way for common psychiatric disorders such as depression and anxiety to manifest themselves in primary care. An international study sponsored by the World Health Organization (WHO) that took place in 14 different countries examined in detail somatic symptoms, as well as various concepts of somatization in primary care, and how these related to depression and other syndromes. The study found that idiopathic physical symptoms defined either as somatoform disorder or abridged somatoform disorder were extremely common (range 53 to 65 percent). No major differences were noted across countries when somatization was defined as either denial of psychological symptoms or reporting of medically unexplained somatic symptoms. However, when somatization was defined according to presenting symptoms, the study found greater variation among the centers. An important observation of this study was confirming that reports on lifetime somatic symptoms are unstable, particularly when somatic symptom thresholds are high (somatoform disorder). Lower symptom counts (abridged somatoform disorder) fared better, but even here, half of symptoms reported at the initial interview were forgotten 1

year later. Because, according to this research, diagnoses based on lifetime recall appear to rest on an unreliable foundation, the use of current rather than lifetime somatic symptoms should be the rule. Idiopathic symptoms may be simply a reflection of the availability of particular health services in a given region, as suggested by the WHO collaborative study mentioned above. There is also some evidence that there are fewer "somatizers" in settings in which a more personalized type of care is common than in those settings in which it is rare.

Treatment

GENERAL OUTLINE FOR TREATMENT OF IDIOPATHIC PHYSICAL SYMPTOMS IN PRIMARY CARE.

The primary care setting has been dubbed the "de facto mental health system," and because this is more readily available across cultures than the more formal mental health system, comparative international studies have often taken place in that setting. Studies in primary care are also of particular importance with regard to somatic presentations because physical symptoms presented at this level can be taken to have clinical importance to the person experiencing them. In the 1920s, Francis W. Peabody, Harvard Professor of Medicine at Boston City Hospital, wrote that teachers and students are at risk of "serious error in their attitude toward a large group of patients who do not show objective, organic pathologic conditions, and who are generally spoken of as having nothing the matter with them." Peabody stated that treatment was the responsibility of the internist and general practitioner, and that its effectiveness resided on the quality of a caring patient–doctor relationship. Although some of these "difficult" patients eventually end up in psychiatric clinics, many remain in primary care. Over the years, the care of these patients has been negatively affected by the lack of an integrated health care system. Given the dimensional characteristics of these disorders, having a broad range of symptoms, symptom severity, and functional impairment, the "one size fits all" approach is grossly unrealistic. Therefore, a personalized, multistage approach should be developed. In primary care, patients with idiopathic symptoms often have symptoms of common mental disorders such as depression and anxiety. The initial step should focus on providing symptom relief for the very complaints that prompted patients to seek treatment in the first place. In the case of these patients, these are somatic complaints that frequently express idioms of distress. One would then expect that a positive effect on physical symptoms would make nonpsychologically minded patients more amenable to psychiatric approaches and interventions. Thus the first steps in this process seem crucial.

Step 1: This starts with detailed scrutiny of medically unexplained physical symptoms that will identify patients in primary care who are likely

to be in need of mental health treatment.

Step 2: This entails the use of simple, brief modules, such as psychoeducation, biofeedback, and stress reduction instructional videos. These should target primarily the type and intensity of physical symptoms with which a given patient presents, under the supervision of primary care providers (e.g., nurse practitioners) who have been trained to assess patients and even provide some simple interventions.

Step 3: If symptoms do not improve with the first two steps, the next one entails greater onsite involvement by mental health specialists. It may include expert assessment and referral to a time-limited, CBT-type intervention administered on site in primary care that focuses on the physical symptoms and has been shown to be beneficial for both physical symptoms and, secondarily, symptoms of depression and anxiety.

Step 4: This and subsequent steps include a menu of options for the continuing care of such patients, including psychotherapy and drug treatments either alone or in combination. These may require referral to mental health clinics and treatment coordinated by a team of mental health clinicians.

Pivotal Study. A 2007 study by Escobar and coworkers tested the effectiveness of a CBT-type intervention delivered in primary care for patients presenting with medically unexplained physical symptoms. Patients were randomly assigned to receive either the new intervention in addition to a consultation letter or "usual clinical care" plus a consultation letter alone. All treatments and assessments took place at the same primary care clinic where patients sought care. The results showed that a significantly higher proportion of patients in the intervention group had their physical symptoms rated as "much–very much improved" compared to those in the other group (60 percent vs. 25.8 percent; odds ratio, 4.1; 95 percent confidence interval [CI], 1.9 to 8.8; $p < 0.001$). The effect of the intervention on unexplained physical symptoms was best noticed at treatment completion, led to significant relief of symptoms in more than half of the patients, and persisted for up to 6 months after the intervention, although its effectiveness did gradually diminish. The intervention also led to significant improvement in self-reported physical symptoms, self-ratings of physical symptoms, and clinician-rated depression scales at the posttreatment assessment. The authors of the study concluded that a time-limited CBT-type intervention significantly ameliorated presenting physical complaints of patients seen in primary care and offer an alternative for managing these common and problematic presentations in primary care settings.

GENERAL GUIDELINES FOR TREATMENT OF SOMATIC SYMPTOM DISORDERS

Reviews of controlled studies have shown good efficacy of psychotherapy,

particularly of CBT type, and some evidence for the efficacy of drug therapy in the treatment of a number of somatoform and functional somatic syndromes. This should dispel the therapeutic nihilism that has surrounded the treatment of these patients for quite some time. Moreover, in a recent update (up to 2006), Kroenke found 34 randomized clinical trials on somatoform disorders, which included more than 1,000 patients. Of these, 13 were CBT trials, 5 were antidepressant trials, and 16 involved other modalities. Overall, CBT showed the highest efficacy, with virtually all reported studies showing a significant effect of CBT on symptoms compared to no treatment. Results were particularly favorable in the case of patients with various functional symptom syndromes and high levels of medically unexplained physical symptoms, as well as for patients with hypochondriasis and body dysmorphic disorder. Antidepressants seem to work in painful syndromes, and the data favor older antidepressants (tricyclics) over new ones (SSRIs). Other interventions, including CBT, the use of a psychiatric consultation letter, brief psychodynamic therapy, exercise, and more recently, biofeedback, have also shown beneficial effects on some of these syndromes. A word of caution should be added, however, to temper the enthusiasm on the efficacy of CBT. Even though the pivotal CBT trials described here were randomized clinical trials, in most instances, the comparison group was restricted to a "medical care as usual" group or to the simple addition of the consultation letter. Thus, it cannot be ruled out that the better outcomes in the CBT-treated patients may be at least partly related to the increased attention and the additional time and effort invested in the case of the experimental (CBT) treatment. Therefore, future trials should rigorously control for this factor.

THE FUTURE OF SOMATIC SYMPTOM DISORDERS'TOWARD A NEW DIAGNOSTIC CLASSIFICATION?

Psychological causation and modulation, as well as behavioral and attitudinal aspects, of the somatic symptom syndromes have been neglected in past diagnostic systems, possibly because they are difficult to define operationally. In an influential book, McHugh and Slavney stated that a somatic presentation ("hysteria" in this instance) "is *not something the patient has; it is something the patient does; that is, it is a behavior*." To some extent, this perspective was incorporated in the new criteria delineated in DSM-5. The tendency to always view certain somatic presentations as the physical expression of a primary mental disorder was evident in criteria such as those proposed by Bridges and Goldberg in England, which emphasized the need for a proper psychiatric diagnosis made by a competent clinician to confirm that the somatic presentations are or are not connected to a psychiatric disorder.

Other interesting models to dissect ambiguous somatic symptom syndromes have been proposed. For example, Richard S. Brown separates

somatic presentations according to three explanatory themes, namely "dissociation, conversion, and somatization." Harold Merskey and Francois Mai distinguish two types of somatic symptom presentations. The first comprises symptoms such as unexplained pain and "functional" cardiovascular or gastrointestinal symptoms, which are involuntary, automatic, related to autonomic arousal, and appear in temporal relationship to the experience of "stress" or jointly with depression/anxiety syndromes. This type seems particularly common in primary care. The other and more ambiguous set depends on thoughts, embodied intentions, or ideas that are not recognized or acknowledged, expressed as unexplained neurological symptoms. This subtype may represent the more severe, polymorphous, treatment-refractory syndromes that are more likely to be seen in specialty mental health environments but are also seen in medical and surgical tertiary care centers. These models make sense intuitively, but they have not been assessed in systematic studies. A European study showed that the expression of physical and psychological complaints is ubiquitous, starting at an early age and that they seem to follow separate paths regarding their expression, evolution, and recognition.

NIMH Research Domain Criteria (RDoC)

This alternative diagnostic approach, developed by Bruce Cuthbert and Thomas Insel at NIMH, relies on behavioral neuroscience, brain "circuits," "cells," and "self-report scales." RDoC is a dimensional approach for use in research, and its main goal is to define clinical characteristics that may reflect the biology of the illness with greater precision, hopefully leading to more specific therapeutic targets compared to the broad overlapping categories of the DSM. This approach may be useful for examining a number of somatic symptom dimensions such as those included in Section III of DSM-5.

Instruments to Assess Somatic Symptom Disorders

The Present State Examination (PSE), a phenomenologically rich semi-structured instrument that requires clinical expertise and significant training and includes excellent symptomatic descriptions, may be a useful tool for assessing DSM-5 categories somatic symptom and related disorders in both clinical and epidemiological studies.

There are numerous widely used scales for assessing somatic complaints, but the diagnosis of somatic symptom disorder requires more than just chronic distressing and impairing somatic complaints. Instead, it requires assessments of excessive dysfunctional cognitions (catastrophizing) and behaviors. A number of inventories such as the Whitley Scale or the Pain Catastrophizing Scale may be helpful, but the anchors and cut points need to derived in careful research lest the

diagnosis by overapplied.

References

*Allen LA, Woolfolk RL, Escobar JI, Gara M, Hamer RM. Cognitive behavior therapy for somatization disorder: a randomized controlled trial. *Arch Intern Med*. 2006;166:1512.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

Bagayogo IP, Interian A, Escobar JI. Transcultural aspects of somatic symptoms in the context of depressive disorders. *Adv Psychosom Med*. 2013;33:64–74.

*Barsky AJ, Ahern DK. Cognitive behavior therapy for hypochondriasis: a randomized controlled trial. *JAMA*. 2004;291:1464.

Barsky AJ, Borus JF. Functional somatic syndromes. *Ann Intern Med*. 1999;130(11):910.

Berrios G. *The History of Mental Symptoms*. Cambridge: Cambridge University Press; 1996.

Bleuler E. *Lehrbuch der Psychiatrie* [Textbook of Psychiatry]. Berlin: Springer-Verlag; 1955.

Briquet P. Traité de l'hysterie. Paris: Bailliere; 1859.

Creed FH, Davies I, Jackson J, et al. The epidemiology of multiple somatic symptoms. *J Psychosom Res*. 2012;72(4):311–317.

Creed FH, Guthrie E, Fink P, et al. Is there a better term than medically unexplained symptoms? *J Psychosom Res*. 2010;68:5–8.

Dimsdale JE, Creed F, Escobar J, et al. Somatic symptom disorder: an important change in DSM. *J Psychosom Res*. 2013;75(3):223–228.

Dimsdale JE, Levinson J. Diagnosis of somatic symptom disorder requires clinical judgment. *J Psychosom Res*. 2013;75(6):588.

Dimsdale JE, Levenson J. What's next for somatic symptom disorder? *Am J Psychiatry*. 2013;170:1393–1395.

Escobar JI, Burnam MA, Karno M, Forsythe A, Golding JM. Somatization in the community. *Arch Gen Psychiatry*. 1987;44:713.

Escobar JI, Cook B, Chen CN, et al. Whether medically explained or not, three or more concurrent somatic symptoms predict psychopathology and service use in community populations. *J Psychosom Res*. 2010;69:1–8.

Escobar JI, Gara M, Cohen Silver R, Waitzkin H, Holman A. Somatization disorder in primary care. *Br J Psychiatry*. 1998;173:262.

*Escobar JI, Gara MA, Diaz-Martinez A, Interian A, Warman M. Effectiveness of a time-limited, cognitive behavior therapy–type intervention among primary care patients with medically unexplained symptoms. *Ann Fam Med*. 2007;5:328.

Ey H. *Manuel de Psychiatrie*. Paris: Masson; 1965.

Fallon B, Petkova E, Skritskaya N, Sanchez-Lacay A, Vermes D. A double-masked placebo-controlled study of fluoxetine for hypochondriasis. *J Clin Psychophamacol*. 2008;28:638.

Feinstein RE, deGruy FV. Difficult patients: personality disorders and somatoform complaints. In: Rakel RE, ed. *Textbook of Family Medicine*. 8th ed. Elsevier; 2011: Chapter 46.

Foucault M. Folie et deraison. In: *Histoire de la folie*. Paris: Librarie Plon; 1961.

Janssens KAM, Oldeehinkel AJ, Verhulst FC, Hunfeld JA, Ormel J, Rosmalen JG. Symptom-

specific associations between low cortisol responses and functional somatic symptoms, the TRAILS study. *Psychoneuroendocrinology*. 2012;37(8):332–340.

Jaspers K. *General Psychopathology*. Baltimore, MD: Johns Hopkins University Press; 1997.

Katsamanis M, Lehrer P, Escobar JI, Gara MA, Kotay A, Liu R. Psychophysiologic treatment for patients with medically unexplained symptoms: a randomized controlled trial. *Psychosomatics*. 2011;52(3):218–229.

Kroenke K. Efficacy of treatment for somatoform disorders: a review of randomized controlled trials. *Psychosom Med*. 2007;69:881.

Lipowski ZJ. Somatization: the concept and its clinical applications. *Am J Psychiatry*. 1988;145:1358.

Lipsitt D, Escobar J. Psychotherapy of somatoform disorders. In: Gabbard GO, Beck JS, Holmes J, eds. *Oxford Textbook of Psychotherapy*. Oxford: Oxford University Press; 2005:247.

McHugh PR, Slavney PR. *The Perspectives of Psychiatry*. Baltimore, MD: Johns Hopkins University Press; 1998:223.

Perley MJ, Guze SB. Hysteria: the stability and usefulness of clinical criteria. *N Engl J Med*. 1962;266(9):421.

Schneider K. *Clinical Psychopathology*. New York: Grune & Stratton; 1959.

Sharpe M. Somatic symptoms: beyond "medically unexplained." *Br J Psychiatry*. 2013;203(5):320–321.

Shorter E. *From Paralysis to Fatigue*. New York: Free Press; 1992.

Shorter E. *From the Mind into the Body: The Cultural Origin of Psychosomatic Symptoms*. New York: Free Press; 1994:90–117.

Simms LJ, Prisciandaro JJ, Krueger RF, Goldberg DP. The structure of depression, anxiety and somatic symptoms in primary care. *Psychol Med*. 2012;42(1):15–28.

Smith GR, Monson RA, Ray DC. Psychiatric consultation in somatization disorder. A randomized controlled study. *N Engl J Med*. 1986;314:1407.

Stekel W. *Die Sprache des Traumes*. Wiesbaden, Germany: Bergmann; 1911.

Teo AR, Du YB, Escobar JI. How can we better manage difficult patients encounters? *J Fam Pract*. 2013;62(8):414–421.

Tomenson B, Essau C, Jacobi F, et al. Total somatic symptom score as a predictor of health outcome in somatic symptom disorders. *Br J Psychiatry*. 2013;203(5):373–380.

*Wessely S, Nimnuan C, Sharpe M. Functional somatic syndromes: one or many? *Lancet*. 1999;354:936.

Zimmerman T, Puschmann E, Ebersbach M, Daubmann A, Steinmann S, Scherer M. Effectiveness of a primary care based complex intervention to promote self-management in patients presenting psychiatric symptoms: study protocol of a cluster-randomized controlled trial. *BMC Psychiatry*. 2014;14:2.

19

Factitious Disorder

DAVIN K. QUINN, M.D., DORA WANG, M.D., SETH POWSNER, M.D., AND STUART J. EISENDRATH, M.D.

INTRODUCTION

Patients with factitious disorder simulate, induce, or aggravate illness to receive medical attention, regardless of whether they are ill or not. These patients may inflict painful, deforming, or even life-threatening injury upon themselves, their children, or other dependents. The primary motivation is to seek attention by playing the role of a patient. Factitious disorder patients, in contrast to malingerers, lack an obvious external reward, such as avoidance of duties, or financial gain.

Factitious means "artificial, false," from the Latin *facticius,* "made by art." The art and artifice of patients with factitious disorder often leads to high hospital drama. Clinicians may exclaim in frustration, "He's not really sick! He's doing it to himself!" and dismiss or refuse to treat patients with factitious disorder. Countertransference can be a major obstacle toward the proper care of patients who are potentially dangerous to themselves or others, and who are likely suffering from a psychiatric disturbance.

Factitious disorders can lead to morbidity or even mortality. Therefore, the medical and psychiatric needs of these patients must be taken seriously, even though presenting complaints are by definition falsified. As an example, an operating room technician, the daughter of a physician, repetitively injected herself with *Pseudomonas,* which caused multiple bouts of sepsis, bilateral renal failure, and ultimately her death. Such deaths are not uncommon in severe factitious disorders.

In factitious disorder imposed on another, also known as factitious disorder by proxy, Munchausen syndrome by proxy, and "fabricated illness in a child," the patient seeks medical attention for another individual, most commonly his or her own preverbal infant. Factitious disorder imposed on another is recognized in all 50 states as child abuse, elder abuse, or abuse of a vulnerable dependent adult. Clinicians must prioritize the safety of the victim, which may involve notifying government agencies responsible for the protection of these groups.

Factitious disorders imposed on self and imposed on another should not be considered diagnoses of exclusion but should be actively pursued, given the potentially high stakes, especially when dependent abuse is

suspected. Clinicians should be mindful to minimize unnecessary invasive medical procedures and to try to steer patients toward the psychiatric care they may need in face-saving, nonthreatening ways. Indeed, factitious illness behavior often represents an underlying psychiatric disturbance, such as a personality disorder, mood disorder, or psychosis. Ironically, even patients presenting with factitious psychological complaints, such as feigned bereavement, usually suffer from another psychiatric condition for which they are not seeking help.

Relatively little empirical knowledge is available about the etiology, epidemiology, course and prognosis, and effective treatment of factitious disorders. Most knowledge comes from case reports, and this information is suspect given that these patients by definition falsify their histories. Methodological problems are inherent in studying deceptive patients. Systematic studies of factitious disorder are few, and federally funded investigation is nonexistent.

It is likely that every clinician will cross paths with a patient with factitious disorder at some point in his or her career. A survey of the literature on factitious disorders may engender conclusions of therapeutic nihilism, given the few descriptions of successful treatment, and the many examples of patients with recurrent factitious behaviors over decades. However, Munchausen syndrome, the prototypical factitious disorder and the first to gain wide medical interest, is now understood to be a chronic severe variant which comprises only a small portion of all cases of factitious disorder. Clinicians may do well to remember that deception is a pervasive, normal, and ubiquitous behavior in human interactions. In medical settings patients may engage in a wide range of deceptive behaviors for a host of reasons, including as a coping strategy under conditions of stress. Conceptualizing deception and abnormal health care seeking on a spectrum of normal to abnormal behaviors will help clinicians understand that not all factitious illness behavior is as refractory or as chronic as that demonstrated by patients with Munchausen syndrome.

Definition

The main clinical feature of factitious disorder is the falsification of physical or psychological signs or symptoms, or the induction, or exaggeration, of injury or disease. It must be associated with identified deception and be evident even in the absence of obvious external rewards. An individual may present oneself as ill, impaired, or injured, in which case the diagnosis is *factitious disorder imposed on self*. If an individual presents another individual as ill, the diagnosis is *factitious disorder imposed on another*. The definitions according to the text revision of the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) are given in Table 19–1. Factitious disorders are now considered part of the Somatic Symptom and Related Disorders category, given that

patients with both disorders manifest predominantly somatic symptoms and both are most often encountered in the medical setting. There are also many areas of overlap between these disorders in comorbidities and treatment.

Munchausen's syndrome, the eponymous term coined by Richard Asher in his landmark 1951 publication, is known as *chronic factitious disorder with predominantly physical signs and symptoms* in the lexicon of DSM-IV. Some practitioners use the term Munchausen syndrome for a chronic, severe, refractory form of factitious disorder, in which deceptive illness behavior becomes a lifestyle, precluding stable relationships or employment. Munchausen syndrome comprises approximately 10 percent of all cases of factitious disorder. Constantly seeking medical care and hospitalization, such patients often assume grandiose false identities, sometimes claiming to be royalty, relatives of celebrities, or figures in important historical events. They travel from hospital to hospital, and when they become well known in one city, take their deception on the road to begin the behavior anew elsewhere. Previous terms applied to these patients included *hospital hoboes, hospital addicts,* and *professional patients*. Two distinguishing features of Munchausen syndrome beyond the simulation of disease are *pseudologia fantastica*—the telling of vague, self-aggrandizing, heroic tales often containing a kernel of truth—and *peregrination*—the tendency to travel widely. There is evidence to support a second durable phenotype of factitious disorder imposed on self, called *common factitious disorder,* in which factitious behavior is confined to one locality and a relatively circumscribed set of complaints. The prototypical patient tends to be young, female, socially connected, employed, and working in health care.

Table 19–1.
Factitious Disorder Imposed on Self

A. Falsification of physical or psychological signs or symptoms, or induction of injury or disease, associated with identified deception.
B. The individual presents himself or herself to others as ill, impaired, or injured.
C. The deceptive behavior is evident even in the absence of obvious external rewards.
D. The behavior is not better explained by another mental disorder, such as delusional disorder or another psychotic disorder.

Specify: single episode
Recurrent episode (two or more events of falsification of illness and/or induction of injury)
Factitious Disorder Imposed on Another (Previously Factitious Disorder by Proxy)

A. Falsification of physical or psychological signs or symptoms, or induction of injury or disease, in another, associated with identified deception.
B. The individual presents another individual (victim) to others as ill, impaired, or injured.
C. The deceptive behavior is present even in the absence of obvious external rewards.
D. The behavior is not better explained by another mental disorder, such as delusional disorder or another psychotic disorder

Note: the perpetrator, not the victim, receives this diagnosis.

Specify: single episode
Recurrent episode (two or more events of falsification of illness and/or induction of injury)

In factitious disorder imposed on another, a person intentionally simulates illness in another individual, often someone who is dependent on the perpetrator for care. Most commonly the perpetrator is a mother feigning or producing illness in her own preverbal infant. Factitious disorder imposed on another may also be committed against an adult, such as an elder or spouse. Epidemics of hospital deaths have been attributed to medical personnel inducing or exacerbating illnesses in patients. Because factitious disorder imposed on another almost always constitutes child abuse or criminal abuse, the forensic terms "perpetrator" and "victim" are used even in the medical literature. Subsequent to the DSM-5, factitious disorder imposed on another replaces the terms factitious disorder by proxy and Munchausen syndrome by proxy in psychiatric nomenclature. There is ongoing debate in the pediatric literature about the appropriate terminology for the condition. *Fabricated illness in a child* is the preferred term of the American Academy of Pediatrics Committee on Child Abuse and Neglect, as it focuses attention on the harm caused to a child and less on the motivations of the perpetrator, consistent with the approach to diagnosing other forms of child maltreatment.

History

In the second century of the Common Era, the Greek-born physician Galen wrote of patients inducing or simulating symptoms such as vomiting or rectal bleeding. The Bible relates accounts of people self-inflicting injury. In the Middle Ages, "hysterics" reportedly put leeches in their mouths to simulate hemoptysis and abraded their skin to reproduce skin conditions. However, judging historical accounts through contemporary mindsets is often problematic, and it is difficult to say whether these accounts actually represent factitious disorder.

In 1838, the Scottish military physician Hector Gavin published an essay, "On the Feigned and Factitious Diseases of Soldiers and Seamen, on the Means Used to Simulate or Produce Them, and on the Best Modes of Discovering Impostors." Although most of Gavin's subjects were malingering to escape duty in the high-casualty Napoleonic Wars, he also noted that the motive of some was simply "to excite compassion or interest" and that "some soldiers, indeed, without any ulterior object, seem to experience an unaccountable gratification in deceiving their officers, comrades, and surgeon."

Jean-Marie Charcot, around 1890, used the term *mania operativa activa* to describe a young girl who continually sought surgery for pain in a knee joint, until her medical care-seeking resulted in a surgeon amputating

the leg. No pathology was ever found in the leg. In 1901, the Swiss physician Henri Secretan lent his name to a peculiar syndrome of nonhealing, traumatically induced edema of the dorsum of the hand. G. Reading, in 1980, confirmed that Secretan syndrome is factitiously produced. In 1934, Karl Menninger described "polysurgical addiction."

Interest in factitious disorder increased markedly when the term Munchausen syndrome was coined by Asher in 1951 to describe three patients with false abdominal complaints, all of whom used multiple identities and sought care at multiple hospitals. In prefacing the case reports he wrote:

> Here is described a common syndrome which most doctors have seen, but about which little has been written. Like the famous Baron von Munchausen, the persons affected have always travelled widely; and their stories, like those attributed to him, are both dramatic and untruthful. Accordingly, the syndrome is respectfully dedicated to the baron, and named after him.

The Baron Karl Friedrich Hieronymus von Munchausen (1720–1797) was a nobleman who served in the Russian army in the war against Turkey (Fig. 19–1). After retirement, he entertained friends with embellished stories of his war adventures. He was turned into a literary figure by the author Rudolph Eric Raspe, a friend of the baron's, who fled Germany for England when caught embezzling from a museum. To pay off debts, Raspe published an account of the baron's tales in 1785.

In 1968, Spiro noted that of the 38 cases of Munchausen syndrome then published, none involved a detailed psychiatric workup and less than half were evaluated by a psychiatrist. He advocated greater understanding of these patients and presented the first detailed psychiatric case study, confirming information with collateral sources. He criticized the "Munchausen" label and recommended the less pejorative term *chronic factitious symptomatology*.

Factitious disorder with psychological symptoms was first described by Alan Gelenberg in 1977, who mused that whereas other factitious disorder patients avoided psychiatrists, his patient, a war veteran, gained admission to more than 30 psychiatric hospitals within a few years, usually feigning depression and suicidality under various pseudonyms.

"Munchausen syndrome by proxy" was first described in 1977 by British pediatrician Roy Meadow. He described one mother who caused salt poisoning in her child and another mother whose repeated fabrications of urinary tract infections in her daughter resulted in 12 hospitalizations, more than 150 urine cultures, 6 examinations under anesthesia, 5 cystoscopies, and 7 major x-ray procedures—all by the age of 6 years. When the girl was admitted for observation, the diagnosis of Munchausen syndrome by proxy was made based on urine samples that were bloody whenever collected by the mother, but normal when collected by the nurse. Of note, the mother had also sought medical treatment for factitiously induced urinary tract infections in herself.

FIGURE 19–1. The Baron Karl Friedrich Hieronymus von Munchausen (1720–1797). **Left:** The Baron wears military armor in this 1750 portrait by G. Bruckner. A nobleman who served the Russian army in the war against the Turks, the baron entertained friends with embellished stories of his war adventures in his retirement. His tales gained fame when published by Rudolph E. Raspe. **Right:** The baron appears as a caricature in this drawing by 19th century artist Gustave Doré. Like the baron, patients with factitious disorders are real persons deserving of respect, even though they often present themselves as caricatures. (Portrait courtesy of Bernhard Wiebel, http://www.Muenchhausen.ch. The actual portrait was lost in World War II. Caricature from Gustave Doré. The Adventures of Baron Munchausen, One Hundred and Sixty Illustrations by Gustave Doré. New York: Pantheon Books; 1944.)

Comparative Nosology

The DSM-III, in 1980, was the first edition of the DSM to recognize factitious disorder. It focused on Munchausen syndrome, calling it the "prototype" of all factitious disorders. Factitious disorder with psychological symptoms was also recognized, as well as a category called atypical factitious disorder that included what later was termed factitious disorder with predominantly physical signs and symptoms.

Subsequent editions of the DSM increasingly recognized the rarity of Munchausen syndrome and placed more emphasis on a greater spectrum of factitious disorders. The DSM-III-R recognized factitious disorder with physical symptoms and factitious disorder with psychological symptoms. The emphasis on Munchausen syndrome was lessened. Factitious disorder with combined physical and psychological symptoms made its appearance under the category of factitious disorder not otherwise specified (NOS).

In contrast, the DSM-IV defined a single category, factitious disorder, with three types: (1) with predominantly psychological signs and

symptoms, (2) with predominantly physical signs and symptoms, and (3) with combined psychological and physical signs and symptoms. Factitious disorder NOS was exemplified by factitious disorder by proxy, a disorder cited as a category deserving of more research, with research criteria listed. The DSM-IV-TR criteria for factitious disorders were unchanged from the DSM-IV criteria.

DSM-5 in 2014 introduced several changes to the diagnostic category. First, it formally recognized factitious disorder by proxy by dividing the general category of factitious disorders into two groups: factitious disorder imposed on self, and factitious disorder imposed on another. The diagnosis of factitious disorder NOS was eliminated. Second, there is no longer a distinction between physical or psychological presentations. Factitious disorders are now specified as being a single episode or recurrent episodes. Third, the criteria no longer require that illness induction be conscious or intentional, recognizing the difficulty inherent in making such a determination. Instead, evidence of deception in the absence of clear material gain is sufficient to establish factitiousness.

The tenth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) lists factitious disorder (F68.1) under the category "other disorders of adult personality and behaviour," emphasizing the strong association of personality disorders and factitious disorders. The diagnostic label is "intentional production or feigning of symptoms or disabilities, either physical or psychological (factitious disorder)." Similar to DSM, the ICD-10 definition emphasizes the lack of obvious external motivation. However, ICD-10 goes further than DSM by stating that "the motivation is obscure and presumably internal with the aim of adopting the sick role." There are four subtypes: (1) unspecified; (2) with predominantly psychological signs and symptoms; (3) with predominantly physical signs and symptoms; and (4) with combined psychological and physical signs and symptoms. Related applicable terms include hospital hopper syndrome, Munchausen syndrome, and peregrinating patient. The ICD-10 excludes "factitial dermatitis" (neurotic skin picking) and "person feigning illness (with obvious motivation)." The ICD-10 makes no mention of factitious disorder by proxy.

Of noteworthy mention in the ICD-10 lexicon is the related category of "elaboration of physical symptoms for psychological reasons," (F68.0) otherwise known as compensation neurosis. This is grouped into the same category of "other disorders of adult personality and behavior" with factitious disorder. Compensation neurosis is the exaggeration of physical symptoms that are due to a confirmed physical disorder, caused by the psychological state of the individual in which distress and anxiety about prolonged or progressive disability feature prominently. It occurs often in the context of legal or bureaucratic procedures to ascertain the extent of disability or liability. It is thought to have both internal and external

motivations, and lies intermediate between somatoform disorders and factitious disorders on the spectrum of intentionality.

The overlapping and poorly defined boundaries of malingering, factitious disorder, and compensation neurosis highlight the limitations in systems of classification that specify symptoms as intentionally or consciously produced. In reality, intent, motive, and awareness are difficult to discern for a clinician with only brief or intermittent interactions with a patient, and each of these occurs on a spectrum. Although the DSM-5 and the ICD-10 require exclusion of malingering to diagnose a factitious disorder, these two conditions easily and frequently coexist: for example, patients who habitually gratify themselves in the sick role may discover that they can also receive disability payments concurrently; or patients who initially fabricated illness to obtain narcotics for intoxication may continue to deceive even when such medications are no longer forthcoming.

Similarly, factitious disorders can coexist, particularly over time, with unconsciously produced illness behaviors such as conversion disorder, somatic symptom disorder, and dissociative disorders. This longitudinal shifting in the production and motivation of somatic symptoms and the difficulty in determining the exact contributors to any one behavior has led to the categorization of factitious disorders as a type of somatic symptom disorder in DSM-5. Krahn et al. in 2008 opined that patients with factitious disorder and patients with DSM-IV somatoform disorders shared many illness dimensions in common, including: (1) the presence of somatic symptoms; (2) difficulty coping with these symptoms; (3) illness as a way of life; (4) maladaptive use of medical care; and (5) refractoriness to palliative and symptom treatment. Seen from the perspective of somatic symptoms as an expression of emotional distress, factitious disorder patients and somatoform disorder patients are both “engagement seeking.”

Criticism of factitious disorder nosology has come from several different perspectives. One group has suggested that in order for factitious disorder to be diagnosed, any autobiographical falsification and self-harm may suffice, excluding requirements of consciousness or intentionality. A second school of thought stated that various types and severities of deception are a commonplace part of human interactions in general and medical encounters specifically, and that an act of deception should not be considered a basis for a diagnosis of a disease. In this school, Bass and colleagues argued that a person’s psychological and social milieu is a greater contributor to most factitious presentations than any specific psychiatric diagnosis. Third, data from neuroimaging studies on deception indicate that there are significant differences in brain activity between individuals with conversion disorders and individuals instructed to feign symptoms. This argues against the similarities that led to the grouping of factitious disorder with the other somatic symptom disorders.

Despite these theoretical controversies, factitious disorder imposed on

another, still known as Munchausen syndrome by proxy in the legal literature, is well recognized in the legal system. All 50 states require the reporting of factitious disorder imposed on a child to protective services as a form of child abuse. Adults who by virtue of a medical or psychiatric condition are mentally or physically disabled, or reliant on another individual for assistance with activities of daily living, are also at risk for becoming victims of abuse via factitious disorder imposed on another. With increasing recognition in the lay press and medical literature of adult dependents as potential victims of an individual with factitious disorder imposed on another, clinicians should be aware that required reporting also extends to elderly patients and to the developmentally delayed.

Neither the DSM-5 nor the ICD-10 accommodates the literature on variants of factitious disorder imposed on another, in which individuals impose their psychiatric disorders or emotional needs upon their children or dependents. These phenomena can range from the mundane (*achievement by proxy*) to the mildly disruptive (*hypochondriasis by proxy* resulting in repetitive pediatric visits and unnecessary tests without fabrication or induction of illness) to the physically harmful (*anorexia nervosa by proxy* and *malingering by proxy*).

Epidemiology

Factitious Disorder Imposed on Self. The prevalence of factitious disorder is unknown, in part because of the role of deception in this population, but is estimated to be approximately 1 percent of the health care–seeking population. Estimates of prevalence are invariably skewed by methodological limitations of the studies, such as the type of data studied (physician surveys, demographic databases), the study population and setting (hospital ward, outpatient subspecialist, disability office), small sample sizes, and recall bias. Studies of specialists and disease-specific tests appear to overestimate prevalence, likely because factitious disorder patients are more likely to see subspecialists and undergo extensive testing, whereas large database studies are likely to underestimate prevalence due to lack of clinician awareness or discomfort in making the diagnosis. There is a large population of people who never present to medical care but engage in factitious disease-simulating behavior, such as those who play the sick role with family, friends, coworkers, social service agencies, Internet communities, or veterinarians and thus would not be included in medical epidemiological estimates. Online support groups and the Internet create ripe conditions for factitious disorder patients to garner attention, owing to the ease of creating false identities, lack of face-to-face interaction, and the ability to broadcast to large sympathetic audiences.

In Sutherland and Rodin's 1990 study of patients referred to a psychiatric consultation service in a general hospital, 10 out of 1,288 patients met criteria for factitious disorder, resulting in a prevalence of 0.8

percent. A review by Kocalevent et al. in 2005 of 18 studies reporting prevalence of factitious disorder found estimates varied from a minimum of 0.032 percent to a maximum of 9.26 percent, with a weighted mean of 0.9 percent. In a 2012 occupational medicine clinic sample of 400 patients tested for maximum grip strength, 32 were found to be deceptive in their performance, but only 3 were diagnosed with factitious disorder (0.75 percent). Fliege et al. in 2007 surveyed 109 senior physicians from internal medicine, surgery, neurology, and dermatology for the estimated prevalence of ICD-10 factitious disorders in their yearly practice. The mean estimated prevalence was 1.3 percent, the modal prevalence was 1 percent, and individual estimates ranged from 0.01 to 15 percent. In 2010, Swanson et al. surveyed 213 physicians regarding their estimates of rates of patients with medically unexplained symptoms. Out of all 4,159 patients seen in 1 day by these physicians, only 1.5 percent of patients were diagnosed with either a somatoform or factitious disorder diagnosis, with the majority likely somatoform disorders.

Psychiatric settings are somewhat less likely to engage factitious disorder patients: Bhugra estimated psychiatric Munchausen syndrome at 0.5 percent of adult psychiatric admissions younger than the age of 65 years, based on four diagnoses of Munchausen syndrome out of 775 admissions. In a study of 4,500 visits to a psychiatric emergency room, only 7, or 0.15 percent, represented factitious disorder. Similarly, a 1994 study estimated the prevalence at 0.14 percent in a community sample. In a reflection of the overlap between somatic symptoms and factitiousness, Fliege prospectively followed 194 patients admitted to a psychosomatic unit and found 9.7 percent (19) manifested factitious illness behavior.

Database studies yield significantly lower rates of factitious disorder. In 2013, the DSM-5 working group on factitious and somatoform disorders examined data from the National Hospital Discharge Survey to determine the assignment rate of various psychiatric disorders. Out of more than 3.8 million discharges tracked, a factitious disorder diagnosis was assigned in approximately 9 cases per 100,000, far below the expected prevalence estimate. It was theorized that making a factitious disorder diagnosis was likely unpalatable to physicians, and could negatively affect insurance reimbursements. However, all psychiatric disorders including depression and generalized anxiety were assigned at rates below their expected prevalence, indicating that the factitious disorder diagnosis is not necessarily less acceptable than any other psychiatric diagnosis. In a study of 28 million military health care visits from 2005 to 2011, Lande and Williams found that only a miniscule number (89 patients) were diagnosed with factitious disorder. However, it is worth noting that in the military, malingering and self-harm are prosecutable offenses.

Profiles of patients with factitious disorder are few. Krahn et al. retrospectively reviewed 93 cases of factitious disorder diagnosed over a

span of 21 years in a general hospital. Seventy-two percent were female, with a mean age of 31. Two-thirds of the female factitious disorder patients possessed health care training or jobs. Follow-up and treatment engagement were sparse in this group, and only 17 percent ever acknowledged simulating illness. Reich et al. found a similar profile in their series of 41 cases over 10 years: predominantly female, unmarried, with an average age of 33 years, working in medical jobs, and manifesting emotional and sexual inhibition. This particular patient profile has been called "common factitious disorder" or "nonperegrinating factitious disorder" by some groups. In contrast, Munchausen syndrome features peregrination or frequent wide-ranging travel to seek medical care, and *pseudologia fantastica*. In the same series described above, the male patients were predominantly middle aged, unemployed, and unmarried, with antisocial traits or histories, and 61 percent engaged in peregrination.

Factitious Disorder Imposed on Another. Factitious disorder imposed on another is most commonly perpetrated by mothers against infants or young children. The estimated incidence is approximately 0.5 to 2.0 cases per 100,000 children under age 16. It accounts for less than 0.04 percent, or 1,000 of 3 million cases, of child abuse reported in the United States each year, although it may well be underrecognized. Of infants brought to an Australian clinic for apparent life-threatening episodes, an estimated 1.5 percent represented factitious disorder imposed on another. A clinical practice at Great Ormond Street Hospital in England reported that in 20 years, 43 children from 37 families were diagnosed as having induced illnesses.

Rosenberg's review of 117 cases of factitious disorder imposed on another revealed valuable information about the typical perpetrators and victims. Male and female children were equally victimized. The mean age of a child at diagnosis was 3.3 years, with the onset of symptoms occurring at a mean of 1.24 years earlier. All perpetrators were mothers, with 98 percent being biological mothers and 2 percent adoptive. Paternal collusion was suspected in only 1.5 percent of the cases. In half of the cases, illness was actively produced in the child, whereas in 25 percent, illness was simulated without direct infliction on the child. In 25 percent of the cases, illness was both simulated and produced.

In the same review, 10 percent of perpetrators were thought to have Munchausen syndrome themselves, and an additional 14 percent showed features of the syndrome. Data collected from the Great Ormand Street Hospital clinic over a 20-year period indicated that about one-third of perpetrating mothers also had a history of factitious disorder imposed on self. Nearly half of them described serious marital problems. About half of the perpetrating caregivers had a history of psychiatric symptoms, and one-third gave histories of emotional neglect or physical abuse. Fathers were

generally absent or peripheral.

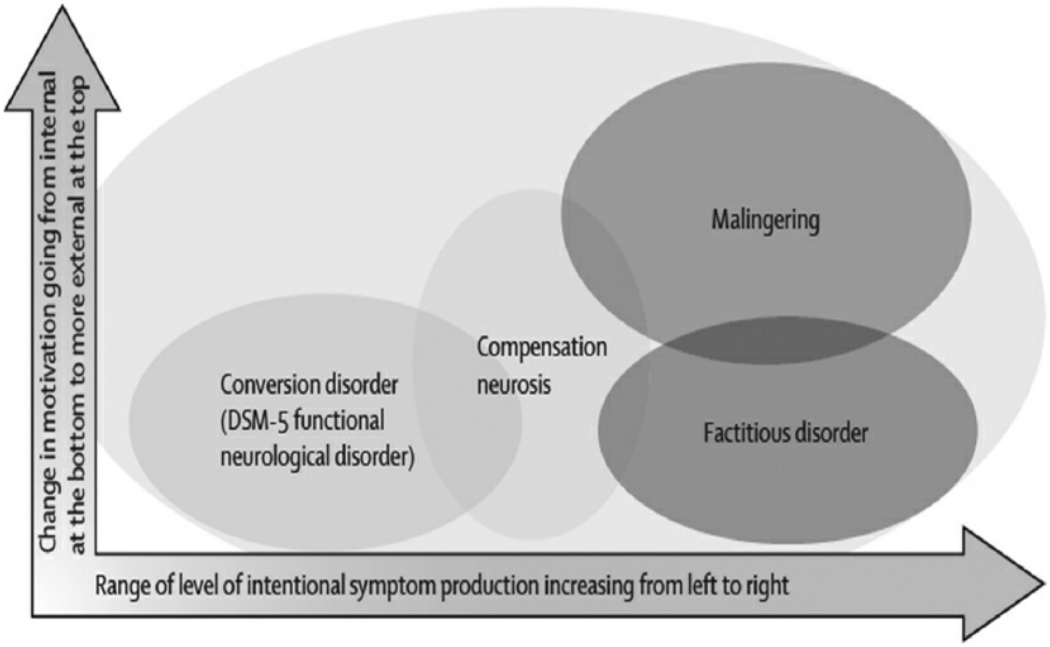

FIGURE 19–2. Intentionality and locus of motivation in disorders of symptom production. (From Bass C, Halligan P. Factitious disorders and malingering: challenges for clinical assessment and management. *Lancet.* 2014;383:1422–1432.)

Bass and Jones examined 28 individuals referred for the treatment for factitious disorder imposed on another. Fifty-seven percent had evidence of a DSM-IV somatoform disorder, and 64 percent had a history of a factitious disorder imposed on self. Thirty-two percent manifested pseudoseizures, and 61 percent engaged in pseudologia fantastica. Early abuse experiences were seen to lead to pathological lying, and finally to fabrication of illness in the victim.

Etiology

Factitious Disorder Imposed on Self. The etiology of factitious disorder is unknown, and a variety of causes likely explain the wide spectrum of factitious illness behavior. The underlying motivations for the behaviors are believed to be unconscious. Two factors underlie most cases of factitious disorder: (1) an affinity for the medical system and (2) poor, maladaptive coping skills.

The various disorders of physical symptoms (somatic symptom disorder, compensation neurosis, factitious disorder, and malingering) can be conceptualized as occupying areas on a two-dimensional graph, in which increasing conscious production of symptoms is on the x-axis, and increasing externality of motivation is on the y-axis. See Figure 19–2.

A majority of factitious disorder patients have medical training, such as in nursing. In a review of 6 case series comprising a total of 165 patients,

Reich and Gottfried found that 60 percent worked in a medical profession. The subjects seemed motivated toward their career choices by genuine lifelong preoccupations with health. Indeed, many with factitious disorder see health care providers as allies, not adversaries. In this case series, many patients had personal ties with their physicians, having worked in their offices or having babysat for them.

Coping deficits are widely noted. Many factitious disorder patients seem to come from large families or to have been neglected as children. They may have grown up without consistent nurturing conducive to the development of mature coping. The poor coping skills of these patients are often symptomatic of an Axis II disorder, such as borderline, narcissistic, dependent, or antisocial personality disorder.

Factitious disorder may be the presenting symptom of an Axis I disorder. Many case reports indicate a lessening or alleviation of factitious illness behavior when major depression is treated. Other case reports point to hypochondriasis as an underlying factor. For example, a 27-year-old physician simulated insulinoma by injecting himself with insulin. When insulin and a syringe were found in a toilet tank, he confessed that he was preoccupied with the thought that he had pancreatic cancer and was trying to provoke further investigation. A 15-year-old boy who was convinced he had a lesion in his urogenital system simulated hematuria to encourage investigation. Substance abuse, psychotic disorders, and mental retardation have also been implicated in factitious illness behavior.

No genetic or familial inheritance pattern has been shown. Intergenerational transmission of susceptibility to factitious disorder is possible, as cases describe families in which a child of a parent with factitious disorder goes on to develop factitious disorder as an adult.

While intact neurological and cognitive functioning is observed in most factitious disorder patients, a neurobiological basis of some factitious behaviors is suggested by reports of abnormal brain scans and deficits in neuropsychological testing in several studies. Pankratz and Lezak in 1988 reviewed 25 cases of factitious disorder in which neuropsychological testing was obtained. Approximately one-third of those patients manifested significant neuropsychological deficits. Lawrie et al. in 1993 described a case of a 34-year-old woman who developed personality change and a full-blown Munchausen syndrome after a cholecystectomy. Electroencephalogram demonstrated focal slowing in the right temporal lobe, and testing revealed attentional and memory deficits with a significant visuospatial constructional apraxia. Single-photon emission computed tomography (SPECT) demonstrated reduced uptake in bilateral frontal and left temporal cortices. Mountz et al. in 1995 reported the results of a SPECT scan in a woman with severe Munchausen syndrome. Regional cerebral blood flow in the right hemithalamus was markedly increased, in line with a theory of right hemisphere abnormalities in factitious disorder

put forth by Pankratz. There is one case report of functional magnetic resonance imaging (fMRI) performed by Spence et al. in a woman convicted of child abuse as a result of factitious disorder imposed on another. She was scanned both while endorsing her version of events in which she is innocent, and while she endorsed the account of her accusers. Longer response times and greater activation of ventrolateral prefrontal areas were observed on repeated scans when she endorsed her accusers' version of events, indicating a change in brain state depending on "truthful" versus "false" endorsement. While this of course cannot be considered a biomarker for innocence or guilt, this finding confirms what has been found in studies using neuroimaging to study simulation. Different and more complex patterns of activation were observed by Stone et al. in bilateral frontal and subcortical areas when patients with conversion weakness attempt to move the paretic limb versus when healthy controls are asked to simulate a weak limb. This line of research suggests that while clinical assessment and treatment of somatic symptom disorders and factitious disorders are converging, deception may involve different brain functions than somatic symptom disorders.

Psychodynamic theories have focused on the concepts of mastery, masochism, and mothering. Striving for mastery may especially apply to factitious disorder patients with predominantly psychological signs and symptoms. For example, patients with factitious psychosis often progress to develop a genuine psychotic disorder, which suggests that the feigning of psychosis may represent a defense or a way of feeling in control of initial psychotic symptoms. Because many patients with factitious physical symptoms seem to have suffered traumatic illnesses as children, adult factitious illness behavior may represent an attempt to master and to feel in control of situations in ways in which they never did as children. They demand or refuse procedures and leave the hospital against medical advice when they feel they are losing control.

Masochism may be a reason that patients repetitively endure painful or deforming surgeries and procedures, such as amputations of limbs and fingers, or exploratory abdominal surgeries that result in extensive scarring. The patient relives childhood physical or emotional abuse at the hands of the medical staff in a repetition compulsion. The physician and the medical system at large become symbolic parents against whom the patient re-enacts dependency, idealization, and anger. The system responds with caring, but also with ostensible physical (e.g., surgery) and emotional (e.g., caregiver anger) abuse and, too often, ultimately with rejection and abandonment. Indeed, for those with chronic factitious disorder, the medical system becomes the main object relation.

Factitious illness behavior often occurs when a patient experiences a loss such as the death of a relative or an occupational loss. Securing the attention of medical clinicians, family, and friends may be a way of

obtaining emotional solace without directly confronting the loss. Dependency and narcissistic needs are fulfilled. Factitious illness behavior may also serve the function of providing some patients a role, identity, or purpose. In the role of a patient, they are masterful orchestrators of medical drama. In practicing *pseudologica fantastica,* they often construct desirable and interesting identities.

Behavioral theories postulate that early in life, these patients received positive reinforcement when sick and perhaps gained nurturing from the medical community that they did not receive at home. Perhaps they learned to see the medical system as a source of caring and emotional support. Alternatively, many of these patients lived in large, neglectful households and became the center of focus only when ill.

Factitious Disorder Imposed on Another. Psychodynamic theories predominately explain factitious disorder imposed on another as objectification of the child to serve the parent's psychological needs. Schreier and Libow called factitious disorder imposed on another a *perversion of mothering* in which a child is dehumanized by the mother and instead serves as a fetishized object through which the mother's dependency needs are met. Eminson and Postlethwaite argued that two axes are at play: The desire to consult, and the inability of the parent to distinguish parental needs from the child's needs.

Rosenberg noted a *disorder of empathy* among perpetrating mothers, along with pervasive themes of loneliness and isolation, often under circumstances of uninvolved or absent husbands. The mother seeks a relationship with the physician, who substitutes for the uninvolved husband. The relationship with the physician is often ambivalent, characterized by idealization and belittling. The more severe the child's alleged illness, the more the mother fulfills her own need for caretaking, both vicariously through the child and more directly in her relationship with the physician.

A common view is that those who impose factitious disorder on another have suffered a loss, such as spousal abandonment or loss of contact with another child. Whether perpetrating mothers tend to have suffered abuse is controversial because many such reports of abuse have subsequently been disqualified by collateral sources. Four factors that suggest a caregiver at risk of fabricating illness in a child include: (1) appearing to need or thrive on attention from physicians; (2) insisting the child cannot cope with the caregiver's ongoing attention; (3) being directly involved in a medical profession or having expertise or familiarity with medical knowledge and terminology; and (4) having a history of factitious or somatic symptom disorder. It should be noted that these characteristics overlap with those of caregivers who are advocates for children with genuine illnesses, and that a caregiver with factitious disorder imposed on another may manifest none

of these traits.

Diagnosis and Clinical Features

Factitious Disorder Imposed on Self. The diagnosis of factitious disorder should be actively pursued and not considered a diagnosis of exclusion. This is recommended for the following reasons: (1) Satisfactory exclusion of all other possibilities may be difficult to accomplish or impossible; (2) early diagnosis can minimize the patient's self-harm, as well as harm through iatrogenic complications; (3) confirmation of deceptive behavior requires positive proof that is usually not forthcoming without proactive investigation; and (4) failure to consider the diagnosis is the most common reason for missing a diagnosis of factitious disorder.

Table 19–2lists reported presentations of factitious disorders along with means used to simulate illness, features suggestive of fabrication, and possible means of detection. The list is a testament to the creativity and resourcefulness of these patients, as well as to their willingness to undergo pain, injury, and inconvenience. Diseases of every organ system have been simulated, limited only by the patient's creativity, knowledge, and access to technological means. Often, the factitious patient's knowledge of a disease exceeds that of the evaluating physician, given the tendency of the patient to be medically trained, as well as the tendency to present to emergency rooms or clinics during evenings or weekends when less experienced resident physicians and staff are on duty.

Factitious symptoms can be (1) *fabricated,* for example, by giving a false history of cancer, acquired immune deficiency syndrome (AIDS), or another illness; (2) *feigned,* for example, by faking symptoms such as pain or seizures; (3) *induced,* by actively producing symptoms through self-infliction of injury or through injection or ingestion; or (4) *aggravated,* such as by manipulating a wound so that it will not heal.

Table 19–3lists clues that should trigger suspicion of factitious disorder. Factitious disorder should be suspected whenever medical signs or symptoms defy conventional medical understanding or when they do not respond to usual medical treatment—for example, when a wound refuses to heal, or when test results show a pattern inconsistent with usual disease presentation. Factitious disorder patients may also demonstrate an exceptional eagerness to undergo invasive or extensive testing. They may deny access to collateral sources of information, refusing to sign releases of information and refusing to give contact information for family, friends, or former physicians. An extensive medical history, evidence of multiple surgeries, and reports of multiple drug allergies are also clues for the astute clinician. These patients often have jobs in a medical profession, have few visitors, and have been known to forecast the progression of their symptoms. They may supply falsified transcripts, test results, medical records, photos, personal documents, and physician reports. There are

even cases of patients bringing in pictures and x-rays of pathological conditions downloaded from the Internet and purported to be their own.

Simply considering factitious disorder is the first important step toward diagnosis. Subsequently, information supporting the diagnosis should be collected. A review of past medical records is mandatory, and may reveal inconsistencies. Collateral information from family, friends, or other health care providers may likewise show inconsistencies. Ward clerks may be the first to notice waxing and waning levels of distress depending on when the patient thinks clinicians are observing. The collection of laboratory specimens under close observation can minimize contamination or manipulation of samples. Laboratory values can provide diagnostic clues, as when microbiology reports reveal unusual pathogens or polymicrobial findings without apparent cause.

Factitious psychological symptoms are more challenging to diagnose because of the lack of clear objective markers for psychiatric symptoms. Methods of confirmation applicable with factitious physical disorders, such as contradictory laboratory results or findings on room searches, are difficult to apply here. Nevertheless, certain features of the patient's presentation can alert the psychiatrist that the patient may be simulating illness. As with physical factitious disorders, the patient may present with unusual symptoms that fail to correspond to any recognizable diagnosis. For example, one patient reported no other psychotic symptoms except seeing the entire cast of a television show emerge from her closet. Other features include worsening of symptoms when the patient is aware of being observed, inconsistencies in the patient's story over time, and the patient's overeagerness to recount symptoms of the illness. Such patients are often suggestible and readily admit to additional symptoms on questioning. They may refuse to cooperate with obtaining collateral information, and untraceable prior health care providers are not unusual. On admission to the ward, patients may reveal familiarity with hospital routine while denying previous hospitalizations. They may exhibit dramatic and unusual reactions to medications. They may demonstrate attention-getting tactics by breaking ward rules. Visitors are usually few or absent.

In contrast to patients with physical factitious disorders, who tend to avoid psychiatric care, these patients actively seek contact with the psychiatric system and readily acknowledge the presence of a psychiatric disorder, even if it may not be the one from which the patient actually suffers. Feigned bereavement and then psychosis appear to be the most common presenting symptoms. Ironically, one patient sought psychiatric admission for his Munchausen syndrome, claiming that he feigned physical illness when in fact there was no evidence of this.

In a series of 20 patients who presented with factitious bereavement, 15 also exhibited a history of factitious physical symptoms. A majority of them met criteria for other psychiatric disorders. They typically reported recent,

dramatic, violent, and often multiple deaths of loved ones, whereas collateral information showed that in fact no deaths had occurred. Another series of 12 cases of factitious bereavement yielded similar findings of complaints of violent, dramatic deaths and referral from medical wards for supposed physical illness. Although the complaints of bereavement were factitious, these patients exhibited prominent symptoms of depression. A theory about factitious bereavement is that patients are expressing the underlying truth of their emotional state, if not the factual truth. In contrast to the stigma associated with psychiatric illness, sympathy and care are usually offered to those in mourning. This may be why factitious bereavement is the most commonly seen factitious psychological disorder. One patient stated that she fabricated the complaint of bereavement to rationalize her depression.

In a series of 219 consecutive cases of psychosis, 9 (0.04 percent) met criteria for factitious disorder. All 9 patients demonstrated severe personality disorder. These patients were doing poorly when followed up 4 to 7 years later, with multiple hospitalizations and poor quality of life. Their outcome was no better than for schizophrenics concurrently followed. In another study of six patients who presented with feigned psychosis at the University Hospital of South Manchester, five of the six were diagnosed with schizophrenia on follow-up 3 months to 10 years later. Consistent with psychodynamic theories about factitious disorders representing an attempt to feel mastery, feigning psychosis may have been an attempt to feel control over early psychotic symptoms.

Table 19–2.
Methods of Factitious Symptom Production, Suggestive Signs, and Confirmatory Tests by Systems

Symptom	Method of Factitious Symptom Manufacture	Signs Suggestive of Factitiousness	Test for Factitious Met
Infectious Disease			
Fever	Injecting infectious material into vein/intravenous line	Higher than 41°C	Monitoring while temp taken
	Ingesting thyroid hormone	Not accompanied by other vital sign abnormalities	Using electronic therm reading skin surface
	Drinking hot fluids	Does not follow diurnal pattern	Recording oral and rec temperatures simult
	Manipulating thermometers with lightbulb or heating pad	No diaphoresis with rapid defervescence	Noting thermometer b names/serial numbe
	Substituting another thermometer		Measuring fresh urine s temperature
	Warm wax/wet cotton in ears		
Bacteremia	Injection of contaminated	Polymicrobial bacteremia	Culture

	substance		
		Absence of urologic/biliary/GI obstruction	
		Stool flora/pet flora noted	
HIV/AIDS	False history/reports	Normal CD4 count/undetectable viral load/Ab negative	Repeat HIV ELISA, west viral load
			Confirm test results wi laboratory
Gastrointestinal			
Diarrhea	Laxative abuse (magnesium, castor oil, phenolphthalein)	Metabolic alkalosis	Detection of laxative ir stool (usually need r tests)
		Acute hyperchloremic metabolic acidosis with normal anion gap	3-day stool collection
		Decreased serum bicarbonate with metabolic acidosis	Urine screen for phenc anthraquinones, bis
		High daily stool volume	Stool screen for magne mmol/L, phosphate
		Melanosis coli on sigmoidoscopy	
		Cathartic colon on barium enema	
		Low urine potassium concentration	
		High fecal fluid potassium concentration	
		Fecal fluid osmolality <290 mOsm/kg	
	Adding water to stool sample	High stool osmolar gap >125 mOsmol/kg	
Vomiting	Ipecac/cathartic abuse	Hypokalemia	Detection of ipecac in s chromatography
		High urine potassium concentration	Serum/urine emetine l
		Low urine chloride concentration	EKG with abnormalities toxicity
		Metabolic alkalosis with increased serum bicarbonate	
Pancreatitis	Spitting into urine sample (salivary amylase)	Hyperamylasuria with normal serum amylase	Obtain monitored sam
GI bleeding	Injection of blood from transfusions	Nasogastric tube shows blood despite normal endoscopy	Radiolabeling transfusi
	Ingestion of NSAIDs/salicylates	Single-stripe sign on colonoscopy	
Obstruction	Ingestion of loperamide		HPLC of serum/stool fc slowing agent

Renal			
Diuresis	Diuretic abuse	High urine potassium concentration	Detection of diuretics l chromatography
Bartter syndrome	Loop diuretic abuse	Hypokalemia, urine chloride low or variable	Renal biopsy for juxtag hyperplasia
	Self-induced vomiting		Screen urine for diuret
Metabolic			
Hypervitaminosis A	Vitamin A abuse	Increased gamma-glutamyltransferase Increased bilirubin	Increased serum/tissue vitamin A, retinoic a
Hypokalemia	Laxative abuse	See above	
Hyperkalemia	Injection of urine into blood sample	Potassium level incompatible with life	Monitoring sample gat normal result
Hypernatremia	Salt load	Fractional excretion of sodium high Gastric aspirate salt concentration >200 mmol/L	
Urologic			
Hematuria	Adding blood from another wound/meat	Lack of red blood cell/hemoglobin casts	"3-tube" urinalysis (mo first tube if urethral
	Traumatizing urethra	Lack of distorted red blood cells	Physical examination
	Inserting foreign bodies into the bladder		Radiography
	Ingestion of anticoagulants		
	Addition of coloring to urine		
Proteinuria	Injection of egg protein in bladder	Large day to day variations in urine protein concentration	Electrophoresis of urin protein/pure album
		Serum albumin concentration remains in normal range	Large albumin band or without transferrin i
		Lack of other signs of nephrotic syndrome	Antibody confirmation albumin
Bacteriuria	Injecting bacteria into bladder or urine specimens		
Urinary calculi	Adding pepper grains to urine		Infrared spectrophoto
	Inserting stones into bladder		Chemical analysis
	Submitting quartz, feldspar as samples		X-ray diffraction
			X-ray crystallography
Hematologic			
Anemia	Self-bloodletting	Decreased serum iron, ferritin, iron-binding capacity	Iron 59 elimination stu
		Decreased bone marrow	Urine/stool iron levels

		iron concentration	
		No evidence of bleeding	Blood typing
Sickle cell crisis		Serum protein electrophoresis normal	Genetic testing
			Hemoglobin electroph
Pancytopenia	Ingestion of chemotherapy		
Hemorrhage/purpura	Ingestion of anticoagulant (rodenticide/coumadin)	Prolonged prothrombin time (PT)	Warfarin/brodifacoum assay
	Heparin injection	Prolonged partial thromboplastin time (PTT) with normal PT	PTT measured every 2 observation
			Reversal with protamir
			Normal reptilase time
			Failure of PTT to corre mixture with norma
			Correction of PTT with removal measures
	Ingestion of quinidine	Purpura with thrombocytopenia	Detection of quinidine antiplatelet antibod
Endocrine			
Hypoglycemia	Injection of insulin	Serum insulin concentration > 100 mU/L	Normal proinsulin leve in insulinoma)
	Ingestion of oral hypoglycemic (sulfonylurea)		High insulin/low C-pep
	Ingestion of tolbutamide		Insulin/C-peptide ratio
	Manipulation of testing strips		Extreme elevations in i
			Insulin antibodies (less human recombinant
			Failure of glucagon to peptide
			Glyburide assay in seru
			Tolbutamide assay in s
Hyperglycemia	Insufficient administration of insulin		
Hyperthyroidism	Ingestion of thyroid hormone	Goiter/eye findings absent	Serum thyroglobulin lo undetectable
		Increased thyroxine or liothyronine levels	Radioactive iodine upt
		Low TSH in thyroid storm	
Cushing syndrome	Injecting/ingesting glucocorticoid	Plasma/urinary cortisol increased	Detection of synthetic glucocorticoids in se
	Adding glucocorticoid to urine	Plasma ACTH low/undetectable	
		Normal corticosterone/high cortisol	Serum cortisol and cor levels
Pheochromocytoma	Injection of epinephrine, isoproterenol	Normal chromogranin A level	Epinephrine/norepinep in serum/urine

	Ingestion of stimulant	Lack of metabolites of epinephrine (metanephrines)	Metanephrine/normet levels in serum/urin
	Injection of epinephrine into urine sample	Lack of norepinephrine increase after glucagon	Provocative (glucagon)/suppress (clonidine) tests
		Normal response to clonidine suppression test	44-metaiodobenzylgua nuclear scan
Hyperaldosteronism	Ingestion of black licorice/glycyrrhizic acid	Treatment-resistant hypokalemia	Serum glycyrrhizic acid
		Metabolic alkalosis	
		Hypernatremia	
Cardiovascular			
Arrhythmia	Rearranging of EKG leads	Unusual pattern on EKG	Supervise placement o
	Ingestion of digitalis, beta blockers, calcium channel blockers		Serum levels of drugs
			Electrochemiluminesce detect BBs in urine
Hypertension	Valsalva maneuver	Office BPs normal, home BPs high	Observe BP checks
	Ingestion of stimulants		Serum/urine assay for pseudoephedrine
Infarction	Beating self with towel to raise creatine kinase	EKGs don't change	Repeat EKGs
		Enzymes do not rise appropriately	Repeat troponins
Dermatologic			
Cheilitis granulomatosa	Self-inoculation with polyvinylpyrrolidone (PVP)		Liver/lymph node biop PVP in histiocytes
Dermatitis Artefacta	Self-mutilation	Distribution of lesions in reachable areas	Skin biopsy shows mec trauma, necrosis, bl
Erythematous lesion	Applying alcohol to lesion		
Nonhealing wounds	Injection of air/contaminants/foreign bodies	Heal when casted	Apply fluorescein to w examine hands/nail fluorescence
		Resistant to treatment	
		Subcutaneous emphysema	
Obstetrical			
Ectopic pregnancy	Injection of hCG	Negative urine b-hCG	Negative ultrasound
		Widely varying b-hCG levels	
Vaginal discharge	Intravaginal insertion	Inconsistent pH	
		Vaginal wall shows abrasion/denies intercourse	
Neurologic			
Movement disorders	Ingestion of neuroleptics to	MRI/EEG/EMG	Serum levels of neurol

	induce parkinsonism	inconsistent Hoover sign/tremor entrainable/distractable Prolactin elevated with antipsychotics	
Multiple sclerosis		Examination inconsistent	CSF protein electrophc
Pseudoseizures		History inconsistent Semiology inconsistent Never observed seizing by others	Video EEG
Pulmonary			
Asthma	Tape interfering with pulse oximeters Inhalation of irritant (talc) Ingestion of allergens	Intractable hypoxia Intractable status asthmaticus	Lung biopsy Serum salicylate level
Cystic fibrosis	Manipulation of sweat test	Normal x-rays Normal examination/lack of clubbing	Sweat potassium
Hemoptysis	Ingestion of anticoagulants Cough abrasion of lung Self-induced trauma to respiratory tract	Normal bronchoscopy/lack of blood	
Respiratory failure		Incongruous arterial blood gas Marked hypocapnia	
Rheumatologic			
Arthritis	Insertion of metal fragments in joints		X-rays Synovial aspirate analy
Lupus	Feigned history/borderline positive ANA	Repeat serum autoimmune tests negative Normal complement in active flare	Repeat serum tests
Vasculitis	Injection of contaminants to create purpura		Tissue biopsy of lesion: foreign material
Ophthalmologic			
Conjunctivitis	Instillation of foreign substance	Purulent discharge purposely left on skin and eyelashes Severity of discharge greater than severity of redness Severity of swelling less than severity of redness Mainly inferior conjunctival involvement	Corneal examination/c biopsy

Anisocoria	Instillation of atropine eye drops	Cornea uninvolved

Table 19–3.
Clues That Should Trigger Suspicion of Factitious Disorder

1. The patient has sought treatment at various different hospitals or clinics
2. The patient is an inconsistent, selective, or misleading informant; he or she resists allowing the treatment team access to outside sources of information
3. The course of the illness is atypical and does not follow the natural history of the presumed disease
4. A remarkable number of tests, consultations, and medical and surgical treatments have been done to little or no avail
5. The magnitude of symptoms consistently exceeds objective pathology or symptoms have proved to be exaggerated by the patient
6. Some findings are discovered to have been self-induced or at least worsened through self-manipulation
7. The patient might eagerly agree to or request invasive medical procedures or surgery
8. Physical evidence of a factitious cause might be discovered during the course of treatment
9. The patient predicts deteriorations or there are exacerbations shortly before their scheduled discharge
10. A diagnosis of factitious disorder has been explicitly considered by at least one health care professional
11. The patient is noncompliant with diagnostic or treatment recommendations or is disruptive on the unit
12. Evidence from laboratory or other tests disputes information provided by the patient
13. The patient has a history of work in the health care field
14. The patient engages in gratuitous, self-aggrandizing lying
15. The patient has been prescribed (or obtained) opiate drugs when not indicated
16. While seeking medical or surgical intervention, the patient opposes psychiatric assessment

In patients who complain of neuropsychiatric or cognitive symptoms, neuropsychological testing can be useful in establishing response bias, unusual patterns of deficits, poor levels of effort, or over-reporting of symptoms in a manner that suggest a motivation to appear ill. As in the assessment of malingering, these tests may come in the form of embedded tests, or stand-alone symptom validity tests, and have been extensively reviewed and approved for this use by the American Academy of Clinical Neuropsychologists. Examples of embedded tests include subscales and indices within the Personality Assessment Inventory (PAI); the Meyers index, which uses seven subscales of the Minnesota Multiphasic Personality Inventory (MMPI); the Fake Bad Scale, drawn from 43 questions on the MMPI that correlate with simulation; reliable digit span tests of memory. Stand-alone tests for simulating symptoms include the Structured Inventory of Malingered Symptomatology (SIMS); the Structured Interview of Reported Symptoms (SIRS); and the Miller Forensic Assessment of Symptoms Test (M-FAST) and the Test of Memory Malingering (TOMM). Important in the use of neuropsychological testing to supplement diagnosis of factitious disorder is an understanding of its limitations: while a set of

given tests can detect a pattern suggestive of simulation, it cannot ascertain motivation or etiology of the pattern. In other words, neuropsychological testing cannot make a diagnosis of factitious disorder—assessment by the clinician is required.

Conclusive confirmation of factitious disorder can be difficult. Covert surveillance, hidden video cameras, or searching the patient's belongings for syringes or illness-inducing substances can be employed after considering the patient's right to privacy and the bounds of the patient–doctor relationship. Legal counsel is advisable, and in some cases, court orders should be obtained. Consultation from a bioethics team can help weigh the benefits and risks of violations of privacy versus morbidity from factitious disorder. This phase of acquiring information to confirm a diagnosis is often a time of conflict for staff, who may have split opinions about the patient and about the means being employed to confirm diagnosis. Regular interdisciplinary meetings are helpful.

However, the obtaining of this collateral and confirmatory information to aid in the diagnosis of a factitious disorder brings up several ethical dilemmas:

1. Privacy: Patients have a right to privacy concerning their personal health information and activities, and providers are bound by biomedical ethical principles such as nonmaleficence (doing no harm) and patient autonomy. On the other hand, providers must also consider ethical principles of beneficence (doing what is good for the patient) and justice (doing what is right for society), and patients are expected to participate in their health care in good faith. When the risk of harm from suspected self-induced behavior becomes great enough, decisions to override patient preferences and seek information from covert searches or videotaping may be made, typically in consultation with legal counsel.
2. Confidentiality: Patients with factitious disorder will frequently refuse to release medical records to their providers, hampering the ability to diagnose and manage factitious behaviors. The US Health Insurance Portability and Accountability Act (HIPAA) allows health care providers to release medical records to one another for treatment purposes without an individual patient's consent, as long as both health care providers had a relationship with the patient and the records pertain to that relationship. Unauthorized disclosing of information to third parties that are not covered entities under the HIPAA law will need legal vetting to ensure compliance. Searching electronic medical records outside of a health care system represents a powerful modern tool to identify evidence of factitious disorder, and should also be informed by consultation with legal counsel.
3. Stigma: Careful workup for genuine organic causes, judicious use of indicated testing, and application of evidenced-based treatment must be

performed even in circumstances where a factitious disorder is heavily suspected. Devising tests for factitious illness and obtaining a resistant patient's consent for such tests requires a skillful approach that is nonconfrontational but truthful. There is no role for deceptiveness on the part of the provider in the diagnosis of factitious disorder, such as misrepresenting the reasons for tests or giving placebos disguised as active treatment. The intense negative countertransference on the part of health care staff deceived by patients with factitious disorder is formidable, and can result in distortions of usual care, violations of protocols and regulations, and even malpractice. If a patient becomes pejoratively labeled as a "faker" or "somatizer" it is very possible that a condition that has a genuine organic basis may go undetected or untreated.

Munchausen Syndrome. Diagnosis is often simpler than in other cases of factitious disorder because of the tendency for dramatic, exaggerated presentations. Treatment and management, however, are usually far more challenging. These patients often are new to a particular area, being prone to peregrination. Munchausen syndrome patients often appear eerily comfortable in hospital settings, immediately talking to nurses, physicians, and medical staff as peers. Some show up for hospital admission wearing surgical scrubs. They demonstrate *pseudologia fantastica,* a specific syndrome of autobiographical lying with four features as defined by King and Ford: (1) the stories are not entirely improbable and are built upon a basis of truth; (2) the stories are enduring; (3) the stories are not told purely for personal gain, and have a self-aggrandizing quality; and (4) they are not delusions, in that the patient can admit to falsehoods when confronted with conflicting facts. Munchausen patients often manifest hostility when their needs are not met or they are confronted by staff regarding suspicions of deception, and will then leave the hospital abruptly. They present repeatedly under different names to different hospitals, employing a repertoire of symptoms and presentations to gain admission and treatment. The cost of this disorder can be tremendous, as illustrated by the British report of a "million-dollar man" who over a 13-year period spent 1,300 days in psychiatric units, 556 days in prison, and 354 days in medical care for 261 hospital admissions. All diagnostic considerations that apply to factitious disorder also apply here. Because the risk of morbidity and mortality may be higher in these patients, it may be prudent to involve bioethics consultations and hospital legal counsel early.

Factitious Disorder Imposed on Another. Previously called *Munchausen syndrome by proxy,* the essential feature of this disorder is the intentional feigning or production of physical or psychological symptoms in another individual who is under the perpetrator's care. The perpetrator's motive is

to assume the sick role by proxy. Mothers of preverbal infants are the most common perpetrators, although fathers, other family members, babysitters, and even medical professionals also have been implicated. Victims can also be spouses, elderly parents, hospital patients, or anyone under the care of a perpetrator. In an unusual case of adult factitious disorder imposed on another, a 34-year-old man drugged his wife with sleeping pills in her coffee, then injected gasoline into her skin to cause lesions from which she eventually died. Subsequently, he hired a female babysitter to care for his children and repeated his actions with her. In each case, he assumed the role of the concerned caretaker in the center of medical drama. He was finally arrested for murder, and in prison, he sought to work in the infirmary. Factitious disorder imposed on another has also been cited as the etiology for epidemics of death at hospitals and nursing homes. Perpetrators have largely been nurses and aides who produced illness through various means, such as injection of insulin, lidocaine, digoxin, or other substances. In Great Britain in 1993, pediatric nurse Beverly Allitt was convicted of killing 4 children under her care and was given 13 life sentences. Florida registered nurse Bobbie Sue Terrell was accused of a rash of murders at a nursing home where she worked, and in 1988, she plead guilty to reduced charges of second-degree murder. Apparently, after learning of her infertility, she began demonstrating self-injurious behavior and factitious disorder.

By and large, however, reported cases involve mothers and their children. The challenges of recognizing this condition are illustrated by unfortunate examples. Nancy Reagan presented a Mother of the Year award in 1988 to foster mother Yvonne Eldridge for her valorous care of ill children. Eldridge was later recognized as a factitious disorder imposed on another perpetrator and was convicted of child abuse in 1996. Hillary Rodham Clinton publicly recognized Kathy Bush's dedication to her daughter. Bush was later recognized as a perpetrator of factitious disorder imposed on another and was sent to prison in 2002 for child abuse that had resulted in 200 hospitalizations, 40 surgeries, and 15 bouts of sepsis for her daughter.

Table 19–4.
Clues Triggering Suspicion for Factitious Disorder Imposed on Self

1. Diagnosis does not match the objective findings
2. Signs or symptoms are bizarre
3. Caregiver or suspected offender does not express relief or pleasure when told that dependent is improving or that dependent does not have a particular illness
4. Inconsistent histories of symptoms from different observers
5. Caregiver insists on invasive or painful procedures or hospitalizations
6. Caregiver's behavior does not match expressed distress or report of symptoms (e.g., unusually calm)
7. Signs and symptoms begin only in the presence of one caregiver

8. Sibling or another dependent has or had an unusual or unexplained illness or death
9. Sensitivity to multiple environmental substances or medicines
10. Failure of the dependent's illness to respond to its normal treatments or unusual intolerance to those treatments
11. Caregiver publicly solicits sympathy or donations or benefits because of the dependent's rare illness
12. Extensive unusual illness history in the caregiver or caregiver's family; caregiver's history of somatization disorders
13. Caregiver seeks other medical opinions when told the dependent does not have illness
14. Caregiver perseverates about borderline abnormal results of no clinical relevance despite repeated reassurance, or refutes the validity of normal results

Sudden infant death syndrome (SIDS) was initially thought to have a strong genetic component because it often occurred in siblings. Today, when siblings die of SIDS, infanticide is suspected. Indeed, a study of 81 children who died of SIDS or natural causes and who later were found to have died at the hands of their parents showed that half of the perpetrating parents had factitious disorder or another somatizing disorder. The comorbidity of factitious disorder and factitious disorder by proxy is estimated to be from 10 to 30 percent. When one child is victimized, there is a high likelihood a sibling will also be victimized, usually serially.

Table 19–4lists clues that should trigger suspicion of factitious disorder imposed on another.

The variety of medical presentations of factitious disorder imposed on another is impressive. In the first comprehensive review of the disorder, published in 1987, Rosenberg described 68 induced or fabricated signs or symptoms. The most common presentations among 117 cases were bleeding (44 percent), seizures (42 percent), central nervous system depression (19 percent), apnea (15 percent), diarrhea (11 percent), vomiting (10 percent), fever (10 percent), and rash (9 percent). Many children had more than one presentation. Twenty-five percent of cases involved simulation of illness, 50 percent involved production of illness, and in the remaining 25 percent, simulation and illness production were concurrent.

Perpetrating caregivers usually appear to be concerned and interested in their children's care. They tend to be exemplary in their interactions with medical staff, enlisting support and sympathy, often crossing professional boundaries by eating meals with staff or helping nurses with duties. They may demonstrate unusual willingness or even excitement at the prospect of invasive procedures for their children.

Table 19–5.
Diagnostic Principles and Tools for Factitious Disorder Imposed on Another

1. Ask: Are the history, signs, and symptoms of disease credible?
2. Ask: Is the child receiving unnecessary and harmful or potentially harmful medical care?
3. Ask: If so, who is instigating the evaluations and treatment?
4. Multidisciplinary evaluation involving medical, psychosocial, child protective, and legal

professionals
5. Consultation with a specialist in child abuse pediatrics/elder abuse
6. Complete review of the medical record
7. Construction of a chronological summary of medical contacts with attention to:
   - ► use of multiple medical facilities
   - ► excessive or inappropriate pattern of utilization
   - ► pattern of missed appointments and discharge against medical advice
   - ► pattern of misrepresentation of opinions of physicians about the child's medical problems
8. Attribution of all medical information to its source (e.g., Did the nurse observe apnea or did the caretaker inform the nurse of it?)
9. Creation of a table with the following elements for each health contact: Name; date; location; reason for contact; reported symptoms as stated by caregiver; objective signs documented by physician; conclusions/diagnoses made; treatment provided; efficacy of treatment.
10. Ask: Does the history provided by the caregiver match the history in the medical record and/or the diagnosis made by the physician?
11. Contact involved physicians and solicit concerns about possible fabrication of illness
12. Have all involved staff meet and develop a consensus management plan
13. Interview the child separately from the caregiver, if verbal, for his or her recollection of any symptoms, including when and where they occurred
14. Take a careful social and family history, including for any unusual or frequent illnesses in the extended family and siblings
15. Perform diagnostic testing for fabricated illness alongside tests for organic etiologies
16. Separate the child from the suspected caregiver and observe for improvement
17. Employ covert video surveillance while the child is hospitalized, with legal and ethical consultation

Factitious disorder imposed on another should not be a diagnosis of exclusion, but confirmatory evidence should be actively pursued with the goal of lessening risk to the victim and assuring the victim's safety. The recommended steps for diagnosing suspected factitious disorder imposed on another by the American Academy of Pediatrics is outlined in Table 19–5.

The gold standard for confirming factitious disorder imposed on another is covert video surveillance that may record a perpetrator causing harm. Examples of documented harms include caregivers suffocating children, causing fractures, and administering poisons or injecting harmful substances into intravenous lines. Such videos have also shown mothers appearing concerned in the presence of staff but behaving indifferently toward their children when unaware of being watched.

Despite the benefits of objective documentation of abuse, the use of covert video surveillance is not without risk: (1) Damage may be done to the child prior to health care intervention; (2) it may interrupt a potential therapeutic alliance; (3) some consider covert taping to violate the parents' right to privacy or represent entrapment; (4) video footage can be difficult to interpret and may result in false accusations; and (5) covert taping can be costly and require constant real-time monitoring to prevent harm. Covert video should only be undertaken after consultation with legal counsel because a court order may be legally required. A bioethics consultation can help weigh the potential benefits to the child versus compromises of privacy for the parent. An institutional protocol should be

in place regarding covert taping, including plans for monitoring and interventions if the child is found to be at risk.

Other means of confirming factitious disorder imposed on another include searching a mother's belongings for illness-inducing agents, reviewing collateral information and past medical records for inconsistencies, gathering information on siblings, recording temporal associations between parental visits and a child's signs and symptoms, observing the child's well-being when he or she is removed from the parent's care, and analyzing specimens taken in the presence of the parent compared to those taken in the parent's absence.

The international literature on factitious disorder imposed on another indicates that signs and symptoms appear to be consistent across the world. Perpetrators are usually mothers, and serial abuse of children is common. This does not appear to be a phenomenon exclusive to medicalized societies.

Pathology and Laboratory Reports. No laboratory or pathology tests are diagnostic of factitious disorders, although they may help to confirm the diagnosis by demonstrating deception. In many cases a laboratory finding may be the only objective sign that factitious behavior is taking place.

See Table 19–2 for laboratory tests that can be helpful in determining that a given sign or symptom has been artificially induced.

Differential Diagnosis

The lynchpin of diagnosis of factitious disorder imposed on self or another is the confirmation of deceptive behavior, perpetrated by the patient, in the service of simulating or exacerbating disease.

An underlying physical or psychiatric disorder is the main consideration in the differential diagnosis of suspected factitious disorder. Unusual, inconsistent, or poorly characterized symptoms must still be explored for nonfactitious etiologies according to best practices, evidence-based guidelines, and clinical indication. For example, a 46-year-old woman with type 2 diabetes manifested multiple episodes of hypoglycemia despite several inpatient hospitalizations to optimize her diet and insulin regimen, and she was suspected of factitious disorder until it was discovered she had developed antibodies to insulin.

Underlying or prodromal psychiatric/neurologic illnesses may initially be mistaken for psychological factitious disorders. Several case reports document the development of full-blown psychotic disorder following the emergence of factitious illness behavior. Cognitive disorders due to cerebral lesions, neurodegenerative disorders, and delirium may lead to self-induced illness and deceptive behavior. In cases of frontotemporal dementia, the deceptions are often quite unsophisticated and the attempts to obfuscate inept. Neuropsychological and functional testing may help to

delineate the cognitive deficits in such cases.

Even when deception is present, comorbid genuine illness is more often the rule than the exception in factitious disorders. Factitious disorder patients frequently aggravate an actual physical illness. Many patients who feign pseudoseizures also have true seizures. Some who manipulate blood sugars to produce symptoms are also diabetic. Munchausen syndrome patients may become addicted to narcotics or learn to seek disability payments and thereby also demonstrate substance abuse and malingering. One area of diagnostic differentiation concerns psychiatric disorders which can present similarly to factitious disorder with unusual or inconsistent physical or psychological symptoms, and may even be comorbid when deceptive behavior is positively confirmed.

In *somatic symptom disorder* there are one or more physical symptoms which cause disproportionate distress manifesting in thoughts, feelings, or behaviors. If the physical symptom does not have a clear etiology or does not conform to typical clinical phenomenology, suspicions may be raised for factitious behaviors, but positive evidence for deceptiveness will be lacking. In *illness anxiety disorder* (formerly hypochondriasis), normal physiologic sensations are perceived as harbingers of dread disease despite negative tests and reassurance to the contrary. It is important to remember that a presentation that does not conform to an identifiable medical or mental disorder, or the diagnosis of factitious disorder, do not necessarily rule out the presence of a true *medical condition not associated with intentional symptom falsification*.

If a symptom is neurological and is not consistent with neurologic pathophysiology, then *conversion disorder* should be diagnosed. For example, a 35-year-old man presented to emergency care with rapidly progressive lower extremity flaccid paralysis following a mild head injury, but workup could not discern any neurologic disorder, including electromyography. There was no evidence of deception or material gain in this patient, but when a random loud noise startled him, the patient leaped off the examination table quite nimbly, and was diagnosed with conversion. The use of video electroencephalographic monitoring allows for the rapid capturing and diagnosis of nonresponsive episodes mimicking absence seizures, complex partial seizures, or syncope.

Dissociative disorders such as *dissociative amnesia* or *dissociative identity disorder* may involve episodes of inability to recall important personal information too extensive to be due to ordinary forgetfulness. These episodes can resemble psychological presentations of factitious disorders in which amnesia or cognitive impairment features prominently. However, they do not involve deceptive behaviors.

A second area of diagnostic differentiation concerns self-inflicted harm. A patient with *borderline personality disorder* may engage in self-destructive, physically injurious, or psychologically damaging behavior

resulting in illness (such as overdosing on acetaminophen or cutting one's arm) for the purposes of stress reduction, pain relief, or conveying distress. If the induction of injury is associated with deception, then factitious disorder may be diagnosed. However, patients who have attempted suicide and caused physical harm may lie about their suicidal ideation, intent, plan, or the aftermath of the attempt. This may be in the service of avoiding detection or hospitalization and carrying out another attempt successfully. This is not factitious disorder nor malingering *per se,* but the expectable behavior of a person bent on suicide, and must be screened for with a high level of suspicion.

If a patient's habitual behaviors contribute to worsening health, such as overeating leading to type 2 diabetes, work-related stress exacerbating hypertension, or cigarette smoking contributing to emphysema, and there is no overt deception in the patient's presentation for care, then the appropriate diagnosis is *psychological factors affecting other medical conditions*. This is a new somatic symptom disorder in DSM-5, and is meant to capture psychological states which complicate the course of a physical disease, interfere with its treatment, and represent a well-established health risk. When assessing whether deception is involved, it is important to remember that patients commonly minimize reporting of behaviors that they expect health care providers to disapprove of, or that are not socially, culturally, or legally sanctioned. This should not automatically trigger a diagnosis of factitious disorder without a full exploration of the reasons for minimization. Likewise, ostensibly deceptive behavior may be a result of a lack of cultural understanding on the part of the health care provider, a language barrier between provider and patient, or due to violence or the threat of violence in the patient's life. For example, a 48-year-old woman with intractable hypertensive urgency endorsed taking her medications faithfully to her primary care physician, but with intensive case management involvement it became clear she did not take her medications as prescribed because she was being forced into prostitution by her husband under threat of death, which she never volunteered to her providers. As Bass and others have recommended, diagnosing a patient's deceptive behavior as a disorder is less important than understanding the underpinnings of the behavior.

Once confirmed, if deceptive behavior is performed to obtain a material gain such as to procure disability payments, narcotics, or an excuse from work, *malingering* is diagnosed. Ganser syndrome, also known as "prison psychosis," presents a related challenge. It was first described in 1897, consisting of twilight states, memory disturbances, and *vorbeireden,* a German term loosely translated as "talking at cross purposes." Ganser syndrome is also known as the "syndrome of approximate answers," in which answers to simple questions are erroneous in a manner that implies knowledge of the correct answer. An example would be answering "3" to

the question of "How many legs does a horse have?" Ganser syndrome may be due to dissociation, malingering, or organic causes.

In actuality, whether a behavior arises from external or internal origins can be difficult to discern. A clinician may be unaware of a patient's upcoming court date and thus potential gain from being sick. For these reasons, Feldman and coworkers argued that these disorders fall along a continuum: "there is little to be gained, and much to be lost, by the use of the current practice of viewing the somatoform disorders, factitious disorder, and malingering as discrete and distinct clinical entities." Conscious versus unconscious production of symptoms can be especially difficult to discern for psychological signs and symptoms, as there are fewer objective tests of mental status. It is for this reason that the DSM-5 no longer requires clinicians to determine that patients are explicitly aware of their deceptive behavior, only that there is deception, and not for evident secondary gain.

Mr. J, a 60-year-old man, presented to his local emergency department complaining of nausea, vomiting, and dizziness. He was admitted for possible bleeding peptic ulcer, and was also found to have a skin rash and depression. His rash was diagnosed as scabies, but Mr. J refused to shower or change out of his dirty garments. He then requested to be discharged. Psychiatric consultation was requested for evaluation of depression.

Mr. J stated he was grieving the deaths of his younger daughter and wife in a car accident 2 and a half years ago, and that his life had been disintegrating since then. He was staying at friends' houses after leaving his home with his older daughter, as he "couldn't stand being there" any longer because of the grief. He had not sought treatment or professional help, he said, because, "I don't want to bother my family or doctors with anything." However, he did endorse having "snapped" and having two "breakdowns" during which he could not remember where he was or who he was, and was hospitalized at a psychiatric unit in another state, but could not recall any details of this incident. He refused to give any contact information for family or friends, citing concerns about bothering them too much, and refused treatment for depression.

His social history included employment as a chemical engineer in his own laboratory for the nearby military base, a PhD in parapsychology and forensic psychology, training as a paramedic, a father who was a fighter pilot, and a mother who was a physician. He wistfully recalled his family being stationed overseas. He related that his daughters both held advanced degrees and his younger daughter was on her way to start a position at a prestigious university hospital when she was killed in the accident.

Inconsistencies between his presentation and his history prompted a search for records in the wider hospital system. Records filed under the same name and date of birth at another facility detailed a significantly different history. When he was approached with these disparate facts he angrily denied them. He was discharged the next day after it was determined he was not a danger to himself or others.

Two weeks later Mr. J presented with worsening altered mental status, shortness of breath, abdominal pain, nausea, cough, dysphagia, and vertigo. He was found to have pneumonia, a posterior inferior cerebellar artery (PICA) stroke, a pulmonary embolus, bilateral lower extremity venous thromboses, and esophagitis causing gastrointestinal bleeding. He was treated with antibiotics, anticoagulation, inferior vena cava filter placement, gastric tube placement. MR brain imaging revealed an old left temporal lobe infarct with encephalomalacia and gliosis, in addition to the PICA infarct.

Psychiatric evaluation after Mr. J's stroke revealed significantly impaired cognition (12/25 on the Montreal Cognitive Assessment) with deficits in short-term memory, attention, and executive function. He could not recall what he was being told by hospital

staff about his condition, and needed frequent repetition of information. He then revealed that he could not recall any remote memories about his early life such as schooling, family possessions, or even where he was born. He continued to refuse to speak with family or friends, however, hospital staff had to contact his daughter so that someone could serve as surrogate decision maker.

Over several weeks Mr. J became more attentive and his short-term memory improved. He revealed that in fact his wife and daughter had not died in a car accident 2 years ago, that this was a story lifted from the life of his friend "Bill." He could offer no explanation for why he had given this as his life history. He also recounted that he was interested in parapsychology but denied having an advanced degree or practicing in this field. He seemed surprised by the details of his life that were obtained from outside hospital records. Interestingly, he complained that the hospital staff was "keeping secrets" from him. He was transferred to the outside hospital familiar with his history for longer-term care.

Three months later Mr. J presented to the local hospital with a supratherapeutic level of anticoagulation and pain in his lower extremities. It was discovered that when Mr. J had been transferred to the subacute facility 3 months prior, the facility received a telephone call from a hospital in another state, stating that Mr. J was in fact not Mr. J, because the "real" Mr. J was admitted at their facility—family positively confirmed that their relative was the person at the other hospital. The patient who had suffered the scabies, cerebellar stroke, and pulmonary emboli was not the person he said he was, nor was he the person the outside hospital thought he was—he was Mr. R.

When Mr. R was interviewed about his identity, he could not recall any episodic history of his life prior to his hospitalizations, nor could he recall his parents' names, wife's name, whether he had children, where he hailed from, or his employment history. He did not recall his previous deceptive stories. When presented with pictures of famous persons and allusions to famous world events, he denied knowing anyone named Hitler, Marilyn Monroe, Ronald Reagan. He stated he recognized Jesus from having read the Bible at the outside hospital, and also knew of Martin Luther of the Reformation from reading. His short-term memory, attention, praxis, and executive function seemed very intact. He was able to recall numerous accurate details about his recent hospitalizations. He refused to give details of where he had been living, and who had been helping him, always out of a concern to "not bother them."

Mr. R was hospitalized on a psychiatric ward, out of concern that his amnesia, whether factitious or organic, was interfering with his ability to manage his medical issues, in particular the anticoagulation. He refused to engage in formal cognitive testing, but accepted treatment for depression and aftercare planning, and was discharged with appropriate appointments and services in place.

After discharge, Mr. R regularly attended anticoagulation clinic, physical therapy, wound clinic, and psychiatric appointments. During his mental health appointments, he endorsed that parts of his remote memory were coming back, including how his family lived overseas, his father served in the military, and his mother was a physician. However, he could not call to mind any memory of his family's appearances or names. He later claimed that at his living situation (a group home, but what he termed "someone's garage") he fixed computers that others brought to him, that others said he was "trained as a paramedic" and that he was "head of an Information Technology department."

DISCUSSION

Mr. R was finally diagnosed with *factitious disorder imposed on self* based on his deceptive behavior (giving a false name and history) in the service of simulating disease (feigned grief and amnesia). Mr. R admitted that he had acted deceptively, making this a conscious willful act, although he never articulated his motivation. His life history contained fantastical elements and he tended toward embellishment and confabulation (*pseudologia fantastica*). No clear material gain was ever established.

This case serves as a cautionary tale against assuming that deceptive behavior implies a lack of organic pathology. Mr. R had an ischemic stroke and evidence of prior strokes causing organic cognitive and memory impairment, but he also gave a false identity.

The available information makes it impossible to determine whether Mr. R's previous amnestic episodes during times of stress were real; one may rightly wonder what effect his

old temporal lobe stroke had on his memory and whether it played any role in his deceptive behavior.

Factitious Disorder Imposed on Another. When considering the differential diagnosis of this disorder, in addition to the considerations for factitious disorder imposed on self, clinicians should keep in mind the wide range of behavior among normal parents who seek medical care for their children. Again, according to Meadow, "exaggeration and mild deception are part of everyday behavior." The possibility of factitious disorder and malingering by children themselves should be considered, especially in older children and teenagers. Other *by-proxy* possibilities should also be considered because children may manifest psychopathologies of their parents. For example, in hypochondriasis by proxy, a hypochondriac mother preoccupied with her child's health may repetitively seek pediatric care leading to unnecessary procedures and iatrogenic illness for the child. In anorexia nervosa by proxy, an anorexic mother may restrict her child's food due to fears of excessive weight in her child. A mother with malingering by proxy may put her child through multiple evaluations to maintain disability or welfare payments. A paranoid father with a history of psychosis feared that his son was being poisoned by breast milk and insisted that the emergency room check his son's hair for mercury. It is important to remember that putting a child through unnecessary tests or treatments which can cause harm is child maltreatment, regardless of the perpetrator's motives.

Other *by-proxy* syndromes described in the literature include *masquerade syndrome* (in which illness fabrication results in the child's increasing dependency on the mother), *mothering to death* (in which the child is confined to a sick role as if the child were ill, while avoiding physicians and agencies), *extreme illness exaggeration* (in which a parent exaggerates their child's symptoms in an effort to increase a pediatrician's attention to the child), and *achievement by proxy* (as in youth sports).

Course and Prognosis

Factitious Disorder Imposed on Self. The course of factitious disorder is difficult to determine, as many patients drop out of treatment once a factitious disorder has been raised. The wide spectrum of this disorder should be kept in mind when considering course and prognosis. At one end of the spectrum, factitious illness behavior could be considered within normal range, as when a patient magnifies his symptoms to seek reassurance from a physician, or when a patient lies about completing her antibiotic course for cellulitis in order to win attention. Further along the spectrum, factitious illness behavior becomes a maladaptive coping mechanism but not a chronic pattern of deception. These less pernicious

cases may respond to a clinician's sensitive inquiry into the precipitants to factitious behavior, and coaching the patient to find more adaptive solutions. Concurrent mood, anxiety, or substance abuse disorders bode a better prognosis; comorbid personality disorders, especially antisocial personality disorder, bode a poorer prognosis. According to some experts, many factitious disorders remit around the age of 40 years, corresponding to the age of remission for many with borderline personality disorder. Unfortunately, Munchausen syndrome tends to have an unremitting, refractory course, and is associated with substantial morbidity and mortality. After efforts to engage a patient in treatment fail, management should be directed toward harm reduction rather than cure.

In Sutherland and Rodin's series of 10 patients diagnosed on a psychiatric consultation service with factitious disorder, only 1 out of 9 patients confronted with the diagnoses admitted to factitious behavior, and only 1 engaged in psychiatric treatment. In Krahn's series, only one in six of confronted patients admitted to factitious behavior, and only 12 percent agreed to psychiatric treatment. These studies would seem to support the contention that confrontation is unlikely helpfully engage a patient. However, in Reich and Gottfried's series of factitious disorder patients, of 33 confronted with the diagnosis, the majority improved, including 4 chronic cases (only 13 ever acknowledged their role). The course of factitious disorder is usually intermittent: isolated events and frequently recurring/unremitting episodes are less common. Onset is usually in early adulthood, often after hospitalization for a medical condition or a mental disorder. In individuals with recurrent episodes of falsification or signs and symptoms of illness and/or induction of injury, this pattern of successive deceptive interactions with medical personnel, may become life-long.

Once clinicians recognize a factitious physical disorder, they should be alert to the possibility of patient-initiated litigation to defend the "realness" of their self-induced illness. Litigation often fulfills the same psychodynamic motivations that led to the original factitious illness behavior. Actions brought against clinicians may provide an additional outlet for the patient's anger, as well as provide an increased level of care or attention. Litigation may be initiated by the patient's family, who may not understand that the patient's illness is self-inflicted. Litigation may be a marker for increased risk of death. In a study of 20 factitious disorder patients with concurrent lawsuits, 4 cases ended in death.

In general, psychological presentations of factitious disorder imposed on self have a poor prognosis comparable to that of Munchausen syndrome. A 1982 study of patients with factitious psychosis showed that 4 to 7 years later, these patients functioned at a level comparable to that of schizophrenic patients in the same study. Other studies found that feigned psychosis is commonly a precursor to actual psychosis, leading Hays to comment, "Acting crazy may bode more ill than being crazy."

Mr. K is a 43-year-old male who presented repeatedly since age 26 to regional tertiary care hospitals across two states. His initial presentations in emergency departments consisted of repeated rattlesnake, cobra, and viper bites, which he said occurred because of working as a volunteer "snake wrangler." His work reportedly entailed catching snakes, raising them, feeding them, and giving them medications, but he also reported bites from picking up snakes he found by the road, sometimes while inebriated.

His past psychiatric history was noteworthy for several hospitalizations as a child for suicidal ideation and behavior, which Mr. K attributed to "wanting to get out of the house while [his mother] was menopausal." He had been given a diagnosis of conduct disorder. Even as a teenager he was noted to frequently sustain snake bites, and to not comply with recommended treatment.

FIRST STAGE

Since his early 30s, Mr. K presented in excess of 30 times for the treatment of snake bite envenomation. He received experimental antibody-fragment antivenin preparations without ill effects so frequently (>150 vials) that his treatment was reported as robust evidence of the safety of the preparation. Complications of his snake bites included anaphylaxis often requiring intensive care, compartment syndromes in both his arms treated with fasciotomies, and finger and toe amputations.

Mr. K denied any intent to be bitten, saying these bites came as a result of his normal snake handling. Yet he expressed no remorse or desire to change his behaviors. He stated during one admission that snake handling was his only source of income, although he also described performing construction work at other times. Over the course of his hospitalizations for anaphylactic symptoms of snake bite, he began to decline antivenin, withhold history, and request escalating amounts of pain medication. During one admission he assaulted a health care worker because he did not feel he had been given enough analgesic.

SECOND STAGE

At age 37, he presented to the emergency department with chandelier lightbulbs in his rectum, stating that they were inserted by assailants in the city park. These were removed fully intact, but he presented again 2 weeks later saying glass shards were retained in his rectum, adamantly denying he had inserted them himself. Proctoscopy revealed glass shards that could not be removed, and colectomy with colostomy was performed.

At age 39, he presented with accidental ethylene glycol ingestion, which he said he mistook for his energy drink while working on an automobile engine with a friend. The next year he presented having ingested 13 shards of glass and 2 razor blades, which he told an admitting physician were shoved down his throat by the same assailants who had inserted the lightbulbs in his rectum. He later told another physician that he had been working construction, holding screws in his mouth, and had accidentally ingested them. X-ray confirmed the presence of several bolts and screws in his stomach.

At age 41 he presented having swallowed fishhooks and screws, and inserted a lightbulb into his rectum. This was the first time he admitted that this was self-inflicted. He endorsed depression and anxiety as factors driving him to perform these acts.

THIRD STAGE

In his 42nd year, he had 15 admissions at a local tertiary hospital for intentionally ingesting nail clippers, bottle tops, nails, screws, scissors, tent stakes, and knives, and inserting lightbulbs into his rectum. He presented with snake bites, but less frequently than earlier in his history. He admitted that he had thoughts of intentionally getting bitten in order to kill himself. He occasionally complained of command auditory hallucinations telling him to ingest objects, such as a letter opener, but this was infrequent. On most admissions he was able to say that negative feelings such as depression, loneliness, stress, or anxiety triggered the impulse to harm himself.

During stages two and three, psychiatric consultations were obtained every admission. The consultants attempted to assess without confrontation, to address reinforcing factors such as pain and sedative medications, to identify and address psychiatric symptoms that

might prompt Mr. K's behaviors, and to build his coping skills. Of note, medication trials including fluoxetine, buspirone, lorazepam, aripiprazole, olanzapine, and trazodone were uniformly ineffective, or Mr. K did not take them. He continually asked for benzodiazepines for anxiety and opiates for chronic pain.

FOURTH STAGE

At age 43, from the fall into that winter, he had no presentations to the local tertiary care hospital. Then he presented having swallowed a tent stake, resulting in esophageal tear and pneumothorax treated with thoracotomy and multiple chest tubes. During that admission, he cited factors such as having an apartment, roommates, and friends caring for him, which kept him from ingesting objects for that span of time. It was evident that social support and positive relationships aided in Mr. K's sense of stability, whereas ruptures in these areas led to further self-destructive behaviors. Of note, interview with one of these friends indicated she thought he was diagnosed with cancer and had been hospitalized with a collapsed lung.

At age 44, during the spring, Mr. K did not present to the tertiary center. In June, he presented after swallowing brush bristles. By this time he had a case worker, disability income, and a representative payee. The precipitant for his relapse was eviction for not following his landlord's rules. Since discharge, he has not presented for 7 months.

Of note, a review of controlled substance prescriptions indicated that at age 43, Mr. K. had visited 13 local providers between April and September, including 8 dentists, obtained 36 prescriptions for oxycodone, hydrocodone, or alprazolam, none written for longer than a week's worth of medication, and utilized 7 separate pharmacies.

DISCUSSION

Mr. K manifested factitious disorder imposed on self. At least initially, there was no identified gain except to play the role of a patient. Later, narcotic analgesics appeared to have become a motivator for Mr. K. His eventual presentations with swallowed and inserted objects went far beyond what is typical for narcotic-seeking malingerers. He eventually admitted to fabricating stories, which confirmed the conscious and willful dimension of his behavior. He fit the typical profile of a patient with Munchausen syndrome: a young male with antisocial traits, who demonstrated a prolonged, chronic course with intermittent hostility, travels to multiple institutions and practitioners (peregrination), and deceptive stories that strained one's credulity (*pseudologia fantastica*).

This case exemplifies how a factitious behavior can be a prequel or sequel of several other disorders during a course of illness. What seemed to start as an occupational hazard (bites while handling snakes) became a source of narcotics (malingering). One behavior became a grab-bag of self-destructive acts (ingesting and inserting objects), about which Mr. K was untruthful (factitious disorder). Finally, when Mr. K admitted to his self-destructive tendencies, these appeared to be stereotyped emotional responses accompanying poor distress tolerance (personality disorder). Mr. K's presentations became less frequent with interventions for personality-disordered patients: stable living and financial arrangements, a sympathetic support system, and case management with frequent contacts. In addition to the phenomenological similarities between factitious disorder, malingering, personality disorders, and somatic symptom disorders, the fact that treatment of factitious disorder bears similarity to treatment for somatization and personality disorder may support the perspective that factitious behavior is simply a dimension of other disorders and not a separate entity.

Factitious Disorder Imposed on Another. Victims of factitious disorder imposed on another are at very high risk. Mortality rate estimates of victims range from 6 to 22 percent, with deaths commonly occurring through suffocation or poisoning. In Rosenberg's 1987 review, 10 of 117 children died, yielding a mortality rate of 9 percent, with the youngest children being the most vulnerable to death. A sobering statistic is that in 20 percent of the deaths, caretakers were confronted with the diagnosis of

factitious disorder imposed on another, but the children were not removed from their custody. Of the survivors, 8 percent suffered long-term morbidity, including destructive joint changes, limp, cerebral palsy, cortical blindness, and psychiatric problems. Approximately 75 percent of the morbidity was precipitated by caregivers' behaviors occurring while the child was hospitalized. In 25 percent of the cases, morbidity was solely iatrogenic, caused by medical procedures and investigations.

Factitious disorder imposed on another is often recurrent. In a 1998 follow-up study by Davis et al. 17 percent of victims who returned home without restrictions were abused again. In cases where the child was poisoned or suffocated, repeat abuse rates were as high as 50 percent, in some cases leading to death. These statistics reinforce the importance of prompt diagnosis, proper intervention, and protection.

Siblings of victims are at great risk. Studies indicate that 9 to 29 percent of siblings die, underscoring the importance of checking and protecting all children of a given perpetrator. Siblings tend to be the focus serially rather than simultaneously, and so clinicians should always be alert for past and future victims. In 1990, Meadow studied 27 young children suffocated by their mothers. Over 13 years, 9 died and 1 suffered severe brain damage. These 27 children had 18 siblings who had died "suddenly and unexpectedly in early life." In another study of 32 children presenting with factitious epilepsy, 21 percent of siblings (7 of 33) had died of sudden infant death syndrome. In a study of suffocation cases, 21 percent of siblings (3 of 14) had died unexpectedly. In another study of 56 victims of factitious disorder imposed on another, 39 percent of siblings were subjected to illness fabrication. Among Rosenberg's review of 117 cases, 10 siblings had died under "unusual circumstances."

Treatment and Management

Guidelines for the treatment and management of factitious disorder are given in Table 19–6.

Factitious Disorder Imposed on Self. Three major goals should guide the treatment and management of factitious disorders: (1) to reduce the risk of morbidity and mortality, (2) to address the underlying emotional needs or psychiatric diagnosis that may be driving factitious illness behavior, and (3) to be mindful of legal and ethical issues.

1. Risk Reduction: The Hippocratic doctrine of "first do no harm" should be foremost. Prompt recognition of factitious disorder can reduce morbidity or mortality. Active pursuit of the diagnosis and timely management are essential. For example, in the case of a woman who reported an extensive family history of breast cancer and sought prophylactic mastectomy, early suspicion of factitious disorder may have prevented the

unnecessary bilateral mastectomies that were performed. Recognition of factitious asthma or arthritis prevents the administration of steroid medications, which pose significant side effects. Invasive, potentially harmful tests, procedures, and treatments can be minimized. Once symptoms are confirmed as factitious, clinicians should administer medical treatment according to their clinical judgment, considering objective evidence and keeping in mind that the patient's complaints and requests can be deceptive.

Table 19–6.
Guidelines for Management and Treatment of Factitious Disorder

Keep in mind that active pursuit of a prompt diagnosis can minimize the risk of morbidity and mortality.
Minimize harm. Avoid unnecessary tests and procedures, especially if they are invasive. Treat according to clinical judgment, keeping in mind that subjective complaints may be deceptive.
Arrange regular interdisciplinary meetings to reduce conflict and splitting among staff. Manage staff countertransference.
Consider facilitating healing by using the double-bind technique or face-saving behavioral strategies such as self-hypnosis or biofeedback.
Steer the patient toward psychiatric treatment in an empathic, nonconfrontational, face-saving manner. Avoid aggressive direct confrontation.
Treat underlying psychiatric disturbances such as Axis I disorders and Axis II disorders. In psychotherapy, address coping strategies and emotional conflicts.
Appoint a primary care provider as a gatekeeper for all medical and psychiatric treatment.
Consider involving risk management and bioethicists from an early point.
Consider appointing a guardian for medical and psychiatric decisions.
As a behavioral disincentive, consider prosecution for fraud.

Management of countertransference is a priority to reduce risk, because a clinician's negative feelings can interfere with appropriate patient care. If possible, psychiatric consultation should be promptly obtained. Good communication among all involved is essential because these patients are prone to cause confusion and splitting of caretakers. Regular interdisciplinary meetings are helpful.

Psychiatrists should be aware of their own countertransference reaction and help medical clinicians cope with their own reactions to these patients who falsify symptoms but who also have genuine needs. Negative countertransference can lead to therapeutic nihilism, breaches of confidentiality, inappropriate treatment, or denial of care. Countertransference feelings of frustration and anger can more constructively be understood as a projective identification and, thus, a window toward understanding the patient's longstanding feelings. It is essential to preserve good staff communication and to formulate a clear management plan involving the whole team to minimize splitting. Allowing staff to express negative views and emotions at team meetings

while providing empathic support and helpful formulations of the patient can reduce unintended retribution against the patient. The psychiatrist or psychotherapist may also be prone to negative countertransference arising from feelings of being "duped" or manipulated and should seek appropriate supervision.

On an outpatient basis, all medical care should be directed through a single primary care physician through whom all care is coordinated. This *gatekeeper* can minimize unnecessary repetition of tests and treatments. However, it should be remembered that the patient may seek care from clinicians in another system, in another city, or even in another state or country. Appointments with the gatekeeper primary care physician should be regular and frequent and not dependent on medical crises. This fosters object constancy while minimizing the patient's need to induce illness to seek medical attention. As a means of last resort, legal disincentives might be pursued. In Arizona, a woman with factitious disorder was prosecuted for fraudulent procurement of medical, psychiatric, and dental services: she pled guilty. Laws in several states specifically address this type of action.

2. Addressing Psychiatric Issues: Allowing the patient to save face is essential toward establishing a therapeutic alliance and toward preventing the patient from simply taking the factitious illness behavior elsewhere. This is especially true because factitious disorder patients usually have immature personalities or personality disorders that make them especially sensitive to narcissistic injury. Direct or aggressive confrontation is generally not effective. The patient usually responds with anger and denial and may leave the hospital against medical advice, only to perpetuate their factitious illness behavior elsewhere. This kind of confrontation may amplify the patient's psychological need for mastery through deception. Clinicians should keep in mind that confession is not a necessary aspect of management or treatment.

   Several behavioral strategies have been successfully used to facilitate the healing of factitiously produced symptoms in face-saving ways. Eisendrath advocated a *double-bind* technique whereby patients are told that if the symptoms are genuine, then they should improve with the treatment administered. If the symptoms do not improve, then they must be factitious. This technique can be used with nonhealing wounds, factitious paralysis, or psychological factitious symptoms that should improve with medications. For example, factitious patients who chronically induce or worsen wounds can be informed, "This treatment should heal your wound. If it doesn't, we have no choice but to conclude that this is due to a factitious disorder."

   Self-hypnosis and biofeedback can allow a patient to relinquish factitious behavior in a face-saving manner. For example, a patient can be told that under self-hypnosis, blood flow to a wound can be increased

and healing promoted. In this manner, the patient can take control of healing the wound, rather than seek control by worsening symptoms. Positive feedback should be given to patients when their efforts result in healing. Biofeedback can be used in a similar way.

Once factitious disorder is confirmed or evidence is deemed sufficient, the patient should be gently and artfully steered toward ongoing psychiatric treatment in a sympathetic manner acceptable to the patient. For example, the patient can be told that his or her factitious illness behavior is an expression of great emotional need or distress. The clinician can make empathic statements about the neglect, abuse, or trauma that the patient may have undergone. The patient should then be engaged in care in a psychiatric setting in which underlying emotional issues can be addressed and where the risk of morbidity and mortality is substantially less than in a medical setting. Because these patients may crave to be understood and to have their emotional needs recognized, this type of sympathetic confrontation that focuses on addressing the patient's genuine rather than factitious needs can be well accepted by patients.

Psychiatric treatment should first focus on underlying or comorbid psychiatric disorders or on the emotional distress that might have precipitated the factitious illness behavior.

Comorbid mental illnesses must be recognized and treated appropriately. For example, most cases of feigned bereavement are comorbid with major depression. One patient claimed to have feigned bereavement to justify depression. Another falsely claimed to have been recently raped and sought help for posttraumatic stress disorder. Although she was not recently raped, she had suffered longstanding childhood sexual abuse. Feigned mental illnesses may express an emotional truth for the patient, if not a factual truth. Comorbid personality disorders are more common than affective or psychotic disorders. Borderline personality disorder is the most common comorbid diagnosis. Antisocial personality traits are also common, especially in patients exhibiting *pseudologica fantastica* and Munchausen syndrome. Pharmacologic and psychotherapeutic treatments should be employed according to the diagnosis. Substance abuse and dependence may require intensive outpatient or residential inpatient treatment. Other than targeting comorbid psychiatric disorders, there is no standard pharmacologic treatment for factitious disorder.

Once in psychiatric treatment, the greatest challenge is getting the patient to engage and commit to long-term therapy. Brief regularly scheduled contact on a time-contingent rather than a distress-contingent basis is the mainstay of therapy. Consistent long-term psychotherapy is aimed at enabling patients to express their feelings,

gain insight and coping skills, and provide a reliable and supportive outlet for communication.

Relapses of factitious behavior should be expected, just as with substance abuse. Clinicians should not despair at relapses, but instead take them as opportunities to further understand their patient. For example, if a patient has a pattern of relapsing under conditions of new romantic involvements or arguments with authority figures, this provides insight into the patient's vulnerabilities, and may be addressed in psychotherapy.

3. Managing Legal and Ethical Issues: Early involvement of risk management experts and bioethicists is prudent in the care of factitious disorder patients. Patients' right to privacy and implicit injunctions against unwarranted searches and seizures are major issues. The once-common practice of alerting all emergency rooms about factitious disorder patients is no longer routinely utilized because of heightened sensitivity about confidentiality rights. Nevertheless, this kind of widespread alerting can be done with the consent of the patient. For example, once a patient acknowledges having factitious disorder, the patient can be asked, "I'd like to minimize harm to you. Can I have your permission to alert emergency rooms and some doctors in the community?"

   Similarly, the patient's right to confidentiality should be respected when gathering collateral information in nonemergent situations. Verbal permission or, preferably, signed releases of information should be obtained prior to contacting collateral sources, except in emergency situations. Clinicians should be careful about revealing information to the patient's employers, friends, or family. The diagnosis of factitious disorder must never be revealed, even to spouses or significant others, without a patient's explicit permission, as this will likely lead to accusations of confidentiality breach. Searching a patient's hospital room or personal belongings for illness-inducing means may facilitate diagnosis and lessen morbidity, but this should only be done after consulting with risk management or other legal counsel because it goes beyond the contractual and fiduciary bounds of the doctor–patient relationship. Covert surveillance should only be undertaken after careful legal consultation. It should also be kept in mind that simply because an action is legal, it is not necessarily ethical. Therefore the involvement of bioethicists can be helpful.

Factitious Disorder Imposed on Another. Protection of the victim is the first priority. Active pursuit of the diagnosis and then prompt intervention is essential for minimizing the risk of morbidity and mortality to the victim, who is usually a vulnerable child. Because young children are the most commonly victimized, Table 19–7 will prioritize the management of

pediatric factitious disorder imposed on another.

Table 19–7.
Pediatric Factitious Disorder Imposed on Another—Basic Principles of Management

Make sure the child is safe.
Make sure the child's future safety is also assured.
Allow treatment to occur in the least restrictive setting possible.
A pediatrician should serve as "gatekeeper" for medical care utilization.
All other physicians should coordinate care with the gatekeeper.
Child protective services should be informed whenever a child is harmed.
Family psychotherapy and/or individual psychotherapy should be instituted for the perpetrating parent and the child.
Health insurance companies, school officials, and other nonmedical sources should be asked to report possible medical abuse to the physician gatekeeper. Permission of a parent or of child protective services must first be obtained.
The possibility should be considered of admitting the child to an inpatient or partial hospital setting to facilitate diagnostic monitoring of symptoms and to institute a treatment plan.
The child may require placement in another family. The perpetrating parent may need to be removed from the child through criminal prosecution and incarceration.

When gathering clinical information and evidence, clinicians should keep in mind that if the mother feels that she is suspected or is losing control of the situation, she is likely to remove the child and to seek care elsewhere. When evidence of factitious disorder imposed on another is felt to be sufficient, both parents should be informed together. In this manner, a nonperpetrating parent, usually the father, might be enlisted as an ally.

Child protective services should be informed and a legal hold should be instituted if clinicians feel that parents might flee with the child. If harm has been done to the child, reporting of factitious disorder imposed on another to child abuse protection authorities is mandatory in all 50 states.

In Rosenberg's review of 117 cases, 20 percent of children who died of factitious disorder imposed on another were sent home with their parents after parents were confronted with the diagnosis. Given the high rate of mortality and morbidity, children should be removed from the parents until further assessment indicates that it is safe for the family to be reunited, in most cases. Child protective authorities usually take responsibility for this action. When this is done, there is a high likelihood that the mother will need intervention for suicidal ideation. Investigation into possible abuse of siblings should also be initiated.

After separation, the child must be treated for ongoing medical problems, as well as for psychological problems. For many children, longstanding refractory medical illnesses will resolve once they are separated from perpetrators. Many such children suffer from posttraumatic stress disorder and should be treated. As is typical of those with

posttraumatic stress disorder, many victims avoid medical care as adults. On the other hand, others develop factitious disorder. By adolescence, many victims of factitious disorder imposed on another collude with parents in perpetrating deception.

The perpetrator, usually the mother, should also undergo treatment to diagnose and address underlying psychiatric conditions. The perpetrator should be evaluated to determine whether it is likely that he or she will ever achieve "good enough" parent status and whether reunification with their child is a realistic goal. Those who actively induce illness in their children are less likely to ever be adequate parents. If reunification is unrealistic, then psychotherapy should focus on the parent's underlying psychopathology, as well as their loss of a child. If reunification is the goal, then treatment should address emotional maturation, an ability to put the child's needs before their own, relinquishing the defense of denial, and learning to seek help and to express themselves in more appropriate ways. The parent should also work on boundary issues aimed at distinguishing their needs from their child's needs. If reunification occurs, careful monitoring should follow. Child protective services should be involved. A primary pediatrician should act as a gatekeeper who monitors and approves all medical care.

Although the literature on factitious disorder imposed on another focuses on the most severe cases, mild cases do occur. Meadow advocated that in these mild cases, parents should be supported and prevented from putting their children through needless investigations and treatments. Psychotherapy should be undertaken. The prognosis in these cases may not be so grim. Factors associated with positive outcomes according to Bools et al.'s study included: (1) continuous positive input from spouse or grandparents; (2) successful short-term foster care before returning to live with the perpetrating caregiver; (3) the perpetrator's long-term relationship with a social worker; (4) successful remarriage for the perpetrator; (5) early adoption of the victim; and (6) long-term foster care placement.

There is little literature on the treatment of factitious disorder imposed on another between adults. In these cases, removal of the victim through legal means is usually not an option if the victim is a legally independent adult, even if he or she might be emotionally dependent on and enmeshed with the perpetrator. Unless the victim has a legally dependent status, adult victims must initiate separation. Psychiatric comorbidity in the victim, if severe enough, may render the individual incapacitated and requiring hospitalization or even guardianship. Adult protective services and/or agencies overseeing the protection of developmentally disabled individuals must be notified when adult factitious disorder imposed on another is suspected.

In cases in which health care providers are suspected of perpetrating

illness among their patients, legal prosecution should be instituted. Simply removing the provider is not adequate because the perpetrator is likely simply to seek another job, where more patients will be at risk.

Future Directions

There is a paucity of research on factitious disorders imposed on self and others, and studies are needed to further elucidate risk factors for developing factitious disorder, best practices for management, and outcomes of specific treatments. Recent changes in the definition of factitious disorder may result in shifting patterns of its recognition and diagnosis, and implications of these changes for clinicians and researchers will need to be explored. With the understanding that deception in the medical setting is commonplace, research into the underpinnings of this behavior may help to identify means of reducing episodes of deception both in patients with factitious disorders and in the general public.

References

Asher R. Munchausen's syndrome. *Lancet*. 1951;1:339.

*Bass C, Halligan P. Factitious disorders and malingering: challenges for clinical assessment and management. *Lancet*. 2014;383:1422–1432.

Bass C, Jones D. Psychopathology of perpetrators of fabricated or induced illness in children: case series. *BJP*. 2011;199:113–118.

Bools CN, Neale BA, Meadow SR. Follow up of victims of fabricated illness (Munchausen syndrome by proxy). *Arch Dis Child*. 1993;69:625–630.

*Davis P, McClure RJ, Rolfe K, et al. Procedures, placement and risks of further abuse after Munchausen syndrome by proxy, non-accidental poisoning, and non-accidental suffocation. *Arch Dis Child*. 1998;78:217–221.

Eisendrath SJ. Factitious physical disorders: treatment without confrontation. *Psychosomatics*. 1989;30:383.

Eisendrath SJ, McNiel DE. Factitious disorders in civil litigation: twenty cases illustrating the spectrum of abnormal illness-affirming behavior. *J Am Acad Psychiatry Law*. 2002;30:391.

Feldman MD, Eisendrath SJ, eds. *The Spectrum of Factitious Disorders*. Washington, DC: American Psychiatric Press; 1997.

*Flaherty EG, MacMillan HL, and the American Academy of Pediatrics Committee on Child Abuse and Neglect. Caregiver-fabricated illness in a child: a manifestation of child maltreatment. *Pediatrics*. 2013;132:590.

Fliege H, Lee JR, Grimm A, Fydrich T, Klapp BF. Axis I comorbidity and psychopathologic correlates of autodestructive syndromes. *Comprehensive Psychiatry*. 2009;50:327–334.

Ford CF. *The Somatizing Disorders: Illness as a Way of Life*. New York: Elsevier; 1983.

Gavin H. *On the Feigned and Factitious Diseases of Soldiers and Seamen, on the Means Used to Simulate or Produce Them, and on the Best Modes of Discovering Imposters*. Edinburgh: University Press; 1838.

Gelenberg AJ. Munchausen's syndrome with a psychiatric presentation. *Dis Nerv Syst*.

1977;38:378.

Hamilton JC, Eger M, Razzak S, Feldman MD, Hallmark N, Cheek S. Somatoform, factitious, and related diagnoses in the National Hospital Discharge Survey: addressing the proposed DSM-5 revision. *Psychosomatics*. 2013;54:142–148.

*Kenedi CA, Shirey KG, Hoffa M, et al. Laboratory diagnosis of factitious disorder: a systematic review of tools useful in the diagnosis of Munchausen's syndrome. *NZMJ*. 2011;124(1342):66–81.

King BH, Ford CV. Pseudologia fantastica. *Acta Psychiatr Scand*. 1988;77:1–6.

Kinns H, Housley D, Freedman DB. Munchausen syndrome and factitious disorder: the role of the laboratory in its detection and diagnosis. *Ann Clin Biochem*. 2013;50:194.

Krahn LE, Bostwick JM, Stonnington CM. Looking toward DSM-V: should factitious disorder become a subtype of somatoform disorder? *Psychosomatics*. 2008;49:277–282.

Krahn LE, Li HZ, O'Connor MK. Patients who strive to be ill: factitious disorder with physical symptoms. *Am J Psychiatry*. 2003;160:1163–1168.

Lawrie SM, Goodwin G, Masterton G. Munchausen's syndrome and organic brain disorder. *BJP*. 1993;162:545–549.

McClure RJ, Davis PM, Meadow SR, Silbert JR. Epidemiology of Munchausen syndrome by proxy, non-accidental poisoning, and non-accidental suffocation. *Arch Dis Child*. 1996;75:57–61.

McCullumsmith CB, Ford CV. Simulated illness: the factitious disorders and malingering. *Psychiatr Clin N Am*. 2011;34:621–641.

Meadow R. Management of Munchausen syndrome by proxy. *Arch Dis Child*. 1985;60:385.

Meadow R. Unnatural sudden infant death syndrome. *Arch Dis Child*. 1999;80:7.

Mountz JM, Parker PE, Liu HG, Bentley TW, Lill DW, Deutsch G. Tc-99m HMPAO brain SPECT scanning in Munchausen syndrome. *J Psychiatry Neurosci*. 1996;21(1):49–52.

Pankratz L, Lezak MD. Cerebral dysfunction in the Munchausen syndrome. *Hillside J Clin Psychiatry*. 1987;9(2):195–206.

Poole CJ. Illness deception and work: incidence, manifestations and detection. *Occupational Med*. 2010;60:127–132.

Reich P, Gottfried LA. Factitious disorders in a teaching hospital. *Ann Intern Med*. 1983;99:240.

Rosenberg DA. Web of deceit: A literature review of Munchausen syndrome by proxy. *Child Abuse Neglect*. 1987;11:547.

Savino AC, Fordtran JS. Factitious disease: clinical lessons from case studies at Baylor University Medical Center. *Proc Bayl Univ Med Cent*. 2006;19:195–208.

*Spence SA, Kaylor-Hughes CJ. Looking for truth and finding lies: the prospects for a nascent neuroimaging of deception. *Neurocase*. 2008;14:68–81.

Spiro HR. Chronic factitious illness: Munchausen's syndrome. *Arch Gen Psychiatry*. 1968;18:569.

Stirling J Jr, American Academy of Pediatrics Committee on Child Abuse and Neglect. Beyond Munchausen syndrome by proxy: identification and treatment of child abuse in a medical setting. *Pediatrics*. 2007;119:1026–1030.

Stone J, Zeman A, Simonotto E, et al. fMRI in patients with motor conversion symptoms and controls with simulated weakness. *Psychosomatic Med*. 2007;69:961–969.

Sutherland AJ, Rodin GM. Factitious disorders in a general hospital setting: clinical features and a review of the literature. *Psychosomatics*. 1990;31:392.

Wallach J. Laboratory diagnosis of factitious disorders. *Arch Intern Med*. 1994;154:1690.

20

Dissociative Disorders

RICHARD J. LOEWENSTEIN, M.D., PAUL FREWEN, PH.D., AND ROBERTO LEWIS-FERNáNDEZ, M.D.

INTRODUCTION

In the *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition (DSM-5), based on recent research, the DSM-5 Anxiety, Obsessive-Compulsive, Posttraumatic, and Dissociative Disorders Work Group expanded the definition of dissociation, and made important changes to some of the diagnostic criteria for the dissociative disorders. The DSM-IV-TR described the "essential features," and the ICD-10 described the "common theme" of dissociation and dissociative disorders. DSM-5 expanded these constructs. Dissociation is now conceptualized as a disruption, interruption, and/or discontinuity of the normal, subjective integration of potentially any aspect of experience and cognition, including behavior, memory, identity, consciousness, emotion, perception, body representation, and motor control. In more acute pathological dissociative presentations, dissociation is primarily related to traumatic and/or overwhelming experiences. In lifelong dissociative adaptations such as dissociative identity disorder, dissociative symptoms can occur both in traumatic contexts and for nontraumatic events in everyday life. Overall, in dissociation, aspects of psychobiological functioning that should be associated, coordinated, and/or linked are not. This does not occur randomly, but due to a potentially explicable psychobiological process or processes. In the literature, dissociation is described as both a dimensional, adaptational process, and as a set of categorically distinct psychopathologies. Generally, contributions to the literature on dissociation do not make this crucial distinction. This has led to conceptual confusion, and, at worst, almost a Tower of Babel effect between different schools of thought writing about dissociation (see the sections on Nosology and Etiology).

The DSM-5 dissociative disorders are dissociative identity disorder; dissociative amnesia; depersonalization/derealization disorder; other specified dissociative disorders; and unspecified dissociative disorder. In the following sections, many issues common to all the dissociative disorders are discussed together, and then each of the dissociative disorders will be reviewed in greater depth.

HISTORY

The history of dissociation and the dissociative disorders traverses the modern history of psychiatry from its very beginning to the present time. Embedded in this history are some of the most complex, important, and, often, controversial issues in psychiatry, and some of the most important figures in the early development the field of psychiatry, like Pierre Janet and Sigmund Freud. Also, from the beginning of the study of dissociative disorders, controversies have also related to larger social, political, cultural, and cross-cultural issues. In addition, even since the early 19th century these disorders have generated media attention and led to awareness of the general public about these disorders, and debates about them. Controversies about dissociation and dissociative disorders will be addressed in later sections of this chapter.

The modern study of dissociation begins at the end of the 18th century with the shift of interest in phenomena of trance and possession from the religious to the medical realm, at least for the individuals in the scientific, medical, and increasingly secularized parts of the culture. Within the Western world, to this day, dissociative trance and possession experiences have continued to be normative in some religious groups and subcultures. In addition from the 19th century until the present time, various forms of the paranormal and the supernatural, including spiritism, mediumship, astrology, and occultism—often associated with dissociative types of experiences—have continued to evoke interest among many members of the general public.

Paracelsus, writing in 1646, is credited with the first medical report of an individual with alternating selves. In 1791, Eberhardt Gmelin described a German woman who alternately exchanged her peasant personality for that of an aristocratic French lady, each amnesic for the other's existence. By the early 19th century, such cases, with diagnoses of dual, double, or duplex consciousness, were being regularly reported on both sides of the Atlantic. In 1811, Benjamin Rush, considered the father of American psychiatry, included a classic example in his medical school lectures. Rush proposed that dual personality reflected a functional disconnection between the two cerebral hemispheres. In the same year, the case of Mary Reynolds, who became the American archetype of dissociative identity disorder for the remainder of the century, was first published.

In Europe, disciples of Franz Anton Mesmer developed the first systematic descriptions of what is now called *hypnosis*. These thinkers developed an interest in *magnetic diseases* (from Mesmer's theories of animal magnetism): Amnesia, fugue, somnambulistic states, and alternating or multiple personality that could be treated with *artificial somnambulism,* that is, hypnosis. These *magnetizers* were mostly outside the mainstream of European academic medicine, and their works are not widely known. Among them, Antoine Despine, a French family doctor,

wrote systematic case studies and reports about several series of multiple personality patients and their treatment, including child and adolescent cases. Most of these 19th century patients also had extensive conversion and somatoform symptoms as part of the clinical presentation.

Psychiatric historians have charted the shift in interest from somnambulism to multiple personality to hysteria and back over the course of the 19th century, with *hysteria* finally seen as a unifying concept for all these conditions. Many 19th century debates about dissociation prefigured those that continue into the 21st century. Paul Briquet, who wrote a famous study of hysteria in the mid-19th century, disputed the prevailing notion that the disorder was caused by female sexual frustration. He viewed hysteria as a disorder of the brain, not the uterus. Briquet documented that traumatic, overwhelming, and grief-engendering experiences were strongly associated with hysteria. Mid-19th century France also witnessed the rise of popular movements such as spiritism, which attempted to study trance states and other dissociative activity (e.g., automatic writing) as forms of spirit communication. The work of Allan Kardec (pseudonym for Hippolyte de Rivail) was disseminated widely, sparking healing movements targeting —and mediated by—spirit possession, particularly across Latin America.

The famous 19th century French neurologist Jean-Martin Charcot synthesized the teachings of the magnetizers with those of the more accepted medical and psychiatric establishment, including Briquet's etiological theories. Charcot influenced major figures in psychiatry and neurology, such as Sigmund Freud, Pierre Janet, Georges Gilles de la Tourette, and Joseph Babinski. As is well known, he gave public demonstrations of female hysterical patients and hypnosis at his Paris hospital La Salpetrière, where the neuropsychiatric wards primarily housed indigent women. Charcot's weekly demonstrations of hysterical symptoms were open to the public and were considered a major social event in Parisian society of the time. Hysterical individuals were depicted in numerous highly popular French plays, novels, short stories, and poetry; some of Charcot's patients became celebrities, as did Charcot himself.

Through these demonstrations, and frequent interaction with Charcot and his associates, hysterical patients were readily prompted in various ways to perform spectacularly for him, since gaining favor with Charcot often led to better treatment by the staff. Despite the relatively poor conditions for patients at La Salpetrière, these were often a considerable improvement over the poverty and misery of the women's outside lives. Some were magnetized (i.e., hypnotized) on the wards or encouraged by sycophantic followers of Charcot to act as dramatically as possible for the great man.

After his death, two of Charcot's most important successors, Babinski and Janet, took divergent views of hysteria. Babinski took the radically nihilistic view that all hysteria was caused by *suggestion* and could be

removed by persuasion or countersuggestion. He coined the term *pithiatism* as a substitute for hysteria. The Nancy school of hypnosis, led by Hippolyte Bernheim, believed that hypnosis was not necessarily a sign of pathology, as Charcot believed, but an effect of suggestion. Babinski and Bernheim argued that the symptoms of Charcot's patients were artifacts of suggestion and contagion, not of bona fide disorders. This view continues today, most prominently in the thinking of Paul McHugh, a vehement critic of the current trauma-based theories of dissociation.

On the other hand, Janet, who remained at La Salpetrière after Charcot's death, emphasized the traumatic origins of hysterical symptoms. Janet's research and clinical theory, in particular the emphasis on the role of traumatic antecedents of dissociation, are widely regarded as the foundation for modern views of dissociation, although, during his lifetime, Janet never developed a school of followers to promulgate his ideas.

During Charcot's time, stage hypnotism began—as a pseudomedical performance with a "doctor" and his "hysterical" subject, including subjects drawn from the audience. Many of the symptoms of Charcot's patients are duplicated today in stage hypnosis demonstrations. Modern hypnosis research has been concerned in a major way with separating what is essential about the hypnotic state, and the hypnotic capacity of the subject, with "demand characteristics," sometimes called sociocognitive aspects of hypnosis. The latter are the expectations subjects bring about how they are to behave while "being hypnotized"—that are different from the core features of the hypnotic state itself, and in large part account for the behavior of individuals in stage hypnotism demonstrations (see sections on Iatrogenic and Sociocognitive Model of Dissociation and Culture, Dissociative Disorders, and Sociocognitive Phenomena; also see Chapter 33.4 on hypnosis).

Trauma and the Patients at La Salpetrière. Recent scholarship has uncovered the medical records of the female patients at La Salpetrière during the 19th and early 20th century, including those studied by Charcot, Babinski, Janet, et al. in their research on hysteria—including some of Charcot's most famous demonstration cases. Supporting Janet's views, the hospital records document histories of multiple forms of severe trauma, loss, and social dislocation among these women.

These women were young, generally between 15 and 30 years old, and came from backgrounds of extreme poverty, illiteracy, violence, and traumatic loss. Many had brutal, incestuously abusive alcoholic fathers or stepfathers. Often, the women were the *de facto* caretakers for their younger siblings, and witnessed the traumatic deaths of many of these children from tuberculosis, other infections, fevers, convulsions, etc., as well as witnessing the deaths of other family members. Some were illegitimate and shuttled between violent homes and religious institutions

characterized by austerity and harsh discipline.

In addition, the later life histories of the hysterical women, reported in the hospital records, detailed rape, beatings, sexual exploitation by employers, and severe physical injuries due to trauma, in addition to earlier life traumas. Also, many had experienced wartime trauma during the 1870 Franco-Prussian war. Many came from working class neighborhoods directly involved in the street warfare in 1871, when the Paris Commune was crushed, a major fire devastated Paris, and thousands of communards were summarily executed in the streets of Paris. In addition, many of the "treatments" of the time were likely more traumatic than therapeutic: ovarian compression (often for hours) with a specific "ovarian compressor," applied electric currents, and cervical cauterization with a hot iron. Drug treatment with opium, ether, chloroform, and/or amyl nitrate was also very common. Finally, based on the records, a number of these women were impregnated by Charcot's students. Nonetheless, the hospital was often a more benign environment for these women than their lives outside La Salpetrière. They were poor, massively traumatized, often uneducated, and, outside the hospital, were only allotted the most despised jobs in Parisian society. Case History 1 describes the history of profound trauma suffered by Blanche Wittmann, *la reine des hysteriques* (the queen of hysterics), as well as her later life. She is also depicted in the classic painting: A Clinical Lesson at the Salpêtrière ("*Une leçon clinique à la Salpêtrière*") by Pierre Aristide André Brouillet (1857–1914) (see Fig. 20–1).

Marie (Blanche) Witmann, *la reine des hysteriques* (the queen of hysterics), had a history of profound medical, emotional, physical, and sexual trauma in her early life. This included severe illnesses and infections before the age of 2; a violent father who was committed for life to a psychiatric hospital when she was 12; losses of five of her eight siblings; and a mother who was "nervous" and died when she was 15. She was apprenticed to a furrier who repeatedly physically and sexually assaulted her. A young man she loved died suddenly when she was 15. She had behavioral problems from early childhood, and could not remain in school; thus, she was uneducated, and only could be employed in the most menial and despised jobs. She suffered from frequent hysterical attacks and was hospitalized as well as working as a "ward girl," someone who did the most reviled and menial tasks in the hospital. She was 18 when admitted to La Salpetrière in 1877, and assumed the name "Blanche" at that time. She became famous in Paris for her symptoms during Charcot's public demonstrations of hypnosis and hysteria. However, in 1888, treated by Jules Janet, he discovered a classic "dual personality" with a second self-state completely free of all the hysterical symptoms and reactions. This identity was grounded, had a positive outlook, and reported co-consciousness (awareness of the thoughts, feelings, and actions) with the hysterical, depressed state. He attempted to increase the dominance of this apparently healthier state, with some success. Subsequent to Charcot's death in 1893, she shifted states again. She denied her whole prior history, and went to work in the radiology laboratory of La Salpetrière.

There she experienced severe radiation burns, and had stoically undergone repeated amputations of her fingers, hand, and arm bilaterally until she died in 1913, of radiation-induced cancer.

FIGURE 20–1. A Clinical Lesson at the Salpêtrière ("*Une leçon clinique à la Salpêtrière*") by Pierre Aristide André Brouillet (1857–1914). Center: Charcot, Joseph Babinski, Blanche Wittmann. Others include Theodule Ribot (far left), Georges Gilles de la Tourette (front in the apron), and Paul Richer (just to the right of Charcot's arm).

History of Dissociation in the 19th and Early 20th Centuries

At the end of the 19th century, Janet in France, Morton Prince and William James in the United States, and others across Europe were engaged in a lively transatlantic discussion about possible psychological and neurological mechanisms underlying cases of multiple personality, amnesia, and fugue. The scientific investigation of dissociation dates to the experiments of Prince in 1908, using a crude polygraph to measure galvanic skin resistance (GSR) across the personality states of a dissociative identity disorder patient—a physiological measure that remains of interest today. Prince reported seeing differential GSR reactivity to words that were emotionally laden for one personality state but not for another. Over the next century, several dozen psychophysiological and cognitive performance studies were added to the literature. Many of these reported findings suggesting differential responses across dissociative identity disorder personality states. Medical models of the time-invoked phenomena such as sleep and dreams, hypnosis and somnambulism, epilepsy, and disconnections between the cerebral hemispheres to explain multiplicity.

As the 20th century began, however, interest in dissociation waned, subsumed under the somatoform hysteria construct. Alternative theories of hysteria, like those of psychoanalysis, began their ascendance. However, clinical interest in dissociative phenomena has reoccurred in every war since the turn of the century with the observation of amnesia, fugues, and conversion symptoms in traumatized soldiers. Subsequently, these observations tended to be forgotten by the psychiatric community, as the

wars themselves faded from memory.

For example, Babinski's ideas about hysteria, renamed "pithiatism"—a disorder defined by symptoms created and removed by suggestion—had considerable influence on European psychiatrists until the onset of World War I. At that time, with the development of the concept of *shell shock,* psychiatrists systematically reported that virtually all of the symptoms of dissociation, amnesia, automatism, and hysteria could be found among battle-traumatized soldiers, plainly contradicting Babinski's ideas and supporting those of Janet.

Janet and Freud

Freud and Joseph Breuer's revered *Studies in Hysteria* (1893) referenced Janet's work, used similar notions of "hypnoid states," fixed ideas, and traumatic etiology, and introduced the cathartic method of cure for hysteria. Anna O, the archetypal Freudian "hysteric," is clearly described as having a dual personality and a plethora of dissociative symptoms, such as complex amnesias for current and historical experience; spontaneous age regressions and trances; depersonalization; fluctuations in handwriting, handedness, and language; and alternating states of consciousness (see section on Diagnosis of Dissociative Identity Disorder).

Freud rejected his early dissociationist ideas, including the notion that hysteria and related conditions were caused by childhood sexual and/or physical violence (the "*seduction theory*"). Later psychoanalysts asserted that Janet dishonestly claimed priority for these ideas related to hysteria, even though Janet's work clearly antedated Freud's, and Janet was referenced specifically in the *Studies in Hysteria*. Janet, for his part, was uncharacteristically publicly angry at this failure to acknowledge his work. Throughout his career, he remained critical of Freud and psychoanalytic ideas.

Despite the Janetian ideas found in early Freudian works on hysteria, until recently, psychoanalytic theory has had difficulty accounting for the phenomenon of dissociation. Freud's disciple Sandor Ferenczi wrote papers that described the sequelae of childhood sexual abuse and resulted in profound "splitting of the personality," analogous to modern descriptions of dissociative identity disorder. Ferenczi's insistence on the importance of actual traumas in psychopathology, as well as other issues, led to his expulsion from institutional psychoanalysis. British psychoanalysts from the Object Relations School, Harry Guntrip and Ronald Fairbairn, both had experience with sexually abused children. They emphasized the "splitting" of the ego as the result of these early traumas. Thus, the psychoanalytic term *splitting* frequently has been equated with dissociation, notwithstanding the criticism among some psychoanalysts of the reification of the splitting metaphor, and the difficulty aligning the range of dissociative phenomena with the splitting concept.

David Rappaport, Merton Gill, Margaret Brenman, Charles Fischer, and Elisabeth Geleerd, among others, wrote important psychoanalytic papers and books on fugue, amnesia, hypnosis, and dissociation in the 1940s and 1950s, with Rappaport, following Janet, presciently enumerating the conceptual and phenomenological differences between repression and dissociation. However, psychoanalytic thinkers have largely ignored this work, as it was hard to align with psychoanalytic theory of the time, including the notions that incest reports were "fantasies," and that hypnosis was anathema in analytic psychotherapy, since Freud rejected this technique early in his work. Recently, many psychoanalysts have recognized the reality of childhood and adult trauma, and have attempted to reintegrate dissociation into psychoanalytic thinking. However, this has been problematic because of the persistence of the conceptual legacy of splitting, and the difficulty of fitting the full spectrum of dissociative psychopathology into this and other psychoanalytic constructs. Psychoanalysts of the relational school, such as Phillip Bromberg, have broadened the dissociation concept so widely that it is seen as characteristic of human psychological processes in general, and of the experience of the psychoanalyst in all psychotherapy, not just therapy of individuals with dissociative disorders.

It is unfortunate that disputes over primacy of ideas led to a schism between Janet and the Freudians. Modern conceptualization of dissociative disorders often involves a synthesis of Janetian ideas and Freudian ones. For example, current theories of treatment include the Janetian notions of systematic work with traumatic memories, using adjunctive hypnotic and imagery techniques, and work with posttraumatic cognitive distortions, using cognitive and behavioral psychotherapy, psychoeducation, and building life skills. The psychoanalytic concepts of transference, countertransference, object relations, projective identification, intrapsychic defense, and conflict, among others, are also essential to understanding modern therapy of dissociative patients.

Based on his work with the severely ill patients of La Salpetrière, Janet viewed hysterics as liable to psychic disintegration with stress. Freud, working with well-to-do, bourgeois patients in his private practice, viewed hysterics as psychologically strong, attempting to survive extreme conflicts leading to symptoms. Both these ideas are relevant to modern conceptualization of dissociative disorders and their treatment. Research has shown that dissociative disorder patients are often highly symptomatic and impaired. However, studies have also shown that the early-life dissociation may be a resilience factor, particularly in the most severe, early-life onset dissociative disorders, like dissociative identity disorder. Early psychological "compartmentalization" of traumatic experiences appears to allow for more normal development in and preservation of cognitive and creative abilities, along with a capacity for humor, hope,

human attachment, and insight, despite reported histories of profound, repeated early life trauma.

Janet continued to be active in psychology and psychiatry until his death in 1947, writing on many different clinical, theoretical, and philosophical subjects. Janet never developed a formal school of psychology nor a group of disciples committed to his theories. He was deeply concerned with the protection of patient confidentiality, perhaps in part as a reaction to the excesses of the Charcot era. Accordingly, he had his library of detailed case histories, including those of the most famous classic hysterics, destroyed at his death. Because of this, and because of the primacy of the psychoanalytic theories in the latter part of the 20th century, his work was largely neglected until the recent resurgence of interest in dissociation.

Janet's contributions to modern dissociation theory include (1) recognition of the causal role of trauma in dissociation in a stress-diathesis model; (2) the taxonometric approach—Janet believed that dissociation was not on a continuum from normal psychological experience; (3) recognition that the underlying mental mechanisms of somatoform and dissociative disorders were similar; (4) the role of *fixed ideas* in dissociative disorders and attempts to find the most fundamental problematic ideas that influence the patient's behavior; and (5) treatment using a multimodal approach with hypnotic and trauma-focused methods combined with cognitive therapy, behavioral therapy, and life-skills building.

Dissociation during and after World War II

Clinicians treating battlefield casualties in World Wars I and II described dissociative symptoms, such as amnesia, fugue, automatisms, and somatoform symptoms, in traumatized soldiers as part of *traumatic war neurosis,* as it came to be known in World War II. The extent of amnesia described in psychiatric battlefield casualties in World War II ranged from relatively brief periods of time to complete amnesia for life history and identity (generalized amnesia), as well as fugue episodes.

Treatment of these World War II amnesia cases included hypnosis, narcosynthesis with a mixture of hypnosis and sodium amobarbital (Amytal), individual and group psychotherapy, or a combination of these. Attempts were made to treat these soldiers on the front lines and return them rapidly to combat. Detailed case descriptions of dissociative amnesia and fugue can be found in clinical studies from that time, as well as studies of large cohorts of patients. Researchers have reported that soldiers from the Korean conflict and the Vietnam war also reported amnesia as part of the posttraumatic syndromes related to combat experiences, and were treated with similar modalities.

Despite the extraordinary interest in multiple personality throughout the 19th century, and the observations of these military psychiatrists, the

study of dissociation mostly waned after the beginning of the 20th century. Authorities suggest that this was related to a variety of factors, including the dominance of Freudian paradigms of hysteria; the disrepute into which hypnosis fell at this time; the rise of the Bleulerian concept of schizophrenia, which may have subsumed dissociative patients (see section on Other Dissociative Disorders); and the loss of interest in the works of Janet, Prince, and others who had been crucial in the development of models of dissociation.

Nonetheless, throughout much of the 20th century, the psychiatric, psychological, and psychoanalytic literature described individual cases or small case series of individuals with dual or multiple personality and/or generalized dissociative amnesia, with and without dissociative fugue. The amnesia cases were often brought to hospitals without recall of their past or personal identity. Despite this, conventional wisdom held that these disorders were extraordinarily rare, especially multiple personality; thus individual case studies, such as the famous *Three Faces of Eve,* were not only published in the professional literature, but entered the popular media as well. Nonetheless, periodic, systematic reviews of the literature continued to support the validity of the diagnostic construct of dissociative identity disorder.

History of Depersonalization/Derealization Disorder. First described by Maurice Krishnaber in 1872, depersonalization was formally named in 1898 by Ludovic Dugas, who sought to convey "the feeling of loss of ego." Freud, Janet, Eugen Bleuler, and other 19th century authorities reported patients with symptoms of depersonalization and derealization. Freud himself described his own symptoms of depersonalization and derealization in 1904 as part of his self-analysis, when he finally was able to visit the Acropolis. A monograph by Paul Schilder in 1914 is regarded as a turning point in psychiatric interest. Classic studies by Wilhelm Mayer-Gross in 1939, Harold Shorvonn in 1946, and Jerome Saperstein in 1949 furthered the delineation of a syndrome of chronic depersonalization. Brian Ackner enumerated the essential features of current diagnostic definitions in 1954. These include "(1) the feeling of unreality or strangeness apropos the self; (2) the retention of insight and lack of delusional elaboration; (3) the lack of affective response ('numbness') except for the discomfort regarding depersonalization; and (4) the unpleasant property that may vary in intensity inversely with the subject's familiarity with the phenomenon."

Later 20th Century History

The interest in trauma and dissociation began anew in the 1970s and 1980s. Authorities cite several social and cultural factors leading to this revival. These include the return of the Vietnam veterans and the systematic academic study of their psychiatric problems; the recognition of

the prevalence of childhood physical and sexual abuse with its associated mandated reporting; systematic research on child abuse and family violence; the rise of feminism with its critique of psychological theories that ascribed sexual abuse reports to fantasy; increased academic rigor in theories about hypnosis, such as Ernest Hilgard's *neodissociation* theory; and greater acceptance of academic hypnosis research; popular interest in multiple personality owing to works such as *Sybil;* and the publication of DSM-III (1980), with its promulgation of diagnostic criteria for the dissociative disorders and posttraumatic stress disorder (PTSD).

Cornelia Wilbur. The contemporary era in the study of multiple personality disorder began with the work of Arnold Ludwig and colleagues, including Cornelia Wilbur, at the University of Kentucky during the 1970s, primarily studying differential findings in the psychobiology of dissociative identity disorder self-states. Cornelia Wilbur, widely identified with the case of *Sybil,* was influential in describing the clinical features of the modern construct of dissociative identity disorder and identifying the role of childhood trauma as a major factor in the etiology of the disorder. Wilbur, a psychoanalyst, described a number of cases of dissociative identity disorder in the mid-20th century, all with more complex self-state systems than the classical dual and duplex personality cases, and all with histories of extreme early life maltreatment, including incest. However, her writings were rejected by psychoanalytic journals, primarily due to the prevailing *zeitgeist* that multiple personality and incest were one-in-a-million conditions, and that female reports of childhood sexual abuse were virtually all Oedipal fantasies. Finally, in frustration, Wilbur published the *Sybil* case in the popular press.

DSM-III and Beyond

Since the publication of the DSM-III in 1980, a large body of systematic research has developed about the dissociative disorders, particularly dissociative identity disorder, dissociative amnesia, and depersonalization disorder. Many academic studies have also investigated dissociation as a quantifiable trait in clinical and nonclinical samples. Studies have also looked at the prevalence and phenomenology of amnesia, particularly as it relates to the question of delayed recall for traumatic experiences.

Beginning in the early 1980s, Frank W. Putnam, Jr. at the National Institute of Mental Health (NIMH) was one of the major figures to develop the scientific study of dissociative disorders and dissociation as a dimensional construct. He undertook the first modern, systematic epidemiological, phenomenological, and psychophysiological studies of dissociative identity disorder, including attempts to repeat the classical studies of reportedly compartmentalized memory and differing psychophysiological responses of the self-states. Among his major findings

were the significantly greater variability among dissociative identity self-states compared with simulating controls on a variety of psychophysiological variables such as event-related potentials (ERPs), 16-lead EEG, and autonomic functioning. He showed that the self-states did not exhibit the kind of test–retest similarity on these measures found in simulating control subjects. This was due to the shifting of states and variably recurrent overlap and interference among states during testing.

This led him to posit that dissociative identity disorder naturalistically existed more as "the state of multiple states," with state shifting leading to overlap, intrusion, and interference among personality states, described as constructed around specific emotions, a sense of self, and a group of state-dependent memories. This was very different from the classical understanding of the disorder as characterized by several distinct, dramatically switching identities that were highly defined, complex, and invariant when studied from one time to the next.

Putnam and his colleague Eve B. Carlson developed the *Dissociative Experiences Scale* (DES), the first screening inventory for trait dissociation, and measured *dissociativity* in many diagnostic groups compared to normal controls, including individuals with PTSD, affective disorders, psychotic disorders, anxiety disorders, personality disorders, and dissociative disorders, among others. The DES has been translated into more than 40 languages, and studies across cultures reveal strong similarities for Western and non-Western samples, attesting to the universality of dissociation. Putnam also contributed major works on diagnosis and treatment of dissociative identity disorder, and was among the first to systematically study dissociation in children and adolescents, developing measures comparable to the DES for younger individuals. This led to his promulgation of the Discrete Behavioral States (DBS) model of dissociation, discussed in depth in the sections on Etiology and on Dissociative Identity Disorder. He also began a groundbreaking, longitudinal case-control study of girls with documented histories of incest, studying a wide range of variables from hypothalamic–pituitary axis (HPA) functioning, dissociation, hypnotizability, development of clinical disorders, functioning at school and socially, and many others. This study has continued for over 25 years, now following the children of the original subjects with an over 90 percent retention rate.

David Spiegel of Stanford University, America's senior expert on dissociation, studied cases of dissociative identity disorder and wrote influential papers on the relationship of hypnosis, trauma, PTSD, dissociation, and dissociative disorders. He and his colleagues also wrote a series of studies on acute dissociation in victims of natural disasters in California. This work led to the DSM-IV diagnostic criteria for acute stress disorder (ASD), a diagnosis now subsumed in the DSM-5 trauma- and stressor-related disorders category. Under Dr. Spiegel's direction, the

DSM-IV changed the names of the dissociative disorders to emphasize their dissociative character and make them more clinically and theoretically relevant. He also took the lead in the changes in diagnostic criteria for the dissociative disorders in DSM-5. He led major studies on dissociation and PTSD, and directed the international studies that were a major factor in the inclusion of a dissociative subtype of PTSD in DSM-5. Also, Dr. Spiegel helped coordinate publication of major papers in support of the trauma model (TM) of dissociation that challenge the sociocognitive model (SCM) of dissociation and of dissociative identity disorder.

International and Cross-Cultural Research on Dissociation and Dissociative Disorders. Beginning in the 1980s, there developed major international research efforts on dissociation and the dissociative disorders. Judith Armstrong, Bethany Brand, Suzette Boon, Phillip Coons, Martin Dorahy, Nel Draijer, Ruth Lanius, Richard J. Loewenstein, Alfonso Martínez-Taboas, Warwick Middleton, Colin Ross, Vedat Sar, Daphne Simeon, Marlene Steinberg, Masahiro Umesue, Onno van der Hart, Eric Vermetten, and other researchers in the United States, Canada, Europe, Australia, New Zealand, Latin America, Turkey, and Japan began systematic research studies on the phenomenology, epidemiology, diagnosis, psychobiology, and treatment of dissociative disorders, especially dissociative identity disorder, in children, adolescents, and adults.

As part of this globalized investigation of dissociation, anthropologically informed research on pathological possession and other trance states began to be conducted cross-culturally with psychometrically sound measures of dissociation. These studies revealed notable similarities between well-characterized cases of pathological possession and some features of dissociative identity disorder, such as loss of personal identity followed by assumption of an alternate (usually spiritual) identity, high indices of trait and state dissociation, and elevated trauma exposure. These research findings paved the way for the inclusion of some pathological possession presentations under the rubric of dissociative identity disorder in DSM-5. Research by Roberto Lewis-Fernández and colleagues on the dissociative nature of some acute reactions to traumatic events, such as *ataque de nervios* (attack of nerves) in Latin America, also led to the inclusion of an acute dissociation category under the other specified dissociative disorders section in DSM-5.

Richard P. Kluft made many contributions to the literature on dissociation and dissociative identity disorder. He wrote extensively about the psychotherapy of dissociative identity disorder and to the phenomenology of the disorder throughout the life cycle. He was the first to identify the high frequency in dissociative identity disorder of apparent *first-rank symptoms* (FRS) of psychosis, albeit experienced quite differently in dissociative identity disorder and indicative of the overlap

and interference of states described by Putnam. He is the only scholar to write extensively about fusion and integration of self-states in the course of treatment of dissociative identity disorder.

The International Society for the Study of Trauma and Dissociation (ISSTD) began under a somewhat different name in 1984. This society has published three editions of Guidelines for the Treatment of Dissociative Identity Disorder in Adults (last updated in 2011) and Guidelines for the Evaluation and Treatment of Dissociative Symptoms in Children and Adolescents (2004). Jennifer Freyd, Ph.D., of the University of Oregon edits the ISSTD's journal, *The Journal of Trauma and Dissociation* (JTD) of the University of Oregon.

Contemporary Study of Depersonalization/Derealization Disorder

Recent systematic research of two large cohorts of well-characterized participants with depersonalization/derealization disorder has yielded strikingly consistent findings. These were reported by Daphne Simeon of Mount Sinai Medical School in the United States and by Anthony David at the Depersonalization Research Unit (DRU) at the Maudsley Hospital, Kings College, London. Clinical features, course and prognosis, and cognitive and neurobiological correlates have emerged that clearly distinguish this disorder from other psychiatric disorders with symptoms of depersonalization. Mauricio Sierra and German E. Berrios, also affiliated with the DRU, developed the most widely, used instrument for screening and diagnosis of this disorder, the *Cambridge Depersonalization Scale* (CDS). Both the US and the British groups have studied treatment and treatment outcome of patients with depersonalization/derealization disorder, including psychopharmacology and psychotherapy (see section on depersonalization/derealization disorder Treatment).

DISSOCIATIVE DISORDERS, MODERN SOCIETY, AND THE MEDIA

The study of the dissociative disorders appears to be inseparable from the social and cultural ambiance in which it occurs. Contemporary academic interest in trauma and its clinical outcomes has been paralleled by social and religious movements focusing on self-identification as a victim of childhood abuse, and the seeking of recovery through a variety of therapeutic and self-help modalities. Furthermore, some feminist theorists argued that healing from childhood sexual abuse mandated confrontation with alleged perpetrators in personal meetings, in the courts, or both.

Legal theorists, concerned that victims of childhood maltreatment did not have standing to sue for damages or to bring criminal actions under usual statutes of limitations, lobbied state legislatures to expand the length of time that victims could seek legal redress. This was based on the idea that many survivors had amnesia for abuse or did not recognize the harm

that had been done to them until many years after the abuse had occurred.

Many states did modify the statute of limitations for childhood sexual abuse, allowing the statute to begin to toll after the person recovered memory of the abuse or, alternatively, after recognizing the harm that had been done whether amnesia was present or not. A number of successful legal actions were brought under these definitions, with courts accepting expert testimony concerning the validity of "repressed memory" for childhood maltreatment. Other state courts have explicitly rejected testimony based on delayed recall of traumatic memory.

In the 1980s, clinicians working with dissociative disorders began to report patient accounts of multi-perpetrator, multi-victim, organized sadistic abuse in the context of supposed underground cult groups. These chilling accounts were amplified by highly publicized daycare abuse cases that described multiple perpetrators, in some cases with elements of ritualized, occultist abuse. Several state governments developed commissions to investigate allegations of "ritual abuse" and claimed to have found evidence of these activities. The media also developed an intense interest in these reports, with publication of popular autobiographical accounts of patients with reported histories of occultist abuse, as well as frequent news stories and television specials on this and related topics.

Unfortunately, some clinicians working with these dissociative patients chose to respond to their patients' reports in a highly publicized and sensationalized manner, rather than subjecting them to dispassionate and rigorous scrutiny. Some of these clinicians painted themselves as crusaders against a vast criminal conspiracy that threatened the nation, presenting themselves and their patients on television and in the newspapers. Patients on some specialized dissociative disorder units shared detailed recollections of trauma in group therapy and in the hospital milieu, with relatively little concern for possible cross-contamination or confusion about the origin of accounts of trauma. Cautionary or critical voices among dissociative disorder researchers and clinicians were ignored or rebuked. Forensic and clinical research was disregarded that cast doubts on the veracity of some of these patients' accounts and suggested alternative mechanisms for many of these ritual abuse narratives. Simultaneously, police were involved in the arrest and prosecution of individuals accused of the sadistic sexual abuse of children in daycare settings.

Backlash. Ultimately, an extensive backlash occurred that included dismissal of all the daycare cases, even those with solid, independent confirmation of the abuse, and increased skepticism about the reliability of children's reports of childhood maltreatment, particularly for childhood sexual abuse. Popular works that encouraged readers to work on recovery from abuse were castigated as suggesting abuse to millions when none had

occurred. Allegations of delayed recall for childhood trauma (unfortunately often termed "*recovered memory*") in clinical and forensic cases were disparaged as based on erroneous views of human memory and cognition, although no one had questioned previously the phenomenon of amnesia related to combat, rape, or other types of trauma. The dissociative disorders construct was dismissed completely as engendered by suggestion and sociocognitive factors (see section on Etiology). Therapy based on working with traumatic memories was discounted as "risky." The reports of an extensive prevalence of childhood abuse in the general population and its actual impact on psychopathology were questioned. Some critics raised doubts about the validity of the diagnostic construct of PTSD.

FALSE MEMORY SYNDROME. The False Memory Syndrome Foundation (FMSF), founded in 1992, produced a significant impact on this debate. The founders of this organization sought out "accused" families, like themselves, some of whom had had successful criminal or civil judgments brought against them because of their children's allegations of abuse. This organization, accompanied by several attorneys, developed a legal theory that there was no such thing as delayed recall for trauma or dissociative identity disorder. The organization's argument went as follows: poorly trained clinicians convinced credulous (usually female) patients that their difficulties were all based on a history of repressed childhood abuse and engaged in risky therapeutic techniques (*guided imagery, hypnosis*) that led patients to believe that these confabulations were accurate. The patients then were mistakenly urged to confront or to stop all contact with family members accused of sexual abuse, ritual abuse, and other forms of maltreatment.

A group of attorneys and academic expert witnesses sought out patients and ex-patients to bring civil cases throughout the country in which therapists were sued for malpractice or were subject to licensing board complaints, or both, based on the FMSF position. Recanting former victims and/or accused parents sued mental health providers alleging malpractice and/or alienation of affections, respectively, based on the notion that there is no such thing as dissociation, amnesia, "repression" of memory, and dissociative disorders.

Many of these cases were settled. Several resulted in jury findings for the plaintiffs with significant damage awards. Others concluded with juries finding for the defendant doctors. A federal mail fraud case against several clinicians ended in a mistrial before the prosecution's case had concluded, with the prosecution declining to retry the case.

Only one published study has looked systematically at retractors of abuse allegations, amnesia, and dissociative identity disorder. It found that virtually all of these individuals had long psychiatric histories, with PTSD and dissociative symptoms predating contact with defendant clinicians.

Essentially all of these patient plaintiffs had significant personality disorders in which they readily took the victim role and looked to forces outside themselves for explanation of their problems. Factitious elements were prominent in many patients' presentations. With treatment, most had actually *improved* with respect to their dissociative psychopathology, but not in their Axis II difficulties. Later, usually after termination of treatment, the patients shifted their perception of victimization from their parents to their prior therapists, now putting their attorneys in the role of rescuer, usually with the promise of large sums of money if they prevailed in their lawsuits. The experience of many experts suggests that some therapies that resulted in false-memory lawsuits likely could have resulted in litigation for treatment below the standard of care for treatment of dissociative disorders.

Despite its popularity in the media, *false-memory syndrome* as a clinical construct has never been operationalized, studied, or validated. As the false-memory fad began to fade, the polarized, extreme views about delayed recall of trauma have been better resolved through research and increased communication between trauma and cognitive researchers, who reject both the extreme false memory view, as well as the extreme credulous view, that is, all reported traumatic memories are accurate (see section on Etiology). A case example illustrates the complexity of this sort of case.

Ms. C brought a malpractice suit against her former psychologist alleging creation of false memories and iatrogenic dissociative identity disorder. In adolescence, Ms. C had been removed from her parental home based on her reports, later recanted, of physical, sexual, and emotional abuse perpetrated by her father. Sporadic contact with her family was usually associated with clinical deterioration. She carried numerous diagnoses, had minimal response to virtually all psychiatric medications, and had required multiple hospitalizations, primarily for self-destructive behavior. She became the patient of Dr. Z, who made the diagnosis of dissociative identity disorder based on switching to defined self-states and the presence of dissociative amnesia. Dr. Z rapidly became over-involved with Ms. C, telling her she "loved" her and considered her as a daughter. Ms. C was not billed for insurance co-payments, was seen outside of sessions in social venues, and engaged in a voluminous, daily email correspondence with Dr. Z. Ms. C began to report a history of multi-perpetrator occultist abuse, based in her family hometown. Some of these accounts appeared to be unlikely to be possible in physical reality. Over several years, Ms. C began to be able to maintain herself safely, stay out of the hospital, and improve her functioning. On the other hand, Dr. Z found herself increasingly enervated by the demands Ms. C continued to place on her. After consultation, Dr. Z attempted to change the therapy framework to improve the boundaries. Ms. C reacted catastrophically to this as a profound rejection; she physically attacked Dr. Z in her office, and therapy was terminated. Ms. C eventually began treatment with a clinician who debunked the dissociative identity disorder diagnosis and "recovered memory" ideas, and encouraged Ms. C to sue Dr. Z. The case was eventually settled out of court for a relatively small sum from Dr. Z and hospitals where Ms. C had been a patient.

PERSISTENCE OF CONTROVERSY ABOUT DISSOCIATION AND DISSOCIATIVE DISORDERS

Despite the extensive body of systematic, international research on

dissociation and dissociative disorders since the 1980s, passionate disagreements remain surrounding the disorders that evolved from classic *hysteria*. These academic and social debates parallel many of those that polarized the field in the days of Babinski and Janet. No other group of disorders in psychiatry can evoke so much heated controversy, conflict, and fervent debate.

Why should there be such a passionate debate? Why should appeals to the results of scientific research not resolve it? More so than other psychiatric conditions, these disorders touch on many complex and contentious aspects of personal, political, philosophical, and even religious beliefs. These include the fundamental nature of the self, the mind, memory, volition, and consciousness, as well as the responsibility of individuals for their behavior.

In psychiatry and psychology, dissociative disorders are the focus of controversies that include the longstanding debates between mentalists and behaviorists, between *psychodynamically oriented* and *behaviorally oriented* clinicians, between various researchers in cognitive psychology, between cognitive researchers and clinical researchers and practitioners, and between different theoretical schools at odds over the nature of hypnosis. Some Western fundamentalist Christian clinicians view *actual* demonic possession as part of the differential diagnosis of dissociative disorders.

Issues here include the significance of early trauma for human psychopathology; the existence of *unconscious* mental life and *intrapsychic defenses,* such as *repression* and *dissociation;* the nature of memory; and the nature of hypnosis, and whether it involves altered states of consciousness or is only a "learned social role." The existence of dissociative amnesias and delayed recall of reported traumatic childhood events raise difficult-to-answer questions about the reliability of traumatic memory.

When viewed within a larger sociopolitical perspective, dissociation theory intersects with many of the most controversial social issues of modern times. Recent systematic research consistently has found a robust relationship between dissociation and traumatic experiences. The role of trauma, particularly intergenerational violence and sexual abuse, crosses into historically taboo subjects, such as rape, incest, child abuse, domestic violence, trafficking of both children and adults in prostitution and/or pornography, and their actual prevalence across societies. This leads into larger legal, social, and cultural questions related to peace and war, the meaning of violence across cultures throughout the world, including differing religious views about the relationship between men, women, and children, and the nature of the family.

The reported failure of many institutions to respond to reports of trauma has also led to intense, polarized debates. Recently these have

included sexual abuse scandals involving not only the Catholic Church, but Christian, Mormon, Jehovah Witness, Buddhist, Hindu, and Orthodox Jewish religious groups as well. In addition to extensive reports of sexual abuse of children by clergy, there have been reported failures of the religious hierarchies to respond appropriately to these abuse reports, sacrificing the well-being of children to that of the accused clergy. Religious communities have been polarized by these abuse allegations.

Similarly, there have been extensive reports of military sexual trauma, rape on college campuses, and failure of institutional hierarchies to respond adequately to the extent of these problems. Also, the returning veterans from the wars in Iraq and Afghanistan have opened up wide discussion, as well as much new research, on the impact of combat trauma. Here again, the issue has not just been the extent of posttraumatic disorders in the returning veterans, but also concerns that the military culture denies and even shuns veterans with posttraumatic disorders, among other issues.

Even so, skeptical pieces about dissociative amnesia, the nature of dissociation, and dissociative identity disorder continue to appear in the media and psychiatric journals. It is unclear why this skepticism specifically focuses on the disorder most associated with severe early life, primarily intrafamilial maltreatment; however, some of these skeptics have also questioned rates of PTSD among Vietnam veterans, and whether PTSD itself should continue as a specific disorder in DSM-5.

DISSOCIATION AND THE MEDIA

The media presentation of dissociative disorders also plays a role in these controversies. From the mid-19th century to the present, the popular media has nurtured public fascination with various forms of dissociation, especially multiple personality, but also fugue, amnesia, and crimes committed under the influence of hypnosis or in somnambulistic states. Accounts of 19th century cases were carried in popular magazines such as *Harpers,* and famous cases, such as Morton Prince's Miss Beauchamps (1905), became the subjects of books and plays. Robert Louis Stevenson's *The Strange Case of Dr. Jekyll and Mr. Hyde* (1886) is the most well-known of these works. Searches of newspaper archives from the mid-19th century onward show many cases of amnesia and fugue described in the popular press (now also on television and the internet), often with the goal of identifying the amnestic individual, or reporting the successful identification of the person and reunion with the family.

In the late 20th century, popular, first person accounts of multiple personality, such as *The Three Faces of Eve, Sybil,* and others, continued in this vein. The widespread use of multiple personality as a fictional plot device—generally to give a bizarre twist to the story—has contributed to the public and professional confusion that surrounds this disorder. Similarly,

amnesia for life circumstances and for trauma is also a recurrent story mechanism in contemporary novels, films, and television—almost never portrayed accurately. Media stereotyping occurs for many forms of mental illness, but the distortions typical in most fictional depictions of dissociative disorders are particularly misleading. Unfortunately, it often seems that most mental health practitioners also have received their training in dissociative disorders from the media, thus leading to misdiagnosis and mistreatment of dissociative patients.

Popular television interview programs have featured therapists and dissociative identity disorder patients, often with interviewers and the audience demanding to see the patients "switch." The ISSTD treatment guidelines specifically enjoin clinicians from involving themselves and their patients in the media. Thus, for the most part only poorly trained and/or exploitative clinicians have been involved in these media debacles. Due to their extensive traumatization, dissociative identity disorder patients are often highly susceptible to later exploitation, including involvement in the media without informed consent about how damaging this can be, as well as the permanent record that will be left in the age of the Internet.

Highly publicized criminal and civil cases, such as the Hillside Strangler, the Franklin murder case, the Ramona case, and many others, have brought additional media attention to claims of multiple personality, amnesia, and recovered memory. These contentious cases, with their warring academic experts and sensationalized media depictions, have also fueled popular stereotyping of dissociative disorders. Media spin about these cases has often been inaccurate and misleading. Unfortunately, for the foreseeable future, because of the complexity of the social, cultural, philosophical, and political issues that these conditions evoke, it is unlikely that debates about them will remain purely in the academic arena.

COMPARATIVE NOSOLOGY

Diagnostic Criteria: DSM and ICD

The modern notion that the dissociative disorders are a distinct group began with the introduction of the DSM-III in 1980. Although the DSM-I distinguished dissociative reactions from other reactions, the DSM-II subsumed dissociative conditions under the superordinate category of hysterical neurosis. The latter was conceptualized as conversion or dissociative subtypes.

In the DSM-III, conditions that developed from the 19th century construct of hysteria were distributed among different diagnostic categories: dissociative disorders, somatoform disorders—now somatic symptom and related disorders in DSM-5—PTSD, and histrionic and borderline personality disorders. The DSM-III, as did the 10th edition of the *International Statistical Classification of Diseases and Related Health*

Problems (ICD-10), took the stance that the label *hysteria* had become so imprecise and variously defined that it had become meaningless. In the DSM-III, these conditions were defined and organized separately, so that systematic empirical research could better clarify their reliability, validity, and possible relationship to each other. The DSM-III dissociative disorders were psychogenic amnesia, psychogenic fugue, multiple personality disorder, and depersonalization disorder.

The ICD-10 classifies the dissociative disorders among the neurotic, stress-related, and somatoform disorders. The ICD-10 dissociative (conversion) disorders include dissociative amnesia, dissociative fugue, dissociative stupor, trance and possession disorder, and a category, dissociative disorders of movement and sensation, that includes dissociative motor disorders, dissociative convulsions, and dissociative anesthesia and sensory loss. This category is roughly equivalent to the DSM-IV-TR and DSM-5 conversion disorder diagnosis. Ganser syndrome and multiple personality disorder are classified under other dissociative disorders. Depersonalization disorder is classified separately and includes both depersonalization and derealization symptoms, whereas from DSM-III until DSM-IV-TR, the DSM system only had a depersonalization symptom criterion. In ICD-10, as in DSM-III, dissociative disorders were described primarily as acute disorders related to traumatic and/or stressful circumstances. The ICD-10 diagnostic criteria for dissociative disorders is found in Table 20–1.

The evolution from the DSM-III conceptualization of dissociation and the dissociative disorders diagnostic criteria through DSM-III-R, DSM-IV, DSM-IV-TR, and DSM-5 has been based on a substantial and increasingly sophisticated body of international research on the phenomenology, course, etiology, treatment, and treatment outcome of the dissociative disorders (see subsequent sections). In particular, these data support the robust relationship of the dissociative disorders with traumatic and/or overwhelming experiences (see sections on History, Epidemiology, and Etiology). This research has been buttressed by the development of a variety of symptom screening measures for dissociation and dissociative disorders, as well as clinician-administered interviews for dissociative symptoms and diagnosis of dissociative disorders. The most important of these measures are detailed in Table 20–2.

The DSM-III criteria for the dissociative disorders were established by expert consensus. By the time DSM-III-R was published, systematic research on the dissociative disorders had increased substantially. Little changed in the diagnostic criteria for depersonalization disorder, psychogenic amnesia, and psychogenic fugue between DSM-III and DSM-III-R. However, the criteria for multiple personality disorder changed considerably to make them more flexible, less reified, and more compatible with research findings about the phenomenology of self-states.

In the DSM-IV (and IV-TR), the names of most of the disorders were changed to more clearly specify their dissociative nature (e.g., psychogenic amnesia became dissociative amnesia and multiple personality disorder became dissociative identity disorder). The criteria for dissociative amnesia now specified that the amnesia usually occurred for traumatic or stressful circumstances. Dissociative fugue now could be diagnosed in the absence of the assumption of another identity, aligning it with ICD-10. Amnesia was added as a criterion symptom for dissociative identity disorder, as is the case in ICD-10. The disorders were structured to reflect pathologies of memory, identity, and perception, respectively.

As will be discussed in more depth below, the DSM-IV (and TR) category of dissociative disorders not otherwise specified (NOS) contained several examples that became particular areas of attention in DSM-5: "Clinical presentations similar to dissociative identity disorder that fail to meet full criteria for this disorder. Examples include presentations in which a) there are not two or more distinct personality states, or b) amnesia for important personal information does not occur;" "Derealization unaccompanied by depersonalization in adults;" and "Dissociative trance disorder . . . [and] Possession trance disorder . . . [where the] . . . dissociative or trance disorder is not a normal part of a broadly accepted collective cultural or religious practice."

DSM-5 Dissociative Disorders and the DSM-5 Trauma- and Stressor-Related Disorders

The DSM-5 Work Group made major changes to the classification of the posttraumatic and dissociative disorders. PTSD, ASD, adjustment disorder, and several other conditions are now placed in a separate category of trauma- and stressor-related disorders. Because of the robust relationship between dissociative disorders and antecedent trauma, dissociative disorders were considered for inclusion in the trauma- and stressor-related category.

The Work Group decided against this inclusion mainly because, unlike all the disorders in this new category, the diagnostic criteria for the dissociative disorders do not require exposure to a stressor or a traumatic event as a diagnostic criterion. However, in order to show the strong relationship between dissociation and trauma, the DSM-5 Work Group placed the dissociative disorders chapter immediately after that of the trauma- and stressor-related disorders. This relationship is explicitly described in the introduction to the dissociative disorders in DSM-5. Also, the DSM-5 diagnostic criteria for PTSD and ASD now specifically describe flashbacks as dissociative. In addition, the PTSD criteria include a new *dissociative subtype,* characterized by depersonalization and/or derealization in response to posttraumatic reminders, often, but not invariably, alternating with the typical intrusive symptoms that

characterize nondissociative forms of PTSD. The ASD diagnostic criteria include dissociative symptoms such as depersonalization, derealization, altered time sense, and dissociative amnesia. Recent studies suggest that there may also be a dissociative subtype of ASD.

Table 20–1.
ICD-10 Diagnostic Criteria for Dissociative (Conversion) Disorders

G1. There must be no evidence of a physical disorder that can explain the characteristic symptoms of this disorder (although physical disorders may be present that give rise to other symptoms).
G2. There are convincing associations in time between the onset of symptoms of the disorder and stressful events, problems, or needs.

Dissociative amnesia

A. The general criteria for dissociative disorder must be met.
B. There must be amnesia, partial or complete, for recent events or problems that were or still are traumatic or stressful.
C. The amnesia is too extensive and persistent to be explained by ordinary forgetfulness (although its depth and extent may vary from one assessment to the next) or by intentional simulation.

Dissociative fugue

A. The general criteria for dissociative disorder must be met.
B. The individual undertakes an unexpected yet organized journey away from home or from the ordinary places of work and social activities, during which self-care is largely maintained.
C. There is amnesia, partial or complete, for the journey, which also meets criterion C for dissociative amnesia.

Dissociative stupor

A. The general criteria for dissociative disorder must be met.
B. There is profound diminution or absence of voluntary movements and speech and of normal responsiveness to light, noise, and touch.
C. Normal muscle tone, static posture, and breathing (and often limited coordinated eye movements) are maintained.

Trance and possession disorders

A. The general criteria for dissociative disorder must be met.
B. Either of the following must be present:
   1. *Trance.* There is temporary alteration of the state of consciousness, shown by any two of the following:
      a. Loss of the usual sense of personal identity
      b. Narrowing of awareness of immediate surroundings or unusually narrow and selective focusing on environmental stimuli
      c. Limitation of movements, postures, and speech to repetition of a small repertoire
   2. *Possession disorder.* The individual is convinced that he or she has been taken over by a spirit, power, deity, or other person.
   3. (1) and (2) of criterion B must be unwanted and troublesome, occurring outside, or being a prolongation of, similar states in religious or other culturally accepted situations.
C. *Most commonly used exclusion clause.* The disorder does not occur at the same time as schizophrenia or related disorders, or mood (affective) disorders with hallucinations or delusions.

Dissociative motor disorders

A. The general criteria for dissociative disorder must be met.
B. Either of the following must be present:
   1. Complete or partial loss of the ability to perform movements that are normally under

voluntary control (including speech)

2. Various or variable degrees of incoordination or ataxia, or inability to stand unaided

Dissociative convulsions

A. The general criteria for dissociative disorder must be met.
B. The individual exhibits sudden and unexpected spasmodic movements, closely resembling any of the varieties of epileptic seizure but not followed by loss of consciousness.
C. The symptoms in criterion B are not accompanied by tongue biting, serious bruising or laceration due to falling, or urinary incontinence.

Dissociative anesthesia and sensory loss

A. The general criteria for dissociative disorder must be met.
B. Either of the following must be present:
   1. Partial or complete loss of any or all of the normal cutaneous sensations over part or all of the body (specify: Touch, pin prick, vibration, heat, cold)
   2. Partial or complete loss of vision, hearing, or smell (specify)

Mixed dissociative (conversion) disorders

Other dissociative (conversion) disorders

This residual code may be used to indicate other dissociative and conversion states that meet Criteria G1 and G2 for dissociative (conversion) disorders but do not meet the criteria for the dissociative disorders listed previously.

Ganser syndrome (approximate answers)

Multiple personality disorder

A. Two or more distinct personalities exist within the individual, only one being evident at a time.
B. Each personality has its own memories, preferences, and behavior patterns and, at some time (and recurrently), takes full control of the individual's behavior.
C. There is inability to recall important personal information, which is too extensive to be explained by ordinary forgetfulness.
D. The symptoms are not due to organic mental disorders (e.g., in epileptic disorders) or psychoactive substance-related disorders (e.g., intoxication or withdrawal).

Transient dissociative (conversion) disorders occurring in childhood and adolescence

Other specified dissociative (conversion) disorders

Specific research criteria are not given for all disorders mentioned previously, because these other dissociative states are rare and not well described. Research workers studying these conditions in detail should specify their own criteria according to the purpose of their studies.

Dissociative (conversion) disorder, unspecified

From World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization; 1993, with permission.

Table 20–2. Reliable and Valid Screening Measures and Structured Diagnostic Interviews for Dissociation and Dissociative Disorders. Dissociation Scales and Diagnostic Interviews

	Characteristics	Structure
Self-report screening measures		
Dissociative Experiences Scale (DES)[a]	Self-Report screening measure to assess trait dissociation in adults including items for amnesia, absorption, depersonalization,	Average score of 28 0–100% of the tin experienced

	derealization; has been translated into more than 40 languages, with strong cross-cultural similarities for western and non-Western populations	
DES-Taxon Scale (DES-T)[a]	Developed through Taxonic Analysis of DES scores of over 1,500 patients with many diagnoses, as well as nonpatient controls	Average score of a s items from DES: Items: 3, 5, 7, 8, 12, DES and DES-T calcu gives the probabi average score) th individual fits the downloaded at http://www.isst-d.org/default.asp contentID=66
Adolescent DES (A-DES)[a]	A version of the DES validated as a screening measure for trait dissociation in adolescents	Average score of 30 on a 0–10 scale fr "always"
Child Dissociative Checklist (CDC)[a]	Parent, caregiver, teacher report for screening of dissociative symptoms in children 5–12 years of age	Average score of 20 from 0 to 2, "not true" of the child previous 12 mont

Somatoform[a] Dissociation Questionnaire (SDQ-20)	SDQ measures somatoform dissociation: dissociative symptoms that phenomenologically involve the body; i.e., somatosensory and conversion symptoms common in dissociative patients. These include motor inhibitions, loss of function, anesthesias and analgesias, pain, and problems with vision, hearing, and smell.	Average score of 20 from 1 to 5 “not a “extremely”
Somatoform Dissociation Questionnaire-5 (SDQ-5)	5 items that screen for somatoform Dissociation	Average score of 5 it SDQ-20, scored fr “not at all” to “ex
Peritraumatic Dissociative Experiences Questionnaire (PDEQ)	Measures the extent of dissociation at the time of a traumatic event. May consist of one or two factors including depersonalization/derealization	Exists in both Rater Report Versions (as well as brief ve questions). Each i from 1 (not at all (extremely true). score is the sum c score above 15 is significant dissoci

Cambridge Depersonalization Scale (CDS)	Assessment measure for depersonalization and derealization symptoms	29 questions rated f and duration ove months. Score >112 indicativ depersonalizatior disorder. Significa than patients witl disorders, mood c partial-complex s Dissociative ident patients often scc
Comprehensive self-report (Diagnostic Measures)		
Multidimensional Inventory of Dissociation (MID)[a] There is also an MID for adolescents.	Diagnostic self-report for dissociative disorders, PTSD, somatic symptom disorders, borderline personality disorder features	218 items self-repor often someone e symptoms from 0 10 (always); only measure that has scales; 30–90 min program available
Multi-scale Dissociation Inventory (MDI). Available at http://www.johnbriere.com/multiscale.htm	Multiscale measure of dissociation with a 5-point Likert format that generates six dissociative symptom subscales and a total dissociation scale. Normed on 444 trauma exposed individuals in the community	30 items; approxima minutes to compl include disengage depersonalizatior derealization; em constriction (num memory disturba dissociation; and dissociation score

Clinician-Administered Dissociative States Scale (CADSS)	Clinician administered inventory that measures current state dissociation; has 28 items rated by the interviewer from 0 (“never”) to 4 (“extreme”)	Administered by a cl in the context of a experimental stuc symptoms of amr depersonalizatior derealization. The proven useful in pharmacological s of the dissociative ketamine in the tr depression), and i stress studies
Structured Clinical Interview for DSM-IV-TR Dissociative Disorders, Revised (SCID-D-R)-being updated for DSM-5	Semi-structured, clinician-administered interview; 276 questions on a 4-point scale; 1–3 hours, if dissociative symptoms are present; much less time for nondissociative patients. Widely seen as “gold standard” for research studies requiring a diagnosis, and for differential diagnosis clinical cases.	Assesses the presen severity of amnes confusion and alt depersonalizatior derealization, anc DSM-IV-TR diagnc dissociative disor acute stress disor
Dissociative Disorders Interview Schedule (DDIS). Has been updated for DSM-5. The DSM-5 version is available at http://www.rossinst.com/ddis.html	Clinician Administered interview 132 items. Strict yes–no format for all questions; may limit utility in situations where feigning is a concern. Requires approximately 30–60 minutes to conduct with dissociative identity disorder patients.	Inquires about, in ac dissociative symp abuse history, sub paranormal exper symptoms of maj somatoform, and personality disorc format limits utili diagnosis of bord personality disorc

[a]DES, DES-T, ADES, CDC, SDQ, MID 6 and MID scoring program, and DES scoring program available from International Society for the Study of Trauma and Dissociation. Available at https://netforum.avectra.com/eweb/DynamicPage.aspx?Site=ISSTD&WebCode=SCREENING

Large studies with general population and clinical samples have documented the cross-national presence of the dissociative PTSD subtype in 15 to 30 percent of individuals with PTSD. Future research will need to clarify the relationship between the dissociative subtype of PTSD and the specific dissociative disorders.

Changes to the Dissociative Disorders in DSM-5

The Work Group made significant changes to the dissociative disorders' criteria in DSM-5. These include changes to the definition of dissociation, the diagnostic criteria for dissociative identity disorder, dissociative amnesia, and depersonalization disorder, and the examples included under other specified dissociative disorders.

Change in the Definition of Dissociation

As noted in the first section, the Work Group combined aspects of both the DSM-IV-TR and the ICD-10 definitions of dissociation, but also went considerably beyond them. The definition had to encompass phenomena that may be relatively acute—such as an episode of dissociative amnesia that is recognized shortly after a soldier is removed from combat—as well as childhood-onset complex dissociative disorders that are frequently chronic, like dissociative identity disorder. As part of the latter, individuals may report or display a plethora of dissociative symptoms, including subjective self-division, chronic complex dissociative amnesia, dissociative fugue, overlap and intrusion among self-states, and depersonalization and derealization, among others. Also, the definition of dissociation had to encompass culturally diverse experiences of pathological dissociation, such as pathological possession (see below).

Accordingly, as described in the introductory section, the DSM-5 definition states that dissociation can affect the normal integration of every aspect of psychobiological functioning. Also, dissociative symptoms can be classified as "negative" and "positive," similarly to the definition of psychotic symptoms. The definition specifies that dissociation can be experienced not only as an absence, failure, or lapse, such as an awareness of dissociative amnesia (a negative symptom). It can also be experienced as a disruption, interruption, and/or discontinuity suffered as recurrent, jarring, involuntary intrusions into executive functioning and sense of self, for instance several minutes of rapid shifting of dissociative identity disorder self-states (a positive symptom). Negative and positive dissociative symptoms may occur sequentially. For example, after several minutes of shifting states the individual with dissociative identity disorder may not recall what has just occurred.

CHANGES IN DIAGNOSTIC CRITERIA TO THE DISSOCIATIVE DISORDERS

Dissociative Identity Disorder

Studies of phenomenology of over 1,000 dissociative identity disorder individuals studied mostly in the United States and Europe with the Structured Clinical Interview for DSM-IV-TR Dissociative Disorders (SCID-D-R), the Dissociative Disorders Interview Schedule (DDIS), or the

Multiaxial Inventory of Dissociation (MID) found that over 90 percent of these individuals endorsed the following symptoms, in order of frequency: amnesia (100 percent); subjective self-alteration (98 percent); symptoms of overlap, interference, and intrusion among self-states, often characterized as variants of the FRS of psychotic individuals (94 percent); awareness of the voices of self-states that are usually experienced within the mind (92 percent); and spontaneous trance behaviors (91 percent). Less commonly, dissociative identity disorder individuals described derealization (63 percent); depersonalization (77 percent); and somatoform symptoms (83 percent). In addition, these studies found that, unless the person was in crisis, only 15 percent of dissociative identity disorder individuals regularly showed clear-cut switching behaviors between distinctly differentiated self-states.

These and other studies found that over 40 percent of DSM-IV-TR dissociative disorder diagnoses were for dissociative disorder NOS, most commonly for Example 1 of the dissociative disorders NOS criteria (see section on DSM and ICD). Indeed, clinicians familiar with dissociative identity disorder patients often used this NOS example as a specific diagnosis: a disorder with the clinical features, course, comorbidities, antecedent developmental trauma history, and treatment response of dissociative identity disorder, but without "distinct" identities that took "control" of the person's behavior through a process of "switching," often a subtle process in and of itself.

DSM-5 diagnostic criteria for dissociative identity disorder moved completely away from the notion of relatively stable ("enduring") personality states that switch ("take control") of behavior to a more user-friendly enumeration of the distinct subjective differences seen in dissociative identity disorder presentations. These can include marked discontinuity in the subjective sense of self, or the self as agent of behavior. In addition, the person may experience and/or display discontinuities in mood and affect; manifest behavior; apparent state of consciousness; availability of autobiographical, procedural, and/or semantic memory; perception of the environment; cognitive schema; and/or sensorimotor functioning. The latter may involve conversion symptoms, especially in some non-Western cultures (see section on Diagnosis and Clinical Features). Also, DSM-5 criteria now allow the diagnosis to be made based on reports by the patient or others of the self-state discontinuities, not only observation by the clinician of these states shifts.

In addition, based on extensive review of the cross-cultural literature, the Work Group moved the DSM-IV-TR dissociative disorder NOS example concerning an experience of pathological possession into the diagnostic criteria for dissociative identity disorder, but left dissociative trance in the examples for other specified dissociative disorders. This decision was based on the phenomenological similarity between pathological possession states

and dissociative identity disorder; their shared dissociative origin; the coexistence in the same patient of identities attributed to either possessing entities or to endogenously generated self-states; the frequent occurrence of amnesia for the experiences during the possession experiences in pathological possession; and shared risk factors, such as high rates of lifetime traumatic experiences. While pathological possession states are often described as more transient and episodic than those of non-possession dissociative identity disorder, further research is needed to assess whether these presentations of pathological possession represent the overt appearance of phenomena that persist in more subtle form over time. Also, as described above most non–possession-form dissociative identity disorder individuals, may have subtle presentations of core dissociative identity disorder symptoms, and/or may have a relapsing/remitting course, and appear to be relatively asymptomatic, unless experiencing a clinical decompensation.

The addition of pathological possession to the DSM-5 dissociative identity disorder criteria indicates that, around the world, dissociative identity disorder presents in two main forms: non–possession-form or possession-form. These are not mutually exclusive: individuals in one cultural setting may experience either form, and a single individual may experience both forms. Typically, in non–possession-form dissociative identity disorder the personality states are experienced as *internal* aspects of the person. In possession-form dissociative identity disorder, by contrast, the alternate identities are typically experienced as *external* possessing agents, usually of supernatural or spiritual origin (e.g., demons, spirits). The main contrast between these two forms of the disorder appears to be shaped by the patient's sociocultural milieu. That is, possession-form dissociative identity disorder arises in contexts where possessing agents are considered to be "real" (e.g., some fundamentalist religions in the United States or traditional South Asian cultures), as opposed to non–possession-form dissociative identity disorder, which draws its alternate identities from individualistically focused experiences, such as stages of life (e.g., childhood) or biographical roles in the patient's life (e.g., protector, perpetrator) (see section on Dissociative Identity Disorder).

DSM-5 is careful to clarify that not all presentations of possession meet criteria for dissociative identity disorder. In addition to meeting criterion A (description of dissociative symptoms), individuals with a presumptive presentation of possession-form dissociative identity disorder must also meet criterion B (amnesia), criterion C (the possession state has to be pathological, that is leading to distress or impairment) and criterion D (not a normal part of a broadly accepted cultural or religious practice). Most instances of possession in the world are normal, embedded in spiritual activities, or do not entail dissociative experiences, functioning instead to communicate distress in a spiritual language (e.g., as a cause of any

number of physical or mental symptoms). The inclusion of possession-form presentations in DSM-5 is intended to improve detection of individuals who otherwise meet dissociative identity disorder criteria but whose identification might be missed due to clinicians' lack of familiarity with this presentation, leading to potentially better treatment response as well as more culturally relevant research studies of disorder prevalence and treatment outcome.

It remains a matter of clinical judgment just how "distinct" personality or possession states have to be, or how "marked" the discontinuities in sense of self and agency, to meet the A criterion for dissociative identity disorder. Despite the use of terms like "distinct" and "marked" in the criteria, the Work Group was clear that all the data on dissociative identity disorder support the subtlety of the symptoms in this disorder. The *distinctness* of self-states does not necessarily refer to the secondary elaboration of the states into dramatically different manifest presentations (see section on Dissociative Identity Disorder). Rather, the term refers to the extent of compartmentalized information manifested by the identity states, accompanied by a subjective sense of self and a capacity for behavioral enactment. Similarly, the term *marked* refers to the disruptiveness of the symptoms to the person, not necessarily to the dramatic presentation of dissociative symptoms.

Finally, based on the phenomenological studies, the DSM-5 diagnostic criteria for dissociative identity modify the amnesia criterion (B). Here, the changes in the amnesia criterion emphasize the complexity of reported memory disturbances, and that patients often report amnesia for ongoing experience (see section on Diagnosis and Clinical Features).

The evolution of the diagnostic criteria for dissociative identity disorder in DSM-III, DSM-IIIR, ICD-10, DSM-IV-TR, and DSM-5 are found in Table 20–3.

Changes to the Diagnostic Criteria for Dissociative Amnesia

There are few systematic studies of dissociative amnesia in the literature. Studies of severe PTSD in traumatized groups have shown very high rates of the PTSD dissociative amnesia criterion. Also, many studies have documented delayed recall of trauma and/or the association of autobiographical memory deficits with trauma. Based on close reading of these studies, many of the subjects would meet diagnostic criteria for dissociative amnesia. However, taking the literature as a whole, the Work Group decided to underscore that dissociative amnesia relates primarily to alteration of *autobiographical memory,* although in some very severe cases of generalized dissociative amnesia, individuals have been shown to lose semantic and procedural memory as well. Also, the criteria now define the most common types of dissociative amnesia: localized, selective, or generalized dissociative amnesia (see Table 20–4 for types of dissociative

amnesia).

Finally, the Work Group changed dissociative fugue from a separate diagnosis to a subtype of dissociative amnesia. There is little systematic literature on dissociative fugue, and the best studies remain detailed descriptive case series from the 1940s and 1950s.

Table 20–3.
Comparison of DSM-III, IIIR, ICD-10, DSM-IV-TR, DSM-5 Diagnostic Criteria for Dissociative Identity Disorder

DSM-III Diagnostic Criteria for 300.14 Multiple Personality Disorder (1980)	DSM-IIIR Diagnostic Criteria for 300.14 Multiple Personality Disorder (1987)	ICD-10 Diagnostic Criteria for (F44.81) Multiple Personality Disorder (1994)	DSM-IV-TR Diagnostic Criteria for 300.14 Dissociative Identity Disorder (1994)	DSM-5 Diagnostic Criteria for 300.14 Dissociative Identity Disorder (2013)
A. The existence within the individual of two or more distinct personalities, each of which is dominant at a particular time.	A. The existence within the person of two or more distinct personalities or personality states (each with its own relatively enduring pattern of perceiving, relating to, and thinking about the environment and self).	A. The existence of two or more distinct personalities within the individual, only one being evident at a time.	A. The presence of two or more distinct identities or personality states (each with its own relatively enduring pattern of perceiving, relating to, and thinking about the environment and self	A. Disruption of identity characterized by two or more distinct personality states, which may be described in some cultures as an experience of possession. The disruption identity involves marked discontinuity in sense of self and sense of agency, accompanied by related alterations in affect, behavior, consciousness, memory, perception, cognition, and/or sensorimotor functioning. These signs and symptoms may be observed by others or reported by the individual.
B. The personality that is dominant at any particular	B. At least two of these personalities or personality states	B. Each personality has its own memories, preferences and behavior	B. At least two of these identities or personality states recurrently	B. Recurrent gaps in the recall of everyday events, important personal information, and/or traumatic

time determines the individual's behavior.	recurrently take full control of the person's behavior.	patterns, and at some time (and recurrently) takes full control of the individual's behavior.	take control of the person's behavior.	events that are inconsistent with ordinary forgetting.
C. Each individual personality is complex and integrated, with its own unique behavior patterns and social relationships.		C. Inability to recall important personal information, too extensive to be explained by ordinary forgetfulness.	C. Inability to recall important personal information that is too extensive to be explained by ordinary forgetfulness.	C. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
		D. Not due to organic mental disorders (F0) (e.g., in epileptic disorders) or psychoactive substance-related disorders (F1) (e.g., intoxication or withdrawal).	D. The disturbance is not due to the direct physiological effects of a substance (e.g., blackouts or chaotic behavior during alcohol intoxication) or a general medical condition (e.g., complex partial seizures). **Note:** In children, the symptoms are not attributable to imaginary playmates or other fantasy play.	D. The disturbance is not a normal part of a broadly accepted cultural or religious practice. Note: In children, the symptoms are not better explained by imaginary playmates or other fantasy play.
				E. The symptoms are not attributable to the physiological effects of a substance (e.g.,

	blackouts or chaotic behavior during alcohol intoxication) or another medical condition (e.g., complex partial seizures).

The Work Group found that there are insufficient data to validate dissociative fugue as a disorder distinct from dissociative amnesia, dissociative identity disorder, or other trauma spectrum conditions. Also, the prior DSM versions of dissociative fugue included a criterion of "assumption of a new identity (partial or complete)." By definition, an individual displaying "a new identity" would meet diagnostic criteria for dissociative identity disorder. The notion of a "partial" or a "complete" new identity is based on older notions of identity alteration in dissociative disorders, as well as being unascertainable by clinical or psychometric assessment. Many individuals who come to clinical attention for an episode of dissociative fugue ultimately meet diagnostic criteria for dissociative identity disorder. Other individuals with dissociative fugue do not differ clinically from those who meet diagnostic criteria for generalized dissociative amnesia, other than for their peripatetic behavior.

Table 20–4.
Types of Dissociative Amnesia

Localized amnesia
Inability to recall events related to a circumscribed period of time
Selective amnesia
Ability to remember some, but not all, of the events occurring during a circumscribed period of time
Generalized amnesia
Failure to recall one's entire life
Continuous amnesia
Failure to recall successive events as they occur
Systematized amnesia
Amnesia for certain categories of memory, such as all memories relating to one's family or to a particular person

Changes to the Diagnostic Criteria for Depersonalization Disorder

The Work Group reviewed the question whether depersonalization disorder should: (a) remain in the DSM-5 Dissociative Disorders section, (b) be moved to another section of the DSM-5, similar to the ICD-10 classification, or (c) be considered a symptom of another disorder or disorders, such as the mood, anxiety, psychotic, or personality disorders, among others. Finally, evidence was examined regarding the nosological accuracy of retaining derealization symptoms as one of the examples of the

DSM-IV-TR dissociative disorder NOS, rather than as a component of a depersonalization–derealization disorder.

Support for the continuing inclusion of DSM-IV-TR depersonalization disorder among the dissociative disorders in DSM-5 is based on findings that depersonalization and derealization are prevalent in other dissociative disorders such as dissociative identity disorder. Depersonalization is one of the quintessential responses to acute trauma (*peritraumatic dissociation*) and has been reported as highly prevalent in those subjected to motor vehicle accidents, emotional/verbal abuse, and imprisonment, among other exposures (see section on Epidemiology). The dissociative subtype of PTSD is based on the symptoms of depersonalization and/or derealization in response to trauma cues.

Several authorities have proposed that depersonalization is a "hard-wired" inhibitory response—part of the human psychobiological survival system—that allows for attenuation of anxiety and psychobiological hyperarousal to help preserve adaptive behavior and conservation of physical resources under conditions of threat and danger. Disorders related to depersonalization may represent generalization and persistence of this response outside the context of immediate threat, just as PTSD fear reactivity can be understood, in part, as a generalization of the adaptive conditioned fear response beyond the immediate context where it was generated.

Cases: Acute Depersonalization

Mr. J, a high functioning professional without psychiatric history, described that, while a college student in a geology class, he was involved in a field trip where he was excavating a small cave. He was working inside the cave when a landslide blocked the entrance to the cave. Although there was complete darkness in the cramped cave, he noticed, with an unemotional sense of detachment, that he could observe himself as if from a distance. He reported having no sense of time. He did not experience fear or concern about his situation. He noted with interest the sounds of his classmates and professor shouting and digging at the front of the cave. He recalls that, as soon as he saw a beam of light come into the cave, he felt that he reentered his body, and suddenly felt anxiety and pain, and noticed bleeding where the rocks from the avalanche had struck his face. He still suffers from a partial facial nerve paralysis due to this event. He denies any subsequent depersonalization/derealization experiences.

Studies of two large cohorts have persuasively documented a consistent nosology for DSM-IV-TR depersonalization disorder. These studies described consistent age of onset, acute and remote antecedents, as well as distinct symptom clusters, cognitive organization, course, treatment response (primarily *lack* of treatment response), and gender ratio. Factor analysis from these studies suggest that DPD symptoms are characterized

by four to five factors: numbing, unreality of self, unreality of other, temporal disintegration, and perceptual alterations, including micropsia/macropsia, out of body experiences, and feelings of unreality of objects and of one's own body. No distinction was found between patients with predominant depersonalization as compared with those with predominant derealization, thus supporting the ICD-10 conceptualization of either symptom as characteristic. Accordingly, the disorder was renamed depersonalization/derealization disorder in DSM-5, and remained among the dissociative disorders.

Dissociative Disorders Not Otherwise Specified Becomes Other Dissociative Disorders in DSM-5

In DSM-5, the examples listed under the category of other dissociative disorders have undergone moderate changes. This category includes four examples, three of which are similar to those in DSM-IV-TR: (a) disorders that are similar to dissociative identity disorder but do not fully meet criteria for the disorder; (b) identity disturbance in response to coercive persuasion; and (c) dissociative trance episodes that are distressing and not part of a culturally or religiously congruent practice. Pathological possession trance episodes and derealization experiences were removed from this diagnostic category and included as aspects of dissociative identity disorders and depersonalization/derealization disorder, respectively.

A new example was added in DSM-5: *acute dissociative reactions to stressful events*. This encompasses usually brief episodes of mixed dissociative symptoms that occur in the context of trauma or extreme stress. They are distinct from brief episodes of monosymptomatic dissociative disorders (e.g., dissociative amnesia) by combining a mix of dissociative symptoms, but without the identity alteration of dissociative identity disorder. These episodes show substantial cultural variation. Some of the cultural concepts of distress included in DSM-5 (e.g., *ataque de nervios*) may meet criteria for this condition. Here DSM-5 returns to the language of DSM-I with its categorization of *reactions,* where current nosology describes *disorders*.

As in the DSM-IV-TR NOS category, these examples may allow more systematic research to consider expansion of the dissociative disorders to include specific diagnostic categories for dissociative trance and acute dissociative reactions, such as a mixed dissociative-pseudopsychotic presentation, usually labeled *dissociative psychosis* (see section on Other Specified Dissociative Disorders.)

VALIDITY OF THE DISSOCIATIVE DISORDERS

As noted in the section on History, despite substantial systematic research

about dissociation and dissociative disorders and their relationship to trauma, debate continues about the validity of the dissociation concept, dissociative amnesia, and dissociative identity disorder. The validity of the dissociation concept, particularly its relationship to trauma will be discussed in the section on Etiology.

Validity of the Dissociative Identity Disorder Construct

Debate about the existence of dissociative identity disorder has waxed and waned for more than a century. However, sufficient data on pathological dissociation, in general, and on dissociative identity disorder, in particular, have accrued to judge this condition by the same standards that are applied to the validity of other psychiatric diagnoses.

The most widely accepted standards are based on a set of criteria first articulated by Eli Robins and Samuel Guze and subsequently refined by others. In essence, a psychiatric diagnosis is considered valid if it satisfies three basic requirements: content validity, criterion-related validity, and construct validity. Content validity requires a detailed clinical description of the disorder that is repeatedly and independently replicated. A large number of clinical phenomenological studies of dissociative identity disorder—including from North and South America, Europe, and Asia—document the presence of a core dissociative psychopathology in dissociative identity disorder patients that fulfills this requirement. Criterion-related validity requires that laboratory tests or reliable psychological tests be consistent with the defined clinical picture. This stipulation is met by the reliable and valid diagnostic interviews and scales that have increasingly been used in dissociation research, as well as psychological testing protocols that discriminate dissociative identity disorder from schizophrenia, depression, nondissociative PTSD and ASD, personality disorders, feigning individuals, and normal controls, among others.

Construct or discriminant validity requires that the disorder can be differentiated from other disorders. By means of structured interviews such as the SCID-D-R, and comprehensive symptom-screening measures such as the MID, dissociative identity disorder patients can be discriminated from normal individuals and patients with Cluster B personality disorders, PTSD, psychotic disorders, and mood disorders, among others, Likewise, psychological test batteries and experimental cognitive and psychophysiological studies have discriminated dissociative identity disorder patients from other subjects, including simulators. Recent psychobiological studies have shown that salivary cortisol, urinary catecholamines, the low-dose dexamethasone-suppression test, and MRI measurements of hippocampal and amygdala volumes discriminate dissociative identity disorder from patients with borderline personality disorder and trauma controls. There were only subtle biological differences

between dissociative identity disorder and PTSD subjects, although the two groups differed significantly in measures of dissociation, such as the Clinician Administered Dissociative States Scale (CADSS) and the DES.

Studies using neurobiological and psychophysiological measurements of dissociative identity disorder subjects compared to simulating controls have invariably shown significant differences between the real subjects and simulators. Also, actual dissociative identity disorder subjects generally show unique patterns of neurobiological and/or psychophysiological variables in these studies, uncharacteristic of other psychiatric disorders. One PET-scan study of dissociative identity disorder found that, when listening to personal trauma scripts, personality states with recall of trauma history as autobiographical memory (Traumatized Identity State), compared with states that did not (Neutral Identity State), showed neuro-circuitry and autonomic responding similar to the dissociative and nondissociative subtypes of PTSD, respectively (see section on Neurobiology).

On the Gudjonsson Suggestibility Scale (GSS), a robust measure of susceptibility to external pressure to produce confabulated narrative accounts, the dissociative identity disorder subjects scored as less suggestible than patients with PTSD, borderline personality disorder, and traumatized controls. All of these findings support the notion that dissociative identity disorder is best conceptualized as a form of posttraumatic psychopathology, not an iatrogenic condition or an epiphenomenon of borderline personality disorder or other disorders.

In summary, dissociative identity disorder satisfies widely accepted standards for psychiatric validity and should be regarded as a legitimate disorder requiring an informed diagnostic and treatment approach.

Validity of Dissociative Amnesia

Studies that address the validity of dissociative amnesia have primarily centered on the existence of the phenomenon, not on the study of the clinical disorder. As noted previously, all of the studies that have attempted to investigate the presence of amnesia for traumatic events have reported that it occurs.

Alternative hypotheses about trauma and memory claim that trauma virtually always causes an indelible inscribing of memory in the brain and that dissociative amnesia does not exist. The debate about the existence of dissociative amnesia has also focused in part on hypothesized psychological mechanisms for the reported amnesia, such as repression or dissociation. Critics have suggested that more mundane processes, such as failure to report, cognitive avoidance, or ordinary forgetting, might account for the apparent delayed recall, although it is difficult to understand how these alternative mechanisms could account for the repeatedly documented findings of acute or delayed amnesia during acute trauma such as combat

or rape. Also, dissociative amnesia remains a criterion symptom for both PTSD and ASD in DSM-5. Most critics of the dissociative amnesia concept do not seem to appreciate that breakdown of dissociative amnesia generally results in intrusive PTSD symptoms: intense, highly disturbing perceptional, somatosensory, and emotional intrusions, and/or full flashbacks, with a fragmented narrative. It does not result in a linear, autobiographical narrative for events, even highly distressing, traumatic, or stressful events.

Data from a variety of research studies show that individuals who report problems with recall of trauma may describe several subjective processes. Some experience inability to recall some or all of traumatic life events or life periods in which the trauma occurred. Others report deliberate attempts to avoid thinking about the traumatic events because of the subjective distress or shame that they experience when they try to do so. Another group describes lack of ability to appreciate the abusive nature of trauma until they were older and/or more emancipated from abusive social milieus. In clinical practice, all of these mechanisms may help account for aspects of dissociative amnesia in an individual patient.

Many studies have found corroboration for the traumatic events recalled by individuals with dissociative amnesia. Types of corroboration have included medical, social service records, and church diocesan records, verbal or written admission by perpetrators, criminal convictions or civil judgments against perpetrators, reports of family members, and information from witnesses or other victims of the same perpetrator. Also, studies have shown no significant differences in the accuracy of delayed recall of trauma compared to accounts of individuals who report continuous memory for trauma. Contrary to stereotype, studies have shown that delayed recall of trauma most frequently occurs outside of treatment contexts, not during psychotherapy.

Reviews of the literature on delayed recall for traumatic experiences have concluded that delayed memories may be shown by corroboration to be generally accurate, partially accurate, or completely confabulated, with no lesser likelihood of accuracy than reportedly “continuous” memories for traumatic events. Reliable corroboration, which may be very difficult to obtain, particularly in cases of reported childhood maltreatment or intimate-partner violence, is the only way to clarify the extent to which a memory is generally accurate (see sections on Epidemiology and Dissociative Amnesia).

EPIDEMIOLOGY OF THE DISSOCIATIVE DISORDERS

Epidemiological studies in North America, Europe, Asia, Australia, Latin America, and Turkey have found pathological dissociation and dissociative disorders to be common in samples of the general population, as well as in samples of psychiatric inpatients, outpatients, emergency department

patients, patients in alcohol and substance abuse treatment, and in special populations with high rates of early childhood trauma, such as women in prostitution. Also, studies of both clinical and community populations with histories of trauma, studied with standardized measures, have shown that there is a high rate of dissociation and dissociative disorders in a subgroup of these individuals.

GENERAL POPULATION STUDIES OF DISSOCIATIVE DISORDERS

Table 20–5 shows the findings for the prevalence of dissociative disorders in three representative general population samples: Winnipeg, Canada, New York State, USA and a female sample in Sivas, a city in central Turkey. As shown in the table, overall prevalence for the dissociative disorders ranged from 9.1 to 18.3 percent. Based on these studies, one researcher has conservatively estimated the lifetime prevalence for dissociative disorders in the community at about 10 percent.

A random sample of 1,055 adults in Winnipeg, Canada, were given the DES and a subsample of 502 respondents, that did not differ significantly demographically from the larger sample, was subsequently reevaluated with the DDIS, findings include a left-skewed distribution on the DES, with only a small subset of the population, 3 to 5 percent, endorsing many forms of dissociative experience, consistent with the construct of *pathological dissociation:* a group that was distinctly different from all other diagnostic categories, with symptoms characteristic of dissociative identity disorder and DSM-IV-TR dissociative disorder NOS. Overall, 12.2 percent of study participants met DSM-IV diagnostic criteria for a lifetime diagnosis of a dissociative disorder.

There was a 9.1 percent 1-year prevalence of dissociative disorders in a longitudinal, community-based study of 658 adults in New York State. Dissociative disorders were associated with clinically significant impairment and, as in other studies, with comorbid mood, anxiety, and personality disorders. Individuals with dissociative identity disorder were the most impaired of the dissociative disorders group.

In a representative sample of 628 women in Sivas, Turkey assessed with the DDIS, the PTSD module of the SCID-I, and the SCID-II, 18.3 percent met criteria for a lifetime DSM-IV diagnosis of a dissociative disorder. Dissociative identity disorder represented about 1.1 percent of the sample; dissociative disorder NOS was the most prevalent dissociative disorder at 8.3 percent, with about half of these meeting partial dissociative identity disorder criteria. Dissociative disorder NOS was also the most common diagnosis in the New York state study. Diagnosis of a dissociative disorder was significantly associated with other comorbid conditions such as PTSD, major depression, somatization disorder, borderline personality disorder, and a history of attempted suicide.

Table 20–5.
Prevalence of Dissociative Disorders in General Population Samples

Study	Ross (1997)	Johnson et al. (2006)	Sar et al. (2007)
Measures	DES and DDIS	Adapted from DES, SCID, SCID-II, and GAFS	DDIS, SCID-PTSD, and SCID-II
Number of subjects	502	658	628 (female)
Diagnosis	Subjects (%)	Subjects (%)	Subjects (%)
Dissociative amnesia	6.0	1.8	7.3
Dissociative fugue	0	0	0.2
Dissociative identity disorder	1.3	1.5	1.1
Depersonalization disorder	2.8	0.8	1.4
Dissociative disorder not otherwise specified (NOS)	0.2	4.3	8.3
Dissociative disorder NOS with multiple personality states			4.1
Dissociative disorder NOS with indirect cues for personality states			2.4
Derealization without depersonalization			1.1
Dissociative trance disorder			0.6
All dissociative disorders	12.2	9.1	18.3

DDIS, Dissociative Disorders Interview Schedule; DES, Dissociative Experiences Scale; GAFS, Global Assessment of Functioning Scale; SCID, Structured Clinical Interview for DSM-IV-TR Dissociative Disorders; SCID-PTSD, Structured Clinical Interview for DSM-IV-TR Dissociative Disorders, Posttraumatic Stress Disorder section.

Data from Ross CA. *Dissociative Identity Disorder: Diagnosis, Clinical Features, and Treatment of Multiple Personality*. New York: John Wiley & Sons; 1997:109; Johnson JG, Cohen P, Kasen K, Brook JS. Dissociative disorders among adults in the community, impaired functioning, and axis I and II comorbidity. *J Psychiatric Res*. 2006;40(2):131; and Sar V, Akyüz G, Dogˇan O. Prevalence of dissociative disorders among women in the general population. *Psychiatry Res*. 2007;149:169.

Pathological Dissociation in the General Population. There are two Finnish general population studies of dissociation: a 3-year prospective study of 2,001 adults and a cross-sectional study of 1,424 adolescents ages 13 to 18. The adult studies used the DES, the DES-Taxon scale (DES-T), the Beck Depression Inventory (BDI), the Somatoform Dissociation Questionnaire (SDQ), the Toronto Alexithymia Scale (TAS), and an instrument based on measures used in the Adverse Childhood Experiences (ACE) study (discussed later in this chapter). A subset of the original adult cohort (N = 1,437) was followed up 3 years later with the DES, DES-T, and BDI. The adolescent participants completed the Adolescent DES (A-DES), and measures of academic performance, competence, social relationships, and alcohol and drug use.

In the adult study, just as in the Winnipeg study, the investigators found

a skewed distribution of the DES, with about 3.5 percent of the general population fitting the pathological dissociation construct. Pathological dissociation was strongly associated with depression, suicidality, alexithymia, and alcohol abuse. SDQ scores increased in a graded fashion with each ACE, such as childhood intrafamilial violence including sexual abuse, physical abuse in childhood, familial alcohol abuse, and poor relationship between the parents. The combination of high SDQ scores and high DES scores differentiated a specific subgroup with higher depression, suicidality, and many life stresses and problems such as work, financial, and health problems.

At 3-year follow-up, membership in the subgroup with high--dissociation scores was more stable than membership in the pathological dissociation taxon. About a third of the original high-dissociation group still showed high-dissociation scores, but the other two-thirds had been replaced by an equal number of different individuals with DES scores above the cutoff of 20. However, only four of the original individuals who fit the pathological taxon remained in the taxon, although the total number of individuals in the taxon remained the same (n = 64; 4.5 percent at 3 years). Improvement in depression and suicidality was associated with lower-dissociation scores, indicating that worsening dissociation may be a response to stressful events, and greater levels of symptoms in several diagnostic domains. Problems with work performance and younger age also predicted consistently higher-dissociation scores.

In the adolescent sample, higher-dissociation scores were correlated with daily smoking, frequent use of alcohol, abuse of legal drugs, cannabis use, social isolation, and poor school performance. There was no assessment of maltreatment in these adolescents. However, in longitudinal studies of abused children, smoking, alcohol and drug use, social isolation and poor school performance all have been associated with childhood adversity and maltreatment (see discussion of the ACE study below).

In the adults, higher-dissociation scores were associated with ACEs and higher depression, suicidality, and life adversities. These findings are consistent with clinical observations that, in many individuals, overt dissociative disorder symptoms may show a remitting/relapsing course, and only fully emerge during times of life stress, trauma, and/or adversity possibly accompanied by depression, PTSD, substance abuse, and self-destructive and suicidal behavior, all of which are associated with posttraumatic outcomes.

Dissociative Identity Disorder

Across these community-based studies, dissociative identity disorder had a prevalence of about 1 to 1.5 percent in the general population, similar to that of schizophrenia. In the New York and Turkish studies, DSM-IV-TR dissociative disorder NOS was the most frequent diagnosis, 4.3 and 8.3

percent, respectively. In the Turkish study, about half of these NOS individuals (4.1 percent of the whole sample) were described as similar to dissociative disorder NOS Example 1 in DSM-IV-TR, "Dissociative disorder NOS with multiple personality states" (see Nosology). Thus, the improved diagnostic precision with DSM-5 diagnostic criteria may increase the prevalence of dissociative identity disorder and decrease that of DSM-5 other dissociative disorder, example 1.

Dissociative Identity Disorder and Gender in the General Population. Developmental studies indicate that the ratio of female to male dissociative identity disorder cases steadily increases from 1 to 1 in early childhood to approximately 8 to 1 by late adolescence in North American samples. Reasons proposed for the increased numbers of female dissociative identity disorder patients relative to male patients include gender-related differences in the types, age of onset, and duration of maltreatment experienced by men and women; differences in clinical presentations, such that male cases are more likely to be missed; and the possibility that more male dissociative identity disorder cases end up in the criminal justice or alcohol and drug treatment systems, or both, rather than the mental health system.

However, this finding has not been replicated in Turkish studies, or in the WHO study of the PTSD dissociative subtype, where male gender was more prevalent. Sociocultural effects are implicated in these discrepant findings. Western/secular depictions of dissociative identity disorder are primarily based on media stereotypes, including popular first-person autobiographical accounts. Here, the stereotypical dissociative identity disorder patient is a bright, creative, histrionic European American female, with dramatic switching among clearly defined, dramatically different, personality states. Unfortunately, this still represents the clinical picture that many mental health providers expect for a person with dissociative identity disorder. Due to this inaccurate diagnostic framework, most female dissociative identity disorder patients are misdiagnosed, often for long periods of time. Similarly, males with dissociative identity disorder are even less likely to be correctly diagnosed.

Epidemiology of Pathological Possession Form of Dissociative Identity Disorder

Currently there are limited data on the prevalence of pathological possession, and it is unclear whether all reported cases would meet the DSM-5 diagnostic criteria for the pathological possession form of dissociative identity disorder. Cases of pathological possession with dissociative symptoms have been reported in many countries, including China, India, Turkey, Iran, the Republic of Singapore, Puerto Rico, and Uganda. Some case series of dissociative identity disorder in the United States and Canada have described patients who have experienced self-

states as possessing entities, or have been judged as possessed by co-religionists and subjected to exorcisms or similar practices to attempt to expel the personality states. In these cases, the phenomenology of dissociative identity disorder was influenced by these individuals' participation in religious groups that emphasized normative possession (e.g., Pentecostals, Charismatics).

A case series of 117 pathological possession cases in Uganda described a significantly higher exposure to traumatic events on standardized measures, compared with 76 randomly selected mentally healthy inhabitants of the same villages. Scores on these trauma scales were strongly correlated with symptoms of pathological possession and with higher levels of dissociative symptoms as measured by the DES and the SDQ.

In rural Indian communities, the period prevalence of pathological possession has been estimated at 0.97 percent (over 6 months) to 3.5 percent (over 1 year), depending on the region, the sample, and the method of assessment. The period prevalence of culturally congruent, nonpathological possession trance contributed an additional 0.2 to 1.8 percent over 6 to 12 months in the same studies.

In the Turkish study of 628 women in Sivas (previously discussed in this chapter), 20.2 percent of the women described paranormal experiences (e.g., precognition; encountering poltergeists or ghosts) and 2.1 percent reported possession experiences. Women with a dissociative disorder reported all types of paranormal and possession experiences significantly more frequently than those without. One subgroup, 7.8 percent of the sample, accounted for 91.6 percent of the reported possession experiences. They were characterized by the most severe trauma in both childhood and adulthood; the most severe dissociative disorders (e.g., dissociative identity disorder and DSM-IV-TR dissociative disorder NOS); and multiple comorbidities commonly associated with trauma, such as major depression, PTSD, somatic symptom disorders, suicidal and self-destructive behavior, and features of borderline personality disorder. About half of the possession experiences were attributed to a dead or living person, and the other half to supernatural entities (e.g., demon, jinn) or forces. In the non–possession-form of dissociative identity disorder in secular psychiatric settings, patients frequently report the experience of "introject" self-states that literally embody other people in the patient's past or current life. These are sometimes experienced as a benevolent person, but mostly they represent prior or current reported perpetrators. This phenomenon may be experienced as so ego-alien that the person reports that they *feel* possessed. However, these individuals usually do not believe that they are *actually* possessed. Future studies will need to address the question of belief in possession as opposed to the dissociative individual *feeling* possessed.

Epidemiology of Dissociative Amnesia and the Dissociative Fugue Subtype of Dissociative Amnesia

Dissociative amnesia, as defined by the DSM-IV-TR, has been reported in approximately 1.8 (1-year prevalence) to 6 percent (lifetime prevalence) of general population samples, with 7.3 percent of the general population sample of Turkish women in Sivas meeting diagnostic criteria for dissociative amnesia in their lifetime. In these studies, there were low rates of the *dissociative fugue subtype:* 0 percent in Winnipeg and New York State, and 0.2 percent in Sivas. The low prevalence of dissociative fugue in general population studies was one of the factors contributing to the decision to change this from a separate diagnosis to a subtype of dissociative amnesia.

Autobiographical Memory Disturbances for Childhood and Trauma

Childhood autobiographical memory disturbances (CAMD) are a common presentation of dissociative amnesia patients, with a selective, localized, and often, systematic amnesia for early life history (see Table 20–4). This finding has empirical support in data from the ACE study. The ACE study, is a collaboration between the Centers for Disease Control (CDC) and Kaiser Permanente HMO in San Diego, CA involving over 17,000 employed patients of a California HMO. It is one of the largest population studies to assess the relationship between childhood adversities and later life medical, psychiatric, and behavioral problems. In the ACE study, during a comprehensive physical examination, patients filled out a questionnaire about 11 different childhood adversities, and about their current health status. This included questions about recurrent childhood contact sexual, physical, and/or emotional abuse; physical or emotional neglect; witnessing domestic violence; presence of a substance-abusing member of the household; parental mental illness; parental incarceration; and parental absence or loss. The ACE study documented that increasing numbers of ACEs were, in a step-wise fashion, associated with a variety of adverse mental, physical, and behavioral health outcomes (see Tables 20–6 and 20–7 for a list of ACEs and their relationship to a variety of clinical disorders and behavioral problems).

As part of the ACE study, CAMD was ascertained in a cross-sectional sub-sample of 9,460 relatively healthy adults, in relationship to 8 different ACEs. Participants were asked: "Are there any large parts of your childhood after age 4 that you cannot remember?" This question is similar to the mental status questions for dissociative amnesia described in the section on clinical diagnosis. Also, the question is congruent with the DSM-5 Criterion A for dissociative amnesia: "An inability to recall important autobiographical information, usually of a traumatic or stressful nature, that is inconsistent with ordinary forgetting."

Table 20–6.
Types of Adverse Childhood Experiences from the ACE Study

- Abuse
  - Physical abuse
  - Sexual abuse
  - Psychological abuse
- Neglect
  - Emotional neglect
  - Physical neglect
- Household dysfunction
  - Mother treated violently
  - Household substance abuse
  - Household mental Illness
  - Parental separation or divorce
  - Incarcerated household member

Based on data from http://www.cdc.gov/violenceprevention/acestudy/

Table 20–7.
Psychiatric, Behavioral, and Medical Outcomes of Childhood Traumas and Adversities: From the Adverse Childhood Experiences Study and Other Studies (Outcomes marked with an asterisk are from the ACE study.)

Psychiatric
Depression*
Posttraumatic stress disorder (PTSD)
Dissociative disorders
Alcohol and substance use disorders*
Non-PTSD anxiety disorders
Suicidal and self-destructive behavior*
Completed suicide*
Somatic symptom disorders
Eating disorders
Personality disorders
Multiple DSM diagnoses*
Hallucinations*
Autobiographical memory disturbances*
Behavioral
Multi-generational victimization and abuse
Victim of repeated sexual assaults
Violent behavior
Intimate partner violence*—both as victim and perpetrator
Smoking*
Impaired school/worker performance*
High-risk sexual behaviors in women*
Earlier sexual behaviors*
Early pregnancy as either father or mother*—teen pregnancy*
Unexpected pregnancy*

>30 Sexual partners in women*
Increased rate of children with autism spectrum disorders
Medical
Reduced health-related quality of life*
Greater healthcare utilization
Heart disease—ischemic heart disease*
Lung disease—COPD*
Liver disease*
Autoimmune diseases*
Sexually transmitted diseases (STD)*
Fetal death*
Morbid obesity*
Early menarche
HPA-axis dysfunction
Headaches*
Gastroesophageal reflux disease
Irritable bowel syndrome

Based on data from Loewenstein R, Brand B, Gilbert LE, et al. Treating complex trauma survivors. *Psychiatric Times.* 2014;31:40–45.

If respondents reported affirmatively to this question, they were then asked to identify periods in 2-year increments, between the ages of 4 and 18, for which they had memory deficits. Overall, the age-standardized lifetime prevalence for CAMD was 18 percent (N = 1,502); the lifetime prevalence was 15 percent among men and 19 percent among women. Strong associations were observed between CAMD and each ACE. For both men and women, as the number of ACEs increased, the prevalence of CAMD increased in a graded fashion. Overall, adults with an ACE score ≥6 had nearly six times the odds of having autobiographical memory deficits compared to adults with an ACE score of zero. Seventy-seven percent of those reporting CAMD identified memory deficits for ages 4 to 6; 57 percent for ages 7 to 9; 32 percent for ages 10 to 12; 15 percent for ages 13 to 15; and 9 percent for ages 16 to 18.

Sexual and physical abuse in childhood was strongly related to autobiographical memory deficits. For women, the odds of reported amnesia in the sexually abused group was approximately double that of the nonabused group. Among abused men, the odds for CAMD was about 1.5 times higher than for the nonabused group. Sexual abuse frequency and severity were significantly related to autobiographical memory deficits, as was a close relationship between victim and perpetrator. Similarly, the presence of childhood physical abuse significantly increased the amnesia reports. Here again, women and men had, respectively, 2 and 1.5 times the odds of reported amnesia compared to individuals not reporting physical abuse. The combination of sexual and physical abuse was associated with a significantly greater extent of CAMD than a single type of abuse. The effects

of comorbid depression could not fully explain the reported amnesia, as nondepressed individuals with higher rates of sexual and/or physical abuse showed significantly higher rates of CAMD than nondepressed, nontraumatized individuals. The authors concluded that ACEs, particularly sexual and physical abuse, have a specific, overarching effect on autobiographical memory, not just an effect on recall deficits for specific traumatic events. They hypothesized, that, in the individuals reporting memory loss, cumulative traumatic experiences were encoded differently than ordinary autobiographical memory.

Factors thought to predispose to trauma-related amnesia are listed in Table 20–8. However, none of these factors has been shown to exclusively distinguish individuals with amnesia for trauma from those without amnesia.

Table 20–8.
Factors Leading to Dissociative Amnesia after Traumatic Experiences

Trauma caused by human assault rather than natural disaster
Repeated traumatization as opposed to single traumatic events
Longer duration of trauma
Fear of death or significant harm during trauma
Trauma caused by multiple perpetrators
Close relationship between perpetrator and victim
Betrayal by a caretaker as part of abuse
Threats of death or significant harm by perpetrator if the victim discloses his or her identity or information regarding the traumatic experience
Violence of trauma (i.e., physical injury caused by the trauma)
Earlier age at onset of trauma

Note: No factor has exclusively been associated with dissociative amnesia.

Adapted from Loewenstein RJ. Dissociative amnesia and dissociative fugue. In: Gabbard GO, ed. *Treatment of Psychiatric Disorders*. 3rd ed. Vol 2. Washington, D.C.: American Psychiatric Press; 2001:1623–1652.

Delayed Recall for Early Life Trauma

Over 70 studies have confirmed the presence of delayed recall of sexual abuse, physical abuse, or both in clinical, community, and forensic populations, including retrospective, prospective, and longitudinal study designs. In various studies of clinical populations and individuals who identified themselves as survivors of childhood abuse, 16 to almost 100 percent reported amnesia for the abuse at some time in their lives, averaging approximately 30 percent across studies, though at a later point they were able to recall the abuse. Lower prevalence was found in outpatients relative to inpatients. In studies of college undergraduates and women in the community, prevalence of amnesia for abuse ranged from 13 to more than 50 percent.

A random sample of 505 men and women in the general population found that approximately 21 percent reported a history of childhood sexual abuse. Of these, 20 percent described full amnesia, and 22 percent reported partial amnesia for the abuse at some time in their lives. Somewhat lower rates of amnesia for physical and emotional abuse have been reported in clinical and nonclinical samples: about 20 percent of participants endorse this finding across studies.

Longitudinal and prospective studies, primarily following up large samples of individuals with childhood sexual trauma documented in court or medical reports, have found that almost 40 percent of respondents failed to recall a forensically or medically documented episode of childhood sexual abuse when interviewed 20 years or more later, even though, in some cases, other episodes of trauma or maltreatment were recalled. Some respondents with documented trauma denied any history of maltreatment at all.

Studies have compared both delayed recall and continuous memory for sexual abuse in a variety of studies. These have included studies of sexually abused individuals whose perpetrators admitted to the sexual abuse, and many of whom had physical injuries related to the abuse. In these and other studies, researchers have concluded that, when corroboration is obtained, there is *no difference in accuracy between trauma memories that are continuous compared to those with delayed recall. Either type of memory may be essentially accurate, partially accurate, or wholly confabulated.*

Epidemiology of Dissociative Fugue Subtype of Dissociative Amnesia. As noted above, community studies have found almost no cases of dissociative fugue. Dissociative fugue is thought to be more common during natural disasters, wartime, or times of major social dislocation and violence, although no systematic data exist on this point. Most of the pre–DSM-III cases are difficult to assess, because the diagnostic conventions are so different. Most cases in the literature describe men with dissociative fugue, primarily in military samples, although historical studies describe an "epidemic" of apparent dissociative fugue in men in late 19th century France. Dissociative fugue is usually described in adults, although 20th century case reports have described adolescents with fugue behaviors. Dissociative fugue as a disorder, or as symptom of other dissociative identity disorder, may be harder to discern in children, as children may have a much more limited life space in which to wander in an altered state of consciousness.

Prevalence of Amnesia in Traumatized Populations

A variety of studies have looked at the prevalence of amnesia for specific types of traumatic experiences: combat, sexual abuse, physical abuse, and

emotional abuse. Other studies have reported amnesia along with other symptoms of PTSD in various traumatized groups, including survivors of genocide.

Across studies of active duty soldiers in World War II, the average point prevalence of amnesia in the immediate aftermath of the combat, ranged from 5 to 35 percent. Only a small percentage of these soldiers were reported to have suffered significant head injuries as a direct cause of their amnesia. In one study of 1,000 soldiers just after the Battle of Dunkirk, 35 percent of those exposed to the most intense combat had amnesia for the battle. On average, 14.4 percent of the soldiers in this sample reported amnesia. Past or family history of dissociation or hysterical symptoms was associated with amnesia in soldiers with minimal battlefield exposure.

Semi-structured interview data using the SCID-D-R have shown significantly higher amnesia scores in Vietnam-era veterans with PTSD compared to combat veteran controls without PTSD. Studies of Nazi Holocaust survivors, Korean War veterans, and Vietnam-era combat veterans with PTSD and adult survivors of childhood sexual abuse have repeatedly documented deficits in explicit memory on standardized memory tests as compared to matched controls.

Forty percent of a sample of 50 randomly selected survivors of the Cambodian Holocaust endorsed amnesia on PTSD diagnostic inventories. The average DES score in this group of severely traumatized individuals was 37.1. Amnesia was reported in 3 to 10 percent of Nazi Holocaust survivors assessed with a PTSD inventory. The highest prevalence was found in tattooed survivors of Auschwitz. Studies of natural disasters, such as the San Francisco earthquake in 1989, have found that 3 to 5 percent of survivors report amnesia for at least some aspects of these events.

Epidemiology of Depersonalization and Derealization Symptoms

Transient experiences of depersonalization and derealization are common in normal and clinical populations. They are the third most commonly reported psychiatric symptoms, after depression and anxiety. Transient symptoms of depersonalization/derealization in the general population are common, with a lifetime prevalence of 26 to 74 percent and 31 to 66 percent at the time of a traumatic event. A survey of a random community-based sample of 1,000 adults in the rural South found a 1-year prevalence of 19 percent for depersonalization and 14 percent for derealization. Several studies have identified age as a significant factor for transient experiences of depersonalization and derealization, with adolescents and young adults reporting the highest rates in normal population samples, although a recent German community sample did not find such a difference. In one study, 46 percent of college students reported at least one significant episode of depersonalization within the prior year.

Depersonalization/derealization can also occur with illicit drug use,

especially marijuana, hallucinogens, ketamine, and 3,4--methylenedioxymethamphetamine (MDMA or Ecstasy). Studies have found that about 15 percent of depersonalization/derealization disorder cases are triggered by drug use. Transient depersonalization has also been described after certain types of meditation, deep hypnosis, extended mirror or crystal gazing, and sensory deprivation experiences. It is common after mild-to-moderate head injury, where there is little or no loss of consciousness, but significantly less likely if unconsciousness lasts for more than 30 minutes. One study estimated that at least 20 percent of minor head injury patients experience significant depersonalization and derealization. Depersonalization/derealization symptoms also occur in patients with epilepsy, both as a manifestation of the aura as well as during the seizure itself, particularly in complex-partial seizure patients.

Depersonalization/Derealization during and after Life-Threatening Experiences. Acute depersonalization/derealization is also common after life-threatening or near-death experiences, with or without serious bodily injury. Several studies of accident victims have found that up to 66 percent of individuals with a life-threatening experience report at least transient depersonalization during the event or immediately thereafter.

As will be discussed further in the section on etiology, measures of depersonalization and derealization increased significantly in studies of general infantry and Special Forces soldiers after undergoing survival school training including experiences of uncontrollable stress, semi-starvation, sleep deprivation, lack of control over hygiene, and external control of movement, communications, and social interactions.

Depersonalization/Derealization and the Dissociative Subtype of PTSD. Individuals with PTSD who experience depersonalization and derealization symptoms in response to trauma cues appear to be suffering from a distinct dissociative subtype of PTSD (see section on PTSD). Studies have found that 15 to 30 percent of individuals with PTSD will meet criteria for the dissociative PTSD subtype. Individuals with the dissociative subtype have specific fMRI and autonomic nervous system patterns when exposed to personal trauma scripts, including activation of medial prefrontal cortex and dorsal anterior cingulate; decreased activation of amygdala and insula; and no change, or even decrease, in heart rate. Individuals with this subtype have been found in a number of studies, including an international WHO study of 16 countries with over 25,000 participants. In this study the dissociative subtype of PTSD was associated with multiple types of earlier trauma, male sex, severe PTSD flashbacks, role impairment, disability, and suicidality, among other factors.

Epidemiology of Depersonalization and Derealization Disorder

Community-based prevalence of DSM-IV-TR depersonalization disorder ranged from a 1-year prevalence of 0.8 percent in New York, to 2.8 percent lifetime prevalence in Winnipeg. The Turkish study found a lifetime prevalence for DSM-IV-TR depersonalization disorder of 1.4 percent, but also included lifetime prevalence for DSM-IV-TR dissociative disorder NOS Example 2: "derealization without depersonalization" (prevalence 1.1 percent) (see Table 20–4). The DSM-5 depersonalization/derealization disorder diagnostic criteria specify that the presence of either or both of these symptoms are sufficient to make the diagnosis (see section on Nosology). Thus, the actual general population prevalence of depersonalization/derealization disorder based on this sample could be around 2.5 percent, similar to that found in Winnipeg. Further evidence for this comes from a representative sample of the German population (*N* = 2,512) where the 6-month prevalence of depersonalization/derealization disorder was 3.4 percent. In this study, there were no age or gender differences in caseness for depersonalization/derealization disorder. The disorder was strongly associated with depression and anxiety. However, principal component analysis clearly showed that depersonalization/derealization disorder is a separate psychopathological entity, and not just a concomitant of anxiety and/or mood disorders.

STUDIES OF DISSOCIATIVE DISORDERS IN CLINICAL POPULATIONS

The frequency of dissociative disorders in clinical populations varies depending on the type of population studied (e.g., state hospital vs. private psychiatric hospital), assessment methodology, and geographic setting. Most studies use a two-tiered approach, where the DES is used to screen the target population, followed by a diagnostic interview (the DDIS and/or the SCID-D-R for individuals above a DES cutoff score, generally 20 to 30). These studies have been criticized for using relatively high cutoff scores that are more likely to capture individuals with the most severe dissociative psychopathology, such as dissociative identity disorder, not those with dissociative amnesia or depersonalization/derealization disorder. Very few cases of dissociative fugue have been found in any of these studies. In general, across all of these studies, high rates of childhood and adult trauma are significantly associated with high-dissociation scores on the DES and dissociative disorder diagnoses. Types of trauma include sexual abuse, physical abuse, emotional abuse, neglect, medical trauma, adult rape, adult physical assault, intimate partner violence, and traumatic bereavement.

Many of these studies employed the DDIS as the diagnostic interview, which uses a simple "yes–no" format to inquire about the presence or absence of depression, somatic symptom disorders, and features of borderline personality disorder, as well as history of various forms of childhood maltreatment. All of these are common comorbidities in patients

with dissociative disorders. However, Ross, the developer of the DDIS, views the diagnosis of borderline personality disorder on the DDIS as more of a marker of general instability caused by severity of posttraumatic and dissociative symptoms, rather than a reliable diagnostic measure for this personality disorder. Subsequent studies of dissociation in college students in Turkey have supported this view.

Inpatient Studies

In studies of over 2,300 adult and adolescent inpatients from Canada, China, Finland, Germany, Switzerland, the Netherlands, Turkey, and the United States, between 12 and 30 percent of patients scored above the DES screening target score. These studies report a wide range in point prevalence of dissociative disorders, although no study failed to find individuals with dissociative disorders: between 1.7 (China) and 40.8 percent (USA) (mean = 15.8 percent; median = 12 percent). Across these inpatient studies, the frequency of dissociative identity disorder ranged from 0.4 percent of admissions to a Swiss hospital to 7.5 percent of admissions in a private psychiatric hospital in the United States (mean = 3 percent; median = 2 percent). In one Turkish inpatient study, almost half of patients diagnosed with DSM-III-R dissociative identity disorder (then, multiple personality disorder), believed themselves possessed by a jinn, a specific type of possessing entity from Turkish folklore.

The point prevalence of dissociative amnesia in psychiatric inpatients has been reported as 0.2 to 11 percent, depending on the sample, with the lowest rate in an inpatient sample in China. The point prevalence of depersonalization/derealization disorder in psychiatric inpatients ranges from less than 1 to 16 percent, depending on the sample and the method of assessment. Earlier studies failed to differentiate carefully between depersonalization/derealization as a symptom of disorders such as schizophrenia or panic disorder and the disorder itself, and reported much higher rates of the symptom.

A recent systematic study of a mixed group of Swiss outpatients and partial hospital patients, using standardized measures of disability, confirmed the finding of the New York State community sample: marked disability in the dissociative disorder group, with greatest disability in patients meeting structured interview, DSM-IV-TR diagnostic criteria for dissociative identity disorder, or dissociative disorder NOS, Example 1. This study concluded that the dissociative disorder patients fit the construct of *severe mental illness*.

Outpatient Studies

In studies of over 800 adult and adolescent outpatients in various settings from the United States, Europe, Turkey, and China, the point prevalence of

dissociative disorders ranged from 12 to 29 percent (mean = 14.8; median = 13.8), with 2 to 6 percent meeting diagnostic criteria for dissociative identity disorder (mean = 5; median = 2.5). The point prevalence of DSM-IV-TR dissociative amnesia has been between 1.3 and 10 percent of adult outpatients. The point prevalence of DSM-IV-TR depersonalization disorder has been reported to range from <1 to 5 percent of adult outpatients.

In a study of 116 adolescents ages 11 to 17 admitted to a Turkish outpatient clinic who were assessed with the A-DES and the SCID-D for DSM-IV-TR, 45.2 percent had a dissociative disorder, 16.4 had dissociative identity disorder, and 28.8 percent had dissociative disorder NOS (primarily Example 1). There were high rates of traumatic experiences among dissociative adolescents, but not significantly higher than among adolescents who did not meet diagnostic criteria for a dissociative disorder.

Dissociative Disorders in Non-Western/Secular Settings. Saxena and colleagues reviewed chart information on more than 4,000 Indian outpatients; 104 had a diagnosis of a dissociative disorder. Only 5 to 10 percent of these outpatients received a specified dissociative disorder diagnosis according to DSM-III or DSM-III-R. By contrast, 85.5 percent of dissociative disorder cases were covered by the specified dissociative diagnoses in ICD-10. Seventy-four percent of patients had dissociative convulsions, 9.5 percent trance and possession disorders, and 2 percent dissociative movement disorder. This supports the observation that, outside of Western/secular settings, dissociative disorders frequently come to clinical attention due to conversion and/or trance/possession symptoms.

Emergency Department. In a study of Turkish emergency department patients, about a third of 43 consecutive patients screened with the DES subsequently met DSM-IV-TR diagnostic criteria for a dissociative disorder when subsequently administered the DDIS. Of these, 13 percent met diagnostic criteria for dissociative identity disorder, 13 percent for dissociative disorder NOS, and 7 percent for dissociative amnesia. These patients also had high rates of childhood trauma and self-destructive and suicidal behavior, the latter frequently leading to treatment at more restrictive levels of care. Failure to recognize and treat these acute presentations of dissociative disorders may lead to extended periods of costly and ineffective treatment (see section on Treatment).

Dissociative Disorders in Substance Use Disorder Populations

Dissociative disorders and childhood and adult trauma are frequently found in studies of inpatient and outpatient substance abuse treatment programs. This is consistent with the high rate of current or past comorbid substance abuse in patients with dissociative disorders from other clinical

settings. In a series of studies in the United States, Canada, Germany, and Turkey, 985 detoxified inpatients and outpatients in substance abuse treatment were screened for pathological dissociation and DSM-III-R or DSM-IV-TR dissociative disorders with the DES and the DDIS and SCID-D, respectively. Males predominated in the samples. Rates of dissociative disorders ranged from 15 to 39 percent (mean = 24.3; median = 21.6); with 6 to 14 percent for dissociative identity disorder (mean = 8.25; median = 6.5); 0.5 to 27 percent for dissociative amnesia or dissociative fugue (mean = 8.6; median = 3.45); and 0.5 to 16 percent for depersonalization disorder (mean = 4.65; median = 1). The remainder received a diagnosis of dissociative disorder NOS. The wide range of findings in these studies likely reflects differing methodologies. These include: varying DES cutoffs, diverse male/female ratios across samples, differences between DSM-III-R and DSM-IV diagnostic criteria, relative rate of alcohol abuse alone (associated with a lower prevalence of dissociative disorders in all studies), and younger age of participants (associated with higher rates of dissociative disorders).

Dissociative disorder diagnoses were strongly associated with a more refractory course—including more frequent treatment discontinuation, substance use, self-destructive and suicidal behavior, and history of childhood abuse and neglect. Women were more likely to receive a dissociative disorder diagnosis and had more severe and earlier forms of childhood trauma. Emotional abuse emerged as a particularly strong factor in the relationship between dissociative disorders and substance abuse. As in other clinical studies using the DDIS, the patients had high rates of comorbid depression, somatic symptoms disorders, and borderline personality features.

Dissociative Disorders in Samples of Women in Prostitution

In his treatise on hysteria, Briquet compared the prevalence of hysteria in nuns, house servants, and prostitutes. He found the highest rates in prostitutes, due, he speculated, to their traumatic and tumultuous lives, with intermediate rates in house servants, and the lowest in the nuns. This comparison supported his view that hysteria was a brain disease, associated with stressful and traumatic life experiences, and not a disease of disordered sexuality.

In modern clinical studies of individuals with dissociative identity disorder, as many as 20 percent reported working as prostitutes at least at some time in their lives. Contemporary studies of women in prostitution have found high rates of childhood abuse: 65 to 90 percent childhood physical abuse, 74 to 99 percent intrafamilial and/or extrafamilial childhood sexual abuse, and 98 percent emotional abuse. Similarly, women in prostitution suffer very high rates of adult rape, life-threatening physical assault, homelessness, and emotional abuse/degradation by pimps, their

associates, and clients.

Not surprisingly, very high rates of PTSD and dissociation have been found in prostitute samples, with about 70 percent meeting diagnostic criteria for PTSD. Consistent with the pathological dissociation construct, in some studies the DES scores of women in prostitution average >30, indicating severe dissociative psychopathology. Also, 67 percent of 40 women in prostitution (including exotic dancers) met diagnostic criteria for a dissociative disorder: 20 percent with dissociative identity disorder, 30 percent with dissociative amnesia, and 17.5 percent with depersonalization/derealization disorder.

DISSOCIATION AND GENETICS

Research investigating genetic contributions to the development of dissociative disorders is only in its infancy and has yielded mixed results to date. Nonetheless, these studies support an incipient epigenetic model for dissociation, and may eventually help develop genetic and epigenetic models for the development of specific dissociative disorders. Twin studies comparing monozygotic and dizygotic twin pairs suggest that genetic factors may explain approximately half or more of the interindividual variance in dissociative symptoms. The remaining variance was accounted for by "nonshared" environmental experiences, that is, traumatic and/or stressful events.

A handful of studies have linked single nucleotide polymorphisms (SNPs) in genes related to HPA axis, serotonin, dopaminergic function, and learning to dissociation and dissociative disorders. An SNP relating to serotonin function (5-HTTLPR) has been associated with increased pathological dissociation symptoms when coupled with depressive symptoms. In another study, childhood trauma predicted increased pathological dissociation in patients with obsessive-compulsive disorder (OCD) homozygous for the short allele of the *5-HTTLPR* gene. The *FKBP5* gene was associated with dissociative symptoms in women with childhood trauma histories. The *COMT* gene has been linked to increased dissociative symptoms in abused children. Finally, a genome-wide association study of dissociative symptoms in persons with combat-related PTSD identified 10 SNPs potentially related to a genetic basis for depersonalization and derealization symptoms, the most significant among which was r263232 located in the adenylyl cyclase 8 (*ADCY8*) gene on chromosome 8 which is known to be involved in long-term potentiation and synaptic plasticity. This finding could potentially be relevant to understanding specific learning and memory disturbances (e.g., dissociative amnesia) found in dissociative disorders.

The genetics of trait hypnotizability may also be relevant as hypnotizability is thought to be partly a genetically mediated ability. Studies have shown high hypnotizability in patients with dissociative

disorders, with those with dissociative identity disorder having the highest hypnotizability compared with other psychiatric disorders, and nonpsychiatric controls. A gene polymorphism relating to dopaminergic function (COMT Val158Met) has been associated with individual differences in hypnotizability in a number of studies. An interaction between the serotonin (5-HTTLPR) polymorphism, attentional ability, and individual differences in hypnotizability has also been reported.

NEUROBIOLOGY OF DISSOCIATION AND DISSOCIATIVE DISORDERS

Neurobiological study into the nature of dissociation does not fit tidily within the diagnostic system for dissociative disorders, but rather can be summarized more effectively along three major theoretical fronts. The first involves investigations into various trauma-related discrete behavioral states (DBS) and trauma-related altered states of consciousness (TRASC), including depersonalization and derealization, that are thought to present as a regulatory response to fear or other extreme emotions in the context of perceived or actual threat (see section on Theories of Dissociation). These states are consistent with animal models of defensive responding and associated with objective behavioral markers including freezing and tonic immobility, cognitive nonresponse (e.g., blank stare or stupor, "trance-like behavior"), and increased pain threshold. However, TRASC are more adequately defined, following the research paradigm of *neurophenomenology,* by additionally appealing to their subjective experiential elements, which can include the 4-D model of trauma-related dissociation, an alteration in time sense (temporality seemingly "standing still"), flashbacks and "reliving" phenomena, voice hearing, depersonalization, and emotional numbing. Acute, bottom-up, state-dependent, stimulus-response processes most easily explain such states across sensorimotor, emotional, and somatosensory systems, including as peritraumatic dissociative responses and in the context of traumatic reminders. However, certain dissociative disorders can also present as chronically evoked defensive states as such.

A second line of research investigates the biological bases of dissociative amnesia, particularly for traumatic life events. Research is beginning to attribute to disturbances in specific neural systems the failed encoding, failed recall, and ultimate recovery of traumatic memories frequently observed in persons with dissociative disorders. The most commonly implicated systems include hippocampal, amygdalar, and prefrontal modulation of amnestic barriers.

Finally, dissociation has also been described as a division or multiplicity of identity or sense of self. Such identity-state dissociation primarily implies a top-down organization of bottom-up processes via higher-order, super-ordinate "self" representations. Putative neurobiological factors involved in TRASC, dissociative amnesia, and identity-state dissociation

are elaborated below.

Neurobiology of Depersonalization and Trauma-Related Altered States of Consciousness

Several theorists have applied animal models as a basis for understanding dissociative responses in the context of peri- and posttraumatic experience. In general, their descriptions are consistent with the early concept of a "defense cascade" as evoked in sequenced fashion proportional to the physical closeness of a predator and the corresponding likelihood of its attack. After the animal orients toward the threat at a distance, a freezing response is usually elicited wherein the animal, while remaining alert of the predator's movements, keeps still to avoid detection. Flight or fight behaviors may ensue, however, when a predator signals impending attack, for example, by charging. However, if an organism determines it is unable to defend itself, it may assume a defensive posture of *submissive immobility,* which some authors term "freeze–fright" and "death feigning" responses. These defensive maneuvers have been regarded by some as the most likely behavioral counterparts to subjective experiences of dissociative TRASC such as including a distortion of temporal processing (altered time perception), depersonalization, emotional and physical numbing, and subsequent amnesia (i.e., limited encoding) for events taking place during the dissociative episode (e.g., mind "going blank"). Such lower-order defensive responses are thought to be adaptive when reserved for dire circumstances during which active defense (fight or flight) is not possible or would likely occasion an increased likelihood of harm. However, these responses become maladaptive particularly when chronically evoked involuntarily as the primary recourse to repeated trauma exposure, such as in circumstances of ongoing domestic violence and childhood abuse. Different stages of the defensive cascade are associated with different neurophysiological responses: those exciting vigilance and motor reactivity to those promoting motor inactivity, reduced cognitive awareness, and anesthesia. As the neurophysiology of early orienting, flight, and fight responses has been described elsewhere including in the chapters focusing on Anxiety Disorders (see Section 14) and PTSD (see Chapter 17), the authors focus on the immobile, submissive responses here.

Maggie Schauer and Thomas Elbert have suggested that traumatic events that provoke marked psychobiological arousal may lead to a prolonged immobile defensive posture in humans, similar to the animal defense of "playing dead." They hypothesize that this leads to functional sensory deafferentation at the level of the thalamus, gating what are otherwise overwhelming incoming visual, auditory, somatosensory, and proprioceptive information. This survival-based psychobiological process may compromise sensory integration and modulation at limbic (amygdala, hippocampus) and higher cortical centers (e.g., frontal cortex, anterior

cingulate, occipital cortex), that is, dissociation at the neurobiological level.

Jonathan Krystal and others have suggested that if these defensive postures are prolonged, they may present phenomenologically as quasi-catatonic states. However, these states are more appropriately understood as indicating a profound level of helplessness marked by hypoarousal, physical analgesia, and a lowering of level of consciousness. Thus, the lack of sensory/perceptual/affective integration, and the disturbance in consciousness leads to unintegrated, fragmented encoding of events. This could underlie the posttraumatic alteration in memory processes described in the ACE study. Phenomenologically, this is experienced as dissociative amnesia (see sections on Epidemiology and Dissociative Amnesia).

Similarly, Steven Porges's *polyvagal theory* may offer another aspect of the psychobiology of these processes. Porges expanded the classical notion of the relationship between the parasympathetic and sympathetic components of the nervous system during the stress response. He differentiates between two vagal systems in the brainstem medulla; that is, between the ventral vagus that regulates cardiac output as a means of facilitating prosocial engagement, and the dorsal vagus that responds to threat with behavioral immobilization, hypoarousal, and analgesia. The latter may be one mediator of the metabolically conservative defensive stance of the immobile, dissociative state. These freezing behaviors are commonly reported by individuals experiencing extreme trauma where they cannot fight or flee, for example, rape, and/or are in situations of captivity. Individuals with severe early life trauma histories with dissociative disorders often report involuntary immobilization and freezing in response to earlier life danger that persists into their later lives. Consistent with this, studies of autonomic response in dissociative disorder individuals, when hearing personal trauma scripts, show that some of these subjects decrease, or do not increase, heart rate, blood pressure, and related autonomic functions.

Also, hypoarousal has been observed in persons with depersonalization/derealization disorder in the form of lower skin conductance levels (SCLs) and slower response-time when viewing unpleasant stimuli, interpreted as autonomic blunting. Persons with depersonalization/derealization disorder were also unable to increase SCLs with real-time biofeedback.

HPA Axis Function and Depersonalization. In contrast, studies of HPA axis functioning in persons with dissociative disorders have evidenced both hyper- and hypoactivation, although discrepant study findings may be attributable to measurement of urinary versus plasma cortisol, as well as measurement of basal HPA axis functioning compared with measurement in response to stimulus-evoked stress responses. For example, in an initial study of persons with depersonalization/derealization disorder symptoms,

dissociation was unrelated to cortisol levels under basal condition, but there were increases in cortisol levels when subjects experienced acutely stressful stimuli. Some researchers have observed increased plasma and urinary cortisol in depersonalization/derealization disorder compared to individuals with PTSD and other disorders, and greater resistance to, and faster escape from, dexamethasone suppression. However, an inverse relationship between severity of dissociation and cortisol stress reactivity has also been found, complicating interpretation of the data. The latter is consistent with Ruth Lanius and colleagues' conceptualization of dissociation as an "overmodulating" stress response that inhibits reactivity to trauma cues. Variable findings may also reflect the state-dependent nature of stress-response system reactivity in persons with dissociative disorders, including as a function of environmental context, as well as the presence versus absence of trauma triggers.

CNS Response and Depersonalization. In the 1950s, Penfield described "queer sensations of not being present and floating away," with stimulation of the superior and middle temporal gyrus in epilepsy patients. He insightfully theorized that these "illusions of unfamiliarity, strangeness and remoteness" involved an "alteration in the usual mechanism of comparison of immediate sensory perceptions with memory records." At the time, he claimed that such perceptual illusions could be produced by cortical stimulation "only in the temporal region, perhaps extending somewhat into the occipital cortex." Subsequently, in studies of persons without psychiatric illness undergoing neurosurgery direct stimulation of the temporoparietal junction (TPJ) (particularly in the right hemisphere) also produced out-of-body experiences. Moreover, decreased response within superior and middle temporal gyrus has been observed in depersonalization/derealization disorder in various studies, as well as decreased activity in orbitofrontal cortex, medial frontal cortex, left frontal-temporal cortex, and left caudate. In contrast, elevated activity has been found in parietal cortex and left occipital lobe, with certain effects contingent on specific symptoms such as negative affect.

In various social cognitive neuroscience studies, the TPJ and extra-striate body area have been implicated in self-embodiment experiences. Accordingly, dysfunction in these areas figures prominently in neuroanatomical models for depersonalization experiences. Decreased volume of middle temporal cortex has also been observed as a function of suppressing emotional responses in depersonalization/derealization disorder, whereas prefrontal cortices may be increased in thickness, possibly associated with enhanced capacity for emotional suppression. Moreover, recent studies have shown that transcranial magnetic stimulation of the left dorsolateral prefrontal cortex (DLPFC), ventrolateral prefrontal cortex (VLPFC), and TPJ lead to decreased symptoms in

depersonalization/derealization disorder patients. On the other hand, during resting state in healthy volunteers, deep transcranial magnetic stimulation over the medial prefrontal cortex induced depersonalization and a lowering of self-awareness. It was hypothesized that this result was caused by introduction of an error signal to cortical midline structures implicated in self-referential processing by neuroimaging studies, these treatments are experimental, they indicate that medial prefrontal, lateral prefrontal, and temporoparietal cortical regions are involved in depersonalization experiences and in a diminished sense of self-embodiment.

Mauricio Sierra and German E. Berrios put forth the "corticolimbic disconnection hypothesis" of depersonalization/derealization, positing bilateral corticolimbic disconnection with prefrontal activation and limbic inhibition, resulting in hypoemotionality and attentional difficulties. Lanius and colleagues also applied this model to understanding the contrast of a state of hypoarousal and "overmodulation" of emotion, in dissociative PTSD, from the hyperarousal and emotional "undermodulation" typically seen in nondissociative presentations of PTSD. This model has been described as a putative neurobiological basis for the dissociative subtype of PTSD.

Further, two distinct neurobiological components of the depersonalization experience have been proposed, subsumed by distinct neurocircuitry: "visual derealization," associated with occipitotemporal dysfunction, and "body alienation," associated with parietal dysfunction. A review of recent neurological literature provides convergent data for significant roles for the temporal and inferior parietal lobes and other transmodal sensory cortical areas in mediating depersonalization-like experiences. In particular, depersonalization is common in seizure patients, especially in temporal lobe epilepsy with left-sided foci. Inferior parietal and angular gyrus tumors can also manifest with depersonalization symptoms. Structural lesions underlying "neglect" syndromes have been found to be concentrated in the right inferior parietal lobule. Patients with parietal lobe epilepsy frequently report somatosensory aurae, disturbances of body image, vertiginous sensations, and visual illusions overlapping with the symptoms of depersonalization/derealization disorder. Differences in the amplitude of distinct EEG oscillatory frequencies have also been found to be gross markers of different arousal states in persons with dissociative disorders. For example, the slower waves of the theta frequency band (4 to 7 Hz) were associated with increasing dissociative states in a control population, however contrasting results implicating decreasing dissociative symptoms with increasing amplitude of theta frequencies were observed in a psychiatric sample with mixed diagnoses. In summary, disturbances in bilateral inferior frontal, medial prefrontal, cingulate, insula, and temporoparietal cortex are all implicated in engendering experiences of

depersonalization/derealization.

Neurochemical Systems and Depersonalization. Dissociative experiences presenting in the form of TRASC have also been proposed to be dependent on specific neurochemical systems. Such hypotheses are informed by a considerable literature describing dissociative experiences in the context of illicit drug use, including depersonalization/derealization and a distorted sense of time. Drugs that have been reported to produce these dissociative symptoms include marijuana, hallucinogens, opiates, cocaine and its relatives, and dissociative anesthetics such as ketamine (Ketalar) and its relatives. Studies also implicate neuromodulators including serotonin and norepinephrine. The array of implicated pharmacological pathways and their putative range of neurochemical actions would suggest that several neurotransmitter systems might be involved in producing dissociative reactions.

A series of studies using ketamine, an *N*-methyl-D-aspartate (NMDA) antagonist that increases glutamate release, has suggested that dysregulation of the NMDA glutamate receptor may play a central role in dissociative symptoms. Ketamine produces dose-dependent increases in dissociative symptoms including slower perception of time, tunnel vision, derealization, and depersonalization. Excess glutamatergic activation may contribute to structural changes in the hippocampus evidenced in MRI studies of patients with dissociative identity disorder. Pretreatment with a benzodiazepine or with lamotrigine, an anticonvulsant that decreases glutamate release reduces by about half the dissociative effects of ketamine. NMDA receptors are widely distributed and have excitatory functions at hippocampal and cortical association areas. They also promote neuroplasticity and potentiate new learning. Krystal has proposed that NMDA-mediated integration of various brain areas may be necessary for cohesive conscious experience, and therefore that dissociation may involve disruption of corticocortical, thalamocortical, amygdalocortical, and hippocampocortical connectivity.

Perceptual changes, derealization, depersonalization, and other acute dissociative symptoms have also been proposed to have a negative relationship with levels of neuropeptide Y (NPY) which is known to have anxiolytic effects. Charles A. Morgan and colleagues performed a series of naturalistic studies of soldiers undergoing high stress training in preparation for entry into the special forces, 95 percent of whom experienced acute dissociative symptoms as measured by the CADDS. They found that NPY was elevated during these stresses and *higher* in Special Forces soldiers after 24 hours of stress than in general recruits. NPY levels were *inversely* correlated with self-reported dissociation severity. NPY was also correlated with cortisol and more adaptive behavioral responses. NPY is a neurotransmitter/modulator found in high concentration in cell bodies

and terminals of the amygdala. In preclinical studies, NPY has shown anxiolytic and stress-lowering effects. Levels of NPY in Special Forces soldiers were inversely correlated with measures of dissociation, but positively correlated with both cortisol levels and behavioral performance under stress. Twenty-four–hour levels of NPY were at normal baseline in Special Forces soldiers, but significantly lower than baseline in the other soldiers. The findings on this stress-hardy group underscore the importance of psychological adaptation to prior trauma on responses to current stressors. The NPY findings may suggest either a neurobiological protective system that antedated stress, or a neurobiological system demonstrating a more resilient adaptation to stress. NPY may have a role in regulation of anxiety and dissociation in high stress situations. Higher levels of dehydroepiandrosterone sulfate (DHEA-S) in plasma were also associated with lower self-reported dissociation and improved behavioral adaptation to stress.

Norepinephrine is a neurotransmitter central to facilitating alertness, selective attention, and enhanced memory encoding under stress. Several studies support the notion of noradrenergic system involvement in depersonalization/derealization symptoms. Daphne Simeon and colleagues observed that decreased urinary norepinephrine is associated with increasing depersonalization symptom severity. In a small sample with chronic depersonalization, 24-hour urine norepinephrine was found to be strongly inversely correlated to dissociation severity. Similar findings have been reported in a sample with acute peritraumatic dissociation. In another study, administration of a norepinephrine reuptake inhibitor–induced depersonalization, while discontinuing the drug resulted in improvement.

Other studies suggest that depersonalization can be induced by serotonin agonists. Neurochemical challenge studies with the 5-HT_{2C} agonist m-chlorophenylpiperazine have demonstrated the induction of depersonalization in subjects with various diagnoses including social phobia, borderline personality disorder, and OCD, as well as the induction of flashbacks and dissociative symptoms in a subgroup of patients with PTSD. A PET imaging study with the hallucinogenic 5-$HT_{1A/2A}$ agonist psilocybin resulted in increased dopamine in striatum that correlated with depersonalization severity.

Other neurochemical systems putatively implicated in depersonalization and derealization include the cannabinoid system. Cannabinoids, such as marijuana, have been consistently shown to induce depersonalization, with a pronounced component of temporal disintegration. In addition to their action at the cannabinoid CB receptors, cannabinoids block NMDA receptors at sites distinct from other noncompetitive NMDA antagonists. A PET study using intravenous tetrahydrocannabinol (active compound in marijuana) found that cerebral blood flow increase in the right frontal and anterior cingulate correlated

with depersonalization severity, while there was a subcortical cerebral blood flow decrease in amygdala, hippocampus, basal ganglia, and thalamus.

The opioid system has also been implicated in depersonalization/derealization disorder states. The endogenous opioid system mediates stress-induced analgesia. In PTSD patients, analgesic response to combat stimuli can be blocked by pretreatment with the opioid antagonist naloxone. The κ opioid agonist enadoline induces a depersonalization-like syndrome in healthy subjects compared to placebo, with perceptual disturbances and a sense of detachment. Along these lines, opiate antagonists have been reported to reduce dissociation, such as high-dose naltrexone in borderline personality disorder, and intravenous naloxone in chronic depersonalization. The opioid antagonist nalmefene has been reported to decrease emotional numbing in veterans with PTSD.

Dissociative Amnesia

As described in the sections on Epidemiology, and Dissociative Amnesia, traumatic events, particularly cumulative traumatic events during development are associated with a variety of autobiographical memory deficits. However, there are few systematic studies of the underlying neurobiology mediating dissociative amnesia per se. Neurobiological findings in clinical populations have been primarily reported in single case studies or small case series, without adequate diagnostic assessment to rule out dissociative identity disorder, history of trauma, or comorbid disorders such as PTSD.

Nevertheless, neuroimaging studies in healthy persons, and lesion studies in neurological patients, strongly implicate crucial roles for the hippocampus, temporal lobe, TPJ, posterior cingulate, precuneus, and medial prefrontal cortex in the formation and recall of episodic memories. As malfunctioning of these structures has variably been implicated in neuroimaging studies of dissociative disorders across various experimental paradigms, such disturbances may partially account for amnesia for traumatic events and an overall impoverishment of memory for childhood in traumatized persons with dissociative disorders. Deficits in encoding of traumatic events, and/or inhibition of retrieval of memory for the events, are important models to help explain dissociative amnesia. Both processes may be relevant to amnesia symptoms in a specific patient. However, for clarity, data on encoding and retrieval models of dissociative amnesia will be described separately.

Peritraumatic Dissociative Amnesia: Encoding. Mildly to moderately emotionally provocative stimuli are known to be associated with a memorial advantage over neutral cues, partially mediated by response in neural networks involved in salience processing that favor their attention

during encoding, for example, via the amygdala, insula, and anterior cingulate. However, highly threatening experiences, such as traumatic events, may cause an individual to orient *away* from the source of threat in fear (e.g., by averting gaze), or to defensively retreat "inward" (i.e., emotional suppression) or "outward" (i.e., out-of-body experience, depersonalization) as automatic peritraumatic emotional self-regulation strategies. Each of these emotion regulatory responses has been associated with a diminution of salience network processing, which in theory should predict reduced attention and encoding. By parsimony, these studies suggest that impoverished encoding is one neural pathway for amnesia for traumatic events.

Consistent with this interpretation, a recent study provided a striking experimental demonstration of the impairment of autobiographical memory by depersonalization, as evoked by a multisensory perceptual illusion during encoding. Study participants evidenced impoverished recall of scripted life events encoded in an "out-of-body" (depersonalized) condition. This was induced by altering subjects' visual and somatosensory experience to give the false impression that their embodied frame of reference emanated from a place 2 m behind or beside their actual physical position. Participants rated events encoded from this depersonalized perspective as less vivid, and recalled these events more poorly. Moreover, a neuroimaging study showed that impoverished recall was associated with reduced left posterior hippocampal activation in early rehearsal trials, with greater activation in later trials. This is a reversal of the observed and expected effect for encoding under natural conditions. The authors concluded that a dissociative out-of-body experience during an acutely stressful event could, by itself, impair the encoding mechanism and produce fragmented, spatiotemporally disorganized memories.

Neuroimaging has also been used to study laboratory paradigms of active memory suppression or repression. For example, an influential study conducted in healthy persons examined neural response during conditions in which participants were either instructed to think about or "not think" about (i.e., suppress memory of) the second member of a previously seen pair of words. Subsequently, cued recall of suppressed words was inferior to recall of baseline words and even generalized to novel test cues. Neural networks implicated in this suppression response involved bilateral prefrontal cortex, paralimbic, subcortical, and parietal integration areas, and reduced activation bilaterally in hippocampal areas. Moreover, hippocampal activation patterns differed from trials involving ordinary forgetting. These were greater in response to suppression-induced forgetting compared to items where suppression failed: that is, items that were successfully recalled despite attempted suppression.

The authors concluded that activation of a broad neural network was involved in memory suppression, including hippocampus, DLPFC, anterior

cingulate cortex, medial-temporal lobe, dorsal premotor cortex, presupplementary motor area, and intraparietal areas. The authors described a neurobiological model for memory control in response to trauma, where, via encoding effects, these networks are involved in producing lasting amnesia in response to traumatic events.

Neurobiology of Dissociative Amnesia: Disturbed Retrieval. Although no neurobiological studies of individuals with dissociative amnesia have employed large samples, case studies of patients with global dissociative amnesia or dissociative fugue subtype with loss of personal identity have shown unusual metabolic and activation patterns in PET scanning and SPECT studies. A PET study of 14 amnestic patients tested with PET in a resting state showed decreased metabolic rate within right inferolateral prefrontal cortex. These findings complemented those in a neuropsychological case series of persons with amnestic syndromes prompted by traumatic and/or stressful events demonstrating deficient response in right frontotemporal cortex. Other findings included, respectively, reversal of the usual patterns of right and left hemisphere activation during autobiographical and semantic recall. These were similar to findings in individuals with traumatic brain damage and related memory deficits, and those with disruptions or functional disconnections between major neural networks.

Replicated neurobiological case studies of dissociative amnesia have shown deficits in activation of hippocampus and medial temporal lobe, with possible inhibition by frontal systems. Reduced occipital involvement during episodic recall has also been documented. These findings are consistent with fMRI studies of posthypnotic amnesia in healthy subjects that showed reduced response in anterior temporal and occipital cortex, with increased response within left rostrolateral prefrontal cortex. These patterns normalized after discontinuation of the hypnotic suggestion for amnesia. Nonhypnotized participants exhibited the expected habituation effect to repeated presentation of previously encoded cues. Participants in the hypnotic amnesia condition had *decreased* neural response during early recognition trials, as compared to later trials following removal of the hypnotic suggestion for amnesia. These findings implicate deficits in retrieval processes, rather than the deficient encoding in depersonalization and dissociative amnesia.

Hans J. Markowitsch has reported similarities between organic and psychogenic (i.e., dissociative) retrograde amnesia, both thought to be caused by a desynchronization or blockade in neural processing ("mnestic block syndrome") related to altered brain metabolism possibly including changes in various neurotransmitter and hormonal systems, such as GABA-agonists, glucocorticoids, and acetylcholine. Current neurobiological theories of dissociation suggest a possible role of glutamate in at least some

dissociative symptoms, warranting further pharmacological investigation.

Michael C. Anderson and colleagues reviewed a large group of studies that attempt to show that memory can be actively suppressed, either at the point of memory encoding or by retrieval inhibition. Many of these studies utilize the laboratory paradigm of "Think, Not Think (TNT)" instructions: to study the variety of neurobiological processes that may be involved, subjects are instructed to recall or forget word pairs, or whole lists of words. In these studies, similar to the model of frontal systems suppressing amygdala and insula in the dissociative subtype of PTSD, DLPFC and VLPFC appeared to be involved directly in suppression of activity in hippocampus, the fusiform gyrus (perceptual awareness of objects), and occipital lobe (visual perceptual priming). Neural networks that account for the findings in the TNT studies are similar to those found in the fMRI study of posthypnotic amnesia. Also, as discussed below, in PET studies of dissociative identity disorder, similar findings of frontal activation and suppression of neural systems involved in autobiographical memory and self-referential processes occurred in self-states that did not experience personal trauma scripts as autobiographical memory.

Neurobiology of Identity-State Dissociation

In contrast to the literature implicating dysregulation of specific neurochemical systems in the experience of depersonalization and derealization, it is unlikely that any single or combination of a few neurotransmitter systems could underlie the highly complex symptom picture of dissociative identity disorder. These include among others, symptoms found in each of the other dissociative disorders: depersonalization/derealization symptoms, and dissociative amnesia symptoms, both for past as well as current life experiences, as well as dissociative fugue experiences. In addition, individuals with dissociative identity disorder experience symptoms of identity confusion and identity alteration; symptoms of overlap, interference, and intrusion of identities; and spontaneous age regression, analgesia, multimodal hallucinations, and negative hallucinations (lack of perception of sensory stimuli) consistent with spontaneous dissociative, posttraumatic, and autohypnotic processes (also, see section on Diagnosis).

Historically, early medical models attempted to explain dissociative identity disorder by invoking phenomena such as sleep and dreams, hypnosis and somnambulism, epilepsy, and disconnections between the cerebral hemispheres to explain the multiplicity of identity and sense of self that characterizes this disorder. The scientific investigation of identity-state dissociation dates to the experiments of Prince in 1908, who, by way of a crude polygraph to measure GSR, demonstrated distinct responses to emotionally valenced words across alternate identity states in a dissociative identity disorder patient. This basic experimental paradigm, involving a

comparison of biological measures in different dissociative identity disorder patients self-states, continues to dominate the literature to this day. Indeed, over the next century dozens of psychophysiological and cognitive performance studies were published along similar lines. These often lacked methodological rigor, however, either due to small sample sizes single case study design, or deficiencies in experimental control. In contrast the work of Frank W. Putnam and Antje A.T.S. Reinders stands out as a notable exception. Their studies included rigorous control groups as well as study of dissociative identity disorder subjects under a variety of conditions such as hypnotic trance state and states of relaxation, among others.

Putnam's research group has demonstrated unique autonomic system (ANS) reactivity between alternate identity states in dissociative identity disorder patients and dysregulation of the EEG during transitions between identity states. Nevertheless, an important observation in these studies was the instability of both ANS and CNS measures not only across but also *within* different identity-state presentations, the magnitude of which was found to be greater than that observed in nonpsychiatric control groups. Rather than implicating a model of sustained differences between alternate identity states in persons with dissociative identity disorder, these studies found general disorganization in both temporal and spatial integration of ANS and CNS. Consistent with this, beyond the gross measures of amplitude of various bandwidths of the EEG typically employed in early studies, more recent literature is also turning to the investigation of functional connectivity and signal complexity as possible EEG markers of dissociative disorders and the general effects of complex trauma histories.

For example, an EEG study of α-wave coherence using five simulator controls and five dissociative identity disorder subjects found significant differences in the latter group between identity states in six brain regions, whereas there were no differences in the control group. A study comparing ERPs (Event-Related Potentials) elicited by words learned in the same or in different identity states for four dissociative identity disorder patients found little support for the existence of amnesia between personality states when compared to controls from another study who deliberately concealed knowledge of previously learned words. However, given that the ERP is highly relevant to the early detection of salience and basic stimulus processing, the current authors think it unlikely that this very early indicator of cognitive processing would strongly distinguish identity states in dissociative identity disorder. The latter are more likely to be distinguished during poststimulus phases typified by cognitive elaboration and higher-level cognition partially overlapping other EEG markers such as the phenomenon of pre- and poststimulus EEG alpha desynchronization (see subsection on Inter-Identity Amnesia in the section on Dissociative Identity Disorder). A QEEG (Quantitative EEG) study, found EEG-β

activity in frontal and temporal lobes varied between different identity states in dissociative identity disorder patients. Other studies also suggest putative roles for primary sensory and motor cortices, and frontal and prefrontal cortex and nucleus accumbens during transitions between identity states.

Reinders' group has also reported elevated heart rate, higher systolic blood pressure, and lower heart rate variability in persons with dissociative identity disorder in response to audio-recorded personal trauma narratives when individuals experienced the narrative as autobiographical in nature (in Traumatic Identity State [TIS]), as compared with when they experienced the event as non–self-referential in a different self-state (Neutral Identity State [NIS]). These researchers have extended the alternate identity-state paradigm to fMRI neuroimaging. Here, they have found that response within regions of the brain previously associated with "self-referential processing" in controls (thinking about oneself, and evaluating whether a stimulus such as a personality trait descriptor is self-referent), most notably including the medial prefrontal cortex and midline posterior cortex (posterior cingulate, precuneus), has differed between TIS and NIS in persons with dissociative identity disorder. Moreover, these effects cannot be simulated by trained nonpsychiatric controls. Differences from controls have also been observed in regions associated with emotional processing (amygdala, insula) and body representation (TPJ). These researchers have further shown that resting state metabolism and functional connectivity differs between identity states in persons with dissociative identity disorder, indicating a level of internal (in)coherence and (de)stabilization that endures beyond the momentary, stimulus-reactive experimental design employed by most previous studies. Additionally, general neural markers for dissociative identity disorder noted in the literature include Sar and colleagues' findings regarding decreased perfusion in the orbitofrontal region coupled with increased blood flow within medial and superior frontal regions, all associated with self-referential processing in controls; these researchers further identified increased perfusion in occipital cortex. Left temporal involvement in dissociative identity disorder also has been implicated in SPECT studies, which may partially underlie auditory pseudohallucinations (voice hearing) frequently reported in such patients. Strikingly, hippocampal volume was also reported to be 19 percent smaller, and amygdalar volume 32 percent smaller, in patients with dissociative identity disorder compared to healthy controls. Reduced volumes in the parahippocampal gyrus were also reported and correlated with severity of both psychoform and somatoform dissociation symptoms.

The understanding of the neurobiological basis of dissociative identity disorder is growing rapidly. However, a factor that has significantly impeded progress in this domain has been the (at least implicit) concern of

many studies to use neurobiological measures in order to cross-validate the existence of the disorder, as opposed to utilizing research methods in order to better understand the unique neurobiological and psychophysiological profiles of this group of patients. In other words, studies have frequently been designed to reveal biological differences between personality states in dissociative identity disorder, often compared with simulator controls. As with attempts to demonstrate out-of-body states in depersonalization/derealization disorder, the focus of much research has been to replicate the finding that such psychological states are indeed real. This methodology would scarcely ever be expected for inclusion in biological studies of other disorders, for example, depression, schizophrenia, PTSD, or borderline personality disorder. Instead, for other disorders, for the most part, the veracity of self-reported symptomatology is taken for granted, with biological studies then designed to interrogate its etiology, rather than merely serving to discount concerns that the self-reported symptoms might be factitious.

One can argue that further studies like these are unlikely to be productive. Rather, what the literature requires is a greater number of comparative studies, for example, between different dissociative disorders, and between dissociative and nondissociative trauma-related disorders. A dimensional perspective may also be useful in this regard. Frontal, limbic, temporal, and parietal cortices are nonspecifically implicated in both depersonalization/derealization disorder and dissociative identity disorder, and may relate to alterations in self-referential processing, emotional numbing, and felt disembodiment common to both disorders. However specific distinguishing clinical features (e.g., voice hearing in dissociative identity disorder) may require study of specific markers appropriate to differential diagnosis, for example, within auditory association cortex. Longitudinal and naturalistic studies utilizing biological measures will be needed, as well as neuroimaging studies of combined psychological and neurobiological interventions (e.g., neurofeedback, TMS) to better appreciate the clinical significance of the biological markers so far identified for the dissociative disorders, and to help develop more effective treatment for these patients.

ETIOLOGICAL THEORIES OF DISSOCIATION AND DISSOCIATIVE DISORDERS

This section discusses etiological theories about dissociation and dissociative disorders. A comprehensive etiological model involves integrating components of several of these models and theories, as well as their cognitive and neurobiological correlates previously discussed.

Theories of Dissociation

In the literature, the term *dissociation* is used in several different, and sometimes diametrically opposed, ways—one researcher counted over 20. Substantial imprecision can result in both clinical and scholarly discussions unless one is clear about which definition of dissociation is being used (see Table 20–9 for a partial list of the often-contradictory ways dissociation has been used in the literature). Indeed, one set of theories classifies dissociation as a basic capacity of the human nervous system. Here dissociation is viewed as a psychobiological filter that permits focused attention on salient phenomena in the internal and external environment, with less or no attention given to peripheral phenomena, for example, background noises, sensations of one's body while sitting, etc. In some theories, this capacity for focus, as well as relative modularity of psychobiological function, can subserve a variety of normal capacities including that of hypnotizability, for example, Ernest Hilgard's *neo-dissociation theory* of mental functioning and hypnotic phenomena.

Also, a major distinction should be drawn between categorical and dimensional theories of dissociation. In studies of dissociative disorders, particularly regarding their treatment (see under *Treatment* of the individual disorders), dissociation is typically used in the sense of *dimensional, adaptational aspects of dissociation* as a response to extreme trauma, and a subsequent intrapsychic defense. In this model it is not surprising, and may even be regarded as positive, that the frequency and severity of some dissociative phenomena may initially increase as treatment progresses. For example, additional dissociative identity disorder self-states may appear, clarifying the patient's symptoms and eventually leading to more integrated functioning. This increase in dissociative symptoms may sometimes signify a resistance to treatment, and needs to be examined carefully (see under *Treatment*). Under other categorical, purely pathological models of dissociation, any increase in dissociative severity may be understood as a sign of worsening pathology. This will be discussed further in the sections on treatment of the dissociative disorders.

Table 20–9.
Different Conceptualizations of Dissociation

1. Dissociation is described as a *normal* part of the perceptual filtering system of the brain/mind that allows for normal functioning by helping the mind focus on relevant stimuli, and filtering out other stimuli.
2. A fundamental aspect of hypnotic experience, related to the capacity for divided consciousness/attention, and/or separation of cognitive and behavioral schemas (e.g., losing track of driving for several minutes).
3. Theories of mind that describe divided consciousness in normal, hypnotic, and pathological states (e.g., Ernst Hilgard's *Neo-Dissociation Theory*).
4. Continuum Model: *A continuum from normal to pathological,* with normal capacities for absorption, focused attention, and imaginal involvement at one and end, and dissociative disorders at the other.
5. Taxon Model: Two *different dimensional continua of dissociative phenomena* that are categorically different: *normal dissociation* and *pathological dissociation.*
6. A *psychobiological process* that is involved in human response to acute and/or chronic trauma.
7. An intrapsychic defense that is understood in psychodynamic terms; e.g., it protects against recall of traumatic experiences, extreme intrapsychic conflict.
8. In the DSM and ICD systems, dissociation is used to describe *categorical distinctions among psychopathological entities:* the dissociative *disorders.*

Taxon Model of Pathological Dissociation

Recent interest in the *taxon model* resurrects a century-old debate about whether dissociation occurs along a continuum proceeding from normal to pathological or whether pathological dissociation, such as dissociative identity disorder, represents a different type (a taxon) of psychological organization. Janet favored the latter, believing that constitutional factors, suggestibility, and powerful emotional events contributed to creating a group of individuals who were fundamentally different from normal individuals. James and Prince argued for a continuum model, ranging from the so-called "normal dissociative phenomena," such as absorption, to pathology, such as amnesias, fugues, and multiple personalities. The continuum theory carried the day, although some hypnosis researchers continued to caution about possible discontinuities between normal consciousness and some deep trance phenomena.

As data with dissociation measures accrued across clinical and community-based populations, it became apparent that there existed a distinct group of high-scoring individuals. Diverse diagnostic groups contained various percentages of these high scorers, yielding different diagnostic group mean scores. The greater the percentage of high scorers, the greater the elevation of the group's mean score relative to the group's modal score. On the DES, for example, high scorers cluster around a mean score of 45, with the rest of the sample clustering around 8.

These observations led to research on the possible existence of a dissociative type of individual who differs significantly on dissociation measures from other individuals, irrespective of their psychiatric status. Using newly developed taxonometric approaches, researchers examined item-response data for manifestations of a latent class variable. They identified eight DES items (Table 20–10), composing the DES-T (Taxon)

subscale, that robustly differentiate dissociative disorder patients from other psychiatric patients and normal controls, who almost never endorse these items.

Table 20–10.
Dissociative Experiences Scale Items Comprising the Dissociative Experiences Scale Taxon

Subjects are asked to rate the percentage of time, from 0 to 100%, that they have the experience.

3. Some people have the experience of finding themselves in a place and having no idea how they got there.
5. Some people have the experience of finding new things among their belongings that they do not remember buying.
7. Some people sometimes have the experience of feeling as though they are standing next to themselves or watching themselves do something, and they actually see themselves as if they were looking at another person.
8. Some people are told that they sometimes do not recognize friends or family members.
12. Some people have the experience of feeling that other people, objects, and the world around them are not real.
13. Some people have the experience of feeling that their body does not seem to belong to them.
22. Some people sometimes find that, in one situation, they may act so differently compared to another situation that they feel almost as if they were two different people.
27. Some people sometimes find that they hear voices inside their head that tell them to do things or comment on things that they are doing.

From Waller NG, Putnam FW, Carlson EB. Types of dissociation and dissociative types: A taxonometric analysis of dissociative experiences. *Psychological Methods.* 1996;1(3):300–321.

Studies have subsequently affirmed that a taxonic approach is a clinically useful way to identify a subgroup of dissociative patients within any given diagnosis (e.g., the dissociative subtype of PTSD), who differ on symptoms and features from the rest of the group. In psychophysiological investigations of dissociation, simply dividing samples into high and low scorers has identified different physiological and cognitive responses to stimuli that serve as traumatic reminders.

The taxon model implies a different developmental scenario than the continuum model, as well as a different approach to treatment. The typological differences between high dissociators and virtually everyone else would seem to lie in a strong genetic predisposition or a fundamentally different early developmental trajectory, or both, that has long-lasting effects on brain connectivity. Recent research has shown a possible link between dissociation and various genetic substrates (see section on *Neurobiology*). As noted above, a twin study showed possible epigenetic effects, where “unshared environment” (i.e., traumatic events) led to expression of the dissociative diathesis in one twin but not the other. Research linking dissociation with type D attachment disturbances (discussed later in this chapter) offers another potential epigenetic factor

that may mediate the relationship between genetic diathesis and expression of pathological dissociation.

Pathological Dissociation and Treatment Response. In a continuum model of dissociation, a positive treatment response would be conceptualized as simply moving a dissociative individual more toward the normal dissociation segment of the continuum. By contrast, a positive treatment outcome in a taxon model implies a more fundamental modification: changing an individual's type from the dissociative to the nondissociative category. The clinical lore on fusion or integration of dissociative identity disorder personality states into a unified personality hints at this possibility. However, beyond clinical accounts, no empirical data exist to confirm that integration produces a taxonic clinical change. The few studies of dissociative identity disorder that include pretreatment and posttreatment dissociation measures show moderate (but significant) decreases in scale scores rather than a significant taxonic shift from the dissociative to the nondissociative category. The taxonic model has proven valuable in spurring researchers to compare high and low dissociators on a variety of measures, but its clinical usefulness as an index of treatment success remains to be demonstrated.

Dissociation as a Response to Trauma

As noted in the historical section, there is a longstanding debate about the relationship of dissociation to antecedent traumatic experiences. Since the 1980s, multiple lines of evidence have linked dissociative disorders with antecedent trauma, building a strong case for its role as a necessary antecedent to the development of pathological dissociation. The most basic set of these lines of evidence involves quantification of Janet's early clinical observations. For each of the DSM-5 dissociative disorders, there now exist multiple independent case series, including non-Western samples, documenting unusually high rates of trauma in dissociative disorder patients, including several case series in which one or more traumatic events could be verified for most patients.

Several hundred peer-reviewed studies have found significantly higher levels of dissociation, as measured by the DES and other well-validated instruments, in traumatized groups compared to nontraumatized clinical and general population samples. The frequency of this finding for diverse forms of traumatic experience (e.g., combat, rape, natural disasters, child abuse, the Cambodian holocaust, among others) and across cultures indicates the universality of the association between trauma and dissociation. Measures of trauma severity (e.g., combat intensity or rape severity scales) are correlated with measures of both PTSD and dissociation severity to an approximately similar extent, indicating a dose–effect relationship for both conditions. However, the range of the correlation

effect is wide (r = 0.025 to 0.40) and weak to moderate, suggesting that other factors may impact the relationship between trauma and dissociation (and trauma and PTSD), and/or the difficulty in quantifying the subjective aspects of traumatic experience. Although these findings link dissociation with antecedent trauma, they are not sufficient to demonstrate that the traumatic event actually *causes* the dissociation.

More sophisticated statistical approaches are uncovering new lines of evidence indicating that increased dissociation is instrumental in shaping outcomes related to traumatic experience and psychophysiological responses to future stress and trauma. Peritraumatic dissociation measured right after the traumatic event (or described retrospectively) has been found to predict subsequent PTSD, which can help identify individuals at risk for PTSD after acute trauma. Research has also demonstrated that dissociation mediates the relationship between antecedent trauma and subsequent psychopathology, including PTSD. This means that—once the level of dissociation is statistically controlled—the strength of the relationship between trauma and psychopathology significantly decreases or disappears. High levels of dissociation also alter the psychophysiological responses of traumatized individuals to traumatic reminders.

In particular, as noted above, multiple lines of evidence support the existence of a dissociative subtype of PTSD (see section on PTSD). This subtype is characterized by earlier and more severe trauma; more severe PTSD symptoms; a specific autonomic and neurobiological circuitry pattern in response to trauma scripts; and possible SNPs associated with the subtype, among other differences.

Studies in Diverse Cultural Settings. The relationship between traumatic exposure and dissociative disorders has been confirmed in diverse cultural settings. In Uganda, for example, a significant association was found between episodes of pathological possession and early life trauma—assessed with the Harvard Trauma Questionnaire and the Traumatic Experiences Checklist—compared with residents of the same villages without pathological possession experiences. In a representative community-based sample of youths ages 11 to 17 in Puerto Rico, pathological dissociative symptoms on the Adolescent DES were significantly higher among participants who reported severe physical abuse and/or were exposed to violence. Compared to individuals with conversion but without a dissociative disorder, individuals in Turkey who also met criteria for a dissociative disorder were more likely to endorse childhood trauma.

At the same time, international studies have noted that some forms of dissociation—pathological possession, psychogenic nonepileptic seizures, conversion disorder with dissociative features, dissociative cultural idioms of distress such as *ataque de nervios* (attacks of nerves)—can be associated

with significant but nontraumatic stressors. For example, these include school and family conflicts among youth, romantic affairs, emotional neglect, unemployment, and insecure attachment. However, not all of these studies systematically assessed for the presence of childhood trauma, which could play a role in predisposing individuals to dissociative disorders once exposed to these nontraumatic stressors. This is similar to descriptions of precipitants to episodes of severe dissociation in the classic cases and case series on generalized dissociative amnesia and dissociative fugue. Here, extreme intrapsychic conflict and/or acute loses or stresses appeared to precipitate most of these events. However, in almost every study, careful reading shows that these individuals had suffered from severe prior traumas such as wartime combat, experiences of violence or sexual assault, and/or childhood sexual abuse and/or physical abuse. Some cultural characteristics—such as the avoidance of strong negative affectivity to maintain equanimity and the nurturing of a more "porous" sense of self open to transpersonal spiritual experiences—may nurture dissociative capacity, which may develop into pathological dissociation upon exposure to severe but nontraumatic stressors.

Studies of Military Survival School Subjects. Remarkable research conducted by Morgan and colleagues with military Special Operations units under extremely stressful conditions are as close as one can ethically come to an experimental model of trauma-induced dissociation. Working with units undergoing forms of simulated combat, such as survival school, Morgan examined the neurobiological and psychological effects of intense stress uniformly applied to drug-free, healthy, nonclinical subjects. Stressors included semi-starvation, exhaustion, sleep deprivation, lack of control over hygiene and bodily functions, and lack of control over one's movements, social contact, and communication. The neuroendocrine stress effects measured in these soldiers were equal to or exceeded those reported for individuals subjected to life-threatening experiences. These stressful experiences were positively associated with dissociation measures. Significant differences were found between pre- and posttest scores on the CADSS overall and for virtually all scale items, with the greatest effects for depersonalization items (e.g., looking at the world through a fog, time slowing, "spacing out"). Life threat from a prior trauma was significantly correlated with pre- and poststress dissociative symptoms. Dissociation scores also accounted for 41 percent of the variance in health complaints after the stress experience. The significant relationship between the pre- and poststress change in dissociation score and somatic symptoms is congruent with many clinical studies reporting linkages between somatization and dissociation.

Morgan's systematic studies are consistent with observational reports of the high frequency of dissociation, especially depersonalization symptoms,

in normal individuals exposed to life-threatening stress. The ubiquity of acutely high levels of dissociation during extreme stress indicates that other factors, perhaps prior life-threatening trauma, are operative in the peritraumatic dissociation–PTSD linkage discussed previously. The highest-dissociation scores were associated with intense subjective appraisal of life threat from past trauma. However, one highly traumatized subgroup, Special Forces soldiers, had *low life-threat appraisals of their past trauma* and, consequently, the *lowest pre- and poststress change in dissociation scores*. As described in the neurobiology section, after interrogation stress during survival training, Special Forces soldiers showed significantly higher levels of Neuropeptide Y (NPY) compared with soldiers in the general forces, suggesting a role of NPY in regulation of anxiety and/or dissociation during high stress situations.

In summary, several independent lines of evidence point to a causal relationship between trauma and dissociation. This relationship is apparent for many forms of trauma and has been found across various cultures in which it has been investigated. Significant levels of dissociation in the immediate aftermath of a traumatic event appear to predispose an individual to developing PTSD and related psychopathology, although additional factors involving the psychological appraisal of prior trauma also appear to be involved. Dissociation may serve as a major mediating process between traumatic experiences and subsequent psychopathology and thus is an important target for prevention and early intervention efforts. Stress-resistant individuals can be characterized, in part, by their lack of dissociative responses to significant stressors.

Iatrogenic, Sociocognitive, and Fantasy Models of Dissociation and the Dissociative Disorders. Despite research data on the relationship between trauma and dissociation, some authorities continue to question the existence of dissociation, the trauma–dissociation relationship, and/or the validity of the dissociative disorders, particularly dissociative identity disorder and dissociative amnesia (see section on History). This skepticism has taken three forms: *the iatrogenic model* (IM), the SCM, and finally, the *fantasy model* (FM*)*.

IATROGENIC AND SOCIOCOGNITIVE MODEL OF DISSOCIATION. The IM posits that dissociative identity disorder and dissociative amnesia are not authentic psychiatric disorders, but instead the product of direct or indirect cuing and suggestion during psychotherapy by credulous therapists to highly "suggestible," hypnotizable, and "fantasy-prone" patients. These are supposed to be characterized by dependent, borderline, and/or histrionic traits, and to be desperate for acceptance and a sense of identity. The therapist, through an ideological belief in dissociative identity disorder or "repressed memory," engages in a "risky recovered memory therapy" (a

nonexistent form of treatment; see section on Treatment) that focuses only on retrieval of as many trauma memories as possible using "suggestive techniques." The therapist thereby implicitly and/or explicitly encourages the patient to undertake the role of the trauma survivor with recovered repressed memories and, in some cases, dissociative identity disorder. The patient develops a real, albeit iatrogenic, factitious disorder, now truly believing in the roles that he or she enacts.

Treatment requires removing the patient from the therapy, ignoring any reference to self-states or trauma memories, educating the patient about the factitious nature of these phenomena, and beginning "real" therapy for other issues. If the patient has become estranged from his/her family of origin due to belief in the false abuse history, reunification with the family is deemed important as well.

The SCM is similar to the IM, but adds a factor related to the broader cultural and social influence—often promulgated by the internet, mass media, self-help books, and/or victim support groups—in which people, primarily in North America, come to accept the possibility of *multiple identity enactments* or a history of forgotten childhood victimization, or both, as part of an established social role. In this view, therapy may not be a necessary factor in the development of a dissociative disorder; a susceptible person may develop the disorder simply through more general social influence. The SCM for dissociation is constructed as an analog to the Sociocognitive Model of Hypnosis, where hypnotizability is not related to an altered state of consciousness or discontinuities in experience, but solely is a learned social role, that of "being hypnotized."

Recent comprehensive reviews of the IM/SCM for dissociative disorders show that virtually no empirical, scientific studies have been performed in clinical populations to examine this model. The original database supporting these constructs consists of a few heterogeneous studies from the 1940s and 1950s, mostly consisting of anecdotal case reports. These demonstrated that some degree of role enactment of an alternative identity occurred with hypnosis, automatic writing, and strong repeated suggestion. Some of the subjects were the students of the researchers and were knowledgeable of the goals of the experiments. The role enactments were limited to the experimental situation. One of the researchers, studying a traumatized soldier, concluded that hypnosis allowed access to authentic, previously dissociated self-aspects, not to artifacts of the procedure.

Research with thousands of subjects undergoing hypnosis over several decades has shown that a subgroup of highly hypnotizable individuals can exhibit limited, temporary apparent self-states after hypnotic suggestion. These states were phenomenologically dissimilar from clinical cases of dissociative identity disorder, and were limited to the experimental situation. The researchers noted that they could not determine whether these procedures uncovered previously existing dissociated aspects of the

mind or created them de novo.

A few studies of college volunteers in various hypnotized and nonhypnotized conditions found that highly hypnotizable, "fantasy-prone" students would enact the role of a dissociative identity disorder murderer with amnesia when given a series of robustly explicit suggestions to do so. This research has been criticized on several grounds, including lack of diagnostic controls (e.g., enacting another psychiatric disorder, such as a murderer with schizophrenia); lack of equivalence of the research design with the clinical situation; conflation of dissociative identity disorder with role enactments, rather than the complex polysymptomatic picture described in the clinical and research literature; lack of attention to strong expectancies and demand characteristics of the research design; and failure to account for preexisting dissociated self-states and traumatic experiences. Finally, the large, international epidemiological studies of dissociative disorders, where clinical and nonclinical samples of dissociative disorder individuals have been identified by validated screening instruments and diagnostic inventories, cannot be accounted for by the IM/SCM postulates (see section on Epidemiology).

The Fantasy Model. The most recent iteration of this skepticism is the FM. In this model, dissociation is not a phenomenon that follows traumatic or overwhelming experiences. Instead it is a personality trait consisting of a mild form of cognitive/executive function impairment that predisposes an individual to being suggestible and generally prone to fantasy, and to have a specific liability to confabulated traumatic experiences. FM proponents aver that the relationship between trauma and dissociation is weak or nonexistent. Being fantasy prone was originally described in nonclinical samples, seen in individuals with high hypnotizability, vivid fantasy lives, and a capacity to spontaneously create intense perceptual, sensory, and imagery experiences, often more vivid than actual reality. Fantasizers also described vivid recall of childhood and multiple religious and paranormal experiences.

Systematic Comparison of the Trauma and Fantasy Model of Dissociation

These forms of skepticism have been criticized, especially most recently with respect to the FM. First, a general critique of the literature on fantasy-proneness posits that the basic psychometric inventories for fantasy-proneness do not measure a specific construct, but rather diverse factors apparently related to potential clinical and normative phenomena, most of which are unrelated to being fantasy prone. Second, fantasy-proneness and dissociation measures have been found to overlap. Third, some studies have found a relationship between fantasy-proneness and childhood trauma, potentially explaining the correlation between dissociation and fantasy-proneness measures.

In 2012, Constance Dalenberg and colleagues published a comprehensive, rigorous meta-analysis that examined eight contrasting predictions from the TM and the FM of dissociation. These findings are summarized in Table 20–11.

Table 20–11.
Eight Predictions of the Trauma Model versus the Fantasy Model of Dissociation

In contrast to predictions associated with the Fantasy Model, The Trauma Model for Dissociation was confirmed on the following eight predictions:

1. The relationship between trauma and dissociation was consistent and moderate in strength across studies.
2. The relationship between trauma and dissociation remained significant when proven, well-supported, and/or objective measures of trauma were used, not just self-report.
3. Dissociation was temporally related to trauma and trauma treatment; symptoms tended to decrease with time, and were responsive to trauma treatment, rather than inflate and increase with treatment.
4. Dissociation was predictive of trauma history when fantasy-proneness was controlled, but not vice versa.
5. Dissociation was not reliably associated with suggestibility on a series of suggestibility paradigms.
   a. Event suggestibility
      i. For nonautobiographical events (post-event misinformation)
      ii. For autobiographical events ("false memory" creation)
   b. Interrogatory suggestibility (IS) (liability to false confessions)
   c. Imagination inflation
   d. Source monitoring errors
6. There was no relationship between dissociation and inaccuracy of documented memory with delayed recall ("recovered memory"); i.e., *both* continuous and delayed recall memories may be accurate, partially accurate, or confabulated.
7. Dissociation negatively related to general measures of narrative cohesion, indicating that dissociation led to difficulties in recall, not inflation of false trauma accounts.
8. Neurobiological research supported the trauma theory of dissociation as a regulatory response to fear or other extreme emotion with measurable biological correlates.

Adapted from Dalenberg CJ, Brand BL, Gleaves DH, et al. Evaluation of the evidence for the trauma and fantasy models of dissociation. *Psychol Bull.* 2012;138(3):550–588.

This review only included studies with sufficient data to calculate effect sizes, sample sizes greater than 50, and whose design was unlikely to restrict the range of the trauma or the dissociation variables. A particularly strong point was a review of the relationship between various suggestibility paradigms and dissociation. Major reviews of the term *suggestibility* cite the multiple definitions and explanations that have been used both for the phenomenon of *suggestion* and the hypothesized trait of *suggestibility*. These have included studies of hypnotic phenomena, unconsciously produced motor behaviors, possession states, false confessions, susceptibility to advertising, responses to psychotherapy, behavioral

conditioning, and memory distortions, among many others. Also, one type of suggestibility may not correlate with other types.

In the Dalenberg et al. review—across a variety of suggestibility paradigms, including those hypothesized to etiologically generate false trauma memories—*dissociation was not reliably related to suggestibility*. Indeed, the average *weighted effect sizes between all forms of suggestibility and dissociation accounted for only 1 to 3 percent of the variance*. Consistent with previous research, the review suggested that a subgroup of dissociative individuals may be more prone to generation of confabulated memory, although studies have not distinguished between false memory, false belief, and acquiescence to an authority figure in giving responses. In addition, studies have not tested *false nonbelief,* that is, subjects who come to deny actual memories (i.e., parallel to false trauma recantation) when suggested to do so by an authority figure. In response to the Dalenberg review, advocates of the IM/SCM/FM models, in a major concession, acknowledged for the first time that some delayed recall of trauma memories may be veridical.

Based on the evidence for the eight predictions of the TM of Dissociation, Dalenberg et al. concluded that there is *strong empirical support for the hypothesis that trauma causes dissociation,* and that dissociation remains related to trauma history after adjusting for fantasy-proneness. They found little empirical support that the dissociation–trauma relationship is due to confabulated memories of trauma, various types of suggestibility, and/or fantasy-proneness.

Dissociation, Dissociative Identity Disorder, and Interrogatory Suggestibility. In a subsequent study, 92 adults recruited from the community were studied with measures of fantasy-proneness, dissociation, trauma history, and the GSS, developed to assess *Interrogatory Suggestibility* (IS), the liability to false confessions. The GSS includes subscales assessing the respondent's liability to yield to leading questions and to shift (confabulate) responses under pressure from the examiner. Participants' dissociation scores correlated significantly with trauma history and fantasy-proneness, but not with GSS scores. Moreover, dissociation was negatively correlated in a sub-study with recall of the events of 9/11. In another study, hospitalized women with recovered childhood sexual abuse memories were significantly less suggestible on the GSS compared with general psychiatric controls from the same hospital.

These data are consistent with preliminary findings from studies of dissociative identity disorder that showed lower GSS scores compared with subjects with PTSD, trauma controls, and individuals with borderline personality disorder. Dissociative identity disorder subjects were *more* resistant than the other groups to interpersonal persuasion. A different study of SCID-D–diagnosed dissociative identity disorder patients failed to

confirm the basic clinical profile posited by the SCM and replicated the finding that dissociative identity disorder patients did not show elevated interrogatory suggestibility. In particular, the dissociative identity disorder patients were highly introversive on the Eysenk personality inventory and showed low histrionic personality features on the Million Multiaxial Clinical Inventory. Most dissociative identity disorder respondents were low or moderately high on fantasy-proneness, with a subgroup showing high scores. As described above, this may indicate that there is a subgroup of dissociative identity disorder individuals with high fantasy-prone scores who may have more difficulty with reality testing and possibly greater liability to forms of suggestibility and inflation of trauma reports.

Culture, Dissociative Disorders, and Sociocognitive Phenomena

Cultural and sociocognitive factors shape the presentation of all psychiatric disorders across cultures. Certain symptoms may be emphasized above others when experiencing or describing disorders (e.g., psychological vs. somatic aspects of depression) and specific experiences may be common in certain settings but rare in others. For example, conversion-like dissociative symptoms are more common in India than in certain affluent settings in the United States. Attributions of pathological experience also vary. For instance, 19th century individuals with paranoid schizophrenia did not have delusions of being controlled over the Internet.

The inclusion in the DSM-5 of the pathological possession form of dissociative identity disorder is based on cross-cultural research that describes sociocognitive differences in the structuring of the disorder between secularized, Western cultures and nonsecularized (often) non-Western cultures. In the literature on dissociative identity disorder, various developmental, social, cultural, intrapsychic, interpersonal, and cognitive factors, including psychotherapy, are hypothesized to account for the *secondary structuring* of the dissociative process into the individualized identities characteristic of a given person in a given culture. Influences may include family factors (e.g., contradictory parental demands for the child to be a boy or a girl); religious, cultural, and social influences (e.g., characters from books, films, television, or video games); the accepted existence of nonmaterial entities (e.g., gods, demons, possessing spirits); and specific traumatic experiences (e.g., being sexually trafficked and having to create identities to fit the fantasies of "purchasers"); among others.

In clinical case series, developmental factors affect the presentation of the dissociative identity disorder self-states in important ways. Usually children with the disorder manifest less overt switching, and self-states are described more as elaborated and autonomous imaginary companions. During adolescence, the person may develop more complex, secondarily structured self-states that relate to many aspects of psychological life. Intelligence and creativity may lead to greater elaboration of identity states

in some individuals with dissociative identity disorder, as does posttraumatic withdrawal from outside reality due to trauma and neglect; these may lead to the creation of more elaborate fantasy worlds in which to "live." Cross-cultural research could help describe childhood forms of pathological possession in cultures that support these belief systems. Future research may trace in greater detail how culture interacts with biology in a developmental timeframe and in the context of an individual's exposure to particular experiences to shape the expression of culturally diverse phenomenologies of dissociative psychopathology (see Developmental Models of Dissociation).

Discrete Behavioral States Model

The DBS model postulates that dissociative disorders belong to a group of psychiatric conditions characterized by rapid—often environmentally triggered—discrete shifts in states of consciousness. Examples include rapid-cycling bipolar disorder, panic disorder, periodic catatonia, and PTSD. Although these disorders differ considerably, a central feature of their pathology is the abrupt switch between two or more distinct states of consciousness, at least one of which is pathological. Panic attacks and the abreactions and flashbacks of PTSD are examples of dysfunctional behavioral states, often triggered by specific stimuli or interactions that rapidly reorganize an individual's emotions, thinking, and behavior. Bipolar patients often cycle through a spectrum of distinct affective states extending from deep depression to psychotic mania.

The concept of *state of consciousness*—as a distinct unit of behavioral organization—has been part of the psychological vocabulary for several centuries. Ribot's speculations that state-dependent learning and memory explained the phenomena seen in dissociative identity disorder arose in the context of widespread interest and clinical experimentation with hypnosis, somnambulism, and trance-like states. In his 1896 Lowell lectures on exceptional mental states, James often referred to altered states of consciousness as underlying multiple personality, fugues, conversion symptoms, and other hysterical symptoms. Modern state theory owes much to psychologist Charles Tart, who advocated a science of discrete states of consciousness. Hilgard's neodissociation theory, which he traces in part to Janet, conceptualizes hypnosis as a set of altered states of consciousness characterized by increased suggestibility, enhanced imagery, decreased planning capacity, and reduced reality testing. Research with drug-induced altered states and meditation has also contributed to current theory and understanding.

It remained, however, for a group of child psychologists and psychiatrists to formally operationalize the study of DBSs of consciousness. Heinz Prechtl, Robert Emde, Robert Harmon, and Peter Wolff empirically validated a multidimensional classification system used to study normal

infants. By using only four to five variables (e.g., heart and respiratory rate, extremity muscle tone, skin perfusion, vocalization, facial expression), researchers could delineate a basic set of DBSs shared by most full-term, newborn infants. With maturation, additional behavioral states are added, and the pathways of connections among the distinct states grow more complex—corresponding to an increased behavioral repertoire on the part of the child. The ability to self-modulate the expression of these DBSs and to integrate the many additional states created with growth and development is regarded as fundamental to healthy emotional regulation.

The DBS model conceptualizes dissociative phenomena as the alternation of distinct states of consciousness that differ in terms of the individual's sense of identity, access to explicit, implicit, and autobiographical memory, and psychophysiological reactivity. Identity disturbances are the nucleus around which other features of the dissociative state organize in much the same fashion as distinct affects serve to organize the DBSs of a bipolar patient. Dissociative disturbances of memory are viewed as extreme examples of state-dependent learning, storage, and retrieval. Alterations in psychophysiological sensitivity in dissociative disorders reflect state-dependent changes in autonomic nervous system and neuroendocrine function. Changes of equal magnitude have been measured in laboratory studies of panic attacks and PTSD flashbacks.

The DBS model of dissociation postulates that early childhood trauma, disturbances in attachment, or both, disrupt the processes essential to self-modulation and the developmental integration of new behavioral states. The DBS model postulates three levels of psychopathology inherent in such *state* disorders. The first is intrinsic to the nature of the pathological states per se, for example, anxiety, depression, mania, catatonia, or depersonalization. The second level results from the individual's inability to self-modulate the expression of these dysfunctional states or to better integrate them into his/her life. The third level results from the consequences of maladaptive attempts to block the often painful experiences of these pathological states (e.g., attempts to self-medicate depression or social withdrawal to avoid triggering flashbacks). The increased psychiatric comorbidity commonly noted in these patients is due to the second and third levels of psychopathology. The DBS model predicts that effective treatment should facilitate better self-modulation of intense or dysphoric states and address the state dependency of dysfunctional perceptions, cognitions, and assumptions.

In the behavioral state model, dissociation is conceptualized as basic to the psychobiology of the human response to life-threatening danger: a protective activation of altered states of consciousness that change perception, pain sensation, time sense, and sense of self. Trauma-related memories and affects are encoded during these altered states. Upon return

to baseline, the person has less access to the dissociated information, which can lead to dissociative amnesia for at least some part of the traumatic events. However, the dissociated memories and affects can manifest themselves in nonverbal forms, as posttraumatic nightmares, reenactments, intrusive imagery, and somatoform symptoms. Not only is there amnesia for the traumatic event, but the person also frequently has dissociated the fact that basic assumptions about the self, relationships, other people, and the nature of the world have been altered by the trauma.

This model does not posit that trauma is forgotten nor does it necessitate psychic pain as the controlling variable in the onset of amnesia, although amnesia can inhibit retrieval of dissociated information. Rather, it suggests that trauma is encoded and remembered but remains verbally inaccessible owing to difficulty in retrieving information across different DBSs. Some authorities have made the analogy between posttraumatic dissociative amnesia and experimentally produced hypnotic amnesia (see section on Neurobiology). In both cases, implicit memory for the material can usually be demonstrated. Following Janet, some have hypothesized that amnesia arises from a decrease in executive control or synthesis among sets of psychic or memory elements and from inhibition of control elements that can give access to information stored in memory.

4-D Model of Trauma-Related Dissociation. The "4-D Model of the Trauma-Related Dissociation" has been articulated by Paul Frewen and Ruth A. Lanius. This framework, based on work from philosophy and consciousness studies, elaborates on the concept of TRASC introduced by the DBS model just described. The 4-D model proposes a four-dimensional neurophenomenology as a means of describing specific experiential states: (1) time and memory, (2) verbal/thought, (3) embodiment, and (4) emotion. The neurophenomenological approach is based on subjects' personal phenomenological descriptions of experience that are then correlated with neurobiological measures. Imaging studies of individuals with PTSD and in community samples have shown differences in fMRI patterns between individuals with and without dissociative symptoms, irrespective of diagnosis. Those without dissociative symptoms have shown patterns of distress associated with the neurobiological normal waking state of consciousness (NWC), as opposed to TRASC, characteristic of distress in individuals with the dissociative subtype of PTSD and with dissociative experiences related to trauma. The 4-D model, consistent with the developmental DBS model, provides a testable model that links neurocircuitry patterns to abnormalities of emotion regulation and self-referential appraisal. These, along with other factors such as attachment disturbances (see sections on Betrayal Trauma and Type-D Attachment) may be important in development of dissociative disorders such dissociative identity disorder and depersonalization/derealization disorder.

Betrayal Trauma

Beyond the multiple associations among various types of traumatic experiences and dissociative sequellae, Jennifer Freyd has recently formulated a specific *Betrayal Trauma Theory* (BTT), derived from social psychological principles and developmental attachment theory. BTT began as an attempt to explain childhood amnesia for trauma as related to the meaning of the negative event as a betrayal by a trusted, needed other, rather than simply being related to the frequency and aversiveness of traumatic experiences. This betrayal is thought to influence how the event is processed and remembered. For example, in this model, a child is more likely to develop amnesia when subjected to abuse by a family member or another person whom the child trusts or on whom the child is dependent, compared to a stranger. The amnesia is protective of the child's developmentally mediated need for attachment, allowing for preservation of emotional and cognitive growth despite the abuse. In order for this to occur, it is hypothesized that psychological schema about the abuse, that is, the betrayal, are subjectively sequestered from mental mechanisms that control attachment and attachment behavior. BTT is consistent with the DBS model and elucidates additional factors that may influence dissociation of state-related information.

In this model, difficulty in being able to "share" the traumatic experiences is also thought to contribute to amnesia for trauma. In addition to the neurobiological factors that inhibit verbal access to such events, they are less likely to be shared with others because of threats, shame, developmental inability to find words for the experience, reliance on adults to define childhood experience, and intense confusion and emotional arousal engendered by the abuse. It is theorized that traumas involving betrayal are processed differently at the cognitive level from other autobiographical experiences, leading to subjective memory deficits. BTT also helps explain the lower prevalence of amnesia for wartime trauma, natural disasters, and the Nazi Holocaust, among others, as compared to intrafamilial childhood sexual or physical abuse.

Several studies have supported the betrayal trauma model. Among undergraduates, abuse by a caretaker was a better predictor of memory impairment for the abuse than its duration or age at the time of the abuse. Stroop tests have shown that, compared to low-DES subjects, high-DES subjects who report more trauma in general, and more betrayal trauma specifically, may actively inhibit threatening information but not neutral information from conscious access. Betrayal trauma and lack of integration of betrayal information—called *betrayal blindness*—have also been studied across diverse types of experience. These include personal betrayals, such as marital infidelity, and cultural and institutional betrayal in situations of racial prejudice and/or lack of institutional support for victims of trauma

(e.g., military sexual trauma, campus rape). Betrayal blindness may be a common human response to these situations, with intense emotional reactions when awareness of the betrayal becomes more integrated and less dissociated from the person's awareness. Research has shown that institutional betrayal trauma contributes significantly to adverse responses to interpersonal trauma, including rape.

A unified information processing and betrayal model predicts the relative likelihood of amnesia under varying circumstances, such as high threat/low betrayal, high threat/high betrayal, and so forth. However, as described above, various factors may lead to disturbances of autobiographical memory for traumatic experiences, although betrayal trauma may be particularly salient as a factor in the lack of integration of trauma-related information. BTT may also help explain liability to revictimization in survivors of childhood trauma. When the survivor encounters potentially exploitative individuals, the "betrayal detection system" becomes relatively unavailable to consciousness. This, as well as the over-learned freeze response to threat developed in childhood (see section on Neurobiology), may explain how dissociation is a risk factor for revictimization.

Type D Attachment

Type D attachment among children is characterized by disorganized and conflicting movement patterns when the primary caregiver returns to the room after a period of enforced separation during the *strange situation,* a standardized procedure for assessing attachment in preverbal children. The child may exhibit contradictory intentions, lack of orientation to the environment, and sudden immobility associated with a dazed expression or a trance-like state, termed *stilling* in the attachment literature. Type D attachment disturbances are common in maltreated infants and toddlers and are thought to result when the child periodically experiences the primary caretaker as frightening or frightened, or alternating between the two. The best predictor of type D attachment is a high DES score in the mother. It is hypothesized that an infant with multiple incompatible models of self and other resulting from intermittent frightening experiences with the caretaker would rapidly switch back and forth between attachment models when confronted with a stressful interaction involving the primary caretaker, such as the strange situation.

In a prospective longitudinal study, 168 high-risk children were followed over a 19-year period to examine whether trauma, sense of self, quality of early mother–child relationship, temperament, intelligence, and dissociation measured at four time points were related to subsequent dissociative symptomatology. Age of onset, chronicity, and severity of trauma were highly correlated and were predictive of dissociation. Two forms of attachment disorders, avoidant and type D attachment, were

strong predictors of subsequent dissociation. Structural equation modeling found that type D attachment mediated between trauma and dissociation, accounting for 15 percent of the variance in dissociation scores (DES) at age 19.

Studies by Karlen Lyons-Ruth revealed two subgroups of mothers in nonclinical populations whose children exhibit type D attachment disturbances: "hostile/self-referential" and "helpless/fearful." The former are characterized by exposure to physical abuse and display "negative intrusive" behaviors such as mocking and/or being physically aggressive with the child. The second group of mothers is characterized by exposure to childhood sexual abuse and display confused, disorganized, and frightened responses to the infant, associated with physical and verbal withdrawal. These patterns were strong, longitudinal predictors of dissociation in their children; mothers with dissociative disorder may show both patterns. Discrete traumatic events, severe repetitive traumas in particular, also may produce dissociative disorders in the children of these disturbed mothers. The relationship of dissociation and type D attachment is consistent with Betrayal Trauma Theory, where the frightening/frightened mother betrays the child's basic developmental needs for secure attachment. Consistent with these childhood studies, a study of 554 healthy adults found a large effect size (d = 1.43) for the combination of verbal/emotional abuse and witnessing domestic violence in childhood in predicting dissociation, which is comparable to the effect of familial sexual abuse.

Developmental Model of Dissociative Disorders

Modern interest in a developmental model of dissociative disorders, especially dissociative identity disorder, can be traced to pioneers who examined children with the condition, such as Cornelia Wilbur and Richard Kluft. Long-forgotten case reports were later discovered in which 19th century clinicians used phrases in Latin to hint at incest and unspeakable familial trauma. Small case series and reviews of childhood dissociative disorders appeared in the mid-to-late 1980s, with more systematic studies coming in the 2000s. As with the adult case series, histories of severe trauma were commonly noted (see section on Dissociative Identity Disorder).

The core phenomenology of dissociative disorders in children is largely similar to that in adults (see section on Culture, Dissociative Disorders, and Sociocognitive Phenomena). However, as with most childhood manifestations of lifelong psychiatric disorders, the day-to-day symptoms vary with the child's age. In general, older children and adolescents are more overtly symptomatic than younger children. In part, this may reflect the ability of older children to better report subjective distress, as well as their greater opportunity to manifest psychopathology in a variety of settings.

In a pooled sample of 177 child and adolescent dissociative disorder cases, amnesias, identity disturbances, and auditory hallucinations increased with age, whereas trance-like and "spacey" behavior was ubiquitous across all age groups. Related psychopathology, such as suicidal ideation, self-mutilation, and somatization, increased with age in parallel with dissociative symptoms. Boys and girls did not differ on dissociative symptoms, but girls were significantly more symptomatic for anxiety and phobic symptoms, PTSD, sleep disturbances, sexual acting out, and somatization. Boys showed more externalizing behaviors such as aggression and violence. These findings closely parallel the gender differences in adult clinical profiles. Here males with dissociative identity disorder report higher rates of violence, criminal involvement, and alcohol and substance use disorders. Children and adolescents diagnosed with dissociative identity disorder were the most symptomatic across all age groups, which is consonant with the clinical belief that dissociative identity disorder is the most severe of the dissociative disorders.

Validated child and adolescent dissociation measures have been administered to various clinical and nonclinical samples (see Table 20–2). Scores were significantly higher for traumatized (typically child abuse) versus nontraumatized respondents across all age groups. Higher levels of dissociation were also significantly associated with more general psychopathology. Among maltreated preschoolers, there were robust correlations with internalizing and externalizing behavior problems for girls and boys, respectively. A comparison of maltreated and nontraumatized preschoolers matched on family constellation and demographics found that the maltreated group had significantly increased levels of dissociation a year later, with the physically abused children accounting for the greatest increase in scores. The controls showed substantial decreases in their dissociation scores over the same period, in line with the often-reported decrease in dissociation scores with age in normal children.

Studies have explored the possible mediating role of dissociation in the development of psychopathology in sexually abused children and adolescents. Measuring psychopathology with a variety of standard measures, the main effect of sexual abuse on behavioral problems disappeared when levels of dissociation were controlled for, as measured by a self-report measure (A-DES) or by observer report (CDC). The possibility that dissociation is a critical mediating variable for subsequent psychopathology has significant implications for early intervention with maltreated and traumatized children. Assessment of dissociation should be a standard part of the evaluation of traumatized children and adolescents, and significant elevations should be addressed as part of the treatment plan.

Search for Developmental Precursors and Substrates of Dissociation

Imaginary Companionship. The focus on child dissociative identity disorder led clinical investigators to initially focus on childhood fantasy phenomena, such as imaginary companionship, as possible developmental precursors for dissociative disorders. Imaginary companions are reported in 20 to 60 percent of normal children, depending on the age of the child and the definition used. Normal imaginary companionship is widely regarded as benign and is commonly considered to be a sign of creativity in young children, but it becomes increasingly suspect in older children and adolescents and is thought to be always pathological in adults. In dissociative children, the rates of imaginary companionship range from 42 to 84 percent, with the highest rates reported for children diagnosed with dissociative identity disorder.

Comparing the characteristics of imaginary companions in maltreated and nonmaltreated boys yields striking differences. The maltreated boys averaged 6.4 entities compared to 2.5 entities for the nonmaltreated boys. The latter first appeared between 2 and 4 years of age and had all disappeared by 8 years of age. The imaginary companions reported by the boys in residential treatment, however, served other functions, including as (1) helpers and comforters, (2) powerful protectors, and (3) family members. In some cases, the boys reported feeling as if they had lost control, and their imaginary companions took over their behavior. The maltreated boys reported that their imaginary companions were still present at a mean age of 10.6 years, well after they had disappeared in the normal boys. The imaginary companions of the maltreated boys often had names such as *God, the devil,* and *guardian angel.* A number of authorities theorize that some of the imaginary companions found in maltreated children eventually evolve into the alter personality states that personify dissociative identity disorder. As yet, no one has documented this transformation, but dissociative identity disorder patients sometimes report that this happened to them.

Developmental Mediation of Psychopathology by Dissociation

The negative effects of elevated dissociation on child development may operate by hindering the development of a sense of self, emotional regulation, impulse control, information processing, and coping with stressors. Children assemble and seek to integrate a complex, multidimensional sense of self over the course of development. Dissociative components, such as autobiographical amnesias, depersonalization, and passive influence experiences, interfere with the integration of self and the development of a unified sense of self-agency.

Dissociative amnesias also disrupt the child's understanding of cause-and-effect sequences, so that negative or risky behaviors may appear

unconnected from their consequences, which may be experienced as "coming out of the blue." This lack of connection can interfere with the ability of dissociative children to learn from their experience. The impact of dissociation on the metacognitive function of self-monitoring one's behavior is thought to disrupt the integration of experience across contexts, further complicating the child's ability to learn and to practice self-control, particularly in the context of stressors. Increased dissociation is well correlated with increased aggression, impulsivity, and poorer social skills in children and adolescents.

Prospective studies of sexually abused girls compared with carefully matched controls show that dissociation is negatively associated with competent learning and overall classroom performance and is strongly predictive of school avoidance. The combination of cognitive dysfunctions and negative school trajectories puts the dissociative child at an academic disadvantage, which has profound implications for adult attainment.

Chronic depersonalization, closely associated with emotional numbing in PTSD patients, is thought to promote a sense of detachment from self that fosters risky and self-destructive behavior, such as self-mutilation. Self-mutilation has been strongly associated with increased dissociation in numerous studies. Dissociative patients often describe self-inflicting pain in attempts to break through profound states of depersonalization. At other times, they report feeling nothing as they cut or burn themselves. Alienation from self is also thought to play a role in the high rate of suicide attempts in dissociative patients. As noted above, dissociation has been implicated in predisposing an individual to revictimization. In aggregate, these experiences expose dissociative individuals to further traumatization, which takes a cumulative toll over the individual's life. Identification of protective environmental factors may help stimulate intervention models that prevent the development of significant dissociation in acutely traumatized children.

Hypnotic Model. In its basic form, the hypnotic model hypothesizes that a traumatized individual uses his or her innate hypnotic capacity to induce autohypnosis as a defense against overwhelming or repetitive traumatic experiences. In dissociative identity disorder, the autohypnotic state is transformed with continuous use into an independent alter personality state. Several lines of evidence may support the autohypnotic hypothesis. First, dissociative patients are highly hypnotizable, especially those with dissociative identity disorder. Second, many of the clinical phenomena associated with pathological dissociation—such as trance states, age regression, auditory hallucinations, and amnesias—can be transiently produced in normal individuals using hypnosis. Finally, two studies suggested that childhood trauma might increase hypnotizability.

More systematic research has dispelled the notion that childhood

trauma generally increases hypnotizability. At least six studies found no differences between traumatized and nontraumatized adult and child samples, nor are there discernible dose effects between measures of trauma severity and hypnotizability in clinical and general population samples. However, a small group of individuals, termed *double dissociators,* score high on measures of dissociation and hypnotizability. These individuals generally have histories of earlier and more severe trauma. This observation is most compatible with the taxonic model of pathological dissociation and may explain the greater hypnotizability of dissociative identity disorder patients. There may be different types of hypnotizability, with high dissociators making up a distinct type. However, other hypnotizability studies have found subgroups of highly hypnotizable people in nondissociative samples. Autohypnosis is not the only pathway to dissociative symptoms, but it remains a possible contributing mechanism in some individuals.

Psychoanalytic Models. As described in the section on History of Dissociation, psychoanalytic theory has had difficulty with the concept of dissociation since the unfortunate schism of Freudian and Janetian ideas in the late 19th and early 20th centuries. In addition, in psychoanalytic writings, the term trauma was broadened to refer to almost any developmental stressor, a conceptual difficulty recognized by some psychoanalytic writers such as Anna Freud, when she studied British children with wartime traumas, such as exposure to bombings. Also, psychoanalytic theory became focused on the intrapsychic appraisal and symbolization of events, rather than on interplay between the severity of traumatic events and subjective appraisal.

Other than in the work of David Rappaport and his associates at the Menninger Clinic in the 1940s, that attempted to integrate Freudian and Janetian concepts, as well as a few papers on "hypnoid states," until recently, most psychoanalytic writing has centered on the concept of "splitting" as equivalent to dissociation. On the other hand, psychoanalyst Steven Marmer, based on his experience treating dissociative identity disorder patients, posited intrapsychic transitional object phenomena as a specific developmental aspect of the creation of dissociative identity disorder self-states. Some theorists of the relational school of psychoanalysis now view dissociation as a general psychological phenomenon that underlies all neurotic, personality disorder, and posttraumatic symptoms. However, in general, psychoanalytic theory appears to have difficulty viewing dissociation as a specific response to trauma, both developmentally and in adults, and as a specific intrapsychic defense against subsequent awareness of trauma and its impact on the person.

Structural Dissociation Theory. Onno van der Hart and Ellert Nijenhuis have attempted to develop a comprehensive theory, *Structural Dissociation Theory* (SDT), postulating that a basic dissociative "split" in personality accounts for a number of posttraumatic manifestations and disorders. SDT is popular in Europe as an etiological theory and treatment model. This theory proposes that dissociative disorders, PTSD, somatoform disorders, and some cases of borderline personality disorder are fundamentally "dissociative" in that severe traumatic experiences led to a split between an *Apparently Normal Personality* (ANP) and an *Emotional Personality* (EP). The hypothesized ANP is constricted and avoidant, unaware of traumatic memories, externally focused on daily life, and phobic of internal experiences. The EP is filled with the pain and distress of unintegrated, unprocessed traumatic memory. In this theory, PTSD results from a single EP and the ANP, and more complex dissociative disorders involve "secondary" and "tertiary" structural dissociations with, respectively, multiple EPs and then multiple ANPs. Tertiary structural dissociation results in the development of dissociative identity disorder, and secondary structural dissociation in less complex dissociative disorders.

SDT is helpful as a heuristic for some aspects of the psychopathology of posttraumatic and dissociative disorders, and attempts to posit a more general conceptualization of dissociative identity disorder self-states, rather than focusing on overt, personified characteristics of self-states (e.g., child, introject, male, female). Aspects of dissociative disorder treatment, particularly internalized phobias of emotional experience, can be easily conceptualized with SDT. Neurobiological research on dissociative identity disorder has used aspects of this theory in a helpful way, such as by comparing "traumatic" (EP) versus "neutral" (ANP) identity states. In PET scan studies, when read personal trauma scripts, the TIS showed neurobiological patterns of activation of amygdala, insula, and emotional centers, with decrease in frontality. In contrast, the NIS showed a pattern of increased frontal activation, decreased activation of amygdala and insula, accompanied by a subjective experience that the trauma narrative was not a personal autobiographical account. Autonomic activation increased for the TIS, but decreased or did not significantly increase in the NIS. These findings show neurocircuitry patterns within the same individual consistent with the dissociative and nondissociative subtypes of PTSD in separate individuals. Simulator controls did not show the same neurobiological and autonomic findings.

Problems with SDT include its essentially reified conception of the mind as a structure, with a concretized personality that is "split." The etiological aspect of the theory appears to be based on reasoning backward from adult dissociative identity disorder, rather than forward from the developmental perspective of the theories described in prior sections. In addition, meta-

analyses of PTSD and dissociation are consistent with the model that dissociation is either best understood as a subtype or a component of PTSD, rather than that dissociation underlies all forms of PTSD. Other problems include the unwieldy, reified, and difficult-to-operationalize conceptual language (e.g., "apparently," "normal," and "personality") that minimizes the notion of dissociation as a "process," or set of processes, in favor of metaphors for "structure." The SDT model of dissociation posits universally rigid amnestic barriers between ANP and EP, and sequential activation of differentiated states, rather than the shifting, variable symptom presentation that characterizes many dissociative individuals, with overlap, interference, and intrusion between states.

DISSOCIATIVE IDENTITY DISORDER

Dissociative identity disorder has been researched most extensively of all the dissociative disorders. It is the paradigmatic dissociative psychopathology in that the symptoms of all the other dissociative disorders are commonly found in patients with dissociative identity disorder: Amnesias, fugues, depersonalization, derealization, and similar symptoms. The DSM-5 diagnostic criteria for dissociative identity disorder are discussed in the section on Nosology, and are listed in Table 20–3.

Etiology of Dissociative Identity Disorder—Trauma

Non–possession-form dissociative identity disorder is strongly linked to extreme, chronic, and early childhood maltreatment in all studies that have systematically examined this question. The rates of reported severe childhood trauma for children and adults with nonpossession dissociative identity disorder range from 85 to 97 percent of cases across a wide variety of studies, in both females and males. Physical and sexual abuse, usually in combination, are the most frequently reported sources of childhood trauma in clinical research studies, although other kinds of trauma have been reported, such as multiple painful medical and surgical procedures during childhood, and wartime trauma. Emotional abuse, neglect, and witnessing family violence have also been cited as traumatic antecedents of dissociative identity disorder, usually in combination with physical and sexual abuse. Studies in the National Comorbidity Study Replication (NCS-R) have found that sexual, physical, and emotional abuse, neglect, intimate partner violence, and severe family dysfunction covary together so extensively in violent and abusive families that no one factor among these can easily be partialed out from any other in attempting to understand the impact of early life maltreatment.

Data on the role of trauma in possession-form dissociative identity disorder is more limited. Marjolein Van Duijl and colleagues conducted one of the first systematic studies of traumatic exposure among 119 possessed

patients who visited traditional healers in Uganda. These individuals scored higher on the Harvard Trauma Questionnaire and the Traumatic Experiences Checklist than 71 randomly selected healthy inhabitants of the same villages, and scores on both trauma scales were strongly correlated with symptoms of dissociation (r = 0.64 to 0.69) and pathological possession (r = 0.51 to 0.64). Nontraumatic stressors, such as romantic disappointments and work or religious conflicts, have triggered other cases of pathological possession. However, the absence of systematic assessment of childhood trauma precludes excluding this exposure from the genesis of possession-form dissociative identity disorder in these populations.

Despite abundant evidence of the relationship between trauma and dissociative identity disorder, critics have raised questions about the validity of dissociative identity disorder patients' self-reports of childhood trauma. Recent studies, including large samples of maltreated children with dissociative disorders and intensively validated case studies, nevertheless have provided rigorous independent corroboration of the patients' reports of maltreatment. These studies continue to strongly support a developmental linkage between childhood trauma and dissociative identity disorder. In general, studies involving adult self-reports of childhood trauma have not shown these reports to be any less valid than self-reports of other clinical phenomena, such as self-reports of a history of smoking.

Dissociative Identity Disorder as the Paradigmatic Posttraumatic Disorder. The posttraumatic etiology of dissociative identity disorder can be supported by data showing that many disorders, symptoms, life experiences, and maladaptive behaviors typical of traumatized individuals are commonly found in dissociative identity disorder. For example, the National Comorbidity Study-1 (NCS-1) found that the most common outcomes of trauma included PTSD, depression, and substance/alcohol abuse. Studies of dissociative identity disorder have found comorbid PTSD in 79 to 100 percent of cases, current or past depression in 83 to 96 percent, and a past or current history of substance abuse in 83 to 96 percent of cases.

Suicidal, self-destructive, violent, and high-risk behaviors are also strongly associated with trauma, and, particularly, with early life trauma and repeated traumatic experiences. Across studies, 66 to 80 percent of individuals with dissociative identity disorder reported a history of suicide attempts, and 92 to 100 percent reported current or past suicidal ideation. Individuals with dissociative identity disorder have higher rates of multiple suicide attempts than individuals with borderline personality disorder, PTSD, and substance abuse disorders who do not have a comorbid dissociative disorder. In a 30-month prospective, longitudinal treatment outcome study described in detail below, 2 patients—1.8 percent of the final

sample of 119—committed suicide. However, this likely is an unrepresentatively low rate, based on participation in a treatment outcome study that demonstrated an overall decrease in suicidality among its participants. Case series also have described completed suicide in dissociative identity disorder patients, including in possession-form cases.

Studies with pre–DSM-5 criteria show that 78 percent of dissociative identity individuals have engaged in self-harm behaviors. These include severe forms of self-injury such as major lacerations requiring sutures or surgical repair to limbs, genitals, and abdomen; severe burns including use of acids and caustic substances; ingestion of caustic substances such as drain or oven cleaner; deliberate fracture of limbs; and severe head-banging resulting in bruising, bleeding, and/or skull fractures, among many other forms of self-destructive behavior.

Clinical case series document that a subgroup of individuals with dissociative identity disorder engage in violent behavior; some have been incarcerated for violent crimes including murder, severe assaults, and rape. Twenty percent of individuals in a case series of male patients with dissociative identity disorder reported committing homicide and 13 percent committing rape. In this study, 10 percent of females with the disorder also reported committing homicide. In a clinical study of mothers with dissociative identity disorder, 16 percent engaged in significant maltreatment of their own children, and 45 percent were compromised or impaired as mothers. Table 20–12 enumerates many of the safety issues that occur in patients with dissociative identity disorder.

Like other victims of early life trauma, persons with dissociative identity disorder have high rates of adult traumatization, in particular rape. In one study, 60 percent of women and 17 percent of men with dissociative identity disorder reported being raped as adults. Nearly 60 percent of women with dissociative identity disorder reported suffering intimate partner violence or other forms of violence in adulthood; 36 percent reported experiencing violence in the past year. These rates of rape and intimate partner violence are notably higher than those for the general US population.

Table 20–12.
Types of Dangerousness to Self and Others Typically Found in Dissociative Identity Disorder Patients

1. Dangerousness to self
   a. Self-destructive behavior
   b. Suicide attempts
   c. Completed suicide
2. Dangerousness to others
   a. Abuse the minor children of the patient
   b. Intimate partner violence

   c. Assault, sexual assault, arson, murder, serial murder, etc.
3. High-risk behaviors, e.g., repeated reckless driving, risky sexual behaviors
4. Substance-related and addictive disorders
5. Eating disorders
6. Enmeshment in/victim of abusive/traumatizing relationships
   a. Enmeshment with abusive family of origin
   b. Intimate partner violence
   c. Being trafficked into prostitution
7. Lack of food, clothing, and/or shelter
8. Lack of access to and/or avoidance of adequate medical care

As described in the Epidemiology section, a subgroup of individuals with dissociative identity disorder report being trafficked into prostitution, often beginning in childhood with victimization in child prostitution and pornography. Many of these individuals also report that they continue to be trafficked into adulthood, sometimes by the same perpetrators who often are members of their families of origin. This phenomenon urgently requires further study, and antihuman trafficking legislation should take such dynamics into consideration.

SOMATIC SYMPTOM DISORDERS, TRAUMA, AND DISSOCIATIVE IDENTITY DISORDER. High rates of childhood abuse and maltreatment characterize subsamples of individuals with somatic symptom and related disorders, especially conversion disorder and somatic symptom disorder. Across studies, 40 to 92 percent of individuals with dissociative identity disorder report somatoform symptoms and 35 to 61 percent have a comorbid somatic symptom disorder. Data from Turkey suggest that pseudoneurological conversion symptoms are the most common presenting symptom of persons with dissociative identity disorder; some of these Turkish patients attributed their alternate identities to possession, often combined with non–possession-form identities. In one study by Vedat Sar and colleagues, nearly half of 38 consecutive psychiatric outpatients diagnosed with conversion disorder met DSM-IV criteria for a dissociative disorder, including 17 percent with dissociative identity disorder and nearly three-quarters with dissociative disorder NOS. A subgroup of North American/European dissociative identity disorder patients come to clinical attention for a primary pseudoneurological conversion disorder, especially nonepileptic seizures and sensory–motor conversion symptoms.

Patients with dissociative identity disorder also frequently meet diagnostic criteria for somatic symptom disorder, often with predominant pain—particularly headaches, abdominal pain, and pelvic pain, but also migratory pain syndromes. In other cases, somatic symptoms can complicate the diagnosis and treatment of physical disorders associated with chronic pain in these patients, such as lumbar disc disease. Childhood sexual abuse survivors with psychophysiological disorders are more likely

than controls to have a lower threshold for experiencing physiological phenomena as noxious or painful.

The ACE study has been discussed in the section on Epidemiology. The authors' clinical experience is that persons with dissociative identity disorder often have—in addition to the psychiatric disorders related to high exposure to ACEs in childhood—strikingly high rates of nonpsychiatric conditions related to multiple ACEs. These include morbid obesity, sexually transmitted diseases (e.g., Hepatitis C), HIV infection, autoimmune diseases, and early pregnancy—including incest pregnancy. Persons with dissociative identity disorder also commonly suffer from gastroesophageal reflux disease (GERD) and irritable bowel syndrome (IBS). Significantly higher rates of childhood sexual abuse have been found in patients with these disorders.

PERSONALITY ORGANIZATION, PERSONALITY DISORDERS, AND DISSOCIATIVE IDENTITY DISORDER. Borderline personality disorder has been reported in 31 to 83 percent of patients with non–possession-form dissociative identity disorder. Some early explanatory theories characterized dissociative identity disorder as a variant of borderline personality disorder or a condition within the psychoanalytic construct of borderline personality organization. However, several studies have shown substantial differences between persons with dissociative identity disorder and individuals with Cluster B personality disorders. Using the Structured Clinical Interview for the DSM-IV Dissociative Disorders (SCID-D; see Table 20–2), patients with dissociative identity disorder and those with Cluster B personality disorder (primarily borderline personality disorder) did not differ significantly on measures of depersonalization and derealization. However, patients with dissociative identity disorder reported significantly more amnesia, identity alteration, and identity confusion. Also, the dissociative patients had significantly earlier and more extensive histories of trauma.

Psychological assessment of large numbers of persons with dissociative identity disorder using the Rorschach test, and other assessment instruments such as the Millon Clinical Multiaxial Inventory-III (MCMI-III), did not show a personality profile typical of borderline or histrionic personality types, or of patients with psychosis studied with the same measures. Instead, on the Rorschach, individuals with dissociative identity disorder showed a unique personality structure characterized by marked traumatic reactivity, obsessional features, introversion, and high complexity of response. Also, these patients showed notable ability for self-reflection, as exhibited by good insight (Rorschach variable FD), reality testing, and capacity for logical as opposed to distorted thinking. Despite early relational trauma, these patients were interested in engaging collaboratively with others. However—when they were overwhelmed by trauma intrusions—their thinking, ability for emotional modulation, and

reality testing tended to become more impaired.

Some individuals with dissociative identity disorder even experienced traumatic flashbacks during testing, treating the Rorschach blot as if it were a traumatizing object. On the other hand, when they start to become overwhelmed by PTSD symptoms, many patients with dissociative identity disorder showed unconscious attempts to distance themselves from trauma responding, reflected in an increase in FD. This may parallel the findings among persons with dissociative PTSD in a study of unconscious and conscious responses to fearful faces. The dissociative PTSD group showed rapid, marked unconscious arousal of fear circuitry to the threatening stimuli. With conscious processing, however, frontal brain systems that regulate emotional arousal in the amygdala were rapidly activated, accompanied by increased dissociative symptoms.

This research shows that early life dissociation may serve a protective, resiliency-inducing function for children growing up in violent, malevolent environments. The dissociative sequestration of traumatic memories and emotions allows for more normal development in other areas of life including capacity for logical thought, mentalization and insight, artistic and intellectual accomplishments, willingness to engage in relationships with others, preservation of a sense of humor, and hope.

Across studies on the MCMI-III, patients with dissociative identity disorder showed highest scores on depression, PTSD, and anxiety. They also frequently endorsed negative personality traits—especially avoidant, self-defeating, passive-aggressive, and borderline traits—similar to patterns found in individuals with chronic PTSD. Consistent with the Rorschach findings, in one study, persons with dissociative identity disorder scored lowest on histrionic traits, followed by low scores on narcissistic, antisocial, and sadistic traits. The MCMI-III findings are consistent with the conceptualization of dissociative identity disorder as a form of posttraumatic psychopathology.

Several researchers using the DDIS that ascertains borderline personality disorder symptoms, view patients endorsing many of these symptoms as a severity indicator for dissociative identity disorder, measuring general instability, rather than a diagnostic indicator of borderline personality disorder. For example, a DDIS diagnosis of borderline personality disorder significantly correlated with more extensive trauma histories, higher scores on the DES, more passive-influence symptoms, increased complex amnesia, and other dissociative symptoms. This is consistent with the Rorschach studies that show deterioration in logical thinking and reality testing in individuals with dissociative identity disorder, when overwhelmed by traumatic intrusions. Clinical case series have also shown that, after stabilization of acute posttraumatic and dissociative symptoms, a minority of patients with dissociative identity disorder exhibit persistent severe Cluster A and B personality disorder

features that complicate clinical progress.

Some individuals with dissociative identity disorder have antisocial, even psychopathic, personality features. This might constitute a subgroup that does not usually come to clinical attention but is regularly seen in social service or law enforcement contexts. Moreover, as noted earlier, a subgroup of individuals with the disorder appears to be perpetrating criminal activities, including the abuse and/or trafficking of their own children. Paranoid personality traits, well beyond the typical posttraumatic suspiciousness of the complex trauma survivor, may be found among this subgroup, including when evaluated in a forensic context.

Explanatory Models: Discrete Behavioral States and Dissociative Identity Disorder

The DBS model, discussed in the prior section on Etiology, is a particularly helpful overarching construct for conceptualizing the development of dissociative identity disorder. The DBS is grounded in data on child development and the ways in which maltreatment and/or severe early stress can change the trajectory of development in sense of self, autobiographical memory, capacity for attachment, emotion regulation, cognitive schema, ability to maintain and/or modulate behavioral states, and beliefs about the perpetrators, the self, and the world, among others.

According to the DBS model, dissociative identity disorder is conceptualized as a childhood-onset posttraumatic developmental disorder in which the traumatized child is unable to complete the normal developmental processes involved in consolidating a core and singular sense of identity. Instead, repeated early trauma disrupts the usual unification of identity, resulting instead in the development of multiple independent or discrete behavioral states. In normal development, adults implicitly recognize state changes in children and attempt to assist the child in reducing distressing states (e.g., hungry, wet, teething, tantrums, ill) and/or in returning to better states (e.g., calm, interested, sleepy). These adult behaviors help the child develop the capacity to recognize and self-modulate his/her own states, and, ultimately, to develop a singular and internally consistent sense of self across states and contexts, including distressing ones.

In comparison, traumatic experiences—particularly repeated maltreatment perpetrated by caregivers—not only produces extreme, overwhelming states, but also leave the child with little or no assistance from adults in modulating the effects of these states. Abusive parents usually provide little or highly inconsistent soothing experiences, especially after having just hurt the child.

Due to disturbed caretaker–child attachment and parenting, repeated early trauma disrupts the development of normal metacognitive processes involved in the consolidation of a unified sense of self across different

relational contexts (e.g., with parents, peers, and others). For example, the child who develops dissociative identity disorder often experiences marked discontinuities between the reality of the chronically abusive, violent home life and the presentation of the family as "perfect" in the outside world (e.g., at church, school, public events). The dissociative child may also experience other areas of life as additional disconnected realities: school, athletics, artistic endeavors, and so on. Thus, the developing self-states may be experienced as encompassing alternate, often contradictory realities, leading some to refer to the disorder as a *multiple reality disorder*.

Over time, the child's developing internal sense of self may come to reflect the disorganization and conflict evident across the various relationships and environments with which he or she is occupied. Different self-states can become subjectively personified and develop along independent developmental trajectories. With adolescence, elaboration of the various self-states may occur, leading to further compartmentalization of the subjective sense of identity and ultimately the phenomenology associated with adult forms of dissociative identity disorder.

The personality states of individuals with dissociative identity disorder can be conceptualized as different DBSs, each organized around a prevailing affect, sense of self (often including a distinct body image), set of state-dependent autobiographical memories, and limited behavioral repertoire, with a capacity for action. This leads to development of multiple, relatively separate centers of information processing in the mind, capable of relatively independent thought, emotion, sense of self, and memory, as well as the capacity to take control of the person's behavior.

One drawback of the DBS model is that it has not been studied in individuals with possession-form dissociative identity disorder. In societies where spiritual agents (e.g., deities, ghosts, animal spirits, demons, *jinn*) are considered ontologically "real," these agents represent alternate sources of identity construction. Individuals who are maltreated may engage these supra-individual identities as figures that represent aversive or supportive elements in their lives. Even in societies where supernatural forces are not generally endorsed, intracultural variation may encourage the assumption of spiritual identities within certain families and social groupings (e.g., highly religious groups). This process would give rise to possessing agents that have a collective existence. For example, in India the same spirit may possess different family members down the generations (or different villagers in the same generation). In societies where the cultural conception of the self is more porous to other identities depending on context ("dividual"), including spiritual identities, the development of DBS may proceed along less individualistic paths, resulting in differently organized alternate identities. On the other hand, clinical data on Western families with multigenerational dissociative identity disorder have found similar self-states across three generations of the same family.

The Four-Factor Theory of Dissociative Identity Disorder. Richard Kluft has proposed a Four-Factor Theory (FFT) of dissociative identity disorder that is complementary to the DBS model. The first factor is the *capacity for dissociation,* likely due to a genetic diathesis (discussed in the section on Neurobiology). The second factor involves *experiences that overwhelm other nondissociative adaptive capacities of the child.* In this model, the child cannot fight or flee from the abusers, but freezes and flees inwardly into altered states of consciousness that provide encapsulation and relief from traumatic experiences. The third factor is the *lack of soothing and restorative experiences,* consistent with the DBS model, where the traumatized child must find ways to calm and soothe him/herself, and, attempt alone to mitigate the impact of these events.

The fourth and final factor of the FFT involves the many posttraumatic, psychodynamic, familial, social, and cultural factors leading to the *secondary structuring, subjective personification, and elaboration of the self-states* (also see section on Culture, Dissociative Disorders, and Sociocognitive Phenomena). Kluft describes how this may include posttraumatic, psychodynamic factors resulting in the first, very early psychological division, the creation of an isomorphic state: one child who is traumatized and one who is not. Further elaboration may include, for example, a female child thinking that, if she were a boy, sexual abuse would not be happening, leading to development of a subjectively male state. Other personified states may attempt to mitigate and sequester the impact of multiple ongoing traumas, leading to development of multiple traumatized states that encapsulate traumatic memories, or parts of one or more memories. Traumatic and/or overwhelming experiences in adolescence or adulthood may also lead to the development of additional traumatized states. Other states may be experienced as without emotions or pain perception to mitigate the impact of traumatic memories and experiences. These states may be experienced as biographical variants of the individual (e.g., a numb child identity)—as is typical in non–possession-form dissociative identity disorder—or as collective, mythical, states that transcend individual biography (e.g., a dead child who cannot feel pain, an invulnerable spirit)—as is seen in possession-form presentations.

Attempts to feel strong and safe may lead to subjective creation of a tough, powerful adult male state or a powerful spiritual ally, such as a deity. Attempts to feel in control of at least some aspect of continual, inevitable, unpredictable trauma may lead to the development of a state that is an internalized representation of the abusive, shaming perpetrator(s), often termed an *introject self-state;* alternatively, this could take the form of a demon or punishing animal spirit. By contrast, people who have been kind to the individual may also be represented as introject states. For example, if a supportive, kind relative dies, the person with dissociative identity

disorder may have difficulty grieving the loss of this person. An introject state develops instead as an internal representation of the deceased individual, forestalling the need for grieving. A case example illustrates this process. In addition, later traumatic or overwhelming events may lead to development of additional self-states.

DISSOCIATIVE IDENTITY DISORDER (DID) CASE 1

Mr. T, a divorced musician in his 30s was admitted to a psychiatry unit. He was suicidal after repeatedly finding himself dressed in women's clothes. He denied recall of buying or putting on these clothes, or for the periods of time that he cross-dressed. He had become panicked, mortified, and intensely suicidal when family members discovered him dressed in female clothes. He had always viewed himself as "a tough, macho kind of guy." Mr. T had been in recovery from drugs and alcohol for a decade and vigorously denied any recent substance use. His family and AA sponsor corroborated Mr. T's having remained sober, and screening for drugs and alcohol was negative. He reported gaps in his recall for many periods of his life. He ascribed the memory problems to substance use, although inexplicable memory gaps were present for long periods of sobriety as well, which he attributed to "dry blackouts." He described sudden disconcerting lapses in his musical abilities, which had created job problems for him. At other times, he described finding that he could play other instruments proficiently without knowing how he had learned them. He reported that his head was "filled with opinions that don't belong to me." On more careful history taking, he described a chaotic early life, being sent to live with violent, abusive relatives. The only consistently positive figure in his development was his maternal grandmother, who had supported and protected him as much as she could. She had died suddenly—just before the episodes of cross-dressing began. As he discussed this, Mr. T reported hearing his grandmother's voice, and broke down in tears stating that he could not bear the idea that she was dead. On sequential evaluation, several self-states were identified, including one based on his grandmother that claimed to only have been created after her death. This self-state acknowledged buying and dressing in the female clothes, "So I could always be with him." Work on grieving for his grandmother resolved the acute episodes of amnesia and pressure to cross-dress. Mr. T was referred to long-term outpatient psychotherapy for dissociative identity disorder.

SHAME AND DEVELOPMENT OF PERSONALITY STATES. Affect theory has been discussed in the section on Etiology. Feelings of shame are likely a concomitant of most, if not all, traumatic experiences. Multiple early life traumas, often involving deliberate, sadistic humiliation and mocking of the victim usually lead to an extreme burden of shame for the individual. Richard Kluft, based on the work on affect theory by Donald Nathanson, has pioneered the conceptualization of self-states as embodying specific shame scripts such as *self-attack* (i.e., repeated self-shaming), *avoidance of shame affect* and anything that might engender it, *withdrawal* from much of life as extreme avoidance of shame, and/or *attacking others* to shame them—a fundamental part of perpetration. The latter is commonly experienced in dissociative identity disorder as negative introject self-states, often felt to be separate beings that shame the person.

OTHER SUBSTRATES SHAPING PERSONALITY STATES. In possession-form presentations, identities are usually perceived to be "real" and external by members of the social group who share the sufferer's perspective regarding the ontological validity of spiritual beings. In non–possession-form

dissociative identity disorder, however, the experience of personality states as literally separate, "outside" people is termed *delusional separateness*. This usually occurs when, early in development, repeated traumatic events have been experienced as so overwhelming, intolerable, damaging, shameful, disgusting, dissonant, and ego-dystonic that the person may attempt to subjectively extrude them altogether, as "not-me" events. In assessment of possible psychosis, it is important for clinicians to take into account the prevailing cultural views of the person and his/her community as well as processes such as hallucinatory phenomena related to self-states and delusional separateness that are best explained by a dissociative and posttraumatic etiology.

Other factors that lead to secondary structuring of dissociative identity disorder identity states can include contradictory parental demands (e.g., father wants a son and dresses the child as a boy; mother wants a girl and dresses the child like a girl). As noted above, persistent imaginary companionship may also be a substrate for the development of self-states. Sociocultural factors that lead to further elaboration of self-states may include books, television, movies, and religious beliefs (e.g., belief in demons). In addition, the intelligence and creativity of the person can influence the complexity of the subjective elaboration of identities. Some individuals with dissociative identity disorder may retreat into highly complex, elaborated inner worlds. This is most likely to occur when the experience of outer reality is perceived as too painful, and especially when the person has not only undergone severe life traumas but also has experienced a limited positive adaptation to the world in his/her adult life. This constellation may be associated with poorer treatment outcome. Similarly, studies have shown that psychotic patients with grandiose delusions are less adherent to medications than those with non-grandiose delusions. In both these clinical situations, the patients attempt to avoid experiencing their painfully limited lives by remaining lost in elaborate dissociative fantasy worlds or grandiose delusional beliefs, respectively.

Psychotherapy can lead to the development of additional personality states, especially if the treatment is experienced as unsafe and/or overwhelming to the person. Therapists may unintentionally reinforce separateness by playing favorites among self-states; engaging in very different behavior with different states (e.g., playing with child self-states); viewing self-states as "good" or "bad," rather than psychological adaptations that need to be understood; refusing to speak to some self-states; and not encouraging a decrease in dissociative barriers by failing to promote internal communication, collaboration, and coordination among self-states—for example, by making separate appointments with self-states, as if with separate people. Sexual or other forms of exploitation by therapists are severe traumatic betrayals and can also lead to development additional self-states.

Unfortunately, the *least important phenomena associated with Factor 4* are those that lead to the most fascination with the patient or the diagnosis, and are emphasized in media portrayals of dissociative identity disorder. These include the person's highly elaborated, differing presentation of personality states, or fantasies about the states ("12th century monk; Samurai Warrior") although even these are factors are depicted in an exaggerated way in the media. In patients, there may be elaboration of the presentational and subjective characteristics of the identities with shifts in wardrobe, grooming, accents, gait, and so on. It can be a poor prognostic sign if the patient has significant narcissistic investment in having dissociative identity disorder and/or in the extensive elaboration of marked presentational differences between personality states.

Diagnosis and Clinical Features

Tables 20–13 through 20–15 describe the symptom clusters that are most commonly found in individuals with dissociative identity disorder and some of the mental status questions that may elicit these symptoms.

Memory and Amnesia Symptoms. Dissociative disturbances of memory usually manifest in a few ways that are frequently observable in clinical settings (Table 20–13). Dissociative amnesia in dissociative identity disorder commonly manifests as a current symptom, not just an experience of being unable to recollect past experiences. As part of the general mental status examination, clinicians should routinely inquire about current or past experiences of "losing time," black-out spells, and major gaps in the continuity of recall for personal information. Patients rarely spontaneously report these experiences, which require active inquiry by the interviewer to uncover the underlying amnesia. Positive responses should be documented with specific examples provided by the patient. In some instances, patients report coming to in the midst of some activity with little or no recall of how they came to be involved in that activity. In other instances, patients find evidence of having done or acquired things for which they have no recall (see DID Case 1). Friends and family members may tell them about significant things that they have said or done that they cannot remember. Patients may find that they have unknowingly traveled some distance (a fugue episode) or that days or even weeks have passed for which the patients cannot account. Dissociative experiences of time loss are too extensive to be explained by normal failures of episodic/autobiographical memory and typically have sharply demarcated onsets and offsets. It is important to establish that such time-loss experiences occur in the absence of intoxication or substance abuse, although high rates of drug and alcohol abuse in dissociative patients may complicate this determination, and some severely dissociative patients would prefer to ascribe their memory

problems to substance use (see DID Case 1).

Table 20–13.
Amnesia and Memory Symptoms in Dissociative Identity Disorder

Blackouts or time loss
Disremembered behavior
Fugues
Unexplained possessions
Inexplicable changes in relationships
Perplexing fluctuations in skills, habits, and knowledge (loss of procedural memory)
Fragmentary recall of entire life history
Chronic mistaken identity experiences
Microdissociations (Brief interruptions of awareness with amnesia)

Mental status examination questions for dissociative amnesia

If answers are positive, ask the patient to describe the event.
Make sure to specify that the symptom does not occur during an episode of intoxication.

1. Do you ever have blackouts? Blank spells? Memory lapses?
2. Do you lose time? Have gaps in your experience of time?
3. Have you ever traveled a considerable distance without recollection of how you did this or where you went exactly? Find yourself somewhere unexpected without knowing how you got there?
4. Do people tell you of things you have said and done that you do not recall?
5. Do you find objects in your possession (such as clothes, personal items, groceries in your grocery cart, books, tools, equipment, jewelry, vehicles, weapons, etc.) that you do not remember acquiring? Out-of-character items? Items that a child might have? Toys? Stuffed animals?
6. Have you ever been told or found evidence that you have talents and abilities that you did not know that you had? For example, musical, artistic, mechanical, literary, athletic, or other talents? Do your tastes seem to fluctuate inexplicably? For example, food preference, personal habits, taste in music or clothes, use of tobacco products, alcohol, drugs, etc.?
7. Do you have gaps in your memory of your life? Are you missing parts of your memory for your life history? Are you missing memories of some important events in your life? For example, weddings, birthdays, graduations, pregnancies, birth of children, etc.?
8. Do you lose track of or tune out conversations or therapy sessions as they are occurring? Do you find that, while you are listening to someone talk, you did not hear all or part of what was just said?
9. What is the longest period of time that you have lost? Minutes? Hours? Days? Weeks? Months? Years? Describe.

Adapted from Loewenstein RJ. An office mental status examination for chronic complex dissociative symptoms and multiple personality disorder. *Psychiatr Clin North Am.* 1991;14:567.

Table 20–14.
Dissociative Identity Disorder Process Symptoms

Presence of dissociative identity disorder self-states
Switching behaviors (identity alteration)
Identity confusion
Passive influence symptoms/interference phenomena between

Mental status examination questions for dissociative identity disorder process symptoms

If answers are positive, ask the patient to describe the event.
Make sure to specify that the symptom does not occur during an episode of intoxication.

1. Do you act so differently in one situation compared to another situation that you feel almost like you were two different people?
2. Do you ever feel that there is more than one of you? More than one

- self-states
  - Made feelings from self-states
  - Made impulses from self-states
  - Made actions by self-states
  - Thought insertion from self-states
  - Thought withdrawal by self-states
  - Self-states' voices commenting on behavior
  - Self-states' voices arguing
- Multimodal hallucinations or pseudohallucinations (may occur in a flashback)
  - Visual
  - Auditory
  - Tactile
  - Olfactory
  - Gustatory
  - Somatic
- Dissociative or posttraumatic thought disorder
- Posttraumatic stress disorder–based cognitive distortions
- Trance logic (toleration/rationalization of logical inconsistency in the hypnotized and/or dissociated state
- Other thought process abnormalities
  - Disorganization due to switching, passive influence
  - Posttraumatic suspiciousness
  - Literal and concrete alternating with abstract
- Linguistic usage: Refers to self as we, they, us, him, her, etc.
- Depersonalization and derealization symptoms
  - Depersonalization
  - Out-of-body experiences (In DID, sometimes experienced as

part of you? Side of you? Do they seem to be in conflict or in a struggle?

3. Does that part (those parts) of you have its (their) own independent way(s) of thinking, perceiving, and relating to the world and the self? Have its (their) own memories, thoughts, and feelings?
4. Does more than one of these entities take control of your behavior?
5. Do you ever have thoughts or feelings, or both, that come from inside you (outside you) that you cannot explain? That do not feel like thoughts or feelings that you would have? That seem like thoughts or feelings that are not under your control (passive influence)?
6. Have you ever felt that your body was engaged in behavior that did not seem to be under your control? For example, saying things, going places, buying things, writing things, drawing or creating things, hurting yourself or others, etc.? That your body does not seem to belong to you?
7. Do you ever feel you have to struggle against another part of you that seems to want to do or to say something that you do not wish to do or to say?
8. Do you ever feel that there is a force (pressure, part) inside you that tries to stop you from doing or saying something?
9. Do you ever hear voices, sounds, or conversations in your mind? That seem to be discussing you? Commenting on what you do? Telling you to do or not do certain things? To hurt yourself or others? That seem to be warning you or trying to protect you? That try to comfort, support, or soothe you? That provide important information about things to you? That argue or say things that have nothing to do with you? That have names? Men? Women? Children?
10. I would like to talk with that part (side, aspect, facet) of you (of the mind) that is called the "angry one" (the Little Girl, Janie, that went to Atlantic City last weekend and spent lots of money, etc.). Can that part come forward now, please? If that feels too distressing, can you listen inwardly and let that part of the mind speak to you, and you tell me what the part is saying?
11. Do you frequently have the experience of feeling like you are outside yourself, inside yourself? Beside yourself, watching yourself as if you were another person?
12. Do you ever feel disconnected from yourself, your thoughts, or your body as if you (your body) were not real? Like a robot? Mechanical?
13. Do you frequently experience the world around you as unreal? As if you are in a fog or a daze? As if it were painted? Two-dimensional? Flat?
14. Do you ever look in the mirror and not recognize yourself? See someone else there?
15. Do you feel like you have lost the capacity for feeling? For emotions? For affection toward others? For hunger? Thirst? Pain?

observing oneself from "inside" oneself.)
Body parts changing in size (metamorphopsia)
Emotional numbing
Loss of connection to the body
Loss of connection to one's own memories
Derealization
Déjà vu
Jamais vu
Déjà vécu
Dream-like states

Adapted from Loewenstein RJ. An office mental status examination for chronic complex dissociative symptoms and multiple personality disorder. *Psychiatr Clin North Am.* 1991;14:567.

Table 20–15. Autohypnotic Symptoms in Dissociative Identity Disorder

Spontaneous trance
Deep enthrallment
Spontaneous age regression
Negative hallucinations (not perceiving aspects of the self or environment that are present)
Hidden observer phenomenon (awareness of what is actually occurring when in the hypnotized state—e.g., experiencing pain in a study of hypnotic analgesia)
Trance logic (tolerance of logical inconsistency during a trance state)
Voluntary analgesia/anesthesia
Eye roll, eye blinking, etc., with switching

Mental status examination questions for spontaneous autohypnotic symptoms

1. Do you frequently space out, trance out, block out, or withdraw from the world around you? Perhaps by putting yourself in a pleasant scene or place in your mind? By focusing your attention on something inside or outside of you? (spontaneous trance)
2. Do you get so wrapped up in a book or a movie that you can completely block out everything else around you? As if the world could end, and you would still be completely engrossed in that activity? (enthrallment)
3. Do you feel that you are different ages at different times? Do you ever feel like you get little? Like you become a child or an adolescent again? When this happens, does it feel like your body changes in size? Do you experience it in your mind only, or does your whole perception of yourself and the world change also? (spontaneous age regression)
4. Do you ever *not* see or hear what is going on around you? Do you or can you block out people and things altogether? (negative hallucinations)
5. Are you able to block out (ignore) physical pain if you want to? Wholly? Partly? Always? Sometimes? (voluntary analgesia)
6. Is there a part of the mind that appears to be observing: aware of and recalls what is occurring when other parts do not? (Hidden observer)

Adapted from Loewenstein RJ. An office mental status examination for chronic complex dissociative symptoms and multiple personality disorder. *Psychiatr Clin North Am.* 1991;14:567.

Patients with severe dissociative memory disturbances also report

perplexing fluctuations in skills, habits, or well-learned abilities, such as fluency in a foreign language or athletic abilities. Patients report drawing a complete blank for skills or knowledge at times whereas, at other times, they easily and reliably access the information in question. This perplexing appearance and disappearance of procedural and semantic memory is believed to be related to dissociative state-dependent disturbances of implicit memory functions.

Dissociative patients often report significant gaps in autobiographical memory, especially for childhood events. These dissociative gaps are usually sharply demarcated and do not fit the normal decline in autobiographical recall over the life cycle. For example, a patient may complain that he or she cannot recall anything between ages 8 and 12 while reporting readily available memories before and after. Another patient may report having no memories available for the first 10 years of life. Available autobiographical memories may have a depersonalized or derealized quality, such that recalled events seem to be memories of a dream; as if the patient had seen them happen to someone else; as if they had read about the events; or as if they heard about the events from someone else. Some dissociative patients report that they only know about major parts of their past from discussions with friends or relatives, or from looking at photographs of themselves in past activities. Other individuals may describe systematized amnesias, where there is lack of memory for certain categories of experience, such as recall of school events, but normal memory for home events; inability to recall a specific family member; and/or failure to recall some everyday experiences (e.g., failure to recall snow despite growing up in a New England state). A case example describes these amnesia symptoms.

DISSOCIATIVE IDENTITY DISORDER CASE 2

Ms. B was seen in consultation in a hospital after becoming confused at work, unable to recall how to do her job, at which she was highly proficient. She reported lack of recall other recent events, such as coming to the hospital. An extensive work up for medical causes of memory pathology was entirely normal and the neurological consultant opined that this was a case of "hysterical" memory loss. On closer questioning about memory, Ms. B noted that she had "always had a lousy memory." For example, she reported that "my memory is like Swiss cheese." She described gaps in her memory for years in grammar school, high school, and college. She could not recall her wedding, but "knew" it happened from pictures. She could not recall her school or teachers in first to fourth grades; recalled school and teachers in fourth to seventh grades; then had limited recall of much of high school, despite many pages devoted to her accomplishments in activities in school. Her memory was reportedly more continuous once she left home for college. She stated, "I have no memory for my older brother, not at all. They [family members] said he was a holy terror, violent, out of control, using drugs . . . He's in prison now for rape." She attempted to be "perfect" but would have puzzling lapses of recall in school, for events friends reported to her, and embarrassing sexual situations that went against her moral values. These problems continue into the present with her being told of excellent presentations she has made at work, for which she denied any recall. She reported lack of any artistic ability, but, at times, she can paint brilliantly, "I do it, I can just do it," moving her hand in an elegant depiction of a brush stroke. Old friends make fun of her difficulties with recall of vacations and other events, even when there are photos of the events. She reported feeling

"different ages at different times, sometimes more than one at the same time." She frequently found herself having driven somewhere without intention to go there, or recall of the drive. She has found many items in her closet that she cannot recall purchasing, including duplicates of clothing she already owns. She has been approached by people whom she does not know, but insist that they have done things with her, gone places with her, and so on. She rationalizes this, and many other memory problems, as "well, isn't this just normal, to have this lousy memory . . . maybe I have a face that's familiar, that's all." After beginning psychotherapy, Ms. B identified a number of self-states that accounted for many of these unrecalled situations. She also recalled memory of violence and chaos in her family, as well as severe, early, repeated sexual abuse by her now imprisoned brother—reportedly ignored by her parents, despite finding evidence of the abuse on many occasions.

COGNITION IN DISSOCIATIVE IDENTITY DISORDER'INTER-IDENTITY AMNESIA. Dysfunctions of memory are a central feature of the dissociative disorders. Dissociative identity disorder, with its apparent web of complex amnesias among personality states, was the focus of early attempts at experimental investigation. At the end of the 19th century, Theodule Ribot first articulated the most parsimonious neuropsychological explanation for these phenomena: that these amnesias are examples of state-dependent learning and retrieval. Inter-identity amnesia has been studied as a paradigm for dissociative amnesia in dissociative identity disorder since the 19th century. With the modern rediscovery of dissociative identity disorder, this paradigm was used in a variety of studies to understand the nature of dissociative amnesia in this disorder.

A 1985 NIMH study assessed nine patients with dissociative identity disorder and ten matched controls, the latter tested as themselves and in a simulated personality state. They acquired data on the separateness of memory between pairs of reportedly mutually amnesic personality states by measuring intrusions from categorically similar word lists learned by the other personality states. Patients with dissociative identity disorder were more likely to compartmentalize the stimuli learned, whereas those mimicking dissociation showed far less evidence of information partitioning.

Conversely, in some recent studies of memory and amnesia in dissociative identity disorder, cognitive researchers have not been able to document amnesia between self-states that are mutually amnestic subjectively by using a variety of implicit and explicit memory paradigms. In one study, feigning control subjects familiar with dissociative identity disorder showed lack of priming in an implicit memory task because they "knew" they were supposed to be amnestic, although actual dissociative identity disorder subjects *did* show normal priming.

Rafaele Huntjens and colleagues compared individuals with dissociative identity disorder who reported mutually amnestic identities with simulator and normal controls. Studies included tests of neutral episodic information, perceptual and conceptual priming, procedural memory, transfer of trauma-related words, and stimulus valence as shown by affective priming.

They found no objective evidence of inter-identity amnesia on any of these tests and concluded that dissociative amnesia in dissociative identity disorder might have more to do with subjective appraisal and "metamemory" than with actual lack of accessibility of memory between identities.

Notwithstanding methodological limitations, individual case studies using autobiographical memory paradigms have found marked alterations in autobiographical memory in dissociative identity disorder, both for global memory of life history and for memory between identities. Accordingly, Huntjens and colleagues also studied autobiographical memory in dissociative identity disorder with reportedly mutually amnestic pairs of self-states. Here, reportedly mutually amnestic personality states were conceptualized using the constructs of apparently normal identity and traumatic identity from the Structural Theory of Dissociation (see section on Etiology), with one state reporting recall of traumatic memories and the other claiming complete amnesia for these. In two standard autobiographical memory paradigms, these researchers could not document inter-identity amnesia for autobiographical memory between the pairs of self-states.

However, in the second study, individuals with dissociative identity disorder were compared with several control groups, including participants with PTSD resulting from multiple episodes of childhood sexual abuse and maltreatment but without dissociative identity disorder. The groups with PTSD and dissociative identity disorder both showed a pattern of *overgeneral autobiographical memory* in which they had difficulty retrieving a specific event from memory that is located in a certain time and place. This pattern may be related to the wish to avoid recalling specific information about traumatic memories that may lead to increased distress, triggering flashbacks, and other intrusive symptoms. This study concluded that an overgeneral memory process is important for maintaining the memory pathology of traumatized individuals, especially those with severe, multiple early life traumas. This supports the view that dissociative identity disorder is a complex posttraumatic psychopathology, related to the dissociative subtype of PTSD.

These studies have major limitations. The notion of relatively stable, fixed "2-way" amnestic identities is based on the classical notion of dissociative identity disorder as a set of relatively unchanging, structured "personalities" whose characteristics can be found reliably on repeated assessment. This contrasts with the modern understanding of dissociative identity disorder as a posttraumatic developmental disorder involving a relatively dynamic self-state system derived from a variety of intrapsychic, interpersonal, and psychosocial needs over developmental time in sociocultural context, and a phenomenology based on overlap, interference, and intrusion, not just switching. Indeed, in another study, using implicit

and explicit memory tasks, researchers found that individuals with reported *co-consciousness* (i.e., the ability to completely share memory and information) actually *did* have inter-identity amnesia on explicit memory tasks between self-states.

Dissociative identity disorder patients may therefore not always reliably report amnesia or co-awareness between self-states, and may not be reliable reporters of the subjective experience of self-states. For example, in a single case study of a person with dissociative identity disorder, the person was randomly signaled by a beeper and filled out mood and activity rating scales in real time, as well as information pertaining to the personality state that was "out." These completed rating scales disagreed with the identities' self-professed mood and activity reports.

Studies of inter-identity amnesia do not test autobiographical memory naturalistically in dissociative identity disorder, in which the patterns of amnesia among self-states are complex, variable, and can differ based on the overlap and interference among states at a given time, current levels of stress, the level of active PTSD symptoms, and even during stressful psychotherapy sessions. Notwithstanding the difficulties involved, it may be more useful to devise studies using autobiographical memory paradigms that more globally and naturalistically evaluate memory problems, state shifting and switching behaviors, and level of PTSD symptoms in individuals with dissociative identity disorder without necessarily focusing on which self-state does or does not report differential recall at a given time. For example, the finding of overgeneral autobiographical memory processes across dissociative identity disorder self-states, unrelated to inter-identity amnesia, provides more useful information about the disorder, its relationship to other posttraumatic disorders, and what to focus on in future research.

In this regard, in a directed-forgetting paradigm between subjectively amnestic dissociative identity disorder identities, there was evidence supporting explicit memory disruption in a state-dependent fashion between identities, indicating not only possible encoding problems but retrieval inhibition between self-states. This disruption was partial, not complete, which may indicate that amnesia between states is not rigidly compartmentalized. This is consistent with the view that in dissociative identity disorder access to memory between self-states may vary in complex ways depending on a variety of factors. In this study, retrieval inhibition began to break down with more emotionally charged stimulus words. This confirmed the clinical observation that the respondent attempted to suppress posttraumatic intrusions, but, ultimately when unable to do so, experienced flashbacks and other intrusive symptoms. These findings are consistent with the psychological assessment and fMRI study of fearful faces described above.

Process Symptoms. Dissociative process symptoms include depersonalization and derealization, dissociative hallucinations, passive influence and interference experiences, and dissociative cognition (Table 20–14). Some authorities include dissociative alterations in identity under this category. Depersonalization and derealization are commonly reported by persons with dissociative identity disorder and may include profound out-of-body experiences, as well as observing oneself from within one's mind. Patients frequently report feeling spaced out or disconnected from themselves and others. The world is perceived as distant or unreal, with a hazy or foggy quality. Patients may report feeling at times as if they exist in a waking dream state. Out-of-body experiences commonly take the form of watching oneself from a distance (inner, as well as outer), as if observing another person, with little or no ability to affect one's actions.

Dissociative Hallucinations. Dissociative hallucinations most commonly take the form of voices heard as originating from within the person, as opposed to coming from the outside, sometimes termed pseudohallucinations. Individual-hallucinated voices often represent a distinctive age and gender, and in possession-form presentations a type of agent, such as a demon or deity. These voices may comment negatively about the patient, argue with each other, command the patient to perform certain acts, discuss neutral topics, or sometimes provide useful information or comfort. The voices may be indistinct, but described with a personified quality, such as being at a party but not being able to make out all the conversations around one.

Patients with non–possession-form dissociative identity disorder generally recognize that the voices are hallucinations and may be reluctant to reveal their existence for fear of being considered psychotic or "crazy." Possession-form patients may share the same fear, but usually attribute the voices to culturally accepted spiritual agents. Many patients report some ability to ignore or disregard the voices unless they are stressed. Ultimately, dissociative voices often are identified with specific personality states or possessing agents. Visual hallucinations often take the form of detailed images with traumatic or frightening content. Other visual hallucinations may be understood as depicting the identities or may even have a complex artistic quality. Tactile, somatosensory, gustatory, and olfactory hallucinations may also occur, leading to misdiagnoses of seizure disorder or other brain disorders. Intrusive posttraumatic flashbacks and images may also be experienced as complex multimodal hallucinations. Musical hallucinations also may occur.

Dissociative identity disorder individuals also report *negative hallucinations*—in which external percepts and stimuli are not consciously registered. This can occur in any sensory modality, including auditory, visual, tactile, gustatory, and olfactory. Dissociative identity disorder

individuals also frequently report the volitional ability to block out or attenuate pain, called *voluntary analgesia*.

Most positive and negative hallucinations in individuals with dissociative identity disorder are likely homologous to phenomena that can be experienced in deep trance states among highly hypnotizable individuals, as well as posttraumatic phenomena, and are not a manifestation of a psychotic process. Unfortunately, in current practice, hallucinatory phenomena of any kind usually are diagnosed as psychotic, ignoring the large literature on nonpsychotic hallucinations in the general population and "hysterical hallucinations" in the older literature. In addition, PTSD flashbacks are defined in DSM-5 as dissociative; flashbacks commonly include hallucinations in all modalities, particularly visual, olfactory, tactile, and somatosensory.

Passive Influence Symptoms and Symptoms of Overlap, Intrusion, and Interference among Self-States. Passive influence and interference symptoms include many apparent first-rank Schneiderian symptoms (FRS), such as audible thoughts, voices arguing with each other, influences playing on the body, thought withdrawal and insertion, and made feelings, impulses, and actions. These symptoms were once considered to be pathognomonic of schizophrenia, but they can also be found in patients with affective, organic, and dissociative disorders. As noted in the section on Nosology, passive influence symptoms are very common in dissociative identity disorder. Studies have shown that these symptoms are more common in individuals with dissociative identity disorder than in those with psychotic disorders. Individuals with non–possession-form dissociative identity disorder typically experience the agents of the passive influence symptoms as internal, not external, as they are in psychotic disorders. Possession-form presentations, by contrast, are usually experienced as external and "real," complicating their differentiation from psychotic conditions. Patients may report strong affects or impulses that they experience without a sense of personal ownership but with a peremptory sense of intrusion and control. Individuals with dissociative identity disorder generally do not have delusional explanations for these experiences. They may feel confused, puzzled, or ashamed of them. Personification is commonly part of the explanation, such as "I feel like someone else wants to cry with my eyes."

Dissociative Thought Disorder. The recognition that dissociative patients frequently manifest subtle, but often clinically significant, cognitive impairments emerge from clinical research with psychological and cognitive test batteries. Research using projective testing finds distinctive cognitive process markers, including evidence of confusing and contradictory responses to the same stimulus. Distinctive responses to

standardized projective testing can often be helpful in distinguishing dissociative patients from other diagnostic groups, such as patients with affective disorders, nondissociative forms of PTSD, personality disorders, psychotic disorders, and factitious disorders. Other forms of dissociative identity disorder thought abnormalities include profound trauma-based cognitive distortions, such as pervasive self-blame for abusive trauma. These posttraumatic and dissociative cognitive distortions may seem so intense as to be virtually delusional at times. However, in most cases, one or more aspects of the mind will report understanding that these attributions are incorrect, or that they would be incorrect if someone else believed them. Dissociative identity disorder individuals may also appear to have derailment and loosening of associations if there is shifting of states that overlap and interfere with each other.

Dissociative Alterations in Identity

NON–POSSESSION-FORM PRESENTATIONS. Clinically, *non–possession-form* dissociative alterations in identity may first manifest through odd first-person plural or third-person singular or plural self-references. Patients may refer to themselves using their own first names or make depersonalized self-references, such as "the body," when describing themselves and others; for example, "The father hurt the body so she was upset. We tried to protect her, but it didn't work." Patients often describe a profound sense of concretized internal division or personified internal conflicts between self-aspects. In some instances, these identities may have proper names or may be designated by their predominant affect or function, for example, "the angry one" or "the wife." Patients may suddenly change the way in which they refer to others, for example, "the son" instead of "my son."

A set of behaviors, collectively referred to as *state shifting or switching behaviors,* may be manifest during evaluation or therapy sessions. These behaviors may be accompanied by intra-interview amnesias, in which the patient does not seem to recall or is confused about the process and content of an assessment or therapy session. These *microdissociative episodes* may manifest via abrupt shifts in train of thought or sudden inexplicable changes in affect or in rapport, and may then quickly shift back again. A variety of physical signs, including pronounced upward eye rolls or bursts of rapid blinking and eyelid fluttering, may occur in conjunction with microdissociative episodes. The patient's tone of voice and manner of speaking, posture, and demeanor may show subtle to marked alteration. When clinicians encounter evidence of possible microdissociative episodes, they should seek to clarify what the patient is experiencing and can recall using nondirective, open-ended questions. Dissociative amnesia is a common concomitant of these episodes; sometimes they may extend for a

few minutes or even the entire psychotherapy session.

These cognitive, behavioral, and physical shifts are manifestations of personality state shifting/switching or overlap and interference between states, or both. Authorities have long cautioned that self-states should not be regarded as separate people. Rather, the personality states are conceptualized as relatively stable and enduring patterns of thought, emotion, and potential behavior that are largely unintegrated with each other and are often in direct conflict.

The set of personality states, usually referred to as the *personality- or self-state system,* constitutes the personality of the individual. Most psychotherapeutic work is directed toward this larger personality system and thus toward the individual as a whole. Much has been made of the psychological and physiological differences among the personality states of individuals with dissociative identity disorder and, as noted above, popular accounts emphasize the presentational differences among them. Laboratory studies support clinical accounts of significant differences; however, as the inter-identity amnesia studies show, much general information and many functions and abilities are shared in common across personality states, and indicate the fundamental unity of the mental processes of the dissociative individual. This provides the foundation for therapeutic efforts directed at the development of a more consciously integrated sense of self in persons with dissociative identity disorder.

Individuals with the disorder show considerable variability in terms of the complexity and therapeutic tractability of their personality state systems. As noted above, several self-state types are commonly reported, including child states; internalized persecutory states, who inflict pain and may attempt to kill the individual; and depleted and depressed personality states, who may function as the apparent primary self-state with respect to the world at large, commonly carrying the legal name of the person. It is a mistake to conceptualize this latter state as the "real person" accompanied by "his/her" personality states. These everyday states are usually created to adapt to functioning in the world, such as in school, work, marriage, and so on.

Different personality states often encompass painful psychological issues for the individual, frequently taking the form of polarized pairs representing antithetical positions, although self-states embodying more neutral and conflict-free processes also commonly occur. In some individuals with dissociative identity disorder virtually every aspect of mental life is personified and structuralized in this form.

POSSESSION-FORM PRESENTATIONS. *Possession-form* presentations may be subtle or dramatic. Onset can be gradual and nonspecific (e.g., diverse somatic complaints) or sudden and specific, in the form of an abrupt transition to the identity state. If the presentation is gradual and

nonspecific, transition to the identity may occur spontaneously or after contact with an indigenous healer specializing in the treatment of pathological possession.

Dramatic onsets may be characterized by semi-purposeful movements (e.g., thrashing, falling to the ground, shrieking); aggressive or violent actions directed at self or others accompanied by derogatory verbalizations; and/or specific gestures, comments or requests denoting the appearance of a known possessing agent by (1) reference to standard attributes of culturally recognizable figures or (2) the name and degree of relation of deceased family members or acquaintances. At times the possessed individual "acts out" some narrative aspect of the possessing agent's story in an effort to communicate to his/her social network the suffering inflicted by the spiritual being or the desires of the possessing agent. For example, a young woman, experiencing herself as afflicted by the spirit of a dead adolescent girl who committed suicide by jumping in a well, engaged in suicidal behaviors that evoke these past events. Alternatively, individuals possessed by high-status spirits may demand deference from family members.

Possession by the spiritual agent is frequently episodic, resulting in alternation between the usual identity of the subject and the altered state. The subject in his/her usual identity appears in a daze, exhausted, distressed or confused about the situation. He/she may report visual or auditory perceptual disturbances regarding the possessing agent, as well as "unconsciousness" and partial or total amnesia for the altered state. States of consciousness may alternate continuously within one episode over hours to weeks, with each state lasting minutes to hours. This pattern may persist intermittently over months or years.

DISSOCIATIVE IDENTITY DISORDER CASE 3

Ms. C is a 23-year-old bilingual woman (Spanish-English) of Hispanic Caribbean descent who was born and raised in a large eastern city of the United States. Her family was active in an Afro-Catholic religious cult involving normative possession states until, when Ms. B was 5 years old, all members converted to Evangelical Protestantism. Since early childhood, she was thought to have spiritual abilities allowing her to predict the future in her dreams and to identify other co-religionists suffering from demonic possession (by seeing "shadows" around them). These were considered spiritual gifts and were highly prized by her family and congregation. Since age 4, Ms. B was the victim of sexual abuse in the form of fondling without penetration by older acquaintances. She also received harsh corporal punishment from her parents leading to bruises but not fractures. Conflicts with her family led to a suicide attempt at age 15 with lethal intent (medicine cabinet overdose). This led to a 3-week inpatient hospitalization and outpatient follow-up with antidepressants for a diagnosis of major depression with panic attacks. After 3 months, however, she self-discontinued treatment. At age 21, Ms. B began to question her faith and to ask God in prayer whether he and the demons she was seeing were real. In this context, she experienced a prolonged possession episode that was involuntary and extremely distressing. Onset came during sleep, when she felt a demon attempt to enter her body but not fully succeed. This occurred for several consecutive nights and was followed during the day by the onset of voices mumbling and calling her name. Several days later, she felt "many spirits coming inside of me" and taking over control of her actions. For 3 months, every day these "demons" spoke through her mouth in a deep and "not human"

voice—"like a monster or several voices at once"—and communicated with her family and other church members who came to pray around her. During most of these episodes, Ms. B was aware of the surrounding events but was unable to control her speech or actions. She reported amnesia for some of the possession episodes, reported recall of most of them. She had no visual hallucinations of the possessing demon. However, she heard him speaking to her, saying that he would "drive her crazy." At those times, Ms. B felt that God sent her a message of reassurance. Once, she felt evil spirits were pulling her soul out of her body. She saw an angel sent by God to protect her standing at the foot of the bed—it was "invisibly white, see-through white" with wings and no face. This helped calm her. Multiple attempts at exorcism conducted by local church leaders were unsuccessful. A prominent pastor was called in from Latin America to assist. The demon claimed to recognize the pastor and alluded to their past encounters during other cases of possession. His exorcisms were also unsuccessful. After 3 months, while she was participating in a religious retreat, she felt God sent her a message, "today is the day." At that point, "the demon trembled and left." She has not suffered another possession state since then, 2 years ago. However, she intensely phobic of the dark and has auditory and visual illusions when hearing background noises in an otherwise silent room ("sound like voices") or if sees herself in a mirror from a distance ("my face is scary, as if I had a demon inside again"). One year ago, she had an episode of conversion hemiparalysis that came on gradually over the course of a day and then disappeared. She continues to suffer from major depression and panic attacks.

Table 20–16compares possession-form and non–possession-form presentations of dissociative identity disorder.

Symptoms Related to Spontaneous Autohypnotic Phenomena. Individuals with dissociative identity disorder commonly exhibit or describe symptoms of involuntary autohypnotic phenomena, consistent with their high scores on standardized hypnotizability scales (Table 20–15). These phenomena include spontaneous trance states that can be clinically disabling, spontaneous or voluntary anesthesia and analgesia, negative hallucinations, spontaneous age regressions, and the *trance logic* or literal-mindedness of the hypnotized subject, as well as tolerance/rationalization of logical contradiction in the hypnotic state. Other similar symptoms overlap with dissociative process symptoms (e.g., hallucinations, thought disorder, child identities) and PTSD symptoms (e.g., behavioral re-experiencing episodes with multimodal hallucinations and age regression to the time of the trauma).

Other Associated Symptoms. As noted above, non–possession-form dissociative identity disorder is conceptualized as a paradigmatic posttraumatic disorder, with most patients meeting DSM-5 diagnostic criteria for PTSD, major depression, current or past substance use disorders, and somatic symptom disorders, as well as commonly engaging in self-destructive and suicidal behavior. It is likely that most individuals with dissociative identity disorder and PTSD will meet the DSM-5 diagnostic criteria for the dissociative subtype of PTSD, but this will need to be assessed in future research. Table 20–17 enumerates associated symptoms of dissociative identity disorder.

Table 20–18lists the most important differential diagnostic

considerations for dissociative identity disorder. In particular, one must differentiate disorders that are comorbid with dissociative identity disorder from those that must be differentiated from dissociative identity disorder. For example, as enumerated above depression, PTSD, substance abuse, and somatic (previously somatoform) symptoms are very common comorbidities with dissociative identity disorder.

Numerous recent studies have shown a robust relationship between certain mood disorders and an antecedent history of trauma, particularly childhood sexual abuse. Major depression is increasingly understood as one of the outcomes following traumatic experiences. There are genetic studies that have proposed an interaction between specific SNPs and childhood sexual abuse that results in depression in later life—albeit there have been substantial critiques of these studies. Most individuals with dissociative identity disorder meet criteria for a mood disorder, usually in the unipolar depression spectrum. Careful assessment of the affective state of the individual commonly finds a variety of emotions and cognitions as part of the "depression." These may include intense shame, profound grief, posttraumatic dysphoria, demoralization, survivor guilt, and distress at one's life situation (see section on Target Symptoms for Medications in the Treatment of Dissociative Identity Disorder-Mood Disorder Symptoms).

Table 20–16.
Comparison of Non–Possession-Form and Possession-Form Presentations of Dissociative Identity Disorder

Characteristic	Non–Possession-Form	Possession-Form
	Internal aspects of the person, usually representing idiosyncratic ages, genders, and functions related to the person's life and intrapsychic needs (e.g., early childhood, adolescence, perpetrators, protectors)	External agents usually of supernatural or spiritual origin following collectively defined identities (e.g., the spirit of a dead person or demons, deities, spiritual beings, animals, or mythical figures)
Behavior of alternate identity	Patterned by characteristics of identity: childlike, aggressive, fearful, sexual	Patterned by collective representations of identity: deity, demon, ghost
Elaborateness of presentation	Typically reticent, inhibited, obsessional	May be dramatic (e.g., shrieking, thrashing) or reticent
Relationship with trauma	Linked systematically to extreme, chronic, and	Associated with traumatic events and nontraumatic stressors (e.g., romantic disappointments). However,

	early childhood maltreatment	systematic studies of childhood trauma are lacking
Onset and presentation to others	Typically gradual and hidden, upon accumulation of traumatic exposures	Gradual and nonspecific (e.g., diverse somatic complaints) or abrupt and specific (e.g., sudden transition to altered identity)
Course	Chronic, in relapsing and remitting pattern	May be transient and episodic, although chronic, relapsing course has been observed
Extent of misdiagnosis	In high-income countries, an average of 3.9 incorrect diagnoses and 7–12 years in care before correct diagnosis	Unknown
Treatment	Multiphasic treatment similar to other complex posttraumatic dissociative disorders	Often exorcism, initiation into spirit devotion cult, and/or specialized role in healing practice, typically supervised by indigenous healer. However, exorcism may fail, and may be experienced as traumatic to the person, with alienation from religion and spiritual life. Then, may respond to phasic treatment as per the non–possession-form, and may experience self-state system phenomena more typical of non–possession-form.

Frequent, rapid mood swings are common, usually experienced as occurring within minutes or hours. These are usually due to posttraumatic and dissociative phenomena, and are much more rapid than even the most rapid cycling documented in bipolar disorder. There may be considerable overlap between PTSD symptoms of anxiety, disturbed sleep, and dysphoria and mood disorder symptoms.

OCD symptoms are also commonly found in individuals with PTSD. Obsessive-compulsive personality traits are common in dissociative identity disorder, and intercurrent OCD symptoms are regularly found in individuals with the disorder, with a subgroup manifesting severe OCD symptoms. OCD symptoms commonly have a posttraumatic quality, for example, checking repeatedly to be sure that no one can enter the house or the bedroom, compulsive washing to relieve a feeling of being dirty because of abuse, and repetitive counting or singing in the mind to distract from anxiety over being abused.

Child and Adolescent Presentations (also, see section on Developmental Model of Dissociative Disorders). A growing clinical research literature documents the diagnosis and treatment of child and adolescent dissociative disorders, including dissociative identity disorder. In many respects, children and adolescents manifest similar core dissociative symptoms and secondary clinical phenomena as adults. Age-related differences in autonomy and lifestyle, however, may significantly influence the clinical expression of dissociative symptoms in youth. For example, dissociative

amnesias and perplexing forgetfulness may be more apparent in school situations than in work or family life. Younger children, in particular, have a less linear and continuous sense of time and are often unable to self-identify dissociative discontinuities in their behavior. Informants, such as teachers and relatives, may be useful in documenting dissociative behaviors.

Several normal childhood phenomena, such as imaginary companionship, elaborated daydreams, or participation in normal possession experiences as part of religious rituals, must be carefully differentiated from pathological dissociation in younger children. For example, preadolescent children with dissociative identity disorder may manifest fewer obvious switching behaviors than adolescents or adults. The clinical presentation may be that of elaborated or autonomous imaginary companions, with these companions taking control of the child's behavior. This is often experienced as passive influences or voices in the child's mind commanding the child to behave in certain ways. A case example illustrates the phenomenology and multigenerational nature of dissociative identity disorder.

DISSOCIATIVE IDENTITY DISORDER CASE 4

Ms. D had been in treatment for dissociative identity disorder for several months when she voiced concern that her 8-year-old daughter was "depressed." One day, to her psychiatrist's surprise, she brought her daughter to therapy. During the session, the child began to fidget, then got up and began to walk around the room. The mother suddenly appeared to shift in demeanor, appearance, and voice tone, shouted at the girl in a harsh and angry voice, grabbed the daughter's arm tightly, and began to angrily slap the child. The little girl attempted to curl up in a fetal position, wailing like a much younger child, while the psychiatrist jumped into the fray, and literally physically separated the mother from beating the child. As he pulled the mother away, the little girl crawled under a table, and said, "I hate you! I hate you!" Both then appeared dazed, the mother wandered back to her chair, the little girl emerged from under the table, looking confused, but appearing more her actual age. When queried about what had happened, both claimed no recall of the events, even when the psychiatrist detailed for them what had just happened. The psychiatrist then arranged for another family member to pick up the little girl and stay with her. He probed the mother until an angry self-state emerged spitting, "the little s–t isn't mine, it's *hers*!" and took responsibility for the events in the office, and admitted to beating her daughter on numerous occasions. The psychiatrist insisted that the patient report herself to the Department of Social Services (DSS) to investigate her ability to care for her daughter. He made the safety of the child a central part of the patient's treatment and assisted with the DSS investigation, resulting in additional supervision for the mother and child, but placement remaining in the home. He also interviewed the father, a passive man, whose work took him away from home much of the time, and who had abdicated child rearing to his wife. A colleague evaluated the daughter with the little girl reporting "people in my head," whom she said talked to her, and "told [her] what to do." The little girl described "2 ladies who are twins" in her mind, "one is a nice mommy, and the other is mean, and you cannot tell them apart, so you never know when one or the other is around." The child said that "a lot of the people in my head are twins, one does it one way, the other the opposite. When that mean lady is around, my twin little ones are there." The psychiatrist learned that the mother's self-state system also was based on many groups of subjectively experienced twin identities, including the maternal and abusive identities, that reportedly called themselves by the same names as the little girl's twin mother and abusive identities. The mother told the psychiatrist that *her* mother also likely shares the diagnosis of dissociative identity disorder, and was actually born with a twin sister who died shortly after birth.

Table 20–17.
Associated Symptoms Commonly Found in Dissociative Identity Disorder

- Posttraumatic stress disorder symptoms
  - Intrusive symptoms
  - Severe dissociative flashbacks and nightmares
  - Avoidance and numbing symptoms
- Persistent and exaggerated negative beliefs/expectations and cognitions about self/world
- Persistent negative emotional state (especially shame/guilt, fear, sadness, anger)
- Detachment/estrangement
- Hypervigilance
- Reckless/self-destructive behavior
- Problems with sleep and concentration
- Symptoms of Somatic Symptom Disorders
  - Conversion and pseudoneurological symptoms, especially, seizure-like episodes
  - Somatic Symptom Disorder
  - With predominant pain symptoms such as:
    - Headache, abdominal, musculoskeletal, pelvic pain
    - Psychological Factors Affecting Other Medical Conditions
    - Medical Disorders Associated with Adverse Childhood Experiences (see Table 20–7)
    - Perimenstrual disorders (sexual abuse)
    - Irritable bowel syndrome (sexual abuse)
    - Gastroesophageal reflux disease (sexual abuse)
    - Morbid Obesity (sexual abuse)
  - Affective symptoms
    - Depressed mood, dysphoria, or anhedonia
    - Brief mood swings or mood lability (within minutes/hours)
    - Suicidal thoughts and attempts
    - Guilt and survivor guilt
    - Profound shame
    - Helpless and hopeless feelings
  - Obsessive-compulsive symptoms
    - Ruminations about trauma and its impact on the person
    - Obsessive counting, singing
    - Arranging
    - Washing
    - Checking
    - Compulsive Body Movements

Table 20–18.
Differential Diagnosis of Dissociative Identity Disorder

- Comorbidity vs. differential diagnosis (see discussion in text of outcomes of childhood trauma and Table 20–7)
- Affective disorders
- Psychotic disorders
- Anxiety disorders
- Posttraumatic stress disorder
- Personality disorders
- Neurocognitive disorders
- Neurological and seizure disorders (e.g., multiple sclerosis, tic disorders, movement disorders)

Somatic symptom disorders
Factitious disorders
Malingering
Other dissociative disorders
Gender dysphoria (comorbid or pseudogender dysphoria based on opposite gender self-states)

Pathology and Laboratory Examination

A clinical mental status examination (Office Mental Status Examination) based on the various symptom clusters described in Tables 20–13 through 20–15 can be helpful in the diagnosis of dissociative identity disorder. Questioning about amnesic, depersonalization/derealization, autohypnotic, PTSD, mood, passive influence, and somatic symptoms usually precedes detailed inquiry about more overt symptoms related to the identities. Careful in-depth questioning about these phenomena may readily bring forth information about process symptoms of dissociative identity disorder. For example, discussion of complex amnesia experiences may readily lead to questions about how purposive behaviors can occur without apparent memory. Discussion of PTSD symptoms may lead to manifestations of identities related to traumatic experiences.

The development and increasing use of screening instruments and standardized diagnostic interviews (see Table 20–2) have contributed to the rising number of identified cases of dissociative identity disorder over the last decade. Before the development of these assessment tools hypnotic or amobarbital interviews were used to diagnose dissociative identity disorder. Due to the current academic and forensic controversies surrounding dissociative disorders and trauma memory, as well as most clinicians' lack of familiarity with these modalities, it is prudent to reserve these interventions for emergency situations when other methods of assessment have failed; for example, in a female patient who is emergently hospitalized after engaging in repeated, dangerous nocturnal fugues and who cannot account for these behaviors despite intensive interviewing. These interventions should optimally be conducted by a clinician experienced in their use and in the differential diagnosis of dissociative disorders. Clear, informed consent should be obtained for use of these interventions for diagnosis of dissociative identity disorder and recall of traumatic experiences. Pharmacologically facilitated interviews (e.g., amobarbital interviews) are now considered conscious sedation by The Joint Commission and can only be performed with the assistance of an anesthesiologist.

Differential Diagnosis

On average, more than 6 to 12 years pass in high-income countries between first psychiatric contact and the diagnosis of dissociative identity disorder. The time lag between symptom onset and referral for treatment in

possession-form presentations may vary, depending on the dramatic or gradual nature of the symptoms. Rather than displaying histrionic flamboyance, it is estimated that over 95 percent of individuals with non–possession-form dissociative identity disorder are reticent about revealing their dissociative symptoms, especially hallucinations, amnesia, and identity divisions. They most commonly present as being relatively inhibited and obsessional, with mood and somatic complaints, and typically acquire an average of three or more psychiatric diagnoses before their dissociative identity disorder is recognized. As noted above, a subgroup of individuals with the disorder shows symptoms and interpersonal dynamics reminiscent of borderline personality disorder and may qualify for this secondary diagnosis once PTSD and dissociative symptoms are stabilized. At the time of initial diagnosis, bipolar disorder characteristics in individuals with dissociative identity disorder are commonly related to instability caused by severe dissociative and PTSD symptoms, not to a core bipolar disorder psychopathology. The presence of auditory hallucinations, disturbed thinking and behavior, confusion due to amnesic gaps, and apparent Schneiderian FRS contributes to the misdiagnosis of psychotic disorders in approximately one-half of dissociative identity disorder patients in high-income countries at some point in their psychiatric histories.

Rapid changes in affect associated with personality state switching may suggest a rapid-cycling mood disorder or schizoaffective disorder. Table 20–18 lists the most common disorders that must be differentiated from dissociative identity disorder.

However, because many of these disorders can coexist with dissociative identity disorder, as well as be mimicked by dissociative identity disorder, differential diagnosis may be a complex process. For example, a subgroup of individuals with dissociative identity disorder may have an intercurrent psychotic disorder with a mixture of psychotic thought derailment, both psychotic and dissociative hallucinations, and, at times, tangential, bizarre associations and persecutory delusions separate from posttraumatic mistrust and hypersensitivity in interpersonal contexts. Their auditory hallucinations often can be differentiated from the subjectively experienced, internalized voices of the personality states (though this would be harder to differentiate in possession-form cases). Dissociative identity disorder hallucinations are hypothesized to result from autohypnotic and posttraumatic processes, as individuals with the highest hypnotic capacity can hallucinate in all sensory modalities in deep trance. More severely ill patients may describe psychotic boundary loss with thought broadcasting and audible thoughts, ideas of influence and reference, and with delusional explanations for these phenomena. Even so, many of these patients may preserve insight that psychotic symptoms are different from posttraumatic and dissociative symptoms. However, those

patients with more severe, intercurrent psychotic disturbances may not have insight about their symptoms. Some of these patients may have a mixture of delusional and posttraumatic phenomena that may be very difficult to entangle.

Individuals with actual intercurrent bipolar disorder may experience sustained hypomanic or manic episodes—different in duration and quality from dissociative mood shifts—and more typical bipolar depression symptoms that are superimposed on the patient's baseline chronic complex posttraumatic mood state. Stabilization of the actual intercurrent bipolar disorder symptoms with mood-stabilizing medication is necessary to allow the patient to participate fully in phasic trauma treatment for dissociative identity disorder. However, depressive symptoms usually are only returned to the patient's more usual chronically dysphoric baseline.

Factitious, Imitative, and Malingered Dissociative Identity Disorder

Concerns about factitious and malingered dissociative identity disorder are common. Recently, reports are increasing of individuals claiming to have dissociative identity disorder who do not meet diagnostic criteria for the disorder when carefully assessed clinically or with structured interviews, such as the SCID-D-R. There may be a mixture of factors leading to this presentation, including misdiagnosis, factitiousness, and assumption of a social role by an abuse victim. Dutch researchers have named this *imitative dissociative identity disorder* when there does not appear to be conscious simulation. Patients may build their lives around their diagnosis and are commonly supported in so doing by concerned others and by their therapists.

Clinical indicators of falsified or imitative dissociative identity disorder may include those typical of other factitious or malingering presentations, such as symptom exaggeration, lies, use of symptoms to excuse antisocial behavior (e.g., amnesia only for bad behavior), amplification of symptoms when under observation, refusal to allow collateral contacts, legal problems, and *pseudologia fantastica*. In contrast, individuals with genuine dissociative identity disorder are usually confused, conflicted, ashamed, and distressed by their symptoms and trauma history. In forensic settings, they actually may attempt to refuse to use an insanity defense, much as many paranoid patients do. The feigning patients frequently show little dysphoria about their disorder and posttraumatic triggers and appear to be gratified by having the diagnosis. These imitative and factitious patients fit the SCM of dissociative identity disorder, with one exception: rigorous diagnostic efforts show that they do not meet diagnostic criteria for the disorder.

The SCID-D-R has been used to help distinguish genuine from feigned/imitative dissociative identity disorder in clinical and forensic contexts. On the SCID-D-R, individuals with fabricated dissociative identity

disorder tend not to describe the typical comorbid conditions, and symptoms associated with the disorder—such as amnesia, identity confusion, and identity alteration—although they may report high levels of depersonalization and derealization. Instead, they often describe symptoms in a jargon-filled way ("I'm dissociating" or "I'm switching") and cannot readily explain their subjective experience on follow-up questions on the SCID-D-R or in clinical interviews. Most frequently, they show Cluster B personality disorder traits on structured diagnostic interviews, but do not meet diagnostic criteria for a dissociative disorder.

There are no absolute clinical indicators to differentiate imitative, factitious, or malingered dissociative identity disorder from the actual disorder. Some persons with the bona fide disorder become invested as their main social role in a dissociative identity disorder identity or that of being a trauma survivor. There are patients with possession-form presentations who show clear evidence of secondary gain, such as a late adolescent newlywed who has recently moved to her husband's family of origin where she is treated badly by her mother-in-law and manages to return home for prolonged periods to treat her possession state. However, most individuals with non–possession-form presentations and many possession-form cases are reticent and secretive about their illness. In some instances of the latter, their possessing identities may remain undisclosed for some time, until a specialist healer exposes them.

There are also patients with the genuine disorder who report factitious histories, produce factitious crises, and create factitious identities. Some patients report that their only happy childhood experiences were when they were hospitalized medically. Like more typical patients with factitious disorder, they seek out and try to remain hospitalized in psychiatric and medical settings. Münchhausen syndrome and Münchhausen syndrome-by-proxy have also been reported in patients with dissociative identity disorder and their first-degree relatives. A case follows.

DISSOCIATIVE IDENTITY DISORDER CASE 5

Ms. F, a 48-year-old divorced mother of three, was referred to a trauma disorder inpatient unit by her managed care company for consultation about diagnosis and treatment. Ms. F had been diagnosed with dissociative identity disorder after doing "inner child work" with her outpatient therapist, and during her outpatient therapist's group for sexually abused women. Ms. F reported an extensive history of childhood neglect, as well as physical and sexual abuse. Ms. F had numerous hospitalizations for suicidal ideation and superficial self-mutilation. Ms. F was on disability and had been investigated several times by her state division of child welfare, who had recently placed her children in foster care. On arrival in the trauma unit, Ms. F proclaimed to patients and staff that, "I have dissociative identity disorder. I was sexually abused." She did not seem distressed by either of these ideas. She repeatedly stated, "I'm switching, I'm switching," although staff members, experienced with dissociative identity disorder patients, did not observe any of the typical signs of dissociative identity disorder switches. Staff noted that they had never observed a dissociative identity disorder patient announce switches in this way. Ms. F became irritable when staff attempted to dissuade her from speaking about this in the milieu. Ms. F was evaluated clinically, observed by nursing staff in the ward milieu, and given a battery of psychological tests, including the DES, the SCID-D, assessments of PTSD,

and personality assessments. On the clinical evaluation, Ms. F did not endorse complex chronic amnesia symptoms, amnesia for her life history, spontaneous self-hypnosis, passive influence symptoms, or inner voices. She endorsed episodes of depersonalization and derealization. She reported only "child and baby alters" but could describe little about them when questioned. She would say things like, "See, now I'm like a child. Can't you tell? That's one of my littles." She did not endorse any sense of fear, shame, conflict, or inner struggle around switching, nor were any of the typical phenomena of switching noted at these times. Ms. F became vague when discussing reported childhood sexual abuse. She became angry when confronted about contradictions in her accounts, the dramatic nature of her presentation, and discrepancies between her history and her emotional reactions. Ms. F appeared uninterested in individual or group psychotherapy that sought to teach her more effective ways to manage symptoms and increase coping. On psychological and diagnostic testing, Ms. F did not endorse most items consistent with clinical dissociation, other than depersonalization and derealization. She did not display typical PTSD or dissociative reactivity to inquiries about the abuse history, nor did she meet diagnostic criteria for PTSD. She did not show behaviors typical of dissociative identity disorder patients during a serendipitous event while being tested, remaining unaffected when another patient suddenly required aggression management on the ward near the testing room, whereas other patients reacted with severe PTSD and dissociative symptoms. Ms. F was diagnosed with DSM-IV-TR factitious disorder with psychological symptoms, mood disorder NOS and personality disorder NOS with histrionic, borderline, and dependent features. However, Ms. F did not appear to be consciously feigning and actually seemed to believe in her self-reports of switching, self-states, and a history of childhood sexual abuse. Both Ms. F and her outpatient provider were resistant to the new diagnostic formulation.

Recently, a group of researchers has explored the standard inventories used in forensic settings for detection of malingering, such as the Structured Interview for Reported Symptoms (SIRS), developed by Richard Rogers. They found that the SIRS misclassified a number of individuals with SCID-D-R–diagnosed dissociative identity disorder as feigning. Rogers replicated this finding and modified the SIRS, now as the SIRS-II, to classify dissociative trauma survivors more accurately.

At this time, assessment of possible factitious or malingered dissociative identity disorder presentations—or factitious or malingered amplification of symptoms in persons with actual dissociative identity disorder—requires a complete clinical and psychometric evaluation, in addition to a review of old records, data from collateral informants, and other forms of assessment. Psychometric instruments, which usually must be administered by a psychologist familiar with how symptoms of dissociation present on standard measures, generally include the DES, MID, SCID-D-R, Rorschach, Thematic Apperception Test (TAT), Trauma Symptom Inventory (TSI), Minnesota Multiphasic Personality Inventory (MMPI)—although this also may misclassify dissociative individuals—Personality Assessment Inventory (PAI), SIRS-II, additional assessment measures of malingering, and assessment measures for depression and PTSD, among others.

Course and Prognosis

Little is known about the natural history of untreated dissociative identity disorder. Small studies of patients with dissociative identity disorder

diagnosed in middle age and in geriatric populations have shown that severe dissociative identity disorder symptoms can persist or can appear in older-age patients. A relapsing and remitting course is common and appears to be the modal pattern in possession-form cases. Also, patients can more or less successfully mask or suppress symptoms for periods of time. Both of these latter phenomena may be mistaken for complete spontaneous remission of the disorder. Possession-form presentations tend to be described in case series as more transient and episodic than non–possession-form cases. It is possible that the emergence of the possession-form type is facilitated by the same porosity of the self that patterns the external attribution of alternate identities, leading to more transient states after a wider range of stressors. However, it is also possible that possession-form cases only seem transient and episodic, but that this is just the overt presentation of phenomena that persist in more subtle form over time. Support for this possibility comes from data on relapse of possession-form presentations. In India, for example, relapse within 15 years was not infrequent, and was also recorded after two to three decades. Four of 15 subjects in one study were found to have suffered previous episodes, and one subject in another series reported three consecutive postpartum attacks. In China, individuals with chronic possession-form illness may be hospitalized after a history of recurrent episodes. Systematic research with larger samples over time is required to clarify this issue.

Some individuals with untreated dissociative identity disorder remain involved in abusive relationships or violent subcultures, resulting in their retraumatization and likely that of their children, with the potential for family transmission of the disorder (see Case 4). Experience with many dissociative identity disorder cases suggests that there are diverse subgroups of afflicted individuals. Some function very well for long periods while others have severely impaired and dysfunctional life trajectories often beginning early in development. The former may only show symptoms when severely stressed or overwhelmed; the latter may be continuously and severely symptomatic and come to clinical attention as children or adolescents. As described in the section on Nosology, there are high rates of disability in dissociative identity disorder patients in the community and in partial hospital and inpatient settings.

Patient presentations and prognosis vary somewhat across the life span. Children with non–possession-form dissociative identity disorder show many dissociative symptoms and behaviors but typically have fewer and less-crystallized personality states that are less invested in their individuality. If diagnosed early, and removed from abusive or neglectful environments, childhood DID may resolve relatively rapidly when appropriate treatment is provided. In adolescence, personality states become more distinct and more invested in their autonomy, as tasks of separation and individuation and sexuality may be highly stressful for

dissociative identity disorder adolescents. Additional self-states may appear that are associated with life stresses, such as academic, athletic, social, or sexual challenges, and the personality system dynamics may become more complicated and polarized. In general, adolescents have a poorer prognosis than children or adults in part because, like other adolescents, they are often not invested in their treatment and are reticent to open up to adults. Better outcome has been reported in adolescents with dissociative identity disorder when their families were successfully engaged in treatment. Young adults typically present in crisis and have a layering of affective, somatic, posttraumatic, and personality disorder symptoms in addition to their core dissociative pathology.

First presentations of dissociative identity disorder in older patients frequently involve a life event, such as the loss of a job, death of a parent, or a revictimization, such as a rape, that reactivates earlier conflicts and destabilizes the self-state system. Achieving sobriety in a substance-abusing patient with dissociative identity disorder may precipitate overt symptoms of the disorder and of PTSD. In adult patients, there appear to be treatment-responsive and treatment-refractory subgroups of individuals. Time course of improvement also may vary, with some patients improving relatively quickly and others requiring intensive treatment over long periods. Some healing traditions for possession-form presentations reformulate the person's suffering into beneficent individual and communal practice via initiation into a spirit devotion cult—such as the Siri cult of South India—or education into the roles of oracle (diviner), exorcist, or, rarely, avatar (divine incarnation).

Prognosis is poorer in patients with comorbid organic mental disorders, actual psychotic disorders, and severe medical illnesses. Refractory substance abuse and eating disorders also suggest a poorer prognosis. Other factors that usually indicate a poorer outcome include significant antisocial personality features, current criminal activity, ongoing perpetration of abuse, and current victimization—including being trafficked, with refusal to leave abusive relationships or situations, whether out of fear or for other reasons. Repeated adult traumas with recurrent episodes of ASD may severely complicate the clinical course. Boundary violations and abuse by a prior treating therapist may prolong the treatment course, because the treatment setting now feels unsafe and precipitates posttraumatic reactivity. Severe personality disorders with overinvestment in multiplicity as a way of life; narcissistic investment in having the disorder and/or the elaborate secondary structuring of self-states; engagement in psychotherapy primarily to seek gratification; and a refusal to take responsibility for behavior, behavior change, and/or symptom management are also generally associated with a poorer prognosis. Number of personality states, however, has only a moderate effect on treatment course.

Treatment of Dissociative Identity Disorder

Current treatment approaches to non–possession-form dissociative identity disorder have evolved considerably. By contrast, the treatment of possession-form presentations remains largely unexplored except in anthropologically oriented observation studies. Available data suggest broad similarities between the treatment of both possession-form and non–possession-form presentations, including direct negotiation with the possessing identities, allowing them to "give voice" to their concerns and assisting them to identify and redress their problems. Specialized indigenous practitioners and ritual therapies are generally available and widely utilized in many low-and-middle-income countries, but psychiatric treatment is typically avoided. Indigenous treatments include neutralization of the conflict or stress via the communal rituals involved in exorcism and/or other longer-term activities, such as integration into a healing community with a specialized function. The evolution of the identities over time, and their eventual change to a more adaptive configuration or unification into a subjectively singular self, constitutes one major basis of improvement that is common to both possession-form and non–possession-form presentations.

The rest of the section on Treatment addresses the care of non–possession-form cases. An in-depth discussion of these current treatment approaches can be found in the *Guidelines for the Treatment of Dissociative Identity Disorder in Adults 3rd Revision* (2011) from the International Society for the Study of Trauma and Dissociation (http://www.isst-d.org). In addition, in 2004, the ISSTD promulgated *Guidelines for the Evaluation and Treatment of Dissociative Symptoms in Children and Adolescents* that can also be found on the ISSTD website.

The current model of treatment of the dissociative identity disorder patient follows a phasic model that is the current standard of care for complex posttraumatic disorders. The phases include (1) development of safety and symptom stabilization; (2) (optionally) focused, in-depth attention to traumatic material; and (3) integration or reintegration, in which the individual moves away from a life adaptation based on chronic traumatization and victimization. These phases are relatively heuristic, and aspects of each may be part of the others.

Treatment Outcome Studies in Complex PTSD and Dissociative PTSD.

Convergent data from treatment outcome studies of patients with complex PTSD, the dissociative subtype of PTSD, and dissociative identity disorder have shown lack of response, or even clinical deterioration, if standard, unmodified Progressive Exposure/Cognitive Behavioral Treatment (PE/CBT) models for PTSD are used to treat these populations. Studies have shown better outcome in individuals with complex and dissociative PTSD from treatment paradigms that give attention to stabilization of

dissociation, emotional dysregulation, and safety issues before, or even instead of, intensive processing of trauma memories. In some studies, highly modified exposure protocols—to "fractionate" the intensity of memories—were required for trauma memory processing in individuals with multiple early life traumas to prevent flooding by trauma memories.

Other studies report significantly higher-than-expected discontinuation rates and lower rates of abstinence after treatment for individuals with dissociative disorders who received standard substance abuse treatment. They also show lack of response to standard treatment for anxiety disorders and higher rates of nonretention from group therapy in sexual abuse survivors. In addition, elevated dissociative symptoms in patients with borderline personality disorder predict limited response to standard dialectical behavior therapy (DBT). Some experts warn that dissociative patients with borderline personality disorder may become overwhelmed by intrusive symptoms if treatment focuses on recall of traumatic memories.

A Norwegian study evaluated the response of dissociation and other symptoms among consecutive admissions to an inpatient program that provided stabilization treatment based on a phasic trauma treatment model for complex trauma survivors. Treatment primarily consisted of group and individual therapy, with adjunctive pharmacotherapy. Patients who met DSM-IV-TR diagnostic criteria for dissociative identity disorder or dissociative disorder NOS (Example 1) only substantially improved when dissociative symptoms were directly addressed in treatment. None of their patients had previously been assessed or treated for a dissociative disorder, and, at the time of this study, the program did not specifically target dissociative disorders in childhood trauma survivors.

Treatment Outcome Studies of Dissociative Identity Disorder

Single-case descriptions of successful treatments for dissociative identity disorder date back more than a century. Systematic outcome studies of non–possession-form cases, however, have only appeared recently. Reviews of clinical case series generally agree that patients with dissociative identity disorder improve with the phasic trauma treatment model. A meta-analysis of eight treatment outcome studies for dissociative identity disorder, including inpatient and outpatient treatments—with treatment provided by both nonexpert and expert clinicians—showed that the phasic model of dissociative identity disorder treatment was associated with improvement across a range of symptoms and comorbidities. These included fewer DSM-IV-TR diagnoses of comorbid axis I and II disorders; decreased dissociation, depression, anxiety, suicidality, and substance abuse symptoms; and reduced general distress. In studies from an inpatient specialty trauma disorder program, gains persisted at 2-year follow-up. Limitations of these studies include data from only a single clinician in some studies; small samples sizes, nonrandom inclusion of

patients, high discontinuation rates, failure to account for nonspecific effects such as regression to the mean, and lack of control groups.

Cost Efficacy Studies of Dissociative Identity Disorder Treatment. Dissociative identity disorder patients have been shown to engender substantial costs to the mental health system in high-income countries, particularly prior to obtaining correct diagnosis and treatment. This is often due to treatment failure for other diagnoses, with frequent utilization of more intensive levels of care—inpatient, residential, and day hospital settings, as well as intensive outpatient treatment.

In a single case study, the annual cost of mental health treatment for a British woman with previously undiagnosed dissociative identity disorder decreased from £29,492 (approximately $47,000) to £10,695 (~$17,000) following correct diagnosis and provision of appropriate phasic trauma treatment, indicating an annual savings of £18,797 (~$30,000), consistent with decreased utilization of more restrictive levels of care.

Data from 55,000 Massachusetts Medicaid patients from 1993 to 1996 showed that inpatient treatment for dissociative identity disorder cost over $3 million during the study period, 33.5 percent of the total Medicaid inpatient costs. This was more than the cost of treatment for patients with panic disorder, major depressive disorder, and bipolar disorder, respectively, although patients with dissociative identity disorder made up only 2.6 percent of the sample. Wives of active-duty military personnel with a dissociative disorder had more outpatient visits than wives seeking treatment for any other psychiatric disorder, although they represented less than 1 percent of the sample.

Two Canadian cost efficacy studies totaling 64 patients with dissociative identity disorder showed substantial reductions in costs for psychiatric treatment once a correct diagnosis was made and appropriate treatment was instituted. Patients averaged 3.9 incorrect diagnoses and 7 to 12 years in treatment before correct diagnosis. Longer time in the mental health system predicted longer delays in reduction of treatment cost, although even highly ill dissociative identity disorder patients with multiple comorbidities showed significant cost savings with correct treatment, although, cost reductions took much longer than for higher functioning patients. Less impaired, less chronically ill subgroups of patients with dissociative identity disorder showed rapid, often dramatic reductions in cost after correct diagnosis. In one study, costs were halved for the less chronically ill subsamples from an average of C$75,000 per year per patient prior to a dissociative identity disorder diagnosis, to an average of C$36,000 per patient per year in the 3 years after correct diagnosis. Notably, in out-years 2 and 3 costs were reduced to an average of C$10,600 per patient per year. Both studies found that average costs remained high for the first year after correct diagnosis. Extrapolating cost savings by

estimating treatment costs if the patients had continued in incorrect treatment for another decade, savings ranged from C$1.35 to C$3.75 million (1989 Canadian dollars).

Longitudinal Treatment Outcome Studies

Bethany Brand and colleagues developed the *Treatment of Patients with Dissociative Disorders Study* (TOP DD Study; http://www.towson.edu/topddstudy/index.asp), an international, prospective, longitudinal outcome study of 280 patients diagnosed with DSM-IV-TR dissociative identity disorder or Dissociative Disorder NOS Example 1, treated by 292 therapists from 19 countries. Treatment progress was assessed prospectively at study entry, and at 6, 18, and 30 months of treatment. Outcome measures were assessed using a five-stage model, with two additional subdivisions in the tri-phasic model. Anchors for stages were stages 1, 3, and 5 corresponding to the three stages in the clinical Phasic Trauma Treatment model of establishing stability and safety (stage 1); processing of trauma memories (stage 2); and integration and reconnection within self and with others (stage 3).

At each time point, therapists completed a data sheet on the patient's stage of treatment (1–5), Global Assessment of Functioning (GAF), and incidence of self-harm, suicide attempts, and hospitalizations. Also, at each time point, clinicians completed the Progress in Treatment Questionnaire (PITQ) that assesses management of impulse control; emotion regulation; capacity to use skills for symptom management; decreases in posttraumatic cognitive distortions and misperceptions; increased internal communication and cooperation between self-states; capacity to work on, and full awareness of, posttraumatic memories; integration of self-states; and increasing capacity to see oneself and others in a realistic, integrated way.

Patient measures were the DES, the Posttraumatic Check List-Civilian version (PCL-C), and the Symptom Checklist-90 (SCL-90-R). Patients also completed a behavioral checklist to assess the frequency over the last 30 days of self-injurious behavior; suicide attempts, "very impulsive" behaviors, alcohol and drug consumption, and high-risk behaviors that could have resulted in death. Also, patients reported the frequency of positive experiences and adaptive behaviors in the last 30 days including: volunteering or attending school, working for pay, using symptom management techniques, socializing, and feeling good "even if for a brief period." Patients also reported the days in inpatient and day treatment programs in the month prior to assessment.

Results from the 30-month follow-up are shown in Table 20–19.

The initial cross-sectional TOP DD results showed that patients in treatment stage 1 had higher levels of dissociation and PTSD, more hospitalizations, higher overall distress, and less adaptive functioning than

patients in stage 5. As reported by both patients and therapists, at the 30-month follow-up, there was decreased dissociation, PTSD, general distress, depression, suicide attempts, self-harm, dangerous behaviors, drug use, physical pain, and hospitalizations, as well as improved functioning, as indicated by higher GAF scores. At the 30-month assessment, more patients were involved in positive activities in volunteer jobs and/or attending school and socializing, and reported feeling well. Consistent with the phasic treatment model, more patients progressed from early stages of treatment to more advanced stages than those that moved in the opposite direction.

Table 20–19.
Findings at 30 Months from the Treatment of Patients with Dissociative Disorders Study (TOPDD Study)

Completion of Treatment, Drop out, and Termination Rates			
Variable	**Patients**	**Therapists**	**# Countries**
Baseline	226	295	19 North America, Latin America, Europe, UK, Australia, New Zealand, China, Singapore, Israel, Slovakia, South Africa
N at 30-month follow-up	119	135	
% Drop out of study (of those who were still in treatment)	36%	48%	–
Successful treatment completers	3% of baseline sample (*n* = 6)	–	–
Terminate treatment for any reason	18% of baseline sample (*n* = 40)	–	–

Summary of significant treatment outcomes (Note: All variables changed in the hypothesized direction, so all significant linear changes indicate improvement over time. If the curvilinear term was significant, it means that the change in that variable flattened out over time; i.e., improvement slowed over time.)

Continuous Variables	**Direction of Change**	**Significant Improvement over Time (Linear Change Significant)**	**Change Flattens over Time (Curvilinear Variable was Significant)**
Patient Reported			
Dissociation	Decrease	X	
PTSD	Decrease	X	X
Distress overall	Decrease	X	X
Depression	Decrease	X	X
Therapist Reported			

GAF	Increase	X	
Adaptive Functioning (PITQ)	Increase	X	X

One Item Dichotomous Variables	Direction of Change	Significant Improvement Over Time (Linear Change Significant)	Change Flattens over Time (Curvilinear Variable was Significant)
Patient Reported: Symptoms			
Self-injurious behavior	Decrease	X	
Dangerous behavior	Decrease	X	
Getting high	Decrease	X	X
Impulsive behavior	Decrease	X	
Inpatient hospitalizations	Decrease		X (hospitalizations went up in the second follow-ups, but decreased by third and fourth follow-ups)
Patient Reported: Adaptive Functioning			
Volunteering/Attending School	Increase	X	
Socializing	Increase	X	
Feeling good	Increase	X	
Therapist Reported: Symptoms			
Self-injurious behavior	Decrease	X	
Inpatient hospitalizations	Decrease	X	

Not significant:

- ► Only single-item variables
- ► Patient reported: suicide attempts, drinking alcohol, using symptom management skills
- ► Therapist reported: day program attendance

Based on Brand BL, McNary SW, Classen CC, et al. The Results of a Longitudinal Naturalistic Study of Treatment Outcome for Patients with Dissociative Disorders (TOP DD Study). Adapted from a paper presented as part of a symposium at the annual convention of the International Society for Traumatic Stress Studies, Montreal, Canada.

The TOP DD patients with the highest levels of dissociation, and with the most severe depression, also showed significant improvements in these symptoms over 30 months. More patients showed “sudden improvement” than “sudden worsening” (i.e., 20 percent increase or decrease in symptoms) at one or more time point(s). Rates of revictimization showed a decreasing trend during the study. Therapists reported less frequent revictimization events and stressors among the patients who improved suddenly relative to those who worsened. Consistent with clinical experience, revictimization and/or severe external stressors were most likely to contribute to worsening in treatment. Only 1.1 percent of patients showed worsening over more than one data collection point, less than overall rates of worsening in treatment outcome studies of general psychiatric patients. As noted above, there were two completed suicides among patients in this study.

Younger patients, despite having elevated symptoms at baseline, were more likely to show improvement, indicating that early treatment of severely dissociative patients may lead to a more positive life trajectory. Patients reported a significant decrease in feeling like "different people" and hearing voices. A trend decrease in reports of dissociative amnesia supports the view that the phasic trauma treatment model helps to improve core dissociative process symptoms, not to increase them as critics of this model have suggested.

In summary, the TOP DD study documented that the phasic trauma treatment model improves a wide range of symptoms and adaptive functioning indices in patients with dissociative identity disorder, and decreases utilization of more restrictive levels of care. Limitations of this study included the lack of a randomized control group, of fidelity compliance assessments for therapists, and of a specified psychotherapy protocol. However, randomized controlled trials with manualized treatment protocols may not be ethical, or clinically feasible, in the treatment of complex patients, such as those with borderline personality disorder and dissociative identity disorder, with multiple severe symptom configurations, and with frequent severe self-destructive and suicidal urges and behaviors.

Expert Consensus Guidelines for Phasic Treatment of Dissociative Identity Disorder

In addition to the TOP DD study, 36 international dissociative identity disorder experts were surveyed to identify treatment phases and evidence-based interventions for the treatment of this disorder. The experts rated the utility of 28 interventions in the treatment of dissociative identity disorder on a 0 (never) to 4 (very often) scale, across the stages of the TOPP DD treatment construct. Across the stages of treatment, the 36 experts found a core set of interventions helpful, as well as particular interventions that are beneficial at various stages. The most frequently recommended interventions for each treatment stage are shown in Table 20–20.

Even in the stage where work on trauma memories is most predominant, the expert treatment recommendations still focus on safety, stabilization, the therapeutic alliance, cognitive work, and symptom management. In the final phase, the therapeutic work is directed toward "reintegration," living well in the present, and relegating traumatic memories to the status of "bad memories"—as opposed to flashbacks, behavioral re-experiencing, or intense posttraumatic reactivity to current situations. These stages are heuristic, since trauma memory may need to be addressed—if only in a more cognitive and distanced manner—in stage 1 and worked through again from a more integrated perspective in stage 3. Safety may be an issue at all stages of trauma treatment. Finally, the entirety of trauma treatment is directed toward helping patients adapt

better to their current circumstances.

Table 20–20.
Expert Recommendations for Treatment of Dissociative Identity Disorder —Stages 1–5

Stage	1	2	3	4	5
Establish safety	✓	✓	✓		
Establish/repair alliance	✓	✓	✓	✓	✓
Teaching/practicing grounding	✓	✓	✓		
Educate about disorders and treatment options	✓	✓			
Diagnose/rule out comorbid psychiatric illnesses	✓				
Teaching/practicing self-care	✓	✓		✓	✓
Developing healthy relationships	✓			✓	
Affect tolerance and impulse control	✓	✓	✓	✓	
Stabilizing from current day stressors/crises	✓	✓	✓	✓	✓
Teaching/practicing containment	✓	✓	✓		
Developing healthy relationships		✓			✓
Cooperation with self-states		✓	✓	✓	
CBT focus on cognition distortions			✓		
Identify/work with self-states			✓		
Awareness of body sensation			✓	✓	✓
Discussing therapeutic relationship				✓	✓
Processing patient's reactions to therapy				✓	✓
Ego strengthening activities				✓	✓
Awareness of emotion					✓
Stabilizing following intrusions from reported perpetrators					✓

Note: Phasic Trauma Treatment Model has three Stages, corresponding to TOPDD stages 1, 3, 5—see text for discussion.

Adapted from Figures 1–5 in Brand BL, Myrick AC, Loewenstein RJ, Classen CC, et al. A survey of practices and recommended treatment interventions among expert therapists treating patients with dissociative identity disorder and dissociative disorder not otherwise specified. *Psychol Res Pract Policy.* 2012;4(5):490–500.

The experts recommended interventions around safety most frequently in all but the last stage of treatment. Safety is a consistent focus because each stage can bring about emotional crises that patients may feel compelled to manage via destructive action toward themselves, or less commonly, others. Developing and repairing the therapeutic alliance was recommended as a top intervention in every stage of treatment, indicating the centrality of the therapeutic alliance in successful treatment of dissociative identity disorder. Recent research suggests that the therapeutic alliance is important in predicting improvement across a range of symptoms as it is in most successful psychotherapies. This may be because most individuals with dissociative identity disorder have suffered betrayal and maltreatment in the context of early dependent relationships. Thus,

they may seek relational connection but also fear betrayal by the persons on whom they are dependent, manifesting covert or overt mistrust of the therapist (see section on Betrayal Trauma).

Other core interventions recommended across all stages of treatment include differential diagnosis and treatment of comorbid psychiatric disorders to differentiate symptoms that are part of a complex posttraumatic dissociative syndrome from genuine comorbid disorders; discussion of pharmacotherapy and psychoeducation about the role of medications in dissociative identity disorder treatment; help becoming more aware of emotions and developing affect tolerance and impulse control; management of daily functioning and current relationships; management of transference particularly traumatic transference and other reactions to therapy; and help coping better with stressful life situations, losses, and additional trauma. Thus, the treatment recommendations in this chapter are based on these expert recommendations, as well as the ISSTD Treatment Guidelines, and the interventions documented in the TOP DD study to lead to improvement in a variety of dimensions.

Stage 1 Phasic Trauma Treatment: Stabilization and Safety

Stabilization of the individual with dissociative identity disorder is vital to successful negotiation of all aspects of treatment. Stabilization focuses on safety and management of core dissociative identity disorder and comorbid symptoms. The vast majority of patients with dissociative identity disorder come to treatment because of overwhelming PTSD and dissociative symptoms and because of some form of self-destructive behavior or violence against others, as shown in Table 20–12.

Before establishing safety, the patient actually or symbolically lives in a world of ongoing trauma, whether perpetrated on the patient by others, and/or by the patient on him/herself due to chronic forms of repeated self-destruction and high-risk behavior. The clinician must logically prioritize basic life and health over other interventions. As previously described, the person with dissociative identity disorder may be affected a wide variety of high-risk behaviors, symptoms, and comorbid disorders. The clinician may need to resort to many different levels and types of interventions to protect the patient including hospitalization, specialized substance use or eating disorders programs, police assistance and shelter for victims of intimate partner violence, social service interventions to protect minor children, and assistance with housing, and access to medical care.

Practical Management of Safety Issues

Most patients with dissociative identity disorder use unsafe and high-risk behaviors as attempts to regulate, manage, and modulate severe PTSD, dissociative, affective, somatic, and cognitive dysregulation. Work on safety generally includes discussion of the antecedents to, and functions of, self-destructive, suicidal, and aggressive behavior toward others, as well as developing safety agreements and crisis management plans. It is critical not to view safety interventions as having the patient "stop bad behavior." There is usually logic behind the dissociative identity disorder patient's self-destructive behavior. For example, there is a highly prevalent cognitive distortion in this population that, "I'm going to get hurt no matter what I do, so control of the timing and intensity of harm, is all the control I get." These beliefs can be reframed as an attempt to achieve a sense of control to survive the helplessness and unpredictability of repeated maltreatment throughout development. The therapeutic goal then becomes to change this "survival skill" to a "recovery skill": developing adult, nontrauma-based strategies for appraising one's current life situation and level of external danger, gaining a repertoire of skills to keep oneself safe in the present, and creating overall safety and self-protection in one's current life.

Usually, in this patient group, self-destructive behavior has an addictive component, and many of these patients have current or past problems with substance misuse. Using addiction metaphors can be helpful: "sobriety" from dangerousness to self involves a commitment to safety, the realization that symptoms likely will *worsen* once the behavior stops, and intensive work to find alternatives to help manage intolerable, often unnamable, affective states such as extreme shame, humiliation, rage, horror, helplessness, and disgust/revulsion. At times, self-injury is experienced as an attempt to "feel something" when in a numb state, and/or to "put the pain on the outside" where it seems more real, or where (the patient believes) others may take it more seriously.

Safety Agreements. Safety agreements may be a useful clinical tool to assist the individual with dissociative identity disorder manage safety issues and to support the patient's ambivalence about being unsafe. Safety agreements are no substitute for the clinician's judgment about the patient's safety from dangerousness to self or others. The clinician must continually assess the patient's safety and, based on clinical judgment, act to protect the patient or others, including initiating certification, despite the patient's protestations that he/she will maintain the agreement.

It is useful to think of safety agreements as *delaying* agreements, where the patient has a hierarchy of interventions to use instead of self-destruction, up to and including going to the Emergency Department for stabilization. Recovering addicts/alcoholics are taught to think in "one-day-

at-a-time" approaches to sobriety. In trauma recovery, patients with dissociative identity disorder may be unable to commit to safety "forever," but may be able to commit to a safety plan for a more manageable period (e.g., from therapy session to therapy session). Some clinicians prefer that the patient write down and sign this type of agreement. The authors' experience is that the crucial factor is the patient's honesty about safety and the ability to maintain him/herself as a safe outpatient, following the safety plan. Many signed safety agreements are not honest, and many verbal commitments are. Some patients feel that writing and signing an agreement reinforces the seriousness of and concretizes the commitment to it. It is not necessary for "all" personality states to sign the safety agreement. Instead, the clinician can ask all states, "known or unknown," if there is any objection to the agreement and to indicate so by a physical signal, such as a head nod, or verbally. If there is an objection, the clinician can ask for some sort of communication in the patient's mind about the concern. Often, the issues that emerge are quite helpful to the patient in the process of working honestly on safety. It may help the patient to realize that the clinician has to trust the patient's honesty regarding safety agreements, and that the patient is not the only one in the therapy who has concerns about trust. If a signed agreement is created, the patient may sign several names and add that "and all [other] known or unknown" self-states agree as well. Occasionally, the patient with dissociative identity disorder will have an idiosyncratic posttraumatic reaction to the words "safety" or "agreement." In this case, the patient may be asked to provide a reasonable alternative, such as the "no-immediate-action pact."

EDUCATION AND INFORMED CONSENT. It is important to educate the individual with dissociative identity disorder about the disorder, its actual and apparent comorbidities, and the course of treatment. Educative interventions help decrease anxiety about symptoms that are often frightening and overwhelming, build a therapeutic alliance, and provide information that is the basis for a meaningful consent for treatment. Furthermore, it is necessary to educate the patient about the contentious and divisive debates that surround the diagnosis and treatment of dissociative identity disorder in contemporary psychiatry and psychology.

Other major issues in informed consent include advising the patient of potential risks and benefits of the treatment and the current evidence base for the treatment model. In addition, patients should be informed about approaches that differ from the phasic trauma treatment model, such as the iatrogenic/sociocognitive/FM that advocates ignoring the self-states and the history of trauma, and the evidence base and potential risks and benefits of these alternatives.

The patient should also be informed about current controversies concerning traumatic memory and its retrieval and explication in therapy.

Additional informed consent should be obtained for formal hypnosis and pharmacotherapy. The patient should be counseled that symptomatic worsening may occur temporarily during treatment, particularly during phases in which memory material is worked with in depth. Simultaneously, however, subjective amelioration of other symptoms may occur, with the patient achieving a better sense of control, mastery, and self-coherence as he/she successfully works through painful material.

Similarly, like many obsessional individuals, patients with dissociative identity disorder often respond positively to learning that their dissociative and PTSD symptoms have a name and an organizing framework. There is often relief at a diagnosis that fits the patient's subjective life experience, especially if there has been a long prior history of unsuccessful psychiatric treatment for other disorders. At the same time, patients may respond with distress at the dissociative identity disorder diagnosis due to media and popular stereotypes, because it brings into sharper relief the meaning of painful symptoms, due to shame reactions such as feeling like a freak, and/or because it brings more focus to the distressing history of early trauma and problematic relationships with family members.

Individuals with dissociative identity disorder often need to be counseled that the purpose of treatment is not to suppress or destroy identities, but rather to allow therapeutic engagement with them as vital aspects of the mind. Self-states may express strong wishes to "get rid" of other identities and are simultaneously terrified that the therapist will attempt to do so. Most experienced clinicians educate patients that treatment is geared toward optimal function and adaptation and that there are usually too many variables at the beginning of treatment to predict whether it will necessarily ultimately result in unification of all self-divisions, a concept that the patient usually cannot really comprehend at the beginning of treatment. At the same time, most experienced clinicians see unification of all self-states as usually providing the best long-term outcome in dissociative identity disorder.

The education and informed consent process should be ongoing throughout treatment. Many of these issues need to be revisited as treatment proceeds and the patient changes in therapy. As with the safety agreement, if a signed informed consent is used, it is not necessary for "all" self-states to sign the agreement, and self-states that are experienced as newly emerging later in treatment may report amnesia for prior psychoeducation. An overly rigid concern about how many states sign the agreement may actually reinforce subjective separateness by treating the identities more as separate people than subjectively structured personified mental processes.

BOUNDARIES AND TREATMENT FRAME. Most patients with dissociative identity disorder report that their traumatic life experiences occur over long periods

of time in the context of close personal relationships with those who have hurt them. These patients often report boundary violations in subsequent relationships with teachers, therapists, medical professionals, and social service workers, among others. For these patients, the role description of significant others does not necessarily predict their behavior, and relationships are often seen as "up for grabs." The patient expects and fears that the clinician will violate boundaries. There may be a subtle (or not so subtle) pull in the transference for the therapist to change the frame of treatment. Patients may experience this as trying to make the inevitable happen, so that they do not have to endure the agony of waiting for things to go wrong. Accordingly, clinicians working with these patients must pay careful attention to issues of the treatment frame and appropriate and consistent boundaries. These include firm limits on the length of therapy sessions, acceptable behavior during sessions, the extent of extra-session contacts, and payment of fees, among many others. Calm, educative, nonblaming interpretation of the patient's fears over the perceived inevitability of boundary violations is preferable to enacting them.

The ISSTD Guidelines, the expert survey, and the TOP DD study all converge to discourage boundary-crossing activities, even if the patient insists that only these interventions will help, and/or particularly help specific self-states, such as child states that are portrayed as needing concrete demonstration of love from someone. Particularly, the therapist may be asked to hold or hug the patient, hold the patient's hand, or other types of physical contact. Rather than acting on this, the clinician should inquire if some personality states are distressed at the idea of this physical contact, as "nice" touching may have evolved into traumatic abuse in the patient's reported past history. Then, the issue is a therapeutic discussion of the meaning of the request, and the patient's profound conflict about it.

It is not recommended that the clinician accept more than a token gift such as a card or a small piece of artwork; give the patient gifts; routinely contact the patient while on vacation; engage in out-of-office contact with the patient, such as going for walks and/or going out for coffee; and screen the patient's voice mailbox for calls from reported perpetrators, among many other potentially boundary-crossing behaviors.

Clinicians who feel overwhelmed by the treatment of these patients are more likely to accede to the patient's insistence on these sorts of actions. It is important to recognize that the patient's reported history of abuse often involved being special to the abuser in some way. Thus, these boundary transgressions recreate this double-edged situation for the patient. Once again, education of the patient about boundaries is the most helpful clinical strategy.

Many of the clinical and conceptual issues that help make treatment workable with these patients are not taught in standard training programs for mental health professionals. The ISSTD has a series of courses on

diagnosis and treatment of dissociative disorders that can be accessed at www.isst-d.org. Consultation with experienced clinicians should be considered by clinicians who are new to working with this patient population, and/or if stalemates or seemingly unsolvable predicaments develop in treatment.

Direct Therapeutic Engagement with Personality States. In order to achieve safety, stability, and overall treatment progress, consensus in the field and the evidence-based data strongly support the view that effective stabilization in most patients with dissociative identity disorder requires psychotherapeutic work with individual personality states. Most patients are unable to stabilize symptoms in the long term if the identities that control these symptoms are not therapeutically engaged. Clinicians new to dissociative identity disorder are frequently uncomfortable or perplexed by the need to work directly with self-states and how to do this without producing a chaotic regression or reinforcing the patient's sense of separateness among self-states.

Certain basic principles are important to understand. No personality state is any more or less real than any other or better or worse than another. All are aspects of a single human being and have adaptive, psychological importance that needs to be heard and respected. *The personality of the patient with dissociative identity disorder is all of the personality states*. The clinician should not engage in interactions that reinforce notions of self-states as separate people. These can include "playing favorites," supporting the notion that some self-states are fundamentally "good" and others "bad," refusing to communicate with certain self-states, and giving separate appointments to self-states as one would with individual patients. However, this does not mean that self-destructive, dangerous, suicidal, violent, high-risk behaviors are condoned. Rather, to help stop these behaviors and develop the person's repertoire of recovery skills, it is important to understand the self-regulatory functions of these behaviors and the trauma-based cognitive distortions that drive them.

It is important that all self-states are held accountable and responsible for the behavior of any other state, even if experienced with amnesia or lack of subjective ownership. Human beings engage in behavior, not personality states. One can attempt to understand the mental state of the person who does the behavior, but explanations are not excuses. Expert consensus and clinical experience has shown that to do otherwise is an invitation for the patient to regress. Because, as described above, patients with dissociative identity disorder have both psychological strengths and liabilities, they do better when held to a standard that underscores their strengths. In the context of psychotherapy, self-states can be understood as developmentally concretized forms of organizing, symbolizing, and mobilizing mental

contents. These subjective entities have many problematic conflicts, cognitions and affects, as well as adaptive attributes. As such, the fullest participation of all mental symbols and sources of knowledge, skills, and adaptations is likely to produce the best outcome in any psychotherapy. This also forms the potential basis for forensic standards around competence and not guilty by reason of insanity (NGRI). This uses both the psychological structure findings of preservation of logical thinking and reality testing, and the notion that dissociative identity disorder patients can inhibit and conform behavior in many contexts, unless overwhelmed by posttraumatic flashbacks.

Concerns about "iatrogenesis" are often raised in this circumstance. Understanding the personality states fundamentally as forms of intrapsychic symbolization reduces this concern to a more routine psychotherapeutic task of understanding the meaning of mental contents, rather than a suppression of some mental states as less authentic than others. At the same time, the clinician must also be able to appreciate the patient's subjective reality in which the self-states are experienced as distinct selves or even as separate persons. The therapist must be able to tolerate engaging with the self-states *as if* they were separate entities, while simultaneously maintaining the understanding that the identities represent the structuring of psychological processes in personified form within a single human mind.

The continuity, stability, and respectful impartiality of the therapist toward the different self-states provide an important therapeutic experience that helps the patient experience, examine, and integrate these dissociated aspects of self. The processing of negative life events, often from the multiple perspectives of different personality states, is an important part of the narrative reorganization of the patient's divided identity into a more coherent whole. The therapist usually is actively involved in negotiating between and among the various self-states to achieve therapeutic goals. These can include ongoing negotiating agreements to stop self-destructive behaviors, facilitating more adaptive inter-identity function and communication, and allowing certain self-states to participate more fully, but safely, in therapy.

INTERVENTIONS FOR SAFETY AND STABILITY: CONTAINMENT, GROUNDING, AND IMAGERY TECHNIQUES. The expert consensus recommends that use and teaching of grounding techniques begin in stage 1 and continue through subsequent stages. Grounding may include a variety of interventions that begin in therapy, but are practiced independently by the patient outside of sessions. Grounding may include having the patient change positions, open his/her eyes, use his/her five senses to become more aware of current circumstances: touch an object, such as the arms of the chair, feet on the floor, and look around the room. If the patient enters a severe dissociative

flashback, the therapist may literally have to orient the patient to the date, place, and circumstances, including the name and profession of the clinician. Grounding can also assist the patient with the confusion, fear, and dysphoria that occurs with overlap and interference among self-states, as well as helping the patient maintain state stability to prevent dysfunctional switching.

In all but the last stage of treatment, expert consensus suggests that the patient develop containment skills for traumatic material. These interventions help the patient use dissociative defenses more adaptively. Patients may oscillate from all "shut down" to overwhelming intrusion of trauma material. The patient can be taught that containment is different than shutting down. It involves a voluntary psychological sequestration of trauma memories or overwhelmed self-states until these can be worked with in a gradual way at a later time.

With the onset of treatment, most expert therapists rapidly introduce grounding and containment skills for symptom management. This necessitates active structuring of sessions to work on containment of severe symptoms and skills-building. Containment techniques include teaching self-hypnosis (only if the clinician is credentialed in the use of hypnosis) and imagery to control the intrusiveness of traumatic material (see discussion on hypnosis). These techniques use patients' high dissociative and hypnotic capacity to assist their dissociative defenses to become more flexible and adaptive. This may involve imagery techniques to attenuate the intensity of PTSD, somatoform, and dissociative symptoms; successful stabilization of identities that embody or are experienced as creating symptoms; and encouragement of collaboration, communication, and empathy among identities. In particular, experienced clinicians work to attenuate the impact of PTSD intrusions, rather than opening up this material prematurely. Imagery interventions also may include ego strengthening (i.e., interventions to promote better overall functioning); calming imagery such as safe-place imagery; imagery for sequestration and/or attenuation of difficult material such as flashbacks; reaffirming statements; relaxation training; and cognitive-behavioral work to change trauma-based cognitions.

Given the consistency of strategies recommended by experts, the authors of the expert survey concluded that these interventions make up the core treatment processes and structure for treating dissociative identity disorder. The consistency of the recommendations in the expert survey and the ISSTD Treatment Guidelines, as well as results of the TOP DD study, indicates that there is now an evidence-based standard of care for the treatment of dissociative identity disorder.

RESPONSIBILITY. The clinician should be clear that the ultimate responsibility for symptom and behavioral management outside of sessions

lies with the patient, who is viewed as an active partner in the treatment. In general, if not immediately, patients with dissociative identity disorder respond well to taking more active responsibility for safety or behavioral control between sessions. The therapist may prescribe homework for the patient to work on symptom management and internal communication between sessions. These strategies are usually essential in countering the patient's tendency toward regression and demoralization. Patients often use journal writing as a helpful adjunctive treatment strategy that allows more controlled expression and communication among identities, fulfillment of homework assignments, and a more attenuated and distanced way of handling overwhelming subjective experiences.

TRAUMA-FOCUSED COGNITIVE-BEHAVIORAL INTERVENTIONS. Individuals with dissociative identity disorder may experience many cognitive errors and distortions based on traumatic life experiences and the ability of dissociation to interfere with reality testing. These include basic misattribution of causality—"I caused my father to abuse me"; de-catastrophizing (minimizing violent physical discipline as appropriate), and catastrophizing (a minor disappointment becomes a matter of survival); trauma-based doublethink (e.g., "sex is love," "anger is violence," "self-assertion is mean," "hurting myself makes me safe," and many others).

Other cognitive distortions include the insistence on delusional separateness; that the patient is helpless to control him- or herself and requires the clinician to manage all difficulties; that the clinician is completely untrustworthy and must not be allowed any access to the patient's mind; since love and sex are the same, and because trauma and abuse are inevitable, it is best to invite them or enact them oneself, so that at least their timing and intensity can be better controlled.

Behind these cognitive distortions frequently lie exceptionally painful realizations and recollections. For example, if the patient gives up the idea of blameworthiness for childhood incestuous experiences, the reality must be faced that a beloved relative has done the patient grievous harm and has treated the patient as a hated object, not a person. Accordingly, many cognitive distortions are only slowly responsive to cognitive therapy techniques, and successful cognitive interventions may lead to additional dysphoria.

TRANSFERENCE AND COUNTERTRANSFERENCE. Patients with dissociative identity disorder often manifest a complex multilayered transference as a whole, as a system of personality states, and as individual personality states. Commonly, transference is dominated by themes of trauma and abuse, with the therapist most commonly experienced as potentially exploitative and abusive or uninvolved or uncaring about the patient's difficulties or as a helpless victim, like the patient faced with an implacable other.

Dissociative identity disorder self-states in a *flashback transference* may literally envision the clinician as the embodiment of an abusive figure from the past, requiring active interventions to help the patient separate the past from the present.

Countertransference responses may vary as well, with over-involvement, detached hostile skepticism, or a sense of being exasperated, overwhelmed, and de-skilled being quite common. Clinicians may experience countertransferential autohypnotic and dissociative phenomena during dissociative identity disorder treatment that need to be recognized and managed. Burnout and secondary PTSD have been reported for therapists and treatment teams working with dissociative identity disorder patients, unless adequate attention is paid to limit-setting, boundaries, and assistance to clinicians adversely experiencing reports of extreme abusive trauma or being overwhelmed by work with patients with dissociative identity disorder.

Affect Regulation/Shame. As described above, innate affect theory is a helpful framework for viewing the experience of the individual with dissociative identity disorder, as well as the structure of the inner world of the identities. The four components or a "compass of shame" include "attack self," "attack other," "avoidance," and "withdrawal." Shame affects and "shame scripts" may dominate the psychic life of patients with dissociative identity disorder. As noted above, Richard Kluft has developed a typology of dissociative identity disorder self-states that links their characteristics to one or more of the poles of shame. Kluft suggests that avoidance of difficult material in dissociative identity disorder is frequently related to hidden shame scripts. In many instances, the dissociative patient is not amnestic, but profoundly inhibited by shame avoidance from discussing important material. Kluft posits that systematic attention to shame allows for progress in negotiating many difficulties and potential impasses in the treatment of dissociative identity disorder.

Stage 2 Phasic Trauma Treatment: Work on Traumatic Memories. For patients who stabilize and form a reasonable working alliance in treatment, longer-term treatment goals involve the detailed, affectively intense, psychotherapeutic processing of life experiences, especially traumatic experiences, and the transformation of the meaning of these experiences for the individual. Authorities emphasize that, in most cases, intensive, detailed psychotherapeutic work with traumatic memories should only be initiated after the patient has demonstrated the ability to use symptom management skills independently, after the self-state system can work together in a reasonably cooperative way, and after a solid therapeutic relationship has been established.

The patient should be able to give informed consent and have a realistic

understanding of the potential risks and benefits of intensive focus on traumatic material. Potential risks may include temporary, acute worsening of PTSD, affective, somatoform, and self-destructive symptoms, and short-term interference with daily activities. Long-term benefits may include significant amelioration of dissociative and PTSD symptoms, decreases in subjective self-division, fusion of identities, and freeing of psychological energy for daily life. The patient must be able to understand that the goal is integration of dissociated thoughts, feelings, recollections, and perceptions, not the exhumation or expulsion of memories per se.

Furthermore, the patient should not be in the midst of an acute life crisis or major life change, comorbid medical and psychiatric disorders should be stabilized, he/she must have the ego strength and psychosocial resources to withstand the rigors of the process, and there must be adequate resources, such as support by significant others, and, potentially, support for additional sessions. Many of these are listed in Table 20–21.

Experienced clinicians attempt to carefully structure affectively intense sessions focused on traumatic material, with attention being given to affect modulation, re-stabilization of the patient before concluding the session, and reasonable availability to assist the patient supportively between sessions. In addition, many sessions may be needed to explicate fully the cognitive and emotional meaning of traumatic events, so that they may become part of the patient's repertoire of nondissociated, ordinary memories for life experience.

Table 20–21.
Therapeutic Considerations for Stage 2 Phasic Trauma Treatment in Dissociative Identity Disorder

Patient has achieved the goals of earlier stages, including
Basic acceptance of diagnoses, trauma issues
Ability to maintain safety despite stress
Basic understanding of the importance of safety
Reasonable internal communication, collaboration/noninterference between self-states
Demonstrated ability to independently use symptom management/containment strategies
Mapping of self-state system gives basic view of major self-state groups/types
Basic stabilization of persecutory/angry self-states
Adequate therapeutic alliance
Severe dysregulated PTSD controlled
Uncontrolled, dysfunctional switching absent, infrequent, or reasonably easily contained
Rational motivation for engaging in the process
Reasonably stabilized life circumstances
Presence of interpersonal supports (and how to manage if limited or nonexistent)
Stable comorbidities, both psychiatric and medical
Adequate ego strength to undergo this process without major decompensation
The therapist is prepared in both availability, skills, and ego strength to work with/tolerate the material

Adapted from Kluft RP, Loewenstein RJ. Treatment of Dissociative Disorders. Gabbard GO, ed. *Gabbard's Treatment of Psychiatric Disorders*. 4th ed. Washington, DC; American Psychiatric Press; 2007.

TRAUMATIC MEMORIES. Media attention and legal cases have led to concerns over the authenticity of traumatic recollections of patients with dissociative identity disorder. Recent rigorous case series have supported the experience of clinicians, who have found corollary information to corroborate these patients' recollections or who have even unearthed data about traumas that the patient does not recall. Based on corollary information, however, some individuals with dissociative identity disorder can misinterpret and misrepresent contemporary information or confabulate aspects of their past history. In addition, dissociative patients are hardly immune to ordinary human emotions, such as greed, envy, wishes for revenge, wishes to evade consequences for misbehavior, and desires to placate significant others. All of these may complicate the patient's veracity, especially in situations of potential financial gain, media attention, legal cases, and possible loss of family contacts, among others. Conversely, collateral informants, such as family members, may be subject to the same sources of unreliability as the index patient, and their input should be weighed accordingly.

Individuals with dissociative identity disorder typically oscillate from regarding their recollections as all true to all false. Specific identities may take opposing positions. The clinician is best served by maintaining a stance of neutrality toward the patient's recollections of his or her life. In this regard, it is generally most helpful for the clinician to identify for the patient his or her internal conflicts over the veracity of recollections and to invite an open airing of all points of view. Repeated discussion of the complex factors that may impact autobiographical recall may need to be part of an ongoing informed consent process as treatment progresses.

The therapist should respond respectfully and thoughtfully to all clinical material brought by the patient. Patients can be helped not to come to premature closure about their views of past events. Some patients with dissociative identity disorder come to understand life events quite differently as they become less dissociative over the course of treatment. Consequently, clinicians must avoid validating memories for the patient or dismissing them out of hand, in the absence of reliable collateral information. Therapy is most successful for the patient if the clinician maintains the role of therapist, not personal advocate, detective, or derisive skeptic.

Stage 3 Phasic Trauma Treatment: Fusion, Integration, Resolution, and Recovery. Over the course of treatment, significant unification of dissociated mental processes may be observed. Self-states lose distinctness and decrease compartmentalization of thoughts, memories, and affects. The patient develops a more unified sense of self. Transference is modified consistent with these changes. Amnesia and switching become less apparent. Fusion of states results in psychological merging of two or more

entities at a point in time, with a subjective experience of loss of all separateness. The term *integration* is sometimes used synonymously with *fusion* but is more generally defined as the process of undoing all forms of dissociative division during treatment. Some patients proceed to what appears to be a complete fusion of all states, with a shift in self-representation from that of an individual with dissociative identity disorder to one with a consistent and continuous sense of self across all behavioral states. Many patients never attain a complete fusion of their personalities but achieve a therapeutic *resolution:* improved communication, collaboration, and cooperation among self-states leading to relative stability and adequate function.

As this integrative process occurs, PTSD symptoms usually improve significantly. Patients often experience a freeing of energy toward everyday life and away from trauma-focused ways of living. Cognitive distortions frequently substantially subside. At the same time, nondissociative and integrative coping strategies must be identified and substituted for dissociative responses to life stressors. Losses must be mourned, and the patient must be helped to connect and cope with the larger world in a more functional manner.

Hypnosis. Hypnosis is a set of techniques that facilitate treatment, not a treatment in and of itself. Because of the high hypnotizability and autohypnotic states of the patient with dissociative identity disorder, all treatment of this disorder involves hypnotic phenomena. Clinicians working with this population should obtain training in hypnosis, if only to recognize spontaneous, autohypnotic processes emitted by the patient. Hypnotherapeutic interventions can often alleviate self-destructive impulses or reduce symptoms, such as flashbacks, dissociative hallucinations, and passive-influence experiences. These interventions are usually far more robust clinically and rapidly acting than medications, particularly as crisis interventions. It is important that patients be taught to use these techniques independently of the therapist, often by having several self-states become "internal" hypnotherapists. In addition, teaching patients self-hypnosis usually gives them a sense of control over lifelong autohypnotic experiences that bewilder and overwhelm them. Hypnosis is also used to create relaxed mental states in which negative life events can be examined without overwhelming anxiety. Clinicians using hypnosis should be trained in its use in general and specifically in trauma populations. Clinicians should be aware of current controversies over the impact of hypnosis on accurate reporting of recollections and should use appropriate informed consent for its use. It is the consensus of experts on hypnotherapy for trauma disorders that it is important to educate the patient on the fact that the memory material that emerges during hypnotic interventions should be viewed no differently than any other memories that

emerge during treatment: they may be essentially veridical, partially accurate, or completely confabulated.

Psychopharmacological Interventions. In general, to date, dissociative memory and process symptoms, as psychopharmacological targets in and of themselves, have proven refractory to medications. Dissociative-like symptoms and behaviors, however, can be induced in some PTSD patients and normal individuals with drugs (e.g., phencyclidine, cannabinoids), suggesting the presence of pharmacologically sensitive neurobiological mechanisms that may lead to future medications that may target dissociative symptoms per se (see section on Neurobiology).

Guidelines for the use of medications with dissociative patients emphasize the need to identify specific treatment-responsive symptoms rather than attempting to treat the dissociation per se. Medications may be helpful in attenuating symptoms to assist the patient in stabilizing during treatment. Patients should be advised that medication response is likely to be partial, devising the best "shock-absorber" system for the patient at a given time. In general, success is more likely if the target symptoms for the medication are present across most or all personality states, rather than confined to one or a few. Because many individuals with dissociative identity disorder lead stressful lives, and psychotherapy for the disorder may also be stressful, rapid symptom oscillations are common. In general, it is not advisable to chase each symptom shift with a medication change. Expert consensus states that it is preferable to change medications in response to shifts in the psychological "climate," not the day-to-day alterations in the psychological "weather." Medications and medication targets are listed in Table 20–22.

Table 20–22.
Medications for Associated Symptoms in Dissociative Identity Disorder

Medications and somatic treatments for PTSD, affective disorders, anxiety disorders, and OCD

Alpha-Adrenergic Blocking Agent, Prazosin (Minipress) Very helpful for PTSD nightmares, and flashbacks, if patient can tolerate BP symptoms. Patients may require up to 20–25 mg daily for maximal effect with twice daily dosing for daytime flashbacks/intrusive symptoms. Doxazosin (Cardura) also useful for this indication and may help with daytime symptoms with single daily dosing—also may require 10–16 mg daily.

Selective serotonin reuptake inhibitors (no preferred agent, except for OCD symptoms)

Fluvoxamine (Luvox) (for OCD presentations)

Clomipramine (Anafranil) (for OCD presentations), if patient is safe enough to take a TCA

Other tricyclic antidepressants

Monoamine oxidase inhibitors (if patient can reliably maintain diet safely)

Mood stabilizers (*more useful for PTSD and anxiety than mood lability*)

- Lamotrigine (Lamictal) (may help with depressive symptoms more than others)
- Carbamazepine (Tegretol)
- Divalproex (Depakote)

Benzodiazepines

- Clonazepam (Klonopin) and lorazepam (Ativan) have best track records

Atypical neuroleptics
- Ziprasidone least likely to cause weight gain: a major issue for many dissociative patients.
- Clozapine for patients with mixture of psychotic and dissociative symptoms; very severe almost delusional cognitive distortions
- Typical neuroleptics (if patient fails trials of atypicals)

β-Blockers (for PTSD hyperarousal symptoms)
Clonidine (Catapres) (for PTSD hyperarousal symptoms/intrusive symptoms)
(Neither β-Blockers or Clonidine as consistently helpful as Prazosin)
If giving both β-blockers and α-blockers, give α-blocker first to prevent reflex tachycardia

Medications for thought disorder
- Atypical neuroleptics preferred

Medications for acute dyscontrol
- Oral or intramuscular neuroleptics
- Oral or intramuscular benzodiazepines

Medications for sleep problems
- Low-dose trazodone (Desyrel)
- Low-dose mirtazapine (Remeron)
- Low-dose tricyclic antidepressants
- Low-dose neuroleptics
- Benzodiazepines (often less helpful for sleep problems in this population)
- Zolpidem (Ambien)
- Anticholinergic agents (diphenhydramine [Benadryl], hydroxyzine [Vistaril])
- Melatonin and congeners sometimes effective

Medications for self-injury, addictions
- Naltrexone (ReVia)

OCD, obsessive-compulsive disorder; PTSD, posttraumatic stress disorder.

INFORMED CONSENT FOR PSYCHOPHARMACOLOGICAL INTERVENTIONS. Individuals with dissociative identity disorder require psychoeducation about which symptoms are likely to be medication responsive, the partial nature of medication effects for most comorbid disorders, and that even relatively robust medication effects may appear to be overwhelmed by severe stress—although without the medications, the patient may fare considerably worse during a crisis. These patients respond best if treated like a partner in evaluating the use of medications and medication response, with the physician as expert consultant, rather than in a more authoritarian stance.

Some individuals with dissociative identity disorder have an addiction-like affinity for medication that they use in an attempt to anesthetize all psychological distress. These patients may be reluctant to accept the shock-absorber model of psychopharmacology and struggle against the psychiatrist's attempt to keep the medication regimen as rational and simple as possible. They may come to treatment on complex, extensive polypharmacy regimens that may include a plethora of psychoactive medications including medications for psychiatric disorders, sleep, pain, and medical disorders. Much effort may be required to simplify these regimens, especially as more effective psychotherapeutic, symptom-management interventions are taught to the patient.

On the other hand, some individuals with dissociative identity disorder

report being drugged during their abuse, and have self-states that are phobic of taking medications. Considerable effort may then be required to educate the patient that medications are tools that potentially could be used for either good or malign purposes; it is how and why they are used that is important to assess. The patient can be asked to risk taking medications as prescribed, and to report honestly about medication effects, so that medication response and adverse effects may be better distinguished from posttraumatic reactivity. Without this process, the psychiatrist may later come to learn that the patient has been partially or fully nonadherent.

If the psychiatrist is working as a medication manager alongside a nonphysician psychotherapist, it is important to maintain communication about the patient. The therapist should let the psychiatrist know about issues that may be affecting symptoms, such as an acute loss or life crisis. In turn, the psychiatrist should inform the therapist about medication changes that may have an impact on the psychotherapy, and that the therapist can assist in assessing. Education of the therapist about the role of medications in dissociative identity disorder is important as well. It is key for the medicating psychiatrist to attend to issues of boundaries and treatment frame. It is vital that the patient have only *one* primary therapist and that the psychiatrist not become an additional therapist who sees the patient frequently (e.g., once weekly), unless major medication changes require more intensive contact. In addition, the psychiatrist should be clear that he/she is only available for medication/medical emergencies, and not be seen as the designated emergency back-up if the patient cannot reach the therapist during a psychotherapy or safety crisis. Also, the medicating psychiatrist should be clear about his/her availability for back-up during the therapist's absences, and only assume this role after specific agreement.

Target Symptoms for Medications in the Treatment of Dissociative Identity Disorder-PTSD Symptoms. The most robust medication treatment effect in this patient population is the response of nightmares to the α_1-adrenergic antagonist prazosin (Minipress) in a wide dose range, as documented in many studies, primarily of patients with combat-related PTSD. As long as the patient tolerates the blood pressure effects, dosages of prazosin as high as 25 mg daily, in single or twice-daily doses, are often remarkably effective for this indication. Daytime dosing may also be effective for control of flashbacks and intrusive symptoms. Recent studies show that Doxazosin (Cardura), a longer-acting α_1-adrenergic antagonist may also be effective for nightmares and intrusive symptoms in PTSD. In addition to prazosin/doxazosin, a variety of medications are helpful as adjunctive agents to help with sleep problems. These are listed in Table 20–22.

Nocturnal PTSD symptoms in this population, and in other patients with complex trauma disorders, may be complicated by posttraumatic

reactivity related to reported nocturnal physical and sexual abuse; then, nighttime, the bedroom, being in bed, and even sleep itself may be experienced as posttraumatic triggers. Also, patients may report that some self-states "stay up" to protect the patient from possible abuse, and/or emerge nocturnally, often in the early morning hours, as this was the only time the patient felt safe enough to allow awareness of internal experience. Because of this, some self-states may attempt, usually successfully, to psychologically override the beneficial effects of medications, or to refuse to take medications at all. These issues require psychoeducation about nocturnal PTSD triggers; cognitive behavioral interventions (e.g., attempts to decrease posttraumatic responses to these phobic stimuli and to help the patient separate past from the present); and symptom management techniques (e.g., finding a "safe place" to sleep for traumatized self-states before the patient as a whole attempts go to bed). Psychological work on sleep should begin several hours before bedtime.

Other medications have shown to be efficacious for PTSD symptoms in double-blind studies of PTSD patients. These include the selective serotonin reuptake inhibitors (SSRIs) sertraline (Zoloft) and paroxetine (Paxil), the tricyclic antidepressant (TCA) desipramine (Norpramin), and the monoamine oxidase inhibitor (MAOI) phenelzine (Nardil). Unfortunately, most medication studies of PTSD do not include patients with suicidal ideation, multiple comorbid disorders, and substance abuse, and thus may have limited ecological validity for complex trauma patients. Experts in psychopharmacology for this population have not found robust efficacy for these medications for comorbid PTSD symptoms, although, as discussed below, these medications may be helpful in ameliorating some of the affective disorder symptoms in these patients.

Other medications for PTSD symptoms that can have efficacy in reducing intrusive symptoms, hyperarousal, and anxiety in patients with dissociative identity disorder include β-blockers such as propranolol (Inderal), other α_1-adrenergic antagonists like clonidine (Catapres), anticonvulsants such as lamotrigine (Lamictal) and carbamazepine (Tegretol), and typical and atypical neuroleptics. The newer atypical neuroleptics, such as risperidone (Risperdal), quetiapine (Seroquel), ziprasidone (Geodon), aripiprazole (Abilify), and olanzapine (Zyprexa), may be more effective and better tolerated than typical neuroleptics for overwhelming anxiety and intrusive PTSD symptoms in patients with dissociative identity disorder. However, weight gain and the metabolic syndrome may be a particularly problematic side effect for these nonpsychotic patients. One small study showed that risperidone was efficacious for intrusive PTSD symptoms in complex PTSD patients. However, since the response of PTSD, anxiety, and panic symptoms to these medications is usually relatively limited in patients with dissociative identity disorder, the psychiatrist and patient together should carefully

balance the potential risks against the potential benefits of these medications. Typical neuroleptics may also be effective for this indication, but without the risk of the metabolic syndrome.

Many patients with dissociative identity disorder have some response to benzodiazepines such as lorazepam (Ativan) or clonazepam (Klonopin) for posttraumatic fear, panic, anxiety, and sleep problems. These patients may tolerate higher doses of these medications than many nondissociative patients. The clinician should educate the patient that severe PTSD involves states of terror, not anxiety. Thus, limited response to anti-anxiety medications is the rule. Careful monitoring of patients on benzodiazepines is necessary to assess tolerance and compulsive drug-seeking behaviors. Gradual tapering of these medications is essential to prevent withdrawal symptoms.

Target Symptoms for Medications in the Treatment of Dissociative Identity Disorder-Mood Disorder Symptoms. As described above, patients with dissociative identity disorder commonly have a complex negative mood state, often described as depression, but with elements of many other negative affective states. These patients commonly describe themselves as having been depressed for their "whole life." Nonetheless, antidepressant medications, including SSRIs, selective serotonin/norepinephrine reuptake inhibitors (SNRI), TCAs, and MAOIs are often important in the partial reduction of depression and stabilization of mood. The psychiatrist should take a careful history of the mood symptoms, and question the patient about other affects such as shame, sadness, and grieving; in addition, the psychiatrist should educate the patient that complete remission of the depressive symptoms is unlikely, and about the shock-absorber nature of medications. In general, expert consensus has shown that no antidepressant is likely to be any more effective than any other in this population. Somatic treatments for mood disorder symptoms in dissociative identity disorder patients may be most effective in a "double depression" clinical picture, although somatic treatments may only return the depressed mood back to the patient's usual chronic negative mood state.

A subgroup of dissociative identity disorder patients show marked obsessive-compulsive symptoms (OCD) and may preferentially respond to antidepressants such as fluvoxamine (Luvox) and clomipramine (Anafranil). During the clinical evaluation, clinicians should always assess dissociative identity disorder patients for the presence of OCD symptoms.

In general, unless the patients describe more clear-cut, sustained, hypomanic or manic episodes, they infrequently respond to mood stabilizers, especially Lithium. If they describe more typical bipolar depression symptoms, usually these too will follow the double depression pattern of treatment response.

OTHER TARGET SYMPTOMS FOR MEDICATIONS IN THE TREATMENT OF DISSOCIATIVE IDENTITY DISORDER. Case reports suggest that aggression may respond to carbamazepine in some individuals with dissociative identity disorder if EEG abnormalities are present. Open-label studies suggest that naltrexone (ReVia) may be helpful to ameliorate recurrent self-injurious behaviors and other compulsive, quasi-addictive high-risk behaviors in a subset of traumatized patients. This medication may be helpful in reducing relapse in dissociative identity disorder patients with significant substance abuse issues.

Questions are often raised about the efficacy of neuroleptic medications, particularly for symptoms such as hallucinations. Expert consensus is that the hallucinations in dissociative identity disorder patients almost never respond robustly to antipsychotics, as they are hypothesized to originate in dissociative, posttraumatic, and autohypnotic processes, not psychosis. Patients with dissociative identity disorder with a true comorbid psychotic disorder can sometimes distinguish between neuroleptic response to the "outside" psychotic voices and report lack of effect on the voices of the self-states. Occasionally, thought-disorganized, overwhelmed, chronically ill patients with dissociative identity disorder, who have not responded to trials of other neuroleptics, respond favorably to a trial of clozapine (Clozaril).

Electroconvulsive Therapy. During their long psychiatric careers of misdiagnosis and treatment failure, many patients with dissociative identity disorder receive unsuccessful trials of ECT. This led to the view that ECT was not an effective treatment in this patient population. One small clinical study of severely depressed dissociative disorder patients, primarily with DSM-III-R dissociative disorder NOS, suggested that, for some patients, ECT was helpful in ameliorating refractory mood disorders and did not worsen dissociative memory problems as measured by the DES. Clinical experience in tertiary care settings for severely ill individuals with dissociative identity disorder suggests that a clinical picture of major depression with persistent, refractory melancholic features across all states may predict a positive response to ECT. In particular, as with medications, response to ECT is more likely in a "double depression" clinical presentation where melancholic symptoms develop as a distinct change from the patient's chronically depressed baseline. However, ECT response is usually only partial, as is typical for most successful somatic treatments in patients with dissociative identity disorder, and returns the patient to his/her more usual mood state.

Inpatient Treatment. As with patients with other psychiatric diagnoses, inpatient hospitalization is most commonly used for treatment of individuals with dissociative identity disorder who are acutely unsafe,

completely destabilized, or both. Less commonly, in the age of managed care, hospitalization provides a safe environment for intensive work with painful affects and memories that cannot be managed on an outpatient basis despite intensive efforts. For several reasons, patients with overt dissociative identity disorder can have disruptive effects on the milieu of general psychiatric units and often stimulate division and conflict among staff about the best way to relate to the patient and the personality states. Typically, the split is between the believers and the skeptics, often paralleling larger societal debates, as well as the subjective conflict within the patient about belief in the dissociative identity disorder diagnosis and the authenticity of traumatic memories. Inpatient treatment on general hospital units is most successful if there is an attempt to bridge different opinions about dissociative identity disorder to focus on a unified approach for ameliorating specific target symptoms to allow the patient to resume outpatient treatment in safety.

Inpatient programs specializing in the treatment of dissociative and PTSD patients are more successful. They provide a unified treatment with a consistent approach to the patient's internal conflicts; rigorous boundaries and firm limit setting; focus on specific, active treatment interventions to assist with dissociative and PTSD symptoms; and an emphasis on helping the patient become proficient in symptom management skills.

In both specialty and general hospital units, the patient should be expected to respond to his or her legal name, or at least one specific name, in the hospital milieu and in all therapeutic interactions, except individual psychotherapy or individual interactions with designated hospital staff. The patient should be expected to be able to present in a reasonably functional adult mode in the hospital, with regressed, dysfunctional, or child states restricted to designated one-on-one encounters or private time in the patient's room or quiet room. The patient is urged to seek his/her "inner adult" in this regard. The patient's activities in the hospital (e.g., group therapy) should be limited to those in which the patient can participate while following these guidelines. Some patients with dissociative identity disorder may be so overwhelmed or overstimulated in the hospital that they need respite in the quiet room. However, this can become a permanent berth for the patient, who may need to be weaned back into more routine hospital activities to prevent a major regression.

Because of past trauma, including prior experiences in hospitals, individuals with dissociative identity disorder may experience everyday aspects of the psychiatric hospital, let alone crisis management strategies as abusive, including the use of the quiet room, all-staff call, or restraints. Staff should proactively help the patient to find reasonable clinical alternatives for possible dyscontrol. Hospital staff may still need to use hands-on, restraints, or similar measures if the patient remains acutely dangerous to self or others and when less restrictive alternatives have failed

to de-escalate the situation. Specialty units generally use restraint much less frequently with these patients than general hospital units. Specialty unit staff is trained to help patients with dissociative identity disorder to use symptom management and containment strategies to handle potential crises and to minimize the use of physical interventions to control behavior. Inpatients with dissociative identity disorder in behavioral dyscontrol may also respond to as-needed oral dosages of atypical or typical neuroleptics, with or without benzodiazepines. The current emphasis on "trauma-informed care" in psychiatric hospital units, with the idea of bringing seclusion and restraint to zero, may be particularly well suited to many hospitalized dissociative identity disorder patients.

If the dissociative inpatient shows a limited or transitory response to psychosocial interventions or oral prn medications for severe dyscontrol, intramuscular medication may help prevent escalation and/or more restrictive intervention. Single doses of intramuscular neuroleptics may help manage acute dyscontrol by inducing sleep to allow a change in state once the patient reawakens. Options include 2 to 5 mg of haloperidol (Haldol); 5 to 10 mg of sublingual or intramuscular olanzapine (Zyprexa); or 20 mg of intramuscular ziprasidone (Geodon); with or without 1 to 2 mg of intramuscular lorazepam (Ativan).

Discharge from hospital usually occurs when the patient with dissociative identity disorder has negotiated the acute crisis; has made adequate gains in symptom management skills and internal communication to handle problems more adaptively; medications have been adjusted to provide better treatment of comorbid disorders; the outpatient psychosocial environment has restabilized sufficiently; and/or a treatment agreement for outpatient safety has been reached that is consistent with discharge to the next level of care.

Partial Hospitalization. Partial hospital treatment can be an effective modality for the dissociative patient if clinical interventions target trauma-based issues and adaptations. Specialized trauma partial hospital programs with groups emphasizing symptom containment, cognitive and behavioral interventions, development of life skills, and psychoeducation can effectively help stabilize more severely and persistently ill dissociative patients, as well as provide an intensive stabilization experience for higher-functioning individuals. Patients with dissociative identity disorder frequently do not respond positively to more usual day treatment programs developed for patients with psychotic and/or mood disorder.

Adjunctive Treatments

GROUP THERAPY. Authorities agree that, in therapy groups including general psychiatric patients, the emergence of self-states can be disruptive

to the group process by eliciting excess fascination or by frightening other patients. Therapy groups composed only of patients with dissociative identity disorder are more successful, although the groups must be carefully structured, must provide firm limits, and should generally focus only on here-and-now issues of coping and adaptation for stage 1 patients. Some stabilized patients may be able to work effectively in psychotherapy process-oriented groups. Even so, it is important that there are firm boundaries and limits on group structure and contacts by patients outside of sessions; the intensity of trauma material that the patients can discuss be monitored; and that patients who become destabilized and unsafe may need to return to a group that focuses on symptom management and stability.

FAMILY THERAPY. Family or couples therapy is often important for long-term stabilization and to address pathological family and marital processes that are common in individuals with dissociative identity disorder and their family members. Education of family and concerned others about the disorder and its treatment may help family members cope more effectively with dissociative identity disorder and PTSD symptoms in their loved ones. Group interventions for education and support of family members have also been found helpful. In particular, family members should be discouraged from interacting with individual personality states, calling them out, and relating to them as separate individuals. Family members should be helped to support the patient's goal of becoming a functional adult in adult relationships, and a functional parent to children. Sex therapy may be an important part of couples' treatment, when the patient is more stabilized, as the patient may be highly triggered into flashback by sexual interactions and thus intensely phobic of intimate contact. Clinicians who work with these patients and their spouses on sexual issues should be trained in work with survivors of childhood sexual abuse and on the specific complexities of these issues with patients with dissociative identity disorder, including the patient's difficulty in excluding traumatized self-states from intimate interactions.

Family therapy with the individual's family of origin is often extremely distressing to patients with the dissociative identity disorder. Such meetings should be set up carefully, because some interactions with family members may lead to an acute decompensation in the patient. These contacts should be carefully structured for defined therapeutic purposes, and the patient should be well prepared for potential adverse reactions. At the same time, such meetings may help clarify or resolve the conflictual emotions that patients frequently experience toward family members. Confrontation of family members about past traumas in an accusatory manner almost invariably has a disastrous outcome for the patient and the relatives. There is usually little place for such interventions in dissociative

identity disorder treatment.

DIALECTICAL BEHAVIOR THERAPY. DBT is a treatment developed by the psychologist Marsha Linehan for borderline personality disorder. Its concepts and techniques may be helpful to structure aspects of stabilization throughout dissociative identity disorder treatment, since many patients with borderline personality disorder report significant trauma histories, and fit the complex trauma construct. Helpfully for the patient with dissociative identity disorder, DBT focuses on the primacy of safety, therapy-interfering behaviors, affect regulation strategies, distress tolerance, and radical acceptance.

However, DBT also may have notable limitations for the patient with dissociative identity disorder. It does not focus on stabilization and management of dissociative symptoms such as depersonalization/derealization and dissociative amnesia, flashbacks, and unconscious reenactments, traumatic transference, and related trauma-based symptoms. DBT does not address stabilization of the self-states' contribution to symptoms, since resolution of many symptoms in dissociative identity disorder treatment can only occur in the context of work with the patient's self-state system.

EYE MOVEMENT DESENSITIZATION AND REPROCESSING. Eye movement desensitization and reprocessing (EMDR) is a treatment for posttraumatic disorders developed by the psychologist Francine Shapiro. It has several components: a comprehensive history of the patient's experiences that could be a focus for EMDR, and assessment of the patient's capacity to manage symptoms and work on specific traumas or problems. Direct work on the traumatic event involves focus on a vivid image, emotions, body sensations, negative beliefs, and a positive belief. During this time, the therapist provides bilateral stimuli, such as moving a finger from side to side, tones, or tapping. During this process, the therapist suggests to the patient that he/she freely associate and allow any thoughts, feelings, sensations or cognitions into awareness. After the patient reports the intensity of the memory has diminished, the patient focuses on the positive belief. The eye movements or bilateral stimuli continue through this process.

In treatment outcome studies for PTSD, EMDR has similar efficacy to cognitive-behavioral therapy. However, patients with suicidality, substance abuse, and multiple comorbidities have generally been excluded from EMDR outcome studies, as in other treatment studies of cognitive-behavioral therapy and medications for PTSD. Dismantling studies have shown that the eye movements that distinguish EMDR from other therapies do not add to its therapeutic effect, although EMDR practitioners dispute this.

There are no systematic studies of EMDR in patients with dissociative identity disorder, although some case reports have been published. Experts in both EMDR and phasic trauma treatment of dissociative identity disorder have written an appendix to the standard text on EMDR, and a section of the ISSTD treatment guidelines. The authors strongly suggest not using this modality in dissociative identity disorder treatment unless the patient is stabilized in stage 1 treatment.

Because EMDR has the significant exposure and free association components described above, in unmodified form this modality can cause significant harm to individuals with dissociative identity disorder. The lead author has considerable experience with adverse outcomes of EMDR when used in these patients during stage 1 of treatment, or before even beginning stabilization in treatment. These outcomes have included severe posttraumatic and dissociative crises, suicide attempts, self-destructive behavior, and acute worsening of PTSD, dissociative, and mood disorder symptoms, often resulting in emergency hospitalization, as well as long-term impairment.

The ISSTD Treatment Guidelines, clinical case series, and the EMDR text appendix state that clinicians may choose EMDR as an adjunctive treatment in some cases to allow the patient to process memories during stage 2 treatment. Specifically, in order for the clinician to practice EMDR with patients with dissociative identity disorder, the clinician must be knowledgeable and skilled in phasic trauma treatment, have completed advanced training in EMDR, and have received supervision in this modality for the treatment of dissociative identity disorder.

SELF-HELP GROUPS. Individuals with dissociative identity disorder usually have a negative outcome to self-help groups or 12-step groups for incest survivors. Most experienced clinicians, as well as the ISSTD guidelines, strongly discourage these patients from participating in these modalities. Complications include intensification of PTSD symptoms due to discussion of trauma material without clinical safeguards, exploitation of the patient with dissociative identity disorder by predatory group members, contamination of the patient's recall by group discussions of trauma, and a feeling of alienation even from these other reputed sufferers of trauma and dissociation.

EXPRESSIVE AND OCCUPATIONAL THERAPIES. Expressive and occupational therapies, such as art and movement therapy, have proven particularly helpful in treatment of patients with dissociative identity disorder. Art therapy may help patients contain and structure severe dissociative identity disorder and PTSD symptoms, as well as allow for safer expression of thoughts, feelings, mental images, and conflicts that may be difficult to verbalize. Movement therapy may facilitate normalization of body sense

and body image for these severely traumatized patients. Occupational therapy may help the patient with focused, structured activities that can be completed successfully and may help with grounding and symptom management. A case example summarizes some of the issues in the treatment of dissociative identity disorder.

DISSOCIATIVE IDENTITY DISORDER CASE 6

Ms. V was a 33-year-old married woman presented to psychiatric attention after discovering her 5-year-old daughter "playing doctor" with several neighborhood children. The patient began to become fearful that her daughter would be molested. She became increasingly obsessed with this idea, much to the bafflement of her husband. The patient was seen by her internist and was treated with antianxiety agents and antidepressants, but with little improvement. Ms. V became increasingly anxious, phobic, depressed, and preoccupied. She sought psychiatric consultation from several clinicians, but repeated, good trials of antidepressants, antianxiety agents, and supportive psychotherapy resulted in limited improvement. After the death of her father from complications of alcoholism, the patient became more symptomatic. He had been estranged from the family since the patient was approximately 12 years of age, owing to his drinking and associated antisocial behavior. Ms. A developed a variety of somatic complaints, including headaches, abdominal pain, menstrual and gastrointestinal problems, back pain, and sleep difficulties. Repeated medical workup was unrevealing, leading to diagnoses such as fibromyalgia, irritable bowel syndrome, and premenstrual tension. Family and marital difficulties increased and work function also deteriorated. Psychiatric hospitalization was precipitated by the patient's arrest for disorderly conduct in a nearby city. She was found in a hotel, in revealing clothing, engaged in an altercation with a man. She denied knowledge of how she had come to the hotel, although the man insisted that she had come there under a different name for a voluntary sexual encounter. On psychiatric examination, the patient described dense amnesia for the first 12 years of her life, with the feeling that her "life started at 12 years old." She reported that, for as long as she could remember, she had an imaginary companion, an elderly African American woman, who advised her and kept her company. She reported hearing other voices in her head: Several women and children, as well as her father's voice repeatedly speaking to her in a derogatory way. She reported that much of her life since 12 years of age was also punctuated by episodes of amnesia: for work, for her marriage, for the birth of her children, and for her sex life with her husband. She reported perplexing changes in skills. The patient reported having repeated experiences of sudden "blasts" of emotion that would appear, and suddenly disappear, primarily anger, fear, and sadness. She described feeling that her thoughts were being removed from her mind, especially when attempting to recall difficult issues in life. She reported experiencing saying things that felt out of her control: "My mouth has a mind of it's own; my mind has a mouth of it's own; my mind has a mind of it's own!" At times, she reported automatic writing, as if her "hand is possessed," and different handwritings appearing without her control. Her husband noted that, at times, she would speak like a child; at times, she would adopt a southern accent; and, at other times, she would be angry and provocative. She frequently had little recall of these episodes. Questioned more closely about her early life, the patient appeared to enter a trance and stated, "I just don't want to be locked in the closet" in a child-like voice. Inquiry about this produced rapid shifts in state between self states that differed in manifested age, facial expression, voice tone, and knowledge of the patient's history. Gradually, with awareness of a number of self states, the patient, recalled a history of family chaos, brutality, and neglect during the first 12 years of the patient's life, until her mother, also alcoholic, achieved sobriety and fled her husband, taking her children with her. The patient began to recall episodes of physical abuse, sexual abuse, and emotional torment by the father, her siblings, and her mother. The self states reported "piggy-backing" emotions, thoughts, impulses, and actions, leading to the experiences of passive-influence. This also explained "taking thoughts out of her head, that she isn't supposed to know," when the patient suddenly blanked on what was being discussed. The patient improved significantly with psychotherapy directed at stabilization of her dissociative identity disorder and PTSD. She responded well to clomipramine (Anafranil) with a significant reduction in obsessive-compulsive and depressive

symptoms. Family sessions with the patient's mother and siblings confirmed many of these reports, with the family recalling episodes of maltreatment that the patient did not recollect. The patient's mother had bid the family never to speak of the earlier difficulties, hoping that everyone would "just forget the whole thing." An older sister, who also had been molested, met diagnostic criteria for dissociative identity disorder. Family therapy was helpful in stabilizing the patient's marriage and helping her husband with the aftermath of the hospitalization and its precipitants. The patient found adjunctive art therapy particularly helpful. At 3-year follow-up, the patient reported fusion of most alters and marked diminution in dissociative, somatoform, and PTSD symptoms, although she still required clomipramine for stabilization of mood and OCD symptoms.

DISSOCIATIVE AMNESIA

As described above, the DSM-5 Work Group clarified the diagnostic criteria for dissociative amnesia to emphasize that the disorder represents a deficit in *autobiographical memory,* particularly for traumatic and stressful experiences. The types of dissociative amnesia have been listed in the section on Epidemiology, found in Table 20–4.

As described in the prior sections, cumulative adverse experiences during childhood can lead to the inability to recall substantial aspects of early life before age 18. The Work Group decided that, based on the extant data, the diagnosis of dissociative fugue described in prior editions of the DSM was better conceptualized as a specifier for dissociative amnesia. In the dissociative fugue presentation of dissociative amnesia, there is apparently purposeful travel, or bewildered wandering accompanied by dissociative amnesia for identity and/or other important autobiographical information, including details of the travel the person has engaged in, when the person comes to clinical attention.

The dissociative amnesia cannot be better explained by a medical, neurological, or substance use condition, including acute intoxication, partial complex seizures, transient global amnesia, or head injury. In addition, the dissociative amnesia cannot be better explained as a symptom of dissociative identity disorder, ASD, PTSD, somatic symptom disorder, a psychotic or cyclic mood disorder, or a neurocognitive disorder.

What Is Dissociative Amnesia?

Dissociative amnesia can be broadly defined as a reversible autobiographical memory impairment in which groups of memories for personal experience that would ordinarily be available for conscious recall cannot be retrieved or retained in the form of usual autobiographical memory (or, if temporarily retrieved, cannot be wholly retained in consciousness). This disturbance may be based on neurobiological changes in the brain caused by traumatic stress. However, the disorder manifests itself as a potentially reversible form of psychological inhibition. Unlike dissociative amnesia, amnesia due to brain injury or cognitive disorders usually cannot be reversed, as is the case with most episodes of amnesia due to substance intoxication. On the other hand, a variety of clinical and

experimental studies have shown state dependent effects for some forms of memory retrieval with a variety of types of intoxicants (e.g., alcohol, barbiturates, amphetamines, marijuana) and with some mood state shifts (e.g., mania vs. depression). In these state dependent/mood state dependent memory studies, reinstatement of context is important in demonstrating the phenomenon. Reinstatement of context, as well as cued retrieval, may be a factor in the recall of memories that have been subject to dissociative amnesia, although, as described below, this usually comes in the form of PTSD intrusive symptoms and flashbacks, not normal autobiographical recall.

The diagnosis of dissociative amnesia generally connotes four factors. First, relatively large groups of memories and associated affects have become unavailable, not just single memories, feelings, or thoughts. Second, the unavailable memories usually relate to day-to-day autobiographical information that would ordinarily be a more or less routine part of conscious awareness: Who a person is, what he or she did, where he or she went, what happened, with whom he or she spoke, what was said, what he or she thought and felt at the time, and so forth. Third, the ability to remember new factual information, general cognitive functioning, and language capacity are usually intact, although in extreme cases the dissociative process can interfere with retrieval of procedural information and registration of new memories. Finally, the dissociated memories often indirectly reveal their presence in more or less disguised form, such as intrusive visual images, flashbacks, somatoform symptoms, nightmares, conversion symptoms, and behavioral reenactments. In most cases, dissociative amnesia should be understood as part of the spectrum of memory dysfunction related to traumatic stress, often alternating with forms of hyperamnesia and flashbacks or a derealized awareness in which the person experiences detachment or estrangement from elements of autobiographical memory, or both.

There are two basic presentations of dissociative amnesia. The first, generalized dissociative amnesia, is a dramatic, sudden disturbance in which extensive aspects of memory for autobiographical information and lived experiences are not available to conscious verbal recall, often including loss of awareness of personal identity. These patients are often seen in emergency departments or general medical or neurological services, because the sudden development of memory loss requires medical assessment. In addition, during an acute amnestic episode, some individuals demonstrate disorientation, perplexity, alteration in consciousness, somatoform symptoms, purposeless or fugue-type wandering, or a combination of these. A case example follows.

DISSOCIATIVE AMNESIA CASE 1

A 45-year-old, divorced, left-handed, male bus dispatcher was seen in psychiatric consultation

on a medical unit. He had been admitted with an episode of chest discomfort, light-headedness, and left-arm weakness. He had a history of hypertension and ischemic chest pain, without a myocardial infarction. Psychiatric consultation was called, as the patient complained of memory loss for the previous 12 years, behaving and responding to the environment as if it were 12 years previously, e.g., he insisted that he was unmarried, denied he had a son, did not recall his current job, and denied recollection of current events (e.g., current president), but correctly recalled historical events from 12 years before. Physical and laboratory examinations were unchanged from the patient's usual baseline; imaging studies of the brain were normal. The patient displayed intact intellectual function other than the 12-year amnesia. He was perplexed by the contradiction between his memory and current circumstances. The patient described a family history of brutal beatings and physical discipline. He was a decorated combat veteran, although he described amnestic episodes for some of his combat experiences. In the military, he had been a left handed golden glove champion. He was educated about his disorder and given the suggestion that his memory could return as he could tolerate it, perhaps overnight during sleep or perhaps over a longer time. If this strategy was unsuccessful, hypnosis or an amobarbital interview was proposed.

At the next examination, the patient reported that his memory had spontaneously returned. Before the amnestic episode, he described that his wife was discussing a separation and had asked him to discuss this with his son. In addition, due to a variety of other family and work problems, he had come to feel completely responsible for his coworkers and for the care of his relatives. He was panicked, overwhelmed, and enraged. He had felt violently angry with his wife but had vowed in the past that he would "beat to death anybody who tried to hurt her." He stated that he would have attempted suicide, but he "couldn't," because he had too many people relying on him. The amnesia developed after he felt a kind of "paralysis" in his left arm while talking with his wife about the separation. Insightfully, he reported that, "I went back to the time, 12 years ago, before I met my wife, had this job and all this family stress. Back then, none of these problems existed."

Despite its relative rarity, generalized dissociative amnesia is featured in the media and in most textbooks as representative of the condition. A far more prevalent form of dissociative amnesia is a deletion from conscious memory of significant aspects of the personal history. Ordinarily, patients do not complain of this, and it is usually only discovered in taking a careful life history. Dissociative amnesia typically has a clear-cut onset and offset, so that the person is subjectively aware of a gap in continuous memory. For example, a patient may report that she does not remember being in third grade despite having a clear memory of other school years. Usually such symptoms are associated with traumatic circumstances: for example, the patient reports being told that during third grade, she was kidnapped by her estranged father in a custody dispute and was held by him and sexually abused for several months. In extreme cases, the patient may deny recall for his or her entire childhood or other major life epochs.

Diagnosis and Clinical Features

There are two major clinical presentations of dissociative amnesia.

Classic Presentation. Classically presenting patients are the textbook cases that form the image of dissociative disorders for most mental health professionals. The classic disorder is an overt, florid, dramatic clinical disturbance that frequently results in the patient being brought quickly to

medical attention specifically for symptoms related to the dissociative disorder. The paradigmatic case is that of the individual who is found without memory for identity or life history, sometimes leading to media reporting of the case. Less extreme forms of amnesia, such as acute amnesia for recent traumatic circumstances (e.g., rape), also fall into this category.

This presentation of amnesia is frequently found in individuals who have experienced extreme acute trauma, such as combat or rape. However, it also commonly develops in the context of profound intrapsychic conflict or emotional stress, for example, a terrified soldier in combat who experiences excruciating conflict over his or her urge to flee despite believing that "all cowards should be shot."

There are very few systematic data about patients with this form of dissociative amnesia, other than individual case studies or small case series. A systematic neuropsychological study assessed memory function in 10 patients with severe retrograde dissociative amnesia, of a median duration of 14 months. They were assessed on tests of retrograde and anterograde amnesia, as well as of autobiographical memory. One respondent admitted to malingering, and 2 acknowledged symptoms consistent with dissociative identity disorder, with the acute dissociative amnesia a manifestation of an amnestic self-state. Other than the malingerer, every patient had a long psychiatric history; dissociative respondents evidenced symptoms that were not feigned, and improved or remitted with treatment. All subjects reported difficulty recalling specific autobiographical memories. In addition, tests of famous faces and past public events revealed that individual patients showed unique, but internally consistent, memory difficulties for this material. These was not thought to represent factitious processes, but rather, the patients' commonsense, individualized notions of how memory works that informed the phenomenology of their presentations.

Patients may present with comorbid somatic and/or conversion symptoms, as in the bus dispatcher case described above. They also may display alterations in consciousness, depersonalization, derealization, trance states, spontaneous age regression, dissociative fugue, and even ongoing anterograde dissociative amnesia. Depression and suicidal ideation are reported in many, but not all, cases. No single personality profile or antecedent history is consistently reported, although a personal or family history of somatoform or dissociative symptoms predisposes individuals to develop acute amnesia during traumatic circumstances. Case reports suggest that many have histories of prior adult or childhood abuse or trauma. However, in wartime cases, as in other forms of combat-related posttraumatic disorders, the most important variable in the development of dissociative symptoms appears to be the intensity of combat.

Factors relating to avoidance of responsibility may be prominent.

Sexual indiscretions, legal difficulties, financial problems, or fear of anticipated combat may be part of the clinical matrix that surrounds the amnesia. There is a possible association in some cases with an antecedent history of head trauma with or without loss of consciousness, although this finding has never been studied rigorously using adequate controls. Unfortunately, data on the prevalence of PTSD symptoms in these patients is lacking.

If carefully questioned, some of these patients will give a history of recurrent episodes of amnesia or fugue, and some will meet criteria for dissociative identity disorder. In the dissociative identity disorder cases, the amnesia occurs when—in an attempt to cope with overwhelming or traumatic circumstances and/or profound intrapsychic conflict—the person begins to experience him- or herself as a new identity state that is unaware of some or all of the past history of the person.

In most acute dissociative amnesia cases, the amnesia resolves within hours to months, spontaneously or through psychotherapy, hypnotherapy, pharmacologically facilitated interviews, or combinations of these modalities. However, in some cases, the amnesia becomes chronic and the patient appears unable to tolerate recall of the events that surrounded its onset. Clinically, chronic dissociative amnesia may be severely disabling, associated with a limited ability to return to family, social, or occupational roles, even when patients "learn" about the missing autobiographical information. This acute presentation should be regarded as a psychiatric emergency, as clinical experience suggests that the longer the period of time before resolution of the amnesia, the more likely that a chronic course will occur. A case example follows:

DISSOCIATIVE AMNESIA CASE 2

Ms. M is a 55-year-old woman was seen in psychiatric consultation because of sudden, complete loss of memory for her identity and entire life history. She had been found by her family after participation with a female friend in a "new age" therapy workshop where, in a session involving picturing oneself going back in time and being reborn, the patient appeared to enter a deep trance state in a fetal position, and had difficulty awakening. When she did, she could not identify who she was, where she was, or anything else about herself or her situation. Several days later, when found by her family, she appeared dazed, confused and disoriented, unable to recognize them or provide any history of recent events. She had difficulty learning new information, seemingly forgetting it as it was provided. She was panicked and complained of not "feeling in my body." She had difficulty recalling how to dress herself, put in her contact lenses, etc. She underwent a complete neurological, neuropsychiatric, and neuropsychological evaluation, with the conclusion that the process was psychological, not a cognitive disorder. When referred for consultation, she had "relearned" information about her identity, family, and her life, but fundamentally had no available autobiographical memory in which to contextualize this information. She complained of ongoing amnesia, spontaneous trances, and periods of intense depersonalization and derealization. She was impaired in performing most daily functions, having no recall of these skills. She had been depressed, and this improved with antidepressant medication, but her memory remained unchanged. Her memory did not return with hypnotic interventions or after an amobarbital interview. In psychotherapy, she gradually was able to recall events around the loss of memory. The "friend" who had accompanied her to the conference had pressured her sexually, and, with great distress, she

finally acceded. This event, experienced with profound shame, had uncanny parallels to episodes of childhood sexual abuse that the patient had described in a prior psychotherapy. After this disclosure, the patient began to have gradual return of her autobiographical memory.

DISSOCIATIVE FUGUE SUBTYPE. The dissociative fugue subtype is usually associated with memory loss typical of the classic presentations of dissociative amnesia. Three types of fugue were described in the 1940s: (1) fugue with awareness of loss of personal identity, (2) fugue with change of personal identity (likely dissociative identity disorder in DSM-5), and (3) fugue with retrograde amnesia. These three types of fugue have been described as each unfolding over the course of three stages. In stage I, there is thought to be an altered state of consciousness appears that still allows an individual to conduct complex activities sometimes over long periods. As originally described by Janet, a single idea that symbolizes or condenses (or both) several important ideas and emotions frequently dominates the patient's thinking in this stage. In stage II, the person becomes aware of the amnesia or loss of personal identity, at which point he or she frequently is brought for treatment. In this stage, amnesia is usually present for the first stage. In stage III, the individual returns to his or her baseline state, usually with amnesia for the first stage and sometimes the second. An alternative view describes stage II as a return to the baseline state with amnesia for stage I or to a state in which there is (1) awareness of loss of personal identity, (2) change in personal identity (dissociative identity disorder), or (3) return to a chronologically earlier period of life, similar to a spontaneous hypnotic age regression.

Dissociative fugues may last from minutes to months. Some patients report multiple fugues; however, in most cases in of this, a more chronic dissociative disorder such as dissociative identity disorder is the likelier diagnosis, as in the classical dissociative fugue cases with change in personal identity.

In some extremely severe cases of PTSD, nightmares may be terminated by a waking fugue in which the patient runs to another part of the house or runs outside. Children or adolescents may be more limited than adults in their ability to travel. Thus, fugues in this population may be brief and involve only short distances. Some cases of children or adolescents who precipitously run away actually may be instances of dissociative fugue, with the child escaping from an abusive or violent home situation, as in the case described below. No systematic research exists on this point.

During a fugue, patients often seem not to be suffering from psychopathology and tend not to attract attention. Some individuals, however, display overtly bizarre, disorganized, or dangerous behavior, such as a soldier in the midst of battle who began a fugue episode by standing up and walking away from the front lines, exposing himself to intense enemy fire.

After the termination of a fugue, the patient may experience perplexity, confusion, trance-like behaviors, depersonalization, derealization, and conversion symptoms, in addition to amnesia. Some patients experience an episode of generalized dissociative amnesia at the end of a fugue state. They may be brought to media attention in an attempt to discover their origin and identity.

As the dissociative fugue patient begins to become less dissociated, he or she may display mood, anxiety or PTSD symptoms, or intense suicidal ideation. In the classic cases, a new identity is experienced under whose auspices the patient lives for a period of time. Today, it is likely that many of these latter cases would be classified as dissociative identity disorder. Indeed, William James described the classic case of Ansel Bourne as one of "dual consciousness." Bourne disappeared from his family, and moved to another state, where he opened a shop as A.J. Brown. After 3 months, his memory returned for his prior life, but not for the time as Brown. Under hypnosis, either identity could be accessed, each with amnesia for the other.

Nonclassic Presentation. Individuals with nonclassic dissociative amnesia can be said to have a *covert* dissociative syndrome, because their primary complaint infrequently relates directly to amnesia. Commonly, patients with the nonclassic presentations of amnesia do not reveal the presence of dissociative symptoms unless directly asked about them. These patients are often uncomfortable when amnesia is inquired about and may minimize the presence or rationalize the importance of the symptom. Yet, chronic, recurrent, or persistent dissociative amnesia, or a combination of these, is usually the most common symptom in these cases, although some individuals may also describe a history of fugue behaviors, consistent with the fugue subtype of dissociative amnesia.

In nonclassic presentations, the amnesia consists of a circumscribed memory gap or series of memory gaps for the life history, often for times when traumatic events occurred, such as childhood or wartime. However, as described in prior sections, more extensive dissociative amnesia for life history also may occur, with the person describing large gaps in recall, and/or fragmentary recall of years of early life. These patients frequently come to treatment for a variety of symptoms, such as depression or mood swings, substance abuse, sleep disturbances, somatoform symptoms, anxiety and panic, suicidal or self-mutilating impulses and acts, violent outbursts, eating problems, and interpersonal problems. Self-mutilation and violent behavior may also be accompanied by amnesia. Amnesia may also occur for flashbacks or behavioral re-experiencing episodes related to trauma. A case example follows:

DISSOCIATIVE AMNESIA CASE 3

Ms. G is a 35-year-old psychologist who was admitted to hospital because of repeated acts of severe self-mutilation. She reported chronic depressed mood, anxiety, and interpersonal problems with inability to form intimate relations. Medications and intensive outpatient psychotherapy had helped moderate some of her symptoms, but she had become increasingly demoralized by her overall failure to improve in treatment. She described virtually no recall for her life between the ages of 7 and about 13 years old, including family holidays, school events, and vacation trips. Her father had been alcoholic throughout her childhood, only achieving sobriety in her early adolescence. When she was 13, her older sister disclosed sexual abuse by the father to their minister, who "did a church intervention," but there was no social service or police referral. Her father acknowledged these episodes at the time, and received treatment for alcoholism and "sex addiction" by church counselors. He denied any other episodes of sexual abuse towards his children. Ms. G denied any feelings about these episodes stating that, "he took care of the problem. Now, he's a great support. I have no reason to be mad at him." She explained her amnesia by saying that, "maybe nothing important happened and that's why I don't remember." Treatment focused on establishing safety from self-injury. Self-harm impulses increased dramatically after phone conversations with her parents. Also, she began to have nightmares, highly disturbing bodily sensations in her genitals and rectum, and would wake up vomiting. Realizing that these were symptoms of PTSD, she found this uncanny, stating, "Nothing bad ever happened to me." Finally, she began to have discrete flashbacks of the father sexually assaulting her, together with her sister at times. In an emotional family meeting, the sister, long estranged from the family, confirmed that these events had occurred. She said, "I had hoped you would never remember, so I never told you."

Only a few studies have examined the clinical characteristics of nonclassic dissociative amnesia patients. In a small, highly selected sample referred for consultation at a specialty clinic for dissociative disorders, most dissociative amnesia patients were women. As in the data from the ACE study, almost all had a prior history of childhood or adult physical, sexual, or emotional abuse and neglect, or a combination of these. The trauma history was less severe than has been reported in patients with dissociative identity disorder. Duration of dissociative amnesia ranged from relatively brief recurrent episodes to years of the life history. Traumatic precipitants of the amnesia were most common, although in 30 percent of cases amnesia was present for problematic behaviors, such as sexual indiscretions and self-mutilation. The most common comorbid conditions were mood disorders, PTSD, and mixed personality disorders. Patients differed from dissociative identity disorder patients in that they had lower rates of substance abuse, self-mutilation, hallucinations, fugues, sexual dysfunction, and somatoform symptoms. Mean DES scores tended to be lower than in dissociative identity disorder patients. Family history was characterized by alcoholism and mood disorders.

A random community sample of 1,301 university students in Sivas, Turkey was assessed for dissociative disorders and borderline personality disorder. Eighty respondents (6 percent) of the sample met DSM-IV-TR criteria for borderline personality disorder, 78 of whom (98 percent) also met DSM-IV-TR criteria for a dissociative disorder. Of these, 15 participants (18.8 percent of the borderline personality disorder group, 1.2 percent of the total sample) met diagnostic criteria for DSM-IV-TR dissociative amnesia. In this study, as in the data from Ross and colleagues

described under *Dissociative identity disorder,* a diagnosis of borderline personality disorder acted essentially as a severity measure, indicating a higher level of childhood trauma exposure and higher levels of dissociation.

Pathology and Laboratory Examination

Dissociative amnesia can be diagnosed with the DDIS or the SCID-D-R, or both, and the SDAQ can be used as screening measure. Dissociative amnesia patients have been reported to have high hypnotizability, as measured with standardized hypnosis scales.

Differential Diagnosis

Autobiographical Memory, Nonpathological Amnesia, and Dissociative Amnesia. Nonpathological forms of amnesia have been described, such as infantile and childhood amnesia, amnesia for sleep and dreaming, and hypnotic amnesia. Most forms of dissociative amnesia are thought primarily to involve difficulties with autobiographical memory—rather than implicit or general semantic memory—although rarely procedural memory may be disrupted. Autobiographical memory may have episodic and semantic components (e.g., "last year I bought a new car" vs. "every 5 years I buy a new car," respectively).

Studies of autobiographical memory support the notion that infantile and childhood amnesia can be experimentally documented. However, there are now data on preverbal learning and memory in young children. Amnesia may result in later years owing to the difficulty in translating this preverbal memory into verbal form. However, experiments designed to overcome this factor in young children have shown that children can later report preverbal memories accurately in verbal form under suitable experimental conditions.

Studies demonstrate that normal adult autobiographical memory has a retention gradient for the last 20 to 30 years of the subject's life, often with a subjective sense of wearing away of memories for the past, in contrast to the subjective gaps in memory typical of dissociative amnesia. Elderly patients in autobiographical memory studies have a relative decrease in recent autobiographical memories and a *reminiscence component* for the events of their youth. Older adults have more semantic autobiographical memories and fewer episodic autobiographical memories, anchored in specific time points. Consistent with this, there may be a shift from the first person (field) to the third person (observer) form of autobiographical memory.

Autobiographical memory has several complex aspects, including emotion, mood, and state-based differentials in memory; for example, negative events may be preferentially recalled in depressed mood states. Schema memory effects may result in the inclusion of a memory of

something one has not experienced into memories of a sequence of experienced events. Due to a variety of factors, individuals may develop pseudomemories for events that did not occur. This can include false confessions, experimentally induced "false memories." and failures to recall events that *did* happen (omissions or "false nonmemory"). It may be difficult to differentiate actual memory alteration in these situations from the belief that something did or did not occur.

Autobiographical memory for repeated routine events, such as going to work or school every day, may not be encoded as discrete memories for each day, but rather as broad memory categories for the events that repeatedly reoccur (*regeneric memory*). This is also thought to occur with some traumatic experiences, such as repeated episodes of childhood sexual abuse that happen recurrently over many years.

Clinically, patients with dissociative amnesia may have a variety of autobiographical memory deficits. For example, a patient with global dissociative amnesia after a rape was evaluated with cognitive testing while amnesic and after memory recovery. As compared to an organically impaired control, the dissociative amnesia patient was able to recall memories from various points in the life history without the temporal gradient that characterized the retrograde amnesia of the organic patient. Recall of autobiographical information during dissociative amnesia seemed related to life events with positive effects that were unconnected with the traumatic events precipitating the amnesia. In this patient, implicit autobiographical memory phenomena were also documented. Similar phenomena have been described in posthypnotic amnesia, with implicit demonstration that the memories for which amnesia has been suggested have been encoded and stored, but without being directly accessible for retrieval. Amnesic patients may also have intense reactions to stimuli that are emotionally significant, without being aware of the reason for the reaction or the significance of the stimulus, such as when a patient with PTSD has a flashback without consciously knowing what triggered it and often without clear recall later of the memory being evoked.

Several studies have confirmed the clinical observation that subjects with dissociative amnesia for their life history can demonstrate implicit autobiographical memory while amnesic. When asked to free associate, to imagine, or to make up a story or when exposed to projective tests, patients with dissociative amnesia include in their productions elements of autobiographical information without necessarily being consciously aware of this.

Memory in Normal Aging, Dementia, Delirium, and Amnesia in Cognitive Disorders. The effects of normal aging on memory are complex, depending on which systems are being evaluated. For example, a study that compared aging air traffic controllers with younger controllers found

typical age-related declines in speed or recall of information (names, dates), but aging and younger air traffic controllers showed equivalent ability on other professional tasks, thought to be related to years of experience with these complex activities. Older adults may have better implicit than explicit memory on some tasks, and may be more vulnerable to distraction effects. Exercise and diet may increase hippocampus size in the aging brain, possibly mediated through increased expression of brain-derived neurotrophic factor (BDNF), its receptor and other related pathways.

There is no single test or examination that can establish absolutely whether a memory disorder has a dissociative, neurocognitive, factious, malingered, or mixed etiology. The clinician evaluating the amnesic patient must have a reasonable index of suspicion about any of these. In ambiguous cases, thorough reassessment on an ongoing basis of the clinical situation is helpful. However, most patients with dissociative amnesia present differently from other disorders with memory impairment (Table 20–23). The evaluation of acute dissociative amnesias is described in Table 20–24.

In patients with dementia, amnestic disorders, and delirium, the memory loss for personal information is embedded in a far more extensive set of cognitive, linguistic, attentional, behavioral, and memory problems. Loss of memory for personal identity is usually not found without evidence of marked disturbance in many domains of cognitive function. Confabulation may be present to various degrees and is usually implausible or bizarre. Causes of neurocognitive amnestic disorders include thiamine deficiency (Korsakoff psychosis), cerebral vascular accident (CVA), postoperative syndrome, infection, anoxia, and transient global amnesia. Electroconvulsive therapy (ECT) may also cause a marked temporary amnesia and, occasionally, persistent memory problems. In ECT-related amnesia, however, memory loss for autobiographical experience is unrelated to traumatic or overwhelming experiences and seems to involve multiple types of personal experience, most commonly those occurring just before or during the ECT treatments.

Posttraumatic Amnesia. In posttraumatic amnesia due to brain injury, there is usually a clear-cut physical trauma, a period of unconsciousness or amnesia—or both—and objective clinical evidence of brain injury. In general, the length of the posttraumatic amnesia is a reasonable predictor of cognitive outcome. Retrograde amnesia may also occur. An extensive retrograde amnesia out of proportion to the head injury may warrant an investigation for dissociative factors.

Table 20–23.
Differential Diagnosis of Dissociative Amnesia

Normal Autobiographical Memory Difficulties
Age-related cognitive decline
Nonpathological forms of amnesia
Infantile and childhood amnesia
Amnesia for sleep and dreaming
Hypnotic amnesia
Dementia
Delirium
Amnestic disorders
Neurological disorders with discrete memory loss episodes
Posttraumatic amnesia
Transient global amnesia
Amnesia related to seizure disorders
Substance-related amnesia
Alcohol
Sedative-hypnotics
Anticholinergic agents
Steroids
Marijuana
Narcotic analgesics
Psychedelics
Phencyclidine
Methyldopa (Aldomet)
Hypoglycemic agents
β-Blockers
Lithium carbonate
Many others
Other dissociative disorders
Dissociative fugue subtype
Dissociative identity disorder
Other dissociative disorders
Acute stress disorder
Posttraumatic stress disorder
Somatic symptom disorder
Psychotic episode/disorders
Lack of memory for psychotic episode when returns to nonpsychotic state
May be difficult to differentiate from thought disordered recall in schizophrenia
Mood disorder episode
Lack of memory for aspects of episode of mania when depressed and vice versa or when euthymic
Factitious disorder
Malingering

Table 20–24.
Evaluation of Acute Amnesia

Complete history (to the extent that this can be gathered)
Medical history
Psychiatric history
Trauma history (combat, violence, childhood maltreatment, etc.)
Developmental history
Collateral informants (if available)
Family and concerned others

Medical, military, police, and social service records
Sequential clinical observation
Physical and neurological examination
Baseline physical and laboratory examination
Full mental status examination
Mini-mental state examination
Dementia workup
Electrocardiogram
Electroencephalogram
Telemetry in unusual cases
Brain imaging
Neuropsychological assessment

Seizure Disorders. In most cases of epileptic seizures, the clinical presentation is different from that of dissociative amnesia, with clear-cut ictal events and sequellae. Patients with nonepileptic seizures, however, may resemble epileptic patients in some ways but may also have cognitive dissociative symptoms, such as amnesia, and an antecedent history of psychological trauma. Rarely, patients with recurrent complex partial seizures may present with ongoing bizarre behavior, memory problems, irritability, or violence, making them hard to distinguish from patients with dissociative amnesia. Some epileptic patients have exhibited wandering or semi-purposeful behavior, or both, during seizures or in postictal states, for which there is subsequent amnesia, which may confuse clinicians trying to exclude a dissociative fugue. However, seizure patients in an epileptic (as opposed to dissociative) fugue often exhibit abnormal behavior, including confusion, perseveration, and abnormal or repetitive movements, as well as experiencing an aura, motor abnormalities, stereotyped behavior, perceptual alterations, incontinence, and a postictal state. Stressful life events may be associated with an increase in seizure frequency in some susceptible epileptic patients, and thus this factor alone is insufficient for differential diagnosis. Serial or telemetric EEGs, or both, usually show abnormalities associated with behavioral pathology.

Substance-Related Amnesia. Several substances and intoxicants have been implicated in the production of amnesia. Common offending agents are listed in Table 20–23. In most cases, a careful history from the patient and ancillary sources, sequential clinical observation, and objective testing clarify the substance-related nature of the amnesia. In some instances of pathological intoxication, in which a small amount of alcohol or of another substance produces a major behavioral disinhibition, the alcohol may be producing its effect by facilitating the onset of a dissociative episode in a susceptible individual. This may be analogous to the disinhibition that occurs in a clinical pharmacologically facilitated interview. Subjects may report amnesia for violent or other out-of-character behavior during such an episode.

The most difficult differential diagnostic problem usually involves patients with a history of substance-induced and dissociative memory problems. Some of these patients may minimize dissociative amnesia and vice versa. Clinically, the relative contribution of the substance abuse and the dissociation may only be fully clarified by sequential clinical observation once the patient has achieved sobriety.

WANDERING BEHAVIOR. Individuals with dissociative amnesia may engage in confused wandering during an amnesia episode. To qualify for the dissociative fugue subtype, the individual must exhibit *purposeful* travel away from home or a customary place of daily activities and is usually preoccupied by a single idea that is accompanied by the wish to escape.

Wandering behavior during a variety of general medical conditions, toxic and substance-related disorders, delirium, dementia, and organic amnestic syndromes could theoretically be confused with the dissociative fugue subtype of dissociative amnesia. However, in most cases, the somatic, toxic, neurological, or substance-related disorder can be ruled in by the history, physical examination, laboratory tests, or toxicological and drug screening. Use of alcohol or other substances may be involved in precipitating an episode of dissociative fugue.

Transient Global Amnesia. Transient global amnesia may be mistaken for dissociative amnesia, especially because stressful life events may precede either disorder. However, in transient global amnesia, there is a sudden onset of complete anterograde amnesia and learning abilities; pronounced retrograde amnesia; preservation of memory for personal identity; anxious awareness of memory loss with repeated, often perseverative, questioning; overall normal behavior; lack of gross neurological abnormalities in most cases; and rapid return of baseline cognitive function, with a persistent short retrograde amnesia. The patient is usually over 50 and shows risk factors for cerebrovascular disease, although epilepsy and migraine have been etiologically implicated in some cases.

Dissociative Disorders. As noted previously, dissociative identity disorder patients can present with acute forms of amnesia and fugue episodes. However, these patients are characterized by a plethora of symptoms, only some of which are usually found in patients with dissociative amnesia. Most dissociative identity disorder patients do report multiple forms of complex amnesia, including recurrent blackouts, fugues, unexplained objects in their possession, and fluctuations in skills, habits, and knowledge. Some patients with other dissociative disorder, particularly acute dissociative reactions to stressful events, may develop micro-amnesias for the duration of their acute episodes, in the context of other

dissociative symptoms (e.g., perceptual disturbances). Individuals with more extensive amnestic episodes may qualify for a dissociative amnesia diagnosis (see section on Other Dissociative Disorders).

Acute Stress Disorder and Posttraumatic Stress Disorder. Many dissociative amnesia patients meet full or partial diagnostic criteria for ASD, PTSD, a somatic symptom disorder, or a combination of these. The DSM-5 stipulates that, to be diagnosed, the dissociative amnesia must be distinct from the course of ASD, PTSD, or somatization disorder. In practice, clinical judgment usually determines whether the extent of the amnesia warrants a separate dissociative diagnosis. As depersonalization/derealization disorder symptoms may occur in dissociative amnesia patients, research will be needed to determine the prevalence of the dissociative subtype of PTSD in dissociative amnesia patients who also meet DSM-5 diagnostic criteria for PTSD.

Other Psychiatric Disorders. State-related memory problems can be observed in psychotic disorders, neurocognitive disorders, and cyclic mood disorders. Upon resolution of an acute psychotic episode, many individuals can report lack of recall of their symptoms, behavior, or life events while psychotic. This lack of recall may also be observed in many individuals who recover from an episode of delirium. In contrast, a minority may fully recall their symptoms of psychosis or delirium and become tormented with fear and shame regarding these experiences.

State-related transitions between mania, depression, and euthymia may be accompanied by an autobiographical memory deficit for mood, behavior, perceptions, and sense of self during the different states. Manic individuals may deny recall of depressive episodes and vice versa. During euthymic states, the individual may have difficulty recalling details of being in the other mood states. Wandering and purposeful travel may occur during the manic phase of bipolar disorder or schizoaffective disorder, with lack of recall of behavior that occurred after return to the euthymic or depressed state. In purposeful travel due to mania, however, the patient is usually preoccupied with grandiose ideas and often calls attention to him- or herself owing to inappropriate behavior. Assumption of an alternate identity does not occur, although the subjective differences in sense of self, emotions, and behavior may lead the person to feel like “different people” during the different mood states. Similarly, peripatetic behavior may occur in some patients with schizophrenia. Memory for events during wandering episodes in such patients may be difficult to ascertain owing to the patient’s thought disorder. However, dissociative fugue patients do not demonstrate a psychotic thought disorder or other symptoms of psychosis.

Malingering and Factitious Amnesia. Feigned amnesia is more common

in patients presenting with the acute, classic forms of dissociative amnesia. However, in one recent forensic case, an adult attempting to sue an admitted abuser using the delayed discovery rule was shown to have falsified delayed recall for trauma in an attempt to overcome the statute of limitations. Investigations showed that the patient had discussed the abuse with others on many occasions before the purported delayed recall. On the other hand, some patients may have secondary amnesia for having remembered and discussed traumatic experiences in the past.

Malingering of dissociative fugue may occur in individuals who are attempting to flee legal, financial, or personal difficulties, and in soldiers attempting to avoid combat or unpleasant military duties. These precipitating factors may be present as well in bona fide dissociative fugue. However, many malingerers confess spontaneously or when confronted. In the forensic context, the examiner should always give careful consideration to the diagnosis of malingering when fugue is reported.

As in other dissociative disorders, there is no test, battery of tests, or set of procedures that invariably distinguish true dissociative symptoms from those that are malingered. Malingerers have been noted to continue their deception even during hypnotically or barbiturate-facilitated interviews. As noted previously, many of the classical cases occurred in a clinical context of financial, sexual, and legal problems or in soldiers who wished to escape from combat.

On the other hand, in clinical case reports, many malingerers have quickly confessed their deceptions spontaneously or when confronted by an examiner. In these nonforensic reports, the malingered amnesiacs were frequently pathetic individuals whose deception was transparent. Nevertheless, it can often be unclear where the conscious deception began and the unconscious defenses ended.

In the current clinical environment, a patient who presents to psychiatric attention asking to recover repressed memories as a chief complaint is most likely to have a factitious disorder, including the imitative form of dissociative disorder described in prior sections. Most of these individuals actually do not describe bona fide amnesia when carefully questioned; but they are often insistent that they must have been abused in childhood to explain their unhappiness or life dysfunction. As noted above, most dissociative individuals are fearful, ashamed, and avoidant of recall of traumatic memories, and minimize and deny the dissociative amnesia. Memory recall in dissociative amnesia does not present as typical episodic, autobiographical memory recall. Rather, these usually are experienced as disturbing intrusive flashbacks with alarming and vivid perceptual, somatoform, and emotional experiences, followed by attempts to "re-dissociate" the material.

Clinical Course and Prognosis

Little is known about the clinical course of dissociative amnesia. Acute dissociative amnesia frequently resolves spontaneously once the person is removed to safety from traumatic or overwhelming circumstances. At the other extreme, there are patients who develop chronic forms of generalized, continuous, or severe localized amnesia who are profoundly disabled and require marked social support, such as nursing home placement or intensive family caretaking (see case example above).

Individuals with nonclassic dissociative amnesia may have a persistent form of the disorder that does not cause them overt distress, and, thus, may not seek clinical attention. Some of these individuals can be understood as being in an episode of PTSD characterized primarily by avoidant symptoms. In some cases, a later traumatic event or stress—even a relatively minor one—precipitates a florid episode of PTSD with alternations of re-experiencing and avoidant or amnestic symptoms.

Studies have shown that recall of previously dissociated memory frequently occurs outside of therapeutic settings, triggered by a variety of stimuli, including the person's children reaching the age at which the person was abused, the person's child being abused, media accounts of trauma or abuse, death of an abusive parent, a variety of sensory cues, and feeling safe in one's life situation, among others.

Treatment

The diagnosis and treatment of dissociative amnesia became controversial in the 1990s due to disputes over delayed recall of childhood sexual abuse. Prior to this, dissociative (psychogenic) amnesia and delayed recall for combat or other forms of trauma had not been considered controversial, although some authors from the World War I and II eras opined that many of these wartime cases were malingered. As described above, controversy exists not only for the existence of dissociative amnesia, but also for how to treat it.

Critics of the dissociative amnesia construct have invented the term *recovered memory therapy* (RMT) to characterize treatment in which clinicians are thought to make aggressive efforts to have patients recall allegedly forgotten traumas as the central focus of treatment. RMT does not represent a known school of therapy or of scholarly research. It may more accurately describe individual fringe practitioners or media or layperson views of trauma treatment. As noted in prior sections, reviews of the literature on trauma treatment report that the predominant model for work on traumatic memories involves a focus on symptom stabilization and safety, integration of posttraumatic memories, beliefs, cognitions, affects, somatic representations, and object relations, not on memory recall per se.

Phase-Oriented Treatment. As for dissociative identity disorder, phase-oriented treatment is the current standard of care for the treatment of

dissociative amnesia, although there are no systematic studies with large cohorts of dissociative amnesia patients. Treatment of the acute classic dissociative amnesia patient follows a similar phasic model. However, here, memory recall is a central issue, because loss of memory for personal identity and large gaps in current autobiographical memory are acutely disabling symptoms that require relatively rapid intervention.

In general, as described in the section on dissociative identity disorder, three basic phases are recognized. This structure is heuristic to some extent, because aspects of each phase may be worked on during another. First, there is a stabilization phase, with a focus on safety, symptom control, containment of affects and impulses, and education about trauma treatment. Once adequate personal safety and clinical stability are established, if indicated, the individual may engage in a second phase that focuses on the integration of traumatic material in greater depth. This processing may involve attempts to overcome persistent amnesia and to resolve material that is not dissociated or less completely dissociated. Finally, there is a third phase of *resolution* or *reintegration,* in which the traumatized person is reconnected to ordinary life. In this phase, the focus is less on the trauma per se and more on the development of a renewed, reinvigorated life that overcomes the lack of freedom imposed by symptoms of the trauma disorder and the dominance of the person's psychology by issues related to traumatization.

In general, studies of treatment outcome in survivors of rape and childhood sexual abuse have shown better outcome when patients directly discuss trauma material in the context of a carefully designed, phasically structured individual or group psychotherapy. Nonetheless, it is a matter of clinical judgment and the patient's individual decision whether the patient has achieved sufficient stability and has sufficient ego strength to move from the stabilization phase of treatment to the phase of memory integration. In this regard, in clinical case series, serious suicide attempts, and even completed suicide, have been reported in patients with dissociative amnesia when amnestic barriers are removed precipitously with inadequate stabilization.

Safety in Acute Dissociative Amnesia. In the case of the patient with an ASD primarily characterized or accompanied by dissociative amnesia, the establishment of the person's physical safety is the first concern. This involves removing the individual from the traumatizing environment (e.g., acute combat), evaluating and treating medical problems, and providing shelter, food, and sleep. The paradigm of *psychological first aid* may be helpful in the initial stabilization of the acutely traumatized, amnesic individual. Sedative medications, such as the benzodiazepines, may be indicated to assist the patient with sleep. Patients with acute amnesia for personal identity or life circumstances, or both, frequently have a relatively

rapid spontaneous remission of symptoms once brought to the safety of the hospital or other protected environment. If immediate spontaneous remission does not occur in patients with acute amnesia after removal from traumatic environments, symptoms may abate later simply in the course of the clinician's taking a psychiatric history or merely with suggestions and reassurance.

Safety in Nonclassic Dissociative Amnesia. Patients with nonclassic, covert amnesia presentations generally should be managed within the framework of a longer-term psychotherapy directed at resolution of the complex psychological sequelae of the events producing the amnesia, usually severe traumatization due to childhood abuse, combat, rape, domestic violence, or other forms of adult victimization. Here, too, the first tasks of treatment are restoration of the patient's physical well-being and safety and establishment of a working alliance. The clinician must be prepared to intervene actively if the patient is acutely dangerous to self or others or is abusing substances in an uncontrollable way. Typically, these patients' difficulties involve suicide attempts, self-mutilation, eating disorders, alcohol or substance abuse, involvement in abusive or destructive relationships, episodes of rage or violence, abuse of the individual's own children or family members, and lack of adequate food, clothing, or shelter. Hospitalization may be necessary for stabilization, as may referral to specialty resources, such as treatment for substance abuse or eating disorders.

In individuals with severe intrusive PTSD symptoms alternating with amnesia, containment and management of intrusive recollections rather than detailed processing of the traumatic material are usually the main goals of the stabilization phase of treatment. This may be accomplished with supportive psychotherapy, pharmacotherapy, imagery or hypnotic techniques for containment and symptom control, and/or dialectical behavioral and cognitive therapy techniques, or a combination of these. Although there is no pharmacological agent that specifically targets dissociative amnesia, specific psychopharmacological treatment of the patient's PTSD, mood, impulsive, psychotic, obsessive-compulsive, or anxiety symptoms with medications may help stabilize severe symptoms that prevent the patient's meaningful participation in psychotherapy.

Long-term treatment for these patients is focused on the manifold dimensions of dysfunction that chronic trauma engenders. These include problems with mood, anxiety, and impulse regulation; disordered attachment schemas engendering troubled relationships and problems with interpersonal boundaries; spontaneous altered states of consciousness and dissociation; memory problems; cognitive distortions and disordered meaning systems; perceptual abnormalities; problems with the sense of self and body image; somatoform symptoms; and self-destructiveness.

Treatment of Amnesia. Treatment of patients with nonclassic forms of dissociative amnesia necessitates that the clinician familiarize him- or herself with the current controversies about trauma and memory to provide adequate informed consent to the patient. Factors that usually contraindicate intensive memory integration work are listed in Table 20–25 (see also Table 20–21 in the section on dissociative identity disorder).

Free Recall. Patients with the acute and chronic forms of dissociative amnesia may respond well to free-recall strategies in which they allow memory material to enter into consciousness. The clinician is supportive and nondirective but focuses on reluctance or resistance to allowing free recall to take place. Classic free-association suggestions are often the most helpful in understanding factors that interfere with recall and in allowing recall to occur at a pace that the patient can tolerate. Most clinicians think that memory accuracy is improved if the clinician asks nonleading questions, whether in a free-recall situation or with methods that facilitate memory recall. Distress at recall may be titrated by techniques for distancing and fractioning emotion. For example, the patient may be asked to give a brief "headline" or two, as if reading a newspaper, but not to go into more detail (e.g., reading the newspaper article in full) until more able to do so.

Transference Interpretations. Studies of transference in patients with combat-related PTSD and severe dissociative disorders indicate that a traumatic transference is usually the predominant initial transference theme in these individuals. This is a set of unconscious perceptions of the clinician, based on relationships formed under traumatic circumstances. For instance, the therapist becomes the buddy who was killed next to the patient in battle, the persecutory abusive parent, the incompetent officer who sent the patient into battlefield disaster, the neglectful relative unconcerned about the patient's abuse, or the patient him- or herself who is subject to the sadistic, abusive behaviors of an implacable other. Identification of the overt patterns of traumatic transference observed by the therapist may be another route to undoing amnesia. A recent study of corroborated trauma memory found that transference-based recall was more accurate than other forms of facilitated recall.

Table 20–25.
Contraindications to the Memory—Integration Phase of Amnesia Treatment (see also Table 20–21)

Has not achieved safety from high-risk behaviors
Current substance abuse or dependence
Has not achieved symptom stabilization
Severe, uncontrolled, intrusive posttraumatic stress disorder symptoms

Severe, uncontrolled dissociative episodes
Dysregulated mood symptoms
Dysregulated anxiety symptoms
Inadequate therapeutic alliance
Current or ongoing abusive relationship
Acute life crisis or times of life transition (divorce, job change, etc.)
Severe physical illness or infirmity
Severe personality disorder symptoms, regression, psychosis
Current involvement in litigation
Impending absence of therapist

Adapted from Loewenstein RJ. Dissociative amnesia and dissociative fugue. In: Gabbard GO, ed. *Treatment of Psychiatric Disorders*. 3rd ed. Vol 2. Washington, DC: American Psychiatric Press; 2001:1633.

Cognitive Therapy. Cognitive therapy may have specific benefits for individuals with trauma disorders. Identifying the specific cognitive distortions that are based in the trauma may also provide an entrée into autobiographical memory for which the patient experiences amnesia. As the patient is more able to correct cognitive distortions, particularly about the meaning of prior trauma, more detailed recall of traumatic events may occur, although the potential for flooding is a significant concern.

Facilitated Recall of Trauma Material. A hierarchy of techniques for facilitation of recall has been described for dissociative amnesia. Research suggests that each of these has the potential to successively increase error rates in recall. The techniques include: (1) context reinstatement (i.e., attempts to focus the patient on time periods for which there is amnesia); (2) state-dependent recall (i.e., intensification and focus on affects or somatic sensations that appear to be related to trauma, such as terror, horror, confusion, rage, or sudden nausea or dizziness); and (3) specialized adjunctive techniques (e.g., hypnosis, relaxation, imagery, pharmacologically assisted interviews). Each of these techniques may facilitate recall of dissociated memories in classic and nonclassic forms of dissociative amnesia. When implementing them, the clinician should use nonleading, nonsuggestive questions to minimize concern about inaccurate or confabulated recall.

Hypnosis for Amnesia. Hypnosis has frequently played an important adjunctive role in the treatment of individuals with dissociative amnesia and dissociative fugue. Hypnosis is not a treatment in itself; rather, it is a set of adjunctive techniques that facilitate certain psychotherapeutic goals. The construct of hypnosis encompasses a wide and complex domain of phenomena and responses that have been applied to many therapies and diverse individuals. Concern about inaccuracies in hypnotically facilitated recall is based on research on memory for nontraumatic information showing that hypnosis increases a person's confidence in what is recalled,

even if it is erroneous. However, critical reviews have noted the complexity of this research problem, the inconsistency of findings among studies, and the many variables related to hypnosis and nonhypnotic factors that appear to influence this and related phenomena. These reviews have emphasized that memory confabulation is related to techniques of social influence on suggestible subjects rather than to hypnosis per se.

The use of hypnosis for memory recall in no way ensures the veracity, or lack of veracity, of the information produced. The clinician should be well aware of the reconstructive nature of memory and should avoid suggestions to the contrary. Thus, in addition to obtaining the patient's informed consent for the use of hypnosis, it is currently considered prudent to also obtain informed consent concerning the reliability or possible lack thereof for autobiographical information revealed during hypnosis. Studies also suggest that educating patients about the controversies regarding hypnosis and memory and informing them of the need to critically evaluate any memory material that they recall during hypnosis, or at any other time in treatment, reduces the likelihood of credulous acceptance of potentially inaccurate information.

Also, clinicians should be aware that, in some states of the United States, individuals who have been exposed to therapeutic hypnosis might be enjoined from giving testimony as a witness in court. This may be so even if the hypnosis had little relationship to the events at legal issue. The clinician should attempt to discuss fully these issues with the patient and his or her legal counsel.

Finally, the use of hypnosis as part of a *forensic* examination should proceed only after obtaining consent from the individual (e.g., plaintiff, defendant, witness) after he or she has consulted with relevant legal authorities and attorneys involved in criminal or civil litigation. Training in forensic hypnosis is mandatory in this context, as there are specific guidelines about use of hypnosis in the forensic setting. Should hypnosis be used in this way, electronic recording of all hypnosis sessions, preferably on video recording, should be used. Failure to follow guidelines for forensic hypnosis may result in invalidation of the expert's testimony.

Hypnosis can be used in a number of different ways in the treatment of dissociative amnesia. In particular, hypnotic interventions can contain, modulate, and titrate the intensity of symptoms; facilitate controlled recall of dissociated memories; provide support and ego strengthening for the patient; and, finally, promote working through and integration of dissociated material. In addition, the patient can be taught self-hypnosis to apply containment and calming techniques in his or her everyday life. Successful use of containment techniques, whether hypnotically facilitated or not, also increases the patient's sense that he or she can more effectively be in control of alternations between intrusive symptoms and amnesia.

In the acutely amnestic patient, a few sessions to help the patient

experience trance successfully and to explore containment and distancing techniques may be sufficient to allow focused hypnotic work on the material for which the patient is amnestic. Clinicians since World War I have recognized how valuable it is for patients to process dissociated material repeatedly over several sessions, often at different levels of affective intensity, to complete the process of integrating the material. In cases of acute amnesia, the first session for memory processing may necessarily be explorative. Here, again, the initial goal is to gain an overview of the information for which the patient is amnestic. As in cases of chronic amnesia, subsequent sessions then focus on reworking the material in greater detail. However, in nonclassic cases of severe, longstanding, posttraumatic disorders, the amnesia itself should be addressed only gradually, in the context of a longer-term psychotherapy in which life stability and function are the basic foci.

It is useful to try to account systematically for different dimensions of the traumatic experience—sensory, affective, cognitive, and behavioral—so as to ensure that all key components have been identified and reconstructed. It is likewise useful to attempt to account systematically for a variety of dysphoric affects that are commonly experienced during traumatic experiences: despair, sorrow, grief, horror, shame, helplessness, rage, guilt, confusion, anguish, etc. Not all of these may be present in a given patient; however, it is useful for the clinician to keep track of which affects the patient most readily describes and which seem less available. Inquiry about these other affects may be quite helpful in resolving the amnesia. In particular, shame, horror, helplessness, and overwhelming confusion are emotions that patients may have the most trouble identifying without assistance from the therapist.

The treatment process is similar when the acute amnesia results not from traumatic experiences, but rather from aspects of the person's current behavior that are in conflict with deeply held moral values or behavioral standards. Treatment in such instances seeks to help the patient tolerate these affects and conflicts without resorting to dissociative defenses. Frequently, these patients describe a developmental history characterized by a rigid family moral code enforced with harsh physical discipline. These experiences often appear to be the traumatic underpinning of the dissociative diathesis. Psychotherapy and hypnotherapy in these patients are directed in part at reducing the patient's brutally unreasonable, and often conflicting, expectations of him- or herself and the guilt and shame that so often accompany acute dissociation. Furthermore, these patients may have great difficulty tolerating anger or violent impulses, because such affects tend to trigger recall of earlier experiences with physical abuse or similar traumas, often as a flashback experience. Thus, psychotherapy may serve not only to focus on the conflicts that led to the acute amnesia, but also to explicate and work through the patient's thoughts, feelings, and self-

perceptions related to the antecedent traumatic events.

Somatic Therapies. There is no known pharmacotherapy for dissociative amnesia other than pharmacologically facilitated interviews. A variety of agents have been used for this purpose, including sodium amobarbital, thiopental (Pentothal), oral benzodiazepines, and amphetamines. To date, no controlled studies have assessed the efficacy of any of these agents in comparison with one another or with other treatment methods. A single placebo-controlled study of a barbiturate-facilitated interview did not find superiority of sodium amobarbital over placebo in producing more clinically useful information. In more systematic World War II studies of barbiturate-facilitated interviews for amnesia and conversion reactions, this treatment was described as leading to more rapid recovery, especially for amnesia, although little overall difference was found in recovery in comparison to subjects treated with psychotherapy alone or with hypnotherapy.

Amobarbital narcosynthesis is a term devised to underscore the need for the patient to process material uncovered in a pharmacologically facilitated interview in his or her usual conscious state. Pharmacologically facilitated interviews are used primarily with acute amnesias and conversion reactions, among other indications, in general hospital medical and psychiatric services. This procedure is also occasionally useful in refractory cases of chronic dissociative amnesia when patients are unresponsive to other interventions. As noted above, The Joint Commission now considers pharmacologically facilitated interviews to be conscious sedation, requiring the presence of an anesthesiologist. This procedure must be performed in settings in which resuscitation equipment is available in case of respiratory arrest, a possible, albeit rare, complication. The interview usually is audiotaped or videotaped to replay for the patient, because amnesia generally is present for the material produced. Although in some cases repeated procedures may be helpful, in others, repeated procedures have led to the patient's dependence on pharmacologically facilitated interviews.

The current controversies over delayed recall for traumatic experiences have also focused on the use of pharmacologically facilitated interviews for patients with reported amnesia for childhood maltreatment. No modern systematic data exist on memory fallibility or accuracy in individuals undergoing pharmacologically facilitated interviews. As discussed previously, pharmacologically facilitated interviews were used extensively in World War II for the treatment of combat-related dissociation. Although wartime memories recovered with pharmacologically facilitated interviews were generally considered to be accurate, some subjects were reported to dissemble completely or withhold crucial information, or both, despite barbiturate treatment.

Given the current controversies, the clinician should obtain informed consent focusing on the nature of memory for pharmacologically facilitated interviews for amnesia symptoms as for patients considering hypnosis. The clinician should emphasize that these drugs are not a truth serum; whatever apparently new information emerges under the drug condition should be regarded no differently with respect to accuracy than any other material that emerges over the course of treatment.

Group Psychotherapy for Amnesia. During World War II, group psychotherapy and group hypnotherapy were among the treatments given to promote recovery in traumatic war-related amnesia. Highly supportive, structured, reassuring, and re-educative approaches were often used by therapists in such cases in an attempt to accomplish return of the patient to functional status and to prevent chronic disability.

Time-limited and longer-term group psychotherapies have been reported to be helpful for combat veterans with PTSD and for survivors of childhood abuse. During group sessions, some authors report that patients recover memories for which they have had amnesia. Supportive interventions by the group members or the group therapist, or both, may facilitate integration and mastery of the dissociated material. On the other hand, concern has been raised about memory contamination in such group therapy settings, as it has for patients' involvement in 12-step or self-help groups for trauma survivors.

TREATMENT OF DISSOCIATIVE FUGUE SUBTYPE. Treatment for dissociative fugue follows the basic structure described above for treatment of dissociative amnesia. Once stabilization is achieved, subsequent therapy is focused on helping the patient regain memory for identity, life circumstances, and personal history. Some dissociative fugue patients may resist uncovering their actual identity even with hypnosis or pharmacologically assisted interviews. Appeals through the local (or even regional) media may not alert the patient's concerned others if the patient has wandered a substantial distance from home. In one case, a patient was asked to randomly select numbers on a phone keypad. The patient tapped a phone number—apparently outside of conscious awareness—that allowed the treating clinician to find the patient's family hundreds of miles away.

Family, sexual, occupational, or legal problems that were part of the original matrix that generated the fugue episode may be substantially exacerbated by the time the patient's original identity and life situation are detected. Thus, family treatment and social service interventions may be necessary to help resolve such complex difficulties. Also, the patient may display other extreme or bizarre dissociative symptoms, such as the Ganser symptom of approximate answers, or attempt to engage in another fugue to escape the situation.

DEPERSONALIZATION/DEREALIZATION DISORDER

As described in the section on Nosology, the DSM-5 Work Group reviewed the relationship between depersonalization/derealization symptoms and psychiatric disorders, especially the dissociative disorders. This review strongly supported continued inclusion of depersonalization/derealization disorder as a dissociative disorder in DSM-5. There was one major change from DSM-IV-TR depersonalization disorder to DSM-5 depersonalization/derealization disorder. No distinction was found between patients with predominant depersonalization as compared with those with predominant derealization, thus supporting the ICD-10 conceptualization of either symptom as characteristic. Accordingly, the disorder was renamed *depersonalization/derealization disorder* in DSM-5.

For many years, the ubiquity of depersonalization/derealization as a psychiatric symptom obscured its broader recognition as a disorder. Consequently, patients with chronic and primary depersonalization/derealization are still misdiagnosed as having solely mood or anxiety disorders. These disorders may be comorbid with chronic depersonalization/derealization, and/or may develop in response to the untreated depersonalization/derealization/derealization disorder symptoms. Recent systematic research by Daphne Simeon and Anthony David on two large cohorts of well-characterized depersonalization/derealization/derealization disorder patients yielded strikingly consistent findings in the United States and Europe. Clinical features, course, prognosis, and cognitive and neurobiological correlates clearly distinguish this disorder from other psychiatric disorders with symptoms of depersonalization/derealization. Also, valid and reliable instruments exist for screening and diagnosis, such as the Cambridge Depersonalization/Derealization Scale developed by Sierra and Berrios (see section on Nosology and Table 20–2).

Etiology

Several models have been proposed to explain depersonalization/derealization disorder. A schematic integration of these is presented in Figure 20–2.

Psychodynamic. Traditional psychodynamic formulations have emphasized the disintegration of the ego or have viewed depersonalization/derealization as an effective response in defense of the ego. These explanations stress the role of overwhelming painful experiences or conflictual impulses as triggering events. The high rates of depersonalization/derealization symptoms in normal adolescents and in patients conceptualized as having borderline or narcissistic personality organizations are cited as evidence that ego immaturity or ego deficits

predispose to these conditions. More recently, attention has been drawn to the similarities between depersonalization/derealization and obsessive-compulsive symptoms. Depersonalization/derealization disorder patients often display obsessive-like behaviors with respect to their symptoms. The split between an observing and a participating self is likened to the division between intellect and emotion in obsessive patients. Both groups respond to serotonin reuptake inhibitors, although the therapeutic response for depersonalization/derealization disorder patients is usually much less robust.

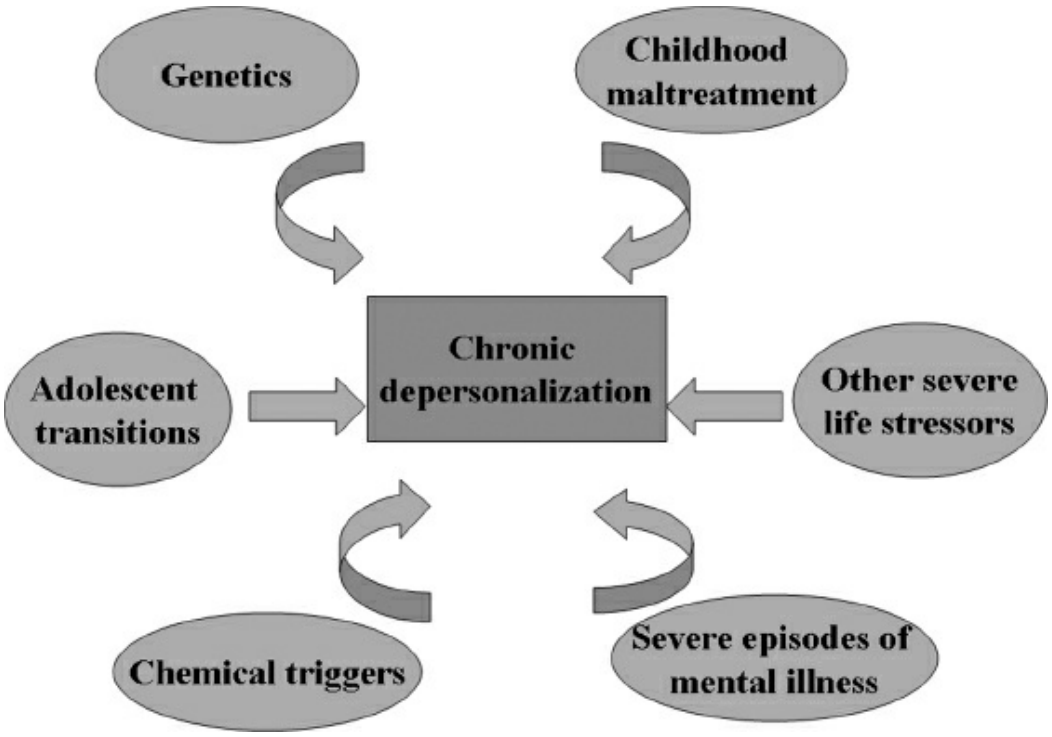

FIGURE 20–2. Schematic integration of proposed models to explain depersonalization disorder.

Traumatic Stress. One-third to one-half of patients included in case series of clinical depersonalization/derealization report past trauma. Depersonalization/derealization symptoms in accident victims and military recruits undergoing high-stress training have been described in prior sections. A controlled clinical study found significantly more childhood trauma, especially emotional abuse, in well-diagnosed depersonalization/derealization disorder patients compared with normal controls. In support of the power of emotional abuse as a traumatic stressor, and its link to dissociation, Martin H. Teicher and colleagues recently showed that parental verbal/emotional abuse was "a potent form of maltreatment" in a community sample of 554 young adults. Verbal abuse experiences resulted in moderate to large increases in measures of dissociation, anxiety, depression, anger-hostility, somatoform symptoms,

and symptoms of "limbic irritability," such as paroxysmal somatic disturbances, brief hallucinatory events, automatisms, and dissociative experiences.

In approximately 20 percent of chronic depersonalization/derealization disorder patients, there was a first-degree relative with a severe psychotic illness, either schizophrenia or bipolar disorder. It was hypothesized that the chronic fear engendered by the relationship with the psychotic relative was etiological in the subsequent development of the depersonalization/derealization disorder. For example, one patient reported that, throughout her childhood, her father and older brother left her alone to handle her violent, schizophrenic mother whenever the mother had psychotic episodes. The patient recalled waiting in a state of terror and dread until the emergency workers came and hospitalized her mother.

DEPERSONALIZATION/DEREALIZATION DISORDER (DDPD) CASE 1

Ms. R a 27-year-old, unmarried, graduate student with a master's degree in biology came to clinical attention for intermittent episodes of "standing back," usually associated with anxiety-provoking social situations. When asked about a recent episode, she described presenting in a seminar course. "All of a sudden, I was talking, but it didn't feel like it was me talking. It was very disconcerting. I had this feeling, 'Who's doing the talking?' I felt like I was just watching. Watching someone else talk. Listening to words come out of my mouth, but I wasn't saying them. It wasn't me. It went on for a while. I was calm, even sort of peaceful. It was as if I was very far away. In the back of the room somewhere—just watching myself. But the person talking didn't even seem like me really. It was like I was watching someone else." The feeling lasted the rest of that day and persisted into the next, during which time it gradually dissipated. She thought that she remembered having similar experiences during high school but was certain that they occurred at least once a year during college and graduate school. Although she said that she usually felt detached and sometimes peaceful during the experience, she was upset at the thought that she would likely have more episodes. She complained that the sudden onset, the eeriness of seemingly watching an almost unrecognizable version of herself from a distance, and the sense of not "being in the world" were almost unbearable in retrospect. As a child, Ms. R reported frequent intense anxiety due to overhearing or witnessing the frequent violent arguments and periodic physical fights between her parents. She remembered lying awake in bed, listening to her parents, imagining the terrible things that were occurring, or about to occur, between them. In addition, the family was subject to many unpredictable dislocations and moves owing to the patient's father's intermittent difficulties with finances and employment. The patient's anxieties did not abate when the parents divorced when she was a late adolescent. Her father moved away and had little further contact with her. Her relationship with her mother became increasingly angry, critical, and contentious. She was unsure if she experienced depersonalization during childhood while listening to her parents' fights.

In general, the trauma reported by individuals with depersonalization/derealization is less severe than that typically reported by other dissociative disorder patients. A large general population study found that individuals with chronic pain were three times more likely to have episodes of depersonalization/derealization, but there was only a weakly significant association with dangerous or upsetting experiences. Many individuals with depersonalization/derealization disorder do not identify a traumatic antecedent and report no clear precipitant to the onset

of their disorder.

On the other hand, nontraumatic stressors—such as severe interpersonal, financial, or occupational losses—have been associated with the onset or exacerbation of depersonalization/derealization disorder. In addition, chemical stressors, most commonly marijuana and the hallucinogens, can sometimes precipitate chronic depersonalization/derealization. These individuals can be conceptualized as having a neurobiological or genetic vulnerability to chronic depersonalization/derealization after drug use, although early life trauma can also play a role.

The following case example describes these complex interactions.

DEPERSONALIZATION/DEREALIZATION DISORDER (DDPD) CASE 2

Mr. A, a 42-year-old actor, was referred for psychiatric evaluation. Since his late teens, he had abused alcohol and a variety of drugs, usually in binges lasting several weeks. After a binge on alcohol, marijuana, and inhaled and smoked cocaine, he awakened with an intense, highly distressing depersonalization experience. He reported feeling completely alien and disconnected from himself, his feelings, and his awareness of the environment. He reported estrangement from his body, not experiencing it as connected to him, feeling as if his limbs were either too far away from him or too close. He reported that the world seemed flat and unreal, almost as if it were a painting. He tried to read scripts that he had been given to prepare for auditions, but felt that the words had no connection with him, and that, when he tried reading aloud, his voice sounded alien, "tinny," far away, and unfamiliar. He denied being fully out of his body, observing himself as if he were another person. However, he reported being "just beside myself, not really in myself, like to one side. It's terrible; I can't get back to me. I don't feel connected to who I am. I know I am feeling things, but I don't know what they are. All I feel is desperate to get back to how I used to be." He described his father as a dictatorial alcoholic who beat him and emotionally berated him throughout growing up, as well as terrorizing the family, including repeatedly beating and berating the patient's mother in front of the family. He reported that several family members have been treated for PTSD, dissociative, and/or conversion (somatic symptom) disorders. He was referred for a complete physical, laboratory, and neurological work up: all were essentially within normal limits. Due to the symptoms, he stopped using alcohol and drugs completely. However, several years later, he has remained chronically depersonalized. He has had a limited response to treatment with most categories of psychotropic medications, and psychotherapy. He has only been able to work in a very limited way.

Cognitive Theories. In the last decade, increasing attention has been drawn to cognitive and behavioral aspects of chronic depersonalization/derealization, essentially positing that an initial, relatively benign, and possibly transient dissociative response is reinforced, maintained, and exacerbated by a vicious cycle of dysfunctional cognitions and behaviors. E.C.M. Hunter and colleagues in the United Kingdom have put forth such a cognitive-behavioral model, proposing that an initial trigger (e.g., trauma, anxiety, depression, stress, fatigue, intoxication) can induce transient symptoms of depersonalization/derealization, which are subsequently cognitively processed by an individual as either situational or catastrophic. If the attributions are situational, and therefore more benign, the depersonalization/derealization symptoms will tend to fade as the

situational factors alleviate. However, if the attributions are catastrophic, they may evoke overwhelming fears such as going mad, losing control, becoming invisible, or having a permanent brain dysfunction. In turn, these fears may lead to an increase in anxiety coupled with a paradoxical decrease in arousal, resulting in increased intensity of the depersonalization/derealization symptoms as the individual enters the maintenance phase. During this phase, individuals may start to avoid situations that they associate with symptom provocation; become preoccupied with safety behaviors (such as acting "normal"); develop cognitive biases resulting in symptom over-monitoring; and a reduced threshold for the perception of threat. Over time, these maintenance factors are thought to perpetuate or exacerbate the symptoms.

Diagnosis and Clinical Features

In clinical samples, the disorder is equally common in men and women. Under-diagnosis has a number of determinants. Patients experiencing depersonalization/derealization symptoms have great difficulty putting their impalpable experiences into words, and may have marked problems with alexithymia. Simultaneously, they fear that they may be taken for "crazy" if they **do** describe their symptoms, and are reluctant to speak about them. They may describe their severe suffering in ways that do not adequately convey their distress to an examiner; for example repeatedly stating that, "I feel dead," "Nothing seems real," or "I'm standing outside of myself." Depersonalization/derealization disorder patients typically endure an intensely unpleasant, and often disabling, subjective experience. Many say that they would gladly exchange the depersonalization/derealization for physical pain that would at least reconnect them with their body. Others say they do not contemplate dying because they already are dead. Some patients may appear remarkably unemotional and robotic, due to profound dissociation of affects. As the patient's complaints do not match their affective presentation, clinicians may not take the severity of this disorder as seriously as they should.

There are a number of distinct components to the experiences of depersonalization and derealization. *Depersonalization* may include a sense of: bodily changes, duality of self as observer and actor, being cut off from others, and being cut off from one's own emotions. The person may suffer from teleopsia, pelopsia, metamorphopsia, micropsia and/or macropsia, with the sense that the body is changing in size and shape; the individual may experience disintegration of time sense with little ability to assess how recently or remotely events occurred; this may include déjà vu, déjà veçu, and jamais vu; the person may describe numbing of emotions and perceptions, including the inability to perceive hunger or thirst; and feeling robotic, detached, and/or without thoughts or feelings. The individual may report out of body experiences, or, as in the case example

above, a highly dysphoric feeling of partial disconnection from one's self. On the other hand, in *derealization,* the world appears strange, foreign, or dreamlike. Objects may appear as if viewed from a great distance and as if they are two-dimensional, without depth or substance. Sounds come from a distance, muffled and distorted. Objects feel strange to the touch. Colors dim and lose their vitality. The faces of others change, becoming unfamiliar and frightening.

Cross-Cultural Aspects of Depersonalization/Derealization Disorder

The anthropologist Richard Castillo has described a set of symptoms of depersonalization/derealization that potentially broadens the cross-cultural validity of depersonalization/derealization disorder. Castillo's cluster of symptoms is similar to the more recent data-derived factor analytic models in Western samples: persistent or recurrent experiences of depersonalization and/or derealization in which familiar experiences of the self and environment become distorted, potentially resulting in physical and emotional numbing; alterations in body image; experiences of being an outside observer of one's mental processes or body; feeling like an automaton; time disturbances; and feeling as if one's body or environment is "unreal" or dreamlike. Visual distortions in depersonalization/derealization included micropsia, macropsia, teleopsia, pelopsia, metamorphopsia, and loss of stereoscopic depth perception. Relevant to diagnosis, the depersonalization/ derealization experience does not occur exclusively during a culturally approved religious or ritual activity and is not considered normal by the cultural standards of the patient's own social group(s).

Differential Diagnosis

The variety of conditions associated with depersonalization/derealization complicate the differential diagnosis of depersonalization/derealization disorder. Depersonalization/derealization may result from a medical or neurological condition, particularly seizure disorders, from intoxication or withdrawal from illicit drugs, or as a side effect of medications. Depersonalization/derealization may be associated with panic attacks, phobias, PTSD, ASD, depression, schizophrenia, or another dissociative disorder. A thorough medical and neurological assessment may be necessary to evaluate the presence of depersonalization/ derealization, including standard laboratory tests, an EEG to rule out partial complex seizures, urine and/or serum drug screens, and brain imaging if symptoms raise the question of structural brain abnormalities. Drug-related depersonalization/derealization is typically transient, but persistent depersonalization/derealization may follow an episode of intoxication with a variety of substances, notably marijuana, hallucinogens, MDMA

(Ecstasy), or ketamine (see Depersonalization/Derealization Disorder [DDPD] Case 2).

A range of neurological conditions, including seizure disorders, brain tumors, postconcussive syndrome, metabolic abnormalities, migraine, vertigo, and Ménière disease, may also need to be ruled out. Depersonalization/derealization caused by organic conditions tend to be primarily sensory without the elaborated descriptions and personalized meanings common to psychiatric etiologies.

In order to establish the psychiatric differential, a thorough and detailed psychiatric history needs to be obtained, especially in the presence of comorbid mood and anxiety disorders. Although the latter have been reported to have a lifetime comorbidity of up to 90 percent with depersonalization/derealization disorder, current diagnostic criteria require that the chronic depersonalization/derealization not occur exclusively in the presence of these other disorders. It is common to obtain a history whereby an earlier and frightening episode of severe anxiety or depression—which conceivably may have acted as an internal stressor and destabilized a person's known sense of self—may precipitate chronic depersonalization/derealization. Subsequently, the latter continues even after the initial anxiety or mood episode remits. In another common scenario, depersonalization/derealization disorder tends to be more refractory, often continuous for periods of many years or decades, while the comorbid disorders go into periods of clear remission, whether with treatment or spontaneously. Conversely, severe chronically ill depersonalization/derealization disorder patients may become unremittingly depressed in response to their primary dissociative disorder leading to mis-diagnosis of a primary affective disorder.

It is essential to examine the intactness of reality testing concerning perceptual alterations, as well as assessment of a psychotic thought process, overvalued ideas, ideas of influence/reference, and Schneiderian symptoms in the differential diagnosis of depersonalization/derealization disorder from psychotic spectrum disorders such as schizophrenia, prodromal schizophrenia, or schizotypal personality disorder. The diagnosis of depersonalization/derealization disorder is based on the "as if" experiential nature of depersonalization/derealization, without distorted, delusional, dereistic cognitive explanations. This is similar to the assessment of a psychotic process accompanying obsessive-compulsive symptoms. A study comparing perceptual alterations and magical ideation in depersonalization/derealization disorder and healthy volunteers found that people with the disorder only had elevated perceptual, but not cognitive, distortions.

Course and Prognosis

Depersonalization/derealization after traumatic experiences or

intoxications commonly remit spontaneously after leaving the traumatic circumstances or the episode of intoxication ends. Depersonalization/derealization accompanying mood, psychotic, or other anxiety disorders commonly remit with definitive treatment of these conditions.

Depersonalization/derealization disorder itself may have an episodic, relapsing and remitting, or chronic course. The latter is most common. Many patients with chronic depersonalization/derealization disorder have a course characterized by severe impairment in occupational, social, and personal functioning. Mean age of onset is thought to be in late adolescence or early adulthood in most cases. Most depersonalization/derealization disorder patients are initially treated for secondary anxiety and mood disorder symptoms. The primary nature of the depersonalization/derealization disorder is often only recognized later on. Traumatic or stressful events may exacerbate depersonalization/derealization disorder symptoms. Symptom exacerbations are commonly related to negative effects, high levels of sensory input, and threatening, stressful, or unfamiliar situations.

Treatment

Clinicians working with depersonalization/derealization patients report them to be singularly refractory to treatment. Several factors may lead to this perception. Due to the typically substantial delay in diagnosis of the condition, clinicians often do not have the opportunity to intervene until years after its onset, at which point the symptoms tend to become constant, nonfluctuating, and less responsive to therapeutic intervention. Clinical reports suggest that individuals with depersonalization/derealization disorder seen earlier in the course of their disorder, and those who continue to have fluctuating or episodic symptoms over the years, may be more responsive to treatment.

Psychotherapy. Many studies describe a range of psychotherapies to treat depersonalization/derealization disorder, including psychodynamic, cognitive, cognitive-behavioral, hypnotherapeutic, and supportive therapies. No systematic data exist that compare these modalities, and so far there is only one systematic study published on the treatment of the disorder.

Psychodynamic treatment for the disorder focuses on underlying threats to self-constancy that give rise to affectively intolerable states. According to psychoanalytic theories, a person in whom the cohesiveness or stability of self-representations is profoundly threatened may resort to depersonalization/derealization (i.e., a disconnection from the self) as a solution. Similarly, psychodynamic theories suggest that depersonalization/derealization can be linked to various levels of character

pathology. In psychotic spectrum character pathology, depersonalization/derealization may be triggered by experiences of impaired self-and-other differentiation. In borderline psychopathology, unstable and switching self-representations may be associated with depersonalization/derealization. In narcissistic pathology, when self-constancy is threatened by the loss (real or imagined) of self-objects serving purposes of object constancy, depersonalization/derealization may arise. Finally, in neurotic psychopathology, de-repressed self-representations associated with overwhelming intrapsychic conflict may trigger depersonalization/derealization. In psychodynamic psychotherapy, the therapist has the opportunity to observe and analyze, often in a moment-to-moment fashion, these and related dynamics as they occur during psychotherapy sessions and/or are described by the patient in his/her life.

More recently, Daphne Simeon has proposed that addressing the fear of affects is a crucial component of psychotherapy for depersonalization/derealization disorder. Dissociation of affect ("not feeling") and alexithymia (difficulty identifying and verbalizing internal emotional states) are core components of the condition. Depersonalized/derealized patients will frequently recount traumatic or stressful life experiences in a highly robotic, matter-of-fact, and emotionless way. The goal of an affect-based therapy is to uncover, experience, label, own, and verbalize intolerable emotions, and to process these emotions in the context of underlying conflicts and disavowed self-representations and self-states. Eventually, such emotions can be gradually integrated with the core sense of self, so that the individual can transition from an "unreal" self to a more "real" owned self. Which affects need to be dealt with vary greatly depending on each person's history and sense of self, ranging from the negative to the positive, and can include anger, grief, sorrow, shame, guilt, excitement, and love. In a simplistic sense, this approach is "opposite" to the affect regulation component of the treatment of trauma-related conditions such as PTSD or dissociative identity disorder, as the goal in the therapy of individuals with depersonalization/derealization disorder is to activate and experience rather than to better modulate and regulate intense affects. On the other hand, as noted in the section on dissociative identity disorder, patients with dissociative identity disorder and the dissociative subtype of PTSD may attempt to mobilize depersonalization/derealization, affect suppression, and state shifting in an attempt to distance and modulate posttraumatic intrusions. Frequently, in the latter patients, this effort at dissociative distancing fails, and patients experience severe flashbacks. They may report, and/or display, oscillations between being all "shut down" and experiencing overwhelming flashback symptoms that can completely overwhelm awareness of current reality when they do break through into consciousness.

In recent years, a disorder-specific cognitive-behavioral therapy approach to depersonalization/derealization disorder has been developed and tested, based on the cognitive-behavioral model outlined earlier. In this type of treatment, the focus is on developing techniques to cognitively reinterpret symptoms as less threatening; to reduce symptom monitoring; and to diminish avoidance and safety behaviors. In an uncontrolled, prospective treatment trial of 21 patients utilizing this approach, significant reductions in ratings of depersonalization/derealization, depression, and anxiety were found. In addition, by the end of the 12-week treatment, 29 percent of participants no longer met diagnostic criteria for the disorder.

A proportion of severely impaired depersonalization/derealization patients may require long-term supportive treatment. The clinician will need to be acutely aware of patient's interpersonal sensitivity, distress, and sense of hopelessness about the condition. Some supportive psychotherapy elements are very important for the effective treatment of all patients with the disorder. These include psychoeducation, such as giving the condition a name and reassuring the patient that it will not evolve into a psychotic disorder and that there is no evidence of permanent brain damage. Also, psychoeducation and supportive interventions can help the patient lift the barrier of secrecy based on shame about symptoms in general, alleviate guilt or shame associated with a drug-induced onset, and provide hope for the future. In this educational/supportive/cognitive-behavioral model, the clinician challenges the patient's experience of the "physicality" of the experience. This can include emphasizing that depersonalization/derealization disorder is a psychological disorder, as physical as it may feel; providing a sense of control over the symptoms that may be perceived as continuous in intensity; using techniques that can help modulate the level of arousal and indirectly soften the symptoms; and reassuring women who are considering pregnancy that the heritability, as far as it is known, appears to be limited. There are no adequate data on the utility of either hypnosis or EMDR in treating chronic depersonalization/derealization. However, one study found that these patients had relatively low hypnotic capacity on standardized hypnotizability measures. This is a potential marker to differentiate individuals with depersonalization/derealization disorder from other dissociative disorder patients, such as those with dissociative amnesia and dissociative identity disorder that demonstrate high hypnotizability on standardized testing protocols.

Pharmacotherapy. Extensive retrospective treatment histories have not provided support for major symptom improvement with any of the known and widely used classes of psychiatric medications, although benzodiazepines, serotonin reuptake inhibitors, and stimulants appear to be partially efficacious for some patients. In past decades typical

antipsychotics were widely used for chronic depersonalization/derealization, often as first-line treatment, as many clinicians viewed depersonalization/derealization symptoms as indicative of a psychotic spectrum condition. There is absolutely no empirical evidence for the efficacy of typical or atypical antipsychotic medications in depersonalization/derealization disorder. They should only be used to target overwhelming anxiety or mood disturbance associated with the depersonalization/derealization. Indeed, these medications may increase a feeling of emotional deadness and lack of emotional response to oneself or the world. Thus, in some individuals with depersonalization/derealization disorder, these medications may actually increase the experience of depersonalization/derealization.

There is some anecdotal support for the use of SSRIs (including clomipramine [Anafranil]) in depersonalization/derealization disorder. However, a recent large randomized, controlled trial of fluoxetine found only a statistically marginal, and clinically minimal, improvement in depersonalization/derealization compared to placebo. Depersonalization/derealization and dissociation scores did not significantly decline in the fluoxetine-treated group, paralleling patients' subjective reports that their symptoms had not changed. On the other hand, patients reported feeling somewhat better and less distressed by symptoms. Within the fluoxetine-treated group, depersonalization/ derealization responders were more likely than nonresponders to show improvement in comorbid anxiety disorder symptoms. This may account for the experience of being less distressed by their symptoms, although the depersonalization/derealization symptoms remained unchanged.

Therefore, it appears that serotonergic antidepressants are unlikely to be of major benefit for depersonalization/derealization disorder symptoms but may potentially ameliorate prominent comorbid anxiety and depressive symptoms. Lamotrigine—a mood-stabilizing anticonvulsant that promotes glutamate release—initially showed promise for depersonalization/derealization disorder symptoms in a small open-label study. However, a subsequent small crossover placebo-controlled trial found that all nine participants were nonresponders. Some have suggested that the combination of SSRIs and lamotrigine may have additive benefit in treating the disorder, but no empirical support exists for this hypothesis.

Opioid antagonists such as naltrexone (Revia), cognitive enhancers, benzodiazepines, and stimulants all have had some benefit in clinical treatment of specific depersonalization/ derealization disorder patients. In uncontrolled and/or proof-of-concept studies, a trial of repetitive transcranial magnetic stimulation (rTMS) for depersonalization/derealization disorder has shown efficacy in some depersonalization/derealization disorder patients. rTMS models have involved stimulation of the temperoparietal junction (TPJ), and the

ventrolateral prefrontal cortex (VLPFC) areas that have been implicated in generation of depersonalization/derealization disorder symptoms.

OTHER SPECIFIED DISSOCIATIVE DISORDERS

The category *other specified dissociative disorders* covers conditions characterized by a primary dissociative response that do not meet diagnostic criteria for one of the four DSM-5 dissociative disorders. This category encompasses a clinically heterogeneous set of dissociative conditions, often with multiple comorbidities, and provides a diagnostic home for several Western and non-Western idioms of distress. Many patients with these conditions have a current or past history of severe trauma. To receive a diagnosis of other specified dissociative disorder, the dissociative presentation must not be accounted for better by ASD, PTSD, or other disorders characterized by dissociative symptoms.

Other specified dissociative disorder in DSM-5 lists four examples to assist clinicians in recognizing conditions in this spectrum. Research findings may support the future inclusion of some or all of these examples as specified dissociative disorders, and/or lead to modifications of the diagnostic criteria for the current DSM-5 dissociative disorders. Future studies should obtain more data on each example to determine whether it represents a stable disorder with a typical set of symptoms, as well as on its antecedent trauma, course, neurobiology, etiology, treatment response, and other outcomes. Some patients who initially receive this diagnosis will ultimately be better conceptualized as having another dissociative disorder, such as dissociative identity disorder.

EXAMPLE 1: ATYPICAL PRESENTATIONS NOT MEETING FULL DIAGNOSTIC CRITERIA FOR DISSOCIATIVE IDENTITY DISORDER

Definition

Example 1 of other specified dissociative disorder covers presentations similar to dissociative identity disorder but with: (a) less-than-marked discontinuities of self and agency and/or (b) alterations of identity or episodes of pathological possession in an individual who reports no dissociative amnesia.

Diagnosis and Clinical Features

As noted in prior sections, a major goal of the DSM-5 Work Group was to reduce the frequency of DSM-IV-TR dissociative disorder NOS diagnoses. Nonetheless, Example 1 is meant to indicate cases where the experience of personality states is not sufficiently distinct, and/or the discontinuities of self insufficiently severe to warrant a diagnosis of dissociative identity disorder (see section on Changes in Diagnostic Criteria to the Dissociative Disorders).

Example 1 can be used as an initial placeholder for a diagnosis of dissociative identity disorder. This allows additional, sequential clinical observation, use of diagnostic screening inventories and interviews, and acquisition of data from collateral informants. Over time, with further diagnostic assessment, most of individuals with this disorder will likely meet diagnostic criteria for dissociative identity disorder. Rarely, clinicians describe patients who fulfill the DSM-5 Criterion A for dissociative identity disorder, but in whom dissociative amnesia is not reported or observed, even on careful questioning. Clinical experience suggests that, with careful sequential observation, most of these individuals ultimately meet the amnesia criterion for dissociative identity disorder. Ultimately, the clinician usually: (a) detects subtle dissociative amnesia symptoms; (b) realizes that the patient manifests dissociative amnesia for episodes of dissociative amnesia (e.g., witnesses the patient not recalling several minutes of a therapy session, with the patient unaware that this has occurred); and/or (c) discovers subtle dissociative amnesia for aspects of the life history, but virtually no contemporary dissociative amnesia symptoms such as ongoing blackouts, fugues, and unexplained belongings. Some of these patients may report continuity of memory inconsistent with dissociative amnesia. However, on careful evaluation, they report that they actually experience substantial parts of their autobiographical memory without an affective, first-hand component: as if it happened to someone else, they heard about it, read about it, and/or have been told by friends and families about past events. Ultimately, the clinician and patient realize that this depersonalized/derealized memory actually conceals the dissociative amnesia, and the patient acknowledges distress about the extent of autobiographical memory deficit.

Some patients with dissociative identity disorder who have achieved partial, but not complete, unification of self-states may continue to display these states, albeit usually in much more attenuated form, and may have significantly improved memory continuity. As treatment progresses, they may no longer meet Criterion A for dissociative identity disorder. Patients with dissociative identity disorder who achieve complete unification do not manifest personality states, or dissociative amnesia, although they may suffer from other more quotidian memory difficulties, such as those associated with advancing age. Under significant stress or later life trauma, some of these individuals display reactivated personality states, but usually without the distinctness that manifested during earlier phases of treatment. Also, they may report dissociative amnesia symptoms for the stressful events that led to the disruption in unification. Technically, some of these patients would meet diagnostic criteria for other specified dissociative disorder, or even no longer meet diagnostic criteria for a dissociative disorder. However, there are no studies that examine the question of classification of dissociative identity disorder in patients who have achieved

stable unification.

Treatment

Based on research on DSM-IV-TR dissociative disorder NOS, most patients with Example 1 of other specified dissociative disorder will require a treatment approach similar to that recommended for dissociative identity disorder.

COERCIVE PERSUASION

Definition, Diagnosis, and Clinical Features

The concept of *brainwashing* or *thought reform* emerged from Western reactions to the Soviet Communist show trials of the late 1940s and alarm over Communist Chinese methods of coercive political re-education. These fears intensified during the Korean conflict after many American prisoners of war made anti-American statements. In 1950, a journalist, Edward Hunter (later identified as a Central Intelligence Agency [CIA] agent) proposed that the Chinese Communists had discovered techniques to modify mental attitudes and beliefs, a process that he called *brainwashing*.

After the Armistice, a team of psychiatrists and psychologists, including Robert J. Lifton and Edgar Schein, interviewed the returning prisoners. Similar studies done for the CIA also obtained information from Communist interrogators and former Soviet and Chinese prisoners. These studies concluded that the apparent brainwashing of these prisoners, as well as of the victims of the Russian show trials, was the result of extreme police interrogation techniques involving severe physical and psychological deprivation, followed by rewarding the severely deprived prisoners with the possibility of the end of their ordeal, as well as food and warm clothing. The reports concluded that the basic attitudes of the American prisoners in Korea had not, in fact, been altered.

However, the subsequent release of another group of prisoners (missionaries, businessmen, doctors, and students) caught in China at the beginning of the Korean War did seem to suggest that some form of thought reform had occurred with these individuals, several of whom continued to falsely insist that they were spies. In a second round of studies, Lief, Schein, and others concluded that changes in an individual's basic beliefs could occur in the context of extreme physical, psychological, and social coercion. In this context, this shift may also have to do with development of dependency of the prisoner on his or her captors. In this kind of imprisonment, the captors/torturers become the prisoner's only human contact, and can bestow torture or respite, based on the prisoner's behavior, leading to the prisoner giving in to the demands of the captor. Judith Herman, in her classic books on trauma, noted the similarity of the captor/captive relationship across various traumatic contexts: intra-

familial child abuse, intimate partner violence, and political torture. Similarly, the use of extreme group persuasion methods by the Chinese Communists on political prisoners may have led to such an outcome. Clinical presentations of individuals subjected to extreme forms of coercive persuasion may resemble those of survivors of other sorts of torture.

One legacy of the fear of Communist brainwashing was the unfortunate misuse of psychiatry and psychology by the CIA and other U.S. government agencies from the late 1940s onward to develop methods of behavioral control, which was revealed in release of government documents and Congressional hearings in the 1970s. In an effort to combat or preempt potential Communist use of mind control, the CIA studied or supported studies of psychedelic drugs and other toxins and intoxicants, radiation, hypnosis, sensory deprivation, repeated ECT treatments, and bizarre conditioning experiments, often on unwitting civilians or military personnel. At least one individual killed himself during an LSD experiment, and other psychiatric casualties have brought suits against the U.S. and Canadian governments because of damage caused them during CIA-funded experiments.

According to reviews of CIA documents, some CIA officials apparently believed in the possibility of using hypnosis to create assassins or spies with hypnotically induced alter identities or amnesia—or both—who could thus resist interrogation and torture. Other CIA officials in charge of mind-control experiments doubted that this could be done successfully. Because CIA officials destroyed many documents relating to these activities, it is unknown whether these experiments were ever attempted. Popularized accounts continue to suggest that the U.S. government attempted to create multiple-personality spies. However, a careful reading of cases purported to demonstrate this show that, based on available information, most bear little similarity to dissociative identity disorder.

Allegations of brainwashing or mind control resurfaced in the late 1960s and 1970s in the context of counterculture religious movements. This theory was also advanced by a group of defense psychiatric experts that included Louis J. West, Lifton, Martin Orne, and Margaret T. Singer in the famous case of Patty Hearst, the heiress who was kidnapped, tortured, sexually abused, and held in prolonged solitary confinement by a radical cult-like group. After many months of this treatment, Hearst was said to have transformed her identity into that of the revolutionary Tania, who aided her captors in criminal acts, including bank robbery and murder. The Patty Hearst jury was not convinced, however, and convicted Hearst despite the arguments of the psychiatric experts for the defense.

Singer subsequently publicized her ideas widely about conditioning techniques that she believed were used by religious cults to render their members incapable of complex rational thought or decision making. In a series of lawsuits by ex-members against various religious cults, she

testified that these techniques were capable of overpowering a person's free will and that the group's control over a member could be total. These often-sensational trials instilled in the public the notion of brainwashing as a common practice in religious cults.

In response, a group of academics, primarily psychologists and sociologists, challenged Singer, pointing out that cult members often came from dysfunctional families and joined these groups to avoid the anxieties and responsibilities of independence. Psychiatrist Marc Galanter reframed the religious cult mind-control debate by advocating the more neutral term *charismatic religious sects* and pointing out that an individual's behavior within such a social group may reflect psychological adaptation rather than psychopathology. Studies of ex–cult members indicated that the manner of leaving the group, voluntarily or by being involuntarily "deprogrammed," was a significant determinant of their subsequent assessment of whether they had been brainwashed, with the latter the most likely to express this belief.

In the context of a particularly sensational case, the American Psychological Association and American Sociological Association submitted *amicus* briefs refuting the concept of brainwashing as lacking scientific validity, which has had the larger effect of essentially nullifying its legal status. Although the academic community has mostly avoided the topic in the last decade, the notion of brainwashing—in which individuals are rendered hyper-suggestible through hypnosis, drugs, or physical stressors and their belief systems are transformed by special conditioning techniques—continues to fascinate the public and to be invoked by a few psychotherapists and others as evidence of occult or government conspiracies (also see the section on History of Dissociative Disorders).

At the same time, some researchers on destructive cults have continued to suggest that cult membership produces an "ego state" or dissociated self-aspect that allows the person to function ego-syntonically within the cult milieu. This entity is not conceptualized as meeting diagnostic criteria for a dissociative identity disorder personality state but suggests a diagnosis that may be analogous to Example 1. Research attempting to fully explicate such issues is likely to be exceptionally difficult to perform.

The military stress studies of Morgan and colleagues prove, however, that combinations of pain, hunger, cold, fatigue, fear, sensory deprivation, and lack of control over bodily functions and social communication can profoundly alter an individual's state of consciousness, producing prominent dissociative symptoms. The accompanying degree of identity alteration induced by such techniques remains to be empirically established, but it is possible that dissociative reactions could account for many of the psychological changes reported in captive individuals subjected to coercive interrogation and persuasion. Additional studies of military survival school participants' show that, under conditions where a

variety of misinformation is provided, subgroups of these soldiers, may show memory distortions for aspects of a highly stressful mock interrogation and/or characteristics of their interrogators, including incorrect photo identification of interrogators. In one study, however, dissociation scores did not predict facial recognition inaccuracy in photo line-ups of interrogators. No subject failed to recall being subject to interrogation as a whole, which included verbal harassment, physical assaults, as well as sleep and food deprivation, although showing memory distortions for aspects of the events and the characteristics of interrogators. Also, military survival school trainees *know* that they are not actually captives, albeit showing extreme levels of stress hormones, dissociation scores, and other measures of distress. However, by implication, actual prisoners who undergo torture and deprivation may be subject to incorporation of substantial misinformation at the behest of their captors, as well as subject to post-event misinformation after being rescued (see section on Studies of Military Survival School Subjects).

Diagnosis and Clinical Features

Individuals subjected to extreme coercive techniques are at risk for persistent depersonalization and, possibly, other dissociative symptoms, including amnesias, trance-like behaviors, and emotional numbing. They may exhibit reduced cognitive flexibility, behavioral regression, and profound changes in values, attitudes, beliefs, and sense of self. In some instances, a type of personality state may emerge that identifies with former tormentors and behaves in a fashion contrary to the individual's prior behavior. This state usually disappears or weakens significantly when the individual is returned to safety and security but may be transiently reactivated by circumstances reminiscent of the traumatic events. Some forensic experts in cases of violent behavior by cult members have posited that the influence of the cult leader results in a "shared psychotic disorder" (now Example 4 in Other Specified Schizophrenia Spectrum and Other Psychotic Disorder). Others have disputed the accuracy of this construct to explain personality and/or behavioral changes in members of cult groups.

Treatment

There are no empirical treatment studies of individuals subjected to extreme coercion applied in the service of indoctrination into a belief system or an attempt at identity alteration. The basic principles of phased trauma treatment would seem to provide an organizing framework for their care, as for victims of other forms of torture. This begins with the creation of safety and symptom stabilization, along with educational information to help normalize the individual's response. This is followed, if appropriate, by a supportive but nonsuggestive exploration of the traumatic experiences

to desensitize their memory and to integrate them better into the individuals' transformed sense of self. Any preexisting psychopathology is likely to be exacerbated by such experiences and should be addressed as the dissociative symptoms abate. Shame over "being broken" and grief over losses experienced because of coercion or cult involvement may be important issues in treatment. Family and social interventions may be necessary as a result of the duress and disruption accompanying the precipitating events and the social effects of the profound changes in the individual's attitudes, behaviors, and beliefs, as well as sequestration from everyday events during the period of captivity.

ACUTE DISSOCIATIVE REACTIONS TO STRESSFUL EVENTS

Definition

Example 3 of other specified dissociative disorder describes acute, transient dissociative symptoms that last 1 month or less, and frequently only hours or days, and occur in reaction to stressful events. Because many of the briefer episodes ended when the patient was brought rapidly for emergency treatment, they might have lasted longer without intervention. Acute dissociative reactions cover a spectrum of severity, ranging from milder presentations of mixed dissociative symptoms to more severe but transient presentations that combine dissociative and psychotic symptoms. The prominence of somatic presentations of dissociation (e.g., conversion symptoms such as analgesia or paralysis) and psychological presentations (e.g., derealization, constriction of consciousness) also varies; a mix of both presentations is typical.

The acute dissociative symptoms may be precipitated by traumatic or nontraumatic stressors and persist for up to a few weeks after the stressor dissipates. Although patients are not always able to identify and report a stressful event as a potential cause of their disorder, a triggering event is usually recognized as precipitating the dissociative reaction once collateral information is obtained. The severity of the stressor varies, including frankly traumatic events such as the unexpected, violent death of a relative, or experiencing a natural disaster. However, less stressful events, such as the loss of a significant relationship may also precipitate this presentation. The Western research literature has tended to subsume immediate dissociative reactions to stress or trauma into the concept of peritraumatic dissociation, defined as dissociative reactions occurring during and after a traumatic event. However, Example 3 does not require the presence of traumatic exposure, which distinguishes this condition from ASD.

Some cultural syndromes and other folk illnesses may fulfill criteria for Example 3, when they are characterized by acute reactions to adversity with mixed but transient dissociative symptoms. For instance, some cases of *ataques de nervios* ("attacks of nerves") in Latino populations would best

be described using this diagnosis. However, cultural concepts of distress do not have a one-to-one relationship with psychiatric disorders; instead they cut across diagnostic clusters and must be translated into psychiatric categories at the level of the individual. For example, some persons with *ataques de nervios* qualify for a dissociative diagnosis—including other specified dissociative disorder—whereas others do not, meeting criteria instead for panic disorder, PTSD, or no psychiatric disorder.

Many of the patients with mixed and transient dissociative and psychotic symptoms would fit the diagnostic construct of *hysterical psychosis,* a condition studied extensively in the 19th century. Janet focused particularly on the traumatic origins of these symptoms, and described their resolution through symptom stabilization, work on traumatic memories and fixed ideas, and, finally, conscious synthesis of the traumatic memories. As described above (section on History), at the end of the 19th century dissociation-based theories of hysteria waned, the Freudian theories of hysterical symptoms became ascendant, and the Bleulerian view of schizophrenia came into prominence, subsuming hysterical psychosis. One dominant view was that individuals with hysterical psychosis had histrionic personality traits, limited ego resources, and were overwhelmed by conflicts, often sexual in nature, and/or overwhelming losses. Other authors maintained that these individuals manifested a type of schizophrenia, and others saw them more in the factitious/malingering spectrum, especially in the context of compensation for injury. More recently, with the rediscovery of Janet's work and the development of the TM of dissociation, individuals with hysterical psychosis have been described as reacting to severe traumatic events, leading to alternative terminology such as *reactive dissociative psychosis.*

These combined dissociation-psychosis presentations must be distinguished from brief psychotic reactions, usually to traumatic stressors, without dissociative symptoms. In the North American literature, many of these patients may be conceptualized under the DSM 5 diagnosis of *brief psychotic disorder,* previously known in DSM-III and DSM-III-R as *brief reactive psychosis.* Symptoms of the latter condition overlap considerably with that of hysterical psychosis, as the diagnostic criteria for brief reactive psychosis specified that precipitating events included severe stressors such as a major loss, or traumatic experiences, such as combat. In DSM-5, brief psychotic disorder has specifiers for occurrence subsequent to severe stress/trauma and without stress/trauma, respectively. Studies of brief psychotic disorder have not used measures of dissociation, and systematic studies of stressful or traumatic antecedents to brief psychotic disorder have not been conducted.

Turkish researchers describe hysterical psychosis as a common, well-known syndrome that is regularly diagnosed by Turkish psychiatrists, often presenting in the emergency department. Some Turkish authorities have

reported case series of patients where hysterical psychosis was the presenting symptom cluster of dissociative identity disorder. The latter was recognized as the patient responded to treatment of the acute symptoms of the hysterical psychosis.

Diagnosis and Clinical Features

The literature identifies several overlapping symptom clusters characteristic of these acute, brief, reactive dissociative conditions. DSM-5 describes symptoms of constricted consciousness, depersonalization/derealization, perceptual disturbances (such as time slowing, metamorphopsia, brief amnestic experiences, transient stupor, and/or conversion sensory–motor deficits. Uncomplicated cases of Example 3 generate disturbances in interpersonal relationships and work-related functioning that usually resolve within a few days. Nevertheless, crisis intervention during the acute episode is often required. This category is also associated with a risk of suicidal behavior.

Acute dissociative reactions may be codified culturally into stereotyped reactions with associated causal attributions, courses, and treatment responses. In Latino communities, for example, cases of *ataque de nervios* may present with prominent dissociative symptoms consistent with Example 3 of other specified dissociative disorder. Patients with *ataques* describe acute fits of emotionality associated with a sense of loss of control, usually precipitated by family stressors. Severity of *ataques,* assessed as lifetime number of episodes, is positively and independently related to self-reported dissociative symptoms measured with the DES and to clinician-diagnosed dissociative disorders, assessed with the Structured Clinical Interview for DSM-IV-TR Dissociative Disorders. In a representative community-based sample of US Latinos, as well as in an urban primary care Latino sample, endorsement of *ataque* on a standardized survey was associated with a higher risk of suicidal ideation and behavior as well has higher morbidity (e.g., greater impairment). A case of *ataque* that meets criteria for Example 3 is included next.

OTHER DISSOCIATIVE DISORDERS CASE 1: CASE OF ATAQUE DE NERVIOS

Ms. C, a 27-year-old woman who migrated to the United States at age 14 from the Dominican Republic, had an episode of acute dissociative and aggressive symptoms while living in her brother's apartment in a large urban center. She had lost her apartment 3 months before, after separating abruptly from her common-law husband, and had been taken in by her brother into his rented apartment, along with her 7-year-old daughter. The morning of her acute episode, her brother informed Ms. C that she could not stay in the apartment any longer because he was being threatened with legal action for violating his lease by taking in additional family members.

After receiving the news, Ms. C reacted initially with apparent calm. However, during a discussion with her brother 2 hours later—when he again raised the topic of her departure and inquired as to her timeline for leaving—she became agitated, started screaming and crying, and attacked her brother with her hands and nails, aiming for his face. The brother was able to restrain her, but she ripped part of his t-shirt off his chest. For approximately one minute,

whimpering softly, Ms. C slowly and repeatedly tore into small pieces the section of the T-shirt.

The episode waned gradually, and Ms. C became coherent again and asked her brother, "Why are you holding me like this?" Helped to lie down on a bed, she denied any memory of the event, recalling only the initial part of the argument. She did recall "an enormous rage" as her episode began, and seeing her brother as "unreal" and "very small and far away" and she herself as "immensely powerful and tall, my head nearly touching the ceiling." She felt she was moving "in slow motion" at first, but reported the absence of recall from that point forward.

Ms. C did not want to seek help for this episode—which she labeled an "*ataque de nervios*" (attack of nerves)—attributing the onset of her acute symptoms of depersonalization, derealization, time slowing, micropsia/macropsia, and micro-amnesia to the impact of the sudden news of her "eviction" by her brother. However, Ms. C's family was very concerned about her amnesia questioning her ability to care for her daughter on her own.

Later that afternoon, Ms. C's relatives escorted her to the psychiatric Emergency Room. She described feeling exhausted after her acute episode, and was irritable during the interview, especially about her family's bringing her to the hospital. She slept on her gurney during most of the next 14 hours. The next day in the ER, Ms. C had two more episodes of sudden agitation, crying, and screaming, during which she attempted to scratch her face and leave the room. Because she responded quickly to verbal intervention and reassuring physical contact, she was not given medications or placed in restraints, but was kept on one-to-one observation. Laboratory test results were unremarkable, as were an electroencephalogram and a lumbar puncture. On examination, she was crying softly, her mood was sad, and she reported transient depersonalization and derealization and ongoing amnesia for her acute episode. She was oriented to time, person, and place; had no psychotic symptoms; and denied current suicidal ideation.

In light of her gradual improvement and lack of evidence of immediate danger to self or others, the ER staff did not feel Ms. C qualified for an involuntary hospitalization, and she refused a voluntary admission. After 36 hours of ER observation, she was discharged with an outpatient appointment for herself and her daughter. However, she did not keep the initial appointment and was lost to follow-up.

Cases of dissociative psychosis usually have a sudden onset; duration is typically 1 to 3 weeks, but in some cases may be longer. The individual's affect is often labile. Due to rapid shifts among different mental states, the individual may exhibit extreme ambivalence—doing and undoing some act in a psychotic or perseverative fashion. The psychotic and dissociative symptoms may have prominent posttraumatic content, which can be embedded in complex perceptual and somatosensory phenomena that occur automatically and uncontrollably (e.g., dissociative flashbacks to the traumatic event). In some cases, functional neurological (conversion) symptoms may be the most prominent presenting sign(s). This condition causes a temporary impairment of daily functioning that typically requires short-term hospitalization. It is also associated with risk of suicidal behavior. A high prevalence of childhood abuse and/or neglect is often associated with dissociative psychosis.

Nonpsychotic cases of Example 3 are differentiated from adjustment disorder, ASD, and PTSD by the predominantly dissociative nature of the presentation and the absence of other symptoms of these disorders (e.g., hyperarousal, avoidance). There is no dissociative subtype of adjustment disorder, and ASD can occur in the absence of dissociative symptoms and, Example 3 presentations do not require a traumatic stressor.

Case descriptions of the symptoms of dissociative psychosis have

included bizarre behavior, such as behaving like an animal, and catatonic-like symptoms. Speech patterns may be disjointed and intermittent, sometimes giving the clinical impression of psychotic thought disorder. However, when the patient can speak coherently, there is no evidence of a psychotic thought disorder.

Differential Diagnosis of Dissociative Identity Disorder and Dissociative Psychosis. Some case series of individuals with dissociative psychosis have described symptoms that appear to be consistent with dissociative identity disorder. For example, a woman diagnosed with hysterical psychosis repeatedly hallucinated a second woman, with more resilient characteristics than herself, who conversed with her supportively about the life difficulties that preceded the psychotic episode.

Reports describe dissociative psychosis patients as experiencing multimodal hallucinosis; age regression; extensive dissociative amnesia; oneroid states; pseudodelirium/pseudodementia; and delusional phenomena. However, in these accounts, the apparent delusions could be understood as related to, and usually symbolically condensing, traumatic experiences and intrapsychic conflicts that have generated the episode. Clinical experience suggests that most of these patients do not have a psychotic disorder, but, with more extensive evaluation, will meet diagnostic criteria for dissociative identity disorder.

Pathology and Laboratory Examination. Research on hysterical psychosis has documented high scores on standard measures of hypnotic capacity in individuals with this disorder. High hypnotizability distinguishes this group of patients from those with psychotic disorders, who generally have low scores on standard measures of hypnotizability.

In contrast to acute organic mental disorder (delirium), persons with dissociative psychosis do not exhibit findings that are characteristic of delirium on EEG, brain imaging, laboratory tests, or neurological examination. They are characterized by dissociative symptoms such as amnesia, dissociative flashbacks, and disruptions of consciousness, rather than by the waxing and waning of symptoms, cognitive deficits, and interpersonal impoverishment typically found in patients with delirium.

Treatment

There are no systematic studies of treatment of cases of Example 3, and treatment principles are derived from case series. *Ataques de nervios* and other folk illnesses fitting Example 3 may be characterized by suicidal or aggressive behaviors in the midst of the acute dissociative episode. Restraining the person may be temporarily necessary. In many cases, the patient or family do not seek professional help. Paradoxically, if the triggering stressors are deemed minor by significant others or there are

concerns about chronicity, the individual may be brought for psychiatric treatment. While comorbid symptoms (e.g., depression) may respond to psychopharmacological interventions, the dissociative symptoms are more likely to respond to supportive-expressive psychotherapy that identifies the traumatic and/or stressful events that preceded the dissociative reaction, as well as the intense psychological conflicts that, almost invariably, have been generated in the context of the stressful events. Usually, supporting the patient as he/she "tells the story" of the stressful event is most therapeutic.

Treatment of dissociative psychosis follows the structure developed by Janet: education to help the patient understand the disorder; supportive interventions to help the patient develop better control over symptoms; expressive interventions to give voice to distress over traumas, conflicts, and losses; and transformation of reenactments into a coherent narrative. Following Janet, modern case series have demonstrated the effectiveness of hypnotic interventions in the treatment of patients with dissociative psychosis who showed high hypnotizability on standardized measures. Clinicians trained in hypnosis may find that assessment of hypnotizability aids in differential diagnosis, as well as hypnotic interventions being effective as an adjunct to psychotherapy for these patients.

Hypnotic interventions give these patients a greater sense of control over symptoms (and autohypnotic states); help the patient to understand symptoms as logical in the context of trauma and/or stressful events; and can provide imagery that may help the patient with affect regulation, symptom control, grief-work, mastery of traumatic material, and integration of these experiences for healing. Patients may also helpfully use self-hypnotic techniques to manage symptoms outside of psychotherapy.

Patients with dissociative identity disorder in an acute episode of dissociative psychosis should be treated for the underlying disorder. The clinician should provide the patient psychoeducation about acute, dissociative psychotic symptoms, to reassure him/her that these are part of the dissociative disorder, and not a manifestation of a comorbid psychotic disorder. The clinician should emphasize that treatment of the episode will follow the phasic trauma treatment model, to understand what has engendered the acute symptoms, and that medication is unlikely to have a major direct impact on symptoms, although medication may help decrease the intensity of severe terror states and assist with sleep (see discussion of psychopharmacological treatments for dissociative identity disorder). The clinician should have an index of suspicion that the dissociative identity disorder patient in an acute dissociative psychotic episode may be responding to a recent traumatic event such as a sexual assault that he/she cannot process in verbal form without psychotherapeutic interventions. A case example explores this kind of presentation in a dissociative identity disorder patient.

OTHER DISSOCIATIVE DISORDERS CASE 2

Ms. M, a 29-year-old dissociative identity disorder patient was admitted to a psychiatric hospital by her therapist. The therapist reported that the patient had been struggling with safety issues, and did not attend an emergency appointment. The therapist called 911 and the police had to forcibly enter the patient's home. The patient was lying in her bed, mute except for moaning, rocking back and forth, and was otherwise unresponsive. She was taken to the emergency department where medical, toxicology, and laboratory examination were within normal limits. The patient had superficial self-inflicted cuts on her arms, bruising on her legs, and cuts on her feet. Examined by the ED psychiatrist, she spoke in a robotic voice, and said she was seeing "lights flashing, colors spinning," and was "hearing so many voices, I can't make them out at all," some voices were experienced inside her mind, but others were "bouncing around the room." She saw "lots of faces," felt "my body just floating around," but also would moan and say, "it hurts" but could not explicate this further. She said her thoughts were 'being sucked out of my mind," but was perplexed about this, yet no delusional explanation was elicited. She denied thought broadcasting or audible thoughts, and was "scared and frightened . . . I think I'm going crazy." At times her thoughts appeared disorganized with blocking, at other times, thought process was coherent and goal directed. She reported intense urges "from somewhere" to "cut myself to pieces and die." She was admitted to the hospital and treated for "acute psychosis" with olanzapine 10 mgs by mouth daily, and lorazepam 2 mg by mouth every 6 hours as needed, with no change in symptoms over a week, other than increased sedation. She was seen by a consultant who attempted to explore the origins of her recent symptom exacerbation. With consistent focus on the patient's making sense of her experience, she reported a recent sexual assault by an "old friend" whom she had encountered at a bar, who may have drugged her drink. Taken to his house, she found herself being assaulted, felt very confused, and attempted to run away, cutting her feet. She reported that his grabbing her legs as she tried to run out of the house caused the bruising. She eventually escaped and found her way home. Then, she cut herself "for punishment for being so stupid." She reported a history of being drugged repeatedly in childhood while being trafficked in child pornography and prostitution. The hallucinations and other perceptual abnormalities, as well as the thought disorganization, were related to specific flashbacks of these experiences, apparently stimulated by the drugging and the rape. After talking about this, the hallucinations abated, and the patient became calmer. She was given psychoeducation about psychotic-like dissociative symptoms. At the time of discharge, antipsychotic medication was discontinued, and the patient appeared grounded, oriented, and without perceptual disturbances or hallucinations, other than the typical internal voices that she experiences from self-states, and from more typical PTSD intrusive symptoms.

DISSOCIATIVE TRANCE DISORDER

Definition

DSM-IV-TR included pathological trance and pathological possession trance states in Appendix B as separate elements of a combined category of dissociative trance disorder, one of the "Criteria Sets and Axes Provided for Further Study." Categories in this section of the manual just fell short of full inclusion in DSM-IV-TR as specified diagnoses but were included in the appendix to stimulate research for possible incorporation in DSM-5. Pathological possession was included in DSM-5 under Dissociative Identity Disorder, leading to its redefinition as two variants: possession-form and non–possession-form, "Cultural Concepts of Distress"). By contrast, there was insufficient data on Example 4, dissociative trance disorder, to include it in DSM-5 as a specific dissociative disorder, or to subsume it under another dissociative disorder.

The goal of making the ICD and DSM nosologies more compatible helped to promote inclusion of pathological dissociative trance under other specified dissociative disorder. Pathological trance states are a component of the ICD-10 Trance and Possession Disorders (F44.3), a specified diagnosis under the Dissociative (Conversion) Disorders in the section on Neurotic, Stress-Related and Somatoform Disorders. This ICD-10 diagnosis has been documented in diverse international settings (e.g., Turkey and India) but most studies do not distinguish the prevalence for trance from that for possession trance.

Diagnosis and Clinical Features

DSM-5 describes dissociative trance disorder as a temporary, marked alteration in the state of consciousness or loss of the customary sense of personal identity without alteration or change of identity. There is often a narrowing of awareness of the immediate surroundings or a selective focus on stimuli within the environment. This may manifest as a profound unresponsiveness or insensitivity to environmental stimuli. In addition, the individual may display stereotypical behaviors (e.g., finger movements) of which the individual is unaware, and/or that the individual experiences as beyond his or her control. Some patients experience transient paralysis and/or loss of consciousness. The dissociative trance cannot be a normal part of a broadly accepted collective cultural or religious practice.

ICD-10 describes pathological trance as a temporary loss of both the sense of personal identity and full awareness of the surroundings. Attention and awareness may be limited to or concentrated upon only one or two aspects of the immediate environment, and there is often a limited but repeated set of movements, postures, and utterances. Only trance disorders that are involuntary or unwanted, and that intrude into ordinary activities by occurring outside (or being a prolongation of) religious or other culturally accepted situations qualify for this diagnosis. The DSM-5 and ICD-10 definitions differ on a key point: pathological alterations of identity are exclusionary in DSM-5, qualifying instead for a diagnosis of dissociative identity disorder or Example 1 of other specified dissociative disorder.

Pathology and Laboratory Examination. There are little systematic data on differential diagnosis of this example of other dissociative disorders. Patients with this condition may require neurological evaluation, EEG, and brain imaging to rule out a seizure disorder, and cardiac workup to rule out hypotension and/or arrhythmias. The differential diagnosis of the abnormal movements in dissociative trance disorder may include Huntington disease, essential tremor, and others. Huntington disease may be associated with significant psychiatric psychopathology. However, it usually does not present with loss of awareness of surroundings until very late in the course, when dementia is prominent. It runs in families, and can

be ruled out by genetic testing, if needed. Children with attention-deficit hyperactivity disorder may be hyperfocused on things that interest them, like video games or television, but do not manifest trance-like behavior apart from this. Many individuals can "tune out" parents, spouses, et al. during family strife. However, this is not the same as a spontaneous, involuntary disorder of pathological trance states.

Patients with dissociative trance disorder may demonstrate high hypnotizability on standard hypnotizability testing, a finding which may be helpful in differential diagnosis.

Treatment

There are no systematic studies of dissociative trance disorder in the absence of pathological possession. In general, treatment should focus on the patient understanding the nature of trance states, and their relationship to more ordinary phenomena such as daydreaming. Assessment of hypnotizability may also allow the clinician to produce the same symptoms the patient is experiencing with hypnotic induction. This may lead to an increased sense of control, and may be sufficient to assist the patient in recognition of autohypnotic states, and having strategies to manage these. In more chronic cases, a supportive-expressive psychotherapy, similar to that for conversion disorder may be most successful, as illustrated by the following case.

OTHER DISSOCIATIVE DISORDERS CASE 3

Ms. M, a married woman in her 50s, was seen in consultation for "passing out attacks." She had had increasingly frequent episodes of grimacing, rapid eye-blinking, and piano-playing finger movements, accompanied by an "eye-roll," showing only the sclera, and falling to the ground in an unresponsive state, often for several hours at a time. Extensive work up was completely normal for seizures, other neurological disorders, cardiac causes, and metabolic causes, among others. Repeated courses of antipsychotic medications and mood-stabilizers for "periodic catatonia" were ineffective and only caused problematic side effects. Anti-depressants helped with the patient's mood, but the episodes persisted. The consultant performed a lengthy diagnostic interview, and gave the patient several screening and diagnostic interviews for dissociative disorders, ruling out dissociative identity disorder and depersonalization/derealization disorder, although the patient reported "watching" herself as if at a distance during the episodes, experiencing them "as if something just comes over me that I cannot control." The consultant discovered many current personal, marital, and family stresses and determined that episodes increased with increased levels of family discord. The patient reported a childhood history of severe verbal violence between her parents and directed at her and her siblings. She coped with this psychologically by "just leaving" and "tuning it all out." The patient was educated about autohypnosis and trance states, and rapidly achieved deep hypnotic trance with a brief induction. Hypnotic interventions were able to reproduce each of the patient's symptoms, as well as a "hidden observer" that remained aware of what was occurring. Over several sessions, the patient was able to learn self-hypnosis to reproduce symptoms and to alleviate them. She was delighted at this improved sense of control and returned to her outpatient providers for work on the family issues. The patient returned to the consultant several years later, disabled by the same symptoms, with unresponsive episodes many times a day. She had stopped doing self-hypnosis, and family problems had also increased. She was admitted to a specialized trauma disorders unit where, initially, she continued to have many episodes a day. Psychotherapy focused on viewing the trance episodes

as a maladaptive coping strategy and challenging the patient to handle problems and difficult affects (anger, shame, sadness) without hypnotic evasions, likening this to an addiction to drugs or alcohol. Milieu interventions involved staff not engaging in efforts to "rescue" the patient from episodes, and helping other patients also not attempt to rescue her by calling staff to intervene. Family therapy focused on psychoeducation and practical solutions to the many issues faced by the family. A strong recommendation was made for outpatient psychotherapy with a clinician proficient in hypnosis, and outpatient family and marital therapy. At the time of discharge, the patient had not had a trance episode for over a week.

References

Brand B, Armstrong JA, Loewenstein RJ. Psychological assessment of patients with dissociative identity disorder. *Psychiatr Clin North Am*. 2006;29(1):145–168.

Brand BL, Loewenstein RJ, Lanius R. Treatment of dissociative identity disorder. In: Gabbard GO, ed. *Gabbard's Treatments of Psychiatric Disorders*. 5th ed. Washington, DC: American Psychiatric Association Press; 2014:439–458.

Brand BL, Loewenstein RJ, Spiegel D. Dispelling myths about dissociative identity disorder treatment: an empirically based approach. *Psychiatry*. 2014;77(2):169–189.

Brand BL, McNary SW, Myrick AC, et al. A longitudinal naturalistic study of patients with dissociative disorders treated by community clinicians. *Psycholo Trauma*. 2012;5(4):301–308.

Brown D, Scheflin AW, Hammond DC. *Memory, Trauma, Treatment, and the Law*. New York: Norton; 1998.

Chalavi S, Vissia EM, Giesen ME, et al. Abnormal hippocampal morphology in dissociative identity disorder and post-traumatic stress disorder correlates with childhood trauma and dissociative symptoms. *Hum Brain Mapp*. 2015;36(5):1692–1704.

Dalenberg CJ, Brand BL, Gleaves DH, et al. Evaluation of the evidence for the trauma and fantasy models of dissociation. *Psychol. Bull*. 2012;138(3):550–588.

Dell PF, O'Neil JA, eds. *Dissociation and the Dissociative Disorders: DSM-V and Beyond*. New York: Routledge; 2009.

DePrince AP, Freyd JJ. Trauma induced dissociation. In: Friedman MJ, Keane TM, Resick PA, eds. *Handbook of PTSD: Science & Practice*. 2nd ed. New York: Guilford; 2014:219–233.

Dorahy MJ, Brand BL, Sar V, et al. Dissociative identity disorder: an empirical overview. *Aust N Z J Psychiatry*. 2014;48(5):402–417.

Hustvedt A. *Medical Muses: Hysteria in Nineteenth-Century Paris*. New York: Norton; 2011.

International Society for the Study of Trauma and Dissociation. Guidelines for the evaluation and treatment of dissociative symptoms in children and adolescents. *J Trauma Dissociation*. 2004;5(3):119–150.

International Society for the Study of Dissociation, Chu JA, Dell PF, et al. Guidelines for treating dissociative identity disorder in adults, third revision. *J Trauma Dissociation*. 2011;12(2):115–187.

Lanius R, Frewen P. *Healing the Traumatized Self: Consciousness, Neuroscience Treatment*. New York: Norton; 2015.

Lanius RA, Vermetten E, Pain C, eds. *The Hidden Epidemic: The Impact of Early Life Trauma on Health and Disease*. Cambridge, UK: Cambridge University Press; 2010.

Loewenstein RJ. An office mental status examination for chronic complex dissociative symptoms

and multiple personality disorder. *Psychiatr Clin North Am*. 1991;14:567–604.

Loewenstein RJ. DID 101: a hands-on clinical guide to the stabilizaton phase of dissociative identity disorder treatment. *Psychiatr Clin N Am*. 2006;29(1):305–332.

Loewenstein RJ. Treatment of dissociative amnesia. In: Gabbard GO, ed. *Gabbard's Treatment of Psychiatric Disorders*. 5th ed. Washington, DC: American Psychiatric Publishing; 2014:471–478.

Putnam FW. *Dissociation in Children and Adolescents: A Developmental Model*. New York: Guilford; 1997.

Simeon D. Treatment of depersonalization/derealization disorder. In: Gabbard GO, ed. *Gabbard's Treatment of Psychiatric Disorders*. 5th ed. Washington, DC: American Psychiatric Publishing; 2014:459–470.

Simeon D, Abugel J. *Feeling Unreal: Depersonalization Disorder and the Loss of Self*. New York: Oxford; 2006.

Spiegel D, Loewenstein RJ, Lewis-Fernandez R, et al. Dissociative disorders in DSM-5. *Depress Anxiety*. 2011;28(9):824–852.

Stein DJ, Koenen KC, Friedman MJ, et al. Dissociation in posttraumatic stress disorder: evidence from the world mental health surveys. *Biol Psychiatry*. 2013;73(4):302–312.

Trickett PK, Noll JG, Putnam FW. The impact of sexual abuse on female development: lessons from a multigenerational, longitudinal research study. *Dev Psychopathol*. 2011;23(2):453–476.

Vermetten E, Dorahy M, Spiegel D, eds. *Traumatic Dissociation*. Washington, DC: American Psychiatric Press; 2007.

21

Normal Sexuality and Sexual Disorders

▲ 21.1 Normal Human Sexuality and Sexual Dysfunctions

VIRGINIA A. SADOCK, M.D.

Sexuality has always been a subject of importance to the medical community. In the classical era, Hippocrates cited the clitoris as the site of female sexual arousal, the first physician historically recorded to have made that assessment. In the middle ages, Islamic physicians recommended coitus interruptus as a form of birth control. At the end of the Renaissance and the beginning of the Reformation, a linen sheath was devised as a condom, not for purposes of birth control, but as protection against syphilis. During the Victorian era, sexologists such as Havelock Ellis (Fig. 21.1–1) and Richard von Krafft-Ebing (Fig. 21.1–2) presented diverging perspectives on sexual behavior. During that same period, Sigmund Freud developed his innovative theories on libido, childhood sexuality, and the effects of the sexual impulse on behavior. In the modern era, the research of Alfred Kinsey, the work of William Masters and Virginia Johnson, and the development of drugs that prevent contraception, aid erection, and replace hormones that decrease with menopause and aging contributed to the development of an era of sexual liberality. Sex has also been a consistent source of motivation and a focus of interest to humankind in general. Depictions of sexual behavior have existed from the time of prehistoric cave drawings through Leonardo da Vinci's anatomical illustrations of intercourse to current pornographic sites on the Internet.

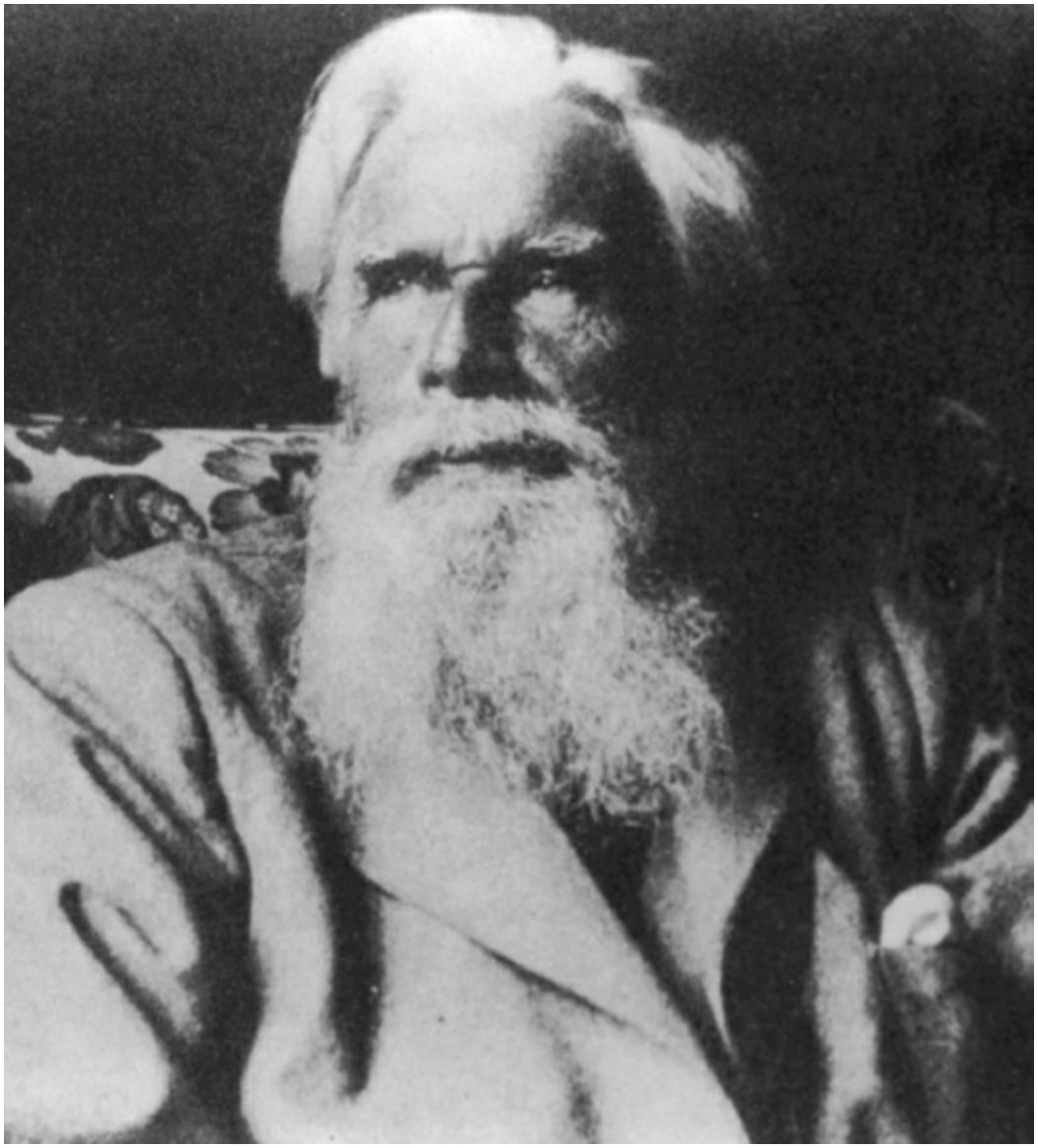

FIGURE 21.1–1. Havelock Ellis (1859–1939). In his book, *Studies in the Psychology of Sex* (1896), Ellis recorded examples of normal and abnormal sexuality. It remains a classic in the field of sexology. (Courtesy of New York Academy of Medicine.)

Sexuality is determined by anatomy, physiology, the culture in which a person lives, relationships with others and developmental experiences throughout the life cycle. It includes the perception of being male or female and private thoughts and fantasies as well as behavior. To the average normal person, sexual attraction to another person and the passion and love that follow are deeply associated with feelings of intimate happiness.

FIGURE 21.1–2. Richard von Krafft-Ebing (1840–1903), a psychiatrist who published a classic text, *Psyhopathia Sexualis* (1898), in which he documented every variation in sexuality, including zoophilia, necrophilia, urolagnia, and lust murder, among others. Case reports were so lurid and detailed that early editions were published in Latin. (Courtesy of New York Academy of Medicine.)

Normal sexual behavior brings pleasure to oneself and one's partner, involves stimulation of the primary sex organs including coitus and often leads to climax; it is devoid of inappropriate feelings of guilt or anxiety and is not compulsive. Recreational, as opposed to relational sex, including casual, premarital and extramarital activity and masturbation is normal in some contexts. Societal understanding of what defines normal sexual behavior is inconstant and varies from era to era, reflecting the cultural mores of the time.

Significant new developments in the field of sexuality include studies of sexual behavior in older adults that have helped breach the taboo against sex in the elderly. Theories on the psychology of sex have examined compulsive sexual behavior—not an official diagnosis in the revised fifth

edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). Clinical reports have demonstrated significant gender differences in response to sexual stimulation. The sequential/linear response model first postulated by Masters and Johnson seems reflective of male sexual response, but studies have indicated it is an insufficient model of the experience of many women. Studies of sex practices in the United States have placed the sexual behavior of Americans in a social context, providing important information both on the ways relationships are likely to form and the types and frequencies of sex practices in this country. Further studies are underway to evaluate the effect of the technology on sex practices and relationships. Additionally, cultural acceptance of homosexuality, first dropped by the APA as a sexual disorder in 1980, has increased substantially, with the majority of Americans currently in favor of same sex marriage which the supreme court legalized in 2015. Also, there is growing acceptance of transgender behavior.

NORMAL SEXUALITY

Anatomical and Physiological Bases

Knowledge about the organs of sexuality and the normal physiological response of men and women is necessary for an informed understanding of the sexual dysfunctions. In the 1960s emphasis was placed on psychological and social contributants to dysfunction; in the 1990s, greater emphasis was placed on the genetic, neuroanatomical, and neurochemical model of human sexuality than on psychological and social factors. Currently, most clinicians believe an interplay of all these factors should be considered in treating sexual dysfunction.

Male Anatomy. The external genitalia of the normal adult man include the penis, scrotum, testes, epididymis, and parts of the vas deferens. Internal components include the vas deferens, ejaculatory ducts, and prostate gland.

Freud referred to the penis as the executive organ of sexuality. Since antiquity, culture has represented the penis in a variety of art forms. In ancient Greece, the cults of Dionysus, Priapus, used the phallus as a recurrent symbol of fertility and rejuvenation. The word *penis* has been traced from the Latin, meaning variously “tail” or “to hang” and refers to the pendant position of the organ in its resting or flaccid state. The size of the penis varies within a fairly constant range, but sex researchers over the years have disagreed on the dimensions of the range. All agree, however, that concern over the size of the penis is practically universal among men. Masters and Johnson report a range of 7 to 11 cm in the flaccid state and 14 to 18 cm in the erect state. Of particular interest was their observation that the flaccid dimension bears little relation to the erect dimension, as the

smaller penis erects proportionally more than the larger one.

Circumcision, a procedure in which the prepuce is surgically removed, has been practiced for centuries as a religious rite by Jews and Moslems and is a common medical procedure in the United States today. The circumcised penis, with its exposed glans, was once believed to be less sensitive because of cornification of the epithelium. In laboratory studies, however, researchers have found no difference in tactile threshold between the circumcised penis and the uncircumcised penis. Intravaginally, the prepuce of the uncircumcised penis remains retracted behind the glans during penile thrusting, dispelling the myth that premature ejaculation may be more common in uncircumcised men because of increased stimulation caused by preputial movements. The most recent recommendation of the American Academy of Pediatrics states that there are clearer medical benefits to circumcision than previously recognized, including lessened ability to transmit cases of AIDS, syphilis and gonorrhea, a lower incidence of penile cancer, and less genitourinary infections in the first year of life. They stop short of recommending circumcision and leave the decision of whether or not to circumcise to parents in consultation with their doctors. However, they encourage insurance companies to cover the procedure.

Ejaculation is the forceful propulsion of semen and seminal fluid from the epididymis, vas deferens, seminal vesicles, and prostate into the urethra. The dilation of the prostatic urethra and the passage of fluid into the penile urethra provide the man with a sensation of impending climax, the emission phase of the ejaculatory process. Indeed, once the prostate contracts, ejaculation is inevitable. The ejaculate is then propelled through the penile urethra by contractions of the striated pelvic and perineal muscles. This phase of ejaculation is essentially under somatic efferent control. The ejaculate consists of approximately 1 teaspoon (2.5 mL) of fluid and contains approximately 120 million sperm cells. It is believed that the larger the ejaculate, the more pleasurable the orgasm, but this belief is highly subjective. The sense of pleasure that accompanies orgasm is thought to be a cortical experience. Men who have had a prostatectomy and therefore no longer ejaculate, report they still experience strong and satisfactory orgasms.

Female Anatomy. The external genitalia of the normal woman, also called the *vulva,* include the mons pubis, major and minor lips, clitoris, glans, vestibule of the vagina, and vaginal orifice. The internal system includes the ovaries, fallopian tubes, uterus, and vagina.

The word *vagina* comes from the Latin word meaning “sheath.” The vagina is usually collapsed, a potential rather than an actual space. Approximately 8 cm long, the vagina extends from the cervix of the uterus above to the vestibule of the vagina or vaginal opening below. In most

virgins, a membranous fold, the hymen, separates the vestibule and opening from the rest of the vaginal canal. The mucous membrane lining the vaginal walls rests in numerous transverse folds. To accommodate the penis during sexual intercourse, the vagina expands in both length and width. After menopause, because circulating estrogen concentrations decrease, the vagina loses much of its elasticity. Masters and Johnson described the clitoris as the primary female sexual organ, because orgasm depends physiologically on adequate clitoral stimulation. Anatomically, the clitoris has a nerve net that is proportionally three times as large as that of the penis.

Alfred Kinsey found that when women masturbate, most prefer clitoral stimulation. That finding was refined further by Masters and Johnson, who reported that women prefer the shaft of the clitoris to the glans, because the glans is hypersensitive if stimulated excessively.

The clitoral prepuce is contiguous with the labia minora, and, during coitus, the penis does not stimulate the clitoris directly. Rather, penile thrusting exerts traction on the minor lips, which, in turn, stimulate the clitoris sufficiently for orgasm. During heightened excitement, just before orgasm, the clitoris retracts under the clitoral hood because of contraction of the ischiocavernosus muscles. Retracting, the clitoris moves away from the vaginal barrel, which makes clitoral–penile contact impossible. The size of the clitoris varies considerably and is unrelated to the sexual responsiveness of a particular woman.

Table 21.1–1.
Responses of Sex Organs to Autonomic Nerve Impulses

	Adrenergic Impulses		Cholinergic Impulses
Effector Organs	**Receptor Type**	**Responses**	**Responses**
Urinary bladder			
Detrusor	β_2	Relaxation (usually)	Contraction
Trigone and sphincter	α_1	Contraction	Relaxation
Ureter			
Motility and tone	α_1	Increase	Increase (?)
Uterus	α_1, β_2	Pregnant: contraction (α_1), relaxation (β_2); nonpregnant: relaxation (β_2)	Variable
Sex organs, male	α_1	Ejaculation	Erection
Skin			
Pilomotor muscles	α_1	Contraction	
Sweat glands	α_1	Localized secretion	Generalized

secretion

Adapted from Goodman Gilman A, Rall TW, Nies AS, Taylor P, eds. *Goodman and Gilman's The Pharmacological Basis of Therapeutics.* 8th ed. New York: Pergamon; 1990.

In 1950, Ernst Graefenberg described an area surrounding the female urethra in the anterior wall of the vagina that has come to be called the *G spot*. Approximately 0.5 to 1.0 cm in size, it becomes engorged during sexual stimulation. Many women report that stimulation of the area is highly pleasurable and, in some, can induce orgasm. Graefenberg believed that the tissue here was analogous to the prostate and might account for the spurt of fluid during orgasm reported by some women, similar to male ejaculation. Some women state that a vaginal orgasm is a different experience for them, providing different sensations than a clitoral orgasm. One explanation is that their pleasure is enhanced by the sense of closeness engendered by coitus. Another hypothesis is that these women climax in response to coital stimulation of the G spot.

Innervation of Sex Organs. Innervation of the sexual organs is mediated primarily through the autonomic nervous system (ANS). Penile tumescence occurs through the synergistic activity of two neurophysiological pathways. A parasympathetic (cholinergic) component mediates reflexogenic erections via impulses that pass through the pelvic splanchnic nerves (S2, S3, and S4). A thoracolumbar, mainly sympathetic pathway transmits psychologically induced impulses. Both parasympathetic and sympathetic mechanisms are thought to play a part in relaxing the smooth muscles of the penile corpora cavernosa, which allows the penile arteries to dilate and causes the inflow of blood that results in penile erection. Relaxation of cavernosal smooth muscles is aided by the release of nitric oxide, an endothelium-derived relaxing factor. Clitoral engorgement and vaginal lubrication also result from parasympathetic stimulation that increases blood flow to genital tissue.

Evidence indicates that the sympathetic (adrenergic) system is primarily responsible for ejaculation. Through the hypogastric plexus, adrenergic impulses innervate the urethral crest, the muscles of the epididymis, and the muscles of the vas deferens, seminal vesicles, and prostate. Stimulation of the plexus causes emission. In women, the sympathetic system facilitates the smooth muscle contraction of the vagina, urethra, and uterus that occurs during orgasm.

The ANS functions outside of voluntary control and is influenced by external events (e.g., stress, drugs) and internal events (hypothalamic, limbic, and cortical stimuli). Considering these influences, it is not surprising that erection and orgasm are so vulnerable to dysfunction (Table 21.1–1).

Endocrinology. From the time of conception, hormones play a major role in human sexual development. Unlike the fetal gonads, which are under chromosomal influence, the fetal external genitalia are very susceptible to hormones. Testosterone mediates the development of the undifferentiated mesodermal wolffian ducts into the male vas deferens, epididymis, and seminal vesicles. Dihydrotestosterone, produced from testosterone, induces the penis and scrotum. The raphe in men corresponds to the anatomical location of the vaginal orifice in women (Fig. 21.1–3).

Exogenous hormonal administration can cause external genital development inconsistent with the fetal sex gland development. For instance, if the pregnant mother receives sufficient exogenous androgen, a female fetus possessing an ovary can develop external genitalia resembling those of a male fetus. Fetal, maternal, or exogenous hormones administered to a pregnant woman may all affect development of the external genitalia of the fetus. Deprived of male and female gonads and the respective hormones, testosterone and estrogen, the human adult does not develop normal secondary sexual characteristics, is incapable of reproduction, and, in the case of the woman, does not develop a menstrual cycle.

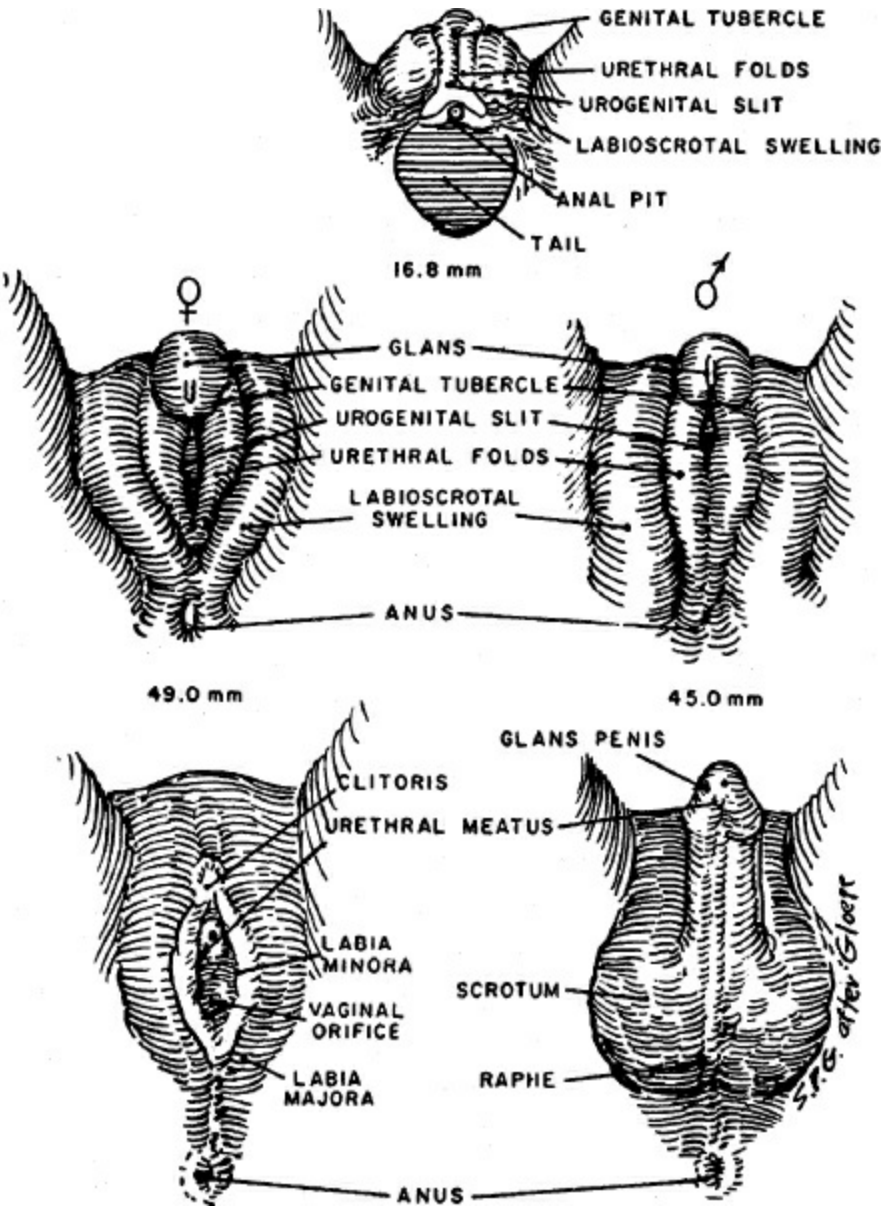

FIGURE 21.1–3. Differentiation of male and female external genitalia from indifferent primordial. Male differentiation occurs only in the presence of androgenic stimulation during the first 12 weeks of fetal life. (From Brobeck JR, ed. *Best and Taylor's Physiological Basis of Medical Practice*. 9th ed. Baltimore, MD: Williams & Wilkins; 1973.)

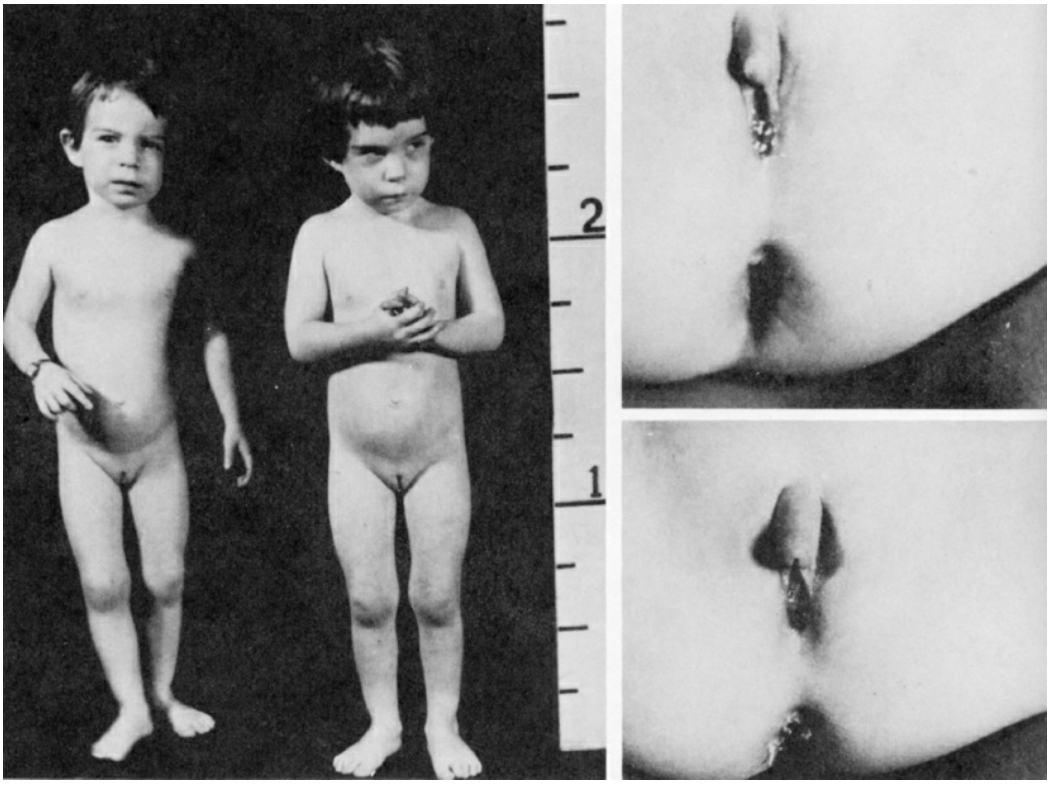

FIGURE 21.1–4. Twins born to a mother who received testosterone during pregnancy. Note the enlarged clitoris in each child. (Courtesy of Robert B. Greenblatt, M.D. and Virginia McNamarra, M.D.)

Testosterone is the hormone believed to be connected with libido in both men and women. In men, stress is inversely correlated with testosterone blood concentration. Other factors, such as sleep, mood, and lifestyle, influence circulating levels of the hormone. The release of testosterone in men is under the control of the hypothalamic–gonadal–pituitary axis. The hormone is secreted in a pulsatile manner and in a diurnal rhythm, with the highest levels occurring in the morning and the lowest levels in the evening. Normal concentrations range from 270 to 1,100 ng/dL. Decreased testosterone concentrations are apparent by age 50 years and decrease at the rate of approximately 100 ng/dL per decade. However, many healthy, aging men never become hypogonadal. It has also been suggested that the sensitivity of androgen receptor sites decreases in aging men.

Recent studies indicate that estrogen is also a factor in the male sexual response and that a decrease in estrogen in the middle-aged male results in greater fat accumulation just as it does in women. Progesterone mildly depresses desire in men and women as do excessive prolactin and cortisol.

Androgen administered to men complaining of loss of potency and loss of libido is usually unsuccessful unless testosterone concentrations are below normal, and administration to women may precipitate disturbing virilization. Many clinicians correct the hormone deficiency of the postmenopausal period with hormone replacement therapy. Testosterone

has been used in combination with estrogen and progesterone in women who do not respond to estrogen alone. The combination is especially useful in treating headache, depression, and reduced libido. Oxytocin is involve in pleasurable sensations during sex and is found in higher levels in men and women following orgasm.

Diethylstilbestrol (DES), an androgenic steroid, was prescribed in the 1950s and 1960s for pregnant women with threatened abortion. However, the drug had untoward effects on the children (especially female children) born to these mothers. Reports of cervical and uterine abnormalities were reported in women and reproductive tract abnormalities were reported in men. The children (called *DES daughters and sons*) organized a National DES Education Program sponsored by the National Cancer Institute to provide information on the potential medical problems confronting those born to DES mothers (Fig. 21.1–4).

Genetics. Genes are involved in gonadal differentiation, resulting in the formation of the bipotential gonad into either a testis or ovary. The best-defined gene in this process is *SRY* (sex-determining region of the Y chromosome), located on the short arm of the Y chromosome. Other genes also play a smaller role, such as *WT1* (Wilms tumor 1), *SF-1* (steroidogenic factor 1), *SOX9, DAX1,* and *MIS-12* (müllerian inhibiting substance 12). Defects or mutations in these genes cause failures in gonadal differentiation that produce clinical syndromes known as *intersex disorders*. Figure 21.1–5 describes some genetic factors involved in the determination of male sex.

Central Nervous System and Sexual Behavior

Cortex. The cortex is involved both in controlling sexual impulses and in processing sexual stimuli that may lead to sexual activity. One study using positron emission tomography (PET) scans to monitor the brain activity of men in their 20s while they were shown various films, including a documentary, a comedy, and a sexually explicit film, found that some areas of the brain were more active during the sex film than they were during the other films. These areas included the orbitofrontal cortex, which is involved in emotions, the left anterior cingulate cortex, which is involved in hormone control and sexual arousal, and the right caudate nucleus, whose activity is a factor in whether sexual activity follows arousal.

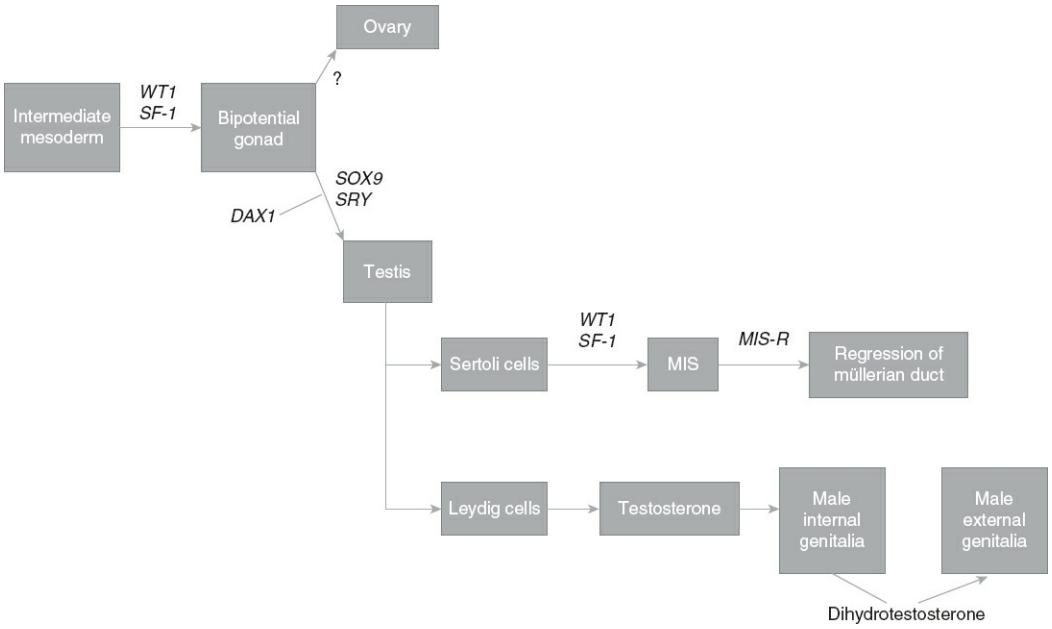

FIGURE 21.1–5. A complex series of steps must occur in gonadal differentiation. A number of genes are critical to appropriate male genital development. *SRY* (sex-determining region of the Y chromosome), a gene on the short arm of the Y chromosome, is a testis-determining factor. The *SOX9* gene is also important in male sexual differentiation. *DAX1,* an orphan member of a nuclear hormone receptor family located on the X chromosome, interacts with steroidogenic factor 1 (*SF-1*). Other genes involved in male gonadal differentiation include the tumor-suppressor gene *WT1* (Wilms tumor 1), and the müllerian inhibiting substance gene (*MIS*) and its receptor, MIS-R.

Limbic System. Experimentation with animals has demonstrated that the limbic system is directly involved with elements of sexual functioning. In all mammals, the limbic system is involved in behavior required for self-preservation and the preservation of the species.

Chemical or electrical stimulation of various sites of the limbic system; the lower part of the septum and the contiguous medial preoptic area, the fimbria of the hippocampus, the mammillary bodies, and the anterior thalamic nuclei have all elicited penile erection.

The hippocampus is believed to influence genital tumescence and affect the regulation of the release of gonadotropins. Stimulation of the amygdala in primates initiates oral (chewing, lip smacking) and then genital (penile erection) behavior. Researchers have stated that the closeness of these functions may derive from the evolutionary fact that the olfactory sense was strongly involved in both feeding and mating. They speculate that the evolution of the third subdivision of the limbic system may reflect a shift in importance from olfactory contact to visual communication in sociosexual behavior.

Studies of women have revealed that those areas activated by emotions of fear and anxiety are notably quiescent when the woman experiences an orgasm.

Brainstem. Brainstem sites exert inhibitory and excitatory control over

spinal sexual reflexes. One site, the nucleus paragigantocellularis, has been identified as important in inhibitory control of climax-like responses in men and is suspected to be involved in female physiology as well. This nucleus projects directly to pelvic efferent neurons in the lumbosacral spinal cord, apparently causing them to secrete serotonin, which is known to inhibit orgasms. The lumbosacral cord also receives strong projections from other serotonergic nuclei in the brainstem, the raphe nuclei, pallidus, magnus, and parapyramidal region.

Brain Neurotransmitters. A vast array of neurotransmitters is produced by the brain, including dopamine, epinephrine, norepinephrine, and serotonin, among others. All affect sexual function. For example, an increase in dopamine is presumed to increase libido. Serotonin produced in the upper pons and midbrain is presumed to have an inhibitory effect on sexual function. Oxytocin, the neurohormone involved in the milk ejection reflex, is released with orgasm and is believed to reinforce pleasurable activities.

Spinal Cord. As described in the innervation of sex organs, sexual arousal and climax are ultimately organized at the spinal level. Sensory stimuli related to sexual function are conveyed via afferents from the pudendal, pelvic, and hypogastric nerves. These afferents terminate in the medial portions of the dorsal horn and in the medial central gray matter of the spinal cord. Several separate experiments suggest that sexual reflexes are mediated by spinal neurons in the central gray region of the lumbosacral segments.

BIOLOGICAL DIFFERENCES BETWEEN THE SEXES

Hormonal Differences

The effects of estrogen in the female and androgen in the male are necessary for the development of ova and spermatozoa, respectively, and for potential fertility. They underlie and define the development of secondary sex characteristics at puberty. Both sexes initially synthesize androgens such as androstenedione or testosterone from acetate or cholesterol in the ovary and testes. Estrogen is then synthesized from the androgen. The two most potent forms of these hormones are estradiol and testosterone.

The defining hormonal difference between women and men is the quantity of these hormones present in the sexes: the testis makes approximately 7,000 ng of testosterone per day and converts one-quarter of 1 percent to estradiol. The ovary makes only 300 ng of testosterone per day and converts 50 percent of it to estradiol. Testosterone production in the ovary may vary during the menstrual cycle, but as testosterone

production increases, so does the percentage of it which is converted to estradiol. Overall, men make 20 times as much androgen as do women and women convert 200 times as much androgen to estradiol as do men. The contrasts are further emphasized because estrogen levels are measured in picograms, while androgen levels are measured in nanograms.

Intracellular hormone metabolism also contributes to sex differentiation. Many cells contain enzymes that can convert androgen to estrogen or testosterone to its more potent metabolite, dihydrotestosterone. Also present in certain cells are proteins which influence the binding of sex hormones to elements affecting transcription sites in relevant DNA.

Ultimately, each tissue can develop an individual androgenic or estrogenic identity which might not be assumed, predicted, or distinguished on the basis of circulating hormone levels. These cellular identities may be particularly important in the brain and may have relevance to sexual behavior, sexual identity, partner choice, and brain dimorphism.

Genetic Differences

Human cells contain 23 pairs of chromosomes for a total of 46. There are 22 pairs of autosomes and 1 pair of sex chromosomes. The sex chromosomes are the X (female) chromosome and the Y (male) chromosome. These chromosomes determine biological sex. Genetically, the normal male has a chromosomal 46XY karyotype and the normal female a 46XX karyotype. The X chromosome carries approximately 1,090 genes, while the human Y chromosome contains only about 80 genes. Further, the Y chromosome contains a unique class of genes in the male-specific region, particularly associated with testicular function. Also, the genes on the Y chromosome have twice the mutation frequency of those on the X chromosome.

The startling difference in the number of genes carried by a 2X woman as contrasted with a single X man is moderated by the fact that the second X chromosome in women is largely inactivated sometime before there are 20 cells in the blastocyst. However, 15 percent of the genes on the second X escape inactivation and another 20 percent are only partially inactivated. The inactivated X chromosome is converted from a linear X into the Barr body, a little ball of genetically inert DNA.

Sex determination is induced by the different karyotypes. In the male, testis growth is induced by the sex determining region of the Y chromosome. In the female, the follicle surrounding her oocytes is stabilized by the second X. Further differences in the ways the XX and XY genotypes influence early differentiation and later development are being studied.

PSYCHOSEXUALITY

Sexuality and total personality are so entwined that it is virtually impossible to speak of sexuality as a separate entity. The term *psychosexual* is therefore used to describe personality development and functioning, as these are affected by sexuality. *Psychosexual* applies to more than sexual feelings and behavior, and it is not synonymous with libido in the broad Freudian sense.

Freud's generalization that all pleasurable impulses and activities are originally sexual has given lay people a somewhat distorted view of sexual concepts and has given psychiatrists a confused picture of motivation. For example, some oral activities are directed toward obtaining food, and others are directed toward achieving sexual gratification. Both activities are pleasure-seeking and use the same organs, but they are not, as Freud contended, both necessarily sexual. Labeling all pleasure-seeking behaviors *sexual* obviates specifying precise motivation. People may also use sexual activities to gratify nonsexual needs, such as dependency, aggression, power, and status. Although sexual and nonsexual impulses may jointly motivate behavior, the analysis of behavior depends on understanding the underlying individual motivation and their interactions. Sexuality is something more than physical sex, coital or noncoital, and something less than all behaviors directed toward gaining pleasure.

Psychosexual Factors

Sexuality depends on four interrelated psychosexual factors: sexual identity, gender identity, sexual orientation, and sexual behavior. These factors affect personality, development, and functioning.

Sexual identity is the pattern of a person's biological sexual characteristics: chromosomes, external and internal genitalia, hormonal composition, gonads, and secondary sex characteristics. In normal development, these characteristics form a cohesive pattern that leaves individuals in no doubt about their sex.

Gender identity is an individual's sense of maleness or femaleness. By the age of 2 or 3 years, most children have a firm conviction that "I am a boy" or "I am a girl." Gender identity results from an almost infinite series of clues derived from experiences with family members, peers, and teachers, and from cultural phenomena. For instance, male infants tend to be handled more vigorously and female infants tend to be cuddled more. Fathers tend to be more aware of their sons' adolescent concerns than of their daughters' anxieties. Boys are more likely to be physically disciplined than girls are. A child's sex may affect parental tolerance for aggression and for reinforcement or extinction of activity and of intellectual, aesthetic, and athletic interests. Physical characteristics derived from a person's biological sex (e.g., physique, body shape, and physical dimensions) interrelate with

an intricate system of stimuli, including rewards, punishment, and parental gender labels, to establish gender goals. Studies of children with intersex conditions have drawn attention to a physiological basis for gender identity. In particular, they have focused on the masculinization or feminization of the fetal brain. Some people claim a gender identity at odds with their biological attributes. Others possess a fluid gender identity or one that does not fit into any recognized category which they call "other."

Gender Role. Related to, and in part derived from, gender identity is gender role behavior. John Money and Anke Ehrhardt described gender role behavior as all those things that a person says or does to disclose himself or herself as having the status of boy or man, girl or woman, respectively. A gender role is not established at birth, but it is built up cumulatively through (1) experiences encountered and transacted through casual and unplanned learning, (2) explicit instruction and inculcation, and (3) spontaneously putting two and two together to make sometimes four and sometimes five. The usual outcome is a congruence of gender identity and gender role. Although biological attributes are significant, the major factor in achieving the role the culture considers appropriate to a person's sex is learning.

Table 21.1–2.
Female Sexual Response Cycle[a]

Organ	Excitement Phase	Orgasmic Phase	Resolution Phase
Mental	Lasts several minutes to several hours; heightened excitement before orgasm, 30 seconds to 3 minutes	3–15 seconds	10–15 minutes; if no orgasm, 0.5–1 day
Skin	Just before orgasm: sexual flush inconsistently appears; maculopapular rash originates on abdomen and spreads to anterior chest wall, face, and neck; can include shoulders and forearms	Well-developed flush	Flush disappears in reverse order of appearance; inconsistently appearing film of perspiration on soles of feet and palms of hands
Breasts	Nipple erection in two-thirds of women, venous congestion and areolar enlargement; size increases to one-fourth more than normal	Breasts may become tremulous	Return to normal in approximately 0.5 hour
Clitoris	Enlargement in diameter of glans and shaft; just before orgasm, shaft retracts into prepuce	No change	Shaft returns to normal position in 5–10 seconds; detumescence in 5–30 minutes; if no orgasm, detumescence takes several hours
Labia	Nullipara: elevate and flatten against	No change	Nullipara: increase to

majora	perineum Multipara: congestion and edema		normal size in 1–2 minutes Multipara: decrease to normal size in 10–15 minutes
Labia minora	Size increase two to three times more than normal; change to pink, red, deep red before orgasm	Contractions of proximal labia minora	Return to normal within 5 minutes
Vagina	Color change to dark purple; transudate appears 10–30 seconds after arousal; elongation and ballooning; lower third constricts before orgasm	3–15 Contractions of lower third at intervals of 0.8 seconds	Ejaculate forms seminal pool in upper two thirds; congestion disappears in seconds or, if no orgasm, in 20–30 minutes
Uterus	Ascends into false pelvis; labor-like contractions begin in heightened excitement just before orgasm	Contractions throughout orgasm	Contractions cease, and uterus descends to normal position
Other	Myotonia	Loss of voluntary muscular control	Return to baseline status in seconds to minutes
	A few drops of mucoid secretion from Bartholin glands during heightened excitement	Rectum: rhythmic contractions of sphincter	Cervix color and size return to normal, and cervix descends into seminal pool
	Cervix swells slightly and is passively elevated with uterus	Hyperventilation and tachycardia	

[a]A desire phase consisting of sex fantasies and desire to have sex often precedes the excitement phase.

Sexual orientation describes the object of a person's sexual impulses: heterosexual (opposite sex), homosexual (same sex), or bisexual (both sexes). The majority of people have a heterosexual orientation. In the United States, 2.8 percent of men and 1.4 percent of women identify themselves as homosexual. These numbers are compatible with figures from Western European countries as well. However, a higher percentage of persons have had at least one same-sex experience in their lives. Additionally, homosexuals congregate in urban areas, so the incidence of homosexuality in some large cities is as high as 8 or 9 percent. A group of people have defined themselves as "asexual" and assert this as a positive identity. Some researchers believe this lack of attraction to any object is a manifestation of a desire disorder. Other people wish not to define their sexual orientation at all and avoid labels. Still others describe themselves as polysexual or pansexual.

Sexual behavior includes desire, receptivity, fantasies, pursuit of partners, autoeroticism, and all the activities engaged in to express and gratify sexual needs. It is an amalgam of psychological and physiological responses to internal and external stimuli.

Physiological Responses

Sexual response is a true psychophysiological experience. Arousal is triggered by both psychological and physical stimuli; levels of tension are experienced both physiologically and emotionally; and with orgasm, normally a subjective perception of a peak of physical reaction and release occurs along with a feeling of well-being. Physical health, psychosexual development, psychological attitudes toward sexuality and attitudes toward one's sexual partner are directly involved with, and affect, the physiology of human sexual response.

Normally, men and women experience a sequence of physiological responses to sexual stimulation. In the first detailed description of these responses, Masters and Johnson observed that the physiological process involves increasing levels of vasocongestion and myotonia (tumescence) and the subsequent release of the vascular activity and muscle tone as a result of orgasm (detumescence). In men an ejaculation is a normal part of the climax.

Refractory Period. Following orgasm, men experience a refractory period during which they cannot be restimulated to further arousal. The length of this period is age and situation dependent, with younger age and novel situations usually resulting in a shorter refractory period. Women do not have a refractory period and can experience multiple, successive orgasms.

Tables 21.1–2 and 21.1–3 describe the physiologic changes in the male and female that accompany sexual arousal, orgasm and, subsequently, resolution back to the pre-stimulation baseline.

It is important to remember that the sequence of responses can overlap and fluctuate. In addition, a person's subjective experiences are as important to sexual satisfaction as the objective physiologic response. Figures 21.1–6 and 21.1–7 illustrate several possible patterns in the phases of the male sexual response and female sexual response, respectively.

Gender Differences in Desire and Erotic Stimuli

Sexual impulses and desire exist in men and women. In measuring desire by the frequency of spontaneous sexual thoughts, interest in participating in sexual activity, and alertness to sexual cues, males generally possess a higher baseline of desire than do women, which may be biologically determined. In the male, a desire phase, consisting of sex fantasies and the desire to have sex usually motivates initiation of the sex act and precedes arousal. Motivations for having sex, other than desire, exist in both men and women, but seem to be more varied and prevalent in women. In women they may include a wish to reinforce the pair bond, the need for a feeling of closeness, a way of preventing the man from straying, or a desire to please the partner. Also, women, particularly those in long-term

relationships, may experience desire synchronously with, or after the onset of, arousal, rather than before it. Receptivity to the sex act seems to be as important a prelude to sexual behavior in women, as desire.

Table 21.1–3.
Male Sexual Response Cycle[a]

Organ	Excitement Phase	Orgasmic Phase	Resolution Phase
Mental	Lasts several minutes to several hours; heightened excitement before orgasm, 30 seconds to 3 minutes	3–15 seconds	10–15 minutes; if no orgasm, 0.5 to 1 day
Skin	Just before orgasm: sexual flush inconsistently appears; maculopapular rash originates on abdomen and spreads to anterior chest wall, face, and neck, and can include shoulders and forearms	Well-developed flush	Flush disappears in reverse order of appearance; inconsistently appearing film of perspiration on soles of feet and palms of hands
Penis	Erection in 10–30 seconds caused by vasocongestion of erectile bodies of corpus cavernosa of shaft; loss of erection may occur with introduction of asexual stimulus—e.g., loud noise; with heightened excitement, size of glans and diameter of penile shaft increase further	Ejaculation: emission phase marked by three to four contractions of 0.8 seconds of vas, seminal vesicles, prostate; ejaculation proper marked by contractions of 0.8 seconds of urethra and ejaculatory spurt of 12–20 in at age 18, decreasing with age to seepage at age 70	Erection: partial involution in 5–10 seconds with variable refractory period; full detumescence in 5–30 minutes
Scrotum and testes	Tightening and lifting of scrotal sac and elevation of testes; with heightened excitement, 50% increase in size of testes over unstimulated state and flattening against perineum, signaling impending ejaculation	No change	Decrease to baseline size because of loss of vasocongestion; testicular and scrotal descent within 5–30 minutes after orgasm; involution may take several hours if no orgasmic release takes place
Cowper glands	2–3 Drops of mucoid fluid that contain viable sperm are secreted during heightened excitement	No change	No change
Other	Breasts: inconsistent nipple erection with heightened excitement before orgasm Myotonia: semispastic contractions of facial,	Loss of voluntary muscular control Rectum: rhythmic contractions of sphincter Heart rate: up to 180 beats/min Blood pressure: up to 40–100	Return to baseline state in 5–10 minutes

abdominal, and intercostal muscles	mm systolic; 20–50 mm diastolic
Tachycardia: up to 175 beats/min	
Blood pressure: rises to 20–80 mm systolic; 10–40 mm diastolic	Respiration: up to 40 respirations/min
Respiration: increased	

[a]A desire phase consisting of sex fantasies and desire to have sex precedes the excitement phase.

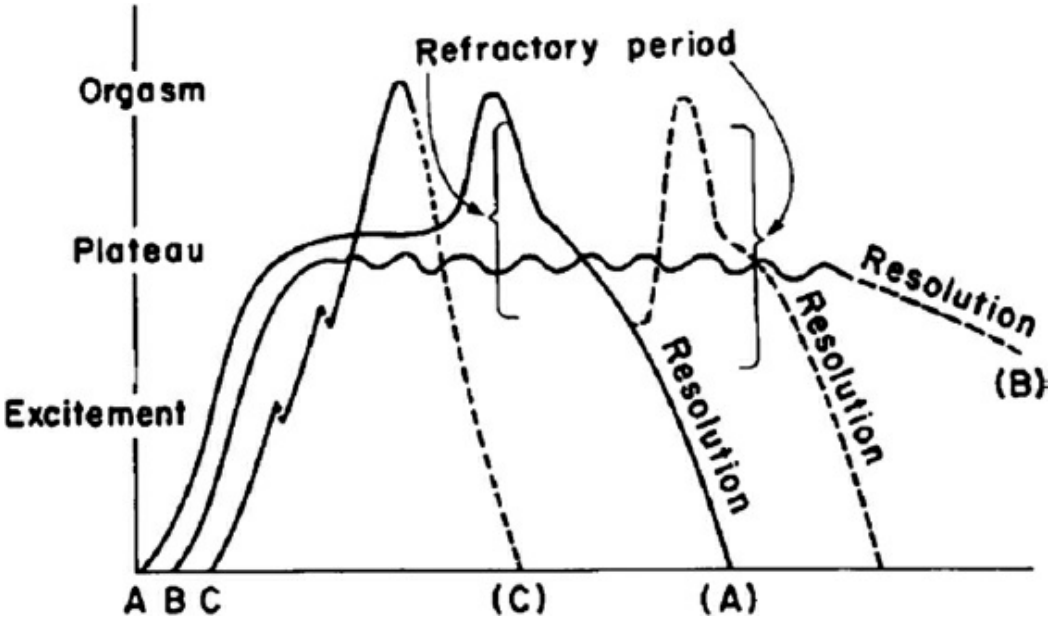

FIGURE 21.1–6. Male sexual response. An individual man may experience any of these three patterns (**A, B,** or **C**) during a particular sexual experience. (From Walker JI, ed. *Essentials of Clinical Psychiatry*. Philadelphia, PA: JB Lippincott; 1985:276.)

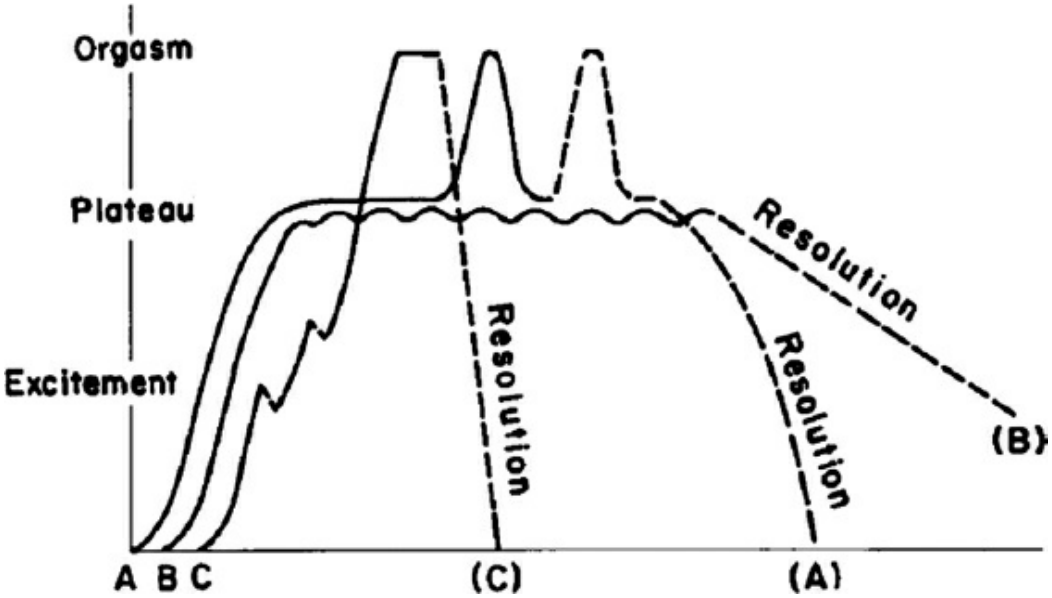

FIGURE 21.1–7. Female sexual response. An individual woman may experience any of these three patterns (**A, B,** or **C**) during a particular sexual experience. (From Walker JI, ed. *Essentials of Clinical Psychiatry*. Philadelphia, PA: JB Lippincott; 1985:276.)

Although explicit sexual fantasies are common to both sexes, the external stimuli for the fantasies frequently differ for men and women. Many men respond sexually to visual stimuli of nude or barely dressed women. Women report responding sexually to romantic stories with a hero whose passion for the heroine impels him toward a lifetime commitment to her. The popular erotic novel, Fifty Shades of Gray, evolves with just such a plot.

A complex component of female sexuality is that a woman's subjective sense of arousal is not always congruent with her physiologic state of arousal. Specifically, her sense of excitement may reflect a readiness to be aroused rather than physiological lubrication. Conversely, she may experience signs of arousal, including vaginal lubrication without being aware of them, or without wanting to act on them. These instances rarely occur in men.

Masturbation

Masturbation usually is a normal precursor of object-related sexual behavior and a form of sexual pleasure that generally lasts throughout a person's lifetime. No other form of sexual activity has been as universally practiced in spite of being severely condemned by many cultures for long periods of time.

In the classical period, Greco-Roman writers and medical authorities, such as Galen, recommended masturbation as a healthful practice for both men and women. The categorization of masturbation as sinful in Western culture derives from Judeo-Christian attitudes. Masturbation is sometimes called *onanism* after Onan in the Old Testament, who was slain for spilling his seed upon the ground. Many biblical scholars today posit that this punishment was not a result of his masturbation, but because he did not obey Jehovah's commandment to take his brother's wife as his own. The prohibition against masturbation was reinforced by Christian Church fathers, particularly by St. Augustine, who preached celibacy and held that sex was appropriate only for purposes of procreation. Ambivalence about masturbation also existed in Eastern cultures. Some Hindus and the ancient Chinese believed that loss of semen resulted in reduction of a vital male essence. However, the prohibitive emphasis was on ejaculation and not on manipulation of the genitalia. Female masturbation was better tolerated or ignored.

For many years, science held attitudes toward masturbation that were as negative as many religious views. What had been viewed as sinful came to be seen as pathological (Fig. 21.1–8). For example, in the 1800s, Krafft-Ebing believed that masturbation could lead to insanity. Currently, and in the second half of the 20th century, a liberal attitude toward sexual behavior, including masturbation, has prevailed. Research by Kinsey in the 1940s into the prevalence of masturbation indicated that nearly all men

and three-fourths of all women masturbate sometime during their lives.

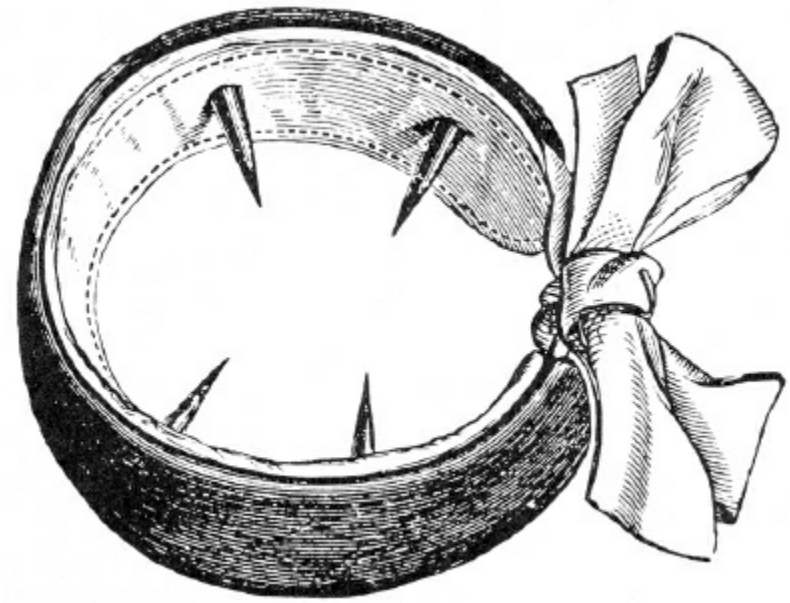

FIGURE 21.1–8. A four-pointed penile ring, which was used to prevent masturbation during the 19th century. This engraving shows how the ring was tied to the penis. (From Sadock BJ, Kaplan HI, Freedman AM. *The Sexual Experience*. Baltimore, MD: Williams & Wilkins; 1976.)

Longitudinal studies of development show that sexual self-stimulation is common in infancy and childhood. Just as infants learn to explore the functions of their fingers and mouths, they learn to do the same with their genitalia. At about 15 to 19 months of age, both sexes begin genital self-stimulation. Pleasurable sensations result from any gentle touch to the genital region. Those sensations, coupled with ordinary desire for exploration of the body, produce a normal interest in masturbatory pleasure at that time. Children also develop an increased interest in the genitalia of others—parents, children, and even animals. As youngsters acquire playmates, the curiosity about their own and others' genitalia motivates episodes of exhibitionism or genital exploration. Such experiences, unless blocked by guilty fear, contribute to continued pleasure from sexual stimulation.

With the approach of puberty, the upsurge of sex hormones, and the development of secondary sex characteristics, sexual curiosity intensifies, and masturbation increases. Adolescents are physically capable of coitus and orgasm, but are usually inhibited to some degree by social restraints. The dual and often conflicting pressures of establishing their sexual identities and controlling their sexual impulses produce a strong physiological sexual tension in teenagers. Masturbation is a normal way to reduce that tension. In general, males learn to masturbate to orgasm earlier than females and masturbate more frequently. An important emotional difference between the adolescent and the youngster of earlier years is the

presence of coital fantasies during masturbation in the adolescent. These fantasies are an important adjunct to the development of sexual identity; in the comparative safety of the imagination, the adolescent learns to perform the adult sex role. This autoerotic activity is usually maintained into the young adult years, when it is normally replaced by coitus.

Couples in a sexual relationship do not abandon masturbation entirely. When coitus is unsatisfactory or is unavailable because of illness, or other unavailability of a partner, self-stimulation can serve an adaptive purpose, combining sensual pleasure and tension release. Masturbation has come to be seen as a separate source of sexual pleasure that is not in conflict with partnered sex.

[Several studies found that in men, orgasm raised the serum prostate-specific antigen (PSA) significantly. Males scheduled for PSA tests should be advised not to masturbate (or have coitus) for at least 7 days prior to the examination.]

Moral taboos against masturbation have generated myths that masturbation causes mental illness, decreased sexual potency, or infertility. No scientific evidence exists to support such claims. Masturbation is a psychopathological symptom only when it becomes a compulsion beyond an individual's willful control. Then, it is a symptom of disturbance, not because it is sexual, but because it is compulsive. It is also symptomatic of a sexual problem when it is the only sexual activity of a person with an available intimate partner. Masturbation is probably a universal aspect of psychosexual development and, in most cases, it is adaptive.

SEXUAL LEARNING AND SEXUAL BEHAVIOR

The sex drive is innate and varies in intensity in different people, but much sexual behavior is learned. Early sexual experiences, particularly those during puberty and early adolescence, can have an imprinting effect. If they are strongly associated with pleasure and release of tension, they are likely to be repeated and the person is conditioned to a particular form of sexual expression. In the normal person, sexual learning and experimentation continue throughout the life cycle, the repertoire of sexual behavior expands, and the behaviors are compatible with cultural norms.

Sexual learning begins in childhood. In a broad sense, learning occurs through parent–child interaction, including the meeting of the infant's needs, cuddling, and the reinforcement or discouragement of gender-associated activities. Cuddling and appropriate physical touching engender emotional security and positive feelings in infants toward their bodies. That physicality lays the groundwork for a healthy body image that is a component of sexual self-esteem. A good body image also derives from mastery of early physical activity and a positive parental approach to tasks such as toilet training.

Genital self-stimulation is a normal activity of babies. It is particularly

pronounced between the ages of 15 and 19 months, and it is part of the general interest of children in their bodies. The activity is reinforced by the pleasurable sensations it produces. Children also learn by watching the interaction of their parents. They observe demonstrations of physical affection (although usually not the sex act itself), and they are sensitive to the sexual undertone of flirtation, bantering, or seductive interchange.

Sexual learning in childhood is adversely affected by exploitative, abusive adults. Incestuous activity, with or without penetration, or sexual abuse by other adults (babysitter, other older relatives, teachers, coaches, etc.) is damaging to the child. It may precipitate sexual dysfunction in adult life, as well as promiscuity and problems with intimacy. It may produce delinquent activities and behavioral problems in adolescence, as well as learning problems in school.

Some children are inappropriately stimulated short of outright incest or genital fondling. They are the objects of extremely seductive parents. Termed *eroticized children,* they frequently become sexually precocious.

Children who view the *primal scene*—the term Freud used to describe a child seeing sex between their parents—may be traumatized. This is particularly the case if they interpret coital sounds and movements as physical aggression between the parents.

Adolescence

With the approach of puberty, the upsurge of sex hormones, and the development of secondary sex characteristics, sexual curiosity is intensified.

In addition to masturbation, adolescents learn sexually through caressing and kissing with partners. In early adolescence, sex play may involve a partner of the same sex for a short period of time for heterosexuals as well as for homosexuals. Conversely homosexuals may experiment with the opposite sex, although this occurs more frequently with lesbians than with homosexual men. Adolescence is also the time when one's body image becomes more definitive and a sense of sexual adequacy and sexual desirability begins to develop. Peer acceptance by the same sex and by the opposite sex is of paramount importance. Engaging in sexual talk and jokes, kissing, stimulating genitalia manually or orally, experimenting with degrees of nudity, and experimenting with different partners or with one partner are part of the process of learning about sexuality. These experiences reinforce the adolescent's sense of being a sexual male or female.

In the past decade, a "hook-up" culture has been described among college students that is replacing traditional dating. It involves casual sex, that is, sex outside a committed or romantic relationship that may occur between friends (friends with benefits), or as a one-night, or occasional-night event. However, a survey contrasting student cohorts from 1988 to

1996 and 2004 to 2012 found respondents from the latter era did not report more sexual partners since age 18, more frequent sex, or more frequent sex partners during the previous year than respondents from the earlier era.

First Coitus

The first coitus is a rite of passage for both sexes. In the United States the majority of people have experienced coitus by young adulthood, by their early 20s. In a study of persons ages 18 to 59, over 95 percent had included coitus in their last sexual interaction.

The young man experiencing intercourse for the first time is vulnerable in his pride and self-esteem. Cultural myths still perpetuate the idea that he should be able to have an erection with no, or little, stimulation and that he should have an easy mastery over the situation, even though it is an act that he has never before experienced. Cultural pressure on the woman with her first coitus reflects remaining cultural ambivalence about her loss of virginity, despite the current era of sexual liberality. This is demonstrated in the statistic that only 50 percent of young women use contraception during their first coitus, and of that 50 percent, an even smaller number use it consistently thereafter. Young women with a history of masturbation are more likely to approach intercourse with positive anticipation and confidence. Studies have shown that young women are more likely to enjoy their first coitus if it occurs in the context of a positive emotional relationship.

Adulthood

By their late 20s, 70 percent of men and 85 percent of women have formed a union, either exclusive cohabitation or marriage. The sexual impulse is catalytic in forming and maintaining adult love relationships. Ideally, a mature sexual relationship encompasses the capacity for intimacy and love for one's partner.

As sexual access ceases to be problematic, more attention is focused on the activity itself. Interference with sexual activity arises from the time and energy required for the pursuit of careers, child rearing, and other family and community obligations. Nonetheless, studies reveal a much higher frequency of sexual interaction among married persons than among single persons. The modal frequency for married persons is three times a month, with many couples relating sexually twice a week. The highest frequency of sexual interaction, four or more times a week, exists among cohabiting couples.

Kissing, intercourse, and masturbation are frequent sexual activities. Many adults have some experience with oral sex. Studies show that, although it is not a regular activity for most couples, many people are likely

to engage in it occasionally throughout their lifetimes. The incidence of the practice of cunnilingus and fellatio are the same; the lifetime prevalence is 75 percent of sexually active people. Anal sex is not part of the repertoire of regular sexual activities for most men and women in the United States. Although a sizable minority (25 percent of men and 20 percent of women) report one experience of anal sex, less repeat the practice. In general, a graph of sexual practices does not form a bell curve. Most curves of sexual behavior are strongly skewed, with many persons indulging in a particular behavior or, conversely, very few people doing so.

Middle Age

During middle age, the frequency of marital intercourse may decline. The rates of interaction depend more on male interest, and the middle-aged man may devote much of his energy to his career at this point. The decline in sexual interaction often reflects de-eroticization of the woman because of her wife–mother role. Familiarity also contributes to a decreased passion, and some men have difficulty connecting marital sexual activity with the erotic scripts they elaborated during their adolescence. Conversely, a couple's experience of each other makes them comfortable and augments their skill at mutual arousal. For women, interest in sexual competence typically increases at this time. A woman's erotic pleasure and sexual commitment are strongly connected to feelings of attachment toward her partner and the security of being in a loving relationship.

With late middle age, the biological drive decreases in intensity. It takes the man longer to reach orgasm, he has a longer refractory period, and he requires more stimulation to achieve an erection. The woman must adjust to the hormonal fluctuations of her perimenopausal years, and she, too, requires more stimulation to become aroused. Although medications are available to address these specific physiological changes, they do not eliminate all need to adjust to the aging process. In terms of family dynamics, children often leave the house for work or further education at this time. This fact plus freedom from concern regarding unwanted pregnancy may enhance sexual activity for persons who have more time to direct attention toward each other and renew their life as a couple.

However, middle age is the period of rising extramarital activity, although the frequencies of such occurrences are low—75 percent of men and 85 percent of women remain faithful throughout the lifetime of their marriages. Daniel Offer has explained that the patterning of extramarital sexual activity for both sexes continues to express earlier patterns of psychosexual development:

> For men, it predominantly has the capacity for detachment that in adolescence was directly related to the pursuit of sexual fantasies and the homosocial validation of masculinity. For women, on the other hand, it resembles a quest for circumstances that justify and confirm a romantic self-image rather than a quest for orgasms.

In the context of a loving, communicative, and committed relationship, a decreased frequency of sexual interaction does not herald the onset of extramarital affairs or threaten the stability of the marriage.

Old Age

In the last two decades, coitus has also been a part of the sexual repertoire of elderly adults, due to the development of sildenafil type drugs, which facilitate erections in men, and hormonally enhanced creams or hormonal pills, which counteract vaginal atrophy in postmenopausal women. Prior to the development of these drugs, and still today, many elderly adults enjoyed gratifying sex play, exclusive of coitus. An aversion to the idea of sex in the elderly persists in today's culture. A man who had coitus with his wife while she was in a nursing home and suffered from a degree of dementia, was accused by his stepchildren of sexual assault and brought to trial. He was found innocent.

An estimated 70 percent of men and 20 percent of women over age 60 years are sexually active. Recent studies have found that three in four people aged 57 to 64, two in four people aged 64 to 75, and one in four people aged 75 to 85 had sexual intercourse in the year before the study. Of interest is that the reported frequency of intercourse, two or three times a month, is the same as the frequency reported by young and middle-aged adults. Expected physiological changes in men include a longer time for erection to occur, decreased penile turgidity, and ejaculatory seepage; in women, decreased vaginal lubrication and vaginal atrophy are associated with lower estrogen levels. Illness and medications can adversely affect sexual behavior.

In studies, approximately half of older sexually active people, aged 59 to 85, report one or more sexual dysfunctions. Women report problems of low desire, difficulty with lubrication, and inability to climax. Among men, erectile disorder is the most commonly reported dysfunction. About 14 percent of older men report using medications or supplements to improve their erections. The factors contributing to a continuing sex life in older adults include the presence of a partner, the absence of severe illness, and a history of sexual activity in young and middle adulthood. Physicians should include sexual behavior in their review of systems with older age patients since the majority of people in this age group will not bring up their concerns unless their doctor initiates the discussion.

LOVE AND INTIMACY

Freud postulated that psychological health could be determined by a person's ability to function well in two spheres, work and love. A person able to give and receive love with a minimum of fear and conflict has the capacity to develop genuinely intimate relationships with others. A desire

to maintain closeness to the love object typifies being in love. Mature love is marked by the intimacy that is a special attribute of the relationship between two persons. When involved in an intimate relationship, the person actively strives for the growth and happiness of the loved person. Sex frequently acts as a catalyst in forming and maintaining intimate relationships. The quality of intimacy in a mature sexual relationship is what Rollo May called "active receiving," in which a person, while loving, permits himself or herself to be loved. May describes the value of sexual love as an expansion of self-awareness, the experience of tenderness, and increase of self-affirmation and pride, and sometimes, at the moment of orgasm, loss of feeling of separateness. In that setting, sex and love are reciprocally enhancing and healthily fused.

Some persons experience conflicts that prevent them from fusing tender and passionate impulses. This can inhibit the expression of sexuality in a relationship, interfere with feelings of closeness to another person, and diminish a person's sense of adequacy and self-esteem. When these problems are severe, they may prevent the formation of, or commitment to, an intimate relationship.

Sex and the Law

Medicine and the law both assess the impact of sexuality on the individual and society and determine what is healthy or legal behavior. Appropriateness or legality of sexual behavior, however, is not always viewed the same way by professionals in both disciplines. The issues at the interface of sexual science and the law often are emotionally charged and reflect cultural divisions about acceptable sexual mores. They include attitudes about abortion, pornography, prostitution, sex education, the homosexual and transgender community, the treatment of sex offenders, and the right to sexual privacy, among other issues. Laws regarding these issues (e.g., criminalization of oral or anal sex by consenting adults, or the need for parental permission by minors who are requesting an abortion) vary from state to state.

Taking a Sex History

A sex history provides important information about patients, regardless of the presence of a sexual disorder or whether that is the patient's chief complaint. The information can be obtained gradually as therapy progresses and through open-ended questions. It is important for the interviewer to be nonjudgmental and sensitive to many patient's embarrassment in discussing sexual issues. As part of the sexual history, information regarding safe sex practices, sexually transmitted diseases, sexual orientation, gender identity, and history of any sexual abuse must be ascertained. The percentage of sexual abuse survivors seen in sex therapy

programs is higher than that in the general population. The outline in Table 21.1–4 provides a guide to areas to be covered and a structure that can be used when time is limited.

Table 21.1–4.
Taking a Sex History

I. Identifying data
   A. Age
   B. Sex
   C. Occupation
   D. Relationship status—single, married, number of times previously married, separated, divorced, cohabiting, serious involvement, casual dating (difficulty forming or keeping relationships should be assessed throughout the interview)
   E. Sexual orientation—heterosexual, homosexual, or bisexual (this may also be ascertained later in the interview)

II. Current functioning
   A. Unsatisfactory to highly satisfactory
   B. If unsatisfactory, why?
   C. Feelings about partner satisfaction
   D. Dysfunctions?—e.g., lack of desire, erectile disorder, inhibited female arousal, anorgasmia, premature ejaculation, retarded ejaculation, pain associated with intercourse (dysfunction discussed below)
      1. Onset—lifelong or acquired
         a. If acquired, when?
         b. Did onset coincide with drug use (medications or illegal recreational drugs), life stresses (e.g., loss of job, birth of child), interpersonal difficulties?
      2. Generalized—occurs in most situations or with most partners
      3. Situational
         a. Only with current partner
         b. In any committed relationship
         c. Only with masturbation
         d. In socially proscribed circumstance (e.g., affair)
         e. In definable circumstance (e.g., very late at night, in parental home, when partner initiated sex play)
   E. Frequency—partnered sex (coital and noncoital sex play)
   F. Desire/libido—how often are sexual feelings, thoughts, fantasies, dreams, experienced (per day, week, etc.)?
   G. Description of typical sexual interaction
      1. Manner of initiation or invitation (e.g., verbal or physical? Does same person always initiate?)
      2. Presence, type, and extent of foreplay (e.g., kissing, caressing, manual, or oral genital stimulation)
      3. Coitus? Positions used?
      4. Verbalization during sex? If so, what kind?
      5. Afterplay? (whether sex act is completed or disrupted by dysfunction); typical activities (e.g., holding, talking, return to daily activities, sleeping)
      6. Feeling after sex: relaxed, tense, angry, loving
   H. Sexual compulsivity? (intrusion of sexual thoughts or participation in sexual activities to a degree that interferes with relationships or work, requires deception, and may endanger the

patient)

III. Past sexual history
   A. Childhood sexuality
      1. Parental attitudes about sex—degree of openness or reserve (assess unusual prudery or seductiveness)
      2. Parents' attitudes about nudity and modesty
      3. Learning about sex
         a. From parents (Initiated by child's questions or parent volunteering information? Which parent? What was child's age?)? Subjects covered (e.g., pregnancy, birth, intercourse, menstruation, nocturnal emission, masturbation)?
         b. From books, magazines, or friends at school or through religious group?
         c. Significant misinformation
         d. Feeling about information
         e. Viewing or hearing primal scene—Reaction?
      4. Viewing sex play or intercourse of person other than parent
      5. Viewing sex between pets or other animals
   B. Childhood sex activities
      1. Genital self-stimulation before adolescence—Age? Reaction if apprehended?
      2. Awareness of self as boy or girl—Bathroom sensual activities (regarding urine, feces, odor, enemas)?
      3. Sexual play or exploration with another child (playing doctor)—Type of activity (e.g., looking, manual touching, genital touching)? Reactions or consequences if apprehended (by whom?)?

I. Adolescence
   A. Age of onset of puberty—development of secondary sex characteristics, age of menarche for girl, wet dreams or first ejaculation for boy (preparation for and reaction to)
   B. Sense of self as feminine or masculine—body image, acceptance by peers (opposite sex and same sex), sense of sexual desirability, onset of coital fantasies
   C. Sex activities
      1. Masturbation—Age begun? Ever punished or prohibited? Method used, accompanying fantasies, frequency (questions about masturbation and fantasies are among the most sensitive for patients to answer)?
      2. Homosexual activities—Ongoing or rare and experimental episodes? Approached by others? If homosexual, has there been any heterosexual experimentation?
      3. Dating—casual or steady; description of first crush, infatuation, or first love
      4. Experiences of kissing, necking, petting ("making out" or "fooling around"), age begun, frequency, number of partners, circumstances, type(s) of activity
      5. Orgasm—When first experienced (may not be experienced during adolescence)? With masturbation, during sleep, or with partner? With intercourse or other sex play? Frequency?
      6. First coitus—age, circumstances, partner, reactions (may not be experienced during adolescence), contraception and/or safe sex precautions used

II. Adult sexual activities (may be experienced by some adolescents)
   A. Premarital sex
      1. Types of sex play experiences—frequency of sexual interactions, types and number of partners
      2. Contraception and/or safe sex precautions used
      3. First coitus (if not experienced in adolescence)—age, circumstances, partner
      4. Cohabitation—age begun, duration, description of partner, sexual fidelity, types of sexual activity, frequency, satisfaction, number of cohabiting relationships, reasons for breakup(s)
      5. Engagement—age; activity during engagement period with fiancé(e), with others; length of

engagement

B. Marriage (if multiple marriages have occurred, explore sexual activity, reasons for marriage, and reasons for divorce in each marriage)
   1. Types and frequency of sexual interaction—describe typical sexual interaction (see above). Satisfaction with sex life? View of partner's feeling.
   2. First sexual experience with spouse—When? What were the circumstances? Was it satisfying? Disappointing?
   3. Honeymoon—Setting, duration, pleasant or unpleasant, sexually active? Frequency? Problems? Compatibility?
   4. Effect of pregnancies and children on marital sex
   5. Extramarital sex—Number of incidents, partner, emotional attachment to extramarital partners? Feelings about extramarital sex
   6. Postmarital masturbation—Frequency? Effect on marital sex?
   7. Extramarital sex by partner—effect on interviewee
   8. Ménage à trois or multiple sex (swinging) (polyamony)
   9. Areas of conflict in marriage (e.g., parenting, finances, division of responsibilities, priorities)

III. Sex after widowhood, separation, divorce—celibacy, orgasms in sleep, masturbation, noncoital sex play, intercourse (number of and relationship to partners), other

IV. Special issues
   A. History of rape, incest, sexual, or physical abuse
   B. Spousal abuse (current)
   C. Chronic illness (physical or psychiatric)
   D. History or presence of sexually transmitted diseases
   E. Fertility problems
   F. Abortions, miscarriages, or unwanted or illegitimate pregnancies
   G. Gender identity conflict—(e.g., transsexualism, wearing clothes of opposite sex)
   H. Paraphilias—(e.g., fetishes, voyeurism, sadomasochism)

Sexual Dysfunctions

The essential features of sexual dysfunctions are inhibition, specifically, an inability to respond to sexual stimulation physiologically or psychologically, or the experience of pain during the sexual act. The pain may have a physiologic, psychological, or mixed etiology. Dysfunction can be defined by disturbance in the subjective sense of pleasure usually associated with sex, or by the objective performance.

According to the 10th revision of the *International Statistical Classification of Diseases and Health Related Problems* (ICD-10), sexual dysfunction refers to a person's inability "to participate in a sexual relationship as he or she would wish." The dysfunction is expressed as a lack of desire or of pleasure or as a physiological inability to begin, maintain, or complete sexual interaction. Because sexual response is psychosomatic, it may be difficult to determine "the relative importance of psychological and/or organic factors."

Definition

In the *Diagnostic and Statistical Manual of Mental Disorders, fifth edition* (DSM-5) sexual dysfunctions include male hypoactive sexual desire

disorder, female sexual female interest/arousal disorder, erectile disorder, female orgasmic disorder, delayed ejaculation, early (premature) ejaculation, genito-pelvic pain/penetration disorder, substance/medication-induced dysfunction, other specified sexual dysfunction and unspecified sexual dysfunction. Sexual dysfunctions are diagnosed only when they are a major part of the clinical picture. If more than one dysfunction exists they should all be diagnosed. Sexual dysfunctions can be lifelong or acquired, generalized or situational, and result from psychological factors, physiological factors, and combined factors, and numerous stressors including prohibitive social mores, health and partner issues, and relationship conflicts. If the dysfunction is attributable entirely to a medical condition, substance use, or adverse effects of medication, then sexual dysfunction due to a general medical condition or substance-induced dysfunction should be diagnosed. In DSM-5, the specification of severity of the dysfunction is indicated by noting whether the symptomatology or the patient's distress is mild, moderate, or severe.

Sexual dysfunctions are frequently associated with other mental disorders, such as depressive disorders, anxiety disorders, personality disorders, and schizophrenia. In many instances the sexual dysfunction may be diagnosed in conjunction with another psychiatric disorder. If the dysfunction is largely attributable to an underlying psychiatric disorder, only the underlying disorder should be diagnosed. Sexual dysfunctions are usually self-perpetuating, with the patient increasingly subject to ongoing performance anxiety, and a concomitant inability to experience pleasure. In relationships, the sexually functional partner often reacts with distress or anger due to feelings of deprivation or a sense that he or she is an insufficiently attractive or adequate sexual partner. In such cases, the clinician must consider whether the sexual problem preceded or arose from relationship difficulties and weigh whether a diagnosis of sexual dysfunction relevant to relationship issues is more appropriate.

The sexual problem offers the psychiatrist a perspective on the person's capacity to give and receive pleasure, to love and be loved, to be psychologically intimate and to maintain some degree of sexuality through the vicissitudes of life.

Similar disorders exist among both homosexual and heterosexual partners, with variations imposed by anatomical differences. For example, although penile–vaginal dysfunction cannot be described among homosexuals, penile–anal dysfunction may occur. Regardless of sexual orientation, the physiological response to sexual stimulation applies equally to same-sex and heterosexual partners, and the methods and principles for treatment are essentially similar.

Rating Scales

Several questionnaires are available to assess sexual function, in addition to the assessments made following the taking of a sexual history. They were developed primarily to evaluate illness or medication-related effects on sexual functioning. The most commonly used scales are the Arizona Sexual Experience Scale, the Brief Sexual Function Questionnaire, the Changes in Sexual Functioning Questionnaire, the Derogatis Sexual Function Inventory, and the Rush Sexual Inventory. The scales vary in length, reliability, validity, and method of administration. Some are rated by the patients themselves, others by the therapist. The formats include structured and semistructured approaches. The scales differ in number of symptoms assessed, the dysfunctions they target, and the time frame they cover.

DESIRE, INTEREST, AND AROUSAL DISORDERS

Male Hypoactive Sexual Desire Disorder

This dysfunction is characterized by a deficiency or absence of sexual fantasies and desire for sexual activity for a minimum duration of approximately 6 months (Table 21.1–5). Men for whom this is a lifelong condition have never experienced many spontaneous erotic/sexual thoughts. Minimal spontaneous sexual thinking or minimal desire for sex ahead of sexual activity is not considered a diagnosable disorder in women, particularly if desire is triggered during the sexual encounter. Also, some men define themselves as being asexual. They are not distressed by a lack of desire and feel they are being erroneously pathologized if this diagnosis is applied to them.

The reported prevalence of low desire is greatest at the younger and older ends of the age spectrum, with only 2 percent of men ages 26 to 44 affected by this disorder. A reported 6 percent of men ages 18 to 24 and 40 percent of men ages 66 to 74, have problems with sexual desire. Some men may confuse decreased desire with decreased activity. Their erotic thoughts and fantasies are undiminished, but they no longer act on them due to health issues, unavailability of a partner, or concern about another sexual dysfunction such as erectile disorder.

A variety of causative factors are associated with low sexual desire. Patients with desire problems often use inhibition of desire defensively to protect against unconscious fears about sex. Sigmund Freud conceptualized low sexual desire as the result of inhibition during the phallic psychosexual phase of development and of unresolved oedipal conflicts. Some men, fixated at the phallic stage of development, are fearful of the vagina and believe they will be castrated if they approach it. Freud called this concept vagina dentata. He theorized that men avoid contact with the vagina when they unconsciously believe that the vagina has teeth. Lack of desire can also result from chronic stress, anxiety, or depression.

The presence of desire depends on several factors: biological drive, adequate self-esteem, previous good experiences with sex, the availability of an appropriate partner, and a good relationship in nonsexual areas with one's partner. Damage to any of those factors may result in diminished desire.

Hypoactive *sexual* desire disorders often become manifest during puberty and may be a lifelong condition. A general medical workup should be conducted to rule out a medical cause, which, if present, would be, according to DSM-5, diagnosed as hyposexual desire disorder due to a general medical condition.

Table 21.1–5.
DSM-5 Diagnostic Criteria for Male Hypoactive Sexual Desire Disorder

A. Persistently or recurrently deficient (or absent) sexual/erotic thoughts or fantasies and desire for sexual activity. The judgment of deficiency is made by the clinician, taking into account factors that affect sexual functioning, such as age and general and sociocultural contexts of the individual's life.
B. The symptoms in criterion A have persisted for a minimum duration of approximately 6 months.
C. The symptoms in criterion A cause clinically significant distress in the individual.
D. The sexual dysfunction is not better explained by a nonsexual mental disorder or as a consequence of severe relationship distress or other significant stressors and is not attributable to the effects of a substance/medication or another medical condition.

Specify whether:
Lifelong: The disturbance has been present since the individual became sexually active.
Acquired: The disturbance began after a period of relatively normal sexual function.

Specify whether:
Generalized: Not limited to certain types of stimulation, situations, or partners.
Situational: Only occurs with certain types of stimulation, situations, or partners.

Specify current severity:
Mild: Evidence of mild distress over the symptoms in criterion A.
Moderate: Evidence of moderate distress over the symptoms in criterion A.
Severe: Evidence of severe or extreme distress over the symptoms in criterion A.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

A 20-year-old college student presented with a complaint of insufficient desire for sex. He stated he was bisexual but complained, "I wish I could be more interested in women." During the course of the interview he revealed that he had sexual fantasies about men dating from his adolescence and that he masturbated regularly to erotic visual stimuli involving men. He had fears of erectile dysfunction prior to coitus with women, but did not experience these fears prior to sexual interaction with men. These interactions involved manual and oral sex play, but no penetration. He connected with men online and had not disclosed his sexual interest in or contact with men to anyone. This patient did not suffer from hyposexual desire disorder. His problems arose from conflicted feelings about his sexual orientation.

A 44-year-old married man had a history of good sexual interaction with his wife of 18 years, until 8 months prior to his consultation. Nine months ago he had sought help for a depression triggered by the loss of employment, financial stress, and lack of immediate prospects for a

new job. He had become lethargic, and withdrawn, and was doing little to pursue new work. He was treated with fluoxetine, 20 mg per day which helped him and currently, he was actively networking in pursuit of employment. However, his sexual interest remained diminished. His diminished libido appeared to be a side effect of his medication, and his diagnosis was pharmacologically induced hyposexual desire disorder.

Female Sexual Interest/Arousal Disorder

The combination of interest (or desire) and arousal into one dysfunction category reflects the recognition that women do not necessarily move stepwise from desire to arousal, but often experience desire synchronously with, or even following beginning feelings of arousal. This is particularly true for women in long-term relationships. As a corollary, women experiencing sexual dysfunction may experience either/or both inability to feel interest or arousal, and they may often have difficulty achieving orgasm or experience pain in addition. Some may experience dysfunction across the entire range of sexual response/pleasure. Complaints in this dysfunction category present variously as a decrease or paucity of erotic feelings, thoughts, or fantasies; a decreased impulse to initiate sex; a decreased or absent receptivity to partner overtures; or an inability to respond to partner stimulation (Table 21.1–6).

A complicating factor in this diagnosis is that a subjective sense of arousal is often poorly correlated with genital lubrication in both normal and dysfunctional women. Therefore, complaints of lack of pleasure are sufficient for these diagnoses even when vaginal lubrication and congestion are present. A woman complaining of lack of arousal may lubricate vaginally, but may not experience a subjective sense of excitement. Some studies, using functional magnetic resonance imaging (fMRI) have revealed a low correlation between brain activation in areas controlling genital response and simultaneous ratings of subjective arousal. Physiological studies of sexual dysfunctions indicate that a hormonal pattern may contribute to responsiveness in women who have arousal dysfunction. William Masters and Virginia Johnson found that women are particularly desirous of sex before the onset of the menses. Other women report feeling the greatest sexual excitement immediately after the menses or at the time of ovulation. Alterations in testosterone, estrogen, prolactin, and thyroxin levels have been implicated in female sexual arousal disorder. In addition, medication with antihistaminic or anticholinergic properties causes a decrease in vaginal lubrication.

Table 21.1–6.
DSM-5 Diagnostic Criteria for Female Sexual Interest/Arousal Disorder

A. Lack of, or significantly reduced, sexual interest/arousal, as manifested by at least three of the following:
   1. Absent/reduced interest in sexual activity.

2. Absent/reduced sexual/erotic thoughts or fantasies.
3. No/reduced initiation of sexual activity, and typically unreceptive to a partner's attempts to initiate.
4. Absent/reduced sexual excitement/pleasure during sexual activity in almost all or all (approximately 75–100%) sexual encounters (in identified situational contexts or, if generalized, in all contexts).
5. Absent/reduced sexual interest/arousal in response to any internal or external sexual/erotic cues (e.g., written, verbal, visual).
6. Absent/reduced genital or nongenital sensations during sexual activity in almost all or all (approximately 75–100%) sexual encounters (in identified situational contexts or, if generalized, in all contexts).

B. The symptoms in criterion A have persisted for a minimum duration of approximately 6 months.
C. The symptoms in criterion A cause clinically significant distress in the individual.
D. The sexual dysfunction is not better explained by a nonsexual mental disorder or as a consequence of severe relationship distress (e.g., partner violence) or other significant stressors and is not attributable to the effects of a substance/medication or another medical condition.

Specify whether:
Lifelong: The disturbance has been present since the individual became sexually active.
Acquired: The disturbance began after a period of relatively normal sexual function.
Specify whether:
Generalized: Not limited to certain types of stimulation, situations, or partners.
Situational: Only occurs with certain types of stimulation, situations, or partners.
Specify current severity:
Mild: Evidence of mild distress over the symptoms in criterion A.
Moderate: Evidence of moderate distress over the symptoms in criterion A.
Severe: Evidence of severe or extreme distress over the symptoms in criterion A.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Factors such as life stresses, aging, menopause, presence of adequate sexual stimulation, must be evaluated before making this diagnosis. Relationship problems are particularly relevant to acquired interest/arousal disorder. In one study of couples with markedly decreased sexual interaction, the most prevalent etiology was marital discord.

A couple with a history of a satisfying sex life together presented with a complaint of significantly decreased sexual interaction over the past 2 years. The man was 68 and the woman was 63. He had been using Viagra with the knowledge and full acceptance of his wife, and she had been on hormone replacement therapy (HRT). The wife reported a decrease in her attraction to him following his retirement 2 years previously. They were financially secure and income and life style per se were not issues. What distressed her was his constant demand for her attention, his reliance on her to plan activities, and a lack of any time for herself. Adjustment to his retirement was difficult for both, and the wife was particularly resentful of his increased dependency on her.

Male Erectile Disorder

Male erectile disorder was historically called *impotence*. The term was dropped for a more medical designation but also because it was considered derogatory and had negative connotations for the man affected with the

problem. However, it describes with accuracy the feelings of powerlessness, helplessness, and resultant low self-esteem men with this dysfunction frequently suffer (Table 21.1–7). A man with lifelong male erectile disorder has never been able to obtain an erection sufficient for insertion. In acquired male erectile disorder, a man has successfully achieved penetration at some time in his sexual life but is later unable to do so. In situational male erectile disorder, a man is able to have coitus in certain circumstances but not in others; for example, he may function effectively with a prostitute but be unable to have an erection when with his partner.

Acquired male erectile disorder has been reported in 10 to 20 percent of all men. Freud declared it common among his patients. Erectile disorder is the chief complaint of more than 50 percent of all men treated for sexual disorders. Lifelong male erectile disorder is rare; it occurs in about 1 percent of men under age 35. The incidence of erectile disorder increases with age. It has been reported variously as 2 to 8 percent of the young adult population. Alfred Kinsey reported that 75 percent of all men were impotent at age 80. There is a reported incidence of 40 to 50 percent in men between ages 60 and 70. All men over 40, Masters and Johnson claimed, have a fear of impotence, which the researchers believed reflected the masculine fear of loss of virility with advancing age. Male erectile disorder, however, is not universal in aging men; having an available sex partner is related to continuing potency, as is a history of consistent sexual activity and the absence of vascular, neurologic, or endocrine disease. Twenty percent of men fear erectile dysfunction prior to their first coitus; the reported incidence of actual erectile dysfunction during first coitus is 8 percent. As Stephen Levine has stated, the first sexual encounter "is a horse race between excitement and anxiety."

Table 21.1–7.
DSM-5 Diagnostic Criteria for Male Erectile Disorder

A. At least one of the three following symptoms must be experienced on almost all or all (approximately 75–100%) occasions of sexual activity (in identified situational contexts or, if generalized, in all contexts):
   1. Marked difficulty in obtaining an erection during sexual activity.
   2. Marked difficulty in maintaining an erection until the completion of sexual activity.
   3. Marked decrease in erectile rigidity.
B. The symptoms in criterion A have persisted for a minimum duration of approximately 6 months.
C. The symptoms in criterion A cause clinically significant distress in the individual.
D. The sexual dysfunction is not better explained by a nonsexual mental disorder or as a consequence of severe relationship distress or other significant stressors and is not attributable to the effects of a substance/medication or another medical condition.

Specify whether:
Lifelong: The disturbance has been present since the individual became sexually active.
Acquired: The disturbance began after a period of relatively normal sexual function.
Specify whether:

Generalized: Not limited to certain types of stimulation, situations, or partners.
Situational: Only occurs with certain types of stimulation, situations, or partners.
Specify current severity:
Mild: Evidence of mild distress over the symptoms in criterion A.
Moderate: Evidence of moderate distress over the symptoms in criterion A.
Severe: Evidence of severe or extreme distress over the symptoms in criterion A.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Male erectile disorder can be organic or psychological or a combination of both, but in young and middle-aged men the cause is usually psychological. A good history is of primary importance in determining the cause of the dysfunction. If a man reports having spontaneous erections at times when he does not plan to have intercourse, having morning erections, or having good erections with masturbation or with partners other than his usual one, the organic causes of his erectile disorder can be considered negligible, and costly diagnostic procedures can be avoided. Male erectile disorder caused by a general medical condition or a pharmacological substance is discussed later in this section.

Freud ascribed one type of erectile disorder to an inability to reconcile feelings of affection toward a woman with feelings of desire for her. Men with such conflicting feelings can function only with women whom they see as degraded (Madonna–Putana complex). Other factors that have been cited as contributing to impotence include a punitive superego, an inability to trust, and feelings of inadequacy or a sense of being undesirable as a partner. A man may be unable to express a sexual impulse because of fear, anxiety, anger, or moral prohibition. In an ongoing relationship, the disorder may reflect difficulties between the partners, particularly when a man cannot communicate his needs or his anger in a direct and constructive way. In addition, episodes of erectile disorder are reinforcing, with the man becoming increasingly anxious before each sexual encounter.

Mr. B presented with erectile disorder of several months duration. He was a 51-year-old man, married for 30 years to his high school sweetheart. His wife had recently discovered that the patient had casual extramarital sexual interactions, usually on business trips. She had confronted him previously with suspicions of his activity, but he had denied his behavior. When she discovered an email chain related to his activity he was forced to admit his extramarital liaisons.

The patient had rationalized his behavior to himself by saying it did not really hurt his family whom he wanted to stay with, and that he loved his wife. Her threats of possible divorce, and his fears of losing her and his children resulted in severe performance anxiety when he tried to resume sexual relations with his wife.

Orgasmic Disorders

Female orgasmic disorder (also known as *inhibited female orgasm,* or *anorgasmia*) is defined as the recurrent and persistent inhibition of the

female orgasm, manifested by the absence or delay of orgasm after a normal sexual excitement phase, or a marked diminution of the intensity of orgasmic sensations. Women who can achieve orgasm with noncoital clitoral stimulation but cannot experience it during coitus in the absence of manual clitoral stimulation are not categorized as anorgasmic. The complaint of anorgasmia is reported by the woman, herself. However, some anorgasmic women are not distressed by the lack of climax and derive pleasure from sexual activity. In the latter instance, a woman may present with this complaint because her partner is troubled by her lack of orgasm.

Physiological research on the female sexual response has demonstrated that orgasms caused by clitoral stimulation are physiologically identical to those caused by vaginal stimulation. Freud's theory that women must give up clitoral sensitivity for vaginal sensitivity to achieve sexual maturity is now considered misleading, although some women say that they gain a special sense of satisfaction from an orgasm precipitated by coitus. Some therapists attribute that to the psychological feeling of closeness engendered by the act of coitus, but others maintain that the coital orgasm is a physiologically different experience. Many women achieve orgasm during coitus by a combination of manual clitoral stimulation and penile vaginal stimulation.

A woman with lifelong female orgasmic disorder has never experienced orgasm by any kind of stimulation. A woman with acquired orgasmic disorder has previously experienced at least one orgasm, regardless of the circumstances or means of stimulation, whether by masturbation or while dreaming during sleep. Studies have shown that women achieve orgasm more consistently with masturbation than with partnered sex. Kinsey found that 5 percent of married women over 35 years of age had never achieved orgasm by any means. The incidence of never having experienced orgasm is reported as 10 percent among all women. The incidence of orgasm increases with age. According to Kinsey, the first orgasm occurs during adolescence in about 50 percent of women as a result of masturbation or genital caressing with a partner; the rest usually experience orgasm as they get older. Lifelong female orgasmic disorder is more common among unmarried women than married women. Increased orgasmic potential in women over 35 years of age has been explained on the basis of less psychological inhibition, greater sexual experience, or both.

Acquired female orgasmic disorder is a common complaint in clinical populations. One clinical treatment facility reported having about four times as many nonorgasmic women in its practice as female patients with all other sexual disorders. In another study, 46 percent of women complained of difficulty reaching orgasm. Inhibition of arousal and orgasmic problems often occur together. The overall prevalence of female orgasmic disorder from all causes is estimated to be 30 percent. A recent twin study suggests that orgasmic dysfunction in some females has a

genetic basis and cannot be attributed solely to cultural differences. That study demonstrated an estimated heritability for difficulty reaching orgasm with intercourse of 34 percent and an estimated heritability in women who could not climax with masturbation of 45 percent.

Numerous psychological factors are associated with female orgasmic disorder. They include fears of impregnation, rejection by a sex partner, and damage to the vagina; hostility toward men; poor body image and feelings of guilt about sexual impulses. Some women equate orgasm with loss of control or with aggressive, destructive, or violent impulses; their fear of these impulses may be expressed through inhibition of arousal or orgasm.

Prolonged marital discord may lead to acquired anorgasmia. Cultural expectations and social restrictions on women are also relevant. Many women have grown up to believe that sexual pleasure is not a natural entitlement for the so-called decent women. Nonorgasmic women may be otherwise symptom free or may experience frustration in a variety of ways; they may have such pelvic complaints as lower abdominal pain, itching, and vaginal discharge, as well as increased tension, irritability, and fatigue.

A 36-year-old woman presented with extreme distress over being anorgasmic. She was unmarried and had recently ended a relationship in which she had been very attracted to the man, but was embarrassed by her lack of orgasm about which she never told him. The patient could not bring herself to orgasm by masturbation and was reluctant to fantasize.

As a child, she had witnessed the primal scene three times that she could remember. Her parents made no attempt to close their door while having sex, and she found herself listening for sounds of coitus.

Treatment focused both on masturbatory exercises with encouragement to fantasize and material to stimulate fantasy, and analytically oriented psychotherapy to deal with oedipal conflicts and inappropriate sexual stimulation during childhood.

Delayed Ejaculation

In male delayed ejaculation, sometimes called *retarded ejaculation,* a man achieves ejaculation during coitus with great difficulty, if at all. The problem is sometimes present with masturbation but appears as a problem primarily during partnered sex. A man with lifelong delayed ejaculation has never been able to ejaculate during partnered sexual activity. The problem is usually most pronounced during coital activity. The disorder is diagnosed as acquired if it develops after previously normal functioning. Some researchers think that orgasm and ejaculation should be differentiated, especially in the case of men who ejaculate but complain of a decreased or absent subjective sense of pleasure during the orgasmic experience (orgasmic anhedonia).

The incidence of male orgasmic disorder is much lower than the incidence of premature ejaculation or erectile disorder. Masters and Johnson reported an incidence of delayed ejaculation of only 3.8 percent in

one group of 447 men with sexual dysfunctions. A general prevalence of 5 percent has been reported. However, an increase in the presentation of this disorder in sex therapy programs has been seen in the past decade. This has been attributed to the increasing use of antidepressants which can have a side effect of delayed ejaculation, and a high use of Internet pornography sites. These sites offer a level of stimulation involving such variety of people and acts that they may inure the man to the stimulation of more typical partnered activity. One study of adolescent males who used these sites frequently, prior to live sexual interaction, reported that these teens do not develop neuronal synapses that will enable them to respond to usual partnered interactions with sufficient pleasure to allow them to achieve climax.

Lifelong delayed ejaculation indicates severe psychopathology. A man may come from a rigid, puritanical background; he may perceive sex as sinful and the genitals as dirty; and he may have conscious or unconscious incest wishes and guilt. He usually has difficulty with closeness in areas beyond those of sexual relations. In a few cases, the condition is aggravated by an attention-deficit disorder. A man's distractibility prevents sufficient arousal for climax to occur.

In an ongoing relationship, acquired male delayed ejaculation disorder frequently reflects interpersonal difficulties. The disorder may be a man's way of coping with real or fantasized changes in a relationship, such as plans for pregnancy about which the man is ambivalent, the loss of sexual attraction to the partner, or demands by the partner for greater commitment as expressed by sexual performance. In some men, the inability to ejaculate reflects unexpressed hostility toward a woman. The problem is more common among men with obsessive-compulsive disorder (OCD) than among others.

A 43-year-old married man presented with delayed ejaculation of 5 months duration. The onset of the dysfunction coincided with his agreeing to try to have a child with his wife. This was the patient's second marriage and he had a 9-year-old daughter from his first marriage who lived primarily with his ex-wife. He and his current wife had discussed having children before they married. Although the patient did not want another child, he agreed to father one, because his current wife, age 39, was childless and he felt it would be selfish to deny her a baby. He was afraid of losing his wife if they did not try to conceive. Therapy-involved behavioral exercises geared to restore the patient's pleasure in sexuality as well as introspective psychotherapy to address his ambivalence about conception.

A 31-year-old male presented with a history of lifelong delayed ejaculation intravaginally. He was engaged to be married and his fiancée was increasingly distressed by his dysfunction. The patient demonstrated obsessive characteristics. For example, he refused to allow his fiancée to hang up his shirts, because he was adamant that they be spaced apart to an exact degree. Additionally, he had been traumatized in his adolescence by seeing his mother miscarry a pregnancy "all over the kitchen floor." The couple was treated with behavioral sex therapy exercises.

Delayed ejaculation may have physiological causes and can occur after surgery of the genitourinary tract, for example, with a prostatectomy. Delayed ejaculation may also be associated with Parkinson disease and other neurological disorder involving the lumbar or sacral sections of the spinal cord. The antihypertensive drugs guanethidine monosulfate (Ismelin) and methyldopa (Aldomet) have been implicated in delayed ejaculation as have the phenothiazines and almost all the antidepressants. Strictly organic cases are diagnosed as medically or pharmacologically induced sexual dysfunction.

Early Ejaculation

In early ejaculation, men persistently or recurrently achieve orgasm and ejaculation before they wish to. The diagnosis is made when a man regularly ejaculates before or within approximately 1 minute after penetration. DSM-5 refers only to "vaginal penetration" in its diagnostic criteria even though it is entirely possible for the disorder to occur in men who are homosexual and do not engage in vaginal penetration. DSM-5 defines the disorder as mild if ejaculation occurs within approximately 30 seconds to 1 minute of vaginal penetration, moderate if ejaculation occurs within approximately 15 to 30 seconds of vaginal penetration, and severe when ejaculation occurs at the start of sexual activity or within approximately 15 seconds of vaginal penetration. A difficulty with these specifiers involves time distortions which patients make in both overestimating and underestimating time from penetration to climax. Clinicians need to consider factors that affect the duration of the excitement phase of the sexual response, such as age, the novelty of the sex partner, and the frequency and duration of coitus. As with the other sexual dysfunctions, premature ejaculation is not diagnosed when it is caused exclusively by organic factors or when it is symptomatic of another clinical psychiatric syndrome.

Early ejaculation is more commonly reported among college-educated men than among men with less education. The complaint is thought to be related to their concern for partner satisfaction, but the true cause of this increased frequency has not been determined. Early ejaculation is the chief complaint of about 35 to 40 percent of men treated for sexual disorders. In DSM-5, the writers state that the disorder, with its newly defined time parameter, would now be an accurate diagnosis for only 1 to 3 percent of men. Some researchers divide men who experience early ejaculation into two groups: those who are physiologically predisposed to climax quickly because of shorter nerve latency time and those with a psychogenic or behaviorally conditioned cause. Difficulty in ejaculatory control can be associated with anxiety regarding the sex act, with unconscious fears about the vagina, or with negative cultural conditioning. Men whose early sexual contacts occurred largely with prostitutes who demanded that the sex act

proceed quickly or whose sexual contacts took place in situations in which discovery would be embarrassing (e.g., in a shared dormitory room or in the parental home) might have been conditioned to achieve orgasm rapidly. With young, inexperienced men, who have the problem, it may resolve in time. In ongoing relationships, the partner has a great influence on a premature ejaculator, and a stressful marriage exacerbates the disorder. The developmental background and the psychodynamics found in premature ejaculation and in erectile disorder are similar.

PAIN DISORDER

Genito-Pelvic Pain/Penetration Disorder

In DSM-5, this disorder refers to one or more of the following complaints, of which any two or more may occur together: difficulty having intercourse; genito-pelvic pain; fear of pain or penetration; and tension of the pelvic floor muscles. Previously, these pain disorders were diagnosed as *dyspareunia* or *vaginismus*. These former diagnoses could coexist with each other or lead to one another and could understandably lead to fear of pain with sex. Thus, it is reasonable to gather them into one diagnostic category. For purposes of clinical discussion, however, the distinct categories of dyspareunia and vaginismus remain useful.

Dyspareunia. Dyspareunia is recurrent or persistent genital pain occurring before, during, or after intercourse. Dyspareunia is related to, and often coincides with, vaginismus. Repeated episodes of vaginismus can lead to dyspareunia and vice versa; in either case, somatic causes must be ruled out. DSM-5 cites that 15 percent of women in North America report recurrent pain during intercourse.

In most cases, dynamic factors are considered causative. Chronic pelvic pain is a common complaint in women with a history of rape or childhood sexual abuse. Painful coitus can result from tension and anxiety about the sex act that cause women to involuntarily contract their pelvic floor muscles. The pain is real and makes intercourse unpleasant or unbearable. Anticipation of further pain may cause women to avoid coitus altogether. If a partner proceeds with intercourse regardless of a woman's state of readiness, the condition is aggravated. There is an increase in reported dyspareunia postmenopausally due to hormonally induced physiologic changes in the vagina; however specific complaints of difficulty having intercourse occur more often in premenopausal women. There is some increase in dyspareunia in the immediate postpartum population, but, it is usually temporary. Dyspareunia may present as any of the four complaints listed under genito-pelvic pain/penetration disorder and should be diagnosed as genito-pelvic pain/penetration disorder.

A 20-year-old woman presented with pelvic-pain/penetration disorder 6 months after her marriage. She and her husband had experienced manual sex play to which she responded with some arousal, but the marriage remained unconsummated. The woman had been brought up in an orthodox Jewish community, where virginity was an imperative and social/sexual interaction between adolescent and young adult men and women was limited. The patient had minimal experience dating, and her husband was similarly inexperienced sexually. Therapy consisted of sex education and behavioral sexual exercises.

Dyspareunia can occur in men, but it is uncommon and usually associated with a medical condition such as Peyronie disease, prostatitis, or gonorrheal or herpetic infections. Vasocongestion during sexual activity without orgasmic release also may lead to discomfort. Rarely, some men experience pain on ejaculation (postejaculatory pain disorder). This pain is caused by an involuntary spasm of the perineal muscles that may be due to psychological conflicts about the sex act or may be an adverse effect of some antidepressant medications. Sexual dysfunction caused by a medication is the diagnosis used when a substance or medication is the sole or major causal factor for the pain.

Sexual Dysfunction Due to a General Medical Condition

The category sexual dysfunction due to a general medical condition covers sexual dysfunction that results in marked distress and interpersonal difficulty when there is evidence from the history, physical examination, or laboratory findings of a general medical condition judged to be causally related to the sexual dysfunction.

Male Erectile Disorder Due to a General Medical Condition. Many studies have focused on the relative incidences of psychological and organic male erectile disorder. Statistics indicate that 20 to 50 percent of men with erectile disorder have an organic basis for the disorder. A physiologic etiology is more likely in men older than 50 and the most likely cause in men older than age 60. Physiologically, erectile dysfunction may be due to a variety of medical causes (Table 21.1–8). In the United States, it is estimated that two million men cannot gain erections because they suffer from diabetes mellitus; an additional 300,000 are dysfunctional because of other endocrine diseases; approximately 1.5 million are dysfunctional as a result of vascular disease; an estimated 180,000 because of multiple sclerosis; approximately 400,000 because of traumas and fractures leading to pelvic fractures or spinal cord injuries; and another 650,000 are estimated to be dysfunctional as a result of radical surgery, including prostatectomies, colostomies, and cystectomies. In addition, the clinician should be aware of the possible pharmacological effects of medication on sexual functioning. The increased incidence of organic causes for erectile dysfunction may partly reflect the increased use of psychotropic and antihypertensive medications. Adverse effects of medication may impair

male sexual functioning in a variety of ways. Castration (removal of the testes) does not always lead to sexual dysfunction, depending on the person. Erection may still occur after castration via a reflex arc that passes through the sacral cord erectile center. It is triggered when the inner thigh is stimulated.

Table 21.1–8.
Diseases and Other Medical Conditions Implicated in Erectile Dysfunction

Infectious and parasitic diseases	Neurological disorders
Elephantiasis	Multiple sclerosis
Mumps	Transverse myelitis
Cardiovascular disease	Parkinson disease
Atherosclerotic disease	Temporal lobe epilepsy
Aortic aneurysm	Traumatic and neoplastic spinal cord diseases
Leriche syndrome	Central nervous system tumor
Cardiac failure	Amyotrophic lateral sclerosis
Renal and urological disorders	Peripheral neuropathy
Peyronie disease	General paresis
Chronic renal failure	Tabes dorsalis
Hydrocele and varicocele	Pharmacological contributants
Hepatic disorders	Alcohol, tobacco and other dependence-inducing substances (heroin, methadone, morphine, cocaine, amphetamines, and barbiturates)
Cirrhosis (usually associated with alcohol dependence)	Prescribed drugs (psychotropic drugs, antihypertensive drugs, estrogens, and antiandrogens)
Pulmonary disorders	Poisoning
Respiratory failure	Lead (plumbism)
Genetic disorders	Herbicides
Klinefelter syndrome	Surgical procedures
Congenital penile vascular and structural abnormalities	Perineal prostatectomy
Nutritional disorders	Abdominal–perineal colon resection
Malnutrition	Sympathectomy (frequently interferes with ejaculation)
Vitamin deficiencies	Aortoiliac surgery
Endocrine disorders	Radical cystectomy
Diabetes mellitus	Retroperitoneal lymphadenectomy
Dysfunction of the pituitary–adrenal–testis axis	Miscellaneous
Acromegaly	Radiation therapy
Addison disease	Pelvic fracture
Chromophobe adenoma	Any severe systemic disease or debilitating condition
Adrenal neoplasia	
Myxedema	
Hyperthyroidism	

The mechanisms involved in the various diseases that lead to erectile disorder include general endothelial dysfunction or dysfunction of the penile endothelium specifically, atherosclerosis, dysregulation of nitric

oxide synthase, decrease of nitric oxide, low testosterone levels due to hyperprolactinemia-influenced changes in the hypothalamic–pituitary axis, vasculopathy, autonomic neuropathy, and disruption of neural proerectile processes.

PHYSIOLOGICAL TESTS. A number of procedures, benign and invasive, are used to help differentiate psychogenic erectile dysfunction. The procedures include monitoring nocturnal penile tumescence (erection that occurs during sleep) normally associated with rapid eye movement; monitoring tumescence with a strain gauge; measuring blood pressure in the penis with a penile plethysmograph or an ultrasound (Doppler) flow meter, both of which assess blood flow in the internal pudendal artery; and measuring pudendal nerve latency time. Neurological impairment of penile function may be indicated by decreased vibratory perception in the penis. Other diagnostic tests that delineate organic bases for erectile disorder include glucose tolerance tests, plasma hormone assays, liver and thyroid function tests, prolactin and follicle-stimulating hormone (FSH) determinations, and cystometric examinations. On physical examination, evidence of hypogonadism such as small testes, gynecomastia, reduced or absent body and facial hair warrant attention. Invasive diagnostic studies include penile arteriography, infusion cavernosonography, and radioactive xenon penography. Invasive procedures require expert interpretation and are used only for patients who are candidates for vascular reconstructive procedures.

MEDICAL VERSUS PSYCHOGENIC CAUSES. A good history is crucial to determining the cause of the male erectile disorder. If a man reports having spontaneous erections at times when he does not plan to have intercourse, having morning erections or only sporadic erectile dysfunction, or having good erections with masturbation or with partners other than his usual one, then organic causes for his disorder can be considered negligible, and costly diagnostic procedures can be avoided. When a medical basis for erectile dysfunction is found, psychological factors often contribute to the dysfunction, and psychiatric treatment may be helpful. Some diabetics, for instance, may experience psychogenic erectile dysfunction.

Dyspareunia Due to a General Medical Condition. An estimated 30 percent of all surgical procedures on the female genital area result in temporary dyspareunia. In addition, 30 to 40 percent of women with the complaint who are seen in sex therapy clinics have pelvic pathology. Organic abnormalities leading to dyspareunia and vaginismus include infected hymenal remnants or congenitally imperforate or unusually thick hymens; episiotomy scars; Bartholin glands infections; various forms of vaginitis and cervicitis; testosterone deficiency; vulvodynia; vaginal treatment, such as radiation or surgery that has caused scarring;

dermatological disorder, such as lichens sclerosis or lichens planus; and endometriosis. Postcoital pain reported by women with myomata and endometriosis is attributed to the uterine contractions during orgasm. Postmenopausal women may have dyspareunia because of thinning of the vaginal mucosa and reduced lubrication. Women who have been on birth control pills for many years may have lessened lubrication and mild vaginal thinning similar to the postmenopausal woman.

Dyspareunia can also occur in men, but it is uncommon and is usually associated with an organic condition such as Peyronie disease, which consists of sclerotic plaques on the penis that cause penile curvature.

Hypoactive Sexual Desire and Arousal/Interest Disorder Due to a General Medical Condition. Desire commonly decreases after major illness or surgery, particularly when body image is affected after such procedures as mastectomy, ileostomy, hysterectomy, and prostatectomy. Illnesses that deplete a person's energy, chronic conditions that require physical and psychological adaptation, and serious illnesses that may cause the person to become depressed can all result in a marked lessening of sexual desire or interest in both men and women. In some cases, biochemical correlates are associated with hypoactive sexual desire disorder (Table 21.1–9). A recent study found markedly lower serum testosterone concentrations in men complaining of low desire than in normal controls in a sleep laboratory situation. Drugs that depress the central nervous system (CNS) or decrease testosterone production can decrease desire.

Other Male Sexual Dysfunction Due to a General Medical Condition

Delayed ejaculation can have physiological causes and can occur after surgery on the genitourinary tract, such as prostatectomy. It may also be associated with Parkinson disease and other neurological disorders involving the lumbar or sacral sections of the spinal cord. The antihypertensive drug guanethidine monosulfate (Ismelin), methyldopa (Aldomet), the phenothiazines, the tricyclic drugs, and the selective serotonin reuptake inhibitors (SSRIs), among others, have been implicated in retarded ejaculation. Delayed ejaculation must also be differentiated from retrograde ejaculation, in which ejaculation occurs but the seminal fluid passes backward into the bladder. Retrograde ejaculation always has an organic cause. It can develop after genitourinary surgery and is also associated with medications that have anticholinergic adverse effects, such as some phenothiazines.

Other Female Sexual Dysfunction Due to a General Medical Condition

Some medical conditions—specifically, endocrine diseases such as hypothyroidism, diabetes mellitus, and primary hyperprolactinemia—can

affect a woman's ability to have orgasms. Several drugs also affect some women's capacity to have orgasms. Antihypertensive medications, tricyclic drugs, SSRIs, and, frequently, monoamine oxidase inhibitors (MAOIs) have interfered with female orgasmic capacity. One study of women taking MAOIs, however, found that after 16 to 18 weeks of pharmacotherapy, the adverse effect of the medication disappeared and the women were able to reexperience orgasms, although they continued taking an undiminished dosage of the drug.

Table 21.1–9.
Neurotransmitter Effects on Sex Function

	Dopamine	Serotonin	Adrenergic	Cholinergic	Clinical Correlation
Erection	+ +	+/–	α, β – +	+/–	Antipsychotics (dopamine receptor antagonists) may lead to erectile dysfunction; dopaminergic drug agonists may lead to enhanced erection and libido; priapism with trazodone (α_1 block); β-adrenergic receptor antagonists may lead to impotence.
Ejaculation and orgasm	+/–	+	+ α_1	+/–	α_1-Adrenergic receptor antagonists, tricyclic drugs, MAOIs, thioridazine may lead to impaired ejaculation; serotonergic agents may inhibit orgasm.

+/–, minimal or no effect; +, facilitates effect; –, inhibiting effect; MAOI, monoamine oxidase inhibitor.

SUBSTANCE/MEDICATION-INDUCED SEXUAL DYSFUNCTION

The diagnosis of substance-induced sexual dysfunction is used when evidence of substance intoxication or withdrawal is apparent from the history, physical examination, or laboratory findings. The disturbance in sexual function must be predominant in the clinical picture. Distressing sexual dysfunction occurs soon after significant substance intoxication or withdrawal or after exposure to a medication or a change in medication use. Specified substances include alcohol, amphetamines or related substances, cocaine, opioids, sedatives, hypnotics, or anxiolytics, and other or unknown substances.

Abused recreational substances affect sexual function in various ways. In small doses, many substances enhance sexual performance by decreasing inhibition or anxiety or by causing a temporary elation of mood. With continued use, however, erectile engorgement and orgasmic and ejaculatory capacities become impaired. The abuse of sedatives, anxiolytics, hypnotics, and particularly opiates and opioids nearly always depresses desire. Alcohol may foster the initiation of sexual activity by removing inhibition, but it also impairs performance. Cocaine and amphetamines

produce the following similar effects: although no direct evidence indicates that sexual drive is enhanced, users initially have feelings of increased energy and may become sexually active; ultimately, dysfunction occurs. Men usually go through two stages: an experience of prolonged erection without ejaculation, then a gradual loss of erectile capability.

Patients recovering from substance dependency may need therapy to regain sexual function, partly because of psychological readjustment to a nondependent state. Many substance abusers have always had difficulty with intimate interactions. Others who spent their crucial developmental years under the influence of a substance have missed the experiences that would have enabled them to learn social and sexual skills.

Pharmacological Agents Implicated in Sex Dysfunction

Almost every pharmacological agent, particularly those used in psychiatry, has been associated with an effect on sexuality. In men, these effects include decreased sex drive, erectile failure, decreased volume of ejaculate, and delayed or retrograde ejaculation. In women, decreased sex drive, decreased vaginal lubrication, inhibited or delayed orgasm, and decreased or absent vaginal contractions may occur. Drugs may also enhance the sexual responses and increase the sex drive, but this is less common than adverse effects. The effects of psychoactive drugs are detailed below.

Antipsychotic Drugs. Most antipsychotic drugs are dopamine receptor antagonists that also block adrenergic and cholinergic receptors, thus accounting for adverse sexual effects (Table 21.1–9). Chlorpromazine (Thorazine), and trifluoperazine (Stelazine) are potent anticholinergics and they impair erection and ejaculation. With some drugs the seminal fluid backs up into the bladder rather than being propelled through the penile urethra. Patients still have a pleasurable sensation, but the orgasm is dry. When urinating after orgasm, the urine may be milky white because it contains the ejaculate. The condition is startling but harmless. Paradoxically, some rare cases of priapism have been reported with antipsychotics.

Antidepressant Drugs. The tricyclic and tetracyclic antidepressants have anticholinergic effects that interfere with erection and delay ejaculation. Because the anticholinergic effects vary among the cyclic antidepressants, those with the fewest effects (e.g., desipramine [Norpramin]) produce the fewest sexual adverse effects. The effects of the tricyclics and tetracyclics have not been documented sufficiently in women; however, few women seem to complain of any effects.

Some men report increased sensitivity of the glans that is pleasurable and that does not interfere with erection, although it delays ejaculation. In some cases, however, the tricyclic causes painful ejaculation, perhaps as the

result of interference with seminal propulsion caused by interference with, in turn, urethral, prostatic, vas, and epididymal smooth muscle contractions. Clomipramine (Anafranil) has been reported to increase sex drive in some persons. Selegiline (Deprenyl), a selective MAO type B (MAO_B) inhibitor, and bupropion (Wellbutrin) have also been reported to increase sex drive, possibly by dopaminergic activity and increased production of norepinephrine.

Venlafaxine (Effexor) and the SSRIs most often have adverse effects because of the rise in serotonin levels. A lowering of the sex drive and difficulty reaching orgasm occur in both sexes. Reversal of those negative effects has been achieved with cyproheptadine (Periactin), an antihistamine with antiserotonergic effects, and with methylphenidate (Ritalin), which has adrenergic effects. Trazodone (Desyrel) is associated with the rare occurrence of priapism, the symptom of prolonged erection in the absence of sexual stimuli. That symptom appears to result from the α_2-adrenergic antagonism of trazodone.

The MAOIs affect biogenic amines broadly. Accordingly, they produce impaired erection, delayed or retrograde ejaculation, vaginal dryness, and inhibited orgasm. Tranylcypromine (Parnate) has a paradoxical sexually stimulating effect in some persons, possibly as a result of its amphetamine-like properties.

General Effects. Because depression is associated with a decreased libido, varying levels of sexual dysfunction and anhedonia are part of the disease process. Some patients report improved sexual functioning as their depression improves as a result of antidepressant medication. The phenomenon makes the evaluation of sexual side effects difficult; also, the side effects may disappear with time, perhaps because a biogenic amine homeostatic mechanism comes into play.

LITHIUM. Lithium (Eskalith) regulates mood and, in the manic state, may reduce hypersexuality, possibly by a dopamine antagonist activity. In some patients, impaired erection has been reported.

SYMPATHOMIMETICS. Psychostimulants, which are sometimes used in the treatment of depression, include amphetamines, methylphenidate, and pemoline (Cylert), which raise the plasma levels of norepinephrine and dopamine. Libido is increased; however, with prolonged use, men may experience a loss of desire and erections.

α -adrenergic and β-adrenergic receptor antagonists. α-Adrenergic and β-adrenergic receptor antagonists are used in the treatment of hypertension, angina, and certain cardiac arrhythmias. They diminish tonic sympathetic nerve outflow from vasomotor centers in the brain. As a result, they can cause impotence, decrease the volume of ejaculate, and produce retrograde

ejaculation. Changes in libido have been reported in both sexes.

Suggestions have been made to use the side effects of drugs therapeutically. Thus, a drug that delays or interferes with ejaculation (e.g., fluoxetine [Prozac]) might be used to treat premature ejaculation.

ANTICHOLINERGICS. The anticholinergics block cholinergic receptors and include such drugs as amantadine (Symmetrel) and benztropine (Cogentin). They produce dryness of the mucous membranes (including those of the vagina) and erectile disorder. However, amantadine may reverse SSRI-induced orgasmic dysfunction through its dopaminergic effect.

ANTIHISTAMINES. Drugs such as diphenhydramine (Benadryl) have anticholinergic activity and are mildly hypnotic. They may inhibit sexual function as a result. Cyproheptadine, although an antihistamine, also has potent activity as a serotonin antagonist. It is used to block the serotonergic sexual adverse effects produced by SSRIs, such as delayed orgasm.

ANTIANXIETY AGENTS. The major class of anxiolytics is the benzodiazepines (e.g., diazepam [Valium]). They act on the γ-aminobutyric acid (GABA) receptors, which are believed to be involved in cognition, memory, and motor control. Because they decrease plasma epinephrine concentrations, they diminish anxiety, and as a result they improve sexual function in persons inhibited by anxiety.

ALCOHOL. Alcohol suppresses CNS activity generally and can produce erectile disorders in men as a result. Alcohol has a direct gonadal effect that decreases testosterone levels in men; paradoxically, it can produce a slight rise in testosterone levels in women. The latter finding may account for women reporting increased libido after drinking small amounts of alcohol. The long-term use of alcohol reduces the ability of the liver to metabolize estrogenic compounds. In men, that produces signs of feminization (such as gynecomastia).

OPIOIDS. Opioids, such as heroin, have adverse sexual effects, such as erectile failure and decreased libido. The alteration of consciousness may enhance the sexual experience in occasional users.

HALLUCINOGENS. The hallucinogens include lysergic acid diethylamide (LSD), phencyclidine (PCP), psilocybin (from some mushrooms), and mescaline (from peyote cactus). In addition to inducing hallucinations, the drugs cause loss of contact with reality and expanding a sense of consciousness. Some users report that the sexual experience is similarly enhanced, but others experience anxiety, delirium, or psychosis, which clearly interferes with sexual function.

CANNABIS. The altered state of consciousness produced by cannabis may enhance sexual pleasure for some persons. Its prolonged use depresses testosterone levels.

BARBITURATES AND SIMILARLY ACTING DRUGS. Barbiturates and similarly acting sedative-hypnotic drugs may enhance sexual responsiveness in persons who are sexually unresponsive as a result of anxiety. They have no direct effect on the sex organs; however, they do produce an alteration in consciousness that some persons find pleasurable. They are subject to abuse and can be fatal when combined with alcohol or other CNS depressants.

Methaqualone (Quaalude) acquired a reputation as a sexual enhancer, which had no biological basis in fact. It is no longer marketed in the United States.

Other Specified Sexual Dysfunction

Diagnoses in this category include dysfunctions that distress the patient but do not meet the criteria for the disorders discussed above. An example would be sexual aversion, in which the patients cannot enjoy the sex act because of a phobic aversion to touching the genitalia.

Unspecified Sexual Dysfunction

This category covers other sexual dysfunctions, that do not fall under the diagnoses discussed above, but that cause distress and interference with sexual connection and pleasure. In these cases, the clinician does not choose to elaborate the differences from the diagnoses discussed above.

Further examples of other specified or unpecified disorders include individuals who experience the physiological components of sexual excitement and orgasm but report no erotic sensation or even anesthesia and the male experience of orgasm with a flaccid penis. The orgasmic woman who desires but has not experienced multiple orgasms can be classified under this heading as well. Disorders of excessive rather than inhibited function, such as compulsive masturbation, might be diagnosed under unspecified dysfunction. These diagnoses also might be used to cover complaints engendered by couple, rather than individual, dysfunction—for example, a couple in which one partner prefers morning sex and the other functions more readily at night or a couple with unequal frequencies of desire.

Compulsive Sexual Behavior. The concept of compulsive sexual behavior (Table 21.1–10), or sex addiction, was developed in the 1980s to describe persons who compulsively seek out sexual experiences and whose behavior becomes impaired if they cannot gratify their sexual impulses. The concept of sex addiction is derived from the model of addiction to

drugs such as heroin or addiction to behavioral patterns such as gambling. Addiction implies psychological dependence, physical dependence, and a withdrawal symptom if the substance (e.g., the drug) is unavailable or the behavior (e.g., gambling) is frustrated.

DSM-5 does not have a diagnosis of *sex addiction* or *compulsive sexual behavior,* nor is this disorder universally recognized or accepted. Nevertheless, the person whose entire life revolves around sex-seeking behavior and activities, who spends an excessive amount of time in such behavior, and who often tries to stop such behavior but cannot do so is well known to clinicians. Such persons show repeated and increasingly frequent attempts to have a sexual experience, and deprivation evokes symptoms of distress. In the author's view, sex addiction is a useful concept heuristically because it can alert the clinician to seek an underlying cause for the manifest behavior.

Table 21.1–10.
Signs of Sexual Addiction or Compulsive Sexual Behavior

1. Out-of-control behavior
2. Severe adverse consequences (medical, legal, interpersonal) due to sexual behavior
3. Persistent pursuit of self-destructive or high-risk sexual behavior
4. Repeated attempts to limit or stop sexual behavior
5. Sexual obsession and fantasy as a primary coping mechanism
6. Need for increasing amounts of sexual activity
7. Severe mood changes related to sexual activity (e.g., depression, euphoria)
8. Inordinate amount of time spent in obtaining sex, being sexual, or recovering from sexual experience
9. Interference of sexual behavior in social, occupational, or recreational activities

Data from Carnes P. *Don't Call It Love.* New York: Bantam Books; 1991.

Persistent Genital Arousal Disorder. Persistent genital arousal disorder has previously been called persistent sexual arousal syndrome.

It has been diagnosed in women who complain of a continual feeling of sexual arousal which is uncomfortable, demands release and interferes with life pleasures and activities. These women masturbate frequently, sometimes incessantly, because climax provides relief. However, the relief is temporary and the sense of arousal returns rapidly and remains. The sense of arousal in these cases is neither pleasurable nor exciting, and the women are not interested in a sexual experience, but in relief from their symptom. Some women have reported masturbating so frequently to alleviate the arousal that they have irritated their genitalia severely. One case of attempted suicide has been reported with this syndrome, with the woman stating that she could no longer tolerate the sensations and that she had masturbated so often that her vulva was raw.

Krafft-Ebing reported a case of a young woman who masturbated

almost incessantly and was unable to control her impulses. She had frequent coitus with many men, but neither coitus nor masturbation sufficed, and she was eventually placed in an institution. Krafft-Ebing referred to the condition as "sexual hyperaesthesia" which he believed could occur in otherwise normal persons.

This case would have to be differentiated from sexual addiction (discussed above). The differentiating factor would be whether the woman desired an orgasm for itself, and looked forward to having sex, or whether she was seeking relief from incessant and intolerable stimulation. There is some speculation that this disorder is due to nerve damage or anomaly, but the etiology is unknown.

Postcoital Dysphoria. Postcoital dysphoria is not listed in DSM-5. It occurs during the resolution phase of sexual activity, when individuals normally experience a sense of general well-being and muscular and psychological relaxation. Some individuals, however, experience postcoital dysphoria—after an otherwise satisfactory sexual experience, they become depressed, tense, anxious, and irritable and show psychomotor agitation. They often want to get away from the partner and may become verbally or even physically abusive. The incidence of the disorder is unknown, but it is more common in men than in women. Its several causes relate to the person's attitude toward sex in general and the partner in particular. It may occur in adulterous sex and with prostitutes, when there is a profound fear of intimacy, or when individuals cannot experience sex without consequent strong feelings of guilt. The fear of sexually transmitted disease causes some persons to experience postcoital dysphoria. Treatment requires insight-oriented psychotherapy to help patients understand the unconscious antecedents to their behavior and attitudes.

Unconsummated Marriage. A couple involved in an unconsummated marriage have never had coitus and are typically uninformed and inhibited about sexuality. Their feelings of guilt, shame, or inadequacy are increased by their problems, and they experience conflict between their need to seek help and their need to conceal their difficulty. Couples present with the problem after having been married several months or several years. Masters and Johnson reported an unconsummated marriage of 17 years' duration.

Frequently, the couple does not seek help directly, but the woman may reveal the problem to her gynecologist on a visit ostensibly concerned with vague vaginal or other somatic complaints. On examining her, the gynecologist may find an intact hymen. In some cases, however, the wife may have undergone a hymenectomy to resolve the problem. Inquiry by a physician who is comfortable in dealing with sexual problems may be the first opening to a frank discussion of the couple's distress. Often, the

pretext of the medical visit is a discussion of contraceptive methods or even a request for an infertility workup. Once presented, the complaint often can be successfully treated. The duration of the problem does not significantly affect the prognosis or the outcome of the case.

The causes of unconsummated marriage are varied: lack of sex education, sexual prohibitions overly stressed by parents or society, problems of an oedipal nature, immaturity in both partners, overdependence on primary families, and problems in sexual identification. Religious orthodoxy, with severe control of sexual and social development or the equation of sexuality with sin or uncleanliness, has also been cited as a dominant cause. Many women involved in an unconsummated marriage have distorted concepts about their vaginas. They may fear that the vagina is too small or too soft, or they may confuse the vagina with the rectum, leading to feelings of being unclean. The man may share in those distortions about the vagina and, in addition, perceive it as dangerous to him. Similarly, both partners may have distortions about the man's penis, perceiving it as a weapon, as too large, or as too small. Many patients can be helped by simple education about genital anatomy and physiology, by suggestions for self-exploration, and by correct information from a physician. The problem of the unconsummated marriage is best treated by seeing both members of the couple. Dual-sex therapy has been markedly effective. However, other forms of conjoint therapy, marital counseling, traditional psychotherapy on a one-to-one basis, and counseling from a sensitive family physician, gynecologist, or urologist are all helpful.

Body Image Problems. Some individuals are ashamed of their bodies and experience feelings of inadequacy related to self-imposed standards of masculinity or femininity. They may insist on sex only during total darkness, not allow certain body parts to be seen or touched, or seek unnecessary operative procedures to deal with their perceived inadequacies. Certain people have difficulty adjusting to narcissistic injuries inflicted by age or illness and withdraw from sex as a result. Body dysmorphic disorder should be ruled out.

Don Juanism. Some men who appear to be hypersexual, as shown by their need to have many sexual encounters or conquests, use their sexual activities to mask deep feelings of inferiority. Some have unconscious homosexual impulses, which they deny by compulsive sexual contacts with women. After having sex, most Don Juans are no longer interested in the woman. The condition is also referred to as *satyriasis* or as a form of sex addiction.

Nymphomania. Nymphomania signifies excessive or pathological desire for coitus in a woman. There have been few scientific studies of the

condition. Those patients who have been studied usually have had one or more sexual disorders, usually including female orgasmic disorder. The woman often has an intense fear of loss of love. She attempts to satisfy her dependency needs rather than to gratify her sexual impulses through her actions. It is sometimes classified as a form of sex addiction.

Fantasies. Other atypical disorders are found in individuals who have one or more sexual fantasies about which they obsess, feel guilty, or are otherwise dysphoric. As indicated in Table 21.1–11, however, the range of common sexual fantasies is broad.

Postcoital Headache. Postcoital headache is a headache immediately after coitus and may last for several hours. It is usually described as throbbing and is localized in the occipital or frontal area. The cause is unknown but may be vascular, due to muscle contraction (tension), or psychogenic. Coitus may precipitate migraine or cluster headaches in predisposed persons.

Table 21.1–11.
Common Sexual Fantasies[a]

Men	Women
Heterosexual	
Replacement of established partner	Replacement of established partner
	Forced sexual encounters with men
Forced sexual encounters with women	Observing sexual activity
	Idyllic encounters with unknown men
Observing sexual activity	
Sexual encounters with men	Sexual encounters with women
Group sex	
Homosexual	
Images of male anatomy	Forced sexual encounters with women
Forced sexual encounters with men	Idyllic encounters with established partner
Sexual encounters with women	
Idyllic encounters with unknown men	Sexual encounters with men
	Memories of past sexual experiences
Group sex	Sadistic imagery

Listed in order of occurrence. A 1994 study found that one in five persons experienced a same-sex sexual fantasy at some time in their lives. Adapted from Masters W, Schwartz M. The Masters and Johnson treatment program for dissatisfied homosexual men. *Am J Psychiatry.* 1984;141:173.

Orgasmic Anhedonia. Orgasmic anhedonia is a condition in which the person had no physical sensation of orgasm, even though the physiological component (e.g., ejaculation) remains intact. Medical causes, such as sacral and cephalic lesions that interfere with afferent pathways from the genitalia to the cortex, must be ruled out. Psychological causes usually relate to

extreme guilt about experiencing sexual pleasure—these feelings produce a type of dissociative response that isolates the affective component of the orgasmic experience from consciousness.

Female Premature Orgasm. Data on female premature orgasm are lacking; no separate category for premature orgasm in women is included in DSM-5. However, in one study, 10 percent of women felt they reached orgasm too quickly.

Masturbatory Pain. Some individuals may experience pain during masturbation. Organic causes should always be ruled out. A small vaginal tear or early Peyronie disease may produce a painful sensation. The condition should be differentiated from compulsive masturbation. People may masturbate to the extent that they do physical damage to their genitals and eventually experience pain during subsequent masturbatory acts.

Certain masturbatory practices have resulted in what has been called *autoerotic asphyxiation,* which is usually classified as a paraphilia (hypoxyphilia). The practices may involve masturbating while hanging by the neck to heighten erotic sensations and the intensity of the orgasm through the mechanism of mild hypoxia. Although they intend to release themselves from the noose after orgasm, an estimated 500 to 1,000 persons a year accidentally kill themselves by hanging. Most who indulge in the practice are men; transvestism is often associated with the habit, and most deaths occur among adolescents. Such masochistic practices are usually associated with severe mental disorders, such as schizophrenia and major mood disorders.

TREATMENT

The treatment of sexual disorders has evolved significantly since the 1970s, when Masters and Johnson focused the attention of the psychiatric community on sexual disorders. Innovations in treatment reflect the results of research and changes in the patient population. For example, in the late 1960s, most cases of erectile dysfunction were considered psychological in origin, with approximately 20 percent of erectile problems having an organic cause. Currently, the numbers are reversed, with the majority of cases considered to result from a physiological problem. Actually, many cases are of mixed origin. Similarly, early patients in sex therapy were assumed to suffer, in part, from lack of sexual information and culturally reinforced negative attitudes toward sex. In the past 50 years, the public has received a great deal of accurate information about sex through the media. Patients now have more knowledge and greater sexual sophistication. Their dysfunctions frequently have a complex etiology involving psychodynamic and relational issues. Problems of desire are seen with increasing frequency and are among the most challenging cases for

therapists. Couple issues involve problems of trust, intimacy, lack of sexual attraction, and struggles for dominance. Finally, along with the rest of psychiatry, sex therapy has experienced medicalization. Biological treatment approaches have developed rapidly. An eclectic approach that allows the use of several techniques sequentially or in combination may be necessary.

Before entering therapy, the patient may need a thorough medical evaluation, including a medical history, physical examination, and appropriate laboratory studies when necessary. If a medical cause for the disorder is found, treatment should be directed toward ameliorating that cause.

Before 1970, the most common treatment of sexual dysfunction was individual psychotherapy. Classic psychodynamic theory considers sexual inadequacy to have its roots in early developmental conflicts, and the sexual disorder is treated as part of a more pervasive emotional disturbance. Treatment focuses on the exploration of unconscious conflicts, motivation, fantasy, and various interpersonal difficulties. Therapy assumes that removal of the conflicts will allow the sexual impulse to become structurally acceptable to the patient's ego and thereby find appropriate means of satisfaction in the environment. Unfortunately, the symptom of sexual dysfunction frequently becomes secondarily autonomous and persists after resolution of the other problems evolving from the patient's pathology. The addition of behavioral techniques is often necessary to cure the sexual problem.

Dual-Sex Therapy

The theoretical basis of the dual-sex therapy approach is the concept of the marital unit or dyad as the object of therapy. The method of dual-sex therapy was originated and developed by Masters and Johnson. Dual-sex therapy does not accept the idea of a dysfunctional half of a patient couple. Both individuals are involved in a relationship in which there is sexual distress, and, thus, both should participate in the therapy program.

The sexual problem often reflects other areas of disharmony or misunderstanding in the marriage. The marital relationship as a whole is treated, with emphasis on sexual functioning as a part of that relationship. Improved communication in sexual and nonsexual areas is a specific goal of treatment. Psychological and physiological aspects of sexual functioning are discussed with an educational attitude. Suggestions are followed in the privacy of the couple's home.

Initial histories are taken to determine suitability for this type of treatment. Evidence of major underlying psychopathology suggests further psychiatric evaluation, and participation in the program may be deferred until the patient seems better able to benefit from it. Concurrent psychotherapy with a psychiatrist while participating in dual-sex therapy is

sometimes recommended.

Each patient is interviewed individually early in the course of treatment. A complete sexual history is obtained, which is later reflected back to the couple to help them understand their present problem. For example, a history of culturally or parentally transmitted inhibition can help relieve the partner of an anorgasmic woman of the idea that he is a bad lover. Similarly, the partner of a man with erectile disorder may be relieved of the concern that his problem is a result of her being insufficiently attractive. The individual sessions also help the therapist understand the patients' lifestyle and allow for suggestions that fit into that lifestyle.

Behavioral Exercises. Treatment is short term and behaviorally oriented. Specific exercises are prescribed to help the couple with their particular problem. Sexual dysfunction often involves a fear of inadequate performance; thus, couples are specifically prohibited from any sexual play other than that prescribed by the therapist. Initially, intercourse is interdicted, and couples learn to give and receive bodily pleasure without the pressure of performance. Beginning exercises usually focus on heightening sensory awareness to touch, sight, sound, and smell.

During these exercises, called *sensate focus exercises,* the couple is given much reinforcement to lessen anxiety. They are urged to use fantasies to distract them from obsessive concerns about performance, which is termed *spectatoring*. The needs of both the dysfunctional partner and the nondysfunctional partner are considered. If either partner becomes sexually excited by the exercises, the other is encouraged to bring him or her to orgasm by manual or oral means. This procedure is important to keep the nondysfunctional partner from sabotaging the treatment. Open communication between the partners is urged, and the expression of mutual needs is encouraged. Resistances, such as claims of fatigue or not enough time to complete the exercises, are common and must be dealt with by the therapist. Genital stimulation is eventually added to general body stimulation. The couple is sequentially taught to try various positions for intercourse without necessarily completing the act and to use varieties of stimulating techniques before they are permitted to proceed with intercourse.

The specific exercises vary with differing presenting complaints, and special techniques are used to treat the various dysfunctions. In cases of vaginismus, for instance, the woman is advised to dilate her vaginal opening with her fingers or with size-graduated vaginal dilators as part of the therapy. She may also be referred to a physiotherapist, specializing in pelvic disorders. In cases of premature ejaculation, an exercise known as the *squeeze technique* is used to raise the threshold of penile excitability. In this exercise, the man or the woman stimulates the erect penis until the earliest sensations of impending orgasm and ejaculation are felt. Penile

stimulation is then stopped abruptly, and the coronal ridge of the penis is forcibly squeezed for several seconds. The technique is repeated several times. A variation is the *stop–start technique,* in which stimulation is interrupted for several seconds but no squeeze is applied. Masturbation to the point of imminent orgasm raises the threshold of excitability to a more tolerant stimulation level. The man is encouraged to focus on sensations of excitement rather than distract himself from them. This makes him more familiar with his excitement pattern and lets him feel in control rather than overwhelmed by sensations of arousal. Communication between the partners is improved because the man must let his partner know his level of sexual excitement so that she can squeeze the penis before the ejaculatory process has started. Sex therapy has been successful with some premature ejaculators; however, a subgroup of dysfunctional men may need pharmacotherapy as well.

A man with sexual desire disorder or erectile disorder is sometimes told to masturbate to demonstrate that full erection and ejaculation are possible. A woman with lifelong female orgasmic disorder is directed to masturbate, sometimes using a vibrator. Kegel's exercises may be introduced to strengthen the pubococcygeal muscles—that is, the woman is encouraged to contract her abdominal and perineal muscles during masturbation and coitus. When a man has erectile disorder, the woman may be instructed to stimulate or tease his penis. The same technique is used with men who suffer from retarded ejaculation, with stimulation sometimes involving a vibrator. Delayed ejaculation is managed by extravaginal ejaculation initially and gradual vaginal entry after stimulation to the point of near ejaculation.

Treatment Goals. The overall goal of treatment is to initiate an educational process, to diminish the fears of performance felt by both sexes, and to facilitate communication in sexual and nonsexual areas. Therapy sessions follow each new exercise period, and problems and satisfactions (both sexual and nonsexual) are discussed. Specific instructions and new exercises geared to the individual couple's progress are reviewed in each session. Gradually, the couple gains confidence and learns (or relearns) to communicate verbally and sexually. Dual-sex therapy is most effective when the sexual dysfunction exists apart from other psychopathology.

Hypnotherapy

Hypnotherapists focus specifically on the anxiety-producing symptom—that is, the particular sexual dysfunction. Successful use of hypnosis helps the patient gain control over the symptom that has been lowering self-esteem and disrupting psychological homeostasis. Patient cooperation is first obtained and encouraged during a series of nonhypnotic sessions with

the therapist, designed to develop a secure doctor–patient relationship and a sense of physical and psychological comfort on the part of the patient and to establish mutually desired treatment goals. During that time, the therapist assesses the patient's capacity for the trance experience. The nonhypnotic sessions also permit the clinician to take a careful psychiatric history and do a mental status examination before beginning hypnotherapy. Treatment focuses on symptom removal and attitude alteration. In a trance state, patients can entertain ideas incongruent with their usual (nonhypnotized) perceptions of reality. Patients are instructed in developing alternative means of dealing with the anxiety provoking situation (i.e., the sexual encounter).

For example, a woman with vaginismus is given the posthypnotic suggestion that she will feel no pain during intercourse and will be able to relax the muscles surrounding her vagina. If compliance with the suggestion is successful, she can deal with the anxiety produced by the sex act. She is also taught new attitudes, such as being entitled to sexual pleasure. Under hypnosis, her fear or anger at sexual contact can be examined, and she learns how her emotions are expressed by involuntary vaginal spasms.

Some patients respond particularly well to the use of self-hypnosis and indirect suggestion. These techniques allow them to retain a greater sense of control over their situation. Typically, patients are instructed to conjure up images and develop ideas antithetical to their dysfunctional responses. For example, a woman with an arousal disorder may first agree to concentrate on imagery that causes her to salivate. She is then told that, just as she has made her mouth water by focusing on stimulating images, she can affect the lubricating response of her vagina by focusing on images she finds erotic or romantic. At the same time, the therapist helps her deal with her anxieties about a positive sexual response. Patients are also taught relaxing techniques to use before sexual relations. With these methods to alleviate anxiety, the physiological responses to sexual stimulation can more readily result in pleasurable excitation and discharge. Hypnosis may be added to a basic individual psychotherapy program to accelerate the impact of psychotherapeutic intervention.

Behavior Therapy

Behavior therapists assume that sexual dysfunction is learned, maladaptive behavior. Behavioral approaches were initially designed to treat phobias. In cases of sexual dysfunction, the therapist sees the patient as phobic of sexual interaction. Using traditional techniques, the therapist sets up a hierarchy of anxiety-provoking situations for the patient, ranging from the least threatening to the most threatening. Mild anxiety may be experienced at the thought of kissing, and massive anxiety may be felt when imagining penile penetration. The behavior therapist helps the patient master the

anxiety through a standard program of systematic desensitization. The program is designed to inhibit the learned anxious response by encouraging behaviors antithetical to anxiety. The patient first deals with the least anxiety-producing situation in fantasy and progresses by steps to the most anxiety-producing situation. Medication, hypnosis, or special training in deep muscle relaxation is sometimes used to help with the initial mastery of anxiety.

Assertiveness training helps teach patients to express their sexual needs openly and without fear. Exercises in assertiveness are given in conjunction with sex therapy, and patients are encouraged both to make sexual requests and to refuse to comply with requests perceived as unreasonable. Sexual exercises may be prescribed for patients to perform at home, and a hierarchy may be established, starting with activities that proved most pleasurable and successful in the past.

One treatment variation involves the participation of the patient's sexual partner in the desensitization program. The partner, rather than the therapist, presents the hierarchical items to the patient. In such situations, a cooperative partner is necessary to help the patient carry gains made during treatment sessions to sexual activity at home.

Behavior therapy techniques have been particularly effective in treating women with severe inhibition of excitement and orgasm when such feelings were accompanied by strong feelings of anxiety, anger, or disgust.

Group Therapy

Methods of group therapy have been used to examine both intrapsychic and interpersonal problems in patients with sexual disorders. The therapy group provides a strong support system for patients who feel ashamed, anxious, or guilty about a particular sexual problem. It is a useful forum in which to counteract sexual myths, correct misconceptions, and provide accurate information regarding sexual anatomy, physiology, and varieties of behavior.

Groups for the treatment of sexual disorders can be organized in several ways. Members may all share the same problem, such as premature ejaculation; members may all be of the same sex and have different sexual problems; or groups may be composed of both men and women who are experiencing different sexual problems. Group therapy may be an adjunct to other forms of therapy or the prime mode of treatment. Groups organized to cure a particular dysfunction usually have a behavioral approach. For example, patients with anorgasmia may participate with others who have the same problem in a short-term, intensive group experience. Sexual histories, feelings of inadequacy, and concerns about body image are shared. Specific physiological information, sometimes with the aid of audiovisual materials, is presented to the group members. Members are given homework assignments (e.g., they may be instructed to

masturbate). A combination of group support and group pressure helps some of the participants complete assignments they might otherwise avoid. As the short-term group process nears termination, members are encouraged to talk about their experiences with their partners.

Groups have also been effective when composed of sexually dysfunctional married couples. The group provides an opportunity to gather accurate information, provides consensual validation of individual preferences, and enhances self-esteem and self-acceptance. Such techniques as role-playing and psychodrama may be used in treatment. These groups are not indicated for couples in which one partner is uncooperative, has severe depression or psychosis, or has a strong repugnance for explicit sexual audiovisual materials or a strong fear of groups.

Integrated Sex Therapy

One of the most effective treatment modalities is the use of sex therapy integrated with supportive, psychodynamic, or insight-orientated psychotherapy. Adding psychodynamic conceptualizations to the behavioral techniques used to treat sexual dysfunctions allows treatment of patients with sex disorders associated with other psychopathology. Also, this type of therapy is appropriate for patients with hypoactive desire disorders, or interest/arousal disorder. Insight-oriented therapy helps them deal with problems in their interpersonal relationships or intrapsychic conflicts that frequently are at the root of the problem. The themes and dynamics that emerge in patients in analytically oriented sex therapy are the same as those that emerge in psychoanalytic therapy—relevant dreams, fear of punishment, aggressive feelings, difficulty with trusting the partner, fear of intimacy, oedipal feelings, and fear of genital mutilation.

The case that follows demonstrates some of these dynamics.

Mr. K, a 62-year-old divorced man, presented for therapy with a chief complaint of erectile dysfunction with his current partner, Miss B. They had been living together for 4 years. In the beginning of their relationship they had had an active and nondysfunctional sex life. He was a successful businessman who met Miss B at a business convention where she was a hostess/welcomer at the convention center. Mr. K was very attracted to her. He also liked her poise and social ease and felt he could be proud of her.

Shortly after they began living together he began seeing other women as well. Miss B eventually found out and insisted that he stop seeing other women. She also demanded that they get married. Mr. K agreed to stop his infidelities and agreed to become engaged. His erectile difficulties with Miss B, which had started somewhat before this confrontation, became persistent and pronounced. Even though he was able to achieve and maintain erections if he used sildenafil, he began avoiding sex with her altogether. He did not think of leaving the relationship (although he broke off the engagement), because he felt he was aging, feared loneliness and lacked confidence that he could find another desirable woman. Also, he had experienced similar problems in his marriage and had enough insight to realize that his pattern of behavior was destructive. (Miss B had her own set of issues which kept her in the

relationship.) She had become depressed, but was more hopeful since Mr. K had agreed to start therapy.

Mr. K had a poised, secure façade, but he was hypersensitive to criticism, frequently felt like an outsider in social groups, relied on reassurance from women for his self-worth, and dealt with his anxieties with denial and sexual impulsiveness.

Mr. K's issues reflected experiential and developmental problems. It had been hard for him to establish peer relationships growing up due to the family's frequent travels and relocations. His father was a moderately successful, demanding man who was competitive with Mr. K and would send him into inappropriate situations. For example, he would send him on errands in a new location where, as a boy, Mr. K did not know his way around, and be critical of Mr. K if he failed in the errand; or he would tell him to "stand up for yourself" with boys who were substantially older and stronger than the patient. At the same time Mr. K was his mother's "darling, her everything." Resulting, unresolved oedipal issues made it difficult for him to direct both tender and passionate feelings toward the same woman. Unresolved dependency needs made it hard for him to discuss differences he had with Miss B in a constructive, assertive way and he denied cumulative resentments (appropriate or inappropriate).

These issues were addressed in introspective psychotherapy. Miss B was also referred for individual therapy, and the couple was seen together for behaviorally oriented sex therapy.

The combined approach of individual and sex therapy is used by the general psychiatrist, who carefully judges the optimal timing of sex therapy and the ability of patients to tolerate the directive approach that focuses on their sexual difficulties.

Biological Treatment Methods

Biological treatments, including pharmacotherapy and surgery, have applications in specific cases of sexual disorder. Advances in biological treatment methods, particularly pharmacological methods of treating sexual dysfunction have significantly augmented the therapist's catalog of therapeutic approaches. Most of the recent advances involve male sexual dysfunctions. Current studies are under way to test pharmacological treatment of sexual dysfunctions in women. A new drug, flibanserin has been approved by the Federal Drug Administration (FDA) to treat low sexual interest in women. Questions about the drug included a report of minimal increase in desire during trials and concerns about side effects, particularly hypotension and syncope. Physicians must undergo a certification process to prescribe flibanserin.

Pharmacotherapy. A variety of drugs have been explored in the treatment of male sexual dysfunction. The major medications used are nitric oxide enhancers such as sildenafil (Viagra); oral prostaglandin (Vasomax); alprostadil (Idex), an injectable phentolamine; and a transurethral alprostadil (MUSE), all used in the treatment of erectile disorder.

NITRIC OXIDE ENHANCERS. Nitric oxide enhancers facilitate the inflow of blood to the penis necessary for an erection. The physiological mechanism of penile erection involves release of nitric oxide in the corpus cavernosum during sexual stimulation. Nitric oxide activates the enzyme guanylate cyclase, which increases cyclic guanosine monophosphate concentration and produces smooth muscle relaxation in the corpus cavernosum that

allows the penile vessels to dilate and admit blood. The first drug developed, sildenafil, enhances the effect of nitric oxide by inhibiting the enzyme that degrades cyclic guanosine monophosphate. Thus, sildenafil augments the natural process involved in gaining and maintaining an erection during sexual stimulation. The drug takes effect approximately 1 hour after ingestion, and its effect can last up to 4 hours. Sildenafil has no effect in the absence of sexual stimulation.

The most common adverse events associated with sildenafil are headaches, flushing, and dyspepsia. Some sildenafil users see things in a blue tint for several hours after taking the medication; because of this, airline pilots have been prohibited from taking the drug so that this visual artifact does not interfere with safe landings. A more serious visual side effect is the loss of visual acuity and nonarteritic anterior ischemic optic neuropathy resulting in blindness. Although the occurrence is extremely rare, phosphodiesterase type 5 inhibitors should not be used with men with serious optic pathology. The use is also contraindicated for people taking organic nitrates. The concomitant action of the two drugs can result in large, sudden, and sometimes fatal drops in systemic blood pressure. The U.S. Food and Drug Administration (FDA) has posted 522 deaths as of 2014, in which sildenafil was listed as an associated medication and advises caution in prescribing sildenafil to men with a recent (6-month) history of myocardial infarction, stroke, life-threatening arrhythmia, significant hypotension or hypertension, cardiac failure, angina, or retinitis pigmentosa. Sildenafil has also been implicated in cases of deafness. One study of Swedish men using phosphodiesterase 5 inhibitors found a modestly significant correlation between use and the presence of malignant melanoma. The researchers stated that this association was not proved to be causal. In association with side effect alerts, the FDA has reiterated that it is a safe drug. Sildenafil is not effective in all cases of erectile dysfunction. It fails to produce an erection rigid enough for penetration in approximately 50 percent of men who have had radical prostate surgery or in those with long-standing, insulin-dependent diabetes. It is also ineffective in certain cases of nerve damage.

Two other nitric oxide enhancers similar to sildenafil—vardenafil (Levitra) and tadalafil (Cialis)—have been developed. Tadalafil has an effective therapeutic window of 36 hours compared to sildenafil and vardenafil, which are effective for approximately 4 hours.

Oral phentolamine has proved effective as a potency enhancer in men with minimal erectile dysfunction. It may prove useful for men with cardiac problems, as sildenafil is contraindicated for men using organic nitrates, but it is not currently approved by the FDA. Apomorphine is also being tested as an oral remedy for erectile dysfunction.

ALPROSTADIL. In contrast to the oral medications, injectable and

transurethral alprostadil act locally on the penis and can produce erections in the absence of sexual stimulation. Alprostadil contains a naturally occurring form of prostaglandin E. Prostaglandins are composed of complex hydroxy fatty acids, and they have wide biological influences. Although some prostaglandins are vasoconstrictive, prostaglandin E_1, found in alprostadil, is a powerful vasodilating agent, especially in local vascular areas, such as the corpus cavernosum of the penis. The drug causes direct smooth muscle relaxation of penile vessels and erectile tissue; this reaction lowers the vascular resistance of the corpus and significantly increases blood flow to the penis. The firm erection produced within 2 to 3 minutes by increased blood flow may last as long as 1 hour. Treatment consists of the patient's self-injection of alprostadil into the corpus before coitus. This technique is easily taught and relatively painless. Infrequent adverse effects include penile bruising and changes in liver function test results, which are readily reversible when a man stops the injections. However, possible hazardous sequelae exist, including priapism and sclerosis of the small veins of the penis. In some men for whom alprostadil was not effective, an injection of Trimix, a combination of prostaglandin E, phentolamine (an alpha l–adrenergic antagonist) and papaverine (a nonspecific phosphodiesterase inhibitor) produced an erection. Another substance being tried is vasoactive intestinal polypeptide (VIP). Intracavernous injection of VIP causes erection and has a parasympathomimetic effect. Some researchers speculate that this substance, which has been found in the hypothalamus and the female genital organs, is the essential factor in male and female arousal. In Europe, phenoxybenzamine (Dibenzyline) is used to produce erections by injection into the penis. Serious adverse effects include priapism and pain accompanying the injection, and the drug is not allowed as a therapy in the United States.

Alprostadil can also be delivered via the urethra, eliminating the need for self-injection. Some men prefer the local, nonsystemic effects of alprostadil to oral sildenafil; others prefer sildenafil because it seems more like a nonpharmacologically aided response to them and their partners. A small trial found a topical cream effective in alleviating erectile dysfunction. The cream consists of three vasoactive substances that are a mixture of ergot alkaloids. Alprostadil is a useful treatment for patients in whom nitric oxide enhancers are contraindicated (e.g., patients who are taking nitrate-containing medication). It is also of use in patients who do not tolerate the side effects of other drugs, and may work for some patients who have found sildenafil ineffective, for example in some postprostatectomy patients.

The pharmacological treatments described above are useful in treatment of erectile dysfunction of various causes: psychogenic, neurogenic, arterial insufficiency, venous leakage, and mixed. The

following case demonstrates the use of pharmacotherapy to treat erectile dysfunction of psychogenic origin.

Mr. J presented for therapy with a chief complaint of erectile disorder. He was a 49-year-old man who had been divorced for 6 years after a marriage of 12 years duration. Although he suffered no erectile dysfunction in his marriage, and in fact experienced strong desire, his wife refused to have sex with him in the last 2 years of their marriage. By that time their relationship had seriously deteriorated. Although it was Mr. J's decision to end the marriage, he felt hurt and humiliated by his ex-wife's sexual rejection and wary of ever committing to someone else. His sense of rejection evoked memories of a painful adolescence when he had felt very inadequate with girls and fearful of being mocked by male peers. After his divorce he had a series of dating relationships with women but suffered from recurrent erectile disorder with all of them.

When he came for therapy he was in the midst of a new relationship. The patient was prescribed sildenafil, 50 mg, but he was reluctant to tell his girlfriend that he was taking a medication to help him function sexually. Instead, he avoided sex with her. Several therapy sessions were devoted to encouraging him to share his sexual problem with her. He finally did so and his girlfriend, a woman with confidence in her sexuality, was accepting of his problem. In fact, she encouraged him to take the sildenafil. After 3 months of an active sex life without erectile problems, Mr. J was told to try intercourse without sildenafil. With the emotional support of his girlfriend he made a successful attempt at coitus. The patient continues to use sildenafil occasionally. The knowledge that he has a "back-up" to help him with his erection if he is feeling very anxious or stressed enables him to maintain a normally active sex life.

Erectile dysfunction of psychological or mixed origin should not be treated by medication alone, even though the dysfunction can be corrected pharmacologically. The drugs should be used in conjunction with, not as a replacement for, sex therapy. In some cases, erectile dysfunction serves as a defense against unconscious conflicts that must be faced when the sexual symptom is removed as result of medication. In some cases, the patient deals with his conflicts by simply not using the medications prescribed for his erectile problem. Also, patients in long-standing or marital relationships require the cooperation of their partners for this treatment to be effective. Although most women are very accommodating to their partner's desire for treatment, several concerns occur with some frequency: the element of romance is an important part of sexual interaction for many women, and pharmacological assistance of erections may eliminate that sense of romance; the woman, or the couple, may bemoan the lack of spontaneity when pharmacotherapy is part of the sex act; and some women feel deprived of feedback about their desirability. Joint sessions help the couple cope with these issues.

SILDENAFIL USE IN WOMEN. The physical sign of sexual excitement in women is vaginal lubrication. That lubrication is a transudate believed to result from increased vasocongestion of the extensive capillary net surrounding the vaginal walls. The same physiological mechanism—vasocongestion—results in erection in the man. Researchers believe that sildenafil may facilitate blood flow in women just as it does in men and, thus, give women with inhibited excitement of psychological or physiological origin a

pharmacological remedy for their dysfunction. Some reports suggest that women find it preferable to apply sildenafil in a cream base to the clitoris, labia, and vagina. However, as in men, pharmacological treatment may need to be combined with psychological modalities to be effective.

Recognition of sexual excitement may be more complex for women than it is for men. For example, in studies of response to pornography, men and women underwent physiological measurements of excitement after visual stimuli. In these studies, men were more accurate in correlating their subjective sense of arousal with physiological arousal (i.e., erection) than were women (i.e., lubrication).

Other Pharmacological Agents

ANTIDEPRESSANTS. The side effects of antidepressants particularly the SSRIs which include delayed orgasm have been used to prolong the sexual response in patients with premature ejaculation. This approach is particularly useful in patients refractory to behavioral techniques or who may have physiologically determined premature ejaculation. Daily treatment is recommended. The SSRIs vary in their ability to delay ejaculation with paroxetine (20 to 40 mg) being the most effective, followed by clomipramine (10 to 50 mg), sertraline (50 to 100 mg), fluoxetine (20 to 40 mg), ditalopram (20 to 40 mg), and escitalopram (10 to 20 mg). On demand use of SSRIs can be tried, but is less effective. However, on demand use of clomipramine (20 to 40 mg) may delay ejaculation after 4 to 6 hours although it may lead to nausea on the same day or the next day. Long-term adverse effects of daily SSRI use include weight gain with its associated risk for type 2 diabetes mellitus. Also, depressed patients must be monitored for suicidal thoughts and bipolar patients monitored for mania. Other on demand treatments for premature ejaculation include topical anesthetics such as lidocaine or prilocaine cream or spray which is applied to the glans of the penis. Use must be titrated to prevent sufficiently strong anesthesia that interferes with erection. Sometimes condom use is necessary to prevent vaginal numbness.

Other Pharmacological Agents. Numerous other pharmacological agents have been used to cure the various sexual disorders. Intravenous methohexital sodium (Brevital) has been used in desensitization therapy. Antianxiety agents may have some application in tense patients, although these drugs can also interfere with the sexual response. The use of antidepressants has been advocated in the treatment of patients who are phobic of sex and in those with a posttraumatic stress disorder (PTSD) after rape. Bromocriptine (Parlodel) is used in the treatment of hyperprolactinemia, which is frequently associated with hypogonadism. Such cases are first worked up to rule out pituitary tumors. Bromocriptine,

a dopamine agonist, may improve sexual function impaired by hyperprolactinemia.

Yohimbine is an alpha-adrenergic receptor antagonist that may cause dilation of the penile artery and improve erections. Recreational drugs, including cocaine, amphetamines, alcohol, and cannabis, are considered enhancers of sexual performance. Although they may provide the user with an initial benefit because of their tranquilizing, disinhibiting, or mood-elevating effects, consistent or prolonged use of any of these substances impairs sexual functioning.

Ginseng has been reported to have androgenic effects. One report described the case of a mother who ingested large amounts of ginseng during her pregnancy, resulting in androgenization of the neonate, who was born with pubic hair and enlarged testes.

Other drugs that have been used by women to alleviate arousal dysfunction include oral phentolamine, topical prostaglandin E, oral oxytocin, ginkgo biloba, and various psychostimulants, including caffeine. Many of these drugs have not been approved for treatment of female sexual dysfunction and must be prescribed with caution.

Dopaminergic agents have been reported to increase libido and improve sex function. Those drugs include L-dopa, a dopamine precursor, and bromocriptine, a dopamine agonist. The antidepressant bupropion has dopaminergic effects and has increased sex drive in some patients. Selegiline, an MAOI, is selective for MAO_B and is dopaminergic. It improves sexual functioning in older persons.

HORMONE THERAPY. Androgens increase the sex drive in women and in men with low testosterone concentrations. Women may experience virilizing effects, some of which are irreversible (e.g., deepening of the voice). This problem is avoided if the woman applies a mild (compounded) testosterone ointment to her labia, inner thighs or abdomen. Ongoing care by an endocrinologist is necessary for women who require more intensive testosterone treatment. Testosterone can be administered to men through injections, transdermal patches, or as gels, creams or pellets. Oral preparations are associated with increased risk of hepatotoxicity. Side effects may include hypertension and prostate enlargement among other side effects. However, recent studies question effects on the prostate.

Clomiphene, usually used as a fertility drug for women, also stimulates testosterone production in men by action on the pituitary-gonad feedback cycle.

Gonadotrophin-releasing hormone (GnRH), also known as *luteinizing hormone-releasing hormone* (LHRH), stimulates the release of luteinizing hormone, which increases testosterone secretion in both sexes. GnRH is used as an inhalant in Europe. It stimulates desire and increases potency. Because GnRH is released normally in a pulsatile fashion, portable infusion

pumps have been developed that simulate pulsatile delivery. An excess of GnRH suppresses estrogen and testosterone; thus, the therapeutic use of GnRH is limited by a narrow therapeutic window.

Women who use estrogens for HRT or for contraception may report decreased libido; in such cases, a combined preparation of estrogen and testosterone has been used effectively. Estrogen itself prevents thinning of the vaginal mucous membrane and facilitates lubrication. It may also increase tactile sensitivity. Vaginal estrogen creams have been helpful to postmenopausal women who cannot take HRT due to their medical or family history.

Testosterone is given to women around the world in the form of tablets, implants, patches, and creams. In the United States, methyltestosterone pills and compounded creams are the primary methods of administration. Testosterone is primarily administered to postmenopausal (occurring surgically or naturally) women, but testosterone deficiency also exists in premenopausal women.

Mrs. M presented for therapy alone with a chief complaint of pain on intercourse dating from the onset of menopause 14 months before coming for treatment. Her husband knew of her visit but did not accompany her, as he was away on one of many routine and frequent business trips.

Sexual intercourse was infrequent because of his heavy travel schedule, but their sex play was varied and, before menopause, Mrs. M had enjoyed sex. Mrs. M found her sexual interactions with her husband gratifying. She had a strong libido and would masturbate when Mr. M was away.

Mrs. M was on oral HRT, which had helped but had not eliminated the pain she had with intercourse. In consultation with her gynecologist, the occasional use of hormone cream was added to her treatment regimen. Additionally, she was instructed to use vaginal dilators routinely when Mr. M was away. This was not for purposes of masturbation, which for her involved clitoral stimulation, but for routine stretching of the vagina. Nonhormonal vaginal cream was prescribed for the couple to use with intercourse. The combination of oral HRT, vaginal creams, and the routine use of vaginal dilators served to eliminate Mrs. M's pain. She was diagnosed with sexual dysfunction due to a general medical condition.

Pheromones are sexual scents that are found in animals and may be present in humans. They produce dramatic sex-seeking behavioral patterns in animals (e.g., male deer following female deer in estrus, mounting behavior in primates). Human pheromones are believed to be short-acting fatty acids present in vaginal secretions and male sweat. In one study, women were consistently more attracted to items impregnated with a chemical derived from male sweat (alpha-androstenol) than to control items. In another study, the sweat of women was preserved in underarm pads they changed daily, and the date each pad was worn was correlated with the women's menstrual cycle. A second group of women was asked to smell these pads as they were rubbed above their upper lips, without knowing what they were and recognizing no scent but the alcohol preservative used on the pads. Depending on whether they were exposed to

pads from the early or late part of the first group's (the wearers) menstrual cycle, the second group (the sniffers) saw their own menstrual cycles shortened or lengthened. This area is still being researched.

ANTIANDROGENS AND ANTIESTROGENS. Estrogen and progesterone are antiandrogens that have been used to treat compulsive sexual behavior in men, usually in sex offenders.

Medroxyprogesterone acetate (Depo-Provera), used primarily as a contraceptive in women, inhibits the secretion of gonadotrophins. It is used in men with compulsive sexual behavior to reduce libido by lowering testosterone levels. Cyproterone acetate is a strong antiandrogen used in Europe to treat sex offenders. At dosages of 100 to 200 mg a day, the sex drive disappears within 2 weeks.

Tamoxifen (Nolvadex) is used to treat breast cancer and has antiestrogenic properties. In some women, increased libido may result from the unopposed testosterone levels. Clomiphene (Clomid) is an ovulatory stimulant used in women with anovulatory menstrual cycles desiring pregnancy. Clomiphene increases release of pituitary gonadotrophins, and some women may report increased libido. Neither of these drugs, however, is used as a treatment for decreased sexual interest in women.

Mechanical Treatment Approaches

STEAL SYNDROME. In male patients with arteriosclerosis (especially of the distal aorta, known as *Leriche syndrome*), the erection may be lost during active pelvic thrusting. The need for increased blood in the gluteal muscles and others served by the ilial or hypogastric arteries takes blood away (steals) from the pudendal artery and, thus, interferes with penile blood flow. Relief may be obtained by decreasing pelvic thrusting, which is also aided by the woman-superior coital position.

VACUUM PUMP. Vacuum pumps are mechanical devices that patients without vascular disease can use to obtain erections. The blood drawn into the penis after the creation of the vacuum is kept there by a ring placed around the base of the penis. This device has no adverse effects, but it is cumbersome, and partners must be willing to accept its use. Some women complain that the penis is redder and cooler than when erection is produced by natural circumstances, and they find the process and the result objectionable.

A vacuum pump with the marketing name *Eros* was devised for women to precipitate a clitoral erection. It works similarly to the vacuum pump used to create penile erections in men, although no band is placed around the clitoris once erection is obtained. The pump is not necessary for intercourse to occur; it was devised to enhance female excitement, but its

effectiveness has not been studied rigorously.

Surgical Treatment

MALE PROSTHESES. Surgical treatment is infrequently advocated, but improved penile prosthetic devices are available for men with inadequate erectile response who are resistant to other treatment methods or who have medically caused deficiencies. There are two main types of prosthesis: a semirigid rod prosthesis that produces a permanent erection that can be positioned close to the body for concealment and an inflatable type that is implanted with its own reservoir and pump for inflation and deflation. The latter type is designed to mimic normal physiological functioning. Placing a penile prosthesis in a man who has lost the ability to ejaculate or to have an orgasm as a result of medical causes will not restore those functions. Men with prosthetic devices have generally reported satisfaction with their subsequent sexual functioning, but their wives report much less satisfaction. Presurgical counseling is strongly recommended so that the couple has a realistic expectation of what the prosthesis can do for their sex lives. Postsurgical counseling may also be necessary to help the couple adapt to their rediscovered ability to have intercourse. They may experience a high level of anxiety if their sex life had been inactive for a prolonged period before surgery. Prosthetic devices have been associated with severe adverse effects, including perforation, infection, urinary retention, and persistent pain.

Some surgeons are attempting revascularization of the penis as a direct approach to treating erectile dysfunction resulting from vascular disorders. Such surgical procedures are indicated in patients with corporal shunts that allow normally entrapped blood to leak from the corporal spaces, leading to inadequate erections (steal syndrome). Limited reports exist of prolonged success with the technique. Endarterectomy can be of benefit if aortoiliac occlusive disease is responsible for the erectile dysfunction.

Another medical treatment being studied for erectile disorders is electrostimulation at the base of the penis. Initial reports indicate minimal physical discomfort in patients receiving this therapy. However, response to treatment is inconsistent, and a problem exists in terms of maintaining erections. At the present time, the treatment seems to have no benefits.

FEMALE PROCEDURES. Surgical approaches to female dysfunctions include hymenectomy in the case of dyspareunia in an unconsummated marriage, vaginoplasty in multiparous women complaining of lessened vaginal sensations, or freeing clitoral adhesions in women with inhibited excitement. Such surgical treatments have not been carefully studied and should be considered cautiously.

Outcome. Demonstrating the effectiveness of the varieties of sex

therapy is just as difficult as assessing the effectiveness of outpatient psychotherapy in treating other psychological problems. As with other disorders, the more severe the psychopathology associated with a problem of long duration, the more adverse the outcome is likely to be.

The most difficult treatment cases involve couples with severe marital discord. Cases involving problems of fear of intimacy, excessive dependency, or excessive hostility are also complex. Other challenges are posed by patients with impulse disorders, unresolved conflicts about sexual orientation or fetishistic defenses. Patients phobic of sex also present treatment difficulties, as do patients diagnosed with lifelong dysfunctions. Disorders of desire or interest are particularly difficult to treat. They require longer, more intensive therapy than some other disorders, and their outcomes are very variable.

When behavioral approaches are used, empirical criteria that are supposed to predict outcome are more easily isolated. Using these criteria, for instance, couples who regularly practice assigned exercises appear to have a much greater likelihood of successful outcome than do more resistant couples or those whose interaction involves sadomasochistic or depressive features or mechanisms of blame and projection. Flexibility of attitude is also a positive prognostic factor. Overall, younger couples tend to complete sex therapy more often than older couples. Couples whose interactional difficulties center on their sex problems, such as inhibition, frustration, or fear of performance failure, are also likely to respond well to therapy.

In general, methods that have proved effective singly or in combination include training in behavioral sexual skills, systematic desensitization, directive marital counseling, traditional psychodynamic approaches, group therapy, and pharmacotherapy. Although most prefer to treat a couple for sexual dysfunctions, treatment of individual persons has also been successful. The frequency of sessions is not a significant factor in treatment success. Thus, whether patients have intensive daily therapy over a period of 2 weeks, weekly therapy, or biweekly therapy appears to have little effect on the outcome of treatment. Also, the use of one therapist to treat a couple instead of a dual-sex cotherapy team is nearly as effective and certainly more practical. A multimodal regimen and an eclectic approach are frequently necessary to treat sexual disorders. Many couples who do not reach their ideal goal of sexual interaction, nonetheless, report considerable improvement and satisfaction in their sex lives and their general sense of well-being.

References

Araoz DL. Uses of hypnosis in the treatment of psychogenic sexual dysfunctions. *Psychiatr Ann*. 1986;16:102.

Bogaert AF. Asexuality: What it is and why it matters. *J Sex Res*. 2011;52(4):362–379.

Brady JP. Behavior therapy and sex therapy. *Am J Psychiatry*. 1976;133:896–899.

Brassard A, Dupuy E, Gergeron S, Shaver PR. Attachment insecurities and women's sexual function and satisfaction: the mediating roles of sexual self-esteem, sexual anxiety, and sexual assertiveness. *J Sex Res*. 2015;52(1):110–119.

Carvalho J, Veríssimo A, Nobre PJ. Psychological factors predicting the distress to female persistent genital arousal symptoms. *J Sex Marital Ther*. 2015;41(1):11–24.

Coady D and Fish N. *Healing Painful Sex*. Berkley, CA: Seal Press; 2011.

*Delgardo PL, McGauey CA, Moreno FA, Laukes C, Gelenberg AJ. Treatment strategies for depression and sexual dysfunction. *J Clin Psychiatry*. 1999;17:22.

Federman DD. Perspective: three facets of sexual differentiation. *N Engl J Med*. 2004;350:323–324.

Goldstein I, Lue T, Padma-Nathatan H, Rosen R, Steers WD, Wicker PA. Oral sildenafil in the treatment of erectile dysfunction. Sildenafil Study Group. *N Engl J Med*. 1998;338:1397–1404.

Grubbs JB, Volk F, Exline JJ, Pargament KI. Internet pornography use: perceived addiction, psychological distress, and the validation of a brief measure. *J Sex Marital Ther*. 2015;41(1):83–106.

Koppelman M, Parry BL, Hamilton JA, Alogna SW, Loreaux PL. Effect of bromocriptine on affect and libido in hyperprolactinemia. *Am J Psychiatry*. 1987;144:1037–1041.

Krafft-Ebing R. *Psychopathia Sexual*. Munich: Matthes & Seitz; 1984.

*Laughman E, Gagnon J, Michael R, Michaels S. *Sex in America*. Chicago: University of Chicago Press; 1994.

Leitenberg H, Detzer M, Srebnik D. Gender differences in masturbation and the relation of masturbation experience in preadolescence and/or early adolescence to sexual behavior and sexual adjustment in young adulthood. *Arch Sex Behav*. 1993;22:87–98.

Linet OI, Ogrinc FG. Efficacy and safety of intracavernosal alprostadil in men with erectile dysfunction. The Alprostadil Study Group. *N Engl J Med*. 1996;334:873–877.

MacLaughlin DT, Donahoe PK. Mechanisms of disease: Sex determination and differentiation. *N Engl J Med*. 2004;350:369.

*Masters WH, Johnson VE. *Human Sexual Response*. Boston: Little, Brown; 1970.

McCarthy B, Pierpaoli C. Sexual challenges with aging: Integrating the GES approach in an elderly couple. *J Sex Marital Ther*. 2015;41(1):72–87.

Negash S, Sheppard N, Lambert N, Fincham FD. Trading later rewards for current pleasure: pornography consumption and delay discounting. *J Sex Res*. 2016;53(6):689–700.

Padma-Nathan H, Hellstrom WJ, Kaiser FE, et al. Treatment of men with erectile dysfunction with transurethral alprostadil. *N Engl J Med*. 1997;336:1–7.

*Purnine DM, Carey MP, Jorgensen RS. Gender differences regarding preferences for specific heterosexual practices. *J Sex Marital Ther*. 1994;20:271–287.

Rhoden EL, Morgentaler A. Risks of testosterone-replacement therapy and recommendations for monitoring. *N Engl J Med*. 2004;350:482–492.

Riley AJ. Life-long absence of sexual drive in a woman associated with 5-dihydrotestosterone deficiency. *J Sex Marital Ther*. 1999;25:73–78.

*Rosen RC, Lane RM, Menza M. Effects of SSRIs on sexual function: A critical review. *J Clin Psychopharmacol*. 1999;19:67–85.

*Sadock VA. The treatment of psychosexual dysfunctions: an overview. In: Grinspoon L, ed. *Psychiatry 1982. The American Psychiatric Association Annual Review*. Washington, DC: American Psychiatric Press; 1982.

Sadock VA. Group psychotherapy of psychosexual dysfunctions. In: Kaplan HI, Sadock BJ, eds. *Comprehensive Group Psychotherapy*. Baltimore, MD: Williams & Wilkins; 1983:286.

Sarin S, Amsel R, Binik YM. A streetcar named "Derousal"? A psychophysiological examination of the desire-arousal distinction in sexually functional and dysfunctional women. *J Sex Res*. 2016;53(6):711–729.

*Semans JH. Premature ejaculation: a new approach. *South Med J*. 1956;49:353–358.

Stein DJ, Hollander E, Anthony DT, Schneier FR. Serotonergic medications for sexual addictions, and paraphilias. *J Clin Psychiatry*. 1992;53:267–271.

Sternbach H. Age associated testosterone decline in men: clinical issues for psychiatry. *Am J Psychiatry*. 1998;155:1310–1318.

Van Houdenlove E, Gijs L, T'Sjoen G, Enzlin P. Asexuality: a multidimensional approach. *J Sex Res*. 2015;52(6):669–678.

Vrangalova Z. Hooking up and psychological well-being in college students: short-term prospective links across different hook-up definitions. *J Sex Res*. 2015;52(5):485–498.

Waldinger MD, Hengeveld WH, Zwinderman A, Olivier B. Effect of SSRI antidepressants on ejaculation: a double-blind, randomized, placebo-controlled study with fluoxetine, fluvoxamine, paroxetine and sertraline. *J Clin Psychopharmacol*. 1998;189:274–281.

▲ 21.2 Homosexuality, Gay and Lesbian Identities, and Homosexual Behavior

JACK DRESCHER, M.D., AND WILLIAM M. BYNE, M.D., PH.D.

Contemporary psychiatry and other mental health professions today use a model of homosexuality that considers same-sex feelings, behaviors, and relationships as normal variations of human sexuality. This normal variant model informs the positions taken by the American Psychiatric Association (APA), the American Psychological Association, and the American Psychoanalytic Association (APsaA) in support of equal rights for homosexual individuals, including marriage equality (same-sex marriage). These positions are an outgrowth of the APA's 1973 conclusion that homosexuality is not a mental disorder and its decision to remove homosexuality per se from its *Diagnostic and Statistical Manual of Mental Disorders* (DSM). Not only did this action play a major role in the cultural normalization of same-sex relationships in the United States, but APA's diagnostic reclassification had an international impact as well. In 1990, the World Health Organization (WHO) made a similar change in the tenth revision of their *International Statistical Classification of Diseases and Related Health Problems* (ICD-10). While homosexuality *per se* is no longer considered a mental disorder in the ICD, at the time of this writing, WHO is planning to delete all remaining diagnoses that pathologize homosexuality such as "ego dystonic sexual orientation."

Once psychiatry and medicine focused on questions like "What causes homosexuality?" or "What can be done about changing homosexuality?" Increasingly today's mental health policy foci are the mental health needs of lesbian, gay, and bisexual (LGB) patients. This shift from historical pathological models of homosexuality to a normal variant model has been accompanied by a dramatically increased social acceptance of openly gay, lesbian, and bisexual individuals in Western societies. This acceptance continues to rapidly evolve in the West. For example, the 8th edition of this chapter noted same-sex marriage was then only legal in Belgium and the Netherlands. By 2009s 9th edition, access to legally sanctioned same-sex marriage had expanded to Canada, South Africa, Spain, and the US states of Massachusetts and Connecticut, while Israel was recognizing same-sex marriages performed in other countries. Today marriage equality is increasingly available in Western Europe, additionally including Denmark, Finland, France, Iceland, Ireland, Luxembourg, Norway, Portugal, Sweden, and the United Kingdom. It is also available in Argentina, Brazil, Colombia, Mexico, New Zealand, and Uruguay. In 2013, in *United States v. Windsor*, the US Supreme Court overturned Section 3 of the discriminatory Defense

of Marriage Act (DOMA) that prevented the federal government from recognizing same-sex marriages in states where such marriages were legal. The Court went a step further in 2015 in *Obergefell v. Hodges*. It overturned thirteen remaining state constitutional bans and made marriage equality legal in all 50 states. Civil unions, although providing fewer legal protections than marriage, are now available in Austria, Chile, Costa Rica, Croatia, Cyprus, Czech Republic, Ecuador, Estonia, Germany, Greece, Hungary, Italy, Lichtenstein, Malta, Northern Ireland, Slovenia, Switzerland, and in some areas of Australia and Taiwan.

Further, gay people serve openly in the military of over 40 countries. In 2011, the US congress repealed "Don't Ask, Don't Tell," the policy that prevented gay people from serving openly in the military.

On the other hand, while there has been growing acceptance of homosexuality in Western countries, this is not the case in the rest of the world. In 2013, Russia passed a law criminalizing the distribution of "propaganda" to minors that supports "nontraditional sexual relationships," generally taken as meaning homosexuality. Other former Soviet states are drafting similar laws. Engaging in homosexual behavior is considered a crime in more than 70 countries and in some of them it is punishable by imprisonment or death. Despite efforts by the United Nations and other international human rights groups to have such legislation repealed, some African countries such as Gambia, Nigeria, and Uganda recently have passed laws increasing criminal penalties for expressions of or even support for homosexuality.

DEFINITIONS

Common terms used to describe human sexuality have a range of meanings, not only in common parlance, but among medical professionals and authoritative texts as well. The definitions of these terms as they are used in this section are presented here. *Sexual orientation* refers to a person's erotic response tendency or sexual attractions, be they directed toward individuals of the same sex (*homosexual*), the other sex (*heterosexual*), both sexes (*bisexual*), or neither (asexual). The terms homosexual and heterosexual can lead to confusion when there are discrepancies between an individual's natal (birth assigned) sex and gender role (as with transsexual individuals, described below). In such situations, sexual orientation may be described less ambiguously as *androphilic* (attracted to men), *gynephilic* (attracted to women), bisexual or *analophilic* (attracted to neither). *Sexual orientation* and *sexual preference* are sometimes used interchangeably; however, the former term has a wider currency in contemporary professional and popular usage. Sexual orientation is the term most experts use, and "sexual preference" is a less precise term as it implies a casual, volitional, or voluntary component. While the decision to engage in sexual behavior is volitional, the

preponderance of evidence suggests that the direction of one's erotic responsiveness is not. In recent years, opponents of gay rights have claimed that paraphilic desires and disorders (see Paraphilias Section 21.3) are also "sexual orientations." However, this is neither a widely held scientific view nor how the term is usually interpreted in either science or in antidiscrimination legislation. Sexual orientation is often described as having three components—desire, behavior, and identity—which may or may not be congruent in an individual; however, while behavior and identity are related to sexual orientation, desire (i.e., one's erotic response tendency) corresponds most closely to the definition of sexual orientation above.

The Kinsey heterosexual/homosexual scale—a 7-point continuous scale, with 0 representing exclusive heterosexuality, 6 representing exclusive homosexuality, and 3 representing equal amounts of both—is the most widely used instrument for describing sexual orientation. The Kinsey scale has been criticized for its unidimensional and bipolar conceptualization of sexual orientation, suggesting a diminished attraction toward one sex may be proportional to an increased attraction to the other sex. Moreover, because fantasies/desires and behaviors are not always congruent, each can be assessed separately. Finally, when compared to desire and behavior which can be described quantitatively, identity is categorical and does not lend itself to quantification with the Kinsey scale (see "Homosexuality and Nongay Identities," below). Finally, the Kinsey scale is a research instrument and has no clinical utility beyond illustrating the finding that sexual desires and behaviors are not necessarily exclusively homo- or heterosexual.

Sexual identities, sometimes referred to as *sexual orientation identities,* are constituted of the subjective experiences and meanings of one's sexual desires or attractions. In the case of *homosexual, lesbian and gay,* or *bisexual* identities, these involve some measure of self-acceptance of one's homosexual desires, and perhaps an identification with a community of similar others by virtue of having same-sex attractions. The acquisition of LGB identities is often conceptualized as a developmental process that occurs over time. Models of LGB identity development usually portray a series of progressive stages involving tasks such as *coming out* (the process of recognizing one's homosexual or bisexual attraction and acknowledging it to either oneself, to others, or both), involvement in sexual minority communities, establishing same-sex romantic relationships, and integration of one's homosexuality and lesbian or gay identity into other aspects of the self.

It is important to note that not all persons who experience homoerotic desire or participate in homosexual behavior follow such a linear path or develop a stable lesbian, gay, or bisexual identity. In addition to orientation, the terms *gay* and *lesbian* may be used to signify identities

based on the recognition of, and identification with, similar others on the basis of sexual orientation. Thus, some individuals may acknowledge their homosexual orientation without identifying as gay, lesbian, or bisexual. In other words, a *sexual identity* or *sexual orientation identity* is not synonymous with a *sexual orientation*. Years may elapse between one's first awareness of same-sex attractions and integration of that awareness into a sexual orientation identity.

For example, an individual with a homosexual orientation may not accept having same-sex feelings or may believe that such feelings are transitory. Such an individual might choose to marry a member of the other sex and to maintain an ostensibly heterosexual lifestyle. Such persons may never lose their homosexual desires, and even if consciously aware of them, still may not identify as gay or lesbian. Such a person may have a history of homosexual encounters and may even continue to engage in homosexual relationships, while at the same time expressing disdain for the "gay lifestyle" and the gay community. Such an individual may seek religious or professional help to change his or her sexual orientation. In recent years, some of these individuals have come to refer to themselves as "ex-gays." Research on whether people can actually change their sexual orientation (see "Sexual Orientation Conversion Efforts," below) is not very persuasive, although in recent years the issue has become one of great cultural controversy in some quarters. It is clear that individuals can change their public identities but less clear that they can change their sexual orientation by volition. Other individuals, some of whom refer to themselves as "nongay homosexuals," may acknowledge their homosexual orientation but resist acting upon their sexual urges and desires.

The clinician must be careful when assessing a patient's sexuality to inquire about all of the components of sexual orientation: desire, behavior, and identity. Again, one cannot assume congruence between all three of these components based solely on information regarding one or two of them. It should further be emphasized that sexual identities are neither diagnostic nor immutable. Although a sexual orientation is highly unlikely to change, sexual orientation identities are more flexible because they reflect how a person deals cognitively and psychologically with the awareness of his or her same-sex attractions.

A *heterosexual identity* refers, in theory, to the subjective experience of those men and women who primarily are aware of attraction to persons of the other sex. Until recently, as a social construct this identity was not as well developed or defined as LGB identities. In today's society, unless one has ongoing contact with nonheterosexual individuals, a heterosexual identity is not consciously experienced as anything other than what is expected or "normal." Where a heterosexual developmental outcome is assumed to be the norm, it is assumed that most people have a heterosexual identity unless they either disavow it or proclaim an

alternative identity. Consequently, there is no parallel in the development of a heterosexual identity to the experience of coming out as gay, lesbian, or bisexual. Moreover, in childhood, most individuals assume that they will become heterosexual adults. Early resistance to adopting a gay or lesbian identity is, therefore, believed to be a common and nonpathological feature of the coming out process.

Homosexuality was first introduced as a medical term in the second half of the 19th century to describe erotic desire for persons of the same sex. *Heterosexuality* was coined later to describe erotic desire for persons of the other sex. A third category of sexual orientation, *bisexuality,* was used to describe an attraction to members of both sexes. The terms homosexuality, heterosexuality, and bisexuality are not applied consistently in theoretical, research, or popular discourse. Depending on the context, these terms have been used to refer to a wide range of differing constructs. These include the categories of sexual desire, gender role attributes, forms of sexual behavior, and personal and social identities. In the 19th century, bisexuality referred to the hypothetical bipotentiality of an organism to physiologically develop as either a male or female of its species. When it was discovered that the embryonic precursor of the human external genitalia did not differentiate as male or female until the 12th week of gestation, it was believed that human beings carried a bisexual potential in them as well. Sigmund Freud, among others, took this paradigm one step further and hypothesized that human beings are psychologically bisexual. Many theorists work from an underlying assumption that sexual orientation reflects some unitary quality or essence, rooted perhaps in biology, that transcends cultures and historical periods. Some academic scholars refer to this belief in this common essence as *essentialism* and those who hold such views as *essentialists*. Essentialist assumptions have sometimes overlooked or dismissed personal and social meanings of sexual desire and relationships that contribute to sexual identities. The contrasting view, that these phenomena are culturally and historically dependent, is referred to as *constructionism*. Constructionist assumptions tend to be dismissive of the role of biology. Although essentialism and constructionism are often viewed as competing schools of thought, that is not necessarily the case. Essentialist theory, in general, deals with sexual orientations, while constructivist theory usually deals with sexual identities. It is certainly possible that individuals throughout history have had sexual orientations as described above, even in cultures where the concept of sexual orientation did not exist and in which sexual orientation had no robust effect on the individual's identity or social role and labeling.

In this section, *homosexuality* refers to an erotic desire for someone of the same gender role and identity and usually of the same natal sex; its presence in an individual does not necessarily imply or dictate the concurrent existence of any other characteristics. When used as an

adjective, *homosexual* is intended to refer to sexual ideation or activity involving members of the same sex or gender identity. Thus, for example, a woman may engage in homosexual behavior, demonstrate typical feminine gender role behaviors, be married to a man, and identity herself as heterosexual; or a man may have homosexual desires and fantasies, have sex only with women, and show no gender role nonconformity.

Bisexuality can also refer to an erotic propensity, an individual identity, or a pattern of sexual behaviors. Bisexuality may occur *sequentially* (expressed as an attraction to or relationship with both sexes, albeit at different times in one's life), or *concurrently* (attraction to or relationships with individuals of both sexes at the same time). Some believe bisexuality to signify a form of identity diffusion or think of it as a transitional state leading toward the development of a more stable heterosexual or gay or lesbian identity. However, some individuals do maintain a stable and persistent bisexual identity. Bisexuals often experience more social stigma than gay people and are often reluctant to declare themselves. Limited research suggests that bisexual women are often perceived by nonbisexual people as "over-sexed" while bisexual men are perceived as being unable to accept their homosexuality. In this section, references to bisexuality and bisexual persons are sometimes included because some parallels exist in the experiences of homosexual and bisexual individuals. A full discussion of the topic of bisexuality, however, is beyond the scope of this section.

In the latter part of the 20th century, the terms *gay, gay man, lesbian,* and *bisexual* emerged in popular usage as a way to refer to men and women whose sexual identity, to some degree, openly recognized their homosexual or bisexual attractions. *Gay* is sometimes also used as a politically inclusive term for both men and women whether homosexual or bisexual. In contrast to the medical term, *homosexual,* which treats one aspect of a person's identity, their sexual attractions, as if it were the sum of the person's entire identity, the terms *lesbian* and *gay* have the advantage of not necessarily connoting sexual acts.

In much of the early psychological and psychoanalytic literature, *homosexual* was used as a noun to designate a person with same-sex desire or behavior; however, the noun is used increasingly less often in scientific and popular literature. For example, recognizing the wide range of sexual identities in those populations at risk of human immunodeficiency virus (HIV) infection, much of the recent literature on HIV and acquired immunodeficiency syndrome (AIDS) refers to *men who have sex with men* (MSM). MSM do not necessarily think of themselves as gay or even as "homosexual" for that matter. In some cultures, only a receptive sexual partner is considered gay or homosexual, while an insertive man who also has sex with women is not culturally defined as either homosexual or bisexual. Thus efforts to categorize people on the basis of particular behaviors often come into conflict with both cultural and subjective

experiences of the meaning of homosexuality.

The following terms do not directly pertain to contemporary theories of homosexuality. However, they are presented here for two reasons. The first is historic; some of these terms were not yet coined in early scientific reports and were therefore conflated with homosexuality. Second, some of these concepts are often blurred in the popular imagination with homosexuality.

Although *sex* and *gender* are often used interchangeably, here *sex* refers to the status of biological variables that can be described as either male-typical or female-typical in normatively developed individuals (e.g., genes, chromosomes, gonads, internal and external genital structures). In most individuals, all of the biological variables of sex are concordant with one another (e.g., all male type or all female type); however, some *disorders of sex development* (DSD) lead to somatic *intersex* conditions in which one or more of the biological variables is discordant with the others or in a form that cannot be classified as either male or female typical. Prior to the mid-1950s the word *gender* had been employed primarily to refer to classes of nouns in languages; however, work with individuals with somatic intersex conditions led John Money and his colleagues to use *gender* in reference to attributes of people; specifically, those attributes that are associated with biological sex but not necessarily determined by it. They observed that most individuals with somatic intersex conditions live and think of themselves as males or females according to how they were assigned in infancy and independent of the status of any particular biological variable of sex. *Gender,* therefore, came to refer to the social categories of males (i.e., boys, men) and females (girls, women) as well as to factors related to living in the social role of a man or a woman. The gender initially assigned to individuals with somatic intersex conditions has been referred to as *natal gender* (often called *natal sex*), a term which has more recently been used to refer to the gender initially assigned to all individuals (particularly in the case of transgender individuals). *Gender identity* came to signify one's persistent inner sense of belonging to either the male or female gender category. *Gender role* was defined as those things that a person says or does to disclose him- or herself as having the status of boy or man, girl or woman, respectively (e.g., general mannerisms, mode of dress, interests, pastimes, deportment, and demeanor; spontaneous topics of talk in unprompted conversation and casual comment). *Gender role* is, therefore, the outward expression of the inner sense of *gender identity*. The terms *experienced gender* and *expressed gender, respectively,* are increasingly replacing gender identity and gender role. Although the initial descriptions of gender role and identity implied two discrete gender categories, gender is now often regarded as more *fluid* reflecting a continuous dimension of masculinity/femininity as opposed to a dichotomy of masculine and feminine. Thus, gender identity/experienced gender may be described as

male (boy/man), female (girl/woman), somewhere in between, or neither, while gender role/expressed gender may be described as masculine, feminine, or mixed. Notions of what is masculine or feminine vary across individuals, cultures, and times.

Many historical discussions of homosexuality erroneously linked a homosexual orientation with an abnormal gender identity. Such theories presumed the cause of homosexuality to be confusion about one's own gender identity, which was then thought to subsequently cause confusion about the sex to which one was attracted. More recently, some biological theoreticians have made the gross distinction of heterosexual versus nonheterosexual without differentiating between homosexuality and transgenderism, viewing them as variations only in the degree of expressions of femininity in males and of masculinity in females. This approach lacks thorough inquiry into the matter, since the two phenomena (sexual orientation and gender identity) are different. *Transgenderism* involves having an experienced gender that is different from one's natal gender and is often associated with a desire to transition to a gender different from the natal gender (i.e., male, female, or other gender category such as *genderqueer* or *eunuch*). On the other hand, homosexuality usually entails no desire to transition gender but merely signifies that one's erotic interests are by and large confined to members of one's own sex and gender role. The term *transsexual* refers to a transgender individual who has, or plans to, employ hormonal or surgical means to modify the body so that it conforms to one's experienced gender. Not all individuals with cross-gender identifications desire, seek, or obtain hormonal or surgical treatments for transition. To find authentic gender expression, some may choose to transition in some respects but not others. That is, transition may involve wearing the clothing or accessories of the nonnatal gender and/or taking hormone supplements to acquire secondary sexual characteristics of the other gender. Not all individuals who transition gender desire gender reassignment surgeries (see Gender Identity Section 21.4). *Transgender* has become an umbrella term and usually refers to someone whose gender identity (or *gender expression*) and natal gender are discordant or not conforming to social norms. Members of the transgender community have coined the term *cisgender* (based on the concept of cis/trans isomerism in organic chemistry) as a way to describe nontransgender individuals whose gender identity and birth assigned sex are in alignment. Some transsexual individuals object to being classified under the transgender umbrella, believing that those who desire to transition hormonally and surgically belong to a category distinct from those who do not.

A natal male who *transitions* to live as a woman is called a *male to female* (MTF) transgender individual, an *MTF transsexual or transwoman*. Conversely, a natal woman who transitions to live as a man is called a *female to male (FTM)* individual, an FTM transsexual or a

transman. Historically, surgical transition was much more common among natal males but transition is increasing among natal females. There is a great deal of popular confusion between transgenderism and homosexuality. Cross-gender identifications, in and of themselves, do not necessarily reveal anything about a person's sexual orientation. For example, a large proportion of MTF individuals are attracted to women (gynephilic).

Paralleling gay activist protests of the mid-20th century, there are now an increasing number of transgender-identified individuals who challenge the characterization of transgenderism and transsexuality as mental disorders. Neither of these terms, however, is used in the current DSM-5. Instead, the presence of a marked incongruence between one's experienced/expressed gender and natal gender are listed among the criteria for the DSM-5 diagnosis of *gender dysphoria (GD) in adolescents and adults*. Importantly, *gender dysphoria,* historically defined as unhappiness or a sense of inappropriateness with one's biological sex, must cause clinically significant distress or impairment in order to satisfy the DSM diagnostic criteria of a disorder. There is also a DSM-5 diagnosis of *gender dysphoria in childhood* (GDC). Longitudinal studies have found that many individuals meeting criteria for GDC do not have GD as adults, but instead identify as homosexual in adolescence or adulthood. However, the reverse is not true and most homosexual adults would not have met criteria for GD as children (see Section 21.4 on Gender Identity, Gender Variance and Gender Dysphoria).

Transvestitism refers to sexual urges and fantasies involving cross-dressing. When the behavior is distressing to a patient, or brought to the attention of a psychiatrist, it may be diagnosed as *transvestic disorder,* which DSM-5 classifies as a *paraphilic disorder* (refer to Paraphilias Section 21.3). Cross-dressing is strongly linked in the popular imagination with homosexuality, however, in DSM-IV-TR, the diagnosis of *transvestic fetishism* could only be made in heterosexual men as most fetishistic cross-dressers are heterosexual men who cross-dress in private. These men are often married to women who may or may not know about their cross-dressing interests. DSM-5 removed the restriction to heterosexual men, opening the diagnosis to women or gay men who cross-dress out of sexual interest. While the change could increase the number of people diagnosed with transvestic disorder, the diagnosis still requires that individuals experience significant distress or impairment because of their behavior. In any event, most gay men and lesbians do not cross-dress. However, there are social venues within gay and lesbian communities that allow for public cross-dressing. These may include either socially performative events (gay pride, Carneval, Halloween, or Mardi Gras parades) or cross-dressing as a form of paid entertainment (drag shows).

THEORIES

Although heterosexuality is often regarded as nature's default setting—and therefore not requiring any explanation—theories of homosexuality's "causes" abound. One exception can be found in Plato's *Symposium,* which tells a creation myth explaining not only male and female homosexuality, but heterosexuality as well. According to this myth, humanity originally consisted of three sexes. Its members were joined in pairs consisting of two men, two women, or a man and a woman (androgyne). To diminish their power and teach them to fear the gods, Zeus divided each pair in half. Thereafter, the split-off aspects were in constant search for their missing halves. Thus men who were originally one-half of the androgyne seek women and vice versa. Women who seek women and men who seek men were instead part of the original male and female beings.

Although male homosexual practices played an important and ritualized role in ancient Babylonia and classical Greece, homosexual acts were prohibited in other cultures. In the Abrahamic religions (Judaism, Christianity, and Islam), prohibitions against homosexuality can be found in religious texts. In the Old and New Testaments of the Bible, these include: Genesis 19, Leviticus 18:7, 22, Leviticus 20:13, Judges 19, I Kings 22:46, II Kings 23:7, Romans 1:27, I Corinthians 6:9, and I Timothy 1:9–10. Historian John Boswell has argued that contemporary intolerance of homosexuality was not an essential factor of Christianity itself, but only became a prevailing church attitude in the second millennium ACE. Although Michel Foucault argued that the idea of identifying *a type or a category of person* based on either homosexual behavior or identity is a relatively new cultural phenomenon that coincides with the medicalization of homosexuality during the 19th century, Boswell, believed that a category of homosexually identified person, although not entirely equivalent to a modern gay identity, has always existed across time and cultures. Plato's androgyne myth, for example, accounts for the origins of such categories. Foucault and Boswell may each have been right in that recognition of individuals defined by their sexual orientation has been discontinuously present across history, and that at any given point in history it is recognized by some cultures and religions but not by others. Today, the study of the forms and meanings of homosexual behavior in different cultures and historical periods is a rich field done by scholars of *lesbian and gay studies* and *queer theory*. Nevertheless, the forms of socialization and identity development associated with homosexuality in contemporary Western society are generally considered to be unique. The history of the development of present forms of homosexual identity is the focus of the following review of theoretical perspectives about homosexuality.

History

The contemporary debates about the meaning of homosexuality can be traced to the middle of the 19th century when European sexologists, including several prominent psychiatrists, began the study of homosexuality from a scientific and medical perspective. Through much of the 20th century, this intellectual enterprise was dominated by a predominance of biological and psychological, particularly psychoanalytic, theories of homosexuality. Psychiatric theories of the first half of the 20th century, with a few exceptions, regarded homosexuality as either a form of psychopathology or as a developmental arrest. In the latter half of the 20th century, however, theories that characterized homosexuality as a normal variant of human sexuality began to predominate. This view emerged from a wide range of social and scientific perspectives outside medicine and psychiatry and eventually supplanted earlier theories of homosexual pathology and immaturity. This modern *normal variant* view began gaining influence around the time of the publication of the two Kinsey volumes on male and female sexuality—in 1948 and 1953, respectively—and culminated in the APA's removal of homosexuality from its list of mental disorders in the DSM in 1973. This action later prompted the WHO to remove homosexuality from the International Classification of Diseases (ICD-10) in 1990. The normal variant model is accepted by today's mental health mainstream.

Medicalization of Homosexuality: From Morality to Pathology

19TH AND EARLY 20TH CENTURY BIOLOGICAL VIEWS. The modern history of homosexuality is often thought to begin in the mid-19th century. In 1869, in response to Bavaria's Paragraph 175, which criminalized male (but not female) homosexual acts, German lawyer and activist Karl Ulrichs began publishing treatises that criticized the law. Ulrichs theorized that some women were born with a man's spirit trapped in their bodies (*Urningins*) and some men were born with a woman's spirit trapped in theirs (*Urnings*). Although he said these people constituted a "third sex," the concept drew on a binary approach to gender that presumed an attraction to men is a feminine attribute and that an attraction to women is a masculine one.

In 1869, a Hungarian writer, Károli Mária Kertbeny, who also wrote critically of laws criminalizing same-sex behavior (*sodomy laws*), first coined the terms *homosexual* and *homosexuality*. The influential German psychiatrist, Richard von Krafft-Ebing, later adopted and popularized the terms in his 1886 classic *Psychopathia Sexualis*.

Krafft-Ebing and the psychiatrist Carl Westphal adopted many of Ulrich's ideas; however, the former two deemed unnatural (*theories of pathology*) the very same traits that the latter had described as normal. Magnus Hirschfeld, a German physician and sexual emancipator, was an

openly homosexual psychiatrist who also believed there was a biological basis for homosexuality. Although he considered homosexuality to be a malformation in development, presumed to have a number of different biological etiologies, Hirschfeld nevertheless asserted that the condition was natural (*theory of normal variation*). Similarly, the English sex reformer Havelock Ellis thought of homosexuality as a biological variation. And although he attributed it to "imperfect sexual differentiation," he nevertheless considered homosexuality to be nonpathological since it did not cause any illness in the affected individual. Other biological bases for homosexuality began to be examined during this period. These included studies of anatomical structures in the brains and nerve structures in the rectums of homosexual men. Endocrinological hypotheses led to experiments in which the supposedly atrophied or abnormal testicles of homosexual men were removed and replaced with what were believed to be the normal testicles of heterosexual men. However, these experiments did not make the homosexual men heterosexual.

Early biological theories, with their emphasis on heredity, endocrinological influences, and anatomical or structural differences between homosexual and heterosexual persons, heralded many of the current efforts to find a biological basis for homosexuality. Biological theories, however, also provided a rationalization for subsequent eugenic efforts in Germany to eradicate homosexual persons. In the 1970s, East German surgeons unsuccessfully tried to eliminate homosexuality in men by operating on the hypothalamus where the "cause" of homosexuality was presumed to lie. Throughout the 20th century, interest in the origins of sexual orientation focused almost exclusively on men, a trend less evident today.

PSYCHOANALYTIC APPROACHES. Freud did not view homosexuality as a sign of illness, by which he meant a symptom arising from psychic conflict. Instead, he saw homosexuality as the unconflicted expression of an innate instinct. He noted that homosexuality could occur in individuals who had no other signs of deviation and no impairment in their functioning. Freud believed in *constitutional bisexuality,* that every individual had a certain component of *masculine* (active) as well as *feminine* (passive) *tendencies*. A homosexual tendency did not necessarily imply a homosexual behavior. Although bisexual tendencies were universal, Freud believed some people were constitutionally endowed with more of one tendency than the other. He believed life experiences, particularly traumatic ones (environmental factors), could have an impact on the development and expression of one's innate instincts (biological factors). Under normal and nontraumatic circumstances, the component instincts that determine the sex of one's final object choice should be consistent with one's anatomical sex. That is to say an anatomic male should ideally express the masculine component

instinct and obtain sexual satisfaction from women. However, Freud also believed that even adult heterosexuals retain the homosexual component, albeit in sublimated form. For Freud, adult homosexuality represented an arrest in development on the road from the instinctual bisexuality of childhood to mature heterosexuality. His *theory of immaturity* comprised an alternative category to religious sin, medical theories of pathology, or normal variant third-sex theories. By maintaining that homosexuality could be a normal part of everyone's experience, Freud allowed for the possibility that the adult homosexual person might sufficiently mature and, if sufficiently motivated, become heterosexual. However, late in his life, Freud doubted the possibility of changing adult homosexuality to heterosexuality in most people.

After Freud's death, Sandor Rado's theories came to hold greater sway on psychoanalytic theorizing about homosexuality than those of Freud himself. In the 1940s, Rado argued that the theory of bisexuality was based on a faulty analogy with anatomical bisexuality. That is, underlying Freud's theory was a 19th century belief in embryonic hermaphroditism, the hypothesis that the potential to become an anatomical man or a woman was present in every embryo. Rado claimed this theory had been disproved, which led to his claim that heterosexuality is the only nonpathological outcome of human sexual development. He viewed homosexuality as an illness, specifically a phobic avoidance of the other sex caused by parental prohibitions against childhood sexuality. Almost all of the mid-20th century psychoanalytic theorists who pathologized homosexuality followed Rado's theory in one form or another. The psychoanalytic community's shift from Freud's model of immaturity (homosexuality as a normal developmental step toward adult heterosexuality) to Rado's model of pathology (homosexuality as a sign of development gone awry) led some analysts to claim they could "cure" homosexuality.

The work of psychoanalyst Irving Bieber and his associates was particularly influential in its portrayal of a pathogenic family type—a detached and rejecting father and a close-binding and domineering mother—that presumably led to homosexuality in the 106 adult homosexual men they studied. This theory of familial etiology, however, has not found support in the research of larger, nonpatient gay male populations. Although about 70 percent of their subjects dropped out during treatment, Bieber and his colleagues claimed a 27 percent cure rate among 29 of the 32 men who remained in the study. The Bieber study was criticized for its methodology—it was not peer reviewed—and the authors were unable to provide any long-term follow-up of their subjects or produce any patients to support their claims of change. Finally, while some discussion of the etiology of female homosexuality existed in the early psychoanalytic literature, the primary emphasis in psychoanalysis, as in the biological sciences, was on male homosexuality; often presentations of homosexuality

in women were simply treated as gender-reversed images of those for male homosexuality.

EARLY THEORETICAL ASSUMPTIONS. Nineteenth and early 20th century theories of homosexuality, be they biological, medical, or psychoanalytic, largely relied on a dichotomous model of categories: Male/female, masculine/feminine, and heterosexual/homosexual. Whether the theorist believed homosexuality to be a normal variant, a form of pathology, or immaturity, the theory usually relied on the assumption that the wide range of human sexuality could be reduced to the two component parts of male and female and that some intrinsic quality of one gender had made its way into a person of another gender. In addition, scientific and medical studies served to remove the responsibility of defining homosexuality from the realms of morality and religion and secured that task within science and medicine. In doing so, science and medicine created a category of person, "the homosexual," in contrast to earlier moral and religious beliefs that homosexuality was a sinful behavior rather than a source of identity. Although science and medicine perpetuated the social stigma of homosexuality, taking it out of the realm of sin and immorality into the realm of pathology and immaturity, these developments would also eventually set the stage for the transition toward normalizing homosexuality that began to take shape in the middle of the 20th century.

Normalization of Homosexuality: From Diagnosis to Normality

INFLUENCE OF RESEARCH FINDINGS FROM SEXOLOGY, CROSS-CULTURAL, ANIMAL, AND PSYCHOLOGICAL STUDIES. A major factor in the shift away from pathologizing theories about homosexuality was the work of sex researchers. Unlike psychiatrists and psychoanalysts, who based their theories on clinical work with people seeking treatment in their offices or on evaluations of prison populations, sex researchers sampled larger, nonpatient populations. Thus, the publication of two studies on male and female sexuality by Alfred Kinsey and his colleagues, in 1948 and 1953, respectively, proved to be a scientific and cultural watershed. Although normal variant theories of homosexuality had been around for almost a century, at the time of its publication, Kinsey's work signaled the beginning of a cultural shift away from viewing homosexuality as pathology toward viewing it as a normal variant of sexual desire and behavior. Despite his studies' methodological flaws, Kinsey's finding of a widespread presence of same-sex feelings and behaviors among his samples of several thousand American men and women changed the general impression that homosexuality was an isolated and pathological phenomenon. His homosexuality/heterosexuality scale, which conceptualized sexual orientation on a continuum, expanded on the binary model used by earlier theorists. As Kinsey put it in his 1948 report, "The world is not to be divided into sheep and goats. Not all things are

black nor all things white. It is a fundamental of taxonomy that nature rarely deals with discrete categories. Only the human mind invents categories and tries to force facts into separated pigeon-holes. The living world is a continuum in each and every one of its aspects."

Cleland Ford and Frank Beach's work, published in 1951, found homosexual behavior to be common across cultures and documented its existence in almost all nonhuman primate species. The interpretation of these findings supported the notion that homosexuality was both natural (as opposed to a traditional view that it was caused by civilized decadence) and widespread. In 1957, psychologist Evelyn Hooker published a ground-breaking study comparing projective test results from 30 nonpatient homosexual men with those of 30 nonpatient heterosexual men. Her study found that experienced judges, unaware of whose test results they were interpreting, could not distinguish between the two groups. Again, the medical and psychiatric literature of that time, which regarded homosexuality as pathological, was mostly based on evaluations of individuals troubled about their sexuality. Hooker's study of *nonpatient subjects* was a serious challenge to the then-prevailing psychiatric and psychoanalytic belief that homosexuality was always associated with mental illness or serious psychopathology. The impact of this significant finding was to draw attention to the a priori assumption of many clinical papers that homosexuality was itself pathological. In the 1960s, psychiatrists like Judd Marmor—a psychoanalyst and future president of the APA—Robert Stoller, and Richard Green gave serious consideration to the nonpsychiatric perspectives offered in these newer research findings. Another psychiatric dissenter was Thomas Szasz, who criticized mental health professionals in general for labeling many unconventional behaviors as signs of illness. He argued that mental illness was a metaphor, rather than an actual illness like a viral infection or a broken leg. Rather than practicing medicine, Szasz accused his fellow psychiatrists of using diagnoses as a way to increase their own power and influence.

DECLASSIFICATION BY THE AMERICAN PSYCHIATRIC ASSOCIATION . Homosexuality had been officially classified as a mental disorder in APA's first edition of the DSM (DSM-I) in 1952. There it was designated as a type of "sociopathic personality disturbance." Classifying homosexuality as a mental illness was not controversial at the time of the DSM-I's publication; societal attitudes of that time coincided with prevailing medical views that homosexuality was an aberration. In 1968, the revised DSM-II still listed homosexuality as a "sexual deviation," but sexual deviations were no longer categorized as a sociopathic personality disturbance.

The DSM's 1968 revision (DSM-II) coincided with a time of great social foment in the United States. In addition to anti–Vietnam War activity, civil rights protests, and emerging feminist voices, gay rights activism also

gained public notice. Gay activism coalesced around the 1969 Stonewall riots in New York City. Efforts by police to raid and shut down a gay bar (The Stonewall), a routine police activity of the time, led to 3 days of rioting in Greenwich Village. Having successfully fought off police and government, the gay rights movement was prepared to challenge psychiatric authority as well. This led to an increasingly vocal opposition to modern psychiatry's classification of homosexuality as a mental disorder. Before that time, many homophile organizations, both in the United States and abroad, had accepted the medical view of homosexuality as a mental disorder. Accepting the view of homosexuality as disease meant treating it as a disability, a category intended to replace moral and religious judgments with scientific, objective, and humane attitudes. However, in most instances, medical and psychiatric portrayals of homosexuality were as problematic as the religious views they supplanted. For example, a homosexual individual could be denied the right to immigrate to the United States on the grounds that the person suffered from a mental disorder. Gay men and lesbians were discharged from the military as medically unfit. To gay activists of that period, the psychiatric designation of homosexuality as a mental disorder had exacerbated antihomosexual societal prejudices rather than ameliorated them. These activists found a receptive audience among an increasing number of psychiatrists familiar with the nonpsychiatric research findings showing homosexuality occurred in large numbers of people, in persons who demonstrated normal psychological adjustment, and across cultures.

Gay activists chose to confront APA about its position on homosexuality —they actually disrupted APA's annual meetings in 1970 and 1971. In response to these protests, APA's leadership took several steps. First, they permitted two educational panels at the APA's 1971 and 1972 meetings. The 1971 panel, titled "Lifestyles of Non-Patient Homosexuals," featured nonpatient gay activists explaining to their psychiatric audience the stigma caused by the "homosexuality" diagnosis. At the 1972 meeting, the activists were on another panel, "Psychiatry: Friend or Foe to Homosexuals: A Dialogue." The latter panel presented the perspectives of the activists, Judd Marmor, and psychiatrist, John Fryer, M.D., who appeared as "Dr. H Anonymous," disguising his true identity from the audience with a rubber mask, a fright wig, and an oversized tuxedo. He spoke of the discrimination gay psychiatrists faced in their own profession, telling his psychiatric audience:

> As psychiatrists who are homosexual, we must know our place and what we must do to be successful. If our goal is academic appointment, a level of earning capacity equal to our fellows, or admission to a psychoanalytic institute, we must make certain that no one in a position of power is aware of our sexual orientation or gender identity. Much like the black man with the light skin who chooses to live as a white man, we cannot be seen with our real friends—our real homosexual family—lest our secret be known and our dooms sealed.

The second step taken by APA in addition to these public panels was to embark upon an internal process of studying the scientific question of whether homosexuality should be considered a psychiatric disorder. APA's Nomenclature Committee was the scientific body charged with addressing this issue. The Committee members reviewed the psychiatric and psychoanalytic literature on the subject as well as the sexology literature. The latter, a subject not usually taught in psychiatric training programs at that time, was unfamiliar to most practicing psychiatrists. Following their study of the issue, which included interviews of psychiatrists who favored and objected to making any changes, the Nomenclature Committee recommended removing homosexuality per se from the DSM. The proposal was approved by the APA's Council on Research and Development, its Reference Committee, and by the Assembly of District Branches (now known as the APA Assembly) before being accepted by APA's Board of Trustees in December 1973. This marked the end of the classification of homosexuality per se as an illness. Other major mental health professional organizations, including the American Psychological Association and the National Association of Social Workers, soon endorsed the APA action. The decision to declassify homosexuality was accompanied by passage and issuing of a position statement by the APA that supported the protection of the civil rights of homosexual persons and that read:

Whereas homosexuality in and of itself implies no impairment in judgment, stability, reliability, or vocational capabilities, therefore, be it resolved that the American Psychiatric Association deplores all public and private discrimination against homosexuals in such areas as employment, housing, public accommodations, and licensing, and declares that no burden of proof of such judgment, capacity, or reliability shall be placed on homosexuals greater than that imposed on any other persons. Further, the APA supports and urges the enactment of civil rights legislation at local, state, and federal levels that would insure homosexual citizens the same protections now guaranteed to others. Further, the APA supports and urges the repeal of all legislation making criminal offenses of sexual acts performed by consenting adults in private.

Some psychiatrists, primarily psychoanalysts, who continued to adhere to pathologizing views of homosexuality, criticized the Board of Trustees and challenged their action by petitioning the organization to hold a referendum in which the entire APA membership could vote. The Board of Trustee's decision to remove homosexuality from the DSM was upheld by a 58 percent majority of voting APA members.

The original scientific and clinical rationale for removing homosexuality from the DSM was based, in part, on a new formulation of the definition of a mental disorder by the Nomenclature Committee's Chair, Robert Spitzer, M.D. In order to be classified as a disorder, a condition had to meet the criteria of either an experience of subjective distress or of a generalized,

objective impairment in social effectiveness or functioning resulting from the condition itself. Neither of these criteria applied to gay people who were satisfied with their sexual orientation and who did not demonstrate impaired functioning. In recognition of the opposition to the diagnostic change, however, the APA did not fully embrace a normal variant model of homosexuality; instead, it made a compromise. The diagnosis of *homosexuality* in the original text of DSM-II was replaced by *sexual orientation disturbance* (SOD) in its seventh printing in 1974. As a result, individuals comfortable with their homosexuality were no longer considered mentally ill. According to SOD criteria, only those who were "bothered by," "in conflict with," or "wished to change" their homosexuality had a mental disorder.

The diagnostic category of SOD had two significant conceptual problems. First, the diagnosis could also apply to heterosexuals, a solution to APA's internal debate that did not concur with clinical reality. There were no reported cases of unhappy heterosexual individuals seeking psychiatric treatment to become homosexual. This problem of diagnostic overinclusiveness was resolved in 1980s DSM-III, where SOD was replaced by *ego-dystonic homosexuality* (EDH). The name change, however, did not resolve a thornier conceptual issue, which was that of making patients' subjective experiences of their own homosexuality *the determining factor* in labeling their homosexual feelings as an illness. To rely on patient subjectivity as the major determining factor was now incongruous with the new evidence-based approach that psychiatry had embraced with DSM-III. This ultimately led, in 1987, to EDH being removed from the revised DSM-III-R. In 1990, the WHO followed suit and removed homosexuality from the tenth revision of the ICD (ICD-10) although it replaced with a diagnosis that resembled SOD: ego-dystonic sexual orientation.

For two decades after the 1973 APA decision, the meetings and publications of the APsaA were among the few American mental health forums where pathological views of homosexuality continued to be expressed. Eventually, however, normal variant views of homosexuality prevailed in the APsaA as well. In 1991, the organization issued a nondiscrimination statement to allow training of gay and lesbian psychoanalytic candidates and the promotion of qualified gay and lesbian individuals to the position of training and supervising analyst. In 1997, the APsaA became the first major mental health organization to support marriage equality. In 2000, the APsaA took a position opposing sexual conversion therapy. However, psychoanalytic views that pathologized homosexuality nevertheless could still be found in the international psychoanalytic community. Then, in 2002, APsaA influence in the International Psychoanalytical Association (IPA) led the latter to issue a nondiscrimination statement allowing for the selection of openly gay and lesbian candidates and the promotion of gay and lesbian faculty.

At the time of this writing, the WHO is revising the ICD with an anticipated publication date of 2018. The WHO Working Group on Sexual Health and Sexual Disorders has recommended removing the F66 diagnoses, including ego-dystonic sexual orientation. The Working Group, reviewing the evidence published since 1990, found no evidence that these diagnoses are clinically useful as they do not contribute to health service delivery or treatment selection nor provide essential information for public health surveillance.

All of these actions represent definitive acknowledgment by the psychiatric, psychoanalytic, medical, and mental health professions that homosexuality, in itself, is not a mental disorder. Not only does this change in the mental health professions' view of homosexuality represent a wider acceptance of a newer scientific view of homosexuality, it also symbolizes a dramatic conceptual shift in the cultural meaning and significance of homosexual behavior within Western society.

FREQUENCY

One must use caution when interpreting studies that address the frequency of homosexuality in a population. First, one must consider the survey methods: What population did the survey aim to assess? Was a random sample of this population ascertained, and was the sample large enough to provide statistically meaningful results? One must clearly identify which aspects of homosexuality were assessed: Same-sex desires, same-sex behaviors, or homosexual identity. Additionally, one must determine whether the numbers reported reflect incidence or prevalence. Do they reflect the proportion of respondents who experienced same-sex desire, who engaged in same-sex behavior, or who identified themselves as homosexual at any time in their lives, within a restricted period of time, at a particular stage of their lives, or in a stable and ongoing pattern?

Beyond the issues of survey and instrument design, the stigma associated with homosexuality may interfere with full disclosure. The impact of stigma on survey results may vary depending on the survey method. Although face-to-face interviews are generally considered preferable, logistical issues may limit the sample size or the geographic region(s) surveyed. Surveys based on random telephone numbers may achieve large representative nationwide samples but may result in underreporting of homosexuality if the subjects do not have privacy at the time of their telephone interview. Moreover, differences in survey methods are sometimes associated with different results that are difficult to explain. For example, a 1998 random survey of 1,672 males found a fourfold increase in the proportion of subjects who reported having ever engaged in homosexual activity when questions were heard over headphones and then answered on a computer as opposed to when subjects used a conventional paper and pencil instrument.

Regardless of the survey method employed, persons who are not accepting of or open about their sexual orientation are less likely to respond fully or accurately to questions about their sexual behavior or identity. Thus, even large-scale random surveys are more likely to result in underestimates rather than overestimates of homosexual phenomena.

Finally, there is the issue of surveyors' unwillingness to even ask specific questions. The 2000 US census, for example, did not ask respondents any questions about their sexual orientation. However, for the first time, the census included a category of "unmarried partner" and found 358,390 same-sex couples who identified themselves as heads of households. As this is the first time the census used the "unmarried partner" designation, there are no previous figures to which these numbers can be compared. The 2010 US census, when marriage equality was only available in five US states and the District of Columbia, found a total of 646,464 gay couples. This included 131,729 same-sex couples who reported being married by checking "husband" or "wife" boxes on their census forms, the first time they could do so, after gay marriage became legal in Massachusetts in 2004.

Given these limitations in existing studies of homosexuality's frequency, it is impossible to determine definitively the number of persons who have (or who have experienced at some time in their lives) same-sex desire, behavior, or identity in the United States today. The 1948 Kinsey study of male sexuality, based on a sample of men between the ages of 16 and 55 years, found 4 percent of the men were exclusively homosexual throughout their lives after adolescence, 50 percent were exclusively heterosexual, and 46 percent were somewhere in between. The report also concluded 10 percent of males are "more or less exclusively homosexual for at least 3 years between the ages of 16 and 55." That conclusion is apparently the basis of the common misperception that 10 percent of the population is homosexual.

The Kinsey study has subsequently been criticized for its sampling errors, which are believed to have led to inflated estimates of the extent of homosexual and bisexual behavior; however, Paul Gebhard, Kinsey's successor as director of the Kinsey Institute for Sex Research, dedicated years to culling purported contaminants from Kinsey's data and concluded that his figures were off by no more than a few percentage points.

Later studies of the frequency of same-sex sexual contacts between men, utilizing results from national surveys conducted in 1970 and 1988, determined that 20 percent of men had sexual contact to orgasm with another man at some time in their lives. It was also found that 5 to 7 percent of men had such behavior in adulthood, but that only one-quarter to one-half of these men also reported having such contact in the 12 months preceding the survey. Because of problems in reporting and with missing data, these figures were considered to be the lower bounds of actual

frequency of same-sex behavior among men. Relatively few studies have addressed sexual orientation in women; however, those that have found the proportion of women reporting same-sex sexual contact or identifying themselves as homosexual is lower than for men, often one-half or less the proportion of men.

A large-scale study of sexuality in the United States published in 1994 by Edward Laumann and his colleagues is generally regarded as providing the most reliable recent estimates of the prevalence of homosexuality in the United States. This study inquired about individual desire, behavior, and identity, and obtained a carefully stratified representative sampling of the general population. This study, based on a total sample of 3,159 men and 1,749 women, found that 2.7 percent of the men and 1.3 percent of the women had participated in same-sex sexual behavior during the preceding year, and 4.9 percent of the men and 4.1 percent of the women had done so since age 18; 7.7 percent of the men and 7.5 percent of the women reported having experienced sexual desire for someone of the same sex; and 2.8 percent of the men and 1.4 percent of the women reported a homosexual or bisexual identity. These figures varied considerably across groups based on age, marital status, education, religion, race, and place of residence. Surveys done over the past 20 years in Canada, Australia, and Europe have produced similar results.

In 2014, the US Center for Diseases Control and Prevention (CDC) released the results of the 2013 National Health Interview Survey. The survey, which asked questions about sexual orientation/identity for the first time, showed that 1.6 percent of adults self-identify as gay or lesbian, and 0.7 percent consider themselves bisexual. The majority (96.6 percent) labeled themselves as straight. An additional 1.1 percent declined to answer, or responded "I don't know the answer" or said they were "something else."

These later studies clearly suggest that the figure of 10 percent (an estimate based on a misinterpretation of Kinsey's findings cited above), popularly stated as the proportion of people who are gay, is inaccurate. In the 12 largest US cities, however, the number of gay and lesbian persons does approach the 10 percent figure, perhaps reflecting their migration to areas with greater tolerance for sexual diversity or sizable sexual minority communities.

Historically, there has been a political dimension surrounding the issue of homosexuality's prevalence in the population. Proponents and opponents of gay rights have attempted to respectively exaggerate or diminish the importance of the gay and lesbian presence in the population at large. Although the exact numbers and proportions remain somewhat elusive, it is nevertheless clear that a larger number of persons report same-sex desire than report the associated behavior, and more report engaging in homosexual sex than report a homosexual or bisexual identity.

For example, a 2006 study of men in New York City found 12.4 percent reported having had sex only with men during the previous year; however, the majority (8.9 percent) identified as "straight," with only 3.3 percent identifying as gay and 0.2 percent identifying as bisexual. For the clinician, it is important to understand that prevalence is dependent on which component of homosexuality and which social or demographic groups are being considered. The validity of estimates of the extent of homosexuality in a given population is likely to be influenced by factors such as the degree of social repression of homosexuality and the cultural meaning of same-sex desire, behavior, and identity.

ORIGINS OF HOMOSEXUALITY

Historically, most research has been done from a heteronormative perspective that treats heterosexuality as normal and therefore not in need of explanation, whereas, homosexuality has been viewed as an abnormality to be explained. Consequently, much research has focused on homosexuality in particular rather than on sexual orientation in general, and often with the aim of elucidating the origins of homosexuality to inform efforts aimed at prevention or "cure." Four areas in which research on the origins of homosexuality have been focused are biology, cross-cultural studies, psychological development, and the acquisition of gay and lesbian identities.

Biological Studies

Since the emergence of the concept of sexual orientation in Western medical discourse, there has been a debate as to whether it is determined, or primarily influenced, by biological as opposed to psychosocial factors. In recent years most scholars have suggested that it emerges from an interaction between both sets of factors. Today, the role of biology often enters into the sociopolitical debates, sometimes referred to as culture wars, regarding whether or not homosexuality is mutable (perhaps even a chosen lifestyle), and whether gay people should be encouraged to change or granted civil protections against discrimination. Weaknesses in the evidence base for a biological contribution to orientation are sometimes employed to support arguments for conversion therapies or discriminatory policies. As shown below, however, even behaviors acquired though experience can be immutable. Furthermore, biological evidence is not needed to demonstrate that homosexuality is impervious to change. That is clearly evident in the psychological literature reporting the ineffectiveness of therapies designed to change it. While sexual orientation may shift in some individuals, few if any of even the most highly motivated have changed their sexual orientation despite subjecting themselves to painful aversion therapies or investing emotional and financial resources in years

of psychoanalysis or reparative therapies. Rather than issues of choice, biological determinism and immutability, the arguments for equal rights and civil protections for homosexuals, like those for religious and racial minorities, are better framed in terms of justice, privacy, and liberty.

Before reviewing the biological research into sexual orientation, this section begins by describing three basic models for conceptualizing the possible contributions of biological factors and their interactions with psychological and social factors.

Biological Models to Explain Homosexuality

MODEL OF PERMISSIVE BIOLOGICAL EFFECTS. In this model biology plays a permissive role by providing the neural machinery through which sexual orientation is inscribed by formative experience. A permissive role could also include delimiting the developmental stage during which the relevant formative experiences must occur. For example, some songbirds must learn their species' song by hearing it during a restricted period of early development. While the song is clearly acquired through experience, biology determines the sensitive period during which that experience must occur. Once a particular song has been acquired, that will be the bird's song for life; it will not be able to unlearn that song or to acquire another. This is not meant to imply that sexual orientation might be acquired by simple mimicry. Instead it shows that phenomena acquired through experience may, nonetheless, be immutable and, therefore, militates against the widespread notion that sexual orientation must be innate since it is impervious to change.

MODEL OF DIRECT BIOLOGICAL EFFECT. In this model, the underlying premise for most biological research regarding sexual orientation, the actions of biological factors such as genes or hormones would directly influence the organization or activity of brain circuits that mediate sexual orientation. The brain is viewed as possessing an intrinsic (i.e., constitutional) sexual orientation or predisposition, the expression of which could nonetheless be subsequently modified by experience.

MODEL OF INDIRECT BIOLOGICAL EFFECTS. This model addresses the contribution of individual differences, some of which may be constitutionally influenced. In this model, biological factors influence sexual orientation only indirectly as the more direct or immediate action of biological factors is upon temperament or other personality traits. From birth, these personality or temperamental traits influence how an individual experiences, interacts with, and modifies the environment. This would include how an individual participates in shaping the relationships and experiences thought by some to influence the development of a sexual orientation. Although similar to the model of permissive biological effects, the model of indirect biological

effects goes further and suggests that the relevant formative experiences may, themselves, be strongly affected by biologically influenced personality variables.

The existing biological data relevant to the underpinnings of sexual orientation are equally compatible with both the direct and indirect models. The distinction between these models can be appreciated in their differing interpretations of three of the more robust findings in the sexual orientation literature as reviewed by William Byne and Bruce Parsons. The first of these findings is that the propensity to engage in rough-and-tumble play during the juvenile period appears to be influenced by prenatal exposure to androgenic (i.e., "male") hormones. Second, compared to heterosexual men, more, but not all, homosexual men recall less childhood interest in competitive rough-and-tumble play. Third, compared to heterosexual men, homosexual men are more likely to recall their fathers as having been distant or rejecting.

In the direct model interpretation, the aversion to rough-and-tumble play represents the childhood expression of a brain that has been prewired for homosexuality. This was the position of Richard Isay, a psychoanalyst who suggested that biological factors wire the brain for sexual orientation and consequently reverse the polarity of the Freudian Oedipus complex. According to this model, in addition to shunning rough-and-tumble play, prehomosexual boys are erotically interested in their fathers during the oedipal period. Fathers might recoil from their prehomosexual sons' gender nonconformity or sexual inclinations. Even if the father did not recoil, Isay speculated that in adulthood gay men might nevertheless recall their fathers as having been cold or distant in order to avoid conscious awareness of their earlier sexual attractions to them.

Alternatively, an indirect model interpretation postulates that a biologically influenced aversion to rough-and-tumble play does not necessarily imply prewiring for homosexuality. Instead, such an aversion becomes a potent factor predisposing to homosexual development in particular environments—perhaps where this behavior is stigmatized as "sissy" behavior and, therefore, causes the boy to see himself as different from his father and male peers. In this scenario, the father's withdrawal from his son would contribute to, rather than result from, his homosexuality. Significantly, an aversion to rough-and-tumble play would arguably have different consequences in environments where this is acceptable, perhaps making no contribution to sexual orientation at all. This example does not suggest that a distant or rejecting father (whether real or projected) is a feature of all pathways to male homosexuality. Based on the indirect model, one might conjecture how any number of biologically influenced personality variants could influence sexual orientation. Theoretically, a given variant might predispose to homosexuality in one environment and to heterosexuality in another, while

making no contribution to sexual orientation in environments in which the variant is perceived to have no valence with respect to perceptions of gender appropriate behavior.

NEUROENDOCRINE RESEARCH. Historically, much biological research has begun with the premise that the constitution of gay men and lesbians reflects some intermediary state between that of their male and female heterosexual counterparts. This assumption led to investigations of various attributes for evidence of sexual atypicality in gays and lesbians. From the 1950s into the 1970s considerable research examined the endocrine system for atypical levels of the so-called sex hormones in gays and lesbians. An overwhelming majority of such studies failed to demonstrate any correlation between sexual orientation and adult hormone levels. In fact, androgens, frequently referred to as "male hormones," although they are present in both sexes, have been found to increase sexual desire in both men and women and yet have no effect on sexual orientation. Moreover, sexual orientation does not change in adults when hormone levels are altered by gonadal malignancies, trauma, or surgical removal. Currently, however, there is much speculation about a potential role of prenatal exposure to androgens and other hormones derived from them.

THE PRENATAL HORMONAL HYPOTHESIS. This hypothesis posits that (1) the brains of heterosexual men and women differ from each other both structurally and functionally, (2) these sex differences in the brain result, at least in part, from prenatal influences on brain development, and (3) some of these hormonally influenced differences determine whether one is sexually attracted to men or to women. This perspective, therefore, views sexual orientation as derivative of a hormonally mediated developmental process leading to sexual differentiation of the brain. Consequently, the brains of homosexual individuals are expected to exhibit neurobehavioral characteristics considered to be more typical of the other sex. This supposition is sometimes referred to as the *intersex hypothesis* of homosexuality.

The prenatal hormonal hypothesis draws extensively on numerous studies and observations of rodents, in which the balance between male and female patterns of mating behaviors is strongly influenced by the amount and timing of early androgen exposure. The period of maximal sensitivity to these organizing effects of androgen varies in different species. The rat has been employed extensively in such research because its period of brain sexual differentiation extends into the early postnatal period (which corresponds to the mid-trimester of human gestation). Thus, hormonal exposure of the rat's brain can be experimentally manipulated by perinatal gonadectomies and injections of various hormones as well as their receptor agonists and antagonists. Such research has revealed that

different aspects of sexual differentiation of the rat brain are mediated by different pathways of androgen action in the brain. Adding to this complexity, the extent to which sexual differentiation of the brain relies on particular androgens and their metabolites appears to vary across species. Thus findings cannot be automatically or easily extrapolated from one animal model to the next, let alone from animal models to humans. This section first reviews what is known about sexual differentiation of the brain in rats. Subsequent sections highlight particular aspects of the rat model that do not apply to primates, including humans. The reader is advised to always be cautious when reading materials about human sexuality that discuss sexual differentiation of the brain generically without identifying the species in which particular research findings were derived.

Normal development of the male rat brain requires completion of two processes: (1) masculinization—induction of the male phenotype, and (2) defeminization—removal of the female phenotype. Both processes, although distinct, are largely directed by the same steroid, estradiol, which is derived from testosterone by the action of aromatase enzymes in the brain. Subsequently, the brain-derived estradiol interacts with estrogen receptors. Various hypotheses have been advanced to explain why the brains of female fetuses are not masculinized and defeminized by estrogens of maternal origin or those that originate from their own ovaries. A possible explanation is that peripheral estrogens are bound by proteins that hinder their crossing of the blood–brain barrier.

The binding of estrogens to their brain receptors activates a variety of downstream cascades, one of which involves the upregulation of prostaglandin E2. This pathway mediates the masculinization of sexual behavior in rats but does not affect defeminization of sexual or maternal behavior. Thus, the intracerebral administration of prostaglandin E2 to female rat pups produces animals that are behaviorally masculinized but not defeminized. That is, they exhibit male-typical mounting behavior in addition to the full range of female sexual and maternal behaviors. Conversely, administration of indomethacin, which blocks the synthesis of prostaglandin E2, to male pups produces animals whose behavior is defeminized but not masculinized. Although estradiol simultaneously defeminizes and masculinizes behavior, those actions are mediated by divergent mechanisms downstream of the estrogen receptor.

Problems inevitably arise when one attempts to draw conclusions about human sexual psychology based on extrapolations from rodent behaviors caused by experimental, prenatal, or perinatal endocrine manipulation. For example, a neonatally castrated male rat that shows a female receptive posture called *lordosis* when mounted by another male is sometimes considered to be "homosexual," as is the perinatally androgenized female rat that mounts other females. The male that mounts another male is sometimes considered to be exhibiting heterosexual behavior, as is the

female that displays lordosis when mounted by another female. In this particular laboratory paradigm, sexual orientation is defined in terms of specific behaviors and postures. In contrast, human sexual orientation is defined not by the motor patterns of copulation, but by one's pattern of erotic responsiveness and the gender of one's preferred sex partner.

Because of the problems in equating rodent behavior with human sexual orientation, researchers have used a variety of strategies to assess partner preference in animals. This is sometimes done by seeing whether a test animal chooses to approach a male or a female stimulus animal placed in opposite arms of a T-maze. Although some unaltered laboratory animals will spontaneously direct most of their sexual behaviors toward their own sex, animal studies of sexual orientation are usually carried out on animals that have been experimentally manipulated. For example, a genetically male rodent may either be castrated as a neonate, depriving his developing brain of androgens, or particular androgen-responsive regions of his brain may be destroyed. However, in order to activate the display of female-typical behaviors and preferences in such male animals, estrogen injections are also required in adulthood. Similarly, androgens must be given again after puberty in order to activate mounting behaviors in female animals that were exposed to androgens prenatally. Because adult homosexual men and women have hormonal profiles that are indistinguishable from those of their heterosexual counterparts, it remains unclear how findings based on such hormonally abnormal animals say anything informative about human sexual orientation. Studies are beginning to examine the neurobiological underpinnings of sexual behavior in male sheep that, without any experimental manipulation, do not mount females but instead mount only males. The males that receive mounts are often weak or incapacitated and unable to fend off the mounting sheep. These so-called gay sheep have not been described as displaying the feminized patterns of behavior that in most animal models are associated with developmental androgen deficiency in males.

Ethical constraints make it impossible to test the prenatal hormonal hypothesis in humans. That is, one cannot experimentally manipulate the prenatal hormonal exposure of human fetuses to study their effects on subsequent sexual orientation. However, there are alternative strategies. These include assessing sexual orientation in individuals with known or suspected prenatal hormonal abnormalities and assessing presumed correlates of prenatal hormonal exposure in known homosexual and heterosexual individuals.

ORIENTATION FOLLOWING DOCUMENTED PRENATAL HORMONAL ABNORMALITIES: HUMAN INTERSEXES. The prenatal hormonal hypothesis would predict that a large proportion of men with medical conditions known to involve prenatal androgen deficiency or insensitivity would be homosexual or transsexual.

The same is true of a large proportion of women exposed prenatally to excess androgens. Because androgens are required for the development of the external male genitalia, such individuals may be born with genitals that are intermediate in morphology between those of normal male and females. Affected individuals, historically referred to as *hermaphrodites* or *pseudohermaphrodites* (depending on their gonadal status), have more recently been referred to as *intersexed.* Somatic intersex conditions are a subset of *DSD,* and the term, DSD, is being used increasingly as if it were synonymous with somatic intersex conditions.

The incidence of homosexuality (relative to the assigned gender) is increased in some intersex conditions, particularly among female-assigned individuals who were exposed to elevated androgens prenatally. Interpretation of these results, however, is complicated by several complex variables. For example, the rearing of individuals born with genital ambiguity may not be unambiguous with regard to the assigned gender, despite the intentions of physicians and parental figures. This may occur when the affected individual has had multiple surgeries in an attempt to construct more normal appearing genitalia. Even surgeries performed so early that the individual has no memory of them may leave physical scars or other anomalies that give rise to concerns about gender. Moreover, parents may remain ambivalent about the gender of the child and communicate this unwittingly and nonverbally to the child. Thus, studies that detect an increased incidence of lesbian or bisexual identity, behavior, or fantasy in female-assigned individuals with somatic intersex conditions do not provide unequivocal support for the direct model of androgenic effects on sexual orientation.

Individuals with *complete androgen insensitivity syndrome* (CAIS or complete AIS) due to mutations of the androgen receptor are an important exception to the rule of increased homosexuality among female assigned intersexes who were exposed to elevated prenatal androgens. These individuals were exposed to high androgen levels in utero; however, their tissues lack the capacity to respond to androgens. Therefore, their external genitalia follow the path of female differentiation, although they are genotypically male and have normally functioning testes. Because of their external appearance, they are assigned as females at birth and reared as such. Until relatively recently, they often remained undiagnosed until puberty. Despite having all of the known metabolic machinery and receptors necessary to both defeminize and masculinize sexual behavior in the rat, all reported adult cases of CAIS have been described as maintaining a female gender identity and exhibiting stereotypically female behavioral patterns and interests, including sexual interest in males. These observations have led some to conclude that in contrast to rodents, male-typical differentiation of the human brain is mediated by the activation of androgen receptors and does not require that androgens be first converted

to estrogens. Evidence from studies of nonhuman primates is consistent with that interpretation. Nevertheless, the possible contributions of the female assignment and rearing of individuals with complete androgen insensitivity should not be overlooked with respect to development of their gender identity and sexual orientation.

PRESUMED CORRELATES OF PRENATAL HORMONAL EXPOSURE IN HOMOSEXUAL INDIVIDUALS.

Another strategy in attempting to establish a link between prenatal hormonal exposure and sexual orientation has been to examine presumed correlates of such exposure in known homosexual and heterosexual individuals. The most obvious correlates would be abnormalities of testicular or ovarian function and abnormalities in the differentiation of the external genitalia. Such variations are rare in homosexual individuals and lend no support to the prenatal hormonal hypothesis; however, androgens could influence such variables and sexual orientation at different periods of development. Alternatively, differentiation of the substrates that mediate sexual orientation could be orchestrated by an as yet unidentified androgen-mediated pathway that does not influence the other variables studied. Animal research suggests two additional correlates of early androgen exposure that have been explored in humans as possible correlates of sexual orientation: The pattern of luteinizing hormone secretion elicited by estrogen and neuroanatomical sexual dimorphisms.

LUTEINIZING HORMONE RESPONSES. In rats, exposure of the developing hypothalamus to aromatizable androgens (i.e., androgens that can be converted to estrogens) determines the signal that the adult's brain will relay to the pituitary gland in response to high levels of estrogen in the bloodstream. If rats do not experience high levels of aromatizable androgens at a certain early phase of development (as in normal females), the adult brain will respond to estrogen by directing the pituitary to increase its secretion of luteinizing hormone, a phenomenon referred to as *positive feedback*. However, if the developing rat brain is exposed to high androgen levels (as in normal males), it will not be able to support this positive feedback response. Thus in rats, the presence or absence of this positive feedback response can be used to infer the pattern of the brain's early androgen exposure. Specifically, the positive feedback response can be elicited in adulthood from normal females and from males that were castrated during the sensitive period of sexual differentiation of the brain, but not from normal males or from females that were experimentally exposed to high androgen levels in early development.

These observations generated speculation that in humans the response of luteinizing hormone to an injection of estrogen would also indicate the sexually differentiated state of the brain and correlate with sexual orientation. Empirical research in this area has produced conflicting results; however, the current consensus opinion is that the feedback

response in humans does not correlate with sexual orientation. Moreover, in humans and other primates convincing research suggests that the brain mechanism regulating luteinizing hormone is the same in both sexes, rather than taking two sexually distinct forms as it does in rodents. If there is no sex difference in the feedback mechanism, it cannot be argued logically that the mechanism should exhibit atypical sexual differentiation in gays and lesbians. Since the positive feedback response is required for cyclical ovarian function and fertility in females, menstruation and pregnancy in lesbians supports this interpretation.

NEUROANATOMICAL SEXUAL DIMORPHISM. Over the past three decades, sex differences have been confirmed in the size of several brain structures in a variety of laboratory animals. These findings have generated speculation concerning the existence of parallel differences in the human brain associated not only with sex, but also with sexual orientation.

Most of the structural sex differences identified involve specific cell groups within a broad region of the rodent hypothalamus that participates in regulating a variety of functions including sexually dimorphic copulatory behaviors. Like the sex differences in copulatory behaviors, several structural sex differences develop in response to sex differences in early androgen exposure: High androgen levels at the appropriate time lead to male-typical anatomy, whereas low levels lead to female-typical anatomy. Consequently, behavioral sex differences are thought to be mediated, at least in part, by the structural differences.

The best-studied anatomical sex difference involves a cell group that straddles the medial preoptic and anterior regions of the hypothalamus—the sexually dimorphic nucleus of the preoptic area (SDN-POA). In the rat, this structure is five to eight times larger in males than in females. Damage to the preoptic region decreases mounting behavior in laboratory animals, whereas electrical stimulation elicits mounting behavior.

The belief that the SDN-POA participates in regulating male sex behavior in rats has led the search for a comparable nucleus in humans. A cell group called *nucleus intermedius* but also known as the *first interstitial nucleus of the anterior hypothalamus* (INAH1) was described by one laboratory as being larger in men than in women and is designated as the *sexually dimorphic nucleus* (SDN) in most current atlases of the human hypothalamus. Of four laboratories that examined this nucleus for variation with sex and sexual orientation, only one found it to vary with sex and none found variation with sexual orientation. Its designation as the SDN, therefore, appears to be premature if not unwarranted.

Another structure, the *third interstitial nucleus of the anterior hypothalamus* (INAH3), is perhaps a more promising candidate for homology with the rat's SDN-POA. Three independent laboratories found this nucleus to be much larger in presumed heterosexual men than in

women, and one laboratory found the nucleus to contain more neurons in men. By extrapolation from animal work, this sex difference is widely believed to reflect sex differences in early hormone exposure, although at present that hypothesis cannot be directly tested in humans. The AIDS epidemic has made it possible to study this nucleus in individuals whose medical records indicated homosexual behavior as the HIV risk factor. (Unless someone dies from complications of AIDS there is usually no documentation of sexual orientation in the medical records available for autopsy studies.) These studies suggest that the volume of INAH3 may be smaller in homosexual men than in heterosexual men, but the number of neurons within the nucleus does not vary with sexual orientation. The suggestion of volume reduction must be viewed skeptically for a variety of technical reasons. For example, tissue shrinks in the process of fixation for histological analysis. This shrinkage influences measures of volume but not measures of cell number. Thus, the finding of equal numbers of neurons in homosexual and heterosexual men may be a more reliable finding than the suggestion of a difference in the volume of the nucleus. Alternatively, the presence of an equal number of cells in a smaller volume could reflect decreased neuropil (e.g., synapses, dendrites) in homosexual men. Decreased neuropil could conceivably reflect low androgens either prenatally or at the time of death due to the antiandrogenic effects of some medications used to combat the opportunistic infections of AIDS patients. A bias could have been introduced into these studies since the majority of heterosexual male subjects with AIDS were intravenous drug users who, compared to gay males, are likely to be less compliant with medication regimens. Additional studies are needed to clarify these issues.

Speculation regarding the function of INAH3 has been based on the assumption that it is the homologue of the rat's SDN-POA. The finding that the size of the SDN-POA correlates positively with the frequency of mounting behavior displayed by male rats established the belief that the SDN-POA participates in regulating male sex behavior. That interpretation, however, is at odds with the fact that the rat's nucleus can be destroyed by electrolytic lesions without any discernible effect on mounting behavior. A better case can be made that the SDN-POA functions to inhibit the expression of a female mating posture by male rats. In rats in which intracerebral prostaglandin E2 was manipulated, the volume of the SDN-POA was found to correlate with defeminization and not with masculinization of copulatory behaviors. Moreover, following lesions that include the SDN-POA, male rats can be induced to display the female copulatory response if they are also injected with estrogen and progesterone.

In addition to the hypothalamus, researchers have sought to identify variation with sex and sexual orientation in the brain commissures, the fiber bundles that connect the left and right hemispheres of the brain. The

rationale underlying these studies was drawn from speculation regarding the neuroanatomical basis of statistical sex differences in the lateralization of particular cognitive functions within the brain. Several lines of investigation suggest that women compared to men are likely to exhibit more bilaterally symmetric hemispheric activation while performing particular cognitive tasks. A popular hypothesis holds that this sex difference in brain activity reflects increased interhemispheric communication in women compared to men, and that this increased communication between the hemispheres is due to more interconnecting fibers between the hemispheres. More than 50 studies have examined the width or area of particular portions of the corpus callosum, the largest of the brain's commissures, for sex differences. Popular accounts to the contrary, the majority of these studies have found no sexual dimorphism. Studies, controlling for variation in the callosum as a function of handedness and aging, however, have found differences in the isthmus of the corpus callosum, which contains axons connecting right and left parietotemporal cortical regions. These regions mediate functions of language and spatial cognition that are asymmetrically represented and that, according to some studies, vary with both sex and sexual orientation. A study, yet to be subjected to replication attempts, found the isthmus to be larger in right-handed homosexual men compared to right-handed heterosexual men. In addition, several studies have examined the anterior commissure, a small bundle of fibers connecting portions of the left and right temporal lobes. These studies have produced conflicting results regarding variation with both sex and sexual orientation. One laboratory found this commissure to be larger in men than in women; another found it to be larger in women and gay men than in heterosexual men; one found no suggestion of a difference based on either sex or sexual orientation.

NEUROPHYSIOLOGICAL MEASURES. A number of neurophysiological measures have been reported to differ between homosexual men and women and their heterosexual counterparts. Several of these studies report that homosexual individuals exhibit neurophysiological responses more typical of heterosexuals of the other sex, but to date none have been adequately replicated by independent laboratories. These include studies of regional patterns of brain activation in response to exposure to putative pheromones. Although positive findings in this area should not be ignored and certainly merit attempts at replication, the existence of a functional vomeronasal organ in humans, thought necessary to respond to pheromones, is itself controversial. In addition, lesbians have been reported to exhibit masculinized otoacoustic emissions, echo-like waves emitted by the inner ear response to brief sounds.

ANTHROPOMETRIC CHARACTERISTICS. Several anthropometric characteristics,

believed by some to reflect androgen exposure, have been explored in relation to sexual orientation. The characteristics examined include measures of height and weight as well as the amount and distribution of facial hair, the ratio of shoulder width to hip width, the size of the genitalia, and more recently dermatoglyphic (finger print) characteristics, and finger length ratios. Most of these studies have been flawed in one or more ways, making the occasional positive findings difficult to interpret. Some of these flaws include reliance on self-reports of small self-selected samples, or on measures obtained by raters who were not blind to the sexual orientation of the subjects.

Finger length ratios continue to be an area of active investigation. The ratio of length of the second to fourth fingers (2D:4D) was first proposed as a marker for prenatal androgen action in 1998. Since then, over 100 studies have been published testing its association with prenatal androgens or using 2D:4D as a putative marker to investigate the association between prenatal androgens and a variety of outcomes including sexual orientation. The validity of digit ratios to serve as a marker of prenatal androgen action, however, is highly controversial. Some but not all studies suggest that adult women tend to have longer index fingers (second digits) than adult men relative to the length of the ring finger (fourth digit). In adults, the correlation between 2D:4D ratio is not high, with 60 percent overlap between males and females. Congenital adrenal hyperplasia treated soon after birth has been the gold-standard method for studying effects of prenatal or perinatal androgens. Examination of the 2D:4D ratio in affected individuals has produced equivocal results. Similarly, studies of the 2D:4D ratio as a function of sexual orientation have produced discrepant results. A decreased ratio has been reported with more consistency in lesbians relative to heterosexual women than has an increased ratio in gay relative to heterosexual men. Nevertheless, the inability to definitively link this ratio to prenatal hormone exposure, together with inconsistent replication studies, does not provide substantial support for the prenatal hormonal hypothesis of sexual orientation in either sex.

PROPOSED MECHANISMS OF ATYPICAL ANDROGEN EXPOSURE. As reviewed above, in men, clinical syndromes involving abnormal androgen function rarely have an impact on sexual orientation, whereas, an impact of elevated prenatal androgen is statistically demonstrable in women in the context of particular clinical syndromes. The *a priori* likelihood that very subtle changes in hormonal profiles and action, not strong enough to be clinically manifest, would affect sexual orientation is, therefore, not large unless some yet to be identified mediators of androgen function are found to specifically influence sexual orientation and no other androgenic functions. Maternal stress has been examined but is no longer widely believed to be a contributory factor. Genetic variations in the mediators of androgen action

have been explored without success; however, the possibility of epigenetic mechanisms (variations in DNA packaging or "epigenetic marking") rather than variation of genes, themselves, remains essentially unexplored). Such mechanisms could alter sensitivity to the actions of androgens and increase or decrease the functional impact of a given level of androgens.

In multiple samples investigated by one group of researchers, homosexual men have been found to have a greater number of older brothers than heterosexual men. According to these researchers, each older brother increases the "relative risk" of being homosexual by 33 to 48 percent, and this antecedent older brother effect accounts for the orientation of one of every seven gay men. The hypothesis advanced by these authors is that the late birth order, with more male siblings born earlier, could lead to a progressive immune response of the mother to products associated with a male, which could impair the defeminization and masculinization of subsequent male fetuses. The existence of such "antimale antibodies" is purely hypothetical as is the male-associated factor that would trigger their production. Androgens have been hypothesized in this regard; however, under usual circumstances they are not antigenic. The Y-linked minor histocompatibility locus has also been hypothesized as the triggering factor. However, how such mechanisms could selectively impair only certain, less well-established processes, such as the organization of the brain, and not others, like the formation of the genitalia, is not explained by this hypothesis. Nor does this theory explain why the majority of boys late in birth order do not become homosexual, even if their elder brother is homosexual. Finally, the effect has been reported in several studies as applying only to right-handed gay men, while left-handed gay men had fewer older brothers than expected. The biological premise of the antecedent older brother effect remains questionable, and attempts at replication by independent groups of researchers is needed.

Genetic Research. As in the neuroendocrinological research, some genetic studies have also been premised on the intersex hypothesis. For example, there have been attempts to show that gays and lesbians have chromosomal material typical of the other sex in their cells and attempts to link homosexuality with genetically controlled aberrations in the process of sexual differentiation. To date these studies have produced negative results. Other approaches, not necessarily wedded to the intersex hypothesis, including heritability and linkage analyses, have also been employed.

FAMILY AND TWIN STUDIES. Although homosexuality has been found to run in families, such findings are not helpful in distinguishing between genetic and environmental influences because most related individuals share environmental influences as well as genes. A first step toward disentangling genetic and environmental influences can be made by studying sexual

orientation in identical and fraternal twins. One such study assessed sexual orientation not only in the identical and fraternal twins, but also in the nontwin biological brothers and the unrelated adopted brothers of homosexual men. The concordance rate for identical twins (52 percent) in that study was much higher than the rate for the fraternal twins (22 percent). Assuming that environmental influences would be the same for all brothers reared together, the higher concordance rate in the identical twins is consistent with a genetic effect, as identical twins share all of their genes, while fraternal twins, on average, share only half of their genes.

It would be a mistake, however, to attribute the increased concordance rate in identical twins to increased gene sharing alone. If there were no environmental effect on sexual orientation, then the rate of homosexuality among the adopted brothers should have equaled the rate of homosexuality in the general population. Recent studies place the rate of homosexuality in men between 2 and 5 percent. That the concordance rate in adopted brothers was 11 percent (two to five times higher than in the general population) suggests a major environmental contribution. The rate for homosexuality among nontwin biological brothers was only 9 percent, a figure statistically indistinguishable from the 11 percent recorded for adopted brothers. If the concordance rate for homosexuality among nontwin brothers is the same whether or not the brothers are genetically related, then the concordance rate cannot be explained exclusively by genetics.

When considered together, the data from the studies of twins and adopted brothers suggest that the increased concordance in the identical twins may be due to a combination of both genetic and environmental influences. Such environmental influences could theoretically be exerted by the prenatal intrauterine environment as well as by the postnatal environment. Environmental factors have been shown to influence the epigenetic marking mentioned above, which can influence gene function. Perhaps the most interesting finding to emerge from twin studies is that approximately 50 percent of identical twins are discordant for sexual orientation even when they are reared together. The discordance for sexual orientation among genetically identical individuals reared in the same family has been consistent across studies and underscores just how little is actually known about the potential contribution of genes or environment to sexual orientation.

LINKAGE STUDIES. The notion of a *gay gene* has captured the popular imagination. It is worth underscoring, however, that gay genes are not required for homosexuality to run in families or for researchers to determine that homosexuality is heritable. *Heritability* has a precise technical meaning and reflects only the *degree* to which a given trait is *associated* with genetic factors. It says nothing about the specific genetic

factors involved or about the mechanisms through which they exert their influence. Furthermore, the concept of heritability provides no information about how a particular trait might change under different environmental conditions. Thus, homosexuality could be highly heritable even if genes influenced sexual orientation entirely through very indirect pathways. The question is often raised as to how genes predisposing to homosexuality could be maintained in the gene pool, assuming that reproduction would be decreased in individuals who carry them. The *Kin Selection Hypothesis* proposes that such genes could be maintained in the population if the costs of not reproducing directly are offset by enhancing the reproductive success of kin, perhaps through altruistic behavior or some other mechanism. For example, one study found increased fecundity among the female maternal relatives of male homosexuals and attributed this finding to genes predisposing to homosexuality in men.

Genes alone cannot directly determine any behavior or psychological phenomenon. Instead, genes direct a particular pattern of ribonucleic acid (RNA) synthesis that, in turn, specifies the production of a particular protein that then may influence behavior. There are necessarily many intervening pathways between a gene and a specific behavior and even more intervening variables between a gene and a pattern that involves both thinking and behaving. The term *homosexual gene* is, therefore, without meaning, unless one proposes that a particular gene, perhaps through a hormonal mechanism, directs the brain to develop in such a way that an individual can only be erotically attracted to others of the same gender. Evidence in support of such a simple and direct link between genes and sexual orientation is lacking.

In 1993, a highly publicized study presented statistical evidence that genes influencing sexual orientation reside on a portion of the X chromosome known as the q28 region. An independent research team was unable to duplicate the Xq28 finding in men using a comparable experimental design, and no group has found evidence that Xq28 is linked to sexual orientation in women. Two full genome scans have been published to date. One failed to produce any statistically significant findings except for a replication of the Xq28 finding. It is important to note that this study included the same subjects upon which the original Xq28 study was based, and that the Xq28 linkage was confirmed only for those subjects. A more recent full genome scan also implicated genes in the Xq28 region in male homosexuality and additionally implicated genes in a region of chromosome 8 that also had been implicated in a previous study. Mechanisms as to how genes in these regions might exert their influence on sexual orientation in men remain speculative. Such speculation includes the possibility that such genes might increase sexual attraction to men in both men and women, leading to increased reproduction in women and, therefore, maintaining the genes in the population. Candidate genes,

including the androgen receptor and aromatase enzyme, have also been investigated for variation with sexual orientation with negative results to date. These genes were chosen for examination because of their role in sexual differentiation of the brain. As noted above, however, evidence for a role of the aromatase enzyme in sexual differentiation of the primate brain is currently lacking.

Directions for Future Research. Biological research on sexual orientation has been driven largely by the prenatal hormonal hypothesis and has focused on the potential role of prenatal androgens. Early androgen exposure clearly organizes the mating behavioral repertoires of laboratory animals; however, it is not clear how the regulation of stereotyped reproductive behaviors in animals pertains to sexual orientation in humans. Nevertheless, it does not seem likely that early androgens influence the adult sexual behaviors of all mammals studied except humans. It is clear, however, that the role of aromatization of androgens to estrogens within the developing brain plays a lesser, if any, role in primates compared to rodents. Although early androgen exposure in human females increases the likelihood of some degree of gynephilia in adulthood, there is no compelling evidence for differences in prenatal androgen levels between heterosexual and homosexual men. One possibility that has yet to be explored is that some yet to be identified mechanism downstream of the androgen receptor plays a role in the development of gynephilia, but no role in other aspects of androgen action such as external genital development. A variation of such a hypothetical mechanism might preclude the development of gynephilia. Females without this hypothetical variation would then be predisposed to gynephilia following relatively mild exposure to elevated androgens prenatally, whereas men with this variation would be predisposed to androphilia despite normal androgen levels prenatally. Currently, relatively little is known regarding the biochemical pathways that are influenced in the primate brain by the binding of androgen to its receptor. Microarray technologies for full transcriptome analysis offer a mechanism for elucidating these pathways in order to identify candidate genes for further study. As noted above, epigenetic variations, which can be environmentally influenced, could lead to such variations in the effects of androgens in the absence of genetic differences (i.e., differences in DNA sequences). Recent technological innovations now allow high throughput examination of epigenetic variations. Large-scale heritability and linkage studies may also provide insights; however, the identification of genes associated with sexual orientation alone will be relatively uninformative in the absence of studies to identify their mechanisms of action. Such mechanisms could prove to be very indirect and mediated via an influence on temperamental characteristics rather than a direct influence on the hypothetical brain

circuits that mediate sexual orientation.

Cross-Cultural Studies

Anthropological studies of human sexuality have demonstrated enormous variation in gender and sexual organization and behavior in different cultures. Ford and Beach's cross-cultural 1951 study, *Patterns of Sexual Behavior,* was a ground-breaking attempt to sort out biological aspects of sexuality, presumed to be universal in all societies, from social components expected to vary from culture to culture. They reviewed the status of homosexuality in 78 cultures of the mid-20th century. Twenty-eight societies, including the United States, disapproved of homosexuality. Forty-eight cultures approved of some forms of homosexuality. One group, the Tswana of Africa, disapproved of homosexuality in men and approved of it in women. Cultures that disapproved of homosexuality commonly reported that homosexual behavior was absent or rare or carried on in secret. In those societies, homosexuality was stigmatized from childhood on into adulthood by ridicule and punishment, and sometimes even the death penalty.

Building on the work of Margaret Mead, Ruth Benedict, and other anthropologists, researchers have been able to classify structural types of sexual development and forms of homosexual organization in different cultures. Gilbert Herdt, in particular, explicates cross-cultural models based on his study of ritualized homosexuality among the Sambia tribe of New Guinea and on his review of findings from other anthropological studies of sexual practices. Although traditional developmental theory posits a presumed continuity in sexual development across the lifespan, Herdt suggests instead the idea of "discontinuity" in sexual development to explain variation in the psychological experience and symbolic meaning of sexuality. He also explores the cultural aspects of fluidity in experience and expression of sexual orientation that has defined much of the theoretical discourse on bisexuality.

Herdt also describes a typology of homosexuality based on a cross-cultural organization of same-sex sexual behavior. The first type is *age-structured homosexuality,* usually involving older and younger males and generally including a sequential pattern of childhood same-sex practices that yield to primarily heterosexual behaviors in adults. An example would be the male homosexual practices of classical Greece (*pederasty*). Phenomena more akin to current conceptualizations of transgenderism have previously been referred to as *gender-reversed homosexuality,* which refers to a reversal in the normative sex-role dress and behavior. An example of such gender-reversed homosexuality is the Native American *berdache,* a man permitted to assume the gender role of a woman, and the *fa'afafine* of modern Samoa. The third type identified by Herdt is *role-specialized homosexuality,* in which same-sex sexual activity is restricted

to certain social roles and positions. This type occurs among the shamans of some Native American tribes who, in following dreams or visions revealing their special status, will cross-dress and engage in homosexual behavior. Role-specialized homosexuality was also found among certain working-class women in 19th century China. Finally, Herdt describes the *modern gay movement,* consisting of egalitarian same-sex practices, as a fourth type of homosexual organization. Although acknowledging that hierarchical relationships still exist in parallel with traditional forms of heterosexual relationships, he concludes that homosexuality in contemporary America is historically unique and that it comprises a sexual orientation, a social identity, and a political movement.

Psychosocial Studies of Development

Classical theories of psychological development posit the origins of adult personality and identity in childhood experience. These theories often lay out presumed, linear pathways of development that progress sequentially from childhood to adulthood. In the later part of the 20th century, however, focus shifted to the distinct processes of development during adulthood itself. The development of a sexual orientation might be better understood by examining psychological and interpersonal events throughout the lifespan that may be characterized by both continuous and discontinuous patterns of development. There is little empirical research to support the traditional view that early childhood behaviors, awareness, and feelings are definitive precursors of adult behaviors, awareness, and feelings. It should be further noted that most psychosocial studies are plagued by problems of sampling and have tended to focus on white, urban, openly identified, and community-affiliated gay men and lesbians. It is not known whether one can generalize the development and experience of the latter group to the experience of individuals coming from other ethnic, racial, or culturally diverse groups.

Findings from psychosocial studies will vary depending on the research methodology. For example, in the mid-20th century retrospective studies of adult homosexual patients in psychoanalysis, the narratives these individuals told of their upbringing were often consistent with cultural notions of homosexuality. In other words, how these individuals remembered their early childhood may have resulted from social learning of popular theories of homosexuality, either before or during their psychoanalytic treatment, with subsequent internalization of the theory as part of their life scripts. In addition, many studies of purported childhood precursors of adult homosexuality often take a narrow focus on selected characteristics (i.e., effeminacy in boys) associated with stereotyped notions of adult homosexuality. Consequently, such studies tend to ignore the enormous individual variability in patterns of development toward and characteristics associated with adult sexual identity.

Historically, children with severely atypical gender behavior (GD of childhood, or GDC), disturbed homosexual adults voluntarily seeking psychiatric treatment, and incarcerated prisoners served as the primary source of statistics about and models for the psychological development of homosexuality. Actual empirical studies of prepubescent homoeroticism in American society are nonexistent, although there are some recent studies of gay, lesbian, and bisexual adolescents. The potential psychosocial understanding that might derive from ethnographic and other social science approaches to studying sexuality in children has been applied only in other cultures and has thus not been utilized as a tool for the study of the development of sexual orientation or sexual identity in American society. Nevertheless, there are theoretical models hypothesizing how adult gay, lesbian, and bisexual identities develop in Western society that are based on a small but growing number of empirical studies of the specific behaviors and events associated with establishing these identities.

Origins of Homosexuality in Childhood and Adolescence. Early attempts to describe childhood psychosocial antecedents of adult homosexuality consisted primarily of psychoanalytic formulations derived from the histories of psychiatric patients. These usually depicted trauma, loss, and disturbed family relationships in childhood as causal mechanisms. Modern sexologists do field studies of larger samples of nonpatient populations. These studies are often guided by essentialist views that assume a particular personality characteristic or specific behavior is a reliable predictor of adult sexual orientation. To date, no single set of characteristics or unitary pathway leading to adult homosexuality (or heterosexuality) has been identified.

One robust finding in studies of childhood factors associated with adult homosexuality is the recollection of early gender nonconformity. That is to say, femininity in boys and masculinity in girls appear to be correlated with the development of adult homosexuality in some individuals. However, this is not true for all adults who identify as gay, lesbian, or bisexual, and any specific relationships between childhood gender nonconformity and homosexuality remain unknown.

Recent studies have focused on the adjustment, health, and mental health issues of young people who self-identify as gay, lesbian, or bisexual who either present to community centers providing services or who are in college. In recent years, adolescents appear to be coming out at a younger age and not all are accepted by their parents. The US Department of Health and Human Services now requires that all organizations serving runaway and homeless youth be equipped to serve LGBT youth, who represent a disproportionate segment of this population.

Acquisition of Gay and Lesbian Identities. It is useful to know something

about the theoretical description of gay, lesbian, and bisexual identity development in an individual as well as the meaning(s) of personal identities at a cultural level. From this social constructivist perspective, the appearance of LGB identities—and the detailed description of their emergence and multifaceted expression within both society and the individual—can be seen as an outcome both of personal growth and development for the individual and of the cultural evolution of the concept of a "homosexual person."

Models of LGB identity acquisition describe an intrapsychic and interpersonal process whereby a subjective experience of oneself as lesbian, gay, or bisexual becomes increasingly congruent with other people's perception of that identity. This process is generally conceived of as occurring over time, in sequential stages, and in a manner that leads to an increasing integration of one's experience of being lesbian, gay, or bisexual into other aspects of one's sense of self. Vivian Cass conceived of identity development as consisting of a series of stages. The first is a period of homosexual self-awareness, a growing awareness and possible confusion about the presence and meaning of same-sex erotic attractions. In people who become gay, lesbian, or bisexual, the period of homosexual self-awareness is followed by a time during which the individual considers accepting these feelings and begins to incorporate them into experiences. In many cases, this phase is followed by an extended phase of growing acceptance of and perhaps even pride in the new identity.

The process of identity development leads to different outcomes for different individuals; it is influenced by interpersonal effects and consequences and feedback from others and may be interrupted or terminated at any point. Most models of gay, lesbian, and bisexual identity development incorporate the completion of several events or tasks: *Coming out,* or becoming aware of same-sex attraction and disclosing it, at some point, to others; forming relationships with other gay and lesbian persons; engaging in same-sex sexual behaviors; becoming involved in gay and lesbian activities and communities; establishing primary relationships with persons of the same sex; and learning how to integrate one's sexual orientation identity into other aspects of the self.

Models of gay and lesbian identity acquisition sometimes present an image of progressive elaboration and synthesis of homosexual orientation with other aspects of the self. However, individuals vary considerably in the degree to which they follow such models. In some persons, the development of a gay, lesbian, or bisexual identity may be inhibited in response to a number of factors. These include a punitive or harsh internal or environmental response to openly expressing aspects of such an identity or a redirection of one's focus on other developmental tasks, such as schooling or career development. Individuals from certain religious communities may eschew adopting a gay or lesbian identity. As models of

gay, lesbian, or bisexual identities have mostly developed based on experiences of urban white people, individuals who live in other types of communities (e.g., rural areas and communities of color) may participate in same-sex sexual activities yet eschew an identity associated with these behaviors.

Individual variations in the extent of development of gay, lesbian, and bisexual identities may result both from prohibitive factors, either internal or external, that interfere with optimal personal growth and from normative psychosocial influences, reflecting diverse individual experiences and different adaptational needs. The tasks for any clinician working with these individuals are to appreciate the individuality of identity development in each patient, to understand their particular psychological and social resources and needs in this area, and to appreciate the risks and benefits in differing pathways of development. Ultimately, each individual will develop her or his own identity; the clinician can only help to optimize the understanding of the meaning of various options and to establish opportunities for maximizing freedom in making choices about often conflicting outcomes.

Homosexuality and Nongay Identities. The complexities involved in establishing a gay, lesbian, or bisexual identity require the clinician's tolerance, respect, patience, and sophistication about sexual matters. The different subjectivities one may encounter in working with gay and lesbian patients can often be understood and appreciated in concepts like hiding/revealing, awareness/unawareness, and acceptance/nonacceptance. Patients with same-sex feelings can be helped to understand that their own sexuality will, in part, be shaped by the meanings attributed to homosexuality by a predominantly heterosexual world. The subjectivity of a sexual identity is shaped by culture and language, often with little or no regard to prevailing scientific categories of sexuality. Some broad subjectivities are identified here to provide clinicians with a way to understand the range of ways in which patients may manage their same-sex feelings (Fig. 21.2–1).

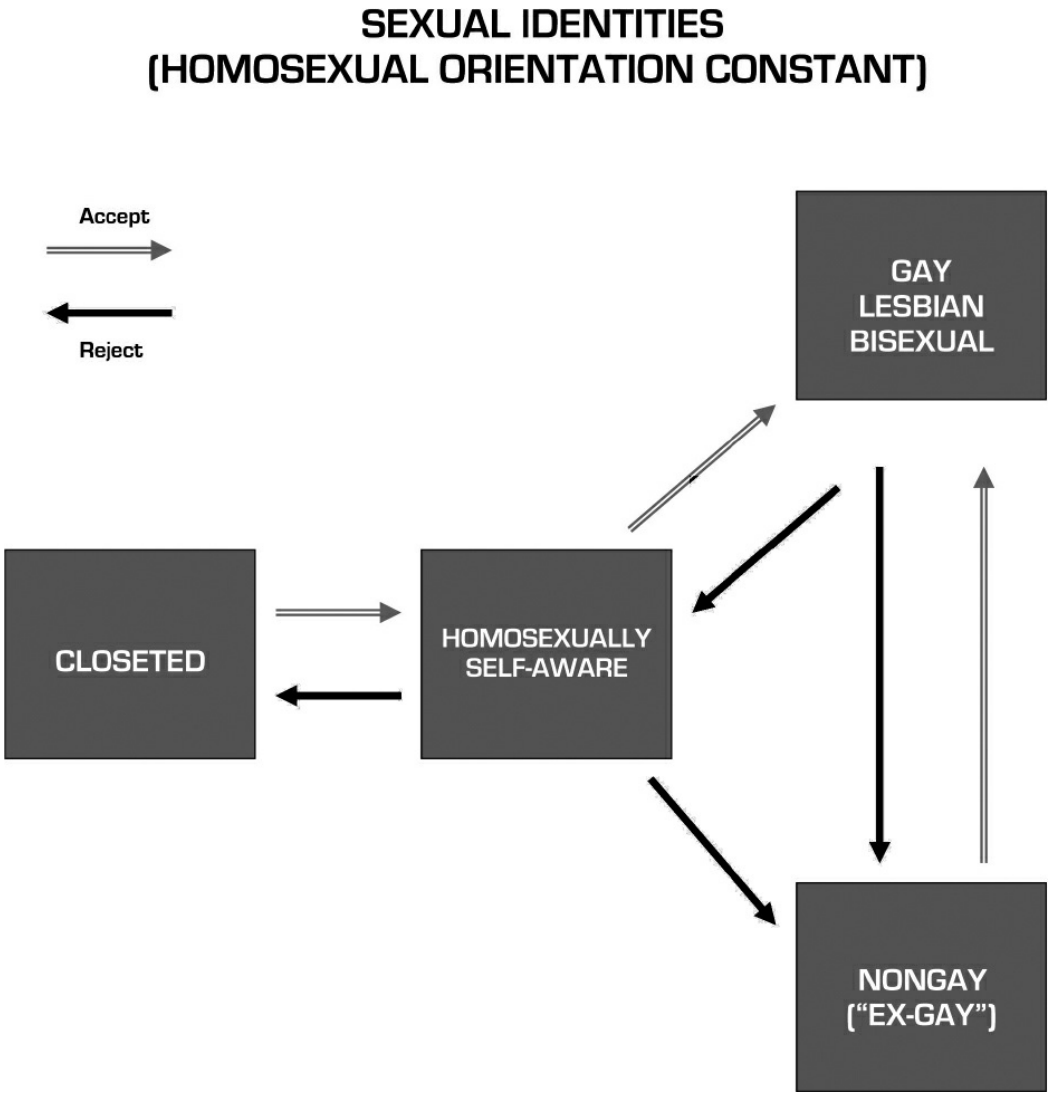

FIGURE 21.2–1. Sexual identities (homosexual orientation constant).

Closeted individuals are unable to acknowledge to themselves or to others that they have homoerotic feelings and fantasies. From a behavior perspective, the popular terms *closeted* or *in the closet* may refer to a person who is actively engaging in homosexual acts but concealing them. However, from the retrospective accounts of many gay men and lesbians, the experience of being in the closet is often more psychologically complex. The feeling of same-sex attraction is often unacceptable and as a result unavailable to direct awareness or integration into the individual's public persona. It may therefore be hidden from the self as well as from others. A closeted individual may not act on the feeling at all, or may do so only in a *dissociated state*. Dissociated feelings are not unconscious, as in the classical concept of latent homosexuality, but they are out of awareness and only sometimes accessible to consciousness, at certain times or in certain situations. Subjectively, a closeted man may tell himself he really does not have same-sex feelings, or, if he is aware of the feelings, he hopes they have some other meaning besides homosexuality. This psychological state of

being closeted is often reported retrospectively by individuals only *after* they have moved on to a later stage. It is also entirely possible for an individual never to leave this stage.

Homosexually self-aware individuals can acknowledge, at least to themselves, that they have homoerotic or same-sex feelings and attractions. However, these subjectivities and behaviors may vary from individual to individual. A person may choose not to act on the feelings or may act on them in a secretive way. Some behaviorally closeted individuals are homosexually self-aware and some are not. Some homosexually self-aware individuals consider the possibility of accepting and integrating these feelings into their public persona. For many individuals, as in Cass's model, this subjectivity may be a normative phase in eventually coming out and affirmatively accepting a gay or lesbian identity. Others may not accept these feelings, as in the example of a homosexually self-aware individual seeking a celibate lifestyle as a way of binding and avoiding a problematic gay or lesbian identity. For a clinician to assume of a patient that "She is gay but doesn't know it" is to make an external judgment that may offer little insight into an individual's subjective experience. Such a phrase could describe a closeted woman who is truly unaware of her homoerotic feelings; it could just as well describe a homosexually self-aware woman who knows she has these feelings, but is trying to prevent others from knowing she has them; and finally, it could describe a lesbian who is out to a limited circle of individuals, recognizes her own sexual feelings, but is selective about revealing them.

A third subjectivity, one that does not yet have a common cultural definition, is that of *nongay-identified individuals*. These people are aware of same-sex feelings, may have acted on them, and may even have come out as gay or lesbian. However, these individuals struggle against accepting any meanings that might naturalize their homosexual feelings. They may enter into heterosexual marriages. They may even engage in homosexual behaviors but still not feel that they have a gay identity. Because almost all gay men and lesbians who come out go through a phase of rejecting their own homosexual feelings before accepting their own gay or lesbian identity, it is sometimes assumed that all individuals will overcome this period of rejection. However, this is not always the case, and some individuals may never accept either their homosexual feelings, a gay identity, or both. These people may seek out ways to change their sexual orientation (see below). In an effort to disidentify with homoerotic feelings and activities, some refer to themselves as either "ex-homosexual" or "ex-gay." In many ways, this is a new cultural subjectivity that has grown out of response to the growing presence of openly gay, lesbian, and bisexual identities.

When conceptualizing sexual identities, they should not be thought of as existing on a developmental continuum or as being associated with more or less psychopathology. It should also be emphasized that this approach to

understanding individual subjectivities does not view sexual identities as offering any diagnostic information about an individual. In addition, these subjectivities are not mutually exclusive; there is overlap between them and differing motivations within them. Nor are these sexual subjectivities immutable. Although it may be difficult if not impossible to change one's sexual orientation, it is entirely possible to change attitudes about one's sexuality. As the subjectivities outlined above are shaped by individual and cultural factors, there may be a wide range of psychosocially constructed attitudes and responses that individuals may develop toward their own homosexuality. Furthermore, the homosexual subjectivities outlined above have no exact heterosexual equivalents, as the need to hide one's interest in the other sex is usually not intrinsic to forming a heterosexual identity.

LIFE STAGE AND CULTURAL ISSUES

The day-to-day lives of lesbians and gay men may often be similar to those of heterosexual cohorts. However, gay people face unique developmental tasks throughout the life cycle. For example, during childhood and adolescence, young people who become aware of significant attractions to someone of the same sex will need to understand what makes them different from the majority of people around them. Some will choose to disclose their homosexuality to others through the process of coming out, some will wait until they are older to do so, and some will never do so. During adulthood, these men and women will confront unique challenges as they try to establish close relationships and to create families. They may do so even without traditional, formal rituals that celebrate their relationships or laws to protect their families. However they navigate the life cycle, gay men and lesbians and bisexual men and women will need to contend both with the adversity of having a stigmatized sexual identity, as well as the potentially life-affirming factors associated with finding one's true self, finding one's true partner, and being part of a community where one can be oneself.

Antihomosexual Attitudes

Antihomosexual attitudes were once considered normative in most social and institutional settings, including in schools, among the police, in the armed services, and within religious organizations. However, as cultural attitudes have shifted, many traditionally antihomosexual groups and organizations have changed their positions. As a consequence of increasing cultural debates about homosexuality, many high schools and universities now offer clubs—gay-straight alliances—for students of all sexual identities to socialize. Many Western countries allow openly gay and lesbian personnel to serve in their armed forces and police departments. Some religious denominations are appointing openly gay and lesbian priests,

ministers, and rabbis to serve their congregations. Despite these changes, antihomosexual attitudes are still widespread, and it is helpful to identify their potential impact on gay, lesbian, and bisexual patients.

The terms *antihomosexual attitudes* or *antigay and antilesbian attitudes* are used here to refer to a wide range of critical and disapproving beliefs and feelings about homosexuality, gay men, and lesbians. The term *homophobia* refers to a fear or hatred of homosexuality and of gay and lesbian persons (*external homophobia*) or the self-hatred that gay people feel for themselves (*internal homophobia*). All LGB persons, to some extent, contend with *internalized homophobia*—they incorporate disapproving societal views of homosexuality and experience these feelings and beliefs as critical self-evaluation. *Transphobia* is a term increasingly being used to describe a range of negative attitudes and feelings toward transsexualism and transsexual or transgender people.

The effects of internalized homophobia on the development of one's sexual identity will differ widely at an individual level. Among other things, however, these feelings can lead to the suppression or closeting of awareness of one's own same-sex attraction, impair acceptance of one's homosexuality, and interfere with the integration of one's homosexual orientation into other aspects of identity. For many gay men and lesbians, internalized homophobia is experienced as a feeling of devaluation and limitation, not only of their homosexuality but of themselves; for some it may play a role in the development of psychiatric and behavioral symptoms, including depression, anxiety, denial, and perhaps even suicide.

The term homophobia has gained widespread popular currency and has been generally, if inexactly, applied to a range of other antihomosexual attitudes. The term has been criticized for suggesting that the existence of antihomosexual attitudes in an individual indicates the presence of a phobia or a formal psychiatric diagnosis. However, cultural attitudes are outside the realm of psychiatric classification. Thus, the occurrence of antihomosexual attitudes does not necessarily represent the existence of a mental illness, distressing as they may be to an individual who holds them. Early research on homophobic attitudes demonstrated that they are correlated with certain demographic and personality characteristics, including sex, geographical location, religious beliefs, and degree of authoritarianism. Studies have also shown that some individuals with antihomosexual attitudes may regard same-sex behaviors as a threat to their cherished personal values.

Heterosexism has been defined as "an ideological system that denies, denigrates, and stigmatizes any nonheterosexual form of behavior, identity, relationship, or community." This system regards heterosexuality as *normal* and *natural* and treats gay and lesbian people as outsiders who are neither. Heterosexism organizes experience from a point of view that naturalizes and idealizes heterosexuality and either dismisses or ignores a

gay subjectivity. Heterosexist beliefs can serve to enforce the maintenance of "compulsory heterosexuality" in society. Some common expressions of heterosexism include valuing heterosexual relationships over homosexual ones, an intensely felt opposition to gay and lesbian parenting—regardless of individuals' actual parenting qualities—and the belief that lesbians and gay men are unsuited for certain professions. Heterosexist attitudes are pervasive in Western cultures where most books, plays, movies, and print and television advertising contain naturalized depictions of heterosexuality. Heterosexism is not necessarily motivated by fear or hatred; in many ways, it is primarily self-referential. However, in the experience of many gay people, heterosexism is frequently experienced as antihomosexual, which is why it is sometimes confused with homophobia.

Heterosexism does not necessarily imply malice or ill will toward lesbians and gay men. Nevertheless, in extremis, it may equate the nonheterosexual with the immoral and unnatural and can rationalize the condemnation of same-sex feelings and behaviors. *Moral condemnations of homosexuality* treat homosexual acts as intrinsically harmful to the individual, to the individual's spirit, and to the social fabric. Those who condemn same-sex activities believe that heterosexist traditions of philosophical, legal, and religious opposition to homosexuality are, in and of themselves, sufficient reasons to forbid its open expression in the modern world. This reasoning led a US Supreme Court majority to uphold the State of Georgia's right to ban same-sex acts in the 1986 case of *Bowers v. Hardwick*. The justices stated that their decision did "not require a judgment on whether laws against sodomy are wise or desirable" but was based on the "ancient roots" of proscriptions against consensual homosexual activity and a longstanding condemnation of those practices in "Judeo-Christian" moral and ethical standards. In other words, their decision justified the censure of homosexuality in the present because it has been censured in the past, and argued that antihomosexual laws are permissible because they have always existed. Ironically, more than a decade later, that same statute was overturned by the Georgia Supreme Court in the case of a heterosexual, male defendant accused of engaging in so-called unnatural sexual practices with a woman. Then, in 2003, the US Supreme Court made a 6 to 3 historic ruling in *Lawrence and Garner v. Texas* when it found the rest of the country's sodomy laws unconstitutional.

Many contemporary, antihomosexual religious authorities speak of embracing the homosexual person, but not the person's homosexuality. This antihomosexual attitude is sometimes referred to as "loving the sinner and hating the sin." It derives from a belief that one *chooses* to be "a homosexual" despite prohibitions against it. As they believe all sinners can be saved, antihomosexual authorities believe they have a religious duty to care about those who make immoral choices. It should be underscored that *choice* is a charged word in moral debates about homosexuality. Similarly,

at the other end of this cultural divide, gay people believe they *choose* to act on their homoerotic feelings in order to avoid the choice of suffering in the closet.

Antihomosexual attitudes—in the form of homophobia, heterosexism, and moral condemnations of homosexuality—serve to shape important aspects of development and day-to-day experience in the lives of lesbians and gay men. Antihomosexual attitudes can ultimately lead to *antigay or antilesbian violence*. Even in the absence of overt violence, however, antihomosexual attitudes impinge on the completion of normative developmental tasks in some gay men and lesbians and may delay or prevent successful adaptation and achievement of an integrated sense of self. If a gay person has been assaulted, the experience may be internalized. Even if they have not been physically assaulted themselves, gay men and lesbians are sensitized to the ways in which members of the dominant culture, either in positions of authority or on its margins, feel they can make threats and do as they please. In addition, many individuals have had verbal assaults directed against them for being gay or lesbian. For some, experiences of verbal abuse can have traumatic consequences as well.

Coming Out and Meaning of Disclosure

Coming out, previously defined as the process of recognizing one's homosexuality and disclosing it to oneself and others, is an essential concept in understanding the lives of lesbians, gay men, and bisexual persons. Coming out, however, is not synonymous with identity development. One way of understanding the difference between these two concepts is to view coming out as one part of a process whose outcome may lead to forming a lesbian or gay identity. Behaviors associated with coming out reflect underlying cognitive and affective transformations that occur as part of identity development. Thus, coming out to oneself—a subjective experience of inner recognition that previously unacceptable feelings or desires are part of one's self—may facilitate disclosure of these feelings to others in the hope of finding social reinforcement and support. Some gay men and lesbians may only come out in certain spheres of their lives, for example, within friendship groups, but remain closeted in other areas, such as in the work setting or within families.

Coming out is an ongoing process because gay and lesbian persons must repeatedly choose whether to inform others of their identity. People are generally assumed to be heterosexual unless they declare themselves to be otherwise. Usually, disclosure of one's homosexuality involves either *telling* someone else that one is gay or lesbian or *displaying* some commonly accepted signal of a homosexual identity. Adoption of certain styles of dress or behavior, residing in neighborhoods where other gay men and lesbians live, entering into gay and lesbian communities, and involvement in relationships with persons of the same sex may serve to confirm the

identity of a gay or lesbian person. For some, these activities are signifiers of coming out, performed to show others that one is gay or lesbian. However, it should also be noted that many young people with stable, heterosexual identities often play with gender ambiguity. Nontraditional gender performances are not necessarily a sign of homosexuality, and interpreting them superficially may lead to confusion about or an inaccurate assessment of a person's actual sexual orientation.

Most empirical studies of coming out describe the ages at which gay men have reported events associated with coming out; fewer studies report on the process for lesbians. However, those studies show, in general, that women tend to have their first sexual experiences and to come out later than men. Younger lesbians may experience stronger prohibitions against coming out because of intense expectations of heterosexual dating for women. More recent studies have also suggested that coming out, much like the age of the first sexual experience, is occurring at progressively earlier ages among adolescents and that there is significant variation in the process based on a variety of factors in addition to sex, such as socioeconomic status, education, and ethnicity. The study of various cultures also demonstrates variation in the coming out process based on the degree of sexual restrictiveness of a society.

A. is a 25-year-old deeply religious man raised in conservative evangelical traditions. He entered treatment complaining of panic attacks and depression, worried that he might have AIDS after a recent, first homosexual experience several months earlier. Although his sexual behavior had not been high-risk, A. was certain he had contracted the virus and repeated negative HIV-tests did not assuage his anxiety. However, he gradually responded to a combination of antidepressant medication, anxiolytics, and twice-a-week psychotherapy.

In the first year of treatment, A.'s sexual identity was rarely discussed as treatment mainly focused on his anxieties about AIDS, his job, and interpersonal relationships. Gradually, therapy did address the question of his identity. A. declared he had had sexual relations with a college girlfriend, so how could he possibly be gay? He found it difficult to reconcile the idea of being gay with his religious beliefs. For months, he would go back and forth on the question of whether or not he was gay, and if he were, how could he possibly accept that and stay true to his religious beliefs. He would cite moral condemnations of homosexual behavior from scripture and modern religious authorities in an effort to dissuade himself (and the therapist) from accepting a gay identity.

A year into treatment, A. casually told the surprised therapist that he had already decided *he was gay* several months earlier. This private, intrapsychic shift in self-identification represented A.'s coming out to himself. He then offered some personal history he had previously omitted, acknowledging not only childhood sexual fantasies about men but that he had fantasized about sex with men while having sex with his college girlfriend.

Following this, a significant part of treatment focused on A.'s anxieties about coming out to his father, whom he described as a religious, conservative man who would *never* accept a gay son. A. had come out to his mother who said not to tell his father, with the warning that "it would kill him." Yet each time he had an argument with his father, A. would contemplate writing a "coming out letter." Asked how his father would react to such a letter, A. acknowledged his wish to use it as a weapon to hurt him. When the therapist asked whether A.'s wish to come out reflected his anger and hostility toward his father rather than a wish for openness and dialogue, he decided not to write the letter.

Eventually, A.'s mother told her husband that their son was gay. He then called A. to say nothing anyone told him would make him stop loving his son. This surprised the patient. The

therapist pointed out that if his father was not as unaccepting as he had imagined, perhaps he had projected his own lack of acceptance onto him. This interpretation allowed the patient to locate the feelings of intolerance within himself with which he struggled as he tried to reconcile a growing sense of himself as a gay man.

Coming out is a complicated process that involves both internal and external dimensions, occurs across a lifetime, and may vary in different domains of a person's life. Coming out can mean increased comfort with one's own feelings, and this is integral to social and psychological development. Greater ease in expressing one's feelings, both to oneself and to others, can lead to enrichment of work and relationships. However, even persons who "appear" to be gay or lesbian may have never verbally disclosed their sexual orientation, and others who provide no stereotypical indications of being gay or lesbian may have disclosed their identity to most of their family, friends, and coworkers. Some persons who are fully aware of their homosexuality may never disclose this fact to others. Furthermore, the time frame between awareness of same-sex attraction and self-labeling as gay or lesbian may range from less than a year to decades. In other words, each individual's coming out process is unique and multidimensional.

Race and Ethnicity

When considering a particular race or ethnic background of a gay, lesbian, or bisexual individual, consideration must be given to two additional issues: Attitudes within that particular racial or ethnic group toward sexuality and homosexuality, and the types of prejudice directed toward that racial or ethnic group by the larger society. These two factors may compound the experiences of stigma and discrimination of gay and lesbian racial and ethnic minority members. Theirs is the experience of being a *double minority*. For example, interactions between racism in the gay community and antihomosexual attitudes within the African-American community may lead to inadequate coping techniques and poor self-concepts among some African-American gay men. Finding support within either the African-American or gay and lesbian communities may be difficult, and integration of potentially conflicting identities based on race and sexual orientation may be an important need for persons from these groups. Similar issues may arise for persons from other racial and ethnic groups. Latino men and Latina women are often influenced by antigay and antilesbian religious attitudes as well as by traditional, heterosexist gender role and family expectations. Native peoples' concepts of gender, sexuality, and sexual orientation often diverge considerably from European notions of these constructs and may incorporate beliefs about spirituality as well. Because Asian American gay men and lesbians come from many differing cultural groups with diverse attitudes toward homosexuality and gay,

lesbian, and bisexual persons, it is impossible to generalize about their experiences. Thus, although membership in a particular group may assist in predicting family, religious, and social reactions to homosexuality, the diversity of backgrounds within many legally designated minority groups requires careful assessment of each individual's experience, level of cultural identity, and degree of acculturation.

The expression of sexual attraction and identity may be strongly influenced by racial and ethnic identity. These identities may be experienced as conflictual, resulting in such strategies as either maintaining complete secrecy about homoerotic feelings, adopting a bisexual rather than a gay or lesbian identity, or even feeling a need to choose between cultural and sexual orientation identities. For some persons, the conflict between identities may lead to profound confusion and alienation from both groups. The integration of potentially oppositional identities may be extremely difficult to achieve, and the difficulty of this struggle must be appreciated and understood by psychiatrists who work among multicultural gay and lesbian populations.

Religion

Individuals from different religious backgrounds may have had exposure to a wide range of attitudes toward homosexuality, and they may or may not maintain ties to the religion of their families of origin. At a social level, religion influences gay men and lesbians who belong to religious groups that present supportive, tolerant, or disapproving views of homosexuality. Religion also has an impact on all gay and lesbian persons because it contributes to shaping cultural and political attitudes toward homosexuality. Many gay men and lesbians attempt to avoid disapproving or condemnatory religious teachings about homosexuality by eschewing any contact with organized religion; others work within their religious groups to try to change traditional, unaccepting attitudes about homosexuality; and still others stay attached to, while perhaps feeling powerless to change, religions that reinforce their own antihomosexual attitudes. Many gay men and lesbians have actively sought to reconcile their sexual identity and spirituality within traditional religious groups; some have been successful in doing so. Others have moved to create new churches and spiritual settings that affirm their sexual identities and their relationships. For example, the Metropolitan Community Church is an international church that does specific outreach to lesbian, gay, bisexual, and transgender people.

Increasingly, religious organizations within American society have been forced to grapple with the meanings of homosexuality, both as a personal issue for their members and their families and as an issue of ongoing theological contention. A growing number of religious groups have moved toward positions that either tolerate or actively embrace gay men and

lesbians; some are ordaining gay priests, ministers, and rabbis. Pope Francis, when asked about his view of gay priests in 2013, responded, "If someone is gay and he searches for the Lord and has good will, who am I to judge?" Other, socially conservative religious organizations have vociferously reaffirmed their traditional opposition to open expressions of homosexuality and criticize the morality of individuals who adopt gay, lesbian, and bisexual identities; they are also critical of heterosexual religious authorities who accept those identities. Regardless of one's personal relationship to religion or one's involvement with a specific religious group, gay men and lesbians can be affected by religious teachings and beliefs about homosexuality.

Disapproving religious views of homosexuality often become internalized in lesbians and gay men and may have an effect on their psychosexual development, particularly inhibiting sexual expression and the development of relationships. Religious beliefs that disapprove of homosexuality constitute substantive, ongoing themes in psychotherapy with some gay and lesbian patients. Many of these themes are intertwined with issues regarding family relationships, particularly in gay men and lesbians who have had strict, fundamentalist, or orthodox religious upbringings. Gay men and lesbians who grow up in communities where social and religious activities are deeply intertwined may stop participating in religious activities and become estranged from their families. They may be deeply concerned about the moral implications of being gay or lesbian and about their alienation from their coreligionists. Spiritual development itself may be affected by an inability to reconcile the conflicting internal and external messages concerning the erotic and spiritual aspects of the self. Gay men and lesbians who are able to overcome disapproving, religiously based self-evaluations may nevertheless find it difficult to be comfortable with their sexual identity. Even when faced with denigrating judgments of their homosexuality, some individuals may not be able nor do they wish to disengage from their religions. Many others struggle to establish a connection to a different, more affirming spiritual community. Conflicts between religion and sexuality are an important focus in clinical work with religious lesbians and gay men.

Interaction of Sexual Orientation with Stage of Development

The 2011 Institute of Medicine report, *The Health of Lesbian, Gay, Bisexual and Transgender People,* endorses a life-course framework which acknowledges that events at each stage of life influence subsequent stages and recognizes that experiences are shaped by one's age cohort and historical context. In general, many of the psychosocial issues confronting gay men and lesbians during different developmental stages are the same as those for heterosexual persons. However, there are issues specific to being gay or lesbian that occur during childhood, adolescence, young

adulthood, midlife, and old age. Developmental tasks across the life stages do not take place at the same age for all persons. Nevertheless, most gay men and lesbians will have to confront certain tasks during each stage of development.

Childhood. Based on retrospective accounts from adults, children who grow up to be gay or lesbian often struggle with feelings of being different; for some of these children, there is an additional and perhaps related issue of atypical gender role behavior. Feeling different can lead to a sense of alienation that may result in social isolation. Boys who are not stereotypically masculine may be shunned, humiliated, derided, and bullied by their peers in school and other social settings, leading to future difficulties with self-esteem. Girls may be less subject to this process, at least until adolescence, as tomboy behaviors in girls are usually better tolerated than the so-called effeminacy in boys. Nevertheless, children with gender atypical behavior are often criticized or otherwise devalued within their own families and sometimes even by mental health professionals. Some parents, believing they might prevent their child from growing up gay, may try to shape developmental outcomes toward more typical gender role behaviors in their children. There is no research supporting the theory that homosexuality can be prevented by forcing children to behave in a gender typical fashion. Other parents may ignore or distance themselves from children who fail to conform to parental expectations of conventional, heterosexual development. These experiences can lead children who grow up to be gay to internalize rejecting reactions of others as they come to devalue aspects of their selves that they associate with atypical sexuality and gender. In cases where this process is accompanied by incidents of either overt discrimination or violence, the child may suffer adult impairment in the capacity for psychological, social, and work adaptation.

Adolescence. Adolescence is a particularly vulnerable time for most young people. It is during this period that they must establish a coherent identity, separate from their families of origin, and practice patterns of relating and working that will be further developed in early adulthood. Many gay and lesbian youth first become aware of their same-sex attraction during puberty when they begin to mature physically. Again, from retrospective accounts, gay and lesbian adults often recall a budding cognitive awareness of their same-sex attractions during adolescence. Most gay and lesbian youth are involved in heterosexual relationships during this time. However, some of them may enter the early stages of establishing a gay or lesbian identity, as they begin to have affectional and physical relationships with persons of the same sex. Some adolescents may go on to develop connections with community structures, such as gay-straight alliances, gay and lesbian youth groups, and adult gay and lesbian

community centers. These activities can be either inhibited or facilitated by the reactions of families, relatives, teachers, religious leaders, and peers.

A 2009 adolescent study by Ryan et al. found family rejection of their LGBT children increased rates of depression, substance abuse, suicidal ideation, and suicide attempts. This may be why many gay and lesbian adolescents, either consciously or unconsciously, delay some aspect of the development of their sexual identity until they are older. Often they obtain a greater degree of safety and support by going away to school, moving to an urban environment that allows more anonymity and acceptance, or creating a hometown network of friends who will be accepting of a gay or lesbian identity. Some adolescents may suppress any awareness of their same-sex attractions; others may defensively attempt to establish a heterosexual identity and relationships, while still others may consciously avoid any outward display of a gay or lesbian identity. Less often, but with increased frequency in urban settings, gay and lesbian youth may openly identify their sexual orientation at an early age, disclose their feelings to others, and complete significant aspects of gay and lesbian identity formation during adolescence. Even outside urban settings, this process has also been greatly facilitated by access to the Internet where gay and lesbian youth can more easily find one another. It remains to be seen whether future generations of gay and lesbian youth will follow in the footsteps of an older generation whose members often entered into unsatisfactory heterosexual relationships and eschewed intimate relationships with persons of the same sex out of fear of disclosure of being gay or lesbian.

Despite the rapid social changes taking place, there is still ongoing stigma associated with being gay or lesbian. Where there are few adult role models and inadequate resources for support, discrimination and violence are often directed against visible gay and lesbian youth in schools. In the absence of recognition and affirmation from parents and other adult figures in their lives, most pregay and prelesbian adolescents undoubtedly continue to experience significant difficulty in developing a stable gay or lesbian identity.

Young Adulthood. The period of transition from adolescence to adulthood is a time when many lesbians and gay men have their first opportunity to come out and to establish gay and lesbian identities independent of their families of origin. Gay men and lesbians in their 20s and 30s confront the same developmental tasks as heterosexual peers. These include developing a career, establishing a social identity, developing intimate relationships, and increasingly for some, raising children. However, for gay men and lesbians, compared to heterosexual cohorts, these tasks may be accomplished in different ways and at different times than for their heterosexual cohorts. Although heterosexual men and

women may be seriously dating, getting engaged, and getting married, closeted gay men and lesbians may be pretending to be heterosexuals as they date and marry members of the other sex as they struggle to avoid the stigma of being homosexual. Such hiding activities may cause a delay in the consolidation of an integrated sexual identity. Others may sublimate their sexual energies entirely and invest greater efforts in careers. The delays and omissions that follow the secrecy surrounding gay adolescence can have developmental consequences. For example, many young gay men and lesbians in their 20s and 30s may appear to behave, from a developmental perspective, more like adolescents. In some urban settings, such individuals form cliques, in-groups, and out-groups with a strong emphasis on style, conformity to standards of dress, a hierarchy of popularity based on looks, athleticism, and affability, as well as opportunities for experimentation with sex and drugs. For some individuals, this "delayed gay adolescence" provides opportunities denied to them when they actually were adolescents: To learn peer and social skills they may not have yet acquired and to experiment with public enactments of their sexual identities. The psychiatrist should understand that these relational patterns are understandable for those who, in adolescence, had to hide their sexual identities and behaviors.

During early adulthood, many heterosexual persons marry and have children. Gay men and lesbians may also form long-term relationships at this age. Many of them do, in fact, enter into primary relationships at some point. In previous generations, a significant proportion of young adult men and women with same-sex attractions adopted a heterosexual identity and entered into conventional marriages; many of these individuals also had children and raised families. However, more relational options are currently available to gay and lesbian people. These include living in gay communities with supportive friendship networks, forming primary same-sex relationships, and choosing to conceive or to adopt children. In recent years, thanks in part to reproductive technologies and the growing access to marriage equality, the number of same-sex couples having children has grown, due in part to a *gay baby or gayby boom* (see below).

Traditionally, gay men and lesbians chose careers that tended to ensure greater acceptance of their sexual identity. Although gay men and lesbians are employed, whether openly or in secret, in every type of profession, often they may feel more comfortable working in fields where they can be visible. Nonetheless, discrimination in hiring and promotion may be a problem for gay men and lesbians, while a lack of legal protection creates considerable work-related stress for many of them. At the time of this writing, 21 US states have passed legislation prohibiting employment discrimination on the basis of sexual orientation while the majority of states offer no such protections.

Midlife. For gay men and lesbians, the midlife decades of the 40s and 50s are characterized, as they are for heterosexual men and women, by consolidation of career identity, concerns about generativity, emerging problems with their own health and aging, and caring for aging parents. Gay and lesbian baby boomers who are currently in middle age and older represent the first generation to have openly come out in significant numbers.

Many lesbians and gay men begin raising children in midlife. These children may be from a previous heterosexual marriage, or they may be adopted or conceived by a single gay parent or two people in a committed same-sex relationship. Although attitudes appear to be rapidly changing, some of these parents may not receive the same support from their extended families as their heterosexual cohort, they may often rely much more heavily on friendship networks and paid child care.

Gay men and lesbians share the same midlife health concerns as their heterosexual cohorts. Some, however, may receive suboptimal health care if they fear disclosing their sexual orientation to health care personnel. For those living in areas with high rates of HIV infection, gay men may be confronted during midlife with issues of death and dying that are usually associated with later stages of the life cycle. Increasingly, as HIV treatments have become more effective, many gay men find themselves living with a chronic illness that creates its own ongoing uncertainties about one's health status.

The Elderly. Individuals 60 years or older have a wide range of experiences in relation to aging; some will develop serious illnesses and become incapacitated, while others will live healthy lives for several more decades. In addition, the commercial worship of youth tends to stigmatize elderly people. All older persons need to deal with the possibility of incapacitating illness and the approach of death. Older gay men and lesbians face additional concerns. For example, they may worry about how they will be treated in nursing homes and skilled nursing facilities. In addition, they must confront a widespread stereotype that older gay people are always lonely and unhappy. Studies, however, have shown older gay men and lesbians to be as well adjusted as their heterosexual counterparts. In addition, they may sometimes be better able than younger gay people to contend with the critical reactions to their sexual identities as they may no longer be concerned about disclosure at work or they may have come to accept their sexuality.

The best predictors of psychological adjustment in older gay men appear to be a stable sexual identity and integration into a gay community. Older gay men generally prefer to associate with their age peers, and many of them report being in relationships. A large majority of older gay men and lesbians report being happy, deny significant concerns about loneliness or

fear of death, and describe good integration into social networks.

Specific problems for older gay men and lesbians include a lack of adequate resources; a widespread reluctance to disclose one's homosexual orientation to health care and other providers; and an almost total invisibility regarding the needs of older gay and lesbian persons, whether single or in relationships, within institutions and agencies that provide care for the aged. Failing to recognize older gay men and lesbians is a function of both the denial of sexuality in older people in general and of an unwillingness to recognize relationships that are not legally sanctioned. Nevertheless, there are a small but growing number of retirement communities that either cater to elderly gay people or are "gay-friendly." Social service agencies like Services and Advocacy for GLBT Elders (SAGE, formerly Senior Action in a Gay Environment) have also emerged to provide services for elderly gay and lesbian individuals.

Families and Relationships

Traditionally, gay, lesbian, and bisexual persons usually grew up in families that did not acknowledge or affirm their children's emerging homosexual attractions or provide them with positive, adult gay and lesbian role models. This is gradually changing, as evidenced by gay adolescents coming out who are living at home or the development of organizations like PFLAG (formerly Parents, Families and Friends of Lesbians and Gays). Nevertheless, not everyone who comes out finds acceptance from his or her family of origin. In some cases, gay people in relationships do not come out to their families at all, particularly if familial nonacceptance poses a threat to the viability of the couple. Alternatively, many gay men and lesbians create what have been called *families of choice* or friendship networks. These provide interpersonal bonds, acceptance, adult friendships, and resources unobtainable from their families of origin. Many gay men and lesbians go on to form new family structures in their ongoing, romantic same-sex relationships, long-term cohabiting relationships, and single or coparent households with children. The findings from a growing body of research have begun to describe these couples and families.

Families of Choice. Many gay men and lesbians grow up in families that either ignore, disapprove, or openly denigrate their homosexual feelings and emerging identities; some report significant physical and emotional harm in families who condemn their homosexual orientation. The reactions of parents and other family members to a young person coming out, whether deliberately or inadvertently, can significantly influence the self-esteem and self-acceptance of gay and lesbian youth. It may irreparably shape the young person's expectations of acceptance or rejection by others outside the family. In the past, gay and lesbian youth could expect no more than tolerance within their families; few would

receive active acceptance or empathic support for the experiences and concerns that occur in relation to their emerging sexualities. Greater social acceptance of homosexuality appears to be changing some parents' attitudes toward an increasing acceptance of their gay and lesbian children, but this is still a distinctly minority position.

Consequently, once most gay and lesbian young people achieve sufficient separation from their families of origin, they often create networks of friends and acquaintances who provide them with the support and respect they either lacked or were denied within their families. For a single gay or lesbian person, families of choice are extremely important; they serve the functions of an extended family. They may help a person pass through developmental crises, are present at celebrations and rituals, and provide a source of comfort and intimacy. In a community that until recently has had little recognition of their primary affectional relationships, these friendships may even be more enduring than romantic relationships, and some even take on greater day-to-day importance than a lover or partner in the life of a particular gay man or lesbian. Understanding and accepting the importance of these friendship networks is essential for the clinician working with gay and lesbian individuals.

Couples. For heterosexual couples, family sanction of marriage provides support and stability. Until recent times, same-sex couples might not find support from their families and in many cases even found opposition. This is a changing social phenomenon.

At the time of this writing, same-sex marriage is legal in the United States, thirteen European countries, Argentina, Brazil, Canada, Colombia, Mexico, New Zealand, South Africa, and Uruguay. An increasing number of countries, and municipalities around the world now legally recognize same-sex relationships through civil unions (twenty-one European countries, Chile, Costa Rica, Ecuador, and in some areas of Australia and Taiwan) or domestic partnerships. The 2010 U.S. census reports a significant increase in the number of people who reported living in households where both members are of the same sex and over 100,000 couples report being married.

Gay and Lesbian Parents and Their Children. Many lesbians have conceived and many gay men have fathered children. In the past, this usually took place within a heterosexual marriage that ended when a closeted gay or lesbian partner came out. However, the 1990s saw the start of a *gay baby boom* in which single or coupled gay men and lesbians began raising children they had either adopted or conceived using reproductive technologies that included donor insemination (using known or unknown donors) or surrogacy. Consequently, in recent years there has been a greater visibility of families with gay and lesbian parents, and many

younger gay men and lesbians today seriously consider the option of having children. It is estimated that anywhere between one and fourteen million children live in households headed by same-sex couples.

There are numerous studies that describe the characteristics of gay and lesbian families and the impact on children of having a gay or lesbian parent. This research was initially undertaken in the late 1970s to evaluate antihomosexual psychological and judicial beliefs about the presumed harm done to children who had been raised by gay and lesbian parents. For example, it was not uncommon—and in some states it is still common—to award custody of children to a nongay parent in a divorce if the other parent comes out as gay or lesbian. These custody decisions were not based on results of empirical studies, but on the belief that gay parents were either by nature unfit or would influence the eventual sexual identity of their own children. However, an overwhelming body of research findings shows that having a lesbian or gay parent has no adverse psychological consequences for children. Nor does having a gay or lesbian parent influence the eventual sexual orientation of children when compared with the children of heterosexual parents. What studies do show is that children with two parents do better than children with one parent, regardless of the sexual orientation of the parents.

Gay and lesbian parents and their children also have to confront the fact of being different from heterosexual families and may experience stigma or antihomosexual bias. In addition, the ability of both same-sex parents to be legally recognized as parents is prohibited in many states. Consequently, the children in families with same-sex parents are often denied parental benefits from the partner with no legal standing, including clear inheritance rights, death benefits, or the right to be covered by a nonbiological parent's health insurance. In response to the needs of these children, in 2001 the American Academy of Pediatrics issued a position statement in support of legal adoption by a second same-sex parent (http://aappolicy.aappublications.org/cgi/content/full/pediatrics; 109/2/339). The APA issued a similar position statement in 2002.

CLINICAL APPROACHES

Although the psychiatric issues of lesbians, gay men, and bisexuals often resemble those of the general population, there are unique aspects to the clinical work with this patient population.

General Considerations

As they have encountered disapproval, discrimination, insensitivity to their concerns, and even abusive treatment within clinical settings, it is not uncommon for gay, lesbian, and bisexual patients, both in private and public settings, to be reluctant about disclosing their sexual orientation

identities to clinicians. They may be unwilling to enter treatment without some clear demonstration by the clinician of a nonjudgmental and accepting attitude. As with all patients, optimal evaluation and treatment requires a respectful and thorough evaluation.

Inclusive and Special Clinical Settings. Gay men and lesbians present for treatment in the same inpatient, outpatient, and residential clinical settings as other patients. In recent years, specialized clinical units and services have been developed, both in the United States and other countries, for working with gay, lesbian, and bisexual persons. These may include community health centers intended to serve the needs of the local lesbian and gay communities, not-for-profit and for-profit clinics, state and municipal government-supported clinics, and specialized services geared to particular problems like HIV infection or substance abuse. Whether services are provided in general or specialized settings, it is always to the patients' benefit when staff members are trained to deliver culturally sensitive care to gay men and lesbians.

The clinical setting and staff with whom the patient comes in contact are often important determinants in creating a climate of inclusiveness for gay, lesbian, and bisexual patients. Explicit and implicit communications of acceptance in these settings can ensure successful psychiatric treatment. For example, the language and behavior of psychiatrists and other staff should convey that they do not automatically presume that all individuals, couples, or heads of families are heterosexual (*heterosexism* or *heteronormativity*). Questions about relationships should be open-ended and nonjudgmental when inquiring about the gender of a patient's partner. It should not be assumed that all married persons are heterosexual or that married patients are sexually exclusive with their spouses. Questions regarding sexuality should distinguish between desire, behavior, and identity, with the recognition that, especially in younger patients, desires and behaviors may be more easily described than identity. Many gays and lesbians report a history of negative experiences in the health care system and may, therefore, enter the clinical setting with a sense of fear and trepidation. The presence of gay-friendly publications or even a rainbow flag in a waiting area and self-reporting forms for patients to fill out that neutrally inquire about diverse sexual orientations and family arrangements can reassure a gay, lesbian, or bisexual person—as well as those uncertain about their sexual orientation—of acceptance in this clinical setting and that they can talk openly about their concerns. Since 2011, The Joint Commission (TJC) has required that health care organizations have an LGBT nondiscrimination statement for accreditation to receive Medicare reimbursement. TJC encourages LGBT-inclusiveness in its publication, *Advancing Effective Communication, Cultural Competence, and Patient- and Family-Centered Care for the Lesbian,*

Gay, Bisexual, and Transgender (LGBT) Community: A Field Guide which also describes how to create an LGBT-inclusive clinical environment.

Another important strategy in clinical work with gay and lesbian patients is to routinely include, when appropriate, references to significant friends and members of nontraditional families during evaluation and treatment. When necessary, same-sex partners and children in these families should be involved in the care of gay and lesbian patients. For example, it may be important to obtain corroborative information from partners, to identify potential sources of a patient's problems within the context of the couple or family, or to include a partner in maintaining compliance. In addition, issues of confidentiality and decision making in working with heterosexual families should also apply to working with gay and lesbian families.

In general, physicians should be conversant with advance directives such as the durable health care power of attorney and the appointment of health care proxies within the states in which they are licensed to practice. These issues may be particularly relevant when working with sexual minorities and alternative families. Historically, same-sex couples did not have marriage rights and found other ways to protect their relationships and families of choice. Even today, some older couples have not availed themselves of the right to marry. In cases where one of the partners becomes incapacitated, courts may grant important medical decision-making rights and privileges to biological family members and exclude the wishes of long-term partners.

The Sexual Orientation of the Psychiatrist. The issues of the psychiatrist's sexual orientation and of disclosing this orientation to the patient are complex. The heterosexist assumption that everyone is heterosexual usually means that a psychiatrist's sexual orientation does not come up in treating most patients. However, this assumption does not hold when treating a gay, lesbian, or bisexual patient. A heterosexual psychiatrist trained to be nondisclosing may refuse to answer questions about his or her sexual orientation and turn the question back on the patient. In general, this might be a technical error with many gay, lesbian, and bisexual patients. Not revealing could be interpreted by a patient as meaning that the psychiatrist is not entirely comfortable speaking about him- or herself or about homosexuality in general, that the psychiatrist might be in the closet, or some other meaning.

In the cases of lesbian or gay patients working with psychiatrists who are openly gay or lesbian, there are several possible benefits and risks to self-disclosure. The benefits of working with a psychiatrist of the same sexual orientation identity may include an easier rapport; greater efficiency resulting from shared knowledge about experiences associated with being lesbian or gay; if the therapist is older, an ability to share knowledge about

the gay community with a younger patient; and affirmative role modeling if the therapist has established a secure gay or lesbian identity. Possible risks for such therapeutic dyads include unwarranted assumptions that the patient and psychiatrist share experiences or personal characteristics based on sexual orientation identity alone; reducing the focus of treatment to sexual orientation identity issues alone; transferential or countertransferential overidentification of the therapist or patient with each other; or a shared, unconscious collusion to avoid discussing painful events or affects associated with gay or lesbian development.

How a gay or lesbian psychiatrist's sexual orientation identity is disclosed to a patient may vary. Sometimes the patient has requested to work with an openly gay or lesbian therapist and has obtained this knowledge before beginning treatment. In other cases, the patient learns this information after treatment has begun, either as the result of self-disclosure by the psychiatrist or through some outside source. When self-disclosure is involved, the timing of the disclosure may have a significant impact on facilitating or inhibiting the work in therapy. In some cases, it may have no effect at all. It is useful for the gay or lesbian psychiatrist to be able to identify those individuals and the clinical situations in which working with an openly gay or lesbian therapist might be most helpful.

In general, the psychiatrist's sexual orientation identity should not be the most significant factor in determining clinical outcomes. Any therapist working with gay men and lesbians needs to maintain conscious awareness of his or her own antigay attitudes and heterosexist assumptions; possess a reasonable knowledge base about LGB issues; apply gay-sensitive and gay-affirmative approaches to psychotherapy; and obtain adequate supervision, peer supervision, or outside consultation when necessary. Having an open gay identity is no substitute for good training, and gay and lesbian mental health professionals should constantly strive to improve their knowledge base and clinical skills.

Sexual orientation may be far less significant than other personal characteristics in determining the effectiveness of a psychiatrist in working with gay, lesbian, and bisexual patients. A gay or lesbian therapist who has not fully come out or established a comfortable identity may have trouble working with patients struggling with similar concerns. Regardless of sexual orientation identity, a therapist who is uncomfortable talking about sexual problems will have more trouble being helpful to gay and lesbian patients with such concerns. Ultimately, the degree of empathy experienced and expressed by the therapist for the patient will, in most cases, be an important factor in determining the effectiveness of treatment.

Specific Clinical Issues

There are clinical concerns that are unique to lesbians, gay men, and bisexuals. The presentation of other concerns that do not directly relate to

sexual orientation identity may also be shaped by the fact that a person is gay, lesbian, or bisexual, and for some persons being a member of a sexual minority group may interfere with, and even prevent, interactions with the health care system.

Societal and Domestic Violence. The problem of *antigay* or *antilesbian violence* is widespread. In 1990, the US government passed a law ordering a study of hate crimes, including attacks on gay men and lesbians. Because of concerns about antigay violence at the time, the law's passage was the first time the federal government made any affirmative mention of sexual orientation.

One of the most common forms of violence is *gay-bashing:* Gangs of young men descend on neighborhoods or gathering places where gay people meet. Armed with bats, clubs, or other weapons, they attack anyone they believe to be gay. The most sensational case of antigay violence in the United States was the 1998 death of Matthew Shepard, a young man from Wyoming, robbed and killed by two other young men because they perceived him to be gay. The trial of his killers, and the increasing publicity surrounding gay-bashing incidents in general, has provoked wider discussion of previously unspoken cultural beliefs. For example, some perpetrators of violence have put forward a *gay panic defense,* arguing that violence is an excusable response by a heterosexual man to an unwanted sexual advance from a gay man. To date, no court has accepted this defense and in 2014, California became the first US state to ban the use of this defense (as well as the *trans panic defense*) in murder trials.

Gay men and lesbians who have been the objects of violence, or sexual or domestic abuse, may have difficulties in reporting their experiences and seeking help; they often fear disclosing their sexual identities or receiving judgmental responses—and even harassment—from police or social service workers. In other words, antigay violence is presumed to be an underreported phenomenon because its victims are often reluctant to come forward and face reprisals from the very authorities to whom such incidents are reported. It has also been argued that antigay violence is the inevitable outgrowth of the unexamined beliefs associated with homophobia, heterosexism, and moral condemnations of homosexuality. Thus, those who marginalize homosexuality may be either directly or indirectly responsible for the brutality aimed at gay men and lesbians. Inevitably, gay men and lesbians may have legitimate concerns about potential physical attacks. The problem of violence directed against lesbians and gay men is widespread and well documented and ranges from verbal abuse to homicide. Violence against gay people may be committed by strangers, by family members, by gangs, by fellow soldiers within the military, and sometimes by authority figures, such as the police.

Although some therapists believe that childhood sexual abuse causes

homosexuality, studies show no causal connection between the phenomenon and the adult development of a gay, lesbian, or bisexual identity. Studies show that sexual abuse of gay men and lesbians during childhood appears to occur at about the same rate as in the general population. Studies of domestic violence and abuse in lesbian and gay relationships demonstrate that physical abuse occurs at about the same rate as in heterosexual relationships.

The impact of violence on gay men and lesbians, as on other victims of violence, includes posttraumatic stress disorder, depression, and a range of other psychological and emotional problems. These include feelings of inadequacy, embarrassment, or shame. In addition, there may be specific symptoms associated with being part of a stigmatized minority group. For example, the victims of antigay or antilesbian violence may hold themselves responsible and believe that their homosexuality caused the assault. The experience of violence can lead an individual to deny a gay or lesbian identity or to avoid other gay and lesbian individuals. Depending on the circumstances, the treating psychiatrist should be prepared to respond to the immediate crisis following an incident of abuse or violence; to the potential long-term effects of these events on the individual; and to the ongoing concerns of all gay men and lesbians about potential violence being directed against them.

The psychiatrist can be helpful by recognizing the influence of exposure to antigay violence on the personalities and experiences of gay, lesbian, and bisexual persons and carefully identifying the intrapsychic and interpersonal wounds they may produce. It is also reasonable for psychiatrists and other clinicians to acknowledge the everyday dangers that gay men and lesbians face or may have faced. A psychiatrist can be most helpful by being sensitive to patient anxieties regarding antihomosexual attitudes, as gay patients can be retraumatized when clinicians deny or minimize the extent to which the phenomenon colors their lives and affects their self-esteem. It should be noted that in this culture, antihomosexual attitudes are ubiquitous, and no one is free of them, not even psychiatrists. This obliges clinicians to understand and be aware of the ways in which these attitudes may have affected the development of their own identities and belief systems. Such awareness can help a psychiatrist to come to a greater empathic appreciation of the impact of antihomosexual attitudes on the gay, lesbian, or bisexual patient they are treating.

Suicide. Determining the rate of suicide attempts and completed suicide among LGB persons is difficult. Factors complicating this research include the absence of information identifying sexual orientation on death certificates, the reluctance of many individuals distressed about their same-sex feelings to discuss them with mental health professionals, an inadequate assessment of sexual orientation in most health care settings,

and the inability to obtain representative samples. As noted in the aforementioned report of the Institute of Medicine, lack of disclosure of homosexuality poses a greater problem for older compared to more recent studies as social acceptance of homosexuality has increased.

In spite of methodological difficulties, recent studies consistently suggest that gay, lesbian, or bisexual youth have an increased risk for attempted suicide. With few exceptions, the increased rate of suicidality among LGB youth in comparison with heterosexual youth is consistent across age groups (i.e., middle school, high school, and young adult populations), gender, race/ethnicity, and differing definitions of sexual orientation (i.e., same-sex attraction, self-identification, and behavior). Suicide rates among gay, lesbian, and bisexual youth may vary on the basis of other risk factors, such as mental illness, substance abuse, rejection by family, and perhaps, degree of gender role nonconformity. For all age groups, an increased risk of suicide may accompany struggles with internalized homophobia, problems with coming out, and difficulty in establishing a firm gay, lesbian, or bisexual identity. At least one study also found suicidal ideation and behavior to vary in sexual minorities as a function of race and ethnicity. Compared to whites, blacks and Latinos were more likely to report a history of a suicide attempt; however, they did not evidence increased rates of mental illness. Most of the suicide attempts came at an early age and were interpreted as probably coinciding with the coming out period and likely to have been associated with increased social disapprobation within the black and Latino communities.

Determination of the exact numbers of gay, lesbian, bisexual, or uncertain people who attempt or complete suicide may be less important than understanding the characteristics that contribute to a risk for suicide. For gay or lesbian youth, enforced invisibility and a lack of adequate resources to address their health and mental health needs can prevent adequate assessment of their concerns or interfere with developing specialized programs for prevention and treatment. Unlike other minority youth, gay and lesbian youth may not typically share their minority status with either parent, and, therefore, may not readily turn to their parents for support. In the United States, this problem has been exacerbated by a socially conservative political climate in which efforts to identify the mental health needs of gay, lesbian, and bisexual individuals are sometimes labeled as an “endorsement” of homosexuality. This has hindered public funding and research efforts to study suicide among gay, lesbian, and bisexual young people as well as efforts to provide them with social support.

In general, any young person in distress should be carefully evaluated for concerns about sexuality and sexual identity. The psychiatrist should take pains to ensure that assessment and treatment do not place young people at further risk of suicide if and when they disclose their sexual

orientation. For example, premature coming out, sometimes precipitated by a conversation with a mental health professional, can increase the possibility of rejection by peers and family. Moreover, all physicians should be aware that many parents are hostile to sexual minorities and should assess the risk that the family's attitudes may pose to the patient. These risks may include verbal and physical abuse, subjection to damaging sexual orientation conversion efforts, rejection of the child, and even homelessness. Appropriate steps must be taken if the parents cannot provide a safe environment for their child. The potential risk of suicide for persons of any age who are struggling with coming out and establishing a gay, lesbian, or bisexual identity should be recognized. Clinical services appropriate for managing suicidal feelings and behaviors and for providing support to suicidal persons should be available as needed.

Alcohol and Substance Abuse. Research findings to date do not reveal the exact prevalence of alcohol and substance abuse disorders in the gay and lesbian community. Controlled studies comparing groups of heterosexuals with lesbians and gay men are absent. Nevertheless, there is a general impression that alcohol and other substance use and abuse, including tobacco, appear to be greater among some segments of the gay and lesbian communities. Some attribute this increased usage to the psychosocial stressors associated with being lesbian, gay, and bisexual. In many cities, bars and clubs where one can enter freely and anonymously are often the first and only point of entry for some individuals into the wider gay and lesbian community.

Recreational or "club drugs" have become increasingly popular among some segments of the gay and lesbian community. Among these are MDMA (3,4-methylenedioxy-*N*-methylamphetamine), colloquially referred to as Ecstasy, which can induce depression and panic attacks in those who use it. Ketamine, also known as K or Special K, is a dissociative anesthetic that induces feelings of unreality and may cause catatonia. Gamma-hydroxybutyrate, known as G or GHB, is a naturally occurring biological substance whose effects mimic alcohol. In modest doses it can cause sleep, coma, and in overdose, some cases have been fatal. The psychiatrist working with gay men and lesbians involved in the club scene or circuit parties should be aware of these particular substances, the lifestyle choices associated with their use, and their potential risks, including potential interactions with prescribed medications.

Mood-altering substances are often used to disinhibit prohibitions on sexual expression, and as a result many lesbians and gay men begin coming out using drugs or alcohol. For some people, their gay or lesbian identities are strongly associated with alcohol or substance abuse. Awareness of this linkage is useful in planning adequate approaches when treating substance abuse among gay men, lesbians, and bisexuals. There are numerous

specialized programs, both inpatient and outpatient, that simultaneously address substance use problems while respecting the sexual identities of gay, lesbian, and bisexual persons. There are also 12-step and other self-help programs that provide special groups for gay men and lesbians and that allow for a more comfortable and open discussion of both sexual identity and substance use issues.

The psychiatrist should be prepared to identify a potential overlap between problems with acceptance of one's sexual orientation identity and substance abuse. Patients with addiction problems, gay, lesbian, and bisexual or otherwise, sometimes show signs of either a poorly integrated or a self-loathing identity. Such persons may exhibit denial, fear, anger, guilt, helplessness, hopelessness, dishonesty, low self-esteem, self-hatred, and social isolation and alienation. Assessment of these traits and an awareness of how addiction and internalized homophobia may reinforce each other are helpful in providing adequate treatment. Furthermore, it may sometimes be difficult to treat substance abuse problems in gay, lesbian, and bisexual patients without addressing concomitant sexual identity problems. As a result, substance abuse programs that are both gay sensitive and gay affirming work best for treating substance abuse disorders among gay men and lesbians. Prevention or reduction of substance abuse among this population must include the development of settings for socializing without alcohol or drugs.

Talking Sex. As psychiatrists and other health professionals are often uncomfortable talking about sexuality in general, dealing with the specific sexual behaviors and concerns of gay men and lesbians may be particularly difficult. Problems may result from either a psychiatrist's unfamiliarity with or disapproval of same-sex behaviors or both. Given the inadequate training in sex therapy in most training programs, it may be difficult for the psychiatrist to discuss the sexual concerns of gay and lesbian patients in an objective and comprehensive manner, to be aware of special diagnostic issues, and to approaches to sex therapy with these men and women. These concerns might include difficulties performing anal intercourse in a committed couple or concerns about HIV infection among gay men and orgasmic difficulties among lesbians. Some gay men with HIV develop low serum testosterone and may seek treatment for loss of sexual drive. Some gay men may struggle with issues of sexual compulsivity, and bisexual men and women may have difficulty mastering safer sex practices.

Patterns of same-sex sexuality have been shown to be different from those in heterosexual couples, with female couples having decreased sexual interaction and male couples initially having increased amounts of sexual interaction in comparison with heterosexual couples. Typologies of male couples have sometimes described them in terms of their degree of sexual exclusivity and openness, and some male couples present with problems

related to sexual exclusivity. In contrast, female couples present more often with problems related to decreased sexual desire. There also appear to be generational differences in openness of relationships, which may be attributable to the growing accessibility to legal protections for same-sex couples via marriage or civil unions.

The psychiatrist working with the sexual problems of gay and lesbian individuals and couples should consider organic, developmental, interpersonal, and social causes. Assessment and treatment can adapt traditional, heterosexually oriented paradigms of sexuality to include consideration of issues related to gay and lesbian development and to the dynamics of same-sex relationships. Special attention should be given to cultural factors that may contribute to sexual problems among gay men and lesbians. These include a lack of readily available and accurate information about same-sex practices as well as the inevitable internalization of judgmental societal attitudes about homosexuality.

Medical Illness. Although often presenting with the same medical problems as the general population, the recent report by the Institute of Medicine found that lesbians and gay men experience unique health disparities as well as barriers to equitable health care. One such barrier is psychosocial: Gay men and lesbians often have to deal with judgmental attitudes from health care personnel which may lead some to avoid or delay seeking care. This may not be as great a problem in some urban settings with large gay and lesbian communities where clinicians, either in private practice or public settings, have experience in caring for lesbian and gay patients. However, even in large urban settings, not all gay and lesbian patients have access to health care personnel who are sensitive and knowledgeable about their concerns or their increased risk for some diseases.

There are no gynecological problems that are unique among lesbians as opposed to among heterosexual women. Lesbians, however, may be at greater risk for undetected disease as studies show they may undergo fewer pelvic examinations and Papanicolaou (Pap) smears and mammograms. Sexually transmitted diseases are rare among women who are sexually active only with other women, and there are few cases of HIV infection in these women. The risk for some types of cancer in women may be affected by factors such as history of sexual intercourse with men, number of pregnancies, age at first giving birth, and breastfeeding. Women who have not had children and have not breast-fed, have not used oral contraceptives, and are older when they first give birth are at increased risk of breast cancer. Historically, lesbians have had a lower number of pregnancies than heterosexual women. However, an increasing number of lesbians are having children using reproductive technologies including anonymous or known donor insemination.

In general, when compared to heterosexual men, gay men who engage in unsafe sexual practices have increased rates of sexually transmitted diseases and HIV infection. Since the 1980s, the influence of HIV infection and AIDS on individual lives and the communities of gay men has been enormous and has affected almost every aspect of the personal and public experience of being gay. Public health campaigns to reduce infection among gay male populations appeared to work for some time in reducing the rate of infection. However, as younger gay men enter the community, educational strategies tailored to older gay men may not reach these new arrivals, and public health responses to their needs have not always been effective. For example, some young gay men may believe that only older gay men are infected with HIV and that they are safe from HIV infection if they only have sex with men their own age. In recent years, there has been a growing phenomenon among some gay men of *barebacking* or *unprotected anal intercourse (UAI)*. Ignoring public health warnings about HIV has been attributed to different causes, depending on the individual: Lack of information, taking calculated risks about who is and is not infected, "AIDS fatigue," and an unwillingness to continue using condoms. Public health measures in this area remain an issue of great concern.

HIV infection—either being infected or knowing someone with the virus—is a potential problem in the life of any gay man who enters the health care system. Health issues related to HIV include testing for infection; ensuring early treatment intervention in infected individuals; preventing further HIV infection; responding to reasonable and irrational fears about contracting the disease; dealing with the chronic stress resulting from either multiple losses, illness, or both; and having to take care of someone else who is infected. For many gay men, an association of HIV infection with one's sexual orientation identity can exacerbate one's own internalized homophobia.

In July 2012, the Food and Drug Administration (FDA) approved a combination of emtricitabine and tenofovir (trade name Truvada) for pre-exposure prophylaxis (PrEP) against HIV infection for individuals at high risk. While the social and ethical complexities of this intervention are currently debated, the Centers for Disease Control and the WHO have published guidelines for PrEP in at-risk populations. These guidelines include the use of condoms in addition to PrEP as some men in clinical studies seroconverted despite apparent adherence with the drug regimen. The message to always use condoms has not proved effective in ending the AIDS epidemic; however, programs including PrEP have the potential to do so. In 2014, New York's governor issued a plan to end the epidemic in his state by 2020. The plan includes PrEP, universal HIV testing and connection to care for people living with HIV since modern treatments have the capacity to lower viral loads to the point that the risk of transmitting the virus is low. Data from the National HIV Surveillance

System suggest that treatment as prevention is being grossly underutilized. Perhaps half of the MSM who are living with HIV are not receiving antiretroviral therapy. The goal of universal HIV testing and connecting those who test positive with treatment is clearly a public health priority. Integrating antiretroviral therapy for HIV-positive partners and PrEP for HIV-negative partners nearly eliminates the risk of HIV transmission in serodiscordant couples. A short antiretroviral regimen initiated within 72 hours of a known exposure to HIV can also prevent seroconversion.

There are generational differences. Gay men under 30 years may be untouched by HIV, having never known anyone infected. Those gay men who were out in the 1980s and 1990s have been affected by the epidemic. Among other things, the psychiatrist can play an important role in differentiating functional and organic symptoms that occur in HIV-infected patients, for example, diagnosing the clinical depression that may occur as a result of low testosterone levels. Psychiatrists may also be called on to treat any psychiatric sequelae of the disease, including depression, bipolar disorder, anxiety disorders, psychosis, and dementia.

Recent advances in treating HIV infection have changed its course from an almost universally fatal diagnosis to a chronic medical condition. However, there are some small, inconclusive studies showing aging-related health problems among AIDS patients such as osteoporosis and chronic obstructive pulmonary disease.

Physical illness in gay men and lesbians may interact with experiences associated with their sexual identity. Lifestyle factors, family structures, and social networks can determine the extent and type of support available. The stress of being ill may evoke psychological reactions that affect capacities for intimacy, affection, and sexuality.

Psychotherapy

Gay and Lesbian Affirmative Psychotherapies. The 1970s saw the introduction of literature on *gay and lesbian affirmative psychotherapy*. This approach, a response to the historical pathologizing of homosexuality, was described by Malyon as a "theoretical disposition [that] regards homosexuality as a non-pathological human potential. The goals of gay-affirmative psychotherapy are similar to those of most traditional approaches to psychological treatment and include both conflict resolution and self-actualization." Gay and lesbian affirmative psychotherapy recognizes the centrality of sexual attraction, behaviors, and identity in a person's life but also appreciates that sexual orientation is only one aspect of personality and, depending on the patient's needs and goals, may or may not be a central focus of therapeutic work.

Acceptance, acknowledgment, and affirmation of gay, lesbian, and bisexual persons are essential elements in psychotherapeutic work with gay men and lesbians. These therapeutic stances reflect the therapist's regard

and empathy for the patient's sexual feelings and for their same-sex relationships. The therapist's attitudes are meant to offer the patient a contrast to the social denial and denigration that the patient may have experienced in the past. The psychiatrist working with gay and lesbian patients should maintain a respectful position that will both help to contain the struggles of the patient and at the same time serve to convey a recognition of the ongoing adverse impact of social prejudice against gay and lesbian persons. The clinician's lack of awareness about or inability to explicitly acknowledge the effects of antihomosexual interpersonal and social forces working against the expression of same-sex feelings and desire may suggest to the patient that the psychiatrist agrees with these attitudes.

As the sexual, romantic, and affectionate expressions of gay and lesbian patients have often either ignored, ridiculed, or rejected, for many gay men and lesbians, speaking about them for the first time with a therapist may represent a unique opportunity. Feeling listened to without judgment, having one's feelings acknowledged as real by a person in authority, and having them affirmed as commonly felt and normal are powerful affirmations. The therapist's ability to tolerate, accept, and validate the patient's feelings allows the patient to express his or her own antihomosexual feelings and doubts, including fears about the critical reactions of others, disapproving feelings about being gay or lesbian, and regret over losses the patient may have sustained after coming out as gay or lesbian.

Contemporary Psychoanalytic Approaches to Gay and Lesbian Patients. Until the late 20th century, psychoanalytic approaches to gay men and lesbians focused on homosexuality first as a developmental arrest and then as a "neurotic" symptom. Starting in the late 1970s, newer psychoanalytic approaches emerged. Rather than trying to get homosexual patients to conform to heterosexual norms, these newer approaches see the therapist as an agent of the gay patient.

Analytic work with gay and lesbian patients has drawn attention to aspects of the therapy process sometimes overlooked in doing psychotherapy with nongay patients. As a result, this approach offers insights into some general principles of psychoanalysis, as well as basic psychoanalytic beliefs and practices including the nature of the psychotherapeutic frame; the values and risks of therapist self-disclosure; the limitations of psychoanalytic data to support theories of etiology; how adherence to theoretical preconceptions will restrict or inhibit a therapist's clinical listening; experience-near versus experience-distant responsiveness from the therapist; the therapist's embeddedness in cultural preconceptions and how they have an impact upon treatment; the role of the patient's and the therapist's subcultural identities and how they co-construct narratives in treatment; and the meanings and therapeutic uses

of countertransference. Developing one's own therapeutic stance in treating gay and lesbian patients depends on specialized training in psychotherapy, continuing education, the therapist's own personal analysis, and ongoing self-analysis. Furthermore, the therapist's stance does not just entail what one knows; it also includes what the therapist does not know, and requires an ability to allow a dynamic interplay of knowing and not knowing in the patient, in the therapist, and in the transitional space between them.

An analytically oriented therapist aims to create a therapeutic holding environment, which is a space where all of a patient's feelings and ideas are allowed to emerge. As in gay and lesbian affirmative psychotherapy, the analyst's job is to remain open to hearing the patient's complaints, sexual or otherwise. In the atmosphere of the holding environment, meanings are discovered, created, and deciphered. All patients, not just gay ones, can benefit from a therapeutic holding environment based on respectful principles. Respect for the patient is essential. But by itself, however, it is not enough. The subject of homosexuality often evokes uncomfortable feelings and denigrated meanings. When these emerge in treatment, they certainly need to be tolerated and respected by the therapist. In addition, a prerequisite for doing psychotherapy with gay men and lesbians is for therapists themselves to be able to accept their patients' homosexuality as a normal variation of human sexuality and to value and respect same-sex feelings and behaviors as well. It should again be noted that for some gay men and lesbians, being treated with such respect by a therapist is a novel experience.

In addition to respect, a contemporary psychoanalytic approach to gay and lesbian patients focuses on the meanings, rather than the origins, of human sexuality. The therapist who searches for the causes of homosexuality will ultimately mistake its meanings to the patient for something else. Alternatively, this approach shifts its focus to the affective meanings of the language of the patient, as well as those of the therapist, and to their transferential and countertransferential implications. From this perspective, it is important to understand there are consequences to a therapist's taking sides in a gay or lesbian patient's interpersonal struggles with family and friends. Ideally, everyone a patient discusses should be treated respectfully, including those whom the patient expresses disapproval. This therapeutic stance allows for a wider exploration of a patient's feelings and identifications with those who accept the patient's sexual identity and those who do not. Patients may try to resolve inner conflicts about being gay or lesbian by selectively inattending to their own antihomosexual identifications. This is sometimes seen in gay men and lesbians who preach rigid doctrine to themselves in order to affirm their homosexuality. Unable to tolerate conflictual feelings about homosexuality, these individuals sound as if they are trying to convince themselves that "It is okay to be gay." However, this strategy only reverses the feelings and

identifications of a former state of mind. In the subjectivity of a closeted identity, heterosexuality is idealized and homosexuality dissociated. After coming out, being gay can become idealized, while disapproving feelings are denied. For example, a patient's internalized religious beliefs may not be limited to condemnations of homosexuality. They also include moral and ethical beliefs that are integral to the adult self. Patients cannot simply rid themselves of self-condemnatory attitudes without severing their attachments to other important identifications. The psychotherapeutic challenge is to integrate patients' adult feelings and understanding of sexuality, ethics, and morality with their internalized childhood beliefs. Therapeutic holding entails being able to contain all those aspects of the patient so that greater integration of all the patient's feelings can take place.

Sexual Orientation Conversion Efforts (SOCE). In recent years, there has been an increase in public attention paid to organizations and individuals who make the claim that sexual orientation is mutable. However, most of the arguments for the position that "change is possible" are not taking place in scientific journals but on religious websites, in the popular press, and in the media. It appears that the goal of these publicity efforts has not been to convince the mental health professions that homosexuality can be changed, but to persuade the public that it can.

The motive for generating such publicity is not surprising. There are, after all, some individuals with same-sex feelings or attractions who do not wish a homosexual identification. These individuals, often for religious reasons, although not always, are unable or unwilling to accept their homoerotic feelings as normal, natural, or moral. Toward that end they may seek support from family members and friends, coreligionists, and professionals who also disapprove of homosexuality. Such individuals may seek clinicians who claim to be able to change through *sexual orientation conversion efforts (SOCE)*. These can include approaches that go by the name of "reparative therapy" or "sexual reorientation therapy" or "sexual conversion therapy." In addition, others unhappy about their homosexuality may seek religious fellowship among the so-called ex-gay groups. According to a 2000 position taken by the APA's Commission on Psychotherapy by Psychiatrists:

> To date, there are no scientifically rigorous outcome studies to determine either the actual efficacy or harm of "reparative treatments." There is sparse scientific data about selection criteria, risks versus benefits of the treatment, and long-term outcomes of "reparative" therapies. The literature consists of anecdotal reports of individuals who have claimed to change, people who claim that attempts to change were harmful to them, and others who claimed to have changed and then later recanted those claims.
>
> With little data about patients, it is [still] possible to evaluate the theories which rationalize the conduct of "reparative" or [sexual] conversion therapies. . . . [These theories] are at odds with the scientific position of the American Psychiatric Association which has maintained, since 1973, that homosexuality per se, is not a mental disorder.

> . . . Until there is such research available . . . ethical practitioners [should] refrain from attempts to change individuals' sexual orientation, keeping in mind the medical dictum to, first, do no harm.

In calling for a moratorium on sexual orientation conversion efforts, organized psychiatry has acted to protect patients who may be harmed by those procedures. Anecdotal reports indicate that unsuccessful conversion therapies may cause depression, intimacy avoidance, and sexual dysfunction. In 2009, the American Psychological Association issued a report of its *Task Force on the Appropriate Therapeutic Response to Sexual Orientation*. The Task Force conducted a systematic review of the peer-reviewed journal literature on SOCE and concluded that efforts to change sexual orientation are unlikely to be successful and involve some risk of harm. Both California and New Jersey have passed laws banning licensed clinicians from the practice of SOCE with minors and similar bans have been introduced in the legislatures of other states. In 2013, one of the largest ex-gay organizations in the world, Exodus International, closed operations, at the same time apologizing for any harm their message of "cure" may have inflicted upon the gay community.

Nevertheless, both APAs recognize that some individuals may still wish to change their sexual orientation for religious or other reasons. Clinical exploration of the irrational aspects of patients' internalized antihomosexual attitudes does not always lead to acceptance of one's homosexuality. Thus, even if homosexuality per se is not a mental disorder, psychiatry and other mental health professions may wish to find a way to help individuals who wish to rid themselves of same-sex feelings. However, this does not and should not require redefining homosexuality as an illness. If psychiatry and other mental health professions are to play a role in assisting such patients, perhaps the field of plastic surgery might serve as a model: Cosmetic plastic surgeons routinely treat nonpathological but socially stigmatized physical conditions.

Plastic surgeons, however, employ standards of care that are not matched by those of sexual conversion therapists, who have yet to develop anything approaching scientific and clinically sound selection criteria for patients. It also is not certain that contemporary conversion therapists can develop more exacting standards of care. For example, after the mental health mainstream endorsed a normal variant model and social acceptance of homosexuality increased, the professional training, credentials, and standing of sexual conversion therapists, in general, have diminished. For example, one study showed that conversion therapists regularly violated professional codes of conduct regarding informed consent, confidentiality, pretermination counseling, and provision of referrals after treatment failure. It is a fact that a field dominated by medically trained practitioners half a century ago is now primarily the province of less rigorously trained and sparsely regulated clinicians, pastoral counselors, and self-help groups.

In addition, as many conversion therapists work primarily within a faith-healing model, it remains to be seen whether they can successfully develop scientific and clinical selection criteria to distinguish individuals who have a reasonable prospect of changing their sexual orientation from those who may be harmed by sexual orientation conversion efforts. Until this can be accomplished, the dictum to "first, do no harm" should be kept in mind for those concerned about the well-being and the quality of care for all patients, regardless of their eventual sexual orientation.

ETHICAL AND LEGAL ISSUES

Principle I of *The Principles of Medical Ethics: With Annotations Especially Applicable to Psychiatry* states that "A physician shall be dedicated to providing competent medical service with compassion and respect for human dignity." The annotation goes on to say, "A psychiatrist should not be a party to any type of policy that excludes, segregates, or demeans the dignity of any patient because of ethnic origin, race, sex, creed, age, socioeconomic status, or sexual orientation." There are several ethical and legal issues that may come up in doing psychiatric work with lesbians, gay men, and bisexuals. Among these are confidentiality of medical records, a potential for abuse in settings that discriminate against gay, lesbian, and bisexual patients, and maintenance of standards of care.

Information about sexual feelings, behaviors, identity, and relationships should be entered into the medical record only when it is relevant to the evaluation and treatment of the patient and must always be treated in a confidential manner. Recent court decisions have upheld the status of confidentiality between a therapist and a patient to be as important as lawyer–client privilege. The US Supreme Court's 1996 *Jaffe v. Redmond* decision in particular holds that while medical records may be part of a patient's public medical record, process notes of private psychotherapy conversations remain privileged. In any event, maintaining the confidentiality of an individual's sexual orientation may be an ongoing concern in some cases, particularly where lesbians, gay men, and bisexuals may face discrimination if the medical records revealed their sexual identity in the workplace, a health care setting, or from insurers. Further complicating matters, in states that discriminate against same-sex relationships, a medical record that finds its way into court could lead to denying a gay, lesbian, or bisexual parent custody privileges.

Although a clinician's knowledge of a patient's sexual identity is essential to delivering quality of care, the recording of this information in an unprotected medical record (as in a public clinic or in a managed care setting) without informing the patient of this fact may become problematic. Nonclinical communications to others (a clinic's clerical staff, managed care case managers, school officials) about a patient's sexual identity or sexual behavior without the informed consent of the patient is always

unethical. When treating patients in settings where total confidentiality cannot be ensured, it is usually best if the psychiatrist discusses with the patient what can and cannot be kept confidential before any personal data are entered into the chart. If or when the patient does not wish personal information to be recorded and if the psychiatrist does not believe this information should be withheld, a referral to another clinician is highly recommended.

There are institutional settings, such as prisons or religious schools and universities that prohibit open expressions of a homosexual orientation identity by students and faculty, where public knowledge of a person's sexual orientation could have adverse consequences for a patient. As in the public sector, the psychiatrist in these settings is employed by a third party and not directly by the patient. Nevertheless, the ethical constraints on the psychiatrist are similar to those in the public mental health sector: The confidentiality of the patient, with the exception of potential harm to self or others, always takes precedence over other institutional needs.

Appropriate standards of care in treating gay, lesbian, and bisexual patients require that psychiatrists monitor their own antihomosexual biases when providing assessment and treatment. In this culture, antihomosexual attitudes are ubiquitous and, consequently, no one is free of them, not even gay and lesbian psychiatrists. This obliges the psychiatrist to understand and be aware of the ways in which these attitudes have had an impact on the development of their own identities. This can help a psychiatrist come to a greater empathic appreciation of the impact of such attitudes on a gay, lesbian, or bisexual patient. Psychiatrists and other practitioners who continue to treat homosexuality as if it were a mental illness engage in such practices outside the mental health mainstream. It is possible that they may risk ethical sanctions, particularly if they violate professional codes of conduct regarding informed consent, confidentiality, pretermination counseling, and provision of referrals after treatment failure. To date, there has been one case of a psychiatrist sanctioned by the APA for offering a patient unwanted sexual orientation conversion efforts. Closeted gay and lesbian psychiatrists struggling with their own identity issues may either need to refrain for a period of time from working with patients who have similar problems or seek supervision or treatment themselves. All psychiatrists, regardless of their own sexual identity, have an ethical obligation to be up to date and informed within their field and should regularly undertake continuing education about the mental health concerns and needs of gay, lesbian, and bisexual patients.

TRAINING AND RESEARCH NEEDS

Conceptions of sexuality, sexual identity, and sexual orientation have shifted significantly in the second half of the 20th century and continue to evolve and become more differentiated as awareness of sexual diversity

increases. Increased awareness of sexual diversity necessitates an expanding lexicon to provide descriptors for newly recognized sexual subgroups and identities. It also necessitates taxonomical and nosological revisions to provide coherent and meaningful systems of classification and nomenclature. A solid grounding in contemporary theories of sexuality will help residents to gain the skills and confidence necessary to interact sensitively and competently with persons of all sexual orientations. Competence in working with sexual minorities also entails an awareness that changing meanings in the vernacular and professional terminology may lead to miscommunication unless unambiguous descriptions are elicited to make sense of a patient's self-labels of identities and behaviors.

In addition to training in contemporary sexology, medical school and psychiatric residency education should include an historical science approach to the origins and evolution of the concept of sexual orientation within Western medical discourse. The mechanism of change and the thinking that went into homosexuality's removal from the DSM-II deserves special emphasis. It should be emphasized that the first DSM work group to exclude homosexuality was also the first to grapple with the issue of what actually constitutes a mental disorder. After careful study, the scientific committee concluded that homosexuality did not meet the emerging criteria for what constituted a mental disorder that would shape the subsequent DSM-III. Teaching residents this history is not only important to learning about the development of scientific thought, it is also essential to countering the erroneous but common misconception that homosexuality was deleted by a simple vote of APA members. That misconception contributes to the erroneous perception that the decision was primarily politically rather than scientifically motivated.

Although homosexuality is no longer in the DSM as a disorder, it is important that faculty and trainees learn about the real mental health consequences of being gay, many of which reflect individual responses to antihomosexual attitudes and social prejudice. Although accepting attitudes toward gay, lesbian, and bisexual individuals are increasing, antihomosexual bias is widespread in society at large and persists within health care systems. The 2011 Institute of Medicine report concludes that the minority stress associated with being a sexual orientation or gender identity minority person contributes to the heath disparities of LGBT people, especially the increased rates of depression, anxiety disorders, and substance abuse. Instruction about the lives and mental health needs of gay men, lesbians, and bisexual persons should not be restricted to classes on sexual minority issues, but should be fully integrated into the cultural competence curriculum as well as the broader curriculum on sexuality. It is important that cultural competence training foster an appreciation of the diversity within the sexual minority community and some familiarity with the common issues and concerns of sexual minorities. At the same time

cultural competence training tries to discourage stereotyping individual patients as each presents with a unique constellation of issues. Recent surveys report that many psychiatry residents neglect to routinely take detailed psychosexual histories or they report discomfort in doing so. Residency training programs need to address this broader issue, while also addressing competence in history taking with sexual minorities.

Theories on the origins of sexual orientation and their role in patient care also need to be addressed. Understanding the history of pathologizing etiological theories and the "therapies" on which they are based gives trainees some insight into the trepidation with which sexual minority individuals may approach mental health professionals. This knowledge also prepares residents to respond appropriately when patients bring this material into treatment. It is important that residents understand that sexual orientation is difficult, if not impossible, to change, appreciate the possible harm caused by efforts to convert a person's sexual orientation, and master the fundamentals of therapies that are affirmative and respectful of an individual's sexual orientation and identity. Given the current prominence of biologically deterministic theories, residents should also know that the decision to declassify homosexuality was not based on reductive presumptions about etiology or resistance to change. Contemporary gay affirmative therapies entail acceptance of homosexuality, regardless of any presumptions about its etiology, as a normal variation of human sexuality. Their focus is on identifying and addressing a patient's problems and conflicts, not revealing why the individual is gay or lesbian. Fear that a heterosexual therapist would focus on their sexual orientation rather than their chief complaints motivates many gays and lesbians to seek gay or lesbian therapists.

Table 21.2–1.
Internet Resources for Gay Men, Lesbian, and Bisexual Patients

American Medical Association LGBT Advisory Committee (public education)	http://www.ama-assn.org/ama/pub/about-ama/our-people/member-groups-sections/glbt-advisory-committee.page?
AGLP (education and professional support for psychiatrists, residents, and medical students)	http://www.aglp.org/
American Psychiatric Association: Facts on Sexual Orientation	http://www.psychiatry.org/mental-health/people/lgbt-sexual-orientation
GLMA (education and professional support for physicians and other health professionals)	http://www.glma.org/
Gay, Lesbian, and Straight Education Network (GLSEN) (support for LGBT youth in the school system)	www.glsen.org/
Group for the Advancement of Psychiatry (GAP): LGBT Mental Health Syllabus (for medical students and residents)	http://www.aglp.org/gap/
Hetrick-Martin Institute (social support for LGBT youth)	http://www.hmi.org/

Lambda Legal (advocacy group for legal protections)	http://www.lambdalegal.org/
PFLAG (support groups for families)	http://www.pflag.org/
SAGE (advocacy group for LGBT elderly)	http://sageusa.org/
Service members Legal Defense Network (advocacy and support of LGBT military personnel)	http://www.sldn.org/
Society for the Psychological Study of Lesbian, Gay, and Bisexual Issues—Division 44 of the American Psychological Association (public education and professional support)	http://www.apa.org/about/division/div44.html
World Professional Association for Transgender Health (WPATH) (professional support and Standards of Care)	http://www.wpath.org/

LGBT, Lesbian, Gay, Bisexual, and Transgender.

Instruction about the mental health needs of sexual minorities must include education concerning obstacles to their health care, many of which are being diminished by recent legislation and policy changes. Nevertheless, sexual minorities, due in part to negative past experiences, may tend to visit health care providers less often. They may also be guarded about discussing their sexual behavior with health care providers, fearing that it will lead to discrimination. Moreover, the use of heteronormative language and failure to inquire about sexuality on the part of health care providers may further discourage them from disclosing their sexual minority status. Yet such disclosure is a prerequisite to discussing how that status affects their lives as well as their particular sexual concerns. Alternatively, an atmosphere that fosters disclosure can be perceived as more welcoming and is likely to enhance compliance with follow-up appointments and continuity of care.

Although all residents will benefit from contact with respected psychiatrists who are openly gay or lesbian, sexual minority residents will benefit from support during their training by the identification of sexual minority faculty as role models. National organizations like AGLP: The Association of LGBTQ Psychiatrists (www.aglp.org) and the Caucus of Gay, Lesbian, and Bisexual Psychiatrists of the APA can serve as such a resource. Medical schools and professional organizations should increase the opportunities available for training about the clinical concerns of sexual minorities and encourage gay and lesbian faculty and other leaders to be visible and active in a wide range of professional activities.

There is still much to be understood about sexual orientation and how the needs of sexual minority individuals vary demographically and across the life cycle. Training should engender curiosity about these unknowns and an awareness of areas in need of further research. Preliminary research has described variations in sexual desire, behavior, and identity associated with a range of demographic and personal characteristics such as sex, age, race, class, geographic location, and religion. Most studies have focused on

white, urban, middle-class men, and there is a need to study sexual diversity in other groups as well. Because most studies have been cross-sectional, there is a need for longitudinal studies to investigate the development and experience of sexual orientation and sexual identity over time and across the lifespan. Compared to cross-sectional studies, longitudinal studies may be more helpful in elucidating the factors that contribute to the development of sexual orientation as well those that shape individual expression of desire and behavior.

Future directions for biological research are discussed briefly in the section on biology above. Current research has moved beyond traditional biological and psychosocial approaches to incorporate constructivist approaches that interpret the role of fluctuating cultural and other historical forces in shaping individual lives and experience.

One resource for a training program is a syllabus developed by the Group for Advancement of Psychiatry, available at http://www.aglp.org/gap/.

INTERNET RESOURCES

The websites listed in Table 21.2–1 offer useful resources for gay men, lesbian, and bisexual patients, their families, and for mental health professionals themselves.

REFERENCES

American Psychological Association Task Force on the Appropriate Therapeutic Response to Sexual Orientation. *Report of the Task Force on the Appropriate Therapeutic Response to Sexual Orientation*. Washington, DC: Author; 2009.

American Psychiatric Association. *Position Statement on Issues Related to Homosexuality*. 2013. Available at https://psychiatry.org/file%20library/about-apa/.../position-2013-homosexuality.pdf. Accessed November 1, 2016.

*Bailey MJ, Vasey PL, Diamond LM, Breedlove SM, Vilain E, Epprecht M. Sexual orientation, controversy and science. *Psychol Sci Public Interest*. 2016;17(2):45–101.

*Bayer R. *Homosexuality and American Psychiatry: The Politics of Diagnosis*. Princeton, NJ: Princeton University Press; 1987.

Boswell J. *Christianity, Social Tolerance, and Homosexuality: Gay People in Western Europe from the Beginning of the Christian Era to the Fourteenth Century*. Chicago: University of Chicago Press; 1980.

*Cabaj RP, Stein TS. *Textbook of Homosexuality and Mental Health*. Washington, DC: American Psychiatric Press; 1996.

Cochran SD, Drescher J, Kismodi E, Giami A, García-Moreno C, Reed GM. Proposed declassification of disease categories related to sexual orientation in ICD-11: rationale and evidence from the Working Group on Sexual Disorders and Sexual Health. *Bull World Health Organ*. 2014;92:672–679.

D'Augelli AR, Patterson C. *Lesbian, Gay, and Bisexual Identities over the Lifespan:*

Psychological Perspectives. New York: Oxford University Press; 1995.

*De Cecco JP, Parker DA. *Sex, Cells, and Same-Sex Desire: The Biology of Sexual Preference*. New York: Haworth Press; 1995.

Drescher J. *Psychoanalytic Therapy and the Gay Man*. Hillsdale, NJ: Analytic Press; 1998.

Drescher J, D'Ercole A, Schoenberg E. *Psychotherapy with Gay Men and Lesbians: Contemporary Dynamic Approaches*. New York: Harrington Park Press; 2003.

*Drescher J, Merlino JP. *American Psychiatry and Homosexuality: An Oral History*. New York: Harrington Park Press; 2007.

Drescher J, Zucker KJ. *Ex-Gay Research: Analyzing the Spitzer Study and Its Relation to Science, Religion, Politics, and Culture*. New York: Harrington Park Press; 2006.

Gartrell, NK, Bos HM, Goldberg NG. Adolescents of the U.S. National Longitudinal Lesbian Family Study: sexual orientation, sexual behavior, and sexual risk exposure. *Arch Sex Behav*. 2011;40:1199–1209.

Gooren L. The biology of human psychosexual differentiation. *Horm Behav*. 2006;50:589–601.

Hooker E. The adjustment of the male overt homosexual. *J Proj Tech*. 1957;21:18–31.

Institute of Medicine. *The Health of Lesbian, Gay, Bisexual and Transgender People: Building a Foundation for Better Understanding*. Washington, DC: The National Academies Press; 2011.

LeVay S. *Gay, Straight, and the Reason Why: The Science of Sexual Orientation*. Oxford: Oxford University Press; 2011.

Levounis P, Drescher J, Barber ME. *The LGBT Casebook*. Washington, DC: American Psychiatric Publishing; 2012.

Magee M, Miller DC. *Lesbian Lives: Psychoanalytic Narratives Old and New*. Hillsdale, NJ: Analytic Press; 1997.

Malyon AK. Psychotherapeutic implications of internalized homophobia in gay men. *J Homosex*. 1981;7:59–69.

Money J, Ehrhardt AA. *Man and Woman, Boy and Girl: The Differentiation and Dimorphism of Gender Identity from Conception to Maturity*. Baltimore: Johns Hopkins University Press; 1972.

Olson LA. *Finally Out: Letting Go of Living Straight, A Psychiatrist's Own Story*. Miami: InGroup Press; 2011.

Pathela P, Hajat A, Schillinger J, Blank S, Sell R. Discordance between sexual behavior and self-reported sexual identity: A population-based survey of New York City men. *Ann Int Med*. 2006;145:416–425.

Plöderl M, Tremblay P. Mental health of sexual minorities: a systematic review. *Int Rev Psychiatry*. 2015;27:367–385.

Reed GM, Drescher J, Krueger RB, et al. Revising the ICD-10 mental and behavioural disorders classification of sexuality and gender identity based on current scientific evidence, best clinical practices, and human rights considerations. *World Psychiatry*. 2016;15:205–221.

Ryan C, Huebner D, Diaz RM, Sanchez, J. Family rejection as a predictor of negative health outcomes in white and Latino lesbian, gay, and bisexual young adults. *Pediatrics*. 2009;123:346–352.

Savin-Williams RC. *The New Gay Teenager*. Cambridge, MA: Harvard University Press; 2005.

The Joint Commission. *Advancing Effective Communication, Cultural Competence, and*

Patient- and Family-Centered Care for the Lesbian, Gay, Bisexual, and Transgender (LGBT) Community: A Field Guide. Oak Brook, IL; 2011.

Worthington RL, Savoy HB, Dillon FR, Vernaglia ER. Heterosexual identity development: A multidimensional model of individual and social identity. *Couns Psychol*. 2002;30:496.

▲ 21.3 Paraphilic Disorders

RENEé M. SORRENTINO, M.D.

DSM-5 defines a *paraphilia* as "any intense and persistent sexual interest other than sexual interest in genital stimulation or preparatory fondling with phenotypically normal physically mature, consenting human partners." The term *paraphilic disorder,* as defined by the *Diagnostic and Statistical Manual of Mental Disorders,* fifth edition (DSM-5), is used to refer to a paraphilia that is currently causing distress or impairment to the individual or a paraphilia whose satisfaction has entailed personal harm, or risk of harm, to others. Paraphilic interests are more common in men and include a wide variety of behaviors typically seen in clinical populations seeking evaluation and treatment for paraphilias. Those who most frequently seek psychiatric treatment for sex offenses are those accused of or involved in child molestation, voyeurism, exhibitionism, fetishism, frottage, and public masturbation. The term *sex offender* is used to refer to an individual who has been legally convicted of a sex offense. The terminology sex offender does not provide any information regarding the offender's reason for committing the sexual offense nor does it indicate the presence or absence of a paraphilia. The psychiatric treatment of sex offenders is distinct from other clinical populations. In many jurisdictions in the United States sex offender legislation has mandated treatment for offenders deemed "sexually dangerous" or "sexually violent" in psychiatric hospitals where they are civilly committed. Such legislation has generated longstanding debates on the diagnosis of paraphilias, the nature of mental illness, and the treatability of sex offenders. The proposed changes to the paraphilia diagnoses in DSM-5 generated debate about how to conceptualize paraphilias and the adverse legal consequences implicit in identifying certain sexual behaviors as pathologic.

DEFINITION

The definition of sexual deviance or sexual perversities depends on the particular society and culture in which the behavior exists. A behavior may be considered deviant by one society and accepted, albeit unusual, by another. The term *paraphilia* is derived from the Greek words "*para*" meaning next to and "*philia*" meaning love. The etymological definition of paraphilia is "next to or alongside of love." The term paraphilia therefore refers to a love that is perverted or "alongside the norm." In addition to the presence of a deviant sexual interest, paraphilias are characterized by the inability to resist an impulse for the sexual acts.

DSM-5 made a distinction between paraphilias and paraphilic

disorders. The term paraphilic disorders refer to a paraphilia that is currently causing distress or impairment to the individual or a paraphilia whose satisfaction has entailed personal harm, or risk of harm, to others. The definition of paraphilias in DSM-5 changed. Previously, paraphilias were defined as intense, recurring sexual fantasies, sexual urges, or behaviors that involve nonhuman objects, children or nonconsenting adults, suffering or humiliation (to self or to others), lasting for at least 6 months. DSM-5 defined paraphilia as "any intense and persistent sexual interest other than sexual interest in genital stimulation or preparatory fondling with phenotypically normal, physically mature, consenting human partners."

DSM-5 includes eight specific paraphilic disorders: exhibitionistic disorder (sexual arousal from exposure of genitals to a stranger), pedophilic disorder (sexual arousal to prepubescent individual, generally 13 years of age or younger), voyeuristic disorder (observing others' sexual activities), fetishistic disorder (use of inert objects, such as female undergarments), transvestic disorder (cross-dressing), sexual sadism disorder (inflicting suffering or humiliation), sexual masochism (being humiliated, beaten, bound, or made to suffer), and frotteuristic disorder (touching, rubbing against a nonconsenting person). Paraphilia not otherwise specified (NOS) has been replaced in DSM-5 with specified paraphilic disorder and unspecified paraphilic disorder as the result of a system-wide standardized DSM-5 change to each residual DSM-IV NOS category.

HISTORY

Societies define normal sexual behavior as that behavior which results in healthy, nonharming, sexual contact. Those sexual practices that do not result in healthy, nonharming, sexual contact are judged as undesirable and in need of change. Those individuals who engage in sexual behaviors outside this definition are subject to criminal action. Historically, society has created sexual prohibitions that, if breached, have serious legal and social consequences. Of all criminal behaviors, those who breach sexual prohibitions are viewed as the most deplorable. Those behaviors that cause obvious harm to others, such as sexual contact with children or bodily punishment to another, are universally defined as abnormal or sexually deviant. However, those sexual practices that do not harm others are not as clearly defined. Some societies have further defined sexual deviance as sexual practices that are maladaptive or deviations from the social norm. This definition would then include those behaviors that are considered evolutionarily maladaptive and uncommon. Richard von Krafft-Ebing, a German psychiatrist credited with formally introducing the study of sexology as a psychiatric phenomenon, identified paraphilias first in his 1886 *Psychopathia Sexualis* (Sexual Psychopathy). This highly influential

psychiatric text laid the groundwork for the development of research and treatment in this area that has taken place over the past century.

The term *paraphilia* was coined by psychologists in the early 1900s to refer to the sexual variations that were just starting to be researched. Because of a lack of understanding, it included transvestism and other nonsexual variations as well. Previously, these were called *perversions*. Paraphiliacs treated by psychologists were seeking help for a behavior they could not control. The word retained this meaning and did not change to include those who were sexually curious.

Sigmund Freud held that human sexuality advanced through stages. He theorized that individuals would progress toward "normal" heterosexuality unless prevented. He saw all paraphilias as infantile, that is, not mature sexual behavior. Because the paraphilias that involved diapers fit well into this theory, they were well documented. Wilhelm Stekel's *Patterns of Psychosexual Infantilism* came from this perspective. Homosexuality, sadism, zoophilia, and many others were included in his examples.

As research advanced, Freud's notion of psychosexual infantilism was replaced with different views. These views dealt individually with the major paraphilias, such as fetishism, masochism, and transvestism. The diaper paraphilias were generally classified into one of these major categories as a result of the author's viewpoints. In 1987, the third revised edition of DSM (DSM-III-R) used *infantilism* to refer to those who had a "desire to be treated as a helpless infant and clothed in diapers" and categorized infantilism under sexual masochism.

COMPARATIVE NOSOLOGY

DSM originally classified sexual deviations with psychopathic personality disturbances. This classification was based on the belief that sexual deviations were unlawful and thereby those individuals engaging in sexual deviations were psychopathic. In DSM-II, sexual deviations were classified with the personality disorders. DSM-III changed the nomenclature of these behaviors from sexual deviation to paraphilia. Paraphilias were classified as psychosexual disorders, which included gender identity disorders, psychosexual dysfunctions, and egodystonic homosexuality. DSM-IV-TR reclassified transvestism from a disorder of gender identity to a paraphilia called *transvestic fetishism*. DSM-5 redefined paraphilias, made a distinction between paraphilias and paraphilic disorders and added course specifiers. The paraphilic disorders are presented in DSM-5 according to the following classification scheme: anomalous activity preferences which include voyeuristic disorder, exhibitionistic disorder, and frotteuristic disorder; algolagnic disorders which include sexual masochism disorder and sexual sadism disorder; anomalous target preferences which include pedophilic disorder, fetishistic disorder, and transvestic disorder.

The question of whether paraphilias are mental disorders was directly

answered in DSM-5. In DSM-5, paraphilias are not considered mental disorders unless they cause distress or impairment (criterion B) to the individual or the paraphilia causes personal harm or risk of harm to others. If a paraphilia causes impairment or harm, it is a mental disorder termed "paraphilic disorder." This change in nomenclature represents an attempt by DSM-5 to preserve the distinction between normative and nonnormative sexual behavior without labeling nonnormative sexual behavior a mental disorder.

DSM-5 was the first edition to add course specifiers to the paraphilic disorders. Similar to the course specifiers in other DSM diagnoses, the course specifiers indicate a change in symptoms. The course specifiers are included in all of the paraphilic disorders with the exception of other specified paraphilic disorder, unspecified paraphilic disorder, and pedophilic disorder. The course specifier named, "in a controlled environment" was introduced to indicate the absence of paraphilic urges in settings where opportunities to engage in paraphilic behaviors are restricted. An individual "in full remission" has not acted on the urges with a nonconseting person and there has been no distress or impairment in social occupational or other areas of functioning for at least 5 years in an uncontrolled environment. The question of whether a paraphilic disorder can be in remission independent of environment is not known.

Previous editions of DSM attempted to address the diagnostic difficulties and assumptions inherent in the diagnostic category of paraphilias. There is a second group of sexual impulse disorders not currently classified as paraphilias because the particular sexual behaviors affected are not considered "deviant" with respect to contemporary cultural norms. These disorders have been designated as paraphilia-related disorders by Kafka based on the following clinical data:

- ► The boundary for social as well as sexual deviance is largely determined by cultural and historical context. As such, sexual disorders once considered paraphilias (e.g., homosexuality) are now regarded as variants of normal sexuality; so too, sexual behaviors currently considered normal (e.g., masturbation) were once culturally proscribed.
- ► Paraphilia-related disorders have been diagnosed in male paraphiliacs, and men selected for studies of paraphilia-related disorders are reported to have paraphilias as well.
- ► Paraphilia-related disorders can produce a pattern of sexual frequency and intensity with concomitant psychosocial impairment that does not readily distinguish them from nonviolent paraphilic disorders.
- ► Analogous with paraphilias, the presence of a single paraphilia-related disorder increases the likelihood of occurrence of a multiplicity of these behaviors in an affected person.
- ► Paraphilias and paraphilia-related disorders can both be ameliorated by

antiandrogens and serotonergic antidepressants, such as SSRIs. Disorders of sexual impulsivity can be pleomorphic and may include sexual arousal to behaviors that are socially "deviant" as well as "normal." In fact, inasmuch as paraphilia-related disorders may be common in paraphiliacs and also occur without the companion of the latter, paraphilia-related disorders may be the more prevalent form of sexual impulsivity. In addition, although the estimated sex difference ratio for paraphilia-related disorders is unknown, protracted promiscuity (e.g., "nymphomania") and compulsive masturbation are not uncommon behaviors described by women.

The paraphilic-related disorders were not incorporated in DSM. However, DSM-5 considered including portions of these disorders by proposing a new disorder, hypersexual disorder. Hypersexual Disorder was defined as nonparaphilic sexual interests with an impulsivity component. Hypersexual disorder was rejected from DSM due to concerns regarding its diagnostic validity.

DSM-5 proposed broadening the pedophilic disorder diagnosis to include attraction to pubescent as well as prepubescent children. The proposed modified diagnosis, pedophebephilic disorder, referred to a distinct group of individuals who are sexually aroused to adolescents. Ultimately pedophebephilic disorder was not incorporated in the DSM-5.

There has been a longstanding disagreement regarding whether rapists meet the criteria for a paraphilia. The category of paraphilic coercive disorder was proposed in DSM-5 for individuals with intense, repetitive urges of 6 months' duration to commit rape, who had either acted on these urges or were disturbed by the urges. DSM has never included rape or paraphilic coercive disorder as a diagnosis.

The changes to paraphilic disorders in DSM-5 represent an attempt to answer the central philosophical question of what constitutes disordered sexual behavior. They include a change in nomenclature and the addition of courses. Previous additions of DSM have attempted to make a distinction between normal and abnormal sexual behaviors. DSM-5 made a distinction between paraphilias and paraphilic disorders, indicating that paraphilias are not ipso facto mental disorders. The basic structure of the diagnostic criteria in DSM-5 is unchanged from DSM-IV-TR.

EPIDEMIOLOGY

Gender Ratio

Paraphilias are predominantly male sexuality disorders. Except for sadism and masochism, the paraphilias are almost never diagnosed in females, although some cases have been reported.

Nonparaphilic Populations. The frequency of paraphilic behavior is

unknown. The majority of epidemiologic data relating to paraphilias derive from clinical populations. It is difficult to obtain the frequency of paraphilic behaviors in the nonclinical population, as individuals are reluctant to report sexual fantasies and behaviors. In one study of the general population, 61.7 percent of men reported fantasies of initiating a young girl into sexuality; 33.0 percent described fantasies of raping adult women; 11.7 percent described masochistic fantasies; 5.3 percent described fantasies of having sex with an animal; and 3.2 percent described fantasies of initiating a young boy into sexuality. A study of college students' sexual behavior revealed that 42 percent reported having participated in voyeurism; 35 percent in frottage; 8 percent in making obscene telephone calls; 5 percent in coercive sexual activity; 3 percent in sexual contact with girls under age 12; and 2 percent in exhibitionism. A total of 65 percent reported having participated in some variant of paraphilic behavior.

Paraphilic Populations. Those individuals who seek evaluation and treatment are a distinct group. It is unlikely that individuals seeking treatment represent the heterogeneous population of individuals with paraphilias. Some paraphilias are more likely to come to the attention of others because they involve another person (such as pedophilia or exhibitionism). However, those paraphilias that do not involve another person, such as the use of fetish objects, are less likely to be treated. Those paraphilias whose behavior is criminal tend to be overrepresented in clinical populations because they frequently lead to mandatory reporting or court-ordered treatment.

Data gathered from more than 90 treatment programs throughout North America regarding 2,129 cases of individuals seeking assessment indicate that 37.1 percent had been involved in child molestation; 20.2 percent in voyeurism; 13.8 percent in exhibitionism; 13.3 percent in fetishism; 11.2 percent in frottage; and 10.7 percent in public masturbation. A study of 561 paraphilic persons who had committed more than 291,000 paraphilic acts against more than 195,000 victims indicated that, of all victims, 37.3 percent were victims of exhibitionism; 28.6 percent of frottage; 13.6 percent of voyeurism; 11.8 percent of molestation against boys outside the home; 3.5 percent of public masturbation; 2.3 percent of molestation of girls outside the home; 1 percent of obscene telephone calls; 0.9 percent of bestiality; 0.5 percent of rape of adult women; 0.2 percent of molestation of girls in the home; 0.2 percent of urolagnia; and 0.1 percent or less of sadism, masochism, fetishism, molestation of boys within the home, obscene mailings, coprophilia, or attraction to odors.

Frequency of Paraphilic Behavior

The presence of a paraphilia does not indicate whether an individual engages in paraphilic behavior or sexual fantasy. Some individuals never

act on paraphilic urges. Rather, they engage in sexual gratification from sexual fantasies. Other individuals carry out paraphilic fantasies. In a study of individuals seeking treatment, the median number of paraphilic acts by paraphilic category are as follows: Nonincestuous molestation of a female victim 1.4; nonincestuous molestation of a male victim 10.1; incestuous molestation of a female victim 4.4; incestuous molestation of a male victim 5.2; rape of adult women 0.9; exhibitionism 50.5; voyeurism 16.5; frottage 29.5; transvestism 25.0; fetishism 3.3; sadism 3.0; masochism 36.0; obscene telephone calls 30.0; public masturbation 50.0; and bestiality 2.2.

Crossing Over

Initially, individuals suffering from a paraphilia were thought to be obsessed with one type of paraphilic behavior. Recent studies suggest that individuals with paraphilias "cross over" from one paraphilia behavior to another. Of paraphilic persons primarily involved with actual touching of victims (frottage, rape, and pedophilia), 30.6 percent had previously or concomitantly been involved in paraphilic acts not involving touching of the victim (voyeurism and exhibitionism). Sixty-four percent of paraphilic individuals primarily involved with nontouching had previously or concomitantly carried out paraphilic acts that involved touching. Of paraphilic persons primarily victimizing outside the family, 29.4 percent had previously or concomitantly carried out paraphilic acts against family victims. Of those who primarily had been involved in paraphilic acts with family members, 65.8 percent had previously or concomitantly carried out paraphilic sexual acts against nonfamiliar victims. Of paraphilic persons who primarily committed acts against female victims, 22.9 percent had previously or concomitantly carried out paraphilic acts against male victims. Of those who had primarily carried out paraphilic acts against male victims, 62.6 percent had previously or concomitantly carried out paraphilic acts against female victims. Those paraphilias with the highest incidence of crossover were sadism, public masturbation, male incest pedophilia, fetishism, masochism, voyeurism, female nonincest pedophilia, exhibitionism, male nonincest pedophilia, transvestism, frottage, female incest pedophile, and rape of adult women. These studies suggest that paraphiliacs tend to cross over between touching and nontouching of their victims, between family and nonfamily members, between female and male victims, and to victims of various ages.

ETIOLOGY

The etiology of paraphilias is unknown. Biological theory, including neuroanatomical and neuroimaging studies of sex offenders, indicate that congenital or acquired brain damage are overrepresented. Some pedophilic perpetrators show structural impairment of brain regions critical for sexual

development. Subtle defects of the right amygdala and closely related structures might be implicated in the pathogenesis of pedophilia and might reflect developmental disturbances or environmental insults at critical periods. Neurological hypothesis of paraphilia propose that sexual deviance is associated with frontal and/or temporal lobe damage. This damage may translate into an individual's inability to control sexual impulse or directly cause paraphilic behavior as a result tissue damage.

A monoamine hypothesis for the pathophysiology of paraphilic disorders was first articulated in 1997 by Kafka. This hypothesis was based on four converging lines of empirical evidence. First, the monoamine neurotransmitters, dopamine, norepinephrine, and serotonin serve a modulatory role in human and mammalian sexual motivation, appetite, and consummatory behavior. Second, the sexual effects of pharmacological agents that affect monoamine neurotransmitter can have both significant facilitative and inhibitory effects on sexual behavior. Third, paraphilic disorders appear to have Axis I comorbid associations with nonsexual psychopathologies that are associated with monoaminergic dysregulation. Last, pharmacological agents that enhance central serotonergic function in particular have been reported to ameliorate paraphilic sexual arousal and behavior.

Psychoanalytical Theory

Psychoanalysts generally theorize that these conditions represent a regression to or a fixation at an earlier level of psychosexual development, resulting in a repetitive pattern of sexual behavior that is not mature in its application and expression. In other words, an individual repeats or reverts to a sexual habit arising early in life. Another psychoanalytic theory holds that these conditions are all expressions of hostility in which sexual fantasies or unusual sexual acts become a means of obtaining revenge for a childhood trauma. The persistent, repetitive nature of the paraphilia is caused by an inability to erase the underlying trauma completely. Indeed, a history of childhood sexual abuse is sometimes seen in individuals with paraphilias.

Behaviorist Theory

Behaviorists suggest that the paraphilia begins via a process of conditioning. Nonsexual objects can become sexually arousing if they are frequently and repeatedly associated with a pleasurable sexual activity. The development of a paraphilia is not usually a matter of conditioning alone; there must usually be some predisposing factor, such as difficulty forming person-to-person sexual relationships or poor self-esteem.

The following are situations or causes that might lead someone in a paraphiliac direction:

- parents who humiliate and punish a small boy for strutting around with an erect penis;
- a young boy who is sexually abused;
- an individual who is dressed in a woman's clothes as a form of parental punishment;
- fear of sexual performance or intimacy;
- inadequate counseling;
- excessive alcohol intake;
- physiological problems;
- sociocultural factors;
- psychosexual trauma.

DIAGNOSIS AND CLINICAL FEATURES

DSM identifies eight categories of paraphilic disorders. All of the paraphilic disorders meet the following DSM criteria: The activity must be the sole means of sexual gratification for a period of 6 months and cause "marked distress or interpersonal difficulty" or "whose satisfaction has entailed personal harm, or risk of harm, to others" to be considered such. The categories of paraphilic disorders are exhibitionism, fetishism, frotteurism, masochism, pedophilia, sadism, transvestic fetishism, and voyeurism. Paraphilias that do not meet the guidelines of these categories are called specific paraphilic disorder or unspecified paraphilic disorder. These terms replaced paraphilia NOS that existed in DSM-IV-TR. Paraphiliacs may choose an occupation or volunteer in an environment that places them in close contact with the desired object. For example, a pedophile may choose to volunteer with young children in an effort to make contact with children.

Diagnostic Considerations

The diagnosis of a paraphilic disorder is based on the individual's report or demonstration of sexual arousal to a specified object or behavior with resultant impairment or risk of harm to self or others. For example, an individual with pedophilia may meet the diagnostic criterion by reporting sexual arousal to prepubescent children. Alternatively, a pedophile may meet the diagnostic criterion for pedophilia by demonstrating sexual arousal to stimuli of prepubescent children (e.g., erection to pictures of children in sexual poses). Paraphilic interests develop shortly after puberty and follow a chronic course.

Exclusive and Nonexclusive Subtypes. Individuals who identify a paraphilic interest to be the only way to achieve sexual gratification are classified as "exclusive" paraphiliacs. A man who is only sexually aroused to prepubescent children is a pedophile, exclusive subtype. Individuals who have paraphilic interest but are able to derive sexual gratification from

other experiences are called "nonexclusive" paraphiliacs. A nonexclusive pedophile experiences sexual gratification from both prepubescent children and adults. Nonexclusive paraphiliacs are more common than exclusive paraphiliacs.

Exhibitionistic Disorder

Exhibitionistic disorder, also known as "flashing" or "indecent exposure," is qualified by recurrent and intense sexual arousal from the exposure of one's genitals to an unsuspecting person persisting for a period of at least 6 months. In order to meet the criteria for a disorder the individual must act on these sexual urges with a nonconseting person, or the sexual urges or fantasies cause clinically significant distress or impairment in social, occupational, or other important areas of functioning. Exhibitionism occurs predominately in men, with a peak age of onset in the 20s. In the United States, most exhibitionists are Caucasian males. About half of exhibitionists are married. Generally, no additional contact with the observer is sought; the individual is stimulated sexually by gaining the attention of and startling the observer. The exhibitionist often masturbates as part of the exposure. Exposing oneself while driving or in a public place are the most common expressions of exhibitionism. DSM-5 added specifiers to identify the source of sexual arousal. The specifiers include the following: sexual arousal by exposing genitals to prepubertal children, sexual arousal by exposing one's genitals to physically mature individuals, and sexual arousal by exposing genitals to prepubertal children and to physically mature individuals.

Exhibitionism is one of the three most common sexual offenses in police records (the other two are voyeurism and pedophilia). Exhibitionists have been found to have a high rate of comorbidity among the paraphilias as a group and between the paraphilias as a group and other mental disorders. In other words, a patient in treatment for exhibitionism is highly likely to engage in other forms of deviant sexual behavior and to suffer from depression or an anxiety or substance abuse disorder.

The diagnosis of exhibitionistic disorder describes individuals, who for at least 6 months experience sexually gratifying urges, fantasies, or behaviors that related to displaying nudity to an unsuspecting individual. The sexual arousal derives from the showing of one's genitals, breasts, or buttocks to a nonconsenting individual. The diagnosis may be made if the sexual urges or fantasies lead to significant distress, impairment in functioning or the actual behavior.

Fetishistic Disorder

Individuals who achieve sexual gratification from the use of nonliving objects or a highly specific focus on nongenital body parts over a period of

at least 6 months with resultant clinically significant distress or impairment in social, occupational, or other important areas of functioning meet the criteria for fetishistic disorder. DSM-5 specifies that the fetish objects are not limited to articles of clothing used in cross-dressing or devices specifically designed for the purpose of tactile genital stimulation. Fetishistic disorder is further defined by specifiers indicating the source of sexual arousal such as "body part(s), nonliving object(s), or other." The use of objects, most commonly women's undergarments, shoes, stockings, or other clothing items, suffer from fetishism. Fetishists often collect the object of sexual gratification. Some of the more common objects are women's lingerie or specific materials like silk, leather, or fur. Partialism refers to fetishes specifically involving nonsexual parts of the body. Fetishism has predominantly been described in men and often exists with other paraphilias.

Fetishistic disorder is used to refer to at least 6-month period of time in which an individual experiences sexual urges, fantasies, or behaviors related to objects, situations, or body parts not traditionally viewed as sexual. The sexual arousal must cause significant distress or impairment in functioning.

Frotteuristic Disorder

Frotteurism refers to the sexual gratification achieved by touching or rubbing against a nonconsenting person. In order to meet the criteria for frotteuristic disorder the individual must experience recurrent and intense sexual arousal over a 6-month period and either acted on the behavior or the sexual urges or fantasies cause clinically significant distress or impairment in social, occupational, or other important areas of functioning. This behavior often occurs in busy, crowded places, such as on busy streets or on crowded buses or subways. Frotteurism is predominantly described in men. A frotteur may fantasize that the women he is rubbing up against is mutually aroused by the behavior.

The diagnosis of frotteuristic disorder is based on a greater than 6-month period of time in which an individual experiences gratification from urges, fantasies, or behaviors related to brushing up against an unsuspecting individual. An individual meets the criteria for the diagnosis if he experiences significant distress or impairment in functioning as a result of the fantasies or has acted on the fantasies.

Pedophilic Disorder

Pedophilia has been included in the American Psychiatric Association DSM as a mental disorder since 1968. In DSM-III and DSM-III-R, pedophilia was defined as involving "recurrent intense sexual urges and sexually arousing fantasies involving sexual activity with a prepubescent child or

children." This criterion excluded individuals who molested children but did not experience recurrent intense sexual urges or fantasies involving sexual activity with a prepubescent child. DSM-IV-TR defined the diagnostic criteria for pedophilia as "recurrent, intense sexually arousing, fantasies, sexual urges, or behaviors involving sexual activity with a prepubescent child or children." These fantasies, sexual urges, or behaviors are pedophilic if they "cause clinically significant distress or impairment in social, occupational, or other important areas of functioning." The pedophile also has to be 16 years of age or older and at least 5 years older than the victim child, although for late adolescents this age discrepancy is not specified. DSM-5 redefined pedophilia as pedophilic disorder referring to the following criteria: over a period of at least 6 months, recurrent, intense sexually arousing fantasies, sexual urges, or behaviors involving sexual activity with a prepubescent child or children (generally age 13 years or younger). The individual has acted on these sexual urges, or the sexual urges or fantasies cause marked distress or interpersonal difficulty. The individual is at least 16 years of age and at least 5 years older than the child or children. The disorder is further defined by specifying whether the sexual fantasies, sexual urges, or behaviors are the exclusive type, attracted only to children, or nonexclusive type, attracted to children and adults. The pedophilic arousal is further identified as sexually attracted to males, females, both, or and whether it is limited to incest.

The term *child molester* is a descriptive, nonclinical term. Some child molesters are driven by pedophilic interests, while others molest children as a result of nonpedophilic interests such as the unavailability of an appropriate sexual partner. Pedophiles typically seek treatment as a result of legal consequences for illegal behavior.

For most pedophiles the onset of attraction to children begins at the time of puberty or adolescence. The course of pedophilia is long term. Individuals who engage in sexual activities with pubescent teenagers under the legal age of consent (ages 16 to 18) are known as *hebophiles* (attracted to individuals at the early stages of puberty) or *ephebophiles* (attracted to pubertal individuals). Most pedophiles do not use physical force to engage a victim in sexual contact; rather they use manipulation and cajoling to elicit the desired response from the child. Pedophiles commonly justify and minimize the actions by stating that the child enjoyed the experience or that the experience was of educational value.

The majority of pedophiles are male. From federal data on sexual crimes, females were reported to be the molester in 6 percent of all juvenile cases. A study of 4,007 child molesters found 1 percent to be female, but the authors of this study suggested the number was low because of the systematic underreporting of women for molestation. Most pedophiles are nonexclusive, meaning that they have sexual attraction to adults in addition to children. Pedophiles are categorized according to type of exclusivity,

gender of child they are attracted to, and whether the child is familial (e.g., incest) or nonfamilial. A larger percentage of pedophiles are homosexual or bisexual in orientation to children. Individuals attracted to females usually prefer children between the ages of 8 and 10 years. Individuals attracted to males usually prefer boys between the ages of 10 and 13 years.

Fifty to 70 percent of pedophiles can be diagnosed as having another paraphilia, such as frotteurism, exhibitionism, voyeurism, or sadism. Pedophiles are approximately two and a half times more likely to engage in physical contact with a child than simply voyeuristic or exhibitionist activities. Typically, pedophiles engage in fondling and genital manipulation more than intercourse, with these three exceptions: intercourse occurring in cases of incest, occurring in pedophiles with a preference for older children or adolescents, and occurring when children are physically coerced.

The question of whether individuals who view child pornography meet the criteria for pedophilic disorder and are at risk for committing hands-on sexual offenses has been the focus of research. Research has not established a clear relationship between viewing child pornography and child sexual offenses. Studies show the majority of men convicted or charged with child pornography offenses have demonstrated sexual arousal to prepubescent stimuli on phallometric testing.

Although preliminary, imaging techniques such as functional magnetic resonance imaging (fMRI), have attempted to identify variations in brain morphology in individuals with pedophilia. fMRI in combination with emotional and/or sexual stimuli showed pedophilia is likely associated with altered brain activity particularly in the frontal brain areas and the temporal lobe.

Research into the etiology of pedophilia has been inconclusive. Reported differences between pedophiles and the general population include lower intelligence, a slight increase in the prominence of left-handed individuals, impaired cognitive abilities, neuroendocrine differences, and brain abnormalities, particularly frontocortical irregularities or differences.

Pedophilic disorder refers to the sexual gratification for at least 6-month period of time from sexual fantasies or behavior related to sexual activity with a prepubescent individual. In order to be considered pedophilic, the individual must be at least 16 years old and the prepubescent child must be at least 5 years older than the individual. The diagnosis is made when the individual's ongoing sexual urges or fantasies have led to significant distress, impairment in functioning or the actual behavior.

Sexual Masochism Disorder

Sexual masochism is a term used to refer to sexual gratification produced

by pain that is inflicted by others or by one's self. Sexual masochistic disorder refers to over a period of at least 6 months, recurrent and intense sexual arousal from the act of being humiliated, beaten, bound, or otherwise made to suffer with resultant clinically significant distress or impairment in social, occupational, or other important areas of functioning. Studies suggest that an estimated 5 to 10 percent of the population has engaged in some masochistic sex play. Research suggests that sexual masochism is a fairly benign activity. Studies of practicing masochists have found them to be reasonably well-adjusted, successful individuals. If studies describe sexual masochism as a benign activity, is sexual masochism a mental disorder? Proponents of sexual masochism as a mental disorder theorize that an individual may become dependent on masochistic activity or may engage in dangerous or injurious behavior warranting a pathologic aspect of the behavior. In most cases masochism does not cause harm to the individuals. However, sexual masochism disorder with asphyxiophilia, which refers to the practice of achieving sexual arousal related to sexual masochism disorder. There is no scientific evidence to support the diagnosis of sexual masochism as pathological. Therefore, if an individual seeks treatment for masochism and there is no harm caused from the behavior, the individual will not benefit from treatment.

Sexual masochism is one of the few paraphilias described in women. As discussed previously, the majority of paraphilias are almost never diagnosed in women. The three main characteristics of masochism are pain, loss of control, and humiliation. One of the psychodynamic theories postulated about masochism is Freud's writings that masochism is derived from sadism: More precisely, the person wishes to dominate and hurt others, but this desire gives rise to guilt, so the person defensively converts it into its opposite, namely a desire to submit to domination by others. However this theory is not entirely supported by the empirical evidence.

Sexual masochism disorder refers to at least 6-month period of time with ongoing sexual arousal from fantasies, urges, or behaviors related to the physical or psychological abuse one is made to suffer. The sexual urges or fantasies cause significant impairment in functioning, significant distress or the actual behaviors.

Sexual Sadism Disorder

Sexual sadism is sexual gratification achieved from the infliction of psychological or physical suffering or both on a partner. Sexual sadism disorder is a period of at least 6 moths with recurrent and intense sexual arousal from the physical or psychological suffering of another person and the individual has acted on these sexual urges with a nonconsenting person or the sexual urges or fantasies cause clinically significant distress or impairment in social, occupational, or other important areas of

functioning. Sadists experience sexual fantasies involving the suffering of another as well as compulsions to enact fantasies. The sadist's partner may consent to the behavior in an effort to satisfy his or her own masochistic fantasies, or the partner may be the victim. Sadism, like masochism, has been described in both men and women. Similarly both sadism and masochism are sexual behaviors that are reported in the context of normal sexuality.

Krafft-Ebing subclassified sexual sadism into several categories including:

1. Lust-murder;
2. Mutilation of corpses or necrophilia;
3. Injury to females (stabbing);
4. Defilement of women;
5. Other kinds of assaults on women—symbolic sadism in which, for example, the perpetrator cuts the hair of his victims rather than harming them directly;
6. Ideal sadism or sadistic fantasies alone without acts;
7. Sadism with other objects, for example, whipping boys;
8. Sadistic acts with animals.

In some individuals sadistic fantasies and compulsions lead to rape or homicide. However, not all rapes or sexual homicides are perpetrated by sadists. Sadistic rapists are aroused by the use of gratuitous violence, whereas the preferential rapist is thought not to use greater force than is necessary for submission of the victim. The sadistic rapist is detected by demonstrating a higher level of arousal to description of physical, but nonsexual, assault on female victims when subject to phallometric assessment. One confounder, however, is that normal men acknowledge sexual arousal to the idea of raping or hurting a woman. Among the general male population, 30 percent report fantasies of tying up and raping a woman.

The majority of sadists report the onset of fantasies in late teen years. Some sadists' accounts of sexual arousal to violence is based on an association between excitement and punishment in childhood. However, the majority of sadists are unable to delineate a historical marker for the sadistic sexual appetite.

The diagnosis of sexual sadism disorder is used to define a period of at least 6 months of sexual gratification from fantasies, urges, or behaviors related to the harm of a nonconsenting individual by psychological abuse or physical abuse. The diagnosis is made when the individual has acted on the fantasies or experienced sexual fantasies or urges that cause significant distress or impairment in functioning.

Transvestic Disorder

The literal meaning of the word transvestism is to wear the clothing of the opposite sex. DSM-5 uses the term transvestism to refer to a male who dresses in women's clothing with resultant sexual arousal. Transvestic disorder is defined as the recurrent and intense sexual arousal from cross-dressing for over a 6-month period of time in which the fantasies, sexual urges, or behaviors cause clinically significant distress or impairment in social, occupational, or other important areas of functioning. If the cross-dresser is specifically sexually aroused by fabrics, materials, or garments, the specifier "with fetishism" should be added. Cross-dressing typically begins in prepubertal children but does not appear to have a sexual component until late adolescence or early adulthood. Individuals with transvestic fetishism are usually heterosexual. However, there is a subgroup of men with transvestic fetishism who report gender dysphoria, which can be defined as discontent with one's biological sex, the desire to possess the body of the opposite sex, and the desire to be regarded by others as a member of the opposite sex. Initially the presence of gender dysphoria ruled out the diagnosis of transvestic fetishism according to DSM-5. However, additional research suggests the presence of gender dysphoria is not exclusionary in the diagnosis of transvestic fetishism. DSM-5 specifier "with autogynephilia" refers to individuals who are sexually aroused by thoughts or images of self as a female.

It is unclear whether transvestic fetishism causes the level of dysfunction typical of mental disorders. Clinical evidence suggests that transvestic fetishism impairs the ability to form intimate pair bonds with other persons in the sense that relationships tend to break down because the individual with transvestic fetishism often has more interest in the nonhuman aspects of the partner (e.g., clothing). However, not all cases of transvestic fetishism exhibit dysfunctional relationships or distress.

The diagnosis of transvestic disorder describes a period of at least 6 months, of ongoing sexual arousal from sexual fantasies, urges, or behaviors related to dressing in the clothing of the opposite gender. The sexual arousal results in significant distress or impairment in functioning.

Voyeuristic Disorder

Voyeurism is sexual gratification achieved from watching people undressing or nude, or observing them during sexual acts without their knowledge or consent. Voyeuristic disorder is qualified by recurrent and intense sexual arousal from observing an unsuspecting person who is naked, in the process of disrobing, or engaging in sexual activity over a period of at least 6 months. In addition, to meet the criteria for voyeuristic disorder, the individual is at least 18 years of age and has acted on these sexual urges with a nonconsenting person or the sexual urges or fantasies cause clinically significant distress or impairment in social, occupational or other important areas of functioning. Voyeurs are usually men aged 20s

and 30s who prefer to watch unknown women. They may masturbate while observing the victims but typically seek no contact with the victim. Some voyeurs report sexual fantasies that the woman they are watching will solicit sex from them. The onset of voyeuristic behavior is usually before the age of 15 years. Voyeurs tend to have multiple paraphilias and to cross over from one paraphilia to another. Although some voyeurs do not commit more serious paraphilic acts, there are some suggestions in the literature that serial rapists and murderers began with histories of voyeurism. Voyeurism is one of the least understood behaviors of the paraphilias.

The diagnosis of voyeuristic disorder refers to a 6-month period of ongoing, sexual gratification from observing an unsuspecting individual who is undressing, naked, or engaging in sexual activity. In order to meet the diagnostic criterion for the diagnosis of voyeuristic disorder, an individual must be at least 18 years old and either engaged in the actual behavior or experienced significant distress or impairment in functioning as a result of the specific sexual arousal.

PATHOLOGY AND LABORATORY EXAMINATION

Plethysmography is the phallometric assessment of physiological sexual arousal by use of equipment that measures penile circumference in males. The test is conducted by placing a small circular strain gauge around the penis. This gauge is calibrated to measure changes in the circumference of the penis. The test subject is shown both audio and visual stimuli depicting various sexual vignettes. As the individual becomes aroused with the stimuli, the erect penis displaces the gauge and a measurement is recorded. Phallometric testing helps elucidate an individual's sexual interest without depending on self-disclosure. One of the difficulties working with individuals who have paraphilias is the individual's reluctance to report his sexual interests. Objective testing (e.g., plethysmography) is one way of addressing the difficulty of self-disclosure in paraphiliacs. In addition to diagnostic evaluations, phallometric testing is helpful in determination of an individual's response to treatment.

DIFFERENTIAL DIAGNOSIS

The diagnosis of a paraphilia begins with the identification of the individual's sexual interest. It is important to understand an individual's sexual preference before the diagnosis of a paraphilia can be made. One of the common examples made in the area of paraphilias is to make a diagnosis based on an individual's behavior rather than on the individual's sexual preference. For example, the differential diagnosis for a child molester may include pedophilia, personality disorder, mental retardation, or substance abuse. In order to determine what diagnosis fits the child molester one has to understand what motivated the molester to have sexual contact with the child. If the molester describes sexual urges and fantasies

to sexual contact with a prepubescent individual, they meet the criteria for pedophilia. However, if the molester reports a history of learning difficulties consistent with mental retardation and explains his molesting behavior on his desire to "have sex with anyone," the diagnosis is mental retardation and possibly personality disorder. The differential diagnosis of paraphilias is simple if one understands what the sexual motivations of the subject are.

COURSE AND PROGNOSIS

Paraphilic behaviors emerge in adolescence and early adulthood. In general paraphilic behaviors are chronic. Although it has been shown that individuals may cross over from one paraphilia to another, the overall course of paraphilic behavior is chronic. Because of the early onset and repudiated biological priming of the paraphilias, they are often difficult to treat.

TREATMENT BACKGROUND

The efficacy of treatment in paraphiliacs is unknown. Empirical evidence for the treatment of individuals with paraphilias is derived from studies of sex offenders. Although studies comparing treated and untreated sex offenders have been done, measurement of outcome is flawed, with recidivism rates underestimating actual recurrence of the problematic behavior. Methodological challenges such as sample selection, study design, and assessment of outcome represent significant problems in treatment efficacy studies. The majority of treatment efficacy studies are conducted on incarcerated or civilly committed offenders. Those individuals who are incarcerated or civilly committed represent a unique population with treatment outcomes that may not generalize to the nonincarcerated or community residing sex offender. Many studies do not differentiate between types of offenders. For example, a study that evaluates child molesters may not differentiate between intra- and extrafamilial offenders. Research suggests that incest offenders recidivate at approximately half the rate of extrafamilial child molesters.

A well-designed study of treatment efficacy in sex offenders would compare treated offenders with a matched, randomly assigned control group. However, the ethical and public safety implications of such a study impede its design. Sex offenders are extremely heterogeneous, and when this is not recognized it is very difficult to make comparisons among treatment efficacy studies.

Biological Treatment

The scientific basis of biological treatment in paraphiliacs is the reduction of sexual behaviors by decreasing testosterone levels. Testosterone is a

steroid hormone produced by the testes that is responsible for the development of secondary sex characteristics in men. Testosterone helps maintain sex drive, the production of sperm cells, male hair patterns, muscle mass, and bone mass. Testosterone strongly influences both male and female sexual drive and the resultant sexual behavior. Animal studies show that male copulatory behavior is almost entirely dependent on circulating levels of plasma testosterone.

The production of testosterone is mediated by FSH and luteinizing hormone (LH), two hormones produced by the pituitary gland. LH signals the testes to produce testosterone. The hypothalamus senses whether the production of testosterone is too much or too little. If the testes are producing too little testosterone, the hypothalamus secrets a hormone called GnRH, which signals the pituitary to increase production of LH, which results in an increase in testosterone production. If the testes are producing too much testosterone, the hypothalamus secretion of GnRH is inhibited. This process, whereby the hypothalamus modulates the production of testosterone, is called negative inhibition.

Surgical Treatment

Surgical treatment for sex offenders consist of two types: Neurosurgery and castration. The neurosurgical procedure involves stereotaxic removal of parts of the hypothalamus to disrupt production of male hormones and decrease sexual arousal and impulsive behaviors. This procedure had significant adverse effects and was considered largely ineffective. Surgical castration is the removal of the testes. The effect of surgical castration is to globally reduce available androgens by the removal of the testes. Sturup followed 107 castrated sex offenders and compared them to 58 who were not castrated over a period of 18 years. The castrated individuals recidivated at a rate of 4.3 percent. The 58 uncastrated individuals recidivated at a 43 percent rate. Postcastration follow-up studies provide the most comprehensive outcome data on the effect of reducing plasma testosterone and the resultant suppression of the sexual drive and deviant sexual behavior. These studies have reported recidivism rates of less than 5 percent with follow-up periods of up to 20 years in a large sample size.

The introduction of antiandrogen and hormonal medications has rendered surgical castration nearly obsolete. The effects of surgical castration can be achieved through chemical castration (i.e., the use of medications to decrease testosterone production), without the invasiveness and irreversibility of surgery. Some states in the United States, such as Texas and California, mandate chemical or surgical treatment of dangerous sexual offenders. Controversy exists as to whether surgical castration should be offered as a treatment, as chemical castration achieves the same results and spares the procedure.

Pharmacologic Treatment

Antiandrogens. Androgens are male sex hormones that promote the development and maintenance of the male sex characteristics. Testosterone is the most common androgen in man. Antiandrogen treatment refers to treatment with drugs used to block production or interfere with the action of male sex hormones. Cyproterone acetate (CPA; Androcur) and medroxyprogesterone (MPA; Depo-Provera) are the two most commonly used antiandrogen medications. CPA is not available in the United States but is widely available in Canada and Europe. Both CPA and MPA are synthetic progesterones that reduce the serum level of testosterone. Reduction of testosterone has been shown to reduce libido, erections, ejaculations, and spermatogenesis. The World Federation of Societies of Biological Psychiatry (WFSBP) concluded in their guidelines for the biological treatment of paraphilias with MPA that there is minimal research based evidence to support this recommendation. They also reported severe side effects observed with MPA, concluding that the benefit/risk ratio did not favor the use of MPA. WFSBP similarly concluded that there is minimal research-based evidence to support treatment with CPA.

The risks associated with the use of antiandrogen agents include side effects such as weight gain, hyperglycemia, hot and cold flashes, liver dysfunction, hypertension, muscle cramps, phlebitis, gastrointestinal complaints, and feminization. In addition, there is little known about the long-term sequelae of antiandrogen treatment. The majority of antiandrogen studies follow patients for 8 years or less. The patient who agrees to antiandrogen treatment should be educated about the absence of long-term data.

Hormonal Agents

Leuprolide (Eligard) and triptorelin (Trelstar) are hormonal agents referred to as long-acting GnRH agonists. These agents inhibit the secretion of LH with a resulting decrease in plasma testosterone levels and libido. They produce a chemical castration in that the hypothalamic–pituitary axis is exhausted and there is a potent inhibition of gonadotropin. Meta-analysis conducted by WFSBP concluded that there were no controlled studies, and some biases were observed. However, the efficacy observed in these open studies was very high and in most cases, subjects were previously treated with psychotherapy or other antiandrogens without efficacy (WFSBP). The WFSBP reported that there is minimal research-based evidence to support the use of hormonal agents in the treatment of paraphilias.

The side effects associated with GnRH agonists include decreased bone mineral density or osteopenia, weight gain, hyperglycemia, diabetes, hypertension, and insomnia. The most commonly reported side effects are

erectile/ejaculatory problems and gynecomastia. In general, patients reported fewer side effects on GnRH agonists when compared to antiandrogens. The preliminary studies of GnRH agonists suggest that leuprolide and triptorelin may be more effective, better-tolerated alternatives to antiandrogen treatment.

Selective Serotonin Reuptake Inhibitors

Although the pharmacologic mechanism is poorly understood, SSRIs have been shown to be effective in reducing paraphilic symptoms. Animal models showed that decreased serotonin levels increased sexual appetitive behaviors while enhanced central serotonin activity reduced them (WFSBP, 2010). Clinical observations of the similarities between paraphiliac fantasies, urges and behaviors with obsessive compulsive symptoms led to hypotheses that SSRIs may be efficacious in paraphilias as they have been in OCD (WFSBP, 2010). Case reports have reported the positive effects of SSRIs in the treatment of paraphilias (Bradford, 1996; Saleh, 2004). Most of the studies examining the efficacy of SSRIs in paraphilias are case reports or small sample sizes and do not include placebo-controlled double blinds studies. The World Federation of Societies of Biological Psychiatry published the Guidelines for the biological treatment of paraphilias in 2010. These guidelines conclude that "a critical analysis of all studies that involved the use of SSRIs in the treatment of paraphilias concluded that the results of psychotropic drug interventions are not favorable (WFSBP, 2010). In summary, although data exist to suggest that SSRIs are effective in the treatment of paraphilias, presently there are insufficient data to conclude that SSRIs are equally efficacious as antiandrogens or hormonal agents.

Psychological and Behavioral Treatment

Psychological approaches to the treatment of sexual offenders focus on direct treatment approaches to modify offenders' cognitions and attitudes. Cognitive approaches described the concept of cognitive distortions (minimizations, excuses, justifications) as one factor that may maintain offender behavior. Relapse prevention and the sexual assault cycle are theoretical models purporting to describe the cognitions and behaviors of sex offenders before, during, and after their abusive acts. Relapse prevention models of treatment were the mainstay of treatment for many years. However in the past decade, these models have largely been abandoned for strength-based rehabilitation programs. The two most common strength-based rehabilitation programs that are currently employed in sex offender treatment are the Good Lives Model and the Risk Need Responsivity Model. Research has consistently demonstrated that structured behavioral or cognitive-behavioral treatment designed to

address offending behavior are more effective in reducing reoffending than didactic, experimental, nondirective or psychodynamic modes of treatment.

The rationale for the implementation of behavioral techniques to control sexual arousal is based on the clinical observation that abusive sexual fantasy is linked with abusive sexual behavior. The factor that most consistently distinguishes male sex offenders from other males is sexual arousal disorder profiles as measured by phallometric assessment. The modification of an individual's behavior is achieved by altering individual reactions to stimuli through positive and negative reinforcements. These reinforcements condition an individual by reinforcing positive behaviors and extinguishing negative behaviors. More specifically, the behavioral techniques used to control sexual arousal employ olfactory aversion conditioning, covert sensitization, and masturbatory satiation to extinguish deviant sexual arousal and reinforce appropriate sexual arousal.

Recidivism rates from cognitive and behavioral treatment programs range from 3 to 31 percent depending on the study. A meta-analysis found that cognitive-behavioral treatment and antiandrogen treatment were comparable in their treatment effects and significantly more effective than behavioral treatment alone.

MODELS OF TREATMENT

The treatment of individuals with paraphilias and sex offenders has primarily focused on two areas: Cognitive-behavior group treatment and pharmacologic treatment. Treatment programs target risk factors empirically associated with sexual offending and recidivism. Meta-analysis has shown that cognitive-behavioral treatment is most effective in reducing sex offender recidivism. Treatment models have changed over the years as research in the field of sex offender management has evolved. Currently cognitive-behavioral treatment remains the most widely accepted and empirically supported model of sex offender treatment. Best practices include the use of actuarial tools, data driven instruments, in the determination of risk. Research demonstrates that actuarial methods are consistent in predicting relative risk of sex offending behavior and are considered more accurate than unstructured professional judgment (Association for the Treatment of Sexual Abusers [ATSA], 2012). A research query asking sex offenders in treatment about acceptable types of treatment revealed the preferred therapies were individual psychotherapy, social skills training, and group therapy; whereas aversion conditioning, castration, and sex drive reducing drug therapy were the least acceptable forms of treatment.

ETHICAL AND LEGAL CONSIDERATIONS

The legal and ethical issues that arise in the course of sex offender

treatment are unique to this population. As such the traditional codes of ethics employed in medical treatment are not applicable to the treatment of sex offenders. A significant percentage of sex offenders receiving treatment have been mandated by the courts to do so as part of sentences of incarceration or release into the community. The concept of mandated or involuntary treatment raises the issue of whether informed consent is possible in sex offender treatment. A condition of informed consent for medical treatment is that the consent must be voluntary. Obviously, court-ordered sex offender treatment is not voluntary. Individuals who reject treatment are subject to punishment imposed by the courts. Individuals who have been court ordered to receive treatment do not have a choice regarding the type of treatment or the treatment provider. Sex offenders are required to complete particular programs, irrespective of any other treatment they might be receiving, in order to gain community release or avoid imprisonment.

In the medical model, patients have a right to refuse various types of treatment. The right to refuse treatment is based on an individual's constitutional right to privacy. The court views sex offenders as incompetent patients, and as such the court becomes the decision maker regarding treatment.

Confidentiality and Privilege

Confidentiality refers to the physician's obligation to keep information learned in a professional relationship private from other parties. Privilege refers to the patient's right to prevent a physician from providing testimony about personal medical information. Both confidentiality and privilege are routinely breached in sex offender treatment. When individuals enter treatment, they are required to give permission for their cases to be discussed with both clinical and nonclinical personnel, correctional officers, members of their family, past and potential victims, and those associated with them and fellow offenders.

In response to the deviations from traditional ethical codes inherent in sex offender treatment, ethical and practice guidelines were developed by the ATSA. ATSA's code of ethics endorses standards of professional conduct that promote competent practice, and as such, they represent a public commitment to clients and society toward the goal of preventing sexual violence. The ATSA guidelines state that the ethical care of sex offenders is achieved by encouraging individuals to take responsibility for their behavior, that is, admission of guilt. ATSA maintains that the identification and collaborative management of risk and safety factors are indeed in the best interests of both sex offender patients and potential victims due to the grave consequences incurred by sexual offender recidivism.

Castration of Sex Offenders

In 1996, California became the first state to authorize the use of either chemical or surgical castration for certain sex offenders who were being released from prison into the community. To date, an additional eight states authorize surgical or chemical castration for sex offenders, including Georgia, Montana, Oregon, Wisconsin, Florida, Iowa, Louisiana, and Texas. The APA stated that "these laws, which predicate release from prison on chemical castration by surgery or antiandrogenic agents, are objectionable because they are not based on adequate diagnostic and treatment considerations. They also improperly link medical treatment with punishment and social control." In some states, informed consent for the procedure is not required, in other states informed consent requires only that the offender be informed regarding the side effects.

FUTURE DIRECTIONS

Further Research

Although studies have demonstrated the effectiveness of some modalities of sex offender treatment, the scientific literature does not identify a definitive effective treatment. The literature suggests that both cognitive-behavioral therapies and antiandrogen medication decrease sex offender recidivism. A decrease in sex offender recidivism, however, does not translate to a decrease in paraphilic behaviors. However, the current sex offender legislation in the United States is not based on the literature. Rather the legislation focuses on the preventive detention of sex offenders without concurrent evidence-based sex offender treatment. The future of research in the treatment of sex offenders and paraphilias lies in the development of evidence-based knowledge that will inform both clinical decisions and public policy. This effort may be afforded by interdisciplinary collaboration between the scientific community, correctional professionals, and policy makers.

Specialized Training

Many psychiatrists are unfamiliar with the fundamentals of the assessment and treatment of individuals with paraphilias. The advent of laws for the civil commitment of sex offenders has resulted in the need for psychiatrists trained in paraphilias. As a result of sex offender legislation, the treatment and release decisions regarding sex offenders has become a focus of attention for psychiatry. Forensic psychiatrists or psychologists must render an opinion as to whether the sex offender has a diagnosed mental disorder and, as such, represents a risk to public safety if released from custody into the community. In order to provide competent care to individuals with paraphilias, psychiatrists should be educated about the

evidence-based approach to the treatment and evaluation of this population.

Education of the Public

The general public perceives sex offenders as a homogeneous group of offenders who are not treatable. However, recent meta-analyses have demonstrated that sex offenders are a highly heterogeneous population with different rates of reoffending and responsiveness to treatment. The successful treatment of sex offenders is dependent, in part, on the support of the public and the courts. In order to implement an evidence-based treatment, the offender must be afforded the opportunity for treatment. The opportunity for treatment is created by an informed society who legislates treatment that is informed by scientific evidence. The education of the public is integral in the successful treatment of sex offenders.

References

*American Psychiatric Association Task Force on Sexually Dangerous Offenders. *Dangerous Sex Offenders: A Task Force Report of the American Psychiatric Association*. Washington, DC: American Psychiatric Association; 1999.

Bancroft J. *Human Sexuality and Its Problems*. Edinburgh: Churchill Livingstone; 1989.

Birgden A, Cucolo H. The treatment of sex offenders: evidence, ethics, and human rights. *Sex Abuse*. 2011;23(3):295–313.

Bradford JM. The neurobiology, neuropharmacology, and pharmacological treatment of the paraphilias and compulsive sexual behavior. *Can J Psychiatry*. 2001;46(1):26–34.

Briken P, Hill A, Berner W. Pharmacotherapy of paraphilias with long-acting agonists of luteinizing hormone-releasing hormone: a systematic review. *J Clin Psychiatry*. 2003;64(8):890.

Briken P, Kafka M. Pharmacological treatments for paraphilic patients and sexual offenders. *Curr Opin Psychiatry*. 2007;20(6):609–613.

Brooks-Gordon B, Bilby C, Wells H. A systematic review of psychological intervention for sexual offenders. I. Randomised control trials. *J Forensic Psychiatry*. 2006;17(3):442–466.

Donovan BT. *Hormones and Human Behavior*. London: Cambridge University Press; 1984.

Grossman LS, Martis B, Fichtner CG. Are sex offenders treatable? A research overview. *Psychiatr Serv*. 1999;50(3):349.

Hanson RK, Bussière MT. Predicting Relapse: A meta-analysis of sexual offender recidivism studies. *J Consult Clin Psychol*. 1998;66(2):348–362.

Kafka M. Axis I psychiatric disorders, paraphilic sexual offending and implications for pharmacological treatment. *Isr J Psychiatry Relat Sci*. 2012;49(4):255–261.

Krueger RB, Kaplan MS. Depot-leuprolide acetate for treatment of paraphilias: a report of twelve cases. *Arch Sex Behav*. 2001;30:409–422.

Laws R, O'Donohue W. *Sexual Deviance: Theory Assessment and Treatment*. New York: Guilford; 1997.

Levenson J, D'Amora J. An ethical paradigm for sex offender treatment: response to glaser. *West*

Criminol Rev. 2003;4(2):143.

Rosler A, Witztum E. Treatment of men with paraphilia with a long-acting analogue of gonadotropin-releasing hormone. *N Engl J Med*. 1998;338:416–422.

Saleh FM. Serotonin reuptake inhibitors and the paraphilias. *Am Acad Psychiatry Law Newsletter*. 2004:12–13.

Scott CL, Holmberg T. Castration of sex offenders: Prisoners' rights versus public safety. *J Am Acad Psychiatry Law*. 2003;31:502–509.

Screenivasan S, Weinberger LE, Garrick T. Expert testimony in sexually violent predator commitments: conceptualizing legal standards of "mental disorder" and "likely to re-offend." *J Am Acad Psychiatry Law*. 2003;31(4):502–509.

Seto MC, Cantor JM, Blanchard R. Child pornography offenses are a valid diagnostic indicator of pedophilia. *J Abnorm Psychology*. 2006;115:610–615.

Sturup GK. Sexual offenders and their treatment in Denmark and other Scandinavian countries. *Int Rev Criminal Policy*. 1953;4:1.

Thibaut F, De La Barra F, Gordon H, Cosyns P, Bradford JM, WFSBP Task Force on Sexual Disorders. The World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for the biological treatment of paraphilias. *World J Biol Psychiatry*. 2010;11:604–655.

Ward T, Salmon K. The ethics of care and treatment of sex offenders. *Sex Abuse*. 2011;23(3):397–413.

▲ 21.4 Gender Identity, Gender Variance, and Gender Dysphoria

JACK DRESCHER, M.D., AND WILLIAM M. BYNE, M.D., PH.D.

Earlier versions of this chapter focused on what previous editions of the APA's Diagnostic and Statistical Manual (DSM) referred to as *Gender Identity Disorder (GID)*. The DSM-5, however, created a new parent diagnostic category known as *Gender Dysphoria*. That change arose, in part, due to public conversations surrounding the DSM-5 revision of the gender diagnoses. Similarly, proposed revisions of the gender diagnoses in the International Classification of Diseases and Related Health Problems (ICD) have stimulated much feedback to the WHO as it prepares to issue ICD-11 in 2018. Questions have been raised not only about what constitutes normal expressions of gender and nonpathological *gender variant* behavior or *gender variance,* but also who ultimately is responsible for deciding what is and is not normal: Psychiatrists? Religious authorities? Civil rights and human rights advocates? Further, not all cultures endorse *binary* notions of gender constituted simply of "men" and "women." Some notable examples include Native American "two spirit people" *(berdache)* or the *hijras* of India. Given these cultural shifts in contemporary discussions of gender, this chapter attempts to provide a more expansive view of gender variance and gender identity in addition to discussing psychiatric aspects of gender dysphoria and its treatment.

DEFINITIONS

Although *sex* and *gender* are often used interchangeably, here *sex* refers to the status of biological variables that can be described as either male-typical or female-typical in normatively developed individuals (e.g., genes, chromosomes, gonads, internal and external genital structures). Prior to the mid-1950s the word *gender* had been employed primarily to refer to classes of nouns in languages; however, John Money and his colleagues used *gender* in reference to attributes of people; specifically, those attributes that are associated with biological sex but not necessarily determined by it. For example, they observed that most individuals with somatic *disorders of sexual development (DSD),* historically referred to as *hermaphroditism* and *pseudohermaphroditism* and later as *intersex conditions,* often live and think of themselves as males or females according to how they were assigned in infancy and independent of the status of any particular biological variable of sex. *Gender,* therefore, came to refer to the social categories of males (i.e., boys, men) and females (girls, women) as well as to factors related to living in the social role of a man or a woman. The gender initially assigned to individuals with a DSD is referred to as *natal gender* (or *natal sex*), a term which has more recently been used to refer to the gender initially assigned to all individuals, particularly in the case of transgender individuals.

Gender identity, a term first coined by Robert Stoller in the mid-20th century, came to signify one's persistent inner sense of belonging to either the male or female gender category. Prior to that, John Money had defined *gender role* as those things that a person says or does to disclose him- or herself as having the status of boy or man, girl or woman, respectively (e.g., general mannerisms, mode of dress, interests, pastimes, deportment and demeanor; spontaneous topics of talk in unprompted conversation and casual comment). *Gender role* is thought of as the outward expression of the inner sense of *gender identity*. However, the terms *experienced gender* and *expressed gender,* respectively, are increasingly replacing gender identity and gender role, particularly in the legal and political arena concerned with antidiscrimination protections. Although the initial descriptions of gender role and identity implied two discrete gender categories, gender is increasingly regarded as more *fluid* reflecting a continuous dimension of masculinity/femininity as opposed to a dichotomy of masculine and feminine. Thus, gender identity/experienced gender may be described as male (boy/man), female (girl/woman), somewhere in between, or neither, while gender role/expressed gender may be described as masculine, feminine, or mixed. Notions of what is masculine or feminine vary across individuals, cultures and times.

Transgender is a popular term coined by Virginia Prince in the early 1970s to refer to people who lived full-time in a gender that was not the one that usually went with their genitals. In recent decades, it has become an

umbrella term that usually refers to someone whose gender identity or gender expression and natal gender are discordant or not conforming to social norms. It may be associated with a desire to transition to or live as a gender different from the natal gender (which could be male, female, or some other gender category such as *genderqueer* or *eunuch*). Members of the transgender community have coined the term *cisgender* (based on the concept of cis/trans isomerism in organic chemistry) as a way to describe nontransgender individuals whose gender identity and birth assigned sex are in alignment.

The concepts of changing the sex of the body and transitioning gender role dates back several millennia, notably found in the Greek myth of Tiresias who was born with the body of a man but changed by the goddess Hera into a woman for 7 years. Of more recent historical vintage, the protagonist of Virginia Woolf's novel *Orlando* goes to sleep as a man and inexplicably wakes up as a woman. Today, modern medical and surgical practices, as well as changing social mores, have shifted the prospect of changing the sexual appearance of one's body from a subject of mythology and fiction to a highly specialized clinical reality. The term *transsexual* refers to an individual who has, or plans to, employ hormonal and/or surgical means to modify the body so that it conforms to one's experienced gender. A natal male who transitions to live as a woman is called a male to female (MTF) individual, an MTF transsexual or transwoman. Conversely, a natal woman who transitions to live as a man is called a female to male (FTM) individual, an FTM transsexual or a transman. Historically, surgical transition was much more common among natal males but transition is increasing among natal females.

Historically referred to as *sex reassignment surgery (SRS),* today surgical transition procedures are referred to as *gender reassignment surgery (GRS)* or *gender confirmation surgery*. Surgical options include *top surgery* (breast augmentation for transwomen, mastectomy and male breast construction for transmen) and bottom surgery (gonadectomy and genitoplasty for transwomen, genitoplasty, gonadectomy, and hysterectomy for transmen). Transsexual individuals may not desire to undergo all of these surgical procedures, while financial or medical contraindications may prevent others from obtaining procedures they desire.

Similarly, not all individuals with cross-gender identifications desire, seek, or obtain hormonal or surgical treatments for transition. To find authentic gender expression, some may choose to transition in some respects but not others. That is, transition may involve wearing the clothing or accessories of the nonnatal gender and/or taking hormone supplements to acquire secondary sexual characteristics of the other gender. Some transsexual individuals object to being classified under the transgender umbrella, believing that those who desire to transition hormonally and

surgically belong to a category distinct from those who do not.

There is a great deal of popular confusion between transgender and homosexuality. However cross-gender identifications, in and of themselves, do not necessarily reveal anything about one's sexual orientation. Further, the terms *homosexual* and *heterosexual* can lead to confusion when there are discrepancies between an individual's natal (birth assigned) sex and experienced/expressed gender. In such situations, sexual orientation may be better described as *androphilic* (attracted to men), *gynephilic* (attracted to women), bisexual or *analophilic* (attracted to neither). For example, a large proportion of MTF individuals are gynephilic.

The presence of a marked incongruence between one's experienced/expressed gender and natal gender are listed among the criteria for the DSM-5 diagnosis of *gender dysphoria (GD) in adolescents and adults*. Importantly, *gender dysphoria,* historically defined as unhappiness or a sense of inappropriateness with one's natal sex or the gender role associated with that sex, must cause clinically significant distress or impairment in order to satisfy the DSM diagnostic criteria of a disorder (see Diagnoses on p. 2028). There is also a DSM-5 diagnosis of *gender dysphoria in childhood* (GDC). Longitudinal studies have found that many individuals meeting criteria for GDC do not have GD as adults. Instead some will identify as cisgender heterosexual adults, while the majority will identify as gay, "homosexual," or bisexual in adolescence or adulthood. However, most homosexual adults would not have met criteria for GD as children. In the forthcoming ICD-11, the gender diagnoses will be known as *gender incongruence of adolescents and adults* and *gender incongruence of childhood*. In addition, the diagnoses will be moved out of the mental disorders section of the ICD into a new parent category: *Conditions Related to Sexual Health* (see Diagnoses).

Transvestitism refers to sexual urges and fantasies involving cross-dressing. When the behavior is distressing to a patient, or brought to the attention of a psychiatrist, it may be diagnosed as *transvestic disorder,* which DSM-5 classifies as a *paraphilic disorder* (refer to Paraphilia Section 21.3). Cross-dressing is strongly linked in the popular imagination with homosexuality, however, in DSM-IV-TR, the diagnosis of *transvestic fetishism* could only be made in heterosexual men as most fetishistic cross-dressers are heterosexual men who cross dress in private. These men are often married to women who may or may not know about their cross-dressing interests. DSM-5 removed the restriction to heterosexual men, opening the diagnosis to women or gay men who cross dress out of erotic sexual interest. While the change could increase the number of people diagnosed with transvestic disorder, the diagnosis still requires that cross-dressing be a source of sexual arousal and that individuals experience significant distress or impairment because of their behavior. In any event, most gay men and lesbians do not cross-dress. However, there are social

venues within gay and lesbian communities that allow for public cross-dressing. These may include either socially performative events (gay pride, Carneval, Halloween, or Mardi Gras parades) or cross-dressing as a form of paid entertainment (drag shows). Some, although not all, cross-dressers identify as transgender.

THEORIES

Scientific and social theorizing about gender variance began in the middle of the 19th century. Although there are theories purporting to explain the origins of any gender identity, at present the manner in which either transgender or cisgender identities develop is unknown.

History

19th and Early 20th Century Theories. The early history of theorizing about gender variance was conflated with theories of what is today referred to as homosexuality. Many theorists routinely conflate homosexuality with transgender identities, relying upon binary beliefs that use conventional heterosexuality and cisgender identities as a frame of reference. For example, in the mid-19th century, German lawyer and activist, Karl Ulrichs, hypothesized that some men were born with a woman's spirit trapped in their bodies. He believed these men constituted a *third sex* and named them *urnings*. While historians of homosexuality unremarkably and routinely seem to regard Ulrichs' urnings as homosexual men, a female spirit in a male body closely resembles the narratives of 20th century theories of transsexualism.

Until the middle of the 20th century, with rare exceptions, transgender presentations were usually classified as "psychopathological." One early exception was Carl Westphal's 1870 account of *Contrary Sexual Feeling,* the case of a young woman who cross-dressed as a boy and who was attracted to women with whom she formed relationships. Richard von Krafft-Ebing, on the other hand, reinforced the psychopathological view in his landmark publication, *Psychopathia Sexualis,* an early psychiatric diagnostic manual. There he documented cases of individuals who desired to live as members of the other sex and those who had been born to one sex and were living as members of the other. In a chapter entitled "General Pathology," he presented what today would be thought of as transgender representations as cases of what he called *Metamorphosis Sexualis Paranoia*. In these examples, there was a significant conflation between homosexuality and gender presentations.

German psychiatrist Magnus Hirschfeld is credited with being first to distinguish the desires of homosexuality (to have partners of the same-sex) from those of transsexualism (to live as the other sex) but which he referred to as *transvestism*. However those distinctions were not broadly accepted

until decades later. Hirschfeld also assisted surgeons in Europe who were already experimenting with gender reassignment surgery in the 1920s.

Mid-20th-Century. The possibility of changing genders seized the popular imagination when American George Jorgensen went to Denmark as a natal man and returned to the United States in 1952 as transwoman Christine Jorgensen. In the wake of much public and professional controversy about Jorgensen's transition, Christian Hamburger and the other Danish physician who performed her surgery published a report of their treatment of her "transvestitism" in the *Journal of the American Medical Association*. The publicity surrounding Jorgensen's transition eventually led to greater popular, medical, and psychiatric awareness of a then little known concept that would eventually come to be known as *gender identity,* as well as recognition of an increasing number of people wishing to "cross over." For those who eventually would come to identify as transsexual, increased public discussions of gender reassignment and gender identity would provide them with a way to put a name to their feelings and desires. As a result, a presentation of gender once considered exceedingly rare gradually became more commonplace.

However, at the time of Jorgensen's 1950s transformation and for several decades thereafter, many professionals were critical of offering hormonal and surgical transition procedures. Many physicians and psychiatrists criticized using surgery and hormones to irreversibly—and in their view incorrectly—treat people suffering from what they perceived to be either a severe neurotic or psychotic, delusional condition in need of psychotherapy and "reality testing." It was in this cultural context that the first two editions of the DSM published in 1952 and 1968, respectively, did not include any gender diagnoses. The diagnosis of trans-sexualism (sic) first appears in 1975 in the ICD-9. In 1980, a newly revamped DSM-III offered new diagnoses based on a growing body of research on child and adult gender identity. Table 21.4–1 outlines the placement of gender diagnoses in the ICD and DSM manuals.

Table 21.4–1.
Gender Diagnoses in the ICD and DSM

Year	ICD	DSM	Parent Category	Diagnosis Name
1948	ICD-6[a]	—	N/A	N/A
1952	—	DSM-I	N/A	N/A
1955	ICD-7	—	N/A	N/A
1965	ICD-8	—	Sexual deviations	Transvestitism
1968	—	DSM-II	Sexual deviations	Transvestitism
1975	ICD-9	-	Sexual deviations	Transvestism Trans-sexualism

1980	-	DSM-III	Psychosexual disorders	Transsexualism Gender identity disorder of childhood
1987	-	DSM-III-R	Disorders usually first evident in infancy, childhood, or adolescence	Transsexualism Gender identity disorder of childhood Gender identity disorder of adolescence and adulthood, nontranssexual type
1900	ICD-10	-	Gender identity disorders	Transsexualism Dual-role transvestism Gender identity disorder of childhood Other gender identity disorders Gender identity disorder, unspecified
1994	-	DSM-IV	Sexual and gender identity disorders	Gender identity disorder in adolescents or adults Gender identity disorder in children
2000	-	DSM-IV-TR	Sexual and gender identity disorders	Gender identity disorder in adolescents or adults Gender identity disorder in children
2013	-	DSM-5	Gender dysphoria	Gender dysphoria in adolescents and adults Gender dysphoria in children Other specified gender dysphoria Unspecified gender dysphoria
2018 (anticipated)	ICD-11	-	Conditions related to sexual health	Gender incongruence of adolescents and adults Gender incongruence of children

[a]ICD-6 was the first iteration of the manual to include mental disorders.

The decisions to include *transsexualism* in the ICD and DSM were based on a growing body of research and clinical observations, most notably the work of John Money, Harry Benjamin, Robert Stoller, and Richard Green.

John Money, a psychologist and sexologist, first began publishing his theories regarding gender identity development in the 1950s. Based on studies of children born with DSDs, he theorized that one's sense of being male or female—what eventually came to be known as one's gender identity—was acquired and that acquisition was primarily determined by external, environmental factors. Citing cases of gender assignment in intersex children born with ambiguous genitalia, Money believed parental attitudes and rearing practices had a strong effect on whether a child would accept the gender category to which it had been surgically assigned. Although Money believed that biology also played a role, he believed that the role of the psychosocial environment was critical and that any biological

predisposition could be overridden if parents unambiguously reinforced the gender assignment. Based on his observations, he suggested that a person's gender identity was fixed by 3 years of age, and considered efforts to change a person's gender identity difficult if not impossible in anyone older. Pessimism about changing an adult's gender identity left only one therapeutic alternative for those dysphoric in their assigned gender: gender reassignment.

In the mid-1960s, in the wake of Money's theoretical work and his clinical and research findings, Johns Hopkins opened the first university-affiliated, multidisciplinary gender clinic offering gender reassignment to transsexuals seeking treatment. More than 40 academic centers in the United States would later open gender clinics as well. More recently, observations of gender identity outcomes in individuals with particular DSD syndromes has suggested that the influence of biological factors on gender identity is stronger than that suggested by Money.

Harry Benjamin, a physician, is credited with both popularizing the term *transsexual* in its current usage and for raising awareness about trans individuals within the medical profession. Benjamin offered transgender individuals hormonal treatment at a time when mainstream psychiatry and medicine regarded gender incongruent individuals as confused homosexuals, neurotics, transvestites, schizophrenic, or some combination thereof. Benjamin believed that transsexualism was a biological disorder arising, for example, in natal males due to brain feminization in utero. Consistent with his essentialist view, he believed psychotherapeutic attempts to change gender identity were futile. Among the first physicians to experiment with hormonal and surgical treatments for aging, he eventually pioneered the treatment of gender dysphoric individuals using sex hormones. Notably, he accomplished this in a private practice setting without either university or academic support. In acknowledgment of his early advocacy for the medical treatment of transsexualism, in 1979 the newly formed Harry Benjamin International Gender Dysphoria Association (HBIGDA) was named in his honor. HBIGDA, which would later change its name to World Professional Association for Transgender Health (WPATH), would go on to develop *Standards of Care (SOC)* for treating gender dysphoric individuals. The WPATH Standards of Care are periodically revised and are presently in their seventh version.

Robert Stoller was an American psychiatrist and psychoanalyst whose theorizing, like Money's, grew out of clinical experience with both transsexual individuals and individuals with somatic intersexuality. Stoller is credited with introducing the concept of *gender identity* into both the psychoanalytic and psychiatric literature. In contrast to Benjamin's essentialist views, Stoller believed that in some cases, childhood family dynamics were responsible for "causing" adult transsexualism. Stoller, influenced by the separation-individuation theories of Margaret Mahler,

opined that gender dysphoria in prepubescent boys was a developmental arrest due to an excessively close and gratifying mother–infant symbiosis that prevented the child from adequately separating himself from his mother's female body and feminine behavior.

Richard Green, who trained with both Money and Stoller, most notably studied cross-gender behavior in prepubescent children. His book, *The "Sissy Boy Syndrome" and the Development of Homosexuality,* was a prospective study that tracked into adulthood the development of 66 gender-atypical boys who stated a wish to be a girl. Seventy-five percent of the children Green studied grew up to be gay men. Green served on the DSM-III Subcommittee on Psychosexual Disorders that recommended including the diagnoses of *transsexualism* and *GID of childhood* in 1980's DSM-III. The childhood diagnosis would subsequently make an appearance in 1990's ICD-10 as well.

Transgender Emergence. Following the publication of a 1979 John Hopkins study by Jon Meyer and Donna Reter citing "no objective improvement" in MTF transsexuals who had undergone gender reassignment surgery, Hopkins closed its gender clinic. Importantly, the conclusion of "no objective improvement" was based on comparison of an "adjustment score" before and after transition. This score was a composite of variables which seem of questionable relevance to resolution of gender dysphoria today including legal and economic considerations as well as marriage or "gender-appropriate" cohabitation. Furthermore, the study concluded that gender reassignment surgery was subjectively satisfying to recipients. Nevertheless, following publication of this study all but 3 of the more than 40 university-affiliated US gender programs closed. This dramatically changed the delivery of medical care for transsexualism in the United States. Prior to the closing of these clinics, in the 1960s and 70s, gender reassignment performed at university-based gender clinics was usually conducted with the aim that the recipient blend into a male/female binary mainstream. In the absence of university protocols, those seeking treatment had what transgender activist and writer Dallas Denny calls an "open market" for the treatment of their gender dysphoria, as those desiring surgical and hormonal transition services were no longer screened for the potential to blend into the cisgender mainstream. Further, transgender individuals began to seek each other out as they were no longer constrained by the privacy requirements of gender clinics that kept them from meeting each other. The mid-1980s saw the creation of support groups and conferences at which they would meet and exchange ideas and experiences. By the 1990s, transgender individuals began to produce a scholarly literature of their own. Also starting in the 1990 s, many of the existing gay, lesbian and bisexual (LGB) civil rights advocacy groups began a process of *trans inclusion* in recognition of the similarities in types of

prejudice and discrimination faced by the entire LGBT community. Today openly transgender individuals, be they actors, models, military personnel or even corporate CEOS, are increasingly in the public eye. In recent years, antidiscrimination protections and/or hate crime laws for gender identity and gender expression have been enacted at the statewide level in California, Colorado, Connecticut, Hawaii, Illinois, Iowa, Maine, New Jersey, New Mexico, Oregon, Pennsylvania, Vermont, and Washington. In 2012, the American Psychiatric Association issued position statements in support of transgender access to care and opposing discrimination against transgender individuals. Over the past 20 years, outcome studies following gender reassignment have focused on subjective variables such as relief of gender dysphoria, subjective sense of wellbeing, and satisfaction with transition as opposed to regret. Such studies have found the process of gender reassignment to be effective in lessening of suffering and increasing subjective satisfaction, with major regrets being rare.

The last two decades have also seen a growing public awareness and acceptance of gender variant minors, be they prepubescent children or adolescents. Where once the predominant psychiatric approach for prepubescent children was to dissuade them from their gender atypical interests, alternative treatment approaches have evolved (discussed later in this chapter). There has also been an increase not only in families allowing their children to socially transition before puberty, but in schools and school boards willing to accommodate those transitions.

Mental Disorder or Normal Gender Variance? Following greater visibility, and paralleling gay activist protests of the mid-20th century, an increasing number of transgender identified individuals and their allies challenge the characterization of transgender, transsexualism, and gender dysphoria as mental disorders. Just as the gay community once argued for removal from an earlier diagnostic nosology (see Homosexuality Section 21.2), the transgender community has adopted similar normalizing arguments to make the case for removal. These include adopting normalizing etiological theories, such as the belief that one is born transgender; adopting a transhistorical approach that connects modern transgender identities to historical figures and cultures; using modern cross-cultural studies to show that antitransgender attitudes are culture bound; using statistics regarding prevalence to refute the notion that transgenderism is rare; underscoring the difficult, if not impossible task of changing a gender identity; and adopting normative language to replace medical terminology.

In response to these social changes, there was much public debate about the retention of gender diagnoses in the DSM-5, starting with the appointment of the DSM-5 Workgroup on Sexual and Gender Identity Disorders in 2008 up until the publication of DSM-5 in 2013. The decision

to retain the diagnoses in DSM-5 was in part based upon the need to maintain access to care for individuals seeking medical and surgical treatment; however, the name was changed to Gender Dysphoria which is more descriptive and less stigmatizing. While the use of V-codes was suggested and considered, this solution was ultimately rejected as limiting access to care since third party payers do not reimburse for V-codes.

Similar concerns about perpetuating the stigma of a mental disorder have been raised as the World Health Organization (WHO) is in the process of revising the *International Classification of Disease* (ICD) with the ICD-11 having an anticipated publication date of 2018. The Working Group on the Classification of Sexual Disorders and Sexual Health (WGSDSH) is charged with reviewing and evaluating clinical and research data informing gender diagnoses. As WHO is part of the United Nations, in addition to its medical mission, it also has a human rights mission. Consequently, many advocates, several countries, the Council of Europe Commissioner for Human Rights and the European Parliament have taken strong positions that issues related to transgender identity should not be classified as mental disorders in the ICD-11.

While reducing stigmatization associated with mental disorders is important, the argument to remove a diagnostic category from the mental disorders section of the ICD simply because mental disorders are stigmatized is neither compelling nor persuasive. Psychiatric illnesses, like "social deviance," often create conflicts between individuals and society. Consequently both the psychiatrically ill and minority groups are subject to stigma. Stigmatization of individuals with psychiatric disorders is a social problem across cultures. WHO, health care policy makers, psychiatrists and other mental health professionals are actively engaged in trying to reduce the stigma of mental disorders in order to increase access to care and to encourage people to avail themselves of mental health services. Further, mental disorders are health conditions which is why they are in the ICD. One unintended consequence of belaboring distinctions between medicine and psychiatry is the perpetuation of existing stigma and prejudices against the mentally ill.

Nevertheless, the combined stigmatization of being transgender and of having a mental disorder diagnosis creates a doubly burdensome situation for this population that contributes adversely to their health status as well as to their attainment of human rights. For example, transgender people are much more likely to be denied care in general medical or community-based settings given the perception that they must be treated by psychiatric specialists, even for conditions that have nothing to do with being transgender. Further, there are unique circumstances in the case of this particular diagnosis that relate to the ability of a person to be viewed as competent to make certain legal decisions. Government agencies in many countries have demonstrated prejudice in extending recognition of change

in legal gender status on identity documents such as passports, birth certificates and drivers' licenses. Courts may often rule against transgender people in child custody decisions in cases where being diagnosed with a mental disorder is used by an ex-spouse to call their competence as parents into question.

It should be underscored that from a historical perspective, the classification of gender identity diagnoses as mental disorders was serendipitous. It appears to have been based more on prevailing social attitudes at the time than of available scientific evidence. For example, if the social narrative at that time had been that transsexualism was related to a "hormone imbalance" rather than being a "sexual deviation," the result could very well have been that transsexualism would be placed in the ICD-10 chapter on "Endocrine, nutritional, and metabolic diseases" today. In fact, the etiology of the condition was unknown when placement decisions were made in the past and remains unknown now. There are no scientifically based criteria to differentiate normal and pathological gender identity and the manner in which *any* gender identity develops remains unknown and a matter of theoretical speculation. The extant scientific database cannot empirically answer the question of whether this diagnosis is purely a "mental" phenomenon as opposed to a biological one. A number of studies posit physiological rather than mental causes of transgender presentations. In addition, many of the treatments prescribed today (e.g., hormonal, surgical, voice therapy, depilation) could not be construed as conventional mental health treatments. That is, today's approaches to gender dysphoria involve changing the body and social role rather than changing the individual's psyche. There is no reason to expect that this approach will change in the foreseeable future.

Yet based on the classification of gender incongruence as a mental disorder, as well as the need to distinguish it from other mental disorders, many countries require psychiatrists and other mental health professionals to act as *gatekeepers* to transition services for transgender people. In some countries, transgender people wishing to transition must prove to psychiatrists that they are the "right" kind of transgender and that they have been so for a sufficiently long period of time in order to qualify for services. This has created some controversy, as many individuals seeking transition do not otherwise have a mental disorder or desire mental health treatment, and the gatekeeping function may be seen as an unnecessarily burdensome requirement. How the gatekeeping role is performed has contributed to considerable tension between mental health professionals and transgender advocacy groups. The gatekeeping role has also sometimes interfered with positive contributions that psychiatrists and other mental health professionals can make to the transition process. Counseling and therapeutic approaches may be very important in helping an individual prepare for and cope with the personal and social effects of transition. In

this regard, the role of mental health professionals is not fundamentally different from their participation as a part of assessment and treatment for individuals undergoing other complex, life-changing medical treatments, such as organ transplantation.

Consequently, in addition to recommending a name change to Gender Incongruence, WHO's Working Group on Sexual Disorders and Sexual Health has recommended that the diagnoses be removed from the ICD-11's section on Mental and Behavioral Disorders. The WGSDSH's preferred recommendation, that the diagnosis be moved to an entirely separate chapter containing no other diagnoses, was not accepted. Instead, the diagnosis will be moved out of Mental and Behavioral Disorders into a new chapter called *Conditions Related to Sexual Health*. A recommendation from advocacy groups that the diagnoses be listed as Z-codes (similar to DSM-5's V-codes) was rejected for similar reasons: third-party payers, including national health services, do not reimburse for the treatment of Z-codes.

The retention of a diagnosis for adolescents and adults has not been as controversial an issue as the retention of a diagnosis for prepubescent children. Even removing the diagnosis from ICD's mental and behavioral disorders section has not satisfied some advocacy groups who now argue that gender variant children are being stigmatized with a *medical diagnosis* where previously they complained of the stigma of a *psychiatric diagnosis*. These arguments have not been persuasive either. Further, children with gender incongruence require access to care that is often complex and involves treatment of both the family and social environment. Consequently, the proposed ICD-11 diagnosis of Gender Incongruence in Childhood seeks to narrowly define the category so that gender atypical presentations, in and of themselves, are not considered diagnosable conditions.

FREQUENCY

The prevalence of transsexualism specifically, or transgender and gender variant identities in general, is unknown for a variety of reasons including concealment of these identities due to stigma as well as methodological problems in defining the populations of interest and designing adequate sampling strategies. Researchers who study incidence and prevalence have tended to focus on transsexual individuals with gender dysphoria who present to specialized gender clinics. Most such prevalence studies have been conducted in European countries and figures range from 0.002 to 0.008 percent for natal males who transition and from less than 0.001 to 0.003 percent for female-to-male (FtM) transsexuals. Since not all who seek hormonal and surgical reassignment attend specialty clinics, the above rates are likely underestimates of transsexual individuals and clearly underestimates of transgender individuals who either do not desire or do

not seek transition services. Studies that employ more inclusive sampling methods not restricted to specialty clinics tend to yield higher prevalence figures. For example, in a Massachusetts random phone survey that included over 28,000 individuals, 0.5 percent identified as transgender. A Dutch population sample found that 1.1 percent of natal males and 0.8 percent of natal females reported an incongruent gender identity while a study of New Zealand high school students reported that 1.2 percent identified as transgender regardless of natal gender.

In nonreferred children and adolescents, prevalence is believed to be less than 1 percent. In most studies of children, sex ratios of boys to girls range from 2.02:1 to 4.49:1. In adolescents, the sex ratio is close to parity; in adults, the sex ratio in most studies favors natal males, with ratios ranging from 1:1 to 6.1:1. In at least two countries, however, the sex ratio appears to favor natal females (Japan: 2.2:1; Poland: 3.4:1). In recent years, gender clinics treating children report a growing number of natal females.

DIAGNOSIS

DSM-5

The recent revision of DSM-5 involved a diagnostic name change from GID and the creation of a new overarching parent category known as Gender Dysphoria (GD). One rationale for this change was to remove the stigmatizing term "disorder" from the diagnostic label. The term "dysphoria" also captures the aversive emotional component of the clinical presentation as well as shifting clinical focus from "identity" per se to patient distress.

In addition, GD was moved out of an earlier parent category of Sexual and Gender Identity Disorders (where it was clustered with paraphilias and sexual dysfunctions) and into a separate category of its own. Other DSM-5 revisions included a new specifier, "with a disorder of sex development (DSD)," for individuals who have intersex conditions and develop symptoms of GD. This is in contrast to previous iterations of the DSM where the presence of a somatic intersex condition was considered an exclusionary criterion for making a GID diagnosis. The DSM-5 also includes a new "posttransition" specifier for older adolescents and adults who have transitioned to full time living in the desired gender. Unlike the DSM-III and IV, DSM-5 removed the sexual orientation specifier as it presently has no clinical relevance in the treatment of GD.

The DSM-5 now includes four GD diagnoses: GD in Adolescents and Adults, GD in Childhood, Other Specified Gender Dysphoria, and Unspecified Gender Dysphoria. The first two are described in greater detail later in this chapter.

Gender Dysphoria in Children. The diagnosis can only be made in

children who meet two criteria, A and B, in the DSM-5. Eight potential manifestations of criterion A are listed, of which at least six, must be met. One of these six criteria must be "a strong desire to be of the other gender or an insistence that he or she is the other gender (or some alternative gender [that is neither conventionally male nor female] different from one's assigned gender)." The other A-symptom criteria include preferring cross-dressing and rejecting conventional attire of one's assigned gender; preferring cross-gender roles in make-believe play; preferring toys, games, or activities stereotypically associated with the other gender; and preferring play with children of the other gender. Natal boys typically reject masculine toys, games, and activities and avoid rough-and-tumble play; natal girls typically feminine toys, games, and activities. These children frequently express a dislike of their sexual anatomy and may express the desire for the sex characteristics, primary and/or secondary, of the experienced gender. In addition to meeting six A criteria, Criterion B must be met which requires the presence of clinically significant distress or impairment in social, school, or other important areas of functioning. While the presence of a DSD was an exclusionary criterion for diagnosing GID in previous editions of DSM, the DSM-5 diagnosis includes a new specifier: "With a DSD." When this specifier is used, both the specific DSD as well as the GD diagnosis should be coded.

The differential diagnosis of GD in childhood includes simple atypical gender behavior in the absence of clinically significant distress or impairment. The diagnosis is not meant to merely describe nonconformity to stereotypical gender role and behavior such as "tomboyism" in girls or "girly" or "sissy" behavior in boys. The majority of such children are comfortable with their bodies and assigned genders and remain so in adolescence or adulthood. GD must also be distinguished from gender concerns that may arise as epiphenomena of coexisting disorders, including psychotic disorders, as well as a stated desire to be the other gender due to some perceived social advantage of being that gender or to escape a perceived disadvantage of the assigned gender (e.g., following sexual trauma). In cases where a child experiences gender dysphoria during periods when psychotic symptoms are in remission, schizophrenia and gender dysphoria may both be diagnosed. GD in Childhood has different criteria sets than GD in Adolescents and Adults and most children with GD in Childhood do not manifest GD in Adolescence and Adulthood.

Gender Dysphoria in Adolescents and Adults. The DSM-5 diagnosis of Gender Dysphoria (GD) in Adolescents and Adults has six A-symptom criteria. To establish the diagnosis, at least two of any of the six are required to be present for at least 6 months. These include a pronounced *incongruence* between one's experienced or expressed gender and the primary/secondary sex characteristics of the *assigned* or *natal gender*. For

prepubescent adolescents or those in puberty's early stages, this symptom is expressed as aversive anxiety to *anticipated* secondary sex characteristics. Another criterion is a strong desire to be rid of the natal gender's primary and/or secondary sex characteristics due to incongruence with one's experienced/expressed gender as opposed to other motives. For prepubescent adolescents or those in puberty's early stages, there is a strong desire to prevent further development of anticipated secondary sex characteristics. A third criterion is a strong desire for primary and/or secondary sex characteristics of the other gender. A fourth is a strong desire to be either of the other gender or some alternative gender, neither conventionally male nor female, rather than one's assigned gender. A fifth is a strong desire to be treated by others as a member of the other gender or some alternative gender. The sixth and final criterion is a strong conviction that one's feelings and reactions are more typical of the other gender or some alternative gender to the assigned gender.

There is also a B criterion of clinically significant distress or impairment in function, in social, occupational or other areas. As in the child diagnosis, there is a new specifier, "With a DSD," that requires coding a specific DSD condition, if present, in addition to the GD diagnosis. There is also a new "posttransition" specifier for individuals who have transitioned to living full-time in the desired gender, regardless of legal name change, who have undergone or plan to have at least one cross-sex medical or surgical procedure.

The differential diagnosis of GD in Adolescents and Adults includes nondiagnosable *atypical gender behavior; transvestic disorder* (a paraphilia in which individuals are sexually aroused by cross-dressing—some individuals may meet criteria for both transvestic disorder and GD); *body dysmorphic disorder* (involving a wish to alter or remove a specific body part perceived to be unattractive or abnormally formed) and *schizophrenia* and *other psychotic disorders* (in which delusions of being another gender may present). Importantly, if the full criteria for GD are met in an individual with a psychotic disorder when the psychosis is in remission, the diagnosis of GD can still be made as is the case if the GD preceded the onset of the psychotic illness. Some men adopt a *male eunuch* identity, a rare condition in which they seek castration to ameliorate psychological or physical effects of androgens yet still identify as male.

ICD-11

Proposals for forthcoming revision of the World Health Organization's (WHO) International Classification of Diseases (ICD) differ from those of the DSM-5 in several ways. First, GID of Childhood will be renamed *Gender Incongruence (GI) of Children* while Transsexualism will be renamed *Gender Incongruence of Adolescents and Adults*. Second, both diagnoses will be removed from the Mental and Behavioral Disorders

section of ICD and will be part of a new section entitled Conditions Related to Sexual Health.

Gender Incongruence of Childhood. ICD-11's Gender Incongruence (GI) of Childhood is characterized by a marked incongruence between an individual's experienced/expressed gender and the assigned sex in prepubertal children. A marked incongruence between the child's experienced/expressed gender and the child's assigned sex must be manifested by the following three indicators: (1) a strong desire on the child's part to be a different gender than the assigned sex, or insistence that he or she is a gender different from one's assigned gender; (2) a strong dislike on the child's part of his or her sexual anatomy or anticipated secondary sex characteristics and/or a strong desire for the primary and/or anticipated secondary sex characteristics that match the experienced gender; and (3) make-believe or fantasy play, toys, games, or activities and playmates that are typical of their experienced rather than their assigned sex. In addition, the incongruence must have persisted for about 2 years so the diagnosis cannot be made before approximately age 5. The diagnosis can only be assigned to children before puberty.

Gender Incongruence of Adolescence and Adulthood. Gender Incongruence (GI) of Adolescence and Adulthood is characterized by a marked and persistent incongruence between an individual's experienced gender and the assigned sex. The diagnosis cannot be assigned prior to the onset of puberty. The diagnosis often leads to a desire to transition, in order to live and be accepted as a person of the experienced gender. Establishing congruence may include hormonal treatment, surgery or other health care services to make the individual's body align, as much as desired and to the extent possible, with the experienced gender.

Diagnostic guidelines include a marked incongruence between the individual's experienced gender and the assigned sex, as manifested by at least two of the following four criteria: (1) a strong dislike or discomfort with one's primary and/or secondary sex characteristics (in adolescents, anticipated secondary sex characteristics) due to their incongruity with the experienced gender; (2) a strong desire to be rid of some of all of one's primary and/or secondary sex characteristics (in adolescents, anticipated secondary sex characteristics) due to their incongruity with the experienced gender; (3) a strong desire to have the primary or secondary sex characteristics of the experienced gender; and (4) a desire to be treated (to live and be accepted) as a person of the experienced gender. The diagnosis cannot be assigned prior to the onset of puberty and the gender incongruence must have been continuously present for at least several months.

ORIGINS OF GENDER DYSPHORIA

Because gender incongruence does not necessarily lead to clinically significant distress, understanding the etiology of gender dysphoria requires understanding both the acquisition of gender identity and how a mismatch between assigned and experienced gender lead to clinically significant distress in some individuals. This section examines theories about the acquisition of gender identity, both congruent and incongruent with natal sex, and then about the origins of distress resulting from gender incongruence in different age groups.

Acquisition of Gender Identity

Studies focusing separately on biological and psychosocial factors suggest that both contribute to gender identity, but the precise role of each, their interactions, and the contribution of the individual to the process remain incompletely understood. Early interest in gender identity focused on the infant/toddler's cognitive abilities during the period of gender identity formation and elucidating the limits of the developmental window of this period. The studies of John Money and his collaborators, briefly discussed previously in this chapter, suggested a period of maximal sensitivity for the acquisition of gender identity that ends at approximately age 2. They found that among children who experienced gender reassignment (usually in the context of a DSD or penile injury) after the neonatal period, 11 of 14 adjusted to the reassignment without complications if the reassignment occurred prior to 27 months of age compared to only one of four who adjusted to the reassignment if it occurred after 27 months. Subsequent work of Money and others suggested that gender identity may be influenced beyond this period but that after some imprecisely defined time, it becomes more or less impervious to change.

Infants begin to categorize gender-relevant attributes prior to the development of expressive language. For example, at 9 months infants can associate male and female faces with male and female voices, respectively, looking longer at a female than a male face when a female voice was presented and vice versa. By 18 months infants associate gender labels such as "lady" or "man" with female and male faces. As language develops, by 15 to 24 months toddlers can label adults as male or female and by 30 months can label peers and self appropriately by gender. As gender labels emerge, in the majority of individuals, a preference for play with peers of the same gender emerges and boys and girls tend to ascribe more positive attributes to their own gender and more negative attributes to the other. Some toddlers, however, show a preference for peers of the other gender, and may even label themselves as the other gender as early as between the ages of 2 and 3.

By age 2, children can recognize their own genitals when shown

anatomically correct dolls. Although by age 3 they may be aware that boys have penises and girls do not, less than half were found to understand this anatomic difference as the basis for gender categorization even by age 5. Between the age of 4 and 6 many children have yet to master the concept of *gender constancy* (an understanding that external changes in appearance or activity do not change one's gender identity) and show confusion when asked whether a change in behavior or clothing (such as a boy playing with a doll and wearing a dress) will lead to change in gender. Similarly, they may have yet to master the concept of *gender stability,* the understanding that for most individuals, gender remains constant throughout life. It is not until age 6 or 7 that nearly all children understand genitals as the basis for gender categorization and that gender for most people is constant and stable. These cognitive limitations leave open the possibility of a world for infants and young children where the feeling that one belongs to a gender other than the natal one or could become that gender would not be immediately countered by awareness that, for most people, one's natal gender is retained for life. For example, until the age of 6 or 7 a natal female who does not understand the association between external genitalia and gender and may believe that she could grow up to be a man.

Role of Psychosocial and Biological Factors. The permissive, direct and indirect/interactional models for conceptualizing the potential roles of biological and psychosocial factors outlined for the development of sexual orientation may apply to gender identity as well as well.

MODEL OF PERMISSIVE BIOLOGICAL EFFECTS. In this model biology plays a permissive role by providing the neural machinery through which gender identity is inscribed by formative experience. A permissive role could also include delimiting the developmental stage during which the relevant formative experiences must occur. For example, some songbirds must learn their species' song by hearing it during a restricted period of early development. While the song is clearly acquired through experience, biology determines the sensitive period during which that experience must occur. Once a particular song has been acquired, that will be the bird's song for life; it will not be able to unlearn that song or to acquire another. This is not meant to imply that gender identity might be acquired by simple mimicry. Instead it shows that phenomena acquired through experience may, nonetheless, be immutable.

MODEL OF DIRECT BIOLOGICAL EFFECT. In this model, the actions of biological factors such as genes or hormones would directly influence the organization or activity of brain circuits that mediate gender identity. The brain is viewed as possessing an intrinsic (i.e., constitutional) gender identity or predisposition, the expression of which could nonetheless be subsequently modified by experience.

MODEL OF INDIRECT BIOLOGICAL EFFECTS. This model addresses the contribution of individual differences, some of which may be constitutionally influenced. In this model, biological factors influence gender identity only indirectly as the more direct or immediate action of biological factors is upon temperament or other personality traits. From birth, these personality or temperamental traits influence how an individual experiences, interacts with, and modifies the environment. This would include how an individual participates in shaping the relationships and experiences thought by some to influence the development of a gender identity. Although similar to the model of permissive biological effects, the model of indirect biological effects goes further and suggests that the relevant formative experiences may, themselves, be strongly affected by biologically influenced personality variables.

The impact of psychosocial factors in the acquisition of gender identity has been documented in a large body of research in developmental psychology on normative development. These factors include verbal labeling (e.g., as a boy or girl) and nonverbal gender-cuing (e.g., mannerisms, clothing, hair length and style) of children by parents and others in their social environment, the shaping of gendered behavior by positive and negative reinforcement and later by explicit statements of gender-role expectations. Problems in parent–child attachment interacting with temperamental dispositions in the child have been suggested to be causally implicated in GD and have been cited as a focus for psychodynamic therapy of the child. Child psychologist, Susan Coates, observed higher levels of psychopathology in clinical samples of primary caregivers and suggested that parental psychological abnormalities may contribute to GD. These observations, however, do not distinguish between cause and effect.

Review of the clinical data on individuals with somatic intersexuality leads to a conclusion similar to that suggested by Money and his collaborators. By and large, individuals with functional androgen receptors and prenatal exposure to male-typical levels of androgens appear to have the capacity at birth to develop either a male or female gender identity depending on the gender of rearing. When assigned female, these individuals nevertheless have an increased likelihood of exhibiting masculinized play preferences, gender dysphoria and some ultimately reject the female identity and role. The proportion that ultimately rejects the female assignment is unknown because this can occur quite late in life, varies across DSD syndromes, and relatively few studies have followed and adequately assessed affected individuals in adulthood. The probability of switching to a male gender identity and role is further increased in androgen-sensitive individuals whose testes are left in place until puberty. These data are open to a variety of interpretations including that androgens act by modifying hypothetical brain circuits that mediate gender identity

(direct model) or that androgens bias behavioral or temperamental propensities in a manner that facilitates the acquisition of a male identity (indirect model). In addition to acting prenatally, the elevated androgen secretion that occurs in the neonatal period and again at puberty might reinforce the effects of prenatal androgens. At puberty, the psychological impact of somatic virilization in response to elevated androgens must be considered in addition to the possibility of physiologic effects on the brain. Since the mid-1950s the majority of individuals with somatic intersexuality have been assigned female. Thus there are inadequate data to base speculations about gender development on studies of individuals with low prenatal androgen exposure who were assigned male.

A hypothesis that appears to be growing in popularity is that cross-gender identification results from a brain-limited form of intersexuality, that is, that in transgender individuals some brain regions or circuits that are involved in gender identity differ between cisgender men and women and are sexually discordant with the rest of the body in transgender individuals. This hypothesis is derivative of Ulrichs' 19th century notion of a spirit of one sex trapped in the body of the other sex. Consistent with this hypothesis, Dick Swaab and collaborators in the Netherlands reported that a subcomponent of a brain structure called the bed nucleus of the terminalis varies in size as a function of gender identity rather than natal sex, a study that has yet to be replicated. Other studies of transsexuals have described other brain differences, both anatomical and neurophysiological, but the evidence at this point is insufficient for categorizing transgender people as having a form of CNS-limited intersexuality. Similarly, the literature on concordance for GD in monozygotic and dizygotic twins, lends some support for a role of genetic factors, especially in natal males. Due to small sample sizes and the absence of genotype studies, however, definitive conclusions cannot be drawn.

Autogynephilia in Natal Males. Psychologist and sex researcher Ray Blanchard has hypothesized two major developmental pathways leading to gender dysphoria in male to female adult transsexuals. In this theory, "homosexual transsexuals" are effeminate, gender atypical natal boys who are attracted to partners of their own natal gender. That is to say they are *androphilic* or attracted to men. The other group consists of *gynephilic* transsexuals (attracted to women) who are usually not overtly gender atypical as children and who often spend much of their lives in relationships with women. The latter group often has a history of cross-dressing that first starts out as a means of sexual arousal and may go on to develop *autogynephilia,* in which a natal male finds the fantasy of being a woman to be sexually arousing. Although also attracted to women, these individuals are often aroused at the fantasy of having sex or actually engaging in sex with men who treat them as women. Cross-dressing,

arousal to cross-dressing or fantasies of being female are not classified as psychiatric disorders unless they cause clinically significant distress or impairment in important areas of functioning (see Paraphilias Section 21.3). In some cases, individuals with autogynephilia later develop gender dysphoria and seek gender reassignment. Posttransition, they may prefer to engage in relationships with cisgender or other transsexual women.

Although the theory of an autogynephilic pathway to transsexualism has been in the research literature for decades, in recent years it has evoked controversy among some segments of the transgender community. They find offensive the linkage of transgender identity with a paraphilia as well as the linkage of the desire to transition gender to sexual arousal. There have also been recent scholarly critiques of the concept. For example, Charles Moser, a sexual medicine expert, surveyed 29 natal women and found 93 percent of them met autogynephilia diagnostic criteria. This finding is at odds with Blanchard's theory that the condition does not occur in natal women and is a type of "male sexuality." Moser argues that the meaning of a sexual interest in one's own body is not well understood for men, women, or transsexuals. Trans activist and biologist Julia Serano notes that a correlation between cross-gender sexual arousal and gender dysphoria does not imply causation. She notes (1) there are exceptions to the two subtype taxonomy (e.g., large numbers of androphilic MTF individuals who experience cross-dressing arousal); (2) that many transgender people experience cross-dressing arousal as a passing phase rather than a pressing paraphilic sexual drive leading to gender dysphoria; and (3) that many nonandrophilic MTF transsexuals experienced the desire to be the other gender long before they experienced cross-gender arousal.

Origins of Distress Associated with Gender Incongruence

An objection to retaining GD as a psychiatric diagnosis is that it identifies gender-variant identity and expression as pathological when the distress required for the diagnosis may be related the reaction of the social environment to the child's gender variance. Since not all who experience a mismatch between assigned and experienced gender experience clinically significant distress, the origins of distress should be considered separately from the origins of the cross-gender identification. One hypothesis, for which there is very limited support, is that the cross-gender identification might, itself, be an expression of distress due to inadequate parenting or loss of a parent (e.g., a boy who loses his mother through death or emotional withdrawal may suddenly show a preference for feminine attire and state that he is or wants to be a girl. Such cases have been interpreted as the boy replacing the lost mother through identification and imitation). Alternatively, psychologist, Diane Ehrensaft, suggests that in most children with GD, the distress arises from stigma and prohibitions within the child's social environment. She bases this suggestion on her observation that

distress often resolves or diminishes when gender dysphoric children are allowed to freely express their experienced gender.

Parents may act to limit social stigma and prohibitions in preschool age children in the home; however, unless preventive measures are taken, starting school and increasing interactions with the outside world may prove traumatic and exacerbate dysphoria. For some, the onset of puberty has been associated with resolution of GD; however, for others the onset of puberty and development of the secondary sex characteristics of the natal gender exacerbate dysphoria and may precipitate an emotional crisis. The foregoing pertains to those in whom GD began in childhood and persisted into adolescence. The origins of GD with later onset are more poorly understood, especially when there was no history of observable gender nonconformity in childhood. In some individuals, as mentioned previously in this chapter, late-onset GD has been hypothesized to be erotically driven (*autogynephilia*) and may entail a history of arousal to cross-dressing prior to the onset of anatomical dysphoria. Alternatively, the authors of a study of military veterans with late-onset GD, suggested that the apparent late onset was due to other psychological factors. Specifically, many veterans reported that they joined the military in an attempt to purge their transgender feelings (*suppression*), believing that the military environment would make men of them, while others arrived at this interpretation after the fact (*repression*).

Directions for Future Research

The origins of gender identity are of broad general and theoretical interest while the causes of the distress and impairment in GD demand research because of their implications for clinical interventions. Due to the low prevalence of GD and the heterogeneity of presentations, research could be facilitated if clinics pool their population samples. Longitudinal studies from childhood through adulthood of such pooled samples, and observing how they change and the factors associated with change (including clinical interventions) could inform neuroscience, genetics, and behavioral approaches with regard to factors that influence experienced gender. Developmental consideration of cognitive, personality, psychosexual, and psychosocial development might be correlated to clinical findings including persistence versus desistence, genomic variations, and evidence of variations in androgen exposure. Regarding the distress and impairment required for the diagnosis of GD, more research aimed at informing clinically appropriate intervention is needed on the timing and contexts in which they initially emerge as well as contexts associated with exacerbation such as entering school and the onset of pubertal changes.

With collaboration among groups, sample sizes might be accrued to provide adequate statistical power to test the hypothesis that the gender identity component of GD entails a CNS-limited form of intersexuality.

Large samples would be required to account for multiple confounding variables including gonadal status (e.g., intact or gonadectomized), age at gonadectomy, and time elapsed since gonadectomy as well as past and present exogenous hormonal regimens. Such research could begin with living subjects and focus on those CNS regions and neurophysiological responses implicated by previous research. In light of the evidence, although controversial, that the origins of GD in natal males may differ according to age at overt onset and sexual orientation, these variables should be assessed and included in future studies. Sexual orientation and gender identity (SOGI) are not currently systematically recorded at death, virtually preventing the accrual of adequate sample sizes for meaningful examination of postmortem specimens for variation associated with GD or transgender status; however, this should change as SOGI fields are added to electronic health records. The work of Dick Swaab's group in the Netherlands is a notable exception, but even that work can at best be regarded as providing leads for future research that will be adequately powered and include adequate control subjects. If enacted, current recommendations for routine clinical assessment of SOGI inclusion in the electronic health record will facilitate classification of gender identity postmortem. In addition, recent changes in the postmortem brain banking endeavors funded by the National Institutes of Health should facilitate accrual of sufficient specimens for postmortem studies of low prevalence conditions, perhaps including GD and transsexuality. It is clear that early androgen exposure affects gender related behaviors and identity, and that this effect is not mediated by androgen levels alone but by the timing of androgen exposure, the presence of functioning androgen receptors, and the signaling mechanisms downstream of the receptors. To date, little work has examined either post receptor signaling mechanisms or how they might be impacted by genetic and epigenetic factors. Recent technological advances now make such studies feasible.

SOCIAL AND CULTURAL ISSUES

Delineating the full extent of social and cultural issues affecting transgender individuals is beyond the scope of this chapter. Among the *stigmatizing* challenges some transgender individuals face are rejection by family (leading to homelessness for some adolescents), estrangement from children produced in pretransition heterosexual relationships, rejection by peers, rejection by religious authorities, lack of acceptance from school authorities, bullying, harassment, job discrimination, lack of adequate medical and mental health care, and violence. In recent years, members of the transgender community have produced a growing literature detailing personal stories of coming out as trans. These include Jennifer Finney Boylan's *She's Not There: A Life in Two Genders* and Jamison Green's *Becoming a Visible Man*.

CLINICAL APPROACHES

The most widely recognized guidance for treating persons with GD come from the WPATH Standards of Care (SOC), now in their 7th revision. The World Professional Association for Transgender Heath Standards of Care (WPATH SOC) endorses psychological evaluation and a staged transition in which fully reversible steps (e.g., presenting as the desired gender) and pubertal suspension (in adolescents), preceding partially reversible procedures (administration of gonadal hormones to bring about the desired secondary sex characteristics), which precede irreversible procedures (e.g., gonadectomy, vaginoplasty in natal males, mastectomy and surgical construction of male-typical chest and phalloplasty in natal females). Hormone therapy can be initiated with a referral from a qualified mental health professional, or other health professional appropriately trained in behavioral health and able to assess GD and the eligibility and readiness for of the patient for such therapy. The WPATH SOC v7 requires that those with mental health issues be referred for appropriate treatment. Previous versions of the SOC specified minimum times of living in the desired gender role prior to the initiation of hormonal therapy but version 7 makes no specification in this regard. The WPATH SOC v7 suggests 1 year of continuous cross-sex hormone therapy (unless contraindicated in an individual), two referral letters prior to genital surgery. The letters must document persistent gender dysphoria, and capacity to make a fully informed decision and give consent for treatment. If any medical or mental health concerns are present, they must be adequately controlled.

The WPATH SOC provides flexible guidelines, not rigid rules. For example, natal females may elect mastectomy and perhaps male breast construction prior to presenting socially as male. Further, not all clinicians follow the WPATH SOC so adults who have capacity to give informed consent may receive gender transition treatments for which they satisfy the qualifying criteria of the treating clinicians. These criteria vary among clinicians and clinics. Reviews that have graded the quality of evidence relating particular components of the WPATH SOC to outcomes have concluded that psychotherapy prior to transition is associated with good outcome.

Treatment of Prepubescent Children

Work with children with GD can be considered to follow one of three approaches, each of which will be discussed in turn. None of these approaches entail hormonal or surgical interventions. The first approach focuses on working with the child and caregivers to lessen gender dysphoria and to decrease cross-gender behaviors and identification. This approach is based upon the assumption that the child's self-esteem can be best served by improved social integration, including positive relationships

with same-sex peers, expected to come with decreased gender atypicality. Since GD that persists into adolescence is more likely to persist into adulthood and manifest as transsexuality, this approach may also be pursued with the goal of inducing desistence prior to adolescence; however, no empirical evidence suggests that is a realistic goal. Persistence is considered to be an undesirable outcome by some but not all clinicians who work in this area. A variety of reasons are cited for considering persistence undesirable including social stigma as well as the likelihood persistence will eventually entail hormonal and surgical procedures with their associated risks and costs. Critics of this approach have likened it so-called *sexual orientation conversion efforts (SOCE)* or *reparative therapies* of homosexuality (see Section 21.2 Homosexuality). However, unlike SOCE practices, this clinical approach to children with GD is accepting of adult homosexual outcomes. It has, nevertheless, been legislatively defined as conversion therapy and banned in many jurisdictions, and the American Psychiatric Association is re-examining its position with respect to the practice.

A second approach, often referred to as *the Dutch approach,* is to remain neutral with respect to gender identity outcome and to not target gender identity and gender atypical behaviors for change. The goal is to allow the developmental trajectory of gender identity to unfold naturally without pursuing or encouraging a specific outcome. This approach entails combined child, parent and community-based interventions to support the child in navigating the potential social risks. This approach is premised on evidence that GD diagnosed in childhood usually does not persist into adolescence and beyond, and on the lack of reliable markers to predict in whom it will or will not persist. This approach aims to avoid damaging the child's self-esteem by risking conveying to him/her that his/her likes and dislikes, behaviors and mannerisms are somehow intrinsically wrong. This approach may also entail assisting the child in mastering the cognitive, emotional, and behavioral tools for coping with the social stigma of gender variance.

A third approach entails affirmation of the child's cross-gender identification by mental health professionals, family members and others. Thus, the child may be supported in transitioning to a cross-gendered role, with the option of age-appropriate endocrinological and surgical interventions if GD persists. The rationale for supporting transition before puberty is the belief that a transgender outcome is to be expected in some children, and that these children can be identified so that primary caregivers and clinicians may opt to support early social transition. Since childhood transition is done solely at a social level without medical intervention until puberty, some assert that children who transition this way can revert to their originally assigned gender if necessary. Reservations regarding this approach include that the peer-reviewed literature does not

support the view that desisters and persisters can currently be reliably distinguished as children as well as concern that transition in childhood may increase the likelihood of persistence. In addition, individual case reports suggest that after transitioning in childhood, reverting to the natal gender may entail complications.

The APA Task Force on Treatment of GID published in 2012 identified the major tasks for mental health professionals working with children referred for gender concerns. These include accurately evaluating the gender concerns that precipitated the referral; accurately diagnosing any gender identity related disorder in the child according to the criteria of the most current DSM; accurately diagnosing any coexisting psychiatric conditions in the child, as well as problems in the parent–child relationship, and recommending appropriate treatment; providing psychoeducation and counseling to the caregivers about the range of treatment options and their implications; providing psychoeducation and counseling to the child appropriate to his or her level of cognitive development; when indicated, engaging in psychotherapy with the appropriate persons, such as the child and/or primary caregivers, or to make appropriate referrals for these services; educating family members and institutions (e.g., day care and preschools, kindergartens, schools, churches) about GV and GID; and assessing the safety of the family, school and community environments in terms of bullying and stigmatization related to gender atypicality, and to address suitable protective measures.

Kim and Jan are identical twins who were assigned male at birth. Their mother's pregnancy and delivery were unremarkable. Starting at age 2, both children began to exhibit strong atypical gender interests. They preferred playing with girls' dolls and rejected stereotypical boys' toys. As they got older, they liked to play dress up with their mother's clothing, sometimes wrapping their heads with towels to simulate long hair. Both insisted that they were girls. They preferred to urinate sitting down. Although they thought their feminine interests would fade away with time, when the children were 5 years old the parents sought professional consultation at a gender clinic in the major metropolitan center where they lived.

Both children were given a DSM-IV diagnosis of GID in children (what would be GD of childhood in DSM-5). The clinic discussed treatment options that ran the gamut from discouraging all cross-gender behavior at one extreme to the possibility of socially transitioning the children as to live and be publicly identified as girls. The parents felt that inhibiting the cross-gender interests of their children was too harsh but were not at all comfortable with the idea of social transition given that it was not clear at the time whether the gender dysphoria of either child would desist or persist. They opted for what is known as the Dutch approach. They would be identified publicly as boys with gender atypical interests and no effort was made to interfere with the unfolding of either child's gender outcome. This entailed much psychological support for the parents, some members of their extended family and occasional consultation with school officials.

At age 10, Kim told the parents that he no longer thought of himself as a girl. At age 14 he came out to his parents as gay. Jan, on the other hand, continued to feel like a girl. At age 12, as puberty began to approach, Jan became very panicky about anticipated bodily changes and her parents approved of treating her with puberty suppressing drugs. Her gender dysphoria never remitted and when she turned 16, Jan was treated with feminizing hormones. At age 18, she began seeking consultation to explore surgical treatments.

Treatment of Adolescents

Management of GD that has persisted from childhood is more straightforward than management of GD with later onset, especially when late onset occurs in the context of recent psychological trauma or co-occurring psychopathology. Adolescents with recent onset should be screened carefully to detect emergence of the desire for gender transition in the context of trauma as well as for any disorders such as schizophrenia, mania or psychotic depression that may produce gender confusion. When present, such psychopathology must be taken into account in assessing the adolescent's desire to transition as well as how supportive to be of that desire. The presence of significant psychopathology does not preclude transition but suggests that one should proceed more slowly in establishing the diagnosis of gender dysphoria as well as insuring that the psychopathology is controlled to the extent that it will not interfere with adherence with a safe transition plan. Even the presence of a disorder with psychotic features does not exclude a gender dysphoria diagnosis if the gender concerns do not remit when the psychosis is treated. In addition, autism spectrum disorders (ASDs) may occur with increased frequency in individuals with gender dysphoria. Impaired abilities to articulate inner experiences and to assess how they are perceived by others, together with cognitive rigidity and an intolerance of ambiguity may present special difficulties in gender identity formation in individuals with ASDs. While these limitations do not preclude gender transition in those with high functioning ASDs, they, nevertheless, present challenges both in the assessment and treatment of gender dysphoria.

With the beginning of puberty, development of the secondary sex characteristics of the natal gender often triggers or exacerbates the anatomical dysphoria of adolescents with GD. In the last two decades, the option has become available for patients approaching puberty with severe GD and minimal, if any, coexisting psychopathology to have puberty suspended medically in order to prevent or to minimize development of unwanted secondary sex characteristics, some of which cannot be fully reversed with subsequent hormonal or surgical gender reassignment therapies. Many adult transsexual persons express regret over the irreversible body changes that occurred during puberty. Withholding medical intervention for adolescents is not a neutral option as it can prolong distress and, by allowing pubertal changes of the natal gender to proceed, contribute to a gender atypical appearance that could provoke abuse and stigmatization. Current evidence suggests that pubertal suspension can be done for a period of up to several years during which time the patient, with the clinicians, can decide whether it is preferable for the adolescent to revert to living in the birth sex or to continue gender transition with cross-sex hormone therapy. There are currently few data

regarding the best timing for initiating cross-sex hormone treatment in adolescents and no studies comparing outcomes when such treatment is initiated in adolescence as opposed to adulthood. In those with DSDs who have been gonadectomized, puberty consistent with the adolescent's experienced gender is generally induced at the time his/her peers are entering puberty. One study on carefully selected individuals in the Netherlands, however, found that, as assessed by satisfaction with surgery and lack of regrets, outcome was generally better in individuals who initiated reassignment as adolescents compared to those who initiated reassignment as adults. A recent longitudinal follow-up study by the same group found that when pubertal suppression was employed in very carefully selected minors, following age-appropriate pubertal induction and gender reassignment surgery in adulthood, subjects scored as well or better than the general population on measures of psychological and social adjustment.

The WPATH SOC v7 requirements for puberty suppressing hormones are that the adolescent has demonstrated a longstanding and intense pattern of gender nonconformity or GD; that GD has emerged or worsened with the onset of puberty; that any coexisting physical or mental conditions that could interfere with treatment have been addressed so that the individual is stable enough to start and to adhere with treatment; and that the adolescent has given informed consent, and particularly when the adolescent has not reached the age for medical consent, the parents or other caretakers have consented to the treatment and will be involved in supporting the adolescent throughout the treatment process. According to the WPATH SOC v 7, genital surgery should not be carried out until patients reach the legal age of majority to consent to medical procedures and have lived continuously for at least 12 months in the gender role congruent with their gender identity. Chest surgery in female to male patients can be carried out earlier than genital surgery, but preferably after "ample time of living in the desired gender role" and after 1 year of testosterone treatment.

Treatment of Adults

Transgender people are sufficiently common that even adult psychiatrists whose practice does not focus on gender variance may encounter patients who are transitioning gender, or contemplating gender transition. Thus even psychiatrists who are not gender specialists should be prepared to provided informed and respectful care to gender variant patients, including accurate diagnostic assessments and culturally attuned treatment plans that include not only provision of immediate care but arrangement of necessary referrals for those who are, or who are contemplating, transitioning gender. All clinicians should be aware of the federal, state, local and regulatory agency policies that pertain to the provision of care to

transgender and gender variant patients in their workplace. Adult patients coming to a gender identity clinic often present with straightforward requests for hormonal and surgical gender reassignment. When gender dysphoria is severe and intractable, gender reassignment is often the best solution.

Mental health professionals who work with individuals who plan to transition using hormonal and/or surgical treatments, or who are in the process of doing so, need to be knowledgeable about these treatments, particularly their mental health implications. All should be aware that cross-sex hormonal regimens may lead to infertility and should receive appropriate counseling and referrals if preservation of reproductive potential is desired. Natal males should be aware that estrogen treatment will cause erectile dysfunction as well atrophy of the testes and penis. As discussed previously in this chapter which, if any, medical and surgical procedures to employ in one's transition plan varies from one individual to the next. Similarly, some seek mental health services for psychological support or to assist in decision making while others do not.

FTM transition usually includes use of androgens, which produce male secondary sex characteristics, including beard growth and male distribution of body hair, deepening of the voice, and coarsening of the facial features and skin. Androgen supplementation also enlarges the clitoris, often to the extent that metoidioplasty becomes feasible. MTF transition often consists of both estrogen supplementation in conjunction with agents that either suppress androgen levels or interfere with androgens' actions. Estrogen effects include breast development and feminizing changes to skin and hair, although for many who transition MTF after completion of male pubertal development, depilation will be needed. Surgical reduction of the laryngeal cartilage or facial surgeries may be needed for a convincing feminine appearance. Use of hormonal preparations is much more effective in "adding" physical sexual characteristics than in "subtracting" those that have already developed with natural puberty. Utilization of either androgens or estrogens carries the potential for both added health risks and, in some cases, physiologic benefits. Psychiatrists may prescribe and monitor the effects of hormonal preparations used in gender transition, but this aspect of transition is often managed by other specialists. All of the treating clinicians should be aware of potential psychological effects including the emotional changes that may occur with use of either androgen or estrogen supplementation. Increased libido usually occurs with androgen use. Some who transition MTF also experience a stronger interest in sex, perhaps due to increased self-confidence after physical feminization. Individuals who are transitioning often benefit from ongoing psychiatric care. In addition to the psychotherapeutic work involved in the context of life-changing experiences, psychiatric care may include monitoring the psychiatric effects

of hormone use along with the prescribing physician and monitoring adherence with medical monitoring. Safer sex information, and instruction in self-protective negotiations in sexual settings, should be provided by the psychiatrist if this has not been done by the prescribing clinician. This information should be tailored to the needs and experiences of transgender persons.

Surgeries for purposes of gender transition include breast and chest ("top") surgeries and genital ("bottom") procedures. The goal of FTM top surgery is not mastectomy, as would be performed for the treatment of cancer of the breast but creation of a natural appearing male chest. Some adults who transition MTF have breast augmentation surgery due to achieving minimal breast development with hormonal treatment alone. Many adults undergoing MTF genital surgery receive penile inversion vaginoplasty with clitoroplasty, labiaplasty, and orchiectomy. FTM genital surgery can consist of either metoidioplasty with limited scrotoplasty, or more extensive surgery, including phalloplasty with grafted tissue from another body site, urethral extension, scrotoplasty and vaginectomy. Hysterectomy and oophorectomy are performed in either case. In order to give truly informed consent, patients should be able to describe the procedures they will receive in reasonable detail including realistic outcomes and risks as well as what, if any convalescent period and medical follow-up will be involved. To demonstrate readiness for each step in transition they must also demonstrate capacity to adhere with its demands, including necessary medical monitoring and postsurgical care.

Maria, a natal female, reported being a tomboy as a child. As a child she had both boys and girls as playmates but her best friends were boys. She enjoyed sports and boy's games and had no interest in dolls. She resisted wearing skirts and dresses and preferred boyish clothing. She reported feeling different from other girls, a feeling that intensified when she entered puberty at which time she wore loose fitting shirts to conceal breast development and became markedly distressed with the onset of menses. She recalled wishing she had been born male. From her earliest awareness of sexual attractions, they were directed toward girls and by the age of fifteen she began to label herself a lesbian and began to socialize with lesbian and gay teens. Nevertheless, she felt different from the other lesbians in a way that she could not quite describe. She wore masculine clothes, baggy shirts, concealed her breasts by binding them and kept her hair cropped short.

In her second year of college she began dating a woman with whom she fell in love and the two began living together as a couple. In the relationship she realized that she experienced herself as a man rather than as a lesbian, began to explore transgender websites and enrolled in a human sexuality course at college. He found that the identity of *transgender man* aptly labeled his experienced gender, begin to think of transitioning, and discussed the possibility with his partner who was supportive. He changed his name to Mark, began to wear only male clothing and requested that family, friends, and colleagues address him as a man. He experienced a level of authenticity living as a man that he had not previously experienced as a lesbian. He referred himself to a medical school affiliated clinic known for offering hormonal and surgical gender transition services. After careful evaluation, a transition plan was formulated and he would have a mastectomy and male breast construction. Testosterone therapy for lowering the voice, beard growth and masculinizing the body was temporally delayed to allow time for Mark's ovaries to be stimulated to harvest eggs for cryoprotection to allow future reproduction (with his partner carrying the pregnancy after fertilization with

donated sperm). He also entered into individual and couple's psychotherapy. Upon starting testosterone therapy menses stopped, the voice deepened, facial and body hair grew, sex drive increased and the clitoris enlarged. As eggs had been harvested, Mark elected to undergo hysterectomy and oophorectomy sooner rather than later. He also elected metoidioplasty. Four years later, Mark had graduated college, was working full time and the couple had their first child.

The APA Task Force on GD identified the following tasks of mental health professionals working with adults with GD: assessing and diagnosing patients' gender concerns according to DSM criteria and assuring that these are appropriately addressed; assessing and correctly diagnosing any coexisting psychopathology and assuring that it is addressed adequately; distinguishing between GD with concurrent psychiatric illness and gender manifestations that are not part of GID but epiphenomena of psychopathology; engaging in psychotherapy with gender variant individuals as indicated. This includes identifying the elements that should be addressed in therapy including the impact of discrimination and stereotyping; ensuring that individuals who are in the process of transitioning, or who are considering or planning to do so, receive counseling from a qualified professional about the full range of treatment options and their physical, psychological, and social implications including both their potential benefits and the full range of potential limitations (e.g., loss of reproductive potential and possibilities for preservation of gametes and assisted reproduction), risks and complications; ascertaining eligibility and readiness for hormone and surgical therapy or locating professionals capable of making these ascertainments to whom the patient may be referred; and educating family members, employers, and institutions about GD.

Treatment of Individuals with DSDs Including Somatic Intersexuality

Some individuals with DSDs, in a proportion that varies greatly with syndrome and assigned gender, become dysphoric in their assigned gender and may reject it. In DSM-5 these individuals can meet criteria for the diagnosis of GD, although they were categorically excluded from the diagnosis of GID in previous versions of the DSM (but could be given the diagnosis of GID Not Otherwise Specified). A variety of issues in the clinical care of individuals with DSDs require the expertise of mental health professionals. The clinical options and decision-making processes that bear on gender transition and reassignment overlap to some extent regardless of the presence or absence of a DSD. When a DSD is present, however, there have been fewer barriers to legal gender reassignment, and the barriers to hormonal and surgical treatments in conjunction with gender reassignment are lower. This is because many individuals with DSDs, particularly those involving somatic intersexuality will have been gonadectomized prior to adolescence as part of routine treatment (e.g., to reduce the risk of

malignancy that is associated with some gonadal abnormalities) and thus will have been rendered sterile. They will also require the age-appropriate administration of exogenous hormones to induce puberty. Thus, concerns of gonadectomy, pubertal induction, and lifelong dependence on exogenous hormones and infertility often do not apply.

Major areas of involvement of psychiatrists and other mental health professionals in the care of individuals with DSDs and gender dysphoria include: the evaluation of gender identity and the assessment of incongruences, if present, between gender identity and assigned gender; decision making regarding gender reassignment; psychotherapy to address significant gender dysphoria in individuals with a DSD who do not transition gender; psychological aspects of the endocrine management of puberty in the context of gender reassignment; and psychological/psychiatric aspects of genital surgery in the context of gender reassignment. Whether or not GD is present, individuals with DSDs, including somatic intersexuality, have often been subjected to a variety of significant stigma due to their condition and may benefit from psychotherapy to address it.

Other Clinical Issues

The 2011 IOM report on the health of LGBT people drew increased attention to the fact that compared to the general population LGBT populations experience increased violence, barriers to healthcare, and poorer physical and mental health. The report further pointed out the tendency of studies to lump LGBT populations together rather than studying them separately, greatly limiting knowledge as to how disparities impact each population individually. This is particularly true for the transgender population, the smallest segment of the LGBT community. The Department of Health and Human Services (HHS) has been testing appropriate survey instruments since 2010 and publication of their first transgender specific data is anticipated in 2015. The LGBT population has been included in the HHS's National Health Disparities Report since 2011, but to date the largest transgender survey published is The National Transgender Discrimination Survey of 2010 by an advocacy organization. The major findings of that survey were that participants reported high levels of postponing medical care when sick or injured due to discrimination (28 percent) or inability to afford it (48 percent), with 28 percent reporting harassment in medical settings and 50 percent believing that providers lacked knowledge about transgender care. Respondents reported a rate of HIV infection four times the national average (2.64 percent vs 0.6 percent), with rates for transgender women at 3.76 percent. For those who were unemployed, the rate was 4.67 percent and for those engaged in sex work 15.32 percent. Over a quarter of participants reported misusing drugs or alcohol specifically to cope with the discrimination they

faced due to their expressed gender. Strikingly, 41 percent reported having attempted suicide compared to 1.6 percent of the general population. These findings are in line with findings of multiple smaller studies.

Violence. Transgender people are often targeted for violence based on their nonconformity with gender norms and/or their perceived sexual orientation. Assaults against transgender people tend to be particularly violent and the Human Rights Campaign has shown that a high proportion of homicides in the United States are antitransgender hate crimes. In the Transgender National Discrimination Survey, 8 percent of respondents reported having been physically attacked or assaulted in a place of public accommodation, with the rate being higher among particular subgroups: African Americans (22 percent), multiracial individuals (13 percent), and the unemployed (12 percent).

As adults, transgender individuals report increased rates of childhood maltreatment and abuse which are predisposing factors for the later onset of psychiatric symptoms and disorders including depression, posttraumatic stress disorder, and substance abuse. In the Transgender National Discrimination Survey, those who expressed a cross-sex identity or gender nonconformity in grades K-12 reported extremely high rates of harassment (78 percent), physical assault (35 percent), and sexual violence (12 percent). Harassment was so severe that 15 percent left a school in K-12 settings or in higher education. Perpetrators of harassment and abuse were reported to be teachers as well as peers, with teacher perpetrators being associated with worse outcomes with respect to HIV infection and substance abuse. Transgender individuals are also among the groups most likely to be impacted by intimate partner violence.

Suicide. There is evidence indicating that transgender populations are at increased risk for suicide. Because gender identity is not systematically recorded at time of death, the number of transgender deaths from suicide (or any other cause) is not known. However, high rates of suicide have been reported among individuals receiving medical interventions for gender transition. Nondisclosure of transgender status is associated with lower risk of suicide attempts, possibly because hiding protects against rejection, discrimination, and victimization, which are otherwise strongly linked to suicide attempts in transgender populations. In surveys, 25 to 43 percent of transgender adults have reported lifetime suicide attempts.

In the National Transgender Discrimination Survey, 41 percent of respondents reported attempting suicide compared to 1.6 percent of the general population, with rates rising for those who lost a job due to bias, had a history of being harassed/bullied in school, had low household income, or were the victim of physical assault or sexual assault. Discrimination was pervasive throughout the entire sample; however, the

impact of antitransgender bias was particularly marked when it was combined with racial discrimination. People of color in general fared worse than white participants, with African American transgender respondents faring worse than all others in many areas examined. Assessment and treatment of suicide risk in transgender individuals requires practitioners to be familiar with the cultural milieu of their lives. Not only should clinicians be cognizant of the impact that antitransgender or *transphobic* attitudes can have on normal psychosocial development, they must also recognize extreme situations that may arise. This requires awareness of and asking about events that a patient—perhaps in the closet or fearing lack of support from a medical/psychiatric authority—may not offer spontaneously. For example, is a child, adolescent or adult being bullied or harassed for gender atypical behavior? Has an adolescent or adult been rejected after coming out to a family member or friend? Has an adolescent been thrown out of his or her home by unaccepting parents? Is a desperate transgender patient unable to access needed medical care for transition? These are just some examples of the kind of culturally sensitive knowledge clinicians need to access if they are to provide support to this at risk patient population.

Mental Health. Survey studies have consistently found transgender men and women to report higher rates of substance abuse, suicidal ideation, suicide attempts and stress-related psychiatric disorders (e.g., anxiety, depression and somatization) compared to national norms. As noted, many report misusing substances to cope with stigma and discrimination due their gender expression. Similarly, self-reported social stigma and discrimination have been found to be correlated with the degree of psychological distress. These findings based on self-report are similar to recent findings based on examining the electronic health records of transgender US veterans who received care through the Veterans Health Administration. Specifically, transgender veterans were found to have an increased prevalence of posttraumatic stress disorder, depression, alcohol abuse, and suicidality. Examination of mortality data revealed that veterans with transgender-related ICD-9 diagnoses have an increased suicide rate, with the average age of suicide being 49.4 years.

Transgender individuals and those with GD may also have psychiatric disorders that are independent of their gender concerns, and while even serious mental illnesses such as major depression, bipolar disorder, borderline personality disorder and schizophrenia are not absolute contraindications to gender transition, they must be adequately treated so as not to interfere with adherence with the transition regimen. Very importantly, the manifestations of coexisting disorders should not be mistaken for manifestations of GD that will resolve with transition. Those in transition should understand that coexisting psychiatric disorders are

very likely to persist and require treatment after transition.

Medical Illness. Little research has focused specifically on physical health disparities in transgender populations beyond sexually transmitted infections among transgender women and the effects of hormone therapy on physical health. Although no clinical trials have been conducted, this research indicates a risk of venous thromboembolic disease and elevated levels of prolactin associated with feminizing hormone therapy; and elevations in liver enzymes, loss of bone mineral density, and increased risk for ovarian cancer associated with masculinizing hormone therapy. Research on increased risk for cardiovascular disease remains inconclusive. Given recent survey data documenting barriers to healthcare for transgender individuals discussed in the foregoing, one would expect health disparities in addition to those associated with hormonal and surgical transition. A recent case control study of disparities in over 5,000 transgender veterans found increased prevalence of 17 out of 18 medical diagnoses including chronic obstructive pulmonary disease, diabetes mellitus and HIV infection. Childhood and adolescence are times of good physical health. Thus studies have exploring the physical health of children with GD is lacking.

Children. Primary modes of therapy utilized in working with children with GD include individual insight-oriented psychoanalytic or psychodynamic psychotherapy; protocol-driven psychotherapy such as behavior modification; parent and peer relations focused therapy; and parent and child therapeutic groups. Additional interventions include support groups for primary caregivers, community education through websites and conferences, school-based curricula, and specialized youth summer camps. The primary focus of intervention is often the primary caregivers. Depending on the treatment approach chosen, work may include parenting support and psychoeducation, careful guidance regarding behavior modification efforts and their potential to damage self-esteem, and instruction in techniques for building self-acceptance and resilience in the child. Some interventions are multifaceted and involve the school and community as well as the child and family. These include diversity education and steps to prevent bullying. As stated, psychotherapeutic efforts to reduce gender dysphoria in prepubescent children have provoked angry responses in the LGBT community, with some arguing that they represent a form of conversion or reparative therapy and with some jurisdictions legislatively defining them as such. Given that most children with gender dysphoria may experience spontaneous remission of their symptoms, it is not clear how, if at all, these therapeutic interventions influence the rate of desistence.

Adolescents and Adults. There is little evidence to indicate that gender dysphoria in adolescents or adults will remit in response to psychotherapeutic interventions. Such approaches are currently not considered effective by mainstream psychiatry and other mental health professions. This lack of psychotherapeutic responsiveness in part underlies the current approach of assisting gender dysphoric individuals with transition rather than trying to change their gender identity. This entails supportive therapy and psychoeducation to ensure that the individual understands and is committed to the long and difficult process of gender transition and has fully explored all options, including not transitioning (and living as a homosexual person if one is sexually attracted to members of the natal gender), transitioning socially without medical or surgical intervention, or selecting which endocrine and surgical alternatives to pursue. Psychotherapy is employed not only to explore issues related to the individual's commitment to living the role of the experienced gender but to provide insight oriented and supportive therapy to assist the individual find authenticity in their experienced and expressed gender.

ETHICAL AND LEGAL ISSUES

Legal Change of Gender

Depending on what country and jurisdiction they live in, transgender individuals face a range of differing regulations if and when they seek a change of gender on legal documents. Some US states (Idaho, Ohio, Tennessee) currently refuse to grant a gender change on birth certificates. A majority of states permit one's name and sex to be changed on birth certificates, either through amending an existing birth certificate or issuing a new one; however, gender confirmation surgery is a prerequisite in many states. All US states allow a legal change of gender on drivers' licenses, although requirements for doing so vary from state to state. Countries without marriage equality may not recognize the marriage of two individuals whose identity documents show they are the same gender. For example, a male to female transsexual who had not changed the gender marker on her identity documents might not be able to marry a man, whereas, if she had changed her identity documents she would not be able to marry a woman. Surviving transsexual spouses have been denied inheritances and surviving spousal rights in many cases.

In the United States, the gender designation on federally issued documents such as passports can now be changed with a single physician's letter that can be issued without requiring either hormonal or surgical treatments. However some countries will not allow legal changes of gender unless the individual has been rendered infertile via surgical means.

Antidiscrimination Protections

Presently, the US federal government does not offer any legal protections against discrimination in housing and employment based on one's gender identity and gender expression. The majority of US states do not provide statewide antidiscrimination protections for transgender individuals. Title VII of the US Civil Rights Act prohibits discrimination on the basis of sex which has been interpreted in several federal cases since 2004 as extending to claims of discrimination based on gender identity or failure to conform to stereotypical notions of masculinity or femininity. In 2012, the Equal Employment Opportunity Commission, the federal agency in charge of enforcing employment discrimination laws, declared unanimously that antitrans bias was sex discrimination under Title VII. In 2014, Attorney General Eric Holder issued a memo clarifying that the Justice Department will no longer assert that Title VII's prohibition of gender discrimination does not encompass gender identity, per se. This memo was intended to foster consistent treatment of claimants and to clarify the Justice Department's ability to file Title VII claims against state and local public employers on behalf of transgender individuals. Nevertheless, most civil rights advocacy organizations believe stronger protections are needed at both the state and federal level.

Gatekeeping

Many countries with national health care systems require mental health professionals to act as *gatekeepers* to transition services for transgender people. This gatekeeping function is ostensibly to reduce the incidence of posttransition regrets, particularly for those undergoing irreversible surgical procedures. Yet there is a somewhat irrational, gender policing quality to this requirement since few other surgical procedures, including those considered high-risk, require psychiatric evaluation and approval. In some countries, transgender people wishing to transition must prove to psychiatrists that they are the "right" kind of transgender and that they have been so for a sufficiently long period of time in order to qualify for services. This has created some controversy, as many individuals seeking transition do not otherwise have a mental disorder or desire mental health treatment, and the gatekeeping function may be seen as an unnecessarily burdensome requirement. How the gatekeeping role is performed has contributed to considerable tension between mental health professionals and transgender advocacy groups.

Access to Care for Incarcerated Individuals

Incarcerated transgender people who have not had genital surgery are imprisoned in facilities according to natal gender. Those who have undergone genital surgery are housed according to their reassigned gender. Access to care for hormone treatment while incarcerated is not guaranteed

and may vary from facility to facility. In recent years, transgender advocacy groups have had some success in lawsuits against prison facilities for denying access to hormone treatment, arguing that denial of such needed medical treatment constitutes "cruel and unusual punishment."

Ethical Issues in Child Treatment Controversies

Clinical approaches to gender variant children raise many as yet unanswered ethical questions. For example, some in the transgender community maintain that clinical attempts to prevent transsexualism in prepubescent gender variant children, no matter how well meaning, are always unethical. It is presently unknown whether prevention of transsexualism is in fact possible, and whether such efforts constitute a benign medical activity or an attack on an individual's identity. There are presently no controlled studies that demonstrate discouraging cross-gender behavior and interests in childhood does in fact reduce persistence or prevent transsexualism. Further, such studies are unlikely to be conducted due to the ethical concerns they would raise together with the fact that parents would be unlikely to allow the treatment of their child to be assigned at random. In addition, there are no reliable indicators to distinguish children whose dysphoria will desist from those who will persist. Since no clinician can accurately predict the future gender identity of any particular child, it is possible that efforts to discourage cross-gender play and identifications may be experienced as hurtful and possibly even traumatic by children whose gender dysphoria persists into adolescence and adulthood. Also, is it not possible that opposing a wish to explore cross-gender expression is harmful to some children, and not just persisters? Whether they persist or desist in their transgender behavior or identity, these children may internalize disapproving attitudes towards atypical gender behavior and expression (transphobia) with possible consequences for adult development. Is it ethical to delay the social transition of persisting children in order to prevent possible psychological harm to those who will desist? Put another way, are the children who will grow up to be trans being subjected to an unnecessary stress in order to preserve the well-being of those who will not? Finally, there is no empirical evidence demonstrating that a prepubescent child who is permitted to socially transition but then desists can simply and harmlessly transition back to the natal gender. Given the complexity involved in the first social transition, should the claim that transition back to the original gender is entirely without risks and pitfalls be questioned? In the absence of empirical studies, is permitting early social transition without a verifiable system of distinguishing persisters from desisters ethical? Further, would conducting such empirical studies even be ethical?

TRAINING AND RESEARCH NEEDS

Training

Transgender presentations are sufficiently common that even adult psychiatrists whose practice does not focus on gender variance encounter patients who are transitioning gender, or contemplating gender transition. However, gender dysphoria and gender transition receive little attention in medical school or psychiatric residency curricula. In a survey of deans of medical education at over 100 medical schools across North America, 34 percent rated coverage of sexual orientation and gender identity (SOGI) minority health care as poor to very poor, with only 30 percent rating it as better than fair. Further, on average, only five hours were devoted to training in caring for LGBT patients though out the 4-year medical school curriculum and only 30 percent of schools provided any training on gender transition.

Table 21.4–2.
Internet Resources for Professionals and for Transgender Individuals and Their Families

American Medical Association LGBT Advisory Committee (public education)	http://www.ama-assn.org/ama/pub/about-ama/our-people/member-groups-sections/glbt-advisory-committee.page?
AGLP (education and professional support for psychiatrists, residents and medical students)	http://www.aglp.org/
Gay, Lesbian, and Straight Education Network (GLSEN) (support for LGBT youth in the school system)	http://www.glsen.org/students/tsr
GLAAD: Transgender 101 (public education)	http://www.glaad.org/transgender/trans101
GLMA (education and professional support for physicians and other health professionals)	http://www.glma.org/
Group for the Advancement of Psychiatry (GAP): LGBT Mental Health Syllabus (for medical students and residents)	http://www.aglp.org/gap/
Hetrick-Martin Institute (social support for LGBT youth)	http://www.hmi.org/
Lambda Legal (advocacy group for legal protections)	http://www.lambdalegal.org/issues/transgender-rights
PFLAG (support groups for families)	http://community.pflag.org/transgender
SAGE (service and advocacy group for LGBT elderly)	http://www.sageusa.org/issues/transgender.cfm
Servicemembers Legal Defense Network (advocacy and support of LGBT military personnel)	http://www.sldn.org/pages/transgender-people-and-military-service
Society for the Psychological Study of Lesbian, Gay and Bisexual Issues—Division 44 of the American Psychological Association (public education and professional support)	http://www.apa.org/about/division/div44.html
World Professional Association for Transgender Health (WPATH) (professional support and	http://www.wpath.org/

In a survey of psychiatrists, many reported difficulty distinguishing gender dysphoria from “gender confusion” that may occur as an epiphenomenon of psychiatric illnesses, notably personality, mood, dissociative, and psychotic disorders. The authors suggested that GD is likely to be underdiagnosed leading to inappropriate treatment. These surveys of medical professionals’ training are consistent with the National Transgender Discrimination Survey in which 50 percent of transgender respondents believed that clinicians were not aware of their unique needs. In recognition of these deficiencies calls for increased clinician training in the provision of care to transgender individuals have been made by several distinguished and authoritative bodies including the Institute of Medicine, the Department of Health and Human Services, and The Joint Commission. Ideally, training in SOGI minority issues, including gender variance, transgender and gender dysphoria should be embedded within the broader curriculum of human sexuality and integrated across all levels of medical school and residency training. Psychiatric training should include: (1) The evolution of concepts of gender, gender variance, and associated terminology; (2) the evolution of medical approaches to gender variance, including the changing roles of mental health professionals in transgender health care as reflected in successive versions of the WPATH SOC and the emergence of informed consent models; (3) common child and adolescent presentations; evaluation of gender variant youth; assessment and management of coexisting psychopathology in minors; treatment options, including pubertal suppression; persistence and desistence of gender dysphoria of childhood; and family concerns; (4) common adult presentations, the process of gender transition and other options for authentic gender expression; assessment and management of concurrent psychiatric illness, and stage of life concerns; (5) mental health assessments for cross-sex hormones and gender-affirming surgery; (6) presentations, evaluation, and management of gender dysphoria in patients with somatic intersexuality; (7) complex presentations, nonbinary gender identities, and the role of the mental health professional in alternative models of treatment; and (8) relevant jurisdictional and institutional policies including those pertaining to respectful treatment of gender variant individuals as well as those that affect access to gender transition services. While the goal of psychiatric residency training is not to produce gender specialists, the general psychiatrist, regardless of specialty or work setting, should be prepared to perform a culturally competent initial assessment of gender variant individuals, including a formulation based on the current version of the DSM, provide initial treatment and provide appropriate referrals.

Research. Some future research directions on the origins of gender identity and gender dysphoria are addressed previously in this chapter. Other areas of research include:

(1) Identification of disparities and barriers to care for transgender individuals is a first step toward addressing them. Most of what is presently known comes from convenience samples which are not fully reliable. The US Department of HHS is now gathering systematic data. Larger-scale studies will become possible if gender identity and natal gender are routinely gathered as demographic variables in health care settings and recorded in EHR as recommended by HHS, IOM, and Joint Commission.

(2) Mental health disparities, including mood and anxiety disorders, substance abuse and suicidality have been detected in studies to date. It is important to research causes of these disparities and to find effective means of intervention. Considerable research suggests minority stress and discrimination, including harassment and bullying, play a significant role. Limited research suggests peer support as a means of ameliorating the effects of minority stress. More studies are needed to elucidate not only effective means for overcoming the effects of minority stress but to identify the causes of intolerance of gender variance and to address them.

(3) There is need to compare quality of life outcomes of diverse treatment approaches to gender variance in childhood. Robust experimental designs, however, such as randomized trials in this area are neither ethical nor feasible. Very few parents would allow the treatment of their child to be randomly assigned. Concern has been expressed that attempts to "normalize" gender expression may have long-term harmful effects on self-esteem. Longitudinal follow-up studies without randomization could potentially address this issue.

(4) More longitudinal studies following pubertal suppression are needed. Recent studies by Dutch researchers found that when puberty suppression was done in very carefully selected minors, following gender reassignment surgery as adults they scored as well or better than the general population on measures of psychological and social adjustment. More studies are needed to determine if these results will generalize to those receiving this treatment in other clinical settings. Since pubertal suppression is a relatively recent intervention, long-term follow-up studies are needed to rule out long-term risks.

(5) Other than gatekeeping, what role should mental health professionals play in gender transition? Follow-up studies are needed to examine how particular mental health interventions influence long-term adjustment and subjective satisfaction. Matching comparison groups—those treated with particular mental health interventions versus those who were not—for potentially confounding variables, however, would be difficult.

INTERNET RESOURCES

The websites listed in Table 21.4–2 offer useful resources for gay men, lesbian, and bisexual patients, their families, and for mental health professionals themselves.

REFERENCES

Benjamin H. *The Transsexual Phenomenon*. New York: Julian Press; 1966.

Blanchard R. Early history of the concept of autogynephilia. *Arch Sex Behav*. 2005;34:439–446.

Bockting W, Coleman E. Deutsch MB, et al. Adult development and quality of life of transgender and gender nonconforming people. *Curr Opin Endocrinol Diabetes Obes*. 2016;23:188–197.

Brown GR, Jones K. Mental health and medical health disparities in 5135 transgender veterans receiving healthcare in the Veterans Health Administration: a case-control study. *LGBT Health*. 2016;3:122–131.

*Byne W, Bradley SJ, Coleman E, et al. Report of the American psychiatric association task force on treatment of gender identity disorder. *Arch Sex Behav*. 2012;41(4):759–796.

Coates S, Wolfe S. Gender identity disorder in boys: the interface of constitution and early experience. *Psychoanal Inq*. 1995;51:6.

Cohen-Kettenis PT, Pfäfflin F. The DSM diagnostic criteria for gender identity disorder in adolescents and adults. *Arch Sex Behav*. 2010;39:499–513.

Dhejne C, Öberg K, Arver S, Landén M. An analysis of all applications for sex reassignment surgery in Sweden, 1960–2010: prevalence, incidence, and regrets. *Arch Sex Behav*. 2014;43:1535–1545.

*Drescher J. Queer diagnoses: parallels and contrasts in the history of homosexuality, gender variance, and the Diagnostic and Statistical Manual (DSM). *Arch Sex Behav*. 2010;39:427–460.

Drescher J. Controversies in gender diagnoses. *LGBT Health*. 2014;1(1):10–14.

Drescher J, Byne W, eds. *Treating Transgender Children and Adolescents: An Interdisciplinary Discussion*. New York: Routledge; 2013.

Drescher J, Cohen-Kettenis P, Winter S. Minding the body: situating gender diagnoses in the ICD-11. *Int Rev Psychiat*. 2012;24(6):568–577.

Drescher J, Pula J. Ethical issues raised by the treatment of gender variant prepubescent children. *Hastings Cent Rep*. 2014;44(Suppl 4):S17–S22.

Drescher J, Cohen-Kettenis PT, Reed GM. Gender incongruence of childhood in the ICD-11: Controversies, proposal, and rationale. *Lancet Psychiatry*. 2016;3(3):297–304.

Fisher AD, Castellini G, Risotri J, et al. Cross-sex hormone treatment and psychobiological changes in transsexual persons: two-year follow-up data. *J Clin Endocrinol Metab*. 2016;101:4260–4269.

Grant JM, Mottet LA, Tanis J, Harrison J, Herman JL, Keisling M. *Injustice at Every Turn: A Report of the National Transgender Discrimination Survey*. Washington, DC: National Center for Transgender Equality and National Gay and Lesbian Task Force; 2011.

Green R. *The "Sissy Boy Syndrome" and the Development of Homosexuality*. New Haven, CT: Yale University Press; 1987.

*Institute of Medicine. *The Health of Lesbian, Gay, Bisexual and Transgender People: Building a Foundation for Better Understanding*. Washington, DC: The National Academies Press; 2011.

Kreukels BPC, Steensma TD, de Vries ALC, eds. *Gender Dysphoria and Disorders of Sex Development*. New York: Springer; 2013.

Levounis P, Drescher J, Barber ME. *The LGBT Casebook*. Washington, DC: American Psychiatric Publishing; 2012.

Meyer-Bahlburg HF. From mental disorder to iatrogenic hypogonadism: Dilemmas in conceptualizing gender identity variants as psychiatric conditions. *Arch Sex Behav*. 2010;39:461–476.

Money J, Ehrhardt AA. *Man and Woman, Boy and Girl: The Differentiation and Dimorphism of Gender Identity from Conception to Maturity*. Baltimore, MD: Johns Hopkins University Press; 1972.

Moser C. Autogynephilia in women. *J Homosex*. 2009;56:539–547.

Serano J. The case against autogynephilia. *Int J Transgenderism*. 2010;12:176.

Steensma TD, McGuire JK, Kreukels BP, Beekman AJ, Cohen-Kettenis PT. Factors associated with desistence and persistence of childhood gender dysphoria: a quantitative follow-up study. *J Am Acad Child Psy*. 2013;52(6):582–590.

*The Joint Commission. *Advancing Effective Communication, Cultural Competence, and Patient- and Family-Centered Care for the Lesbian, Gay, Bisexual, and Transgender (LGBT) Community: A Field Guide*. Oak Brook, IL: Joint Commission Resources; 2011.

*World Professional Association for Transgender Health (WPATH). *Standards of Care for the Health of Transsexual, Transgender and Gender Non-Conforming People, 7*th *Version,* 2011. Available at http://www.wpath.org/.

Wylie K, Knudson G, Sharful IS, et al. Serving transgender people: clinical care considerations and service delivery models in transgender health. *Lancet*. 2016;388:401–411.

Zucker KJ. The DSM diagnostic criteria for gender identity disorder in children. *Arch Sex Behav*. 2010;39:477–498.

Zucker KJ, Cohen-Kettenis PT, Drescher J, Meyer-Bahlburg HF, Pfäfflin F, Womack WM. Memo outlining evidence for change for gender identity disorder in the DSM-5. *Arch Sex Behav*. 2013;42:901–914.

▲ 21.5 Sexual Addiction

AVIEL GOODMAN, M.D.

A disorder in which a person's driven sexual behavior is poorly controlled and continues despite harmful consequences has a documented history of more than two millennia. The Pali Canon, the standard scriptures of Theravada Buddhism that were compiled from oral tradition and written in 29 BCE, includes the following passage:

> Then early in the morning the Blessed One, having put on his robes and carrying his bowl and outer robe, went into Savatthi for alms. He saw that most of the people in Savatthi were excessively attached to sensual pleasures, that they lived delighting in, addicted to, infatuated with, fastened to, absorbed in sensual pleasures. Then, on realizing the significance of that, the Blessed One on that occasion exclaimed:
>
> Blinded by sensuality, covered by the net, veiled with the veil of craving, bound by the bond of heedlessness like fish in the mouth of a trap, they go to aging and death like a suckling calf to its mother. [Udana VII.4 Ka¯mesu Satta Sutta, tr. Thanissaro Bhikkhu]

A similar condition was described in 1886 by the psychiatrist Richard von Krafft-Ebing:

> [H]is sexual appetite is abnormally increased to such an extent that it permeates all his thoughts and feelings, allowing of no other aims in life, tumultuously, and in a rut-like fashion demanding gratification without granting the possibility of moral and righteous counter-presentations, and resolving itself into an impulsive, insatiable succession of sexual enjoyments. . . . This pathological sexuality is a dreadful scourge for its victim, for he is in constant danger of violating the laws of the state and of morality, of losing his honor, his freedom and even his life.

Krafft-Ebing identified this condition as *hypersexuality*, which is generally defined as excessive sexual interest (or urges) or activity. It was an advance over its predecessors, *satyriasis* and *nymphomania*, which are defined as uncontrollable or excessive sexual desire in a man and in a woman, respectively. The definition of *hypersexuality* is more closely related to sexual behavior, which is more objectively observable and quantifiable than is sexual desire. In addition, hypersexuality is gender-neutral.

In 1945, the psychoanalyst Otto Fenichel coined the term *sexual addiction* to designate the condition of driven sexual behavior that is under discussion. Whereas excessiveness in the definition of hypersexuality represents a quantitative dimension, the term *sexual addiction* directs attention toward indications of a sexual behavior's adverse consequences and a person's impaired control over the behavior. It also implies a relationship between this disorder and other disorders the names of which also include the word *addiction*.

The next section addresses terminological and diagnostic matters that

pertain to this condition. Before proceeding, three characters are introduced whose stories give human faces to the clinical abstractions, provide examples of how a condition of driven sexual behavior may look in a person's life, and illustrate the potentially tragic seriousness of the condition. (The characters are fictional and are presented as illustrative examples, not case reports.)

Ronnie, an electrician in his mid-20s, had used alcohol and other drugs heavily in high school. He dropped out of college in his freshman year and went to work in his uncle's electrical contracting business. After a serious accident on the job while he was intoxicated, he quit recreational use of psychotropic substances.

Over the next several months, he found that his sexual fantasies and urges to masturbate were becoming more frequent and more intense than they previously had been. Since his middle teens, Ronnie had masturbated nearly every night before going to sleep, unless he was having sex with a girlfriend or a woman he had picked up. Some mornings during his first year of being drug-free, Ronnie would get strong urges to masturbate. If he did not act on these urges, he would remain "horned up" all day and would feel restless, distracted, and irritable. Consequently he began to masturbate in the mornings as well, even though doing so sometimes resulted in his arriving to work late.

Some months later, Ronnie began to get urges to masturbate when he was at work, particularly if he saw a woman at the job site. Again, he found that if he did not masturbate, he would become uncomfortable and unable to focus on his work. No longer sufficiently excited by his fantasies, Ronnie began to watch free online pornographic videos while he masturbated. He also discovered peer-to-peer file-sharing sites from which he could at no charge (but illegally) download better quality, longer, and more varied porn videos.

Before long, Ronnie was masturbating three or four times a day and even more when he felt under stress. He was chronically tired from lack of sleep, had received a second warning about tardiness and inattentiveness at work, and felt disgusted with himself. He tried several times to stop masturbating. But each time, sexual fantasies accompanied by arousal would intrude into his consciousness throughout the day and he would feel as though he were going to blow up. The only way he knew to free himself from the building pressure was to masturbate. At the beginning of a new sexual relationship, he would delete his collection of pornography and resolve to quit masturbating. Within a few months, though, masturbation would creep in and take over, his sexual interest in his girlfriend would fade, and the relationship would fall apart.

As Ronnie became increasingly bored with familiar porn videos and new ones with the same old plot formulas, he began to accompany his masturbation with increasingly sadomasochistic videos. He also started using telephone sex and online sex-cam services, neither of which were free of charge. When he had "maxed out" his credit cards, he obtained new ones that he soon ran up to their limits. He got behind on his rent and the power company threatened to cut him off unless he paid his bill.

Socially isolated, deeply in debt, and about to lose his job, Ronnie realized that his preoccupation with masturbation and pornography was ruining his life. But he felt helpless. Meanwhile, he also was beginning to sense that pornography, phone sex, and sex-cams no longer excited him as much as they used to. He started going to shopping malls specifically for the purpose of looking at women and girls, whom he would later fantasize about raping while he masturbated. When he then began to feel compelled to study the patterns of the women who worked in the malls and to compile schedules of their comings and goings, he became scared that he was in danger of acting on his fantasies. At that point, he sought help.

Lori was an MBA graduate and a highly effective manager. Yet she was unable to hold a job for more than a year. When she began a new job, her performance was exemplary. After about 2 to 6 months, she would begin to act as though she were distracted and less engaged. The quality of her work would gradually decline and her work relationships would become strained. Eventually she would be asked to resign and would comply.

Lori was in love with being in love. She loved the boost to her self-esteem and self-confidence; the soft ongoing sexual arousal that made her feel fully alive; the sexual excitement that thrilled her when she just thought about her lover; and the "mind-blowing" orgasms she experienced during sex with him. Being in love filled her emptiness, made her whole, kept the pain and terror out of her awareness, and gave her a reason to live.

When Lori was in love with a man, nothing else in her life felt as important to her as did having sex with him. She would want sex desperately. If circumstances prevented her from having sex with him, she would become restless, then agitated, then frantic. She could interrupt this sequence by masturbating during the restless stage before she became agitated. Otherwise, she could terminate such episodes only by drinking herself into a stupor. When she was in love she gradually neglected everything else in her life, including her work. When she was not in love, she would search desperately for a man and typically would find someone with whom to fall in love and have sex. Often, her feeling of being in love did not last. Often, the man lost interest before she did.

On several occasions Lori was reminded that her trouble with employment was related to her love life. When she was in love, each time she tried to hold back or defer sexual activity—for example, to complete a work assignment or attend a meeting—she experienced the aforementioned sequence of restlessness, agitation, and frenzied panic. When she was not in love and refrained from her desperate quest—primarily by staying home during evenings and weekends—she experienced abandonment feelings that intensified, threatened to overwhelm her, and were increasingly accompanied by suicidal thoughts. Twice she had attempted to kill herself.

After Lori's 30th birthday she found herself increasingly drawn to "bad boys," men with an air of danger who did not follow rules. One evening, some friends of a boyfriend who had moved in with her joined them for dinner. An hour later, a dozen FBI agents stormed in and placed everyone under arrest. Two kilos of uncut cocaine were later found in Lori's house. Lori and her boyfriend were arrested for possession with intent to distribute and operating a continuing criminal enterprise.

Lori was allowed to plead guilty to a minor charge in exchange for her cooperation with prosecution. A year or so after the trial, she began to feel hints of the familiar craving to fall in love and be swept into a delicious maelstrom of sex and passion. She had by then come to realize that many of the misfortunes in her life had resulted from following where that craving led. When the hints became stronger, she sought help.

Wayne, a college dropout in his middle 40s, had difficulty holding a job despite being a capable computer programmer and a gifted musician. He had no close friends and spent much of his free time by himself, either working on his computer or playing his electric piano through headphones. Sometimes he would go to a casino, which he said that he did in order to be around other people.

Wayne had been involved in two brief relationships with women, both of whom he drove away by his demanding possessiveness and paranoid jealousy. He tended to develop obsessive crushes on women whom he hardly knew and then would feel driven to watch and follow them. Complaints of harassment by female coworkers had led to his being fired from two jobs and once he was arrested for stalking.

Wayne masturbated several times a day, even at work. At night, he sometimes called random telephone numbers until he found a woman who would answer his questions about her sex life while he masturbated. When he had an urge to masturbate and did not act on it, he experienced a disorganizing panic that would intensify until he felt that he was about to burst. He found that he could not resist these urges, even when to succumb could jeopardize his job or get him into trouble with the law.

The strongest urge that Wayne experienced was to expose his penis to a woman and masturbate while she looked on in shock and fear. For this reason, he particularly liked to expose himself in elevators, when he and a woman were the only passengers. Even after two arrests, one conviction, and one occasion on which a woman's boyfriend beat him severely, Wayne felt driven to continue riding elevators and hunting for unaccompanied women. One time, Wayne targeted a woman who had once been raped. Having vowed that she would never be victimized like that again, she carried a handgun. When Wayne turned toward her with his

penis exposed, she panicked and shot him. He survived, thanks in part to the woman's immediate 911 call. Six months after being discharged from the hospital, he was arrested again for exposing his penis to a woman in an elevator.

While the three persons who are described in the case examples differed in many respects, they had one trait in common. All engaged in some form of sexual behavior in a pattern that was characterized by two key features: recurrent failure to control the behavior and continuation of the behavior despite significant harmful consequences.

DIAGNOSIS AND RELATED MATTERS

Terminology

Agreement on the name of this disorder of driven sexual behavior is not universal. The primary contenders have been *sexual compulsivity* (or *compulsive sexual behavior*) and *sex addiction* (or *sexual addiction*). *Sexual compulsivity* is the name given to this condition by investigators who argue that the symptomatic sexual behavior is compulsive because it functions to reduce anxiety and other painful affects. While phenomenological characteristics and treatment research support classification of this condition as an addictive disorder, the clearest argument against describing its symptomatic behavior as compulsive appears in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) and also in previous editions of the DSM. The definition of compulsion (compulsive behavior) in DSM-5 includes its function of providing relief from anxiety and other painful affects. However, it also specifies that compulsions are not performed to produce pleasure. Although reduction of anxiety or distress contributes to a person's motivation to engage in the sexual behavior that characterizes this disorder, pleasure or gratification also contributes. The relationship between compulsion and the condition being considered here is explicitly clarified in the differential diagnosis of OCD:

> Certain behaviors are sometimes described as "compulsive," including sexual behavior (in the case of paraphilias), gambling (i.e., gambling disorder), and substance use (e.g., alcohol use disorder). However, these behaviors differ from the compulsions of OCD in that the person usually derives pleasure from the activity and may wish to resist it only because of its deleterious consequences.

A suitable name for the condition under discussion would designate behavior that, like compulsive behavior, (1) feels driven, (2) entails harmful or unpleasant consequences, and (3) functions to reduce anxiety or other painful affects; and also that—unlike compulsive behavior—(4) functions to produce pleasure or gratification. *Addiction* is the term that best fits this description. Drug addiction, for example, has the characteristics of feeling driven, entailing harmful consequences, and functioning both to relieve painful affects and to produce pleasure.

However, determination of whether "sex addiction" is the most appropriate designation for the condition of driven sexual behavior that is being considered here requires a clear and meaningful definition of addiction.

Definition of Addiction

As noted in DSM-5, neither tolerance nor withdrawal is necessary nor are both together sufficient for a diagnosis of substance use disorder. The characteristics that seem to be both necessary and sufficient for identifying a pattern of psychoactive substance use as a substance use disorder are (1) recurrent failure to control the use of one or more psychoactive substances and (2) continuation of psychoactive substance use despite significant harmful consequences. ("Recurrent failure to control" does not mean that addicted individuals invariably lose control when they use drugs. It means that their predictions that they would refrain from or limit their drug use have repeatedly proven to be unreliable.)

The preceding discussion of terminology identified two critical characteristics of addictive behavior. It can reduce anxiety and other painful affects, as does compulsive behavior. In addition, it produces pleasure or gratification (that may diminish with repetition). Combining this distinctive dual motivational function—positive as well as negative reinforcement—with the key features that distinguish drug addiction from simple drug use yields a workable, behaviorally nonspecific definition of addiction, as follows. *Addiction* is a condition in which a behavior that can function both to produce pleasure and to reduce anxiety or other painful affects is employed in a pattern that is characterized by two key features: (1) recurrent failure to control the behavior and (2) continuation of the behavior despite significant harmful consequences.

The definition of a specific addictive disorder can be derived from the definition of addiction by replacing the phrase, "a behavior that can function both to produce pleasure and to reduce painful affects" with the name of a behavior that can perform both of these functions (e.g., alcohol consumption, gambling, eating, sexual behavior). Accordingly, sex addiction is defined as a condition in which some form of sexual behavior is employed in a pattern that is characterized by two key features: (1) recurrent failure to control the behavior and (2) continuation of the behavior despite significant harmful consequences. In other words, sex addiction is a condition in which some form of sexual behavior relates to and affects a person's life in such a manner as to accord with the definition of addiction. A sex addict is a person who engages in some form of sexual behavior addictively—that is, in a pattern that is characterized by recurrent failure to control and continuation despite significant harmful consequences.

In a reprise of terminology matters, the definition of addiction as a

pattern of behavior indicates that addiction is accurately identified not as something that a person has, like cancer, but as something that a person does. Also inaccurate is the characterization of a person as addicted to a substance or behavior, which would imply that the person is a passive victim over whom the substance or behavior has control. The engine of addiction is not in a substance or behavior. It is in the person. The link between a person and an addictive disorder or pattern of behavior is most accurately described by a statement to the effect that a person engages in one or more behaviors addictively. Some common phrases such as "addicted to porn" are inaccurate not only for the reasons just mentioned but additionally because they misidentify the pathognomonic behavior. In an addictive disorder that involves pornography, the behavior that is used addictively is not the viewing of pornography. It is masturbation, for which looking at pornography serves as a facilitator or enhancer. Terminology was considered worthy of further discussion because imprecise use of "addiction" and related terms is sometimes exploited as an opportunity to discredit the term, however much the imprecision resides not in the term but in its user.

Diagnostic Criteria for Sex Addiction

DSM-5 excludes several behavioral addictions, including sex addiction, on grounds that peer-reviewed evidence is insufficient to establish their diagnostic criteria. Meanwhile, the generation of peer-reviewed evidence depends on investigators of this disorder having a valid and reliable set of provisional diagnostic criteria by which to identify subjects who are appropriate for their studies.

Provisional diagnostic criteria for sex addiction that follow the DSM format can be derived from the DSM-5 criteria for any substance use disorder, the paradigmatic addictive disorder, by replacing the name of the substance and its use—for example, *alcohol* and *alcohol use*—with the term *sexual behavior*, replacing *taken* with *engaged in*, and replacing "a characteristic withdrawal syndrome for the substance" with "a characteristic biopsychological withdrawal syndrome of physiologically and/or psychologically described changes upon discontinuation of the sexual behavior." The set of criteria that is presented here demonstrates that diagnostic criteria for sex addiction can be developed that are consistent with the form and content of the current DSM, are expressed in the same descriptive language, and are likely to be as reliable and valid as are the DSM-5 diagnostic criteria for substance use disorder. To increase specificity and reliability, a diagnosis of sex addiction requires that at least four of the diagnostic criteria be met. (A diagnosis of substance use disorder requires only that two criteria be met.)

Sex Addiction

A maladaptive pattern of sexual behavior that has led to *clinically significant impairment or distress*, as manifested by the occurrence within a 12-month period of four (or more) of the following:

1) the sexual behavior is often engaged in over a longer period, in greater quantity, or at a higher level of intensity than was intended
2) efforts to cut down or control the sexual behavior have repeatedly been unsuccessful
3) a great deal of time is spent in activities necessary to prepare for the sexual behavior, engage in the behavior, or recover from its effects
4) the subject experiences a craving (strong desire or urge) to engage in the sexual behavior
5) engaging in the sexual behavior has resulted in failure to fulfill significant responsibilities at work, school, or home
6) important social, occupational, or recreational activities are given up or reduced as a result of engaging in the sexual behavior
7) the sexual behavior continues despite persistent or recurrent social or interpersonal problems that its effects have caused or exacerbated
8) the sexual behavior continues despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by the behavior
9) the sexual behavior is recurrently performed in situations in which doing so is physically hazardous
10) tolerance, as defined by either of the following:
    a) producing the same levels of desired effects as on earlier occasions requires a frequency or intensity of sexual behavior that is markedly greater than had been required in the past
    b) engaging in the sexual behavior at the same frequency and intensity as on earlier occasions produces levels of desired effects that are markedly lower than had been produced in the past
11) withdrawal, as manifested by either of the following:
    a) a characteristic biopsychological withdrawal syndrome of physiologically and/or psychologically described changes upon discontinuation of the sexual behavior
    b) the same (or a closely related) sexual behavior is engaged in to relieve or avoid withdrawal symptoms

These diagnostic criteria are far from the final word on diagnosing sex addiction. They are provisional and are likely to need revision in the light of further developments in research or theory. For example, investigators might find that the number of criteria required (four) is too low or that the required time frame (12 months) is unnecessarily long. They might determine that adding (or substituting) items from the DSM-5 diagnostic criteria for gambling disorder—such as preoccupation, engaging in the

behavior to relieve dysphoric moods, restlessness or irritability when attempting to decrease or discontinue the behavior, or lying to conceal the extent of the behavior—enhances their discriminative utility. At the very least, these criteria provide a starting point for research. The research function of diagnostic criteria is critically important, particularly for newly recognized conditions. The inclusion of a condition in the next DSM depends on scientific research that documents its legitimacy. Meanwhile, a condition can be researched only if clear, valid, and reliable diagnostic criteria enable researchers to recognize it. Multiple varied sets of diagnostic criteria limit the generalizability of research results as well as undermining claims that the condition is a valid diagnostic entity.

Significantly, no form of sexual behavior in itself constitutes sex addiction. Whether a pattern of sexual behavior qualifies as sex addiction is determined not by the type of behavior, its object, its frequency, or its social acceptability, but by how the behavior relates to and affects a person's life. Any sexual behavior has the potential to be engaged in addictively but constitutes an addictive disorder only to the extent that it occurs in a pattern that meets the diagnostic criteria or accords with the definition. The key features that distinguish sex addiction from other patterns of sexual behavior are: (1) the person is not reliably able to control the sexual behavior and (2) the sexual behavior has significant harmful consequences and continues nonetheless.

Sex addiction can now be distinguished from the diagnostic entity in DSM-5 that is most similar to it, paraphilic disorder. DSM-5 defines *paraphilia* as "any intense and persistent sexual interest other than sexual interest [sic] in genital stimulation or preparatory fondling with phenotypically normal, physically mature, consenting human partners." It defines *paraphilic disorder* as "a paraphilia that is currently causing distress or impairment to the individual or a paraphilia whose [sic] satisfaction has entailed personal harm or risk of harm to others." The most apparent distinction is that sex addiction can be diagnosed when its diagnostic criteria are met whether the symptomatic behavior is paraphilic or non-paraphilic, while paraphilic disorder is limited to paraphilias. A second distinction is that impaired ability to control the symptomatic behavior is prominent in the diagnosis of sex addiction but is mentioned in neither the definition of paraphilic disorder nor the diagnostic criteria for any of the paraphilic disorders. Third, risk of harm to others as a result of symptomatic behavior is sufficient to warrant a diagnosis of paraphilic disorder but not a diagnosis of sex addiction. Fourth, DSM-5 defines paraphilia as an interest, not a behavior. Diagnostic criterion B for each of the paraphilic disorders is "The individual has acted on these sexual urges with a nonconsenting person, *or the sexual urges or fantasies* [author's italics] cause clinically significant distress or impairment in social, occupational, or other important areas of functioning." The definition of

and diagnostic criteria for sex addiction depend on objectively observable events: sexual behavior, comparison between addicts' predictions that they would refrain from or limit their sexual behavior and their subsequent sexual behavior, and consequences of the sexual behavior. Sexual interests, urges, or fantasies *per se* are not relevant to a diagnosis of sex addiction.

Clinical Examples. A few clinical examples are presented to illustrate the considerations that are involved in making a diagnosis of sex addiction.

Dale, a dealer of fine art, began psychoanalysis to help him get past his lifelong depression and inability to sustain an intimate relationship. Ever since he quit drinking 6 years earlier he had been desperately lonely. Consciously he wanted nothing more than to find a partner, someone with whom he could share his life. However, he was very anxious in social interactions. He feared that he would not be able to hold up his end of a conversation and would then be humiliated. He avoided situations in which conversation was likely. Instead he would go to parks or adult theaters where he could meet other men for anonymous sex. After these sexual encounters he felt dirty and disgusted with himself. Often he would suspect that he had picked up HIV. He also had been beaten several times and robbed twice. But when he was alone or even when he anticipated being alone—for example, when he was driving home after work or after an evening activity—he would experience an intense anxiety that felt like he was coming apart in a billion pieces. He could not tolerate this feeling, and sexual activity relieved it. Masturbating to pornography at home took the edge off his anxiety and usually enabled him to sleep. However, he increasingly was finding that it did not help as much as it once had done. On numerous occasions he tried to get involved in cultural or religious activities where he would be likely to meet educated, stable gay men. Every time he soon found himself back at a park or a theater.

Steve was a software developer in his 40s who was in treatment for depression, anxiety, and marital difficulties. Sex was among the sources of tension in his marriage. He wanted to have sex with his wife every day, sometimes two or even three times a day, and he became upset when she did not comply. When his wife declined to have sex with him he felt desperate. He feared that she did not love him, that he was not good enough for her, that she was tired of him and was preparing to leave him. On such occasions, he usually withdrew and immersed himself in work. He rarely masturbated and he did not seek sex elsewhere. When his wife's lupus flared up, Steve cared for her sensitively and she expressed appreciation and gratitude. At such times, he felt needed and valued, and he rarely thought of sex. Steve's desire for sex occasionally offended his wife, who felt then that he would rather have sex than talk with her. Sometimes when she complied with Steve's requests for sex she resented him. Further inquiry did not reveal other harmful consequences from Steve's sexual behavior, nor did it suggest that he had any difficulty controlling his sexual behavior.

Jolene was a married emergency room nurse in her early 30s who was "in love" with two extramarital men with whom she was sexually involved. Meanwhile she loved her husband Ed and could not stand the thought of leaving him. She trusted him and felt that he understood her. He knew of Jolene's recurrent pattern of intense infatuation with a new man, resisting, giving in to overpowering sexual attraction, and then becoming bored. When she was not sexually involved with at least one man outside her marriage she felt desperate, panicky, and empty. Typically she would begin a new involvement as soon as she sensed that she was beginning to become bored with her current extramarital man. Sure, she noted, she had tried to rein it in, many times. After all, she continued, it had cost her two marriages, her kids, her place in medical school, a whole lot of time, and a whole lot more emotional hell. She had tried meditation, counseling, and women's support groups. She had joined a fundamentalist church. She even had taken excessive doses of medroxyprogesterone (Provera) to decrease her sex

drive until she developed a pulmonary embolus from which she almost died. Jolene had become involved with Ed while she still was married to her second husband. She lived in constant fear that Ed too would get fed up with her and throw her out. Now one of her lovers was talking about killing himself if Jolene did not marry him. The other lover was talking about confronting Ed. Jolene felt overwhelmed and had developed symptoms of major depression and panic disorder.

Charla, a hairstylist and fashion model in her 20s, agreed to accompany her boyfriend Mike to a session with his psychotherapist. Mike was upset that Charla was having sex with other men and he thought that she might be a sex addict. Charla said that she had told Mike shortly after they met that sex was part of her career and if he did not think he could stand it they should not get involved. Having sex with wealthy and powerful men advanced her success as a model, she explained, as well as bringing her money and expensive gifts. She acknowledged that what she was doing was like prostitution; but it was clean and legal and she was in charge. She stated that she certainly did not do it for the sex, which did nothing for her. She was not orgasmic with these men. She likened having sex to cutting hair. Her rewards were the finished product, a satisfied customer, and a nice tip. She also enjoyed her sense of control over these wealthy and powerful men. Charla stated that she had never engaged in more sexual behavior than she intended and that she had never tried or wanted to decrease her sexual behavior. She also denied that she had experienced any kind of withdrawal symptoms or desire to engage in substitute behavior when she was not sexually active. She recognized that her sexual behavior was bothering Mike but she said that it was his problem. She was making ten times as much money as her father ever did and was not about to give that up because Mike's fragile male ego could not deal with it. Mike acknowledged that she seemed to be in control of her sexual behavior and that he was not aware of any problems that resulted from her sexual behavior, other than its effect on him.

The clinical example of Dale seems to meet all of the criteria except 1. The example of Steve might meet criterion 3 but does not seem to meet any of the other criteria. The example of Jolene meets criteria 2, 3, 6, 8, and probably 11. Charla probably meets criterion 5. She might or might not meet criterion 6, depending on the importance (to her) of the activities with Mike that she gives up to have sex with other men. Accordingly, Dale and Jolene merit diagnoses of sex addiction according to the diagnostic criteria while Steve and Charla do not. (The conclusion is not that Charla's behavior is acceptable nor that she is psychiatrically healthy, only that she does not merit a diagnosis of sex addiction.) An informal assessment based on the definition of sex addiction reaches the same conclusions. Dale and Jolene demonstrate patterns of sexual behavior that are characterized by recurrent failure to control and continuation despite significant harmful consequences while Steve and Charla do not.

Differential Diagnosis

A number of other conditions can at first glance look like sex addiction. The hypersexual activity and paraphilic behaviors that often characterize sex addiction can occur also as manifestations of underlying brain pathology and occasionally are its earliest or most prominent symptoms. Paraphilic or hypersexual behavior can occur in the context of a seizure disorder, especially in association with temporal lobe epilepsy. More broadly, any

disorder that is associated with impairment of cerebral functioning can weaken normal inhibitory controls and thereby allow the expression of sexual behaviors that ordinarily are suppressed. Disinhibited sexual behavior has been reported particularly in patients with frontal lobe lesions and in elderly patients with dementia. Hypersexual behavior can occur also as a side effect of medication, particularly antiparkinsonian agents. Finally, sexually aggressive behavior has been associated with elevated levels of testosterone. Distinguishing these conditions from sex addiction is usually facilitated by the presence of additional symptoms or circumstances that suggest the underlying organic disturbance, although altered sexual behavior may be the earliest manifestation in some cases of brain pathology. Clues that invite an organic evaluation include onset in middle age or later, regression from previously normal sexuality, excessive aggression, report of auras or seizure-like symptoms prior to or during the sexual behavior, and impaired perceptual or motor skills. The most useful instrument for determining whether a case of paraphilic or hypersexual behavior represents sex addiction is the set of diagnostic criteria for sex addiction. Patterns of paraphilic or hypersexual behavior that are not symptoms of sex addiction are generally characterized neither by a persistent desire to cut down or control the behavior nor by tolerance or biopsychological withdrawal symptoms on discontinuation of the behavior (usually affective discomfort, irritability, or restlessness). The diagnostic criteria for sex addiction are useful also in distinguishing sex addiction from nonaddictive patterns of exploitative or aggressive sexual behavior that can occur with antisocial personality disorder.

Obsessions and compulsions with sexual content can occur in OCD. Sexual obsessions were reported in 32 percent of the OCD patients in a large study. However, the content of these obsessions consisted most often not of sexual fantasies but of fears of acting on sexual impulses or fears of being a pervert. Sexual fantasies are typical in sex addiction. As noted earlier, the symptoms of sex addiction differ from sexual obsessions and compulsions primarily in that the former are associated with sexual arousal and sexual pleasure while the latter are typically accompanied not by sexual arousal but by anxiety.

A syndrome that meets the diagnostic criteria for sex addiction can occur in the context of other psychiatric disorders, including bipolar disorder, schizophrenia, personality disorders, and psychoactive substance use disorder. When the diagnostic criteria for both sex addiction and another psychiatric disorder are met both diagnoses are warranted, regardless of whether sex addiction might be secondary to the other psychiatric disorder. The diagnosis of sex addiction is a descriptive designation of how a pattern of sexual behavior relates to and affects a person's life. It does not presume anything about how the pattern of sexual behavior developed nor is it precluded by the presence of other conditions

that may have influenced its development.

Epidemiology

Two early studies reported that 3 to 6 percent of the American population suffer from sexual addiction and 5 percent meet diagnostic criteria for sexual compulsivity. Slightly higher percentages of sexually compulsive respondents were reported by Swedish studies of individuals who used the Internet for sexual purposes. A smaller Canadian study found that 5 of 40 community subjects (12.5 percent) were sex addicts. A more recent and larger Canadian study reported that 50 of 114 (43.9 percent) of incarcerated sex offenders and 21 of 117 (17.9 percent) of socioeconomically matched community males reached the addiction criterion in the Sexual Addiction Screening Test, which was used also in the study that reported a prevalence of 3 to 6 percent. The authors attributed the difference to the higher socioeconomic class of the subjects in the earlier study.

About 80 percent of individuals who use sexual behavior addictively are male. Gender differences in prevalence are greater for paraphilic than for nonparaphilic sex addictions. Within the nonparaphilic arena males tend to focus more on physical sexual gratification and thus more often engage addictively in masturbation, use of prostitutes, and impersonal sex. Females tend to focus more on the emotional aspects of sexuality and thus more often engage addictively in romantic relationships. Sex addiction is usually a chronic disorder with remissions and exacerbations.

Addictive use of sexual behavior typically begins in the teens, peaks between the ages of 20 and 40, and then gradually declines. The "modal subject" in a cohort that responded to an advertisement for evaluation and treatment of sex addictions and compulsive sexual behaviors was characterized as a 34-year-old, white, married, Catholic college graduate earning a middle-class income.

Comorbidity and Familial Patterns

A review of comorbidity findings is prefaced by the observation that little research has actually been conducted specifically with individuals who use sexual behavior addictively. For a variety of reasons, most of the reviewed research studied paraphilic subjects and sex offenders. While the group of sex addicts overlaps with the group of sex offenders and more so with the group of paraphilic subjects, these groups are not identical.

Researchers have noted significant comorbidity among different forms of sex addiction: Individuals who engage addictively in one form of sexual behavior are likely also to engage addictively in other forms of sexual behavior. High frequencies of comorbidity have been observed not only within the category of paraphilic sex addictions but also between paraphilic and nonparaphilic sex addictions. A variety of psychiatric disorders have

been reported to be comorbid with sex addiction at a rate that is higher than is their prevalence in the general population. The strongest trend that the research reveals is significant comorbidity of sex addiction with substance use disorder, gambling disorder, bulimia nervosa, and compulsive buying. Significant comorbidity between sex addiction and mood disorders, anxiety disorders, attention deficit disorder, and personality disorders also has been reported. The one family history study that has been published found a significantly increased prevalence of sex addiction and other addictive disorders (substance use disorder, eating disorders, and gambling disorder) in the parents of sex addict probands. More research is needed in the areas of sex addiction comorbidity and family history, particularly in light of the usefulness of such findings in understanding etiology and in formulating treatment plans.

Course of Illness

A recent search of PubMed and Google Scholar online resources found no publications that reported on the course of illness for sex addiction (or for sexual addiction, sexual compulsivity, compulsive sexual behavior, or paraphilia). In DSM-5, the course of illness descriptions for the paraphilic disorders mention possibilities but provide no information.

That sex addiction shares the addictive process with other addictive disorders suggests that its course of illness may be similar to those of the others. Thus far, the preceding statement is but a hypothesis in search of empirical investigation.

THE ADDICTIVE PROCESS

Other disorders in addition to sex addiction and the substance use disorders fit the definition of addiction and fulfill the diagnostic criteria for addictive disorder, including gambling disorder, binge eating disorder, bulimia nervosa, and kleptomania. In addition, they share a number of clinical and phenomenological features that are characteristic of addiction: (1) characteristic course—the disorder typically begins in adolescence or early adulthood and follows a chronic course with remissions and exacerbations; (2) behavioral features—narrowing of behavioral repertoire, continuation of the behavior despite harmful consequences; (3) individuals' subjective experience of the condition—sense of craving, preoccupation, excitement during preparatory activity, mood-altering effects of the behavior, sense of loss of control; (4) progressive development of the condition—craving, loss of control, narrowing of behavioral repertoire, and harmfulness of consequences all tending to increase as the duration of the condition increases; (5) experience of tolerance—as the behavior is repeated, its potency to produce reinforcing effects tends to diminish; (6) experience of withdrawal phenomena—psychological or physical

discomfort when the behavior is discontinued; (7) tendency to "relapse"—to return to harmful patterns of behavior after a period of abstinence or control has been achieved; (8) propensity for behavioral substitution—when the behavioral symptoms of the disorder have come under control, tendency for addictive engagement in other behaviors to emerge or intensify; (9) relationship between the condition and other aspects of affected individuals' lives—neglect of other areas of life as the behavior assumes priority; and (10) recurrent themes in the ways individuals with these conditions relate to others and to themselves—including low self-esteem, self-centeredness, denial, rationalization, and conflicts over dependency and control.

That these six conditions with diverse symptomatic behaviors—sex addiction, the substance use disorders, gambling disorder, binge eating disorder, bulimia nervosa, and kleptomania—meet the same set of diagnostic criteria for addictive disorder and share such an array of clinical features suggested a hypothesis that they—and, by extension, addictive disorders in general—share an underlying biopsychological process. This hypothesis has been supported by extensive research findings that demonstrated (a) significantly elevated lifetime comorbidity among these conditions and between them and other psychiatric disorders, (b) significant family history relationships, (c) similar responses to psychiatric medications, (d) similar results on psychometric assessments, and (e) similar neurobiological correlates. Comprehensive reviews have concluded that variation in the liability to a substance use disorder correlates with personality phenotypic variations that predate the initiation of substance use. Prospective studies have reported the temporal and predictive primacy of anxiety, mood, and attention deficit disorders when one of them is found to be comorbid with a substance use disorder. Subsequent studies found that deficits in affect- and self-regulatory functioning usually precede and increase the risk for development of substance use disorders, bulimia nervosa, gambling disorder, and sex addiction. That the documented abnormalities of psychological and neurobiological functioning are temporally primary to and predictive of the behavioral syndromes that have been under consideration here suggests that the underlying biopsychological process that these conditions share precedes their onset and is not simply a consequence of the behavior or lifestyle that characterizes them.

Definition

Addictive process is the term that designates the core biopsychological process that addictive disorders are hypothesized to share. It develops prior to the emergence of addictive patterns of behavior and constitutes a biopsychological predisposition to develop an addictive disorder. It can be defined descriptively as an enduring, inordinately strong tendency to

engage in some form of pleasure-producing behavior in a pattern that is characterized by impaired control and continuation despite harmful consequences. It is not what makes cocaine, ice cream, or sex pleasurable for people in general but what makes the drive for cocaine, ice cream, or sex so much more inexorable for those who use it addictively. The class of addictive disorders includes substance use disorders, gambling disorder, binge eating disorder, bulimia nervosa, kleptomania, sex addiction, and other enduring conditions in which a behavior that can function both to produce pleasure and to reduce painful affects is employed in a pattern that is characterized by two key features: (1) recurrent failure to control the behavior and (2) continuation of the behavior despite significant harmful consequences. The group of addictive disorders is most usefully considered as not a collection of distinct disorders but an underlying process that can be expressed in one or more of various behavioral manifestations.

Thus, the development of an addictive disorder is shaped by two sets of factors: those that contribute to the underlying addictive process and those that guide the selection of a particular behavior as the one preferred for addictive use. Whereas the type of behavior that is being used addictively is the most readily noticeable manifestation of an addictive disorder, the addictive process is the component of the disorder that leads to the behavior being used in ways that are pathological. In other words, the addictive process is the underlying disease and an instance of addictive behavior is a symptom of the disease. Addictive behavior is behavior that the person under consideration is using or is likely to use addictively: that is, with impaired control and despite harmful consequences.

That the addictive process is a biopsychological process does not mean that it represents an interaction between biological and psychological, nor that it has psychological properties that emerge out of relationships within its biological system. Biological and psychological are not distinct forms of reality but discrete sets of terms that represent distinct ways of experiencing and understanding a unitary human reality. The addictive process is a unitary biopsychological process that can be described and understood in both biological and psychological terms. Brief neurobiological and psychological formulations of the addictive process follow.

Neurobiological Formulation

From the neurobiological standpoint, the underlying vulnerability that all addictive disorders share can be understood to involve impairments in three interrelated functional systems: motivation-reward, affect regulation, and behavioral inhibition.

Impaired motivation-reward exposes addicts to unsatisfied states of restless anhedonia, irritable tension, and subjective emptiness. In the context of impaired motivation-reward function, behaviors that are

associated with activation of the reward system are more strongly reinforced (via both positive and negative reinforcement) than they otherwise would have been. Motivation-reward impairment may additionally distort the relative reward values among varieties of experience.

Impaired affect regulation renders addicts chronically vulnerable to painful affects, affective hyperresponsivity, and emotional instability. In the context of impaired affect regulation, behaviors that are associated with escape from or avoidance of painful affects are more strongly reinforced (via negative reinforcement) than they otherwise would have been. Combined impairments of motivation-reward and affect regulation particularly enhance the salience of drug self-administration and other behaviors that are associated with both (a) activation of the reward system and (b) escape from or avoidance of painful affects.

Impaired behavioral inhibition increases the likelihood that urges for some form of reinforcement in the short term (negative, positive, or both) will override consideration of longer term consequences, both negative and positive. The combination of impaired motivation-reward, impaired affect regulation, and impaired behavioral inhibition produces a condition in which urges to engage in behaviors that both provide pleasure or gratification and reduce painful affects are extraordinarily difficult to resist despite the harmful consequences that they might entail.

Impairments in motivation-reward, affect regulation, and behavioral inhibition result from interactions among genetic and environmental influences. They can be characterized as vulnerability factors, biological markers, or endophenotypic dimensions of addiction. The diathesis-stress model provides an etiological schema that guides the organization of research findings. Genetic predispositions interact with adverse experiences during critical phases of development to result in a phenotype that is neurobiologically vulnerable to the effects of stress later in life. This phenotype thus lowers an individual's threshold for developing an addictive disorder upon further stress exposure. Factors that have been found to contribute to the development of an addictive process include genetics, maternal gestational (prenatal) stress, caregiving deficiencies during infancy, and adverse childhood events or conditions.

Genetics. Researchers now generally believe that the predisposition to develop an addictive disorder results from the interaction of multiple genes. A large number of genes contribute to the risk for substance use disorder but no single gene displays such a large magnitude of effect that it accounts for a major fraction of the genetic influence. The same holds for gambling disorder. Meanwhile, most of the genetic liability to develop addictive use of any one psychoactive substance is shared among the substances, as distinct from liability that is specific to a particular

substance or class of substances. Along similar lines, the genetic risk for alcohol use disorder accounts for much of the risk for gambling disorder, and genes that increase the risk for developing gambling disorder increase the risk for developing other impulsive-compulsive and addictive patterns of behavior. Many addiction-related genetic variants (polymorphisms) do not fit into categories that are defined by substances or behaviors and are more usefully organized into categories that are defined by the brain functions that they alter. One such set of categories is motivation-reward, affect regulation, and behavioral inhibition.

MOTIVATION-REWARD. The homozygous *11* genotype of the dopamine (DA) D_1 receptor is associated with increased risk of alcohol use, cigarette smoking, use of other drugs, gambling, compulsive shopping, and compulsive eating. The Taq *A1* allele of the D_2 receptor gene predicts alcoholism, cigarette smoking, addictive use of psychoactive substances, pathological gambling, and exaggerated reward value of food. The Taq *A1* allele is associated with reduced D_2 receptor density, hypersensitive presynaptic D_2 receptors, and decreased general responsiveness of the reward system to rewarding stimuli, along with heightened responsiveness after events that increase levels of DA in the synapse between the ventral tegmental area and the ventral striatum. Such an increased reward effect against a baseline of diminished reward could promote the development of an addictive disorder by intensifying a person's motivation to repeat behaviors that increase intrasynaptic DA in the midbrain motivation-reward system. Behaviors that do so include self-administration of a hedonic substance, eating (especially sweets), gambling, and engaging in sexual behavior. The Val(158) allele of the COMT gene that codes for catechol-*O*-methyl transferase (an enzyme that inactivates DA and norepinephrine [NE]) is associated with alcoholism, methamphetamine abuse, heroin addiction, and polysubstance abuse.

Variants of the gene that codes for the κ-opioid ligand prodynorphin occur more frequently in alcoholics, cocaine abusers, and methamphetamine addicts than in controls. Polymorphisms of the gene for the cannabinoid CB_1 receptor are associated with increased risk for addiction to alcohol, cocaine, amphetamines, cannabis, and multiple substances.

AFFECT REGULATION. The gene that encodes the serotonin (5-HT) transporter protein has a variant site (*5-HTTLPR,* for 5-HT transporter-linked polymorphic region) that is characterized by 2 alleles, 1 short (*s*) and the other long (*l*). The *s* allele, and especially the homozygous *s/s* genotype, is associated with alcoholism, heroin addiction, cigarette smoking, pathological gambling, bulimia, and binge eating disorder. It is associated also with affective instability, anxiety-related personality traits, and

heightened sensitivity to mild stressors.

The *A* allele of the gene that encodes brain-derived neurotrophic factor (BDNF) is associated with alcoholism, methamphetamine abuse, heroin abuse, tobacco smoking, and all subtypes of eating disorder. It is specifically associated with binge eating behavior, a phenotypic trait that occurs in both bulimia and binge-eating disorder. The long allele of another BDNF gene variant, in the gene's promoter region, is associated with vulnerability to polysubstance use disorder.

Alcoholism is associated with three polymorphisms of the gene that codes for neuropeptide Y (NPY) and with two haplotypes of the galanin gene. The mediating link in both cases is thought to be impaired regulation of anxiety.

BEHAVIORAL INHIBITION. Whereas DA D_1, D_2, and D_3 receptors play central roles in the motivation-reward system, D_4 and D_5 receptors are involved primarily in behavioral inhibition and attentional processing. The long allele of a polymorphism in the D_4 receptor gene is associated with impulsive personality traits and is a risk factor for adolescent alcohol abuse, adolescent "hard drug" use, heroin abuse, cue-elicited heroin craving, greater severity of alcohol and opiate addiction, pathological gambling, binge eating, and cue-elicited craving for food.

The common 148 bp variant of the D_5 receptor gene is correlated with substance abuse and also with attention deficit hyperactivity disorder (ADHD) and novelty seeking. By itself, the 10-repeat allele of the DA transporter (DAT) gene *DAT1* is not associated with increased consumption of alcohol. However, the correlation between novelty seeking and increased alcohol consumption is significantly greater with this allele than with the normal genotype.

Studies have reported associations between variants of the gene *GABRA2*, which encodes the α-2 subunit of the $GABA_A$ receptor, and both alcoholism and polysubstance use disorder. Risk alleles of this gene are associated also with conduct disorder in children and with antisocial personality disorder in adults. Polymorphisms of the gene that encodes the α-3 subunit of the $GABA_A$ receptor are associated with alcoholism.

Although impaired behavioral inhibition or impulsivity is a key feature of addiction, research has linked variants of genes that code for $5\text{-}HT_{1B}$ and $5\text{-}HT_{2A}$ receptors with greater impulsivity among alcoholics and bulimics, respectively, but not with increased risk for any addictive disorder. In at least some instances, genotypic variations seem to map onto disorder-relevant intermediate phenotypes (i.e., trait-level variations) within a population that is affected by a particular disorder more closely than they map onto gross phenotypes, such as the disorders themselves.

Maternal Gestational (prenatal) Stress. Research with rodents and monkeys has indicated that maternal gestational stress (MGS) amplifies the subsequent reinforcing effects of ethanol, psychostimulants, and morphine on the offspring. It accelerates sensitization to amphetamine, increases motor-responsiveness to cocaine, and potentiates cocaine-primed reinstatement of cocaine-seeking behavior. In addition, it intensifies physiological-emotional responses to postnatal stress, retards the learning of both active and passive means of managing stress, and increases impulsivity. Prospective studies with human subjects have found MGS to be associated with high affective/behavioral reactivity, negative affect, difficulty being consoled, and disinhibitory traits in the offspring. The consensus among investigators is that exposure to MGS results in a general susceptibility to psychopathology rather than increased risk for any particular form of psychopathology.

Caregiving Deficiencies during Infancy. Considerable development of the human central nervous system occurs after birth, particularly during the first 2 years of life. Postnatal brain development is highly responsive to conditions in the environment. Epidemiological, clinical, and preclinical data indicate that early exposure to adverse environments engenders a vulnerability to distortion of the physiological stress response and subsequent impairments in affect regulation, motivation-reward, and behavioral inhibition. With respect to postnatal brain development in mammals, the most important constituent of the environment is the mother (or mother-substitute). Research has investigated the potential consequences in adulthood of two mother-related variables that may occur during infancy: maternal separation (MS) or deprivation and deficiencies in the quality of maternal care.

MATERNAL SEPARATION. Compared to normally reared rats, adult rats that were exposed to MS during infancy are characterized by increased ethanol and sucrose consumption; amplified sensitivity to the neurochemical and behavioral effects of psychostimulants and opiates; increased susceptibility to the rewarding properties of psychostimulants, opiates, and food; enhanced acquisition and maintenance of operant responding for drug or food reinforcer; elevated risk of developing addiction-like patterns of consuming or self-administering ethanol, cocaine, morphine, and sucrose solution; and greater tendency to overeat and gain weight when given free access to food. On a fasting/refeeding schedule, compensatory hyperphagia during refeeding days diminishes gradually in rats that were raised normally but continues in MS rats. A prospective study with human subjects found that compared to offspring who had been breast-fed for more than 1 month, offspring who had been weaned at age 1 month or earlier had a 1.47-fold greater risk of being hospitalized in a psychiatric unit

with an ethanol-related diagnosis.

MS during infancy is associated also with other behavioral abnormalities. Most common are behaviors that suggest impaired affect regulation: increased anxiety-like behaviors in response to aversive environmental stimuli, intensified behavioral and visceral reactivity to stress, increased startle response, increased fearfulness in the face of novelty, increased depression-like behavior, decreased social interaction, reduced coping behavior in stressful environments, and increased vulnerability to develop anxiety-like or depression-like behavior when subsequently exposed to chronic stress. The amount of ethanol consumed by MS animals is highly correlated with the magnitude of the glucocorticoid (GC) response to a mild stressor and with measures of anxiety-like and fearful behaviors in both rats and monkeys.

Behaviors that are consistent with impaired motivation-reward function have been reported in MS research with rats and with monkeys. These include reduced social motivation, blunted perception of differences in reward value, and reduced operant responses on progressive ratio reinforcement schedules, which suggest either low motivation or diminished sensitivity to reward. MS rats' low motivation and diminished sensitivity to reward have been normalized by administration of fluoxetine.

Other studies of MS animals found increased impulsive behavior, locomotor hyperactivity, reduced prepulse inhibition of the startle response, deficits in shifting the attentional set, lower-risk assessment and higher-risk taking, and impaired retention of inhibitory avoidance. Children who were adopted from deprived institutional settings before age 43 months were found to display high degrees of inattention and overactivity at 6 and 11 years of age. They were found also to demonstrate deficits in executive functions of planning, inhibition, set-shifting, working memory, generativity, and action monitoring, which were most severe in children who had experienced more than 6 months of institutional maternal deprivation.

MATERNAL CARE DEFICIENCIES. A group of studies has explored the relationship between naturally occurring variations in maternal care of infant rats and the latter's capacity for affect regulation as they develop to adulthood. Maternal care is represented by mother rats' licking or grooming and arched-back nursing (LG-ABN) of their newborn progeny—simply licking and grooming (LG) in some recent studies. These studies have provided a credible experimental model for the relationship between the maternal care of human infants and their ensuing development.

Differences in the frequency of these maternal care behaviors have been found to be predictive of individual differences in affect regulation in the offspring when they become adults. Compared to the offspring of mother rats that show high levels of LG-ABN, the offspring of low-LG-ABN dams

exhibit heightened anxiety-like behavior, exaggerated responses to stress, and less time spent in social interaction. Female offspring of low–LG-ABN mothers exhibit earlier puberty, increased sexual receptivity, and greater vulnerability to stress-induced binge-eating than do their high–LG-ABN counterparts.

To test the hypothesis that the observed characteristics of low–LG-ABN mothers' offspring had resulted from lower levels of maternal care, as distinct from genetic factors that also produce low–LG-ABN behavior, rat pups were "cross-fostered"—raised from birth by mothers whose LG-ABN frequency was opposite from that of the mothers to whom they were born. Pups that were born to high–LG-ABN dams and raised by low–LG-ABN mothers become hyperemotional, high-strung adult rats. Meanwhile, pups that were born to low–LG-ABN dams and raised from birth by high–LG-ABN mothers become mellow, well-adjusted adult rats. These findings support the hypothesis that the lower levels of maternal care that these rats received during their early postnatal lives was a critical determinant of the observed differences in adult rat functioning.

Variations in maternal licking/grooming influence the development of stress-responsiveness primarily through central processes that stimulate—via amygdalar corticotropin-releasing hormone (CRH) —or inhibit—via gamma-aminobutyric acid (GABA)—the noradrenergic systems that activate endocrine and behavioral responses to stress. Relative to offspring of high–LG-ABN mothers, adult offspring of low–LG-ABN mothers exhibit increased CRH expression in the central amygdala (CeA), increased CRH receptor density in the locus ceruleus (LC, primary site of NE synthesis), decreased density of hippocampal GC receptors (and hence diminished GC feedback sensitivity), decreased expression of glutamic acid decarboxylase (the rate-limiting enzyme in GABA synthesis) in the medial prefrontal cortex (PFC), reduced GABA affinity of benzodiazepine (BZD)-$GABA_A$ receptor complexes in the LC and nucleus tractus solitarii, and decreased BZD-$GABA_A$ receptor complex densities in the basolateral amygdala, CeA, and LC. Consequently, the offspring of low-LG-ABN mothers display higher stress–induced increases in paraventricular nucleus levels of NE than do the offspring of high–LG-ABN mothers. Noradrenergic input to the paraventricular nucleus provides the primary source of activation for synthesis and release of the CRH that initiates that hypothalamic–pituitary–adrenal (HPA) axis cascade. Compared to adult offspring of high–LG-ABN mothers, adult offspring of low–LG-ABN mothers also show decreased mGluR1 mRNA and protein in the hippocampus, decreased type I metabotropic glutamate receptor (mGluR1)-induced hippocampal long-term depression, and increased paired-pulse depression (PPD). PPD is an inhibitory feedback process that prevents excessive glutamate release during high-frequency stimulation.

At the heart of these differences in neurotransmitter/neuromodulator

and receptor functioning are differences in gene expression. Maternal care of infants regulates the expression of the hippocampal glucocorticoid receptor (GR) gene by influencing two processes: acetylation of histones H3-K9 and methylation of the NGFI-A consensus sequence on the exon 1(7) promoter. Adult offspring of low–LG-ABN mothers were found to have lower levels of DNA methylation than do adult offspring of high–LG–ABN mothers. These differences emerged over the first week of life, were reversed with cross-fostering, persisted into adulthood, and were associated with altered histone acetylation and transcription factor (NGFI-A) binding to the GR promoter.

Central infusion of either a histone deacetylase inhibitor or a methyl donor removed the group differences in NGFI-A binding, GR expression, and HPA and behavioral responses to stress. These findings suggest a causal relation among epigenomic state (the sum of all epigenetic changes [DNA methylations and histone modifications] to an organism's genome), GR expression, and the effect of maternal infant care on subsequent stress responses in the offspring. Moreover, they demonstrate that an epigenomic state of a gene can be established through behavioral programming and is potentially reversible.

The demonstration by this body of studies that naturally occurring differences in maternal care of infants are correlated with subsequent differences in the behavior and neurobiology of adults constitutes an important evidence-based bridge between neuroscience and psychodynamic/psychoanalytic understanding.

Adverse Childhood Events or Conditions. Children who have been sexually or physically abused manifest abnormal baseline and stressor-responsive cortisol levels, either abnormally elevated or abnormally flat. Childhood abuse may sensitize the stress response system, thus rendering persons who were abused during childhood particularly vulnerable to stress and increasing their risk for stress-related disorders. This vulnerability may result in hypersecretion of CRH when these individuals are stressed.

Upon exposure to a standardized (laboratory) psychosocial stressor, women with a history of childhood sexual or physical abuse demonstrated markedly increased pituitary–adrenal and autonomic responses. The mean adrenocorticotropic hormone response was highest in women with a history of childhood abuse and current major depression, who also showed significantly increased cortisol responses to psychosocial stress. These findings suggest that HPA axis hyperreactivity, probably due to hypersecretion of CRH, is a persistent consequence of childhood abuse that may contribute to a diathesis for later development of addiction and other psychiatric disorders.

A study screened 120 healthy college students for quality of parental care during childhood. Five from the top end of the distribution and five

from the bottom end were invited to participate in a PET study during which they were asked to complete a stressful psychosocial task. The scans indicated that the psychosocial stressor elicited a significant release of DA in the ventral striatum of subjects in the low but not the high parental care group. The low parental care group also showed higher baseline cortisol levels and higher increases of cortisol during the stressful task. The magnitude of the cortisol response to stress was highly correlated with the magnitude of the ventral striatum DA response. These findings suggest that the chronic stress of neglectful or abusive parental care during childhood results in an HPA axis that is hypersensitive to psychosocial stressors and a midbrain that is hyperresponsive to triggers of DA release—conditions that are highly conducive to the development of addictive disorders.

In preclinical studies, reinstatement of drug-seeking behavior after extinction of its association with drug delivery can be elicited by brief exposure to a stressor. In human subjects, exposure to stressors and presentation of stress-related imagery have been identified as potent triggers for drug-seeking and recurrence of drug use. Most studies on drug self-administration have reported increases in responding after repeated or prolonged exposure to stress levels of GCs. GCs are understood to enhance drug responding by selectively facilitating DA transmission in the nucleus accumbens shell. During chronic stress, repeated increases in GCs and DA result in sensitization of the reward system. This sensitized state, which can persist after the end of the stress response, renders subjects more responsive to pleasurable substances or behaviors that are associated with release of mesolimbic DA and consequently more vulnerable to the development of addiction.

Thus, addicts' vulnerability to stress-induced reinstatement of drug-seeking and thence to drug use depends on the sensitivity of their HPA axis and the responsiveness of their mesolimbic dopaminergic system, in addition to the frequency and intensity of stressor exposures. Both elevated baseline cortisol levels and higher increases of cortisol in response to stressors are common sequelae of adverse conditions or recurrent adverse events during childhood.

Stress and disorders that are related to chronic stress tend to increase dendritic atrophy, accelerate neuronal degeneration, and subvert neuronal regeneration in the hippocampus and PFC. Mediating factors may include chronically elevated levels of GC, decreased expression of BDNF, stunted sprouting of serotonergic axons from insufficient availability of BDNF, glial cell loss, and decreased arborization and density of noradrenergic axons. In the developing brain, elevated levels of catecholamines and cortisol may lead to structural deviations or deficits through accelerated loss of neurons, delays in myelination, abnormalities in developmentally appropriate synaptic pruning, and inhibition of neurogenesis. Sustained GC exposure truncates dendrites and impairs neurogenesis in the hippocampus and

PFC. Reduced baseline and hyporesponsive cortisol levels also can cause neuronal damage.

Together, these processes feed the downward spiral by potentiating hippocampal and cortical atrophy. Neuronal degeneration in the hippocampus and PFC diminishes the capacity of these regions to modulate or inhibit amygdalar stress/fear pathways and undermines hippocampal inhibitory control over cortisol release by the HPA system. Evidence of neuronal loss in the anterior cingulate cortex (ACC) has been reported in children, adolescents, and nonhuman primates with histories of adverse experiences early in life. In addition, adult survivors of early abuse have been found to have changes in hippocampal structure and function. Thus, these stress-induced processes may lead to compromised executive function, impaired affect regulation, and a greater incidence of impulsive behaviors.

Psychological Formulation

The central component of the addictive process as psychologically formulated is impaired self-regulation. Associated character pathology emerges from the same genetic-environmental matrix in which the self-regulatory impairments originate and develops in response to the interactive effects of impaired self-regulation and adverse personal history.

Impaired Self-Regulation. The descriptive definition of the addictive process that was presented earlier identifies objectively confirmable features that observers of the person in question might similarly apprehend. The addictive process also has a subjective definition; that is, what it feels like from the inside. Defined subjectively, the addictive process is an enduring tendency to feel driven to use some form of pleasure-producing behavior as a consciously or unconsciously intended means of relieving or avoiding painful affects and regulating one's sense of self. A driven need to relieve painful affects and to regulate one's sense of self suggests that addicts' inherent systems for regulating affects and sense of self are impaired. Dependence on externally directed behavior to relieve painful affects and regulate one's sense of self suggests a relative incapacity to utilize other, more adaptive means of managing. The subjective definition of the addictive process suggests that impaired self-regulation is at the nucleus of the addictive process.

Self-regulation is a comprehensive term for the biopsychological processes that regulate one's subjective states and behavior. The subjective states that are most affected when self-regulation is impaired are affects and sense of self. Volition, sensory experience, and thought processes are often also affected but less dramatically. Impaired internal regulation of people's subjective states leads them to depend on externally directed actions to regulate their affects and sense of self. Under such

circumstances, intense and potentially traumatic affects may be experienced not as emotions but as urges to act. Impaired behavioral regulation limits people's capacity to resist urges to engage in actions that modify their affects or sense of self and that seem to promise, however illusorily, to protect them from emotional trauma.

The self-regulation system is a set of processes that are interrelated both functionally and developmentally. It can be understood to consist of three groups of functions: affect regulation, self-care, and self-governance.

Affect regulation functions are the psychobiological processes that regulate the subjective, physiological, behavioral, and interpersonal aspects of one's emotional states. These functions include self-soothing, self-enlivening, keeping affects' intensity within a manageable range, maintaining a balance when affects conflict, and tolerating affects without being overwhelmed.

Self-care functions concern the organization and regulation of one's perceptual, cognitive, executive, and behavioral processes in the service of self-protection and self-nurturing. Self-protection involves recognizing, assessing, warning oneself of, and defending oneself against dangers. Self-nurturing involves recognizing, assessing, setting priorities among, and taking care of one's needs. A critical component of self-care is the ability to use affects as signals or guides for appropriate action.

Self-governance functions include values, standards, self-esteem, self-punishment, and maintenance of a stable and cohesive sense of self. Self-governance thus subsumes (1) what psychoanalytic theory identifies as the superego—in its encouraging, approving, and rewarding aspects as well as its prohibiting, condemning, and punishing aspects; (2) the ego ideal or self-ideal, a fantasied representation of oneself that embodies all one's ideals; and (3) the processes that are involved in regulating one's sense of self.

The core self-regulation system develops during the first years of life out of the reciprocal interactions between the expression of infants' genes and infants' relationships with their primary caregivers. Considerable development of the human brain occurs after birth, particularly during the first 2 years of life, and its postnatal development is highly responsive to conditions in the caregiving environment. In healthy development, "good enough" genetic endowment interacts with caregivers' more or less adequate provision of functions that regulate the infants' somato-affective inner world. As their physiologies and prepsychological anlagen are challenged but not overwhelmed, infants gradually become able to provide for themselves the regulatory functions for the provision of which they had been depending on their caregivers. The metaphorical shorthand for this developmental process is *internalization*. However, various combinations of genetic and environmental factors can interfere with internalization and disrupt the development of the self-regulation system. Later in life,

adaptive self-regulatory functioning, particularly in the areas of self-care and self-governance, depends on effective management of one's inner conflicts—which is predicated on awareness of them. When awareness of inner conflicts is blocked, their operation in the dark can distort the functioning of the self-regulation system—as can the processes that block such awareness.

To the extent that individuals' affect regulation is impaired, their emotional responses are easily provoked, unusually intense, and slow to recede. Their affect states tend to be unstable, intense, and disorganizing. Especially threatening are abandonment-separation anxiety, claustrophobic anxiety (anxiety associated with the general idea of being closed in, trapped, or smothered), depressive affects, the "superego affects" of guilt and shame, and affects that are associated with narcissistic injury. But almost any intense affect, including excitement or love, can be experienced as threatening when affect regulation is impaired.

To the extent that individuals' self-care is impaired, they are deficient in their abilities to avoid danger, to protect themselves, and to take care of their needs. The dangers to which they expose themselves include physical, emotional, and narcissistic dangers. ("Narcissistic" means related to self-esteem and sense of self.) Similarly, the needs to which they do not adequately attend include physical, emotional, and narcissistic needs. As a result, they are particularly susceptible to getting hurt and to being chronically dissatisfied on all three levels.

To the extent that individuals' self-governance is impaired they lack a consistent sense of self, a stable self-esteem, and a clear sense of direction. They may have skewed or fragmented values, ideals, and sense of meaning. Any threat to their sense of being in control can potentially trigger intense anxiety about being overwhelmed by helplessness. Their behavior tends to oscillate between extremes and they may swing from actions that violate their values to actions that are self-injurious or invite punishment from the environment.

When individuals' self-regulation is impaired, they live in a world of primitive pain, terrifying anxieties, and an unstable sense of self. They tend to respond with overwhelmingly intense affects to events that most others seem to take in stride. They experience the risk of being thus affectively overwhelmed and losing their self-coherence as threats to their survival, so they do what they can to avoid affective states that are potentially traumatic. Many seek desperately for something that can perform for them the self-regulatory functions that people with intact self-regulation systems are able to do autonomously. Often, they develop fantasies (usually unconscious) that the solution for their self-regulatory problems lies in something outside their selves: a substance, a person, a nonhuman object, a bodily state, or an institution. Many of them learn to ward off traumatic affects and self-states by engaging in a rewarding activity such as eating,

taking a mood-altering substance into their bodies, gambling, or engaging in some form of sexual behavior. They then tend to become dependent on this activity to provide sham versions of self-regulatory functions that they have not been sufficiently able to provide for themselves.

When external or internal cues that are associated with intense affects or with loss of self-coherence become conditioned to particular craving sensations or discriminative stimuli for particular addictive-behavioral responses, these cues can trigger or be experienced as urges to engage in symptomatic behavior. Since these behaviors are addicts' desperate attempts to ward off traumatic affects and self-states that threaten their survival, their urges to enact the behavior typically are so imperative that their patterns of engaging in it are likely to be characterized by impaired control and by relative imperviousness to the restraining influence of harmful consequences.

Associated Character Pathology. *Character pathology* is an inclusive term that refers to both (a) enduring, pervasive, relatively inflexible patterns of inner experience and behavior that begin in adolescence or early adulthood and lead to significant functional impairment or subjective distress, and (b) the underlying psychobiological processes of which the inner experience and behavior are observable manifestations (inner experience being observable to the subject). Its inclusion of (b) distinguishes it from the more widely used term *personality disorder*.

The character pathology that tends to be associated with addictive disorders originates under the same genetic and environmental circumstances as do the just discussed self-regulatory impairments. Its development is then shaped by infants' and children's needs to protect themselves from the overwhelming affects and loss of sense of self to which their personal history and impaired self-regulation expose them. Accompanying character pathology that tends to be associated with addiction (each to varying degrees in different individuals) includes: characteristic self-protective processes, impaired ability to symbolize affects, impaired management of the basic conflict, and reliance on the primal fantasy.

CHARACTERISTIC SELF-PROTECTIVE PROCESSES. The self-protective processes (often referred to as defenses) that people who engage addictively in one or more forms of behavior most typically employ are denial or disavowal, externalization, and the processes that constitute the narcissistic character/personality system.

Denial or disavowal is an unconsciously motivated failure either to recognize the reality of what is perceived or to appreciate the personal significance or implications of what is perceived. Its function is to support dissociation, an unconsciously intended removal from the realm of

awareness something that if consciously experienced could precipitate overwhelming emotional pain or self-fragmentation. In addictive disorders, denial or disavowal most often occurs to support dissociation of: (1) painful affects, particularly guilt, shame, and affects that evoke guilt or shame; (2) one side or the other of an inner conflict, particularly the basic conflict and its derivatives; (3) narcissistic vulnerability and limitations; and (4) impulses that have been associated with severe trauma.

Externalization is an unconscious process whereby an aspect of one's self is experienced as an aspect not of oneself but of something external to oneself. It is the self-protective process that is most characteristic of addiction. The basis for people's transition from simply engaging in a rewarding activity to engaging in that activity addictively is their externalization of their self-regulatory functions to that rewarding activity or to the object of that activity. Their mental representations of their capacity to regulate their affects and sense of self shifts from their mental representations of their selves to their mental representations of that activity or object. In the course of this shift, their mental representations of their self-regulatory functioning are infiltrated by fantasies of omnipotence and become more effective in people's mental representations of the rewarding activities or objects than they ever could have been in real life.

The scope of externalization in addicts' lives extends beyond their self-regulatory functions. For example, to support the denial of an inner conflict, addicts (or anyone else) can assign one side of the conflict to an external agent and thereby transform the inner conflict into an external conflict. Or, in an attempt to take magical omnipotent control over their inner lives, they may externalize a passively experienced, traumatic inner state to a tangible external problem that they believe can be resolved through effort and action. A consistent feature of externalization is its direction of attention and energy toward action in the external world.

The DSM-5 diagnostic criteria for narcissistic personality disorder (lightly abridged) are: (1) grandiose sense of self-importance; (2) preoccupation with fantasies of unlimited success, power, brilliance, beauty, or ideal love; (3) belief that one is unique and special; (4) abnormal need for admiration; (5) sense of entitlement; (6) interpersonal exploitativeness; (7) lack of empathy; (8) envy of others or belief that one is envied by others; and (9) arrogant behaviors or attitudes. The use in common parlance of *narcissistic* and *narcissist* as terms of disparagement reflects the magnitude of people's antipathy toward narcissistic behaviors and attitudes (or perhaps of their motivation to disavow their own narcissistic problems). It also reflects people's failure to appreciate that these behaviors and attitudes arise as desperate measures to avoid annihilation and persist because they are unconsciously believed to be required as protection against that same danger.

The narcissistic character system is a variable set of self-protective

processes that function to preserve one's sense of self from annihilation or fragmentation and to sustain an artificial self-esteem that supports the dissociation of overwhelming shame, humiliation, vulnerability, sense of inadequacy, and helplessness. Grandiosity is the most overt aspect of the artificial self-esteem. It supports dissociation of the aforementioned affective experiences by trumpeting their opposite. Deficient empathy (ability to sense the emotions that another person is consciously or unconsciously experiencing) can result from genetic factors or from isolation in early infancy. However, it most commonly originates early in life as a self-protective deadening of one's empathic sense under circumstances in which experiencing a primary caregiver's emotions entails a threat of annihilation (e.g., by becoming affectively overwhelmed, or by losing a sense of self-other boundary and subjectively ceasing to exist as a distinct self). One's capacity for empathy can be further undermined by impaired ability to recognize or name affects.

A sense of entitlement can originate in grandiose fantasies. More often, it is based in reality. During the first years of life, most if not all persons who later develop narcissistic character pathology did not receive enough of the caring, nurturing, or soothing that they needed for healthy development—and to which they as mammals were biopsychologically entitled by virtue of having been born. In the context of that history, lingering feelings of need, emptiness, deprivation, and entitlement are understandable. Grandiosity may then seep into and swell the entitlement to drive away feelings of need, vulnerability, and helplessness. In general, the problem with entitlement is not entitlement as such but people's unconscious fantasies that they will receive restitution from people in their current lives for what they were entitled to but did not get in their early childhoods. The main function of this fantasy, like that of the primal fantasy from which it is an offshoot, is to support the dissociation of traumatic early experiences of privation or abandonment.

Individuals whose functioning is structured by the narcissistic character system are unconsciously terrified by the ongoing threats to their psychological survival. They desperately seek and try to control potential sources of narcissistic supplies—whatever might help preserve their beleaguered sense of self or sustain their prosthetic self-esteem. Since they unconsciously experience their survival to be at stake, they tend to relate to other persons primarily in terms of the latter's capacity to provide or interfere with access to such supplies. Addictive behavior patterns can be understood as a natural expression of the characteristic modes of adaptation that individuals with narcissistic character pathology employ to deal with their underlying vulnerabilities and impairments.

In addictive disorders, denial, externalization, and narcissistic self-protective processes support and promote one another. In effect, they operate together as one system of character or personality.

IMPAIRED ABILITY TO SYMBOLIZE AFFECTS. People who struggle with addictive disorders share with psychosomatic disorder patients and posttraumatic stress disorder (PTSD) patients an impaired ability to symbolize affects. They have difficulty recognizing, naming, and verbalizing their affects, which they typically experience as physical bodily states rather than meaningful emotions. These disturbances often are accompanied by a lowered capacity for empathy and diminished emotional involvement in interpersonal relationships. The ability to recognize and identify emotions is a critical component of affect tolerance, which accordingly is diminished when the ability to symbolize affects is impaired. Affects that arise under these circumstances consequently are more often experienced as potentially overwhelming, as incipient threats to the stability of the self. The experience of affects as physical bodily states moreover disposes people to address them by physical means, by action—and particularly by action that directly affects the body.

Episodes of addictive behavior are usually precipitated by painful affects that addicts at some level perceive as threatening to overwhelm them. However, they often are consciously aware of neither the sense of being threatened nor even the affects themselves. Preceding an episode of addictive behavior, they typically experience a nonspecific sense of dysphoria or discomfort—a vague restlessness, an uneasiness, an unpleasant sense of tension. Some addicts do not experience the emotional discomfort at all but are aware only of an urge or a craving to engage in symptomatic behavior.

An impaired ability to symbolize affects can result from one or more of the following factors: (1) a primary developmental deficit; (2) regression to a more primitive kind of symbolic and/or affective functioning (regression is reversion to a developmentally earlier mode of being that usually occurs as an unconsciously intended means of self-protection from traumatic emotional pain); (3) a broad-spectrum self-protective process that is directed against affects or symbolic processes in general; and (4) a "walling off" of affect regulation functions, unconsciously reserving them for one's caregiver or, more accurately, for the caregiver as represented in one's mind.

THE BASIC CONFLICT. The basic conflict is between two fundamental sets of drives and fears that emerge during the first 2 years of life. On one side is a drive for attachment, connection, or merger with the caregiver, which is combined with fear of being abandoned by or isolated from the caregiver. On the other is a drive for separation, autonomy, and mastery, which is combined with fear of being engulfed, swallowed up, or controlled by the caregiver. In short, the basic conflict can be understood as being between a basic pull toward attachment-union with another person and a basic push toward being a separate, autonomous self.

The importance of the basic conflict derives first of all from its ubiquity. It is not a form of psychopathology but a part of the human condition that pulsates in the background throughout people's lives. How it manifests depends on the life context and on where a person is developmentally. Examples of forms it can take include conflicts between wishes to be taken care of and fears of being controlled, and between wishes to be competent and fears of being abandoned. In addition to being ubiquitous, the basic conflict is perpetual and unavoidably painful. Both connections with other persons and existence as a separate self are necessary for an individual's psychological survival, and a move toward one side can evoke painful feelings of threat or loss related to the other. When the basic conflict is not effectively managed, it intensifies any inner conflict that can be mapped onto it. (Management of a conflict does not mean that it has been solved and settled but that a relatively stable dynamic equilibrium is being maintained.)

A number of factors can make the basic conflict difficult to manage. Most often, people have difficulty resolving the basic conflict when they unconsciously believe that it is associated with more anxiety than they can tolerate. Perhaps their primary caregivers gave them messages that their connections with their caregivers were contingent on their not being separate persons and that being separate would result in being abandoned. Or their history with their primary caregivers led them to expect that connecting with another person would be too likely to lead to being devastatingly hurt, being humiliated, or losing their identity. Or they unconsciously refuse to become separate and masterful because to do so would imply recognition that their unmet early childhood needs would never be met, which would expose them to traumatic depressive pain. One or more of these patterns characterize the early history of most addicts. A more general factor is impaired self-regulation, which can intensify anxiety or emotional pain to the point of feeling as though it were threatening to overwhelm and destroy the subject. The final common pathway is that affected persons unconsciously believe that awareness of the basic conflict would be accompanied by overwhelming, shattering anxiety. Consequently they avoid that awareness by keeping one side of the conflict out of conscious view. This process of dissociation or splitting is often sustained by unconsciously assigning (i.e., externalizing) the dissociated side of the conflict to another person. The other person is then experienced as though he or she were pushing for that side of the conflict.

For example, a fictional patient Blake unconsciously wants his psychotherapist to take care of him totally but also fears being taken over and losing his sense of self. Assigning the former side of the conflict to the therapist could lead Blake to experience his therapist as trying to seduce or trick him into becoming dependent while Blake experiences himself as trying to hold onto his autonomy. Conversely, assigning the latter side to the therapist could lead Blake to experience himself as trying to develop a closer relationship with his therapist while the

therapist is holding back or pushing him away.

THE PRIMAL FANTASY. The primal fantasy is an unconscious fantasy that can develop during the first 2 years of life when some combination of deficient caregiver responsiveness and a child's genetically or epigenetically based impairments interfere with fulfillment of the child's needs to be cared for, held, nurtured, and soothed. The fantasy is that those fundamental needs that were not met during early childhood will someday be met. While it is defined by anticipated gratification, its power derives primarily from its capacity to sustain the dissociation of overwhelmingly painful early experiences of privation or abandonment and thus to provide protection from traumatic pain. The belief that unmet needs of early childhood might yet be met enables avoidance of the realization that the opportunity to have those needs met is, along with early childhood, lost and gone forever.

A common form of this fantasy begins with children's implicit belief that the reason their needs were not met was that they were bad or not good enough and thus did not deserve to have their needs met. Their minds gravitate toward this explanation in preference to others that locate the cause of their privation in the primary caregiver or an unpredictable universe. Believing that they themselves are to blame for their needs not being met is painful, but believing that they are helpless and can do nothing to influence this condition is even more so. Self-blame sustains illusions of having control. The fantasy then takes over with the belief that if they were to be good enough or do well enough they would get their needs met. In addition to fostering a sense of efficacy, this belief inoculates the primal fantasy against empirical test. Individuals' failure to get their needs met after being good or doing well does not disprove the fantasy. It just indicates that they were not good enough or did not do well enough.

When people rely on the primal fantasy to protect themselves from the pain of early privation or abandonment, they have a hard time separating from their parents and becoming authentic, self-responsible persons. The fantasy is about meeting the needs of early childhood, so the subject of the fantasy must in significant ways remain in early childhood. As long as people do not grow up, they can maintain the fantasy that the unmet needs of their early childhood still can be met. Growing up, becoming separate and authentic persons, means letting go of the fantasy. The consequent annihilation anxiety leads some individuals to avoid separating from their parents, taking care of themselves, or assuming adult responsibilities.

People typically externalize the primal fantasy onto another person. They at some level believe that important needs that were not met during their early childhood will be met by this person. During their subsequent childhood this fantasy can increase their vulnerability to being sexually abused. Later in life it can lead them repeatedly into dysfunctional

relationships. Sexualization of the fantasy can turn them toward addictive involvement in romantic relationships or other kinds of sexual behavior. The psychodynamics of addictive disorders in general usually involve the primal fantasy. In such cases, people unconsciously perceive the drug or food or other object of the addictive behavior as the fantasy object that will relieve their pain, fill their emptiness, make them whole, and meet in distorted form the unmet needs of their early childhoods.

Development of the Addictive Process. The development of the addictive process in psychological terms can be condensed to: (1) due to some combination of genetic and environmental factors, individuals' self-regulatory functions are impaired; (2) impaired self-regulation leaves them abnormally vulnerable to being overwhelmed by intense emotional states and loss of self-coherence; (3) in attempts to protect themselves from the danger of being thus overwhelmed, they externalize fantasy-tinged mental representations of their self-regulatory functions onto nonself activities or objects over which they believe that they have control. When the behavior that results occurs in a pattern that is characterized by recurrent loss of control and continuation of the behavior despite significant harmful consequences, they can be characterized as "addicts," whatever the behaviors in which they addictively engage.

SELECTION OF SEXUAL BEHAVIOR

The addictive process represents a predisposition to engage in one or more behaviors that can provide both pleasure and relief in a pattern that is characterized by recurrent failure to control and continuation despite harmful consequences. Given a sufficient degree of predisposition, what determines the specific behavior(s) in which a person addictively engages? More particularly, what might lead a person to choose (unconsciously) a form of sexual behavior rather than, for example, using cocaine or eating?

The development of a particular addictive disorder seems to have three components: (1) by virtue of the addictive process, individuals are predisposed to depend on some kind of pleasure-producing behavior to regulate their (dysregulated) affects and sense of self; (2) they take a mood-altering substance or engage in some other rewarding behavior; and (3) the effects of the substance-taking or other behavior—positively reinforcing, negatively reinforcing, and aversive—and the individuals' expectations about its effects in the future combine to determine the likelihood that they will select to take that substance or engage in that behavior as their addictive behavior of choice. These effects and expectations are determined by both genetic and experiential factors.

The foregoing observation suggests that sex addiction develops in addictively predisposed individuals who experience some form of sexual stimulation and learn to expect that sexual behavior will not only provide

pleasure but also relieve the affects that they experience as most dangerous, through means that are congruent with their characteristic modes of adaptation.

Experiences that sexual stimulation provides both pleasure and relief begin early in life. Newborn infants' first and most fundamental experience of pleasure and soothing is not suckling a breast or ingesting milk. It is skin-to-skin contact. This bodily contact supplies warmth, which relieves the first and (usually) most extreme distress of early extrauterine life. It also soothes somatic-affective discomfort and thereby provides a foundation for the subsequent development of affect-regulation. Skin precedes the mouth as the first erogenous zone. As infants develop, their responses to skin-to-skin contact begin to differentiate and become specific to particular areas of contact. Physiological and behavioral changes that indicate intense pleasure in response to bodily contact in the genital region becomes increasingly distinct from the pleasure responses to skin-to-skin contact with other areas of the body.

Infant observation studies have reported that at some time between 15 and 19 months of age, both male and female infants normally demonstrate an upsurge in genital self-manipulation. In time this takes on a more recognizably masturbatory pattern. Infants actively produce warm, pleasurable feelings in their own bodies which earlier they only passively experienced through nurturing maternal care. The genital self-stimulation seems from the outset to be mother-directed. During masturbation children often engage their mothers visually and smile at them. Later they use mother-connected objects, such as nursing bottles or transitional objects, for masturbatory activities. Children's early sexual interest and activity not only provide pleasure and tension release but moreover (in healthy development) facilitate the consolidation of self and object representations.

From the beginning of life, skin-to-skin sensations and then sexual sensations are intimately associated with survival, mother, and soothing, the last of which is the foundation of affect-regulation. Genital self-manipulation in the second year of life is mother-oriented and influences the development of not only mental representations of emotionally significant others but also sense of self and self-regulation. Significantly, infants are able to engage in sexual self-stimulation and provide the associated pleasurable and soothing feelings for themselves, independent of others' availability and willingness to soothe them. Not until they are older are they able to obtain and consume food on their own. A subjective drive to provide regulation of one's inner states by external means that seem to be within one's own control is at the core of addiction. Freud characterized masturbation as "the primal addiction" from which all other addictive disorders derive.

Early experiences of sexual stimulation that provide both pleasure and

relief are practically universal among neurologically intact infants. Yet not all individuals who are addictively predisposed become sex addicts. The eventual development of sex addiction in addictively predisposed individuals typically is associated with either seduction (overt or covert) or maternal sexualization. Sexualization is the infiltration of unconscious sexual fantasy into ordinarily nonsexual behaviors that hence can be used as vehicles for fantasied sexual gratification, which functions not so much to provide pleasure as to protect oneself from potentially overwhelming affect or self-fragmentation.

(In this paragraph and the next, "mother" and "maternal" are intended to refer to the mothering person, whether she be the child's mother or a nonmaternal primary caregiver.) When a woman who unconsciously relies on sexualization for psychological survival becomes a mother, sexualization is a constituent of the matrix in which the mother–child relationship begins. The influence of maternal sexualization on that relationship and hence on the child's early development is typically greater when the child is the same sex as the woman's fantasized sexual object. During infancy the emotionally and psychologically deprived child experiences the gratifying aspects of mother's sexualization and overt or covert seduction as though they were the vitally needed but missing nurturance, holding, and soothing. The child comes to identify with mother's sexualization and seductiveness as a way of preserving the tie to her. Sexual seductiveness may eventually become the child's predominant mode for relating to others and for expressing his or her desperate hunger for a fulfilling connection with another person.

In short, if the etiology of sex addiction were to be summed up in one brief equation, it would be: sex addiction = the addictive process + maternal sexualization.

TREATMENT

A core element of the medical way of thinking is the distinction between a disease and its symptoms and signs—observable indications of a disease that often result from the body's response to it. According to the understanding of addiction that is presented here, the addictive process is the underlying disease and instances of addictive behavior are symptoms and signs of the disease. Among the corollaries of the disease-symptom distinction is that treating the symptoms—in this case, instances of addictive behavior—without treating the underlying disease—the addictive process—is likely to be of limited effectiveness. Meanwhile, a distinctive feature of addictive disorders is the frequency with which external and internal consequences of the symptoms exacerbate the disease or interfere with its treatment. An implication of these conditions is that treatment for an addictive disorder is most likely to be effective when it addresses both the symptomatic behavior and the underlying addictive process. Treatment

modalities that are directed primarily toward addressing symptomatic behavior are discussed below under the heading, Symptomatic Behavior Management. Those that are directed primarily toward addressing the addictive process are discussed subsequently under the heading, Healing of the Addictive Process.

Symptomatic Behavior Management

A broad range of sexual behaviors can be used addictively, from those that in themselves are natural and benign aspects of healthy life, such as masturbation and consensual adult sex, to those that violate other people and inflict deep and lasting damage, such as sexual abuse of children and rape. The goals of symptomatic behavior management in the treatment of sex addiction vary with the degrees of violation and damage that are associated with the behavior.

At the more violating-damaging end of the continuum, the goal is to prevent any occurrence of symptomatic sexual behavior. At the less violating-damaging end, the ultimate goal is to facilitate addicts' development of capacities to distinguish between healthy and unhealthy sexual behaviors (the latter being those that if engaged in by this particular person are likely to be engaged in addictively and lead to harmful consequences) and between healthy and unhealthy occasions of engagement in a sexual behavior that is not in itself unhealthy for this person. In general, engagement in a healthy sexual behavior is unhealthy when its primary motivation (usually unconscious) is to keep some heretofore unconscious mental content from entering awareness.

The middle region is occupied by sexual behaviors that have been identified as unhealthy for this person, either initially or as a result of his or her having learned how to identify unhealthy behaviors. In this broad area, the goals of symptomatic behavior management are twofold: (1) to prevent the occurrence of symptomatic sexual behavior; and (2) if symptomatic behavior does occur, to bring about its discontinuation before it leads to significant harmful consequences. Symptomatic behavior management consists primarily of behavior mastery training, other cognitive-behavioral techniques, and dialectical behavior therapy (DBT). On occasion, it may also require behavioral modification or endocrinologic medication. While support groups do not constitute treatment, they can be valuable adjuncts that sustain people's efforts to refrain from engaging in symptomatic behavior.

While discontinuation of all sexual behavior is not a goal of treatment, an initial period of 1 to 3 months during which addicts refrain from any kind of sexual behavior is often a useful launch into treatment. Few if any are able yet to reliably distinguish healthy from unhealthy with respect to either sexual behavior or occasions of engagement in it. During this period of refrainment, self-perpetuating behavioral, neurobiological, and

psychodynamic patterns will have a chance to cool down, and subjective states that symptomatic behavior had functioned to exclude from awareness may begin to bubble in. During such a period of complete sexual refrainment, any instance of sexual behavior can be explored to identify vulnerability factors, triggers, and what the trigger evoked internally that overwhelmed or undermined the person's control system.

Behavior Mastery Training. Behavior mastery training is a set of techniques that sex addicts can use on their own to help themselves refrain from enacting symptomatic behavior. The three primary techniques are risk management, urge handling, and symptom containment.

RISK MANAGEMENT. Symptomatic sexual behavior is more likely to occur under some conditions than others. Risk management teaches sex addicts to identify factors that increase the likelihood of symptomatic behavior. They then are trained to initiate protection against the influence of these factors before urges to engage in the behavior overwhelm their ability to organize an effective defense.

The first step is to review specific instances of symptomatic behavior in detail while focusing on the thoughts, feelings, and circumstances that led up to the act. Such reviews often can illuminate sequences of steps that culminate in episodes of symptomatic behavior, key features or patterns that different sequences share, and triggers that precipitate the sequences. Once identified, high-risk situations can be avoided or entered only when accompanied. Behaviors that increase susceptibility can be discontinued or restricted. Recognition of risk factors is also facilitated when the addict practices self-monitoring: for example, by keeping a daily journal of situations, times of day, moods, people, activities, and thoughts that are associated with urges to engage in symptomatic behavior.

The earlier in the process leading up to symptomatic behavior that a person's healthy resources can be mobilized, the less likely is he or she to engage in symptomatic behavior. Consequently, a critical component of risk management is identification of behavioral warning signs. These are behaviors that occur early in a sequence that culminates in symptomatic behavior and thus can be read as signals that such a sequence has already begun (and must be interrupted).

URGE HANDLING. During the course of healing from sex addiction, urges to engage in symptomatic behavior are inevitable. Urge-handling comprises behavioral and cognitive skills that help people who use sexual behavior addictively manage these urges without reverting to symptomatic behavior.

Behavioral skills are actions that can be put into practice. Beyond avoiding an expanded range of risk factors, they include specific protective responses that people can enact when they are exposed to a risk factor—for example, physically leaving a situation, engaging in an activity that

precludes the next step toward sexually acting out, or contacting a peer and talking about the situation. Behavioral skills are most effective when planned in advance and rehearsed, since the concentration that formulation and performance of an unfamiliar behavioral plan requires may not be available in the midst of a strong addictive urge.

Cognitive skills are methods that people can employ to modify their own thoughts. They include (1) accepting urges to engage in symptomatic sexual behavior as natural accompaniments of the healing process and hence not as signs of inadequacy; (2) recognizing that urges are not imperatives to act but surface manifestations of affects; (3) reviewing the benefits of refraining from symptomatic behavior and the potential negative consequences of engaging in that behavior; and (4) when fantasies of engaging in the behavior persist, prolonging the fantasies beyond the pleasurable experiences and imagining the harmful consequences and accompanying guilt, shame, or humiliation that could ensue. Particularly in the early phases of healing, cognitive skills are most likely to be effective in practice when they have been verbally rehearsed with another person.

SYMPTOM CONTAINMENT. An episode of symptom recurrence begins with an instance of symptomatic sexual behavior and ends when the addict resumes refraining from such behavior before it produces significant harmful consequences. If the behavior continues to the point that it generates harmful consequences, the event is not a recurrence (of symptoms) but a reversion (of health status). Symptom-containment skills are developed in order to prevent a recurrence from becoming a reversion. The goal is to reinstate refrainment before the symptomatic behavior generates significant harmful consequences.

A discussion of symptom-containment skills presupposes that symptomatic behavior can recur and thus seems to imply that recurrences can to some degree be tolerated. Such tolerance may make sense when the symptomatic behavior is visiting strip clubs or masturbating to pornography. But some sexual behaviors, such as rape and sexual contact with children, are so harmful to others that zero tolerance of the behavior is the only reasonable treatment stance. When people who use sexual behavior addictively are not reliably able to restrain themselves from acting on urges to engage in such seriously harmful behavior, treatment begins with behavior modification or endocrinologic pharmacotherapy or both. These interventions are continued until such time as the patients are reliably able to restrain themselves. If a sex-addicted person who is not reliably able to restrain him- or herself from acting on urges to engage in seriously harmful sexual behavior does not accept a recommendation for behavior modification and/or endocrinologic pharmacotherapy, a clinician would be well advised not to undertake the treatment of this person. The risks are considerable and if the patient is not sufficiently motivated to

cooperate with a recommendation of this type, the likelihood of significant progress in treatment is too small to justify such risks. The discussion of symptom-containment that follows thus applies only to the treatment of sex-addicted individuals whose sexual behaviors, while far from benign, are not so seriously harmful to others.

Even when motivation is genuine and therapeutic management is competent, symptomatic behavior can recur. However much a sex addict may consciously wish to refrain from symptomatic behavior, a strong unconscious motivation to resume the behavior persists—fueled by the same motivations and self-regulatory impairments that originally drove the addictive use of the behavior. Some combination of emotional stress, narcissistic injury, weakening of internal behavioral controls, and gaps in the external support structure can converge momentarily to shift the motivational balance in favor of resuming the behavior. When such a motivational shift happens to coincide with the availability of an opportunity to engage in the behavior, symptom recurrence is likely.

Like urge-handling skills, symptom-containment skills consist of behavioral and cognitive skills. Behavioral skills are actions that limit the progression of symptomatic behavior and reinstitute healing, such as communicating with one's therapist, getting together with a supportive friend, or attending a support group meeting. Similar to the behavioral skills of urge-coping, those of symptom containment are most effective when they are planned in advance and rehearsed.

Cognitive skills for handling symptom recurrence involve reframing the recurrence or reconceptualizing what having engaged in symptomatic behavior means. A recurrence is most usefully understood not as a sign of failure or the beginning of reversion but as a signal that a change in the addict's self-management program is needed. It represents a combination of affect and availability for which the addict had not yet developed an effective coping strategy. A recurrence thus reframed becomes an opportunity to learn and to improve the therapeutic program. A problem-solving approach is ultimately more effective in containing symptomatic recurrences and promoting healing than is judgmental reproof or self-reproach. In the case of both behavioral and cognitive skills for symptom containment, the sooner after a recurrence that an intervention occurs, the greater the probability of preventing reversion.

Behavior mastery training conceptualizes urges to engage in symptomatic sexual behavior as signals of disruptive affects that addicts need to develop healthier, more adaptive means to manage. In thus shifting the focus from controlling the behavior to understanding the affects, behavior mastery training provides a natural bridge from behavior management to psychodynamic psychotherapy. It can be imparted in therapeutic groups or integrated into individual psychotherapy.

Other Cognitive-Behavioral Techniques. Cognitive-behavioral therapy (CBT) is a psychotherapeutic approach that conceptualizes psychological problems mainly in terms of maladaptive learning. Treatment focuses on identifying and correcting inaccurate or distorted thoughts as a means of changing maladaptive behaviors and emotional states. Interventions typically are directive, practical, task-oriented, and educative in nature. Through CBT, people acquire new information, develop new ways of thinking and reacting, and learn specific skills that they can use for the rest of their lives.

A variety of cognitive-behavioral techniques in addition to behavior mastery training can be helpful in the treatment of sex addiction. Patients can be helped to identify distorted thoughts that enable them to dehumanize people who are hurt by their symptomatic sexual behavior or to avoid becoming aware of the emotional, relational, and existential ways in which their sexual behavior harms themselves.

Victim empathy training helps patients become aware of how their behavior affects other people's emotional states. The training focuses not only on direct victims on whom the sexual behavior was perpetrated (when it was inflicted on others) but also on indirect victims: people in the patients' lives who have been hurt or damaged as a result of the patients' sexual behavior, other behavior that prepared for or concealed the sexual behavior, and (eventually, perhaps) behavioral symptoms of the patients' character pathology. Patients are helped to become aware of how their symptomatic behavior has caused people about whom they care to have felt or to feel painful emotional states: abandonment feelings, intense anxiety, loss of self-esteem, and so on. The empathically experienced pain and associated feelings of remorse, guilt, or self-hatred can then function as the aversive affects that are associated with symptomatic sexual behavior in the aversive affect induction procedure. The potential usefulness of this variant is limited to patients who are capable of consciously accessible empathic perception.

Most other cognitive-behavioral techniques focus not on the symptomatic behavior itself but on other aspects of a person's life that predispose to reliance on symptomatic behavior as a means of coping with painful affects and unmet needs. Examples of skills training that helps patients to more effectively manage their affects and meet their needs include anger management, assertiveness training, problem solving, communications training (e.g., learning to speak in the format "When you did x I felt y" rather than accusing or blaming), and lessons in how to initiate and maintain healthy sexual relationships.

Dialectical Behavior Therapy. DBT is a variant of CBT that was developed to enable people who suffer from borderline personality disorder (BPD) to manage their emotions more effectively and thereby gain control

over their self-harming and self-endangering behaviors. It has been applied also in the treatment of substance use disorders, eating disorders, and sex addiction. DBT understands the core vulnerability in BPD to be emotional dysregulation. Its fundamental objective is to help patients develop skills that enhance emotional regulation and implement them in their daily lives. These skills are organized into four modules: mindfulness skills, distress tolerance skills, emotion regulation skills, and interpersonal effectiveness skills.

Mindfulness skills in DBT are cognitive-behavioral adaptations of Eastern meditation practice and philosophy. They are identified as observation, description, and participation in the present moment effectively and without judgment. Patients are instructed to nonjudgmentally focus their attention on their inner and outer environments and then to describe their observations. They also are taught to act with focused attention and a nonjudgmental mind. A version of meditation typically is taught as a means of developing mindfulness skills.

Distress tolerance skills enable patients to experience painful and disruptive feelings that they cannot immediately change without impulsively engaging in maladaptive, self-harming actions in a desperate reach for relief. At their core is learning to practice nonevaluative acceptance of experienced reality: one's current emotional state(s), the current situation, oneself, the world in which one lives. The goal is to help patients become capable of calmly recognizing painful or threatening situations and their impact rather than hiding from them or being overwhelmed. They then are more able to make wise decisions about whether to act and if so, how.

Emotion regulation skills enable patients to reduce their maladaptive emotional reactions to situations in day-to-day life and their tendency to respond to such emotional reactions with impulsive behavior. Patients are taught to (1) identify and describe their emotions, (2) understand the functions of their emotions, (3) reduce their vulnerability to the triggering of impulsive actions by intense emotions, (4) identify and challenge cognitions that lead to maladaptive emotional reactions, (5) identify obstacles to modifying their maladaptive emotional reactions, and (6) increase the frequency of positive emotional experiences in their lives.

Interpersonal effectiveness skills in DBT range from basic social skills to assertiveness and goal-oriented interpersonal problem solving. They include effective strategies for asking for what one wants, declining requests to do something one does not want to do, and negotiating interpersonal conflict. The skills taught are intended to maximize the chances that a person's goals in a specific situation will be met while at the same time not damaging either the relationship or the person's self-respect.

Patients' histories usually indicate general areas where skills are deficient and thus where skills training may be helpful or necessary. More

specific skill deficits and their consequences can be elucidated by chain analysis, a CBT technique that facilitates identification of the factors that promoted or permitted an occurrence of maladaptive behavior. When an instance of self-harming or self-endangering behavior has occurred, the sequence of cues, emotional responses, and cognitions that led up to and followed the behavior is reviewed in detail. The therapist and patient then attempt to determine the specific vulnerability factors and antecedent events that increased the likelihood that the behavior would occur and the consequences that reinforced the behavior. The goal of chain analysis is to identify the patterns of variables that tend to control the patient's dysfunctional behaviors. This information then guides training in skills that reduce vulnerabilities in the future and lower the risk that precipitating events will recur. Chain analyses in DBT often focus on the various ways in which emotions are related to the initiation and maintenance of dysfunctional behaviors. Attention is directed toward the intense or aversive emotional states that may precede target behaviors as well as the potential emotion regulatory function that such behaviors may serve.

To the extent that DBT helps patients develop skills that enhance their emotional regulation, it contributes not only to symptomatic behavior management but also to the healing of the addictive process.

Treatments to Modify Paraphilic Behavior. As noted earlier, paraphilic disorders and sex addictions are different kinds of disorder that share an area of overlap. For many of the individuals who occupy this area of overlap, failure to address the paraphilic nature of their arousal patterns, urges, and sexual behavior can undermine the effectiveness of treatment for sex addiction. In some of these individuals, those who are paraphilic predators, treatment for sex addiction outside a locked facility may entail too much risk of harm to others unless their paraphilic behavior and sometimes their capacity for sexual arousal are shut down. Worthy of note is that the class of paraphilias includes sexual behavior in which one person coerces another into participation. Such behavior is motivated more by sadism than by craving for sexual gratification. In DSM-5, sexual sadism disorder is included among the paraphilic disorders.

BEHAVIOR MODIFICATION. Behavioral methods are oriented toward helping paraphilic individuals reduce their erotic arousal to paraphilic objects/subjects or shift the balance of erotic arousal potential from paraphilic to nonparaphilic interests. The primary forms of behavioral treatment for paraphilic disorders are aversion conditioning, covert sensitization, masturbatory training, imaginal desensitization, and aversive affect induction. All rely on classical or Pavlovian conditioning, in which the repeated pairing of two stimuli increasingly enables one of them to elicit physiological and affective responses that previously had been elicited

only by the other.

In aversion conditioning, unwanted patterns of sexual behavior are linked repeatedly with unpleasant stimuli, most often electric shock or foul odors. This straightforward classical conditioning procedure is intended to transform stimulus features of the sexual behavior into conditioned stimuli or triggers for the aversive responses to the unpleasant stimuli. A number of uncontrolled studies found that aversion conditioning yielded positive results, while one found no significant effect. Another study found that the success of aversion conditioning depended on the availability of appropriate sexual outlets and the absence of severe nonsexual pathology.

Covert sensitization was developed as an imaginary form of aversion conditioning, in which fantasies of paraphilic arousal are paired with fantasies of aversive events in order to promote a learned association between paraphilic themes and unpleasant feelings. In covert sensitization training, patients are first led through a relaxation exercise, which enhances their receptiveness to vivid mental imagery and helps them to focus their attention on the imagery rather than on external events. They then are instructed to visualize themselves engaging in some aspect of their paraphilic behaviors and are encouraged to feel personally involved. At that point, patients are told to imagine either an unpleasant event, such as nausea and vomiting, or an aversive consequence that could follow the behavior in reality, such as public humiliation or shame at being discovered by someone whose esteem deeply matters to them. The aversive valence of the imagined noxious scene can be reinforced by the presentation of a nauseating odor. Studies have found covert sensitization to be effective in the treatment of exhibitionism and pedophilia. It neither inflicts physical discomfort nor requires expensive technology and patients can be taught to self-administer the treatment. Consequently, many behavior therapists prefer to begin with covert sensitization, turning to aversion therapy only for patients who do not respond to covert sensitization.

Masturbatory training attempts to shift paraphilic patients' arousal patterns in the conventional direction by controlling the visual stimuli they view or the fantasies they conjure when they masturbate. The two main forms of masturbatory training are satiation and fading. Satiation involves instructing male patients to masturbate to orgasm with conventional sexual stimuli or fantasies and then to continue masturbating while visualizing their deviant objects. The idea is that masturbating after orgasm will extinguish erotic arousal to paraphilic fantasies. The aim of the fading technique is to shift patients' sexual fantasies from deviant to conventional during periods of sexual arousal. Patients may be provided with a video or sequence of slides on which to focus while they masturbate. As they masturbate, the visual presentation progressively fades from paraphilic to nonparaphilic sexual images. Alternatively, patients may be instructed to begin masturbating to paraphilic fantasies and then gradually evolve the

fantasies to nonparaphilic sexual scenarios. One study of masturbatory satiation reported mild success, but other therapeutic modalities were provided at the same time and no posttreatment follow-up was mentioned.

The technique of imaginal desensitization is based on the theory that once an individual has developed a sequence of sexual activity that begins with a sexual thought or fantasy and ends with overt behavior, interruption of the sequence results in anxiety, which then motivates completion of the sequence. Imaginal desensitization uses the methods of systematic desensitization to diminish the anxiety that is aroused by interruption of that sequence. The procedure begins with the patient describing a scene in which he or she is stimulated to carry out the paraphilic sexual activities for which he or she sought treatment. After relaxation is induced, the patient visualizes performing a sequence of behaviors that leads to noncompletion of the act. Relaxation is allowed to develop with each visualized behavior before the patient proceeds to the next behavior in the sequence. The patient is thus trained to visualize not completing the sexual act while remaining relaxed. One study reported imaginal desensitization to be more effective than covert sensitization.

Aversive affect induction involves focusing on the negative emotional consequences of symptomatic sexual behavior while the urge to engage in such behavior is active. Specific cues that patients have identified as being associated in their minds with symptomatic sexual behavior (e.g., the home page of a web site that features sadistic pornography, the sounds of children playing) are presented to evoke an urge to engage in that behavior. Patients are then asked to describe the emotions that they experience before, during, and after an episode of symptomatic sexual behavior—for example, excitement, power, sexual release, shame, self-loathing. The intent is that associations develop in the patient's mind between urges to engage in symptomatic sexual behavior and narcissistically gashing emotions, such as shame and self-loathing. Aversive affect induction has the potential to be helpful in nonparaphilic sex addiction but its use in that capacity has not been studied.

ENDOCRINOLOGIC AGENTS. Endocrinologic agents are used primarily to reduce sex drive in paraphilic men. Those that are in use at this time include progestogens and gonadotropin-releasing hormone (GnRH) agonists. Estrogens have been studied for use in controlling sexual behavior but were found to have too many undesirable effects.

The progestogens that currently are being used in the treatment of paraphilias are cyproterone acetate (CPA) (Androcur) and medroxyprogesterone acetate (MPA) (Cyrin, Depo-Provera, Provera). Their action as progestogens blocks the secretion of the pituitary gonadotropins, luteinizing hormone (LH) and follicle-stimulating hormone (FSH). CPA has strong progestational and antigonadotropic activity. In addition, it

competes with androgens for androgen receptors and thus is an antiandrogen. CPA has been reported to decrease sexually driven aggression and to reduce arousal responses to pedophilic and coercive sexual stimuli but not to mutually consenting adult sex scenes. Side effects of weight gain, depression, and feminization have been noted. MPA exhibits no antiandrogenic activity. Its efficacy in treating paraphilias results primarily from its induction of the enzyme testosterone reductase in the liver, which decreases circulating levels of testosterone. A number of clinical studies have demonstrated that MPA can reduce sex drive, sexual thoughts, sexual behavior, and aggressiveness. Unfortunately, its administration is frequently associated with unpleasant side effects and occasionally with serious adverse events such as deep vein thrombosis, pulmonary embolism, hepatic dysfunction, and precipitation or exacerbation of diabetes mellitus. Noncompliance also is a significant problem. One study reported a dropout rate of greater than 50 percent. Both progestogens are available in oral and IM formulations.

Analogs of GnRH have been developed that have higher potency and longer duration of action than does naturally occurring GNRH. Triptorelin (Trelstar) is injected once a month and leuprolide acetate (Lupron) is injected once every 3 months. Initial administration of these agents raises serum testosterone by stimulating pituitary gonadotrope cells to increase production of LH and FSH. LH then stimulates testicular Leydig cells to synthesize more androgens. During the first 2 to 4 weeks after initiation of treatment, sex drive may actually increase. However, continuous administration of long-acting GnRH analogs produces downregulation of GnRH receptors on the pituitary gonadotropes. GnRH receptor downregulation leads to decreased secretion of LH and FSH, a consequent decrease in the synthesis of testosterone, and thus a substantial diminution of sex drive. Several reports of uncontrolled open-label trials of GnRH agonists in the treatment of paraphilias and hypersexual disorders have been published and all demonstrated significant positive effects. The main side effects were erectile dysfunction, hot flashes, and decrease in bone density. These results suggest that GnRH agonists could prove to be a more effective, safer, and less noxious alternative to the progestogenic agents. More study of these promising agents is warranted.

A recent review article recommended that endocrinologic agents be used in combination with serotonergic antidepressants, not as monotherapies, and that progestogens be tried before LHRH (GnRH) agonists. Another review reported that patients who had taken CPA, MPA, or selective serotonin reuptake inhibitors (SSRIs) and subsequently took a GnRH agonist reported greater therapeutic benefits from the latter.

Although these medications do not change the direction of the paraphilic person's sexual interest, they decrease the intensity of sexual drive, so the person is more in control and less likely to engage in

paraphilic behavior. The primary therapeutic function of these medications is to reduce sex drive to manageable levels in those individuals whose ability to control their behavioral impulses is so impaired as to put them at risk either to injure themselves or others or to render them unresponsive to psychological interventions. Endocrinologic agents can lower the risk of problematic sexual behavior during the interval between the initiation of treatment and the consolidation of the changes that affect-regulating agents, behavior modification, DBT, and/or psychodynamic psychotherapy can induce. A recent review article on the treatment of paraphilias concluded that the combination of pharmacological therapy and psychotherapy is associated with greater efficacy than is either treatment alone.

A number of other medications that were developed for other purposes have side effects of decreased sex drive or erectile dysfunction and are on occasion used informally—often by patients on their own initiative and sometimes without informed medical direction—in attempts to reduce the symptoms of sex addiction. These include antihypertensive agents, anticholinergics, antihistamines, disulfiram (Antabuse), antipsychotic medications, and hair growth stimulants. None of these have been studied for use in controlling sexual fantasies, urges, or behavior.

Support Groups. Support groups are distinguished from treatment groups in that the effective agent is neither a therapeutic method/technique nor a person who leads the group but the support that group members provide for one another. Support groups for sex addicts typically provide a safe, minimally judgmental sanctuary where they can talk about shame- and guilt-laden matters with others whose presence in the group and self-identification as sex addicts indicates that they too struggle with comparable shame- and guilt-laden matters. In such groups sex addicts can learn from one another's experience, identify with others who manage their addiction more effectively or make good decisions more often, benefit from others' encouragement and positive responses to their successes, and feel compassion for others' stumbles when they are not yet ready to begin feeling compassion for themselves. When they or others in the group engage in symptomatic behavior (recurrence or reversion) and are welcomed back into the group with neither reprimand nor willingness to hear excuses, they can learn to be more self-accepting and also more honest with themselves. More broadly, sex addiction support groups provide opportunities to experience commonality (the "I'm not the only one" experience), empathic bonding, group cohesiveness, and sense of belonging. They thereby can facilitate the development of abilities to make meaningful connections with others and to turn to people in times of need instead of turning to addictive behavior.

The most widespread and easy-to-access support groups for sex addicts

are the 12-step groups that have adapted the model of Alcoholics Anonymous to sex addiction. They offer membership at no charge for as long as a person wishes, around-the-clock support, and connection with a worldwide support network. They present a coherent framework for approaching addictive problems and life in general that includes self-awareness, acceptance of what is, and the centrality of human relationships. In addition, 12-step groups offer (but do not impose) a nonjudgmental, nondogmatic spiritual foundation. Meanwhile, a consequence of the lack of formal leadership structures and internal controls in 12-step groups is that the characteristics of a particular group are shaped primarily by the individuals who attend it. How much a person benefits from and enjoys being in a group is a function of his or her goodness of fit with the other participants in the group.

Healing of the Addictive Process

The addictive process can be addressed through psychiatric medications that improve affect regulation (which includes regulation of anxiety) or reduce impulsivity and psychodynamic psychotherapy. Meditation and yoga can be helpful adjuncts to the primary treatment modalities. Since the addictive process underlies all addictive disorders, the material in this section is not specific to sex addiction but applies to other addictive disorders as well.

Psychiatric Pharmacotherapy. Psychiatric pharmacotherapy is direct intervention to improve emotional and behavioral self-regulation as well as to treat other symptoms of comorbid psychiatric disorders. In sex addiction, craving and urges to act out are expressions of dysregulated emotional states. Such urges are more likely to be acted out when behavioral regulation is impaired. Consequently, enhancement of affect regulation tends to diminish the frequency and intensity of addictive urges while enhancement of behavioral regulation reduces the likelihood that urges will lead to an episode of symptomatic behavior. Medications that improve affect regulation often also reduce impulsivity, either directly or through their effects on affect regulation.

AFFECT-REGULATING AGENTS. A number of case reports and open-label studies have provided evidence for the efficacy of affect-regulating agents, particularly the SSRIs, in the treatment of paraphilias and nonparaphilic sex addictions, even in the absence of major depression. Some studies found that paraphilic patients with comorbid depression showed a concurrent decrease in paraphilic behavior when their depressive symptoms improved. Agents that have been found to be effective include fluoxetine (Prozac, Sarafem), sertraline (Zoloft), citalopram (Celexa), paroxetine (Brisdelle, Paxil, Pexeva), fluvoxamine (Luvox), nefazodone

(Serzone), imipramine (Tofranil), desipramine (Norpramin), clomipramine (Anafranil), lithium carbonate (Eskalith, Lithane, Lithobid, Lithonate), topiramate (Qudexy, Topamax, Trokendi), carbamazepine (Carbatrol, Equetro, Tegretol), divalproex (Depakote), buspirone (BuSpar), and electroconvulsive therapy. Most of these studies reported a positive response rate in the range of 50 to 90 percent. One investigator reported anecdotally that augmentation of an SSRI with bupropion (Budeprion, Forfivo, Wellbutrin, Zyban) or a psychostimulant can further reduce sexual fantasies, urges, and behavior, particularly when symptoms of attention-deficit disorder (ADD) are present. Others have found that divalproex or lamotrigine (Lamictal) can be helpful for sex addiction symptoms that arise in the context of atypical bipolar conditions or "emotionally unstable character disorders." Antidepressant medications, especially the SSRIs, can produce diminished libido as a side effect, but a number of studies noted that antidepressants reduced the drive for symptomatic sexual behavior without decreasing the drive for healthy sexual behavior. The only double-blind, placebo-controlled study in this series found that citalopram reduced the frequency of sex addicts' symptomatic sexual urges, masturbation, and use of pornography while having no significant effect on partnered sexual behaviors. The large number of positive findings is encouraging but the confidence with which conclusions can be drawn from them is limited by the paucity of controlled studies.

That the symptoms of sex addiction respond to affect-regulating medications in a pattern that resembles the response pattern of affective disorder symptoms is consistent with the earlier description of the addictive process that underlies sex addiction. As formulated neurobiologically, the addictive process originates in impaired affect regulation, impaired behavioral inhibition, and impaired functioning of the motivational-reward system. Addictive urges can be recognized to represent both (1) expressions of dysregulated affect, often a nonspecific tense or irritable dysphoria that is experienced consciously less as an affect than as an urge to do something, usually to engage in symptomatic behavior; and (2) conditioned stimuli for coping responses (responses to cope with dysregulated affect) that have been learned in the context of impaired behavioral inhibition and impaired motivational-reward function. Interventions that enhance affect regulation and behavioral inhibition can thus be expected to diminish addictive craving and symptomatic behavior.

In nontechnical terms, the impairment of affect regulation that underlies addiction can be seen as an imbalance between the intensity of patients' affects, particularly their painful affects, and the capacity of their built-in coping systems. Painful affects such as anxiety and dysphoria are not in themselves pathological or "negative." They are signals that people send to themselves to call their attention to something that is important for their physical or psychological survival. Painful affects become a problem

when their intensity overwhelms or threatens to overwhelm a person's coping system. They are then experienced as extreme dangers that must be escaped or avoided and are no longer used as signals. Often patients will not consciously experience the painful affects or even the threats of these affects but will feel only a vague unease. Or they may be closed off to their affects and may experience only bodily states that they interpret as urges to engage in some kind of symptomatic behavior. People who have learned that certain self-initiated behaviors can quell such dangerous affects and prevent the system from being overwhelmed will experience strong urges to engage in these behaviors whenever such affects threaten. This process forms the core of addictive and compulsive patterns of behavior.

Psychiatric medication works by restoring or reinforcing the brain chemistry that is associated with patients' built-in coping systems. Although psychodynamic psychotherapy also can enhance the capacity of their coping systems, patients' accessibility to psychotherapy and to what it can offer may be limited when they are emotionally overwhelmed. To the extent that they shut down or shut out subjective experience because they unconsciously perceive that it threatens their psychological survival, their access to their inner lives will be compromised. Moreover, states of being overwhelmed can disorganize or disrupt cognitive functioning, so they not only are less aware of material with which they can work, they also are less able to work with the material of which they are aware. When psychiatric medication is effective in establishing a more stable balance between patients' affects and the capacity of their coping systems, the coping systems are less likely to be overwhelmed and patients can better tolerate experiencing their painful affects as affects (rather than as states of emergency). Consequently, they will have more conscious access to their inner states and their cognitive functions will be more organized and more free to operate. These developments provide them with more opportunity to work with their inner states in psychotherapy. Subsequent psychotherapeutic work can then help them to develop healthier, more adaptive ways of regulating and responding to their inner states, which may eventually enable them to manage without the psychiatric medication.

MU-OPIOID RECEPTOR ANTAGONISTS. The mu-opioid receptor antagonist naltrexone has been reported to reduce symptoms of "compulsive sexual behavior" as an adjunct to affect-regulating medications, primarily SSRIs. Other studies have found naltrexone by itself to reduce urges and symptomatic behavior in gambling disorder and kleptomania as well as in substance use disorders. Some reports indicate that it reduces the pleasure that is associated with symptomatic behavior but not the pleasure that occurs in the course of normal living.

Psychodynamic Psychotherapy. Psychodynamic (or psychoanalytic)

psychotherapy is a circumscribed dialogue in the context of an intimate relationship between two people, the explicit goal of which is to help one of the participants become more fully self-aware in order to master emotional distress and remedy impairments in social or occupational functioning. The central role of developing self-awareness distinguishes psychodynamic psychotherapy from other forms of psychological treatment. Mastering emotional distress and remedying functional impairment identify this process as treatment, above and beyond the educational and growth-promoting functions that it serves. In short, psychodynamic psychotherapy is a dialogue that fosters self-awareness in order to heal the personality. Its functional core is constituted by three interacting processes: understanding, integration, and internalization.

UNDERSTANDING. Understanding as a component of psychodynamic psychotherapy is verbally mediated knowledge that the psychotherapist imparts in the course of treatment. It is a primarily cognitive process that is not specific to psychodynamic psychotherapy. What is more specific (in addition to the content) is the recognition that patients' difficulties in understanding or in the practical application of what has been understood may reflect motivational factors that operate outside their conscious awareness. Moreover, psychodynamic psychotherapy regards these difficulties not as obstacles but as opportunities to explore the nature, function, and history of the underlying motivational factors. Both general and specific understanding are useful.

General understanding concerns information about a condition from which a person may be suffering and factors that characteristically contribute to its development. At the beginning of treatment this process often enhances a person's capacity and motivation to participate in psychodynamic psychotherapy.

For example, early in the treatment of sex addicts, the psychotherapist can briefly describe in simple terms the psychodynamic processes that underlie sex addiction and other addictive disorders. Such information first of all provides an organized array of concepts that can help people name their experiences, find order in them, and integrate them into their cognitive networks. Having words to describe inner states and processes gives the conscious mind a handle by which to apprehend them. Moreover, naming inner realities provides perspective. The part of the mind that names in so doing steps back from the part of the mind that feels. Painful emotions and self-states no longer occupy the totality of a person's subjective psychic space, so he or she becomes less likely to be overwhelmed by them.

As noted earlier, emotions are important self-generated signals. But when they become or threaten to become overwhelming, people experience them as threats to their survival. Their survival instincts then drive them to

ward off or shut down their emotions. Insofar as they do so, they are unable to benefit from their emotions' signal function and the information that it conveys. To the extent that people get perspective and become less overwhelmed, they can begin to experience their emotions as signals rather than as states of emergency. Significant progress down this path usually depends on the integration and internalization components of psychodynamic psychotherapy.

General understanding also offers benefits that are more tangibly practical and specific to addiction. Learning that addictive behavior represents an attempt to ward off painful or disorganizing inner states that threaten to overwhelm the person attenuates the shame and self-loathing that addicts feel in relation to their symptomatic behavior and reduces associated self-destructive inclinations. Hearing that urges to engage in symptomatic behavior are conscious manifestations of anxiety that surges when such threats arise shift's addicts' focus from fighting the urge to understanding the threat. In consequence, behavioral symptom management becomes less effortful and addicts become more capable of using psychodynamic psychotherapy effectively.

General information about developmental factors that may have contributed to the genesis of sex addiction provides a framework within which the condition becomes comprehensible, often thereby diminishing addicts' anxiety. Moreover, it reduces the self-hatred or self-punishment that people typically experience when they blame themselves for the emotional pain and functional impairment that are associated with sex addiction.

Specific understanding concerns the patient him- or herself and what actually is going on in his or her life. The primary focus of specific understanding is the dynamic process that has led up to a particular instance of symptomatic sexual behavior. This process is typically of the following form: (1) impaired regulation of emotional states develops out of the interaction between genetic predispositions and the infant's relationship with his or her early caregiving environment; (2) specific vulnerabilities develop as the emotions that are evoked by painful or scary childhood events (or less commonly, postchildhood events) are hyperintense to the point of being traumatic due to impaired affect regulation; (3) later in life a current event taps into one of those vulnerabilities in the context of deficient affect regulation and thereby triggers a disproportionately intense emotional response that threatens to overwhelm the person; and (4) a desperate reflexive attempt to protect oneself from this threat precipitates an occurrence of symptomatic behavior.

In general, behavioral symptoms arise as (1) expressions of painful or disturbing emotional states that an impaired affect regulation system is not effectively managing, (2) unconsciously directed self-protective processes

that operate to keep such emotional states and inner conflicts out of one's awareness, or (3) a combination of both. In addictive disorders, symptomatic behavior may be the only conscious manifestation of dysregulated emotional states that at an unconscious level threaten to burst into a patient's awareness and overwhelm him or her. Thus, in addition to being distressful or debilitating, symptomatic behaviors are conveyors of information about patients' emotional states, affect regulation systems, and self-protective processes.

The likelihood of accessing this information is greatest when inquiry focuses on the timing of changes in the symptomatic behavior—when it emerged out of quiescence, got worse, got better, or changed shape. The emergence or exacerbation of symptomatic behavior indicates that some event activated a painful or disturbing emotional state that the affect regulation system was unable to manage. The point of information access is temporal contiguity. Unless the timing of symptoms is random, the occurrence of an instance of symptomatic behavior implies an antecedent change in contextual conditions that (by evoking a potentially traumatic emotional state) led the behavior to occur when it did rather than at some other time. Specification of when the symptomatic behavior occurred guides an investigation of what was going on in the person's life just before the symptomatic act, with an eye toward identifying the event(s) that may have precipitated it.

The process of specific understanding typically is initiated when a patient reports an instance of symptomatic behavior. In the treatment of sex addiction, the initial focus is on symptomatic sexual behavior. As treatment progresses and as the frequency of symptomatic sexual behavior decreases, the focus may expand to other symptomatic or maladaptive behaviors. The process begins with a review of circumstances and interactions that preceded the episode of symptomatic behavior. During this review antennae are tuned for any event that might have triggered an intense emotional response. Identification of the triggering event usually leads to identification of the emotional response, though on occasion the emotional response is identified first. The goal is to help patients increase their capacity to be aware of disproportionate intensity in their emotional responses before they act, so they can factor it out of their decisions about what actions to take.

Guided by the hypothesis that symptomatic behavior characteristically occurs when an event taps into an area of emotional vulnerability, the investigation readily expands in the direction of exploring the contributory factors that were shaped by personal history and intensified beyond the range of ordinary manageability by impaired affect regulation. As it does so, the therapeutic process shifts beyond understanding to include integration, the next component of psychodynamic psychotherapy to be discussed.

INTEGRATION. Integration refers to a process through which the patient's personality or self becomes increasingly integrated as its various aspects coalesce into a coherent whole. Psychologically, many people are characterized by deficits in integration on several levels. Without conscious intent or awareness they reflexively engage in a variety of mental processes that function to protect them from emotional trauma. Most of these self-protective processes can be understood as ways of keeping out of people's awareness material that they unconsciously imagine would lead to overwhelming pain or self-fragmentation were they to become aware of it. Such material can include emotions, needs, wishes, fears, inner conflicts, core beliefs, or the self-protective processes themselves. Ultimately, what they are protecting themselves from most of the time are emotions that they fear would overwhelm and psychologically destroy them were they allowed to enter their consciousness.

For example, when people automatically push a need or a wish out of their awareness, they do so usually not because the need or wish is dangerous in itself but because past experiences of other people's responses to their expressing that need or wish have led them to associate it with potentially overwhelming emotions such as disappointment, loss, shame, guilt, rage, or humiliation. Similarly, when people automatically keep an inner conflict unconscious—for example, by allowing only one of its components into awareness at a time—they do so because the conflict is fraught with unbearable anxiety, which in turn develops because choosing one of the alternatives and giving up the other is associated with painful emotions that could be traumatic whichever alternative is chosen. Finally, the self-protective processes that people keep out of their awareness typically are not themselves infused with overwhelming emotion. But if someone were aware that a self-protective process was operating, it would be less likely to succeed in concealing the emotion-laden material that it was unconsciously intended to conceal. Thus, people also (unconsciously) hide their self-protective processes from their inner view.

Self-protective processes are not obstacles to self-awareness that must be torn down for healing to occur. Nor do they constitute a "false self" that must be dismantled or peeled away to liberate and reveal a person's "true self." People's means of protecting themselves from emotional trauma are no less true than are the mental processes and material that they keep from consciousness and no less important as potential or actual objects of self-awareness. Exploration of self-protective processes often yields treasure troves of information. In addition, a person's psychological survival can at times depend on inner processes that protect him or her from emotional trauma. The goal of treatment is not to get rid of them but to increase the influence of patients' conscious reason on when and how they operate.

People's subjective experience and behavior are influenced by conscious thoughts, feelings, and reasoning; and also by emotions, needs, wishes,

fears, inner conflicts, core beliefs, and self-protective processes of which they are not consciously aware. When such influences operate outside a person's awareness, they are not susceptible to being modulated or contextualized by his or her conscious reason.

The process of integration occurs through consciousness. The critical step is becoming aware of feelings, thoughts, and fantasies that heretofore had remained unavailable to conscious awareness. Consciousness is primarily a verbally mediated process, so becoming aware of what had been going on in one's mind unconsciously involves the development of associative links between unconscious mental processes and their verbal representations. Such links are the basic elements of integration. They typically develop as a result of temporal contiguity between the occurrence of a heretofore unconscious mental process and the conscious registration of its verbal designation or description.

Psychodynamic psychotherapy promotes integration (1) by providing an intimate interpersonal situation that tends to stir up people's emotions, needs, wishes, fears, inner conflicts, core beliefs, and self-protective processes; (2) by emphasizing verbal communication and meaning, with both patient and therapist working together to create verbal renditions of the patient's inner world; and (3) by facilitating identification and working through of self-protective processes that interfere with the flow of experience or communication. It thus provides optimal conditions for patients to link their unconscious mental processes with the co-created verbal descriptions of them and thus to develop the self-awareness that leads to integration.

Psychodynamic psychotherapy is less structured and less directed by the therapist than are other forms of psychotherapy. Either from the very beginning or after an initial period during which background information is gathered, the steering wheel is left in or returned to the patient's hands. Meanwhile, the therapist has at hand an auxiliary set of controls, as in a driver education car, so he or she can take over if for whatever reason the patient becomes temporarily incapable of driving safely. Recognition of this safety feature, which often remains implicit, helps enable patients to take more risks in exploring realms of their mental world with their therapists than they would have done on their own.

The patient's primary instruction (which may or may not be explicitly stated) is to say whatever occurs to his or her mind without directing or censoring it, even if it seems irrelevant, trivial, offensive, or embarrassing. Psychodynamic factors outside people's awareness operate all the time, influencing what they think, say, and do. During most of the time that they are awake, their conscious mental processes also are operating and they dominate the conscious field of view. The influences of unconscious psychodynamic processes are difficult to discern except perhaps when indirectly embodied in slips of the tongue or symptomatic acts. To the

extent that the dominance of conscious reason over patients' verbal output is softened, the influence of unconscious psychodynamic factors over the content and form of what they say becomes more discernable. As patients' words increasingly reflect the operation of psychodynamic factors, inferences about these factors that are based on what patients say become more productive.

In general, psychodynamic psychotherapists do not set agendas for therapy sessions or direct the flow of the dialogue but try to discover and bring to light their patients' underlying emotions, needs, wishes, fears, inner conflicts, core beliefs, and self-protective processes, as inferred from what they observe in the patients and in themselves. Their primary charge is to offer comments and questions that are likely to expand their patients' self-awareness, directly or indirectly. Comments or questions that can expand patients' self-awareness indirectly include those that identify ways in which patients unconsciously constrain or sabotage the expansion of their self-awareness—usually to protect themselves from emotions that are associated with something of which they (unconsciously) fear to become aware.

Whereas the aspects of a person's self that can become integrated in treatment extend over many years of history, the process of integration occurs in the present and operates with present experience, whatever subject matter is being discussed. The closer the subject matter under discussion is to present experience, the more active and salient the associated emotions are, the more available are underlying psychodynamics to conscious exploration, and the more likely is integration to be potentiated by the discussion. Consequently, psychotherapeutic dialogue tends to be most effective in promoting integration when it addresses issues as they arise in the here-and-now. In the psychotherapeutic situation, the most here-and-now issues are those that concern the relationship between the patient and the therapist. Hence, these issues often are the most productive to address.

Early in life, people begin to develop inner models of self-and-other-in-relationship that organize how they perceive, experience, and respond to other people and also how they perceive and experience themselves in relation to others. These models or mental representations are more than audiovisual recordings. They encompass the somatic and behavioral residues of the complex of emotions, needs, wishes, fears, inner conflicts, core beliefs, and self-protective processes that individuals experienced during their interactions with other people or developed in response to such interactions. Their experiences with others in the first years of their lives have the most profound and enduring influence on their models for two reasons that probably are related. Humans are helpless at birth and during their early years depend on their caregivers for their survival. Since what they learn about their caregivers and how to respond to them is

literally a matter of life or death, natural selection has favored those who learned quickly and held on tightly to what they learned. Meanwhile, humans are born with central nervous systems that are only partially developed. Most of the remaining development occurs epigenetically during the first two years of life and is shaped by the young humans' interactions with their environment, which during those early years consists primarily of their caregivers. The developmental plasticity of human central nervous systems enables them to be much more specifically adapted to each individual's particular environment than would anything that the genetic code could engineer by itself. As a corollary, the people who constitute a person's early environment have a profound and enduring influence on how his or her central nervous system develops and subsequently functions.

People's inner models of self-and-other-in-relationship are usually of great value in their early relationships with their primary caregivers, on whom they depend for their survival. Their inner models guide them to act in ways that increase the likelihood that their needs will be met, decrease the likelihood that they will experience traumatic emotional-somatic pain, and in general enhance the chances that they will survive. As people's interpersonal world expands, their inner models continue to be adaptive to the extent that (1) the individuals with whom they are now interacting are similar to those who were around when the models initially developed, and (2) the models are sensitive to the differences between past and present people and can evolve to accommodate to the latter.

Sometimes, however, people's models of self-and-other-in-relationship are not flexible enough to accommodate to the differences between the caregivers whose interactions with them shaped the models and the individuals with whom they now interact—particularly when the models are suffused with anxiety about threats to survival. Under these conditions, people's inner models may guide them to perceive, experience, and respond to significant individuals in ways that are maladaptive—that cause them emotional distress or impair their social or occupational functioning. Often, such emotional distress and functional impairment are what lead people to seek help.

In psychodynamic psychotherapy, the relationship between the patient and therapist is the stage on which the patient's inner models reveal themselves. As patient and therapist spend time together, interacting at a level more intimate than ordinary discourse, the emotional significance of their relationship deepens and the patient's inner models of self-and-other-in-relationship become increasingly activated. A process develops through which the emotions, needs, wishes, fears, inner conflicts, core beliefs, and self-protective processes that are associated with these inner models and that unconsciously influence the patient's other significant relationships increasingly affect how the patient perceives, experiences, and responds to

the therapist.

The therapeutic relationship, however, differs from other close relationships in a crucial respect. To the extent that their own personal vulnerabilities permit, psychodynamic psychotherapists refrain from reacting with defensiveness, retaliation, withdrawal, sadism, or exploitation to their patients' words and actions. They recognize that patients' responses to them always reflect some degree of accurate perception of the therapist and some degree of distortion or disproportion due to the activation of patients' inner models that are incongruous with the current situation. To the extent that patients' inner models have participated, opportunities become available for both patient and therapist to learn how those models operate. The therapeutic relationship then provides a safe environment in which the two of them together can explore in real time the meaning of the patient's responses in terms of his or her emotions, needs, wishes, fears, inner conflicts, core beliefs, and means of self-protection. This collaborative exploration gives words to the patient's subjective, not-quite-conscious experience of these processes while they still are fresh, thus expanding the patient's awareness of them and promoting their integration into the patient's living narrative. Bringing together in consciousness the various disintegrated aspects of a person's psychic processes gradually heals the personality and enables it increasingly to function as an integrated whole. The net result is more conscious choice, more flexibility, and more freedom.

The interaction between a patient's inner models and his or her relationship with the therapist is more complex than was its description in the preceding paragraphs. The activation of the patient's inner models that manifests as distorted or disproportionate responses not only provides information and opportunities to expand awareness. It also opens the inner models to modification through the patient-therapist relationship. This process is an aspect of the third component of psychodynamic psychotherapy, internalization, which is discussed in the next section.

INTERNALIZATION. Internalization is the least understood component of psychodynamic psychotherapy. Originally it is the process whereby young children gradually become able to provide for themselves the regulatory functions that had been provided for them by their caregivers. The development of self-regulatory functions occurs gradually over the course of numerous child-caregiver interactions in which the caregiver responds more or less appropriately to the child's needs for nurturing, soothing, and holding. What is internalized over the course of these interactions is not only self-regulatory functions but moreover a relationship in which a caregiver provides regulatory functions for another who needs regulation. The term *internalization* is a metaphor that represents a combination of procedural learning and classical conditioning that interacts dialectically

with people's genetically guided (and epigenetically adjusted) maturation. The capacity for developmental internalization of self-regulatory functions is greatest during early childhood but it continues throughout life.

The function of internalization in psychodynamic psychotherapy derives from its role in the development of self-regulation. A primary means by which psychotherapy promotes the healing of impaired self-regulation is by providing new opportunities for people to internalize self-regulatory functions that they did not adequately internalize during their childhood. Development of affect regulation functions can be facilitated by internalization of therapists' provision of a holding environment, their soothing and enlivening functions, and their treatment of emotions as indications of inner dynamics that deserve attention. Development of self-care functions can be facilitated by internalization of the therapist's nurturing and protective functions, accompanied by internalization of the message that the patient is valued as a person and believed to be worth taking care of. Development of healthy self-governance functions can be facilitated by internalization of the therapist's integrity, respectfulness, nonjudgmental acceptance, and general stabilizing function. Development of self-mindfulness—a central factor in affect regulation, self-care, and self-governance—can be facilitated by internalization of the therapist's actively receptive attention to the patient's words, actions, and empathically discernable emotions. Raw material for internalization is provided not only by the therapists' relationships with their patients but also by how they manage or respond to everything else, including their own affects, intrusions of external circumstances into a session, minor mishaps, medical problems, and personal crises. Worthy of note is that patients' perception of potentially internalizeable material occurs through empathy as well as through the five primary senses.

Psychodynamic psychotherapy is not the only context in which self-regulatory functions can be internalized. Internalization can occur in any significant relationship and also in some group situations. What distinguishes psychodynamic psychotherapy is (1) the depth of the relationship; (2) the therapeutic frame of a dyadic relationship with a person who combines caregiving attributes and authority attributes; (3) the therapist's ability to facilitate the patient's personality integration and thus to minimize the extent to which the patient's projections distort the internalization process; and (4) (as patients have a right to expect) the therapist's level of mental-emotional health.

In psychodynamic psychotherapy with people who suffer from addictive disorders, internalization can be facilitated by these individuals' tendencies to experience their self-regulatory functions as being located in the external world and to seek them there. The very process that drives their maladaptive externalizing behavior may thus be an important facilitator of therapeutic internalization in psychotherapy.

After a detailed discussion of the neurobiological development of the addictive process, a section on treatment that devotes much more space to talk therapies than to pharmacology may seem to be a non sequitur. However, psychodynamic psychotherapy, DBT, and behavior mastery training are neurobiological as well as psychological processes. Every psychological event is also a neurobiological event—a unitary event that can be known and described in either of two ways, psychological/subjective or neurobiological/objective. An early study reported that fluoxetine and CBT were almost identical in their effects on both symptoms and caudate glucose metabolism in OCD patients. Advances in neuroimaging technology have enabled more recent studies to report more comprehensive and detailed findings on the neural correlates of psychotherapy and of changes that occur in the course of psychotherapy.

Meditation and Yoga. A few decades ago, meditation and yoga were generally regarded as culturally alien fringe activities. Since then they have become almost mainstream. More importantly, they have been scientifically investigated for potentially beneficial effects on the symptoms of a variety of psychiatric and non-psychiatric medical conditions. Thus far, research has not been conducted on the effects of meditation or yoga on sex addiction. However, their well-documented effectiveness as adjuncts in the treatment of other addictive disorders and related psychiatric conditions suggests that they would likely be similarly effective as adjuncts in the treatment of sex addiction. Similar to psychodynamic psychotherapy, the primary goal of yoga and of the meditation described below is self-awareness. All three modalities also share a set of accompanying values: mindfulness, acceptance (of self, of what is), integration, emotional balance, and resilience.

MEDITATION. Published studies have reported meditation to be associated with decreased use of alcohol, cocaine, cannabis, heroin, and tobacco; reduced craving and withdrawal symptoms; prevention of substance use disorders; and decreased binge-eating and emotional eating. It has been reported to be an effective adjunct also in the treatment of depression, anxiety disorders, PTSD, and insomnia. Meditation has been found to correlate with improved self-esteem and increased scores on self-related character scales, high scores on which are associated with a lower risk of personality disorder. It has been reported to improve self-regulation over the life span and especially to improve affect regulation. Research findings indicate that it also enhances executive function, insightful problem solving, cognitive flexibility, and attentional processing. While meditation is most often practiced individually, one study found that group meditation was more effective than was individual meditation in increasing mindfulness and reducing cognitive-behavioral avoidance and impulsivity.

Neuroimaging studies have revealed that meditation improves activation and connectivity in brain areas that are related to self-regulation. It has been reported to increase activation of the ventrolateral PFC, functional connectivity between the PFC and the amygdala, and fractional anisotropy in white matter pathways that connect the ACC to other brain areas. Long-term meditators were found to have larger hippocampal volumes than did non-meditators, particularly in the subiculum, which plays a significant role in stress regulation. In beginning meditators, meditation was found to induce down-regulation of the left amygdala during emotional processing, which indicates a "top-down" form of affect regulation. In experienced meditators, it was found to induce deactivation of default mode network areas (DMN; medial prefrontal and posterior cingulate cortices, precuneus, medial temporal and inferior parietal lobes) and to weaken functional connectivity between them, but not to influence responses in brain regions that are involved in emotional reactivity during emotional processing—which indicates a "bottom-up" form of affect regulation. Meditation has been reported to stimulate GABA release by the thalamic reticular nucleus and thereby to promote GABAergic cortical inhibition, which has been implicated in improved cognitive performance and enhanced emotional regulation.

Meditation is practiced in a variety of forms. The majority involve an intentional focus of attention—for example, one's own breathing, one's own movements (e.g., walking meditation), a word or idea, a sound of internal or external origin, a particular sensation. Patients who have practiced some form of meditation and found it to be helpful are encouraged to continue or resume it. Others can be taught the form of meditation that the Buddha practiced: *Anapana-sati* Yoga, the yoga of breath awareness. Practitioners begin by finding a comfortable upright seated position, usually with their legs crossed. Once settled, they try not to move at all during the entire meditation period (usually 20 to 60 minutes) other than to breathe. The suggested focus of attention is the sensation of their own breath coming in and going out through their noses—more specifically, what the skin between their upper lip and nostrils feels as each inhale or exhale brushes across it. Thoughts and emotions are not banished. When they arise, meditators notice but do not name or describe them, in a manner analogous to noticing clouds that pass across the sky while one's attention is focused on something closer. When attention wanders to such thoughts or emotions, meditators gently bring it back to their breath. This form of meditation is suggested here because it is simple, direct, and unencumbered. Moreover, it facilitates an experience in which the mind can let go of naming or holding on to what is noticed and remain in the immediacy of the moment.

YOGA. Research findings that concern the effects of yoga on psychiatric

disorders and mental functioning are remarkably similar to those that concern the effects of meditation. Yoga has been found to decrease the use of alcohol, tobacco, and cannabis; reduce the risk of developing a substance use disorder; reduce the use of opioids for chronic pain; decrease binge-eating; and decrease eating disorder symptoms in a mixed cohort with diagnoses of anorexia, bulimia, or eating disorder NOS. A review article concluded that as an adjunct in the treatment of eating disorders, yoga can be an effective method for increasing self-awareness, reflection, and the ability to self-soothe. It has been reported also to be an effective adjunct in the treatment of anxiety, depression, obsessive-compulsive disorder (OCD), PTSD, and ADD, and to improve quality of life in schizophrenia. Yoga practice has been reported to be associated with increased plasma levels of oxytocin in schizophrenia patients, larger hippocampal volumes, and increased serum levels of BDNF.

That physical exercise in itself is beneficial raises a question about whether the benefits of yoga are due to a particular characteristic of yoga or more generally to the exercise that it provides. One study found that 12 weeks of yoga was associated with greater improvements in mood and anxiety than was a metabolically matched walking exercise. In the yoga group, improved mood and decreased anxiety were positively correlated with thalamic GABA levels. Another study reported that following eight weeks of yoga practice, participants in the yoga intervention group showed significantly greater improvement in performance on the executive function measures of working memory capacity and efficiency of mental set shifting and flexibility compared with their stretching-strengthening counterparts.

Yoga as a whole consists of physical yoga, the active or intentional component, and mental yoga, the body of yogic knowledge and principles. The latter includes philosophy, ethics, science, and spirituality. Physical yoga includes the poses [*asanas*] that are yoga's most visible manifestation. In a yoga practice, movements into and out of poses are guided by the breath–each is initiated with either an inhalation or an exhalation. The sequence of poses and movements promotes relaxation, inner integration, and receptivity as well as improving flexibility, balance, and strength. The developmental goal of a yoga practice is to increase the practitioner's bodily self-awareness, breath awareness, and sense of body–mind integration. Its immediate goal is to bring the practitioner to a body-mind state that is conducive to meditation. (The origin of meditation is often ascribed to Buddhism. However, Gautama Buddha was a yogi, not a Buddhist. Prior to his awakening [*bodhi*], he had practiced intensively for years under two consecutive hermit teachers of yogic meditation.)

The form of meditation that was described above can be learned and practiced without much instruction. Some people might find even the foregoing brief description of the process sufficient as a starting point.

Learning and practicing yoga require considerably more instruction. Classes and individual teaching are widely available. In addition, considerable information about yoga can be obtained from books or online, and audio-visual instruction is available online and via DVD.

EFFECTS ON THE ADDICTIVE PROCESS. The documented effectiveness of meditation and (separately) yoga in decreasing the use of a variety of psychoactive substances and the frequency of eating disorder symptoms (primarily of binge eating) suggests that both of these adjunctive treatment modalities are effecting positive change at the level of the addictive process. Reported changes that address addictive process-related impairments include improved self-regulation (meditation), increased self-awareness and ability to self-soothe (yoga), and improved affect regulation (both). Both meditation and yoga also have been associated with improved executive function, which would support healing of impairments in self-care and self-governance. The beneficial effect of yoga as an adjunct to the treatment of ADD and OCD suggest that it may improve behavioral inhibition.

Development of a Healthier Lifestyle. A lifestyle is constituted by a profusion of elements, each of which varies along a number of dimensions. Among the more prominent elements are nutrition, sleep, exercise, love relationships, social relationships, work (activity that is motivated by anticipation of its wanted consequences), play (activity that is motivated by enjoyment or pleasure that is inherent to the activity), and creative activity. Many would add spirituality. To illustrate how the term *dimension* is being used here, examples in the area of nutrition include caloric intake, protein intake, frequency of meals, and eliminated foods such as gluten or (for vegetarians and vegans) meat. The most relevant dimensions and the optimal values for each dimension vary from person to person. Complex lifestyle elements concern balances between simple elements, such as the balance between work and play or between social time and solitary time. What constitutes being in balance also tends to vary from one person to another.

People can begin to work on reducing stress in their lifestyles by scanning a span of time for activities, situations, and people that are associated with increased levels of stress. Once a "stressogen" has been identified, they can decide whether to eliminate it, modify it, or leave it as is (when according to their own calculations its benefits outweigh its costs).

Sex addicts' symptomatic behavior has continued because it was reinforced by its consequences, both gratification and alleviation or evasion of painful affects. Their ongoing capacity to refrain from such behavior is enhanced when their lifestyles include nonharmful activities that produce similarly reinforcing consequences or strongly reinforced activities that

preclude symptomatic behavior (or both).

Developing a healthier lifestyle requires more than information about changes that would promote physical or mental-emotional health. It requires deciding to act, taking action, and continuing the action. When this sequence breaks down, the interference may have originated in either maladaptive thoughts and ways of thinking or emotional dynamics that were stirred up by something related to the potential change. Under such circumstances, a brief cognitive-behavioral or psychodynamic psychotherapeutic intervention (respectively) may help patients resume their paths toward more health and balance in their lives.

Prognosis

To date, no outcome studies have been published that evaluate the effectiveness of integrated treatment for sex addiction. Empirical research on almost every aspect of sex addiction is lacking: neurobiology, psychometrics, family history, diagnostic criteria (reliability, coverage, and predictive validity), course of illness, and response to treatments. This deficit may have been due to the unavailability until recently of clear and meaningful diagnostic criteria for sex addiction. It also may have been due to reluctance by many to consider sex addiction as a fit subject for scientific study. Perhaps this chapter and similar publications will stimulate the empirical research that this new field so desperately needs.

Although the long-term response of sex addiction to various types of treatment cannot at this time be reliably predicted, a number of factors have been identified that seem to influence the prognosis. These factors can be clustered in three groups: (1) illness factors (that tend to worsen the prognosis), which include early age of onset, high frequency of symptomatic sexual behavior, concomitant use of alcohol or other drugs, and absence of anxiety or guilt about the behavior; (2) healing support factors (that tend to improve the prognosis), which include a stable job, a stable primary relationship, a supportive social network, and availability of appropriate sexual outlets; and (3) personality factors (that tend to improve the prognosis), which include intelligence, creativity, self-observatory capacity, sense of humor, capacity to form interpersonal connections, and motivation for change—which is a function of people's ability to experience themselves as responsible as well as their degree of subjective distress. Prognosis is also influenced by comorbidity with other psychiatric disorders, by the degree of associated character pathology, and by the kind of relationship that develops between an addict and his or her therapist (or whoever else is most involved in the addict's treatment).

In general, treatment for sex addiction is likely to be most effective when the number of different clinicians (and hence of opportunities for splitting) is kept to a minimum. When the psychodynamic psychotherapist is also the clinician who conducts DBT or symptomatic behavior

management and (when indicated) prescribes psychiatric medication, patients' psychodynamically based (conscious or unconscious) undermining of the latter modalities is more likely to be effectively managed. At another level, assigning different therapeutic modalities to different clinicians is inconsistent with the integrative intent of the treatment.

Significant healing from sex addiction, as from any other addictive disorder, takes time. Addiction interferes with normal development and corrodes interpersonal relationships. The resulting problems remain after symptomatic behavior has ceased and only then can they begin to be effectively addressed. More fundamentally, as described in the preceding section, the principal constituent of healing from an addictive disorder is healing of the addictive process. The initial phase of this healing, which in addition to the overall goals of treatment focuses specifically on affect regulation and stabilization of sense of self, is estimated to require 3 to 5 years.

Healing is not cure. Sex addiction—like other addictive disorders, most psychiatric conditions, and the majority of diseases in medicine—is not an acute disease that effective treatment can cure but a chronic disease that effective treatment can, with patients' active collaboration, bring into remission. And remission in chronic disease is a provisional condition the continuation of which depends on patients' maintenance of a healthy lifestyle. However much addicts (of any behavioral type) may have benefited from treatment, most will to some extent remain vulnerable to being overwhelmed by intense affects and loss of self-coherence in situations that would be neither traumatic nor disorganizing for individuals whose self-regulation systems are intact. Sex addicts are particularly likely to experience threats of being overwhelmed by these internal states as urges to engage in sexual behavior. Even after stable remission has been achieved, ongoing health-promoting activity may be required indefinitely to prevent recurrence of symptomatic behavior and preserve the progress that has been made. Healing does not make addicts' underlying pain disappear. It helps them become increasingly able to accept and manage their pain while freeing them to more fully and joyfully live their lives.

References

Champagne DL, Bagot RC, van Hasselt F, et al. Maternal care and hippocampal plasticity: evidence for experience-dependent structural plasticity, altered synaptic functioning, and differential responsiveness to glucocorticoids and stress. *J Neurosci*. 2008;28:6037.

Coen SJ. Sexualization as a predominant mode of defense. *J Am Psychoanal Assoc*. 1981;29:893.

Feldstein Ewing SW, Chung T. Neuroimaging mechanisms of change in psychotherapy for addictive behaviors: emerging translational approaches that bridge biology and behavior. *Psychol Addict Behav*. 2013;27:329.

Frascella J, Potenza MN, Brown LL, Childress AR. Shared brain vulnerabilities open the way for

nonsubstance addictions: carving addiction at a new joint? *Ann N Y Acad Sci*. 2010;1187:294–315.

Garcia FD, Thibaut F. Sexual addictions. *Am J Drug Alcohol Abuse*. 2010;36:254–260.

*Goodman A. Organic unity theory: An integrative mind-body theory for psychiatry. *Theor Med*. 1997;18:357.

*Goodman A. *Sexual Addiction: An Integrated Approach*. Madison, CT: International Universities Press; 1998.

*Goodman A. Neurobiology of addiction: an integrative review. *Biochem Pharmacol*. 2008;75:266.

*Grant JE, Potenza MN, Weinstein A, Gorelick DA. Introduction to behavioral addictions. *Am J Drug Alcohol Abuse*. 2010;36:233.

Guay DR. Drug treatment of paraphilic and nonparaphilic sexual disorders. *Clin Ther*. 2009;31:1.

Gutman DA, Nemeroff CB. Persistent central nervous system effects of an adverse early environment: clinical and preclinical studies. *Physiol Behav*. 2003;79:471.

Harned MS, Banawan SF, Lynch TR. Dialectical behavior therapy: an emotion-focused treatment for borderline personality disorder. *J Contemp Psychother*. 2006;36:67.

Hyman SE, Malenka RC, Nestler EJ. Neural mechanisms of addiction: the role of reward-related learning and memory. *Ann Rev Neurosci*. 2006;29:565.

Kendler KS, Jacobson KC, Prescott CA, Neale MC. Specificity of genetic and environmental risk factors for use and abuse/dependence of cannabis, cocaine, hallucinogens, sedatives, stimulants, and opiates in male twins. *Am J Psychiatry*. 2003;160:687.

Khan MMR. *Alienation in Perversions*. London: Hogarth; 1979.

Khantzian EJ. Self-regulation and self-medication factors in alcoholism and the addictions: similarities and differences. *Rec Dev Alcohol*. 1990;8:255.

Koob GF, LeMoal M. *Neurobiology of Addiction*. London: Academic Press; 2006.

Marchand WR. Neural mechanisms of mindfulness and meditation: evidence from neuroimaging studies. *World J Radiol*. 2014;6:471.

Messina I, Sambin M, Palmieri A, Viviani R. Neural correlates of psychotherapy in anxiety and depression: a meta-analysis. *PLoS One*. 2013;8 e74657.

Moffett MC, Vicentic A, Kozel M, Plotsky P, Francis DD. Maternal separation alters drug intake patterns in adulthood in rats. *Biochem Pharmacol*. 2007;73:321.

*Myers W. Sexual addiction. In: Dowling S, ed. *The Psychology and Treatment of Addictive Behavior*. Madison, CT: International Universities Press; 1995.

Nestler EJ. Is there a common molecular pathway for addiction? *Nature Neurosci*. 2005;8:1445.

Reid RC, Carpenter BN, Spackman M, Willes DL. Alexithymia, emotional instability, and vulnerability to stress proneness in patients seeking help for hypersexual behavior. *J Sex Marital Ther*. 2008;34:133.

Scharfman MA, Vogel SA. Current concepts of the development of sexuality. *J Amer Psychoanal Assn*. 1989;37:787.

Skowronek IB, Mounsey A, Handler L. Clinical inquiry: can yoga reduce symptoms of anxiety and depression? *J Fam Pract*. 2014;63:398.

Stein DJ, Hollander E, Anthony DT, Schneier FR, Fallon BA. Serotonergic medications for sexual

obsessions, sexual addictions, and paraphilias. *J Clin Psychiatry*. 1992;53:267.

Volkow ND, Fowler JS, Wang G-J. The addicted human brain viewed in the light of imaging studies: brain circuits and treatment strategies. *Neuropharmacology*. 2004;47[Suppl 1]:3.

Volkow ND, Wise RA. How can drug addiction help us understand obesity? *Nature Neurosci*. 2005;8:555.

Weaver IGC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S. Epigenetic programming by maternal behavior. *Nature Neurosci*. 2004;7:847.

Zhang TY, Labonté B, Wen XL, Turecki G, Meaney MJ. Epigenetic mechanisms for the early environmental regulation of hippocampal glucocorticoid receptor gene expression in rodents and humans. *Neuropsychopharmacology*. 2013;38:111.

22

Feeding and Eating Disorders

CHRISTINE C. CALL, A.B., EVELYN Attia, M.D., AND B. TIMOTHY WALSH, M.D.

INTRODUCTION

As their name suggests, feeding and eating disorders are characterized by persistent disturbances in feeding and eating behaviors that significantly interfere with the afflicted individual's life, often by negatively and severely impacting physical health or medical status, impairing psychosocial functioning, and causing intense distress. The psychopathology defining these illnesses manifests most saliently in disordered behaviors, but disturbances in thoughts, attitudes, and emotions are also prominent. Because of the medical complications inherent to feeding and eating pathology, these illnesses are serious and often dangerous, making early identification and intervention particularly important, although, unfortunately, it is not uncommon for them to go unrecognized and untreated.

While the syndromes described within the feeding and eating disorders category center around disordered feeding and eating behaviors, they are also each sufficiently different in etiology, clinical presentation, course, prognosis, and treatment needs to warrant distinct diagnostic criteria. The "Feeding and Eating Disorders" chapter of the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5) includes six diagnoses: anorexia nervosa, bulimia nervosa, binge eating disorder (BED), avoidant/restrictive food intake disorder (ARFID), pica, and rumination disorder. As discussed in greater detail later in this chapter, these disorders appear in one cohesive chapter for the first time in DSM-5, as feeding disorders and eating disorders were previously considered separate diagnostic categories. Feeding disorders have received significantly less research attention than eating disorders, but it is hoped that the merging of the two categories will result in further investigation of these syndromes. Because of the very limited data available, this chapter will not discuss aspects of feeding disorders that remain poorly understood, including epidemiology, course, and treatment.

DEFINITIONS

The definitions of feeding and eating disorders provided below are meant to introduce the reader to each syndrome, but diagnostic criteria and

clinical characteristics are discussed in detail in the "Diagnosis and Clinical Features" section of this chapter and outlined in Tables 22–1 through 22–3.

Anorexia Nervosa

Anorexia nervosa is marked by significantly low weight based on developmental trajectory and height. Individuals with this disorder exhibit a persistent and relentless drive for thinness, leading to behaviors that cause weight loss or (particularly in later phases of the illness and for adolescents who are still growing) prevent weight gain. Most prominently, individuals with this illness significantly restrict their food intake, but may also engage in other behaviors that impact weight, including purging (e.g., self-induced vomiting, laxative, diuretic, or enema use) and intense and often compulsive exercise. Individuals with anorexia nervosa demonstrate a strong fear of gaining weight or becoming fat and typically describe body image distortions, believing that they are fat or large even at a low body weight. Weight and shape tend to become essential in self-evaluation, taking over the majority of the individual's self-schema. In addition, individuals with this illness often fail to recognize the medical implications of their current weight status.

Bulimia Nervosa

While sharing certain core features with anorexia nervosa, bulimia nervosa is defined by recurrent episodes of binge eating followed by compensatory behaviors, which might include self-induced vomiting, laxative, diuretic, or enema use, intense exercise, fasting, and medication misuse or abuse. Central to the conceptualization of bulimia nervosa is the definition of binge eating, a term that refers to the consumption of an unusually large amount of food accompanied by a sense of lack of control, or the feeling that one cannot stop eating. Individuals with bulimia nervosa maintain their weight at or above normal range for age and height but, similar to those with anorexia nervosa, exhibit intense preoccupation with shape or weight. They often rely almost entirely on these factors in self-evaluation, which typically leads to harsh self-judgment.

Table 22–1.
DSM-5 Diagnostic Criteria for Anorexia Nervosa

A. Restriction of energy intake relative to requirements, leading to a significantly low body weight in the context of age, sex, developmental trajectory, and physical health. *Significantly low weight* is a weight that is less than minimally normal, or, for children and adolescents, less than that minimally expected.
B. Intense fear of gaining weight or becoming fat, or persistent behavior that interferes with weight gain, even though at a significantly low weight.
C. Disturbance in the way in which one's body weight or shape is experienced, undue influence of

body shape or weight on self-evaluation, or persistent lack of recognition of the seriousness of current low body weight.

Specify whether:

Restricting type: During the last 3 months, the individual has not engaged in recurrent episodes of binge eating or purging behavior (i.e., self-induced vomiting or the misuse of laxatives, diuretics, or enemas). This subtype describes presentations in which weight loss is accomplished primarily through dieting, fasting, and/or excessive exercise.

Binge eating/purging type: During the last 3 months, the individual has engaged in recurrent episodes of binge eating or purging behavior (i.e., self-induced vomiting or the misuse of laxatives, diuretics, or enemas).

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Table 22–2.
DSM-5 Diagnostic Criteria for Bulimia Nervosa

A. Recurrent episodes of binge eating. An episode of binge eating is characterized by both of the following:
   1. Eating, in a discrete period of time (e.g., within any 2-hour period), an amount of food that is definitely larger than most people would eat during a similar period of time and under similar circumstances
   2. A sense of lack of control over eating during the episode (e.g., a feeling that one cannot stop eating or control what or how much one is eating)
B. Recurrent inappropriate compensatory behaviors in order to prevent weight gain, such as self-induced vomiting; misuse of laxatives, diuretics, enemas, or other medications; fasting; or excessive exercise.
C. The binge eating and inappropriate compensatory behaviors both occur, on average, at least once a week for 3 months.
D. Self-evaluation is unduly influenced by body shape and weight.
E. The disturbance does not occur exclusively during episodes of anorexia nervosa.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Binge Eating Disorder

Individuals with BED, like those with bulimia nervosa, experience repeated episodes of binge eating. However, different from the behavioral pattern of bulimia nervosa, these episodes are *not* followed by compensatory behaviors. While individuals with BED may, and often do, attempt some level of dietary restriction outside of binge eating, they do not engage in inappropriate behaviors meant specifically to compensate for an episode of overeating, like purging or fasting. Individuals with this disorder also maintain their weight at or above the normal range for age and height (and, often, weight is in the overweight or obese category).

Table 22–3.

DSM-5 Diagnostic Criteria for Binge Eating Disorder

A. Recurrent episodes of binge eating. An episode of binge eating is characterized by both of the following:
   a. Eating, in a discrete period of time (e.g., within any 2-hour period), an amount of food that is definitely larger than most people would eat during a similar period of time and under similar circumstances
   b. A sense of lack of control over eating during the episode (e.g., a feeling that one cannot stop eating or control what or how much one is eating)
B. The binge eating episodes are associated with three (or more) of the following:
   a. Eating much more rapidly than normal.
   b. Eating until feeling uncomfortably full.
   c. Eating large amounts of food when not feeling physically hungry.
   d. Eating alone because of feeling embarrassed by how much one is eating.
   e. Feeling disgusted with oneself, depressed, or very guilty afterward.
C. Marked distress regarding binge eating is present.
D. The binge eating occurs, on average, at least once a week for 3 months.
E. The binge eating is not associated with the recurrent use of inappropriate compensatory behavior as in bulimia nervosa and does not occur exclusively during the course of bulimia nervosa or anorexia nervosa.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Avoidant/Restrictive Food Intake Disorder

ARFID, referred to in DSM-IV as feeding disorder of infancy or early childhood, is characterized by persistent disturbances in eating behavior, leading to clinically significant consequences (e.g., nutritional deficiencies, psychosocial impairment, etc.). The restrictive eating that marks this disorder may arise for a variety of reasons (e.g., concern or disgust with texture/consistency/smell of specific food items, anxiety around eating due to a history of choking, vomiting, or a pattern of extremely stressful mealtimes). Notably, shape or weight concern must not motivate the restrictive behaviors as it does in anorexia nervosa and bulimia nervosa. Likewise, a diagnosis of ARFID should not be made if the relevant behaviors occur solely in the context of a separate disorder (e.g., development delays), unless the behaviors are severe enough to warrant specific clinical attention. Finally, feeding or eating behaviors that are considered culturally or socially appropriate should not be considered in making a diagnosis of ARFID.

Pica

Individuals with pica repeatedly consume nonfood items (e.g., paper). As with ARFID, this disturbance must occur outside of the context of culturally sanctioned practices or pre-existing medical or psychiatric

conditions that would cause this behavior (e.g., developmental delays) unless the behavior is so impairing that it requires additional, specific clinical attention.

Rumination Disorder

Individuals with rumination disorder repeatedly regurgitate food and then spit it out, re-chew, or re-swallow it. Rumination disorder cannot be diagnosed in the presence of another eating disorder or ARFID, although the behavior of rumination often occurs in the context of these syndromes. Likewise, individuals with other medical or psychiatric conditions should not receive a diagnosis of rumination disorder unless the behaviors cause substantial impairment sufficient to necessitate specialized intervention.

Other Specified and Unspecified Feeding and Eating Disorders

Finally, DSM-5 includes two broad, heterogeneous categories meant to capture clinically significant feeding or eating disturbances that do not clearly fall within one of the diagnostic categories outlined above. Other specified feeding or eating disorder (OSFED) provides descriptions of disorders that are not yet formally recognized due to insufficient research describing and characterizing them, or those that are similar to the formal eating disorder categories without meeting the frequency or duration criteria needed for these diagnoses. For example, purging disorder, in which an individual engages in repeated purging behavior without the binge episodes that mark bulimia nervosa, is briefly outlined within the category OSFED. Also outlined within this category are atypical anorexia nervosa, bulimia nervosa of low frequency and/or limited duration, BED of low frequency and/or limited duration, and night eating syndrome. Unspecified feeding or eating disorder (USFED), on the other hand, exists to capture individuals whose behaviors and symptoms are not adequately described by a formal feeding or eating disorder, or by the examples provided in OSFED.

HISTORY

Although the specific diagnostic criteria describing feeding and eating disorders continue to shift over time as research findings help guide the current understanding of these illnesses, historical accounts of disordered eating date back as far as the 4th century. Some of these early descriptions appear remarkably similar to the currently accepted definitions of feeding and eating disorders. Still, several of the syndromes now classified as feeding and eating disorders were not officially recognized until much more recently, with BED, only formally accepted as an eating disorder in DSM-5 in 2013, being an excellent example.

Anorexia Nervosa

Anorexia nervosa is the oldest recognized feeding or eating disorder, perhaps because the characteristic low weight that defines the illness has made it more easily discernible over history. The disorder was first formally characterized and named in the 1860s and 1870s by Sir William Gull in London, who coined the name anorexia nervosa, and Dr. Charles Laségue in France, who referred to it as anorexia hysterique. Both physicians described cases of individuals presenting at markedly low body weights and engaging in self-starvation behaviors.

Although Sir William Gull and Dr. Charles Laségue were the first to clearly document reports of the illness, instances of anorexia nervosa likely occurred well before the mid-1800s. In 1689, Richard Morton described a young woman with amenorrhea and significant weight-loss, which he ascribed to her mental state. Evidence of self-imposed starvation, often for religious purposes but nonetheless quite severe, date back to the 4th century, when the gnostic sect of Christian men, believing that material objects including food were evil, engaged in strict and intense fasting practices. Similarly, the term "Holy Anorexia" refers to the practices of numerous women during the Middle Ages (many canonized by the Catholic Church as saints in the 14th and 15th centuries) who strove for holiness by denying themselves food, the self-starvation seeming to become self-sustaining and often resulting in premature death. While it is unclear whether the motives behind these accounts of self-starvation were at all related to body image or feelings about shape or weight, it is nonetheless interesting to note that the most prominent behavior in the current conceptualization of anorexia nervosa, severe self-imposed food restriction, has occurred throughout history.

Since Sir William Gull and Dr. Charles Laségue presented the first clear clinical descriptions of the syndrome, the search to understand this pernicious illness has taken a variety of turns. In the late 1800s, anorexia nervosa was temporarily confused with postpartum pituitary necrosis, leaving a lasting, even if disproved, assumption that some primary abnormality of endocrine function was implicated in the illness's origin. At the beginning of the 20th century, early psychoanalytic hypotheses proposed that the illness might develop from fears of oral impregnation and related themes. Starting in the 1950s, Hilda Bruch formulated a more contemporary psychodynamically oriented hypothesis. Finally, in the 1960s, researchers including Gerald Russell, Arthur Crisp, and Pierre Beaumont initiated the first empirical studies of anorexia nervosa, launching the modern day understanding of the syndrome.

Bulimia Nervosa

Bulimia nervosa was formally delineated much later than anorexia nervosa,

but it, too, was likely present well before its first official documentation. Although historical accounts of the behaviors associated with bulimia nervosa are not nearly as common as those of the self-starvation inherent to anorexia nervosa, the culturally sanctioned practice of overeating and vomiting in purgatoriums in Ancient Rome suggests that relevant behaviors, whether culturally condoned or disordered, have existed for some time. Moreover, given the secrecy that surrounds the syndrome and the lack of easily observable physical indications of the illness (as in anorexia nervosa), it would not be surprising if cases went unnoticed for quite some time before the disorder was formally recognized. Even today many cases of bulimia nervosa go unrecognized because the afflicted individuals maintain normal weight throughout the duration of the disturbance and often hide their disordered behaviors out of shame.

In 1979, Gerald Russell gave a name to the behavioral pattern that was being reported with increasing frequency at the time—episodes of binge eating followed by inappropriate compensatory behaviors. At first, bulimia nervosa, as he named it, was viewed as a variant of anorexia nervosa, but it was quickly recognized that, although the disorders overlap in certain ways, the two are distinct. In 1980, bulimia nervosa was included alongside anorexia nervosa in DSM-III.

Binge Eating Disorder

Interestingly, the behavioral pattern that defines BED, recurrent episodes of binge eating without any inappropriate compensatory behaviors, was clearly characterized in 1959 by Albert Stunkard, two decades before Russell described bulimia nervosa. Despite being documented much earlier than bulimia nervosa, BED began to receive attention and was formally recognized much later. In fact, BED did not appear in the DSM until its fourth edition (DSM-IV) in 1994, when it was described in the appendix as an example of an eating disorder not otherwise specified (EDNOS). Significant research focus on the disorder over the following two decades led to its formal recognition as an eating disorder in DSM-5.

Avoidant/Restrictive Food Intake Disorder, Rumination Disorder, and Pica

As described in greater detail in the next section, the predecessor to ARFID along with rumination disorder and pica were categorized as feeding and eating disorders of infancy and early childhood in DSM-IV. Their shift into the DSM-5 “Feeding and Eating Disorders” chapter demonstrates the growing awareness that these disorders can occur throughout the lifespan, not just in infancy and early childhood. The modifications made to ARFID between DSM-IV and DSM-5 are described in the next section.

COMPARATIVE NOSOLOGY

Currently, two major systems are widely utilized in the diagnosis of mental disorders: The DSM, published by the American Psychiatric Association (APA) and considered the standard classification system in the United States, and the International Classification of Diseases (ICD), published by the World Health Organization and used by a significant number of countries worldwide. The fifth edition of the DSM (DSM-5) was released in May of 2013, and the upcoming ICD-11 is due for release in 2018. The current edition of the ICD (ICD-10) overlaps substantially with DSM-IV in terms of the diagnosis of feeding and eating disorders, with some minor differences. It is expected that ICD-11 will feature many modifications to feeding and eating disorders that are quite similar to those that appear in DSM-5. Thus, the following discussion will focus on the changes that occurred between DSM-IV and DSM-5.

Changes from DSM-IV-TR to DSM-5

Published in 2013, nearly 20 years after the initial release of its predecessor, DSM-5 includes several revisions to the DSM-IV conceptualization of feeding and eating disorders. Most significantly, the DSM-5 "Feeding and Eating Disorders" chapter merges two previously distinct categories: Feeding and Eating Disorders of Infancy or Early Childhood (part of the DSM-IV chapter "Disorders Usually First Diagnosed During Infancy, Childhood, or Adolescence," which was removed in DSM-5) and Eating Disorders. The final chapter includes three feeding disorders (ARFID, pica, and rumination disorder), and three eating disorders (anorexia nervosa, bulimia nervosa, and BED). Of note, BED appears as a formally recognized eating disorder for the first time in DSM-5, as it was previously described within DSM-IV's broad, heterogeneous category, EDNOS.

In addition to these larger changes, DSM-5 features several minor clarifications and modifications to the diagnostic criteria of feeding and eating disorders that aim to increase their utility. In particular, the changes are meant to decrease the frequency with which individuals are assigned to the heterogeneous unspecified feeding and eating disorder category. Using DSM-IV diagnostic criteria, over 50 percent of individuals presenting for eating disorder treatment were assigned to this broad category, which lacks the clinical utility of the formally recognized diagnoses. Preliminary epidemiological studies of DSM-5 feeding and eating disorders suggest that the new criteria substantially reduce the reliance on the unspecified category.

Anorexia Nervosa. The core features that characterize anorexia nervosa remain unchanged in DSM-5, but minor clarifications and modifications to the diagnostic criteria were made. Perhaps most notably, the DSM-IV criterion that requires an individual to exhibit amenorrhea for

at least 3 months to qualify for a diagnosis of anorexia nervosa was removed in DSM-5. The current wording now allows women who report regular or irregular menstrual activity to receive a diagnosis. This change came after several research studies found that women who met all other diagnostic criteria for anorexia nervosa but continued to menstruate regularly did not differ significantly from women meeting full DSM-IV criteria for the syndrome in terms of course, prognosis or illness severity.

In addition to the deletion of the amenorrhea criterion, DSM-5 features a clarification to the language describing weight status in anorexia nervosa. This small change aims to eliminate the misunderstanding that "low weight" is defined by one particular body mass index (BMI) or percent ideal body weight (IBW) cutoff. Because assessing weight status is crucial to making a diagnosis of anorexia nervosa, DSM-IV provides 85 percent of IBW as an example of low weight within the diagnostic criteria for anorexia nervosa. Unfortunately, this specific number was often interpreted as the definition rather than an example of low weight, leading individuals who may have qualified for a diagnosis of anorexia nervosa but presented at greater than 85 percent of IBW to instead receive a diagnosis of EDNOS. DSM-5 has therefore removed the reference to 85 percent IBW and has added additional information to the text about determining whether the individual's weight is significantly low by taking into consideration factors including age, developmental trajectory (i.e., where a child or adolescent falls on his or her individual growth curve), weight history, and medical complications.

Finally, the diagnostic criterion describing fear of weight gain was modified so that individuals who do not endorse or articulate this facet of the illness, but nonetheless demonstrate persistent behaviors that interfere with weight gain or maintenance, may receive a diagnosis of anorexia nervosa. This clarification is meant to capture children and adolescents who may not express this fear due to their age and developmental status, or individuals who, for other reasons (including cultural differences), deny fat-phobia.

Overall, research suggests that these changes, particularly the removal of the amenorrhea criterion, decrease the frequency of EDNOS/OSFED cases without leading to an overdiagnosis of anorexia nervosa. That is, the individuals who are reassigned from the DSM-IV unspecified eating disorder category to the DSM-5 anorexia nervosa category seem to clinically resemble those who meet full DSM-IV criteria for the disorder. The major risk to the modifications discussed above is that they have made the diagnostic criteria more flexible, thus requiring a greater reliance on clinical judgment, for example in assessing weight status or determining whether an individual behaviorally demonstrates fear of weight gain.

Bulimia Nervosa. The changes to bulimia nervosa appearing in DSM-5

are minor, but are aimed to reassign individuals classified as EDNOS in DSM-IV into a more appropriate and descriptive category that better captures their illness. The primary modification involves the frequency of binge eating and purging behaviors. To meet DSM-IV criteria for bulimia nervosa, individuals must engage in binge eating and inappropriate compensatory behaviors at least twice per week on average for 3 or more months. Since the publication of DSM-IV, research has shown that individuals who exhibit these behaviors once per week on average do not differ clinically from those who experience them more frequently. Thus, in DSM-5, the threshold frequency of binge eating and compensatory behaviors has been reduced to once per week for a diagnosis of bulimia nervosa. In addition to this change, the subtypes previously used to delineate bulimia nervosa, purging versus nonpurging type, no longer appear in DSM-5 as they seemed to have limited clinical utility.

Binge Eating Disorder. Between DSM-IV and DSM-5, BED received a substantial amount of research focus, ultimately allowing for its shift from an example of EDNOS in DSM-IV to a formally recognized eating disorder in DSM-5. Besides this significant move, the diagnostic criteria describing the syndrome were modified slightly to become more consistent with the DSM-5 conceptualization of bulimia nervosa. To receive a diagnosis of BED in DSM-IV, an individual must exhibit binge eating episodes on 2 or more days per week on average for at least 6 months. In DSM-5, the frequency and duration of the behavior were altered such that an individual may obtain a diagnosis of BED if he or she demonstrates episodes of binge eating at least once per week on average for 3 or more months. These changes, particularly the formal recognition of BED as an eating disorder, are expected to substantially reduce the frequency of EDNOS/OSFED and to more appropriately label clinically significant cases.

ARFID. ARFID represents a reconceptualization and broadening of the DSM-IV syndrome feeding disorder of infancy or early childhood. As with all three DSM-5 feeding disorders, the move from the DSM-IV chapter "Disorders Usually First Diagnosed During Infancy, Childhood, or Adolescence" to the DSM-5 chapter "Feeding and Eating Disorders" demonstrates the growing understanding that individuals at all points in the lifespan may develop ARFID. In DSM-5, an individual of any age may receive a diagnosis of ARFID.

Besides the removal of the age requirement, DSM-5 includes several modifications that further broaden the diagnostic criteria of ARFID's predecessor. The DSM-IV definition of feeding disorder of infancy or early childhood focused considerably on weight as a marker of clinical significance. To meet criteria for the DSM-IV disorder, an individual must demonstrate significantly restrictive eating behaviors that lead to weight

loss or interfere with appropriate weight gain or growth. DSM-5 acknowledges that while weight often provides useful information about the clinical significance of restrictive eating, some individuals may adequately maintain their weight in the presence of behaviors that nonetheless warrant clinical attention. For example, a child engaging in severely restrictive eating may maintain his or her weight through nutritional supplementation. Thus, in DSM-5, an individual may receive a diagnosis of ARFID even outside the context of weight loss or growth impairment provided that the problematic eating behaviors significantly interfere in some other way.

DSM-5 also expands the definition of this disorder by allowing individuals with a comorbid medical condition to receive a diagnosis of ARFID if their eating behavior and subsequent impairment warrant it. In DSM-IV, diagnostic criteria for feeding disorder of infancy or early childhood specify that a general medical condition must not cause the restrictive eating behaviors associated with ARFID. However, because it can be difficult to determine whether a medical condition leads to abnormal eating or vice versa, DSM-5 permits a diagnosis of ARFID in the presence of medical symptoms so long as the eating behavior and associated symptoms require separate clinical attention.

Finally, the DSM-5 criteria for ARFID recognize motivations for the restrictive eating behavior not previously delineated in DSM-IV's definition of feeding disorder of infancy or early childhood. In addition to the reasons for food avoidance or restriction described in DSM-IV, such as staying away from foods with certain sensory characteristics or demonstrating a lack of interest in eating in general, DSM-5 describes concerns with aversive consequences of eating as another explanation of the behaviors observed in ARFID. This concern may stem from experiencing a negative food-related event (e.g., choking or observing someone choke while eating) or from an emotional disturbance surrounding eating (e.g., a history of particularly problematic and difficult mealtimes due to a poor relationship with the caregiver).

Pica and Rumination Disorder. Both pica and rumination disorder, like ARFID, appear in the "Feeding and Eating Disorders" chapter of DSM-5, having moved from the DSM-IV chapter "Disorders Usually First Diagnosed During Infancy, Childhood, or Adolescence." As with ARFID, based on the modifications appearing in DSM-5, an individual of any age may receive a diagnosis of pica or rumination disorder. Besides this change, pica and rumination disorder remain essentially unchanged from DSM-IV to DSM-5.

EPIDEMIOLOGY

Feeding and eating disorders cut across demographic divides, affecting

males and females of all ages, races, and cultural backgrounds, although certain demographic characteristics impart higher risk for each syndrome. With the broadening of diagnostic criteria in DSM-5, more individuals fall into a formal feeding or eating disorder category than before, as a notable portion has been reassigned from the broad heterogeneous other and unspecified categories that previously accounted for over 50 percent of eating disorder cases. Nonetheless, feeding and eating disorders are relatively rare when compared to more common psychiatric conditions, like major depressive disorder. Although a large proportion of the population will, at some point, experience body dissatisfaction and engage in dieting behaviors, a much smaller percentage will struggle with clinically relevant problematic eating and weight-related cognitions and behaviors, and an even smaller subset will meet full criteria for a feeding or eating disorder. Table 22–4 provides an overview of the epidemiology of eating disorders.

Anorexia Nervosa

Although a substantial proportion of those diagnosed with subthreshold anorexia nervosa in DSM-IV now receive a formal diagnosis in DSM-5, combined rates of subthreshold and threshold anorexia nervosa have remained relatively stable since the 1970s. Various studies have, however, documented a significant increase in incidence rates (the number of new cases in the population over a set period of time) in the high-risk group of 15- to 19-year-old females in recent years. It remains unclear if this increase signifies an earlier age of illness onset or earlier recognition and intervention. Across all age groups and genders, a review of recent epidemiological studies found that lifetime prevalence (the proportion of individuals who have the illness at any point in their life) of DSM-5 anorexia nervosa is between 2.4 and 4.3 percent, approximately double the rate of cases diagnosed using DSM-IV criteria. Among girls and women, point prevalence (prevalence at a specific point in time) of DSM-5 anorexia nervosa is between 0.6 (in a study of adolescents) and 0.7 percent (in a study of college students).

Although epidemiological studies indicate that it may occur as many as 10 times more frequently in females, anorexia nervosa does affect males. Recent research suggests that the illness, which continues to be perceived by many as a female disorder, may be even more common in men and boys than previously believed. It may go unrecognized in men for a variety of reasons. In particular, men may be reluctant to seek treatment out of shame, and clinicians may be less likely to recognize the syndrome in male versus female patients. Additionally, while anorexia nervosa has been thought of historically as a disease of white, wealthy women, it impacts individuals of all racial and socioeconomic backgrounds.

Table 22–4.
Epidemiological Features of DSM-5 Eating Disorders

	Anorexia Nervosa	Bulimia Nervosa	Binge Eating Disorder
Approx. Lifetime Prevalence	2–4%	2%	2–6%
Mortality Risk	6× greater than in general population (1 in 5 from suicide)	Reports of elevated morality rates but additional research controlling for lifetime history of anorexia nervosa needed	No clear increased mortality risk, but increased risk of obesity
Gender	Women > men (10:1)	Women > men (10:1)	Women > men but much less gender divided (1.75:1)
Typical Age of Onset	Early-mid Adolescence	Late adolescence-young adulthood	Less clear, but later than anorexia or bulimia nervosa

Bulimia Nervosa

As with anorexia nervosa, rates of bulimia nervosa have increased with changes to diagnostic criteria. However, prior to the recent modifications appearing in DSM-5, studies suggested that the incidence of the illness has decreased in recent years. Using DSM-5 criteria, it is estimated that lifetime prevalence of bulimia nervosa in women is approximately 2 percent, and point prevalence is close to 0.6 percent. Average age of onset also seems to have decreased, although this finding may be an artifact of earlier detection. Like anorexia nervosa, the disorder is more common in women than in men.

Binge Eating Disorder

BED is the most common eating disorder and the least gender-divided of the three. In the United States, lifetime prevalence of BED has been estimated at 3.6 percent for women and 2.1 percent for men, substantially higher than in European countries where lifetime prevalence is closer to 1.9 percent in women and 0.3 percent in men (although these European studies used the more stringent DSM-IV criteria, which likely accounts for some, but not all, of the difference). Rates of BED are particularly high in obese and overweight individuals.

ETIOLOGY

Our current understanding of the specific causes of feeding and eating disorders is limited. Research regarding the etiology of eating disorders suggests that a confluence of factors, including ones genetic, biological, psychological, and environmental in nature, likely contributes to the onset

of these illnesses. However, the precise mechanisms of causation remain elusive. Most studies investigating the etiology of feeding and eating disorders examine individuals who are already afflicted. Yet, the disorders themselves have such a profound impact on biology and psychology that separating cause from effect proves difficult. Moreover, the relatively low frequency of these disorders makes longitudinal studies that examine individuals before and after illness onset challenging, as a very small portion of the original sample will go on to develop a feeding or eating disorder. Despite these obstacles, studies have allowed investigators to identify risk factors and form research-informed theories about the cause of these syndromes.

Current theories posit that individuals with eating disorders, in particular, possess a set of predisposing traits (e.g., biological, genetic, and/or personality vulnerability) which, when triggered by a precipitating event (e.g., the stress of puberty, the decision to go on a diet), result in the illness. Once the illness sets in, several factors work to maintain it (e.g., the social rewards of weight loss, the effects of the starvation state in anorexia nervosa). For example, a young woman with a family history of anorexia nervosa (predisposing factor) may decide to go on a diet (precipitating event) with her friend. While the friend, with no genetic predisposition to the illness, may successfully lose a few pounds before her weight stabilizes, the young woman's initial weight loss may quickly spiral into a state of severe, continued restriction and marked weight loss. Once at a low weight, the obsessional thinking common to the starvation state, along with other factors, may serve to sustain the cycle of restrictive eating (maintaining factor). A discussion of the current understanding of the etiology of feeding and eating disorders, and the gaps within, is detailed below.

Anorexia Nervosa

Genetic and Biological Factors. Although no specific gene has been conclusively implicated in anorexia nervosa, several lines of evidence suggest that genetic vulnerability plays an important role in the development of the illness. Family studies, in particular, have provided useful information about genetic susceptibility, demonstrating that individuals with a family history of anorexia nervosa are much more likely than those with no family history to receive a diagnosis during their lifetime. In fact, the risk of developing the illness is up to 11 times greater in individuals with a first-degree relative who has experienced the illness. Risk is also increased if a relative has a history of a different eating disorder. Twin studies have demonstrated that concordance rates of anorexia nervosa are substantially higher in monozygotic (identical) twins as compared to dizygotic (fraternal) twins, suggesting that something more than family environment is at play in increasing the odds of illness development. Studies of the serotonin transporter offer preliminary

evidence that this gene may interact with environmental stressors to play a role in the development of anorexia, further emphasizing that a convergence of elements is necessary for illness development. Finally, once the illness sets in, the biological and psychological changes to the body that occur in the starvation state, including low mood and increased obsessionality, may help to maintain the illness.

Developmental Factors. Anorexia nervosa tends to develop during adolescence, suggesting that factors particular to this period may put individuals at risk for the syndrome. Adolescence represents a time of increased biological, psychological, and social change. The experience of going through puberty and experiencing changes to body shape and weight may serve as a major stressor for some, triggering or worsening body dissatisfaction and low self-esteem. Research suggests that this trend may be particularly true for individuals who enter puberty earlier than their peers. Weight-related teasing and bullying may become increasingly prevalent at this time, as well, exacerbating the issue. Some studies also suggest that hormonal changes during puberty, specifically in estrogen levels, may play a role in eating disorder development, although this has not been firmly established. Finally, a number of major social and psychological transitions occur throughout adolescence, including identity and role formation, increasing independence from parents, and the initiation of romantic relationships. These stressors and others may work to catalyze the eating disorder.

Psychological Factors. Several psychological factors appear to confer added risk for developing anorexia nervosa. Certain personality traits including high levels of perfectionism, self-discipline, harm-avoidance, and self-criticism are common in individuals with the illness. Individuals with the restricting subtype, in particular, exhibit low impulsivity and are much more likely to delay rewards than individuals without the illness. Cognitive inflexibility is usually prominent, as well. In a subset of individuals, mood and anxiety disorders or symptoms precede the development of the anorexia nervosa. Obsessive compulsive disorder (OCD) and obsessive-compulsive personality traits also appear to serve vulnerability factors.

Environmental and Social Factors. Environmental and social factors small and large, from experiences in the family or school setting to cultural ideals, may play a role in the development of anorexia nervosa. The influence of family functioning style as a potential predisposing factor remains controversial. No single specific family functioning style appears to be either a necessary or sufficient requirement for developing an eating disorder. As in most psychiatric disorders, various family dysfunctional styles appear to act as nonspecific vulnerability factors and also hamper

recovery.

Participation in certain activities in which weight is emphasized may increase the probability of developing anorexia nervosa or another eating disorder. Ballet, gymnastics, modeling, and weight-restricted sports like wrestling or light-weight rowing may lead to preoccupation with body form and unhealthy attempts to control weight. As many of these activities likely select for individuals who are higher in perfectionism, another risk factor for anorexia nervosa, they may be even more likely to contribute to illness onset.

Although it is a common misconception that Western media causes eating disorders, the cultural ideal of thinness, which is certainly perpetuated by the media, may fuel the overvaluation of shape and weight that marks both anorexia nervosa and bulimia nervosa. Indeed, studies in Fiji, where Western media was not available until the 1990s, showed a significant increase in cases of eating disorders following the introduction of Western television shows to the country. Nonetheless, classic anorexia nervosa remains rare, and research has firmly established that media alone does not cause eating disorders.

Bulimia Nervosa

Like anorexia nervosa, there is still much to learn regarding the etiology of bulimia nervosa. This syndrome likely results from the confluence of several factors, as well, many of which overlap with those implicated in the development of anorexia nervosa, but others that differ substantially.

Genetic and Biological Factors. As in anorexia nervosa, genetics likely play a role in the development of bulimia nervosa, as evidenced by twin and family studies. Individuals with a family history of bulimia nervosa, mood disorder, substance abuse/dependence, or obesity are at higher risk for developing the syndrome.

Neurobiological disturbances are also present in individuals with bulimia nervosa and may increase the probability of binge eating. In particular, individuals with the disorder tend to display delayed gastric emptying, enlarged stomach capacity, and reduced secretion of cholecystokinin (CCK), a peptide hormone released by the small intestine that helps to signal satiety during food consumption. Together, disruptions to these neurobiological processes may put individuals at elevated risk for overeating.

Developmental Factors. The same psychological and social stressors of puberty and adolescence that put individuals at risk for anorexia nervosa also confer risk for bulimia nervosa. Research suggests that hormonal changes during this time, especially shifts in estrogen levels, may be an even greater risk factor in bulimia nervosa than in anorexia nervosa.

Psychological Factors. Unlike individuals with anorexia nervosa, those with bulimia nervosa tend to exhibit high levels of novelty seeking and impulsivity. They also tend to display elevated levels of harm avoidance, negative emotionally, and stress reactivity. Afflicted individuals are more likely than others to experience a comorbid substance use disorder and engage in self-harm, leading to speculation that a subset of individuals with bulimia nervosa may have a propensity toward impulsivity across a range of problematic behaviors. Research findings regarding the possibility of multi-impulsivity as a risk factor for bulimia nervosa remain mixed.

Mood and anxiety disturbances appear to serve as risk factors for bulimia nervosa, as well. In a related vein, research suggests that strong negative emotions may play a role in triggering individual binge episodes in individuals with the disorder, as well, thus acting as maintenance factors in the illness.

Just as the starvation state works to maintain anorexia nervosa, the cycle of binge eating and purging often becomes self-sustaining in bulimia nervosa. Individuals with bulimia nervosa tend to restrict their food intake outside of binge episodes, which puts them at increased risk for episodes of overeating. In turn, after experiencing a binge episode and the resulting compensatory behaviors, individuals often renew their commitment to restrictive eating in an attempt to avoid future overeating episodes. Thus, this cycle tends to become self-perpetuating and works to maintain the disorder.

Environmental and Social Factors. Individuals with bulimia nervosa are likely affected by similar environmental and social stressors as those with anorexia nervosa.

Binge Eating Disorder

Even less is known about the etiology of BED than of anorexia or bulimia nervosa. Research thus far has suggested that a history of childhood obesity, mood disorder, and certain negative family dynamics may put individuals at higher risk for developing the illness. In some individuals, dietary restriction appears to play a role in precipitating episodes of binge eating, as in bulimia nervosa. Many individuals with BED, however, seem to experience binge eating outside of the context of dietary restriction. For these individuals, binge episodes may instead be provoked by intense emotions.

DIAGNOSIS AND CLINICAL FEATURES

In DSM-5 nomenclature, feeding and eating disorders exist within a diagnostic hierarchy. An individual may receive a diagnosis of pica in the presence of another feeding or eating disorder, but the remaining five

diagnoses are mutually exclusive and may not be assigned concurrently. In order of precedence, a diagnosis of anorexia nervosa trumps the other feeding and eating disorders, followed by bulimia nervosa, and then BED. A diagnosis of ARFID can be assigned to individuals who do not meet criteria for anorexia nervosa or bulimia nervosa. Finally, rumination disorder cannot be assigned in the presence of any eating disorder (other than pica). Of note, while these diagnoses may, in general, not occur simultaneously, diagnostic crossover is relatively common, with an individual who previously met criteria for one feeding or eating disorder (e.g., anorexia nervosa) later meeting criteria for a different feeding or eating disorder (e.g., bulimia nervosa). Despite diagnostic crossover and certain shared features, feeding and eating disorders differ substantially from one another in terms of symptoms, behaviors, course, and treatment, as described throughout this chapter. See Tables 22–1 through 22–3 for DSM-5 diagnostic criteria.

Anorexia Nervosa

Anorexia nervosa is characterized by three core features, the most salient being significantly low body weight based on age and height. No specific weight-loss standard defines the illness, as it is important to consider several factors in addition to age and height in determining what constitutes significantly low body weight for an individual. These factors include gender, developmental trajectory, weight history, and current physical health. Nonetheless, drawing from CDC recommendations, DSM-5 notes that the usual lower limit of normal body weight in adults is a BMI of 18.5 kg/m^2. Of note, this standard does not apply to children and adolescents who are still undergoing development; for such young people, it is critical to assess their growth trajectory. To meet criteria for anorexia nervosa, low weight and the symptoms described below must be present for at least 3 months.

Per DSM-5 standards, individuals with anorexia nervosa must demonstrate an intense fear of gaining weight or becoming fat. For individuals who do not articulate or explicitly endorse a fear of weight gain, a diagnosis of anorexia nervosa may still be made in the presence of persistent behaviors that interfere with achieving adequate weight. The behavioral component of this criterion is particularly relevant to children and adolescents, whose age and developmental status may lead to difficulties in verbalizing motivations for engaging in eating disordered behaviors. Additionally, this aspect of the criterion allows individuals who may not endorse fear of fat for cultural reasons to nonetheless obtain a diagnosis of anorexia nervosa.

Finally, anorexia nervosa is marked by disturbances in the experience of body shape or weight. This criterion typically manifests in the way in which an individual with anorexia nervosa perceives his or her body, believing

that his or her shape is too large or the number on the scale too high despite being at a significantly low weight. Individuals with anorexia nervosa may also place a disproportionate emphasis on thinness in self-schema, such that body shape or weight becomes among the most important aspects in evaluating themselves. Often individuals with this illness do not recognize the severe and dangerous medical risks associated with their current low body weight, even after family, friends, and treatment providers repeatedly emphasize their concern.

Within anorexia nervosa, two subtypes, restricting and binge eating/purging subtype, exist. Individuals with the restricting subtype keep weight off by limiting their daily food intake, often by counting calories, controlling portion size by measuring or weighing food, avoiding specific food items (e.g., cake, red meat) or classes of foods (e.g., fats, desserts), only eating at certain times of the day or a certain amount of time after the last meal, and following other rigid rules that govern eating behaviors. They may also engage in intense and compulsive exercise to lose or maintain weight.

Individuals with the binge eating/purging subtype typically engage in a similar pattern of restrictive eating and may also exercise to keep weight off. However, in addition to these behaviors, individuals who fall under this subtype exhibit recurrent eating binges and/or purging through self-induced vomiting, laxative, or diuretic use. Individuals with these clinical features differ from those with bulimia nervosa because of their low weight status.

Bulimia Nervosa

Bulimia nervosa is marked by repeated episodes of binge eating followed by compensatory behaviors. Based on DSM-5 criteria, binge eating and compensatory behaviors must occur on average at least once per week for 3 or more months to warrant a diagnosis of bulimia nervosa. As mentioned in the previous section, unlike individuals with anorexia nervosa, those with bulimia nervosa must be at or above normal weight during the occurrence of these behaviors. Finally, individuals with bulimia nervosa place undue influence on shape and weight during self-evaluation.

Binge Eating. Central to the diagnosis of bulimia nervosa is the definition of a binge eating episode. In general, an episode of eating is considered a binge when two factors are present: (1) the individual consumes an objectively large amount of food, that is, an amount that most would deem unusually large taking into account context, and (2) the individual experiences a sense of loss of control, such as feeling that he or she cannot stop or resist eating at some point during the episode.

While this definition provides a useful framework for thinking about an episode of eating, determining whether an amount of food is objectively

large can be difficult. It is useful for the clinician or researcher to ask the patient to describe a typical episode in which eating feels out of control, generate a list of specific foods consumed, and then determine the usual quantity of each food item eaten. Because context may impact whether an amount is considered objectively large, several questions that probe the specific circumstances surrounding the episode should be considered. For example: Was the individual eating alone or with others? If eating with others, how much was everyone else eating? Did the eating occur at a time when overeating might be expected, like at Thanksgiving or a buffet? How much time was spent eating? Did the individual stop eating for an extended period of time during the episode? Table 22–5 provides an overview of questions that may be useful in determining the presence of binge eating and eliciting a typical binge episode.

Although the current conceptualization of bulimia nervosa requires clinicians and researchers to assess whether the individual is regularly consuming unusually large amounts of food, as described above, some controversy exists surrounding the importance of the size of the binge. Some experts believe that the amount of food consumed may be less important than the experience of loss of control over eating in making a diagnosis of bulimia nervosa. Research has yet to determine whether individuals who regularly experience loss of control while consuming a normal or small amount of food differ clinically from those who experience objectively large binge episodes, although some studies suggest that binge size does not correlate to psychopathology or impairment. Nevertheless, DSM-5 requires that individuals have objectively large binge episodes to qualify for a diagnosis of bulimia nervosa.

Table 22–5.
Assessing for Binge Eating

Determining Size of Binge: Is it Objectively Large?

1. Ask about the specific food items typically consumed (e.g., donuts, pizza, chocolate bars)
2. Determine amount of food consumed (e.g., 5 large donuts, 3 slices pizza, 5 fun size chocolate bars)
3. Assess duration of episode, asking about any breaks in eating to determine whether example is better characterized as one or multiple episodes (e.g., 1 hour with no breaks vs. 3 hours with hour-long break in eating)
4. Ask about context:
   a. Was individual alone or with others?
      ► If with others, what were they eating at the time?
   b. Did eating occur in a context in which overeating might be expected (e.g., buffet, holiday gathering)?

Probing for Loss of Control

Did you feel like you were unable to stop eating, even if you wanted?
How strong was the drive to continue eating?
If you were interrupted by an external event, would you be able to stop eating?
Did you feel like you could not stop eating until the episode reached its own, natural end point?

Compensatory Behaviors. Individuals with bulimia nervosa engage in inappropriate behaviors to compensate for binge episodes and to control shape and weight. In clinical samples, these individuals most commonly report self-induced vomiting or laxative use. Other compensatory behaviors may include diuretic or enema use, intense exercise, fasting, and medication abuse or misuse (e.g., abusing thyroid medication or, in type I diabetes, purposefully restricting insulin use to aid in weight loss). In addition to these behaviors, which must be specifically intended to compensate for episodes of overeating, individuals with bulimia nervosa also tend to restrict or limit their food intake outside of binge episodes. This regular restriction likely plays an important role in the cycle of engaging in binge eating and compensatory behaviors, as described in the previous section on etiology. Specifically, strict rules regarding eating often become impossible to maintain, making the individual more vulnerable to binge eating episodes. Binge episodes, in turn, lead to compensatory behaviors and a renewed desire to restrict food intake to prevent future binge eating episodes, although this logic typically fails.

Binge Eating Disorder

Similar to bulimia nervosa, a diagnosis of BED requires that an individual experiences binge eating episodes at least once per week for 3 or more months. Unlike those with bulimia nervosa, individuals with BED do not engage in regular inappropriate compensatory behaviors. Moreover, in addition to the subjective sense of loss of control that marks binge episodes, behavioral evidence of loss of control must characterize these episodes in BED. Specifically, during binge episodes, individuals with BED typically eat much more quickly than usual, continue eating until uncomfortably full, and consume large quantities of food when not hungry. Individuals with BED may also eat alone due to embarrassment about the amount of food that they consume. After an episode, they may feel depressed, disgusted, and/or very guilty. To warrant a diagnosis of BED, the individual must also report extreme distress about the binge episodes.

Avoidant/Restrictive Food Intake Disorder

ARFID manifests as continual disturbances in eating behavior, which result in clinically significant impairment, such as failure to maintain appropriate weight for age and height, significant nutritional deficiencies, reliance on nasogastric feeding or nutritional supplements to maintain developmentally appropriate weight and/or stable nutritional status, or psychosocial impairment (e.g., inability to eat in public). Individuals with this disorder may exhibit restrictive eating patterns for a variety of reasons, including intense discomfort or dislike of food with certain sensory characteristics (e.g., texture, consistency, smell), fear of eating following an

anxiety-provoking, eating-related event (e.g., fear of swallowing after choking or watching someone else choke; fear of eating after an episode of vomiting), or inability to eat normally due to emotional disturbances surrounding eating (e.g., extreme stress surrounding mealtimes due to unpleasant interactions with caregiver). Of note, the reason motivating the restrictive or avoidant food intake must not be weight- or shape-related, as in anorexia or bulimia nervosa, culturally sanctioned, or caused by another disorder (e.g., developmental delay).

Pica

Pica refers to the recurrent eating of nonfood, nonnutritive items by individuals who are at a developmental stage at which this behavior is inappropriate. Individuals who consume nonfood items as part of a culturally sanctioned practice should not receive a diagnosis of pica. Pica most often occurs in individuals with developmental delays or in pregnant women, but when a medical or psychiatric condition such as these is present, a diagnosis should only be made if the behavior is so impairing or harmful that it requires separate clinical attention. As mentioned previously, a diagnosis of pica may be assigned in addition to another feeding or eating disorder.

Rumination Disorder

In rumination disorder, individuals repeatedly regurgitate the food that they have consumed before spitting it out, re-chewing, or re-swallowing it. Although the behavior of rumination may occur in the context of other eating disorders, specifically anorexia or bulimia nervosa, a diagnosis of rumination disorder cannot be given in addition to these illnesses. In other words, to warrant a diagnosis of rumination disorder, the individual must not meet criteria for another feeding or eating disorder (other than pica).

PATHOLOGY AND LABORATORY EXAMINATION

Medical complications are common and often quite serious in both feeding and eating disorders and require close monitoring by the clinical team. Depending on the nature of the disorder, pathophysiological disturbances may result from starvation and malnutrition, the methods utilized to lose or maintain weight (e.g., restrictive eating, over-exercising, self-induced vomiting, laxatives, diet pills, diuretics, or other medication abuse), binge eating, and (in pica) eating of nonfood items. Mortality rates are elevated in both anorexia nervosa and bulimia nervosa due in large part to medical consequences of the disordered behaviors, making medical monitoring particularly important.

When a clinician suspects the presence of a feeding or eating disorder, a full physical examination should be conducted, including height,

postvoiding weight in a hospital gown (observing for hidden weights), vital signs (with particular attention to heart rate and orthostatic blood pressure changes), a complete examination of skin, muscle, and subcutaneous fat (noting the degree of starvation), and a neurological examination (which is, in general, completely normal). Laboratory tests are helpful for assessing the severity of feeding and eating disorders and their medical consequences, although it should be noted that, in many cases, laboratory tests are normal even in the presence of a serious feeding or eating disorder. Abnormal laboratory findings may also be helpful in enhancing the patient's motivation, particularly for individuals who minimize their disorder and its medical implications.

While many of the medical consequences associated with feeding and eating disorders are not imminently dangerous, some require urgent medical evaluation, such as a prolonged QT or QTc (rate corrected) interval on electrocardiogram (ECG) or marked hypokalemia. Medical specialist consultation should be regularly utilized in the management of patients with such abnormalities. Table 22–6 lists the suggested laboratory tests for patients with anorexia nervosa or bulimia nervosa, with judgment regarding the full extent of the workup for individual patients based on weight loss, severity of illness, subtype, and comorbidity. For example, for individuals with anorexia nervosa binge eating/purging type and bulimia nervosa, laboratory tests often include determination of electrolytes and serum amylase, which are more likely to be abnormal in those who purge than in the restricting type of anorexia nervosa. Table 22–7 summarizes common medical consequences of anorexia nervosa, bulimia nervosa, and BED. The following sections provide additional details regarding pathophysiology that may be specific to each eating disorder.

Table 22–6.
Suggested Laboratory Workup for Patients with Anorexia Nervosa and Bulimia Nervosa

For all patients:

- Complete blood count
- Electrolytes
- Blood urea nitrogen, creatinine
- Thyroid-stimulating hormone, free thyroxine
- Total protein and prealbumin
- Fasting glucose
- Amylase if purging occurs
- Serum phosphate
- Electrocardiogram

If underweight or experiencing amenorrhea:

- Bone mineral density (dual energy x-ray absorptiometry)

Anorexia Nervosa

Patients with anorexia nervosa may present with certain physiological abnormalities that are readily apparent even without an extensive evaluation. These may include significantly low body weight, lanugo (soft, fine hair that grows on the body in the starvation state), hair loss or thinning, and carotenemia (yellowing of the skin). Patients may also present with medical complaints that are often confirmed by further examination. They may, for example, report constantly feeling cold and, indeed, their body temperature is often lower than normal, secondary to hypometabolic activity associated with starvation. Reports of weakness may be due in part to muscle atrophy. Many patients report lightheadedness or fainting and are often orthostatic by vital signs, in addition to exhibiting bradycardia. Reports of gastrointestinal discomfort may be related to delayed gastric emptying and constipation, which commonly accompany anorexia nervosa.

Many malnourished patients with anorexia nervosa develop the euthyroid sick syndrome, in which decreased levels of thyroid-stimulating hormone (TSH), total thyroxine (T_4), and total tri-iodothyronine (T_3) may be seen, free T_4 and free T_3 are usually unchanged, and reverse T_3 (rT_3) is elevated. Generally, these endocrine abnormalities represent energy-conserving consequences of starvation that improve spontaneously with refeeding. They do not require immediate treatment but should be followed to determine if improvement occurs with refeeding.

The hypothalamic–pituitary–gonadal axis is virtually always affected in anorexia nervosa. Low estrogen and progesterone levels are common in women, and low testosterone levels are common in men. Additionally, low levels of luteinizing hormone (LH) and follicle-stimulating hormone (FSH) are seen. Patients with anorexia nervosa have low levels of serum leptin that correlate with low body weight and low body fat, and mediate changes in the hypothalamic–pituitary–gonadal axis.

Table 22–7.
Common Medical and Physical Complications Associated with Eating Disorders

	Anorexia Nervosa	Bulimia Nervosa	Binge Eating Disorder
General	► Low weight ► Loss of body fat ► Muscle Atrophy ► Hypothermia	► Weight within normal-overweight/obese range	► Weight often in overweight/obese range (with associated complications not listed here)
Skin	► Lanugo (fine hair growth on body) ► Hair loss or thinning ► Carotenemia	► Russell's sign: Calluses on hands from teeth scraping during self-induced vomiting (not	

	(yellowing of skin)	extremely common)
Cardiovascular	► Arrhythmias ► Bradycardia ► Hypotension ► Orthostasis ► Prolonged QTc ► Peripheral edema	► Prolonged QTc ► Myocardial toxicity from emetine (ipecac)
Gastrointestinal	► Delayed gastric emptying ► Constipation	► Gastric distention ► Swollen parotid and salivary glands ► Elevated serum amylase levels ► Erosion of dental enamel, multiple caries
Electrolyte/Metabolic	► Hypokalemia (low potassium) ► Hyponatremia (low sodium) ► Hypochloremia (low chloride) ► Alkalosis ► Hypercholesterolemia (high cholesterol) ► Hypoglycemia (low blood sugar) ► Elevated liver enzymes	► Hypokalemia (low potassium) ► Hyponatremia (low sodium) ► Hypochloremia (low chloride) ► Alkalosis
Endocrine	► Low luteinizing hormone ► Low follicle-stimulating hormone ► Low estrogen or testosterone ► Low thyroxine ► Elevated cortisol ► Amenorrhea ► Decreased bone mineral density (can lead to osteopenia or osteoporosis)	► Oligomenorrhea or amenorrhea
Hematological	► Leukopenia ► Anemia	

Related to these abnormalities, bone mineral density is typically low and patients are at risk for osteopenia and osteoporosis, likely due to a combination of low levels of estrogen, high levels of cortisol, and malnutrition. Dual-energy x-ray absorptiometry (DEXA) scans can be helpful in alerting patients to the realistic medical consequences of their disorders. The mechanisms of osteopenia and osteoporosis in anorexia

nervosa are more complex than in postmenopausal osteoporosis, and treatments useful in postmenopausal women do not necessarily help in anorexia nervosa. Orally administered estrogen, calcium, and vitamin D supplementation in anorexia nervosa have not been shown to offer benefit in bone mineral density while the patient is still low in weight. While almost all other medical consequences of anorexia nervosa resolve with weight restoration, bone health may not fully recover.

The structural and functional neuropathology of anorexia nervosa is increasingly well documented by magnetic resonance imaging (MRI), computed tomography (CT), fMRI, positron emission tomography (PET), and single photon emission CT (SPECT) scans. Decreases in total brain volume and sulcal complexity and increases in brain ventricular size are usually seen. The neuropathology of anorexia nervosa reverses substantially with refeeding and weight gain. Approximately 40 percent of typical anorexia nervosa patients score in an abnormal range on two or more neuropsychological tests in standard batteries, suggesting that these patients may have difficulties benefiting from cognitive-behavioral therapies and other psychotherapies in these states. The extent to which long-term improvement in neuropsychological abnormalities may occur in proportion to weight restoration and cessation of binge/purge activity over a prolonged period of stability is not yet clear, and persistent abnormalities in neuropsychological performance following nutritional rehabilitation have been associated with poor prognosis. In usual clinical practice, neuroimaging is not necessary, but in patients with atypical features, especially if other findings suggest that a brain lesion may be present, imaging studies may be of value.

Bulimia Nervosa

Individuals with bulimia nervosa are mainly at risk for medical consequences related to inappropriate compensatory behaviors. Self-induced vomiting, laxative or diuretic abuse, and other compensatory behaviors can cause electrolyte disturbances with serious health risks. Of note, individuals with the binge–purge subtype of anorexia nervosa are also at risk for many of the medical sequelae discussed below.

Hypokalemia (low serum potassium), a particularly dangerous electrolyte imbalance that can cause cardiac arrhythmias, commonly occurs in individuals with recurrent vomiting. Thus, potassium should be regularly monitored, especially in individuals with self-reported purging. Likewise, ECGs are recommended to assess for arrhythmias and for prolongation in QT and QTc, which may even occur in individuals without electrolyte abnormalities. Hypomagnesemia (low magnesium) may co-occur with hypokalemia and must also be addressed as it inhibits prolonged normalization of potassium.

Hyponatremia (low serum sodium) also commonly occurs in patients

with bulimia nervosa (and often in anorexia nervosa) due either to increased water intake, which dilutes sodium concentration in the blood, or to the development of a syndrome of inappropriate antidiuretic hormone release (SIADH). Individuals with eating disorders may consume excessive amounts of water to aid in appetite suppression and subsequent food restriction or, in the context of treatment that involves weigh-ins, to falsely inflate weight ("water loading"). Hyponatremia increases seizure risk and thus must be addressed quickly. Regardless of the etiology of the hyponatremia, fluid restriction is the treatment of choice. When fluid restriction does not sufficiently improve sodium levels, intranasal arginine vasopressin (ddAVP) should be considered.

As in anorexia nervosa, many individuals with bulimia nervosa may present with dehydration (often due to fluid loss from inappropriate compensatory behaviors). Dehydration may result in tachycardia, orthostatic hypotension, and laboratory abnormalities related to prerenal azotemia, including elevated creatinine and blood urea nitrogen (BUN) levels. Although these issues typically resolve with the normalization of food and fluid intake, intravenous hydration may be necessary in the short term for speedier symptom relief and the reversal of laboratory abnormalities.

Binge Eating Disorder

Obesity and its consequences are the primary health concerns for individuals with BED. Whether obese individuals with BED are at greater risk for medical complications than obese individuals without an eating disorder remains unknown.

DIFFERENTIAL DIAGNOSIS

Differentiating feeding and eating disorders from other disorders is typically not difficult when a clinician is able to assess current symptoms and behaviors and obtain a full history. Eating disorders in particular have been researched extensively and are therefore more easily separated from other syndromes. Because much less research describing the presentation, course, prognosis, and treatment response of feeding disorders exists at the moment, differential diagnosis for these syndromes may prove more complicated.

Anorexia Nervosa

Because of the prominent physical and behavioral symptoms (i.e., significantly low body weight and restrictive eating) that mark anorexia nervosa, this disorder is typically easy to identify. In certain cases, however, differentiating the disorder from other syndromes may prove more difficult and clinical judgment may become key in making a diagnosis.

Differential diagnosis may be particularly difficult for individuals who have reached a low body weight but deny or do not articulate fear of gaining weight or becoming fat. By DSM-5 criteria, these individuals may still warrant a diagnosis of anorexia nervosa provided that they demonstrate behaviors that persistently interfere with weight gain. Children and individuals from certain cultural backgrounds, in particular, may not acknowledge a fear of becoming fat, but may demonstrate behavioral manifestations of the illness. In these cases, especially in the evaluation of children, collateral information from parents and others may assist in creating a clear picture of the behaviors that the afflicted individual engages in, thus elucidating the diagnosis. Obtaining a full history of the weight loss and any related behaviors often helps to clarify whether the individual should receive a diagnosis of anorexia nervosa, as well.

Throughout the evaluation process, particularly when certain aspects of the illness (such as fear of weight gain) are not readily apparent, clinicians should consider whether a different disorder better explains the low weight status. Medical conditions that may interfere with weight gain, including gastrointestinal problems and thyroid conditions, should be included in the differential diagnosis. When considering these syndromes, however, clinicians should bear in mind that medical sequelae, including gastrointestinal and thyroid abnormalities, are common in eating disorders and their presence does not necessarily exclude a diagnosis of anorexia nervosa.

Individuals with certain psychiatric conditions may also present with weight loss, suppressed appetite, or even fear of eating that can be distinguished from anorexia nervosa. Individuals with psychotic symptoms may demonstrate delusions relating to food (e.g., that it has been poisoned) that lead to restricted intake. Mood and anxiety disorders may also include symptoms of weight loss or restrictive eating. For some individuals, a major depressive episode may be accompanied by weight loss and lack of appetite. In certain cases (i.e., when the diagnostic criteria become present), an additional diagnosis of anorexia nervosa may be warranted, but, for many, the weight-related symptoms remain attributable to the episode of depressed mood. Likewise, certain anxiety disorders may lead to weight loss or restrictive behaviors. For example, individuals with a specific phobia of vomit may restrict their eating out of concern that eating could lead to an episode of vomiting, and individuals with obsessive-compulsive disorder may develop compulsions related to eating behavior (e.g., only taking a bite of food every minute, or eating foods in specific order depending on color). Unless they also begin to manifest other characteristics of anorexia nervosa (including low weight and preoccupation with shape and weight), these individuals should not receive a diagnosis of anorexia nervosa. Clinicians should also contemplate a diagnosis of ARFID, which is distinct from

anorexia nervosa in that psychopathology related to shape and weight concern is absent in this illness.

Under most circumstances, distinguishing anorexia nervosa from bulimia nervosa should be relatively straightforward. The two are differentiated largely on the basis of weight status, with low weight individuals qualifying for a diagnosis of anorexia nervosa regardless of whether or not they engage in episodes of binge eating and compensatory behaviors. The distinction may become more complicated when an individual with recurrent binge eating and compensatory behaviors presents at a body weight that borders on low, but is not clearly in the normal or low range (e.g., a BMI of 18.5 kg/m²). In these cases, obtaining a detailed weight history may help determine if the individual has lost weight in the context of the disordered eating behaviors (making a diagnosis of anorexia nervosa more likely), or if the individual has continued to maintain preillness weight throughout the course of the eating disorder (making a diagnosis of bulimia nervosa more likely).

Individuals who have recently lost a significant amount of weight and demonstrate other characteristics of anorexia nervosa, including fear of weight gain, but who remain within a normal weight range may meet criteria for atypical anorexia nervosa, an example of OSFED in DSM-5. Diagnostically, atypical anorexia nervosa essentially mirrors classic anorexia nervosa, with the exception of the low weight criteria. However, little is known about course, prognosis, or effective intervention strategies for individuals with this diagnosis.

Bulimia Nervosa

When able to obtain a report of current eating behaviors, differentiating bulimia nervosa from other disorders should be relatively straightforward. As discussed earlier in the chapter, perhaps the most challenging aspect inherent to the assessment of bulimia nervosa revolves around determining whether an individual experiences objective binge episodes. This factor is key, as individuals who do not experience objective binge episodes should not, according to DSM-5, receive a diagnosis of bulimia nervosa. Determining whether an individual is experiencing objective binge episodes and compensatory behaviors at the threshold required for a diagnosis of bulimia nervosa (one time per week for 3 or more months) may be particularly important in differentiating bulimia nervosa from other eating disorders. Specifically, individuals who engage in regular compensatory behaviors (e.g., self-induced vomiting, laxative, diuretic, or enema use) but do not report recurrent, objective binge eating episodes may fall under the OSFED category, purging disorder. Alternatively, those who engage in both objective binge eating episodes and compensatory behaviors, but do so less than once per week or for fewer than 3 months should be assigned to the OSFED category, bulimia nervosa of low

frequency and/or limited duration. Individuals who report recurrent binge eating without compensatory behaviors should receive a diagnosis of BED. Finally, as mentioned above, bulimia nervosa is distinguished from anorexia nervosa, binge eating/purging type based on weight status.

Some individuals experiencing a major depressive episode may present with episodes of overeating secondary to their mood diagnosis. If, however, their episodes are objectively large, accompanied by a sense of loss of control, followed by compensatory behaviors, and have occurred at or above the required frequency and duration thresholds, an additional diagnosis of bulimia nervosa may be warranted.

Binge Eating Disorder

BED should be relatively easy to recognize based on the patient's self-reported pattern of eating. As described above, BED can be distinguished from bulimia nervosa in that recurrent compensatory behaviors are absent in this diagnosis. Mood should also be assessed as individuals experiencing a major depressive episode may report overeating episodes that could be attributable to the depressive episode, or may warrant a separate diagnosis of BED.

Distinguishing BED from the patterns of eating seen in overweight or obese individuals without an eating disorder may be particularly difficult. Many individuals presenting above the normal weight range may describe consuming large amounts of food on a regular basis. Only a subset of this population, however, will meet full criteria for BED. In overweight or obese individuals with episodes of overeating, it is particularly important to probe for behavioral evidence of lack of control while eating. Individuals with BED typically describe feeling as if they are unable to stop or resist eating, even if interrupted by an external cue, like their phone ringing. Additionally, individuals with BED must demonstrate marked distress over their binge eating.

Avoidant/Restrictive Food Intake Disorder

ARFID should be relatively easy to distinguish from other feeding or eating disorders. While individuals with this disorder may present at low weight and often exhibit eating behaviors that appear quite similar to those seen in anorexia nervosa, individuals with ARFID lack any excessive concern with shape or weight. As mentioned in the discussion of anorexia nervosa, differentiating between these disorders may require particular thought when evaluating children, young adolescents, and individuals from certain cultures who may not explicitly endorse fear of weight gain but who may, nonetheless, demonstrate repeated attempts to avoid weight gain. For this reason, clinicians should not immediately exclude anorexia nervosa from the differential upon hearing an individual deny shape or weight concern. If

evidence of attempts to lose or maintain low weight are present, the individual may warrant a diagnosis of anorexia nervosa rather than ARFID.

Importantly, ARFID should not be diagnosed in the presence of another medical or psychiatric condition that explains the restrictive food intake unless the behaviors or symptoms are severe enough to require separate, specialized clinical attention. If eating behaviors are so restrictive that the individual is at risk for medical or nutritional complications, can only maintain health with the assistance of nutritional supplements or nasogastric feeding, or experiences significant psychosocial impairment specifically due to problems with food intake, a diagnosis of ARFID should be made even in the presence of another pre-existing condition. Pre-existing conditions may include medical issues that impact food intake, like gastrointestinal problems that make eating uncomfortable, or psychiatric illnesses, particularly anxiety disorders (e.g., obsessive compulsive disorder, specific phobia of choking or vomiting), autism spectrum disorders, and developmental delays.

Pica

Pica is the sole feeding or eating disorder that may be diagnosed in addition to another feeding or eating disorder. For this reason, regular consumption of nonfood items should be assessed even when the presence of anorexia nervosa, bulimia nervosa, BED, ARFID, or rumination disorder has already been established. Like ARFID, pica should not be diagnosed in addition to a pre-existing medical or psychiatric condition unless the feeding or eating behaviors cause substantial impairment to require specialized intervention. Because pica as a behavior often occurs in the context of developmental delays or certain medical conditions (e.g., pregnancy), it is particularly important to determine the severity of the symptoms and whether they necessitate additional care.

Rumination Disorder

Although the behavior of rumination sometimes occurs within an episode of another feeding or eating disorder, particularly anorexia nervosa or bulimia nervosa, rumination disorder as a syndrome cannot be diagnosed in the presence of anorexia nervosa, bulimia nervosa, BED, or ARFID. Rumination disorder should also be distinguished from the OSFED category, purging disorder, as the behaviors present in these disorders, while similar to some degree, are fundamentally distinct. Rumination refers to the regurgitation of food that has been consumed. While individuals who engage in this behavior may spit out food, this pattern is distinct from purging through self-induced vomiting.

As with ARFID and pica, rumination disorder should not be diagnosed in addition to another disorder unless the behaviors necessitate specialized

treatment. Individuals with certain medical conditions, particularly gastrointestinal issues like acid reflux, may report the symptom of rumination, but will usually describe that this behavior occurs involuntarily and that it resolves with treatment of the underlying medical condition. Individuals with developmental delays may also engage in rumination, but only warrant a separate diagnosis if clinically indicated based on symptom severity.

COURSE AND PROGNOSIS

Anorexia Nervosa

Anorexia nervosa typically develops in adolescence, although illness onset may occur at almost any point during the lifespan, with increasingly frequent accounts of it preceding puberty. Most often, the disorder begins when a normal or slightly overweight individual (commonly an adolescent girl or young woman) starts a diet with the intention of losing a few pounds. The decision to begin dieting usually signifies some level of pre-existing body dissatisfaction and is often spurred by a recent stressor like a breakup, transition into a new school, or a negative comment about shape or weight made by a peer, family member, coach, or dance instructor. For some individuals, initial weight loss is caused not by a diet but rather by an incident entirely outside of the individual's control, like an acute gastrointestinal condition. Either way, after losing a few pounds, individuals tend to report experiencing positive reinforcement, often in the form of compliments from those around them, that encourage continued weight-loss efforts.

Through this point in the illness trajectory, these individuals may resemble any other girl or woman who loses a few pounds and finds the experience rewarding. However, for individuals vulnerable to developing anorexia nervosa, the illness typically intensifies rather quickly after the initial weight loss. Dietary restriction becomes increasingly severe, with individuals limiting portion size or restricting calories to dangerously low levels, avoiding entire food groups (particularly high-fat foods), and skipping meals. Some individuals also begin to limit fluid intake. Exercise may begin to acquire a compulsive and intense quality, and binge eating and purging (through various means) may emerge. As individuals reach a state of malnourishment and low weight, idiosyncratic ritualistic eating behaviors (e.g., cutting food into small pieces, only eating at set times, concocting strange mixtures of foods, etc.) become prevalent. Anxiety and depressive symptoms worsen, preoccupation with body shape and weight intensifies, and individuals tend to become increasingly withdrawn from friends and family.

The precise course of the illness varies substantially, although certain patterns have emerged in the literature. Follow-up studies of patients with

anorexia nervosa reveal that, at the time of assessment, approximately 30 to 50 percent have achieved full recovery, 10 to 20 percent remain chronically ill, and the remainder have improved but continue to struggle with certain disordered behaviors. The chronically ill group often requires multiple hospitalizations, with a small number of patients being in and out of the hospital for the majority of their lives. Because research on anorexia nervosa tends to examine individuals who have presented for specialized treatment, little is known about the frequency with which individuals experience a brief illness and achieve recovery without eating disorder specific care.

Of particular note, anorexia nervosa has a mortality rate as high as that associated with any psychiatric illness. Compared to the general population, individuals with the illness are up to six times more likely to die. The majority of deaths are attributable to medical complications of low weight and malnourishment but a smaller, yet significant, proportion of deaths (approximately 1 in 5) are due to suicide.

Although additional research is needed to identify predictors of outcome in anorexia nervosa, studies have found that adolescents with a shorter duration of illness tend to have a better prognosis, emphasizing the importance of early detection and intervention. Studies examining inpatients with anorexia nervosa just prior to and following discharge have also identified certain predictors of posttreatment outcome. In particular, individuals who achieve full weight-restoration on an inpatient unit and maintain their weight in the first month after discharge are, perhaps unsurprisingly, more likely to remain at a healthy weight up to a year after treatment. Lower BMI at discharge and weight loss in the first month after treatment, on the other hand, predict poorer long-term outcome. Additionally, individuals who demonstrate an ability to consume a diet that is high in variety and energy density (i.e., a greater concentration of kcal/g) prior to discharge seem to do better after treatment.

Bulimia Nervosa

Bulimia nervosa usually develops slightly later than anorexia nervosa, in the late teens or early 20s. Although a substantial portion of individuals report binge eating as their first behavioral symptom of bulimia nervosa, in just under half of all cases restrictive eating and dieting appear to precede the onset of binge eating. Diagnostic crossover from anorexia nervosa to bulimia nervosa is relatively common and, in some cases, restriction may lead to low weight initially, with individuals passing through an episode of anorexia nervosa prior to the development of the bulimia nervosa. In other cases, individuals will remain at normal weight throughout the duration of the dieting. Either way, for those whose illness begins with restrictive eating, this behavior eventually becomes unsustainable, and the individual begins to experience binge episodes. In some cases, the individual may

experiment with or even regularly engage in purging before the binge episodes emerge, but more commonly purging starts as a way to compensate for food binges. The cycle of restriction, binge eating and purging tend to become self-maintaining, as discussed previously.

In general, individuals with bulimia nervosa are more likely than those with anorexia nervosa to attain full or partial recovery, particularly if they have received treatment for the illness. A review of 79 studies of outcome in bulimia nervosa found that, on average, around 45 percent of individuals achieve full recovery and 23 percent remain chronically ill, with the remainder experiencing some degree of improvement. Notably, research finds that bulimia nervosa often goes undetected by health care providers. Several reports suggest an increased mortality risk in this illness. However, study samples typically include a subgroup with a lifetime history of anorexia nervosa who may falsely elevate the mortality risk specifically associated with bulimia nervosa, therefore limiting the interpretation of these data.

Research has yet to identify any clear predictors of outcome in bulimia nervosa. Studies have examined a variety of potential prognostic factors, including duration of illness, age of onset, illness severity, comorbid diagnoses, and personality characteristics, with mixed results. In individuals who receive treatment for bulimia nervosa, studies have found that rapid symptom reduction predicts better outcome.

Binge Eating Disorder

Compared to anorexia nervosa and bulimia nervosa, less is known about the course of binge eating disorder. In general, the disorder appears to affect an older age group, with patients often presenting for treatment well into adulthood. In most cases, it appears that the episodes of binge eating emerge outside of the context of dietary restriction.

Individuals with BED have a much better prognosis than those with anorexia or bulimia nervosa. A substantial number of afflicted individuals appear to experience spontaneous remission without any treatment. One long-term follow-up study found that over 80 percent of individuals who initially met criteria for BED achieved full recovery 5 years later. The vast majority of individuals in this sample did not receive any treatment over the course of the study. Shorter-term follow-up studies have also found impressive recovery rates, with over 50 percent of the sample experiencing full remission after just 6 months. Nonetheless, rates of obesity and overweight remain high in individuals with BED, putting these individuals at risk for weight-related health complications.

TREATMENT

Treatment of feeding and eating disorders remains challenging, with

empirically supported interventions still lacking for certain populations, like adults with anorexia nervosa, and rates of treatment nonresponse or relapse remaining relatively high even in populations for which evidence-based treatments exist, like bulimia nervosa. Interventions for feeding disorders, in particular, have received very little research attention, although the merging of feeding and eating disorders into one cohesive category in DSM-5 will hopefully prompt further investigation into treatment strategies for these syndromes. Although additional research efforts are needed to identify effective interventions for feeding and eating disorders, the sections below will discuss commonly utilized treatments that represent the current understanding of best clinical practices.

Treatment Goals

Although each feeding and eating disorder is substantially different from the others and therefore responds well to different interventions, certain goals of treatment remain the same across the board. As behavioral disorders, one of the primary goals of treatment for any feeding or eating disorder is to disrupt disordered behaviors, specifically, problematic patterns of feeding or eating. For disorders that put individuals at risk for weight problems (i.e., low weight in anorexia nervosa or ARFID, overweight or obesity in BED and, sometimes, bulimia nervosa), another key goal of treatment is to normalize and stabilize weight. For eating disorders in particular, these two initial goals (normalization of eating behaviors and weight) complement each other, as the normalization of one often promotes the normalization of the other. For example, in anorexia nervosa, weight gain and nutritional rehabilitation often lead to a decrease in ritualistic eating behaviors, and as ritualistic eating behaviors diminish, weight gain and maintenance become more attainable.

A secondary, but often equally important, goal of treatment for eating disorders in particular consists of altering the problematic cognitions, attitudes, and beliefs that may help maintain the disordered behaviors. In anorexia nervosa and bulimia nervosa especially, it is common for many distorted thoughts to decrease in intensity as weight and/or behaviors normalize. In anorexia nervosa, weight stabilization often leads to reduction in obsessive thinking. In bulimia nervosa, breaking the cycle of binge eating, purging, and restriction often allows individuals to recognize the irrational nature of their fears about weight gain. In either case, after disrupting problematic behaviors, additional therapy is typically necessary to work through some of the more deeply ingrained beliefs, such as those that tie weight and shape to self-worth.

Finally, eating disorders have high rates of psychiatric comorbidity and addressing these problems often becomes important after the eating disorder is stabilized (although, in some cases, the comorbidities may take precedence). For many individuals, certain comorbid symptoms, particular

depressive and anxiety symptoms, substantially diminish following the stabilization of the eating disorder. As the state of starvation is associated with depressed mood and higher levels of anxiety, this is particularly true for individuals with anorexia nervosa who restore their weight. Likewise, as discussed in greater detail below, certain treatments (especially the use of SSRIs and other antidepressants) for bulimia nervosa and BED can simultaneously improve eating disorder symptoms and mood. For other individuals, however, comorbidities require additional, separate treatment even after eating disorder remission.

Level of Care

Important to the discussion of treatment of eating disorders, in particular, is the concept of level of care. Level of care refers to the range of treatment settings that exist (in this case, for eating disorders), from less restrictive options like outpatient therapy, to more intensive interventions like residential or hospital-based treatment programs. In choosing an appropriate option for an individual with an eating disorder, illness severity, including medical status, psychosocial impairment, and the degree of behavioral disturbance, should be considered. Treatment setting should assure safety without being any more restrictive than necessary. Because eating disorders are serious psychiatric conditions that may be associated with medical or psychiatric comorbidity, intensive treatment options often become appropriate or even necessary.

Anorexia Nervosa

The egosyntonic nature of anorexia nervosa (i.e., affected individuals often view some of their symptoms as aspects they wish to continue, or as parts of themselves) can make treatment of this syndrome particularly difficult. Individuals are often ambivalent about recovery. While they may identify certain thoughts or behaviors that they would like to resolve (e.g., distressing obsessive thoughts about food, shape or weight; objective or subjective binge eating episodes; weakness and fatigue), or certain areas of their life that might improve without the eating disorder (e.g., relationships, concentration), they typically remain very afraid of and resistant to the idea of gaining weight. This ambivalence can make engaging patients in treatment particularly challenging. Therefore, an initial goal of treatment should be for the clinician to establish a positive, supportive relationship with the patient and to begin to involve him or her in treatment planning. Psychoeducation, including a full discussion of the medical risks associated with low weight, often plays an important role in motivating the patient, particularly in the initial phases of treatment.

Because of the complex nature of the illness and the high frequency of serious medical and psychiatric comorbidities, individuals with anorexia

nervosa typically require a treatment team, or a group of clinicians with different specializations who work together to help the individual make progress. This team may deliver care on an outpatient basis or in a more structured setting, and typically consists of a group of multidisciplinary clinicians, including a therapist and/or a psychiatrist, nutritionist, and general medical doctor. In more structured settings, the team often includes additional members, including nursing staff, mental health aids, recreational or occupational therapists, and social workers. For children and adolescents, in particular, parents or other family members may take on an active role in refeeding and weight restoration. Each member plays a different specialized role, but it is imperative that the team regularly communicates regarding treatment goals, progress, and plan.

For adults with anorexia nervosa, no particular treatment has emerged as clearly superior based on empirical studies. For adolescents and children with anorexia nervosa, certain treatments, particularly family-based therapies (discussed below), have demonstrated efficacy in research trials. Despite the limited research support for existing interventions, a variety of treatment types and modalities can promote recovery for individuals with the illness.

Nutritional Rehabilitation and Weight Restoration. Nutritional rehabilitation and weight restoration are primary treatment goals that may be achieved at any level of care, although inpatient hospitalization typically allows for the most efficient refeeding process because of the close supervision with medical monitoring that is available. The following section will discuss weight restoration and nutritional rehabilitation in a structured setting, although many of these principles may be applied to less structured treatment settings, as well.

Structured treatments are indicated for individuals who have not achieved weight gain in outpatient settings, or whose behaviors or symptoms are severe enough to require a higher level of monitoring or care than is possible on an outpatient basis. Individuals presenting with serious medical complications or suicidal ideation or behavior are typically candidates for inpatient treatment. Several options that vary in intensity exist under the umbrella of structured treatment, but all share the primary goal of weight restoration. Most structured settings employ behavioral protocols aimed at normalizing weight through calorie prescriptions, meal plans, structured and supportive meals, disruption of compensatory behaviors (including purging and exercise), and expectations of regular weight gain. A treatment team also provides a range of individual and group therapies, some of which are discussed later in this section.

A privilege or level system is often employed, such that individuals enter the program at the lowest and most restrictive level but can increase their level and subsequently gain privileges and autonomy as they progress

through treatment. Pro-recovery behaviors (e.g., eating 100 percent of food on the meal plan, abstaining from compensatory behaviors, meeting weight gain goals) are rewarded with positive reinforcements, which may mean gaining access to different group activities, receiving less monitoring and supervision, taking a more active role in meal planning, etc. Behaviors that cater to the eating disorder (e.g., restricting food intake, purging or exercising, failing to meet weight gain expectations) are met with negative reinforcements, or a reduction in privileges and an increase in supervision. For individuals who continuously fail to gain weight, activity is restricted to decrease calorie expenditure, and bed rest may be employed.

In most structured facilities, calorie prescription starts between 1,500 and 1,800 kcal/day (at times even higher), consumed across three to four supervised meals and snacks. Macronutrient content is also prescribed, so that patients receive the necessary daily requirements. Because this diet represents an increase in food intake compared to pretreatment consumption for almost all patients with anorexia nervosa, medical status should be monitored closely for signs of refeeding syndrome, a rare but potentially fatal medical complication that involves fluid, electrolyte, and mineral shifts in the body (see Table 22–8 for common signs of this syndrome). Despite the need for careful medical monitoring during this time, a diet of 1,500 to 1,800 kcal/day has been shown to be a safe starting point for patients with anorexia nervosa across various research studies. Additionally, due to the medical complications associated with low weight and starvation, it is typically viewed as safer and more ethical to increase calories (and therefore weight) relatively quickly.

Table 22–8.
Signs of Refeeding Syndrome

Lab abnormalities
Hypophosphatemia
Hypokalemia
Hypomagnesemia
Thiamine deficiency
Glucose intolerance
EKG abnormalities
Dysrhythmias
Torsades de pointes
Physical symptoms
Edema
Respiratory difficulty
Muscle pain and weakness
Gastrointestinal disturbances (e.g., constipation, diarrhea)

Most programs will maintain calories around 1,500 to 1,800 for several days, until the patient is clearly medically stable. At this point, calories are typically increased in a predetermined sequence, often by several hundred

calories every other day, until a diet consisting of approximately 3,500 to 4,000 kcal/day is achieved. To meet these calorie needs, patients may consume high calorie nutritional supplements, like Ensure®, as part of the meal plan. For individuals who refuse to consume solid foods or nutritional supplements, some facilities utilize nasogastric feeding (i.e., tube feeding) to achieve weight gain, although this option is generally viewed as less ideal as it is important for patients to practice normal eating as part of the recovery process.

While calorie prescriptions on structured units are typically quite high, it is important to consider that in order to gain 1 lb, an individual must exceed his or her energy requirements by roughly 3,500 kcal. Thus, the calorie prescriptions described above typically result in gaining between 2 to 5 lb per week. Goal weights are determined at the start of the treatment, but vary across facilities and are often impacted by insurance coverage. Generally speaking, for an individual to achieve full weight restoration, BMI should be increased to greater than 18.5 kg/m^2 or 85 percent IBW, and most clinicians agree that a BMI above 20 kg/m^2 or 90 percent IBW is healthier and more likely to prevent relapse. For children or adolescents, goal weight should be determined based on pediatric growth curves and, for adults, it is often useful to consider preillness weight history to determine if a goal weight substantially higher than a BMI of 18.5 kg/m^2 or 85 percent IBW should be recommended. Despite these clinical goals, health insurance companies may limit reimbursement, covering weight restoration to weights or BMI ranges that are often below ideal.

While weight restoration often resolves many of the depressive and anxiety symptoms associated with the starvation state, additional therapy is almost always indicated after weight has been normalized to prevent relapse, continue to focus on disordered thoughts and behaviors, and work through any underlying issues or comorbidities. Some structured treatments utilize stepped care systems, in which individuals can transition from higher levels of care (e.g., inpatient or residential) to lower levels of care (e.g., day treatment, evening program, outpatient) within the same system as they progress through treatment. Other programs depend on case managers or other staff members to set up discharge plans that aim to help individuals maintain treatment gains. Throughout structured treatment and after, a variety of psychotherapy options are available.

Psychotherapy. Several psychotherapies have been evaluated for the treatment of anorexia nervosa, both in structured and unstructured settings. For adults, no particular psychotherapy has shown clear superiority, although several are regularly used and recommended for clinical management of the illness. Perhaps the most extensively studied and utilized treatment for anorexia nervosa is cognitive-behavioral therapy (CBT), which often focuses on helping the individual alter disordered

behaviors regarding eating, exercise, and other compensatory behaviors, challenge related distorted thoughts, and identify negative core beliefs. Interpersonal psychotherapy (IPT), which focuses on dysfunction in interpersonal relationships, social role, and communication, is also often used for the treatment of anorexia nervosa with similar outcome to that of CBT. While most psychotherapies for anorexia nervosa that have been empirically evaluated are rooted in cognitive-behavioral principles, more recently focal psychodynamic therapy has been examined in a controlled trial, and again, seems to fare similarly to other therapies.

For children and adolescents with anorexia nervosa, family-based therapies have garnered substantial research support. As their name suggests, these therapies focus on involving parents and other family members in the treatment of the eating disorder. The "Maudsley Method," in particular, has demonstrated effectiveness across a number of research studies. In this approach, parents take on the responsibility of refeeding their eating disordered child with the therapist's coaching and support. The parental role often mirrors that of staff members on an inpatient unit who are helping a patient achieve weight restoration. Parents help to enforce meal plans, caloric intake, and weight gain by providing empathic but firm support and supervision, particularly during mealtimes. As the child progresses through treatment, control over eating is ultimately transitioned back to him or her. The final stage of treatment focuses on ways in which the eating disorder has impacted the child or adolescent's development, with the goal of ultimately establishing a healthy adolescent identity.

Psychopharmacology. Although many medications have been studied for the treatment of anorexia nervosa, no clear empirical support has been established for any. The treatment of anorexia nervosa using antidepressants, particularly selective serotonin reuptake inhibitors (SSRIs), has received significant research attention. Nonetheless, studies have consistently demonstrated that, when compared to placebo, antidepressants do not significantly aid in weight gain, alter eating pathology, or reduce depressive or anxiety symptoms in underweight individuals with anorexia nervosa. Rather, weight restoration itself has been shown to reduce symptoms of depression and anxiety. A small number of studies have also examined the efficacy of SSRIs after weight restoration and similarly found no significant impact on relapse prevention or psychological outcome. Despite these findings, some individuals whose depressive symptoms do not diminish with weight restoration may benefit from antidepressant treatment following weight restoration.

Atypical antipsychotic medications have also been studied in the treatment of anorexia nervosa. In normal weight individuals without eating pathology, atypical antipsychotic medication often results in extreme and unwanted weight gain, although the precise mechanism of this change

remains unknown. In patients with anorexia nervosa, one atypical antipsychotic, olanzapine, has shown promise in promoting weight gain and reducing distressing psychological symptoms, particularly the obsessionality that often marks the illness. Of note, treatment studies of olanzapine in patients with anorexia nervosa have demonstrated much more modest weight gains than are observed in noneating disorder psychotic patients taking the medication. Olanzapine for the treatment of anorexia nervosa remains under study, with large placebo-controlled trials currently underway.

Bulimia Nervosa

Interventions with strong empirical support exist for bulimia nervosa, although there is still substantial room for improvement as a significant portion of individuals continues to experience symptoms after receiving evidence-based treatment. Many individuals with bulimia nervosa respond well to and remain medically stable during the course of established outpatient treatments and, thus, inpatient care is indicated far less often than it is for anorexia nervosa.

Psychotherapy. Several psychotherapy approaches for bulimia nervosa have received research attention, with CBT emerging as the treatment of choice for the illness. CBT has demonstrated efficacy in both individual and group formats, with 30 to 50 percent of individuals reporting symptom abstinence following treatment and a higher percentage reporting symptom reduction. CBT for bulimia nervosa has also been adapted into pure and guided self-help formats, with both showing greater effectiveness than a waitlist control condition (although research has yet to determine how CBT self-help formats compare to the standard face-to-face modality).

CBT for bulimia nervosa consists of several phases, each meant to disrupt problematic behaviors (e.g., binge eating, purging, restricting) and identify and restructure distorted thoughts and beliefs that may be maintaining the unhealthy behaviors (e.g., “Eating sugary foods will make me fat”). In general, disordered behaviors are tackled first and, once they have improved, cognitive distortions become the focus of treatment. The initial phase of treatment often involves psychoeducation and working with the patient to increase the regularity of eating and resist urges to binge or purge. Self-monitoring and recording are key strategies in CBT and may rely on pen and paper or utilize web-based applications, which have become increasingly accessible and seemingly useful. In addition to self-monitoring and recording, patients receive various homework assignments to help identify and alter unhelpful behaviors and dysfunctional attitudes and beliefs. Patients are also taught to identify stressors or triggers and to employ more adaptive coping strategies. Finally, after symptoms have abated, relapse prevention strategies are used to reduce the likelihood of a

return of symptoms by anticipating and preparing for stressful situations and setbacks likely to be encountered in the future.

In addition to CBT, several other psychotherapies have emerged as helpful in the treatment of bulimia nervosa. IPT, in particular, has been shown to reduce binge eating and purging, although research suggests that CBT does so more rapidly. Specifically, one study found that, at the end of treatment, a significantly higher percentage of individuals in the CBT group experienced symptom abstinence as compared to the IPT group (45 percent in CBT vs. 8 percent in IPT). However, individuals in the IPT group make additional gains after treatment and, 1 year later, essentially catch up to the CBT group in terms of rate of symptom abstinence (40 percent in CBT vs. 27 percent in IPT). As in anorexia nervosa, IPT for patients with bulimia nervosa focuses on dysfunctional interpersonal relationships.

While they have received less research attention, dialectical behavioral therapy (DBT) and integrative cognitive-affective therapy (ICAT) also show promise for the treatment of bulimia nervosa. Both draw from principles of CBT, but differ in certain ways. DBT focuses on helping patients develop mindfulness, emotion regulation, and distress tolerance in an attempt to regulate distressing emotions that may contribute to maladaptive eating behaviors. ICAT, in addition to focusing on distorted cognitions and problematic behaviors as in CBT, emphasizes the role of dysfunctional interpersonal interactions on the eating disorder. Additional research is needed to further clarify the efficacy of these therapies in the treatment of bulimia nervosa.

Psychopharmacology. Although psychopharmacology currently has no clear role in the treatment of anorexia nervosa, research has identified several medications that effectively reduce symptoms in many patients with bulimia nervosa, particularly when administered in conjunction with evidence-based psychotherapy. Notably, research suggests that a course of drug therapy alone is inferior to a course of CBT alone, but some studies suggests that a combination of the two interventions may be more helpful than either alone. Thus, medication as the sole therapy for bulimia nervosa is typically not the first choice recommendation, but medication as an adjunctive treatment for bulimia nervosa has a strong evidence base.

Antidepressants in particular have amassed substantial research support for the treatment of bulimia nervosa and, of note, are equally effective for patients with and without depressive symptoms. Although several antidepressants appear to significantly reduce binge eating and purging, the SSRI fluoxetine has been studied the most extensively and is viewed as the medication of choice for this illness. Although the medication is traditionally prescribed at a dose of 20 to 40 mg per day for the treatment of depressive disorders, studies have found that higher doses (e.g., 60 mg per day) are more effective in disrupting binge eating and

purging. Even at this higher dose, fluoxetine has fewer side effects than some of the other medications described below that have demonstrated efficacy in treating bulimia nervosa. Research has yet to determine the optimal length of fluoxetine treatment for bulimia nervosa as most trials have been relatively short.

Tricyclic antidepressant medications (TCAs), monoamine oxidase inhibitors (MAOIs), and atypical antiepileptic agents such as topiramate have also demonstrated efficacy in randomized-controlled trials of bulimia nervosa. TCAs and MAOIs, however, are not viewed as a first-line treatment due to significant side effects and the risk of fatal overdose. Across several studies, topiramate has resulted in significant weight loss in patients with bulimia nervosa, a complicated side effect for an eating disorder patient that should be carefully considered prior to treatment initiation. Finally, bupropion is contraindicated for bulimia nervosa due to concerns regarding an increased risk for seizures in patients who are actively purging.

Binge Eating Disorder

Because both disorders are marked by recurrent episodes of binge eating, much of the treatment research for BED has focused on interventions that are known to effectively treat bulimia nervosa. However, the substantial portion of patients with BED who are overweight or obese may require additional treatment to achieve a healthier body weight, although clinicians tend to agree that binge eating behavior should be addressed before weight reduction is attempted. Additional studies examining interventions for BED are currently underway.

Psychotherapy. As in bulimia nervosa, CBT has emerged as a highly effective treatment for BED with over 50 percent of individuals attaining abstinence from binge eating following treatment. CBT for BED is structured quite similarly to CBT for bulimia nervosa, without the focus on inappropriate compensatory behaviors. Despite its efficacy in reducing binge eating in BED, studies have found that CBT does not effectively result in weight loss for overweight or obese individuals.

IPT may be even more useful for BED than it is for bulimia nervosa. In fact, research examining IPT versus CBT for BED suggests that the two interventions may be equally efficacious both at the end of treatment and in follow-up. DBT, too, may be effective in the treatment of BED but additional studies are needed to determine this therapy's impact on this population.

Psychopharmacology. Many of the medications that are useful for bulimia nervosa also seem to reduce binge eating in BED, although long-term studies have yet to examine whether these treatment gains persist

over time. Specifically, SSRIs and TCAs compare favorably to placebo in eliminating binge eating, but do not seem to aid in weight loss. Because it both reduces binge eating and causes weight loss, topiramate appears to serve as an effective treatment for BED, as well. Weight loss medications have not been extensively studied in the treatment of BED, although one, orlistat, seems to promote weight loss but does not reduce binge eating. Other appetite suppressing and weight reducing medications have been considered for this population, including stimulants. Most recently, lisdexamfetamine has been approved by the FDA for short-term treatment of BED. Of note, long-term data are not yet available regarding the effectiveness and safety of this treatment.

Weight Loss Treatment. Treatment of the obesity associated with BED requires both nutritional modification with calorie reduction and an increase in exercise. For severe obesity, especially class III, medical supervision and, on occasion, gastric surgery may be recommended.

ARFID, Pica, and Rumination Disorder

Very little research has investigated effective treatments for ARFID, pica, or rumination disorder. Behavioral strategies that have been identified as useful for other disturbances of eating behavior are often "borrowed" for treatments that aim to normalize behaviors in individuals with ARFID, pica, and rumination disorder.

Case Histories

Anorexia Nervosa. Ms. T, a 20-year-old woman, was admitted to an inpatient eating disorder unit at a height of 5′7″ and a weight of 105 lb (BMI: 16.4 kg/m^2; 75 percent IBW). At the time of her admission, Ms. T described a history of disordered eating starting 2 years earlier, soon after entering college. The transition to college was stressful for Ms. T who felt homesick and struggled to make friends. She ran cross-country throughout high school, maintaining a BMI around 19.5 kg/m^2 until college, when, in the context of discontinuing her exercise regimen and adjusting to buffet meals in the cafeteria, she gained approximately 10 lb. Ms. T experienced significant distress surrounding her weight gain and began eliminating snacks and desserts from her diet, successfully losing 5 lb. She enjoyed compliments from acquaintances regarding her weight loss and decided to resume regular exercise, eventually increasing her running to 5 to 10 miles daily. Her exercise quickly began to feel driven, and she noted feelings of intense guilt and increased dietary restriction when she missed a day. She also began to skip meals and avoided eating in the cafeteria. Within 2 months, she lost 20 lb (BMI: 18.0 kg/m^2; 82 percent IBW), her menstrual periods became irregular, and she struggled with low mood. She became intensely afraid of becoming fat and weighed herself twice daily to ensure that her weight had not increased. At this point, Ms. T was subsisting on two meals per day totaling approximately 500 to 700 kcal and consisting of foods that she deemed "safe" (e.g., fat-free Greek yogurt, grilled chicken, vegetables). She ate at a set time each day and weighed and measured each food item before consuming it, keeping a detailed food log. When her parents became aware of the eating disorder, Ms. T began to see a therapist, nutritionist, and medical doctor on an outpatient basis. She increased her weight by several pounds and her exercise and food rituals loosened. However, the following year, an ankle injury prevented her from running and Ms. T began to restrict her food intake more severely, losing to a BMI of 16.4 kg/m^2 (75 percent IBW). Subsequently, her treatment team recommended inpatient care

with the goal of restoring her weight and becoming reacquainted with healthy eating habits.

ARFID. An 11-year-old K was referred to an eating disorder specialist after her weight fell below the 7th percentile for her age and height. Although K had received treatment for anxiety throughout much of her childhood, she did not manifest worrisome eating until 18 months prior to presentation when she refused to eat during a family vacation because "all the foods tasted weird." When she returned home, she began to refuse food, citing an intense fear of vomiting. K's parents were encouraged to wait out the period of change since children often go through phases of picky eating. Yet, K's eating behaviors became increasingly extreme, and K continued to eat very small amounts. She refused eating in restaurants and avoided going to friend's houses or even eating at school, stating that she was less likely to vomit when at home. When she did eat, K would push her food around on the plate, crying when asked to try more. K had been thin throughout childhood, but her weight had hovered around the 25th percentile prior to age 9. When her weight dropped below the 10th percentile and the pediatrician explained the dangerous consequences associated with low weight, K grew tearful and stated that she wanted to gain weight. When asked more about her food restriction, she spoke of an event that occurred years earlier, when she had eaten a meal, grown nauseous, and vomited. She adamantly denied any concern with her appearance, restating her desire to be at a healthy weight.

References

Agras WS, Walsh BT, Fairburn CG, Wilson GT, Kraemer HC. A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. *Arch Gen Psychiatry*. 2000;57(5):459–466.

Arcelus J, Mitchell AJ, Wales J, Nielsen S. Mortality rates in patients with anorexia nervosa and other eating disorders: a meta-analysis of 36 studies. *Arch Gen Psych*. 2011;68(7):724–731.

*Attia E, Walsh BT. Behavioral management for anorexia nervosa. *NEJM*. 2009;360(5):500–506.

Becker AE, Burwell RA, Herzog DB, Hamburg P, Gilman SE. Eating behaviours and attitudes following prolonged exposure to television among ethnic Fijian adolescent girls. *Br J Psychiatry*. 2002;180(6):509–514.

Call C, Walsh BT, Attia E. From DSM-IV to DSM-5: changes to eating disorder diagnoses. *Curr Opin Psychiatry*. 2013;26(6):532–536.

Culbert KM, Racine SE, Klump KL. Research review: what we have learned about the causes of eating disorders—a synthesis of sociocultural, psychological, and biological research. *J Child Psychol Psychiatry*. 2015;56(11):1141–1164.

Eddy KT, Dorer DJ, Franko DL, Tahilani K, Thompson-Brenner H, Herzog DB. Diagnostic crossover in anorexia nervosa and bulimia nervosa: implications for DSM-V. *Am J Psychiatry*. 2008;165(2):245–250.

Fairburn CG, Doll HA, Welch SL, Hay PJ, Davies BA, O'Connor ME. Risk factors for binge eating disorder: a community-based, case-control study. *Arch General Psychiatry*. 1998;55(5):425–432.

Franko DL, Keshaviah A, Eddy KT, et al. A longitudinal investigation of mortality in anorexia nervosa and bulimia nervosa. *Am J Psychiatry*. 2014;170(8):917–925.

Garber AK, Sawyer SM, Golden NH, et al. A systematic review of approaches to refeeding in patients with anorexia nervosa. *Int J Eat Disord*. 2016;49(3):293–310.

*Hay P. A systematic review of evidence for psychological treatments in eating disorders: 2005–2012. *Int J Eat Disord*. 2013;46(5):462–469.

Hay PP, Claudino AM, Kaio MH. Antidepressants versus psychological treatments and their combination for bulimia nervosa. *Cochrane Database Syst Rev*. 2001;4:1–21.

Heaner MK, Walsh BT. A history of the identification of the characteristic eating disturbances of bulimia nervosa, binge eating disorder and anorexia nervosa. *Appetite*. 2013;65:185–188.

Kaplan AS, Walsh BT, Olmsted M, et al. The slippery slope: Prediction of successful weight maintenance in anorexia nervosa. *Psychol Med*. 2009;39(06):1037–1045.

*Klein DA, Walsh BT. Eating disorders: clinical features and pathophysiology. *Physiol Behav*. 2004;81(2):359–374.

Klump KL. Puberty as a critical risk period for eating disorders: a review of human and animal studies. *Horm Behav*. 2013;64(2):399–410.

Lock J, Le Grange D. *Treatment Manual for Anorexia Nervosa: A Family-Based Approach*. New York: Guilford Press; 2012.

Mehler PS, Crews CK. Refeeding the patient with anorexia nervosa. *Eat Disord*. 2001;9(2):167–171.

Nagl M, Jacobi C, Paul M, Beesdo-Baum K, Höfler M, Lieb R, Wittchen HU. Prevalence, incidence, and natural course of anorexia and bulimia nervosa among adolescents and young adults. *Eur Child Adoles Psy*. 2016;11:1–6.

Pollice C, Kaye WH, Greeno CG, Weltzin TE. Relationship of depression, anxiety, and obsessionality to state of illness in anorexia nervosa. *Int J Eat Disord*. 1997;21(4):367–376.

Schebendach J, Mayer LE, Devlin MJ, Attia E, Walsh BT. Dietary energy density and diet variety as risk factors for relapse in anorexia nervosa: a replication. *Int J Eat Disord*. 2012;45(1):79–84.

*Smink FR, van Hoeken D, Hoek HW. Epidemiology of eating disorders: incidence, prevalence and mortality rates. *Curr Psychiatry Rep*. 2012;14(4):406–414.

Steinhausen HC. The outcome of anorexia nervosa in the 20th century. *Am Psychiatry*. 2002;159(8):1284–1293.

Steinhausen HC, Weber S, Phil C. The outcome of bulimia nervosa: findings from one-quarter century of research. *Am J Psychiatry*. 2009;166(12):1331–1341.

Stefano SC, Bacaltchuk J, Blay SL, Appolinario JC. Antidepressants in short term treatment of binge eating disorder: systematic review and meta-analysis. *Eat Behav*. 2008;9(2):129–166.

Thomas JJ, Eddy KT, Murray HB, et al. The impact of revised DSM-5 criteria on the relative distribution and inter-rater reliability of eating disorder diagnoses in a residential treatment setting. *Psychiat Res*. 2015;229(1):517–523.

Uher R, Rutter M. Classification of feeding and eating disorders: review of evidence and proposals for ICD-11. *World Psychiatry*. 2012;11(2):80–92.

van Elburg A, Treasure J. Advances in the neurobiology of eating disorders. *Curr Opin Psychiatry*. 2013;26(6):556–561.

*Watson HJ, Bulik CM. Update on the treatment of anorexia nervosa: review of clinical trials, practice guidelines and emerging interventions. *Psychol Med*. 2013;43(12):2477–2500.

Yilmaz Z, Hardaway JA, Bulik CM. Genetics and epigenetics of eating disorders. *Adv Genomics Genet*. 2015;5:131.

23

Sleep Disorders

MAX HIRSHKOWITZ, PH.D., AND AMIR SHARAFKHANEH, M.D., PH.D.

To understand sleep and its disorders, it helps to begin with the sleep's three basic characteristics: (1) Sleep is a process required for proper brain function. Failure to sleep impairs thought processes, mood regulation, and a host of normal physiological functions. (2) Sleep is not a single process; there are several distinct types of sleep. These different types of sleep differ both qualitatively and quantitatively. Each type of sleep has unique characteristics, functional importance, and regulatory mechanisms. Selectively depriving one particular type of sleep produces compensatory rebound when an individual is allowed to sleep *ad lib*. (3) Sleep is not a passive process; sleep can be associated with a high degree of brain activation and metabolism.

Several physiological mechanisms regulate sleep and when these systems go awry, sleep disorders occur. Initial interest in sleep disorders emanated from psychiatrists, psychologists, and neurologists. Medical research during the past four decades expanded sleep medicine into a multidisciplinary field. Research illustrating sleep-disordered breathing's medical consequences attracted many pulmonary, internal medicine specialists, and cardiologists to the field. Perioperative risks associated in patients with sleep-disordered breathing attracted the attention of anesthesiologists and treatment options for breathing disorders now involve surgeons and dentists. Sleep–wake cycle-related endocrinology and circadian rhythm research has migrated from the laboratory bench to the bedside. Nonetheless, the seriousness of sleep disorders remains poorly recognized by the general public and many clinical practitioners. Progress, however, is being made. In the past decade, *Sleep Medicine* became a recognized medical specialty and ACGME accredited Sleep Medicine Fellowship Training programs followed.

Sleep disorders are both dangerous and expensive. Hypersomnolence is a serious, potentially life-threatening condition affecting not only the sleepy individual but also their family, coworkers, and society in general. In fact, sleep-related motor vehicle accidents represent a major public safety concern with some states enacting criminal statues to deter sleepy driving. Investigations link many major industrial catastrophes to sleepiness. Research shows that sleep-disordered breathing contributes to

hypertension, heart failure, and stroke. Sleep disorders' direct cost per annum in the United States is estimated at 16 billion dollars with indirect costs ranging upward to more than 100 billion dollars.

SLEEP DISORDER CLASSIFICATIONS

Three different nosologies provide classification systems for sleep disorders: (1) the Diagnostic and Statistical Manual of Mental Disorders (DSM), (2) the International Classification of Sleep Disorders (ICSD), and (3) the International Classification of Diseases (ICD). All three have been updated since the last edition of this textbook; current sleep–wake disorder classifications are the DSM-5, ICSD-3, and ICD-10, respectively.

The American Psychiatric Association's DSM-5 represents the standard diagnostic system used by most physicians who are not Sleep Medicine Specialists. Sleep Medicine specialists often prefer using the ICSD-3 (which is cross-referenced to the required DSM-5 coding). The World Health Organization's ICD-10 is not as popular (in the United States). The three systems organizational schemas differ. Some lump multiple nosologic entities into a single diagnostic classification while others split disorders, providing separate designations. In the present chapter, the DSM-5 is used as an organizational guideline because (1) it represents a diagnostic standard, (2) its coding system is widely used, and (3) it is the parsimonious classification system, describing essential sleep–wake disorders. ICSD-3 is more comprehensive; however, diagnostic differentiation sometimes resembles research diagnostic criteria rather than clinical practice criteria. Where appropriate, ICSD-3 diagnostic entities will be discussed as clinical subtypes. ICSD-3 is the direct descendent of the first comprehensive sleep disorders nosology. In 1979, the Association of Sleep Disorders Centers (ASDC) commissioned a task force to officially codify sleep disorders. Approximately 20 years later, the ASDC had become the American Sleep Disorders Association (ASDA) and the organization sponsored creation of an updated classification system called the International Classification of Sleep Disorders (ICSD-1). After major revision, a second edition (ICSD-2) was published in 2005 (by the organization now known as the American Academy of Sleep Disorders). Recently, the current (third edition) was published in 2014 (see Table 23–1). Nonetheless, for major and common sleep disorders, diagnostic criteria for DSM-5 and ICSD-3 are quite similar (and sometimes identical). Finally, DSM-5 was chosen to organize this chapter because, after all, this is the Comprehensive Textbook of Psychiatry.

INSOMNIA DISORDER

Previously, DSM-IV defined insomnia as difficulty initiating sleep, maintaining sleep, or nonrestorative sleep for 1 month or more. *The overall conceptualization of insomnia has changed dramatically, beginning with*

the dissatisfaction with sleep quantity or quality serving as the primary element. Three other paradigm shifts in thinking also include the following. First, nonrestorative sleep and its consequent sleepiness no longer represent a key feature for insomnia. Nonrestorative sleep is now considered an attribute of hypersomnolence. Second, whereas insomnia was previously considered a symptom, secondary to some other underlying problem, it is now recognized as a possible independent condition. Third, the hitherto insomnias considered secondary are now considered comorbid. In the past, practitioners were admonished to treat insomnia's cause rather than the symptoms. There was an implicit notion that by doing so, the sleep problems would improve. Clinical experience suggested otherwise. Consequently, current therapeutics favor providing relief and managing symptoms. To some extent, this change in thinking parallels getting away from monotherapy. In the past it was argued that if insomnia was related to depression, treating the insomnia would mask the depression and thereby interferes with antidepressant treatment regimens. This does not appear to happen.

Table 23–1.
Outline of the *International Classification of Sleep Disorders*, Third Edition

I. Insomnia
   A. Chronic Insomnia Disorder
   B. Short-Term Insomnia Disorder
   C. Other Insomnia Disorder
   D. Isolate Symptoms and Normal Variants
      1. Excessive Time in Bed
      2. Short Sleeper

II. Sleep-Related Breathing Disorders
   A. Obstructive Sleep Apnea Disorders
      1. Obstructive Sleep Apnea, Adult
      2. Obstructive Sleep Apnea Disorder, Pediatric
   B. Central Sleep Apnea Syndromes
      1. Central Sleep Apnea with Cheyne–Stokes Breathing
      2. Central Apnea Due to a Medical Disorder without Cheyne–Stokes Breathing
      3. Central Sleep Apnea Due to High-Altitude Periodic Breathing
      4. Central Sleep Apnea Due to a Medication or Substance
      5. Primary Central Sleep Apnea
      6. Primary Central Sleep Apnea of Infancy
      7. Primary Central Sleep Apnea of

IV. Circadian Rhythm Sleep–Wake Disorders
   A. Delayed Sleep–Wake Phase Disorder
   B. Advanced Sleep–Wake Phase Disorder
   C. Irregular Sleep–Wake Rhythm Disorder
   D. Non–24-Hour Sleep–Wake Rhythm Disorder
   E. Shift Work Disorder
   F. Jet Lag Disorder
   G. Circadian Sleep–Wake Disorder Not Otherwise Specified

V. Parasomnias
   A. NREM-Related Parasomnias
      1. Confusional Arousals
      2. Sleepwalking
      3. Sleep Terrors
      4. Sleep-Related Eating Disorder
   B. REM-Related Parasomnias
      1. REM sleep behavior disorder
      2. Recurrent Isolated Sleep Paralysis
      3. Nightmare Disorder
   C. Other Parasomnias
      1. Exploding Head Syndrome
      2. Sleep-Related Hallucinations

Prematurity
8. Treatment-Emergent Central Sleep Apnea

C. Sleep-Related Hypoventilation Disorders
1. Obesity Hypoventilation Syndrome
2. Congenital Central Alveolar Hypoventilation Syndrome
3. Late-Onset Central Hypoventilation and Hypothalamic Dysfunction
4. Idiopathic Central Alveolar Hypoventilation
5. Sleep-Related Hypoventilation Due to a Medication of Substance
6. Sleep-Related Hypoventilation Due to a Medical Disorder

D. Sleep-Related Hypoxemia Disorder
1. Sleep-Related Hypoxemia

E. Isolated Symptoms and Normal Variants
1. Snoring
2. Catathrenia

III. Central Disorders of Hypersomnolence
A. Narcolepsy Type 1
B. Narcolepsy Type 2
C. Idiopathic Hypersomnia
D. Kleine–Levin Syndrome
E. Hypersomnia Due to a Medical Disorder
F. Hypersomnia Due to a Medication or Substance
G. Hypersomnia Associated with a Psychiatric Disorder
H. Insufficient Sleep Syndrome
I. Isolated Symptoms and Normal Variants
1. Long Sleeper
3. Sleep Enuresis
4. Parasomnia Due to Medical Disorder
5. Parasomnia Due to a Medication or Substance
6. Parasomnia, Unspecified
7. Isolated Symptoms and Normal Variants
a. Sleep Talking

VI. Sleep-Related Movement Disorders
A. Restless Legs Syndrome
B. Periodic Limb Movement Disorder
C. Sleep-Related Leg Cramps
D. Sleep-Related Bruxism
E. Sleep-Related Rhythmic Movement Disorder
F. Benign Sleep Myoclonus of Infancy
G. Propriospinal Myoclonus at Sleep Onset
H. Sleep-Related Movement Disorder Due to Medical Disorder
I. Sleep-Related Movement Disorder Due to a Medication or Substance.
J. Sleep-Related Movement Disorder, Unspecified
K. Isolated Symptoms and Normal Variants
1. Excessive Fragmentary Myoclonus
2. Hypnagogic Foot Tremor and Alternating Leg Muscle Activation
3. Sleep Starts (Hypnic Jerks)

VII. Other Sleep Disorder

Diagnostic Criteria

Insomnia should be evaluated according to eight criteria. DSM-5 sets forth these as follows: (1) how it adversely affects sleep; (2) how it influences subsequent wakefulness; (3) its frequency; (4) its chronicity; (5) whether it occurs in conjunction with some other sleep disorder; (6) medication or substance use; or (7) some other psychiatric, neurologic, or medical condition. Also, (8) the sleep difficulty must occur notwithstanding an adequate opportunity to sleep.

The primary criteria descriptively relates to how the insomnia affects sleep. Specifically, one or more of the following must be present: (a) difficulty initiating sleep, (b) difficulty maintaining sleep, or (c) early-morning awakening with inability to resume sleep. In the case of children, insomnia can manifest as resistance to the caregiver designation of bed time and/or difficulty sleeping without caregiver intervention. The

influences on subsequent wakefulness in DSM-5 include "significant distress or impairment of social, occupational, educational, academic, behavioral, or other important areas of functioning." ICSD-3 also considers fatigue; malaise; attention, concentration, or memory impairment; mood disturbance; irritability; behavioral problems (especially in children) including hyperactivity, impulsivity, or aggression; reduced motivation, energy, or initiative; and error or accident proneness.

Frequency and chronicity criteria stipulate sleep disturbance (with associated symptoms) three times per week (or more) for 3 months, or longer. The insomnia must not be better explained by another sleep, psychiatric, neurological, or medical disorder; must not be attributable to a medication or substance use; and must persist when adequate opportunity for sleep occurs.

When the diagnosis Insomnia Disorder is made, the clinician should also specify if there are comorbid conditions (be they sleep disorder, psychiatric, neurologic, or medical) and whether the insomnia occurs episodically (symptoms last one, or more, months but less than 3 months), persistently (3 months or more), or recurrently (two or more episodes in a 12-month period).

Insomnia Subtypes

Neither DSM-5 nor ICSD-3 formalizes diagnostic criteria for clinical and pathophysiological insomnia subtypes. Nonetheless, ICSD-3 describes each briefly. These include psychophysiological insomnia, idiopathic insomnia, paradoxical insomnia, inadequate sleep hygiene, behavioral insomnia of childhood, insomnia comorbid with mental disorder, insomnia comorbid with medical condition, and insomnia due to drug or substance. ICSD-3 authors caution that "Despite the heuristic appeal of these various subtypes, it is often difficult to discriminate among them in clinical practice." While this may be the case, paying attention to and understanding the overlapping complications of these subtypes can be essential for creating an effective treatment plan. The insomnia subtypes are described here to alert the clinician of possible insomnia nuances and complications.

Psychophysiological Insomnia. Psychophysiological Insomnia (PPI) involves conditioned arousal associated with the thought of sleeping. Objects related to sleep (e.g., the bed, the bedroom) likewise have become conditioned stimuli that evoke insomnia. Daytime adaptation is usually good (in contrast to impaired function in patients with psychiatrically related insomnia); however, there can be extreme tiredness and the individual can become desperate. PPI often occurs in combination with stress and anxiety disorders, delayed sleep phase syndrome, and hypnotic drug use and withdrawal. Other features characteristic of PPI include: (1)

excessive worry about not being able to sleep, (2) trying too hard to sleep, (3) rumination—inability to clear one's mind while trying to sleep, (4) increased muscle tension when getting into bed, (5) other somatic manifestations of anxiety, (6) ability to fall asleep when not trying to (e.g., when watching television), and (7) sleeping better away from own bedroom (including the sleep laboratory). In one sense, the individual has developed a performance anxiety concerning the ability to sleep and autonomic nervous system conditioning reinforces the situation. Stimulus Control Therapy is well suited for treating PPI (see behavioral treatments for insomnia). This nonpharmacological approach is designed to improve the association between going to bed and being able to fall asleep. Because many patients with PPI have developed poor sleep habits, improving sleep hygiene is usually beneficial. If muscle tension and rumination at bedtime are prominent features, relaxation therapy is a useful ancillary treatment.

Idiopathic Insomnia. Idiopathic Insomnia characterizes patients with a lifelong inability to obtain adequate sleep. The insomnia predates any psychiatric condition and other etiologies must be ruled out or treated, including psychophysiological insomnia, environmental sleep disturbances, and practices constituting poor sleep hygiene. If the insomnia persists then it is assumed that there is a defect in the neurological mechanisms that govern the sleep–wake system. That is, the sleep homeostatic process is dysfunctional. Sleep restriction therapy may provide some benefits and will also give the clinician a clue to the limits of the patient's homeostatic process. Sleep restriction is a technique by which homeostatic drive is maximized by sleep schedule compression (see behavioral treatments). Unfortunately, patients with idiopathic insomnia will likely never have normal sleep. Treatment is difficult; however, improved sleep hygiene, relaxation therapy, and use of hypnotic medicines are reportedly helpful. Although controversial, this group may require chronic pharmacotherapeutic intervention.

Paradoxical Insomnia. Paradoxical Insomnia, at its core, involves a dissociation between sleep and its usual attendant unconsciousness. In Paradoxical Insomnia, an individual thinks he or she is awake and having insomnia even though brain electrophysiological activity pattern is consistent with the correlates of normal sleep. That is, Paradoxical Insomnia is diagnosed when a patient complains of difficulty initiating or maintaining sleep and no objective evidence of sleep disruption is found. For example, a patient sleeping in the laboratory reports it taking more than an hour to fall asleep, awakening more than 30 times, and sleeping less than 2 hours the entire night. By contrast, the polysomnographic tracing showed sleep onset occurring within 15 minutes, few awakenings, a 90 percent sleep efficiency, and total sleep time greater than 7 hours.

Paradoxical Insomnia can occur in individuals who are apparently free from psychopathology; however, it may represent a somatic delusion or hypochondriasis. Some patients with Paradoxical Insomnia have obsessional features with respect to somatic functions. Cognitive relabeling and/or diffusing the worry about being unable to sleep can help. Interestingly, anxiolytics can profoundly reduce the perception of sleeplessness without remarkably changing sleep physiologically.

At various times, paradoxical Insomnia was previously called *pseudoinsomnia, sleep state misperception,* and *insomnia without objective findings*. Such appellations imply invalidation of the complaint and place the clinician at odds with the patient. Making matters worse, this dissociation between the perception and physiology is very common in patients with schizophrenia. However, a study has shown there is an increased metabolic rate and maximum O_2 uptake in patients with *Paradoxical Insomnia* compared to normal controls. Interestingly, most individuals readily acknowledge dreaming and understand that dreams constitute mental activity; nonetheless, there is a general mistaken impression that mental activity ceases during sleep. One might arrive at this conclusion because based on the inability to recall anything between the bed time and morning arising. However, this confuses the amnesic property of sleep with lack of mentation. While it is true that under normal circumstances sleep and unconsciousness are coupled, in some cases these phenomena dissociate. Anxiety can provoke sleep-state misperception and ruminative worry about not sleeping adds fuel to the fire. Thus if one lies in bed attempting to sleep and has ruminative thoughts, the individual may perceives themselves as being awake, even though brain activity is consistent with the electrophysiological correlates associated with normal sleep. A psychoeducational approach that explains the dissociability between brain activity and the perception of sleep can be helpful. Furthermore, the ruminations (which seem so urgent at the time but lose their importance, if remembered at all in the morning) may actually be a useful marker of sleep onset. If an individual can interpret the rumination as an indicator that they are asleep and it belays their worrying about not being able to sleep, the paradoxical insomnia often diminishes or even disappears.

Inadequate Sleep Hygiene. Inadequate Sleep Hygiene refers to insomnia produced by behaviors that are not conducive to good sleep. ICSD-1 defined inadequate sleep hygiene as “a sleep disorder due to the performance of daily living activities that are inconsistent with the maintenance of good quality sleep and full daytime alertness.” Many behaviors can interfere with sleep. Some of these behaviors increase arousal. For example, consuming caffeine or nicotine at night or engaging in excessive emotional or physical stimulation within a few hours of

bedtime. Other behaviors interfere with sleep architecture including daytime naps and a large variation of the daily sleep–wake schedule. It is widely believed that improving sleep hygiene can help relieve insomnia (see Table 23–2). It must be realized, however, that many activities related to sleep hygiene are life style issues and effecting life style change is difficult. Treatment should focus on only one or two problem areas at a time. Overwhelming the patient with too many lifestyle changes or a complex regimen seldom succeeds.

Table 23–2.
Do's and Don'ts for Good Sleep Hygiene

	DO	DON'T
Maintain regular hours of bedtime and arising	✓	
If you are hungry, have a light snack before bedtime	✓	
Maintain a regular exercise schedule	✓	
Give yourself an approximately an hour to wind down before going to bed	✓	
If you are preoccupied or worried about something at bedtime, write it down and deal with it in the morning	✓	
Keep the bedroom cool	✓	
Keep the bedroom dark	✓	
Keep the bedroom quiet	✓	
Take naps		✓
Watch the clock so you know how bad your insomnia actually is		✓
Exercise right before going to bed in order wear yourself out		✓
Watch television in bed when you cannot sleep		✓
Eat a heavy meal before bedtime to help you sleep		✓
Drink coffee in the afternoon and evening		✓
If you cannot sleep, smoke a cigarette		✓
Use alcohol to help in going to sleep		✓
Read in bed when you cannot sleep		✓
Eat in bed		✓
Exercise in bed		✓
Talk on the phone in bed		✓

Behavioral Insomnia of Childhood. Behavioral Insomnia of Childhood results from the child's dependency on specific stimulation, objects or setting for initiating sleep or returning to sleep. Without presence of the teddy-bear or parent, sleep onset is significantly delayed. Alternatively, without adequate limit setting by the caregiver, bedtime stalling ("Dad, I'm thirsty—can I get some water?"), or bedtime refusal ("I'm not tired! I don't want go to sleep!") can ensue. *Limit-setting Sleep Disorder* occurs when a child takes advantage of an inadequately enforced bedtime schedule. Childhood insomnia, as it is sometimes called, involves a child refusing to or delaying going to bed when the caretaker stipulates. If unresolved, when the child reaches adolescence and can set his/her own schedule, inadequate

sleep hygiene, insufficient sleep syndrome, or irregular sleep–wake pattern may become the problem. Delay strategies employed by the child include getting up to use the bathroom, requests for another drink, wanting another bedtime story, or complaints of hunger. Sometimes children will claim that they are afraid of the dark, monsters under the bed, or some other factitious creation. The well-intentioned parent may be manipulated out of emotional concern. The problem often begins when the child can climb out of the crib or has been moved to a bed. The caregivers' knowledge and ability to properly set limits is critical; therefore, education of caregiver can provide effective treatment.

Insomnia Comorbid with Mental Disorder. Insomnia Comorbid with Mental Disorder is the most common form of disorders of initiating and maintaining sleep. Thirty-five percent of patients seen in sleep disorders centers with insomnia as their chief complaint had a mental, and half of these patients had a mood disorder.

Ninety percent of patients with Major Depressive Disorder (MDD) have insomnia. Insomnia is a risk factor (or marker) for depression on 3-year follow-up, conferring an odds ratio of 3.95 compared to individual without insomnia. Insomnia is also an independent risk factor for suicide in patients with MDD, and 1 month after a trauma, insomnia is predictive of who will develop PTSD 1 year after a trauma. It has also been suggested that treating insomnia may help prevent mental disorders, especially depression. Sleep disturbances using EEG criteria in MDD include: (1) generalized sleep disturbance (increase sleep latency, increased nocturnal awakenings, and early morning awakenings), (2) slow-wave sleep decrease in the first nonrapid eye movement (NREM)-rapid eye movement (REM) cycle with delta activity shifts to the second NREM period, (3) latency to REM sleep shortened, (4) REM sleep occurring earlier in the night, and (5) REM density increase (especially early in the night) (see Fig. 23–1). Selective REM sleep deprivation alleviates depression. Furthermore, this antidepressant action can persist several days, or more. Interestingly, most antidepressant medication suppresses REM sleep (including: selective serotonin reuptake inhibitors, tricyclic antidepressants, and monoamine oxidase inhibitors). Arecoline (a cholinergic agonist) infusion will induce REM sleep in depressed patients and Scopolamine withdrawal mimics depression in normal subjects (including EEG sleep changes). These data support the Cholinergic-Aminergic Imbalance theory of Major Depression. While it still holds that any drug capable of suppressing REM sleep has potential antidepressant properties, a couple of newer antidepressants do not suppress REM sleep (e.g., bupropion and nefazodone). Thus, a piece is missing from the theory. Restated, REM suppression is sufficient but not necessary for antidepressant action. Treating insomnia while treating MDD may improve antidepressant effect. In a study on patients meeting DSM-IV

criteria for both MDD and insomnia (n = 545), each received a morning dose of fluoxetine and were randomized to nightly eszopiclone 3 mg (ESZ + FLX) or placebo (PBO + FLX) for 8 weeks. Subjective sleep and daytime function were assessed weekly. Patients in the ESZ + FLX group had significantly decreased sleep latency, wake time after sleep onset, increased total sleep time, sleep quality, and depth of sleep. Eszopiclone cotherapy also produced significant and progressive improvement in mood when compared to fluoxetine alone. And the improvements of adjunctive insomnia therapy persisted for 2 weeks after the eszopiclone treatment was discontinued; fluoxetine was continued.

Bipolar disorder profoundly adversely affects sleep. Patients with mania and hypomania seldom complain of sleep problems even though they sleep only for a short duration (2 to 4 hours per night), have very prolonged latency to sleep, and sometimes have reduced slow-wave sleep. When they enter the depressed phase, they are more likely to be seen at a sleep disorders center. Clinically, it is also noted that patients with bipolar disorder, if they break from their sleep–wake routine and stay up very late or all night, may provoke a manic episode.

Patients with Schizophrenia have no consistent EEG sleep changes except that they do not have REM sleep rebound in response to REM sleep deprivation. The only other consistent finding is that patients with schizophrenia frequently deny having slept even though EEG-EOG-EMG sleep appears normal. Patients with anxiety and personality disorders often have sleep onset, and sleep maintenance insomnia. Short REM sleep latencies and reduced slow-wave sleep are reported in some studies; however, the finding is not specific to schizophrenia and is not consistent.

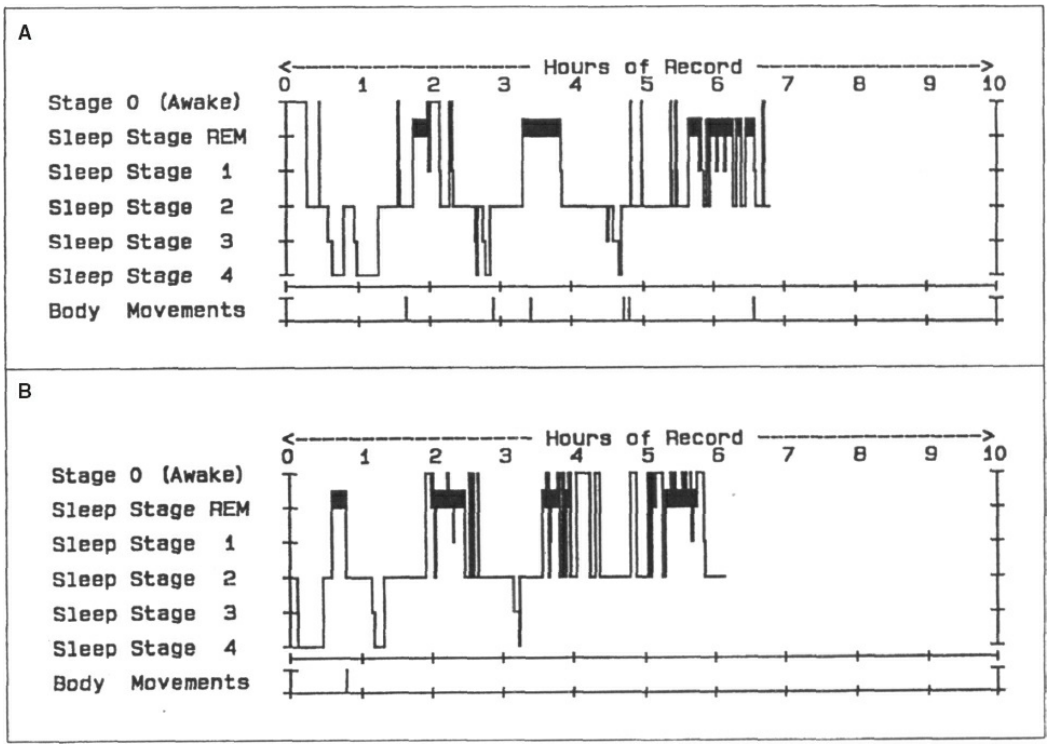

FIGURE 23–1. Sleep histograms showing normal sleep and insomnia. Sleep stage histograms comparing normal sleep (**Panel A**) with that found in a patient with major depressive disorders (**Panel B**). Difficulty maintaining sleep and early morning awakenings are common complaints in patients with depression. Panel B illustrates the electrophysiological correlates of these complaints beginning, in this case, after approximately 2 hours of sleep. Sleep continuity becomes disrupted as morning approaches. Also present is a markedly reduced latency to REM sleep. This sleep feature is characteristic of this patient population and is thought by some to reflect cholinergic–aminergic imbalance.

Patients with Generalized Anxiety Disorder worry excessively about multiple events, for more days than not, for at least 6 months; have difficulty controlling the worry; and the worrying causes at least three of the following: disturbed sleep, muscle tension, irritability, difficulty concentrating, restlessness, easily fatigued. Although insomnia is one of the key symptoms, data addressing nonsedating anxiolytic's efficacy for treating disturbed sleep are sparse. In a couple of randomized, placebo-controlled studies, the benzodiazepine receptor agonist (BZRA) zopiclone improve sleep onset and daytime anxiety when compared to benzodiazepines (BZDs). However, zopiclone is not available in the United States.

Insomnia Comorbid with Medical Condition. Insomnia accompanies many medical and neurological conditions. Given pain's potential for disturbing sleep, all medical conditions producing pain can (and usually do) disturb sleep. Unfortunately a synergy exists between pain and sleep such that poor sleep lowers pain threshold. This vicious cycle can present a difficult treatment challenge. However, reducing pain can also improve

sleep and improving sleep can reduce pain. Thus, a two-pronged approach to pain and sleep can prove beneficial. In other medical conditions, the sleep disturbance appears to be secondary. For example, sleep-related gastroesophageal reflux disease (GERD) is often associated with insomnia. Treating the reflux results in sleep improvement but insomnia treatments seldom, if ever, relieve nocturnal GERD. Patients with chronic obstructive lung disease commonly suffer from both sleep onset and sleep maintenance insomnia. Treating insomnia does not appear to ameliorate COPD. Neurodegenerative diseases are frequent associates of sleep disorders, including insomnia, hypersomnolence, parasomnias, and circadian rhythm disorders. Discussing specific medical and neurologic disorder associations with sleep disorders are beyond the scope of this chapter and the reader is referred to Chokroverty (2009) for details.

Insomnia due to Drug or Substance. Many prescription drugs, even when taken properly, can disturb sleep. This include antidepressants (i.e., fluoxetine), antihypertensives, corticosteroids, hypolipidemic medications, antiparkinsonian drugs, decongestants (i.e., pseudoephedrine), anorectics, stimulants, and antiepileptic medications. In addition, some sedating medications cause rebound insomnia when discontinued.

Alcohol and hypnotic use initially promote sleep onset because of their sedating properties. A problem occurs when sleep quality is adversely affected, tolerance develops after chronic use, or withdrawal begins. Alcohol may relax a tense person and thereby decrease latency to sleep; however, sleep later in the night, will usually be fragmented by arousals. As tolerance develops to the alcohol, greater amounts or more frequent dosing are needed to sustain the effects. Furthermore, during withdrawal or after tolerance develops, insomnia may rebound to a level more severe than the initial disturbance. If sleep is considered a behavior, and this behavior becomes dependent on the use of alcohol, then by definition this is a form of alcoholism. Interestingly, for many patients with alcohol dependence, insomnia predates alcohol use and persists even during prolonged periods of abstinence.

Caffeine (the active ingredient in coffee) and theobromine (the active ingredient in chocolate) are methyl xanthenes and they act as psychostimulants in the mammalian central nervous system (CNS). Psychostimulants increase sleep latency, reduce sleep efficiency, and decrease total sleep time. Caffeine's half-life is 3 to 7 hours and may interfere with sleep when consumed in large quantities throughout the day or even in smaller portions closer to bedtime. Some individuals are hypersensitive to methyl xanthenes and any coffee or chocolate can trigger difficulty falling asleep or awakening after a couple of hours of sleep with difficulty getting back to sleep. Finally abuse of illicit substances, particularly stimulants (such as cocaine and amphetamines), interfere with

sleep onset and sleep maintenance. Unlike with alcohol, discontinuation of these substances will cause hypersomnolence.

Short Sleeper. People who fall under this category require less than 5 hours of sleep per 24-hour period in order to maintain normal daytime functioning and mood. The patient may be distressed about not sleeping. However, if the short sleep duration does not provoke adverse daytime consequences, functional impairment, or is associated with relevant comorbid conditions, full Insomnia Disorder criteria are not met. Being a short sleeper appears to run in families but the specific genes are yet unidentified.

Treating Insomnia

Pharmacological Treatment. For many years, BZDs were the most commonly prescribed sedative hypnotic medications for treating insomnia. BZRAs replace BZDs as the standard for sedative-hypnotic medications used to treat insomnia. Subsequent development of a low-dose antihistamine (Silenor), a melatonin receptor agonist (ramelteon), and an orexin antagonist (Suvorexant) have more recently augmented the armamentarium of medications approved for treating insomnia. However, unapproved off-label (and not recommended) use of sedating antidepressants is common. Additionally, a variety of over-the-counter (OTC) sleep aids are available. Nonprescription formulas include sedating antihistamines, protein precursors, and other substances. L-tryptophan was popular and readily available at health food stores until an outbreak of eosinophilia led to it being pulled off the shelves. Melatonin is the leader among self-administered food additives believed by some to alleviate sleeplessness. Essentially, low-dose melatonin can improve sleep when used to align or re-entrain the circadian rhythm with the desired sleep–wake schedule.

Prescription medicines are rigorously tested in clinical trials; therefore, they hold an advantage over the virtually untested OTCs. To attain FDA approval as a hypnotic, a medication must be safe and effective. Most hypnotic medications are approved for short-, not long-term, use. Exceptions include zolpidem-modified release, eszopiclone, and ramelteon; all of which are approved for long-term therapy. When properly used, hypnotics can provide immediate and adequate relief from sleeplessness. Insomnia, however, usually returns upon discontinuation of dosing.

Nonpharmacological Treatment with Cognitive-Behavioral Therapy for Insomnia (CBTi). This treatment modality combines behavioral and cognitive techniques to overcome dysfunctional sleep behaviors and misperceptions, distorted, disruptive thoughts about sleep. The therapist begins with a careful clinical interview to assess the insomnia's etiology,

chronicity, severity, associations, and comorbid conditions. A treatment plan is then designed using cognitive and behavioral techniques deemed relevant and appropriate. These may include: universal sleep hygiene, stimulus control therapy, sleep restriction therapy, relaxation therapies and biofeedback, cognitive therapy, and occasionally, paradoxical intention.

Studies repeatedly show significant, sustained improvement in sleep symptoms with CBTi. Improvements include reduced wakefulness duration, decreased number of awakenings, and shorter latency to sleep onset. Short-term benefits are similar to that of medication but CBTi tends to have lasting benefits, even 36 months after treatment. With cessation of the medication insomnia frequently returns and is sometimes accompanied by rebound insomnia. CBTi has not been shown to produce any adverse effects. There are no established "best practice" guidelines for length or quantity of sessions. CBTi, however, is not without limitations. Most data do not compare the efficacy of the individual components (described below) of CBTi. However, sleep hygiene education alone does not appear to improve sleep. Intuitively, the multi-component approach addresses many of the variables contributing to insomnia. Other limitations of CBTi include lack or unavailability of trained specialists, cost of sessions, noncoverage by insurance carriers.

The effects of CBTi take longer to emerge than effects of medications. Individuals are often desperate when they finally seek treatment of their insomnia. This makes it difficult to convince them to try a therapy that may require several weeks before it provides relief. Also, patients must be active participants in this type of therapy. Many individuals not only want a "quick fix" but they also want to undergo a procedure or have something administered rather than be involved in the therapeutic process. For optimal CBTi, patients must commit to multiple sessions and also be open to the idea that modifying thoughts and behaviors about sleep can improve the symptoms of insomnia. The "quick fix" model is common in primary care; by contrast, psychiatrists are accustomed to delayed response from their experience with antidepressants and other psychotropic medicines. Consequently, psychiatrists may be more comfortable recommending CBTi than other practitioners.

Although firmly focused on cognitive and behavioral issues, it helps to extend CBTi just slightly into the psychodynamic sphere. For some patients with longstanding difficulty sleeping, being an insomniac becomes an important part of their identity. There may be primary or secondary gain to such identification. It is the negative emotional response (i.e., anger at the inability to control one's sleep, feeling like a failure because they can't sleep) to insomnia that contributes to its chronicity. In general, these individuals tend to internalize rather than express emotion, feel a heightened need for control, experience interpersonal difficulties, and have significant discontent with past events. For this subset of people, if the

emotional response is not addressed, there is more likely to be a limited response to CBTi or a relapse of insomnia over time. The clinician who is attuned (i.e., to patient's tendency to view something as a failure rather than a challenge), will be better able to intercept barriers to treatment. "As chronic insomnia is an extremely psychologically painful condition, a high level of motivation and willingness to explore oneself, at least with a specific focus, is often present."

UNIVERSAL SLEEP HYGIENE. The focus of universal sleep hygiene is on modifiable environmental and lifestyle components that may interfere with sleep as well as behaviors that may improve sleep. Especially because some of these behaviors are difficult to change, only one or two items that are collaboratively chosen by the patient and clinician should be addressed at a time. This will give the patient the best chance at a successful intervention. Sleep enhancing directives are enumerated in Table 23–2. Often a few simple alterations in a patient's habits or sleep environment can be effective. The clinician, however, needs to spend time reviewing both the patient's routine and its irregularity. In some respects, the essence of insomnia is its variability. The day-to-day changes in behavior and the changing severity of sleeplessness can obscure the factors responsible for the problem. A carefully explained program of sleep hygiene, with follow-up, represents a fairly inexpensive but effective intervention. Furthermore, improving sleep habits can enhance sleep even when the major cause of insomnia is physical.

STIMULUS CONTROL THERAPY. Stimulus control therapy is a deconditioning paradigm developed by Richard Bootzin and colleagues at the University of Arizona. This treatment aims to break the cycle of problems commonly associated with difficulty initiating sleep. By attempting to undo conditioning that undermines sleep, stimulus control therapy helps reduce both primary and reactive factors involved in insomnia. The rules attempt to enhance stimulus cues for sleeping and diminish associations with sleeplessness. The instructions are simple; however, they must be followed consistently. The *first rule* is to go to bed only when sleepy in order to maximize success. *Second,* use the bed only for sleeping. Do not watch television in bed, do not read, do not eat, and do not talk on the telephone while in bed. *Third,* do not to lie in bed and become frustrated if unable to sleep. After a few minutes (do not watch the clock), get up, go to another room, and do something nonarousing until sleepiness returns. The goal is to associate the bed with rapid sleep onset. Rule three should be repeated as often as needed. The *fourth and final instruction* attempts to enhance the mechanisms underlying the circadian and sleep–wake cycles; that is, awaken at the same time every morning (regardless of bedtime, total sleep time, or day of week) and totally avoid napping. Stimulus control therapy

does work; however, you may not see results during the first few weeks or month. If continually practiced, the bouts of insomnia lessen in both frequency and severity.

SLEEP RESTRICTION THERAPY. Sleep restriction therapy is a strategy designed to increase sleep efficiency by decreasing the amount of time spent awake while lying in bed. Developed by Arthur Spielman, this therapy specifically targets those patients who lay awake in bed unable to sleep. Restricting time in bed can help consolidate sleep. If the patient reports sleeping only 5 hours of a scheduled 8-hour time in bed, reduce the time in bed. It is advised, however, not to reduce bedtime to less than 4 hours per night and to warn the patient about the hazards of daytime sleepiness. Sleep at other times during the day must be avoided, except in the elderly who may take a 30-minute nap. The clinician then monitors sleep efficiency (time asleep as a percentage of the time in bed). When sleep efficiency reaches 85 percent (averaged over five nights), time in bed is increased by 15 minutes. Sleep restriction therapy produces a gradual and steady decline in nocturnal wakefulness.

RELAXATION THERAPY AND BIOFEEDBACK. The most important aspects of relaxation therapy are that it be performed properly and overlearned. Self-hypnosis, progressive relaxation, guided imagery, and deep breathing exercises are all effective if they produce relaxation. The goal is to find the optimal technique for each patient; but not all patients need help relaxing. Progressive muscle relaxation is especially useful for patients who experience muscle tension. Purposefully tense (5 to 6 seconds) and then relax (20 to 30 seconds) muscle groups beginning at the head and ending at the feet. Appreciate the difference between tension and relaxation. Guided imagery has the patient visualize a pleasant, restful scene, engaging all of their senses. Breathing exercises are practiced for at least 20 minutes per day for 2 weeks. Once mastered, use the technique once at bedtime for 30 minutes. If it does not work, try again another night. It is important that the technique not become associated with failure to fall asleep. Abdominal breathing is described below. The patient must become comfortable with the each step before moving on to the next. *First,* in the supine position, breathe normally through your mouth or nose, whichever is more comfortable, and attend to your breathing pattern. *Second,* while maintaining that rhythm, begin to breathe more with your abdomen and less with your chest. *Third,* pause for a half second after each breath cycle (in and out) and evaluate the breath. How did it feel? Was it smooth? Eventually each breath will become uniform and smooth. *Fourth,* find a place where you can best feel the air move in and out. Concentrate on that spot and on the air moving in and out. *Fifth,* visualize intrusive thoughts as floating away; if there are too many thoughts stop practicing and try again

later. Biofeedback provides stimulus cues for physiological markers of relaxation and can increase self-awareness. The machine can measure muscle tension in your forehead or finger temperature. Finger temperature will rise when you become more relaxed. Patients require careful and adequate training; simply giving them an instruction tape is not especially helpful. Techniques are ideally mastered during the day for several weeks before application to the sleep problem; this is best achieved outside of the bed. By the time the techniques are applied in bed, the skill should be automatic. Relaxation techniques readily lend themselves to being combined with sleep hygiene and stimulus control therapies. Sometimes, they make for good distractions from a person thinking about their inability to sleep. The ruminations fuel the insomnia and if the ruminator can be distracted, then they may sleep better.

COGNITIVE THERAPY. This effective, validated treatment for a variety of psychiatric conditions including major depression and generalized anxiety has been adapted for use with insomnia. The cognitive aspect of insomnia treatment targets the negative emotional response to an appraisal of a sleep-related situation. The negative emotional response is thought to produce emotional arousal which in turn contributes to or perpetuates insomnia. People who have maladaptive cognitions tend to exaggerate the negative consequences of insomnia: "there must be something really wrong with me if I can't fall asleep in 40 minutes." They also tend to have unrealistic expectations about their sleep requirements: "if I don't sleep 8 hours a night then my whole day will be ruined." The first step is to identify these cognitions, next challenge their validity and finally substitute them with more adaptive cognitions. This is most easily illustrated by example: DH is a 42-year-old male with a 5-year history of insomnia. He identifies being fired from his job and the birth of a colicky baby as precipitating factors in his inability to sleep. However, even after finding a new position with better hours and pay and now that the child sleeps through the night, he continues to experience difficulty falling and staying asleep. Perpetuating factors for JS include low back pain and a spouse with periodic limb movement disorder. He reports spending 8 to 9 hours in bed each night and sleeping only 4 to 5 hours intermittently. He watches 1 hour of TV in bed before turning out the light for bedtime. He spends hours watching his minutes tick away. He does not wake feeling rested and when his alarm goes off he is frequently already awake and has the thoughts "I hardly slept at all last night. I should be able to get more sleep. There must be something wrong with me. Great, I'll be too tired to concentrate on anything today."

Examples of maladaptive thoughts: *"I should be able to get more sleep."* This is a faulty appraisal of sleeping ability may relate to a need for control over sleep. This need for control interferes with having a more laissez-faire

attitude about a few missed hours of rest. Should thoughts also lead to feelings of frustration and anger. *"Great, I'll be too tired to concentrate on anything today."* This is a misattribution of daytime impairment due to poor sleep. DH is also magnifying the negative/discounting the positive and black-and-white or all-or-nothing thinking. Could DH be too tired to concentrate on some things but not all things? Might his inability to concentrate be due to a myriad of other factors? *"There must be something wrong with me [if I can't get enough sleep]."* Catastrophizing and emotional reasoning—just because you have a feeling, doesn't mean that the thought/feeling is true. A strongly held belief that sleeplessness negatively impacts physical and mental health can set off catastrophizing (see Table 23–3).

Table 23–3.
Catastrophizing and Alternative Thoughts Scenario

Situation	Alarm clock goes off at 6:02 AM Monday
Thought	"I hardly slept at all last night. I should be able to get more sleep. There must be something wrong with me. Great, I'll be too tired to concentrate on anything today."
Behavior/Feeling	Hit snooze to try to get 9 more minutes of sleep. Despair 70% Irritated 75%
Evidence	"Too tired to concentrate" For: I usually feel tired on Mondays. I lost my keys last week. Against: I just got a promotion, so I must be doing something right
Alternative balanced thought	"I did get some sleep last night and I usually can concentrate on my work."
Feeling	Despair 50% Irritated 60%

PARADOXICAL INTENTION. This is a cognitive technique with conflicting evidence regarding its efficacy. In clinical practice compliance is often a barrier, but it does work for a limited number of patients. The theory is that performance anxiety interferes with sleep onset. So when the patient tries to stay awake for as long as possible, rather tries to fall asleep, performance anxiety will be reduced and sleep latency will improve.

HYPERSOMNOLENCE DISORDER

The term *hypersomnolence* broadly refers to excessive sleepiness, excessive sleep quantity of sleep, and drowsiness with reduced vigilance during usual periods of wakefulness. Excessive sleepiness can be a serious, debilitating, potentially life-threatening condition. It affects not only the afflicted individual but also their family, coworkers, and the public at large. Although excessive sleepiness can arise from insufficient sleep, disrupted

sleep, or desynchrony with one's circadian rhythm, the Hypersomnolence Disorder is thought to stem from some basic neurologic sleep regulation dysfunction.

Hypersomnolence adversely affects attention, concentration, memory, and higher order cognitive processes. Serious results of sleepiness include failure at school, loss of employment, motor vehicle accidents, and industrial disasters. The transportation industry is particularly prone to sleep-related accidents, including trucking, railroad, marine, and aviation. There are many sleep disorders associated with excessive daytime sleepiness; however, Obstructive Sleep Apnea Hypopnea (OSAH) is by far the most common dyssomnia seen in sleep disorders centers. By contrast, Hypersomnolence Disorder is a relatively rare condition.

Diagnostic Criteria

The Hypersomnolence Disorder should be evaluated according to six criteria. DSM-5 sets forth these as follows: (1) how it influences wakefulness notwithstanding adequate sleep time (7 hours or more for the adult); (2) its frequency; (3) its adverse effect on functioning or emotional stability; (4) that the sleepiness is not better explained by whether it occurs in conjunction with some other sleep disorder; (5) medication or substance use; or (6) some other psychiatric, neurologic, or medical condition. Other conditions may be present (and should be specified); however, they must not better explain the hypersomnolence. All criteria must be met.

The primary criteria descriptively relate to how the hypersomnolence effects wakefulness. Specifically, one or more of the following must be present: (a) recurrent sleep episodes or lapses into sleep within the same day; (b) a prolonged (9 hours or more) sleep episode that is not refreshing; (c) difficulty achieving full wakefulness upon abrupt awaking (sometimes called, *sleep inertia*). The frequency criteria stipulates that hypersomnolence must occur three times (or more) weekly for at least 3 months. The influence on wakefulness must be associated with "significant distress or impairment in cognitive, social, occupational, or other important areas of functioning." To diagnose Hypersomnolence Disorder must not be attributable to a medication or substance use.

When the diagnosis Hypersomnolence Disorder is made, the clinician should specify if there are comorbid conditions (be they sleep disorder, psychiatric, neurologic, or medical) and whether the hypersomnolence duration is acute (less than 1 month), subacute (1 to 3 months duration), or persistent (more than 3 months). The diagnosis should also include specification for severity based on the patient's difficulty maintaining alertness while socializing, inactively sitting, reading, watching television, working, driving, or engaging in other activities. The difficulty can result from extreme drowsiness, nearly irresistible sleepiness, or actual lapses into sleep. The clinician should specify "mild" if difficulties occur 1 to 2

days per week, “moderate” if 3 to 4 days per week, and “severe” if 5 to 7 days per week.

Characteristics and Hypersomnolence Subtypes

Hypersomnolence Disorder presents in several forms. It may be associated with very long sleep periods after which the individual remains sleepy (see Fig. 23–2). The disorder can also occur with adequate but not long sleep periods. Sleep in patients with Hypersomnolence Disorder is usually well preserved and sleep efficiency remain high even in forms associated with extended sleep schedules (12 hours, or more). Furthermore, the patient readily falls asleep if given an opportunity to nap the following day. Sleep sometimes has elevated slow-wave sleep; however, the EEG sleep pattern is essentially the same as that found in normal individuals who are sleep deprived. Unlike a sleep-deprived individual, the sleep pattern continues in this profile even after several nights of extended sleep. The etiology of Hypersomnolence Disorder is not known; however, a CNS cause is presumed. Some individuals with Hypersomnolence Disorder are HLA Cw2 positive, have autonomic nervous system dysfunctions, and have other affected family members. In some cases (approximately 10 percent) begin with a viral infection and quickly evolve into more permanent hypersomnolence (e.g., infectious mononucleosis, viral pneumonia, Guillain-Barre syndrome). There also appears to be a familial form of Hypersomnolence Disorder with autosomal-dominance inheritance.

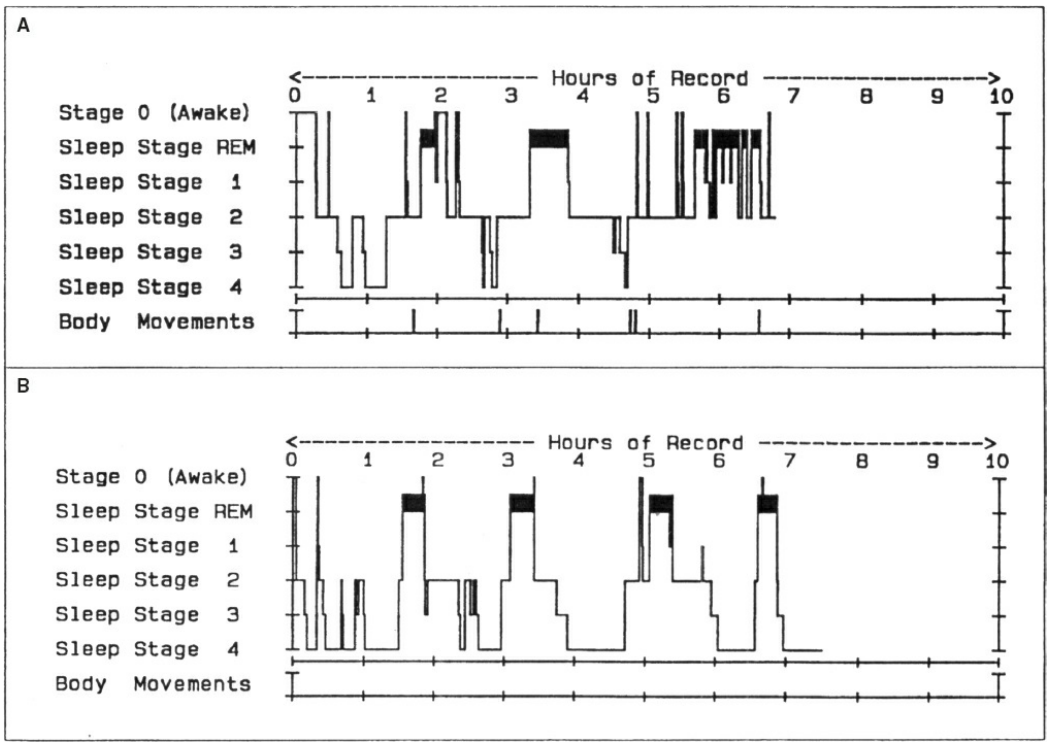

FIGURE 23–2. Sleep stage histograms comparing a normal sleep profile with that of a patient with hypersomnolence. **Panel A** illustrates the sleep pattern recorded in a healthy

young adult. His sleep was interrupted several times by awakenings associated with other concurrent laboratory procedures; however, all parameters were within normal limits. **Panel B** depicts the sleep pattern of a patient with severe daytime sleepiness. The patient did not have narcolepsy or apnea. In the laboratory, sleep onset was rapid, sleep continued virtually without interruption, and he was extremely difficult to arouse. This recording was terminated after 7.5 hours to begin daytime testing; however, the patient's sleep revealed little indication of progressing toward a lighter state. In contrast to the normal sleep pattern, slow-wave sleep duration was elevated and it continued to emerge in alternation with REM sleep throughout the night. The patient awoke feeling unrefreshed and not satisfied with his sleep.

DSM-5 does not include descriptions of Hypersomnolence Disorder subtypes; however, ICSD-2 and ICSD-3 discuss some of these variants. Perhaps the intriguing of these is Kleine–Levin syndrome. Kleine–Levin predominantly afflicts males in early adolescence; however, it can occur later in life and in females. In its classic form, the recurrent episodes are associated with extreme sleepiness (18- to 20-hour sleep periods), voracious eating, hypersexuality, and disinhibition (e.g., aggression). Episodes typically last for a few days up to several weeks and appear once to 10 times per year. A monosymptomatic hypersomnolence form can occur. The frequency of the HLA DQB1*02 is increased in patients with this syndrome.

Medical condition known to cause hypersomnolence includes head trauma, stroke, encephalitis, Parkinson disease, inflammatory conditions, tumors, genetic diseases, and neurodegenerative diseases. Recurrent hypersomnolence can also occur related to the menstrual cycle and appears in months of menarche. The symptoms typically last for 1 week and resolve with menstruation. Treatment with oral contraceptives is effective and therefore the disorder is thought to be secondary to a hormone imbalance.

Of course, sleepiness can be provoked by many prescription medications and recreational substances. Use or abuse of sedative-hypnotics, sedating antihistamines, sedating antidepressants, antiepileptics, neuroleptics, and opioid analgesics can cause sleepiness. Hypersomnia may also be provoked by withdrawal from cocaine, amphetamines, caffeine, and nicotine.

Treating Hypersomnolence Disorder

Hypersomnolence Disorder is usually managed pharmacologically. Although there are no medications indicated for its treatment, stimulants are commonly used with varying degrees of success. The wake promoting substance modafinil or traditional psychostimulants (amphetamines and their derivatives) sometimes help manage the sleepiness. In addition to drug therapies, the overall therapeutic approach usually includes lifestyle adjustment, psychological counseling, drug holidays to reduce tolerance (if stimulants are used), and careful monitoring of refills, general health, and cardiac status.

NARCOLEPSY

For decades, Narcolepsy was characterized by a tetrad of symptoms. These included (1) excessive sleepiness, (2) cataplexy, (3) sleep paralysis, and (4) hypnagogic hallucinations. Hypersomnolence and cataplexy always served as the key symptoms. Cataplexy involves the sudden loss of muscle tone, usually with continuing full or partial consciousness. Cataplexy is commonly precipitated by laughter. However, other strong emotions can "trigger" cataplexy. Cataplexy severity ranges widely from transient weakness in knees to total paralysis while the patient is fully conscious. Episodes may last from several seconds to minutes. Usually, the patient is unable to speak and may fall to the floor. According to DSM-5, in children (or during the first 6 months of onset), cataplexy can be recognized by "spontaneous grimaces or jaw-opening episodes with tongue thrusting or a global hypotonia, without any obvious emotional triggers."

Sleep paralysis and hypnagogic (or hypnopompic) hallucinations are nonspecific symptoms and may occur in isolation or in conjunction with other sleep disorders. Polysomnographic sleep studies typically revealed abnormal sleep architecture in which REM sleep occurs soon after sleep onset both at night and during daytime naps (see Fig. 23–3). This, in connection with the symptom tetrad suggested narcolepsy involved REM sleep intrusion, presumably resultant from dysfunction of REM sleep generator gating mechanisms. The tetrad features resembled REM sleep characteristics. Sleep paralysis is similar to the muscle atonia occurring during REM sleep and hypnagogic hallucinations resembling dreaming. Nocturnal sleep is often fragmented and there can be considerable sleep disturbance. Patients may experience depression associated with narcolepsy, especially when it is not treated. Social isolation, difficulty with academics and employment, and fear of driving contribute to a sense of loss experienced by patients with narcolepsy.

The discovery that Narcolepsy is strongly associated with a hypocretin (orexin) deficit radically changed diagnostic practice. Narcolepsy diagnostic criteria now include variants without cataplexy. ICSD-3 actually has separate criteria for Narcolepsy with and without cataplexy. DSM-5 handles these variants with specific specifications.

Diagnostic Criteria

Patients with Narcolepsy have irresistible sleep episodes, lapses into sleep, or napping at least 3 days per week for at least 3 months. This hypersomnolence profile must occur in conjunction with one or more of the following:

1. Cataplexy occurring at least a few times per month.
2. Cerebrospinal fluid (CSF) hypocretin-1 deficiency (110 pg/mL [or less] or less than one-third of the assay's control normative values. [The CSF

hypocretin deficiency must not be in conjunction with acute brain injury, infection, or inflammation.])

3. An overnight polysomnography revealing an REM sleep latency of 15 minutes (or less) or a multiple sleep latency test (MSLT) on which mean sleep latency is 8 minutes (or less) and REM sleep occurs on two (or more) nap opportunities.

DSM-5 also indicates that the clinician should specify whether the diagnosis is: (a) Narcolepsy without cataplexy but with hypocretin deficiency, (b) Narcolepsy with cataplexy but without hypocretin deficiency, autosomal dominant cerebellar ataxia, deafness, and narcolepsy (because these conditions can be associated with low or intermediate CSF hypocretin levels), (c) Autosomal dominant narcolepsy, obesity, and type 2 diabetes (because patients with obesity and type 2 diabetes can have low CSF hypocretin levels), or (d) Narcolepsy secondary to another medical disorder (because some diseases or head trauma can be associated with hypocretin neuron destruction).

The variant of narcolepsy, previously called *Narcolepsy Due to Medical Condition,* often exhibits cataplexy. The cataplexy can be provoked by tumors or sarcoidosis of the hypothalamus; multiple sclerosis plaques affecting the hypothalamus; paraneoplastic syndrome with anti-Ma2 antibodies; Neiman Pick type C disease (especially in children younger than 5 years); and possibly Coffin–Lowry syndrome. Cataplexy is usually absent in narcolepsy associated with multiple sclerosis, myotonic dystrophy, Prader–Willi syndrome, Parkinson disease, and multiple system atrophy. Narcolepsy due to medical condition included what used to be called Posttraumatic Hypersomnia which may or may not present with cataplexy.

Treating Narcolepsy

Currently, Narcolepsy cannot be cured; however, symptoms can be managed with either the wake-promoting substance modafinil (first-line treatment) or traditional psychostimulants like amphetamines and their derivatives (if modafinil fails). For managing cataplexy, clinicians often administer REM sleep suppressing drugs (e.g., many antidepressants). This approach capitalizes on the anticholinergic REM sleep suppressant properties of these drugs. Since cataplexy is presumably an intrusion of REM sleep phenomena into the awake state, the rationale is clear. Many reports indicate that imipramine and protriptyline are quite effective for reducing or eliminating cataplexy. Selective serotonin reuptake inhibitors have gained popularity because they are associated with fewer side effects than the tricyclic antidepressants. More recently, sodium oxybate (Xyrem) has proven extremely effective for reducing cataplexy, even in cases where the cataplexy was thought to be intractable. Studies also suggest that sodium oxybate helps improve sleep and relieves some of the sleepiness

associated with narcolepsy. Although drug therapies are the treatment of choice, the overall therapeutic approach should include scheduled naps, lifestyle adjustment, psychological counseling, drug holidays to reduce tolerance (if stimulants are used), and careful monitoring of refills, general health, and cardiac status.

SLEEP-RELATED BREATHING DISORDERS

Overview

Sleep-related breathing disorders includes disorders ranging along several different dimensions, including those involving (1) large airway obstruction during sleep, (2) breathing cessation resulting from central respiratory mechanisms, and (3) hypoventilation without breathing cessation. A sleep-related breathing cessation (sleep apnea) is defined as a cessation of breathing for 10 seconds or more during sleep. A reduction in breathing is termed a hypopnea. These sleep-related breathing impairments are most often caused by airway obstruction; however, sometimes respiratory reduction results from central (brainstem) changes in ventilatory control, heart failure, or metabolic factors.

Cessation or reduced breathing during wakefulness is not intrinsically pathophysiological. One can voluntarily hold his or her breath to avoid noxious odors. When speaking, an individual reduces airflow to the lungs. Neither of these activities constitute dysfunction. However, when asleep, breathing cessations or reduction typically provokes CNS arousal, significant oxyhemoglobin desaturation, or both. Thus, the apnea or hypopnea *induces* a pathophysiological event.

Each sleep apnea can be classified as obstructive, central, or mixed. Obstructive apnea results from airway collapse notwithstanding continued respiratory effort. In a central apnea, the breathing cessation results from decreased or absent respiratory effort. Mixed apnea typically begins with a central apnea but when respiratory effort resumes, airway collapse prevents ventilation. Figure 23–4 illustrates obstructive, central, and mixed apnea episodes. A partial airway obstruction that produces a 4 percent (or more) oxyhemoglobin desaturation is scored as an obstructive hypopnea, according to Central Medicare Services (CMS) criteria. By contrast, partial airway obstruction not meeting the CMS desaturation criteria but provoking a brief CNS arousal can be scored as a respiratory effort related arousal (RERA) according to AASM criteria. Figure 23–5 shows a polysomnographic tracings for obstructive hypopnea and RERA.

Obstructive Sleep Apnea Hypopnea

Diagnostic Criteria. Obstructive sleep–related breathing ranges across the spectrum from upper airway resistance syndrome to severe obstructive sleep apnea (OSAH). Adult Diagnostic Criteria: An adult meets criteria if an

overnight polysomnographic sleep study a combined total of predominantly obstructive apnea and hypopnea of 15 (or more) per hour of sleep. However, if the patient has relevant symptoms or comorbid conditions, only 5 (or more) events per hour are required. Relevant symptoms include any of the following: hypersomnolence (given that the individual has adequate opportunity to sleep); witnessed cessation of breathing during sleep; gasping or snorting; or habitual snoring. Medicare and/or ICSD-3 also considers relevant comorbid conditions include hypertension, mood disorder, cognitive dysfunction, coronary artery disease, stroke, congestive heart failure, atrial fibrillation, or type 2 diabetes reasons to use the lower apnea + hypopnea criteria. ICSD also defines Pediatric Diagnostic Criteria as a child with overnight polysomnographic sleep showing a combined total of one (or more) obstructive apnea, mixed apnea, and hypopnea per hour of sleep. An alternate polysomnographic finding of obstructive hypoventilation (25 percent [or more], of total sleep time with hypercapnia [$PaCO_2$ > 50 mm Hg] in association with flattened nasal pressure tracings during inspiration. Either of these two polysomnographic findings when associated with snoring; paradoxical movement of the chest wall and abdomen during sleep; sleepiness; behavioral or learning problems; or hyperactivity makes the diagnosis.

Comorbidities. In adults, obstructive sleep apnea is associated with many serious and chronic illnesses. A review of over four million records from the Veterans Health Administration (VHA) from 1998 to 2001 found a 2.91 percent (118,105) prevalence of sleep apnea in that population. Comorbid diagnoses included hypertension (60.1 percent), obesity (30.5 percent), diabetes mellitus (32.9 percent), cardiovascular disease including angina and myocardial infarction (27.6 percent), heart failure (13.5 percent), and stroke including transient ischemic attacks (5.7 percent). Psychiatric comorbidity in the sleep apnea group was significantly higher ($p < 0.0001$) than in the nonsleep apnea group for the diagnoses of mood disorders, anxiety and posttraumatic stress disorder, psychosis, and dementia. There are several theories about reason for this association. Psychiatric disorders may be a consequence of sleep apnea (and disturbed sleep and hypoxemia). Conversely, psychiatric disorders may predispose people to developing sleep disturbances such as sleep apnea. Pediatric obstructive sleep apnea is associated with neurobehavioral and cardiovascular problems. Obesity and growth abnormalities are also reported.

Treatments. Interventions available for treating OSA, include (a) positive airway pressure (PAP), (b) oral appliances, (c) positional therapy, (d) surgical intervention, and (e) weight loss. Finally, drug therapies have

been tried for OSA but without success. Medroxyprogesterone acetate was originally thought to be helpful but is seldom used now. Likewise, tricyclic antidepressants sometimes decrease apnea severity by reducing REM sleep, the stage of sleep in which obstructive apnea is usually more frequent. Theophylline was reported to reduce apnea but further study is needed. There were also animal studies suggesting mirtazapine and similar serotonin presynaptic affecting compounds improved breathing; however, studies in humans were disappointing. The only drug therapy approved for use in patients with OSAH is the wake promoting substance modafinil. Modafinil, however, does nothing to treat the pathophysiology of airway occlusion but rather is used as an adjunct for treating the residual sleepiness that persists in about 8 to 12 percent of patients who are otherwise well treated with and adequately use PAP therapy.

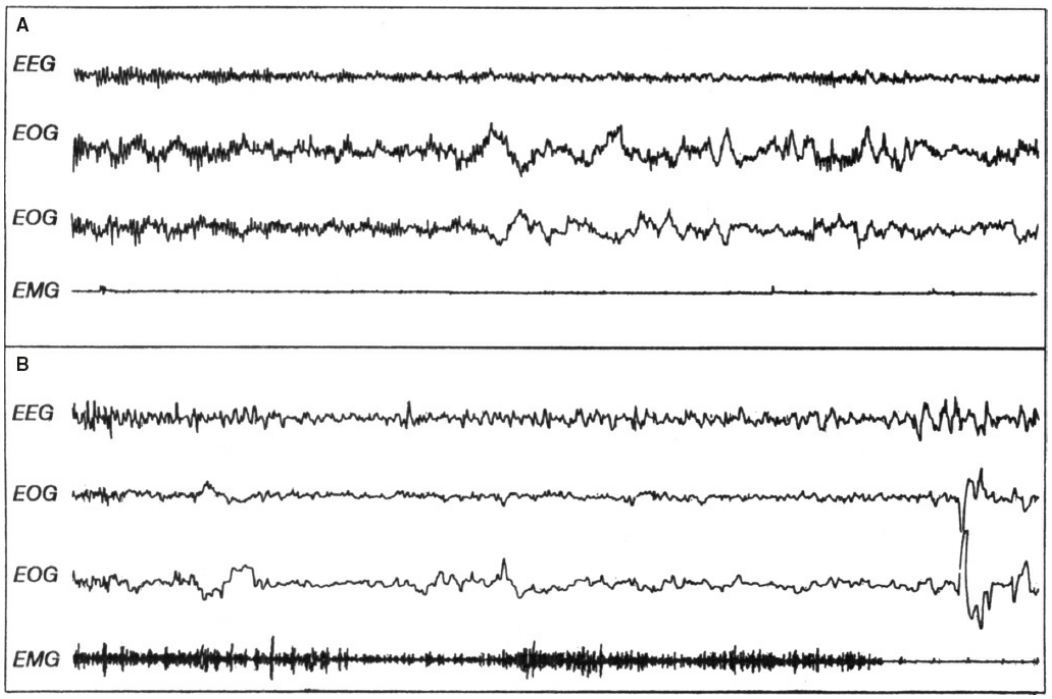

FIGURE 23–3. Polysomnographic tracing comparing normal sleep onset with that of a patient with narcolepsy. Each panel illustrates approximately 30 seconds of polysomnographic recording beginning with relaxed wakefulness. **Panel A** (normal sleep progression) shows a reduction of EEG alpha activity and development of slow rolling eye movements. **Panel B** shows the normally expected abatement of EEG alpha activity associated with increased theta activity and the appearance of a few slow eye movements. However, within 25 seconds (far right of figure) a swift loss of muscle tone occurs accompanied by rapid eye movements. This appearance of sleep-onset REM sleep (SOREM) characterizes narcolepsy and is part of the diagnostic criteria.

POSITIVE AIRWAY PRESSURE. Currently, PAP is the preferred treatment for OSAH. The PAP apparatus consists of a fan-driven blower, a nasal mask or oronasal interface, and tubing connecting the two. The airflow through the interface (usually a mask or nasal pillows) provides a positive pressure that offsets oropharyngeal collapse produced by inspiratory negative thoracic

pressure. In this manner it acts as a pneumatic splint, thereby maintaining the airway. When the pressure is properly titrated, even the most severe sleep apnea can usually be alleviated. Results are often dramatic. PAP devices come in several varieties. The most common are systems providing a preset continuous positive airway pressure (CPAP). For individuals who find it difficult to exhale against a continuous pressure, bi-level positive airway pressure (BPAP) may provide a solution. BPAP devices have different inspiratory and expiratory pressure settings. More recently, systems that sense the patient's changes in airway resistance and automatically adjust the positive airway pressure (APAP) have been gaining popularity. Such APAP systems should theoretically be able to adapt to changes in pressure requirements produced by sleep deprivation, medications, weight change, sleep depth, illness, and aging. Finally, timed bi-level and servo-ventilation systems have also been developed but fall into the category of non-invasive positive pressure ventilation (non-NIPPV) systems that are more appropriate for treating other pulmonary diseases and breathing problems in neuromuscular diseases. The tremendous efficacy and remarkable safety of PAP therapies have made it the standard of care for patients that can tolerate sleeping with the machine (see Fig. 23–6). The major therapeutic challenge is utilization. Patient education and systematic follow-up are crucial. When problems with the interface, the pressure, nasal stuffiness, and other barriers to routine, nightly use arise, they must be remedied quickly to assure therapeutic adherence. When used properly, PAP therapy's success has rendered surgical intervention a secondary option, resorted to mainly after PAP failure, rejection, or nonadherence.

ORAL APPLIANCE. Oral appliances represent another therapeutic option that is gaining popularity. A variety of oral appliances have also been developed to treat snoring and can provide benefit for patients with mild-to-moderate OSAH. The general approach is to manipulate the position of the mandible, lift the palate, and/or retain the tongue. Randomized trials indicate that some oral appliances sufficiently improve airway patency. However, in patients with severe OSA, improvement does not always reach satisfactory levels; therefore, follow-up evaluations are critically important.

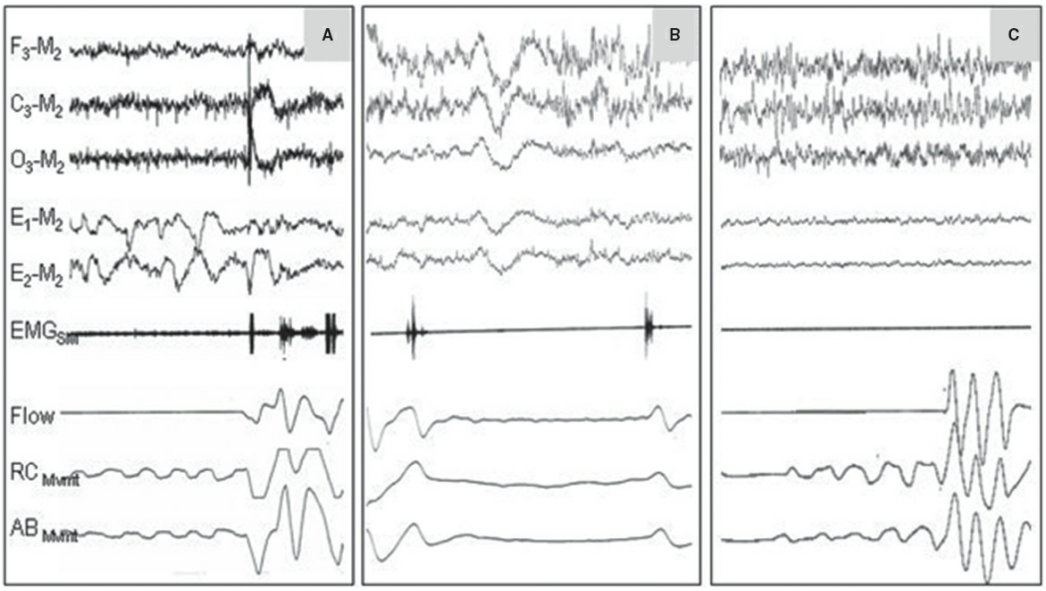

FIGURE 23–4. Polysomnographic tracings illustrating sleep apnea episodes. **Panel A** illustrates an obstructive sleep apnea, **Panel B** shows a central sleep apnea, and **Panel C** depicts a mixed sleep apnea. Note the presence of respiratory effort (rib cage and abdominal movement), notwithstanding cessation of airflow, throughout the obstructive apnea and during the latter portion of the mixed apnea. By contrast, respiratory effort is absent during the flow cessation throughout the central apnea and during the early portion of the mixed apnea. Notation: F_3 (left frontal scalp derivation); M_2 (right mastoid electrode location); C_3 (left central scalp derivation); O_3 (left occipital scalp derivation); E_1 (left outer canthus eye electrode location); E_2 (right outer cantus eye electrode location); EMG_{SM} (surface electrode electromyogram from submentalis muscle); Flow (airflow measured with nasal–nasal–oral thermistors); RC Mvmt (rib cage movement measured with inductive pleythsmography); AB Mvmt (abdominal movement measured with inductive pleythsmography). (From Hirshkowitz M, Sharafkhaneh A. Diagnostic assessment methods in adults. In: Kryger MH, Avidan AY, Berry RB, eds. *Atlas of Clinical Sleep Medicine*. 2nd ed. Elsevier Saunders: Philadelphia, PA; 2014, p. 378.)

POSITIONAL THERAPY. In some cases, obstructed breathing occurs only when the patient sleeps supine. In such situations, preventing patients from sleeping on their backs may produce beneficial results. Tennis balls sewn onto or placed into pockets on the back of the nightshirt or foam wedges may accomplish this goal. Although such interventions are found useful in clinical practice, large-scale systematic clinical trials are needed.

SURGICAL INTERVENTION. Aggressive surgical treatments evolved soon after OSAH's pathophysiological and potentially life-threatening consequences were recognized. The earliest surgical intervention was designed to create a patent airway; thus, in the late 1970s tracheostomies were performed on individuals with severe apnea. There is little doubt that tracheostomy succeeds in creating an airway. Although no longer the preferred treatment, it remains a standard against which newer, more refined therapies are judged. Second-generation surgical approaches attempted to correct airway obstructions and malformations. Early studies of uvulopalatopharyngoplasty (UPPP) suggested that modification of soft

palate effectively relieved most sleep apnea. Later follow-up results were less impressive. Approximately 30 to 50 percent of patients with sleep apnea benefit from UPPP. These patients are likely those with oropharyngeal obstruction; thus, careful attention to selection criteria presumably improves outcome. However, if obstruction occurs in the posterior airway space (PAS), maxillomandibular surgery may be appropriate. In retrognathic patients or patients with cephalometrics revealing compromised PAS, moving the jaw forward can achieve impressive normalization of breathing during sleep.

WEIGHT LOSS. Weight loss is known to help many patients. However, because losing weight and maintaining weight loss is difficult and unreliably achieved, the prudent clinician should recommend weight loss but rely on other therapies.

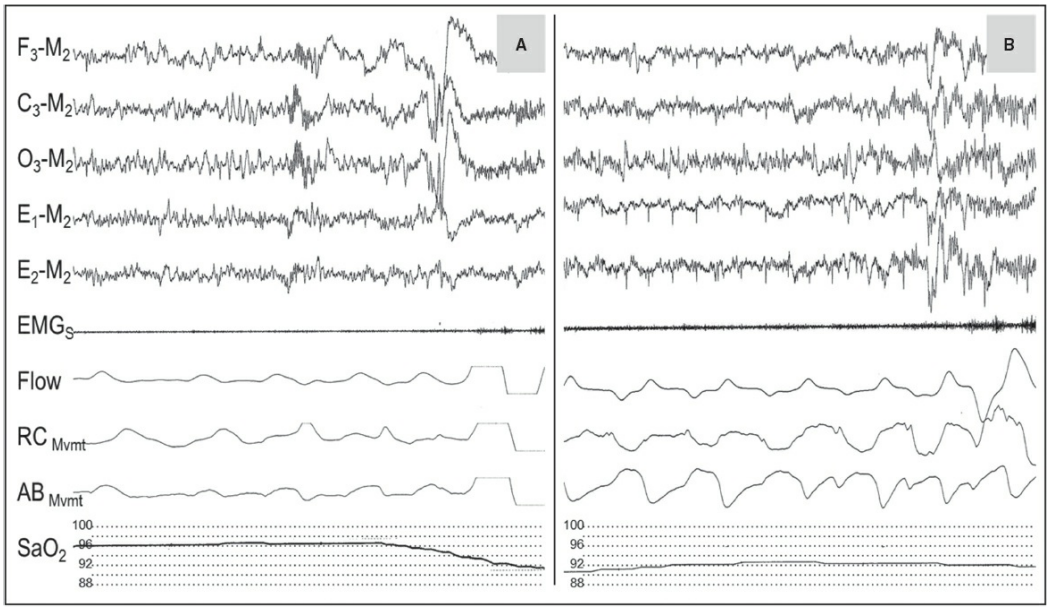

FIGURE 23–5. Sleep hypopnea and respiratory effort–related arousal (RERA) episodes. **Panel A** illustrates an obstructive hypopnea, **Panel B** shows a RERA. Note the presence of respiratory effort (rib cage and abdominal movement), notwithstanding reduced of airflow. The hypopnea episode is accompanied by significant oxyhemoglobin desaturation as the event progresses. By contrast, the RERA does not provoke desaturation; however, a CNS arousal occurs followed by increasing ventilation. Notation: F_3 (left frontal scalp derivation); M_2 (right mastoid electrode location); C_3 (left central scalp derivation); O_3 (left occipital scalp derivation); E_1 (left outer canthus eye electrode location); E_2 (right outer cantus eye electrode location); EMG_{SM} (surface electrode electromyogram from submentalis muscle); Flow (airflow measured with nasal–nasal–oral thermistors); RC Mvmt (rib cage movement) measured with inductive pleythsmography; AB Mvmt (abdominal movement) measured with inductive pleythsmography; SaO_2 (pulse oximetry) shows oxygen saturation percentage.

Central Sleep Apnea

The key characteristic linking the various central sleep apnea syndromes is that breathing cessations are not provoked by airway flow limitations or

obstructions. Breathing pauses due to lack of respiratory effort. Several different etiologies can produce respiratory effort, including: heart failure, high altitude, brainstem lesions, metabolic conditions, CNS depressant, congenital abnormalities, and PAP treatment.

Diagnostic Criteria. For diagnosing Central Sleep Apnea, DSM-5 requires polysomnographic evidence of central apneas per hour of sleep. Also, the disorder must not be better explained by another concurrent sleep disorder. Clinicians should specify whether the central apnea is associated with (a) Cheyne–Stokes Breathing, (b) comorbid opioid use, or (c) no evidence of airway obstruction. ICSD-3 describes eight diagnostic entities, as follows: (1) Central Sleep Apnea with Cheyne–Stokes Breathing, (2) Central Apnea Due to a Medical Disorder without Cheyne–Stokes Breathing, (3) Central Sleep Apnea Due to High Altitude Periodic Breathing, (4) Central Sleep Apnea Due to Medication or Substance, (5) Primary Central Sleep Apnea, (6) Primary Central Sleep Apnea of Infancy, (7) Primary Central Sleep Apnea of Prematurity, and (8) Treatment-Emergent Central Sleep Apnea. These conditions are described in the Central Apnea Subtypes section below.

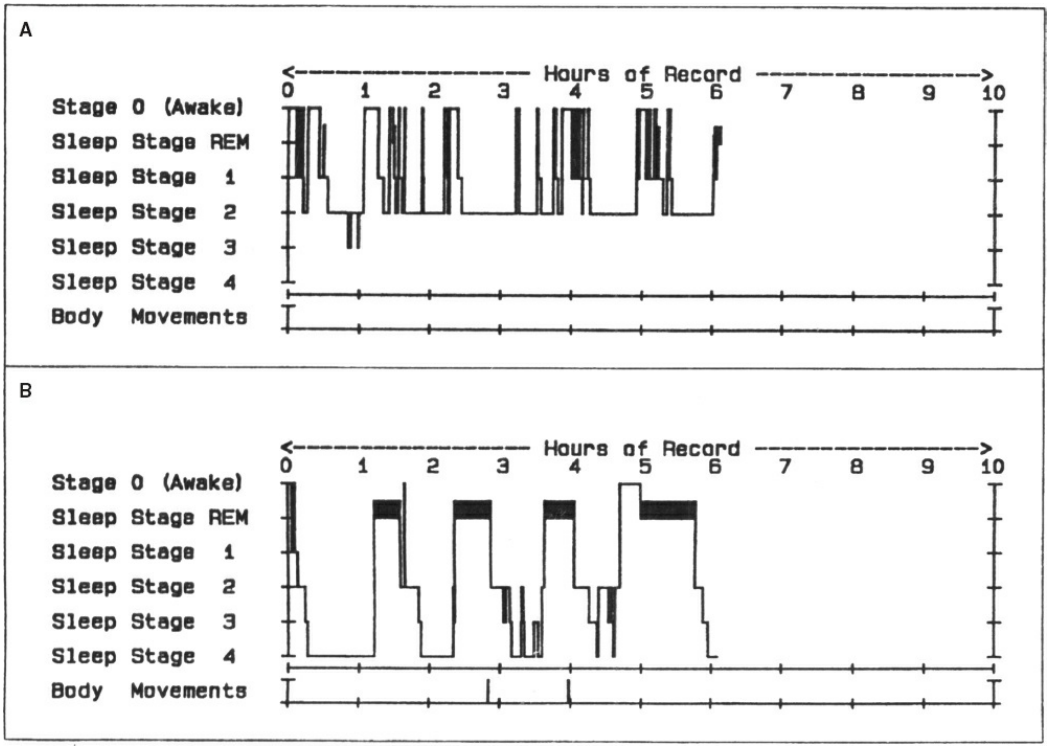

FIGURE 23–6. Sleep histogram showing CPAP-related sleep improvement. Sleep stage histogram illustrating the immediate, dramatic improvement in sleep architecture produced by treating obstructive sleep apnea with continuous positive airway pressure (CPAP) therapy. **Panel A** illustrates the abnormal sleep pattern on a night when the patient had more than 200 episodes of obstructive sleep apnea. Sleep is disturbed by frequent awakenings while REM and slow-wave (stages 3 and 4) sleep are nearly absent. **Panel B** shows data from the same patient being treated with CPAP on the next night. Normalization of sleep continuity with a massive rebound in REM and slow-wave sleep is

evident.

Central Apnea Subtypes Described in ICSD-3

CENTRAL SLEEP APNEA WITH CHEYNE–STOKES BREATHING. Cheyne–Stokes Breathing is characterized by a unique breathing pattern consisting of prolonged hyperpneas during which tidal volume gradually waxes and wanes in a crescendo–decrescendo fashion. The hyperpneas alternate with apnea and hypopnea episodes that are associated with reduced ventilatory effort. This pattern is most common in older men with congestive heart failure or stroke. As with Primary Central Sleep Apnea, the patients present with daytime sleepiness, insomnia, and awakening short of breath. Polysomnography reveals 5, or more, central apnea and hypopnea episodes per hour of sleep.

CENTRAL APNEA DUE TO A MEDICAL DISORDER WITHOUT CHEYNE–STOKES BREATHING. This form of central sleep apnea is usually caused by a brainstem lesion associated with a wide range of variable etiologies. Cardiac and renal disorders can also cause central apnea. Diagnostic criteria require polysomnographically verified rate of 5, or more, central apneas and hypopneas per hour of sleep.

CENTRAL SLEEP APNEA DUE TO HIGH ALTITUDE PERIODIC BREATHING. Central apnea at sleep onset is universal at elevations above 7,600 m but can occur at 1,500 m in some individuals (especially with a recent or rapid ascent). Periods of central apnea alternate with periods of hyperpnea in a 12- to 34-second cycle. This is an extension of normal respiratory control at sleep onset where medullary pH receptors raise their set-point and require lower pH to respond. At high altitudes, hyperventilation causes a hypocapnic alkalosis that reduces ventilation during sleep. Sleep architecture may suffer with increased duration in stages one and two and less slow-wave sleep. REM sleep may not be affected. This condition can be treated with acetazolamide which lowers serum pH and increases the respiratory drive. Acetazolamide side effects include metabolic acidosis, electrolyte imbalance, anaphylaxis, Stevens–Johnson syndrome, toxic epidermal necrolysis, and agranulocytosis. Common reactions include but are not limited to fatigue, anorexia, taste changes, polyuria, diarrhea, melena, tinnitus, and photosensitivity.

CENTRAL SLEEP APNEA DUE TO MEDICATION OR SUBSTANCE. Central apnea episodes can be provoked by a variety of drugs or drug combinations, most notably, long-acting opiates. Diagnostic criteria are a central apnea index (number of episodes per hour) of 5, or more per hour.

PRIMARY CENTRAL SLEEP APNEA. This is an idiopathic form of central sleep

apnea. Patients typically have low normal $PaCO_2$ while awake and have a high ventilatory response to CO_2. They may present with daytime sleepiness, insomnia, and/or awakening short of breath. Respiratory cessations during sleep occur independent of ventilatory effort. Polysomnography reveals 5, or more, central apneas per hour of sleep.

PRIMARY CENTRAL SLEEP APNEA OF INFANCY. This form of central sleep apnea involves prolonged apneas or hypopneas with concomitant hypoxemia, bradycardia, or both. This condition afflicts preterm neonates presumably because their brainstems are not fully developed. The condition may be exacerbated by other medical problems that further compromise the infant's physiological and developmental status.

PRIMARY CENTRAL SLEEP APNEA OF PREMATURITY. Central apnea, cyanosis, significant oxyhemoglobin desaturation, or bradycardia may be detected during postnatal hospital monitoring in infants less than 37 weeks conceptual age. Polysomnography or home monitoring must show either (i) recurrent 20 seconds, or longer, central apnea or (ii) periodic breathing for 5 percent, or more, of total sleep time.

TREATMENT-EMERGENT CENTRAL SLEEP APNEA. In patients with obstructive sleep apnea who are treated with PAP, central apnea episodes sometimes emerge. If the central apnea persists over time and lowering the pressure permits obstructions to re-occur (i.e., lower pressures are not adequate to treat the obstructive apnea) the patient is considered to have treatment-emergent central sleep apnea.

Sleep-Related Hypoventilation

Sleep-related hypoventilation is commonly associated with chronic obstructive pulmonary disease, obesity, and neuromuscular disease. However, other comorbid medical conditions may contribute. ICSD-3 differentiates hypoventilation disorders from hypoxemia disorders and further subdivides each. In hypoventilation disorders, ICSD-3 includes (1) obesity hypoventilation syndrome, (2) congenital central alveolar hypoventilation syndrome, (3) late-onset central hypoventilation and hypothalamic dysfunction, (4) idiopathic central alveolar hypoventilation, (5) sleep-related hypoventilation due to a medication or substance, and (6) sleep-related hypoventilation due to a medical disorder. ICSD-3 also lists (7) sleep-related hypoxemia as a separate nosologic entity. Snoring and catathrenia (sleep-related groaning) are considered isolated symptoms and normal variants.

Diagnostic Criteria. Polysomnography must demonstrate either (a) episodes with elevated CO_2 levels and decreased respirations and/or (b) if

CO_2 is not monitored, episodes of persistently low oxyhemoglobin levels that are not provoked by apnea or hypopnea. Clinicians should specify if the hypoventilation is attributable to (i) congenital central alveolar hypoventilation, (ii) comorbid sleep-related hypoventilation (e.g., COPD), or if it is (iii) idiopathic hypoventilation. Severity should also be specified related to evidence of related diseases (e.g., right heart failure) or blood gasses (if available).

CIRCADIAN RHYTHM SLEEP–WAKE DISORDERS

This collection of sleep disorders shares the same basic underlying etiology; that is, desynchrony between an individual's internal circadian biological clock and the desired sleep–wake cycle. The circadian (circa + dias, approximately 1 day) pacemaker is located in the suprachiasmatic nucleus (SCN). SCN firing oscillate with an almost sinusoidal pattern, the period of which is 24 hours and the output correlates with the daily fluctuations in core body temperature. SCN firing patterns persists even in *isole* preparations. Mismatched circadian clock and desired schedules can arise from improper phase relationships between the two, travel across time zones, or dysfunctions in the basic biological rhythm. DSM-5 provides six specific subtype coding for disorders listed as separate ICSD-3 diagnostic entities. Jet Lag Disorder is also listed in ICSD-3 but not in DSM-5.

Diagnostic Criteria

To diagnose a Circadian Sleep–Wake Disorder, an individual must (1) have an alteration in or misalignment between their endogenous circadian rhythm; (2) the alteration or misalignment must disrupt sleep producing insomnia and/or hypersomnolence, and (3) this circumstance must either elicit significant distress or social, occupational, or other functional impairment. The clinician should specify if the Circadian Sleep–Wake Disorder subtype constitutes (a) delayed sleep phase type, (b) advanced sleep phase type, (c) irregular sleep–wake type, (d) non–24-hour sleep–wake type, (e) shift work type, or (f) unspecified type. The diagnosis should also indicate whether the problem is (a) episodic with symptoms lasting 1 month or more but less than 3 months; (b) persistent with symptoms lasting 3 months, or more; or (c) recurrent with two (or more) episodes within a 1-year period.

Circadian Rhythm Sleep–Wake Disorder Subtypes

Delayed Sleep Phase Type. The Delayed Sleep Phase circadian disorder occurs when the biological clock run slower than 24 hours or is shifted later than the desired schedule. This produces a phase delay in the sleepiness–alertness cycle. Individuals with delayed sleep phase are more alert in the evening and early nighttime, stay up later, and are more tired in the

morning. These individuals are commonly referred to as *night owls*.

Advanced Sleep Phase Type. Advanced sleep phase is when the circadian rhythm cycle is shifted earlier. Therefore, the sleepiness cycle is advanced with respect to clock time. Individuals with advanced sleep phase are drowsy in the evening, want to retire to bed earlier, awaken earlier, and are more alert in the early morning. Individuals with this pattern of advanced sleep phase are sometimes called *early birds or larks* (sometimes jokingly called *surgeons* in the medical center).

Irregular Sleep–Wake Type. The irregular sleep–wake pattern occurs when the circadian sleep–wake rhythm is absent or pathologically diminished. The sleep–wake pattern is temporally disorganized and the timing of sleep and wakefulness is unpredictable. Individuals with this condition will have a normal amount of sleep during a 24-hour period; however, it will be fragmented into three or more episodes that occur irregularly. Long daytime naps and inappropriate nocturnal wakefulness will occur. Except in unusual circumstances, activities of daily life are significantly impaired.

Non–24-Hour Sleep–Wake Type. When the circadian sleep–wake pacemaker has a cycle length greater or less than 24 hours and is not reset each morning, a patient may develop this type of circadian rhythm disorder. Under normal circumstances, resynchronization of circadian rhythm occurs daily in response to the light-dark cycle. Problems occur incrementally when internal and environmental clocks become more and more out of phase. If the circadian clock's period is longer than 24 hours and does not reset each day, the patient experiences progressively worsening sleep-onset insomnia and daytime sleepiness. Sleep problems peak when circadian and environmental clocks are 12 hours out of phase and then begin to lessen, emulating progressively resolving advanced sleep phase. Eventually the clocks correlate and sleep–wake cycle is normal for a few days, after which the insomnia–hypersomnia cycle begins again. For this reason, non–24-hour sleep–wake disorder has been called periodic insomnia and periodic excessive sleepiness.

Shift Work Type. Many service industries require 24-hour operation (e.g., transportation, health care). Similarly, as cultures became more capital intensive, mining and manufacturing became around-the-clock enterprises. The number of individuals doing shift work has been increasing steadily for decades. Shift workers commonly suffer from insomnia, excessive sleepiness, or both. Some individuals require only a short time to adjust to a shift change while others have great difficulty. Frequent shift rotation adds to the problems. Also, to meet social demands,

shift workers will often adopt a nonshifted sleep–wake schedule on weekends and holidays. Even those individuals who try to stay shifted usually retain an unshifted circadian rhythm. The result can be severe insomnia when attempting to sleep and excessive sleepiness when attempting to remain awake. The result is profound sleep deprivation as circadian rhythm continues to be poorly entrained to the sleep–wake schedule. The natural low point in the normal sleep–wake rhythm occurs at approximately 3:00 to 5:00 AM. Interestingly, this is precisely the time frame during which transportation and industrial accidents commonly occur as a direct consequence of sleepiness.

Jet Lag (Not Included in DSM-5). With the advent of high speed air travel, an induced desynchrony between circadian and environmental clocks became possible. Thus, the term jet lag came into use. When an individual rapidly travels across many time zones they induce either a circadian phase advance or a phase delay, depending upon the direction of travel. Typically, one or two time zone translocation will not produce a sustained problem; however, overseas travel can be marked by great difficulty in adjusting one's sleep–wake routine. Individuals who frequently travel for business can find themselves quite impaired at the time they need to make important decisions. Furthermore, "owls" will experience greater difficulty adjusting to eastward travel because resynchronization requires phase advance. Similarly, larks theoretically will have more difficulty with westward travel. The number of time zones crossed is a critical factor. Normally, healthy individuals can easily adapt to one to two time zone changes per day; therefore, natural adjustment to an 8-hour translocation may take 4 or more days.

Circadian Rhythm Sleep–Wake Disorder Unspecified Type. During illnesses that keep patients bedridden, during hospitalizations, and in some forms of dementia, individuals often sleep *ad lib*. The resulting chaotic sleep–wake pattern adversely affects the circadian rhythm. The breakdown in the sleep–wake cycle may be further exacerbated by medication with sedative properties. Sleep in patients in the ICU is disturbed by noise, light, and the therapeutic and monitoring procedures being performed. The resulting disorganized sleep–wake pattern can produce a significant sleep disorder. Additionally, abuse of recreational street drugs (e.g., methamphetamine and ecstasy) is associated with individuals remaining awake overnight or continuously for several days at a time. These episodes of prolonged wakefulness ultimately produce periods of profound hypersomnia.

Treating of Circadian Rhythm Sleep–Wake Disorders

Under normal circumstances the internal circadian pacemaker is reset each

day by bright light, social cues, stimulants, and activity. In cases where these factors fail to reentrain the circadian rhythm, circadian sleep–wake disorders occur.

Light Therapy. Research indicates that light exposure can reset the circadian pacemaker. This seems especially the case when the light is bright (greater than 10,000 lx) or in the blue spectrum. With precise timing of bright light exposure, the biological clock can be reset. Exposure to light modifies the set-point of the biological clock. Using core body temperature as a physiological marker, bright lights produce phase delay when presented before the temperature nadir. By contrast, light exposure after temperature nadir evokes phase advance. The closer one presents light to the point of inflection (temperature nadir), the more robust the response in altering the cycle. Thus, early morning bright light therapy can be used to phase advance individuals with delayed sleep phase syndrome. Similarly, exposure to bright light in the evening can help patients with advanced sleep phase syndrome. More recently, it was discovered that the blue spectrum of light is the crucial ingredient in phase setting and shifting. Therefore, blue light exposure (or limitation with blue-blocking eyewear) at specific times can provide therapeutic benefit. Clinicians increasingly relies on light therapy to shift circadian sleep–wake rhythms of individuals with delayed and advanced sleep phase; shift workers; astronauts; and individuals experiencing jet lag.

Medication. Administering melatonin to individuals who are blind-from-birth successfully entrains their circadian sleep–wake rhythm. Researchers posit that melatonin secretion acts as the biological substrate for the internal circadian oscillator. Under normal circumstances, melatonin levels begin to rise at dusk and remain elevated until dawn. Bright light suppresses the release of melatonin. Melatonin, in a sense, is the signal of darkness in the brain. As such, it can be used clinically to manage sighted patients with disturbed sleep–wake cycles. Melatonin is available OTC in the United States. A prescription synthetic melatonin agonist (Ramelteon) is also available (FDA approved for treating patients with sleep-onset insomnia). In Europe, a 2-mg sustained-release melatonin (Circadin) is available in many European countries, Australia, and some places in Asia. Interestingly, the wake-promoting compound modafinil is FDA approved for treating sleepiness in patients with shift work type circadian rhythm sleep–wake disorder.

Chronotherapy. Chronotherapy involves progressively phase delaying a person until the circadian oscillator is synchronized with the desired sleep–wake schedule. Most young and middle-aged adults have a propensity to progressively phase delay their circadian sleep–wake rhythm

anyway. Thus, phase delaying each night by 2 or 3 hours is thought to be easier than phase advancing because it capitalizes on a natural tendency. Halting the phase delay at the appropriate moment and maintaining the desired synchrony can be a challenge. The patient also has to cope with an odd sleep–wake schedule for the better part of a week during therapy (which can interfere with school or work). For these reasons, the development of light therapy and pharmacological interventions has overshadowed chronotherapy in the past two decades.

PARASOMNIAS AND SLEEP-RELATED MOVEMENT DISORDERS

DSM-5 includes only three of the ten specific parasomnias described in ICSD-3. Furthermore, DSM-5 includes Restless Legs Syndrome (RLS) which in ICSD-3 is classified as one of the seven specific Sleep-Related Movement Disorders.

Parasomnias are sometimes referred to as disorders of partial arousal. In general, the parasomnias are a diverse collection of sleep disorders characterized by physiological or behavioral phenomena that occur during or are potentiated by sleep. One conceptual framework posits many parasomnias as overlap or intrusion of one basic sleep–wake state into another. Wakefulness, NREM sleep, and REM sleep can be characterized as the three basic states that differ in their neurological organization. During wakefulness, both the body and brain are active. In NREM sleep, both the body and brain are inactive. REM sleep, however, pairs an atonic body with an active brain (capable of creating elaborate dream fantasies). Regional cerebral blood flow, MRI, and other imaging studies confirm increased brain activation during REM sleep. It certainly appears that in some parasomnias there are state boundary violations. For example, the arousal disorders (sleep walking and sleep terrors) involve momentary or partial wakeful behaviors suddenly occurring in NREM sleep. Similarly, the ICSD-3 nosologic entity Isolated Sleep Paralysis is the persistence of REM sleep atonia into the wakefulness transition while REM sleep behavior disorder is the failure of the mechanism creating paralytic atonia such that individuals literally act out their dreams.

The occurrence frequency of significant parasomnias varies. Clinical significance relates more to the medical consequences or the patient's evoked level of distress than with how often the abnormal events occur. Parasomnias' irregularities in occurrence make them difficult to document in the sleep laboratory. Sleep studies, however, are indicated for ruling out that the unusual behavior is secondary to a seizure disorder or a sleep-related breathing disorder.

Nonrapid Eye Movement Sleep Arousal Disorders

Sleepwalking. Sleepwalking in its classic form, as the name implies, is

a condition in which an individual arises from bed and ambulates without fully awakening. Sometimes called somnambulism, individuals can engage in a variety of complex behaviors while unconscious. Sleepwalking usually occurs during NREM slow-wave sleep and characteristically begins toward the end of the first or second slow-wave sleep episodes. Sleep deprivation and interruption of slow-wave sleep appear to exacerbate, or even provoke, sleepwalking in susceptible individuals. Sleepwalking episodes may range from sitting-up and attempting to walk to conducting an involved sequence of semi-purposeful actions. The sleepwalker often can successfully intact with the environment (e.g., avoiding tripping over objects in his path). However, the sleepwalker will often interact with the environment inappropriately, which sometimes results in injury (e.g., stepping out of an upstairs window or walking into the roadway). There are cases in which sleepwalkers have committed acts of violence. An individual who is sleepwalking is difficult to awaken. Once awake, the sleepwalker will usually appear confused. It is best to gently attempt to lead sleepwalkers back to bed rather than attempting to awaken them by grabbing, shaking, or shouting. In their confused state, the sleepwalker may think they are being attacked and may react violently to defend themselves. Sleepwalking in adults is rare, has a familial pattern, and may occur as a primary parasomnia or secondary to another sleep disorder (e.g., sleep apnea). By contrast, sleepwalking is very common in children with peak prevalence between ages 4 and 8 years. After adolescence it usually disappears spontaneously. Nightly to weekly sleepwalking episodes associated with physical injury to patient and others are considered severe.

Sleep Terrors. A sleep terror is characterized by a sudden arousal with intense fearfulness. They may begin with a piercing scream or cry (see Fig. 23–7). Autonomic and behavioral correlates of fright typically mark the experience. An individual experiencing a sleep terror usually sits up in bed, is unresponsive to stimuli, and if awakened is confused or disoriented. Vocalizations may occur but they usually are incoherent. Notwithstanding the intensity of these events, amnesia for the episodes usually occurs. Like sleepwalking, these episodes usually arise from slow-wave sleep. Fever and CNS depressant withdrawal potentiates sleep terror episodes. Unlike nightmares in which an elaborate dream sequence unfolds, sleep terrors may be devoid of images or contain only fragments of very brief but frighteningly vivid but sometimes static images. Sometimes called *Pavor nocturnus,* incubus, or night terror, a familial pattern has been reported. Like other slow-wave sleep parasomnias, sleep deprivation can provoke or exacerbate sleep terrors. Psychopathology is seldom associated with sleep terrors in children; however, a history of traumatic experience or frank psychiatric problems is often comorbid in adults with this disorder. Severity ranges from less than once per month to almost nightly occurrence

(with injury to patient or others).

Diagnostic Criteria. For Sleepwalking, DSM-5 requires repeated episodes of rising from bed and walking. The sleepwalker usually has a blank expression, is generally nonresponsive, and is difficult to awaken. For Sleep Terror, there must also be recurrent episodes where the individual awakens with intense fear and autonomic signs of distress. The patient is typically so panicked as to be inconsolable.

The clinician should specify if the NREM Parasomnia is: (a) sleepwalking, (b) sleepwalking type with sleep-related eating, (c) sleepwalking type with sleep-related sexual behavior (sexsomnia), or (d) sleep terror type.

Nightmare Disorder

Nightmares are frightening or terrifying dreams. Sometimes called dream anxiety attacks, they can produce sympathetic activation and ultimately awaken the dreamer. Nightmares occur in REM sleep and usually evolve from a long, complicated dream that becomes increasingly frightening. Having aroused to wakefulness, the dream content is typically remembered (in contrast to sleep terrors). Some nightmares are recurrent and reportedly when occurring in association with posttraumatic stress disorder they may be recollections of actual events. Common in children ages 3 to 6 (prevalence estimates range from 10 to 50 percent), nightmares are rare in adults (1 percent or less). Frequent and distressing nightmares are sometimes responsible for insomnia because the individual is afraid to sleep. In Freudian terms, the nightmare is an example of the failure of dream-process that defuses the emotional content of the dream by disguising it symbolically, thus preserving sleep. Most patients afflicted with nightmares are free from psychiatric conditions. Nonetheless, individuals at risk for nightmares include those with schizotypal, borderline, and schizoid personality disorders, as well as schizophrenia. Hartmann (1998) posits that nightmares are more common in individuals with “thin boundaries” who are open, trusting, and who often have creative or artistic inclinations. Having thin boundaries makes these individuals more vulnerable; furthermore, they may be at risk for schizophrenia. Traumatic events are known to induce nightmares, sometimes immediately but at other times delayed. The nightmares can persist for many years. Several medications are known to sometimes provoke nightmares, including L-DOPA, beta-adrenergic blockers, and withdrawal from REM suppressant medications. Finally, drug or alcohol abuse is associated with nightmares.

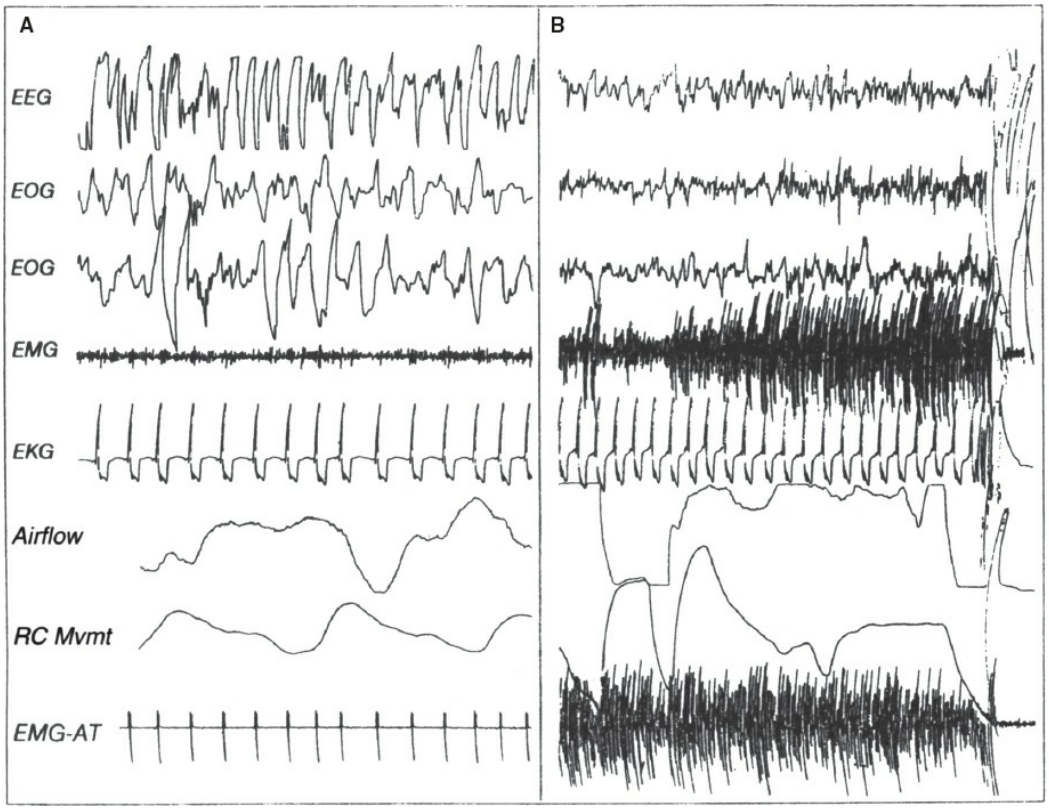

FIGURE 23–7. Polysomnographic tracing illustrating a sleep terror. **Panel A** shows approximately 14 seconds of tracing occurring immediately before the sleep terror. Prominent EEG slow-wave activity and other characteristics of stage 4 sleep are seen. **Panel B** shows the awakening, accompanied by tachycardia and movement. EEG activity is ambiguous and the patient eventually disconnected his electrodes as he thrashed about in bed (visible at far right of figure). Although the patient was screaming and greatly agitated, there was no report of dreaming. In the morning, there was little recollection of anything having occurred during the night.

Frequently occurring nightmares often produce a "fear of sleeping" type of insomnia. In turn, the insomnia may provoke sleep deprivation, which is known to exacerbate nightmares. In this manner, a vicious cycle is created. Treatment using behavioral techniques can be helpful. Universal sleep hygiene, stimulus control therapy, lucid dream therapy, and cognitive therapy reportedly improve sleep and reduce nightmares. BZD and BZRAs are sometimes use as treatment. Additionally, evidence for the use of prazosin—a CNS alpha-1 receptor antagonist, in the treatment of PTSD-related nightmares is growing. Prazosin significantly increased total sleep time, REM sleep time and significantly reduced trauma-related nightmares and distressed awakenings.

Rapid Eye Movement Sleep Behavior Disorder

REM sleep behavior disorder involves a failure of the patient to have atonia (sleep paralysis) during stage REM sleep. The result is that the patient literally enacts their dreams. Under normal circumstances, the dreamer is immobilized by REM-related hypopolarization of alpha and gamma motor

neurons. Without this paralysis or with intermittent atonia, punching, kicking, leaping, and running from bed during attempted dream enactment occur. The activity has been correlated with dream imagery and unlike during sleepwalking the individual seems unaware of the actual environment but rather is acting on their dream sensorium. Thus, a sleepwalker may calmly go to a bedroom window, open it, and step out. By contrast, a person with REM sleep behavior disorder would more likely dive through the window thinking it is a dream-visualized lake. Patients and bed partners frequently sustain injury, sometimes serious (e.g., lacerations, fractures). A wide variety of drugs and comorbid conditions can precipitate or worsen RBD (see section on Parasomnia due to Drug or Substance Use and Parasomnia due to Medical Conditions). In animals, presumed RBD can be produced with bilateral peri-locus coeruleus lesions. In humans, there is a suggestion that RBD may result from diffuse hemispheric lesions, bilateral thalamic abnormalities, or brainstem lesions. Clonazepam has been used successfully to treat RBD.

Restless Legs Syndrome

RLS is characterized by the irresistible urge to move the legs when at rest or while trying to fall asleep. Patients often report crawling feelings in their legs. Moving the legs or walking around helps alleviate the discomfort (see Fig. 23–8). Thus, as the individual is lying in bed and relaxing they are disturbed by these sensations. They then move their legs and again try to fall asleep. This cycle sometimes continues for hours and results in profound insomnia.

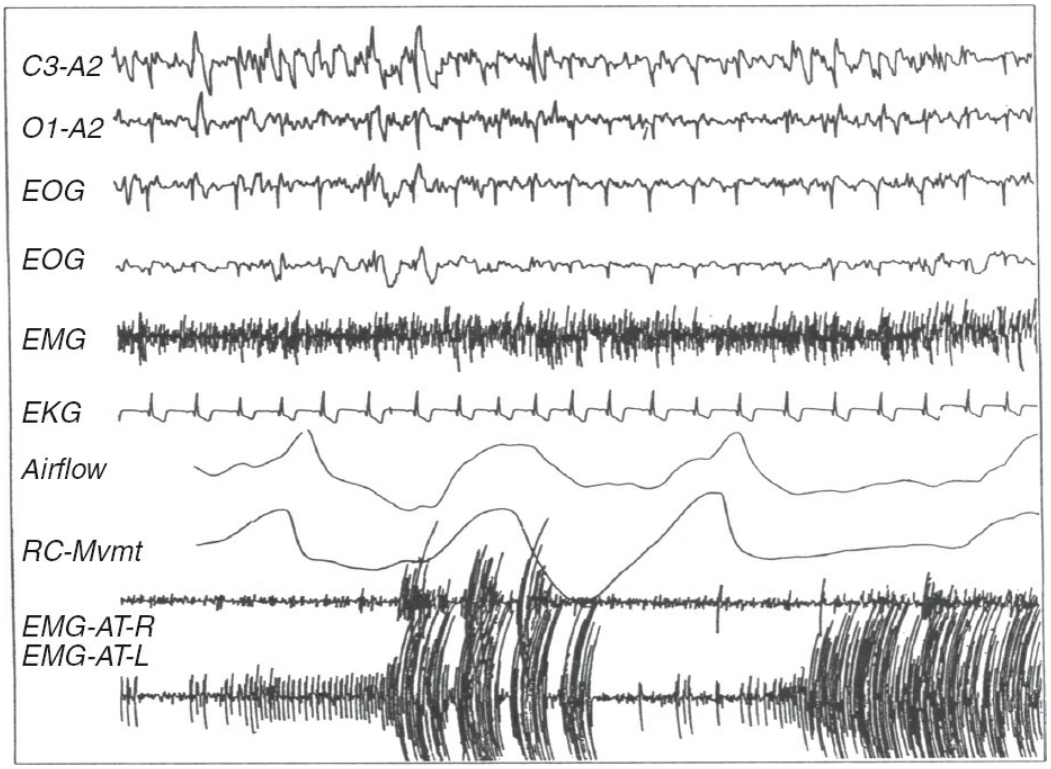

FIGURE 23–8. Restless legs syndrome. This patient presented with complaints of uncomfortable, crawling sensations in the legs when trying to fall asleep. Patients commonly report an urge to move the leg in order to dispel the sensation. This figure shows a bilateral pattern of leg EMG activity; however, the discharge is more pronounced in the left anterior tibialis (EMG-AT-L) than the right (EMG-AT-R). This pattern continued for more than an hour as the patient attempted to fall asleep; note that the sharp activity in central and occipital EEG (C3-A2 and O1-A2, respectively) and EOG is EKG artifact and not an EEG abnormality.

Diagnostic Criteria

(a) The patient must (i) feel an initial or worsening urge to move the legs during rest or inactivity; (ii) feel some or total relief of the sensation when moving the legs; and (iii) feel these urges more (or only) in the evening or at night compared with during the day.

(b) The symptoms above must occur at least 3-times weekly for 3, or more, months.

(c) The symptoms (i) produce significant distress or impairment; (ii) cannot be explained another medical, psychiatric, or medical condition; and (iii) are not drug (or substance) related.

An NIH workshop established criteria for the diagnosis of RLS (see Table 23–4). Uremia, neuropathies, and iron and folic acid deficiency anemias can produce secondary RLS. RLS is also reported in association with fibromyalgia, rheumatoid arthritis, diabetes, thyroid diseases, and chronic obstructive pulmonary disease. Detailed history and physical examination are important part of RLS workup. In addition, ferritin level should be checked.

Pharmacologically, the dopaminergic agonists pramipexole, rotigotine,

and ropinirole are FDA approved and represent the treatment of choice. These two compounds have evidence-based support. Nonetheless, other agents used to treat RLS include dopamine precursors (e.g., levedopa), BZDs, opiates, and antiepileptic drugs (e.g., gabapentin). Sometimes the alpha-2 agonist clonidine is also used. Nonpharmacologic treatment includes avoiding alcohol or nicotine use close to bed time, massaging the affected parts of the legs, taking hot baths, applying hot or cold to the affected areas, and moderate exercise.

Other ICSD-3 Parasomnias and Sleep-Related Movement Disorders

Confusional Arousals. ICSD-3 defines Confusional Arousals as a milder form of NREM Sleep Parasomnias. It is common in young children. The child will typically partially awaken from sleep and sit up. The episodes are marked by confusion but usually the child lies back down and resumes sleep. It is thought that confusional arousals, sleepwalking, and sleep terrors lie on a continuum.

Table 23–4.
International Study Group Diagnostic Criteria for Restless Legs Syndrome

Essential features	An urge to move legs, usually accomplished or caused by uncontrollable or unpleasant sensations in the legs
	The urge to mover, or unpleasant sensations that begin or worsen during periods of rest or inactivity such as lying or sitting
	The urge to move or unpleasant sensations are partially or totally relieved by movement, such as walking or stretching, at least as long as the activity continues
	The urge to move or unpleasant sensations are worse in evening or night than during the day or only occur in the evening or night
Common features	Positive family history: 3–5 times higher prevalence of RLS among first-degree relatives
	Response to dopaminergic therapy
	PLM during sleep or wakefulness
	Clinical course: middle-age onset and usually with progressive course
	Sleep disturbance resulting in sleep-onset insomnia or daytime sleepiness
	No abnormality in primary form of the syndrome

Sleep-Related Eating Disorder. This diagnostic entity is subsumed under the DSM-5 classification NREM Sleep Parasomnias and specified as "sleepwalking type with sleep-related eating." In adults, nocturnal eating can be conditioned to awakening. Eating may become obsessional and several small meals may be eaten during the course of one night. The individual may be unaware of the activity and weight gain can become a problem. Essentially, this syndrome's ICSD-3 diagnostic criteria are as follows:

(a) The patient has recurrent sleep-related eating episodes.

(b) The patient consumes peculiar foods, food combinations, or toxic substances; creates a hazard or is injured during food preparation (e.g., starting a fire on the stove); or the patient suffers adverse health consequences from the sleep-related eating (e.g., substantial weight gain).
(c) The eating is associated with partial or complete amnesia and general lowering or loss of conscious awareness.
(d) The eating is not better explained by another disorder, medications, or other substance use.

Recurrent Isolated Sleep Paralysis. Sleep paralysis is, as the name implies, an inability to make voluntary movements during sleep. ICSD-3 lists it as a parasomnia when it occurs at sleep onset or upon awakening, a time when the individual is partial conscious and aware of the surroundings. This inability to move can be extremely distressing, especially when it is coupled with the feeling that there is an intruder in the house or when hypnagogic hallucinations are occurring. Sleep paralysis is one of the tetrad of symptoms associated with narcolepsy; however, it is known to occur (with or without hypnagogia) in individuals that have neither cataplexy nor excessive daytime sleepiness. Although sometimes frightening, sleep paralysis is a feature of normal REM sleep briefly intruding into wakefulness. The paralysis may last from one to several minutes. It is interesting that the occurrence of sleep paralysis with hypnagogia may account for a variety of experiences where the sleeper is confronted or attacked by some sort of creature. The common description is that a "presence" was felt to be near, the individual was paralyzed, the creature talks, attacks, or sits on the sleeper's chest, and then vanishes. Whether it is called incubus, "Old Hag," vampire, ghost oppression (kanashibari), witch riding, or alien encounter, elements common to sleep paralysis are seen. Irregular sleep, sleep deprivation, psychological stress, and shift work are thought to increase the likelihood of sleep paralysis occurring. Occasional sleep paralysis occurs in 7 to 8 percent of young adults. Estimates of at least one experience of sleep paralysis during the lifetime range from 25 to 50 percent. Improved sleep hygiene and assuring sufficient sleep are first-line therapies. Sometimes, if the individual voluntarily makes very REMs or is touched by another person, the episode usually will terminate.

Exploding Head Syndrome. Individuals with this parasomnia "hear" a loud imagined noise or a sense of a violent explosion in the head just as they are about to fall asleep or during a nocturnal awakening. The experience can occur just once or recurrently. There is no pain associated with the noise but the individual may be concerned that they are having a stroke or that something is very wrong. Even a single episode can trigger

for severe insomnia. There are no known neurological consequences to this syndrome.

Sleep-Related Hallucinations. Sleep-Related Hallucinations are typically visual images occurring at sleep onset (hypnagogic) or on awakening (hypnopompic) from sleep. Sometimes difficult to differentiate from dreams, they are common in patients with narcolepsy. Complex hallucinations are rare and usually happen with abrupt awakening and without remembrance dreaming. Images tend to be vivid and immobile and persist for several minutes (usually disappearing when a light is turned on). The images can be frightening.

Sleep Enuresis. Sleep enuresis is a disorder in which the individual urinates during sleep while in bed. In adults, sleep enuresis often accompanies severe obstructive sleep apnea and in such cases it would fall into the category of parasomnia related to a medical disorder. In children, bedwetting, as it is commonly called, has primary and secondary forms. In children, primary sleep enuresis is the continuance of bedwetting since infancy. Secondary enuresis refers to relapse after toilet training was complete and there was a period where the child remained dry. Usually, after toilet training bedwetting spontaneously resolves before age 6 years. Prevalence progressively declines from 30 percent at age 4, to 10 percent at age 6, to 5 percent at age 10, and 3 percent at age 12 years. Parental primary enuresis increases the likelihood that the children will also have enuresis. A single recessive gene is suspected. Secondary enuresis in children may occur with the birth of a sibling and represent a "cry for attention." Secondary enuresis can also be associated with nocturnal seizures, sleep deprivation, and urological anomalies. In adults, sleep enuresis is sometimes seen in patients with sleep-disordered breathing. In most cases, embarrassment, shame, and guilt are the most serious consequence. Nonetheless, if sleep enuresis is not addressed it may leave psychosocial scars.

Treating sleep enuresis involves medication, behavioral interventions, or both. Among medications, desmopressin acetate is the preferred choice. An evidence-based medicine review of 47 randomized trials found efficacy with desmopressin. Other medications used to treat sleep enuresis include imipramine and oxybutynin chloride. Behavioral treatments, including bladder training, using conditioning devices (bell and pad), diet modification, and fluid restriction have reportedly good success when properly administered. Other treatments include psychotherapy, motivational strategies, and hypnotherapy.

Parasomnia due to a Medical Disorder. Seizure disorder should always be on the top of a differential diagnosis list for most parasomnias. In fact,

the AASM practice guidelines concerning the indications for polysomnography includes using sleep testing to rule out seizures when diagnosing sleep terror, sleep walking, RBD, nightmares, and other parasomnias (see Fig. 23–9). Sleep-related breathing disorders are also known to trigger sleep walking, enuresis, sleep terror, confusional arousal, and nightmares. RBD is associated with a variety of neurological conditions, including Parkinson disease, dementia, progressive supranuclear palsy, Shy–Drager syndrome, narcolepsy, and others.

Parasomnia due to Drug or Substance Use. Many drugs and substances can trigger parasomnias. Particularly those compounds that lighten sleep; however, alcohol is notorious for producing sleepwalking (especially in individuals who have taken sleeping pills). RBD can be provoked or worsened by biperiden, tricyclic antidepressants, monoamine oxidase inhibitors, caffeine, venlafaxine, selegiline, and serotonin agonists. RBD may also occur during withdrawal from alcohol, meprobamate, pentazocine, and nitrazepam. Medications known provoke nightmares include L-DOPA and beta blockers. Nightmares can also be caused by drug-induced REM sleep rebound (e.g., withdrawal from REM suppressing drugs like methamphetamine), beta blockers, and alcohol abuse or withdrawal.

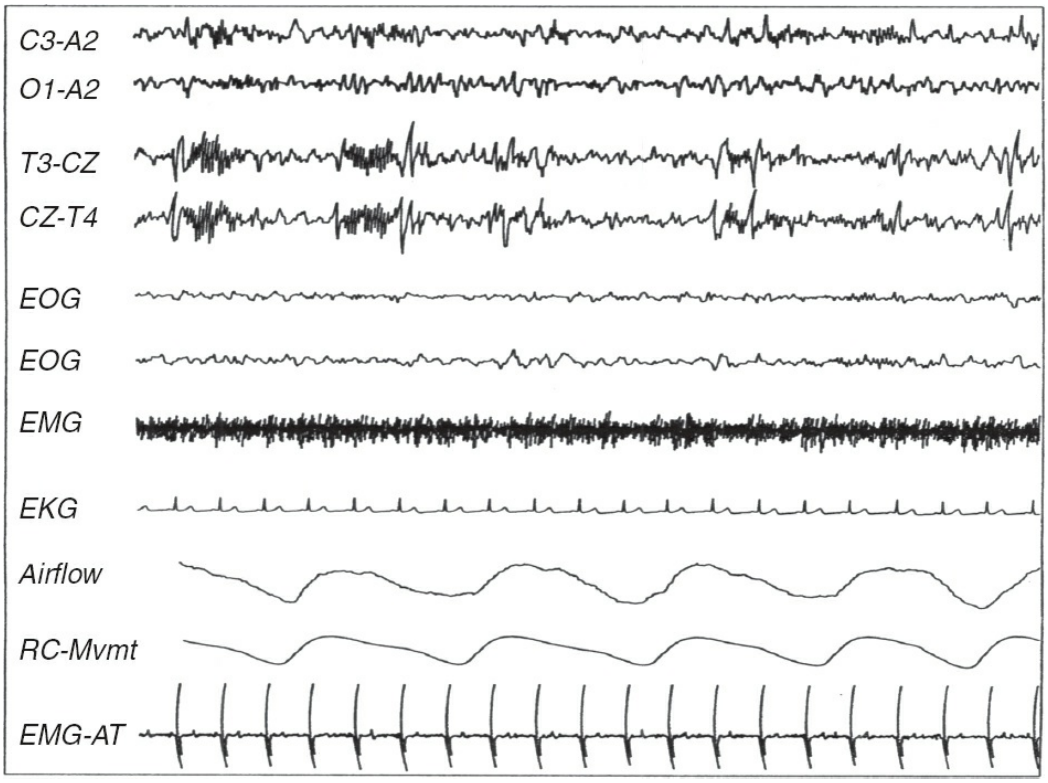

FIGURE 23–9. Polysomnographic tracing from a patient with complex partial seizures (CPS). EEG sleep spindles of excessive magnitude and duration mixed with sharp activity are visible, especially in temporal lobe derivations. The slow paper speed used in sleep studies makes it difficult to classify this activity precisely; however, the abnormality is apparent.

Such a finding should be confirmed with standard clinical EEG techniques both during wakefulness and sleep.

Periodic Limb Movement Disorder. Periodic Limb Movement Disorder (PLMD) previously called nocturnal myoclonus, involves brief, stereotypic, repetitive, nonepileptiform movements of the limbs, usually the legs. It occurs primarily in NREM sleep and involves an extension of the big toe. A partial flexion of the ankle, knee, and hip may also occur. These movements range from 0.5 to 5 seconds in duration and occur every 20 to 40 seconds. The leg movements are frequently associated with brief arousals from sleep and as a result can (but do not always) disturb sleep architecture (see Fig. 23–10). The prevalence of PLMD increases with aging and can occur in association with folate deficiency, renal disease, anemia, and the use of antidepressants. Pharmacotherapy for PLMD associated with RLS is same as RLS. Clinical trial of pharmacotherapy for other forms of PLMD is lacking. However, BZDs, especially clonazepam, reportedly improve sleep in patients with PLMD.

Sleep-Related Leg Cramps. Nocturnal leg cramps are much like leg cramps that occur during wakefulness. They usually affect the calf and are painful muscle contractions. The pain awakens the sleeper and thereby disrupts sleep. Metabolic disorders, mineral deficiencies, electrolytic imbalances, diabetes, and pregnancy are known precipitators. The reason that some individuals have repeated leg cramps during sleep and not during the day is not known.

Sleep-Related Bruxism. Sleep-Related Bruxism is diagnosed when an individual grinds or clenches their teeth during sleep (see Fig. 23–11). ICSD formerly classified sleep bruxism as a parasomnia but now lists it as a sleep-related movement disorder. Sleep bruxism can produce abnormal wear on the teeth, damage teeth, provoke tooth and jaw pain, and/or make loud unpleasant sounds that disturb the bed-partner. Sometimes atypical facial pain and headache also result. More than 85 percent of the population may brux at one time or another; however, it is clinically significant in only about 5 percent. Teeth grinding occurs in any sleep stage but appears to be most common at transition to sleep, in stage 2 sleep, and during REM sleep. Some evidence indicates that teeth-grinding during REM sleep is more commonly associated with dental wear or damage. Sleep bruxism does not appear to be exacerbated by dental malocclusion but rather worsens during periods of stress. Researchers studying sleep bruxism find that many patients seem to have less frequent teeth-grinding when sleeping in the laboratory; therefore, repeated study may be needed to document the disorder. By contrast, bruxism frequently appears on polysomnographic recordings made for other purposes and may occur

secondary to (a) sleep-related breathing disorders, (b) the use of psychostimulants (e.g., amphetamine, cocaine), (c) alcohol ingestion, and (d) treatment with some serotonin reuptake inhibitors. Differential diagnosis should rule out nocturnal seizures. Sleep bruxism can occur infrequently (monthly), regularly (weekly), or frequently (nightly). Severity is judged on the basis sleep disruption, consequent pain, and dental damage. Usual treatment involves having the patient wear an oral appliance to protect the teeth during sleep. There are two basic types of appliances used. The soft one (mouth guard) is typical used in the short term whereas the hard acrylic one (bite splint) is use longer term and requires regular follow-up. Relaxation, biofeedback, hypnosis, physical therapy, and stress management are also used to treat sleep bruxism. A variety of drug therapies have been tried (BZDs, muscle relaxants, dopaminergic agonists, and propranolol); however, reliable outcome data are not available.

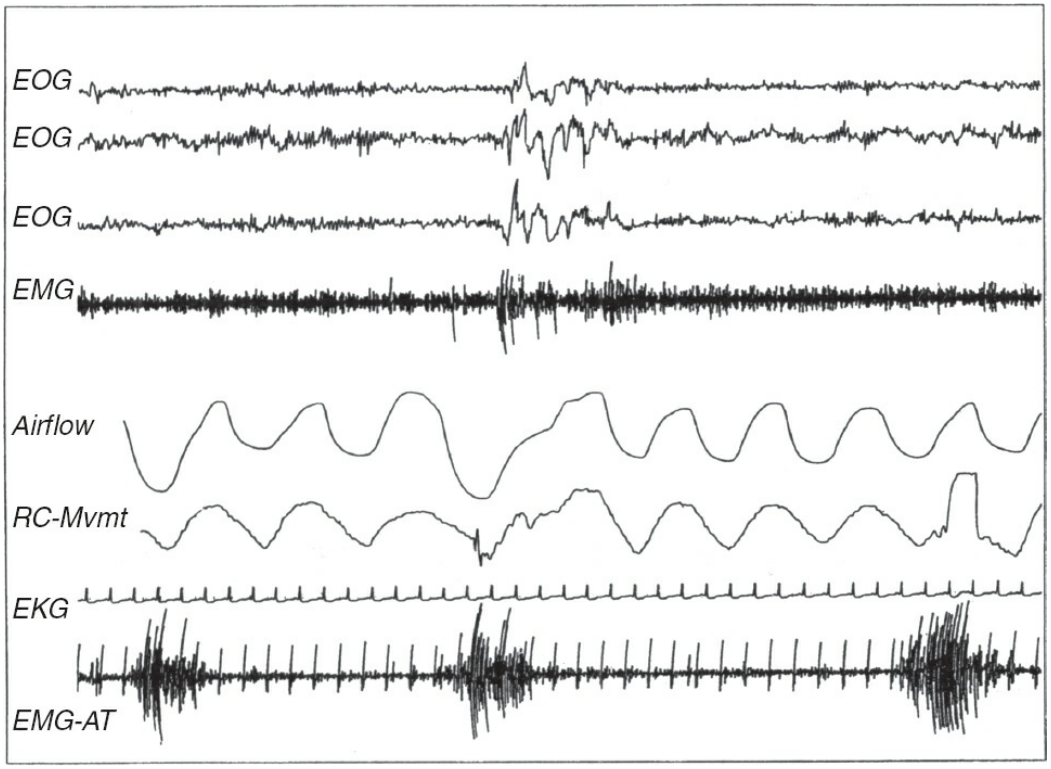

FIGURE 23–10. Polysomnographic tracing illustrating periodic leg movements in sleep. This figure shows periodic leg movements during sleep. Approximately 33 seconds of recording shown here clearly reveal a pattern of electromyographic activity in anterior tibialis (EMG-AT). The activity pattern repeats at regular intervals and sometimes produces transient arousals (note EEG-EOG-EMG) after movement at the center of the figure. This patient, whose chief complaint was nonrefreshing sleep with sleep maintenance insomnia, had more than 250 such leg movements during each night he slept in the laboratory.

Sleep-Related Rhythmic Movement Disorder. This sleep disorder is marked by repetitive, rhythmic movements usually involving the head and neck. Usually occurring at the transition from wakefulness to sleep, this

movement disorder may also continue during light sleep. ICSD formerly classified this nosologic entity as a parasomnia but is now listed as a sleep-related movement disorder. The Sleep-Related Rhythmic Movement Disorder has been called by many names, including *jactatio capitis nocturna,* head banging, head rolling, body rocking, and *rhythmie du sommeil.* A majority of infants will body rock. Some clinicians think body rocking develops from the soothing effect of vestibular stimulation. If the rhythmic movement persists into childhood and involves head banging, the risk of injury increases. Male to female ratio is 4:1. Severity ranges from less than one episode weekly to nightly episodes producing injury.

Benign Sleep Myoclonus of Infancy. Previously called benign neonatal sleep myoclonus, this disorder is characterized by asynchronous jerking of limbs and trunk during quiet sleep in neonates. This benign, apparently rare, parasomnia usually begins within the first week of life and may last a few days or several months. No treatment is recommended.

Propriospinal Myoclonus at Sleep Onset. This is a spinal cord–mediated movement disorder, sometimes associated with spinal cord lesions. Movements appear during times of wakeful relaxation and then may interfere with sleep onset. They start in the abdominal and truncal muscles and then progress to the neck and proximal muscles of the limbs. Treatment with clonazepam or anticonvulsants may be effective.

Sleep-Related Movement Disorder due to a Medical Disorder. Neurologic diseases that are associated with daytime movement disorders can likewise be associated with sleep-related movement disorders. Stress, anxiety, and sleep deprivation may contribute to bruxism.

Sleep-Related Movement Disorder due to Medication or Substance. A variety of medications and substances can produce or exacerbate sleep-related movement disorders. Stimulants can produce rhythmic movement disorders and bruxism. Antidepressants (including most tricyclics and selective serotonin reuptake inhibitors), antiemetics, lithium, calcium channel blockers, antihistamines, and neuroleptics can provoke Restless Legs symptoms and Periodic Limb Movement Disorder.

Sleep-Related Movement Disorder, Unspecified. ICSD-3 provides this nosologic classification for coding other unusual behaviors and otherwise unexplained movement disorders occurring during sleep.

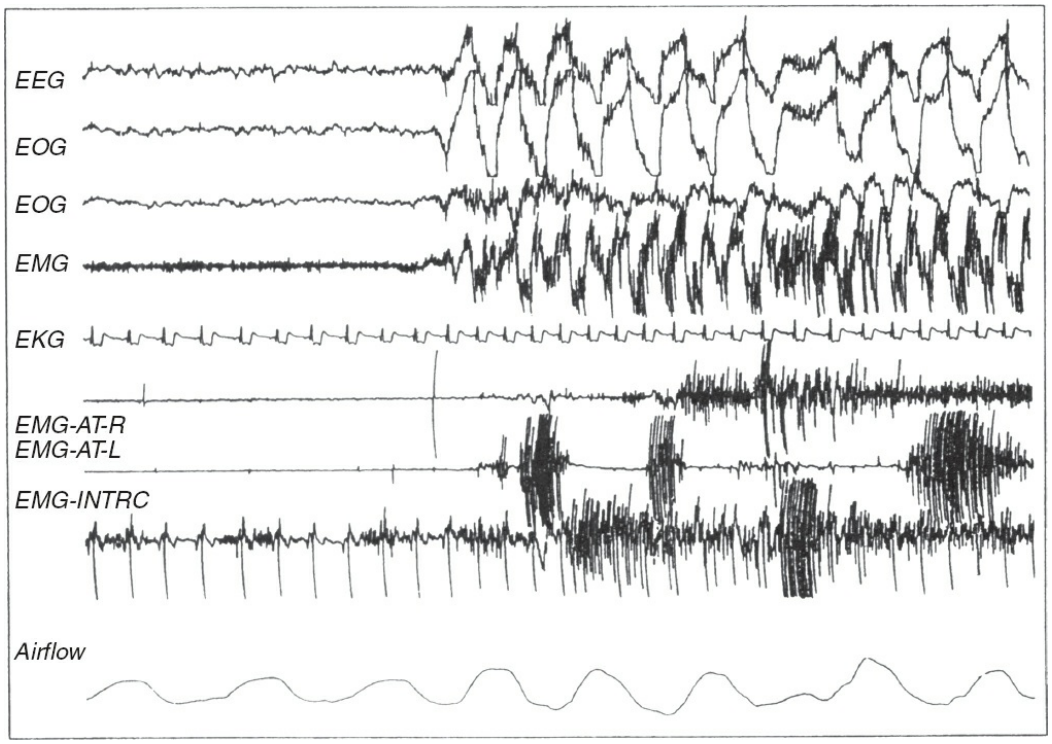

FIGURE 23–11. Polysomnographic tracing showing sleep-related bruxism. This figure showing approximately 25 seconds of tracing was obtained from a patient during an episode of bruxism. Bruxism can occur during any stage of sleep or wakefulness. The interference pattern of EEG-EOG-EMG channels is typical and reflects the rhythmic jaw movement and grinding of the teeth. This patient had many such episodes, some that caused awakening. Readily observable tooth damage and jaw pain were noted.

Isolated Symptoms and Normal Variants

Sleep Talking. As the name implies, sleep talking in classic form involves unconscious speech during sleep. It is seldom recognized in an individual unless it annoys the bed partner. It can be induced by fever, stress, or conversing with the sleeper. Somniloquy may accompany sleep terror, sleepwalking, confusional arousals, OSA, and REM sleep behavior disorder.

Excessive Fragmentary Myoclonus. These small movements or muscle fasciculations of the fingers, toes, or corners of the mouth are involuntary and can occur during wakefulness or sleep. Although little or no visible movement is present, the patient is typically aware of the twitching. In patients with apneas, twitching may worsen during periods of hypoxemia.

Hypnagogic Foot Tremor and Alternating Leg Muscle Activation. Hypnagogic foot tremor occurs at sleep onset or during stages 1 and 2 of sleep. It consists of a rhythmic movement of the toes or feet for seconds to minutes. Alternating Leg Muscle Activation consists of brief activation of the anterior tibialis in one leg and then the other.

Sleep Starts (Hypnic Jerks). Sleep starts are sudden, brief muscle contractions that occur at the transition between wakefulness and sleep in 60 to 70 percent of adults. The contractions commonly involve the legs; however, sometimes there is movement in the arms and head. This "hypnic jerk," as it is sometimes called, is usually benign. The sleep start; however, can interfere with the ability to fall asleep and may be accompanied by sensations of falling, a hallucinated flash of light, or a loud crackling sound. In severe cases the sleep start produces profound sleep-onset insomnia.

SUBSTANCE/MEDICATION-INDUCED SLEEP DISORDER

Medications and substances associated with or thought to provoke sleep disorders are discussed in sections: III.B.8. Insomnia Due to Drug or Substance; IV.B. Characteristics and Hypersomnolence Disorder Subtypes; VI.C.2.a. Central Sleep Apnea Due to Medication or Substance; V.E.8. Parasomnia Due to Drug or Substance; and V.E.16. Sleep-Related Movement Disorders Due to Medication or Substance.

TOOLS IN SLEEP MEDICINE

Clinical Interview

A careful and thorough clinical interview remains one of the most informative parts of a patient workup for sleep disorders. The habitual bedtime and arising times for both weekdays and weekends; the frequency, duration, and refreshingness of naps; and overall level of sleepiness and satisfaction with sleep are good places to begin. Specific sleep problems relating to difficulty initiating and maintaining sleep are important, including whether there is rumination at bedtime, fear of not being able to sleep, or excessive worry when attempting to sleep. Leg movement, leg sensations, leg cramps, teeth grinding, dream enactments (with or without injury), and other movements should be queried. Morning headaches, morning dry mouth, nocturnal reflux, hyperhidrosis, nocturia, enuresis, nocturnal tongue biting, nightmares, sleep terrors, and other sleep-related problems should be reviewed. The presence of family pets and whether they sleep in the bedroom (or bed) can be important in some cases.

Polysomnography

Polysomnography is the continuous, attended, comprehensive recording of physiological activity during sleep. A polysomnogram is typically recorded at night and is 6 to 8 hours in duration. Brain waves activity, eye movements, submentalis electromyography activity, nasal–oral airflow, nasal pressure, respiratory effort, oxyhemoglobin saturation, hearth rhythm, and leg movements during sleep are recorded. Body position is usually noted and snoring sounds may be recorded. Brain wave activity, eye movements, and submentalis electromyogram are important for identifying

sleep stages and CNS arousals. Muscle tension and movements subside with deeper sleep and can also be useful in the diagnosis of PLMD and RLS. Nasal airflow, nasal pressure, respiratory effort, and oxyhemoglobin saturation are instrumental in diagnosing sleep apnea and other sleep-related breathing disorders.

Indications for polysomnography include: (1) Diagnosis of sleep-related breathing disorders, (2) PAP titration and assessment of treatment efficacy, (3) evaluation of sleep-related behaviors that are violent or may potentially harm the patient or bed partner, (4) differentiating narcolepsy from other hypersomnolence disorders, and (5) differentiating parasomnias from nocturnal seizures. Polysomnography can also be used to diagnose atypical parasomnias, sleep-related problems secondary to neuromuscular disorders, periodic limb movement disorder, and arousals secondary to seizure disorder. Also, patients with excessive daytime sleepiness, or those who wake up gasping or choking, should be referred for polysomnography. A sleep study is not needed to diagnose restless leg syndrome.

Polysomnography is also recommended to assess sleep quality and quantity on the night just prior to an MSLT being conducted to diagnose narcolepsy.

Home Sleep Testing

Home sleep testing (HST) involves recording a limited number of cardiopulmonary parameters to assess patients for sleep-related breathing disorders. HST is much less expensive than polysomnography. Usually, airflow, respiratory effort, heart rhythm, snoring sounds, and oximetry are recorded. HST devices will detect sleep apnea in patients with moderate-to-severe pathophysiology. Negative studies are problematic because HST is less sensitive than full laboratory polysomnography. Patients with negative tests, notwithstanding obvious symptoms or comorbidity, should be scheduled for laboratory polysomnography. Additionally, HST does not check for the full spectrum of sleep disorders; therefore, residual symptoms, after a breathing disorder diagnosed in this manner and treated, need careful follow-up.

Multiple Sleep Latency Test

The MSLT is indicated for diagnosing narcolepsy. Every 2 hours, beginning 2 hours after morning awakening, a 20-minute nap opportunity is provided. The patient is instructed to let him- or herself fall asleep; that is, to not resist falling asleep. Electroencephalographic, electrooculographic, and submentalis electromyography activity are recorded in order to determine sleep stage. The latency to sleep is used to assess hypersomnolence and the appearance of REM sleep on two or more nap opportunities confirms narcolepsy (or on one nap if the previous night's

polysomnography revealed REM sleep within 15 minutes of sleep onset). If the patient falls asleep on a given nap opportunity, the nap is terminated 15 minutes after initial sleep onset. If the patient does not fall asleep, the session is terminated after 20 minutes of recording. A total of five nap opportunities are provided at 2-hour intervals across the day.

Maintenance of Wakefulness Test

Similar procedurally to MSLT, the maintenance of wakefulness test (MWT) provides 40-minute test sessions at 2-hour intervals across the day. However, the patient is instructed to try and remain awake and resist sleep rather than to not resist sleep. This technique is used to assess treatment outcomes and is sometimes required as part of "fit for duty" testing. Patients sit in a comfortable chair or on the bed with a bolster pillow in a darkened room while recordings are made. The first epoch of stages 2, 3, 4, or REM or three consecutive epochs of stage 1 mark unequivocal sleep onset. Falling asleep on MWT indicates some level of sleepiness. Fifty-nine percent of volunteers (presumably normal) remain awake for the entire 40 minutes across all four trials (using unequivocal sleep criteria). Any sleep latency <8 minutes is abnormal. Sleep latency ranging from 8 to 40 minutes is of unknown significance. The mean sleep latency (first sleep epoch) is 30.4 ± 11.2 minutes and the upper 95 percent confidence interval is 40 minutes.

Actigraphy

An actigraph is a device that measures and record movement. It is usually worn on the wrist (like a watch) and can be used to measure of the sleep–wake cycle. Depending on the model and the settings, it can make continuous recordings for days or weeks. It can be especially useful for assessing insomnia, circadian rhythm disorders, movement disorders, or an assortment of rare events. Research grade actigraph recorders are validated test instruments. Recently, the public's increased awareness and focus on fitness has spawned a market for consumer electronic devices that are essentially actigraphic recording devices. Many of these products purport to measure sleep and wakefulness; however, the precision and accuracy levels are currently unknown.

SLEEP DISORDERS CASES

CASE STUDY 1 (HYPERSOMNIA)

DESCRIPTION

Mr. J is a 28-year-old single African American male with an approximately 10-year history of fatigue and sleepiness in the daytime. He began to recognize the daytime sleepiness as a problem in his freshman year of college, when he would fall asleep in class or in the dormitory. He admits that his sleep–wake schedule was disrupted during college due to taking long naps

and then having to stay up until 1:00 or 2:00 AM to complete his studies. His grades and social life suffered and he describes himself as depressed, isolated, and hopeless about his planned future as a certified public accountant (CPA).

As a child, Mr. J says he slept "normally." In high school he felt best with 10 hours of sleep per night and was able to function well in the daytime. Mr. J denies abuse of alcohol or drugs. He does not use tobacco and drinks about 8 to 10 cups of coffee per day. Family history is negative for known sleep or psychiatric disorders. Physical Examination Findings are noncontributory except for body mass index of 0.29. Routine labs were normal, including TSH.

Mr. J's excessive sleepiness has continued to the present day, notwithstanding some improved sleep hygiene. Improvements include more consistency in bedtime, trying not to nap, and a torturous month-long trial without caffeine. He remains dysphoric and discouraged about his future, blaming his chronic sleepiness as the continuing impediment to his life plans. "I'm just tired of being tired," he says.

Currently, his bedtime is between 10:00 and 10:30 PM; his wake-up alarm is set for 6:30 am. He oversleeps at least once a week on work days and sleeps from 10:30 PM until 10:00 AM on weekends in an attempt to "catch up." He has difficulty awakening and feels unrefreshed or mildly refreshed. By drinking six to eight cups of coffee in the morning, he is usually able to avoid dozing during the morning. Luckily, he is an independent CPA and can schedule client appointments during this relatively alert time. After lunch he routinely falls asleep at the computer while working. He sleeps for 20 to 60 minutes and is usually awakened by his secretary. He then drinks another two cups of coffee and continues with his work. Unexpected napping can also occur later in the afternoon or evening and he has "nodded off" while driving. He sleeps alone; however, he has been told that he snores loudly. He does not awaken gasping or choking. He denies hypnagogic hallucinations and sleep paralysis but thinks he may feel weak after the rare occasions when he participates in a heated argument.

DISCUSSION OF CASE 1

Mr. J has one of the hypersomnias. Most consistent with his history are obstructive sleep apnea syndrome, idiopathic hypersomnia, sleep deprivation in a long sleeper, dyssomnia associated with mood disorder, and narcolepsy. The ancillary symptoms of narcolepsy are absent, with the possible exception of cataplexy. When cataplexy occurs clearly, the diagnosis of narcolepsy is strongly indicated. However, Mr. J's possible infrequent weakness during heated arguments is equivocal for cataplexy. His persistent desire for a 10-hour sleep period would be unusual for a patient with narcolepsy.

A long sleeper or an individual with idiopathic hypersomnia requires prolonged sleep periods and may awaken groggy as does Mr. J. The main differentiating feature is that whenever consistently given a chance to have a full nightly sleep period (usually 10 to 12 hours), the long sleeper does not experience excessive daytime sleepiness. Furthermore, there may be associated autonomic nervous system dysfunction or polysomnographic evidence of elevated slow-wave sleep percentage in patients with idiopathic hypersomnia.

Sleepiness associated with a low-grade chronic depression may be difficult to differentiate from other causes of hypersomnolence. Polysomnography, psychiatric interview, and psychometric testing can be helpful. Mr. J relates his dysphoria to sleepiness and not vice versa; nonetheless, a dyssomnia associated with mood disorder should be considered.

Obstructive sleep apnea syndrome is a strong possibly. Mr. J is overweight (BMI = 0.29) and snores loudly. Many patients are unaware of gasping or choking for breath. Often, family members witness cessation of breathing during sleep and urge patients to seek treatment. However, Mr. J lives and sleeps alone.

Polysomnography is recommended in patients suspected of obstructive sleep apnea syndrome, narcolepsy, or idiopathic hypersomnia. These disorders usually require lifelong treatment and have significant morbidity and mortality if untreated.

CASE STUDY 2 (PARASOMNIA)

DESCRIPTION

Ms. R is a 20-year-old white woman who was referred with symptoms of talking, mumbling, and crying out during sleep. At least twice per week she screams in her sleep. She is bothered

with excessive sleepiness and falling asleep inappropriately, such as during a conversation. When inactive, she is tired and sleepy, even after a full 8-hour night of sleep. However, she has energy when motivated and leads vigorous life. Once, she awakened outside of her apartment and her roommate had to let her back in because she had locked herself out. She does not recall the sleepwalking episode or other nocturnal wanderings but sometimes remembers yelling. From the history, crying seems to occur in light sleep but she rarely recalls any sleep-related thoughts or dreams. However, there is a history of occasional nightmares and bruxism. The patient uses an oral appliance to protect her teeth. Leg-kicking and mild snoring without gasping or choking are noted. The patient also complains of leg-kicking during sleep. Her sleep–wake schedule is irregular and she averages between 5 and 7 hours of sleep per night. She occasionally awakens with a headache in the morning.

Previous health history includes a hospitalization for febrile convulsions during infancy, ophthalmologic surgery for strabismus during childhood, and tonsillectomy as a teenager. Health is otherwise excellent. The patient does not smoke tobacco or drink alcohol.

DISCUSSION OF CASE 2

By history Ms. R has one or more of the parasomnias. Sleep talking alone does not require a sleep study, but this patient has nocturnal wanderings. Polysomnography, with clinical EEG are indicated to rule out unrecognized nocturnal seizure disorder or other organic factors inducing sleep walking. Sleepwalking is common and not necessarily considered abnormal in young children; however, in the adult it is rare and merits careful evaluation. Ms. R's excessive daytime sleepiness is likely due to insufficient sleep (5 to 7 hours per night) and possibly parasomnia-related disruption. Interestingly, many parasomnias are exacerbated by sleep deprivation as is nocturnal seizure disorder.

Sleep studies were performed using comprehensive, attended, laboratory polysomnography. Prior to the overnight study, a clinical EEG was performed. The clinical EEG study did not reveal any significant abnormal EEG activity during baseline, photic stimulation, or hyperventilation. An extended EEG montage was used during the sleep study. Overall sleep quality was within the normal range. Sleep efficiency was 96 percent and latency to sleep was 1 minute. REM sleep percentage was elevated (31 percent) and latency to REM sleep was less than normal (57 minutes). Slow-wave sleep was normal in percentage but EEG delta activity was very high amplitude. The overall macroarchitectural sleep pattern suggested rebound from sleep deprivation.

By contrast, sleep microarchitecture contained many abnormal features. We observed high amplitude paroxysmal EEG bursts. Excessively prolonged sleep spindles were noted and rhythmic K complexes were observed (see Fig. 23–12). There was one arousal out of slow-wave sleep with rhythmic EEG discharges alternating with sharp waves. Sharps and spikes occurred several times; however the focus was difficult to localize (possibly right temporal lobe). There were frequent body movements and full body jerks, most of which occurred during NREM sleep. There were episodes of moaning during slow-wave sleep and laughing during stage 2 sleep that was followed by high amplitude theta bursts and REM sleep. We observed frequent movements and arousals from REM sleep, but no REM-related spikes or sharp waves. Seizure-like EEG activity was noted during the night and occurred predominantly during slow-wave sleep. However, the patient did not attempt to sleepwalk. Sharp wave and spike activity increased during the final 45 minutes of the sleep study.

The patient did not have any sleep-related breathing impairment and SaO_2 nadir was 90 percent. She had no periodic limb movements during sleep and polygraphic features associated with RLS were absent.

CASE STUDY 3 (INSOMNIA)

DESCRIPTION

Ms. W is a 41-year-old, divorced white female presenting with a two and one half year complaint of sleeplessness. She has some difficulty falling asleep (30- to 45-minute sleep-onset latency) and awakens every hour or two after sleep onset. These awakenings may last 15 minutes to several hours and she estimates approximately 4.5 hours of sleep on an average night. She rarely takes daytime naps notwithstanding feeling tired and edgy. The patient

describes her sleep problem with the following words. “It seems like I never get into a deep sleep. I have never been a heavy sleeper, but now the slightest noise wakes me up. Sometimes I have a hard time getting my mind to shut down.” She views the bedroom as an unpleasant place of sleeplessness and states, “I tried staying at a friend’s house where it is quiet, but then I couldn’t sleep because of the silence.”

At times, Ms. W is unsure whether she is asleep or awake. She has a history of clock watching (to time her wakefulness) but stopped doing this when she realized it was contributing to the problem. Reportedly the insomnia is unrelated to seasonal changes, menstrual cycle, or time-zone translocation. Her basic sleep hygiene is good. Appetite and libido are unchanged. She denies mood disturbance, except that she is quite frustrated and concerned about sleeplessness and its effect on her work. Her work involves sitting at a microscope 6 of a 9-hour working day and meticulously documenting her findings. Her final output hasn’t suffered, but she must now “double check” for accuracy.

She describes herself as a worrier and a Type-A personality. The patient does not know how to relax. For example, on vacation she continually worries about things that can go wrong. She will not even begin to unwind until she has arrived at the destination, checked in, and unpacked. Even then, she is unable to relax.

Medical history is unremarkable except for tonsillectomy (age 16 years), migraine headaches (current), and diet-controlled hypercholesterolemia. She takes naproxen as needed for headache. She does not currently drink caffeinated beverages, smoke tobacco, or drink alcoholic beverages. She does not use recreational drugs.

The problem with insomnia began after relocation to a new city and place of employment. She attributes her insomnia to the noisy neighborhood in which she now lives. She first sought treatment 18 months ago. Her family practice physician diagnosed depression and she was started on fluoxetine that made her “climb the walls.” Antihistamines were tried next with similar results. She was then switched to low-dose trazodone (for sleep) and developed nausea. After these medical interventions, she sought medical care elsewhere. Zolpidem, 5 mg was prescribed but it made her feel drugged and upon discontinuation she had withdrawal effects. Another family practice physician diagnosed “nonspecific anxiety disorder” and began buspirone; an experience she describes as “having an alien try to climb out of my skin.” Buspirone was discontinued. Paroxetine was tried for 8 weeks with no effect. Finally, a psychiatrist was consulted, who diagnosed adult attention-deficit disorder (without hyperactivity) and suggested treatment with methylphenidate. At this point, the patient was convinced that a stimulant would not help her insomnia and demanded referral to a sleep disorders center.

DISCUSSION OF CASE 3

Ms. W’s symptoms fall into the broad category of an insomnia and the symptoms began after having moved from one city to another. Environmental sleep disorder (noise) and adjustment sleep disorder (new job, city, and apartment) are likely initial diagnoses. However, a more chronic, endogenous problem has become operative. What is it? Ms. W is a “worrier” and meticulous, but doesn’t reach diagnostic criteria for personality or anxiety disorders. Dyssomnia associated with mood disorder should be considered in any patient with sleep maintenance problems and early morning awakening insomnia. However, this patient does not have other significant signs of depression. Unfortunately, many patients are misdiagnosed with depression or “masked depression” on the sole basis of an insomnia complaint and unsuccessfully treated with antidepressant medication. Ms. W’s job demands long hours with focused concentration. Her job performance has been superior for many years notwithstanding insomnia. Thus, a diagnosis of attention-deficit disorder is unlikely. Idiopathic insomnia implies a childhood complaint, which Ms. W denies.

The likely working diagnosis is PPI. There may be some sleep state misperception (sometimes unclear of whether she is awake or asleep), but this cannot adequately account for the constellation of symptoms. An initial treatment plan should include further documentation of the sleep pattern using a sleep log. Behavioral treatments will likely benefit this patient. Medications with sedative effects are sometimes useful during initial treatment of PPI. However, thus far in this patient they have done more harm than good. She is likely to be a challenging patient to treat.

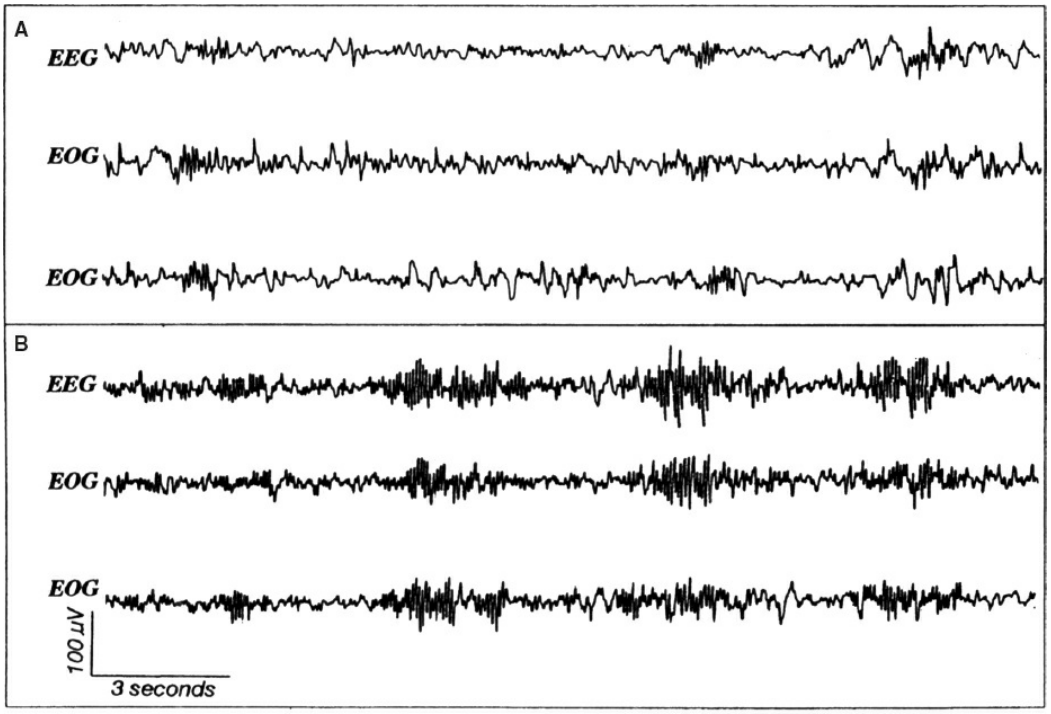

FIGURE 23–12. Polysomnographic tracings illustrating normal and excessive spindle activity. Polysomnographic tracings comparing normal sleep spindle EEG activity during stage 2 sleep (**Panel A**) with that of a patient chronically using a benzodiazepine (**Panel B**). This patient, who had been treated with benzodiazepines for more than a decade (and currently taking an extremely high dose) before being seen at the sleep disorders center. Sleep was grossly abnormal. The most obvious aberration was the tremendous increase in the frequency, magnitude, and duration of EEG spindle activity (see Panel B). Additionally, slow-wave sleep was absent, stage 2 was grossly elevated, and spindles even intruded into REM sleep.

SUMMARY

Sleep disorders medicine is an important, rapidly evolving discipline. There is great diversity in presenting complaint, etiology, course, and treatment of sleep disorders. Several nosological systems exist for classifying sleep disorders; however, the ICSD-2 appears to be the most comprehensive. Sleep disorders significantly alter an individual's health, mood, quality of life, and cognitive ability. Sleep disorders may be primary; however, many are associated with other medical, neurological, and psychiatric conditions. Disturbances in sleep or wakefulness may also result from substance use or abuse. For psychiatrists, sleep disorders are particularly important because insomnia and hypersomnia commonly occur in a wide range of psychiatric conditions. With appropriate treatment, most sleep disorders can be effectively managed. The extent of improvement in a patient's general well-being after treatment is usually dramatic. A good night of sleep is as important to general health as nutrition and exercise.

REFERENCES

American Academy of Sleep Medicine. *International Classification of Sleep Disorders*. 3rd ed.

Darien, IL: American Academy of Sleep Medicine; 2014.

American Academy of Sleep Medicine. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology, and Technical Specifications*. www.aasm.org. Darien, IL: American Academy of Sleep Medicine; 2014.

Ancoli-Israel S, Cole R, Alessi CA, et al. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*. 2003;26(3):342–392.

Chokroverty S. *Sleep Disorders Medicine: Basic Science, Technical Considerations, and Clinical Aspects*. 3rd ed. Philadelphia, PA: Saunders Elsevier; 2009.

Hirshkowitz M, Rose MW, Sharafkhaneh A. Neurotransmitters, neurochemistry, and the clinical pharmacology of sleep. In: Chokroverty S, ed. *Sleep Disorders Medicine*. 3rd ed. Philadelphia, PA: Saunders Elsevier; 2009:67–79.

Hirshkowitz M, Sharafkhaneh A. Diagnostic assessment methods. In: Kryger MH, ed. *Atlas of Clinical Sleep Medicine*. Philadelphia, PA: Saunders Elsevier; 2010:261–282.

Kryger MH, Roth T, Dement WC, eds. *Principles and Practice of Sleep Medicine*. 6th ed. Philadelphia, PA: Saunders; 2015.

Kushida CA, Littner MR, Morgenthaler T, et al. Practice parameters for the indications for polysomnography and related procedures: An update for 2005. *Sleep*. 2005;28(4):499–521.

Maquet P. Positron emission tomography studies of sleep and sleep disorders. *J Neuro*. 1997;244(4 Suppl 1):S23–S28.

Nishino S, Mignot E. Pharmacological aspects of human and canine narcolepsy. *Progr Neurobiol*. 1997;52:27–78.

Ohayon MM, Dauvilliers Y, Reynolds CF III. Operational definitions and algoithms for excessive sleepiness in the general population: implications for DSM-5 nosology. *Arch Gen Psychiatry*. 2012;9(1):71–79.

Ohayon MM, Reynolds CF III. Epidemiological and clinical relevance of insomnia diagnosis algorithms according to the DSM-IV and the International Classification of Sleep Disorders (ICSD). *Sleep Medicine*. 2009;10(9):952–960.

24

Impulse-Control Disorders

▲ 24.1 Intermittent Explosive Disorder

DIVYA RAMESH, PH.D., SAMEER HASSAMAL, M.D., AND FREDERICK GERARD MOELLER, M.D.

DIAGNOSIS AND EPIDEMIOLOGY

Intermittent explosive disorder (IED) in DSM-5 represents a disorder of recurrent, problematic, reactive (i.e., affective or impulsive), aggressive behavior that, over the lifetime, affects about 5 to 8 percent of individuals in the United States. Key diagnostic features of IED in DSM-5 include the following: (1) Aggressive outbursts have a rapid onset, typically without any premeditation, last for less than 30 minutes, and usually occur in response to provocation by a known individual. (2) Less severe episodes (verbal or noninjurious physical) are often interspersed with more severe episodes. (3) The disproportionate impulsive-aggressive response to a psychosocial stressor. (4) Aggressive outbursts are impulsive or provoked by anger and cause significant distress or impairment in psychosocial functioning. (5) The individual is above the age of 6 years or comparable developmental level. (6) The aggressive acts are "not better accounted for" by other DSM-5 disorders or drugs or a documented general medical condition. As described in DSM-5, differential diagnosis section on pages 468–469, a diagnosis of IED should not be given to individuals if the aggressive symptoms are better explained by another disorder such as disruptive mood regulation disorder, anti-social personality disorder or borderline personality disorder, delirium, major neurocognitive disorder, personality change due to another medical condition (aggressive type), substance intoxication or withdrawal, ADHD, conduct disorder, oppositional defiant disorder, or autism spectrum disorder. "In addition, children ages 6 to 18 years should not receive this diagnosis when impulsive aggressive outbursts occur in context of an adjustment disorder." However, the DSM-5 also suggests that a diagnosis of IED can be made in addition to the diagnosis of ADHD, conduct disorder, oppositional defiant disorder, or autism spectrum disorder if the recurrent impulsive aggressive outbursts

are in excess of those usually seen in these disorders and require independent clinical evaluation. Clinical history that is helpful in determining whether the aggressive behavior is better accounted for by another disorder includes whether the aggressive behavior is primarily impulsive and not for secondary gain, does not only occur in the context of a mood disorder, and does not occur only when the individual is intoxicated. If this is not the case, the aggressive behavior is more likely related to another condition.

The new DSM-5 criteria incorporates criteria that have been considered research criteria previously because studies show that these criteria better identify individuals with elevated levels of aggression, impulsivity, familial risk of aggression, and abnormalities in neurobiological markers of aggression. Another significant change in DSM-5 IED criteria is the type of aggressive outbursts that should be considered; physical aggression was required in DSM-IV, whereas verbal aggression and nondestructive/noninjurious physical aggression also meet criteria in DSM-5. In addition, DSM-5 also provides more specific criteria defining frequency needed to meet criteria and specifies that the aggressive outbursts are impulsive and/or anger-based in nature, and must cause marked distress, cause impairment in occupational or interpersonal functioning, or be associated with negative financial or legal consequences. Furthermore, because of the lack of research on this disorder in young children and the potential difficulty of distinguishing these outbursts from normal temper tantrums in young children, a minimum age of 6 years (or equivalent developmental level) is now required. Finally, especially for youth, the relationship of this disorder to other disorders (e.g., ADHD, disruptive mood dysregulation disorder) has been further clarified.

The age of onset for IED is in adolescence to early adulthood, ranging from 12 to 21 years of age, and is more common in males than females. There is a substantial comorbidity between IED and other psychiatric disorders, with 81 percent of individuals with IED meeting criteria for at least one other psychiatric disorder, with the most common comorbid disorders being substance use disorders, disruptive behavior disorders, PTSD, ASPD, borderline personality disorder, anxiety disorders, and major depression. Studies have also shown that those with IED are at increased risk for engagement in self-harm behaviors and global emotion regulation deficits.

ETIOLOGY

The well-documented case of Phineas Gage points out the importance of the prefrontal cortex in impulsive and aggressive behavior. Gage was a railroad worker who survived after a metal rod was blown through his skull in an accident. Prior to the injury he was an upstanding member of the community and a reliable worker. After the injury he was “fitful, irreverent,

indulging at times in the grossest profanity (which was not previously his custom), manifesting but little deference for his fellows, impatient of restraint or advice when it conflicts with his desires."

In a review of the literature on frontal lobe dysfunction and violent behavior Brower and Price concluded that the literature supports an association between frontal lobe dysfunction and increased aggressive behavior, with focal orbitofrontal injury being especially associated with increased aggression. For an individual patient, if an injury were clearly associated with the aggressive behavior, then by DSM-5 criteria the diagnosis of IED would not be made.

However, many violent patients have a history of head trauma due to the nature of their behavior that is not clearly associated with the onset of aggression. It is also unclear what degree of injury is necessary to increase aggressive behavior, making the diagnosis of personality change associated with a general medical condition (head trauma) challenging to separate from IED. A recent functional magnetic resonance imaging (fMRI) study in patients with IED showed higher error activity and increased activation in a dorsolateral prefrontal cortex subregion during a color-word Stroop task, possibly indicating a specific neural correlate of clinically high anger expression. Although more research is needed, the literature clearly supports the association between prefrontal cortex brain dysfunction and impulsive aggressive behavior as seen in IED.

A second line of biological evidence supports alteration of neurotransmitter function and impulsive aggression, with the largest body of evidence supporting a role for the neurotransmitter **serotonin** in aggression. A number of animal studies associate reduced serotonin function and increased aggression. Reduction in the serotonin precursor tryptophan, lesions of the serotonin cell bodies in the dorsal raphe, and serotonin gene knockout mice all increase aggression in animal models. In humans lower levels of the serotonin metabolite 5-hydroxyindole acetic acid (5-HIAA) are associated with violent suicide attempts and impulsive aggression; however, other studies have found no association between 5-HIAA levels and aggression. Further evidence supporting a role for reduced serotonin function leading to aggression in humans comes from studies using behavioral laboratory aggression paradigms. Studies using these paradigms have examined the role of serotonin using a drink deficient in the serotonin precursor tryptophan, which produces a significant reduction in plasma tryptophan levels and a reduction in brain serotonin. Several studies using this method of reducing serotonin have shown that there is an increase in aggression in human subjects after tryptophan depletion. The role of central serotonin function in IED has been examined in a study in IED patients where the authors found that physiologic responses to central 5-HT stimulation with d-FEN were reduced as a function of aggression (but not generalized impulsivity) in human subjects. The central 5-HT deficits in

these individuals can be attributed to abnormalities involving 5-$HT_{2A/2C}$ postsynaptic receptors in the limbic hypothalamus. This notion is supported by results from previous studies where selective 5-$HT_{2A/2C}$ antagonists completely blocked the prolactin response to d-FEN challenge in healthy volunteers. Likewise, at least some studies have found a reduction in impulsive aggressive behavior after treatment with serotonin reuptake inhibitors in patients with prominent histories of impulsive aggressive behavior, although other studies have not shown this. Another line of evidence suggests a significant inverse relationship between platelet 5-HT transporter levels and aggression in individuals with IED, further supporting the notion that reductions in brain 5-HT function are associated with aggressive behavior, particularly impulsive aggression.

Other factors which have been associated with impulsive aggression include childhood exposure to violence, maltreatment, and neglect. A recent study found that there was significant association between IED and aversive childhood parenting environment comparable to other psychopathologies. Another study by the same group found that physical abuse in childhood was independently associated with IED, with impulsivity and aggression as the mediators of the relationship between physical abuse and IED. Others have shown that IED is associated with childhood exposure to interpersonal traumatic events. It has been suggested that it is possible that a gene–environment interaction may exist where individuals who have certain serotonin gene polymorphisms may be more prone to impulsive aggressive behavior after exposure to **childhood maltreatment**.

TREATMENT

Pharmacotherapy has been a mainstay of the treatment of aggressive behavior in hospital emergency rooms for many years. However, this type of pharmacotherapy targets the immediate threat of aggression, where the issue of specificity of the treatment is not as important. For the treatment of IED, where impulsive aggression is a longstanding pattern of behavior, it is not clinically acceptable to sedate the patient in order to reduce the aggressive behavior over the long term. For pharmacotherapy for IED, the target is reduction in aggression without reduction in other behaviors.

A number of medications are known to reduce aggression and prevent rage outbursts, including antidepressants (namely selective serotonin reuptake inhibitors, or SSRIs), mood stabilizers (lithium and anticonvulsants), and antipsychotic drugs. Placebo-controlled studies targeting long-term treatment of impulsive aggression with pharmacotherapy date to the work of Sheard, where he found that violent young inmates showed a significant reduction in aggressive behavior after treatment with **lithium** compared to inmates treated with placebo. Other

studies have replicated this finding in adolescents with severe conduct disorder. The primary difficulty with the use of lithium for the treatment of impulsive aggression is its relatively narrow therapeutic window and side effects, including nausea, vomiting, and polyuria.

A number of anticonvulsants have been shown to reduce impulsive aggression in patients without seizure disorders. **Phenytoin** has been shown to reduce aggressive acts in inmates with impulsive aggression in placebo-controlled studies at blood levels that would be considered low for the treatment of epilepsy. **Valproic acid** in various preparations has also been shown to reduce impulsive aggression. Placebo-controlled trials showed that valproate and divalproex significantly reduce aggressive acts in individuals with personality disorders and conduct disorder. For individuals with personality disorders, the effect on aggression was greatest in individuals with the highest trait impulsivity, pointing out the importance of impulsivity in the pharmacotherapy of aggression. Other anticonvulsants with some evidence to support an effect on impulsive aggression include **carbamazepine and topiramate**. To date the evidence supporting carbamazepine as an anti-aggressive agent are mixed with some positive studies and a larger negative trial. Topiramate has also been shown to reduce aggression in several studies carried out by the same research group.

Because of the link between lower serotonin function and aggression, a number of trials of selective serotonin reuptake inhibitors (SSRIs) for impulsive aggression have been undertaken. Coccaro et al. administered **fluoxetine** or placebo to 40 patients with personality disorders and impulsive aggression. Subjects with current major depression, bipolar disorder, or schizophrenia were excluded. Results of that study were a significant reduction in irritability and aggression on the Overt Aggression Scale Modified for Outpatients (OAS-M) after months 2 and 3 of treatment. A similar study found that the SSRI sertraline considerably reduced irritability and aggression in in a small open-label trial with 11 personality disordered patients with impulsive aggression. Another study in 2009 on 100 IED patients found that those who took fluoxetine for 12 weeks experienced statistically significant reductions in impulsive-aggressive behavior compared with those who took a placebo. Despite the robust results during the study, fewer than half the patients taking fluoxetine achieved a full or partial remission.

As mentioned previously, the use of high-dose antipsychotic agents as a sedative is not clinically acceptable for chronic treatment of aggression due to the lack of specificity and the side effects. However, lower doses of newer generation antipsychotic agents appear to have a specific anti-aggressive effect that is independent of the antipsychotic or sedative effect. Although no controlled studies have specifically examined the efficacy of newer-generation antipsychotic medications in intermittent explosive disorder, a

number of studies have shown a reduction in aggression in children and adolescents with disruptive behavior disorders and autism as well as in adults with dementia. In children and adolescents with disruptive behavior disorders and sub-average intelligence, doses of 0.02 to 0.06 mg/kg/day of risperidone are associated with reduction in a variety of symptoms including irritability, lethargy/social withdrawal, hyperactivity/noncompliance, and aggression. However, studies in this population have also shown higher side effects in the risperidone-treated subjects, including sedation, extrapyramidal symptoms, and weight gain.

In elderly patients with dementia, at doses of 0.8 to 2.0 mg/day, **risperidone** has shown efficacy at reducing aggression and agitation and is generally well tolerated. Nevertheless, the U.S. Food and Drug Administration has determined that there is an increased risk of death in elderly patients with dementia-related psychosis treated with risperidone of between 1.6 and 1.7 times that seen in placebo-treated patients, due to a variety of causes including cardiovascular and infectious etiologies. This has led to a "black box" warning for risperidone related to these potential side effects. The risk of risperidone compared to other medications should be kept in mind in light of a recent study that found no difference in improvement in agitation and aggression between risperidone and citalopram in patients with dementia.

Psychostimulants, particularly **methylphenidate**, have been shown to be effective at reducing impulsive aggression in conduct disorder and ADHD. However, their use for the treatment of impulsive aggression outside conduct disorder or ADHD is limited due to the abuse potential of these medications as well as the potential for inducing mania in subjects with Bipolar Disorder.

Psychotherapy

Cognitive–behavioral therapy and contingency management have both shown evidence of reducing aggressive behavior in both adults and children. However the literature on the treatment of IED per se with psychotherapy is limited. A randomized controlled trial compared group and individual CBT for the treatment of IED (n = 45) with a wait-list control group. After 12 weekly sessions, group and individual cognitive–behavioral therapy reduced aggression, anger, hostile thinking, and depressive symptoms, while improving anger control relative to wait-list participant patients, with improvements persisting at the 3-month follow-up. Those who attended individual therapy sessions also reported an improvement in their overall quality of life.

References

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders:*

DSM-5. Washington, DC: American Psychiatric Publishing, Inc.; 2013.

Brower MC, Price BH. Neuropsychiatry of frontal lobe dysfunction in violent and criminal behaviour: A critical review. *J Neurol Neurosurg Psychiatry*. 2001;71(6):720–726.

Brown GL, Linnoila MI. CSF serotonin metabolite (5-HIAA) studies in depression, impulsivity, and violence. *J Clin Psychiatry*. 1990;51(Suppl):31–41; discussion 42–33.

Coccaro EF. Intermittent explosive disorder as a disorder of impulsive aggression for DSM-5. *Am J Psychiatry*. 2012;169(6):577–588.

Coccaro EF, Astill JL, Herbert JL, Schut AG. Fluoxetine treatment of impulsive aggression in DSM-III-R personality disorder patients. *J Clin Psychopharmacol*. 1990;10(5):373–375.

Coccaro EF, Fanning JR, Phan KL, Lee R. Serotonin and impulsive aggression. *CNS Spectr*. 2015;20(3):295–302.

Coccaro EF, Lee R, Kavoussi RJ. Aggression, suicidality, and intermittent explosive disorder: Serotonergic correlates in personality disorder and healthy control subjects. *Neuropsychopharmacology*. 2010;35(2):435–444.

Coccaro EF, Lee RJ, Kavoussi RJ. A double-blind, randomized, placebo-controlled trial of fluoxetine in patients with intermittent explosive disorder. *J Clin Psychiatry*. 2009;70(5):653–662.

Coccaro EF, Lee R, Kavoussi RJ. Inverse relationship between numbers of 5-HT transporter binding sites and life history of aggression and intermittent explosive disorder. *J Psychiatr Res*. 2010;44(3):137–142.

Coccaro EF, Siever EJ. Pathophysiology and treatment of aggression. In: Davis KL, Charney D, Coyle JT, Nemeroff C, eds. *Neuropsychopharmacology: The Fifth Generation of Progress*. Illadelph, PA: Lippincott Williams & Wilkins; 2002:1709–1723.

Cueva JE, Overall JE, Small AM, Armenteros JL, Perry R, Campbell M. Carbamazepine in aggressive children with conduct disorder: A double-blind and placebo-controlled study. *J Am Acad Child Adolesc Psychiatry*. 1996;35(4):480–490.

De Deyn PP, Buitelaar J. Risperidone in the management of agitation and aggression associated with psychiatric disorders. *Eur Psychiatry*. 2006;21(1):21–28.

Fanning JR, Meyerhoff JJ, Lee R, Coccaro EF. History of childhood maltreatment in intermittent explosive disorder and suicidal behavior. *J Psychiatr Res*. 2014;56:10–17.

Fehon DC, Grilo CM, Lipschitz DS. A comparison of adolescent inpatients with and without a history of violence perpetration: Impulsivity, PTSD, and violence risk. *J Nerv Ment Dis*. 2005;193(6):405–411.

Fettich KC, McCloskey MS, Look AE, Coccaro EF. Emotion regulation deficits in intermittent explosive disorder. *Aggress Behav*. 2014.

Goodall EM, Cowen PJ, Franklin M, Silverstone T. Ritanserin attenuates anorectic, endocrine and thermic responses to d-fenfluramine in human volunteers. *Psychopharmacology (Berl)*. 1993;112(4):461–466.

Goodman M, New A, Siever L. Trauma, genes, and the neurobiology of personality disorders. *Ann N Y Acad Sci*. 2004;1032:104–116.

Harlow JM. Recovery from the passage of an iron bar through the head. *Publ Mass Med Soc*. 1868;2:327–347.

Hollander E, Swann AC, Coccaro EF, Jiang P, Smith TB. Impact of trait impulsivity and state

aggression on divalproex versus placebo response in borderline personality disorder. *Am J Psychiatry*. 2005;162(3):621–624.

Jenkins AL, McCloskey MS, Kulper D, Berman ME, Coccaro EF. Self-harm behavior among individuals with intermittent explosive disorder and personality disorders. *J Psychiatr Res*. 2015;60:125–131.

Jones RM, Arlidge J, Gillham R, Reagu S, van den Bree M, Taylor PJ. Efficacy of mood stabilisers in the treatment of impulsive or repetitive aggression: Systematic review and meta-analysis. *Br J Psychiatry*. 2011;198(2):93–98.

Kessler RC, Amminger GP, Aguilar-Gaxiola S, Alonso J, Lee S, Ustun TB. Age of onset of mental disorders: A review of recent literature. *Curr Opin Psychiatry*. 2007;20(4):359–364.

Lee R, Meyerhoff J, Coccaro EF. Intermittent Explosive Disorder and aversive parental care. *Psychiatry Res*. 2014;220(1–2):477–482.

Malone RP, Delaney MA, Luebbert JF, Cater J, Campbell M. A double-blind placebo-controlled study of lithium in hospitalized aggressive children and adolescents with conduct disorder. *Arch Gen Psychiatry*. 2000;57(7):649–654.

McCloskey MS, Noblett KL, Deffenbacher JL, Gollan JK, Coccaro EF. Cognitive-behavioral therapy for intermittent explosive disorder: A pilot randomized clinical trial. *J Consult Clin Psychol*. 2008;76(5):876–886.

Moeller FG, Barratt ES, Dougherty DM, Schmitz JM, Swann AC. Psychiatric aspects of impulsivity. *American Journal of Psychiatry*. 2001;158(11):1783–1793.

Moeller FG, Dougherty DM, Swann AC, Collins D, Davis CM, Cherek DR. Tryptophan depletion and aggressive responding in healthy males. *Psychopharmacology (Berl)*. 1996;126(2):97–103.

Moeller SJ, Frobose MI, Konova AB, et al. Common and distinct neural correlates of inhibitory dysregulation: Stroop fMRI study of cocaine addiction and intermittent explosive disorder. *J Psychiatr Res*. 2014;58:55–62.

Nickel MK, Nickel C, Mitterlehner FO, Tritt K, Lahmann C, Leiberich PK, et al. Topiramate treatment of aggression in female borderline personality disorder patients: A double-blind, placebo-controlled study. *J Clin Psychiatry*. 2004;65(11):1515–1519.

Nickerson A, Aderka IM, Bryant RA, Hofmann SG. The relationship between childhood exposure to trauma and intermittent explosive disorder. *Psychiatry Res*. 2012;197(1–2):128–134.

Pandina GJ, Aman MG, Findling RL. Risperidone in the management of disruptive behavior disorders. *J Child Adolesc Psychopharmacol*. 2006;16(4):379–392.

Pollock BG, Mulsant BH, Rosen J, et al. A double-blind comparison of citalopram and risperidone for the treatment of behavioral and psychotic symptoms associated with dementia. *Am J Geriatr Psychiatry*. 2007;15(11):942–952.

Reardon AF, Hein CL, Wolf EJ, Prince LB, Ryabchenko K, Miller MW. Intermittent explosive disorder: Associations with PTSD and other Axis I disorders in a US military veteran sample. *J Anxiety Disord*. 2014;28(5):488–494.

Rinne T, van den Brink W, Wouters L, van Dyck R. SSRI treatment of borderline personality disorder: A randomized, placebo-controlled clinical trial for female patients with borderline personality disorder. *Am J Psychiatry*. 2002;159(12):2048–2054.

Sheard MH, Marini JL, Bridges CI, Wagner E. The effect of lithium on impulsive aggressive behavior in man. *Am J Psychiatry*. 1976;133(12):1409–1413.

Sinzig J, Dopfner M, Lehmkuhl G, et al; German Methylphenidate Study Group. Long-acting methylphenidate has an effect on aggressive behavior in children with attention-deficit/hyperactivity disorder. *J Child Adolesc Psychopharmacol*. 2007;17(4):421–432.

Stanford MS, Helfritz LE, Conklin SM, et al. A comparison of anticonvulsants in the treatment of impulsive aggression. *Exp Clin Psychopharmacol*. 2005;13(1):72–77.

Stanford MS, Houston RJ, Mathias CW, Greve KW, Villemarette-Pittman NR, Adams D. A double-blind placebo-controlled crossover study of phenytoin in individuals with impulsive aggression. *Psychiatry Res*. 2001;103(2–3):193–203.

Steiner H, Petersen ML, Saxena K, Ford S, Matthews Z. Divalproex sodium for the treatment of conduct disorder: A randomized controlled clinical trial. *J Clin Psychiatry*. 2003;64(10):1183–1191.

▲ 24.2 Pyromania

SAMEER HASSAMAL, M.D., DIVYA RAMESH, PH.D., AND FREDERICK GERARD MOELLER, M.D.

DIAGNOSIS AND EPIDEMIOLOGY

The diagnosis of pyromania requires a pattern of behavior with multiple episodes of deliberate and purposeful fire setting on multiple occasions. The fire-setting behavior can cause shame to the individual, with one study finding that 38.1 percent of individuals with pyromania considered committing suicide to stop their fire-setting behavior. There is tension or emotional arousal before the act and relief after the act, similar to OCD and kleptomania. Individuals with pyromania are curious about fires and are fascinated with fire paraphernalia, institutions, and people affiliated with the fire departments. There is pleasure or relief when setting fires, or when watching or engaging in its aftermath. In comparison to arson, the fire setting is not done for monetary reasons, to hide a crime, express a sociopolitical idea, vengeance, or in response to psychosis or judgment that is impaired by major neurocognitive disorder, intellectual disability, or substance intoxication. The diagnosis of pyromania is not better explained by conduct disorder, a manic episode, or antisocial personality disorder. Pyromania is highly comorbid with other disruptive, impulse-control, and conduct disorders, substance use disorders, gambling disorders, and bipolar disorders.

It is important to differentiate pyromania from other causes of fire setting. Fire setting is a behavior that can be accidental or intentional, and is not always a symptom of an underlying psychiatric illness. Fire setting (e.g., playing with matches, lighters) may also occur during childhood due to curiosity and may be part of the developmental process. Assessments for fire setting in children and adolescents include the Children's Fire Setting Inventory (CFI), the Fire Setting Risk Interview (FRI), the Fire Setting Incident Analysis Parent Version (FIA-P), and Fire Setting Incident

Analysis Child Version (FIA-C). A diagnosis of pyromania is excluded if the fire setting is better explained by conduct disorder, mania, antisocial personality disorder, major neurocognitive impairment, intellectual disability, substance intoxication, or in response to delusions, hallucinations, or a medical condition such as epilepsy.

There is a lack of epidemiologic data on pyromania, but there is data on fire setting without the diagnostic restrictions of pyromania. According to DSM-5, the prevalence of pyromania is unknown but lifetime fire setting in a sample of US residents 18 and older was found to be about 1.0 percent. In this sample, Caucasian males aged 18 to 35 living in the western region of the US had higher rates of fire setting. The most common comorbidities with fire setting include lifetime alcohol and marijuana use disorders, conduct disorder, and antisocial and obsessive-compulsive personality disorders. The most common motive for fire setting in general is revenge and anger. In contrast, pyromania, for which revenge and anger are excluded as causes of fire setting, appears to be quite rare. A study found that only 3.3 percent of arson recidivists met DSM-IV criteria for pyromania, largely because the majority of the arsonists had committed the act under the influence of alcohol, which is an exclusion criterion in DSM-IV and DSM-5. In fact, most arsonists do not meet criteria for pyromania. A study of 204 inpatient adults on a psychiatric unit found that 3.4 percent endorsed current pyromania symptoms and 5.9 percent had lifetime pyromania symptoms. What literature does exist about the prevalence of pyromania relates to children and adolescents, where prevalence estimates range from 2.4 to 3.5 percent, with 60 percent of all fires in large US cities being lit by individuals between 11 and 18 years of age.

Etiology

Pyromania has been described in literature for over 200 years; however, the etiology of pyromania is poorly understood. Freud described pyromania as consequence of abnormal psychosexual development. Literature on the origins of pyromania has suggested factors ranging from temperament to environmental factors and parental psychopathology. Other authors have suggested pyromania and the other impulse-control disorders may share a neurobiological link with substance use disorders. One case describes the onset of pyromania with the use of escitalopram. A single case report found a perfusion deficit in the left inferior frontal lobe on Single-photon emission tomography (SPECT), which resolved after treatment with cognitive-behavioral therapy (CBT) and topiramate. Other support for frontal inhibitory dysfunction in pyromania comes from neuropsychological testing of one subject with pyromania, which found attentional, visual memory, and executive function impairments. Earlier limited studies suggest that pyromania may be associated with lower concentration of norepinephrine and serotonin metabolites 3-methoxy-4-

hyroxyphenylglycol (MHPG) and 5-hydroxyindoleacetic acid (5-HIAA).

There is a lack of data on the development and course of pyromania. Most individuals with pyromania start setting fires in adolescence or young adulthood. The frequency and intensity of fire setting may increase over time or may wax and wane. The course of pyromania is unknown but smaller studies suggest it may be chronic. Pyromania occurs more frequently in males with poorer social skills and learning difficulties.

Treatment

There are no controlled clinical trials in the treatment of pyromania, and there are no FDA approved medications for the disorder. Case reports on pharmacologic management of pyromania have been sparse and results have been inconclusive. Some case reports have shown a reduction in pyromania urges and behaviors with SSRIs, lithium, naltrexone, stimulants, topiramate, valproic acid, carbamazepine, anti-androgen medication, clonazepam, and olanzapine. Other reports have shown no improvement in pyromania with pharmacotherapy.

What little literature does exist suggests CBT and other psychotherapies may be helpful. Other successful behavioral therapies for fire setting include aversive therapy, education, relaxation techniques, positive reinforcement, graphing therapy, and fire safety education. A study found that CBT, fire safety education (FSE), and a home visit from a firefighter (HVF) improved fire setting behaviors in children who had recently set a fire. However, CBT was the most efficacious in reducing fire setting, interest in playing with matches, and fire-related behaviors.

References

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Washington, DC: American Psychiatric Association; 2013.

Bennett BK, Gamelli RL, Duchene RC, Atkocaitis D, Plunkett JA. Burn Education Awareness Recognition and Support (BEARS): A community-based juvenile firesetters assessment and treatment program. *J Burn Care Rehabil*. 2004;25(3):324–327.

Burton PR, McNiel DE, Binder RL. Firesetting, arson, pyromania, and the forensic mental health expert. *J Am Acad Psychiatry Law*. 2012;40(3):355–365.

Ceylan MF, Durukan I, Turkbay T, Akca OF, Kara K. Pyromania associated with escitalopram in a child. *J Child Adolesc Psychopharmacol*. 2011;21(4):381–382.

Dell'Osso B, Altamura AC, Allen A, Marazziti D, Hollander E. Epidemiologic and clinical updates on impulse control disorders: A critical review. *Eur Arch Psychiatry Clin Neurosci*. 2006;256(8):464–475.

Franklin GA, Pucci PS, Arbabi S, Brandt MM, Wahl WL, Taheri PA. Decreased juvenile arson and firesetting recidivism after implementation of a multidisciplinary prevention program. *J Trauma*. 2002;53(2):260–264; discussion 4–6.

Freud S. The acquisition of power over fire. *Int J Psychoanal*. 1993;13:405–410.

Grant JE. SPECT imaging and treatment of pyromania. *J Clin Psychiatry*. 2006;67(6):998.

Grant JE, Levine L, Kim D, Potenza MN. Impulse control disorders in adult psychiatric inpatients. *Am J Psychiatry*. 2005;162(11):2184–2188.

Grant JE, Won Kim S. Clinical characteristics and psychiatric comorbidity of pyromania. *J Clin Psychiatry*. 2007;68(11):1717–1722.

Koles MR, Jenson WR. Comprehensive treatment of chronic fire setting in a severely disordered boy. *J Behav Ther Exp Psychiatry*. 1985;16(1):81–85.

Kolko D. Firesetting in children and youth. In: Van Hasselt V, Hersen M, eds. *Handbook of Psychological Approaches with Violent Offenders: Contemporary Strategies and Issues*. New York, NY: Kluwer Academic; 1999.

Kolko DJ. Efficacy of cognitive-behavioral treatment and fire safety education for children who set fires: Initial and follow-up outcomes. *J Child Psychol Psychiatry*. 2001;42(3):359–369.

Lindberg N, Holi MM, Tani P, Virkkunen M. Looking for pyromania: Characteristics of a consecutive sample of Finnish male criminals with histories of recidivist fire-setting between 1973 and 1993. *BMC Psychiatry*. 2005;5:47.

Milrod LM, Urion DK. Juvenile fire setting and the photoparoxysmal response. *Ann Neurol*. 1992;32(2):222–223.

Parks RW, Green RD, Girgis S, Hunter MD, Woodruff PW, Spence SA. Response of pyromania to biological treatment in a homeless person. *Neuropsychiatr Dis Treat*. 2005;1(3):277–280.

Rix KJ. A psychiatric study of adult arsonists. *Med Sci Law*. 1994;34(1):21–34.

Roncero C, Rodriguez-Urrutia A, Grau-Lopez L, Casas M. [Antiepileptic drugs in the control of the impulses disorders]. *Actas Esp Psiquiatr*. 2009;37(4):205–212.

Vaughn MG, Fu Q, Delisi M, et al. Prevalence and correlates of fire-setting in the United States: Results from the National Epidemiological Survey on Alcohol and Related Conditions. *Compr Psychiatry*. 2010;51(3):217–223.

Virkkunen M, Nuutila A, Goodwin FK, Linnoila M. Cerebrospinal fluid monoamine metabolite levels in male arsonists. *Arch Gen Psychiatry*. 1987;44(3):241–247.

▲ 24.3 Kleptomania

SAMEER HASSAMAL, M.D., DIVYA RAMESH, PH.D., AND FREDERICK GERARD MOELLER, M.D.

DIAGNOSIS AND EPIDEMIOLOGY

Kleptomania is one of the least studied of the disruptive, impulse-control, and conduct disorders. As conceptualized in DSM-5, the key diagnostic feature of kleptomania is that the items stolen are not needed for personal use or for their monetary value and not in response to anger, vengeance, a delusion, or a hallucination. The diagnosis of kleptomania also includes a sense of tension immediately before the theft and a sense of relief at the time of the theft, similar in structure to obsessive-compulsive disorder (OCD). The individual may hoard or return the stolen objects. Usually, the thefts are not preplanned and very little attention is given to the

consequences of the theft. The urge to steal is egodystonic and distressing. Subsequently, people with kleptomania experience guilt, depression, and suicide. The emotional stress, jail time, and marital and relationship conflict from shoplifting has been shown to decrease quality of life in individuals with kleptomania. As with other impulse-control disorders, the diagnosis of kleptomania is not better explained by conduct disorder, a manic episode, or antisocial personality disorder.

DSM-5 lists the prevalence of kleptomania in the general population to be 0.3 to 0.6 percent. The prevalence of kleptomania in those arrested for shoplifting ranges from 3.8 to 24 percent. A study of 204 consecutive admissions of psychiatric inpatients found a current prevalence of 7.8 percent and a lifetime prevalence of 9.3 percent, implying that the disorder may be more common than previously thought. The course of kleptomania is chronic with waxing and waning of symptoms. As the disorder is egodystonic, patients are embarrassed to discuss their symptoms and thus the disorder is usually underdiagnosed. A cross-cultural comparison of kleptomania between participants in the USA and Brazil found that both groups sought treatment 10 to 20 years after the illness onset and had high rates of suicide. Kleptomania may begin in childhood, adolescence, or adulthood but commonly begins during adolescence. Kleptomania occurs more frequently in women, with about two-thirds of kleptomania patients being women. Females with kleptomania are more likely to be married, shoplift household items, store stolen items, shoplift at a later age, or have a comorbid eating disorder or other impulse-control disorder. Kleptomania is highly comorbid with compulsive buying, mood disorders, anxiety disorders, eating disorders, personality disorders, and substance use disorders, as well as other disruptive, impulse-control, and conduct disorders. The rate of OCD in kleptomania has ranged from 6.5 to 60 percent. Due to the high comorbidity between OCD and kleptomania, it has been suggested that kleptomania may be a variant of OCD.

Kleptomania is associated with increased impulsivity, with a study showing that patients with kleptomania had higher scores on the Barratt Impulsiveness Scale, greater cognitive impulsivity, and greater sensation seeking in comparison to a psychiatric group. Furthermore, individuals with a greater severity of kleptomania symptoms have been shown to demonstrate more executive functioning impairment on the Wisconsin Card Sorting Test (WCST), greater cognitive impulsivity, and greater sensation seeking.

Etiology

The Swiss physician Mathey described kleptomania as early as the 19th century but little is known about the etiology of kleptomania. There have been case reports of kleptomania being associated with head trauma, frontal lobe lesions, cortical atrophy, dementia, and hypoglycemia

secondary to insulinoma. On diffusion tensor imaging, there is reduced fractional anisotropy in the inferior frontal white matter in patients with kleptomania, implying that the disorder may be associated with subtle white matter pathology. Another study found that although, as a group, patients with kleptomania did not show deficits on neuropsychological testing, the severity of illness was associated with impairment in executive functioning.

Data are lacking and limited on the genetics of kleptomania. A study found that individuals with kleptomania had more first-degree relatives with alcohol use disorders compared to controls. Serotonin and dopamine as well as the opioid and glutamatergic neurotransmitter systems have been shown to play a role in kleptomania. Additionally, there are higher rates of OCD in relatives of individuals with kleptomania compared to the general population.

Treatment

Due to the guilt associated with the behavior, most individuals with kleptomania seek treatment after being arrested. Pharmacological studies for kleptomania have been limited to mostly case reports and case series. Selective serotonin reuptake inhibitors have been tried based on the notion that kleptomania may be an obsessive-compulsive spectrum disorder. However, case reports have shown mixed results for SSRIs, with some reports of improvement but at least one report describing patients who developed kleptomania during treatment with SSRIs. An open label trial found that the mu-opiate receptor antagonist naltrexone reduced urges to steal and stealing behavior in kleptomania. In another open label trial, n-methyl-d-aspartate (NMDA) receptor antagonist memantine reduced urges to shoplift and shoplifting behavior along with improving impulsivity, mood, anxiety, and psychosocial functioning in kleptomania. In a placebo-controlled trial in the treatment of kleptomania there was a significant effect of naltrexone on kleptomania symptoms with 66 percent of the patients that received naltrexone achieving remission whereas 7.7 percent of those receiving placebo reported symptom remission. The only controlled trial of antidepressants in the treatment of kleptomania found no benefit of escitalopram in the treatment of kleptomania. There have also been case reports of lithium, topiramate, and methylphenidate reducing urges to steal in kleptomania.

There is little literature regarding the treatment efficacy of psychotherapy for kleptomania. Case reports and case series suggest that cognitive-behavioral therapy may be helpful.

Although the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) is not a validated rating scale for kleptomania, it has been used to monitor treatment response in kleptomania. The Kleptomania Symptom Assessment Scale, a self-report measure, has been shown to have adequate

psychometric properties in assessing changes in behaviors, urges, and cognition during treatment.

References

Aizer A, Lowengrub K, Dannon PN. Kleptomania after head trauma: Two case reports and combination treatment strategies. *Clin Neuropharmacol*. 2004;27(5):211–215.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Washington, DC: American Psychiatric Association; 2013.

Bayle FJ, Caci H, Millet B, Richa S, Olie JP. Psychopathology and comorbidity of psychiatric disorders in patients with kleptomania. *Am J Psychiatry*. 2003;160(8):1509–1513.

Chong SA, Low BL. Treatment of kleptomania with fluvoxamine. *Acta Psychiatr Scand*. 1996;93(4):314–315.

Dannon PN. Kleptomania: An impulse control disorder? *Int J Psychiatry Clin Pract*. 2002;6(1):3–7.

Grant JE. Understanding and treating kleptomania: New models and new treatments. *Isr J Psychiatry Relat Sci*. 2006;43(2):81–87.

Grant JE, Correia S, Brennan-Krohn T. White matter integrity in kleptomania: A pilot study. *Psychiatry Res*. 2006;147(2–3):233–237.

Grant JE, Kim SW. An open-label study of naltrexone in the treatment of kleptomania. *J Clin Psychiatry*. 2002;63(4):349–356.

Grant JE, Kim SW. Clinical characteristics and associated psychopathology of 22 patients with kleptomania. *Compr Psychiatry*. 2002;43(5):378–384.

Grant JE, Kim SW. Quality of life in kleptomania and pathological gambling. *Compr Psychiatry*. 2005;46(1):34–37.

Grant JE, Kim SW, Odlaug BL. A double-blind, placebo-controlled study of the opiate antagonist, naltrexone, in the treatment of kleptomania. *Biol Psychiatry*. 2009;65(7):600–606.

Grant JE, Levine L, Kim D, Potenza MN. Impulse control disorders in adult psychiatric inpatients. *Am J Psychiatry*. 2005;162(11):2184–2188.

Grant JE, Odlaug BL, Kim SW. Kleptomania: Clinical characteristics and relationship to substance use disorders. *Am J Drug Alcohol Abuse*. 2010;36(5):291–295.

Grant JE, Odlaug BL, Medeiros G, Christianine AR, Tavares H. Cross-cultural comparison of compulsive stealing (kleptomania). *Ann Clin Psychiatry*. 2015;27(2):150–151.

Grant JE, Odlaug BL, Schreiber LR, Chamberlain SR, Won Kim S. Memantine reduces stealing behavior and impulsivity in kleptomania: A pilot study. *Int Clin Psychopharmacol*. 2013;28(2):106–111.

Grant JE, Odlaug BL, Wozniak JR. Neuropsychological functioning in kleptomania. *Behav Res Ther*. 2007;45(7):1663–1670.

Grant JE, Potenza MN. Gender-related differences in individuals seeking treatment for kleptomania. *CNS Spectr*. 2008;13(3):235–245.

Gurlek Yuksel E, Taskin EO, Yilmaz Ovali G, Karacam M, Esen Danaci A. Case report: Kleptomania and other psychiatric symptoms after carbon monoxide intoxication. *Turk Psikiyatri Derg*. 2007;18(1):80–86.

Kaplan Y. Epilepsy and kleptomania. *Epilepsy Behav*. 2007;11(3):474–475.

Kindler S, Dannon PN, Iancu I, Sasson Y, Zohar J. Emergence of kleptomania during treatment for depression with serotonin selective reuptake inhibitors. *Clinical Neuropharmacol.* 1997;20(2):126–129.

Koran LM, Aboujaoude EN, Gamel NN. Escitalopram treatment of kleptomania: An open-label trial followed by double-blind discontinuation. *J Clin Psychiatry*. 2007;68(3):422–427.

Lepkifker E, Dannon PN, Ziv R, Iancu I, Horesh N, Kotler M. The treatment of kleptomania with serotonin reuptake inhibitors. *Clin Neuropharmacol.* 1999;22(1):40–43.

McElroy SL, Hudson JI, Pope HG, Keck PE. Kleptomania: Clinical characteristics and associated psychopathology. *Psychol Med.* 1991;21(1):93–108.

Mendez MF. Pathological stealing in dementia. *J Am Geriatr Soc.* 1988;36(9):825–826.

Odlaug BL, Grant JE, Kim SW. Suicide attempts in 107 adolescents and adults with kleptomania. *Arch Suicide Res.* 2012;16(4):348–359.

Presta S, Marazziti D, Dell'Osso L, Pfanner C, Pallanti S, Cassano GB. Kleptomania: Clinical features and comorbidity in an Italian sample. *Compr Psychiatry*. 2002;43(1):7–12.

Segal M. Letter: Psychiatric aspects of shoplifting. *Br Med J.* 1976;1(6008):523–524.

Talih FR. Kleptomania and potential exacerbating factors: A review and case report. *Innov Clin Neurosci.* 2011;8(10):35–39.

Yilmaz S, Bilgic A. The possible effect of methylphenidate on kleptomania in a school-age girl with attention-deficit/hyperactivity disorder. *Eur Child Adolesc Psychiatry*. 2014;23(5):361–363.

25

Adjustment Disorders

JEFFREY W. KATZMAN, M.D., AND CYNTHIA M. A. GEPPERT, M.D., PH.D.

INTRODUCTION

The diagnostic category of Adjustment Disorders (AD) is widely used among clinicians in practice. The new DSM-5 classification of AD among a group of stress response syndromes underscores that the central heuristic feature of AD is the linkage between the development of symptoms to an external stressful event. The event often involves financial, legal, or relationship difficulties or a new medical diagnosis. The adjustment disorder construct is often criticized for a lack of scientific support, and there are few controlled studies of the adjustment disorders in the literature to date, though new interesting data has emerged over the past few years that will be briefly summarized in this chapter.

The diagnosis of adjustment disorder strays from the general phenomenological approach of the DSM. The fifth edition of the diagnostic and statistical manual does not provide observable symptom criteria such as the familiar list of signs and symptoms that characterize the other major diagnostic categories. In the great debate over a categorical versus a dimensional formulation of mental disorders, DSM-5 underscores the dimensionality of not just AD but of the entire nascent class of trauma- and stress-related disorders that represent an etiological model linking a stressor to symptom formation. The linkage of an environmental stressor and symptom formation is consistent with paradigms utilized by many clinicians and even with the health beliefs of much of the public that stress and illness are connected, and this may account somewhat for the appeal of this diagnostic category for patients and practitioners alike.

The development of symptoms as a result of difficult events is widely considered part of the general human experience. For this reason, the diagnosis is often seen as far less stigmatizing than other psychiatric disorders as symptoms develop in response to an event and are often short-lived. Furthermore, medical insurance carriers do not consider the diagnosis as a preexisting illness in the same way as other psychiatric conditions. This distinction frequently leads clinicians to use the diagnosis to protect patients from future discrimination in medical, life, and disability insurances applications. However, this may be of less importance with changes in the structure of health care under the Affordable Care Act.

Whether the DSM-5 declaration that AD is a stress disorder akin to PTSD will alter these putative advantages is unknown.

COMPARATIVE NOSOLOGY: DSM

Despite the clinical appeal of this category of diagnoses, adjustment disorders have been seen as problematic for a number of reasons and the reconfiguration of DSM-5 has done little to address these problems. The diagnostic criteria describe a syndrome in which a stressful event leads to the development of a symptom complex. However, within the diagnostic construct, there are no criteria to qualify or quantify the stressor for an adjustment disorder in any way. Similarly, the symptom complex that develops has been criticized as lacking specificity. The temporal course between the stressor and the development of symptoms lacks rigorous scientific evidence. Furthermore, it is difficult in clinical practice to link an event to the development of a symptom complex. These issues have raised questions about the diagnosis in general and have no doubt contributed to the lack of academic investigation. Perhaps the greatest dilemma in considering this diagnostic category involves the question as to why some individuals develop symptoms in response to a stressful event while others do not, making AD a perfect vehicle for studies of the often underappreciated role of resilience in the face of trauma.

The diagnosis of AD has been criticized as problematic on multiple levels and at times has been termed a “wastebasket” or “afterthought” diagnosis. During the construction of DSM-5 there was scholarly consideration of the need to clean up the diagnosis of AD but the final version leaves the wastebasket full. The criticisms have focused on three main nosological weaknesses. The most serious objection is that the diagnosis is yet another manifestation of the increased medicalization of life-problems as well as the hegemony of biological models in psychiatry at the expense of social, psychological, cultural, and spiritual conceptualizations and responses to illness. The second critique points to the diagnosis’ failure to present a consistent, coherent, and clear explication of the essential relationship of an external stressor and internal diathesis. While the criteria require the stressor to precede the developments of symptoms in a causative manner, no specific symptoms, duration, nature, or magnitude of stressor has been delineated. It is extremely difficult in clinical practice to infer causality from the stressor to the development of symptoms. A clinician must then assess whether an individual’s symptoms are above and beyond what would be considered normal for that particular stressor. This raises one of the central philosophical issues in postmodern psychiatry: where normality ends and pathology begins. The literature points to the potential subjective nature of this evaluation, and the history of the diagnosis underscores the somewhat arbitrary designation of temporal aspects of the nosology.

The third "wastebasket" criticism refers to the use of adjustment disorder for clinical presentations that fail to meet the criteria of other psychiatric diagnoses—particularly major or minor depressions. Similarly, the predictive validity of the diagnosis has been criticized in that many initial diagnoses of adjustment disorder are later changed to another primary affective, substance use or personality disorder. Frequently the main factor in that change is duration of disturbance and not degree of severity. Once the temporal threshold has been met for another anxiety or mood or even trauma disorder, the diagnosis is changed, yet the arbitrariness of these time periods is well recognized. Claiming AD suddenly is transformed into another entity after 2 weeks or 2 months is not sound science. However, defenders of the diagnosis argue that the nosological problem lies more with the vague, abstract and over-general formulation of the criteria than the actual concept of a maladaptive reaction to stress. Despite these deficiencies, the diagnosis has had an enduring clinical and heuristic value in both ICD and DSM that warrants a more nuanced understanding of its utility and meaning.

HISTORY

Adjustment disorders as a personal reaction to an environmental or social stressor are likely universal phenomena, which have occurred in all historical epochs and cultures. A recent author suggests that the first recognizable clinical description of an adjustment disorder and its appropriate treatment is in the 11th century writings of the Islamic physician–philosopher Avicenna. Historically, DSM has held a place for a diagnostic category involving an acute psychological response to an environmental stressor. The initial impetus for this came through documentation of severe wartime stress seen in World War II as well as through the evolution of crisis intervention theory and practice. DSM-I in 1952 described the category of Transient Situational Personality Disorder. Within this category were the subtypes of gross stress reaction, adult situational reaction, adjustment reaction of infancy, adjustment reaction of childhood, adjustment reaction of adolescence, and adjustment reaction of late life. It is notable that with even the first appearance of this diagnosis, there was an overlap in the diagnostic symptoms between the concepts of a stressor, depressive and anxious affect, and one's personality.

The diagnosis was modified somewhat in DSM-II, which changed the concept to transient situational disorder. This category was reserved for "more or less transient disorders of any severity (including those of psychotic proportions) that occur in individuals without any apparent underlying mental disorders and that represent an acute reaction to overwhelming environmental stress." DSM-II provided similar developmental subtypes to DSM-I; however, the subtypes of gross stress reaction and adult situational reaction were eliminated.

DSM-III introduced the diagnosis of adjustment disorder. The developmental periods of the earlier diagnostic systems were eliminated. Instead, the subtypes of adjustment disorder were categorized based on the predominant affective experience. These included adjustment disorder with depressed mood, anxious mood, mixed emotional features, disturbance of conduct, mixed disturbance of emotions and conduct, work inhibition, withdrawal, and atypical features. DSM-III-R retained these diagnostic subtypes and added an additional one involving physical complaints. DSM-III-R also specified that symptoms of an adjustment disorder could not exceed 6 months.

DSM-IV modified the diagnosis of adjustment disorder in several ways. The subtypes of mixed emotional features, work inhibition, withdrawal, and physical complaints were eliminated. The stressor was allowed to persist for an indefinite period of time. A descriptor of chronicity was specified, whereby symptoms persisting for greater than 6 months were deemed chronic.

DSM-5

While DSM-5 does not make substantial changes to the actual diagnostic criteria for adjustment disorder, a key change in classification has occurred that, in the view of one author, "represents a major shift in thinking and is the most significant change since its introduction in the DSM-III." AD is now classified with reactive attachment disorder, disinhibited social engagement disorder, posttraumatic stress disorder, acute stress disorder, other specified trauma- and stressor-related disorders, and unspecified trauma- and stressor-related disorders in a newly delineated category of Trauma- and Stressor-Related Disorders. In some measure this reorganization settles at least for the time being the long historical debate about whether AD belongs in the family of mood, anxiety, or stress disorders. The authors of DSM-5 have chosen to emphasize most strongly the nature of AD as one of a range of human responses to stress and trauma and to conceptualize the disturbances of mood, anxiety, and conduct that constitute the established subtypes of AD as varied aspects of the stress–trauma response spectrum.

In prior versions of DSM, AD was coded on Axis I, with the specific stressor coded on AXIS III or IV of the problematic system that DSM-5 has eliminated. The diagnosis of AD could not be assigned in the presence of an Axis I disorder, which generated further obfuscation of the diagnostic picture, which in turn could lead to inappropriate treatment recommendations. It is hoped that the relocation of AD—long without a conceptual home—to a family of disorders thematically united as reactions to an external stressor will facilitate a higher quality of research that could lead to more evidence-based treatments than have hitherto been available for AD. AD was located in the trauma-related category in ICD-10 and it will

reportedly remain there in the 11th version of the criteria.

DEFINITION

The adjustment disorders are a diagnostic category characterized by an emotional response to a stressful event. The stressors involved for the other diagnoses in the DSM-5 category have criteria, which are fairly well delineated. There are currently no such criteria for the stressor of an adjustment disorder. Typically, the stressor involves financial issues, a medical illness, or a relationship problem. The symptom complex that develops may involve anxious or depressive affect or may present with a disturbance of conduct. There is little specificity to qualify these symptoms other than an excess of what would be expected from the particular stressor or a significant social or occupational impairment. By definition, the symptoms must begin within 3 months of the stressor, and must remit within 6 months of removal of the stressor. If the symptom complex is less than 6 months in duration it is deemed acute. Symptoms lasting beyond 6 months of the initial event are coded as a chronic adjustment disorder. The symptom complex must not qualify for another Axis I condition. Normal bereavement is an exclusion. The subtypes of adjustment disorder have not significantly changed in DSM-5. These include adjustment disorder with depressed mood, anxious mood, mixed anxiety and depressed mood, disturbance of conduct, mixed disturbance of emotions and conduct, and unspecified type.

STRESSOR CRITERIA

DSM-5 criteria for AD rely first and foremost on the presence of a stressor. "The presence of emotional or behavioral symptoms in response to an identifiable stressor is the essential feature of adjustment disorder." The explanatory material that accompany the criteria in the manual try to sketch the nature of the stressor but the contours remain unclear, especially when compared with the relatively sharper shape of the trauma of PTSD or the obvious stressor in bereavement. The text contains multiple examples of this ambiguity. For instance, the stressor can be single, like the loss of a significant relationship, or multiple, such as combined employment and marital problems. The stressor can be recurrent, such as a construction worker who cannot find work during the winter, or continuous, most commonly with AD from a chronic illness. In contrast with the individualistic focus of most of the disorders in DSM, AD can be familial or communal in its impact, as can PTSD. The implicit developmental triggers of AD—well known to clinicians—are made explicit, such as the "empty nest syndrome" or a child going off to college.

This leaves the clinician to determine whether any particular stressor could lead to the development of a stress response. Bereavement, for instance, is not an AD if it is normal but *is* an AD if "the intensity, quality or

persistence of grief reactions exceeds what normally would be expected." The vast range and variance in these characteristics is quite large, and one would not reasonably believe that a divergent group of human experiences would lead to a similar psychological experience. One might consider that an event involved in the development of an adjustment disorder is critical yet not uncommon. Such events effect most people in the course of a lifetime, yet are infrequent against the timeframe of an entire life.

Similarly, the implication of the current criteria is that there is an expectable reaction to a particular stressor. Clinicians are left to determine what response might be "expectable" to any identifiable event. Any such assessment of the response would be influenced by the personal history, values, and culture of the individual examiner and the patient. In the case of bereavement the criteria acknowledge that the "cultural, religious and age-appropriate norms must be taken into account."

TEMPORAL COURSE

The time course of AD has not changed in DSM-5. The onset of symptoms of AD must be within 3 months of a stressor and the resolution of symptoms within 6 months of the termination of the stressor unless another major psychiatric diagnosis preempts AD. There is some evidence to support the idea of symptom development closely following the time of the stressor. Two studies have demonstrated the close temporal relationship of stressor and symptom formation. In one study of children with a new diagnosis of insulin-dependent diabetes, nearly one-third developed adjustment disorders. The highest risk period was in the first month following diagnosis, and 31 of 33 children had developed symptoms within the first 3 months following the diagnosis of diabetes.

However, one could imagine that symptom formation in some individuals might come with considerable time lag. The "with delayed expression" specifier for PTSD allows for this development if symptoms are present 6 months or more from the time of the stressor. DSM-5 includes a similar exception under the "Other Specified Trauma- and Stressor-Related Disorders" that lists a presentation of AD with delayed onset 3 months after a stressor. The literature confirms the potential temporal variation in symptom formation from the onset of the stressor. One study examined the timing and number of life events occurring in the year before the onset of emotional disorders in school-aged children in an outpatient clinic. Thirty percent of cases did not develop symptoms until 6 months to 1 year following the undesirable event. Children who experienced multiple negative life events were more likely to have less time between stressor and symptom formation.

The diagnosis of adjustment disorder specifies that the symptoms do not persist for more than 6 months once the stressor—or its consequences—have terminated. There is a new diagnosis included in the Other Specified

Trauma- and Stressor-Related Disorder classification for an AD-like disorder with a duration of symptoms prolonged for greater than 6 months without continuation of the stressor. It is often extremely difficult for a clinician to gauge when a stressor is no longer a stressor. In many cases, stressors leading to this diagnosis such as chronic illness or marital tensions do not go away. Rather, the meaning of the stressor to the individual may change over time, for instance, with a support group or couples therapy. The literature supports the notion that emotional symptoms can linger for quite some time. One study compared the long-term mental health effects in adults of two different community stressors (the Three Mile Island accident and widespread unemployment due to layoff). Symptoms remained elevated for as long as 42 months afterward. It is unclear, however, if this data shows that an event can lead to lingering responses, or whether the event itself lingers and continues to impact an individual, therefore engendering ongoing emotional symptoms.

A study at the University of Michigan captures the enduring nature of stress. The research identified increased stressors including breakup of a marriage relationship, occurrence of marital stressors, and financial issues as lasting symptoms following the September 11, 2001 terrorist attacks in a city 600 miles from the actual events. In addition, there was an increased rate of medically serious suicide attempts in the 2-year period following the attacks compared to the previous 2 years. These results speak to the potential chronicity of stress following a single traumatic event and may be applicable to the development of chronic adjustment disorder as well.

RELIABILITY

The literature contains mixed information about the reliability of the adjustment disorder construct. The results of the UK-World Health Organization (UK-WHO) study of reliability of the ninth revision of International Statistical Classification of Disease categories in children and adolescents were consistent with this. The inter-rater reliability for adjustment disorders was 0.23; this was lower than for many other categories. The clinical relevance of poor reliability was underscored in a study that examined the concordance rate between an initial diagnosis of adjustment disorder given by a house officer and a subsequent diagnosis by a consulting psychiatrist. Fifty percent of the inpatients with adjustment disorder actually had delirium.

VALIDITY

Despite the poor reliability of the diagnosis, there is a growing body of research that supports the content and predictive validity for the diagnosis. Content validity has been demonstrated through several studies that differentiate adjustment disorder from major depression on the one hand and normal reactions to stress on the other. For instance, a study of

consultation cases in the medical setting found that patients given the adjustment disorder diagnosis also had higher ratings on Axis IV, lower illness severity, and were more likely to be improved at time of discharge. It should be noted that results from the European Outcome of Depression International Network (ODIN) research failed to demonstrate content validity in that it identified no variables that might distinguish adjustment disorder from depression.

Predictive validity was demonstrated through a study involving a group of 2,078 male Navy-enlisted personnel diagnosed with transient situational disturbance. This proved to be a less disabling diagnosis in terms of chronicity, length of hospitalization, and ability to return to a previous job than any other psychiatric condition. The predictive validity of the diagnosis of adjustment disorder was demonstrated in a large study of adult inpatients. The investigators sought to determine whether patients with adjustment disorder shared a similar prognosis at 5 years. Seventy-nine percent of adults in the study were well at 5-year follow-up. The diagnosis was not as predictive for adolescents with 57 percent well at 5 years. The authors suggest that the good prognosis shared by the adult cohort implies specific characteristics of the group diagnosed with adjustment disorder, and that the diagnosis carries with it a prediction of recovery at 5 years. The cohort of children in the sample, however, does not share this predictive validity.

A study of Danish conscripts supports both content and predictive validity. Subjects diagnosed with adjustment disorder showed a steady increase in symptoms of depression, sleep, personality problems, and somatization from time of admission to the service until discharge. Fourteen days after discharge, scores on all symptom domains were comparable to those at enrollment into the military and 1 year later, and except for a higher depression score, subjects were comparable to controls. Subjects also differed from general conscripts on personality profile at enrollment having more adverse family events, depression, and personality problems.

One recent study raises questions regarding the validity of the subtypes of adjustment disorder in particular. In an outpatient sample, no differences in sociodemographics, comorbid diagnoses, or lifetime history of the disorder could be found in groups diagnosed with the anxious and depressive subtypes of adjustment disorder. The author raised the question of collapsing these subtypes into one diagnostic entity.

CULTURAL PERSPECTIVES

An interesting discussion is posited regarding the utility of the adjustment disorder diagnosis through the lens of cultural anthropology. DSM-5 cautions the clinician attempting to decide whether or not a given response is disproportional to the expected reaction to a stressor to “take into

account the external context and the cultural factors that might influence symptom severity and presentation." This perspective considers the evidence that most cultures have an entity named for a process whereby an individual is stressed to the point of symptom development. This parallels the lay concept of a "nervous breakdown" within a particular cultural setting. In this sense, the adjustment disorders share features with other culture-bound syndromes such as Susto, Koro, and Arctic Hysteria in that they all describe the development of a set of symptoms following a particular stressor. DSM-5 includes these and other culture-bound syndromes in the category of Other Specified Trauma- and Stressor-Related Disorders and specifically mentions *Ataque de Nervios*. The particular cultural presentation of an adjustment disorder will be affected by: (1) the nature, intensity, and meaning of the stressor in question; (2) the nature of the modal personality configuration of the people involved—which includes style and/or rules about behavior and emotional expression; (3) idiosyncratic features of the self in question; and (4) the meaning that adjustment disorder has in the culture. The individual is first treated within the family and other social support networks. When this fails, the event is inevitably medicalized in some fashion.

The universality of a maladaptive reaction to disturbing but less-than-catastrophic traumas is shown in studies in non-Western countries. A study of adolescents in Bahrain suggests cultural utility if not validity in that the 72 youths diagnosed with adjustment disorder had characteristics which differentiated them from controls, including stressors such as disappointment in relationships with family or friends of opposite sex, and were more likely to have premorbid family dysfunction.

Studies with refugees, immigrants, victims of torture, and expatriate populations suggest that cultural conflict, problematic acculturation, social or economic displacement, or political disenfranchisement or persecution are all contemporary stressors that may lead to the development of adjustment disorder in a global context.

Modern medical culture often uses adjustment disorder as a means of simultaneously destigmatizing and legitimatizing psychic distress. Adjustment disorder is uniquely suited as is no other entity in DSM to communicate both to patients and third parties that the dysfunction in questions is the result of an understandable if exaggerated reaction to external events rather than an internal disease state whether biological or psychosocial. As the case below illustrates, third parties including health and life insurance carriers, employers, and family members often look at this diagnosis as a nonrecurring phenomenon with a generally good outcome. Adjustment disorder enables a clinician using the current language of health care delivery to indicate that the patient is significantly impaired to warrant clinical attention and hence payment for services rendered. It also underscores the assumption in Western medical culture

that maladaptive forms of distress necessarily represent diagnostic entities.

A 46-year-old pediatric nurse receives a negative evaluation at work from her nursing supervisor and is put on probation. She had always been successful in her career but the development of a back injury and one of her sons being diagnosed with bipolar disorder had impaired her ability to perform adequately. After the evaluation she became unable to get out of bed and often called in sick. Her mood was irritable and depressed and she was unable to concentrate or make decisions either at work or home. When her children found her crying one afternoon, she realized she needed help and sought assistance from employee health. Her therapist indicated she felt the patient was suffering from major depression. The patient was quite upset and requested the therapist assign a less serious diagnosis as she felt depression would provide further evidence that she was not fit for duty. The therapist agreed to use the DSM-5 criteria for adjustment disorder in her clinical notes and for billing purposes.

ICD

The 10th revision of *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) also contains a category of Adjustment Disorders. The diagnosis is similar to the DSM entity in outlining the development of psychological symptoms following a stressor. However, in ICD-10, the symptoms must appear within 1 month of the stressor rather than the 3-month temporal course of DSM-IV-TR. The ICD-10 criteria require that symptoms must not persist for greater than 6 months after the removal of the stressor. DSM-5 removed the acute and chronic specifiers for the subtypes of AD. However, ICD-10 includes a category of Prolonged Depressive Reaction for AD-like symptoms that persist for up to 2 years.

The ICD-10 states that symptoms of an adjustment disorder may be any of those found in the major affective, neurotic, stress-related, somatoform, or conduct disorders provided that the symptoms are not sufficient to qualify for the particular diagnostic entity. It goes on to specify specific diagnostic subtypes of adjustment disorder based on the nature, degree, and temporal course of symptoms involved. These subtypes track the DSM-5 subtypes fairly well, in that they involve symptoms of depression, anxiety, and conduct. However, there are some issues that are unique to the subtypes of ICD-10. For example, the ICD-10 subtype of adjustment disorder with predominant disturbance of other emotions is somewhat similar to the DSM category of Mixed Emotional Features. However, this ICD-10 category is also used to code specific childhood regressive reactions including bedwetting and thumbsucking.

EPIDEMIOLOGY

Few studies have examined the prevalence of adjustment disorder in community samples, underscoring its marginal and ambiguous status in the outpatient diagnostic repertoire. Neither the Epidemiological Catchment Area nor the National Comorbidity Study included adjustment disorder in the survey instruments. One exception to the exclusion of AD in epidemiological research is the European Outcome of Depression

International Network (ODIN) study which was a cross-sectional two-phase community assessment using first the Beck Depression Inventory and then the Schedule for Clinical Assessment in Neuropsychiatry (SCAN), one of the few structured instruments to include AD. Less than 1 percent of the subjects were diagnosed with AD, and the rate varied from country to country. The discrepancy between these findings and almost all previous studies raises further questions regarding the reliability and validity of current diagnostic criteria, not just for AD, but also for all the depressive disorders bordering AD.

Further epidemiological information on adjustment disorder can be gleaned by an examination of specific populations exposed to stressful events. One of the largest studies looking at the prevalence of adjustment disorders in a clinical population examined patients on entry to the clinical services of the Western Psychiatric Institute. Intake involved a diagnostic assessment of over 11,000 individuals of all ages using a semi-structured assessment form and DSM-III criteria sheets. Ten percent of the sample was found to have adjustment disorders, making it the second largest diagnostic category. In children and adolescents under 18 years of age, over 16 percent had adjustment disorders. In adults, the female-to-male ratio was approximately two-to-one. In children, there was only a slight preponderance of females with the diagnosis.

Among patients known to have endured stressful experiences, the prevalence of an AD diagnosis is far greater. For example, the prevalence of a stress response disorder was analyzed among Danish-born citizens in Denmark who sought treatment at an inpatient or outpatient facility between 1995 and 2011 and received a diagnosis related to a stressor. Among this group, 90.5 percent received a diagnosis of severe stress or adjustment disorder as the primary diagnosis. The study cites an incident rate of 97 for adults and 37.6 for children, with a higher representation of women receiving this diagnosis. This lends further support for the current move to include adjustment disorders in the DSM-5 category of Stress Response Syndromes.

Decades of research have found a higher prevalence of adjustment disorder in patients in the medical setting—especially patients with cancer and those receiving palliative care. In a study of 107 patients with head and neck cancers, 14 (12.1 percent) had a diagnosis of adjustment disorder. In another study of 55 women following a first recurrence of breast cancer, 19 (35 percent) were diagnosed with adjustment disorder.

Other specific unfortunate life events put individuals at particular risk for the development of an adjustment disorder. One such group is described in a Danish study examining the prevalence of adjustment disorder in women who have experienced a stillbirth—loss of a fetus anytime following 22 weeks. This large-scale study showed that women had a transient increased risk for psychiatric hospitalization within the first

month following fetal death. The most likely diagnosis in this group was an adjustment disorder. Another such group involves individuals who have been incarcerated.

A further group at risk involves individuals who have experienced what may be termed bullying, harassment, or mobbing in the workplace. One German study used the World Health Organization and International Labour Office definitions of "repeated and over time, offensive behavior through vindictive, cruel or malicious attempts to humiliate or undermine an individual or groups of employees." This typically involves social isolation at work, intrusion into privacy, verbal attacks, or intimidation. The study reports multiple physical and psychiatric sequelae of the experience of harassment. While the specific psychiatric diagnoses are not specified, harassed versus nonharassed peers are over four times as likely to experience depressive affects, and nearly three times as likely to feel sustained anxiety.

One contributor to this discrepant data has been the lack of a structured diagnostic instrument for AD as is available for most other psychiatric disorders, although a new instrument has been developed. A recent study described the development of a 29-item self-assessment—the Diagnostic Interview for Adjustment Disorders (DIAD)—used in two sample populations of patients diagnosed with adjustment disorder. The assessment tool identified three major characteristics of patients with AD: intrusive memories of the event, an avoidance of feelings, and a failure to adapt. The authors conclude that this tool will likely facilitate future research in the area of adjustment disorders. The DIAD covers the domains of stressful life events, onset and duration of the stressor, and symptoms of and distress and impairment from the stressor, along with relationship between these and the stressor. Further studies will of course be needed to establish psychometric validity but the development of the DIAD is certainly a step forward in diagnostic rigor for AD.

ETIOLOGY

Current knowledge regarding the relationship between a stressful life experience and the development of symptoms remains complex. Unresolved questions persist in the literature regarding the etiology of adjustment disorder. Many authors in the field of adjustment disorders criticize the idea of linking a single stressor to a symptom complex. As discussed earlier, if a stressor does impact an individual, the model of linear causality has less utility than other complex models of causality. Recent literature has considered symptom formation following a stressor from a variety of paradigms. These include psychological models of processing stressful life experiences; psychodynamic models of affect attunement, the impact of early life experiences and ongoing psychiatric issues on the vulnerability to stress, and biological vulnerability to stressful

life experiences.

One investigator describes the integration of stressful life experiences as a series of responses. The first of these involves the realization that a stressful event has occurred. This is then followed by a period of suppression of this information. There is then alternating intrusion of the event and attempts at suppressing it, and a working through of the information which results in an integration of the material into one's cognitive schema or in symptom formation.

Another author has elaborated a theory of associative network formation following a significant stressful experience. The network consists of three elements: stimulus information about the stressor; information about cognitive, behavioral, and physiological reactions to the stressor; and meaning elements representing basic assumptions and their violation. Symptoms occur when one is reexposed to an element of the associative network. The literature provides an example of meeting one's former boss, and how that may trigger thoughts about the unexpected dismissal from work and why this happened. Treatment would be aimed at modifying this associative network.

In another model proposed by a group of researchers, the etiology of symptoms comes not only from a representation of the event in one's memory, but also from individual differences in the meaning and understanding of the stressor and its sequelae. If the meaning of a stressor to an individual greatly disturbs one's autobiographical narrative, it will lead to more disturbing psychological consequences. In such cases, it becomes more difficult to reorganize one's understanding of previous experiences to reestablish a stable sense of self.

Contemporary psychodynamic theory focuses on the context in which an event occurs, and how this leads to symptom formation. The intersubjective paradigm delineates the critical nature of attunement from others following a significant stressor. A stressor leads to the generation of an affective experience, which in turn leads to the desire for understanding of this affect. If others fail to attune to this affect but rather criticize, reject, or neglect the individual, he is left to cope on his own, leading to symptom formation. Though generally supported, this idea has been questioned in some studies which infer that excessive social support can in fact lead one to feel incapable of coping unaided, in turn leading to lower self-esteem and distress.

Psychodynamic theory also posits that it is the lack of an experience of affect or feelings about the stressor that generates problems. Individuals may not have the opportunity to experience their feelings about an event or toward a person because the emergence of these feelings represented a danger or threat for whatever reason. Perhaps it threatened a caregiver, perhaps there was not sufficient time, or perhaps this was already a learned style. The emergence of these feelings signals danger and a sense of anxiety.

Individuals develop psychological defenses to mitigate this anxiety which in themselves can cause problems. For example, anger at an individual may be turned against the self, resulting in symptoms of depression and demoralization.

Some evidence exists that individuals with greater deprivation early in life are at greater risk for developing symptoms following a life stressor. Current literature points to the importance of childhood experiences in the vulnerability to symptom formation in response to later life stress. Evidence from a variety of sources suggests that early experience with diminished control may foster a cognitive style characterized by an increased probability of interpreting or processing subsequent events as out of one's control. This may represent a psychological vulnerability to the development of anxiety. Research in the area of adverse childhood events further underscores the contribution of early life experiences in the generation of depressive affect later in life. For many, the experience of adversity in childhood is associated with the development of further stressful life events over the course of development and adulthood. Both early adverse childhood events as well as these associated later adult experiences of adversity seem to contribute partially and significantly to the development of the experience of the affect of depression.

Recent literature has furthermore examined the question of resilience to stressful life experiences. Most of this research has focused on resilience in the face of adversity during childhood and adolescence. Studies of children exposed to stressors including war, family violence, poverty, and natural disasters have revealed a consistent pattern of individual characteristics associated with successful adaptation. These include good intellectual functioning, effective self-regulation of emotions and attachment behaviors, a positive self-concept, optimism, altruism, a capacity to convert traumatic helplessness into learned helpfulness, and an active coping style in confronting a stressor.

Biological theorists have postulated the role of numerous neurotransmitters in the generation of pathology in individuals most at risk at the time of a stressful life event. Following a review of eleven neurochemical, neuropeptide, and hormonal systems, one author postulates that changes following acute stressful experiences may lead to long-term risk factors for the future capacity to cope with stress. This "allostatic load" of neurochemical networks may relate to individuals resilience in the face of stressful life events. One prediction is that individuals in the highest quartile for measures of HPA axis, corticotrophin-releasing hormone (CRH), locus coeruleus-norepinephrine, dopamine, and estrogen activity, and the lowest quartile for dehydroepiandrosterone (DHEA), neuropeptide Y, galanin, testosterone, 5HT1A receptor, and benzodiazepine receptor function will have the highest index for psychobiological allostatic load and increased risk for

psychopathology after exposure to stress. These neurotransmitters act through brain regions including the amygdala, hippocampus, locus coeruleus, and prefrontal cortex, deemed critical in the modulation of a stressful experience.

DIAGNOSIS AND CLINICAL FEATURES

A precipitating experience in the form of a stressful event(s) or life situation is the sine qua non for diagnosis of adjustment disorder. It is characterized by the development of emotional or behavioral symptoms in the context of identified psychosocial stressor(s). The symptoms vary widely in form and severity and the person may or may not be aware of the stressor as being significant. In contrast to other DSM-5 disorders, the adjustment disorder criteria have no clear and specific symptom profile that defines the condition.

It is deemed clinically significant if accompanied by marked distress "out of proportion to the severity or intensity of the stressor" or by significant impairment in social and occupational functioning. The challenge for the clinician in these situations is to differentiate a reasonable and expected response to psychosocial stressors from that which might indicate a diagnosis of adjustment disorder, thus a reaction within the normal range of human experience must be included in the differential.

Many individuals who seek psychiatric care following a significant life event may desire help in the form of psychotherapy or psychopharmacology, yet their response may not in fact be outside of what would be expected. One author notes that there is no way to categorize "problems of daily living" in the current system, yet, with more reliance on mental health practitioners for a role that may have once belonged to families, community support, and religious affiliations, it is critical to understand that such reactions may not be pathologic.

Adjustment disorders may occur in any age group. Some studies have identified school problems as the most frequent precipitant in adolescents while marital, financial, and health issues are more common in adults. Several studies have described the presenting features of adjustment disorders in clinical samples. Common symptoms of adjustment disorders identified in a large study of adjustment disorder included depressed mood, low self-esteem, suicidal behavior, increased motor activity, hypervigilance, impulsivity, and substance use. In most of these samples, depressive symptoms have been the most prominent.

The literature also points to the experience of embitterment following significant life events. In such cases, the emotional response following a significant life event is embitterment and feelings of injustice. There is an impairment in emotional modulation and patients may find joy in thoughts of revenge. Additional symptoms may include feelings of helplessness, self-blame, rejection of help, suicidal ideation, dysphoria, aggression, down-

heartedness, seemingly melancholic depression, unspecific somatic complaints, loss of appetite, sleep disturbance, pain, phobic symptoms in respect to the place or to persons related to the event, and reduced drive. One author proposes a concept of Posttraumatic Embitterment Disorder to characterize this experience following a negative life event. Studies of adjustment disorders using structured diagnostic instruments have reported high levels of comorbidity. The degree of comorbidity may be an important determinant of the level of impairment and risk for poorer outcomes n persons with adjustment disorder. In one study of Medicare patients, 54 percent of patients diagnosed with adjustment disorder also displayed abnormal illness behavior when using the structured interview for Diagnostic Criteria for Psychosomatic Research (DCPR). Symptoms of illness behavior documented included disease phobia, health anxiety, tanatophobia, nosophobia, and illness denial. One-third also met criteria for somatization. Demoralization, alexithymia, and unstable mood were diagnosed in nearly 30 percent of patients.

Several studies have reported a significant association of adjustment disorders with suicidal ideation. In one study of patients who had attempted self-harm and presented to the emergency room, those with adjustment disorders had more impulsive acts, were more often associated with alcohol abuse, and attempts at self harm occurred earlier in the course of the disorder than for those patients with a diagnosis of major depression. These findings underscore the seriousness of adjustment disorders in a subset of individuals; subthreshold conditions may carry high morbidity including the risk of self-harm and death.

Subtypes

Six subtypes of adjustment disorders have been described historically based on the predominant symptom presentations. DSM-5 continues the subtypes described in DSM-IV-TR. These do not identify specific symptom profiles for each of the subtypes; however, attempts have been made to more accurately characterize them. Notable symptoms for individual adjustment disorders subtypes are (1) depressed mood: depression, hyposomnia, low self-esteem, and suicidal indicators; (2) anxiety: generalized anxiety, increased motor activity, and situational anxiety; (3) disturbance of conduct: impulsivity, lack of insight, and violent behavior; and (4) mixed disturbance of emotions and conduct: excessive alcohol ingestion, suspiciousness, hostility, defrauding behavior, and homicidal ideation. A study in a population of primary care patients found that the anxious subtype was the most prevalent of the adjustment disorders, twice as prevalent as the depressive subtype, and far more prevalent than other forms.

There was a movement to include two additional subtypes in DSM-5. The first of these involved abnormal grief, but this was ultimately included

in a separate section of DSM-5. The second involved a subtype of PTSD/Acute Stress Disorder. This was to include individuals with symptoms that don't rise to the full diagnostic criteria of either of these disorders but met many of the criteria. Both subtypes underscore the relationship between these diagnoses, as all involve a stressor and ensuing symptomatology. Ultimately, however, the PTSD/ASD subtype was not specifically included in the DSM-5 but was acknowledged through the inclusion of the category of Other Trauma- and Stressor-Related Disorders (Table 25–1).

A 39-year-old married woman, in good health, with no previous psychiatric difficulties presented to the emergency room reporting that she had overdosed on a handful of antihistamines shortly before she had arrived. She described her problems as having started 2 months earlier, soon after her husband unexpectedly requested a divorce. She felt betrayed after having devoted much of her 10-year marriage to being a wife, mother and homemaker. She describes a stressful year involving a car accident that she was in 9 months prior, ultimately leading to a miscarriage of an early, unexpected pregnancy. She was sad, tearful at times, and occasionally had difficulty sleeping. Otherwise she had no vegetative symptoms and enjoyed time with her children, parents, and friends. She describes feeling desperate and suicidal after she realized "he no longer loved me." Following crises intervention in the emergency setting, she responded well to individual psychotherapy over a 3-month period, focused on her anger at her husband and over time how she has developed a mechanism of experiencing anxiety, sadness, and withdrawal preferential to her actual feelings toward people. The therapy facilitated grief over the end of her marriage as well as her facing feelings of loss she had never considered following the car accident and miscarriage. She occasionally required benzodiazepines for anxiety during the period of treatment. By the time of discharge from care, she had returned to her baseline function, though felt like a stronger and more authentic person in general. She came to terms with the possibility of life after divorce, and was exploring her best options under the difficult circumstances.

Table 25–1.
DSM-IV-TR Subtypes of Adjustment Disorders

With Depressed Mood
With Anxiety
With Mixed Anxiety and Depressed Mood
With Disturbance of Conduct
With Mixed Disturbance of Emotions and Conduct
Unspecified

DIFFERENTIAL DIAGNOSIS

A wide range of psychiatric disorders frequently occurs in or is exacerbated by stressful life events. The largest number of studies investigating the differential diagnosis of adjustment disorder has compared AD, usually depressed mood subtype, to the upper threshold of major depression and the lower threshold of patients without a diagnosable illness according to DSM. This research has supported the middle or transitional position of AD in terms of symptom severity, impairment of function, course of illness,

and prognosis. However, the resemblance of patients with AD to “normal” patients raises the interesting question of whether the syndrome is most appropriately understood as a psychiatric illness or as “problems of living” that become overwhelming as a result of an individual’s lack of internal resources and external supports.

The presence of a stressor in the adjustment disorder diagnostic criteria places acute and posttraumatic stress in the logical differential diagnosis. The nature of the stressor and the accompanying constellation of affective and autonomic symptoms have been far better defined and characterized for both stress disorders. In contrast, the stressor in adjustment disorder can be of any severity, with a wide range of possible symptoms. When the response to an extreme stressor does not meet the acute stress or posttraumatic disorder threshold, the adjustment disorder diagnosis would be appropriate, reinforcing the logic of placing them under one diagnostic heading.

Some authors have suggested that adjustment disorder within a stress context may be a *fruste forme* PTSD for patients lacking either the severity of the stressor or the required number and type of symptoms for posttraumatic diagnosis. Acute stress disorder is even more closely allied to AD in that it also has a limited time course and often a better outcome. Identifying the factors of genetics, environment, or their combination, which account for the different developmental course of stress-related disorders is an area warranting further research. With the placement of the adjustment disorders in the category of Stress Response Syndromes, this is more likely to happen in future years.

The adjustment disorder subtypes need to be distinguished from subthreshold types of the major mental disorders, and the so-called not otherwise specified (NOS). Affective disturbance, anxiety, or conduct problems associated with an identifiable stressor render the diagnosis adjustment disorder rather than the NOS categories in DSM-5. There is some evidence that these diagnoses do in fact represent distinct entities. In one study at the Rhode Island Hospital involving 3,400 patients diagnosed with either depression NOS or adjustment disorder, important differences were identified. The patients with depressive disorder NOS were more often diagnosed with social phobia and a personality disorder. The patients with depressive disorder NOS reported more anhedonia, increased appetite, increased sleep, and indecisiveness, whereas the patients with adjustment disorder reported more weight loss, reduced appetite, and insomnia. There was no significant difference between the groups in overall level of severity of depression or impaired functioning. The patients with depressive disorder NOS had a nonsignificantly elevated morbid risk of depression in their first-degree relatives. An additional review comparing adjustment disorders and depressive episodes in clinical practice focused on the role of personality disorders. Though the rate of personality

disorders in patients with adjustment disorders is fairly high, the rate in patients with depressive disorders is substantially higher. Interestingly, if the symptoms of adjustment disorder persist for over 6 months after the stressor or its consequences have resolved, the diagnosis would have to change to the appropriate NOS category if symptoms remain subthreshold or to a more specific psychiatric diagnosis.

Psychosomatic physicians have long noted the difficulty diagnosing a mood disorder in hospitalized or medically ill patients. The differentiation of adjustment disorder from a mood disorder due to a general medical condition or demoralization is complex. While the medical condition may be the obvious stressor, affective symptoms may arise as a direct physiological consequence of some medical conditions or the stress of prolonged hospitalization or disability. In medical and other settings, care must be taken to distinguish between the reasonable and the expected response to psychological stressors, which many authors believe includes demoralization and the inordinate response that may suggest an adjustment disorder. For the clinician it may be most useful to see the disorders along a dynamic continuum of affective stress responses where the intensity of distress and nature or degree of functional impairment are more determinative of the need for psychiatric intervention than the actual diagnosis. Normative data on stressor response patterns that differentiate pathological from abnormal responses are not available in any rigorous fashion and, as the case below illustrates, clinical and value judgments will continue to play a significant role in the diagnosis of adjustment disorders.

A 59-year-old unmarried man has been hospitalized for over 3 months with complications from a perforated gallbladder that required repeated surgeries. He has spent most of the time in the intensive care unit with a nasogastric tube and in severe pain. At times, the surgeons tell the patient that he can recover and return to his job as a security guard while, at other points, the surgical team informs him that the best they can hope for is to discharge the patient to a nursing home. As the patient goes into his fourth month of hospitalization, the nurses notice that for several days a week the patient is less communicative, seems alternatively sad and then irritable, sometimes refuses to work with physical therapy, and expresses doubts he should ever have agreed to the initial surgery. The surgeon is concerned the patient is "giving up" and requests a psychiatric consultation "for depression and whether the patient would benefit from medication." The resident on the psychosomatic medicine service, after reviewing the chart and interviewing the patient, tells the attending, "I would have bad days too if I had been this sick in the hospital this long." The attending agrees and informs the surgeons that "the patient's behaviors are within the normal range of human experience when confronted with protracted serious illness."

COURSE AND PROGNOSIS

Historically, adjustment disorder has been viewed as a transitional diagnostic category and by definition is not an enduring diagnosis. The proposed etiology of the disorder assumes that it would not occur without a stressor and that symptoms do not persist beyond 6 months after the stressor or its consequences have terminated. If they do persist, another

psychiatric diagnosis would presumably be assigned. The symptoms therefore have two prognostic courses: either they resolve or progress to a more serious illness. Yet a study with children and adolescents showed that up to one-third had symptoms of adjustment disorder which persisted over 6 months, and one-fourth still remained symptomatic after a year. Furthermore, the youth had significant morbidity and poor outcome in contrast to the benign profile found in other research. The majority of studies in diverse settings do provide evidence regarding a shorter duration of illness and hospitalization of patients with adjustment disorder compared to other Axis I conditions.

Attempts have been made by a number of investigators to address the question of whether an initial episode of adjustment disorder predicts the subsequent development of more specific and severe pathology. Multiple studies have shown a fairly benign course for adults, with a lower likelihood of recovery in adolescents. A 5-year follow-up study at the University of Iowa showed a recovery rate of 71 percent in adults versus 44 percent in adolescents. An additional 8 percent of adults were well at follow-up but had intervening problems, compared to 13 percent of the adolescent group. Almost 70 percent of the adolescent group was ill during the 5-year follow period, while less than 30 percent of the adult group experienced psychiatric disorders in the study period. While most of the adults developed major depressive disorder and alcohol abuse, the adolescents developed a wider range of major psychiatric disorders including schizophrenia, bipolar disorder, antisocial personality disorder, drug abuse, and major depressive disorder. Chronicity of symptoms and behavioral disturbance were the best predictor of a poor outcome. The specific adjustment disorder alone did not predict the follow-up diagnosis. In studies following disasters such as Hurricane Katrina in the Gulf, specific factors have been found to decrease the risk of symptoms. These included excellent support systems, high income and education level, successful mastery of past traumatic events, timely information, availability of recovery services, compassionate understanding from health care providers, and good follow-up.

Research over the last 5 years has disclosed a risk for suicide, especially in adolescent patients with adjustment disorder, not previously fully appreciated. A recent study of 119 patients with adjustment disorder indicated that 60 percent had documented suicide attempts in the past and 96 percent had been suicidal during their admission to the hospital. Fifty percent had attempted suicide immediately prior to their hospital admission. Comorbid diagnoses of substance abuse and personality disorder contributed to the suicide risk profile.

Psychological autopsies of 19 adolescent suicides found that one fifth of the victims had an adjustment disorder but over half had a personality disorder as well. The patients who completed suicide were dealing with

multiple psychosocial pressures and stressful life events and few were receiving treatment, suggesting that inadequate coping may have played a role in the deaths. Autopsy results also discovered that the suicidal process was much shorter and more rapid in its evolution in adolescents with adjustment disorder than in those with other diagnoses and that most of the victims did not have a history of emotional disturbance. Together, these and other studies serve as a cautionary tale to clinicians to reconsider the conceptualization of adjustment disorder as one of significant suicidal risks deserving of careful monitoring of symptoms and prompt intervention.

A 17-year-old high school senior experienced rejection in his first serious relationship. In the weeks after the end of the relationship he began to exhibit dysphoric mood accompanied by anxiety and psychomotor agitation. He had received counseling in junior high school when his parents divorced, began using alcohol and marijuana, and had been suspended in his freshmen year for fighting. A month after the break-up, he began to tell his parents that life was no longer worth living without his former girlfriend. Two months later his parents came home from work and found him hanging in the garage with a note stating he could not go on alone.

TREATMENT

The adjustment disorders may present with subthreshold symptomatology across multiple symptom domains; thus, there is no single treatment intervention approach for the heterogeneous clinical manifestation of the disorder. There has been very little systematic study of the treatment individuals with adjustment disorder and other subthreshold conditions. Nonetheless, adjustment disorders are common diagnoses in clinical practice and attention to this and other subthreshold conditions may potentially forestall the development of more severe psychiatric morbidity.

The clinician is often faced with the question of whether there is sufficient psychiatric morbidity to warrant treatment intervention. The treatment of individuals with adjustment disorder requires the careful assessment of the nature and severity of symptoms with consideration of the risk factors associated with poor outcome, such as poor premorbid functioning and persistent stressors. It is essential to understand the meaning of the stressor(s) to the patient and why it is associated with the presenting symptoms.

The primary goals of treatment are to relieve symptoms and the achievement of a level of adaptive functioning that is comparable, or in some situations better than the level of premorbid functioning. It is important to understand and facilitate those factors that mitigate the pathological response to stress, as well as the patient's level of vulnerability and capacity for adaptation. Treatment interventions should be designed to minimize the impact of the stressors on day-to-day function and to mobilize adaptive stress-coping mechanisms.

Specific treatment interventions that may be considered in individuals with adjustment disorder include supportive psychological approaches,

cognitive–behavioral, problem-solving techniques, and psychodynamic interventions. Concrete guidance with the resolution of stressors is often helpful. Short-term treatment may be adequate for many patients; however, more extended treatment may be appropriate in situations where individual characteristics predispose the patient to stress intolerance. Depending on the level of acuity and existing supports, crisis intervention and case management may be necessary on a short-term basis to preclude rapid progression, suicide attempts, and perhaps avoid hospitalization. Individual psychotherapy tends to be supportive in nature, rather than exploratory, though recent literature challenges this assumption. It is important to help the patient to better understand the meaning of the stressor and why it may have overwhelmed their coping mechanisms.

A recent study demonstrated large effect sizes in terms of therapeutic change in a group of patients primarily diagnosed with adjustment disorder. Though the proposed mechanism of change involved the development of a reliance on more mature defense mechanisms, only a moderate change in defensive functioning was observed. The authors suggested that longer-term psychodynamic psychotherapy or psychoanalysis might be helpful in these responsive patients to further develop a less defensive style. Another study shows promising results for the use of meditation training for patients with both anxious and depressive subtypes of adjustment disorders. One group has demonstrated the use of an online virtual reality cognitive–behavioral therapy for use at home as homework assignments for patients with adjustment disorders. Interestingly, in a Cochrane meta-analysis, only problem-solving therapies showed a reduction in the time for patients with the diagnosis of adjustment disorder to return to part- or full-time work.

Family, couples, and group therapies may have utility depending on the presenting circumstances. Many individual modalities may well have utility in a group. One study demonstrated good efficacy on the capacity to regulate emotions as a result of a group therapy with multiple modalities for women recently diagnosed with breast cancer.

The Dutch have actually developed practice guidelines for managing adjustment disorders in occupational and primary health care which aim to reduce the suffering, sick-leave, disability, and economic losses. The guideline recommendations are based on cognitive–behavioral therapy, stress-inoculation training, and graded activity with the goal of improving the problem-solving ability of workers. Recommendations have also been made for the management of patients with adjustment disorder who are actively deployed. These recommendations can be generally taken as minimizing the stressor, teaching psychological skills, and working with the meaning of the stressor, with ongoing suicide risk assessment. Medications should be used for symptoms of insomnia and anxiety in the short run, and antidepressants for more moderate-to-severe forms of depression.

There is a small but emerging clinical trial literature on the biological treatment of adjustment disorders, often with alternative agents. A randomized placebo-controlled trial of Gingko biloba in patients with generalized anxiety disorder, the anxious subtype of adjustment disorder, and cognitive decline for 4 weeks found that anxiety, aggression, and global clinical impression scales all showed statistically significant benefit. A naturalistic study in 52 adjustment disorder patients treated over 56 days with trazodone (Desyrel) found that nearly 65 percent obtained remission and another 26 percent responded to treatment. The medication was well tolerated and its efficacy was due to amelioration of insomnia, anxiety, and depressed mood. Obviously the results of these small and early trials require rigorous replication but they do offer some pharmacological options to complement the primarily psychosocial treatment of the disorder.

There is plenty of less-rigorous evidence from psychosomatic medicine and clinic settings suggesting it is reasonable to consider the judicious use of medications to treat the specific symptoms of with adjustment disorders. Typically, these are antidepressants and anxiolytics. It is often helpful and important to treat such issues as insomnia and severe anxiety with pharmacological interventions in the short run. The treatment of depressive symptoms with pharmacotherapy is less clear. Adjustment disorders represent the greatest growth area in the prescription of antidepressants, yet the evidence for the utility in improving depressive symptoms in adjustment disorders is scant. Selective serotonin reuptake inhibitors have been found to be useful in treating some subthreshold depressive symptoms and may benefit certain subtypes of adjustment disorders. It should, however, be emphasized that psychosocial strategies remain the mainstay of treatment with pharmacological intervention being a complimentary treatment modality.

FUTURE DIRECTIONS

Creating an interactive model that takes into account both stress and resiliency factors that are responsive to individual and cultural differences remains challenging. Questions abound regarding the adjustment disorders; for instance, what is the relationship between adjustment disorder and other stress-related disorders such as poststress disorder and acute stress disorder? Are subtypes of adjustment disorders more appropriately conceptualized as minor disorders of mood, anxiety, and behavior?

Some authors have suggested a better understanding of the adjustment disorders might require a radical choice. Either the diagnosis of adjustment disorder is not based on a definition that lends itself to rigorous investigation—thus, it should be reconsidered—or its pragmatic utility, despite methodological problems, needs to be recognized as a reason for

further research, particularly study of its epidemiological and therapeutic aspects. The latter option would require continued efforts to improve the specificity, reliability, and validity of the diagnostic category. The placement of the adjustment disorders in the Stress Response section of DSM represents a first step in this direction. Further research must also be directed towards developing optimal treatment approaches using randomized controlled trials and rigorous cost–benefit analyses. Ultimately, the usefulness of any diagnostic category is largely dependent upon the strength of its prognostic and therapeutic implications.

REFERENCES

Andreasen NC, Hoenck PR. The predictive value of adjustment disorders: A follow-up study. *Am J Psychiatry*. 1982;139:584–590.

Arbus C, Herqueta T, Duburcq A, et al. Adjustment disorder with anxiety in old age: Comparing prevalence and clinical management in primary care and mental health care. *Eur Psychiatry*. 2014;29(4):233–238.

*Arends I, Bruinvels DJ, Rebergen DS, et al. Interventions to facilitate return to work in adults with adjustment disorders. *Cochrane Database Syst Rev*. 2012;12:CD006389.

Baumeister H, Maercker A, Casey P. Adjustment disorders with depressed mood: A critique of its DSM-IV and ICD conceptualization and recommendations for the future. *Psychopathology*. 2009;42(3):139–147.

Cameron LD, Booth RJ, Schlatter M, Ziginskas D, Harman JE. Changes in emotion regulation and psychological adjustment following use of a group psychosocial support program for women recently diagnosed with breast cancer. *Psychoncology*. 2007;16(3):171–180.

Caneo C, Marston L, Bellon JA, Kng M. Examining the relationship between physical illness and depression: is there a difference between inflammatory and non inflammatory diseases? A cohort study. *Gen Hosp Psychiatry*. 2016;43:71–77.

Casey P. The "afterthought diagnosis": Rehabilitating adjustment disorders. *Expert Rev Neurother*. 2006;6(2):145–151.

*Casey P. Adjustment disorders: New developments. *Curr Psychiatry Rep*. 2014;16(6):451.

Casey P, Maracy M, Kelly BD, et al. Can adjustment disorder and depressive episode be distinguished? Results from ODIN. *J Affective Disord*. 2006;92:291–297.

Charney DS. Psychobiological mechanisms of resilience and vulnerability: Implications for successful adaptation to extreme stress. *Am J Psychiatry*. 2004;161:195–216.

Cornelius LR, Brouwer S, de Boer MR, Groothoff JW, van der Klink HH. Development and validation of the Diagnostic Interview Adjustment Disorder (DIAD). *Int J Methods Psychiatr Res*. 2014;23(2):192–207.

Courtwright AM, Salomon S, Lehmann LS, et al. The association between mood, anxiety and adjustment disorders and hospitalization following lung transplantation. *Gen Hosp Psychiatry*. 2016;41:1–5.

Einsle F, Kollner V, Bly S, Maercker A. Development and validation of a self-report for the assessment of adjustment disorders. *Psychol Health Med*. 2010;15(5):584–595.

Fernandez A, Mendive JM, Salvador-Carulla L, et al; DASMAP investigators. Adjustment

disorders in primary care: Prevalence, recognition and use of services. *Br J Psychiatry*. 2012;201:137–142.

*Fielden JS. Review: Management of adjustment disorder in the deployed setting. *Mil Med*. 2012;177(9):1022–1027.

Gorschenek N, Schwab R, Eckert J. Psychotherapy of adjustment disorders. *Psychother Psychosom Med Psychol*. 2008;58(5):200–207.

Gradus JL, Bozi I, Antonsen S, et al. Severe stress and adjustment disorder diagnoses in the population of Denmark. *J Traum Stress*. 2014;27:370–374.

Grassi L, Mangelli L, Fava G, et al. Psychosomatic characterization of adjustment disorders in the medical setting: Some suggestions for DSM-V. *J Affect Disord*. 2007;101:251–254.

Jacobson JC, Maytal G, Stern T. Demoralization in medical practice. *Prim Care Companion J Clin Psychiatry*. 2007;9(2):139–143.

Jones R, Yates WR, Williams S, Zhou M, Hardman L. Outcome for adjustment disorder with depressed mood: Comparison with other mood disorders. *J Affect Disord*. 1999;55:55–61.

Kramer U, Despland JN, Michel L, Drapeau M, de Roten Y. Change in defense mechanisms and coping over the course of short-term dynamic psychotherapy for adjustment disorder. *J Clin Psychol*. 2010;66(12):1232–1241.

Kryzhanovskaya L, Canerbury R. Suicidal behavior in patients with adjustment disorders. *Crisis*. 2001;22:125–131.

LaNoue M, Graeber D, de Hernandez BU, Warner TD, Helitzer DL. Direct and indirect effects of childhood adversity on adult depression. *Community Ment Health J*. 2012;48(2):187–192.

Linden M. Posttraumatic embitterment disorder. *Psychother Psychosom*. 2003;72:195–202.

Lorenz L, Bachem RC, Maercker A. The adjustment disorder – new module 20 as screening instrument: cluster analysis and cut-off values. *Int J Occup Environ Med*. 2016;7(4):215–220.

Maercker A, Forstmeier S, Enzler A, et al. Adjustment disorder, post-traumatic stress disorder, and depressive disorder in old age: Findings from a community survey. *Compr Psychiatry*. 2008;49(2):113–120.

McEwen BS. The neurobiology of stress: From serendipity to clinical relevance. *Brain Res*. 2000;886:172–189.

O'Donnell ML, Alkemade N, Creamer M, et al. A longitudinal study of adjustment disorder after trauma exposure. *Am J Psychiatry*. 2016;173(12):1231–1238.

Pelkonen M, Marttunen M, Henriksson M, Lonnqvist J. Suicidality in adjustment disorder: Clinical characteristics of adolescent outpatients. *Eur Child Adolesc Psychiatry*. 2005;14:174–180.

Polyakova I, Knobler HY, Ambrumova A, Lerner V. Characteristics of suicidal attempts in major depression versus adjustment reactions. *J Affect Disord*. 1998;47:159–167.

Skruibis P, Eimontas J, Dovydaitiene M, Mazulyte E, Zelviene P, Kazlauskas E. Internet based modular program BADI for adjustment disorder: protocol of a randomized controlled trial. *BMC Psychiatry*. 2016;16:264.

Starkman MN. The terrorist attacks of September 11, 2001 as psychological toxin: Increase in suicide attempts. *J Nerv Ment Dis*. 2006;194:547–550.

*Strain JJ, Friedman MJ. Considering adjustment disorders as stress response syndromes for DSM-5. *Depress Anxiety*. 2011;28:818–823.

van der Klink JJ, van Dijk FJ. Dutch practice guidelines for managing adjustment disorders in occupational and primary health care. *Scan J Work Environ Health*. 2003;29:478–487.

Zimmerman M, Martinez JH, Dalrymple K, Martinez JH, Cheminsky I, Young D. Is the distinction between adjustment disorder with depressed mood and adjustment disorder with mixed anxious and depressed mood valid? *Ann Clin Psychiatry*. 2013;25:257–265.

26

Personality Disorders

C. ROBERT CLONINGER, M.D., PH.D., AND DRAGAN M. SVRAKIC, M.D., PH.D.

Major advances in the current understanding of the psychobiology of human personality have occurred during the past 5 years and have led to substantial advances in the diagnosis and treatment of personality disorders (PDs). Until recently, access to inner mechanisms of the human mental functions was severely limited. Psychiatrists often skipped efforts to describe and understand healthy mental functioning and focused only on mental disorders. Psychologists assumed that they could describe the components of healthy functioning by qualitative studies or by factor analysis of traits used to describe people's observed behavior, but they lacked a realistic model of the complex interactions occurring within a person. As a result, alternative models of the psychobiological processes underlying personality expression and its disorders were often based on authoritative opinions and questionable assumptions rather than rigorous science that could promote progressive advance based on a solid foundation of facts.

Now methods have been developed that do provide insight into the fundamental mechanisms of healthy and unhealthy personality functioning using person-centered approaches to functional brain imaging and analysis of the interactions among multiple genetic and environmental causes of individual differences in personality. As a result, the diagnosis and treatment of personality and psychopathology can now be solidly based on simultaneous assessments of objective laboratory findings, mental status, and clinical history. Psychiatrists cannot depend on only clinical variation, as in DSM-5, or on only molecular variation as some continuous promise will suffice in the future. Both causal and clinical information can be utilized simultaneously to understand the complex developmental processes that influence personality and psychopathology.

For the first time, models of human temperament and character traits that were originally based on animal studies have now been confirmed in human beings by results of functional brain imaging and person-centered analyses of the multiple genetic and environmental causes of personality in the general population of people from various cultures. Such studies of healthy people provide insight into both their strengths and their vulnerabilities. In other words, the study of personality provides a way to

understand healthy functioning as well as its disorders, much as medical students study healthy functioning before they study disease.

From this scientific foundation, both neuropharmacological and psychosocial approaches to treatment of PDs and cultivation of healthy functioning have been developed to promote well-being and reduce the burdens of ill-being. Person-centered approaches to health promotion help people to cultivate the habits, goals, and values that facilitate self-awareness of what brings and maintains well-being, thereby reducing the burden of acute and chronic disease. The lifestyle that people cultivate and the situations in which they live shape epigenetic changes that alter the expression of genes of people, leading to differences in gene expression that are substantial even in monozygotic twins.

The person-centered approach to understanding personality and its disorders is what distinguishes psychiatry fundamentally from all other branches of medicine. A person is a self-aware human being, not a machine-like object that lacks self-awareness. Personality refers to all the characteristics that distinguish a constantly developing, self-organizing human being from a predictable machine-like object. In other words, personality refers to all the ways someone shapes and adapts in a unique way to an ever-changing internal and external environment.

Knowledge of the personalities of people provides the psychobiological context in which their behavior and relationships can be adequately understood as a complex adaptive process, which is crucial for all diagnosis and treatment. As a result, no psychiatric assessment is adequate without a description of the patient's personality and its development across the life span. The neglect of thorough personality assessment and its implications for health leads to people developing many chronic diseases, which are the source of most of the burden of disease in affluent societies. More than 70 percent of the burden of disease overall could be prevented by universal health care guided by effective personality assessment and treatment.

Even if a psychiatrist takes a reductive, rather than a holistic, view of psychological medicine, the assessment of PD is still essential in clinical practice because PDs are such common conditions in contemporary society. PDs occur in 10 to 20 percent of the general population and in about half of psychiatric inpatients or outpatients. People with PD have chronic impairments in their ability to work and love, tend to be less educated, drug dependent, single, unemployed, and to have marital difficulties. They consume a large portion of community services, social welfare benefits, public health, and prison resources.

About one half of all psychiatric patients have PD, which is frequently comorbid with other clinical syndromes. PDs and associated traits, such as low self-directedness, are predisposing influences on the full range of other psychiatric disorders (e.g., substance use, suicide, affective disorders, schizophrenia and other psychotic disorders, impulse control disorders,

eating disorders, and anxiety disorders). Comorbid PDs interfere with treatment outcomes of patients and increase personal incapacitation, morbidity, and mortality of these patients.

On the contrary, recovery of a positive quality of life in people with mental disorders is most strongly predicted by positive personality developments, such as increased self-directedness. Positive personality traits, such as high self-directedness and cooperativeness, serve as protective and recovery factors for many mental and other medical disorders and often improve the treatment outcome of the comorbid conditions, including psychoses. Both psychotherapy and medication management are benefited by the recognition and understanding of personality and its disorders. Optimal psychiatric treatment requires that recovery address the needs of the person for a positive quality of life in a hopeful and compassionate manner, rather than merely working to reduce symptoms and risk of harm. A person-centered psychiatry must be directed toward improving positive mental health as well as reducing negative symptoms of mental disorder.

At the turn of the 20th century, when Freud was pioneering his understanding of human nature, neuroses and neurotic problems were the dominant "psychopathology of everyday life." At the time, the central psychological task people were facing was to find socially acceptable ways to express their "asocial" impulses, predominantly aggression and sexuality. Obviously, behavior codes of present society differ substantially from the ethical conservativism at the turn of the twentieth century. Both aggression and sexuality have become much easier to express in the current ethically more liberal society. Instead of struggling with morality, people in contemporary society are faced with questions of identity, meaning, and choice. Now at the turn of the twenty-first century, one frequently encounters patients who seek help but cannot precisely describe their problems; who are more disappointed than anxious; who struggle with questions of purpose rather than guilt; who are ambivalent rather than inhibited; who feel empty rather than sad; and who manifest a peculiar inability to learn from experience as they tend to repeat maladaptive behaviors over and over again.

Despite the great importance of personality and its disorders, clinicians seldom have the time or inclination to ask all the questions specified in DSM-5 criteria for PDs. Most psychiatrists find the official criteria of little utility except for rare and extreme prototypic cases. Official criteria have failed to provide the practicing clinician a way to address personality assessment in an efficient and systematic way that is needed for optimal patient care.

The failure of DSM-IV-TR is largely the result of its reliance on categorical criteria despite the absence of evidence that PD is composed of discrete entities. Also, most dimensional systems for personality describe

the differences between people, rather than the intrapsychic processes by which an individual can grow and develop in mental health. As a result, most categorical and dimensional systems for personality have been stigmatizing and are not useful prescriptions for positive personality development. Persons with PD are frequently labeled as "aggravating," "demanding," or "parasitic" and are generally considered to have poor prognosis for emotional growth and maturation. Alternatively, they may be seductive or dependent and elicit inappropriate blurring of professional boundaries, such as sexual interest or the urge to rescue. These patients challenge the limits of physician's theoretical knowledge, clinical skills, and even personal maturity. In general, practitioners with low narcissism, high energy, and high tolerance are optimal for treating patients with PD. Recognizing and treating disorders of personality helps clinicians to develop their personal maturity and clinical skills.

Fortunately, it is now possible to carry out a brief assessment of personality and its disorders as part of routine psychiatric examinations in an efficient way that may actually make assessments of both personality and other psychopathology more reliable and valid by identifying risk and protective factors for psychopathology. Such assessments also serve to establish a therapeutic alliance because everyone likes to feel that he or she is appreciated and understood as a person, not a diagnostic case. People are impressed when a clinician can predict their motivation and behavior with understanding and accuracy. Therefore, the first part of this chapter discusses some basic definitions and then provide the essentials for clinical assessment of personality as an integrated part of the mental status and psychiatric history. Later sections expand upon the essentials to provide a fuller understanding of disorders of personality and their measurement, etiology, development, and treatment.

BASIC DEFINITIONS AND TERMINOLOGY

The terms personality, temperament, motivation, character, and psyche are often used interchangeably. This is both misleading and unjustified. In what follows, these concepts are reviewed to distinguish them with more clarity.

PERSONALITY

Personality is both complex and unique: on one hand, people differ greatly from one another in multiple components of behavior, and, on the other hand, each person expresses only one of his or her many potential lifestyles. The common characteristic of all existing definitions of personality is that they are functional, that is, they focus on questions related to motivation and mental adaptation of the organism. Specifically, Allport defined personality as the "dynamic organization within the individual of those psychophysical systems that determine his/her unique

adjustment to his/her environment."

Allport elaborated on this definition by explaining that the expression "dynamic organization" emphasizes that personality is an organized system ("unitas multiplex") that is constantly evolving and changing. The phrase "within the individual" means that personality refers to intrapsychic processes, not judgmental comparisons of one person to another. The term "psychophysical" means that personality is neither exclusively mental nor exclusively neural but a combination of the two. The expression "unique adjustment to the environment" has both functional and evolutionary significance pointing to personality as a mode of survival and, more generally, learning and adaptation, which is unique to each individual. Allport's use of the verb "determine" suggests that personality traits are determining dispositions that limit a person's behavioral repertoire.

It is now known that people are self-aware and have enough freedom of will to change their behavior in directions that are not necessarily predictable or determined by their past. Accordingly, Allport's definition has been modified as follows: human personality is *the dynamic organization of the psychobiological systems by which a person shapes and adapts in a unique way to a changing internal and external environment*. In other words, personality is a way of describing the dynamical processes that occur within the person to shape and adapt to life experiences by self-aware learning and proactive planning. The maturation and integration of human personality involves one's growing in self-awareness through experiences across a wide range of situations.

It has been widely accepted that personality develops through the interaction of hereditary dispositions and environmental influences. Waddington's notion of genetic canalization (or "epigenetic landscape") has been revised to include reciprocal interactions among genetic endowment, environmental stimulation, and changes in self-awareness across the life span. Genetic differences account for about half of the variance in differences between people for most normally distributed temperament traits. The remaining 50 percent of the variance in differences between people is attributed to environmental influences. Most of the environmental influences are unique to each individual, particularly for temperament traits, but parental rearing and sociocultural norms have a substantial influence on a person's character development. Of note, adoption studies suggest somewhat lower heritability of about 30 percent for personality traits. The higher heritability estimates from twins are the result of nonadditive genetic influences (e.g., higher-order interaction among alleles at each locus or among loci), which is the same for monozygotic twins but contributes little to the resemblance of other relatives. Epistasis (i.e., interactions among multiple gene loci) is so strong for both temperament and character that it was difficult to identify specific genes for personality until methods were developed to take epistasis into

account.

However, heritability estimates refer only to genetic influences on the differences between people at one point in time. People are able to change from one time to another as a result of growth in self-awareness in ways that are only partly constrained by their past genetic expression. For example, intelligence quotient (IQ) is highly heritable, but there has been increase of about 3.3 IQ points per decade that requires frequent restandardization of IQ tests. The increase in IQ scores in successive birth cohorts is called the "Flynn effect." The increase in IQ is attributable to skill in problem solving on the spot ("fluid intelligence"), such as recognition of similarities, rather than skills dependent on past individual education, such as vocabulary size ("crystallized intelligence"). Likewise, the differences between people in Self-Transcendence at one point in time is 49 percent heritable, but the prevalence of self-transcendent experiences increased from 48 percent in 1987 to 76 percent in 2000 in English people according to studies of the zoologist David Hay. These rapid changes in fluid intelligence and self-transcendence cannot be explained by genetic evolution. Character and fluid intelligence can develop rapidly within individuals in response to cultural developments that influence human society as a whole.

From the structural standpoint, personality can be decomposed into temperament, character, and psyche. Roughly speaking, temperament involves basic emotions, character involves rational concepts about self and interpersonal relations, and the psyche involves intuitive self-awareness and fluid intelligence. Self-awareness and fluid intelligence influence personality development substantially, so measures of temperament and character provide an incomplete understanding of personality development. Basic functions of personality are to feel, think, and perceive, and to incorporate these into purposeful behaviors.

TEMPERAMENT

Temperament refers to the body's biases in the modulation of conditioned behavioral responses to prescriptive physical stimuli. Behavioral conditioning (i.e., procedural learning) involves presemantic sensations that elicit basic emotions, such as fear or anger, independent of conscious recognition, descriptive observation, reflection, or reasoning. Pioneering work by A. Thomas and S. Chess conceptualized temperament as the stylistic component (HOW) of behavior, as differentiated from motivation (WHY) and content (WHAT) of behavior. Modern concepts of temperament, however, emphasize its emotional, motivational, and adaptive aspects. Specifically, four major temperament traits have been identified and subjected to extensive neurobiological, psychosocial, and clinical investigation in humans and other animals: Harm Avoidance, Novelty Seeking, Reward Dependence, and Persistence. It is remarkable

that this four-factor model of temperament can, in retrospect, be seen as a modern interpretation of the ancient four temperaments: individuals differ in the degree to which they are melancholic (Harm Avoidance), choleric (Novelty Seeking), sanguine (Reward Dependence), and phlegmatic (Persistence). However, the four temperaments are now understood to be genetically independent dimensions that occur in all possible combinations within the same individual, rather than as mutually exclusive categories.

Psychobiology of Temperament

Temperament traits of Harm Avoidance, Novelty Seeking, Reward Dependence, and Persistence are defined as heritable differences underlying one's automatic response to danger, novelty, social approval, and intermittent reward, respectively. These four temperament traits are closely associated with the four basic emotions of fear (Harm Avoidance), anger (Novelty Seeking), attachment (Reward Dependence), and ambition (Persistence).

Individual differences in temperament and basic emotions modify the processing of sensory information and shape early learning characteristics, especially associative conditioning of unconscious behavior responses and preattentive components of perception. Temperament is conceptualized as heritable biases in emotionality and learning that underlie the acquisition of emotion-based, automatic behavioral traits and habits observable early in life and relatively stable over one's life span.

Each of the four major dimensions is a normally distributed quantitative trait, moderately heritable, observable early in childhood, relatively stable in time, and moderately predictive of adolescent and adult behavior. The four dimensions have been shown to be genetically homogeneous and independently inherited from one another in large, independent twin studies in the United States, Australia, Sweden, Korea, and Japan. Temperamental differences, which are not very stable initially, tend to stabilize during the second and third years of life. Accordingly, ratings of these four temperament traits at age 10 to 11 years were moderately predictive of personality traits at ages 15, 18, and 27 years in a large sample of Swedish children. Whereas character traits develop in a specific direction toward socially favored cultural norms, the number of people who increase in Harm Avoidance, Reward Dependence, and Persistence is about the same as the number who decrease, so that the average of these traits remains the same from ages 20 to 45 years. Novelty Seeking decreases about 10 percent per decade as a person's impulse control matures.

The four dimensions have been repeatedly shown to be universal across different cultures, ethnic groups, and political systems on every inhabited continent. In summary, these aspects of personality are called temperament because they are heritable, manifest early in life, are

developmentally stable, and are consistent in different cultures. Temperament traits are like crystallized intelligence in that they do not show the rapid changes within increasing age or across birth cohorts that are observed for fluid intelligence and character traits. Table 26–1 summarizes contrasting sets of behaviors that distinguish extreme scorers on the four dimensions of temperament. Note that each extreme of these dimensions has specific adaptive advantages and disadvantages, so neither high nor low scores inherently mean better adaptation.

Table 26–1.
Descriptors of Individuals Who Score High and Low on the Four Temperament Dimensions

	Descriptors of Extreme Variants	
Temperament Dimension	**High**	**Low**
Harm avoidance	Pessimistic	Optimistic
	Fearful	Daring
	Shy	Outgoing
	Fatigable	Energetic
Novelty seeking	Exploratory	Reserved
	Impulsive	Deliberate
	Extravagant	Thrifty
	Disorderly	Orderly
Reward dependence	Sentimental	Detached
	Open	Aloof
	Warm	Cold
	Affectionate	Independent
Persistence	Industrious	Lazy
	Determined	Spoiled
	Ambitious	Underachiever
	Perfectionist	Pragmatist

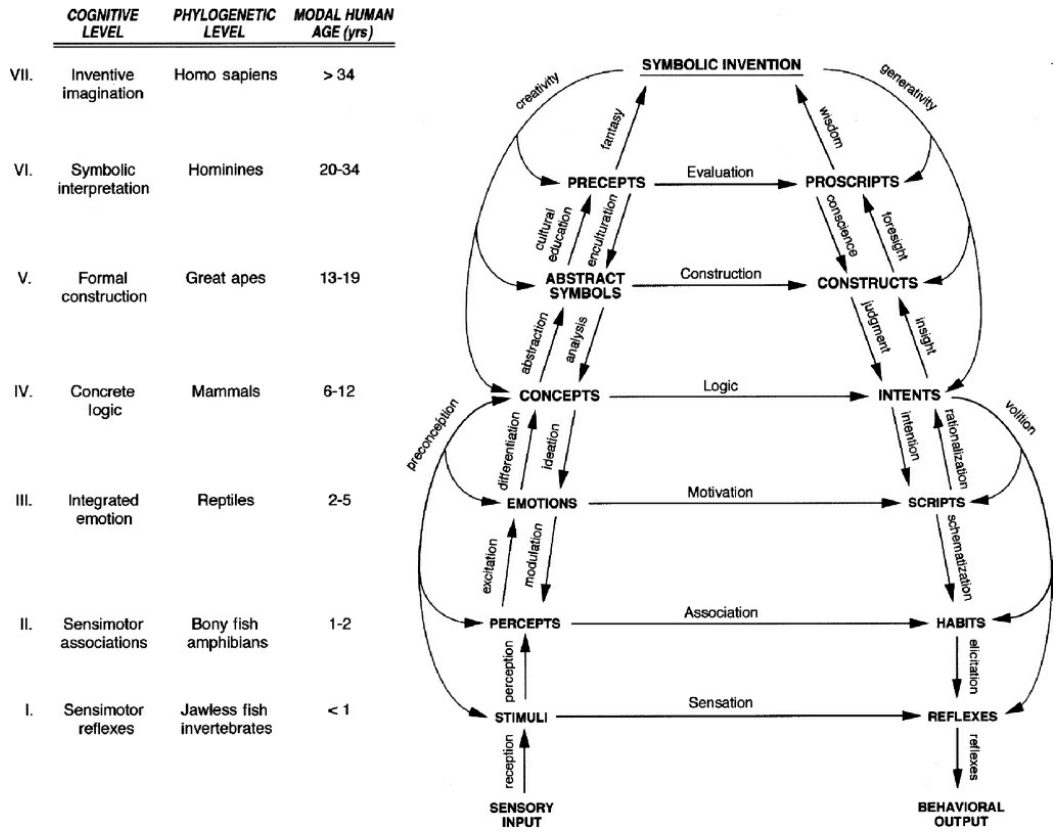

FIGURE 26–1. A hierarchical model of learning.

The component traits ("facets") for each of the four temperament dimension have distinct learning characteristics and correlate more strongly with one another than with other components of temperament. They share a common source of covariation that is strong and invariant regardless of changes in the environment and past experience. Each of the four temperament dimensions has unique genetic determinants according to family and twin studies, as well as studies of genetic associations with specific DNA markers.

The four dimensions of human temperament correspond closely to those observed in other mammals. Multiple levels in the phylogeny of learning abilities in animals from invertebrates to man indicate that learning has multiple component processes that are hierarchically organized and interact extensively throughout development (Fig. 26–1).

In Figure 26–1, temperament corresponds to the processes of sensation, association, and motivation that underlie the integration of skills and habits based on emotion.

Figure 26–1 is useful to specify the hierarchical phylogenetic and ontogenetic organization of learning and its relevance to the concepts of temperament and character, not to compare learning in humans at a particular age to learning in lower animals. Specifically, temperament and character are conceptualized on the basis of two types of memory and learning that have been described in humans and primates. These are

called procedural memory (i.e., behavioral conditioning) and semantic memory (i.e., long-term propositional or declarative memory). Temperament (the "emotional core" of personality) involves procedural memory regulated by complex distributed circuits in the cortico-striato-limbic system, primarily the sensory cortical areas, amygdala, and the caudate and the putamen. Procedural memory underlies associative learning and involves presemantic perceptual processing of visuospatial information and affective valence that can operate independently of abstract conceptual and/or volitional processes. In contrast, semantic learning involves higher cognitive functions of abstraction and symbolization. These two systems of learning and memory can be dissociated functionally from one another and also from a third memory system that is unique to human beings, as discussed later.

Neurobiological studies of animals have been highly informative about the functional organization of brain systems underlying procedural learning and temperament. The explicit animal models that were originally described on the basis of extensive work in rodents and other nonprimates provided hypotheses that have recently been confirmed by extensive work in humans using modern techniques of brain imaging, neurochemistry, and neurogenetics. The most comprehensive neurobiological model of learning in animals that has been systematically related to the structure of human temperament is summarized in Table 26–2. This model distinguishes four dissociable brain systems for behavioral inhibition (Harm Avoidance), behavioral activation (Novelty Seeking), social attachment (Reward Dependence), and partial reinforcement (Persistence).

The psychobiology of Harm Avoidance, Novelty Seeking, Reward Dependence, and Persistence is briefly described later.

Table 26–2.
Four Dissociable Brain Systems Influencing Stimulus–Response Patterns Underlying Temperament

Brain System (Related Personality Dimension)	Principal Neuromodulators	Relevant Stimuli	Behavioral Response
Behavioral inhibition (Harm avoidance)	GABA Serotonin (dorsal raphe)	Aversive conditioning (pairing CS and UCS) Conditioned signals for punishment and frustrative nonreward	Formation of aversive CS Passive avoidance Extinction
Behavioral activation (Novelty seeking)	Dopamine	Novelty CS of reward CS or UCS of relief of monotony or punishment	Exploratory pursuit Appetitive approach Active avoidance Escape

Social attachment (Reward dependence)	Norepinephrine Serotonin (median raphe)	Reward conditioning (pairing CS and UCS)	Formation of appetitive CS
Partial reinforcement (Persistence)	Glutamate Serotonin (dorsal raphe)	Intermittent (partial) reinforcement	Resistance to extinction

CS, conditioned stimulus; GABA, γ-aminobutyric acid; UCS, unconditioned stimulus.

Adapted from Cloninger CR. A systematic method for clinical description and classification of personality variables. *Arch General Psychiatry.* 1987;44:573–588.

Harm Avoidance. Harm Avoidance involves a heritable bias in the inhibition of behavior in response to signals of punishment and frustrative nonreward. It is observed as fear of uncertainty, shyness, social inhibition, passive avoidance of problems/danger, rapid fatigability, and pessimistic worry in anticipation of problems even in situations that do not worry other people. Adaptive advantages of high Harm Avoidance are cautiousness and careful planning when hazard is likely. The disadvantages occur when hazard is unlikely but still is anticipated, which leads to maladaptive inhibition and anxiety. People low in Harm Avoidance are carefree, courageous, energetic, outgoing, and optimistic even in situations that worry most people. The advantages of low Harm Avoidance are confidence in the face of danger and uncertainty, leading to optimistic and energetic efforts with little or no distress. The disadvantages are related to unresponsiveness to danger or unrealistic optimism with potentially severe consequences when hazard is likely.

The psychobiology of Harm Avoidance is complex. In animal studies, ascending serotonergic projections from the dorsal raphe nuclei to the substantia nigra inhibit nigrostriatal dopaminergic neurons and are essential for conditioned inhibition of activity by signals of punishment and frustrative nonreward. Benzodiazepines disinhibit passive avoidance conditioning by gamma-aminobutyric acid (GABA)-ergic inhibition of serotonergic neurons originating in the dorsal raphe nuclei. The anterior serotonergic cells in the dorsal raphe nucleus intermingle with the dopaminergic cells of the ventral tegmental area and both groups innervate the same structures (e.g., basal ganglia, accumbens, amygdala), providing opposing dopaminergic–serotonergic influences in the modulation of approach and avoidance behavior. The anterior serotonergic projections from the dorsal raphe to striatum, accumbens, amygdala, and frontal cortex are usually associated with 5HT2 receptors. Individuals who are high in both Harm Avoidance and Novelty Seeking are expected to have frequent approach–avoidance conflicts, as seen in cycles of binging and purging in bulimia. More generally, excessive behavioral inhibition (i.e., high Harm Avoidance) predisposes individuals to anxiety, depression, and low self-esteem. Effective antidepressant treatment lowers scores in Harm Avoidance, but higher scores in Harm Avoidance predict poorer responses

to antidepressants, including tricyclics and selective-serotonin reuptake inhibitors.

Recent neuropsychological studies confirm that Harm Avoidance is associated with individual differences in classical aversive conditioning, whereas other dimensions of personality are uncorrelated (Table 26–3). Harm Avoidance, and not other dimensions of temperament, is also replicably associated with individual differences in startle responses: individuals who are high in Harm Avoidance have a greater startle response when they are viewing unpleasant stimuli, whereas those who are low in Harm Avoidance have a reduced startle response when they are viewing pleasant stimuli. Other neuropsychological tests also confirm that Harm Avoidance is associated with behavioral inhibition, as measured by Mismatch Negativity, an indicator of early preattentive sensory processing and automatic attention in orienting responses. Hence, the average effects of Harm Avoidance on startle obscure the important nonlinear interactions between predisposition and mental state. Likewise, Harm Avoidance is associated with the Posner validity effect (the Posner task uses a detection reaction time paradigm with discrete presentation of simple visual stimuli in the periphery). Cues that correctly direct attention to the spatial location where the target stimulus will appear are called valid cues, whereas those that direct attention to incorrect locations are called invalid cues. In three large samples of healthy college students, those higher in Harm Avoidance had greater slowing of their reactions after invalid cues, or less benefit from valid cues, than others. Harm Avoidance is associated with greater sensitivity to the perception of fearful stimuli, which is mediated by strong functional connectivity between the amygdala and the anterior cingulate cortex. Even at rest, sustained fearful reactivity is sustained in people with the stable trait of Harm Avoidance, which is associated with reductions of slow oscillatory activity of the dorsal part of the medial prefrontal cortex in several independent studies.

Carriers of a gene that influences the activity of the serotonin transporter have increased Harm Avoidance, increased reactivity of the amygdala to fear, and increased risk of depression compared to others. Harm Avoidance scores explain 30 percent of the variance in the functional connectivity between the amygdala and perigenual cingulate, which are involved in perceptual processing of fearful stimuli. There is also a correspondence between structure and function: healthy human volunteers who are high in Harm Avoidance tend to have a greater volume of their right anterior cingulate gyrus and smaller volume of cerebellar white matter and cortex bilaterally (Table 26–3). Likewise, individual differences in Harm Avoidance are significantly negatively correlated with functional differences in activity in the dorsal part of the medial prefrontal network at rest in several independent studies. This includes significant negative correlations between regional blood flow and the score for Harm Avoidance

in paralimbic regions, such as the left parahippocampal gyrus and the left orbitoinsular junction, and various neocortical regions in the frontal, parietal, and temporal cortex. Positron emission tomography at NIMH with 18-FDG in 31 healthy adult volunteers during a simple continuous Performance Task showed that Harm Avoidance was associated with increased activity in the anterior paralimbic circuit, specifically the right amygdala and insula, the right orbitofrontal cortex, and the left medial prefrontal cortex. Subsequent fMRI studies have confirmed that individuals with high levels of Harm Avoidance have increased sensitivity to fearful stimuli mediated by activation of the amygdala and its connectivity with the perigenual cingulate cortex and the dorsal part of the medial prefrontal cortex. This stress sensitivity network also interacts with the Insular Salience Network, which involves the right anterior insula and its connections with the anterior cingulate and the dorsomedial prefrontal cortex. This activation pattern corresponds well to the 5HT2 terminal projections of the dorsal raphe. However, 5HT2 has been measured only in platelets (see Table 26–3).

Table 26–3.
Psychobiological Findings about Harm Avoidance

Variables	Observed Effects of High Harm Avoidance
Neuropsychology	
Aversive conditioning	Greater associative conditioning with punishment (r = 0.4)
Startle reflex	Potentiation of eye-blink response to aversive stimulus when HA is high and reduction when low
Fear perception	Greater sensitivity to fearful stimuli from stronger Amgydalar–Anterior Cingulate connectivity (r = 0.5)
Fear sustainment	Reduced slow low-frequency oscillations at rest in dorsomedial prefrontal cortex associated with greater amygdalar connectivity
Mismatch negativity	Greater amplitude of inhibiting early orienting response to deviant sounds in frontal cortex (r = –0.4)
Posner validity effect	Greater slowing of responses after invalid cues (r = 0.3)
Neuroimaging	
Insular salience network	Increased connectivity of right anterior insula with anterior cingulate cortex and dorsomedial prefrontal cortex
Amygdala activation	Increased sensitivity to fearful stimuli mediated by activation of amygdala and its connectivity with perigenual cingulate and dorsomedial prefrontal cortices
Volume of anterior cingulate cortex	Anterior cingulate cortex volume is larger in right and smaller in left hemisphere
Volume of cerebellar cortex	White matter and cortex smaller bilaterally

Mean diffusivity of striatum	Increased in pallidum bilaterally
Neurochemistry	
Platelet 5HT-2 receptor	Fewer receptors ($r = -0.6$)
Plasma GABA	Lower level ($r = -0.5$)
Neurogenetics	
5HT Transporter promoter gene	Greater reuptake activity
Tryptophan hydroxylase	Slower serotonin synthesis, especially in women reared in a hostile parental environment

Higher-plasma GABA levels have also been correlated with low Harm Avoidance. Plasma GABA has also been correlated with other measures of anxiety proneness and correlates highly with brain GABA levels. Finally, a gene on chromosome 17q12 that regulates the expression of the serotonin transporter has been found to account for 4 to 9 percent of the total variance in Harm Avoidance in most, but not all, of the tests of this relationship. These findings support a role for both GABA and serotonergic projections from the dorsal raphe underlying individual differences in behavioral inhibition as measured by the trait of Harm Avoidance.

Novelty Seeking. Novelty Seeking reflects a heritable bias in the initiation or activation of appetitive approach in response to novelty, approach to signals of reward, active avoidance of conditioned signals of punishment, and escape from unconditioned punishment (all of which are hypothesized to covary as part of one heritable system of learning). This trait is called Novelty Seeking because its most distinctive feature is exploratory approach behavior, as confirmed by fMRI studies showing increased striatal dopamine release activated by novel stimuli regardless of whether they predict reward or not in people high in Novelty Seeking (Table 26–4). Novelty Seeking is observed as exploratory activity in response to novelty, impulsiveness, extravagance in approach to cues of reward, and active avoidance of frustration. Individuals high in Novelty Seeking are quick-tempered, curious, easily bored, impulsive, extravagant, and disorderly. Adaptive advantages of high Novelty Seeking are enthusiastic exploration of new and unfamiliar stimuli, potentially leading to originality, discoveries, and reward. The disadvantages are boredom susceptibility, impulsivity, angry outbursts, potential fickleness in relationships, and impressionism in efforts. Persons low in Novelty Seeking are slow tempered, uninquiring, stoical, reflective, frugal, reserved, tolerant of monotony, and orderly. Their reflectiveness, systematic effort, and meticulous approach are clearly advantageous when these features are adaptively needed. The disadvantages reflect tolerance of monotony and lack of enthusiasm, potentially leading to prosaic routinization of activities.

Table 26–4.
Psychobiological Findings about Novelty Seeking

Variables	Observed Effects of High Novelty Seeking
Neuropsychology	
Reinforcement learning	More approach to novel stimuli and slower to react to neutral stimuli
Neocortical reactivity	Greater cortical reactivity to somatosensory (MEG) and novel stimuli (ERP)
Striatal–mediated behavioral activation	Greater striatal dopamine release activated by novel stimuli that do not predict reward
Impulsive decision making	Greater ventromedial prefrontal cortical activity at rest leads to reduced delay of gratification
Cortisol response to stress	Reduced at rest and after social stress
Neuroimaging	
Connectivity of ventral striatum	Greater connectivity with hippocampus, amygdala, and orbitofrontal cortex is activated by novel stimuli
Insular salience network	Reduced activity
Posterior cingulate volume	Greater gray matter volume in left hemisphere
Inferior frontal cortex volume	Smaller gray matter volume bilaterally
Neurochemistry	
Ventral striatal dopamine synthesis	Greater synthesis capacity (r = 0.8)
Midbrain dopamine autoreceptors	Reduced autoreceptor availability in substantia nigra/ventral tegmental area of midbrain, leading to more rapid sensitization to stimulants
Dopamine DRD2 receptor	Negative correlation of Novelty Seeking with DRD2 levels in insula bilaterally in parkinsonism
Neurogenetics	
Dopamine transporter gene	Greater reuptake activity
Dopamine DRD2 receptor gene	Moderates development of higher adult Novelty Seeking in response to harsh discipline in childhood

Mesolimbic and mesofrontal dopaminergic projections have been shown to have a crucial role in incentive activation of each aspect of Novelty Seeking in animals. For example, dopamine-depleting lesions in the nucleus accumbens or the ventral tegmentum of rats and mice lead to neglect of novel environmental stimuli and reduce both spontaneous activity and investigative behavior. Behavioral activation by dopaminergic agonists is dependent on integrity of the nucleus accumbens but not the caudate nucleus. In human studies, individuals at risk for Parkinson disease have low premorbid scores in Novelty Seeking, but not other dimensions of personality, supporting the importance of dopamine in

incentive activation of pleasurable behavior. The initiation and frequency of hyperactivity, binge eating, sexual hedonism, pathological gambling, Internet addiction, drinking, smoking, and other substance abuse, especially stimulants, are each associated with high scores in Novelty Seeking.

The psychobiological correlates reported for Novelty Seeking are summarized in Table 26–4. High scores correlate with greater capacity for synthesis of dopamine in the ventral striatum and reduced D2/D3 autoreceptor availability in the substantia nigra and ventral tegmental area in the midbrain, thereby leading to accentuated dopaminergic responses to novelty and other conditions that induce dopamine release. Human functional neuroimaging studies also show that the high Novelty Seekers have greater functional connectivity with the hippocampus, amygdala, and orbitofrontal cortex in response to novel stimuli than other people. Structural imaging studies indicate that people with higher Novelty Seeking have larger left posterior cingulate gray matter volumes and smaller inferior frontal gray matter volumes.

Evoked potential studies confirm that Novelty Seeking is associated with greater cortical reactivity and augmentation of stimulus intensity in response to somatosensory and novel stimuli. Novelty Seekers are also sensitive to both sedatives and stimulants: they are easily oversedated by benzodiazepines and overstimulated by amphetamines, leading to deterioration in their information processing. Their reaction times are slow to neutral stimuli, but they show greater approach to novel stimuli whether these predict reward or not.

The association of increased striatal dopamine release in response to novel stimuli with high Novelty Seeking is more specifically associated with higher density of the dopamine transporter, suggesting that Novelty Seeking involves increased reuptake of dopamine at presynaptic terminals, thereby requiring frequent stimulation to maintain optimal levels of postsynaptic dopaminergic stimulation. Novelty Seeking leads to impulsive decision making and various pleasure-seeking behaviors, including cigarette smoking, which may explain the frequent observation of low platelet MAO B activity because cigarette smoking has the effect of inhibiting MAO B activity in platelets and in brain.

Studies of candidate genes involved in dopamine neurotransmission (e.g., dopamine transporter and D4DR) have provided evidence of association with Novelty Seeking and no other dimension of temperament. The dopamine transporter, which is responsible for presynaptic reuptake of dopamine and a major site of action of drugs including stimulants such as methylphenidate, is encoded by a gene locus on chromosome 5p. Polymorphisms at this gene locus are associated with Attention Deficit Disorder and other disorders related to variation in Novelty Seeking. Likewise, polymorphisms at the DRD4 locus have been associated with

Attention Deficit Disorder, opioid dependence, and other traits related to Novelty Seeking.

Reward Dependence. Reward Dependence reflects a heritable bias in the maintenance of behavior in response to cues of social reward. Reward dependence is characterized by as sentimentality, social sensitivity, attachment, and dependence on approval by others. Individuals high in Reward Dependence are tender-hearted, sensitive, dedicated, dependent, and warmly sociable. One of the major adaptive advantages of high Reward Dependence is the sensitivity to social cues, which facilitates affectionate social relations and genuine care for others. The disadvantage is related to suggestibility and loss of objectivity frequently encountered with people who are excessively socially dependent. Individuals low in Reward Dependence are practical, tough-minded, cold, socially insensitive, irresolute, and indifferent if alone. The advantages of low Reward Dependence are personal independence and objectivity not corrupted by efforts to please others. Its adaptive disadvantage is related to social withdrawal, detachment, and coldness in social attitudes.

Noradrenergic projections from the locus coeruleus and serotonergic projections from the median raphe are proposed to influence such reward conditioning (Table 26–2). In animals, stimulation of the noradrenergic locus coeruleus or its dorsal bundle, or direct application of norepinephrine, decreases the firing rate of terminal neurons and increases their sensitivity to other afferents, so that targeted stimuli can stand out from nontargeted stimuli. In humans, short-term reduction of norepinephrine release by acute infusion of the alpha-2 presynaptic agonist clonidine selectively impairs paired-associate learning, particularly the acquisition of novel associations. The noradrenergic locus coeruleus is located at the same posterior level of the brainstem as the serotonergic median raphe, and both of these posterior monoamine cells innervate structures important to formation of paired associations, such as the thalamus, neocortex, and hippocampus. Neurophysiological studies show that the anterior temporal lobe decodes social signals, such as facial images and social gestures. Consequently, individuals high in Reward Dependence are expected to be particularly sensitive in their social communication, whereas those low in Reward Dependence are expected to be socially aloof. The “love hormone” oxytocin is released in response to sexual orgasm and its plasma level is positively correlated (r = 0.5) with individual differences in Reward Dependence but not other dimensions of personality.

The psychobiological correlates reported about Reward Dependence are summarized in Table 26–5. As predicted, Reward Dependence is associated with individual differences in formation of conditioned signals of reward. This is also supported by its association with individual differences in paired associate learning. People who are high in Reward Dependence also

show greater discrimination of facial emotions and social cues, which is mediated by their having larger oxytocinergic regions in the hypothalamus. In individuals with high Reward Dependence, the oxytocinergic regions of the hypothalamus have reduced connectivity with the amygdala, septum, and dorsal anterior cingulate in response to perceptual processing of facial emotions and social cues. Reward Dependence is also associated with greater opioid receptor availability in the ventral striatum bilaterally. Consequently, people who are high in Reward Dependence have readily form warm and caring attachments to other people.

Table 26–5.
Psychobiological Findings about Reward Dependence

Variables	Observed Effects of High Reward Dependence
Neuropsychology	
Reward conditioning	Enhanced associative pairing with rewards and not novelty without reward
Paired associates	Better learning of paired associations
Posner validity effect	Faster responses after valid cues
Perceptual processing of facial emotions and social cues	Better discrimination of facial emotions mediated by variability of regional oxytocin receptor activity
Mismatch negativity	Greater latency of preattentive MEG response to changes in tonal duration in left hemisphere
Neuroimaging	
Corticostriatal system	Greater gray matter volume and connectivity of striatum (caudate) with orbitofrontal cortex and temporal lobes
Caudate volume and activity	Greater size and activity
Oxytocinergic system	Larger oxytocinergic regions in hypothalamus with reduced connectivity to amygdala, septum, and dorsal anterior cingulate, which is a network activated by perceptual processing of facial emotions and social cues
Neurochemistry	
Striatal opioid availability	Greater opioid receptor availability in ventral striatum bilaterally
Plasma oxytocin	Higher oxytocin levels (r = 0.5)
Plasma cortisol	Higher morning cortisol when depressed (r = 0.3)
Neurogenetics	
5HT2 c receptor	Allelic association (effect size 2.0)
Epidermal growth factor	Epidermal growth factor A61G G/G polymorphism associated with higher RD in men

Persistence. Persistence reflects a heritable bias in the maintenance of behavior despite frustration, fatigue, and intermittent reinforcement. It is observed as industriousness, determination, ambitiousness, and

perfectionism. Highly Persistent people are hardworking, perseverant, and ambitious overachievers who tend to intensify their effort in response to anticipated reward and perceive frustration and fatigue as a personal challenge. High Persistence is an adaptive behavioral strategy when rewards are intermittent but contingencies remain stable. When the contingencies change rapidly, perseveration becomes maladaptive. Individuals low in Persistence are indolent, inactive, unstable, and erratic; they tend to give up easily when faced with frustration, rarely strive for higher accomplishments, and manifest a low level of perseverance even in response to intermittent reward. Accordingly, low persistence is an adaptive strategy when reward contingencies change rapidly and may be maladaptive when rewards are infrequent but occur in the long run.

Persistence can be objectively measured by the partial reinforcement extinction effect (PREE) in which persistent individuals are more resistant to the extinction of previously intermittently rewarded behavior than other individuals who have been continuously reinforced. Earlier work in rodents showed that the integrity of the PREE depends on projections from the hippocampal subiculum to the nucleus accumbens. This glutaminergic projection may be considered as a short circuit from the behavioral inhibition system to the behavioral activation system, thereby converting a conditioned signal of punishment into a conditioned signal of anticipated reward. This connection is probably disrupted in humans by lesions of the orbitomedial cortex that may have a specific antipersistence effect of therapeutic benefit to some patients with severe obsessive-compulsive personality disorder. Bilateral cingulotomy, which reduces Harm Avoidance only, is less effective in reducing persistent compulsive behavior than cingulotomy combined with orbitomedial lesions.

Table 26–6.
Psychobiological Findings about Persistence

Variables	Observed Effects of High Persistence
Neuropsychology	
Partial reinforcement extinction effect (PREE)	Greater resistance to extinction (quitting effort to obtain a goal) despite frustration from infrequent reward with high persistence
Positively biased reward perception	Neutral stimuli perceived as rewards
Cortical reactivity	Reduced cortical excitability at rest
Phonemic mismatch negativity	Greater strength of preattentive response to infrequent vowel sounds
Rey word list learning	Better learning without reinforcement

Intelligence and achievement	Higher academic achievement and slightly higher IQ
Neuroimaging	
Reward-seeking system	Greater activation of circuit involving ventral striatum, lateral orbitofrontal cortex (BA 47), and anterior cingulate in response to intermittent rewards, giving rise to the PREE
Striatal-mOFC connectivity	Greater connectivity of striatum and medial orbitofrontal cortex (mOFC)
Hippocampal–accumbens pathway	Glutaminergic projection from subiculum to nucleus accumbens is essential to PREE
Neurochemistry	
Salivary alpha amylase activity	Greater noradrenergic arousal on average (whether at rest or stressed)
Neurogenetics	
Interaction of DAT1 and 5HT2a	Greater persistence is associated with interaction of dopamine transporter (DAT1) and serotonin receptor 5HT2A
Brain-derived neurotrophic factor	Association with Val66Met polymorphism modulates the effects of parental rearing

These findings in animals suggested that the regulation of Persistence involved the ventral striatum, which includes the nucleus accumbens. This has recently been tested in humans and confirmed (Table 26–6). Individual differences in persistence are strongly correlated (r = 0.8) with responses measured by fMRI in a circuit involving the ventral striatum, orbitofrontal cortex/rostral insula, and dorsal anterior cingulate cortex (BA 32). Subjects low in persistence exhibited relative decreases in activity within this circuit whereas those high in persistence exhibited relative increases. Persistence scores also correlated with apparent selection bias such that subjects with high persistence scores made relatively more pleasant judgments at the expense of neutral judgments when viewing pictures from the International Affective Picture System.

MOTIVATION

Survival and reproduction, the basic drives for all animal species, are expressed in humans primarily through their experiential derivatives, that is, affects and emotions. In contrast to the limited motivational spectrum of the basic drives, emotions have an independent motivational power that makes them not only the primary motivational system for many people but also the personality processes that give meaning at least to rational materialists, who make up most of contemporary society.

Temperament traits of Harm Avoidance, Novelty Seeking, Reward Dependence, and Persistence, with their respective primary emotions of fear, anger, attachment, and ambition, are observable early in development. Research in children has demonstrated that, during the first

several months of life, fear and anger differentiate from the disposition to distress (a tendency to become upset and autonomically aroused easily and intensely). During this period, which is characterized by active regressive changes in the organization of neuronal circuits and density of synapses, especially in frontal, temporal, and limbic cortical areas, many other complex and social functions in human behavior emerge. These new functions are, however, organized around an early set of enduring behavioral and emotional dispositions referred to in this work as temperament traits.

A relatively limited spectrum of basic emotions is associated with the four temperament traits (Table 26–7). Depending on whether a particular temperament trait is high or low, certain emotions will tend to dominate one's motivation, perception, and behavior.

As can be inferred from Table 26–7, the same external stimulus is likely to elicit responses via activation of multiple temperament dimensions. For example, novel or unfamiliar stimuli elicit interest in approach in proportion to Novelty Seeking, as well as inhibition of approach in proportion to Harm Avoidance. Each temperament trait clearly involves an integration of multiple emotional drives that may be conflicting (competitive) or facilitatory (synergistic). Such shared environmental effects mean that the genetic and phenotypic structure of personality cannot be assumed to be the same.

Temperament involves a relatively small set of emotions associated with one's basic needs, for example, safety (so-called primary motives). Excessive fear and anger, associated with temperament, are motivationally monopolistic and take over the personality by altering perception, learning, and behavior in a biased way. However, under normal circumstances, after survival needs are met, the goals of normally developing personality change to include not only the integrity of the physical self but also the integrity of mental self (e.g., self-esteem). Normal personality development also adapts to numerous social goals (e.g., education, occupation, family) and a rich spectrum of secondary (social) emotions (such as shame and compassion). These "secondary," "social," or "growth" motives are closely functionally related to character development. Specifically, basic emotions of fear, anger, and excitement are transformed into more complex secondary emotions, such as carefulness, assertiveness, and joy, through the interaction with increasingly more complex internalized concepts associated with character. Although some basic character components develop early in life, such as trust and confidence, it is the completion of self-object differentiation ("Me" vs. "Not me") between 18 months and 3 years of age that sets the stage for the development of character traits and secondary emotions such as empathy.

Table 26–7.

Effects of Positive (+) and Negative (–) Reinforcement on Emotional State of Four Temperaments

Temperament	High Scorers		Low Scorers	
Dimension	+	–	+	–
Harm avoidance	Anxious (agitated)	Depressed (retarded)	Cheerful	Fearless
Novelty seeking	Euphoric	Angry	Placid	Stoical
Reward dependence	Sympathetic	Disgusted	Aloof	Indifferent
Persistence	Enthusiastic	Steadfast	Unstable	Discouraged

Table 26–8.
Effects of Positive (+) and Negative (–) Reinforcement on Emotional State of Three Characters

Character	High Scorers		Low Scorers	
Dimension	+	–	+	–
Self-directed	Hopeful	Resourceful	Vain	Shameful
Cooperative	Loving	Forgiving	Scornful	Revengeful
Self-transcendent	Joyful	Peaceful	Greedy	Miserable

The secondary emotions take over as primary motivators of further character development and maturation. Of note, they are not as monopolistic as the basic emotions and thus motivate development of more flexible and adaptable personality traits. As shown in Table 26–8, each of the three character traits is associated with a typical pattern of secondary emotions.

In conclusion, abnormal (deviant, immature) motivation derives from two or three dominant, monopolistic elementary emotional needs associated with survival. In contrast, mature motivation develops after basic needs are met and the person is freed to experience numerous secondary motives for growth in character and in self-awareness. This explains the motivational inflexibility and poverty of deviant personality and accounts for all the rich motivational diversity and flexibility of mature personality.

CHARACTER

Character refers to the mind, that is, the conceptual core of personality. It involves individual differences in self-concepts and object relations that reflect personal goals and values. In other words, character is what individuals make of themselves intentionally. Character is rational and volitional. Whereas temperament involves basic emotions such as fear and anger, character involves secondary emotions such as purposeful moderation, empathy, and patience, and in even more mature individuals, hope, love, and faith. As a result, character can be described as one's mental self-government, which involves executive, legislative, and judicial

functions. In what follows, the concept of character is outlined on the basis of the seven-factor psychobiological model of personality. The psychodynamic concept of character, which has contributed many illuminating clinical and theoretical formulations on the subject, is discussed in some detail as well.

Psychobiology of Character

Character (or the "conceptual core" of personality) involves higher cognitive functions, which include abstraction, symbolic interpretation, and reasoning. These higher cognitive functions are instantiated in complex distributed networks in the brain, which involve encoding of cognitive schemas by the hippocampus with long-term storage as semantic memories in neocortex. Such symbolic memory functions interact with temperament through cognitive processing of emotionally ridden sensory percepts regulated by temperament. This temperament–character interaction leads to the development of mature, realistic internalized concepts about the self and the external world. The executive, legislative, and judicial functions of mental self-government can be measured as three distinct character traits, which are called Self-Directedness, Cooperativeness, and Self-Transcendence, respectively. These character traits are adaptive, but their low ends are less advantageous because of a limited spectrum of circumstances in which immaturity, especially low Self-Directedness and Cooperativeness, are adaptive. Self-Directedness quantifies differences in the executive competence of individuals. A highly self-directed person is self-sufficient, responsible, reliable, resourceful, goal-oriented, and self-accepting. The most advantageous summary feature of self-directed individuals is that they are realistic and effective, that is, they are able to adapt their behavior in accord with individually chosen, voluntary goals. Individuals low in Self-Directedness are blaming, helpless, irresponsible, unreliable, reactive, and unable to define, set, and pursue internally meaningful goals. Such poor executive function, manifest as unrealistic behavior and lack of internal guidance, is rarely advantageous to the individual.

Cooperativeness quantifies differences in the legislative functions of individuals. That is, Cooperativeness allows people to make the rules and principles by which they get along with other people. Highly cooperative people conceptualize themselves as integral parts of human society. Such highly cooperative persons are described as empathetic, tolerant, compassionate, supportive, and principled. These features are advantageous in teamwork and social groups but not for individuals who must live in a solitary manner. People who are low in Cooperativeness are self-absorbed, intolerant, critical, unhelpful, revengeful, and opportunistic. They primarily look out for themselves and tend to be hostile and inconsiderate of other peoples' rights or feelings.

Self-Transcendence quantifies individual differences in the judicial functions of people. In other words, Self-Transcendence allows one to recognize when various rules apply in a particular context (i.e., in the narrative of intersecting lives). Self-Transcendence reflects the extent to which people conceptualize themselves as an integral part of the universe as a whole. Self-transcendent individuals are described as judicious, insightful, idealistic, contemplative, and spiritual. These traits are adaptively advantageous when people are confronted with ultimate situations such as suffering, illness, or death, which is inevitable with advancing age. They may appear disadvantageous in most modern societies where idealism, modesty, and meditative search for meaning might interfere with the acquisition of wealth and power. People low in Self-Transcendence tend to be pragmatic, objective, materialistic, controlling, and pretentious. Such individuals appear to fit in well in most Western societies because of their rational objectivity and materialistic success. However, they consistently have difficulty accepting suffering, failures, personal and material loses, and death, which leads to lack of serenity and adjustment problems particularly with advancing age. Contrasting sets of descriptors that distinguish high and low scorers are shown in Table 26–9, High and low scorers in each character dimension are distinguished by behavior traits that arise from differences in concepts that are each internally consistent but not logically falsifiable. For example, people low in Self-Transcendence live in the material world, skeptical of whatever they cannot prove objectively and use practically. In contrast, for highly Self-Transcendent individuals, the meaning of life goes beyond material things and includes intuitive awareness of what is beautiful, true, and good, to which materialists may be insensitive. Each set of belief appears to be internally consistent to those who hold them, who also regard their approach to life as realistic. Furthermore, spiritual people cannot prove that a materialist is wrong, or vice versa, because each has different intuitions of reality on which they base their concepts.

Character matures in a stepwise manner in incremental shifts from infancy through late adulthood. The timing and rate of transitions between levels of maturity are nonlinear functions of antecedent temperament configurations, systematic cultural biases, and experiences unique to each individual, which depend on individual differences in episodic memory or intuitive self-awareness that enables human beings to remember past experiences. The developing character traits (i.e., newly internalized concepts about one's self and the external world) optimize adaptation of temperament (i.e., early emotionality) to the environment by reducing discrepancies between one's emotional needs and norm-favoring social pressures.

Table 26–9.

Descriptors of Individuals Who Score High and Low on the Three Character Dimensions

	Descriptors of Extreme Variants	
Character Dimension	**High**	**Low**
Self-Directedness	Responsible	Blaming
	Purposeful	Aimless
	Self-accepting	Vain
	Resourceful	Passive
	Spontaneous	Conflicted
Cooperative	Tolerant	Intolerant
	Empathic	Insensitive
	Helpful	Selfish
	Compassionate	Revengeful
	Principled	Opportunistic
Self-Transcendent	Self-forgetful	Self-preoccupied
	Transpersonal	Self-centered
	Spiritual	Skeptical
	Contemplative	Objective
	Idealistic	Pragmatic

In Figure 26–1, character corresponds to the processes of logic, construction, and evaluation of abstract symbols that are based on conceptual representation of information and are well-developed only in some mature humans. These processes are related to executive functions (predominantly logic), legislative functions (predominantly construction), and judicial functions (predominantly evaluation). Early fMRI research in healthy human volunteers by Debra Gusnard and Marc Raichle demonstrated strong correlation between Self-Directedness and a cortical circuit involving the medial prefrontal cortex, which is known to regulate executive function (see Cloninger, 2004). Earlier work also showed that psychophysiological markers of neocortical processing, such as the P300 event-related potential and Contingent Negative Variation, are correlated with measures of character but not temperament. Namely, Self-Directedness, but not temperament traits, correlates moderately with the evoked potential P300 (r = 0.4, P < 0.002). Likewise, Cooperativeness correlates with the Contingent Negative Variation (this was particularly obvious for the Empathy subscale, r = 0.4). Individuals with Parkinson disorder differ in temperament from others (they are lower in Novelty Seeking premorbidly) but not in character. These early fMRI and ERP findings have been confirmed and extended by recent structural and functioning neuroimaging findings demonstrating that character involves higher cortical functions instantiated in specific neocortical networks of learning and memory (Table 26–10). The executive and legislative functions related to Self-Directedness and Cooperativeness involve conceptual processing in components of the semantic learning system,

which recognize and appraise the feelings and intentions of one's self and of others. In contrast, Self-Transcendence involves the functioning of the autobiographical memory system that allows self-aware (autobiographical) learning and memory functions with a holistic narrative perspective including mental time-travel, vivid recollection, and imagination of the past and future as a continuity, and identity formation as a creative process.

Table 26–10.
The Relationship between Character Traits, Higher Cognitive Functions, and Brain Structures Based on Voxel-Based Morphometry

Character Dimension	Higher Cognitive Function	Associated Brain Structures
Self-Directedness	Executive functioning Evaluation of internal cues Self-reflection Appraisal of the self and personal intentions	Executive brain network including Anterior cingulate cortex, Posterior cingulate cortex Medial prefrontal cortex Temporal cortex (for self-appraisal)
Cooperativeness	Legislative functioning Empathy Metacognition Appraisal of others' feelings and intentions	Legislative brain network including Medial frontal cortex (white matter) Precentral gyrus (white matter) Superior temporal gyrus (gray matter, BA22)
Self-Transcendence	Judicial functioning Autobiographical recollection Mental time-travel Holistic perspective shifting (time, place, person) Imagination of possible future events in vivid detail in place and time Transpersonal identification	Judicial brain network involves autobiographical memory and learning, including Middle temporal gyrus (greater GMV of BA 21) Inferior parietal gyrus (greater GMV of BA 40) Superior frontal gyrus in medial prefrontal cortex (smaller GMV of BA 10)

Gardini S, Cloninger CR, Venneri A. Individual differences in personality traits reflect structural variance in specific brain regions. *Brain Res Bull*. 2009;79(5):265–270; and Van Schuerbeek P, Baeken C, De Raedt R, De Mey J, Luypaert R. Individual differences in local gray and white matter volumes reflect differences in temperament and character: a voxel-based morphometry study in healthy young females. *Brain Res.* 2011;1371:32–42.

Consequently, it is useful to distinguish three major aspects of human personality: temperament, character, and identity. The characteristics of temperament, character, and identity are summarized in Table 26–11. A person's identity is cultivated and formed in autobiographical memory, which integrates a person's emotional drives (temperament) with their emerging goals and values (character). Identity is the narrative self in the context of a self-aware individual developing over time in a variety of situations. Identity, character, and temperament are continuously influencing one another as personality develops over the life span and are also responsive in more or less adaptive ways to changes in life situations

and events, including changing social roles with age, marriage, and work, as well as traumatic and therapeutic experiences.

Genomic Approaches to Understanding Temperament and Character. Personality development, including identity formation in particular, involves the adaptive development and integration of temperament and character in the changing situations and roles of a person's life. How do temperament and character develop? What are the mechanisms that allow adaptive learning across the life span in relation to temperament and character? These basic questions have been matters of speculation in the past, but recently, Cloninger, Svrakic, Zwir, and their associates have begun to uncover the hidden mechanisms by carrying out person-centered Genome-wide Association Studies (GWAS) of temperament and character to learn what genes are associated with temperament only, character only, and the interactions of both temperament and character together. The method of person-centered GWAS allows for gene–gene and gene–-environment interactions. Different profiles of temperament and of character were associated with different sets of functionally interacting genes detected through single-nucleotide polymorphisms. The broad heritability of the various traits of personality are all around 50 percent, but this depends on substantial interactions among multiple genes and environmental variables. The average effects of individual genes explains less than 1 percent of the variability in personality, but allowing for gene–gene interactions and gene–environmental interactions accounted well for the broad heritability of temperament and character. The whole genome was screened for specific genes using about one million single-nucleotide polymorphisms in more than 4,000 people representative of the general populations of Finland, Germany, and Korea. Personality was found to be highly polygenic with the associated genes distributed across all the human chromosomes. Figure 26–2 shows the number of genes per chromosome that are strongly associated with clusters of people with particular temperament profiles (i.e., temperament only), clusters of people with particular character profiles (i.e., character only), and clusters of people with profiles of both temperament and character traits (i.e., temperament and character). Sets of multiple genes were identified independent of information about the personality traits, and the clusters of people were identified without information about the genetic profiles. The gene sets, therefore, represent naturally occurring combinations of genes with distinct functional interactions. The clusters of personality traits represent attractor states in a complex adaptive system, allowing people to function in particular situation more or less well, as illustrated in the figures that follow next. The sets of genes and the personality profiles were very strongly associated, but the relationships were complex in the sense that the same gene set could be expressed in multiple ways (i.e., multifinality,

pleiotropy), and different sets of genes could be expressed in the same way (i.e., equifinality, heterogeneity).

Table 26–11.
Key Differences between Temperament, Character, and Identity

Variable Properties	Temperament	Character	Identity
Awareness level	Automatic	Intentional	Self-aware
Memory form	Percepts Procedures	Concepts Propositions	Intuitions Narratives
Learning principle	Associative conditioning	Logical/conceptual analysis	Automatic Intelligence
Role of subject in mental activity	Passive Reproductive	Active effort and logical construction	Effortless and fluid creativity
Key brain system	Limbic systems Corticostriatal habit system	Semantic memory (frontotemporal neocortex)	Autobiographic/autonoetic memory system
Form of mental representation	Stimulus–response sequences varying additively in strength	Logical hierarchies following from assumptions, prior goals, and values	Narrative schema and interactive networks varying qualitatively in configuration

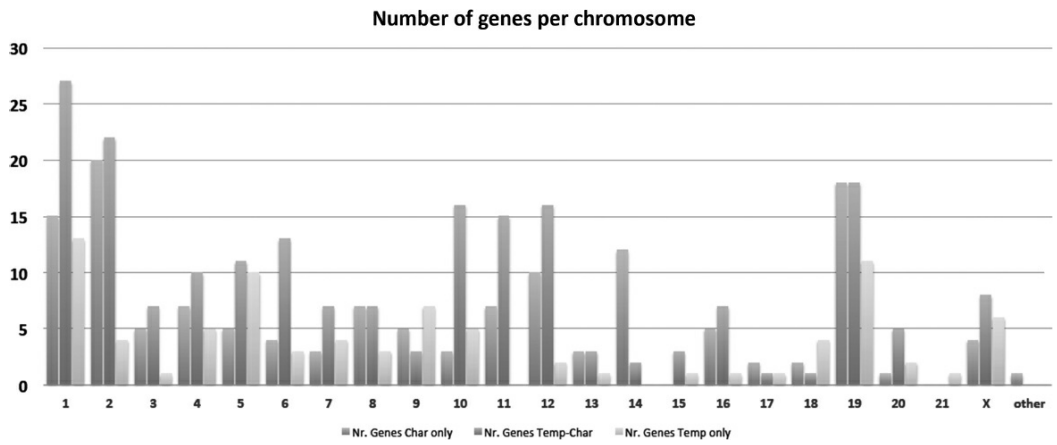

FIGURE 26–2. Number of genes per chromosome strongly associated with temperament profiles only (temp only), character profiles only (char only), and the interaction of both temperament and character (temp and char) in person-centered GWAS of the temperament and character inventory. (Used with permission of the Center for Psychobiology of Personality at Washington University in St. Louis.)

The complex development of personality profiles from interacting sets of multiple genes and environmental variables will be described in more detail later. Now it is fundamentally important to recognize that most of the genes that influence personality have regulatory functions, rather than being protein-coding genes (Fig. 26–3). In older genetic studies of personality, nearly all attention has been focused on protein-coding genes related to neurotransmitter function or neurotropin functions liked those mentioned for temperament traits in tables. However, it is the regulatory

genes that help one to understand change and development of human personality. Most of the regulatory genes are long non-coding RNAs (lncRNAs), other shorter non-coding RNAs, and pseudogenes. These non-coding RNAs (ncRNAs) are functional RNA molecules that are transcribed from DNA but are not translated into proteins. They can regulate gene expression at both the transcriptional and posttranscriptional levels. They are involved in epigenetic processes that can be divided into two main groups: the short ncRNAs of less than 30 nucleotides and the long ncRNAs of more than 200 nucleotides. The major classes of short ncRNAs are microRNAs (miRNAs), short interfering RNAs (siRNAs), and piwi-interacting RNAs (piRNAs). Both long and short RNAS are known to play a role in heterochromatin formation, histone modification, DNA methylation targeting, and gene silencing. However, lncRNAs and the shorter ncRNAs act differently.

Nearly 30 percent of the genes associated with character only involve lncRNAs whereas the percentage is only 20 percent in the other two groups, as shown in Figure 26–3. In contrast, protein-coding genes are more often associated with temperament only than with the groups involving character. The more prominent role of lncRNas in character is consistent with the greater importance of epigenetic change during adulthood for the development of character compared to temperament. This genomic finding supports findings from longitudinal studies that character shows much greater change during young adulthood than does temperament. LncRNAs form extensive networks of ribonucleoprotein (RNP) complexes with numerous chromatin regulators. These chromatin regulators target these enzymatic activities to the appropriate locations in the genome. Apparently, lncRNAs can function as modular scaffolds to specify higher-order organization in RNP complexes and in chromatin states. LncRNAs have several functions in the regulation of gene transcription, including the regulation of specific gene transcription (either activation or repression), posttranscriptional regulation, differential spicing and regulation of epigenetic modifications, including histone and DNA methylation, histone acetylation and sumoylation, and remodeling of chromatin domains. These findings suggest the hypothesis that character development involves epigenetic processes involving regulation of gene expression and chromatin modification. LncRNA dysfunction is also often observed in common chronic diseases related to stress and lifestyle regulation, including cancer, cardiovascular disease, neurological diseases, and immune-mediated diseases.

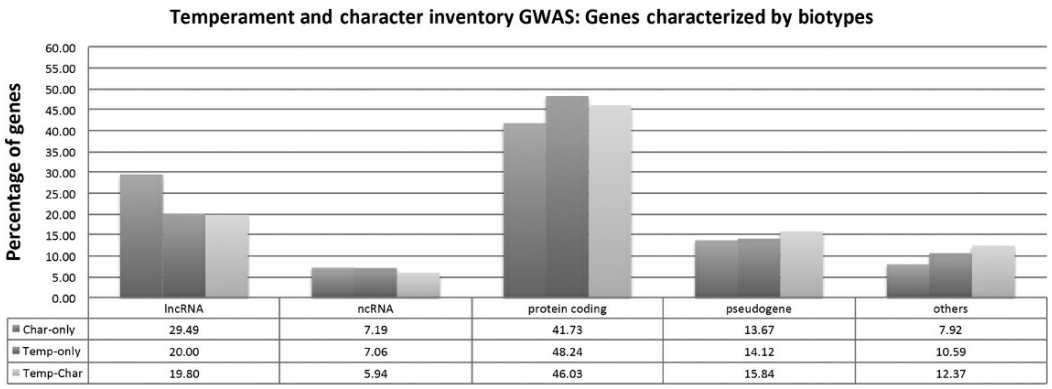

FIGURE 26–3. Percentage of different types of genes strongly associated with temperament profiles only (temp only), character profiles only (char only), and the interaction of both temperament and character (temp and char) in three independent samples (Finland, Germany, and Korea). (Used with permission of the Center for Psychobiology of Personality at Washington University in St. Louis.)

In contrast, the shorter ncRNAs are mostly involved in distinct temporal expression patterns during embryogenesis, acting in specific tissues to regulate expression at a posttranscriptional level. MicroRNAs usually bind to a specific target messenger RNA with a complementary sequence to induce cleavage, degradation, or to block translation, and they have crucial roles in epigenesis. Silent-interfering RNA function in ways similar to miRNAs to mediate posttranscriptional gene silencing as a result of messenger RNA degradation and to induce heterochromatin formation. Piwi-interacting RNAs are so named due to their interaction with the piwi family of proteins. They are involved in chromatin regulation and suppression of transposon activity during embryogenesis. The percentage of ncRNA genes associated with personality is only 5 percent and does not distinguish the temperament clusters from the character clusters.

As shown in Figure 26–3, the distribution of the types of genes associated with character differs from that of other groups. This adds support to findings from longitudinal studies of development and from functional brain imaging that temperament and character are distinct domains of personality with distinct functions in different brain regions. The prominence of regulatory functions for genes associated with personality provides a mechanism for understanding the way learning and memory processes influence personality development as a complex adaptive process. It is clear from the genomic results that personality involves dynamic self-organization that is more or less adaptive depending on both nature and nurture. The convergent findings from neurogenetics and psychosocial development provide a foundation for renewed appreciation of earlier clinical work on personality from diverse fields including psychodynamics.

Psychodynamic Concept of Character

The psychodynamic concept of character derives from the concept of defense mechanisms. The latter are defined as automatic, unconscious psychological processes (both cognitive and emotional), which protect against anxiety generated by intrapsychic conflicts and external stressors. All defense mechanisms are grouped into three levels, for example, mature (sublimation, anticipation, humor, altruism), neurotic (centered around suppression), and immature (centered around splitting). Psychodynamic concepts (Anna Freud, Otto Fenichel) define mature character traits as residues of previous mature defense mechanisms that have become stable behavior patterns. In other words, normal character reflects one's capacity to postpone immediate gratification and process internal needs through sublimation, anticipation, altruism, and humor. Behaviors reflecting these particular defense mechanisms correspond to the description of Self-Directedness (especially sublimation), Cooperativeness (especially altruism), and Self-Transcendence (especially anticipation and cheerful humor). Psychodynamic theories have also recognized that the Ego not only protects from, but shifts, organizes, and reacts to internal impulses and external stimuli. The pattern of these integrative functions and the ways in which one combines various functions to satisfy multiple external and internal pressures also constitute character. Again, Ego-strength corresponds to the description of Self-Directedness. Psychodynamic theories describe two forms of character disorders: neurotic character and character neurosis. Neurotic character traits are postulated to derive from neurotic defenses (e.g., suppression, reaction formation, projection, repression, undoing), which have dissociated from their original conflict and have become inflexible, pervasive, and ego-syntonic traits of everyday behavior. A. Freud called this inflexibility "the armor plating" of character. As extensions of previous conflicts between one's emotions and opposing social pressures, neurotic character traits protect a person from being involved in such situations again. For example, excessive scrupulousness, a frequently observed character trait of obsessive personalities, is a "reaction formation" against their usually strong aggressive impulses. By being rigidly hyperscrupulous, obsessive persons protect themselves from conscious experience of aggression and from intrapsychic conflicts that would have been triggered by such an experience.

Character neurosis is observed when an inflexible neurotic character trait interferes with healthy parts of personality. For example, character neurosis is present when excessive cleanness (an ego-syntonic character trait observed with obsessive individuals) frequently interferes with the need to interact freely with others (generated by healthy parts of obsessive personality); instead of being perceived as natural, cleanness becomes so excessive that it is perceived as frustrating. Of note, character neurosis has been included in the *International Classification of Diseases,* 10th revision, but not in DSM-IV.

Individual differences in the maturity of defense mechanisms (i.e., normal, neurotic, immature) tend to covary with the severity of psychopathology (i.e., normality, anxiety disorders, PD, respectively). Patients with PD, who in general have immature defenses, also tend to "favor" certain immature mechanisms, which, in turn, shape their behavior style and clinical presentation. These dominant defenses are, for example, projection for paranoid PD, acting out for antisocial PD, fantasy for schizoid PD, and so forth. Note, however, that in addition to these favorite defenses, PDs also use other immature defenses that contribute to their irresponsibility, such as their tendency to blame others and act opportunistically in a selfish manner. Reflecting these and similar observations about the importance of defense mechanisms for the overall understanding of deviant personality, the DSM-IV classification enabled clinicians to classify the use of a particular, dominant defense mechanism on Axis II, but the multiaxial system and description of defense mechanisms have been omitted in DSM-5 as psychiatrists operating in managed care settings have increasingly limited the practice to pharmacotherapy without a systematic approach to psychotherapy.

The psychodynamic and the psychobiological descriptions of character traits are similar. Both concepts underscore the adaptive function of character in the overall effort to satisfy one's emotional needs in a variety of situations and constraining influences. However, the two concepts differ in other etiological and classificatory aspects of character. The psychobiological model explains individual differences in character as quantitative variation on the three character traits. In contrast to this dimensional approach, the psychodynamic concept distinguishes categorically between character subtypes (e.g., neurotic character is further divided into anal, oral, phallic) and discrete character disorders (i.e., neurotic characters and character neurosis). The psychobiological model underlines the conscious nature of character, whereas the psychodynamic concept understands character traits, with the possible exception of anticipation, as predominantly unconscious processes. The psychobiological model is based on replicable neurobiological findings that were unavailable to Freud and his colleagues but were part of Freud's hope for the scientific future of personality assessment and treatment. Most importantly, the psychodynamic concept does not provide guidelines for further etiological and pathogenetic studies of character. Nevertheless, the psychodynamic concepts are important treatment tools and are used to reveal and revise conceptual biases inherent in immature defenses and deviant character traits.

Psyche and Identity

Psychology is literally the study of the psyche, but this word is frequently misunderstood. Psyche refers to a person's consciousness, self-awareness,

or spirit. More precisely in measurable scientific terms, individual variability in the psyche can be measured in terms of differences in a person's level of intuitive self-awareness, which is observable in terms of the scope of a person's identity. The word "identity" is derived from the Latin *identitas,* which is based on the Latin word *idem* for "same." The material contents of a person's body change over time, but a person's identity has a nonmaterial basis that can be the same throughout the narrative of a person's life and can even extend beyond the local reality of his or her self at a particular time and place. Intuitive self-awareness is unique to human beings, but humans do show variation in their level of self-awareness in ways that have a strong influence on susceptibility to PDs or the capacity for wisdom and well-being. In Figure 26–1, psyche involves the unique human process of symbolic invention, which is based on intuitive self-awareness, leading to the human invention of art, science, and spirituality. The growth of self-awareness is crucial to the development of full coherence of personality, which is manifest as creativity, well-being, and wisdom. Self-Transcendence is a measurable construct that is closely related to concepts of the psyche because the human capacity for self-transcendence and self-awareness depends on a distinct system of learning and memory called the autonoetic (i.e., self-aware) memory system. The scope of a person's identity can be quantified in terms of the extent to which a person views his or her self as a separate localized object or as a self-transcendent being through the process of transpersonal identification.

Psychobiology of the Psyche

Detailed studies of human learning and memory have distinguished procedural, semantic, and episodic memory. As described in relation to temperament, procedural memory underlies associative learning and involves presemantic perceptual processing of information from the physical senses that can operate independently of abstract conceptual and/or volitional processes. In contrast, intuitive learning is based on immediate recognition from a single episode of observation. One type of memory based on immediate recognition is knowledge of facts, which has been called "semantic memory" or "long-term declarative memory," as discussed in relation to the registration of cognitive schema underlying personality traits. The second type of recognition memory is awareness of one's own intentions in a spatiotemporal context, which has been called "episodic memory." In other words, episodic memory involves self-awareness and recall of events in a context that gives personal meaning to the "when" and "where" of life experiences. Initially, this type of recognition memory was defined in terms of materials and tasks, such as intentions (i.e., one's self in a spatiotemporal context). Subsequently, the concept of episodic memory has been refined and elaborated in terms of awareness of one's self, the sense of subjective time, and a conscious state

of self-awareness of the continuity in one's past, present, and future, which is different from dreaming or imagination. Tulving now describes episodic memory as self-awareness or "autonoetic consciousness," which is consistent with other descriptions of ordinary human self-awareness as based on immediate recognition by the intuitive senses. Episodic memory is also called "recollection" to distinguish it from factual knowledge that has no personal context in space and time. Tulving notes that there is no evidence that such intuitive self-awareness is present in any animal species except human beings.

Procedural, semantic, and episodic memory can be functionally dissociated from one another. For example, individuals with Parkinson disorder, characterized by striatal lesions, exhibit deficits in procedural learning but not in semantic or episodic memory. Individuals with an amnestic syndrome, characterized by lesions in the medial temporal lobe, have deficits in semantic learning of new facts but not in procedural learning or episodic consciousness. Individuals with amnestic syndrome have an anterograde amnesia for new facts but may have no deficit in their immediate sense of self-awareness. Bilateral lesions in the human hippocampus lead to deficits in self-awareness (i.e., episodic consciousness or the intuitive recognition of one's self in a spatiotemporal context). Episodic memory is subserved by a widely distributed network of cortical and subcortical brain regions that overlaps with, but also extends beyond, the networks subserving other memory systems. Key regions in this network include the hippocampus, regions of the prefrontal cortex (including the frontal poles and medial prefrontal cortex), and anterior cingulate cortex.

Essentially, the human episodic memory system acts like a mirror that provides spontaneous self-awareness to humans, whereas other primates require provision of an external mirror to even recognize themselves. This mirroring function depends on hemispheric specialization in the encoding and retrieval of information. Left prefrontal cortex is more involved in encoding episodic memories and retrieval of semantic memories, whereas right prefrontal cortex is more involved in the retrieval of episodic memories. The episodic retrieval mode is a state of consciousness (in which a person is actively recollecting or remembering his or her past) that involves the strong activation of sites in the right prefrontal cortex and the weak activation of sites in the left prefrontal cortex and medial anterior cingulate. The ability to recall past events in vivid detail is closely related to the ability to imagine future events in future detail. Hence, the human capacity for recollection and imagination is component of self-aware narrative learning. People who have been abused or traumatized are often unable to recall events in their life around the time of the trauma and are likely to be unable to imagine their future with sufficient detail to function adaptively in creating a satisfying way of living.

The experimental evidence that human beings have an intuitive form of learning and memory as the basis for their unique capacity for self-awareness is now strong but its significance has yet to be widely appreciated. Current evidence suggests that the hippocampus and other parts of the episodic memory system, such as prefrontal cortex, are necessary for encoding of episodic memories but not for the acquisition of factual knowledge. The hippocampus processes all sensory information available to a person about both one's external and one's internal milieu. Moreover, this pan-sensory information is constantly updated, allowing the registration and recall of ongoing life experiences in a personal context that is unique to each individual. The hippocampus is also known to maintain its plasticity throughout life, including the ongoing generation of new neurons as it adapts to ongoing experience. In fact, genes associated with impaired neurotrophin functions that promote healthy neuronal formation in the hippocampus (e.g., *NTRK3*) are associated with disturbances of self-awareness, as is seen in PDs and the schizophrenias. In contrast, long-term declarative memories, such as factual knowledge and self-reported character traits, are stored in brain networks that show greater stability and less plasticity than the hippocampus.

Likewise, character traits are static concepts, which are abstractions from which the spatiotemporal context has been reduced to a factual statement by a fixed choice that eliminates the intentional flexibility inherent in self-aware consciousness. Essentially, self-aware consciousness is "living" with many possible outcomes that can be voluntarily influenced, whereas factual knowledge is "dead" or "frozen" like a collapsed quantum wave function. These findings about self-aware consciousness suggest a possible explanation of the importance of factors unique to the individual in personality development and of the substantial resistance to change of character and related cognitive schema when cognitive-behavioral or psychodynamic strategies are used. The plasticity of the episodic memory system and its role in self-awareness also suggests a crucial role for growth in self-awareness as a means of experiential transformation of personality.

Development and Assessment of the Psyche and Identity

Episodic memory is a recently evolved memory system that appears to be unique to human beings, just as components of the episodic network, such as prefrontal cortex, are recent evolutionary developments. As expected for a recently evolved function, episodic memory develops later than procedural or semantic learning. Children younger than 4 years do not yet have a mature episodic memory system, so people generally have no direct recollection of their past before 4 years of age. In contrast, basic emotions are present at birth and self-object differentiation occurs between 18 months and 3 years of age.

Cloninger has observed five distinct levels of identity resulting from

varying degrees of intuitive awareness in human beings. There is a hierarchy of levels of awareness that can be measured in terms of what are sometimes called the innate human ideas or intuitions about being and human nature. Assessment requires an empathic understanding of the way another person is conscious of being, freedom of will, beauty, truth, and goodness, which provide markers of a hierarchy of levels of awareness. These five aspects of consciousness depend on one's intuitive awareness of the world. The awareness of facts and intentions are only two of the empirically observable levels of awareness.

For example, some patients with severe PDs complain of "emptiness," which is a fearful feeling of isolation, separateness, annihilation, or lack of being. Such patients lack a stable awareness of their being, so they are sometimes described as having a borderline level of personality organization. This is characterized by difficulty in sublimation, intolerance of distress, distortions of reality testing, and poor self-acceptance. Their unstable sense of being (i.e., lack of spiritual fullness, identity, and vitality) makes them so emotionally unstable and hopeless that they often mutilate themselves or attempt suicide. About 1 percent of people are deficient in their intuitive sense of being and experience life with a borderline organization chronically. Another 1 percent of people experience such disorganization and emptiness intermittently under stress. In contrast, the vast majority of people have a well-developed awareness of being all the time. This level of awareness involves memory of existence with a stable sense of being, that is, without fear of annihilation. This level of awareness is based on the recognition and memory of "facts," or what has traditionally been called "semantic memory."

Second, other people usually have a hopeful confidence in their being but feel that they are slaves to extrinsic influences and trapped in time like a hell or prison from which they have no means of escape. In other words, such people have a lack of flexibility in their actions, or little freedom of will. The lack of awareness of freedom of intention (i.e., will) also distorts the sense of a free flow in time. The lack of consciousness of free will in time has been described as "chronesthesia (mental time travel)" in patients who have brain lesions in their episodic memory system. Even in people with no brain lesions, the lack of awareness of free will can be pervasive. For example, some people are fatalistic materialists and claim that freedom of will is an illusion or something in which they are lacking. Their own intuition contradicts that of humanists and dualists who feel that rational responsibility and freedom of will are what give human beings their dignity. About 10 percent of people experience lack of freedom and the urgency of time commitments to external influences chronically throughout their lives. This corresponds closely to traditional definitions of PD, which involve lack of freedom or flexibility in adaptive behavior. In contrast, most people have a stable intuitive sense of their freedom and responsibility for their

intentions. More generally, this level of awareness involves a person's awareness of his or her intentions in a personal spatiotemporal context. This is the second of Tulving's types of intuition-based types of memories, but it does not exhaust the levels of human awareness.

Third, some people have no esthetic sensibility. That is, some human beings have no intuitive sense of what is lovely or beautiful. Cloninger has observed that those who had attained each level always had awareness of the lower levels but not the higher levels of consciousness. For example, those with an awareness of beauty also had awareness of being, freedom, but not necessarily an awareness of truth or goodness. In other words, these indicators of consciousness define a hierarchy of levels of awareness. It is shocking to recognize that some patients with PDs have no recollection at any time in their lives of feeling a sense of wonder and admiration for the beauty of any poem, painting, or nature scenes! Try to understand what that means for the quality of their lives. When such unaware individuals see a rainbow, they do not feel it is beautiful or wonderful! They just see physical colors and shapes with no conscious recognition of beauty. The complete absence of any sense of love and beauty is rarely acknowledged, occurring in only about 3 percent of individuals, but many more show little intensity in their appreciation of beauty. Although they know what opinions are fashionable, it has little or no impact on their ordinary awareness. Individuals with an unstable sense of beauty are consequently unstable in their capacity for love and emotional intimacy in relationships. Approximately half of marriages in the United States end in divorce, which is an objective indicator of instability in the sense of what is beautiful and loved. In contrast, the stable intuitive sense of what is beautiful and loved is characteristic of the other half of the general population, who have sometimes been designated as "romantics." In particular, esthetic sensitivity is a highly developed intuitive sense in romantic poets and novelists, such as Emerson and Thoreau, and in romantic musicians, such as Schubert.

Fourth, still other individuals have a sense of being, freedom, and beauty at least some times in their lives but never have any intuition of absolute truth. This is typical of individuals who describe themselves as pragmatists or objectivists. Such individuals also are usually agnostic or atheists because they have no intuition of their participation in a unity of consciousness, which gives rise to faith. About 10 percent of individuals in the United States are atheists or agnostics, which is consistent with recent nationwide Gallup polls. According to Gallup polls (2014), about nine of 10 (86 percent) Americans believe in God, but the majority do not feel that institutional religions are adequately inspiring and satisfying their spiritual needs. Eight percent of Americans do not believe in God but believe in a universal spirit or higher power. Only 5 percent of people are atheists who also deny belief in a universal spirit or higher power. However, it is

important to note that many people who describe themselves as atheists or agnostics actually have a well-developed spiritual sensitivity, as has been described by David Hay. In fact, some atheists may be more spiritual in their dedication to truth and tolerance than people who regularly attend religious services out of custom and convention.

Individuals who have attained each level usually have had awareness of the lower levels but not the higher levels of consciousness. For example, individuals with a sense of emptiness were also lacking in awareness of freedom, beauty, truth, and goodness. Those with an awareness of truth also had awareness of being, freedom, and beauty but not necessarily an awareness of goodness. Hence, awareness of truth is an intuitive sense that is well developed in elevated levels of consciousness. The intuitive sense of truth is prominent in many scientists, such as Teilhard de Chardin and Carl Rogers. It is even more fully developed in highly self-aware sages, such as Hegel, Leonardo da Vinci, and Gandhi, who viewed their lives as experiments in truth.

Fifth, still other individuals may have had a sense of being, freedom, beauty, and truth at times in their lives but had never had any intuition (i.e., direct awareness) of the intrinsic goodness of all things despite the perversions and corruption that exist in the world. Unfortunately, this group includes nearly everyone at times when they wonder about the existence and nature of evil or seek retribution and revenge against other human beings who threaten or attack us, or perhaps only frustrate us. Deep awareness of the universal unity of being and consciousness is required to love one's enemies. The intuitive sense of goodness is well developed only rarely. There is a partial awareness of it in individuals who are highly developed in all three character traits, including Self-Transcendence, such as the American transcendentalists such as Emerson and Thoreau. However, a well-developed awareness of goodness occurs only rarely, as in wise people such as Plato and Gandhi.

Finally, it is important for clinicians to realize that character development is modular and stepwise, and there can be blind spots or blockages in any particular module. As a result, there are infrequent exceptions to the rule that high levels of awareness imply awareness at all the lower levels. Such exceptions are particularly striking in high-functioning patients, such as alcoholic physicians and other professionals. Such patients may have mature character development overall but have specific blind spots in their insight and judgment that lead to serious impairment in particular aspects of their lives, such as romantic relationships or various forms of impulse control.

PERSONALITY: AN INTEGRATED VIEW

Recent reports about complex and fundamental processes underlying neurophysiology, development, and phenomenology of personality provide

powerful guidelines for the formulation of a comprehensive psychobiological model of personality as the coherent organization of the human body, mind, and spirit. The model integrates earlier phenotypic observations about behavior with contemporary concepts based on neuroimaging, biology of behavior, genetics, psychometric advances, and nonlinear modeling of normal and deviant personality development. Personality is conceptualized as a complex adaptive system involving a multidimensional interaction among temperament, character, and psyche. The coherent maturation of personality requires awareness of the natural unity of being, which, in turn, depends on the interplay of the procedural, semantic, and episodic systems of learning and memory (Fig. 26–4).

Through the interaction of temperament and character, different aspects of internalized concepts of the self and the external world modify the significance and the salience of sensory percepts and affects regulated by temperament and vice versa. In other words, temperament regulates what one notice and, in turn, character modifies its meaning, so that the salience and significance of all experience depend on both one's temperament and character. In turn, the development of character is derivative of individual differences in intuitive self-awareness. Essentially, personality development depends on the joint interactions among three systems of learning and memory. These are called the procedural, semantic, and episodic memory and are related to temperament, character, and identity (i.e., self-aware psyche), respectively. Reliable measures of the functioning of these three systems are provided by temperament (procedural learning of emotional drives), Self-Directedness and Cooperativeness (semantic learning of object relationships), and Self-Transcendence (self-aware learning in episodic or autobiographical memory). The least plasticity is in temperament (procedural memory) and the most plasticity is in the psyche (episodic memory), with character (long-term declarative or semantic memory) intermediate to the other two in its malleability.

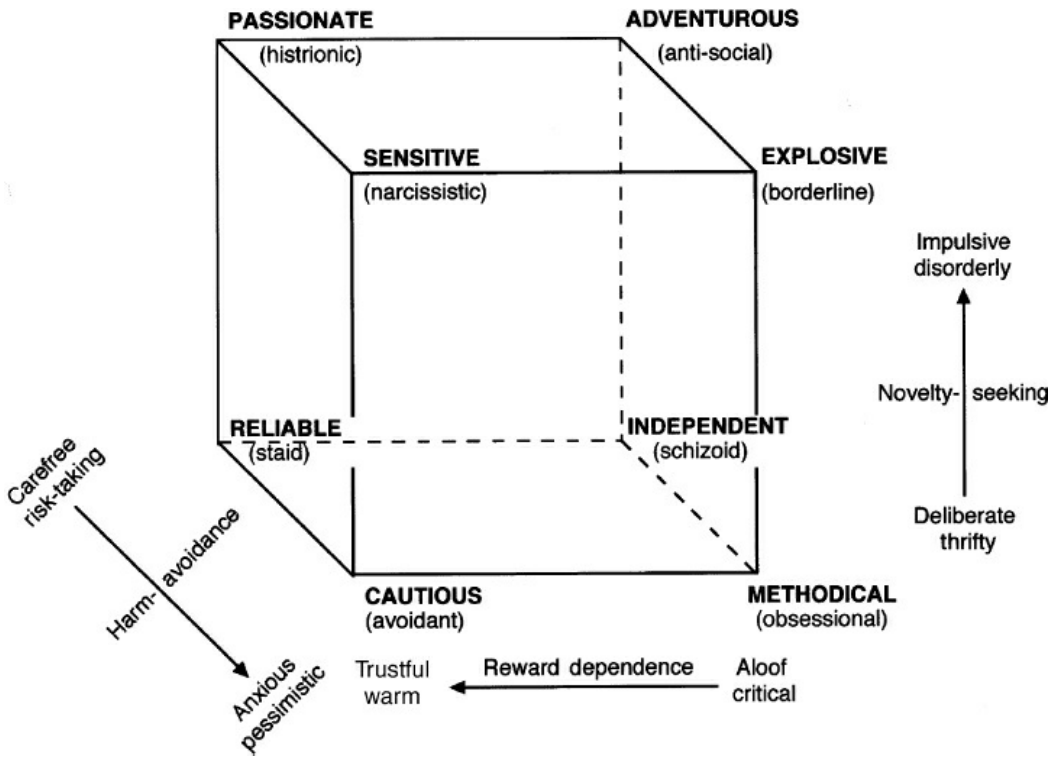

FIGURE 26–4. Temperament configurations.

Consequently, treatment efficacy depends substantially on whether treatment is focused on temperament, character, or identity (psyche). Somatic therapies, medications, and behavioral conditioning target temperament, which usually changes little beyond specifically conditioned situations. Cognitive and psychodynamic approaches seldom eliminate or transform maladaptive cognitive schemas (i.e., transform character) but may improve the management of conflicts by means of prolonged treatment of more than a year. Only treatment focused on intuitive awareness is likely to lead to actual cure of PD by transformation of the level of intuitive awareness, and the evidence in support of this suggestion by the authors remains anecdotal.

Any effective treatment has to allow the patient to be ready to see the world more coherently. Extreme temperament traits make PD patients inflexible, biased, and reactive. These “armor-like” features make it difficult for these patients to elevate their self-awareness so that they can see or feel or intuitively recognize world of beauty and truth that seems “hidden” to them. Everyone has become able to recognize how beautiful the truth is only after they were able to achieve a level of personal acceptance. Medication and cognitive therapies can help a person to be “ready” for higher levels of maturity, but the transformation always involves a change in perspective. Therapists must be flexible and alert to opportunities to help patients see through the fog of their fears and other negative emotions, which cloud their perspective. The key to change is often quite unique and unexpected. Accordingly, there is no recipe or algorithm by

which individuals always progress, and effective therapy requires ingenuity and flexibility because of the uniqueness of each person's strengths and weaknesses.

ESSENTIALS OF PERSONALITY ASSESSMENT IN CLINICAL PRACTICE

The assessment of personality can be made during routine assessment combining information from mental status and the lifetime psychiatric history. Integrating this information means that the clinician must recognize and understand the invariance of the organization of human thought regardless of timescale: mental status observations span milliseconds (e.g., a transient smile, fear, or anger) to minutes (e.g., perceptual awareness, fluidity of speech, resourcefulness, and speed in problem solving) whereas the lifetime history spans months (e.g., episodes of depression, gambling, substance abuse) to decades (e.g., repeated divorces, achievements). Human thought is dynamic with an average of 10 different thoughts per second, involving sudden shifts in attention, emotion, and intention, but it varies in ways that can also be observed to develop over the entire life span of the same individual. To be able to describe and understand this variability coherently, it is helpful for a clinician to have an integrated model of the structure of thought and personality.

Human personality has five layers or "planes of being" according to Cloninger. The five layers of personality are concerned with human adaptations in situations that are perceived to be predominantly concerned with reproduction and sexuality ("sexual plane"), practical everyday activities concerned with power and possessions ("material plane"), emotional bonds and social attachments ("emotional plane"), communication and culture ("intellectual plane"), and understanding what is beyond individual human existence ("spiritual plane"). The brain systems that make it possible for human beings to adopt these different perspectives have evolved in a hierarchy of discrete steps over time to allow the survival of progressively more flexible, aware, and creative organisms.

The development of well-being and the regulation of potential conflicts among emotional drives in self-aware consciousness involve three branches of mental self-government. These branches of self-government correspond to the higher cognitive processes related to the three character traits and should be systematically assessed during mental status examination. The assessment of insight and judgment is often trivialized to a subjective judgment about whether patients agree with their doctor ("insight") and do what they are told to do ("judgment"). It is more useful to evaluate insight as a combination of reasoning ability (self-directedness) and intuition (self-transcendence), which corresponds well to concepts of insight as rational intuition that date back at least to the writings of Plato. Judgment has come to mean something like the legislative function measured by

Cooperativeness, providing a capacity for empathy, teamwork, and principled action. However, the rating of these processes can be made much more informative and precise by recognizing how the three character traits are expressed in different situations. Each of these three cognitive functions has modules for adapting to each type of situation or layer of personality, which can be measured reliably by specific facets of character.

The three branches of mental self-government and their components are summarized in Table 26–12, along with the specific subscales of the Temperament and Character Inventory (TCI) designed to measure these functions. Self-Directedness in the TCI measures the executive functions, which can be considered the strength component of insight. It is expressed in various situations, including a person being self-directed in the sexual plane (i.e., resourceful), in the material plane (i.e., purposeful), in the emotional plane (i.e., self-accepting), in the intellectual plane (i.e., responsible), and in the spiritual plane (i.e., self-actualizing with few bad habits that interfere with achieving one's goals). On mental status examination, a person may be impartial and realistic in his or her thinking, high in self-esteem, and resourceful and far-sighted in actualizing his or her plans if he or she is highly self-directed, or he or she may be emotionally vulnerable, biased and unrealistic in his or her thinking, low in self-esteem, and short-sighted in planning. Remember that the rating of this one characteristic (executive functioning expressed as self-directedness) reveals more about the level of the person's maturity, his or her vulnerability to psychopathology, and his or her prognosis for recovery with a positive quality of life than any other feature of the mental status. Self-directedness is the expression of a life guided by a sense of hope that allows a realistic acceptance of reality as a path to understand the wonders and mysteries of life, even if they appear inconsistent with past expectations. Self-directedness is not a fixed trait—it can be temporarily reduced during stress, trauma, intoxication, or psychosis, and it can be enhanced by psychotherapy or life experiences such as opportunities to gain in self-directedness by one's accomplishments from responsible work. Everyone can grow in self-directedness under appropriate therapeutic conditions, such as reduced stress, nonviolence, and sobriety, especially when encouraged by the hope and compassion of others.

Table 26–12.
Descriptions of Three Branches of Mental Self-Government (Executive, Legislative, and Judicial Cognitive Functions) That Regulate the Emotional Drives Relevant to Each of the Five Layers of Human Personality, Which Are Defined by the Predominant Focus of the Person's Perspective in a Situation[a]

Cognitive Function (Character	Characteristics of the	Characteristics of the	Characteristics of the	Characteristics of the Intellectual	Characteri the Spiritu

Dimension)	Sexual Layer	Material Layer	Emotional Layer	Layer	Layer
Executive Function (Self-Directedness)	Resourceful vs. Inadequate (SD3)	Purposeful vs. Aimless (SD2)	Self-accepting vs. Vain (SD4)	Responsible vs. Blaming (SD1)	Self-Actual vs. Hope (SD5)
Legislative Function (Cooperativeness)	Tolerant vs. Prejudiced (CO1)	Helpful vs. Hostile (CO3)	Empathic vs. Inconsiderate (CO2)	Principled vs. Opportunistic (CO1)	Compassio vs. Reve (CO4)
Judicial Function (Self-Transcendence)	Idealistic vs. Pragmatic (ST5)	Self-forgetful vs. Self-Preoccupied (ST1)	Transpersonal Joy vs. Alienated (ST2)	Contemplative vs. Objective (ST4)	Spiritual vs Skeptica

[a]Within each layer of personality, maturation involves increasing each of the three character dimensions. Integration of the whole person requires working through these functions in each of the layers of personality. TCI subscales measuring these processes are indicated within parentheses.

Adapted from Cloninger CR. *Feeling Good: The Science of Well-Being.* New York: Oxford University Press; 2004:219–225.

Cooperativeness measures judgment as the legislative function, which involves making rules that allow people to get along in a reasonable and flexible manner with one another in each plane of their life. Cooperativeness can be recognized in the sexual plane as tolerance instead of prejudice, in the material plane as being helpful instead of hostile, in the emotional plane as empathy instead of self-centeredness, in the intellectual plane as being principled instead of opportunistic, and in the spiritual plane as a sense of compassion and fairness instead of vengefulness. In brief, good judgment is shown by a person's flexibility and ability to get along with others with values guided by compassion and principle. Poor judgment is shown by prejudice, lack of impulse control, insecure or hostile attachments, difficulty with teamwork, and opportunism. Comparison of a person's interaction during mental status examination with his or her life history is important to understand the range of cooperativeness he or she has demonstrated at different times during his or her life. For example, a person who is usually cooperative may temporarily become inflexible and hostile under stress or the aftermath of violent trauma. Alternatively, a person who has been uncooperative in the past may develop better judgment. A person may also have a blind spot in his or her awareness that elicits uncooperativeness in a recurrent but specific type of situation but allows good judgment in most other situations.

Self-Transcendence measures the depth component of insight. Depth of insight allows people to know intuitively when legislative rules apply in a particular situation. The depth of a person's insight is shown in the sexual plane by a person's being idealistic rather than selfishly pragmatic. Depth of insight is shown in the material plane by being self-forgetfully absorbed in what people are doing rather than self-preoccupied. In the emotional

plane, insight is shown by awareness of one's transpersonal connections with nature and other people, which leads to a sense of abundance, joyful exuberance, and being gracious. In the intellectual plane, depth of insight is shown by being contemplative rather than materialistically objective. In the spiritual plane, insight is shown by a sense of connectedness to what is divine or beyond human existence, expressed as faith or spiritual acceptance rather than skepticism. In general, poor depth of insight is shown by shallowness of awareness or magical ideation, whereas depth of insight is shown by wisdom, creativity, faith, and depth of understanding. In most people, the depth of insight is indicated by the frequency of peak experiences of transpersonal union, inseparable connectedness with nature, or oceanic feelings. However, reports of such peak experiences are sometimes the product of magical thinking in people who are very low in self-directedness. Hence, it is important to evaluate both the strength and depth components of insight (i.e., both self-directedness and self-transcendence).

All personality modules involve person by situation interactions that are regulated by a set of dynamical nonlinear systems, which allows human beings to be purposeful, flexible, and self-aware in their adaptation to life. *Human personality is not adequately characterized as a set of linear traits because the components of personality are nonlinear in their functional effects and relationships with one another*. Human thought involves frequent rapid adaptation of a person to different situations or different perspectives on the same experience. The connections between different brain modules change suddenly and synchronously with changes in thought.

The full three-dimensional structure of personality can be visualized as a three-dimensional spiral with five planes and increasing in height, width, and depth from the sexual to the spiritual plane. The height of the spiral indicates a person's strength of insight, as measured by his or her level of TCI Self-Directedness, which indicates a person's strength to face and adapt to reality rather than to blame, deny, or repress unpleasant facts. The width of the spiral indicates a person's broadness of social concern and flexibility of judgment, as measured by his or her level of TCI Cooperativeness. The depth of the spiral indicates a person's depth or shallowness of insight, as measured by his or her level of TCI Self-Transcendence. Essentially, a person's thoughts can shift rapidly throughout his or her possible range, spiraling down with increasing negative emotions and other indicators of ill health when stressed (e.g., preoccupied with fears, threats, or desires) or spiraling up with increasing positive emotions and other indicators of well-being when calm and secure (e.g., feeling hopeful, kind, and aware). A psychiatrist is often presented with a patient in a state of distress in which his or her thoughts have fallen to a low point in the range. The first step in treatment is to create

conditions that allow the patient's thoughts to ascend in the spiral of self-aware consciousness through validation, encouragement, medication, and mental exercises that allow a person to regain his or her more hopeful, kind, and coherent perspective toward life. Later work can help understand and reduce his or her vulnerabilities for relapse and recurrence.

Although human thought is highly dynamic with a wide range, one can use averages or modes of one's thoughts and relationships as central indicators of states of being that are metastable, as is typical of equilibrium states in nonlinear dynamical systems. Frequent metastable states of being are often described as personality traits, even though they are neither fixed nor linear in their functions and relationships. Accordingly, personality is a way of describing the dynamical processes that occur within the person to shape and adapt to life experiences by learning and planning. The maturation and integration of human personality involves one's growing in self-awareness through experiences across this wide range of situations.

As a result, personality is a complex expression of nonlinear interactions among a whole hierarchy of learning systems that develop over time as a complex adaptive process. Each personality module has three aspects, each regulated by dissociable but overlapping systems of learning and memory: a somatic or emotional aspect regulated by learning of habits and skills, a cognitive aspect regulated by learning of facts and propositions, and a spiritual aspect regulated by self-aware learning of one's continuity with and participation in a whole greater than oneself. These three systems can be dissociated but normally interact through their extensive interconnections. For example, a person can know how to do something (using procedural memory) but not know the fact that he knows the skill (a deficit in semantic memory). A person can know a particular fact and report it to others (using semantic memory) but not have memory of when and where he or she learned the fact (a deficit in episodic memory). The systems evolved and mature in sequence but operate simultaneously from early in childhood. As a result, the same symptom (e.g., feeling rejected) can be the result of different vulnerabilities (e.g., temperaments with high Reward Dependence, character with low Cooperativeness, poor intrapsychic integration leading to low well-being). Therefore, the effective clinician must look beyond presenting symptoms to understand and treat the underlying processes that produce symptoms, which is not possible with categorical diagnoses or multidimensional systems designed to compare and judge one person against another. Nevertheless, it is possible to use information about clinical syndromes to help identify the underlying processes that are most likely.

Table 26–13lists the personality traits that are most consistently associated with different clinical disorders and states of mental health. This is not meant to be an exhaustive listing of all the personality correlates of different mental disorders but only the traits that are so strongly related to

different disorders that a clinician can use the information in the differential diagnostic process. For example, high Harm Avoidance is commonly associated with most forms of psychopathology, but it has a particularly strong relationship as a cause of mood and anxiety disorders. Therefore, high Harm Avoidance should be suspected as a predisposing trait in mood and anxiety disorders. The presence of Harm Avoidance can be quickly evaluated by inquiring about child and adolescent traits such as being a worrier, pessimist, or shy. Likewise, high Novelty Seeking should be suspected if a person is alcoholic, drug abuser, gambler, or bulimic. Novelty Seeking can be quickly evaluated by inquiring about a person having a quick temper in his or her youth, being impulsive, extravagant, or disorganized, instead of slow-tempered and liking order and structure in what he or she does. On the contrary, mature character development confers well-being and resilience to psychopathology. In addition, positive character development facilitates recovery from mental illness, promoting the development of a sense of coherence, personal mastery, self-esteem, and life satisfaction. Thus, the distinction between temperament traits, which are value neutral, and character traits, whose development can facilitate recovery and well-being, provides a perspective that is encouraging and not pejorative.

Table 26–13.
Personality Traits[a] That Are Highly Likely to Be High (+) or Low (–) in Different Mental Disorders and States of Mental Health

Mental Disorder	HA	NS	RD	PS	SD	CO	ST
Schizophrenia	+		–		–	–	+
Mood disorder							
Depression	+				–		
Bipolar	+	+		–			+
Anxiety disorder							
GAD, panic	+						
OCD	+	–					
Social anxiety	+		+				
Eating disorder							
Anorexia	+			+			
Bulimea	+	+					
Personality disorder							
OCPD	+	–		+	–	–	
Cluster A			–		–	–	+
Cluster B		+			–	–	
Cluster C	+				–	–	
Mental order							
Recovered life quality					+		
Well Sibling of depressive					+		
Well Sibling of schizophrenic					+	+	

Resilience	–	+	+	+	+	+	+

[a]The personality traits are symbolized as HA (harm avoidance), NS (novelty seeking), RD (reward dependence), PS (persistence), SD (self-directedness), CO (cooperativeness), and ST (self-transcendence).

CONCEPTUAL AND PSYCHOMETRIC ISSUES

Personality Traits: Person versus Situation Debate

Personality traits are neuropsychic structures with the capacity to render many stimuli functionally equivalent and to initiate and guide equivalent (consistent) forms of adaptive and expressive behavior. Accordingly, the DSM defines personality traits as "enduring patterns of perceiving, relating to, and thinking about oneself and the environment," that is, other people and the world as a whole. The major value of personality traits, therefore, lies in their usefulness in identifying predictable regularities in an individual's behavior.

The stability of personality traits in time and across situations has been the central issue in personality theory for decades, as it bears direct relevance to the question whether internal dispositions or external situations determine behavior. The "person vs. situation" debate was initiated in the early 1900s by two groups of scientists ("situationists" vs "personologists"); the former viewed behavior as highly situationally specific, the latter as centrally organized and purposive. As noted by Epstein and O'Brien, the dilemma remained unresolved for decades because it was not realized that behavior can be situationally specific at the item level and cross-situationally general at the aggregate level. Single items of behavior have limited reliability and generality. They must be aggregated over situations and occasions in order to reveal broad and stable traits, which allow moderately accurate predictions of behavior tendencies (without having to specify the eliciting situations) but only weak prediction of behavioral acts in particular situations.

On the contrary, global personality traits also do not explain behavior completely because people do not exhibit unmodulated consistencies in behavior across time and situations (complete invariance in behavior is associated more with psychopathology than with normality). Biogenetic factors influence how an individual adapts to experience, and the latter modifies adaptive tendencies. Hence, personality traits are not fixed. This current understanding of traits differs from the conservative one of 20 years ago. The emerging consensus is an interactionist position allowing room for both situational (external) and dispositional (internal) determinants of behavior in a variety of complex combinations, as described in Tables 26–11 and 26–12 for temperament and character. Fortunately, it is now possible to specify facets of personality traits that are relevant to particular situations and also to aggregate these facets into

aggregate scales that are moderately predictive across a broad range of situations. In comparison of many popular personality models, the TCI predicts behavior and clinical disorders as well or better than other models.

Phenotypic and Developmental Trait Personality Models

The major difficulty in relating observed variation in personality to its underlying biological processes is that the observed phenotypic structure that distinguishes one person from another differs from the underlying psychobiological organization of adaptive processes that occur within the individual. Factor analysis describes only the differences between people, not the number or content of the processes that occur within individuals as they learn to adapt in time. Hence, the *phenotypic trait personality models,* which are usually factor analytically derived (e.g., Eysenck's tridimensional model, Costa and McCrae's Five-Factor Model) account for the much of the observed phenotypic variance but may not correspond to the underlying intrapsychic processes. For example, Neuroticism is a composite of both high Harm Avoidance and low Self-Directedness. However, Harm Avoidance and Self-Directedness have unique genetic determinants and are regulated by different brain circuits. For example, Harm Avoidance, and not Self-Directedness, is moderately correlated with Mismatch Negativity in evoked potentials, which is a measure of behavioral inhibition related to the orienting response, automatic attention, and early unconscious sensory processing. Furthermore, Self-Directedness, but not Harm Avoidance, is strongly correlated with activity with the late-evoked potential P300 and with activation of the medial prefrontal cortex in fMRI studies during executive tasks. This double dissociation of Harm Avoidance and Self-Directedness with early and late-occurring evoked potentials is consistent with the phylogenetic history of Harm Avoidance evolving early and Self-Directedness evolving late.

In contrast, *developmental trait personality models* take into account both the underlying biological dispositions to observable behaviors and individual differences in responses to experience during personality development. The most recognized dimensional models are those of Sjobring, Zuckerman, and Cloninger. Despite some limitations, Sjobring's developmental model in the late 1940s is a heuristic attempt to take personality research beyond the level of phenotypic observations and factor analysis. The model defines three dimensions of latent variation predisposing to personality: "validity" (i.e., the degree of energy available in adaptation to experience), "solidity" (the consistency of adaptation), and "stability" (referring to the maximum potential of a person to develop skillful habits in adapting to routine). These three genetic dispositions were postulated to be independent from one another and from the factor of general intelligence (referred to as "capacity") and also to define one's susceptibility to behavior disorders. Validity is similar in description to low

Harm Avoidance, Solidity to low Novelty Seeking, and Stability to low Reward Dependence.

Personality has often been described in terms of a linear sequence of qualitatively discrete developmental stages and structural types. However, the frequent occurrence of heterotypic individuals with intermediate or mixed features calls into question the basic assumption that personality structures are limited to a finite number of types that are qualitatively discrete and homogeneous. Likewise, it is becoming increasingly clear that there is not only variation in the rate and direction of personality development as a whole but also asynchrony and inconsistency among its major components that are reorganized through time. The following section outlines a more general nonlinear quantitative theory of personality development that does not make the restrictive assumptions about a finite number of qualitatively discrete structural types.

PERSONALITY DEVELOPMENT: FUNDAMENTALS OF A SELF-ORGANIZING PSYCHOBIOLOGICAL COMPLEX

Svrakic and colleagues formulated a quantitative model of normal and deviant personality development as a complex dynamic system that is self-organizing and partly molded by familial and sociocultural influences. The model allows for nonlinear interactions among etiologically distinct components of personality. It also accounts for the frequent, but not invariant, development of stage-like periods of moderately stable personality configurations (similar to personality types) punctuated by abrupt transitions in which there is structural reorganization and emergence of qualitatively new features. The model is based on a sophisticated mathematical framework that can be implemented and tested with readily accessible data. More recently, the developmental interactions among temperament, character, and identity in terms of interactions among procedural, semantic, and episodic systems of learning and memory have been able to be interpreted.

The central ideas behind this model are the following:

- Personality is a dynamical multidimensional system comprising more elementary operating components (traits) that are organized in an interdependent way critical to carrying out a particular function. Such a system is characterized by particular rules of operation, for example, the principles of associative learning within the temperament function, long-term semantic learning within the character function, and episodic intuitive awareness within the identity function. The satisfaction of such multiple constraints results in nonlinear dynamics, which is characteristic of all systems involving growth and development in biology, neuroscience, psychology, and sociology. Such multidimensional dynamical systems are usually called complex

adaptive systems.

- ► The correlations observed among multiple personality traits can be used to explain the spontaneous organization of stable multidimensional configurations (i.e., personality types); these types develop in a stage-like fashion with successive periods of prolonged stability punctuated by rapid transitions. However, the progression of each individual is unique because differences in intuitive awareness result in substantial variability that is unique to each individual.
- ► The model accounts for functional interactions among multiple types of influence, including genetic factors, family environment, sociocultural norms, and experiences unique to each individual.
- ► Based on the specified basic differences between temperament, character, and identity (Table 26–10), this model takes into account all aspects of information processing from associative conditioning of habits and skills to enculturation about goals and values within a unique psychological context.

Personality development is presented as a walk on an adaptive (or fitness) landscape with two or more hills (representing high adaptive values) separated by valleys (representing low adaptive values). Fitness is defined as the ability to produce change in personality. In general, a complex system subjected to constraints responds to these constraints by optimizing fitness, that is, by adaptive changes in personality. As change in personality is motivated by optimization of fitness, people will most likely and most rapidly move in the direction of the greatest increase in adaptive value (i.e., the nearest hill). Once they have reached the peak of the hill, they will stay there for a relatively long time because they would first have to decrease fitness (i.e., descend into the valley) in order to find a hill with a higher peak (i.e., better adaptation). Such adaptive development is called U-shaped.

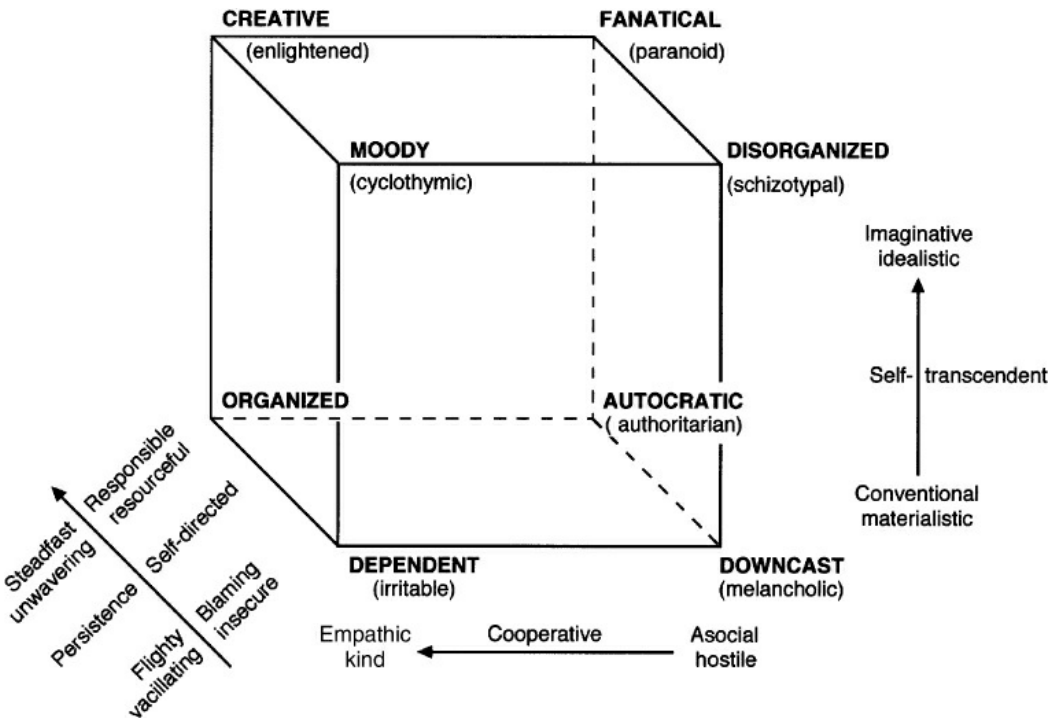

FIGURE 26–5. Character configurations.

As noted, temperament and character traits are etiologically distinct but functionally related. Such an interactive system has the property of being self-organizing as a result of the collective dynamics among its multiple components. In other words, the organism is spontaneously driven to find patterns of behavior that result in coherence of all information about both the external and the internal milieu. For example, Harm Avoidance interacts with both Novelty Seeking and Reward Dependence, inhibiting approach to novel stimuli and inhibiting social attachments by increasing fear of the unfamiliar and sensitivity to social criticism. Accordingly, consistent negative correlations of Harm Avoidance with Novelty Seeking and Reward Dependence have been found. In contrast, Novelty Seeking and Reward Dependence are positively synergistic, facilitating sociability and seeking of social approval. Such interactions influence the stability of each of the eight possible multidimensional configurations or "profiles" of temperament traits (Fig. 26–5), making some more stable than others.

In the fitness landscape, the valleys (or the adaptive minima) are unstable because even small perturbations, such as random events, will drive the system away from these points of the lowest fitness. In contrast, hills (i.e., adaptive maxima) are stable because they act as "attractors" for all lower points in the neighborhood. Consequently, personality interactions tend to self-organize into such an attractor state and to remain in that configuration in the absence of external pressure or maturational processes. For the understanding of both the chronicity and treatment resistance of PDs, it is of critical importance to realize that a particular personality configuration may be highly stable, even though it may not be

the most adaptive behavior possible for that individual (to be discussed).

We are mainly concerned with patterns of change of character traits as a function of relatively stable temperament traits and variation in the level of self-awareness, as discussed later. The three character traits interact to produce eight possible character configurations or profiles (Fig. 26–6).

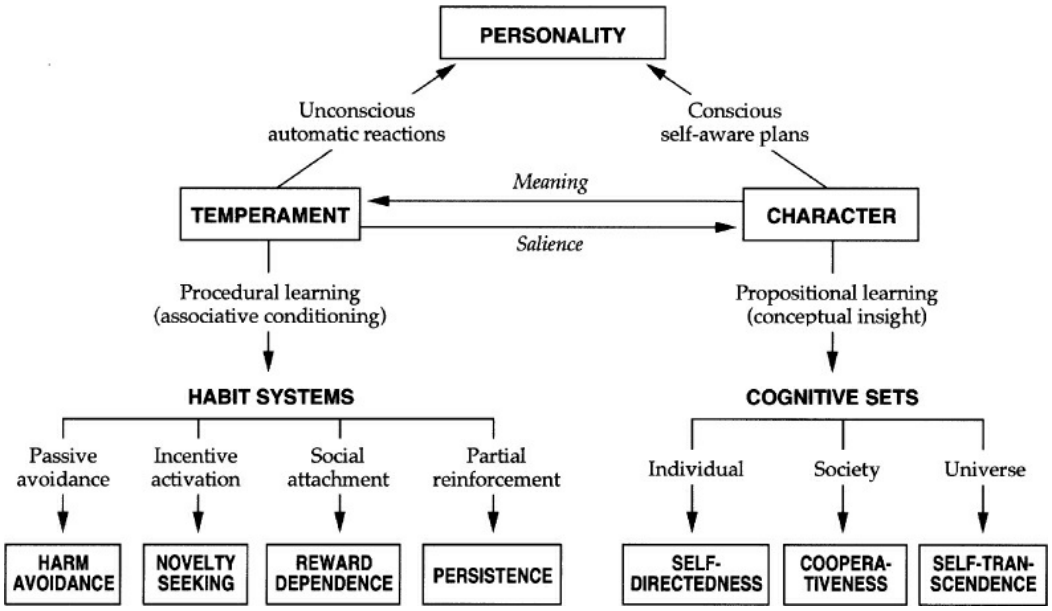

FIGURE 26–6. Personality: An integrated psychobiological schema.

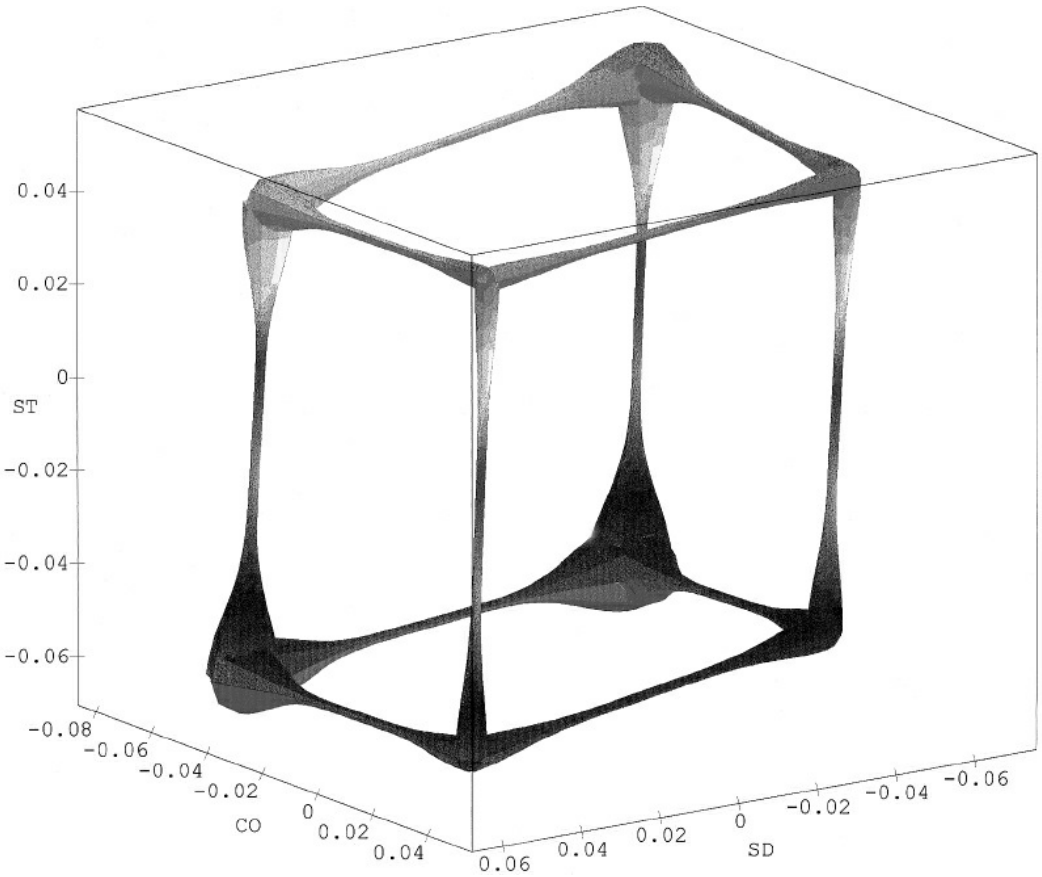

FIGURE 26–7. Most probable character outcomes for the adventurous antecedent temperament traits. CO, cooperativeness; SD, self-directedness; ST, self-transcendence.

Any change in these character configurations can be predicted from the initial temperament configuration taking into account temperament–character interactions and the effects of sociocultural norms, random events unique to the individual, social learning within the family, and genetic factors. However, whether there is change depends on self-awareness, which is unique to each individual and unpredictable, that is, underdetermined or free. All these sources of individual differences influence the likelihood for different stable character configurations in the outcome. The sociocultural norms can be additionally accounted for as the systematic bias in development associated with demographic variables, such as subculture, gender, race, age, and occupation. These norms and social education in the family create environments that are supportive or oppressive of specific character development. With all this taken into account, it turns out that a single initial temperament configuration may lead to several different stable character configurations; this aspect of development is referred to as multifinality. These results are graphically represented in Figures 26–7 through 26–9, with four different initial temperament configurations and their possible character outcomes.

Three dimensions of character (Self-Directedness—SD, Cooperativeness—CO, and Self-Transcendence—ST) are drawn along three orthogonal axes, and each character configuration is represented by a point in this abstract space. Eight possible high-low combinations of three character dimensions are represented by the corners of the cube. Observe that certain temperament types may and do give rise to several possible stable character outcomes, each with a certain probability, a feature characteristic of nonlinear systems. The size of the bulging at a particular character configuration is proportional to the probability of that outcome. Figures 26–5 through 26–7 illustrate this for adventurous (Hi NS, Lo HA, Lo RD), explosive (Hi NS, Hi HA, Lo RD), and sensitive (Hi NS, Hi HA, Hi RD) temperament types. Similar results can be obtained for the remaining five temperament configurations.

Individuals are born with one temperament profile with several possible character outcomes. This multifinality reflects the fact that the fitness landscape in personality development has both objective and subjective components. The sense of identity, feelings, and internalized concepts unique to the individual do influence the evaluation of the "objectively" worthwhile goals and values. The observed discrepancies between the natural directions influenced by antecedent temperament traits and actually achieved character configurations point to the importance of external events such as social learning and random fluctuations in the final character outcome.

As a result of the presence of points of high- and low-adaptive values in the fitness landscape, personality development is characterized by periods of relative stability alternating with more rapid transitions to new adaptive levels. The subscale structure of the TCI was formulated to specify character in terms of 15 component steps of its development, as shown in Table 26–14.

The 15 steps in character development are a theoretical ideal, corresponding to the modal pathway that leads to full character development with high scores on all three character dimensions, which is not optimal for everybody. This 15-step developmental sequence is consistent with prior qualitative descriptions of developmental stages by Piaget, Freud, and Erikson (Table 26–14) but allows for the actual nonlinearity in development that depends uniquely in each individual's identity. One's identity can extend beyond the present to influence the trajectory of the development along a life project that can be either creative or self-destructive. Observe that successive steps in character developmental form a spiral pattern, where each revolution around the spiral (presented as each of the five TCI tiers in Table 26–14) introduces a new set of developmental issues associated with the new set of component facets of the three major character traits. Furthermore, each revolution of the spiral depends on different levels in a hierarchy of intuitive awareness, as described earlier. For example, in early character development, during the first revolution around the spiral, a person encounters problems of trust versus mistrust, self-respect versus shame, and moderation versus indulgence; these are associated with Tolerance or Social Acceptance (facet 1 for Cooperativeness), Resourcefulness (facet 1 for Self-Directedness), and Idealism (facet 1 for Self-Transcendence). This depends substantially on a person's intuitive sense of being, whereas later developmental tasks depend on the intuitive senses of freedom and still later beauty, truth, and goodness. As can be inferred from Table 26–14, as they move along the spiral, people successively face new developmental tasks associated with new component facets in Self-Directedness, Cooperativeness, and Self-Transcendence. This spiral pattern of character development provides an opportunity to correct "developmental errors" that were made at any of the previous steps within the same trait. For example, as character develops around the spiral to reach a new step of Self-Directedness (e.g., responsibility), one gets "in line" historically with issues of self-acceptance, purposefulness, and resourcefulness (also steps of Self-Directedness but encountered during earlier revolutions) (Table 26–14). This alignment facilitates retrospective revisions of errors at these earlier developmental levels. The possibility of "retrospective healing" of old character errors is important in planning the psychotherapy of patients with PDs. Namely, when addressing problems related to a certain developmental step, these patients are more susceptible to changes related to previous steps within

the same character traits. In other words, learning about character in any one of the five planes of being can facilitate growth in self-awareness of similar processes in the other planes of being. This process of increasing maturation and integration of character continues throughout life because of the intrinsic human drive for coherence that is expressed as hope, love, and joy.

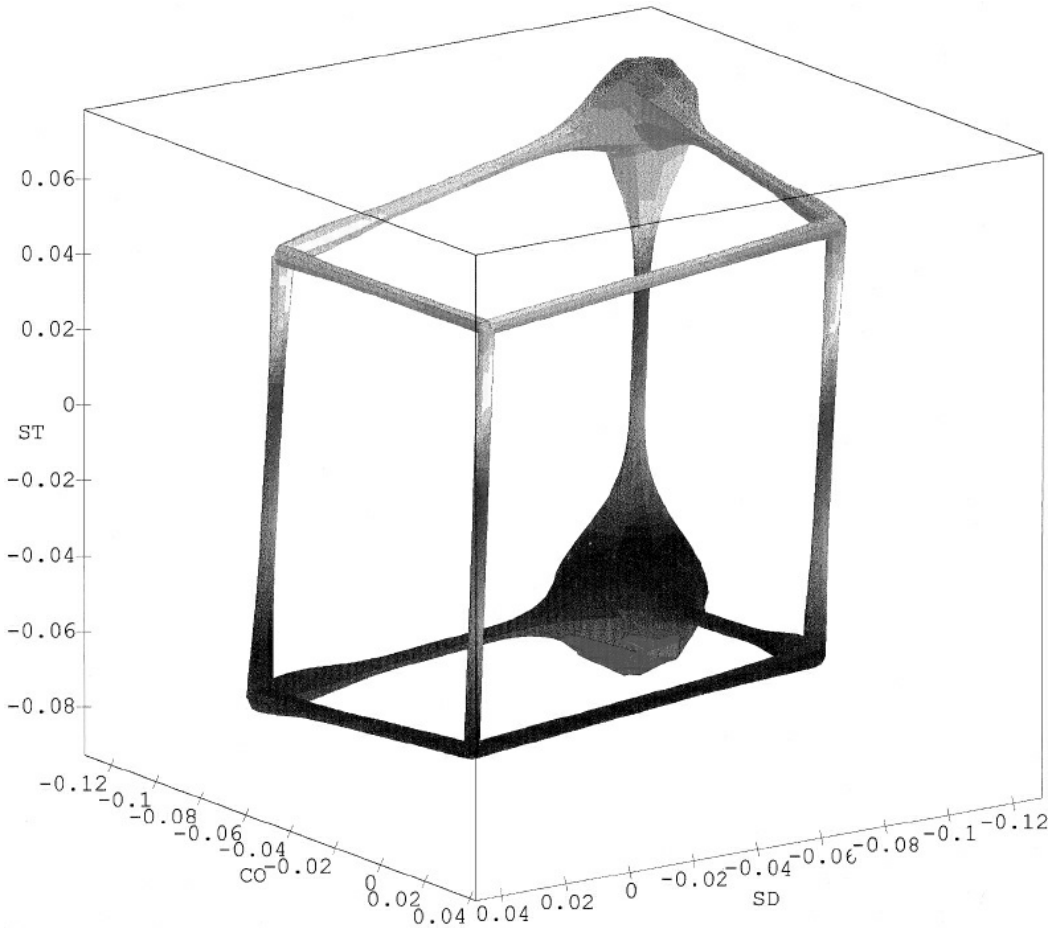

FIGURE 26–8. Most probable character outcomes for the explosive (borderline) antecedent temperament traits. CO, cooperativeness; SD, self-directedness; ST, self-transcendence.

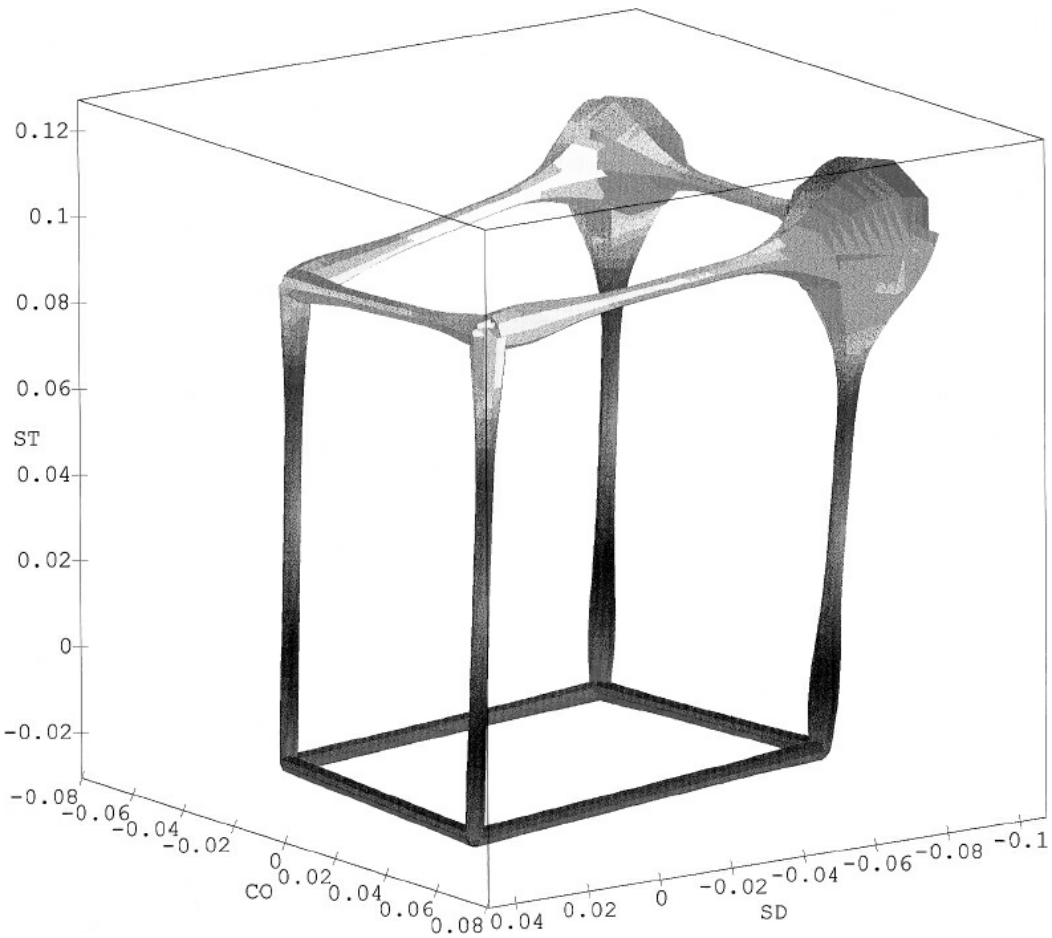

FIGURE 26–9. Most probable character outcomes for the sensitive antecedent temperament traits. CO, cooperativeness; SD, self-directedness; ST, self-transcendence.

Table 26–14. Comparison of Different Descriptions of Character Development and Its Spiral Path

Stage of Character Development	Stage of Piaget	Stage of Freud	Stage of Erikson	Legislative Function (Cooperative)	Judicial Function (Self-Transcendent)	Executive Function (Self-Directed)
0	Reflexive					
1	Enactive			Tolerance		
2		Oral	Trust		Trustful idealism	
3	Intuitive	Anal	Autonomy			Resourcefulness
4	Concrete operations	Phallic	Initiative	Helpfulness		
5	Abstract operations	Latency	Industry		Self-forgetful absorption	
6		Early genital				Purposefulness
7				Empathy		
8					Transpersonal identification	
9			Identity			Self-acceptance

10	Adult genital	Intimacy	Conscience		
11		Keeper of meaning		Contemplation	
12		Integrity			Responsibility
13	Sublimated genital		Compassion		
14				Spiritual acceptance	
15					Self-actualization

Computer simulations of the self-organized development of character have been carried out beginning with average initial temperament traits and taking into account all of the above contributing factors. These simulations found that children first increase in Self-Transcendence (i.e., become imaginative and enjoy fantasy), then increase in cooperativeness (i.e., become conforming and rule-based), and only later does self-directedness increase (i.e., behavior becomes increasingly self-reliant and autonomous). This prediction is consistent with the description of the stages of Ego development by Loevinger and her colleagues. This early sequence does not exclude the subsequent further development of other character dimensions in response to demands of changing social roles with age and/or changes in self-awareness. It is this subsequent development in response to external social pressures and ultimate existential questions that may explain adult self-actualization and moral development as described by Erikson, Kohlberg, and others. Consequently, intellectual components of the character traits are conscience, contemplation, and responsibility (Table 26–14). Likewise, the ultimate spiritual components of the character traits are compassion, spiritual acceptance, and self-actualization (also called enlightened second nature).

PERSONALITY DISORDER

Normal versus Deviant Personality

Normal personality is usually defined (i) directly, using criteria of health ideals, or (ii) indirectly, as the opposite to deviant personality, or most frequently (iii) statistically, by behaviors that are most common in the given environment. However, the distinction between “normal” and “abnormal” personality is inherently relative as it relies on arbitrary cutoff points on the continuum between two extremes (very low and very high) of any behavior. This distinction is also context dependent as the same behavior, manifested in different situations, could be viewed as normal or maladaptive (e.g., invariant cautiousness when danger is unlikely and the same trait when danger is likely). However, relying solely on the social or situational context to establish the diagnosis is problematic because PD

involves many noninterpersonal traits as well (e.g., narcissistic persons satisfy many aspects of their grandiosity in fantasy). Here, personal deviance alone also does not reliably distinguish between normal behavior and PD (e.g., some individuals are very socially withdrawn without impairment in professional functioning or signs of personal suffering and distress). In other words, both personal and social aspects are needed to fully account for the symptoms of PD.

Classification of Personality Disorders

The leading classifications of PDs are the International Classification of Diseases (ICD) of the World Health Organization, and the DSM of the American Psychiatric Association. The ICD classification of PDs, 9th edition, relied extensively on Kurt Schneider's book "Psychopathic Personalities" and his description of 10 discrete socially deviant personality types. The 10th revision of ICD (ICD-10), published in 1987, corresponds more closely to the DSM system, however. Anglo-American concepts of PD originate in Pritchard's description of moral insanity, which was later termed "sociopathy" and finally became antisocial PD in DSM-II in 1968. However, other classified PDs in the DSM system can be traced to the work of Kurt Schneider. This chapter describes the DSM-5 classification of PD, which is the same as DSM-IV, originally published in 1994. It is, however, descriptively similar to ICD-10 and, for practical purposes, much of the following section on DSM-5 holds for ICD-10 as well. A noteworthy exception to this is that ICD-10 classifies schizotypal PD on Axis I, among Schizophrenic disorders, whereas DSM-5 keeps this disorder on Axis II, among PDs.

It is also useful to be aware of proposal for the forthcoming 11th edition of the ICD (ICD-11), as well as alternative DSM-5 criteria that were rejected by the APA because of their unwieldy content and lack of scientific consensus. Both of these interesting criteria sets involve two major steps, as proposed by Cloninger (2000): first, rating the presence of PD and its severity in terms of levels of adaptive functioning, and second, describing the specific traits that are prominent in the person. In the ICD-11, personality is first rated as mild, moderate, or severe PD. Prominent personality difficulties can also be recognized by clinicians but are not considered to be disorders. In contrast, in the alternative DSM-5 criteria, the presence of PD is based on character traits such as Self-Directedness and Cooperativeness, which provide a description of personalities that are healthy or disordered. In the ICD-11 criteria, the specific traits are proposed to be like the four temperament dimensions and described by "the four As": asthenic/anxious (i.e., negative affective features such as high Harm Avoidance, as in the DSM anxious cluster C and Galen's melancholic temperament), adventurous/antisocial (i.e., dissocial and disinhibited features such as high Novelty Seeking, as in the DSM

impulsive cluster B and Galen's choleric temperament), asocial (i.e., detached features such as low Reward Dependence as in the DSM aloof cluster A and unlike Galen's sanguine temperament), and anankastic (i.e., perseverative and perfectionistic features such as high Persistence as in obsessional personalities and Galen's phlegmatic temperament). Thus, ICD-11 does not include a domain of psychoticism like the alternative model of DSM-5 that is designed as an abnormal form of the Five-Factor Model. Thus, the ICD-11 proposal is more like the temperament structure of Cloninger's psychobiological model than the Five-Factor Model, but some advocate of the Five-Factor Model push for separating dissocial features from disinhibition even though the features usually overlap. In any case, all these alternative models overlap extensively. It is difficult to choose between alternative descriptive models because factor analysis is a statistically indeterminate method. For a reasonable scientific choice among alternative models, it is essential to take into account data about the neurobiology and genetics of personality, as is done in the psychobiological model of personality but not the purely descriptive model of DSM-5.

These personality traits can also occur in combination with one another, so the produced categories are really not mutually exclusive, just as is also true of the corresponding DSM clusters. More fine-grained descriptions can be used as in DSM, but the international classification needs to be efficiently applied around the world in places with limited resources, so ICD-11 requires only ratings of the severity of PD and may not require the descriptors for subtypes. Even more lengthy criteria like the DSM-5 may be too unwieldy and burdensome. The alternative DSM-5 criteria also involve categorical ratings that are not systematically derived from psychobiologically validated dimensions of personality, so the American Psychiatric Association judged it premature to endorse them. The official DSM-5 criteria, which are the same as in DSM-IV, will be described in detail here because of a wealth of clinical experience and research that is available about them.

DSM-5 Classification

According to DSM-5, the critical criterion for distinguishing deviant personality traits is the presence (evidence) of long-term maladaptation and inflexibility that are manifested as either subjective distress or socio-occupational functional impairment. DSM-5 defines PDs as follows:

> (Personality disorders are) An enduring pattern of inner experience and behavior that deviates markedly from the expectations of the individual's culture. The pattern is manifested in two (or more) of the following areas:
>
> *Cognition* (i.e., ways of perceiving and interpreting self, other people, and events)
> *Affecti*vity (i.e., the range, intensity, lability, and appropriateness of emotional response)
> Interpersonal functioning; and
> Impulse control

The pattern is stable and of long duration and its onset can be traced back at least to adolescence or early adulthood. It is inflexible and pervasive across a broad range of personal and social situations and leads to clinically significant distress or impairment in social, occupational, or other important areas of functioning.

It is unfortunate that these general criteria are rather vague because the most important assessment about personality is the level of maturity of character functions. The most commonly diagnosed PD is "not otherwise specified." As a result, Cloninger has proposed that clinicians first evaluate the dimensions of character, particularly Self-Directedness and Cooperativeness, to determine whether there is any PD or not. As a second step, the features that distinguish subtypes can be described, but this is of secondary importance because the main focus of treatment is character development in every case.

Personality disorder subtypes classified in DSM-5 are Schizotypal, Schizoid, Paranoid (Cluster A), Narcissistic, Borderline, Antisocial, Histrionic (Cluster B), and Obsessive–Compulsive, Dependent, and Avoidant (Cluster C). In addition to these 10 "standard" disorders, the DSM-5 recognizes "personality change due to another medical condition," "other specified personality disorder" such as people with mixed features of more than one personality category, and "unspecified personality disorder" for presentations in which the clinician chooses not to be specific or has insufficient information to be specific. Two other disorders, Passive-aggressive and Depressive, that were recognized for further study in DSM-IV are omitted in DSM-5 but can be diagnosed as: "other specified PD." Likewise, ICD-11 proposes to drop the popular diagnosis of Borderline PD. The behaviors associated with borderline, passive–aggressive, and depressive adaptations are regarded by many experts as nonspecific expressions of severe character dysfunction.

The DSM-5 arranges categorical PDs into three clusters, each sharing some clinical features (see Table 26–15): Cluster A includes three disorders with odd, aloof features, such as Paranoid, Schizoid, and Schizotypal. Cluster B includes four disorders with dramatic, impulsive, and erratic features, such as Borderline, Antisocial, Narcissistic, and Histrionic. Cluster C includes three disorders sharing anxious and fearful features, such as Avoidant, Dependent, and Obsessive–Compulsive. The ICD-11 may further separate the obsessive compulsive cases from the avoidant and dependent cases. Several studies have supported the construct validity of these clusters except that the symptoms for compulsive disorder tend to form a separate, fourth cluster, as proposed for ICD-11 and indicated in Table 26–15. Note that the three dimensions underlying Clusters A, B, and C (i.e., detachment, impulsivity, and fearfulness) correspond closely to normal temperament traits (i.e., low Reward dependence, high Novelty Seeking, and high Harm Avoidance, respectively), suggesting that variation in these temperament traits might be significant in distinguishing among

the three clusters of disorders.

Table 26–15.
Qualitative Clusters and Subtypes of Personality Disorders According to the American Psychiatric Association (DSM-IV, 1994)

Cluster	Subtype	Discriminating Features
Odd/eccentric		
	Schizoid	Socially indifferent
	Paranoid	Suspicious
	Schizotypal	Eccentric
Erratic/impulsive		
	Antisocial	Disagreeable
	Borderline	Unstable
	Histrionic	Attention seeking
	Narcissistic	Self-centered
Anxious/fearful		
	Avoidant	Inhibited
	Dependent	Submissive
	Obsessive	Perfectionistic
Other specified		
	Passive–aggressive	Negativistic
	Depressive	Pessimistic

INDIVIDUAL PERSONALITY DISORDERS

General Diagnostic Guidelines

DSM-5 introduces the following general guidelines for the diagnosis of PD:

Many of the features of various PDs may be also seen during an episode of another mental disorder. The diagnosis of PD is made only when the features are typical of long-term functioning and are not limited to a discrete episode of another mental disorder. Likewise, when maladaptive behavior features are due to the direct psychological effects of another mental disorder including substance use (e.g., various psychoactive substances including medication), diagnosis of PD is not warranted. However, it is possible to diagnose personality change due to another medical condition such as temporal lobe epilepsy; in this case, the medical condition must be specified along with the associated diagnosis of personality change.

Judgments about personality functioning must take into account one's ethnic, social, and cultural background. Personality disorder should not be confused with acculturation following immigration or with the expression of customs and religious and political values characteristic of one's culture of origin.

Diagnosis of specific PD may be made in children or adolescents when observed maladaptive personality traits are pervasive, persistent, and unlikely to be limited to a particular developmental stage or an episode of an Axis I disorder. To diagnose a PD in an individual younger than 18 years, the features must be present for more than 1 year. The only exception to this is Antisocial Personality Disorder, which cannot be diagnosed in individuals younger than 18 years.

Clinical experience points to a potential sex bias in diagnosing PDs. Certain PDs are diagnosed more frequently in men (e.g., Antisocial and Schizoid), whereas some disorders are more frequently diagnosed in women (e.g., Borderline, Histrionic, and Dependent). Although it is likely that there exist real gender differences in the prevalence of these disorders, clinicians are cautioned not to diagnose certain PDs in males and females because of social stereotypes about typical gender roles and behaviors.

Observed maladaptive behavior traits must be pervasive, that is, manifest in a wide range of personal and social contexts (i.e., at home, at work, with family and friends), not isolated aspects of the person's life. The collection of data from collateral informants is thus considered critical to ensuring the high-quality personality assessment, diagnostic reliability, and validity.

In addition to "official" diagnostic criteria, DSM-5 specifies a group of "associated features." These features are not given in the form of explicit criteria (as their "official" counterparts) but rather as loose descriptions of

optional, but fairly frequent, behaviors intended to help clinicians in cases in which the diagnosis is not certain.

Cluster A Personality Disorders

Paranoid Personality Disorder

CLINICAL CRITERIA. The hallmarks of Paranoid PD are excessive suspiciousness and distrust of others expressed as a pervasive tendency to interpret actions of others as deliberately demeaning, malevolent, threatening, exploiting, or deceiving. Diagnostic features also include at least four of the following:

- Hypersensitivity to and unforgiveness of insults, slights, and rebuffs
- Unwarranted tendency to question the loyalty of friends, or fidelity of spouse or sexual partners
- Reluctance to confide in others because of unwarranted fear that the information will be used against them
- Reads hidden demeaning or threatening meanings into benign remarks or events
- Preoccupation with "conspirational" explanations of and hidden demeaning or threatening meanings into benign events or remarks
- Persistently bears grudges (i.e., is unforgiving)
- Unwarranted tendency to perceive attacks on their character or reputation with angry reactions or counterattacks

Some of the associated features include the following:

- Excessive need to be self-sufficient and strong sense of autonomy
- Pervasive inability to relax and compromise
- Frequent involvement in legal disputes
- Frequently impress others as "fanatics"
- Tendency to form closed groups or "cults" consisting of people with similar beliefs
- Peculiar ability to generate fear in others

COMPLICATIONS. Brief reactive psychosis, particularly in response to stress.

COMORBIDITY. These patients are at increased risk for Major Depression, Obsessive-Compulsive Disorder, Agoraphobia, and Substance Abuse or Dependence. The most common co-occurring PDs are Schizotypal, Schizoid, Narcissistic, Avoidant, and Borderline.

Paranoid PD has been postulated to be a premorbid antecedent of Delusional Disorder, paranoid type.

IMPAIRMENT. Frequently only mild and typically includes occupational

and social difficulties.

SEX RATIO. According to DSM-5, this disorder is more commonly diagnosed in males.

EPIDEMIOLOGY. Prevalence rates of 0.5 to 4.4 percent in the general population, 10 to 30 percent for psychiatric inpatients, and 2 to 10 percent for psychiatric outpatients are reported in DSM-IV-TR and DSM-5.

FAMILIAL PATTERN AND GENETICS. Some studies have demonstrated increased prevalence of this PD among relatives of probands with chronic Schizophrenia and Delusional Disorder, paranoid type.

DIFFERENTIAL DIAGNOSIS. Paranoid PD is distinguished from Schizophrenia (especially paranoid type), Delusional Disorder, paranoid type, and Affective Disorder with Psychotic features based on periods with positive psychotic symptoms, such as delusions and hallucinations, in the latter. When a brief reactive psychosis with delusions complicates clinical picture of Paranoid PD, this distinction is far more difficult. The duration of the latter and its frequent association with stress is usually sufficient for differential diagnosis.

Paranoid PD is sometimes difficult to distinguish from the following PDs:

- Schizotypal (which includes magical thinking, unusual perceptual experiences, oddities in speech, appearance, and thought processes)
- Obsessive–Compulsive, Schizoid, Borderline, and Histrionic (all with no prominent paranoid ideation)
- Avoidant (which includes fear of embarrassment)
- Antisocial (which includes personal gains in antisocial behavior)
- Narcissistic (which includes fear of having their "hidden" imperfections and flaws revealed)

Schizoid Personality Disorder

CLINICAL CRITERIA. The hallmarks of schizoid PD are pervasive pattern of social detachment and a restricted range of expressed emotions in interpersonal settings beginning by early adulthood and present in a variety of contexts, as indicated by four or more of the following:

- Indifference to praise and criticism
- Preference for solitary activities and fantasy ("loners")
- Lack of interest for sexual interactions
- Lack of desire or pleasure in close relationships
- Emotional coldness, detachment, or flattened affectivity
- No close friends or confidants other than family members

Pleasure experienced in few, if any, activities

Some of the associated features include the following:

Difficulty in expressing anger, even in response to direct provocation, which contributes to the impression of flattened affect
Passivity in adverse circumstances
Severe lack of social skills

COMPLICATIONS. Very brief reactive psychosis, particularly in response to stress.

COMORBIDITY. This PD sometimes appears as the premorbid antecedent of Delusional Disorder, Schizophrenia, or rarely Major Depression. The most common co-occurring PDs are Paranoid, Schizotypal, and Avoidant.

IMPAIRMENT. Frequently severe problems in social relations; occupational problems develop when interpersonal involvement required, solitary work sometimes favorably affects overall performance.

SEX RATIO. According to DSM-5, this disorder is more commonly diagnosed in males and may cause more impairment in them.

EPIDEMIOLOGY. Prevalence rates varying of 3.1 to 4.9 percent have been reported (DSM-5).

FAMILIAL PATTERN AND GENETICS. An increased prevalence among the relatives of probands with Schizophrenia or Schizotypal PD has been reported.

DIFFERENTIAL DIAGNOSIS. Schizoid PD is distinguished from Schizophrenia, Delusional Disorder, and Affective Disorder with Psychotic features based on periods with positive psychotic symptoms, such as delusions and hallucinations, in the latter. When a brief reactive psychosis complicates clinical picture of Schizoid PD, this distinction is far more difficult. The duration of the latter and its frequent association with stress is usually sufficient for differential diagnosis.

Schizoid PD is distinguished from Autistic Disorder and Asperger's Disorder by more severely impaired social interactions and stereotypic behaviors and interests in the latter two disorders. Schizoid disorder is distinguished from the following PDs:

Schizotypal (which includes magical thinking, unusual perceptual experiences, oddities in speech, appearance, and thought processes)
Avoidant (adequate emotionality in the latter, also social isolation due to the fear of embarrassment, not to indifference)
Obsessive-Compulsive (adequate capacity for intimacy, despite

sometimes excessive isolation due to perfectionism and workaholic attitudes)
- Paranoid (which includes suspiciousness, ideas of reference, and guarded facade)

Schizotypal Personality Disorder

CLINICAL CRITERIA. Schizotypal PD is characterized by social and interpersonal deficits as indicated by pervasive discomfort with reduced capacity for close relationships as well as cognitive and perceptual distortions and eccentric behavior (not severe enough to meet criteria for Schizophrenia). Diagnostic features also include at least five of the following:

- Ideas of reference (not delusions)
- Odd beliefs and magical thinking (superstitiousness, beliefs in clairvoyance, telepathy, "sixth sense"; in children and adolescents, bizarre fantasies or preoccupation)
- Unusual perceptual disturbances (illusions including bodily illusions, sensing a presence of a person nearby)
- Paranoid ideation and suspiciousness
- Odd, eccentric, peculiar behavior
- Lack of close friends, except family members
- Odd thinking and speech without coherence (e.g., vague, metaphorical, overelaborated, stereotypic)
- Inappropriate or constricted affect
- Social anxiety that does not diminish with familiarity and is associated with paranoid fears

PREDISPOSING FACTORS. Empirical findings about this disorder are difficult to interpret as there appears to be several clinical subtypes of this disorder, with potentially differential relationship to the schizophrenic spectrum. Adoption, family, and twin studies demonstrate an increased prevalence of schizotypal features in the families of schizophrenic patients, especially when schizotypal features were not associated with comorbid affective symptoms. Small anticipatory saccades that disrupt Smooth Pursuit Eye Movement, hypothesized to be a marker of genetic vulnerability to schizophrenia, have been also detected in schizotypal and introverted personalities. The nature of the relationship of this PD and Schizophrenia is still controversial, with some genetic studies not finding the relationship and some studies finding a pattern of gradation in multifactorial liability from schizotypal personality (mild) to broad schizophrenia (moderate) to narrow schizophrenia (severe).

COMPLICATIONS. Transient psychotic episodes, particularly in response to

stress. Symptoms sometimes become so significant that meet criteria for Schizophreniform Disorder, Delusional Disorder, and Brief Psychotic Disorder.

COMORBIDITY. More than a half of these patients have had at least one episode of Major Depression, and 30 to 50 percent have Major Depression concurrent with this PD. The most common co-occurring PDs are Schizoid, Paranoid, Avoidant, and Borderline.

IMPAIRMENT. Typically includes occupational and social difficulties.

SEX RATIO. Unknown; frequently diagnosed in females with fragile X syndrome.

EPIDEMIOLOGY. Prevalence rates of 3.9 to 4.6 percent in the general population are reported in DSM-5. Earlier reports suggested a range between 1 and 6 percent. It may be less common in clinical populations than in the general population according to DSM-5.

FAMILIAL PATTERN AND GENETICS. An increased prevalence of this PD among the first-degree relatives of probands with Schizophrenia. Also, an increased prevalence of Schizophrenia and other psychoses in the relatives of probands with Schizotypal PD. The disorder itself tends to aggregate in families (DSM-5).

DIFFERENTIAL DIAGNOSIS. Schizotypal PD is distinguished from Schizophrenia, Delusional Disorder, and Affective Disorder with Psychotic features based on periods with positive psychotic symptoms, such as delusions and hallucinations, in the latter. When a brief reactive psychosis with delusions complicates the clinical picture of Schizotypal Personality, this distinction is far more difficult.

Schizotypal PD is difficult to distinguish from the heterogeneous group of solitary, odd children whose behavior is characterized by social isolation, eccentricity, and peculiarities in language seen in Autistic Disorder, Asperger's Disorder, and Expressive and Mixed Receptive–Expressive Language Disorder. Communication disorders might be distinguished by the primacy and the severity of the disorder in language accompanied by compensatory efforts of the child to communicate by other means and also by specialized language assessment. Autistic Disorder and Asperger's Disorder are distinguished on the basis of more severely impaired social interactions and stereotype behaviors and interests in the latter two disorders. Schizotypal disorder is distinguished from the following PDs:

Schizoid and Paranoid (which rarely have magical thinking, unusual perceptual experiences, or oddities in speech, appearance, and thought processes)

Narcissistic (with predominant sense of grandiosity, fragile self-esteem, and fear of having their "hidden" imperfections or flaws revealed)
Avoidant (which rarely has oddities in appearance and behavior, and here fear of embarrassment, not disinterest and detachment, causes social avoidance and isolation)
Borderline (characterized by affective instability and stormy relationships as well as impulsive and manipulative behavior)

Cluster B Personality Disorders

Antisocial Personality Disorder

CLINICAL CRITERIA. The hallmarks of antisocial PD are pervasive disregard for and violation of rights of others occurring since the age of 15 years and continuing into adulthood. A person has to be 18 years of age or older, and there has to be evidence of conduct disorder before the age of 15 years (conduct disorder involves a repetitive and persistent pattern of behavior in which the basic rights of others or major age-appropriate social rules are violated; the examples include aggression to people and/or animals, destruction of property, deceitfulness or theft, serious violation of rules). Diagnostic features also include at least three of the following:

Failure to conform to social norms (resulting in frequent arrests)
Deceitfulness, including lying and conning others for personal profit or pleasure
Impulsivity or failure to plan ahead
Irritability and aggressiveness, including repeated physical fights or assaults
Recklessness, with disregard for safety of self and others
Irresponsibility, indicated by the failure to honor financial obligations or to sustain consistent work behavior
Lack of remorse, indicated by indifference or rationalizing having hurt, mistreated, or stolen from others

Some of the associated features include the following:

Promiscuity and inability to sustain a monogamous relationship
Lack of empathy, cynicism, contempt for feelings, rights, or suffering of others
Inflated and arrogant self-appraisal
Abusiveness and irresponsibility toward children

COMPLICATIONS. Dysphoria, tension, low tolerance for boredom, depressed mood, and premature violent death.

COMORBIDITY. These patients are at increased risk for Impulse Control Disorders, Major Depression, Substance Abuse or Dependence, Pathological Gambling, Anxiety Disorders, and Somatization Disorder. The most common co-occurring PDs are Narcissistic, Borderline, and Histrionic.

IMPAIRMENT. Extremely variable, typically includes social difficulties.

SEX RATIO. According to DSM-5, this disorder is more commonly (3:1) diagnosed in males.

EPIDEMIOLOGY. Prevalence rates of 3 percent for males and 1 percent for females in the general population, 3 to 30 percent in clinical settings with even higher rates for forensic samples and substance abusers, have been reported. As noted in DSM-5, there has been some concern that this disorder may be underdiagnosed in females, given the emphasis on aggressive items in diagnosing Conduct Disorder.

High frequency of Antisocial PD is associated with low socioeconomic status and urban settings.

COURSE. After the age of 30 years, both the most flagrant antisocial behaviors (promiscuity, crime) and the less severe behaviors and substance use tend to decrease. Even after the severe antisocial behavior "burns out," people diagnosed with antisocial PD usually continue to be irritable, impulsive, and detached.

FAMILIAL PATTERN AND GENETICS. Antisocial PD is more frequent among the first-degree biological relatives of probands with this disorder. Biological relatives of females with Antisocial PD are at increased risk for the same disorder than biological relatives of males with Antisocial PD. Genetic studies have suggested familial transmission of Antisocial PD, Substance Use, and Somatization Disorder, the former two being characteristic of males, the latter of females in the same family.

Adoption studies have shown that both genetic and environmental factors contribute to the risk for this disorder.

Both adopted and biological children of parents with Antisocial PD are at increased risk for this disorder. Conduct Disorder (before the age of 10 years) and accompanying Attention Deficit Hyperactivity Disorder increase the likelihood of developing antisocial personality in adult life. Conduct Disorder is more likely to develop into Antisocial Disorder with erratic parenting, neglect, or inconsistent parental discipline.

DIFFERENTIAL DIAGNOSIS. Antisocial PD is distinguished from Bipolar Disorder, manic on the basis of episodic course and euphoric mood of the latter. Antisocial PD is distinguished from the following PDs:

- Narcissistic (which rarely manifests serious criminality, aggression, and deceit, and is characterized by excessive need for admiration and envy of others)
- Histrionic (which includes seductiveness, attention seeking, superficiality, and rarely serious criminality and aggressiveness)
- Borderline (which includes manipulativeness to gain nurturance, and affective instability). Contrary to the common belief that borderline personalities rarely commit crime, individuals with explosive or borderline temperaments (high Harm Avoidance, high Novelty Seeking, and low Reward Dependence) frequently manifest antisocial behaviors ("secondary" psychopathy). Secondary psychopathy is distinguished from antisocial personality proper (or "primary" psychopathy) as the latter has a different temperament profile: high Novelty Seeking with low Harm Avoidance and low Reward Dependence, corresponding to the adventurous temperament profile.
- Paranoid (which includes suspiciousness, guarded attitude, rarely serious antisocial behaviors)
- Adult Antisocial Behavior (with no personality pathology in the background)

Narcissistic Personality Disorder

CLINICAL CRITERIA. The hallmarks of Narcissistic PD are pervasive sense of grandiosity (in fantasy or in behavior), need for admiration, lack of empathy, and chronic intense envy. Diagnostic features also include at least five of the following:

- Grandiose sense of self-importance and specialness
- Preoccupation with fantasies of unlimited success, power, brilliance, beauty, or ideal love
- Sense of entitlement
- Regards self as "special" and unique
- Interpersonal exploitativeness such as taking advantage of others to achieve own needs
- Lack of empathy
- Excessive need for admiration and acclaims
- Intensive and chronic envy
- Arrogant and haughty attitude

Some of the associated features include the following:

- Fragile self-esteem (which exclusively depends on external admiration) with hypersensitivity to criticism
- High achievements more frequent than in any other PD

Strong feelings of shame and humiliation
Exhibitionism (behavior motivated by the pleasure of being looked at)
Fear of having their "hidden" imperfections and flaws revealed

COMPLICATIONS. Social withdrawal, depressed mood, and Dysthymic or Major Depressive Disorder in reaction to criticism or failure.

COMORBIDITY. These patients are at increased risk for Major Depression and Substance Abuse or Dependence (especially cocaine use). The most common co-occurring PDs are Borderline, Antisocial, Histrionic, and Paranoid.

IMPAIRMENT. Frequently severe, typically includes marital problems and interpersonal relationships in general.

SEX RATIO. According to DSM-5, this disorder is more commonly diagnosed in males (50 to 75 percent of diagnosed cases are males).

COURSE. Chronic. However, narcissistic symptoms tend to diminish after the age of 40 years, when pessimism usually develops.

EPIDEMIOLOGY. Prevalence rates of 2 to 16 percent in the clinical population, less than 1 percent in the general population are reported in DSM-IV-TR and DSM-5.

PREDISPOSING FEATURES. There may be a higher risk for this PD in the offsprings of narcissistic parents who impart on their children an unrealistic sense of grandiosity. In addition, most narcissistic persons are realistically talented, beautiful, or highly intelligent, as these features serve as the nucleus around which the sense of specialness is further organized.

DIFFERENTIAL DIAGNOSIS. Narcissistic PD is distinguished from Manic or Hypomanic episode by the episodic course, euphoria, and functional impairments in the latter two. Narcissistic personality is distinguished from the following PDs:

Antisocial (narcissistic exploitativeness is more driven by the wish to establish one's dominance than by material gains, history of Conduct Disorder, no excessive need for admiration)
Borderline (which includes unstable self-concept, chaotic behaviors, self-destructive gestures, chronic anxiety, and rarely high achievements)
Histrionic (which includes capacity for empathy, emotional display, and rarely unscrupulousness and exploitation of others)
Obsessive-Compulsive (which includes inflexibility, detail orientedness, and social isolation in addition to perfectionism and belief that others cannot do things as well)

Paranoid and Schizotypal (which include suspiciousness and social withdrawal).

Histrionic Personality Disorder

CLINICAL CRITERIA. The hallmarks of histrionic PD are pervasive and excessive self-dramatization, excessive emotionality, and attention seeking. Diagnostic features also include at least five of the following:

Inappropriate sexual seductiveness or provocativeness
Excessive need to be in the center of attention
Rapidly shifting and shallow expression of emotions
Suggestibility
Physical appearance used for attention-seeking purposes
Impressionistic speech lacking detail
Self-dramatization, theatricality, exaggerated expression of emotions
Relationships considered more intimate than they really are

Some of the associated features include the following:

Difficulties in achieving emotional intimacy in romantic or sexual relationships
Craving for excitement and stimulation
Promiscuity or complete sexual naiveness
Low tolerance for delayed gratification

COMPLICATIONS. Frequent suicidal gestures and threats to coerce better caregiving. Interpersonal relations are unstable, shallow, and generally ungratifying. Frequent marital problems secondary to the tendency to neglect long-term relationships for the excitement of new relationships.

COMORBIDITY. These patients are at increased risk for Major Depression, Somatization Disorder, and Conversion Disorder. The most common co-occurring disorders are Narcissistic, Borderline, Antisocial, and Dependent.

IMPAIRMENT. Frequently only mild and typically includes personal romantic relationships.

SEX RATIO. There seems to be a general agreement that this disorder occurs far more frequently among women. According to DSM-5, the disorder might be equally frequent among men and women.

EPIDEMIOLOGY. Prevalence rates of about 2 percent in the general population, 10 to 15 percent for psychiatric inpatients and outpatients are reported in DSM-IV-TR and DSM-5.

FAMILIAL PATTERN AND GENETICS. This disorder tends to run in families. A

genetic "link" between Histrionic and Antisocial PD and Alcoholism has been suggested.

DIFFERENTIAL DIAGNOSIS. Histrionic PD is distinguished from the following PDs:

- Antisocial (which includes antisocial behaviors and crime to gain profit, power, or some other material gratification, history of Conduct Disorder, no excessive self-dramatization, no exaggerated emotional expression)
- Borderline (which includes unstable self-concept, chaotic behaviors, self-destructive gestures, chronic anxiety, identity disturbance)
- Narcissistic (which includes fear of having their "hidden" imperfections and flaws revealed, sense of grandiosity and specialness)

ANTISOCIAL, NARCISSISTIC, AND HISTRIONIC SPECTRUM. Narcissistic, Antisocial, and Histrionic PDs qualify for the so-called spectrum disorders. The three disorders have been shown to aggregate in the same family and to co-occur in the same person. Phenotypically distinguishable disorders may be referred to as spectrum disorders if they meet the above two conditions. Spectrum disorders sometimes reflect differential expression of the same liability. Antisocial, Narcissistic, and Histrionic PDs exemplify the spectrum in which proneness to impulsivity and aggression (associated with high Novelty Seeking) interferes with character development and maturity. Empirical data show that symptoms of the three disorders tend to group around impulsivity, aggression, and dramatic affects. A spectrum disorder may also reflect a less deviant form of the other on an underlying liability scale. In that regard, the underlying antisocial behaviors increase in severity as one proceeds from Histrionic Personality (manipulation) via Narcissistic Personality exploitation) and Antisocial Personality (violent and property crime).

Borderline Personality Disorder

CLINICAL CRITERIA. The hallmarks of Borderline PD are pervasive and excessive instability of affects, self-image, and interpersonal relationships as well as marked impulsivity. Diagnostic features also include at least five of the following:

- Frantic efforts to avoid real or imagined abandonment. **Note:** Do not include suicidal or self-mutilating behavior covered in Criterion 5
- Unstable and intense interpersonal relationships, with alternating between idealization and devaluation
- Markedly and persistently unstable self-image or sense of self
- Impulsivity in at least two potentially self-damaging areas (spending,

sex, substance abuse, binge eating, reckless driving). **Note:** Do not include suicidal or self-mutilating behavior covered in Criterion 5

Recurrent suicidal behavior, gestures, threats, or self-mutilating behaviors

Instability of affect due to marked reactivity of mood

Chronic feelings of emptiness

Inappropriately intense anger or difficulty controlling anger

Stress related, transient paranoid ideation or dissociative symptoms

Some of the associated features include the following:

Tendency to undermine self when close to realizing a goal

Feeling more secure with nonhuman objects (pets, inanimate objects) than in interpersonal relationships

PREDISPOSING FACTORS. Numerous studies have pointed to early traumatic experiences in the etiology of this PD. Recently, a tripartite etiopathogenetic model, including childhood trauma, vulnerable temperament, and a series of triggering events, has been formulated. Both dynamic and biological psychiatry agree that a combination of early traumatic events and certain biological vulnerabilities (mostly in the emotional domain) represent primary etiological factors for this disorder. Familial aggregation of Borderline PD has been repeatedly demonstrated.

COMPLICATIONS. Psychotic-like symptoms (hallucinations, body image distortions, hypnagogic phenomena, ideas of reference) in response to stress; premature death or physical handicaps from suicide and suicidal gestures, failed suicide, and self-injurious behavior.

COMORBIDITY. These patients are at increased risk for Major Depression, Substance Abuse or Dependence, Eating Disorder (notably Bulimia), Post-Traumatic Stress Disorder, and Attention Deficit Hyperactivity Disorder. Borderline PD co-occurs with most other PDs.

IMPAIRMENT. Frequent and severe; frequent job losses, interrupted education, broken marriages.

SEX RATIO. According to DSM-5, this disorder is more commonly diagnosed in females (75 percent of diagnosed cases are females).

COURSE. Variable. Most commonly follows a pattern of chronic instability in early adulthood, with episodes of serious affective and impulsive dyscontrol. The impairment and the risk of suicide are the greatest at the young adult years and gradually wane with advancing age. In the fourth and fifth decades, these individuals tend to attain greater stability in their relationships and functioning.

EPIDEMIOLOGY. Prevalence rates of about 2 percent in the general population, 10 percent for psychiatric outpatients, 20 percent for psychiatric inpatients, and 30 to 60 percent among patients with PDs are reported in DSM-5. The diagnosis is more common in younger than in older samples, suggesting a natural tendency toward maturation and remission.

FAMILIAL PATTERN AND GENETICS. Physical and sexual abuse, neglect, hostile conflict, and early parental loss or separation are more common in childhood histories of patients with this disorder. Borderline PD is five times more common among relatives of probands with this disorder than in the general population. It also increases familial risk for Antisocial PD, Substance Abuse, and Mood Disorders.

DIFFERENTIAL DIAGNOSIS. Borderline PD is distinguished from Mood Disorder, Dysthymic Disorder, and Cyclothymia (with depression and mood swings mimicking borderline affective problems) based on efforts to avoid abandonment, unstable relationships with alternating between idealization and devaluation, identity disturbance, impulsivity in potentially self-damaging areas, chronic feelings of emptiness, and inappropriately intensive anger or difficulty controlling anger (these symptoms are rarely observed in Mood Disorder, Dysthymia, or Cyclothymia). Borderline disorder shares many features and is difficult to distinguish from all other PDs, frequently as an exclusion diagnosis based on typical clinical symptoms for other PDs. Borderline disorder is distinguished from Identity Problems (the latter is limited to a developmental stage).

Cluster C Personality Disorders

Avoidant Personality Disorder

CLINICAL CRITERIA. The hallmarks of avoidant PD are pervasive and excessive hypersensitivity to negative evaluation, social inhibition, and feelings of inadequacy. Diagnostic features also include at least four of the following:

- Avoidance of occupational activities that involve significant interpersonal contact because of fears of criticism, rejection, or disapproval
- Unwillingness to be involved with others unless certain of being liked
- Restraint in intimate relationships because of the fear of being shamed or ridiculed
- Preoccupation of being criticized or rejected in social situations
- Inhibition in new social situations because of feelings of inadequacy
- Reluctance to take personal risks or to engage in any new activities

because they may prove embarrassing

- ► Views self as socially inept, personally unappealing, or inferior to others

Some of the associated features include the following:

- ► Fearful and tense demeanor
- ► Fear of blushing or crying in front of others in response to criticism
- ► Social isolation accompanied by craving social relations and fantasizing about ideal relationships with others

COMPLICATIONS. Social phobia.

COMORBIDITY. These patients are at increased risk for Mood and Anxiety Disorders (especially Social Phobia, generalized type). The most common co-occurring disorders are Schizotypal, Schizoid, Paranoid, Dependent, and Borderline.

IMPAIRMENT. Can be severe and typically includes occupational and social difficulties.

SEX RATIO. According to DSM-5, this disorder is equally frequent in males and in females.

EPIDEMIOLOGY. Prevalence rates of 0.5 to 2 percent in the general population and 10 percent for psychiatric outpatients are reported in DSM-IV-TR and DSM-5.

COURSE. Frequently begins in childhood with shyness and fear of strangers and new situation. Disfiguring illness and shyness in childhood predispose children for this PD.

DIFFERENTIAL DIAGNOSIS. Avoidant PD is very difficult to distinguish from Social Phobia (many authors believe that these are alternative labels for the same or similar condition). In Social Phobia, specific situations rather than interpersonal contact are avoided. Panic Disorder with Agoraphobia also manifests avoidance but usually after the onset of panic attacks. Avoidant disorder is distinguished from the following PDs:

Schizotypal and Schizoid (social isolation of avoidant personalities is accompanied by the desire for social relations, which is not observed in Schizoid and Schizotypal PD)

Paranoid (which includes guarded attitude, preoccupation with hidden meanings and conspirational explanations of being events)

Dependent (which is focused on being taken care of rather than on the fear of negative evaluation)

Dependent Personality Disorder

CLINICAL CRITERIA. The hallmarks of Dependent PD are pervasive and excessive need to be taken care of that leads to clinging behavior, submissiveness, fear of separation, and interpersonal dependency.

Diagnostic features also include at least five of the following:

- Difficulty in making everyday decisions without excessive reassurance and advice from others
- Needs others to assume responsibility for major areas of his or her life
- Difficulties expressing disagreement with others because of fear of loss of support or approval. Note: Do not include realistic fears of retribution
- Lack of initiative
- Unrealistic preoccupation with fears of being left to take care of self
- Urgent search for another relationship as a source of care and support when a close relationship ends
- Extensive efforts to obtain nurturance and support from others (to the point of volunteering to do unpleasant things)
- Uncomfortable and helpless feelings when alone because of exaggerated fears of being unable to take care of self pessimism, self-doubt, self-devaluation.

Some of the associated features include the following:

COMPLICATIONS. Mood Disorders, Anxiety Disorders, Adjustment Disorder, Social Phobia. Low socioeconomic status, poor family, and marital functioning.

COMORBIDITY. These patients are at increased risk for Major Depression, Anxiety Disorders, and Adjustment Disorder. The most common co-occurring disorders are Histrionic, Avoidant, and Borderline.

IMPAIRMENT. Frequently only mild and typically includes interpersonal relationships and occupational functioning if independency is required.

SEX RATIO. Equally frequent in males and females (DSM-5).

EPIDEMIOLOGY. Prevalence rates of 0.5 to 0.6 percent are reported in DSM-5 for this specific diagnosis, even though dependency is a prominent feature of many PDs.

FAMILIAL PATTERN AND GENETICS. Unknown. Chronic physical illness or Separation Anxiety Disorder may predispose for Dependent PD.

DIFFERENTIAL DIAGNOSIS. Dependent disorder is distinguished from dependency seen in Mood Disorders, Panic Disorder, Agoraphobia, and as a result of general medical condition. Dependent disorder is distinguished from the following PDs:

- Borderline (which includes unstable, stormy relationships and the reaction to abandonment with rage, emptiness, and demands as opposed to increasing appeasement and submissiveness seen with dependent personalities)
- Histrionic (which includes gregarious flamboyance with active demands for attention)
- Avoidant (which includes social isolation because of the fear of negative evaluation as opposed to clinging and submissive behavior of dependent personalities)

Obsessive-Compulsive Personality Disorder

CLINICAL CRITERIA. The hallmarks of Schizoid PD are pervasive and preoccupation with orderliness, perfectionism, and mental and interpersonal control at the expense of flexibility, openness, and efficiency. Diagnostic features also include at least four of the following:

- Preoccupation with details, rules, lists, order, procedures, organization, or schedules to the extent that the major point of activity is lost
- Perfectionism that interferes with task completion
- Excessive devotion to work and productivity to the exclusion of leisure activities and friendships (not accounted for obvious economic necessity)
- Overconscientiousness, scrupulousness, and inflexibility about matters of morality, ethics, or values (not accounted for by religion or culture)
- Inability to discard worn off or worthless objects with no sentimental value
- Reluctance to delegate tasks or work with others unless they submit exactly to his or her way of doing things
- Stinginess (money viewed as something to be hoarded for future catastrophes)
- Rigidity and stubbornness

Some of the associated features include the following:

- Decision-making difficulties when no strict rules or established procedures dictate correct action
- Anger and frustration when not able to maintain control of physical or interpersonal environment (anger typically not expressed directly)
- Excessive attentiveness to their relative status in dominance–submission relationships
- Controlled and restricted emotional expression, reserved personal style
- Formal and serious quality of everyday relationships

PREDISPOSING FACTORS. It has been repeatedly demonstrated that Obsessive-Compulsive PD and Obsessive–Compulsive Disorder frequently coexist. Obsessions and compulsions have been repeatedly linked to high central serotonergic function, which is associated with anxiety in general, supporting the hypothesis that obsessions and compulsions represent psychological and behavioral mechanisms reflecting underlying anxiety.

COMPLICATIONS. Distress and difficulties when confronted with new situations that require flexibility and compromise. Myocardial infarction (secondary to features typical of type A personalities, such as time urgency, hostility, competitiveness).

COMORBIDITY. These patients are at increased risk for Major Depression and Anxiety Disorder. Equivocal evidence for increased risk of Obsessive-Compulsive Disorder.

IMPAIRMENT. Frequently severe, typically includes occupational and social difficulties.

SEX RATIO. According to DSM-5, this disorder is twice as common in males than in females.

EPIDEMIOLOGY. Prevalence rates of 2 to 8 percent in the general population and 3 to 10 percent for psychiatric outpatients are reported in DSM-IV-TR and DSM-5.

FAMILIAL PATTERN AND GENETICS. Some studies have demonstrated familial aggregation of this disorder.

DIFFERENTIAL DIAGNOSIS. Obsessive–compulsive PD is distinguished Obsessive–Compulsive Disorder based on true obsessions and compulsions in the latter. This PD is distinguished from the following PDs:

- Schizoid (which includes lack of capacity for intimacy and social isolation secondary to emotional detachment, as opposed to devotion to work, discomfort with emotions)
- Antisocial (which includes material goals in antisocial behavior and criminality as opposed to hypermorality of obsessive personalities)
- Narcissistic (which includes sense of grandiosity, self-aggrandizement, exhibitionism, and fear of having their "hidden" imperfections and flaws revealed)

Other Personality Disorders

Personality Change due to Another Medical Condition. This diagnosis is made when persistent change in someone's personality is due to another medical condition, such as temporal lobe epilepsy. The disturbance must be

marked and the direct result of the other medical condition, which must be specified. The disturbance cannot be better explained by another mental disorder and does not occur exclusively during the course of a delirium. It must meet the general criteria for a PD in terms of subjective distress or impairment in functioning. Several subtypes can be specified on the basis of the predominant feature:

Labile type
Disinhibited type
Aggressive type
Apathetic type
Paranoid
Other
Combined (more than one of the above features)
Unspecified (e.g., if change is due to substance abuse and features are not specified)

Other Specified Personality Disorder

This disorder is diagnosed when the DSM-5 criteria for dysfunctional personality (i.e., general criteria for a PD) are met, but no specific criteria for any of the subtypes are observed. For example, a patient may not meet criteria for any of the classified disorders but still manifest features from more than one PD ("mixed PD") which, together, cause distress or impairments in social or professional functioning. This category is also used when the clinician judges that a specific PD that is not included in DSM-5 is appropriate (e.g., depressive personality or passive aggressive personality).

Unspecified Personality Disorder

This diagnosis is applied when the general criteria for a PD are met and the clinician has insufficient information to make a specific diagnosis or chooses not to communicate the specific reason that the presentation does not meet the criteria for a specific diagnosis.

Summary Description of the DSM-IV Personality Disorders

Qualitative features that are diagnostic of a PD as classified in DSM-5 are summarized in Table 26–16. Cloninger has also provided practical criteria to detail these general features so that a clinician may first decide whether a person has any PD before proceeding to diagnose a specific cluster or category of illness. The general features of all PDs are the character traits, whereas the clusters and categories are determined by additional temperament features.

Table 26–16.
Quantifiable (Dimensional) Features of Personality Disorder

Consistent features
Low self-directedness
Irresponsible, blaming
No mature goals
Resourceless, helpless
Poor self-esteem
Undisciplined
Low cooperativeness
Intolerant of others
Lack of empathy
Unhelpful
Revengeful
Unprincipled
Variable features
High persistence (obsessive-compulsive symptoms only)
Low reward dependence (odd cluster only)
High novelty seeking (erratic cluster only
High harm avoidance (anxious cluster only)

Disadvantages of DSM-5 Classification

In addition to problems shared by all categorical models of personality (discussed later), the DSM classification has its specific limitations.

First, DSM-5 is an etiologically "atheoretical" classification of PD based solely on descriptive criteria derived from observations of prototypical cases. This "atheoretical" approach was expected to stimulate work on the etiopathogenetic understanding of deviant behaviors classified as PDs. Unfortunately, this has not been the case. The introduction of descriptive criteria in DSM-III has inspired extensive research of some basic psychometric aspects of the PD diagnosis but rarely studies of etiology, underlying dynamics, or pathogenesis. While this emphasis on psychometrics obviously reflects an attempt of this new field to delineate its subject, it is also reflective of the purely descriptive approach, which provided no etiopathogenetic models or hypotheses for testing. For example, the DSM criteria have been of little or no help in developing the psychobiological model described here. In contrast, the experimental animal and twin studies on which the psychobiological model of personality was built have now been confirmed and extended by extensive work in human beings using functional brain imaging and person-centered genomic–phenomic analyses.

Second, the DSM-IV describes as pathognomonic for PD maladaptive response patterns that are "enduring," that is, appearing before adulthood with long-term duration. In practice, however, it can be difficult to distinguish long-term maladaptation indicative of PD from chronic

personality changes caused by other factors, such as other mental disorders (e.g., chronic depression), long-term situational factors (e.g., job-related timidity), and by other medical disorders (e.g., irritability associated with hyperthyroidism).

Third, the DSM definition of PDs is inherently imprecise, as it requires that maladaptive behaviors cause **clinically significant** subjective distress or **clinically significant** impairment in social and occupational function, or both. **"Subjective distress"** usually presents as low self-esteem, anxiety, guilt, depression, and hypochondriasis. These symptoms are frequently diagnosed as another mental disorder while the background PD is overlooked, especially now that the multiaxial system of DSM-IV that distinguished personality variation from other disorders has been dropped in DSM-5. The quantifier "significant" is an arbitrary diagnostic element that can hardly be made objective, as shown by allowing a clinician's judgment without qualification or justification to determine what is significant.

CLINICAL AND PSYCHOMETRIC ISSUES

Egosyntonic and Alloplastic Nature of Personality Disorder

Patients with PD typically blame other people or unfavorable circumstances for their own problems. This externalizing of responsibility is a result of the following two characteristics. First, most of these patients perceive their own deviant behaviors as appropriate and adequate; in other words, their symptoms are egosyntonic. The exceptions to this are patients with Dependent and Avoidant PD. Both disorders are characterized by prominent anxiety associated with maladaptive behaviors, and this frequently causes the patients to perceive their symptoms as disturbing, that is, egodystonic.

Second, patients with PD try to change others, not themselves; in other words, their attitude is alloplastic. The exceptions to this rule are patients with Avoidant and Schizoid PD. Avoidant persons usually try to improve their own performance to avoid negative evaluation by others. Schizoid personalities are socially detached to the extent that they are usually indifferent to what others might do, think, or feel.

Personality Disorder: Social or Clinical Diagnosis?

One of the general DSM-5 requirements for PD is that maladaptive behavior "deviates markedly from the expectations of the individual's culture." This requirement introduces a significant social connotation to what was meant to be a clinical classification. Indeed, clinicians frequently diagnose PD based solely on the extent to which certain behavior is accepted within the local society.

Sociocultural pressures are always norm-favoring, promoting

phenotypes that fall within the range of accepted behavior norms. As shown later in the text, most temperament profiles with high Reward Dependence have a low incidence of diagnosed PD, whereas those with high Harm Avoidance and/or high Novelty Seeking increase this risk. In most Western societies, the normative phenotype is not the one with average personality traits (as one would expect given the adaptive flexibility of such configuration) but the one with high Reward Dependence (high Reward Dependence also means much easier conditioning of socially accepted behaviors than is the case with other temperament traits).

In a similar way, child psychiatry classifies three broad groups of mental disorders in children: one group is characterized by high Novelty Seeking (e.g., externalizing disorders such as Attention Deficit Hyperactivity Disorder, Conduct Disorder), one group by high harm Avoidance (e.g., internalizing disorders such as Depressive and Anxiety Disorders), and one group by low Reward Dependence (asocial disorders such as Autism). Children with these extreme temperaments are diagnosed and treated to change their behavior, whereas children with high Reward Dependence (sociable, attached, dependent) are selectively permitted to develop with no corrective intervention. This raises a question about the role of medicine and psychiatry in "engineering" psychological profiles of young children by "cherry-picking" the favorable profiles, and suppressing the unfavorable ones, which is beyond the scope of this chapter. Suffice here to say, biological adaptation appears to have become less critical for human survival than behavioral and spiritual adaptation. Natural selection is not targeting human morphological features any more but rather, human personality features and level of awareness. Through its suppression of certain temperament traits and its norm-favoring influence on character development, modern human society appears to have an important role in this process (see Josefsson et al., 2013).

Categorical versus Dimensional Approach to Personality Disorder

Medical diagnosis has historically been categorical and most clinicians tend to "think" categorically for two reasons. First, one of the most important functions of diagnosis is to prescribe appropriate corrective action, that is, treatment. The decision whether a person is "affected" or "unaffected" ("sick" or "not sick") is functionally equivalent to the decision whether or not that person needs treatment (hence, categorical diagnosis facilitates treatment decisions in practical work). Second, categorical systems simplify professional communication as they describe prototypical cases that meet all or most of the diagnostic criteria. Just one category, for example, antisocial personality, can communicate a great deal of vivid clinical information about the patient. However, categorical descriptions convey useful information only about prototypical and severe cases but little or no information about atypical, mixed, or mild cases (see later).

Categorical Approach. The categorical approach adopts the medical diagnostic model that a PD is either "present" or "absent" and individuals either "affected" or "unaffected." Categorical models are optimal for discontinuous variables (e.g., marital status, fractured humerus). However, when distinguishing characteristics of a disorder are variants of normality, such as personality traits, any categorical decision about the presence or absence of a disorder is arbitrary. In other words, one of the major disadvantages of categorical models of PD is that they establish arbitrary cutoff points or thresholds for continuous behavior traits. This becomes clear especially in mild cases close to the cutoff point, which cannot be well classified neither as "affected" or "unaffected." In an attempt to minimize this problem, DSM-5 now allows clinicians to classify "maladaptive personality traits" that do not meet the threshold for a PD. Nevertheless, the categorical approach is not optimal for patients who do not perfectly fit their "pigeonhole" in the classification, so an additional wastebasket category for "atypical" or "mixed" cases had to be established, and this often is used to describe most cases of PD. DSM-5 calls this wastebasket "Other specified personality disorders" but that does not make it more than the wastebasket it was when it was frankly acknowledged to be "not otherwise specified." Finally, the categorical approach to personality does not even establish a prescriptive relationship between diagnosis and treatment because of the significant overlap of diagnostic criteria for PD subtypes, and categorical systems usually yield multiple personality diagnoses for individual patients. In such cases, treatment decisions and priorities are not easy to establish.

Dimensional Approach. Dimensional models define a number of graded and continuous behavior dimensions and specify individual differences as quantitative variations along these dimensions. These models are optimal for features that vary quantitatively (e.g., as personality does) and have different adaptive significance in different situations. These models account for the fact that everyone has multiple personality traits, more or less prominent and adaptive, rather than being simply present or absent. This eliminates the need for description of multiple and overlapping categorical diagnoses. By quantifying one's position on a continuous personality trait, dimensional models conserve information about individual patients more than categorical models. Personality dimensions are as easily applied to the common atypical and mild cases as they are to the rare prototypical and severe cases. Finally, by using a dimensional system, one can easily manipulate complex data mathematically and include many variables.

Research has shown that normal personality traits and deviant personality traits observed in PDs represent a continuum where deviant traits are extreme variants of normal, adaptive behaviors. Note, however,

that dimensional models do not answer the question what is it about some traits that makes them disorders. Natural breaks or points of rarity on the continuum of normal and maladaptive personality traits have not been detected. The extreme standing alone is not sufficient to be dysfunctional (e.g., people with low Reward Dependence can be well adapted in their environment). Dimensional models have to introduce cutoff points for continuous traits to make them more compatible with categorical diagnoses and thus more useful for treatment decisions or, as proposed here, to use probability estimates for categorical PD diagnosis based on the severity of measured temperament and character traits.

Integrated View of the Importance of Personality Subtypes

There is widespread dissatisfaction with the current official criteria for PD subtypes. The most common clinical diagnosis of PD using DSM-IV was "not otherwise specified." There is no evidence that the putative categories are truly discrete entities that are either present or absent. So should clinicians rely on dimensional or prototypical approaches? Dimensional approaches can consider configurations of scores on multiple dimensions that correspond to particular subtypes of PD. Prototypical approaches allow clinicians to rate the degree to which a particular case matches with the criteria of a particular subtype, now regarded as a prototype or model. What these suggestions ignore is that complex adaptive systems, such as personality, do have moderately stable configurations and that such configurations have modest separation from neighboring conditions. Furthermore, these metastable, quasi-discrete configurations are often maintained by particular gene–environmental interactions that are most predictable when a specific configuration of traits are specified. What is most important is for the clinician to realize that personality subtypes are configurations of component processes that interact as expressions of a nonlinear dynamical system. They are neither discrete taxons nor arbitrary configurations of dimensions. However, they tend to be stable over time which is mostly reflecting the fact that most (but not all) people with PD do not "spontaneously" mature but need professional intervention or some major life events (either extremely good or extremely bad or tragic) to "jump start" the process. It is useful to recognize traditional subtypes of PD as long as you realize they are metastable, quasi-discrete expressions of complex psychobiological processes that are adaptive and dynamic.

In other words, the advantages of categories and dimensions can be combined by recognition that in complex adaptive systems, it is the configuration of multiple dimensions into a personal profile that determines the functionality of the system. Such multidimensional profiles identify specific individuals with specific characteristics so that they have the advantages of categorical diagnoses without falsely implying that they represent a discrete disorder. For example, a person with antisocial

personality is specified in terms of high Novelty Seeking combined with low Harm Avoidance and low Reward Dependence in people with immature character. The person-centered approach to profile analysis provides clusters of people whose psychobiological characteristics and behaviors are relatively stable but can be modified by changing circumstances, age, or treatment. Consequently, clinicians are not misled to slip into the fallacies of thinking that PDs are discrete disease entities or that people with PDs cannot change.

Measurement of Personality Disorders

Self-Report or Interview? Interviews are often considered by psychiatrists to be more reliable and valid diagnostic tools than self-reports for measuring PD, whereas psychologists often advocate psychometric testing. In fact, no data for a definite advantage of one of these formats have been established. People in general are biased when self-reporting deviant behaviors (in PDs, these behaviors are ego-syntonic, perceived as "OK"), so interviewing the patient reduces this problem because unclear, inconsistent, or defensive responses can be clarified and the patient's demeanor and appearance observed. On the negative side, interviews tend to be affected by systematic biases, ideological orientation, experience, and rating idiosyncrasies of interviewers. In self-reports, various response sets (e.g., careless inconsistency, defensiveness, and exaggeration) and validity items can be measured to help in interpretation. Recently, a set of performance scales, independent of the item content, based on the observed regularities in the technical pattern of responding to questions has been defined for validity analysis of the TCI. These performance scales are used to test the validity of self-reported personality traits. Specifically, self-report to questions, regardless of the content of the questions, tends to follow specific patterns that correlate with personality traits. For example, the number of endorsed items that are otherwise rarely found in the general population correlates with low Reward Dependence. These performance scales are used to predict scores on each of the temperament and character dimensions, which are then compared with self-reported, content-dependent, scores on these same traits. Substantial score differences between "subjective-" and "performance"-based scores on personality traits indicate either significant bias or intentional misrepresentation. Training in the use and interpretation of the TCI is available online (https://anthropedia.org).

Both self-reports and interviews depend on the patient's accuracy, honesty, and level of insightfulness. Hence, the collection of data from collateral informants and expert ratings is usually considered critical to ensuring high-quality personality diagnosis.

State-Trait Effect. Most major clinical disorders have their less severe

variants in the form of PDs (Table 26–17). The "state-trait" distinction between symptoms reflecting long-term personality traits and those reflective of transient clinical states can be difficult. Indeed, transient variation in one's emotional state can influence both self-report and clinical evaluation of long-term personality characteristics. The susceptibility to such state effects has been demonstrated for some categorical measures of PDs. The state effect can be reduced, although not eliminated, when subjects are repeatedly reminded to describe their usual, not their current, behavior, feelings, and attitudes. In contrast to categorical diagnosis, only those dimensions of personality that underlie susceptibility to depression and anxiety, such as neuroticism and Harm Avoidance, tend to covary with mood and anxiety states. Also, increases in anxiety or depression lead to mild and transient decrease in Self-Directedness and Cooperativeness. Consequently, a person with no enduring PD may act immaturely when depressed or under stress. Self-reports of other dimensional traits, Novelty Seeking, Reward Dependence, and Persistence in particular, are largely independent from and unaffected by comorbid mood and anxiety states.

Categorical Tests for Personality Disorder. The majority of categorical tests for PDs are based on the DSM criteria and share all the advantages and disadvantages of that categorical approach. In the DSM-III field trials, the interrater reliability was 0.61 and the test–retest reliability was 0.54 for any PD. The reliability is usually lower for individual subtypes than for the presence versus absence of any PD because of the difficulty distinguishing among the subtypes. In the DSM-5 field trials, reliability was questionable for most diagnoses and the reliability for Borderline PD was only 0.34. As a result, the ICD-11 is considering a simpler approach that clinicians can implement more reliably.

Table 26–17.
Phenotypic Similarity between Major Clinical Syndromes and Personality Disorders

Clinical Syndromes: Major Variant	Personality Disorders: Minor Variants	Shared Symptoms
Schizophrenia spectrum	Schizoid PD Schizotypal PD	Negative and positive
Delusional disorder	Paranoid PD	Suspiciousness
Major depression	Depressive PD Borderline PD	Proneness to depressed mood; proneness to self-destruction
Cyclothymia Bipolar disorder Mania	Narcissistic PD Histrionic PD Antisocial PD	Mood swings; impulsivity
Obsessive-compulsive disorder	Obsessive-compulsive PD	Hypochondriasis; inflexibility

Social phobia	Avoidant PD	Shyness; avoidant behavior
Panic disorder with agoraphobia	Dependent PD	Dependency

Among the **categorical self-reports**, the two most frequently used are The Personality Diagnostic Questionnaire (PDQ R), and the MMPI scales for PDs (MMPI PD). The PDQ R is a self-report version of the DSM criteria for PDs, along with the "Impairment/Distress" scale to quantify the significance of impairment or subjective distress required by DSM criteria. In practice, this instrument has been shown to overdiagnose PDs when compared to expert clinical diagnosis.

The MMPI PD scales were derived from the MMPI item pool to evaluate symptoms for DSM-III PDs. The MMPI PD scales are incorporated into MMPI so that, in addition to evaluating PD and other clinical syndromes, clinicians and researchers can also use other MMPI scales (e.g., Lie, Validity, Defensiveness) for validity and response set analyses.

Several semistructured interviews have been developed to diagnose PDs. These typically take more than an hour to conduct, so they have limited use in clinical practice. Among the **interviews that categorically diagnose** PDs, the most commonly used is The Structured Interview for DSM-IV PDs (SIDP IV). The original SIDP was developed for PDs described in DSM-III, and it has been regularly updated with each new DSM revision and edition (SIDP R for DSM-III-R, SIDP IV for DSM-IV). To increase its diagnostic validity, the SIDP-IV requires that answers for selected questions be obtained from a collateral informant. This test can be combined with a 10-minute SIDP-IV screen for PD, which can considerably reduce testing time in populations with many individuals who do not have PD. The screening instrument alone tends to overdiagnose cases but is useful to screen out people with no PD.

Other popular interviews for categorical classification of PDs include the Structured Clinical Interview for DSM-IV PDs (SCID II), The Diagnostic Interview for Personality Disorders (DIPD), and the Personality Disorder Examination (PDE). The PDE is available in an international version from the World Health Organization in several languages and makes diagnoses according to the International Classification of Diseases in addition to DSM-IV. The interviews for DSM-IV PD remain the same for DSM-5 because the official criteria have not changed. There is a questionnaire for the alternative DSM-5 criteria called the Personality Inventory for DSM-5, but this involves questions about features that are not sanctioned by either ICD-11 or DSM-5 officially.

Dimensional Assessment of Personality Disorders. Historically, personality has been studied and evaluated using dimensions or traits, and PD categorically, as discrete entities of psychopathology. Hence, conceptual

advances in the field of normal behavior and personality have had little or no effect on research in the field of PD and vice versa. For example, if the features associated with Eysenck's model of personality, that is, neuroticism, extraversion, and tough-mindedness, are combined, the resulting combinations do not correspond closely to traditional categories of PD.

The popular Five-Factor Model of normal personality has been recently advocated to account for underlying dimensions of PD as well. However, the ability of the Five-Factor Model to account for the underlying structure of PD has been tested primarily in nonclinical samples and normal individuals. The few studies that tested the Five-Factor Model in clinical samples showed that neuroticism, extraversion, and low agreeableness predict PD symptoms. After controlling for age and depression, only high neuroticism and low agreeableness remained as significant predictors of PD symptoms. These two dimensions define PD in a nonspecific way, namely, as a general predisposition to psychopathology (i.e., high neuroticism) accompanied by an antagonistic behavior facade (i.e., low agreeableness).

Neuroticism also tends to be confounded with nonspecific factors (such as depression or anxiety) that reduce its ability to distinguish PD, other psychopathology (e.g., mood and anxiety disorders), and well-adjusted individuals with high neuroticism, leaving open the question of its specificity for PDs.

In the authors' work, they found that some common, core features, distinct from other forms of psychopathology that causes personal dysfunction, characterize PD subtypes classified in DSM-IV and DSM-5 as discrete categories. For example, DSM-IV and DSM-5 arrange all PDs into three Clusters (A, B, and C) each comprising a number of PD subtypes with similar clinical features. Cluster A includes subtypes with predominantly odd and eccentric symptoms (e.g., paranoid and schizoid PD), Cluster B includes subtypes with dramatic, erratic, and impulsive symptoms (e.g., Histrionic, Antisocial, Borderline PD), and Cluster C includes subtypes with fearful and anxious symptoms (Dependent, Avoidant, Obsessive PD). These clusters represent three independent dimensions underlying PD subtypes within each of the clusters. However, the ability of these three features to discriminate subjects with and without PD or to distinguish between PD subtypes has not been tested.

Dimensional tests of PDs are formatted as self-reports (except for Tyrer's Personality Assessment Schedule). The most frequently used are the Millon Clinical Multiaxial Inventory (MCMI), the earlier mentioned NEO PI that evaluates the Five-Factor Personality Model, and the TCI that evaluates the Seven-Factor Psychobiological Personality Model. These tests have been originally designed to evaluate normal personality (the NEO PI), psychiatric patients (the MCMI), and both normal personality and deviant

personality (the TCI). The NEO PI is discussed previously vis-a-vis the homogeneity of its component traits and its usefulness in clinical work with PD. The other two tests are briefly discussed later.

The MCMI III (designed to match DSM-IV PD categories) evaluates long-term behavior traits systematized as 10 Basic Personality Patterns (dependent, avoidant, schizoid, passive–aggressive, narcissistic, antisocial, hysterical, compulsive, aggressive, and self-defeating) and three Pathological PDs (i.e., Borderline, Schizotypal, and Paranoid). The latter are used to determine the severity of the 10 Basic Personality Patterns. The test also evaluates nine clinical syndromes. Some reports suggest that MCMI III tends to overdiagnose PD in comparison to expert clinicians.

In the literature, the number of proposed dimensions that underlie deviant personality ranges from 2 to 24 in the literature, although, in most models, can be reduced to four dimensions (aloof, anxious, anankastic, and adventurous). The optimal number of underlying dimensions and the content of dimensions that purport to describe personality and its disorders most efficiently are still vigorously debated. This ambiguity arises because factor analyses of the behavior of unrelated individuals are always compatible with an infinite number of rotations of the inferred phenotypic factors. Fortunately, recent GWAS have been replicated in multiple independent samples with results that identify unique sets of genes underlying each of the facets of the TCI using a generalized clustering method that does not make arbitrary assumptions about the number of personality factors or their relationships with one another in individuals. In what follows, the distinction between temperament and character is used to improve both the etiological understanding and the diagnosis of PD. In the treatment section of the chapter, the concepts of temperament and character are used to outline treatment guidelines for patients with PD. First, the TCI is described briefly.

The Temperament and Character Inventory. The TCI evaluates four major temperament dimensions (Harm Avoidance, Novelty Seeking, Reward Dependence, and Persistence), and three major character dimensions (Self-Directedness, Cooperativeness, and Self-Transcendence). Each of these main temperament and character dimensions is composed of component facets (or subscales) to evaluate response patterns elicited by specific stimuli (e.g., Harm Avoidance is observed in different settings as worry and pessimism, and/or fear of uncertainty, shyness, or excessive fatigability). The seven major dimensions and their correlations with the number of symptoms of personality disorders are shown in Table 26–18. The TCI is a family of tests with several specialized versions designed for varying types of informants (self-report, peer ratings, interviewers), varying age groups (for 3 to 5 years the preschool TCI, 7 to 14 years the Junior TCI, for 12 to 18 years the adolescent TCI, and 15+ years the Adult TCI), scope of

information (temperament, character, both), and level of clinical detail (from 60 items for major dimensions only to 240 items for multiple facets of each dimension). The tests for various age groups overlap in their range so that exactly the same test can be administered to a person as they develop across the life span. The test measuring the temperament dimensions only was originally called the Tridimensional Personality Questionnaire (TPQ). In the TPQ, Novelty Seeking, Harm Avoidance, Reward Dependence, and Persistence were all measured, but Persistence was originally scored as a component of Reward Dependence. The name of the test was changed to TCI when the character scales were added and Persistence was recognized as a fourth, separately inherited temperament dimension in twin studies in Australia and United States. The self-report version of the TCI is a 240-item test available in both true/false and five-point Likert formats. Its psychometric properties are presented in the TCI Manual.

The ability of the TCI to predict categorical diagnoses of PDs was clinically tested in samples of psychiatric inpatients and outpatients with and without PDs and varying mood and anxiety states. Correlation and regression analyses consistently demonstrate that low scores on character dimensions, especially Self-Directedness and Cooperativeness, are associated with high symptom counts for any PD, for each of the DSM clusters of PD (Tables 26–18 and 26–19), and for each clinical subtype of PD, with or without comorbid clinical syndromes and independent of age. As shown in Table 26–19, the TCI character scales of Self-Directedness, Cooperativeness, and Self Transcendence predict the number of PD symptoms, after controlling for age, anxiety, and depression.

Table 26–18.
Correlations between Temperament and Character Inventory Scales and the Total Number of Symptoms for Personality Disorders, Cluster A, Cluster B, and Cluster C

	Total No.PD Sx	Cluster A Sx	Cluster B Sx	Cluster C Sx
Novelty Seeking	**0.22**[a]	0.02	**0.44**[b]	–0.06
Harm Avoidance	**0.31**[c]	**0.23**[c]	0.08	**0.43**[b]
Reward Dependence	–0.14	**–0.37**[b]	–0.08	–0.04
Persistence	0.00	–0.07	0.04	–0.01
Self-Directedness	**–0.56**[b]	**–0.35**[b]	**–0.43**[b]	**–0.50**[b]
Cooperativeness	**–0.44**[b]	**–0.44**[b]	**–0.40**[b]	**–0.28**[c]
Self-Transcendence	0.02	–0.08	0.03	0.04

Personality Disorders Inpatients (*N* = 136).

[a]<0.01.

[b]<0.0001.

[c]<0.001.

Clearly, low scores on character traits represent a core feature (common denominator) of all PD clusters and subtypes.

In differential diagnosis, PD subtypes are distinguished on the basis of temperament scores without any overlap because each subtype correlates with the TCI temperament dimensions in a unique way (Fig. 26–10).

Table 26–19.
Stepwise Multiple Regression

Total Number of PD Symptoms		
	Partial R^2	*P*
Total number of interview PD symptoms predicted by TCI (controlling for age and depression; inpatients [*N* = 136])		
Age and Depression	0.35	0.0001
Novelty Seeking	—	NS
Harm Avoidance	—	NS
Reward Dependence	—	NS
Persistence	—	NS
Self-Directedness	0.13	0.0001
Cooperativeness	0.03	0.0076
Self-Transcendence	0.02	0.0376
Cumulative R^2	0.18	0.0001
Model R^2	0.53	0.0001
Total number of self-report PD symptoms predicted by TCI (controlling for age, depression and anxiety; outpatients [*N* = 109])		
Age, Anxiety, and Depression	0.45	0.0001
Novelty Seeking	—	NS
Harm Avoidance	—	NS
Reward Dependence	—	NS
Persistence	—	NS
Self-Directedness	0.20	0.0001
Cooperativeness	0.05	0.0001
Self-Transcendence	0.05	0.0001
Cumulative R^2	0.30	0.0001
Model R^2	0.75	0.0001

The aforementioned results have important diagnostic and treatment ramifications for PD. Diagnosis is discussed later; the treatment aspects are reviewed in the treatment section.

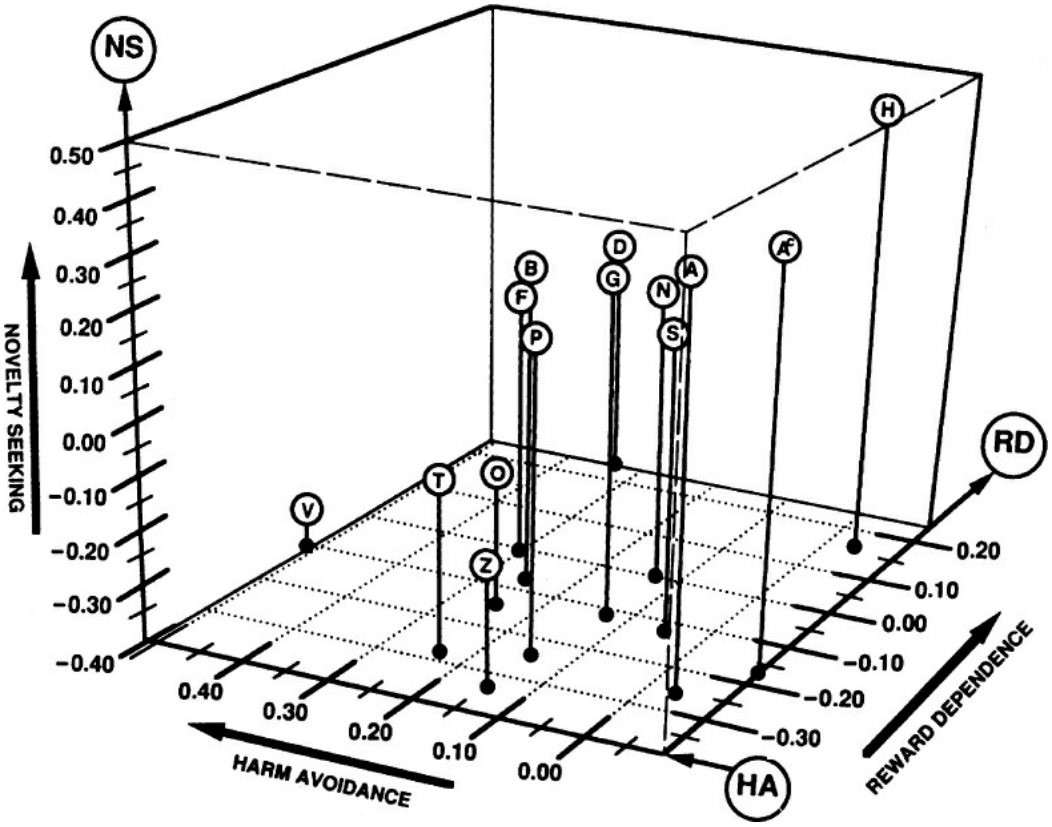

FIGURE 26–10. Correlations between individual personality disorder subtypes and temperament traits of harm avoidance (HA), novelty seeking (NS), and reward dependence (RD). A, antisocial; Ac, antisocial children; B, borderline; D, dependent; F, self-defeating; G, passive aggressive; H, histrionic; N, narcissistic; O, obsessive; P, paranoid; S, sadistic; GT, schizotypal; V, avoidant; Z, schizoid.

DIAGNOSIS OF PERSONALITY DISORDERS. Character and temperament traits delineate both the core features and the distinguishing features of PDs (Tables 26–18 and 26–19 Fig. 26–10). Poorly developed character traits, especially Self-Directedness and Cooperativeness, are a common denominator extending across all subtypes of PD and are used in diagnosis to predict categorical PD symptoms (Tables 26–18 and 26–19). Clinically, low character scores explain chronic difficulties in accepting responsibility, setting long-term goals, fragile self-esteem, and other features associated with low Self-Directedness that are so characteristic of individuals with PD as described in DSM-5. Usually, but not always, these patients are also uncooperative as well (i.e., revengeful, opportunistic, self-centered, socially intolerant, unhelpful to others, and lacking in empathy, compassion, or principles), as is also described in DSM-5.

High Self-Transcendence correlates with schizotypal and paranoid symptoms (depicting primary process thinking and fanaticism, respectively) and with borderline, histrionic, and narcissistic symptoms (depicting proneness to dissociation), much as predicted in Figure 26–6. In addition, recent data indicate that high Self-Transcendence might be an important component in one's susceptibility to psychosis (when other two

character traits are low) or, on the contrary, it may predispose to high creativity (when other two character traits are high). Recent psychobiological studies suggest that the healthiest and happiest people are those who are high in all three character dimensions. Now that the human population's use of resources is exceeding what the planet can replenish, Self-Transcendence has become crucial if the human species is to avert mass extinction. Human health and survival depends on cultivation of self-transcendent values and behavior that are altruistic and creative. In addition, the maturity and integration of all three character traits promotes physiological health, happiness, and fulfillment in ways that are indicated objectively by measures of psychophysiological coherence and markers of longevity, such as optimal heart rate variability, vagal tone, and DHEA/cortisol ratio.

High Self-Directedness alone is not always protective against PD. Some narcissistic and antisocial persons can be self-sufficient, resourceful, and successful but still maladapted because of low Cooperativeness (i.e., intolerance of others, revengefulness, low empathy, etc.).

Once the probability for the presence of PD is established on the basis of character, temperament traits are used for differential diagnosis. Membership in Cluster A (aloof, eccentric personalities) is most clearly determined by low Reward Dependence (described as social detachment). Membership in Cluster B (dramatic, erratic personalities) is most clearly determined by high Novelty Seeking (described as impulsivity). Membership in Cluster C (anxious, fearful personalities) is determined by high Harm Avoidance (described as fearfulness) (Table 26–20). Recent data indicate that Persistence predicts Obsessive–Compulsive traits.

Individual subtypes of PD are distinguished on the basis of their unique pattern of correlations with temperament traits without any overlap (Fig. 26–10). In other words, clinical presentation of PD varies along four underlying dimensions, corresponding to the four temperament traits of Harm Avoidance, Novelty Seeking, Reward Dependence, and Persistence (Table 26–16).

The authors' model does not require arbitrary cutoffs for the presence or absence of PD. Personality disorder can be quantified in terms of degree of severity or probability of diagnosis can be estimated. As shown in Table 26–21, the probability for the presence or absence of PD based on the Self-Directedness scores alone ranges from very low (cca. 10 percent) to very high (cca. 95 percent) in clinical samples.

The authors' model "reconciles" some aspects of categorical and dimensional approach to PD. It specifies seven temperament and character *dimensions* to account for individual differences in behavior. These dimensions interact to create temperament and character composites (types, profiles, or configurations), which correspond to **categorical** DSM subtypes of PD (Figs. 26–4 and 26–5). These multidimensional profiles

also correspond to the clusters of people with particular profiles of traits associated with sets of genes identified in person-centered GWAS, as noted earlier. In other words, this model is dimensional but still can be used for categorical clinical diagnosis, because each temperament and character composite profile corresponds to one clinical category of PD (in practice, this is useful for purposes of description and prediction of course). These multidimensional profiles are related to underlying causal variables strongly but in complex ways.

Table 26–20.
TCI Scales and Subscales

I. Temperament:

Harm Avoidance (HA)
HA1: Worry and pessimism vs. uninhibited optimism
HA2: Fear of uncertainty
HA3: Shyness with strangers
HA4: Fatigability and asthenia

Novelty Seeking (NS)
NS1: Exploratory excitability vs. stoic rigidity
NS2: Impulsiveness vs. reflection
NS3: Extravagance vs. reserve
NS4: Disorderliness vs. orderliness

Reward Dependence (RD)
RD1: Sentimentality
RD2: Sociability vs. aloofness
RD3: Attachment vs. detachment
RD4: Dependence vs. independence

Persistence (PS)
PS1: Eagerness of effort vs. laziness
PS2: Work hardened vs. spoiled
PS3: Ambitiousness vs. underachieving
PS4: Perfectionism vs. pragmatism

II. Character:

Self-Directedness (SD)
SD1: Responsibility vs. blaming
SD2: Purposefulness vs. lack of goal direction
SD3: Resourcefulness vs. helplessness
SD4: Self-acceptance vs. self-striving
SD5: Spontaneity vs. lack of congruent second nature

Cooperativeness (CO)
C1: Social acceptance vs. social intolerance
C2: Empathy vs. social disinterest
C3: Helpfulness vs. unhelpfulness
C4: Compassion vs. revengefulness
C5: Fairness/principles vs. self-serving opportunism

Self-Transcendence (ST)
ST1: Self-forgetfulness vs. self-preoccupation

ST2: Transpersonal identification vs. alienation
ST3: Spiritual acceptance vs. skepticism
ST4: Contemplation vs. objectivity
ST5: Idealism vs. practicality

Table 26–21.
Predictions of Personality Disorders from Self-Directedness Scores[a]

Self-Directedness Scores	Predicted with PD (%)	Observed with PD f/N	Observed with PD %
44–39	13	4/18	22 ± 0.10
38–34	23	5/17	29 ± 0.11
33–30	33	4/16	25 ± 0.11
29–27	44	8/20	40 ± 0.11
26–23	55	8/19	42 ± 0.11
22–20	66	6/12	50 ± 0.14
19–16	70	14/16	88 ± 0.08
<16	85	17/18	94 ± 0.05

[a]Predicted % = 100 $e^y/(1 + e^y)$, where $y = -3.37 + .13$ (SD score).

PATHOGENESIS OF PERSONALITY DISORDER

Only a few developmental models of PD are available. Of those, the psychodynamic model focuses on the effects of early traumatic experiences and a constitutionally high disposition to anxiety and anger, whereas the spectrum hypothesis suggests that PDs may be a minor form of particular major clinical syndromes. A growing body of evidence points to a nonlinear pathogenesis with multiple factors, including genetic, physiological, psychological, social, and spiritual influences, which are fully incorporated into the psychobiological model of PD.

The Spectrum Disorder Hypothesis

Echoing Kretchmer's work from many decades ago, some experts still consider subtypes of PDs as "minor" variants of "major" clinical syndromes. Indeed, most major clinical disorders have their less severe variants or "representatives" as PDs. For example, Schizophrenia is represented by Schizoid PD, Delusional Disorder by Paranoid PD, Social Phobia by Avoidant PD, Cyclothymia by the group of emotionally unstable Cluster B PDs, and so forth (Table 26–17).

These major and minor disorders are symptomatically similar and frequently differentiated based only on the severity of their clinical presentation. The etiopathogenetic nature of this major–minor relationship has not been established, aside from the obvious symptomatic similarity. Personality disorders might be predisposing factors, premorbid features, or subclinical expression of Axis I disorders. Their common etiopathogenesis,

however, is still an open question. So far, a close etiological relationship has been solidly supported only for Schizotypal personality and the spectrum of Schizophrenia.

Psychodynamic Model of Personality Disorder

The central idea behind this model is that PD subtypes, classified as separate nosologic units, reflect different behavioral expressions of the same core deficit in personality. The so-called "borderline level of personality organization" described by Kernberg has been postulated to represent the nucleus shared by most subtypes of PD.

As defined by Kernberg (1975), the borderline level of personality organization is characterized by *nonspecific manifestations of Ego weakness* (such as lack of impulse control, lack of anxiety tolerance, and lack of sublimatory potentials), *specific Ego defects* (manifested as partially blurred self-object boundaries, mild to moderate primary process thinking, periodically distorted reality testing), *partial object relations* (manifested as alternations between "all good" and "all bad" perception of the self and external objects), *primitive defense mechanisms* (centered around splitting), and *identity disturbance*. In addition, an inadequately developed Super-Ego is either sadistic (e.g., obsessive personalities) or rigid in some but completely absent in other areas, permitting conflict-free expression of impulses and behaviors.

Immature defenses and a fragmentary self-concept (which are very similar to the above-described borderline personality organization) are a normal early phase in mental development. In the case of PD, however, this early developmental phase pathologically persists in later periods. Etiologically, this pathological persistence of immaturity is postulated to reflect either constitutional factors (e.g., poor anxiety tolerance, high aggressivity, genetic vulnerability to certain affects) or environmental traumatic factors (e.g., early separation, loss of a parent, physical or mental abuse, neglect).

According to the theory, because of such traumatic etiological factors, aggressively charged, negative representations of the self and external objects are predominately incorporated into the internal world. This, in turn, interferes with the crystallization of the early Ego, which normally occurs around positive, libidinally cathected representations of the self and external objects. In other words, normal early motivation and growth are centered around positive primary emotions, particularly interest and joy. The relative predominance of strong negative emotions, particularly anger and fear, interferes with normal character development. As a consequence, primitive defenses (splitting and the related defenses), which are normally predominant in this early phase of life, pathologically persist in the inner world and interfere with normal development. More mature defenses, such as repression, require more energy for their operation than splitting (e.g.,

repression requires at least as much energy for its operation as is contained in unacceptable impulses). In contrast, splitting simply keeps opposite "good" and "bad" self and object representations apart from each other. As Kernberg pointed out, splitting and early Ego weakness mutually perpetuate each other. By protecting positive self and object representations, that is, keeping them apart from their negative analogues, primitive defenses ensure at least some development around positive experiences but interfere with neutralization of aggression and fear. These "pregenital" emotions and their related impulses (or a chaotic combination of pregenital and genital ones) persist as dominant motivators of behavior and significantly interfere with normal personality development.

The Borderline Core of Personality Disorders. Kernberg describes the borderline personality organization as follows:

1. Chronic free-floating anxiety
2. Polysymptomatic neurosis (e.g., multiple phobias related to one's body or appearance, bizarre conversion symptoms, dissociative reactions, hypochondriasis)
3. Polymorphous perverse sexual trends (e.g., coexistence of genital and pregenital elements, bizarre forms of perversions involving aggression)
4. Poor impulse control and addictions
5. Shift toward primary process thinking ("magical" thinking, primitive fantasies)
6. Partially impaired reality testing

Note, however, that the aforementioned symptoms are not unique to individuals with borderline personality but are shared by other subtypes of PD. Reich and Frances have shown that "the diagnosis of borderline personality organization is essentially equivalent to the diagnosis of PD (any subtype). Likewise, a hierarchical cluster analysis of identical and fraternal twins, their siblings and parents, showed that 12 DSM PD subtypes can be grouped into eight heritable syndromes, which can be further aggregated into one group closely corresponding to the above-described borderline personality organization.

Perry and Vaillant wrote: "Just as fever, pus, and callus formation are the body's physiological reactions to insults of disease, personality disorders reflects persons' efforts to heal themselves." Indeed, a PD reflects a person's attempt to overcome its underlying fragility through maladaptive but purposeful behaviors. These maladaptive behaviors are referred to as "purposeful" because they compensate for the fragility and proneness to decompensation associated with the borderline nucleus. Persons with PD become less fragile because the personality deviation (e.g., narcissism) takes over as the principal motivator of behaviors but more abnormal because an already immature personality is further impaired by additional

distortions of self- and object representations.

These compensatory behaviors can have different types of symptomatic setup (e.g., avoidant, paranoid, antisocial, schizoid, dependent). Note that compensatory behavior styles help to classify a syndrome regarding its descriptive label (e.g., avoidant, antisocial) but tell little about the real character of the disorder underlying the descriptive label. Compensatory behavior styles develop as a function of both internal factors (e.g., fearfulness causes avoidant symptoms) and external factors (e.g., social class shapes antisocial behaviors). In other words, depending on certain constitutional and environmental factors, persons with borderline level of personality organization "create" the clinical picture of their PD.

A subgroup of borderline patients manifest the core symptoms, for example, stormy affects, vague identity, unstable relationships, in a rather stable way ("stable instability"). These patients are fragile and prone to fragmentation but do not develop any compensatory behavior facade to protect their fragility. Some of them can periodically "mimic" behaviors of any PD subtype (avoidant, antisocial, obsessive) as a temporary, "chameleon-like" solution to their chronic internal problems.

The described compensatory process generates a two-level personality structure with the dominant compensatory self-concept and the split-off real self-concept. Consequently, clinical expression of PD is dominated by overt, prototypical behaviors, but a set of subtle and covert symptoms and signs associated with the underlying borderline nucleus are regularly observed in the clinical picture. These difficult-to-detect symptoms are overlooked in the DSM system as diagnostic criteria of PD. For example, clinical presentation of narcissistic personality is dominated by the grandiose self, a deviant self-concept that generates an unrealistic sense of one's specialness, entitlement, arrogance, and other prototypical narcissistic behaviors (included in the DSM description of this disorder). The hidden real self of a narcissistic person is recognizable through chronic feelings of inferiority, hypochondriasis, envy, and pessimism that are not included in the DSM description of this disorder, even though they regularly accompany and sometimes even dominate clinical presentation of pathological narcissism. Svrakic has established similar "overt" and "covert" clinical features for other PDs as well.

Some Conceptual and Methodological Issues. The borderline level of personality organization can be reliably distinguished from the schizophrenic spectrum disorders but less reliably from affective syndromes. The same immature personality processes that increase the susceptibility to deviant behaviors may also increase the susceptibility to mood disorders. In many cases, mood disorders and PD are comorbid and can be interwoven to the extent that no meaningful distinction between them can be made (as implied in the concepts of depressive personality,

hysteroid dysphoria, or characterological depression).

An increasing number of psychometric tests to improve the reliable measurement of psychodynamic constructs are available (e.g., the Defense Style Questionnaire, the aforementioned Kernberg's Structural Interview). A satisfactory diagnostic stability and discriminant validity in patients diagnosed as "borderline personality organization" has been demonstrated.

Both the psychodynamic model and the psychobiological model define a common denominator ("the core feature"), which extends across discrete subtypes of personality syndromes. This core feature is the borderline personality organization in the psychodynamic model, and poorly developed character traits, mostly Self-Directedness, in the psychobiological model. Symptomatically, low Self-Directedness and borderline personality organization are for all practical purposes identical. For example, individuals with low Self-Directedness are described as blaming, helpless, irresponsible, unreliable, with no internal direction, and with difficulties in defining and pursuing stable goals (rather, they tend to experience numerous short-term, frequently mutually exclusive motives, none of which can develop to the point of real personal significance). It has been shown that low Self-Directedness correlates highly with the use of immature defense mechanisms that are central in the psychodynamic concept of borderline personality. Etiologically, both concepts postulate that excessive negative emotionality, such as fear and anger (or high Harm Avoidance and high Novelty Seeking), interferes with developmental processes and increases the risk of PD. Clearly, the psychodynamic concepts can be incorporated into a comprehensive, psychobiological understanding of PD.

Psychobiological Summary of Personality Disorder

Psychodynamic understanding of PD, as shown previously, was a useful first step in the understanding of human behaviors. However, the psychodynamic model neither accounts for nor provides directions for expanding the current understanding of the full etiological complexity of PD. In other words, as an etiological theory, the psychodynamic model, addresses only a portion of the iceberg representing deviant behaviors. However, as a treatment tool, its unique understanding of emotional and cognitive processes behind deviant behaviors is effectively used in the therapy of PD, especially in reorganizing internalized concepts about the self and the external world. Psychodynamic strategies are an integral part of the comprehensive psychobiological approach to the treatment of PD presented in this chapter.

Modern psychobiology of behaviors provides an integrative, multifactorial, and developmental etiological and developmental model of personality and its disorders. Psychobiological approach is based on four fundamental pathogenetic perspectives: genetic, neurophysiological,

learning, and phenomenological that interact in a nonlinear way to produce phenotypic differences in cognitive styles and behavior traits. Psychobiology of temperament and character traits is discussed throughout this chapter (see Tables 26–2 through 26–10 for review). The adaptive self-organization of character is discussed in the section on development. Diagnostic and treatment implications are discussed in the corresponding sections. The following section summarizes the psychobiological concept of PD.

Multiple genetic and environmental factors interact in complex ways to influence the risk of PD. Available genetic observations about twins, adoptees, and families are explained by the hypothesis that there is quantitative inheritance of underlying personality dimensions that influence the risk of PD, rather than separate inheritance of individual PD subtypes. As a reminder, more than one half of the variance in the four major temperament traits is inherited. GWAS have now identified the specific combinations of genes that account for this heritability. It is the interactions of multiple genes with one another and with environmental variables, not the average effects, that account for variability in personality. These temperament traits determine one's susceptibility to specific neurochemical processes that in turn underlie individual differences in basic emotions and influence early learning characteristics. These temperament traits have their dissociable genetic, neurobiological, and phenomenological correlates (Tables 26–2 through 26–6).

Table 26–22.
Relative Risk of Personality Disorder as a Function of Temperament Type in a Sample from the General Community

Temperament Type	Configuration	N	Immature (%)
High risk			
Borderline	NHr	39	72
Obsessional	nHr	44	59
Antisocial	Nhr	25	48
Passive–aggressive	NHR	30	40
Average risk			
Average profile	—	15	33
Low risk			
Avoidant	nHR	30	17
Schizoid	nhr	31	16
Histrionic	NhR	50	12
Reliable	nhR	36	6
Total		300	33

Lower case letters (h, n, r) and capital letters (H, N, R) indicate low and high values for temperament traits.

The antecedent temperament factors, along with systematic cultural biases and random life events, critically influence character development, represented as internalized attitudes and concepts about the self and the external world. As shown in Table 26–22, various temperament types affect differentially one's risk of immature character and PD. Some configurations, mostly those with high Reward Dependence, are protective against PD, whereas some increase this risk (e.g., the explosive or borderline profile with low Reward Dependence, and high Novelty Seeking and Harm Avoidance). Contrary to the common belief, average scores on the temperament dimensions do not protect against maladaptation and immaturity. People with average temperament traits have an average (not a decreased) risk of PD (Table 26–22).

Extreme temperament variants are not necessarily indicative of PD. They are expected to be associated with long-term personal, social, and/or occupational impairments that warrant the PD diagnosis *only when accompanied by low character traits*. In other words, poorly developed character is what makes some behavior traits maladaptive and increases the risk of PD. An individual high in Novelty Seeking and low in Harm Avoidance may have an impulsive PD if he or she is low in Self-Directedness and Cooperativeness, or may be an inquisitive scientist without PD if he or she is self-directed and cooperative. Mature character traits (i.e., mature concepts about one's self and the external world) optimize adaptation of temperament (i.e., basic emotionality) to the environment by reducing discrepancies between one's emotional needs and norm-favoring social pressures. In PD, immature character traits and extreme temperament configurations mutually perpetuate each other.

The concepts of temperament and character are essential to decompose the symptoms of PD into the *common features,* shared by all PDs, and *distinguishing features* that permit the discrimination among PD clusters and individual subtypes. This is not possible for models that confound temperament and character. This also indicates that PD reflects deviations in both temperament and character, that is, in both emotional and conceptual personality processes.

Extreme temperament traits and their associated primary emotions of fear, anger, and detachment are motivationally monopolistic, that is, they persist as dominant motivators of behavior and interfere with character development and maturation. In some cases, personality stabilizes at this level of PD (this is observed as "maladaptation"). Personality disorder represents a point of developmental stability for many individuals, as it involves maladaptation, that is, deviant or poor adaptation (but still some adaptation) to the local environment. In other words, PD represents a suboptimal developmental outcome, with established, although maladaptive, personal and social roles, support network, and so forth. These suboptimal outcomes, which are clearly maladaptive relative to the

possible maximum for that individual, tend to be stable (to resist change). The stability of outcomes that are not optimal (i.e., they are maladaptive relative to the global adaptive optimum possible for that individual) occurs because each step in personality development must increase the adaptive fitness of the organism so that it is more successful in balancing multiple internal and external constraints. In other words, search for higher adaptive maxima is discouraged by the necessity of initially decreasing in overall fitness (the earlier mentioned U-shaped developmental pattern).

Behavior genetic research has demonstrated the importance of gene–environment interactions in personality development. For example, prospective studies of birth cohorts in Finland have demonstrated that specific genetic variants of the dopamine DRD2 and DRD4 receptor influence the development of Novelty Seeking but only in individuals exposed to harsh and inconsistent discipline by emotionally distant parents. Behavior geneticists have demonstrated that the effect of sociocultural factors on personality appears less specific than that of genetic factors, influencing success in adaptation rather than its form or personality style. Hence, psychosocial disorganization in the rearing environment of a child has a substantial influence on the risk of PDs. This is essential for preventive strategies as even temperament configurations with high risk of PD might be overcome in homes and communities that provide security and limits on behavior in a warm, compassionate manner, as well as encouraging self-directed choice and the value of respect for other people.

Personality disorder is associated with younger age. This indicates maturation (i.e., remission of deviant behaviors) with increasing age. In general, three dimensions of personality change substantially with age. Novelty Seeking decreases with age by approximately 18 percent, so that older individuals become less impulsive (more reflective), less rule-breaking (more orderly), and less quick-tempered (more stoical). Cooperativeness increases markedly in most children during school age and then increases by 12 percent on average after the age of 18 years. Self-Directedness increases markedly in most people during adolescence and young adulthood, increasing on average by 9 percent after the age of 18 years. Prospective studies with four measurements from the ages of 20 years to 45 years show that Self-Directedness and Cooperativeness increase substantially with increasing age whereas Self-Transcendence decreases between the ages of 20 to 45 years before increasing again as people must face ultimate situations like their own mortality.

The decreasing prevalence of PD with age is attributable to the increased development of both Self-Directedness and Cooperativeness with age. The additional tendency for Novelty Seeking to decrease with age explains the finding that patients with impulsive PDs show more improvement than those with anxious or eccentric PDs. The best

documented finding about change in deviant behaviors is the remission of criminal behavior in individuals with Antisocial PD. These individuals nearly always remained impulsive (high Novelty Seeking), risk-taking (low Harm Avoidance), and aloof (low Reward Dependence) but became mature enough to maintain work and family life in a stable manner.

TREATMENT

Stages in the Development of Self-Awareness and Well-Being

A full assessment of personality requires consideration of a person's level of self-awareness and well-being, not just his or her impairments. Health and well-being are more than the absence of deviant traits. Well-being depends on a person's level of self-awareness and leads to the expression of human virtues and positive emotions that go beyond what is average in contemporary society.

There are three major stages of self-awareness along the path to well-being, as summarized in Table 26–23, based on extensive work by many people. The absence of self-awareness occurs in severe PDs and psychoses in which there is little or no insightful awareness of the preverbal outlook or beliefs and interpretations that automatically lead to emotional drives and actions. Lacking self-awareness, people act on their immediate likes and dislikes, which is usually described as an immature or "childlike" ego state.

The first stage of self-awareness is typical of most adults most of the time. Ordinary adult cognition involves a capacity to delay gratification in order to attain personal goals but remains egocentric and defensive. Ordinary adult cognition is associated with frequent distress when attachments and desires are frustrated. Hence, the average person can function well under good conditions but may frequently experience problems under stress. At this stage of self-awareness, a person is able to make a choice to relax and let go of his or her negative emotions, thereby setting the stage for acceptance of reality and movement to higher stages of coherent understanding.

The second stage of self-aware consciousness is typical of adults when they operate like a "good parent." Good parents are allocentric in perspective—that is, they are "other-centered" and capable of calmly considering the perspective and needs of their children and other people in a balanced way that leads to satisfaction and harmony. This state is experienced when a person is able to observe his or her own subconscious thoughts and consider the thought processes of others in a similar way to observing his or her own thoughts. Hence, the second stage is described as "metacognitive" awareness, mindfulness, or "mentalizing." The ability of the mind to observe itself allows for more flexibility in action by reducing dichotomous thinking. At this stage, a person is able to observe himself and

others for understanding, without judging or blaming. However, in a mindful state, people still experience the emotions that emerge from a dualistic perspective, so mindfulness is only moderately effective in improving well-being.

Table 26–23.
Three Stages of Self-Awareness on the Path to Well-Being

Stage	Description	Psychological Characteristics
0	Unaware	Immature, seeking immediate gratification ("childlike" ego state)
1	Average adult cognition	Purposeful but egocentric; able to delay gratification but has frequent negative emotions (anxiety, anger, disgust) ("adult" ego state)
2	Metacognition	Mature and allocentric; aware of own subconscious thinking; calm and patient, so able to supervise conflicts and relationships ("parental" ego state, "mindfulness")
3	Contemplation	Effortless calm; impartial awareness; wise, creative, and loving; able to access what was previously unconscious as needed without effort or distress ("state of well-being," "soulfulness")

Adapted from Cloninger CR. *Feeling Good: The Science of Well-Being.* New York: Oxford University Press; 2004.

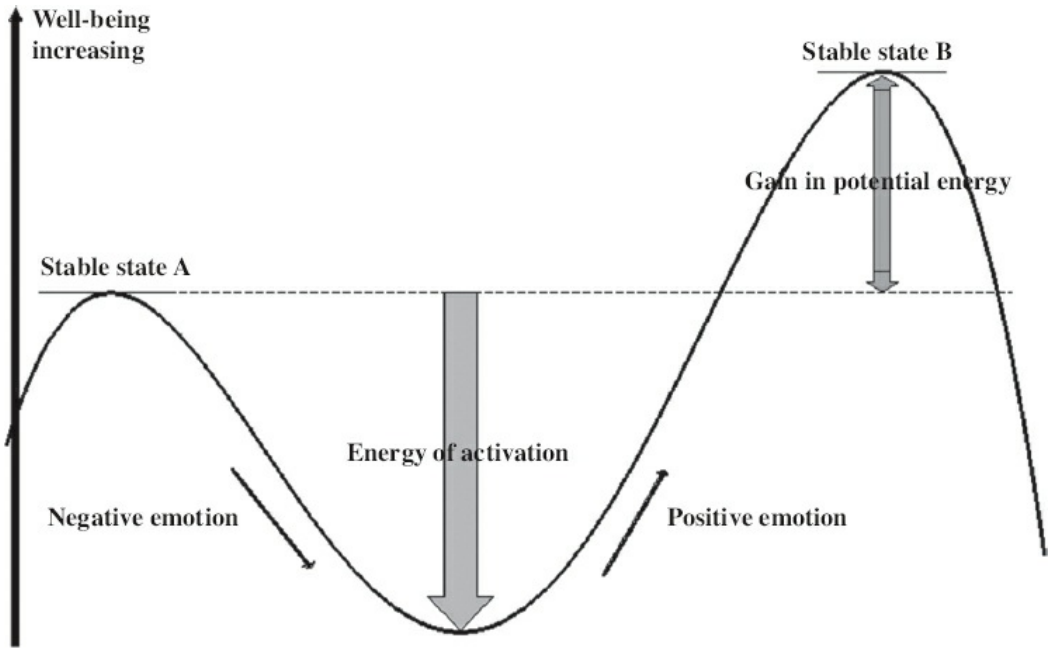

FIGURE 26–11. The emotional dynamics of well-being as successive transitions in a complex adaptive system. (Used with permission of Center for Well-Being at Washington University.)

The third stage of self-awareness is called contemplation because it is direct perception of one's initial perspective—that is, the preverbal outlook or schemas that direct one's attention and provide the frame that organizes one's expectations, attitudes, and interpretation of events. Direct awareness of one's outlook allows the enlarging of consciousness by accessing

previously unconscious material, thereby letting go of wishful thinking and the impartial questioning of basic assumptions and core beliefs about life, such as "I am helpless," "I am unlovable," or "faith is an illusion." In other words, contemplation is reversal of the process of repression: contemplation enlarges consciousness whereas repression reduces it. The third stage of self-awareness can be described as contemplative because in this state, a person becomes aware of deep preverbal feelings that emerge spontaneously from a unitive perspective, such as hope, compassion, and reverence. Contemplation is much more powerful in transforming personality than is mindfulness, which often fails to reduce feelings of hopelessness.

Extensive empirical work has shown that movement through these stages of development can be described and quantified in terms of steps in character development or psychosocial development, as in the work of Vaillant on Erikson's stages of ego development. As previously mentioned, such development can be visualized as a spiral of expanding height, width, and depth as a person matures or increases in coherence of personality. Likewise, the movement of thought from week to week or month to month has the same spiral form regardless of the timescale. Such "self-similarity" in form regardless of timescale is a property characteristic of complex adaptive systems, which are typical of living processes in general, whether they are psychosocial or neurobiological. The clinical utility of this property is that therapists can teach people to exercise their capacity for self-awareness, moving through each of the stages of awareness just described. Their ability to do so, and the difficulties they have, reveals the way they are able to face challenges in life. Cloninger (2004) has developed an exercise, called the "Silence of the Mind" meditation, with explicit instructions to take people thorough each of the stages of awareness as well as they can. The first phase of this meditation results in a relaxed state in the first stage of self-awareness. The second phase facilitates entry into the second stage of self-awareness, and the third phase into the third stage of self-awareness, if the person is able to do so. Using this and a way of observing thought during mental status examination, mental health professionals can assess a person's thought and its level of coherence in a way that is constructive, easy, and precise without being judgmental.

Each step of growth in self-awareness and well-being requires particular psychobiological conditions. The emotional dynamics of growth in well-being is depicted in Figure 26–11 as a transition within a nonlinear dynamical system. The transition from one state of lower well-being (state A) to a state of higher well-being (state B) requires growth in self-awareness. Each such transition requires a person to accept the limitations of the assumptions he or she has been making about what is satisfying and valuable in order to begin adapting in ways that are more coherent and flexible. However, such growth requires a person to let go of what is

familiar that so there is a temporary increase in negative emotions before there is further growth in positive emotions and other aspects of well-being. Therefore, an effective clinician must help provide appropriate conditions to facilitate this transition ("the energy of activation") with his or her own hope, patience, and awareness, realizing that the patient must ultimately be the one who does the work to grow in self-understanding. Each step along the pathway to well-being requires different procedures to facilitate growth because different aspects of the being are involved in the different steps in character development. These specific techniques for the various steps in development are described in what follows.

General Principles of Treating Personality Disorder

Individuals with PD do not recognize that they are ill and seldom seek help unless others people (such as a spouse or parents) are insistent. This usually happens when maladaptive behaviors create marital, family, and career problems, or when other mental symptoms (e.g., anxiety, depression, substance abuse), or somatic symptoms (e.g., obesity, heart disease, COPD), complicate their clinical picture. In general, patients with PD require a multifaceted treatment plan that often combines psychotherapy and pharmacotherapy.

Bidirectional Interaction of Psychotherapy and Pharmacotherapy

Psychotherapy and pharmacotherapy are almost invariably combined in the treatment of individuals with PD. Psychotherapy is the primary treatment, as it promotes the maturation of character and ultimately the development of the patient's capacity to develop better adaptive solutions. However, psychotherapy is very hard to implement in the setting of unstable affects and risky or self-destructive behaviors, all very common in the initial stages of treatment. During these early stages, pharmacotherapy achieves a relatively prompt control of affects and behaviors. Most importantly, this nonspecific improvement sets a more suitable platform for the work in psychotherapy.

Psychotherapy and pharmacotherapy have a bidirectional relationship, as both modalities, in addition to their independent effects, also interact in a positive feedback manner. As the first step, pharmacotherapy stabilizes affects and ensures safety of the patient, which interferes with psychotherapy. Psychotherapy promotes maturation of internalized object relations and defense mechanisms, ultimately leading to the development of integrated, less fragmented, and more realistic perceptions of self and objects. In turn, this achievement tones down excessive behaviors and acting outs associated with primary emotions, primarily anger and fear. The latter are the main etiological cause for PD and the main reason for the persistence of the underlying immature character later in life. Toward the

end of treatment, it is expected that the patient will develop more mature object relations, stable identity, and a more realistic perception of self and the surrounding world. These are necessary in order to develop better adaptive solutions, with the stable sense of internal direction, sustained motivations, and prosocial attitudes and behaviors. Such successful outcomes are not a wishful utopia, as they are much more frequent than commonly believed.

Primary Pathology and Secondary Symptoms of Personality Disorders

Common assumption is that pharmacotherapy "works" only at the biological level and psychotherapy works only at the psychological level. However, psychotherapy can modulate structural CNS characteristics, such as receptor activity and expression in brain tissue through epigenetic mechanisms. On the contrary, pharmacological improvement of depressed mood or symptoms of psychosis can restore complex psychological phenomena, such as integration of identity, in Major Depression or Postpartum psychosis. As a brief aside, it is ironic that after years of downplaying its efficacy, the strongest support for psychotherapy comes from basic neurosciences.

However, effects of psychotherapy and pharmacotherapy are not completely interchangeable. In comparison to psychotherapy, medication is more effective in treating mood or modulating aggression then in changing internalized concepts about self and the world. Unstable mood, intensive but unstable interpersonal relationships, or risky and self-destructive behavior is clinically impressive but nonspecific for PDs. These and similar symptoms are secondary, that is, they represent phenotypic expressions of the underlying immature personality structure, such as identity diffusion and lack of sublimatory channels, which is the primary pathology of PD. Modern pharmacotherapy has not achieved the level to be able to dissect precise molecular substrates of these complex subjective phenomena. Most of them are the so called "emerging properties," that is, features that unpredictably emerge from combinations of lower-order components. Biological dissection of such properties may not be possible even with most advanced technologies, and psychotherapy may remain as the only approach. Psychotherapy approaches its target problems "top-to-bottom" (from phenotype to biogenetics) and hence can correct complex phenomena such as identity diffusion or immature character without having to know their precise neuromodular structure. In contrast, in order to be specific, pharmacotherapy would have to approach its target symptoms "bottom-to-top," that is, from molecular mechanisms to complex subjective phenomena. Such highly specific pharmacotherapy of subtle subjective experiences may well be an unattainable utopia. For many other psychiatric disorders, however, specific pharmacotherapy is a realistic ideal. Recent advances in the current understanding of the

phenotypic–genotypic architecture of the group of Schizophrenias may provide important guidelines for the development of more specific medications.

The Role of Antipsychotics in the Treatment of Personality Disorder

Medications commonly used to treat psychosis are called "antipsychotics" or "neuroleptics." This group of drugs is effective in treating the so-called "positive symptoms" of most forms of psychosis, such as delusions and hallucinations. Their main indication is most forms of psychosis and they are very rarely, if ever, used to treat nonpsychotic symptoms. This mostly reflects the fact that antipsychotics administered to nonpsychotic individuals have major sedative effects and a number of other undesired and/or lifelong or irreversible adverse effects.

During the last three decades, a number of "second-generation" antipsychotics have been developed. These are called "atypical" because of a much more favorable side effect profile and less sedation. Many authors prefer the broader term "psychotropics" for these medications to accentuate that in addition to antipsychotic, these medications also have antidepressant, antimanic, mood stabilizing, and anxiolytic effects. Indeed, Food and Drug Administration–approved indications for the group of psychotropics include psychosis, bipolar disorder, and depression (either as monotherapy or as augmenting agents). Clozapine has been for decades considered to be a "heavy gun" used exclusively for the group of Schizophrenias, with major side effects and potentially lethal adverse effects. An increasing number of clinicians are beginning to capitalize on clozapine's efficacy in atypical depression, refractory depression, mood stabilization, or debilitating chronic anxiety. Recent reports indicate that in addition to its serotonergic–dopaminergic antagonism, clozapine has a number of other effects. One of these is histone deacetylase inhibition, an epigenetic mechanism that "unlocks" *GABA* genes in the frontal cortex, leading to increased production of GABA and electrophysiological stabilization of cortical neurons. Another example is that clozapine stimulates production of brain neurosteroids, an effect that may explain its efficacy not only in psychosis but also in chronic anxiety and vulnerability to stress. With the above in mind, clozapine is increasingly used to treat eating disorders (anorexia), treatment refractory depressions, bipolar disorder, and so forth, Similarly, aripiprazole, originally marketed as an antipsychotic, is far more frequently used as antidepressant (either as monotherapy or for augmentation). In fact, many consider long-acting injectable aripiprazole to be the first long-acting injectable antidepressant ever.

Preventable Errors in Treating Individuals with Personality Disorders

There are three major barriers to effective treatment of PD, but all are preventable errors within the control of the health care provider. The first is the frequent loss of professional objectivity, signaled by the development of strong emotions (positive or negative) and called positive or negative countertransference. This inappropriate personal involvement is a red flag to reassess diagnosis and treatment. Frequent discussions and counseling with colleagues are useful because even strong countertransference feelings sometimes persist unrecognized.

The second preventable error in PD management is to believe the myth that PDs cannot be treated effectively. This myth is partly generated by countertransference problems of some professionals and then sustained by a failure to consider facts showing the effectiveness of treatment. Belief in the untreatability of a patient creates a self-fulfilling prophecy. For example, many controlled studies indicate that even severe PDs, such as borderline or antisocial PDs, can be effectively treated with an appropriate condition, such as a cooperative therapeutic alliance.

The third preventable error in PD management is to give direct advice on personal and social problems. This is counterproductive in patients with PD because they usually become dependent, noncompliant, or resentful. Occasionally, direct advice may be offered to some antisocial, narcissistic, and schizoid patients who are at low risk of developing dependency and need precise structure and direction initially. However, it is most beneficial to provide guidance and support without giving direct advice. All that are usually required in supportive therapy are compassionate attention, respect for the dignity of other human being despite his or her flaws, and reinforcement of the patient's existing coping mechanisms that are adaptive.

When tempted to give direct advice to patients, remember that change in personality requires more than common sense and logic. If the relationship leads to frequent advice giving, then referral to a psychiatrist or a psychologist may be indicated. People change if they become self-aware, usually from personal recognition of dissatisfaction with themselves and their relationships. Personal growth thus arises from new insights about oneself and the environment. Direct advice robs the patient of the opportunity to develop new insights and to learn from his or her mistakes. In summary, supporting these patients involves joint evaluation of options and encouragement to practice skills in solving problems.

Substantial personality change, which is invariably needed by people with PDs, involves an extensive reorganization of internalized concepts and coping mechanisms and thus requires precise diagnostic analyses, specific treatment strategies, and expert training. The expert treatment may include any of the several available psychotherapy approaches and is usually combined with pharmacotherapy. The major points relevant to psychotherapy and pharmacotherapy of PDs are summarized later.

As already mentioned, individuals with PD have a peculiar capacity to elicit strong emotions from other people. They are often described as aggravating, unlikable, difficult, or bad. Alternatively, they may be seductive or dependent and elicit inappropriate emotions or actions, such as sexual interest or the urge to rescue. Even professionals may have difficulty treating them with respectful objectivity because of a blurring of personal boundaries. Such loss of objectivity occurs because the patient's deeply felt assumptions about other people may often elicit interpersonal responses that are appropriate to the patient's assumptions. People's assumptions about themselves and others often become self-fulfilling prophecies because of automatic mechanisms of affect transfer. If someone smiles at you, communicating appreciation, it is natural to experience feelings of social attachment and to smile back automatically. Likewise, if someone frowns, communicating anger, it is natural to feel defensive in preparation for his or her angry attack. For example, many patients with PD are suspicious and hostile about others' motives. This distrustful attitude is communicated in many verbal and nonverbal ways and often elicits disagreement or frank hostility from others. These uncooperative responses reinforce the original negative assumptions of the patient, which in turn leads to further alienation.

This vicious cycle of affect transfer can be interrupted only by professional objectivity combined with patience and compassionate respect for the patients' disability. Such objectivity arises from recognizing the overall meaning and implications of their pattern of interpersonal signals, so that their verbal and nonverbal communication takes on diagnostic and therapeutic, rather than personal, significance. In optimal therapeutic relationships, "patients" should be patiently hopeful and physicians should be compassionately realistic. Whenever professionals become aware of strong positive or negative emotions toward a patient (so-called "countertransference" reactions), this should help alert them to the possibility that the patient has a PD.

As many patients with PD do not recognize or admit their psychopathology, they resist and resent psychiatric diagnoses and any form of mental health treatment. Accordingly, it is prudent to let the patient define his or her treatment goals and then jointly evaluate the likelihood of successful outcome until treatment goals that both patient and the therapist agree upon can be identified. Initially, these goals should be as simple and concrete as possible (e.g., "to develop social skills," or "to reduce alcohol use"). In many, but not all cases, successful completion of this initial phase will motivate the patient to define other, more complex treatment goals and to continue treatment.

A psychiatrist should keep in mind that there is a natural succession of stages in the treatment of patients with PDs. Each has different goals and requires different methods. The complete psychiatrist should be prepared

to guide the patient along these stages, ever ready to advance to the next stage if the patient is interested and prepared to do so.

Four Major Stages in Treatment of Personality Disorders

The four stages in the treatment of a patient with PD can be described as (1) crisis management and stabilization, (2) awakening of a positive perspective and personal values in life, (3) other-centered awareness, and (4) integrated intelligence. The initial stage of crisis management and stabilization deals with the presenting problem and stressors in order to help the patient get into a calm enough state and a working alliance with the psychiatrist. The second stage involves elevating a person's outlook on life so that he or she can experience things he or she enjoys and values under relaxed conditions. This involves a spiritual awakening that has often been neglected in strictly cognitive-behavioral or psychodynamic approaches but without which there is little capacity for fundamental change in the quality of life. The third stage of metacognitive (other-centered) awareness involves increases in self-awareness and capacity for contemplation that elevate a person's usual thoughts, feelings, and relationships in a wide range of conditions. The fourth stage of integration of reason and love in action allows a person to be mature and happy even under conditions that were previously stressful. Patients with PDs can pass through these stages on their own (i.e., remit spontaneously) or be guided through these stages in treatment facilitated by a scientifically designed set of physical, personal, social, cognitive, and spiritual exercises.

Stabilization Phase. What is done in the first stage of treatment depends greatly on individual patient and his or her presenting situation. This initial stage may involve stabilization of the patient with medications if they are indicated and the patient is interested in such treatment. Medications are often helpful, but not everyone wants such treatment because they always carry some risk of side effects. The advantages and disadvantages must be carefully weighed to respect the patient's wishes and to help him or her be calm and organized enough for further growth in self-awareness.

A useful approach to the initial stage of treatment is to focus primarily on the chief complaints that bring them to treatment. These complaints are often related to a person's work, relationships, and/or general health, which often offer an adequate basis for the diagnosis and initial treatment of PD. For example, a systematic focus on a person's work, school performance, or relationships can identify attitudes, feelings, and behaviors that help him or her recognize the barriers to his or her success and happiness. However, this can bring up many sensitive issues that they prefer to minimize or ignore or that lead them to lie and distort information about their real situation. An alternative initial approach,

particularly if they are entering treatment at the request or demand of someone else, is to focus on healthy lifestyle choices, which provides a nonthreatening basis for the evaluation of a patient's goals, values, habits, and skills (i.e., his or her personality). Choices about diet, weight control, exercise, smoking, drinking, and ways of relaxing and managing stress are appropriate for discussion with one's physician and do not threaten or stigmatize the patient. Discussion of these choices with a patient can provide a guiding stimulus for developing more self-direction and constructive planning about life. Discipline in working toward chosen goals is an indicator of maturity. Success in accomplishing even simple goals builds self-confidence and often generalizes to other forms of self-directedness. Lack of success stimulates learning about goal setting and personal growth by, for example, breaking a problem into smaller steps to be taken one at a time. In this process, patients have the opportunity to learn from experience, which is essential for patients with PDs. In many cases, such goal setting and problem solving leads to the ability to admit faults and to recognize one's strengths and limitations. Respect for one's self from accomplishment requires acceptance of responsibility and leads to trust of others. Self-respect and respect for others usually progress hand in hand. Such patient guidance is a simplified and nonthreatening form of what is usually called cognitive-behavioral therapy (CBT) and can be safely practiced in a busy office practice with short but regular sessions.

CHOICE OF MEDICATIONS FOR STABILIZATION. During the initial stage of treating patients with PDs, medications are often used to target specific symptoms of their disorders with the goals of relieving subjective distress, risky or self-destructive behaviors, and/or conflict with others, thereby preparing them for later stages of treatment that require calmness, safety, and nondefensiveness to facilitate growth in self-awareness.

One should keep in mind that the vast majority of affective symptoms of PDs (such as dysthymia, unstable affects among others), their chronic anxiety, and most of their behavior symptoms (e.g., impulsivity) are shared by all classified clinical subtypes. Hence, symptomatic pharmacotherapy is not focused on individual subtypes of PD but rather on the following four domains shared by all subtypes: (i) mood and anxiety dysregulation, related most strongly to Harm Avoidance, (ii) aggression and impulse control, related most strongly to Novelty Seeking, (iii) social and emotional detachment, related most strongly to Reward Dependence, and (iv) psychotic symptoms and cognitive distortions, related most strongly to intellectual reasoning and Persistence. These four correspond roughly to symptoms most strongly related to the four temperament dimensions. Hence, modern pharmacotherapy of the aforementioned four domains shared by all subtypes of PDs brings closer the paradigms of causal and symptomatic pharmacotherapy. Recommendations about

pharmacotherapy of symptom domains shared by most PDs are summarized in Table 26–24.

(i) Symptom domain of mood dysregulation and anxiety

This symptom domain includes mood instability, typical and atypical depression and/or dysphoria (dysphoria and dysthymia are used here as synonymous terms—Latin and Greek for low mood), and chronic cognitive and somatic anxiety.

Emotional instability and mood swings are usually responsive to lithium or valproates (for patients with frequent episodes of euphoria) or lamotrigine (for patients with more frequent depressive episodes). Low-dose atypical psychotropics may be used to stabilize mood or improve depression as the second-line medications. Tricyclic antidepressants (TCAs), like imipramine, sometimes increase impulsivity and anger in emotionally unstable patients (e.g., borderline, narcissistic, histrionic, dependent). This paradoxical effect has been hypothesized to reflect cognitive-behavioral disorganization caused by the so-called "catecholamine stress," that is, a dramatic increase in catecholamine levels during TCA treatment. As a general rule, antidepressants may cause an episode of mania if not combined with mood stabilizers in persons with emotional lability or cyclothymia. TCAs are extremely dangerous in an overdose, so these drugs ought to be used with caution in patients with PD.

Atypical depression and dysphoria are very frequently observed in patients with any subtype of PD. The "first-line" medications used here are selective serotonin reuptake inhibitors (SSRIs), serotonin norepinephrine reuptake inhibitors (SNRIs), monoamine oxidase inhibitors (MAOIs), or low-dose atypical psychotropics, such as aripiprazole, ziprasidone, lurasidone, or clozapine. TCAs are not recommended as at least half of the PD subjects suffering from atypical depression worsen on TCAs. Again, antipsychotics are used only after careful consideration of the risk–benefit ratio in voluntary patients.

Typical depression, including Major Depression may complicate any PD. These syndromes are treated with antidepressants, including heterocyclics, in doses suggested for primary major depression.

Patients with PD often present with both cognitive anxiety (anticipatory worry) and somatic anxiety (concerns about bodily pains and psychophysiological reactions). Treatment of choice for chronic cognitive anxiety is psychotherapy, especially various forms of CBT. With respect to pharmacotherapy, chronic cognitive anxiety is most responsive to SSRIs, MAOIs, SNRIs, GABA analogues (such as valproates, pregabalin, gabapentin), and (with caution) long half-life benzodiazepines. Chronic somatic anxiety is more responsive to MAOIs, SNRIs (e.g., venlafaxine or duloxetine), and buspirone. Low doses of TCAs are very effective for somatic anxiety in some patients, but MAOIs are more often effective if the

required dietary regimen can be followed or patch systems (transdermal selegiline) are used. Avoidant traits can be also effectively treated with either SSRIs or MAOIs. Some components of somatic anxiety, such as sweating, palpitations, diarrhea, and tremor, can be treated with beta-blockers. Severe, psychotic-like anxiety responds to low-dose neuroleptics, especially drugs with low D2 affinity (e.g., quetiapine). Despite relative safety of novel atypical drugs, caution about prolonged use is necessary, as noted earlier, because there is still substantial risk of serious side effects (e.g., for risperidone, risks per person per year are 0.6 percent for tardive dyskinesia, 0.35 percent for sudden cardiac death, and 0.02 to 2.44 percent for neuroleptic malignant syndrome). Therefore, the risk of some serious and irreversible side effects is about 1 percent per person per year and requires careful consideration of risk–benefit ratio with voluntary informed consent.

Recent reports indicate that a number of medications used to treat depression and anxiety, for example, SSRIs, stimulate neurogenesis in the hippocampus and thereby restore hippocampal control of the HPA axis. This mechanism is believed to underlie reduced vulnerability to stress in susceptible individuals treated with antidepressants.

(ii) Aggression

It is useful, though sometimes difficult, to distinguish different types of aggression. However, diagnosis of different types of aggressive behaviors has important diagnostic implications as well as crucial guidelines for pharmacotherapy. The most common form of aggression occurs when a quick-tempered person is provoked by frustration or threats. Here, the very act of aggression does not have a discernible long-term goal or secondary gain. This is often called “affective aggression” and is frequent in Cluster B impulsive–aggressive individuals who are high in novelty seeking and low in harm avoidance. Aggression that appears to be unprovoked sometimes occurs in patients with cerebral instability documented by an abnormal electroencephalogram (EEG) and is often called “ictal aggression” regardless of any associated personality traits. Predatory aggression or “cruelty” is also known as “organized aggression” because it involves planning, social motives, and/or secondary gains. This type of aggression involves hostile revengefulness and taking pleasure in victimizing others, usually with intact impulse control. Such predatory aggression is most frequent in individuals who are very low in cooperativeness and who are also emotionally dissociated and disengaged, such as Antisocial and Schizoid PD. Finally, “organic-like” aggression is often accompanied by poor social judgment and disinhibition; it is best distinguished from other impulsive–aggressive syndromes by prominent distractibility, inattention, and emotional lability, all characteristic of patients with frontal lobe lesions. These patients also manifest high somatic anxiety with panic and cardiorespiratory symptoms, muscular tension, and motor restlessness.

Table 26–24.
Choice of Drugs According to Target Symptoms of Personality Disorders

Target Symptom	Drug/Treatment of Choice	Not Recommended
I. Mood dysregulation and anxiety		
Anxiety		
Chronic cognitive	PSYCHOTHERAPY SSRIs, SNRIs, MAOIs LOW-DOSE NOVEL PSYCHOTROPICS (aripiprazole, quetiapine) Valproates and other GABA analogs clonazepam, buspirone	Benzodiazepines and ethanol (risk of abuse/addiction)
Chronic somatic	MAOIs, SNRIs (duloxetine, milnacipran) Pregabalin and other GABA analogs TCAs, beta-blockers	If used—benzodiazepines with long half-life and short trials preferred
Obsessions	SSRIs, PSYCHOTROPICS (quetiapine) TCAs (clomipramine) Mild NMDA antagonists (riluzole, memantine)	
Acute and severe	MIRTAZAPINE, NOVEL PSYCHOTROPICS (quetiapine, aripiprazole, clozapine) TCAs, clonazepam, valproates, lithium	
Depression		
Atypical depression/dysphoria	MAOIs, SSRIs, SNRIs, ARIPIPRAZOLE Lurasidone, ziprasidone, quetiapine	TCAs
Classical depression	STANDARD ANTIDEPRESSANTS TCAs (males) SSRI (females) Atypical psychotropics (as monotherapy or augmentation)	
Emotional lability/rapid cycling	LITHIUM, LAMOTRIGINE, VALPROATES Lower-dose novel psychotropics (olanzapine, aripiprazole, clozapine, ziprasidone)	TCAs (“catecholamine stress”) Standard antidepressants (risk of switching to mania)
II. Behavior dyscontrol		
Aggression/impulsivity		
Affective aggression “Hot temper” with normal EEG	LITHIUM, SSRIs, ANTICONVULSANTS Low-dose novel psychotropics	Benzodiazepines (disinhibition)
Predatory aggression (cold blooded revenge/cruelty)	NO EFFECTIVE PHARMACOLOGICAL Tx Novel psychotropics, lithium, valproates, beta-blockers	BENZODIAZEPINES (disinhibition)
Organic-like aggression (traumatic brain injury)	BETA-BLOCKERS, VALPROATES, QUETIAPINE, CARBAMAZEPINE TCAs, cholinesterase inhibitors (donepezil)	BENZODIAZEPINES (disinhibition, delirium)
Ictal aggression (abnormal EEG)	CBMZ, DIPHENYLHYDANTOIN, VALPROATES Benzodiazepines (clonazepam)	TCAs LOW-POTENCY TYPICALS (both increase risk of seizures)
III. Social and emotional detachment		
Chronic asociality and disinterest	LOW-DOSE PSYCHOTROPICS (aripiprazole, olanzapine, low-dose	

Blunted affect	clozapine, sulpiride)
IV. Cognitive-perceptual distortions/psychotic symptoms	
Acute and brief psychotic episodes	NOVEL PSYCHOTROPICS (Risperdal, olanzapine) Typical neuroleptics (for the duration of psychosis)
Chronic and low-level psychotic-like symptoms	NOVEL PSYCHOTROPICS

Used with permission of the Center for Well-Being at Washington University.

Multiple double-blind trials have shown efficacy of lithium carbonate in the treatment of affective aggression. Lithium salts help impulsive–aggressive individuals to be more reflective, that is, to think about consequences before acting on impulse. To a lesser extent, it may be helpful in reducing cruelty and lack of cooperativeness, but this may be an indirect result of reducing impulsivity, which often is a predisposing influence in the development of hostility and revengefulness. Likewise, low-dose atypical neuroleptics may be useful in setting the stage for the patient to modify old habits and assist him or her in reducing affective aggression. The decisions to use long-term neuroleptics requires consideration of potential side effects, such as tardive dyskinesia, and should be made with carefully informed consent of the patient. Anticonvulsants, such as valproates, lamotrigine, carbamazepine, and oxcarbazepine (to mention only those most frequently used), reduce both the intensity and the frequency of unprovoked angry outbursts in many patients regardless of normality of their EEG. Double-blind trials have shown that psychostimulants and catecholamine agonists, such as methylphenidate, are often beneficial in the treatment of inattentive and hyperactive adults who are impulsive and aggressive, especially when the symptoms have begun in early childhood.

In accord with postulated serotonergic mechanisms in aggressive behaviors, antidepressants (particularly SSRIs and SNRIs) are considered by many to be beneficial for certain subtypes of impulsive PD (e.g., borderline, histrionic). Finally, MAOIs are effective in some dysphoric states with somatic anxiety and hostility.

There is no effective pharmacotherapy for predatory, premeditated, and cold-blooded aggression. One treatment with some chances of success is psychotherapy, usually CBT. Low-dose traditional neuroleptics (haloperidol) or low-dose novel psychotropics (Risperdal, quetiapine) may help tone down predatory aggression. Of note, antipsychotic-level doses have not been shown more effective here and are not recommended. When accompanied by impulsivity, predatory aggression may improve if treated with beta-blockers.

There are some relative contraindications for the aforementioned listed

drugs. Lithium should not be given to antisocial persons without aggression and impulsivity because it does not diminish nonaggressive antisocial behaviors (such as lying, cheating, and stealing) and is poorly tolerated by anxious schizoid individuals. Thus, lithium exposes such patients to risk of neurological, renal, and thyroid toxicity without providing much benefit. Likewise, benzodiazepines and alcohol have disinhibiting effects on violence, reduce conditioned avoidance behavior ("loosen inhibitions"), and further impair passive avoidance learning in impulsive antisocial persons. The use of benzodiazepines seems appropriate only in nonaggressive asocial behaviors, for example, patients with schizoid PD. As a general rule, typical and novel psychotropics are used with extreme caution, and the patient is closely monitored, so that early signs of complications and adverse effects are recognized and prevented, usually by terminating the offending agent.

(iii) Symptom domain of social and emotional detachment

Emotional detachment, cold and aloof emotions, and disinterest in social relations ("chronic asociality") are most frequently observed with schizoid and paranoid PD and, to a lesser extent, with antisocial PD. In some cases, emotional detachment or disinterest may respond to novel psychotropics, such as aripiprazole, risperidone, quetiapine, clozapine, olanzapine, or ziprasidone. These medications may help reduce social withdrawal and other features of "aloof" PDs with less risk of extrapyramidal symptoms than with typical neuroleptics. However, dose adjustment is crucial to maintain compliance because patients with PD often have little tolerance for side effects. In cases in which emotional disinterest reflects an underlying depression, antidepressants (SSRIs or MAOIs) frequently help. One should be cautious with TCAs in schizotypal PD, because they may worsen and/or trigger psychosis.

(iv) Cognitive-perceptual distortions; psychotic symptoms

Cognitive-perceptual distortions observed in patient with PD primarily include mild cognitive symptoms, such as bizarre fantasies, eccentric attitudes, suspiciousness, egocentric perception of reality ("everything revolves around me"), nonpsychotic thought disorder (such as ideas of reference and magical thinking), unusual perceptual experiences (such as illusions), and eccentric and bizarre behaviors, among many others. Dissociative phenomena are also frequently observed, such as déjà vu, depersonalization, and derealization. These and similar symptoms can be occasionally observed with all subtypes of PDs but are more frequent in Cluster A and to a lesser extent in Cluster B PDs. These chronic, low-level, psychosis-like but essentially nonpsychotic symptoms may improve indirectly, after pharmacological stabilization of disturbing affects, and/or after psychotherapeutical correction of cognitive underlying beliefs, and/or after reduction of realistic external stress. For example, a narcissistic person can be "paranoid" that other people have conspired to prevent his or

her success. Psychotherapy is the treatment of choice for this kind of nonpsychotic paranoia. Reflecting the underlying operation of splitting (a primitive defense mechanism that prevents a realistic perception of self and others), the same narcissistic person may at other times perceive these same "enemies" in a diametrically opposite way, as friends or supporters, and even idealize them. In cases in which these symptoms persist and interfere with everyday functioning despite psychotherapy, medications are occasionally needed, usually novel atypical psychotropics (such as aripiprazole, quetiapine).

Some chronic cognitive disturbances, such as mild ideas of reference or suspiciousness, tend to subside when the background emotional tension is reduced. For example, alprazolam has been found to be beneficial in patients with borderline personality, particularly those with a history of drug abuse and suspiciousness. However, long-term use of benzodiazepines is associated with high risk of drug dependence, particularly in patients with PDs, so benzodiazepines like alprazolam should be prescribed only after careful consideration of the risk–benefit ratio and their use carefully monitored for evidence of abuse or dependence.

Chronic cognitive and behavior oddities, such as eccentric attitudes, bizarre behavior, odd beliefs, or extrasensory perceptions are seen with Schizotypal PD. There is substantial empirical evidence that Schizotypal symptoms ought to be classified among other Schizophrenia Spectrum Disorders. Acute, brief reactive psychoses may complicate most subtypes of PD. These are treated symptomatically, according to accepted pharmacological practices, most frequently with second-generation atypical psychotropics due to much better safety and tolerability.

Again, acute psychotic symptoms requiring medication may subside when environmental stressors are brought under control; thus one should be ready to lower the dose or discontinue the medication.

Second stage of personality transformation: Neither medications nor cognitive-behavioral approaches, alone or in combination, are usually adequate to transform a person's personality in a fundamental way. Individuals who radically change their perspective on life usually attribute the change to getting a good job that provides a sense of self-respect, marrying a loving and trusted spouse, or experiencing a religious conversion. These kinds of life experiences change a person's initial perspective on life, which in turn transforms his or her thoughts, feelings, and behavior. Cognitive-behavioral and psychodynamic therapies often leave a patient in a tense inner struggle with himself or herself unless treatment provides experiences that allow a reevaluation of basic assumptions about life. Otherwise, a person cannot transcend the conflicts among his or her emotional drives, so he or she remains in constant or recurrent struggles among parts of himself or herself.

Growth in self-awareness is possible only once a person is calm enough

to face and accept unpleasant realities about himself or herself. As a result, therapy usually proceeds by a series of small steps that lead to progressively more mature and integrated ways of viewing and coping with the challenges of life. Figuratively speaking, a person who feels fearful, hungry, and rejected cannot appreciate beauty and wonders while struggling to gratify basic needs for safety, food, and comfort. However, it is difficult to specify the necessary or sufficient conditions for the progress of any particular person because either internal or external resources can satisfy a person's needs. There are no procedural recipes that assure treatment outcomes because growth in self-understanding depends on the dynamics of nonalgorithmic processes such as insight, judgment, and foresight. What a person desires depends as strongly on his or her worldview as on his or her external circumstances. A person's worldview influences what he or she finds important, meaningful, and satisfying. For example, one person may react to a disaster with posttraumatic stress whereas another finds meaning in the opportunity to help others, thereby growing in self-awareness and well-being. In fact, internal psychological resources, rather than external material and social supports, satisfy a person's needs to an increasing degree as he or she matures. People with PDs have limited internal psychological resources, so their psychotherapy must begin to awakening values that will promote well-being and a willingness to work for greater self-understanding.

Table 26–25.
Awakening a Positive Perspective: Four Major Principles of the Science of Well-Being

	Experiences and Activities	
Positive Approaches to Well-Being	**Recommended**	**Not Recommended**
Letting go of struggles	Acts of hope and self-direction; accepting responsibility Silence of mind, phase 1[a] (calm reflection)	Any violence, fighting, complaining, or blaming others consumerism feeding greed/addiction
Working in the service of others	Acts of kindness/cooperation; purposeful giving of oneself Union in nature[a] Silence of mind, phase 2[a] (mindful meditation)	Divisions feeding fear and hate; possessiveness and hoarding; seeking power and dominance
Growing in awareness	Acts of faith and humility; transcendent problem solving Union in nature[a] Listening to the heart[a] Silence of mind, phase 3[a] (contemplation)	Criticism or praise of self and others, depending on external direction, denial of what one does not understand

Understanding thought processes	Calm reflection, mindful meditation, contemplation (silence of mind,[a] all three phases)	Seeking justification, depending on external advice, depending on intellect alone

[a]Further literature and programs for cultivation of well-being are available through a nonprofit educational foundation (http://anthropedia.org).

Used with permission of the Center for Well-Being at Washington University.

A systematic approach to personality transformation without tension or conflict has been described by Cloninger as coherence therapy in his book on the Science of Well-Being. The second stage of treating PD involves the awakening of the positive outlooks on life that are needed for well-being, as described in Table 26–25. The basic principles of well-being are summarized in Table 26–25 along with therapeutic experiences and activities for patients that are designed to help them value the dignity of their life and that of others as human beings as a result of self-awareness of each person's body, mind, and psyche. Psychiatry literally means the "healing of the psyche," and the psyche refers to a creative, self-aware consciousness that is more than analytical reasoning. Recovery from mental illness most consistently involves a transformation of a person's spirit or basic perspective about life. However, psychiatrists often become narrowly focused on reduction of symptoms or risk of harm and forget the importance of helping the patient to maintain self-respect and hope. So often the insight that psychiatry is the "healing of the psyche" has been neglected or denied as a result of the errors of materialistic and reductionistic thinking. What is meant here by an awakening and healing of the psyche is that the patient becomes directly aware that his or her worldview (i.e., his or her outlook on life) has an impact on his or her thoughts, emotions, and actions. The human outlook on life is what makes people vulnerable to mental disorders, so people must become aware of the assumptions implicit in their initial perspectives. Without some degree of awareness of the consequences of one's initial outlook, it is not possible to transcend the conflicts and contradictions in the thoughts and emotions that derive from these more or less coherent outlooks. As a result, psychotherapies that neglect the stage of spiritual awakening inevitably leave their patients locked in an inner struggle among parts of themselves from which there is no escape, as described poignantly by Freud about himself in *Civilization and Its Discontents*. Further information about coaching or therapy to cultivate well-being is available for interested clinicians (http://anthropedia.org).

Once there is an awakening of self-awareness to a metacognitive level, then a patient can proceed to the advanced stages of treatment that are briefly summarized in Table 26–26. Notice now that the patient is behaviorally stable and self-aware, then it is possible to focus on the *causes*

of the symptoms that were targets of initial intervention during the stabilization phase. For example, the causes of anxiety and dysregulated mood are rooted in a patient's lack of self-respect and trust of others, as expressed in his or her feelings of separateness, catastrophe, and victimization (see Table 26–26). The therapeutic strategies for addressing each of these sets of causes include psychoeducation, physical and other nonverbal therapies (e.g., psychomotor activities, acupressure), emotional skills training, cognitive skills training, and spiritual exercises. As a result, no one form of therapy (such as behavioral, cognitive, interpersonal, or psychodynamic) is really comprehensive, even with addition of modules from positive psychology or mindfulness training.

Notice in Tables 26–24 and 26–26 also that the symptomatic targets for psychobiological treatment correspond closely to the skill training modules that have been shown to be moderately effective in the treatment of severe PDs by Marsha Linehan and others except that the focus is shifted from a behavioral approach to the cognitive perspectives or schemas that lead to the behavioral problems. In addition, a nondualistic approach is taken of a stepwise path of character development to well-being and emphasis is placed on reconciliation of emotional conflicts that can be measured by the extremes of each of temperament dimensions. In essence, the extremes of each temperament are transcended by the development of particular forms of spiritually elevated thoughts—namely, self-respect reconciles the extremes of Harm Avoidance, self-mastery reconciles the extremes of Novelty Seeking, secure attachments reconcile the extremes of Reward Dependence, and the cultivation of virtues and transcendent meaning reconciles the extremes of Persistence and the limits of the finite human intellect. Both extremes of each temperament have advantages and disadvantages, and transcendence of the underlying conflict resulting from these disadvantages allows a person to live without tension or conflict about these issues as a result of a more holistic initial perspective.

Table 26–26.
Physical, Personal, Social, Cognitive, and Spiritual Therapeutic Procedures for Advanced Character Development (Stages 3 and 4 of Coherence Therapy for Personality Disorders)

Target Problems	Character Goals	Recommended Experiential Activities
1. Mood dysregulation (feeling separate, intolerant/hateful, self-critical, catastrophizing/distressed, victimized)	Trust and self-respect	Psychoeducation about health, sex, stress Physical exercise for relaxation and fitness Reconciliation of anxiety versus risk-taking Recognition of immature automatic thoughts

		Teamwork to build trust and respect
2. Aggression/impulsivity (being violent, angry, frustrated, greedy/jealous, proud, selfish)	Impulse control and self-mastery	Psychoeducation about nutrition, biorhythms Self-efficacy about cravings/substance abuse Reconciliation of impulsivity versus rigidity Nonviolent assertive communication Social service and acts of forgiveness
3. Social dysregulation (feeling rejected, detached, insecure or cold, unappreciated, unfriendly)	Empathy and secure attachments	Psychoeducation about social signals Empathy training and active listening skills Reconciliation of approval-seeking versus aloofness Social problem solving Therapeutic touch and appeasement
4. Cognitive distortion (feeling meaningless, judgmental, dualistic, afraid of death, lacking faith)	Mindfulness and meaning	Psychoeducation about self-transcendence Practice of mindfulness meditation Reconciliation of persistence by virtue Reflection on mysteries of God, life/death Creative works (art, music, writing, etc.)
5. Emptiness (seeking fulfillment, satisfaction, positive emotions, well-being coaching[a], virtue, well-being)	Virtue and well-being	Psychoeducation about well-being Recognizing triggers of negative emotion

[a]Literature and training available through http://anthropedia.org.

Used with permission of the Center for Well-Being at Washington University.

A major practical advantage of therapy modules that focus on the reconciliation of both extremes of each temperament is that heterogeneous groups of patients with widely different personality profiles can be treated together. The focus is on transcending emotional conflicts, coherent character development, and well-being for everyone, not particular personality subtypes. The stigma linked to PD and mental disorders in general is mitigated by facing the facts that everyone is imperfect but can learn to live without fear along the path to well-being. The design of therapies has often failed to recognize the path of development of character and well-being and the crucial role of spirituality in transcending and sublimating emotional conflicts. Psychotherapies have almost entirely focused on thought and have generally ignored the body as well as the psyche, as a result of the influence of errors in dualistic, behavioral, materialistic, or psychoanalytical thinking. Such an intellectual stance is arrogant and ultimately self-defeating. Alternative therapies have focused on methods that emphasize the body and ignore the benefits of medications or psychodynamics, which is equally arrogant and self-defeating. The coherence therapy described here is based on an integrative psychobiological approach that attends in a balanced way to all three

aspects of one's being—body, mind, and psyche. There is no incompatibility between biological, psychological, or spiritual aspects of the treatment of PD. In fact, there is a crucial synergy between the components of treatment directed at each aspect because rational brains do not function well when the emotional brain (i.e., limbic system) is distressed. There is much to be learned about the treatment of PD and related mental disorders, but the approach outlined here provides a paradigm that will allow the clinicians to integrate what they know about people and bring that to help and guide their treatment of every patient they encounter. It should be emphasized again that advanced psychotherapy or use of psychotropic medications for treatment of PD requires expert training or close supervision by an experienced clinician. This is a rapidly advancing area of psychiatry and psychology. Continuing education about the treatment of PD is likely to be especially rewarding for both psychiatrists and their patients because it offers a crucial gateway into the path to well-being and reduced disability from the full range of mental disorders.

References

Allport G. *Personality: A Psychological Interpretation*. London: Constable; 1937.

American Psychiatric Association. *Diagnostic and Statistical Mental of Mental Disorders (IIIrd ed, III-R, IV, IV-TR, 5)*. Washington: APA Press; 1980, 1987, 1994, 2000, 2013.

Arnedo J, Svrakic DM, del Val C, et al. Uncovering the hidden risk architecture of the schizophrenias: confirmation in three independent genome-wide association studies. *Am J Psychiatry*. 2015;172(2):139–153.

*Cloninger CR. A systematic method for clinical description and classification of personality variants. A proposal. *Arch Gen Psychiatry*. 1987;44:573–588.

Cloninger CR. A practical way to diagnose personality disorder: a proposal. *J Pers Disord*. 2000;14:99–108.

*Cloninger CR. *Feeling Good: The Science of Well-Being*. New York: Oxford University Press; 2004.

Cloninger CR. Evolution of human brain functions: the functional structure of human consciousness. *Aust NZ J Psychiatry*. 2009;43(11):994–1006.

*Cloninger CR, Svrakic DM, Przybeck TR. A psychobiological model of temperament and character. *Arch Gen Psychiatry*. 1993;50:975–990.

Cloninger CR, Zohar AH. Personality and the perception of health and happiness. *J Affective Disorders*. 2011;129(1–2):24–32.

Cloninger CR, Zohar AH, Hirschmann S, Dahan D. The psychological costs and benefits of being highly persistent: personality profiles distinguish mood disorders from anxiety disorders. *J Affective Disorders*. 2012;136(3):758–766.

Eysenck HJ. *The Biological Basis of Personality*. Springfield, IL: Charles Thomas; 1967.

Gardini S, Cloninger CR, Venneri A. Individual differences in personality traits reflect structural variance in specific brain regions. *Brain Res Bull*. 2009;79(5):265–270.

Hansenne M, Pinto E, Scantamburlo G, et al. Harm avoidance is related to mismatch negativity

(MMN) amplitude in healthy subjects. *Pers Individ Dif.* 2003;34:1039–1048.

Hirvonen J, Hietala J, Kajander J, et al. Effects of antidepressant drug treatment and psychotherapy on striatal and thalamic dopamine D2/3 receptors in major depressive disorder studied with [11C]raclopride PET. *J Psychopharmacol.* 2011;25(10):1329–1336.

Josefsson K, Jokela M, Cloninger CR, et al. Maturity and change in personality: developmental trends in temperament and character in adulthood. *Dev Psychopathol.* 2013;25:713–727.

Karlsson H, Hirvonen J, Kajander J, et al. Psychotherapy increases brain serotonin 5-HT1 A receptors in patients with major depressive disorder. *Psychol Med.* 2013;40(3):523–528.

Kernberg O. *Borderline Conditions and Pathological Narcissism.* New York: Aronson; 1975.

Markett S, Weber B, Voigt G, et al. Intrinsic connectivity networks and personality: the temperament dimension harm avoidance moderates functional connectivity in the resting brain. *Neuroscience.* 2013;240:98–105.

Mulder R, Joyce P. Temperament and structure of personality disorder symptoms. *Psychol Med.* 1997;27(1):99–106.

Nakao T, Matsumoto T, Shimizu D, et al. Resting state low-frequency fluctuations in prefrontal cortex reflect degrees of harm avoidance and novelty seeking: an exploratory NIRS study. *Front Syst Neurosci.* 2013;7:115.

Perry C, Vaillant G. *Personality disorders.* In: Kaplan HI, Sadock BJ, eds. *Comprehensive Textbook of Psychiatry.* 5th ed. Baltimore, MD: Williams and Wilkins; 1989:1352–1387.

Pezawas L, Meyer-Lindenberg A, Drabant EH, et al. 5HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. *Nat Neurosci.* 2005;8(6):828–834.

Quilty LC, Godfrey KM, Kennedy SH, Bagby RM. Harm avoidance as a mediator of treatment response to antidepressant treatment of patients with major depression. *Psychother Psychosom.* 2010;79:116–122.

Qureshi IA, Mehler MF. Emerging roles of non-coding RNAs in brain evolution development, plasticity, and disease. *Nat Rev Neuroscience.* 2012;13(8):528–541.

Reich J, Frances A. The structural interview method for diagnosing borderline disorders. *Psychiatr Q.* 1984;56:229–235.

Schneider K. *Psychopathic Personalities.* London: Casell; 1956.

*Sperry L. *Cognitive Behavior Therapy of DSM-IV-TR Personality Disorders.* 2nd ed. New York: Routledge; 2006.

Strack S. *Differentiating Normal and Abnormal Personality.* 2nd ed. New York: Springer Publishing Company; 2006.

*Svrakic DM, Whitehead C, Przybeck TR, Cloninger CR. Dimensional diagnosis of personality disorders by the seven factor model of temperament and character. *Arch Gen Psychiatry.* 1993;50:991–999.

Svrakic NM, Svrakic DM, Cloninger CR. A general quantitative theory of personality development: fundamentals of a self-organizing psychobiological complex, development and psychopathology. 1996;8:247–272.

Tulving E. Episodic memory: from mind to brain. *Ann Rev Psychol.* 2002;53:1–25.

*Tyrer P, Reed GM, Crawford M. Classification, assessment, prevalence, and effect of personality disorder. *Lancet.* 2015;385(9969):717–726.

Van Luyn B, Akhtar S, Livesley WJ, eds. *Severe Personality Disorders: Everyday Issues in Clinical Practice*. Cambridge: Cambridge University Press; 2007.

Van Schuerbeek P, Baeken C, De Raedt R, De Mey J, Luypaert R. Individual differences in local gray and white matter volumes reflect differences in temperament and character: a voxel-based morphometry study in healthy young females. *Brain Res*. 2011;1371:32–42.

27 Psychosomatic Medicine

▲ 27.1 History and Current Trends

CAROL L. ALTER, M.D., AND STEVEN A. EPSTEIN, M.D.

Psychosomatic medicine has been a specific area of focus within the field of psychiatry for more than 50 years. The term *psychosomatic* is derived from the Greek words *psyche* (soul) and *soma* (body). The term literally refers to how the mind affects the body. Unfortunately, it has come to be used, at least by the lay public, to describe an individual with medical complaints that have no physical cause and are "all in your head." In part due to this misconceptualization, the American Psychiatric Association's (APA) *Diagnostic and Statistical Manual of Mental Disorders* (DSM), in 1980, deleted the nosological term psychophysiological (or psychosomatic) disorders and replaced it with "Psychological Factors Affecting Physical Conditions" in DSM-5. Nonetheless, the term continues to be used by researchers and is in the title of major journals in the field (e.g., *Psychosomatic Medicine*, *Psychosomatics*, and *Journal of Psychosomatic Research*). It is also used by the two major national organizations in the field (the Academy of Psychosomatic Medicine and the American Psychosomatic Society) as well as international organizations (e.g., the European Association of Psychosomatic Medicine). In 2003, the American Board of Medical Specialties and the American Board of Psychiatry and Neurology approved the specialty of psychosomatic medicine. That decision recognized the importance of the field and also brought the term *psychosomatic* back into common use in the field.

ANCIENT HISTORY

According to Hippocrates (460 to 370 BC), who selected from the Egyptian and the Greek traditions, disease originated within the body and was due to an imbalance in fluid matter. That imbalance could be related to or even caused by a similar imbalance in the patient's external environment. Hippocrates warned the visiting doctor to consider the altitude, the wind direction, the purity of water supply, and the season of the year before

making any diagnosis. He pointed out that citizens of cities in particular environments would run to phlegmatic or sanguine temperaments as a reflection of the city's coldness or heat. Physical or fluid imbalance could be caused by emotional upset, too. Hippocrates reported that fear produced sweat and that shame brought on palpitations of the heart; he urged young physicians to look at patients with kindly expressions and never with impatience, because impatience could inhibit the return of health. He is said to have cured a king of an intestinal lesion by analyzing a dream. It was Hippocrates who wrote "in order to cure the human body, it is necessary to have a knowledge of the whole of things."

In *Timaeus*, Plato remarked that trouble in the soul could bring trouble to the body, and, in *Charmides*, Plato quoted Socrates, who had himself attributed the words to a Greek king: "As it is not proper to cure the eyes without the head, nor the head without the body, so neither is it proper to cure the body without the soul." Aristotle (384 to 322 BC) observed that the emotions of anger, fear, courage, and joy affect the body, and Areteus (first century bc) pinpointed a disturbance of the emotions as one of the six major causes of paralysis.

MODERN HISTORY

As Edward Shorter discusses in detail in his history of psychosomatic illness, ways of presenting illness have varied over history, since patients unconsciously select symptoms that are thought to represent true somatic illnesses. Prior to 1800, physicians did not conduct clinical evaluations and could not distinguish somatic from psychogenic illness. As a result, the diagnoses of hysteria and hypochondriasis could easily be made erroneously in the presence of true medical illnesses and did not suggest any specific disease presentations. According to Shorter, "cultural shaping" of patients' symptoms began in the middle of the 19th century with the popular diagnosis of spinal irritation. That diagnosis was made when a physician believed that spinal cord irritation caused symptoms, and the patient complained of spine sensitivity as well as peripheral symptoms. As was the case for previous explanations of illness, spinal irritation was either an invalid diagnosis for true medical illness or a "code word" for psychosomatic symptoms. The term "psychosomatic" was first used by Johann Christian Heinroth in 1818.

In the mid-19th century, after much popularity in Europe and the United States, the diagnosis of spinal irritation went out of favor. It was replaced by "reflex neurosis" as an explanation for somatization. That diagnosis was based on the view that any affected organ could cause irritation in any other organ in the body. The conceptualization was based on reflex theory, the notion that nervous connections throughout the body regulate all bodily organs, independent of free will. The Berlin neurologist Moritz Romberg was an influential proponent of the view that the uterus

was the cause of hysteria, a reflex neurosis. In this view, hysteria, which encompassed globus hystericus, hysterical seizures, and other psychogenic disorders, was due to irritation of the uterus and the ovaries. He believed that male genitals could also cause hysteria, but that they do so less commonly because they are less frequently irritated. In the late 19th century, reflex theory was one of the leading models of nervous illness. Since physicians promulgated the theory to their patients, it is not surprising that they found many of their patients to be preoccupied with their genital organs. Gynecologist August Rheinstadter of Cologne commented skeptically in 1884:

> But now, a true hysteromania, a Furor uterinus, has unfortunately arisen from the previous reluctance of women to undergo a gynecological examination and from the indifference which doctors once displayed towards gynecological disease. So that every woman who suffers from migraine, stomach cramps, or palpitations believes herself to have a uterine illness and indeed will find some physician willing to indulge her in the treatment of this presumed cause.

In the late 19th century, many physicians in Europe and the United States believed that gynecologic problems produced hysteria and other forms of mental illness. As a result, gynecologists performed surgery to treat hysteria and insanity, including removal of healthy ovaries. Widespread availability of gynecologic surgery in the United States also led women to request such surgery to relieve psychosomatic symptoms. In his history of psychosomatic illness, Shorter notes that some Americans' "addiction to surgery" was due to practitioners convincing the public that vague physical symptoms were due to reflex phenomena from peripheral organs. The nose was the final organ addressed by reflex theory in the late 19th century. One of its chief proponents was Wilhelm Fliess, a Berlin general practitioner, who greatly influenced Freud.

In the 19th century, hysterical fits were the most common form of motor hysteria. There were two primary forms: catalepsy and uncontrolled motor activity such as thrashing about. The psychiatrist Pierre Briquet described over 400 women with hysterical conversion disorders whom he had evaluated at the Pitie Hospital in Paris from 1849 to 1859. Many had multiple unexplained somatic complaints; the diagnosis of Briquet syndrome ultimately developed into the diagnosis of somatization disorder and is most similar to the current DSM-5 diagnosis of somatic symptom disorder.

Hysterical paralysis, often due to stressful life events, was apparently quite common through the 19th and early 20th centuries. Louis Verhaeghe, a spa physician in Belgium, wrote in 1850:

> How often do we see patients with paralyzed arms, legs, the side of the face or even a more limited area such as one or two fingers, problems attributable to an unequal distribution of nervous fluid. It is individuals of a pronounced nervous temperament, subject to hysterical fits, somnambules, persons subject to shameful habits or who abuse the pleasures of love, who are most likely to experience these paralyses.

Astasia-abasia was also seen frequently; Silas Wier Mitchell in Philadelphia referred to it as "hysterical motor ataxia." One of the most prominent physicians of the late 19th century, Jean-Martin Charcot promulgated a highly influential theory of hysteria. He believed that it was a chronic inherited functional disease of the nervous system. "In the [hysterical] fit," he said in 1882, "nothing is left to chance, that to the contrary everything unfolds according to the rules, which are always the same and characterize what we see in our outpatients as well as inpatients; they are valid for all countries, for all epochs, for all races, and are, in short, universal." Although his conception of hysteria was widely disseminated, some physicians who were aware of his ability to suggest symptoms to vulnerable patients were dubious of his theory. The Swiss neurologist Paul Dubois wrote in 1904: "Endowed with the spirit of authority, [Charcot] . . . suggested to them their attitudes and their gestures. Example is contagious even in sickness, and in the great hospitals of Paris, at La Salpetriere, all cases of hysteria resemble each other. At the command of the chief of the staff, or of the interns, they begin to act like marionettes, or like circus horses accustomed to repeat the same evolutions." It was not until near the end of his life that Charcot's theories became more psychologically focused.

Within a decade after he died in 1893, Charcot hysteria had virtually disappeared. Many physicians began to believe that some of his patients most likely had organic illness. Joseph Babinski's discovery in 1896 of the Babinski sign was extremely helpful in distinguishing organic disease from hysteria. In 1901, he put forward a new conceptualization of hysteria: It was defined as any symptom that could be induced by suggestion and eliminated by persuasion.

As the reflex model faded in prominence in the late 19th century, the most dominant model became one in which central nervous disease was held responsible for nervous symptoms. In the early 20th century, psychological theories of somatization became popular, but they became dominant only in the latter half of the 20th century. "Neurasthenia," or tired nerves, was a new diagnosis at the end of the 19th century that was brought to prominence by George Beard, a New York electrotherapist. A term applied broadly, the diagnosis encompassed depression, anxiety, somatization, male hysteria, and chronic fatigue. Wilhelm Stekel, a Viennese psychoanalyst, coined the term "somatization" in the 1920s.

Sigmund Freud was the principal theoretician to bring psyche and soma back together. He demonstrated the importance of the emotions in producing mental disturbances and somatic disorders. His early psychoanalytic formulations detailed the role of psychic determinism in somatic conversion reactions. Picking up where the Greeks and Romans had left off, Freud redignified the study of the emotions as a separate study and pointed to their relationship with the soma in the new field of

psychiatry. Using Freud's insight, a number of workers in the early decades of the 20th century tried to expand the understanding of the interrelationship of psyche and soma. The influence on adult organ tissue of various unresolved pregenital impulses was proposed by Karl Abraham in 1927, the application of the idea of conversion reaction to organs under the control of the autonomic nervous systems was described by Sandor Ferenczi in 1926, and the attaching of a symbolic meaning to fever and hemorrhage was suggested by George Groddeck in 1929.

In the 20th century, somatization symptoms changed from predominantly neurologic (e.g., hysterical paralysis) to other symptoms such as fatigue and chronic pain. Edward Shorter attributes this change to three causes: (1) improvements in medical diagnostic techniques made it easier to rule out organic causes for neurologic disease; (2) the central nervous system (CNS) paradigm faded; and (3) social roles changed (e.g., the disappearance of the historical notion that "weak" women would be expected to have fainting spells and paralysis).

Although hysterical neurologic symptoms have remained relatively less common in the 21st century, CNS explanations of chronic pain and fatigue are gaining prominence. For example, functional brain research has demonstrated brain dysfunction and possibly genetic contributions among some individuals with fibromyalgia and chronic fatigue syndrome. Those syndromes, while still thought by some to represent somatization variants, are currently established medical diagnoses.

HISTORY OF PSYCHOSOMATIC MEDICINE RESEARCH

Since the mid-20th century, research in psychosomatic medicine and consultation-liaison psychiatry has taken two interconnected paths. Psychosomatic medicine research has generally focused on understanding psychophysiologic mechanisms underlying mind–body relationships. Consultation-liaison research, however, has largely been directed at understanding psychiatric problems among clinical populations with medical illnesses.

In the early 20th century, two theoretical trends developed, one suggesting that specific emotions led to specific cell and tissue damage, and the second that generalized anxiety created the preconditions for a number of not-necessarily predetermined diseases. Cases of shell shock during World War I and new endocrine studies provided observations for the swell of interest in theory evidenced in the 1930s.

In the early 1930s, Franz Alexander, who had left Berlin to move to Chicago where he founded the Chicago Institute for Psychoanalysis, described how he believed specific psychological conflicts caused specific somatoform symptoms. Alexander saw conflict as a stress and suggested that, when conflict presents itself, the individual may suppress the stress and produce, through the voluntary nervous system, a reaction such as the

conversion reaction described by Freud. However, after suppressing stress, the individual may, through the autonomic system, keep his or her sympathetic responses alert for heightened aggression or flight or parasympathetic responses alerted for heightened vegetative activity. According to these theories, prolonged alertness and tension can produce physiological disorders and eventual pathology of visceral organs.

For instance, Alexander postulated that, in a passive-dependent person without someone to satisfy his or her dependence, stress is created. That particular stress may stimulate and keep alert the parasympathetic nervous system, which means that too much gastric acid is secreted, and gastric hypermotility results, all of which may lead to a peptic ulcer. Another dependent person with a different genetic set may, in repressing conflict, stimulate parasympathetic overfunctioning through pathways leading to colitis or asthma. Still other dependent persons, in seeking to move beyond dependency, incorporate the stress; such a move entails overstimulation of the sympathetic system, and the resulting chronic alertness produces migraine, hypertension, or arthritis.

Although there were multiple case reports supporting Alexander's views, it has not been demonstrated that the same specific conflicts are present in all cases of the same disease. Early methodologically flawed reports were rooted in psychoanalytic explanations of "psychosomatic" illnesses. This work had numerous problems, including selection bias, small sample size, and lack of control groups.

A number of investigators developed the pathway concept of constellations into other theories involving the whole personality. In 1954, Flanders Dunbar proposed that the ambitious, hard-driving man would be prone to coronary occlusion. That personality type is similar to the competitive, restless, time-haunted, and coronary-bound type A person proposed by M. Friedman and R. H. Rosenman in 1959. Modern research has supported the notion that hostility, one component of type A personality, is a risk factor for the development of coronary artery disease.

A second trend in psychosomatic medicine that began in the middle of the 20th century has been to investigate what happens to a person in a nonspecific way when faced with stress. In the 1950s, Harold Wolff and Stewart Wolf observed that chronic hyperfunction or chronic hypofunction in the vascular and secretory activities of the mucosa of the gastrointestinal (GI) and respiratory systems can produce pathology. Overfunctioning of the mucosa was correlated with hostility, and underfunctioning was correlated with fear or sadness. The patient's entire reactive patterning and his or her life history account for whether he or she reacts to stress by hyperfunctioning or hypofunctioning.

Other workers in the nonspecific group described various possible mechanisms by which psychologically induced stress may cause organic disease in humans and animals. In 1950, Hans Selye thought that the

hypophyseal–adrenocortical axis responded to various types of physical and psychic stress with hormonal changes that can ultimately cause a variety of organic diseases, such as rheumatoid arthritis and peptic ulcer. Selye viewed such diseases as a by-product of the body's attempt to adapt to stress from any source.

Experimental psychologists with a learning theory conception of behavior studied the effects of chronic unrelieved anxiety in humans and animals and found that gastric hydrochloric acid production increases under such circumstances. Because such acidity is a precursor of peptic ulcer, they concluded that chronic anxiety is the variable intervening between the behavioral and physical events involved in psychosomatic illness. Modern research does continue to support the role of emotional factors in peptic ulcer disease (e.g., the work of Levenstein [see below]), indicating that anxiety leads to poorer ulcer healing.

Beginning in the late 1960s, T. H. Holmes and R. H. Rahe showed that life crises often precede illness and that there is some correlation between the intensity of the crisis and the severity of the illness that follows; the crisis ranked as the most intense by subjects was the death of a spouse. Other works suggested a giving-up–given-up complex, in which individuals feel powerless to change their environments or themselves and eventually lose the will to try, becoming mentally or physically ill, perhaps because of changes in the immunological and neuroendocrine systems.

As exemplified by the long history of psychosomatic research in pain and GI disease, the quality of research in the field has evolved significantly over the past 70 years. Early research efforts were psychoanalytically oriented case studies of conditions such as ulcerative colitis. Such research often led to invalid conclusions. For example, one report by A. Johnson and colleagues in 1947 of a series of 33 patients with rheumatoid arthritis led to the preliminary conclusion that women with this disorder have "a chronic inhibited hostile aggressive state as a reaction to the earliest masochistic dependence on the mother that is carried over to the father and all human relationships, including the sexual." Nonetheless, even those authors were aware of potential methodological limitations of their work: "but these factors . . . are found so commonly in patients who do not suffer from arthritis that additional etiologic factors, still unknown, must be postulated."

In the mid-20th century, research on the classic psychosomatic illnesses such as rheumatoid arthritis and peptic ulcer disease used projective testing such as the Rorschach to identify specific personality traits. Later studies used structured clinical interviews and objective instruments such as the Minnesota Multiphasic Personality Inventory (MMPI). In one leading journal in the field, *Psychosomatic Medicine*, psychosomatic research on upper GI syndromes has changed its focus over the past 70 years. For example, psychodynamic and personality research peaked in the

mid-20th century. However, research on stress and interactions between psychological and medical factors have remained at fairly steady levels over that time period.

Psychophysiological research began in the first half of the 20th century. Researchers used techniques such as electromyogram (EMG), electroencephalogram (EEG), and galvanic skin response to examine physiologic parameters among individuals with a variety of psychosomatic illnesses. Research has evolved to the present day to include areas such as heart rate variability and stress, and the role of the CNS in illnesses such as fibromyalgia and chronic fatigue syndrome. Pain treatment research has evolved from less-controlled studies to randomized-controlled studies.

As early as the 1960s, researchers reported on the prevalence of various psychiatric problems in inpatient and outpatient medical settings. Methodological variability led to wide variation in estimates of comorbidity. Early work also reported on the characteristics of patients referred for psychiatric consultation. R. J. Kahana's and G. L. Bibring's classic work in 1964 emphasized the importance of understanding personality types to the effective treatment of one's patients.

Consultation-liaison psychiatry research in the mid-20th century also elucidated the most common reasons for psychiatric consultation. While many of those reasons hold true to this day, psychoanalytic influences on physician consulting behavior have clearly waned in recent years. Consider Zbigniew Lipowski's summary in 1967 of the two most common diagnostic questions asked of the consulting psychiatrist:

1. Is this particular bodily complaint (or set of complaints) explicable as an expression of psychological ill health, that is, is it partly or wholly psychogenic?
2. Is the inner experience of psychic change reported by the patient or his observable behavior explicable as the direct results of, or a psychological reaction to, organic disease?

However, reasons for requests for assistance in patient management have seemingly changed very little since Lipowski summarized them in 1967, and will be familiar to all currently practicing psychosomatic medicine psychiatrists:

1. Suicidal attempt or threat
2. Grossly disturbed behavior, for example, delirium and psychosis
3. Excessive emotional reactions, for example, depression and anger
4. Refusal to cooperate
5. Delayed convalescence
6. Conflict between patient and personnel
7. Patient with psychiatric history
8. Psychiatric side effects of drugs
9. Selection and/or preparation of patients, for example, for surgery

ɔ. Disposition

Although such epidemiologic work has been useful for educational and staffing purposes, it did not contribute significantly to the advancement of understanding of mind–body relationships (Table 27.1–1).

Table 27.1–1.
Major Conceptual Trends in the History of Psychosomatic Medicine

I. Psychoanalytic

Sigmund Freud (1900): Somatic involvement occurs in conversion hysteria, which is psychogenic in origin—e.g., paralysis of an extremity. Conversion hysteria always has a primary psychic cause and meaning; i.e., it represents the symbolic substitutive expression of an unconscious conflict. It involves organs innervated only by the voluntary neuromuscular or the sensorimotor nervous system. Psychic energy that is dammed up is discharged through physiological outlets.

Sandor Ferenczi (1910): The concept of conversion hysteria is applied to organs innervated by the autonomic nervous system; e.g., the bleeding of ulcerative colitis may be described as representing a specific psychic fantasy.

George Groddeck (1910): Clearly organic diseases, such as fever and hemorrhage, are held to have primary psychic meanings; i.e., they are interpreted as conversion symptoms that represent the expression of unconscious fantasies.

Franz Alexander (1934, 1968): Psychosomatic symptoms occur only in organs innervated by the autonomic nervous system and have no specific psychic meaning (in contrast to conversion hysteria) but are end results of prolonged physiological states, which are the physiological accompaniments of certain specific unconscious repressed conflicts. Presented first conceptualization of the biopsychosocial model.

Helen Flanders Dunbar (1936): Specific conscious personality pictures are associated with specific psychosomatic diseases, an idea similar to Meyer Friedman's 1959 theory of the type A coronary type.

Peter Sifneos, John C. Nemiah (1970): Elaborated the concept of alexithymia. Developmental arrests in the capacity to identify and express conflict-related affect result in psychosomatic symptom formation. Concept of "alexithymia" modified later by Stoudemire, who advocated the term "somatothymia" that emphasized cultural influences on use of somatic language and somatic symptoms to express affective distress.

II. Psychophysiological

Walter Cannon (1927): Demonstrated the physiological concomitants of some emotions and the important role of the autonomic nervous system in producing those reactions. The concept is based on Pavlovian behavioral experimental designs.

Harold Wolff (1943): Attempted to correlate life stress to physiological response, using objective laboratory tests. Physiological change, if prolonged, may lead to structural change. He established the basic research paradigm for the fields of psychoimmunology, psychocardiology, and psychoneuroendocrinology.

Hans Selye (1945): Under stress a general adaptation syndrome develops. Adrenal cortical hormones are responsible for the physiological reaction.

Meyer Friedman (1959): Theory of type A personality as a risk factor for cardiovascular disease. The basic concept was introduced by Helen Flanders Dunbar as early as 1936.

Robert Ader (2007): Beginning in the 1970s, established the basic concepts and the research methods for the field of psychoneuroimmunology.

III. Sociocultural

Karen Horney (1939), James Halliday (1948): Emphasized the influence of culture in the development of psychosomatic illness. They thought that culture influences the mother, who, in turn, affects the child through her relationship with the child—e.g., nursing, child rearing, and anxiety transmission.

Thomas Holmes, Richard Rahe (1975): Correlated the severity and the number of recent stressful life events with the likelihood of disease.

IV. Systems theory

Adolph Meyer (1958): Formulated the psychobiological approach to patient assessment that emphasizes the integrated assessment of developmental, psychological, social, environmental, and biological aspects of the patient's condition. Basic concept of the biopsychosocial model is implicit in his approach.

Zbigniew Lipowski (1970): A total approach to psychosomatic disease is necessary. External (ecological, infectious, cultural, environmental), internal (emotional), genetic, somatic, and constitutional factors as well as past and present history are important and should be studied by investigators working in the various fields in which they are trained.

George Engel (1977): Coined the term "biopsychosocial" derived from general systems theory and based on conceptual ideas introduced much earlier by Alexander and Meyer.

Leon Eisenberg (1995): Contemporary psychiatric research demonstrates that the mind–brain responds to biological and social vectors while being jointly constructed of both. Major brain pathways are specified in the genome; detailed connections are fashioned by, and consequently reflect, socially mediated experience in the world.

CURRENT TRENDS

The practice of psychosomatic medicine has evolved considerably since its early clinical origins and has come to focus on psychiatric illnesses that occur in the setting of physical health care. In large part this evolution has occurred as a result of the increased complexity of medicine, the increased understanding of the relationship of medical illness to psychiatric illness, and the greater appreciation of mind and body as one. A key outcome of this has been the granting of subspecialty status for psychosomatic medicine. Clinical care is now delivered in a variety of health care settings and utilizes an ever expanding set of diagnostic tools as well as many effective somatic and psychotherapeutic interventions. Research in the area has progressed to include a greater understanding of the relationship between chronic medical conditions and psychiatric disorders and has examined the pathophysiologic relationships, the epidemiology of comorbid medical and psychiatric disorders, and the role that specific interventions play in physiologic, clinical, and economic outcomes.

The extent of the burden caused by co-occurring mental and physical disorders represents a tremendous public health problem. This was highlighted by the report from U.S. President George W. Bush's New Freedom Commission, *Achieving the Promise: Transforming Mental Health Care in America*, which called for closer coordination of the medical and mental health systems. Further, the Institute of Medicine (IOM) study, *Quality of Healthcare for Mental and Substance Use Disorders* (2005), states: "there is no health care without mental health care." There has been increasing awareness that mental disorders may be both risk factors for and poor prognostic indicators in chronic medical illnesses such as cardiovascular disease, diabetes, and cancer.

Psychiatric morbidity is very common in patients with medical

conditions, with a prevalence ranging from approximately 20 to 65 percent, depending on the illness. Patients in the general hospital have the highest rate of psychiatric disorders when compared to community samples or patients in ambulatory primary care. For example, compared to community samples, depressive disorders in the general hospital are more than twice as common, and substance abuse is two to three times as common. Delirium occurs in approximately 20 percent of inpatients. Psychiatric morbidity has serious effects on medically ill patients and is often a risk factor for exacerbation of their medical conditions. It is well established that depression is both a risk factor and a poor prognostic indicator in coronary artery disease. Psychiatric illness worsens cardiac morbidity and mortality in patients with a history of myocardial infarction, diminishes glycemic control in patients with diabetes, and decreases return to functioning in patients experiencing a stroke. Depressive and anxiety disorders compound the disability associated with stroke. In the context of neurodegenerative disease such as Parkinson disease or Alzheimer disease, depression, psychosis, and behavioral disturbances are significant predictors of functional decline, institutionalization, and caregiver burden. Hospitalized patients with delirium are significantly less likely to improve in function compared to patients without delirium. Delirium is associated with worse outcomes after surgery, even after controlling for severity of medical illness.

In addition, depression and other mental disorders significantly impact quality of life and the ability of patients to adhere to treatment regimens (e.g., in patients with diabetes mellitus). Psychiatric disorders are linked to nonadherence with antiretroviral therapy, adversely affecting the survival of human immunodeficiency virus (HIV)-infected patients, and they worsen the prognosis and quality of life of cancer patients. Finally, they are linked to nonadherence with safe sex guidelines and with use of sterile needles in HIV-infected injection drug users, thus having major public health implications.

Failure to identify, evaluate, diagnose, treat, or achieve symptom resolution of psychiatric morbidity in medical care settings results in significantly increased service utilization. For example, depression, dementia, and delirium are associated with higher utilization of medical care, both in the hospital (longer lengths of stay) and after discharge. When delirium is unrecognized and untreated in the hospital, it may result in unnecessary placement in nursing homes instead of discharge to home. Untreated depression is associated with higher medical utilization after hospital discharge and with higher mortality and morbidity in coronary disease, hypertension, diabetes, and stroke.

THE CLINICAL PRACTICE OF PSYCHOSOMATIC MEDICINE

The primary objective for psychosomatic medicine is the diagnosis and

treatment of psychiatric disorders in patients with complex medical conditions. Thus, psychosomatic medicine focuses on a variety of clinical problems that occur in patients being treated in medical settings, including primary psychiatric disorders such as delirium and dementia, which occur as a *result* of a medical condition; anxiety or depression, which occur *in the setting* of chronic medical conditions but may also be due to medical causes; and somatic symptom disorders (formerly somatoform disorders). In addition to the care of specific conditions, psychosomatic medicine also encompasses a variety of other clinical activities that occur in the medical setting, including evaluation of decision-making capacity, attention to quality of life and symptom management in chronic and terminally ill individuals, provision of liaison services to medical staff, and assistance with management of primary psychiatric conditions such as bipolar disorder or schizophrenia. Table 27.1–2 summarizes the types of clinical problems addressed in psychosomatic medicine.

Psychosomatic medicine is practiced in a variety of settings and encompasses a combination of consultative, treatment, and liaison activities. Historically, the majority of this work occurred within the general hospital setting. Largely this has been due to the fact that the greatest risk factor for requiring psychiatric intervention has been the severity of the medical illness. However, as more illnesses become chronic in nature (e.g., cancer and HIV), there has been an increase in delivery of psychosomatic medicine services in outpatient settings as well. Regardless of where care is delivered, the basic approaches are similar: To conduct a psychiatric assessment, provide psychotherapeutic, behavioral, or pharmacologic interventions to patients, and to work closely with other medical professionals in a liaison capacity.

Table 27.1–2.
Summary of Clinical Problems in Psychosomatic Medicine

Type of Clinical Problem	Example
Psychiatric symptoms secondary to a medical condition	Delirium, dementia
Psychiatric symptoms as a reaction to medical condition or treatments	Anxiety related to chemotherapy, depression related to limb amputation
Psychiatric complications of medical conditions and treatments	Depression secondary to interferon treatment
Psychological factors contributing to medical symptoms	Somatic symptom disorders
Medical complications of psychiatric conditions or treatment	Neuroleptic malignant syndrome, acute withdrawal from alcohol or other substance

Co-occurring medical and psychiatric conditions	Recurrence of depressive disorder in setting of cancer treatment (conditions occur independently); schizophrenia in a patient with end-stage renal disease.
Psychiatric/psychosocial assessment	Capacity evaluation; evaluation prior to organ transplantation

EVALUATION PROCESS IN PSYCHOSOMATIC MEDICINE

Psychiatric assessment in the medical setting includes a standard psychiatric assessment as well as a particular focus on the medical history and context of physical health care. In addition to obtaining a complete psychiatric history, including past history, family history, developmental history, and a review of systems, the medical history and current treatments should be reviewed and documented. A full mental status examination should be completed, and focused neurologic and physical examinations are often indicated depending on the nature of the presenting problem.

Several components of the process should be emphasized. First, it is critical to speak with the referring clinician. In many cases, the reason for consultation is not clear, and it is often helpful to understand why the treating team believes the patient might have a psychiatric problem. Such information can guide the assessment and may provide information about other concerns the treatment team might have, including the patient's behavior and over or under use of treatment, such as pain medication. It may also reveal the team's anger or frustration with a particular patient. Second, collateral information from family members and other health care providers, including mental health professionals, is important in understanding any prior history of psychiatric problems and helping the team plan for ongoing care. Third, while the interview of the patient should be comprehensive, it also needs to be sensitive to the patient's medical status. In many cases, patients in the general medical setting are experiencing pain or discomfort, distracted by medical problems, or cognitively compromised. Thus, the interview may need to be focused on the reason for consultation.

A full mental status examination, including specific focus on cognitive functioning, should be completed. The mental status should address the level of consciousness, attention, concentration, memory, executive function, language, praxis, mood, affect, perception, judgment, and insight. The Mini-Mental State Examination or Montreal Cognitive Assessment is often a helpful component because it provides an objective score that can serve as a baseline for retesting, especially when delirium is a concern. It is also important to be sure that physical and neurological examinations have been completed. It is appropriate for the psychiatric consultant to either complete limited components of these examinations or request the referring clinicians to perform additional examinations.

The psychiatric consultant should also ensure that appropriate laboratory and other diagnostic evaluations have been conducted. Basic chemistry, hematologic, and cardiovascular workup can be supplemented with other tests, including thyroid function, liver function, vitamin B_{12}, folate, toxicology, serological tests for syphilis, and HIV testing. Additional diagnostic studies, including computed tomography (CT) or magnetic resonance imaging (MRI) scans of the brain, EEG, or lumbar puncture, may also be indicated.

Another important objective of the psychiatric evaluation is to gain an understanding of the patient's experience of his or her illness. In many cases, this becomes the central focus for both the psychiatric assessment and interventions. It is often helpful to develop an understanding of the patient's developmental and personal history as well as key dynamic conflicts, which in turn may help to make the patient's experience with illness more comprehensible. Such an evaluation can include use of the concepts of stress, personality traits, coping strategies, and defense mechanisms. Observations and hypotheses that are developed can help to guide psychotherapy with patients aimed at diminishing distress and may also be helpful for the primary medical team in their interactions with the patient.

Finally, a full report synthesizing the information should be completed and should include specific recommendations for additional evaluations and intervention. Ideally, the report should be accompanied by a discussion with the referring physician.

TREATMENTS USED IN PSYCHOSOMATIC MEDICINE

A host of interventions have been successfully utilized in psychosomatic medicine. Specific consideration must be given to medical illness and treatments when making recommendations for psychotropic medications. Psychotherapy also plays an important role in psychosomatic medicine and may vary in its structure and outcomes as compared to therapy that occurs in a mental health practice.

Psychopharmacologic recommendations need to consider several important factors. In addition to targeting a patient's active symptoms, considering the history of illness and treatments, and weighing the particular side effect profile of a particular medication, there are several other factors that must be considered that relate to the patient's medical illness and treatment. It is critical to evaluate potential drug–drug interactions and contraindications to the use of potential psychotropic agents. Since the majority of psychotropic medications are metabolized in the liver, awareness of liver function is important. General appreciation of side effects, such as weight gain, risk of development of diabetes, and cardiovascular risk, must be considered in the choice of medications. In addition, it is also important to incorporate knowledge of recent data that

outline effectiveness and specific risks involved for patients with co-occurring psychiatric and physical disorders. For example, a greater understanding of the side effects of antipsychotic medications has raised concerns about the use of these medications in patients with dementia.

Psychosocial interventions also require adaptation when used in this population. The methods used and the goals of psychosocial interventions in the medically ill are often determined by the consideration of disease onset, etiology, course, prognosis, treatment, and understanding of the nature of the presenting psychiatric symptoms in addition to an understanding of the patient's existing coping skills and social support networks. There are ample data that psychosocial interventions are effective in addressing a series of identified problems and that such interventions in many cases are associated with a variety of positive clinical outcomes.

Every major psychosocial intervention has been used in psychosomatic medicine, including dynamic, supportive, and behavioral interventions, delivered in family, group, and individual settings. The goals for psychosocial interventions vary widely depending on the nature of both the psychiatric issues and the medical illness. In addition, the specific intervention will vary based on the illnesses and goals. For instance, while a goal generally present for all medical illness may be addressing changes in role and function resulting from illness, addressing concerns regarding death and dying will be a key concern for those in the terminal phases of an illness such as cancer. Coping with changes in physical function and pain are key issues that can be addressed for patients with arthritis. Addressing barriers to adherence to treatment becomes critical to psychosocial interventions in patients with diabetes.

An extensive body of research has demonstrated positive clinical outcomes associated with a variety of interventions. At minimum, psychosocial interventions have been shown to improve quality of life and psychiatric conditions; in addition, there is some evidence that psychosocial interventions may positively impact treatment adherence and medical outcomes. Interventions aimed at depression in patients with diabetes have shown not only decreases in depressive symptoms but also reduction in medical costs of care.

MENTAL DISORDERS IN PRIMARY CARE

An area of considerable focus has been the successful identification and management of patients with mental disorders treated in the primary care setting. Although a variety of interventions aimed at education of primary care physicians, screening initiatives, and implementation of quality improvement programs may have short-term benefit, they have not had a long-term impact on the improvement in detection and treatment of depression. However, programs that incorporate access to mental health

professionals, combined with education of primary care physicians and use of a care manager to provide appropriate monitoring and care strategies, have been shown to be very successful. These programs, known collectively as "collaborative care," provide direct evidence-based treatment (e.g., psychotherapy and psychopharmacology) with coordination of care between mental health professionals and the primary care team. A recent meta-analysis of randomized-controlled trials of collaborative care models for treating depression in primary care settings showed that they significantly improved the quality of care for depression, depression outcomes, patient and provider satisfaction, and increased patient adherence to the treatment regimen. In addition, similar disease management programs in medical settings have been shown to effectively treat panic disorder, late-life depression, depression in individuals with cancer, and depression in obstetrics and gynecology clinics. This model of intervention is consistent with the New Freedom Commission on Mental Health Report, which recommended that "collaborative care models for identification and treatment of mental disorders across the lifespan should be widely implemented in primary care settings and reimbursed by public and private insurers." Both the APA and the Academy of Psychosomatic Medicine have recently made integrated care initiatives a high priority for the field.

RESEARCH IN PSYCHOSOMATIC MEDICINE

Research in psychosomatic medicine has focused on several key areas that have contributed to the current understanding of the burdens of co-occurring psychiatric and medical disorders. Important advances in psychosomatic medicine research have identified common pathways between psychiatric illness and physical illnesses, such as cardiovascular disease or diabetes, bidirectional risk associated with the presence of psychiatric disorders or medical conditions such as stroke, and the identification of clinical interventions that can impact both medical and psychiatric outcomes. However, research in psychosomatic medicine has been limited by the complexity of research questions that require investigation of both medical and psychiatric outcomes and the existing "division" of the research enterprise into separate psychiatric and medical domains.

Various conceptual models of the relationship between medical and physical illness underlie the core research areas. For example, a large body of epidemiologic work has examined the relationship of depressive disorders to poor outcomes among patients with cardiovascular disease. In contrast, in the field of psychoneuroimmunology, basic science techniques are used to examine the relationship between psychosocial stressors and immune functioning. Table 27.1–3 presents an overview of conceptual areas and specific current research areas.

CURRENT PSYCHOSOMATIC MEDICINE RESEARCH

Psychosomatic medicine is now involved with a wide, ever-expanding spectrum of investigations examining the medical illness–psychiatry relationship. Important contributions have occurred in acquired immunodeficiency syndrome (AIDS), cancer, transplantation, cardiac, neurological, pulmonary, renal, GI disease, and obstetrics and gynecology. In each of these areas, first-generation studies identified the extent and nature of psychiatric morbidity associated with the most common diseases. More sophisticated second-generation cross-sectional epidemiological studies have established the prevalence rates of a broader range of psychiatric disorders in several medical conditions, as well as newer psychiatric syndromes, such as mild traumatic brain injury.

In some cases, a new body of research has extended the knowledge about genetic, neurochemical, and behavioral factors contributing to the development of psychiatric disorders among complex medically ill populations. This work has shown that many different mechanisms are involved. Examples include the relationships between depression and diabetes, the location of stroke and major depression, and specific autoantibodies and limbic encephalitis. This research has led to some intervention studies focused on reducing the rate of occurrence of these disorders, for example, hospital programs that reduce postcardiotomy delirium.

Of long-standing interest to the psychosomatic medicine field has been the relationship between emotional factors and disease development. Recent studies have demonstrated the dramatic impact of coexisting psychiatric morbidity on the course or outcome of specific medical illnesses. Studies have now been extended to examine the effect of psychiatric comorbidity on health costs. Comorbid psychiatric disorders have been shown to increase the length of hospital stay and rehospitalization rates in numerous illnesses. Several cost-effectiveness studies have demonstrated the value of psychosomatic medicine interventions in hospital and outpatient populations. In addition, psychobiologic research continues to focus on emotional factors and illness development in individuals who do not necessarily have psychiatric illnesses. For example, there continues to be important work in the relationship between emotional stressors and the functioning of the immune system.

The future research agenda for the psychosomatic medicine field will likely continue to revolve around these themes. Table 27.1–3 gives examples of the types of studies that are anticipated in the next few decades in the psychosomatic medicine field. They are grouped along types of studies, with specific examples provided within each group. Psychosomatic medicine research is likely to continue to move away from descriptive studies and focus increasingly on studies of mechanism, while at the same

time continuing with key treatment studies, until mechanism studies lead to the development of "designer" therapies. Thus, there will be increasing application of newer research methods from genetics, brain imaging, animal models, pharmacology, psychology, epidemiology, and clinical trials.

THE FIELD OF PSYCHOSOMATIC MEDICINE

The term psychosomatic medicine was first used by Felix Deutsch in the early 1920s. Organized psychosomatic medicine and consultation-liaison psychiatry had their origins in the publication of the journal *Psychosomatic Medicine* in 1939, the founding of the American Psychosomatic Society in 1947, and the journal *Psychosomatics* in 1960. The Academy of Psychosomatic Medicine was founded in 1953 and continues to be the organization for consultation-liaison psychiatrists. In 1968, John Schwab published the first consultation-liaison psychiatry textbook.

Table 27.1–3.
Current and Potential Research in Psychosomatic Medicine

Types of Studies	Examples of Current Research	Examples of Potential Studies
Epidemiologic and clinical epidemiologic studies		
Occurrence, nosology, and impact of different types of psychiatric morbidity over the course of general medical conditions	Occurrence and phenomenology of depression over the course of diabetes PTSD in individuals with traumatic brain injury	Relationship of psychiatric problems to disability and quality of life through the course of Parkinson disease Long-term effects of psychiatric conditions on quality of life after organ transplantation
Mechanistic studies		
Understanding the role of psychosocial, biological, and genetic factors in the causation of psychiatric and medical morbidity	Social isolation leading to poorer health outcomes Affective circuitry in chronic pain Family adversity and inflammation	Interaction of social isolation, genetic predisposition, and brain damage in the development of major depression after traumatic brain injury Gene–environment interactions in the development of chronic stress-related illnesses Genetic association studies of depressive symptoms after the onset of chronic medical illness
Studies of the social, psychological, biological, and genetic factors involved in the expression of abnormal illness behavior, somatic symptom disorders, and chronic pain	Personality traits as predictors of hypochondriasis Functional brain imaging abnormalities in chronic fatigue	Genetic risk factors for the development of chronic pain syndromes The role of health care system factors in contributing to the persistent utilization of health care

	and conversion	
Studies of the effects of medical illness and their treatments on the brain and on the development of secondary ("organic") psychiatric disorders	Interferon-induced depression Paraneoplastic limbic encephalitis	Animal models to understand brain mechanisms involved in delirium Brain imaging and neuropsychological studies of the brain effects of various cancer chemotherapeutic agents
Mechanistic studies to understand how psychiatric morbidity affects medical outcomes in several contexts	Depression as a predictor of poor adherence to diabetes treatment	Brain imaging and genetic studies of the role of depression in the development of cognitive decline Blood and cerebrospinal fluid biomarker studies of endocrine, immune, and other factors involved in the development of postpartum depression
Treatment studies		
Treatment studies of the efficacy and safety of psychiatric treatments as applied to the complex medically ill	Cognitive behavioral therapy in somatic symptom disorders The efficacy and safety of α-2 adrenoreceptor agonists in treating delirium	The efficacy and safety of emerging antidepressant therapies for depression and anxiety after organ transplantation The efficacy and safety of cholinesterase inhibitors or memantine in treating postchemotherapy cognitive impairment
Treatment studies of how treating psychiatric conditions affect medical outcomes in the complex medically ill	The effect of treating depression on hemoglobin A_{1c} levels in diabetic patients	The effect of treating depression after coronary artery bypass graft surgery on the recurrence of coronary artery disease The effect of intensive depression reduction on the occurrence of retinopathy, neuropathy, and other long-term complications of diabetes
Development and study of "designer" psychiatric treatments based on an understanding of the causation of psychiatric morbidity in the complex medically ill	None	The potential benefits of deep brain stimulation for depression in the medically ill. The potential benefits of amyloid or tau reducing therapy on the occurrence of Alzheimer's associated mood symptoms
Service delivery studies		
Studies of the delivery of efficacious psychiatric treatments for the complex medically ill across different service settings, such as primary care, home-based care, assisted living, nursing homes, and other long-term care settings	Use of collaborative care strategies in primary and specialty care settings	The benefits of implementing "dementia care" interventions in primary care and assisted living settings Development of new ways to deliver manualized psychotherapies in primary care Cost-effectiveness of integration of

The origins of consultation-liaison psychiatry fellowships can be traced to George Engel's and John Romano's fellowship program in psychosocial skills at Rochester University College of Medicine. This program, first offered in 1946, was offered to nonpsychiatric physicians as well as psychiatrists. The Massachusetts General Hospital established its first consultation service in 1954.

In 2003, the American Board of Medical Specialists approved psychosomatic medicine as a subspecialty field of psychiatry. Historically, psychosomatic medicine was known as consultation-liaison psychiatry, and it represented the care delivered by psychiatrists to patients with co-occurring medical and psychiatric problems who were treated primarily in medical settings. The approval of the subspecialty was both a response to and an opportunity for the field to highlight and expand the foundation in clinical care, research, teaching, and other scholarly activities related to the care of this special patient population. Although psychosomatic medicine specialists do not provide all, or even most, of the care to this patient population, the presence of the subspecialty clearly has had a positive impact on general training, research, and academic psychiatry.

Psychosomatic medicine, as defined and characterized by the subspecialty, is practiced by psychiatrists. The Academy of Psychosomatic Medicine has over 1,200 members and approximately 3,000 psychiatrist members of the APA state that they are interested in or spend a substantial amount of time providing consultation in a general hospital setting. Other mental health professionals also provide care for this population. Psychologists have long served in the general medical setting providing assessments, psychotherapeutic and behavioral interventions, and conducting research. The health psychology section of the American Psychological Association has over 3,500 members. In addition, social workers and nurses participate in care for these patients. In many settings, psychosomatic medicine psychiatrists function as part of multidisciplinary teams, which include psychiatry, psychology, nursing, and social work. In most cases, these teams work within medical specialty settings that provide care to special populations such as those with HIV, cancer, cardiovascular disease, transplant, and diabetes.

The APA formed the Council on Psychosomatic Medicine in 2004 in response to the approval of the subspecialty. The APA Council and the Academy of Psychosomatic Medicine, as well as other organizations, have participated actively in the development of clinical standards, addressed issues related to access to care and reimbursement, and provided education to members, trainees, medical professionals, and the public. Care of special populations such as those with cancer, HIV, transplant, and obstetrics and

gynecology is the focus of other organizations (e.g., American Psychosocial Oncology Society and North American Society for Psychosocial Obstetrics and Gynecology). Research is a core interest for some organizations, including the American Psychosomatic Society. In addition to the organizations in the United States, there has been growing interest in the field internationally (e.g., The European Association of Psychosomatic Medicine).

Education and training in psychosomatic medicine are integrated throughout the psychiatry curriculum in medical school and general psychiatric training. Starting in 1961, the Accreditation Council for Graduate Medical Education (ACGME) addressed the need for psychiatric residents to have contact with other services besides psychiatry, including medicine, pediatrics, and neurology. By the 1980s, the ACGME's Residency Review Committee (RRC) required all psychiatry residency programs to provide teaching and experience in consultation liaison psychiatry. The current RRC requirements require at least 2 months of full-time work in which residents consult under supervision on other medical and surgical services.

Currently, there are 58 ACGME-accredited psychosomatic medicine fellowship programs in the United States, an increase of 50 percent in the past 8 years. The ACGME program requirements for psychosomatic medicine fellowships can be found on their website (www.acgme.org). Table 27.1–4 lists the clinical activities in which psychosomatic medicine fellows must demonstrate competence. In addition, there are also 36 residencies that combine psychiatry with internal medicine, family practice, pediatrics, and neurology.

ACCESS TO PSYCHOSOMATIC MEDICINE CARE: IMPLICATIONS FOR HEALTH POLICY

Despite an abundance of evidence that a variety of psychiatric and psychosocial interventions have positive economic and clinical outcomes for patients with co-occurring psychiatric and medical problems, access to and reimbursement for psychiatric and other mental health services are severely limited. Reimbursement for these services is complex due to the separation of mental health and physical health benefits in most private and public insurance coverage. Even in areas where there is reimbursement for mental health services, it is often administered through a mental health carve-out that separates reimbursement for these services from the medical care reimbursement, resulting in patients not being able to receive care in the medical setting or have access to professionals with the appropriate expertise to treat psychiatric conditions in the context of complex medical illnesses.

Table 27.1–4.
Patient Care Experiences in Psychosomatic Medicine Required of All Psychosomatic Medicine Fellows

1. Evaluation of psychiatric complications of medical illnesses
2. Evaluation of psychiatric complications of medical treatments, including medications, traditional and new surgical or medical procedures, transplantation, and experimental therapies
3. Evaluation of typical and atypical presentations of psychiatric disorders that are due to medical, neurological, and surgical illnesses
4. Evaluation and management of delirium, dementia, and psychiatric disorders due to medical illness
5. Evaluation and management of acute pain, chronic pain, and somatic symptom disorders
6. Evaluation and management of palliative care and end-of-life issues
7. Evaluation and management of issues in adjusting to the emotional stresses of medical illness
8. Assessment of capacity to give informed consent for medical and surgical procedures in the presence of cognitive impairment
9. Provision of psychosocial interventions, including psychotherapeutic interventions, appropriate for the medically ill
10. Appropriate use of psychoactive medication in medical, neurologic, obstetrical, and surgical conditions
11. Assessing and managing suicidality and other high-risk behavior in the medical setting
12. Interactions between psychotropic medications and the full-range of medications used for a variety of medical and surgical conditions
13. Collaboration with other physicians and other members of the multidisciplinary treatment team
14. Teaching other physicians and other members of the multidisciplinary team how to recognize and respond to various psychiatric disorders
15. Leading an integrated psychosocial health care team in the medical setting

Adapted from ACGME Program Requirements for Graduate Medical Education in Psychosomatic Medicine. Available at http://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/409_psychosomatic_med_2016_1-YR.pdf

Recent developments, largely driven by components of the Affordable Care Act (ACA), may provide support to evidence-based psychiatric interventions in the medical setting. The ACA requires mental health be included in the "essential benefit" and therefore extends the reach of mental health parity across the majority of health plans. It also includes a number of components that promote mental health care in primary care and other medical settings, such as the patient centered medical home, which includes requirements for behavioral health screening and interventions. In addition, new financing arrangements such as accountable care organizations (ACOs), value based purchasing and pay for performance create incentives for organizations to incorporate mental health care more effectively. In the ACO setting, where there is payment for a population rather than on a fee for service basis, psychiatrists who provide care across the health system can demonstrate both economic and clinical value and can be paid accordingly. However, there is still a need to better characterize these economic models and define the role of the psychiatrist in quality measures in order to assure full access to these

services. Recent attention to these types of issues by the APA, including a commissioned Milliman report, highlighted the economic impact of integrated medical-behavioral health care. Expanded use of collaborative care models, including adoption of reimbursement strategies to pay for the components of care, illustrates some of the increased interest in the field. In summary, there has been an increasing call to action to improve mental health care for the medically ill and there continues to be room for improving current delivery systems to achieve that aim.

Editor's note: The founding editor of this textbook, Harold I. Kaplan, M.D., had a long-standing interest in psychosomatic medicine and the history of psychiatry. He wrote the section on the history of psychosomatic medicine for the first edition of this textbook and for each subsequent edition thereafter until his death in 1998. This chapter includes some of his original work.

References

American Psychiatric Association, Academy of Psychosomatic Medicine. Dissemination of integrated care within adult primary care settings: the collaborative care model. Report, 2016. https://www.psychiatry.org/psychiatrists/practice/professional-interests/integrated-care/collaborative-care-model

*Ader R, ed. *Psychoneuroimmunology*. 4th ed. New York: Elsevier; 2007.

Alexander F. *Psychosomatic Medicine: Its Principles and Application*. New York: W.W. Norton; 1950.

*Blumenfield M, Strain JJ. *Psychosomatic Medicine*. Philadelphia, PA: Lippincott Williams & Wilkins; 2006.

Cannon WB. *The Wisdom of the Body*. New York: W.W. Norton; 1932.

Deutsch F. The choice of organ in organ neurosis. *Int J Psychoanal*. 1939;20:1.

Dunbar F. *Emotions and Bodily Changes*. New York: Columbia University Press; 1954.

Eisenberg L. The social construction of the human brain. *Am J Psychiatry*. 1995;152:1563.

Engel GL. The need for a new medical model. *Science*. 1977;196:129.

Freud S. Fragment of an analysis of a case of hysteria. In: *Standard Edition of the Complete Psychological Works of Sigmund Freud*. Vol 7. London: Hogarth Press; 1953:40.

Friedman M, Byers S, Rosenman RH. Coronary-prone individuals (type A behavior pattern): some biochemical characteristics. *JAMA*. 1970;2:1030.

Gitlin DF, Levenson JL, Lyketsos CG. Psychosomatic medicine: a new psychiatric specialty. *Acad Psychiatry*. 2004;28:4.

Gluhm S, Goldstein J, Loc K, Colt A, Liew C, Corey-Bloom J. Cognitive performance on the Mini-Mental State Examination and the Montreal Cognitive Assessment Across the Healthy Adult Lifespan. *Cogn Behav Neurol*. 2013;26(1):1–5.

Grinker R, Robbins F. *Psychosomatic Case Book*. Philadelphia, PA: Blakiston; 1954.

Holmes TH, Rahe RH. The social readjustment rating scale. *J Psychosom Res*. 1967;11:213.

Horney K. *The Neurotic Personality of Our Time*. New York: W.W. Norton; 1937.

Institute of Medicine. *Improving the Quality of Health Care for Mental and Substance-Use Conditions*. Washington, DC: Institute of Medicine; 2005.

*Kaplan HI. History of psychosomatic medicine. In: Sadock BJ, Sadock VA, eds. *Kaplan and Sadock's Comprehensive Textbook of Psychiatry*. 8th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2005:2105.

Katon WJ, Lin EHB, Von Korff M, et al. Collaborative care for patients with depression and chronic illnesses. *N Engl J Med*. 2010;363:2611–2620.

Keefe FJ, Lumley MA, Buffington ALH, Carson JW, Studts JL. Changing face of pain: evolution of pain research in psychosomatic medicine. *Psychosom Med*. 2002;64:921.

Levenson JL. *The American Psychiatric Publishing Textbook of Psychosomatic Medicine: Psychiatric Care of the Medically Ill*. 2nd ed. Washington, DC: American Psychiatric Publishing; 2010.

Levenstein S. Bellyaching in these pages: upper gastrointestinal disorders in psychosomatic medicine. *Psychosom Med*. 2002;64:767.

Lichtman JH, Froelicher ES, Blumenthal JA, et al. Depression as a risk factor for poor prognosis among patients with acute coronary syndrome: systematic review and recommendations. *Circulation*. 2014;129:1350–1369.

Lipowski ZJ. Review of consultation psychiatry and psychosomatic medicine: I. General principles. *Psychosom Med*. 1967;29:153.

*Lipsitt DR. Consultation-liaison psychiatry and psychosomatic medicine: the company they keep. *Psychosom Med*. 2001;63:896.

Melek SP, Norris DT, Paulus J. Economic impact of integrated medical-behavioral healthcare: implications for psychiatry. Milliman, Inc. American Psychiatric Association Report. 2014, monograph.

Novack DH. Realizing Engel's vision: psychosomatic medicine and the education of physician-healers. *Psychosom Med*. 2003;65:925.

*Shorter E. *From Paralysis to Fatigue: A History of Psychosomatic Illness in the Modern Era*. New York: Free Press; 1992.

Stern TA, Fricchione GL, Cassem NH, Jellinek MS, Rosenbaum JF. *Massachusetts General Hospital Handbook of General Hospital Psychiatry*. 6th ed. Philadelphia, PA: Saunders; 2010.

▲ 27.2 Cardiovascular Disorders

PETER A. SHAPIRO, M.D., AND ADAM R. CRITCHFIELD, M.D.

Cardiovascular disorders are the leading cause of death in the United States and the industrialized world, and ischemic heart disease is one of the leading causes of worldwide disability burden. Psychiatric problems contribute to the development of cardiovascular disease and develop as complications of it. Comorbidity poses a treatment challenge for cardiology and psychiatry. Behavioral risk factors, such as smoking, failure to exercise, and failure to adhere to treatment and lifestyle recommendations, are clearly exacerbated by psychological stress, depression, and anxiety, and may benefit from psychiatric treatment. Moreover, psychiatric disorders are common in many cardiac patients and require intervention not only to

improve adherence to and outcomes of medical therapy, but also to relieve suffering due to the psychiatric illnesses themselves. From an epidemiological standpoint, the most important categories of cardiovascular disease include coronary artery disease (CAD), congestive heart failure (CHF), arrhythmias, and stroke; the chapter also includes some discussion of other conditions of interest. Here, brief presentation of background information on these cardiovascular conditions is accompanied by discussion of psychological conditions and psychiatric disorders as risk factors for their development and course. The discussion then turns to psychiatric comorbidity with established cardiovascular disease and relevant treatment issues.

CORONARY ARTERY DISEASE

Definition and Comparative Nosology

Atherosclerotic CAD is a progressive illness characterized by the formation of atherosclerotic plaque in one or more of the coronary arteries. Angina pectoris is the major symptom of coronary disease. Acute episodes of illness occur owing to formation of thrombus on the surface of a disrupted atherosclerotic plaque, leading to total or subtotal occlusion of the affected coronary artery, abrupt myocardial ischemia, and, if not rapidly corrected, myocardial cell death. The resulting acute coronary syndromes—unstable angina, myocardial infarction (MI), and cardiac arrest—lead to death or progressive disability due to recurrent pain, cardiac arrhythmias, and CHF. Atherosclerosis generally begins early in life but is not clinically evident until mid-adult life, and behavioral factors clearly affect the development of CAD.

Epidemiology

Approximately one-third of all adults older than 35 years of age ultimately die of cardiovascular disease, most often of complications of CAD. In the United States, the incidence of CAD is more than 5 million cases per year, and more than 600,000 persons per year have a first MI. One-fifth of survivors of an acute infarction die within 1 year. Public health education campaigns and more effective treatments have reduced CAD mortality, but not its incidence. Established risk factors for coronary disease include family history, male sex, hypertension, hyperlipidemia, diabetes, sedentary lifestyle, obesity, and smoking. Premenopausal women have reduced risk compared to men, but the incidence of clinically significant disease rises after menopause to match that of men. A variety of additional metabolic, nutritional, and inflammatory factors have been identified as risk factors for incident coronary disease. Psychological factors have increasingly been recognized to contribute to the risk of incident coronary disease and for disease progression. Notably, depression is independently associated with

an almost doubled risk of the development of CAD.

Diagnosis and Clinical Features

The prototypical feature of CAD is chest pain, typically on exertion, but often there are no symptoms until an acute coronary event occurs. The typical presentation of myocardial ischemia includes substernal chest pain, often described as pressure or burning, with radiation to the shoulders, back, neck, jaw, or left arm; diaphoresis; nausea; light-headedness; or palpitations; and occurs with or after exertion, eating, or psychological arousal. Many patients, especially women (for unknown reasons), have atypical presentations, often without chest pain but with abdominal pain, fatigue, anxiety, shortness of breath, or dizziness as the presenting complaint. Characteristic abnormalities may be evident on the resting electrocardiogram (ECG), such as ST segment depression, left bundle branch block or left ventricular hypertrophy, or, in acute coronary syndromes or completed MI, ST segment elevation, pathological Q waves, and T wave inversions, but the resting ECG is often normal. Further diagnostic evaluation typically may include exercise stress testing, with or without echocardiography or radioisotope imaging, and cardiac catheterization. The diagnosis of acute MI generally rests on a combination of acute symptoms or characteristic ECG changes along with new elevation in serum levels of cardiac enzymes (troponin and creatine phosphokinase [CPK]).

Differential Diagnosis

The differential diagnosis of CAD is extensive. Shortness of breath occurs in primary lung disease, CHF without significant coronary disease, pericarditis, and psychiatric disorders including anxiety and somatic symptom disorders. Chest pain occurs in pericarditis, peptic ulcer disease, gastroesophageal reflux disease, hiatal hernia, functional heartburn, and esophagitis; in pulmonary embolism, pleuritis, pleural effusion, and pneumonia; after rib fracture; and also, again, in psychiatric disorders. Often, no diagnosis can be established to account for chest pain, and musculoskeletal pain or costochondritis is invoked, especially if the pain can be reproduced by manual pressure on the costochondral junction. Palpitations occur in cases of CAD but also in mitral valve prolapse, in patients with arrhythmias due to other causes, in hyperthyroidism, and in many patients with somatic symptom preoccupation, panic, and anxiety disorders. Ectopic beats also occur in normal individuals with no evidence of medical or psychiatric disorder.

Course and Prognosis

One-third of patients experiencing their first MI die within 1 hour of the

onset of symptoms. Evolving trends in coronary care have reduced the short-term mortality of patients with acute coronary syndrome who survive long enough to be hospitalized. In survivors, recurrent cardiac events occur with an incidence of 10 to 20 percent per year, and the 1-year mortality after MI is 20 percent. Current practice in the treatment of acute coronary syndromes attempts urgent thrombolysis or revascularization by angioplasty, stent placement, or coronary artery bypass graft (CABG) surgery; β-adrenergic blockade; statin therapy to stabilize plaque; and aspirin or antiplatelet drugs.

Treatment

Preventive measures in treatment of coronary disease include the use of aspirin as an antiplatelet drug, exercise, lipid-lowering therapy with statin drugs, abstinence from or cessation of smoking, blood pressure control, and maintenance of normoglycemia. Standard pharmacotherapy for CAD generally includes β-adrenergic blockade, aspirin, statins, and angiotensin-converting enzyme (ACE) inhibitors or angiotensin receptor blockers, as well as primary preventive measures. Complications of acute or progressive disease, such as arrhythmias and CHF, must also be treated.

Psychological Factors Affecting Coronary Artery Disease

Depression, vital exhaustion, anxiety, type A behavior, hostility, anger, and acute and chronic mental stress have been evaluated as risk factors for the development and expression of coronary disease. Negative affect in general, low socioeconomic status, and low social support have been shown to have significant relationships with each of these individual psychological factors, and some investigators have proposed these latter characteristics as more promising indices of psychological risk. Data from the Normative Aging Study demonstrate a dose–response relationship between negative emotions (a combination of anxiety and depression symptoms) and incident coronary disease. Posttraumatic stress disorder (PTSD) is increasingly recognized as a risk factor for heart disease.

Depression. Depression is strongly associated with increased risk for the development and progression of CAD. Numerous large-scale, prospective, epidemiological studies of nonclinical community-dwelling subjects yield converging estimates of increased relative risks of incident MI and cardiac-related mortality of approximately 1.5 to 2.0 in association with depression. This finding holds even after controlling for smoking, a potent risk factor for cardiovascular morbidity and mortality that is far more prevalent in those with depression than in the population at large, and the effect of depression continues to be present even in long-term follow-up over 10 to 20 years. Meta-analyses of numerous studies of

patients with pre-existing CAD also demonstrate a doubling of risk of adverse coronary disease-related outcomes, including MI, revascularization procedures for unstable angina, and death, in association with depression. A study of 1,002 stable outpatients with coronary disease showed that depression more strongly predicts impairment in quality of life and physical functioning than reduced ejection fraction or exercise-induced myocardial ischemia. A dose–response relationship appears to exist between the severity of depression symptoms after acute MI or unstable angina and the risk of death over 5-year follow-up, even after controlling for other significant prognostic factors (Fig. 27.2–1), an observation replicated in many studies. Depression immediately after CABG surgery, marked depression 6 months after CABG surgery, and persistence of even moderate depression symptoms beginning before surgery at 6-month postoperative follow-up, all predict increased risk of death over 5- to 10-year follow-up.

Major adverse cardiac events associated with depression in patients with established CAD include recurrent infarction and sudden cardiac death. The co-occurrence of depression and frequent premature ventricular contractions after MI appears to substantially increase the risk of ventricular tachyarrhythmias; the mortality associated with these arrhythmic events may be reduced in contemporary cardiology practice because of the increased use of implantable defibrillators for both secondary and primary prevention of tachyarrhythmia-related sudden cardiac death. Although several trials have failed to demonstrate a significant effect of treatment of depression on mortality and cardiac outcomes, a recent meta-analytic review found that serotonin reuptake inhibitor treatment does modestly reduce the risk of recurrent cardiac events and mortality.

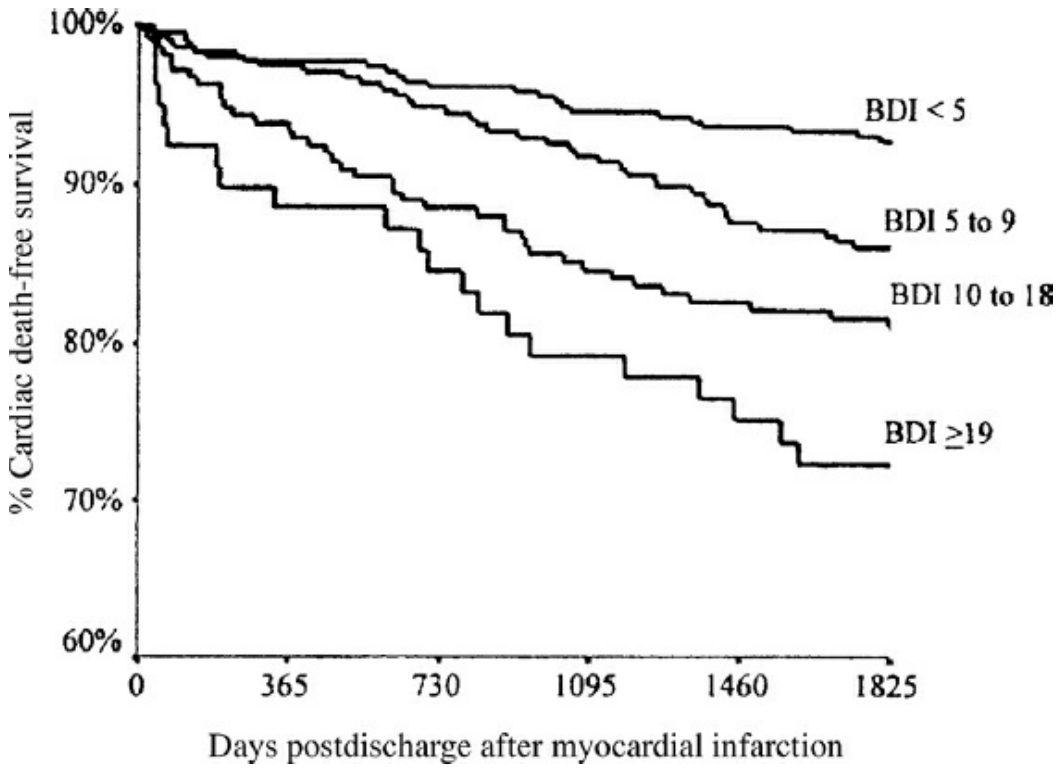

FIGURE 27.2–1. Long-term survival after myocardial infarction in relation to Beck Depression Index (BDI) score during hospitalization. (From Lesperance F, Frasure Smith N, Talajic M, Bourassa MG. Five-year risk of cardiac mortality in relation to initial severity and one-year changes in depression symptoms after myocardial infarction. *Circulation*. 2002;105:1049, with permission.)

Conversely, poor response to antidepressant treatment after an MI is associated with higher risk of subsequent cardiac events and death. Secondary analyses of three large trials of depression interventions in patients with CAD have found that patients who do not experience improvement in depression symptoms have significantly higher risk of cardiac death than those whose depression symptoms improved.

POSSIBLE MECHANISMS. The mechanisms by which depression may increase coronary disease risk have not been established, but potential mechanisms may be organized into behavioral and biological pathways. Behavioral mechanisms include social isolation, physical inactivity, poor adherence, and smoking. The relationship between cigarette smoking and mental illness, particularly depression, has been well documented. History of depression is associated with lifetime risk of smoking, current smoking, and lower successful quit rate. As the prevalence of smoking trends downward in the United States, these associations have strengthened. Biological pathways include links of depression with abnormal neuroendocrine and autonomic function, smoking, and platelet dysfunction.

Hypothalamic–pituitary–adrenal (HPA) axis dysregulation in depression is manifest by elevated levels of circulating cortisol and loss of

the usual diurnal variation in cortisol level. Elevated cortisol has toxic effects on the coronary artery endothelium and plays a role in plaque development.

Autonomic dysregulation with diminished cardiac vagal modulation and increased sympathetic nervous system activation occurs in depression. This may provide a substrate for increased arrhythmic activity and sudden death. Heart rate variability, an index of cardiac autonomic control, is reduced in depression, and reduced variability has been shown to be a predictor of sudden death in cardiac patients.

Disordered platelet aggregation, leading to increased thrombus formation, may also play a role in increasing risk of coronary events in depression. Measures of in vivo platelet activation and platelet aggregation show that depressed patients exhibit greater prothrombotic activity at baseline and greater platelet activation by orthostatic challenge. Serotonergic antidepressants, especially those with high affinity for the serotonin reuptake transporter, reduce platelet activation, and several studies indicate that they reduce the likelihood of first MI.

Type A Behavior Pattern, Anger, and Hostility. The relationship between a behavior pattern characterized by easily aroused anger, impatience, aggression, competitive striving, and time urgency (type A) and coronary heart disease dominated studies in psychosomatic cardiology in the 1970s and 1980s. Several large prospective epidemiological studies found the type A pattern to be associated with a nearly twofold increased risk of incident MI and coronary disease-related mortality. Hostility as a core component of the original type A concept has received considerable empirical support as a predictor of coronary heart disease outcomes. Low hostility is associated with low coronary disease risk in studies of workplace populations. High hostility is associated with increased risk of death in 16-year follow-up of survivors of a previous MI. In addition, hostility is associated with several physiological processes that, in turn, are associated with coronary disease, such as reduced parasympathetic modulation of heart rate, increased circulating catecholamines, increased coronary calcification, and increased lipid levels during interpersonal conflict. Conversely, submissiveness has been found to be protective against coronary disease risk in women. Adrenergic receptor function is down-regulated in hostile men, presumably an adaptive response to heightened sympathetic drive and chronic overproduction of catecholamines due to chronic and frequent anger. A recent meta-analysis of psychological interventions for patients with CAD found that those that targeted type A behavior had the strongest effects on cardiovascular outcome. A 2011 randomized clinical trial in post-MI patients found that a 1-year, 20-session group cognitive-behavioral therapy (CBT) stress management program that included a strong emphasis on reduction in anger dramatically reduced

fatal and nonfatal recurrent cardiovascular events, recurrent MI, and all-cause mortality over an 8-year follow-up period. This finding is consistent with prior studies showing that group therapy for type A behavior modification was associated with reduced reinfarction and mortality in post-MI patients and that type A behavior modification therapy reduces episodes of silent myocardial ischemia.

Vital Exhaustion. *Vital exhaustion*, a state of fatigue associated with loss of energy, increased irritability, and demoralization (e.g., decreased interest in work, pessimism, and burnout) has been identified as a risk factor for incident coronary disease and cardiac events in patients with CAD in European, predominantly Dutch, studies. Because the vital exhaustion concept overlaps with depression, its independent contribution remains uncertain.

Acute Mental Stress. Mental stress induces arterial endothelial dysfunction, with impaired flow-mediated vasodilatation; paradoxical vasoconstriction during stress occurs in atherosclerotic arterial segments. Acute mental stress has a significant effect on coronary artery blood flow that may be of significance in patients with pre-existing coronary disease. States of fear, excitement, and, especially, acute anger reduce blood flow through atherosclerotic coronary segments, provoke coronary spasm, and are associated with abnormal left ventricular wall motion and ECG evidence of myocardial ischemia. Mental stress-induced ischemia is associated with increased risk of subsequent cardiac events in patients with known coronary disease and may occur even in patients who do not demonstrate evidence of ischemia during exercise stress testing. Relaxation training can alter autonomic activation during mental stress, implying a potential therapeutic role for such training in stress-induced ischemia. There is also evidence that sertraline treatment reduces myocardial ischemia induced by mental stress.

A Clinical Vignette

A 55-year-old man with a history of coronary disease and major depressive disorder reported daily episodes of angina associated with anger provoked by interpersonal conflicts. Angina ceased after he began treatment with sertraline.

Triggering of acute cardiac events, including acute MI and sudden cardiac death, by acute mental stress has been demonstrated in epidemiological studies of earthquakes and other natural disasters, sporting events such as World Cup football matches, bombings, and acts of terrorism such as the destruction of the World Trade Center in 2001. Ecological momentary assessment studies with periodic ratings of mood, stress, and affect correlated with ambulatory ECG recordings, and

retrospective reviews of events preceding acute coronary events also indicate that acute emotional stress is a frequent trigger of MI. Between 1 and 4 percent of MI and between 20 and 40 percent of sudden cardiac deaths may be precipitated by acute emotional stressors.

Chronic Mental Stress. Chronic mental stress, such as job or marital strain, also contributes to the development and the progression of CAD. One study of over 9,000 British civil servants found that persistently difficult close social relationships independently contributed to the risk of new CAD. In a study of over 900 men and women who returned to work after their first MI, job strain (high demand plus low decision latitude) independently doubled the risk of recurrent cardiac events in the next 6 years. In the largest case control study of predictors of MI, the INTERHEART study found that among 11,119 cases of MI in 52 countries, psychosocial factors, defined as a combination of stress and depression, ranked as the third-highest predictor of MI, raising the risk of MI by an odds ratio of 2.67, similar to smoking and diabetes.

Posttraumatic Stress Disorder. The impact of PTSD on the development and course of cardiovascular and cerebrovascular disease has become a topic of considerable recent investigation, along with current recognition of the high prevalence of PTSD in recent combat veterans. Military veterans with PTSD have increased risk of the development of CAD, heart failure, and peripheral vascular and cerebrovascular disease. The effects remain significant after adjustment for smoking, depression, and other CAD risk factors. Both behavioral and physiological processes are likely to be involved as mediators of risk of cardiovascular disease, including smoking, difficulty adhering to healthy lifestyle, and autonomic dysfunction with abnormal sympathetic nervous system activation. Even after adjustment for medication adherence, PTSD is associated with increased morbidity and mortality.

Schizophrenia and Bipolar Illness. Persons with severe mental disorders such as bipolar disorder, schizophrenia, and schizoaffective disorder tend to die at a younger age than the general population, with doubling of mortality risk due to cardiovascular disease. In bipolar disorder, notwithstanding the high rate of death by suicide and accidents, cardiovascular disease is the leading cause of death. Onset of cardiovascular morbidity is earlier in people with bipolar disorder at least in part as a result of higher rates of unhealthy behaviors, common medication side effects, and problems with access to medical care.

Individuals with severe mental illness are more likely to engage in several behaviors that are known to increase the risk of cardiovascular disease. Smoking is highly prevalent in individuals with severe mental

illness, which leads to vascular endothelial damage and systemic inflammation. Negative symptoms in patients with schizophrenia are associated with physical inactivity that itself raises cardiovascular disease risk and is related to poor diet and obesity. For patients with bipolar and schizoaffective disorder, hypomanic and manic episodes are associated with the increased risk of substance use, which in turn increases cardiovascular disease risk. Interestingly, the association between episodes of mania and hypomania and cardiovascular risk may not depend solely on altered behavior; lipid metabolism in individuals with bipolar disorder is an area of ongoing research into a potential link between the pathophysiology of bipolar disorder and cardiovascular disease.

Various medications used to treat persons with severe mental illness can contribute to cardiovascular disease risk. Many second-generation antipsychotics, including clozapine, olanzapine, and risperidone, are associated with rapid weight gain that can lead to metabolic syndrome (hypertension, dyslipidemia, increased abdominal fat, insulin resistance, and hyperglycemia). Mood stabilizers such as valproic acid and lithium are also associated with weight gain.

It has been noted that patients with serious mental illness who experience an acute MI are less likely than the general population to receive appropriate treatment, including both medications and revascularization procedures. Most individuals with bipolar disorder and schizophrenia receive inadequate medical care to mitigate their increased risk of cardiovascular morbidity and mortality. Those with severe mental illness are less likely to have a primary care physician and more likely to have difficulty maintaining continuous insurance coverage. Despite numerous practice guidelines with agreed-upon recommendations for screening of modifiable cardiovascular risk factors, many patients are not screened or treated adequately. For example, among the patients with chronically treated schizophrenia entering the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) study, 88 percent of patients with dyslipidemia, 62 percent of patients with hypertension, and 30 percent of patients with diabetes were not being treated with medications for these conditions. These data underscore the need for better coordination between mental and physical health services, potentially through systems of integrated care and/or colocalized psychiatric and medical services. Baseline assessments for these patients should include a thorough medical history, anthropometric assessments, thyroid testing, fasting plasma glucose, fasting blood lipid panel, baseline ECG, and resting blood pressure.

CARDIAC ARRHYTHMIAS AND SUDDEN CARDIAC DEATH

Definition and Comparative Nosology

Among the many subtypes of cardiac arrhythmia, of greatest importance to psychiatrists are sinus node dysfunction and atrioventricular (AV) conduction disturbances resulting in bradyarrhythmias, atrial fibrillation, and tachyarrhythmias. Bradycardia, defined by a heart rate of less than 60 beats per minute, may occur in normal subjects, especially trained athletes, but also occurs pathologically in hypothyroidism, severe liver disease, and vasovagal syncope, and during a variety of metabolic disturbances. Idiopathic primary sinoatrial node degeneration and pharmacologically induced dysfunction are the most common causes of bradycardia. Symptomatic patients typically present with dizziness or syncope. Bradyarrhythmias also occur in patients with heart block due to impaired conduction through the AV conduction system, especially AV nodal block. Sinus node dysfunction including sinus arrest is a known adverse effect of lithium and can result in clinically significant bradycardia and syncope. Selective serotonin reuptake inhibitors (SSRIs) may impair sinus node function, and at high doses tricyclic antidepressants may impair AV nodal conduction, leading to complete heart block. Serotonin reuptake inhibitors taken together with β-adrenergic blockers or vagomimetic compounds may have additive effects to slow the heart rate, sometimes exacerbated if the serotonergic agent also inhibits the hepatic metabolism of the beta-blocker, augmenting its effect.

Atrial fibrillation is the most common cardiac arrhythmia in clinical cardiology practice; it is an irregularly irregular heart rhythm characterized by disorganized depolarization of atrial myocytes, so that the AV node is stimulated in a rapid irregular manner. The ventricular response and resulting heartbeat can be extremely rapid, with palpitations and symptoms of reduced cardiac output. Disorganized atrial systole reduces the efficiency of ejection of atrial contents into the ventricles, further reducing cardiac output ("loss of 'atrial kick'"). Inefficient atrial emptying promotes thrombus formation within the atrium, with resulting fivefold increased risk of embolic stroke and twofold increased risk of death. Atrial fibrillation can occur as a consequence of hyperthyroidism, in mitral valve disease, after cardiac surgery, after MI, and in CHF. Atrial fibrillation is treated with rhythm or rate control using beta-blockers, calcium channel blockers, cardioversion, or ablation. Anticoagulation reduces the risk of embolic events but increases the risk of bleeding. Sadness, anxiety, acute stress, panic disorder, and anger-hostility have been associated with increased risk of incident atrial fibrillation; in one study, self-reported happiness was negatively associated with risk.

Ectopic atrial premature beats and ventricular premature beats are common, occurring in more than 50 percent of adults during 24-hour ECG monitoring. In isolation, they are of no pathological significance but may be perceived as flutters in the chest, skipped beats, palpitations, or abnormally vigorous beats. These sensations exist in a circular relationship with

anxiety: Awareness of unusual heartbeats often provokes vigilance and anxiety about body sensations, and anxiety is a potent stimulus for increased frequency of premature beats. Paroxysmal supraventricular tachycardia due to an ectopic atrial pacemaker frequently presents with anxiety associated with tachycardia and palpitations and may be misdiagnosed as a panic attack.

Sustained ventricular tachycardia and ventricular fibrillation are lethal arrhythmias; patients with CAD or completed MI who have frequent premature ventricular depolarizations are at risk of cardiac arrest due to one of these forms of arrhythmia. About half of all cardiac deaths are sudden cardiac deaths, and half of cases had no previously known heart disease.

Pathophysiology of Ventricular Tachyarrhythmias. Malignant ventricular arrhythmias may occur as an electrical accident in individuals with no structural heart disease but more often occur in those with underlying heart disease. Because conduction of myocardial depolarization may be delayed in regions of myocardial tissue damage, regions distal to the affected area may be stimulated to depolarization while adjacent tissue is undergoing repolarization, setting off a premature depolarization of the adjacent tissue that may precipitate ventricular fibrillation. The rate of local repolarization and readiness to initiate the next cardiac cycle is modulated by local innervation of the heart by branches of the sympathetic and parasympathetic nervous systems. In general, sympathetic fibers predominate in the left ventricular myocardium and have the effect of increasing irritability and stimulating depolarization. Parasympathetic fibers predominate in the innervation of the specialized cardiac conduction system and have the effect of slowing heart rate and reducing automaticity.

Lethal ventricular tachyarrhythmias are more likely to occur in the setting of structural heart disease (e.g., after an MI) but also are related to genetic disorders regulating ion channels that control ventricular depolarization and repolarization, such as congenital long QT syndrome and Brugada syndrome. These latter disorders are sometimes clinically silent until “unmasked” by medications including lithium, tricyclic antidepressants, and phenothiazines.

Diagnosis

Diagnosis of cardiac arrhythmias depends on the ability to record the rhythm disturbance. Because the resting ECG is most often normal or shows isolated premature contractions in patients with episodic rhythm disturbances, ambulatory ECG monitoring may be needed to establish the diagnosis. Provocative electrophysiological testing may be used to attempt to induce arrhythmias; depending on the nature of the resulting arrhythmias, if present, therapies may include ablation of the focus of

ectopic activity, implantation of an automatic cardiac defibrillator, or antiarrhythmic drug therapy. Indications for defibrillator implantation have been broadened to include not only secondary prevention for patients with previous episodes of sudden cardiac death/ventricular tachycardia/ventricular fibrillation but also primary prevention for patients with reduced left ventricular ejection fraction due to ischemic or nonischemic heart disease.

Psychological Factors Affecting Arrhythmia

Because autonomic cardiac modulation is profoundly sensitive to acute emotional stress, such as intense anger, fear, or sadness, it is not surprising that acute emotions can stimulate arrhythmias. Instances of sudden cardiac death related to sudden emotional distress have been noted throughout history in all cultures. In community samples depression, phobic anxiety and mixed anxiety symptoms increase the risk of sudden cardiac death. After the destruction of the World Trade Center in New York in 2001, patients with implanted defibrillators had sharply increased frequency of ventricular tachycardia events. In post-MI patients with ventricular ectopy, comorbid depression is associated with significantly increased rates of ventricular tachycardia–ventricular fibrillation. Psychiatric conditions treated with antipsychotic drugs and citalopram are also associated with increased risk of ventricular tachyarrhythmias, due to the adverse effect of the psychiatric medication on cardiac conduction. The effect of psychotropic drugs on the QT interval and arrhythmia risk is discussed below. Depression is associated with increased mortality in patients who have received an implanted defibrillator for the prevention of sudden cardiac death.

CONGESTIVE HEART FAILURE, HEART TRANSPLANT, AND ASSIST DEVICES

Definition and Comparative Nosology

CHF is an end result of many forms of heart disease and represents the most common final pathway toward disability and mortality. Heart failure with reduced ejection fraction (systolic dysfunction) occurs after MI. Heart failure with preserved ejection fraction (diastolic dysfunction), with impaired ventricular filling or elevated filling pressure, is increasingly recognized and may be due to infiltrative disease, hypertension with resultant chronic ventricular hypertrophy and increased ventricular wall stiffness, or other causes. Shortness of breath, reduced exercise tolerance, and peripheral edema are typical symptoms of heart failure. Hepatic congestion and cardiac cirrhosis occur as complications of elevated central venous pressure. Renal dysfunction occurs secondary to poor renal perfusion and impaired venous drainage, and may be exacerbated by diuretic-induced dehydration. Heart failure patients may struggle with

oscillation between the extremes of being "too wet" or "too dry." Treatment of heart failure patients usually includes ACE inhibitors or angiotensin receptor blockers, β-adrenergic blocking agents, spironolactone, and diuretics. Diastolic dysfunction tends to be more refractory to treatment. Heart transplantation and left ventricular assist devices (LVADs) are options for refractory cases. Apart from their evaluation as heart transplant candidates, there has been relatively little study of the psychiatric problems of heart failure patients. Clinically, it appears that adjustment disorders, depression, and anxiety are relatively common problems in this population. Disorders of alertness and cognition occur in patients with hypotension, hyponatremia, hepatic failure, and renal failure secondary to heart failure. Diagnosis of psychiatric disorders in patients with CHF requires clinical judgment, as symptoms of heart failure may confound the psychiatric presentation. Dyspnea and orthopnea, for example, may resemble panic, and orthopnea, in particular, may interfere with sleep. Hopelessness, decreased self-esteem, suicidal ideation, and other cognitive aspects of affective illness are not normally present in heart failure patients and may help make the distinction.

Epidemiology

Recent data continue to demonstrate that more than 500,000 people develop CHF per year in the United States, and heart failure is cited as the principal cause in 40,000 deaths and as a contributing factor in 250,000 deaths per year.

Etiology

Neurohormonal activation, with a chronic state of sympathetic nervous system activation, is a hallmark of the development of heart failure and promotes ventricular remodeling, the dilatation of the left ventricle associated with reduced ventricular contractility. MI, hypertension, and mitral and aortic valve disease are among the most common known causes of heart failure, but dilated cardiomyopathies have many other etiologies, including viral cardiomyopathy, toxicity of adriamycin, cocaine and other agents, genetic disorders, and inflammatory disorders, such as giant cell myocarditis. Almost one half of all nonischemic cardiomyopathy is idiopathic. Restrictive (e.g., constrictive pericarditis), obstructive (e.g., asymmetrical septal hypertrophy and hypertrophic subaortic stenosis), and infiltrative diseases (e.g., systemic amyloidosis) also result in heart failure.

Diagnosis and Clinical Features

Patients typically present with complaints of shortness of breath or dyspnea on exertion, diminished exercise tolerance, fatigue, or edema. Rales, abnormal heart sounds (especially S_3 gallop), displacement of the

point of maximal impact of the heart on the chest wall, neck vein distention, hepatic enlargement, and lower extremity edema may be present on physical examination. Echocardiography or ventriculography demonstrates ventricular dilatation (in many cases) and decreased contractility. Cardiopulmonary exercise stress testing demonstrates decreased capacity to utilize oxygen. B-type natriuretic peptide levels mark severity of physiological derangement.

Differential Diagnosis

Hypothyroidism, pulmonary disease, renal disease, and liver disease may present with symptoms resembling those of heart failure, such as shortness of breath, edema, and fatigue. Psychiatric disorders must also be included in the differential diagnosis of heart failure. Fatigue and low energy may be presenting complaints in depression and generalized anxiety disorder, and shortness of breath may be the presentation of somatoform and panic disorders.

Course and Prognosis

CHF is a progressive condition with an approximately 50 percent 5-year mortality. Classification schemes for staging of CHF are shown in Table 27.2–1. To some degree, prognosis depends on the underlying cause and the possibility of reversing the pathophysiological derangement. Some forms of infectious and inflammatory myocarditis are completely reversible. Ventricular contractility may be impaired by reversible coronary vascular insufficiency; angiography and positron emission tomography may identify patients with viable myocardium who will experience improved ventricular function after a revascularization procedure. For most patients, however, who have idiopathic dilated cardiomyopathy or MI with ventricular dysfunction, the course tends to be one of stable or gradually downhill functional status punctuated by acute episodes of fluid overload or cardiac arrhythmia, precipitating admission. Comorbid depression is associated with higher morbidity and mortality.

Treatment

Mainstays of therapy for CHF are diuretics, ACE inhibitors or angiotensin receptor blockers, and β-adrenergic blocking agents; the role of digoxin has diminished, as some studies suggest that it increases mortality. Many trials of agents with positive inotropic effects have failed to establish a survival benefit. Cardiac arrhythmias occur at increased frequency in patients on inotropic agents. Nevertheless, intravenous (IV) dobutamine (Dobutrex), milrinone (Primacor), and similar medications are sometimes prescribed for severely ill patients who would otherwise succumb to low-output states. Heart transplantation is an option limited by donor supply to a small

fraction of those who might benefit from it. In the absence of significant improvement in survival with new medications in recent years, and given the limited number of patients who can receive transplants, the survival benefit provided by ventricular assist devices has led to their increased use not only for "bridge to transplantation" but also as "destination" therapy.

Table 27.2–1.
Heart Failure Classification

New York Heart Association	American Heart Association
Functional classification	Prognostic classification
Class I: Asymptomatic	Stage A: Presence of risk factors
	Stage B: Structural heart disease without symptoms
Class II: Mild symptoms with moderate exertion	Stage C: Structural heart disease with current symptoms or previous symptoms that have responded to therapy
Class III: Mild exertion limited by symptoms	
Class IV: Symptoms at rest	Stage D: Symptoms are refractory to conventional treatment and require assist device, transplant or palliative care

Heart Transplantation. Heart transplantation is available to between 2,000 and 2,500 patients annually in the United States. It provides approximately 75 percent 5-year survival for patients with severe heart failure, who would otherwise have a less than 50 percent 2-year survival. Candidates for heart transplantation typically experience a series of adaptive challenges as they proceed through the process of evaluation, waiting, perioperative management, postoperative recuperation, and long-term adaptation to life with a transplant. These stages of adaptation typically are accompanied by anxiety, depression, elation, and working through of grief. Mood disorders are common in transplant recipients.

Left Ventricular Assist Devices. LVADs are approved as a bridge to heart transplantation and also for long-term treatment of CHF. First- and second-generation devices consist of a conduit that drains the left ventricle of the heart, connected to a mechanical pump implanted in the abdomen, which, in turn, drains into a conduit connected to the ascending aorta. Some newer devices use a small axial flow pump implanted inside the pericardium that drains the left ventricle and empties into the ascending aorta. The power for the pump is supplied by an external power source, such as a battery that can be worn on a vest, permitting the patient to ambulate freely. It has not yet been possible to produce a totally implantable device with no need to traverse the skin. Patients with LVADs may recover from the clinical complications of left ventricular functional impairment, with improved renal, hepatic, cerebral, and peripheral circulation leading to improved exercise tolerance, functional capacity, and

general well-being. They are subject to complications, such as bleeding, infection, anorexia, and stroke. Depression and organic mental syndromes occur frequently. Prior cerebrovascular disease appears to be a risk factor for psychiatric and neuropsychiatric complications in patients with LVADs. Caregiver burden is substantial for family members of assist device patients when they are discharged from the hospital. The neuropsychiatric and quality-of-life effects of VAD therapy depend on pump design and are likely to continue to evolve over the near future.

Psychological Factors Affecting Congestive Heart Failure

Relatively little attention has been paid to the role of psychological factors in heart failure. However, depression has been recognized as a risk factor in heart failure incidence and mortality, and the role of psychological factors in outcome after heart transplantation has been explored. Poor self-care related to depression has been highlighted as a factor in poor outcome in heart failure, along with fatigue, denial, lack of knowledge, and impaired cognitive function.

Depression and Prognosis in Congestive Heart Failure. In a prospective, longitudinal study of 4,538 hypertensive patients older than 60 years of age, with a 4.5-year follow-up, CHF developed in 8.1 percent of patients with high depression symptoms versus 3.2 percent of patients with low depression symptoms. Adjusting for demographic and medical variables predicting development of heart failure, including incident MI during the follow-up period, depression was associated with almost threefold increased risk of the development of heart failure.

Several studies, involving more than 1,000 patients in total, have demonstrated the substantial prevalence and prognostic significance of depression symptoms and major depressive disorder in patients with CHF. In one study, comorbid major depression increased 1-year mortality more than threefold in patients with New York Heart Association class III and IV heart failure. In a second study, 14 percent of hospitalized patients with CHF met criteria for major depression, and 39 percent had Beck Depression Index scores of greater than 10. One-year mortality was 16 percent; major depression was associated with more than doubling of mortality at 3-month and 1-year follow-up, even after adjusting for other prognostic factors. A retrospective study with 5-year follow-up of 396 patients with CHF due to nonischemic dilated cardiomyopathy identified 21 percent as clinically depressed; of these, 60 percent were receiving antidepressants. Mortality at 5 years was 36 percent in depressed patients versus 16 percent in nondepressed patients, with a statistically significant, threefold increased risk of death for the depressed patients after adjusting for other factors. As with CAD, depression hastens the development of and the mortality associated with CHF. In a randomized trial, sertraline

treatment of depression in patients with CHF did not result in improved depression or survival; in a smaller study, paroxetine was superior to placebo for improvement in depression symptoms. In a recent trial, cognitive behavioral therapy improved depression but not heart failure self-care in a heart failure patients with depression. A recent animal study found that, independent of its antidepressant effects, paroxetine protects against pathological changes in left ventricular structure and function in the setting of heart failure due to MI.

Effect of Psychosocial Factors on Compliance and Survival in Heart Transplant Patients. An emerging body of evidence has demonstrated that preoperatively assessed psychosocial variables predict not only psychiatric, but also medical, outcomes after heart transplantation. Personality disorders, substance abuse disorders, dimensional measures of coping skills and social support, and clinically assessed compliance history have been linked to the increased risk of poor postoperative compliance, rejection episodes, and increased mortality. The use of psychiatric assessment of transplant candidates may identify opportunities for intervention before transplantation with subsequent improvement in outcome. Depression and PTSD after heart transplantation are associated with heightened mortality.

HYPERTENSION

Definition and Comparative Nosology

Hypertension is defined by elevation of blood pressure, typically (but arbitrarily) with a cutoff of 120 to 140 mm Hg, for systolic blood pressure, or more than 80 to 90 mm Hg, for diastolic blood pressure, on repeated measurements, at rest, over time. Blood pressure normally increases during exertion and falls during sleep ("nocturnal dipping"). Typically, hypertension is an asymptomatic chronic condition that eventually results in increased risk of CAD, MI, cardiomyopathy, renal failure, peripheral and cerebral vascular disease, and stroke.

Epidemiology

The prevalence of hypertension in the United States is high, in the range of 20 to 50 percent of adults, depending on the population examined and cutoff criteria for diagnosis employed. The highest incidence is in the third through fifth decades of life. Epidemiological risk factors for the development of hypertension include male sex, family history, nonwhite race, and high dietary sodium intake. The role of psychosocial risk factors in the development of hypertension is controversial. Anxiety, acute mental stress, chronic anger, and job strain have been associated with risk of hypertension, but epidemiological studies have not yielded consistent

associations. Acute mental stress is associated with transient increased blood pressure and underlies the phenomenon of *white-coat hypertension*, the propensity to manifest elevation of blood pressure during physical examination, although blood pressure is normal at other times; the patient's anxiety about the result of the blood pressure measurement has the effect of momentarily increasing blood pressure. For this reason, blood pressure should be recorded repeatedly to allow habituation to the situation before hypertension is considered to be present.

Etiology

More than 90 percent of patients with hypertension have idiopathic or essential hypertension. Familial and genetic influences are evident. Pathophysiologically, sympathetic nervous system activation and renin-angiotensin system activation are important determinants of elevation of blood pressure. Vasoconstriction mediated by the smooth muscle media layer of the vessel walls of small arterioles is the primary moment-to-moment regulator of blood pressure. Renal function—the regulation of sodium, potassium, and free water excretion—provides more long-term regulation of blood pressure.

Differential Diagnosis

In addition to essential hypertension, renal, vascular, and endocrine disorders are the most common causes of hypertension. Renal artery stenosis, adrenal adenomas (Cushing syndrome), and pituitary adenomas (Cushing disease) result in hypertension, with characteristic findings on neuroendocrine testing and clinical examination; in these illnesses, hypertension responds to correction of the lesion. In pheochromocytoma, paroxysmal hypertension results from surges of epinephrine secretion from the neuroendocrine tumor.

Course and Prognosis

Although hypertension is often clinically silent for decades, it may eventually lead to stroke—ischemic or hemorrhagic—cardiac disease, or renal failure. Symptoms of hypertension, when present, may include chest discomfort, headache, and dizziness.

Treatment

Mainstays of pharmacotherapy of hypertension include diuretics, ACE inhibitors and angiotensin II blockers, β-adrenergic blockers, and calcium channel blockers. Low-salt diet, weight loss, and exercise help reduce blood pressure.

Psychological Factors Affecting Hypertension

"Does tension cause hypertension?" Psychodynamic theories of hypertension from the 1950s have been discredited, and empirical studies of psychological factors contributing to the development of hypertension have yielded mixed results. Laboratory and ambulatory monitoring studies clearly demonstrate that acute mental stress results in transient elevation of systolic and diastolic blood pressure. Some studies have found that these transient hemodynamic responses predict increased blood pressure and incident hypertension over long-term follow-up. Some recent, large, population-based cohort studies, examining stress, negative affect, and anxiety as predictors of incident hypertension, have found a positive association, with 1.5- to 3-fold increased risk, while others have found none, or even a negative association. A small study of individualized stress management for patients with hypertension showed improvement in blood pressure as a result of the intervention, with reduction in blood pressure correlated with reduced stress and improved coping with anger.

SYNCOPE

Definition and Comparative Nosology

Syncope is defined as a sudden, transient loss of consciousness with associated loss of postural tone, followed by spontaneous recovery, and is due to temporary reduction in cerebral blood flow. Long-term follow-up from the Framingham Heart Study found the incidence of first report of syncope to be 6.2 cases per 1,000 person-years. Approximately 3 percent of emergency room visits and 1 to 6 percent of hospital admissions are for syncope. Mechanisms of syncope include disruption in vascular tone or inadequate blood volume, heart rhythm disorders, perfusion failure in aortic stenosis or severe pulmonary hypertension, or primary cerebrovascular insufficiency (usually vertebrobasilar insufficiency). Disruption in autonomic tone is most common and vasovagal syncope, also referred to as *vasodepressor* or *neurocardiogenic syncope*, and postural hypotension account for 30 to 50 percent of all cases of syncope. Syncope can occur as a single episode or can be recurrent and chronic. Approximately 30 to 40 percent of syncope is idiopathic. It is most important to detect bradyarrhythmias and ventricular tachyarrhythmias as an underlying cause, because these are usually associated with underlying structural heart disease and carry increased mortality risk.

Diagnosis and Clinical Features

Physical examination, including supine and standing blood pressure, and ECG are the basic examinations for patients with syncope. Abnormal ECG or structural heart disease often dictates stress testing, echocardiography, ambulatory monitoring, or electrophysiological (EP) study. In the absence of structural heart disease, the most likely cause of syncope is one of the

neurally mediated syndromes. Tilt testing with or without isoproterenol infusion or sublingual nitroglycerine helps demonstrate orthostatic hypotension and neurocardiogenic syncope, but results of tilt testing are often irreproducible, and specificity and sensitivity of tilt testing are disappointing.

Differential Diagnosis

Drop attacks, dizziness, and vertigo do not cause loss of consciousness. Seizures can be difficult to distinguish from syncope, but a preceding aura, prolonged loss of consciousness for more than 5 minutes, and rhythmic movements during loss of consciousness are characteristic of seizures. Pain, micturition, defecation, exercise, and stress as precipitating events are more characteristic of syncope than of seizures.

Course and Prognosis

Heart disease is an important prognostic indicator in syncope. In particular, patients with syncope with associated left ventricular dysfunction and CHF have significant 1-year mortality risk. Older age, ventricular arrhythmias, abnormal ECG, and heart failure contribute additive mortality risk in patients with syncope. Long-term follow-up data indicate no increase in mortality or MI risk for patients with vasovagal syncope or orthostatic hypotension but increased risk of death and MI for patients with syncope due to underlying heart disease, neurological disease, or syncope of unknown cause.

Treatment

Treatment addresses the underlying cause of syncope, when possible. For the majority of patients with idiopathic syncope or neurocardiogenic syncope, orthostatic hypotension, or vasovagal syncope, or a combination of these, behavioral interventions include instruction about avoiding precipitants and lying down when premonitory symptoms arise. Alcohol consumption, sleep deprivation, fasting, dehydration, and prolonged standing should be avoided.

Both beta-blockers and β-adrenergic agonists have been used for pharmacological treatment of syncope, but clinical trials are lacking or have shown at best equivocal efficacy. Permanent pacemakers may be the best option for patients with recurrent syncope with significant functional impairment, particularly if cardiac monitoring or tilt testing has demonstrated bradycardia. Orthostatic hypotension can be treated with compression stockings, α-adrenergic agonists, and salt supplements or fludrocortisone.

Psychological Factors Affecting Syncope

Although anxiety and acute emotional stress are recognized as precipitants of syncope, the prevalence of these factors in syncope is unknown. Anxiety and panic disorder are common in patients with recurrent syncope, but whether they were present before syncope or only after has not been well established. One study found no difference between syncope patients with and without positive tilt table tests in rates of panic and generalized anxiety disorder, but identified depression as a predictor of recurrent syncope over 3-year follow-up. It should be noted that some clinicians distinguish *hysterical fainting* from syncope by the absence of pallor, hypotension, or bradycardia during the hysterical faint, but the prevalence of this condition and its relationship to psychological factors or psychiatric diagnoses are not clear.

TAKOTSUBO CARDIOMYOPATHY

Takotsubo cardiomyopathy, also known as apical ballooning syndrome, stress cardiomyopathy, or broken-heart syndrome, is an unusual disorder, but notable on account of the frequently noted role of acute psychological stress in most cases. Cases typically present with sudden onset of chest pain and left ventricular dysfunction, do not have angiographic or enzyme evidence of an MI, and demonstrate characteristic ballooning of the left ventricular apex with echocardiography or angiography. The history typically reveals sudden psychological shock immediately preceding the event, such as receiving news of the death of a loved one, surprise birthday parties, or witnessing a traumatic event. Retrospective case-control studies suggest that premorbid social isolation, anxiety, depression, and family history of anxiety or depression are more common in patients with stress cardiomyopathy than in age- and sex-matched MI patients or community dwelling nonpatients. Left ventricular function usually recovers over a few days to weeks. The transient changes in cardiac function have been attributed to a surge in circulating catecholamines and sympathetic nervous system activation associated with the sudden psychological stressor.

CONGENITAL HEART DISEASE

Congenital heart defects occur in 1 percent of live births. Since the 1970s, advances in cardiac surgery have enabled patients who would previously have died in childhood to reach adulthood. Most of these patients have residual abnormalities of circulation, due to uncorrected problems or to surgical modification of circulation. Late complications include right to left shunts with cyanosis, sinus and AV node dysfunction, arrhythmias, heart block, valvular dysfunction, and risk of endocarditis. Ventricular dysfunction can also occur with right or left heart failure. Adjustment and developmental problems appear to be common in patients with congenital heart disease. Cognitive problems may occur as a result of congenital right-

to-left shunting with resulting systemic cyanosis and impaired cerebral oxygenation.

VALVULAR HEART DISEASE

The relationship between valvular heart disease and psychiatric disorder has been a matter of considerable interest over the past several decades. In panic disorder, mitral valve prolapse is detected in 10 to 25 percent of patients studied with echocardiography. However, prolapse also occurs in a substantial portion of the population without panic disorder, and the nature of the relationship remains uncertain. The subjective experience of valve prolapse (e.g., fluttering and chest pressure) may be a trigger for panic sensations; alternatively, the association may be purely coincidental. Obsessive-compulsive disorders (OCDs), tic disorders, and Tourette syndrome are associated with poststreptococcal immune system–mediated inflammatory responses that are similar to those leading to glomerulonephritis and rheumatic heart disease, a finding stimulated by consideration of the tic-like quality of movement disorders occurring in rheumatic fever. This association would lead to the expectation of an excess of rheumatic valvular heart disease in patients with the OCD spectrum disorders.

STROKE

Stroke is the second leading cause of death worldwide, and the fourth leading cause in the United States, accounting for 5 percent of deaths. Most strokes are due to atherosclerotic disease, with risk factors similar to those for coronary disease. Embolic events and primary intracranial hemorrhage are other causes of stroke. Study of the relationship between stroke and psychiatric illness has focused primarily on depression as an effect of stroke; the incidence of poststroke depression appears to be in the range of 20 to 50 percent at 1 month to 1 year after the index event. However, other questions of interest include the effect of psychological factors on the risk of stroke incidence and recovery from stroke, the relationship between stroke lesion location and incidence of depression, and possible prophylaxis.

Depression is a risk factor for stroke. In NHANES I, depression predicted stroke in models adjusting for age and other stroke risk factors in white men and blacks of both sexes, with relative risks ranging from 1.68 to 2.60. For white women, the effect of depression was of borderline significance. In the Women's Health Initiative Study, SSRI antidepressants increased the risk of hemorrhagic stroke in postmenopausal women.

A previous history of depression, previous history of other psychiatric illness, dysphasia, functional impairment, living alone, and social isolation are the most important predictors of depression after stroke. Lesion laterality is not a good predictor of depression risk. Worsening of

depression symptoms over time after stroke is associated with subcortical white matter lesions, and persistent symptoms with small basal ganglia lesions or large cerebral cortical lesions.

Depression after stroke appears to have a negative effect on functional outcome, independent of the severity of the stroke itself as judged by other measures. In the short term, depression impairs reacquisition of mobility and capacity for activities of daily living, even in patients receiving antidepressant treatment. This effect persists in studies looking at outcome as long as 15 months after stroke.

Given the high incidence, prevention of depression in poststroke patients has been a matter of investigation. Some studies have demonstrated prophylactic benefits from sertraline, fluoxetine (Prozac), and nortriptyline. Treatment of depression in poststroke patients with nortriptyline and serotonin reuptake inhibitors is safe and effective but has not been shown to reduce poststroke mortality. In a randomized controlled trial, fluoxetine was superior to placebo in its effect on motor recovery over 6 months after stroke. This effect was independent of an effect on depression.

CLOZAPINE MYOCARDITIS

Clozapine treatment is associated with myocarditis, with reported incidence varying from <0.05 to 1 percent. Signs and symptoms include flu-like illness, chest pain, tachycardia, hypotension, palpitations, and dyspnea. Arrhythmias and heart failure may occur. About 50 percent of cases recover completely, while 10 percent have significant persistent heart failure symptoms or other complications. Fever during the first 6 weeks of clozapine treatment may be a sign of early myocarditis. Occurrence is idiosyncratic and may involve allergic or autoimmune responses.

FUNCTIONAL CARDIAC SYMPTOMS

Somatic symptom disorder, panic disorder, anxiety, and depression can all present with somatic complaints and represent a substantial issue in ambulatory and emergency cardiology practice. In studies of patients presenting with the chief complaint of palpitations, these diagnoses account for approximately 30 percent of cases. In this population, psychiatric disorder is associated with more frequent recurrent symptoms, emergency room visits, hypochondriacal concerns, and impairment in activities of daily living. Similarly, chest pain is a common presentation for panic disorder. In a study of 1,999 consecutive patients presenting with chest pain, panic disorder was diagnosed in 38 percent. Patients without CAD were twice as likely to receive a panic disorder diagnosis. Panic patients were younger, more likely to be female, and more likely to be unemployed. In another study of individuals presenting to the emergency room with chest pain, variables associated with panic disorder included

absence of CAD, female sex, atypical quality of chest pain, younger age, and high level of self-reported anxiety.

PSYCHIATRIC PROBLEMS IN PATIENTS WITH CARDIAC DISEASE

Psychiatric disorders frequently occur as complications or as comorbid conditions in individuals with cardiovascular disease. Depression, anxiety, delirium, and cognitive disorders are especially prevalent problems. Surveys of ambulatory cardiology practice patients with documented heart disease indicate a point prevalence of 5 to 10 percent with anxiety disorders, predominantly panic attacks and phobias, and 10 to 15 percent with affective disorders, predominantly major depressive episodes, minor depression, or dysthymia. Major depressive disorder occurs in 15 to 20 percent of patients after MI. PTSD is increasingly recognized as a complication and comorbidity in patients with heart disease.

Psychiatric Complications of Arrhythmias

Ventricular arrhythmias may be asymptomatic or may cause palpitations, light-headedness, dizziness, syncope, or sudden cardiac death. Patients who experience life-threatening rhythm disturbances are prone to secondary adjustment, mood, and anxiety disorders. The cardiovascular symptoms may lead to profound disruption of social roles and capacity for autonomous functioning. Even patients without symptoms may be counseled to avoid activities such as driving, which may be hazardous in the event of an arrhythmic event. Psychodynamically, because of the recurring, unpredictable, and sudden quality of the course of illness, fear of exertion, issues of dependence on others, and loss of control are especially prominent, as well as anxiety about death itself.

Treatment of cardiac rhythm disorders can engender psychiatric complications. Many antiarrhythmic agents have psychiatric side effects, including delirium, hallucinations, paranoid ideation, and mood disturbance. Automatic implanted cardioverter-defibrillators (AICDs) have supplanted chronic antiarrhythmic medical therapy in most patients who have experienced ventricular tachycardia–ventricular fibrillation. Unfortunately, inappropriate shocks are commonplace. Cardioversion by an AICD may occur in a conscious patient, who experiences the sensation of one or more powerful shocks or blows to the chest. Although most patients with AICD seem satisfied with their devices, approximately 10 percent of patients develop clinically significant anxious or depressed mood symptoms, including panic and agoraphobia and symptoms of posttraumatic stress associated with AICD discharges. Avoidance behaviors and anxiety in the setting of discharge-associated stimuli reduce quality of life for patients who have experienced shocks during ordinary activities such as sexual relations, showering, and low-intensity exercise. Risk factors

for psychiatric complications in patients with AICD include defibrillator discharges early after implantation, repetitive discharges, and a high cumulative number of discharges, along with patient factors such as neuroticism and premorbid anxiety. PTSD in cardiac patients is discussed further below.

Psychiatric Complications of Cardiac Surgery

Valve replacement was among the first of the open-heart surgeries to be developed, and pioneering work in psychiatric aspects of cardiovascular surgery demonstrated a high prevalence of delirium in early postcardiotomy patients. This delirium could be attributed to toxic or metabolic processes in many cases; prolonged exposure to the intensive care unit (ICU) environment with sleep deprivation, sensory stimulation, and simultaneous monotony led to the phenomenon of delirium after a lucid interval (so-called ICU psychosis), but shortened ICU stays and better attention to maintenance of day–night cycles, noise reduction, and the provision of privacy have greatly reduced the extent of this phenomenon.

CABG surgery is one of the most frequently performed operations in the United States. Except for patient with disease of the left main and left anterior descending coronary arteries, CABG does not have an advantage over medical management of CAD with respect to survival, but surgery may be more effective than medication for amelioration of angina symptoms and improvement in exercise tolerance, functional capacity, and overall quality of life. Psychiatric complications after CABG have been noted since its inception and persist despite improvements in cardiopulmonary bypass, anesthetic, and surgical technique. Delirium after surgery is common and is discussed below. Persistent, subtle memory and cognitive impairment may occur after CABG. Predictors of stroke, death, and adverse neuropsychiatric outcome include older age, proximal aortic atherosclerosis, prior history of neurological disease, systolic hypertension, pulmonary disease, and excessive alcohol use. Adverse cerebral outcomes substantially increase hospitalization duration, in-hospital mortality, and the need for discharge to institutional supportive care. Although embolic events associated with cardiopulmonary bypass have long been associated with adverse cerebral outcomes, cardiac surgery performed “off-pump,” that is, without cardiopulmonary bypass, has not been found to result in significant improvement in the rate of cognitive impairment or neuropsychiatric adverse outcome.

Delirium

Delirium is a common problem in severely ill cardiac patients. Three main categories of patients are at risk: patients with severe CHF, patients receiving antiarrhythmic agents for tachyarrhythmias early after MI or

cardiac surgery, and patients who have had cardiac surgery. In patients with CHF, delirium commonly results from hypoxia, hyponatremia, liver failure, or azotemia, as pulmonary congestion and poor end-organ perfusion progress. Patients receiving lidocaine and procainamide—now rarely employed–may appear psychotic or delirious even at nominally therapeutic blood levels. In postoperative patients, the clinical picture may include all of these elements; in addition, cerebrovascular hypoperfusion during surgery, infection, sedatives, and narcotics may contribute to delirium. Management relies on correcting the underlying abnormality while treating psychosis or agitation with antipsychotic agents. Other sedatives should generally be avoided, although the agitated patient may benefit from concurrent administration of lorazepam (Ativan) with haloperidol. IV haloperidol can be administered frequently, provided that hemodynamic and ECG monitoring is in place. High doses, however, may be associated with arrhythmias, including torsade de pointes (TdP), due to prolonged cardiac conduction. (Conduction disturbance associated with psychotropics and resulting risk of ventricular tachycardia–ventricular fibrillation is discussed below.) It should be noted that no antipsychotic drug has an indication for the treatment of delirium, and that the U.S. Food and Drug Administration (FDA) has mandated a warning label about the risk of ventricular tachyarrhythmias associated with antipsychotic drug use in the management of agitation in the elderly. Nevertheless, off-label use of antipsychotic agents remains widespread in the absence of any apparently better alternative therapy for the agitated postoperative patient. Use of dexmedetomidine (in place of opiates, propofol, and benzodiazepines) for sedation and pain control, sedative interruption, and early mobilization are management options that appear to reduce incidence of delirium after cardiac surgery.

Depression

About 20 percent of patients after acute coronary syndromes, with stable coronary disease, and with CHF meet criteria for major depressive disorder; more have elevated depressive symptoms. Although some clinicians erroneously believe that depressed mood is normal in patients with medical illness and wanes spontaneously over time after an acute event, the syndrome of major depression persists in a large portion of patients who appear depressed after an acute coronary event. Persistence of depression confers elevated risk of adverse cardiac events and mortality.

Mild-to-moderate degrees of depression occur in approximately one-third of patients after coronary bypass surgery but may remit within weeks to months. Depressive symptoms are present in almost 40 percent of patients with CABG 6 months after surgery, but the presence of depression in the early postoperative period is not well correlated with depression at 6-month follow-up. Persistent or worsening depression is associated with

increased mortality at 3-year follow-up. Subtle changes in cognition are frequently associated with depression after coronary bypass surgery. The linkage of cognitive and affective disturbances after CABG may be attributable to small vessel cerebrovascular changes seen in many elderly with depression.

Vascular Depression. Middle-aged and older depressed patients with no clinical history of stroke have a high rate of cerebral white matter infarcts, leading to the "vascular depression" hypothesis that some cases of depression are caused by multiple small infarcts that do not lead to obvious motor or sensory impairments. A diagnosis of vascular depression requires evidence of vascular disease and late-onset depression (over 50 years old) or a change in the course of depression in people who had early onset depression. Unlike typical major depressive disorder, vascular depression is not epidemiologically associated with female sex or family history of affective disorder. Some evidence suggests that vascular depression is less likely to respond to antidepressant medications than typical major depression.

Treatment of Depression. Tricyclic antidepressants cause orthostatic hypotension and cardiac conduction disturbances. In combination with diuretics, vasodilators, or benzodiazepines, their effect on blood pressure is exaggerated. In toxic levels, they may precipitate ventricular arrhythmias, but, at therapeutic doses, they exhibit type 1A antiarrhythmic properties, similar to those of quinidine (Cardioquin). Studies of type 1 antiarrhythmic therapy for patients with MI and ventricular ectopy demonstrated that antiarrhythmic therapy was associated with increased mortality. Consequently, tricyclic agents are no longer recommended as first-line drugs in the treatment of depression in patients with ischemic heart disease. Cardiovascular effects of serotonergic antidepressants in patients with impaired left ventricular function, including patients with CAD and prior MI, are lowering of heart rate by 2 to 4 beats per minute and, possibly, increased left ventricular ejection fraction. Except for dose-related increased ventricular repolarization time associated with citalopram, they have no significant effect on cardiac conduction, blood pressure, or arrhythmias. Compared to side effect rates observed with nortriptyline treatment, side effects occur less frequently and result in lower rates of treatment discontinuation in depressed heart disease patients treated with serotonin reuptake inhibitors. Bupropion in high doses may be associated with increased blood pressure and ventricular ectopic activity. Trazodone (Desyrel) and nefazodone (Serzone) may be associated with occasional ventricular ectopy and orthostatic hypotension. Venlafaxine (Effexor) in higher doses is associated with sustained elevations of blood pressure (Table 27.2–2).

In patients with acute coronary syndromes (unstable angina and MI), the sertraline antidepressant heart attack randomized trial (SADHART) demonstrated that sertraline was a safe and effective treatment of major depression. Sertraline was indistinguishable from placebo with respect to hemodynamic, ECG, and inotropic effects but was associated with a higher rate of response to treatment, better quality of life, and a trend toward fewer adverse cardiac events. The antidepressant effect was most pronounced in patients with more severe depression, with prior episodes of depression before the index cardiac event, and with onset of the current episode before the index event. Several other small trials have tended to show a beneficial effect of sertraline on depression after MI. The CREATE trial showed that citalopram, like sertraline, is effective when combined with clinical management for the treatment of major depression in patients with CAD; in contrast, interpersonal psychotherapy was not more effective than clinical management without psychotherapy. In another trial enrolling patients with depression up to 12 months after MI, mirtazapine was not more effective than placebo as measured by change in the Hamilton Rating Scale for Depression, but was modestly more effective as determined by self-rating depression scales. Recently, a large Korean study demonstrated safety and efficacy for escitalopram treatment of depression in the first 3 months after an acute coronary syndrome. Several recent collaborative care model studies, employing stepped-care interventions and incorporating patient preferences in the choice of initial treatment, have demonstrated good results with respect to patient satisfaction, mood, and quality of life outcomes, in patients with depression following admission for CAD or CHF, in patients with recent acute coronary syndromes, and in post-CABG patients (Bypassing the Blues Trial). In contrast, randomized placebo controlled trials of sertraline and escitalopram for depression in heart failure patients have not shown benefit with respect to either cardiovascular or mood outcome.

Table 27.2–2.
Cardiovascular Effects of Psychotropic Drugs

Agent	Effects
Tricyclic antidepressants	Orthostatic hypotension; delayed cardiac conduction, heart block; type 1A antiarrhythmic effect; ventricular arrhythmias in overdose; tachycardia
Selective serotonin reuptake inhibitors	Bradycardia Citalopram: Dose-related QTc prolongation
Monoamine oxidase inhibitors	Hypotension; Peripheral edema
Bupropion (Wellbutrin)	Hypertension
Nefazodone (Serzone), trazodone (Desyrel)	Hypotension
Venlafaxine (Effexor)	Hypertension, especially at doses >300 mg/day

Lithium (Eskalith)	Sinus node dysfunction, including sinus arrest
First-generation antipsychotic agents	QT interval prolongation; torsades de pointes
Second-generation antipsychotic agents—except aripiprazole (Abilify)	QT interval prolongation

Aerobic exercise has also been demonstrated to significantly reduce depression symptoms in patient with depression and CAD and depression and CHF, and rehabilitation and stress management interventions that include exercise have demonstrated beneficial effects on recurrent cardiovascular morbidity and mortality in patients with CAD and CHF, whereas psychotropic medications and psychological interventions without exercise components have not.

Experienced psychosomatic medicine psychiatrists often use stimulants as (off-label) primary or adjunctive treatment of medically ill patients with depressed mood and psychomotor retardation; this may include patients who have had cardiac surgery and patients with coronary disease or heart failure. Unstable myocardial ischemia, uncontrolled hypertension, and tachyarrhythmias are obvious contraindications. Even with doses of 20 to 30 mg/day, the effects of methylphenidate on heart rate and blood pressure are minimal. Recent studies find no evidence of increased cardiovascular morbidity associated with stimulant use in large cohorts of young adults treated with stimulants for attention deficit disorder.

Anxiety

Although the prevalence of anxiety symptoms is high in patients with acute MI, established CAD, heart failure, and arrhythmias, the treatment of anxiety disorders has received little attention in randomized trials. Psychological interventions to reduce anxiety symptoms in patients with cardiovascular disease were found to be beneficial in recent meta-analyses. By and large, these studies do not specifically target persons with high anxiety symptom levels or anxiety disorders. A common anxiety problem in patients who have had acute coronary events, heart failure, cardiac surgery, or implanted defibrillator placement is fear of resumption of sexual activity; in fact, sexual activity is safe for most patients with heart disease, and the American Heart Association has published guidelines for counseling, including advice about energy requirements, situations, and positioning most conducive to comfortable and safe sexual activity in the setting of heart disease. In the absence of randomized trials for anxiety specifically in cardiovascular disorder, it is reasonable to utilize treatments of anxiety established in other clinical settings, with an appreciation of the potential adverse cardiovascular effects of psychotropic medications (discussed below). As patients may be reluctant to initiate discussion of anxiety-provoking topics such as sexuality, open-ended questions

exploratory questions by the physician, psychiatrist, or other health professional about fears and anxiety may be welcomed as "giving permission" for discussion.

Posttraumatic Stress Disorder

Acute coronary events and their treatment can themselves be traumatic events; dramatic, painful, and intensely frightening experiences such as the abrupt onset of intense physical symptoms or emergency resuscitation procedures may be the stimulus for intrusive flashback experiences. Up to 15 percent of patients are likely to develop PTSD in the year after an MI or cardiac surgery, most commonly between 3 and 6 months after the cardiac event. Affected patients may also experience significant depression symptoms and may be less likely to adhere to medications and behavioral care. PTSD is also associated with subjective stress and functional impairment, failure to return to work, and poor quality of life. Clinically significant PTSD symptoms induced by acute coronary syndromes are associated with twofold increased risk of recurrent cardiac events and mortality.

Risk factors for the development of cardiac-related PTSD include prior history of PTSD, depression, or generalized anxiety disorder. Some studies suggest that subjective perceptions of danger or potential for life-threat are not as important as perceptions of control or mastery, especially a general sense of control over life events, in determining the risk of the development of PTSD in cardiac patients. Social support during convalescence may reduce the incidence of PTSD.

Cardiac events such as MI differ from other traumatic events in important ways. The life-threatening experience comes from within rather than outside the body, so that instead of directing vigilance to external sources of danger, the cardiac patient may instead become hypervigilant about and preoccupied with somatic signals of disease, resembling patients with the somatosensory amplification experience seen in patients with panic and illness anxiety. Cardiac patients with PTSD have higher rates of somatic symptoms than those without it and are more likely to make repeated visits to emergency departments. Anxious apprehension about future events may appear similar to the uncontrollable worry associated with generalized anxiety disorder. Cardiac medicines themselves may be reminders of the traumatic event; adherence problems may represent the avoidance of cues to intrusive flashbacks, a hallmark PTSD symptom. Furthermore, PTSD symptoms may co-occur with depression, anger, anxiety, and substance use disorder problems, complicating treatment.

Treatment of PTSD in cardiac patients follows the general guidelines for the treatment of PTSD with psychotherapy, focusing on review of the traumatic events, and use of SSRIs, with appreciation for adverse effects of medication as discussed below. A recent trial found telephone-based CBT

reduced PTSD symptoms in ICD patients, but did not result in an effect on the rate of ICD discharge.

Heart Transplantation

Although quality of life improves in most respects for most patients who undergo heart transplantation, myriad difficulties can occur. Common medical problems of transplant survivors include infections, graft rejection, graft CAD, renal insufficiency, hypertension, and adverse effects of immunosuppressive drugs, including osteoporosis. Return to work is problematic, and other social role functioning may remain limited, even after physical recovery from heart transplant surgery. Family adjustments to shifting patterns of chronic illness and recovery may be chaotic and upsetting. Sexual dysfunction affects one-third or more of heart transplant recipients. Psychiatric difficulties in heart transplant patients include primary and secondary mood disorders, delirium, pain, addiction to analgesic and anxiolytic agents, and anxiety. Tacrolimus and cyclosporine may induce a variety of neuropsychiatric problems including peripheral neuropathy pain, headache, altered mental status, and seizures. Serotonergic antidepressants and nortriptyline are generally well tolerated in heart transplant recipients with depression.

Patients with Congenital Heart Disease

The psychiatric status of patients with congenital heart disease has received little systematic study, but the clinical experience of cardiologists involved in their care suggests that many experience enduring psychological sequelae. MRI studies of neonates with congenital heart disease before surgery demonstrate abnormalities of structure and function (periventricular leukomalacia, infarcts, and elevated lactate levels), and follow-up after surgery demonstrates new lesions in more than two-thirds of patients. The significance of these early brain lesions for subsequent cognitive and emotional development is uncertain. Children with congenital heart disease have been reported to demonstrate mild cognitive deficits, increased anxiety, more body-image concerns, and lower self-confidence than age-matched peers. Exclusion from peer-group activities, maternal anxiety, cyanosis, and pain has been linked to higher levels of anxiety and lower quality-of-life ratings. For late adolescent and young adult patients, sexuality, childbearing, risk of having offspring with congenital heart disease, employability, and health insurance issues are dominant concerns. Surveys of adult patients attending congenital heart disease clinics demonstrate self-report of quality of life to be worse than population norms, particularly for those with cyanotic conditions, but one study found that psychopathology was lower in adult patients with congenital heart disease than in a matched sample of patients with

orthopedic disorders.

Stress Management

Psychological intervention including various stress management techniques is widely included in cardiac rehabilitation programs, without selection for patients with significant psychopathology. Overall, these programs appear to have beneficial effects on indices of psychological well-being; findings vary with respect to the effects on recurrent cardiac events and mortality. Although a 2000 meta-analysis of 23 trials including over 3,000 subjects noted improvement in both psychological and somatic disease outcomes in association with psychosocial treatments, a 2013 review concluded that psychological treatments without exercise rehabilitation components did not improve somatic disease outcomes. Sex-segregated stress management programs for patients with CAD may improve outcomes.

Other Issues in Treating Psychiatric Disorders in Patients with Cardiovascular Disease

Psychiatric Side Effects of Cardiovascular Medications. A few drugs used in the treatment of cardiovascular disorders have psychiatric side effects that should be kept in mind when assessing differential diagnosis of psychopathology in patients with cardiovascular disease. Some of these effects are listed in Table 27.2–3. Amiodarone deserves special mention because it is widely used for patients with arrhythmia control and can cause profound hypothyroidism in just a few months. Symptoms can include cognitive dulling and depressed mood and may be misattributed to depression or underlying heart failure; correction of the hypothyroid state may be neglected. In addition, amiodarone inhibits hepatic drug metabolism via cytochromes P450 3A4, 2D6, and 2C9, potentially increasing blood levels of many psychotropic agents metabolized through these enzyme systems, and increasing the likelihood of toxicities including QT prolongation and arrhythmic events. The calcium channel blocker diltiazem is also a strong inhibitor of the 3A4 system with potential for toxic drug interactions.

Cardiac Side Effects of Psychiatric Medications. Two especially important issues in use of psychotropic medications in cardiac patients are QT interval prolongation and risk of malignant ventricular arrhythmias, and bleeding risk associated with SSRIs in patients using antiplatelet agents or warfarin.

QT INTERVAL PROLONGATION AND VENTRICULAR TACHYARRHYTHMIAS. The QT interval represents the duration of ventricular depolarization and repolarization. The QT interval, corrected for heart rate (QTc), is normally less than 450

msec. Prolongation of the QTc interval is associated with progressively increasing risk of TdP, a form of malignant ventricular tachycardia; QTc interval of 500 msec is associated with double the risk of QTc of 400 msec. Risk factors for abnormal QT interval prolongation include congenital long QT syndrome, hypokalemia, hypomagnesemia, AV nodal block, low left ventricular ejection fraction, alcohol and cocaine use, advanced age, and female sex. A history of unexplained syncope, seizure, or arrhythmia, or family history of sudden cardiac death, should provoke careful scrutiny for congenital long QT syndrome. The QTc interval is an imperfect index of risk of TdP and sudden cardiac death: some drugs that prolong the QTc interval have never been associated with cases of torsades; QTc varies according to the method used to calculate it; and there is natural diurnal variation in QTc. However, several first-generation antipsychotic drugs prolong the QTc and are clearly associated with increased risk, including, thioridazine (Mellaril), mesoridazine (Serentil), haloperidol (Haldol) (particularly in IV formulation), and pimozide (Orap). Among second-generation antipsychotics, aripiprazole uniquely does not prolong the QTc, while most other second-generation agents do, to varying degrees. Ziprasidone (Geodon) is associated with the greatest QTc prolongation of the second-generation antipsychotics, yet has been associated with TdP in only two case reports. For patients who do not have heart disease, none of the second-generation antipsychotics has a well-established risk of TdP when taken at recommended doses; only when other QT-prolonging conditions are present do these medications seem to pose increased risk. These other conditions include the risk factors already listed above such as structural heart disease, advanced aged, and electrolyte disturbances; the combination of a second-generation antipsychotic with other drugs known to prolong the QT interval (see Table 27.2–4); pharmacodynamic interactions that inhibit the metabolism of the QT prolonging antipsychotic thereby increasing its blood level; and medication overdose. Drugs that inhibit metabolism of these antipsychotic agents may exacerbate QT interval prolongation and increase the risk of arrhythmia. Cardiac patients being treated with antipsychotic medications that prolong the QT interval should be monitored with periodic ECGs.

Table 27.2–3.
Some Psychiatric Side Effects of Cardiovascular Drugs

Medication	Effects
Digoxin (Lanoxin)	Visual hallucinations
Angiotensin-converting enzyme inhibitors	Mood elevation
Thiazide diuretics	Mood elevation, erectile dysfunction
β-Adrenergic blockers	Fatigue, sexual dysfunction
Lidocaine (Dalcaine)	Delirium

Mexiletine (Mexitil)	Delirium
Amiodarone (Cordarone)	Hypothyroid state, with associated depression

Among serotonin reuptake inhibitors, citalopram stands out in that increasing dose of citalopram is associated with increasing QTc interval; the FDA requires a label warning against doses above 40 mg/day due to concern about TdP, although large population studies show that the absolute risk of arrhythmias due to high-dose citalopram is very low.

BLEEDING. SSRI antidepressants inhibit platelet storage of serotonin and may prolong bleeding in occasional patients via an effect on platelets. They do not have a systematic effect on coagulation, as measured by the prothrombin time test, even in patients treated with warfarin (Coumadin). Bleeding risk is increased in patients treated with concomitant SSRIs, antiplatelet agents, and warfarin. There is about a 1.5- to 2-fold increased risk of hospitalization for bleeding events when SSRIs are combined with antiplatelet agents or warfarin, and an estimated 36 treatment-years needed to cause one additional hospitalization. The benefit of SSRI therapy may justify this risk in selected patients.

Table 27.2–4.
Substances That Prolong the QT Interval

Alcohol	Halofantrine
Amiodarone	Ketoconazole
Azithromycin	Levofloxacin
Chloroquine	Methadone
Cisapride (withdrawn due to TdP)	Moxifloxacin
Citalopram	Pentamidine
Clarithromycin	Procainamide
Cocaine	Quinidine
Disopyramide	Quinolones
Dofetilide	Sotalol
Erythromycin	Tamoxifen
Fluconazole	Terfenadine (withdrawn due to TdP)
Furosemide	Vandetanib

Additional cardiovascular effects of psychotropic drugs are listed in Table 27.2–2.

Drug Interactions and the Cytochrome System. Pharmacodynamic effects may be additive or offsetting; for example, both SSRIs and beta-blockers slow heart rate, and the combination may result in symptomatic bradycardia. Pharmacokinetic effects can increase or decrease drug effect or toxicity; for example, lithium (Eskalith) levels may be affected by the use of diuretics affecting the proximal renal tubule (thiazides) and by ACE

inhibitors, leading to lithium toxicity.

Fluoxetine, fluvoxamine, nefazodone, and bupropion are cytochrome inhibitors that may affect metabolism of cardiac medications, causing toxicity; for example, cytochrome P450 2D6 inhibitors such as bupropion, fluoxetine, and paroxetine may interfere with beta-blocker metabolism. Sertraline and paroxetine have weaker cytochrome inhibiting effects. Conversely, barbiturates and carbamazepine are enzyme inducers and may lower the effectiveness of relevant substrates. Online resources are readily available to search for pharmacokinetic and pharmacodynamic drug interactions.

A Clinical Vignette

An 86-year-old man with recurrent atrial fibrillation, mitral insufficiency, parkinsonism, mild cognitive impairment, and a history of MI and hip fracture was treated for recurrent depression with sertraline 150 mg daily. Concomitant medications included metoprolol (Lopressor), propafenone (Rythmol), hydrochlorothiazide (Esidrix) and triamterene (Dyrenium), atorvastatin (Lipitor), warfarin, levothyroxine (Levoxyl), levodopa and carbidopa (Sinemet), lansoprazole (Prevacid), tamsulosin (Flomax), and donepezil (Aricept). A permanent pacemaker had been implanted. After several months of remission of depression symptoms, sertraline was withdrawn due to sexual side effects. When depression symptoms subsequently recurred, he was treated with bupropion sustained-release tablets, in gradually escalating doses. He developed severe light-headedness on standing, and his blood pressure fell to 95/60, with paced heart rate of 60 beats per minute. Despite decreased bupropion dose, he continued to experience hypotension and dizziness until bupropion was discontinued, leading to complete resolution of symptoms and normalization of blood pressure. The psychiatrist hypothesized that bupropion, a potent inhibitor of the cytochrome 2D6 system, had interfered with metabolism of metoprolol, increasing the hypotensive effect of the beta-blocker.

Electroconvulsive Therapy. Electroconvulsive therapy (ECT) is a highly effective treatment of severe depressive and manic episodes. During ECT, initial vagally mediated bradycardia is followed by substantial increases in heart rate and blood pressure. Patients with ischemic heart disease or ventricular arrhythmias are at increased risk of MI or ventricular tachyarrhythmias. Atropine (AtroPen) is commonly used to reduce airway secretions during ECT and may reduce the extent of bradycardia but may exacerbate tachyarrhythmias. β-Adrenergic blocking agents, conversely, may precipitate heart block. Few studies have systematically controlled concurrent treatment and cardiac diagnoses in assessing effects of ECT in cardiac patients, but available information suggests that patients with cardiac disease are more likely than comparison patients without heart disease to experience cardiovascular complications during ECT. Most complications are transient, however, and do not necessarily preclude use of this modality. ECT has been used successfully in patients with heart transplants.

REFERENCES

Afilalo J, Rasti M, Ohayon SM, Shimony A, Eisenberg MJ. Off-pump vs. on-pump coronary artery bypass surgery: an updated meta-analysis and meta-regression of randomized trials. *Eur Heart J*. 2012;33:1257–1267.

Angermann CE, Gelbrich G, Stork S, et al. Effect of escitalopram on all-cause mortality and hospitalization in patients with heart failure and depression: the MOOD-HF randomized clinical trial. *JAMA*. 2016;315:2683–2693.

*Beach SR, Celano CM, Noseworthy PA, Januzzi JL, Huffman JC. QTc prolongation, torsades de pointes, and psychotropic medications. *Psychosomatics*. 2013;54:1–13.

Brojmohun A, Lou JY, Zardkoohi O, Funk MC. Protected from torsades de pointes? What psychiatrists need to know about pacemakers and defibrillators. *Psychosomatics*. 2013;54:407–417.

Castro VM, Clements CC, Murphy SN, et al. QT interval and antidepressant use: a cross sectional study of electronic health records. *BMJ*. 2013;346:f288.

Cochran KA, Cavallari LH, Shapiro NL, Bishop JR. Bleeding incidence with concomitant use of antidepressants and warfarin. *Therap Drug Monit*. 2011;33:433–438.

Cooper WO, Habel LA, Sox CM, et al. ADHD Drugs and Serious Cardiovascular Events in Children and Young Adults. *New Engl J Med*. 2011;365:1896–1904.

Davidson KW, Bigger JT, Burg MM, et al. Centralized, stepped, patient preference–based treatment for patients with post–acute coronary syndrome depression: CODIACS vanguard randomized controlled trial. *JAMA Intern Med*. 2013;173:997–1004.

Dickens C, Cherrington A, Adeyemi I, et al. Characteristics of psychological interventions that improve depression in people with coronary heart disease: a systematic review and meta-regression. *Psychosom Med*. 2013;75:211–221.

Doyle F, McGee H, Conroy R, et al. Systematic Review and Individual Patient Data Meta-Analysis of Sex Differences in Depression and Prognosis in Persons With Myocardial Infarction: A MINDMAPS Study. *Psychosom Med*. 2015;77:419–428.

Edmondson D, Cohen BE. Posttraumatic Stress Disorder and Cardiovascular Disease. *Prog Cardiovasc Dis*. 2013;55:548–556.

Eisen A, Giugliano RP, Braunwald E. Updates on acute coronary Syndrome. A review. *JAMA Cardiol*. 2016;1:718–1730.

Gottlieb SS, Kop WJ, Thomas SA, et al. A double-blind placebo-controlled pilot study of controlled-release paroxetine on depression and quality of life in chronic heart failure. *Amer Heart J*. 2007;153:868–873.

Hasnain M, Vieweg WV. QTc interval prolongation and torsade de pointes associated with second-generation antipsychotics and antidepressants: a comprehensive review. *CNS Drugs*. 2014;28:887–920.

Huffman JC, Mastromauro CA, Beach SR, et al. Collaborative care for depression and anxiety disorders in patients with recent cardiac events: the Management of Sadness and Anxiety in Cardiology (MOSAIC) randomized clinical trial. *JAMA Intern Med*. 2014;174:927–935.

Kim JM, Bae KY, Stewart R, et al. Escitalopram treatment for major depressive disorder following acute coronary syndrome: a 24-week double-blind placebo-controlled trial. *J Clin Psychiatry*. 2015;76(1):62–68.

Labos C, Dasgupta K, Nedjar H, Turecki G, Rahme E. Risk of bleeding associated with combined use of selective serotonin reuptake inhibitors and antiplatelet therapy following acute myocardial infarction. *CMAJ*. 2011;183:1835–1843.

Lesperance F, Frasure-Smith N, Talajic M, Bourassa MG. Five-year risk of cardiac mortality in relation to initial severity and one-year changes in depression symptoms after myocardial infarction. *Circulation*. 2002;105:1049.

*Lichtman JH, Froelicher ES, Blumenthal JA, et al. Depression as a risk factor for poor prognosis among patients with acute coronary syndrome: systematic review and recommendations: a scientific statement from the American Heart Association. *Circulation*. 2014;129:1350–1369.

Low CA, Thurston RC, Matthews KA. Psychosocial factors in the development of heart disease in women: current research and future directions. *Psychosom Med*. 2010;72:842–854.

McMurray JJV, Packer M, Desai AS, et al. Angiotensin–neprilysin inhibition versus enalapril in heart failure. *N Engl J Med*. 2014;371:993–1004.

Meijer A, Conradi HJ, Bos EH, et al. Adjusted prognostic association of depression following myocardial infarction with mortality and cardiovascular events: individual patient data meta-analysis. *Br J Psychiatry*. 2013;203(2):90–102.

Mozaffarian D, Benjamin EJ, Go AS, et al. Heart disease and stroke statistics—2016 update: a report from the American Heart Association. *Circulation*. 2015.

Murray D, Weiner M, Prabhaker, M, Fiedorowicz JG. Mania and mortality: why the excess cardiovascular risk in bipolar disorder? *Curr Psychiatry Rep*. 2009;11:475–480.

Peyre H, Hoertel N, Hatteea H, Limosin F, Dubuc C, Delorme R. Adulthood self-reported cardiovascular risk and ADHD medications: results from the 2004–2005 National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry*. 2014;75:181–182.

Pizzi C, Rutjes AWS, Costa GM, Fontana F, Mezzetti A, Manzoli L. Meta-analysis of selective serotonin reuptake inhibitors in patients with depression and coronary heart disease. *Am J Cardiol*. 2011;107:972–979.

Riegel B, Lee CS, Dickson VV. Self-care in patients with chronic heart failure. *Nat Rev Cardiol*. 2011;8:644–654.

Rosengren A, Hawken S, Ounpuu S, Sliwa K, Zubaid M. Association of psychosocial risk factors with risk of acute myocardial infarction in 11,119 cases and 13,648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet*. 2004;364:953.

Rutledge T, Redwine LS, Linke SE, Mills PJ. A meta-analysis of mental health treatments and cardiac rehabilitation for improving clinical outcomes and depression among patients with coronary heart disease. *Psychosom Med*. 2013;75:335–349.

Schelleman H, Brensinger CM, Bilker WB, Hennessy S. Antidepressant-Warfarin Interaction and Associated Gastrointestinal Bleeding Risk in a Case-Control Study. *PLoS ONE*. 2011;6(6):e21447.

Schumacher SM, Gao E, Zhu W, et al. Paroxetine-mediated GRK2 inhibition reverses cardiac dysfunction and remodeling after myocardial infarction. *Science Translat Med*. 2015;7(277):277ra31.

*Shapiro PA. Depression treatment and coronary artery disease outcomes: time for reflection. *J Psychosom Res*. 2013;74:4–5.

Steinke EE, Jaarsma T, Barnason SA, et al. Sexual counseling for individuals with cardiovascular

disease and their partners: a consensus document from the American Heart Association and the ESC Council on Cardiovascular Nursing and Allied Professions (CCNAP). *Circulation*. 2013;128:2075–2096.

Steptoe A, Brydon L. Emotional triggering of cardiac events. *Neurosci Biobehav Rev*. 2009;33:63–70.

Summers MR, Lennon RJ, Prasad A. Premorbid psychiatric and cardiovascular diseases in apical ballooning syndrome (tako-tsubo/stress-induced cardiomyopathy) potential pre-disposing factors? *J Am Coll Cardiol*. 2010;55:700–701.

Tulloch H, Greenman PS, Tassé V. Post-Traumatic Stress Disorder among Cardiac Patients: Prevalence, Risk Factors, and Considerations for Assessment and Treatment. *Behav Sci (Basel)*. 2015;5:27–40.

Wallerstedt SM, Gleerup H, Sundstrom A, Stigendal L, Ny L. Risk of clinically relevant bleeding in warfarin-treated patients—influence of SSRI treatment. *Pharmacoepidemiol Drug Saf*. 2009;18:412–416.

Welton NJ, Caldwell DM, Adamopoulos E, Vedhara K. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. *Am J Epidemiol*. 2009;169(9):1158–1165.

Whalley B, Rees K, Davies P, et al. Psychological interventions for coronary heart disease. *Cochrane Database Syst Rev*. 2011;(8):CD002902.

Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation*. 2013.

*Zivin K, Pfeiffer PN, Bohnert AS, et al. Evaluation of the FDA warning against prescribing citalopram at doses exceeding 40 mg. *Am J Psychiatry*. 2013;170:642–650.

▲ 27.3 Gastrointestinal Disorders

CLAUDE ANDREW BOTHA, PH.D., JAMES K. RUFFLE, B.Sc., M.B.B.S., AND ADAM D. FARMER, PH.D.

INTRODUCTION

There is an intimate relationship between the GI tract and the psyche. For example, stress or anxiety may modulate GI function and result in disturbances of function such as diarrhea, bloating, nausea, and discomfort. Indeed, the influence that anxiety may exert on the GI tract is often reflected in phrases that are in common parlance such as "butterflies in my stomach" and "gut wrenching." However, it was not until the 19th century that such observations began to be objectively examined. William Beaumont (1785 to 1853), an early pioneer of GI physiology, isolated hydrochloric acid from gastric secretions and described the influence of emotions on gastric acid secretion. While the techniques used to interrogate the anatomy and physiology of the GI tract function have advanced, the relationship between the psyche and GI function, both in health and disease, remains prominent in modern clinical practice and research. Research from a diverse array of disciplines, coupled with technological advances particularly in functional neuroimaging, has provided an increasing body of evidence for the proposal of a gut–brain axis; a bidirectional intercommunication between the gut and the brain, which has gained widespread acceptance as the germane construct for providing an explanation of the normal, acute, and chronic alterations in digestive tract function. Moreover, this model of circuitous communication has provided a biological construct to underpin the biopsychosocial concept of GI disorders by facilitating the integration of many contributing factors whether they are biological, psychological, or social in nature.

In many ways, considering the essence of the construct of the gut–brain axis, it should come as no surprise that many GI disorders are comorbid with psychiatric disorders. The aims of this section are to provide the reader with a summary of both inflammatory bowel disease (IBD) and the so-called functional GI disorders, itself comprising of a number of distinct disorders, as many of these patients are routinely referred to the psychiatrist. There has been an appreciable interest into developing a mechanistic-based understanding of the pathophysiology of functional GI disorders over the recent past, largely as a consequence of the clinical burden that such patients exert. Within secondary care, more than one-third of new patients are diagnosed as having one or more functional GI disorders. The functional GI disorders are a heterogeneous group consisting of in excess of 25 separate disorders, defined according to the Rome multinational consensus as "variable combinations of chronic or

recurrent GI symptoms which are not explained by structural or biochemical abnormalities." Functional GI disorders are characterized by chronic abdominal pain and discomfort in the absence of any physical, biological, or anatomical abnormalities. The most prevalent functional GI disorder is irritable bowel syndrome (IBS), with a reported prevalence in excess of 10 percent in the adult population, peaking in the second, third, and fourth decades of life with a female preponderance. The Rome process, which has developed in an iterative fashion in a manner not dissimilar to the *Diagnostic and Statistical Manual of Mental Disorders,* has generated considerable debate among gastroenterologists who are used to making diagnoses on the basis of observable pathological abnormalities. Regrettably, this often translates into clinical practice that symptoms based upon demonstrable abnormalities, visualized perhaps through endoscopic or cross-sectional imaging, are "real," whereas other complaints may be dismissed as functional with a large psychiatric component. The psychiatrist must remain mindful that this message may have been transmitted, unwittingly, to patients prior to referral, making some patients reticent when asked about many areas of their psychiatric history. Nevertheless, despite protestations from the gastroenterologist, the psychiatrist does have a pivotal role in the management of patients with GI disorders, and those who are particularly refractory to standard medical therapy.

The chapter will be structured around examining both "organic" and "functional" disorders of the digestive tract. The seemingly arbitrary split of gastroenterological disorders in "organic" or "functional" based upon the presence or absence of a particular biomarker remains the dominant model in specialist practice, although such dualism has marked limitations. For instance, peptic ulcer disease was considered to be a psychosomatic disorder prior to the identification of *Helicobacter pylori.* Furthermore, while the identification of colonic inflammation may be consistent with a diagnosis of gastroenteritis or IBD, many patients continue to remain symptomatic even in the context of mucosal healing and resolution of inflammation. This has led to the proposal that IBD and IBS in reality represent a similar disease at differing ends of a spectrum where central and peripheral factors are involved in symptom generation. This chapter will be structured around describing the epidemiology, putative pathogenesis, and broad management of both "organic" and "functional" GI disorders. Each disorder will be further illustrated with a clinical case to illustrate important practice points as well as a focus on the pivotal role of the psychiatrist in successfully managing such patients.

INFLAMMATORY BOWEL DISEASE

The two major types of IBD encountered in clinical practice are Crohn disease (CD) and ulcerative colitis (UC), with the former potentially

affecting any portion of the GI tract whereas the latter is confined solely to the colon. The prevalence of CD is approximately 25 to 100 per 100,000, in contrast to UC that is slightly more prevalent at 80 to 150 per 100,000. While both disorders have a global representation, the highest prevalence and incidence rates are reported in Europe and North America. Although the etiopathogenesis of IBD is incompletely understood, evidence suggests that this disorder is caused by a complex interaction between an exaggerated immunological response to enteric microorganisms in a genetically susceptible host. The symptoms of IBD largely revolve around changes in bowel habit, abdominal pain, and rectal bleeding, and as such sufferers can feel stigmatized leading to severe impairment in health-related quality of life. Moreover, medical interventions such as systemic corticosteroids or immunosuppressive drugs may result in secondary health problems compounding the severity.

The relationship between psychiatric disorders and IBD is incompletely understood. Evidence suggests that anxiety and depression are more common in patients with IBD and that the symptoms of these affective disorders are more prominent during periods of disease relapse. Nevertheless, there is a paucity of studies that objectively investigate the hypothesis that anxiety and depression are risk factors for IBD. However, comorbid anxiety and depression seem to worsen the natural history of the IBD once it is developed.

IBD is a relapsing and remitting disorder. Psychological stress has long been reported anecdotally to increase disease activity in IBD, with a number of studies confirming that depression, adverse life events, and chronic stress increase the probability of relapse in patients with quiescent IBD. Within the evolving area of psychoneuroimmunology, the biological mechanisms by which the nervous system can influence immune function, both at the level of the GI mucosa and systemically, are increasingly being appreciated, particularly with animal models. However, such mechanisms are likely to be numerous and interactional with one another. For example, the contents of mast cells, various cytokines and chemokine, may activate primary afferent neurons thereby sensitizing them and facilitating increased afferent input into central pain centers.

Psychological Therapies in IBD

Depression and anxiety may affect symptoms and functioning through the development of heightened pain sensitivity of the intra-abdominal organs (termed visceral hypersensitivity), exacerbation of bowel symptoms, and reduced self-motivation to overcome the difficulties associated with IBD. Depression may adversely impact health-related quality of life because of depressive coping (social withdrawal and irritable behavior toward others as well as brooding, self-pitying, and pessimistic thinking) and hopelessness. Symptoms and functioning can potentially improve if the

depression and anxiety are treated successfully.

A recent study reported by Szigethy et al. compared the efficacy of CBT to supportive nondirective therapy (SNDT) in treating adolescents with comorbid IBD and depression. While both psychotherapies resulted in significant reductions in depression scales, SNDT was associated with a greater reduction in IBD activity in the absence of health-related quality of life improvements. Hypnotherapy may also have a role in treating IBD. Miller and Whorwell administered gut-focused hypnotherapy to 15 patients with severe IBD and demonstrated improvements in quality of life and a decreased reliance on corticosteroids (more than half of the patients stopped corticosteroid therapy completely). They concluded that hypnotherapy might represent a promising adjunctive treatment of IBD with steroid-sparing effects, although there is a need for well-designed controlled trials in the area. A recent meta-analysis identified seven randomized controlled trials for mind/body interventions such as lifestyle modification, hypnotherapy, relaxation training, and mindfulness, which may display a degree of efficacy in treatment, although the results were heterogeneous and at significant risk of bias.

FUNCTIONAL GI DISORDERS

Visceral pain is common and remains a leading worldwide cause of health-care usage. Moreover, unexplained abdominal pain is the sixth most common cause of acute presentation to secondary care service in women and the tenth most common cause in men. In the United Kingdom, it has been estimated that nonspecific abdominal pain exerts an annual cost in excess of £100 million. The functional GI disorders are a group of heterogeneous disorders that represent one of the great unmet needs in gastroenterology and are characterized by recurrent unexplained abdominal pain. Functional GI disorders exert a significant socioeconomic impact. Most marked among these disorders is absenteeism, which has been estimated to be in the order of 13 days per annum, compared to 5 days in healthy age and gender-matched controls. Furthermore, it is estimated that between 0.1 and 0.5 percent of total health-care expenditure is spent on functional GI disorders in health-care economies of industrialized countries. Although all functional GI disorders are associated with a reduction in health-related quality of life, they are not associated with any mortality per se. For the sake of brevity, this chapter will discuss three particularly prevalent functional GI disorders, namely IBS, functional abdominal pain (FAP), and functional chest pain (FCP) of presumed esophageal origin.

IRRITABLE BOWEL SYNDROME

The most frequently encountered functional GI disorder is IBS with a reported global prevalence of >10 percent and has a marked socioeconomic

burden. IBS is also associated with a significant diminution in work productivity and health-related quality of life. The main clinical features of IBS are recurrent abdominal pain and/or discomfort, associated with an alteration in stool frequency or consistency and relief with defecation (Table 27.3–1). Extra-GI features are also prevalent and include lassitude, headache, dysmenorrhea, and dyspareunia. While symptoms characteristically wax and wane, patients typically report pain or discomfort on approximately 3 days per week, which often occurs in clusters. IBS can be usefully subdivided based upon predominant bowel habit and is grouped as IBS with constipation, IBS with diarrhea, IBS with alternating bowel habit, or IBS unclassified (see Fig. 27.3–1).

Table 27.3–1.
The Rome III Definition of Irritable Bowel Syndrome

Rome III Criteria for the Diagnosis of Irritable Bowel Syndrome
Recurrent abdominal pain or discomfort at least 3 days a month in the past 3 months, associated with two or more of the following: ► Improvement with defecation ► Onset associated with a change in frequency of stool ► Onset associated with a change in form (appearance) of stool
*Criteria fulfilled for the past 3 months with symptom onset at least 6 months before diagnosis. ''Discomfort'' means an uncomfortable sensation not described as pain.

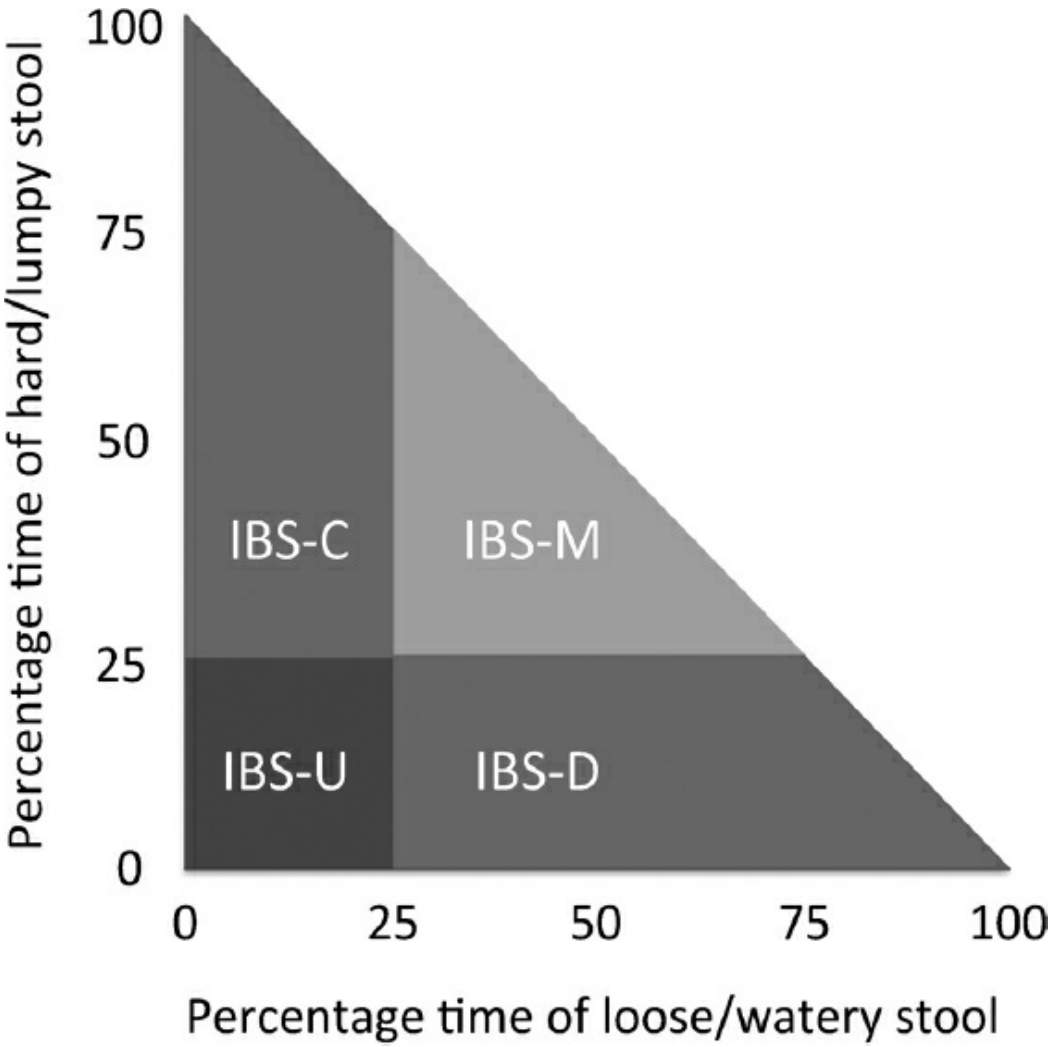

FIGURE 27.3–1. The different subtypes of IBS are classified according to the predominant bowel habit: IBS with constipation (IBS-C), IBS with diarrhea (IBS-D), IBS with mixed bowel habit (IBS-M), and IBS-unclassified (IBS-U).

Psychological comorbidity such as depression, anxiety, and hypochondriasis is common and it has been estimated that over half of patients with IBS suffer from these disorders. In animal models, studies have shown that adverse early life events, such as maternal separation, represent independent risk factors for the development of IBS and visceral hypersensitivity in adulthood. In humans, there is evidence that a history of sexual abuse, especially in childhood, can alter visceral pain sensitivity. Furthermore, the psychological context in which GI symptoms are interpreted by an individual may predict the development of IBS following an episode of gastroenteritis. Given the marked heterogeneity of the disorder, the efficacy of pharmacotherapeutic interventions is, at best, limited. However, meta-analytic evidence has demonstrated that psychological treatments, as a class of interventions per se, are effective in symptom reduction in IBS. Behavioral therapies, including CBT, dynamic psychotherapy, and hypnotherapy all alleviate the symptoms associated

with IBS. The role of CBT was specifically addressed by Lackner et al. where patients with moderate-to-severe IBS were randomized to receive psychotherapist delivered CBT, self-administered CBT, or usual care. Those who received therapist-administered CBT and self-administered CBT had superior reporting rates of relief from abdominal pain in comparison to those receiving usual care (60.9, 72 vs. 7.4 percent, respectively). Gut-focused hypnotherapy, where the aim is to reassure the patient regarding their symptoms, with the aid of hypnosis, improves GI symptoms through cognitive changes in function, thereby influencing sensitivity or spasm. The use of hypnotherapy has been evaluated in a number of studies. For example, Whorwell et al. demonstrated significant improvement in abdominal pain in a cohort of patients with refractory IBS, and in a recent Cochrane review hypnotherapy was concluded to be more efficacious than the current standard of care. Given the relatively high burden of side effects associated with more traditional pharmacological therapies, hypnotherapy represents an excellent treatment option. However, limited therapist availability often restricts accessibility for patients.

Table 27.3–2.
The Rome III Diagnostic Criteria for the Diagnosis of Functional Abdominal Pain

Rome III Diagnostic Criteria for Functional Abdominal Pain
All the following criteria must be fulfilled, with symptom onset at least 6 months prior to diagnosis:-
► Continuous or almost continuous abdominal pain
► No relation to physiological events
► Some loss of daily functioning
► The pain is real and not feigned
► Insufficient symptoms to meet the criteria for another functional GI disorder

FUNCTIONAL ABDOMINAL PAIN

FAP syndrome is characterized by frequent or continuous abdominal pain associated with a degree of loss of daily activity in the absence of a change in bowel habit (see Table 27.3–2). It is likely that FAP represents a relatively uncommon entity with reported prevalence ranging from 0.5 to 2 percent with a female preponderance.

There is an absence of international consensus with respect to the optimal clinical management of FAP, largely due to the paucity of well-designed randomized-controlled trials for therapeutic interventions. Within contemporaneous practice therefore, interventions are largely based from evidence and anecdotal experience derived from other chronic pain syndromes.

FUNCTIONAL CHEST PAIN OF PRESUMED ESOPHAGEAL ORIGIN

FCP of presumed esophageal origin is characterized by recurrent

unexplained chest pain felt in the midline. The Rome III diagnostic criteria include at least 3 months of symptoms, with onset at least 6 months prior to diagnosis, in the absence of another cause such as gastroesophageal reflux disease or esophageal dysmotility. Notwithstanding significant reductions in quality of life, patients often have disproportionately high health-care utilization and often manifest in recurrent negative investigations across medical specialties, such as cardiology, and are frequently recalcitrant to standard pharmacotherapies.

The pathophysiological mechanisms that underpin FCP are incompletely understood; however, three broad mechanisms have been suggested. First, heightened sensitivity to esophageal balloon distension, first described in 1986 by Richter et al. and subsequently confirmed by others, is considered to be a pathophysiological feature although without clinically applicable specificity and sensitivity. Second, the stress responsive physiological systems, encompassing the autonomic nervous system (ANS) and the HPA axis, are implicated in a number of functional disorders including FCP. These stress responsive systems act as brain-body interfaces and are critical components of the physiological adaption in response to stress. Finally, up to 61 percent of patients with FCP have a demonstrable degree of psychological comorbidity, which itself may enhance esophageal perception.

THE ROLE OF THE PSYCHIATRIST IN MANAGING FUNCTIONAL GI DISORDERS

Psychiatrists have traditionally not done well with the clinical management of patients with functional GI disorders. While this observation is almost certainly multifactorial, gastroenterologists frequently provide feedback of unsatisfactory clinical outcomes following referrals to local psychiatric services. It is for this reason important to highlight the role of the psychiatrist in the comanagement of functional GI disorders using psychiatric interventions alongside those that are medical. This section provides background information on the functional GI disorders for psychiatrists, discusses recent research relating to the biopsychosocial mechanisms that contribute to the illness experience, provides practical aspects of diagnosis, and describes a rationale for the treatment of patients with functional GI disorders for psychiatrists. Most of the treatment options for these conditions involve the use of psychotropic medications and psychotherapeutic methods familiar to psychiatrists. There is hence considerable benefit for psychiatrists in engaging more actively in the care of this often misunderstood and undertreated patient group.

THE SOMATIC PRESENTATIONS OF DEPRESSION IN FUNCTIONAL GI DISORDERS AND THE FAILURE TO DIAGNOSE IT BECAUSE OF "PRESENTATION BIAS"

When a patient primarily presents with physical symptoms and the apparent absence of psychiatric symptoms, any subsequent failure to

identify and diagnose psychiatric symptoms may be attributed to "presentation bias." A consultation that is empathic, supportive, and appropriately scheduled often pays considerable dividends for both patient and clinician. More often than not, the maxims of "going slow, is going fast" or "don't do something, stand there" are helpful strategies to help the patient feel trusting enough to disclose their psychosocial associations. In particular, the validation of physical symptoms and a vigilance not to be blinded by presentation bias are of critical importance in achieving mutually satisfactory outcomes for both patient and clinician. In an international study of 1,146 patients with major depression, only 69 percent reported physical symptoms. A similar study showed that in patients presenting with solely somatic complaints, only 22 percent were correctly diagnosed with anxiety or depression. This is in contrast to patients who presented with psychosocial problems, where 77 percent were correctly diagnosed with psychiatric conditions. Kroenke et al. observed that with an increase in the number of physical symptoms, there is a marked rise in affective problems, although this goes frequently unnoticed (see Table 27.3–3).

Largely as a consequence of furthered understanding of functional GI disorders and their pathophysiological determinants, there has been a significant improvement in available treatment options. Previously, abnormal motility within the GI tract was highly influential in treatment choices. However, dysmotility fails to fully explain the full symptomology and pathogenesis of functional GI disorders. It is now widely accepted that dysregulation at any level of the gut–brain axis can lead to chronic symptoms. It is within the construct of dysregulation within the gut–brain axis that one can better appreciate the symptom complex as a whole. Novel neurotransmitter and functional brain imaging research implicate dysregulation of the gut–brain axis at the peripheral, spinal, and cerebral levels, all of which may contribute toward the symptoms of functional GI disorder. Neurotransmitters such as serotonin (5-HT), norepinephrine (NE), corticotropin-releasing factor (CRF), and opioids modify both motility and sensation in the gut. It is for this reason that centrally acting treatments such as psychotropic and psychotherapeutic agents have become more commonly used. Due to their effect on the 5-HT and NE pathways, these treatments directly modulate all levels of the gut–brain axis, in addition to their traditional use of managing any psychological and psychiatric disturbances that are commonly associated with these disorders.

Table 27.3–3.
Physical Symptoms Are a Potential Surrogate Marker of Psychological Distress. The Higher the Number of Physical Symptoms, the Greater the

Degree of Psychological Distress

Symptom Counts	Number of Patients	Mood or Anxiety (%)
0–1	106	4
2–3	131	18
4–5	129	31
6–8	96	52
9+	37	78

Adapted from Kroenke K, Jackson JL, Chamberlin J. Depressive and anxiety disorders in patients presenting with physical complaints: clinical predictors and outcome. *Am J Med*. 1997;103:339–347.

The use of pharmacotherapeutic agents for functional GI disorders has grown considerably over the last two decades. Recent evidence suggests that up to one in eight patients with a functional GI disorder are offered an antidepressant. It is relatively common that when a patient presents to tertiary care, secondary care gastroenterologists have already commenced an antidepressant. A recent pharmacy database study in the United Kingdom has shown that patients prescribed ongoing therapy for presumed IBS (antispasmodic medicines for a continuous 3-month period or more) are two to four times more likely to be prescribed a drug acting on the CNS than age- and gender-matched controls. These centrally acting agents included antidepressants (35.4 percent), anxiolytics (27.8 percent), antipsychotics (9.8 percent), and hypnosedatives (32.7 percent). In a Swedish study, after antiacids, antidepressants were the most commonly used drug category reported by patient with IBS. However, it remains a challenging strategy as of insufficient understanding regarding the complex nature of these disorders. This is further compounded by a paucity of well-designed drug studies, and lastly variability among the treatment efficacy end points and indeed outcomes at the individual patient level.

Psychosocial Factors and the Biopsychosocial Model in Functional GI Disorders

One of the most clinically useful ways to develop a broadened understanding of the pathophysiology underpinning the functional GI disorders is via the biopsychosocial model, first conceptualized by the late George Engel in 1977. In this model, the influence of the CNS (both at the spinal and cerebral level), ANS and HPA axis may culminate in sensory and motor dysfunction of the GI tract. Furthermore, via the autonomic and peripheral nervous system, the GI tract can communicate with the CNS and thus provide a means of bidirectional communication, termed the gut–brain axis (see Fig. 27.3–2). Threats to homeostasis, whether it be peripheral (e.g., infection and abdominal surgery), central (e.g., history of abuse), or indeed psychosocial (e.g., alexithymia, catastrophization, and ongoing work stress and life events), frequently hold an important role in the perpetuation and clinical manifestation of these disorders through centrally media ted pathways. There is proportionality between the

importance of psychosocial factors and the severity of the functional disorder, to the degree that these factors are of major significance at the severe end of the functional spectrum. The inability to incorporate and successfully deal with these factors at present leads to more gastroenterology referrals and needless special investigations at great cost, both financially and in quality of life.

Considering the natural history of functional GI disorders, psychosocial factors have been demonstrated to significantly influence symptom generation and illness behavior during all stages: predisposing, precipitating, and perpetuating. It has been shown that up to 75 percent of patients seeking care at a tertiary referral center also fulfilled the psychiatric diagnostic criteria for anxiety or depression. This represents a marked increase in prevalence compared to patients seen in primary care or even in other general gastroenterology practice. Patients with functional GI disorders commonly give personal histories of major stressful life events, including sexual abuse, personal losses, separation, or early histories of deprivation. Patients, particularly those at the severe end of the spectrum, may also perpetuate their symptoms by means of maladaptive illness behavior such as catastrophizing that often "drives" and maintains hypersensitivity within the viscera (see Fig. 27.3–3).

Independent of pre-existing psychiatric diagnoses, upsetting feelings and distress with regard to chronic GI conditions has a negative effect on a patient's mental health and physical state. In the biopsychosocial conceptual model, bidirectional influences constantly occur. For instance, psychiatric disorders and stress can enable functional GI disorder symptoms. Likewise, chronic functional GI disorder symptoms can lead to physiological effects. It is known that stress aggravates motility, lowers pain thresholds, and increases gut inflammation. Gut distension in rats activates the locus coeruleus and leads to anxiety behavior; in humans it can activate the anterior midcingulate leading to greater affective pain experience. In this manner, the process both creates and perpetuates symptoms and worsens overall cognition, leading to further examinations and referrals. Repeated physical examinations, poor patient understanding, incomplete and often minimal or over technical explanations from physicians is in many cases a further cause of anxiety and fear that perpetuates the severity of symptoms. Chronic stress of this kind, by modulating the ANS and activation of the HPA axis, can have a sustained effect on inflammation and the gut–brain axis as a whole and explains conditions such as postinfective (PI)-IBS. Psychosocial distress at the time of infection has been shown to be an independent predictor of later PI-IBS development.

It has been suggested that 15 to 45 percent of patients with functional GI disorders additionally fulfill the diagnostic criteria for somatization disorder. In these patients, typically those afflicted with severe and

refractory symptoms, non-GI symptoms and comorbid conditions such as fibromyalgia, chronic fatigue, and chronic generalized pain are also reported. These patients are typically the "nonresponders" in a number of pharmacological intervention research studies, and it is possible that these individual set a lower threshold for symptom reporting. It is postulated that these patients have central dysregulation of their pain regulatory pathways (central sensitization). As many of these pathways are activated by neurotransmitters in common with psychiatric conditions (such as 5-HT, NE and opiates), this understanding has important implications for using centrally acting treatments. This can be done as monotherapy or in combinations with other pain medications and/or psychotherapeutic treatments (e.g., CBT and hypnotherapy). It is interesting to note that this could be effective even in the apparent "absence of psychiatric symptoms," as the pathways are now the target of treatment and not the psychiatric syndrome.

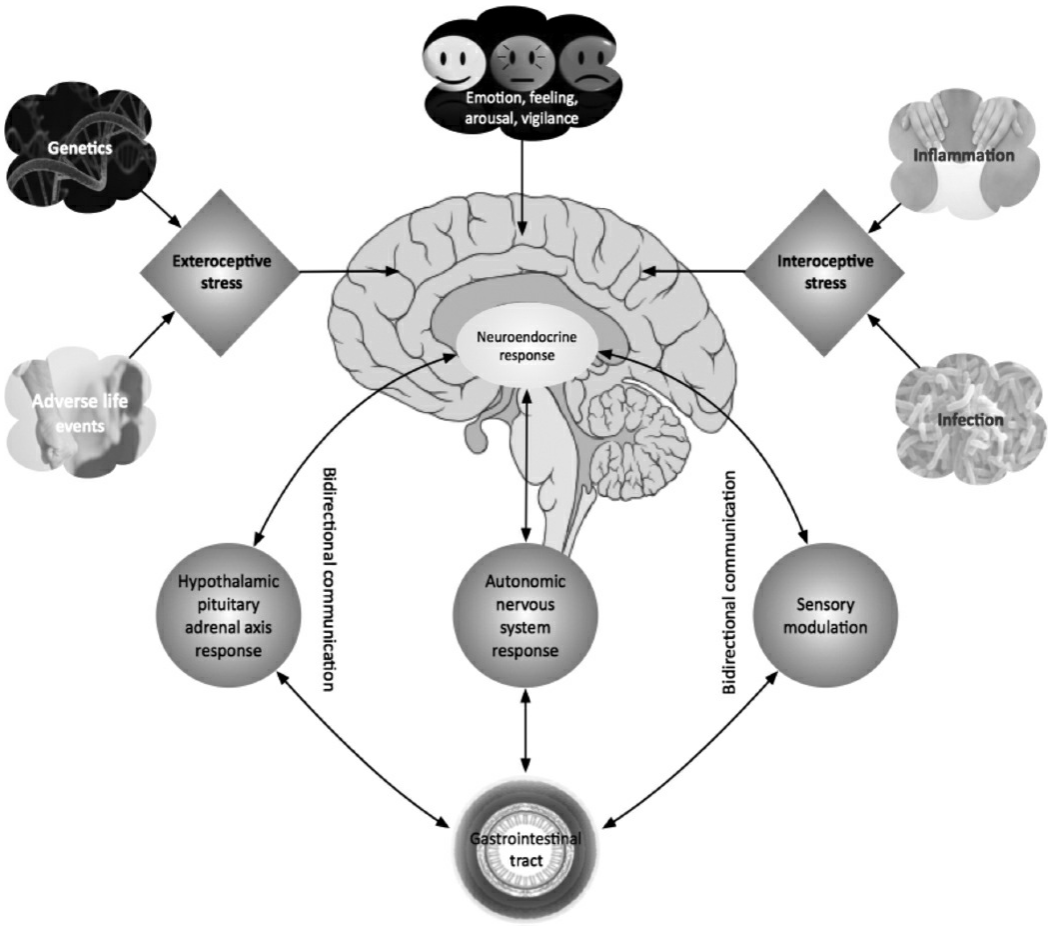

FIGURE 27.3–2. A highly schematic representation of the gut–brain axis. Exteroceptive or interoceptive stressors may activate a complex neuroendocrine response, which may influence the stress responsive systems thereby influence GI tract motility and sensory sensitivity. This response feeds back into the central nervous system giving to the concept of bidirectional communication.

GENERAL PSYCHIATRIC APPROACH AND MANAGEMENT FOR FUNCTIONAL GI

DISORDERS

Therapeutic Relationship

This can be developed when the psychiatrist (a) elicits and acknowledges the patient's beliefs, concerns, and expectations; (b) offers empathy when needed; (c) clarifies misunderstandings; (d) provides education; and (e) negotiates a plan of treatment with the patient. Some patients are initially unwilling to accept the role of psychosocial factors in the illness and the need for psychiatric intervention; this difficulty is commonly seen in patients who have suffered severe developmental trauma such as sexual abuse.

Associating Bowel Symptoms with Psychosocial Factors

In order to adequately assess the link between bowel symptoms and psychosocial factors, the authors advocate that the patient is asked to keep a daily record of their symptoms along with the frequency and consistency of bowel movements and the timing of menstruation. Such a diary can be usefully complemented with records concerning dietary, lifestyle changes, and psychosocial stressors. This information can provide the opportunity for psychoeducation and laying the basis for cognitive behavioral strategies.

Reassurance

Reassurance can only be achieved after the referring physician fully understands the patient's particular concerns, and it is here where the psychiatrist can act as a mediator, especially with emotionally complicated or aroused patients. The overall favorable prognosis for the functional GI disorder can then be stressed.

Accept the Adaptations of Chronic Illness

When the symptoms are chronic, the patient should be educated as to the nature of the illness as a troublesome cluster of symptoms, rather than as an indication of an undiagnosed underlying pathology. In some patients, chronic illness may provide attention from others, release from usual responsibilities, and social and financial compensation. In these situations, appreciable clinical improvement may take a long time, but may be advanced if the physician places more focus on improving the patient's function in the presence of illness, rather than attempting to "cure." The psychiatrist must help to minimize diagnostic studies and symptomatic treatments and work toward reinforcing health-promoting behavior in the patient.

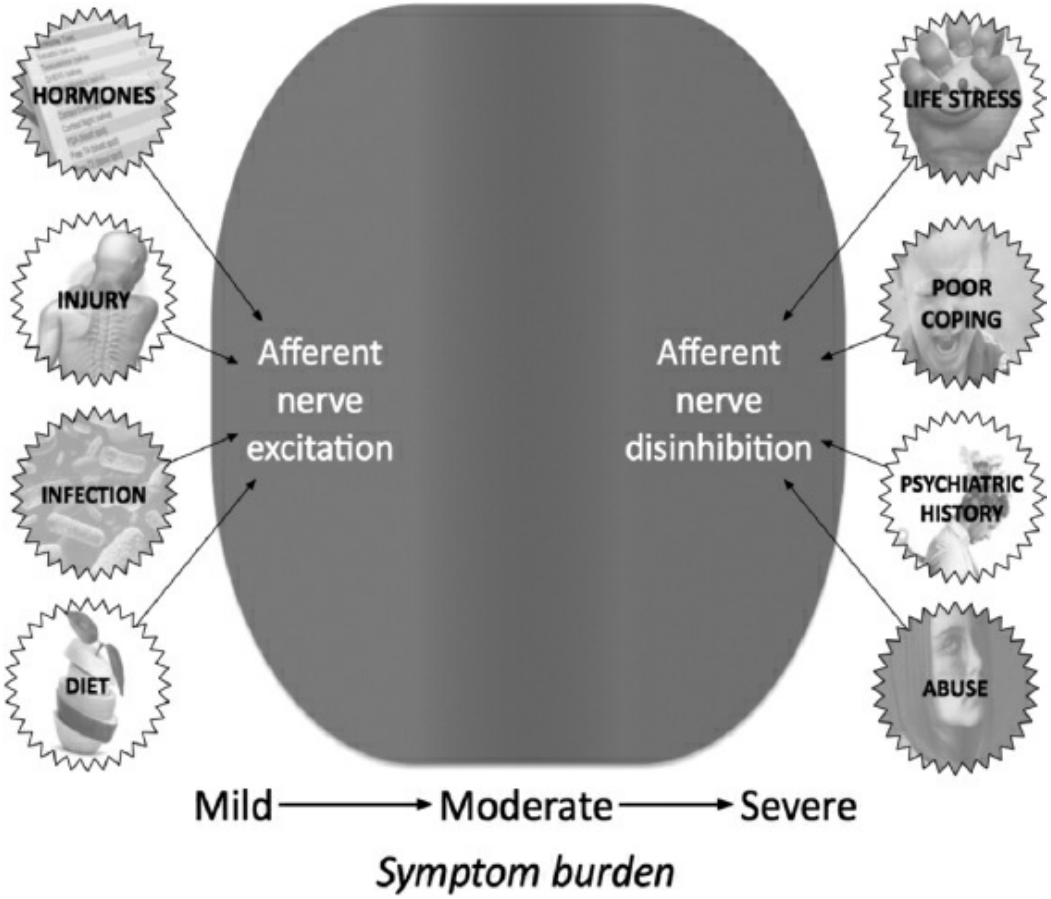

FIGURE 27.3–3. A number of physiological factors such as hormones, injury, infection, and diet can alter and enhance visceral afferent nerve signaling. Chronic life stress, poor coping, a psychiatric history, or abuse can cause long-term dysregulation or disinhibition of visceral afferent nerve leading to hyperexcitability. (Adapted from Drossman DA, Chang L, Bellamy N, et al. Severity in irritable bowel syndrome: a Rome Foundation Working Team report. *Am J Gastroenterol*. 2011;106:1749–1759; quiz 60.)

SPECIFIC PSYCHIATRIC APPROACH AND MANAGEMENT FOR FUNCTIONAL GI DISORDERS

The specific psychiatric approach consists primarily of the diagnosis and treatment of psychiatric comorbidity and includes (1) psychiatric disorders (e.g., major depression and panic disorder) that require specific treatments (e.g., antidepressants, CBT, or other psychotherapy); (2) a history of abuse that comes to light during consultation and may be interfering with adjustment to the current illness; (3) serious impairment in daily function that requires specific treatment to improve coping skills; and (4) somatization, where multiple symptoms lead to numerous consultations across specialties. In order for the gastroenterologist to refer a patient to a mental health professional, he or she must acknowledge the relevance of the psychosocial aspects to the patient's presenting problem. Otherwise, the patient may resist the referral because of perceived stigma or a rejection ("the workup is negative, it must be in your head") by the gastroenterologist. Continuing care by the gastroenterologist may also be

necessary.

PSYCHOPHARMACOLOGICAL PRINCIPLES AND MANAGEMENT FOR FUNCTIONAL GI DISORDERS

It is particularly with the use and clinical management of psychotropic medication that the role of the psychiatrist becomes crucial as part of the multidisciplinary team in the treatment of functional GI disorders. The rationale for using psychotropic agents lies in the high comorbidity: roughly half of patients with functional GI disorders also suffer from depression and/or anxiety disorders, which may respond to psychopharmacological intervention. In addition, data continues to support both the efficacy of antidepressants in the relief of chronic pain and functional GI disorders. At least five studies have shown efficacy in functional GI disorders that are independent of a change in mood.

Antidepressants

The most commonly used psychotropic medications in functional GI disorders are the antidepressants. Specifically, tricyclic antidepressants (TCAs) are commonly used as of their dual central and peripheral pain controlling effects, as well as being well-tolerated and, for some, effective at low doses. Numerous reviews and meta-analyses indicate the potential benefits of antidepressants in functional GI disorders, producing both reduction in pain and global improvements in quality of life. They also indicate that the treatment with antidepressants is associated with a significantly decreased relative risk of persistent IBS symptoms [36]. An odds ratio of between 2.6 and 4.2 has been estimated for overall improvement in IBS, translating into a number needed to treat for antidepressants of 3.2.

In an open-label physician-rated study, up to 80 percent of patients with IBS showed improvement in antidepressants, and adherence to treatment was higher than for any other treatments. However, the therapeutic effects of antidepressants on visceral hypersensitivity remain unclear as study results to date have been mixed, although it appears that the peripheral effects on the gut may be of secondary importance as most of these patients responding to antidepressants have failed treatment with conventional gut-acting agents. TCAs and SSRIs may improve global measures of well-being without having a direct correlation with pain ratings. Especially with the use of SSRIs, it seems that improvements in global measures of well-being have the most benefit to patients, as opposed to the direct effect on pain ratings. Because morbidity associated with functional GI disorders is linked to the global distress of the patient, as seen with the impairment in social interaction, work absenteeism, and functional limitations, the improvements brought about by SSRIs are still

deemed therapeutic and their use in this context is seen as good practice. Furthermore, the classes of antidepressants have added differential benefits in the treatment of functional GI disorders. A generally accepted guideline is that the TCAs and serotonin–norepinephrine reuptake inhibitors (SNRIs) are used for pain regulation, particularly where diarrhea is the predominant bowel habit, while the SSRIs should be used for somatization associated with anxiety and phobias where constipation predominates. Furthermore, mirtazapine is often helpful in the context of nausea and for boosting appetite.

Antipsychotics and Mood Stabilizers

Looking at other classes of psychotropic medications, antipsychotics like risperidone, olanzapine, and quetiapine have been found to be useful at low doses for visceral pain conditions associated with high arousal states and agitation, for instance in conditions such as globus, aerophagia, regurgitation, and postprandial distress syndrome. Mood stabilizers such as carbamazepine, sodium valproate, and pregabalin have been found to be helpful in chronic visceral pain disorders. On the whole, the general approach is to first use monotherapies at doses that can be tolerated by the patient. Should this be unsuccessful to remediate symptoms, combinations of the same or different classes could be used. It has been reported that a combination at lower doses could have a better side effect profile then for instance monotherapy at a higher dose.

Anxiolytics

Anxiolytic agents are effective for reducing anxiety in the short term, but their CNS depressant effect, including mild transient cognitive dysfunction and the risk of addiction with long-term benzodiazepines, limits their use in functional GI disorders as they are mostly chronic in nature. Prescribing benzodiazepines beyond initial short-term stabilization should therefore only be considered after persistent refractory responses to other pharmacological options (e.g., psychotropic or mood stabilizers). Alternative strategies for the treatment of anxiety should be adopted.

Psychotropic Drugs and Neuroregeneration

Probably the most striking rationale for the use of centrally acting treatments in recent years has been in its connection with the concept of neuroplasticity. This regards the observed loss of cortical neurons following psychiatric trauma, but subsequent neurogenesis (i.e., the regrowth of neurons) following treatment with antidepressants and possibly psychotherapy. Functional neuroimaging studies have shown reduced neuronal density in cortical brain regions involved in emotion and pain regulation in patients with pain disorders and IBS, just as with major

depression and PTSD. Therefore, pain and psychological trauma (and especially their combination) can be neurodegenerative, much like Alzheimer or Parkinson disease. It is in these psychological and pain conditions that the use of antidepressants and other centrally targeted agents and methods might offer some remedy by stimulating an increase in the levels of brain-derived neurotrophic factor (BDNF) following treatment. BDNF stimulates neurogenesis, and this is observed with prolonged periods of treatment that correlate with the degree of recovery from depression. Lastly, it is observed that the duration of antidepressant treatment correlates with a decreased frequency of relapse and recurrence in depression. These findings provide insight into neuronal growth regulation in key areas of the central pain matrix and provide new and important opportunities for research and patient care using antidepressants for the treatment of functional GI disorders.

PSYCHOLOGICAL THERAPIES FOR FUNCTIONAL GI DISORDER

Concerning the use of psychological therapies, experience and empirical research suggests that no single particular treatment is superior for treating functional GI disorders. The most important aspect of treatment is the patient's acceptance of the need for treatment and his or her motivation to engage in it. This can be enhanced if the gastroenterologist and psychiatrist (psychologist/therapist) help the patient accept the treatment as a necessary part of an overall plan of care. These range in their depth and ambition, from a limited attempt to control symptoms with simple behavior therapy, to hypnosis or insight-oriented psychotherapy. It should be recognized that the availability of the more time consuming techniques is limited in the public health sector and their use should therefore be restricted to only the most difficult cases.

Relaxation Therapy

Relaxation therapy is the simplest form of "psychotherapy" that can easily be taught to patients. This is usually accessed on a self-help basis through a number of "apps," CD's/audiotapes, or "classes" at local community centers or gyms. It is thought that if stress causes functional GI disorders, reducing autonomic arousal by relaxation may reduce symptoms and induce a sense of well-being, allowing the patient to feel more confident and greatening their autonomy. Patients are taught to exclude sources of tension and to increase the degree of relaxation. Arousal reduction training includes a variety of different methods to teach patients to counteract the physiological sequelae of stress or anxiety. The most widely used arousal reduction techniques include (1) progressive muscle relaxation training; (2) biofeedback for striated muscle tension; (3) autogenic training; (4) mindfulness-based stress reduction (MBSR); and (5) Tai-Chi,

transcendental, or Yogic meditation. One study showed that relaxation training reduced symptoms and, most importantly, the number of medical consultations during a 40-month follow-up, and that 10 sessions of progressive muscle relaxation over 8 weeks indicated a reduction in symptoms compared with a placebo control of symptom monitoring visits alone. Unfortunately, to date there are few well-designed controlled studies concerning short- and long-term efficacy.

Cognitive Behavioral Therapy

CBT is based on the assumption that functional GI disorders in some patients are a behavioral disease generated by responses to life events. In addition, patients who may benefit from this type of treatment may have (a) a high prevalence of anxiety, depression, and assertion difficulties; (b) a high need for social approval; and (c) perfectionistic attitudes; in patient with functional GI disorder, which are all amenable to CBT. CBT consists of a wide range of strategies and procedures designed to bring about alterations in patients' perceptions of their situation. It facilitates their ability to control GI symptoms by learning new ways of thinking and behaving through personal experience and practice. Treatment essentially involves helping the patient to recognize maladaptive patterns of thinking and behavior, which produces a greater sense of control and autonomy in the patient. It insists that the patient takes some responsibility for the illness and helps him/her find a more healthy way of dealing with the underlying problem.

Interpersonal Therapies

This form of therapy focuses on resolving difficulties in interpersonal relationships that underlie or exacerbate abdominal symptoms in functional GI disorders. It consists of the identification and mental processing of key problem areas that include unresolved grief or loss, role transitions, and relationship discord. Initial focus is on the patient's abdominal symptoms, which are explored in significant detail. Emotional distress and abnormal feeling states arising from or linked to physical symptoms are then identified. Key problem areas in relationships and their link to physical and psychological symptoms are understood. This leads to the targeting of maladaptive relationship patterns, which may have developed after key childhood experiences such as trauma or sexual abuse. Solutions to these interpersonal difficulties are firstly tested out in therapy, before subsequently being implemented outside of the therapy environment. Treatment is essentially an exercise in identification and solving of interpersonal problems, which facilitates a greater sense of control and autonomy and therefore reduces the overall degree of sustained arousal in interpersonal difficulties, producing improvement in symptoms.

Psychodynamic Therapy

Psychodynamic psychotherapy is similar to interpersonal psychotherapy and requires a close relationship between the patient and the therapist, in which difficulties in interpersonal relationships can be highlighted. As the patient understands these problems, he or she may act upon these insights that may lead to a reduction in symptoms. Analytical psychotherapy attempts to provide the patient with an insight into why particular symptoms have developed and what they might mean/represent in the light of changes in key relationships. There is an implicit assumption that this insight will cause long-lasting changes in attitude and behavior. Symptoms often seem to stem from significant life changes (often a loss of a relationship) set against the background of a fragile personality and difficulty coping with separation. The patient works through his or her interpersonal difficulties through the safe and contained relationship with the therapist. There is undoubtedly an important placebo (nonspecific) effect, which the therapist seeks to exploit.

Hypnotherapy

Hypnosis is used to induce a state of relaxation with heightened suggestibility, in which the hypnotherapist uses progressive muscular relaxation with suggestions of relaxation to reduce striated muscle tension and to relax GI smooth muscle. This then alters the underlying abnormalities of gut motility and/or sensation. This is firstly initiated in the presence of the therapist, with the ultimate aim of enabling patients to control symptoms on their own. Patients are asked to practice autohypnosis at home using an audiotape, with the goal of being able to self-administer suggestions of relaxation. Unfortunately, the success of therapy can be dependent on the enthusiasm and skill of the therapist and is both time consuming and expensive to provide. Studies, however, confirm that long-term efficacy is cost-effective in severe refractory cases, and suggest that treatment is more likely to be successful with younger patients and those without serious psychopathology.

Biofeedback Therapy

This therapeutic approach depends on feedback as a measure of visceral function to show the patient a disturbance in physiology so that they themselves can learn to correct it. Heart rate variability feedback in relaxation and stress reduction is a growing application in functional GI disorders, but the most common use of biofeedback is in the treatment of incontinence and constipation. The therapy aims to make the patient more sensitive to rectal sensation and avoid inappropriate straining. It also provides a detailed explanation of normal physiology and a re-education concerning the optimum defecatory patterns. A supportive relationship

that the therapist can offer such patients undoubtedly contributes to the overall positive effect.

CLINICAL CASES AND LEARNING POINTS

Clinical Case of Inflammatory Bowel Disease

A 62-year-old woman, with a known history of UC, presented to the gastroenterology outpatient with a 2-week history of rectal bleeding, abdominal discomfort, and profuse diarrhea. On clinical examination, she was tachycardic and her left iliac fossa was tender in the absence of guarding. Apart from raised inflammatory markers, her hematological, biochemical, and microbiological markers were normal. She was admitted to the clinic and a colonoscopy was undertaken, whose appearances are shown in Figure 27.3–4. Following the colonoscopy, a diagnosis of an acute exacerbation of UC was made and the patient was commenced on oral 5-aminosalicates and intravenous hydrocortisone 100 mg qds. After 48 hours of therapy, her GI symptoms began to improve rapidly. On the third day of her admission, she began to become confused and agitated. Hematological and biochemical parameters were normal and an infection screen did not demonstrate any abnormalities. A computerized tomography (CT) scan of her head was normal. Over the next 3 days, she began to become increasingly anxious and suspicious of staff members, suggesting that they belonged to the secret service. A liaison psychiatric was called and a diagnosis of steroid induced psychosis was made. She was commenced on regular haloperidol and the intravenous steroids were rapidly tapered and discontinued. She was commenced on the immunosuppressive agent azathioprine. Her psychiatric symptoms improved over the next 10 days and she was discharged home.

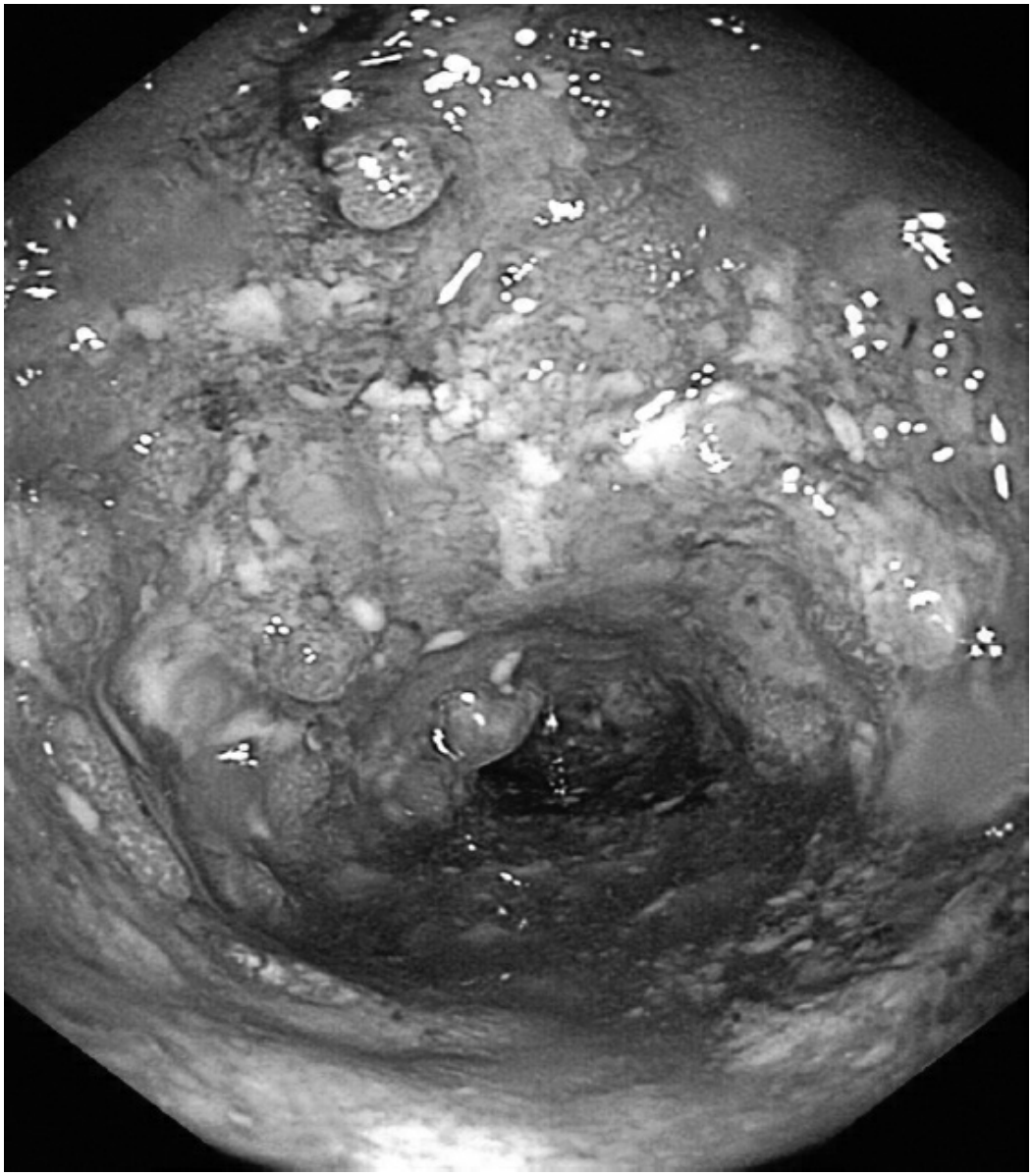

FIGURE 27.3–4. Appearances of the rectum at colonoscopy. The mucosa is friable with marked erythema, absent vascular pattern, and deep ulceration. These features are consistent with severe disease.

Learning points

1. Steroid-induced psychosis is uncommon but can present in patients with IBD.
2. In such cases, discontinuation of steroids and appropriate management of psychosis is needed.
3. A number of factors have been postulated to confer a higher risk of steroid-induced psychosis including female gender and higher doses of steroids.

Clinical Case of Irritable Bowel Syndrome

A 35-year-old woman television producer presented with a 3-year history of increasing

abdominal pain, distension, and alternating bowel habit. She described her pain as intermittent and cramping in nature. The intensity was 8/10 (10, being the worst pain possible), to the degree that it was influencing her ability to perform socially and at work. Her pain was temporarily relieved by defecation. She gave no past psychiatric history and denied any affective symptoms. She stressed that her only problem was her GI symptoms and her not being able to effectively cope with them at work. Her clinical examination was unremarkable as were standard hematological and biochemical parameters. Her antitissue transglutaminase antibodies were normal. Colonoscopic examination of the lower GI tract was macroscopically and microscopically normal, although she found the procedure very painful. Standard medical therapies did not improve her symptoms. Her gastroenterologist referred her to the local liaison psychiatrist service. She was not anticipating benefit from this referral, as she believed it was not relevant to her GI symptoms.

After taking time to reassure her that the consultation was primarily aimed at dealing with her GI symptoms and sufficiently validating her physical problems, she relaxed and the rapport improved. On account of this deepening in rapport, she then felt confident enough to volunteer more information about the degree her physical symptoms has been affecting her. She described easy fatigability, anhedonia, sleep disturbance, weight loss, and low mood. She then discussed how she was experiencing marital and work stress "as a result of her physical problems" of late, and how her "perfectionism and self-criticism" had always been a problem for her in similar situations in the past. The consultation then logically continued to make connections between stress, personality, and somatic symptoms, and how best to treat these.

Learning points

1. Patients with IBS often report discomfort during lower GI endoscopic investigations as a result of visceral hypersensitivity.
2. Standard medical therapies often fail to improve symptoms in IBS.
3. In many cases, patients are often unwilling to engage with psychiatric services.
4. Chronic functional GI symptoms are frequently associated psychopathology.

Clinical Case of Functional Abdominal Pain Syndrome

A 19-year-old woman was referred to the gastroenterology clinic with long-term chronic abdominal pain that had initially begun during her early teenage years. She had seen five different gastroenterologists in the past. She had abdominal pain most of the time and described the pain as burning in nature. The pain was not associated with a change in bowel habit and was not relieved by defecation. There was no association with her menstrual cycle. She also had multiple tender points across her shoulders and back, which her general practitioner had diagnosed as fibromyalgia. She has been admitted to hospital in excess of 30 times over the past 12 months. These acute presentations had been in the context of multiple and repeated normal hematological, biochemical, metabolic, immunological, endoscopic, and radiological investigations. Of specific note, she had had a normal CT scan of her abdomen and pelvis and ileocolonoscopy with histology in the last 12 months. The gastroenterologist made a clinical diagnosis of FAP.

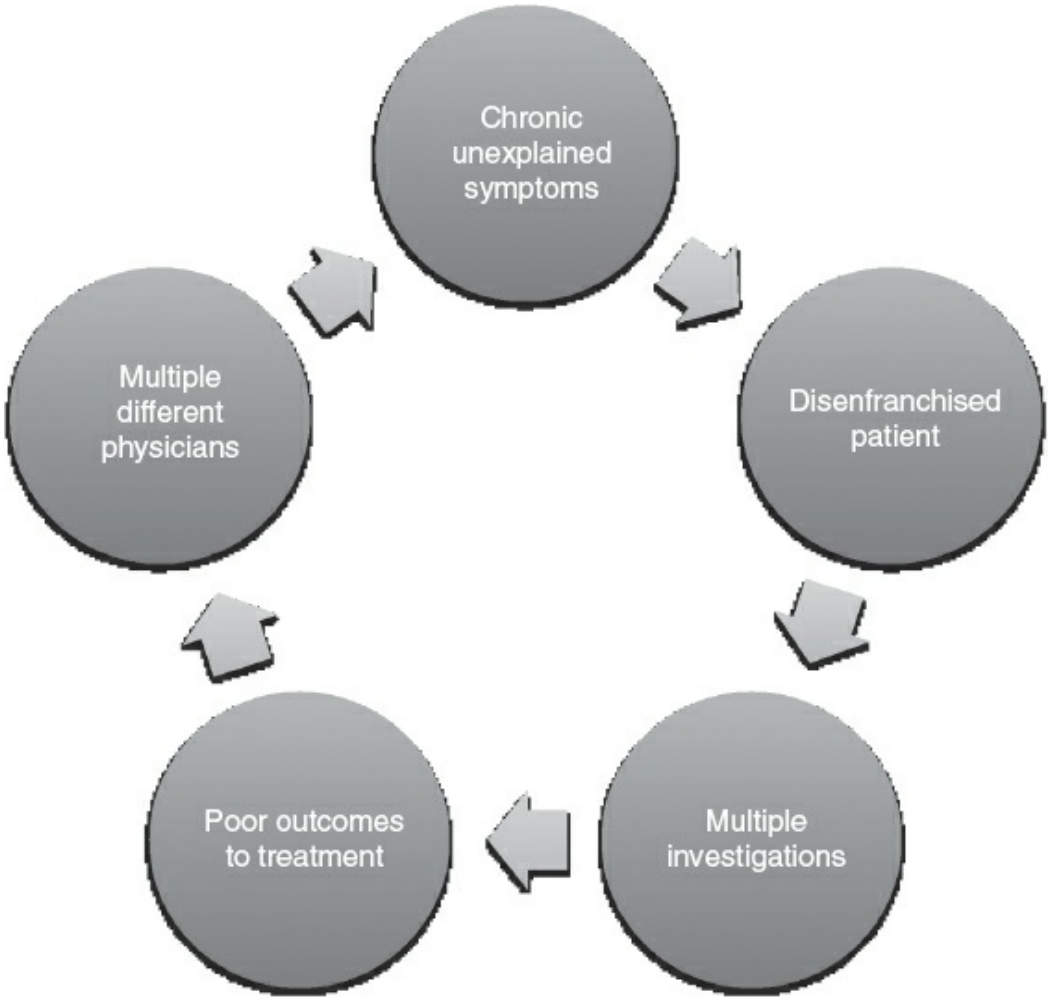

FIGURE 27.3–5. A patient presenting with chronic unexplained symptoms is frequently investigated multiple times, across multiple physicians leading to poor outcomes. Such a cycle can only be interrupted with time, empathy, and validation of the patient's symptoms.

Learning points

1. Patients with one functional GI disorder often suffer with other chronic somatic pain syndromes, such as fibromyalgia.
2. There is often a delay in making the diagnosis of a functional GI disorder.
3. Patients with chronic unexplained symptoms are often disenfranchised with the medical profession. The lack of a specific diagnosis often leads to multiple investigations, which in itself may heighten a patient's anxiety. Frequently, the consequences of this are poor outcomes to treatment and onward referral to another physician and thus the cycle begins again (see Fig. 27.3–5). This cycle can be broken with time, empathy, and validation of the patient's symptoms.

Clinical Case of Functional Chest Pain

A 47-year-old man, who had had a myocardial infarction and coronary artery bypass graft 17 years prior to presentation, presented with recurrent episodes of severe onset central chest pain. He had presented on multiple occasions to his local accident and emergency department, where standard testing did not demonstrate a cardiac cause. Onward referral to a cardiologist, who had performed a coronary angiogram, did not demonstrate recurrent ischemic heart disease. The cardiologist arranged referral to a gastroenterologist who performed an upper GI endoscopy with biopsies, barium meal, and abdominal ultrasound, which were all found to be within normal limits. Further assessment included esophageal manometry and 24-hour pH testing that again was normal. A therapeutic trial of proton pump inhibitors did not improve the symptoms. A diagnosis of FCP of presumed esophageal origin was made. After a period of 15 months, the gastroenterologist, frustrated with the lack of clinical progress, referred him to the psychiatrist. The psychiatrist found him depressed and tearful. He had also developed panic attacks. The psychiatrist diagnosed major depressive disorder with panic attacks. After a detailed discussion with a cardiologist about the possible risks of antidepressants, he was treated with an SSRI antidepressant and improved rapidly after 4 or 5 weeks.

Learning points

- ► Functional GI disorders may occur in males.
- ► Functional GI disorders frequently occur concurrently with organic disease.
- ► Numerous bodily symptoms may attract numerous investigations for physical diseases. This led to a profound delay in the detection of this patient's depressive disorder and, therefore, prolonged suffering and impairment of health-related quality of life.
- ► Treatment of this patient's depressive disorder was relatively straightforward, but there needed to be a joint decision with the cardiologist regarding the safety of antidepressant treatment in patients with severe heart disease.

CONCLUDING REMARKS

The gut and the brain are intimately linked and any disorder in one can manifest in the other. Therefore, psychiatry can play a pivotal therapeutic role in the management of patients with GI disorders.

REFERENCES

Abreu MT. The genetics and pathogenesis of inflammatory bowel disease. *Gastroenterol Hepatol (NY)*. 2013;9:521–523.

Barreau F, Ferrier L, Fioramonti J, Bueno L. New insights in the etiology and pathophysiology of irritable bowel syndrome: contribution of neonatal stress models. *Pediatr Res*. 2007;62:240–245.

Blanchard EB, Greene B, Scharff L, Schwarz-McMorris SP. Relaxation training as a treatment for irritable bowel syndrome. *Biofeedback Self Regul*. 1993;18:125–132.

Canavan JB, Bennett K, Feely J, O'Morain CA, O'Connor HJ. Significant psychological morbidity occurs in irritable bowel syndrome: a case-control study using a pharmacy reimbursement database. *Aliment Pharmacol Ther*. 2009;29:440–449.

Canavan C, West J, Card T. Review article: the economic impact of the irritable bowel syndrome.

Aliment Pharmacol Ther. 2014;40:1023–1034.

Chang L, Lembo A, Sultan S. American Gastroenterological Association Institute Technical Review on the Pharmacological Management of Irritable Bowel Syndrome. *Gastroenterology*. 2014;147:1149–72 e2.

Clouse RE, Mayer EA, Aziz Q, et al. Functional abdominal pain syndrome. *Gastroenterology*. 2006;130:1492–1497.

Drossman DA. *Rome III. The Functional Gastrointestinal Disorders*. McLean, VA: Degnon Associates; 2006.

Drossman DA, Chang L, Bellamy N, et al. Severity in irritable bowel syndrome: a Rome Foundation Working Team report. *Am J Gastroenterol*. 2011;106:1749–1759; quiz 60.

Drossman DA, Li Z, Andruzzi E, et al. U.S. householder survey of functional gastrointestinal disorders. Prevalence, sociodemography, and health impact. *Dig Dis Sci*. 1993;38:1569–1580.

Drossman DA, Whitehead WE, Toner BB, et al. What determines severity among patients with painful functional bowel disorders? *Am J Gastroenterol*. 2000;95:974–980.

Engel GL. The need for a new medical model: a challenge for biomedicine. *Science*. 1977;196:129–136.

Faresjo A, Grodzinsky E, Johansson S, Wallander MA, Faresjo T, Timpka T. Self-reported use of pharmaceuticals among patients with irritable bowel syndrome in primary care. *J Manag Care Pharm*. 2008;14:870–877.

Ford AC, Talley NJ, Schoenfeld PS, Quigley EM, Moayyedi P. Efficacy of antidepressants and psychological therapies in irritable bowel syndrome: systematic review and meta-analysis. *Gut*. 2009;58:367–378.

Graff LA, Walker JR, Bernstein CN. Depression and anxiety in inflammatory bowel disease: a review of comorbidity and management. *Inflamm Bowel Dis*. 2009;15:1105–1118.

Gwee KA, Leong YL, Graham C, et al. The role of psychological and biological factors in postinfective gut dysfunction. *Gut*. 1999;44:400–406.

Hungin AP, Whorwell PJ, Tack J, Mearin F. The prevalence, patterns and impact of irritable bowel syndrome: an international survey of 40,000 subjects. *Aliment Pharmacol Ther*. 2003;17:643–650.

Kirmayer LJ, Robbins JM, Dworkind M, Yaffe MJ. Somatization and the recognition of depression and anxiety in primary care. *Amer J Psychiatry*. 1993;150:734–741.

Kroenke K, Jackson JL, Chamberlin J. Depressive and anxiety disorders in patients presenting with physical complaints: clinical predictors and outcome. *Am J Med*. 1997;103:339–347.

Kuiken SD, Tytgat GN, Boeckxstaens GE. The selective serotonin reuptake inhibitor fluoxetine does not change rectal sensitivity and symptoms in patients with irritable bowel syndrome: a double blind, randomized, placebo-controlled study. *Clin Gastroenterol Hepatol*. 2003;1:219–228.

Lackner JM, Jaccard J, Krasner SS, Katz LA, Gudleski GD, Holroyd K. Self-administered cognitive behavior therapy for moderate to severe irritable bowel syndrome: clinical efficacy, tolerability, feasibility. *Clin Gastroenterol Hepatol*. 2008;6:899–906.

Langhorst J, Wulfert H, Lauche R, et al. Systematic review of complementary and alternative medicine treatments in inflammatory bowel diseases. *J Crohns Colitis*. 2015;9:86–106.

Long MD, Drossman DA. Inflammatory bowel disease, irritable bowel syndrome, or what? A

challenge to the functional-organic dichotomy. *Am J Gastroenterol*. 2010;105:1796–1798.

Lydiard RB. Irritable bowel syndrome, anxiety, and depression: what are the links? *J Clin Psychiatry*. 2001;62 Suppl 8:38–45; discussion 6–7.

Mawdsley JE, Rampton DS. Psychological stress in IBD: new insights into pathogenic and therapeutic implications. *Gut*. 2005;54:1481–1491.

Mikocka-Walus A, Turnbull D, Moulding N, Wilson I, Andrews JM, Holtmann G. Psychological comorbidity and complexity of gastrointestinal symptoms in clinically diagnosed irritable bowel syndrome patients. *J Gastroenterol Hepatol*. 2008;23:1137–1143.

Miller V, Whorwell PJ. Treatment of inflammatory bowel disease: a role for hypnotherapy? *Int J Clin Exp Hypn*. 2008;56:306–317.

Nahon S, Lahmek P, Durance C, et al. Risk factors of anxiety and depression in inflammatory bowel disease. *Inflamm Bowel Dis*. 2012;18:2086–2091.

Palsson OS, Drossman DA. Psychiatric and psychological dysfunction in irritable bowel syndrome and the role of psychological treatments. *Gastroenterol Clin North Am*. 2005;34:281–303.

Richter JE, Barish CF, Castell DO. Abnormal sensory perception in patients with esophageal chest pain. *Gastroenterol*. 1986;91:845–852.

Shivaji N, Ford AC. Prevalence of functional gastrointestinal disorders among consecutive new patient referrals to a gastroenterology clinic. *Frontline Gastroenterol*. 2014;5:266–271.

Simon GE, VonKorff M, Piccinelli M, Fullerton C, Ormel J. An international study of the relation between somatic symptoms and depression. *NEJM*. 1999;341:1329–1335.

Sirois F. Steroid psychosis: a review. *Gen Hosp Psychiatry*. 2003;25:27–33.

Spiller R, Aziz Q, Creed F, et al. Guidelines on the irritable bowel syndrome: mechanisms and practical management. *Gut*. 2007;56:1770–1798.

Szigethy E, Bujoreanu SI, Youk AO, et al. Randomized efficacy trial of two psychotherapies for depression in youth with inflammatory bowel disease. *J Am Acad Child Adolesc Psychiatry*. 2014;53:726–735.

Tougas G, Spaziani R, Hollerbach S, et al. Cardiac autonomic function and oesophageal acid sensitivity in patients with non-cardiac chest pain. *Gut*. 2001;49:706–712.

Tyler K, Moriceau S, Sullivan RM, Greenwood-van Meerveld B. Long-term colonic hypersensitivity in adult rats induced by neonatal unpredictable vs predictable shock. *Neurogastroenterol Motil*. 2007;19:761–768.

Valet M, Gundel H, Sprenger T, et al. Patients with pain disorder show gray-matter loss in pain-processing structures: a voxel-based morphometric study. *Psychosom Med*. 2009;71:49–56.

Whitehead WE, Levy RL, Von Korff M, et al. The usual medical care for irritable bowel syndrome. *Aliment Pharmacol Ther*. 2004;20:1305–1315.

Whorwell PJ, Prior A, Colgan SM. Hypnotherapy in severe irritable bowel syndrome: further experience. *Gut*. 1987;28:423–425.

Whorwell PJ, Prior A, Faragher EB. Controlled trial of hypnotherapy in the treatment of severe refractory irritable-bowel syndrome. *Lancet*. 1984;2:1232–1234.

▲ 27.4 Psychosomatic Medicine: Obesity

DEFINITION, NOSOLOGY, AND EPIDEMIOLOGY

Obesity is a global epidemic with physical, economical, and psychological costs. The prevalence of obesity in the United States is 78 million adults, with approximately 69 percent of US adults in the overweight range, and 35 percent of Americans meeting criteria for obesity according to the 2011 to 2012 National Health and Nutrition Examination Survey (NHANES) (Table 27.4–1). In the United States, approximately 17 percent of children are obese. In both adults and children, there is a higher rate of obesity among minority groups including Hispanics and African Americans. Rates of obesity have shown an alarming trend upward since the 1980s, with a leveling off among adults between 2003 to 2004 and 2010 to 2011.

Obesity is defined primarily by the body mass index (BMI). BMI is measured as weight in kilograms divided by height in meters squared (kg/m^2). A normal BMI is 18.5 to 24.5, and individuals are considered overweight with BMIs from 25 to 30. BMIs above 30 are in the obesity range, with a BMI above 40 considered morbidly or severely obese (Table 27.4–2).

Table 27.4–1.
Prevalence of Overweight and Obesity among US Adults, Age 20 Years and over, NHANES 2011–2012

	Total	Non-Hispanic White	Non-Hispanic Black	Non-Hispanic Asian	Hispanic
Overweight or obese (BMI greater than or equal to 25)					
All	69.0	68.5	76.3	38.2	77.1
Men	71.6	72.7	69.1	42.4	77.9
Women	66.5	64.6	82.1	34.4	76.2
Obese (BMI greater than or equal to 30)					
All	35.1	33.4	47.8	10.9	42.0
Men	33.7	33.1	37.0	10.4	40.7
Women	36.5	33.7	56.7	11.4	43.3

NHANES, National Health and Nutrition Examination Surveys 2011–2012; BMI, body mass index.

Obesity raises the risk of morbidity from hypertension, dyslipidemia, type 2 diabetes, coronary heart disease (CHD), stroke, gallbladder disease, osteoarthritis, sleep apnea and respiratory problems, and some cancers. Obesity is also associated with increased risk in all-cause and cardiovascular disease mortality. In addition, obesity has a negative impact on quality of life, mobility, mental health, and stress level.

The economic costs of obesity and its associated health problems place a

heavy burden on the US health-care system. The costs of preventing, diagnosing, and treating obesity are directly observable, whereas the indirect costs may include loss of income, decreased productivity, restricted activity, and absenteeism. Costs may include the value of future income lost by premature death. According to the Centers for Disease Control, the estimated medical cost of obesity is $147 billion yearly in the United States.

DIAGNOSIS, CLINICAL FEATURES

BMI has been widely used to diagnose obesity. BMI charts classify a patient's weight as underweight if less than 20, healthy if 20 to 25, overweight if 25 to 30, obese if 30 to 40, and severely or morbidly obese if greater than 40. Another indicator of obesity is the measurement of waist circumference. Waist circumference greater than 35 inches for women or greater than 40 inches for men is considered obese. High waist circumference may elevate risks for obesity-related disease. Obesity is not included as a diagnosis in DSM-5, although binge eating disorder (BED) has been formally accepted as a diagnosis as discussed later.

Table 27.4–2.
Classification of Obesity and Recommended Treatment

	Obesity Class	Body Mass Index (kg/m^2)	Recommended Treatment
Underweight		<18.5	Nutrition/eating disorder treatment
Normal		18.5–24.9	Maintenance
Overweight		25–29.9	BWM
Obese	I Mild	30–34.9	BWM + MWM
	II Moderate	35–39.9 (plus comorbidities)	BWM + MWM or BWM + WLS
	III Severe	>40	BWM + WLS

BWM, behavioral weight management; MWM, medical weight management; WLS, weight loss surgery.

Obesity in childhood is diagnosed using an age- and sex-specific percentile for BMI, due to the variations in body composition with age and gender during childhood. Children with a BMI between the 85th and 95th percentile for children of the same age and gender are considered overweight, and children with a BMI above the 95th percentile are considered obese.

COURSE AND PROGNOSIS

Childhood obesity is a risk factor for adult obesity and related to the severity of adult obesity. Children with obesity are more likely to develop cardiovascular risk factors such as high blood pressure and high cholesterol; impaired glucose tolerance, insulin resistance and type 2 diabetes; breathing problems such as sleep apnea and asthma; joint and pain problems; fatty liver disease, gallstones, and gastroesophageal reflux; and social and psychological problems, such as stigma and poor self-esteem

(Fig. 27.4–1).

In adults, associated medical comorbidities of obesity include congestive heart failure, coronary artery disease, venous state ulcers, insulin resistance, polycystic ovarian syndrome, diabetes mellitus, hypertension, abdominal hernia, gall stones, gastroesophageal reflux disease, nonalcoholic fatty liver disease, stress urinary incontinence, deep venous thrombosis, pulmonary embolism, gout, plantar fasciitis, depression, anxiety, pseudotumor cerebri, stroke, infertility, asthma, obstructive sleep apnea, hypercholesterolemia, osteoarthritis, cancers, and others (Fig. 27.4–2). The risk of mortality and morbidity increases with BMI above 35 and rises exponentially with the severity of BMI. Obesity and overweight together are the second leading cause of preventable death in the United States. An estimated 300,000 deaths per year are due to the obesity epidemic.

Behavioral and lifestyle interventions can effectively result in modest weight loss, which has a positive effect on comorbidities. However, the rate of recidivism is high, and weight loss is difficult to maintain for a variety of reasons. For patients with severe obesity, weight loss surgery appears to be an effective long-term intervention that results in significant weight loss and resolution of comorbidities.

PATHOLOGY AND LABORATORY CONSIDERATIONS

When working with patients with obesity, clinicians may want to consider the following adjunctive information: routine blood work for hyperlipidemia, hypertension, hypothyroid function, and hyperglucosemia; sleep studies to identify obstructive sleep apnea, and vitamin level panel to identify vitamin deficiencies.

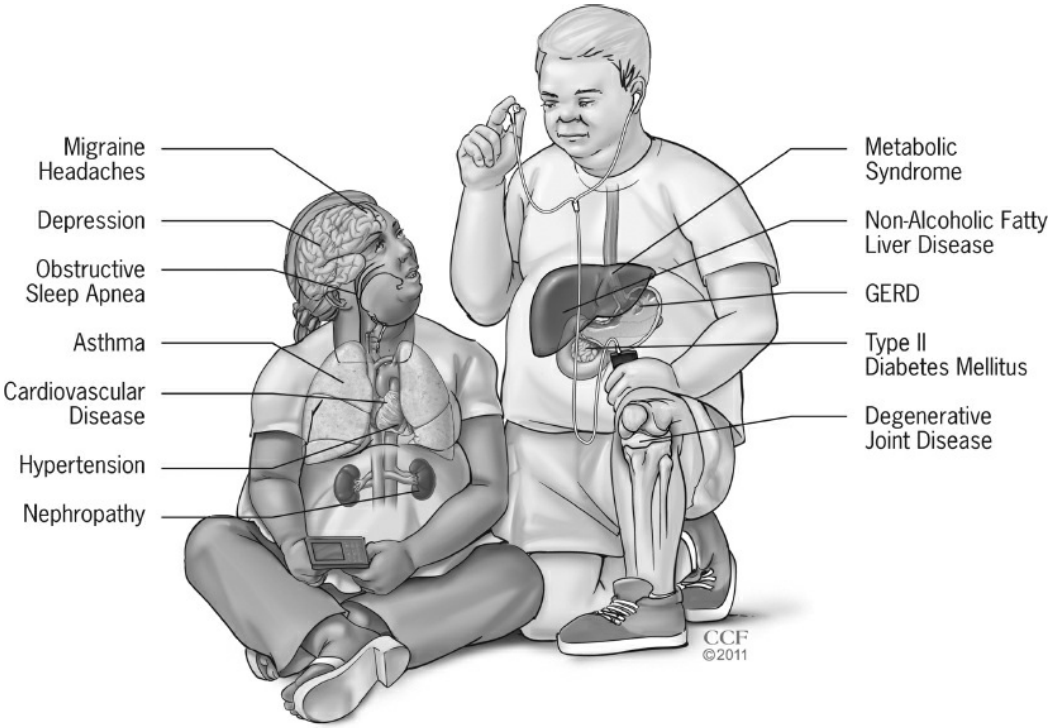

FIGURE 27.4–1. Medical comorbidities of obesity in adolescents. (Reprinted with permission, Cleveland Clinic Center for Medical Art & Photography © 2006–2015. All Rights Reserved.)

ETIOLOGY

Fundamentally, obesity is the result of an imbalance between energy intake and expenditure. However, the development of obesity is far more complex than this simplified view portrays. Given the alarming rise in obesity rates, researchers have increasing sought to elucidate the multifaceted causes of obesity in hopes of finding strategies to combat this epidemic.

Genetics, Biology, and Obesity

From an evolutionary perspective, the ability to store energy as adipose tissue would have yielded a selective advantage for survival and reproduction. Using this concept, Neel (1962) proposed the notion of thrifty genotypes. This theory hypothesizes that thrifty genotypes code for effective energy storage. Through natural selection, genes with the ability to save and store energy likely overrode other genes, resulting in many thrifty genes undergoing such selection. In fact, researchers have found over 600 single gene mutations and genomic and chromosomal regions potentially related to obesity. Thus, due to natural selection, many individuals may be genetically susceptible to developing obesity.

In addition, twin and adoption studies have underscored the role of genes in obesity. This body of literature suggests that heritable genetic differences account for 55 to 85 percent of weight variance. Interestingly, several twin and adoption studies also highlight how most of the

nongenetic factors influencing weight may stem from nonshared environmental contexts unique to a given person (e.g., school environment and peer relationships). Contrary to common assumption, the evidence suggests that the shared family environment may have a minimal effect on BMI. Therefore, genes and individual nonshared experiences may have a more significant impact on weight than family environment.

Increasingly, adipose tissue and hormones have also been implicated as biological contributors to obesity. More specifically, there is growing evidence that enlarged adipocytes in obese individuals create inflammation and result in the dysregulation of metabolic homeostasis; impacting levels of leptin, insulin, and adiponectin; thus creating metabolic syndrome. Researchers have just begun to understand how such changes can impact weight, as well as exercise, food intake, and overall hormonal control.

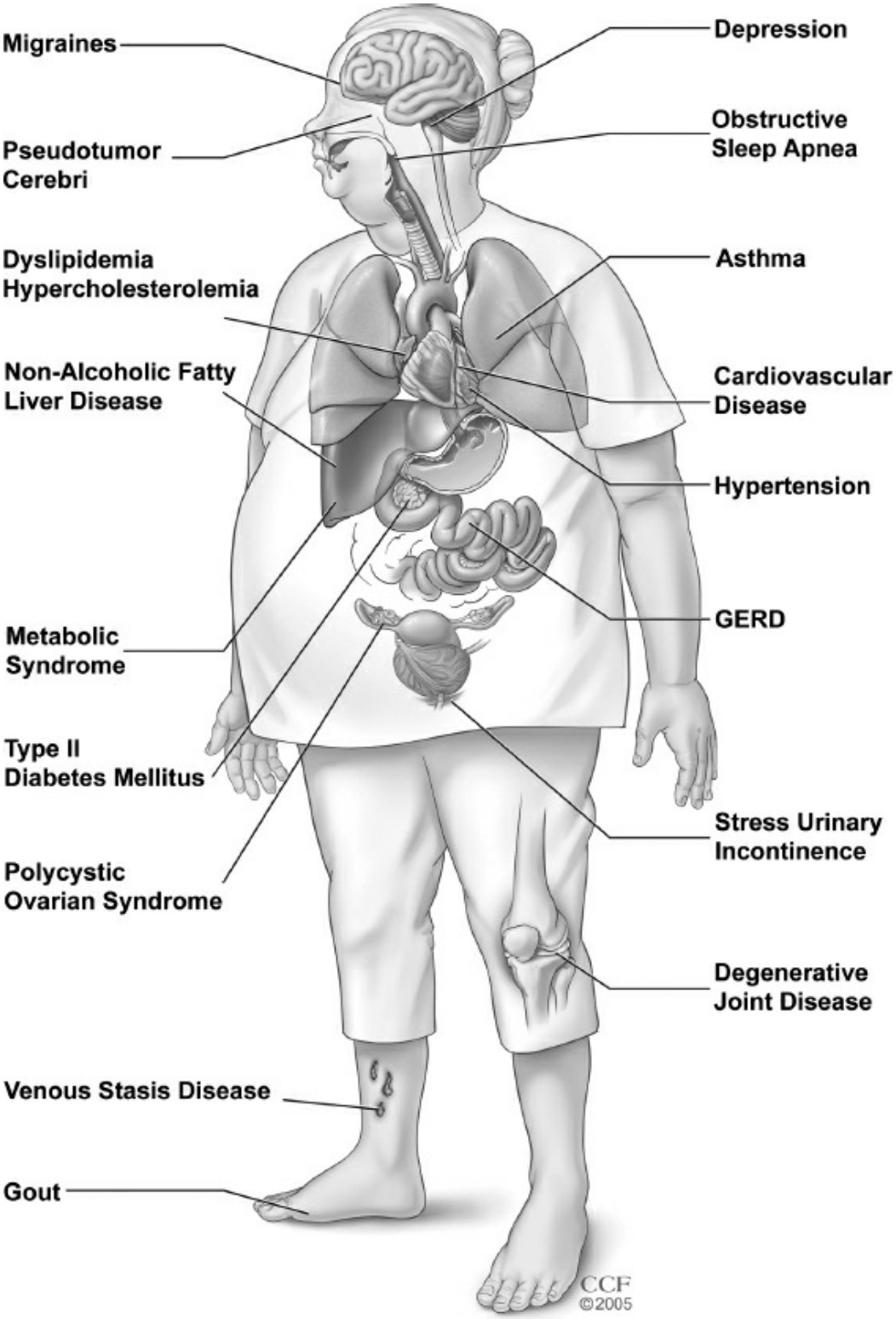

FIGURE 27.4–2. Medical comorbidities of obesity in adults. (Reprinted with permission, Cleveland Clinic Center for Medical Art & Photography © 2006–2015. All Rights Reserved.)

Environment and Obesity

However, genetics in and of themselves do not explain the rising rates of obesity globally. A large body of research has underscored the importance

of global environmental influences on eating behavior. For most of history, humans have struggled to combat food shortage, disease, and hostile surroundings. However, the modern industrialized environment stands in stark contrast to this type of setting. Instead, the current environment is overflowing with calorically dense foods, along with energy-saving devices. To capture these changes, Swinburn (1999) developed the term "obesogenic environment." Researchers contend that the environments of most industrialized nations promote a positive energy balance through the availability and affordability of highly palatable, energy dense foods, as well as numerous opportunities for inactivity (e.g., car ownership, television, sedentary job environments, and the supermarket system). In the current environment, expending an ample amount of energy typically requires individuals to engage in deliberate, consistent healthy lifestyle choices. Unfortunately, making a purposeful choice to be active is at odds with humans' biological propensity to conserve energy. Therefore, humans' biological predisposition makes them extremely vulnerable to overeating and inactivity; particularly in an environment with endless food and inactivity cues. The obesogenic environment has thus created an ideal setting for the growth of obesity. Many researchers and lobbyists have argued for national and global food reform, including taxation on sugary foods and drinks, as a primary means of combating obesity.

Food Addiction and Obesity

Despite a common food environment, research indicates that the susceptibility to food cues varies across individuals and serves as a predictor of BMI, the ability to resist cravings, and dietary compliance. Within this framework, researchers contend that food can become an addiction akin to other substance use disorders. More specifically, studies suggest that recurrent exposures to highly palatable foods result in diminished dopamine levels in response to eating. As a result, individuals must consume larger amounts of food to feel satisfied; resulting in a cycle of overeating. The term *food addiction* (FA) has been used to capture the resulting pattern of excessive consumption. It is thought that FA to highly palatable, processed foods could be a key contributor to the development of obesity. Therefore, the availability of calorically dense foods, combined with a potential for physiological addiction to these foods, may play a central role in the increasing rates of obesity.

The Yale Food Addiction Scale (YFAS) was specifically developed to measure FA by applying DSM-IV criteria for substance dependence to eating behaviors. To assess the presence of FA further, a 2014 review aimed to systematically examine studies that used the YFAS. Using meta-analysis, the researchers found that the weighted mean prevalence of FA was 19.9 percent across adult samples. Higher BMI was associated with higher rates of FA diagnosis and number of symptoms. In fact, rates of FA were doubled

in overweight/obese populations compared to healthy controls (24.9 vs. 11.1 percent, respectively). Additionally, in disordered eating samples, the mean prevalence of FA was extremely high (57.6 percent) compared to samples with no eating disorder diagnoses (16.2 percent). Taken together, these results highlight the relationships between FA, weight, and disordered eating.

Table 27.4–3.
DSM-5 Diagnostic Criteria for Binge Eating Disorder

Recurrent episodes of binge eating (eating a larger than normal amount with loss of control at least once per week for 3 months) accompanied by at least three of the following:

- Eating very rapidly
- Eating until uncomfortably full
- Eating when not hungry
- Eating alone or hiding eating from others
- Feeling guilty or disgusted after a binge

No recurrent compensatory purging, exercising, or fasting
Marked distress

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Behaviors, Cognitions, and Obesity

Many individuals report eating in response to negative emotions. Certain internal experiences, such as anxiety, sadness, negative thoughts, and cravings, are frequently felt to be unpleasant. Individuals vary in their ability to tolerate such internal distress. Some individuals may take great effort to escape these experiences via psychological (i.e., suppression) or behavioral (i.e., isolation) avoidance. It is common for individuals to use food as a means of temporarily suppressing emotional distress. In fact, a significant number of patients report eating regularly in response to stress, boredom, fatigue, anxiety, and depression. For many, this pattern may begin in childhood and sustain throughout adulthood. Eating can become a learned response to emotional distress and for some may become a means of coping. Research suggests emotional eating is positively associated with weight and BMI. Thus, emotional eating can develop into a primary maladaptive affect regulation tool, resulting in weight gain.

Dimensions of Personality and Obesity

There is also a growing body of research focusing on the influence of personality traits on weight. In particular, many studies suggest that neuroticism (manifested in greater anxiety, depression, impulsivity, emotional vulnerability, hostility, and anger) is a risk factor for becoming overweight or obese; particularly for women. It is hypothesized that

individuals with higher levels of neuroticism may be more susceptible to using food as a dysfunctional coping strategy. In addition, the literature indicates that there are higher rates of impulsivity in obese populations with and without disordered eating (unhealthy eating patterns often characterized by distress and/or loss of control (see Tables 27.4–3 and 27.4–4). Similarly, high reward sensitivity appears to be another risk factor for obesity. Researchers postulate that individuals with higher sensitivity to reward have a proclivity toward uncontrolled eating, resulting in sustained positive energy balance. However, higher levels of conscientiousness (regulation of internal urges) and self-control have been shown to be protective in regard to weight gain. Given this literature, it may be important to consider personality traits when working with patients struggling with obesity.

PSYCHIATRIC COMORBIDITIES OF OBESITY

Eating Disorders

Eating disorders are common psychiatric comorbidities of obesity and include BED, bulimia, and syndromes such as night eating, graze eating, and loss of control consumption of high-calorie liquids.

Table 27.4–4.
Proposed Criteria for Night Eating Syndrome

1. Core criteria (One of the following):
   A. Evening hyperphagia and/or
   B. Nocturnal awakenings with food ingestion (at least 2×/week)
2. Three of five associated symptoms:
   A. Insomnia: Sleep initiation and maintenance insomnia
   B. Morning anorexia: Either in the form of lack of desire to eat in the morning or skipping breakfast more days than not
   C. Urges to eat at night
   D. Depressed evening mood
   E. A belief that one must eat to sleep

Symptoms must be associated with distress and occur for 3 or more months

Data from Allison KC, Lundgren JD, O'Reardon JP, et al. Proposed diagnostic criteria for night eating syndrome. *Int J Eat Disord*. 2010;43(3):241–247.

BED has been included in the DSM-5 as a stand-alone disorder. BED was previously included in the DSM-IV under the category Eating Disorder Not Otherwise Specified (Table 27.4–3). BED is characterized by eating large amounts in a short period of time with a feeling of loss of control. Associated symptoms include feelings of guilt, eating quickly, eating when not hungry, hiding eating patterns from others, and eating until uncomfortably full. The binge episodes must be distressing and occur at

least once per week for 3 months to meet criteria for BED. The prevalence of BED among patients with obesity is estimated to be 30 percent, and patients with BED tend to be obese. BED appears to be the most common eating disorder, with estimated population rates of 2 percent.

Other eating syndromes sometimes associated with obesity include graze eating and night eating syndrome (NES). Graze eating syndrome is less well defined than NES. Graze eating typically involves eating large amounts slowly over the day or continuous eating. Graze eating is often associated with feelings of loss of control, distress, and eating for reasons other than hunger. A high-calorie liquid type of grazing has been identified as a pattern of drinking high-calorie liquids continuously throughout the day.

In addition, NES is a relatively understudied disorder involving a circadian delay in food consumption. According to the 2010 proposed criteria, the diagnostic criterion for NES requires the presence of evening hyperphagia and/or nocturnal awakenings with food ingestion. Additional symptoms include morning anorexia, urges to eat at night, insomnia, low evening mood, and a belief that one must eat to sleep (Table 27.4–4). There appears to be a relationship between NES and weight in some, but not all, studies. Within the research, NES has been associated with several correlates of distress, including depressed mood, anxiety, low self-esteem, poorer body image, and decreased quality of life. Not surprisingly, there are higher rates of Axis I disorders in NES populations, including current and lifetime major depression. It is also important to note that more distress has been associated with patients who wake up and eat in the middle of the night versus patients with only evening hyperphagia. Given these findings, some consider patients with nocturnal eating episodes to be more severe than patients without nocturnal ingestions. Bulimia nervosa is characterized by binge episodes followed by compensatory behavior such as vomiting, laxative use, or over exercise. Binge/purge episodes must occur at least once per week for 3 months to meet criteria for the disorder, consistent with BED. Patients with bulimia may be obese but are more typically of normal weight.

Depression

Depression and obesity often overlap as disease processes. This is perhaps unsurprising as increased appetite and low energy are among the diagnostic symptoms for depression (DSM-5). There is a high incidence of depression among obese patients, with estimates ranging from 19 to 66 percent. The relationship between obesity and depression appears to be reciprocal, with 80 percent of studies suggesting obesity contributes to depression, and 53 percent of studies suggesting that depression leads to obesity. Obesity may lead to depression by negatively affecting health and body image, increasing functional impairment, decreasing quality of life

and physical activity, and adding to stigma. Depression may lead to obesity by contributing to an elevated stress reaction, impairing immunity, negative cognitions, eating behaviors, and poor adherence to treatment. Factors moderating the relationship between depression and obesity include gender, age, socioeconomic status, and severity of obesity; with females, increasing age, and lower income more likely to experience depression with obesity. Other factors that increase the likelihood of depression with obesity include binge eating behaviors and experiencing stigma.

Anxiety

Anxiety disorders are the most common psychological disorders affecting 25 percent of the population. Obesity may lead to anxiety by causing distress from weight-related stigma. In addition, the physical comorbidities and negative effect of obesity on quality of life may lead to stress and cause anxiety disorders. However, the prospective evidence for obesity causing anxiety is weak. A review of the literature suggests there is a moderate amount of evidence for a weak but positive association between anxiety and obesity. There appears to be a stronger relationship between severe obesity and anxiety, and among anxiety disorders, specific phobia, social phobia and PTSD may have the strongest associations with obesity. Anxiety disorders may contribute to obesity through the HPA axis dysfunction, and anxiety may increase cravings for high sugar and high fat foods. Anxiety may also be related to obesity through negative childhood events such as abuse and eating disorders, with patients developing anxiety in response to trauma and those with anxiety having a higher risk of eating disorder behaviors.

Severe Mental Illness

Patients with severe mental illness (SMI), such as schizophrenia and bipolar disorder, are at increased risk of obesity, with identified rates up to 60 percent. Although weight gain side effects of psychotropic medications appear to play a major role, there is some evidence of being overweight among medication naïve patients. Patients with SMI also have an increased risk of medical conditions related to obesity including type 2 diabetes, cardiovascular disease, and metabolic syndrome. Individuals with schizophrenia and bipolar disorder have a shorter lifespan and an increased risk of death from medical causes. Weight gain side effects of psychotropic medications are a cause for obesity in persons with SMI, with atypical antipsychotics associated with more weight gain than typical agents, and patients treated with a combination of atypicals and mood stabilizers showing the highest weight gain. In persons with SMI, binge eating behaviors and female gender were also associated with higher rates

of obesity. Clinicians should closely monitor weight in patients with SMI, and include regular measurement of BMI, waist circumference, lipid profile, and plasma glucose level. Consideration of adjunctive therapy such as cognitive behavioral weight management and diet/exercise/lifestyle changes can mitigate weight gain effects in persons with SMI. Other options include choosing the medication that best treats the patient's mental disorder with the least weight gain profile, switching antipsychotics to promote weight loss in patient who experience significant weight gain, and considering adjunctive treatment with medications that promote weight loss.

Psychiatric Medications and Weight Gain. Weight gain may contribute to nonadherence to psychotropic medications, and the majority of antidepressants, antipsychotics, and mood stabilizers contribute to weight gain (Table 27.4–5). Antipsychotics appear to have some of the highest rates of weight gain, including olanzapine (2.3 kg/month), quetiapine (1.8 kg/month), and clozapine (1.7 kg/month).

Weight gain due to antidepressants appears to be under-recognized and an important consideration given the high prescribing volume of these medications. Tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs) appear to have higher rates of weight gain than serotonin selective reuptake inhibitors (SSRIs). Weight gain is associated with mirtazapine, whereas bupropion may be the only antidepressant agent that contributes to weight loss. On average, antidepressants result in a 1 to 3 kg weight gain. The short-term weight loss seen with fluoxetine appears to be transient, with later weight gain offsetting this effect. Weight gain is a common side effect of mood stabilizers, with divalproex, carbamazepine, and lithium appearing to cause the most gain (0.9 to 15 kg). Lamotrigine appears to be more weight neutral, and zonisamide may promote weight loss.

The causes of weight gain with psychotropic use may include increased calorie intake, slowing of the metabolism, and increased appetite. Patients should be alerted to weight gain as a potential side effect and closely monitored with routine blood work for hyperlipidemia, hypertension, and hyperglucosemia. Patients being prescribed psychotropics with weight gain side effects should be given recommendations simultaneously for diet and exercise as a preventative measure. Careful consideration to the weight gain profile of a medication should be given, especially in patients already with increased BMI and comorbidities. In cases where the benefits of the medication on the psychiatric symptoms appear to outweigh the risks of weight gain, consideration of adjunctive weight loss medications may be considered.

Table 27.4–5.

Psychotropic Medications and Weight Gain

Psychotropic Medication	Weight Gain Profile
Antipsychotics	
Olanzapine (Zyprexa)	+++
Quetiapine (Seroquel)	+++
Risperidone (Risperdal)	++
Ziprasidone (Geodon)	+
Asenapine (Saphris)	+
Iloperidone (Fanapt)	++
Paliperidone (Invega)	0
Aripiprazole (Abilify)	+
Clozapine (Clozaril)	+++
Antidepressants	
Imipramine (Tofranil)	+
Amitriptyline (Elavil)	+
Trazodone (Oleptro)	+
Fluoxetine (Prozac)	0/+
Citalopram (Celexa)	+
Fluvoxamine (Luvox)	0
Duloxetine (Cymbalta)	+
Escitalopram (Lexapro)	+
Bupropion (Wellbutrin)	–
Mirtazapine (Remeron)	+
Paroxetine (Paxil)	++
Sertraline (Zoloft)	+
Venlafaxine (Effexor)	0
Mood Stabilizers/Anticonvulsants	
Divalproex (Depakote)	++
Lamotrigine (Lamictal)	0
Carbamazepine (Tegretol)	++
Lithium (Lithobid)	++
Zonisamide (Zonegran)	–
Topiramate (Topamax)	–

–, Weight loss; 0, Weight neutral; +, Weight gain (mild); ++, Weight gain (moderate); +++, Weight gain (significant).

Substance Use Disorders

Researchers have found similarities in the neurobiology between obesity and substance use disorders, suggesting that these common paths might also have treatment indications. Some conceptualize obesity as an FA with compulsive overeating as previously discussed. Drugs and addictive foods may activate similar brain regions involved in reward processing and decision making.

There appears to be a positive relationship between nicotine

dependence and obesity. Epidemiological studies also point to a relationship between alcohol use disorders and obesity, although no consistent relationships between cocaine/stimulant use and obesity have been found. The co-occurrence of substance use disorders and obesity suggests that treatments such as CBT, diet/exercise, and medications targeting underlying biological mechanism in the reward circuitry of the brain may be useful for both conditions. Considering that substance use remission often results in weight gain, awareness of weight management in treating substance use disorders is important.

TREATMENT

Treatments for obesity are tailored to the risk level and individual needs of the patient (Table 27.4–2). The basis for all weight management treatment is behavioral lifestyle intervention including behavioral modification, diet, and exercise. Patients in the overweight range (BMI between 25 and 30) are encouraged to engage in behavioral weight management. Patients with mild obesity (BMI >30) are directed to add medical weight management, and patients with severe obesity (BMI >40) are candidates for weight loss surgery.

Benefits of Weight Loss

A large body of literature calls attention to the physical and psychological benefits of even small weight reductions. For instance, a weight loss of 5 to 10 percent of initial body weight has demonstrated clinically significant improvements across physical domains, including blood pressure and lipid levels. Similarly, a 3 percent decrease in weight has been associated with decreased metabolic risk factors. Weight loss of only 1 to 9 lb has been correlated with decreased mortality in individuals who are overweight and have diabetes. In addition, studies have consistently shown a positive relationship between weight loss efforts and psychological variables, including body image, health-related quality of life, self-esteem, and depression. Arguably, this body of evidence supports the concept that a weight loss goal is never too small to result in benefit.

Commercial Diets

Despite the complex etiology of obesity, individual lifestyle and behavioral change can play a critical role in addressing factors influencing weight. Many individuals may initially attempt weight loss through commercial diets. For individuals who struggle with their weight, there are numerous commercial diets available with various food and activity goals (see Table 27.4–6 for details and comparisons of diets). Significant weight loss has been observed with many low-carbohydrate or low-fat diets. Generally, weight loss differences between diet types are minimal. Studies suggest that

behavioral support and exercise may play key roles in weight loss differences versus the diet regimen itself. Similarly, in comparing several popular diet programs (Atkins, Weight Watchers, Ornish, and Zone diet), research indicates that adherence is the biggest predictor of weight loss, regardless of type of diet. Therefore, adherence appears to be the critical factor in positive outcomes. Unfortunately, due to gradual nonadherence, it is not uncommon for patients to have attempted numerous weight loss programs to be at their starting weight or higher.

Table 27.4–6.
Types of Diets

Type of Diet	Calorie Deficit	Weight Loss	Summary
Low-calorie diet (LCD)	800–1,500 cals/day	0.5–1 kg/wk	Carbohydrates 55% Protein 15% Fat <30%
Very low-calorie diet (VLCD)	500–800 cals/day	1.5–2.5 kg/wk	70–100 g of protein/day via liquid meal replacements
South Beach diet/Atkins/Zone diet	<30 g of carbohydrate/day	Rapid initial weight loss followed by steady weight loss	Low carbohydrate Some promote high fiber, lean meats, and low-glycemic carbohydrates
Weight Watchers	Promotes balanced food choices via a points system based on the amount of fats, carbohydrates, protein, and fiber in foods.	1–2 lb/wk	Mod fat, balanced nutrition Focus on activity Group/online support
Jenny Craig/Nutrisystem	Prepackaged low-calorie food	1–2 lb/wk	Meal replacements (1,200–1,500 calories/day) with gradual return to grocery foods in maintenance phase Individual consultant and/or online support

Behavioral Weight Loss

Although some individuals may experience success with commercial diets, many patients may need additional specialized support. In considering psychological treatment options for obesity, numerous studies demonstrate the efficacy of lifestyle and behavioral weight loss (BWL) interventions. In fact, based on a systematic review of the literature, the updated 2013 Guidelines for Managing Overweight and Obesity in Adults encourage clinicians to offer a rigorous, multicomponent behavioral intervention to patients with a BMI >30 or >25 with at least one comorbidity or risk factor. The basic tenets of BWL stem from Learning Theory, which assumes that behavior can be altered by modifying the context in which the behavior

occurs. Applying these principles to weight loss, treatment aims to reduce maladaptive behaviors, such as the consumption of calorically dense food and sedentary behaviors, while reinforcing healthy alternatives (e.g., the selection of healthy food choices and implementation of exercise). Through psychoeducation and skill-based interventions, such treatments encourage patients to make permanent changes to their food intake, activity level, and environment. In an effort to promote new, enduring habits, patients are encouraged to take a balanced approach to food choices and exercise. For example, primary dietary goals often included reducing calorie and saturated fat intake, while increasing the consumption of complex carbs, vegetables, and other nutritionally dense foods. Thus, BWL supports flexibility over rigid adherence.

In addition to general dietary and activity goals, BWL treatment focuses on learning behavioral strategies to facilitate weight loss. For example, self-monitoring and goal setting are incorporated into treatment to help patients observe their calorie targets and exercise objectives. It can be particularly important for clinicians to help patients set short-term, reasonable behavioral goals that can be monitored over time. In addition, BWL protocols typically focus on teaching stimulus control techniques; modifying cues in the environment to increase the potential that a healthy behavior will occur. For example, to decrease the likelihood of an overeating episode, individuals may limit their access to trigger foods by purchasing a cone of ice cream versus a gallon container. Similarly, to increase the probability of being active, individuals may keep their workout clothes in their car so that they go straight to the gym after work before returning home. Behavioral strategies, such as stimulus control, have consistently been shown to be an effective component for decreasing sedentary behavior in the context of a weight loss.

Commonly, BWL treatment is disseminated in a group format with 4 to 6 months of weekly sessions. Due to the skill-based nature of treatment, information is typically delivered in a psychoeducational format with topics presented as stand-alone modules. The LEARN (Lifestyle, Exercise, Attitudes, Relationships, and Nutrition) Program is considered the gold standard for BWL group-based treatment. The protocol has been utilized in numerous clinical and research settings. Similar to standard BWL programs, LEARN encourages patients to implement gradual lifestyle changes. Following the federal nutritional and activity guidelines, the program focuses on moderate caloric restriction and the gradual increase in physical activity. The program is structured in a series of additive steps with some repetition to facilitate mastery.

On average, BWL interventions produce a 5 to 10 percent weight reduction over a 6-month period. An evaluation of the literature conducted by the American College of Cardiology, the American Heart Association in Dallas, Texas, and the Obesity Society found that comprehensive

behavioral interventions involving a modification of diet, increased physical activity, and behavioral therapy skills, resulted in an average weight loss of 8 kg; with most patients' weight stabilizing after 6 months of treatment. Despite these typical outcomes, both clinical experience and research suggest that many overweight and obese individuals seeking weight loss begin treatment with goals that are two to three times greater than the expected weight loss range. While some research suggest no association or a positive relationship between initial weight loss goals and results, several studies have found that greater weight loss expectations are related to poorer weight loss outcomes. Thus, it may be important to discuss realistic expectations with patients prior to beginning BWL.

Weight Loss Maintenance

Although many patients can successfully lose weight during the initial phase of treatment, obesity treatments demonstrate limited effectiveness in regard to long-term weight loss maintenance. Unfortunately, a large body of literature suggests that following the completion of weight loss programs, most individuals regain a majority of their weight 3 to 5 years after treatment. In fact, it is estimated that less than 5 percent of patients who lose weight will have maintained their losses after 4 to 5 years. While some research suggests that metabolic changes may account for a small percentage of weight regain, the disappointing long-term results of standard behavior treatment are often attributed to decreased compliance with behavioral, dietary, and physical activity guidelines. In fact, it is posited that at the time of greatest weight loss, many individuals have already displayed marked deviations from their initial healthy behavior changes. Thus, the trend for individuals to regain weight underscores how difficult it can be to alter habits long term.

In understanding the difficulties with sustained weight loss, it can be helpful to understand how some individuals are able to achieve weight loss maintenance. Across studies, when comparing individuals who have maintained weight loss long term to individuals who have regained weight, maintainers typically report lower dietary fat consumption, higher levels of physical activity (specifically vigorous activity), increased use of general behavioral strategies, and more frequent self-weighing. Taken together, this body of literature indicates that sustained behavioral changes are associated with more long-term positive weight outcomes. Moreover, the National Weight Control Registry is a particularly useful database of individuals followed prospectively over time who have on average lost more than 65 lb and sustained their weight losses for more than 5.5 years. Using this registry, one study found that several factors predicted weight regain versus maintenance over a 1 year period. Risk factors for weight gain included a larger initial weight loss, heavier maximum weight, more significant history of weight cycling, shorter period of maintaining current

weight, greater depressive symptoms, more disinhibition, and a greater desire to continue to lose weight. These results indicate that even among individuals who have lost a significant amount of weight and kept the weight off for a period of time, there are still several risk factors that make some individuals more vulnerable to regaining weight.

Numerous studies have shown that dietary self-monitoring is associated with both greater initial weight loss and sustained maintenance of lost weight. Unfortunately, akin to other health behaviors, compliance with self-monitoring often decreases over time. Given this trend, one study sought to identify what aspects of self-monitoring are essential for long-term success. Researchers examined the effects of frequency (total number of food records), consistency (number of weeks with >2 records), and comprehensiveness of self-monitoring on weight loss maintenance in 220 women at 12 months into a weight loss journey. Results revealed that individuals who continued to self-monitor frequently and consistently at 12 months sustained weight losses more effectively. At the same time, the findings demonstrated that increased comprehensiveness of monitoring did not result in any added benefit above and beyond the effect of frequency. These findings indicate that continuing to complete food diaries frequently and consistently may help individuals to sustain a level of awareness and adherence, though the detail of self-monitoring appears to have less of an impact on outcomes.

Psychological Treatment of Obesity

Cognitive Behavioral Therapy for Weight Loss. In addition to general lifestyle and behavioral treatments, CBT is an effective, empirically supported psychological intervention for weight management in both youth and adults. CBT for obesity typically includes psychoeducation (especially around dietary guidelines and physical activity), behavioral skills (e.g., self-monitoring and stimulus control), and cognitive strategies. CBT can be utilized as a stand-alone treatment or can be used to improve behavioral lifestyle treatment components. Many of the tenets and skills taught in CBT treatment overlap with standard weight loss interventions, including the use of psychoeducation, self-monitoring, stimulus control, and problem solving. The literature suggests that CBT programs result in similar, if not slightly higher, weight losses than BWL; with most patients losing 5 to 20 percent of their initial weight. CBT may enhance traditional treatments by focusing on cognitive interventions targeting the maladaptive thinking that often triggers overeating, poor body image, and inactivity.

More specifically, cognitive restructuring is an important component of CBT that involves identifying and challenging negative thought patterns. For example, after eating a cookie, an individual may think “I blew it. I might as well give up and eat the whole package.” As a result of this thought, the person may binge on the box of cookies. In an attempt to

reduce such problematic behavioral patterns, patients are trained to identify cognitive distortions (irrational thoughts) and learn techniques to challenge beliefs that lead to unhealthy behaviors, such as identifying evidence for and against a thought. Clinicians often utilize Socratic questioning to help patients learn to challenge automatic thoughts and underlying beliefs. For instance, in the example above, clinicians may ask, "Why do you think it was difficult for you to accomplish your eating goals this week?" Similarly, clinicians may recommend that patients engage in behavioral experiments to challenge their negative assumptions. For example, patients can remain inactive as a result of thinking "I'm too tired after work to do anything." Such patients are asked to directly confront this thought by taking a walk despite fatigue. By the end of treatment, it is hoped that patients can elucidate and challenge dysfunctional thoughts and generate alternative thoughts in-the-moment.

Emotional distress often serves as a trigger for these negative thought patterns. Therefore, using CBT, clinicians may also focus on stress-reduction methods outside of food. In particular, clinicians may teach relaxation strategies, such as guided imagery, diaphragmatic breathing, and progressive muscle training, to help patients more effectively manage problematic emotional states. In high stress situations, relaxation techniques are intended to directly decrease emotional distress, while indirectly facilitating patients' ability to identify negative thoughts. Ultimately, it is anticipated that relaxation skills and other stress-management techniques can help patients reduce emotional distress and focus on actions consistent with weight loss goals.

Mindfulness-Based Approaches. Recent developments in mindfulness and acceptance-based interventions have led to increased attention on alternatives to more traditional BWL approaches. Mindfulness and acceptance-based interventions focus on changing patients' relationship with unpleasant thoughts, feelings, and bodily sensations. For example, acceptance and commitment therapy (ACT) targets experiential avoidance (the avoidance or attempt to control internal events, such as difficult thoughts, emotions, or bodily sensations), with a primary aim of helping patients increase their willingness to experience negative emotional experiences through mindfulness, cognitive defusion, and values techniques. Presumably, by becoming more willing to observe the present moment, patients are able to reduce maladaptive patterns used for avoidance (i.e., eating) and focus on value-driven behaviors (i.e., engaging in activities in the service of health and quality of life). Within this framework, ACT contends that experiential avoidance is the core of both emotional eating and eating in response to environmental stimuli. More specifically, emotional eating typically occurs in response to emotional distress and serves as an attempt to reduce internal negative experiences.

Although eating in response to external triggers does not necessarily stem from negative emotions, individuals often experience uncomfortable thoughts and feelings when trying to resist overeating in response to these cues. ACT seeks to promote healthy behaviors consistent with patients' values, while reducing avoidance through mindfulness and acceptance techniques.

The research examining mindfulness-based interventions for weight loss is in its infancy. Small studies have supported the use of ACT and general mindfulness techniques to address disordered eating patterns, physical activity, food cravings, and body image dissatisfaction. In addition, a small body of research demonstrates the positive effect of ACT on behavioral change in the context of weight loss and weight maintenance. One 2008 exploratory study examined the effectiveness of a brief ACT-based group treatment on 68 females randomized to ACT treatment or a control condition. Although no significant differences were initially found in BMI and mental health, intervention participants reported significantly increased physical activity levels at 6 months compared to controls. In addition, when participants who reported never using skills were excluded from the analysis, the ACT intervention participants showed significantly greater reductions in BMI. Findings suggested that decreases in BMI were mediated by declines in binge eating. This study, along with other research, provides initial evidence for the utility of ACT and mindfulness-based techniques in addressing obesity.

Motivational Interviewing. As the weight loss maintenance research suggests, not all patients seeking weight loss are successful. Some patients may struggle to maintain behavioral change long term, while others may never even begin to change before dropping out of treatment. The Transtheoretical Model (TTM) first proposed by Prochaska and DiClemente helps to explain why some patients experience difficulty initiating small changes, while others keep food diaries and make numerous behavioral modifications from the start. Although the TTM was initially developed to explain failure in smoking cessation interventions, it is now used to conceptualize a wide range of health behaviors, from condom use to exercise adoption and overeating. Regardless of the specific behavior, the TTM postulates that adopting a new behavior involves an evolution through six distinct stages of readiness to change (precontemplation, contemplation, preparation, action, maintenance and termination) (Table 27.4–7). Within this framework, behavior change is theorized to be a process that develops over time.

Prior to initiating lifestyle interventions for obesity, it is critical to assess and understand a presenting patient's readiness for change. Patients in the precontemplation stage of change may have limited appreciation that their obesity is a problem and only coming to treatment due to pressure

from others. It is not uncommon for patients in this stage to have little intention to change their lifestyle and therefore feeling threatened by treatment. Although patients in the contemplation stage are aware that a problem exists (i.e., "I need to lose weight in order to improve my diabetes"), they often remain stuck between a desire to change versus the perceived obstacles of change. Patients in this stage are sincerely contemplating change, but have yet to make a commitment to action. One study indicates that patients commonly remain in the contemplation stage for at least 2 years prior to proceeding to the next stage of change. In the subsequent stage (preparation stage), patients are typically intending to take measurable action within the month and may have even unsuccessfully attempted changes over the past year. In addition, a patient in the preparation stage may describe currently making small changes in preparation for broader change. For example, patients may report that they have obtained a gym membership in anticipation of starting to exercise in the near future. Arguably, many patients coming for a medical weight management evaluation are in this stage of change. Often patients may be seeking professional support in order to take overt steps toward a lifestyle change.

Table 27.4–7.
Stages of Change

Stage of Change	Description
Precontemplation	► No intention to change in foreseeable future ► May be unaware of problem ► Often present to therapy because of pressure by others ► Usually feel coerced or threatened by treatment
Contemplation	► Aware problem exists ► Seriously considering change but no commitment to action ► Struggle with positive evaluation of problem behavior and effort it might take to overcome it
Preparation	► Intending to take action in the next month ► May have unsuccessfully taken action in the past year ► May report small changes ("baby steps") ► Some reductions may be evident
Action	► Modifications in behavior, experience, and/or environment ► Overt behavior change ► One day to 6 months
Maintenance	► Prevent relapse and consolidate the gains ► Remaining free of the problem behavior and/or consistently engaging in incompatible behavior for >6 months ► Can be considered to last a lifetime for some behaviors
Termination	► Change process complete ► No longer working to prevent relapse ► Total confidence and self-efficacy across all high risk situations

► Exited from the cycle of change

Data from Prochaska JO, DiClemente CC, Norcross JC. In search of how people change. Applications to addictive behaviors. *Am Psychol*. 1992;47(9):1102–1114.

When patients proceed into the action stage, they will typically begin to alter their behavior, environment, and/or experience in an attempt to address their obesity. For instance, patients may change their home environments by throwing away trigger foods and replacing these items with healthier purchases. Within the TTM, individuals are considered to be in the action stage if they have changed a dysfunctional behavior (e.g., inactivity and high-calorie food choices) from as little as 1 day up to 6 months. In the addictions literature, behavior change is typically defined as the absence of a given behavior (i.e., remaining sober). In weight management, the notion of elimination is less applicable and instead behavioral modification or reshaping may be more relevant. Overall, changes during this period are often the most visible. Subsequently, individuals frequently receive the most external recognition during this stage. In particular, it is not uncommon for patients to receive compliments about their weight loss, which can often serve to reinforce lifestyle change. Following 6 months of consistent change, patients enter the maintenance stage where they work to consolidate gains and prevent relapse. Individuals are thought to be in the maintenance stage when they remain free of a behavior and/or implement a new, incompatible behavior for more than 6 months. In regard to lifestyle changes related to weight management, many consider the maintenance stage to be lifelong. However, technically, patients can enter the termination stage where they have exited the cycle of change. In order to achieve this stage, patients must no longer be working to prevent relapse and have complete confidence in all high-risk situations.

Most individuals engaging in behavior change proceed through the stages of change in a spiral versus linear pattern. Relapse is the rule rather than the exception. Therefore, many patients will not maintain changes on their first attempt. In fact, as they modify their behavior, it is quite common for patients to regress back to an earlier stage and consequently cycle back through the stages. For instance, patients engaging in BWL may have consistently engaged in healthy eating patterns for a few months. However, while at a holiday party they may have a binge eating episode and subsequently feel ambivalent about continuing to make positive behavior changes. At the beginning of treatment, it is essential to educate patients about the spiral nature of change and to normalize these common experiences.

Furthermore, given the varying levels of readiness for change, when presented with a patient seeking treatment of their obesity, it is important to consider stage-matched treatment. For instance, if patients are in the contemplation stage giving them action tasks will likely result in failure or

drop-out. Unfortunately, most treatments, including weight management programs, are action oriented and assume that patients are in the action stage. Treatment is the most effective when clinicians match treatment goals to patients' stages of change. In fact, treatments focused on helping patients move just one stage a month can double the chances that patients will engage in change on their own in the future.

Movement through the stages of change has been shown to be influenced by mediating variables, including the decisional balance and self-efficacy. More specifically, decisional balance is conceptualized as the balance between the perceived comparative gains and losses (pros and cons) of making a change versus the status quo. Typically, patients in the earlier stages (precontemplation and contemplation) identify more disadvantages than advantages to behavior change. The decisional balance tends to shift as patients move through the stages of change, with patients often perceiving more advantages than disadvantages to change in the preparation, action, and maintenance stages. For example, in the contemplation stage patients likely perceive many barriers to engaging in consistent activity, while those in the preparation stage may begin to recognize that increased activity results in increased energy and overall improved well-being. Similarly, self-efficacy (individuals' confidence in their ability to change) appears to be a key variable in movement through the stages. Studies suggest that self-efficacy typically increases linearly across the stages of change.

Motivational interviewing (MI) is a treatment approach based on the stages of change model. MI is intended to enhance patients' motivation for change and thus support patients' movement through the stages of change. MI is considered a patient-centered counseling style with a heavy emphasis on collaboration and partnership between patients and clinicians. The treatment postulates that each patient holds an intrinsic motivation for change. In line with a collaborative approach, clinicians seek to enhance patients' motivation for change through eliciting personal goals, values, and perceptions. Within this framework, the primary goal of MI is to help patients explore and resolve ambivalence around behavior change (see Table 27.4–8 for details regarding MI principles and techniques).

Table 27.4–8.
Strategies Using Motivational Interviewing Based on Stage of Change

Stage of Change	Potential Clinical Strategies
General techniques	*Express empathy *Increase self-efficacy *Roll with resistance *Develop discrepancy
Contemplation stage	*Identify pros and cons of change vs. status quo *Normalize ambivalence

	*Explore barriers to change *Increase commitment *Support self-efficacy
Preparation stage	*Clarify goals *Explore options *Reinforce sense of personal choice *Practice skills *Develop a plan
Action stage	*Support self-efficacy *Engage in problem-solving *Reinforce commitment to change and success thus far

Adapted from Erol S, Erdogan S. Application of a stage based motivational interviewing approach to adolescent smoking cessation: the transtheoretical model-based study. *Patient Educ Couns*. 2008;72(1):42.

Given the high rates of weight regain, MI has increasingly been examined and utilized as an alternative and/or addition to more traditional skill-based weight management techniques. Two large-scale weight loss studies (the Diabetes Prevention Program and the Look AHEAD trial) utilized MI as a part of their lifestyle intervention. In addition, a 2011 meta-analysis of randomized controlled trials using MI found that MI techniques resulted in a moderate effect size for reducing BMI above and beyond control interventions. Across studies using weight as a primary outcome, MI significantly enhanced weight loss outcomes, with participants who received MI losing 1.47 kg more on average than participants receiving control treatments. Of note, in studies tracking change in BMI, MI produced improved but nonsignificant reductions in BMI over controls. Taken together, this body of literature underscores the potential utility of MI in addressing patients' motivation for change when engaging in weight loss treatments.

Case Example

The case of "Robin" exemplified some of the obstacles often seen in BWL treatment. Robin was a 24-year-old Caucasian female self-referred for weight and occasional binge eating episodes. At a BMI of 40, Robin identified several ways obesity negatively impacted her life. In particular, she discussed how she often avoided social gatherings as a result of poor body image and social anxiety related to her weight. Robin further acknowledged that she often used food to cope with stress and anxiety. At the initial session, Robin verbally endorsed a strong desire to lose weight and change her behaviors. However, across treatment sessions, Robin did not complete food monitoring forms and experienced difficulty meeting her exercise goals. Her clinician initially tried to address these difficulties by helping her identify and problem solve barriers to change. Although barriers may have played a role in Robin's noncompliance, it is likely that she was struggling with ambivalence around change and may have had low self-efficacy regarding her ability to engage in consistent healthy behaviors. Instead of continuing to focus on action, the clinician decided it was important to explore Robin's ambivalence further. In particular, the clinician highlighted to Robin the discrepancies between her words and her actions. In addition, they spent time identifying the pros and cons to change. Once ready, Robin, along with her clinician, clarified goals and discussed several options moving forward. Through using MI techniques and increasing her intrinsic motivation for change,

Robin was able to move forward into the action stage of change. Fortunately, the clinician realized by using MI that continuing to focus on action-oriented tasks, as most treatments promote, was likely to result in drop-out.

Psychological Treatment of Eating Disorders. Several psychological treatments are supported in treating eating disorders typically associated with obesity. The research examining inventions for BED highlights the various treatment options currently available for several disordered eating patterns. Given a strong evidence base, CBT is currently considered a gold standard treatment of BED. The treatment is typically disseminated in 12 to 19 weekly group sessions. Across studies, CBT has demonstrated marked improvements in binge eating episodes, symptoms of psychopathology, psychosocial functioning, and modest reductions in weight. In fact, research indicates that patients may experience a clinically significant response by the fourth session. Similar outcomes have also been shown in research examining interpersonal psychotherapy (IPT), though results typically are not observed as quickly. While CBT focuses on dysfunctional behaviors and cognitions related to eating and body image, IPT aims to treat the interpersonal difficulties related to disordered eating patterns. IPT identifies four potential focus areas, including grief, interpersonal role conflicts, role transitions, and interpersonal deficits. In a collaborative, supportive environment, clinicians often utilize clarification, exploration, emotional expression, behavioral modification, and communication analysis to help patients address their interpersonal difficulties. Long-term follow-up in randomized control trials have further supported the sustained treatment effects of both CBT and IPT. For example, one study compared the efficacy of CBT and IPT on BED and associated symptoms 4 years after the cessation of treatment. Across treatments, 64.4 percent of patients continued to demonstrate full remission from binge eating, supporting the lasting efficacy of both interventions. While the IPT group showed improvements in binge eating symptoms over time, the CBT group demonstrated worsening of symptoms, though treatment outcomes were not significantly different at any time point. The authors propose that these differences in recovery trajectories likely reflect a leveling-off between-treatment effects, given the rapid benefits typically observed with CBT. Further, for both treatments, patients continued to report significant improvements in associated eating disorder symptoms and overall psychopathology at 4 years, though a few variables, including binge episode days, restraint and anxiety, trended toward relapse. Overall, these results support the continued efficacy of CBT and IPT several years following treatment of BED.

In addition to CBT and IPT, clinicians have utilized dialectical behavior therapy for BED (DBT-BED). DBT-BED focuses on treating emotion dysregulation associated with disordered eating patterns. In a group

setting, DBT typically covers mindfulness, distress tolerance, and emotional regulation skills. Patients are further encouraged to use mindful eating and disordered eating-specific diary cards. Increasingly, research has supported the efficacy of DBT for binge eating. One study examined potential variables moderating the effect of DBT-BED treatment outcomes. Interestingly, the researchers found that higher baseline psychopathology was associated with better treatment results. Findings suggested that avoidant personality disorder and the early onset of overweight and dieting (<15 years old) moderated the effect of treatment on posttreatment binge eating episodes. It is possible that patients who have a longstanding history of using food as a avoidant coping mechanism may specifically benefit from a treatment focused on adaptive emotional regulation. However, additional research is needed to elucidate these findings. In addition, CBT and IPT hold the strongest support for the psychological treatment of bulimia nervosa. Similar to outcomes with BED, patients appears to have a quicker resolution of binging and purging behaviors with CBT, though outcomes between CBT and IPT are typically similar around 1 year.

Less is known regarding the behavioral treatment of NES. However, there is strong emerging evidence for the use of CBT for NES. Treatment typically focuses on behavioral techniques (e.g., closing the kitchen at a specific time and delaying getting out of bed), increased awareness of thoughts and behaviors associated with night eating (i.e., "If I don't eat I won't fall back to sleep"), cognitive restructuring, and sleep hygiene behaviors. A pilot study of a 10-session CBT intervention for NES demonstrated significant decreases in evening caloric intake, nocturnal ingestions, number of awakenings per week, and weight. Improvements were also seen in mood and quality of life. Although initial research is promising, continued investigation is needed in this specialized area of eating disorder treatment.

Medical Treatments

Medications for weight loss should be considered in patients with a BMI >30 or a BMI >27 with serious comorbidities (Table 27.4–2). Behavioral and lifestyle interventions including a diet and exercise plan are essential adjunctive therapies when using medications for weight loss. Medication treatment of obesity has a varied history including concerns about addiction and serious side effects. Amphetamines were used for weight loss in the early 20th century but led to substance abuse. In the late 20th century, combination treatment of di-fenfluramine and phentermine ("Fen-Phen") became popular until valvular heart disease side effects were discovered. Over-the-counter (OTC) treatments such as ephedrine with caffeine were linked with reports of death and pulled from the market. Research on the use of leptin to treat weight problems has been disappointing, and mood changes associated with potential weight loss

drugs such as ecopipam and rimonabant have prevented their approval for use. Sibutramine (Meridia) was approved for several years, but withdrawn after increased blood pressure and adverse cardiovascular events were discovered.

Current medications approved for weight loss include orlistat (Xenical, Alli), lorcaserin (Belviq), phentermine/topiramate (Qysmia), naltrexone/bupropion (Contrave), phentermine (Adipex) and liraglutide (Saxenda) (Table 27.4–9). Orlistat acts as a lipase inhibitor, thereby reducing caloric absorption and is available in prescription formulation (120 mg three times daily before meals) and an OTC formulation (60 mg). Orlistat typically results in an 11 percent weight loss, although oily stools and GI distress are common side effects. Orlistat may also cause reduction in absorption of fat-soluble vitamins and in rare cases liver disease. Lorcaserin selectively targets serotonin 2c receptors; weight loss at 1 year is approximately 7 kg. Lorcaserin appears to be well tolerated with less cardiovascular risks than previous medications, but should not be used with SSRIs or MAOIs because of the risk of serotonin syndrome.

Phentermine/topiramate is approved for the long-term use in doses of 3.75 to 15 mg phentermine and 23 to 96 mg extended release topiramate and appears to work through reduction in appetite. Weight loss with the medication appears to be 9 to 10 percent. Side effects may include paraesthesias, dizziness, altered taste, insomnia, constipation, and dry mouth. This medication, as with all weight loss medications, is contraindicated in pregnancy. Of note, other potential risks include suicide, mood, and sleep disorders.

Table 27.4–9.
Medications Approved for Weight Loss Treatment

Medication	Trade Name	Weight Loss	Approved Use	Side Effects/Contraindications
Orlistat	Xenical, Alli (OTC)	9–11% body weight	Long-term	GI distress, oily stool, vitamin malabsorption
Lorcaserin	Belviq	3–10% body weight	Long-term	Cannot be used with SSRIs or MAOIs
Phentermine–topiramate ER	Qsymia	9–11% body weight	Long-term	Paraesthesias, dizziness, altered taste, insomnia, constipation, and dry mouth. Suicidal thoughts, mood, and sleep disorders.
Naltrexone–bupropion	Contrave	5–10% body weight	Long-term	Nausea and potential interactions with SSRIs and MAOIs
Phentermine	Adipex, many others	4–6% body weight	Short-term	Dry mouth, insomnia, anxiety, and increased heart rate and blood pressure. Not recommended for patients with a history of cardiovascular disease

Liraglutide	Saxenda	5–10% body weight	Long-term	Thyroid T-cell tumors, increased heart rate, renal disease, hypoglycemia, pancreatitis, suicidal ideation

Naltrexone/bupropion was recently brought to market and appears to reduce food intake by acting on adrenergic and dopaminergic receptors in the hypothalamus. Significant weight loss and reduction in comorbidities were achieved on the medication. Side effects include nausea and potential interactions with SSRIs and MAOIs.

Phentermine is approved for short-term use in treating patients with obesity. Because of the potential for abuse, its use is typically limited to 12 weeks. Phentermine is the most often prescribed medication for weight loss in the United States, but there is little data to determine its long-term efficacy. Effects of the medication appears to be dose related, and side effects can include dry mouth, insomnia, anxiety, and increased heart rate and blood pressure. Phentermine is not recommended for patients with a history of cardiovascular disease.

Liraglutide is the newest addition to the drugs approved for weight loss. Its action on GLP-1 may regulate areas of brain involved in appetite. Weight loss results are promising, but side effects/risks include thyroid t-cell tumors, acute pancreatitis, increased heart rate and suicidal ideation.

Further evaluation of antidiabetic drugs for weight loss indications is an area of future research in the field. There are concerns about the extended use of weight loss medications, as they typically do not produce permanent weight loss (i.e., when the drug is stopped, the weight is regained.) In addition, the history of approved weight loss medications with later emerging serious side effects is important to consider for future research.

Surgical Treatments

Weight loss surgery is currently the most effective long-term treatment of severe obesity (Table 27.4–2). Patients with severe obesity (BMI >40) or moderate obesity (BMI >35) with associated obesity related disease are candidates for weight loss surgery. Candidates should be well informed, motivated, psychologically stable, have good social supports, have no substance abuse, and be an acceptable operative risk. Patients applying for weight loss surgery should have documented nonsurgical weight loss attempts and be willing to make long-term lifestyle changes.

The current surgical treatments include both restrictive and malabsorptive procedures (Fig. 27.4–3). Laparoscopic adjustable gastric banding (LAGB) is a procedure that involves an adjustable device that creates a small gastric pouch and narrow outlet thus restricting the amount of intake and inducing satiety. LAGB is reversible and has low early postoperative risks, but the weight loss tends to be lower than with other procedures. In addition, late postoperative risks such as erosion and

slippage are common. Laparoscopic sleeve gastrectomy (LSG) involves the removal of 70 percent of the stomach, restricting intake and appears to result in greater weight loss than the LAGB. Malabsorptive procedures, including the biliopancreatic diversion and duodenal switch, bypass a segment of the small intestine so less food is absorbed. These procedures are rarely performed because of the high risk of nutritional deficiencies. The Roux-en-Y gastric bypass (RYGB) is a combination procedure using both restriction and malabsorption to promote weight loss. The most common procedure performed currently, RYGB results in a 60 to 80 percent loss of excess body weight.

The majority of bariatric surgeries are performed laparoscopically which reduces risk of infection, incisional hernia, postoperative pain, and wound complications. Benefits of weight loss surgery include weight loss of 40 to 80 percent excess body weight and resolution of comorbidities such as metabolic syndrome, hypertension, sleep apnea, liver steatosis, hypercholesterolemia, and gastroesophageal reflux disease. Other benefits include increased life expectancy and improved quality of life. Potential risks of weight loss surgery include bleeding, anastomic leak, wound infection, thromboembolism, anastomic strictures, marginal ulcers, bowel obstruction, cholelithiasis, and nutritional deficiencies. The risk of perioperative death for RYGB is approximately 0.5 percent; with risks increasing for older patients, male patients, and lower surgeon volume.

Psychological Considerations for Bariatric Surgery . Psychological risks after weight loss surgery include increased risk of alcohol abuse, increased suicide rates, and possible problems with absorption of psychiatric medication. Researchers have established an increased risk of the development of alcohol use disorders in patients following RYGB and LSG that may be related to increased absorption. There appears to be a threefold increase in suicide risk for patients following weight loss surgery, with patients in the first 3 years after surgery and men being at a higher risk of suicide. Another possible concern for mental health providers is the apparent decrease in absorption of SSRIs after Roux-en-Y. Given the high number of bariatric patients on antidepressants, recommendations include close psychiatric monitoring after surgery and movement to shorter acting as opposed to extended release formulations.

Other considerations are the stresses of adjusting to a new lifestyle, changes in family relationships due to role changes from weight loss, body image issues and coping with excess skin. Research also suggests concern regarding eating disorder behaviors returning after surgery, such as loss of control eating, bulimia, and anorexia. Adjunctive counseling and psychotherapy can be helpful in monitoring psychological symptoms after surgery and helping patients adjust to a new lifestyle.

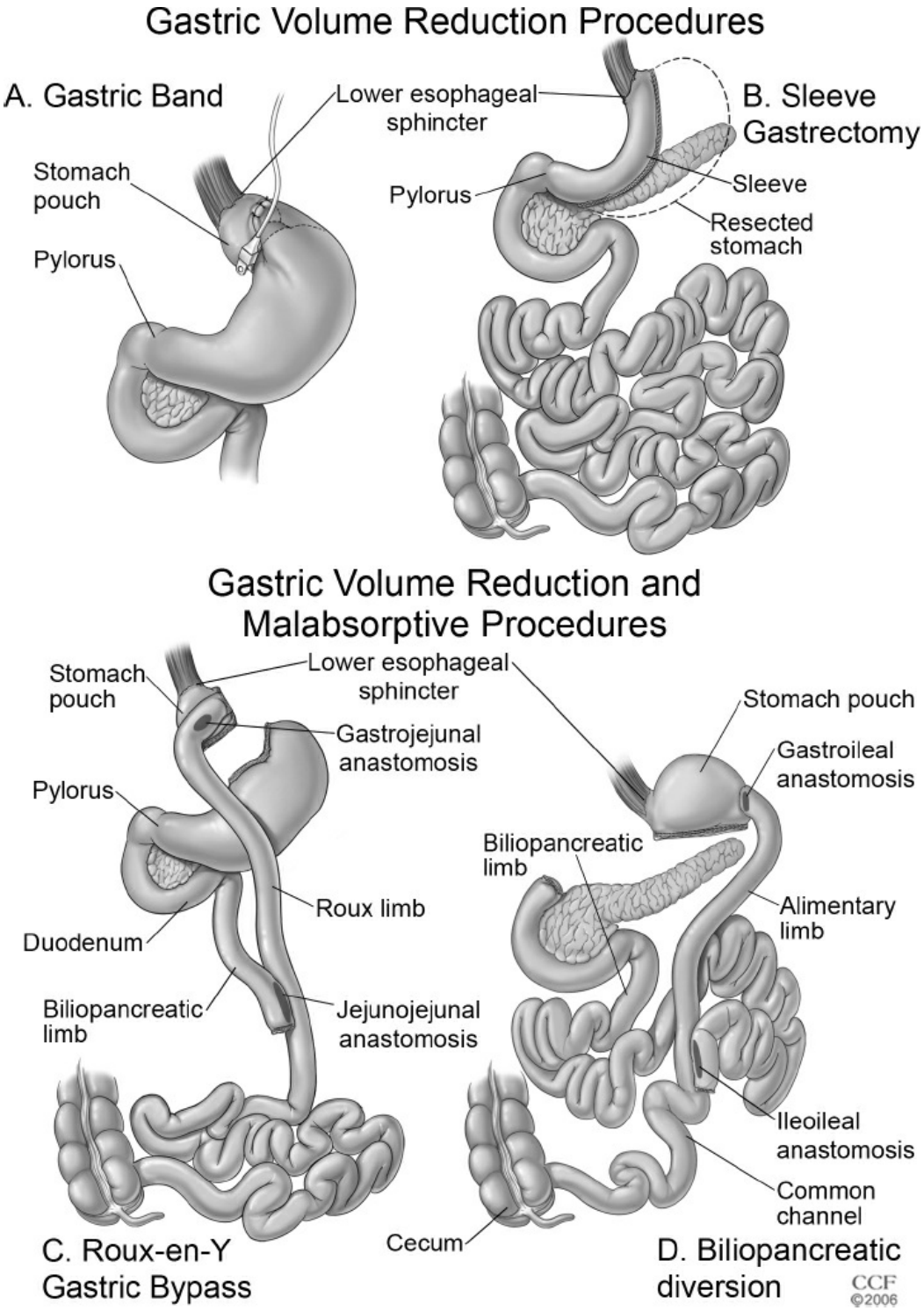

FIGURE 27.4–3. Restrictive, malabsorptive, and combination weight loss surgery procedures. (Reprinted with permission, Cleveland Clinic Center for Medical Art & Photography © 2006–2015. All Rights Reserved.)

Psychological treatment of the bariatric candidate ideally involves a behavioral health professional such as a psychologist or psychiatrist who is fully integrated into the bariatric team. The rates of psychiatric disorders, such as depression, bipolar depression, BED, and anxiety disorders, are high among bariatric surgery patients. For example, one study of

presurgical bariatric patients found a 68 percent lifetime incidence of Axis I disorders, and 38 percent of patients currently met criteria for an Axis I diagnosis. Presurgical evaluation of the bariatric patient is standard and typically includes psychological testing and a psychosocial interview. Elements of the evaluation should include assessment of the patient's motivation, knowledge of surgery and lifestyle changes needed, weight loss and other expectations, mental health, substance use, social supports, stressors, coping strategies, and eating pattern. Psychological contraindications for surgery include unstable depression or psychosis, suicidal ideation, active bulimia, and active substance abuse.

Behavioral health specialists may also play a role in optimizing patients for surgery by engaging them in mental health treatment to address instability and treating eating disorder behaviors presurgically. The psychological team may work with inpatient bariatric patients showing anxiety or depressed mood in the perioperative period, and follow-up after surgery may help to monitor mood and behavior change as the patients adjust to their new lifestyle. The behavioral health specialist often plays a leading role in providing support groups as part of the bariatric program.

FUTURE DIRECTIONS

There is a well-established relationship between obesity and psychiatric disorders, but more needs to be understood about the common causes of obesity and psychological disorders and how best to provide treatment to patients with both obesity and mental health disorders. The field would benefit from additional clinical trials that utilize combination treatments such as medications with behavioral weight management/CBT to help mitigate weight gain from psychotropic medications and optimally treat mental health disorders with the lowest side effect profile for patients with obesity. The role of the behavioral health specialist in the treatment of obesity is likely to grow as health care moves toward a more integrated approach to patient treatment and understanding mind-body health. There are benefits to adjunctive psychological treatment at all stages of obesity treatment including prevention, diagnosis, behavioral, medical, and surgical treatment. Future research should establish data to support the integrated treatment of obesity in the psychiatric field.

REFERENCES

Allison KC, Lundren JD, Moore RH, Stunkard AJ. Cognitive behavior therapy for night eating syndrome: a pilot study. *Am J Psychother*. 2010;64(1):91.

*Amstrong MJ, Mottershead TA, Ronksley PE, Sigal RJ, Camobell TS, Hemmelgarn BR. Motivational interviewing to improve weight loss in overweight and/or obese patients: a systematic review and meta-analysis of randomized controlled trials. *Obes Manag*. 2011;12:709.

Anastasiou CA, Karfopoulou E, Yannakoulia M. Weight regaining: From statistics and behaviors to physiology and metabolism. *Metabolism*. 2015;64(11):1395–1407. doi: 10.1016/j.metabol.2015.08.006. Review.

*Bray GA. Medical treatment of obesity: the past, the present and the future. *Best Pract Res Clin Gastroenterol*. 2014;28:665.

Centers for Disease Control. *Overweight and Obesity*; 2014. control.:http://www.cdc.gov/obesity/index.html

Erol S, Erdogan S. Application of a stage based motivational interviewing approach to adolescent smoking cessation: the transtheoretical model-based study. *Patient Educ Couns*. 2008;72(1):42.

Gariepy G, Nikita D, Schmitz N. The association between obesity and anxiety disorders in the population: a systematic review and meta-analysis. *Int J Obesity*. 2010;34:407.

Gerlach G, Herpertz S, Loeber S. Personality traits and obesity: a systematic review. *Obes Rev*. 2015;32(16):32.

Heinberg LJ, Ashton K, Windover A. Moving beyond dichotomous psychological evaluation: the Cleveland Clinic Behavioral Rating System for weight loss surgery. *Surg Obes Relat Dis*. 2010;6(2):185–190.

Hilbert A, Bishop ME, Stein RI, et al. Long term efficacy of psychological treatments for binge eating disorder. *Br J Psychiatry*. 2012;200:232.

King WC, Chen JY, Mitchell JE, et al. Prevalence of alcohol use disorders before and after bariatric surgery. *JAMA*. 2012;307:2516.

Lasikiewicz N, Myrissa K, Hoyland A, Lawton CL. Psychological benefits of weight loss following behavioral and/or dietary weight loss interventions. A systematic research review. *Appetite*. 2014;72:123.

Lee WJ, Almulaifi A. Recent advances in bariatric/metabolic surgery: appraisal of clinical evidence. *J Biomed Res*. 2015;29(2):98–104.

Lillis J, Kendra KE. Acceptance and commitment therapy for weight control: model, evidence, and future directions. *J Contextual Behav Sci*. 2014;3(1):1.

MacLean PS, Wing RR, Davidson T, et al. NIH working group report: innovative research to improve maintenance of weight loss. *Obesity*. 2015;23(1):7.

Marek RJ, Lavery M, Heinberg L, Rish J, Ashton K. A review of psychological assessment instruments for use in bariatric surgery evaluations. *Psychol Assessment*. 2016;28(9):1142–1157.

*McElroy S. Obesity in patients with severe mental illness: overview and management. *J Clin Psychiatry*. 2009;70:12.

Michael JC, Belendiuk KA, Kuchera AM, Rofey DL. Cognitive behavioral therapy for weight management. *Behav Therapist*. 2013;36(2):28.

*Nihalani N, Schwartz TL, Siddiqui UA, Megna JL. Obesity and psychotropics. *CNS Neurosci Ther*. 2012;18:57.

Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011–2012. *JAMA*. 2014;311:806.

Peterhänsel C, Petroff D, Klinitzke G, Kersting A, Wagner B. Risk of completed suicide after bariatric surgery: a systematic review. *Obes Rev*. 2013;14(5):369–382. doi: 10.1111/obr.12014.

Review.

Peterson ND, Middleton KR, Nackers LM, Medina KE, Milsom VA, Perri MG. Dietary self-monitoring and long-term success with weight management. *Obesity (Silver Spring)*. 2014;22(9):1962.

Preiss K, Brennan L, Clark D. A systematic review of variables associated with the relationship between obesity and depression. *Obes Rev*. 2013;14:906.

Prochaska JO, DiClemente CC, Norcross JC. In search of how people change. *Am Psychol*. 1992;47(9):1102.

Pursey KM, Stanwell P, Gearhardt AN, Collins CE, Burrows TL. The prevalence of food addiction as assessed by the Yale Food Addiction Scale: a systematic review. *Nutrients*. 2014;6. Available at www.mdpi.com/journal/nutrients.

Robinson AH, Safer DL. Moderators of dialectical behavior therapy for binge eating disorder: results from a randomized controlled trial. *Int J Eat Disord*. 2012;45(4):597.

Tapper K, Shaw C, Ilsley J, Hill AJ, Bond FW, Moore L. Exploratory randomized controlled trial of a mindfulness-based weight loss intervention for women. *Appetite*. 2009;52:396.

VanBuskirk KA, Potenza MN. The treatment of obesity and its co-occurrence with substance use disorders. *J Addict Med*. 2010;4:1.

Yen YC, Huang CK, Tai CM. Psychiatric aspects of bariatric surgery. *Curr Opin Psychiatry*. 2014;27(5):374–379.

Yska JP, van der Linde S, Tapper VV, et al. Influence of bariatric surgery on the use and pharmacokinetics of some major drug classes. *Obes Surg*. 2013;23:819.

▲ 27.5 Respiratory Disorders

THOMAS RITZ, PH.D., CHELSEY A. WERCHAN, M.A., AND E. SHERWOOD BROWN, M.D., PH.D.

DEFINITION AND COMPARATIVE NOSOLOGY

Psychiatric symptoms and syndromes in respiratory disorders are less extensively investigated than those in other common medical illnesses. However, numerous clinical features of these illnesses may increase the risk of developing a psychiatric disorder. Respiratory disorders are frequently associated with dyspnea, insomnia, and hypoxia that can increase the risk for a variety of psychiatric illnesses. Furthermore, severe respiratory disorders are associated with substantial functional impairment that can increase the risk of depression. Many of the medications given for lung diseases can have significant central nervous system effects.

The role of the psychiatrist in respiratory disease includes (1) diagnosis and treatment of comorbid psychopathology; (2) addressing individual psychosocial triggers of exacerbation; (3) assessment and treatment of emotional, neurocognitive, and psychosocial consequences of chronic respiratory disease and its management; and (4) supporting the management of the disease (adherence, motivation for rehabilitation, devise management strategies, improvement in comanagement by families and partners).

ASTHMA

Asthma is a very common respiratory illness. Some of the disease-related factors that may increase vulnerability to psychiatric disorders in patients with asthma also apply to other respiratory disorders. The composite case vignette below illustrates the complex relationship between asthma and the brain.

A 34-year-old woman presents with a history of episodes of severe and prolonged coughing since the first years of life. In childhood, symptoms typically started following upper respiratory infections and resulted in persistent coughs that made it difficult for the patient to sleep through the night. Exercise also left her wheezing and coughing at the end of the day, particularly during summer. Asthma was diagnosed formally around 6 years of age, after a particularly severe exacerbation that required an emergency room visit. The patient was prescribed a low dose of inhaled corticosteroids as maintenance medication and a β-adrenergic bronchodilator for as-needed use. During her teens, she became highly self-conscious about her asthma symptoms and her medication adherence greatly decreased, resulting in poorer symptom control and increased bronchodilator use. During an emergency room visit at age 18, she received a burst of high-dose oral corticosteroids for a week. Her energy greatly increased and she developed severe insomnia. She spent one sleepless night cleaning her room. She was also very irritable and had a verbal altercation with a classmate at school. These symptoms resolved within a week of stopping the corticosteroids, but similar symptoms emerged with subsequent corticosteroid courses prescribed for asthma exacerbations.

During college, the patient was provided with a brief asthma management training, which

included a plan to avoid overexposure to outdoor air pollutants during critical days in summer and improvement in indoor environment by mold and house dust mite eradication. Because she also reported anxiety and depression, a psychiatric consultation was obtained, which revealed major depressive disorder, generalized anxiety disorder, and occasional panic attacks. Feelings of helplessness were accompanied by hyperventilation, which led to strong symptoms of shortness of breath, chest tightness, and dizziness. Symptoms sometimes spiraled out of control as the patient feared suffering from an asthma attack. Bronchodilators were only partly effective in relieving these symptoms. The treating psychiatrist initiated pharmacological treatment, with an SSRI and cognitive behavioral training that also included slow and shallow breathing exercises to reduce hyperventilation.

Historical Perspective

Observations of labored breathing and asthma have a long history and can be traced back to ancient texts from Egypt and Greece. Already at that time, a connection between the brain and lungs was suspected. Hippocrates viewed health as a balance of healthy quantities of humors and an excess of humors was linked to illness. If one of these, phlegm, was produced excessively, it would condense in the brain and travel down to the lower organs, among these the lungs. Hence, the term *catarrh*, translated from Greece as “flowing” or “flowing down.” The psychoanalytic tradition viewed asthma as a vegetative manifestation of passive–regressive impulses that stem from unresolved conflicts related to maternal attachment, famously culminating in the interpretation of asthma as “suppressed cry for the mother.” With progress in elucidating molecular and cellular aspects of asthma pathophysiology, the psychological study of asthma became more heavily oriented toward illness management behavior in the 1970s and 1980s. Only recently, large-scale epidemiological studies and psychophysiological laboratory experimentation have again produced more convincing empirical evidence for influences of psychosocial factors and psychopathology, on both etiology and pathophysiology of the disease.

Disease Manifestation and Pathophysiology

Asthma is a chronic, inflammatory airway disease that leads to episodic symptom exacerbations, which extoll a substantial burden on the patients’ well-being and can be life threatening if not controlled adequately. Pathophysiological changes include obstruction, inflammation, hyper-responsiveness, and remodeling of the airways. Patients experience unpleasant breathing sensations (dyspnea), which is characterized by shortness of breath, chest tightness, coughing, and/or wheezing. Additionally, mood and energy levels are affected in symptomatic episodes, commonly leading to fatigue, irritability, anxiety, fear, and/or panic. The airways are hyper-responsive in asthma, with stronger than normal narrowing of the bronchi in response to a range of triggers, including allergens, physical exercise, air pollutants, odors, smells and irritants, cold air, respiratory tract infections, certain drugs (e.g., aspirin), gastric reflux,

or psychological stress and emotions.

There is no single diagnostic indicator for asthma. Evidence comes from multiple sources, including disease history and manifestations (including symptom characteristics and precipitating factors); physical examination focused on upper respiratory tract, chest, and skin; spirometry; demonstration of reversibility of airway obstruction by bronchodilator application; airway hyper-reactivity testing with methacholine, histamine, cold air, or exercise; and skin or blood tests for allergic sensitization.

Prevalence and Societal Burden

Approximately 300 million people worldwide suffer from asthma, including 25.7 million in the United States. Asthma is one of the most common chronic diseases, and it has continued to increase in prevalence over the past several decades. Individuals of all ages can be affected by asthma, although a clear diagnosis is difficult in the first 5 years of life. In children, asthma is the most common chronic disease, affecting 7.1 million children under 18 years in the United States. Children from minority populations and socioeconomically disadvantaged backgrounds are disproportionally affected by asthma, in particular in African-American and Puerto Rican communities. With close to 4,000 deaths from asthma per year in the United States, asthma-related morbidity and mortality highlight the need for better control and prevention methods. The annual economic costs of asthma in the United States have recently been estimated by the Center for Disease Control, with approximately $56 billion in direct health-care costs and indirect costs, including lost productivity.

Pharmacological Treatment and Self-Management

The goal of asthma pharmacological management is to achieve optimal control of asthma with the right type and dose of agents. Anti-inflammatory medication, in particular inhaled corticosteroids, is the first line of treatment and is required as daily maintenance medication. Short-acting bronchodilators (typically β-adrenergic) are used as rescue medication on an as-need basis and provide relief from symptoms for approximately 6 hours. Long-acting β-adrenergic bronchodilators help prevent bronchoconstriction for approximately 12 hours, particularly throughout the night. Other anti-inflammatory agents are sometimes used as maintenance medication, such as leukotriene modifiers, which target specific inflammatory mediators and also have bronchodilatory properties. Additional immunoglobulin-E inhibitors are sometimes used in severe allergic asthma. Severe cases or disease exacerbations typically require high doses of systemic corticosteroids.

Consistent efforts are required by patients with asthma to self-manage their condition, which includes adherence to medication schedules, action

plan-guided self-adjustments of medication in exacerbations, self-monitoring symptoms of lung function, and control of critical environmental and endogenous asthma triggers. Typically, 40 to 60 percent of patients show suboptimal medication adherence, which contributes to low asthma control and frequent exacerbations. Comprehensive educational programs in self-management improve asthma outcomes and are generally recommended by guidelines. Designing and implementing interventions based on theories of behavior change, learning theory, or motivational interviewing techniques can enhance their effects. Involving families in self-management of children and young adolescents is essential, and additional training in family problem-solving skills or parent coaching can benefit asthma-related quality of life.

Stress, Emotion, and Asthma

A long history of clinical and epidemiological literature links asthma exacerbations to patients' reports of a variety of emotional precipitants, typically various negative affective states, but also ambivalence or conflicts over emotional expression, or even positive states of excitement and mirth. Beyond patients' recollections, which can be subject to multiple biases, laboratory studies have clearly established that bronchoconstriction can be elicited by stress and emotion-induction techniques. Individuals with asthma vary in the extent to which they are susceptible to such triggers, with an estimated 20 to 40 percent of patients responding with clinically relevant airway obstruction. Major physiological pathways for obstruction due to psychological stimuli are parasympathetic excitation, hyperpnea (increased depth of breathing), and hypocapnia. Observational studies have also expanded the current knowledge, showing that psychosocial stressors, negative life events, and adversity, such as low socioeconomic status, are linked to lung function decline, airway inflammation, and changes in immune status favoring asthma exacerbations. Consequently, emotions and stress are now listed among the well-established triggers of asthma symptoms, as reflected in current asthma management guidelines. Elevated anxiety and problems communicating emotions, a facet of alexithymia, have been found in patients who had experienced a near-fatal asthma exacerbation.

Evidence is also strong for an association between anxiety, depression, and patients' perception and report of asthma symptoms. Given that asthma symptoms are typically episodic, the patient's ability to detect and communicate symptoms effectively is key to successful management of the disease. Anxious patients tend to overestimate or over-report airway obstruction. Consequently, they use more rescue inhalers and access more professional care than necessary, including emergency room treatments. Patients who develop an intense fear of asthma exacerbations may be treated with more oral corticosteroids than the less fearful patients. Recent

neuroimaging studies have demonstrated that brain structures and pathways typically involved in emotional processing and psychopathology, in particular limbic structures, including right anterior insular, are also involved in the perception of dyspnea, thus improving the mechanistic understanding of emotional influences on airway symptoms. Furthermore, many patients with asthma have deficits in recognizing airway obstructions, and thus underestimate the severity of their disease state. As a consequence, they show insufficient adherence to maintenance medication and delay necessary treatments, and thereby risking life-threatening exacerbations.

Comorbidity with Psychiatric Disorders

Epidemiological studies have shown that asthma is associated with an elevated risk of psychiatric conditions, including anxiety disorders (agoraphobia/panic, generalized anxiety disorders, post-traumatic stress disorder, and social phobia), depressive disorders, alcohol use disorders, schizophrenia, and suicide. Symptoms patients with asthma experience during exacerbations can include anxiety, panic, irritability, frustration, and fatigue, closely mimicking some of the psychiatric symptomatologies. The unsettling experience of the respiratory disease may lead to secondary psychiatric problems, or conversely, psychopathology may facilitate neuroendocrine and/or inflammatory/immunological pathways involved in asthma pathogenesis. Thus, asthma symptoms experienced in childhood are associated with greater levels of anxiety and shyness, and they predict the future development of anxiety disorders. Subsequent onset of panic disorder has been interpreted as evidence for a suffocation-alarm theory of panic, which views panic as an exaggerated response to CO_2-increases and posits that changes in chemosensitivity to rising CO_2 levels due to respiratory disease are a critical factor in the development of panic. Beyond direct pathways, third variables, such as shared genetic risk factors or specific childhood adversities leading to disease manifestations at different time points, also need to be considered. In support of such assumptions, elevated levels of a range of psychiatric conditions are typically found in first-degree relatives of adolescents with asthma, although these could also be secondary to the burden of asthma care in the family.

Existing psychiatric disorders and asthma can influence each other in a number of ways. Overlap of symptoms can cause challenges in the diagnosis of asthma. Prominent symptoms of a psychiatric condition, such as schizophrenia, can distract from comorbid asthma and lead to its underdiagnosis and undertreatment. Other complications can affect the course and treatment of asthma. Strong emotional states associated with psychopathology can affect asthma pathophysiology directly through pathways discussed earlier. In the case of psychiatric disorders, these states

occur repeatedly over a prolonged time period and thereby constitute stressors that affect organ function chronically. Thus, hyperventilation associated with panic disorder can worsen lung function and potentially exacerbate inflammation. Anxious or depressed mood can also affect disease management, for example, by over- or underuse of rescue inhalers, lack of adherence to prescribed maintenance medication, neglect of trigger control, problematic health behaviors affecting the airways (e.g., smoking), or delayed treatment seeking. Earlier research has highlighted problems in patients who are prone to intense states of "panic-fear," most likely indicating clinically significant anxiety disorders. Such patients were found to have excessive rates of hospitalization for asthma exacerbations. Asthma-related quality of life is typically reduced with comorbid psychopathology. Conversely, asthma can also amplify the symptomatic appearance of psychiatric conditions and complicate their treatment. Symptoms of asthma can be a source of panic or contribute to the fatigue and exhaustion experienced in mood disorders. Internalizing disorders in children seeking cognitive behavior treatment are more severe in those with comorbid asthma.

Asthma may also be associated with an increased risk of cognitive disorders. Large studies suggest that asthma is associated with an increased risk of mild cognitive impairment, and that an asthma diagnosis increases the risk of later development of dementia. The reasons for this association between asthma and cognitive disorders are not known, but could be related to reduction in respiratory function, inflammatory processes, or medications.

Treatment of Comorbid Psychopathology in Asthma. The psychiatric treatment of comorbid asthma patients requires specific considerations and precautions. However, despite findings of high comorbidity rates, the development of specifically adapted treatments for comorbid asthma patients has been lagging. Medication treatments for both conditions can have mutual adverse side effects. Psychotropic medication for anxiety can affect respiration adversely and thus may complicate asthma symptomatology. High doses of benzodiazepines or hypnotics can lead to respiratory depression and are related to a greater risk of death from asthma. Some data supports the use of SSRIs for depression. There is some emerging evidence that cognitive behavior therapy for anxiety disorders, and perhaps also depression, adapted for asthma is beneficial in reducing anxiety and bronchodilator use and improving quality of life. Treatments typically combined generic anxiety-related treatment elements, such as cognitive restructuring, interoceptive exposure to feared bodily sensations, progressive muscle relaxation, and psychoeducation, with asthma relevant elements such as asthma education and breathing retraining, and eliminated treatment aspects potentially detrimental to patients with

asthma, such as voluntary hyperventilation as interoceptive exposure exercise. Although there is a clear need for targeting asthma-related depressive deficits, including low self-esteem, social isolation, and reduced medication adherence and management behaviors, empirical support from published psychological intervention studies is still largely missing.

Adjunctive Cognitive and Behavioral Treatment Modalities for Asthma

Adjunctive cognitive and behavioral treatment modalities have been developed or adapted for asthma, which are not specifically directed at treating comorbid psychiatric conditions but address psychiatrically relevant deficits in patients with asthma, such as maladaptive breathing patterns, inaccurate perception of airway obstruction, or suffering from excessive stress. They can significantly add to pharmacotherapy. Certain habitual breathing patterns that support hypocapnia or hyperpnea are targeted by dedicated training modules of slow, abdominal, shallow, more regular and/or nasal breathing. Although breathing is under voluntary control, characteristics of the person's breathing pattern are typically quite stable and systematic, and instructed efforts are needed to change it permanently. Training of slow, abdominal breathing has been shown to reduce symptoms and improve patients' quality of life. Mechanisms underlying such improvements are largely unexplored but may include a better gas exchange efficiency or reduction in high flow rates that can irritate and dry out the hyper-reactive airways. However, exaggerated deep breaths in slow breathing can also lower PCO_2 levels, thus making hypoventilation training a safer approach. Successful interventions train prolonged respiratory pauses or slower and shallower breathing with capnometry-assisted biofeedback. More sustained symptomatic improvements have been noted. Feedback can also be provided to increase respiratory sinus arrhythmia by slow breathing, which has been associated with some of the same benefits. Although various forms of relaxation training, meditation, or yoga have considerable face validity for addressing the distress associated with a chronic condition, findings on benefits for patients with asthma have remained quite variable and high-quality large-scale clinical trials are lacking. This is not surprising given that the rationale for such techniques is unclear. Relaxation aims to reduce sympathetic activity and increase parasympathetic activity, both of which support bronchoconstriction.

Interoception training can address a potentially dangerous lack of awareness of airway obstruction in patients with asthma. It is often administered in protocols of systematic exposure to added resistive loads that vary in the extent to which they impede airflow. Patients learn to better distinguish different loads or to lower their perceptual thresholds for just noticeable loads.

CHRONIC OBSTRUCTIVE PULMONARY DISEASE

Chronic obstructive pulmonary disease (COPD) is a group of progressive and irreversible lung diseases (e.g., emphysema and chronic bronchitis) that lead to progressive reduction in airflow in the lungs. Common symptoms include cough and *dyspnea,* the unpleasant sensation of breathlessness that is continuous, often frightening, and is associated with considerable limitations on daily activities. COPD is not only strongly associated with tobacco smoking but also environmental exposures, in particular fumes from biomass fuel stoves in developing countries. The prevalence of COPD in adults age 40 or older is approximately 10 percent. It is currently the fifth leading cause of death worldwide with an increasing trend.

Pharmacological treatment of COPD uses bronchodilator medication, antibiotics, or inhaled corticosteroids. Oxygen therapy, lung volume reduction surgery, or lung transplantation may be necessary in severe cases. Multidisciplinary rehabilitation is not only focused on physical exercise but also includes education and self-management skills, nutritional components, and psychosocial support. It is associated with improved exercise tolerance, reduced exacerbations, and improved dyspnea and quality of life.

The prevalence of depression in COPD ranges from about 25 to 50 percent. Greater disease severity, dyspnea, female sex, and disability appear to be risk factors for depression in COPD. When depression is present in COPD, it is associated with decreased quality of life and poorer adherence to pulmonary rehabilitation programs. The presence of depression may be more strongly associated with functional impairment in COPD than lung disease severity.

Although the results are somewhat mixed, depression treatment appears to improve depressive symptoms in people with COPD but have modest effects on disease-related outcomes. Pulmonary rehabilitation, particularly if it includes stress reduction or psychotherapy, appears to be associated with improvement in depression in people with COPD.

Anxiety disorders, particularly panic disorder, are much more common in patients with COPD than in the general population. As with depression, anxiety appears to be associated with the poor functional status in people with COPD. The treatment of anxiety in COPD has been the topic of few studies. A potential concern with the use of benzodiazepines in people with COPD is decreased respiratory drive. Some data suggest that buspirone (BuSpar) or SSRIs might be useful alternatives to benzodiazepines in this population.

Cognitive impairment (e.g., working memory and attention) is also common in people with COPD. The cognitive deficits appear to worsen with pulmonary disease progression. Some data suggest improvement in cognition with pulmonary rehabilitation.

In psychiatric assessments of people with COPD, it is important to consider that the majority have a significant smoking history. Tobacco use is associated with increased rates of a variety of psychiatric disorders. The increased risk of both mood and anxiety disorders in people with COPD appears in large part to be explained by nicotine dependence. Furthermore, rates of tobacco use are much higher in people with major psychiatric disorders, particularly schizophrenia, than in the general population. Thus, people with psychiatric illnesses are at an increased risk of developing COPD. Therefore, the etiology of the increased rates of psychiatric disorders in COPD may be particularly complex and may reflect effects of hypoxia, functional impairment, feelings of breathlessness, as well as pre-existing vulnerability related to tobacco use.

INTERSTITIAL LUNG DISEASE

Interstitial lung disease includes a diverse group of illnesses characterized by progressive scarring and fibrosis of the lungs. Progressive worsening of dyspnea is a characteristic feature. These illnesses can be idiopathic, familial, or result from a variety of environmental exposures or certain medications. The prognosis is poor. No treatments are presently available that substantially alter the course of interstitial lung diseases.

Little is known about the nature or prevalence of psychiatric disorders in interstitial lung disease. High rates of depression and anxiety that are related to the severity of dyspnea are reported in patients with idiopathic pulmonary fibrosis.

CYSTIC FIBROSIS

Physiology and Manifestations

Cystic fibrosis (CF) is an autosomal recessive disease that is characterized by lung and digestive symptoms and affects approximately 70,000 individuals worldwide. CF is the most common genetic disease among Caucasians and it is caused by a mutation on chromosome 7 that leads to malfunction of the cystic fibrosis transmembrane regulator (CFTR) protein, which facilitates movement and concentration of chloride ions across the cell membranes of the sweat glands, sinuses, lungs, pancreas, gastrointestinal tract, and reproductive organs.

Individuals with CF experience a wide range of symptoms, including recurrent cough and lung infections, wheezing or shortness of breath, salty skin, sinus infections, pancreatic insufficiency, poor nutrition and growth, and difficulty with digestion and bowel movements. There can be significant variations in individual presentations of the disease. Over time, recurrent respiratory infections and bacterial colonization result in gradual loss of lung functioning, and respiratory failure is the most common cause of mortality. In adulthood, individuals are at increased risk of the

development of CF-related diabetes due to deterioration of the pancreas and osteopenia or osteoporosis as a result of poor nutrition and bone density.

Diagnosis and Treatment

Diagnosis of CF is commonly made through the results of a sweat test that examines the salt concentration of an individual's sweat. Due to the malfunction of the CFTR protein, individuals with CF have abnormally high concentrations of chloride in their sweat. Additionally, genetic testing can be used to make a definitive diagnosis and to identify the individual's specific mutations.

Currently, there is no cure for CF. Treatment is aimed at slowing the progression of the disease and preventing infection. Since discovery of the CFTR protein in the late 1980s, medical treatment of CF has progressed rapidly and life expectancy for those with the disease has likewise increased dramatically over the years, with almost 50 percent of the CF population currently over the age of 18 years.

Treatment of CF involves an extremely complex and time consuming daily regimen that may include airway clearance, aerosolized medication, nutritional support, dietary supplements, and other medications, such as antibiotics and anti-inflammatories. Individuals in North America are often followed by a CF center, accredited by the Cystic Fibrosis Foundation, which involves care from a team of specialists, including pulmonologists, nurse practitioners, respiratory therapists, dieticians, physical therapists, and social workers. Individuals with CF are routinely monitored every 3 to 4 months through their clinics in order to adjust treatment approaches as necessary.

Self-Management and Psychosocial Considerations

As with many chronic illnesses, self-management in CF is a vital component of medical maintenance of the disease as well as the patient's overall quality of life. Effective self-management involves many skills, including proper diet and nutrition, exercise, adherence to airway clearance and other respiratory treatments, monitoring of symptoms of exacerbations, and regular coordination of care with the CF care team. Research suggests that adherence to different aspects of self-management in CF varies with the lowest rates of adherence often found with chest physiotherapy recommendations.

Psychosocial considerations for health-care professionals treating patients with CF will include a variety of issues, depending on the patient's overall level of functioning, medical status, and phase of life. Coordination of care during childhood will necessitate consideration of family dynamics and the mental health of adult caregivers. Overall family functioning has

been found to significantly impact quality of life and health status in patients with CF.

Emerging Adulthood/Transition

The growing adult population has necessitated the development of adequate protocols for transition of patients to adult CF care. This transitional phase has yet to be thoroughly addressed by existing research and available recommendations are uncommon and often vague. Available commentaries suggest that the stage of transition may be difficult for many patients who have had prolonged care with a single multidisciplinary team for a number of years. Continuous and effective communication between the medical providers and the patient is recommended during the stage of transition in order to ensure proper continuation of care and well-being of the patient.

Fertility Issues. Fertility issues and family planning may become matter of concern as patients with CF enter adulthood. A majority of males with CF lack a vas deferens, causing infertility, and a significant portion of females may also experience fertility problems. Additionally, genetic counseling may be warranted as patients consider whether their offspring will be carriers of the disease or potentially have the diagnosis itself.

Transplantation and End of Life Issues. As the severity of the disease progresses, lung transplantation may be considered as a treatment option and will greatly impact the overall self-management and functionality of the patient. With greater decline in health status, patients also need to make decisions regarding end of life care.

Psychiatric Comorbidities

According to the International Depression Epidemiological Study on the prevalence of anxiety and depression in patients with CF and their caregivers, rates of depression vary by age and type of screening tool being used (Hospital Anxiety and Depression Scale [HADS] vs. Center for Epidemiologic Studies Depression Scale [CES-D]). On average, 5 to 19 percent of adolescents and 13 to 29 percent of adults with CF reported symptoms of depression, with higher rates of depression found using the CES-D. Overall, prevalence rates for symptoms of anxiety range from 14 to 35 percent across different countries, with an average of 30 percent of patients reporting symptoms. Additionally, a high degree of comorbidity occurs between anxiety and depression, with those reporting anxiety symptoms being from 13.64 times (adults) to 14.97 times (adolescents) more likely to also report symptoms of depression than those without elevations on anxiety.

Elevated rates of depression and anxiety are also reported in caregivers of patients with CF, with prevalence rates ranging from 48 (mothers) to 36 percent (fathers) for anxiety and from 37 (mothers) to 31 percent (fathers) for depression. When caregivers are experiencing elevated symptoms of depression or anxiety, adolescents with CF are more likely to report symptoms.

In patients with CF, higher rates of depression and anxiety are negatively associated with health status and may have a number of negative outcomes, such as poor treatment adherence, increased health-care costs, lower lung functioning, and decreased quality of life. Therefore, careful and regular monitoring of psychological functioning is recommended in order to provide appropriate intervention or mental health referrals as needed.

Prescription of any psychiatric medication would require careful consideration, particularly with regard to side effects, dosages, and potential interactions with other medications that patients are prescribed.

Psychological Interventions

Psychological interventions for patients with CF may target a number of different outcomes, including education, emotional functioning, quality of life, treatment adherence, airway clearance techniques, and dietary intake. Interventions have primarily been educational or behavioral in nature and target specific treatment concerns rather than broader psychosocial functioning, with limited significant effects. Interventional research in this area has lagged behind in comparison to the medical advances in the disease. Increased attention is needed on the development of effective interventions that improve treatment adherence, quality of life, and independent self-management in this population.

RESPIRATORY FAILURE

Over one million persons in the United States are admitted to intensive care units annually due to respiratory failure. Respiratory failure can be a consequence of a variety of lung diseases. Both hypoxia and hypercapnia can occur in respiratory failure, and in severe cases, mechanical ventilation is needed. Therefore, the psychiatric syndromes observed with respiratory failure include those of severe lung disease as well as depression and post-traumatic stress disorder or other anxiety disorders related to mechanical ventilator support and the intensive care unit. Some data suggest that the administration of corticosteroids may decrease the risk of developing post-traumatic stress disorder during an intensive care unit admission.

SOMATOFORM RESPIRATORY DISORDERS

A variety of dysfunctional and distressing respiratory regulation abnormalities have been identified in adults and children under various

terms in past decades. Typically, diagnosis has relied heavily on exclusion of somatic diseases and a lack of demonstrable pathophysiological processes, which can produce most of the same symptoms. The adequacy of these disorders as diagnostic entities has often been debated. Efforts have been made more recently to distill common characteristics of these conditions and formulate positive diagnostic criteria under the umbrella of *somatoform respiratory disorders*. These would distinguish subgroups of functional disorders of respiratory regulation (e.g., vocal cord dysfunction), stereotypic respiratory disorders (e.g., habit cough), and psychogenic respiratory disorders (e.g., suffocation anxiety). However, the label "psychogenic" has become increasingly obsolete as contemporary neuroscience demonstrates tight associations between mental, neurophysiological, and peripheral autonomic and somatic processes in health and disease. Common to all conditions is that a psychiatric background and specific behavioral interventions and/or psychotherapy are often suggested, but the available evidence is mostly limited to case reports and uncontrolled studies. Although the patients' quality of life is often reduced, the prognosis is good to excellent with adequate treatment. Reassurance that the symptoms are not indicative of a somatic disease and that the symptoms are treatable already provides significant relief for patients and their families.

Chronic Hyperventilation

Hyperventilation is defined as breathing in excess of metabolic demand, leading to a fall in arterial PCO_2 levels. The respiratory pattern is often deep and/or fast, leading to exaggerated minute ventilation relative to demands from physical activity. As a consequence, patients complain about symptoms of dizziness, headaches, chest pain, peripheral paresthesia (lips and limbs), and anxiety. Despite overbreathing, patients have paradoxical sensations of air hunger and shortness of breath. Levels of $PCO_2 < 30$ mm Hg have been proposed as an indicator of unequivocal chronic hypocapnia, but many of patients' symptoms are less tightly linked to PCO_2 levels than previously thought, which is not surprising given the complexity of regulation of respiratory sensations. Whereas acute respiratory hyperventilation leads to disturbances in acid–base balance with high pH levels (acute respiratory alkalosis), pH levels in chronic hyperventilation are often near-normal due to compensatory bicarbonate changes.

Triggers of hyperventilation are often emotional, and hyperventilation is most frequently observed in panic disorder. However, chronic hyperventilation is also found without evidence of anxiety disorders (idiopathic hyperventilation). Breathing training focused on reducing hypocapnia (hypoventilation training) is the mainstay of therapy. Slow, shallow, abdominal, and nasal breathing is ideally supported by

measurement and feedback of actual PCO_2 values. Rebreathing techniques using bags are discouraged because they reduce oxygenation and may trigger panic by raising CO_2 levels. Patients profit from identifying individual triggers, decatastrophizing and reattribution of benign symptoms, and education about hyperventilation and anxiety.

Vocal Cord Dysfunction

Patients with vocal cord dysfunction suffer from sudden, intermittent adduction (narrowing or closure) of the vocal cords during inspiration. The inspiratory occurrence is most common and paradoxical, but it can also be observed during expiration or both phases of the breathing cycle. Proposed positive diagnostic criteria are presentation with sudden onset of dyspnea, problems with inspiration, intermittent symptoms that resolve spontaneously after a few minutes, and symptoms self-localized in the throat. Diagnostic confirmation in the acute stage can be obtained by flexible laryngoscopy (closure of vocal cords >50 percent) or abnormal inspiratory and/or expiratory features of the flow-volume loops in spirometry.

Episodes are often associated with considerable fear of suffocation and panic, and anticipatory anxiety of an additional attack is common. A variety of triggers can be identified, including stressful life events, anxiety, exercise, irritants, gastroesophageal reflux, respiratory infections, postnasal drip, or olfactory stimuli. A psychiatric background can be found in up to 75 percent of patients and anxiety disorders, a history of early life traumatic experiences, or abuse are not uncommon. The condition can co-occur with asthma or mimic it, complicating differential diagnosis and treatment and leading to extensive and costly measures. Unresponsiveness to bronchodilator therapy and lack of airway hyper-reactivity can help exclude asthma. The recommended treatment is ideally multidisciplinary. Underlying somatic conditions such as reflux, infections, or nasal allergies require adequate medication treatment. Among behavioral treatments, speech therapy has been particularly promising.

Chronic Cough Disorder (Habit Cough and Psychogenic Cough)

Cough is an integrated evoked, reflexive, or voluntary motor pattern that has important functions in airway protection and clearance. It is elicited by irritation of the vocal cords or airways and is a nonspecific symptom of a range of diseases and conditions, including respiratory infections, asthma, COPD, gastroesophageal reflux, allergic rhinitis, and postnasal drip. Chronic cough that is persistent, without evidence for any of these causes, has often been labeled as habit cough or psychogenic cough. It is most frequently reported in children and adolescents. The disruptive nature of the phenomenon can have negative social consequences. Reduction or

cessation during the night or with distraction has also been frequently observed, although use of these clinical presentations as defining criteria has been debated. A recent expert panel advocated replacing habit cough with "tic cough," aligning it with tic disorders as it shares their central features of suppressibility, distractibility, suggestibility, and variability, as well as the presence of a sensation of urge (premonitory sensation), regardless of whether the cough is the only tic or one of many. The panel also favored the term *somatic cough disorder* to replace psychogenic cough, placing it within the DSM-5 framework of somatization disorder, in which a psychological distress is expressed as a physical symptom. The diagnosis should only be given if a tic disorder and other causes can be excluded, and the patient meets criteria for DSM-5 somatization symptom disorder. Hypnosis and suggestion therapy are the most successful treatment modalities for somatic cough disorder and tic cough, in particular in children.

Sighing Dyspnea

A sigh breath is characterized by a sudden increase in depth, and often also length, of the breathing cycle. It occurs intermittently and is usually limited to a single breath. There is no clear convention as to the magnitude of volume change necessary, with some sources requiring a full inspiration to inspiratory capacity and others requiring 2 to 2.5 times the regular tidal volume. The clinical manifestations of sighing dyspnea include a more frequent rate of sighing, which is often episodic and accompanied by sensations of breathlessness or chest pain, and sometimes also yawning. Most observations have been in children, adolescents, or young adults. Proposed positive criteria are self-limited sighing with intensified inspiration, normal respiration rate, perceived respiratory effort greater than sensation of respiratory distress, and no nighttime awakening from sighing. Although associations with anxiety have been reported, this has not been confirmed empirically. Elevated sigh breathing has also been observed in adult anxiety patients, particularly in panic, and is thought to maintain hypocapnia. Breathing training focused on regularity of breathing (without suppressing occasional sighs) and reducing hyperventilation appears as the most relevant primary treatment. For psychotherapy and anxiolytic drug therapy, a psychiatric background should be clearly established in the individual case.

PSYCHIATRIC SYMPTOMS ASSOCIATED WITH RESPIRATORY DISEASE MEDICATIONS

As illustrated by the representative case vignette above, medications used to treat respiratory illnesses can have effects on the brain and the patients' experience. Oral corticosteroids such as prednisone are frequently

prescribed for patients with lung diseases. The use of these medications ranges from brief "bursts" for asthma exacerbations or other acute and reversible pulmonary changes to chronic therapy that can last for years for several lung diseases. Long-term use of corticosteroids is associated with a variety of systemic side effects, including suppression of the hypothalamic–pituitary–adrenal axis, glaucoma, cataracts, weight gain, infections, and osteoporosis. These side effects are sometimes termed iatrogenic Cushing syndrome because of the frequent development of truncal obesity, moon faces, and buffalo hump. Inhaled corticosteroids are also frequently used for some lung diseases. While these corticosteroid preparations are designed to be locally acting, when used at high doses for extended periods, they can be associated with many of the side effects observed with oral corticosteroids.

Shortly after the widespread use of corticosteroids began in the early 1950s, reports of psychiatric symptoms from their use began to appear in the literature. A Hollywood movie, "Bigger than Life," was even released that provided a fictionalized account of the potential for corticosteroids to induce psychiatric side effects and even foster abuse.

The use of corticosteroids is associated with an increased risk of mood symptoms that include depression, mania, and mixed features. Acute, higher dose corticosteroid therapy, as with bursts for asthma exacerbations, may be more strongly associated with mania or hypomania, while depression may be more common with chronic, lower dose therapy. Psychotic features, such as delusions and hallucinations, can also occur with corticosteroids.

Perhaps even more common than mood symptoms are changes in cognition. In animal models, the hippocampus appears to be particularly sensitive to the effects of corticosteroids. Consistent with effects on the hippocampus, memory appears to be a cognitive domain that shows changes with corticosteroid therapy. In healthy controls, even a single high-dose exposure to corticosteroids is associated with a decline in declarative and working memory performance. Structural brain changes, including a reduction in hippocampal volume, appear to accompany memory changes, even during very brief corticosteroid administration. More generalized cognition changes, as with delirium, also occur during corticosteroid therapy, especially in older persons.

The most clearly established risk factor for psychiatric side effects with corticosteroids is dose. Thus, the lowest effective dose of these medications should be used when treating lung diseases or other conditions. If the dose cannot be reduced or if this approach is ineffective, the treatment, like the symptoms, is similar to that of bipolar disorder. Although, placebo-controlled trials have not been conducted, mood stabilizers such as lithium, anticonvulsants, and antipsychotics seem to ameliorate mood changes and psychosis with corticosteroids.

Other medications used for lung diseases may also have psychiatric side effects. Anxiety, agitation, and depression have been reported with theophylline and β-2 adrenergic agonists. Theophylline may also be associated with an increased risk of suicidal ideation. Post marketing reports suggest a possible link between leukotriene inhibitors and a variety of psychiatric symptoms, including agitation, insomnia, anxiety, depression, and suicidal ideation. However, the data does not appear to support an association of these medications with completed suicide.

FUTURE DIRECTIONS

The psychiatric symptoms associated with respiratory illnesses have been less extensively studied than those in other common medical conditions, such as coronary artery disease. The underlying mechanisms are not well understood but are likely multifactorial. Furthermore, little is known about the treatment of depression, anxiety, and other disorders in this population. The increasing prevalence of respiratory illnesses, in particular of COPD and asthma, and current population trends in aging pose considerable challenges to scientists and clinicians. Interdisciplinary collaboration will be needed to further elucidate the interaction of psychosocial factors with respiratory illness and to develop and implement adequate pharmacological, cognitive, and behavioral interventions.

REFERENCES

Abbott J, Elborn JS, Georgiopoulos AM, et al. Cystic Fibrosis Foundation and European Cystic Fibrosis Society survey of cystic fibrosis mental health care delivery. *J Cyst Fibros*. 2015;14:533.

Brown ES, Jeon-Slaughter H, Lu H, et al. Hippocampal volume in healthy controls given 3-day stress doses of hydrocortisone. *Neuropsychopharmacology*. 2015;40:1216.

Brown ES, Vigil L, Khan DA, Liggin JD, Carmody TJ, Rush AJ. A randomized trial of citalopram versus placebo in outpatients with asthma and major depressive disorder: a proof of concept study. *Biol Psychiatry*. 2005;58:865.

Caldera-Alvarado G, Khan DA, Defina LF, Pieper A, Brown ES. Relationship between asthma and cognition: the Cooper Center Longitudinal Study. *Allergy*. 2013;68:545.

Capel LH. A history of the pulmonary catarrhs and airways obstruction. *Br J Dis Chest*. 1966;60:113.

Chen MH, Li CT, Tsai CF, et al. Risk of dementia among patients with asthma: a nationwide longitudinal study. *J Am Med Dir Assoc*. 2014;15:763.

Gada E, Khan DA, DeFina LF, Brown ES. The relationship between asthma and self-reported anxiety in a predominantly healthy adult population. *Ann Allergy Asthma Immunol*. 2014;112:329.

Glasscoe CA, Quittner AL. Psychological interventions for people with cystic fibrosis and their families. *Cochrane Database Syst Rev*. 2008(3):CD003148.

Goodwin RD, Lavoie KL, Lemeshow AR, Jenkins E, Brown ES, Fedoronko DA. Depression,

anxiety, and COPD: the unexamined role of nicotine dependence. *Nicotine Tob Res*. 2012;14:176.

Goodwin RD, Scheckner B, Pena L, Feldman JM, Taha F, Lipsitz JD. A 10-year prospective study of respiratory disease and depression and anxiety in adulthood. *Ann Allergy Asthma Immunol*. 2014;113:565.

Grüber C, Lehmann C, Weiss C, Niggemann B. Somatoform respiratory disorders in children and adolescents-proposals for a practical approach to definition and classification. *Pediatr Pulmonol*. 2012;47:199.

*Holland AE, Fiore JF Jr, Bell EC, et al. Dyspnoea and comorbidity contribute to anxiety and depression in interstitial lung disease. *Respirology*. 2014;19:1215.

*Jain A, Lolak S. Psychiatric aspects of chronic lung disease. *Curr Psychiatry Rep*. 2009;11:219.

Janssens T, Verleden G, De Peuter S, Van Diest I, Van den Bergh O. Inaccurate perception of asthma symptoms: a cognitive-affective framework and implications for asthma treatment. *Clin Psychol Rev*. 2009;29:317.

*Judd LL, Schettler PJ, Brown ES, et al. Adverse consequences of glucocorticoid medication: psychological, cognitive, and behavioral effects. *Am J Psychiatry*. 2014;171:1045.

Kaptein AA, Klok T, Moss-Morris R, Brand PL. Illness perceptions: impact on self-management and control in asthma. *Curr Opinion Allergy Clin Immunol*. 2010;10:194.

Kaugars AS, Klinnert MD, Bender BG. Family influences on pediatric asthma. *J Pediatric Psychol*. 2004;29:475.

Lehrer P, Feldman J, Giardino N, Song HS, Schmaling K. Psychological aspects of asthma. *J Consult Clin Psychol*. 2002;70:691.

Meuret AE, Ritz T. Hypocapnia in anxiety and respiratory disease: empirical evidence and clinical strategies. *Int J Psychophysiol*. 2010;78:68.

*Pfeffer PE, Pfeffer JM Hodson ME. The psychosocial and psychiatric side of cystic fibrosis in adolescents and adults. *J Cyst Fibros*. 2003;2:61.

Pumar MI, Gray CR, Walsh JR, Yang IA, Rolls TA, Ward DL. Anxiety and depression—important psychological comorbidities of COPD. *J Thora Dis*. 2014;6:1615.

Quittner AL, Barker DH, Snell C, Grimley ME, Marciel K, Cruz I. Prevalence and impact of depression in cystic fibrosis. *Curr Opin Pulm Med*. 2008;14:582.

Quittner AL, Goldbeck L, Abbott J, et al. Prevalence of depression and anxiety in patients with cystic fibrosis and parent caregivers: results of The International Depression Epidemiological Study across nine countries. *Thorax*. 2014;69:1090.

Ritz T. Airway responsiveness to psychological processes in asthma and health. *Front Physiol*. 2012;3:343.

Ritz T, Meuret AE, Trueba AF, Fritzsche A, von Leupoldt A. Psychosocial factors and behavioral medicine interventions in asthma. *J Consult Clin Psychol*. 2013;81:231.

Trojan TD, Khan DA, Defina LF, Akpotaire O, Goodwin RD, Brown ES. Asthma and depression: the Cooper Center Longitudinal Study. *Ann Allergy Asthma Immunol*. 2014;112:432.

Trueba AF, Ritz T. Stress, asthma, and respiratory infections: pathways involving airway immunology and microbial endocrinology. *Brain Behav Immun*. 2013;29:11.

*Verma S, Cardenas-Garcia J, Mohapatra PR, Talwar A. Depression in pulmonary arterial hypertension and interstitial lung diseases. *N Am J Med Sci*. 2014;6:240.

Vertigan AE, Murad MH, Pringsheim T, et al. Somatic Cough Syndrome (Previously Referred to as Psychogenic Cough) and Tic Cough (Previously Referred to as Habit Cough) in Adults and Children: Chest Guideline and Expert Panel Report. *Chest*. 2015;148(1):24–31.

von Leupoldt A, Fritzsche A, Trueba AF, Meuret AE, Ritz T. Behavioral medicine approaches to chronic obstructive pulmonary disease. *Ann Behav Med*. 2012;44:52.

▲ 27.6 Diabetes: Psychosocial Issues and Psychiatric Disorders

WAYNE KATON, M.D., AND PAUL Ciechanowski, M.D.

Diabetes is one of the most complex chronic medical conditions, requiring a high degree of patient self-management within a supportive social network to ensure optimal outcomes. Psychiatric illness can adversely affect patients' self-care behaviors and their abilities to interact with others. Psychiatric illnesses, such as depression, may also lead to neurohormonal changes such as changes in the hypothalamic–pituitary axis, which may adversely affect glucose control.

DIABETES: THE NATURE OF THE ILLNESS

Diabetes results from insufficient insulin secretion or resistance to insulin. Persons with type 1 diabetes produce no or very limited insulin and the onset is typically early in life. Type 1 diabetes accounts for 5 to 10 percent of all patients with diabetes. Type 1 diabetes is due to an autoimmune, predominantly T-cell process that selectively destroys insulin-producing pancreatic β cells. This autoimmune process likely has a multifactorial origin with both environmental and genetic factors. The onset of type 2 diabetes is more common among adults, although childhood onset is increasing. Individuals with type 2 diabetes usually have insulin resistance, but as they age, they have decreasing production of insulin, with many eventually becoming insulin dependent. Genetic vulnerability coupled with obesity are the most common causes of this condition, and dietary restriction and exercise are often sufficient treatments in early stages. In later states, oral hypoglycemic medications and insulin are usually required for adequate control.

Individuals with diabetes are at significantly increased risk of: heart disease, stroke, peripheral vascular disease, kidney disease, lower extremity ulcers and amputations, retinal disease, neuropathy, infections, digestive disease, and periodontal disease. The total direct and indirect costs in the United States due to diabetes mellitus have been estimated at $245 billion per year. Diabetes is the seventh leading cause of death in the United States, with 234,000 death certificates listing diabetes as an underlying or contributing cause of death annually.

Obesity Epidemic

There is an epidemic of obesity and type 2 diabetes around the world, including the United States. (See Section 27.4 on obesity.) Approximately 20 million Americans have type 2 diabetes, which is more than triple the prevalence in 1980. Approximately 35 percent of adults meet the criteria for obesity (BMI more than 30 kg/m^2), which is double the prevalence from 25 years ago. High caloric intake and lack of exercise are fueling this worldwide increase in the prevalence of obesity and type 2 diabetes. There are differences in genetic predisposition for type 2 diabetes among different ethnicities. For example, an estimated one in three Caucasians and as many as one in two Hispanic Americans born after 2000 will develop diabetes in their lifetime. Asian Americans also appear to develop type 2 diabetes with less weight gain compared to Caucasians.

The epidemic of obesity in the United States has also caused a marked increase in the prevalence of metabolic syndrome. This pre-diabetic syndrome includes central obesity, elevated blood glucose, elevated blood pressure, low–high-density lipoproteins (HDL), and elevated triglycerides. The elevated blood glucose is related to increased insulin resistance, and in the United States over 80 million individuals are considered to have "pre-diabetes" based on elevated blood glucose. Prospective studies have shown that metabolic syndrome raises the risk for development of coronary vascular disease by over 50 percent and the risk for type 2 diabetes by approximately threefold compared to individuals without metabolic syndrome.

DIABETES: PSYCHIATRIC SYMPTOMS OFTEN OCCUR AT KEY DEVELOPMENTAL STAGES IN PATIENTS WITH DIABETES

Patients and their family members are most likely to exhibit psychiatric symptoms at key stages of diabetes management, including reactions to initial diagnosis, at the time when a complication (i.e., neuropathy, retinopathy) develops, or when intensification of a treatment regimen is needed such as starting insulin. Conversely, chronic stressors and psychiatric symptoms may markedly interfere with adherence to self-care regimens (e.g., diet, exercise, taking medications, cessation of smoking) and diabetic disease control. Family conflict regarding diabetes tasks such as following a new dietary regimen or increasing exercise may also interfere with treatment.

Opportunities for clinicians to screen for psychosocial difficulties and specific psychiatric disorders include at the time of initial diagnosis of diabetes, during regular scheduled diabetes visits, during visits with diabetes nurses and/or educators, during hospitalization or development of a diabetes complication, and when discussing problems with glucose control or adherence to lifestyle or medication regimens. Clinicians may

want to consider use of validated brief screening tools for depression and anxiety such as the Patient Health Questionnaire [PHQ-9] and Generalized Anxiety Disorder [GAD-7], respectively, or use the Problem Areas in Diabetes (PAID) instrument, as a way of better understanding diabetes-specific emotional distress.

Patient-Provider Relationship: Effect of Psychiatric Comorbidity and Maladaptive Attachment Styles

Effective self-management of either type 1 or type 2 diabetes requires a close working relationship with a physician, other health care providers, and with family members. This partnership with the health care provider allows the patient and his or her providers to work in an interactive, iterative fashion with the goal of achieving ideal medication management that maximizes benefits and minimizes side effects while implementing important lifestyle changes (e.g., losing weight, increasing exercise, cessation of smoking).

Satisfaction with the quality of care of a patient's physician has been linked with enhanced adherence to medical regimens in patients with diabetes. Depressive disorders and maladaptive attachment styles have been shown to be associated with decreased ratings of satisfaction with care in patients with diabetes. Moreover, primary care effectiveness studies testing interventions that were shown to improve major depressive outcomes showed that enhanced depressive outcomes were associated with increased satisfaction with care.

Depression and anxiety disorders in adulthood are also associated with higher rates of having experienced childhood adversity (abuse and neglect) as well as with maladaptive attachment styles, which are associated with lower levels of trust of others. Maladaptive attachment styles may make it difficult for a patient to develop trusting relationships with others, including health care providers. Two of the styles, *dismissing* and *fearful attachment style*, are associated with difficulty trusting others and are the predominant relationship styles in 48 percent of adult primary care patients with diabetes. Adults with dismissing attachment style develop strategies early in life in which they become highly independent and uncomfortable trusting others, while adults with a predominantly fearful attachment style may have an initial desire for social contact at times of distress that is inhibited by fear of rejection, resulting in approach-avoidance behavior. Among over 4,000 primary care patients with type 2 diabetes, dismissing and fearful attachment styles were associated with 46 and 64 percent more missed primary care appointments, respectively, and both styles were associated with less satisfaction with care as compared to those with secure attachment style. In the same sample, compared to patients with secure attachment style, those with dismissing attachment style had lower levels of exercise, foot care, adherence to diet, adherence to

oral hypoglycemic medications, and higher rates of smoking. Other studies have also shown that in individuals with diabetes, dismissing and fearful attachment styles are associated with decreased glucose monitoring, nonadherence with insulin injections, higher glucose levels, and with higher mortality over a 5-year period. The adverse impact of dismissing and fearful attachment described above was independent of the effect of depression.

Depression may also increase the severity and impact of maladaptive attachment style traits. It has recently been shown that decreased depressive symptoms over a 10-month period was associated with a transition to more adaptive attachment behaviors and greater trust of others. This suggests that affective illness increases maladaptive interpersonal behaviors, potentially making collaboration with the physician more difficult and contributing to patients with these disorders receiving less adequate preventive medical care and having poorer diabetes outcomes.

Type 1 Diabetes: Developmental and Family Issues

Having a child with a chronic illness often causes a great deal of stress for families and may interfere with normal childhood development. Research suggests higher divorce rates in families where one child has a significant chronic illness or developmental disability, and children with chronic illnesses such as diabetes have twofold higher rates of affective illness and anxiety disorders.

Type 1 diabetes may be a particularly taxing illness for children and their families. Family-based care of diabetes requires checking blood glucose several times a day and altering insulin doses depending on level of blood glucose. This illness also requires following careful dietary regimens and getting regular levels of exercise. When children develop this illness early in life much of the responsibility for diabetes care depends on parents and may add significantly to parental stress. Recurrent hospitalization, inability to participate in normal childhood outings (particularly those away from home), and parental overprotectiveness can all delay normal childhood progression toward mastery and autonomy. Moreover, as the child moves into adolescence, there may be greater conflict as the child tries to normally separate and individuate. Acting out behavior such as dietary indiscretions, missing insulin doses, and avoiding checking blood sugars can all lead to poor glucose control during adolescence. Individual and family therapy approaches may be necessary to improve the quality of diabetes care and to help adolescents and their families adjust to new developmental phases.

Psychiatric Comorbidity

Thomas Willis, an early British physician, stated in the 17th century that diabetes is caused by "sadness or long sorrow and other depressions and disorders." Indeed, in recent years, researchers have shown that the rate of diabetes is approximately 1.5- to 2-fold higher in patients with affective disorders (major depression and bipolar disorder) and schizophrenia compared to the general population. The increased risk of diabetes in patients with depression has also been shown to be present within all racial/ethnic groups in the Multiethnic Study of Atherosclerosis (MESA). Looked at another way, it has been shown that there is an elevated risk of having diabetes (odds ratio [OR] = 2.6) in Medicaid beneficiaries with psychiatric disorders as compared to those without these disorders.

Chronic stress may have a role in the development of obesity-related illnesses by causing sustained activation of the HPA axis. A biologic correlate of chronic stress is activation of the HPA axis, due to either excess activity of corticotrophin-releasing hormone or defective feedback activity due to glucocorticoid receptor resistance, leading to hypercortisolemia. Chronic stress associated with mental illness that leads to hypercortisolemia may cause centripetal obesity and metabolic syndrome. Studies have shown that depression is associated with glycemic dysregulation in cross-sectional and longitudinal studies. Depression is associated with increased levels of glucocorticoids, catecholamines, and growth hormone; changes in glucose transport function; and secretion of inflammatory cytokines, which could lead to insulin resistance and, ultimately, be causal factors in development of diabetes as well as complications of diabetes. High salivary cortisol levels have been shown to be associated with decreased insulin sensitivity in unmedicated depressed patients. Adverse health habits such as smoking, poor diet, and sedentary lifestyle also occur in a higher prevalence in patients with mental illness compared to those without, which may also increase the incidence of obesity, metabolic syndrome, and type 2 diabetes.

Recent studies have found that obesity was associated with an approximately 25 percent higher risk of having mood (major depression, bipolar disorders) and anxiety disorders. Longitudinal studies that followed adolescents into adulthood also found that depression in adolescence significantly increased the risk for obesity in early adulthood, especially in females. A recent meta-analysis found that depression earlier in life increased the risk of development of type 2 diabetes by approximately 60 percent. In elderly community respondents (65 years or older), depression was shown to raise the risk of developing type 2 diabetes by approximately 50 percent. Diabetes earlier in life also increases the risk of developing subsequent depression by approximately 15 percent.

A meta-analysis of 39 studies reported a 12 percent prevalence of major depression in patients with diabetes. This meta-analysis also showed that in controlled studies there was a twofold increased prevalence of

depression in patients with diabetes compared to controls. A recent, large health maintenance organization–based study found that 12 percent of approximately 4,800 adult patients with predominantly type 2 diabetes (96 percent) met criteria for major depression and 8.5 percent for minor depression. The group with major depression and diabetes compared to those with diabetes alone were younger, more likely to be female, were less educated, had significantly higher medical comorbidity, had more diabetes complications, had a longer duration of diabetes, had higher hemoglobin A1c (HbA1c) and BMI values, and were more likely to be smokers. The patients with depression and diabetes developed diabetes an average of 5 years earlier as compared to those with diabetes alone.

Depression tends to be either a chronic or recurrent condition in patients with diabetes. Data from the large HMO study described above showed that 63 percent of patients with diabetes and major depression had 2 years or longer of significant depressive symptoms. Another study found that 79 percent of patients with diabetes who had major depression relapsed over a 5-year follow-up period with a mean of four episodes per patient.

Less research has focused on the prevalence of depression among youth with type 1 diabetes. A recent study examined the rate of depression in 2,672 youth (aged 10 to 21 years), most of whom had type 1 diabetes. Based on a depression screening measure, 14 percent were mildly depressed and 8.6 percent were moderately to severely depressed. Participants with depression and diabetes had significantly higher HbA1c levels and a higher number of emergency room visits. Another study from a large national database found that youth with type 1 diabetes who were hospitalized with diabetes who had a comorbid depression diagnosis were twice as likely to be rehospitalized in the next 12-month period compared to youth with a diabetes diagnosis alone.

Type 2 diabetes has been shown to be associated with a greater prevalence of cognitive impairment, incidence of cognitive decline, and incidence of Alzheimer disease. A recent review of 25 prospective studies found that people with diabetes had a 1.2 to 1.5 greater risk of cognitive decline as compared with people without diabetes. Patients with type 2 diabetes have reduced sensitivity to insulin (i.e., increased insulin resistance), and reduced neuronal sensitivity to insulin has also been found. It is posited that reduced neuronal insulin sensitivity exacerbates neurodegeneration, causing cognitive decline and dementia. Diabetes is also associated with a higher risk of vascular disease including stroke, which also may increase risk of dementia, particularly vascular dementia. A complication of treatment with glucose lowering drugs, particularly insulin, is hypoglycemia. Serious hypoglycemic episodes associated with emergency room visits or hospitalizations have been found to increase risk of dementia in patients with diabetes. The presence of comorbid depression can further

contribute to the association between diabetes and cognitive decline. In the Diabetes and Aging Study, which was a cohort investigation examining risk factors for dementia in individuals with diabetes, over 19,000 patients were examined over a 3- to 5-year period. Patients with comorbid depression had a 100 percent increased risk of dementia during the observation period.

Cognitive decline is likely to interfere with self-management of diabetes, potentially worsening insulin resistance and hyperglycemia. Therefore, screening patients with diabetes using tools like the Mini-Mental Status Examination or the Montreal Cognitive Assessment (MoCA) is essential to ensure early recognition. Enlisting family support for patients with cognitive impairment and diabetes is essential to slowing progression of diabetes.

A recent meta-analysis examined the prevalence of eating disorders among patients with type 1 diabetes versus controls. The prevalence of anorexia nervosa was not significantly different in patients with type 1 diabetes versus controls (0.27 vs. 0.06 percent) while the prevalence of bulimia nervosa (1.73 vs. 0.69 percent) and anorexia nervosa and bulimia nervosa combined (2.00 vs. 0.75 percent) were significantly higher in patients with type 1 diabetes. Some patients with insulin-dependent diabetes also try to lose weight by skipping or reducing their insulin injections—a phenomenon called "insulin withholding"—predisposing them to poor glucose control and future complications.

Effects of Depression on Diabetes Symptom Burden

Five studies (N = 5,094 patients) have examined the association of comorbid depression and anxiety in patients with diabetes on physical symptoms characteristic of poor glucose control (i.e., polyuria, polydipsia). These studies have shown that depression and anxiety disorders were more highly associated with self-reported diabetes symptoms than were HbA1c levels or number of diabetes complications. A study by Ludman and colleagues (Fig. 27.6–1) demonstrates this finding by showing that, after controlling for potential confounders, patients with diabetes and depression were significantly more likely to report ten diabetes symptoms compared to patients with diabetes alone (odds ratios between 2 and 5 for each symptom). Interestingly, patients with two or more diabetes complications compared to those with fewer than two complications had a significantly higher likelihood of having nine of ten of the diabetes symptoms, but odds ratios were only between 1.3 and 2.0. These findings demonstrate that depression is more highly associated with diabetes symptom burden than number of diabetes complications. A study that compared collaborative care versus usual primary care in treating depression in 387 patients with diabetes found that not only was collaborative care associated with better depression outcomes but also

significantly decreased diabetes symptom burden compared to usual care despite not having a significant effect on glycated hemoglobin levels.

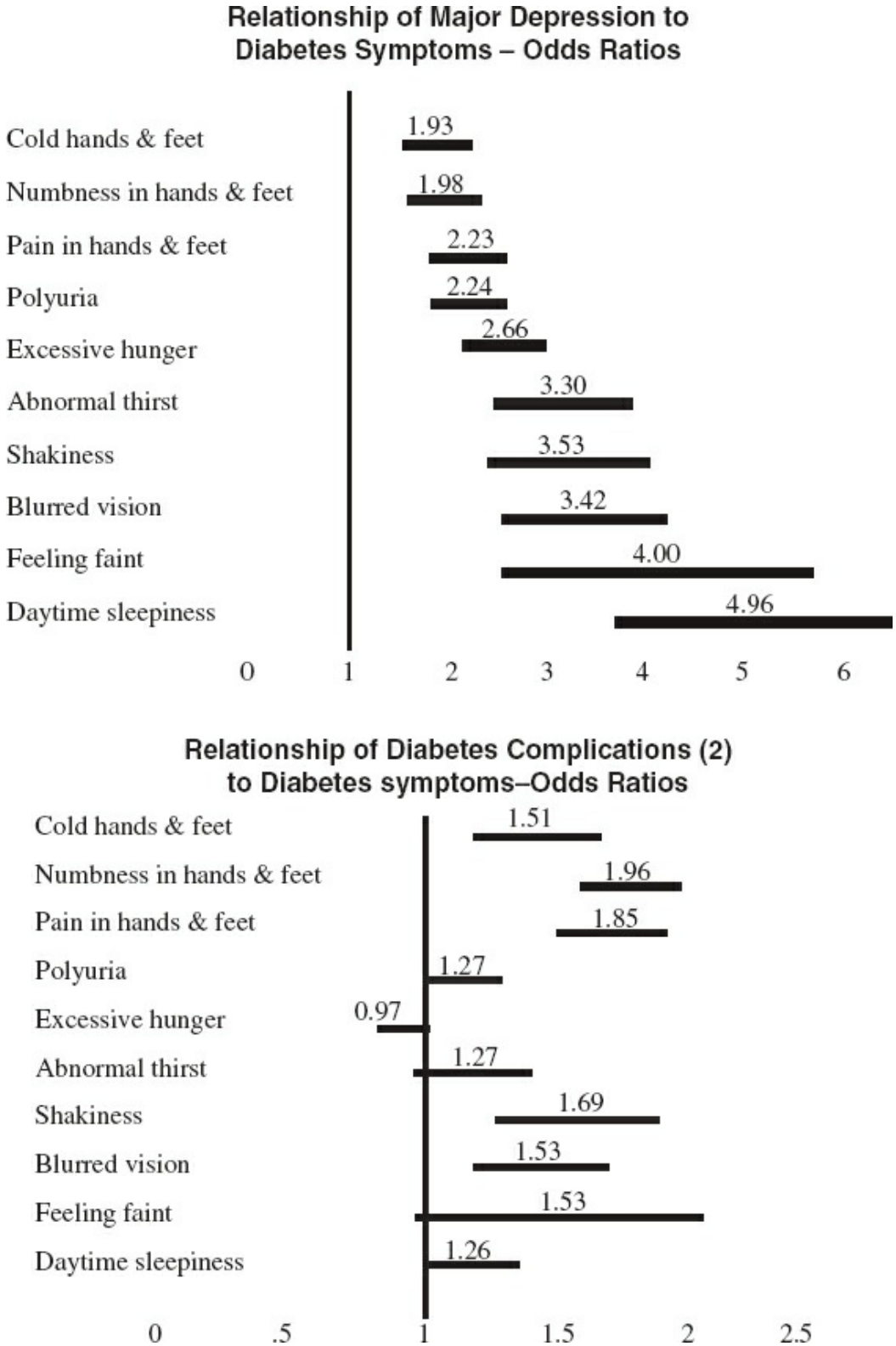

FIGURE 27.6–1. Relationship of major depression to diabetes symptoms (*top*) and diabetes complications to symptoms (*bottom*). (Data from Ludman EJ, Katon W, Russo J, et al. Depression and diabetes symptom burden. *Gen Hosp Psychiatry*. 2004;26(6):430–436.)

Some symptoms of diabetes such as chronic pain due to neuropathy may lead to bidirectional adverse experiences with depression. Depression has been shown to lead to amplification of pain complaints, thus preventing habituation to aversive symptoms. Recent biologic evidence suggests that the depletion of serotonin and norepinephrine that occurs in depression may lead to dysregulation of descending pain pathways that normally dampen aversive symptom perception. Longitudinal studies have also shown that chronic pain is associated with a higher incidence of depression. Antidepressant medications, such as venlafaxine and duloxetine, that increase norepinephrine and serotonin have been found to be effective in treating diabetic neuropathy, even in non-depressed patients.

Effect of Depression and Anxiety on Functioning

Multiple studies have shown that depression and anxiety lead to additive social and vocational disability in patients with chronic medical illness. The Pathways Epidemiologic Study has longitudinally surveyed over 4,000 patients in a large health maintenance organization with diabetes and shown that comorbid major depression in patients with diabetes was associated with a 10-fold increase in likelihood of functional impairment on the World Health Organization (WHO) Disability Scale and a fourfold increased odds of having 20 or more days of reduced ability to do housework in the past 90 days after controlling for diabetes complications, medical comorbidity, and glycemic control. As shown in Figure 27.6–2, diabetes complications had much less impact on functional disability than depression. Among 1,642 nonelderly persons in the Pathways Epidemiologic Study (excluding homemakers and retired persons), comorbid depression was associated with a fourfold increased likelihood of missing more than 5 days of work in the prior month, greater difficulty performing work tasks, and greater levels of unemployment after controlling for comorbid chronic medical illness, diabetes complications, and glycemic control.

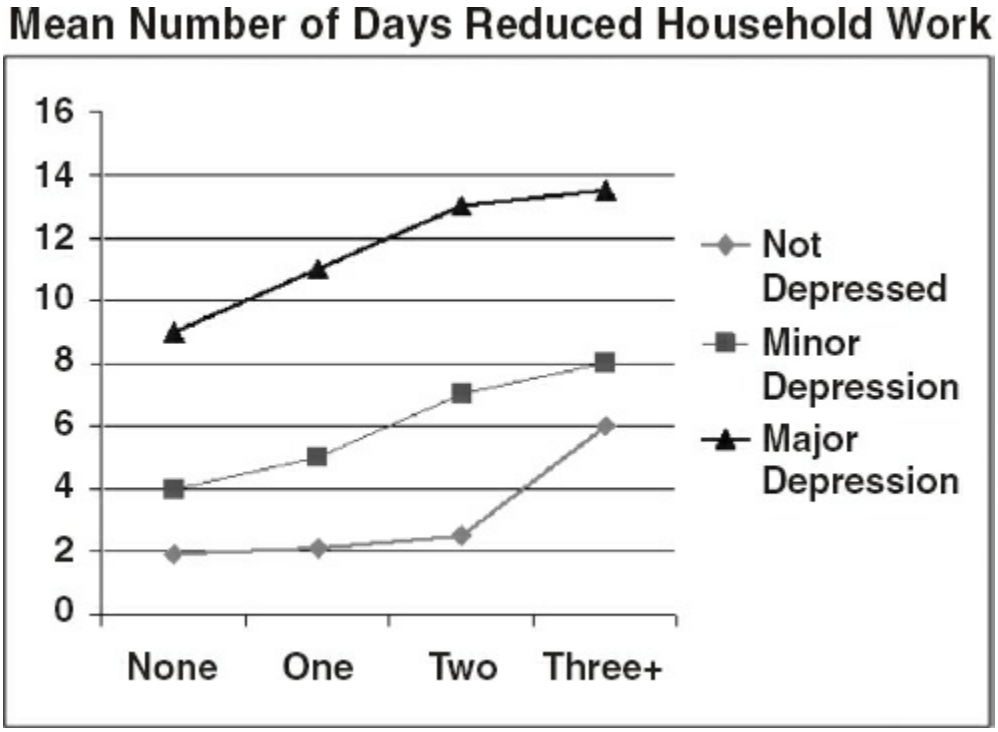

FIGURE 27.6–2. Relationship of comorbid depression to disability in patients with diabetes. (Data from Von Korff M, Katon W, Lin EH, et al. Potentially modifiable factors associated with disability among people with diabetes. *Psychosom Med*. 2005;67(2):233–240.)

A large longitudinal epidemiologic study that followed approximately 2,800 elderly Hispanic Americans in the Southwest over 7 years found that the interaction of depression and diabetes was synergistic in predicting greater incidence of decrements in activities of daily living, even after controlling for sociodemographic characteristics. Both depression and diabetes independently predicted increased incidence of decrements in function and, when they were comorbid, there was more than an additive adverse effect on functioning.

A study that compared a collaborative care intervention to usual primary care in enhancing the quality of depression care in over 400 elderly patients with diabetes found that the intervention patients not only had improved depressive outcomes, but also had improved physical functioning over a 12-month period.

Effect of Depression and Other Psychiatric Comorbidities on Disease Self-Management

Patients with diabetes must optimally follow complex self-care regimens including changing their diet, increasing exercise, checking blood glucose and altering medication dosages based on glucose fluctuations, stopping smoking, and taking multiple disease control medications. Depression may adversely affect these self-management activities by decreasing energy, concentration and memory, and motivation. A recent meta-analysis of over

20 studies of patients with chronic medical illness found that depression and anxiety were associated with a threefold higher rate of nonadherence to medical regimens.

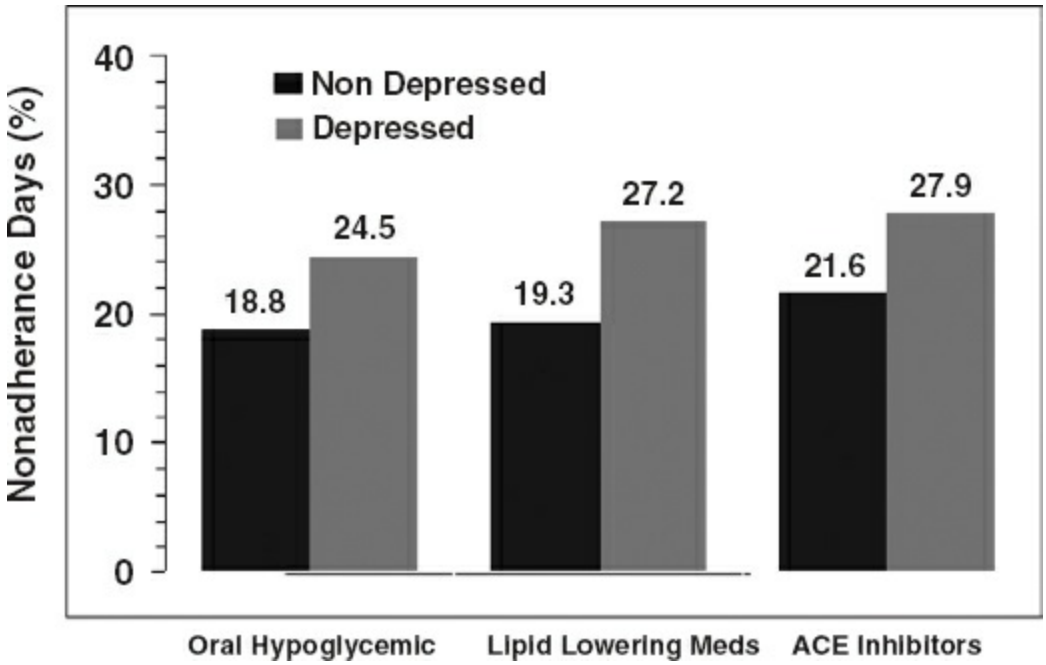

FIGURE 27.6–3. The association of depression with medication adherence in patients with diabetes. (Data from Lin EH, Katon W, Von Korff M, et al. Relationship of depression and diabetes self-care, medication adherence, and preventive care. *Diabetes Care*. 2004;27(9):2154–2160.)

In the Pathways Epidemiologic Study, patients with major depression and diabetes compared to those with diabetes alone were less likely to eat a healthy diet, more likely to eat high fat foods, less likely to exercise, and more likely to smoke. Depressed patients were also more likely to have lapses in refilling prescriptions for oral hypoglycemic, lipid lowering, and antihypertensive medications (Fig. 27.6–3). These lapses in taking medications may lead to a higher risk of diabetes complications.

Other psychiatric illnesses such as bulimia, night eating syndrome, binge eating syndrome, and chronic psychiatric illnesses such as schizophrenia and bipolar illness have also been shown to be associated with poor self-management of chronic illnesses such as diabetes as well as with poor quality of medical care. Patients with schizophrenia and bipolar disorder are twice as likely to die from cardiovascular disease as the general population. Research has shown that patients with diabetes and schizophrenia are treated less aggressively for their cardiovascular risk than patients with diabetes without a mental disorder.

Depression and Diabetes: Cardiac Risk Factors

Approximately 50 percent of mortality in patients with diabetes is related

to myocardial infarction or stroke. Comorbid depression may increase this risk because of the association of depression with multiple cardiac risk factors as well as the association of depression with biologic factors such as elevated sympathoadrenal activity, elevated cortisol, increased platelet adhesiveness, increased inflammatory markers, and endothelial dysfunction. The patients at highest risk of death from these two macrovascular complications have multiple cardiac risk factors. A longitudinal study of approximately 93,000 women found that those community respondents with depression compared to those without were significantly more likely to have multiple cardiac risk factors, including smoking, obesity, sedentary lifestyle, high cholesterol, diabetes, and hypertension. In a study of over 4,000 patients with diabetes, those with comorbid depression were found to be almost twice as likely to have three or more of eight Framingham cardiac risk factors as those with diabetes alone.

Both major depression and depressive symptoms have been associated with glycemic dysregulation in cross-sectional and longitudinal studies. As shown in Figure 27.6–4, there was a dose–response relationship between depression and HbA1c levels in the Pathways Epidemiology Study. Patients with both comorbid minor and major depression and diabetes had significantly higher HbA1c levels compared to patients with diabetes alone. Poor glycemic control is a risk factor for both macrovascular (i.e., heart attack, stroke) and microvascular (i.e., nephropathy, neuropathy, and retinopathy) complications. A large observational study reported that for every half-point decrease in the HbA1c, there is an 11.5 percent reduction in risk of diabetes-related complications.

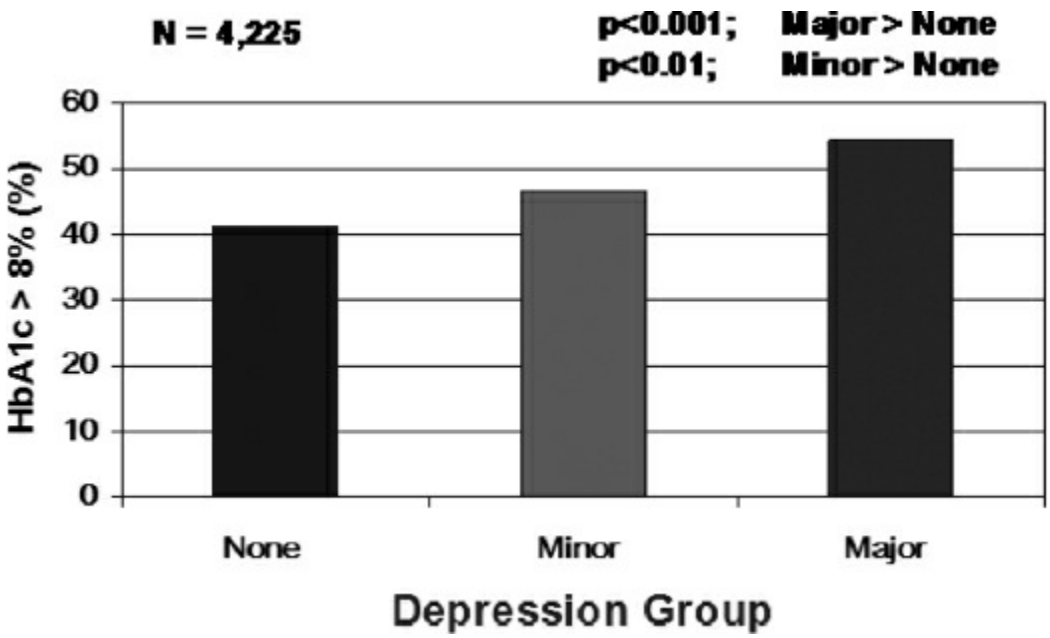

FIGURE 27.6–4. The association of depression with hemoglobin A1c (HbA1c) levels. Adjusted for demographics, medical comorbidity, diabetes severity, diabetes type and

duration, treatment type, and clinic. (Data from Katon W, Von Korff M, Ciechanowski P, et al. Behavioral and clinical factors associated with depression among individuals with diabetes. *Diabetes Care*. 2004;27(4):914–920.)

Patients with chronic mental illnesses such as bipolar illness and schizophrenia have life spans that are shorter by approximately 10 to 15 years compared to controls. Recent data from a large, national representative sample in the United Kingdom found that patients with severe mental illness who were less than 50 years of age had a threefold increased death rate from CHD and in those 50 to 70 years of age, there was a twofold increased risk of death from CHD and stroke. The increased risk of death also occurred after controlling for use of antipsychotic medication, although patients on a higher prescribed dosage of these medications had a greater risk of mortality. Patients with schizophrenia and bipolar disorders have a higher prevalence of most medical illnesses and often receive poorer quality medical care from those illnesses. In the Clinical Antipsychotic Trials in Intervention Effectiveness (CATIE) trial, which tested the effectiveness of a large number of antipsychotic medications in patients with schizophrenia, researchers examined the quality of medical care. Only 70 percent of patients with diabetes were taking oral hypoglycemic medication or insulin, only 38 percent of patients with hypertension were taking antihypertensive medications, and only 12 percent of patients with dyslipidemias were taking lipid- or triglyceride-lowering medications. In addition to poor quality of medical care, high rates of adverse health behaviors, such as smoking, obesity, and sedentary lifestyle, are associated with a higher risk of adverse macrovascular outcomes in patients with chronic mental illness and diabetes. Recent studies have shown that integrating medical care into community mental health centers in an organized approach can improve preventive care and physical functioning.

Psychiatric Illness: Diabetes Complications

A meta-analysis of 27 studies showed a significant association between depression and each of six diabetes complications (heart attack, stroke, nephropathy, neuropathy, retinopathy, peripheral vascular disease). There are likely to be bidirectional effects with depression associated with increased risk of complications due to poor self-care and the biologic abnormalities associated with depression such as increased cortisol levels. Complications of diabetes that cause functional decline, such as retinopathy or amputation, may also lead to secondary depression. Several longitudinal studies have also shown that depression was associated with increased complications. In a large longitudinal study of Hispanic elderly respondents in the southwestern part of the United States, elderly patients with diabetes who were initially depressed versus those with diabetes alone

were three times as likely to develop macrovascular and microvascular complications over a 7-year period. In the Pathways Epidemiologic Study that followed over 4,000 patients with diabetes over 5 years, those with depression and diabetes versus those with diabetes alone were 36 percent more likely to develop microvascular complications and 24 percent more likely to develop macrovascular complications. This same study also found that patients with depression and diabetes compared to those with diabetes alone had twofold increased risk of developing a foot ulcer (which is often a precursor to development of a below knee amputation), were about 15 percent more likely to develop retinopathy and were 40 percent more likely to have a serious hypoglycemic episode leading to an emergency room visit or hospitalization. A large longitudinal study of children with type 1 diabetes also showed that the risk for developing retinopathy was independently associated with three risk factors: duration of diabetes, time spent with poor glucose control, and time spent in major depression.

Two studies have also found that patients with comorbid depression and type 2 diabetes compared to those with diabetes alone had an approximately twofold greater risk of developing dementia. A study of approximately 29,000 patients with type 2 diabetes developed a risk prediction model for dementia and identified depression as an important clinical risk factor. Another recent study that analyzed data from a large randomized trial that aimed to optimally control glucose and cardiovascular risk factors in patients with diabetes, found that those with comorbid depression and diabetes versus those with diabetes alone had a greater risk of cognitive decline on neuropsychological tests over a 40-month period.

Effect of Depression on Medical Costs

Given the adverse effect of depression on somatic symptom burden, functional impairment, poor adherence to self-care, and increased complications, it is not surprising that several studies have shown that total health service costs of patients with diabetes and depression are significantly higher than those with diabetes alone. Studies that have carefully controlled for demographic factors, severity of diabetes, and medical comorbidity have found that total health service costs are approximately 70 percent higher for individuals with major depression and diabetes than diabetes alone. This is important for health care systems because patients with diabetes have approximately two- to fourfold higher medical costs compared to patients without diabetes. Therefore, a 70 percent increase in costs of these already expensive patients adds a disproportionate amount of total costs to a health care system. Recent studies have also shown that patients with depression and diabetes compared to those with diabetes alone have an over twofold risk of an intensive care unit (ICU) admission after controlling for sociodemographic

factors and severity of diabetes. ICU admission costs are about double the costs of usual hospital admissions. Another recent study showed that those with depression and diabetes compared to those with depression alone were about 40 percent more likely to have an ambulatory care sensitive hospitalization (a potentially avoidable hospitalization if patients received higher quality of primary care).

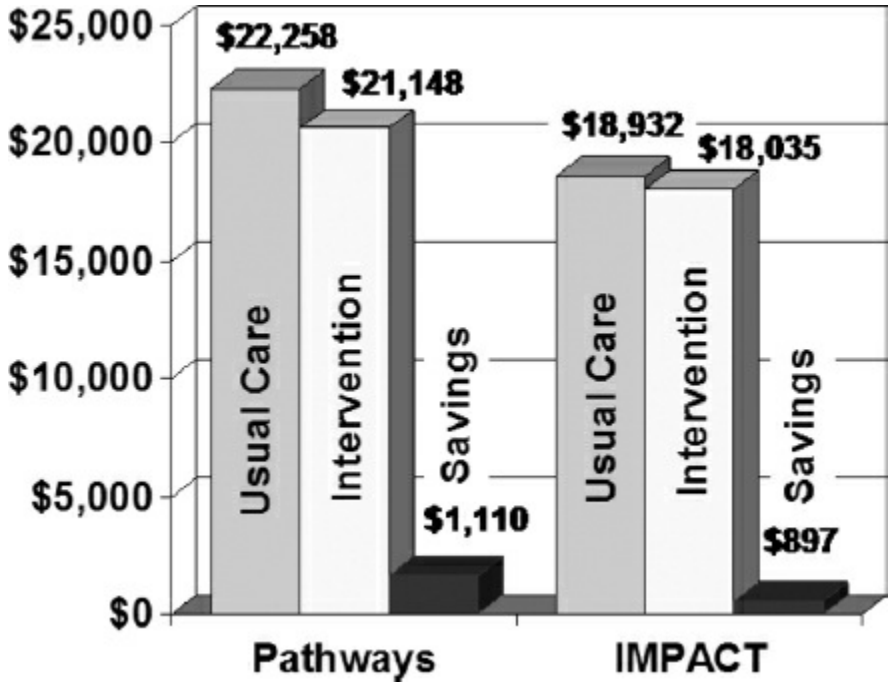

FIGURE 27.6–5. Enhanced treatment of depression lowers total health care costs over 2-year period.

Studies have also shown that the increase in costs in patients with diabetes and depression compared to those with diabetes alone is not explained by an increase in mental health care costs. Rather, there is an increase in costs in every category that is measured including primary care visits, medical specialty visits, nonpsychiatric prescriptions, x-ray and lab costs, emergency room and urgent care visits, inpatient medical care and mental health visits, and prescriptions. Two large primary care–based studies (the Pathways and IMPACT [Improving Mood—Promoting Access to Collaborative Treatment for Late Life Depression] studies) that tested collaborative care interventions compared to usual primary care in patients with depression and diabetes showed that improving quality of care and outcomes of patients with depression results in cost savings in medical costs that offset the additional costs of the collaborative care mental health intervention (Fig. 27.6–5).

Psychiatric Illness and Mortality in Patients with Diabetes

A recent systematic review of 10 longitudinal studies found that comorbid

depression in patients with diabetes was associated with a 50 percent increased risk of mortality. One study showed that both depression and diabetes in elderly patients increased the risk of mortality by almost one-third, and respondents with both disorders had a fourfold increase in mortality compared to elderly with neither disorder.

Among patients with diabetes in the Pathways Epidemiology Study, after controlling for age, gender, and socioeconomic factors, comorbid major depression was associated with over a twofold increase in mortality over a 3-year period compared to nondepressed patients. Minor depression also increased the risk of mortality by 67 percent. Even when diabetes, medical disease severity, and adverse health behaviors (smoking, obesity, sedentary lifestyle, HbA1c) were controlled for, major and minor depression were still associated with an approximately 50 percent increase in mortality.

Treatment Trials: Diabetes and Depression

Three large primary care–based randomized trials tested collaborative care treatment of depression among patients with diabetes: 418 patients with diabetes and depression age 60 and over in the IMPACT trial; 329 mixed age patients with diabetes and depression with a mean age of 59 years in the Pathways trial, and 387 predominantly Hispanic patients in which 52 percent of patients were 55+ years old. These trials tested a stepped-care nurse intervention for depression versus usual primary care. All three trials showed marked intervention versus usual care effects in quality of depression care and significant improvements in depression outcomes (effect sizes between 0.47 and 0.68), but no significant changes in self-care such as diet or checking blood glucose and no significant difference in HbA1c levels. Approximately half of usual primary care controls were treated with antidepressants by their primary care physician during the 12-month trials.

Three efficacy studies that randomized patients with major depression and diabetes to antidepressant versus placebo or CBT versus a diabetes education group found significant improvement in depression in all studies (effect sizes in antidepressant trials between 0.62 and 0.63, and 1.16 in the CBT trial) but only one study showed a beneficial effect on HbA1c levels.

Among patients with depression and diabetes in the above three large trials, 60 to 80 percent were found to have double depression (i.e., concurrent major depression and dysthymic disorder). Double depression predicts treatment resistance and recurrence or relapse. Comorbid chronic medical illness in patients with depression is associated with a more chronic course, with slower recovery and less treatment response. Given the high rates of relapse and chronicity in patients with depression and diabetes, researchers have begun to test the effect of maintenance antidepressant treatment. Several antidepressant trials that tested the

effect of maintenance treatment have suggested that long-term improvements in treatment that resulted in enhanced rates of recovery and remission are associated with improved glycemic outcomes. A recent maintenance trial randomized 152 patients who had recovered with an open label trial of an SSRI to either SSRI or placebo for up to 52 weeks. Patients who received maintenance treatment with the SSRI had a significantly greater depression-free interval compare to those treated with placebo. Both depression recovery with the SSRI as well as sustained remission with or without active treatments were associated with improvements in HbA1c levels for at least 1 year.

A second maintenance trial that treated 93 patients with type 2 diabetes with acute phase bupropion (Wellbutrin) offered maintenance-phase bupropion treatment to patients who remitted (N = 63) at the dosage that was associated with remission. BMI, total fat mass, and HbA1c values decreased significantly, composite diabetes care improved over the initial acute phase, and these effects persisted through the maintenance phase. Reduction in both BMI and depression severity with bupropion treatment predicted lower HbA1c levels after acute phase therapy, but only a reduction in depression predicted lower HbA1c levels during the maintenance phase of treatment.

Diabetic neuropathy can lead to chronic pain in patients with diabetes and secondary depression. Research studies have shown that venlafaxine (Effexor), duloxetine (Cymbalta), gabapentin (Neurontin), and pregabalin (Lyrica) are all more effective than placebo in treating pain associated with diabetic neuropathy whether or not patients were depressed (Table 27.6–1). Prior older research has also shown that tricyclic antidepressants were more effective than placebo in treatment of diabetic neuropathy. Given that tricyclic antidepressants are associated with weight gain in about 20 percent of patients and with a risk of cardiac complications, use of newer antidepressants and anticonvulsants are better first choices in patients with diabetic neuropathy.

Table 27.6–1.
Percentage of Patients with Diabetic Neuropathy Who Had at Least Moderate Improvement in Pain in Randomized Controlled Trials

	Active Medication (%)	Placebo (%)
Venlafaxine (Effexor)	56	34
Duloxetine (Cymbalta)	53 (120 mg dosage) 43 (60 mg dosage)	27
Amitriptyline (Elavil)	74	41
Desipramine (Norpramin)	61	41
Gabapentin (Neurontin)	60	33
Pregabalin (Lyrica)	40	14.5

This series of randomized trials suggests three conclusions. First, organized depression treatment programs such as collaborative depression care are clearly effective for patients with diabetes. Second, the effects of depression treatment may be smaller for patients with chronic medical illness than in general primary care samples. Third, short-term treatment programs limited to improving depression treatment may not have clear or consistent effects on medical outcomes such as HbA1c values. The maintenance trials tend to show that patients who have remitted in acute phase treatment and who remain on antidepressants over the next year have improved glycemic control, whereas acute phase trials of antidepressants have inconsistent effects on HbA1c levels.

Piette et al. have proposed a model to explain how depression treatment may improve the outcomes of comorbid chronic disease, which they applied to diabetes. In this model, effective treatment of depression may be more difficult because of the adverse health behaviors (smoking, obesity, sedentary lifestyle) as well as higher symptom burden and functional disability associated with depression. Adverse health behaviors do not necessarily improve with effective depression treatment alone and are negative prognostic factors for depression outcomes (e.g., obesity is associated with negative social feedback, lower self-esteem, and less exercise, which all affect mood). New treatment models that provide both interventions aimed at improving quality of depression care and aimed at optimizing positive health behaviors (exercise, diet, self-management) and treatment of chronic medical illness by optimizing medication treatment in order to avoid complications have shown improvement in depression and medical disease control outcomes. Furthermore, studies show that a significant contributor to suboptimal outcomes in primary care is *clinical inertia*, i.e., lack of timely treatment intensification by the primary care physician or endocrinologist such as titrating medications for reducing blood glucose, blood pressure and cholesterol. As an example, among 161,697 Kaiser Permanente patients with diabetes, 50 percent were above target for A1c, blood pressure and LDL cholesterol. Lack of treatment intensification by health care teams was found to be a more significant contributor to out-of-target values as compared with poor patient treatment adherence (Fig. 27.6–6).

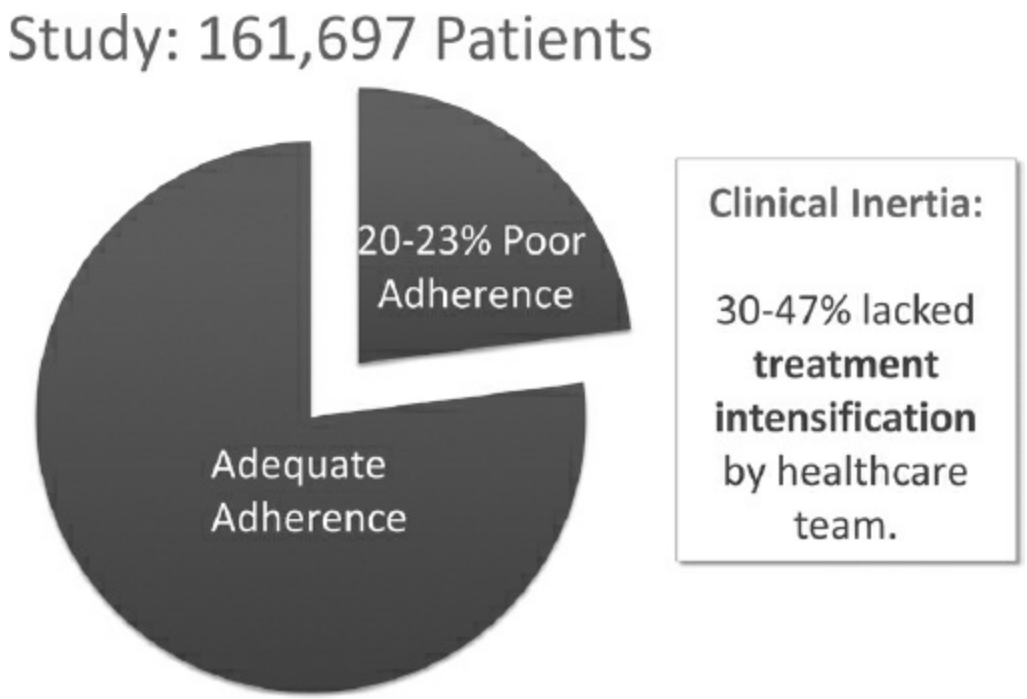

FIGURE 27.6–6. Clinical inertia and suboptimal outcomes in primary care. (Created with data from Schmittdiel JA, Uratsu CS, et al. Why don't diabetes patients achieve recommended risk factor targets? Poor adherence versus lack of treatment intensification. *J Gen Intern Med*. 2008;23(5):588–594.)

Using Multicondition Collaborative Care to Reduce Clinical Inertia. The TEAMcare clinical trial was a landmark study in that it expanded the collaborative care model from primarily focusing on treating depression to addressing multiple chronic conditions simultaneously using one case manager per patient, who focused on "A-B-C-D" i.e., A1c, blood pressure, cholesterol (LDL) and depression. This clinical trial randomized 214 patients with depression and with poorly controlled diabetes or heart disease (i.e., patients had to have either a HbA1c level of at least 8.5 percent, a systolic blood pressure >140 mm Hg or a low-density cholesterol level (LDL) of >130 mg/dL) to a *multicondition collaborative care intervention* aimed at improving quality of medication treatment of depression and these medical illnesses and improving self-care (exercise, diet, cessation of smoking and taking medications as prescribed). In addition to treat-to-target and systematic case review methods, this intervention also included motivational enhancement, problem solving treatment and behavioral activation clinical methods by front line nurse case managers. The intervention was associated with significant improvements in depression, HbA1c, SBP, and LDL control along with improvement in exercise.

Another randomized trial of 180 patients, prescribed medications for both diabetes and depression in primary care, tested an intervention that utilized care managers integrated into primary care to offer education, guideline-based treatment recommendations and monitoring of adherence

and clinical status. This intervention was significantly more likely than usual primary care to achieve HbA1c levels <7 percent and remission of depression. Another recent study evaluated the impact of a telephone delivered cognitive behavioral therapy intervention targeting depressive symptoms, a pedometer-based walking program aimed at increasing physical activity levels and improving diabetes-related outcomes in 291 patients with type 2 diabetes and significant depressive symptoms. Although no differences were found between intervention patients and usual care patients on the primary outcome of HbA1c levels, patients in the intervention group had significantly greater reductions in depressive symptoms, systolic blood pressure and step counts.

These trials show that to improve both depression and medical outcomes, programs need to address clinical inertia by combining interventions for depression as well as interventions aimed at improving adherence and intensification of diabetes disease control medications, and self-care activities (e.g., providing pedometers to encourage increased exercise).

Psychiatric Medications, Obesity, and Type 2 Diabetes

The widespread use of atypical antipsychotic medications for patients with bipolar illness, schizophrenia, and to a lesser extent treatment-resistant depression has led to concerns about these psychiatric medications causing obesity, metabolic syndrome, and type 2 diabetes. Table 27.6–2 describes the data on risk of obesity associated with the six FDA approved atypical antipsychotic medications. As shown in the table, clozapine (Clozaril) and olanzapine (Zyprexa) have the highest risk, quetiapine (Seroquel), risperidone (Risperdal), iloperidone (Fanapt), sertindole (Serdolect), and asenapine (Saphris)have intermediate risk, and ziprasidone (Geodon), aripiprazole (Abilify), and lurasidone (Latuda) the lowest risk. Numerous case reports have also described the less common occurrence of acute hyperglycemia after initiation of these medications before marked weight gain had occurred. Some but not all cases report hyperglycemia resolved after the medication was discontinued. These data suggest a direct role of atypical antipsychotic medications in increasing insulin resistance and causing type 2 diabetes in addition to an indirect role through causing obesity.

Table 27.6–2.
Drugs Weight Gain Effects

- Clozapine (Clozaril) and Olanzapine (Zyprexa): High
- Risperidone (Risperdal), Quetiapine (Seroquel), Iloperidone (Fanapt), Paliperidone (Invega), Sertindole (Serdolect), and Asenapine (Saphris): Moderate
- Ziprasidone (Geodon), aripiprazole (Abilify), and lurasidone (Latuda): Low

Data have shown two- to fourfold increased rates of metabolic syndrome in patients with schizophrenia treated with typical or atypical antipsychotic agents. An estimated 16 to 25 percent of patients with schizophrenia have type 2 diabetes. However, the risk of development of diabetes with atypical antipsychotic agents appears to be higher. Data from a large Veterans Administration study of over 38,000 patients treated with atypical versus typical antipsychotic agents show an approximately 10 percent increased rate of diabetes associated with treatment with atypical antipsychotic agents. The large scale CATIE trial also showed significantly more weight gain associated with atypical antipsychotic medications such as olanzapine and clozapine compared to those treated with the perphenazine (Trilafon), a typical antipsychotic agent.

Given the increasing concerns about the atypical antipsychotic agents causing obesity, metabolic syndrome, and type 2 diabetes, monitoring protocols have been recommended and describe the need to elicit a personal and family history about diabetes, heart disease, and other risk factors such as smoking, as well as regular monitoring of weight, waist circumference, blood pressure, fasting blood glucose, and lipid profiles.

It is important to emphasize that other psychiatric medications can also lead to weight gain. Use of anticonvulsant agents such as valproic acid (Depakene, Depakote) and carbamazepine (Tegretol) as well as lithium carbonate (Eskalith, Lithobide) in patients with bipolar disorder are also frequently associated with weight gain. Lamotrigine (Lamictal), another anticonvulsant medication that has been increasingly used in both bipolar disorder and treatment resistant depression, is generally not associated with weight gain. Although topiramate (Topamax) has not been FDA approved for use in bipolar disorder, psychiatrists often use lower dosages of this medication as an adjunctive medication due to its effect on decreasing appetite and inducing weight loss. Metformin (Glucophage) has also been shown to induce weight loss in patients taking atypical antipsychotics.

Among antidepressant agents, both tricyclic antidepressant medications and mirtazapine (Remeron) cause weight gain in a high percentage of patients. SSRIs appear to induce weight gain in only 5 to 10 percent of patients compared to placebo. Bupropion is the only antidepressant showing no weight gain or slight weight loss in placebo-controlled trials, and a recent acute and maintenance phase trial of bupropion treatment in patients with depression and diabetes also found that this antidepressant was associated with weight loss.

A Clinical Approach to Treating Depression in Patients with Diabetes

Table 27.6–3 outlines a diabetes-focused clinical approach to treating depression and other psychiatric problems in patients with diabetes. Because of the increased prevalence of depression in patients with diabetes,

screening tools such as the PHQ-9 should be used to recognize depression in diabetes settings. Frequently, when patients with diabetes and comorbid depression present to a health care provider, they are unaware of their depressive symptoms. Instead, they may experience physical symptoms of diabetes such as polyuria and fatigue. Commonly, they present with high levels of frustration, resignation, or guilt related to recent increased difficulties in managing their glucose and other diabetes-specific tasks. This may be because of the effects that depression has on the HPA axis, but also because of the aforementioned negative effects of depression in patients with diabetes such as amotivation, fatigue, anhedonia, lack of concentration, and hopelessness that make diabetes self-management a difficult task.

Table 27.6–3.
Diabetes-Focused Depression Care

In addition to standard evaluation of current (and past) episodes of depression:

1. Assume depressed patient is experiencing "loss of control" of disease self-management or outcomes resulting in frustration, bewilderment, resignation, etc.
2. Assess for understanding of bidirectional link between stress and suboptimal disease self-management and outcomes.
3. Define depression and how it overlaps with (and is distinct from) "stress."
4. Review symptoms of depression and how these symptoms overlap with or mimic diabetes symptoms. Discuss depression-related symptom amplification.
5. Consider "biological" sequelae resulting from comorbid depression:
   - e.g., increased glucose levels (increased glucometer mean + SD, higher HbA1c)
   - e.g., increased neuropathy pain, i.e., increased intensity, functional impairment
6. Consider "behavioral" sequelae resulting from comorbid depression:
   - e.g., decreased self-care (refilling prescriptions for medications and monitoring strips; glucose monitoring; exercise; diet). Look for helplessness/"giving up" or sense of being overwhelmed around disease self-care
   - e.g., missed appointments or high health care utilization (labs, visits, urgent care)
   - e.g., decreased satisfaction with care
   - e.g., decreased ability to rely on others
7. Screen for comorbid anxiety ± panic attacks and look for inability to differentiate anxiety symptoms from diabetes symptoms (e.g., hypoglycemia).
8. Screen for associated eating concerns (e.g., emotional eating in response to sadness/loneliness/anger, etc.; binge eating ± purging; night eating).
9. Break down tasks (self-management of diabetes, depression, heart disease, other illnesses) and help patient prioritize in order of importance
10. For patients with comorbid anxiety, consider using SSRI or SNRI
11. For patients with sexual dysfunction, consider using bupropion or mirtazapine (consider weight gain with mirtazapine)
12. For patients with significant neuropathy, consider bupropion, venlafaxine or duloxetine
13. Consider adjunctive brief psychotherapy for:
    - emotional eating (CBT)
    - breaking down problems (problem solving therapy)
    - improving treatment adherence (motivational interviewing)
    - improving effective patient-provider interactions (patient coaching)

SD, standard deviation; HbA1c, hemoglobin A1c; SSRIs, selective serotonin reuptake inhibitors; SNRI, serotonin and norepinephrine reuptake inhibitor; CBT, cognitive-behavioral therapy.

The clinician treating individuals with diabetes and depression, may benefit from the concept of "diapression" that recognizes that individuals with diabetes who develop depressive symptoms may experience the following characteristics: (1) depression can mimic and amplify diabetes symptoms, (2) depression is associated with poor diabetes self-management and treatment adherence, (3) depression can lead to adverse lifestyle habits, (4) depression is associated with changes in health care utilization patterns, and (5) symptoms of depression reduce levels of trust and satisfaction with care. A powerful way of engaging patients experiencing depressive symptoms may be to validate their experience of recent "loss of control" of their diabetes in a nonjudgmental way. Furthermore, this approach is useful in that it provides opportunities to educate patients about the interaction between stress, depression and anxiety on one hand and diabetes outcomes on the other. Patients are often relieved to hear that it is not their fault that their glucose control has become suboptimal.

Screening for comorbid anxiety is critical, since symptoms of anxiety and panic disorder may worsen control of diabetes. Furthermore, for patients on insulin who may have frequent hypoglycemic episodes because of poor self-management, it is often difficult to disentangle anxiety symptoms from hypoglycemic symptoms. Patients may eventually ignore important hypoglycemic symptoms or may catastrophize when anxious because they misinterpret anxiety symptoms as hypoglycemic symptoms. Using screening tools for anxiety such as the GAD-7 can be helpful in working with patients with diabetes.

Asking about eating behaviors and emotional regulation through eating is important since patients with diabetes and depression have a higher risk of bulimia, binge eating disorder, and night eating syndrome. A subgroup of patients eat a significant proportion (e.g., more than 25 percent) of their caloric intake after supper time, and such "night eating" behaviors are associated with higher HbA1c levels, more complications, and higher BMI. A nonjudgmental approach that includes optimally treating underlying anxiety or depression and potentially using CBT techniques for difficult eating behaviors can be very effective in relieving emotional eating in patients with diabetes.

Other strategies that are useful for patients with depression include problem-solving treatment, behavioral activation, CBT for depression, and use of medications, with care to choose medications that may add benefits for pain control in patients with neuropathy pain (about one-third of patients with diabetes have neuropathy).

FUTURE DIRECTIONS

The United States is facing an epidemic of type 2 diabetes, and psychiatric illness comorbidity with this medical illness is very high. It is important to screen patients with diabetes for psychiatric illness and to screen patients with psychiatric illness for diabetes since comorbidity is associated with adverse medical and psychiatric outcomes.

REFERENCES

American Diabetes Association, American Psychiatric Association, American Association of Clinical Endocrinologists, North American Association for the Study of Obesity. Consensus development conference on antipsychotic drugs and obesity and diabetes. *Diabetes Care*. 2004;27:596–601.

Anderson RJ, Freedland KE, Clouse RE, Lustman PJ. The prevalence of comorbid depression in adults with diabetes: A meta-analysis. *Diabetes Care*. 2001;24(6):1069–1078.

Black SA, Markides KS, Ray LA. Depression predicts increased incidence of adverse health outcomes in older Mexican Americans with type 2 diabetes. *Diabetes Care*. 2003;26(10):2822–2828.

Bogner HR, Morales KH, de Vries HF, Cappola AR. Integrated management of type 2 diabetes mellitus and depression treatment to improve medication adherence: a randomized controlled trial. *Ann Fam Med*. 2012;10(1):15–22.

Ciechanowski PS. Diapression: an integrated model for understanding the experience of individuals with co-occurring diabetes and depression. *Clin Diabet*. 2011;29(2):43–49.

Ciechanowski PS, Katon WJ, Russo JE, Walker EA. The patient-provider relationship: attachment theory and adherence to treatment in diabetes. *Am J Psychiatry*. 2001;158(1):29–35.

Ciechanowski P, Russo J, Katon W, et al. Influence of patient attachment style on self-care and outcomes in diabetes. *Psychosom Med*. 2004;66(5):720–728.

de Groot M, Anderson R, Freedland KE, Clouse RE, Lustman PJ. Association of depression and diabetes complications: a meta-analysis. *Psychosom Med*. 2001;63(4):619–630.

Hasnain M, Fredrickson SK, Vieweg WV, Vieweg WV, Pandurangi AK. Metabolic syndrome associated with schizophrenia and atypical antipsychotics. *Curr Diab Rep*. 2010;10(3):209–216.

http://www.diabetes.org/diabetes-basics/statistics/.

http://www.cdc.gov/diabetes/statistics/prev/national/figadults.htm.

http://www.cdc.gov/obesity/data/adult.html.

Jacobson AM. The psychological care of patients with insulin dependent diabetes mellitus. *N Engl J Med*. 1996;334(19):1249–1253.

Katon W, Lin EH, Kroenke K. The association of depression and anxiety with medical symptom burden in patients with chronic medical illness. *Gen Hosp Psychiatry*. 2007;29(2):147–155.

Katon WJ, Lin EH, Russo J, et al. Cardiac risk factors in patients with diabetes mellitus and major depression. *J Gen Intern Med*. 2004;19(12):1192–1199.

Katon WJ, Lin EH, Von Korff M, et al. Collaborative care for patients with depression and

chronic illnesses. *N Engl J Med*. 2010;363(27):2611–2620.

Katon W, Lyles CR, Parker MM, Karter AJ, Huang ES, Whitmer RA. Association of depression with increased risk of dementia in patients with type 2 diabetes: the Diabetes and Aging Study. *Arch Gen Psychiatry*. 2012;69(4):410–417.

Katon W, Unützer J, Fan MY, et al. Cost-effectiveness and net benefit of enhanced treatment of depression for older adults with diabetes and depression. *Diabetes Care*. 2006;29(2):265–270.

Katon WJ, Von Korff M, Lin EH, et al. The Pathways study: a randomized trial of collaborative care in patients with diabetes and depression. *Arch Gen Psychiatry*. 2004;61(10):1042–1049.

Lawrence JM, Standiford DA, Loots B, et al. Prevalence and correlates of depressed mood among youth with diabetes: the SEARCH for Diabetes in Youth study. *Pediatrics*. 2006;117(4):1348–1358.

Lin EH, Katon W, Von Korff M, et al. Relationship of depression and diabetes self-care, medication adherence, and preventive care. *Diabetes Care*. 2004;27(9):2154–2160.

Lustman PJ, Griffith LS, Freedland KE, Kissel SS, Clouse RE. Cognitive behavior therapy for depression in type 2 diabetes mellitus. A randomized, controlled trial. *Ann Intern Med*. 1998;129(8):613–621.

Lustman PJ, Williams MM, Sayuk GS, Nix BD, Clouse RE. Factors influencing glycemic control in type 2 diabetes during acute- and maintenance-phase treatment of major depressive disorder with bupropion. *Diabetes Care*. 2007;30(3):459–466.

Mezuk B, Eaton WW, Albrecht S, Golden SH. Depression and type 2 diabetes over the lifespan: a meta-analysis. *Diabetes Care*. 2008;31(12):2383–2390.

Nasrallah HA, Meyer JM, Goff DC, McEvoy JP, Davis SM, Stroup TS. Low rates of treatment for hypertension, dyslipidemia and diabetes in schizophrenia: data from the CATIE schizophrenia trial sample at baseline. *Schizophr Res*. 2006;86(1-3):15–22.

Park M, Katon WJ, Wolf FM. Depression and risk of mortality in individuals with diabetes: a meta-analysis and systematic review. *Gen Hosp Psychiatry*. 2013;35(3):217–225.

Piette JD, Richardson C, Himle J, et al. A randomized trial of telephonic counseling plus walking for depressed diabetes patients. *Med Care*. 2011;49(7):641–648.

Piette JD, Richardson C, Valenstein M. Addressing the needs of patients with multiple chronic illnesses: the case of diabetes and depression. *Am J Manag Care*. 2004;10:152–162.

Schmittdiel JA, Uratsu CS, Karter AJ, et al. Why don't diabetes patients achieve recommended risk factor targets? Poor adherence versus lack of treatment intensification. *J Gen Intern Med*. 2008;23(5):588–594.

Sernyak MJ, Leslie DL, Alarcon RD, Losonczy MF, Rosenheck R. Association of diabetes mellitus with use of atypical neuroleptics in the treatment of schizophrenia. *Am J Psychiatry*. 2002;159(4):561–566.

Simon GE, Katon WJ, Lin EH, et al. Diabetes complications and depression as predictors of health care costs. *Gen Hosp Psychiatry*. 2005;27(5):344–351.

Simon GE, Katon WJ, Lin EH, et al. Cost-effectiveness of systematic depression treatment among people with diabetes mellitus. *Arch Gen Psychiatry*. 2007;64(1):65–72.

Sullivan MD, Katon WJ, Lovato LC, et al. Association of depression with accelerated cognitive decline among patients with type 2 diabetes in the ACCORD-MIND trial. *JAMA Psychiatry*. 2013;70(10):1041–1047.

Wassertheil-Smoller S, Shumaker S, Ockene J, et al. Depression and cardiovascular sequelae in postmenopausal women. The Women's Health Initiative (WHI). *Arch Intern Med.* 2004;164(3):289–298.

Williams JW Jr, Katon W, Lin EH, Nöel PH, Worchel J. The effectiveness of depression care management on diabetes-related outcomes in older patients. *Ann Intern Med.* 2004;140(12):1015–1024.

▲ 27.7 Endocrine and Metabolic Disorders

KATHERINE E. WILLIAMS, M.D., AND NATALIE L. RASGON, M.D., PH.D.

INTRODUCTION

Endocrine disorders are common endocrine conditions that the psychiatrist is likely to frequently encounter in clinical practice. This chapter will focus on the most prevalent endocrine illnesses with neuropsychiatric symptoms, as well as those that are frequently comorbid with primary psychiatric disorders. Recently, there has been considerable progress in understanding the intimate and complex relationship between endocrine and psychiatric disorders, and there is increasing evidence that for many conditions a bidirectional relationship exists with respect to etiology and treatment. For instance, insulin resistance has been associated with the onset of major depression, and major depression has been associated with the onset of insulin resistance. Treatment of insulin resistance with insulin sensitizing agents has in fact been shown to treat not only the insulin resistance, but also depressive symptoms. Research is currently underway to elucidate the complex relationship between endocrinopathies and psychiatric illnesses.

GENERAL PRINCIPLES

Endocrinology is the study of hormones, or biochemical mediators that travel from their site of production to influence the function of other tissues. They exist in the form of peptide, steroid, or amines. Endocrinopathies result from excessive or deficient amounts of a hormone or from resistance to a hormone's action, for example, when target tissues do not respond to a normal level of the hormone. These latter cases are primarily hereditary and are characterized by normal or elevated levels of a hormone, despite an absence or deficiency of the hormone's action. Elevated levels result from the lack of normal negative feedback on the hormone's production, a consequence of the failure or deficiency of the hormone's action.

Hormone production is regulated largely by the levels and activity of the hormones themselves and additionally by hormones from other sites. For example hypothalamic releasing factors regulate the production of pituitary

hormones. The pituitary gland, situated in the sella turcica at the base of the skull, is divided into the anterior and posterior lobes. The anterior pituitary produces six hormones: thyroid-stimulating hormone (TSH), which regulates the production of thyroid hormones; growth hormone (GH), which promotes growth and development; prolactin (PRL), which stimulates lactation; luteinizing hormone (LH) and follicle-stimulating hormone (FSH), which are necessary for gonadal function in men and women; and adrenocorticotropic hormone (ACTH), which controls adrenal cortex production of glucocorticoid. The posterior pituitary stores two hormones: antidiuretic hormone (arginine vasopressin [AVPJ]), which regulates fluid balance, and oxytocin, which is necessary for milk let-down during breastfeeding. Both are produced by the hypothalamus. The hypothalamus additionally produces releasing factors that influence pituitary hormone production. These are growth hormone–releasing hormone (GHRH) and growth hormone release–inhibiting factor (GIH; also known as somatostatin), which influence production of GH; gonadotropin-releasing hormone (GnRH), which influences production of LH and FH; prolactin-inhibiting factor (PIF) (i.e., dopamine) and prolactin-releasing factor (PRF), influencing PRL production; and thyrotropin-releasing hormone (TRH) and corticotropin-releasing hormone (CRH), which influence TSH and ACTH release, respectively.

In addition to the hypothalamus and the pituitary, the major endocrine organs include the adrenal, thyroid, and parathyroid glands and the ovaries and testes. Endocrinopathies may produce a heterogeneity of psychiatric symptoms, as hormones affect a variety of organ systems. Patients are more likely to present primarily with psychiatric symptoms early in the course of endocrine disease and experience a worsening of the psychiatric condition and an emergence of physical symptoms with the progression of the disorder. Endocrinopathies should also be considered as a possible cause of treatment-resistant psychiatric disorders. Because endocrine disorders are common medical conditions, an evaluation for endocrine function is an essential part of the initial diagnostic workup. To identify potential endocrine etiologies, a laboratory evaluation should include serum electrolytes, fasting blood glucose, TSH, and calcium concentrations. Until the underlying endocrine disturbance is corrected, psychiatric symptoms will not resolve and may worsen with psychotropic agents. When psychiatric disorders appear to result from a medical condition, they are categorized in the fifth edition of the Diagnostic and Statistical Manual (DSM-V) as XX.

THYROID DISORDERS

The thyroid hormones, triiodothyronine (T3) and thyroxine (T4), are produced from the thyroid gland and are necessary for regulation of metabolic activity and protein synthesis. They are formed from a common

precursor thyroglobulin, stored in the thyroid follicles. In extrathyroidal tissues, T4 is converted to T3, the metabolically active hormone. TSH (thyrotropin), produced by the anterior pituitary, promotes the synthesis and release of these hormones and is, in turn, regulated by TRH, produced in the hypothalamus. Low serum levels of thyroid hormones stimulate release of TRH and TSH. The central nervous system (CNS) neurotransmitter dopamine decreases TSH secretion, as do glucocorticoids.

T3 and T4 are tightly protein bound to thyroxine-binding globulin (TBG), so that only 0.3 percent of T3 and 0.02 percent of T4 are free. Total serum T4 levels vary with TBG levels; for example, exposure to estrogens (pregnancy and oral contraceptive) raises TBG and, therefore, serum T4, whereas androgens have the opposite effect. Free T4 index (FT4I), which measures the percentage of free T4 (FT4), produces a more accurate determination of thyroid function. FT4I is elevated in hyperthyroid and low in hypothyroid patients and is an essential test in the assessment of thyroid status. Measures of TSH levels are also necessary in the workup of thyroid disorders. Levels are elevated in thyroidal hypothyroidism and are usually low in pituitary-hypothalamic hypothyroidism. The range of normal levels of TSH is 0.3 10 4.7 mIU/L and free T3 index (FT3I) range from 5 to 11 mIU/L and 75 to 175 mIU/L respectively.

T3 is the main determinant of thyroid hormone that is available to peripheral organs whereas T4 is the main determinant of thyroid hormone that is available to the brain. T3 acts as a neuromodulator by enhancing γ-aminobutyric acid (GABA) release and calcium influx, which subsequently facilitate neurotransmitter release. T3 levels have been shown to predict relapse in unipolar depression, and exogenous T3 supplementation appears to potentiate or accelerate the antidepressant response.

Many, but not all, studies of major depressive disorder have reported an acute activation of the hypothalamic-pituitary-thyroid (HPT) axis at multiple levels, including increased cerebrospinal fluid (CSF) TRH, TRH-stimulated thyrotropin (TSH) blunting, nocturnal TSH blunting, and relative increases (i.e., within the normal range) in iodothyronines, most notably hyperthyroxinemia (i.e., serum FT4, total T4 [TT4], and reverse T3). Neurobiologically, this thyroid activation has been proposed to represent a compensatory mechanism to a pathological process, that is, depression, or the primary pathology itself. Successful treatment of depression has been associated with a decrease in these thyroid indices, again primarily, but not exclusively, T4. Furthermore, subsequent T4 decreases have been greater in responders, in comparison to nonresponders, to a number of treatment interventions for primarily unipolar depression, including carbamazepine, TCA, cognitive psychotherapy, and T3 augmentation of desipramine.

The TRH stimulation test (TRH infusion test) is sometimes used in the workup of thyroid function. After the patient's baseline TSH level is

obtained, TRH is administered intravenously (IV), and TSH levels are then measured at regular intervals. A normal response is a rise in TSH of 5 1-1U/mL or more above the baseline level. In hypothyroid patients, the response is exaggerated, whereas in hyperthyroid patients, it is blunted. With the development in the 1990s of highly sensitive TSH testing, the TRH stimulation test is seldom used clinically. Its use as a diagnostic tool in the psychiatric assessment of depression has similarly waned, as its specificity and sensitivity are poor. For example, although some patients with depression experience an elevation of TSH after TRH infusion, others have a blunted response. Tests of radioactive iodine (RAI) uptake are used to determine the activity of the thyroid gland. A patient receives a small amount of oral RAI, and a scan then determines the amount taken up by the thyroid gland. Disorders of the thyroid gland fall into two general categories: hyperthyroidism or hypothyroidism. Of all the endocrine disorders, disorders of the thyroid are the most common.

Hypothyroidism

Definition and Comparative Nosology and Epidemiology. Hypothyroidism results from inadequate production of thyroid hormone. In overt hypothyroidism, serum TSH levels are elevated and serum T4 levels are low compared to normal levels within general population. Often, but not always, patients will experience signs and symptoms of hypothyroidism. When TSH levels are elevated but T4 levels are normal, subclinical hypothyroidism (SCH) is diagnosed.

The most common cause of hypothyroidism in adults worldwide is iodine deficiency. In countries, such as the United States, where iodine is added to salt, Hashimoto thyroiditis (chronic lymphocytic thyroiditis) is the most common cause. Lithium is another common cause of hypothyroidism: 20 percent of patients develop overt hypothyroidism with a goiter after long term lithium use, and 20 percent develop subclinical hypothyroidism. This risk is elevated in women and in patients with antithyroid antibodies. The longer the exposure to lithium, the greater the risk of developing either overt or subclinical hypothyroidism. Other causes of hypothyroidism are hypopituitarism (due to tumor or postpartum hemorrhage), previous thyroid surgery, or antithyroid drugs.

Hypothyroidism must be distinguished from euthyroid sick syndrome, which is associated with medical illnesses. In this condition, conversion from T4 to T3 is decreased, leading to normal TSH, but lower T3 levels.

Etiology. Antibodies against the thyroid are the most common cause of hypothyroidism.

Lithium directly interferes with thyroid hormone release, leading to hypothyroidism. Thyroid hormones have widespread actions on the central nervous system, including genomic and nongenomic effects that underlie

the changes in mood and cognition seen in hypothyroidism. The symptoms of hypothyroidism may result from alterations in receptor functioning, neurotrophin production or changes in intracellular signaling that occur due to alternations in thyroid hormone levels. For instance, decreased postsynaptic B receptor number and function is associated with low levels of thyroid hormones. Brain synthesis of serotonin, 5-HT, is also reduced in hypothyroidism and the PRL and cortisol responses to the 5-HT agonist fenfluramine are blunted compared to euthyroid controls, suggesting reduced central 5-HT activity. Serum 5-HT levels correlate with serum T3 levels. Thyroid hormone is associated with reduced sensitivity of 5-HT1A autoreceptors in raphe nuclei, and increased 5-H2 sensitivity. Regional cerebral blood flow studies have found decreased regional blood flow in bilateral amygdala, hippocampus, and anterior cingulate cortex in hypothyroid patients, and treatment with thyroid hormone reduced neuropsychiatric complaints and restored regional blood flow, especially in ventral anterior cingulate cortex.

Epidemiology

General Population. Overt hypothyroidism is more common in women than men in the general population, affecting 2 percent women and less than 0.1 percent in men. Subclinical hypothyroidism also predominates in women, occurring in approximately 7.5 percent of women compared to 3 percent of men. The prevalence of hypothyroidism increases with age, and subclinical hypothyroidism is especially common in postmenopausal women, occurring, as studies report 20 percent prevalence rates. The progression of subclinical hypothyroidism to overt hypothyroidism is approximately 5 to 8 percent per year. Risk factors for conversion include higher baseline TSH, female sex, age, presence of antithyroid antibodies, and family history of autoimmune disorders.

Pregnancy. Overt hypothyroidism is estimated to occur in 0.3 to 0.5 percent of pregnancies and subclinical hypothyroidism in 3 to 5 percent. The presence of elevated anti-TPO, anti-TG titers are also common in euthyroid pregnant women, and higher titers are associated with an increased risk of developing postpartum thyroiditis, which occurs in 7.5 percent of women postpartum. Twenty to forty percent of women with this condition develop permanent hypothyroidism. Women with recurrent spontaneous abortion have higher rates of overt and subclinical hypothyroidism.

Psychiatric Populations. In psychiatric patient populations, hypothyroidism is even more common and the gender difference persists. Bipolar disorder patients not on lithium have high prevalence rates of hypothyroidism, 25 percent in women and 12 percent in men.

Hypothyroidism is especially common in rapid cycling bipolar disorder. It is estimated that 20 percent of patients treated with lithium develop hypothyroidism and up to 50 percent may develop a goiter. Females are five times more likely than men to develop hypothyroidism with lithium treatment. Older age and pretreatment presence of thyroid antibodies are also risk factors for hypothyroidism with lithium use. Due to the high prevalence of postpartum thyroiditis in women with postpartum depression, the American Thyroid Association recommends that women with postpartum depression have TSH, free T4, and anti-thyroid peroxidase screening.

Diagnosis and Clinical Features

Physical Symptoms. Physical symptoms associated with overt and subclinical hypothyroidism are diverse, and reflect the widespread influence of the thyroid hormones on metabolic, gastrointestinal, cardiac, neurological, reproductive, and even skin functioning. Symptom onset is usually insidious and classic symptoms include weight gain, constipation, cold intolerance, muscle cramps, dry skin, brittle and thinning hair, especially lateral eyebrows, slowed deep tendon reflexes (DTRs), paresthesias, and menstrual disturbances, including amenorrhea and menorrhagia. Subclinical hypothyroidism has been associated with ischemic heart disease and cardiomegaly. Laboratory abnormalities associated with hypothyroidism include elevated cholesterol, triglycerides, and normochromic, normocytic anemia.

Neuropsychiatric Symptoms. Neuropsychiatric symptoms are classically associated with overt hypothyroidism; however, recently, mood and cognitive symptoms have been found in patients with subclinical hypothyroidism, as well as euthyroid antibody positive patients. The ongoing autoimmune process underlying the majority of cases of hypothyroidism may be associated with nonspecific symptoms of decreased energy and decreased motivation seen in other autoimmune conditions. Depressed mood, apathy, psychomotor retardation, and slowed cognition with memory and concentration complaints occur in overt hypothyroidism. Neuropsychiatric testing confirms the subjective complaints, as hypothyroid patients demonstrate decreased scores on attention, memory —especially verbal memory—and information processing speed. Psychosis, “myxedema madness,” is now rare; the presentation is varied and includes delusions, auditory and visual hallucinations, thought disorders, and paranoia.

Neuropsychiatric changes seen in patients with subclinical hypothyroidism are less pronounced and prevalence rates vary between studies, depending on many conditions, including the sensitivity of tests and the age of participants. In some patients with subclinical

hypothyroidism, mild depressive symptoms are found, as well as subtle deficits in working memory and executive function. Both overt and subclinical hypothyroidisms are associated with poorer antidepressant treatment response in patients with patients with major depression.

Postpartum thyroiditis may present with hypothyroidism alone (43 percent), with hyperthyroidism alone (32 percent), or with an initial phase of hyperthyroidism followed by hypothyroidism (25 percent). The most common symptoms of the hyperthyroid state are anxiety, palpitations, and fatigue; the hypothyroid state is associated with decreased energy, depressed mood, and impaired concentration. Even patients who are euthyroid, but who are thyroid antibody positive, have been found to have higher prevalence of depressive symptoms in some but not all studies.

Ms. A, a 54-year-old woman with no previous psychiatric history experienced depressed mood, lethargy, apathy, weight gain, and problems with memory and concentration over the past year. Initially she attributed her symptoms to menopause, but her symptoms have not improved after six months of hormone replacement therapy. A TSH level was obtained and was elevated at 14 mIU/L. The patient was started on levothyroxine (Levoxyl) 0.05 mg/d and within 8 weeks she experienced an improvement in energy, interest in her activities, and subjective memory. A TSH drawn 8 weeks after initiation of thyroid replacement was normal at 3.3 mIU/L.

Ms. K, a 32-year-old woman with recurrent spontaneous abortion and recurrent major depression presented to the ER at 14 weeks postpartum for major depression with passive suicidal ideation. Ms. K had discontinued her antidepressant in the first trimester of pregnancy and remained euthymic until a few weeks before delivery when she became increasingly anxious. Her pregnancy was complicated by gestational diabetes, and she delivered prematurely at 36 weeks. She restarted her serotonin reuptake inhibitor at her prepregnancy dose within a week of delivery, but her mood continued to deteriorate. She had new onset of tachycardia, and increasing insomnia, agitation, irritability, inability to focus, and complaints of panic. After approximately a month, her symptoms then switched to a severe depression with lethargy and apathy. Lithium 600 mg HS was added for treatment resistant depression but her symptoms did not improve. When she experienced a decrease in milk supply, increasing problems with memory and concentration, feelings that she could not care for her infant, and suicidal ideation, she was voluntarily admitted. Laboratory tests were as follows: TSH 8 mIU/L, free T4 0.5 ng/dL, T3 normal. Anti-TPO and antithyroid antibody titers were elevated. Physical examination revealed a painless goiter and ultrasound hypoechogenicity. Ms. K was diagnosed with postpartum thyroiditis and Ms. K and husband responded well to psychoeducation about autoimmune thyroiditis and its role in her treatment resistance. Synthroid was initiated and lithium was discontinued. Bupropion was also added to her SSRI and, within 4 weeks, Ms. K experienced significant improvement in her mood and cognition. At 6 weeks, Ms. K's TSH and T4 were normal and she continued on Synthroid, with plans for TFTs again in 6 months and a possible trial taper off the Synthroid after a year and thyroid testing 1 to 2 months later.

Hyperthyroidism

Definition and Comparative Nosology. Hyperthyroidism, or thyrotoxicosis, results from overproduction of thyroid hormone by the thyroid gland. The most common cause is Graves disease. Toxic nodular goiter causes another 10 percent of cases among middle-aged and elderly

patients. Other causes include hyperfunctioning solitary thyroid adenomas, thyroiditis, use of exogenous thyroid hormone, TSH-producing pituitary adenoma, pituitary resistance of suppression of TSH secretion by thyroid hormone, thyroid carcinoma, choriocarcinoma, hydatidiform moles, and struma ovarii.

Subclinical hyperthyroidism describes a condition in which circulating thyroid hormone levels are normal, but the TSH response to TRH is blunted or absent. Patients may experience symptoms of hyperthyroidism, such as nervousness, irritability, fatigue, and tachycardia. Subclinical hyperthyroidism may progress to overt hyperthyroidism.

Thyroid storm (thyrotoxic crisis) is a life-threatening syndrome that is usually precipitated by illness or injury but may also develop after withdrawal of antithyroid drugs (ATDs) or after therapy with radioactive iodine (RAI). Symptoms include marked tachycardia, weakness, fever, and altered mental status.

Epidemiology. In the general population, the prevalence of hyperthyroidism is approximately 0.5 percent. Graves disease is the most common cause, accounting for 80 percent of cases. Graves disease occurs four times more frequently in women than men and occurs primarily in the third and fourth decades of life. The prevalence of thyroiditis is estimated to occur in 5 to 9 percent of women in the postpartum period, usually resolving spontaneously by 1 year after delivery.

Etiology. Patients with Graves disease produce thyroid-stimulating antibodies and frequently have a family history of other autoimmune disorders, such as systemic lupus erythematosus and myasthenia gravis. Hyperthyroidism from toxic nodular goiter results from the lack of response of thyroid nodules to normal feedback regulation. In the case of thyroiditis, hyperthyroidism results from the inflammation-induced damage to the thyroid gland, which then spills the stored colloidal thyroid hormones.

Diagnosis and Clinical Features. Signs and symptoms of hyperthyroidism include increased pulse, arrhythmias, elevated blood pressure, fine tremor, heat intolerance, excessive sweating increased appetite, weight loss, palpitations, tachycardia, frequent bowel movements, menstrual irregularities, muscle weakness from catabolism of muscle protein, exophthalmos, lid lag, infrequent blinking, and hyperactive deep tendon reflexes. In cases of thyroiditis, the patient may experience pain and tenderness in the thyroid region. Psychiatric features include nervousness, fatigue, insomnia, mood instability, and dysphoria. Speech may be pressured, and patients may demonstrate a heightened activity level. Neuropsychiatric symptoms include a short attention span, impaired

recent memory, and an exaggerated startle response. In severe cases, there may be visual hallucinations, paranoid ideation, and delirium. Although some symptoms of hyperthyroidism resemble those of a manic episode, or mixed manic episode, an association between hyperthyroidism and mania has rarely been observed.

Pathology and Laboratory Examination. Laboratory examination reveals elevated levels of TT4 and FT4. Levels of total and free T3 are usually elevated but are not necessary to make the diagnosis of hyperthyroidism. TSH level is abnormally low, as its production is suppressed by the elevated thyroid hormones, and the TSH response to TRH is blunted. The uptake of RAI by thyroid hormone is increased, except in cases of hyperthyroidism due to thyroiditis or ingestion of thyroid hormone.

Differential Diagnosis. Panic disorder, generalized anxiety disorder, and social and simple phobias present with symptoms that resemble hyperthyroidism, including tachycardia, diaphoresis, and tremulousness. Other conditions that produce symptoms of autonomic hyperarousal include hypoglycemia, diabetes, dehydration, and pheochromocytoma. Chronic obstructive pulmonary disease may produce anxiety, diaphoresis, dyspnea, and tachycardia. In some cases, hyperthyroidism may produce severe agitation resembling a manic state. Certain medications, for example, appetite suppressants, stimulants, and street drugs, such as cocaine and amphetamines, may also produce these symptoms. Laboratory testing distinguishes between hyperthyroidism and other medical conditions.

Course and Prognosis. Rates of remission of Graves disease depend on the type and length of treatment. Patients treated with radioiodine rarely experience a recurrence of hyperthyroidism. However, permanent hypothyroidism is a common complication of RAI. Patients treated with ATDs for 6 months have approximately a 30 percent rate of remission as compared to a rate of 80 percent in patients treated for 2 years. Patients who undergo surgical ablation have a relapse rate of approximately 10 percent; these patients are also at risk for permanent hypothyroidism. Hyperthyroidism due to thyroid nodular disease usually remits with RAI therapy. Thyroiditis is usually a short-lived condition, lasting only a few weeks to months, and resolving spontaneously. Treatment of thyroiditis includes 'supportive measures' such as B-adrenergic agents for tachycardia, and close follow up of TFTs, since patients with thyroiditis, especially postpartum thyroiditis, are at an increased risk of hypothyroidism.

Treatment. Treatments for Graves disease include ATDs, RAI, and surgical thyroidectomy. The ATDs, propylthiouracil (PTU), and

methimazole (Tapazole) inhibit synthesis of thyroid hormone. PTU additionally inhibits the monodeiodination of T4 to T3. Methimazole has the advantage of single-day dosing. Initial doses of methimazole range from 10 to 40 mg and, for PTU, from 75 to 100 mg three times a day (t.i.d.). The doses should be adjusted as the patient returns to a euthyroid state. Usual maintenance doses of methimazole are 5 to 10 mg per day, and doses of PTU are 50 to 100 mg per day. A rare and serious side effect of these medications is agranulocytosis. Other side effects include pruritus, rash, and fever. Length of treatment varies from 6 months to 2 years. These medications may not produce a therapeutic effect for several weeks because of the long half-life of T4. Some clinicians combine the ATDs with T4 replacement to prevent iatrogenic hypothyroidism.

Although some patients with Graves disease respond to ATDs, most require RAI. A RAI uptake test done before administration of iodine-131 allows for determination of the appropriate dose. Hypothyroidism may be a complication of RAI, and patients may subsequently require thyroid hormone replacement.

Subtotal thyroidectomy is an option for patients who have an allergy to ATDs or are unwilling to undergo RAI therapy. Patients should receive ATDs or potassium iodide or B-adrenergic antagonists, so that they are euthyroid before surgery, to reduce the risk of postoperative thyrotoxic crisis. Subtotal thyroidectomy has several potential complications, such as hypoparathyroidism and injury to the recurrent laryngeal nerve. Patients may have recurrent hyperthyroidism after surgery, if insufficient thyroid tissue is removed. In general, thyroidectomy is used only if other forms of treatment are not possible.

B-blockers, such as propranolol (Inderal), metoprolol (Lopressor), atenolol (Tenormin), or nadolol (Corgard), can provide symptomatic relief while treatment is in progress. Atenolol (50 to 100 mg per day) and nadolol (80 mg per day) have the advantage of once-daily dosing. These medications should be used cautiously in patients with asthma or heart failure. Iodine (given as a solution of potassium iodide or as Lugol's solution) inhibits thyroid hormone. As the duration of its antithyroid action is brief (a few weeks), its use is limited to pretreatment before surgery, management of thyroid storm, and after radioiodine therapy to speed the reduction in thyroid hormone levels. Treatment for thyroid nodular goiter consists of B-adrenergic antagonists and radioiodine. Treatment of thyroiditis consists of a brief course (a few weeks) of B-adrenergic antagonists, as this condition is short lived. For patients with psychotic symptoms, antipsychotic agents can be used. However, certain antipsychotic agents are best avoided—for example, those that are most likely to produce tachycardia (e.g., low potency antipsychotics) or QT interval prolongation. Similarly, TCAs should be used with caution, if at all, in these patients. In general, the psychiatric symptoms resolve with

successful treatment of the hyperthyroidism.

ADRENAL DISORDERS

The adrenal gland produces three major steroid hormones: glucocorticoids, mineralocorticoids, and sex steroids. Of these, the main steroids produced by the adrenal gland are cortisol, aldosterone, testosterone, estrone, and estradiol. The function of glucocorticoids is to promote gluconeogenesis, lipolysis, and glycogenesis, to increase catecholamine activity during stress, and to inhibit inflammatory reactions. Aldosterone's effect is primarily to increase renal sodium reabsorption and potassium excretion. Only a small proportion of circulating sex steroids are produced by the adrenal glands, as their major site of production is the gonads.

Production of cortisol is under control of ACTH from the anterior pituitary, which, in turn, is regulated by CRH from the hypothalamus. Negative cortisol feedback occurs at pituitary, hypothalamic, and suprahypothalamic levels. CRH secretion follows a circadian pattern, subsequently producing circadian patterns of ACTH and cortisol levels, with their levels peaking in early morning and reaching a nadir in the evening. CRH release is also influenced by emotional and physical stress.

Hypercortisolism

Definition and Comparative Nosology. Spontaneous Cushing syndrome results from either excess production of cortisol from the adrenal glands, such as a cortisol producing adenoma, or from excessive secretion of ACTH, which stimulates the adrenal gland to produce cortisol. Cushing disease, the most common form of spontaneous Cushing syndrome, results from excessive pituitary secretion of ACTH, usually from a pituitary adenoma, or rarely ectopically from neuroendocrine tumors. Cushing syndrome is ACTH-dependent if the syndrome results from the effect of ACTH on the adrenal gland, and non–ACTH-dependent in cases of excessive cortisol regardless of ACTH levels, for example, as a result of an adrenal adenoma. The hypothalamic–pituitary–adrenal (HPA) axis responsivity to stressors may be influenced by stress in infancy and even during in utero development. Studies in primates and rodents show that prenatal stress may induce long-lasting effects on the HPA axis regulation: offspring whose mothers were stressed during pregnancy show greater HPA axis response to stressors compared to control offspring.

Epidemiology and Etiology. The prevalence of Cushing syndrome is approximately 40 per one million, and it occurs far more frequently in women, particularly in the third and fourth decades of life. The most common etiology of spontaneous Cushing syndrome is an ACTH-producing pituitary adenoma. Other causes are autonomous ACTH-producing tumor,

such as oat cell carcinoma of the lung, islet-cell pancreatic tumors, or thymomas. An adrenal adenoma or adrenocortical carcinoma and nodular adrenal hyperplasia are ACTH-independent etiologies of elevated levels of cortisol.

The relationship of stressful life events with the onset of Cushing syndrome is controversial, with some studies finding a greater frequency of adverse life event before the development of the disorder, whereas others have found no such association. Similarly, the relationship between depressive symptoms and the onset of Cushing syndrome remains inconclusive. Some studies have reported a high frequency of depression preceding the physical manifestations of the illness, whereas others have found no greater rate of depressive symptoms than noted in control groups.

The etiology of cognitive impairment in hypercortisolemic patients is believed to result from the effect of corticosteroids on the hippocampal formation. Prolonged exposure to elevated corticosteroid levels is associated with hippocampal damage and smaller hippocampal volume. The hippocampus is involved in selective attention, learning, and memory, and reduced hippocampal volume is associated with lower scores on memory tests. In electrophysiological studies, the response to relevant, but not irrelevant, stimuli is reduced after acute cortisol administration. This finding suggests that cortisol impairs selective attention and discrimination of important information. Several researchers have suggested that the cognitive impairments in patients with clinical depression, many of whom have HPA axis dysregulation, may result, in part, from cortisol effects on the hippocampus.

The exogenous use of steroids may also produce psychiatric sequelae. The likelihood of psychiatric side effects increases in direct proportion to prednisone dosage. A dosage of less than 40 mg per day is associated with less than a 2 percent incidence of psychiatric symptoms, but this rate rises to almost 20 percent when the dosage exceeds 80 mg per day.

Diagnosis and Clinical Features. Weight gain is a common nonspecific symptom in hypercortisolemia. Other more unique features of Cushing syndrome include a characteristic moon facies or rounded face from accumulation of adipose tissue around the zygomatic arch, and truncal obesity, as well as the “buffalo hump” appearance from cervicodorsal adipose tissue deposition. The catabolic effect of cortisol on protein produces muscle wasting, slow wound healing, and easy bruising, as well as thinning of the skin, leading to abdominal striae. Bones become osteoporotic, sometimes resulting in pathological fractures and loss of height. Diabetes may occur in some patients as a result of the gluconeogenesis stimulated by the high level of glucocorticoids. Chronic ACTH stimulation of the adrenal glands may lead to excess androgen production, with secondary hirsutism, acne, and amenorrhea.

Hyperpigmentation may also occur, as ACTH is a precursor of a-melanocyte-stimulating hormone, which produces a skin-darkening effect. Although cortisol's mineralocorticoid effects are mild, at high concentrations, it produces sodium retention, potassium loss, metabolic alkalosis, and edema. Hypertension occurs in the majority of patients.

Psychiatric symptoms in Cushing syndrome patients are frequent, diverse, and often severe. Approximately 75 percent of patients report depressed mood. Of these, approximately 60 percent experience moderate or severe depression. The severity of depression does not appear to be influenced by the etiology underlying Cushing syndrome. Depressive symptoms occur more commonly in female patients than male patients with Cushing syndrome. Hypersensitivity to stimuli and emotional lability are common, especially irritability, which is increased in 86 percent of patients and is often an early manifestation of the illness. Anxiety is also prominent and characterized as generalized anxiety, although panic attacks can occur. Appetite is dysregulated and may be either increased or decreased, and the majority of patients experience fatigue. Over 65 percent patients will complain of sleep problems, and early morning awakening is common. Patients with Cushing syndrome have disrupted sleep patterns similar to patients with major depression: lower sleep efficiency, less total sleep time, shortened time to onset of rapid eye movement sleep (REM). Somatic symptoms and elevated neuroticism scores on Eysenck Personality Inventory have also been reported, with significant improvements after reduction of cortisol levels. Social withdrawal may develop as a result of shame regarding one's physical appearance. Paranoia, hallucinations, and depersonalization are estimated to occur in 5 to 15 percent of cases.

Manic and psychotic symptoms occur much less frequently than depression, at a rate of approximately 3 to 8 percent of patients, but rising to as much as 40 percent in patients with adrenal carcinomas and iatrogenic steroids. In cases of iatrogenic hypercortisolemia, the risk of mania and psychosis appears to be dose dependent, with doses over 80 mg per day of prednisone, resulting in 18 percent of patients suffering severe psychiatric symptoms. The psychiatric disturbances in prednisone-treated patients tend to appear within the first few days to 2 weeks of treatment and occur more commonly in women than men.

The withdrawal of steroid can also produce psychiatric disturbances, particularly depression, weakness anorexia, and arthralgia. Other steroid-induced withdrawal symptoms include emotional lability, memory impairment, and delirium. Withdrawal symptoms have been noted to persist for as long as 8 weeks after corticosteroid withdrawal.

Cognitive changes are common in Cushing syndrome, with approximately 83 percent of patients experiencing deficits in concentration and memory. Declarative and verbal memory deficits predominate. The severity of these deficit correlates with plasma cortisol and ACTH levels.

However, severity of cognitive changes does not correlate with severity of depression; hypercortisolemia appears to have independent effects. Cognitive deficits are also common in patients exposed to high doses of exogenous steroid hormones, and steroid-induced reversible dementia syndromes have been reported. Neuropsychiatric studies in healthy volunteers have confirmed that acute high dose steroids may associated with reversible declarative memory, episodic and working memory deficits, while chronic hypercortisolemia is associated with impaired spatial memory, arousal, attention, and executive functioning.

Hypercortisolemia is associated with glucose intolerance in 30 to 60 percent of patients, and diabetes develops in 20 to 50 percent. Dyslipidemia is also common; patients have high triglyceride levels and elevated cholesterol.

Ms. A, a 42-year-old marathon runner, was diagnosed with tendinitis prior to her 10-km race, and began prednisone 60 mg per day. Within a day of starting the medication, she developed insomnia, and sleep decreased from 8 hours to 6 hours per night. Her speech was pressured, and her energy level increased. She reported that even though she had been advised to decrease her activity, she was unable to resist cleaning her house "from top to bottom." She also described increased irritability and anxiety. Within 2 weeks of discontinuing the prednisone, her mental state returned to baseline.

Pathology and Laboratory Examination. In Cushing syndrome, plasma and urine levels of cortisol are elevated, and normal circadian rhythms of cortisol secretion are blunted or absent. Administration of metyrapone (Metopirone), which blocks cortisol synthesis, is normally followed by increased cortisol production. It produces no such increase in cases of adrenal or ectopic origin, whereas, in cases of pituitary origin, postmetyrapone response is augmented. Another test that distinguishes pituitary from adrenal etiologies of Cushing syndrome is the dexamethasone-suppression test (DST). In cases of Cushing from adrenal origin, administration of 1 mg of the synthetic steroid, dexamethasone (Decadron) does not lead to suppression of cortisol. In contrast, it suppresses cortisol in pituitary Cushing syndrome. The overnight DST is considered the procedure of choice for screening for Cushing syndrome. It involves administration of 1 mg dexamethasone at midnight, and a plasma cortisol level is obtained the following morning. Plasma cortisol levels of greater than 200 nmol/L indicate a high likelihood of Cushing syndrome. The overnight 1 mg DST is simple to administer and produces greater specificity and sensitivity than other screening procedures, such as measurement of urinary free cortisol. As plasma cortisol levels vary considerably across the day, an isolated measurement of plasma cortisol does not provide an adequate screening for Cushing syndrome. DST may also show nonsuppression in major depression, alcoholism, pregnancy, or during acute physical or emotional stress.

Differential Diagnosis. Patients with Cushing syndrome commonly have mood changes that are similar to those of patient with a major depressive episode. Both group of patients have a shortened rapid eye movement (REM) latency, elevated levels of cortisol, and abnormal findings on DSTs. The physical manifestations of Cushing syndrome help distinguish between the two diagnoses but may be absent early in the course of the illness. Other distinguishing symptoms of Cushing syndrome are greater irritability and mood lability than in most patients with unipolar depression, and greater tendency to feel best in the morning hours, in contrast to patients with major depression, melancholic subtype, who tend to feel worse at that time.

Treatment and Course. The treatment of choice for pituitary ACTH-producing tumors (Cushing disease) involves surgical resection, transsphenoidal adenotomy, and if this is unsuccessful, pituitary irradiation. Adrenal adenomas and carcinoma are removed surgically, and, in the case of carcinomas, chemotherapy is instituted subsequently. Medications that inhibit cortisol production, such as ketoconazole or metyrapone, are used when patients are awaiting surgery or radiation, or if these first line treatments have been unsuccessful. These drugs are limited by side effects, such as hepatotoxicity (Ketoconazole). Mifepristone (RU486), a progesterone and glucocorticoid receptor antagonist is a new addition to the medical treatment of Cushing Syndrome. It received FDA approval in 2012 for the treatment of hyperglycemia in patients who are not surgical candidates or who have not achieved remission after surgery. Mifepristone has also been studied in patients with psychotic depression, which has been associated with HPA axis dysregulation and abnormalities in glucocorticoid receptors.

After treatment for hypercortisolemia, either with surgery or with medications, most patients experience improvement in hypertension and correction of electrolyte imbalances. The physical stigmata and subtle cognitive deficits may take months or years to resolve. After successful treatment of the Cushing syndrome, most patients experience significant improvement in depressive and anxiety symptoms. For patients undergoing prednisone treatment, lithium and typical (haloperidol) or atypical medications during the course of treatment help prevent the development of manic or psychotic symptoms. Steroid psychosis resolves after discontinuation of the medication. Prednisone should be tapered rather than discontinued abruptly, to allow for the pituitary gland, which has been suppressed, to resume its production of ACTH.

Adrenocortical Insufficiency

Definition and Comparative Nosology. Adrenocortical insufficiency results from inadequate production of three major steroid hormones—

glucocorticoids, mineralocorticoids, and sex steroids—by the adrenal gland. In cases of primary adrenal insufficiency, adrenal hypofunction results from adrenal gland disease (e.g., autoimmune disease, infection, idiopathic atrophy, and metastatic tumor). ACTH and CRH levels rise in response to low levels of adrenal steroids. Secondary adrenal insufficiency results from deficient ACTH secretion because of pituitary disease, and tertiary adrenal insufficiency refers to deficient hypothalamic secretion of cortisol-releasing hormone (CRH). Adrenal aldosterone production tends to be less affected by secondary or tertiary forms of adrenal insufficiency.

Epidemiology. The prevalence of Addison's disease is estimated at 110 cases per 1 million population, with an incidence of 5.6 cases per 1 million per year.

Etiology. Autoimmune destruction of the adrenal cortex produces more than 90 percent of cases of primary adrenal insufficiency disease. Secondary and tertiary adrenal insufficiency result primarily from withdrawal of exogenous corticosteroids or ACTH, as their administration suppresses the HPA axis. Adrenal insufficiency secondary to HPA axis suppression is now the most common cause of adrenal insufficiency as these medications are now so widely used in multiple medical conditions, from asthma to organ transplantation. Less common etiologies include pituitary tumors, trauma, infectious disease (e.g., human immunodeficiency virus [HIV]), infarction, and surgical ablation.

Diagnosis and Clinical Features. Clinical symptoms usually do not become apparent until more than 90 percent of the adrenal cortex is damaged. Due to the insidious onset and generally nonspecific early symptoms, adrenal insufficiency frequently goes undiagnosed for many years and many patients undergo extensive medical tests before a final diagnosis is confirmed. Many patients are initially misdiagnosed as suffering from primary psychiatric or gastrointestinal illnesses.

Adrenal insufficiency can be life threatening, particularly in cases of acute adrenal insufficiency, such as after adrenal hemorrhage from infectious process or use of anticoagulants. Physical signs and symptoms include weakness, hypoglycemia, hyponatremia, hyperkalemia, nausea, diarrhea, fever, and hypotension. Chronic adrenal insufficiency produces more subtle symptoms, including loss of libido, fatigability, salt craving, weight loss, vitiligo, nausea, loss of body hair, muscle cramps. ACTH is metabolized to alpha-melanocyte–stimulating hormone, which promotes melanocyte activity, so hyperpigmentation of the skin may develop as ACTH levels rise.

Psychiatric symptoms of adrenal insufficiency include apathy, irritability, anxiety, and depression, ranging from mild to severe.

Symptoms may develop slowly over months or years. Even patients receiving corticosteroid replacement therapy are at increased risk of developing severe affective disorders. While psychosis is rare in adrenal insufficiency compared to hypercortisolemia, case reports do exist.

A diagnosis of adrenocortical insufficiency is made primarily on the basis of plasma cortisol measures during the ACTH stimulation test. After IV administration of 250 ug cosyntropin, cortisol levels normally peak in 30 to 90 minutes. A normal test is defined as a baseline cortisol level of at least 5 ug/dL, a rise between baseline and stimulated cortisol levels of greater than 7 ug/dL, and a stimulated cortisol level of greater than 18 ug/dL. Low cortisol responses to this test indicate adrenal insufficiency.

Pathology and Laboratory Examination. Common laboratory findings include low serum levels of sodium and elevated levels of potassium and low or normal plasma cortisol levels. The diagnosis of adrenal insufficiency is particularly likely if plasma cortisol levels drawn in the early morning (when cortisol levels should be peaking) are low (less than 3 ug/dL). In cases of primary adrenal insufficiency, ACTH levels are elevated, and cortisol levels do not rise after ACTH stimulation. In cases of adrenal insufficiency resulting from pituitary disease, ACTH levels are low or normal, and the response to CRH is blunted. If the adrenal insufficiency results from hypothalamic disease, the ACTH response to CRH administration is adequate, but cortisol levels may fail to rise because of adrenal atrophy after prolonged ACTH deficiency.

Differential Diagnosis. The symptoms of chronic adrenal insufficiency overlap with symptoms of depression. The symptoms of easy fatigability, muscle cramps, anorexia, and lethargy may also lead to a diagnosis of chronic fatigue syndrome. In cases of acute adrenal insufficiency, such as after abrupt discontinuation of steroid administration, patients may develop a psychotic or delirious syndrome.

Course and Prognosis. Physical and mental symptoms respond quickly to steroid replacement. A patient's prognosis depends largely on the underlying cause of the adrenal dysfunction. In cases of chronic adrenal insufficiency, patients have an unremitting illness requiring lifelong steroid treatment.

Treatment. Acute adrenal insufficiency requires immediate treatment with IV hydrocortisone (Hydrocortone), in addition to fluid replacement with saline solution and potassium supplementation. In cases of primary adrenal insufficiency, a mineralocorticoid (e.g., fludrocortisone [Florinef]) is also required. In cases of chronic adrenal insufficiency, prednisone or hydrocortisone is administered orally as maintenance treatment. Patients

should be encouraged to carry a card or bracelet identifying their diagnosis, medications, dosages, and physician.

Recently, treatment of adrenal insufficiency has focused on prevention of adrenal insufficiency due to steroid withdrawal. Adrenal suppression risk may be increased with longer-acting steroids and evening dosing, but large studies identifying risk factors and confirming tapering protocols are not yet available. Currently, it is recommended that glucocorticoid therapy be tapered, not abruptly discontinued, even if it is topical or inhaled.

DIABETES MELLITUS

Diabetes is a major health problem affecting over 29 million Americans and it is estimated that over 8 million of these people are undiagnosed. It currently ranks as the seventh most common cause of mortality in the United States. Most of the morbidity and mortality of diabetes is caused by its long-term complications. In the general population, the increasing economic (currently exceeding 245 billion dollars annually) and medical burdens of treating diabetes have made this condition a focus of intense public health concern. Furthermore, the socioeconomic cost of diabetes is expected to escalate with the increase in its prevalence due to the aging population, the increase in visceral obesity in American society, a trend toward physical inactivity, and a growing population of susceptible African Americans, Native Americans, and Asian Americans. Furthermore, diabetes is of great concern in the psychiatric population, as recent research has shown that there is an increased prevalence of non-insulin-dependent diabetes mellitus (NIDDM) in psychiatric patients compared to the general population, and that many psychiatric medications increase the risk of development of this disorder.

Diabetes mellitus is a syndrome that includes a heterogeneous group of disorders, all of which are characterized by hyperglycemia and absolute or relative insulin deficiency. Type I diabetes (DMI) accounts for approximately 10 percent of the cases, whereas type II diabetes (DMII) accounts for more than 85 percent.

Type I Diabetes Definition and Comparative Nosology

Type I (insulin-dependent) diabetes, previously known as juvenile-onset diabetes, occurs because the pancreatic Islets of Langerhans beta cells are damaged due to autoimmune attack. Pancreatic islet cell antibodies are present in 85 percent of new-onset DMI patients. Over time, insulin deficiency develops and glucose cannot be mobilized into the body's cells for use as energy, resulting in hyperglycemia. While the damage to insulin-producing cells in type I diabetes occurs over a period of years, the symptoms of type I diabetes may occur over a period of days to weeks and patients may present in hyperglycemic crisis.

There is considerable evidence that genetic and environmental factors are major determinants of type I diabetes. Although most people in the population have a 0.4 percent chance of developing type I diabetes, the risk increases to approximately 2 percent if the mother has diabetes, and 6 percent if the father or siblings have type I diabetes. Concordance rates for DMI reach 50 percent in monozygotic twins. In terms of genetic markers, type I diabetes is primarily determined by genes in the HLA region of chromosome 6. The HLA-DQ locus appears to be the best single marker of susceptibility, particularly among whites. However, recent genome screens have identified at least 15 additional loci that may also contribute to disease risk. Although type I diabetes is likely a polygenic disorder, epidemiological patterns of type I diabetes, including seasonal and temporal changes in incidence, suggest that environmental factors are involved. With the exception of a possible role for viruses and infant nutrition, the environmental factors that initiate or precipitate the onset of type I diabetes have not been established.

Epidemiology and Etiology

Type I DM is the most common form of diabetes in people younger than 20 years of age, but it can occur at any age. The majority of cases occur before 35 years of age, with peak onset during the second decade of life. Type I diabetes affects approximately 700,000 people in the United States and is the most common chronic metabolic disorder to affect children. It is most common in white populations, especially those in Scandinavia, and is rare in people of Asian or African descent.

Type II Diabetes

Definition and Comparative Nosology. Unlike people with DMI, people with type II diabetes produce insulin. DMII is generally preceded by a clinically silent phase of impaired fasting glucose or impaired glucose tolerance. Insulin resistance, defined as a defect in the ability to drive glucose to its main target tissues, is the primary pathophysiological feature in the majority of cases. The pathological mechanisms behind insulin resistance are still not fully understood, despite intensive research. A variety of cellular and molecular defects have been implicated in the condition. These include dysfunctional insulin receptors, aberrant receptor signaling pathways, and abnormalities in glucose transport or glucose metabolism. Hyperinsulinemic individuals maintain normal or near-normal blood glucose levels, despite a continuing decline in insulin sensitivity. In time, however, as insulin resistance gradually increases, the pancreas is less able to compensate by increasing insulin secretion, and progressive hyperglycemia develops. Not all people with insulin resistance develop diabetes. Beta-cell function also plays an essential role. Whether

beta-cell dysfunction is a primary or secondary factor remains unclear. Impaired beta-cell function may result from exhaustion—long-term exposure to insulin resistance—and prolonged, increased demand for insulin secretion that eventually exceeds beta cell capacity. Minimal elevation of plasma glucose during deterioration of glucose tolerance may also impair beta-cell function–glucose toxicity. Some people may be able to maintain high levels of sustained insulin secretion, whereas others may have acquired or inherited an abnormality that makes beta cells susceptible to exhaustion of glucose toxicity. Such defects could include reduced islet-cell mass, limited replicative capacity, and more rapid age-related functional or structural abnormalities.

Epidemiology and Etiology. Type II diabetes affects approximately 8 percent of the US population. It is most common in people older than 40 years of age who are overweight, although the incidence of DMII among young obese individuals has been rising. Obesity appears to be the strongest risk factor for development of DMII, especially among women. Among the psychiatric population, obesity may be iatrogenic in nature, specifically with antipsychotic use. Lack of exercise, smoking, and a poor diet are also risk factors for development of type II diabetes. Type II diabetes is most prevalent among African Americans, Latinos, Asian-Pacific Islanders, and Native Americans.

Diagnosis, Clinical Features, and Laboratory Examination. Classic symptoms of DM I and II include polyphagia, polyuria, and polydipsia. Other symptoms include blurred vision, fatigue, weight loss, poor wound healing, dry mouth, dry or itchy skin, or recurrent infections (e.g., vaginal yeast infections, groin rash, and external ear infections). Patients with DM I or II may also present with acute metabolic compensation and dehydration, nausea, vomiting, abdominal pain, confusion, or coma. Individuals with DMI usually seek medical help when these symptoms arise, but they often experience weight loss developing over several days that is associated with the onset of their diabetes. The onset of these first symptoms may be fairly abrupt or more gradual. Many patients with DMII remain asymptomatic for years with subtle signs and symptoms, such as neuropathy. Three methods of laboratory assessment may be used to diagnose diabetes: a fasting plasma glucose test and a value of 126 mg/dL or greater; a nonfasting plasma glucose value of 200 mg/dL or greater in people with symptoms of diabetes; an abnormal oral glucose tolerance test, with a 2-hour glucose value of 200 mg/dL or greater. Recently, the hemoglobin A1c test (HgA1c) has become a standard method of diagnosing DM I and II as well as monitoring treatment over time. The HgA1c reflects glycemic control over the preceding 2 to 3 months; two values over 6.5 percent are diagnostic for diabetes.

Diabetes and Psychiatric Illness. Eating disorders (ED) occur fairly often in adolescents with type I diabetes and present challenges in diagnosis and management. A meta-analysis of recent studies, reports that ED occur in 8.2 percent of DMI patients compared to 2.8 percent of controls. Insulin omission is a unique purging behavior found in up to 40 percent of DMI patients; risk factors include female sex and higher BMI, as insulin is omitted in order to lose weight. Eating disorders should be considered in patients with poor glycemic control or multiple episodes of diabetic ketoacidosis.

DMII and major depression appear to have a bidirectional relationship. While the risk estimates vary depending upon populations studied, measures, and diagnostic criteria, studies consistently report that diabetes is associated with an increased risk for depression, and depression is associated with an increased risk for insulin resistance and DMII. Shared lifestyle factors and histories, as well as biological mechanisms are believed to underlie this complex relationship. Insulin resistance is increasingly recognized as an important regulator of mood and cognition. Recent reports show a significantly lower insulin sensitivity index in depressed patients than in most control subjects. Insulin resistance is associated with elevations in cytokines, and it is a proinflammatory state. Inflammation has been linked to the development of symptoms of depression in both animal models and humans, and patients with major depression with elevated inflammatory markers have been found to be less responsive to antidepressants and lithium. Insulin sensitizing agents, such rosiglitazone and pioglitazone, have been found to improve mood in some nondiabetic patients with treatment-resistant depression. Depression has been demonstrated to be a major factor in causing hospital admissions and death in persons with diabetes mellitus. Furthermore, antidepressant therapy with selective 5-HT reuptake has been reported to ameliorate insulin sensitivity in type II diabetes patients.

Among patients with bipolar disorder and schizophrenia, the prevalence of DMII has been reported to be two to three times that of the general population, and it is expected that these rates underestimate the true prevalence of diabetes in this population since medical problems in people with severe mental illness often go undiagnosed. As in major depression, the high comorbidity is likely multifactorial and complex. For instance, there may be shared genetic mechanisms, and several genes have been identified that may play a role in both conditions, such as the gene that encodes glycogen synthetase kinase 3 or dopamine 2 receptor. Shared lifestyle such as sedentary lifestyle and diets rich in saturated fat and low in vitamin D may contribute as well. The use of psychotropic medication, in particular the atypical antipsychotics (SGAs) increases the risk of diabetes in this population. Olanzapine and clozapine have been associated with highest risk, while aripiprazole, ziprasidone, and lurasidone appear to have

lower risks. While the weight gain and risk for DMII differs among the SGAs, the American Diabetes Society, the American Psychiatric Association, and the International Society for Bipolar Disorder all have published guidelines that call for a thorough screening of all patients prior to initiation of SGAs that includes weight, blood pressure, fasting glucose and lipids, and documentation of family history of diabetes and cardiovascular disease. Consensus guidelines recommend metabolic monitoring at regular intervals in order to intervene early in development of metabolic syndrome.

Insulin resistance has recently been found to be associated with greater mood instability in bipolar patients, while the insulin-sensitizing agent pioglitazone has been shown to improve depression.

Treatment. Type I diabetic patients require insulin treatment. Some people with type II diabetes can manage it by controlling their weight, watching their diet, and exercising regularly. Others may also need to take an oral glucose-lowering medication (e.g., sulfonylureas, meglitinides, biguanides, thiazolidinediones, and alpha-glucosidase inhibitors). Generally, one agent is chosen and titrated to maximum dose, unless side effects develop. If one agent is ineffective in reaching hemoglobin A1c goal, a second agent is added. Some patients require therapy with three oral hypoglycemic. Insulin treatment is indicated for type II diabetics when lifestyle changes and oral medications fail. Insulin can be combined with oral agents.

At the time of diagnosis, 20 percent of people with type II diabetes already have other risk factors that contribute to the high incidence of chronic complications, which include blindness, renal failure, neuropathy, peripheral vascular disease, stroke, and myocardial infarction. It is therefore recommended that patients with type II diabetes undergo a complete assessment of the adequacy of their metabolic control, as well as risk factors for vascular complication and extensive education on self-management. Because 80 percent of patients with type II diabetes are overweight, special emphasis should be placed on nutritional education and exercise. Because of the complications associated with diabetes, frequent follow-up visits should be scheduled, depending on the severity of complications that a patient may be experiencing. At a minimum, diabetic patients should have an annual complete physical examination, including a foot examination and examination of any systems found to be abnormal previously. A dilated eye examination should be performed annually by an ophthalmologist. Hemoglobin A1c should be checked at least every 6 months in patients who have met glycemic goals and quarterly in all others. A fasting lipid profile and testing for microalbuminuria should be performed annually.

Polycystic Ovary Syndrome

Polycystic Ovary Syndrome Definition and Clinical Features. PCOS is one of the most widespread endocrine disorders in women. Based on the Rotterdam criteria, which was adopted by the National Institute of Health in 2012, PCOS occurs in up to 15 percent of women in the general population of reproductive-age women. Rates reach 30 to 75 percent when based on women presenting for treatment of menstrual irregularities or hirsutism. After exclusion of other causes of hyperandrogenism (congenital hyperplasia, Cushing Syndrome, androgen secreting tumors), PCOS is diagnosed if 2 out of the 3 conditions or symptoms are present: oligoovulation and/or anovulation, clinical or biochemical signs of hyperandrogenism, ultrasound documented polycystic ovaries. Notably, normal-appearing ovaries no longer rule out this diagnosis if the other two criteria are met, as PCOS is now recognized as primarily an endocrine disorder. Clinical features of PCOS include hirsutism, coarse hair distributed in a male pattern such as the chin, upper lip, and periareolar breast area, as well male pattern alopecia and acne. Irregular menstrual cycles and infertility may be the presenting symptom in some women. BMIs over 25 kg/m (overweight) and 30 kg/m (obese) are common in PCOS, affecting 20 to 80 percent of patients. However, some women with PCOS are normal or even underweight, and weight is not a diagnostic criterion. Sleep apnea is also increased in PCOS.

Etiology. The fundamental defect in PCOS remains unclear, but theories implicate dysfunction of the hypothalamic–pituitary–gonadal (HPG) axis. GnRH appears to be insensitive to ovarian steroids, leading to a failure in normal negative feedback mechanism. A corresponding increased LH pulse amplitude and frequency results in increased thecal cell androgen synthesis. There is also a relatively impaired FSH synthesis, which leads to impaired follicular development in the granulosa cells. One of the hallmark laboratory findings in PCOS is an increased LH:FSH ratio. Normally, FSH regulates aromatase activity, which leads to production of estrogen from androgen precursors in the ovary, but in PCOS, the impaired FSH secretion leads to less estrogen production and more androgen production.

Insulin resistance is also a critical pathogenic mechanism in PCOS and women with PCOS have high rates of glucose intolerance regardless of weight. Hyperinsulinemia is a compensatory increase in insulin secretion due to the peripheral insulin resistance. Approximately 50 percent of women with PCOS are obese, which may contribute to insulin resistance and other symptoms of PCOS. A strong association between insulin resistance and obesity has been recognized for years, yet the mechanism by which insulin resistance may provoke obesity, and vice versa, are still

unclear. Obesity itself increases the risk of insulin resistance, although insulin resistance can also occur in the absence of obesity. Insulin sensitivity is decreased by 35 to 40 percent in PCOS patients, independent of obesity. Furthermore, increased androgen levels are as associated with insulin resistance and obesity. Insulin is known to increase ovarian androgen production in women with PCOS via stimulation of the thecal cell of the ovaries. Furthermore, increased insulin levels lead to decreased sex hormone–binding globulin production by the liver, which leads to more free-circulating testosterone.

Clinical and Laboratory Examination. Central obesity is common, but many women with PCOS are normal weight.

On ultrasonography the ovaries are commonly enlarged, with multiple cystic follicles. Normal-appearing ovaries, however, do not rule out this condition. Because of the absence of ovulation, estrogen is unopposed by progesterone, and vaginal ultrasound may reveal endometrial hyperplasia and carcinoma.

LH levels are usually elevated, whereas FSH levels are low or normal and androstenedione, testosterone, and estrone levels are elevated. Fasting blood sugar and hemoglobin A1c may be elevated later in the course of the illness, due to insulin resistance. The prevalence of undiagnosed diabetes mellitus type II is 7 to 10 times higher in women with PCOS than in controls of similar age.

Differential Diagnosis. Rapid onset of marked virilization, such as low voice or clitoromegaly, should raise the suspicion of an adrenal or ovarian androgen secreting tumor, and serum testosterone and DHEAS are significantly increased. Hyperprolactinemia and thyroid disorders are associated with menstrual irregularities, and should always be ruled out through serum prolactin levels and thyroid function tests. Congenital adrenal hyperplasia caused by 21-hydroxylase deficiency is a rare cause of virilization and menstrual abnormality.

PCOS and Psychiatric Illness. Women with PCOS have an increased prevalence of major depression compared to age matched controls, and a meta-analysis controlling for BMI found rates as high as four times greater than controls. The etiology of this higher rate of depression is controversial, since studies are conflicting and total and free testosterone and androstenedione levels do not consistently differ between women with PCOS with depression compared to PCOS women without depression. Insulin resistance may be a common pathological link between depression and PCOS (Fig. 27.7–1). Insulin resistance has been found to be higher in PCOS women with higher depression scores compared to PCOS women with lower scores and the insulin-sensitizing agent metformin has been

shown to decrease both depression and insulin resistance in PCOS patients. However, improvement in insulin resistance does not appear to be the only mechanism by which insulin-sensitizing drugs improve mood in PCOS patients. Treatment with pioglitazone, a thiazolidinedione (TZD) insulin sensitizer with a different mechanism of action than metformin, was associated with a greater reduction in Hamilton depression scores than metformin in patients with MDD and PCOS and change of depression scores but homeostatic model of insulin resistance, HOMA-R did not differ between the groups.

Anxiety symptoms are also higher in women with PCOS compared to controls, but to date few studies have examined prevalence of anxiety disorders using standardized measures. Psychological reactions to the physical changes associated with PCOS, including virilization and obesity, have been proposed as contributing to depressive and anxiety symptoms.

Bipolar disorder has been associated with PCOS, but controversy remains as to whether the higher rates of PCOS are associated with mood-stabilizing agents or underlying comorbidity and/or shared pathologic mechanisms. For instance, genetic predisposition likely plays a role in the development of insulin resistance in women with bipolar disorder who themselves develop insulin resistance with exposure to mood stabilizers, as women who developed menstrual cycle irregularities had significantly worse metabolic biomarkers than healthy control women with family history of diabetes. Menstrual abnormalities have been found in some bipolar disorder patients to precede initiation of medication, further suggesting an underlying predisposition to abnormal menstrual cycles in this population. The antiepileptic drug valproic acid is associated with an increased risk of PCOS both in patients with epilepsy and in patients with bipolar disorder. Rates of new onset PCOS vary across studies; in the largest study to date of bipolar disorder patients, 10.5 percent of women on valproic acid developed PCOS compared to 1.4 percent of women on other mood stabilizers.

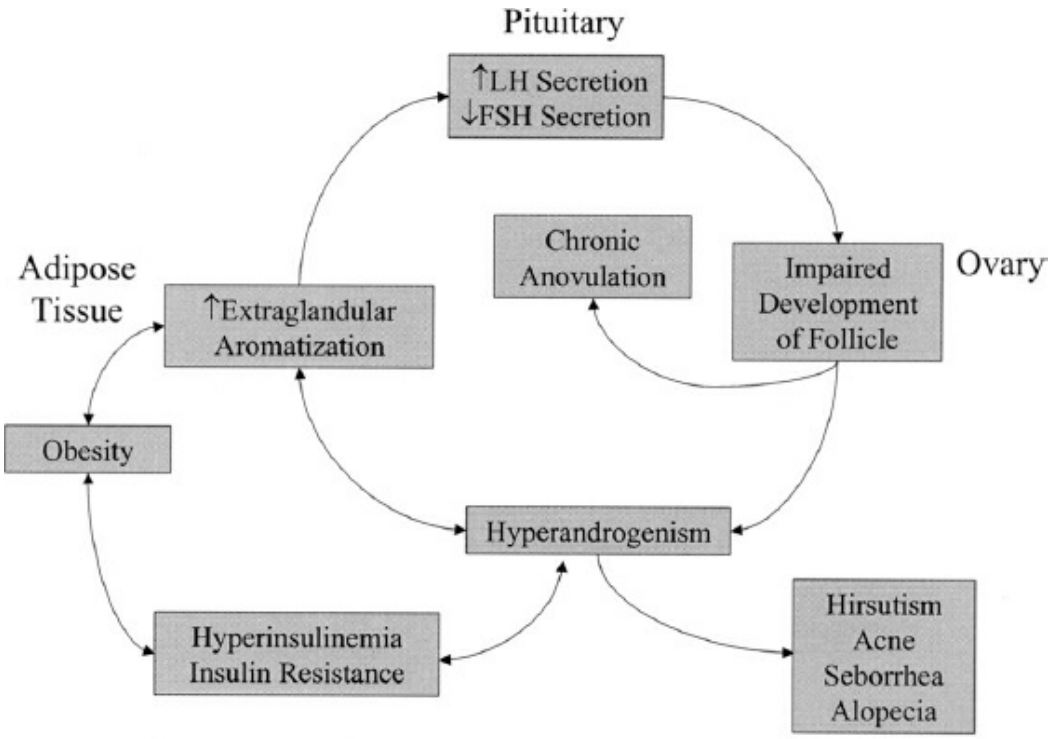

FIGURE 27.7–1. Proposed mechanism for the perpetuation of peripheral endocrine disorders. FSH, follicle-stimulating hormone; LH, luteinizing hormone.

Ms. KB was a 30-year-old white woman diagnosed with PCOS who had a history of treatment-resistant major depression, most recently treated with venlafaxine (Effexor XR). Her serum glucose and insulin levels were 83 mg/dL and 19.0 uIU/mL, respectively. Homeostatic model insulin resistance (HOMA) measured her insulin resistance at 3.9. HOMA values greater than 2.3 are indicative of insulin resistance. Ms. KB subsequently began metformin (Glucophage) (850 mg per day, titrated to 2.550 mg per day) and spironolactone (Aldactone) (100 mg per day). After 2 weeks of these medications, Ms. KB's mood had improved, and, at a follow-up visit 3 months later, she reported feeling euthymic, had lost approximately 10 lb in 3 months, and had begun menstruating approximately every 33 days. She continued to experience low mood, decreased energy, and irritability during the week before menses; however, these symptoms resolved with the onset of menses. Because Ms. KB was not significantly impaired by these symptoms, no antidepressant medications were prescribed. Laboratory tests showed that her serum glucose level was 89 mg/dL, and her serum insulin level at this time was 15.0 uIU/mL. Her insulin resistance, as measured by HOMA, had decreased to 2.

Treatment. Insulin-sensitizing agents are used in the treatment of PCOS. Metformin is a medication that inhibits hepatic glucose production and increases peripheral insulin sensitivity, although it does not modify pancreatic insulin secretion. The decrease in insulin resistance may be achieved via decreasing gut absorption of glucose, improving glucose uptake by tissues, or increasing the number of insulin receptors, or a combination of these. Metformin is widely used to treat insulin resistance in patients with type II diabetes mellitus. In treating PCOS, metformin has been shown to decrease insulin resistance, acne, hirsutism, and total and bioavailable testosterone. A significant decrease in body mass index and waist to hip ratio, as a result of treatment with metformin, has also been

observed, as well a restoration of ovulation.

Several agents are used to treat the hyperandrogenism associated with PCOS. Low-dose oral contraceptives (containing estrogen/progesterone) are frequently first line treatments for acne and hirsutism. They decrease LH secretion, which leads to decreased thecal cell androgen synthesis. Oral contraceptives are contraindicated in some women because of increased risk of deep venous thrombosis, so some women are prescribed spironolactone—an antimineralocorticoid, diuretic, and antiandrogen. It was originally used to treat hypertension and has also been shown to be a weak inhibitor of testosterone biosynthesis. Spironolactone is an effective medication to reduce the hirsutism often seen in PCOS women. It has also been shown to decrease insulin resistance and fasting insulin levels in PCOS patients, and there has been an observed decline in testosterone levels. Other treatments include antiestrogens, such as clomiphene citrate (Clomid), which is first line treatment for PCOS women who desire pregnancy, and laparoscopic surgery. Weight reduction is strongly recommended in the over 50 percent of PCOS women who are obese. Hirsutism may respond to spironolactone, an androgen receptor antagonist, but it may be difficult to tolerate, as it may produce mood changes, decreased libido, and fatigue.

HYPERPROLACTINEMIA

Definition and Comparative Nosology

PRL is produced by the anterior pituitary and functions to stimulate milk production from the breast and to modulate maternal behavior. Its production is inhibited by dopamine, also known as PJF, which is produced by the tuberoinfundibular neurons in the arcuate nucleus of the hypothalamus. Normal levels in women vary from 5 to 25 ng/mL and in men from 5 to 15 ng/mL. Levels fluctuate during the day and peak during sleep. Exercise and emotional stress can elevate PRL level.

Epidemiology and Etiology

Hyperprolactinemia occurs most frequently as a result of PRL-secreting pituitary adenomas (prolactinomas) and idiopathic hyperprolactinemia, two conditions that occur more commonly in women. Typical neuroleptic medications can raise PRL level as much as 20-fold and do so in a dose-dependent fashion, with PRL level reaching plateau. While all typical neuroleptics appear equally likely to raise PRL levels, Haloperidol, because of it strong D2 receptor blockade, appears to be especially associated. The atypical antipsychotics also increase prolactin levels but, in general, less frequently than the typical antipsychotics, and the rise is much less, except in the case of risperidone. The rate of hyperprolactinemia in patients on risperidone is very high, 70 to 100 percent in studies. Quetiapine and

olanzapine are less likely to raise prolactin levels, and studies report 10 to 40 percent, while lurasidone and asenapine have not been found to raise prolactin levels. Studies have shown that switching patients with risperidone-induced prolactinemia to aripiprazole, a partial D2 agonist, results in decrease in prolactin levels. Clinicians, especially those who are working with infertility patients, should be aware of the rarer hyperprolactinemia associated with antidepressants. All the serotonin reuptake inhibitors (SSRIs) have been associated in case reports. Only one study has been completed, which reported that in patients with normal prolactin levels, 4.5 percent of men and 22 percent of women developed hyperprolactinemia after 12 weeks of fluoxetine treatment. Other medications that may elevate PRL levels include oral contraceptives, estrogens, TCAs, metoclopramide, verapamil, and propranolol. Hypothyroidism results in elevated levels of PRL since TRH stimulates PRL release. Acromegaly is also associated with hyperprolactinemia as a result of the lactogenic properties of GH. Hypothalamic disease, such as craniopharyngioma, is also associated with hyperprolactinemia, as the hypothalamus exerts an inhibitory influence on PRL secretion. Physiological hyperprolactinemia occurs in pregnant and breast-feeding women; nipple stimulation also produces an elevation of PRL levels.

Traumatic childhood experiences, such a separation from parents or living with an alcoholic father, have been reported to produce a greater predisposition to hyperprolactinemia. Stressful life events are also associated with galactorrhea, even in the absence of elevated PRL levels.

Diagnosis, Clinical Features, and Laboratory Examination

Patients with elevated PRL levels may be asymptomatic, although, as PRL levels reach 60 to 100 ng/mL, symptoms usually develop. Women may experience amenorrhea as PRL suppresses normal menstrual cycling and men may develop gynecomastia. Women and men may experience low libido decreased fertility, and galactorrhea (inappropriate lactation), although the latter occurs primarily in women. Headache and visual impairment may result in cases of pituitary adenoma. Women with pathological hyperprolactinemia have been reported to experience elevated levels of depression, anxiety, and hostility compared to controls, which lessen with normalization of PRL levels.

Treatment

A laboratory workup is the initial step in treatment and should include thyroid function tests, pregnancy test, and a chemistry panel. A brain imaging study, either an MRI with gadolinium or CT with contrast, should be obtained. Bromocriptine (Parlodel), a dopamine agonist, reduces the synthesis and secretion of PRL. In approximately 80 percent of cases,

bromocriptine at dosages of 2.5 to 5 mg twice a day (b.i.d.) or t.i.d. is effective treatment. However, many patients cannot tolerate the side effects of bromocriptine, which can include severe nausea and dizziness. Alternative medications include quinagolide (Norprolac), pergolide (Permax), and cabergoline (Dostinex). Medication use must continue indefinitely, as medication discontinuation usually leads to a relapse of hyperprolactinemia. For patients who are refractory to medication treatment, radiotherapy is another treatment strategy. Surgical removal through transsphenoidal surgery is used far less often now than before the 1990s because of the high rate of recurrence of prolactinomas.

REFERENCES

Altshuler L, Bauer M, Frye M, et al. Does thyroid supplementation accelerate antidepressant response? A review and meta-analysis of the literature. *Am J Psychiatry*. 2001;158:1617.

Amial A, Joffe H, Nachtigall LB. Psychotropic-induced hyperprolactinemia: a clinical review. *Psychosomatics*. 2014;55:29–36.

Aronson R, Offman HJ, Joffe RT, Naylor D. Triiodothyronine augmentation in the treatment of refractory depression. A meta-analysis. *Arch Gen Psychiatry*. 1996;53:842.

Bauer M, London E, Silverman D, Rasgon NL, Kiesslinger J, Whybrow P. Thyroid, brain, and mood: novel insights from molecular research and functional brain imaging studies. *Pharmacopsychology*. 2003;36:S215–S221.

Bauer M, Glenn T, Pfennig A, Whybrow PC. Gender differences in thyroid system function: relevance to bipolar disorder and its treatment. *Bipolar Disorders*. 2014;16:58–71.

Bauer M, Silverman DH, Schlagenhauf F, et al. Brain glucose metabolism in hypothyroidism: a positron emission tomography study before and after thyroid replacement therapy. *J Clin Endocrinol Metab*. 2009;94:2922–2929.

Calkin CV, Ruzickova M, Uher R, et al. Insulin resistance and outcome in bipolar disorder. *Br J Psychiatry*. 2015;206(1):52–57.

Davis JD, Tremont G. Neuropsychiatric aspects of hypothyroidism and treatment reversibility. *Minerva Endocrinol*. 2007;32:49–65.

Degner D, Haust M, Meller J, Ruther E, Reulbach U. Association between autoimmune thyroiditis and depressive disorder in psychiatric outpatients. *Eur Arch Psychiatry Clin Neurosci*. 2015;265:67–72.

Dinan T. Novel approaches to the treatment of depression by modulating the hypothalamic-pituitary-adrenal axis. *Psydropharmacol*. 2001;16:89.

Dokras A. Mood and anxiety disorders in PCOS. *Steroids*. 2012;77:338–341.

Dorman JS, Bunker CH. HLA-DQ locus of the human leukocyte antigen complex and type I diabetes mellitus: A HuGE review. *Epidemiol Rev*. 2000;22:218.

Horacek J, Kuzmiakova M, Hosch C, Alldel M, Bahbonh R. The relationship between central serotonergic nativity and insulin sensitivity in healthy volunteers. *Psychoneuroendocrinology*. 1999;24(8):785.

Kashani L, Omidvar T, Farazmand B, et al. Does pioglitazone improve depression through insulin-sensitization? Results of a randomized double-blind metformin controlled trial in

patients with polycystic ovarian syndrome and comorbid depression. *Psychoneuroendocrinology*. 2013;38:767–776.

Kornegay CJ, Vasilakis-Scaramozza C, Jick H. Incident diabetes associated with antipsychotic use in the United Kingdom general practice research database. *J Clin Psychiatry*. 2002;63:758.

McElroy L, Frye MA, Suppes T, et al. Correlates of overweight and obesity in 644 patients with bipolar disorder. *J Clin Psychiatry*. 2002;63:207.

McIntyre R, Mancini DA, McCann S, Srinivasan J, Kennedy SH. Valproate, bipolar disorder and polycystic ovarian syndrome. *Bipolar Disord*. 2003;5:28.

Messinas IE, Messini C, Anifandis M, Dagoupolous K. Polycystic ovaries and obesity. Best practice and research. *Clin Obstet and Gynecology*. 2015;29(4):479–488.

O'Donovan C, Kusumakar V, Graves GR, Bird DC. Menstrual abnormalities and polycystic ovary syndrome in women taking valproate for bipolar mood disorder. *J Clin Psychiatry*. 2002;63:322.

Oomen HA, Schipperjn AJ, Drexhage HA. The prevalence of affective disorder and in particular of a rapid cycling of bipolar disorder in patients with abnormal thyroid function tests. *Clin Endocrinol (Oxf)*. 1996;45(2):215–223.

Ozerdem A, Tunca C, Hidiroglu C, Ergor G. Female vulnerability for thyroid function abnormality in bipolar disorder: role of lithium treatment. *Bipolar Disorders*. 2014;16:72–82.

Pinhas-Hamiel O, Hamiel U, Levy-Shraga Y. Eating disorders in adolescents with type 1 diabetes: challenges in diagnosis and treatment. *World J Diabetes*. 2015;6(3):519–526.

Rasgon NL, Altshuler LL, Gudeman D, et al. Medication status and PCO syndrome in women with bipolar disorder: a preliminary report. *J Clin Psychiatry*. 2000;61(3):173.

Rasgon NL, Bauer M, Glenn I, Elman S, Whybrow PC. Menstrual cycle related mood changes in women with bipolar disorder. *Bipolar Disord*. 2003;5(1):48–52.

Rasgon NL, Carter MS, Elman S, Bauer M, Love M, Korenman SG. Common treatment of polycystic ovarian syndrome and major depressive disorder: case report and review. *Curr Drug Targets Immune Endocr Metabol Disord*. 2002;2(1):97–102.

Rasgon NL, Kenna HA, Wiliams KE, Powers B, Wroolie T, Schatzberg AF. Rosiglitazone add-on in treatment of depressed patients with insulin resistance: a pilot study. *Scientific World Journal*. 2010;10:321–328.

Reynolds-May MF, Kenna HA, Marsh W, et al. Evaluation of reproductive function in women treated for bipolar disorder compared to healthy controls. *Bipolar Disord*. 2014;16(1):37.

Samuels MH. Psychiatric and cognitive manifestations of hypothyroidism. *Curr Opin Endocrinol Diabetes Obes*. 2014;21(5):377–383.

Schwartz SA, Weissberg-Benchell J, Perlmuter LC. Personal control and disordered eating in female adolescents with type I diabetes. *Diabetes Care*. 2002;25:1987.

Sobrinho LG. The psychogenic effects of prolactin. *Acta Endocrinol*. 1993;129(Suppl 1):38–40.

Starkman MN. Neuropsychiatric findings in Cushing syndrome and exogenous glucocorticoid administration. *Endocrinol Metab Clin North Am*. 2013;42(3):477–488.

Thomsen AF, Kvist TK, Andersen PK, Kessing LV. The risk of affective disorders in patients with adrenocortical insufficiency. *Psychoneuroendocrinology*. 2006;31:614–622.

Trikudanathan S. Polycystic ovarian syndrome: *Med Clin North Am*. 2015;99(1):221–235.

Wartofsky L. Update in endocrinology. *Ann Med*. 2001;135:601.

Young K, Preskorn SH. Neuroscientific basis of corticosteroid induced changes in human cognitive and emotional processing: implications for affective illness. *J Psychiatr Pract*. 2012;19(4):309–314.

▲ 27.8 Psycho-Oncology

SEAN P. HEFFERNAN, M.D., YESNE ALICI, M.D., WILLIAM S. BREITBART, M.D., AND MARGUERITE S. LEDERBERG, M.D.

Psycho-oncology, or the study of the psychological aspects of cancer along the continuum from prevention to cure, has emerged slowly since 1975 as a subspecialty within oncology as well as psychiatry and psychosomatic medicine. Psycho-oncology addresses the psychological response to cancer of patients, their families, and health care providers. It also addresses the psychological, behavioral, and social factors that influence cancer risk, detection, and survival. Psycho-oncology training programs, clinical researchers, and practitioners exist throughout the world and are represented by over 50 national organizations as well as an international organization. The American Psychosocial Oncology Society (APOS) is the multidisciplinary national organization representing the subspecialty of psycho-oncology in the United States. The International Psycho-Oncology Society (IPOS) is the international organization that was formed in 1984 in recognition of the need for communication among professionals with interests in psycho-oncology worldwide. Most cancer centers have an identified member or members responsible for providing psychosocial support. Clinicians and researchers from psychiatry, psychology, social work, nursing, and clergy constitute the members of multidisciplinary psycho-oncology teams. These teams provide consultation to patients and their families and liaison services to oncology staff, and also collaborate with investigators studying the impact of therapeutic interventions on quality of life and patient-reported outcomes. Clinical practice guidelines have been established based on evidence from existing research. The Institute of Medicine (IOM) published the first policy statement regarding psycho-oncology in 2008. The IOM reported that to achieve quality cancer care, routine oncology practice must integrate psycho-oncology. This was further reinforced by the American College of Surgeons Commission on Cancer. The Commission on Cancer now requires accredited cancer centers to have a psychosocial distress screening program in place. Communication skills training programs have been developed by psycho-oncologists for oncology staff, as effective communication impacts the care of patients with all cancer types, at all stages, and at each step of treatment. The amount and variety of research in psycho-oncology grows apace. Researchers are more actively exploring social, psychological, and behavioral contributions to cancer prevention, early detection, and survival. Research examining the mechanism of cytokines in producing “cytokine-induced sickness behavior” may provide a biological basis and novel treatment options for common

symptoms of fatigue, depression, anxiety, and cognitive changes in cancer patients. Several new evidence-based psychotherapies applicable for different stages of cancer have been developed such as meaning-centered psychotherapy (MCP) for advanced cancer patients.

This section describes the history of psycho-oncology, the psychosocial challenges experienced by cancer patients, the common psychiatric disorders in cancer patients, the psychiatric syndromes specific to tumor site and cancer treatment modality, the therapeutic interventions available, pediatric and geriatric psycho-oncology, issues specific to cancer survivors, patients at genetic risk for cancer, and psychiatric aspects of cancer screening.

HISTORICAL BACKGROUND

For centuries, physicians were reluctant to discuss a diagnosis of cancer with patients and their families, as cancer represented inevitable death due to lack of effective treatments and to reveal a diagnosis of cancer was regarded as cruel and destructive. With the advent of anesthesia and antisepsis, curative surgical resection of early-stage tumors became possible. The American Cancer Society (ACS) was formed in 1913 to educate the public of the warning signs of cancer and to fight fatalism that interfered with early presentation for treatment. In 1937 the National Cancer Institute was established to seek etiologies and treatments for cancer. In the 1950s the addition of chemotherapy to the combined treatment modalities resulted in cure of several childhood tumors. By the early 1970s, with improvements in survival, oncologists became more comfortable in discussing cancer diagnoses with their patients, and the patients' reluctance to identify themselves as cancer patients diminished. Alongside, hospice programs were developed to improve pain management and palliative care with increasing interest in delivering the best supportive care to patients at the end of life. Clinicians' enhanced comfort with communicating a diagnosis of cancer, increased concern for palliative symptom control, and the growing interest in quality of life and patient rights emphasized the need for supportive and psychological aspects of care. In the 1980s, psycho-oncology units began to develop in larger cancer centers. First, prevalence studies of psychiatric and psychological sequelae in cancer were reported. Special units were developed throughout Europe and the United States during the 1970s, followed by societies such as the British Psychosocial Oncology Group (1983), the International Psycho-Oncology Society (1984), and the Japanese Psycho-Oncology Society (1985). Conferences, journals (*Journal of Psychosocial Oncology*, 1982; *Psycho-Oncology*, 1992), textbooks, and training programs followed thereafter. In the 1990s, behavioral research in changing habits (e.g., smoking), diet, and lifestyle improved education of the public on cancer prevention. Health-related quality-of-life assessments and more recently

patient-reported outcomes have become part of outcome measures in clinical trials. In the early 21st century, psycho-oncology, a discipline almost 40 years old, continues to grow, with the development of novel psychotherapy modalities for advanced cancer patients; intervention trials to improve symptom control in the terminally ill; increased awareness of the role of communication skills training; researchers' efforts to understand and conceptualize the effects of cancer therapies on the CNS; and recognition of the special needs of elderly cancer patients with the rapidly growing elderly population worldwide.

It is important to emphasize that despite all the developments outlined, historical attitudes toward cancer have contributed to the reluctance of patients and families to identify their emotional problems to the clinicians, even today, especially in underserved populations, different cultures, and several parts of the world.

INCIDENCE OF CANCER AND CANCER MORTALITY RATES IN THE UNITED STATES

In the United States the lifetime risk of developing cancer is slightly less than one in two for men and slightly more than one in three for women. It is estimated that more than 1.6 million new cases of cancer will be diagnosed in 2015. Cancer is the second most common cause of death after heart disease, accounting for one in every four deaths. The 5-year relative survival rate for all cancers diagnosed between 2004 and 2010 is 68 percent; this has improved from 49 percent between 1975 and 1977. Prostate and breast cancers are the most common cancer type in men and women, respectively, but lung cancer is responsible for the highest rates of mortality in both groups. Lung cancer accounts for about 15 percent of cancer diagnoses. The incidence rates of lung cancer have been declining in men since the 1980s; the incidence rates in women are approaching a plateau after a long period of increase. Incidence rates of prostate cancer have changed substantially over the past 20 years, rapidly increasing from 1988 to 1992, declining sharply from 1992 through 1995, remaining stable from 1995 to 2000, and decreasing since 2000. These trends largely reflect increased prostate cancer screening with prostate-specific antigen (PSA) blood testing. Death rates from breast cancer in women have decreased since 1990s due to a combination of earlier detection and improved treatment. Colorectal cancer is the third most common cancer in both men and women. The incidence of colorectal cancer has decreased over the past two decades with screening colonoscopies, through removal of polyps before they become cancerous, and also by detecting a higher proportion of cancers at early and more treatable stages. Given these incidence and mortality figures, it is rare that a family is spared the experience of cancer in one of its members, and many live through it repeatedly, given shared environmental or genetic factors.

COMMON PSYCHOLOGICAL RESPONSES TO CANCER

The diagnosis of cancer creates a crisis requiring the patient to adapt quickly to catastrophic news. Major concerns are fear of death, dependency, disfigurement, disability, and abandonment, as well as disruptions in relationships, role functioning, and financial status. The diagnosis of cancer induces distress, and individuals show an initial characteristic response of shock and disbelief that usually lasts less than 1 week. Patients try to control the level of emotional distress while making crucial treatment decisions. Presence of a relative or friend can help with the processing of important information. Research has shown that the way the news is conveyed by health care professionals can influence a patient's beliefs, emotions, and attitudes toward the future as well as the medical staff. There are guidelines and recommendations on how to deliver the diagnosis, treatment plan, and prognosis.

The second phase of response to diagnosis is characterized by a period of commotion with mixed symptoms of anxiety and depression, irritability, insomnia, poor concentration, and inability to function. These symptoms usually begin to resolve with support from family and friends and from the physician who outlines a treatment plan that offers hope. This phase usually lasts 1 to 2 weeks.

During the third phase, the patient adapts to the diagnosis and treatment. The patient returns to previously used coping strategies that are helpful in reducing stress. These coping strategies are influenced by the previous level of maturation and adaptation.

FACTORS IN ADAPTATION TO CANCER

Patient responses are modulated by factors derived from society, the patient, and cancer type. The society-derived factors reflect society's attitudes toward cancer and its treatment as well as perceptions and knowledge of cancer at a given time. The public nowadays is better informed about cancer, making the illness less threatening and stigmatized. The legal requirement of informed consent has improved communication between the physician and the patient about illness, treatment options, prognosis, and participation as part of the treatment. However, this represents an additional burden to some patients due to the awareness of the severity of their illness. After patients complete their treatment they are constantly monitored for recurrence, which may precipitate fear and anxiety.

Another hardship that became popular is the belief that stress, major life events such as grief, depression, or flaws in personality cause cancer. During the past 10 years, more than 15 reviews of psychosocial risk factors for cancer have been published. All the studies published so far have limitations, making it difficult to reach a definitive decision about causality.

However, the majority of well-designed studies do not confirm that depression, personality, stress, or major life events cause cancer.

The patient-derived factors that modulate adaptation to cancer originate from three sources: intrapersonal factors, interpersonal factors, and socioeconomic and social class. The *intrapersonal factors* include pre-existing character style, coping ability, ego strength, developmental stage of life, and the impact and meaning of the cancer at that stage. *Interpersonal factors* refer to the social support derived from others. *Socioeconomic status* influences the third set of patient-derived factors. Data have shown that lower socioeconomic status can be a barrier to access health care due to lack of education and fewer resources.

The cancer-derived factors that affect adaptation to cancer are related to characteristics of the disease itself, such as stage of the disease, symptoms, site, prognosis, type of treatment, and the impact in functionality.

EPIDEMIOLOGY OF PSYCHIATRIC DISORDERS IN CANCER PATIENTS

Most psychiatric disturbances in patients with cancer relate to the cancer itself or treatment side effects. The prevalence is highest among patients with advanced disease and poor prognosis. The prevalence of psychiatric disorders in cancer patients is approximately 50 percent. More than two-thirds of those represent adjustment disorders, 10 to 15 percent major depression, and about 10 percent delirium. Inpatient studies show a higher incidence of both depression (20 to 45 percent) and delirium (rising from 15 to 75 percent with advancing disease). Studies of psychiatric consultation data reveal that treatable syndromes, such as major depression and delirium, continue to be underdiagnosed and undertreated, despite their known prevalence in cancer patients.

Table 27.8–1.
Symptoms in Patients with Cancer

Symptom	% Positive
Fatigue	74
Worry	71
Sadness	66
Pain	63
Drowsiness	61
Dry mouth	56
Insomnia	54
Poor appetite	45
Nausea	44
Bloating	39
Difficulty in concentration	38
Change in taste	36
Constipation	33

Cough	30
Sexual dysfunction	24
Incontinence	12
Nightmares	11

Table 27.8–1 lists the most common symptoms found in a group of cancer patients with advanced disease. Five of the top seven symptoms are in the psychological realm and are found in more than 50 percent of patients. Three of the next ten symptoms are present in 12 to 45 percent of patients. Other studies have given comparable results. Recognizing the magnitude of the symptom burden of cancer patients has led researchers to an increased interest in identifying patients by their symptoms rather than by their diagnosis.

There are many instruments, the simplest of which is the distress thermometer, in which a patient is asked to rate his or her level of distress in several domains. It has been developed by the National Comprehensive Cancer Network (NCCN) and can be used in other countries and cultures.

Further studies using a wider range of validated instruments, and more focused subpopulations have documented the high prevalence of moderate to severe distress in different subtypes of cancer. Breast cancer remains the most studied site, due to effective activism by the affected population. Prostate cancer is being intensely investigated, and patients with gynecological, colorectal, and lung cancers, patients and families at genetic risk, caregivers, long-term survivors, and geriatric and pediatric age groups are all receiving increased attention. Each subgroup has its own profile of distress and unmet needs.

GENERAL PRINCIPLES OF ASSESSMENT AND TREATMENT IN PSYCHO-ONCOLOGY

The hallmark of treatment in psycho-oncology is therapeutic activism with the simultaneous use of several modalities in an aggressive attempt for rapid relief of symptoms. More than ever, growing understanding of cancer and cancer treatments as traumatic events makes this approach not only humane, but also essential to improving subsequent adjustment. The physician must also address the caregiver and the family, as well as the caregiver and/or family–staff interface, and must remain active through all stages of the illness. Psychiatrists must be prepared for the existential plight of the cancer patient, which intensifies the expectable countertransference reactions. They must resist the urge to distance themselves from a painful situation, while guarding against overinvolvement. They must be aware of the patient's physical condition and yet must not allow acute physical or emotional distress, legitimate despair, and bars to communication, such as tracheostomies, to prevent a thorough evaluation. This requires overcoming fear and guilt and engaging the patient as deeply as one's interviewing skills allow.

Areas of Assessment

Assessment of cancer patients begins with the standard psychiatric history, which should also include an exploration of the patient's understanding of the illness and its meaning. Of particular importance is the prognosis: What is the patient's understanding of his or her disease? What are the patient and family being told? What are they absorbing? There can be major discrepancies in this area among patients, family members, and staff, which leads to unnecessary stress and disruption. The next area is the impact of medical factors on mental state, and the last area is the role of the environment, especially the patient's family and their interface with the health care system. The recurrent demands of cancer treatments promote intense relationships between the patient and family and their medical caregivers, which, if dysfunctional, can lead to psychiatric symptoms and impairment of care.

Treatment Modalities

Psychopharmacology. Because there is frequently a physiological cause or component to psychological symptoms, the psycho-oncologist must encourage an etiological work-up and must recommend medical treatments, such as correction of electrolyte abnormalities and removal of offending drugs. However, symptomatic relief must be provided without waiting for the results. Relief from pain and insomnia goes a long way toward relieving psychological distress. Insomnia is recognized but is often mistreated with hypnotics or antihistamines, which cause habituation or daytime sedation and may worsen symptoms of depression and delirium. The depressed patient may benefit from a low-dose sedating antidepressant, such as mirtazapine (Remeron), or trazodone (Desyrel); or the mildly confused or anxious patient may benefit from a low dose of antipsychotic, such as olanzapine (Zyprexa), risperidone (Risperdal), or quetiapine (Seroquel) at bedtime. Short- to medium-acting benzodiazepines are good hypnotics for intermittent use in the absence of complicating factors, such as delirium.

Fatigue is now getting the investigative attention that it deserves. It can be caused by a psychiatric disorder; the cancer itself; more often, the cancer treatments; or, most often, a combination of the three. Cancer-related fatigue is extensively reviewed later in this section.

Given the interrelatedness of physical and psychological symptoms, it is more important than ever that psychiatrists be comfortable discussing the treatment of symptoms such as nausea, vomiting, diarrhea, dry mouth, and shortness of breath.

Medication management is discussed in detail under specific psychiatric conditions later in this section.

Psychotherapeutic Interventions. Telepsychiatry is being increasingly used as an effective form of intervention, particularly for isolated homebound patients, patients in isolation, or children who need to maintain a sense of connection to a hospitalized parent. DVDs for patient education are also being used more often, but they are most effective when used in combination with face-to-face counseling, rather than as standalone tools.

The Internet has created a new universe of opportunities for patient education, counseling, group interventions, and informal patient-to-patient contacts through the chat room format. The web is heavily used and contains some rich and reliable sites for patients to educate themselves. However, the Internet also contains uncounted commercial sites and impassioned personal sites whose content has been found to be biased and unreliable. They are essentially unsupervisable. There have been some questions about the ethical propriety of researchers entering chat rooms or other informal sites to study them. However, the sites are open to the public, and as long as anonymity is carefully maintained and there is no misrepresentation or interference, data gathering is no different from reading published materials. It is increasingly common for patients to come into the physician's office having done research on their disease. Exploring the patient's Internet experiences can be a productive interaction as long as the psychiatrist maintains an open mind without relinquishing his or her professional stance.

Reviews of the many available psychotherapeutic modalities, such as cognitive–behavioral therapy, crisis intervention, problem-solving techniques, group psychotherapy, and supportive psychotherapy with an educational and a psychodynamic focus, describe a beneficial effect for all of them. Structured interventions are efficient, showing benefit after a brief course of treatment. Some studies have demonstrated persistent effects after 1 year. Working with individual patients or groups, the psychotherapies lower distress, increase self-esteem and self-image, decrease anxiety, and give patients a better sense of control. Dynamically oriented therapy still has a place in individuals whose personal issues overshadow the cancer, a situation that is more frequent than one might imagine. However, it should remain flexible to reflect the patient's shifting condition.

Referrals often begin with a crisis intervention but rapidly move into other modes. Many cancer patients are happy to leave psychotherapy when they feel better and return only episodically, when in relapse or crisis. This is not treatment resistance; it is inherent in the original treatment contract. Two major issues for patients are the reality of living a shortened life span and the coming of death. Despite their close connection, these are often dealt with separately, the first as a negotiable issue on a continuum with other life problems, the second as the mysterious and frightening absolute

that it is. Psychiatrists must overcome their own reflexes to stay with the first issue and to avoid the second. Later in their course, many patients come to an ambivalent acceptance of and even a yearning for death. It is important to recognize both sides of the ambivalence and to help families and staff tolerate and accept it as well. Families may be feeling ambivalent at this time and are also in need of support. If the patient dies, one is in a unique position to offer bereavement counseling.

Families should be routinely involved in the treatment of ill or physically dependent patients; hence, family-focused therapy and family groups are now more common and are under intensive study. Caregiver groups have developed in recognition of their heavy burden. Cancer survivors are being studied and treated in the family context, understanding that the entire family is permanently altered and that the welfare of several individuals is at stake.

Patient groups provide education, problem-solving exercises, mutual support, ventilation, validation of feelings, reassurance and help about side-effect management, and help in negotiating the doctor–patient relationship and the interface with the medical system. These effects are so positive that groups have multiplied. When a patient describes a negative response to a support group, the reaction should be respected but explored, because it usually reveals important psychological concerns.

Groups can be led by mental health professionals who facilitate psychological work and oncology staff members who represent an authoritative source of information. Patients appreciate both kinds of help, and combination groups with a mix of structured and unstructured meetings satisfy the greatest number of patients. Self-help groups increase patients' sense of empowerment and can be helpful, but, over time, they may be overtaken by individual members and make some patients feel excluded or minimized. Putting in democratically defined structure and guidelines can increase their longevity. Fifteen years after its publication, efforts to duplicate a group intervention that prolonged survival in breast cancer patients have been unsuccessful. However, even if the effect on survival is in doubt, the improved quality of life, coping skills, self-esteem, pain management, and interpersonal relations were never questioned and remain as important as ever.

MCP is a novel modality that is designed to help patients with advanced cancer to sustain or enhance a sense of meaning, peace, and purpose in their lives even as they approach the end of life. Inspired primarily by the works of Viktor Frankl and further informed by the contributions of Irvin Yalom, Breitbart and colleagues adapted Frankl's concepts involving the importance of meaning in human existence and created MCP. While MCP relies heavily on Frankl's concepts of meaning, it also incorporates other fundamental existential concepts related to the search, connection, and creation of meaning. Using didactics and experiential exercises in each

session, the therapist and the patient(s) work together to help patients understand the importance and relevance of sustaining, reconnecting with, and creating meaning in their lives through common and reliable sources of meaning. Furthermore, patients are taught that these various sources of meaning may serve as *resources* during especially difficult times to help diminish despair near the end of life. Both individual and group formats of MCP have been developed and tested in a number of randomized controlled studies. These studies provide strong support for the efficacy of MCP as a treatment for psychological and existential/spiritual distress among patients with advanced cancer.

Educational and support programs can be built into the medical routine, as in presurgery education groups, postsurgery rehabilitation groups, and even waiting room groups. The growing number of survivors has prompted the development of innovative posttreatment group programs. Behavioral techniques, such as hypnosis, self-hypnosis, active and passive relaxation, desensitization, and distraction, are useful in a wide range of situations in oncology, including relief of acute and chronic anxiety related to surgery and painful procedures; relief or prevention of conditioned symptoms, such as anticipatory nausea and vomiting (ANV) associated with chemotherapy; and, with age-appropriate modification, improved management of almost all pediatric problems. Their spectrum of use can be as varied as the expertise and initiative of available practitioners.

Social Services. Social workers who are available in large centers or community institutions are a first line of psychological support. They may provide extensive psychotherapeutic support, given the training and the opportunity, but, at the very least, they perform a triage function, alerting staff to patients with serious psychopathology. Their classic role is to help patients and families find their way to available resources, which is often the most helpful intervention that patients identify. If social workers are not available, or it is not part of their mandate, psychiatrists must do their best to fulfill this function by at least having a good file of local resources and a readiness to refer patients to them.

Chaplaincy Services and the Interface of Religion and Psychiatry. As many as 80 percent of cancer patients describe themselves as believing in God, and a large proportion of those state that faith is an important, if not primary, contributor to their coping ability. There is a large and growing literature on the role of spiritual belief in health behavior. Most of the studies are small and methodologically weak, and there may be a publication bias against negative results. Many more studies have found an association between religious beliefs and psychological variables. Several studies have shown that religious beliefs are associated with mature and

active coping methods. A few studies have focused on negative effects, such as an irreducible sense of betrayal, a sense of being punished for personal failings, or strong feelings of guilt about anger and loss of faith.

There is some controversy in the pastoral community about the usefulness, or even the propriety, of doing research on religious phenomena and the appropriateness of doctors intruding on religious practices, especially in prescriptive ways. However, it is important for psychiatrists to become knowledgeable about the interface between spiritual beliefs and psychological experiences without a negative bias toward the former, especially with ill patients in whom existential concerns are universally present.

The psychiatric consultant should inquire about faith, its meaning, associated religious practices, and its impact on the coping response. It can be a source of strength or guilt at all stages of the disease, ranging from the earliest "What did I do to cause this?" to "Will God give me only what I can carry?" in midcourse, to the poignant life review of the late stage. It is often a primary modifier of attitudes toward suicide and terminal care decisions. Mental health professionals should be prepared to deal with these areas in an unself-conscious, uncondescending manner, enabling the patient to fully integrate this aspect of his or her personality into the therapeutic relationship and the current crisis. However, this should never be mistaken for spiritual counseling. Some centers are fortunate to have chaplains available to patients on request. Psychiatrists should work with them, should let the patient know about their availability, and should arrange a visit if the patient shows interest. They should encourage the use of other spiritual guides in the community as well. Where specialized chaplains are not available, the psychiatrist should become familiar with selected religious leaders from the major faiths that appear in their practice, so as to obtain advice and to make referrals.

The Wellness Movement. Wellness programs are a growing group of programs that demonstrate multiple and original uses of group modalities to help cancer patients and their families. The Wellness Community (TWC) and Gilda's Club are good prototypes, but there are fine local programs throughout the country, some of which provide additional personal services, such as transportation or accompaniment to doctors' appointments. They are nonprofit organizations, are free of charge, and are devoted to helping cancer patients and their families cope and grow through the illness. They create an upbeat, welcoming environment for the whole family and position themselves as complementing standard medical treatment. They encourage patients to take a more active role in their care and coach them to interact more effectively with their medical caregiver. Although volunteers play a large role, and patients themselves are welcome to help, the specialized offerings are all run by trained professionals.

Volunteerism. The programs described previously, created and maintained by dedicated local and national volunteer workers and philanthropists, are only a small part of the support provided to cancer patients through volunteer activities. The ACS, the Leukemia Lymphoma Society of America (LLSA), and numerous institutions and family foundations have developed patient services along with their major research funding. "I Can Cope" is one of the largest. Cancer Care runs numerous groups and has established telephone and Internet groups. One organization fulfills the wishes of extremely ill children (e.g., a trip to Disneyland or a visit from a famous sports hero). Another provides free transportation to medical centers in executive jets. Most of them originated with grateful survivors or bereaved families trying to give meaning to their loss. This does not include the hundreds of thousands of individuals who volunteer their time in hospitals and clinics and are willing to do whatever is asked to make the patient's experience more tolerable. For the volunteers who provide the help, it is a source of increased self-esteem and life purpose, whether they have a personal cancer experience or not. The magnitude and variety of offerings are heartening for the medical caregivers and a reminder to all that selfless initiatives are still an important aspect of society.

DIAGNOSIS AND MANAGEMENT OF PSYCHIATRIC DISORDERS IN CANCER PATIENTS

Delirium, Dementia, and Other Cognitive Disorders

Cognitive disorders due to chemotherapy, radiation, cancer, and paraneoplastic syndromes will be reviewed separately later in this section. Dementia will be reviewed under geriatric psycho-oncology section.

Delirium. Delirium is the most common and often serious neuropsychiatric complication in the management of cancer patients characterized by an abrupt onset of disturbances in level of consciousness, awareness, attention, cognition, and perception that tend to fluctuate over the course of the day, precipitated by an underlying medical condition. Delirium is a medical emergency that needs to be identified and treated vigorously. Unfortunately, delirium is frequently underdiagnosed and untreated in the medical setting, which leads to increased morbidity and mortality, interference in the management of symptoms such as pain, an increased length of hospitalization, increased health care costs, as well as distress for patients and their caregivers. Delirium is associated with increased morbidity in the terminally ill, causing distress in patients, family members, and staff. Fifty-four percent of patients were found to have recalled their delirium experience after recovery from delirium in a study of hospitalized cancer patients. Factors predicting delirium recall included the

degree of short-term memory impairment, delirium severity, and the presence of perceptual disturbances. The most significant factor predicting distress for patients was the presence of delusions. Patients with hypoactive delirium were just as distressed as patients with hyperactive delirium.

EPIDEMIOLOGY. In cancer patients, delirium is present in 25 to 85 percent depending on the stage of illness. Prevalence rates of delirium range from 15 to 30 percent in hospitalized cancer patients; they are higher in the last weeks of life, at 40 to 85 percent. The reported prevalence of delirium in hospice may approach as high as 85 percent. Elderly patients who develop delirium during a hospital stay have an estimated mortality rate of 22 to 76 percent during that hospitalization.

ETIOLOGY. Frequently, a combination of factors (three or more etiologies) contributes to the development of delirium. Delirium is irreversible 30 to 40 percent of the time, especially in the terminally ill. In the event that the etiology of the delirium is unidentifiable or irreversible, relief from the distressing symptoms of delirium may be provided to the patient. Predisposing factors including age, functional impairment, and the nature and severity of illness increase the risk of experiencing delirium during hospitalization. Dementia is an important independent risk factor for developing delirium. Physical impairment, sensory impairment, low BMI, cachexia, and malnutrition have been associated with increased risk for delirium. With an increase from zero to three risk factors, the incidence of delirium rises from 2 to 50 percent, bringing longer hospital stays and poorer outcomes. In patients with advanced cancer, delirium can be due either to the direct effects of cancer on the CNS or to indirect CNS effects of the disease or treatments (e.g., medications, electrolyte imbalance, dehydration, major organ failure, infection, vascular complications, paraneoplastic syndromes). Given the large numbers of drugs cancer patients require and the fragile state of their physiological functioning, even routinely ordered hypnotics are enough to tip patients over into a delirium. Narcotic analgesics such as levorphanol (Levo-Dromoran), morphine sulfate (Roxanol), and meperidine (Demerol), are common causes of confusional states, particularly in the elderly and terminally ill. Chemotherapeutic agents known to cause delirium include methotrexate (Rheumatrex), ifosfamide (Ifex), fluorouracil (Adrucil), vincristine (Oncovin), vinblastine (Velban), bleomycin (Blenoxane), BCNU (carmustine, Gliadel), cisplatin (Platinol-AQ), asparaginase (Elspar), procarbazine (Procanbid), and corticosteroids. Most patients receiving chemotherapeutics will not develop prominent CNS effects unless dexamethasone (Decadron) or prednisone (Deltasone) is included in the regimen. In addition, drug withdrawal may be a common cause of delirium, especially in postoperative patients whose drugs have been interrupted for

surgery and have not been promptly restarted (Table 27.8–2).

CLINICAL FEATURES AND DIAGNOSIS. Clinically, the diagnostic gold standard for delirium is the clinician's assessment utilizing the *DSM-5* criteria. Clinical features of delirium include abrupt onset, fluctuating course, attention disturbances, altered level of alertness, awareness, and arousal, increased or decreased psychomotor activity, disturbance of sleep–wake cycle, affective symptoms, altered perceptions, disorganized thinking and incoherent speech, disorientation, and memory impairment. Neurological abnormalities including cortical (aphasia, apraxia, dyssomnia, dysgraphia) and motor symptoms (tremor, asterixis, myoclonus, changes in muscle tone and deep tendon reflexes) can also be present. It is this protean nature of the symptoms and the fluctuation of clinical findings that make delirium so difficult to diagnose and treat.

Table 27.8–2.
Causes of Delirium in Cancer Patients

Direct
Primary brain tumor
Metastatic spread to the brain
Indirect
Hypoxia
Metabolic encephalopathy due to organ failure
Electrolyte imbalance
Alcohol and drug withdrawal states
Treatment side effects from
Chemotherapeutic agents, steroids, biological response modifiers
Radiation
Opioids
Anticholinergics
Antiemetics
Benzodiazepines
Infection
Hematological abnormalities
Nutritional deficiencies
Vitamin deficiencies (thiamine, folate, vitamin B_{12}, and niacin)
Paraneoplastic syndromes
Endocrine disorders
Tumors of the pituitary, thyroid, parathyroid, and adrenal
Multiple endocrine neoplasia syndromes
Carcinoid tumors, mainly of the ovary and lung

Delirium is classified into three clinical subtypes, most commonly based on psychomotor behavior, including the hyperactive, the hypoactive, and the mixed subtype. Approximately two-thirds are either of the hypoactive or mixed subtype. The hyperactive form is most often characterized by

hallucinations, delusions, agitation, and disorientation, while the hypoactive form is characterized by confusion and sedation, but is rarely accompanied by hallucinations or delusions.

Many of the clinical features and symptoms of delirium can also be associated with other psychiatric disorders such as depression, mania, psychosis, and dementia. For instance, delirious patients, not uncommonly, exhibit mood disturbances such as anxiety, fear, depression, irritability, anger, euphoria, apathy, or lability. Delirium, particularly the hypoactive subtype, is often initially misdiagnosed as depression. In distinguishing delirium from depression, particularly in the context of advanced cancer, an evaluation of the onset and temporal sequencing of depressive and cognitive symptoms is particularly helpful.

A number of scales or instruments have been developed that can aid the clinician in rapidly screening for cognitive disorders, or in establishing a diagnosis of delirium, including the Delirium Rating Scale–Revised 98, the Confusion Assessment Method, the Abbreviated Cognitive Test for Delirium, and the Memorial Delirium Assessment Scale.

The Memorial Delirium Assessment Scale (MDAS) is a 10-item delirium assessment tool, validated among hospitalized patients with advanced cancer and acquired immunodeficiency syndrome (AIDS). The MDAS is a delirium diagnostic screening tool as well as a reliable tool for assessing delirium severity among patients with advanced disease. A cutoff score of 13 is diagnostic of delirium. Recently, MDAS was further examined for clinical utility and validation in a population of advanced cancer patients in a palliative care unit. These investigators found the MDAS to be useful in this population, and found that a cutoff score of 7 of 30 yielded the highest sensitivity (98 percent) and specificity (76 percent) for a delirium diagnosis in this palliative care population.

ASSESSMENT AND MANAGEMENT OF DELIRIUM. The standard approach to managing delirium in the medically ill, and even in those with advanced disease, includes a search for underlying causes, correction of those factors, and management of the symptoms of delirium. The desired and often achievable outcome is a patient who is awake, alert, calm, cognitively intact, not psychotic, and communicating coherently. Studies in patients with earlier stages of cancer have demonstrated the potential utility of a thorough diagnostic assessment. When such diagnostic information is available, specific therapy may be able to reverse delirium. The diagnostic evaluation includes an assessment of potentially reversible causes of delirium. A full physical examination should assess for evidence of sepsis, dehydration, or major organ failure. Medications that could contribute to delirium should be reviewed. A screen of laboratory parameters will allow assessment of the possible role of metabolic abnormalities. Imaging studies of the brain and assessment of the cerebrospinal fluid may be appropriate

in some instances.

In the terminally ill patient who develops delirium in the last days of life (referred to as terminal delirium), the management of delirium is unique, presenting a number of dilemmas, and the desired clinical outcome may be significantly altered by the dying process. When confronted with delirium in the terminally ill or dying patient, a differential diagnosis should always be formulated as to the likely etiology or etiologies. There is an ongoing debate as to the appropriate extent of diagnostic evaluation that should be pursued in a dying patient with a terminal delirium. Most palliative care clinicians would undertake diagnostic studies only when a clinically suspected etiology can be identified easily, with minimal use of invasive procedures, and treated effectively with simple, interventions that carry minimal burden or risk of causing further distress. Diagnostic work-up in pursuit of an etiology for delirium may be limited by either practical constraints such as the setting (home, hospice) or the focus on patient comfort, so that unpleasant or painful diagnostics may be avoided. Most often, however, the etiology of terminal delirium is multifactorial or may not be determined. Etiology is discovered in less than 50 percent of terminally ill patients with delirium. When a distinct cause is found for delirium in the terminally ill, it is often irreversible or difficult to treat. One study found that 68 percent of delirious cancer patients could be improved, despite a 30-day mortality of 31 percent. Another found that one-third of the episodes of cognitive failure improved following evaluation that yielded a cause for these episodes in 43 percent. In a recent prospective study of delirium in patients on a palliative care unit, investigators reported that the etiology of delirium was multifactorial in the great majority of cases. Even though delirium occurred in 88 percent of dying patients in the last week of life, delirium was reversible in approximately 50 percent of episodes. Causes of delirium that were most associated with reversibility included dehydration and medications. Hypoxic and metabolic encephalopathies were less likely to be reversed in terminal delirium.

TREATMENT. The standard approach to the management of delirium includes identification and elimination of factors contributing to delirium, in addition to nonpharmacological and pharmacological treatment interventions. Nonpharmacological approaches include creating a calm, comfortable, well-lit environment, using orienting objects such as calendars and clocks, early mobilization, having family members around, limiting room and staff changes, and allowing patients to have uninterrupted periods of rest at nights to improve the sleep–wake cycle. One-to-one nursing observation may also be necessary and useful. Fluid and electrolyte balance, nutrition, and vitamins may be useful.

Supportive techniques alone are often not effective in controlling the symptoms of delirium, and pharmacological treatment is necessary in most

cases. Antipsychotics constitute the primary pharmacological intervention. Antipsychotics such as haloperidol (Haldol) are frequently utilized as antiemetics in the medical setting; however, only 0.5 to 2 percent of hospitalized cancer patients receive haloperidol for the management of the symptoms of delirium. In terminally ill populations, only 17 percent receive an antipsychotic for agitation or psychological distress, despite an estimated prevalence of delirium ranging from 25 percent in the hospitalized cancer patient to 85 percent in the terminally ill.

Haloperidol (a "typical" antipsychotic) is the gold-standard medication for treatment of delirium among cancer patients, due to its efficacy and safety (e.g., few anticholinergic effects, lack of active metabolites, and availability in different routes of administration). Haloperidol in low doses (1 to 3 mg per day) is usually effective in targeting agitation and psychotic symptoms. In general, doses of haloperidol need not exceed 20 mg in a 24-hour period; however, some clinicians advocate higher doses in selected cases. In severe agitation related to delirium, clinicians may add lorazepam to haloperidol. This combination may be more effective in rapidly sedating patients and may help minimize any extrapyramidal adverse effects of haloperidol. However, there is evidence that benzodiazepines can worsen delirium, and switching to a more sedating antipsychotic is another option.

Oral or intravenous (IV) chlorpromazine is considered to be an effective alternative to haloperidol (with or without lorazepam) when increased sedation is required, especially in the ICU setting where close blood pressure monitoring is feasible, and for severe agitation in terminally ill patients to decrease distress for the patient, family, and staff. It is important to monitor chlorpromazine's anticholinergic and hypotensive side effects, particularly in elderly patients.

In a double-blind, randomized comparison trial (n = 30) of haloperidol, chlorpromazine, and lorazepam, Breitbart and colleagues demonstrated that lorazepam alone, in doses up to 8 mg in a 12-hour period, was ineffective in the treatment of delirium, and in fact sometimes worsened it. Both haloperidol and chlorpromazine, in low doses (approximately 2 mg of haloperidol equivalent per 24 hours) were effective in controlling the symptoms of delirium and in improving cognitive function in the first 24 hours of treatment. Hyperactive and hypoactive subtypes of delirium were equally responsive to treatment with haloperidol or chlorpromazine.

A Cochrane review on drug therapy for delirium in the terminally ill concluded that haloperidol was the most suitable medication for the treatment of patients with delirium near the end of life, with chlorpromazine being an acceptable alternative.

The FDA has issued a warning against the risk of QTc prolongation and *torsades de pointes* with IV haloperidol; thus, monitoring QTc intervals daily among medically ill patients receiving IV haloperidol has become the standard clinical practice.

Atypical antipsychotics, namely, olanzapine (Zyprexa), risperidone, (Risperdal), quetiapine (Seroquel), ziprasidone (Geodon), and aripiprazole (Abilify) are being increasingly used in the treatment of delirium due to their favorable side-effect profile. Several studies have been conducted with atypical antipsychotics, including olanzapine, risperidone, quetiapine, and ziprasidone in treatment of delirium. In a double-blind comparative delirium intervention study assessing the efficacy of haloperidol versus risperidone, researchers demonstrated, in a small sample of 24 oncology patients, that there was no significant difference in clinical efficacy or response rate. The mean risperidone dose was 1.02 mg, and the mean haloperidol dose was 1.71 mg. MDAS scores improved significantly in both groups. No significant difference was observed in adverse effects. However, despite the double-blind design in this study, authors acknowledged that they were not able to obtain identical looking tablets of haloperidol and risperidone. A relatively recent open-label trial using liquid risperidone in 29 delirious cancer patients (mean age 69) referred to their psychiatric consult service found reduction of delirium severity in 79 percent of patients, response (defined as at least 25 percent reduction in DRS-R-98 severity score) in 48 percent of patients, and remission in 38 percent of patients after seven days. Risperidone has also been compared to olanzapine in a small (n = 32, primarily oncology patients) randomized, single-blind trial. Researchers in this study used the DRS-R-98 to assess change over seven days and reported no difference in response rates between the two groups. An (n = 82) open trial of olanzapine for the treatment of delirium in hospitalized patients with advanced cancer showed that olanzapine was effective in the treatment of delirium, resolving delirium in 76 percent of patients with no incidence of extrapyramidal side effects. Several factors were found to be significantly associated with poorer response to olanzapine treatment for delirium, including age over 70, history of dementia, and hypoactive subtype of delirium. The average starting dose was in the 2.5 to 5 mg range, and patients were given up to 20 mg per day of olanzapine. Sedation was the most common side effect. Preliminary work suggests that olanzapine could be given subcutaneously if needed. Good tolerability of subcutaneous olanzapine was demonstrated in an open-label study of 24 advanced cancer patients with agitated delirium. No injection site toxicity was observed, but there were four possible systemic adverse events recorded (severe hypotension, seizure, diabetes insipidus, and paradoxical agitation). Further work is needed to assess the safety and efficacy of the subcutaneous route of administration. A small open-label trial of 12 patients (most with leukemia) showed good effectiveness and tolerability of quetiapine at an average dose of 94 mg. Mean time to stabilization was approximately 6 days. From a prospectively collected database of delirious inpatients with cancer, researchers compared 21 patients who received

aripiprazole with 21 case-matched patients who received haloperidol. They found no significant differences in effectiveness between the two medications, and haloperidol caused more extrapyramidal side effects.

A Cochrane review comparing the efficacy and the incidence of adverse effects between haloperidol and atypical antipsychotics concluded that, like haloperidol, selected newer atypical antipsychotics (risperidone, olanzapine) were effective in managing delirium. Haloperidol doses greater than 4.5 mg per day were more likely to result in increased rates of EPS compared with the atypical antipsychotics, but low-dose haloperidol (i.e., less than 3.5 mg per day) did not result in a greater frequency of extrapyramidal adverse effects.

APA practice guidelines provide directions for the use of antipsychotics in the treatment of delirium, and growing evidence supports their use. The APA guidelines for treatment of delirium recommend the use of low-dose haloperidol (i.e., 1 to 2 mg PO every 4 hours as needed, or 0.25 to 0.5 mg PO every 4 hours for the elderly) as the treatment of choice in cases where medications are necessary.

In the light of existing literature, risperidone may be used in the treatment of delirium, starting at doses ranging from 0.25 to 1 mg and titrated up as necessary, with particular attention to the risks of EPS, orthostatic hypotension, and sedation at higher doses. Olanzapine can be started between 2.5 and 5 mg nightly and titrated up, with sedation being the major limiting factor, which may be favorable in the treatment of hyperactive delirium. The current literature on the use of quetiapine suggests a starting dose of 25 to 50 mg and a titration up to 100 to 200 mg per day (usually at twice-daily divided doses). Sedation and orthostatic hypotension are the main dose-limiting factors. Findings to date suggest a starting dose of 10 to 15 mg daily for aripiprazole, with a maximum dose of 30 mg daily.

Important considerations in starting treatment with any antipsychotic for delirium may include EPS risk, sedation, anticholinergic side effects, cardiac arrhythmias, and possible drug–drug interactions. The FDA has issued a “black box” warning of increased risk of death associated with the use of typical and atypical antipsychotics in elderly patients with dementia-related psychoses. Clinicians should attempt to use low doses only for short periods of time, especially when treating elderly patients with delirium. However, leaving delirium untreated may impose a greater risk of morbidity and mortality (Table 27.8–3).

Several researchers have studied both pharmacological and nonpharmacological interventions in the prevention of delirium among older patient populations, particularly in surgical settings with mixed results. The applicability of these interventions to the prevention of delirium in cancer patients has not been studied.

Anxiety Disorders

Anxiety is the most common response in the setting of cancer. It is a natural human response mechanism when confronted by a threat, but it can become maladaptive when the response is excessive. For example, in the course of treatment a patient sometimes will learn that their treatment has failed or that the disease has recurred. These situations are often met with disbelief and denial followed by emotional upheaval, characterized by mixed symptoms of depression and anxiety, together with loss of appetite, insomnia, and lack of concentration. These symptoms typically decrease gradually and resolve over a 7- to 10-day period. Anxiety may represent an acute reaction to cancer; however, some patients have premorbid psychological issues that are exacerbated by the diagnosis of cancer. This premorbid psychological functioning is influenced by developmental state, personality, and prior psychiatric history.

Table 27.8–3.
Antipsychotics in the Treatment of Delirium

Generic Name	Approximate Daily Dose Range (mg)	Route[a]
Haloperidol	0.5–2 mg every 2–12 hrs	po, iv, sc, im
Chlorpromazine	12.5–50 mg every 4–12 hrs	po, iv, im
Olanzapine	2.5–10 mg every 12–24 hrs	po, ODT[b]
Risperidone	0.25–2 mg every 12–24 hrs	po, ODT[b]
Quetiapine	12.5–200 mg every 12–24 hrs	po
Ziprasidone	10–40 mg every 12–24 hrs	po
Aripiprazole	10–30 mg every 24 hrs	po, ODT[b]

po, orally; iv, intravenous; sc, subcutaneous; im, intramuscular.

[a]Risperidone and aripiprazole are available in liquid formulations.

[b]Orally disintegrating tablet (ODT) forms are available for olanzapine, risperidone, and aripiprazole.

Symptoms and Classification of Anxiety Disorders. Anxiety is manifested by a broad array of physical signs and measures of autonomic activation, changes in thinking (i.e., intrusive thoughts), and behaviors. The most common anxiety disorders in cancer patients are standardized in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5).

Generalized anxiety disorder is characterized by excessive anxiety across different settings. The worrying is difficult to control, evoking fatigue, irritability, muscle tension, and sleep dysfunction.

Panic disorder entails recurrent unexpected panic attacks, marked by intense fear or discomfort associated with palpitations, diaphoresis, dyspnea, sense of choking, nausea, dizziness, or derealization. The attacks may occur from a calm or anxious state and are followed by worry, concern, and behavior changes related to the attack.

Specific phobias are characterized by persistent and excessive fear elicited by the presence or anticipation of a specific object or situation. Phobias of blood, needles, hospitals, magnetic resonance imaging machines, and radiation simulators may complicate treatment adherence. Patients undergoing chemotherapy have nausea and vomiting and about 25 to 30 percent develop ANV despite adequate antiemetic regimens. Patients experience nausea when they are exposed to stimuli that remind them about drug infusions.

The classification "anxiety disorder due to a general medical condition" is meant to describe cases in which the symptoms are a direct physiological consequence of the illness. Anxiety is ubiquitous and can be found with many of the conditions that cause other psychiatric disorders. Anxiety in cancer can be caused by different medical conditions including pulmonary embolism, hypoxia, hypoglycemia, and cardiac disorders. Hormone-secreting neoplasms and paraneoplastic syndromes can also cause anxiety (Table 27.8–4).

Anxiety due to substance withdrawal is recognized and treated by most clinicians. Nicotine withdrawal is less commonly recognized. Many patients stop smoking abruptly on receiving their diagnosis and may experience severe anxiety symptoms, sometimes associated with dissociative symptoms and bizarre behaviors. Behavioral smoking cessation programs are an indispensable adjunct and are commonly offered to family members as well, taking advantage of what is known as a window of opportunity, when defenses are shaken and motivation is high.

Table 27.8–4.
Medical Conditions Associated with Anxiety

Metabolic	Hyperkalemia
	Hyperthermia
	Hypoglycemia
	Hyponatremia
	Vitamin deficiencies
Cancer	Central nervous system neoplasms
	Carcinoid syndrome
	Lung cancer
Endocrine	Adrenal abnormalities
	Thyroid abnormalities
	Parathyroid abnormalities
	Pituitary abnormalities
	Pheochromocytoma
Cardiovascular	Arrhythmia
	Congestive heart failure
	Myocardial infarction
	Angina pectoris
	Valvular disease

Pulmonary	Pulmonary embolism
	Asthma
	Chronic obstructive pulmonary disease
	Pneumothorax
	Pulmonary edema
Medication/toxic conditions	Corticosteroids
	Neuroleptics
	Thyroxine
	Sympathomimetic agents
	Serotonergic agents
	Antihistamines
	Withdrawal states (alcohol, opioid analgesics, sedative-hypnotics, caffeine)
	Antibiotics (cephalosporins, acyclovir, isoniazid)
Other symptoms	Pain

Substance-induced anxiety is a common cause of anxiety in cancer patients. They include: bronchodilators, frequently used in lung cancer patients in whom they compound the anxiety caused by air hunger; corticosteroids, the mainstay of many protocols; anti-inflammatory cytokines, which are being increasingly used for treatment; and antipsychotics, including some of the ones that are used as antiemetics, such as metoclopramide (Reglan). Akathisia is often mistaken for anxiety by patient and staff; thyroxine, psychostimulants, sympathomimetic agents, serotonergic agents, antihistamines, and certain antibiotics produce symptoms of anxiety.

Treatment. Treatment aims include reducing the patient's overall level of emotional distress as well as reducing specific target symptoms that may impair social or occupational function. Treatment of anxiety in cancer patients depends on the etiology and the timing of onset of symptoms. In some situations, anxie ty symptoms can represent disorders that started before the onset of a medical illness. In that case the treatment approach should be directed to the premorbid condition. In other instances anxiety can be the consequence of a medical condition or its treatment. In such cases the treatment should target the underlying condition.

Current or past delirium can contribute to anxiety in cancer patients. Delirious patients can present with early symptoms of restlessness, irritability, altered sleep patterns, and impaired concentration. Patients who have been delirious may have persistent worry related to their memory of this experience or their behavior during episodes of delirium.

PHARMACOLOGICAL TREATMENT. Medications used to treat anxiety in cancer patients are also used in primary anxiety disorders. These medications are benzodiazepines, antidepressants, and atypical antipsychotics.

NONPHARMACOLOGICAL TREATMENT. Nonpharmacological treatment involves several approaches: psychotherapy, education, and supportive measures. There are different types of psychotherapy including but not limited to interpersonal psychotherapy, supportive psychotherapy, and cognitive–behavioral therapy. MCP should be considered in advanced cancer patients with anxiety symptoms. Several behavioral methods have been used such as progressive muscle relaxation, breathing exercises, meditation, biofeedback, and guided imagery.

Trauma- and Stressor-Related Disorders

Trauma- and stressor-related disorders are a group of diagnoses that entail both a traumatic or stressful exposure and a subsequent change in emotions and behavior. These include adjustment disorders, acute stress disorder, and PTSD.

Adjustment disorders with depressed mood, anxious mood, and mixed emotional features represent the largest group of psychiatric diagnoses found in cancer patients. Diagnosis of cancer may invoke a normal reaction to stress where mood can be affected in cancer patients. They may use their coping mechanisms and continue with their life without it affecting their functioning. Patients with cancer are on a continuum, with acute stress responses at one end and mood and anxiety disorders at the other. Adjustment disorders represent a maladaptation in one or more of the psychosocial areas. This disorder may resolve when the stressor is over but it may become chronic requiring medications along with counseling. The prevalence of adjustment disorder in oncological care settings is 19.4 percent, but often co-occurs with anxiety and mood disorders.

Acute stress disorder (ASD) follows exposure to a traumatic event. Having cancer is perceived as a life-threatening event. The symptoms of ASD include the combination of one or more of intrusion, dissociative, avoidance, or arousal symptoms and negative mood. Dissociative symptoms include an altered sense of the one's self or surroundings, or an inability to remember aspects of the event. These symptoms typically begin immediately after the trauma and persist for 3 days to 1 month after exposure; PTSD should be considered for symptoms that have persisted for more than 1 month after such an event. The prevalence of ASD has not been established in cancer patients.

The essential feature of PTSD is the development of problematic symptoms following the experience of a traumatic event that involves threat of death or injury. Syndromal presentation includes intrusion symptoms, negative cognitions and mood, persistent avoidance of stimuli associated with the trauma, and marked alteration in arousal and reactivity associated with the trauma. PTSD differs from ASD in that the disturbance lasts for longer than 1 month. Prevalence of PTSD has been reported as

high as 37 percent among survivors of non-Hodgkin lymphomas at a median of 12.9 years after diagnosis. It is difficult to assess classical PTS symptoms in cancer patients. Evaluation of intrusive symptoms is nonspecific since cancer patients often report fears related to the future, and rarely flashbacks or intrusive memories. Avoidance behavior is difficult to ascertain, as patients are frequently confronted with the disease and its treatment with potential trauma-related stressors, and there have been reports of patients who felt worse because of premature interventions directed toward re-experiencing the trauma so as to integrate it. Acute posttraumatic denial needs to be approached cautiously, if at all. PTSD can be acute (immediate and lasting less than 3 months), chronic (lasting more than 3 months), delayed (if starting more than 6 months after the event), and recurrent (when a recurrence of symptoms is precipitated by another event that is reminiscent of but not necessarily as severe as the initial event). This latter type is more frequent than reported and should be sought out, because patients themselves do not make the connection. In a paradoxical case, a Vietnam veteran was unaware of the striking parallels between his severe depressive response to being told that he did *not* have lung cancer and his reaction on returning stateside many years earlier. He experienced both events as undeserved reprieves from death that aroused profound survivor guilt about his dead buddies and the living ones that he had left behind.

Because avoidance of people and places associated with cancer, such as a clinic, is part of PTSD, this syndrome may interfere with optimal medical care delivery. Therapies used to treat PTSD include crisis intervention to lessen symptoms and improve functioning, cognitive–behavioral therapy to change thinking and behavior, and support groups for further support and to normalize the presentation. Patients with severe symptoms may benefit from medications, such as antidepressants, anxiolytics, and neuroleptics.

Cancer-related posttraumatic stress (PTS) is a syndrome much like PTSD, though presentation is generally less severe. Patients with new diagnosis of cancer may experience recurrent frightening thoughts, distractibility, insomnia, and a sense of detachment or depersonalization. Helplessness is also a common symptom. Cancer-related PTS can occur anytime during or after treatment. Risk factors for cancer-related PTS can include diagnosis of cancer that recurs, longer course of treatment, and past history of anxiety or trauma. Good social support and clear communication are thought to be protective against PTS symptoms. Treatments used for PTS are similar to those for PTSD, including early identification, crisis intervention, and relaxation training. Support groups may be helpful. For patients with severe symptoms of PTS, antidepressant and anxiolytic medicines may be used, especially in cases of comorbid mood and anxiety disorders.

Depressive Disorders

In DSM-5, depressive disorders include major depressive disorder, depressive disorder due to another medical condition, and substance/medication-induced depressive disorders. This section will principally review depressive disorders in cancer patients. Depressive symptoms are common in the oncology setting. Depression is a common psychiatric complication of cancer and is an important risk factor for suicide. Cancer patients are vulnerable to these symptoms at all stages of the illness, from appearance of the first symptoms of cancer to the time of diagnosis, during treatment, palliative care, and even after remission or cure. Evaluating the nature and intensity of depressive symptoms is important. For the clinician, the relevant clinical question is how to identify the point when normal sadness or distress associated with the cancer has become a depressive disorder. Diagnosis of depression in the context of cancer is challenging. Recognizing the risk factors for depression, with careful attention to the signs and symptoms of depression, leads to improved recognition and treatment of depressive disorders, thus increasing adherence to cancer treatment, improving quality of life, and reducing serious consequences such as desire for hastened death and suicide.

Prevalence of Depression in Cancer Patients. Many studies have documented the prevalence of depression in cancer patients. The prevalence estimates of major depressive disorder in cancer patients have varied widely in different studies, from a low of 1 percent to a high of 50 percent. The use of different diagnostic measures and cutoff criteria impacts prevalence estimates. The site of cancer, the physical symptoms, and the stage of cancer are additional factors that contribute to differences in prevalence. In a study of terminally ill cancer patients, a symptom threshold consistent with DSM-IV-TR criteria was associated with a depression diagnosis in 13 percent of patients. A relatively minor reduction in the symptom severity threshold elevated the depression diagnosis to 26.1 percent of patients. A 2011 meta-analysis that stratified for various diagnostic systems and stages of disease reported a 25 percent prevalence rate for depressive or adjustment disorders and 29 percent for all types of mood disorders among cancer patients.

Risk Factors for Depressive Disorders in Cancer Patients. There are several risk factors for depression in cancer patients. The site of cancer and several chemotherapy regimens tend to increase the risk. Vinblastine, vincristine, interferon, procarbazine, asparaginase, tamoxifen (Nolvadex), cyproterone (Androcur), and corticosteroids are associated with greater risk. A higher prevalence of depressive disorders has been found among patients with pancreatic, head and neck, breast, and lung cancers, with

relatively lower rates observed among those patients with lymphoma, colon, and gynecological cancers. Depression is a common symptom of pancreatic cancer, with some early data suggesting that the mood disturbance is mediated by alteration of brain serotonergic function through proinflammatory cytokines (Table 27.8–5).

Depression may also develop secondary to organ failure or from nutritional, endocrine, and neurological complications of cancer. Other risk factors for depressive disorders are advanced disease stage and physical disability, presence of other chronic medical illnesses, a previous history of depression, family history of depression, uncontrolled pain, low social support, social isolation, recent experience of a significant loss, and low self-esteem. Although it is well established that depression is more prevalent among women than men in the general population, the gender difference is not evident among cancer patients. Older individuals are at greater risk for depression even in the absence of medical illness.

Evaluation and Diagnosis of Depression in Cancer Patients. Diagnosis of depression is challenging in cancer patients due to the neurovegetative symptoms that mimic many symptoms caused by cancer or its treatment, such as loss of appetite, fatigue, sleep disturbances, psychomotor retardation, apathy, and poor concentration. The assessment of depressive symptoms in cancer patients should focus on the presence of dysphoria, anhedonia, hopelessness, worthlessness, excessive or inappropriate guilt, and suicidal ideation. Presence of these symptoms helps distinguish depression from cancer-related symptoms. Changes in appetite, sleep disturbances, fatigue, and loss of energy can still be used to diagnose the cancer patient who is physically well, although one must bear in mind that chemotherapy- and radiation therapy–induced fatigue and other vegetative disturbances continue for a long time. Poor memory and impaired concentration are more likely to be the initial chief complaints in elderly depressed patients. Although rare, delusions or hallucinations may accompany depression. In medically ill depressed patients, presence of delusions and hallucinations might be reflective of a diagnosis of delirium, which should be ruled out first. If present, the diagnosis of delirium precludes the diagnosis of a mood disorder with depressive features.

Table 27.8–5.
Medical Causes of Mood Disorders in Cancer Patients

Electrolyte disturbances
Vitamin B_{12}, folate, or other vitamin deficiencies
Cancer-related
Primary CNS tumors
CNS metastasis
Paraneoplastic syndromes

	Pancreatic cancer, small cell lung cancer, breast cancer, carcinoid tumor, adrenocortical tumors
	Neurological diseases
	Cerebrovascular disease
	Dementia
	Cerebral trauma
	CNS infections
Systemic disorders	
	Autoimmune disorders
	Inflammatory disorders
	Infections
Endocrine abnormalities	
	Hyperthyroidism or hypothyroidism
	Adrenal insufficiency
Medications	
	Corticosteroids
	Interferon and interleukin-2
	Cardiac and antiarrhythmic drugs
	Antibiotics, antivirals
	Psychotropic drugs (antipsychotics, benzodiazepines, other sedative/hypnotics)
	Opiates
	Cimetidine
	Levodopa, methyldopa
	Pentazocine
	Tamoxifen
	Analgesics and anti-inflammatory drugs
	Some chemotherapeutic agents: vincristine, vinblastine, procarbazine, asparaginase, tamoxifen, cyproterone, mithramycin, L-asparaginase
Other conditions	
	Cardiorespiratory disease
	Renal disease, anemia, uncontrolled pain

CNS, central nervous system.

Four conceptual approaches have been proposed to evaluate depression in the medically ill: inclusive, exclusive, etiological, and substitutive approaches. No single approach is inherently superior to the others. When selecting a conceptual approach, it is important to consider whether a diagnosis of depression is being made for clinical or research purposes. The inclusive approach, in which all symptoms of depression are counted regardless of their suspected etiology, is highly sensitive, making its use more appropriate for diagnoses in clinical settings. The substitutive approach replaces somatic symptoms of depression (i.e., fatigue, sleep disturbance, change in weight or appetite, difficulty concentrating) with psychological symptoms such as depressed appearance, social withdrawal, brooding, self-pity, pessimism, and anhedonia. This approach is conceptually reasonable, but there is little evidence for its superiority. The etiological approach requires that the clinician include a symptom as part of a clinical depression only if it is clearly not a result of medical illness. It

is theoretically sound; however, it is difficult to apply this reliably in oncology settings due to the overlap of symptoms. The exclusive approach eliminates the somatic symptoms from the diagnostic criteria. The high specificity associated with the exclusive approach makes it valuable for research purposes, but it is not desirable in clinical practice due to the risk of missing a diagnosis of depression.

Several subcategories of DSM-5 depressive disorders diagnoses are found in the context of cancer. A diagnosis of depressive disorder due to another medical condition is the appropriate diagnosis when the depressive disorder is due to an underlying cancer, such as pancreatic cancer. When a medication (such as interferon) is the underlying cause of a depressive disorder, the diagnosis of substance/medication-induced depressive disorder is used. Corticosteroids are used in a wide range of situations and can result in depression to the extent that patients may become suicidal or psychotic. Establishing the etiology of symptoms of depression in cancer patients is difficult, and usually multiple factors contribute to depressive symptoms in most of these patients. Self-report symptoms inventories, such as the Patient Health Questionnaire-9 (PHQ-9) or the Beck Depression Inventory (BDI), are useful for screening purposes but do not discriminate between a major depressive disorder and depressive disorder due to another medical condition. Major depressive disorder is defined as an episode of clinically significant persistent and pervasive depressed mood or anhedonia, accompanied by cognitive and behavioral symptoms. The DSM-5 criteria for major depression are consistent with the etiological approach in which symptoms that are deemed to be due to a medical condition or medication do not count toward the diagnosis. Patients who have had a major depressive episode in the past are at risk of a recurrence in the context of cancer. Patients may have a pre-existing dysthymic disorder, which may become more manifest or worse in cancer patients.

Management of Depression in Cancer Patients. Management of depression in cancer patients requires a comprehensive approach that addresses the evaluation, treatment, and follow-up. The initial management begins with the establishment of a therapeutic alliance with the clinician and recruiting support from family or friends. It is important to note that a relationship with the oncologist is the key component of support for the cancer patients. Maintaining ongoing contact with the depressed cancer patient, especially terminally ill patients, ensures that continual evaluation and caring are available.

The NCCN has applied the APA practice guidelines for the treatment of depressive disorders in physically healthy individuals to the treatment of depression in cancer patients. There are several pharmacological and psychotherapeutic strategies available. Prior to selecting an appropriate treatment, the site of cancer, current cancer treatment, comorbid medical

conditions, and medications should be taken into consideration, any of which may contribute to depressive symptoms. Reversible conditions such as thyroid function abnormalities should be ruled out. If the depressive disorder is thought to be secondary to medical condition or caused by a drug, the clinician should treat the underlying condition or change the drug; however, antidepressants are usually started concurrently to relieve the patient's suffering as quickly as possible.

Pharmacological Treatments. The use of antidepressant medications in cancer patients creates unique challenges. A rapid onset of action is preferable in cancer patients, especially in the terminally ill; however, antidepressants may take several weeks to have a therapeutic effect due to their delayed onset of action. An appropriate antidepressant should be selected based on the potential side effects of each antidepressant, drug–drug interactions, the patient's prognosis, primary symptoms of depression, and comorbid condition. Antidepressants should be started at low doses and titrated up slowly in medically frail cancer patients, especially in the elderly.

SSRIs have become the first line of treatment for depressive disorders in the medically ill. They are efficacious, generally well tolerated, and are not as toxic in overdose as TCAs. Some SSRIs, such as fluoxetine (Prozac) and fluvoxamine (Luvox), are inhibitors of cytochrome P450 isoenzymes. It is therefore important to monitor for the possibility of drug–drug interactions. Sertraline (Zoloft), citalopram (Celexa), and escitalopram (Lexapro) are less protein-bound and may have a lower risk of drug interactions with the P450 system. Citalopram (Celexa) causes dose-dependent QT interval prolongation and has an FDA-restricted maximum dose of 20 mg daily for subgroups of patients, including those taking a cytochrome P450 2C19 inhibitor and those older than 60 years. Many of the SSRIs are available in liquid forms, making it easier for patients who cannot swallow pills.

SNRIs are a growing class of antidepressants, which includes venlafaxine (Effexor), desvenlafaxine (Pristiq), duloxetine (Cymbalta), and levomilnacipran (Fetzima).They are generally well tolerated, with a benign side-effect profile similar to SSRIs. However, norepinephrine reuptake inhibition may result in palpitations and hypertension. Therefore, blood pressure monitoring is recommended for patients on an SNRI. It is important to note that venlafaxine mainly inhibits serotonin reuptake at low doses; its effect on norepinephrine reuptake inhibition is seen at doses higher than 150 mg per day. Duloxetine shows serotonin and norepinephrine reuptake inhibition at starting doses. All four medications have low P450 effect and moderate plasma protein binding, which minimize drug–drug interactions and allows for their common use in depressed cancer patients. These medications have proven effective for

treatment of neuropathic pain and are commonly prescribed as adjunct pain medications.

TCAs have been around for many years and are less expensive than many of the SSRIs. Because of their anticholinergic, antiadrenergic, and antihistaminergic side effects, they are less frequently used in cancer patients. Anticholinergic side effects should be especially avoided in elderly patients at risk for anticholinergic delirium or patients at risk for GI dysmotility or obstruction. Their role as adjunct pain medications, especially for neuropathic pain, has become their most common indication for use in cancer patients.

MAOIs are rarely used in the treatment of cancer patients with depression, due to the risk of fatal hypertensive crisis when concurrently used with tyramine-rich food or sympathomimetic drugs.

Of the other antidepressants, bupropion (Wellbutrin) acts primarily on the dopamine system and may have a mild stimulant effect, which can be beneficial for cancer patients with fatigue or psychomotor retardation. It is generally tolerated well in the medically ill. Bupropion is associated with an increased risk of seizures at higher doses and should be used with extreme caution in individuals with CNS tumors or seizure disorders. Nefazodone (Serzone) and trazodone (Desyrel) block postsynaptic serotonin 5-HT_2 (5-hydroxytryptamine) receptors. Nefazodone is uncommonly used due to concerns about hepatic failure. Trazodone is often used in lower doses as a sleep aid, rather than a primary antidepressant because its most common side-effect is sedation. Vilazodone (Viibryd) is a serotonin reuptake inhibitor with 5-HT_{1A} receptor partial agonism; it has less adverse impact on sexual function than other antidepressants. Mirtazapine (Remeron) acts by blocking the 5-HT_2, 5-HT_3, and α-2 adrenergic receptor sites. Its side effects of sedation and weight gain are beneficial for many cancer patients with insomnia and weight loss. It also has antiemetic properties. It is also available in a dissolvable tablet form, which is particularly useful for patients who cannot swallow or who have difficulty with nausea and vomiting. Mirtazapine has minimal drug–drug interactions. Vortioxetine (Brintellix) is a serotonin modulator and stimulator that is approved for the treatment of MDD. It is rarely used in patients with cancer due to high risk of serotonin syndrome when used alone or in combination with other antidepressants, opioids, or CYP2D6 inhibitors.

Psychostimulants and wakefulness-promoting agents are helpful to treat depressed cancer patients' symptoms of fatigue, psychomotor retardation, and poor concentration. Psychostimulants exert dopaminergic effects. Psychostimulants used in cancer patients include dextroamphetamine (Dexedrine), methylphenidate (Ritalin or Concerta), and lisdexamfetamine (Vyvanse). They have a major advantage over antidepressants due to their rapid onset of action and their role in relieving

fatigue, promoting wakefulness, and countering opioid-related sedation. Side effects may be anorexia, anxiety, insomnia, euphoria, irritability, and mood lability. However, side effects are not common at low doses and can be avoided by slow titration. Hypertension and cardiac complications can occur; thus, it is advisable to monitor cardiac function. For depressed cancer patients with short survival expectancy, psychostimulants provide rapid relief from distressing depressive symptoms. Modafinil (Provigil) and armodafinil (Nuvigil) are known as wakefulness-promoting agents. They produce increased alertness, wakefulness, and energy. They are presumably better tolerated than the psychostimulants, but may also cause anxiety, restlessness, and insomnia. They should be used with caution in patients with poorly controlled hypertension.

Psychotherapy. Several psychotherapeutic techniques have been successfully employed with depressed cancer patients. Psychotherapy is often combined with a pharmacological intervention. Supportive psychotherapy and cognitive–behavioral therapy are commonly used in cancer patients. MCP is a novel therapy modality that has been shown to improve depressive symptoms among advanced cancer patients both in individual and group therapy settings.

Group therapy can be helpful to improve social networks, connecting the patient with others who have the same diagnosis or treatment, decreasing the patient's sense of isolation. Supportive–expressive and cognitive–existential group psychotherapies have been studied and used successfully for cancer patients.

Electroconvulsive Therapy. ECT is an effective treatment modality for depressed patients. ECT should be considered in patients who are refractory to psychopharmacological treatment (i.e., at least two trials with antidepressant medicines from different pharmacological classes that are adequate in dose, duration, and adherence), have severe weight loss secondary to depression, exhibit acute psychosis, or have a high suicide risk. Although CNS tumors and cardiac problems are not absolute contraindications, ECT is used with caution in these individuals.

Suicide Assessment and Management in Cancer Patients. The incidence of suicide is higher in cancer patients compared with the general population. Studies suggest that although a small number of cancer patients commit suicide, the relative risk of suicide in this population far exceeds that of the general population, particularly early after diagnosis. Compared to cancer-free persons, the relative risk of suicide after cancer diagnosis is 12.6 in the first week and 3.1 during the first year; the risk elevations lower rapidly in subsequent years after diagnosis. Suicide may also occur in advanced cancer patients with escalating depression,

hopelessness, and the presence of poorly controlled symptoms, particularly pain. Suicidal thoughts in patients with advanced disease, poor prognosis, or poorly controlled symptoms should not be viewed as rational. It is important to note that those patients may have a treatable major depressive episode influencing their suicidal ideation. Clinicians should evaluate for hopelessness and a diagnosis of depression in terminally ill patients with persistent desire for death or suicidal intention (Table 27.8–6).

Prior history of psychiatric illness, previous history of depression or suicide attempts, recent bereavement, history of alcohol or other substance abuse or dependence, male gender, family history of depression or suicide and lack of family or social support, and recent losses are common risk factors for suicide. Untreated delirium may lead to unpredictable suicide attempts due to impaired judgment and impulse control. Older patients and individuals with head and neck, lung, breast, urogenital, GI cancers, and myeloma seem to have an increased risk of suicide. An international population-based study from Denmark, Finland, Norway, Sweden, and the United States has shown a small, but statistically significant, increased long-term risk of suicide, 25 or more years after a breast cancer diagnosis.

Table 27.8–6.
Suicide Risk Factors in Cancer Patients

Depression
Hopelessness
Uncontrolled pain
Extreme fatigue
Anxiety
Delirium
Substance abuse or dependence
Feeling alone in the world
Objective lack of social support
Stressful family relations
Intolerance of loss of control, dependence, and loss of dignity
Severe financial problems
Positive family history
History of severe trauma
Presence of active statements and plans
Personality disorder
Previous history of depression and suicide attempts

Evaluation of suicidal thoughts should take into account the disease stage and prognosis. It is helpful to consider the issue of suicidality from four perspectives: (1) the suicidal thoughts that occur transiently in many patients with cancer, (2) suicidal thoughts in patients who are in remission with a good prognosis, (3) suicidal thoughts in patients with poor prognosis

or poor symptom control, and (4) patients in terminal stages. Suicidal ideation or plans to commit suicide in cancer patients with good prognosis or those in remission also require careful assessment. A Finnish survey of suicides in 1 year revealed that 4.3 percent had cancer, half of whom were in remission at the time of suicide. Patients were noted to have had a history of psychiatric illness prior to cancer diagnosis, in particular substance abuse.

It is important to recognize and aggressively treat high-risk patients for depression and address suicidal risk with acute psychiatric hospitalization, if necessary. Maintaining a supportive relationship, symptom control (e.g., pain, nausea, depression), and involving the family or friends are the initial steps in management of a suicidal patient. A recent study examining the suicidal ideation and past suicidal attempt in adult survivors of childhood cancer found a strong correlation between physical health and suicidality, which underscores the importance of symptom control in cancer patients. Early psychiatric involvement with high-risk individuals can often avert suicide in cancer patients. A careful evaluation includes an exploration of the reasons for suicidal thoughts and the seriousness of the risk. The clinician should listen empathically, without appearing critical or judgmental. Allowing the patient to discuss suicidal thoughts often decreases the risk of suicide despite common belief to the contrary. Patients often reconsider and reject the idea of suicide when the physician acknowledges the legitimacy of their option and the need to retain a sense of control over aspects of their death. Rarely, a competent patient presents a difficult problem by continuing to threaten suicide or to ask for help in dying. Patients must be given intense ongoing support, including open discussions about treatment options. Medical staff and family may need to be reminded of the competent patient's right to refuse all treatments, even life-saving ones.

Bipolar Disorder

Cancer patients with bipolar disorder are most likely brought to the attention of consultation psychiatrists due to the complexity of their management. Acutely manic patients must be brought under control as rapidly as possible. The rigors and dangers of chemotherapy and radiation treatment require a level of consistent cooperation that manic patients cannot be trusted to sustain, even when they understand their situation and earnestly desire treatment. These are difficult cases that require patience and multidisciplinary cooperation. Psychiatrists should carefully assess for safety and level of functioning to determine the best treatment setting. It is also important to screen for substance abuse or dependence and to assess whether or not the patient is suicidal or a risk to others (homicidal or violent ideation). When the patient cannot be rapidly controlled, a brief psychiatric admission may be necessary, preferably

followed by a direct transfer from the psychiatric ward to the oncology ward.

In an acutely manic or hypomanic cancer patient (especially in the elderly), it is important to complete a comprehensive medical assessment to rule out medical conditions or medications as the possible culprit of the mood disturbance. Corticosteroids can cause hypomania or even mania with psychotic features especially in patients with prior history of mood disorder or a family history of bipolar disorder. The spectrum of mental disturbances related to steroids includes minor mood lability, affective disorders (mania or depression), cognitive impairment, and delirium. The incidence of these disorders range from 28 to 55 percent in noncancer populations, and they occur most commonly on higher doses. Symptoms usually develop within the first 2 weeks on steroids, but in fact can occur at any time, on any dose, even during the tapering phase. These disorders are often rapidly reversible upon dose reduction or discontinuation. Manic symptoms can also be a part of postictal psychosis most likely in patients with temporal lobe epilepsy or after temporal lobectomy. An EEG may be helpful for diagnostic clarity. Thyroid function abnormalities should also be ruled out, especially in patients with thyroid carcinomas or those who have undergone radiation to the neck.

Cancer patients with bipolar disorder in remission can be maintained on lithium (Eskalith) or other mood-stabilizing medications. Renal insufficiency is a contraindication to lithium use, but end-stage renal disease can be managed with low-dose lithium after each dialysis. Medically ill patients must be monitored for lithium toxicity in the presence of vomiting and dehydration (as during chemotherapy), hypercalcemia, renal impairment (as with cisplatin [Platinol] treatment), possible hypothyroidism, or when there is an increased risk of cardiac arrhythmia (as in preexisting heart disease or with the use of digoxin [Lanoxin], bronchodilators, or cardiotoxic chemotherapy agents, such as doxorubicin [Adriamycin]). Radiation therapy is not a contraindication, except for cranial irradiation, whose side effects mimic lithium toxicity, making monitoring difficult. Patients should be monitored carefully during postoperative period for development of nephrogenic diabetes insipidus (NDI), an uncommon but severe side effect. Patients may not report their symptoms of polyuria, polydipsia, or nocturia, as the insidious onset of the condition leads patients and families to believe that the increased intake and output is normal. The problem may go unrecognized for long periods if physicians do not specifically question their patients about the existence of polyuria, polydipsia, or nocturia. Serum lithium levels might be in the therapeutic range despite polyuria. Patients may present with severe hypernatremia during acute illness or after fluid deprivation in perioperative period. Annual measurement of 24-hour urine volume may help prevent this potentially fatal condition in cancer patients using lithium

for treatment of bipolar disorder.

Cancer-Related Fatigue

Cancer-related fatigue is a persistent, subjective sense of tiredness related to cancer or cancer treatment that interferes with usual functioning. Fatigue is a common and highly distressing symptom of cancer, associated with reduced quality of life, as well as considerable psychological and functional morbidity. Fatigue is more distressing than pain or nausea and vomiting, which are manageable by medications in most patients. Cancer-related fatigue is usually refractory to sleep and rest. The reported prevalence of cancer-related fatigue ranges from 4 to 91 percent depending on the specific cancer population studied and the methods of assessment. Fatigue is present at the time of diagnosis in about 50 percent of cancer patients. It occurs in up to 75 percent of patients with bone metastases. Fatigue is also prevalent in long-term cancer survivors and has a serious impact on a person's quality of life.

A group of expert clinicians proposed a set of diagnostic criteria that have been included in the 10th edition of *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) (Table 27.8–7). A standardized interview guide has been designed and studied for its reliability and validity to assess the proposed clinical syndrome, with results suggesting its utility in identifying patients experiencing clinically significant cancer-related fatigue. Fatigue is difficult to quantify. There are a variety of standardized self-report scales, most of which have been developed in the context of cancer. Different scales may measure fundamentally different aspects or even potentially distinct conceptions of fatigue. The challenge facing the clinician or researcher is to choose a reliable and valid tool for the measurement of fatigue that is most adequately suited to their purposes. Given the multifactorial nature of fatigue, accessory scales (e.g., depression scales) and measurements of certain biological parameters should be used in addition to fatigue assessment tools in order to obtain the most complete evaluation of a patient's fatigue.

Table 27.8–7.
ICD-10 Criteria for Cancer-Related Fatigue

A. Six (or more) of the following symptoms have been present every day or nearly every day during the same 2-week period in the past month, and at least one of the symptoms is (A1) significant fatigue:
   - A1. Significant fatigue, diminished energy, or increased need to rest, disproportionate to any recent change in activity level
   - A2. Complaints of generalized weakness or limb heaviness
   - A3. Diminished concentration or attention
   - A4. Decreased motivation or interest to engage in usual activities

A5. Insomnia or hypersomnia
A6. Experience of sleep as unrefreshing or nonrestorative
A7. Perceived need to struggle to overcome inactivity
A8. Marked emotional reactivity (e.g., sadness, frustration, or irritability) to feeling fatigued
A9. Difficulty completing daily tasks attributed to feeling fatigued
A10. Perceived problems with short-term memory
A11. Postexertional malaise lasting several hours

B. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
C. There is evidence from the history, physical examination, or laboratory findings that the symptoms are a consequence of cancer or cancer therapy.
D. The symptoms are not primarily a consequence of comorbid psychiatric disorders such as major depression, somatization disorder, somatoform disorder, or delirium.

From World Health Organization. *International Statistical Classification of Diseases and Related Health Problems.* 10th revision, 2nd ed. Geneva: World Health Organization; 2004, with permission.

Fatigue and depression may coexist in cancer patients, and there is considerable overlap of symptoms of these two conditions. It is necessary to clarify the relationship between depression and fatigue in order to effectively evaluate and treat cancer-related fatigue. There are many symptoms common to both fatigue and depression, such as decreased energy and motivation, sleep disruptions, diminished concentration and attention, and problems with short-term memory. The clinician seeking to differentiate between depression and fatigue should remember that depressive symptoms due to fatigue are typically less severe, and that patients tend to attribute such symptoms to the consequences of fatigue. Depression, on the other hand, is more likely in the presence of hopelessness, feelings of worthlessness or guilt, suicidal ideation, and a family history of depression. The nature of any causal relationship between cancer-related fatigue and depression is unclear.

Cancer-related fatigue is caused by a number of physical and psychosocial mechanisms, including tumor by-products, opioids or other drugs (such as antidepressants, β-blockers, benzodiazepines, antihistamines), hypogonadism, hypothyroidism, cachexia, anemia, chemotherapy, radiation therapy, bone marrow transplantation, and treatment with biological response modifiers. Cancer-related fatigue has been related to pain, depression, emotional distress, sleep deprivation, and reduced physical activity. Cytokines (interleukin-1 and -6 [IL-1, IL-6], tumor necrosis factor-α [TNF-α]) have a role in the development of cancer-related fatigue, which has led researchers to consider cytokine-antagonist drugs such as TNF receptor etanercept (Enbrel) and TNF-α antagonist thalidomide (Thalomid) to improve tolerability of chemotherapy regimens and potentially to treat cancer-related fatigue and cachexia. Genetic variables have also been implicated to play a role in the development of fatigue among cancer patients. Advanced colorectal cancer patients with two variant forms of the dihydropyrimidine dehydrogenase (*DPYD)* gene—

an enzyme that is involved in the breakdown of excess nucleotides involved in DNA, RNA, and ATP production—were significantly less likely than those patients with a form of the gene known as *DPYD*5* to report fatigue following treatment with a chemotherapy regimen of 5-fluorouracil, irinotecan, and oxaliplatin.

All patients should be screened for fatigue at their initial visit, at regular intervals during and following cancer treatment, and as clinically indicated. The NCCN practice guidelines on cancer-related fatigue recommend the use of numerical self-report scales or verbal scales to assess the severity of fatigue. If moderate or severe (a score of 4 to 10 on a scale of 0 to 10 with higher numbers indicating increased severity), a focused history and clinical examination with evaluation of the pattern of fatigue, associated symptoms, and interference with normal functioning are recommended. Potentially reversible causes of fatigue (such as pain, emotional distress, sleep disturbance, anemia, hypothyroidism) should be identified and treated, and nonessential centrally acting drugs should be eliminated. If anemia is the main cause of fatigue, the physician should determine the necessity of a transfusion in severely symptomatic patients. Clinical trials have shown that anemic patients have improved energy and less fatigue after erythropoietin treatment.

Nonpharmacological approaches have been recommended by NCCN guidelines for the treatment of cancer-related fatigue. Activity enhancement and psychosocial interventions (i.e., education, support groups, individual counseling, stress management training) have been well supported by research. Dietary management, attention-restoring therapy, sleep therapy, and mind–body interventions have also been recommended in the treatment of cancer-related fatigue.

A number of pharmacological interventions are being evaluated in clinical trials for the treatment of fatigue in cancer patients (http://clinicaltrials.gov/). Psychostimulants, wakefulness promoting agents, and antidepressants are the psychotropic medications used in the treatment of cancer-related fatigue.

Psychostimulants have also been used in treatment of fatigue-related conditions such as pain, depression, and cognitive impairment. Methylphenidate has been studied most and appears to be relatively effective and well tolerated despite relatively common side effects such as insomnia and agitation. A randomized, double-blind, placebo-controlled trial of methylphenidate and pemoline (Cylert) for the treatment of fatigue in ambulatory patients with human immunodeficiency virus (HIV) disease has found that both medications were equally effective and significantly superior to placebo in decreasing fatigue severity with minimal side effects. Improvement in fatigue was also associated with improved quality of life, decreased depression, and decreased psychological distress. A pilot study examined the effects of exercise and methylphenidate on fatigue, functional

ability, and cognitive function in patients with melanoma and compared results with historic controls who received usual care only while on interferon α. Patients were instructed to take 20 mg per day sustained-release methylphenidate and follow an exercise program 4 days a week for 15 to 30 minutes. Fatigue was lower for the exercise and methylphenidate group than historic controls. The researchers concluded that the combination of aerobic exercise and methylphenidate may have a positive effect on fatigue, cognitive function, and functional ability. An open-label pilot study using patient-controlled methylphenidate for the management of cancer-related fatigue has shown that methylphenidate was temporarily associated with improvement in fatigue, overall well-being, and depression. Following the open-label study, the same group of researchers conducted a double-blind, randomized, placebo-controlled trial comparing patient-controlled methylphenidate with placebo. Fatigue intensity improved significantly on day 8 in both the methylphenidate and placebo groups. However, there was no significant difference in fatigue improvement between the intervention (n = 52) and placebo (n = 53) groups. Researchers concluded that in the absence of a placebo-controlled group, it is difficult to determine whether these findings reflect an independent or an extension of placebo effect. Dexmethylphenidate is the d-isomer of methylphenidate, with a longer duration of action (about 6 hours) compared to methylphenidate. It has been found to be well tolerated and more effective than placebo in improvement of fatigue in adult cancer patients.

Dextroamphetamine is the d-isomer of amphetamine and is more potent than methylphenidate. Dextroamphetamine has not been studied in the treatment of cancer-related fatigue; however, it is commonly used by clinicians for this indication. It has been studied in the treatment of HIV-related fatigue with favorable results.

A meta-analysis of pharmacological treatment options for cancer-related fatigue has concluded that methylphenidate might be effective for treating fatigue. A significant placebo response has been observed in randomized trials. There was also evidence that treatment with hematopoietic agents relieved fatigue due to chemotherapy-induced anemia.

Modafinil is a wakefulness-promoting agent, approved by the FDA in December 1998 for treatment of excessive daytime sleepiness in patients with narcolepsy. Compared to other psychostimulants, modafinil has a novel mechanism of action and less abuse potential. It presumably enhances activity in hypothalamic wakefulness center (i.e., tuberomammillary nucleus), activates tuberomammillary nucleus neurons that release histamine, and also activates other hypothalamic neurons that release orexin/hypocretin. It is currently used as a first-line agent for the treatment of severe fatigue in multiple sclerosis. Modafinil is commonly used for the treatment of severe fatigue in multiple sclerosis and has been

studied as a treatment option for cancer-related fatigue with improvement of fatigue. The results are supportive of its use among cancer patients with severe fatigue. Further well-designed, randomized, controlled clinical trials are necessary to clarify the role of modafinil and armodafinil in the treatment of cancer-related fatigue.

Table 27.8–8.
Psychotropic Medications Used in Cancer-Related Fatigue

Medication	Starting Dose	Dose Range	Comments
Psychostimulants			
Methylphenidate	2.5–5 mg daily or twice daily	5–30 mg/d, usually divided as twice daily	Longer-acting forms are available. Capsule forms can be sprinkled in food.
Dextroamphetamine	2.5–5 mg daily or twice daily	5–30 mg/d, usually divided as twice daily	Longer-acting formulations are available. Capsule forms can be sprinkled in food.
Wakefulness promoting agents			
Modafinil	50–100 mg daily	50–400 mg daily, may be divided as twice daily	Favorable side-effect profile
Antidepressants[a]			
Selective serotonin reuptake inhibitors			
Well tolerated. Citalopram, escitalopram, and sertraline seem to have the fewest drug–drug interactions.			
Fluoxetine[b]	10–20 mg/d	10–60 mg/d	
Paroxetine	10–20 mg/d	10–40 mg/d	
Citalopram[b]	10–20 mg/d	10–40 mg/d	
Escitalopram[b]	5–10 mg/d	5–20 mg/d	
Sertraline[b]	25–50 mg/d	25–200 mg/d	
Serotonin–norepinephrine reuptake inhibitors			
Well tolerated. Monitor blood pressure, especially at higher doses.			
Venlafaxine	37.5–75 mg/d	37.5–225 mg/d	
Duloxetine	20–30 mg/d	20–60 mg/d	
Norepinephrine–dopamine reuptake inhibitor			
Doses higher than 300 mg/d should be used twice daily to minimize the risk of seizures.			
Bupropion	75 mg/d	75–450 mg/d	
Tricyclic antidepressants			
Drug–drug interactions are more common compared to SSRIs.			
Secondary Amines			
Desipramine	25–50 mg/d	25–200 mg/d	
Nortriptyline[b]	10–25 mg/d	10–150 mg/d	
Tertiary Amines			
Amitriptyline	10–25 mg/d	10–150 mg/d	
Doxepin[b]	25–50 mg/d	25–300 mg/d	

Imipramine	10–25 mg/d	10–200 mg/d
α-2 antagonist/5-HT_2/5-HT_3 antagonist		
Mirtazapine	7.5–15 mg/d	7.5–45 mg/d

[a]Antidepressant medications have not been proven to be efficacious in the treatment of cancer-related fatigue in the absence of a diagnosis of clinical depression.

[b]Available in liquid formulations.

Underlying depression should be treated with SSRIs, which are generally better tolerated than TCAs in cancer patients. When prescribing such medications, clinicians should remember that, as fatigued patients are generally more sensitive to antidepressant side effects, treatment should be initiated with small doses. Drug–drug interactions should also be carefully monitored. Studies have examined the role of paroxetine, sertraline, and bupropion in the treatment of cancer-related fatigue. Paroxetine and sertraline were effective in improving fatigue among cancer patients with comorbid depressive symptoms. Bupropion was found to be effective and well tolerated in both depressed and non-depressed cancer patients in open-label trials. However, controlled studies are required to determine whether the effect of bupropion on fatigue is independent of its antidepressant effects.

Corticosteroids have been used in the treatment of cancer-related fatigue. Current evidence on the use of corticosteroids is anecdotal and responses to corticosteroids appear to be temporary. Due to concerns for toxicity associated with long-term administration, use of corticosteroids should be restricted to the terminally ill, with concomitant fatigue and anorexia, and/or those with pain related to bone or brain metastases.

Megestrol acetate (Megace) is a progestational agent that has been found to improve appetite in cancer-related cachexia and has also been suggested to have a role in the treatment of cancer-related fatigue. The effects of megestrol acetate on fatigue are not clear, but they probably involve anticytokine and corticosteroid-type effects (Table 27.8–8).

Personality Disorders

Personality disorders are chronic character styles that lead to clinically significant interpersonal and behavioral problems. Personality affects the way an individual copes with different life events. This aspect is important in the medical setting because it will influence how the patient will cope with cancer and its concurrent stressors such as hospitalization and treatment. During the evaluation, it is important for the consultant psychiatrist to identify maladaptive responses to stressors, defense mechanisms that the patient might be using to deal with cancer. Patients' symptoms are chronic and ego-syntonic, and it is the family and staff who feel stressed by a worsening of the patient's usual dysfunctional modes of

adaptation and who ask for help because of increased conflict and disrupted treatment. Personality-disordered patients do have acute stress syndromes, which may be overlooked because of the chronic issues. Evaluation should search for other psychiatric disorders. Comorbid anxiety or depression should be actively treated. Environmental measures should be implemented to decrease conflict and to improve patient care, including family and staff involved in the intervention. Patients with personality disorders are also distressed by unbearable emotional conditions and might require attention to basic needs such as sleep deprivation, hunger, and necessary medical and comfort care. Medications such as benzodiazepines or antipsychotics can be used to stabilize agitation or emotional upheaval. The psychiatrist should establish a working relationship by being empathetic with the patient and also by being aware of transference and countertransference issues. The dynamics of difficult patients and ways of dealing with them have been well described in the psychiatric literature. The feelings of anger and guilt that patients with personality disorders routinely induce in mental health staff also occur in oncology staff. The latter, especially when inexperienced, may find these feelings harder to deal with, because they are not trained to recognize these feelings as part of the patient's pathology and are disarmed by the genuine pathos of the patient's condition. Explanations, support, and guidance are helpful.

Schizophrenia and Other Psychotic Disorders

Patients with schizophrenia may appear less often in doctors' offices, because their disease interferes with the early and consistent seeking of care. They give idiosyncratic meanings to their symptoms and may tolerate them without complaint, often because of what is called a *high pain threshold*, although a more accurate description would be a different central processing mechanism. The patients' unreliability continues throughout the treatment course and may complicate postoperative and long-term medical management. When treating hospitalized patients with schizophrenia, it is important to ensure continuity of care. Ongoing medications can usually be continued, except for a brief operative period. The medical staff needs help in handling bizarre behaviors and reassurance about intimidating mannerisms. They need a precise protocol and written orders for the management of aggressive or disruptive episodes, as well as a clear understanding of the legal and ethical considerations governing the patients' right to refuse medical treatment and to leave the hospital. Decisional capacity is situation specific and independent of the presence of a psychiatric disorder; hence, psychiatric consultation should be called to help establish its presence or absence in any conflictual situation.

Substance Use Disorders

In 2012, an estimated 23.9 million Americans aged 12 or older (9.2 percent of the population) reported they had used an illicit drug or misused a prescribed medication (opioid, stimulant, or benzodiazepine/tranquilizer) and 17.0 million Americans (6.5 percent) reported heavy drinking in the past month. While prevalence data for substance use disorders in cancer patients is limited, one can infer that a significant proportion of patients with new cancer diagnosis have problematic drug and alcohol use.

Oral, pharyngeal, laryngeal, esophageal, breast, liver, and rectal cancers are more common in alcohol users than in non-alcohol users. Smokers who also drink are at much higher risk. Alcohol withdrawal syndrome can greatly affect a patient's cancer treatment. During acute hospitalization, alcohol withdrawal may commence on the second or third day and elicit a psychiatric consultation for anxiety, poor coping, or postoperative delirium. Thoracic and head and neck cancer patients are at high risk for this and for nicotine withdrawal, given their high incidence of smoking. The symptoms of nicotine withdrawal are those of anxiety, irritability, and insomnia. Transdermal nicotine patches are commonly used to ease the distress of patients forced to stop smoking suddenly. A clinician must be mindful that cigarette smoke includes polycyclic aromatic hydrocarbons that contribute to induction of several cytochrome P450 enzymes (specifically CYP1A1, CYP1A2, and CYP2E1) and induction of glucuronidation of some drugs. Abrupt cessation of smoking can affect medication levels; this effect is independent of nicotine and not countered with nicotine patches.

Managing cancer pain in addicted patients requires a multidisciplinary approach with involvement of a psychiatrist, a palliative care physician, social workers, and nurses. The psychiatrist should manage psychiatric comorbidities and the countertransference of the team. It is also important to set limits to control harm with substantial emotional support. Pain management in addicted cancer patients is challenging and should be focused on the level of pain, with attention to the addiction. In treating patients with past or current opioid use disorders, the clinician must be mindful of the effect of tolerance and opioid-induced hyperesthesia, in which cases the patient will require higher doses for pain control. Patients will benefit from long-acting opioids such as methadone (Dolophine) and by limiting the use of short-acting or breakthrough doses. Furthermore, adjuvant medications such as SNRIs or antiepileptics may be helpful. Scheduled urine toxicology screens with clear consequences and contingency plans are used to monitor patient's illicit drug usage.

Somatoform Disorders

The prevalence of somatoform disorders in cancer patients is unknown. The development of cancer in patients with a somatoform disorder can create difficult diagnostic problems, because patients may incorporate the presenting signs and symptoms of cancer into their psychiatric disorder–

associated illness behaviors. This results in complex syndromes with both oncological and psychogenic components. Physicians may feel manipulated and frustrated by these presentations, which may lead to delays in treatment or suboptimal care through over- or under-treatment. A psychiatric evaluation and collaborative effort may allow the psychiatrist to work with the oncology staff to treat the patient appropriately through clarification of presentation, addressing psychiatric comorbidity, and setting limits on maladaptive behaviors.

PSYCHIATRIC SYNDROMES AND TUMOR SITE

CNS Cancers

Brain tumors are classified as primary or metastatic tumors. Brain metastases actually occur up to 10 times more frequently than primary brain tumors. Seventy percent of brain tumors are supratentorial. Primary brain and CNS tumors are relatively uncommon, accounting for approximately 23,380 (1.4 percent) new cancer diagnoses, and 14,320 (2.4 percent) cancer-related deaths annually. Meningiomas, which are usually non-malignant, are the most common tumor type. Glioblastomas are the most common primary malignant brain tumors, accounting for approximately 16 percent of all new primary brain tumors and 45 percent of primary malignant brain tumors. Unfortunately, they are generally associated with dismal prognoses, with 5-year survival rates of <5 percent.

Psychiatric symptoms of brain tumors are influenced by their anatomic location, interconnection of different areas of the brain, rate of tumor growth, premorbid level of functioning, tumor type, and increased intracranial pressure. Psychiatric symptoms are present in almost all patients with supratentorial tumors and are the presenting symptom in 25 percent of the patients. Frontal lobe tumors are associated with behavioral symptoms in 90 percent of patients. Mood symptoms can include anxiety, irritability, depression, apathy, and euphoria. There have been reported changes in personality in 70 percent of patients. These changes include irresponsibility, disinhibition, and inappropriate sexual behavior. Psychotic symptoms including delusions and hallucinations have been reported in 10 percent of patients. Bifrontal disease typically causes impairment of memory, lability of mood, urinary incontinence, and frank dementia, with prominent grasp, suck, and snout reflexes. Psychotic symptoms are common in patients with tumors of the temporal lobe and pituitary gland. Tumors affecting the temporal lobe can present with auditory and olfactory hallucinations as well as abnormal behavior (apathy, irritability expansive mood with mania or hypomania). Nondominant temporal lobe tumors can cause minor perceptual problems and spatial disorientation. Dominant temporal lobe tumors can present with impaired perception of verbal commands and sensory aphasia. Parietal lobe tumors affect sensory and

perceptual functions more than motor modalities. Tumors in either parietal lobe can produce sensory disorders ranging from mild sensory extinction or astereognosis to a more severe sensory loss such as hemianesthesia. Poor proprioception in the affected limb is common. Homonymous inferior quadrantanopsia, an incongruent hemianopsia, or visual inattention also may occur. Nondominant parietal lobe tumors cause contralateral neglect and, in severe cases, anosognosia and apraxia. Dominant parietal lobe tumors lead to alexia, dysgraphia, and certain types of apraxia. Occipital lobe tumors can produce contralateral homonymous hemianopsia or complex visual aberrations affecting perception of color, object size, or location. Bilateral occipital disease, rarely seen at presentation, can produce cortical blindness.

Delirium is a common psychiatric complication in patients with CNS tumors. The main causes of delirium in this population are CNS insults, medications (corticosteroids), and metabolic changes. Dementia and cognitive impairment are also symptoms and complications of treatment of CNS tumors.

Management of Psychiatric Symptoms in Patients with Brain Tumors. Brain tumor patients will benefit from psychopharmacological, psychotherapeutic, and psychosocial interventions. The psychiatrist should work closely with the medical and surgical team to help patient's functioning and improve their quality of life.

The data that support the pharmacological treatment of patients with psychiatric symptoms come from clinical experience and case reports. Medications must be initiated at low doses due to frequent brain insults caused by the cancer and its treatment. Antipsychotics are indicated for treating delirium, anxiety, and organic mood disorders. The psychiatrist should be cognizant of several side effects in this population. Antipsychotics can lower the seizure threshold. Antipsychotics such as clozapine (Clozaril) and chlorpromazine have the greatest potential to lower seizure threshold; thus, they should be avoided. High-potency antipsychotics are indicated if sedation is not desired. Atypical antipsychotics should also be considered in treating psychotic and mood symptoms.

Psychostimulants such as methylphenidate and modafinil can improve attention, mood, and cognitive function. These medications should also be started at low doses and be used with caution due to the potential of causing hallucinations and decreasing the seizure threshold.

Antidepressant medications are effective in treating depression and anxiety. The SSRIs are frequently used due to their low anticholinergic activity. Benzodiazepines for treating anxiety symptoms can induce disinhibition and delirium in patients with brain tumors.

Mood lability can be treated with lithium, carbamazepine (Tegretol),

valproic acid (Depakene), and atypical antipsychotics.

Head and Neck Cancers

There were approximately 52,000 men and women in the United States diagnosed with head and neck cancers in 2012; this accounted for almost 3 percent of all cancer diagnoses. Head and neck cancers are diagnosed almost twice as often in men as in women. Time of diagnosis is more often after age 50, except for salivary gland and nasopharyngeal tumors, which may occur in younger age groups. Cigarette smokers have an increased risk for multiple head and neck primary cancers as well as for lung cancer. Alcohol use has also been implicated as a causative factor for certain head and neck cancers, and the effects of alcohol and tobacco use seem to be additive. Infection with some types of human papillomavirus (HPV), particularly HPV-16, is a risk factor for head and neck cancers, especially oropharyngeal cancers that involve the base of the tongue and tonsils. In the US, the incidence of oropharyngeal cancer associated with HPV infection is increasing. Patients with head and neck cancers have an increased risk for developing esophageal and lung cancers.

Surgery and radiation therapy are curative treatments for carcinoma arising in the head and neck. Chemotherapy is useful in the adjuvant setting; used alone, it is not curative.

The diagnosis of head and neck cancer brings many psychosocial challenges. Patients are confronted with fears of death, pain, and loss. Patients also are concerned with disfigurement and disability from cancer treatment. In anticipation for surgery, patients experience distress and anxiety. An incidence of adjustment disorders as high as 36 to 50 percent has been reported. Recent studies have also demonstrated prevalence rates of clinically significant depression diagnosed by various self-reporting and structured interview techniques of between 17 and 30 percent specifically in head and neck cancer patients. Both pharmacological and psychological interventions are effective in treating depression in these patients.

Another challenge is managing the sensitivity to stigma associated with HPV-positive oropharyngeal cancers as a sexually transmitted disease. Established risk factors for developing HPV-positive oropharyngeal cancers include having greater than six lifetime oral sex partners, concurrent anogenital squamous cell carcinoma, and HIV infection. Furthermore, husbands of women with cervical cancer are at a 3-fold increased risk of HPV-positive oropharyngeal cancer. Psychotherapy may focus on addressing perceived shame of transmission method.

In a recent suicide study among cancer patients, unmarried men with head and neck cancers were found to have the highest incidence of suicide. Other risks factors for suicide are personality disorders, alcoholism, and treatment sequelae such as facial disfigurement.

Postoperative patients are at risk of developing delirium due to alcohol

withdrawal, metabolic disturbances, and other medical complications. Extensive reconstructive operations affect swallowing physiology, leading to aspiration events and situational anxiety.

GI Cancer

Newly diagnosed cancers of the oral cavity, pharynx, and the digestive system account for 21 percent of total cancer incidence and 25 percent of total cancer mortality in the United States.

Patients who suffer from GI tumors have to adjust to several psychosocial changes. Patients have to deal with severe eating disorders, weight loss, abdominal discomfort, and diarrhea or constipation, as well as other disease-related issues such as ostomy and repeated operations.

The psychological issues in patients with colorectal cancer are related to the cancer itself and the impact of treatment of bowel function in relation to social and sexual activity. Patients who require ostomies may interpret this procedure as a significant loss. Different strategies have been recommended to patients with ostomies at different phases of recovery. Patients may feel anxious, hopeless, and helpless when an ostomy is recommended. It is important to involve the family and provide adequate teaching of ostomy procedural management and troubleshooting to minimize distress. Patients may experience shame, low self-esteem, anxiety, or sexual undesirability following their operation. Community ostomy support groups can provide guidance and counseling to patients and their families.

Surgical procedures involving the pelvic area may lead to sexual dysfunction through removal of the posterior wall of the vagina, weakening of muscles of the pelvic floor, or damage to the pudendal and associated nerves. Determining the etiology of sexual dysfunction is challenging, but dictates the potential treatments.

Lung Cancer

Lung cancer is the leading cause of cancer mortality in men and women in the United States, accounting for an estimated 180,000 deaths in 2012. Lung cancer accounts for 20 percent of all cancer deaths worldwide. In the United States, lung cancer is the leading cause of cancer-related death and exceeds deaths from breast, colon, and prostate cancers combined. Tobacco is the primary risk factor for lung cancer. The overall 5-year survival for lung cancer is approximately 18 percent and has improved only marginally over the past 30 years. The major reason for poor survival is the lack of early detection and treatment measures for lung cancer. The most common histological type of cancer is non–small cell (NSCLC) followed by small cell lung cancer (SCLC), which accounts for about 20 percent of all lung cancers.

Patients with lung cancer experience higher emotional distress than patients with other cancers. Emotional distress appears during diagnosis and persists during the course of treatment. Disease-related symptoms such as pain, fatigue, and dyspnea are detrimental to the emotional well-being of patients.

In a study of 333 patients with lung cancer, all patients completed the Distress Thermometer and an associated Problem Symptom List. Over half of the patients reported significant distress. Problems associated with higher level of distress were in the area of emotional functioning, family relationships, and lack of information about the diagnosis, physical functioning, and cognitive functioning. In a study that included 987 patients, the prevalence of depression in lung cancer patients was found to be as high as 33 percent, and in half of those, the depression persisted at the first follow-up visit, 14 to 35 days after beginning chemotherapy for SCLC and 21 to 56 days after starting radiotherapy for NSCLC.

Cognitive disorders have been associated with lung cancer, especially SCLC. These disorders may be related to brain metastases or other comorbidities and complications of the illness and its treatment. Paraneoplastic syndromes, brain radiation, and chemotherapy have been associated with cognitive impairment. Depression can also lead to cognitive impairment, and patients should be screened and treated for this condition.

Smoking is common in lung cancer patients and some studies show that smoking cessation at the time of diagnosis of lung cancer may reduce the rate of development of tumors. One study found that the relative risk of a second lung cancer developing following curative intent therapy of SCLC was lower for those who stopped smoking. Because smoking cessation remains a challenge for those patients, they should be offered intensive tobacco cessation programs, including counseling, behavioral therapy, use of sustained-release bupropion and nicotine replacement, and telephone follow-up, to increase successful abstinence.

Genitourinary Cancers

Prostate Cancer. Prostate cancer is one of the leading causes of cancer death among men in the United States. Prostate cancer screening is a controversial issue, with many arguing that the digital rectal examination (DRE) and PSA may lead to unnecessary and possibly harmful consequences. There has been no established mortality benefit form DRE or PSA screening. The United States Preventive Services Task Force and the Centers for Disease Control and Prevention recommend against the PSA. The American Society of Clinical Oncology and the American College of Physicians discourage screening in those expected to live less than 15 years and recommend others decide after weighing their potential risks and benefits. Treatment options for prostate cancer are surgery and external

beam radiation or radioactive seed implants (brachytherapy). Hormonal therapy, chemotherapy, and radiation are offered for metastatic disease and as supplemental therapies for early-stage disease.

Testicular Cancer. Testicular cancer comprises a morphologically and clinically diverse group of neoplasms, most of which are germ cell tumors (GCTs) and are the most common solid tumors in men between the ages of 15 and 35 years. The pathognomonic presentation of a primary testicular tumor is a painless testicular mass that may range in size from a few millimeters to several centimeters. A radical inguinal orchiectomy minimizes local tumor recurrence and aberrant lymphatic spread. It is the only acceptable therapeutic and diagnostic procedure. GCT is classified into two major subgroups: Seminomas and nonseminomas. Treatment for early stage of seminomatous tumors is orchiectomy and radiation therapy. For moderate and advance disease the treatment is combined radiation and chemotherapy. Nonseminomatous tumors are treated with orchiectomy and combined chemotherapy for more advanced disease.

Bladder Cancer. Bladder cancer is two and a half times more common in males than in females and more common in whites than in blacks. Approximately 53,000 new cases occur per year in the United States, which is a 20 percent increase from 20 years ago. The incidence increases with age and peaks in the sixth and seventh decades of life. More than half of men and one-third of women with bladder cancer have a history of smoking cigarettes. The clinical stage of the primary tumor is determined by transurethral resection of the bladder tumor (TURBT). Surgery, alone or in combination with other treatments, is used in more than 90 percent of cases. Superficial, localized cancers may also be treated by chemotherapy or immunotherapy directly into the bladder. Chemotherapy alone or with radiation before cystectomy has improved some treatment results.

Kidney Cancer. Each year in the United States, there are approximately 31,900 cases of kidney and upper urinary tract cancers, resulting in more than 11,900 deaths. Renal carcinoma may remain clinically occult for most of its course. Approximately 30 percent of patients with renal carcinoma present with metastatic disease, 25 percent with locally advanced renal carcinoma, and 45 percent with localized disease. Surgery is the only known effective therapy for localized renal carcinoma. Adjuvant or palliative nephrectomy is frequently performed in patients with metastatic renal carcinoma.

MANAGEMENT OF PSYCHOLOGICAL ISSUES IN PATIENTS WITH GENITOURINARY CANCERS. The causes of psychological distress in men with genitourinary cancers are related to the diagnosis of cancer, treatment options, potential sexual

dysfunction, urinary incontinence, fatigue, weakness, pain, and other side effects from the cancer and its treatment. Patients in this population have difficulties coping with changes in sexuality, relationships, and lifestyle. Management of these psychological issues includes education, psychotherapy, couple therapy, behavioral interventions, and medications.

GCTs, bladder cancer, and kidney cancers are some of the most common primary cancer sites that can spread to the brain. Metastasis must be on the clinician's differential diagnosis if a patient with one of these cancers presents with delirium or altered mood, personality, or behavior.

Breast Cancer

Breast cancer ranks second among cancer deaths in women. Breast cancer accounted for 29 percent of all new cancer cases in the United States in 2014. Death rates have declined by an average of 2.3 percent per year since the 1990s in all women, with larger decreases in younger (less than 50 years) women. Risk factors for breast cancer include age, inherited genetic mutations (*BRCA1* and *BRCA2*), personal or family history of breast cancer, high breast tissue density, and radiation to the chest. Reproductive factors include a long menstrual history, use of oral contraceptives, never having children, and having first child after age 30. Other factors are obesity, postmenopausal hormone therapy, physical inactivity, and consumption of one or more alcoholic beverages per day. Treatment of breast cancer involves lumpectomy or mastectomy with removal of axillary lymph nodes, radiation therapy, chemotherapy, or hormone therapy such as aromatase inhibitors. These treatments are often used in combination. Breast cancer has a complex psychological impact on women. Three factors have been reported to contribute to this psychological impact: the sociocultural context in which treatment options are offered, the psychosocial and psychological factors, and the medical factors. Breast cancer also has an impact on lifestyle due to physical discomfort, marital or sexual disruption, and disturbed body image. Women with breast cancer are likely to suffer from anxiety and depression. Prevalence of depression among early breast cancer women occurs twice as often as in the general female population, especially during the first year after diagnosis. In a study that included 4,496, the prevalence of depression among patients with breast cancer was estimated to be around 52.65 percent. Management of psychological distress includes providing patients with the skills to cope with their illness and to improve their quality of life. Antidepressant medications and psychotherapy are the gold standard treatments for mood and anxiety disorders in women with breast cancer. Antidepressants are also beneficial in the treatment of hot flashes in breast cancer. Approximately two-thirds of women with a history of breast cancer experience hot flashes due to cytotoxic or hormonal therapies that can cause premature menopause.

The current pathophysiological hypothesis for hot flashes is the dysfunction of hypothalamic thermoregulatory processes due to low estrogen levels. Norepinephrine and serotonin are the two most recognized neurotransmitters involved in hypothalamic thermoregulation. In a study that included breast cancer survivors and menopausal women, each group was assigned to receive one of three different dose levels of venlafaxine (37.5 mg, 75 mg, and 150 mg daily) or placebo for 4 weeks. A dose-related diminution in average hot flash scores from baseline was noted (27 percent in the placebo group versus 37 percent, 61 percent, and 61 percent for the three venlafaxine groups, respectively).

Consideration should be given to potential drug interactions, including the potential for an antidepressant to affect hepatic metabolism of cancer drugs via cytochrome p450 enzyme induction or inhibition in all cancer patients. In breast cancer patients, tamoxifen has been widely studied in this context, because of early reports that indicated an association between antidepressants with strong 2D6 inhibition (such as fluoxetine and paroxetine) and reduced efficacy of tamoxifen in maintaining remission in breast cancer. While several studies have confirmed that antidepressant drugs that have strong 2D6 inhibition can reduce the conversion of tamoxifen to its active metabolite, more recent studies indicate that this effect does not appear to have an impact on breast cancer recurrence rates or survival. However, until further research resolves the debate, it remains prudent to avoid potent CYP2D6 inhibitors when possible in preference for antidepressant options that have weak 2D6 inhibition (such as venlafaxine or citalopram).

Gynecological Cancers

The group of gynecological cancers includes those of the ovary, fallopian tube, uterine corpus, cervix, vagina, and vulva. Cancer of the endometrium is the most common gynecological cancer, with 54,870 new cases in the United States in 2015, as per estimates from the ACS. The ACS estimates about 21,290 women will receive a new diagnosis of ovarian cancer and 14,180 women will die from ovarian cancer in 2015, the highest mortality of the gynecological cancers. Over the last 30 years, the cervical cancer death rate has gone down by more than 50 percent due to increased use of the Pap test. Furthermore, the HPV vaccine will continue to decrease the number of women affected by the cancer-causing pathogen. Vaginal and vulvar cancers are less common. They are often detected earlier and thus more responsive to treatment. Women may sometimes delay seeking evaluation because the symptoms are mild or are attributed to other problems. For instance, abdominal bloating may be mistaken for premenstrual syndrome.

Treatment for these cancers includes surgery, chemotherapy, and/or radiation. The cancer and its treatment have significant impact on

hormones, sexuality, fertility, and bladder and bowel function, causing psychological distress.

Hematological Malignancies

Hematological malignancies are associated with significant psychosocial distress in many patients, particularly those undergoing stem cell transplantation. Leukemia, lymphoma, and multiple myeloma (MM) account for approximately 9 percent of all cancers, with an estimated annual combined incidence in the United States of about 150,000. Blood dyscrasias can occur during childhood or early adulthood, thus impacting patients during a normally productive and active time in their lives. Complicated treatment regimens with intensive multicomponent protocols have greatly increased cure rates and have prolonged survival.

Distress is common before, during, and after treatment for hematological malignancies and occurs along a continuum of severity that is elevated compared to the general population. Adjustment, mood, and anxiety disorders are particularly common. Individuals undergoing HSCT encounter major disruptions to social, family, and work life. They are separated from supports during transplant or chemotherapy induction and hospitalization. Challenges prior to allogeneic transplant may emerge as a result of stem cell donor selection, family stressors, and financial, housing, and other practical concerns. Patients are more vulnerable to distress during periods when prognosis or treatment course change (e.g., graft failure), severe adverse effects develop (e.g., GVHD), rehospitalization occurs, or changes occur to clinician follow-up (e.g., discharge from transplant center). Periods of immunological compromise limit social contact, even after discharge. Illness-associated stigma and job loss can occur. Patients who have psychiatric issues such as anxiety or depression are at risk for poorer health outcomes, longer hospital lengths of stay, and higher mortality. Insomnia, delirium, fatigue, depression, and anxiety are common psychiatric symptoms and syndromes encountered during treatment course. Survivors commonly experience PTSD, depression, anxiety, and cognitive impairment long after treatment completion. Psychosocial support is an essential component of treatment of patients with hematological malignancies at all stages of illness.

Paraneoplastic Syndromes

The term *paraneoplastic* refers to a group of disorders defined as indirect, remote, nonmetastatic effects of cancer. In 50 percent of the cases, symptoms precede the diagnosis of cancer and run an independent course. The syndromes may be due to (1) production of substances by the tumor that directly or indirectly cause distant symptoms, (2) depletion of normal substances that lead to a paraneoplastic manifestation, or (3) host response

to the tumor that results in the syndrome.

Paraneoplastic Neurological Disorders. Any part of the CNS and peripheral nervous system can be affected. Most paraneoplastic neurological disorders begin acutely or subacutely, progress over weeks to months, and then stabilize. Many of these disorders are associated with serum and cerebrospinal fluid antibodies formed against antigens that are also expressed within the tumor. Dementia, particularly memory loss (amnesia), is a common feature in several variants, which are grouped under the name *carcinomatous encephalomyelitis*. Most of these patients have SCLC and anti-Hu antibody.

Limbic encephalitis is another paraneoplastic syndrome that usually presents with abrupt changes in mood, personality, and with severe impairment in recent memory. Patients may appear agitated or confused and may experience hallucinations as part of complex partial seizures. The neoplasms associated with limbic encephalitis include testicular cancer, SCLC, thymoma, breast cancer, and Hodgkin lymphoma. The antibodies associated with limbic encephalitis are anti-Hu, anti-Ma2, CRMP5, and amphiphysin.

Opsoclonus-myoclonus ataxia syndrome consists of uncontrolled arrhythmic eye movements. It is also associated with diffuse or focal myoclonus, truncal titubation with or without ataxia, or other cerebellar signs. Ataxia, irritability, insomnia, and dementia often accompany the opsoclonus. The incidence of this disorder is less than 2 percent. In adults, it occurs in the presence of breast, gynecological, or other types of cancers. In children, it can be due to neuroblastomas or viral infections. It is associated with anti-Ri antibody.

Paraneoplastic syndromes that affect the motor neurons can also present with psychiatric symptoms in patients with degeneration of cortical neurons, resulting in frontotemporal dementias. *Neuromyotonia* consists of several disorders that affect the neuromuscular junction. These include neuromyotonia, the cramp fasciculation syndrome, and Morvan syndrome. The disorder is autoimmune and affects voltage-gated potassium channels. Morvan syndrome consists of memory loss, insomnia, changes in mood, and hallucinations. These disorders are associated with thymoma, SCLC, and Hodgkin lymphoma.

The treatment for paraneoplastic syndromes consists of treating the underlying tumor and suppression of the immune response using corticosteroids, plasma exchange, and intravenous immunoglobulin.

Metabolic and Endocrine Paraneoplastic Syn dr omes. Production of cytokines, hormones, or hormone precursors by the tumor may result in development of endocrine syndromes in cancer patients. Tumors of endocrine glands may be physiologically active and produce anxiety, mood

disorders, confusional states, and psychosis.

Hypercortisolism (Cushing syndrome) can be caused a pituitary adenoma, hypothalamic tumors, ectopic adrenocorticotropic hormone (ACTH) syndrome, and exogenous corticosteroids. Hypercortisolism presents with mood instability, depression, insomnia, fatigue, cognitive deficits, psychosis, and delirium.

A syndrome of inappropriate antidiuretic hormone secretion can be associated with SCLC, cerebral metastases, medications, and infections. This syndrome results in hyponatremia; the severity of the symptoms depends on the sodium level and the rate of drop in sodium levels. Patients can present with symptoms of nausea, vomiting, lethargy, stupor, delirium, cognitive impairment, seizures, agitation, focal neurological deficits, and coma.

Carcinoid tumors cause anxiety, depression, and delirium. Carcinoma of the pancreas has been reported to be associated with an elevated incidence of depression, sometimes as the presenting symptom. Metabolic disturbances including hypoglycemia, hyperglycemia, hypophosphatemia, and hypomagnesemia due to different types of cancers and can cause delirium, mood disturbances, and anxiety.

PSYCHIATRIC PROBLEMS DUE TO CANCER TREATMENT

Radiation and most of the drugs used in cancer treatment have significant side effects and are even carcinogenic in tumoricidal doses. Patients may become transiently more ill from the treatment than from the disease, developing side effects such as hair loss, which have great emotional impact. Steady compliance requires a level of courage and endurance of which the medical staff is not fully aware. Patients may become upset about the side effects, convinced that they represent disease progression, depressed about the paradoxical nature of treatment, and exhausted to the point at which they develop clinical depression and anxiety. If untreated, they delay or stop cancer treatment, moving to alternative treatments and jeopardizing their survival. Later adaptation is also impaired. Many patients, however, come to feel dependent on continued treatment as a means of holding the cancer at bay. There is a rise in anxiety level at the end of treatment associated with this fear and with the difficulty of making a transition to the survivor status. This can be assuaged by acknowledging it, exploring its cause, and making concrete plans for continued medical supervision and psychological support.

Radiation Therapy

About half of all the patients with cancer are treated with radiation therapy, either alone or in combination with other types of cancer treatment. Radiation therapy may be external or internal and may be curative or

palliative. External radiation comes from a machine outside the body and is usually given on an outpatient basis. Internal radiation is implanted into or near the tumor in small capsules or containers. It may require a hospital stay. If the patient will have external radiation, the radiation oncologist uses a process called simulation to define where to aim the radiation. During simulation, the patient lies still on an examining table while the radiation therapist uses a special x-ray machine to define the treatment port or field. Simulation may also involve imaging studies (CT scans) to help the radiation therapist plan on how to direct the radiation. Patients with anxiety disorders, claustrophobia, or pain have difficulties with simulation and radiation treatment. The early effects of radiation may be seen a few days or weeks after treatments have started and may continue for several weeks after treatments are completed. Late effects may not be seen until months, or even years, later. Side effects depend on the site, dose, and size of the radiation fields and the duration of treatment. The skin may be red, tender, and sensitive during the first 2 weeks. Dryness and peeling may occur in 3 to 4 weeks. Patients may suffer from fatigue, sore throat, and anorexia. Patients who receive total body irradiation are at greatest risk for having nausea. Without effective prophylaxis or treatment, more than 90 percent will experience nausea and vomiting acutely. Affected patients may suffer from various adverse complications, including dehydration, electrolyte disturbances, malnutrition, and aspiration pneumonia. Treatment of nausea includes adequate nutritional support and hydration. Medications to treat nausea include treatment with 5-HT_3 receptor antagonists (5-HTRAs) with or without corticosteroids, dopamine antagonists, antihistamines, or anticholinergics. Sexual impairment is common after radiation to the pelvic area, in the form of impotence in men and dyspareunia in women. The latter, related to vaginal fibrosis, is improved by vaginal dilatation. Acute, transient GI problems, such as diarrhea and food intolerance, are common; severe, irreversible GI problems are rare. Fatigue is reported in 80 percent of patients and is sometimes the most debilitating and demoralizing. It overlaps with symptoms of depression but exists independently of it. The normal response to brain radiation has been divided into three phases based on the time of initiation of symptoms. Acute effects occur during the first weeks of treatment. Symptoms include headache, drowsiness, nausea, vomiting, and worsening of pre-existing focal symptoms. Treatment with dexamethasone usually relieves symptoms. Early delayed or subacute encephalopathy appears at 1 to 6 months after completion of radiation therapy with headache, somnolence, fatigue, and deterioration of previous deficits. Subacute encephalopathy is caused by demyelination. It responds to steroids and is usually self-limited. The third is an irreversible late, delayed (months to years) form that is localized to the white matter and is probably secondary to vascular injury, demyelination, and necrosis. Late effects are

also known as focal radiation necrosis, diffuse white matter injury, and combined therapy leukoencephalopathy. Symptoms of focal radiation necrosis include seizures, signs of increased intracranial pressure, and symptoms of localized lesions. Symptoms of diffuse white matter changes vary from lassitude to significant memory loss to severe dementia. Symptoms of combined therapy leukoencephalopathy include ataxia, confusion, dysarthria, seizures, dementia, or death. In long-term survivors (mainly from Hodgkin lymphoma and childhood leukemias) who have undergone whole brain radiation, a short-term memory deficit has been observed, as well as a decrease in intelligence quotient (IQ) of 10 to 25 points. Even in the presence of a normal IQ, neuropsychological problems and learning disabilities have been documented. The elderly and the young are the most affected.

Chemotherapy

Nausea and Vomiting. The immediate effects of chemotherapy include nausea and vomiting, which occur within 2 to 6 hours of treatment and may last up to 24 hours. The most emetogenic drugs cause symptoms in virtually all patients; the less emetogenic ones are variable. Some result in delayed vomiting, usually 8 to 24 hours after treatment. Chemotherapeutic agents and radiation cause vomiting by different mechanisms. Chemotherapy stimulates the chemoreceptor trigger zone directly, which is functionally outside the blood–brain barrier. This area when exposed to toxins in the bloodstream and cerebrospinal fluid can stimulate vomiting. This effect is mediated by 5-HT_3 and neurokinin type 1 receptors. Chemotherapy damages the GI mucosa and causes release of neurotransmitters including 5-HT_3, stimulating peripheral pathways via splanchnic and vagal nerves. It has also been reported that there is a neurohormonal component mediated by arginine vasopressin and prostaglandins.

Classical conditioning has been shown to be an important mechanism in the causation of several reactions to chemotherapy: ANV, which had been documented in 25 to 75 percent of patients before the use of better antiemetics; anticipatory immune suppression (AIS), the clinical significance of which has not been established; and learned food aversions (LFAs), which can be minimized by avoidance of staples and favorites. Patients develop a conditioned response to sights and smell reminiscent of the chemotherapy experience. After termination of treatment, these symptoms decrease; however, one study found that 6 to 14 months after treatment, 63 percent of patients still reported nausea, 5 percent reported vomiting, and 80 percent still reported anxiety with specific stimuli.

Because all of these reactions are presumed to represent conditioned responses, treatment approaches have been formulated accordingly.

Prophylactic antiemetic and antianxiety treatments are used to prevent or decrease the initial episode or conditioned response and have greatly minimized what used to be a dreaded side effect. With the advent of highly specific, centrally acting antiemetic agents, such as the 5-HT_3 receptor blockers, ondansetron (Zofran) and granisetron (Kytril), the success of antiemetic regimens has greatly increased and new ones with different mechanisms continue to appear. The evidence support that these agents are effective in chemo- and radiation-induced nausea and postoperative nausea. Other medications that are indicated for treatment of nausea are dopamine 2 (D_2) receptor antagonists that block the chemoreceptor trigger zone. The literature has shown that for the most common etiologies of nausea and vomiting at the end of life, 5-HT_3 antagonists are no more effective than the less expensive D_2 antagonists. Tardive dyskinesia is rare due to short-term use of these medications. Akathisia occurs in 10 to 20 percent of patients on metoclopramide, but the coincident use of lorazepam lowers the incidence to less than 3 percent. Mild sedation is acceptable or even welcome. Hypotension may limit the use of prochlorperazine (Compazine) and chlorpromazine. Since some of the nausea is also anxiety-related, benzodiazepines are effective in preventing anticipatory nausea. In an open-label trial, mirtazapine, a potent antagonist of 5-HT_2 and 5-HT_3 receptors, was found to decrease nausea. The commercially available cannabinoids, nabilone (Cesamet), and dronabinol (Marinol) target ubiquitous cannabinoid (CB) receptors in the central (CB1) and peripheral (CB1 and CB2) nervous systems. Dronabinol is no more effective than other drugs and has a similar side-effect profile, plus a significant incidence of dysphoric reactions in older patients. Corticosteroids, such as dexamethasone, are generally included in most antiemetic regimens for their potential effect on tumor-associated inflammation.

Behavioral methods are important, effective ways to lessen the reactions to chemotherapy. They are most effective when instituted before the conditioned response is established. They include hypnosis and progressive muscular relaxation with guided imagery or relaxation with cognitive distraction to block the conditioned stimulus and the conditioned response; systemic desensitization, which extinguishes the conditioned response; and cognitive distraction, which blocks perception of the conditioned stimulus. Patients are also helped by reassurance about the frequency and normalcy of their responses.

CNS Toxicity. Chemotherapeutic agents may cause neuropsychiatric problems in one of the following ways: Indirectly by causing metabolic abnormalities or other organ failure and directly by causing CNS toxicity. The CNS, comprised of the brain and spinal cord, is responsible for the neurological functions of mental status, level of consciousness, sensory and

motor functions, and cerebellar and cranial nerve functions. The blood–brain barrier determines whether a cytotoxic agent can reach the nervous system. Chemotherapeutic agents that can penetrate the blood–brain barrier to the CNS cause alterations in the cranial nerves, as well as cognitive, cerebellar, and motor and sensory functions. Patients themselves are keenly aware of the cognitive impairment and talk to each other about the so-called *chemo-brain* in the waiting room and on the Internet.

Initially, mood changes, personality changes, and subtle cognitive changes may predominate, leading to a psychiatric consultation. Only later does the encephalopathic nature of the condition becomes manifest. Since the mid-1970s, cross-sectional and longitudinal studies have found an association between chemotherapy and cognitive changes. Cognitive functioning is objectively assessed by performing a battery of neuropsychological tests that allow comparison to a control group or normative data. The research data have shown that there is a discrepancy between objective measures of cognitive impairment and the patients' self-report of their cognitive functioning. Patients' perception of the dysfunction appears to be more deteriorated than what is detected by the objective evaluation and is associated with anxiety or depression. This discrepancy between patients' self-report and objective measures of cognitive function might be due to the inability of the neuropsychological tests to evaluate cognitive functioning in real-life situations and the fact they are performed under optimal conditions. Another explanation for the discrepancy is that patients tend to attribute their premorbid cognitive impairment to cancer or to depression or anxiety. It is imperative to evaluate premorbid cognitive functioning that serves as a parameter for comparison with cognitive impairment posttreatment. Although acute cognitive changes during chemotherapy are common, studies in survivors have shown long-term posttreatment cognitive changes in 17 to 34 percent of patients. The cognitive changes occur at different areas of cognition, including working memory, attention/concentration, executive function, and information processing speed. The cognitive functioning is reduced but often remains in the normal range.

Studies that have used imaging techniques have shown a reduction in the volume of brain structures. Changes in the integrity of white matter tracks that connect brain structures have been implicated in decreased cognitive functioning.

Possible mechanisms by which chemotherapy can cause cognitive deterioration include direct neurotoxic effects, oxidative damage, immune dysregulation with the release of cytokines, decreased delivery of oxygen secondary to anemia and blood clotting, and hormonal changes. Genetic predisposition with variability of genes that regulate neural repair or plasticity, such apolipoprotein E (APO E) and BDNF, may increase patients' vulnerability to chemotherapy-induced cognitive side effects.

Certain drugs, such as the vinca alkaloids, especially vincristine, cause neurotoxicity at all doses; others, such as cisplatin, cause neurotoxicity only at higher doses. In some drugs, toxicity is mild and transient; in others, it is delayed, irreversible, or both. Toxicity is also usually, but not always, dose-related and cumulative. Certain drugs, such as methotrexate, when given intrathecally, may cause several CNS side effects, such as reversible encephalopathy, delayed irreversible encephalopathy, dementia, or developmental delay when coadministered with radiation therapy. Several studies have found an association between cognitive impairment and the number of cycles and higher doses of chemotherapy. Chemotherapy regimen that consists of methotrexate, cyclophosphamide (Cytoxan), and 5-fluorouracil (Adrucil) appears to be highly associated with cognitive impairment.

Several treatment interventions have been initiated and evaluated with methylphenidate, erythropoietin, ginkgo biloba, and cholinesterase inhibitors, but the data are inconclusive and further research is needed. Large longitudinal studies are needed to evaluate the duration of cognitive impairment, which remains uncertain (Table 27.8–9).

Table 27.8–10 lists agents that predominantly cause mood or psychotic symptoms. All of these chemotherapy agents and many more can also cause delirium. Delirium due to chemotherapy agents is often reversible and remits spontaneously if the drug is stopped. This may take only a few days or up to several weeks. If the offending drug cannot be discontinued, symptomatic treatment should be instituted.

Disorders of special senses include ototoxicity, associated with cisplatin and several commonly used antibiotics; visual disturbances, associated with mitotane (Lysodren) and tamoxifen; gustatory and olfactory hallucinations, described with vinblastine. These disorders may occur slowly or suddenly and reversibly or permanently and often result in acute psychological distress requiring psychiatric consultation. Long-term effects of ototoxicity add considerably to the educational and neuropsychiatric impairment of these patients. The diagnosis is often missed in younger children, in whom early intervention is essential to minimize later developmental impairment.

Ifosfamide is a highly active chemotherapy agent, but its neuropsychiatric toxicity is a challenging problem, occurring in 5 to 30 percent of patients with cancer. Clinical symptoms can range from mild confusion to stupor, coma, and death. Somnolence, confusion, disorientation, fatigue, hallucinations, agitation, blurred vision, aphasia, seizures, asterixis, extrapyramidal symptoms, mutism, stupor, coma, and death have all been described. Delirium is the most commonly reported neuropsychiatric toxicity. Psychotic disorders with delusions or hallucinations in the presence of a clear sensorium and mood disorders with psychotic features induced by ifosfamide have also been described.

Chloroacetaldehyde (CAA) is a cytotoxic ifosfamide metabolite. The principal effect of CAA is the depletion of intracellular glutathione, which results in neurotoxicity due to impaired detoxification of neurotoxic substances. CAA is structurally similar to metabolites of ethanol (acetaldehyde) and chloral hydrate (trichloroacetaldehyde) with CNS depressant effects. Discontinuation of ifosfamide, and intravenous hydration, are part of the initial supportive care. Methylene blue, an agent FDA approved for the treatment of methemoglobinemia and lower genitourinary tract discomfort, was first reported for treatment and prophylaxis of ifosfamide neuropsychiatric toxicity in 1994. Intravenous methylene blue 50 mg every 4 hours is recommended until the symptoms of ifosfamide neuropsychiatric toxicity resolve. Methylene blue is hypothesized to reverse ifosfamide neuropsychiatric toxicity through several different mechanisms, namely restoring mitochondrial respiratory chain function by serving as an alternative electron acceptor, restoring hepatic gluconeogenesis by reversing nicotinamide adenine dinucleotide plus hydrogen (NADH) inhibition, preventing the formation of CAA. Psychotropic medications were not found to be helpful in controlling psychiatric symptoms induced by ifosfamide in the authors' experience, as also reported by other researchers.

Table 27.8–9.
Neurological Complications of Selected Chemotherapeutic Drugs

Acute encephalopathy (delirium)
Corticosteroids
Methotrexate (Folex, Mexate) (high dose IV, IT)
Cis-platinum
Vincristine (Oncovin)
Asparaginase (Elspar)
Procarbazine (Matulane)
5-Fluorouracil (Adrucil) (± levamisole [Ergamisol])
Ara-C (Cytosar-U)
Nitrosoureas (high dose or arterial)
Cyclosporine (Neoral)
Interleukin-2 (Proleukin)
Ifosfamide (IFEX)/mesna (MESNEX)
Interferons
Tamoxifen (Nolvadex)
VP-16 (Toposar) (high dose)
N-phosphonacetyl-L-aspartate (PALA)
Chronic encephalopathy (dementia)
Methotrexate
BCNU (BiCNU)
Ara-C
1-Hexyl-5-fluorouracil (Carmofur; Mifurol)

Fludarabine (Fludara)

Visual loss

Tamoxifen
Calcium nitrate
Nitrosoureas (intra-arterial)
Cis-platinum

Cerebellar dysfunction/ataxia

5-Fluorouracil (± levamisole)
Ara-C
Phenytoin (Dilantin)
Procarbazine
Hexamethylmelamine (Hexalen)
Vincristine
Cyclosporine

Aseptic meningitis

Trimethoprim and sulfamethoxazole (Co-trimoxazole) (Bactrim)
IV immunoglobulin (Iveegam)
NSAIDs
Levamisole
Monoclonal antibodies
Metrizamide (Amipaque)
OKT3 (anti C-3 monoclonal antibody)
Ara-C
Carbamazepine (Tegretol)
Methotrexate

Headaches without meningitis

Retinoic acid (Avita)
Trimethoprim and sulfamethoxazole
Cimetidine (Tagamet)
Corticosteroids
Tamoxifen

Seizures

Methotrexate
VP-16 (high dose)
Cis-platinum
Vincristine
Asparaginase
Nitrogen mustard (Mustargen)
BCNU
Dacarbazine (DTIC-Dome) (intra-arterial or high dose)
PALA
m-AMSA (AMSA P-D; Amsacrine)
Busulfan (Busulfex) (high dose)
Cyclosporine
Metronidazole (Flagyl)
Misonidazole
β-Lactam antibiotics
Iodinated contrast material (IV or IT)

Myelopathy (IT drugs)

Methotrexate
Ara-C
Thiotepa (Thioplex)

Peripheral neuropathy

Vinca alkaloids
Cis-platinum
Hexamethylmelamine
Procarbazine
5-Azacytidine (5-Aza-C)
VP-16
VM-26 (Vumon)
Misonidazole
Methyl-C
Ara-C
Paclitaxel (Taxol)
Suramin (Bayer 205)
Mitotane (Lysodren)

BCNU, bischloroethylnitrosourea; IT, intrathecal; IV, intravenous; NSAIDs, nonsteroidal anti-inflammatory drugs.

Table 27.8–10.
Chemotherapy Agents Associated with Mood and Psychotic Symptoms

Dacarbazine (DTIC-Dome): depression and suicide reported, especially when used with hexamethylmelamine (Hexalen).
Vinblastine (Velban): frequent reversible depression
Vincristine (Oncovin): 5% incidence of hallucinations; depression noted.
L-Asparaginase (Elspar): reversible depression noted.
Procarbazine (Matulane): monoamine oxidase inhibitor; concurrent tricyclics and selective serotonin reuptake inhibitors are contraindicated; associated with mania and depression; potentiates alcohol, barbiturates, and phenothiazines.
Hydroxyurea (Hydrea): hallucinations reported.
Interferon: anxiety, depression with suicidal ideation common at high doses.
Steroids: frequent alterations of mental state, ranging from emotional lability through mania or severe, suicidal depression to frank psychosis
Tamoxifen (Nolvadex): rare depression or delusional syndrome, or both
Leuprolide (Lupron): depression

Antineoplastic Hormones and Biological Modifiers

The hormonal agents and biological modifier agents are used in an increasing number of cancers. They are associated with depressive symptoms but also with many adjustment disorders related to their sexual side effects, such as feminization, masculinization, and disorders of sexual functioning. Besides the reversible effects already described, interferon can cause an acute reversible encephalopathy, with confusion, lethargy, and

hallucinations. One case of delayed progressive encephalopathy has been reported. IL-2 (Proleukin) causes headache, agitation, disorientation, and confusion, which are dose related and reversible. A rare severe encephalopathy has also been reported. These drugs have become frightening to patients who hear of the multiple side effects. However, symptomatic treatment is effective for all but the rarest severe reactions. A randomized prospective study of melanoma patients given prophylactic paroxetine has raised some controversy, with some researchers decrying the indiscriminate use of an SSRI in a group of patients who may not all need it. A previous psychiatric history was a strong risk factor, as was the presence of baseline symptoms. It is easy for psychiatric consultants to help oncologists put in place a screening program asking patients about their psychiatric history and their baseline mood symptoms. Simple questionnaires, such as the Hospital Anxiety and Depression Scale (HADS), which has been developed for medically ill patients, have identified a 10 to 15 percent sample that would qualify for prophylactic treatment. Ongoing monitoring for depressive symptom would allow early treatment and would spare patients from unnecessary suffering. Once a medical team has been educated and a routine has been set in place, ongoing psychiatric consultation is not necessary for any but the new-onset or complicated cases.

Tailor-made vaccines against various parts of the malignant cell are being developed in increasingly sophisticated ways, in particular, for melanoma. They remain difficult to make, the responses are usually measured in months, and they are limited to research-based programs, but they are of great scientific interest and have much unexplored potential.

Surgery

Surgery continues to play a central role in the treatment of cancer. Several psychological and psychiatric issues arise around the time of surgery, such as an inability to give consent, preoperative anxiety, exacerbation of pre-existing psychiatric disorders, postoperative delirium, drug–drug interactions, drug withdrawal concerns especially for patients on benzodiazepines, poor compliance with postsurgical rehabilitation recommendations, and many others. In addition to perioperative reactions common to different types of surgical procedures, psychological reactions specific to the surgery site and the related functional loss should be taken into consideration. Many studies on the psychological effects of mastectomy, hysterectomy, colostomy, and prostatectomy report the high cost paid in quality of life and recommend that more emphasis be placed on this dimension. Laryngectomy patients and those who undergo other extensive head and neck procedures experience considerable psychiatric morbidity; the loss of voice is catastrophic for many, and the blow to self-image is severe. Patients who have had curative limb amputations can have

a good psychological outcome. There is little systematic information about patients who undergo unusual and extreme procedures, such as pelvic exenteration.

Stem Cell and Bone Marrow Transplants

Stem cell transplantation is being used for an increasing number of childhood and adult cancers. It is commonly recommended for leukemias, lymphomas, multiple myeloma, and myelodysplastic syndromes. It is a potentially curative treatment for otherwise lethal disorders and succeeds best when given to patients in remission, with as little previous treatment as possible. Yet, it has an uncomfortable course, and a significant incidence of graft-versus-host disease. Undertaking this type of treatment, especially when the patient is doing relatively well, can be stressful for patients, families, and caregivers. However, successful transplants are indeed lifesaving. Comprehensive psychiatric assessment has become a part of the routine pretransplant evaluation in major cancer centers to identify patients at high risk. Patients sometimes need psychological intervention before the procedure and often need it if their posttransplant course is complicated. Delirium is common during acute periods. Anxiety and depression symptoms are common due to the rigors of the experience. Nevertheless, there are enough patients who leave the hospital after a few weeks, with variable degrees of emotional and physical depletion, but essentially cured of an incurable disease.

New Supportive Technologies

The use of life supports and nutritional supports in the critically ill patient has created a spate of psychological, legal, and ethical issues that are widely discussed in the lay press. Less well known are developments in the nutritional support of cancer patients who can no longer take in food because of physical obstruction, as in head and neck cancers, or because of physiological failure, as in radiation enteritis. In the former, percutaneous endoscopic gastrostomy (PEG) or percutaneous endoscopic jejunostomy (PEJ) and, in the latter, home parenteral nutrition (HPN) allow a significant improvement in quality of life. These home-based technologies create major changes in family-structuring routines, such as the preparing and sharing of meals, and are time and labor intensive, causing adaptational strains that may bring them to psychiatric attention.

PSYCHOLOGICAL CONCERNS IN PATIENTS WITH CHRONIC CANCER

Cancer has become curable in a growing number of cases, especially with improved screening techniques and early diagnosis. However, in many cases it becomes a chronic disease, at which point treatment goals focus on management. While these efforts may be effective for many years, the

patient's experience is nonetheless challenging and can involve navigating secondary prevention guidelines, episodes of acute disease, and sometimes coping with indolent tumors.

Patients vary in how they deal with these difficult life situations, and doctors vary in the style and content of their communications. Communication skills training has been identified as a neglected part of medical education and is being addressed by researchers and educators. All physicians communicate not only through their words, but also with voluntary and involuntary nonverbal behaviors, which are similarly subject to interpretation, scrutiny, and sometimes distortion. In the case of cancer care, oncologists face the challenge of balancing the details of diagnosis, prognosis, and realistic expectations with installation of hope. When treating patients with cancer, the psychiatrist must have an understanding of the medical reality, including burden of illness and staging, treatment options and plan, and long-term ramifications of both the disease and treatment. Patients with cancer frequently ask their psychiatrists to clarify details of their treatment. This may happen for a number of reasons, ranging from having a better therapeutic relationship with their psychiatrist than their oncologist, to limitations during oncology appointments such as anxiety, surprise/shock, or time.

Cancer Pain

The overall incidence of pain in cancer patients is approximately 50 percent but increases in frequency and severity with advancing disease. There are many pain syndromes that have been defined in cancer patients. Their variety illustrates the importance of evaluating and diagnosing the sources of pain in cancer patients, because the treatments for each of them are so different and, in many cases, specific. The psychiatrist has a special role to play throughout the course of the illness, especially as pain specialists can be difficult to access, and physicians still display reluctance to treat the pain of cancer patients aggressively enough. The primary goal of cancer pain management is effective symptom control to allow patients to participate in the necessary diagnostic and therapeutic approaches to treat their cancer and to achieve an acceptable quality of life at all stages of the disease.

The debilitating and demoralizing effects of pain give rise to increased incidence of psychiatric symptoms and disorders at any stage of the disease. Pharmacological treatment is essential. Patients and families may be comforted to know that pain treatment with opioids very rarely results in addiction, despite the emergence of physiological dependence evident in tolerance and withdrawal. Addiction, or psychological dependence, is a pattern of behavioral and cognitive symptoms that indicate continued use of substance despite significant substance-related problems. These symptoms include impaired control, social impairment and failure to fulfill

major role obligations, and risky use. Patients' and physicians' fears of addiction are a major barrier to adequate cancer pain management. Patients with poorly managed or undertreated pain may mimic the signs of addiction, such as requesting higher and more frequent doses or exaggerating pain complaints; this behavioral presentation is sometimes referred to as pseudo-addiction.

Identification of the nature of the pain (somatic, visceral, and neuropathic) and the specific pain syndrome can facilitate the development of an approach for pain relief in patients with pain and cancer. It is well recognized that pharmacological approaches represent the mainstay of therapy, but any therapeutic strategy must also implement the use of psychological, behavioral, mechanical, anesthetic, and neurosurgical approaches as needed. The guiding principles of a therapeutic strategy for cancer pain must include (1) detailed assessment of the patient's pain; (2) individualization of the therapeutic approach; (3) the assurance of available expertise to provide therapeutic strategies to patients; (4) continual reassessment of the degree of pain relief and impact on mood, functional status, patient and family acceptance, and the patient's overall quality of life; (5) choosing the simplest approach prior to the use of complicated and extensive techniques; (6) ongoing communication between the physician and patient in defining the options for therapy and the potential risk and benefit ratios of any of the therapeutic approaches; and (7) defining the goals of pain management at or near the end of life.

Critical to the development of a successful therapeutic strategy is the physician's ability to establish a trusting relationship with the patient. This requires a complete history of the pain complaint, including the patient's description of the site of pain, its quality, exacerbating and relieving factors, its radiation if any, its exact onset and temporal pattern, and to what extent the pain interferes with activities of daily living, work, and social life, as well as the degree to which the pain has had a psychological impact. Multiple pain complaints, common in patients particularly with advanced disease, need to be prioritized and classified. In some instances, verifying the history from a family member may be necessary because the patient is either unable or unwilling to detail the pain symptom appropriately. Some patients may deny or minimize their pain; in these cases, family members may be more objective in assessing the disability of patients who underreport their symptoms. At the time of the initial assessment, the patient's psychological state should be evaluated, including the patient's psychiatric history, current level of anxiety or depression, suicidal ideation, and degree of functional incapacity. Numerous studies support the significance of psychological factors in accounting for differences in pain experiences in patients with cancer. In a study of the incidence of psychiatric disorders in patients with cancer, 39 percent of those patients who received a psychiatric diagnosis had significant pain. In

contrast, only 19 percent of patients who did not receive a psychiatric diagnosis had significant pain. In this group with pain, the psychiatric diagnoses included adjustment disorders with depressed or mixed mood and major depression. Uncontrolled pain is a major factor in cancer-related suicide. Understanding the meaning of pain to the cancer patient is a critical component of a successful assessment strategy. It is not uncommon for patients during active therapy to endure significant pain for the promise of a successful outcome. With advanced disease, when curative therapy is no longer an option, it is common for patients and families to request that, if nothing else can be done, at least the patients' pain should be adequately managed.

Chronic cancer patients may have a constant source of pain or one that worsens slowly. The patient may become habituated to a given level of opiate and need increasing levels of opioids at a given level of disease and as the disease progresses. Persistent patient complaints should be heeded, even if scans and MRI do not reveal a new etiology. Chronic pain management modalities are important. Nonpharmacological methods such as meditation, meditation-based stress reduction (MBSR), behavioral methods, group support, and supportive therapy focused on improving quality of life all have a role to play.

The principles of acute postoperative pain relief emphasize aggressive immediate treatment, often with the use of patient-controlled analgesia (PCA) pumps. They should be carefully used in confused patients, and a bowel regimen should always be instituted at the same time that any opiate is prescribed. Sufficient initial relief is associated with an easier postoperative course and, in some studies, a shorter hospital stay. Pre-emptive anesthesia is known to prevent phantom pain, and there is one report of its beneficial effect in prostate cancer surgical patients.

Psychiatrists should be familiar with drug–drug interactions related to use of opioids analgesics and adjuvant pain medications (i.e., antidepressants, anticonvulsants). Tramadol (Ultram) is a commonly prescribed, weak mu-opioid receptor agonist that also inhibits serotonin and norepinephrine reuptake. As such, patients taking tramadol concurrently with antidepressants should be monitored for serotonin syndrome. Opioid analgesics, especially those with long elimination half-lives, increase the risk of delirium in cancer patients. Rotation of opioids, continual assessment, and prescribing the lowest effective dose are recommended approaches to minimize adverse effects.

Physicians should also remain alert to the possibility that cancer patients may experience pain from other causes. Through longitudinal evaluation of quality and persistence of complaints, psychiatrists play an important role in a collaborate approach to pain management.

Sexuality

Many patients hide their sexual concerns, because they are embarrassed and perceive that their doctors are embarrassed. However, they are usually eager for information that is seldom provided. These complaints may vary from reports of sexual dysfunction and decreased libido to sudden onset of menopause symptoms.

Concerns for sexual function vary in younger and older cancer patients. Younger patients are more sexually active, and fertility remains an important issue. Research efforts have revealed conflicting results regarding the incidence of sexual dysfunction based strictly on site of primary cancer. A large sample of men with testicular cancer found no significant differences in dysfunction from population-based controls. However, a second study found that long-term male survivors of lymphoma had significantly more sexual dysfunction than matched controls. The sexual impairments associated with prostate, bladder, and rectal cancers match the amount of postsurgical sympathetic nerve damage rather than gross anatomical change. Both radical prostatectomy and radiotherapy for prostate cancer as well as adjuvant hormone treatment are often followed by severe long-term erectile dysfunction; sexual recovery is seldom achieved. Any course of radiation or chemotherapy can result in some sexual dysfunction, with radiation causing a high rate of dyspareunia.

Young female breast cancer patients often experience long-term lack of sexual interest. The attitude of their partners toward their body and femininity is very important for their sexual well-being. Women treated for gynecological cancers may develop lack of interest, vaginal dryness, and dyspareunia. Treatment-induced menopause is common to surgical, radiologic, and chemotherapeutic modalities of treatment. The symptoms are the classically described ones, but anecdotal evidence points to a higher severe symptoms when menopause is induced by treatment. Hot flashes are among the most troubling to treat because estrogens are contraindicated in many cancer patients, especially those with estrogen receptor–sensitive tumors.

The incidence of sexual dysfunction increases over time, and spontaneous remissions do not occur, even in the presence of stable or improved physical states. Both men and women treated for colorectal cancer can have persistent sexual dysfunction four years after diagnosis. Dysfunction is greater in younger patients across all diagnoses and is particularly high for cancers of breast, head, and neck, and other physically disfiguring cancers. In an ideal world, sexual concerns should be routinely addressed before treatment, with a baseline assessment that models open communication from the start. A detailed sexual history can be used to explore all possibilities and make subsequent discussions easier. Patients need detailed information about what problems to expect and about their reversibility. The information should be combined with reassurance that some form of sexual expression is always possible. If there is a spouse or

sexual partner, early joint involvement may be useful.

Early discussions are particularly necessary because sexual concerns play an important role in patient decision making. Both men and women may decide against the optimal treatment because of the fear of sexual impairment, which is not always justified. Then again, they may select very aggressive treatment in ambiguous situations without realizing the full impact it will have on their sexual functioning.

During the treatment, patients may experience direct pain and injury about which they may not grieve openly. Even in the absence of direct pain and injury, there is usually an interruption of sexual activity, which may be very threatening for patients who are uncertain about their partners' tolerance for the interruption. The extent of recovery needs to be clarified so that realistic coping can begin.

After the initial treatment period, patients should again be queried about sexual concerns. Patient workshops, support groups, and information sessions may be useful tools to complement individual treatment. A flexible format, with a mix of information, emotional ventilation, and focused problem solving has been well received in several types of cancer.

The period after the end of active treatment is a critical one for future adjustment. The patient usually cannot resume premorbid functioning even in the absence of any direct injury, but fatigue, pain, disappointment, and fear about the future are maximal. It is also a time of uncertainty and, hence, a window of opportunity. Dysfunctional patterns have not yet crystallized and, given the opportunity, couples can work together toward a new equilibrium.

Recent studies reinforce the observation that psychological and relational factors play a large (often the largest) role in determining the degree of posttreatment recovery of sexual functioning. Given the lack of attention paid to sexual issues that is characteristic of much of medical treatment, psychiatrists and other mental health professionals can occupy an important role in evaluation of sexual issues and facilitating treatment. They need to know the basic anatomy and physiological potential of a patient's condition and should refer the patient for further work-up and interventions if needed. Psychological treatment includes couple psychotherapy and sex therapy. Sexual rehabilitation involves an adaptation of sex therapy principles and can be helpful. It is based on the desire-arousal-orgasm model of sexual response and requires a phase-specific diagnosis. Late-phase dysfunction is quite effective if there is no anatomical interdiction. Early-phase dysfunction is more difficult to treat and benefits from a specialized referral.

Cancer Survivors

There are more than 14 million Americans living with a history of cancer,

according to the ACS. Almost two-thirds of cancer survivors were diagnosed 5 or more years ago, and nearly half of all cancer survivors are older than 70 years. The rising number of cancer survivors is a tribute to the human intellect and resiliency; however, it is important recognize that successful treatment does not come without times of distress and often long-lasting problems.

According to the most recent estimates from the ACS, the most frequent sites for invasive primary cancer among survivors were breast cancer among women (22 percent), prostate cancer among men (20 percent), and colorectal cancer (9 percent). More than half of cancer survivors had a history of one of these cancers. Among male cancer survivors the three leading types of cancer are prostate cancer (43 percent), colorectal cancer (9 percent), and melanoma (8 percent). Among female cancer survivors, the three leading types of cancer are breast cancer (41 percent), cancer of the uterus (excluding the cervix, 8 percent), and colorectal cancer (8 percent). Survivors of childhood cancer who are under age 20 make up a small fraction of all cancer survivors (1 percent). The survival rates are higher for white patients and lowest for African Americans, Native American/Alaska Native, and Hispanics. Cultural factors, including language, values, traditions, and trust in providers, can influence underlying risk factors, as can access to treatment, health behaviors, beliefs about illness, and approaches to medical care.

In early self-report studies, survivors were believed to be free of psychological residue, because many of them used denial and emphasized positive beliefs, such as having learned deeper values and being grateful for life. However, studies of leukemia/lymphoma survivors and breast cancer patients, using the Brief Psychiatric Symptom Inventory, showed a level of psychiatric symptomatology one standard deviation greater than the norm, or halfway between normal controls and psychiatric patients. Many subsequent studies confirm this association. Data from an American cohort of 9,535 long-term survivors of childhood cancer showed a 43.6 percent incidence of self-reported impairment in one or more of six domains: general health, mental health, functional impairment, activity limitations, pain resulting from cancer treatment, and anxiety or fears directly related to cancer and cancer treatment. Mental health concerns were the most frequently mentioned, and risk factors included female gender, low educational attainment, and low household income. A Dutch study found lower educational achievement, a lower rate of employment, a lower rate of marriage, and a higher incidence of men living with their family of origin as compared with a healthy control group. A history of brain tumor or cranial irradiation was the strongest independent predictor but did not account for all of the variation. A Danish study confirmed this finding; although long-term survivors had more psychological symptoms, they did not have an increased rate of psychiatric hospitalizations unless they had had brain

tumors or brain irradiation.

Delayed Consequences of Treatment

Four categories of delayed consequences of treatment have major psychological impact.

Infertility problems, premature ovarian failure, and interference with sexual response are particularly distressing if patients have not been educated and prepared to face them. However, there is evidence of persistent patient denial even in the face of preparation, such as sperm banking. Intervention should include endocrinological, gynecological, or urological evaluation and treatment. Psychotherapy, couples therapy, group therapy, veteran patient programs, and self-help programs can all be used to good effect.

Pituitary, thyroid, and other endocrine failures occur as a consequence of radiation therapy. Patients with these problems often present with psychiatric symptoms. Patients treated before or during puberty have short stature due to lack of growth hormone. Chronic fatigue is a common symptom that calls for patient and family education, because it is often given an unjustified psychological label.

Increased vulnerability to late organ failure and secondary malignancies and to the chronic anxiety that this vulnerability generates is another challenge for survivors. If such a complication comes to pass, it finds the survivor less able to cope and more prone to a posttraumatic response.

Social rejection, employment discrimination, and serious insurance problems may be the lot of survivors in a frightened, suspicious world. Survivors well enough to return to work plateau in their careers, being passed over for advancement by employers who fear a recurrence, or because they are averse to moving due to having become uninsurable. Despite the Americans with Disabilities Act (ADA), which bars employers from firing patients, there are cases of summary dismissal, and the appeal process is complex and uncertain. However, there are also cases in which sick employees are given extensive sick leave, warm support, and reassurance about their position. Also, the stigma of cancer is slowly diminishing as there are more examples of public figures undergoing treatment and continuing to lead productive lives.

Predictors of Poor Coping. Predictors of poor coping include previous psychiatric problems, lack of social supports, low educational and economic status, and stage of disease at diagnosis. In treating the cancer survivors psychotherapeutically, one should strive for a detailed understanding of the patient's experience. When the content is anxiety-provoking to the therapist, it is sometimes overlooked, leading to poorer outcomes even in the hands of seasoned therapists. Full knowledge is

required to understand where the trauma response is primary and where it is a shield for deeper psychodynamic issues. Lingering too sympathetically on the trauma may deprive the patient of the opportunity to grow beyond it. Yet, failing to be sufficiently empathic in understanding its impact is obviously unproductive. There are many survivor programs that offer workshops, lectures, and groups for patients, couples, and families. The National Coalition of Cancer Survivors (NCCS) is emblematic of the possibilities for self-awareness and psychological and political empowerment.

Patients at Genetic Risk for Cancer

The advent of deoxyribonucleic acid (DNA) technology and sophisticated epidemiological studies has permitted the identification of many healthy individuals at high genetic risk for developing cancer. Sources of genetic information include information from the patient's family medical history, physical examination, and medical records. There is a wide availability of information for the public on the Internet, which can cause concerns among patients and may lead them to consult with their doctors. Once patients are identified as being at risk of developing cancer, they should be referred for genetic counseling. After an individual is found to have an increased risk of having cancer, there are potential interventions with the goal of reducing the risk, such as having a colonoscopy to detect colon cancer. Information about familial cancer has psychological effects. High-risk individuals have chronic anxiety, lowered self-esteem, and problems surrounding marriage and childbearing. Self-estimates of risk are usually much higher than their real risk, even for those in the highest risk bracket. Previously unsuspecting carriers do show increased distress, whereas decliners of testing remain anxious throughout. High risk and high anxiety often go hand in hand with denial and poor health maintenance behaviors, such as avoiding breast self-examination and surveillance testing. A group therapy intervention with educational and expressive components has been effective in reducing psychological distress, normalizing risk estimation, and improving health behaviors. Individual educational and supportive interventions have shown promise. Some form of psychological intervention should be a regular part of the surveillance programs that have been developed in many centers. Similarly, psychological support is crucial to deciding whether to enter surveillance trials, to have prophylactic mastectomy, and, above all, to have testing. One study has shown that although only 71 percent of high-risk women would undergo testing for breast cancer, as many as 42 percent of average-risk women wished to undergo it as well. The consequences of testing go far beyond the psychological distress of the individual. Cooperation of unwilling family members may be required, and unwelcome knowledge may be forced on them. Marital, sibling, and parent–child relations may suffer, at which

point family-focused interventions come into play.

PSYCHOLOGICAL ASPECTS OF CANCER SCREENING

Screening for genetic risk is not the only area in which scientific and technological advances have created ambiguous situations. There are now many tumor markers, molecules associated with cancer, that can be used to (1) screen for the presence of cancer, (2) diagnose specific cancers, (3) determine prognosis in a given patient, and (4) monitor the course of the disease. No test meets all of these requirements, but there is a tendency to overuse them. PSA for men is the most reliable. It is specific to prostate, but an elevated level can signal prostatitis as well as cancer. It is extensively used both for screening and for monitoring prostate cancer and has resulted in many early diagnoses for a type of cancer that is found at autopsy in a large percentage of elderly men in whom it never grew to cause problems. The new patients are faced with difficult choices between watchful waiting, which some cannot tolerate, and treatment, which usually consists of either radiation or surgery preceded or followed by antiandrogenic hormonal therapy. All of these have significant and emotionally loaded side effects, such as impotence, incontinence, feminization, or depression. Cancer antigen 125 (CA125) is useful for following a known ovarian cancer but is not specific enough to be used for diagnosis, although it is often drawn for that purpose. Carcinoembryonic antigen (CEA) is even less specific for colon cancer.

Despite the potential for false positives, a recent study found that patients did not regret being tested. The author concluded that this enthusiasm would lead to overuse of other screening methods, especially powerful, expensive techniques such as CT scans, MRIs, and positron emission tomography (PET) scans, unless physicians have a higher threshold before ordering them without clear-cut criteria.

These patients may request counseling to cope with the many ramifications of test results and to help them decide on treatment. Asymptomatic patients who are told that they have cancer have an added burden when they undertake an arduous treatment. A part of them is never quite sure that there was really anything wrong with them. They continue to experience their illness as due to the treatment, and the more fragile patients with a tendency to denial, projection, mistrust, or paranoia may begin to doubt the reality of the cancer and the integrity of their physicians.

ETHICAL ISSUES

Psychiatry–Ethics Interface

In today's contentious social climate, ethical and legal dilemmas multiply with increasing stress, often resulting in psychiatric consultation. If inappropriate psychiatric labeling has occurred, psychiatrists must remove

the label, redefine the situation, bring out the underlying ethical issues, and set a decision-making process in motion. To be effective, they must be clear about the difference between psychiatric issues on the one hand and ethical and legal issues on the other hand. They must diagnose and act decisively in the former, be restrained and humble in the latter, and educate others about the distinction. If they understand the ethical issues, have a precise knowledge of the legal constraints operating in the case, and call for appropriate consultation, their interviewing skills and experience in conflict reduction can make them valuable participants. Selected ethical issues in the care of cancer patients are discussed in the following sections, with emphasis on their psychological correlates.

Truth Telling

At the extremes, cultures that put a high value on autonomy consider the need for truth telling as self-evident, whereas cultures in which the patient is tightly embedded in a wider family unit consider truth telling as an exercise in maleficence. The spreading breakdown in close kinship ties and the steady improvement of modern cancer treatments are combining to increase truth telling in growing segments of the world. However, local cultural deviance continues to arouse strong objections to it and may generate the need for outside consultation for cultural and religious guidance.

Informed Consent

In the United States, informed consent is legally required for conventional and experimental treatment. Given the toxicity and noncurative nature of much cancer treatment, this does not seem unreasonable and is spreading to other Western countries. However, it comes at some psychological cost. Anxiety reactions and occasional psychiatric decompensations are seen when patients feel overburdened by demands to make decisions. These often occur in predisposed individuals, even those with a good premorbid adjustment. Patients respond to active interventions, including support, medication, behavioral techniques, and active environmental manipulation of home and treatment setting. These interventions are usually sufficient to help patients make decisions on their own behalf and to undergo treatment, but posttreatment relapses may occur and are resistant to therapy.

Investigational Treatment

Investigational treatment is common in cancer. Most patients in academic centers will probably be considered for eligibility in an investigational protocol. Since the Nuremburg trials and the Helsinki Declaration, the moral requirement for informed consent from research subjects has been

clearly spelled out and has been buttressed by regulations and institutions in the United States. Many critics feel it is honored in letter more than in spirit, but competent patients are known to systematically suppress information given to them to create the illusion of an old-fashioned paternalistic relationship; hence, the situation remains ambiguous, despite the efforts of lawyers, bioethicists, and well-intentioned physicians to tame it.

Randomized Clinical Trials

Critics of clinical trials claim randomization is incompatible with the doctor's basic responsibility to the patient. Overall, patient reactions are determined less by the protocol than by their relationship to the physician. They have an irreducible need to believe that their physician thinks that the study is good for them. This is not completely false, because no treatment reaches human trials if it has not shown some promise, and a physician-researcher is morally obligated and scientifically well advised to remove a patient from a protocol that will harm him or her. However, career pressures and local power relationships may mold physician behaviors, so institutional ethics are just as important as physician ethics but are much more complex and difficult to change. A good institutional review board and transparency in the presence of large medical teams are additional safeguards.

Some patients find randomization intolerable and refuse to participate or later demand to break the code. Most patients find trials with a placebo versus treatment design emotionally more difficult to understand and tolerate than one with two treatment arms. Many patients seek experimental treatment with desperate intensity, peddling influence to enter limited protocols or pleading to use untested agents that have not shown enough efficacy to warrant clinical trials. A few patients do participate in clinical trials out of strong altruistic motives but, most of the time, these are secondary to the hope of personal benefit.

Alternatives to randomized clinical trials (RCTs) have been suggested, but they are unwieldy and less powerful than the trials. The results of large, multicenter RCTs are supposed to establish the standard of treatment, but there are inevitable intercenter differences even in the absence of any misconduct. In the case of the mastectomy-lumpectomy controversy of the 1970s and 1980s, RCTs from several countries gave consistent results that did indeed change standard medical practice. However, in the case of combined hormone replacement therapy (HRT), there have been contradictory trials, and only recently has there been a preponderance of results to establish that HRT not only lacks the cardiac and osteoblastic benefits that had been believed to offset their moderate but well-established increased breast cancer risk, but also, in fact, comes at a higher risk for all of them.

PSYCHOLOGICAL PROBLEMS OF FAMILIES

Research on the family of the cancer patient is an active and exciting field, especially due to the integration of family therapy concepts. Some findings are disease specific, but many of them can serve as a paradigm for understanding the impact of any major illness on the family.

Psychological Reciprocity

Awareness of the powerful reciprocity of emotional states and behaviors between cancer patient and family members is well established from numerous studies of parents, spouses, and siblings that document levels of distress and functional impairment that can rival or even exceed that of the patient. More disturbingly, primary caregivers have been noted to worsen over time in the face of stabilization or improvement in the patient. Finally, studies of families after bereavement or cancer survival show a significant incidence of impaired functioning over extended periods of time and, again, a worsening over time in the distressed families. These findings have some urgency, given that most families experience cancer at some time and that it will probably be a drawn-out, chronic ordeal. High-risk families can be identified early and can be significantly helped by professional intervention.

Family Evaluation

Families can be evaluated from several points of view. The most traditional one is to examine how well the family is taking care of the patient. Medical staffs are quite sensitive to this and often call for help when families fail in this area. Another approach is to scan for psychopathology in family members. If the family member is acutely symptomatic and visible to the treatment team, a consultation follows, but absent members go untreated.

Areas for which the family is responsible include:

1. *Providing emotional support and containment*. Family members are expected and expect themselves to contain their own reactions and to support the patient.
2. *Sharing responsibility for decision making*. Family members often carry the lion's share in researching and selecting options, with or without patient approval. This goes hand in hand with an enormous and universally reported *need for information*.
3. *Concrete care taking*. In today's budget-conscious climate and vanishingly short hospital stays, an ever-increasing amount of care is being thrown onto the family. This is not acknowledged; hence, families receive little help or training.
4. *Meeting financial costs*. Even under socialized medicine, families absorb massive expenses and suffer significant downward mobility.
5. *Absorbing social costs*. In one study, 25 to 30 percent of primary

caregivers lost their jobs, and more than half of the remaining primary caregivers moved to lesser paying jobs to accommodate their need for flexibility.

6. *Maintaining stability*. The patient's role and contribution must be replaced, a difficult task in the case of young mothers or breadwinners. The new demands must be folded into ongoing functioning. Some members develop increased emotional needs that must be met. All of the standard family functions, such as emotional nurturing, feeding, clothing, sheltering, educating, and socializing, must continue unabated.
7. *Adapting to change*. It has been said that change may be the only constant during a cancer illness. However, it is useful to divide events into acute and chronic phases:
   a. *Characteristics of the acute phases*: An acute event is met with effective mobilization of family members but, psychologically, they may be reacting quite asynchronously, with the denial of some and the pessimism of others causing significant friction. Members may protect themselves, under the guise of protecting each other, by colluding in a conspiracy of silence, which can rigidify and can be hard to change later. There is a universal need for information, to decrease uncertainty, to help the family problem-solve effectively, and to serve as a conduit for ensuring better communication, cohesion, and emotional containment.
   b. *Characteristics of the chronic phases*: Now, the family must return to developmental and maintenance tasks, which may expose conflicts between the patients' continuing needs and the needs of others. Initial mobilization is replaced with a paradoxical increase in psychological symptoms. Families isolate themselves, even when outside support is available, and may go into neutral gear, putting major decisions on hold.

Family-Centered Approaches

New and powerful methods of evaluation use family-centered approaches to examine how the family sustains itself and copes with the illness, while still meeting the needs of all its members. Four concepts are important to understanding the impact of the illness and to devising interventions that are not instantly neutralized for the sake of family homeostasis.

Developmental Stages. Families go through developmental stages with specific tasks at each stage. Cancer disrupts the tasks at any point, sometimes permanently, if no help is offered. Any intervention must be compatible with helping the family carry out its phase-specific tasks.

Family Structure. One must understand how a family is organized to

understand the caregiving dynamics and to intervene in ways that reinforce the family, rather than threaten and destabilize it.

Family History, Beliefs, and Myths. Every family is permeated by stories that dictate unconscious responses to illness, dependency, loss, and other fundamental life experiences. They underlie behaviors that seem idiosyncratic or bizarre to untrained observers and must be factored into treatment. Explaining them to the treatment team minimizes scapegoating and polarization.

Family Style of Interaction. In past studies, when family researchers were studying family functioning outside the medical setting, good communication among family members was found to be the best predictor of outcome across several theoretical models. Other factors that were studied include cohesiveness, expressiveness, conflict, hostility, distance versus enmeshment, and adaptability, to name the most common. Cohesiveness in earlier models was thought to have a bimodal distribution, with too much cohesiveness having a negative impact on the development of individual members. Yet, that was not found to be the case when high cohesiveness was culturally determined. The study of families undergoing severe illness or death of a member has found cohesiveness to be an even better predictor of good functioning than communication.

Communication in the medical setting has a very specific content and function. The truth-telling debates bear witness to the fact that medical information is not viewed as an unalloyed blessing. Perhaps cohesive families can put medical information to better use and become even closer as they process it, whereas it may further stress and disrupt non-cohesive families. Cohesiveness has also been found to be the common factor in families that show strain during crises but heal themselves over time.

Nontraditional Families. Divorced or blended families, single-parent families, homosexual families, and families that are ethnically, culturally, or religiously discrepant from the ambient medical culture experience many of the problems described previously but also have specific characteristics that are beginning to be systematically recognized.

Staff–Family Interface

During a life-threatening illness like cancer, the medical staff plays an enormous role in the family dynamics, invading the family system to create a new ad hoc hybrid system. That hybrid is the milieu in which consultation requests are generated. It must be understood and used for any intervention to be effective.

Guidelines for Intervention. Although most families show generosity,

courage, and resiliency, as many as 25 to 30 percent would clearly benefit from intervention. Of those, a much smaller number requires formal family therapy, which must be done with a realistic understanding of the medical facts and the medical milieu. Most families manage with short-term crisis interventions designed with an accurate understanding of the family factors described previously and a commitment to the family as a whole, not just to selected members. In family-focused grief therapy, the patient and family meet together to process the events and anticipate the upcoming loss. It continues for a few sessions after the patient's death, helping the family to consolidate the positive achievements made while the patient was still with them. The staff–family conference deserves special mention. It has the potential for being a powerful mutative event and has been known to resolve what had seemed to be knotty problems.

PSYCHOLOGICAL ISSUES FOR STAFF

Anatomy of Stress

The stress on caregivers has been documented in studies across many settings and many countries. Communication problems between doctors and patients lead not only to patient dissatisfaction, but also to lower job satisfaction and self-esteem for the physician. The following stressors are more particular to the oncology setting:

1. *The negative aura* that still surrounds it.
2. *The toxic nature of treatment* and its inevitable side effects.
3. *The occupational hazards* surrounding chemotherapeutic and radioactive agents and the resistant infections to which immunocompromised patients are prone.
4. *The ambiguity of many treatment decisions*, which often breeds conflicts between staff members.
5. *The unsettled social and ethical climate* that surrounds palliative care and terminal care decisions.
6. *The extreme emotional reactions* engendered by cancer and projected onto staff in the form of overvaluation, ingratiation, magical expectation, and covert bargaining, followed by disillusionment and anger.
7. *Conscious and unconscious responses of staff to cancer*. At a conscious level, caregivers depend on their intellectual grasp of the situation and make a point of rejecting myth and superstition. They are proud of their ability to function under pressure with reliable emotional controls and to offer something in all situations, even hopeless ones. Yet, at an unconscious level, there are covert adaptations that are idiosyncratic and irrational, often determined by historical precedent and resistant to change. They may include superstitious beliefs and stubborn fears and feelings. Their eruption into awareness makes the worker feel embarrassed and inadequate. Hence, the feelings are kept secret; the

caregivers may believe that they are alone in having them. Thus, staff workers isolate themselves from one another.

3. *Social isolation*. Outside of the workplace, staff isolation is reinforced by family, friends, and a society that shuns reminders of a frightening topic. Caregivers are not singled out personally, but they may be discouraged from talking about their work and risk being shunned if they do not accept this taboo.

Clinical Spectrum of Staff Responses

Normal Adaptation and Coping. Staff undergo a developmental process when they begin to work intensively with cancer patients. A high initial level of dysphoria, anxiety, sadness, and numbing recedes over the first few months, displaced by the need to demonstrate competence and the ability to survive. Over the next few months, existential issues are allowed to surface, and staff members shape a deeper adaptation that engages their whole personality. Formal support is not necessary, but informal support and peer group interactions are important to a successful resolution.

Adjustment Reaction. Adjustment reactions may occur during the formative stages of providing cancer care, but they also overtake staff members at any point in their careers when a personal loss resonates with their work (cancer in self or loved ones, in particular). These reactions are acute but time-limited and respond well to support or a brief change of venue. Rarely is psychiatric intervention needed. Sometimes adjustment does not improve with time, exposing an underlying psychiatric disorder, often pre-existing and precipitated by work stresses. Psychiatric evaluation is then indicated and may result in transfer to a less emotionally stressful situation.

Supportive Interventions for Staff

Staff support interventions generally fall into three categories: Organizational, individual, and group. Organizational issues include leaders who model realistic expectations and monitor overinvolvement in themselves, as well as others. A team approach encourages the sharing of difficult tasks, prevents a sense of being indispensable, encourages staff cohesion, and eases time off and vacations. Open communication and good orientation procedures are important. Individual responses include peer support, supervisor support, access to an employee assistance program, and other confidential psychological support or psychiatric evaluation. There are many published reports about the benefits of staff support groups in many forms and many settings, but they are anecdotal. These groups are difficult to run but, when they work, participants value them. The group must not be perceived as undermining unit norms or leadership.

Structure must be provided, as needed, to alleviate staff anxiety, while still allowing unscheduled topics and feelings to be aired. Therapy issues must be avoided, and discussions should clearly flow from work-based topics. An active area of new and productive intervention is the development of group-based workshop formats for improving communication. Many have been described and found to be helpful; more are under development.

CANCER IN THE PSYCHOTHERAPIST

Unlike degenerative diseases, with their linear course and associated impairments, cancer may have an episodic course with no loss of cognitive functioning. It inevitably forces a blurring of psychological and, later, overt boundaries in therapy. An all-but-intolerable burden is placed on doctors and patients when therapists who have cancer continue treatment too long because of their own need to deny their fate. This suggests that, difficult as it is, the therapist's most ethical course of action is to set a fixed termination date that allows the patients to transfer to someone else while the therapist is still alive and well enough to negotiate the termination. This may be asking too much of an individual in whom denial is both the main coping mechanism and the only impairment. The help of colleagues is invaluable at this time.

PEDIATRIC PSYCHO-ONCOLOGY

Magnitude of the Problem

The year 2015 will bring an estimated 10,380 new cases of cancer in children under 15 years old, according to the ACS, a projected 13 percent increase from only nine years previously. Leukemias are the most common cancers in children, accounting for 30 percent of all new cases. Brain and CNS tumors are the second most common, making up 26 percent of childhood cancer cases. Caucasian children are more likely to develop cancer than any other ethnic group. The combined 5-year survival rates for all childhood cancers improved from 56 percent in the 1970s to 80 percent currently. However, approximately 1,250 children under the age of 15 are expected to die of cancer in 2015.

Child Patients

Sick children are benefiting from social and legal trends toward empowering them. They are being informed more honestly, and their cooperation is being thoughtfully enlisted. It has become evident that, like adults, children know or intuit a great deal without being told and are relieved by open communication. Its provision is complicated by the need to be age-appropriate and is more painful because of adult vulnerability to suffering in the young. As a result, adult denial often prevails, and children respond by increased behavioral symptoms or withdrawal.

Discussion about Diagnosis and Prognosis

A child can hardly remain oblivious to their diagnosis, given the endless clinic visits, repeated chemotherapies, flagrant symptoms such as hair loss, repeated exposure to other patients, and statements of people around them. Studies note that immediate perception outweighs distant reality; hence, intrusive procedures can loom larger than prognosis, especially in younger children. The developmental stages of death awareness in healthy children are well described but are probably not applicable to children with cancer or to other children who have been exposed to death in an untimely and personally threatening way that results in earlier and keener awareness. Thus, discussions are clearly indicated, but they must follow the child's lead, looking for and responding to his or her sources of anxiety and remaining reassuring about care and comfort if he or she cannot be so about outcome. It has been found that children who received truthful information had less subsequent anxiety than those who did not. However, parents often request and benefit from psychological guidance about how to proceed, because truth telling to young children is strongly molded by the child's circumstances and level of possible understanding. This is not always obvious to even the most loving parents. But they are usually eager to learn and very responsive to guidance.

Psychiatric Diagnosis and Treatment

In approaching psychiatric diagnosis of a child, it is important to know the child's previous developmental profile and to assess current symptomatology, arrests, or regressions. Pre-existing disorders are aggravated by cancer and must be monitored and treated throughout. The stress of cancer can make a mild disorder more severe.

Anxiety, depression, and delirium occur routinely in children and respond well to treatment. Behavioral interventions are the first line of treatment for anxiety, but some situations may require the use of anxiolytic medications. Benzodiazepines are sometimes used but can cause a disinhibition syndrome that may be mistaken for more anxiety. Mild depressive symptoms are common, and often complicated by demoralization. After paroxetine was reported to predispose children to suicide, warnings were issued against the use of SSRIs in children. But experts point out that the evidence is not strong enough to warrant a categorical prohibition and suggest cautious use with close monitoring. Desipramine (Norpramin) is contraindicated owing to several reports of sudden death.

Children suffer from delirium during critical illness, similar to adults. Pediatric delirium occurs in at least 30 percent of critically ill children. As with adults, children with delirium have worse clinical outcomes that those unaffected by delirium, including days admitted to the pediatric intense

care unit (PICU), memory deficits, and longer school absences. Haloperidol has been used for behavioral symptoms associated with delirium in children, such as agitation. There are no randomized placebo-controlled trials of atypical antipsychotics in children with delirium; however, risperidone (Risperdal), olanzapine (Zyprexa), and quetiapine (Seroquel) appear to be effective tools for safe management of behavioral symptoms during delirium. Pediatric psycho-oncologists should be quicker to use medications for behavioral symptoms than child psychiatrists. Starting with small doses, using lowest effective dose for shortest possible time, and diligent monitoring for side effects are advised. Further studies in this area are certainly needed.

Pain Management

Children are at high risk of receiving insufficient pain management. Contributing factors include misplaced beliefs about childhood pain perception, inappropriate use of as-needed dosing, unclear communication of pain experience from young children, and complaints made in the context of normal regression, which is sometimes used to disqualify the complaints. Exhausted, critically ill children, given relief of suffering, may enjoy opiates in a way that upsets some parents and caretakers, leading to lowering of dosing.

Recent findings indicate that the experience of pain in early life may lead to long-term behavioral and psychological consequences. It has also been found that chronic pain, including neuropathic pain, is far more common in children than was previously believed. These findings still wait to be extensively disseminated and translated into proactive and even pre-emptive pain treatment in infants and children.

Management of Stressful and Painful Procedures

Recent work on PTSD in children reinforces the principle that aggressive prophylactic relief during stressful procedures is important not only for current comfort but also for long-term adaptation. Poorly managed painful procedures not only brutalize the child but also result in anticipatory fear of pain, loss of trust and cooperativeness, phobias, and posttraumatic symptoms that do not improve over time. Unlike adults, children are anesthetized for bone marrow biopsies. Multimodal behavioral methods may be taught to parents, which can be extremely helpful during procedures.

Family Involvement

The families of pediatric patients are even more vulnerable than those of adults because of the child's total dependency and the biological, emotional, and social factors that make families uniquely responsible. No

intervention has any hope of success if it does not address the family as much as the child. In one study, mothers had a greater rate of posttraumatic symptoms than their children.

Childhood Cancer Survivors

Child survivors face daunting tasks. They must move beyond lengthy hospital stays or repeated day hospital visits, painful procedures, and severe side effects. They may have endured separation from their family, incurred resentment from their siblings, and missed as much as a year of school. They may emerge with physical handicaps or disfigurements, but the most insidious problems are the cognitive and verbal deficits associated with chemotherapy and radiation therapy. These may be compounded by drug-induced hearing losses. Children need to be monitored and given remedial help promptly.

It has been noted that the child's own perception of risk and the intensity of cancer treatment is more predictive of posttraumatic stress than the objective ratings of health professionals. Hence, it is useful to monitor perceived health status. There is a role for psychological evaluation and support throughout the illness as well as during the post-illness phase.

Children and their families may also feel a sense of victory and relief that helps the children's self-esteem. Nonetheless, it is still important for families to be supported and queried for difficulties when they return for the follow-up appointments, as they have been found to function less effectively than non-affected families. Contacting the child's school counselor and teachers is an effective intervention, because the latter are eager to be educated about special needs and reassured about how to help the child. The ability to reintegrate comfortably into school life and school friendships is a major factor in the child's adjustment.

Consent in Pediatrics

All pediatric consents are third-party consents and carry the possibility of conflicts between parent and child. The parent may make a decision that is not in the interests of the child, as in the refusal of life-saving treatment. The courts have tended to overrule parental authority in such cases. Parent and child may disagree, especially adolescents, although disagreements occur with school-age children as well, presenting staff with painful problems. Psychiatric intervention may help resolve the conflict if the consultant does not impose personal beliefs. Legal recourse is used only after all attempts at conflict resolution have been unsuccessful.

Genetic Testing of Minors

Decisions about whether to offer genetic testing or screening should be determined by the best interest of the child. Genetic counselors discourage

the use of genetic testing to abate parental curiosity and anxiety.

Staff Problems

The maintenance of staff morale, the prevention of dysfunctional norms, the support of faltering members, and the defusing of unconscious competitive feelings toward parents are important ways of contributing to the well-being of all participants. The need of staff members to acknowledge their grief with the repeated loss of children whom they came to know well has given impetus to the custom of memorial services for one pediatric oncology ward, where the staff members gather every few months to remember each child that died in the interim. There is also more attention being paid to families during the mourning process, with outreach offers of support and provision of support groups. Staff members may derive indirect benefit from this greater acknowledgment.

Team Approach

The multidisciplinary approach to patient care is most widely implemented in pediatric cancer patients. Psychiatrists who specialize in child psychology can play a useful role and, where the model is not in place, are in good position to help organize it. Local teachers and clergy should be involved, and community resources should be used to help families reintegrate when they return home.

GERIATRIC PSYCHO-ONCOLOGY

The population of older people is growing rapidly throughout the world. Currently, there are about 35 million people in the United States over the age of 65. That number is expected to double by 2030 and comprise 20 percent of the population according to the US Census Bureau. Cancer is a leading cause of morbidity and mortality in the aging population, with approximately 60 percent of all cancers and 70 percent of cancer mortality occurring in people aged 65 years or older. Of the 24 million cancer survivors worldwide, 61 percent are 65 or older. Psychiatric care of older persons with cancer presents a unique challenge for the clinicians due to differences in phenomenology, assessment, and treatment in this population. Differences in psychological sequelae or specific psychological needs of older cancer survivors have not been well studied, despite their growing numbers. Capacity assessments, informed consent, advanced directives, and surrogate decision making are commonly encountered issues in caring for older cancer patients and should be handled based on the specific needs of the elderly.

Several cross-sectional studies have found lower or similar rates of psychological distress in elderly cancer patients compared to the younger adult population. Longitudinal studies either found no difference or a higher psychological distress in younger adults. It is important to note that elderly cancer patients are at risk for underdetection and undertreatment of their psychiatric diagnoses due to decreased self-report of psychiatric symptoms, cognitive deficits, comorbid medical conditions, poor access to mental health care, and differences in phenomenology. In palliative care settings, there is a need to appreciate the individual differences in physical and psychosocial needs in older adults. The elderly with cancer at the end of their lives are at greater risk for untreated pain and depression and are more likely to commit suicide compared to their younger counterparts.

The goal of identifying and treating these distressing problems is to improve the overall quality of life and the functional status of older people living with cancer. This section will review the assessment and management of common psychiatric conditions in elderly patients with cancer.

Depression in Elderly Cancer Patients

The prevalence of depression is 3 percent in elderly people living in the community, and 10 to 15 percent for those living in nursing homes. Estimates of depression in cancer patients vary widely with averages between 17 to 25 percent. Although many early studies of depression in cancer patients included older patients, there was not often a description of findings based on demographic data such as age. Several studies have

reported that older patients with cancer have lower rates of depressive disorders or self-reported psychological distress than younger patients. Risk factors for depression in elderly cancer patients include loss of spouse, functional disability, inadequate social support, uncontrolled pain, poor physical condition, advanced illness, previous history of depression, other life stresses or losses, family history of depression or suicide, and medications known to cause depression. A study of 420 elderly cancer patients showed that higher levels of symptom severity, lower levels of prior physical functioning, and greater physical functioning deficits predicted higher levels of depressive symptomatology. The prevalence rate for depression was 18 percent. Cancer site, gender, stage of disease, and treatment type had no significant effect in predicting depressive symptomatology. In a longitudinal analysis of the course of depressive symptomatology in elderly breast, prostate, colon, and lung patients, the prevalence of depression was 26.1 percent 2 to 4 weeks after diagnosis, compared to 15.8 percent 1 year after the diagnosis. Lung cancer patients had higher rates of depression likely due to higher symptom severity. In a group of elderly colorectal cancer patients, a longitudinal study examining depressive symptomatology over the first year following cancer diagnosis has shown that gender, race, comorbid conditions, physical functioning, social functioning, and symptoms were significant predictors of depressive symptomatology. Female patients, African Americans, and patients with two or more comorbid conditions exhibited more depressive symptomatology. The highest prevalence rates for depression were at the time of initial assessment, 18 percent. A prospective study in head and neck cancer patients comparing elderly and younger cancer patients has shown that 28 percent of younger and 30 percent of elderly patients had depression, both group of patients had higher rates of depression 1 year after diagnosis. There was no significant difference between younger and elderly cancer patients regarding the prevalence of depressive symptoms.

Assessing depressed mood in elderly cancer patients is difficult. Neurovegetative symptoms of depression such as poor sleep, poor appetite, and decreased energy can be related to the cancer or the treatment; however, the elderly tend to have more somatic symptoms of depression than younger patients. Assessing the cognitive symptoms of depression such as hopelessness, worthlessness, guilt, and suicidal thoughts in addition to the somatic symptoms is important in geriatric cancer patients. Geriatric Depression Scale (GDS), the Center for Epidemiologic Studies Depression Scale (CES-D), and the HADS are commonly used in screening for depression in elderly patients. These scales need to be validated in elderly cancer patients. The NCCN Older Adult Oncology Guidelines have recommended the use of the Distress Thermometer to assess distress in cancer patients.

Differential diagnosis is an important part of the depression assessment

in elderly cancer patients. Although a cancer patient may respond with an appropriately sad mood in the face of initial diagnosis or setbacks such as recurrence, it is important to note that depression is not a normal part of aging or of having cancer. Late-onset first episode depression presents a diagnostic challenge to the clinician as some of these patients may be in the preclinical stages of Alzheimer disease or vascular dementia. Poor memory or impaired attention may be the initial presenting symptom of depression in older persons. Delirium should be ruled out in the presence of disturbance of level of alertness, attention deficits, and psychotic symptoms. Medical conditions should be considered as possible causes for depression.

Several psychotherapy techniques (i.e., interpersonal, supportive, cognitive-behavioral, group, psychodynamic psychotherapy, hypnotherapy, relaxation, and biofeedback) have been used effectively to help elderly cancer patients who are suffering from depression. The efficacy of antidepressants for the treatment of depression in cancer patients has been well established. There is also good evidence that depression in elderly patients with chronic medical illness can be treated successfully. Elderly patients should be started at lower doses of the medications and the dose titration should be slow. SSRIs are usually the first line of pharmacological treatment in depression. Studies in geriatric patients have found SSRIs to be equally effective as TCAs in the treatment of depression. Although TCAs have been the most extensively studied medications for the treatment of geriatric patients, they have been replaced by the newer antidepressants because of a preferable side-effect profile. In treating elderly cancer patients who have depression, it is important to carefully review medications for possible drug–drug interactions.

Of the newer antidepressants, bupropion can have a stimulant-like effect that may be helpful for elderly cancer patients with significant fatigue. Mirtazapine is associated with sedation and weight gain, which may be a beneficial effect in elderly cancer patients. Psychostimulants may be most helpful in elderly cancer patients where the depression is accompanied by psychomotor slowing, decreased energy, and apathy due to their rapid onset of action. ECT is a safe and effective treatment for depression in the elderly. However, it has not been specifically studied in elderly cancer patients.

Suicide in Elderly Cancer Patients

Older persons account for 18 percent of all suicide deaths. The ratio of attempted to completed suicide is 4 to 1 among the elderly, compared to a ratio of 200 to 1 in a younger population. The incidence of suicide is also higher in cancer patients compared with the general population. Identification and prevention of suicide require the definition of risk factors. Physical illness is a common risk factor for suicide in late life.

Elderly cancer patients, and individuals with head and neck, lung, breast, urogenital, and GI cancers, and myelomas, seem to have an increased risk of suicide. It is important to recognize and aggressively treat high-risk patients for depression and address suicidal risk.

Anxiety in Elderly Cancer Patients

Anxiety is not well studied in the geriatric cancer population. In patients with cancer the prevalence of anxiety ranges from 21 to 28 percent. Anxiety may be a component of pain, delirium, or depression. The differential diagnosis of anxiety in elderly cancer patients can be difficult due to several factors contributing to anxiety. Steroids, antiemetics such as prochlorperazine and metoclopramide, alcohol, or benzodiazepine withdrawal are associated with anxiety.

As in younger patients, elderly patients may benefit from meditation, guided imagery, progressive relaxation, and reframing of negative, irrational thoughts. Other psychotherapeutic techniques such as insight-oriented, supportive, and group psychotherapy may also help reduce anxiety. Benzodiazepines are generally useful for short-term relief of anxiety symptoms. The clinician should be cautious in the elderly population. Lorazepam, oxazepam (Serax), and temazepam (Restoril) are better tolerated by the elderly. However, with all benzodiazepines, side effects such as drowsiness, confusion, and motor incoordination must be monitored. One of the greatest morbidities is their potential to cause imbalance in the elderly patient, leading to falls. Patients with dementia or brain injury may have a paradoxical reaction to benzodiazepines and become disinhibited, irritable, or agitated. Many of the antidepressants listed above are also useful in the treatment of anxiety disorders. Stimulants and stimulating antidepressants (i.e., bupropion) should be avoided in patients suffering from severe anxiety. Buspirone (BuSpar) is a nonbenzodiazepine antianxiety medication that can be a helpful anxiolytic in the elderly.

When patients cannot tolerate benzodiazepines or when their anxiety is not well controlled, the use of a low-dose atypical antipsychotic can be considered. Atypical antipsychotics do not have the adverse effects on respiration as do benzodiazepines. None of the antipsychotic medications have been FDA approved for use in the treatment of anxiety disorders in the general adult or elderly population. It is also important to note that the FDA has recently issued a warning against the use of antipsychotics for behavioral disorders in elderly patients with dementia due to increased mortality from cardiac and cerebrovascular events.

Fatigue in Elderly Cancer Patients

Fatigue is a common symptom in older cancer patients. A recent study of

fatigue prevalence has found this symptom to be universal in the elderly age group. In 84 percent of the cases, patients' fatigue interfered with the general level of activity. Fatigue was the most reported symptom in a quality-of-life study among elderly cancer patients. Fatigue remained the most common symptom both at the time of diagnosis and 3 months after diagnosis.

In elderly patients, the common etiologies of cancer-related fatigue are immobility, pain, depression, sleep disorders, anemia, metabolic derangements, undernutrition, infections, side effects of cancer treatment, centrally acting drugs, and direct effects of the tumor. All of these conditions should be considered in evaluation of fatigue. There are several fatigue assessment measures, such as the Brief Fatigue Inventory, the Multidimensional Fatigue Inventory, and Piper Fatigue Scale, none of which have been validated specifically for use in the elderly cancer population. The relationship between pain, fatigue, and depression is important. A statistically significant relationship was found between fatigue and depression, pain, number of medications, and physical function in a study of fatigue in the elderly. Another study has shown that 70 percent of patients with fatigue had a psychiatric comorbidity. Due to the significance of this relationship, patients reporting fatigue should be assessed for pain and comorbid psychiatric conditions, particularly depression.

Treatment of fatigue in elderly cancer patients has not been well studied. The first approach should be to address any potentially identifiable etiologies, such as anemia, depression, insomnia, or any metabolic abnormalities. Exercise has been used as a treatment for fatigue in cancer patients receiving chemotherapy but has not been well studied as a treatment for fatigue in elderly cancer patients. Medications have been used to treat fatigue in cancer patients. Studies on the treatment of fatigue in elderly cancer patients have not been published. Psychostimulants such as methylphenidate or wakefulness-promoting agents such as modafinil have been used successfully for cancer-related fatigue. Modafinil may be a gentler, better tolerated medication for the elderly. Antidepressants with some stimulating effects such as bupropion or fluoxetine may be helpful to improve energy, especially when there are also concomitant symptoms of depression.

Delirium in Elderly Cancer Patients

Delirium is a disturbance in level of alertness, awareness, attention, cognition, or perception that develops over a short period of time (hours to days) and fluctuates during the course of the day. The elderly are particularly susceptible to the development of delirium. An estimated 30 to 50 percent of elderly patients develop delirium during a medical hospitalization. Elderly patients who develop delirium during a hospitalization have an estimated 22 to 76 percent risk of mortality during

that admission. Major risk factors for delirium appear to be pre-existing cognitive impairment, advanced age, and severity of comorbid medical illness.

The etiology of delirium in elderly cancer patients is often multifactorial. Mini-Mental State Examination (MMSE) is a useful, commonly used screening tool for cognitive impairment but it does not distinguish between delirium and dementia. The Confusion Assessment Method (CAM), the DRS, and the Memorial Delirium Assessment Scale (MDAS) are other tools that can be used in the assessment of delirium, as detailed above.

In assessment of delirium in the elderly cancer patient, one of the most challenging problems is to distinguish delirium from dementia. Further complicating the issue, patients with dementia can develop a superimposed delirium. Patients with dementia appear relatively alert and have more stable cognitive deficits, as opposed to the fluctuating symptoms in delirium. The onset of symptoms in dementia is slower, with a chronically progressive decline in cognitive function rather than the relatively rapid onset of symptoms seen in delirium. The management of delirium in elderly cancer patients involves addressing the symptoms of the delirium while investigating and ultimately treating its underlying causes. The treatment of delirium in cancer patients, including the elderly, is extensively reviewed above.

Dementia in Elderly Cancer Patients

There are an estimated 1.8 million elderly individuals in the United States who have advanced dementia requiring extensive care. DSM-5 classifies dementia syndromes under major neurocognitive disorders and allows for continued use of the term dementia. Patients with severe dementia often exhibit emotional and behavioral disturbances that make their care more difficult. Studies have shown that geriatric patients with cancer who are cognitively impaired have a worse performance status, higher incidence of depression, and are at greater risk for death. Alzheimer dementia is the most common type of dementia. It accounts for 60 to 70 percent of all patients with dementia. Another 15 to 25 percent of patients have vascular dementia and approximately 10 to 15 percent will have a mixed dementia. The remaining patients have other diagnoses such as Parkinson disease, dementia of Lewy body disease, and frontotemporal dementia. The recognition rates of cognitive impairment range from 25 to 90 percent depending on the health care setting and the degree of impairment. Recognition of cognitive impairment in cancer patients is especially important due to issues of capacity and decision making in older adults. A diagnosis of dementia does not automatically determine a lack of capacity. However, when cognitive impairment is identified, it should prompt further evaluation. Often a psychiatric evaluation will be necessary to

determine the patient's ability to participate in the decision-making process. When a patient is deemed not to have capacity, a health care proxy must be determined. Given the potentially poor quality of life for an older cancer patient with dementia, the extent of aggressive treatment should be carefully determined in a patient with severe cognitive deficits. Cognitive impairment may present special problems in the control of pain and other symptoms due to limitations in patient's self-report or active participation in assessment tools.

Dementia has also been noted to develop secondary to cancer treatment, including chemotherapy, radiation treatment, immunotherapy, hormonal therapy, and supportive medications. It is not well understood how chemotherapy or radiation therapy might affect cognition and whether older patients are at greater risk. A study of lung cancer patients has shown that the treatment of chemotherapy-related anemia could increase the cognitive capacity of elderly patients. A recent review of the literature has concluded that there was not sufficient information on cancer treatment–related cognitive impairment in the elderly due to the small number of individuals studied, the variety of cancer types and treatments surveyed, the low number of longitudinal studies, and the limited battery of neuropsychological tests employed.

CANCER RISK AND SURVIVAL

The role of psychosocial factors in altering cancer risk and survival fascinates the public but the discussion is often blighted by wishful thinking and genuine confusion. Cancer risk and survival are often conflated, and great emphasis is placed on the direct role of the individual's psychoneuroimmune mechanisms in preventing or curing cancer, while ignoring many psychologically mediated effects that are far stronger and better established. There are many distorted beliefs and overinterpretations on the subject that psychiatrists need to know the current state of the art to better understand and educate their patients.

Cancer biology: The natural history of a given cancer begins when changes in a cell's DNA establishes the potential for this cell to become malignant. This *cancer initiation* precedes by many years (anywhere from 3 to 30 years) the second event, known as *cancer promotion*, in which growth suppression is lost and malignant multiplication begins. This again precedes by a period of months to years the appearance of clinical symptoms. Finally, there is a variable period between the appearance of any clinical symptoms and the seeking of care. Cancer risk can relate to the risk of initiation or the risk of promotion. Cancer survival is measured from the time of diagnosis, but this ignores the important factor of delayed diagnosis, for which initial disease staging is a partial, but incomplete, solution.

Psychosocial factors: One attempt at clarifying the multifaceted impact

of psychosocial factors on cancer is to think in terms of: (1) The *external loop*, which refers to ways in which psychological, behavioral, and socioenvironmental factors affect the person's behavior and lifestyle and, hence, cancer risk. Even there, the actual behaviors must be teased out from their psychological determinants. (2) The *internal loop*, which refers to the ways in which the person's psychological status affects cancer risk and survival via changes in the internal milieu. This effect can be mediated in many ways, such as neurological, immune, and endocrinological mechanisms, individually and in varying relations to each other.

Psychologically Mediated Environmental Factors

Cancer Risk. By far the largest role in the development of cancer is played by inescapable environmental insults, such as radiation, viruses, and chemicals, acting cumulatively over time on each individual's predetermined complement of potentially cancer-causing genes. The individual's exposure to injury is in part out of his or her control and in part mediated by social and psychological factors over which he or she has variable control. By far the greatest volitionally determined contributors of psyche to cancer incidence are personal lifestyle and habits. Smoking cessation, abstinence from alcohol, more exercise, and better eating habits would probably lead to a substantial decrease in the incidence of cancer in the United States.

Cancer Survival. The greatest contributor to decreased survival is poor use of screening, because early diagnosis is still the best predictor of prolonged survival for most cancer types. Many psychologically determined factors can play a role in delayed diagnosis. Patients with cognitive impairment, mental retardation, and psychotic disorders are especially at-risk populations for delayed presentation to health care professionals, which makes cancer screening efforts immensely important in those groups.

Research Review

Methodological Considerations. To be reliable, the studies of the role of psychosocial variables should involve large populations over long periods of time. In commonly used case-control studies, the index population is usually unrepresentative of broader populations, and the matching is incomplete; hence, it is unlikely that one can generalize the findings to larger unselected populations unless the effects are so large as to be inescapable. In retrospective studies, the index population is equally unrepresentative, but these are essential pilot studies pointing the way to later larger ones. In prospective cohorts, researchers can control for the confounding variables, but the task remains methodologically daunting.

Current Findings. All this notwithstanding, more than 50 psychosocial factors have been claimed to play a role in cancer initiation or promotion. A wave of recent reviews has shown that few of them are valid, in many cases because the quality of the studies is poor, but also because the question is difficult to answer, given that cancer initiation begins 20 years before there is any clinical evidence. The lack of impact of bereavement has been demonstrated with recent, large, population-based studies. The role of external stress had been thought to be significant, but later studies have disproved this.

On the other hand, several recent studies have shown that social support is a significant positive survival factor, a fact that has important implications for mental health professionals. Social class has a variable effect on cancer risk, with some cancers labeled *cancers of poverty* and others labeled *cancers of affluence*. However, low social class has an unequivocal negative effect on cancer survival. One must take into account the fact that these individuals are less knowledgeable about and less able to avoid toxic exposures and unhealthy lifestyles and, more to the point, may lack access to decent, not to mention excellent, medical care.

Race is also a strong predictor of decreased survival, but the effect is much decreased or even disappears after correction for social class and shows great variability within racial groups. Two recent studies of race and ethnicity were published back to back, one showing significantly poorer results for nonwhite children with acute lymphoblastic leukemia, and the other showing that, given equal access to effective state-of-the-art therapy, children of both races could expect the same rate of cure. The experiment was possible because acute lymphoblastic leukemia has an acute onset and a well-defined course of treatment. Given access to the treatment, parents could be depended on to bring the child for a life-saving treatment.

The effectiveness of interventions aimed at improving lifestyle and health behaviors has often been disappointing, usually because the complexity of the problem is underestimated or because unrealistically short-term results are being demanded, or both.

Internally Mediated Risk and Survival Factors

Recent developments in psychoneuroendocrinology or psychoneuroimmunology have confirmed the feedback loop between feelings and perceptions, on the one hand, and neurological, endocrine, and immune changes, on the other. These are, in small part, an updating of classic psychosomatic concepts. More urgently, this is the area of research in which recent developments lend theoretical support to current popular beliefs about the power of the mind over the body.

Personality Traits, Feeling States, and Coping Styles. There are weak and conflicting results on the role of personality in cancer prevalence and

survival, beyond what is mediated through risk-avoidant behavior. Studies are handicapped by the absence of comprehensive prognostic indicator profiles. Meanwhile, no strong relationship has been established between personality style and disease outcome. Feeling states, such as level of distress or anxiety, have not been found to have any predictive value. A weakly significant association between passive coping styles and poor disease outcomes has been demonstrated in several studies using measures of so-called fighting spirit, helplessness and hopelessness, stoic acceptance, and anxious preoccupation. The role of denial, which can be viewed as the operative factor, remains an open question and points again to the complexity of what is being measured.

Immune Mechanisms. Numerous studies have shown small, but statistically significant, differences in bereavement, disease states, and many kinds of stress. In cancer, a conditioned negative immune response has been found with chemotherapy. An improved immune response has been seen following a group intervention in melanoma patients, and differences have been found in the families of cancer patients depending on their coping effectiveness. However, all of these studies are small and many are methodologically poor. Furthermore, there is a strong likelihood of a publication bias against negative results, so no claims can be supported at this time. Nevertheless, because immunological studies allow for well-defined dependent and independent variables, better methodology is achievable, and it is one of the areas believed to warrant further study.

Endocrine Mechanisms. The discovery that breast cancers with estrogen-sensitive receptors are associated with a better outcome is of clinical use when receptor typing is done on biopsy specimens to guide subsequent management. Because the receptors are responsive to emotional mediation through the HPA axis, a relationship to emotional states and psychological adjustment can be postulated, although studies have shown conflicting results.

Effect of Psychological Interventions

The direct effect of psychological interventions on cancer risk through internally mediated mechanisms has not been studied and may never be fully disentangled from the large, well-known externally mediated effects on lifestyle and health behaviors. However, the few studies that showed some effect on survival have not been duplicated, despite several efforts. Meanwhile, psychosocial interventions have been repeatedly shown to have a positive effect on quality of life, whether it is the direct outcome variable or is measured through mediating variables, such as improved mood, self-esteem, interpersonal relations, and level of functioning. Given the unequivocal improvement in quality of life shown by psychosocial

intervention studies, their continued use and further development is important.

However, the premature and overenthusiastic acceptance of a major role for internally mediated psychological factors has many negative implications for patients. They feel anguish over their inability to have a positive attitude, and families conspire to worsen their guilt by demanding an upbeat stance from patients who are far too ill to do much except endure. Given the current state of the art, it would seem best to work actively to improve patients' quality of life, while at the same time encouraging them to enjoy these improvements for their own sake, independent of their belief about any effect on survival time. Group interventions are particularly desirable because they tap into the well-demonstrated protective effect of social supports, even as they address their other stated goals. It is also important to educate patients about the fact that several recent reviews have concluded that there is no good evidence at this time that cancer survival can be affected by any psychological factors other than those that mediate the presence of social supports, financial resources, and access to early and competent care.

ALTERNATIVE TREATMENTS

Complementary and alternative medicine (CAM) is often defined as any medical system, practice, or product that is not thought of as standard.

Respect for the role of diet in cancer risk is now well accepted. These are viewed as effective, safe, and inexpensive adjuncts to traditional treatment. Unproven methods are often provided in the setting of an intense charismatic relationship, with great emphasis on faith, support, and encouragement. Patients regard this experience alone as valuable. The motives that drive providers range from mercenary ones to idealistic beliefs about the power of mind over body. The motives that drive patients include universal concerns, such as the often justified fear of failure of traditional treatment, the hope for a miraculous cure, and the desire to achieve mastery and control. The stigma of cancer propels people irresistibly into magical thinking. The toxicity of traditional treatments also makes patients so anxious that other options become compelling. The amount of out-of-pocket expenses spent on nontraditional treatment is in the billions. In one study, a large percentage of patients at a tertiary medical center were seeking alternative treatment while still in traditional treatment. They did so early in their course of treatment and took both simultaneously, dropping the traditional treatment before they dropped the alternative one. Educated and wealthy patients were best represented. The main treatments were physical ones to cleanse and to strengthen the body; special diets to purify and fortify the body, as well as avoiding harmful substances; immune therapies; faith healing; psychological methods to improve mental attitudes and to reinforce the body's defenses; and many Eastern

modalities such as yoga, tai chi, and meditation that promote improved discipline and decreased anxiety.

Many such treatments give patients a sense of purpose and control. The last category, improving attitudes, is most troubling to the psychiatrist, because it leads so readily into blaming the victim. Patients are told that they can influence their cancer by mental efforts. The unspoken corollary is that further disease will be due to inadequate efforts and, worse yet, that they caused the cancer in the first place. This plays directly on attribution mechanisms to which cancer patients are already universally prone. One can also see desperate family members grasping at these explanations and trapping the patient ever deeper in a web of guilt and responsibility as a means of coping with their own grief and frustration or an angry patient blaming a family member for causing the cancer.

The psychiatrist must remember that many patients are secretly involved in alternative care that makes them embarrassed and distant if they cannot share it. As time passes, if their condition worsens, the burden of failure and guilt makes a greater contribution to psychic distress. Hence, clinicians must encourage disclosure and must help patients extract something positive from the experience. Active cooperation from physicians in encouraging their patients to use the nonharmful modalities, especially the ones for which there is a consensus about benefit, increases the patient's trust and adds to the placebo effect and synergy of all the patient's treatments. This is difficult to do if patients and families retreat into anger at the establishment and reduce cognitive dissonance by clinging desperately to their alternate beliefs.

Given how much cachet these therapies had achieved, recent surveys in breast cancer patients suggested that between 48 and 66 percent of patients are trying some form of CAM. A recent review showed that the current data available from RCTs fail to demonstrate the efficacy of CAM therapy as curative. A small number of CAM therapies may help patients feel better and recover faster. Aromatherapy and massage enhance the well-being of cancer patients.

One way in which the positive effects of complementary medicine are being brought into the mainstream, while discouraging the dubious offerings, is through the creation of complementary or integrative medicine centers within traditional medical centers. They shun herbal remedies but offer acupuncture, nutrition counseling, massage, meditation, gentle yoga, music therapy, and other disciplines geared to help patients with stress and pain. Patients are generally pleased that the help they had been seeking surreptitiously is now accepted by their physician. Herbs and so-called food supplements remain unregulated and can be dangerous. *Aristolochia fangchi* was found to cause urothelial cancer in a group of Belgian patients, and patients who use *St. John's Wort* to medicate their depression without telling their physician are, in effect, taking a tricyclic antidepressant drug

without supervision, and one whose side effects have been better demonstrated than its efficacy. The National Cancer Institute (NCI) and the National Center for Complementary and Alternative Medicine (NCCAM) are sponsoring a number of clinical trials at medical centers to evaluate CAM therapies for cancer.

FUTURE DIRECTIONS

The study of the psychological aspects of cancer continues to lead to many fundamental human issues, starting with the basic existential confrontations with death and the sense of self. Despite continuing progress, the public continues its fascination with the role of psychological factors in altering the course of cancer, while ignoring the 50 percent decrease in cancer occurrence that increased exercise, better eating habits, and stopping smoking and drinking would achieve. The issues associated with genetic advances illustrate the double edge of scientific progress, great promise, and great psychological stresses for which physicians are not prepared.

Psycho-oncology is one of the disciplines whose vitality and scientific productivity encouraged the American Board of Psychiatry and Neurology to make psychosomatic medicine a formal subspecialty. Integration of a psychiatrist into an oncology team which is familiar with the psychological aspects of cancer, which can identify vulnerable cancer patients at risk, and which can recognize and treat the common psychiatric disorders has become fundamental. Psycho-oncologists can look forward to further research to improve the treatment of psychiatric syndromes in cancer patients, to maximize the ability to identify high-risk patients and to offer them support, and to come closer to translating basic science insights into clinical practice.

WEBSITES

The American Psychosocial Oncology Society (APOS) has a website at http://www.apos-society.org/clinical/default.asp. There is also an APOS hotline at 1-866-APOS-HELP, where individuals can call with cancer-related psychological problems and questions, and are given appropriate referrals.

The International Psycho-Oncology Society (IPOS) has a website at http://www.ipos-society.org/index.htm. A multilingual core curriculum on psycho-oncology for professionals is available at the website.

The NCCN has a website at http://www.nccn.org and provides the complete library of NCCN Clinical Practice Guidelines in Oncology. Besides treatment guidelines, they include detailed supportive care guidelines (including distress) and the latest detection, prevention, and risk guidelines.

The National Cancer Institute (NCI) maintains an extensive website at

http://www.cancer.gov, and the ACS has one at http://www.cancer.org. Both of these sites contain reliable information for the public, patients, families, survivors, and professionals. The NCI maintains a list of current research protocols.

References

*Ahles TA, Root JC, Ryan EL. Cancer- and cancer treatment-associated cognitive changes: an update on the state of the science. *J Clin Oncol*. 2012;30:3675–3686.

American Cancer Society. *Cancer Facts and Figures 2012*. Atlanta, GA: American Cancer Society; 2015.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

*Breitbart W, Alici Y. Agitation and delirium at the end of life: "We couldn't manage him." *JAMA*. 2008;300(24):2898–2910.

Breitbart W, Alici Y. Evidence-based treatment of delirium in patients with cancer. *J Clin Oncol*. 2012;30(11):1206–1214.

*Breitbart W, Alici Y. *Psychosocial Palliative Care*. New York: Oxford University Press; 2014.

Breitbart W, Marotta R, Platt MM, Weisman H, Derevenco M. A double-blind trial of haloperidol, chlorpromazine, and lorazepam in the treatment of delirium in hospitalized AIDS patients. *Am J Psychiatry*. 1996;153(2):231.

*Breitbart W, Poppito S. *Individual Meaning-Centered Psychotherapy for Patients with Advanced Cancer*. Oxford: Oxford University Press; 2014.

Breitbart W, Poppito S, Rosenfeld B, et al. Pilot randomized controlled trial of individual meaning-centered psychotherapy for patients with advanced cancer. *J Clin Oncol*. 2012;30(12):1304–1309.

Breitbart W, Rosenfeld B, Gibson C, et al. Meaning-centered group psychotherapy for patients with advanced cancer: a pilot randomized controlled trial. *Psycho-Oncology*. 2010;19(1):21–28.

Breitbart W, Rosenfeld B, Pessin H, et al. Depression, hopelessness, and desire for hastened death in terminally ill patients with cancer. *JAMA*. 2000;284(22):2907–2911.

Breitbart W, Rosenfeld B, Pessin H, Applebaum A, Kulikowski J, Lichtenthal WG. Meaning-centered group psychotherapy: an effective intervention for improving psychological well-being in patients with advanced cancer. *J Clin Oncol*. 2015;33(7):749–754.

*Chochinov HM, Wilson KG, Enns M, Lander S. Depression, hopelessness, and suicidal ideation in the terminally ill. *Psychosomatics*. 1998;39(4):336.

*Dalton SO, Boesen EH, Ross L, Schapiro IR, Johansen C. Mind and cancer: do psychological factors cause cancer? *Eur J Cancer*. 2002;38:1313.

Den Ousten BL, Traa MJ, Thong MS, et al. Higher prevalence of sexual dysfunction in colon and rectal cancer survivors compared with the normative population: a population-based study. *Eur J Cancer*. 2012;48:3161–3170.

Fang F, Fall K, Mittleman MA, Sparen P, Ye W, Adami H, Valdimarsdottir U. Suicide and cardiovascular death after a cancer diagnosis. *N Engl J Med*. 2012;366:1310–1318.

*Holland JC, Breitbart WS, Jacobsen PB, Loscalzo MJ, McCorkle R, Butow PN. *Textbook of*

Psychooncology. 3rd ed. New York: Oxford University Press; 2015.

*Holland JC, Golant M, Greenberg DB, et al., eds. *Psycho-Oncology: A Quick Reference on the Psychosocial Dimensions of Cancer Symptom Management*. 2nd ed. New York: Oxford University Press; 2015.

*Holland JC, Weiss T, Nelson CJ, Roth AJ, Alici Y. *Geriatric Psycho-Oncology*. New York: Oxford University Press; 2015.

Hudson MM, Mertens AC, Yasui Y, Hobbie W, Chen H. Health status of adult long-term survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. *JAMA*. 2003;290:1583.

*Kangas M, Henry JL, Bryant RA. Posttraumatic stress disorder following cancer: a conceptual and empirical review. *Clin Psychol Rev*. 2002;22:499.

Kissane DW, Love A, Hatton A, et al. Effect of cognitive-existential group therapy on survival in early-stage breast cancer. *J Clin Oncol*. 2004;22(21):4255–4260.

Members NC-RFP. National Comprehensive Cancer Network (v.1.2014). Cancer-related fatigue. Available at http://www.nccn.org/professionals/physician_gls/pdf/fatigue.pdf. Accessed May 15, 2015.

Mitchell AJ, Chan M, Bhatti H, et al. Prevalence of depression, anxiety, and adjustment disorder in oncological, haematological, and palliative care settings: a meta-analysis of 94 interview-based studies. *Lancet Oncol*. 2011;12:160–174.

National Cancer Institute. *PDQ® Cancer-Related Post-traumatic Stress*. Bethesda, MD: National Cancer Institute; 2015. Available at http://cancer.gov/cancertopics/pdq/supportivecare/post-traumatic-stress/Patient.

*National Comprehensive Cancer Network (NCCN). *Clinical Practice Guidelines in Oncology: Older Adult Oncology, version 1.2015*. http://www.nccn.org/professionals/physician_gls/PDF/senior.pdf.

*National Comprehensive Cancer Network (NCCN). *Clinical Practice Guidelines in Oncology: Adult Cancer Pain, version 2.2015*. https://www.nccn.org/professionals/physician_gls/f_guidelines.asp#pain

*National Comprehensive Cancer Network (NCCN). *Clinical Practice Guidelines in Oncology: Distress Management, version 1.2015*. https://www.nccn.org/professionals/physician_gls/f_guidelines.asp#distress

Silver G, Kearney J, Traube C, Atkinson TM, Wyka KE, Walkup J. Pediatric delirium: evaluating the gold standard. *Palliat Support Care*. 2014;13(3):513–516.

Smith SK, Zimmerman S, Williams CS, et al. Post-traumatic stress symptoms in long-term non-Hodgkin's lymphoma survivors: does time heal? *J Clin Oncol*. 2011;29:4526–4533.

▲ 27.9 End-of-Life and Palliative Care

MARGUERITE S. LEDERBERG, M.D., XIOMARA ROCHA-CADMAN, M.D., AND MINDY GREENSTEIN, PH.D.

Physicians spend many years working to forestall death, and when a patient's death is close at hand, it is their task not only to ease it but also to remember that *a patient who is dying is a patient who is still alive!* There have been times when this was easier said than done.

THE CHANGING WORLD

In the 20th century, especially after World War II, science developed beyond imagination, adding some 20 years to the life span in developed countries. Medicine today is a far cry from the doctors who came with a black bag but provided most of the care through compassionately shared explanations and predictions. Their reliable presence earned them trust and gratitude from patients and families. While payment was necessary, it was really the shared feelings that sustained their sense of meaning.

In contrast, medicine today requires mastering ever more sophisticated knowledge, skills, and procedures. While medical science is still driven by a fundamental desire to help, its relentless expansion makes demands on caregivers, even as it arms them for the fight. This tension can be particularly difficult to navigate when caring for patients at the end of life, as Atul Gawande describes so passionately in his acclaimed book, *Being Mortal*. He particularly emphasizes the importance of caring about patients' *total well-being* rather than focusing primarily on illness. This chapter will shortly discuss the many ways in which physicians and other health care providers have used their ingenuity to improve this situation.

DEATH, LIVING, AND DYING

Death follows many trajectories, among which peaceful old age has never been the most common. It erupts in the midst of health, in war and peace, or creeps along an episodic or deteriorating course. Different trajectories bring different experiences, and any death radically alters the future *life* trajectory of the bereaved. Death from chronic illness forces sufferers and their loved ones to live in its presence for a prolonged period. *Untimely* deaths take their own special toll on the survivors, especially when suicide was the chosen method.

The concept of *timely* deaths is more problematic. Patients of any age can undergo such deterioration and suffering that death is viewed as liberation. Some elderly people accept death, but many elderly patients find death premature, and many families wish desperately to extend the life of a beloved parent. In such cases, only the medical caregivers think in terms of

timeliness, and care must be taken lest the concept become a back door through which ageism creeps into geriatric care or, more commonly and painfully, lest it widen the empathy gap that so troubles patients and families.

There is also a distinction of great psychological consequence between physical and social death. While the bearer endures physical death, the entire community experiences social death. In a car accident, the person dies an immediate physical death but can be said to have died a social death only after being buried and mourned and after the family begins to cope with the void left behind. Contrast this with a neurodegenerative disease, whose sufferers lose more and more of the roles that defined their social existence before dying a physical death, having been watched for years by grieving caregivers.

Thus, good end-of-life care must address the physical, psychological, spiritual, familial, social, and cultural characteristics that define the unique life that will be extinguished. The term *dignity* has been suggested as an organizing principle, perhaps because it points to the awesomeness of the event and the respect that it should induce in every witness. Developed further under the name "dignity therapy," its goal can be described as "care that honors and protects—indeed cherishes—those who are dying, that conveys by word and action that dignity resides in people, not physical attributes, and that helps people to preserve their integrity while coping with unavoidable physical insults and losses." This requires medical caregivers who derive satisfaction from working toward this goal. A smile, a few moments of peace, scenes of reconciliation and loving farewells—these are precious gifts from those who are cherished.

AWARENESS OF DEATH

Dying has been described as a physical process that begins at birth, but, more importantly, living is a state of mind that endures to the last breath. Thus, the goal of good palliative care should be the same as any care, namely helping the patient *live* to the fullest under any condition. Yet, there is an emotional chasm between thinking about one's death and seeing its actual contours. It has been said, most notably by Ernest Becker, that denial of death is at the root of many, if not all, social and cultural accomplishments, and it is undeniable that the strength of many religious beliefs is buttressed by their effectiveness in helping humans cope with suffering and death. Yet, Sigmund Freud believed that death anxiety represented fear of abandonment and loss while fear of death arose because the human psyche could not conceive of its own non-being. That is technically true, but incomplete. A number of people accept their death with equanimity. Some are heard to say they feel comfortable joining Mother Earth, some are comforted by the better life they believe will follow, bypassing altogether the question of being or non-being. Difficulty coping

with it is usually associated with long standing psychological issues, which had never been tackled effectively, but these are not necessarily major diagnoses. Some existential therapists believe that helping the living grow in the face of their mortality is at the core of strength. This may be useful to some patients, but not likely for the ones whose mortality is already their relentless and increasingly intimate companion.

Theories aside, the involved emotions and grief can be so difficult to manage that many patients and families have a shifting perspective in which awareness alternates or coexists with denial. This paradox is characteristic of ill patients and may confuse medical staff trying to meet the patient's needs, in different ways. It is also seen when families make decisions for a permanently unaware relative. Ultimately, confrontations with death generate over-determined responses in *everyone*: patients, families, and sometimes health professionals, each playing out a personal script on a personal timetable. Discrepancies and misunderstandings are difficult to avoid and often become intense under such stress, all of which are the physician's responsibility to manage! One would be hard put to find discussions of what this responsibility really entails in medical texts of 20 years ago, but recent decades have seen major changes.

HUMANISM IN MEDICINE

Medicine used to be a tightly closed shop, and still is from the strictly medical point of view. It now looks appreciatively on many other disciplines that were eager to help patients and that both studies and doctors' own experience have proven to be helpful. In addition to nurses, psychologists, social workers, and physical therapists, all of whom have always been there, there are now art therapists, music therapists, and writers who volunteer to help patients write their stories, to name a few. There are also support groups of all stripes to help patients and family members, most of them originated according to the needs of a given ward or hospital, and the creativity of the group facilitators. Participants have voted with their feet, and described improved mood and enriched lives. Groups have also been helpful to residents and fellows in many guises, including T-groups and journal clubs.

Humanism has extended further into medical education, which now includes expanded communication skills and ethics education, as well as cultural and religious sensitivity. In a large number of major hospitals and medical schools, senior faculty members have spontaneously produced films to help doctors understand illness from the patient's perspective. The most moving are the ones that describe a young doctor's terminal illness. The community is also exposed to expert panels of patients, former patients, and family members, as well as films and plays, selected for their applicability. The questions-and-answer sessions that follow provide meaningful exchanges between caregivers, patients, and family members.

Even medical students are eager to participate. Stanford students, feeling that sexuality was not adequately covered in their schedule, arranged a well-attended evening course taught by a professor from a science department on the university campus. It was well attended and subsequently became part of the curriculum. Yale second year students, eager to help the nearby poor community, organized a Saturday clinic, backed by professors impressed by their desire to help. It went well, and by year's end, they presented the dean with a full roster of students ready to take over the following year.

Palliative Care

"Palliative care" once was a way of saying that there was no more treatment to offer. It was more an announcement than a treatment. It has, however, undergone a welcome change. Like many important social changes, the current interest in palliative care goes back a long way. Two popular innovators contributed to a sea change. In 1969 Elisabeth Kübler-Ross published a very popular book, *On Death and Dying*, which vividly described the psychological experience of dying patients. Suddenly, they became people with an inner life, a meaningful journey, and important intrapersonal and interpersonal experiences. Next, the work of Dame Cecily Saunders at St. Christopher's brought a fully elaborated model of good palliative care to world attention in the 1970s. Her insistence on the uses of opiates at whatever dose was needed to relieve patient suffering aroused medical disapproval and widespread public approval, leading to expansion of the hospice movement in the United Kingdom and a lively grassroots movement in the United States.

Since then, there has been a steady stream of dedicated physicians who devoted themselves to developing palliative care within hospitals, such as Katherine Foley, who pointed out that "competence by itself is not enough, it must be paired with compassion." The value of palliative care was quickly recognized, especially since it enhanced quality of life for patients of any age, with any kind of diagnosis, and during curative treatment as well as terminal care. It became established as a subspecialty of medicine in 2006, and trains physicians, psychologists, advance care nurses, social workers, and chaplains, as well as many other dedicated specialists and volunteers. The demand for them has become so great that there is a shortage of them.

The Introduction of Medical Ethics into Hospitals

The initial rapprochement between ethicists and hospital physicians was difficult without at least one of them already able to function as a bridge. However, in keeping with the flexibility of the era, ethics committees did become well integrated in their hospitals and are now mandated by law. Committee chairs are selected for extensive and appropriate medical or

nursing experience, a thorough knowledge of the ethical issues, and a high level of experience. They continue to keep the members up to date and bring in speakers as needed.

Most important, they staff a highly valued consultation service that can support a physician in especially demanding conditions. The ability to prolong both life and the dying process in the context of individual patients' and family members' disparate values, goals, and levels of capacity can put physicians into an ethical bind, which also has psychological consequences for their own well-being. Doctors must recognize that managing them is a collaborative effort. In addition to nurses, there are also ethics committees who can mentor them about the various ethical issues involved.

Since terminal situations are the most tragic, the consultation service is often called and may support both family members and members of the team. That support may be a greater factor in resolution than any ethical argument and confirms the value of pairing competence with compassion.

PSYCHOLOGICAL REACTIONS OF PHYSICIANS

All humans are equal in the face of death, but physicians are expected to put aside their personal reactions to remain focused on the care of any dying patient, as well as the family. Yet, there is evidence that doctors may have more death anxiety than others and that some become physicians to conquer that fear. Regardless, long exposures to disease and death can breed anxiety, despite the appearance of familiarity, and, since patients and families cannot see their doctor falter, the practice of medicine demands high intellectual coping skills. Fortunately, this work is profoundly rewarding for most seasoned physicians, who had enough resilience and compassion to get through all the hard moments and benefit from the learning opportunities. Time and experience also facilitate resilience.

However, today's physicians, especially those still in graduate training, are dealing with a society with increasing expectations. Patients and families read a wide range of medical articles—sometimes useful, sometimes inflammatory—and may demand more ineffective treatments. Others may manifest uncontrolled grief or anger, sometimes becoming threatening. In the middle of the hectic pace and complex demands typical of today's hospitals, facing many such confrontations with death can arouse threatening emotions, especially in younger physicians. Rare, but perhaps deepest, is an inchoate fear that can be so rapidly submerged that it is barely experienced, much less explored! Basic self-protective behaviors such as tamping down one's emotional responses are inevitable. If recurrent, they will eventually lead to feelings of inadequacy and anxiety which may feed upon themselves.

Also, when facing "bedrock" human pain, tired physicians may suddenly realize that some personal experience has been awakened. They tamp that down as well and correctly focus on tackling the situation at hand.

However, recurrent experiences like this can ultimately impede their ability to be truly present with their patients. In all these cases, if physicians in this position avail themselves of supportive resources, constructive discussions can liberate them and help them feel stronger.

Table 27.9–1.
Risk Factors for the Development of Aversive Reactions in Physicians

The physician:
Identifies with the patient: looks, profession, age, character, etc.
Identifies the patient with someone in his own life.
Is currently dealing with a sick family member.
Is recently bereaved or dealing with unresolved loss or grief issues.
Feels professionally insecure.
Is fearful of death and disability.
Is unconsciously reflecting feelings felt or expressed by the patient or family.
Cannot tolerate high and protracted levels of ambiguity or uncertainty.
Carries a psychiatric diagnosis, such as depression or substance abuse.

Adapted from Meier DE, Back AL, Morrison RS. The inner life of physicians and care of the seriously ill. *JAMA*. 2001;286:3007.

This kind of denial is not universal. Many rise to the demands of their life choices, to the benefit of patients, trainees, and colleagues. However, many of today's students have not previously been exposed to this kind of suffering, since death has moved out of the home and into the hospital. This is why it is particularly important for institutions to support young physicians as they navigate this difficult terrain. Even those who usually navigate it well might need help from time to time. Fortunately, the last 25 years have seen welcome developments in the profession's awareness of the need for support and training of staff. Risk factors are summarized in Table 27.9–1.

Helping the Helpers

Help for medical professionals and staff must address environmental, interpersonal, and intrapsychic dimensions, especially in wards with prolonged illness and high mortality.

A supportive environment acknowledges the stressors and provides:

- Leaders who validate emotional realities and make clear but understandable demands
- A team approach that provides backup, prevents isolation, and precludes feeling too irreplaceable
- The use of practice interviews for a range of difficult situations, with a trained leader using actors or real patients
- The presentation of seminars and case conferences to discuss difficult

cases, the ethical and existential issues that they arouse, or any reaction that a trainee wishes to discuss

- The use of a group for first-year trainees with an outside leader familiar with the hospital environment they face
- The availability of confidential psychological support without prejudice

Doctors can be more authentic with patients if they are allowed to be authentic with themselves and each other, within the proper bounds of the workplace. This extends across wider groups as long as they have shared goals, as proven by some of the many humanistic programs offered to the wider hospital community in many medical centers, such as Memorial Sloan Kettering's *Art of Medicine* series. The result is a more comfortable atmosphere, easier relationships, and, above all, patients and families who benefit from the staff's ability to affirm shared goals.

The trainee, or any physician who is troubled or underfunctioning in any way, should be approached for an individual discussion with a trusted supervisor, and, if indicated, referred to a seasoned psychiatrist who is familiar with the setting.

One approach, originally introduced by hospices, is to hold staff memorial services at regular intervals, during which all patients who had died in the interim are named and remembered with stories, tears, and chuckles. These have now been instituted on many hospital services and are highly valued by staff. Some wards invite family participation, which is deeply valued.

COMMUNICATION BETWEEN DOCTOR AND PATIENT

Good communication skills are essential to the success of any clinical encounter. Poor skills can create vicious cycles, as the physician cannot identify the patient and family values that should influence medical decisions, cannot develop a solid treatment plan, and cannot assuage uncertainties. This generates chronic anxiety and dissatisfaction in patient and family and increases the likelihood of inconsistent information from other members of the team. Finally, without effective communication, the patient and family experience the inconsistencies that often accompany their sense—right or wrong—of vague treatment goals, which lead them to demand yet more end-of-life health resources.

Guidelines to Good Communication

Communication skills can be learned in many ways, including experiential exercises, retreats, and workshops. Physicians who undergo training in this area report greater comfort and increased feelings of self-worth as they provide terminal care. Interactive teaching modules with role-plays are particularly effective. The experiential, interactive, and group aspects help clinicians detoxify their own insecurities and find support in one another.

There are also didactic teaching modules and articles with clear and useful guidelines about breaking bad news. Despite the challenges posed by these tasks, the literature indicates that the courses increase confidence and improve communication. For an immediate example, a scenario follows:

What does the patient know? Always inquire about what the patient and family know before providing new information. "We are here to discuss the further plans for your care, but, before that, I would like to ask what concerns or questions you have about your illness and the way things are going." Physicians are often startled to see gaps shaped by lack of information as well as culture, education, socioeconomic class, and religion. When ignored, these latter factors generate serious misunderstandings. Clarify cultural expectations if indicated.

How much does the patient want to know? Historically, families have generally "protected" loved one from knowing the future. This is still the case in many nations, but most western countries view this as an abuse. Patients have the full right to receive all information, but are not deprived of the right *not* to know. The corresponding legal terms in the US are *informed consent* and *informed refusal.* The language used in medical care is more detailed:

- Full Disclosure is for alert, expectant patients, who may pay a psychological price for their knowledge, but will keep up an active dialog and maintain good access to the doctor.
- Nondisclosure: Total nondisclosure applies to the very frightened, dependent patient, or an incapacitated one; both will need various supports, management, and a proxy.
- Individualized disclosure is for patients who only want to know so much. They should be given the right to *informed refusal,* which should be documented, and they should be asked to name a surrogate or a health proxy. "Are you saying that you would rather that I did not discuss the details of your illness with you?"

Sharing the information: The physician should have a few selected points to make, because patients do not process much information when stunned by painful news. "I have given you a lot to process all at once. We can talk tomorrow when you have had a chance to think more about it." When continuing earlier discussions, the focus should be on adapting the goals of care to a changing situation. "We cannot cure the cancer, but we have some treatments that may slow its growth." It is helpful to include familiar staff members. The tone must be one of openness and realism but must always contain a message of hope, even when the situation is so dire that hope lies in the promise of continued caring, support, and symptomatic care. "It is true that the treatment is no longer working, but there are many other measures to discuss, and, no matter what you choose, we will work closely with you to keep you comfortable." Feedback must

always be obtained to ensure that the patient has understood and agrees. Patients should be encouraged to express their feelings and explore their level of understanding. “Before we end, I want to make sure that we all agree about the next steps.”

Responding to patient and family feelings: The physician should be prepared to witness outbursts of intense emotion. Unfortunately, these may be the consequences of previous rosy promises. Alternatively, patients may have been given realistic information, but subsequently comforted themselves with unrealistic hopes. The opposite is also heard: “You have taken away hope!” This can be minimized by an affirmative, engaged attitude about the many ways in which care will be continued to the end. Ambivalence is always present, and disbelief may predominate. It can be gently confronted, but there is no need to insist, because the conversation will be resumed many times in the days to follow, and the passage of time will effect a change. Physicians must acknowledge the emotions, validate their normalcy, and reflect on the patient’s and family’s observed strengths, while indicating an awareness of the difficulties that they are facing. “This is difficult news to hear, even with all of the support and comfort that I have seen you give each other.”

Planning and follow-up: These are the keys to a successful interchange. Patient and family must be left with a clear sense of what care is being provided and why. Further meetings must be planned; end-of-life decisions are psychologically and morally so stressful that repetition is essential. Their understanding and acceptance must be ensured, and multiple support sources must be explored. However, under so much emotional fear and pain, and in complex families, the responses may impair patient care and psychiatrists may be consulted.

The tenor of these discussions will vary with each physician, but its understanding and respectful tenor should always be there.

Consultation: Communication is always enhanced by consulting with other members of the medical team. It is to everyone’s benefit—whether senior attendings or young physicians—to maintain good relationships with seasoned nurses, whose keen understanding of the patient’s family situation helps greatly. Social workers are often sophisticated family therapists, a very important type of therapy not familiar enough to young physicians. Other resources available include psychologists, who are both therapists and have specialized skills, patient representatives, physical therapists, or any other staff playing a role in patient care.

Psychiatric Contributions

When asked for an inpatient consultation in a new ward, the psychiatric consultant must recognize that the ward is a complex interactive system which affects every patient case. Taking some time to learn the ropes while working on a case pays off for that day and on future consults. Staying in

touch with staff is essential.

The consultant should evaluate a responsive patient alone first, then with input from family members. Next, sharing one's findings with the physicians and involved nurses will probably include additional questions and information exchanges on both sides. The results should be a manageable plan. With very difficult cases, the extra time is well rewarded. The psychiatrists must reframe family dynamics in ways the staff finds more understandable, or, if not, more pointed and open to different management.

The consultant is obliged to avoid unnecessary details, even—or especially—if they might become gossip. However, an anecdote that brings shared pleasure into the picture, like the patient's tennis prize, should definitely be shared. Turning to staff members, if psychiatric consultants observe something like exaggerated negative judgments, they can interpret some of the patient and family dynamics in a more understanding way. If consultants identify particular individuals who have become too involved positively or negatively, they must note this and review it privately with the individual. Over time, trusted psychiatrists may be called on by staff members for personal reasons, and obtaining the best referral for them is an important contribution to their work.

Table 27.9–2 summarizes the content and illustrates the tone of physician communications that are most likely to allay the patient's and family's fear of abandonment and to enable them to best marshal their coping resources.

PHYSICAL AND PSYCHOLOGICAL SYMPTOMS IN PATIENTS WITH ADVANCED DISEASE

Turning now from the needs of the doctors to the needs of the patients, studies have shown that dying patients experience more symptoms than they can describe. In a cancer patient study, outpatients averaged more than 9 symptoms, and inpatients more than 13 symptoms, many of them unrecognized or undertreated. Unexpectedly, when psychological symptoms and existential concerns—such as nervousness, sadness, insomnia, fear of dependence, or loss of dignity—are included, they are significantly more prevalent than the physical ones. Many recent studies confirm that physical, psychological, neurological, hormonal, and existential concerns interact with each other in ways that are only beginning to be understood. What this relationship indicates is that all symptoms must be attended to. Sometimes this needs to be emphasized with physicians untrained in palliative care.

Table 27.9–2.
Seven Promises a Physician Should Make to a Dying Patient

You will have the best medical treatment, aiming to prevent exacerbation, to improve function and survival, and to ensure comfort.

You will never have to endure overwhelming pain, shortness of breath, or other symptoms.

Your care will be continuous, comprehensive, and coordinated.

You and your family will be prepared for everything that is likely to happen in the course of your illness.

Your wishes will be sought, respected, and followed whenever possible.

We will help you consider your personal and financial resources, and we will respect your choices about their use.

We will do all that we can to see that you and your family will have the opportunity to make the best of every day.

Adapted from Mitka M. Suggestions for help when the end is near. *JAMA*. 2000;284:2441.

Symptom Assessment

Any intervention must be examined in the light of the patient's total illness. How does it relate to general quality of life and the goals of care? For example, metastatic bone pain responds to radiation, but this should only be done if the benefit outweighs the discomfort of moving the patient and if the patient prefers it to intensified opiate treatment.

The use of validated instruments, such as the Memorial Pain Assessment Scale (MPAS), allows more reliable evaluations, making it harder to dismiss symptoms. It asks for the level of pain, from 0 to 10. Over time, patients serve as their own controls, because their scores are internally consistent. Pain evaluation has become a vital sign used in routine reassessments, insuring that the patient's voice is heard by overworked caregivers. Other scales that can measure distress and quality of life include the Hamilton Depression Rating Scale (HDRS), the Hamilton Anxiety Rating Scale (HADS), the Distress Thermometer, and the Quality of Life questionnaire (QOL). Internal consistency is true for them as well.

Assessment and Treatment of Pain

Incidence. Pain is common in dying patients but can be relieved or at least managed until the very end. In cancer, its overall incidence of 51 percent rises to 74 percent in advanced disease, of which 40 to 50 percent is moderate to severe, and 20 to 30 percent is said to be excruciating if untreated. Pain is also frequent in AIDS and advanced multiple sclerosis (MS) and can be expected in almost any advancing illness, unless patients slip into coma, as seen in hepatic failure or untreated hypercalcemia. Pain specialists are widely available. In the hands of a skilled physician, adequate relief should be obtainable in the majority of cases, and some alleviation obtainable in the rest, but there are often significant side effects. Many of them can be corrected or minimized, but patient and physician must decide together which symptoms or side effects are most acceptable. The psychiatrist can work with psychological symptoms now known to be intimately connected to the experience of pain.

Types of Patients. It is useful to recognize patients in several categories.

Acute pain patients respond well to treatment but can develop tolerance to any analgesic, requiring more medication if the pain lasts beyond a few days or weeks. This needs to be explored without prejudice and judgment in very ill or dying patients. Nor should patients who will live long enough to improve and continue to use the medications be viewed prematurely as *addicted* or *drug-seeking*. While the very success of recent treatments (bringing more cures or longer survival) has been accompanied by increased addiction and all its problems, this does not warrant decreasing pain treatment.

Chronic pain patients: The treating physician should avoid minimizing the severity of patients' pain over time or losing sight of the progression of disease, even if it is slow. The use of opiates in chronic pain is still controversial, especially in nonmalignant pain. However, regardless of controversy, there is no justification for withholding relief in a patient with a predictably limited future, even one that may extend over many months.

Chronic pain patients in palliative care are already dependent and may begin to use opiates as a release from existential suffering, as well as from physical pain. This should not be denied to a dying patient. However, if indicated, a psychiatric consultant can help explore and differentiate between these conditions, and can recommend specific interventions. Improved spiritual and psychological well-being occasionally result in less opiate use, but effective relief of distress should always be the guiding principle.

Palliative care patients with a substance abuse problem: If actively abusing or on methadone (Dolophine), these patients are already tolerant and must be given a correspondingly larger amount of medication for adequate pain control. Prejudices about such individuals must be put aside, and they should receive the same compassionate pain relief as others.

Multiple Pains. Patients with advanced disease commonly experience several different kinds of pain (Table 27.9–3). The distinctions are important because they call for different treatment strategies: somatic and visceral pains are more responsive to opiates, while neuropathic and sympathetically maintained pain often responds better to non-opioid adjuvant medications. If neuropathic pain is an integral part of the condition, it may respond poorly to treatment. Treatments to consider are glucocorticoids (dexamethasone) or specific anticonvulsants and antidepressants such as pregabalin (Lyrica) or gabapentin (Neurontin). A comorbid depression may be helped by analgesic antidepressants, for example, a tricyclic drug such as desipramine (Norpramin), or a serotonin-norepinephrine reuptake inhibitor such as duloxetine (Cymbalta). Therefore, the psychiatric consultant should seek out a sophisticated pain

history and ensure that all of the patient's different pains receive individual attention, until the last breath. Again, pain consultants should be consulted as needed.

Diagnosis and Prognostication. A careful evaluation may require an etiological work-up; certain pains can be eradicated even in terminally ill patients. Even so, one should always weigh the benefits of better pain relief against the pain and intrusiveness of diagnostic procedures. Clinical decisions are hindered by the inaccuracy of medical prognostication at the end of life. Studies show that the majority of DNR orders were written within less than a week of death. This over-optimism leads doctors to be more proactive with treatment and, often, less generous with opiates. It also worsens the family's ordeal and increases their later regrets, which can create lasting wounds.

Table 27.9–3.
Types of Pain

Nociceptive Pain	
Somatic pain	Usually, but not always, constant, aching, gnawing, and well localized: e.g., bone metastases
Visceral pain	Usually, but not always, constant, deep, squeezing, and poorly localized, with possible cutaneous referral: e.g., pleural effusion leading to (1) deep chest pain and (2) diaphragmatic irritation referred to shoulder
Neuropathic pain	Burning dysesthetic pain with shock-like paroxysms associated with direct damage to peripheral receptors, afferent fibers, or the central nervous system, leading to loss of central inhibitory modulation and spontaneous firing: e.g., phantom limb pain; can involve sympathetic somatic afferents.
Psychogenic pain	Variable characteristics, secondary to psychological factors in the absence of medical factors: vanishingly rare as a pure phenomenon in patients with advanced conditions, in which pain is often present, but often is an additional factor in the presence of organic pain.

Treatment of Pain. Aggressive multimodal treatment should be the norm. Drugs are the mainstay but must be part of a broader approach, including ancillary programs such as relaxation, guided imagery, massage, hypnosis, acupuncture, use of splints, and exercises. Music has always been soothing, to the point of helping regulate breathing in semicomatose patients.

PHARMACOLOGICAL TREATMENT OF PAIN. The World Health Organization's (WHO) three-step ladder for cancer pain control is a drug-treatment method known to be effective in many cultural and national settings. The use of drugs begins with nonnarcotic analgesics, moves to the opiates used for moderate pain, then to opiates used for severe pain, culminating in a continuous drip with adequate *rescue* doses for breakthrough pain. Many

opiates can be escalated without any upper limit, as long as side effects are well managed. Nonnarcotic analgesics and other adjuvant medications should be used as adjuncts to narcotics at any stage.

CLASSES OF PAIN MEDICATION

Nonnarcotic Analgesics. Nonnarcotic analgesics include nonsteroidal anti-inflammatory drugs (NSAIDs) and acetaminophen. The latter is a good analgesic for mild to moderate pain, with few side effects, except rare hepatotoxicity, especially in the presence of alcohol abuse or hepatic disease. All NSAIDs have a ceiling effect; hence, their use in fixed combination with a narcotic means that the narcotic cannot be used to maximum effect, unless more of it is added separately. NSAIDs are said to help bone pain, in particular, and are good anti-inflammatory agents, but side effects often limit their use in patients with advanced disease. Contraindications include renal disease, a history of peptic ulceration, hypertension, bleeding diathesis, and encephalopathy.

Narcotic Analgesics. Some narcotics are effective for moderate to moderately severe pain and should be used until they are no longer effective. Opioids include morphine (Duramorph), fentanyl (Sublimaze), oxycodone (OxyContin), hydromorphone (Dilaudid), oxymorphone (Opana), levorphanol (Levo-Dromoran), and methadone (Dolophine). Their main side effects are sedation, dizziness, and nausea, which can be dose-limiting, but confusion, delirium, and hallucinosis are also seen. The narcotics that work for severe to very severe pain are divided into pure agonists and mixed agonist-antagonists. Pure agonists have no ceiling effect and can be increased indefinitely, as long as side effects are controlled. Somnolence responds to stimulating agents such as dextroamphetamine (Dexedrine), methylphenidate (Ritalin), and amphetamine and dextroamphetamine (Adderall) in both immediate and delayed release preparations. Modafinil (Provigil), a nonstimulant-activating drug developed for use in narcolepsy, is effective in medically ill patients. Confusion, delirium, and organic hallucinosis respond to antipsychotic agents, and nausea can be helped by a variety of antiemetics (e.g., prochlorperazine [Compazine]). Mixed agonists-antagonists should be left in the hands of skilled practitioners because they have a ceiling effect, can reverse positive effects, or can precipitate an abstinence syndrome in patients already dependent on pure agonists. They also have a higher incidence of psychotomimetic effects, such as confusion and hallucinations. A frequent mechanism of psychotoxicity is the accumulation of drug or metabolites whose duration of analgesia is shorter than their plasma half-life, notably morphine (Duramorph), levorphanol (Levo-Dromoran), and methadone. Using drugs such as hydromorphone (Dilaudid) with half-lives closer to their analgesic duration can relieve the problem without loss of pain control. Cross-

tolerance is incomplete between opiates, so patients should be tried on several drugs and the dosage should be lowered when switching drugs. Knowing equianalgesic doses of drugs and different routes of administration (Table 27.9–4) is important so as to avoid accidental undermedication or overmedication. For example, to change a patient from intramuscular (IM) to oral morphine, one should multiply the IM dose by two to six times (depending on the chronicity of usage) to avoid causing the patient pain and provoking what others will call drug-seeking behavior. Dying patients who must choose between pain and alertness face a poignant dilemma, but somnolence can be countered by several safe psychostimulants and can be used in ways that meet shifting needs.

Table 27.9–4.
Routes of Opioid Administration

Route	Comment
Oral	Preferred in cancer pain management
Sublingual	Buprenorphine (Buprenex) is effective but not available in the United States. Efficacy of highly lipid soluble drugs, such as fentanyl (Actiq), is likely, but there are no studies and little clinical experience.
	Efficacy of morphine (Duramorph) is controversial.
Rectal	Available for morphine, oxymorphone (Numorphan), and hydromorphone (Dilaudid). Customarily used as if dose were equianalgesic to oral dose. Absorption is variable, however, and relative potency may be higher than expected.
Transdermal	Available for fentanyl. Dosing interval is 2–3 days. Empirical indications include problems with oral drug, desire to offer trial of fentanyl, compliance problems with oral dosing, or possibility of improved quality of life if oral therapy is avoided.
Oral transmucosal	Formulation using fentanyl is currently undergoing trials for breakthrough pain.
Subcutaneous	
Repetitive bolus	Ambulatory infusion pumps can provide continuous infusion with any parenteral opioid formulation. More advanced pumps can also provide patient-controlled analgesia. Clearest indication is inability to tolerate oral route.
Continuous infusion	
Continuous infusion with patient-controlled analgesia	
Intravenous	
Repetitive bolus	Continuous infusion possible if permanent venous access device is available.
Continuous infusion	
Continuous infusion with patient-controlled analgesia	
Epidural	

Repetitive bolus	Clearest indication is pain below arms and dose-limiting side effects from systemic opioid. Often coadministered with local anesthetic.
Continuous infusion using percutaneous or implanted system	
Intrathecal	Usually administered via a totally implanted infusion pump. May be cost-effective for those patients with clear indication for intraspinal therapy and long life expectancy.

Adapted from Portenoy RK, Osta BE, Bruera E. Physical symptom management in the terminally ill. In: Chochinov HM, Breitbart W, eds. *Handbook of Psychiatry in Palliative Medicine*. 2nd ed. New York: Oxford University Press; 2009.

Dosing Instructions. The need for regular *by-the-clock* administration, as opposed to *as-needed* (PRN) administration, cannot be overemphasized. It is counterintuitive, but a maintenance schedule achieves faster and better pain control with a smaller total dose, because it is an early and often pre-emptive response. PRN orders only provide a slightly or very delayed response. The patient may wait until pain is significant before calling for medication and then wait again for it to be given, during which time increasing pain is worsened by anxiety and growing resentment. The patient starts to ring repeatedly, often annoying overburdened nurses. The patient has been set up! To some extent, so have the nurses.

Rescue doses should be directly available to the patient for breakthrough pain, and their repeated use should signal the need to raise the maintenance dose. Any patient on breakthrough doses should always be sent home or to a hospice with the same prescriptions, since new doctors may not realize the patient's condition.

Adjuvant Drugs. Adjuvant drugs should be used simultaneously throughout. Many of them are psychotropics well known to the psychiatrist, but their analgesic effect has often been shown to be separate from their primary psychotropic effect. They are particularly important in neuropathic and sympathetically maintained pain, where they can be the mainstay of treatment, but they also introduce new side effects that need to be recognized.

Surgical and Anesthetic Modalities. Other modalities include palliative radiation and anesthetic and neurosurgical techniques that can all give regional relief where other methods have failed. Their intrusiveness is warranted only if they are painless once installed and can provide significant relief, as they often do. Other methods should be considered in imminently dying patients.

Coexistence of Pain and Psychiatric Symptoms. Patients with pain have a significantly higher incidence of depression and anxiety. It had been

suggested that: (1) chronic pain may be a depressive equivalent, (2) pain may cause psychiatric syndromes, and (3) pain may coexist with psychopathology in vulnerable subjects. Given the steady stream of interactions now recognized as integral to the human organism, it can be said that *all* suffering patients are vulnerable and treatment should proceed from *any* angle. Psychiatric interventions along with primary medical ones, selected integrative medicine options, and thoughtful personal care all play a role. The result will be more than the sum of its parts.

Furthermore, in the face of advanced disease, the role of personality factors and pre-existing diagnoses becomes less and less significant. Psychiatrists mention them in their notes, but must also comment that *they are no longer in play*. This is important in preventing the staff from psychologizing the patient's delirium and related behaviors.

However, there are some patients who manifest clear psychopathology, are more difficult to evaluate, and whose complaints are more easily disqualified. Hence, the consultant must remain attentive and help the staff treat their pain as aggressively as that of other patients. The psychiatric diagnosis and its impact on illness must be explained and treatment recommendations made, while generating as much genuine sympathy for the patient as possible.

Undertreatment of Pain. The incidence of psychiatric complications is particularly high when pain is undertreated by the staff, which happens for many reasons. The daily use of the pain scale was driven by studies showing that when doctors and patients were separately asked to rate pain, their responses matched well for mild and moderate pain, but the higher the patient's pain rating went, the slower the physician raised his/hers. Inadequate management and unawareness of resources are both possible reasons, but often it appears to be an unconscious avoidance of the treater's own helplessness in the face of the patient's complaints. Nurses and aides might also share it, especially if the patient is genuinely unlikable, but they may tend to recognize it sooner because of their more frequent and intimate interactions.

However, it can also be patients and/or their families who have false beliefs about medications, such as destructive effects of drugs or side effects or addiction. These misconceptions are often a metaphor for the fear of death. In its most severe form, a dependent spouse whose mate is suffering but not yet moribund, or the mother of a young child who is still conscious but exhausted, cannot tolerate the loss of reactivity that comes with pain relief. They cannot allow the patient much needed relaxation or sleep, as they experience this separation as the beginning of their final loss. Psychiatrists can be called in, but familiar, trusted physicians can often handle similar situations.

Medical Staff Vulnerabilities. Medical staff barriers to pain treatment can be better managed if better understood. At an unconscious level, doctors and nurses, living through many patient experiences, develop an incompletely acknowledged, internal standard of how much pain is *expectable* in a given situation. The standard is personal, not universal, but tends to lean toward the so-called "good patient," that is, the brave and uncomplaining one. Each patient's complaints are unconsciously measured against both other patients' behaviors and the caregiver's own standard; many come up wanting. Nurses are more likely to make a correction for each other, but young, untested doctors who still feel immortal may have more difficulty. They will not have the same understanding as a mature physician who has had a hard turn at being a patient, or a senior physician who just lost a spouse after a long illness. In the final analysis, the tendency to underestimate pain is intimately related to every physician's personal history and stage of life, as well as their deepest motivation and coping mechanisms in doing this difficult work. They are capable of learning, but attention should be paid to it. They must also protect themselves against too much empathic identification, as well as guilt. An occasional lapse into ignoring or blaming the victim is an all-but-irresistible response to so many difficult situations, and another good reason for close teamwork and the support and insight that come with it.

Psychiatric Interventions. Psychiatrists can help optimize pain management by reorienting staff responses in a supportive, nonjudgmental way, emphasizing the need to respect the patient's statements, and reaffirming the now well-accepted importance of daily review of patients' pain scales. Psychiatrists can give the team insight into the patient's psychological state, emphasizing the patient's problems in ways that undermine a team's unconscious pain standard. They can sympathize with the team's difficulties and uncertainties, normalize their feelings, ensure adequate pain management, and re-educate them about the low addiction risk and its now irrelevant role in dying patients.

Psychiatrists' role in bringing about optimal pain relief is a powerful intervention in itself, but it is not complete unless integrated into a multipronged program that includes supportive psychotherapy and some of the many focused therapies that deal with terminal care (Table 27.9–5). Patients using behavioral interventions report greater sense of control and decreased anxiety and pain. In one study, doctors underestimated the value of these methods compared to the patient's own estimate. Hence, it is useful for psychiatrists to recommend them. Psychotherapy must help the patient differentiate between pain and suffering (which includes psychological symptoms but also prominent religious and existential concerns!) and must facilitate appropriate responses to each. Most hospitals have chaplains, but a trusted religious familiar is also desirable.

Table 27.9–5.
Therapy in the Terminally Ill Patient

Psychotherapy (individual, family, or group)
Crisis intervention
Family-focused grief therapy
Group supportive psychotherapy
Meaning-centered psychotherapy
Dignity therapy
Cognitive–behavioral therapy
Behavioral therapy
Supportive–expressive therapy
Provision of information
Life review
Facilitating communication
Addressing unfinished business
Leaving a legacy
Provision of information
Reframing
Focusing
Problem solving
Relaxation
Meditation
Mindfulness-based stress reduction
Hypnosis
Desensitization
Rehearsal
Music therapy

Fatigue

Fatigue is one of the most common symptoms reported by patients, and its effect on quality of life is profound. In one study, 62 percent of cancer patients rated it as worse than pain, 12 percent said that it made them want to die, and only 27 percent reported that their doctor had suggested a treatment. Yet, fatigue is hard to define, even harder to measure, and often goes unnoticed. It is a multidimensional concept with physical, cognitive, affective, and motivational components that differentiate it from symptoms such as weakness, somnolence, and tiredness after exertion. It is integral to acquired immunodeficiency syndrome (AIDS), cardiac, respiratory, hepatic, renal, and neurological end-stage disease.

If internists accept fatigue as an inevitable correlate of major illness, psychiatrists may over-psychologize it. Fatigue is, after all, a defining criterion of the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* diagnosis of depressive and anxiety disorders. Psychologically driven fatigue may be incapacitating and is known to amplify somatic symptoms. Furthermore, the incidence of psychological disorders in very ill patients is high. However, as mentioned before, these distinctions are no longer clear; hence, teasing out psychologically from physically determined fatigue may not be possible because the two coexist

with such regularity.

Table 27.9–6 outlines some causes and interventions for end-of-life fatigue. It should be evaluated in multiple dimensions before deciding on treatment. It can be cruel to require a dying patient to move, but this can improve quality of life in a patient who has more time left. Both patients benefit from being taught energy-conserving techniques. Psychostimulants used purely for fatigue have been observed to be helpful and can interrupt a vicious cycle of fatigue and despair, although different patients will have different risks from side effects.

Table 27.9–6.
Possible Etiologies of Fatigue in the Medically Ill

I. Medical/physical conditions
   - Associated with the underlying disease itself
   - Associated with treatment for the disease
   - Chemotherapy
   - Radiation
   - Surgery
   - Biological response modifiers
   - Associated with intercurrent systemic disorders
   - Anemia
   - Infection
   - Pulmonary disorders
   - Hepatic failures
   - Renal insufficiency
   - Malnutrition
   - Neuromuscular disorders
   - Hypothyroidism
   - Hypogonadism
   - Hypoadrenalism
   - Associated with sleep disorders
   - Associated with immobility and lack of exercise
   - Associated with chronic pain
   - Associated with use of centrally acting drugs (e.g., opioids)

II. Psychosocial factors
   - Associated with anxiety disorders
   - Associated with depressive disorders
   - Stress related
   - Related to environmental reinforcers

Adapted from Portenoy RK. Physical symptom management in the terminally ill: an overview for mental health professionals. In: Chochinov HM, Breitbart W, eds. *Handbook of Psychiatry in Palliative Medicine*. 2nd ed. New York: Oxford University Press; 2009.

Nonstimulant-activating drugs, such as modafinil (Provigil) and dextroamphetamine/amphetamine (Adderall XR), are increasingly used in ill patients, although they are not, at this time, FDA-approved indications.

Insomnia

Sleep hygiene is important to the healthy and the sick, and yet hospitals can be blatant offenders in their lack of concern for patients' sleep. Dying patients may be desperate for the blessed release of "getting away" for a few hours. Of course, very ill patients may require care and procedures at any time of day, but such simple measures as minimizing noise and maintaining optimum light are neglected in environments that are otherwise attentive to patient needs. Patients and their caregivers need education about maintaining as close to a normal circadian rhythm as possible and preceding sleep time with a quiet routine. Long-acting pain medication is critical to prevent middle-of-the-night breakthroughs, and many patients need help in understanding the use of sustained-release preparations, such as morphine (MS Contin) or oxycodone (OxyContin). Despite recent publicity about abuse, these long-acting pain medications are not as readily abused by very ill patients as by patients who may get better for some time. They have an important role to play in late-stage patients and should not be avoided because they are too long-lasting. Sedatives are equally important. Zolpidem (Ambien) and zaleplon (Sonata) are well tolerated, and the latter is short-acting enough that it can be taken in the middle of the night. Benzodiazepines have an antianxiety effect as well, which often makes them the drugs of choice in nondelirious patients, but are contraindicated in delirious ones. Several antidepressants are so predictably sedating that they are often used as hypnotics even in the absence of a diagnosis of depression. They include trazodone (Desyrel), nefazodone (Serzone), mirtazapine (Remeron), amitriptyline (Elavil), imipramine (Tofranil), and other tricyclics. Antipsychotics, such as olanzapine (Zyprexa), quetiapine (Seroquel), and risperidone (Risperdal), can be used the same way. The side effects of all the psychotropics should be evaluated, but the habituation that accompanies the use of many hypnotics generally does not occur.

Terminal Anorexia and Cachexia

Anorexia and the wasting state known as cachexia are common near the end of life, but, in some conditions, such as cancer and AIDS, they appear well before the end of life, causing great distress to family and friends. If they have not been educated about its involuntary nature, unprepared family members become angry that medical staff do not share their intense concern. They may try to relieve their sense of helplessness by fixating on making the patient "try a little harder" to eat. This generates guilt and anguish in the patient at a time when peaceful human contacts are the most precious. Families have been known to request a psychiatry consultation for an eating disorder, wanting to believe, against all evidence, that better nutrition would change the inevitable. This example of making patients responsible for their illness is one of the most poignant and frequent demonstrations of love and desperation gone astray. Families, and indeed

the public as a whole, need to understand the involuntary and overwhelming nature of terminal anorexia and, even more importantly, that the patient is not experiencing hunger and will not benefit from forced nutrition.

In fact, the opposite is often true. When anorexia is part of the active dying process, total parenteral nutrition (TPN) has been shown to worsen the patient's quality of life. However, in cases in which the patient can still experience hunger—for example, in ovarian cancer in which a nonfunctional GI tract may exist in the presence of functional liver and kidneys—TPN is available and has been known to sustain patients for long periods. (Its wisdom can be disputed, but that is a different issue.) Thirst, another family concern, is more related to dry mucous membranes and can be relieved with frequent small sips of water. Intravenous (IV) hydration and transfusions are counterproductive in the face of failing heart or kidneys. Table 27.9–7 outlines some recognizable causes and the treatment of anorexia for patients who are not yet actively dying. The treatment plan should address multiple etiologies.

Other GI Symptoms

Nausea and Vomiting. Nausea is near the top of lists of unpleasant symptoms, and unfortunately is one of the most common. Many patients say, "I would be okay if it were not for the nausea!" The act of vomiting itself, although it gives temporary relief from nausea, is unpleasant and can be painful in certain conditions. It can be difficult to eradicate even when its etiology is known and all the more so when one cannot pinpoint the cause—a not infrequent event.

Available antiemetic medications are helpful in controlling more vomiting and nausea. A careful assessment may help to establish a better approach and more effective treatment for both. Asking the patient specific questions, such as onset, frequency, time of the day, quality of content, and its association with other symptoms, will help the clinician to choose a more effective medication. For instance, escalation of pain medication could be associated with nausea as well as medical interventions, procedures, and certain treatments such as chemotherapy or radiation. Common causes of nausea and vomiting in the medically ill are related to opioids, NSAIDs and centrally acting drugs. Intra-abdominal neoplasm and antineoplastic medications are common causes of nausea and vomiting as are intracranial neoplasm and leptomeningeal disease. Other causes include psychological processes and metabolic disturbances such as hepatic failure or renal failure and electrolytes abnormalities.

Table 27.9–7.
Causes and Treatment of Anorexia

Causes
Disease related
Nausea, fever, anemia, weakness, dehydration, pain
Sepsis, other acute malaise, electrolytes abnormalities
Weakness or damage in muscles of deglutition
Altered cognitive states
Altered metabolism: uremia, hepatic failure
Abnormal response to nutritional intake
Treatment related
Chemotherapy or other drug effect (nausea, diarrhea, altered taste or smell)
Surgery
Radiation therapy effects
Dialysis side effects
Transplant rejection, graft-versus-host disease
Cancer-related symptoms
Psychological
Anxiety syndromes, depression, delirium, specific phobias
Pre-existing or relapsed psychiatric disorders
Nonpharmacological intervention
Education
Dietary management
Discuss etiology (often multifocal) and characteristics (involuntary) with patient and family
Reassurance
Increase physical activity and psychosocial intervention
Alternative therapies (massage, muscle relaxation, mindfulness-based therapy, yoga)
Pharmacological intervention
Megestrol acetate (Megace)
Steroids
Appetite-inducing psychotropics (olanzapine, quetiapine)
Psychostimulants (methylphenidate, dextroamphetamine)
Wakefulness-promoting agents (modafinil)
Selective serotonin reuptake inhibitors (fluoxetine, paroxetine, citalopram, escitalopram, sertraline)
Selective norepinephrine reuptake inhibitors (venlafaxine and duloxetine)
Norepinephrine-dopamine reuptake inhibitor (bupropion)
Alpha-2 antagonist/5HT2/5HT3 antagonist (mirtazapine)

Diarrhea and Constipation. Diarrhea is common in AIDS and in certain cancers. It is usually associated with local pain, burning, and often abdominal cramping. It also leads to fecal incontinence, one of the most humiliating and unpleasant symptoms a patient must endure. Rapid cleaning and good skin care by a relaxed, well-related person is important to shorten the experience and minimize pain, skin irritation, and breakdown. Drugs that slow down intestinal motility can be helpful, so the anticholinergic side effects of many psychotropics are welcome. However, it can be a very resistant symptom when caused by opportunistic organisms and/or *Clostridium difficil*e.

Constipation is also common and can cause significant abdominal cramping, bloating, distention, nausea, and vomiting. Ideally, it should be prevented or aborted by adequate hydration, optimal mobilization when possible, and, most of all, early institution of an aggressive and consistent

bowel regimen. The use of pain medications could contribute to constipation if not given with stool softeners. Enemas or disimpaction may be instituted if the patient desires. However, in someone close to death, increased opiates may be preferable, especially as normal bowel function is frequently absent during the terminal phase.

Respiratory Symptoms

Respiratory distress is one of the most difficult symptoms to endure—and to watch. Although measures have been developed to assess dyspnea, assessment of the ill patient should be based on self-report, because objective measures, such as oxygen saturation and arterial blood gas measurements do not often correlate with subjective experience. It is easy to treat because opioids and benzodiazepines decrease the anxiety and the urge to breathe within minutes, allowing the patient to rest and to relax. However, in the actively dying patient, a significant decrease in the effort to breathe is associated with increased hypoxia and eventual carbon dioxide (CO_2) narcosis. Parenthetically, this is one of the most painless ways to die, and is the subject of debates about the "double effect," in which an action has an intended (patient comfort) and an unintended effect (hastening of death). If dose escalation is too sudden, it may hasten the moment of death, but when the drugs are prescribed appropriately, respiratory depression is gradual; death still comes gradually but more comfortably.

As with pain patients, the dyspneic patient is anxious and hyper-alert and may have already been exposed to opiates or benzodiazepines (the latter is now contraindicated). Therefore, the gradual escalation of drugs to give consistent relief of dyspnea does not result in sudden death. In fact, family members who have waited at a dying patient's side comment on "how long it takes for a heart to give up." One of the more common signals that the end is near is agonal breathing, where one hears only acute gasping breaths that become less and less frequent. If family members are keeping a long watch, it is useful to check on them, and explain what is and will be happening. The dying patient can still display involuntary movements that are experienced as a communication by a watching grieving spouse or child. These last moments should be managed with kindness and delicacy because they will never be forgotten.

Distress due to persistent coughing is exhausting and interferes with communication. Opioids turn off the cough reflex promptly, and, if the patient is strong enough to cough repeatedly, opioids can and should be used without risk of respiratory depression.

Sedation of Patients. Neuromuscular blocking agents are used routinely in intubated patients, to prevent them from fighting the respirator. Agitation can precipitate accidental extubation, leaving a conscious patient without any ability to breathe, a terrible experience. For

that reason, patients in critical care units are often kept sedated while on the ventilator. When the patient is receiving only comfort measures, such a condition must be re-evaluated, but, whatever the case, *the planned withdrawal of life supports must always include adequate pre-sedation*.

This needs affirmation because there are now more critical care unit (CCU) patients stabilized on respirators who must be transferred to intermediate units, to allow the treatment of newly acute patients. This is where final controversies may play out. There have been times when end-of-life management occupied the time of politicians whose duties lay elsewhere, or when individual caregivers did not yet understand that the rule of care took priority over their personal convictions while it was being argued in appropriate venues. All of these scenarios with patients in pain and respiratory distress involve the double effect. This partly describes every drug ever given to any patient, but if the side effect is hastened death, the controversy should be brought up and carefully reviewed before proceeding. This will be discussed further in the section on ethical issues.

PRIMARY PSYCHOLOGICAL SYNDROMES AND SYMPTOMS IN TERMINALLY ILL PATIENTS

Psychological symptoms are nearly universal at the end of life. Psychiatric syndromes occur with an increased, definable frequency and may have a different age and gender distribution. For example, anxiety and depression are as common in men as in women. The DSM-5 classification remains an important framework on which to base clinical observations, but it was not designed with dying patients in mind. Hence, for terminally ill patients, it is pragmatically useful to focus on three major conditions: anxiety, depression, and confusion—which, in fact, frequently coexist and overlap. Rarely, specific phobias of needles, enclosures, and the like may interfere with comfort into the terminal period, and should be addressed to ease the patient's journey. The consultant should remain alert to any precipitant factors that can cause new discomforts and work to defuse them. Major psychotic disorders may become submerged under the increasing symptomatology of the active dying process. However, they require specific attention when the patient is not actively dying and when the psychotic symptoms are clearly separate from, and superimposed on, the symptoms of either the disease or the treatment.

Prevalence

Much of the research on prevalence has been done in cancer and AIDS and shows a marked increase of psychiatric conditions near the end of life. Severe depressive symptoms rose from 25 to 77 percent in a sample of hospitalized cancer patients, although stricter criteria find 15 percent with major depression and another 15 percent having severe depressive

symptoms. The prevalence of delirium rises from a range from 25 to 40 percent to as high as 85 percent with increasingly advanced disease. The association of psychiatric symptoms with pain was demonstrated in one of several consistent studies in which 39 percent of patients with a psychiatric diagnosis reported significant pain, compared to 19 percent for those without a psychiatric diagnosis. In a sample of AIDS inpatients, 65 percent had an organic mental disorder, and 27 percent had major depression. The financial cost of psychiatric disorders can be inferred from a study in which patients with a psychiatric diagnosis remained in the hospital 60 days longer than those without a psychiatric diagnosis.

General Treatment Principles

Improved quality of life is the *immediate* goal, so all available modalities that can bring symptom relief should be instituted rapidly while a plan for psychological and family interventions is designed and set into motion. Since psychiatric syndromes are intertwined with medical conditions, understanding the medical condition is necessary. Good communication with the treating physician yields a fuller analysis and better coordination with medical treatment goals.

Most commonly, terminally ill patients demonstrate an intertwining of anxiety and depression. These are difficult to tease apart, and the term *negative affect* has been suggested to define them as a symptom complex. It is hard not to feel grief at so much loss and fear about the dark unknown. Individuals whose deep faith in an afterlife that animates their spirit are an exception, but, even among them, many describe coexisting regret at the loss of their temporal life with its rich palette of life events.

However, current symptoms may be an exacerbation of a pre-existing disorder, in which case, symptom control still remains the primary response. Other than medical complications, a personal or family history is the best predictor of psychiatric problems in advanced disease.

Death Trajectories

There are many different death trajectories with different conditions, of which a few will be briefly mentioned. Death from pulmonary insufficiency is slow, and is terrible for both the patient and the family. Mild sedation eases the pain and may let the patient rest while slightly shortening the time to death (the so-called double effect). Psychiatrists may be needed when there is debate among participants, as might the Ethics Committee.

Neurological disorders are multiple. Many cause a slow decline until, far into the disease, only shreds of the patient's sense of self remain intact. In such cases, both patient and family caregivers should be offered early support, the patient some tranquilization, and the family a chance to speak about their reactions. Under such conditions, there may be no previous

psychopathology, but the late events can create it, as happens in accidents and disasters.

Every cancer is an unpredictable disease and can have a long and complex course, hence an unpredictable mode of death. It is also characterized by presenting several possible options, not only throughout the course, but often during the final months. Families live through this unpredictability at a cost. Complications and disagreements cause stress for everyone, including medical staff. In addition to the usual chart review and examinations, psychiatric consultants invited at this time must also learn from the staff the "story" of the course of treatment. This often opens routes for improvement.

Anxiety in Patients with Advanced Disease

Characteristics. Anxiety can be the presenting symptom for almost all medical disorders and is often a side effect from medications. However, in patients with advanced disease, it usually presents somatically, such as restlessness, hyperactivity, tachycardia, GI distress, nausea, insomnia, shortness of breath, numbness, or tremor. It lowers the threshold for pain, worsens functional impairment, and increases distress. Anxiety often interferes with the patient's ability to cooperate with other treatments or to relate optimally with loved ones. Patients refer to fear, apprehension, or ruminations more often than to anxiety. It is also associated with deep existential concerns, as outlined in Table 27.9–8. Individuals with pre-existing anxiety disorders commonly have an exacerbation.

Mr. S, a 50-year-old physical therapist with newly diagnosed unresectable lung cancer, had panic symptoms when his wife left his bedside to attend to other needs. He hyperventilated, felt short of breath, became restless, was unable to concentrate, and felt overwhelmed with morbid ruminations about his future. His symptoms were exacerbated by his discomfort about his dependence on her. Mr. S. was taught special breathing and relaxation exercises, and given clonazepam (Klonopin), which brought about a marked resolution of his anxiety. He felt calmer, more resilient, and able to handle periods of solitude with less difficulty.

Table 27.9–8.
Causes and Treatment of Anxiety

Causes	Treatment
Medical causes (psychiatric diagnosis: anxiety secondary to medical condition)	**Search for an etiology**
	If a medical etiology is present, treat it or minimize it.
	Use a multimodal approach.
	Should always include provision of immediate relief along with other interventions.
Untreated delirium	**Anxiolytic drugs**
Central nervous system pathology	Benzodiazepines
Hormone-secreting tumors	Lorazepam (Ativan), oxazepam (Serax), and temazepam (Restoril) are metabolized through conjugation and hence are more usable in hepatic insufficiency.
Metabolic or electrolyte abnormalities	

Pulmonary embolus	Alprazolam (Xanax), clonazepam (Klonopin), diazepam (Valium), and chlordiazepoxide (Librium) are benzodiazepines which are metabolized through oxidative pathways and can be used with renal impairment.
Respiratory insufficiency, hypoxia, dyspnea	Nonbenzodiazepines (beware of drug interactions)
Sepsis	Selective serotonin reuptake inhibitors
Poor pain control	Escitalopram (Lexapro)
Side effect of bronchodilators, steroids, β-adrenergic drugs, and other medications	Serotonin and norepinephrine reuptake inhibitor
Substance withdrawal	Venlafaxine (Effexor) and duloxetine (Cymbalta)
Psychiatric disorders	Neuroleptics (alone or with a benzodiazepine)
Acute stress response	Haloperidol (Haldol), olanzapine (Zyprexa), quetiapine (Seroquel)
Adjustment disorder with anxiety	Tricyclics
Adjustment disorder with mixed anxiety and depressive mood	Hydroxyzine (Atarax)
Panic disorder[a]	Buspirone (BuSpar)
Phobic disorder[a]	**Nonpharmacological interventions**
Generalized anxiety syndrome[a]	Behavioral methods
Depressive disorders[a]	Hypnosis, meditation, deep relaxation, with or without guided imagery (selected for induction of peacefulness)
Existential causes	Supportive psychotherapy
Awareness of oncoming death	Explores patient fears, reduces cognitive uncertainty, and reassures about accompaniment
Fear of death itself	Family sessions
Fear of the dying process	To improve communications, to reduce tensions, and to say goodbye
Mourning of losses	Spiritual counseling
Interpersonal regrets	
Continuing conflicts	
Guilt over abandonment	

[a]Often exacerbations of pre-existing symptoms.

Treatment. Anxiety is so common that it is sometimes discounted, but an etiology should be sought, and treatment should be instituted promptly. Anxious patients are more likely to be aroused enough for supportive therapy, cognitive–behavioral techniques, and complementary modalities, such as deep relaxation, meditation, and adapted yoga. If a patient brings up an issue that had never been resolved, discussing it is not only appropriate, but a major gift, contrary to some therapists' feeling that "the work won't go anywhere." Patients' inner lives are their last personal possession; sharing and enriching them is a powerful intervention, one of the last times they will reflect upon themselves in a meaningful dialogue with another human being. It is perhaps the *most* meaningful of all.

MEDICATIONS. Several classes of effective antianxiety medications can and should be used even in the presence of hepatic or renal impairment, given careful drug selection, dosage adaptation, and ongoing monitoring. Benzodiazepines add to the depressant effects of other sedating drugs and

can worsen the symptoms of confusional states, making them a poor long-term choice for the elderly. However, if the patient has been taking them for a long time with good effect, continuing them is desirable because their action is immediate and, with careful dosage, generally beneficial. They must be chosen with the patient's hepatic function in mind. All are metabolized by the liver, except for lorazepam (Ativan), oxazepam (Serax), and temazepam (Restoril), which are metabolized by conjugation and, hence, less dependent on liver function. Lorazepam and diazepam (Valium) can be given parenterally and sublingually. The first is short-acting and is preferred by patients who want more control over their ability to interact. The second is long-acting and tends to accumulate unless monitored carefully, but it is effective and has the added benefit of being an excellent muscle relaxant, which, for some patients, gives significant pain relief in itself. Clonazepam lasts 12 hours and does not accumulate. Habituation can require dosage escalation of all benzodiazepines, but this is acceptable if the drug improves the quality of these patients' limited life.

Depression in Patients with Advanced Disease

Prevalence. Depressive symptoms are also common in advanced disease and associated with the same existential factors as anxiety. Studies have found a prevalence ranging from 9 to 58 percent, depending on criteria strictness. Risk factors again include previous episodes or a family history. Table 27.9–9 outlines the common causes and treatments for major depression.

Diagnosis. Differentiating somatic effects of disease from the neurovegetative criteria of major depression is especially daunting in the terminally ill. Since the DSM-5 criteria (developed for primary psychiatric patients) include somatic symptoms that may be due to their medical disease, a *substitutive approach* has been developed, which replaces physical symptoms with nonsomatic symptoms. Table 27.9–10 describes the Endicott Substitution Criteria, which is the substitutive approach in primary use, because it has been found to perform as well as the DSM-IV-TR criteria and the stricter Research Diagnostic Criteria. It also highlights the important clinical observation that classically described depressive thoughts and feelings are not universal in terminally ill patients and, when present, reflect depression just as they do in physically healthy patients.

Treatment. Current studies and a large body of clinical experience with medically ill patients have shown that the pharmacological treatment of depression can be useful through the last days of life, even in the face of definable medical causes.

Mrs. V, a diabetic patient with end-stage renal disease, was on dialysis for 2 years, and was

diagnosed with depression and severe insomnia. Started on a daily oral dose of 15 mg mirtazapine, she immediately slept better and felt less depressed within 3 weeks. Four weeks later, she was admitted to the hospital in congestive heart failure, for what would be her terminal admission. She did not wish to give up the antidepressant but now felt oversedated and wished to be more alert during the day to interact with her family. Twice-daily 5-mg doses of methylphenidate made her more engaged and communicative. After her death, her family expressed deep gratitude for feeling so connected with her until the end.

Table 27.9–9.
Causes and Treatment of Depression in Terminally Ill Patients

Causes	Treatment
Medical causes: resulting in mood disorder secondary to general medical condition	Search for an etiology; if a medical etiology is present, treat it, or minimize it
Medications	Psychopharmacological management
Steroids	
Interferon and interleukin-2 (Proleukin)	Choose an antidepressant based on the patient's review of symptoms
Methyldopa (Aldomet)	
Reserpine (Serpalan)	Tricyclics antidepressants in depressed patients with pain and insomnia
Barbiturates	Monoamine oxidase inhibitors
Propranolol (Inderal)	
Some antibiotics (amphotericin B [Amphocin])	Mirtazapine (Remeron) in depressed patients with insomnia and anorexia
Some chemotherapeutic agents	
Disease conditions	Bupropion (Wellbutrin) in sedated depressed patients
Metabolic and electrolyte abnormalities	
Endocrine abnormalities	Psychostimulants effective in treatment-resistant depression and sedated patients
Hormone-secreting tumors	
Nutritional deficiencies	Selective serotonin reuptake inhibitors in depressed patients with cardiac risk factors
Anemia	
Uncontrolled pain	Serotonin norepinephrine reuptake inhibitors
Psychological causes	
Adjustment disorder with depressive mood	Mood stabilizers
	Nonpharmacological interventions
Adjustment disorder with mixed anxiety and depressive mood	Individual
Comorbid with various anxiety disorders	Supportive psychotherapy
	Cognitive-behavioral therapy
Depressive disorder not otherwise specified	Family therapy
Existential causes (all held with morbid intensity impervious to reassurance)	Group therapy
Awareness of oncoming death	Spiritual counseling
Mourning of losses	
Interpersonal regrets	
Continuing conflicts	

Guilt about burdening family
Guilt about betraying loved ones
Decreased self-worth
Self-criticism about perceived failed relationships and life achievements

Table 27.9–10.
Endicott Substitution Criteria for Depression

Physical-Somatic Symptoms	Psychological Symptom Substitute
Change in appetite or weight, or both	Tearfulness, depressed appearance
Sleep disturbance	Social withdrawal, decreased talkativeness
Fatigue, loss of energy	Brooding, self-pity, pessimism
Memory and concentration deficits, indecisiveness	Lack of reactivity

Antidepressants are the first line of treatment in patients who can be expected to survive long enough for the delayed onset to take effect. If there is no time or if increased energy and ability to interact with loved ones are urgently desired, patients can be treated with psychostimulants. The side effects of antidepressants can be very variable, including sedation and excess energy. For this reason, antidepressants should be used empirically, and their side effects monitored carefully.

Cognitive–behavioral techniques are less applicable because very ill patients often cannot concentrate well enough. However, for those who can, deep relaxation associated with guided imagery can be powerful enough to transport them, even briefly, to a better place.

Ms. D was a 17-year-old high school student with recurrent, unresponsive acute lymphocytic leukemia (ALL). Her loving parents encouraged her to do more than she could and were counseled separately to minimize their expectations. When seen alone the first time, she was very aware, thoughtful, and saddened by her parents' requests, yet showed remarkable understanding and maturity. She gladly selected deep relaxation and guided imagery instead of further "talking," and mentioned that she loved hiking. The therapist saw her several times before she died, "taking her" on several favorite hikes. This created safe moments for Ms. D, who described them as "really peaceful." Relaxing like this enabled her to stay connected to her own strengths.

Supportive therapy of depressed terminally ill patients is one of the more difficult therapeutic tasks. Every therapist has moments of frustration instead of empathy with the imperviousness of depressive thinking and affect. When found at the door of death itself, the therapist's own feelings of helplessness and fear may generate the urge to flee. Taking an active role in symptom control, psychopharmacological treatment, and behavioral interventions can relieve the doctor and help the patient. Beyond that, if the patient seems to welcome the therapist's presence, then staying with the patient is paramount; a faithful presence at the end is profoundly meaningful to someone who is dying. Some therapists experience such

moments as gifts to them as well.

Confusional States in Patients with Advanced Disease

Prevalence. The prevalence of delirium rises to 85 percent in patients with advanced disease. If one includes the last hours of life, it nears 100 percent, unless the patient lapses rapidly into coma or dies of an acute event, like pulmonary embolus. It is always unexpected and traumatizes the survivors, even though they knew full well that the patient was near death. Some patients slip into an irreversible coma, leaving families in a bedside vigil, which may actually give them a period of adaptation before the final instant of death. However, these events only occur 15 to 25 percent of the time.

Diagnosis. Patients frequently experience some disorientation, impaired memory, concentration, and altered arousal as they become increasingly ill, especially if they are elderly. These symptoms may remain mild or may be the harbingers of a full delirium, as described in Table 27.9–11. Figure 27.9–1 illustrates the treatment principles, and Table 27.9–12 summarizes the causes and treatment of delirium.

The DSM-5 diagnosis excludes the presence of dementia as an etiological factor, but, in the growing elderly population, not only is pre-existing dementia common, it is a strong risk factor for a superimposed delirium and makes the delirium much more resistant to treatment. While it is harder to tease out how much symptomatology is reversible, this does not alter the management.

Table 27.9–11.
Subtypes of Delirium

	Hyperactive	Hypoactive
Type	Hyperalert, agitated	Hypoalert, lethargic
Symptoms	Hallucinations, delusions	Sleepy, withdrawn
	Hyperarousal	Slowed
Examples	Withdrawal syndromes: alcohol, benzodiazepine	Hepatic, renal encephalopathies; benzodiazepine intoxication
Pathophysiology	Elevated or normal cerebral metabolism	Decreased global cerebral metabolism
	EEG: fast or normal	EEG: diffuse slowing
	Reduced activity in GABA systems	Overstimulation of GABA systems

EEG, electroencephalogram; GABA, γ-aminobutyric acid.

From Breitbart W, Lawlor PG, Friedlander M. Delirium in the terminally ill. In: Chochinov HM, Breitbart W, eds. *Handbook of Psychiatry in Palliative Medicine*. 2nd ed. New York: Oxford University Press; 2009, with permission.

Since Paula Trzepacz's important paper, delirium is now divided into:

- A hyperactive hallucinatory delirium, easily recognized by both physicians and families although the early symptoms are often mistaken for depression, anxiety, and poor coping. They may not get the diagnosis right, but they do bring a real problem to medical attention.
- A hypoactive delirium must be differentiated from the more easily recognized hyperactive state (Table 27.9–11). The patients are hypoactive, withdrawn, with varying levels of somnolence and unresponsiveness. They are quiet and undemanding, which may explain why the diagnosis of this reversible condition can be long overlooked. This is highly regrettable since a recent group of patients who recovered after treatment described continued awareness and had acutely unpleasant memories of the experience. Hence, physicians should remain alert to diagnosing hypoactive delirium and treat it as aggressively as the hyperactive form.

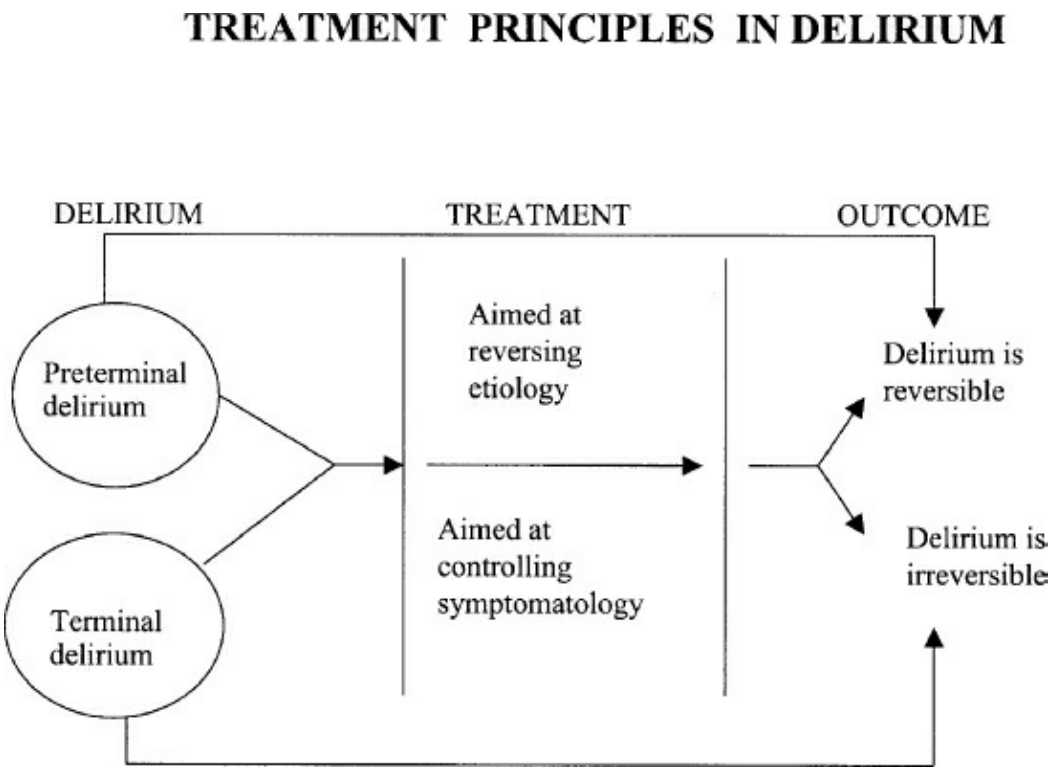

FIGURE 27.9–1. Treatment principles in delirium. Overview of management of delirium in preterminally and terminally ill patients. (From Chochinov HM, Breitbart W, eds. *Handbook of Psychiatry in Palliative Medicine*. New York: Oxford University Press; 2000, with permission.)

Table 27.9–12.
Causes and Treatment of Delirium in Terminally Ill Patients

Causes	Treatment
Opioids	Medications commonly used in managing delirium
Some hallucinatory drugs such as LSD and PCP	Typical neuroleptics

Benzodiazepines	Haloperidol (Haldol)
Anticholinergic drugs	Chlorpromazine (Thorazine)
Sedative-hypnotics	
Steroids	Atypical neuroleptics[a]
Cardiac glycosides, antihypertensives	Olanzapine (Zyprexa)
Anticonvulsants	Risperidone (Risperdal)
Antihistamines	Quetiapine (Seroquel)
Bronchodilators	Ziprasidone (Geodon)
Benzodiazepine and alcohol withdrawal	Aripiprazole (Abilify)
Metabolic and nutritional deficiencies	
Hypoxia	Benzodiazepines may be combined with neuroleptics to calm severely agitated delirious patients.
Fluid, electrolyte, osmolarity imbalance	
Vitamin deficiencies (thiamine, folate, vitamin B_{12}, niacin)	Lorazepam (Ativan)
Infections	Midazolam (Versed)
Liver and renal failure	
Vascular disorders	Anesthetics
Endocrine and hormonal disorders	Propofol
Endocrine tumors	
Carcinoid tumors: mainly ovary and lung	
CNS disorders, metastatic disease, and neoplasm	

A psychiatric consultation was sought to evaluate depression in Mr. K, a 56-year-old practicing attorney with pancreatic cancer. His moderately severe back pain was apparently responding to morphine, but he was noted by the inpatient staff to be more withdrawn, disengaged, quiet, with poor eye contact, and somnolent through the day. The consultant found Mr. K to have disturbed arousal, with mild confusion and disorientation, slow speech, and disorganized thought process. Mr. K admitted to intermittently experiencing visual hallucinations that he had been too embarrassed to report to nursing staff. Diagnosed with a hypoactive delirium secondary to opioid medications, he was treated with a low dose of olanzapine, 2.5 mg at bedtime, improving his sensorium dramatically. He became alert, fully oriented, better related, and without perceptual disturbances or thought disorder, all without decreasing his much-needed pain medications.

- A mixed delirium includes patients with fluctuating activity levels. They may also have normal levels of psychomotor activity, but their attention and awareness are disturbed.

Management. In a preterminal delirium, multiple etiologies are the rule and should be pursued. However, when a patient develops a terminal delirium in the last days of life, symptom control must again be the primary goal. Its general management includes security and orientation measures, pharmacological management, and a search for treatable etiologies. Eighty to ninety percent of delirious patients respond to symptomatic treatment,

even when close to death.

Haloperidol (Haldol) and chlorpromazine (Thorazine) are the main IV medications, but the atypical antipsychotics are effective in patients who use other routes of administration. New routes of administration are being actively developed for several of them. Given alone, benzodiazepines worsen delirium and are contraindicated, but they are useful adjuncts to antipsychotic drugs, providing sedation for persistently agitated patients. When cognitive impairment and early delirium are present, antipsychotics alone can be effective in controlling symptoms. The weight gain and type 2 diabetes seen in chronic psychiatric patients are not a problem here, but dyskinesia, akathisia, tremors, parkinsonian rigidity, and, rarely, tardive dyskinesia do occur. Parkinsonian symptoms can occasionally affect muscles of deglutition and intercostal muscles, leading to impaired swallowing and respiration that are mistakenly ascribed to other causes.

Delirium is so common that medical staff deal with it quite matter-of-factly. However, a hyperactive one can frighten family members, who see it as a drastic change for the worse. They fear a permanent change in the patient, an especially upsetting one, given the content of hallucinations, their commonly paranoid cast, the appearance of anguish and terror, the frequent aggressiveness, and the loss of dignity associated with many delirious behaviors. In such situations, family education is critical; in fact, explanations and reassurances should be routinely given at the first hint of any symptoms to minimize their trauma. It usually falls to the nurses, who are well acquainted with it, closest to the patient's condition, and well aware of the needs of the family and visitors. Hypoactive deliria, however, are often referred to psychiatry as depressions.

PSYCHOLOGICAL ACCOMPANIMENT OF THE DYING PATIENT

When facing a bedrock human experience with an imminently dying patient, the usual language of therapy should be altered so as to avoid pathologizing tragedy. Accompaniment is the goal of interactions with dying patients, even those who are less distressed, and even while aggressive efforts are made to treat all treatable diagnoses and symptoms. The psychiatric consultant must be familiar with the patient's medical condition. The patient–psychiatrist boundary should not be as well-defined with terminally ill patients as with healthier ones. The doctor should be alert and responsive to signs of increased discomfort. Touching bedridden terminal patients calls for judgment but need not have negative connotations. They are weak and sometimes unheard and unable to call for help. Helping them call for help is itself a strong intervention. Untangling or pulling up a sheet the patient cannot reach can be offered casually. Helping to contact the nurse for more complex needs can always follow. Psychiatrists are well trained to respect boundaries, but this is a time when some flexibility is indicated.

Family involvement must be explored and encouraged, and staff members queried for their views. The latter are, especially at this time, a part of the operant group and must be included in understanding and resolution of difficult situations.

Sharing the Patient's Subjective Experience

Dying patients should be engaged for as long as they are able to respond. Giving them a chance to share feelings, memories, and achievements honors their existence and is a precious gift to them. A patient who is talking with an empathic listener feels herself/himself as *still living*, not dying, during that conversation. Eliciting and listening are more important than anything that the interviewer has to say. Although supportive responses are appreciated, unrealistic reassurance aborts authentic conversation and is cruelly isolating.

Many dying patients are deeply troubled by unresolved personal issues, and often keep silent in response to covert messages from those around them. The therapist should try to explore them with as much care as with any patient. Besides the painful presence of their disease, they may be still full of thoughts about the meaning of their lives, the values by which they have lived, the relationships that they feel they have honored (or betrayed), about death, about life after death, about their faith or lack of it, and about their relationship to God. The conversation opens the door to sharing burdens and shedding tears. The relief it can bring is like letting fresh air into the room.

Patients are also grateful when helped to reconnect to happy memories, meaningful achievements, joys, and laughter. Consultants can judiciously share some achievements with the team, who will be glad to discuss them with the patient, rather than only drug doses. Patients close to death can still enjoy mundane activities that bring them closer to loved ones or to treasured aspects of themselves. A late life review may reveal positive meanings and connections, helping patients value their lives with more depth, solidity, and self-esteem.

Some patients may have late emotional changes of desires or feelings. It has been shown in several settings that, as patients become sicker, they accept an increasingly limited quality of life as worth living and opt for more onerous, less promising treatments if they offer even a small chance of help. It is reassuring that IRB committees make sure that only procedures with some potential are allowed for such helpless patients.

Disease-Specific Considerations

Different illnesses routinely bring special issues. For example, a patient on dialysis can opt for death rather than treatment three times a week and is more vulnerable to acting on feelings of depression, anger, hopelessness,

and reactions to family neglect or conflict. The cancer patient must acknowledge possible death during complex treatments, while hoping for a cure or remission. Treatment decisions in the face of disease that is less and less treatable become increasingly difficult and can cause acute anxiety. Stem cell transplant patients experience high levels of anxiety and depression because they are getting a last chance with high stakes and high risks. The availability of organ transplants has created a large population of patients who know they may die while waiting for a cure that could be just around the corner. A patient who ignores medical advice to amputate a limb, and is later diagnosed with metastasis, suffers a great deal and benefits from added support and understanding. Other patients who have lost limbs may have lost a very important part of their identity. Neurodegenerative disorders are associated with increasing physical disability and dependence. If cognition is maintained, a number of patients and their caregivers can manage well until the end. When loss of cognitive faculties is present, the problems can be difficult for the caregivers.

Group Therapies

Given enough survival time, group interventions for patients with advanced disease are helpful. They provide a forum in which patients can "detoxify death" by relating to each other's concerns, talking about dying without fearing that they are upsetting loved ones, and practicing getting acquainted with dying. Many other kinds of approaches have been developed, including groups that use rational–emotive therapy, *meaning-based* groups based on Viktor Frankl's work, *spiritually based* groups that use a nonsectarian but more faith-based vocabulary, groups that use mindfulness-based meditation techniques, and many free support groups run by nonprofit organizations, each focused on a particular condition. Various kinds of family involvement and support have also been given of benefit to both family and patient. While this subject deserves further attention, it goes beyond the scope of this chapter, and is well documented elsewhere.

Other Modalities

Hypnosis, guided meditation, deep relaxation techniques, meditation, minimal movement yoga, music therapy, and techniques that address body and spirit, such as soothing the body and mind with forms of touch and gentle massage, are all helpful to patients who can participate. The consultant can make a personalized tape for the patient during which gentle relaxation and breathing are followed by a detailed evocation of a place or experience chosen by the patient. If the result is relaxation or peaceful sleep, the patient will continue to use it at home. Music therapy in skilled hands that follow the patient, not the score, has been observed to

ease respirations and calm agitation in very ill patients. All of these methods deserve the growing attention they have been getting.

SPIRITUAL CONCERNS

More than 70 percent of patients turn to religious beliefs for comfort, and many others yearn for spiritual sources of meaning. The therapist must engage this side of the patient's inner life in a comfortable, respectful way. The consultant should not presume to replace a chaplain but should integrate this aspect into psychological concerns, as most patients have learned that their religious or spiritual concerns do not find a home in the medical environment. Does the patient have religious beliefs? Did he or she have them in youth? What is his or her relationship to them? To a personal God? To a sense of a transcendent reality? Ned Cassem suggests that the patient's relationship to God be viewed as one of the patient's many important relationships. Some patients describe themselves as not religious but spiritual. If so, can they describe how they experience their spirituality? Can it be expressed through the values by which they have defined themselves? Through belief in a greater power or life force? For people with environmentalist convictions, the unbroken chain of life in nature may be a link to immortality.

The psychiatric consultant—and everyone else—should facilitate a connection to appropriate religious guides for any willing patient. Their presence can offer healing that no layperson can duplicate. Patients may lack the energy or the will to ask for them, and many do not know that they are available.

As mentioned earlier, patients with a well-elaborated faith, especially one that includes a reunion with God in the afterlife, tend to do better at the end. They may experience the illness as a test of their faith, may view suffering as having redemptive value, and may gain strength from the belief that God will not send them more than they can handle. However, they also often ask, "Why me?" and may struggle with anger, fear, a sense of betrayal or abandonment, disappointment, self-imposed guilt, and, at worst, a loss of faith that leaves them truly desolate. If the patient wants to continue to speak with the therapist, it is appropriate, but referral to an experienced religious authority should also be on the table.

Decision-Making Issues

Many patients base medical decisions about life-sustaining interventions on religious beliefs. The belief that life belongs to God and must be nurtured as long as possible, or a belief in miracles, or both, are often the reason given for demanding aggressive life-prolonging measures. This may be quite sincere, but patients and families may use religion as camouflage for dynamic conflicts or as a way to control end-of-life treatment decisions

because they lack trust in the staff or are unable to face death. Mistrust is a frequent reason for rejecting medical recommendations and often crystallizes around inadequate DNR discussions. Today's clinicians are trained to respect religious beliefs and can also recognize multiple agendas, but it may still take psychiatric input to untangle them and promote conflict resolution.

PATIENT–FAMILY UNIT, THE FUTURE SURVIVORS

The intensity of family relationships becomes even greater during the terminal period. Yet, that is also when one may see conspiracies of silence. With an unconscious patient, it is less necessary to intervene, but the saddest silence is that of a loving family, tense and quiet at the bedside because they are protecting the patient by not talking about dying, and the patient is quiet for fear of upsetting them. Instead of closeness, expressions of gratitude, apologies, reminiscences, and farewells, there is distance. Although surrounded, such patients feel as though they are dying alone.

Family-Focused Grief Therapy

The stakes go beyond the patient. The last days of a dying patient are indelibly imprinted on the memories of the survivors. Events perceived to have gone wrong will resonate for years. Research has shown that families with good communication endure the experience better and have a more normal grieving process. Later, they are better able to function as a family, with better outcomes for individual members too. In one study, families that had discussed dying openly were initially more upset but, in the end, were less distressed than families that had been protective. It is likely that the patient died still feeling loved and valued.

Psychiatrists can use family sessions to open patient–family dialogues that can identify discrepant views of the illness or treatment, and explore concerns regarding an absent member, which may be undermining patient support and medical management. A major crisis, such as the imminent loss of a member, destabilizes the family structure, creating an opportunity for the psychiatrist to promote adaptive change and reconstitution. For example, family-centered grief therapy, which includes the patient and continues after the patient's death, provides a natural setting for such interventions.

When an older adolescent dropped out of college to take over her dying mother's duties with greater effectiveness, the family subtly discouraged her return to school after her mother's death. Psychiatric intervention while the mother was still alive allowed the patient to play a part in reorganizing family roles in a way that made it more possible for her daughter to continue her studies without facing too much hostility. The patient herself had wanted it settled and the family responded lovingly.

The Last Vigil

As patients become sicker, consciousness waxes and wanes. It cannot be known how aware they are of their surroundings, but families usually keep watch and the patient's comfort measures should not be changed unless it is clear they have become counterproductive. As patients fade, family members often feel as though they are being abandoned. This can be unexpectedly painful, even in relatives who had felt prepared for the event. Involuntary movements are often overinterpreted as last messages. The consultant should recognize these dynamics, help the family grasp the reality, and prevent inappropriate requests for more life-sustaining measures. Comatose or moribund patients may look uncomfortable during agonal breathing, but their level of awareness is minimal. It is difficult for exhausted family members, who may feel guilty for wishing it was over. *These negative feelings should be recognized, validated, and normalized.* Taking turns can be encouraged. As death becomes imminent, nursing staff will be attentive, calling for a physician and supporting the family.

After Death: The Final Farewell

Responses range from dramatic to subdued. The family is given time alone, crying and comforting each other, sometimes singing hymns or songs. Sometimes a spouse or parent clings to the body and refuses to leave. A private leave-taking should be respected and the grievers are usually ready to separate before institutional post-death routines are interfered with. Some surveillance always continues by nursing staff, who will call psychiatry if an occasional acutely distraught member refuses to leave, if suicide is mentioned, or if anything else concerns them.

All of the patient's doctors should be informed about the death so that they can respond personally. A phone call or a handwritten note is highly valued by family members. The consultant should offer bereavement counseling as needed. Family members sometimes call back years later. Hospices automatically provide bereavement counseling. Some hospitals also have bereavement programs, sending letters to the family at intervals during the first year, with condolences and information about bereavement groups. The response is usually very positive.

Unattended Deaths

Many traumatic or unexpected deaths become known from a catastrophic phone call. The patient is gone, but the traumatized family needs help to absorb the loss, cope with the circumstances, and come to some kind of closure. An unexpected death, even an understandable one, such as a heart attack in a cardiac patient, leaves an aching sense of unfinished business over and above the expected grief. If the person was the victim of a crime, obsessive thoughts may be difficult to repress, and grief may turn to rage, with profound psychological disruption. It is also difficult, if not

impossible, for survivors to fully make peace with suicide.

The absence of a body or a locale is also troubling. Soldiers take enormous risks to retrieve bodies. Families clamor for remains or a shred of irrefutable knowledge or proof. Emergency rooms, police departments, and religious and community institutions should be equipped with a list of resources to help survivors of traumatic deaths. Psychiatric input can include program development and consultation to a wide range of professionals and individuals. Families must be helped to construct their own ritual whereby they acknowledge to themselves and others the finality of their loss and perhaps create a place, if not a grave, around which to center their memorializing.

Attended Deaths

Patients can die in an acute hospital, in nursing homes, in hospices, and at home, with or without hospice support. Most patients still die in acute care hospitals, having received active care until shortly before they die. This may occur because death is sudden or because the family or patient needed to be in a place where "everything is being done." Fortunately, a growing number of hospitals have palliative care teams that provide appropriate care in the acute hospital setting and home care programs are still increasing. Many patients die in nursing homes without the benefit of special care.

This unfortunate situation could be remedied by bringing formal hospice care into nursing homes, but funding sources and turf issues need to be settled before it can, if ever, become routine practice. Inpatient hospice care was the first model of care to be developed and is warmly remembered by grateful families whose multiple human needs and those of their loved ones had been met in ways that they had not been trained to expect. As hospice care has gained acceptance, the insufficiency of inpatient beds has encouraged the development of home hospice care. This has, in turn, encouraged more families to elect to keep dying loved ones at home.

In a home hospice program, the patient is evaluated and accepted in the usual manner but stays at home. The patient and family receive extensive instruction about what to expect, including the fact that if they become truly overwhelmed, the patient can be re-admitted to the hospital. They are helped to obtain necessary materials, taught how to use them, and get as much home help as they need. All the while, they receive medical supervision, nursing care, and emotional support with 24-hour phone availability and routine daily contacts.

Without this kind of help, a good death at home can be difficult to achieve. With it, patients feel safe from the abandonment they commonly fear, and families feel safe from the terror of an unmanageable event. The families of these patients work hard, but are more likely to feel competent and in control.

However, this rosy picture should not become a guilt-provoking

expectation. Some families feel unable to manage at home and, indeed, many do lack the physical and emotional resources. When a patient wishes to die at home, the family may concur while harboring serious qualms. When a patient pleads and a family refuses, the clinician must resolve a painful conflict. Sometimes a visit to a hospice or an interview with a hospice worker eases the impasse, leaving the patient more willing to go into hospice or the family more willing to take a chance at home, as long as transfer to inpatient hospice is a guilt-free option.

SURROGATE BURDENS AND FAMILY COPING DURING THE END OF LIFE

The difficulties faced by family surrogates are sometimes discounted by medical staff because the need for their decisions may displace other concerns. However, end-of-life decisions are momentous, and their memory remains as permanently vivid as classic traumatic memories. Laypeople severely overestimate the practical consequences of end-of-life decisions. Proxies feel that they have decided the "when" of a patient's death, while all they did was to minimally alter the "how." Health professionals should be sensitive to these distortions and should educate and support surrogates well before the actual moment of decision. Early, extensive communication can positively impact the surrogate's ability to grieve and to reach a healthy sense of resolution.

Three months after her 83-year-old demented father's death, Mrs. X consulted the hospital psychiatrist because she was haunted by the decision that she had been asked to make about stopping antibiotics during his terminal pneumonia. His baseline pulmonary function was getting poorer, which led the physician to ask the question in good conscience, but she experienced her final act on her father's behalf as a betrayal.

Grief counseling focused on two areas. The first area was her relationship to her father before and during his years of dementia and how they led to the final decision. Had her family relationships been seriously dysfunctional, or had she been more vulnerable, the counseling might have required a longer course. This bifurcation is a common technical issue for psychiatrists dealing with cases of complicated bereavement. If the physician wishes to make the option available, it should be openly discussed, and the goals of treatment should be altered through a mutual decision process. The second area was her relationship to his caregivers, who, while generally supportive, had underestimated her vulnerability and her need for their affirmation of any decision she made.

Medical caregivers may not realize how little some families know, and how much their words and behaviors are overvalued. Family members bring intense feelings and possibly contradictory agendas to the bedside. There may be jostling for authority, or determination to replay an important family scenario one last time (e.g., favorite child, supportive sibling, or patriarchal authority), or confrontations previously held at bay. Formal proxy or informal surrogates are at added risk for being either criticized by family members or pressured for decisions by the staff. Psychiatric consultation should be called as early as possible when this constellation is identified.

Primary Caregivers

Primary caregivers are a special case. Their physical and psychological distress has been found to be, in some cases, as great or greater than that of the ill, but not dying, patient. Studies have shown that they experience a 25 to 30 percent incidence of job loss, and more than half move to lower paying jobs to accommodate the need for flexibility. When the patient nears death, some caregivers feel ready or even eager to let go, which may cause some guilt. Sometimes they cling desperately to an individual whose continued existence has become the raison d'être for their own. This should not deny the finding that almost all caregivers also describe feelings of joy and transcendence while carrying out their task, as described in a study of HIV patients and their partner caregivers. However, they still come to the end in a state of depletion, and their welfare should be of concern to clinicians.

Ms. T and Ms. M were sisters who had lived together after they were widowed, although both had children. Ms. T had cared for Ms. M as the latter's coronary artery disease had worsened, and was the health care proxy. When Ms. M suffered a stroke and became comatose but not brain dead, her son began to discuss goals of care with the physician, who had to remind him that it was not his decision. At the same time, she had to urge Ms. T to overcome her reluctance to confront the severity of her sister's condition and to protect her from her nephew's growing resentment. Aware that the closeness between the sisters had created a distance from the next generation, the physician focused on fostering rapprochement between them. As the son acknowledged his appreciation for the care that his mother had received, his sister felt more supported and could begin processing her imminent loss.

LEGAL ISSUES: ADVANCE DIRECTIVES, PROXIES, AND SURROGATES

This section reviews information that does not always affect an end-of-life case, but needs to be familiar to physicians and other staff.

Advance Directives

Advance directives (ADs) are instructions left by patients about their wishes regarding medical interventions once they can no longer make their own decisions. Legally binding in all 50 states, they are dictated by the Federal Patient Self-Determination Act of 1992. ADs refer to the right to *refuse*, not to demand, treatment. This distinction is sometimes lost on patients and even more on families, opening the door to stressful confrontations. Although the use of ADs is increasing, the majority of patients still do not have them, leaving family members to make decisions.

Living Will. Patients may leave instructions about what care they do or do not wish to receive after they become incapacitated. These can be extremely helpful to family members and medical caregivers. However, they are often vaguely worded and cannot predict the exact form that terminal events will take, so there is often an element of uncertainty about

their interpretation. This can be lessened if there have been family discussions and/or legal advice.

Health Proxy. Patients may formally designate any individual to function as a health proxy and make any and all health decisions on their behalf. The law dictates that choices be made according to the following hierarchy: (1) the patient's own wishes; (2) if the patient's wishes are unknown, the patient's best interests; (3) if that is ambiguous, the decision can be justified as being the most reasonable one; and (4) lastly, state interest can be invoked; the latter is currently applied mainly to pregnant women, in which case the welfare of the fetus has increasingly been seen as competing and having priority over that of the mother, but that is another story. The proxy cannot overrule any patient's written instructions, especially with respect to life-sustaining measures, such as feeding, hydration, and life supports.

Family members are believed to have the most reliable knowledge of patients' wishes and are deemed natural surrogates. Yet, numerous studies have found that agreement between patients and their family members about end-of-life decisions is no better than random. Indeed, their intimacy with the patient creates ambiguous motivations. Some proxies approach the task without any previous discussions with the patient and do not feel confident about understanding the patient's wishes. Others harbor different opinions and carry out their own wishes, despite legal instructions to the contrary. Both scenarios engender significant guilt, which endures long after the patient's death. In general, the psychological and ethical burden of proxies, as well as their need for ongoing education and support, is profoundly underestimated.

Surrogate Decision Maker. When patients have left no formal instructions, the nearest relative becomes the decision maker by custom and by default. It is usually the one who is most involved, but, occasionally, the lot falls to one who is emotionally or geographically distant, uninformed, or unwilling. This leads to problems, mentioned above.

Durable Power of Attorney. Durable power of attorney is a broader legal power that includes the health proxy along with financial powers. It has been legally available for a long time but was less widely used.

Brain Death Determination. Brain death is defined as two successive determinations of a flat EEG at least 6 hours apart and is accepted as a legal alternative to death by cardiac arrest. Many families have difficulty accepting that a loved one with warm flesh and a beating heart is truly dead. Often, life supports are continued for a period of time, while the family is helped to come to terms with the situation. Some families invoke

religious beliefs to refuse it altogether, creating stalemates that often go to institutional ethics committees with unpredictable results. The growing use of organ transplantation has been a factor in the legal acceptance of brain death, because it enables a better chance for the transplanted organ to work well.

Autopsy and Organ Donation. A patient's advance interdiction of autopsy or organ donation will not be overridden. However, in the absence of instructions, families are asked to make the decision. It can be upsetting for some families, especially the process of harvesting organs from a body that is brain dead and still appears alive. Even without pre-existing religious objections, some families find it difficult to allow. Ethicists have weighed concerns about the moral acceptability of pre-death procurement as decided by surrogates in the absence of clear patient instructions. Although there is an appropriate firewall between those with a stake in organ transplantation and those with a direct responsibility to the patient, more and more people are aware that refusal deprives unknown individuals of a chance to live. A few sensitized family survivors may experience painful conflicts and may harbor residual guilt.

Table 27.9–13 summarizes this information and points out the extent to which all of these categories are subject to interpretation and argument. True to current society, only the instructions about financial and legal matters are unequivocally binding. The thicket of decisions confronting families today would bewilder a time traveler from as little as 50 years ago, who would glance at the patient and know it was time to mourn—nothing more, nothing less.

Table 27.9–13.
Advanced Directives

Directives	Patient	Family
Medical decisions		
Living will	Often causes problems if lacking details about withholding and withdrawal of life supports, feeding, nutrition, dialysis, total parenteral nutrition, blood products.	Changing circumstances are associated with ambiguities; the family may deny the current applicability of the living will.
Health proxy (or durable power of attorney)	Same problems as with living wills	
Legal guidelines		
Substituted judgment		If there are no written instructions, the proxy may act freely but can be opposed by the medical staff.
Best interests		
Reasonable		

decision		
State interest		
De facto proxy behaviors		Conflicts may be referred to:
Proxy desires		An ethics committee
Staff desires		The courts (rare)
Brain death determination	Legally defined	Families may still demur on vaguely worded religious grounds.
Autopsy	Usually no patient instructions	Family demands are respected.
Organ donation and transplantation	Patient's desires should be honored. Formal documents are binding.	Patient's choice may be overridden by family, absent formal directives; none of these decisions can be overturned without recourse to legal intervention.
Financial and legal decisions		

Table 27.9–14.
Evaluation of Cultural Issues: The ABCDE Model

*A*ttitudes: Elicit *a*ttitudes about illness, family care, and responsibilities.

*B*eliefs: Explore *b*eliefs and ritual practices.

Ask direct questions about faith and religious experience; follow up on cues. "What does that mean to you right now?"

Do not respond to statements of religious and spiritual values with facts.

Reflect them respectfully.

Do not deny or pretend to go along with expectations such as miracles but use "I hope" or "I wish" statements that indicate a shared desire for the best outcome that can be arranged.

*C*ontext: Inquire about the cultural *c*ontext in which the patient and family operate.

*D*ecisions: Determine their *d*ecision-making style.

*E*nvironment: Learn about their current *e*nvironment.

Adapted from Koenig BA, Gates-Williams J. Understanding cultural difference in caring for dying patients. *West J Med*. 1995;163:244.

CULTURAL, SPIRITUAL, AND RELIGIOUS ISSUES

Culture and religion shape the ways in which people make meaning out of suffering, loss, and death, and the ways in which they understand the sick role, the caregiver role, and every other basic role, such as parent, partner, child, or sibling. When clinicians share the same unconscious assumptions, decisions can be negotiated more easily. Major differences still occur but play out against the same background music. The consultant should be attuned to the differences—not only intercultural differences, but also the ways in which given patients or family members deviate from their intracultural norms. It is a tall order but well worth the effort in decreased psychological stress, smoother decision making, and improved care. Furthermore, the knowledge gained in one case later eases the management of other cases with the same background. Many institutions have formal seminars or protocols to educate staff about recurrent cultural or religious scenarios. Table 27.9–14 describes a model for the evaluation

of cultural issues.

Cross-Cultural Barriers

Cross-cultural problems can be manifested in multiple ways, the most important ones of which are discussed here.

Trust. It is more difficult to establish trust across wide cultural barriers. Minority groups are thought to want more aggressive care, use fewer DNR orders, and participate less in organ donation and research because they do not trust medical caregivers to treat them as well as they treat the majority group. Some mistrust is justified. There are significant differences in quality of communication, access to care, and quality of care that result in well-documented poorer health outcomes for minority populations. However, when trust is present, the differences are often minimized.

Truth Telling. The requirement for truth telling is so embedded in US law and custom today that it may be hard to remember how foreign it is to other cultures. Surveys of Chinese, Korean, Japanese, Ethiopian, Mexican American, and Native American groups found that only 35 to 52 percent saw end-of-life truth telling as desirable. Furthermore, truth in prognostication is not well practiced even in western medicine. Not all physicians prognosticate truthfully and some may be using good judgment by respecting a strong family culture. It is also useful to remember that nonverbal communications are more highly valued in other cultures. The ability to intuit what the other person is thinking, and to address concerns without too direct a confrontation, is appreciated.

Family Involvement in Decision Making. Families are always involved in decision making but, in some cultures, there may be a strong duty to completely relieve the patient of that burden. This custom is independent of socioeconomic status and, in a study of Japanese American families, was independent of the degree of acculturation as well. A family whose request for nondisclosure is ignored may feel upset, angry, and disrespected. If the patient gives informed refusal, the family can be accommodated. However, if the patient is ambivalent or appears interested, a psychiatric consultation should be requested and must include an evaluation of cultural and religious factors as well as patient and family dynamics. The family should be educated about the patient's situation, and their concerns about disclosure should be explored nonconfrontationally with indirect questions about the consequences. "How will you feel if you cannot answer your father's questions?" "What will your mother think as she becomes sicker?" When the physician feels strongly that the patient must be asked to participate, it may be wise to also call an ethics committee consultation, for

which the consultant's analysis is a useful document.

Mrs. M came from a traditional Chinese family and did not speak English. When she was found to have advanced carcinomatosis, her son and husband refused to tell her about it and were indignant at the thought that anyone would upset her with this information. Consultation with an acculturated Chinese physician from the same subgroup confirmed that their reaction was the standard one. He interviewed the patient and found that she had decisional capacity, had no interest in her medical condition, and shared her son and husband's views on the nature of family support. This was considered adequate informed refusal and, although she was periodically revisited by the same physician, her demeanor never altered.

Such cases arouse discomfort in medical staff well indoctrinated in the rules of informed consent. However, assuming that a westernized moral climate applies to everyone has its dangers. It is sometimes wiser to respect family mores.

Religious Issues

Every religion has beliefs about death and every culture has highly valued rituals to honor them. They play a major role in family expectations and reactions. Every hospital or hospice staff should be familiar with such customs in the major religions that they will encounter. They should have access to consulting appropriate clergy, to enable them to understand and respect family reactions. The humanizing trends described above have made this more commonplace.

STAFF–FAMILY INTERFACE

In any inpatient setting, medical caregivers are an integral part of the terminal scenario. Their words and nonverbal cues are scrutinized, giving their conscious and unconscious biases an exaggerated impact on events. Staff members can become overinvolved with a patient and family, can scapegoat them, or can overestimate or underestimate their coping abilities, all of which negatively impact care.

In the hospice setting, patients, families, and staff are already all aware of the outcome and do not bear the burden of the many transitions and secrets that characterize the modern death trajectory in hospitals. Nevertheless, continued motivation without burnout and without developing dysfunctional norms requires vigilant attention, and hospices have had built-in psychological support from their inception which probably explains their lower turnover rate of personnel.

In hospital settings, however, staff members must make constant transitions from day to day and from patient to patient on a given day. They must repeatedly shift set and help patients and families struggle through painful changes. Fortunately, in today's setting, psychological support is available to staff as well as patients, often within case discussions that now include psychological as well as medical issues.

PEDIATRIC END-OF-LIFE CARE

The death of a child always feels like a crime against nature. The

biologically determined instincts that drive some adults to care for sick children may endow pediatric wards with a particular intensity that must be channeled with experience and training, but it also provides the children with the responsiveness they need. Medical caregivers must also be aware of the child's level of development and request psychiatric help early if there is a question of cognitive or emotional problems, since they will interfere with the child's ability to handle the treatment and further course.

Their parents will remain distraught, although some may acquire strength in the process, while others will require ongoing attention. In all cases, *a cooperative, supportive connection between caregivers, the family members and the child must be maintained throughout*. This tripartite connection helps the child feel more secure, while the family appreciates the consideration and support, and the caregivers can be, and feel, more effective.

ETHICAL RAMIFICATIONS OF THE TRIPARTITE CONNECTION

The requirements of the tripartite connection define the entire course of treatment, and hinge on three values that must always be present: *trust, honesty, and confidentiality*. Their importance revolves mostly, though not exclusively, around three main issues: (1) The parents are facing the loss of a child to whom they gave life, which, along with grief, induces a sense of irrational guilt that cannot be fully erased and may drive them to make extreme demands against the child's interests. (2) Legally, pediatric patients are always third parties, ranging from helpless infants to 18-year-olds who may be wiser than their parents, yet not free to make their own decisions. (3) The medical world had, until recently, not focused much on the fact that adolescent maturation is as relentless as breathing and continues its struggles to survive under a dark sky. Parents able to help younger children may find it more difficult to honor the adolescent's needs and require more information, explanations, and support.

Trust and Honesty

This creates the need for what was well characterized in this quote: "The goal from the outset is to establish a dual treatment relationship with both parent and child with a common goal of understanding the child's thoughts and feelings and improving the child's well-being. *It is necessary to establish a three-way message of trust."* (see Kearney, J.) Physicians who interact regularly with the parents and child carry a special burden. The situation demands that they earn everyone's trust and honesty when meeting with child and parents, separately or together.

Confidentiality

Negotiating confidentiality is a delicate matter, and is dependent on the

doctor's communication skills. Having addressed the need for full honesty, she or he is obligated to begin by explaining to the parents the legal limits to confidentiality when safety is at risk. Next he must show the kind of understanding and support that convinces parents to trust that they will not lose touch with their child, even if they are not told all the thoughts the child has shared. He further explains that he *needs* that trust because the child will only trust him if he can promise to use appropriate discretion when talking to his parents. Finally, the doctor describes his own determination to always keep in mind the priority of sustaining the parent–child relationship.

These meaningful issues must be addressed from the beginning of treatment, as gradually as necessary, when there is time for the parents to adapt incrementally to the more difficult times that lie ahead. It may require a psychiatrist, since parents are such an important part of the patient's treatment that they need the same level of attention, for the good of the child. Another potentially difficult position arises when parents, who have the legal right to make many decisions, are obstructing the best course. This is one more reason why creating a trusting collaboration from the beginning is extremely important.

DEVELOPMENTAL STAGES

Infants and Toddlers

It is hard for anyone to watch the suffering of infants and toddlers, but while staff can move on more easily, the pain of parents is inherent in their parenthood. Giving birth implied giving life, and they are now failing and helpless. They need support to remain stable and keep providing the loving presence the child expects and needs. The medical team will do their part for the child's comfort and will continue to sustain the parents, although further help is often needed.

School-Age Children through Preadolescence

Within this age limitation, children vary widely in their capacity for absorbing what is happening. Adults often underestimate children's understanding, particularly in a crisis that draws out their most protective responses. However, the reality is that, like sponges, children listen intently to what they hear, soak up the emotions that flow around them, and subject everything to an active process of interpretation, organized around finding out why they are being left out. Punishment is often their frame of reference, and they elaborate in various ways. When this happens, the child has unwittingly been made to carry what should be the parents' burden.

These suffering parents should be helped gently from the beginning to understand more about their child's response to their situation, and should learn to speak to the child in age-appropriate ways, surrounded by

reassuring affection. A knowledgeable individual must review the child's level of understanding so as to acquaint parents with appropriate ways of responding, explaining to them that these interactions are as fundamental to parenting as the care they had given their infants. On today's wards, parents must be given special support, which is one reason why many pediatric wards have their own support staff, including psychiatrists and psychologists.

The Child's Cognitive Understanding of Death and Dying

Understanding children's developmental stages and their level of insight about terminal illness has practical implications for the clinician. It provides a basis for effective communication with children about their illness and guidelines for improving their coping skills. Although children's reactions depend on their personality and level of development, they are also strongly influenced by their emotional family milieu, and must be understood within that context.

Children's reactions to their own dying or to the death of others follow a developmentally determined conceptual framework, enriched at times by their own interpretations of familial conversations. Accurate comprehension of death requires the mastery of four basic concepts: irreversibility, finality, causality, and inevitability (or universality). These are detailed in Table 27.9–15, and should be familiar to adults dealing with children facing severe illness and death, because the children's attempts at self-explanation make their course worse than it needs to be.

During the first few years, children have almost no concept of death other than as a disappearance or separation. Most children under the age of 5 years do not think of death as final and believe that the departed person is somewhere else. The most commonly offered explanation is heaven. Worried about being separated from their caregivers, young children may ask, "When I die, will Mommy come with me?" However, even children under the age of 5 develop a precocious understanding of death if they experience a traumatic event like the death of a parent or sibling or any event exposing them to death. Children between the ages of 6 and 9 years understand that death is final and entertain concrete thoughts about the process of death and dying. For example, an 8-year-old child may conceptualize death in terms of damage to a particular injured body part. Stories of devils and ghosts may frighten children, and they may have vague ideas about the soul continuing to live on after death. A 7-year-old child suffering from terminal Ewing's sarcoma asked his mother, "When I die, will my leg fall off and stop hurting?" Children aged 10 years and older have a fairly good understanding of death. They begin to see patterns in nature, recognize life cycles, and often become aware of their own mortality. Some are quite precocious and could qualify as young adolescents.

Table 27.9–15.
Concepts of Death and Implications of Incomplete Understanding for Adjustment to Loss

Irreversibility: Death is seen as a permanent phenomenon from which there is no recovery or return.

Example of incomplete understanding: the child expects the deceased to return, as if from a trip.

Implication: failure to comprehend this concept prevents the child from engaging in the process of detachment from the deceased, a necessary first step in successful mourning.

Finality (nonfunctionality): Death is seen as a state in which *all* life functions cease completely.

Examples of incomplete understanding: the child worries about a buried relative being cold or in pain; the child wishes to bury food with the deceased.

Implication: may lead to preoccupation with physical suffering of the deceased and may impair readjustment.

Inevitability (universality): Death is seen as a natural phenomenon that no living being can escape indefinitely.

Example of incomplete understanding: the child views significant individuals (i.e., self, parents) as immortal.

Implication: if the child does not view death as inevitable, he or she is likely to view death as a punishment (either for actions or thoughts of the child or the deceased), leading to excessive guilt or shame.

Causality: The child develops a realistic understanding of the causes of death.

Examples of incomplete understanding: the child who relies on magical thinking is apt to assume responsibility for the death of a loved one by assuming that bad thoughts or unrelated actions were causative.

Implications: tends to lead to excessive guilt that is difficult for the child to resolve.

From Schonfeld D. Crisis intervention for bereavement support: a model of intervention in the children's school. *Clin Pediatr*. 1989;28:29.

Younger children will resort to themes of retribution and magical thinking when reacting to their illness. Perceiving their suffering as a punishment for real or imagined misdeeds, they accept the burden of personal shame or guilt to account for it. They view their illness in an egocentric manner and express their fears accordingly: "Did I cause it? Who is going to look after me? Am I going to be abandoned?" An older child or adolescent, aware of the irreversibility of death, may exhibit anger, anxiety, and depression while trying to cope with the terminal illness. However, some children inspire respect by their courage and stoicism in the face of death.

Communication with Severely Ill Children

Children quickly learn in whom they can trust and confide. Whether or not individual children should be told about their prognosis is not a simple matter. Parents have psychological, moral, and legal authority. Their wishes must be explored, earlier rather than later. Some parents do not want their children to be told they are going to die. It is useful to ask them *how they would respond if the child asked them if he or she was dying*. It should be the opening of an ongoing conversation, aiming to help them

understand more about their particular child's response, the fears, the sadness, and the need for *authentic* reassurance. When some families are willing to tell the child, one individual should be selected to talk to the child, but only after discussions with the physician. Sometimes, the physician will be asked to be present during such conversations, or might even be asked to be the one to speak to the child. The same information must then be conveyed to everyone close to the child because inconsistent explanations are confusing and even emotionally crippling. In such cases, an experienced child psychiatrist or psychologist can be helpful early in the process.

When the child asks a question about death, it is very important to first clarify what the child already knows and the reasons behind the question, since they are often different from what adults expect. Some children are responding to the anxiety and grief projected by their parents; others are preoccupied with pain, discomfort, mutilation, abandonment, and loneliness. The child should be given ample time and opportunity to express concerns, and any frightening misconceptions should be clarified. Even when children are told their illness is fatal, they should be comforted and reassured that the physicians will do everything possible to keep them comfortable and that they will always be loved and cared for. If the physician expects a difficult session, it may be useful to ask a child therapist to attend. In fact, early referrals are frequent and desirable.

Communication with Parents

Communicating with the parents of a very sick child follows the principles used in families of adult patients, save for one crucial difference: the parents' deep sense of responsibility for the child's welfare, and the attendant feeling of failure. Some parents are too overwhelmed to function. Others throw themselves into research, taking the child for multiple opinions and treatments. Still others never leave the child's bedside, which is helpful for younger ones, but may dominate the child's relationships with other caregivers. Constant tasks leave them little time for thinking or feeling. A few may actually remove themselves. Fortunately, pediatric wards still witness remarkable feats of loving accompaniment in the face of increasing grief and exhaustion.

Parents often need guidance and support, particularly about whether they should treat the child normally or with special attention, especially when they have other children. Intensely protecting a young child beyond what the illness requires confuses the child, who has no framework to explain this novel behavior. The older child and the adolescent understand its basis; some welcome this protection, and some crave and require it. Others prefer more freedom and privacy, often acting out a fantasy of being well or finding a means of denying the fatality of their disease. The advice and support of an experienced child psychiatrist can help parents strike the

right balance.

Parents may demand aggressive medical treatment because they cannot bear the thought of future regret over any omission. These feelings may not be openly voiced to the physician. They should be sought for and discussed openly if any other option is to be considered.

Finally, as parents learn to accept that their child will die, they reach a state of resignation that spares them the previous bouts of activity, ushering in a different kind of suffering and anticipatory grieving. In this exhausted state, one or both parents may find themselves wishing the child dead. These unacceptable secret thoughts burden them with excruciating guilt. They should be *sought for and clearly normalized*. It is a major gift to relieve them frequently of these attacks against themselves.

PSYCHIATRIC SYNDROMES IN THE TERMINALLY ILL CHILD

Although psychiatric problems in dying children are similar to those seen in adults, their presentation varies depending on the child's stage of development. The young child's ability to report symptoms of physical or emotional discomfort is generally undeveloped or unreliable. Other young children may not wish to disclose their suffering out of unwarranted shame or guilt from real or imagined wrongdoing. Clinging to magical thinking, these children may fear that voicing their complaints will worsen their symptoms. Others may be the opposite, and will use a lot of energy on loud complaints, in which the line between physical and emotional is hard to distinguish. A few children may assume that if they wish to be healthy and think "good thoughts," the problems will go away. All this leads to withholding of information and compromised care. Older children may not be much more forthcoming. An astute physician knows to look for symptoms even if the child is not vocalizing them. Consulting with child therapists is useful in such cases.

Symptoms of delirium, depression, and anxiety occur in children with terminal illness, and their causes are physiologically the same as in adult patients. Children with pre-existing psychopathology are more vulnerable to developing psychiatric syndromes as they confront terminal illness, but the symptomatic treatment is still evaluated the same way as with adults.

Delirium in Terminally Ill Children

As in adults, symptoms of mild delirium may be overlooked or mistaken for regressive behavior in a young child. Because a young child's world normally fluctuates between fantasy and reality, symptoms of visual or auditory hallucinations may be viewed by the parents as the child's fantasy rather than as a delirious state of altered perceptions. Disoriented states may be overlooked or minimized because children's attention to the items measured is variable at best. In response to this, Paula Trzepacz has

adapted her delirium scale to children and adolescents.

As with adults, treatment of delirium requires immediate institution of symptom management, along with a search for the etiology and correction of the underlying cause if possible. Neuroleptics, such as haloperidol or droperidol (Inapsine), as well as the newer antipsychotics, may be used to treat the symptoms. Benzodiazepines are usually avoided because of the associated disinhibition. In addition to this treatment, reassuring and reorienting the child by placing familiar objects and people in the room is helpful. Psychoeducation of parents can relieve fears and misconceptions and allow them to better support the child.

Depression in Terminally Ill Children

Depressive symptoms are expected in children battling a life-threatening illness. One may see anticipatory grief and mourning for the perceived loss and impending death or behavioral problems, such as regression, opposition, aggression, withdrawal from family and friends, and noncompliance with medical care. Very few children develop a full-blown major depressive disorder that requires pharmacological intervention. As with adults, vegetative symptoms, such as fatigue, anorexia, and insomnia, are so common in terminal illnesses and major depression that it is impossible to tease them apart.

With adolescents, it is important to explore suicidal ideations and fantasies overtly. In their struggle with a bleak and uncertain future, some harbor suicidal thoughts as a means of gaining control over their destiny. To judge the level of risk, the consultant must explore the youth's understanding of what suicide means, the personal motivation for it, and the presence or absence of a plan. Known risk factors include the presence of previous or comorbid psychiatric disorders, persistent and intractable pain, and poor social supports.

Treatment of depressive symptoms begins with helping the families make the children feel confident that they will not be abandoned and that their comfort will always be a top priority. Maladaptive behavior patterns can be explained to the family as the child's way of coping with anger and grief. Some parents can help the child process the myriad emotions that assail him or her in the face of impending death. Exploring the child's feelings through play or drawings can be liberating. Providing children with an open forum to talk about their concerns, beliefs, and fears about death helps them emotionally process what is happening and feel less alone. Some courageous children decide to whom they want to leave their belongings and say personal goodbyes to friends and family in a variety of ways. This authentic and generous engagement can be empowering, but it *must* come entirely from the child; suggesting it could be very upsetting to an unprepared child.

Antidepressants may be necessary for a minority of children with major

depressive disorder, but children should also be receiving active treatment for pain, fatigue, insomnia, and any other treatable symptoms. The etiology should be sought while treatment is instituted.

Anxiety in Terminally Ill Children

Anxiety in the face of death is ubiquitous. It may stem from physiological issues such as intractable pain, compromised respiratory state, or medications, as well as from psychological issues such as separation from parents, unfamiliarity of the hospital environment, and fear of death. One should always treat physiological symptoms, while helping parents treat the symptoms due to the child's own observations, experiences, and interpretations. Ensuring consistent contact with parents, avoiding prolonged separation, addressing the emotional states of caregivers, and making the hospital milieu more child-friendly go a long way toward assuaging the child's anxieties. More severe symptoms may require the use of anxiolytics. Short-acting benzodiazepines are preferred to minimize sedation, respiratory compromise, and disinhibition. Child psychiatrists and other child therapists should be effectively used.

Pain Management in Terminally Ill Children

Pain assessment in children is more complicated than in adults. Children who lack verbal skills cannot communicate pain to caregivers. Young verbal children may distract themselves from pain with play, some recoil into sleep, and some remain withdrawn and anxious. Adolescents may not be forthcoming about pain due to the belief that they are invincible or due to a misplaced fear of addiction.

Yet, with the use of appropriate rating scales, children can be reliable reporters. Validated, age-appropriate scales do exist, and more physicians are now familiar with them. As with adults, it has been found that professionals using clinical judgment assess the child's pain level as lower than do both the parents and the child. Worse yet, the greater the child's pain, the greater the discrepancy.

Most parents recognize their child's pain and are eager to see it relieved, but a few mistrust the medication to the point of forbidding it. Some are worried about addiction, but accept education and support. A few are unable to tolerate seeing their child relax and become less related, because they do not experience it as relief for the child, but oncoming loss for themselves. With PCA, they express fear of overmedication. Yet, it has been demonstrated that children old enough to understand the technique showed significantly diminished use of opioids with PCA. Family education around these issues is important, but often a consultant must ease the parents into realizing how much their reactions are a way of avoiding the truth of their child's condition. Only after they confront the reality can they

wholeheartedly cooperate with optimal pain management.

COMMUNICATION WITH ADOLESCENTS AND YOUNG ADULTS (AYA)

Introduction

One of medicine's strengths is its focus on the individual and the positive attention it brings to each patient's care. The earlier discussion of the close interaction between medical and psychological management for children and adults need not be repeated here, because adolescence overlaps both categories, as will be discussed below. This section will discuss new developments that specifically focus on adolescents and young adults.

Current Interface of Hospitals and Adolescents

Only in the last 40 years have medical hospitals thought about special arrangements for adolescents, since they could still be assigned to either pediatric or adult wards. Yet, human societies have always been aware of the adolescent trajectory as a meaningful period for them and society. There have been increasing numbers of specialists over time, but only in the last two decades have hospitals introduced dedicated spaces and trained individuals to work with them. They have been well utilized by terminally ill adolescents who suddenly have to integrate the most profoundly threatening experiences while making meaningful relationships in the process.

AYAs' Development during Their Illness

Adolescents and young adults, now referred to as "AYAs" by some current researchers, are a group of patients that should not be defined solely by age but by level of development. It could include a precocious 12-year-old or a solitary 25-year-old. An early hospital study found that parents' protective instincts were as powerfully aroused with older as with younger children, which made them cling more strongly to their parental decision making. This observation encouraged caregivers to help parents understand and respect their adolescents' growth. It confirmed what was being increasingly noted, that relentless exploration towards adulthood continued throughout their illness to the very end. This means recognizing that their grim situation does not stop them from thinking about making new peer relationships in the hospital, still focusing on sexuality or at least intimacy, evaluating the life goals lost, maybe finding an appropriate new one, and lastly, despite—or because of—their misfortune, occasionally achieving unexpected levels of intellectual and moral development.

Adolescent psychiatry now plays a well-established role in medicine. Large medical centers provide gathering places where only adolescents, including gravely ill ones, can share their experiences with each other, and where adolescent psychiatrists, psychologists, and teachers have roles to

play. The main purpose is to allow adolescents, including dying ones, to still experience adolescent development. This is an excellent change, but it still is not always available in small hospitals and rural areas, where the smaller number of patients might require a different approach.

It is best to begin as early as possible to help AYA parents help their children to remain themselves. This help includes assessing AYA patients' readiness to discuss their disease, with a plan to help them understand and participate more and more—to the extent that they wish. Over time, their ability or willingness to confront their impending death should be assessed and explored, and if any of them show any ability to deal with it, they should be helped and encouraged. For some adolescents, it is deeply meaningful to write letters, leave gifts, or give instructions about future events such as funerals and posthumous remembrances of their birthday. Others simply hang on to close connections, leaving life on a note of fundamental love and attachment.

A group supported by the National Institute of Mental Health has been studying this area at length. They published the document *Voicing My CHOiCES*™ based on their research as well as feedback from AYA patients themselves. It guides health care providers to help patients and their families navigate this painful and complex process, using individual sections which AYAs can fill out with the support of health care providers, according to their comfort level. *Voicing My CHOiCES*™ provides a thoughtful and thorough discussion of how AYA patients can be helped to explore the important issues, although it also makes them aware of options at every point. When queried, over 90 percent of patients reported feeling positive about participating in the process.

PSYCHOLOGICAL ISSUES OF HEALTHY SIBLINGS

Studies of healthy siblings have found increased levels of regressed behaviors, somatic complaints, altered peer relations, and a drop in scholastic achievement. They have also found higher anxiety levels and lower self-esteem. Children are at the mercy of their environment, and it is easy for anguished adults to forget that siblings still require the same nurturance and now also have additional needs. Caregivers who notice any problems should help parents develop constructive responses. Professional consultation may be required for complicated situations or dysfunctional families.

Parental Response to Healthy Siblings

The illness and death of a sibling is a world-transforming event. It permanently changes family dynamics and sometimes leaves the family in financial straits that diminish the surviving child's future opportunities. Acutely, it profoundly disrupts the family routines and captures the adults'

time and attention, leaving them emotionally depleted. Some parents are strong enough to remain available to their other children, but many react by protecting them (i.e., keeping them uninformed and away from the family tragedy). Sadly, the patient's siblings are marginalized at a time when they most need to feel like integral members of the family. Keeping healthy children engaged in the family drama requires an understanding of the child's developmental level and emotional state, so as to know what will help them feel loved, trusted, and included. Optimally, the staff caring for the patient will help parents maintain a closer and more open relationship with their other children throughout the illness.

Children's Understanding of a Sibling's Terminal Illness

Siblings 5 years old and younger do not understand the terminal nature of the illness and may react negatively to the parents' emotional state. If they sense withdrawal and sadness, they become frightened and may regress or act out, hoping to gain parental attention. Children between the ages of 5 and 10 usually understand their sibling's condition but may become anxious about being abandoned by a distraught parent as well as fearing that harm will befall them. Some children are plagued with guilt because they have been jealous and angry about the attention given to the patient and remember past times when they harbored angry feelings and even wished the sick child was dead. Younger children may believe that this wish was magically fulfilled and thus blame themselves for the course of events. On the positive side, given both reassurance about being loved and age-appropriate information such as well-related options on how to relate to the ailing sibling, many children become eager to show their support by making drawings, bringing special toys and treats to the hospital. With parents' approval and affection, this fruitful participation in family events significantly enhances the sibling's self-esteem and sense of security, and may help the sick child as well.

Adolescent siblings can become supportive, to the point of assuming parental roles toward younger siblings and taking over household tasks. When the family members begin to take it as a given, it may put these youngsters in a position that will interfere with their development. However, in today's world, it is increasingly likely for adolescents to withdraw from the family, acting out or just spending more time with their friends. In any case, they too will need help, whether to prevent them from continuing their parentified role beyond the death, or to help them return into the family fold in time to manage their feelings toward their dying siblings (with help, if needed), and experience the shared affection as well as the grief.

STAFF DYNAMICS IN END-OF-LIFE CARE

The family–staff interface in pediatrics and the associated trust and communication needs are the same as with adult patients, but there are a few special issues associated with the staff. Pediatricians and nurses choose to work with children because they love them. During an increasingly severe illness, they come to know the child well and may become very attached. The expression of affection toward a pediatric patient by caregivers need not have the inappropriate implications it has for adults, as long as medical judgment is unimpaired. Ideally, the staff and the family remain united by their common dedication to the child, but occasionally the attachment is so strong that it creates conflicts. For example, a nurse's maternal instincts can engender an unstated competition with the child's mother which makes her less supportive and affects other staff as well. The mother is usually keenly aware of it but sometimes afraid to mention it. This situation requires a change of assignments for the nurse, with special support and training to prevent recurrences.

Sometimes, staff members are too quick to pass judgment on families that are not meeting their expectations. Once crystallized, these judgments are very hard to revise. Negative stereotypes and competitive feelings are very destructive. They cannot be hidden from parents, who are exquisitely sensitive to staff behaviors and utterances. It is critical that team leaders set the tone and create a culture in which *both* children and parents are equally empathically treated. Consultants must remind the staff how much the children want to feel securely embedded in their families at all stages of illness. Children's profound dependency on their families is such that they are enslaved by faithfulness, even in the face of parental indifference or abuse. *Any staff behavior that threatens the unity in a family is unacceptable no matter how well intended.*

Staff overinvolvement also creates intrastaff rifts. In one case, an oncology fellow became so involved with a child that he argued heatedly over the attending's choice of treatment. In another case, two nurses became so attached to a patient that they arranged their work schedule to have one of them on duty every day, because they felt their colleagues would not care for him well enough. It is not that unusual for young pediatricians to share some of the parents' feelings and, occasionally, to become more involved than they had intended. These doctors must learn to exercise control with the parents and would benefit greatly by sharing their feelings with a trusted colleague. In general, this is why group work and open discussions have become a larger part of the climate when working with terminally ill patients.

END-OF-LIFE DECISIONS IN PEDIATRICS

Assent and Consent

As children mature and are better able to process information about their

illness, they can play a more active role in decision making. However, with the exception of emancipated minors, the child can only assent, while it is the parent who can give legally binding consent. This is usually not a problem because most parents respect the wishes of the dying child, but there may be times when discrepancies arise.

One scenario is that of adolescents refusing treatment that their parents, and sometimes their physicians, would like to administer. This situation calls for psychiatric consultation to allow all participants to explore their feelings and to help achieve consensus about the care plan. Consultation with an ethics committee may be necessary when agreement cannot be reached.

Another scenario is that of an exhausted or frightened parent who refuses treatment that the staff believes a child would accept were he or she not afraid to disagree with the parent. It is more likely to happen if the staff has not achieved an easy supportive relationship.

The single mother of a 9-year-old boy with relapsed leukemia refused a bone marrow transplant on religious grounds. The nursing staff dealing with the child felt that, at age 9, he was aware and would not have minded the treatment, but was afraid to disagree with his mother. The consulting psychiatrist found that the mother felt overwhelmed and unsupported by the staff. Her religious feelings, although genuine, were also a way for her to escape the situation. The consultant urged the physicians and nurses to hold off on requesting any decision, while being very supportive of the mother and validating her apprehensions and pain. Within 3 days, she was ready to proceed with the treatment, which the child accepted readily.

Neonatal and Infant Issues

Neonatologists can now save premature babies who weigh little more than 1 lb and are at high risk for intracranial hemorrhage or respiratory insufficiency. They can also salvage severely handicapped newborns who will always need intrusive medical care and whose quality of life will remain poor. Some people believe life-sustaining interventions should always be used, others believe they should never be used, and a large group thinks that it depends on circumstances. Among physicians who are not absolutely opposed to withholding life-sustaining measures, the relevant quality-of-life issues are: (1) the extent of physical impairment, such as severe neurological damage; (2) the burden the child will place on the family; and (3) the ability of the child to derive pleasure from existence, such as having an awareness of being alive or being able to form relationships. From a philosophical point of view, these are very disparate arguments, but they derive from the pragmatic and casuistic method that has characterized valid medical thinking.

WHERE WE STAND NOW

Decision Making

Truth telling and obtaining informed consent are legal obligations. The intellectual requirements for having capacity to consent are well elaborated. First, the patient must be provided with information about diagnosis, prognosis, and treatment options, as well as a review of the risks, benefits, alternatives to treatment, and the consequences of no treatment at all. Processing this information can be difficult, especially under the dual burden of severe disease and the nearness of death. Given the high incidence of cognitive clouding in very ill patients, every dying patient's right to make decisions might be imperiled if one focused only on cognitive criteria when establishing decisional capacity. Consultants must come when patients are most alert and lucid and *must look for an internally consistent sense of self and a general awareness of their condition.* If present, these should weigh more heavily than the patient's ability to perform on a standard mental status examination. This is not to imply that the consultant should be less aggressive in correcting the patient's cognitive state by all the usual means.

Informed Consent for Experimental Treatment. It is difficult for patients in relatively good health to absorb the basic facts about clinical trials. They may agree to the trial to please their physician or may believe that the physician is offering them a place in the trial because it is good for them. Reality #1: doctors continue to have a personal duty to the patient even if they enroll them in a trial; they must remove the patient from the trial if the patient's welfare is incompatible with the demands of the protocol. Reality #2: patients need to keep believing their doctor is serving their needs whether they are in a trial or not; this is even more intense for terminally ill patients who have become known as *desperate volunteers,* and who enter trials "because" it is their only chance—and indeed, their desperation cannot be assuaged! Its presence places an even greater ethical burden on both physician-researcher and Institutional Review Board (IRB). If both parties do their jobs, the psychological and ethical ambiguities inherent in working with desperate volunteers should not deprive the patients or society of the potential benefits of valid research.

An article on the ethics of phase I oncology trials observed that the risk–benefit ratio of phase I trials is not significantly worse than those approved by the FDA for clinical trials of chemotherapy. It identified three important issues relating to informed consent in terminally ill volunteers: (1) disclosure, which available data show to be adequate; (2) understanding, which also has been found acceptable in more than 70 percent of patients; and (3) voluntariness, which was initially strongly denied by moral philosophers. The authors argued convincingly that patients with advanced cancer have a different perspective than the healthy bioethicists who worry about them and feel they are being abused. Many dying patients are known to be truly yearning to try more low-yield treatments the closer they are to

death.

Ethical Issues about End-of-Life Treatment

There are possibilities of management that are unique to end-of-life care. They include withholding treatment, withdrawing treatment, and invoking the double effect.

All of these were legitimized in 1982 when the Belmont Report, produced by a presidential commission and supported by major Catholic and Protestant doctrinal statements, affirmed the moral acceptability of withholding extraordinary measures, or of stopping them later. It also described acceptable uses of the double effect.

However, it has proved to be surprisingly difficult to define *extraordinary* versus *ordinary* measures in the context of withholding or withdrawing treatment. If food and hydration are *ordinary*, what of nasopharyngeal tube feedings? What of IV fluids or percutaneous feedings? Perhaps everyone can agree that TPN is extraordinary. Cardiopulmonary life supports are generally accepted as *extraordinary*, but what of intubating a patient in worsening respiratory distress without cardiac arrest? What of raising vasopressors? What of the so-called temporary addition of dialysis? What of blood products? The persistent controversies can be understood better if one asks two fundamental questions:

1. *From what do patients die?* Do they die from their disease? Or because dialysis was discontinued? Or because they received morphine to relieve respiratory distress? Or because they were not resuscitated? Or because life supports were discontinued in the face of irreversible organ failures?
2. *To whom does life belong?* If it belongs to a higher being, many believers feel that it is held in trust and cannot be shortened. The most extreme of these vitalists feel that all available methods of treatment must be used, even when survival would be miraculous. However, if it belongs to the individual, or even if the family, tribe, or state has a claim, decisions can be made according to personal preferences, social mores, or legal criteria.

Table 27.9–16.
Range of Terminal Care Treatment Decisions

Never diminish comfort measures unless they have become uncomfortable!
Stopping
Curative treatment
Invasive procedures and treatments
Withholding
All diagnostic procedures
Blood products
Antibiotics

Total parenteral nutrition or dialysis, or both
Cardiopulmonary resuscitation
Invoking the double effect, if indicated
Withdrawing
The previously mentioned treatments
Life supports
Tube feedings
Intravenous hydration
Oral feedings

The fundamental incompatibility between these views is illustrated by looking at the range of end-of-life measures in Table 27.9–16. The vitalists may object to every measure, whereas those who believe that life belongs to the bearer may view them all as a matter of individual choice. A dedicated vitalist will believe that death resulted from the immediate action, such as discontinuation of dialysis, rather than from the irreversible disease process. He or she would refer to withholding treatment as *passive euthanasia* and to withdrawing it as *active euthanasia*. On the other hand, the nonvitalist would be comfortable with the use of slowly escalating doses of morphine to relieve terminal distress, even if it hastens death.

As the debate has become more heated, loaded terms have been used, such as *slow euthanasia*, which has been used to describe opiates for effective pain control, or *passive euthanasia*, which has been used to describe the withholding of invasive life-sustaining treatments. Thus, the vitalist position blurs the distinction between assisted death practices and good palliative care as it is currently understood. On the other side, dedicated proponents of assisted death have supported actions and policies that cause uneasiness even in those who are sympathetic to their point of view. Because physician aid in dying is being legalized in a growing numbers of jurisdictions, it is crucial to avoid polarization of a debate in which good palliative care is ultimately the casualty.

Physician Aid in Dying

The nature of patients' desire for death, outlined in Table 27.9–17, ranges from transient passive thoughts of suicide, through the suicidal preoccupations of major depressive episodes, to active requests for physician aid in dying. Physician aid in dying ostensibly addresses only the last category, patients whose pain is uncontrolled or whose suffering is intolerable and who ask for help in ending their lives. However, it has been demonstrated that, indeed, many patients requesting help are clinically depressed and must be given the benefit of a consultation and attempt at treatment.

Physician aid in dying includes physician-assisted suicide (the provision of enough medication to enable a patient to commit suicide) and active voluntary euthanasia (bringing about the death of another person with

their consent). Involuntary euthanasia (bringing about the death of another person without their consent) is murder in the United States.

Table 27.9–17.
Desire for Death in Terminally Ill Patients

Transient passive thoughts of death
Easily dissolved with improved quality of life
"It would be OK if I did not wake up."
Suicidal ideation in the context of major depression
Persistent, obsessive thoughts of dying and desire for death, associated with excessive guilt, feelings of self-hatred, and worthlessness
Not necessarily associated with a plan
Not responsive to medical and environmental measures
Request for physician aid in dying: assisted suicide and euthanasia
Commonly associated with fear and with hopelessness, with and without depression
Often has a strong component of more or less conscious anger

Physician-assisted suicide and euthanasia were legalized in the Netherlands after a long period of having been openly tolerated and regulated. Opponents have pointed to longitudinal studies that show the incidence of euthanasia rising in proportion to physician-assisted suicide and found some cases of euthanasia of unconscious patients in which overt patient requests were not documented, leaving one to wonder about the dangers of the so-called slippery slope.

In the United States, two federal Circuit Courts of Appeals supported assisted suicide, one on the basis of equal protection under the law, equating it to the available right to discontinue life-preserving treatment, and the other on the basis of liberty and privacy rights. Both decisions were halted by injunctions and were reviewed by the US Supreme Court. In June 1997 the Court refused to affirm a constitutional right to assisted suicide, but stated a need for further social exploration. In particular, Justice Sandra Day O'Connor emphasized a positive right to excellent pain and symptom management, giving hope to many that palliative care would gain impetus, as indeed it has since then.

While euthanasia remains unacceptable, physician-assisted suicide was legalized in Oregon in 1997 after angry voters and attorneys general faced off over this issue. In response to physicians' reasonable fear of harassment, the California Board of Professional Medical Conduct was one of the first states to publish guidelines to protect physicians. Other professional bodies have done the same. Subsequent history in Oregon has not borne out the fears of a slippery slope. As of this writing, there are at least five states that have legalized physician-assisted suicide, with a large number on the way.

However, it is important for physicians to be aware of psychiatric and

medical issues which are sometimes overlooked. Delirium, for example, is associated with impulsivity and must be treated. Also at risk are patients with poor pain control, a sense of helplessness and loss of control, and a previous psychiatric diagnosis. Hopelessness as part of a pessimistic coping style has been found to be a better predictor of suicidal wishes than depression without hopelessness. To complicate matters, the *will to live*, as measured with a visual analog scale, was found to fluctuate rapidly, correlating with anxiety, depression, shortness of breath, and other painful or frightening symptoms. Treatment of pain and other symptoms often diminishes or dissolves the desire for help in dying. Provision of social and spiritual support is also powerful. Understanding these issues is more important to the patient's welfare than public hand wringing about the slippery slope.

Voluntary Refusal of Food and Water

An Oregon report from Dr. Linda Ganzini found that almost one in three nurses had cared for patients who died a peaceful death after refusing food and water. Compared to the assisted suicide group, these patients were less likely to have cancer, in less acute pain, less preoccupied with control, a decade older, and more ready to die. This interesting study highlights a mode of death that is universally available to any competent patient. However, many patients requesting aid in dying feel out-of-control and afraid to die in acute pain and solitude. They are anguished and yearn for loving care to heal both pain and suffering. Lacking that, they demand an ultimate act of nurturance that heals by bringing oblivion. By placing a heavy burden on others, they may express both yearning and resentment. Others, however, plan ahead, discussing those plans with loved ones, and finding physicians with whom they negotiate their request in a cooperative mode. Additionally, one should not ignore cases in which patients are driven by the realistic fear of losing the ability to express their wishes.

Caregiver Responses

In the same study, nurses stated that all staff members found it difficult to care for patients who elect to forgo fluids and nutrition. Similarly, Oregon physicians who had responded to a request for physician-assisted suicide described feeling very uncomfortable. They had discussed technical issues with colleagues but kept their conflicted feelings to themselves. They usually concluded that they had no regrets and believed that they had become more comfortable in dealing with end-of-life issues. Some may wonder whether this represents improved palliative care skills or desensitization to death. In any case, the development of palliative care has given physicians an option to deal with such patients to their professional and ethical satisfaction.

Societal Implications

Proponents of legalization emphasize dignity and autonomy, pointing to Dutch studies that name this as the most frequent reason for help-in-dying requests. Opponents point to the lack of good pain management and other palliative care, citing American and Canadian studies that found depression and poor symptom control as the major factors underlying patient suicidal tendency. Given severe retrenchment of health care expenditures, with politicians, insurance companies, and managed care corporations setting the policies, there is cause for concern even among individuals who have no philosophical or moral objection to assisted suicide. African Americans and other minorities—groups that have historically known inferior medical treatment or inferior representation in medical studies—are far more opposed to legalizing assisted suicide than the rest of the population. Associations of disabled individuals also oppose it because they believe that many of them would not be alive had it been acceptable to terminate lives deemed not worth living. All of these groups insist that moral opprobrium and legal sanctions are essential to hold dark human tendencies in check. Independent of religious or philosophical convictions, historical and contemporary examples demonstrate that they cannot be ignored. But the authors do experience relief at the wholehearted adoption of palliative care and the gentle invasion of humanizing trends in medicine today.

The Moral Burden of Physicians

Medical training is inherently casuistic and teaches that each case must be decided on its own merits. Primary care physicians and hospitalists have depended on their professional values to guide them and to prevent abuse more often than not. However, this process worked as well as it did because it was buttressed by a social climate that protected individual life. Many of the choices physicians made were transparent. Decisions about withdrawing treatments and about nutrition and hydration were de facto shared by family members and other caregivers, giving social morality an opportunity to mold physician behavior. Medicating dying patients with drugs that invoke the rule of double effect is less controllable and presents doctors with the easiest opportunity for assisted suicide. Contrary to public myths about doctors' enjoyment of power, most physicians have not welcomed it, and are delighted with an unambiguous mandate for good palliative care, free of moral and political shoals.

Given the complex motives of terminally ill patients and the fact that palliative care specialists are still in short supply, it seems dangerous to risk desensitizing medical caregivers in particular, and society in general, to the inadequacies of palliative care and, ultimately, to the complexities of helping a patient die. Yet, the legalization of physician-assisted suicide could lead to such a development. Palliative care specialists *are still in a*

race against time, as they try to educate the medical profession *and* the public about the provision of state-of-the-art care before legalization of aid in dying provides excuses for continued avoidance of faithful attention to the needs of dying patients. The fact that state-of-the-art palliative care for suffering patients is not yet routine and that, even when it is available, can become mired in political controversy, is tragic.

Whatever the outcome of the debate, psychiatrists should remember that there is no room for political or philosophical controversy at the bedside. There is just one reality that they and their medical colleagues should keep uppermost in their minds: a dying human being whose physical, social, emotional, and spiritual needs are all being effectively attended is far less likely to request help to commit suicide, and the family members—given proper help and support—derive a deep sense of peace from having helped their loved one to die feeling loved and secure.

References

Agrawal M, Emanuel EJ. Ethics of phase 1 oncology studies: re-examining the arguments and data. *JAMA*. 2003;290(8):1075–1082.

American Academy of Hospice and Palliative Medicine. Position statement on physician-assisted death. *J Pain Palliat Care Pharmacother*. 2007;21(4):55.

Bakitas M, Ahles TA, Skalla K, Brokaw FC, Byock I. Proxy perspectives regarding end-of-life care for persons with cancer. *Cancer*. 2008;112(8):1854.

Billings JA, Block S. Integrating psychiatry and palliative medicine: the challenges and opportunities. In: Chochinov HM, Breitbart W, eds. *Handbook of Psychiatry in Palliative Medicine*. 2nd ed. New York: Oxford University Press; 2009:15.

Breitbart W, Rosenfeld B. Depression, hopelessness, and desire for hastened death in terminally ill patients with cancer. *JAMA*. 2000;284:2907.

Brown EA. Epidemiology of renal palliative care. *J Palliat Med*. 2007;10(6):1248.

*Cherny N, Fallon M, Kaasa S, Portenoy R, Currow DC. *Oxford Textbook of Palliative Medicine*. New York: Oxford University Press; 2015.

*Chochinov HM. *Dignity Therapy: Final Words for Final Days*. New York: Oxford University Press; 2012.

*Chochinov HM, Breitbart W. *Handbook of Psychiatry in Palliative Medicine*. 2nd ed. New York: Oxford University Press; 2009.

Connolly S, Galvin M, Hardiman O. End-of-life management in patients with amyotrophic lateral sclerosis. *Lancet Neurol*. 2015;14(4):435–442.

Crawley LM. Racial, cultural, and ethnic factors influencing end-of-life care. *J Palliat Med*. 2005;8(Suppl 1):58–69.

Ferrell BR, Coyle N, Paice J. *Oxford Textbook of Palliative Nursing*. 4th ed. New York: Oxford University Press; 2014.

Field MJ, Behrman RE, eds. *When Children Die: Improving Palliative and End-of-Life Care for Children and Their Families*. Committee on Palliative and End-of-Life Care for Children and Their Families. Washington, DC: Institute of Medicine/National Academy Press; 2003.

Fraser J, Harris N, Berringer AJ, Prescott H, Findlay F. Advance care planning in children with life-limiting conditions—the Wishes Document. *Arch Dis Child*. 2010;95:79.

Fried TR, Bradley EH, Towle VR, Allore H. Understanding the treatment preferences of seriously ill patients. *N Engl J Med*. 2002;346:1061–1066.

Fromme E, Billings JA. Care of the dying doctor: on the other end of the stethoscope. *JAMA*. 2003;290:2048–2055.

Ganzini L, Dobscha SK, Heintz RT, Press N. Oregon physicians' perceptions of patients who request assisted suicide and their families. *J Palliat Med*. 2003;6:381.

Ganzini L, Goy ER, Miller LL, Harvath TA, Jackson A. Nurses' experience with hospice patients who deliberately hasten death by voluntary refusal of food and fluids. *N Engl J Med*. 2003;349:359.

*Gawande A. *Being Mortal: Medicine and What Matters in the End*. New York: Henry Holt; 2014.

Hammes BJ, Briggs L. *Respecting Choices: Building a Systems Approach to Advance Care Planning*. LaCrosse, WI: Gundersen Health Systems; 2011.

Institute of Medicine. *Dying in America: Improving Quality and Honoring Individual Preferences near the End of Life*. Washington, DC: National Academies Press; 2014.

*Kearney JA, Lederberg MS. Ethical issues in pediatric oncology. In: Wiener LS, Pao M, Kazak AE, Kupst MJ, Patenaude AF, eds. *Pediatric Psycho-Oncology: A Quick Reference on the Psychosocial Dimensions of Cancer Symptom Management*. New York: Oxford University Press; 2015:313–324.

Kissane DW, Bloch S. *Family Focused Grief Therapy: A Model of Family-Centered Care during Palliative Care and Bereavement*. Philadelphia, PA: Open University Press; 2002.

Lo B, Ruston D, Kates LW, Arnold RM, Cohen CB. Discussing religious and spiritual issues at the end of life: a practical guide for physicians. *JAMA*. 2002;287:749–754.

McCurdy D. Ethical spiritual care at the end of life. *Am J Nurs*. 2008;108(5):11.

Meier DE, Isaacs SL, Hughes R. *Palliative Care: Transforming the Care of Serious Illness*. San Francisco, CA: Jossey-Bass; 2010.

Morita T, Hirai K, Sakaguchi Y, Tsuneto S, Shima Y. Family-perceived distress from delirium-related symptoms in terminally ill cancer patients. *Psychosomatics*. 2004;45:107–113.

Old JL. Discussing end-of-life care with your patients. *Fam Pract Manag*. 2008;15(3):18.

*Postovsky S, Ben Arush MW. Care of a child dying of cancer: the role of the palliative care team in pediatric oncology. *Pediatr Hematol Oncol*. 2004;21:67.

Qaseem A, Snow V, Shekelle P, Casey DE Jr, Cross JT Jr. Evidence-based interventions to improve the palliative care of pain, dyspnea, and depression at the end of life: a clinical practice guideline from the American College of Physicians. *Ann Intern Med*. 2008;148(2):141.

Turkel SB, Braslow K, Tavare CJ, Trzepacz PT. The delirium rating scale in children and adolescents. *Psychosomatics*. 2003;44:126.

*Zadeh S, Pao M, Wiener L. Opening end-of-life discussions: how to introduce Voicing My CHOiCES™, an advance care planning guide for adolescents and young adults. *Palliat Support Care*. 2014;13:1–9.

▲ 27.10 Stress and Psychiatry

PAUL S. NESTADT, M.D., TRACI J. SPEED, M.D., PH.D., FRANCIS J. KEEFE, PH.D., AND JOEL E. DIMSDALE, M.D.

Prior to the 1930s, the term "stress" was primarily used in engineering and simply referred to the physical strain on a material object. The modern psychological construct of stress was formulated in the 20th century but has rapidly become a fundamental concept of psychiatry. Patients will readily describe their chronic and acute stressors without much prompting and much can be learned from their particular way of recounting, prioritizing, and reportedly overcoming (or succumbing to) each instance of stress. All people share an existential vulnerability to life's difficulties; as the classical Greeks would have it, "we are all but a heartbeat away from disaster." Stress has been invoked as a cause of major psychopathology, a precipitator or trigger of psychiatric illness, and a contributor to considerable mental anguish. Similar roles have been cast for stress due to illness throughout medicine. Many of the suggested links are intuitively appealing. After all, who has not experienced somatic distress in the face of unpleasant events and thoughts? Such associations between stress and diverse symptoms are obvious. Whether stress causes an illness is much more difficult to demonstrate.

The slipperiness of the term stress is both an advantage and a disadvantage. It is an advantageous term because it is all-embracing and allows a synthesis of many aspects of life's experience. It is disadvantageous because its referents are diverse. Stress refers both to a subject and a predicate, an event and the consequences of that event. For clarity, it is probably wise to eschew use of the word stress altogether and to use different terms for its various meanings. Thus, it is much clearer to refer to "stressful" events or "stressors" and the resulting "strain" or "stress responses" incurred. Throughout this section, such terms will be utilized, but occasionally the broader term of stress will be utilized.

HISTORY

The term *stress* was coined by Hans Selye (1907–1982). He observed that many highly diverse ways of perturbing the organism resulted in common physiological responses. The concept was unusual because the prevailing ideology of medical research emphasized specificity, that is, that specific pathogens elicited unique pathological responses. Selye found that highly diverse stimuli seemed to result in common changes. From animal studies he reported that there is a "general adaptation system," a systemic well-orchestrated adaptive response to diverse stressors. Selye invoked the

adrenocortical system as the crucial responder to stressful stimulation. He observed that any novelty or perturbation was associated with an elevation of adrenocortical activity, at least transiently, and that if stressors were unremitting, diverse pathological changes would be evident.

Like all pioneers, Selye was guilty of oversimplification, a weakness that has become abundantly clear with the accumulation of additional empirical data. Not all organisms respond to all perturbations with the general adaptation system. However, his insights were terribly important in shaping the field. As a result of his prolific scholarship, the field of psychology has been blessed (or cursed) with the term stress, which has become so popular that it is used constantly, even when researchers would prefer more precision. Stress is a term much like neurosis or hysteria, ripe with years of reference, conveying both information and vagueness.

The other groundbreaker in the field of stress research was Walter Cannon (1871–1945). Although he made major contributions in gastroenterology and roentgenology, certainly his long-abiding interest was in the physiology of emotion and the autonomic nervous system. If Selye is known for his explorations of the adrenocortical system, Cannon is acclaimed for his methodical investigations of the other great pathway of stressful responses, the sympathetic nervous system. Cannon also focused on more immediate or short-term responses to stressors.

Whereas the term stress and the accumulating empirical research reflect the 20th century and the intellectual progeny of Selye and Cannon, there are earlier referents for the concept. For centuries physicians and patients have made the association between adverse life events and illness. In law, for example, there is a long history of tort cases alleging that emotional arousal secondary to a stressful event causes miscarriage. The links between emotion and sudden cardiac death have been noted in sources as diverse as the Bible, anthropology, and clinical experience. These types of links were repeatedly observed by crucial figures in the history of medicine. William Harvey in the 17th century and William Osler in the 19th century frequently alluded to the relationship between adverse life events and illness onset. Many contemporary cultures regard illness as the outcome of being “out of balance” with the environment and its demands, a manifestation of inadequate coping with diverse stressors.

The evidence for this relationship, as will be discussed below, is hardly unanimous. Physicians who see their patients briefly and from a highly specialized perspective are rarely familiar with the nuances of their patients’ lives, and thus it is not surprising that many specialists dismiss this sort of link between stressors and illness as not in their clinical experience. Researchers, on the other hand, criticize the oversimplification of most studies examining stress and illness.

The confusion between stress as “cause” versus “trigger” of illness has hindered progress. The allegation that a stressor caused an illness is

difficult to demonstrate in most human studies, which impose a short-term stressor and examine the resultant physiological changes. Illnesses such as myocardial infarction usually become evident only after a long accumulation of coronary artery disease. As a result, how can one assert that "stress caused a myocardial infarction"? There is, however, an alternative hypothesis that provides very strong evidence indeed for a link between stressors and illness. That argument emphasizes that stressors trigger illness in individuals whose underlying illness is latent or subclinical. This sort of case is demonstrated routinely in psychophysiology and in psychosomatic medicine studies. For instance, studies of the physiological response to stressors in normal individuals demonstrate large amplitude changes in physiological levels of blood pressure, gastric motility, and hormonal levels. When such psychophysiological studies are extended to patient groups, one sees breathtaking evidence for the role of stressors as trigger events in eliciting pathophysiological phenomena. For example, in vulnerable individuals, even discussing emotional conflicts can elicit life-threatening arrhythmias, and exposure to a catastrophic stressor can be accompanied by unique and reversible cardiomyopathy (Fig. 27.10–1). Similar effects of stressors on blood pressure have been repeatedly demonstrated in warfare and in natural disasters. For example, a 54-year-old man whose office was in the vicinity of the World Trade Center had his blood pressure recorded twice a day from September 4 to September 21, 2001. His blood pressure was dramatically increased during the week after the World Trade Center attack (Fig. 27.10–2).

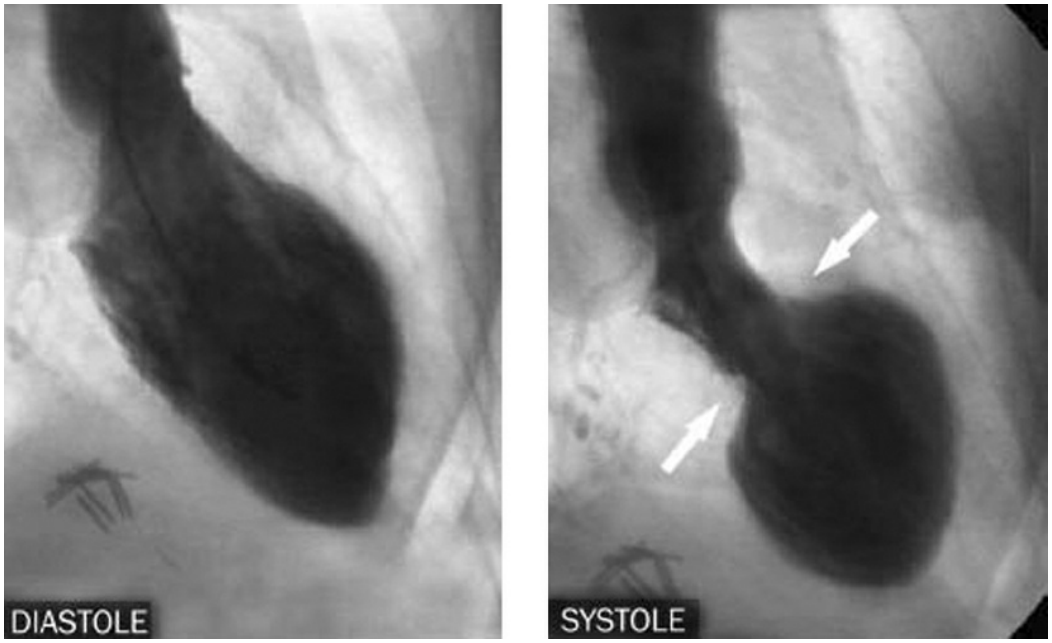

FIGURE 27.10–1. Impact of stress on cardiac function. A reversible cardiomyopathy can be elicited by psychologically stressful events. This is characterized by systolic dysfunction affecting the distal left ventricular chamber. (From Sharkey S, Lesser J, Zenovich A, Maron M, Lindberg J. Acute and reversible cardiomyopathy provoked by stress in women from the United States. *Circulation*. 2005;111:472, with permission.)

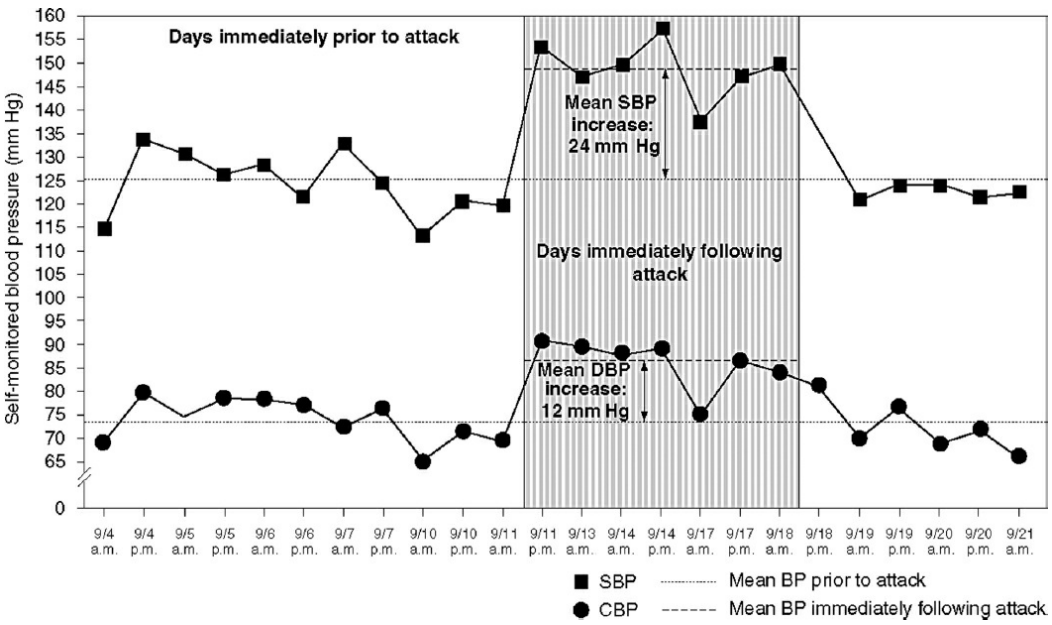

FIGURE 27.10–2. Impact of stress on blood pressure (BP). DBP, diastolic blood pressure; SBP, systolic blood pressure. (From Lipsky S, Pickering T, Gerin E. World Trade Center disaster effect on blood pressure. *Blood Press Monit*. 2002;7:249, with permission.)

VULNERABILITY AND RESILIENCE

In addition to the confusion between stress as a trigger and stress as a response, another key conceptual problem revolves around the issue of vulnerability versus resilience or stress versus coping. Researchers initially focused on the stressor, treating it like a virulent pathogen that would infect all individuals despite their varied life experiences. Thus, the initial stress research primarily focused on the physiological consequences of stressor imposition. The variability in response patterns then motivated investigators to study the subtleties of how people interpreted the stressful challenges. Initially, these interpretations were framed in terms of "defensive" styles, how individuals guard themselves against painful emotional stressors. Subsequently, however, the inquiry broadened and investigators examined individuals' defensive behaviors, information-seeking behaviors, affiliative behaviors, and all-round problem-solving behaviors. This heterogeneous combination is referred to as *coping behavior*.

Genetic and environmental factors are also being studied for their impact on resilience or vulnerability to stress. For instance, a New Zealand study prospectively followed approximately 1,000 subjects aged 3 to 26 in an effort to learn more about stress effects on onset of major depression. The promoter region of the serotonin transporter gene had a powerful effect on vulnerability to stress. Individuals who were homozygous for the short allele of the promoter region of the serotonin transporter gene were more likely to become depressed and suicidal in response to stressors. The role of environmental factors on stress is also being studied. Further

research demonstrated an association between the short allele and increased vulnerability to stress in studies of childhood maltreatment and certain medical conditions.

As stress-related psychiatric disorders are influenced by both genetic predisposition and environmental factors, in the past decade researchers have begun studying epigenetic mechanisms that may account for environmental influences on development of stress-related disorders. Recent research shows that prenatal factors such as stress or mood disorders in mothers are risk factors for psychiatric disorders in offspring.

MEASURING STRESS PSYCHOLOGICALLY

Regardless of whether stress is a cause or trigger and regardless of whether one focuses on stress or the factors that affect resilience to stress, the key step is to recognize and in some way quantify its extent. Doctors have always asked about the stressors in their patients' lives. Obtaining such information is, after all, at the core of history taking. "What types of things aggravate your symptoms? Are you facing any unusual burdens or challenges?" With certain symptoms such as angina, the link between stressor and illness is immediately obvious. However, it appears that many illnesses not traditionally thought of as "stress-sensitive" are aggravated by stressful conditions. As will be discussed later, the mechanisms for this link are complex.

Approximately 80 years ago, Adolf Meyer suggested that psychiatrists employ a "life chart" in evaluating their patients. The idea is elegant in its simplicity. The psychiatrist plots a patient's life on a line, beginning with birth and ending in the present. He or she then pencils in when various important life events have occurred and also indicates the onset or worsening of health problems. The life chart turns out to be a very effective way of eliciting a clinical history. Sometimes, recurrent patterns of association between life events and illness episodes are suggested (Fig. 27.10–3).

The next step in the evolving field of life events was developed by the fruitful collaboration of Thomas Holmes and Richard Rahe. They felt that some kind of list of commonly encountered stressors would be more useful than taking an unstructured history. After considerable research they developed a list of roughly 40 life events ranging in severity from death of a spouse to receipt of a traffic ticket. The checklist or Schedule of Recent Events can be self-administered and requires very little time to complete. It is noteworthy that the form tallies both positive and negative life events. The underlying assumption is that all life events, be they positive or negative, impose a demand for adaptation, and that such demands are stressful. While at first blush this may seem counterintuitive, positive life events are not totally benign; getting married brings with it a host of stressors. Another characteristic of the Holmes–Rahe approach is that the

items are weighted in terms of their adaptational demand. Most people rank order stressors similarly, and thus it is possible to assign a specific "number" of life change units to the occurrence of each specific stressor. This technique has been used in hundreds of studies and has demonstrated that numerous illnesses follow on the heels of increases in life stress. It is not that the high-scoring individual is at risk, but rather that the individual with an *increase* in life stress units is at risk for any number of health consequences.

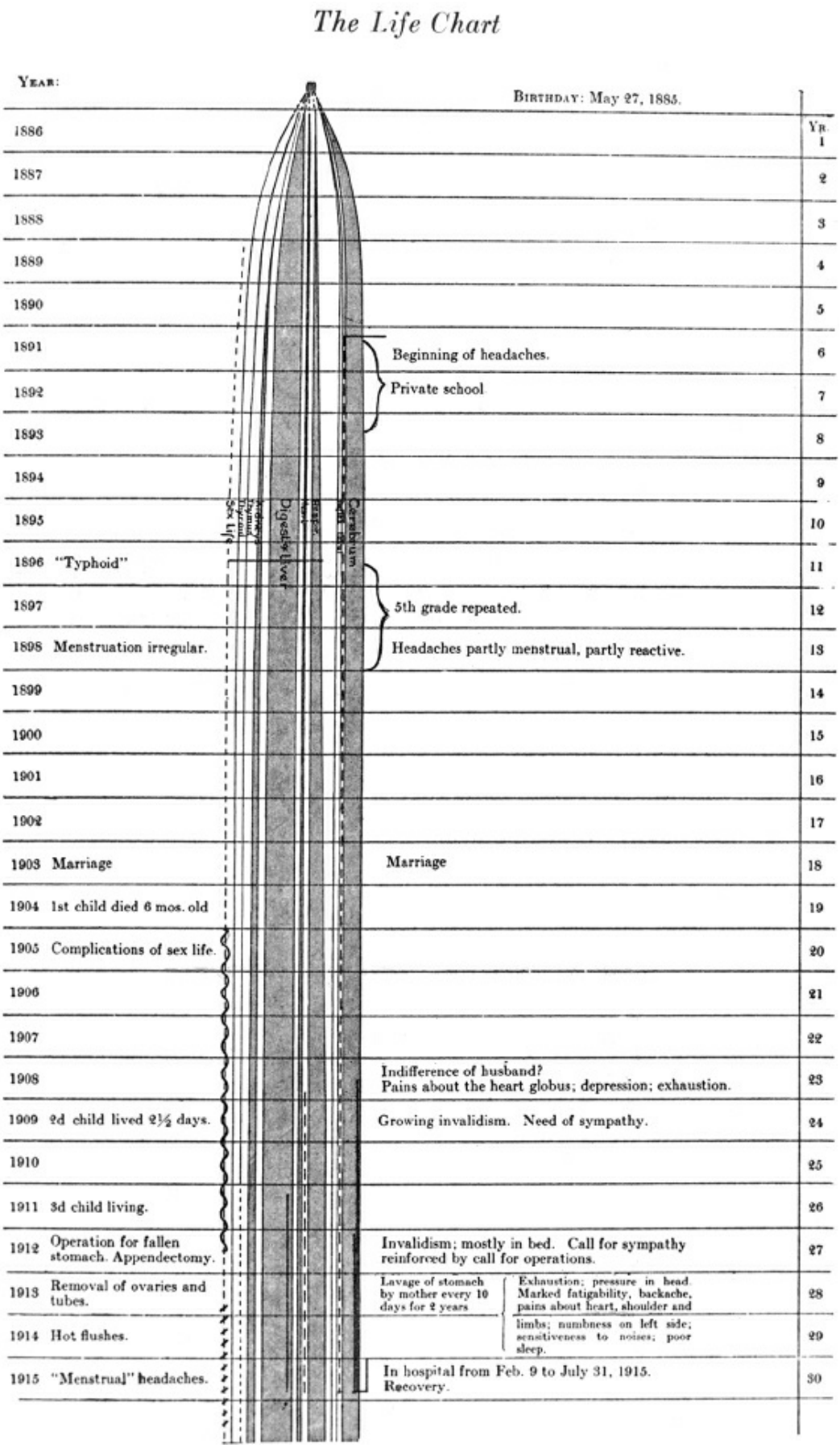

FIGURE 27.10–3. Line drawing of the life chart by Adolf Meyer. (From *Contributions to*

Medical and Biological Research, Dedicated to Sir William Osler, bart. M.D., F.R.S., in Honour of his Seventieth Birthday, June 12, 1919, by His pupils and Co-workers. Paul B. Hoeber, New York, 1919. Credit to Wellcome Library, London. Creative Commons License.)

There are many criticisms of this approach and, of course, alternative ways of measuring stressful life events. Some investigators feel that negative life events carry with them *far* greater anguish and adaptational pressure, and that the positive life events may be ignored. Some question the utility of adding all the life change units together and coming up with one number to represent the total stressful exposure of the individual. Whereas the Holmes–Rahe approach may be extremely useful for epidemiological research, for clinical use, it may be more helpful to have subscales that list, for example, the stressor scores associated with job pressure, home pressure, and so forth into separate categories. Some theorists feel that inventories of relatively rarely occurring stressful events are of limited utility and that more is to be gained by measuring the accumulation of the more common microstressors that individuals encounter frequently. Other investigators doubt the validity of a universal weighting of stressful events and favor allowing individuals to apply their own idiosyncratic assessment of emotional travail elicited by events. Finally, common to all scales that measure stressful events is the problem of retrospective distortion. Patients forget, disavow, deny, and repress—all words that reflect a splitting off of accurate recall. As a result, when patients are asked to retrospectively report life events, they tend to forget events that are shrouded in the past. On the other hand, patients' beliefs sometime cause them to "look harder" for events if they have sustained an illness they feel is caused by stress.

SELECTED PHYSIOLOGICAL RESPONSES TO STRESSORS

The body's reaction to a stressor (be it real, symbolic, or imagined) is to initiate a set of responses that seek to diminish the impact of the stressor. A stressor disrupts an organism's equilibrium, and the *stress response* consists of the initiation of physiological adjustments that seek to react to the stressor, bring about an adaptive response and restoring homeostasis. Much is known about the physiological response to *acute stress*, and considerably less is known about the response to *chronic stress*. This represents a major gap in physicians' ability to relate the effects of stress to the occurrence of psychiatric illness in humans. Many stressors occur over a prolonged period of time or have long-lasting repercussions (e.g., loss of a spouse is followed by months or years of loneliness; a violent sexual assault is followed by years of apprehension and worry about walking outside at night). Animal models that can capture this reverberation or "echo" of stress are limited. Still, what is known about the neuroendocrine and immune response to laboratory stressors is expanding and will ultimately

inform the understanding of why and how stress has an adverse effect on humans.

Neurotransmitter Responses to Stress

Stressors activate a release of catecholamines from the autonomic nervous system including noradrenaline, dopamine, and serotonin. The stressor type influences the brain regions and mediators that engage in the adaptive response. Physical stressors (e.g., blood loss, trauma, hypothermia) recruit the brainstem and hypothalamic brain regions. Psychological stressors (e.g., embarrassment, work-related stress) engage brain regions involved in emotion (amygdala), learning and memory (hippocampus), and decision making (prefrontal cortex). The duration and timing of stress also affects neuronal pathways differently. Chronic stress has enduring effects on neuronal firing and gene activation. Recent animal studies demonstrate that varying isoform expression of tryptophan hydroxylase, the rate-limiting enzyme in serotonin biosynthesis, is affected by adverse early life events. These changes may have implications in the pathophysiology of stress-related disorders, including major depression and anxiety.

Over the past two decades, it has become clear that amino acid and peptidergic neurotransmitters are also intricately involved in the stress response. For example, corticotropin-releasing factor (CRH).

Endocrine Responses to Stress

Endocrine response to stress is secreted from the hypothalamus into the hypophysial-pituitary-portal system in response to stress. CRH acts at the anterior pituitary to trigger release of adrenocorticotrophic hormone (ACTH). Once ACTH is released, it acts at the adrenal cortex to stimulate the synthesis and release of glucocorticoids. Glucocorticoids themselves have myriad effects within the body, but their actions can be summarized in the very short term as promoting energy utilization, increasing cardiovascular activity (in the service of the flight-or-fight response), and inhibiting functions such as growth, reproduction, and immunity.

This HPA axis is subject to tight negative feedback control by its own end products (i.e., ACTH and cortisol) at multiple levels, including the anterior pituitary, the hypothalamus, and the hippocampus. It is now known that the HPA axis is also influenced by multiple extrinsic systems, which should not be surprising given the coordinated nature of the stress response. For example, in addition to CRH, numerous secretagogues (i.e., substances that elicit ACTH release) exist that can bypass CRH release and act directly to initiate the glucocorticoid cascade. Examples of such secretagogues include catecholamines, vasopressin, oxytocin, and orexin. Interestingly, different stressors (e.g., cold stress versus hypotension) trigger different patterns of secretagogue release, again demonstrating that

the notion of a uniform stress response to a generic stressor is an oversimplification.

A relevant feature of the HPA axis response to stress is that it varies depending on whether the organism (most of these data come from rats) is exposed to the stressor on a single occasion, more than once, or over a prolonged period of time. Interestingly, it has not yet been possible to rely on a model that can predict whether "sensitization" (an enhanced response to the same stress over time) or "desensitization" (a reduced response to the same stress over time) will occur for a given stressor. Nonetheless, the fact that the stress response can change dramatically depending on the frequency and duration of exposure highlights the necessity of specifying these elements when trying to understand the role of stress in clinical psychiatric illness.

PSYCHONEUROIMMUNOLOGY

Researchers have come to appreciate that the immune system is also a key player in stress physiology. There are numerous bidirectional interactions between brain, behavior, and the immune system, and these interactions are studied in the field of psychoneuroimmunology. The immune system provides the body's defense against invading external pathogens such as viruses and bacteria and from abnormal internal cells such as tumors. Innate immunity refers to the body's resistance to pathogens; this form of immunity operates in a nonspecific way without recognition of the different nature of various pathogens. Macrophages and granulocytes are nonspecific immune cells that induce immune activation by the release of proinflammatory cytokines (TNF-α; IL-6) and react to tissue damage by consuming debris and invading organisms. Specific immunity is acquired in response to the identification of non–self-molecules called antigens. Each T cell or B cell is genetically programmed to attack a specific target by secreting antibodies (B cell) or by killing cells of the body that harbor a virus (T cell). Both innate and specific immunity is orchestrated by the release of interleukins or cytokines from immune cells, and this cytokine network aids in the differentiation of the immune response and in the coordination of its magnitude and duration. However, the immunoregulatory processes cannot be fully understood without taking into account the internal and external milieu in which innate and specific immune responses occur, as well as the interactions of the brain, autonomic nervous system, and neuroendocrine axis with the immune system (Fig. 27.10–4).

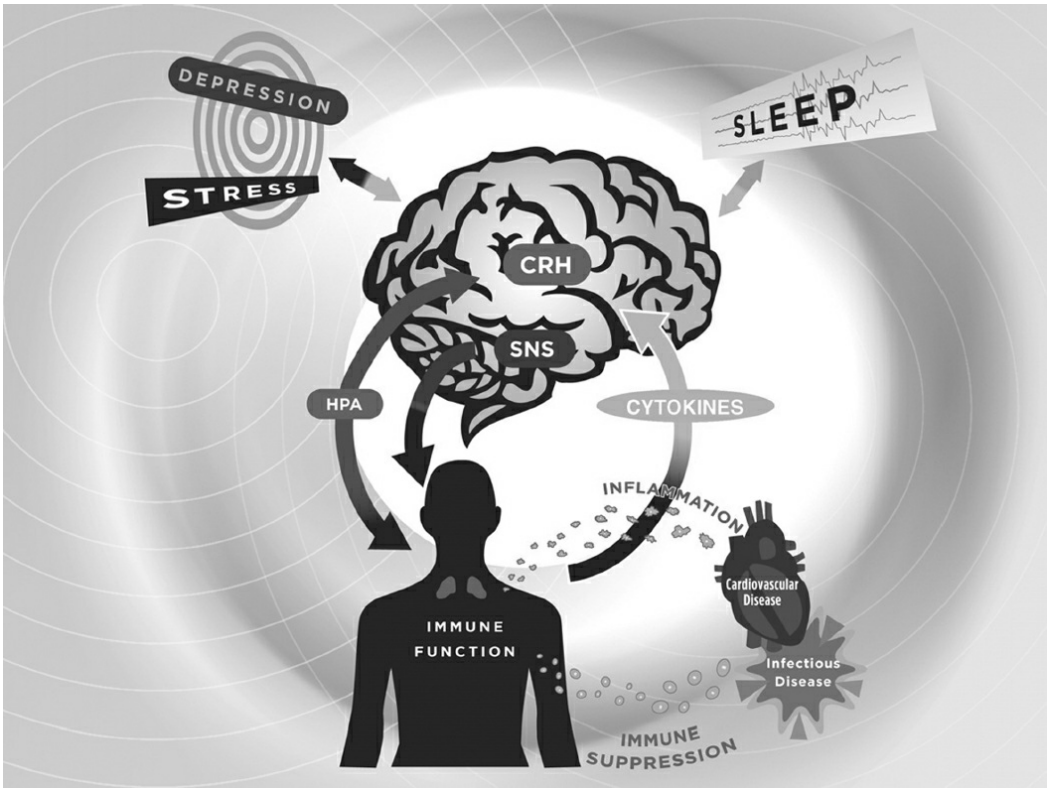

FIGURE 27.10–4. Model of human psychoneuroimmunology interactions and clinical implications. The model depicts the bidirectional interactions between the brain and the immune system, and the clinical implications of immune alterations due to stress, depression, or sleep disturbance. CRH, corticotropin-releasing factor; SNS, sympathetic nervous system; HPA, hypothalamic–pituitary–adrenal axis. (From Irwin MR. Human psychoneuroimmunology: 20 years of discovery. *Brain Behav Immun*. 2008;22(2):129, with permission.)

Biological Connections between the CNS and Immune System

Autonomic Nervous System. The CNS and the immune system are linked by two major physiological systems, the HPA and the autonomic nervous system composed of sympathetic and parasympathetic branches. The sympathetic nervous system (SNS) is a network of nerve cells running from the brainstem down the spinal cord and out into the body to contact a wide variety of organs including organs where the immune system cells develop and respond to pathogens. Sympathetic release of norepinephrine and neuropeptide Y, together with receptor binding of these neurotransmitters by immune cells, serves as the signal in this hard-wired connection between the brain and the immune system. In addition, sympathetic nerves penetrate into the adrenal gland and cause the release of epinephrine into the bloodstream, which circulates to immune cells as another sympathetic regulatory signal.

Under both laboratory and naturalistic conditions, sympathetic activation suppresses the activity of diverse populations of immune cells. Sympathetic activation can also enhance other aspects of the immune response such as the production of antibodies by B cells and the ability of

macrophages to release cytokines and thereby signal the presence of a pathogen. Additional studies indicate that sympathetic activation can also shunt some immune system cells out of circulating blood and into lymphoid organs (e.g., spleen), while recruiting other types of immune cells into circulation (e.g., natural killer [NK] cells). In general, SNS activation can reduce the immune system's ability to destroy pathogens that live inside cells (e.g., viruses) via decreases of innate and cellular immune responses, while sparing or enhancing the humoral immune response to pathogens that live outside cells (e.g., bacteria). Together, these observations are a cornerstone for understanding fundamental, neuroanatomic signaling between the autonomic nervous and immune systems.

Neuroendocrine Axis. The other way in which the brain can communicate with the immune system is via the HPA system. Cortisol exerts influence on the actions of various cells involved in an immune response by suppressing the cellular immune response. Indeed, a synthetic analog of cortisol is often used to suppress excessive immune system responses. However, cortisol can also prompt some immune cells to move out from circulating blood into lymphoid organs or peripheral tissues such as the skin. It is important to note that immune cells can produce neuroendocrine peptides (e.g., endorphin, ACTH), which suggests that the brain, neuroendocrine axis, and immune system use the same molecular signals to communicate with one another.

Central Modulation of Immunity. This converging evidence of brain–immune system interactions suggests that brain has a physiological role in the regulation of immunity. Indeed, release of CRH in the brain alters a variety of immune processes, including aspects of innate immunity, cellular immunity, and in vivo measures of antibody production. Peripheral immune measures also change following lesioning of the brain (e.g., hypothalamus) or in response the stimulation of certain brain regions. The brain exerts control over immune cells in lymphoid tissue just as it controls other visceral organs through a complex series of coordinated autonomic and neuroendocrine pathways.

Behavioral and Psychological Influences on Immunity

Given that psychological responses are expressed in neural activity with accompanying changes in neuroendocrine and autonomic function, it is not surprising that behaviors and emotions are capable of altering immunity. One seminal example is classical conditioning of immune responses. Conditioned immune responses have also been found to retard disease progression in an animal model of autoimmune disease. In the clinical setting, immunosuppressive conditioning occurs in cancer patients who

receive chemotherapeutic drugs, and conditioning processes are thought to contribute to biological effects of placebo.

Stress and Immunity in Animals. In animals, acute exposure to stressors produces alterations of immunity that cannot be accounted for by the physical effects of the stress. Psychological components such as conditioned fear and uncontrollability induce altered immunity, and these effects are related in a dose-dependent manner to the severity of the stressor. Almost every component of the immune system has been found to respond to stress.

Acute Stress and Immunity in Humans. Acute laboratory stressors (e.g., mental arithmetic) produce profound and rapid changes in the immune system due to the redistribution of immunoregulatory cells from lymphoid organs such as the spleen into the vascular space. Such acute stressors elicit decreases in cellular immune responses and increases in markers of inflammation (e.g., IL-6), which are thought to be mediated by release of sympathetic neurotransmitters and β-adrenergic receptor activation. Individuals who are elderly or who are undergoing chronic stress show exaggerated immune responses to acute stress and are likely to take longer to recover from the administration of stress, which may put them at increased risk for infectious disease. Finally, among depressed patients, acute psychological stress leads to exaggerated increases of inflammatory cytokine activity and greater activation of the inflammatory signaling pathway, nuclear transcription factor-κB.

Chronic Stress, Depression and Immunity. In contrast to the effects of laboratory stress, chronic or naturalistic stressors such as bereavement or caregiving, as well as depression, are associated with reliable decreases of cellular and innate immunity, along with increases in proinflammatory cytokine activity, due possibly to a downregulation of glucocorticoid receptor signaling. When the stress remits or depressive symptoms resolve, a normalization of natural and cellular immune function occurs.

In the past decade, researchers have found an association between chronic psychological stress and compromised telomere maintenance. Telomeres are repeats of DNA sequences at the ends of chromosomes bound by specialized protective proteins. Telomere shortening is a sign of cell senescence and longevity. In a study of mothers who were caretakers of chronically ill children, there was an association of severity and chronicity of stress with shorter telomere length. Furthermore, in a prospective study of over 200 children, those who suffered two episodes of abuse showed greater telomere shortening. Numerous studies have shown that individuals with major depression have shorter telomere lengths. These findings elucidate one of the many mechanisms of immune dysfunction in

response to chronic stress.

Heterogeneity in the effects of stress and depression on immunity can be accounted for by factors such as age, gender, ethnicity, health behaviors (e.g., smoking, alcohol consumption), and coping. Depressed patients who have comorbid alcohol abuse or tobacco smoking show exaggerated declines of natural and cellular immune responses. Personal characteristics such as coping and personality (e.g., positive affect), which moderate neuroendocrine and sympathetic activity, also contribute to individual differences of immune responses to psychological stress as well as vaccine responses.

Sleep, Cytokines, and Immunity

Disordered sleep and loss of sleep are thought to adversely affect resistance to infectious disease and alter inflammatory disease progression. Studies show that sleep deprivation impairs influenza viral clearance and increases rates of bacteremia. In humans, normal sleep is associated with a redistribution of circulating lymphocyte subsets, increases of NK activity, IL-2, IL-6, and the transsignaling IL-6 receptor agonist (IL-6R), and a relative shift toward Th1 cytokine expression, which is independent of circadian processes. Conversely, sleep deprivation suppresses NK activity and IL-2 production and induces decreases in specific antibody production to infectious challenge (e.g., vaccination to influenza or hepatitis B virus). In contrast, measures of innate immunity and proinflammatory cytokine expression, including cellular and genomic markers of inflammation, are enhanced following sleep loss. In clinical populations who show disordered sleep (e.g., as a result of depression, bereavement, alcoholism), alterations of natural and cellular immune function coincide with sleep loss and disturbances of sleep architecture. Furthermore, decreases of sleep continuity or increases of REM sleep are associated with increases in the nocturnal and daytime expression of IL-6, possibly with consequences for daytime fatigue. Laboratory pain testing methods are now being used to understand individual differences in pro-inflammatory cytokines reactivity. One recent study showed greater levels of pain-evoked IL-6 among older adults with insomnia disorder compared to those without insomnia. Bidirectional actions of cytokines on sleep have also been identified. In animals, cytokines have both somatogenic and inhibitory effects on sleep, depending on the cytokine, plasma level, and circadian phase. In humans, less is known about the sleep regulatory effects of cytokines. Expression of the anti-inflammatory cytokine IL-10 prior to sleep predicts amounts of delta sleep during the nocturnal period. In contrast, peripheral administration of the proinflammatory cytokine IL-6 reduces delta sleep, and nocturnal levels of both IL-6 and TNF temporally correlate with increases of REM sleep, particularly during the late part of the night.

Cytokine Influences on the CNS and Behavior

The nervous and immune systems convey information to each other, ultimately regulating neuroendocrine and autonomic outflow and the immune response. With peripheral immune activation, the inflammatory cytokines, TNF, IL-1, and IL-6, activate the HPA system via central release of CRH. Immune activation leads to changes of peripheral physiology that are similar to a stress response. Coincident with these physiological changes, animals show reductions in activity, exploration of novel objects, social interactions, food and water intake, and sexual behaviors, a response pattern that has become known as *sickness behaviors*. Human studies have shown similar links between peripheral cytokines and behavioral changes. Healthy individuals develop anxiety and depression in response to low-dose cytokines. Even patients with the common cold show sickness behaviors such as fatigue, lack of interest, low mood, and cognitive impairment. Patients with chronic inflammatory diseases such as hepatitis C infection have increased rates of fatigue and depression with associated changes in metabolism of central neurotransmitters in discrete brain regions. Large doses of cytokines, given as immunotherapy for cancer or hepatitis C, frequently worsen depressive symptoms, which can be attenuated by pretreatment with antidepressant medications.

Considerable recent research has focused on how cytokines signal the brain, given their large molecular size and inability to readily cross the blood–brain barrier. In the past decade, researchers have demonstrated that bidirectional communication between the nervous and gastrointestinal systems is another route of mediating response to stress (Fig. 27.10–5). The brain acts on gastrointestinal and immune functions that shape the gut microbiome while gut microbes make neurotransmitters and metabolites that interact with the immune system. Interestingly, gut microbes control leakage through the intestinal lining and the blood–brain barrier. Animal studies show that disruptions to the intestinal microbiome induced anxiety- and depressive-like behaviors in mice. In humans, chronic disorders of the gut are associated with intestinal microbiome dysbiosis and behavioral changes such as depressive and anxiety symptoms.

Clinical Implications of Psychoneuroimmunology and the Cytokine–Brain Link

Increasing evidence suggests that abnormalities in cytokine expression and regulation that occur during depression or stress might exacerbate the course of many chronic diseases. The following paragraphs provide several pertinent examples in which cytokines and psychoneuroimmunology processes are thought to contribute to disease processes.

Cardiovascular Disease. Depression is associated with increased risk of coronary artery disease, and increased morbidity and mortality after

myocardial infarction, congestive heart failure, or mitral valve replacement. Atherosclerosis is now thought to be an inflammatory process that involves a series of steps, each of which can be impacted by stress or depression. Activated macrophages within the vasculature secrete proinflammatory cytokines such as IL-1, IL-6, and TNF, which in turn leads to expression of cellular adhesion molecules. With recruitment of immune cells to the vascular endothelium and the release of inflammatory cytokines, endothelial activation occurs, facilitating further binding of immune cells. Importantly, psychological and physical stressors lead to an increased release of proinflammatory cytokines and expression of adhesion molecules that tether and bind immune cells to the vascular endothelium. Depressed mood, sleep disturbance, inflammation, and arterial repair might form a vicious cycle that promotes cardiovascular disease. In a study of 1,794 Canadians, depression and elevations in inflammatory markers predicted an increased risk of coronary artery disease. In another study of patients with coronary heart disease, severity of depression was associated with elevated CRP and IL-6. This link was driven by behavioral health factors, including physical inactivity, smoking, and obesity suggesting other mechanisms including epigenetic factors may promote the link between depression and cardiovascular disease.

Infectious Disease Risk. Compelling evidence has shown that inescapable stress, a putative animal model of depression, increases susceptibility to viral diseases such as herpes simplex and infection via alterations in immune function. In humans, prospective epidemiological studies and experimental viral challenge studies show that persons reporting more psychological stress have both a higher incidence and a greater severity of certain infectious illness such as Epstein–Barr virus infections and the common cold. Moreover, experimental vaccinations have probed the disease-specific and integrated in vivo action of the immune system in relation to psychological stress.

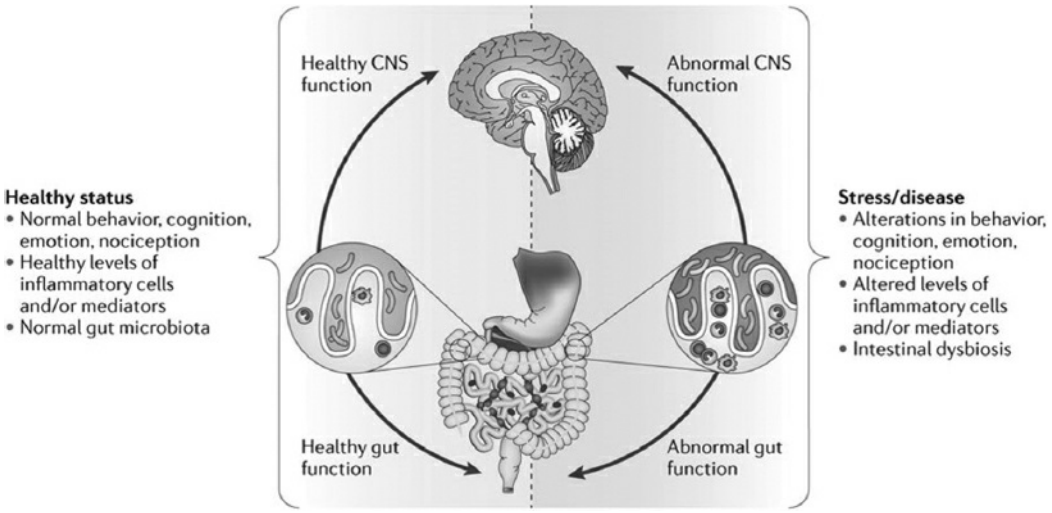

FIGURE 27.10–5. Model of the gut–brain axis. Normal gut physiology contributes to appropriate signaling to the central nervous system, and helps regulate healthy physiological interactions. Intestinal dysbiosis alters gut physiology and leads to adverse consequences for central nervous system functions and results in disease states (*shaded area*). CNS, central nervous system. (Adapted from Cryan, J, Dinan T. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat Rev Neurosci*. 2012;13:701.)

Atopic Dermatitis. Atopic dermatitis is a chronic and relapsing skin disease characterized by impaired epidermal skin barrier function, inflammatory infiltrates, and excessive pruritus. Psychological stress contributes to the course of atopic dermatitis through its direct and indirect effects on immune response. Skin biopsies from patients with atopic dermatitis have elevated levels of Th2 cells, cytokines, and eosinophils with elevated IgE production in response to stress. Patients with atopic dermatitis and high trait anxiety also have elevated levels of IgE. Numerous studies demonstrate that psychological stress aggravates pruritus. Both antidepressants and psychological interventions including CBT improve pruritus and psychogenic itch.

Stress, Depression, and Rheumatoid Arthritis. In a negative feedback loop, proinflammatory cytokines stimulate the HPA axis, which results in the secretion of glucocorticoids, which in turn suppress the immune response. However, in autoimmune disorders such as rheumatoid arthritis, it is thought that the counterregulatory glucocorticoid response is not fully achieved. Rheumatoid arthritis patients show a relative hypofunctioning of the HPA axis despite the degree of inflammation. Data suggest that chronic stress provokes increases in the production of the proinflammatory cytokine IL-6, which correlates with symptoms of disease including fatigue, pain, and functional limitations. Moreover, the presence of depression in rheumatoid arthritis patients undergoing stress is associated with exaggerated increases of IL-6, a biomarker predictive of disease progression. Tocilizumab (TCZ), an interleukin 6 (IL-6) receptor blocker used in treatment of rheumatoid arthritis, was noted by patients to improve pain, quality of life, physical functioning, and fatigue.

Cancer and Psychoneuroimmunology. In recent history, cancer has been reconceptualized from an acute disease model to a chronic disease with interest shifting towards intraindividual differences that affect clinical outcomes. In particular, researchers are beginning to conceptualize the psychoneuroimmunological pathways that affect the quality of life of cancer patients. Elevated levels of cytokines have been associated with cancer stage in melanoma, breast cancer, and bone sarcoma, as well as posttreatment fatigue in breast cancer survivors. Recent data suggests that memory deficits in newly diagnosed breast cancer patients are related to an

activated immunological response. Stressful life events and depression are associated with higher mortality rates across a variety of cancers. In a study of ovarian cancer patients, those with a greater sense of social attachment had a lower likelihood of death. Furthermore, perceived social isolation or loneliness predicted morbidity and mortality risk across different ages. While data remains limited on psychoneuroimmunology's relationship to cancer, there is continued interest in the discovery of immune mechanisms to improve clinical cancer care.

STRESS AND PSYCHIATRIC ILLNESS

Historical Conceptualization

Mental disorders have been ascribed to an imbalance in bodily humors, to the influence of external spiritual or other supernatural forces, and to moral or somatic deficiency. In the early-to-mid-1800s, a school of thought emerged, led by Philippe Pinel (1745–1826) of the Salpêtrière in Paris and Amariah Brigham (1798–1849) in America, which held that the expression of mental illness was affected by life circumstances and, more broadly, by societal factors. These various influences converge in the biopsychosocial model proposed by Meyer. This orientation held that an individual's response to stress is modified by a number of intrinsic factors (e.g., genes, premorbid personality) and extrinsic factors (e.g., social support). The model incorporated individual temperamental and experiential characteristics, such as potential vulnerability (or resiliency) factors or stressful life events, as initiating or exacerbating factors, and a support network as a potentially modifying factor for the occurrence of mental illness. For the most part, this multidetermined model of mental illness dominates modern psychiatric thinking.

It is interesting to see how stress is handled in the diagnostic nomenclature of the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). This classification does not emphasize the biopsychosocial model of the nature and genesis of psychiatric disorders because of the decision to have the DSM-5 (similar to previous DSM editions) expound a *descriptive*, phenomenological typology, one that would be least governed by theoretical or other subjective interpretations. Thus, in DSM-5, "major depression" is "major depression," whether it appears in the context of job loss, divorce, war trauma, or for no apparent reason. Historically, there have been inconsistencies in this approach. Until the 2013 release of the DSM-5, major depression was not considered major depression if it occurred within 2 months of the death of a loved one. Such grief was somehow considered mutually exclusive of major depression, making it difficult to diagnose depression during bereavement, barring extreme symptoms such as suicidality or psychosis. The newest edition removes this bereavement exemption and replaces it with

recommendations to consider true major depression even in the context of significant loss, if the reaction is inappropriate to the stressor. In a welcome call for thoughtfulness, the text further allows the reader to exercise clinical judgment in making the assessment of appropriateness of reactions.

Stress can clearly contribute to pathology, so which diagnostic options are afforded by the DSM-5 to explain a depressive syndrome following a major stressor? For example, what if an individual becomes depressed following the divorce of their parents? According to DSM-5, if such a person develops a depressive syndrome, it might be considered an adjustment disorder (if short-lived) but, if it is persistent, then a diagnosis of major depression applies. Following a traumatic stressor (e.g., sudden loss of a spouse in a motor vehicle collision), might one better diagnose PTSD? Not necessarily. Although PTSD is another exception to the DSM-5 rule that the nomenclature should be atheoretical (it is one of the very few disorders in DSM-5 for which the occurrence of a stressful life event is pivotal for diagnosis), it has its own characteristic symptoms (e.g., re-experiencing symptoms) that are not equivalent to the depressive syndrome (although comorbidity with major depression is extremely common in PTSD). Further complicating this diagnostic dilemma is the suggestion that traumatic loss can result in a bereavement syndrome (suggested for further study in the DSM-5 as "persistent complex bereavement disorder") that is not well captured by either the major depression disorder or PTSD diagnoses, although it has elements of both. Thus, like those who struggle to understand the influence of life events on the emotional functioning of their patients, the DSM-5 continues to struggle with the best way to classify and incorporate stressful life events into diagnosis. In the face of this ongoing turbulence, however, much research has been conducted that serves to sharpen the thinking about the relationship between stress and psychiatric illness.

Stress and Psychotic Disorders

There is little reason to suspect that stress plays a role in the pathogenesis of schizophrenia. Most current etiological theories of schizophrenia focus on genetic and prenatal environmental factors, or a combination thereof. There is, however, ample evidence that adverse life events and stressful social or familial milieu play an important role in determining the course of illness in general and episodes of relapse in particular.

Chronic interpersonal stress, usually studied in the context of the family, is an important risk factor for relapse in schizophrenia. Studies of "expressed emotion," a proxy for familial tumult, have demonstrated that persons with schizophrenia who reside in homes with high-expressed emotion tend to relapse at twice the rate of those who live with families who have low-expressed emotion. Other stressors that plague the lives of many patients with schizophrenia (e.g., poverty, homelessness, and

criminal victimization) also powerfully affect the clinical course of their illness. Combined with evidence that schizophrenia renders the individual more susceptible and sensitive to the negative effects of even minor stressors, it is obvious that stress plays a major role in the course of schizophrenia. Numerous studies have examined the effects on relapse rates of psychosocial interventions that seek to enhance schizophrenic patients' ability to cope with stressful events.

Affective Disorders

Stressful life events and affective illness share a complicated relationship. As discussed earlier, studies of stress face common methodological challenges. This is particularly clear in studies of stress and depression because there is a tendency for depressed patients to selectively recall negative life events and to associate these events with their current difficulties. There is nonetheless evidence that depressed patients really do experience more negative life events than nondepressed patients. This difference may be accounted for predominantly by an excess of "dependent" events (i.e., events that may conceivably have been *caused* by the individual's actions while depressed, such as losing a job due to poor performance or failure in school), as opposed to those independent of the depression (e.g., illness of a family member).

Another way of thinking about the relationship between stress and depression is to focus not on specific life events, per se, but rather on the effects of chronic stress. Studies in this genre are few (e.g., examining rates of depression in mothers of chronically ill children or in Alzheimer's caregivers compared to control subjects), but those that do exist tend to support a relationship between chronic stress and the propensity for major depression. To the extent that most animal models of depression (e.g., chronic restraint, forced swim) are subchronic or chronic stress paradigms, the study of chronic stress effects on depression in humans offers the possibility for convergence and translation between the preclinical, epidemiologic, and clinical knowledge bases.

There is a fundamental similarity between the neuroendocrine effects of stress on humans and the neuroendocrine abnormalities in major depression. For 40 years major depression has been recognized as associated with overactivity of the HPA axis, as inferred by elevated cerebrospinal CRH levels, increased plasma cortisol and corticotropin (ACTH) levels, increased urinary free cortisol, increased cerebrospinal fluid cortisol, and cortisol resistance to dexamethasone suppression. There is also evidence that some features of psychotic depression, in particular, may be mediated by hypercortisolemia. Some studies suggest that childhood maltreatment, which is a strong risk factor for affective and other forms of psychopathology, leaves a neurohormonal "scar" that leads to CRH hypersecretion under stress and may be a vulnerability factor for mood

(and anxiety) disorders.

Biopsychosocial models posit that risk factors on different levels (e.g., experiential and biological) could interact to influence susceptibility to major depression. Although these models have held sway for many years, it is only recently that data to support them have been reported. Researchers are beginning to discover genetic variants within the stress hormone system that increase risk for depression and other psychiatric disorders. Dysfunction of the glucocorticoid receptor pathway is reported in MDD and in individuals exposed to early stressors. A recent study shows that genetic variants alter transcriptional responsiveness of glucocorticoid receptor target genes to the GR. This study suggests that the risk of developing MDD after stressful life events may be influenced by an individual's sensitivity to cortisol release.

Increasingly, studies suggest that neither the gene nor the life stress alone significantly influenced risk; rather, it was the *interaction* of these factors that explained the variation in depressive outcomes. Recent advances in epigenetics may elucidate the mechanisms linking external stresses to affective psychopathology more specifically. Stress may effect gene expression through chromatin remodeling, histone acetylation or methylation, or DNA methylation. For instance, it has been noted that chronic social defeat stress in mice will increase the transcription of a particular DNA methyltransferase (*Dnmt3a*) in the nucleus accumbens, which in turn will increase depressive behaviors after mild social defeat exposures. This effect can be reversed by selectively blocking DNA methylation in that region.

Thus, contemporary genetic and epigenetic research is yielding empirical evidence of what has been viewed as a truism in contemporary psychiatric teachings: Genes and stressful life events operate in concert to elicit depressive episodes. It must also be considered, however, that these types of gene-by-environment interactions may explain risk not only for major depression but also for an array of other stress-influenced mental disorders (e.g., anxiety disorders).

Anxiety Disorders (other than PTSD)

In comparison to the voluminous literature on stress and depression, there is less empirical research on the relationship between stress and anxiety disorders. It is nonetheless apparent that panic disorder frequently has its onset or recrudescence in the context of stressful life events. In particular, there is evidence to suggest that either interpersonal conflict or serious illness (in a significant other) may trigger the onset of panic disorder in susceptible individuals. Given the fact that patients with panic disorder are often somatically preoccupied and likely to interpret physical sensations as life-threatening, it is not surprising that the occurrence of illness in someone close to these patients might precipitate the onset or an

exacerbation of their panic.

One difficulty in assessing the relationship between stressors, stress responses, and anxiety disorders may be that generalized anxiety disorder itself may include a pathological prioritizing of stressors, distorting reported stress burdens for study. That said, at least one major long-term study has identified recent adverse life events as a risk factor for late-onset generalized anxiety disorder. Another found that stressors in the domains of loss and danger were particularly prone to predict onset of anxiety syndromes. In terms of the effects of early life stressors, there is growing evidence that certain adverse early life events such as childhood emotional, physical, or sexual abuse are risk factors for the later development of panic disorder, particularly in women.

Posttraumatic Stress Disorder

DSM-5 recognizes the existence of a group of disorders that are, by definition, stress related. Included here are PTSD and ASD. ASD has its onset following particularly traumatic (often life-threatening) events, such as violent assault or serious accidents, and is denoted by the presence of prominent dissociative symptoms (e.g., derealization, numbing). When ASD occurs following trauma, it identifies a subset of individuals who are at several-fold increased risk for the subsequent development of PTSD (and major depression).

Whereas PTSD was not long ago considered to be a disorder of war veterans (15 to 20 percent of US Vietnam veterans and US soldiers deployed to the more recent Afghanistan and Iraqi conflicts may suffer from this disorder), it is now recognized as a highly prevalent disorder affecting a much larger segment of the population. Among the traumatic stressors that can elicit PTSD are relatively rare events such as natural disasters and frighteningly common events such as criminal victimization. In prior versions of the DSM, it was noted that in order to qualify for a diagnosis of PTSD, the traumatic event needed to be "beyond the range of normal human experience." Epidemiological research has shown that, as much as it is hoped it could be otherwise, horrible things happen to many people and must be considered part of the human experience. Consequently, although the debate continues, it is now generally accepted that any life-threatening (or, perhaps, life-altering) event, however common (e.g., a serious motor vehicle collision), can be considered sufficiently traumatic that it is capable of eliciting PTSD. With this definition in place, recent surveys place the lifetime prevalence of PTSD in the general US population at around 8 to 10 percent, with recognition that certain subpopulations (e.g., soldiers, immigrants from war-torn countries, Native Americans living on reservations) may have rates that are several-fold higher.

The occurrence of PTSD following severe psychological trauma provides

the most persuasive evidence in favor of certain kinds of stress causing psychiatric illness. It is important to note, however, that the neuroendocrine profile of patients with PTSD is not what one might expect to see following chronic stress, nor is it akin to what is observed in major depression. Whereas the latter two conditions would be expected to be associated with hypercortisolism, many studies of patients with PTSD have instead demonstrated *hypocortisolism*. Although the neurobiological implications of this finding remain to be elucidated, the message for stress researchers is clear: it is unlikely that a uniform neuroendocrine "stress response" that explains the heterogeneity of psychiatric disorders encountered in the wake of psychological stress will be seen.

Most severe psychological stressors *do not* result in the development of PTSD (or other forms of psychiatric illness). Even the terror of the attacks on September 11, 2001, resulted in PTSD in only approximately 1 in 10 residents of Manhattan, New York; these data stand as testimony to the resilience of the human psyche. In light of the natural course of timely recovery, it is important for mental health workers to avoid rapidly pathologizing normal reactions to trauma, and to debrief cautiously. Certain characteristics of the traumatic stressor or the immediate postexposure response seem to increase the propensity to develop PTSD following exposure. Some risk factors include the magnitude of and proximity to the stressor, the occurrence of physical injury (including even mild traumatic brain injury), and the occurrence of peritraumatic dissociation. Individual differences also contribute to the variance in PTSD expression following trauma exposure; these include female gender, genetic vulnerability, and a history of exposure to early childhood trauma.

There has also been a surge of interest in recent years in the phenomenon of emotional (or psychological) resilience. Whereas serious traumatic stressors almost inevitably result in short-term symptoms and a decline in functioning, some individuals rebound especially quickly from such insults and go on to function extremely well, even in the face of ongoing chronic stress. Determination of individual differences in personality that characterize psychological resilience, and the neurobiological mechanisms that underlie these observed differences, are a subject of intense research scrutiny at the present time.

PHARMACOLOGICAL APPROACHES TO STRESS MANAGEMENT

As discussed above, stressors worsen most psychiatric conditions. When major psychiatric conditions like depression or psychosis are present, the most appropriate medications are those for the primary psychiatric diagnosis. However, more typically, the picture is primarily one of anxiety or insomnia, perhaps not of as great a magnitude as that found typically in anxiety disorder, but still very distressing to the patient. In such settings, pharmacological therapy with anxiolytic medications or short-term

hypnotics may be helpful, but carries risk. The field is of two minds on this topic and has historically oscillated between extreme viewpoints. Besides the risk for drug dependence, some feel that benzodiazepines may impair the efficacy of psychotherapy by offering a temporary alternative to developing coping skills. On the other hand, depriving patients of such medication when they are in crisis is unconscionable. Short-term use of low-dose benzodiazepines is of considerable utility in helping patients struggle with unusual stressors. This intervention in association with supportive psychotherapy is extremely effective in crisis intervention.

COGNITIVE–BEHAVIORAL APPROACHES TO STRESS MANAGEMENT

William James famously suggested that the greatest weapon against stress is one's ability to choose one thought over another. CBT methods are increasingly used to help individuals better manage their responses to stressful life events. These treatment methods are based on the notion that cognitive appraisals about stressful events and the coping efforts related to these appraisals play a major role in determining stress responding.

Lazarus et al. have developed models of stress and coping that serve as the conceptual foundation for most CBT stress management protocols. These models maintain that there are two types of cognitive appraisal that are especially important in mediating reactions to stress. The first, *primary appraisal*, refers to the way in which a person evaluates the significance or meaning of a given event. When events (e.g., being diagnosed with a chronic illness such as rheumatoid arthritis) are appraised as harmful, threatening, or a loss (e.g., "Because of my rheumatoid arthritis, I will never be able to do anything I really care about"), the individual is more likely to become anxious, depressed, and withdrawn. However, if the same event is viewed as challenging (e.g., "Although my arthritis prevents me from doing some physically demanding activities, there are many other important and meaningful activities that I can do and enjoy"), then more positive outcomes are more likely to occur. *Secondary appraisal* refers to the process of evaluating what can be done about the stressful event and whether or not what can be done will work. Several factors can influence such secondary appraisals, including their learning history, breadth of individuals' repertoire of coping skills, their mastery of specific coping skills, and their expectation that their skills will be effective.

CBT approaches to stress management have three major aims. The first aim is to help individuals become more aware of their own cognitive appraisals of stressful events. The second aim is to educate individuals about how their appraisals of stressful events can influence negative emotional and behavioral responses and to help them reconceptualize their abilities to alter these appraisals. The third aim is to teach individuals how to develop and maintain the use of a variety of effective cognitive and behavioral stress management skills.

The Efficacy of CBT in Stress Management

A number of controlled studies have documented the efficacy of CBT stress management protocols. Three recent controlled studies provide good examples of this type of research.

Individuals who are in the early stages of coping with recovery from breast cancer surgery not only must cope with treatment-related symptoms but also with major disruptions in their work and family life. Investigators conducted a study of newly surgically treated breast cancer patients. Approximately 200 patients were randomly assigned to one of two groups: (1) a comprehensive CBT program focusing on stress management skills, or (2) an educational seminar control condition. Patients in the CBT condition met in groups weekly for 10 weeks and received training in relaxation, problem solving, and assertion training. At the end of 3 months, patients in the CBT condition showed significantly greater improvements in intrusive thoughts about cancer, cancer-related anxiety, and emotional distress than patients in the control condition. Recently these patients were recontacted and found to have maintained significantly lower depressive symptoms and better quality of life up to 15 years post intervention.

HIV is another incurable, high-mortality, highly stigmatized disease which carries a heavy stress burden both from the disease itself and the imperfect treatment options. Several studies have applied similar cognitive–behavioral stress management interventions to HIV-positive patients around the world. One recent study randomized 71 low-income, racial-minority, HIV-positive and HPV-positive women to either CBT stress management or a psychoeducation control group. The investigators found that the CBT group demonstrated lower depression scores as well as stronger immune function post intervention than the control group. Several studies of HIV-positive patients in the US, Iran, Africa, and elsewhere have similarly demonstrated the effect of CBT stress management to significantly decrease depressive symptoms, increase well-being and sleep quality, and in fact often improve biomarkers such as viral load and rates of cervical intraepithelial neoplasia. The therapy has also been found to alleviate functional dyspepsia symptoms, improve well-being in prostate cancer survivors, and decrease symptom severity and health care utilization costs in patients with chronically unexplained medical symptoms.

Stress Management Training

When CBT is used for stress management, training is provided in a wide range of stress management skills. Although space limitations preclude a discussion of all of these skills, it is important to highlight five skills that form the core of almost all CBT stress management programs: Self-observation, cognitive restructuring, relaxation training, time management,

and problem solving.

Self-Observation. One of the most effective ways of helping individuals become more aware of how they respond to problem situations is to have them keep a daily record of their behavior. A daily diary format (Table 27.10–1) is often used with patients who are asked to keep a record of how they responded to challenging or stressful events that occurred each day. Entries are made in three columns: antecedents, behaviors, and consequences. In the antecedents column, patients record a specific environmental event that they perceived as stressful. The event might be an interpersonal stressor such as having an argument with a spouse or a confrontation with a coworker, or an intrapersonal stressor such as a flare in pain symptoms or a major natural disaster. To assess appraisals of the stressor, patients are asked to rate how stressful they found this event on a 0 to 100 scale (e.g., 0 = not at all stressful, to 100 = the most stressful event ever experienced). In the column marked behavior, patients are instructed to record their behavioral reactions to the stressful event. The term *behavior* is broadly defined so that it not only encompasses overt behavioral responses but also more covert cognitive, affective, and physiological responses. Thus, in the behavioral response column individuals might indicate specific coping behaviors they engaged in (e.g., avoidance), cognitive responses to the stressor (e.g., "I am overwhelmed and can't deal with this right now"), negative emotions (e.g., anxiety, depression), and aversive physiological responses (e.g., increasing sweating, heart rate, or nausea). In the column marked consequences, patients record the outcomes resulting from their behavior in this situation.

Self-observation is effective for several reasons. First, it makes individuals more aware of behaviors that they usually fail to notice. Individuals whose style of dealing with social anxiety is avoidance, for example, may avoid potentially stressful situations in a more or less automatic fashion (e.g., routinely refuse invitations), use very subtle forms of avoidance (e.g., arriving or leaving a party early), or arrange their lifestyle (e.g., work alone) so that they have very few chances to confront their fears. By keeping a record of their own behavior, individuals often begin to recognize these behavioral patterns. A second important feature of self-observation is that it helps individuals see how their cognitive–behavioral responses are linked to situational antecedents and consequences. Participants who keep daily diaries are often struck by the variations in their own reactions across different events. For example, while they may cope effectively with one stressful event (e.g., meeting an important work deadline), they find themselves unable to deal with another stressful event (e.g., an argument with their teenage son or daughter). Self-observation highlights the role that coping efforts and cognitive appraisals, as well as emotional and physiological responses, play in explaining these

very different responses to stressful situations.

Table 27.10–1.
A Daily Diary Format for Recording Stress-Related Behaviors

Antecedents (Briefly describe the stressful event)	Stress Rating (Rate how stressful the event was: 0 = not at all, 100 = most stressful event ever experienced)	Behaviors (Indicate how you responded to the event)	Consequences (What was the outcome?)
A minor traffic accident: I was hit from behind at a stop light. Hurt my neck slightly, no damage to the car.	Rating = 70	Thoughts: This kind of thing always happens to me, my boss is really going to be angry at me for missing the morning meeting. Feelings: Frustrated, anxious. Physiological response: Clenching my teeth, stomach churning. Behaviors: Started shouting at the driver who hit me and then went back to my car to calm myself down and wait for the policeman.	I was able to regain control over my emotions. The policeman came quickly and I was able to make the morning meeting on time.

With the ubiquity of smartphones, the authors have seen growing interest in daily record keeping of both behavioral and biological data collected for individual patients. Apps such as Ginger.io and Mood 24/7 have made it easy to document symptoms and situational factors (precipitating events, social context, etc.) in real time. This approach, called *ecological momentary assessment*, enables a person to capture stress and coping processes much closer to their occurrence and can increase the accuracy and validity of self-observation. Automatically graphing the timing of anxiety exacerbations, for instance, can aid patients in identifying patterns in order to better understand and change their responses daily stressors.

Cognitive Restructuring. A hallmark of CBT is its insistence that cognition plays a central role in the stress and coping process. In CBT, cognitive appraisals about stressful events are considered to be the key factor in determining stress-related responding. Given this emphasis on cognition, it is not surprising that a major thrust of CBT approaches to stress management is on helping participants to become aware of and change their maladaptive thoughts, beliefs, and expectations.

Cognitive–behavioral therapists have developed protocols for helping distressed individuals restructure dysfunctional or maladaptive thinking patterns. Such cognitive restructuring begins by educating patients about the cognitive model and the notion that underlying cognitions are major

determinants of emotional responding. According to the cognitive model, emotional reactions that are excessive or prolonged are often the result of cognitions that are distorted or dysfunctional.

The second step in cognitive restructuring is monitoring and analyzing dysfunctional thoughts. Table 27.10–2 lists a number of stress-related thoughts that might be identified through daily diary records. As can be seen, each of these thoughts can be linked to a negative emotion. A careful analysis of each thought also reveals an inherent error in logic. Certain types of errors, or cognitive distortions, occur quite frequently, and individuals can, thus, be taught to recognize them.

Table 27.10–2.
Stress-Related Negative Thoughts

Stressor	Cognition	Emotion
Painful medical procedure	I know I won't be able to tolerate the pain.	Fear
Argument with a friend	She will never speak to me again.	Discouragement
Coworker injured on the job	It is my fault. I should have warned him about the danger.	Guilt
Late for a deadline	It's hopeless, I'm always messing things up.	Depression
Spouse criticizes you	He just doesn't get it. He will never understand what I am trying to do.	Anger

Table 27.10–3.
Key Questions for Challenging Negative Automatic Thoughts

Is there an alternative explanation?
What is the evidence that this thought is true?
What will be the effect of continuing to think this way?
What is the best outcome, worst outcome, and most realistic outcome?
What is the likelihood that this will happen?

Adapted from Beck JS. *Cognitive Therapy: Basics and Beyond*. New York: Guilford; 1995.

The third step in cognitive restructuring is to challenge and change cognitive distortions. Participants are often provided with a list of "key questions" (Table 27.10–3) that can be applied to their own negative thoughts. They are asked to select one or two questions that are useful in identifying the underlying logical problems in their thinking. Finally, participants develop a rational response that represents a more accurate and helpful cognitive response to the situation.

Table 27.10–4 illustrates how one might change various overly negative stress-related thoughts to more adaptive rational responses. If the rational response is to be effective, it must be used frequently, ideally each time the negative thought occurs. To ensure that this occurs, patients are given home practice assignments in which they are asked to expose themselves to

stressful events, monitor their negative thoughts, and respond with a more adaptive thought pattern. They are also asked to keep records of the resulting consequences, both in terms of changes in emotions and behaviors.

Cognitive restructuring is receiving increasing attention among researchers interested in stress management. Some of the most interesting topics receiving recent attention include how cognitive restructuring alters brain processes related to stress and coping and also how changes in cognition relate to sudden gains achieved during treatment.

Relaxation Training. Relaxation skills can be very helpful in managing stress. When individuals learn to relax, their overall muscle tension is reduced, as does their overall level of autonomic arousal. Individuals who are able to relax are also more likely to be able to think more rationally and to restructure negative cognitions when faced with stressful events. They also are more likely to rest and sleep better, thereby reducing the fatigue that often accompanies stress. Finally, relaxation skills may be helpful in interrupting maladaptive behavior patterns. Progressive relaxation training usually consists of a series of structured exercises involving slowly tensing and relaxing major muscle groups. The exercises are introduced during treatment sessions, and patients are encouraged to practice them daily at home. One of the most important tasks in relaxation training is helping patients learn how to generalize their skills in relaxation from home practice sessions to stressful and demanding daily life events. Several techniques are used to enhance generalization. The first is a method called *differential relaxation*. In differential relaxation, the trainee is instructed to engage in a daily task and only use the muscles that are necessary for carrying out the task. For example, while writing one's name there needs to be activity in muscles of the hands and arms, but the muscles in the face, lower trunk, legs, and feet can be deeply relaxed. Differential relaxation is initially practiced in relatively easy situations (e.g., sitting or standing) where the demands are minimal. Over time, however, patients learn how to use differential relaxation in situations that are challenging and stressful. A second generalization method is the *mini-practice*. This consists of a brief relaxation procedure in which the trainee takes a deep breath, and then while slowly exhaling focuses on sensations of relaxation flowing downward from the muscles of the face to those of the neck, shoulders, trunk, and legs. The mini-practice takes approximately 30 seconds and thus can be done repeatedly over the course of a day. Individuals are encouraged to develop innovative reminders to cue themselves to perform frequent mini-practices every day. As individuals increase this frequency, they often report that relaxation becomes more or less an automatic and habitual response.

Table 27.10–4.
Negative Thoughts and Rational Responses

Negative Thought	Rational Response
I have no control over this situation.	It's true that I can't change what happened, but I can control how I respond.
I can't return to my old job, I am worthless.	Even though I may not be able to do the same kind of work, there are a number of other jobs that I may be able to do.
It's not fair that I am being treated this way.	Life is not always fair. Sometimes things go my way, other times they don't.
I am a burden on my family.	My family has had to give me more support lately, but I do a lot of things for them as well.
No one really cares about me.	There are people in my life (e.g., spouse or child) that I can turn to for support and love.

Relaxation training is frequently used in managing stress in persons suffering from chronic medical disorders such as cancer, cardiovascular disease, and chronic pain. For some people, relaxation training and its variant of meditation can have powerful effects on both physiological and psychological responses to stressors. The heart rate response to intravenous isoproterenol (Isuprel) is a venerable tool of autonomic physiology with highly predictable results. A serendipitous case observation found that the heart rate response to pharmacological doses of isoproterenol was blocked in a patient who spontaneously started to meditate during the infusion (Fig. 27.10–6).

Time Management. Time management methods are designed to help individuals restore a sense of balance to their lives. The first step in training in time management skills is designed to enhance awareness of current patterns of time use. To accomplish this goal, individuals might be asked to keep a record of how they spend their time each day, noting the amount of time spent in important categories such as work, family, exercise, or leisure activities.

The second step in time management is designed to help individuals set their priorities. In setting priorities, an important distinction can be drawn between urgency and importance. To explore this distinction, it is helpful to envision a table that has along one axis the dimension of importance and along the other axis the dimension of urgency. Patients are then encouraged to enter into each quadrant of the table examples of activities from their own life and to reflect on the patterns they observe. Individuals may note that much of their time is spent doing activities that are in the urgent-not-important quadrant or the not-urgent-not-important quadrant. Alternatively, they may report that they spend little or no time engaging in activities in the not-urgent-but-important quadrant. Patients are encouraged to discuss the benefits that can come from increasing the priority they give to such important, but less urgent, activities such as

exercising, resting, relaxing, or doing volunteer work. They are also encouraged to try to match the time spent engaging in such meaningful activities with the priority they place on the activity.

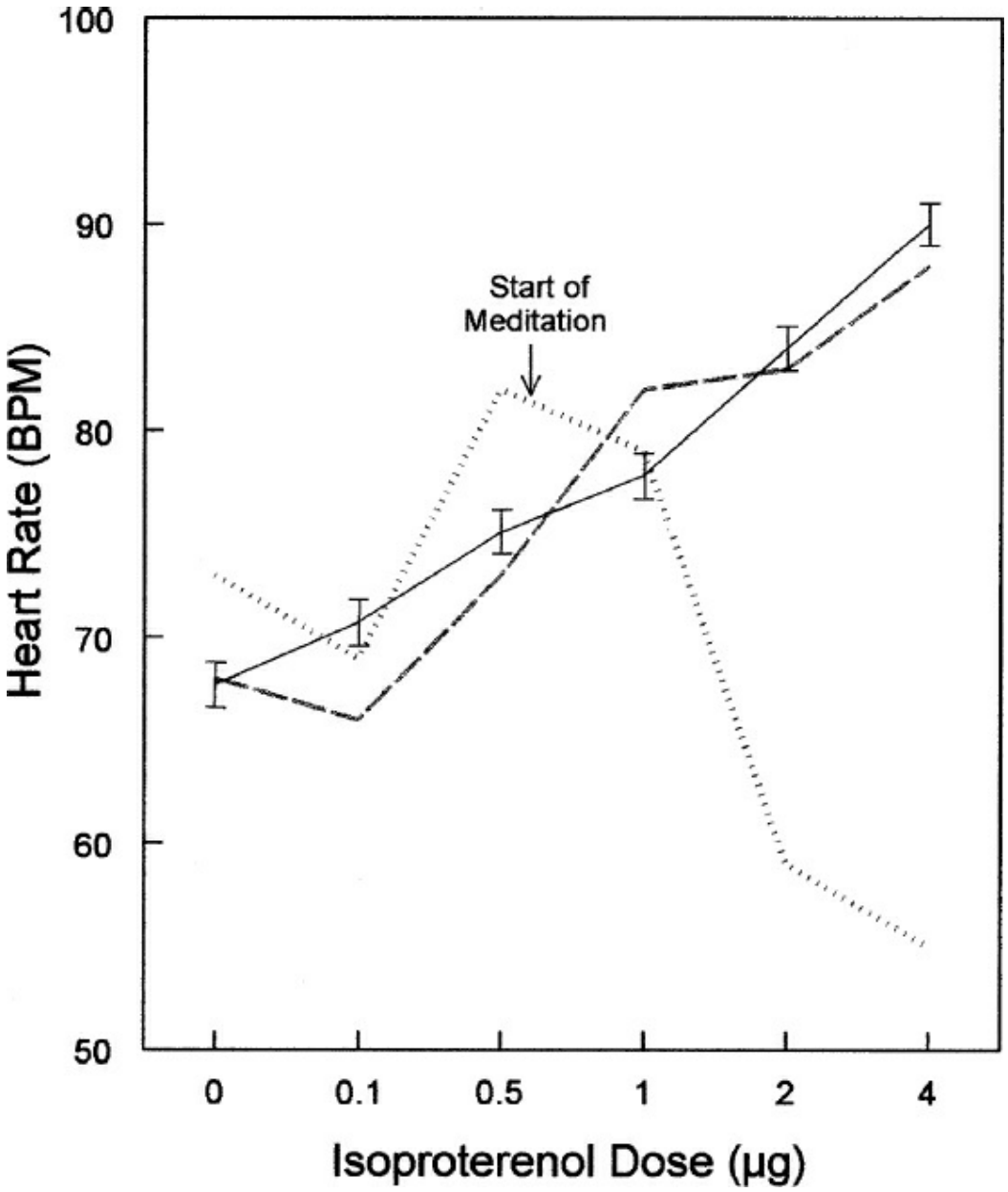

FIGURE 27.10–6. The effect of meditation on heart rate response to isoproterenol. The patient (*dotted line*) began spontaneously meditating and her heart rate dropped. The solid line presents the expected heart rate response. The patient was retested and instructed not to meditate (*dashed line*) and her response was then typical. BPM, beat per minute. (From Dimsdale J, Mills P. An unanticipated effect of meditation on cardiovascular pharmacology and physiology. *Am J Cardiol*. 2002;90:908.)

Goal setting focuses on specific behavior change targets that fit with an individual's long-term priorities. First, the individual is asked to identify a goal to be accomplished in the next week. The goal, for example, might be to lose 2 lb or to sign up for an exercise program. Second, patients are

asked to critically analyze the goal to ensure that is reasonable, specific, and personally meaningful. In the third step of goal setting, individuals are asked to reconfirm a time frame (e.g., 1 week) in which they will meet the goal. This step is important in reducing procrastination, a common obstacle to effective time management. Finally, the individual's efforts in reaching the goal are reviewed. Successful completion of a goal that has been long neglected can be very rewarding.

The evidence base for time management is growing. One meta-analysis of 32 empirical studies of time management found that the increased use of time management behaviors was significantly related to better health, lower levels of stress, and higher levels of job satisfaction.

Problem Solving. Problem solving involves several basic steps. The first is problem identification. In this step, an individual tries to identify the key problematic aspects of a stressful event such as problematic behaviors, thoughts, feelings, and physiological responses. For example, dealing with the aftermath of a hurricane may be perceived as stressful because it is associated with: (1) particular behavioral problems, e.g., having to clean up while simultaneously caring for one's children; (2) overly negative thoughts, e.g., "We will never recover from this event"; (3) negative feelings, e.g., fear, anger, and depression; and (4) physiological responses, e.g., muscle tension, fatigue, and sleep problems. Helping individuals understand the various facets of their problematic responses is often helpful because it suggests what skills might be particularly useful. For example, if the problem is mainly a somatic one (increased tension), then relaxation-based strategies may be the most effective.

The second step in problem solving is generating alternatives. Brainstorming is used to generate a wide range of alternative problem solutions. The rules of brainstorming include: (1) individuals offer as many solutions as they can, even if the solutions do not immediately seem reasonable, (2) no criticism is allowed, and (3) there should be an attempt to mix and match solutions in creative ways.

The third step in problem solving involves evaluating the alternatives and selecting the best solution. To begin this process, a participant can be asked to rate the likelihood that each solution will have a positive outcome. It is often helpful to discuss some of the factors that influence how the participant makes these ratings (e.g., cost, availability of resources, effects of the solution on others). Typically, one solution stands out as the best and most practical.

The final step of problem solving involves implementing the solution. Participants basically try to apply the best solution in the problem situation and then review their progress with the therapist.

There is growing recognition that the development of problem-solving skills is crucial to the prevention of relapse and maintenance of the effects

of CBT. Participants who learn how to approach a problem situation and formulate and implement effective coping options benefit in several important ways. First, they are often able to anticipate problems and make plans to avoid them. Second, they cope with setbacks more effectively and are able to get their coping efforts back on track after a setback so as to prevent a major relapse. Finally, they are better prepared for termination of therapy because they have learned to engage in a form of "self-therapy" in which they can pinpoint and solve their own problems.

FUTURE DIRECTIONS

A 51-year-old single woman was seen in psychiatric consultation for help with severe anxiety and occasional panic 8 months after having been stabbed by a stranger in an apparent gang initiation assault. She survived without permanent injury, but had required a chest tube and a short ICU stay at the time, and had been afraid to leave her home unaccompanied since the assault. Over the next few months, she began to experience nightmares, insomnia, loss of appetite, and episodes of severe shortness of breath, flushing, and chest pain lasting 15 to 20 minutes about once a week, sometimes leading to emergency department visits. An emergency physician prescribed her low-dose benzodiazepines, which helped her sleep, but had begun to create a dependency.

In the consultation clinic, she was jumpy, fretful, and underweight. She reported feeling "stupid" for having been out alone at night. An SSRI was started and her appetite, hypervigilance, and insomnia improved. She no longer required benzodiazepines. However, her anxiety regarding leaving the house remained, as did her fear of panic attacks and her self-blame.

Cognitive–behavioral therapy helped her to see this self-blame as a cognitive distortion and she practiced countering this with affirmations that reassigned the guilt to her attacker. She practiced breathing and muscle relaxation exercises that reduced the fear associated with her panic attacks, and they eventually abated. Her psychiatrist helped her to develop a plan of gradual exposure aimed at getting her out of the house unaccompanied, first briefly in the daytime, then ultimately after dark alone. After 3 months of treatment, she had essentially returned to her full level of functioning.

By way of conclusion, the case study provided demonstrates many of the issues discussed in this section—the confluence of adverse life events, physical illness, and emotional distress, and the pharmacological and psychotherapeutic treatment of stress-related psychopathology.

How would this case be changed 10 years from now? How would it be conceptualized in future taxonomies of psychiatric illness? What sorts of additional information would the clinician of tomorrow try to obtain to treat this patient? In other words, how will "stress" be evaluated and treated by psychiatrists in the future?

Crime and violence are not going to go away. Indeed, with increased population density and volatile economies, such transgressions may afflict ever-greater populations. Increasing numbers of patients whose lives are disrupted by violence may present for psychiatric care.

Perhaps, in the future, physicians may be better able to forecast individuals' vulnerabilities to such stressors. The extensive epidemiological literatures from military psychiatry and disaster psychiatry are consistent

in their finding of a spectrum of risk among those who are exposed to major stressors. This has led to an emphasis on intervening in those at higher risk. Typically, this has meant focusing on the relatively socially isolated individuals, as social isolation is repeatedly observed as a risk for poorer outcomes after trauma exposure. With increasing research, physicians may also find that interventions are particularly important for individuals with certain lifetime experiences or certain genetic legacies. With advances in understanding of individual vulnerabilities comes the opportunity for early identification and intervention. Future research will no doubt test the efficacy of early treatment with psychopharmacologic, cognitive–behavioral, or combined treatments.

To date, most approaches to managing stress have focused solely on the individual. However, with changes in the way health care is being delivered, there will be a greater need to involve family members and caregivers in the treatment process. Stress occurs in a social context and, for this reason, involving significant others in the process of both assessment and treatment makes sense. Significant others provide an important role as informants about the patient's status; they also can play an active role in helping patients better understand their adaptation to stress and in encouraging the adoption and continued use of effective stress management interventions.

Looking into the future, it is also obvious that medical management of physical injury and diseases will change. Patients always have difficulties coping with unfamiliar and invasive treatments. It takes time for such treatments to become less upsetting. Similarly, the new medical treatments will bring with them unanticipated consequences and side effects, which psychiatrists will need to consider in their management of patients. If the patient's mood worsens in the hospital, is the decline best attributed to the stress or to a side effect of some new (or old) drug given in the intensive care unit? It is clear that in the future, psychiatrists will need to be more seamlessly embedded in somatic treatment settings.

Psychopharmacological interventions will assuredly also be different in the future. The days of prescribing barbiturates to help patients deal with stress are long gone. These dated interventions have disappeared, along with telegraphs and slide rules, to be replaced by more effective and safer agents. As more is learned about the fundamentals of stress, anxiety, and depression, it is only a matter of time before benzodiazepines and SSRIs will be supplanted by pharmacological agents with different actions and (hopefully) stronger benefits and fewer risks. Moreover, if the promise of personalized medicine is fulfilled, then the ability to match treatments to patients will also improve.

As the current understanding advances and interventions evolve, psychiatrists will continue to provide patients with the tools and therapies required not just to overcome, but actually to understand their stress.

Looking ahead and to the future, it is important to not forget the insights already recognized at the dawn of stress research. While stress can be a hazard or even a risk factor, it can also be a potent constructive force in patients' lives.

References

Antoni MH. Psychosocial intervention effects on adaptation, disease course and biobehavioral processes in cancer. *Brain Behav Immun*. 2013;30 Suppl:S88–S98.

Beck AT, Weishaar M. Cognitive therapy. In: Wedding D, Corsini RJ, eds. *Current Psychotherapies*. 10th ed. Belmont, CA: Brooks/Cole Cengage Learning; 2013:231.

Boland TA, Lee VH, Bleck TP. Stress-induced cardiomyopathy. *Crit Care Med*. 2015;43(3):686–693.

Cozzolongo R, Porcelli P, Cariola F, et al. Serotonin gene polymorphisms and lifetime mood disorders in predicting interferon-induced depression in chronic hepatitis C. *J Affect Disord*. 2015;183:90–97.

Epel ES, Blackburn EH, Lin J, et al. Accelerated telomere shortening in response to life stress. *Proc Natl Acad Sci USA*. 2004;101(49):17312–17315.

*Feder A, Nestler EJ, Charney DS. Psychobiology and molecular genetics of resilience. *Nat Rev Neurosci*. 2009;10(6):446–457.

Golbidi S, Frisbee JC, Laher I. Chronic stress impacts the cardiovascular system: animal models and clinical outcomes. *Am J Physiol Heart Circ Physiol*. 2015;308(12):H1476–H1498.

*Gold PW. The organization of the stress system and its dysregulation in depressive illness. *Mol Psychiatry*. 2015;20(1):32–47.

Goyal M, Singh S, Sibinga EMS, et al. Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. *JAMA Intern Med*. 2014;174(3):357–368.

*Haglund ME, Nestadt PS, Cooper NS, Southwick SM, Charney DS. Psychobiological mechanisms of resilience: relevance to prevention and treatment of stress-related psychopathology. *Dev Psychopathol*. 2007;19(3):889–920.

Hirohata S, Abe A, Murasawa A, Kanamono T, Tomita T, Yoshikawa H. Differential effects of IL-6 blockade tocilizumab and TNF inhibitors on angiogenesis in synovial tissues from patients with rheumatoid arthritis. *Modern Rheumatology*. 2016:1–22. doi:10.1080/14397595.2016.1259717

*Irwin MR. Why sleep is important for health: a psychoneuroimmunology perspective. *Annu Rev Psychol*. 2015;66:143–172.

Irwin MR, Opp MR. Sleep health: reciprocal regulation of sleep and innate immunity. *Neuropsychopharmacology*. November 2016. doi:10.1038/npp.2016.148

Karg K, Burmeister M, Shedden K, Sen S. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. *Arch Gen Psychiatry*. 2011;68(5):444–454.

Kendler KS, Hettema JM, Butera F, Gardner CO, Prescott CA. Life event dimensions of loss, humiliation, entrapment, and danger in the prediction of onsets of major depression and generalized anxiety. *Arch Gen Psychiatry*. 2003;60(8):789–796.

Laplant Q, Vialou V, Covington HE, et al. Dnmt3a regulates emotional behavior and spine

plasticity in the nucleus accumbens. *Nat Neurosci*. 2010;13(9):1137–1143.

Lopez CR, Antoni MH, Pereira D, et al. Stress management, depression and immune status in lower income racial/ethnic minority women co-infected with HIV and HPV. *J Appl Biobehav Res*. 2013;18(1):37–57.

McHugh PR, Treisman G. PTSD: a problematic diagnostic category. *J Anxiety Disord*. 2007;21(2):211–222.

Murray C, Griffin ÉW, O'Loughlin E, et al. Interdependent and independent roles of type I interferons and IL-6 in innate immune, neuroinflammatory and sickness behaviour responses to systemic poly I:C. *Brain Behav Immun*. 2015;48:274–286.

Musliner KL, Seifuddin F, Judy JA, Pirooznia M, Goes FS, Zandi PP. Polygenic risk, stressful life events and depressive symptoms in older adults: a polygenic score analysis. *Psychol Med*. 2015;45(8):1709–1720.

Patel SK, Wong AL, Wong FL, et al. Inflammatory biomarkers, comorbidity, and neurocognition in women with newly diagnosed breast cancer. *J Natl Cancer Inst*. 2015;107(8). pii: djv131.

Post RM, Leverich GS. The role of psychosocial stress in the onset and progression of bipolar disorder and its comorbidities: the need for earlier and alternative modes of therapeutic intervention. *Dev Psychopathol*. 2006;18(4):1181–1211.

Quartana PJ, Finan PH, Page GG, Smith MT. Effects of insomnia disorder and knee osteoarthritis on resting and pain-evoked inflammatory markers. *Brain Behav Immun*. 2015;47:228–237.

Réus GZ, Fries GR, Stertz L, et al. The role of inflammation and microglial activation in the pathophysiology of psychiatric disorders. *Neuroscience*. 2015;300:141–154.

Sampson TR, Mazmanian SK. Control of brain development, function, and behavior by the microbiome. *Cell Host Microbe*. 2015;17(5):565–576.

Shaw W, Dimsdale J, Patterson T. Stress and life events measures. In: A John Rush Jr. ed. *Handbook of Psychiatric Measures*. Washington, DC: American Psychiatric Association; 2000:221.

Stagl JM, Bouchard LC, Lechner SC, et al. Long-term psychological benefits of cognitive-behavioral stress management for women with breast cancer: 11-year follow-up of a randomized controlled trial. *Cancer*. 2015;121(11):1873–1881.

Ulrich-lai YM, Herman JP. Neural regulation of endocrine and autonomic stress responses. *Nat Rev Neurosci*. 2009;10(6):397–409.

*Zannas AS, West AE. Epigenetics and the regulation of stress vulnerability and resilience. *Neuroscience*. 2014;264:157–170.

▲ 27.11 Psychiatric Morbidity Following Critical Illness

O. JOSEPH BIENVENU, M.D., PH.D., AND DIMITRY S. DAVYDOW, M.D.

INTRODUCTION

Mr. Williams was a healthy, busy 35-year-old family man when he developed a cough with fever, diaphoresis, and pleuritic chest pain. At a local emergency department, he was started on

oral antibiotics for "walking pneumonia." However, his wife had to bring him back 2 days later as he had increasing shortness of breath and weakness and he felt like he was dying. Because of inadequate oxygenation with noninvasive methods (respiratory failure), he was intubated and started on mechanical ventilation. He was diagnosed with the acute respiratory distress syndrome and required intensive care for 12 days.

While in intensive care, Mr. Williams did not know where he was and had visual hallucinations of children with animal heads and circular mazes dripping with blood. He was quite agitated and required physical restraints, and his endotracheal tube and Foley catheter were apparently quite uncomfortable. He developed a belief that his penis had been excised and that he would have to pay a large sum to have it reattached. He thought his nurses were variously ignoring and trying to kill him.

After he was extubated and transferred to a general hospital ward, Mr. Williams required physical therapy until he could move around better. However, he remained easily fatigued for months, and his strength returned quite gradually, with substantial effort. He also had trouble getting things done, with his expectations greatly exceeding his abilities. He had trouble with his memory (e.g., where he had put his keys) and difficulty completing sequential tasks (e.g., putting furniture together). The loss of his fast-paced, high-responsibility job had a profound demoralizing effect on him.

Mr. Williams also became very fearful, particularly of infections, as he was told his critical illness stemmed from an infection. He used a mask when cleaning, changed his sheets frequently, and took vitamins and minerals meant to boost his immune system. He feared overdoing it at the gym, and his wife became embarrassed when he frequently spoke of illness to others. When he became ill with a cough or a runny nose, he stayed home or made unnecessary emergency room visits, and he avoided playing with his young son when his son had a cold. He had intrusive memories of wearing an oxygen mask in the emergency department and the sound of the ventilator when he was critically ill. He felt distant from his loved ones, as well as unsupported, and also that no one understood what he had been through. He no longer felt he would have a long life. He felt like a burden, particularly because he could not do much around the house. He wondered if he had done something wrong, why this had happened to him. He was frustrated by his slow recovery and avoided others, not wanting to be bothered. Sometimes he had difficulty falling asleep and would anxiously check his pulse and breathing, especially when concerned he was coming down with something. He was also irritable, especially when people did not show respect or empathy.

Post–Intensive Care Syndrome

With advances in intensive care medicine, many more patients are surviving critical illnesses like severe sepsis and the acute respiratory distress syndrome. Unfortunately, many critical illness survivors suffer with posttraumatic stress, anxiety, and depressive phenomena, as well as cognitive impairment and physical weakness, and their family members are also often quite distressed. To increase awareness of these common problems, a group of clinical outcomes researchers and other stakeholders coined the term "post–intensive care syndrome." This section will touch on all of these problems, as they are often relevant to critical illness survivors and their families, but this section will focus particularly on neuropsychiatric problems.

PSYCHIATRIC MORBIDITY IN CRITICAL ILLNESS SURVIVORS AND THEIR FAMILY MEMBERS

What Is Stressful about Critical Illness?

In many ways, critical illness–associated distress is easy to understand. Patients who are critically ill with acute respiratory failure experience pain with endotracheal intubation and suctioning, as well as invasive monitoring and therapeutic interventions such as arterial and central intravenous lines, Foley catheters, and chest tubes. In addition, severe sepsis is associated with systemic inflammation and dysfunction of many organs, including the brain and the adrenal glands (leading to a need for endogenous and exogenous catecholamines to maintain adequate blood pressure). With acute brain dysfunction from infection, organ failure, and sedative medications, patients are often unable to process what they're experiencing (e.g., manipulation of a Foley catheter can be misperceived/misinterpreted as rape), and frightening hallucinations and delusions are common. All of these stresses occur while patients have lost their autonomy and ability to communicate normally. If patients survive, they are often weak and cognitively impaired; thus, their work, social, and family lives are additionally disrupted by these factors. Similarly, family members face a number of stressors, including uncertainty about their loved ones' survival and ultimate functioning, financial difficulties with disrupted jobs and the need to provide informal care, and additional disruptions in work and social lives, including changing roles within the family.

Adult Critical Illness Survivors

Adult critical illness survivors have substantially higher prevalences of psychiatric morbidity than expected, compared to the general population. For example, ~25 percent of critical illness survivors suffer with clinically significant PTSD symptoms, and a similar proportion suffer with clinically significant depressive symptoms. Studies employing clinician interviews suggest lower but still substantial prevalences of PTSD and depression-related diagnoses. Importantly, psychiatric morbidity may explain, at least in part, the frequently low health-related quality of life (HRQOL) in survivors.

Risk factors for psychiatric morbidity in critical illness survivors include prior psychiatric illness, in-ICU benzodiazepine/sedative doses, mechanical ventilation, and post-ICU memories of frightening psychotic experiences. In addition, early post-ICU distress strongly predicts later psychiatric morbidity.

Pediatric Critical Illness Survivors

Pediatric critical illness survivors also have higher rates of psychiatric disorders than their general population counterparts. Patient and parent reports indicate that substantial PTSD symptoms occur in 10 to 28 percent of child and adolescent critical illness survivors, and clinician interview

results suggest frank PTSD in 0 to 21 percent. Patient and parent reports indicate that substantial depressive symptoms occur in 7 to 13 percent of this patient population, although, as in adults, major depressive disorder appears less common. Also as in adults, clinically important psychiatric symptoms are associated with worse HRQOL.

As in adult critical illness survivors, a pre–critical illness history of psychiatric and/or developmental problems appears to be the most potent predictor of substantial psychiatric symptoms in children after intensive care. Also in keeping with adult critical illness survivors, early post-illness psychiatric symptoms present appear to be predictors of later, more persistent substantial PTSD and depressive symptoms in children surviving critical illnesses. Notably, parental PTSD, anxiety, and/or depressive symptoms after a child's critical illness and the child's post-ICU psychiatric morbidity appear correlated.

Family Members of Critically Ill Patients

As mentioned, critical illness survivors are not alone in facing substantial distress. Family members of critically ill patients also have high rates of clinically important psychiatric symptoms, with concomitant poorer functioning and quality of life. Approximately one-third of parents of critically ill children experience substantial acute stress symptoms. Studies using questionnaires suggest that anywhere from 13 to 56 percent of relatives of critically ill patients have substantial PTSD symptoms, while 8 to 42 percent have substantial depressive symptoms. Furthermore, 21 to 56 percent of family members of critically ill patients have substantial nonspecific anxiety symptoms. As noted above, substantial PTSD, depressive, and nonspecific anxiety symptoms tend to cooccur in family members of critically ill patients.

Women with a critically ill family member may be at greater risk of developing substantial PTSD and/or depressive symptoms than men. In addition, family members of patients who die or almost die in the ICU, and those family members who have additional stressors beyond their relative's critical illness, appear at greater risk of clinically important psychiatric symptoms. For family members of patients who die in the ICU, discordance between their preferred medical decision-making role and their actual role, and being present at the time of the loved one's death, may increase risk. Finally, wives of men with substantial functional limitations after a severe illness may be at increased risk of developing substantial distress compared to husbands of women with similar post-illness limitations.

COGNITIVE DYSFUNCTION IN CRITICAL ILLNESS SURVIVORS

Prevalence

A growing body of literature suggests that critical illnesses can affect

cognitive functioning substantially and persistently. In studies of acute respiratory distress syndrome (ARDS) survivors, 70 to 100 percent had evidence of cognitive impairment at hospital discharge, and 25 to 47 percent had clinically meaningful cognitive impairment at 2-year follow-up. In studies of general ICU survivors, a wider range (10 to 71 percent) showed evidence of cognitive impairment 1 year after hospital discharge. Notably, the degree of long-term cognitive impairment present in these patients can be comparable to that seen in patients with moderate traumatic brain injury or mild Alzheimer disease.

Affected Cognitive Domains and Neuroimaging Findings

While critical illness survivors can exhibit impairment in all domains of cognitive functioning, the domains most severely affected are attention/concentration, memory, and executive functioning. Neuroimaging studies of critical illness survivors have correlated these phenomena with cerebral atrophy, ventricular enlargement, and hippocampal volume reductions, as well as connectivity disturbances.

Potential Risk Factors

Although older age is a risk factor for persistent cognitive impairment after critical illness, younger and middle-aged adult survivors can also be affected. Pre-existing cognitive impairment and depression appear to increase risk.

Several clinical factors appear to be causally related to, or at least markers of risk of, long-term cognitive dysfunction in critical illness survivors. For example, severe sepsis appears causally related, and delirium in the ICU at least a marker of risk, for persistent cognitive dysfunction. In-ICU hypoxia, hypotension, and glucose dysregulation have all been associated with persistent cognitive impairment, and other clinical, genetic, and treatment factors may be relevant.

PHYSICAL DYSFUNCTION IN CRITICAL ILLNESS SURVIVORS

Neuromuscular weakness after critical illness/intensive care is very common, particularly early in the course of recovery (during the first year). “ICU-acquired weakness” encompasses critical illness polyneuropathy and myopathy, as well as disuse atrophy. Polyneuropathy can recover particularly slowly (over years, as opposed to months).

Apparent risk factors for critical illness polyneuropathy and myopathy include hyperglycemia, the systemic inflammatory response syndrome, sepsis, and multi-organ dysfunction. The causal factor for disuse atrophy is immobility (bed rest).

Importantly, physical functioning (including the ability to perform instrumental activities of daily living) can remain impaired longer than

muscle weakness. Some apparent risk factors include ICU administration of systemic corticosteroids, ICU-acquired illnesses, slow resolution of lung injury, depression, and older age.

INTERVENTIONS FOR PSYCHIATRIC DISORDERS IN CRITICAL ILLNESS SURVIVORS

Interventions for Post–Critical Illness PTSD and Depression

Adults. One of the most promising interventions to reduce long-term distress in critical illness survivors is ICU diaries. Unlike normal diaries, ICU diaries are written *to* patients who are critically ill. Clinicians, often intensive care nurses, write in plain language about the patient's condition throughout the ICU stay, including diagnostic and therapeutic procedures. Family members are also encouraged to write about their observations, thoughts, feelings, and occurrences outside of the hospital. Sometimes the diaries include photographs of the ICU room (+/– the patient) and of staff. The rationale for ICU diaries is to help patients create a more coherent autobiographical narrative than they may otherwise have due to acute brain dysfunction while critically ill. Patients can process this information when they're ready and able, at their own pace, sometimes sharing with loved ones. Studies suggest that ICU diaries are acceptable to patients and families and that they reduce long-term distress, especially PTSD symptoms, in critical illness survivors. ICU diary programs are becoming fairly common, especially in Europe, though the field still has much to learn regarding optimum practices. Other interventions being investigated include in-ICU psychological support for patients, families, and staff. Intriguingly, stress doses of hydrocortisone in the ICU also appear to reduce the incidence of post-ICU PTSD phenomena, though glucocorticoids are not prescribed routinely for this purpose.

Children. To date, there has been limited research on interventions to prevent psychiatric morbidity in pediatric critical illness survivors. A randomized controlled trial investigated the efficacy of an emotional regulation and anxiety reduction program for mothers of critically ill children; the children whose mothers received the intervention exhibited fewer behavioral problems following their critical illness. Notably, ICU diaries have also been adapted for use in children.

Families. Results of two studies (one small randomized trial, one before-and-after design cohort study) suggest that ICU diaries prevent long-term PTSD symptoms in relatives. In addition, studies have found that enhanced communication strategies (e.g., informative brochures, longer care conferences) between critical care health care workers (e.g., physicians and nurses) and family members of dying critically ill patients reduce psychological distress, including PTSD, depressive and general

anxiety symptoms, in family members.

Interventions for Post–Critical Illness Cognitive Dysfunction. There have been few studies of interventions to prevent persistent cognitive dysfunction in critical illness survivors. Preliminary evidence supports a role for in-home cognitive and physical rehabilitation, though larger randomized controlled trials are needed.

Several interventions to reduce or prevent delirium in the critically ill have been evaluated. Dexmedetomidine, a central alpha-2 adrenergic agonist used for sedation with mechanical ventilation, is less deliriogenic than benzodiazepines. Early physical rehabilitation in the ICU, including during mechanical ventilation, reduces the incidence and duration of delirium. However, it remains unclear whether such delirium reductions will lead to better long-term cognitive outcomes.

Interventions for Post–Critical Illness Physical Dysfunction. The most promising interventions to reduce post-ICU weakness involve minimizing sedation and maximizing physical rehabilitation as early as possible, including while patients are critically ill. Such interventions reduce the need for prolonged mechanical ventilation and ICU stays and, as noted, delirium. Though, to the authors' knowledge, the effects of these interventions on long-term neuropsychiatric outcomes are unknown, they seem unlikely to be harmful and likely to be beneficial.

References

Bienvenu OJ, Colantuoni E, Mendez-Tellez PA, et al. Depressive symptoms and impaired physical function after acute lung injury: a 2-year longitudinal study. *Am J Respir Crit Care Med*. 2012;185:517–524.

Bienvenu OJ, Gellar J, Althouse BM, et al. Post-traumatic stress disorder symptoms after acute lung injury: a 2-year prospective longitudinal study. *Psychol Med*. 2013;43:2657–2671.

Colville GA, Pierce CM. Children's self-reported quality of life after intensive care treatment. *Pediatr Crit Care Med*. 2014;14:e85–e92.

Davidson JE, Jones C, Bienvenu OJ. Family response to critical illness: post-intensive care syndrome—family. *Crit Care Med*. 2012;40:618–624.

Davydow DS, Desai SV, Needham DM, Bienvenu OJ. Psychiatric morbidity in survivors of the acute respiratory distress syndrome: a systematic review. *Psychosom Med*. 2008;70:512–519.

Davydow DS, Gifford JM, Desai SV, Bienvenu OJ, Needham DM. Depression in general intensive care unit survivors: a systematic review. *Intensive Care Med*. 2009;35:796–809.

Davydow DS, Gifford JM, Desai SV, Needham DM, Bienvenu OJ. Posttraumatic stress disorder in general intensive care unit survivors: a systematic review. *Gen Hosp Psychiatry*. 2008;30:421–434.

Davydow DS, Hough CL, Langa KM, Iwashyna TJ. Pre-sepsis depressive symptoms are associated with incident cognitive impairment in survivors of severe sepsis: a prospective cohort study of older Americans. *J Am Geriatr Soc*. 2012;60:2290–2296.

Davydow DS, Hough CL, Langa KM, Iwashyna TJ. Depressive symptoms in spouses of older patients with severe sepsis. *Crit Care Med*. 2012;40:2335–2341.

Davydow DS, Richardson LP, Zatzick DF, Katon WJ. Psychiatric morbidity in pediatric critical illness survivors: a comprehensive review of the literature. *Arch Pediatr Adolesc Med*. 2010;164:377–385.

Desai SV, Law TJ, Needham DM. Long-term complications of critical care. *Crit Care Med*. 2011;39:371–379.

Fan E, Dowdy DW, Colantuoni E, et al. Physical complications in acute lung injury survivors: a two-year longitudinal prospective study. *Crit Care Med*. 2014;42:849–859.

Friedman JI, Soleimani L, McGonigle DP, Egol C, Silverstein JH. Pharmacological treatments of non-substance-withdrawal delirium: a systematic review of prospective trials. *Am J Psychiatry*. 2014;171:151–159.

Garrouste-Orgeas M, Coquet I, Périer A, et al. Impact of an intensive care unit diary on psychological distress in patients and relatives. *Crit Care Med*. 2012;40:2033–2040.

Gries CJ, Engelberg RA, Kross EK, et al. Predictors of symptoms of posttraumatic stress and depression in family members after patient death in the ICU. *Chest*. 2010;137:280–287.

Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability in survivors of severe sepsis. *JAMA*. 2010;304:1787–1794.

Jones C, Bäckman C, Capuzzo M, et al. Intensive care diaries reduce new-onset posttraumatic stress disorder following critical illness: a randomised, controlled trial. *Crit Care*. 2010;14:R168.

Jones C, Bäckman C, Griffiths RD. Intensive care diaries and relatives' symptoms of posttraumatic stress disorder after critical illness: a pilot study. *Am J Crit Care*. 2012;21:172–176.

Kross EK, Engelberg RA, Gries CJ, Nielsen EL, Zatzick D, Curtis JR. ICU care associated with symptoms of depression and posttraumatic stress disorder among family members of patients who die in the ICU. *Chest*. 2011;139:795–801.

Melnyk BM, Alpert-Gillis L, Feinstein MF, et al. Creating opportunities for parent empowerment: program effects on the mental health/coping outcomes of critically ill young children and their mothers. *Pediatrics*. 2004;113:e597–e607.

Needham DM, Davidson J, Cohen H, et al. Improving long-term outcomes after discharge from ICU: report from a stakeholders' conference. *Crit Care Med*. 2012;40:502–509.

Pandharipande PP, Girard TD, Jackson JC, et al. Long-term cognitive impairment after critical illness. *NEJM*. 2013;369:1306–1316.

Pandharipande PP, Pun BT, Herr DL, et al. Effect of sedation with dexmedetomidine vs. lorazepam on acute brain dysfunction in mechanically ventilated patients: the MENDS randomized controlled trial. *JAMA*. 2007;298:2644–2653.

Parker AM, Sricharoenchai T, Raparla S, Schneck KW, Bienvenu OJ, Needham DM. Posttraumatic stress disorder in critical illness survivors: a meta-analysis. *Crit Care Med*. 2015;43:1121–1129.

Schelling G, Roozendaal B, Krauseneck T, Schmoelz M, De Quervain DJ, Briegel J. Efficacy of hydrocortisone in preventing posttraumatic stress disorder following critical illness and major surgery. *Ann N Y Acad Sci*. 2006;1071:46–53.

Wilcox ME, Brummel NE, Archer K, Ely EW, Jackson JC, Hopkins RO. Cognitive dysfunction in ICU patients: risk factors, predictors, and rehabilitation interventions. *Crit Care Med*. 2013;41:S81–S98.

▲ 27.12 Psychocutaneous Disorders

ADARSH K. GUPTA, M.D., AND MOHAMMAD JAFFERANY, M.D.

INTRODUCTION

The field of psychodermatology or psychocutaneous medicine focuses on the interaction between the mind, the brain, and the skin. Skin is not only the largest organ of the body but also a metabolically active biological interface between the individual and the environment. Skin has a vital role in the development of the infant's maternal attachment at birth, has a major effect on body image, and remains a powerful organ of communication throughout the life cycle. Skin is an organ of expression, and responds to emotions with blushing, pallor, piloerection, and perspiration. Brain and skin originate from the same germ layer, the embryonic ectoderm, and are affected by the same hormones and neurotransmitters. Psychiatry is focused on the "internal invisible disease" and dermatology is focused on "external visible disease." The interaction between mind, brain, and skin is variable: psychopathological factors can play an etiological role in the development of skin disorders or exacerbate pre-existing skin disorders, while dermatology patients may suffer the psychosocial consequences of disfigurement. Various studies have reported significant psychiatric and psychosocial comorbidity in at least 30 percent of dermatological patients. Psychodermatology as a discipline encourages a comprehensive and humanistic approach to the management of dermatology patients with psychiatric comorbidity. The purpose of this chapter is to provide clinicians a background using a biopsychosocial model for clinical assessment and management of psychosomatic aspects of dermatological disorders. The chapter reviews the following: (1) the current classification of psychocutaneous conditions along with changes in new DSM-5; (2) findings regarding psychosocial stress in patients with skin disorders; (3) relevant aspects of psychoneuroimmunology (interactions between "stress" and the skin involving alterations of nervous, endocrine, and immune systems); (4) common psychocutaneous disorders; and (5) use of psychotropic medications in psychocutaneous disorders.

CLASSIFICATION

Although there is no single universally accepted classification system of psychocutaneous disorders, the one most widely accepted was devised by John Koo. It includes five different groupings listed in Table 27.12–1, which also includes the text from the revised fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) classification of psychocutaneous disorders. *Psychophysiological disorders* are bona fide

skin disorders exacerbated by stress. The term *primary psychiatric disorders* refers to cases without "real" skin disease, but where patients present with serious psychopathology and visible skin lesions are self-induced. This category is the most typical of psychodermatology disorders that gets a dermatologist's attention. The term *secondary psychiatric disorders* refers to cases in which patients develop psychological problems as a result of skin disease and associated disfigurement. Some skin conditions, such as acne and alopecia areata, although not life-threatening, may have severe psychosocial impact on patient's quality of life. The term *cutaneous sensory disorders* refers to cases in which patients have unpleasant sensations on the skin, such as itching, stinging, burning, or crawling, with no proven skin-based etiology, and in which a psychiatric diagnosis may or may not be evident. These are considered to be equivalents of chronic pain syndromes. These subcategories of psychocutaneous disorders are not mutually exclusive. For example, acne can be precipitated by stress, a psychophysiological condition, but acne can also cause social anxiety and depression due to disfigurement, which would be categorized as a secondary psychiatric condition. Koo's classification includes a separate category to describe the *use of psychotropic medications* in treating certain skin conditions, which may be more efficacious than traditional dermatological treatments. For example, the antidepressant doxepin (Sinequan) is a more powerful antipruritic agent than traditional antihistamines. For dermatologists it may be very important to make such a distinction between the above diagnostic categories, which can help them to select an optimal treatment approach.

PSYCHOSOCIAL STRESS AND COPING IN SKIN DISORDERS

The clinical experience of dermatologists suggests that most patients are able to make appropriate adjustment to their skin disorder and do not encounter undue distress. However, some patients with chronic skin illnesses and appearance-altering conditions run the risk of social, psychological, and physical distress. The patient's experience of psychosocial distress is variable and depends on (1) the characteristics of skin disorder itself, (2) the individual characteristics of patients and his or her life situations, and (3) cultural attitudes related to skin disease (often expressed as stigma).

Table 27.12–1.
Classification of Psychocutaneous Disorders

Koo's Classification	DSM-5 Classification
Psychophysiological disorders	**(1) Psychological factors affecting other medical condition (316)** Atopic dermatitis, psoriasis, alopecia areata, urticaria and angioedema, acne vulgaris

Primary psychiatric disorders	**(2) Delusional disorder, somatic type (297.1)** Delusional parasitosis, delusions of a defect in appearance, delusions of a foul body odor **(3) Obsessive-compulsive and related disorders (OCRD)** Psychogenic excoriation (698.4), trichotillomania (312.39), body dysmorphic disorder (300.7) **(4) Unspecified OCRD (Body-focused repetitive behavior disorder) (300.3)** Onychophagia **(5) Factitious disorder (300.19):** Dermatitis artefacta, psychogenic purpura
Secondary psychiatric	(6) Anxiety disorder due to another medical condition (293.84) (7) Generalized anxiety disorder (300.02) (8) Major depressive disorder (296.3)
Cutaneous sensory disorders	**(9) Unspecified somatic symptom and related disorder (300.82)** Chronic idiopathic pruritus, idiopathic pruritus ani, vulvae, scroti **(10) Somatic Symptom Disorder, with predominant pain (300.82)** Idiopathic glossodynia, dysesthetic (essential) vulvodynia
Psychotropic medication for nonpsychiatric indications	

From American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013, and Koo J. Psychodermatology: a practical manual for clinicians. *Curr Probl Dermatol*. 1995;7:203–225.

Characteristics of the Dermatological Disorder

The emotional reaction to a particular skin condition is variably affected by the patient's understanding of its *origin*. A congenitally acquired condition such as port wine stain may invoke a different reaction, compared to that evoked by an acquired skin disorder such as a sexually transmitted disease, contact dermatitis, lichen simplex, and prurigo nodularis. For example, the latter two are induced by scratching and may invoke a sense of responsibility and guilt. In genetically predisposed conditions like atopic dermatitis and psoriasis, the patient's response may include blaming parents, as well as feeling unfortunate, frustrated, or helpless. The external *appearance of skin lesions* can lead to different degrees of disfigurement. Skin lesions might be red, bullae, oozing, crusted, or scaling, or might have leathery thickening. *Associated symptoms* of persistent intolerable itching, burning, or pain may add to distress, insomnia, anxiety, and depression. The *location of skin lesions* might affect the patient's self-consciousness and how he or she is perceived by others. Lesions directly visible on the face and hands usually cause the greatest concern. Lesions present on the trunk and limbs can be a source of concern in public areas such as a gym, health clubs, or the beach. Genital lesions can be particularly embarrassing and interfere with sexual intimacy. Hair loss can diminish self-confidence and affect social poise. Lesions involving the nails and hands may be

disabling in jobs involving transactions with other people, as in banks, bakeries, or retail stores. The *timing of skin disease onset* can effect psychological development, especially in children. Congenital conditions like port wine stain or childhood atopic dermatitis influence the child's self-image and general emotional development. In adolescents, chronic skin conditions can be a significant blow to self-esteem and image, affecting the maturation of self, sexual identity, and personality. The *course of the skin illness*, with a chronic relapsing condition and treatment-related side effects, also has a major bearing on patients and their families on a day-to-day basis.

Individual Characteristics

Age and sex are obviously important when considering a person's ability to cope with a chronic illness. A young woman suffering from psoriasis is likely to be far more troubled emotionally than an older man with a similar condition. Personality can influence a patient's reactions and subjective experience of the illness and coping. Patients with a narcissistic personality may have "an expectation of admiration and attention, sense of entitlement, and excessive concern about how others regard them." Such patients may be humiliated by skin disease that changes their personal appearance. Patients with borderline personality disorder may consider skin disease as a threat to self-image and autonomy and react with unstable moods, severe anxiety, and fear of abandonment. If a person has an obsessive-compulsive trait, the skin disorder may invoke feeling of lack of control over his or her body and excessive disgust with feelings of being unclean or dirty due to skin lesions and topical treatments. The presence of an Axis I psychiatric disorder such as depression, obsessive-compulsive disorder (OCD), or psychosis can influence the patient's cognition, thoughts, and beliefs about the skin lesions. A patient with delusions of parasitosis may have delusional beliefs about exposure to parasites and specific details about the infestation. Major depressive disorder is one of the most commonly encountered psychiatric disorders in dermatology patients. Depression may increase the sensation of itching in pruritic skin disorder such as atopic dermatitis, psoriasis, and chronic idiopathic urticaria (Fig. 27.12–1).

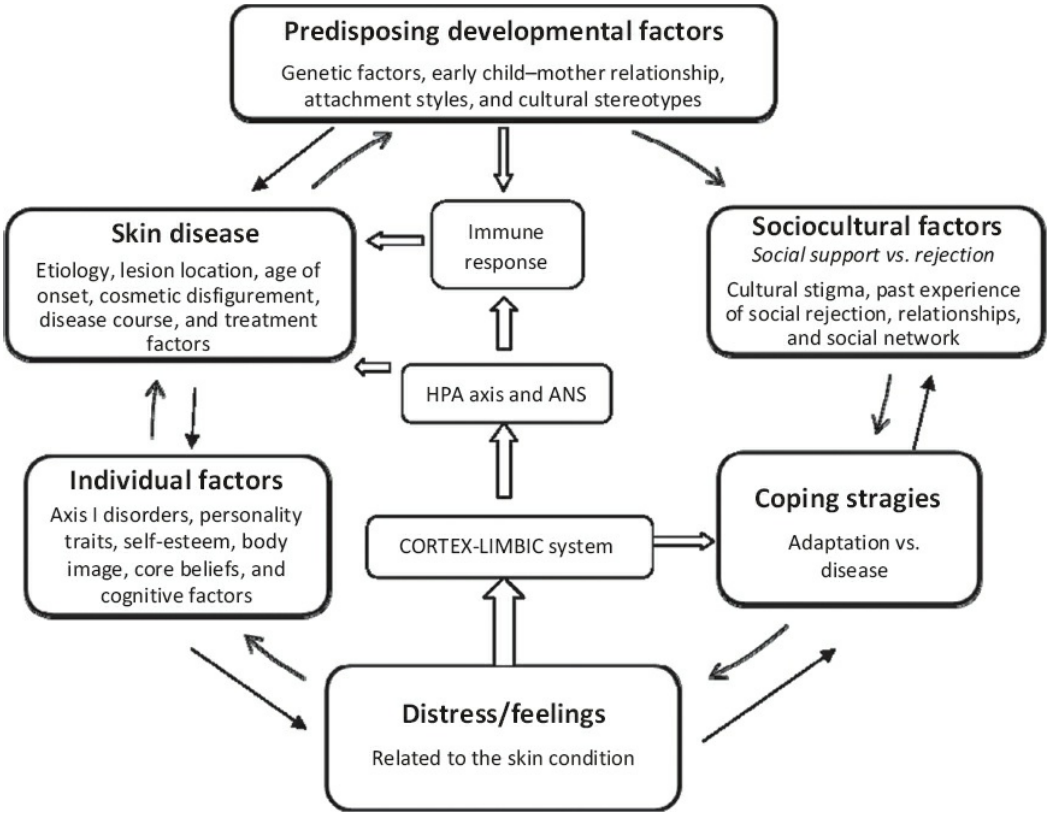

FIGURE 27.12–1. Model of psychosocial factors implicated in origin and maintenance of skin-specific affect/distress, and systemic adaptation response. (Adapted and modified from Thompson A. Coping with chronic skin conditions. In: Walker C, Papadopoulos L, eds. *Psychodermatology: The Psychological Impact of Skin Disorders*. Cambridge, UK: Cambridge University Press; 2005:61.)

Disease-Related Stress in Childhood. Cutaneous stimulation plays an important role for cell growth and differentiation and CNS maturation throughout infancy. One controlled study reported that human preterm neonates who received tactile and kinesthetic stimulation gained more weight, had more mature neurological reflexes, and were more alert than controls. In another study with monkeys, Harry Harlow showed that deprivation of touch and physical comfort in early life contributed to the development of aggressive behaviors and other psychological problems. These and other findings emphasize the role of cutaneous perception and stimulation in physical and psychological development and possible implications of skin diseases in child–parent interactions and child's early development.

Childhood skin diseases like atopic dermatitis can influence a child's psychological development and personality. An infant with unhealthy skin may experience two contrary emotional stimuli from the skin: the nurturing attention gained from stroking and massaging with skin ointments, and at the same time the sensations of pain, itching, and burning from skin lesions (caused by too much pressure on skin lesions and other environmental discomforts). The presence of the mother and her

applying ointment is experienced as the only source of relief and comfort, leading to fixation on the maternal object, so that separation from the mother is experienced as a threat and leads to exacerbation in perceptions of pain, itching, and discomfort. Psychosomatic practitioners have termed this psychodynamic model as *allergic object relationship*. The sick child afflicted with severe itching also suffers from disturbance in sleep rhythms, special dietary restrictions, and lack of cuddling and rocking from the mother. In their desperation to comfort their child some parents may become excessively protective and indulgent, turning the child into a little tyrant. The restless behavior of children and their poor frustration tolerance often strains the family. At school small children with dermatological problems may often face rejection based on visible skin lesions and feel harassed by teasing from other children.

The Impact of Skin Disease on the Mother–Child Relationship. Early studies on mother–child relationships in atopic dermatitis indicated that maternal rejection was a common experience for the children, defining the rejecting mother as one "who considered the child as a burden and whose behavior is affected by conscious or unconscious desire to get rid of the child." While recent studies acknowledge that mothers of children with atopic dermatitis react less emotionally and spontaneously to the child's emotion than mothers in comparison groups, they have not been able to confirm the earlier hypothesis that the "atopic" child's experience of maternal rejection is a result of "pathogenic" maternal personality. Similar studies have led to the conclusion that limited communication between mother and child was a consequence of skin disease rather than a cause. H. P. Langfeldt and K. Luys studied the mother's attitude of chronically ill, recently afflicted, and healthy children. They concluded that "a hostile rejecting attitude towards childbearing" can best be interpreted not as a cause of onset of disease, but more likely as its consequence because the temporal relationship and chronic nature of some skin disease can overtax the adaptive capacity of the mother and modify her spontaneous responses.

Body Image and Self-Schema. T. F. Cash and T. Pruzinsky concluded that *conceptualization of body image* is multifaceted and consists of perceptions, feelings, and associated thoughts about the body and that these are intertwined with core feelings about the self. Perceptions related to body image also have a social component with interpersonal meaning and cultural physical aesthetics. Self-schema has many reference points, including socially represented idealized body (ideals represented from media, culture, peers, and family), objective body and internalized ideal body (a compromise between the objective body and the socially represented ideal). The greater the discrepancy between the perceived body and idealized body, the more body image dissatisfaction occurs, which is

associated with lower self-esteem. The term *cutaneous body image* (CBI) refers to individual mental perception of one's skin, hair and, nails. CBI is a highly subjective construct consisting of developmental, psychosocial, cultural, and psychiatric factors. Clinical assessment of CBI dissatisfaction can provide valuable information about a patient's quality of life, psychiatric comorbidity like body dysmorphic disorder, self-induced dermatoses, and risk for suicide. In patients with a chronic disfiguring skin disorder, perception of CBI may vary depending upon the episodic nature, appearance, and visibility of skin lesions. Avoidant and concealment behaviors, compulsive rituals, depression, and social anxiety may also result. Poor CBI is also implicated in lack of adherence to treatment in chronic skin disorders, as psychosocial burden and distress may cause patients to lose confidence in treatments.

Skin Disease and Relationships. A focus on how a person looks or want to appear in the eyes of the others, can powerfully influence the chemistry of relationships, especially in their earlier stages. The onset of skin disease can result in physical disfigurement, discomfort, embarrassment, and social stigma, which can impact the individual functioning and change the matrix of social interaction. Feelings of shame and increased self-consciousness can alter an individual's sense of self, and jeopardize security in a relationship. Ideally partners should support each other in health and sickness, but in the context of chronic illness couples are often prone to more emotional conflicts, lack of support, and relationship dissatisfaction. Coping with the complexities of a skin disorder can test the commitment that couples have towards each other by challenging both partners to work within new parameters.

ATTACHMENT STYLES. Original attachment theory by John Bowlby, which explained the emotional bond between infants and caregivers, was further developed by Cindy Hazan and Philips Shaver in 1987. They noted that major attachment styles described in infant literature, namely "secure, avoidant and anxious-ambivalent" were similar to how adults related in intimate relationships. Like infants, similar behavioral patterns were observed in adults when they perceived themselves to be psychologically separated from their partner. They stated that romantic love can be conceptualized as an attachment process, but unlike infant attachment it involves reciprocal behavioral systems of care giving and sex. Attachment styles are considered to be *working models* that shape cognitive, emotional, and behavioral responses to others. These working models operate largely outside awareness, are considered relatively stable and resistant to change in 70 percent of individuals, and center around regulation and fulfillment of attachment needs. In comparison to insecure attachment, couples with secure attachment are found to have higher levels

of trust, as well as greater ability to communicate effectively and support each other. The attachment framework elucidates why some individuals with a skin disorder use an intimate relationship as a support to deal with disease-related stress, like feelings of shame and sexual problems, while other individuals struggle to communicate or avoid sharing their distress. S. Kellet described patient responses to disfigurements as a specific form of body shame and termed it "dermatological shame." The focus is on the appearance of skin and relative attractiveness. In individuals with secure attachment the dermatological shame may be disease-specific, mostly confined to the disease itself, and not generalized to self-schema (belief about self). Such individuals may be more accepting of the disease and accept their physical and psychosocial issues in a healthy way. The anxious-ambivalent person may experience more pervasive generalized shame, which merges with self-schema to create an inner experience of self-disgust, negative beliefs and social stigma. Avoidant persons are more likely to minimize display of emotional distress in order to hide their shame.

COMMUNICATION DIFFICULTIES. Recent studies on the relationship between attachment style and self-disclosure in couples indicate that secure and anxious-ambivalent individuals report more self-disclosure than avoidant adults. The communication patterns of individuals with secure attachment show greater flexibility and emotional reciprocity over a range of issues. In individuals with skin disease, the change in a couple's social network may occur through reduced interactions with others and increase in homebound activities. The avoidance may have a protective function but it may also lead to loss of friends and limited social network. A British study reported that 84 percent of 186 psoriasis patients conveyed that social isolation was the worst aspect of their illness, based on their social difficulties.

SKIN DISEASE AND SEXUAL INTIMACY. The relationship between body image and sexual intimacy is a good illustration of the reciprocal nature of interpersonal influences in patients with skin disease. A negative body image can damage the perception of self as sexual being and thus disrupt intimacy in relationship. A recent study focused on sexual activity of 120 patients suffering from psoriasis and found that 40 percent experienced a decline in sexual activity and had higher scores on depression than non-affected persons. In 2003, a US marketing firm surveyed 502 patients with psoriasis and reported that 38 percent had difficulties with sex and intimacy.

Stigma about Skin Diseases

In patients with dermatological disorders, the analogy of "skin failure" (comparable to failure of other organ systems in the body like heart or

kidney failure) implies failure of skin as an organ of physical expression, thus leading to rejection, shame, isolation, and psychosocial impairment. I. H. Ginsberg and B. G. Link explored the feelings of stigmatization in patients with psoriasis, and a factor analysis of their data implied that belief about stigmatization might be grouped into six dimensions: *anticipation of rejection* (e.g., "I feel physically and sexually unattractive when my psoriasis is bad"); *sensitivity to opinions* of others (e.g., "Sometimes I feel outcast because of my psoriasis"); *feeling of being flawed* (e.g., "I believe other people think that psoriasis patients are dirty"); *guilt and shame* (e.g., "My psoriasis is a source of deep humiliation and embarrassment for me and my family"); *secretiveness* (e.g., "I try not to share with family members away from me, that I have psoriasis"); and *a more positive attitude* which is unaffected by negative reactions of others (e.g., "If my child has to develop psoriasis, he could develop his potential as though he did not have it"). A previous experience of rejection was considered to be strong predictor of such beliefs. Some individuals with skin disorders often believe that their main difficulties arise from others' reactions to their disease, rather than the disease itself.

People with skin disease tend to conceptualize themselves as having no control over the surface of their bodies**.** Psychoanalytic studies have described skin as the boundary of the self, which encloses whatever is supposed to remain closed off inside. In unconscious terms, skin is supposed to wall off dirt, emerging from the body either in reality or symbolically. Bodily products are perceived as dangerous and destructive. Sigmund Freud referred to dirt as "matter out of body." Those affected with skin lesions in form of bumps, redness, scales, or exudates are unable to control their own "dirt" and symbolically their own "dirty impulses." Infections associated with visible skin disease are assumed to be sexual in nature. The unconscious guilt about sexual urges may imply symbolically lack of control over inner dirt or "dirty thoughts and wishes" and skin lesions perceived as sign of impurity, contagion, danger, and punishment.

It appears part of normal narcissism to "want to be looked at and admired." However in individuals with skin lesions, to be looked at with surprise, fear, or disdain evokes feelings of painful shame and humiliation. People with normal skin are commonly inclined to turn away physically or psychologically from those suffering from skin diseases. Iona Ginsburg suggested that this may involve "a fleeting identification" with the afflicted person, followed by a negative empathic response. The visible skin lesion symbolizes how defective and vulnerable one can be, and when coupled with observed feelings of shame and embarrassment, invokes a momentary flash of empathic identification. The healthy person proceeds to block out and dismiss from their awareness any compassion for the sufferer's pain, and then turns away and distances themselves from the person with skin disease. A similar phenomenon is observed in psoriasis patients, who

typically report that they would avoid other people with psoriasis if they did not have active disease.

PSYCHONEUROIMMUNOLOGY OF SKIN DISEASES

The use of the term psychoneuroimmunology has gained increasing attention since R. Adler and N. Cohen demonstrated behaviorally conditioned immunosuppression. The association between psychosocial stresses and exacerbation of dermatological conditions is well documented in medical literature. A. Picardi and D. Abeni reviewed much of the literature with scientific rigor to examine the available evidence for the role of stress in dermatological conditions. They found that evidence for the role of stress in exacerbation of psoriasis, atopic dermatitis, urticaria, and alopecia areata was highly convincing. More recent work from F. S. Dhabhar and others has recognized how severity of stress interacts across the "stress-immune spectrum" and leads to genesis of immunoprotection, immunopathology and immunoregulation in psychocutaneous disorders. This section will review the functional relationships between stress and skin disorders, mediated by the brain through a complex integration of endocrine, autonomic, and immune systems. The field of psychoneuroimmunology intends to validate the "psychosomatic" link in cutaneous disorders by understanding the potential pathophysiological mechanisms underlying the role of stress in the exacerbation of skin diseases.

Stress and Adaptation Response

Stress is a ubiquitous part of life and generally considered to have a negative connotation. There are numerous definitions of the word "stress," but an integrated definition states that stress is a constellation of events consisting of exposure to stimulus (stressor) that precipitates a reaction in the brain (stress perception), which in turn activates physiological fight-or-flight systems in the body (stress response). Hans Selye described the term "general adaptation response" in 1936, who proposed that organisms have the ability to adapt to acute changes and maintain homeostasis (e.g., fight-or-flight response). However, the chronicity and severity of stress lead to exhaustion of adaptation mechanisms and ultimately a disease response. Adaptation in the face of stressful situations and stimuli involves activation of neural, neuroendocrine, and neuroendocrine-immune mechanisms. This adaptation has been called "*allostasis*" or "maintaining stability through change," which is an essential component of maintaining homeostasis (B. S. McEwen). "*Allostatic load*" refers to the price the body pays for being forced to adapt to adverse psychosocial or physical situations, and it represents either the presence of too much stress or the inefficient operation of the stress hormone response system, which must be turned on

and then turned off again after the stressful situation is over. Conditions of high allostatic load are likely to result in dysregulation or suppression of the immune function. (The specific nature of stress in individuals with cutaneous disorders is described earlier under psychosocial aspects of skin diseases.)

"*Resilience*" refers to the capacity of psychophysiological systems to recover from challenging conditions and allow the organism to efficiently return to its health maintenance equilibrium (allostasis). *Psychological resilience* is mediated by factors such as early experiences, genetics, coping mechanisms, optimism, social support, and learning. *Physiological resilience* may be mediated by genetics, neuroendocrine reactivity, sleep, environment, nutrition, etc. (Fig. 27.12–2). The duration and magnitude of physiological stress response is determined by individual's stress perception and processing by the brain and mechanisms mediating psychological and physiological resilience. Psychogenic stressors are very important in human beings because they can generate stress response long after stressor exposure (e.g., PTSD, chronic psychosocial conflicts). Psychological resilience mechanisms are crucial in humans because they can limit the duration and magnitude of chronic stress response. Stressor exposure and its processing by the brain leads to a "physiological stress response," which may consist of acute or chronic physiological activation (neurotransmitters, hormones, and their molecular, cellular, organ-level, and systemic effects). These result in psychophysiological states with different effects on immune function and overall health.

The Stress-Immune Spectrum

It is important to acknowledge that a psychophysiological stress response is one of nature's fundamental survival mechanism and stress response can have salubrious adaptive effects in the short run. However, chronic and long-lasting stress can be harmful and affect health adversely. To reconcile these seemingly dichotomous effects of stress, F. S. Dhabhar proposed that stress response and its effects on immune function be viewed in the context of a "stress spectrum." Acute stress or *eustress* is represented on one end of this spectrum, where conditions of short duration stress result in immunopreparatory or immunoenhancing physiological conditions. Eustress is characterized by rapid physiological stress response triggered by a stressor, followed by a rapid shutdown of response upon termination of the stressor. The opposite end of this stress spectrum is characterized by chronic stress or *distress*, where repeated or prolonged stress may result in dysregulation or suppression of immune function. This results in overall increased exposure to stress hormones because physiological stress response either persists long after cessation of stressor, or it is activated repeatedly over a longer duration. This model proposes that acute or chronic stress is generally superimposed on a psychophysiological health

maintenance equilibrium (Fig. 27.12–2).

The following sections focus on the stress-induced body's response at various levels, including the skin's sensory perception, the neuroendocrine and autonomic systems, and immunomodulation.

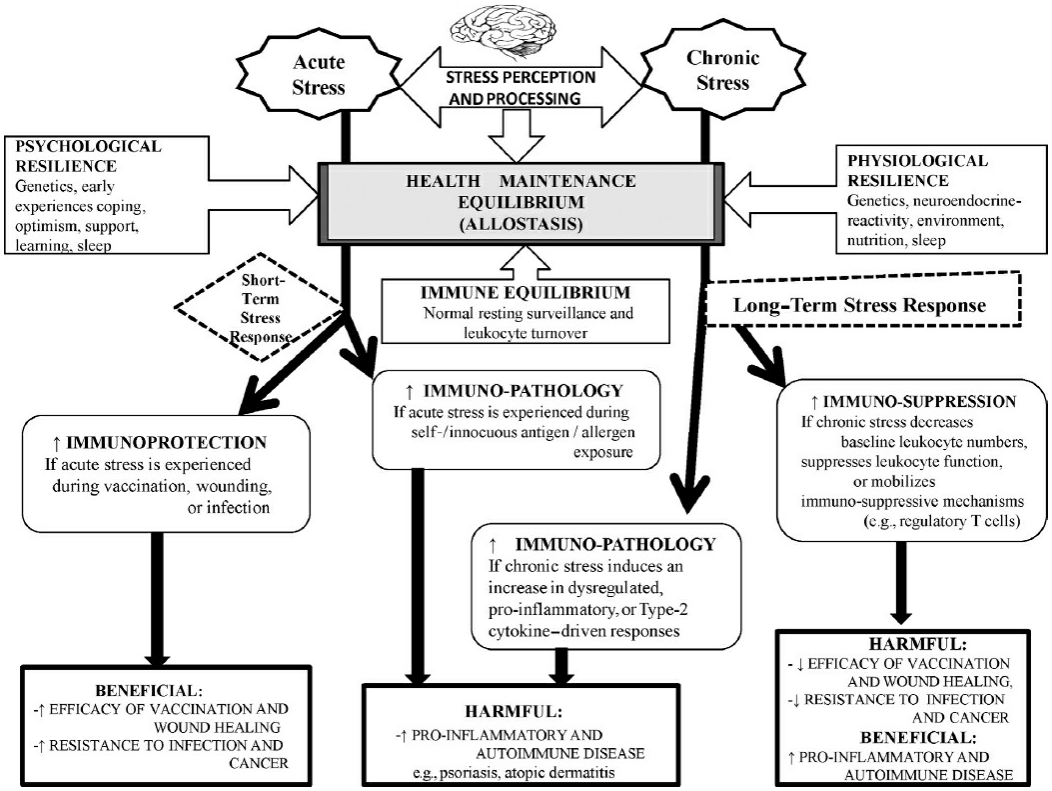

FIGURE 27.12–2. The stress-spectrum model describes basic allostasis and resilience mechanisms, and effects of acute versus chronic stress on immune and health outcomes: immunoprotection, immunopathology, and immunoregulation/suppression. (Adapted and modified from Dhabhar FS, Psychological stress and immunoprotection versus immunopathology in the skin. *Clin Dermatol*. 2013;31:18–30.)

Skin and Sensory Perception

Skin has the same embryological origin from ectoderm as the CNS. It employs large populations of cells in a coordinated manner to respond to both external and internal challenges. It consists of epidermal cells, keratinocytes, dendritic cells, melanocytes, and so forth, which function along with dermal structures such as macrophages, mast cells, leucocytes, and dermal dendritic cells to provide the immune response. *Langerhans cells* are dendritic cells which originate in bone marrow and migrate to the epidermis. They act as frontline defense of the immune system and present invading antigen to the lymphoid system. *Merkel cells*, derived from neural crest, are specialized for perception of light touch. The sensory perception of pain, pressure, and temperature is conveyed via *afferent sensory bodies* like Meissner corpuscle, Pacinian corpuscles, and free nerve endings to afferent sensory fibers. These sensations are relayed via unmyelinated C

fibers and myelinated A fibers to axons ending in the dorsal root ganglia and then via ascending spinal pathways to the thalamus and sensory cortex in the postcentral gyrus. Unmyelinated C fibers are also specialized to serve efferent function by secreting neuropeptides (NPs). NPs are simple compounds of 40 or fewer amino acids that contribute to neurogenic inflammation (vasodilatation, plasma exudation, and migration of leukocytes, as a result of antidromic stimulation of dorsal root) by their action on specific neuropeptides receptors and on various immune cells. The major NPs are *substance P* (SP), *calcitonin gene-related peptide* (CGRP), and *vasoactive intestinal peptide* (VIP). Other significant NPs are neuropeptide Y, somatostatin, neurokinin A & B, and the opiomelanocortins, which include endorphins. Regions with a high density of nerve endings include the genitalia, finger pads, lips, and other regions of the face innervated by the three branches of the trigeminal nerve. It is noteworthy that the face and scalp tend to be some of the most frequent sites for body-focused repetitive behavior disorders such as skin picking and trichotillomania. A brief summary of major NPs and their specific target cells and cytokine responses are indicated in Table 27.12–2.

Neuroendocrine and Autonomic Nervous System's Response

Primary sensory stimuli are integrated with corticolimbic input and produce an active recognition at cognitive and subconscious levels (stress perception). The stress response is determined by the individual's interpretation of stimuli as distressful and not by the nature itself of stimuli. A perceived stressful stimulus triggers the neuroendocrine pathway via the HPA axis. The parvicellular neurons in the hypothalamic paraventricular nuclei secrete corticotrophin-releasing hormone (CRH) and arginine vasopressin (AVP) into the hypophyseal portal system. These travel to the anterior pituitary and release adrenocorticotrophic hormone (ACTH). Pituitary ACTH stimulates the zona fasciculate cells of the adrenal cortex to secrete cortisol into the circulation. Cortisol, through a classic feedback mechanism, suppresses ACTH, CRH, and AVP. The autonomic nervous system is activated concurrently and modulates the activity of the HPA axis. The locus ceruleus neurons in the brain stem maintain reciprocal connections with CRH and AVP, producing neurons in paraventricular nuclei. The locus ceruleus plays a major role in this activation through: (1) control of the autonomic output by noradrenergic brain stem nuclei, which then activates the paraspinal sympathetic ganglia to produce norepinephrine and neuropeptides, and (2) stimulation of adrenal medulla to secrete catecholamines into the circulation. The autonomic stress system is activated by serotonergic and cholinergic input and inhibited by GABA, benzodiazepine systems, and pro-opiomelanocortin (POMC) peptides. The POMC neurons in the arcuate nucleus have reciprocal connections with CRH-secreting neurons in the hypothalamus (thus modulating the HPA

axis), and also pain control areas in the hindbrain and spinal cord (thus enhancing analgesic effect in stressful states). The endocrine and increased sympathetic output response to stress invokes a multisystem reactivity to engage in flight-or-fight response.

Table 27.12–2.
Neuropeptides and Their Effects on the Various Skin Cells

	Neuropeptides		
Target Cells	**SP**	**CGRP**	**VIP**
Endothelial	Proliferation, ↑ permeability	Proliferation, IL-8 secretion	Proliferation
Keratinocyte	Stimulates IL-1, IL-8, comitogen with CGRP, LTB4	Proliferation in association with SP	Mitogenic proliferation
Lymphocyte	Stimulates proliferation and synthesis of IL-2, ↑ IgA and IgM, B-cell differentiation	Chemotactic T-cell proliferation	Inhibits NK-cell activity
Macrophage	Stimulates IL-1, IL-6 secretion	Impairs antigen presentation	—
Mast cell	Proliferation, histamine and TNF-α release	Degranulation: histamine release	Histamine release
Monocyte	↑ Chemotaxis, phagocytosis, arachidonic acid synthesis	—	Enhanced migration
Neutrophil	Stimulates chemotaxis, phagocytosis	Enhances chemotaxis	—

↑, increase; SP, substance P; VIP, vasoactive intestinal peptide; CGRP, calcitonin gene-related peptide; Ig, immunoglobulin; IL, interleukin; NK, natural killer; TNF-α, tumor necrosis factor α.

Immunomodulation: Brain–Cytokines Link

The Immune Triad: Immunoprotection, Immunopathology, and Immunoregulation. Immune responses have been categorized in terms of principal cellular and molecular components—e.g., innate, adaptive, TH1, TH2, and cytokines. In addition to these categories, F. S. Dhabhar proposed to define immune responses in terms of their end-effects. Accordingly, immune responses can be categorized as immunoprotective, immunopathological, and immunoregulatory/suppressive. *Immunoprotective responses* tend to promote wound healing, eliminate viral infection and cancer, and mediate vaccine-induced immunological memory. *Immunopathological responses* are those directed against self (autoimmune disease like psoriasis, multiple sclerosis, lupus, etc.) or other antigens (skin hypersensitivities, allergies, asthma), as well as nonresolving inflammation. Immunopathology can also result in low-grade chronic increase in local and/or systemic inflammatory mediators. "*Immunoregulatory/suppressive*" responses are defined as those involving immune cells and cytokine factors that regulate/inhibit the function of other immune cells (Fig. 27.12–2).

There are significant individual variations in stress perception and in processing and coping mechanisms, which can have variable effects on kinetics, peak levels, and duration of circulating stress hormones. The factors that determine outcome of stress-immune reaction and potential health consequences include: (a) duration of stress; (b) effect of stress on leukocyte distribution in the body; (c) differential effects of physiological versus pharmacological concentration of glucocorticoids; (d) timing of stressor or stress hormones exposure in relation to time of activation of immune response; and (e) other factors including age, genetics, circadian cycle, nature of immunizing antigen, etc.

Acute Stress Response: Immunoenhancement. The stress-induced neuroendocrine and autonomic changes have profound effects on immune regulation by modulatory effects of glucocorticoids, catecholamines, and neuropeptides. Because virtually every cell in the body expresses receptors for one or more of these factors, stress hormones can induce changes in almost all cells and tissues and inform them about the presence of a stressor. Glucocorticoid and catecholamine hormones are major mediators of acute stress and induce rapid and significant changes in *leukocyte distribution* by mobilizing them from lymphoid stores (spleen, lung, marginated pool, and other organs) to blood vessels and lymphatics. This results in an increase in blood leukocyte numbers, especially the natural killer (NK) cells and granulocytes. As the stress response continues, activation of the HPA axis results in the release of glucocorticoid hormones that induce leukocytes to exit the blood and take position at potential "battle stations" (skin, mucosal lining of gastrointestinal and urinary-genital tracts, lung, liver, and lymph nodes) in preparation for immune challenges which may be imposed by the actions of the stressor. Such a redistribution of leukocytes results in a decrease in blood leukocyte numbers and leads to enhancement of immune function in skin and subcutaneous tissues. Depending on the primary chemoattractants driving an immune response, acute stress may selectively mobilize specific leukocyte subpopulations into sites of surgery, wounding, or inflammation. Such a stress-induced increase in leukocyte trafficking may be an important mechanism by which acute stressors alter the course of different (innate versus adaptive, early versus late, acute versus chronic) protective or pathological immune responses.

F. S. Dhabhar and B. S. McEwen investigated skin immune reactions during stress in animals and reported on the role of interferon gamma (IFN-γ) in immunoenhancement. Accordingly, the acute stress hormones and catecholamines modulate the effects of IFN-γ on monocytes, resulting in enhanced cell-mediated immunity (CMI). In addition to IFN-γ, other cytokines like TNF-α, MCP-1, MIP-3α, IL-1, and IL-6 have also been associated with a stress-induced enhancement of skin cell–mediated

immunity. Recent animal studies found that short-term stress experienced at the time of immune activation enhances dendritic cell, neutrophil, macrophage, and lymphocyte trafficking, maturation, and function, and has been shown to augment innate and adaptive immunity. Leukocyte infiltration was driven by increased levels of the type-1 cytokines, IL-2 and IFN- γ, and TNF-α, observed at the site of antigen re-exposure in animals that had been stressed at the time of primary immunization. These studies demonstrated that a relatively mild behavioral manipulation can enhance an important class of immune responses that mediate harmful (allergic dermatitis) as well as beneficial (resistance to certain viruses, bacteria, and tumors) aspects of immune function (Fig. 27.12–3).

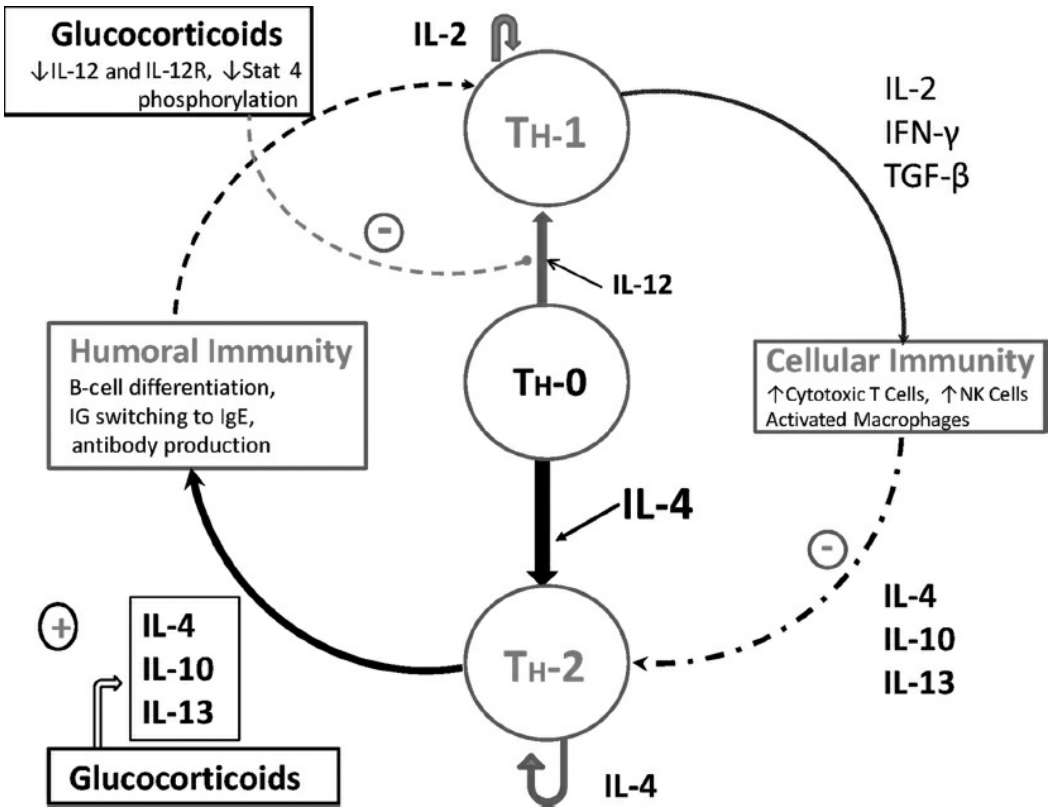

FIGURE 27.12–3. Glucocorticoids modulation of humoral and cellular immunity. Glucocorticoids modulate a cytokine shift in favor of humoral immunity. Glucocorticoids inhibit the expression of interleukin (IL)-12 and its receptor (IL-12R) by T and natural killer (NK) cells and suppress IL-12–induced phosphorylation of Stat4, a key translation factor in the expression of TH1 cytokines interferon γ (IFN-γ). Glucocorticoids also potentiate the secretion of key TH2 cytokines, such as IL-10. (From Koo JYM, Lee CS, eds. *Psychocutaneous Medicine*. New York: Marcel Dekker; 2003.)

Chronic Stress Response: Immunosuppression. At physiological levels, glucocorticoids contribute to immunoenhancement by causing a shift in cytokines, which promotes a change from CMI to enhanced humoral immunity. However, during chronic psychological stress, glucocorticoids are the key mediators of overall immunosuppressive response. During chronic stress, steroids have wide-ranging anti-inflammatory effects such

as leukocyte adhesion, margination and migration, B-cell maturation and plasma antibody production, macrophage activation and antigen presentation, and T-cell activation and cytotoxicity. Glucocorticoids contribute to this immunological dichotomy by mediating the key cytokines IL-10 and IL-12, which modulate T-helper cell response. T-helper lymphocytes are subclassified into TH0 (non-antigen-educated), TH1 (promotes CMI), and TH2 (promotes humoral immunity). The cytokine IL-12 stimulates TH0 cells to differentiate into TH1 cells, which secrete IL-2, IFN-γ, and transforming growth factor β (TGF-β). These factors promote a CMI response, consisting of increase in cytotoxic T cells, NK cells, and activated macrophages. IL-2 stimulates clonal expansion of TH1 cells, and IL-12 and IFN-γ tend to suppress the TH2 humoral immune response. The TH0 cells can be stimulated by cytokine IL-4 to differentiate into TH2 cells, which promote the humoral immunity by secretion of IL-4, IL-10, and IL-13. The humoral immunity changes involve stimulation of B-cell differentiation, immunoglobin (Ig) switching to IgE, and antibody production. IL-4 and IL-10 inhibit the TH1 immune response. The glucocorticoids mediate the switch from TH1 to TH2 by inhibiting the production of IL-12, downregulation of IL-12 receptor expression in T and NK cells, and inhibiting IL-12-induced phosphorylation of Stat4 (a key transcription factor in the expression of TH1-produced cytokine IFN-γ). Cytokine IL-12 plays the key role in promoting TH1 response and inhibiting TH2 at the same time. Glucocorticoids also promote the secretion of IL-10 from TH2 cells. In addition to the above effects of glucocorticoids, they also have direct effects on cutaneous permeability barrier homeostasis. This may be another contributing factor in the pathogenesis of inflammatory skin diseases like psoriasis and atopic dermatitis. The catecholamines also cause a similar shift by inhibiting IL-12 and enhancing IL-10. They exert a tonic inhibition of TH1 immunity mediated by β-adrenoreceptors on TH1 cells.

In summary, chronic stress suppresses or dysregulates innate and adaptive immune responses by altering the type-1/type-2-cytokine balance, inducing low-grade chronic increases in proinflammatory factors, and suppressing numbers, trafficking, and function of immunoprotective cells. This may exacerbate immunopathology (e.g., psoriasis, atopic dermatitis) if the enhanced immune response is directed against innocuous or self-antigens, or if the system is chronically activated, as seen during long-term stress. The stress spectrum, taken together with the preceding discussion, shows that the duration, intensity, and timing of exposure to stressor are critical for determining whether stress will enhance or suppress/dysregulate immune function. The model suggests that an effective strategy for health maintenance would be to stay within the “good stress” and psychophysiological equilibrium regions of the stress spectrum while minimizing the time spent in chronic stress.

SLEEP, CIRCADIAN RHYTHMS, AND SKIN

Sleep is important for growth and renewal of multiple physiological systems. Chronic poor sleep quality is associated with increased signs of intrinsic skin ageing, diminished skin barrier function, and lower satisfaction with appearance. It is during deep sleep that "slow-wave sleep" stimulates the release of growth hormone and promotes normal tissue repair and the patching up of the wear and tear of the day. The human sleep–wake cycle is usually tightly coupled to the circadian time course of *core body temperature*. The circadian thermoregulation of heat loss in the evening, via distal skin regions, is intimately associated with sleepiness and the ease to fall asleep. The preoptic anterior hypothalamus (POAH) region in the brain is the main integrator of sleep and thermoregulatory information. If thermoregulatory function is impaired as a result of a skin disorder and heat is not dissipated normally, sleep onset may be prolonged and the patient may experience a decline in restorative sleep and poor health outcomes.

Circadian rhythmicity and the sleep–wake cycle are also associated with the *skin barrier function of stratum corneum* via control of transepidermal water loss (TEWL), skin surface pH, and skin temperature, with skin permeability being higher in the evening and night than in the morning. Higher TEWL in the evening with a compromised epidermal barrier function is associated with greater itch intensity, which is typically higher in the evening before bedtime. The majority of patients with inflammatory dermatoses report exacerbation of pruritus during the night. Good examples are atopic dermatitis, psoriasis, and chronic idiopathic urticaria. This diurnal pattern of pruritus is observed in various skin disorders and it is mediated by complex circadian-mediated factors including lower cortisol levels, increased skin temperature, and decline in epidermal barrier function.

Sleep disorders and skin disease: *Pruritus* (including nocturnal pruritus) is one of the most common and bothersome symptoms of dermatological disease, and impairs quality of life in patients and their family in atopic dermatitis. Scratching during sleep, which may vary during different stages of sleep depending upon sympathetic nervous system activity, usually occurs more frequently during non-rapid eye movement (NREM) stages 1 and 2 and REM sleep, and least in deep sleep (stage N3) with lowest sympathetic tone. Sleep deprivation and/or restriction can cause a heightened proinflammatory state (e.g., increased NK-cell activity, elevated levels of IL-1-β and TNF-α) and potentially lead to exacerbation of various dermatological disorders. *Repeated arousals* in sleep disorders like obstructive sleep apnea and primary insomnia are associated with high sympathetic tone and enhanced proinflammatory state, and may contribute to exacerbation and perpetuation of chronic inflammatory dermatoses such as psoriasis. In psychiatric disorders like PTSD, autonomic hyperarousal

and insomnia have been related to nocturnal hyperhidrosis, idiopathic urticaria, and nocturnal pruritus. The patients with atopic dermatitis tend to have increased arousals from sleep because of underlying hyperarousability and overactive sympathetic response to itch and scratching. Recognition of such a bidirectional relationship between sleep and skin disorders is an important consideration in management of psychocutaneous disorders.

ATOPIC DERMATITIS

Epidemiology and Clinical Features

Atopic dermatitis (AD) is a chronic inflammatory disease of skin, characterized by pruritus, mild to severe erythema, scaling, excoriation, and lichenification of the skin. AD often begins in infants and toddlers, presenting as erythematous and pruritic maculopapular eruptions. Subsequent rubbing and scratching lead to lichenification and further itching and scratching cycle (itch-scratch cycle). The pruriginous lichenoid papules and lichenification in flexural areas and hands is more prominent in adolescents and adults. Individuals with AD may have a personal or family history of atopic disease group (allergic rhinitis, allergic conjunctivitis, bronchial asthma, or hay fever). It is estimated that 1 to 3 percent of adults and 5 to 20 percent of children suffer from AD over a worldwide distribution. The AD manifests in first year of life in about 60 percent of cases, and only 30 percent in the subsequent 1 year. The clinical course of AD improves by puberty, and its initial manifestation is rare after puberty. Both sexes are affected, but boys are somewhat more afflicted during the infantile phase. Its incidence has been observed to increase in the past few decades, which has been attributed to early exposure to environmental factors such as outdoor pollution, food additives, household allergens, and infections. The hygiene hypothesis proposes that a western lifestyle with early exposure to psychosocial stress and reduced microbial exposure (from near sterile food preparations and large-scale immunizations) affects the differentiation of T cells toward an atopic profile of predominant TH2 cells. A combination of early onset severe dermatitis, comorbid asthma, or allergic rhinitis is often associated with disease chronicity (Fig. 27.12–4).

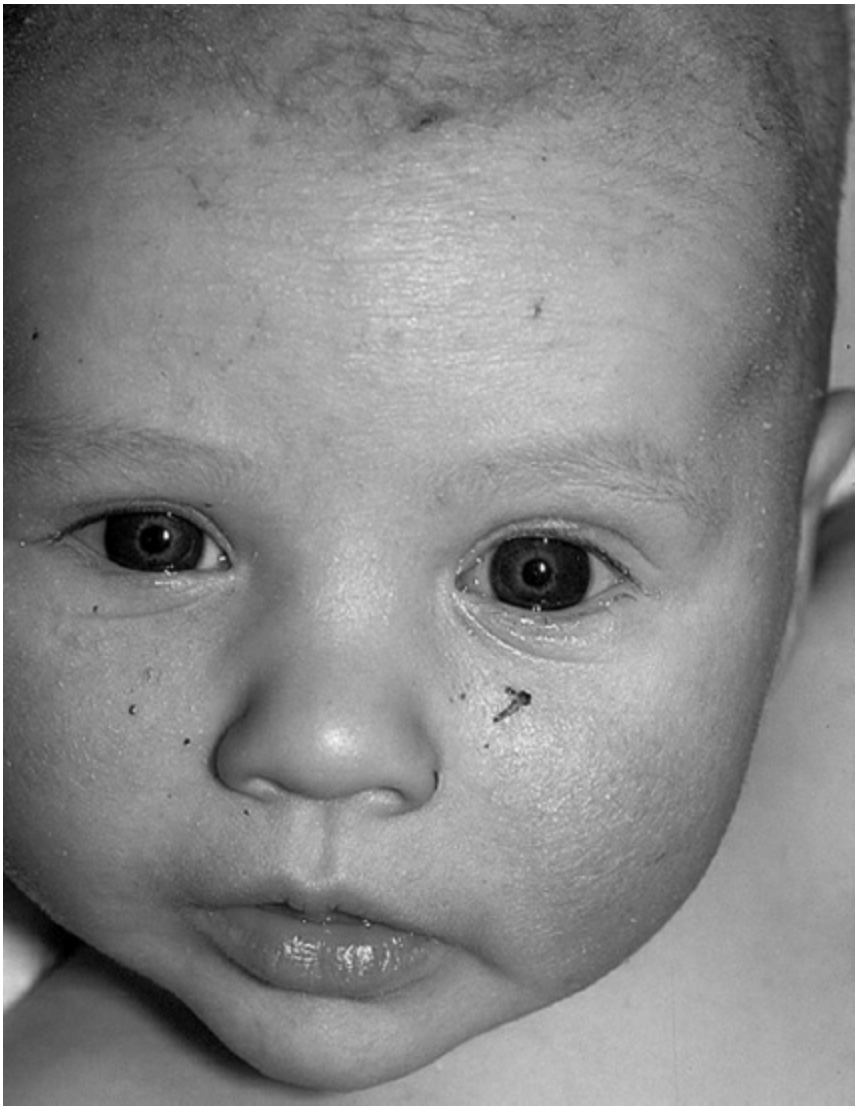

FIGURE 27.12–4. Atopic dermatitis. Infantile-confluent erythema, microvesiculation, scaling, and crusting on face. (From Goodheart HP. *Goodheart's Photoguide of Common Skin Disorders*. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2003.)

Pathophysiology of Atopic Dermatitis

AD's pathophysiology involves multiple factors including genetic, environmental, and psychosocial factors, which interact with underlying psychoneuroimmunological mechanisms, leading to multifactorial pathogenesis. AD provides a good model for the integrated understanding of "biological psychosomatics" (Fig. 27.12–5).

Genetic Predisposition. Patients with AD have a genetic predisposition, and about two-thirds of them have a positive family history of AD. The concordance rates for monozygotic twins (0.75) is higher than dizygotic (0.25) twins. The inheritance is considered to be polygenic and multifactorial. More than 20 genes are described in the development of allergic diseases, and chromosomes 5q331-33 have genes for several important cytokines (IL-3, IL-4, IL-5, IL-13). There is some evidence for

existence of an atopy gene in chromosomal region 11q13. The chromosomal inheritable traits in AD have been implicated for HPA axis abnormalities and imbalance of immunoregulatory T-lymphocyte system, leading to defective cell-mediated immunity and increased IgE production.

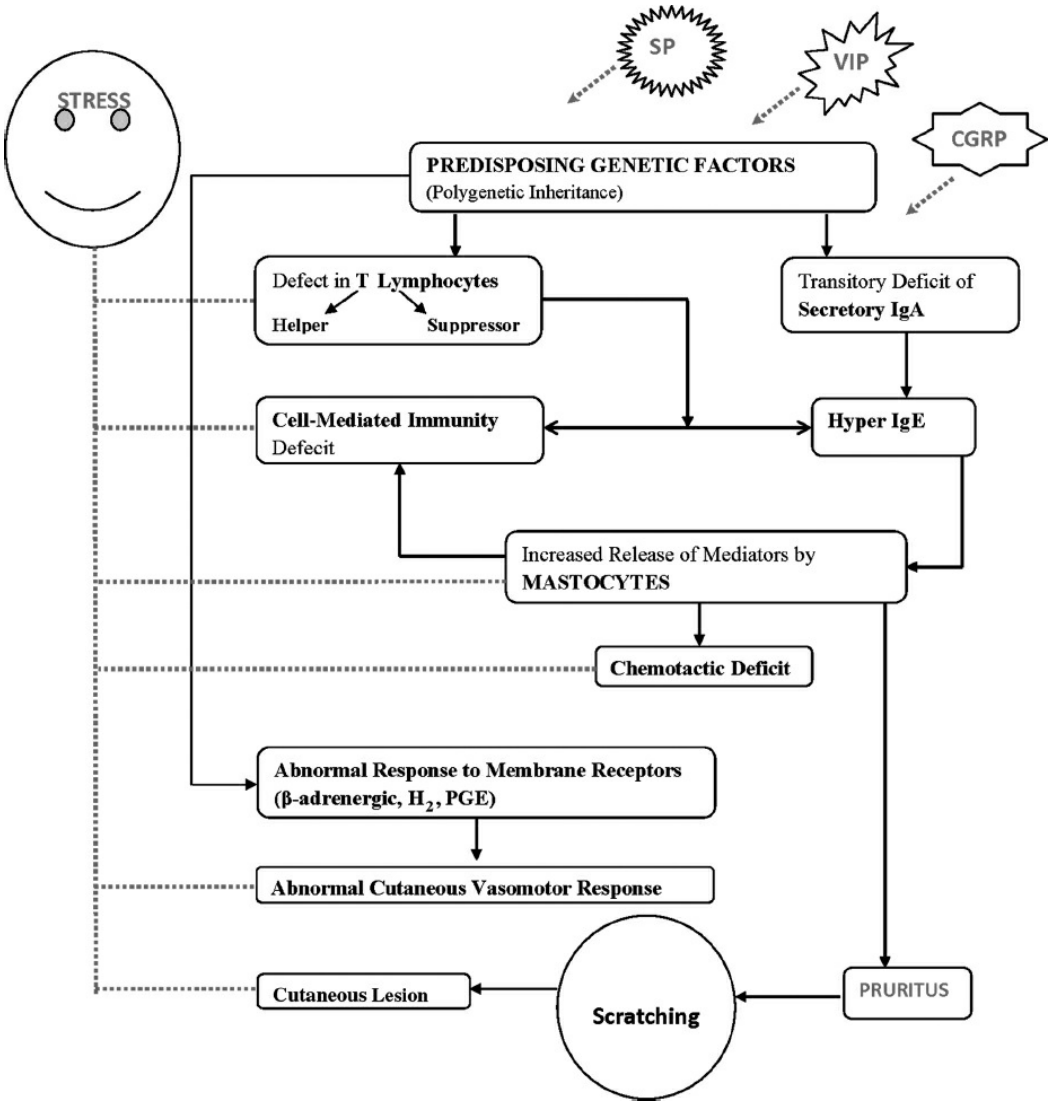

FIGURE 27.12–5. Hypothetical outline of the pathogenesis of atopic dermatitis correlated with the possible influence of emotional stress (*dotted lines*). SP, substance P; VIP, vasoactive intestinal peptide; CGRP, calcitonin gene–related peptide; Ig, immunoglobulin; H_2, type 2 receptors for histamine; PGE, prostaglandin E. (From Koo JYM, Lee CS, eds. *Psychocutaneous Medicine*. New York: Marcel Dekker; 2003.)

Psychosocial stress associated with AD is discussed earlier in this section and it includes perceived maternal rejection from lack of early cutaneous stimulation and emotional separation from the mother, altered psychodynamics in mother–child relationship, temperamental changes in AD infants, problems with self-esteem and relationships in childhood, and development of conflicted personality styles (Fig. 27.12–1).

Acute and Chronic AD (Role of IgE, Cytokines, TH1/TH2 Concept). D. Leung and colleagues proposed the widely recognized pathophysiological mechanisms of AD in uninvolved, acute, and chronic stages. In *uninvolved skin* the cutaneous microenvironment favors a humoral immunity with predominance of mast cells, eosinophils, IgE, and circulating TH2 lymphocytes. In the *acute immune response*, serum IgE plays a major role in AD pathogenesis and binds to mast cells, basophils, and Langerhans cells. An acute response is triggered by IgE-mediated inciting antigen (food items such as milk, eggs, wheat; aeroallergens like dust, mites, ragweed pollen; bacterial infection with *Staphylococcus aureus*), which is presented to T cells. T cells are activated to switch to TH2 lymphocytes. TH2 cells are then stimulated to expand their population, secrete TH2 cytokines such as IL-4 and IL-5, promote elevation of IgE levels, and mediate degranulation of mast cells. IL-4 promotes maturation of plasma cells into IgE-producing B cells, and IL-5 is associated with development of eosinophils. A persistence of irritative mechanical (scratching) and inflammatory stimuli can result in the *chronic phase*. The chronic CMI response is mediated by TH1 cells, which produce cytokine IFN-γ. There is an inverse relationship between IFN-γ and IgE levels in patients with AD: In the acute stage IL-4 levels are high and IFN-γ is low, and in the chronic stage there is a decrease in IL-4 and an increase in IFN-γ and IL-5.

Pruritus and Cutaneous Nervous System (Histamine, NPs and Endorphins). Individuals with AD are more vulnerable to experience itching in response to minor stimuli and emotional upsets, resulting from CNS arousal, which can enhance the vasomotor and sweat responses in skin and lead to lowered itch threshold, which triggers the scratch response. Some studies have reported increased density and hypertrophy of immunoreactive unmyelinated C fibers with extensive innervation in skin. These fibers, when sensitized by histamine, also produce local inflammatory reactions (axon reflex erythema). The sensation of pruritus is modulated by itch-specific C receptors on unmyelinated sensory nerve fibers and is mediated by histamine, cytokines, leukotrienes, and neuropeptides. The acute vascular stage of neurocutaneous immune response is mediated by histamine released from tissue mast cells and triggered by CRH, neurotensin, and SP. Several NPs, such as SP, CGRP, and VIP, have been identified in AP lesions, and the specific effects of these NPs are summarized in Table 27.12–2. The work of L. Singh and colleagues has shown that stress-induced release of CRH from the hypothalamus can directly affect mast cell degranulation, thus affecting the pathogenesis of AD by raising the level of cutaneous histamine. Their studies also reported that blocking of neurotensin (NT) and SP also decreased mast cell degranulation. Another study revealed that opiate neuropeptide β-endorphins are increased in AD patients with over 20 percent of skin

involvement (endorphins reinforcing the endogenous reward from the itch-scratch cycle). Its release in the AD lesions may be triggered by direct CNS to peripheral nerve stimulation or through neuroendocrine pathway via release of CRH and ACTH.

Impaired Glucocorticoids Response and Blunted HPA Axis. As discussed earlier, glucocorticoids act on the immune system to induce the shift from TH1 to TH2 state. Recent research has focused on the role of glucocorticoids resistance in atopic group diseases. Current investigators postulate that increased levels of IL-2 and IL-4 may induce the expression of a defective T-cell glucocorticoids ligand receptor (GCR-β), which has decreased DNA–binding affinity, thus leading to abnormal cellular responses. Various studies imply that patients with AD have a blunted HPA axis response and produce lower cortisol levels in response to stress and glucocorticoids resistance. A compensatory mechanism in AD may include upregulation of glucocorticoid receptors on peripheral lymphocytes, leading to a hyperreactive state in response to glucocorticoids. In spite of a blunted HPA axis in AD, stress-induced cortisol release can stimulate upregulated lymphocytes to a hyperreactive response and a cytokine shift from TH1 to TH2 state, leading to the acute inflammatory stage of AD. In addition, glucocorticoid-induced impaired cutaneous barrier function increases the susceptibility to pathogenic allergens, bacteria, and viruses, which are potential triggers for acute AD. There is strong evidence that patients with AD have abnormal HPA axis activity, leading to decreased ACTH and cortisol response with CRH stimulation. Thus, stress-induced release of CRH from hypothalamic-pituitary vasculature can stimulate mast cell degranulation and trigger a local inflammatory response. An altered HPA axis with lower cortisol levels leads to a vigorous immune response and exacerbation of AD.

Psychological stress can increase catecholamine levels through an overreactive sympathetic response. Catecholamines also contribute to secretion of TH2 cytokines through the activity of intracellular type 4 phosphodiesterases (PDE4) via β-adrenoreceptors. PDE4 degrades intracellular cyclic adenosine monophosphate (cAMP), leading to secretion of IL-13 and IL 4, the potent mediators of TH2 state. The mechanisms of stress-induced exacerbation of AD are complicated, multifactorial, and not completely understood. However, stress-induced glucocorticoids play a major role by mediating the cytokine shift from TH1 to TH2 and also increasing the susceptibility to allergens and pathogens by impairment of cutaneous barrier function. These effects are coupled with systemic and local effects of catecholamines, neuropeptides, and neurotrophins, leading to stress-induced exacerbation of AD.

Psychopathology

Individuals with AD have been found to have increased anxiety and depression. Using the Freiburg Personality Inventory and the State-Trait Anxiety Inventory, Scheich and colleagues reported that AD patients have higher traits of excitability and inadequate stress coping mechanisms. They also found that AD patients with serum IgE levels greater than 100 IU/mL have increased traits of excitability and inadequate coping compared to other AD patients. M. Hashiro and M. Okumura examined the immunological markers and found that NK-cell activity and IL-4 levels are lower in AD patients with depression or anxiety, and IFN-γ levels are higher in AD patients than in controls. By further analysis, they found that decreased NK-cell activity is more strongly correlated with higher anxiety levels. The severity of pruritus was directly correlated with severity of depressive symptoms in AD patients, which is consistent with findings that depression and anxiety can magnify the itch perception and enhance the scratching behavior. Higher anxiety levels in AD patients are correlated with ease of conditioning to develop itch-scratch response. The emotional state in AD children is closely related to the severity of the dermatitis. Adult patients with AD are found to have chronic anxiety disorders and a tendency to internalize anger in conflicted relationships.

Treatment

The role of psychiatric treatments in patients with AD includes reduction of itching or scratching, treatment of associated anxiety and depressive symptoms, and improvement in conflicted relationships. Behavioral modalities focus on interrupting the vicious cycle of itching and scratching (habit reversal training). Controlled trials of psychotropic drug treatments found that topical 5 percent doxepin cream was effective in reducing pruritus because of its potent antihistaminic action. Another antidepressant, trimipramine (Surmontil), with similar histamine antagonism, improved sleep quality and reduced scratching during the night in AD patients.

In a meta-analysis various psychotherapeutic treatments were examined with respect to skin condition and subjective well-being. These included cognitive–behavioral therapy, relaxation training, meditation, and stress management. These treatments were clearly more effective than standard somatic-medical treatments used alone. The psychological treatments were effective in reducing anxiety, depression, and neuroticism, thus providing stable adjunctive treatment response to standard medical care of AD. Combinations of psychotherapeutic interventions resulted in reduction of the itch-scratch cycle, less use of local and systemic steroids, and longer periods of remission. Research into the effectiveness of psychotherapy in treating AD is still largely in early stages.

PSORIASIS

The writer John Updike, who had psoriasis himself, so poignantly said about being a person with psoriasis: "I am silvery, scaly. Puddles of flakes form wherever I rest my flesh. Lusty, though we are loathsome to love. Keen-sighted, though we hate to look upon ourselves. The name of the disease, spiritually speaking, is Humiliation." (Fig. 27.12–6)

FIGURE 27.12–6. Psoriasis vulgaris, chronic stable type. Multiple large scaling plaques on the legs. (From Goodheart HP. *Goodheart's Photoguide of Common Skin Disorders*. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2003.)

Epidemiology and Clinical Features

Psoriasis is a chronic, relapsing skin disease presenting with erythematous, scaling papules and indurated plaques, arising preferentially on the elbows, knees, and scalp. The clinical presentation takes various forms, some causing nail dystrophies and arthritis. It affects 1.5 to 2 percent of the population in western countries, and there are 3 to 5 million people with psoriasis in the United States. Psoriasis is hereditary, with an age of onset mostly in the early 20s. Childhood onset predicts a severe and chronic course and is usually associated with a positive family history. Common trigger factors include physical trauma (Koebner phenomenon: rubbing and scratching stimulate the psoriatic proliferative process), acute bacterial or viral infections, and drugs like lithium (Eskalith), antimalarials,

interferon, corticosteroid withdrawal, and β-adrenergic blockers. Alcohol ingestion is also a known trigger. Psychological stress has been reported to trigger exacerbations of psoriasis. A large study reported that 33 percent of patients noticed new lesions at times of distress. A study found that in 42 percent of 536 patients, worrying was associated with exacerbation of psoriasis. The majority of these patients felt that it was the disease-related stress from cosmetic disfigurement and social stigma, rather than stressful major life events, that led to exacerbation of symptoms. The psychosocial aspects of psoriasis and impairment of quality of life have been described earlier.

Pathophysiology of Psoriasis

Psoriasis plaques are produced largely by the activity of hypertrophic keratinocytes (multiplying approximately 28 times faster than normal skin) and inflammatory changes in dermis. Histological markers include capillary dilatation, sterile microabscesses, parakeratosis, and, more recently discovered, hyperproliferation of cutaneous nerves. Farber and colleagues are proponents of the *neurogenic inflammation* hypothesis of psoriasis. According to this theory, neuropeptides like SP and nerve growth factor (NGF) play a central role in pathogenesis. SP and other NPs are released from the unmyelinated terminals of sensory fibers in skin and generate local neurogenic inflammation in genetically predisposed individuals. In support, one study found increased numbers of neural filaments in both lesional and symptomless psoriatic skin. The development of histochemical staining has led to identification of nerves containing neuropeptides like SP, VIP, and CGRP in psoriatic skin. The specific effects of NPs on target cells in psoriasis are listed in Table 27.12–2.

Stressful life events are associated with higher levels of SP in the central and peripheral nervous systems of animal models. Descending autonomic paths via the dorsal root ganglia in the spinal cord can stimulate release of cutaneous NPs through SP-containing neurons. The descending autonomic fibers also have connections to opioid interneurons in the dorsal horn. Stress-related stimulation of the HPA axis results in increased levels of ACTH, glucocorticoids, and adrenalin. There is evidence for higher NGF levels in response to stress, which plays a central role in regulating skin innervation and upregulating NPs. Several studies indicate that increased expression of NGF in keratinocytes may be an early event in the pathogenesis of psoriasis. NGF stimulates T-lymphocyte proliferation and mast cell degranulation and induces expression of a potent chemokine RANTES, which has chemotactic effects on resting $CD4^+$ memory T cells and memory T cells.

Recent literature has also focused on the role of T lymphocytes in the

pathogenesis of psoriasis. However, the role of autoimmune inflammation in the pathogenesis of psoriasis remains controversial. The T-cell paradigm alone does not explain some of the important clinical features of psoriasis (e.g., Koebner phenomenon, symmetrical distribution of lesions, increased nerve density and NP expression in psoriatic lesions, and clearance of psoriatic lesions in anesthetic areas following a sensory nerve).

Psychopathology and Treatment

Patients with psoriasis have been reported to have a high prevalence of generalized anxiety disorder and major depressive disorder (anxiety, depression) and comorbid personality disorders. Suicidal ideation was observed to be prevalent among 7.2 percent of severely affected psoriasis inpatients, versus 2.5 percent among less severely affected outpatients. Psoriasis-related stress has been associated with greater psychiatric comorbidity. Social deprivation and the experience of being stigmatized were found to be correlated with higher depression scores. The adult psoriasis patients who experienced touch deprivation in social interactions had higher depression scores compared to those who did not perceive the touch deprivation. In psoriasis, the severity of pruritus is associated with higher depression scores and a greater risk for suicide. An improvement in pruritus led to a decrease in depression scores. Early onset psoriasis has been associated with greater difficulties with expression of anger, a personality trait that may add to a patient's vulnerability to stress and depression. The prevalence of psychiatric comorbidity and emerging evidence for stress-induced exacerbations of psoriasis has led to the development of psychosocial intervention as an important adjunctive therapy in the management of psoriasis.

The psychotherapeutic modalities with evidence for significant clinical improvement include *arousal reduction techniques* like hypnosis, mindfulness meditation and guided imagery training. *Cognitive–behavioral therapy* may be of specific value in improving individual coping strategies by addressing illness perceptions, beliefs about control and curability of illness and interpersonal issues. Screening for psychological distress would detect undiagnosed cases with anxiety and depression and enable patients to get psychological support and treatment. There is limited data on controlled trials of *psychopharmacological treatments* of anxiety and depression in psoriasis patients and its effect on the course of psoriasis. TNF-α antagonists are safe and effective in improving both the physical severity of psoriasis and quality of life. Patients with depression also have increased levels of TNF-α and it is speculated that TNF is related to fatigue and sleepiness and may account for coexistence of depression with fatigue. TNF-α antagonists (Adalimumab) may reverse depressive symptoms and fatigue in patients with psoriasis. Reduction in anxiety and depression scores have also been reported with use of ustekinumab (an

anti-IL-12/23 agent used to treat psoriasis). There is need for caution about use of psychotropic medications in patients with psoriasis because of several conflicting case reports of induction or exacerbation of psoriasis with use of fluoxetine and bupropion. Use of mood stabilizer lithium is commonly associated with precipitation or aggravation of psoriasis.

ALOPECIA AREATA

Alopecia areata (AA) is a common disease involving localized loss of hair in round or oval areas, without visible inflammation of the skin in hair-bearing areas (Fig. 27.12–7).

Epidemiology and Clinical Features

Breakage of the hair shaft presents as characteristic exclamation-mark hairs. Any hair-bearing surface may be affected, but it frequently occurs on the scalp (70 to 90 percent), eyebrows (4 percent), and eyelashes and beard (28 percent in males). About 1 percent of the US population has at least one episode of AA by age 50. AA represents 2 percent of new dermatological outpatient visits in the United States. Incidence is equal in both sexes. Sixty percent of patients with AA present with their first lesion before 20 years of age. AA is a partial loss of hair; alopecia totalis is total loss of scalp hair; and alopecia universalis is total loss of hair on scalp and body. The course of AA is variable; 30 percent have complete recovery and 20 to 30 percent never recover from the first episode. Poor recovery process is associated with prepubertal onset, atopy, rapid evolution of hair loss, nail dystrophy, and extensive hair loss.

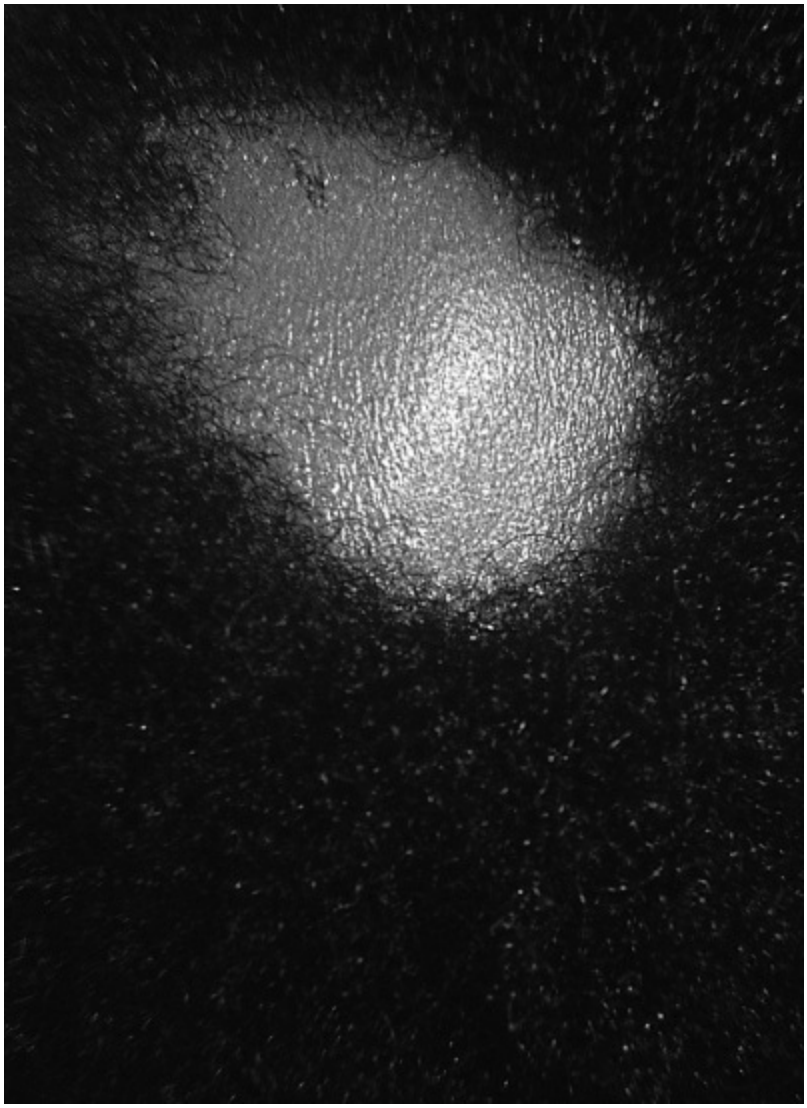

FIGURE 27.12–7. Alopecia areata of scalp. Sharp outline, without scaling, erythema, atrophy, exclamation point hair. (From Fleisher GR, Ludwig S, Baskin MN. *Atlas of Pediatric Emergency Medicine*. Philadelphia, PA: Lippincott Williams & Wilkins; 2004.)

Pathophysiology

The causes of AA remain unknown. Some preliminary evidence suggests that AA is a T-cell–mediated autoimmune condition that is most likely to occur in genetically predisposed individuals. There are associations between AA and autoimmune disorders such as Hashimoto thyroiditis (prevalence varies from 8 to 11.8 percent). The autoantibody response is heterogeneous and targets multiple structures of the anagen phase hair follicle. Whether these antibodies play a direct role in the pathogenesis or are epiphenomena is not known yet. Several studies propose how stress may precipitate AA. Acute stress leads to CRH release and activation of overexpressed type 2-β CRH receptors around the hair follicles, leading to an intense local inflammation. M. Toyoda and colleagues postulate a role of NPs in the pathogenesis of AA. SP is released by dermal nerve fibers

around hair follicles, and their study reveals that SP and its degrading enzyme neural endopeptidase (NEP) have been strongly expressed in affected hair follicles in the acute progressive as well as chronic stable phase of the disorder. It is possible that different pathogenic mechanisms may be involved in different AA, leading to a variable prognosis.

Psychopathology

In a study of 294 patients with AA, prevalence of major depression was 8.8 percent. Another smaller study of 31 patients with AA revealed 39 percent prevalence of major depression, and 79 percent lifetime prevalence of one or more psychiatric disorders. A study of 32 patients with AA reported 66 percent prevalence of psychiatric comorbidity, including generalized anxiety disorders, adjustment disorders, and major depressive episodes. Patients with highly stress-reactive AA report more depressive symptoms than patients who deny an association of stress with their AA.

Treatment

The standard treatment of AA involves topical and intralesional steroid injections as well as systemic glucocorticoids. There are few studies involving psychiatric treatment of AA. In one randomized, double-blind, placebo-controlled trial, patients given imipramine (Tofranil, 75 mg per day) had significantly more hair regrowth than the untreated group. AA patients with comorbid major depression and anxiety disorders showed more improvement in AA symptoms when they were treated with selective serotonin reuptake inhibitor (SSRI) agents. Use of relaxation techniques, stress management, and psychotherapy has been an effective adjunctive treatment for AA, but results are not backed by controlled studies.

ACNE VULGARIS

Epidemiology and Clinical Features

Acne is an inflammation of the pilosebaceous glands in certain parts of the body (mainly involving face and trunk) that most frequently occurs in adolescence and manifests itself as comedones (comedonal acne), papulopustules (papulopustular acne), or nodules and cysts (nodulocystic acne and acne conglobata). Possible complications such as pitted, depressed, or hypertrophic scars may result from all types, but especially in nodulocystic and conglobata acne. It is a very common skin disorder among young people and usually begins at 10 to 17 years of age in females and 14 to 19 years in males. The course is usually self-limiting, and most individuals improve in several months to years. Acne may persist into the second or third decade, especially in women, whilst in men it tends to take a more severe form. The genetic predisposition is multifactorial, and the majority of patients with cystic acne have a family history of severe acne.

Severe acne may be associated with XYY syndrome (Fig 27.12–8).

FIGURE 27.12–8. Acne vulgaris, nodulocystic type, with nodules, cysts, and pustules. (Image provided by Stedman's.)

Pathophysiology

Acne results from changes in the keratinization pattern, followed by dense keratin blocking of the secretion of sebum from sebaceous glands. The keratin plugs are called *comedos*, which cause plugging and collections of sebum within pilosebaceous units. This is followed by complex interaction of androgens and bacteria (*Propionibacterium acne*) within the plugged sebaceous glands. Androgens stimulate the sebaceous glands to produce more sebum. Bacteria containing lipase break down lipids into fatty acids and produce proinflammatory mediators like IL-1 TNF-α. Fatty acids and proinflammatory cytokines cause a sterile inflammatory response; if distended, follicle walls break, the contents (sebum, lipids, fatty acids, keratin, and bacteria) enter the dermis, invoking an inflammatory foreign body response. This results in formation of pustules, nodules, and scarring. Patients with acne vulgaris frequently experience stress-induced exacerbation of their acne lesion. The postulated mechanism involves stress-induced activation of the HPA axis followed by increased secretion of glucocorticoids and androgens, which exacerbates the inflammatory response within the pilosebaceous unit. Higher anxiety levels in patients with cystic acne have been associated with higher blood catecholamine levels, which decrease with treatment of acne. The psychological stress from acne might play a role in further exacerbation of acne lesions through

increase in catecholamine levels.

Psychopathology

Disfigurement resulting from acne can lead to serious emotional consequences, including depression, social phobia, anger, and low self-esteem. One study reported a 5.6 percent prevalence of suicidal ideation in adolescent and young adults with noncystic facial acne. In a large study of 4,376 patients with dermatological disorders, 53 percent of the acne patients reported that there was close chronological association (average latency was 2 days) between emotional stress and exacerbation of their acne lesions. In a study of 317 students aged 14 to 16 years, those with significant lesions (more than 12 lesions) had higher levels of emotional and behavioral difficulties than those with minimal acne. Teenagers can be extremely sensitive to issues of peer acceptance and overly conscious of their appearance. They are vulnerable to develop social phobia and can feel inhibited in their lifestyle. Acne can significantly interfere with their social interactions, sports activities, academic performance, and dating behavior.

Depressive symptoms associated with acne are often in reaction to body image concerns in teenagers and young people. There is a highly significant positive association between acne and poor self-image. In one study, patients consistently rated their acne as more severe than did the clinician, and they rated their satisfaction with treatment outcome less than that of the treating dermatologist. In contrast to psoriasis, the severity of acne does not necessarily correlate with severity of depression, and even mild-to-moderate acne has been associated with depression, suicidal ideation, and completed suicide. Often, the presence of a primary psychiatric disorder such as OCD, delusional disorder, or eating disorder can add to the severity and chronicity of acne. A good example is *acne excoriée*, where the patient's need to persistently pick on acne lesions results in severe disfigurement from scarring.

Treatment

Since younger patients with acne vulgaris have a higher risk of developing depression, psychosocial stress, and suicidal ideation, it is important to educate and monitor each patient carefully. Treatment options for depression, anxiety, and social phobia in patients with acne include cognitive–behavioral therapy, relaxation training, self-hypnosis, and pharmacotherapy if indicated. One study used cognitive imagery and biofeedback relaxation as an adjunct to the medical treatment of acne. The study reported greater improvement when psychological treatments were used, and discontinuation of such interventions led to a relapse in acne severity. There are case reports of improvement of acne after treatment of depression with paroxetine (Paxil). There has been extensive media

attention because of reports of treatment of acne with isotretinoin (Accutane) and its association with depression, suicidal ideations, and suicide attempts. Recently the FDA and isotretinoin's manufacturer added another warning to the psychiatric section of package insert about the possible development of aggressive or violent behaviors. However, a large epidemiological study failed to demonstrate a higher prevalence of depression and suicide in patients treated for acne with isotretinoin versus antibiotics. Design limitations of this and other studies preclude firm conclusions about a possible causal relationship between isotretinoin and psychiatric complications. A careful history will reveal whether or not exacerbation of depressive symptoms is temporally related to initiating isotretinoin treatment. The depressed acne patients should always be assessed for suicide risk.

URTICARIA AND ANGIOEDEMA

Epidemiology and Clinical Features

Urticaria is composed of wheals (hives), characterized by transient, well-circumscribed edematous papules and plaques, usually pruritic, and involves superficial (papillary) dermis. Angioedema is a large, deep, ill-defined edematous area that involves dermis and subcutaneous tissue. Urticaria and angioedema are the same edematous process involving different levels of the cutaneous vascular plexus. They may be acute recurrent or chronic recurrent. Fifteen to 23 percent of the population may have this condition during their lifetimes. The course of urticaria is chronic in about 25 percent of the cases. *Acute urticaria* has sudden onset with frequent recurrence within 30 days. It usually presents with large wheals, often associated with angioedema. Acute urticaria is IgE-dependent (in patients with atopic background) and associated with common triggers such as alimentary parasites, infection, and intolerance to some drugs or food. Serum sickness–like reactions can result from transfusions involving whole blood or immunoglobulins. This process might be complement mediated. Urticaria reactions are mediated by vasoactive mediators in the skin, such as histamine from mast cells and basophils. *Chronic urticaria* has a recurring pattern, over more than 30 days, and presents with both small and large wheals. These wheals are *rarely* IgE-dependent but are often due to anti-FcεR and/or anti-IgE autoantibodies, suggesting autoimmune pathogenesis. The etiology is unknown in 80 percent of cases; thus, the condition is considered idiopathic. Chronic urticaria affects adults predominantly and is twice as common in women compared to men. About 40 percent of patients with chronic urticaria lasting more than 6 months continue to have urticaria 10 years or longer, which seriously compromises their quality of life.

Pathophysiology

Urticaria and angioedema are classified on the basis of etiology into various subcategories, including *immunological* (IgE-mediated, complement-mediated, autoimmune, immune contact type), *physical* (dermographism, cold urticaria, solar, cholinergic, pressure angioedema, vibratory angioedema), *idiopathic*, and *hereditary angioedema*. Of particular interest is adrenergic urticaria, presenting as typical "halo-hives" (papules surrounded by white halo). This develops in response to acute emotional stress, usually associated with increased plasma levels of adrenaline and noradrenaline. Adrenergic urticaria improves after treatment with β-adrenoreceptor blockers such as propranolol (Inderal) and can be induced by intradermal injection of noradrenaline. The role of psychological factors in chronic idiopathic urticaria is unclear. One study indicates higher prevalence of depression and anxiety in women with chronic urticaria. Historically, angioedema has also been known as angioneurotic edema because of frequent clinical observation that stressful events preceded the onset of angioedema. The physiological basis for stress-induced urticaria and angioedema is likely to be secretion of neuropeptides, such as VIP and SP, which can lead to vasodilatation and urticarial wheal formation.

Treatment

In urticaria caused by a known afflicting agent, treatment includes prevention of attacks by eliminating the use of etiological chemicals, drugs, and food additives, or by preventing infections. Controlled trials have indicated that antidepressants like doxepin (10 mg three times a day) are more effective than diphenhydramine (Benadryl) (25 mg three times a day) in controlling chronic angioedema. There are anecdotal reports of improvement in urticaria symptoms with use of SSRI agents in patients with urticaria comorbid with panic disorder. One controlled study found that use of psychological treatments like hypnosis and relaxation reduced pruritus, but did not affect the number of wheals.

DELUSIONAL DISORDER, SOMATIC TYPE: DELUSIONAL PARASITOSIS

Clinical Features

Patients suffering from delusional parasitosis (DP) have the fixed belief that they have been infested with parasites. G. Thibiérge first described the condition in 1890s and called the syndrome *acarophobia*. Other terms used in the literature are parasitophobia, entomophobia, monosymptomatic hypochondriacal psychosis, and Ekbom syndrome. Delusional disorder (parasitosis), using the DSM-5 term, occurs as a single somatic delusion with no impairment of thought processes. Patients often describe the sensation of bugs crawling on or inside the skin; some report

seeing bugs or hearing buzzing sounds; and some "bring the parasites" for further examination (usually pieces of hair, skin, or cloth). This behavior is so characteristic that it has been referred to as "the matchbox sign." Often a patient with DP may report a specific precipitant, such as a history of infection with a contagious organism or real infestation. Their attempts for self-treatment include repeated washing, checking, and cleaning; excoriation of skin with needles, knives, or fingernails; discarding household items; and excessive use of insect repellents or insecticides. Relatives may share the delusion (*folie à deux*), and the fear of contaminating others is common. DSM-5 no longer differentiates delusional disorder from shared delusional disorder. Most patients seek dermatological treatment because their delusion involves skin invasion. It is a relatively common syndrome in dermatological and parasitology practice, but rarely seen in psychiatry. *Morgellons disease*, a pattern of dermatological symptoms very similar, if not identical, to those of delusions of parasitosis, was first described many centuries ago, but has recently been given much attention on the internet and in the mass media. An Internet-based foundation has developed a working case definition on its web site with specific common signs and symptoms for diagnosis. There may be potential benefit of using the term as a means of building trust and rapport with patients to maximize treatment benefit (Fig. 27.12–9).

Case Report 1. Delusional Parasitosis

Diagnosis of DP

Symptoms of DP can be confused with phobia or OCD, and accurate distinction between these conditions has obvious relevance to treatment. The differential diagnosis of DP includes a number of medical, neurological, and psychiatric disorders, which are listed below in the section on psychogenic pruritus. The effects of certain drugs can mimic a delusion; cocaine and amphetamines especially may produce a tactile hallucination referred to as "cocaine bugs" or "Magnan sign," where the affected person believes he or she has or feels parasites crawling under the skin. Patients with severe alcohol withdrawal may experience visual hallucinations of insects or tactile hallucinations that bugs are crawling over their skin (formication).

FIGURE 27.12–9. Delusions of parasitosis. Lesions are mostly excoriations. Patient picks on skin with nails and needles. (Image provided by Stedman's.)

Pathogenesis

The psychopathology of DP involving specificity of content and fixed belief has been the focus of many empirical studies in the past few decades. The pathogenesis of DP has been considered a multifactorial process in which psychic, cognitive, and social circumstances act as disorder predisposing, triggering, and maintaining factors, respectively. The psychic factors may add predisposition to a particular delusional theme and involves complex interactions of age, gender, social situation, and other key pathoplastic experiences, as suggested by K. Jaspers. Among DP patients, typically elderly, socially isolated women, tactile phenomena constitute key experiences. Most patients feel sensations of movement in and under the skin, making them believe the sensation could be caused by parasites. The fixation of the false belief is likely affected by an altered cognitive frame from organic disorders, schizophrenia spectrum thought disorder, or affective psychosis. Delusions are considered a dynamic process that only persists if disorder-maintaining social factors (isolation, alienation, avoidance behaviors) are operative.

Magnetic resonance imaging (MRI) in patients with DP is suggestive of disturbed functioning in brain regions associated with interpretation of perceptions, involving putamen and the associated dorsal striato-thalamo-cortical loop. This is reported to be linked to somatic delusions and tactile misperceptions, while frontal dysfunction is hypothesized to be responsible

for disturbed reality testing. The therapeutic efficacy of D2-dopamine antagonists in DP indicates dopaminergic dysfunction in above brain regions. Patients with intoxication from substances influencing dopamine transmission (e.g., cocaine, amphetamines, etc.) may experience delusional infestation, which further validates the imaging data.

Delusional Parasitosis: Three Subgroups. Earlier studies considered DP as a particular manifestation of monosymptomatic hypochondriasis. However, the results of empirical psychological studies of DP patients and psychopathological analyses of large numbers of single cases indicate that DP should not always be considered to be hypochondriacal psychosis. In fact, this evidence suggests that there are three subgroups of DP:

1. Patients with *predominantly hypochondriacal traits*, who are convinced that they have an incurable physical illness. This is referred to as hypochondriacal delusion of parasitosis or simply DP. In some cases, those not suffering from organic psychosis, schizophrenia, or affective disorder may be considered in the group of monosymptomatic hypochondriacal psychosis.
2. Patients with *paranoid delusions* and without hypochondriacal traits. They tend to fight against parasites which impair their existence. This is referred as delusions of infestation. Such patients consult parasitologists, entomologists, and even pest control companies but seldom see dermatologists.
3. Patients with signs of both *hypochondriacal as well as paranoid delusions*. They are referred as hypochondriacal delusions of infestation. Such patients usually get repeated dermatological consultation, like those in the first group, hopelessly trapped in their beliefs.

Treatment

The above distinction of three groups has relevance to treatment. Patients with DP are reluctant to seek psychiatric treatment. It is important to establish rapport and build a stable relationship based on trust. This is accomplished by listening to patients' stories, showing empathic understanding, asking how the condition affects their lives, and trying to reduce their sense of isolation. It is often a mistake to inform patients in the first meeting that they are suffering from a delusional disorder. It is best to acknowledge patients' complaints as an experienced reality and leaving open the correctness of their interpretation of the experience. Treatment strategies include psychotropic medications as well as psychotherapeutic and sociotherapeutic methods. By implementation of such an integrative approach, improvement in delusional symptomatology can be expected in about two-thirds of patients. Psychopharmacological treatment of DP has most commonly advocated the use of pimozide (Orap),

a potent neuroleptic. In one double-blind, crossover-controlled study of treatment of DP with pimozide, 10 of 11 patients improved after 6 weeks of 1 to 5 mg per day. A follow-up study reported that pimozide treatment lasting for 5 months resulted in long-term remission in many patients. Full remission of DP has been reported in 50 percent of the patients receiving pimozide. Other antipsychotic medications (risperidone [Risperdal], haloperidol [Haldol], trifluoperazine [Stelazine], chlorpromazine [Thorazine]) have also been effective in uncontrolled trials, and there is no conclusive evidence that pimozide is superior to other antipsychotics. However, pimozide is also a potent opiate antagonist, which has an antipruritic action and may contribute to improvement seen in DP patients. Pimozide treatment carries an increased risk of QT prolongation, which precludes its use in patients with cardiac dysrhythmia. Most authors recommend the use of second generation antipsychotics (risperidone, olanzapine, amisulpride) due to improved side-effect profile. The latter has response rate of 75 versus 60 to 100 percent with first generation antipsychotics. There are anecdotal case reports of treatment response from SSRIs, tricyclic antidepressants, and electroconvulsive therapy. One case report of successful treatment with naloxone (Narcan) and naltrexone (ReVia) indicates a possible disorder in endogenous opiate system, which may also explain the efficacy of pimozide.

PRURITUS AND CHRONIC IDIOPATHIC PRURITUS

Psychosomatic Aspects of Pruritus

Pruritus is a common unpleasant sensation that provokes a desire to scratch. In the brain, pruritus and psyche are intertwined in a complex manner, and appears to have bidirectional effect on each other. The skin and the brain are polar terminal differentiations from the same embryonic neuroectoderm, and pruritus as a symptom demonstrates the intricate link between these two organs.

Psychological Aspects of Pruritus

It is a common observation that “itch” can be induced simply by thinking about it. In a study involving atopic dermatitis patients who were shown video recording of other people scratching, the anticipation of developing itch caused the subjects to experience intense itch. Many patients with chronic pruritus experience scratching as an automatic response to the sensation of itch. Scratching provides immediate relief of discomfort and may perhaps serve to reduce inner tension, and this negative reinforcement leads to conditioned scratching. A vicious cycle of itch-and-scratch eventually results; and continuation of this situation can be perceived as a loss of control and helplessness, and it is often accompanied by a sense of despondency and guilt.

The concept of "ego-skin and somatoform dissociation" may be helpful in understanding why psychological conflicts may be translated into skin symptoms as itch. Sigmund Freud observed that during early development, the ego is rooted in the body, especially the skin, which is an important organ of communication, and that bodily sensations and experiences form the core around which the ego develops. The ego partly remains identified with the skin even as the person becomes an adult, and disturbance of the normally integrated state of the ego can result in symptoms focused on the skin.

Pathophysiology of Pruritus

As described earlier under the section in atopic dermatitis, the epidermis is heavily innervated by axonal nerve endings with nociceptive (pain sensing) unmyelinated C-fibers. The sensation of an itch to the CNS is mediated by the same nociceptive unmyelinated C-fibers, but they are functionally different, and modulated by itch-specific C-receptors. D. Andrew reported a separate itch pathway by identifying a 'second-order histamine-specific itch neuron' in the dorsal horn of the spinal cord, implicated in neuropathic itch syndromes. From the dorsal horn, spinothalamic neurons project to the thalamus and anterior cingulate cortex and insula. In recent studies, the cingulate cortex was found to be an important area in the processing of itch, and became significantly activated in patients with AD, after administration of histamine and its activation correlated with disease severity. Also the cingular cortex was found to be deactivated after scratching in healthy controls. The insula is the seat of *interoceptive representation* and provides the basis for subjective feelings from all the sensory input from the body. Activation of insular cortex has been reported with sensations of pain from skin, itch, touch, temperature and prickly sensations.

Anterior cortex has been known to be involved in modulation of emotional and cognitive activity(such as reward anticipation). Based on these observations, it is conceivable how patients having difficulties with emotional regulation and interoceptive awareness can engage in body-based repetitive behaviors, and thus provide the physiological basis of how emotional factors can affect the perception and processing of itch.

A number of chemicals can provoke or enhance pruritus. The best-known are histamine, prostaglandins, proteinases, substance P, neuropeptides, cytokines (IL-2), and bile salts. Some of these substances act directly on the free nerve endings, while others act through action on mast or other cells. Among prostaglandins, PGE-2 does not itself cause itch, but lowers the threshold and potentiates the itch provoked by histamine. Opioids, such as endorphin, met-enkephalin, and leu-enkephalin, can play a role in causing itch, acting centrally or peripherally. A recent investigation by PET demonstrated that intracutaneous injection

of histamine caused activation of the premotor area and inferior parietal lobule, regions that activated planned movement. Probably there are synaptic connections to the motor area of the cortex, which induce the need for scratching. In a large population study in Oslo there was a significant correlation between mental stress and itch as well as pain. Psychological factors such as depression can modulate histamine-induced pruritus by lowering the itch threshold, and severity of pruritus has been correlated with degree of depressive symptoms.

Etiology of Chronic Idiopathic Pruritus

It can be caused by numerous cutaneous, medical, neurological, and psychiatric disorders. Evaluation of a patient with pruritus should always involve a careful examination and medical work-up to rule out any condition that may be treatable.

- *Medical disorders associated with pruritus* include chronic lymphocytic leukemia, polycythemia vera, malignant melanoma, syphilis, congestive heart failure, diabetes mellitus, vitamin B_{12} deficiency, pellagra, myxedema, hyperparathyroidism, chronic renal failure, hepatic disease, acquired immunodeficiency virus (AIDS), drug toxicity, and old age.
- *Associated neurological conditions with pruritus* include dementia, multiple sclerosis, Parkinson disease, Huntington disease, neurofibromatosis, aneurysm, cerebral infarctions, and CNS tumors.
- *Psychogenic pruritus* is often associated with anxiety disorders, OCD, major depression, or psychosis. Chronic idiopathic pruritus, idiopathic pruritus ani (scratching in the anal area), vulvae (itching in the vaginal area), and scroti (itching in the scrotum) may be the result of a central mechanism. Itch in patients with psychiatric diseases is not uncommon. Among psychiatric inpatients, idiopathic itch was found to affect 36 to 42 percent of the patients and was more frequent in those who exhibited anger trait, angry temperament, and ruminative catastrophization.
- Several neurological causes of pruritus occur in the absence of primary skin lesions and should be considered in the diagnosis of psychogenic pruritus. Neurological pruritus is characterized by: (1) lack of sudden onset; (2) a chronic course; (3) greater severity or intensity; (4) unilateral or bilateral location; (5) association of pruritus with other sensory phenomena such as allodynia, dysesthesia, and hyperpathia; (6) paroxysmal course, often starting and ending abruptly, lasting seconds to minutes, and recurring frequently; (7) pruritus accompanied by paroxysmal constant pain in the same area; or (8) awakenings from sleep or insomnia. In contrast, psychogenic pruritus is characterized by: (1) temporal association with psychiatric symptoms; (2) unlikely occurrence at night; and (3) paroxysmal nature with increased severity, sudden onset and resolution, and intervening symptom-free periods. Both

psychogenic and organic pruritus may coexist, resulting in mixed clinical presentation of pruritus.

Treatment. Antidepressants such as doxepin (Sinequan) have been used effectively in chronic idiopathic pruritus. Amitriptyline (Elavil) has the best documentation in literature of its analgesic property and may work well in itching associated with pain. The use of TCAs amitriptyline and doxepin is limited nowadays in view of the better efficacy and safety profile of the SSRIs. The SSRI antidepressants that have been used to treat pruritus consist of paroxetine, sertraline, fluvoxamine, and fluoxetine. Mirtazapine (Remeron) is the only noradrenergic and specific serotonergic antidepressant (NaSSA) that has been used as an antidepressant, anxiolytic, and antipruritic agent. A trend toward a lower prevalence of idiopathic pruritus was also observed in patients treated with TCAs compared with those on other antidepressants. Gabapentin (Neurontin), a structural analog of the neurotransmitter GABA, has been shown to be effective for prurigo nodularis, neuropathic itch, and uremic pruritus. Pregabalin (Lyrica) is a compound similar to gabapentin and both these medications may have antidepressant and anxiolytic effects in addition to their antipruritic effects. Aprepitant, a neurokinin receptor (NKR)-1 antagonist that has been used as an antiemetic, was recently reported to be effective for chronic refractory pruritus in a case series. NKR-1 is a receptor for substance P, an important itch mediator. Psychological approaches, such as cognitive–behavioral therapy and habit reversal training, can prevent the long-term complications of itching (such as lichen simplex chronicus) by interrupting the itch-scratch cycle.

Table 27.12–3.
Classification of Cutaneous Dysesthesias

Scalp/Head	Urogenital	Known Pathogenesis
Glossodynia	Vulvodynia	Erythromelalgia
Burning mouth syndrome	Orchiodynia	Trigeminal neuralgia
Scalp dysesthesia	Urodynia	Postzosteric neuralgia
Trichodynia	Urethral syndrome	Proctalgia fugax
	Phallodynia	
	Prostatodynia	
	Coccygodynia	
	Perineal pain syndrome	
	Anodynia	
	Proctodynia	

CUTANEOUS SENSORY DISORDERS

Cutaneous sensory disorders (also called cutaneous dysesthesia syndrome) are conditions in which clinical symptoms include abnormal skin

sensations (itching, burning, crawling, stinging, biting, pain) without the presence of primary skin lesions that can be associated with dermatological, neurological, or medical disorders. Dermatologists often consider cutaneous sensory disorder a diagnosis of exclusion due to the absence of any apparent skin lesions and a negative medical work-up. These sensations can occur in any body region but tends to develop in areas with greater density of epidermal innervation, most commonly involving face, scalp, or perineum. The classification of cutaneous dysesthesia is summarized in Table 27.12–3, and is typically referred to either by its affected anatomy or a specific known pathogenesis.

Glossodynia

Patients with *glossodynia* present with chronic altered sensations of pain or burning affecting the tip and sides of the tongue (or other areas inside the oral cavity). Other associated sensory changes may include paresthesia, changes in taste and smell, and mouth dryness. Glossodynia typically affects women in their 50s and is observed in over 5 percent of patients seen by dentists. The condition is usually idiopathic, but glossodynia may be caused by vitamin B deficiencies, diabetes mellitus, candida glossitis, hormonal changes around menopause, salivary disturbances, and problems related to dentures or dental fillings. The causal relationship with psychiatric factors is uncertain, but higher rates of anxiety and depression are reported in patients with glossodynia. The duration of symptoms in glossodynia is variable and may last for years.

Vulvodynia

The term vulvodynia is defined as "chronic vulvar discomfort with abnormal sensations" in the absence of skin disease or infection. The prevalence of vulvodynia is 15 percent in gynecological outpatient practices. Most women with this disorder are between 20 and 50 years of age and have had pain for several years by the time they present for treatment. Affected women with vulvodynia are usually sensitive to touch or light pressure in the vestibule of the vagina, have pelvic floor musculature with altered contractile characteristics, and may also have heightened pain sensitivity in theirs arm or legs. There is growing consensus that vulvodynia is a chronic pain disorder. Vulvodynia patients experience more sexual dysfunction as a result of their primary symptoms. The associated symptoms of allodynia (pain elicited by a nonpainful stimulus) and hyperdynia (when a stimulus causes greater pain than expected) also suggest a neuropathic cause of the pain of vulvodynia. The term *dysesthetic vulvodynia* is used to describe increased pain perception in the vulvar area without a known cause, and may be of central or nerve root origin. This condition often occurs in perimenopausal or postmenopausal women, and

they have fewer symptoms of dyspareunia and less point tenderness on examination.

Somatization appears to be an important factor in cutaneous sensory syndrome patients and higher rates of anxiety and depression are found in this population. However, cutaneous sensory disorders are not specified in DSM-5 under "Somatic symptom and related disorders." A comprehensive medical, neurologic, and psychiatric work-up should be performed to exclude any organic pathology. Once the diagnosis is made, the physician should acknowledge and validate the patient's experiences and symptoms. Direct confrontation of symptoms by clinicians may not be helpful. Pharmacological interventions target the sensations of pain both peripherally and centrally. TCAs (amitriptyline, desipramine), SSRIs, gabapentin, and pregabalin have been effective in therapeutic trials. There are several case reports and uncontrolled treatment trials, suggesting the efficacy of venlafaxine and duloxetine in treating neuropathic pain of chronic vulvodynia. Augmentation with antipsychotics may be used in refractory cases. Controlled studies of cognitive-behavioral treatment and of biofeedback training to reduce pelvic floor spasm have indicated that these treatments are effective in symptom reduction.

OBSESSIVE-COMPULSIVE AND RELATED DISORDERS: BODY-FOCUSED REPETITIVE BEHAVIORS

There are important changes in DSM-5 in relation to psychodermatology. Body-focused repetitive behaviors like excoriation (skin picking) disorder (SPD), trichotillomania (hair pulling disorder) (TTM), and body dysmorphic disorder (BDD) are included under a new chapter *Obsessive-Compulsive and Related Disorders* (OCRD). Epidemiological studies and clinical observations suggest that more than one body-focused repetitive disorder, SPD, TTM, and/or onychophagia may be encountered in the same patient. Table 27.12–4 outlines how these OCRD disorders are categorized under DSM-5 and DSM-IV-TR.

Excessive/compulsive manipulation of skin by patients is common in a large number of psychocutaneous disorders. Skin regions with greater density of epidermal innervation (e.g., face, scalp, perineum) are more susceptible to disagreeable sensations (itching, stinging, burning, etc.). Interestingly, face and scalp are some of the most frequent targets for SPD and TTC. Self-induced dermatoses (TTC, SPD, dermatitis artefacta, onychophagia) caused by excessive manipulation of integument, can contribute to significant morbidity and complicate the course of a primary skin condition like acne (in acne excoriée) and pruritic dermatosis.

Self-induced dermatoses are known to be comorbid with mood and anxiety disorder, including OCD spectrum disorders. M. A. Gupta and others have reported the role of "emotional dysregulation and dissociation" in self-induced dermatoses. TTM and dermatitis artefacta have been

reported to occur in conjunction with other trauma/dissociative disorders like acute stress disorder, PTSD, and multiple personality disorder. Dissociated symptoms are more common in TTM and dermatitis artefacta than in SPD. Dissociation (disruption in integrated functions of consciousness, memory, identity, and perception) is a symptom of emotional dysregulation, which occurs in the context of severe stress coping and extremes of emotional tolerance (window of emotional tolerance). Dissociation is associated with increased threshold for pain perception and numbness of the skin and is often a key factor in self-induced dermatoses, which has significant clinical and treatment implications (Fig. 27.12–10).

Table 27.12–4.
DSM-5 Changes Relevant to Psychocutaneous Disorders

Psychocutaneous Disorders	DSM-5 Classification	DSM-IV-TR Classification
SKIN PICKING DISORDERS Neurotic excoriations Psychogenic excoriations Acne excoriée	**Excoriation (skin picking disorder) (698.4)** A new separate diagnostic group in DSM-5 under "Obsessive-Compulsive and Related Disorders" (OCRD)	**Stereotypic movement disorder with self-injurious behavior (307.3) under major heading of "Other Disorders of Infancy, Childhood, or Adolescence"** Impulse control disorder not otherwise specified (312.30); skin picking used as an example of disorder
TRICHOTILLOMANIA (TTM) Hair pulling disorder Trichotemnomania Trichoteiromania Trichorrhizophagia	**Trichotillomania** (hair pulling disorder) (312.39) classified under **OCRD**	**Trichotillomania** (312.39) Classified under "Impulse Control Disorders Not Elsewhere Classified"
NAIL BITING (Onychophagia) Onychotillomania, lip chewing	Classified as an example of a "Body-Focused Repetitive Behavior Disorder" under Unspecified OCRD (300.7)	No specific diagnostic category
BODY DYSMORPHIC DISORDER (BDD) Dysmorphophobia, dermatological nondisease	**Body dysmorphic disorder (300.7)** Classified under "OCRD"	**Body dysmorphic disorder (300.7)** Classified under "Somatoform Disorders"

As indicated in Figure 27.12–10, skin responds to stress with increased sweat gland activity and increased *skin conductance*, which is a direct measure of general sympathetic nervous system activity. In the context of prolonged or repeated stress, increased sympathetic tone can lead to a state of *hyperarousal*. Such a state of hyperarousal can render the skin more vulnerable to injury (because of increased glucocorticoid levels and poor skin–barrier function) from excessive manipulation. Acute stress is also

associated with increased histamine levels from mast cell degranulation, making skin more vulnerable to itching and scratching by the patient. Chronic stress stimulates manipulation of already more vulnerable skin and contribute to self-induced dermatoses. Traumatic stress stimulates the autonomic nervous system resulting in either sympathetic hyperarousal (fight-or-flight response) or parasympathetic hypoarousal (submission and freezing response), both of which are the basis of defensive survival responses. The later can also follow states of high arousal. Between extremes of sympathetic hyperarousal and parasympathetic hypoarousal, there is the "window of tolerance," in which emotional arousal can be experienced as tolerable and normal homeostasis maintained. This window of tolerance can be quite narrow in an individual dealing with prolonged severe or repeated stress.

Excoriations (Skin Picking) Disorder

The term neurotic excoriation was first introduced by Erasmus Wilson in 1875. Later, in 1889, M. L. Brocq used the term *acne excoriée des jeunes filles* to describe the excoriations caused by self-inflicted lesions in young "neurotic" women with pre-existing acne. A variety of other synonyms have been used to describe skin picking syndrome such as pathological skin picking, neurotic excoriation, dermatillomania or psychogenic excoriation. Excoriation (skin picking) disorder (SPD) is one of the new diagnoses in DSM-5. Diagnostic criteria for SPD includes: (1) recurrent skin picking resulting in skin lesions; (2) repeated attempts to decrease or stop skin picking; (3) the skin picking causes clinically significant distress or impairment in social, occupational, or other important areas of functioning; (4) the skin picking is not attributable to the physiological effects of a substance (e.g., cocaine) or another medical condition (e.g., scabies); and (5) the skin picking is not better explained by symptoms of another mental disorder (e.g., delusions or tactile hallucinations in a psychotic disorder, attempts to improve a perceived defect or flaw in appearance in body dysmorphic disorder, stereotypies in stereotypic movement disorder, or intention to harm oneself in non-suicidal self-injury).

Epidemiology and Clinical Features. SPD is a common skin disorder characterized by excessive picking, scratching, digging, rubbing, gouging, or squeezing the normal skin or previously diseased skin (mostly juvenile acne). Patients suffer from irresistible urges to excoriate their normal-appearing skin or visible pre-existing abnormalities like acne, papules, scabs, scars, or insect bites. The classic lesions of psychogenic excoriation are characterized by clean, linear erosions, scabs, and scars that can be hypopigmented or hyperpigmented. The distribution of excoriations is mostly confined to accessible areas like face, arms (mostly extensor side),

and anterior thigh, with sparing of middle and upper lateral back (butterfly sign). This distribution of skin lesions can give the clinician a valuable clue, but it is not diagnostic of psychogenic excoriation because any condition that has pruritus as a primary symptom (such as hepatic, uremic, or idiopathic pruritus) can result in a similar distribution of skin lesions. Once the patient begins the process of self-picking, a focal itch sensation in the uninvolved skin may initiate scratching, or the tissue damage caused by scratching may trigger further itching. The itch-scratch cycle may become part of the clinical picture and promote the severity and chronicity of psychogenic excoriation. Lichen simplex chronicus is a chronic pruritic lichenoid plaque with thickening of the skin areas, which is provoked and maintained by scratching or rubbing. Before making a diagnosis of psychogenic excoriation, other conditions that can cause pruritus or excoriation, such as the dermatological, medical, and neurological illnesses listed earlier must be ruled out. In the general population, life-time prevalence of SPD in adults is 1.4 percent or higher. Three-quarters or more of the individuals are female. Psychogenic excoriation is reported to occur in 9 percent of dermatology outpatients with pruritus. Most studies report a mean onset of age between 30 and 45 years, but acne excoriée can begin in adolescence and continue into adulthood.

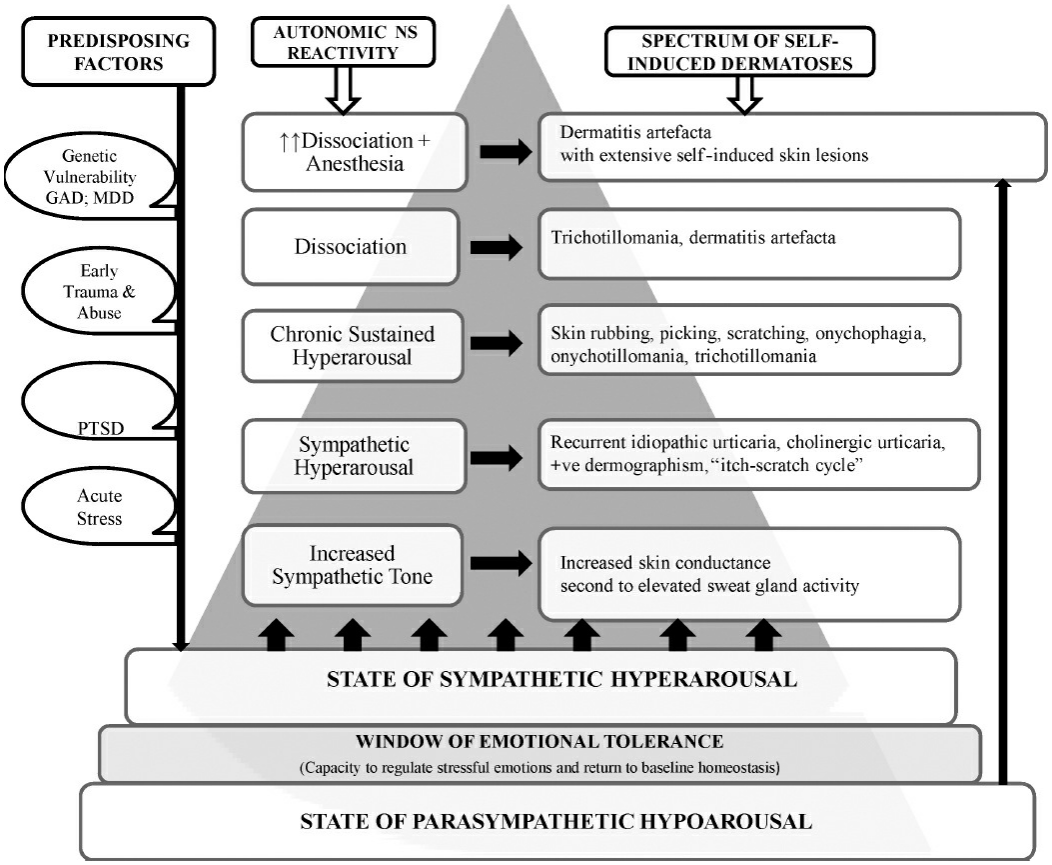

FIGURE 27.12–10. "Window of Tolerance" model for autonomic nervous system reactivity and symptoms of self-induced dermatoses. (Adapted and modified from Gupta MA. Emotional regulation, dissociation, and the self-induced dermatoses: clinical features

and implications for treatment with mood stabilizers. *Clin Dermatol*. 2013;31(1):110–117.)

Psychiatric Comorbidity. Skin picking may be preceded or accompanied by various emotional states. It may be triggered by feelings of anxiety or boredom, may be preceded by an increasing sense of tension (either immediately before picking the skin or when attempting to resist the urge to pick), and may lead to gratification, pleasure, or a sense of relief when skin or scab has been picked. In SPD some individuals engage in skin picking that is more focused, i.e., with preceding tension and subsequent relief, features that are consistent with obsessive-compulsive symptoms. Some individuals may engage in more automatic picking, with the picking seeming to occur without full awareness and without preceding tension. Although the DSM-5 does not specifically mention that such patients may be dissociative, some authors postulate that patients engaging in skin picking without preceding tension or full awareness are likely to have higher levels of emotional dysregulation and dissociation without necessarily meeting the criteria for a dissociative disorder. They propose that recognition of the dissociative component may be of clinical importance, as the standard approaches to treating obsessive-compulsive skin picking (with habit reversal therapy and SSRI antidepressants) are not sufficient when high levels of dissociation are present (validated by scores on a dissociative rating scale).

DSM-5 Axis I disorders have been found to occur in the lifetimes of 54.5 to 100 percent of individuals with SPD, as reported in various studies. The most common psychiatric disorders seen in these patients are OCD, trichotillomania (TTM), generalized anxiety disorder, and major depressive disorder. Generalized anxiety disorder, panic disorder, social phobia and posttraumatic disorder are present in 41 to 65 percent of patients. Mood disorders, including major depressive disorder, dysthymia, and bipolar disorder are present in 48 to 68 percent of the patients with psychogenic excoriations. Other comorbid OCRD disorders with SPD include body dysmorphic disorder, TTM, compulsive nail biting, and eating disorders. In one study, 45 percent of patients with BDD met the criteria of SPD. Patients with OCD may have obsessions about contamination and skin abnormalities or may be preoccupied with having smooth skin, flawless complexion, and cleanliness. This preoccupation with appearance might be severe enough to meet criteria for an OCD. “OCRD due to another medical condition” refers to obsessive-compulsive symptoms best explained by the direct pathophysiological consequence of another medical condition excluding acne and includes a specifier “with skin picking symptoms.” This diagnosis is applicable in a wide range of dermatological disorders, e.g., pruritic lesions. While pruritus may occur with psychogenic excoriation, compulsive and impulsive behaviors are characteristic of SPD diagnosis.

Treatment. There are few controlled data on the pharmacological treatment of psychogenic excoriation. Most evidence comes from case studies, open trials, and small double-blind studies, and supports treatment with SSRIs. Fluoxetine (Prozac) is the most studied SSRI in the treatment of psychogenic excoriation. In one double-blind, 10-week study, at a mean dose of 55 mg per day, fluoxetine was superior to placebo in reducing the psychogenic excoriations. Other open trials and case studies report benefit with treatment with other SSRIs, such as sertraline (Zoloft), fluvoxamine (Luvox), and the tricyclic antidepressants (TCAs) doxepin and clomipramine (Anafranil). SSRIs may be particularly effective in patients with prominent compulsive features, comorbid mood disorders, and anxiety disorders. Naltrexone, an orally active opiate antagonist, was effective at a dose of 50 mg per day for reducing the psychogenic excoriation with pruritus. However, many patients with excoriation experienced pain with naltrexone treatment, which limits its value. Other case reports suggest antipsychotic medications such as pimozide, aripiprazole (Abilify), and olanzapine (Zyprexa) may be of benefit. Patients with predominantly impulsive feature or comorbid bipolar disorder are likely to benefit from traditional anticonvulsants such as lamotrigine (Lamictal) or divalproex (Depakote), or from atypical antipsychotics such as olanzapine. In one study, lamotrigine was associated with improvements in two-thirds of subjects with pathological skin picking. A recent systematic review and meta-analysis of psychiatric treatments for SPD found strong effect size for SSRIs, lamotrigine and behavioral therapy. The authors suggest behavior therapy may be the first-line treatment for SPD with potential for long term therapeutic effects.

Trichotillomania

François Hallopeau coined the term trichotillomania (TTM) in 1889, from the Greek words *tricho* (hair), *tillo* (pull), and *mania* (fury). Hair pulling has been described since antiquity. *Trichobezoar* (hair ball in stomach) was first described in 18th-century medical literature. Hippocrates advised physicians to assess hair pulling as part of a routine mental status examination and described a case of hair pulling in the apparent context of depression (Fig. 27.12–11).

Case Report 2. Trichotillomnia

Clinical Features and Epidemiology. Trichotillomania (hair pulling disorder) is currently classified in DSM-5 under OCRD. The DSM-5 diagnostic criteria include: (1) recurrent pulling of one's hair resulting in hair loss; (2) repeated attempts to decrease or stop hair pulling; (3) the hair pulling causes clinically significant distress or impairment in social, occupational, or other important areas of functioning; (4) the hair pulling or hair loss is not attributable to another medical condition (e.g., a

dermatological condition); and (5) the hair pulling is not better explained by the symptoms of another mental disorder (e.g., attempts to improve a perceived defect or flaw in appearance in body dysmorphic disorder).

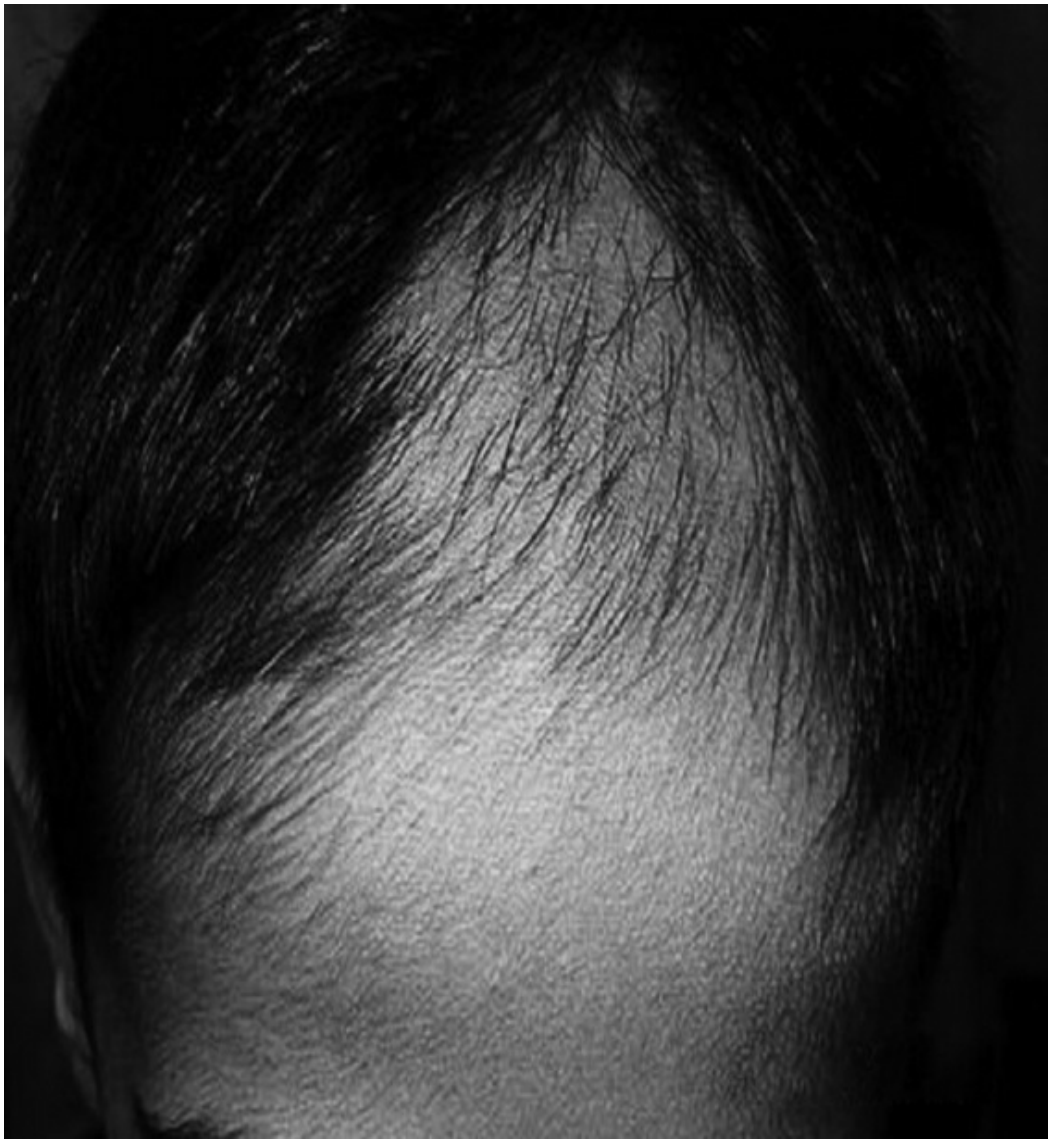

FIGURE 27.12–11. Picture of patient with Trichotillomania with alopecia of frontal scalp. Areas of different hair lengths with blunt or tapered ends, areas of black dots on the surface of the scalp. The hair does not pull out easily (negative "pull test").

Clinically, patients present with areas of different hair lengths with blunt or tapered ends, or areas of black dots on the surface of the scalp. The hair does not pull out easily (negative 'pull test'). TTM affects women much more commonly than men (10 to 1 ratio).

TTM Subtypes. Hair pulling typically involves scalp, while other common sites are eyebrows and eyelashes. Two types of hair pulling have been described. *Focused pulling* is viewed as use of an intentional act to control unpleasant personal experiences, such as an urge, bodily sensation

(e.g., itching or burning), or thought. This behavior is thought to be a compulsion in that the patient resists the act of pulling and feels relieved after pulling. In contrast, *automatic pulling* seems to occur outside the person's awareness and most often during sedentary activities. It is akin to an impulse control disorder. Many patients have a combination of these types of hair pulling. As pointed out earlier, some studies have recognized that lack of awareness in hair pulling is associated with possible "dissociation and emotional dysregulation" in patients with TTM.

Pathophysiology. Patients with TTM may have comorbid mood, anxiety, and eating disorders. Major depression was the most commonly associated mood disorder, and generalized anxiety disorder was the most common anxiety disorder in a large series of patients with TTM. Comorbid cluster B (borderline 18 percent, histrionic 26 percent) and cluster C (obsessive-compulsive 27 percent, avoidant 14 percent) personality disorders are commonly seen in patients with TTM.

Familial association between TTM and OCD has been found in some studies. Patients with TTM often have a family history of OCD or other anxiety disorders. OCD is thought to be related to overactivity of cortico-striatal-thalamic-cortical (CSTC) circuits, which is also hypothesized to play a role in a range of dysfunctional repetitive behaviors, including TTM. One magnetic resonance brain imaging study reported decreased putamen volume (normal caudate volume) in patients with TTM compared to healthy controls. These findings are suggestive of CSTC circuit involvement. A PET study of patients with TTM showed that treatment response to clomipramine was negatively correlated with anterior cingulate and orbitofrontal metabolism. In another study of patients with TTM, high baseline CSF 5-hydroxyindoleacetic acid (5-HIAA) levels predicted a better response to SSRI treatment, consistent with what is seen in OCD.

Treatment. It is recommended to consider the age of onset to decide what intervention may be most helpful. In preschool children, where TTM is more often automatic type, guiding the parents and simply raising their awareness is most effective. Preadolescents and young adults are often not aware of their hair pulling, and require establishing a diagnosis, as well as reassurance that hair is normal and will regrow when it is left alone. Behavioral modification treatments are recommended if the condition becomes chronic. Adult-onset TTM is usually associated with psychiatric disorders and requires psychopharmacological and behavioral interventions. In an earlier double-blind crossover study, clomipramine was found to be more effective than desipramine for treatment of TTM. More recent open trials of SSRIs (fluoxetine, fluvoxamine, paroxetine, sertraline, citalopram [Celexa], and venlafaxine) for the treatment of TTM suggest positive treatment response with short-term and long-term

benefits. Given their superior tolerability compared to TCAs, the SSRIs arguably remain the first-line choice medication treatment for TTM. Various case reports have also documented the potential efficacy of atypical antipsychotics (olanzapine, risperidone, and aripiprazole) in treatment of TTM as monotherapy or augmentation with SSRIs. Mood stabilizers (lithium, oxcarbazepine, topiramate, and Lamictal) have been used in TTM patients with mood disorders.

Few studies have examined the effects of Naltrexone on hair pulling. Blocking the opioid binding may decrease the positive reinforcement (as in addictive behaviors) from hair pulling. An open label study of 14 children with TTM found Naltrexone reduced hair pulling; another study involving adults (DBPC) reported Naltrexone had some improvement in hair pulling behavior. The most recent development in pharmacotherapy of TTM involves use of the glutamate modulator *N*-acetylcysteine (NAC). In a 12-week double-blind placebo-controlled trial, NAC was used in a dosage of 1,200 to 2,400 mg per day and reported that 56 percent of TTM patients showed significant improvement. Other studies using NAC reported treatment outcomes comparable to behavioral intervention like CBT. Notably, NAC is not an FDA-regulated product, and it is readily available in health food stores.

Most recent literature on treatment of TTM supports the use of CBTs that include the "habit reversal training" as the first-line treatment option. *Habit reversal training* has three core elements: first, "awareness training" (self-monitoring), to become aware of urge that precedes pulling; second, "stimulus control," to reduce the likelihood that pulling behavior begins; and third, "competing response training," to engage in a behavior, at the earliest sign of pulling, that is physically incompatible with hair pulling. Preliminary evidence suggests that combined use of behavioral therapies and medications may be more effective than either treatment alone.

Onychophagia

Onychophagia (or chronic nail biting) is a common behavior among children, beginning as early as 4 years of age and having a peak during the teens. Patients with onychophagia have clinical features of repetition, resistance, and relief, comparable to hair pulling and skin picking. Patients with severe and chronic onychophagia can experience significant dental and medical problems, such as hand infections and mandibular deformities. A combination of thumb sucking is also frequent. Onychophagia behavior is usually increased during period of unresolved conflicts or stressful situations. Behavioral treatments like habit reversal interventions have also been efficacious.

Body Dysmorphic Disorder

Prior to DSM-5, BDD (formally known as *dysmorphophobia*) has been considered under somatoform disorders, but changes in DSM-5 have included BDD under the group of OCRD. The changes from DSM-IV-TR criteria mainly focus on increasing evidence of a relationship of BDD to obsessive-compulsive spectrum disorders.

DSM-5 Criteria for Diagnosis of BDD. DSM-5 Criteria for Diagnosis of BDD includes the following:

(1) Preoccupation with one or more perceived defects or flaws in physical appearance that are not observable or appear only slight to others. (2) At some point during the course of the disorder, the individual has performed repetitive behaviors (e.g., mirror checking, excessive grooming, skin picking, or reassurance seeking) or mental acts (e.g., comparing one's appearance with that of other people) in response to the appearance concerns. (3) The preoccupation causes clinically significant distress or impairment in social, occupational, or other important area of functioning. (4) Appearance preoccupation does not meet diagnostic criteria for an eating disorder. In DSM-5, a specifier indicates degree of insight regarding BDD: "with good or fair insight," "with poor insight," and "with absent insight/delusional beliefs." In contrast to the DSM-IV-TR, in DSM-5, the delusional variant of BDD (individuals completely convinced that their perceived flaws or defects are truly abnormal appearing) is no longer coded as both a delusional disorder, somatic type, and BDD. In DSM-5 this presentation is coded only as BDD with the "absent insight/delusional beliefs" specifier.

Clinical Features. Patients present with imagined or slight flaws of one or many body areas, most commonly the facial skin (e.g., perceived acne, scars, lines, wrinkles, paleness, vascular markings, or complexion changes), hair (e.g., "thinning" hair or "excessive" body or facial hair), or nose (e.g., size or shape) or other facial features. The other body regions involved may include genitals, breasts, buttocks, abdomen, upper and lower extremities, overall body size, build, and muscularity. BDD usually manifests during adolescence and has equal sex distribution. Patients with BDD engage in various time-consuming activities to improve the perceived defect such as camouflaging, mirror checking, comparison of defects with the same body parts on others, questioning, and reassurance seeking. Some patients go through several cosmetic and surgical procedures for their perceived defects. The functional impairment and distress expressed by the patients with BDD include severe depression and low self-esteem, anxiety, withdrawal from relationships, social isolation, and work-related problems. BDD is associated with a high rate of suicidal ideation and suicide attempts, especially among the adolescent age group, more often those

having severe acne. The shame and humiliation associated with BDD plays a major role in the underreporting of this disorder.

Treatment. The prevalence of BDD in dermatology patients is 9 to 15 percent, and 7 to 8 percent among cosmetic surgery patients in the United States. This has important implications in dermatology, as clinical treatment guidelines for dermatological disorders increasingly emphasize the importance of assessing psychosocial comorbidity, including aspects of CBI (cutaneous body image), in order to achieve optimal treatment outcomes. Patients with BDD can be challenging to engage in psychiatric treatment, and an empathic understanding and establishment of good therapeutic rapport on the part of the clinician is crucial. The efficacy of cognitive–behavioral therapy with focus on cognitive restructuring has been demonstrated. SSRIs are the drug of choice in the treatment of BDD. A large study of tricyclic antidepressants has shown a significantly better efficacy of clomipramine over desipramine, but less favorable side effects limits their use.

FACTITIOUS DISORDERS: FACTITIOUS DERMATITIS

Factitious dermatitis or *dermatitis artefacta* (DA) is a disorder in which skin is the target of self-inflicted injury and the patient uses more elaborate methods than simple excoriation to self-induce skin lesions.

Epidemiology and Clinical Features

The morphology of factitious dermatitis can imitate most cutaneous disease: "typical is what is atypical." Factitious dermatitis can present as an aggravation of dermatosis, targeting a variety of skin lesions including blisters, ulcers, erythema, edema, purpura, and sinuses. The morphology of factitious dermatitis lesions is often bizarre and linear, with clear-cut, angulated, or geometric edges. Presence of completely normal, unaffected skin adjacent to the horrific-looking lesions is a clue to the diagnosis of factitious dermatitis. The skin damage can be extensive with full thickness skin loss and severe scarring requiring plastic surgery. The patient's description of history of the skin lesions is usually vague and hollow and lacks detail about the appearance and evolution of the lesions. There is no evidence of natural progression of lesions through different stages of development, as one may expect in a typical dermatosis. Factitious dermatitis lesions are usually located in areas easily reached by the dominant hand. Patients with factitious dermatitis are seemingly unmoved by the disfiguring painful lesions and typically deny the self-inflicted nature of the disorder (Fig. 27.12–12).

FIGURE 27.12–12. Picture of a patient with dermatitis artefacta: self-inflicted sharp, rounded lesions on the forearm made with a razor during times of stress and emotional dysregulation.

Case Report 3. Factitious dermatitis

Factitious dermatitis is seen in 3 percent of dermatology patients. It is mostly gender-specific, with a ratio of female-to-male as high as 8 to 1. It can present at any age, but occurs most frequently in adolescents and young adults. The onset of self-inflicted behavior often follows severe psychosocial stress involving isolation, loss or threatened loss, and relational difficulties. It is not uncommon to get a history of prior affiliation within the health care field. Factitious dermatitis patients often have a history of other psychosomatic illnesses and have had multiple consultations with physicians and been treated with "ineffective" treatments. These patients need to assume a sick role is supported by the following: a passive concern and indifference to illness, center of getting medical attention, aim of becoming a difficult case in terms of diagnosis and treatment, or secondary gains like pursuit of sickness leave or benefit. The course of this disorder varies from a brief episode to a lifelong behavior; one study reported that 30 percent of factitious dermatitis patients continued to produce lesions over more than 12 years.

Psychopathology

Dermatitis artefacta is associated with several psychiatric disorders (besides a primary factitious etiology) such as schizoaffective psychosis, adjustment disorder, major depressive disorders, PTSD, delusional parasitosis, mental retardation, and pervasive developmental disorders, as well as with pathological intoxication with alcohol, cocaine, or lysergic acid

diethylamide (LSD). The most common comorbid psychiatric condition associated with factitious dermatitis is borderline and histrionic personality disorder. Patients with factious disorders may have a history of contact with multiple health care services, especially psychiatry, and carry a volume of medical notes. These patients are often aware of their impulsive acts in the context of emotional and relational instability, which may resemble obsessive-compulsive spectrum disorder. Factitious dermatitis may have a connection with body image disturbance as a subtype of BDD, in which the individual is preoccupied with skin-related personal ugliness, cosmetic defect, or deformity. A patient with severe BDD, focused on skin, may pursue requests for incisive dermatological treatments (laser surgery or transplants) or cosmetic surgery, and when denied access to such treatments, may take initiative to "self-treatment." Psychiatric disorders are often the cause of self-mutilation in dermatology and may also be the consequence of the visible disfigured lesions, resulting in anxiety, social phobia, diminished self-esteem, and depression. A connection between DA patients' self-inflicted behaviors has been highlighted in literature to the extent that DA may represent a marked suicidal behavior.

Treatment

There are no controlled trials of treatment of factitious dermatitis. Patients usually are resistant to accept a referral to a psychiatrist, and direct confrontation may jeopardize the doctor–patient relationship. An effort to develop an empathic therapeutic relationship between the dermatologist and the patient is critical to the management of these patients. Once such a therapeutic relationship is established, some patients may accept appropriate therapeutic approaches and psychiatric referral.

PSYCHOGENIC PURPURA

Psychogenic purpura (Gardner–Diamond Syndrome), also known as autoerythrocyte sensitization syndrome, is an extremely rare condition typically noted in women with psychiatric comorbidity. Psychogenic purpura is regarded primarily as an autoimmune vasculopathy with sensitization to phosphatidylserine, a component of erythrocyte stroma. Manifestation usually occurs via spontaneous development of painful edematous skin lesions, predominantly in the extremities, which may be isolated or multiple, progressing to ecchymosis over the next 24 hours. Development generally follows severe stress or emotional trauma or is comorbid with certain psychiatric psychopathology. Associated initial symptoms include burning and pain, followed by erythema, swelling, and pruritus, ultimately evolving to ecchymosis. Medical work-up, including hemostatic and blood coagulation tests, is normal. The age of onset is variable; women are affected more frequently. The proposed mechanisms

of psychogenic purpura include autoerythrocyte sensitization, conversion reaction, and factitious disorder. The evidence for autoerythrocyte sensitization is very limited. The psychogenic cause of purpura has been supported by some studies because of the association between purpura onsets and psychosocial stresses. Most patients also have multiple other unexplained somatic symptoms, as well as psychiatric comorbidity, including generalized anxiety disorder, major depressive disorder, and personality disorders (borderline personality disorder). It is difficult to differentiate convincingly between conversion and factitious purpura because patients often have features of both.

Table 27.12–5.
Psychotropic Medications Used in Treatment of Psychocutaneous Disorders

PRURITUS	**SSRIs and TCAs** (amitriptyline and doxepin): commonly useful in pruritus **Doxepin 5% topical cream**: FDA-approved for short-term management of pruritus in AD **Gabapentin, pregabalin**: effective in prurigo nodularis, neuropathic pain, and fibromyalgia **Aprepitant** (neurokinin receptor [NKR]-1 antagonist): effective in chronic refractory pruritus. NKR-1 is a receptor for substance P, an important itch mediator. **Naltrexone** (opioid receptor antagonist): used in Rx of OCRD
BODY-FOCUSED REPETITIVE BEHAVIORS (under OCRD): SPD, TTM, BDD, dermatitis artefacta	**SSRIs** (fluoxetine, paroxetine, sertraline, citalopram, escitalopram, and fluvoxamine): considered first-line treatment for OCRD **Clomipramine** (TCA): found to be effective in TTM, BDD, OCD; use caution with TCAs (side effects). **Naltrexone** (opioid receptor antagonist): In one study reduced pruritus by 50% in OCRD, body repetitive disorders; caution if patient is opiates abuser. **N-Acetylcysteine** (NAC): treatment efficacy comparable to CBT in one study, promising treatment for TTM **Olanzapine and aripiprazole**: +ve case reports for Rx of SPD and TTM **Mood stabilizers** (lamotrigine, lithium, valproic acid): few positive studies; indicated for Rx of emotional dysregulation/dissociation and comorbid mood disorder
CUTANEOUS DYSESTHESIAS: Glossodynia, vulvodynia, trichodynia/scalp dysesthesia urogenital and rectal pain syndrome	**TCAs** (doxepin, nortriptyline, imipramine, desipramine, and particularly amitriptyline): long h/o use, effective Rx **SSRIs**: considered first-line Rx **Mirtazapine**: useful when anticholinergic side effects of TCAs cannot be tolerated **Duloxetine**: has efficacy in pain, neuropathic disorders, and fibromyalgia. **Gabapentin**: useful for neuropathic pain and pruritus, safe, minimal drug interaction, and no abuse potential.
DELUSIONAL	**Pimozide**: a potent neuroleptic, most commonly advocated, 60–

PARASITOSIS	100% response rate; monitor risk of QT prolongation. **Risperidone, olanzapine, amisulpride**: 75% response rate; preferred over pimozide due to improved side effect profile.

Rx, treatment; AD, atopic dermatitis; SSRIs, selective serotonin reuptake inhibitors; TCA, tricyclic antidepressants; TTM, trichotillomania; SPD, skin picking disorder; OCRD, obsessive-compulsive and related disorders; BDD, body dysmorphic disorder.

PSYCHOPHARMACOLOGY OF PSYCHOCUTANEOUS DISORDERS

Psychiatric therapies, including psychotropic agents such as antidepressants, antianxiety, and antipsychotic drugs, may be used as adjunctive therapies in three major clinical situations: treatment of a dermatological symptom that represents an underlying psychiatric disorder, such as delusions of parasitosis; management of psychiatric disorders associated with a primary dermatological disorder, such as major depressive disorder; and when certain pharmacological properties of the psychotropic agents are desired, for example, the antihistaminic effect of the tricyclic antidepressants. Please refer to section on psychotropic medications in this textbook for detailed understanding of use of antidepressants, mood stabilizers, and antipsychotics. Table 27.12–5 provides a summary of psychotropic medications used in treatment of psychocutaneous disorders, based on various studies and clinical guidelines.

Adverse Cutaneous Reactions of Psychiatric Medications

Dermatological side effects of psychotropic medications are fairly common and the majority of them are benign and easily treatable. Most drug reactions are hypersensitivity reactions with an underlying immunological mechanism; however, some reactions are attributed to cumulative dose toxicity, drug overdose, and drug–drug reactions. Dermatological side effects of psychotropic medications are summarized in the Table 27.12–6.

FUTURE DIRECTIONS

Psychocutaneous medicine involves several different types of disorders at the interface between dermatology and psychiatry. Psychocutaneous disorders are a diverse group, including some in which psychopathology plays the primary etiological role in the absence of a skin disease, others in which psychological factors exacerbate skin disorders, and yet others in which patients develop psychological complications as a result of disfigurement caused by skin disease. The psychology of the dermatological patient is an under-researched area. There is need for better understanding of certain psychological factors, such as the impact of skin disorders on child development, self-image (cutaneous body image), early relationships, psychosocial implications, and associated stigma and role of stress in exacerbation of skin disorders. An increased awareness of these factors

may help dermatologists appreciate the biopsychosocial aspects of individuals' skin diseases (biopsychosocial model, as emphasized by Engel), rather than a limited focus on skin lesions only. Findings from psychoneuroimmunology are beginning to shed light on the role of emotional stress in the pathogenesis of some skin disorders (e.g., atopic dermatitis, psoriasis), which bodes well for the further development of the biological aspects of psychosomatic medicine. More recent research explores the potential mechanisms that may mediate the bidirectional effects of stress on skin immune function, and provide targets for clinical interventions that may be designed to dampen or eliminate stress-induced exacerbation of skin immunopathology.

Recent changes in DSM-5 under the group of OCRD have direct implications in treatment of psychocutaneous disorders. Recurrent body-focused repetitive behaviors in SPD and TTM have both compulsive and dissociative features, the latter being more prevalent in TTM than SPD. The integumentary system is frequently the focus of tension-reducing behaviors in self-induced dermatoses. Targeting the treatment of "autonomic hyperarousal and dissociation" can improve therapeutic outcomes in SPD, TTM, and dermatitis artefacta. Recent studies suggest a bidirectional relationship between sleep and inflammatory skin disorders such as psoriasis and atopic dermatitis. The sleep–wake cycle is mediated by thermoregulation in skin via endogenous circadian rhythms. Sleep deprivation and repeated arousal in many sleep disorders lead to a heightened proinflammatory state. Thus, management of sleep disorders in patients with psychocutaneous disorders can potentially improve their skin condition and quality of life.

Patients with psychocutaneous disorders can be particularly challenging for dermatologists because they often do not agree to psychiatric referrals. However, they may be more willing to agree to psychiatric evaluation from a psychiatrist working within the dermatology clinic. Multidisciplinary collaboration between psychiatrists, psychologists, nurses, and dermatologists is likely to improve treatment outcomes and quality of life in patients with psychocutaneous disorders.

Table 27.12–6.
Dermatological Adverse Effects of Psychotropic Medications

Drugs	Skin Adverse Effects
MOOD STABILIZERS	Hair loss, acne, nail pigmentation, exacerbation of psoriasis
Lithium	Hair loss, hair color change, lupus, scleroderma, skin vasculitis
Valproic acid	Stevens–Johnson syndrome, toxic epidermal necrolysis, angioedema
Lamotrigine	Pruritic rash, hair loss, S–J syndrome, hypersensitivity reaction
Carbamazepine	
ANTIDEPRESSANTS	Allergic reaction (hives, urticaria), excessive sweating, pruritus; hair loss
SSRIs	reported with fluoxetine
Bupropion	Erythroderma, erythema multiforme (rare)

Venlafaxine	Erythroderma, erythema nodosum
Duloxetine	Erythema multiforme
Mirtazapine	Erythroderma
TCAs	Photosensitivity, erythroderma
Vortioxetine	Pruritus
Vilazodone	Drug rash
ANTIPSYCHOTICS	Drug hypersensitivity
Haloperidol	Skin pigmentation (blue-gray discoloration)
Phenothiazine	Erythema multiforme, S–J syndrome, drug hypersensitivity
Clozapine	Erythema multiforme, erythroderma
Risperidone	Drug rash with eosinophilia
Olanzapine	Erythroderma
Ziprasidone	Drug rash (rare)
Aripiprazole	Drug rash, pruritus, angioedema (rare)
Lurasidone	Drug rash
Paliperidone	
ANXIOLYTICS	Photosensitivity
Alprazolam	

References

Anthis L. Skin disease and relationship. In: Walker C, Papadopoulos L, eds. *Psychodermatology: The Psychological Impact of Skin Disorders*. Cambridge, UK: Cambridge University Press; 2005.

Bender BG, Ballard R, Canono B, Murphy JR, Leung DY. Disease severity, scratching, and sleep quality in patients with atopic dermatitis. *J Am Acad Dermatol*. 2008;58:415–420.

Cash TF, Pruzinsky T. *Body Images: Development, Deviance and Change*. New York: Guildford Publications Inc.; 1997.

Corrigan FM, Fisher JJ, Nutt DJ. Autonomic dysregulation and the Window of Tolerance model of the effects of complex emotional trauma. *J Psychopharmacol*. 2011;25:17–25.

Craig AD. How do you feel—now? The anterior insula and human awareness. *Nat Rev Neurosci*. 2009;10(1):59–70.

*Dhabhar FS. Psychological stress and immunoprotection vs. immunopathology in the skin. *Clin Dermatol*. 2013;31:18–30.

Dhabhar FS, McEwen BS. Bidirectional effects of stress and glucocorticoid hormones on immune function: possible explanations for paradoxical observations. In: Ader R, Felten DL, Cohen N, eds. *Psychoneuroimmunology*. 3rd ed. San Diego, CA: Academic Press; 2001:301–338.

Dixon LJ, Lee AA, Viana AG, McCowan, Brodell RT, Tull MT. Anxiety sensitivity in dermatological patients. *Psychosomatics*. 2016;57(5):498–504.

Gieler U, Niemeier V, Kupfer J, Bsrosig B. Psychophysiological aspects of atopic dermatitis. In: Koo JY, Lee CS, eds. *Psychocutaneous Medicine*. New York: Marcel Dekker; 2003.

Ginsberg IH. Psychosocial aspects of skin disease on the patient. In: Koo JYM, Lee CS, eds. *Psychocutaneous Medicine*. New York: Marcel Dekker; 2003.

Goldstein DS, McEwen B. Allostasis, homeostats, and the nature of stress. *Stress*. 2002;5:55–58.

Grant JE, Chamberlain SR, Redden SA, Leppink EW, Odlaug BL, Kim SW. N-acetylcysteine in the treatment of excoriation disroder: a randomized clinical trial. *JAMA Psychiatry*. 2016;73(5):490–496.

Grant JE, Odlaug BL, Chamberlain SR, Kim SW. A double-blind, placebo-controlled trial of lamotrigine for pathological skin picking. Treatment efficacy and neurocognitive predictors of response. *J Clin Psychopharmacol*. 2010;30:396–403.

*Gupta MA. Emotional regulation, dissociation, and the self-induced dermatoses: clinical features and implications for treatment with mood stabilizers. *Clin Dermatol*. 2013;31(1):110–117.

Gupta MA, Gupta AK. Evaluation of cutaneous body image dissatisfaction in the dermatology patient. *Clin Dermatol*. 2013;31:72–77.

Hashiro M, Okumura M. The relationship between the psychological and immunological state in patients with atopic dermatitis. *J Derm Sci*. 1998;16:231–235.

Hirotsu C, Rydlewski M, Araujo MS, Tufik S, Andersen ML. Sleep loss and cytokines levels in an experimental model of psoriasis. *PLoS One*. 2012;7(11):e51183.

*Hong TL, Wallergren MD, Yosipovitch MD. Psychosomatic factors in pruritus. *Clin Dermatol*. 2013;31:31–40.

Hundley JL, Yosipovitch G. Mirtazapine for reducing nocturnal itch in patients with chronic pruritus: a pilot study. *J Am Acad Dermatol*. 2004;50:889–891.

Jafferany M, Stoep AV, Dumitrescu A, Hornung RL. Psychocutaneous disorders: a survey study of psychiatrists' awareness and treatment patterns. *South Med J*. 2010;103(12):1199–1203.

Koo J. Psychodermatology: a practical manual for clinicians. *Curr Probl Dermatol*. 1995;7:203–225.

Musalek M. Delusional parasitosis. In: Koo JYM, Lee CS, eds. *Psychocutaneous Medicine*. New York: Marcel Dekker; 2003.

Nowak DA, Wong SM. DSM-5 update in psychodermatology. *Skin Therapy letter*. 2016;21(3).

Panconesi E, Hautmann G. Stress and emotions in skin diseases. In: Koo JYM, Lee CS, eds. *Psychocutaneous Medicine*. New York: Marcel Dekker; 2003.

Patel T, Ishiuji Y, Yosipovitch G. Nocturnal itch: why do we itch at night? *Acta Derm Venereol*. 2007;87(4):295–298.

Picari A, Abeni D. Stressful life events and skin diseases: disentangling evidence from myth. *Psychother Psychosom*. 2001;70:118–136.

Selles RR, McGuire JF, Small BJ, Storch EA. A systematic review and meta-analysis of psychiatric treatments for excoriation (skin-picking) disorder. *Gen Hosp Psychiatry*. 2016;41:29–37.

Singh L, Pang X, Alexacos N, Letourneau R, Theoharides T. Acute immobilization stress triggers skin mast cell degranulation via corticotropin releasing hormone, neurotensin, and substance P: a link to neurogenic skin disorders. *Brain Behav Immune*. 1999;13:225–239.

Tey HL, Wallengren J, Yosipovitch G. Psychosomatic factors in pruritus. *Clin Dermatol*. 2013;31(1):31–40.

*Thompson A. Coping with chronic conditions: factors important in explaining individual variation in adjustment. In: Walker C, Papadopoulos L, eds. *Psychodermatology: The Psychological Impact of Skin Disorders*. Cambridge, UK: Cambridge University Press; 2005.

Yang YW, Kang JH, Lin HC. Increased risk of psoriasis following obstructive sleep apnea: a longitudinal population-based study. *Sleep Med*. 2012;13(3):285–289.

Yosipovitch G, Ishiuji Y, Patel TS, et al. The brain processing of scratching. *J Invest Dermatol*.

2008;128(7):1806–1811.
*Zane LT. Psychoneuroendocrinimmunodermatology. In: Koo JYM, Lee CS, eds. *Psychocutaneous Medicine*. New York: Marcel Dekker; 2003.

▲ 27.13 Organ Transplantation

ANDREA FRANCES DIMARTINI, M.D., MARY AMANDA DEW, PH.D., AND CATHERINE CHANG CRONE, M.D.

Although solid organ and other tissue transplants had been attempted since the early 1900s, it was not until 1954 that the first successful organ transplant was accomplished by Joseph Murray using an identical twin brother as a kidney donor. Despite this surgical breakthrough, routine transplantation of organs would not be possible until the challenge of adequate suppression of the immune system was resolved. In the early 1960s, the development of chemical immunosuppression allowed the first deceased donor kidney, liver, lung, and pancreas transplants to be performed, followed by the first heart transplant in the late 1960s. Two decades later the discovery and development of cyclosporine (Neoral and others), a much more powerful and specific immunosuppressant drug, provided selective immunoregulation of T cells without excessive toxicity. Cyclosporine's approval by the FDA in 1983 for use in solid organ transplantation led to a dramatic increase in the numbers of transplants performed, decreases in recipient morbidity, and increases in posttransplant survival rates. This drug and others that followed changed organ transplantation from an experimental procedure to the standard of care for many types of advanced organ disease.

In 1984 the US Congress enacted the National Organ Transplant Act (NOTA), establishing a regionalized system of equitable organ distribution, and the US Department of Health and Human Services contracted with the United Network of Organ Sharing (UNOS) to administer the nation's Organ Procurement and Transplantation Network (OPTN). A central feature of NOTA was a prohibition on the buying or selling of human organs for transplantation, heralding some of the controversies to come in this complex field. In addition to the OPTN's primary goals of increasing the efficiency of organ sharing and the supply of donated organs, UNOS collects extensive data on all solid organ transplants in the United States (e.g., wait lists counts, survival rates) and serves as an educational resource and a policy-making organization. UNOS data, maintained by the Scientific Registry of Transplant Recipients (SRTR), are available for public access on the SRTR Web site (http://srtr.transplant.hrsa.gov). Although transplantation has become a life-sustaining procedure for those with advanced organ disease, the stresses inherent in the process—including the

severe medical illness of the candidate, lack of available organs, and the uncertainty of surviving to transplantation, risk of surgery, and the challenges of postoperative recovery—place enormous demands on transplant candidates, recipients, and their caregivers. Additionally, the organ shortage—as well as extensive health care resources and the degree of personal responsibility required of transplant patients for successful outcomes—has created an environment where transplant programs attempt to select candidates deemed to have the best chance of optimal posttransplant outcomes. Due to these issues, psychosomatic medicine specialists and other mental health professionals are frequently involved in the evaluation, selection, treatment, and long-term care of transplant candidates and recipients. Knowledge of the organ-specific disease process is important, but one should be additionally familiar with transplant-specific issues in order to care for these complex and unique patients. This section will review the essential elements of adult solid organ transplantation, including the epidemiology of transplantation, the common types of organ disease leading to transplantation, and the course and prognosis for the transplant patient. Within the course and prognosis the emphasis will be on the assessment of candidates and the specific phases of the transplantation process, as these areas must be understood in order for the mental health professional to provide appropriate patient care. In addition to the more common psychological factors and psychiatric disorders affecting the process and outcomes of organ transplantation, this section will also review issues particular to transplantation including bridge-to-transplant technology, living organ donation, treatment issues specific to transplantation, and neuropsychiatric immunosuppressive medication side effects.

ORGAN TRANSPLANTATION

Definition and Comparative Nosology

Organ transplantation includes most of the solid organs either singly or in combination: kidney, pancreas, kidney–pancreas, liver, heart, lung, heart–lung, intestine, multivisceral (e.g., multiple abdominal organs or heart–liver). A new area of organ transplantation recognized by the US Department of Health and Human Services in 2014 is vascular composite allograft (VCA) transplantation. VCAs are a heterogeneous category of "organs" that include face, limb, long bones and soft tissues. Because they are rare in the United States, the remainder of this section will focus on other types of organ transplantation.

Transplantation can be performed with deceased donor organs, using either complete or partial organs (i.e., split liver, single lung, partial intestine, or pancreatic islet cells), or can otherwise involve a living donor who provides either a complete organ (i.e., single kidney) or partial organ

(i.e., portions of the liver, lung, pancreas, and intestine). In 2001, the number of living donors exceeded the number of deceased donors (6,526 vs. 6,081) for the first time, with the majority of organs coming from kidney donors. However, following a highly publicized liver donor death, the numbers of partial hepatectomy living liver donors decreased from a peak of 520 in 2001 to slightly under 300 per year. Even a heart can be donated through a domino procedure in which a recipient who requires a heart–lung transplant but has a sufficiently healthy heart can donate his or her heart to another. Although rarely performed in the United States, domino hearts have been predominantly from patients with cystic fibrosis or pulmonary hypertension. Whole organ xenotransplantation has been unsuccessfully attempted in the past using pig and baboon organs. Currently, only pig islet cells are being investigated for transplantation in patients with pancreatic disease.

Epidemiology

The greatest obstacle to receiving a transplant is the shortage of donated organs. The number of wait-listed individuals is increasing far beyond the availability of donated organs. In 2003, the Department of Health and Human Services created the Organ Donation Breakthrough Collaborative to replicate and promote practices by the national organ procurement organizations that increase access to available organs. During the first year of the Organ Donation Breakthrough Collaborative, the average numbers of organs transplanted increased 6 percent, with an additional 8 percent increase the following year. By the end of 2014, over 500,000 individuals had received a transplant in the United States. However, while the collaborative has successfully increased the numbers of donor organs, the numbers of wait-listed patients has steadily risen (Fig. 27.13–1). The limited organ availability has increasingly led to the consideration and use of living organ donors. In kidney donation, to deal with incompatible living donor–recipient pairs, paired donation exchanges and, later in 2005, domino chains (non-directed donor who initiates a sequence of kidney transplants) using computer algorithms to match successive donor–recipient pairs were initiated to efficiently maximize the available living donor pool. Although the average amount of time on the wait list depends on the type of organ required, the severity of illness, and the blood type of the candidate, for all major organ types, over 40 percent of US wait-listed candidates waited 2 years or more for an organ. However, 10 to 18 percent die while on the wait list.

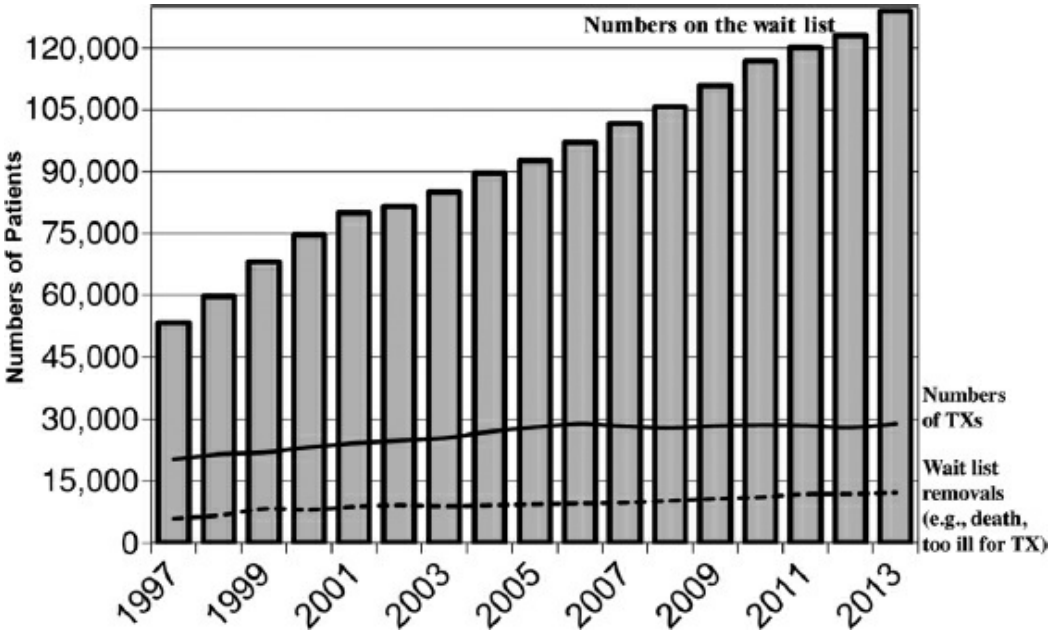

FIGURE 27.13–1. Wait list statistics in the United States: 1997 to 2013. (Data available on the Scientific Registry of Transplant Recipients, Health Resources and Services Administration website, Healthcare Systems Bureau, Division of Transplantation, Rockville, MD. Available at http://srtr.transplant.hrsa.gov)

Following transplantation, recipients of living kidney and living liver grafts show the highest long-term survival rates 10 years posttransplantation (78 percent and 73 percent respectively) (Fig. 27.13–2). Recipients of deceased donor kidneys, deceased donor livers, and heart recipients have somewhat lower 10-year survival (63 percent, 61 percent, and 58.5 percent, respectively) and lung and intestine recipients have the poorest 10-year survival (30 percent and 47 percent, respectively). Even so, these data are from transplants performed over 10 years ago, and with advances in technology, immunosuppression, and medical care the survival rates continue to improve. Graft survival rates are generally lower than patient survival rates, indicating that many transplant recipients eventually require retransplantation. Additionally, many immunosuppressive medications are nephrotoxic, and liver, heart, lung, and intestine transplant recipients may eventually develop renal failure requiring kidney transplantation.

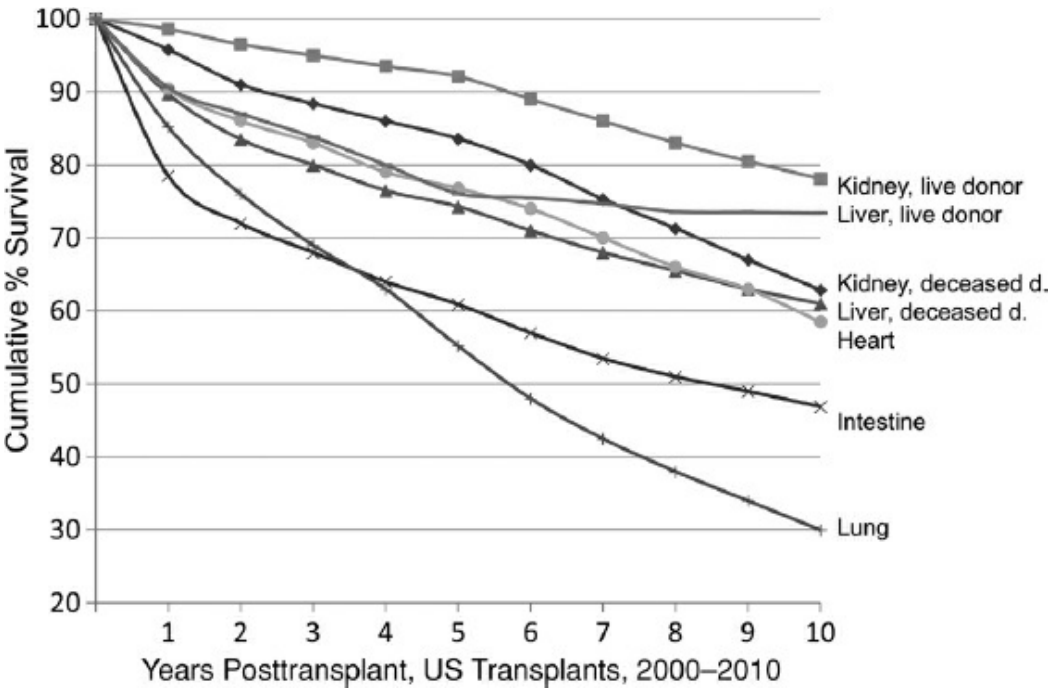

FIGURE 27.13–2. Transplant patient survival by organ type. (Data available on the Scientific Registry of Transplant Recipients Health Resources and Services Administration website, Healthcare Systems Bureau, Division of Transplantation, Rockville, MD. Available at http://srtr.transplant.hrsa.gov)

Etiology

The diseases that typically result in solid organ transplantation are organ-specific and are listed in Table 27.13–1 by organ type and frequency. Advanced organ disease can occur acutely as in fulminant hepatic failure (e.g., acetaminophen overdose, acute viral hepatitis) or can be a chronic progressive disease process (e.g., primary sclerosing cholangitis, chronic obstructive pulmonary disease). The disease could involve the organ parenchyma (e.g., diabetic nephritis), the vascular flow to the organ (e.g., Budd–Chiari syndrome), or transplantation could be required due to cancer in the organ (e.g., hepatocellular or renal cell carcinoma) or traumatic injury (e.g., blunt force trauma, inhalation/burns injury). In liver transplantation, recent shifts in the epidemiology of liver diseases, from an increase in the prevalence of nonalcoholic fatty liver disease (often from morbid obesity) and curative treatments for hepatitis C (often incurred through injected drug use), are changing the referral diseases and reasons for pretransplant psychiatric consultation.

Course and Prognosis

Transplantation can be envisioned as a continuum demarcated by specific phases. The development of advanced organ disease heralds the beginning of the process. From evaluation at the transplant program, to waiting on the national list, then early and longer term recovery and adaptation to life

as an organ recipient, each phase is associated with different stressors and requires differing skills and resources from patients, caregivers, and the mental health clinicians caring for them (Fig. 27.13–3). Mental health professionals are perhaps most frequently called on to aid in the evaluation of the prospective transplant candidate. Because these evaluations require a focus on transplant-specific areas somewhat different from a typical psychiatric consultation, they are addressed in detail below.

Clinical Evaluation for Transplantation. Due to the limited availability of donor organs; the considerable medical, social, and financial resources necessary for successful transplantation; the degree of personal responsibility of recipients; and the needed involvement of their family and other caregivers, careful assessment of transplant candidates and their caregivers is required. Pretransplant evaluations typically involve two components: a medical and a psychosocial assessment. During the medical assessment, patients are evaluated to determine whether there is a clear physiologic indication for transplantation based on UNOS criteria and to rule out significant medical comorbidities that pose additional risks or diminish the possibility of successful transplantation. For example, concurrent diabetes mellitus, morbid obesity, or a history of malignancy are conditions that can considerably raise the risk for posttransplant morbidity and mortality. Whether these conditions or others make a patient ineligible or less suitable for transplantation will partly depend on the severity of the comorbid diseases, but also on the transplant team's level of experience, ability, and comfort working with high-risk patients.

Table 27.13–1.
Diseases Resulting in Listing for Transplantation in the United States as of December 2012

Organ Type	Percentage	Organ Type	Percentage
Kidney, *n* = 92,885		**Liver, *n* = 15,308**	
Diabetes mellitus type I and II	34%	Hepatitis C	30%
Hypertensive nephrosclerosis		Alcoholic	24%
and malignant hypertension	25%	Cholestatic disease	8%
Chronic glomerulonephritis	14%	Hepatocellular carcinoma	7%
Polycystic	8%	Acute necrosis	2%
Other	19%	Hepatitis B	3%
Heart, *n* = 3,108		Other	26%
Cardiomyopathy	51%	**Lung, *n* = 1,616**	
Coronary artery disease	37%	Obstructive lung disease	
Congenital	5%	(e.g., COPD/emphysema)	45%
Valvular disease	2%	Restrictive lung disease	36%
Other	5%	Pulmonary vascular disease	6%

Intestine, *n* = **252**		Cystic fibrosis and immunodeficiency disorders	11%
Other short gut syndrome	34%	Other	2%
Necrotizing enterocolitis	12%	**Pancreas,** *n* = **417**	
Congenital short gut syndrome	15%	Diabetes mellitus	
Pseudo-obstruction	6%	Type I	80%
Enteropathies	1%	Type II	4%
Other	32%	Other	16%

Fireman M, DiMartini A, Crone C, (2017). Psychopharmacology in Organ Transplantation. In: Ferrando and Levenson (Eds), *Psychopharmacology of the Medically Ill*, second edition, Arlington, VA: American Psychiatric Publishing, pp. 597–621,

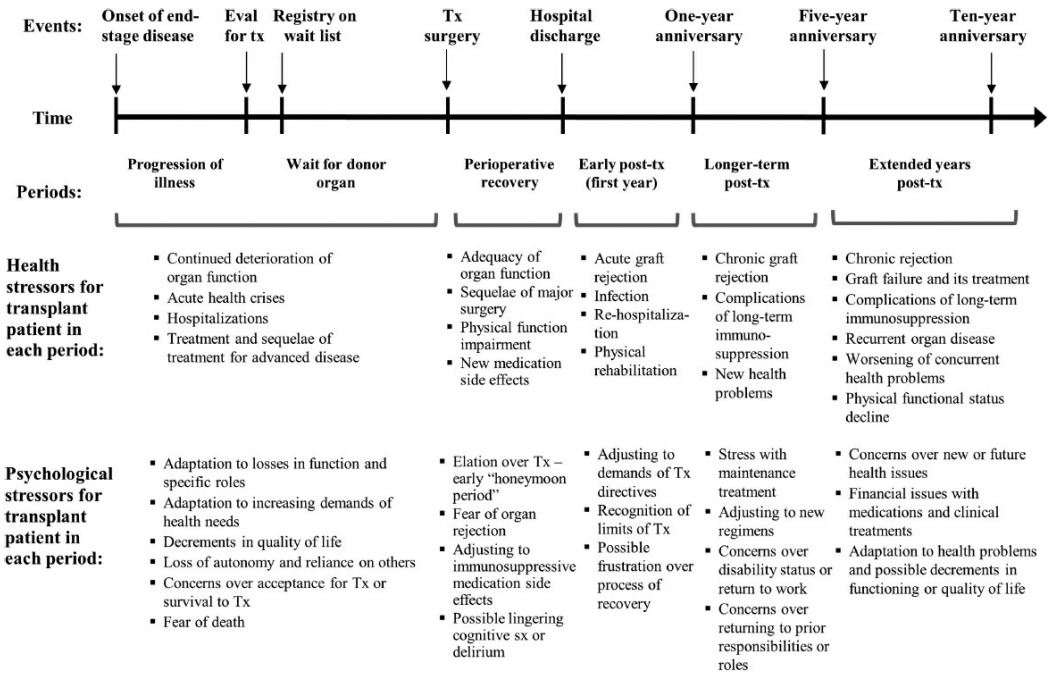

FIGURE 27.13–3. Organ transplant timeline: critical events, time periods for transplant patients. (From Dew MA, DiMartini AF, Kormos RL. Organ Transplantation, Stress of. In: Fink G, ed. *Encyclopedia of Stress*. Vol. 3. 2nd ed. Oxford, UK: Academic Press (Elsevier); 2007:33, with permission.)

Table 27.13–2.
Purposes and Goals of the Pretransplant Psychosocial Evaluation

1. Develop individualized treatment plans for patients including requirements for transplant listing.
2. Establish working relationships with patients.
3. Assess patients' coping strengths and weaknesses/vulnerabilities.
4. Diagnose and provide treatment options for active psychiatric disorders.
5. Assess past treatment adherence and the patients' ability to make lifestyle/behavior changes.
6. Evaluate amount, stability, and availability of social support systems.
7. Ensure the informed consent process for patients.
8. Educate and prepare patients for transplantation.
9. Provide input to transplant candidacy process.

RELATIVE IMPORTANCE OF PSYCHOSOCIAL CRITERIA. Although there is general consensus about the medical assessment and the criteria used for transplant candidacy, there are considerable differences among teams regarding the purpose of psychosocial evaluations, who performs these evaluations, the detail of the evaluation, and the use of psychosocial information obtained. This is perhaps due to the fact that transplant teams have differing opinions on the relevance and significance of psychosocial criteria. Additionally, there are no national psychosocial criteria for candidate listing in the United States although there are published guidelines on the basic components for the psychosocial evaluation (see Tables 27.13–2 and 27.13–3 for the goals and content of the psychosocial evaluation). Thus, each transplant program must decide on the pertinent psychosocial criteria and the relevance of these criteria for individual patients, often with input from mental health professionals. In general, of the organ types, heart transplant teams are more likely to view psychosocial issues as contraindications, while kidney programs are the most tolerant. The relative leniency of kidney programs is perhaps due to the larger experience (larger numbers of patients, longer periods of follow-up, and greater availability of kidneys) and thus possibly greater comfort with complex patients at these programs, as well as the cost benefits of transplantation compared to dialysis that tend to promote the surgery. Differences in approaches toward substance use and addictive disorders and whether a specified period of abstinence is required prior to evaluation or listing varies substantially and not all substances are considered equally across different types of organ transplant programs. The psychosocial evaluator should therefore be knowledgeable of their transplant team's approach to psychosocial issues so that the needs of the patient and transplant team, including their mandatory requirements for transplant listing, are met.

UNIQUE ROLE OF THE PSYCHOSOCIAL EVALUATOR. Due to these differences, the mental health professional must be knowledgeable of their team's approach to psychosocial issues and may need to educate the team on the prognosis and treatment of psychiatric disorders. Thus, mental health professionals have dual functions in the transplant evaluation process, serving the needs of both the transplant team and the transplant patient. Balancing these two roles may be complicated, but the mental health professional can both advocate for the patient and communicate and interpret the transplant team's requirements back to the patient. Information gathered during the psychosocial evaluation is essential to the transplant team's complete understanding of the patient. The communication back to the patient provides education on specific transplant-related expectations, the team's requirements for listing the patient as a transplant candidate, and the importance of these requirements to the success of transplantation.

Assistance with an appropriate referral for recommended psychiatric or psychological treatment can be provided. Thus the mental health professional can facilitate the development of a long-term collaborative relationship necessary between the patient and the transplant team.

Table 27.13–3.
Content of a Comprehensive Pretransplant Psychosocial Evaluation

Medical history	Course of advanced organ disease (acute vs. chronic), prior hospitalizations, treatments, complications, level of functioning/disability, quality of life, current symptoms, understanding of illness
Attitude toward transplantation	Circumstances leading to referral for transplantation, understanding of reasons for referral, receptiveness to transplant, prior experience or knowledge about transplantation, concerns/fears/reservations, expectations about transplantation
Knowledge about transplantation	Patient's initial education about the pre-to-posttransplant process, additional research/reading of his or her own on transplant, understanding of the risks and benefits involved with transplant, understanding of the long-term commitment to ongoing medical follow-up, immunosuppressants, monitoring, and lifestyle/behavioral changes required
Treatment adherence	Past and current history of adherence to medication regimens, self-monitoring (e.g., checking blood sugar, blood pressure), diagnostic testing, other treatment interventions (e.g., operations, supplemental oxygen, pulmonary rehabilitation), diet, fluid restrictions, exercise, abstinence from alcohol/drugs/tobacco
Coping with illness	Expression of acceptance or denial about prognosis and treatment options, utilization of active vs. passive coping skills for dealing with his or her illness (e.g., self-education, adjusts activity level vs. avoidance, heavy reliance on spouse, etc.), expressions of distress or feeling overwhelmed with his or her illness, methods for managing these feelings
Past medical history	Prior illnesses, hospitalizations, operations, experiences with health care system
Psychiatric history	Past and current history of mood disorders, anxiety disorders, or other disorders, symptom severity, course of symptoms, chronicity of symptoms, experience and response to psychiatric treatment, hospitalizations, treatment adherence
Substance use/abuse	Drugs/alcohol/tobacco frequency of use, amount, duration, diagnosable disorder, level of impairment to health/work/relationships, legal issues, knowledge of and insight into his or her addiction, commitment and ability to remain abstinent including prior attempts to reduce usage or prior periods of abstinence, prior and number of attempts at formal rehabilitation/counseling, attendance at 12-step meetings, sponsors, willingness to get therapy when needed, supports for abstinence, family history of substance abuse
Current medications	Knowledge of current medications including dose and frequency, understanding the purpose of each medication
Family history	Current and past history of medical, psychiatric, and substance abuse disorders
Social history	Developmental history, childhood experiences, educational background, history of trauma, significant relationships, work experience, current living situation, financial status, dependents, concurrent life stressors (e.g., family dysfunction)
Support system	Availability, stability, and extent of support available to help care for patient before and after transplantation, including family members, friends, neighbors, religious organizations, and coworkers
Mental status examination	Level of consciousness, appearance, behavior, speech, presence of psychomotor agitation or retardation, mood, affect, thought process/content, insight into

organ disease/transplant, judgment, attention, concentration, memory, and executive functioning

SCREENING AND CONTENT OF THE PSYCHOSOCIAL EVALUATION. At most centers, psychosocial evaluations are performed predominantly by clinical social workers. Typically, this evaluation serves as a screening evaluation, with a more comprehensive psychiatric assessment requested for patients identified as having prior or current psychiatric histories or psychosocial issues requiring additional attention. This two-stage approach can be an efficient use of resources because the sheer number of candidates may preclude a comprehensive psychiatric evaluation of all patients and additionally, with thorough screening, may not be indicated.

A comprehensive pretransplant psychosocial evaluation contains all the areas of a typical comprehensive psychiatric assessment, with the addition of transplant specific areas (e.g., coping with illness, adherence, and other health or illness behaviors) (Tables 27.13–2 and 27.13–3). Standardized transplant specific assessment tools (i.e., Psychosocial Assessment of Candidates for Transplantation, Transplant Evaluation Rating Scale, Stanford Integrated Psychosocial Assessment for Transplantation, and Structured Interview for Renal Transplantation) can be used alongside the clinical interview to ensure that relevant topics are covered. These instruments should be completed by a trained clinician and provide useful organization of information pertinent transplantation. Non-transplant-specific instruments assessing psychiatric symptoms or cognitive status are sometimes used in the evaluation if the situation warrants. Collateral information from family members, partners, medical and pharmacy records, toxicology screening, attendance at clinic appointments, and current or past health care providers to determine adherence and treatment engagement is often essential to substantiate or clarify information obtained from patient interviews. This type of expanded medical record review is being facilitated by the increasing use of electronic medical records connecting both intra- and inter-medical system data.

SPECIAL AREAS OF TRANSPLANT ASSESSMENT

Social Supports. Assessment of the presence of an adequate, stable, and available support system to help the patient through transplantation is another area requiring careful evaluation. Patients often require assistance with activities of daily living, medications, dressing changes, and transportation both before and after transplantation. Particularly while awaiting transplant, some patients may be too ill to physically care for themselves. Others may develop cognitive impairments that require constant monitoring for safety. Caregivers often feel the need to present themselves to the transplant team as being a capable and concerned support system for the patient. However, the best evidence of adequate

support is consistent behavior demonstrating this ability. Thus, the patient's caregivers should be present during the evaluation process so that their knowledge, ability (cognitive and physical), and willingness to take on these responsibilities can be assessed.

Treatment Adherence. Nonadherence to the posttransplant regimen and immunosuppressive medications in particular has been associated with poor clinical outcomes, contributing substantially to late acute rejection and graft loss. For these reasons, transplant teams are highly invested in a patient's ability to adhere to the long-term medical, behavioral, and lifestyle requirements necessary to optimize outcomes and survival. Following transplantation, patients must: take medications on a daily basis; monitor critical parameters such as vital signs, blood sugar, and pulmonary functions; come to clinic appointments; undergo diagnostic tests; follow dietary restrictions; participate in physical rehabilitation or exercise programs; and adhere to restrictions on use of illicit drugs, tobacco, and often alcohol as well. Lifelong adherence to immunosuppressants is needed to reduce the risk of organ rejection, and other drugs such as antibiotics, antifungals, antihypertensive agents, and hypoglycemic agents are often required lifelong as well. Despite these requirements, a review of kidney recipients found a prevalence of 28 percent for self-reported nonadherence to immunosuppressants (Denhaerynck, 2005). Similarly, a comprehensive meta-analysis of adherence studies across all organ types identified a rate of 23 cases of immunosuppressant nonadherence per 100 transplant recipients per year (Dew, 2007). Rates of nonadherence to other posttransplant requirements vary with higher rates for home self-care behaviors (e.g., diet, exercise, home monitoring of vital signs), and lower rates of nonadherence in the areas of attending clinic appointments and use of restricted or banned substances. Identification of the reasons and risk factors for nonadherence is one goal of the pretransplant evaluation. Pretransplant factors such as a history of nonadherence, substance use history, or poor social supports, appear to be most consistently linked to nonadherence posttransplantation. In addition to commonly assumed behavioral reasons for poor adherence (e.g., poor organizational skills, failure to accept personal responsibility for one's health, etc.), other social issues such as type of insurance coverage or transportation availability or even ambivalence about treatment should be considered. These factors and their impact on current adherence should be assessed during the psychosocial evaluation. Patient education on the importance of adherence and plans for pretransplant monitoring may be required. Additionally, a posttransplant plan for monitoring and education should be considered because patient adherence often worsens with time since transplantation.

Substance Use. Substance use and abuse is an area of special concern for

transplant teams, especially in liver transplantation where the use of drugs or alcohol often leads to the need for transplantation. Inquiries about substance use or abuse should be made with all patients, regardless of the primary cause of their advanced organ disease, and transplant programs may perform toxicology screening on all candidates at evaluation. Substance abuse may not be assessed by other transplant team members for patients whose organ failure was not directly linked to alcohol or drugs, and the mental health professional may be relied on to address this area (see Table 27.13–3 for areas of substance use assessment). In the United States, cannabis use has increased with the passing of medical prescribing laws and legalization in certain states. While smoked marijuana has known health risks, there is a lack of agreement over the recreational use of marijuana among transplant teams, although most cardiothoracic and some abdominal transplant programs exclude patients who are actively smoking marijuana. In addition to drugs and alcohol, the cessation of tobacco use (both smoked and smokeless) has been a requirement for heart and lung transplant listing and is increasingly seen as a contraindication to listing for abdominal transplantation as well (see tobacco section below). Some programs may not allow listing while a patient is using nicotine replacement therapy for smoking cessation, so the specific allowances of the transplant team must be known. Transplant programs attempt to select patients deemed unlikely to relapse posttransplant due to concerns that substance use will diminish patient or graft survival (e.g., patients and transplanted organs at risk from direct toxicity of the substance, increased risk of infection due to risky behaviors, or increased morbidities) or that substance use relapse would lead patients to become nonadherent (see substance use outcomes below).

Informed Consent in Terminally Ill Transplant Candidates. The importance of informed consent for organ transplantation cannot be overemphasized given the need for patient commitment to long-term responsibilities along with the potential serious risks and complications that can arise in both the early- and late-term following transplantation. Psychosocial evaluations help to address how well patients understand their illness, prognosis, treatment options, and the transplant process itself, thereby facilitating the informed consent process. However, evaluating patients' understanding of their options may be problematic under the circumstances in which transplant patients are evaluated. Often, patients with advanced organ disease referred for transplant evaluation are facing life-or-death decisions, with transplantation being the only means to extend their survival. For other patients, alternative treatments such as hemodialysis have become too difficult to tolerate or their quality of life has become unacceptable. Whether their circumstances are truly desperate or only perceived to be so, patients may have limited ability to fully appreciate, comprehend, and

consider all aspects of transplantation. Their illness may have resulted in acute or chronic cognitive impairments that also reduce their ability to comprehend and make informed decisions (see additional discussion of impairments below). Given these factors, a careful psychosocial assessment with attention to informed consent should be performed, balanced with realistic expectations about a patient's behavior under such emotionally difficult conditions. The evaluation can identify areas in which the provision of additional information on risks versus benefits of transplantation is required before an informed decision can truly be made by the patient.

Emergency Evaluations. Requests for emergency psychosocial transplant evaluations may occur for patients with fulminant liver failure or acute ischemic or idiopathic cardiomyopathy who have not been previously evaluated by the transplant team. These patients may be stuporous, on mechanical ventilation, or be in primary or drug-induced coma. Under these circumstances, the majority if not all of the pretransplant evaluation information gathered will depend on thorough interviews with family, friends, and health care providers, and attempts to obtain previous medical and psychiatric records. The informed consent process remains a concern, but the focus will be on the ability of the family or surrogate decision maker to understand the patient's medical condition, the transplant process, and maintain respect for the patient's wishes. Establishing whether a stable support system exists for these patients is especially important as both patient and family may not have had the opportunity to prepare themselves for transplant or the posttransplant regimen.

Patients with fulminant hepatic failure due to acetaminophen-induced liver injury require special evaluation for psychopathology. Determining whether the ingestion was a mistake made by a patient attempting to treat pain or cold symptoms or an intentional act of suicide will lead to different recommendations and care plans. As with any suicide attempt, the nature of the attempt (e.g., impulsive vs. planned), current stressors, presence of character pathology, and risk factors for subsequent suicide attempts or gestures need to be assessed. Whether specific patient characteristics could interfere with the patient's ability to develop a long-term collaborative relationship with the transplant team and consistently follow medical directives must be considered.

Living Donors. As noted earlier, organ transplantation increasingly often relies on living donation, particularly in kidney transplantation. Donor safety is of utmost concern, given the altruistic nature of living donation and the fact that donors have no medical need for surgery. As donation has expanded beyond biologically or emotionally close relatives of transplant patients to donors with little if any pre-existing relationship to the recipient, the medical and psychosocial evaluation of prospective donors

has become more extensive. Currently, the evaluation process at most US centers is specified by OPTN policy. The evaluation is as detailed and exhaustive in prospective donors as in prospective transplant candidates (Table 27.13–4). Mental health professionals must assess the psychiatric history of prospective donors and determine their current level of functioning and their cognitive capacity to make an informed decision about donating. Consensus opinions and both OPTN and Centers for Medicare and Medicaid Services policies also require that all potential living donors be provided with an independent living donor advocate (ILDA) who is not responsible for the transplant candidate's care or the decision to list the candidate for a transplant. The ILDA also evaluates the donor, assists them in gaining additional education about possible donation, and advocates on their behalf.

Table 27.13–4.
Areas of Assessment for Living Donor Evaluation

Mental health history and current status	Donor's psychological health and history, including psychiatric disorders, personality disorders, pain syndromes, prior psychological trauma/abuse, substance use
Motives	Motivations including those potentially related to urgency of the medical condition of recipient, lack of other donors, guilt, psychological expectations (e.g., desire to be hero, special, etc.), expectations about postdonation relationship with recipient. Exploration of ambivalence.
Relationships	Relationship between donor and recipient; attitudes of significant others and family toward the donation; availability of social supports; evidence of coercion/inducement
Knowledge about surgery	Knowledge related to the donation and transplantation procedures; risks to self; risks and potential benefits to recipient; expectations about donation/recovery; understanding of need for medical follow-up with transplant program after donation
Financial resources	Donor and family resources; donor's understanding of financial implications of donation
Psychosocial factors	Work- and/or school-related issues (if applicable); timing of the donation relative to work, school, or other important life events
Infectious disease status	Presence of any behaviors that may increase risk for disease transmission to the organ recipient
Need for interventions	Need for any educational or therapeutic interventions before a decision can be made regarding suitability to donate

Under most circumstances, current active psychiatric disorders or substance abuse or dependence are contraindications to donation. However, exceptions may arise under certain circumstances, such as a parent donating to a critically ill child. Therefore, psychiatric management of prospective donors and follow-up to continue psychiatric treatment after donation may be required.

Outcomes following living donation are given special considerations later in this section. The case study presented here highlights some of the complexities in donor evaluations.

A 30-year-old woman seeks to be evaluated as a living related kidney donor for her older brother, who has advanced renal failure secondary to diabetes mellitus. She states that she first learned about her brother's need for transplantation 6 months ago when she was contacted by their mother. Before this, she had been estranged from the family for several years due to their disapproval of her lifestyle, which included drug abuse and unstable relationships with various men. She reports that she has tried to re-establish ties with her family before, but they have not been receptive to her efforts even when she has tried to show them she is leading a more stable lifestyle. When asked about her motivation to donate to a brother who she has not spoken to for years, she emphasizes that no one is pressuring her to donate. However, she admits feeling that her brother has been more responsible than she has been and a "model son" when it came to helping their parents through difficult financial times. She expresses hope that donating a kidney will help her family to see that she is now acting as a responsible adult and family member. In this case the evaluator should be concerned over her motivations to donate. Issues to explore would include her expectations of her relationship with her family and especially her brother following donation, feelings of guilt, a desire to do something meaningful with her life, and a desire to restore her status in the family.

COMMUNICATION REGARDING TRANSPLANT CANDIDATE EVALUATION RESULTS.

Communication about patient suitability for transplantation and recommendations to the team should be plainly stated with minimal psychiatric jargon, providing straightforward guidance. However, caution must be taken regarding comments on a patient's suitability for transplant, partly because there is lack of consensus on what psychosocial issues would make a patient unacceptable for transplant, and because third-party payors may review the psychosocial assessment as part of their process to determine whether to financially approve a patient for transplant coverage. In some cases, health insurers may use information from the psychosocial evaluation to set forth more restrictive psychosocial requirements for the patient to fulfill before they will agree to cover transplantation costs. Fortunately, most transplant teams do not reject candidates outright but allow an opportunity for psychosocial issues to be addressed or resolved, followed by a re-evaluation for candidacy. Areas of preparation that are of particular importance to transplantation are addressed below.

Preparation for Transplantation. In addition to addressing the transplant team's concerns about psychiatric or psychosocial issues and aiding in candidate eligibility determinations, the pretransplant psychosocial evaluation can be used to develop an individualized treatment plan to educate and prepare candidates for transplantation and optimize posttransplant outcomes (Table 27.13–2). For example, the concerns, fears, or even ambivalence about the surgery that patients express during the evaluation can be useful to allow the evaluator to target appropriate interventions. Optimally, candidates should be referred for any

psychotherapeutic interventions early in their illness before medical or cognitive deterioration (e.g., encephalopathy) that will limit their ability to participate in or benefit from such treatment.

A 54-year-old woman with familial cardiomyopathy is referred for cardiac transplantation. Despite agreeing with her cardiologist about his referral, she fails to contact the transplant center until he reminds her during a follow-up appointment. When she begins her evaluation, she cancels some of the appointments and tests she is asked to complete. The transplant coordinator asks about these cancellations, and the patient states they were due to scheduling difficulties at her job. When seen by the team's social worker, she admits to being hesitant about transplant due to her cousin's experience several years prior. Her cousin developed recurrent episodes of rejection and died 5 years after her transplant. The patient is afraid the same fate awaits her and only agreed to an evaluation upon her family's insistence. Once her fears are uncovered, the transplant team assures her that she would not be pressured to be listed for transplant. The coordinator meets with the patient and her family to better educate them about transplantation. During this meeting the patient is able to gather information about her specific situation, the risks of surgery, and life as a transplant recipient. The patient also meets with the team's psychiatrist for several sessions to address her fears in greater depth and to reinforce her coping skills. Her anxiety lessens over time and she eventually decides she wants to pursue transplantation.

BEHAVIORAL AND CONTINGENCY CONTRACTING. For patients with significant psychosocial risk factors such as substance abuse, treatment nonadherence, or other psychiatric disorders requiring ongoing treatment, transplant teams may decide to use behavioral contracts to stipulate requirements for listing or monitoring during the waiting period for transplantation. These contracts can positively encourage behavioral change, minimize miscommunication, foster collaboration between patient and team, and reinforce the need for patients to take active responsibility for their health care. Although not a legal document, behavioral contracts clearly define what is expected of patients, how their adherence with the contract will be assessed, and what consequences they will face if they cooperate with or fail to cooperate with listed expectations. Successful adherence to these requirements would result in listing for transplantation; nonadherence would lead to removal from the waiting list. For example, patients with psychiatric disorders may be required to receive ongoing outpatient psychiatric care and be periodically reassessed by the team to determine whether the psychiatric disorder is under adequate control. Patients with a history of treatment nonadherence may be asked to demonstrate adherence to various aspects of their care (e.g., medications, appointments, tests, dialysis, dietary or fluid restrictions) for a period of time prior to listing and continue to be monitored during the waiting period leading up to transplantation.

A 58-year-old man is referred for cardiac transplantation due to a long history of ischemic cardiomyopathy. He suffered his first myocardial infarction in his mid-30s but was able to return to an active lifestyle after several months. He followed up with a cardiologist for a few years but eventually stopped as he was feeling well, had stopped smoking, and was exercising

regularly and watching his diet. Despite his efforts, he suffered a second infarction in his early 40s and a third at age 50. Since his last infarction he has developed steadily worsening symptoms of congestive heart failure. He has followed up with a cardiologist regularly and is on medications, but he argues with his physician whenever a new medication is added or a dose increase is advised. He has occasionally tapered the doses of some of his medicines to see if he could get by with less. He complains that each medication adds new side effects and he hates being a "walking pharmacy." He is very interested in transplantation as he feels his quality of life is poor, but he knows he will have to be on a large number of medications posttransplant. In preparation for transplant listing he is required to demonstrate consistent medication compliance and begin psychotherapy. In counseling he is able to identify previously unresolved issues about his vulnerability, lack of control over his condition, and anger toward health care providers as they remind him of his debility. He is able to address these issues and develops a realistic approach to his health and health care.

SUBSTANCE USE–DISORDERED PATIENT. Requirements for length of sobriety and addiction rehabilitation vary greatly from program to program and among different organ types even within a transplant program. Although stable abstinence is measured in years, not months, many liver transplant programs—for example—focus on the length of abstinence as the main criterion, with 6 months of abstinence being a benchmark for listing. Except for short abstinence, no single risk factor strongly predicts future use although poor social support, family history of addiction, polysubstance use, and failed rehabilitation have been associated with increased risk. Addiction counseling remains the best treatment option for providing education about the illness and strategies for monitoring and avoiding relapse, lifestyle, and behavioral changes, as well as for identifying appropriate supports to maintain abstinence. For those with short abstinence (less than 1 year) and no prior therapy, addiction counseling is highly recommended as a prerequisite to transplant listing. Patients with some period of abstinence but no addiction rehabilitation may not perceive a need for addiction treatment. Many have stopped actively using due to being ill but have not incorporated the skills or strategies to stay sober and prepare for continued abstinence posttransplant. Motivational interviewing may be needed to assist those reluctant to engage in treatment. However, for some, outright mandating of addiction treatment as a requirement for transplant listing may be needed. The intensity of addiction treatment may need to be tailored or adjusted over time to the medical debility of the patient. A recent controversy over liver transplantation for acute alcoholic hepatitis, where there is an expected rapidly deteriorating course and typically very short abstinence, if any, has led some programs to consider and transplant such patients. Mental health professionals should be prepared to evaluate and provide treatment recommendations for short-abstinence patients, keeping in mind that limited time and potential encephalopathy and/or physical debility could impede engagement in treatment and shorten the time frame to make listing decisions. As mentioned above, some centers use behavioral contracts stipulating the frequency and duration of addiction counseling. Others require random

toxicology screening to demonstrate that substances are not being used. Even after wait listing, ongoing periodic reassessment and monitoring are recommended because of the continued risk for relapse.

Transplant programs may also require patients on methadone maintenance therapy (MMT) to discontinue methadone either as a prerequisite for transplant listing or in the immediate postoperative period. However, the process of transplantation is very stressful for patients and is not an optimal time to attempt the additional challenge of a methadone taper. Although progressive medical illness may necessitate a decrease in the methadone dose, requiring a patient to discontinue MMT specifically for transplant listing could be considered unethical because a premature taper off methadone may precipitate a relapse, which would then prohibit the patient from transplantation.

REASSESSMENT OF WAIT-LISTED CANDIDATES. Periodic psychosocial follow-up with wait-listed candidates, either in the hospital or outpatient setting, is desirable because the candidate's status may change with respect to affective symptoms, stress, coping, cognitive functioning, adherence, substance use, social support, and financial issues. In some cases (i.e., resumption of substance use) reassessment may lead a candidate to be removed from the transplant list. More often, however, ongoing pretransplant reassessment not only provides continuity of care but can also improve the team's understanding of the candidate, help identify problems that need to be resolved pretransplant, foster a collaborative relationship with the team, and enhance adherence. Some candidates may not return until the point of transplantation, and ongoing reassessment at the transplant program may not be geographically or financially feasible for the candidate and family. Appropriate psychiatric services should be identified for those at a distance from the transplant program if issues arise, and information on the treatment or resolution of the problem should be conveyed back to the transplant team. Psychiatric clinicians and other physicians in the community, who may be providing services to transplant patients but are not familiar with transplantation, may benefit from a discussion with the psychosocial evaluator regarding the transplant team requirements for treatment goals, or from the receipt of available written educational materials on transplantation.

Patient Adaptation through the Transplant Time Course

EVALUATION AND WAITING PERIOD. Patients face significant adaptive challenges as they proceed through the pre- to-posttransplant course. Successful completion of the transplant evaluation and the individualized requirements to be accepted for listing on the national organ transplant candidate list is only the first hurdle. Many patients and their families perceive the waiting period to be even more psychologically stressful. This

is especially the case if patients are newly adapting to their diagnosis of terminal advanced organ disease and considering for the first time the transplant surgery and a future life as a transplant recipient. The patients' level of adjustment at the point of transplant evaluation may depend on how sick they are, how long they have lived with their illness (acute vs. chronic), and how long they have been anticipating (either positively or negatively) the issue of eventually requiring organ replacement. With chronic illness there can be progressive debility, with loss of vitality and physical and social functioning, changes in or loss of family or social roles, loss of independence, and financial hardships. Adapting to these changes can elicit anxiety, depression, avoidance, and denial and may require working through the grief of these losses. Unless candidates have a living donor and can therefore choose the timing of their surgery, they must endure the uncertainty of the wait time, the extent of their medical deterioration, and loss of functioning before the transplant surgery occurs. Additionally, wait-listed candidates may develop medical contraindications to transplantation (i.e., infection, serious stroke, progressive organ dysfunction), and both patient and family should be made aware that their eligibility might change for many reasons over time. These inherent stresses in the transplant process will tax the character style, coping, and social support resources of patients and caregivers, and these stresses can increase over time. Psychological counseling for patients and families can aid in successfully negotiating these transitions and prepare them for either transplantation or death.

PERIOPERATIVE PERIOD. For patients fortunate enough to survive the waiting period and undergo transplantation, the immediate perioperative period is a time of recovery and adjustment to new challenges. In some cases, recovery will be rapid and uncomplicated, while in others, it may be prolonged and marked with multiple complications. Except for renal transplantation, patients begin their postoperative recovery in the intensive care unit. Common postoperative complications such as bleeding, infections, rejection episodes, immunosuppressive medication side effects, and perfusion injury to the donor organ delay recovery may be perceived as setbacks by the recipient, and can lead to growing discouragement and frustration. Assisting patients through this period typically involves a combination of adequate pain control, education, reassurance, and careful monitoring for early psychiatric and psychological complications. Delirium is common during this period (see cognitive disorders and neuropsychiatric side effects of immunosuppressant sections below). Without complications a typical posttransplant length of hospital stay can be as short as 3 to 4 days for kidney recipients, 7 to 10 days for liver and heart recipients, and approximately 2 weeks for lung recipients.

EARLY POSTTRANSPLANT ADAPTATION. The immediate postdischarge phase places tremendous burdens on recipients and their families and caregivers now responsible for following complex medical regimens posttransplantation. In addition to new combinations of multiple medications, patients and caregivers must attend clinic appointments and get laboratory tests frequently, sometimes several times a week for at least the first several months after hospital discharge. Complications that arise may necessitate additional procedures and tests at the hospital or emergency room. At home, patients and caregivers must monitor vital signs, follow specific diets, and adhere to restrictions on physical activity (e.g., lifting weight, driving) and on use of alcohol, drugs, and tobacco. As recovery progresses, the regimen may become simpler and clinic visits and laboratory tests less frequent, and some restrictions may be removed. Eventually recipients are encouraged to return to physical activity and even employment if indicated.

TRANSPLANT-RELATED ILLNESS BEHAVIOR. Maladaptive illness behavior can interfere with successful psychological adaptation to the demands of transplantation. Experience working with adult transplant recipients suggests two major types of such illness behavior. First, for patients with a long history of living with chronic disease, transplantation may represent a transition from one state of illness to another, with the transplant process being a familiar part of life as a patient. However, some chronically ill patients may psychologically adapt to the sick role, developing coping strategies that perpetuate their dependency on being ill. These patients can have difficulty adjusting to or transitioning into a healthier state of being. They may feel that the transplant team is expecting too fast a recovery or pressuring them to get better. They may develop unexplained somatic complaints, such as chronic pain, or become nonadherent with posttransplant care recommendations. They may experience concern, anxiety, or fear over being seen by the transplant team less frequently as their postoperative course improves. The second pattern of maladaptive behavior can be observed in patients who become acutely ill and undergo emergency transplantation (e.g., due to fulminant hepatic failure or acute drug toxicity). For these individuals, with only the briefest period of pretransplant infirmity, the transplant can be an unwelcomed event. Awareness of their organ failure, their life-and-death circumstances, and their need for a donor organ may not occur until the perioperative period. These patients may feel a heightened sense of vulnerability and may have difficulty accepting or be in denial over their medical situation. They may attempt to return to normal functioning as quickly as possible posttransplant, recovering faster than the transplant team expects, yet possibly suffering later from pushing themselves too fast (i.e., returning to work before they are physically ready). They may resent having received a

transplant with all the inherent restrictions, lifestyle changes, and regimens, and they may act out their anger or denial in episodes of nonadherence.

LONG-TERM OUTCOMES. Long-term physical functioning depends in part on recipients' initial state of health, postoperative complications, other medical comorbidities, the degree of adherence to medical directives, and their motivation to stay fit. Most transplant recipients experience a multiyear period following the first anniversary of the transplant in which they achieve maximal functioning of the transplanted organ, with a relatively low level and severity of complications and infrequent episodes of acute organ rejection. Their relatively good health during this period appears to contribute to improved emotional well-being. With time, however, other problems with the organ will develop (e.g., chronic graft rejection) and patients will develop long-term immunosuppressive medication complications including renal insufficiency or failure, diabetes, hypertension, and cancers. In fact 10 to 20 percent of nonrenal transplant recipients will experience renal insufficiency or failure in the five years following transplantation. Additionally, some patients will have recurrence of or ongoing organ damage from their primary disease that can eventually destroy their transplant. Overall, transplantation extends life and improves health and quality of life compared to pretransplant functioning, but for most transplant recipients their health and quality of life remains somewhat lower than that of a normal healthy person (see Figs. 27.13–2 and 27.13–3 on survival and long-term health issues posttransplant).

PSYCHIATRIC DISORDERS AFFECTING ORGAN TRANSPLANTATION

Similar to other populations with chronic physical illnesses, organ transplant candidates and recipients are at elevated risk for elevated psychiatric symptom levels and diagnosable psychiatric disorders. Mood- and anxiety-related disorders are the most common psychiatric illnesses observed both pre- and posttransplant, although acute cognitive impairment and delirium are often seen shortly pretransplant and perioperatively due to metabolic perturbations and the effects of surgery. Some residual cognitive impairments may remain in the long-term posttransplant, although whether or not transplant recipients are at increased risk for dementias is unknown. In subpopulations of transplant recipients with histories of substance abuse or dependence (e.g., patients with alcoholic liver disease), risk for relapse remains a concern both pre- and posttransplant. Psychosis in transplant populations, although rare, is a significant concern because of its potential impact on health and well-being if not carefully managed. There is mounting evidence that each of these classes of psychiatric disorders can affect patient health outcomes after organ transplantation.

Mood Disorders

The most common mood disorders observed in transplant populations are major depressive disorder, dysthymia, and adjustment disorders. Up to 20 percent of kidney recipients, 30 percent of liver recipients, and 63 percent of heart recipients experience these mood disorders during the first several years after transplant. Elevated levels of subclinical depressive symptomatology have also been noted in most studies. The risk for depression appears to be greatest during the first year posttransplant. Indeed, rates of mood disorders have been observed to be higher during the first year posttransplant than during either the waiting period for transplant or during patients' lifetimes before they became transplant candidates. The heightened risk during the first year after the transplant has been linked to stressors associated with physical recovery and rehabilitation; adaptation to immunosuppressive medications and a new medical regimen; and psychosocial and role changes as the transplant recipient transitions from critical illness to a less illness-focused lifestyle. After the first year, the likelihood of additional episodes of mood disorders appears to decline for a period of several years. Although there is little empirical evidence on prevalence of mood or other disorders beyond about 5 years posttransplant, anecdotal accounts suggest that the prevalence of these psychiatric disorders, depression in particular, may begin to rise in the face of the late-term transplant-related health changes noted earlier.

A growing number of studies indicate that the presence of either diagnosable mood disorders or elevated depressive symptomatology increases the risk of poor health outcomes in transplant candidates and recipients. In a recent meta-analysis of the transplant literature, Dew et al. demonstrated that depression, either pre- or posttransplant, has a strong negative impact on posttransplant survival, increasing the relative risk of mortality by 65 percent (2015). Risk does not appear to differ by type of transplant. Although morbidities are infrequently studied, depression also appears to increase the risk of death-censored graft loss. To date, however, no studies have examined the mechanism by which depression may lead to poor posttransplant health outcomes, although it is speculated that depression might result in poor health behaviors (e.g., nonadherence) and/or pathophysiologic abnormalities (e.g., neurovegetative symptoms or immune mechanisms) that could contribute to morbidities and mortality. Regardless of the mechanism, the potential for depression to at least predict morbidity and mortality, combined with the fact that depression contributes to poor quality of life and functional disability in both patient and nonpatient community populations, suggests the importance of timely treatment of this condition in transplant recipients. Screening for depression, especially in the early posttransplant phase, merits consideration.

Anxiety Disorders

Anxiety disorders appear to be less prevalent than mood disorders both before and after transplantation. They have been observed in up to one-third of transplant recipients, although there have been too few studies to specify prevalence rates of each type of anxiety disorder. Some anxiety disorders (e.g., panic disorder) appear to be more common in patients with advanced lung disease than patients with other types of advanced organ diseases, and these disorders may be more common after transplantation in these patients as well. Posttraumatic stress disorder (PTSD) has received increasing attention in transplant recipients and may account for a significant proportion of all cases of anxiety disorder, especially during the first year posttransplant. Factors contributing to PTSD include actual events surrounding the transplant experience (e.g., critical medical events during the waiting period, the surgery itself, and the perioperative recovery period), as well as patients' experiences of hallucinated events and delusions during delirium. As for mood disorders, the prevalence of anxiety disorders appears to decline beyond the first year posttransplant, although stressors associated with the development of medical complications in the very late years posttransplant may precipitate new episodes of anxiety disorders.

Anxiety disorders may have the capacity to predict health outcomes in transplant recipients, although research remains sparse and the findings are mixed. One study found that PTSD experienced during the first year after heart transplantation predicted mortality during the subsequent 2-year follow-up period. However, persistent non-PTSD anxiety symptoms were unrelated to either mortality or morbidities in the cohort. The mechanism by which PTSD predicted mortality is not clear, and whether treatment of anxiety disorders would improve transplant recipients' health outcomes is unknown. Moreover, other studies have found little impact of other anxiety disorders such as generalized anxiety disorder or panic disorder. However, the potential quality-of-life benefits of treating such conditions emphasize the importance of monitoring transplant recipients for the development of anxiety.

Psychotic Disorders and Mania

Psychosis, including conditions such as schizophrenia, schizoaffective disorder, mania, and psychotic depression, is rare in any population, including transplant recipients. Moreover, unlike the depressive and anxiety-related disorders discussed above, which occur in transplant recipients both with and without earlier lifetime histories of such disorders, psychosis emerging beyond the postoperative recovery period (and which can therefore not be ascribed exclusively to acute metabolic perturbations associated with surgery and marked medication changes) appears to occur

almost exclusively in individuals with pretransplant histories of such illness. In fact, individuals with a history of psychosis are at high risk for posttransplant recurrences. A survey found that 37 percent of such patients experienced manic or psychotic episodes in the first several years posttransplant (Coffman, 2002). Nevertheless, these patients can do well posttransplant if good social supports, adherence to medical and psychiatric directives, and good control of psychotic symptoms are established prior to transplant. Under these circumstances these patients do not have rates of nonadherence that differ from other transplant recipients. The potential for good management posttransplant, plus the very negative effects if psychoses are not well managed, make them important conditions to consider during the pretransplant assessment. Despite common beliefs, it is not clear that a prior psychiatric history makes a patient more susceptible to the potential of immunosuppressive medications, especially steroids, to produce or exacerbate psychotic symptoms or mania (i.e., "steroid psychosis"). The appropriate use of antipsychotic medication is usually adequate to prevent or manage these symptoms if they emerge, and care must be taken that these medications are not inadvertently discontinued in the early posttransplant phase.

Alcohol and Other Substance Use Disorders

Mental health professionals are mostly likely to be called on to evaluate patients for these disorders in the context of liver transplantation. Alcoholic liver disease and hepatitis C (most often incurred due to injected drug use) are the two most common indications for liver transplantation in the United States (Table 27.13–1). Alcohol use disorders, predominantly in liver transplant patients, have thus received the most attention with respect to candidacy issues, policy, and pre- and posttransplant outcomes. Although the use of other substances by liver as well as other solid organ transplant patients is equally important, it appears less prevalent and has received much less consideration.

There is not a direct concordance between the amount of alcohol consumed, a diagnosis of alcohol abuse or dependence, and either the development of alcoholic cirrhosis or alcoholic cardiomyopathy. Evidence from the natural history of alcoholic cardiomyopathy and alcoholic liver disease suggests that the risk for these diseases increases in a dose-dependent fashion with 10 years or more of daily consumption. For alcoholic liver disease, the habitual intake of 20 g of ethanol per day (i.e., two standard drinks) for women and 80 g per day for men can cause cirrhosis.

The rates of problematic alcohol and other substance use are relatively low posttransplant, perhaps due to rigorous selection criteria. Over the long term following liver transplant for alcoholic liver disease (including projections out to 10 years posttransplant), although 30 to 60 percent may

drink (including those who have had only a single drink), some who drink will begin or return to addiction counseling, and only a smaller percentage, 10 to 25 percent, will continue to drink in a pattern potentially harmful to their health or graft functioning. However, since even a single drink can lead to a heavier or hazardous pattern of alcohol use, the identification of any use is reason for concern and an opportunity to provide education and referral for addiction counseling. At both US and European programs posttransplant alcohol consumption that results in premature death or graft failure is uncommon, and long-term survival of patients with alcoholic liver disease is equal to those transplanted for other causes, even among those who drink the heaviest. Nevertheless, among the heaviest drinkers, alcohol can contribute to medication nonadherence, which in combination with alcohol-related liver damage can lead to loss of graft function and death. Other nonalcohol substance use is uncommon. In a meta-analysis of studies reporting posttransplant substance use, the relapse to illicit drugs for transplant recipients with a pretransplant history of such drug use was 3.7 cases per 100 patients per year (PPY) (Dew, 2008). The rate was significantly lower in liver transplant recipients (1.9 cases per 100 PPY) compared to other organ types (6.1 cases per 100 PPY). These favorable results suggest that the most important task for transplant teams is to continue ongoing monitoring for substance use posttransplant, not to condone or dismiss any use, and to provide resources for psychiatric assessment, intervention, and treatment as soon as any substance use is discovered.

Characteristics that predict relapse to use in the general population also appear to predict posttransplant relapse to use, including a diagnosis of substance dependence compared to diagnosis of abuse, family history of addiction, failed rehabilitations, multiple substances of abuse, poor social support for abstinence, and short pretransplant length of abstinence. However, these risk factors have been largely studied in relation to alcohol relapse; risk factors for other substance use have been mostly unexplored.

A few studies have addressed outcomes in MMT transplant recipients. These studies show little evidence that MMT participants have problems with posttransplant relapse or noncompliance. Although postoperative pain management may be more complicated for these patients in the immediate posttransplant period, methadone can be used as their primary analgesic. Close coordination with and transition back to their outpatient MMT program is essential when the patient is medically and physically capable. In one study of MMT, liver recipients mostly were eventually returned to their original methadone dose, with only a few requiring a modest dose increase.

Nicotine Dependence: Smoked and Smokeless Tobacco

There is substantial evidence of the contribution of tobacco use to

posttransplant vascular complications, health problems, cancers, graft loss, and death. Transplant patients may be at specific risk of tobacco-induced cancers due to the effects of immunosuppression. For example, in one study, compared to nonimmunosuppressed individuals, alcoholic liver disease transplant recipients had a fourfold increased risk for lung cancer and a 25-fold increased risk for oropharyngeal cancer, risks presumed due to tobacco use (Jain, 2000). As noted earlier, all types of solid organ transplant programs are increasingly likely to suggest or require candidates to be tobacco free for transplant listing. Nevertheless, significant rates of tobacco use resumption occur after transplantation. Tobacco use posttransplant is particularly common in patients transplanted after histories of alcohol or illicit drug use, with a rate of 10 cases per 100 PPY, compared to the rate of 3 cases per 100 PPY in the general transplant population (Dew, 2008). Thus, posttransplant clinicians should be vigilant to monitoring for tobacco use and counseling for cessation.

Cognitive Disorders

Assessing cognitive function in transplant candidates is essential to identifying deficits that could compromise patients' ability to follow medical directives and capacity to comprehend and participate in decision making before surgery. Caregivers may need to assist the patient with following medical directives or even with basic activities of daily living if deficits are severe. Patients often experience cognitive impairments while awaiting transplantation due to the effects of advanced organ disease. For example, in heart failure patients, low cardiac output and hypoperfusion or reduced cerebral blood flow, impaired cerebrovascular reactivity, ischemia in the absence of cerebrovascular events, and even medications such as inotropic agents can contribute to cognitive impairments. Patients on ventricular assist devices (VADs) pretransplant may have microemboli to the brain, leading to temporary or permanent cognitive impairment (see VAD section below). Cognitive impairment is also relatively common in lung transplant candidates due to ineffective pulmonary gas exchange. Both elevated arterial partial pressure of carbon dioxide (PCO_2) and decreased partial pressure of oxygen (PO_2) contribute to these deficits (Parekh, 2005). Oxygen therapy may improve cognitive functioning in certain candidates, but whether hypercapnia or hypoxemia will lead to permanent brain damage is unclear.

Prior to transplantation, it is critical to differentiate between the fluctuating course of a delirium that is potentially reversible posttransplantation and more persistent cognitive deficits that may represent a pre-existing dementia or a static cognitive impairment. If possible, therapies to reverse or alleviate the metabolic or physiologic effects of organ disease should be implemented and re-examination

performed. Existing impairments can persist or worsen posttransplant. Even if reversible, posttransplant symptoms of cognitive impairment may take months to years to resolve.

Although organ dysfunction–related cognitive impairment may be reversed by transplantation, postoperative delirium or neuropsychiatric side effects of immunosuppressants may worsen cognitive symptoms. Delirium is the most common cognitive disorder associated with transplantation, and most if not all patients will experience delirium at some point either due to pretransplant medical conditions or postoperative recovery. Hepatic and uremic encephalopathies are two specific commonly encountered areas of pretransplant delirium, separately addressed in more detail below. For transplant recipients there are multiple factors in the perioperative period that may contribute to the development of delirium (Table 27.13–5). Postoperative medications, especially immunosuppressants, can also contribute to delirium (see section on neuropsychiatric side effects of immunosuppressants below). Patients with pre-existing delirium (e.g., hepatic or uremic encephalopathy) often have persisting symptoms of delirium postoperatively, and these patients may be more vulnerable to the deliriogenic effects of other factors. Patients with other pretransplant CNS insults may similarly be more vulnerable to delirium. In a prospective study of cognitive impairment in lung transplant patients, 57 percent showed neurocognitive impairment at discharge from the transplant hospitalization while only 45 percent were impaired prior to transplant (Smith, 2014). Although impairment associated with delirium tended to improve at follow-up 3 months later, 80 percent of those who were impaired pretransplant remained impaired at follow-up with older age and lower education associated with poorer neurocognitive performance. In a cohort of liver transplant patients, cognitive impairment characteristic of hepatic encephalopathy largely resolved after transplant (Tryc, 2014). However, new onset cognitive impairments emerged by 1 year posttransplant in 70 percent of recipients and were not related to a history of encephalopathy, primary liver disease, immunosuppression regimen, or prior alcohol use.

Table 27.13–5.
Perioperative Causes of Delirium

Anesthesia	
Hypoxia	Anemia, hypotension, pulmonary emboli, pulmonary insufficiency, ARDS
Infection	Aspiration pneumonia, peritonitis, pneumonia, sepsis, stasis/decubitus ulcer, wound infections, urosepsis
Neurologic events	Abscess, cerebrovascular accident, central pontine myelinolysis, intracerebral hemorrhage, meningitis, posterior leukoencephalopathy syndrome, seizures
Metabolic imbalance	Acidosis, alkalosis, hypo- and hypercalcemia, hyponatremia, hypo- and hyperglycemia, hypo- and hyperkalemia, hypo- and hypermagnesemia

Nutritional deficiency	Protein-calorie, vitamin (B_{12}, folate, niacin, thiamine)
Medications	
Immunosuppressants	Cyclosporine, mycophenolate mofetil, steroids, tacrolimus
Antibiotics	Aminoglycosides, cephalosporins, fluoroquinolones, sulfonamides
Antifungals	Itraconazole, voriconazole
Antivirals	Acyclovir, famciclovir, ganciclovir, valaciclovir
Cardiovascular	ACE inhibitors, β-blockers, calcium channel blockers, clonidine
Gastrointestinal	Metoclopramide, proton pump inhibitors, ranitidine
Narcotics	Morphine sulfate, dihydromorphinone, etc.
Respiratory	Albuterol, montelukast, pentamidine
Sedatives	Benzodiazepines, diphenhydramine, zaleplon, zolpidem
Drug withdrawal	Narcotic or benzodiazepine withdrawal

ARDS, acute respiratory distress syndrome; ACE, angiotensin-converting enzyme.

Hepatic Encephalopathy in Liver Disease. Hepatic encephalopathy (HE) is classified by the fourth revised edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR) as a delirium and it is commonly encountered in patients with significant liver disease. It is characterized by a constellation of signs and symptoms that can include alteration of consciousness (including stupor or coma), cognitive impairment, confusion or disorientation, affective or emotional dysregulation, psychosis, behavioral disturbances, bioregulatory disturbances, and disturbances of the motor system represented by asterixis, tremor, increased tendon reflexes, increased muscle tone, ataxic gait, bradykinesia, slurred speech, or incoordination. The identification of hepatic encephalopathy is important as it impacts patient quality and quantity of life. Minimal hepatic encephalopathy predicts overt hepatic encephalopathy, and fulminant hepatic encephalopathy is associated with intracranial hypertension, cerebral edema, seizures, and death pretransplant. Persistent hepatic encephalopathy is rare and is observed predominantly in patients with extensive portacaval collateral circulation or after surgical or transjugular portosystemic stent shunting procedures.

Hepatic encephalopathy reflects a continuum of symptoms, with subclinical or minimal hepatic encephalopathy lying at the minimal end of the spectrum. Even minimal hepatic encephalopathy is clinically important as it can affect up to 60 percent of cirrhotic patients, can impair patient safety (such as the ability to drive a motor vehicle), job functioning, and quality of life, and is associated with persistent cognitive deficits after liver transplantation. The earliest signs of hepatic encephalopathy are often sleep disturbances and subtle behavioral changes, which are not obvious on clinical examination but should be elicited from the patient and caregivers. Minimal hepatic encephalopathy is primarily associated with abnormalities in attention, memory, and response inhibition and may not be discernable on a typical clinical examination. Additional neuropsychological tests assessing psychomotor speed, praxis, concentration, and attention, such as

Trail-Making Tests A and B and the Digit Symbol and Block Design Tests from the Wechsler Adult Intelligence Scale (revised), are commonly used to identify minimal hepatic encephalopathy impairment. Although an EEG may show common abnormalities, such as generalized slowing of dominant rhythm, cognitive testing is more efficient and perhaps more sensitive to minimal hepatic encephalopathy.

Although hepatic encephalopathy most likely has a multifactor pathogenesis, the predominant treatment strategy is to reduce the production of ammonia from protein in the intestinal tract and reduce its absorption. However, a variety of other compounds and metabolites (e.g., mercaptans, false neurotransmitters, aromatic—phenylalanine, tyrosine, and tryptophan—and branched chain amino acids, manganese, endogenous benzodiazepines, and increased concentrations of CNS GABA) as well as processes such as edema, inflammation, and increased permeability of the blood–brain barrier have also been implicated in hepatic encephalopathy. Hepatic encephalopathy can be precipitated by gastrointestinal hemorrhage, uremia, use of some psychoactive medications or diuretics, significant protein intake, dehydration, or electrolyte imbalance, and these problems should be corrected first. Medications that can contribute to symptoms of hepatic encephalopathy or slow intestinal motility (anticholinergic drugs, tranquillizers, narcotics, and sedatives) should be avoided.

Uremic Encephalopathy in Renal Disease. Chronic renal failure frequently causes multiple pathophysiologic changes involving catabolic, metabolic, and endocrinologic processes that contribute to uremic encephalopathy. In uremia the accumulation of organic substances that act as neurotoxins (e.g., urea, uric acid, guanidine compounds, hippuric acid, various amino acids, indoleacetate, and others) are believed to be related to the encephalopathy, but no single metabolite has been identified as the sole cause. Other pathophysiologic changes also implicated in uremic encephalopathy include hormonal elevations, electrolyte imbalances (e.g., acidosis, hyponatremia, hyperkalemia, hypocalcemia, hypermagnesemia), anemia, malnutrition, and CNS factors (increased calcium and decreased GABA and glycine activity). Reduced levels of thiamine, a water-soluble vitamin, have been noted in hemodialysis patients, and cases of Wernicke's encephalopathy have been mistaken for uremic encephalopathy.

The manifestations of uremic encephalopathy typically fluctuate and can begin insidiously, with patients reporting mild symptoms of irritability, insomnia, or cognitive, attentional, and memory impairments. Patients may complain of physical symptoms of slurred speech, muscle twitches, or restless legs. Symptoms can progress slowly or rapidly to confusion, lethargy, seizures, overt delirium, psychosis, catatonia, and stupor or coma. An EEG can aid in the diagnosis (typically showing generalized slowing of

the dominant rhythm) and can evaluate for seizures or even nonconvulsive status epilepticus, which can occur in uremia and could be mistaken for uremic encephalopathy.

Uremic encephalopathy has also been associated with a cliniconeuroradiologic syndrome termed posterior reversible (leuko) encephalopathy syndrome (PRES) (see section on PRES in the neurotoxicity of immunosuppressive medications below). Characteristic radiographic findings on CT or MRI are seen in the posterior cortical and subcortical white matter.

Removal of uremic toxins by hemodialysis, correction of electrolyte imbalances and anemia, and treatment of malnutrition can diminish the symptoms of encephalopathy and improve cognition. Seizures may be treated with anticonvulsants. Drugs that are sedating or require elimination of the parent compound or active metabolite by renal clearance should be avoided or dosages adjusted to account for the degree of impaired renal functioning.

Dialysis encephalopathy (also called dialysis dementia) is a progressively deteriorating fatal process seen in long-term hemodialysis patients. Its etiology is controversial but speculated to be attributed to aluminum toxicity from the water used in the dialysate. Due to the elimination of aluminum in the dialysate and the discontinuation of aluminum-containing phosphate binders for dialysis patients, cases of dialysis dementia are now rare.

SPECIAL ISSUES IN TRANSPLANTATION

Life-Sustaining Technologies and Their Effects

Dialysis. Most renal transplant candidates are on maintenance dialysis prior to transplantation, although about 2 percent receive kidney transplants as their first advanced renal disease therapy. Without a living donor, a US adult kidney candidate can expect a median wait time of 3 to 5 years. Thus survival on dialysis until transplant remains an important issue. According the US Renal Data System statistics, dialysis survival after 1 year is 80 percent and only 35 percent after 5 years, but these rates are dependent on age, primary disease, and medical debility. Although there is no specific age limit for renal transplantation, advanced age in combination with medical comorbidity make transplantation less of an option. Cardiovascular diseases are common reasons why patients are ineligible for transplantation and are also common causes of death on dialysis.

In addition to medical criteria, candidates should demonstrate good adherence to dialysis and follow-up with their nephrologist. Pretransplant adherence to dialysis is a useful indicator of posttransplant adherence. Patient must not only comply with dialysis sessions but also with dietary and fluid restrictions as well. Attendance at dialysis and missed or shortened sessions should be evaluated. If nonadherence is identified, the reasons should be explored and addressed. Comorbid depression can affect adherence to dialysis and, if treated, can improve adherence. Uremic encephalopathy can impair functioning and quality of life during the wait period (see section on uremic encephalopathy above). Following transplantation, recipients often report good overall quality of life, with distinct improvements in physical, mental, and social functioning compared to pretransplantation.

VADs in Heart Transplantation. Progress in the development of VADs from external or paracorporeal devices to implantable devices has dramatically improved both the physical and psychological health of patients with advanced heart failure. These devices are used primarily as a bridge to transplantation, although they are now offered as “destination” therapy for some patients ineligible for transplant. The newest VADs now include implantable left ventricular or biventricular versions that have improved mobility, easy of wearability, and quality of life, and routinely allow discharge from the hospital. Portable pneumatic drivers are compact and lightweight and can be worn on a shoulder strap or towed on a luggage-type carrier. Patients must be able to comply with postoperative care of the device and drive line site, as well as anticoagulation or other medical regimens. Following discharge, patients can engage in light-to-moderate physical activity (including walking, driving, dancing, and even

work). Most patients can achieve New York Heart Association functional status I or II on a VAD. In addition, while supported on a VAD, patients can achieve significant gains in physical and physiologic rehabilitation and rebuild muscle mass, potentially stabilizing their cardiac condition. With the patient in a less urgent status, the transplant team can wait for an optimal donor organ. These benefits notwithstanding, the logistics of arranging for a VAD patient to live outside of the hospital require that all who are involved in the patient's care be extensively trained, with a medical team who are available at all times as well as trained local emergency paramedical personnel available near the patient's home. Although this can cause significant patient, caregiver, and medical system burden, the outpatient arrangements can be economically advantageous compared to prolonged hospitalization.

Although VADs can improve quality of life and allow mobility and functioning for most activities of daily living, psychological and cognitive problems are not uncommon. Adjusting to the VAD can be psychologically difficult for patients, and incorporating the machinery into their bodies can evoke feelings of a damaged body image and sense of self, intense vulnerability, and fear of a VAD malfunction or infection. Psychotherapy to address these issues for both patients and caregivers may ease the transitions onto a VAD and help them prepare for eventual transplant. Patients bridged to transplantation with a VAD have similar and sometimes quicker physical recovery after transplant compared to those not requiring a VAD before transplant, and their posttransplant levels of emotional well-being are also similar. However, VAD patients have been found to have poorer residual cognitive functioning posttransplant (Dew, 2001). Cognitive impairments may in part be due to the higher risk of neurologic events (e.g., thromboembolism) while supported on the VAD, although continued improvements in VAD design have reduced the risk of such events.

Living Organ Donor Outcomes

Living liver and lung donations are much more surgically complex, invasive, and potentially more dangerous procedures than kidney donation. In fact, since laparoscopic donor nephrectomy has been increasingly used, this procedure has significantly improved kidney donor outcomes with less postoperative pain, shorter hospital stays (as little as 2 to 3 days), overall quicker recovery times, and more favorable cosmetic results. Although mortality rates have been less than 1 percent for both kidney and liver donors, at least one-third of liver donors have perioperative complications, with serious complications occurring in 14 percent of donors. Although no lung donor deaths have been reported, nearly 20 percent of lung donors will have perioperative complications and will lose 10 to 20 percent of their predonation lung functioning (Bowdish, 2004). Kidney donors typically

require 4 to 6 weeks of recovery before returning to work. Liver donors require 8 to 12 weeks of recovery or longer especially if the job involves heavy lifting.

Donor quality of life and emotional well-being do not appear, on average, to be adversely affected by organ donation. In fact, there is literature dating back to the mid-1960s showing that donors can accrue significant personal benefits from donation in terms of feelings of self-esteem and self-worth. Extremely low percentages of donors regret having donated or state that they would not donate if they had it to do over again. For the majority of donors, there is little evidence that mental health, physical health, or quality of life is worsened by donation. However, adverse outcomes remain a possibility and have been observed in donors. The development of mood and anxiety disorders has been noted, and several case studies have documented completed or attempted suicide (although the extent to which the suicides were attributable to donation has usually been unclear). There remains a dearth of long-term follow-up of donors many years postdonation, particularly in liver and in lung donors. In a multi-center follow-up study of kidney donors who averaged 17 years postdonation (ranging from 5 to 48 years), donors' health-related quality of life exceeded that of the general population, although 13 to 20 percent of donors showed quality of life impairments in one or more areas (Gross, 2013). A history of psychiatric difficulties increased these donors' risk of experiencing mental health problems postdonation.

Treatment Issues Specific to Transplantation

Psychotherapeutic Treatment Modalities. For the pretransplant period, several randomized controlled studies have examined a variety of cognitive–behavioral and education-based therapeutic interventions to address issues specific to this phase—uncertainty of wait or surviving to transplant, decline in health and increasing dependence on others, losses of roles and functioning, decrements in quality of life. Some have included caregivers as well. Results have been mixed but mostly demonstrate improvements in quality of life and reduction in psychological distress among candidates in the active treatment arms.

Following transplantation, individual psychotherapy appears beneficial in addressing three recurring psychological themes: (1) fear of organ rejection, (2) feelings of paradoxical loss posttransplant despite successful transplantation, and (3) psychological integration of the organ. Group therapy and other innovative strategies have also been employed to deal with specific issues of transplantation and to also address logistics and staffing resources. Group therapy is a frequently and successfully used technique for organ transplantation patients and family members. Educational groups are often used to provide information to transplant candidates about the transplant process, systems issues, and the

responsibilities of transplant recipients and caregivers. Depending on the available resources and number of patients near to the program location, support groups may also be comprised of pre- or posttransplant patients, patients or caregivers, and organ-specific or cross-organ type participants. Group therapy participants report decreases in negative affect, increases in positive affect and happiness, less illness intrusiveness, improved quality of life, and decreased need for contact from their transplant coordinators both in clinic and by telephone for social support. Additional support can be found through the Transplant Recipient International Organization (www.trioweb.org), a nonprofit international organization that provides support, education, and advocacy for transplant recipients worldwide. The member chapters usually have regular meetings of transplant recipients, and they provide group support for transplant candidates, recipients, and their families.

Using an innovative approach, the University of Toronto created a mentoring program for heart transplant candidates using peer mentorship by an already transplanted recipient to augment patient care by the staff. Mentors provided information and support from a peer perspective. Although patient satisfaction with the program was high, the coordinators felt early introduction of the mentor and individualizing mentors according to demographics and clinical course could improve the experience further. Gross et al. (2004) tested a multisession group mindfulness-based stress reduction approach and observed benefits to patients' psychological adjustment and quality of life. Dew and colleagues developed an Internet-based intervention to deal with the logistical problem of recipients living at a distance from the transplant program. The multicomponent intervention included stress and medical regimen management workshops on the website, monitored discussion groups, access to electronic communication with the transplant team, and information on transplant-related health issues. Compared to patients without access to the website, heart recipients who used the Internet intervention had significant decreases in depressive and anxiety symptoms, their quality of life in the social functioning domain significantly improved, and their caregivers had decreases in anxiety and hostility symptoms. Additionally, the subgroup using the website's medical regimen workshop showed significantly better adherence to multiple elements of the medical regimen.

Interventions for Adherence to the Posttransplant Regimen. As discussed earlier, lifelong adherence to the posttransplant regimen can be difficult. Many factors contribute to posttransplant nonadherence, including medication side effects, behavior problems (e.g., poor planning, poor health responsibility), and social issues (e.g., insurance coverage, copays, transportation). Eliciting patient-specific barriers will guide selection or development of the best approach for intervention. Patient involvement in

selection of adherence intervention and allowing them, with support, to make decisions about their care is important to success. Trials designed to minimize posttransplant nonadherence show no single intervention to be superior, and a combination of interventions might be the best approach. Re-education about the regimen and its rationale can be helpful and educational interventions delivered by pharmacists have been effective. Problem solving, cognitive–behavioral therapy, or other behavioral therapies may be useful. Patients traumatized by the transplant experience may have associated adherence issues that could be targeted with psychotherapy. Identification and treatment of depression and/or substance use may improve adherence.

Interventions have included a variety of approaches: case management/coaching and educational interventions, multicomponent interventions, patient-tailored interventions with a multidisciplinary approach and with widespread use technology, cognitive–educational interventions making use of the Internet and electronic feedback (including real-time feedback). Technology-based interventions hold promise for improving patient self-monitoring and adherence to posttransplant care, ideally increasing patients' involvement in their care. A small literature suggests that, following the intervention, booster or maintenance sessions may be required to maintain optimal adherence.

Psychopharmacologic Therapy. Pharmacotherapy is complicated during the pretransplant period due to the presence of advanced organ disease. Pharmacokinetic alterations caused by changes in perfusion, fluid status, protein binding, hepatic drug metabolism, and excretion are common in advanced organ disease and require careful selection of medications and dosage. Additionally, pretransplant patients are more susceptible to adverse side effects (e.g., encephalopathy, arrhythmia) and potential drug–drug interactions, as it is not uncommon for them to be on a number of medications prior to transplantation. Some psychotropic agents should be avoided in advanced organ disease, particularly when advanced cirrhosis is present. In general, pharmacotherapy in pretransplant patients calls for using lower starting dosages and slower titration of drugs until they reach clinical effectiveness.

Following transplantation, pharmacokinetic alterations caused by advanced organ disease usually resolve but considerations about potential drug–drug interactions arise. In particular, many of the immunosuppressive agents (tacrolimus [Prograf], cyclosporine [Neoral and others], sirolimus [Rapamune], mycophenolate mofetil [Cellcept, Myfortic]) are primarily metabolized by CYP 450 3A4; thus, psychotropic medications that strongly inhibit CYP 450 3A4 may pose a risk. Such psychotropic drugs include, in decreasing order of inhibition: fluvoxamine, nefazodone > fluoxetine > sertraline, tricyclic antidepressants, paroxetine >

venlafaxine. Although fluoxetine and tricyclic antidepressants have been used to treat depressed transplant recipients without significant changes in immunosuppressant levels, nefazodone has resulted in toxic immunosuppressant levels with serious complications. Although this suggests that the amount of CYP 450 3A4 inhibition may not create clinically meaningful drug interactions for all patients, these drugs should be used cautiously and with frequent monitoring of immunosuppressant levels. Most immunosuppressants have significant metabolic side effects (e.g., weight gain, hyperlipidemia, glucose intolerance), which must be considered with choosing psychotropics to prescribe. Table 27.13–6 summarizes primary considerations about psychotropic agents pre- and posttransplantation.

Due to its narrow therapeutic window, the use of lithium before and early after transplantation can be complicated due to fluctuations in fluid status (e.g., edema, dehydration) and renal function causing unstable lithium levels and potential toxicity. Additionally, lithium can cause renal tubular damage that might be problematic when combined with the nephrotoxic effects of immunosuppressant medications. If lithium is necessary, monitoring of lithium levels, fluid status, electrolytes, renal functioning, and creatinine clearance should be closely followed and lithium discontinued if renal insufficiency develops.

The use of antipsychotic medications may be necessary posttransplant to manage delirium, marked immunosuppressant-induced mood changes, or a pre-existing psychotic disorder. Symptomatic treatment of agitation, hallucinations, or paranoia usually calls for the use of atypical antipsychotic agents or traditional neuroleptics (e.g., haloperidol). Due to the risk of QT prolongation associated with these medications, along with the likely presence of other medications or conditions that have also been linked to torsades de pointes, electrocardiogram (ECG) monitoring is recommended.

Benzodiazepines are effective for anxiolysis before and after transplant, but care is needed when selecting these agents. Oversedation, increased cognitive impairment, and reactivation of a past substance use disorder are potential risks. Pretransplant, impaired hepatic oxidation can delay clearance of benzodiazepines and increase risk for adverse side effects. Use of agents that only require conjugation (i.e., lorazepam, oxazepam, temazepam) would be preferable. Buspirone can also be beneficial for anxiety and may also have respiratory stimulating properties especially beneficial for patients with lung disease or those attempting to wean from a ventilator.

Medications that may reduce cravings and potentially diminish the risk for relapse to alcohol (e.g., acamprosate, ondansetron, naltrexone) or opioids (naltrexone) have not been studied in transplant patients. Naltrexone has a small risk of hepatotoxicity (contained in a black box warning) and is not recommended for patients with active liver disease,

hepatitis, or liver failure, making it less suitable for liver transplant patients. At doses higher than recommended (more than 300 mg per day), naltrexone can be a direct hepatotoxin. Disulfiram blocks the oxidation of alcohol at the acetaldehyde stage and can create severe nausea, vomiting, and hemodynamic instability. It requires hepatic metabolism for conversion into an active drug. In addition, a metabolite of disulfiram is an inhibitor of CYP 3A4, and posttransplant may interfere with immunosuppressant metabolism.

For tobacco users, in addition to smoking cessation counseling and nicotine replacement therapies (gum, patches, lozenges, and inhaled), the use of medications (bupropion or varenicline) may aid in cessation. Bupropion should not be used in those with a history of a seizure disorder or electrolytes disturbances that could contribute to a seizure. Varenicline is excreted by the kidneys, and dose adjustment is required for patients with moderate to severe renal impairment. Both medications can increase the likelihood of success and should be continued for 6 to 8 weeks after tobacco cessation to prevent relapse.

Neurotoxicity of Immunosuppressant Medications

While transplant teams tend to focus on the neurotoxic effects of corticosteroids (e.g., delirium, agitation, insomnia, depression), the impact of these agents on transplant recipients is often transient and occurs early posttransplant when higher doses are utilized. Of greater significance are calcineurin inhibitors (CNI), which continue to serve as a mainstay of immunosuppressive medication regimens in solid organ transplantation. Tacrolimus and cyclosporine have similar neurotoxic side effect profiles with as many as 40 to 60 percent of transplant recipients experiencing mild neurologic symptoms such as tremulousness, headache, restlessness, agitation, insomnia and vivid dreams, paresthesias, and mood changes. Moderate-to-severe neurotoxic side effects (e.g., delirium, seizures, cortical blindness, catatonia, coma) occur less often but can reach 20 to 30 percent in the early postoperative period. Neurotoxicity is likely more common with tacrolimus, but can occur with either agent even at therapeutic drug levels. Liver transplant patients appear to be more susceptible to these neurotoxic side effects than heart and lung transplant recipients.

The actual mechanism of CNI neurotoxicity is unclear but possibly multifactorial. Investigations into the etiology of more serious neurotoxic side effects suggest associations with biochemical and electrolyte derangements including higher plasma levels, intravenous administration, hypocholesterolemia, and hypomagnesemia. However, these features are not present in every case. Neurotoxicity of CNIs may involve drug-induced vascular changes or direct action at the CNS neuronal level. Studies have revealed that CNIs increase sympathetic outflow and nerve activity, alter mitochondrial function, may be selectively toxic to white matter, and are

capable of disrupting the integrity of the blood–brain barrier. In some cases, correcting the metabolic disturbance or reducing the CNI drug level will lead to resolution of neurotoxic side effects, though severe symptoms may require discontinuation of the CNI.

While CNIs are highly effective immunosuppressants, their toxic side effects, particularly nephrotoxicity, has led to the search for alternate immunosuppressant agents that might serve as substitutes or allow reductions in CNI dosage. Within recent years, a number of medications have been developed including mycophenolate mofetil, sirolimus, everolimus, belatacept, and newer monoclonal antibodies, such as basiliximab and daclizumab. Monoclonal antibodies are used for treatment of acute rejection and as induction therapy during the first days or weeks after transplantation, but not for maintenance therapy. These offer reduced nephrotoxicity and also generally demonstrate less neurotoxic side effects (Table 27.13–7). However, there are infrequent complications that can arise with either CNIs or some of the newer immunosuppressants—namely, posterior reversible encephalopathy syndrome (PRES) or, rarely, progressive multifocal leukoencephalopathy (PML). PRES usually arises in the first month posttransplantation, though it can arise years later, and presents with a range of symptoms including seizure, headache, visual changes, confusion, ataxia and other neurologic symptoms depending on the location of the lesion. Diagnosis via MRI usually reveals T2-weighted and fluid-attenuated inversion recovery (FLAIR) MRI images showing hyperintense lesions typically in the subcortical and cortical regions of the bilateral parieto-occipital lobes, but lesions can occur in other areas and in the brainstem. Discontinuation of CNI or sirolimus usually leads to resolution of clinical and radiologic abnormalities. PML is a rare and severe demyelinating disease caused by reactivation of JC virus that can develop in immunocompromised patient populations.

Table 27.13–6.
Pharmacotherapy Considerations in Transplant Patients

Medication	Specific Considerations for Transplant Patients
Antidepressants	
SSRIs	
Citalopram	Preferred, lack of drug–drug interactions
Escitalopram	Preferred, lack of drug–drug interactions
Fluoxetine	Reduce dosage or lengthen dosing interval in HI; risk of interaction with immunosuppressants.
Fluvoxamine	Consider lower initial dose and slow titration in HI; risk of interaction with immunosuppressants.
Paroxetine	Reduce initial dose and slow titration in moderate-to-severe RI and HI. Anticholinergic side effects.
Sertraline	May consider lower or less frequent dosing in HI.

SNRIs	
Venlafaxine	Reduce total daily dose in moderate-to-severe RI/hemodialysis/HI; risk of elevated blood pressure.
Desvenlafaxine	Dose should not exceed 50 mg/d in moderate RI; no more than 50 mg every other day in severe RI.
Duloxetine	Avoid use in HI and in severe RI.
Vortioxetine	Not recommended in severe HI/cirrhosis
Others	
Bupropion	Limit dose to 150 mg of XL formulation every other day in severe HI; 100 mg/day of SR formulation; reduce initial dose in RI.
TCAs	Cardiovascular side effects are undesirable in cardiac patients. Anticholinergic side effects.
MAOIs	Significant risk of drug–drug interactions make this undesirable in transplant patients.
Mirtazapine	Clearance reduced in moderate-to-severe RI/HI. Consider lower initial dose and slower titration, sedative, and weight gain side effects.
Vilazodone	Not recommended in severe RI
Nefazodone	Avoid use due to risk of drug–drug interaction with immunosuppressants.
Anxiolytics	
Benzodiazepines	Lorazepam, oxazepam, temazepam do not require hepatic oxidation. Safer choices in presence of HI; risk of respiratory depression.
Buspirone	Not recommended in severe RI or HI
	All pose risk of QT prolongation; atypical agents also increase risk of metabolic syndrome.
Antipsychotics	
Atypical	
Aripiprazole	No dose adjustment needed in RI/HI
Asenapine	Use contraindicated in severe HI
Clozapine	Risk of seizures, agranulocytosis, orthostatic hypotension complicate use in transplant.
Lurasidone	Moderate HI and moderate-to-severe RI: starting dose 20 mg/d, total dose not to exceed 80 mg/d; severe HI: total dose not to exceed 40 mg/d
Iloperidone	Use not recommended in severe HI
Olanzapine	Poses greater risk of metabolic syndrome.
Paliperidone	Clearance reduced in RI; for moderate-to-severe RI: starting dose 1.5 mg/d, total dose maximum 3 mg/d
Risperidone	Severe RI: start dose at 0.5 mg b.i.d. Slow titration is recommended.
Quetiapine	No dosage adjustment needed in RI/HI
Ziprasidone	No dosage adjustment needed in RI/HI; greater risk of QT prolongation
Typical	
e.g., haloperidol	No dosage adjustment needed in RI/HI
Mood stabilizers/anticonvulsants	
Lithium	Moderate RI: use 50–75% of usual dosage hemodialysis-supplemental dose of 300 mg once after each dialysis session.
	May be difficult to use due to shifting fluid status, unstable levels, risk of toxicity; cyclosporine reduces Li clearance.
Carbamazepine	CYP 3A4 inducer effects of CBZ lower levels of immunosuppressants, risk of organ rejection
Oxcarbazepine	Avoid extended release formulation in severe HI; severe RI: reduce initial dose to 300 mg/d, titrate slowly, CYP 3A4 inducer drops immunosuppressant levels.

Lamotrigine	Severe HI w/ascites: reduce initial, escalation, and maintenance dose by 50%; without ascites: reduce by 25%.
Topiramate	Moderate-to-severe RI: reduce dose by 50%; cognitive side effects may be problematic for transplant patients.
Valproic acid	Avoid use in HI if possible due to reduced clearance; reduced serum albumin can raise free VPA levels.
Gabapentin	Moderate RI: 600 mg/d; severe RI: 150–300 mg/d; hemodialysis: may require both maintenance and supplemental dose after dialysis session.
Psychostimulants	
Atomoxetine	Moderate HI: reduce initial and total dose by 50%; severe HI: reduce initial and total dose by 75%.
Methylphenidate	No dosage adjustment needed in RI/HI; use in cardiac patients limited by cardiovascular side effects
Amphetamine	No dosage adjustment needed in RI/ HI; use in cardiac patients limited by cardiovascular side effects
Modafinil	Severe HI: reduce dose by 50%.
Armodafinil	Severe HI: reduce dose due to reduced clearance; severe RI: lack of dosing info prevents safe use.
Anticraving medications	
Buprenorphine	Moderate-to-severe HI: use with caution and at lower dose; may have longer exposure and activity.
Naltrexone	Use in severe RI/HI has not been studied—thus, safety is unknown.
Methadone	RI/HI: lower initial dose and slow titration advised; can prolong QT interval; use with caution in cardiac patients.
Varenicline	Severe RI: initial dose of 0.5 mg/d, may increase to 0.5 mg bid; risk of neuropsychiatric side effects.
Acamprosate	Moderate RI: initial dose of 333 mg t.i.d.; severe RI: use is contraindicated.
Disulfiram	HI: use with caution; may cause hepatotoxicity even without alcohol use; a metabolite is a 3A4 inhibitor so could raise immunosuppressant levels.

HI, hepatic insufficiency; RI, renal insufficiency; moderate RI, creatinine clearance 30–50 mL/min; severe RI, creatinine clearance <30 mL/min; moderate HI, Childs-Pugh class B; severe HI, Childs-Pugh class C.

Table 27.13–7. Neurotoxicity of Immunosuppressant Medications

Medication	Side Effect(s)
Calcineurin inhibitors	Headache, tremor, paresthesia, visual disturbances, insomnia, delirium, seizures, apraxia, parkinsonism, cortical blindness, anxiety, mood changes, cognitive impairment, agitation, PRES, PML
Corticosteroids	Euphoria, depression, anxiety, agitation, insomnia, delirium, irritability, hallucinations, psychosis, cognitive impairment
Sirolimus	Headache, dizziness, tremor, depression, insomnia, PRES, PML
Everolimus	Tremor, somnolence, dizziness, paresthesia, hypoesthesia
Mycophenolate mofetil	Headache, insomnia, anxiety, dizziness, PML
Belatacept	Headache, Guillain–Barré syndrome, PML
Azathioprine	PML
Monoclonal antibodies	
Basiliximab	Headache, tremor, insomnia, anxiety, depression, fatigue

Daclizumab	Insomnia, fatigue
Rituximab	Anxiety, depression, delirium, hallucinations
Alemtuzumab	Headache, dizziness, light-headedness, weakness, fatigue

PRES, posterior reversible encephalopathy syndrome; PML, progressive multifocal leukoencephalopathy.

REFERENCES

Abecassis M, Adams M, Adams P, Arnold RM, Atkins CR. Consensus statement on the live organ donor. *JAMA*. 2000;284:2919.

Bowdish ME, Barr ML, Schenkel FA, Woo MS, Bremner RM. A decade of living lobar lung transplantation: perioperative complications after 253 donor lobectomies. *Am J Transplant*. 2004;4(8):1283.

Coffman K, Crone C. Rational guidelines for transplantation in patients with psychotic disorders. *Curr Opin Organ Transplant*. 2002;7:385.

De Bleser L, Matteson M, Dobbels F, Russell C, De Geest S. Interventions to improve medication-adherence after transplantation: a systematic review. *Transpl Int*. 2009;228:780–797.

*De Geest S, Burkhalter H, Bogert L, Berben L, Glass TR, Denhaerynck K; Psychosocial Interest Group; Swiss Transplant Cohort Study. Describing the evolution of medication nonadherence from pretransplant until 3 years post-transplant and determining pretransplant medication nonadherence as risk factor for post-transplant nonadherence to immunosuppressives: the Swiss Transplant Cohort Study. *Transpl Int*. 2014;27:657–666.

Denhaerynck K, Dobbels F, Cleemput I, Desmyttere A, Schäfer-Keller P. Prevalence, consequences, and determinants of nonadherence in adult renal transplant patients: a literature review. *Transplant Int*. 2005;18:1121.

*Dew MA, DiMartini AD, De Vito Dabbs A, Myaskovsky L, Steel J. Rates and risk factors for nonadherence to the medical regimen after adult solid organ transplantation. *Transplantation*. 2007;83(7):858.

*Dew MA, DiMartini AF, Kormos RL. Organ transplantation, stress of. In: Fink G, ed. *Encyclopedia of Stress*. Vol. 3. 2nd ed. Oxford, UK: Academic Press (Elsevier); 2007:33.

Dew MA, DiMartini AF, Steel J, De Vito Dabbs A, Myaskovsky L, Greehouse J. Meta-analysis of risk for relapse to substance use after transplantation of the liver or other solid organs. *Liver Transplantation*. 2008;14(2):159.

Dew MA, Jacobs C, Jowsey SG, Hanto R, Miller C. Guidelines for the psychosocial evaluation of living unrelated kidney donors in the United States. *Am J Transplant*. 2007;7(5):1047.

Dew MA, Kormos RL, Winowich S, Harris RC, Stanford EA. Quality of life outcomes after heart transplantation in individuals bridged to transplant with ventricular assist devices. *J Heart Lung Transplant*. 2001;20(11):1199.

*Dew MA, Rosenberger EM, Myaskovsky L, et al. Depression and anxiety as risk factors for morbidity and mortality after organ transplantation: a systematic review and meta-analysis. *Transplantation*. 2015;100(5):988–1003.

DiMartini A, Chopra K. The importance of hepatic encephalopathy: pre- to post transplant. *Liver Transplant*. 2009;15:121–123.

*DiMartini A, Crone C, Fireman M, Dew MA. Psychiatric aspects of organ transplantation in critical care. *Crit Care Clin*. 2008;24(4):949.

DiMartini A, Crone C, Dew MA. Alcohol and substance use in liver transplant patients. *Clin Liver Dis*. 2011;15(4):727–751.

*DiMartini A, Day N, Dew MA, Javed L, Fitzgerald MG. Alcohol consumption patterns and predictors of use following liver transplantation for alcoholic liver disease. *Liver Transplant*. 2006;12:813.

Dobbels F, Verleden G, Dupont L, Vanhaecke J, De Geest S. To transplant or not? The importance of psychosocial and behavioural factors before lung transplantation. *Chronic Respir Dis*. 2006;3(1):39.

Gross CR, Kreitzer MJ, Russas V, Treesak C, Frazier PA. Mindfulness meditation to reduce symptoms after organ transplant: a pilot study. *Adv Mind Body Med*. 2004;20:20.

*Gross CR, Messersmith EE, Hong BA, et al. Health-related quality of life in kidney donors from the last five decades: results from the RELIVE study. *Am J Transplant*. 2013;13(11):2924–2934.

Jain A, DiMartini A, Kashyap R, Youk A, Rohal S. Long-term follow-up after liver transplantation for alcoholic liver disease under tacrolimus. *Transplantation*. 2000;70:1335.

Jowsey S, Schneekloth T. Psychosocial assessment of adult liver transplant recipients. In: Busuttil RW, Klintmalm GK, eds. *Transplantation of the Liver*, 2nd ed. Philadelphia, PA: Elsevier Saunders; 2005:395–404.

Liu L, Schiano T, Lau N, O'Rourke M, Min A. Survival and risk of recidivism in methadone-dependent patients undergoing liver transplantation. *Am J Transplant*. 2003;3:1273.

Maldonado JR, Dubois HC, David EE, et al. The Stanford Integrated Psychosocial Assessment for Transplantation (SIPAT): a new tool for the psychosocial evaluation of pre-transplant candidates. *Psychosomatics*. 2012;2:123.

*Olbrisch ME, Benedict SM, Ashe K, Levenson J. Psychological assessment and care of organ transplant patients. *J Consult Clin Psychol*. 2002;70:771.

Parekh PI, Blumenthal JA, Babyak MA, LaCaille R, Rowe S. Gas exchange and exercise capacity affect neurocognitive performance in patients with lung disease. *Psychosom Med*. 2005;67(3):425.

Popoola J, Greene H, Kyegombe M, MacPhee IA. Patient involvement in selection of immunosuppressive regimen following transplantation. *Patient Prefer Adherence*. 2014;8:1705–1712.

Skotzko CE, Strouse TB. Solid organ transplantation. In: MG Wise, JR Rundell, eds. *The American Psychiatric Publishing Textbook of Consultation-Liaison Psychiatry: Psychiatry in the Medically Ill*. Washington, DC: American Psychiatric Press; 2002:623.

Smith PJ, Rivelli S, Waters A, et al. Neurocognitive changes after lung transplantation. *Ann Am Thorac Soc*. 2014;11(10):1520–1527.

Tan HP, Marcos A, Shapiro R, eds. *Living Donor Transplantation*. New York: InformaHealthcare; 2007.

Tryc AB, Pflugrad H, Goldbecker A, et al. New-onset cognitive dysfunction impairs the quality of life in patients after liver transplantation. *Liver Transplant*. 2014;20(7):807–814.

▲ 27.14 Psychiatric Care of the Burned Patient

MARTHA C. GAMBOA, M.D., JULIANNE K. SUOJANEN, D.O., EJINE OKOROAFOR, M.D., AND MICHAEL BLUMENFIELD, M.D.

INTRODUCTION

Burn injuries represent a major global public health burden and constitute the fourth most common type of trauma worldwide, following traffic accidents, falls, and interpersonal violence.

Across the world, fire-related burns account for more than 300,000 deaths per year, with additional deaths resulting from other types of burns. The vast majority of these deaths (95 percent) occur in low- and middle-income countries. In the United States alone, around 3,400 patients who have suffered flame and smoke inhalation injuries die each year, more than 400,000 burn patients require medical attention, and nearly 40,000 require hospitalization annually.

Major medical and surgical advances have greatly increased the burned patient's survival. The 2014 National Burn Repository (NBR) reviewed the combined data set of more than 190,000 acute burn admissions for the time period between 2004 and 2013 from 102 burn centers (96 from the United States, 4 from Canada, and 2 from Sweden) and found that during the last 10 years the mortality rate from burn injuries has decreased from 3.4 to 2.7 percent for males and from 4.6 to 3.3 percent for females, and that the average length of stay of burn patients has also declined from 9 to 8 days. Increased survival rates translate to a larger number of burn patients that require physical, psychological and social rehabilitation. Burn patients often suffer severe pain, major psychological distress, and very complicated psychosocial disruption to their lives. Psychiatrists and other mental health professionals working in the medical or surgical setting have a major role to play in the care of these patients.

RISK AND ASSOCIATED FACTORS FOR DEVELOPING BURN INJURIES

Burns result from the destruction of one or more cell layers of the skin or other organ tissues in contact with hot liquids (scalds), solids (contact burns), or flames (flame burns). Injuries due to ultraviolet radiation, radioactivity, electricity, or chemicals and respiratory damage resulting from smoke inhalation are also considered to be burns.

According to the WHO, some of the major risk factors for sustaining burn injuries around the world are alcohol and smoking, particularly in combination. Contributory factors include: local cultural practices, such as floor-level stoves, bedside fires, loose flammable attires, and fire walking; socioeconomic status (for both developed and developing countries);

overcrowded living conditions; lack of proper safety measures; and insufficient parental supervision. There is notably an increased risk of burn injuries in vulnerable populations, including children, women, and the elderly. According to the WHO, globally, more than 95,000 minors die from burns each year, or approximately 262 children per day.

The 2014 NBR reported that in the United States the two most common etiologies for burn injuries across all age groups are flame burns and scalds, accounting for almost 8 out of 10 reported burn cases. Scalds are more prevalent in the younger population, especially among those younger than 5 years of age, while fire/flame burns prevail in the remaining age categories. The vast majority of burns seen at burn centers were considered accidental, with only 14.1 percent of these related to work, and only 3.6 percent specified as nonaccidental. Home continues to be the most common place of burn injury occurrence (73 percent). According to the 2014 NBR, males constitute 69 percent of all burn victims. Males also predominate in almost all age categories in comparison to females. Children younger than 16 years of age accounted for 29 percent of all patients, and those under the age of 5 constituted 19 percent of the cases. Patients age 60 or older represented 13 percent of the cases. The age group with the highest burn prevalence was that between 20 and 60 years old, accounting for greater than 50 percent of all patients. The mean age for all cases was 32 years.

The risk of burns in children usually results from the interplay between environmental, developmental, behavioral, and familial factors. Some of the environmental risk factors include lack of supervision of children, proximity to burning substances, clothing made of flammable materials, congested housing, low socioeconomic status, and poor social supports. Behavioral factors include risk-taking tendencies, fire setting, as well as abuse of alcohol or illicit drugs. Familial factors including parental illiteracy, parental psychopathology, and parental stress or neglect also contribute to burn injuries in children. Child neglect or abuse causes about 6 to 20 percent of pediatric burns, with scalds as the prevalent etiological factor.

The average rate of burn injuries decreases from age 4 until adolescence. However, after age 15, the incidence rate of burn injuries begin to increase again, presumably because of greater exposure to hazards, experimentation, and risk taking, as well as employment. Psychiatric diagnoses usually associated with increased risk of burn injuries include childhood depression, learning disorders, intellectual disabilities, substance use disorders, oppositional defiant disorder, conduct disorder, attention deficit hyperactivity disorder, and bipolar disorder.

The risk factors for burn injuries in adults include chronic mental illness, neurocognitive disorders, alcohol and substance use disorders, character disorders, history of suicide attempt, and history of abuse or

homicide. Factors like prior physical or sexual abuse, substance use disorders, affective disorders, psychotic disorders, and personality disorder significantly contribute to the risk of self-inflicted burns. Conduct disorder, personality disorders, bipolar disorder, major depression, schizophrenia, factitious disorder, and substance use disorders are some of the psychiatric diagnoses associated with increased risk for self-inflicted burns. Female gender, black ethnicity, unemployed, single, and homeless status and Medicaid funding are common characteristics found in the majority of patients with assault-related burn injuries.

Some medical and neuropsychiatric conditions interfere with intact sensory, motor, or cognitive functions, and in turn decrease the patient's ability to properly recognize and respond to situations that can represent a risk for burn injury. In the setting of insufficient or inappropriate care giver supervision, early onset neurological impairments including cerebral palsy, spina bifida, attention deficit hyperactivity disorder, intellectual developmental disorder, and autism increase the risk to sustaining scalds, flame burns, and contact burns during activities related to cooking, eating, and bathing. The lack of temperature control settings in the shower for this group of patients is an additional risk factor.

A 26-year-old single woman with a history of newly diagnosed multiple scleroses (MS) and no prior psychiatric history was at home, cooking on her gas stove, when she accidentally slipped on the floor, her loose clothes catching fire. She was found to have sustained 30 percent TBSA partial thickness flame burns, as well as mild diplopia and monoparesis of the left lower extremity. During psychiatric assessment, she had endorsed mild depressive symptoms a month prior to her burn trauma, including sad mood related to her new MS diagnosis, decreased appetite, and poor concentration. While in the burn unit, her depressive symptoms had worsened and citalopram (Celexa) was commenced. Her anxiety and insomnia were effectively managed with low-dose clonazepam (Klonopin) and brief supportive psychotherapy sessions. After several weeks of hospitalization and treatment she was transferred to a rehabilitation unit where she had a home safety assessment before her discharge. She was referred for follow-up to a local outpatient mental health clinic.

ALCOHOL, SUBSTANCE ABUSE, AND BURN INJURY

The association of burn injuries and alcohol is widely studied and accepted. It has both a direct and an indirect causal effect. Approximately 50 percent of burn patient fatalities have a history of alcohol use. Alcohol use is also known to increase the morbidity and mortality of hospitalized burn-injured patients as a result of increased susceptibility to bacterial infection and suppression of cellular immunity. Burn patients admitted with either alcohol intoxication or a history of alcohol use tend to have increased percentage of total body surface area (TBSA) burned, higher rate of inhalation injury and mortality, longer length of hospital stay, and consequently increased medical costs in comparison to burn patients who are nonalcoholic users. Alcohol exposure is frequently associated with deaths from cigarette-ignited fires. Furthermore, an association exists

between alcohol and/or substance use with suicide by burning and other intentional burn injuries, including assault and self-mutilation.

Illegal drug manufacturing and the use of peculiar drug paraphernalia pose added and somewhat novel risks for various types of thermal injury. For example, free-basing cocaine may cause burns to the face and hands, smoking crack has led to accidental inhalation of crack pipe "brillo" screens resulting in burns to trachea, mouth, and esophagus. Methamphetamine laboratory explosions have been shown to inflict blast injury and smoking in bed while intoxicated can cause flame burn and inhalation injury. Ignited alcoholic drinks are vastly becoming a popular fad, whereby mixed high-proof alcohol drinks are ignited before serving and pose significant risks for facial, chest, esophageal, and other burn injuries.

In the review of agitated or confused patients in the burn unit, a cautious clinician should always rule out alcohol or other substance intoxication. Blood alcohol level and urine toxicology screens should be standard tests in evaluation of burn patients in order to provide them with optimal care. Timely intervention and initiation of treatment for withdrawal symptoms undoubtedly helps to save lives, time, and money. Uncontrolled withdrawal poses significant problems for both patient and staff and utilizes a staggering degree of hospital resources as patients may require any combination of services, including one-to-one observation due to multiple periods of agitation or attempts to leave against medical advice, frequent (sometimes hourly) assessment of vital signs, frequent administration of intravenous (IV) medications, close observation for effects of CNS depression secondary to benzodiazepine treatments, and intubation if withdrawal progresses to alcohol withdrawal delirium and is accompanied by seizures. Fluctuating fluids levels, multiple pain medications, anoxic injury, and CNS anesthetics greatly complicate the use of any standard alcohol withdrawal protocol. These patients must be assessed frequently for withdrawal and their treatment adjusted as needed. This treatment may be required for days to weeks depending on the severity and duration of the withdrawal, which is affected by the age of the patient, the number of drinking years, the level of sustained drinking, the existence and severity of comorbidities, and a history of withdrawal symptoms. Drug or alcohol abuse or dependence can profoundly impact all phases of burn recovery, including the rehabilitation phase. Chronic alcohol use can: (1) depress the immune function, leaving compromised patients less able to battle infection; (2) cause malnutrition, when adequate nutrition is one of the most important factors in healing from burn injury and graft surgery; and (3) damage multiple organs such as liver, heart, and brain.

Negative transference reactions to the agitated, withdrawing patient can evolve in staff members who are repeatedly and directly involved with a difficult patient's care. It is important to address these issues as they arise

to avoid any potential acting out behaviors from staff.

DELIRIUM

Delirium is a neurocognitive disorder caused by dysregulation of the baseline neuronal activity secondary to systemic disturbances. It is characterized by an alteration in the level of attention and awareness and it also affects other areas of cognitive functioning. Delirium develops over a relatively short period of time, represents a change from the patient's baseline, and can have a rapid and fluctuating course. Agitated delirium in the critically ill burn patient can complicate the treatment and even be life threatening. Severely agitated patients may pull out arterial lines or tubes, pull off dressings, and shear skin grafts, lengthening healing time for wounds and potentially exposing them to infectious agents. Diagnosing delirium in the burn patient, as in any patient, should lead to a search for the underlying etiology and appropriate treatment. In burn patients, delirium may be due to the systemic inflammatory response, sepsis, and multiorgan dysfunction syndrome. The manifestations of delirium may be mistakenly attributed to mental illness in burn patients with a known history of psychiatric problems. The most commonly identified causes of delirium in burn patients are opiates, cerebral anoxia, skin infections with or without associated sepsis, fluid or electrolyte disturbances, and withdrawal from opiates or benzodiazepines.

Vital signs, including pulse oximetry along with ECG and chest radiography, constitute crucial information in the evaluation of a delirious patient. Appropriate laboratory tests include vitamin B_{12}, complete blood count, and blood chemistry, including calcium, magnesium, phosphorous, blood urea nitrogen, ammonia, uric acid, liver function tests, drug levels, urinalysis, and thyroid-stimulating hormone. Prolactin and troponin levels should be checked if clinically indicated. Other tests such as CT scan, MRI, EEG, cerebrospinal fluid examination, lower extremities Doppler ultrasound for suspected vein thrombosis, arterial blood gases, and lung perfusion tests for suspected pulmonary embolus should be pursued when clinically relevant.

Treatment of Delirium in the Burn Patient

Treatment of delirium includes the identification and correction of the underlying medical problems. The first priority in the treatment of any patient with delirium is safety. If confusion, agitation, and inappropriate behavior are predominant, the patient requires protection from falls, wandering, hurting others, or pulling out venous accesses, tubing, or other necessary equipment. The use of side rails on the bed or soft restraints may be adequate measures for some patients, but others will need constant observation in a one-to-one fashion. Completion of the medical history by

interviewing the patient and obtaining information from collateral sources, including history of present illness, past psychiatric history, past medical and surgical history, social history, substance abuse history, and examination of the mental status and vital signs help determine the best pharmacological choices for treatment of agitation.

In burn patients, as in patients with other medical conditions, antipsychotics are used to control agitated delirium unless the cause of delirium is benzodiazepine or alcohol withdrawal, in which case benzodiazepines constitute the main treatment. Before using antipsychotics, it is important to check the QT interval corrected for heart rate (QTc) to reduce the risk of torsade de pointes and ventricular tachycardia. It is important and useful to inquire about a history of cardiac, liver and kidney disease, dystonic reactions, neuroleptic malignant syndrome, seizures, hypotension, diabetes, and hyperprolactinemia in order to choose the preferred pharmacological agent.

In adults, dystonia and akathisia are uncommon with the use of IV haloperidol. Although the use of IV haloperidol is not approved by the FDA, in the burn unit and other critical care settings, IV haloperidol is used under close cardiac monitoring to effectively and rapidly control hyperactive delirium in patients without oral access. Starting doses of IV haloperidol in burn adults and older children are 0.5 to 2 mg. Incremental and frequent doses may be necessary, while the patient is closely observed and on a cardiac-monitored intensive burn unit bed. Rapid relief of delirium can be achieved within minutes or hours but can persist for days or weeks, especially if the etiology is not correctly identified and addressed. Patients with delirium and extreme anxiety or agitation may benefit from the addition of low-dose benzodiazepines, such as Lorazepam. However, caution is necessary because the addition of benzodiazepines can increase the risk of delirium. Therefore, if used, benzodiazepines should be tapered off quickly.

In children, it is especially important to differentiate insomnia, agitation without delirium, and anxiety from delirium as the former two are effectively treated with benzodiazepines. One retrospective study in which 26 critically ill burn children and acute pediatric patients suffering from burn wounds received haloperidol, mostly intravenously, found an association of this drug with dystonia, hypotension, hyperpyrexia, and respiratory and renal failure, leading to death in one child. Despite these observations, haloperidol continues to be often used parenterally to control agitation in older burn children for 2 to 3 days with rapid relief. It is essential that such treatment be very carefully monitored. Atypical antipsychotics can be used in adolescents and adults for the treatment of agitated delirium if oral access is available, although their use has not been studied in the burn patient.

Along with the appropriate psychopharmacology, an important aspect

in the treatment of delirium is to facilitate and promote orientation to the environment. Allowing the patient who wears glasses and hearing aids to have them available and encouraging him or her to wear them, placing a visible calendar in the patient's room, and giving the patient simple statements regarding the procedures that will be done and the things that are expected from him or her are simple things that can help the patient to be oriented to place and time. Avoiding unnecessary changes in the room setting and the nursing staff who care for the patient may be helpful measures to help orient the delirious patient to an unusual and possibly perceived threatening environment.

PAIN MANAGEMENT

A patient in significant pain has little tolerance for a meaningful conversation; therefore, it is not possible to properly assess psychological issues such as the reaction to the injury, depression, guilt, body image, or cognitive deficits in such a state. Any effort to talk to the patient without addressing pain status will likely be experienced as insensitive and irrelevant. The proper treatment of burn injuries usually requires debridement, daily wound care, skin graftings, and retraction release surgeries, followed by sometimes weeks or months of physical therapy. These treatments can be extremely painful. Patients have different thresholds for pain, different coping abilities, and different physiological responses to treatment. For these reasons it is important to inquire about the adequacy of pain control during each contact with the burned patient and address it accordingly.

Background pain refers to the underlying pain from the initial injury. It is ongoing and present even in the absence of activity or procedures. Management of this type of pain is generally addressed through the use of long-acting analgesic agents. Breakthrough pain is more intense, episodic, and associated with activities of daily living and other minor activities that require movement or manipulation of injured areas. It is usually managed with short-acting agents. Procedural pain is associated with invasive procedures and ongoing daily burn care, such as wound cleansing, dressing changes, and with physical and occupational therapy. A structured approach to pain management should aim to control all three types of pain. Acute pain disrupts normal physiological functions and increases stress. Poor pain control causes sleep disturbances and increased anxiety, which further exacerbate pain. It has been shown that uncontrolled pain has physiological effects that impede wound healing. In addition to the pain that arises from the burned tissues and donor sites, burned patients may experience additional pain from wound infections, concomitant injuries suffered at the time of the burn, pre-existing conditions, or abdominal pathology related to sepsis or due to heterotopic calcifications.

Psychological addiction is infrequently caused in the context of burn

care, but some patients treated with high opiate doses for long term can develop physiological dependence, and acute tolerance has been documented in children during the initial phases of intensive care. There is no evidence that opioid addiction occurs more often in burn patients compared to other patients requiring opiates for acute pain. Active substance abusers who are victims of burn injuries require higher doses of opiates than their counterparts to achieve adequate analgesia. Frequent assessment of the patient's pain and general condition by experienced burn unit staff and the collaboration of the pain service consultants help to elucidate the appropriate time for changes in the medications' dosages and forms of administration, in an effort to effectively control the pain and at the same time avoid unnecessary exposure to high opiate dosing. Despite the goal of adequate pain control, patients are frequently undermedicated for pain. This can lead to the full range of psychiatric symptoms, which include anxiety, depression, and PTSD.

Although quantifying pain is always difficult, a burn injury must be considered the most painful trauma that a person can sustain. Methods to measuring pain include subjective reports, as determined by rating scales and questionnaires, and more objective measures such as physiological parameters. The visual analog scale, the numerical pain scale, and the Wong–Baker faces pain scale appear among the most used. The Initial Pain Assessment Tool and the Brief Pain Inventory are two useful questionnaires that provide options for initial evaluation of acute or chronic pain by a comprehensive set of questions on the nature and effects of pain. It has been accepted that on a 0 to 10 scale (where 10 equals the maximal pain) the pain is well controlled if it is patient-self-rated between 3 and 5 or less at rest and between 5 and 7 or less during procedures. An unmarked 100-mm scale can also be used, where the score is determined by measuring the distance between the two poles of "no pain" and "the most pain you can imagine." This visual analog scale translates the subjective experience of pain to a specific number from 1 to 100. An electronic device can also be used to do this. There are other clinical indicators of pain. Sweating, increased pulse, elevated blood pressure, agitation, and increased respiratory frequency can be due to pain, but they also could be due to hypoxia or alcohol withdrawal, to name a few examples.

Pharmacology of Pain Management

Opioid administration is the foundation of effective burn pain management. Opiates and benzodiazepines are often given together for the management of pain and pain-associated anxiety control in the burn patient. It is often observed that burn patients require exceptional doses of opiates and benzodiazepines to relieve pain and anxiety, compared to the standards in nonburn patients, to compensate for their hypermetabolic

state. The preferred routes of opiate administration are IV and oral, taking into account the fact that a normal intestinal function is needed to ensure effectiveness of the oral route. Morphine (MS Contin) administered parenterally or enterally is very effective in relieving acute burn pain, and with appropriate ventilatory assistance its dose should be titrated to the level needed for analgesia. Benzodiazepines usually enhance opiate response. For pain-associated anxiety, midazolam (Versed) and lorazepam are the benzodiazepines more commonly used. For adults on ventilatory support, starting doses of midazolam are 5 to 10 mg per hour and for children 0.05 to 0.1 mg/kg per hour. These doses will need titration according to anxiety ratings and the clinical situation. As the pain decreases and after a taper of 10 percent of IV morphine and midazolam, gradual transition to lorazepam, oxycodone (OxyContin), codeine, or acetaminophen (Tylenol) is usually made. Lorazepam added to opioid analgesics during procedural pain in patients with high baseline pain and state anxiety appears to have an analgesic effect with significant reduction in pain ratings.

Meperidine hydrochloride (Demerol) is not used due to its short half-life and because normeperidine, its major metabolite, can cause agitation, hallucinations, and seizures and may accumulate to dangerous levels with repeated dosing. Methadone (Dolophine) is a long-acting opioid with properties similar to morphine. Its *N*-methyl-D-aspartate (NMDA) antagonist properties are postulated to moderate opioid-induced tolerance, hyperalgesia, and neuropathic pain. Fentanyl (Actiq), a synthetic opioid around 100 times more potent than morphine, is available in injectable, transmucosal, and sustained transdermal presentations. The oral transmucosal form appears promising in addressing anxiety and pain in burned children. Anesthetic agents such as ketamine (Ketalar), propofol (Diprivan), and nitrous oxide are being used outside the operating room more often to facilitate the performance of burn care.

Psychotherapeutic Strategies for Pain

Psychological approaches to pain management can increase the patient's sense of control over the painful burn care procedures. Any psychotherapeutic approach that successfully addresses the patient's anxiety, depression, guilt, acute stress syndrome, or other psychological issues stands a good chance of reducing the patient's pain experience. Some randomized prospective small-sample studies using adjunctive psychological techniques in the burned patient utilizing pain and anxiety self-report and as well as observer-rated measures have been conducted.

In order to use cognitive interventions for the adjunctive treatment of pain, it is important for the clinician to recognize the patient's coping skills. Patients who cope with avoidance will be most likely to benefit from

distraction during short procedures and deep relaxation and distraction imagery during longer procedures. Patients who focus on the procedure may benefit from reappraisal techniques. They should be encouraged to differentiate sensory from affective components of pain and evaluate its meaning. They usually benefit from being told that pain sensation is a positive sign of wound healing. Distraction therapy using pleasant videos along with administration of analgesics during burn dressing changes has demonstrated significant reduced pain and anxiety. There is also some supporting evidence that the use of virtual reality provides analgesia with minimal side effects and little impact on the physical hospital environment. Virtual reality takes distraction one step further by using a visual stimulus instead of the power of suggestion. Visual stimulus is provided through a headset that covers the eyes. Auditory stimulus can be provided through associated earphones. A controlled study using virtual reality showed a reduction in pain, nausea, and anxiety during physical therapy. The management of burn pain is multidisciplinary and dynamic as the burn wound pain is essentially multifaceted. Several psychological approaches have been proven effective as adjuvants of the pharmacological treatment, but larger studies are still needed. Pain management needs to be individualized to each patient's needs and continuously monitored for optimum control.

Hypnosis and Pain

Hypnosis can help to control pain during burn wound debridement and is a viable adjunct treatment for burn pain. Hypnosis can be used in the short term. It has been reported that in less severe burns, hypnosis may help preventing the post-burn inflammatory response and reduce burn severity. Hypnosis is also helpful as an adjuvant in the treatment of pain and pain-related anxiety. During burn dressing changes, anxiety appears to be better controlled by hypnosis than by stress-reducing strategies. Hypnosis is particularly valuable when there are medical contraindications to adequate narcotic control of pain.

Relaxation Techniques

It is widely accepted that relaxation enhances the effectiveness of analgesics by reducing muscle tension and increasing blood flow. Breathing exercises, guided imagery, and deep muscle relaxation are all examples of practical bedside interventions to alleviate tension and enhance the patient's sense of control and well-being.

An elderly Hispanic married couple was admitted to a burn unit after burning down their trailer for the second time while intoxicated and smoking cigarettes. They were successfully treated for second- and third-degree burns of less than 30 percent total body surface. On initial psychiatric assessment prior to their developing alcohol withdrawal symptoms, they were

found mildly intoxicated but orientated, able to recount the burn event, and nonpsychotic. However, within another 12 hours of hospitalization, they had experienced significant withdrawal symptoms involving psychotic phenomenon, tremulousness, unstable vital signs, and agitation. They required close observation as they persisted in trying to leave the hospital against medical advice in their paranoid, agitated, and confused state. They had received lorazepam (Ativan) tailored to withdrawal progression or response. The husband required small doses (0.25 to 0 .5 mg) of haloperidol (Haldol) approximately 7 days into his admission for persistent psychotic symptoms, and had developed severe extrapyramidal symptoms and dystonia. The wife on the other hand made an uneventful recovery with resolution of her withdrawal symptoms within 7 days and was discharged from the hospital. Her husband's delirium on the other hand, had persisted for extra 2 weeks, complicated by treatment for urinary tract infection and anemia. This case helps to exemplify the risks of substance use, individual variations in treatment response and duration of alcohol withdrawal, and potential behavioral issues arising from the treatment of trauma patients in withdrawal.

An alcoholic homeless 76-year-old white male was admitted to the burn unit of a large teaching hospital after accidentally burning down his old boat house, sustaining a 45 percent TBSA burn. His protracted hospital stay involved an amputation, recurrent infections, pneumonia, and alcohol withdrawal, superimposed on a dementia illness per psychiatric assessment. He was consistently confused, confabulating, and paranoid but without alternating levels of consciousness, a hallmark of typical delirium. His discharge disposition or placement was further prolonged by his homeless and amputee status after he was fairly recovered. While awaiting placement, his confusion and confabulation were notably resolving. He was soon able to read the papers, hold coherent conversations, and tell jokes with his caregivers. This case exemplifies an elderly deconditioned, acutely unwell alcoholic male patient undergoing extended recovery from both burn injury and alcohol withdrawal. He had required substantive amounts of benzodiazepines titration and adjustment to his condition, as well as thiamine and B12 for probable Korsakoff psychosis.

BURN INJURY AND PSYCHOPATHOLOGY

Through the stages of acute treatment to rehabilitation and reintegration, burn patients contend with multiple stressors, including severe pain, disfiguring scars, loss of functionality, and other distress. Some of these factors lead to enduring psychosocial and psychological difficulties. Greater psychosocial distress during acute hospitalization for burn injury has been shown to result in a more difficult rehabilitation and recovery. As recent major medical and surgical advances have greatly increased the survival rates of burn patients, there has also been increasing interest in their psychiatric and functional sequel.

Depression and PTSD, prevalent in 13 to 23 percent and 13 to 45 percent of cases respectively, have been the most common areas of research in burn patients. In the case of depression, risk factors have been found to include premorbid depression and female gender in combination with facial disfigurement. For PTSD, premorbid depression, type and severity of baseline symptoms, anxiety related to pain, and visibility of burn injury are all contributing factors. It is also noted that the emergence of acute stress disorder (ASD) and/or PTSD symptoms in the first few weeks of hospitalization can be a predictor of the subsequent development of chronic PTSD. However, PTSD symptoms can singularly occur months after burn injury has occurred in a significant minority of patients who had never exhibited symptoms of ASD or PTSD during acute hospitalization.

Both neuroticism and avoidant coping style are important personality dimensions linked to the development of PTSD as well.

Neuropsychological difficulties have been associated mainly with electrical injuries. Great levels of acute pain have been associated with negative long-term psychological effects such as acute stress disorder, depression, suicidal ideation, and posttraumatic stress disorder for as long as 2 years after the initial burn injury. Subjective reports of feelings of "alienation and anxiety" experienced during hospitalization are also linked with distress for up to 2 years in burn patients.

Burn patients can experience difficulties in sexual life and social interactions. Even burn patients who never develop chronic psychological sequelae following burn injury will frequently experience, especially early in recovery, episodic expressions of stress, depression, dysthymia, acute or PTSD, other anxiety disorders, or psychosis. The quality of life of burn patients initially seems to be lower in comparison to the general population; however, over a period of many years, quality of life normalizes or improves.

The treatment of psychopathologies associated with burn injuries does not differ greatly from the treatment of the general population with similar ailments. The psychiatrist on initial assessment must screen for underlying risk factors including premorbid psychiatric disorders, personality characteristics, peritraumatic and posttraumatic factors as well as proneness to shame, body image problems, and self-esteem issues in case of patients with disfigurement. Psychotherapy (cognitive, behavioral) and pharmacological treatments (SSRI, SNRI, etc.) are valuable in the treatment of depression. Exposure therapy and eye movement reprocessing and desensitization have been successful in treatment of PTSD as well as available pharmacological treatments. Additional measures, including social skills training, community interventions, sexual health promotion, and counseling have all been found to be helpful in enhancing the quality of life of these patients.

PSYCHOTHERAPEUTIC INTERVENTIONS

There has been a wide variety of psychotherapeutic modalities used in treating patients who have suffered burn injury. They include cognitive-behavioral, supportive, brief and short-term psychodynamic and exposure therapies, debriefing, eye movement desensitization, and psychoeducational interventions, all of which have reportedly been used with varying levels of success. Authors reviewing general trauma treatment were rather unimpressed by any of the myriad modalities, including medication therapy. However, exposure therapy garnered substantial research results as a relatively effective treatment for ameliorating some PTSD symptoms. Pain control, as previously discussed, may be one of the most crucial treatment modalities for adults and children because, if

effective, it may reduce the potential for developing posttraumatic symptoms. The choice of therapeutic intervention is often dependent on the stage of recovery from burn injury.

Acute Phase

During the acute phase, the burned patient faces multiple immediate stressors, including fear of death, acute pain, being in an acute care environment such as the emergency room and intensive care unit (burn unit), and concerns for other victims of the same burn event. The patient may experience delirium manifested by fluctuations in the level of consciousness, confusion, agitation, and disorientation. The delirium could be related to the acute physiological effects of the burn itself, alcohol withdrawal, metabolic complications, infection, medication or substances of abuse. At this stage, psychological interventions focus on immediate concerns such as pain control, anxiety, management of delirium, and sleep. The priority at this stage is survival; therefore, in-depth psychological interventions are not indicated at this point.

The initial psychiatric assessment offers patients an opportunity to recount the burn event if they are psychologically prepared to do so. As patients reveal their understanding of their injuries, the consulting psychiatrist in concert with the treating surgical team orients patients and families to burn unit protocols and prepares them for the journey that lies ahead. A crucial part of providing psychiatric care to a burned patient is building a therapeutic alliance with their family or support system. If the patient is unable to provide information, the initial history can be provided by collaterals, which can serve as an opportunity to facilitate the family's understanding of the injury and its treatment course. The presence of distressed family members may enhance the patient's anxiety. Efforts to moderate caregiver stress may also serve to lower a patient's anxiety and stress reaction.

Chemical neurological embedding of traumatic memories is thought to involve sympathetic nervous system release of norepinephrine. Adequate treatment of acute pain and anxiety symptoms after burn injury may prevent some psychological sequelae by moderating the serum levels of norepinephrine, and possibly other neurotransmitters and endogenous hormones associated with the stress response. Historically, treatments in the acute phase have also involved pharmacological methods such as oral and IV administration of opiates and benzodiazepines for pain and anxiety, SSRIs, tricyclic and monoamine oxidase inhibitor antidepressants, β-blockers, and D-cycloserine in conjunction with exposure therapy. The antihypertensive prazosin (an alpha adrenergic blocker) is being used to treat traumatic nightmares related to various types of trauma, including combat events. It is reported to improve sleep onset and, if taken during

the day, has been reported to reduce irritability and other hyperarousal symptoms in some. In one literature review, prazosin doses varied between 1 mg and 16 mg, though higher doses have been needed to achieve remission of nightmares and have been used with positive effects. Researchers continue to examine the role of neuroactive steroids in the development and treatment of PTSD and the biochemical or genetic individual differences that may underlie variable responses to pharmacotherapy. Benzodiazepines have not been found to be effective in treating symptoms of PTSD, although they are widely prescribed for anxiety and agitation. Some evidence is cited in a recent literature review to support the use of pharmacological intervention in reducing the risk for adults of developing PTSD after trauma, specifically hydrocortisone administered immediately after trauma exposure. In the same review, researchers found no supporting data for propranolol, escitalopram, temazepam, or gabapentin in the prevention of PTSD. Critical stress debriefing used in the early treatment of traumatic stress has been shown to have a neutral or harmful effect.

A 55-year-old male patient with no visible wounds sustained an electrical injury from contact with an underground power line. A psychiatric consult was ordered on day 1 of this patient's hospitalization for repeated signs and symptoms of anxiety and panic, including increased blood pressures, diaphoresis, subjective report of nightmares, and awakening in a cold sweat, "feeling in a daze." The patient associated the onset of symptoms with the sounds of the arrival and departure of the hospital medical transport helicopter. He had been transported to the hospital by helicopter after suffering electrocution. The patient was relocated within the burn unit and lorazepam was ordered to be given as needed for anxiety. Interestingly, once the source of his reaction was identified, he reported a reduction in anxiety while awake, even without medication. By subjective report, benzodiazepines helped this patient achieve restful sleep with a decrease in startled nighttime awakenings and nightmares.

A 34-year-old man came to the burn unit following a chemical burn on his hands and legs that occurred at work. He complained bitterly about his care and the various policies of the burn unit to the embarrassment of his wife. The staff frequently defended the care that he received and argued on behalf of the doctors. Negative feelings about the patient were discussed at psychosocial rounds with the mental health team. This patient believed that the statements that he was making were true, and he would increase his attacks if the caregivers attempted to argue with him. The approach with this patient was to acknowledge his beliefs but not to offer an opinion about them. Occasionally asking the patient for his patience with difficult situations steered him away from being defensive and put him in a more cooperative mode. As soon as possible, he was taught to do his own dressing changes and allowed to take showers to help him feel more in control and reduce his reliance on staff members, whom he did not trust. The family was counseled in an effort to help them understand that the patient's complaints may have stemmed more from his fear and sense of helplessness than from the quality of care.

Intermediate Phase

Once acute medical issues have been stabilized, increasing the likelihood of survival, the patient enters the intermediate phase. The focus of treatment shifts toward restorative care. Most of the burn wounds have been grafted

and are in various stages of healing, and ventilator support, sedation, and severe infections become less prevalent. However, patients continue to undergo painful procedures. They become more aware of the physical and functional impact of their injuries, and at the same time are more psychologically ready to participate in a psychiatric diagnostic assessment. Anxiety- and depression-related symptoms often begin to appear in this phase. Sleep disturbances can be related to anxiety, depression, pain, and the noises and routines of the intensive care setting. Symptoms and behaviors related to premorbid psychopathologies such as personality disorders and substance abuse may surface in this stage. A dysfunctional pattern of coping of the individual can have an adverse impact on outcomes, including increased length of hospital stay and more serious psychopathologies after the injury. During the intermediate phase, the grieving process is more evident. Burn patients may have lost family members, friends, pets, their home, or personal property as well as aspects of their former life: work, appearance, and physical abilities.

Optimum pain control facilitates the patient's ability to participate in treatment protocols and allows for the development of a sense of hope for recovery. As discussed above, psychological interventions are helpful adjuvants to appropriate psychopharmacological treatment in the management of pain and anxiety. CBT and hypnosis have proven to be effective in treating procedural pain. Virtual reality utilizing special computerized programs is used as a distraction technique during painful procedures. A recent pilot study using a manualized cognitive-behavioral treatment protocol, developed specifically for treating PTSD and psychosocial sequelae following burn injury, resulted in improvements in PTSD and depressive symptoms, self-image, and aspects of reintegration.

Brief supportive psychotherapy to treat depression and anxiety in combination with appropriate psychopharmacological interventions is also helpful in this phase in managing depression and anxiety. Informing patients that nightmares are common and usually subside over time can help reduce stress and anxiety, increasing patients' understanding of the psychological healing process. Other forms of psychotherapy utilize cognitively oriented psychoeducation. Large burn centers may offer a variety of rehabilitation support groups for adults, children, adolescents, and parents, using various modalities, including expressive and CBTs. The focus of interventions during the intermediate phase is the assessment of emotional and cognitive adaptation to the injury, treatment and surgical procedures, and readying patients for transfer to rehabilitation care, or discharge home and return to school or work. The importance of psychoeducation and preparation for future care and other postdischarge issues must be highlighted. Knowledge of and engagement in the planning of upcoming surgeries, for example, gives patients a sense of control as they anticipate and prepare for the experience. Supportive treatments, and

normalizing the emergence of psychiatric symptoms related to the trauma can be very helpful during this and all phases. Appropriate mental health care referrals should be made at discharge for surveillance and treatment of current or evolving psychological concerns.

Recovery Phase

The recovery phase usually begins at the time of discharge from the hospital, when the patient begins to reintegrate into society, and is marked by the promise of survival. Patients with severe burns may still require long-term outpatient physical rehabilitation, including painful procedures such as dressing changes and surgeries. Ongoing sufficient nutritional intake is still imperative for healing skin and grafts. During the recovery stage, patients have significantly more control over their treatment than while acutely hospitalized. The timing and extent of plastic and reconstructive surgeries are collaborative decisions made by the patient and surgeon. The duration of this phase depends on the extent of impairment and the need for corrective surgeries. A patient may require multiple surgeries over a span of several years to release scars and improve function or restore appearance. Forty-six percent of patients presenting for these reconstructive surgeries were found to suffer from mild-to-moderate depression. Body image dissatisfaction was predictive of increased depression scores.

Psychological challenges in this stage include regaining a sense of competence while adjusting to the practical limitations of the injury. Psychological considerations during this phase include body image changes, self-esteem issues, adjustment to functional impairments, depressive and anxiety disorders, and particularly PTSD and phobias. Ongoing issues of grief, guilt, and premorbid psychiatric conditions continue to require psychotherapeutic support.

Social stressors in this stage include vocational rehabilitation, readjustment to work, family strains, and lack of social support, among others. Appropriate interventions of this phase involve psychiatric outpatient follow-up, social skills training, support groups, peer counseling, vocational counseling, and substance abuse relapse prevention.

Functional and cosmetic outcomes following burn injury have dramatically improved over the past few decades. Evolving techniques of cell culturing have resulted in improved skin grafts, and advances in plastic and reconstructive surgery have reduced permanent disfigurement. With the advent of specialized burn and trauma units, climbing survival rates of even massively burned individuals has placed significant emphasis on the recovery phase.

Factors Influencing Recovery. Patients influence the rate of progress during rehabilitation by their degree of participation and cooperation. If

they are, or become, unmotivated, anxious, or depressed, outcomes will be less than optimal. One study related larger burns and more psychological distress (during acute hospitalization) to more physical impairment and a slower recovery. In addition, a greater subjective psychological burden was associated with slower psychosocial recovery and more post burn psychological impairment. Patient participation in recovery and rehabilitation programs can be maximized by adequate attention to pain and anxiety control. Treatment plans are more successful if premorbid physical, developmental, and psychiatric status is assessed in order to set realistic goals for recovery.

Three important factors affecting burn recovery are pain, brain injury, and overall psychological adjustment. Uncontrolled pain has been linked to increased anxiety and can limit motivation, affecting optimal recovery. Brain injury potentially limits full participation in rehabilitation efforts and can involve periods of impaired impulse control and agitated delirium. Psychological distress also affects motivation and participation. In this regard, whenever appropriate, there should be some effort to explore the question that nearly all patients ask themselves: "Why did this happen to me?" Burn injuries, like many traumatic experiences, can be a turning point for some people, demanding major internal and external shifts in priorities, self-image, relationships, and life goals. Useful psychological information can be obtained by asking the patient "Do you have any thoughts or fantasies about why this might have happened?" The answer to the question may present a starting point in the therapeutic process.

Although many burn victims sustain only small physical scars, due either to limited injury or to the skills of plastic and reconstructive surgeons, some studies have found that the size of burn is not directly related to the degree of periburn emotional trauma or later development of psychopathology. For example, small burns have been shown to result in disproportionately severe psychological trauma, especially if the injury involved hands (decreasing functional capacity), face (altering self-image and increased visibility of scarring or disfigurement), or perineum (complicates elimination and sexual activity/identity).

A 20-year-old Hispanic female with prior history of PTSD and substance abuse suffered 65 percent total body flame burns primarily involving her upper body including face and hands. Drug dealers had set the house on fire as she slept inside. Acute hospitalization lasted for approximately 1 year and she required chemically induced coma for over 1 month while fragile skin grafts healed and plastic and reconstructive surgeons attempted to repair her ears and nose. Several years after discharge, she continued to undergo reconstructive and release surgeries to remove keloid scars. During this recovery phase, the plastic surgery department, physical therapy office, mental health clinic, and Alcoholics Anonymous (AA) became alternate homes for her, where acceptance, hope, and structure were provided. At a point in her recovery, she gave meaning to this traumatic experience by using it in a constructive way. She began speaking to groups of adolescents about the realities and dangers of substance abuse. This case exemplifies much of what has been discussed in this section, including the relationship of burns to substance abuse and premorbid psychiatric disorders, advancements

in trauma and surgical treatments and survival of massive burn injury, importance of psychiatric involvement at all phases of burn injury, psychological adjustment to new body image, and finding meaning in a life-altering traumatic event.

THERAPEUTIC APPROACH TO GRIEF, REGRESSION, AND ANGER IN THE BURN PATIENT

Grief

Grief work is necessary, as any patient who suffered a burn has sustained some kind of loss: appearance, function, role, independence, sense of safety in the world, possessions, or even loved ones. In facilitating grief work, it is useful to have in mind certain principles:

1. *Acknowledging the loss*: It is useful to prompt the patient to look at the losses by bringing up the topic that family and others often avoid. Most patients are willing to talk about their losses and providing the patient with an opportunity to do so may promote an outpouring of emotions that the listener should be prepared to hear.
2. *Allow the patient to be with pain*: The pain of the loss is not eliminated with words. It is preferable to avoid rationalizations about "the good that will come of the loss," which often makes the caregiver feel better but not the patient. More helpful interventions are focused on providing the patient with space and time to convey the full power of his or her emotions and to feel heard by an empathic listener. After the patient has amply expressed his or her feelings, it is useful to provide reassurance about gradual subsiding of the emotional pain with time.
3. *Affirm the individual's experience as real and part of normal grief*: Somatization, preoccupation with the lost object, guilt, hostility, and changes in normal patterns of conduct are anticipated components of a normal grief reaction.
4. *Provide assurance by exploring past coping patterns*: Burn patients may perceive their loss as unbearable and may be helped to use their coping skills by being encouraged to think of how they dealt with other losses in the past if any. This process requires careful probing as the cumulative pain of prior losses not fully experienced may be rekindled.
5. *Monitoring danger signals*: Persistent thoughts of self-harm, neurovegetative symptoms of depression, failure to make progression in the grieving process, and exacerbation of previous mental illness should prompt pertinent interventions. Those include increasing the level of monitoring, adjusting psychotherapeutic and psychopharmacological approaches, and, if necessary, admitting the patient to an inpatient psychiatric treatment.

Regression

A common behavior noted among burn patients is regression. The patient,

feeling overwhelmed by the physical and psychological distress, may attempt to return to the comfort of earlier stages in life. Staff and family members may have negative reactions to this phenomenon. In managing regression it is useful to:

1. Recognize the regressed behavior as a coping mechanism.
2. Examine the caregivers' responses and provide them with necessary feedback.
3. Determine if the regressed behavior interferes with the immediate patient care goals and decide the proper intervention accordingly. Behavioral modification techniques with rewards for achieved goals are effective in some patients.
4. Develop a consistent plan to handle regressed behavior. Such plan needs to be clearly explained to the patient and his or her family allowing frequent communication.

Anger

When anger begins surfacing during interactions with medical or mental health staff it is useful to acknowledge this expressed emotion with a supportive approach instead of attempting to talk the patient out of any feelings he or she has. It is useful to pay attention to the interactions with family members, who usually feel uncomfortable with expressed negative emotions and try to redirect the patient to avoid talking about them. Some family members may express their own anger, fueling the patient's emotions.

Although nursing staff can handle and provide opportunities for the patient's expression and abreaction of angry feelings, it is sometimes advisable that a psychotherapist talk with the patient about his or her feelings and meet with the patient regularly even for a short period. On the surface patients' anger is directed outward toward causes of the accident, missing work, being in pain, being hospitalized, and perceived inconsiderate treatment by hospital staff. Patients often avoid expressing anger toward the hospital staff as they depend on them. The ultimate object of the anger may be "self." Patients blame themselves for being in the situation in which they were burned whether or not there is validity to this idea. This is not easy to accept and the patient may displace his or her anger toward themselves or toward family members. They may push away their family and even their health care team by utilizing acting-out behavior, projection, and projective identification. A differential diagnosis in terms of potential causes of mood changes or mood disorders needs to be considered and pharmacological interventions implemented when needed. Patients are really angry at themselves, for their role in the burn injury (real or imaginary), their lack of ability to avoid the situation, the losses they "caused," not recovering fast enough, needing pain medication,

not tolerating the hospitalization, and so forth. As these feelings are unconscious and unacceptable to the patient, they are displaced onto external objects, sometimes in excess.

In this setting, the most successful interpretation is that which allows patients to realize that they are really anger at themselves while blaming others. However, it is very important to be aware that taking away the defense against inwardly directed hostility can elicit profound depression and even suicidal ideation. Precautionary measures need to be in place. Supportive techniques may be used here to express respect for the patient and for their acknowledgment of their true feelings. Another technique is that of an inexact interpretation that can be used to diffuse the displacement of anger toward staff. Using objects directly or indirectly related to the burn injury such as "the malfunctioning furnace," "the drunk driver," and so forth, rather than focusing on the patient's guilt and self-blame, is more likely to be accepted by the patient.

Expressive and abreactive techniques, time-limited psychotherapy, and psychodynamic life narrative treatment approaches may be successfully applied in the treatment of burn patients to allow verbalization of emotions, examination of separation, and acceptance of the meaning of the injury in the context of patient's life trajectory.

Special Psychopharmacological Considerations

The pharmacological treatment of burned patients with acute stress disorder, PTSD, reactive anxiety, depression, and psychotic syndromes follows the guidelines for treatment of other populations. Nevertheless, the physiopathological alterations caused by the burn injury as well as the medical and surgical treatments provided to burn patients may affect psychopharmacology decisions, although there is a paucity of information regarding the clinical implications of these changes.

Usually, cautionary steps, "start low and go slow," are granted in making clinical decisions in the treatment of a seriously injured patient. In the burn patient, agitation and pain are often treated with continuous IV benzodiazepines and opiates, especially during the acute and intermediate phases. The prolonged exposure to these medications could lead to withdrawal if tapered quickly. Some clinicians find that an adequate weaning rate from these medications appears to be 10 percent a day. The use of clonidine (Catapres) and dexmedetomidine (Precedex) appears helpful in the treatment of opioid withdrawal. The use of benzodiazepines and opiates when given in the context of burn-related anxiety and burn pain does not usually cause subsequent drug abuse. Attention should be paid to potential drug–drug interactions when considering the addition of a psychopharmacological agent to the usually complex medication regime of a burn patient. For example, concomitant use of the antibiotic linezolid (Zyvox) can interact with adrenergic and serotonergic medications, causing

serotonin syndrome. Careful monitoring of medication side effects should always be done especially when two or more medications have similar known adverse reactions profile. Potentiation of CNS depression by concurrent use of opiates and benzodiazepines needs to be assessed routinely. Judicious use of neuroleptics in severely medically ill is important. Although it is rare, neuroleptic malignant syndrome (NMS) can develop with the use of antipsychotics or other dopamine blockers such as metoclopramide (Reclomide). Oxandrolone (Oxandrin), an anabolic steroid commonly used for burn patients in a catabolic state, has the potential of causing habituation, excitation, insomnia, depression, and changes in libido.

ASD and PTSD usually require a combination of antidepressants and anxiolytics. SSRIs are first-line treatment for PTSD in other trauma populations. Other antidepressants, benzodiazepines, anticonvulsants, antipsychotics, β-blockers, and α-2 adrenergic agents such as clonidine or dexmedetomidine have also been used. Careful consideration should be given to the use of MAOIs, as their concomitant use with opiates that have serotonin reuptake inhibition properties (meperidine, tramadol [Ultram], methadone, dextromethorphan [Benylin], and propoxyphene [Darvon]) may precipitate dangerous serotonin toxicity. SSRIs constitute the first pharmacological line of treatment of burned patients with major depressive disorder. TCAs are considered second-line treatment due to potential side effects.

PLASTIC SURGERY AND FACE GRAFTS

The overall incidence of burn injuries has declined during the recent decades in the United States, but the proportion of survivors of large burn injuries has increased, carrying with it more severe aesthetic and functional sequelae. Facial and body disfigurement as a result of burns are known causes of profound shame, stigmatization, and isolation. Deep burns frequently result in loss of functionally and damage of cosmetically important body areas. Many of these burn victims suffer long-term morbidity as a result of functional impairment and severe facial defects.

Deep burns can involve not only the skin but also adjacent tendons, cartilage, and bone structures, contributing to complex deformities. The defects that result from burn injuries include hypertrophic scars, contractures, disorders of pigment, and loss of anatomic parts. Burn reconstructive surgery attempts to help the burn survivor to recover, as much as possible, normal appearance and functioning. Unfortunately, there is still a need to identify and develop novel strategies to treat burn deformities in order to obtain improved outcomes. Early closure of the wounds translates to less severe scarring, lower mortality from infection, lower mortality from other complications, decreased stress, decreased hyper metabolic response, and decreased bacterial load.

Burn patients require multiple procedures arranged over months or years to address the deformities. Empathy from staff with the experience of being burned can assist the patient in adapting to a new appearance. Burns may profoundly affect body image, self-esteem, and interpersonal relations as well as ego functioning. Skin allografts transplanted between genetically nonidentical individuals of the same species mostly come from cadaveric sources but can also come from close family members. Close relative intermingled skin allografts along with autografts have been satisfactorily used in a small number of child and adult burn victims. Among other advantages, this technique allows active participation of family members in the treatment of the burn victim, creating other interesting dynamics and ethical issues. Synthetic acellular skin substitutes have been used in the treatment of burns, which have the advantage of being readily available for immediate coverage of the wound.

A 28-year-old Hispanic engaged woman was admitted with 18 percent total body partial thickness burns to both of her hands, arms, and forearms as a consequence of an accident at work. Her hands had been trapped in the press iron she was working on. Burn staff reported their concerns regarding the functionality outcome of the injury, as it was suspected that her tendons were likely compromised. Consult was called because she was tearful and quiet. When interviewed, she stated that she felt ashamed of being unable to take care of her most basic needs due to the immobilization of her upper extremities. Her spoken English was rudimentary. She was interviewed in Spanish, her native language. The patient tearfully revealed that her deepest fear was to be unable to perform as a future wife and mother due to the potential sequelae from the burns. Remarkably, and after a series of autograft skin transplantations and resolution of edema, she started to slowly move her fingers again. At the time of discharge she concealed the change in appearance of her hands and forearms with the hope of complete functional recovery.

Face Reconstructive Surgery

The human face is a functional, dynamic, and esthetic organ with critical roles in the individual development since birth. There is evidence supporting that newborns have an innate ability to recognize individual human faces shortly after birth. The face is ingrained in the concept of one's identity and comprises a very individualized set of tools that allows the human being to effectively communicate and relate to others. In a patient with facial disfigurement, effective communication is impaired due to self-perception but also as a result of verbal and nonverbal responses of the nondisfigured. Additionally, the face includes basic bodily functions, such as breathing and eating. The optimal result of facial soft tissue reconstruction should achieve aesthetic balance and symmetry, distinct facial planes, skin texture appropriate for camouflage, and facial animation, as well as functionality. The deformity caused by massive facial burns is an indication for the consideration of facial allograft transplantation. Partial face transplantation became a reality in recent years, with the first two partial facial transplants performed in France (2005) and China (2006).

The first total facial transplant took place in Spain (2010). To date, there have been 28 facial transplantations completed in the world. Eight of the facial transplant recipients were burn victims and the others had suffered other types of injuries, ranging from animal attacks to gunshot wounds. Four of the burned patients received their facial transplants in Boston, MA. Facial transplantation mortality is estimated at 11 percent so far.

Facial transplantation requires the presence of a multidisciplinary team composed of plastic and transplant surgeons, psychiatrists, psychologists, ethicists, social workers, radiologists, infectious diseases specialists, and nephrologists. From the pretransplant phase to long-term posttransplant support, psychiatry has a very important role as a part of the facial transplant multidisciplinary team. The pretransplant evaluation identifies the patient's motivations and expectations; obtains the premorbid psychiatric history; describes the patient's psychological strengths, coping style, and self-esteem; determines the patient's body image and identity issues; studies the functionality of the support network; assesses the capacity to consent for the surgery; and attempts to predict the likelihood of adherence to treatment. Additionally, the psychiatrist diagnoses and treats using potential conditions psychopharmacological and psychotherapeutic interventions before, during, and after the face transplantation.

FAMILY SUPPORT

Family members and loved ones of critically injured burn patients face great strain in all phases of treatment. During the acute or initial phase, it is important to provide the family with updates on the patient's condition and evolving treatment plans as information becomes available. If possible, burn unit staff should describe to patient's loved ones any changes in the patient's appearance resulting from burn injury or treatment. This will help them prepare and process the information related to the extent of the burn injury, and the impact of the injury on the patient's functioning in the future.

Patients with extensive burn injury can display disconcerting behaviors related to states of pain or delirium, and they may be agitated, anxious, or confused during family visits. In response, loved ones may make unrealistic demands on burn unit staff. The restrained patient often elicits great concern from family members. Taking the time to listen to their concerns and to describe the reasons for restraints (or one-to-one supervision), the expected course of treatment and recovery, and the likely etiologies of the patient's current states may help to ease their distress. Once the family understands the nature of the injury and its treatment and begins to feel supported and reassured that staff is doing all they can, the family can often help calm the patient. Denial and other defensive mechanisms can be employed by patients and loved ones in the setting of acute and critical

injury, and should be respected as a method of psychological protection and a phase of healing. The presence of mental health staff can be helpful to family members in all phases of recovery. Psychiatric consultants should be present and available and introduced to patients and family members as part of the treatment team. This approach may help to destigmatize mental health involvement on the burn unit.

SELF-INFLICTED BURNS

Self-immolation remains a rare form of self-injury within developed countries. Risk for self-inflicted burns includes the presence of a psychiatric disorder (most often associated with psychotic and depressive states), alcohol and substance abuse, stress, impulsivity, previous suicide attempts (not necessarily by burning), low socioeconomic status, single marital status, and chronic medical illness. Ignition of clothing and use of fluid accelerants are the most frequently employed methods of self-immolation. Most individuals who attempt suicide by burning in the Western world carry a psychiatric diagnosis. Although percentages reported in the literature have varied widely in the past, one recent review (including postmortem subjects from Germany, UK, Denmark, Italy, France, USA, Canada, and Turkey) reporting 84.7 percent of individuals committing suicide by self-immolation had a psychiatric history (Gauthier et al., 2014). The same study reported previous suicide attempts in 93.2 percent of their study population. The authors of a large retrospective study in the United Kingdom reported an alcohol or substance abuse history in 25 percent of patients. Drug and alcohol use may serve as an emotional and behavioral accelerant, effecting poor judgment and impulsivity.

The prevalence of self-immolation increases in some undeveloped countries and within certain cultures, though the incidence can vary within continents, countries and cultures. An association between literal interpretations of religious scripture (often connected with fundamentalist beliefs) and acts of self-immolation has been suggested. Earlier work by Stoddard and coworkers with adolescent patients who attempted suicide by burning led to the proposal of religiosity as a risk factor for self-immolation in children under 18 years of age. In Iran, over 80 percent of female suicides are reported to be by burning and most victims are young, illiterate, impoverished, married women. *Suttee* is practiced by Hindu widows when they throw themselves on their husband's funeral pyre, an act far less common today than in the past. It has been shown that Asians who burn themselves are a decade younger than Europeans.

Survival rates of intentional burns depend on several factors, including availability of advanced burn care (making survival less likely in developing countries), the individual's intention (to die or harm), and the method of burning (use of accelerant). Compared to survivors of unintentional burns, those with self-inflicted burns suffer more severe injury with higher rates of

morbidity and mortality and require long periods of acute hospitalization. Injuries usually affect the upper body, including face, chest, arms, hands, and ventral thighs. The severity of these burns may be due to deliberate use of liquid accelerants on clothing, which keeps the skin in contact with the fire. In addition, the person who intends to die by fire is less likely to extinguish the flame, compared with victims who suffer burns by assault or accident.

Patients who attempt suicide, especially by such unequivocal and aggressive acts, can evoke extreme emotional responses from others, including from medical providers and loved ones, the individuals responsible for the direct care and support of the patient. Some caregivers may even feel it is the right of the patient to die given the extremes to which they have gone. These emotional responses, among others such as anger and anxiety, may affect the care and survival of the patient. Treatment of the psychiatric and emotional needs of the patient, the family, and at times the staff, are imperative.

The process of discharging patients with a recent suicide attempt and severe burns challenges all hospital services involved in arranging placement and aftercare. Inpatient psychiatric units vary in their ability to absorb severely burned patients, given the complex nature of the injury and the need for medical and surgical follow-up care. The creativity and adaptability of the burn unit staff, including psychiatry, may make the difference between a timely discharge and the patient needlessly remaining on the burn unit for months. Inpatient psychiatric units are often reluctant to manage the medical follow-up care of the severely burned patient. Rehabilitation facilities are reluctant to admit patients with complex psychiatric needs and a recent suicide attempt. Establishing secure psychiatric units within rehabilitation centers has been suggested.

Self-inflicted burn injury continues to present a complex medical and psychiatric challenge to all involved and requires a coordinated and multidisciplinary approach. Recent studies acknowledge the paucity of quality research data related to self-inflicted burns and suicide attempts by self-immolation. Larger, controlled, prospective studies are needed to further elucidate factors that might lead to prevention and treatment modalities. There is evidence that, with appropriate treatment, patients with self-inflicted burns can go on to live prodcsuctive lives.

A 46-year-old white, left-handed, domiciled alone, employed, separated, childless woman was admitted to the burn unit with self-inflicted 61 percent TBSA flame burns to her head, neck, upper and lower extremities, thorax, and abdomen. She also suffered inhalation injury.

The initial psychiatric consultation was called 9 days after her admission, when the primary team requested an evaluation of the patient's capacity to withdraw care. The patient's brother, acting as her surrogate, had requested to withdraw life support and refused to sign consents for skin grafting after learning about the likelihood of a prolonged hospital treatment and her guarded prognosis. At that point the patient, despite being intubated and on ventilatory support, was able to communicate by nodding yes or no. Upon further investigation, it was

found that the patient's brother based his decision on his own life experience. Three years earlier, his wife was involved in a serious motor vehicle accident, resulting in respiratory failure and a traumatic brain injury–related major neurocognitive disorder. She had a very complicated hospital course and become ventilator dependent until her death 1.5 years later.

On initial assessment, except for a portion of her face, the patient was completely covered with dressings. Both upper extremities were immobilized and kept at 90-degree angle abduction. Foley catheter and fecal management system were in place. She was on a cardiac monitor, intubated, and receiving IV fluids and medications. She appeared somnolent; however, she consistently made eye contact after hearing her name called. She dozed off intermittently during the interview. She reported her pain being 8/10; nursing staff was made aware and the patient quickly received 1 mg of hydromorphone intravenously as a breakthrough, with good effect. Due to the limitations, most of the interview was accomplished by asking her to answer yes-no questions. The patient was aware of being in a hospital, and she was oriented to the year only. Her vigilance and attention were only mildly impaired. She confirmed that she had set fire to herself with the intention of ending her life. However, she denied wanting to die at the time of the examination. Despite delirium, she consistently and clearly expressed her wish to live. The patient was capable of choosing to continue life support. However, when asked if she was willing to proceed with excisions and skin graftings, the patient offered inconsistent answers. It was recommended to obtain an ethics consultation given the patient's apparent inability to communicate a choice and given that the primary team and the patient's brother had conflicting plans for her. The following day, prior to the ethics consultation, the patient's mental status improved and she seemed to understand the risk and benefit of having versus not having the skin grafting procedures in order to properly treat her injuries and keep her alive. She was found to have capacity to consent for the excisions and skin graftings. The primary team seemed relieved and proceeded with their treatment plan as per her burn care needs.

The primary team reported that since her admission the patient had remained calm and cooperative without periods of agitation. Given her level of immobility it was decided that a one-to-one status was not necessary at the time, and that frequent psychiatric assessments were necessary in order to ensure the patient's safety given the severity of her suicide attempt. Her initial laboratory results showed a urine toxicology negative for illicit drugs, positive for benzodiazepines, and a blood alcohol level of less than 10.0 mg/dL. The patient's brother provided the patient's history and it was found that her ex-husband used to be verbally and physically abusive towards her. They had separated 1 year earlier. Following the separation the patient experienced one episode of increased energy and elevated mood. At that time she had painted her entire house in 4 days. Afterwards her mood progressively became depressed and she started drinking alcohol in increasing amounts and frequency and taking diazepam pills prescribed by her primary care physician for anxiety. One month prior to her suicide attempt she called her brother and told him that she was being observed by video cameras placed in her house. She subsequently took an overdose of unknown number of diazepam pills, and cut her wrists superficially. She became drowsy and lost consciousness. Upon waking up she took herself to the closest emergency room and she was admitted psychiatrically for the first time in her life. Per discharge records obtained from her psychiatric hospitalization, it was found that the patient did not require stitches or advanced treatments related to her overdose status. She was diagnosed with bipolar disorder, alcohol use disorder, and sedative use disorder and discharged on valproic acid and diazepam 8 days later. The patient did not take the valproic acid. The patient's brother also reported that their mother suffered from a serious mental illness seemingly with psychotic features, requiring prolonged psychiatric hospitalizations and leading to major functional impairment and inability to take care of her children.

Once the patient's respiratory failure, rhabdomyolysis, and electrolyte imbalances resolved, her cognitive status and her ability to communicate effectively improved. She was extubated and transitioned to a tent mask, and eventually to a nasal cannula. Her initial delirium completely resolved. She confirmed the information provided by her brother. She started developing depressive symptoms months prior to the separation from her husband. She confided having used cocaine rarely since that time. She stated that the day of her burns, she woke up early in the morning with the pressing thought that she needed to "do something." She lit her first cigarette of the day and decided to take a motorcycle ride. The tank was almost empty. She picked up her emergency gasoline container and instead of filling her tank, poured

it on herself, and lit herself on fire with her cigarette, hoping to die. She denied experiencing hallucinations at the time but acknowledged having contemplated killing herself by burning weeks earlier. She did not have an explanation for choosing this aggressive method of suicide. She regretted her suicide attempts particularly as she increasingly recognized her physical losses and functional impairments. She required multiple excisions, graftings, regraftings, releases and amputations. She lost her left auricle, her first, second, and third left fingers, most of her scalp, and both breasts. She initially reported feeling "ashamed" due to the suffering she caused much to her family and friends. The patient reported her mood as "fine" and her affect ranged from appropriately sad to appropriately bright on good days. At one point she reported being watched by cameras in her room but declined to elaborate further. She received a quetiapine titration, which was helpful for her insomnia. The patient also received lorazepam on an as-needed basis for anxiety, and in the initial phase lorazepam was recommended for potential signs and symptoms of alcohol or benzodiazepine withdrawal. She did not develop abstinence symptoms and took lorazepam very rarely. The patient received supportive psychotherapy approaches and elements of cognitive–behavioral therapy. Although aware of her physical disfigurement, the patient for the most part avoided speaking about it, and instead focused on the progress of her mobility. She found strength in her two cats. Thinking of meeting with them again was a strong motivation for the patient, and so visualization exercises were often used with good results.

The patient did not require a one-to-one level of care at any point during her treatment in the burn unit. She consistently denied suicidal ideation and her behaviors were consistent with her verbalizations. The burn staff expressed concern for the patient on occasions when she was tearful and a team meeting often ensued upon a psychiatric reevaluation the same day. The patient cried when experiencing headaches or breakthrough pain and when feeling frustrated due to discharge plan postponements. She was discharged to an acute physical rehabilitation facility equipped with a staff psychiatrist fully aware of her untapped and unresolved psychological issues and comfortable treating her in that institution.

CHILDREN AND BURN INJURY

Table 27.14–1 lists multiple factors that contribute to the risks of childhood burn injury.

Children account for a large percentage of burn unit admissions around the world, according to the WHO. Nationally, thermal injury is the third leading cause of injury-related child death from ages 1 through 9 and the fourth leading cause from ages 10 through 14. Just over half of all burned children in the US are Hispanic (53 percent vs. Caucasian 31 percent, African American 13 percent, Asian 1 percent, and Other 3 percent) and two-thirds are boys. Age less than 4 confers a higher likelihood that a burn will be fatal when compared with older children. The most common causes of burns in children are fire or flame (58 percent) followed by scalding (26 percent). Scald burns in children under 2 years old represent 28 percent of all scald injuries across all ages and 70 percent of all reported injuries in children. The home remains the most likely place to sustain burn injury at every age. Kitchens and bathrooms are particularly dangerous for young children because of their tendency to explore, coupled with exposure to flame (gas or wood burning stoves and matches), hot liquids (boiling or tap water and hot oil), objects or surfaces.

Table 27.14–1.
Factors That Contribute to the Risks of Childhood Burn Injury

Physical Environment
Structural materials of residence (substandard housing)
Heating mechanisms (space heaters), objects or flame (candles, stoves, bathtub)
Presence of smoke/carbon monoxide detectors/alarms, fire extinguishers (half of all home fire deaths occurred in homes without alarms)
Awareness of fire safety precautions and services, including adequate means of egress
Rural areas
Fall and winter months
Family, Cultural, and Social Structure
Level of supervision of children
Overcrowding
Low socioeconomic status
Traditional practices (e.g., open flame ovens)
Developmental/Gender (Chronological, Emotional, and Intellectual Stage)
Very young children, especially under 4 years old
Boys greater than girls (boys shown to exhibit more risk-taking behaviors and use of flammable materials)
Increased levels of risk-taking behavior and exploration
Epilepsy
Learning disabilities and mental retardation
Psychological/Behavioral Disturbances of Child or Caregiver
Alcohol/substance abuse and smoking
Childhood/adult depression
Violence/abuse
Oppositional defiant disorder, attention deficient/hyperactivity disorder, bipolar disorder

Table 27.14–2.
Typical Signs of Physical Abuse in Burned Children

Bilateral burns to feet, buttocks, or back by dipping in hot water bath
Unburned lower extremity flexion creases, as child reflexively tries to avoid pain
Cigarette-sized, round lesions
Posterior, deep burns with sharp margins from sustained or forced contact with hot object (child prevented from withdrawing)
Illogical, inconsistent explanations of the burn event provided by caregiver
Failure to thrive
Physical numbness to pain
Depression, unresponsiveness, intense fear reaction
Multiple injuries at different stages of healing, including fractures

Child abuse is a major consideration in burn-injured children and must be ruled out during evaluation and treatment of the child. Burning and scalding are some of the most common types of abusive behaviors resulting in death of the child. A WHO literature review and recent research reports (1) that burn injury was involved in approximately 6 to 20 percent of all child abuse cases and (2) that the incidence of child abuse among children hospitalized for burns ranged from 6 to 35 percent. Underreporting of abuse is likely. An adult other than the perpetrator of the abuse often presents the child for treatment. Child victims of abuse are reluctant to present for treatment for fear of legal, physical, or emotional repercussions.

Children generally want to remain with their primary caregivers at all costs. Table 27.14–2 lists the signs of abuse in children who present with burns. Suspicions should also be raised when a young child presents with large tap water burns, care is delayed, and the history and description do not appear to follow the injury pattern.

Psychological Sequelae of Burn Injury in Children

Postburn symptoms of anxiety and depression, behavioral and developmental disturbances, and the deleterious effects on caregivers of burned children have been well documented. Inadequate pain control during lengthy periods of hospitalization, coupled with long periods of separation from primary caregivers, can exacerbate the developing, often fragile, psychological and ego stability of children. A study assessing 101 young adult males and females who survived childhood burns demonstrated an increased prevalence of major psychiatric illnesses (females greater than males), particularly anxiety disorders, which included social anxiety, specific phobias, and PTSD. However burn camp attendees scored significantly lower for anxiety than those who did not attend burn camp.

The psychological development of burned children can be affected by a decrease in their ability to form adaptive behaviors. Adaptability is a necessary emotional tool for adjusting to life's challenges at any age and can ultimately have an effect on maturation, which may in turn affect behavior. Common assumptions exist about the negative effects of disfiguring burns on a child's general attitudes. One study evaluating attitudes of adolescents burned as children contradicted these beliefs and found them to have positive attitudes about their appearance and about how others viewed them, as well as good quality-of-life assessments.

Acute Stress and PTSD Symptoms in Children

Children under age 5 account for just under half of all pediatric burn injuries. A large number of young burned children present with acute stress symptoms during hospitalization. Risk factors for the expression of acute stress symptoms have included size of burn, high pain level, increased pulse rate of the child or the parent, and poor body image of the child. Separate links have been made between (1) burn size, heart rate, and acute stress and (2) level of the child's pain, parental stress, and acute stress symptoms of the child. Two other biologically separate routes have been proposed in the development of PTSD in burned children. One involves separation anxiety as it relates to increased pain levels and the other connects the expression of dissociation with burn size. The importance of these recent findings lies in the possibility of preventing the development of PTSD by detecting and treating early signs and symptoms of acute stress.

Medical and Psychiatric Treatment of Burned Children

Adequate pain control is an ongoing issue during the care of burned children. Studies continue to support that unresolved pain can have negative short- and long-term psychological outcomes. Despite specialized pain management departments, an arsenal of pain medications, and alternative methods of treatment, including hypnosis and guided imagery, too often pain remains difficult to treat, or is undertreated. The treatment of anxiety has been substantially improved with increased efforts to allow the constant presence of caregivers on burn and trauma units, the advent of child life specialists, and administration of medications such as benzodiazepines, selective serotonin reuptake inhibitors, and tricyclic antidepressants. Medications such as benzodiazepines, β-blockers, and oxandrolone can be useful in treating physical manifestations of the hypermetabolic responses to severe burn injury, and may also have positive effects on anxiety and mood. The parental level of coping with the stress of caring for a burned child influences the child's recovery. Treatment of the caregiver may be essential for optimum recovery of the child. Additionally, educating parents about the psychological effects of burn injury on their child will help them to be more attuned to the child's needs and stages of healing. A recent study noted that parents underestimate their burned child's anxiety level. Helping the parent to remain present physically and emotionally with the child will decrease the sense of isolation and separation.

Janey, a preteen girl with 70 percent total body burns, was treated for over 1 year in an acute setting before being transferred to a specialized long-term rehabilitation center. During her acute treatment phase, and as survival become more certain, Janey was more capable of experiencing normal adolescent feelings. She became more attached to staff, expressing her boredom and becoming somewhat stubborn and rebellious. A laptop computer provided an opportunity to connect with the outside world, as she began to communicate regularly via the Internet with her friends. After many months of avoiding her appearance, she also eased her way into looking at her reflection in the window of her burn unit room and later a mirror. Once she had processed her dramatically altered appearance, she began to allow friends to visit.

Developmental Considerations

The rates of premorbid dysfunction are high in burned children and can reach 50 percent. The presence of learning difficulties, emotional and behavioral disturbances, and developmental delay put children at risk for burn injury. Because of the complexities of burn treatment and often lengthy recovery times, children who suffer burn injury can fall significantly behind in their physical and emotional development. Preburn physical, intellectual, and emotional levels will help determine realistic rehabilitation goals as much as will the extent and impact of the burn injury. Health care providers involved in a child's or adolescent's transition to rehabilitation must take premorbid and postinjury developmental status

into consideration when assessing the goals of the child, family, and caregivers during discharge planning.

Burns incurred during childhood can have long-lasting physical and psychological effects. More children than ever before are surviving even large total body burns. Reducing caregiver stress can have direct results on the well-being of the patient, as shown in several studies focused on the effect of stress level in parents of injured or ill children. A child's level of resilience is likely due to some combination of preinjury emotional stability, level of support, family construction, socioeconomic status, temperament or genetics, and certain aspects of after-burn care and rehabilitation.

STAFF SUPPORT

Hospital burn unit staff are repeatedly exposed to severely traumatized bodies and minds. Patients and their families depend on the staff for support, clinical updates on the patient's condition, and glimmers of hope that "everything will be okay." When patient outcomes are positive, these responsibilities can provide staff with great satisfaction and the sense that they have been able to adequately minister to the patient. As on any critical care unit, morbidity and mortality rates can be high. Factors that contribute to burn unit staff burnout and increased levels of staff stress include: (1) high acuity and the labile nature of burn-injured patients' physiological and psychological states; (2) understaffing of the unit or lack of organizational support and guidelines to reduce work stress; (3) physical isolation of the burn unit from other areas of the hospital and avoidance by other services; (4) "secondary traumatization" via repeated visual exposure to the disfigurement and severely burned body parts of patients and by repeated descriptions in rounds of the actual burn event; (5) reactions and demands of traumatized patients and their families; (6) pressure to make life-and-death ethical decisions; and (7) managing patients' ongoing pain.

Many burn units are tucked away in relatively remote areas of the hospital. The isolation of the physical plant may mirror and enhance the psychological isolation felt by the staff when handling critically ill or multiple acute patients. Caring for a burned patient may require repeated visual impressions of maimed, grafted, or amputated body parts, or of the patient wrapped in gauze with only eyes and mouth exposed. Many burned patients are protected from infection by isolation precautions and live behind a glass window. One must gown and glove with each visit. The patient may have suffered or required amputation(s), causing disfigurement and altering their life in many ways. The heavy weight of these burdens can create feelings of helplessness and despair in staff and loved ones as well as the patient. In order to avoid these painful emotions, staff and family may struggle with urges to avoid the patient at any stage of treatment or recovery. In addition, consultants may try to avoid coming to

the burn unit, another factor that contributes to feelings of isolation among burn unit staff.

Studies have shown an increase in mistakes made by caregivers with increased levels of stress. Acute and chronic stress in the workplace leads to staff with lowered morale, impaired performance, increases in turnover, and absenteeism. Training and supervision of intensive care unit nursing staff has been shown to reduce absenteeism and turnover and improve patient care. At particularly high risk of increased stress is burn unit staff in charge of critical decision making, as they attempt to balance the needs of the patient and the unit and implement national guidelines and local standards of care with the resources available to the unit, hospital, and community.

The psychiatric consultant can play an integral role in reducing staff stress by making daily rounds on the burn unit, making recommendations for the treatment of patients' psychological, neurological, and pain issues. Liaison work with families and staff conveys a sense of connection and support.

END OF LIFE AND ETHICS

Consultation liaison psychiatrists are regularly called upon to help evaluate multiple ethical queries arising from the management of burn patients. Patients with massive burns are now more likely to survive than in the past. Those patients can fall into protracted shock or unresponsiveness, and their management raises questions regarding capacity for medical decision making, continuation of treatment, appointment of health care proxy, do-not-resuscitate orders, initiation of comfort measures as opposed to standard treatment, etc.

Autonomy, beneficence, nonmaleficence, and justice constitute four leading principles of biomedical ethics. However, there is no simple algorithm one can follow in order to solve the ethical challenges that can present in the management of the severely burned patient. The patient's exercise of autonomy is usually central to the resolution of ethical dilemmas often posed in their care. Ideally, patients should be asked the type of care that they would wish or not wish to receive at the start of the initial resuscitation, a period in which they are usually lucid. The consulting psychiatrist must evaluate the medical indications for treatment, including: risks, benefits, and potential burdens; the patient preferences based on patient's autonomy; quality-of-life issues; and contextual features, including socioeconomic, family, and legal issues. Patients' surrogates may aid in decision making and must be adequately informed regarding the extent of the burn injury, treatment options, likely outcomes, and prognosis. The hospital's bioethics committee is sometimes involved in the resolution of very complex ethical situations.

An often-arising intricate ethical query is withholding or withdrawal of

care in severely burned patients. Even in cases when this is the patient's wish or when the staffs feel death is inevitable, making this decision can still cause high anxiety in both staff and family members. A "consensus conference" can be used to help the burn team discern the judgment of the group. All involved staff members discuss and resolve their differences of opinion and unify their position prior to presenting to patient or their family. This group decision helps remove the burden of responsibility on individual staff members as well as on family members, who may otherwise bear the burden of making a decision that may ultimately end with the patient's death.

Pain control and management is of paramount importance in all burn patients as this might be an underlying element behind a burn patient's refusal of further treatment, dressing changes, and other related procedures, as well as in terminally ill burn patients where adequate comfort and sensible care should be ensured. A consulting psychiatrist must evaluate for pain control and advocate for optimal pain control in these patients. Comforting words may be of help even in seemingly unresponsive patients. All involved staff should also help patients and their families with anticipatory grief and with concrete plans in order to help alleviate their distress.

Several years ago, at a conference, one of the authors (MB) was on a panel discussing ethical issues in the burn patient. Another guest panelist was a nonphysician professional who had suffered a major burn injury that had severely scarred his face and body, causing blindness. He had subsequently gone on to graduate school, got married, and had children, and reported that he enjoyed having his family very much. However, his point to the panel was that, during his acute hospitalization, he had asked numerous times that his treatment be stopped and that he be allowed to die because he found the treatment intolerable. He believed that it was wrong that his wishes were not respected, and was advocating that this should not be done in the future with other patients in this situation. He was asked by this author whether he had adequate pain control during his treatment and replied, "Definitely not." He also agreed that he might feel differently if there had been better pain relief of his severe pain.

Ethical Considerations of Face Transplantation

The face is one of the most important parts of the human body and plays a central role in the expression and recognition of emotions, identity, and fulfillment of basic functions, such as eating and breathing. Facial burns treated with traditional plastic surgery techniques often lead to poor functional and aesthetic outcomes. Emerging advances in plastic and reconstructive surgery have revolutionized the management of patients suffering from mutilating and disfiguring facial burn injuries. Facial transplantation, since its inception in 2005, is fast gaining popularity as a life-changing and viable treatment option for these patients. Recent reports indicate that between 2005, when the initial facial transplant was reported, and 2014, there have been about 28 facial transplants worldwide.

Advances in immunomodulatory and immunosuppressive protocols,

microsurgical techniques, and computer-aided surgical planning have enabled broader clinical application of facial transplantation to patients. Despite significant advances, facial transplants remain experimental procedures with multiple unresolved ethical questions including the rationale behind the exposure of individuals to the sequelae of lifelong high-dose multidrug immunosuppression, the development of waiting lists for nonvital organ transplantation, the risk of potential immunological rejection of the facial graft with its potential devastating consequences, the consent of the donor's family, the cost to society of such a highly technical procedure, and the potential damage caused to other organs of the donor during the face harvesting period. Results, however, show that from a clinical, technical, and immunological standpoint, facial transplantation has achieved functional, aesthetic, and social rehabilitation in severely facially disfigured patients. Early involvement of a hospital bioethics committee is advocated to ensure compliance with ethical standards in this procedure that requires research in order to establish the risk of experimental interventions.

References

*Amos T, Stein DJ, Ipser JC. Pharmacological interventions for preventing post-traumatic stress disorder (PTSD). *Cochrane Database Syst Rev*. 2014;(7):CD006239.

Anwar MU, Majumder S, Austin O, Phipps A. Smoking, substance abuse, psychiatric history, and burns: trends in adult patients. *J Burn Care Rehabil*. 2005;26(6):493.

*Arno A, Barret JP, Harrison RA, Jeschke MG. Face allotransplantation and burns: a review. *J Burn Care Res*. 2012;33:561.

*Blumenfield M, Schoeps M (eds.). Psychological interventions. *Psychological Care of the Burn and Trauma Patient*. Baltimore, MD: Williams & Wilkins; 1993.

Blumenfield M, Schoeps M. Reintegrating the healed burned adult into society: psychological problems and solutions. *Burn Rehabil Reconstr Clin Plast Surg*. 1992;19(3):599.

Brubaker Rimmer R, Bay C, Alam NB, et al. Burn-injured youth may be at increased risk for long term anxiety disorders. *J Burn Care Res*. 2014;35:154.

Brubaker Rimmer R, Bay RC, Sadler IJ, Alam NB, Foster KN, Caruso DM. Parent vs. burn-injured child self-report: contributions to a better understanding of anxiety levels. *J Burn Care Res*. 2014;35:296.

Cukor J, Wyka K, Leahy N, Yurt R, Difede J. The treatment of post-traumatic stress disorder and related psychosocial consequences of burn injury: a pilot study. *J Burn Care Res*. 2014;36(1):184.

Davis CS, Esposito TJ, Palladino AG, et al. Implications of alcohol intoxication at the time of burn and smoke inhalation injury: an epidemiologic and clinical analysis. *J Burn Care Res*. 2013;34(1):120.

Davydow DS, Katon WJ, Zatzick DF. Psychiatric morbidity and functional impairments in survivors of burns, traumatic injuries and ICU stays from other critical illnesses: a review of the literature. *Int Rev Psychiatry*. 2009;21:53.

Faucher L, Furukawa K. Practice guidelines for the management of pain. *J Burn Care Res*. 2006;(5):659.

Fratianne RB, Brandt C, Yurko L, Coffee T. When enough is enough? Ethical dilemmas on the burn unit. *J Burn Care Rehabil*. 1992;13(5):600.

Gauthier S, Reisch T, Bartsch C. Self-burning—a rare suicide method in Switzerland and other industrialised nations—a review. *Burns*. 2014;40(8):1720–1726.

*Hahn AP, Jochai D, Coufield-Noll CP, et al. Self-inflicted burns: a systematic review of the literature. *J Burn Care Res*. 2014;35:102.

Khalifian S, Brazio PS, Mohan R, et al. Facial transplantation: the first 9 years. *Lancet*. 2014;384:2153.

Kung S, Espinel Z, Lapid MI. Treatment of nightmares with prazosin: a systematic review. *Mayo Clinic Proc*. 2012;87(9):890.

Lawrence JW, Fauerbach JA. Personality, coping, chronic stress, social support and PTSD symptoms among adult burn survivors: a path analysis. *J Burn Care Rehabil*. 2003;24(1):63.

Malic CC, Karoo ROS, Austin O, Phipps A. Burns inflicted by self or by others—an 11-year snapshot. *Burns*. 2007;33(1):92.

Miller PL. Application of a model to guide ethical decision making in burn treatment. *J Burn Care Rehabil*. 2000;21(5):433.

Mock C, Peck M, Peden M, Krug E. *A WHO plan for burn prevention and care*. Geneva: World Health Organization; 2008.

Morris P, Bradley A, Doyal L. Face transplantation: a review of the technical, immunological, psychological and clinical issues with recommendations for good practice. *Transplantation*. 2007;83:109.

*National burn repository 2014 Report, dataset version 2.0. Available at http://www.ameriburn.org/2014NBRAnnualReport.pdf. Accessed February 2015.

National Trauma Institute. Trauma statistics. 2014. Available at http://www.nationaltraumainstitute.org/home/trauma_statistics.html. Accessed on December 2014.

Pope SJ, Solomons WR, Done DJ, Cohn N, Possamai AM. Body image, mood and quality of life in young burn survivors. *Burns*. 2007;33(6):747.

Ratcliff SL, Meyer WJ, Cuervo LJ, Villarreal C, Thomas CR. The use of haloperidol and associated complications in the agitated, acutely ill pediatric burn patient. *J Burn Care Rehabil*. 2004;25:472.

Rosenberg L, Blakeney P, Thomas CR, Holzer CE, Robert RS. The importance of family environment for young adults burned during childhood. *Burns*. 2007;33(5):541.

Saxe GN, Stoddard F, Hall E, Chawla N, Lopez C. Pathways to PTSD, part I: children with burns. *Am J Psychiatry*. 2005;162:1299.

Stoddard FJ Jr, Fricchione GL. Burn patients. In: Stern TA, Fricchione GL, Cassem NH, eds. *Massachusetts General Hospital Handbook of General Hospital Psychiatry*. 5th ed. Philadelphia, PA: Mosby; 2004.

*Stoddard FJ, Levine JB, Lund KK. Burn injuries. In: Blumenfield M, Strain J, eds. *Psychosomatic Medicine*. Philadelphia, PA: Lippincott Williams & Wilkins; 2006:603.

Stoddard F, Saxe G, Ronfeldt H, Drake J, Burns J. Acute stress symptoms in young children with

burns. *J Am Acad Child Adolesc Psychiatry*. 2006;45(1):87.

Suhrabi Z, Delpisheh A, Taghinejad H. Tragedy of women's self-immolation in Iran and developing communities: a review. *Int J Burns Trauma*. 2012;2(2):93.

Tcheung W, Robert R, Rosenberg L, Rosenberg M, Villarreal C. Early treatment of acute stress disorder in children with major burn injury. *Pediatr Crit Care Med*. 2005;6(6):676.

Waldmann CS. Suffering and coping: the challenges for staff. In: Ridley S, ed. *Critical Care Focus: The Psychological Challenges of Intensive Care, 12*. Boston, MA: Blackwell; 2005.

Wibbenmeyer L, Liao J, Heard J, Kealey L, Kealey G, Oral R. Factors related to child maltreatment in children presenting with burn injuries. *J Burn Care Res*. 2014;35:374.

World Health Organization. Burns. *Fact Sheet No. 365*. Updated April 2014.

▲ 27.15 Management of Chronic Pain

J. GREGORY HOBELMANN, M.D., AND MICHAEL R. CLARK, M.D.

DEFINITION AND COMPARATIVE NOSOLOGY

Pain is a complex experience integrating cognitive, behavioral, and affective factors serving as a warning of impending or actual injury. The International Association for the Study of Pain defines pain as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage." This definition recognizes that pain may be experienced in the absence of tissue damage and is, therefore, a subjective experience. As such, the patient's report of pain should be considered valid in the absence of evidence to the contrary.

Table 27.15–1.
Types of Painful Experiences

Allodynia	A painful response to a normally non-noxious stimulus
Deafferentation pain	Pain resulting from loss of sensory input into the central nervous system
Dysesthesia	An abnormal pain that can be spontaneous or evoked
Hyperalgesia	An exaggerated painful response to a normally noxious stimulus
Hyperesthesia	Increased sensitivity to stimulation that excludes the special senses
Hyperpathia	An exaggerated painful response evoked by a noxious or non-noxious stimulus
Hypoesthesia	Diminished sensitivity to stimulation that excludes the special senses
Nociception	Detection of tissue damage by transducers in skin and deeper structures and the central propagation of this information via A delta and C fibers in the peripheral nerves
Paresthesia	An abnormal nonpainful sensation that is either spontaneous or evoked
Sensitization	Lowered threshold and prolonged/enhanced response to stimulation

There are many terms used to describe different types of painful experiences (Table 27.15–1). Pain can be either acute or chronic. Acute pain

is typically the result of actual tissue damage secondary to injury, surgery, or the exacerbation of chronic disease. It has survival value in that it promotes behaviors that minimize repeat injury. If pain persists on a daily basis for a month beyond what would be considered necessary to heal the underlying injury, it is generally considered chronic. Chronic pain seems to serve no purpose and is much more complex and difficult to manage than acute pain. The intensity of pain can be classified as mild, moderate, or severe and is usually measured on a 10-point numerical rating scale.

Pain is further classified as either nociceptive or neuropathic (Table 27.15–2). Nociceptive pain results from injury or disease and can be divided into somatic or visceral pain. Somatic pain is usually well localized and results from injury or disease of the skin, muscle, bone or related structures. Examples of somatic pain conditions are arthritis, fibromyalgia, and myofascial pain. Visceral pain arises from internal organ dysfunction as the result of inflammation, ischemia, or organ distention and is usually poorly localized and difficult to characterize. Examples of visceral pain conditions are endometriosis, appendicitis, and interstitial cystitis. The quality of nociceptive pain is often described as throbbing, aching, or pressure-like pain.

Neuropathic pain is pain arising from a lesion or disease of the somatosensory system and is often accompanied by maladaptive changes in the nervous system. The distinction is made between peripheral, as in peripheral neuropathy and postherpetic neuralgia, or central, as in multiple sclerosis and central poststroke pain. The quality of neuropathic pain is usually described as shooting, stabbing, or electric-like pain. Conditions such as headache and cancer pain can involve both nociceptive and neuropathic pain. Over time, all types of chronic pain produce more difficult to reverse changes in functioning of the CNS as well as the structure of its components termed central sensitization.

Table 27.15–2.
Neuropathic and Nociceptive Pain Conditions

NEUROPATHIC	
Peripheral	**Central**
Peripheral neuropathy	Spinal cord injury
Postherpetic neuralgia	Central poststroke pain
Trigeminal neuralgia	Parkinson disease
Chronic postsurgical pain	Multiple sclerosis
Phantom limb pain	Seizure disorder
Complex regional pain syndrome	
Radiculopathy	
Nerve entrapment	
NOCICEPTIVE	
Somatic	**Visceral**
Arthritis	Irritable bowel syndrome

Myofascial pain	Endometriosis
Fibromyalgia	Interstitial cystitis
Connective tissue disorders	Ulcers
Surgical/trauma related pain	Appendicitis
Burn pain	Cholecystitis

EPIDEMIOLOGY

Pain is the number one reason that patients present to their physician, and the societal burden of pain is considerable. In 2008, about 100 million adults in the United States reported a chronic pain condition. Chronic back pain is the most common condition, followed by severe headaches, arthralgias and neck pain. Patients with chronic pain utilize about twice as many health care resources as the general population. People with moderate to severe pain lose 8 days of work every 6 months on average. The total health care costs secondary to pain ranged from $560 billion to $635 billion in 2010, eclipsing the cost of heart disease, diabetes or cancer.

Many factors here been linked to the development of chronic pain. According to the available literature, the effect of age is inconsistent, with some studies reporting an increase in prevalence of chronic pain with age and others reporting a decrease. Most studies report that women present with more severe, frequent, and longer lasting pain than men. There appears to be in inverse relationship between socioeconomic status and the prevalence of chronic pain, with a higher prevalence in patients with lower income, lower levels of education, and who are unemployed. The most reliable factors for the development of chronic pain are a history of a previous episode of pain and more severe intensity of acute pain.

ETIOLOGY AND PATHOPHYSIOLOGY

The neurobiology of pain is complex and involves peripheral, dorsal horn, ascending tract, and descending inhibition mechanisms. Pain results from the activation of peripheral receptors, or nociceptors, by a stimulus. Ongoing stimulation causes changes in peripheral nerves, spinal cord structures and supraspinal structures, which can result in the development of chronic pain.

Peripheral mechanisms begin with primary afferent nociceptors responding to a mechanical, chemical, or thermal stimulus. Myelinated A-delta fibers transmit mechanothermal information with a sharp, prickling quality. Unmyelinated C-fibers are polymodal and transmit pain with a burning, aching quality. Usually, stimulation activates only high-threshold nociceptors. In conditions of nerve injury or inflammation, peptides such as substance P and neurokinin A are released, causing neurogenic inflammation. This results in nerve fibers becoming more excitable, vascular structures dilating with extravasation of plasma proteins, and inflammatory cells releasing a variety of chemical mediators. These

chemicals can alter the response of the high-threshold nociceptors by reducing the response threshold called peripheral sensitization. In peripheral sensitization, there is a zone of allodynia around the site of injury in which low-intensity stimulation can produce pain.

The primary afferent nociceptors terminate in the dorsal horn of the spinal cord. The dorsal horn is the site of action of many neurotransmitters and neuromodulators that act at several receptors including *n*-methyl-D-aspartate (NMDA), kainite, glutamate, opioid, and GABA, among others. The second order neurons can be classified into nociceptive- specific/high-threshold or wide dynamic range neurons. Allodynia occurs when wide dynamic range neurons become sensitized and fire at an increased frequency. Central sensitization is a state of increased sensitivity to central pain signaling neurons and also occurs in the dorsal horn. It refers to several modifications resulting from prolonged activation of nociceptive afferent fibers. These modifications include windup (progressive increase in neuronal activity during stimulation), facilitation (magnification and prolongation of the neuronal response), action potential threshold reduction, receptive field expansion, oncogenic induction, and long-term potentiation (strengthening of the synaptic transmission after stimulation). NMDA receptor activation by glutamate is thought to be an important mechanism for the production of central sensitization.

Second-order neurons project to supraspinal structures in the ascending tracts of the spinal cord. These ascending tracts are located in the contralateral anterolateral spinal cord (spinothalamic, spinoreticular, spinomesencephalic) and latent ipsilateral pathways. The ventroposterior nuclei of the thalamus represent temporal and spatial aspects of pain, while the medial nuclei are involved in the affective-motivational features of pain. Other subcortical structures, such as the basil ganglia, hypothalamus, amygdala, and cerebellum are thought to function in the transmission of nociception and perception of pain. The role of cortical structures in pain perception is not well understood, but it is postulated that the parietal lobes and somatosensory cortex contribute to the sensory-discriminative component while the cingulate cortex is involved with the affect component.

Pain can be decreased by descending inhibition. This is a complex mechanism involving many structures and neurotransmitters throughout the central nervous system. Treatment modalities involving electrical stimulation attempt to activate descending inhibition to reduce the activation of nociceptive mechanisms.

DIAGNOSIS AND CLINICAL FEATURES

Pain is a subjective experience and is difficult to assess, especially in people with cognitive limitations or other chronic degenerative diseases. Complaints of pain should always be taken seriously. There are no

diagnostic tests that can ascertain the existence of pain or measure the severity of pain. The evaluation of pain should always be individualized and begins with a thorough history and physical examination. A comprehensive history should include the quality and location of the pain, associated signs and symptoms, exacerbating and alleviating factors, temporal aspects, interference with daily functioning, and responses to prior treatments. Acute pain is usually the result of an inciting insult such as trauma or surgery, but the etiology of chronic pain is usually more insidious and often the etiology of chronic pain cannot be determined. As a result, chronic pain conditions result in more psychological sequelae than do acute pain conditions.

The severity of pain can be assessed using a variety of rating scales including verbal, numerical scales from 0 to 10 and the visual analog scale (VAS). The VAS consists of a 10-cm line with “no pain” on the left and “worst possible pain” on the right. For consistency, the same scale should be used over time in individual patients. For patients that have difficulty using these scales, such as young children or cognitively limited patients, substitutes such as the facial expression scale have been validated. In addition to severity, other important aspects of pain should be assessed including functional capacity, psychological and emotional function, satisfaction ratings and adverse treatment effects. Distinguishing between neuropathic and nociceptive pain is also important as treatment approaches to each vary considerably.

A psychological history must be included in the evaluation as well. A focused exam should include prior psychiatric diagnoses, work history, legal history, substance abuse history and prior or current suicidality. Up to two-thirds of chronic pain patients have comorbid psychopathology. Depression is the most common comorbidity, followed by anxiety disorders, somatoform disorders, and substance use disorders. It is important to elicit these disorders and treat them concurrently, as they have been associated with poorer treatment prognosis. In addition, people seeking care for chronic pain may also be more likely to have personality vulnerabilities which can act as a barrier to effective treatment.

Physical examination of the patient should include all body systems because pain is often a manifestation of systemic disease. However, a focal neurologic and musculoskeletal examination may be appropriate in patients with limited complaints. It is important to evaluate for sensory changes, motor changes, passive and active range of motion, and postural and gait abnormalities. Acute pain is typically associated with autonomic activation (altered vital signs or facial grimacing), but chronic pain is usually not. The examination findings are rarely pathognomonic, but rather serve to select patients for further imaging studies or invasive diagnostic procedures to focus on a specific etiology.

Pathology and Imaging Studies

Various imagining techniques have become commonly utilized in the search for the cause of chronic pain. However, they have several drawbacks and are limited in their utility. The correlation between MRI findings and severity of pain is poor. Greater than 50 percent of asymptomatic patients will have abnormalities on lumbar, thoracic, or cervical films and early imaging for back pain does not appear to improve outcomes or affect decision making. MRIs are indicated in patients with serious or progressive neurologic deficits, bowel or bladder dysfunction, suspected metastatic disease, or infection.

For nonspinal conditions, MRI is best for inflammation and soft tissue disease and CT scanning is better for detecting bone disease and bleeding. Plain films can be useful to detect gross bone abnormalities and joint degeneration. Ultrasound is infrequently used but has the advantages of being inexpensive and does not expose patients to radiation. In most instances, imaging serves to alert the practitioner to seek further invasive diagnostic procedures.

Electromyography and nerve conduction studies (NCS) can be helpful in detecting lesions of large nerves. They do not detect impairment of smaller nerves, however, and are associated with significant false negative and false positive rates. For peripheral neuropathies, skin biopsies are sensitive tests that demonstrate decreased density of epidermal nerve fibers.

Recent studies have shown that the brains of people with chronic pain are different from those of people without pain complaints. There are morphologic alterations such as decreases in the gray matter in the areas involved in the perception of pain. This appears to be a result of the pain and can be reversed with successful treatment. MR-based brain imaging methods are developing rapidly and will likely be utilized more in the diagnosis and treatment of pain in the future.

DIFFERENTIAL DIAGNOSIS

Acute pain is usually the result of an injury or trauma and determining the underlying cause is straightforward. Chronic pain, however, often has no identifiable etiology, and management must be focused on the suppression of symptoms. The more common chronic pain syndromes (Table 27.15–3) are discussed in the next section.

Low Back Pain

Low back pain (LBP) is one of the most common physical complaints and the most expensive condition when health care costs and loss of productivity are considered. Psychological factors such as depressed mood and somatization are highly correlated with LBP and predict the transition

from acute to chronic pain. The most powerful predictor of chronicity is poor functional status 1 month after seeking treatment. Both economic and social rewards have been associated with higher levels of disability and depression in patients with chronic LBP.

The presence of a depressive disorder increases the risk of developing chronic LBP and musculoskeletal pain in general. Depression has been associated with a nearly fourfold increase in the likelihood of seeking treatment for new onset LBP. On the other hand, chronic disabling LBP can result in significant psychopathology that affects outcome and treatment response.

Chronic LBP is a condition that exemplifies the complexity of chronic pain treatment. Patients with chronic LBP present with symptoms representing numerous diagnoses that require multimodal treatment approaches. Treatments can often reduce symptoms, but there is conflicting evidence about their ability to improve functional status and return patients to work. Almost all treatment modalities, including conservative interventions, have produced inconsistent results. A thorough evaluation and individualized treatment plan for each patient presenting with LBP is essential.

Peripheral Neuropathic Pain

Peripheral neuropathy is the result of damage to peripheral nerves and often causes pain, numbness and weakness. It usually occurs in the hands and feet, but can occur in other areas of the body. There are numerous etiologies of peripheral neuropathy, including diabetes mellitus and excessive alcohol consumption. Approximately 25 percent of patients with diabetes will experience painful diabetic neuropathy, and longer duration of disease and poor glycemic control increase the risk. The pain from peripheral neuropathy is often described as burning and lancinating and can be constant or episodic as well as spontaneous or evoked. The pain is primarily the result of axonal degeneration, segmental demyelination, and sites of ectopic impulse generation, and can occur anywhere along the peripheral nerve up to the dorsal root ganglion. Voltage-dependent sodium channels contribute to hyperexcitability, and central sensitization amplifies and sustains pain by a variety of mechanisms. Pharmacologic treatment can help to reduce symptoms, but reversal of the pathology is rarely attainable.

Postherpetic Neuralgia

Postherpetic neuralgia (PHN) is responsible for about 15 percent of referrals to pain clinics. It is defined as pain that persists or recurs at the site of shingles at least 3 months after the onset of an acute varicella zoster rash. PHN occurs in about 10 percent of patients with acute herpes zoster

and in over 50 percent of patients over the age of 65. Risk factors for the development of PHN include advanced age, cancer, diabetes mellitus, immunosuppression, longer duration of prodromal symptoms, greater pain and severity of rash, sensory impairment, and psychological stress. Although it generally improves over time, 25 percent of patients with PHN continue to complain of pain for longer than 1 year.

Table 27.15–3.
Common Chronic Pain Conditions

HEADACHE

- Migraine
- Tension-type headache
- Cluster headache
- Headache associated with trauma
- Chronic daily headache
- Cervicogenic headache
- Postdural puncture headache

OROFACIAL PAIN

- Trigeminal neuralgia
- Geniculate neuralgia (Ramsay Hunt Syndrome)
- Glossopharyngeal neuralgia
- Superior laryngeal neuralgia
- Tolosa–Hunt syndrome

LOW BACK PAIN

- Spondylosis
- Spondylolisthesis
- Spondylolysis
- Disc herniation
- Foraminal stenosis
- Facet arthropathy
- Spinal stenosis
- Fracture
- Musculoligamentous
- Discogenic pain
- Congenital disease

BUTTOCK PAIN

- Sacroiliac dysfunction
- Piriformis syndrome

PERIPHERAL NEUROPATHIES

- Painful peripheral neuropathies (e.g., diabetes)
- Focal entrapment neuropathies
- Postsurgical syndromes
- Traumatic brachial plexus avulsion
- Ischemic cerebrovascular injury
- Syringomyelia
- Arachnoiditis
- Complex regional pain syndrome
- Meningoradiculopathies
- Epidural spinal cord compression
- Herpes zoster and postherpetic neuralgia

PELVIC PAIN

- ► Endometriosis
- ► Infection
- ► Uterine pain
- ► Pelvic congestion
- ► Pelvic adhesions
- ► Ovarian remnant syndrome
- ► Cancer pain

CENTRAL PAIN STATES

- ► Spinal cord injury pain
- ► Poststroke pain
- ► Parkinson disease
- ► Multiple sclerosis

FIBROMYALGIA
MYOFASCIAL PAIN SYNDROME
PHANTOM PAIN
SICKLE CELL ANEMIA

Degeneration and destruction of motor and sensory fibers of the dorsal root ganglion characterize acute herpes zoster. In PHN, further neurologic damage, including inflammation of the spinal cord, myelin disruption, axonal damage, and a decrease in the number of nerve endings in the affected area, have been demonstrated. Studies have suggested the role of both peripheral and central sensitization in PHN. Early detection and treatment of herpes zoster with antiviral medication and analgesics may have potential for the prevention of PHN. In addition, a live attenuated virus vaccination for the prevention of herpes zoster is now available and has reduced the burden of illness in adults older than 60 years.

Migraine and Chronic Daily Headache

The lifetime prevalence of migraine is 18 percent in women and 6 percent in men. The peak incidence occurs between the third and sixth decades of life. Common migraine is unilateral and pulsatile and may be accompanied by nausea, vomiting, photophobia and phonophobia. Prodromal symptoms, such as scintillating scotomata, are typically present in the classic form of migraine. Complicated migraine includes focal neurologic symptoms such as cranial nerve palsies or other neurologic deficits, and is often described by the deficit, such as hemiplegic or basilar migraine. The definitive pathogenesis has not been described, but several theories exist involving the trigeminovascular system and protein extravasation, membrane-stabilizing effects through action at voltage-sensitive calcium channels, and increased levels of substance P. Management of migraine starts with the identification and control of comorbidities and risk factors that predispose individuals to migraines. Beta blockers, flunarizine, valproate, topiramate, and CBT appear to be the best prophylactic agents and several studies have demonstrated the efficacy of nonsteroidal anti-

inflammatory medications and triptans in the treatment of acute attacks.

The most common reason for lost productivity in the US workforce is headache and about 5 percent of the population suffers from chronic daily headaches. Common headache conditions include constant (transformed) migraine, chronic tension-type headache, medication overuse headache, and hemicranias continua. Patients with chronic daily headaches are more likely to have psychiatric comorbidities such as depression and anxiety and report functional disability. They also tend to overuse analgesic medications leading to rebound headaches. Patients with chronic daily headaches are a difficult group to treat and studies suggest that treatment with medications and CBT together is better than either treatment alone.

Central Poststroke Pain and Spinal Cord Injury

Lesions of the central nervous system, such as stroke or spinal cord injury (SCI), are commonly associated with pain. Symptoms of central pain are generally poorly localized and vary over time. Patients may experience allodynia, hyperalgesia, muscle and visceral pain, and lancinating pain. The pain can be described as burning, aching, lacerating, or pricking but tends to vary widely among patients. The pain may not occur until 1 month or longer after the initial insult. There are likely many processes involved in the generation of the pain and excitatory amino acids have been implicated in the generation of central sensitization after a stroke or SCI.

Central poststroke pain and SCI pain can be severe and disabling, but are commonly underreported and often overshadowed by neurologic complications such as weakness, spasticity, and aphasia. Studies indicate that treatment with analgesics, adjuvant medications, and in some cases electrical stimulation can significantly reduce pain in these conditions. The studies have been small, however, and more research is needed to determine which treatments are most efficacious.

Fibromyalgia

Fibromyalgia is characterized by widespread musculoskeletal pain in all four extremities as well as the trunk, exaggerated tenderness, and stiffness. It is diagnosed in 2 to 8 percent of the population, but this may be an overestimation as criteria for the condition have been debated Fibromyalgia is characterized by the manifestations of central sensitization of nociception with concomitant fatigue, poor sleep, cognitive difficulties, and depression. Musculoskeletal pain is the most prominent feature of fibromyalgia, but it should be suspected in any patient who has multifocal pain that is not explained by a definable etiology. Because pain is amplified and can occur anywhere in the body, sensory hyperresponsiveness, chronic headaches, somatic pain, and visceral pain are quite common. Studies suggest that muscle relaxants, membrane stabilizers, selective

serotonin/norepinephrine reuptake inhibitors, and tramadol are effective in reducing pain associated with fibromyalgia.

Complex Regional Pain Syndrome

Complex regional pain syndrome (CRPS) was formally called reflex sympathetic dystrophy (RSD) and causalgia. It is described as an array of painful conditions characterized by continued pain that is precipitated by trauma or immobilization that is usually associated with allodynia or hyperalgesia. Approximately 3 to 5 percent of patients who have had a distal radial fracture subsequently develop CRPS. The pain is regional but is not limited to a specific dermatome or peripheral nerve. In acute stages, inflammatory symptoms prevail. In chronic stages, symptoms can include hyperalgesia, sensory loss, motor changes, sudomotor dysfunction, edema, blood flow abnormalities, perception disturbances, and complications of disuse and physical deactivation. These changes are due to central neuroplasticity and likely involve neurogenic inflammation and neuropeptides such as calcitonin gene-related peptide, cytokines, nerve growth factor, and substance P. Sympathetically maintained pain is often present and may result from coupling of the sympathetic and sensory neurons. Patients with sympathetically maintained pain can receive temporary relief from sympathetic blockade. Mood, anxiety, and substance abuse disorders are fairly common in patients with CRPS. These disorders are generally considered a consequence, rather than a cause, of the condition.

Pharmacotherapy for CRPS has limited efficacy, but there is evidence supporting the use of anti-inflammatory drugs, gabapentin, pregabalin, carbamazepine, tricyclic antidepressants, and opioids. Physical therapy, graded motor imagery, and graded pain exposure can help to overcome central reorganization. If a psychiatric comorbidity is present, the patient should be referred for specialized treatment. Invasive treatment, such as electrical stimulation and surgical sympathectomy, should only be offered after psychosomatic assessment and careful patient selection. The prognosis for recovery from CRPS is now thought to be better than was historically assumed if treatments facilitate the rehabilitation of the affected limb with activated movement and use.

Phantom Limb Pain

Phantom limb pain is a type of neuropathic pain that originates in the distal territory of an amputated limb. It can also occur after the surgical removal of other body parts, such as breast, rectum, penis, eye, or tongue. It occurs in 50 to 80 percent of amputees and is generally described as stabbing, throbbing, aching, burning, or cramping. While studied extensively and easy to diagnose, phantom pain can be challenging to treat.

Pharmacologic treatments and neural blockade seem to have limited efficacy, but morphine, calcitonin, and mexiletine have been shown to reduce pain lasting more than 6 months suggesting a role for the usual pharmacologic treatments of neuropathic pain and central sensitization. The limited success with pharmacologic treatments has spawned interest in psychological interventions, such as CBT and biofeedback, for the treatment of phantom limb pain. The limited available evidence demonstrating the effectiveness of these treatments has shown various degrees of success.

Orofacial Pain

Trigeminal neuralgia, or tic douloureux, is a chronic pain condition with severe, paroxysmal, recurrent, lancinating pain in the distribution of the facial nerve (cranial nerve V). It is usually unilateral and involves the mandibular or maxillary branch of the nerve. Episodes of pain are often spontaneous but can be evoked by nonpainful stimuli or environmental conditions. Motor and sensory deficits are not typically present. Uncontrolled pain with frequent or severe attacks increases the risk of developing anxiety, depression, insomnia, weight loss, and social withdrawal.

In most cases of trigeminal neuralgia, there is evidence of trigeminal nerve root compression and demyelination by blood vessels, mass lesions, or other diseases such as multiple sclerosis or herpes zoster. Compression and demyelination causes hyperactivity of the trigeminal nucleus. Carbamazepine is the first-line treatment and there is evidence to support the use of oxcarbazepine and lamotrigine. If pharmacologic approaches fail, there are several invasive procedures such as microvascular decompression, percutaneous gangliolysis, and stereotactic radiosurgery that can be considered.

Temporomandibular disorder (TMD) refers to conditions that involve pain from the temporomandibular joint, muscles of mastication, or other orofacial musculoskeletal structures. Most commonly, pain arises from the muscles of mastication and is precipitated by movement of the jaw when chewing or talking. The pain is generally sharp, sudden, intense, and localized to the preauricular area, and joint sounds such as “popping” or “clicking” are common. Other symptoms include feelings of muscle fatigue, weakness, tightness, and changes in bite or ability to open or close the jaw. TMD is often associated with psychological stress, which is responsive to behavioral treatments. Treatment of patients with TMD should initially be based on the use of conservative and evidence-based therapeutic modalities, with more invasive interventions, such as surgery, being reserved for refractory pain.

Burning mouth syndrome (BMS) is an array of conditions characterized by pain in the oral and pharyngeal cavities that is often associated with

dryness and taste alterations. Middle-aged and postmenopausal women are most frequently affected. Although most cases are idiopathic, it can be associated with other conditions such as bruxism, poorly fitting dentures, oral candidiasis, hypothyroidism, diabetes mellitus, and neoplasm. Psychiatric comorbidities are fairly common and include anxiety, depression, traumatic life events, and somatization. Treatment begins with identifying potential underlying etiologies such as anxiety and depression, nutritional deficiencies (iron, folate, and B-vitamins), maladaptive oral habits, and iatrogenic causes. Symptomatic treatments include tricyclic antidepressants, anticonvulsants, soft desensitizing oral appliances, vitamins, hormonal supplements, and habit awareness counseling.

There are a number of less common orofacial pain conditions. These include cluster headache, short-lasting, unilateral, neuralgia-form pain with conjunctival injection and tearing (SUNCT), Tolosa–Hunt syndrome, and atypical facial pain.

Parkinson Disease

The majority of patients with Parkinson disease have pain. The pain can be musculoskeletal, neuropathic, or dystonic and is typically described as cramping and aching in the low back and extremities. It is more common in women. The pain is often reduced when the patient is treated with levodopa. This suggests a central origin that may be produced by the lack of dopaminergic input. Dopamine is now recognized as having a role in the endogenous pain modulating system. In addition to levodopa, analgesics and adjuvant medications can significantly reduce the pain associated with Parkinson disease. Deep brain stimulation is also an option when symptoms are refractory to other treatments.

COURSE AND PROGNOSIS

Pain is a complex perceptual experience that is not only affected by the activation of nociceptors, but also emotions, cognition, expectations, and social and cultural context. In acute pain, the course and prognosis is fairly predictable with resolution of pain when the initial insult has resolved. In chronic pain, however, the course and prognosis varies widely and is difficult to predict. Certainly, psychiatric comorbidities and their treatments influence the course and prognosis greatly. Below are some of the psychiatric comorbidities that are frequently associated with chronic pain conditions.

Somatic Symptom Disorder

When a somatic cause for pain cannot be identified, practitioners often begin to look for psychological causes. Pain caused by emotional factors was first introduced in the DSM-II under psychophysiological disorders

and has gone through several modifications in subsequent editions. The DSM-5 has a new category name, Somatic Symptom and Related Disorders, and its prototype is the somatic symptom disorder (SSD). The criteria for SSD have been modified substantially from its predecessor, somatoform pain disorder (SPD). The most radical change is the abolition of the distinction between medically explained and medically unexplained somatic complaints. This will likely increase considerably the number of patients with SSD compared to somatoform pain disorder. There is little research on SSD, as it was just introduced into the DSM-5, so this section will focus on SPD.

In SPD, pain is the chief complaint, but the experience of pain is increased by psychological factors. It is often associated with "psychogenic" pain with no "real" cause, but this is a misconception. Neuroimaging studies have demonstrated decreases in gray matter in the areas of the cortex that modulate the perception of pain. It is important to appreciate how the psychological and physical domains interrelate with each other rather than conceptualizing them as "either/or" concepts.

Despite the fact that SPD is typically associated with multiple pain complaints, it is uncommon in patients with chronic pain. Subsyndromal forms of somatization disorder, such as multisomatoform disorder, are more common in patients with multiple unexplained painful symptoms. These patients are more likely to think that physicians believe their pain is imaginary, engage in catastrophic thinking, express feelings of loss of control and believe the cause of their pain to be from a mysterious disease. They are also at risk of having excessive diagnostic tests performed, inappropriate medications prescribed and undergoing unnecessary surgery. Commonly, somatoform disorders and affective disorders overlap and patients with somatization experience functional disability and role impairment independent of psychiatric and medical comorbidity. The modifications in the DSM-5 reflect the concept that somatization is more likely a dimensional process based on somatic distress and care-seeking behavior that is persistent across the entire population rather than affecting only a small subset of people.

Substance Use Disorder

The prevalence of severe substance use disorder (formerly substance dependence) in patients with a chronic pain condition is estimated to be 7 to 20 percent. The essential criteria to note for substance use disorder in patients with chronic pain are continued use of a substance despite negative consequences, loss of control of substance use, and excessive preoccupation with using the substance. Prescription medication abuse has dramatically increased over the last two decades. Americans account for only 5 percent of the world's population but consume 80 percent of the global opioid supply.

Aberrant medication-taking behaviors are often assumed to be abuse or addiction. However, chronic, persistent pain can lead to an increased focus on taking opioid analgesics. Understandably, patients fear the reemergence of pain or withdrawal if they run out of medication and often take measures to assure that they have an adequate supply. Furthermore, medication-seeking behavior may be the result of a patient trying to maintain previous levels of pain control or improve pain control of a partial, but inadequate, dose of analgesic. This can lead to pseudoaddiction behaviors that may mimic those of a patient with addiction, but which result from therapeutic dependence and true or potential undertreatment of pain. In pseudoaddiction, medication-seeking behaviors abate when pain is controlled with adequate analgesic therapy, but in addiction they do not.

There are several patterns of misuse of or nonadherence to prescribed medications that can occur when patients are concerned about medication. In general, patients take more medication than prescribed. This occurs when patients are concerned about withdrawal, tolerance, addiction, scrutiny of medication by others, or a greater perceived need for medication. Having these concerns is a risk factor for developing a substance use disorder. Patients are at greatest increased risk for substance use disorders during the first 5 years of chronic pain treatment, and the risk is higher if they have a history of substance abuse or psychiatric comorbidity.

Aberrant medication-taking behaviors are present in about 50 percent of chronic pain patients who are prescribed opioid analgesics. However, studies suggest that only about 10 percent of patients with aberrant medication-taking behaviors do so as the result of a substance use disorder.

Patients with a history of a substance use disorder are at increased risk for relapse when prescribed opioid analgesics for pain, but the process is not well understood and is likely multifactorial. The cycle of pain followed by relief after taking medication is reinforcing and certainly plays a role. In patients who develop a new substance use disorder, the prescribed medication is commonly involved. Risk prediction tools such as the Opioid Risk Tool (ORT), Screener and Opioid Assessment for Patients with Pain (SOAPP), and Current Opioid Misuse Measure (COMM) are available to offer guidance, but they are self-assessment tools which have limitations.

Conversely, patients with substance use disorders are at increased risk of having a chronic pain condition. This can lead to stigmatization and undertreatment of their pain. This is potentially a difficult population to treat because opioid-dependent patients with chronic pain have relatively high rates of substance use disorders. Most chronic pain patients who enter a treatment facility for a substance use disorder reported that they received a prescription for opioids from a clinician. They also frequently reported obtaining opioids illegally and altering the route of administration, which suggests that they had progressed to a severe substance use disorder.

Frequent follow-up care is necessary for patients with chronic pain who are prescribed opioids. Prior to the initiation of opioids, the risks and benefits of the medication should be discussed with the patient and having the patients sign an opioid consent form is recommended. Close monitoring with intermittent, random toxicology testing should also be performed. Integrating care for chronic pain with innovative stepped-care models of substance abuse could improve outcomes by tailoring the treatments and their intensity to the needs of the individual.

Depression and Affective Illness

There is an intimate and bidirectional relationship between pain and depression. Patients with major depression often have physical symptoms and about 60 percent of depressed patients complain of pain. It has been estimated that patients with depression have about a fourfold increase in the prevalence of chronic pain conditions. Baseline symptoms of depression have been found to predict future pain and disability better than initial ratings of the pain itself.

There are higher rates of lifetime major depression in patients with chronic physical complaints. Of patients that present to pain clinics, up to 50 percent meet criteria for major depressive disorder. In patients with chronic low back pain, the rate of depression is greater than three times the rate in the general population. In patients with chronic pain, depression is associated with greater severity and persistence of pain, greater interference from pain, more pain behaviors observed by others, and early retirement from working. These patients are also more likely to use medication, including analgesics and sedative-hypnotics, daily. Studies indicate that comorbid pain and depression almost doubles the health care costs of managing the conditions.

Patients with chronic pain syndromes such as migraine, chronic abdominal pain, and orthopedic pain conditions also report greater suicidality. Rates of suicidal ideation, suicidal attempts, and suicide completions are two to three times greater in these patients. There tends to be a decrease in self-efficacy in chronic pain patients that contributes to feelings of hopelessness. Depression is the most consistent and strongest predictor of suicide in patients with chronic pain, although other psychosocial variables likely play a role.

Comorbid pain can result in depression that is more resistant to treatment. Depression should not be viewed simply as an expected outcome in patients with pain and should be treated aggressively. Successful treatment of depressive symptoms often results in the reduction of pain and improvement in functional status and quality of life. This effect is independent of the direct treatment of pain and can likely be explained by the fact that the neurobiology of pain and depression overlap.

Anxiety Disorders, Fear, and Anger

Anxiety symptoms are present in almost half of patients with chronic pain conditions. Up to 30 percent of these patients meet the criteria for an anxiety disorder such as generalized anxiety disorder, panic disorder, agoraphobia, and PTSD. PTSD is increasingly recognized as a comorbid condition in patients with medical illnesses, especially chronic pain. It is associated with greater levels of pain, emotional distress, interference in function, and disability. On the other hand, anxiety is associated with high levels of somatic preoccupation and physical symptoms. Pain complaints are present in up to two-thirds of patients with panic disorder, and pain intensity is greater in patients with symptoms of anxiety. Pain is associated with higher levels of anxiety and more frequent panic symptoms.

Fear of pain, movement, and reinjury can result in the avoidance of activities that can promote the transition to and maintenance of chronic pain. Restriction of activities can result in physiologic changes, such as weight gain and muscle atrophy, and functional deterioration. This process is reinforced by feelings of low self-efficacy, catastrophic interpretations, and expectations of failure in rehabilitation. The fear-avoidance model of pain incorporates components such as fear of and attention to pain, catastrophizing, escape or avoidance behavior, disability, disuse, and vulnerability to help explain the transition from acute to chronic pain. Fear avoidance beliefs have been shown to be one of the most significant predictors of failure to return to work. They can reinforce disability through operant conditioning if the avoidance provides short-term benefits, such as relieving anxiety or unwanted responsibilities. Decreasing work-specific fears can be more important than addressing the general fears of activity when trying to improve physical capability to work.

Catastrophic thinking can be described as the amplification of threatening information. It involves elements of cognitive rumination, symptom amplification, expectations of pessimism, and feelings of helplessness. Catastrophizing interferes with the patient's ability to remain involved in productive activities and intensifies the experience of pain, emotional distress, and perceived disability. Cognitive distortions such as catastrophizing and fear-avoidance beliefs predict poor coping and adjustment to chronic pain better than objective factors such as disease status and physical impairment. Catastrophizing is also a predictor of suicidality, independent of pain severity and depressive symptoms.

TREATMENT

The primary aims in the treatment of pain are elucidating a cause, alleviating suffering, and restoring function. Biological, psychological, and social factors all play a role in the perception of pain and each should be addressed. Pain reduction, functional improvement, minimizing

psychological comorbidities, and improving social interactions are all important goals. A strong consensus exists that an interdisciplinary rehabilitation approach to the treatment of pain is beneficial.

Pharmacologic Treatment

There are numerous medications used to treat chronic pain. These medications target the mechanisms of peripheral and central nervous system sensitization such as sodium and calcium channel upregulation, descending modulation, loss of inhibitory interneurons, spinal hyperexcitability, and aberrant sympathetic-somatic nervous system interactions. Antidepressants, anticonvulsants, and antipyretic analgesics are the best-studied medications and are recommended as first-line agents. Opioids are also well studied and frequently used in the treatment of pain, but are not without significant drawbacks and long-term effectiveness has not been demonstrated in rigorous research. There is no simple medication algorithm that provides a straightforward approach to the complexities encountered in the treatment of chronic pain.

Antipyretic Analgesics

Aspirin, acetaminophen, and nonsteroidal anti-inflammatory drugs (NSAIDs) are the backbone of pharmacologic management of acute pain. Aspirin is the most widely used analgesic in the world. These drugs inhibit cyclooxygenase, which is the rate-limiting enzyme in the production of prostaglandins. Prostaglandins regulate inflammation and sensitize nociceptors. Acetaminophen works primarily in the central nervous system, lacks anti-inflammatory effects, and is a weaker analgesic than NSAIDs. NSAIDs work both centrally and peripherally, making them effective orally and topically. Antipyretic analgesics are more effective in the treatment of nociceptive pain than they are for neuropathic pain. They are appropriate treatments for mild to moderate pain, but they have a ceiling effect and usually cannot be used as standalone agents in the treatment of severe chronic pain. They work synergistically with opioids and have proven opioid-sparing effects.

Antipyretic analgesics, in particular NSAIDs, have significant side effects including bleeding, gastrointestinal ulceration, renal toxicity, and increased risk of cardiovascular events. The use of cyclooxygenase-2 selective inhibitors, such as celecoxib, and acid reducing agents, such as proton pump inhibitory, may reduce the risk of gastrointestinal complications, but nothing can attenuate the renal and cardiovascular risks. As such, they should be prescribed at the lowest effective dose for the shortest duration possible. When appropriate, topical NSAIDs can be used and have comparable efficacy for regional pain conditions.

Opioid Analgesics

Opioid analgesics interact with opioid receptors in both the peripheral and central nervous systems to reduce the sensory and affective components of pain. They are the cornerstone of treatment for cancer pain, but controversy surrounds their long-term use for chronic nonmalignant pain. Several reviews have indicated that they provide prolonged relief and improvement in function in few individuals and their superiority over other analgesics is inconclusive.

There are several available guidelines for the use of opioids in chronic nonmalignant pain, and careful attention to balancing the beneficial effects with undesirable adverse effects is necessary. Rational use requires an understanding of the clinical pharmacology of opioids including potency, duration of action, bioavailability, pharmacokinetics and pharmacodynamics. The side effects of opioids are numerous and include sedation, nausea, constipation, respiratory depression, pruritus, urinary retention, and endocrine dysfunction. Important steps in the managing of chronic pain patients on opioid therapy are a thorough evaluation, detailed documentation, a function-based treatment plan with an exit strategy for discontinuing opioids, patient education, informed consent, frequent follow-up, and close monitoring. Patients who are appropriate for opioid analgesia include those with moderate-to-severe pain lasting for more than 3 months that adversely affects function or quality of life. Prior to starting opioid therapy, factors such as the patient's particular pain condition, response to therapy, and potential for aberrant drug-related behaviors should be considered. Treatment outcomes, such as degree of analgesia, activities of daily living and adverse events should be routinely assessed.

Long-term opioid use is associated with tolerance and dependence. Cross-tolerance among opioids is incomplete and alternating opioids may result in up to a 50 percent reduction in equianalgesic dose, so opioid rotation is a strategy that may be employed when tolerance is suspected. Reported addiction rates in chronic pain patients who are prescribed opioids vary widely and range from 5 to 50 percent. A synthesis of evidence suggests that about 50 percent of patients will exhibit aberrant drug-related behaviors, 20 percent will abuse the medications, and 10 percent will become addicted.

Table 27.15–4.
Clinically Available Opioids Used in Chronic Pain Conditions

Drug	Equianalgesic Dose	Duration of Action
Pure agonists		
Morphine	30 mg	3–6 hours[a]
Oxycodone	20 mg	3–6 hours[a]
Hydrocodone	30–60 mg	3–6 hours[a]

Hydromorphone	3–6 mg	3–6 hours[a]
Oxymorphone	10 mg	3–6 hours[a]
Fentanyl	12.5 mcg/h (TD)	72 hours (TD)
	800–1,000 mcg (TM)	1.5–3 hours (TM)
Codeine	200 mg	3–6 hours[a]
Methadone	2–20 mg	6–12 hours for pain
Agonist/Antagonists, Partial agonists		
Buprenorphine	0.3 mg (SL)	6–8 hours (SL)
	5–70 mcg/h (TD)	7 days (TD)
Butorphanol	1 mg (IN)	3–4 hours
Pentazocine	90 mg	8 hours
Weak, Dual-Action Agonists		
Tramadol	150–300 mg	4–6 hours[a]
Tapentadol	75–110 mg	4–6 hours[a]

[a]Long-acting formulations available. TD, transdermal; TM, transmucosal; SS, sublingual.

Clinically available opioids include: naturally occurring compounds, such as morphine and codeine; synthetic opioids, such as fentanyl, methadone, and tramadol; and semisynthetic derivatives, such as oxycodone, hydrocodone, hydromorphone, and oxymorphone (Table 27.15–4). Morphine has poor oral bioavailability due to its hydrophilicity and delayed central nervous system absorption and onset of action. This delay prolongs the analgesic effects of morphine, which decreased the potential for accumulation and toxicity with repeated dosing. Oxycodone is a popular analgesic used to treat chronic pain patients. It has a higher bioavailability, faster onset of action, and more predictable plasma levels compared to morphine. These qualities may increase the risk of abuse and addiction. It has similar analgesic efficacy to morphine, but causes less histamine release and fewer hallucinations. Hydrocodone is similar to oxycodone with regards to its rapid oral absorption and onset of action, but is a less potent analgesic. It is metabolized to hydromorphone, which has properties similar to those of morphine, but has lower rates of side effects. Fentanyl is highly hydrophilic with affinity for neuronal tissues. It is frequently used intravenously in the treatment of acute surgical pain, but can also be delivered via transdermal and transmucosal routes for the treatment of chronic pain.

Methadone for use in the treatment of chronic pain warrants special consideration because of its stigma, low cost, high bioavailability, rapid onset of action, long half-life, multiple receptor affinities (NMDA), lack of neurotoxic metabolites, and incomplete cross-tolerance with other opioids. It has significantly greater risk of overdose because of the longer time needed for adaptation and its long and variable half-life (15 to 120 hours). Methadone is also associated with QTc prolongation and torsades de pointes at high doses (>100 mg per day) as well as increased rates of all-

cause mortality at all doses. Although once-daily dosing is adequate to prevent withdrawal symptoms due to its extensive tissue distribution, its analgesic effects only last about 6 hours due to biphasic elimination. Repeated dosing can increase the analgesic period to 8 to 12 hours because it accumulates in tissues, but this same phenomenon can also increase the risk for delayed respiratory depression, overdose, and death.

Tramadol and tapentadol are a newer class of analgesic drugs with weaker analgesic effects and a dual mechanism of action. They are weak opioid agonists that also inhibit the reuptake of serotonin and norepinephrine. Adverse interactions can occur when administered with coumadin, selective serotonin reuptake inhibitors, and serotonin-norepinephrine reuptake inhibitors, and they are associated with tolerance, dependence, and abuse.

Butorphanol, nalbuphine, and pentazocine are opioid agonist–antagonist drugs. Buprenorphine is a partial agonist at the mu-receptor and antagonist at the other opioid receptors. These drugs should be used with caution in patient taking pure agonist drugs because they can attenuate the effects of the pure agonists and even precipitate withdrawal. Buprenorphine (in combination with naloxone) is available orally for treatment of opioid addiction and as a transdermal patch for the treatment of chronic pain.

Antidepressants

The analgesic effects of antidepressants are thought to be independent of their antidepressant effects and are mediated primarily by the reuptake inhibition of serotonin and norepinephrine. Increased levels of these neurotransmitters enhance the activation of the descending inhibitory neurons in the dorsal horn of the spinal cord. They may also produce antinociceptive effects through a variety of other mechanism as well including modulation of monoamines, interaction with opioid systems, inhibition of ion channel activity, and NMDA, histamine, and cholinergic receptors. This suggests that all antidepressants would be effective in the treatment of chronic pain, but the analgesic properties of these medications remain underappreciated.

Tricyclic Antidepressants

Studies have confirmed that TCAs are the most effective agents for the treatment of neuropathic pain. They are also effective in the treatment of headache syndromes, central pain syndromes, PHN and other painful neuropathies, and postmastectomy pain syndrome. They have not been shown to be effective in treating SCI pain, phantom limb pain, or painful HIV neuropathy. All TCAs have equal analgesic effects, but the secondary amines, such as nortriptyline, are better tolerated than the tertiary amines,

such as amitriptyline. TCAs may produce analgesia at lower doses than needed for the treatment of depression but treatment should be guided by titration to efficacy. Advantages of TCAs include cost-effectiveness and that optimal titration can be monitored with serum levels to insure adequate trials and the avoidance of toxicity.

Serotonin-Norepinephrine Reuptake Inhibitors

SNRIs include duloxetine, venlafaxine, desvenlafaxine, and milnacipran, and they inhibit the presynaptic reuptake of serotonin, norepinephrine, and, to a lesser extent, dopamine. They are associated with fewer side effects than TCAs and can be titrated quicker, but titration cannot be monitored by serum levels. Duloxetine more potently blocks serotonin and norepinephrine than does venlafaxine. It has been shown to be effective in the treatment of fibromyalgia and painful diabetic neuropathy. Venlafaxine has been shown to reduce neuropathic pain breast cancer treatment, decrease hyperalgesia and allodynia in neuropathic pain and atypical facial pain, and prevent migraines. Milnacipran reduces hyperalgesia mediated by monoamine and opioid systems and is effective in the treatment of fibromyalgia.

Selective Serotonin Reuptake Inhibitors

The efficacy of selective serotonin reuptake inhibitors in the treatment of chronic pain conditions has been inconsistent in clinical trials. For some conditions, such as migraine, SSRIs do not appear to be more efficacious than placebo. For other conditions, such as fibromyalgia, abdominal pain associated with irritable bowel syndrome, and painful diabetic neuropathy, there is evidence that their use can reduce pain, independent of changes in affective disorders. Because of these inconsistencies, SSRIs are not recommended as first-line agents. Instead, they should be used as alternatives to TCAs or SNRIs when they are ineffective or not well tolerated.

Novel Antidepressants

The neurobiology of pain suggests that all antidepressants would be effective for the treatment of pain, but there is little evidence to support the use of novel antidepressants such as mirtazapine, bupropion, trazodone, vortioxetine, and vilazodone. Mirtazapine has been shown to decrease the duration and intensity of chronic tension type headaches and bupropion has been effective in the treatment of neuropathic pain by decreasing pain intensity and interference of pain with quality of life. Several reports have suggested that trazodone is effective for treating chronic pain, but side effects can hinder titration to effective doses. Vortioxetine and vilazodone are fairly new antidepressants that selectively inhibit serotonin reuptake

and also have action at 5-HT3 and HT1A receptors. They have been shown to be effective antidepressants, but their efficacy in treating pain has not yet been established.

Anticonvulsants

Anticonvulsants reduce pain by inhibiting excessive neuronal activity. They block voltage-gated sodium channels, modulate calcium channels, inhibit excitatory amino acid neurotransmitters, and enhance amino acid–mediated inhibitory neurotransmitters. They are effective for the treatment of diabetic neuropathy, trigeminal neuralgia, PHN, and migraine recurrence.

First-generation anticonvulsants have been used in the treatment of pain since the 1940s when phenytoin was first reported to be successful for the treatment of trigeminal neuralgia. Carbamazepine is the most widely studied anticonvulsant in the treatment of neuropathic pain and is still the first-line treatment for trigeminal neuralgia. Valproic acid is commonly used in the prophylaxis of migraine, but has also been shown to be effective for other neuropathic pain conditions as well.

Second-generation anticonvulsants include gabapentin, pregabalin, lamotrigine and topiramate. Gabapentin and pregabalin decrease the influx of calcium, modulate the release of substance P and glutamate, and activate the descending inhibitory of nociception. They are effective in the treatment of painful diabetic neuropathy, PHN, fibromyalgia, phantom limb pain, and SCI-associated pain. Using flexible titration schedules has been associated with fewer discontinuations, higher final dose, and slightly better pain relief compared to fixed schedules. Lamotrigine is effective in treating HIV-related neuropathy and central poststroke pain, but has been disappointing in treating other neuropathic pain conditions. Topiramate has been effective in migraine prophylaxis, chronic LBP, lumbar radiculopathy, and painful diabetic neuropathy. It has the advantages of minimal hepatic metabolism, few drug interactions, and a long half-life; it also has the unusual side effect of weight loss.

Oxcarbazepine is a carbamazepine derivative with an improved safety profile that is considered a next-generation anticonvulsant. Studies have demonstrated its efficacy in painful diabetic neuropathy and refractory PHN. Other next-generation anticonvulsants include tiagabine, vigabatrin, retigabine, levetiracetam, and zonisamide. They have a wide spectrum of pharmacologic actions and antinociceptive effects in animals, but studies in humans are sparse. They are also associated with significant cognitive impairment and are not recommended as first-line treatments for chronic pain conditions. Combinations of anticonvulsants with complimentary mechanisms of action should always be considered as they may increase effectiveness and decrease adverse effects of high dose monotherapy.

Other Adjunctive Drugs

Antipsychotics have been used in diabetic neuropathy, PHN, headache, facial pain, musculoskeletal pain, and pain associated with AIDS and cancer, and there is increasing evidence to support their effectiveness. Compared to typical antipsychotics, atypical antipsychotics offer a broader therapeutic spectrum, have lower rates of extrapyramidal side effects, and have greater potential for augmentation in the treatment of depression. However, they are associated with metabolic disruptions including weight gain, new-onset diabetes, hyperlipidemias, and cardiac arrhythmias, which offset their benefits.

Oral and topical local anesthetics have long been used in the treatment of pain. Topical lidocaine is approved for the treatment of PHN, but has been shown to be ineffective in the treatment of HIV neuropathic pain. It does not produce significant serum levels. Oral mexiletine is effective in the treatment of diabetic neuropathy, peripheral nerve injury, and alcoholic neuropathy. There are mixed reports for its use in phantom limb pain and it appears to be ineffective in treating cancer-related pain. The analgesic effects of mexiletine did not correlate with serum levels.

Calcium channel blockers have also been used in the treatment of chronic pain syndromes. Verapamil is the most commonly prescribed of the calcium channel blockers and is effective in the treatment of migraine and cluster headaches. Ziconotide is a neuron-specific calcium channel blocker that is derived from the toxin of the cone snail, Conus magus. It has been approved for intrathecal use in the treatment of refractory cancer or AIDS pain and has potent analgesic effects that can be synergistic with morphine. However, it produces common side effects including dizziness, nausea, confusion, and headache that often make it difficult to tolerate.

Benzodiazepines are commonly prescribed for chronic pain patients to treat insomnia, anxiety, and spasticity, but practitioners must use them with caution in this population. Their use in treating target symptoms is not supported in the literature. In fact, benzodiazepine use has been associated with decreased activity levels, higher rates of health care visits, depression, increased domestic instability, and more disability days. In combination with opioids, their use can be counterproductive and potentially dangerous. In studies of opioid related mortality, concomitant benzodiazepine use was frequently implicated.

Psychological Treatments

Cognitive-Behavioral Models. Many psychological treatments have been used in the treatment of pain and its associated distress and disability. Fordyce pioneered the psychological treatment for chronic pain using an operant conditioning behavioral model. The behavioral approach focuses on understanding pain occurring in a social context. The behavior of

chronic pain patients is not only reinforced by others, but also reinforces the behavior of others, including physicians. Pain behaviors, such as grimacing, guarding, and taking pain medications are indicators of perceived pain severity and functional disability. It has been shown that these pain behaviors alone are predictors of whether or not opioid analgesics were prescribed to patients. The behavioral model assumes that pain and disability will persist if these behaviors persist. In CBT, healthy behaviors are targeted for reinforcement to replace extinguished pain behaviors.

The cognitive-behavioral model of chronic pain is based in the theory that individual beliefs, attitudes, and expectations affect emotional and behavioral reactions to life experiences. Pain and the resulting pain behaviors are influenced by biological, psychological and socio-environmental variables. Patients will experience more negative affective responses, increased pain, and decreased functioning if the patient believes that pain is inevitable and uncontrollable. This cycle of disability can be interrupted by CBT approaches including relaxation, biofeedback, guided imagery, meditation, motivational interviewing, external reinforcement, cognitive restructuring, and the development of coping strategies. Patients are taught to be active participants in the management of their pain by using these methods to minimize distressing thoughts and feelings. The goal of CBT is to increase activity, independence, and resourcefulness regardless of the particular technique used. CBT has been proven to be effective in reducing pain severity, pain behaviors, distress, and depression and improving coping mechanisms. These benefits can persist for longer than 6 months after the completion of treatment.

A heterogeneous group of behavioral treatments has been termed "third-wave psychotherapies." These include acceptance and commitment treatment, behavioral activation, cognitive-behavioral analysis system psychotherapy, dialectical behavioral therapy, mindfulness-based cognitive therapy, and metacognitive therapy. These therapies incorporate a variety of psychological and social theories that emphasize the context and function of psychological phenomena rather than just form. The goal is to create strategies for creating change in context and experience to develop broad, flexible, and effective repertoires to improve quality of life. These therapies have added to the spectrum of empirically supportive treatments for mental disorders and could certainly be useful in the treatment of pain disorders.

Beliefs

Beliefs can be conceptualized as the thoughts that patients have about their personal pain problem. Psychosocial dysfunction has been correlated with believing emotions are related to pain, receiving overly solicitous responses from others, and attributing dysfunction to pain. Patients have also

endorsed the belief that pain signifies injury and that, therefore, activity should be avoided. A change in perceived control over pain has been shown to be a predictor of the beneficial effects of CBT in the treatment of chronic pain.

Self-efficacy, outcome expectations, and locus of control are cognitive variables derived from the social learning theory. Self-efficacy is a belief about one's ability to perform a specific behavior and outcome expectation is a belief about the consequences of performing the behavior. Patients are more likely to engage in coping efforts that they believe are within their capabilities and will result in a positive outcome. Patients who score higher on self-efficacy measures or have an internal locus of control have reported lower levels of pain, higher pain thresholds, increased exercise performance, and more positive coping efforts.

Acceptance of pain is associated with multiple domains of the experience of pain including lower pain intensity, less pain-related anxiety and avoidance, less depression, less disability (physical and psychological), more daily uptime, and better work status. It predicts a patient's adjustment to the illness and is independent of catastrophizing, coping skills, and pain-related beliefs and cognitions.

Coping

Coping has been defined as an individual's cognitive and behavioral efforts to solve personal or interpersonal problems that are appraised as taxing or exceeding the person's resources. Coping strategies support the cognitive-behavioral model of chronic pain, whether they are active, passive, adaptive or maladaptive. However, patients who remain passive or continue to use maladaptive coping strategies such as catastrophizing, ignoring or reinterpreting pain sensations, or diverting attention from pain have higher disability.

The effectiveness of particular coping strategies is dependent on many aspects of the patient's experience with pain. For example, reinterpreting pain as not being signs of ongoing injury had been found to be useful in reducing experimentally induced pain. However, focusing on pain and disability for greater amounts of time, as is needed in CBT to help patients reinterpret pain sensations, can be detrimental. It can divert patients from engaging in social activities and healthy behaviors. When patients perceive an increase in the effectiveness of a new coping skill, pain and disability can be decreased and their use of maladaptive coping strategies can be reduced.

Placebo Response

Placebo effects are complex phenomena that are similar to those of active treatments and can be biologically measured. The literature supporting the placebo response has been criticized as flawed and overrated. Clinically, it

is difficult to distinguish improvements from treatment and placebo responses. In the past, placebo interventions were a part of the paternalistic practitioner's treatment armamentarium. Currently, however, controversy exists over the evidence of clinically important effects that justify the use of placebo treatments. A positive placebo response does not prove that the patient's pain is psychogenic or that the patient would not benefit from active treatment, and use of a placebo to determine if the patient's pain is "real" is dishonest and counterproductive. Such an intervention can result in the loss of trust in the practitioner and render future treatments less effective.

Multidisciplinary Rehabilitation

Chronic pain patients have greater reductions in physical, psychological, and social well-being and lower rated quality of life than do those with any other medical condition. The goal of treating chronic pain is to ease suffering and end disability that prevents people from returning to productive activities. Evidence suggests that early failures in single-modality treatment for chronic pain can have devastating consequences for future treatment effects. Furthermore, analysis of multidisciplinary programs that use comprehensive assessments, step-care treatments, and rehabilitation goals result in significant reductions in pain along with functional and quality-of-life improvements. Evidence-based practice guidelines emphasize interdisciplinary rehabilitation, integrated treatment, and patient selection criteria. Furthermore, interdisciplinary pain rehabilitation programs (IPRPs) are efficacious in returning patients to work and ultimately can reduce the economic burden of chronic pain.

Invasive Treatments

Nerve Blocks. In the treatment of chronic pain, injections can be performed for diagnostic and/or therapeutic purposes. Injections with local anesthetics may work by enhancing blood flow, releasing entrapped nerves, and interrupting the processes involved in central sensitization. Adding corticosteroid to local anesthetics may provide the added benefits of blocking the inflammatory cascade and inhibiting the synthesis of prostaglandins which can sensitize nociceptors.

Although nerve blocks are rarely curative, they may provide pain relief for a time and facilitate rehabilitative therapy. Peripheral nerve blocks that can provide relief in properly selected patients include trigger point injections for myofascial pain, intra-articular injections for osteoarthritis, and direct nerve blocks for entrapment syndromes. Spinal injections include epidural steroid injections for radicular pain and facet injections, sometimes with radiofrequency denervation, for facet joint pain. Other neuroablative procedures using alcohol or glycerol, such as celiac plexus

neurolysis for pancreatic cancer and hypogastric plexus neurolysis for pelvic pain, can provide pain relief for several months. However, they can be associated with severe rebound pain and should be avoided in patients with nonmalignant pain conditions.

Electrical Stimulation

Spinal cord stimulation (SCS) is a reversible treatment that uses low-voltage electrical pulses to manage chronic, intractable pain conditions. The procedure includes the implantation of a pulse generator, extension cable, and leads, which are placed in the dorsal spinal column near the nerve root that is thought to be associated with the pain. An electrical current induces a paresthesia, or a tingling sensation that is meant to mask the pain. SCS was developed based on the gate-control theory, which postulates that activation of peripheral sensory A-fibers can attenuate pain signaling by C-fibers. SCS is typically used in pain conditions that have been refractory to other treatments including failed back surgery syndrome and CRPS.

Surgery

Surgical interventions should be reserved for chronic pain patients who have failed to respond to more conservative measures. It can be challenging to decide when operative procedures are appropriate because of the difficulty involved in establishing a causative relationship between the targeted pathologic process and pain. In many individuals, the targeted disease may not be the primary cause of pain. Rarely does removing a painful body part result in long-term pain relief. Furthermore, the sequelae of surgery itself can result in worsening pain, such as when scar tissue or adhesions develop. Careful patient selection is a necessity in deciding who should receive surgery for relief of a chronic pain condition.

Conclusion

Chronic pain is a significant public health problem and negatively impacts those who are affected by it. The traditional view that pain is either caused by tissue pathology (nociception) or psychological states (suffering) is too simplistic and outdated. It is now widely recognized that the relationship between pain and nociception is complex and must be understood in terms of the situation of the whole individual. There are high rates of psychiatric comorbidity in patients with chronic pain and a psychiatric evaluation should be part of any comprehensive pain evaluation. Psychiatric diagnosis and treatment can add an essential and often neglected component to the conceptualization and treatment of chronic pain disorders. Care for patients with chronic pain should occur in a medical treatment setting, as this is the most effective way to reassure patients that the somatic elements

of their pain are not neglected and it allows effective integration of somatic and psychological treatments.

References

Bair MJ, Robinson RL, Katon W, et al. Depression and pain comorbidity: a literature review. *Arch Intern Med*. 2003;163:2433–2445.

Balasubramaniam R, Klasser GD. Orofacial pain syndromes: evaluation and management. *Med Clin North Am*. 2014;98(6):1385–1405.

Birklein F, O'Neill D, Schlereth T. Complex regional pain syndrome: an optimistic perspective. *Neurology*. 2015;84(1):89–96.

Chou R, Qaseem A, Snow V, et al; Clinical Efficacy Assessment Subcommittee of the American College of Physicians; American College of Physicians; American Pain Society Low Back Pain Guidelines Panel. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med*. 2007;147(7):478–491.

Clark MR. Pain. In: Coffey CE, Cummings JL, eds. *Textbook of Geriatric Neuropsychiatry*. 3rd ed. Washington, DC: American Psychiatric Press; 2011.

Clark MR. Pain management. In: Levenson JL, ed. *Textbook of Psychosomatic Medicine*. 2nd ed. Arlington, TX: American Psychiatric Publishing, Inc.; 2011.

Clark MR, Owen JA. Pain management. In: Ferrando S, ed. *Clinical Handbook of Psychopharmacology in the Medically Ill*. Arlington, TX: American Psychiatric Publishing, Inc.; 2010:613–656.

Clauw DJ. Fibromyalgia: a clinical review. *JAMA*. 2014;311(15):1547–1555.

Cohen SP, Mao J. Neuropathic pain: mechanisms and their clinical implications. *BMJ*. 2014;348:f7656.

De Rossi SS, Greenberg MS, Liu F, Steinkeler A. Temporomandibular disorders: evaluation and management. *Med Clin North Am*. 2014;98(6):1353–1384.

Finnerup NB. A review of central neuropathic pain states. *Curr Opin Anaesthesiol*. 2008;21:586–589.

Flaster M, Meresh E, Rao M, Biller J. Central poststroke pain: current diagnosis and treatment. *Top Stroke Rehabil*. 2013;20(2):116–123.

Guzman J, Esmail R, Karjalainen K, et al. Multidisciplinary bio-psycho-social rehabilitation for chronic low-back pain. *Cochrane Database Syst*. 2007;(2):CD000963.

Haibach JP, Beehler GP, Dollar KM, Finnell DS. Moving toward integrated behavioral intervention for treating multimorbidity among chronic pain, depression, and substance-use disorders in primary care. *Med Care*. 2014;52(4):322–327.

Henschke N, Kamper SJ, Maher CG. The epidemiology and economic consequences of pain. *Mayo Clin Proc*. 2015;90(1):139–147.

International Association for the Study of Pain. IASP taxonomy. Availble at http//www.iasp-pain.org.

Kahla KG, Wintera L, Schweiger U. The third wave of cognitive behavioural therapies: what is new and what is effective? *Curr Opin Psychiatry*. 2012;25:522–528.

Lee MC, Tracey I. Imaging pain: a potent means for investigating pain mechanisms in patients.

Br J Anesthe. 2013;111(1):64–72.

Manchikanti L, Abdi S, Atluri S, et al. American Society of Interventional Pain Physicians (ASIPP) guidelines for responsible opioid prescribing in chronic non-cancer pain: part I—evidence assessment. *Pain Physician*. 2012;15(3 Suppl):S1–65.

Manchikanti L, Abdi S, Atluri S, et al. American Society of Interventional Pain Physicians (ASIPP) guidelines for responsible opioid prescribing in chronic non-cancer pain: part 2—guidance. *Pain Physician*. 2012;15(3 Suppl):S67–116.

Marnitz U, Weh L, Muller G, et al. Multimodal integrated assessment and treatment of patients with back pain: pain related results and ability to work. *Schmerz*. 2008;22:415–423.

McCleane G. Antidepressants as analgesics. *CNS Drugs*. 2008;22:139–156.

Molton IR, Graham C, Stoelb BL, Jensen MP. Current psychological approaches to the management of chronic pain. *Curr Opin Anaesthesiol*. 2007;20(5):485–489.

Niraj S, Niraj G. Phantom limb pain and its psychologic management: a critical review. *Pain Manag Nurs*. 2012;15(1):349–364.

Pontell D. A clinical approach to complex regional pain syndrome. *Clin Podiatr Med Surg*. 2008;25:361–380.

Rief W, Martin A. How to use the new DSM-5 somatic symptom disorder diagnosis in research and practice: a critical evaluation and a proposal for modifications. *Annu Rev Clin Psychol*. 2014;10:339–367.

Seidel S, Aigner M, Ossege M, et al. Antipsychotics for acute and chronic pain in adults. *Cochrane Database Syst Rev*. 2008;(4):CD004844.

Stein C, Reinecke H, Sorgatz H. Opioid use in chronic non-cancer pain: guidelines revisited. *Curr Opin Anaesthesiol*. 2010;23:598–601.

Turner JA, Holtzman S, Mancl L. Mediators, moderators, and predictors of therapeutic change in cognitive-behavioral therapy for chronic pain. *Pain*. 2007;127:197–198.

Verdu B, Decosterd I, Buclin T, et al. Antidepressants for the treatment of chronic pain. *Drugs*. 2008;68:2611–2632.

Wasan AJ, Sullivan MD, Clark MR. Psychiatric illness, depression, anxiety, and somatoform pain disorders. In: Fishman SM, Ballantyne JC, Rathmell JP, eds. *Bonica's Management of Pain*. 4th ed. Baltimore, MD: Lippincott Williams &Wilkins; 2009:19–42.

28

Additional Conditions That May Be a Focus of Clinical Attention

▲ 28.1 Malingering

FRANK JOHN NINIVAGGI, M.D.

The condition termed *malingering* was formally introduced into psychiatry in 1980 in the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III). DSM-5 denotes it as the intentional production of false or grossly exaggerated physical or psychological symptoms motivated by external incentives and states that aspects of it may represent adaptive behavior. The symptoms of malingering are under voluntary control. This reflects intentionally contrived statements and behaviors that have both conscious and nonconscious components, all of which are intended to mislead the evaluator, usually in medical and legal contexts. Emphasis is given to the conscious awareness that the malingerer has of the malingering presentation. Malingering as a form of illness behavior suggests dysfunction but not true psychopathology. This section discusses the psychiatric condition and clinical indicators suggestive of malingering primarily as they present themselves in the psychiatric assessment. Neuropsychological investigations are often used as one part of comprehensive evaluations. Legal considerations and determinations are not part of the psychiatrist's role.

Misconceptions regarding malingering abound. Some consider it rare; some believe it to be common. Some equate all medical deception with malingering and correlate this feigning with wholesale antisocial personality disorder. Studies, however, show that those with antisocial personality disorder are not more likely to malinger or to be more successful at doing so than the general population. Mild exaggeration of symptoms and efforts at "impression management" or reasonable attempts to put oneself in a good light do not constitute malingering. When malingering is present, it is thought to be a static response with a stable base rate; this is not true. Some think of it as a diagnosis with clear-cut

criteria; rather; it is a condition with broad parameters. Malingering does not preclude other genuine disorders, such as posttraumatic stress disorder (PTSD), depression, traumatic brain injury (TBI), and others to be discussed later. While DSM-5 gives criteria for malingering, some have suggested that these require further precision to ensure greater validity.

The 2001 edition of the *New Oxford American Dictionary* denotes *malingerer* as one who exaggerates or feigns illness to escape duty or work. It dates the origin of the term in English to the very early 19th century, describing it as "back-formation from *malingere*, apparently from French *malingre*, perhaps formed as *mal-* 'wrongly, improperly' + *haingre* 'weak,' probably of Germanic origin." The term *malingerer* is said to have first appeared in 1785 in a French dictionary and to have denoted a soldier who feigns sickness or induces or protracts an illness to avoid his duty. This 200-year history suggests only relatively recent attention to an individual problem with important social implications. Malingering is socially problematic behavior.

The relative newness of this DSM entity has brought with it a plethora of conceptual difficulties. In the last 25 years, the psychiatric, psychological, and sociological literature on this topic has been increasing. Contributions from these varied sources have been replete with a wide variety of terms attempting to capture not only core meanings but also attendant nuances. In an effort to hone the concept in a relatively lucid fashion and establish a tentative yet uniform lexicon for this section, the following definitions will be used.

1. *Response style* denotes the way one individual who is being evaluated communicates self-disclosures reflecting personal goals (e.g., to reveal or to deceive), intended consequences, and predicted utility in a specific instance and a particular context. Response styles include all linguistic and paralinguistic communications.
2. *Dissimulation* is the broadest term that encompasses all forms of deception specifically in regard to psychological symptomatology. *Dissimulate* denotes concealing, hiding, or disguising perceived truth under a cover of deliberate, voluntarily controlled pretense. In the vernacular, this concept resembles that of *hoax*, an act intended to deceive or produce a facade of deliberate trickery intended to gain an advantage.
3. *Dissimulation* concepts are encompassed in three principal subcategories:
   a. Malingering,
   b. Defensiveness or *faking good*, and
   c. Markedly inconsistent, evasive responses.

Dissimulation in psychiatry correlates with malingering, *faking good*, and factitious disorder.

Malingering is medical deception, asserting that one has significantly distressing symptomatology in the actual absence of such. Malingering, however, spans a spectrum from complete misrepresentation to partial shifts in emphasis that skew the presentation to favor the patient's obtaining some desired goal. Its range of shaded meanings all point to a core denotation of deliberate illness misrepresentation. In this sense, the *gross* exaggeration of feigned symptoms for external gain in ordinary circumstances is typically not merely a plea for help. Nuanced terminology about such situations includes feigning, fabricating, disguising, deceiving, lying, misleading, masquerading, tricking, and faking. The terms "ruse," "subterfuge," and "wile" also capture the intention to deceive. Some researchers, however, restrict the denotation of "feigning" to mean deliberate fabrication or gross exaggeration of symptoms without any assumption about its goals. In a manner of speaking, the malingerer may be considered deliberately to take on the role of impostor, one whose *conscious motivation* seeks to achieve an intentionally misleading appearance.

Malingering has also been termed *faking bad*, which denotes consciously creating false, untrue, yet apparently distressing symptoms. *Faking bad* may also include intentionally over-reporting, embellishing, amplifying, or elaborating symptoms that are present but not as problematic as the individual tries to convey in such grossly exaggerated fashion. Symptom education by way of Internet-derived information is commonplace today.

Defensiveness, or *faking good*, by contrast, is a dissimulation by way of omission. It denotes consciously minimizing or denying symptoms that are present and perceived to have negative consequences, hence their intentional suppression. *Faking good* is a guarded attempt to portray oneself in a better light than is actually the case. For example, pedophiles might use *faking good* strategies, referred to as a "dance of denial" to minimize incriminating behaviors when being examined. In child custody cases, a parent may also use *faking good* intentionally to minimize perceived negative behaviors that might jeopardize that parent's chances of gaining or maintaining custody. Faking good like faking bad is an intentionally motivated form of response bias, that is, reporting material thought to be the "right" answer to achieve a goal. Because the term *defensiveness* has been used in the malingering literature in a rather loose manner, it would be better to replace it with the more precise phrase *consciously intended denial or gross minimization of real symptomatology*.

The term *simulation* has been used as a synonym for malingering but will not be used in this way here. *Simulation* more precisely denotes a consciously contrived pretended situation, an imitation, or a representation (not a consciously deceptive misrepresentation) of one process or situation

by using a different process or situation in an illustrative and instructive context. For example, simulation or analogue research protocols examining dissimulation or contexts in which faking is being studied use a quasi-experimental design with participants randomly assigned to an experimental group (participants asked to dissimulate and deceive) or to a control group (participants asked to be honest).

In the last decade, the condition of malingering within the field of psychiatry has been a focus of concern in assessments in several different medicolegal contexts, for example, insurance claims, disability requests, physical injury litigation, harassment suits, posttrauma cases, neurocognitive dysfunction, chronic pain, and criminal culpability situations. Not only attorneys and courts but also hospital inpatient services request such assessments and consultations. The forensic evaluator (e.g., psychiatrist, psychologist, or attorney)—especially in criminal contexts—acts primarily as fact finder, discovering facts based on credible evidence, often on a very short-term basis. This consultant role is different from the therapeutic relationship that the longer-term treater establishes in a clinically oriented helping situation. More traditional, forensically related general psychiatry assessments were in the realms of suicide, boundary violations, informed consent, capacity and competency, civil commitment, confidentiality, domestic violence, and abuse and neglect.

The primary goal of the psychiatrist in evaluations is first to establish the real presence of actual disorders, and then discern whether and in which areas genuine symptoms are present and whether the condition of malingering might exist. Malingering is the presence of consciously produced symptoms that are not genuine. The role of *projective identification* will be addressed because input from patients, legal systems, and the psychiatric community may unwittingly and implicitly act in the coauthoring of symptoms that influence the psychiatrist's mental processes trying to be objective. Making sense of a presentation with possibly strong emotional components requires a working knowledge of emotional intelligence. An intelligent understanding of emotions and complex feeling states deepens psychiatric perspectives and interpretive options. The wide array of individual differences in response styles may have crucial emotional determinants needing deliberation in assessments. One of the strengths that distinguishes psychiatric assessment and the interpretation of complex, often ambiguous, information is an explicit recognition of unconscious psychodynamics, individual differences, and an appreciation of subtle individual distinctions. Consideration of the presence of what malingering connotes, especially inconsistencies and discrepancies, in any patient's presentation (e.g., self-reports, complex response styles, direct observation, and especially in collateral data) will guide the astute clinician in proper treatment planning, especially the rational prescription of

medications.

DEFINITION AND COMPARATIVE NOSOLOGY

DSM-5 describes the condition of *malingering* as follows:

> The essential feature of malingering is the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs. Under some circumstances, malingering may represent adaptive behavior—for example, feigning illness while a captive of the enemy during wartime. Malingering should be strongly suspected if any combination of the following is noted:

1. Medicolegal context of presentation (e.g., the individual is referred by an attorney to the clinician for examination, or the individual self-refers while litigation or criminal charges are pending).
2. Marked discrepancy between the individual's claimed stress or disability and the objective findings and observations.
3. Lack of cooperation during the diagnostic evaluation and in complying with the prescribed treatment regimen.
4. The presence of antisocial personality disorder.

Malingering differs from factitious disorder in that the motivation for the symptom production in malingering is an external incentive, whereas in factitious disorder external incentives are absent. Malingering is differentiated from conversion disorder and somatic symptom-related mental disorders by the malingerer's intentional production of symptoms and by the obvious external incentives associated with it. Definite evidence of feigning such as clear evidence that loss of function is present during the examination but not at home would suggest a diagnosis of factitious disorder if the individual's apparent aim is to assume the sick role. Deception would suggest malingering if the goal were to obtain an incentive, such as money.

Malingering is coded in DSM-5 as Z76.5. This denotes a condition (not a mental disorder) that may be a focus of clinical attention. DSM-5 differs from DSM-IV-TR in that it adds under 1. "or the individual self-refers while litigation or criminal charges are pending." This addition emphasizes the prevalence of suspected malingering in this population. DSM-5 also includes a new discussion of malingering, factitious disorder, and somatic symptom-related disorders and their distinctions. In factitious disorder, the voluntariness of the pursuit of the sick role as a psychological conclusion is emphasized in contrast to malingering in which the voluntariness of symptom production is in pursuit of financial or other external goals yielding a behavioral conclusion.

The tenth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) lists malingering in Chapter XXI of this text as code Z76.5, page 244. It codes the following:

- Malingering (conscious simulation)
- Person feigning illness (with obvious motivation)

The ICD-10 chapter that contains malingering states that these categories are provided for occasions when circumstances other than disease, injury, or external cause classifiable to categories A00–Y89 (more clear-cut medical and psychiatric disorders) are recorded as "diagnoses" or "problems." This can arise when a person who may or may not be sick encounters the health services for some specific purpose or when some circumstance or problem is present that influences the person's health status but is not in itself a current illness or injury.

The DSM-5 guidelines used to indicate the condition of malingering are minimal and fairly broad. There are no suggestions regarding chronological age range or clear-cut demographic or epidemiological findings because malingering has not been studied in the rigorous, scientific manner in which the mental disorders have been researched. In addition, because malingering denotes patient deception and unreliable information, it is difficult, if not impossible, to organize findings in a consistently coherent fashion. Clinical judgment, especially about inferred motivation and conscious intention typically to avoid perceived distress and achieve desired goals, must be sensitive to ambiguous, indeterminate presentations. Use of careful and well-chosen terminology, therefore, indicating doubt or suspected lack of credibility (e.g., feigning, fabrication, noncredible, and dissimulation short of malingering) when present is essential in diagnostic evaluations and documentation.

HISTORY

Examples of malingering can be found in literature from the story of Ulysses feigning insanity to escape the Trojan War to the biblical story of King David. The image of street beggars and vagabonds who feigned disease to gain alms is a common one in early Western and Eastern cultures. It was not until about 1880 in Germany and about 1910 in the United Kingdom that malingering moved from the political, military, and social spheres to the aegis of the medical profession. World War I reinforced the phenomenon of what already was known as *compensation neurosis*. At the time, medically prescribed removal from duty was added to address war fatigue, shell shock, and war trauma, particularly physical distress. It was at this point that physicians took or were given a leading role in the management of malingering because it presented as a symptomatic picture in the absence of typically associated material causative features. Because many symptom profiles were strongly suggestive of "mental" and emotional distress, psychiatrists and psychoanalysts then took a lead in trying to sort out genuine from feigned illness.

From its introduction in the DSM-III in 1980 to about 2005, there was an almost threefold increase in the scientific literature on malingering. Historical accounts have been written, and clinical and empirical case reports have appeared, but evidence-based research is still in its infancy. Hence, as a scientifically justifiable construct, malingering remains more in the context of discovery. The research phase in which evidence is brought to bear on objectively testable hypotheses remains in statu nascendi.

EPIDEMIOLOGY

The actual incidence, prevalence, and distribution features associated with malingering are imprecise. This is because malingering has not achieved status as a formal psychiatric diagnosis with attendant studies examining validity and reliability in the context of clinically defined populations. Speculations suggest that base rates can vary from 7.5 to 33 percent. In ordinary clinical practice, some have suggested a 1 percent prevalence; in military populations, a 5 percent occurrence; and in litigious, criminal situations, up to 20 percent. A 2002 study of 131 clinicians of the American Board of Clinical Neuropsychology gave the following breakdown of base rates for different contexts of malingering: 29 percent for personal injury, 30 percent for disability or worker's compensation, 19 percent in criminal cases, and 8 percent in medical or psychiatric cases.

ETIOLOGY

The etiological factors contributing to the spectrum of malingering presentations are broad and related to much larger questions concerning the structure of and motivations within human nature. Issues of cognitive development and refinement, introspection, insight, accurate self-analysis, ego defense mechanisms, adaptation and survival, moral development, openness to self-disclosure, honesty, and the capacity deliberately to lie all play into an assessment of the mechanisms contributing to malingering. A basic knowledge of child and adolescent development, especially the capacity to lie, is useful. Malingering may be construed to carry a pejorative judgment, and so it poses significant professional and legal implications for the doctor if it is challenged. Hence, diligent, parsimonious, and perspicacious approaches to a patient's motivations are essential ingredients leading to fair and reasonable conclusions.

A thought-provoking 1994 study by Richard Rogers Ph.D., a recognized leader in the malingering field, proposed consideration of three explanatory models of malingering, using a prototypical analysis and subjects who were trained psychologists with extensive experience with malingerers. Etiological determinants showed that adaptational attributes such as making rational decisions based on expected rewards, attempting to cope with very difficult circumstances, weighing alternatives, trying to make the best of a bad situation, and being faced with an unsympathetic or

adversarial system were first-rank factors that explain the behavior of response styles. Antisocial and criminological factors were in the second rank, and a psychopathological model was favored by only 3.4 percent of the forensic experts. This study emphasized that the choice of explanatory model may have important implications for treatment recommendations. For example, the adaptational and pathogenic model carried with it the sense of a more positive motivation for recommending treatment.

The biopsychosocial model in psychiatry, both theoretically and applied to individual cases, may be one of the best ways to consider possible etiological factors until further research reveals more objectively verifiable causes. This includes data from the clinical interview and collateral information: biological factors such as temperament and medical condition, psychological factors such as defense mechanisms, coping, adaptive styles, and social factors such as current stressors, interpersonal, occupational, and financial considerations, medical and legal matters, and military obligations.

DIAGNOSIS AND CLINICAL FEATURES

For the psychiatrist, clinical presentation includes one or more interviews with the patient during which an interpersonal and social field is established. Within this context, all aspects of the patient's demeanor, and not only self-reports and response styles admittedly complex and often nuanced, are carefully observed and explored; some aspects of cognition, memory, social and emotional responsivity, and behavior may be tested. Issues around the manner (e.g., passive and helpless versus more aggressive and grandiose) in which the patient views and may express envy and greed—with their implicit spoiling and destructive characteristics—have clinically relevant value. Initially, what the patient produces may be considered raw data that only gradually undergo an evaluative and interpretive analysis and subsequent synthesis. All the patient's productions are initially taken at face value; some of these include sets of partial intangibles. Conclusions can be drawn only after diligent consideration and further collation using collateral information (e.g., neuropsychological testing, medical reports, laboratory investigations, past medical history, past psychiatric history, and psychosocial and legal history). Short of the only genuine gold standard for establishing malingering, which is the patient's full disclosure of deliberate intention to malinger, coding for malingering remains a matter of clinical judgment and professional opinion. A judicious, prudent, and parsimonious approach is to list the evidence suggesting malingering but not to reach a definitive conclusion unless the evidence is overwhelming and supported by clear-cut proof.

Projective Identification

Projective identification is a fundamental unconscious and implicit information processing mechanism that was first described by Melanie Klein in 1946 and expanded upon by Wilfred Bion during the 1960s and 1970s. This theoretical construct denotes the mobilization of partial aspects of an individual's unconsciously conflicted emotions, thoughts, and wishes that then become split off and intrusively projected (information transfers) into an unwitting receiver. Material communicational indicators, that is, response styles—language, gestures, demeanor, facial expression, even silence—are the medium of these data. The goal is a subtle yet aggressive manipulation persuading the receiver to believe the sender. The astute clinician, sensitive to the possible intersubjective meaning of this resonance, can parse out this important *additional* factor as a subtle source of subliminal data in assessing all malingering presentations, when formulating clinical judgments, and in treatment. Further aspects of *projective identification* are discussed in the treatment and management subsection.

Truth and Malingering

Malingering depends on the clinician's conception and assessment of truth —the truthfulness and veracity of the patient's clinical presentation, especially verbal descriptions of symptoms. This rests, in part, on the patient's openness to accurate self-disclosure. The assessment of malingering challenges the trust inherent in the doctor–patient relationship. Detection of true from feigned symptoms and disorders is made by comparisons to accepted diagnostic standards such as found in the DSM. The normal development of a child's reality sense and capacity to intentionally misdirect ("lying") is found in many texts on child development. Consistency in meeting standards, taken together with understandable atypical variations, reflects the fidelity of clinical assessment. Marked discrepancies, mismatches, deceptive shifts of diagnostic frame, and blatant excursions from established norms suggest deception. A high index of suspicion is raised when the patient demands medication, divulges symptoms too eagerly or dramatically, and makes blatant conditional threats of self-harm, violence, or litigation. Research has suggested that the following are representative detection indicators for feigned mental disorder: rare symptoms (a very important indicator), improbable combinations, symptoms that rarely occur together, spurious patterns, erroneous stereotypes, endorsing a very high proportion of symptoms (an important detection indicator), endorsing a wide array of minor psychological problems, reporting symptoms of grossly extreme severity, and evidencing greater impairment on reported symptoms versus observed behavior. Gross exaggeration is a detection indicator when its magnitude is strikingly amplified. When consciously deliberate intention and tangible gain can be assigned to illness deception, then a diagnosis of

malingering can be entertained.

Table 28.1–1.
Contexts in Which Malingering May Present

1. Avoiding responsibility: Military service, child or adolescent school avoidance
2. Criminal forensic contexts: Avoiding punishment, criminal liability, or seeking transfer from prison to hospital (the base rate incidence in criminal cases is about 19%)
3. Medicolegal contexts: To obtain financial gain, compensation cases (personal injury, 29%; disability, 30%)
4. Emergency department or primary care outpatient clinic visits to gain admission to hospital as may be found in homeless persons seeking shelter
5. Outpatient settings: Medical requests to obtain drugs such as controlled substances, opioids, minor tranquilizers, or psychostimulants (e.g., ADHD medications: methylphenidate, dextroamphetamine, etc.)

Diagnosing the condition of malingering is based on comprehensive considerations of the best evidence available and astute clinical judgment. When reaching clinical conclusions, particular attention to configurations of factors associated with malingering is useful. When psychological testing is used as collateral evidence, neuropsychological profiles are constructed using multiple tests, each tapping a different area. There are no established unambiguous criteria. From the patient's side, the issue of good faith, openness to self-disclosure, and well-intentioned honesty are presumed to be present but are examined for possible discrepancies. The extent to which such discrepant information is conveyed in what is suspected to be a deliberately deceitful manner lends itself to a high index of suspicion of malingering. Table 28.1–1 summarizes contexts in which malingering may present itself.

CONTEXTS IN WHICH MALINGERING MAY PRESENT

Avoiding Responsibility

Avoiding responsibility has developmental and social determinants. The current presence of international unrest and war makes the context of military service a continuing impetus for some to use malingering to avoid personal threat. Avoiding responsibility may also appear in child and adolescent contexts in the frequently encountered condition of school avoidance. Although an anxiety disorder may be the underlying cause of such behaviors, many youngsters also intentionally feign illness to avoid going to school. Malingering may not be an essential feature in such cases because truancy may be associated with more salient disorders such as conduct disorder, learning disorders, or understandable environmentally based problems such as bullying or an impaired family situation in which neglect, abuse, or both exist.

Criminal and Forensic Contexts

Criminal and forensic contexts in which adults, as well as adolescents, find themselves in situations in which malingering offers an option to avoid unwanted legal prosecution or confinement. Feigning psychiatric illness may offer a person convicted of a crime the opportunity to go to a psychiatric hospital rather than to jail. The issues surrounding criminal behaviors and psychiatric illness are complex and require the skills of those trained in forensic psychiatry and psychology for appropriate evaluation and recommendations. Malingering has been seen in immigrants caught for illegal entry into a country and threatened with deportation back to their country of origin.

A 15-year-old boy was referred by the court to a juvenile justice psychiatric evaluation clinic. He was being held at a detention center for multiple charges: possessing cannabis and crack cocaine, setting fire to his family's mobile home, expressing violence toward a parent, and resisting arrest. The arresting officers described him as uncooperative, at times belligerent, and communicatively unclear. They had known him for several years before this episode since he and his mother had involved authorities because of an allegation of physical abuse by one of his mother's live-in boyfriends. At the point of arrest, he was brought to a local emergency department, where he was examined and found to be in good health; he became quiet and did not appear to exhibit any clear-cut psychiatric symptomatology. Shortly thereafter he was brought to the clinic on a daily basis for a series of psychiatric and psychosocial examinations along with psychological testing. The patient was reticently compliant and demonstrated minimal expressive communication. He did admit to illicit substance use over a period of several years, a history of school avoidance and failure, and a chronic history of family conflict. When asked about the fire setting, he did not reply. Family interviews revealed a chaotic household; the biological father was incarcerated, the mother had multiple male companions and a history of depression with suicidal attempts, and there was an overall lack of structure and home supervision. Psychological testing showed a full-scale intelligence quotient (IQ) of 75, with pervasive deficits in adaptive and academic functioning, and projective tests suggested possible psychosis but no depression. He and his mother wanted to avoid incarceration and emphasized a history of abuse. The patient then claimed to be depressed; his mother vociferously supported this. From all the available past history, current interviews, and psychological testing, diagnoses of psychotic disorder unspecified (substance abuse disorders), borderline intellectual functioning, and possible PTSD were made. Although both mother and son intentionally tried to establish a primary diagnosis of clinical depression alone, the evaluator did not concur. Adding malingering to the diagnostic profile, moreover, was thought to be unhelpful in the context of this already complex group of psychiatric disorders. Recommendations were made for the long-term therapeutic residential placement and consideration of psychotherapy and psychotropic medication for PTSD and psychosis.

Medicolegal Contexts

Medicolegal contexts can suggest malingering when issues of financial compensation are at stake. This may involve personal injury cases, as well as disability claims for psychiatric or developmental problems. Because these complex cases have salient medical and neurocognitive components, their assessment requires the combined efforts of medical specialists to assess neurological and musculoskeletal features, as well as psychiatric and psychological evaluations.

Police brought a 16-year-old boy to the emergency department. He was wildly out of control, moving about in a disorganized, hyperkinetic fashion and speaking incoherently; thinking showed loose associations. Initial physical examination was unremarkable, but laboratory investigations showed elevated liver function tests; drug screening was negative. The patient's mother said that he had no prior medical or psychiatric history and that she was unaware of any indication of substance abuse or cigarette smoking. The patient was admitted to the psychiatric unit, where he was moderately agitated, demonstrated a catatonic-like excitement, and reported visual hallucinations. Atypical neuroleptics were begun; after several days, his motor behavior began to normalize, and he was more coherent. His cognitive functioning cleared but was slow. He denied the use of illicit substances but said that he had recently begun a summer job in a small paint factory; this appeared to coincide with his symptoms. A tentative diagnosis of inhalant-induced psychotic disorder was made. The patient was discharged on no psychotropic medication and seen 1 week later. He appeared to be stable and in no acute distress. Thinking was clear with no reports of hallucinations. Sleep, appetite, and energy level were normal. Blood work was within normal limits at this time. He returned to school in the fall and was reported to be doing as well as he had been doing prior to his decompensation. His mother then requested that the hospital physicians support her contention that her son was currently disabled, thereby certifying him to be eligible for long-term financial compensation. After further interviews, no evidence of changed or limited functioning from prehospital baselines could be found to support disability claims. The mother's attorneys brought the matter to a hearing. It was agreed that the paint factory would make a small settlement. This case demonstrates how patients may try intentionally to shift evidence of a real disorder and its resolution toward exaggerated claims of continued impairment.

Emergency Facilities

Emergency departments and primary care facilities may be the only resource to which those who are socially and occupationally impaired have relatively easy access. Such persons may have complex chronic medical and even psychiatric illnesses, which nonetheless do not warrant immediate hospitalization. Careful diagnostic assessment of emergent, urgent, and longer-term needs is critical. The assistance of social services, particularly if immediate shelter is needed, can be an important resource. Whether malingering, in addition to other problems, is believed to be present, clinical judgment must dictate the appropriate disposition of these cases.

Outpatient Settings

A common situation in which malingering may present itself can be seen in typical outpatient practice. This obtains not only for the psychiatrist but also for all physicians. Patients may malinger about pain symptoms that are either untrue or deliberately exaggerated. The request for controlled substances such as opioids in excess of what the physician deems reasonable always needs careful attention. Patients may be habituated to or dependent on benzodiazepines and so demand, ongoing prescriptions or excessively high daily doses. The physician's clinical judgment again must weigh reasonable risks and benefits, especially if malingering is suspected. A not infrequent request by adolescents and by some adults is for psychostimulant medications such as methylphenidate and its derivatives and dextroamphetamine and its derivatives to treat cases of supposed

attention-deficit/hyperactivity disorder (ADHD) or attention-deficit disorder (ADD). Because these disorders are frequently diagnosed, careful clinical assessment is needed to discern legitimate cases from those persons that wish to obtain controlled substances for purposes such as weight loss or nonmedically justified stimulation, diversion, and abuse. It has been almost commonplace for adolescent girls and boys to be overly concerned about body weight and to request psychostimulant medications under the ruse of suffering from ADHD when, in fact, the underlying intent is to use the psychostimulant drug for effective weight loss. Child and adolescent psychiatrists as well as pediatricians face this issue on an almost daily basis. In addition, careful histories and assessments of the family context, particularly members with illicit drug histories or criminal backgrounds, are essential to rule out abusive use or diversion into the street for illicit drug trade.

A 35-year-old attorney requested a psychiatric evaluation for the treatment of what he claimed was his adult ADD. This was the third such consultation that he had requested in the last 6 months. The patient was a middle-class male, married with two children, and a practicing attorney for the last 6 years. He had a reported past history of intermittent, moderate depressions since age 20 years, for which he received periodic brief psychotherapy and antidepressant medication. He stated that he had not felt depressed for the last 5 years, although he described himself in terms he said his wife used: "a man who fears success." This, he explained, was related to multiple attempts at jobs in various law firms that quickly ended with his impulsive resignation. He now felt that he was troubled with difficulties concentrating, staying focused, and completing tasks that required attention, concentration, and perseverance. He said that he was able to diagnose himself and believed that he suffered from ADD. He sought a prescription for a stimulant medication.

After several sessions of careful history taking, diagnoses of anxiety and mild mood instability were suggested, especially based on the past history. After agreeing with the psychiatrist's "skillful insights," as he described them, he continued to assert that the assessment was essentially wrong and missed understanding the core of his dilemma. He said that, although partly right, the assessment was "perfectly flawed." He insisted, however, that he would come for further sessions to "show the doctor" how wrong he was, and that the putative ADD required a prescription that "only the doctor had the power and authority to order. Whether or not I meet criteria for ADD, I feel I can benefit from the drug." Personality disorders were considered, but such diagnoses were deferred.

During the course of about six sessions, he behaved in a similar and repetitive fashion. When it became clear that no psychostimulant medication would be prescribed, he said, "You just won't open up and listen to me. It's really very clear, I know myself better than you do. Even though I didn't go to some fancy medical school, I'm competent enough to see things as they are." He would then add, "I guess we're on different wavelengths." Any sort of interpretive attempt to clarify his presentation, however tactful, was thwarted, and the patient declared that an "impasse" had been reached and that he would not return.

CASE DISCUSSION

The nature of the patient's motivation was complex. Characterological issues were principal. Envy of the therapist was an obvious feature. He certainly intended to persuade the psychiatrist both of his ADD and his need for drugs, but the extent to which he feigned or believed these to be true is unclear. His almost arrogant insistence on his independently arrived at, correct self-diagnosis and his inability to consider another opinion, along with the constant comparisons he made between himself and the psychiatrist, evidenced rigidity and a quality of recalcitrance that in this context suggested pathological, unconscious envy. His partial acknowledgment of the therapist's skill, yet his insistence on the assessment as

ultimately being "perfectly flawed," pointed to his attempt at envious spoiling of what was sensed to be partially skillful and insightful—that is, ideal to some degree. This patient's aggressive attempt to influence the psychiatrist can be viewed as a display of defensive omnipotence and the use of *projective identification* as a way aggressively to gain active control to ward off and defend against feelings of inferred inadequacy and helplessness. His choice to leave the therapeutic field due to what he called an "impasse" further attested to the usual blocking of forward progress that results when excessive envy and narcissism are major underlying factors in the therapeutic impasse. The boundary between malingering and factitious disorder as seen in this case is thin.

CLINICAL PRESENTATIONS OF MALINGERING

Physical Symptomatology

Pain and Medical Presentations. Malingered symptomatology may take various forms. Table 28.1–2 summarizes clinical presentations of malingered symptomatology. They may be divided into physical symptomatology (pain, medical with cognitive incapacity) and psychological symptomatology. Whereas factitious disorders often present with physical symptomatology and signs, malingering appears in psychiatric settings to present with psychological symptomatology suggestive of psychiatric disorders. Perhaps the presentation of pain and distress in suspected malingering is most often encountered in contexts in which financial or disability compensation is at stake. This can be related to physical injury or degenerative disease states. Forensic and litigation disputes may be venues for this. If pain is malingered, it may typically take the form of an embellishment of the severity of pain and subsequently what appears to be significantly impaired functional capacity. Ordinarily, some degree of real pain is present, but the excessive degree to which it is intentionally exaggerated may suggest malingering. Because this is an extremely complex area and one in which malingering has been shown to be less prevalent than once thought, such presentations need careful consideration. When physical symptoms are in question, medical specialists are needed to rule out neurological, infectious, musculoskeletal, and other primarily somatically based complaints. In addition, one must carefully discern the presence or absence of pain disorder or the condition of pain associated with a general medical condition, the code for which is specifically selected based on the associated medical condition or the anatomical site.

Pseudoseizures. Pseudoseizures (aka, psychoseizures, psychogenic nonepileptic spells, hystero-epilepsy) resemble epileptic seizures. Prevalence is suggested to be about 5 percent of presentations in emergency departments. Patients exhibiting pseudoseizures are unusually suggestible. They may also have true seizure disorders or a family history of epilepsy. Although they resemble tonic–clonic convulsions, no tonic phase is present; there is an asynchronous thrashing of the limbs typically lasting

more than a few minutes, unlike in true seizures. Consciousness may be normal or lost, and typical postictal phase or loss of continence, tongue bite, and headache are variable. Laboratory studies show no electrocerebral changes. Serum prolactin, ordinarily increased between 15 and 30 minutes after a true tonic–clonic seizure, remains unchanged from baseline. Antiepileptic drug treatment has little, typically no, therapeutic effect. In some cases, hysterical conversion reactions associated with acute emotional stress or chronic physical and sexual abuse produce pseudoseizures. Pseudoseizures also have been seen in malingering. When there is clear evidence of incompatibility with neurological disease and when malingering has been ruled out, DSM-5 terms pseudoseizures to be a functional neurological symptom disorder (conversion disorder).

Table 28.1–2.
Clinical Presentations in Which Malingered Symptomatology May Present

1. Physical:
   - Physical pain and incapacity due to injury or chronic medical condition
   - Pseudoseizures
2. Psychological:
   - Posttraumatic stress disorder
   - Depression
   - Amnesia
   - Psychosis
   - Neurodevelopmental:
   - Intellectual disabilities
   - Borderline intellectual functioning
   - Miscellaneous disorders: Attention-deficit/hyperactivity disorder; Ganser syndrome
3. Neurocognitive:
   - Traumatic brain injuries
   - Amnesia
   - Dementia
   - Delirium
   - HIV

Neurocognitive Presentations

Neurocognitive functions include levels of consciousness and all information processing abilities–attention, concentration, memory, reasoning, problem solving, executive functions, and language. Cognitive functions are closely linked to neural pathways and cortical networks. Impairments in these—DSM-5 "neurocognitive disorders (NCD)"—represent a decline from a previously achieved level of functioning. NCDs were referred to in DSM-IV-TR as dementia, delirium, and other cognitive disorders. They may result from central nervous system injury, moderate-to-severe TBI with significant declines in previous levels of performance, concussion whose traumatic effects may be transient, mild traumatic brain injury (mTBI), toxicity, disease, or deterioration. Clinical presentations

may be classified as major NCD or minor cognitive disorder and may present themselves as delirium, dementias, involve amnestic disorders (DSM-5 "major or minor NCD due to another medical condition"), headache, HIV, and so forth. They may also appear as a variety of intellectual and memory impairments as, for example, occurring after motor vehicle accidents or war trauma following concussion and TBI. Overlap with psychological symptoms is common. Trauma and injury events—head injuries—typically present themselves in the ruling out of possible malingering. Examples of typical venues are personal injury litigation, workmen's compensation, and disability claims. Neuropsychological testing and the assessment of symptom validity are essential in these cases. Cognitive assessment relies on the comparative measurement of responses: incorrect, partially correct, and correct responses. Estimations of valid effort, which may be difficult to assess, help in interpretation. Unlikely presentations, excessive impairment, and magnitude of test error are useful detection strategies when malingering is suspected. Pain, depression, and PTSD represent the most common differential diagnoses when considering mild TBI or modest decline in previous level of performance. Various forms of the Mini Mental State Examination can be used as a screen for dementia, although a number of factors may render results unreliable: imprecise testing of executive functions and performance variations based on patient's educational level and motivation. The expertise of an experienced forensic neuropsychologist is essential in determinations and includes the use of standardized scales assessing levels of cognitive response and diagnostic category.

Psychological Symptomatology: Clinical Presentations

Psychological symptomatology and the suggestion of mental disorders are presentations that the psychiatrist encounters, often by referral or consultation, in ruling out suspected malingering. A high index of suspicion is raised when symptoms are clearly atypical, grossly unusual, nontraditionally bizarre, extraordinarily exaggerated, embellished, extravagantly elaborated, or alleged to occur in the future if some incentive or personal gain in the present is not met. Laypersons, for example, may decide to exhibit dramatic symptoms or to portray some florid clinical picture based on what has been seen in the popular media. Professionals who malinger might consult psychiatric texts and attempt to feign a bona fide mental disorder by using every feature and criterion, listed as well as their precisely given time frames. Such direct textbook presentations rarely present in day-to-day clinical contexts. Detection strategies include close attention to the patient's description of when and how symptoms arose, as well as to their sequence. Clinical incongruities in symptom presentation—for example, an individual describing an anxiety disorder in a calm and composed manner at every interview—suggest feigning. Finally, markedly

atypical juxtaposition of symptoms may suggest malingering. Some examples of psychological symptomatology are discussed in the following.

Posttraumatic Stress Disorder. PTSD is a common presentation in which malingering may be a consideration. Cases involving personal injury, automobile accident, and those exposed to terrorism and traumas associated with military service and combat may present with features of PTSD. Combat-related and civilian PTSD needs to be differentiated. Although those in the military are prepared for the unexpected, they too may experience genuine feelings of helplessness, vulnerability, and surprise. Suicide patterns in the military have increased over the last decades. Military suicides are at their highest rate in 10 years. Twenty percent of national suicides are completed by veterans. Mega-organizational complexity, time pressures, and at times feeling overwhelmed to act efficiently and reliably may be contributing stress factors. More than half the suicides occur in those who have never been deployed. Those who have been exposed to combat may experience a variety of postcombat stress-related sequelae. PTSD is a factor and is also associated with the breakup of relationships and subjective feelings of isolation. "Moral injury," which is different from PTSD, has been described in those military personnel who experience feelings of guilt and shame that might be related to killing others or having friends killed in combat. Genuine PTSD has been correlated with the patient's minimizing combat as causative, blaming self versus others, experiencing dreams with themes of helplessness, general reluctance to actively speak about combat memories, feeling survivor guilt and shame and anger. Feigned PTSD diverts care from those military personnel who genuinely have suffered trauma and require services. Currently, genuine PTSD is common.

The most common motivation for malingering PTSD in the general population is thought to be financial gain. Symptom presentations in feigned cases by contrast often appear dramatic and highly embellished. Dramatic positive symptoms such as flashbacks and nightmares are emphasized. For example, nightmares might be described as relentlessly persistent, whereas traumatic nightmares often dissipate over time. Feigned experiences are reported in ways that are inconsistent with DSM definitions; they may reflect unrealistic caricatures seen in the media. Subtle features and generally incapacitating anxiety are often absent. For example, in an effort to avoid occupational duties, psychological and physical incapacity (e.g., disability) may be claimed. Often, on further examination, this incapacity is found not to extend to other, more pleasant activities, such as social events, recreation, travel, vacation, or sports. A thorough history may reveal past occupational difficulties or frequent job changes. The general demeanor of uncooperativeness, withholding information, belligerence, or personality disorders (e.g., antisocial,

borderline, or histrionic), may also suggest—though not confirm—malingering. A bland refusal to undergo treatment and to consider future re-employment is also suggestive. The use of all available collateral information is important with regard to substantiating or disconfirming self-reports and also detecting discrepancies.

Depression. Depression may be malingered. Because depressive disorder and depressive variations (e.g., melancholic, atypical, and bipolar) have been so extensively researched, the psychiatrist is usually perceptive in detecting faking in this area. For example, the malingerer may be extremely fluent in descriptions of "feeling depressed" and "not sleeping" but may not report the typical details of impairment (diminished specific hobbies; dysfunctional interpersonal relationships that include distressing envy, jealousy, guilt, and the destructive trends common to these; limitations in everyday activities of daily living; and altered vegetative functions) that the genuinely depressed person may report. Appetite and weight remain constant. The very loquaciousness that often attends such presentations is inconsistent with depression. In addition, malingered depressions may not pervade all aspects of the person's life and functioning. For example, the malingerer may continue to engage in going out to dinner, hobbies, and regular sports activities. Fatigue, loss of energy, degrees of anhedonia, and inappropriate guilt are absent.

Amnesia. Amnesia with its neurocognitive and emotional dimensions may be a form of malingering seen in forensic settings and is usually related to alleged crimes. It has been suggested that up to one-third of homicide perpetrators claim amnesia. Because this is a grave situation, specialists in forensic psychiatry in collaboration with other specialists are best suited to discern such presentations. Head injury after accidents also presents with degrees of amnesia. Such cases were historically referred to in the literature as *accident neurosis* or *compensation neurosis* and seen in a negative light until recent research suggested that genuine cognitive deficits might be present in large numbers of these cases. In assessing amnesia, highly selective, localized memory deficits, especially in the areas where the patient's motivation and liability are in question, may suggest feigning. Typically, genuine amnesic patients demonstrate impaired new learning and short-term memory, but procedural memory and premorbid semantic memory are preserved. If deficits are produced in immediate recall, digit span, autobiographical and self-memory, and new learning, this may suggest feigning, if the wider context is also corroborative. In genuine disorders (DSM-5 major or minor NCDs) with pathophysiologic concomitants, disorientation may be to place and to time. There may be lack of insight into circumscribed memory deficits, and the patient may explicitly deny the presence of severe memory impairment. No progressive

deterioration is seen; apathy and emotional blandness, however, are common. Chronic substance abuse and alcoholism need to be ruled out. Dissociative states and histrionic personality disorder tend to lessen the probability of malingering. In dissociative amnesic disorder, the patient does demonstrate inability to recall important personal information usually of a traumatic or stressful nature. Patients with true dissociative amnesia usually score high on standard measures of hypnotizability, unlike the case with malingerers. Antisocial personality disorder is a feature of malingering according to the DSM-5 guidelines and needs to be considered in forensic contexts. Of course, neurological examination and appropriate radiographic studies (e.g., electroencephalogram [EEG] and magnetic imaging resonance [MRI]) should be considered.

A 25-year-old female was seen for consultation on a forensic neurological in-patient unit. The question of whether malingering was present was raised by her neurologist. She had set fire to her boyfriend's trailer after a physical altercation during which he struck her on the head with a blunt object several days ago; she was admitted for observation and holding. Neighbors called police and firefighters. Trailer was demolished, and boyfriend sustained serious burns and was unable to speak with authorities. The patient never lost consciousness; she sustained minor injuries. Her neurological examination, EEG, CT, MRI, and blood work were noncontributory. The patient claimed amnesia for the event, and asked for an attorney. After extensive interviews and in the absence of a history of psychosis or mental illness, probable malingering was seriously entertained. When her boyfriend recovered after several weeks, he described her setting the fire. When she was told, she then "broke down" in tears, and admitted to the act.

Psychosis. Psychosis can be malingered—for example, in criminal defendants attempting to avoid trial or to alter prison sentences, in the homeless who present to the emergency department in an effort to gain entry into the hospital for temporary shelter and food, or in those whose motivation is to avoid some clear-cut accountability that they find undesirable, such as imprisonment for crimes. This may be a more common malingered presentation in less sophisticated persons. For the experienced clinician, it may not be very difficult to differentiate true psychosis, as, for example, in schizophrenia, or bipolar mania, from deceptive illness. Persons with schizophrenia are often withdrawn and minimally verbal, not open to fluent descriptions of their symptoms. Individuals feigning psychosis often produce reports of very dramatic, perhaps melodramatic, hallucinations with bizarre content that is more characteristic of science fiction than what is ordinarily reported by those with genuine psychosis. Hallucinations may be uncharacteristically not accompanied by associated delusions. Feigned hallucinations are reported to be constant, not intermittent; irresistibly commanding to action with no chastising content, always terrifying rather than manageable; and, if visual, seen in black and white or conversely in dramatic colors in cinematic fashion. Malingerers may speak of delusions in a forthright, open manner, unlike the guardedness often accompanying descriptions of genuine

delusions. These persons may describe their delusions as having all-or-nothing features rather than the subtler and continuous tone characteristic of truly psychotic delusions. In addition, malingerers feign only the contents that accompany psychotic states and often demonstrate no irregularities in the form (e.g., derailment, looseness of associations, and disorganization) of their thinking processes, another important characteristic of psychosis. Also, negative symptomatology (e.g., affect blunting, avolition, and alogia) may be absent. Chronic thought disorder is typically distressing.

A 35-year-old male pharmacist was referred to an outpatient clinic by his internist for medication evaluation of his long-standing bipolar disorder. He was in apparently good health and working more than 60 hours per week but was experiencing anxiety, dysphoria, and increasing disorganization. He reported that the last year had been turbulent in his relationship with a girlfriend of 3 years. Sleep, appetite, and job performance were beginning to suffer. The patient had been on lithium and an atypical neuroleptic for 4 years with good control of symptoms. When his symptoms had recently escalated, the patient suggested to his internist that his ADHD and his bipolar disorder were not being properly addressed. A once-per-day, long-acting psychostimulant was then added. Although the patient stated that he felt that he could function more efficiently on the stimulant, his observable behaviors dramatically deteriorated, speech became pressured, and thinking became progressively more disorganized. When he demanded benzodiazepines for anxiety, escalating irritability, and insomnia, his prescribing physician made the referral. At interview, the patient was thin and agitated; he appeared to be highly intelligent and articulate. He at first asked for increases of his lithium and neuroleptic, but then demanded increases in his stimulant and also the addition of a benzodiazepine. Then, it was apparent that he was hypomanic and that his judgment was becoming progressively impaired. After several sessions of attempting to target the escalating bipolar presentation and obtain blood work (which the patient failed to procure), a recommendation for inpatient hospitalization was made. He protested and reiterated that his ADHD was his primary problem and that stimulant medication was the only answer. A second psychiatric opinion was obtained, and it confirmed bipolar disorder with little or no evidence of ADHD. Again, hospitalization was strongly recommended, at which point the patient abruptly left the clinic. Attempts to contact him failed. After several weeks, it was discovered that he had left the state, and his whereabouts were unknown. Six months later a family member contacted the clinic to request records, stating that the patient had been found dead in suspicious circumstances and that an investigation was in progress.

Intellectual Disabilities. Serious cognitive limitations, now referred to as "intellectual disabilities," must be considered and identified. Intellectual disabilities connote substantially reduced cognitive capacities, that is, a significantly limited age-appropriate capacity and use of thinking skills, memory, reasoning, abstraction, judgment, and decision-making. Such intellectual limitations have traditionally been referred to as ranging from the extreme of what had previously been termed "mental retardation" to well below average or borderline intellectual functioning. These are not declines but relatively stable levels of cognitive functioning. It is also not very common for an adult to feign such conditions as intellectual disability because past history, especially of academic performance, vocational functioning, activities of daily living, and social adaptation, ordinarily appear in quite distinct ways as substantially impaired over long periods in

genuine cases. These impairments pervade all sectors of the personality and often require obvious caregiving whether in childhood, adolescence, or adulthood. The clinician may be presented with a case of malingering by proxy, which suggests that the real malingerer (usually a parent) is alleging that his or her child has substantial intellectual disabilities that demand financial and other compensation. In these cases of malingering imposed upon another, careful assessment of the past history of both child and caregiver is essential. Insofar as differential diagnoses to be considered, dementia (not formally an "intellectual disability") represents a significant intellectual decline from a previously higher cognitive level; there is a decrease in expressive language and marked perseveration; this is usually not a malingered condition. A person with mild dementia (mild NCD) may erroneously hold a false belief in a nonexistent illness and claim monetary compensation; this is not considered malingering.

Borderline Intellectual Functioning. Borderline intellectual functioning is a condition that denotes an IQ between about 70 and 85 accompanied by significantly impaired adaptive functioning. Overall functioning is well below average. Such persons often present as *slow learners* who may appear to be feigning their cognitive limitations in an effort to achieve an external gain—needed help from others. The often-genuine needs of these persons should be differentiated from the deceptions of malingering. In addition, persons of any intelligence level may innocently attribute an unlikely causative factor (e.g., allergies, job stress, or Lyme disease) to what they genuinely believe is a disorder. Such innocent misattributions should be distinguished from consciously motivated malingering. Cognitive decline, however, after head injury, toxin exposure, or chronic substance abuse suggests a clear traumatic precipitant. Past history, medical examination, and laboratory screening for illicit substances and liver functions can help rule out legitimately understandable causes of true illness even if the patient misattributes cause and effect.

The single mother of an 18-year-old boy just graduating from a residential school for adolescents with academic, behavioral, and emotional problems requested documentation verifying that her son was disabled on the grounds of learning disabilities in order to obtain financial assistance for entrance into college. The mother was an immigrant from Europe who had been living in the United States for more than 10 years; the father had died in an automobile accident 5 years previously. There were no other children, and mother was unable to work due to disability secondary to depression. The boy had been enrolled in the school about 3 years previously because of school failure. Psychological testing showed his intelligence to be below average (Full scale IQ 80), and there was no evidence of primary attentional problems or specific learning disorders. His academic failure was deemed due to primary situational conflicts between him and his mother, in addition to probable borderline intellectual functioning. For the last 3 years, his school performance and behavior were approaching normal parameters, but his relational situation did not improve. After comprehensive retesting and assessment of psychological and social factors, it was judged that he was not disabled. While mother's request may have been well-intentioned, it exaggerated her son's level of impairment with the goal of procuring resources based on partially plausible

data. Attendance at a supportive college that would accommodate his learning needs away from home was suggested.

Ganser Syndrome. Ganser syndrome is classified as a DSM-5 other specified dissociative disorder. It is denoted as the giving of approximate answers (*Vorbereiden*) to questions when not associated with dissociative amnesia or dissociative fugue. Dissociative disorders connote disruptions in the integration of consciousness, memory, identity, or perception of the environment. Ganser syndrome was originally described in prisoners who presented as confused. Although the utility of this diagnostic classification has been questioned, such presentations require ruling out genuine intellectual disabilities, borderline intellectual functioning, and forms of delirium, as well as malingering. Considerations of both inferred motivation and assessing the prospect of external gain need to be taken into account.

Attention-Deficit/Hyperactivity Disorder. ADHD is a commonly made diagnosis. It is characterized by significant inattention, hyperactivity, impulsivity, several symptoms before age 12 years, and functional impairment in two or more settings such as in school and at home. Task avoidance and low motivation, which may be primary in such presentations, must always be ruled out in assessing attentional difficulties. Because one aspect of the comprehensive educational, psychosocial, and psychopharmacological treatment of ADHD requires Class II controlled psychostimulant medication, adolescents and adults have been known to feign symptomatology to gain these drugs for their perceived pleasurable and excitatory effects, which is drug abuse, or for diversion and financial gain. In *malingering by proxy*, a caregiver might attempt to persuade the physician that the child or adolescent has ADHD symptomatology in the known absence of such to procure these controlled substances for personal use or for illicit diversion.

An 8-year-old boy was brought to the clinic by his mother for medication management of his 2-year history of ADHD, which had been medicated by the family pediatrician until then. The referring physician reported that he was unwilling to continue to prescribe controlled substances because the mother had been unwilling to involve the child or herself in psychosocial interventions that included psychotherapy, parent management skills, and adjunctive school interventions. Mother and child were interviewed both together and alone on several occasions. The mother reluctantly agreed to the release of school records. Previous psychological and academic testing showed the child to have IQ scores that were low but within normal limits and there was adequate performance in all subjects except a clear-cut problem with reading. School personnel regarded this as probable dyslexia, not ADHD, and appropriate remediation was in place. The child was significantly below average for height and weight. He had multiple facial grimaces and tics and a history of insomnia and difficulty awakening in the morning. His medication regimen included dextroamphetamine in high doses and clonidine at night for sleep. His mother also requested that the clinic verify the child's need for disability compensation based on his ADHD. The multidisciplinary clinic team assessment included social work home visits and an intensive assessment of the mother's history. It was found that

she was on probation for a history of financial scams, writing bad checks, and involvement with the illicit drug trade. She gave informed consent for the release of her probation records. It was also discovered that state social services were in the process of investigating alleged child neglect. Before the psychiatric evaluation was completed, the clinic was informed that mother had been arrested for possession of and dealing in illicit substances, some of which included stimulant medications. The child was put in custody of his grandparents. Recommendations were made for comprehensive psychosocial services for the child and his grandparents, who agreed that all medications should be tapered so that a new baseline of overall functioning could be obtained.

Child and Adolescent Illness Falsification

The importance of taking into account the developmental characteristics of the child or adolescent when considering whether genuine illness is present cannot be overemphasized. Because individuals younger than the age of 18 years fail to meet some of the guidelines used to code for malingering (e.g., antisocial personality disorder), the phrase *illness falsification* is best used to denote situations in which youth appear to feign a medical or psychiatric illness having some ulterior personal gain as a goal. Often, such *illness falsification* is associated with symptomatic lying in this age group. In addition, lying in youth may have temporary adaptive value as, for example, under the duress of threatened abuse and trauma.

Does Illness Falsification Present Differently at Different Developmental Stages? In childhood, up to about 12 years of age, separation anxiety, school avoidance, and PTSD may be the leading disorders associated with a child's feigning illness (e.g., stomachache, headache, forgetfulness and intentional suppression of information of abuse, pseudoseizures) to obtain a desired goal such as staying home from school, avoiding consequences for negative behaviors, or wishing to receive gifts or increased attention for being sick. Children who are bullied in school might also use *illness falsification* to avoid exposure to the context of trauma. Symptom coaching, sometimes a consideration when a child falsely attributes illness to a parent, may occur in the context of an acrimonious divorce and child custody dispute. One parent may coach the child to describe the other parent in an unfairly negative light. This is an example of *illness falsification* by proxy.

In adolescence, about age 13 to 18 years of age, the emotions of anxiety, fear, and guilt typically prompt the teenager toward *illness falsification*. Anxiety and fear of attending classes may be related to anxiety disorders, social phobia, autism spectrum disorders, and bullying. School avoidance may result. As is discussed in the section on attentional problems, it is not uncommon for adolescents to claim ADD and ADHD as an excuse to rationalize both poor motivation and intentional task avoidance and so excuse excessively poor school performance. Juvenile offenders, often carrying a diagnosis of conduct disorder, which strongly connotes the

presence of lying and deception, have been noted to use *illness falsification* to avoid sentencing to jail, to lighten sentencing severity, or to divert their adjudication from jail to the mental health system for rehabilitation and treatment.

Antisocial Personality Disorder and Malingering

The *criminological model* as described by Richard Rogers in 1990, 1994, and 2008 proposes that this DSM-articulated framework is used, perhaps inaptly, in Rogers' view, to explain the primary motivations underlying malingering. It includes the diagnostic variable of antisocial personality disorder, the contextual variable of a medicolegal evaluation, and two presentational variables—uncooperativeness and discrepancies with objective findings. The strong association between antisocial personality and deceptiveness is emphasized in this implied risk model. Although there are obvious correlations between these, some studies have questioned their validity. The prejudicial role of confirmatory bias must be taken into account in all assessments.

The Nonmalingerer Dilemma. A very important consideration in any assessment of malingering is the need to be aware of the patient whose symptomatic presentation is real, that is, genuine and not feigned. Ordinarily such patients have genuine disease processes, bona fide diagnoses, obvious impairments, limitations, and restricted functioning. Such functional incapacity typically affects all areas of their life, including home, occupational, and social spheres. Their biopsychosocial functional capacity is significantly diminished, and true disability leaves them handicapped. The context in which this appears may be work related, injury related, associated with degenerative disease, or a mental disorder.

In the process of assessing a patient's genuine functional incapacity and current needs, the extensive and rigorous interviewing, at times perceived as an interrogation, and the implicit level of accusation that questions the patient's credibility and forthrightness are often sources that strain the patient's already compromised emotional resilience. The integrity of the working relationship between examiner, caregivers, and patients may become unduly encumbered. In such cases, iatrogenic factors may appear in the patient's presentation and confound correct evaluation and management. It is here that the adaptive component of the individual's resources may be mobilized. Considerations of these factors are useful guidelines to help temper any tendencies to conduct evaluations in an inappropriately aggressive manner that, in itself, may be experienced as a trauma.

CODING MALINGERING

When malingering is considered to be present, it is useful to indicate its

inferred type.

- *Full malingering* denotes complete fabrication.
- *Pure malingering* denotes the feigning of disease when it does not exist.
- *Positive malingering* denotes the feigning of symptoms that do not exist.
- *Partial malingering* denotes the exaggeration of existing symptoms or the statement that extinct symptoms are still present; this level of malingering may be common.
- *False imputation* is a more nebulous type of malingering. The malingerer deceptively shifts the frame of the diagnostic focus by ascribing actual symptoms to a cause consciously recognized to have no relationship to those symptoms. Such obfuscation darkens and obscures diagnostic clarity, yet may be often overlooked.

It may also be helpful to place malingering in a context that connotes degree of certainty:

- Indeterminate
- Strongly suspected
- Provisional
- Probable
- Deemed likely present

The American Medical Association has stated that confirmation of malingering is extremely difficult. Table 28.1–3 summarizes a range of degrees of certainty in coding for malingering. Finally, the inadvertent misattribution of linking symptoms with unrelated causes, as, for example, found in those with cognitive limitations, should not be considered malingering.

Table 28.1–4 summarizes DSM guidelines indicative of malingering and provides a clinical decision model for coding the condition. Although the presence of these entire threshold-screening indicators is not essential, factors 1 and 2 are fundamental and should be accompanied by some of the other typical features. Although DSM-5 does list the presence of antisocial personality disorder in these coding guidelines, some have questioned the strength of this association. Others have called attention to the connection between the clinical findings of deliberate misrepresentation and lying, for instance, commonly seen in persons with antisocial personality disorder, as well as in persons with borderline personality disorder, and similar findings of intentional deception in malingering.

Conclusions should derive from five considerations:

1. Is the clinical evaluator acting as a short-term consultant in a setting that implies a medicolegal or, perhaps, an adversarial context?
2. Is the patient's presentation reasonably consistent with the presentation of genuine illness?

3. Is there a medical or psychiatric disorder that plausibly explains the signs and symptoms presented?
4. Are the final diagnostic conclusions based on evidence of consistency or on marked inconsistency?
5. What is the best professional opinion about the patient's intentions and goals?

Table 28.1–3.
Degrees of Certainty for Coding Malingering

1. *Indeterminate*: Findings suggest the possible presence of malingering according to DSM-5 guidelines but are too ambiguous, insufficient, and unsupported by collateral evidence.
2. *Strongly suspected*: Findings are consistent with basic DSM-5 guidelines but lack typically associated features and contexts; collateral support is equivocal.
3. *Provisional*: Findings are compatible with primary DSM-5 guidelines, although they are partial or too brief; yet a strong presumption exists that further indicators will be brought to light over time.
4. *Probable*: Findings are consistent with basic DSM-5 guidelines; associated features and contexts are present, but there is a lack strong of confirmation from collateral sources, especially psychological testing.
5. *Present*: Findings are consistent with basic DSM-5 guidelines, associated features and contexts, and are supported by a preponderance of evidence from most collateral sources. If expressed, acknowledgment by the subject is definitive.

Table 28.1–4.
Clinical Decision Model for Coding Malingering

1. Intentional production of false or grossly exaggerated physical or psychological symptoms.
2. Motivation by external incentives:
   - Avoiding military duty
   - Avoiding work
   - Obtaining financial compensation
   - Evading criminal prosecution
   - Obtaining drugs (e.g., for recreational use, abuse, financial gain, and diversion)
   - Adaptive behavior (e.g., feigning illness while a captive during wartime)
3. Medicolegal context
4. Marked discrepancy between a person's claimed stress or disability and the objective findings
5. Lack of cooperation during the diagnostic evaluation or in complying with treatment
6. The presence of antisocial personality disorder

Genuine medical and psychiatric illnesses have several features in common. Reported symptoms are typically not in excess of signs and measurable findings. Real illnesses demonstrate a regular consistency across an adequate database that includes findings from different examiners, varied contexts (e.g., work, home, and recreation), and clear presence when the patient is distracted. Disability specialists, for example, weigh these heavily in impairment determinations. It is understandable that both forensic examiners and clinicians may be reluctant to give an opinion of malingering for a variety of reasons, which may include

litigation fear, concerns about error, and the pejorative nature of the term. Partial malingering adds to the condition being far from clear-cut. Careful documentation of objective findings and facts in a nonpejorative manner is essential.

COLLATERAL DIAGNOSTIC DATA

The finding of malingering may also be supported by obtaining information from collateral sources. These include data from interviews with people who know the patient, relevant records, and psychological testing. Informed consent issues need to be considered and documented. Interviews with family members and significant others, employers, past therapists, jail or prison personnel, and probation or parole officers may provide important information. Past medical records, psychiatric records, occupational and educational reports, court evaluations, and prior disability claims are also useful. There are no physical examination techniques that have been demonstrated objectively to prove the presence of malingering. Internists and family practice physicians stress, however, that warning flags include persistent noncompliance, striking inconsistency between physical findings and stated symptoms, and a referral from an attorney or an insurance company. Although the psychiatrist has an important clinical role in an overall assessment, the collaborative efforts of a team of medical, psychological (e.g., using neuropsychological testing), and legal specialists may be needed to provide corroborative evidence and hence a consensus to support psychiatric findings.

Psychological testing involves the use of standardized psychometric instruments by a trained and experienced clinical psychologist. Such tests are often used in the assessment of competency to stand trial, child custody evaluations, criminal responsibility, violence risk assessments, and personal injury litigation. To be admissible in legal contexts, tests must show validity and reliability, as well as adequate sensitivity to identify true positives and adequate specificity to identify true negatives. These federally recognized criteria are known as the *Daubert* standard and are derived from the 1993 U.S. Supreme Court case *Daubert v. Merrell Dow Pharmaceuticals, Inc.* Some of the more commonly employed tests with these features used to detect malingering will be reviewed. It is acknowledged that there is no one gold standard that is considered definitive. The use of multiple test measures and other collateral data aids in providing converging supportive evidence toward the finding of malingering.

The Minnesota Multiphasic Personality Inventory-2 (MMPI-2) is one of the most commonly employed instruments used in detecting malingering of mental illness. Some of its subscales are particularly useful. The F scale detects symptom exaggeration or severe psychological disturbance, deliberately exaggerated psychopathology. It is considered to be the best

available discriminator of honest versus malingered symptom profiles. It is not sensitive, however, to detecting neurocognitive dysfunction. The K scale is considered good at detecting *faking good*, or minimizing what the respondent perceives as incriminatingly negative symptomatology; this is mildly negative symptomatology that an ordinary person would openly admit experiencing. The F-K scale indicates a tendency to exaggerate symptoms. The F-psychopathology scale detects truly rare symptoms, and the D scale and FBS scale expose erroneous stereotypes of illness. MMPI-2-RF (2008/2011) is the restructured version currently used.

The Wechsler Intelligence Scales also are used to determine strengths and weaknesses in areas of intellectual functioning from verbal, perceptual reasoning, working memory, and processing speed perspectives. Gross abnormalities in intellectual level such as forms of intellectual disability and striking discrepancies in subtest areas can be detected. Wechsler Memory Scales are also frequently used. Results are supplementary to the other tests of malingering described in this section.

The Structured Inventory of Reported Symptoms (SIRS), a very popular instrument gaining in use, is used to test for feigned mental disorders and truly rare symptoms. Some consider it the best measure or gold standard of whether someone is faking psychotic symptoms. The two underlying dimensions in this test are the Spurious Presentation, which is used to assess symptom constellations that are unlikely to be true under any circumstances, and the Plausible Presentation, which is used to assess some symptoms that genuine patients are likely to endorse. SIRS-2 (2010) is currently used. Another instrument is the Personality Assessment Test (PAI), which is also used to test for truly rare symptoms and unlikely patterns of psychopathology. It can be completed in less than 1 hour. The Validity Indicator Profile (VIP) is a psychometric instrument that uses verbal and nonverbal items to test for cognitive malingering; it takes less than 1 hour. The VIP contains very easy and difficult items. Malingered profiles show easy items missed and difficult items responded to at chance levels by malingerers. Nonmalingers typically pass these items. It should be noted that such tests are not suitable for those with significant intellectual disabilities because such cognitively limited persons may not try as hard or may give up easily and produce what are recognized to be careless and invalid profiles.

Task avoidance, low motivation, and poor effort must always be considered in assessing the results of psychometric tests. Any test result, for that matter, may be wrongly skewed in the presence of real pain, headache, lack of sleep, fatigue, and so forth. Not having subaverage cognitive ability and substantial neurocognitive impairment, the malingerer may produce deceitful reports by using dishonest, invalid, and intentionally motivated poor effort in supplying data.

The Test of Memory Malingering (TOMM) is a newer, more

sophisticated malingering test. Also widely used, it assesses exaggerated memory complaints and the dissimulation of cognitive impairment, and it detects subtle forms of malingering. Its classification is based on the "floor effect," which is a testing strategy that relies on the nonmalingering individual's easily passing most test items. Studies of its validation indicate excellent specificity and modest sensitivity. It has a very high accuracy rate, and its validity has been supported in a forensic psychiatric population. The Millon Clinical Multiaxial Inventory (MCMI-III) is a test of malingering that detects tendencies to overstate emotional and personal problems. The Fifteen Item Test (FIT) is a test of recall; it evaluates malingered neurocognitive deficits. The Miller Forensic Assessment of Symptoms Test (M-FAST), a relatively new instrument, uses a structured interview and is considered most efficient at distinguishing feigned from bona fide schizophrenia. This tool is a screening device with highly validated studies confirming that it has sound psychometric properties. The majority of its items represent psychotic or highly unusual or improbable symptoms.

The use of the polygraph or lie detector, drug-assisted (e.g., amobarbital) interviews, and hypnosis are not considered standardized in an adequately scientific fashion and so are not ordinarily used to detect malingering. Ultimately, the only acknowledged gold standard that is undisputed in confirming malingering is full disclosure by the subject's stating a deliberate intention to malinger.

DIFFERENTIAL DIAGNOSES

The condition of malingering may coexist with true mental disorders, such as major depressive disorder, anxiety disorders, bipolar disorders, and personality disorders. Careful assessment must discriminate genuine mental disorders and personality disorders from the pretense of malingering. Presentations are usually complicated and typically not straightforward. Often, more than one diagnosis and condition may exist together. In addition, one must screen for intact cognitive status and rule out intellectual disabilities and dementias because these may inadvertently cause the patient incorrectly to describe what seem to be symptoms. Such mistaken beliefs do not carry the motivation of intent to deceive and so must be distinguished from malingering.

Malingering has been contrasted with somatic symptom mental disorders, notably somatic symptom disorder, conversion disorder (functional neurological symptom disorder), and factitious disorder. DSM-5 has a new category, "somatic symptom and related disorders," that includes these and other related disorders, and states that their initial presentation occurs mainly in medical rather than psychiatric settings. Table 28.1–5 summarizes some aspects of the differentiation of malingering from factitious and somatic symptom disorders.

Table 28.1–5.
Differentiating Malingering from Factitious and Somatic Symptom Disorders

	Malingering	Factitious	Somatic Symptom Disorders
Presentation	Psychological/physical symptoms	Psychological/physical signs and symptoms	Physical symptoms
Deliberate feigning	Yes	Yes	No
Inferred motivation	Conscious desire for tangible gain	Sick role; to receive medical care	Multiple determinations

Factitious disorders imposed on self are characterized by physical, psychological, or mixed symptoms plus signs of illness that are intentionally produced, feigned, or aggravated. This may or may not be associated with an unconscious psychological need to assume the sick role; yet, it is an apparent wish to receive medical care. Emphasis on diagnosis is on the objective identification of falsification, rather than on an inference about intent or possible underlying motivation. This diagnosis always implies psychopathology. Its prevalence is uncertain but said to be about 1 percent. Individuals with factitious disorder present their history in a dramatic fashion but tend to be vague regarding details. Pathological lying (i.e., pseudologia fantastica) may be present. If frustrated by caregivers, these individuals may become uncooperative or abusive. Intentionally feigned symptoms may include depression, amnesia, PTSD caricatures, infections, pain, fevers, or neurological symptoms. Pain and requests for analgesic medications are very common. The most severe and chronic form of factitious disorder primarily with physical signs and symptoms has been referred to as *Munchausen syndrome*, which includes recurrent hospitalizations, peregrination (traveling), and pseudologia fantastica. "Factitious disorder imposed on another" (*Munchausen's by proxy*) refers to a form of maltreatment in which an individual in the parental or caregiving role fabricates or produces physical, psychological, or combined illness in another to assume the vicarious sick role. Other actual mental disorders and borderline personality disorder may also accompany this.

Malingering differs from factitious disorder in that the malingerer's intentionally feigned symptom presentation is secondary to a clear-cut conscious motivation regarding a recognizable external reward, payment, or prize. The malingerer seeks a tangible external incentive, which is usually material, not an abstract satisfaction. The patient with factitious disorder, unlike the malingerer, can be shown to create an identifiable falsification to being ill, impaired, or injured in the absence of obvious external rewards. Dangerous, self-injurious, and self-handicapping signs often accompany feigned symptoms. Malingering and factitious disorder

may not be diagnosed together. In psychodynamic terms, the intentionally feigned signs and symptoms of factitious disorder satisfy the unconscious need that the *primary gain* from an illness serves, that is, the unconsciously determined transformation of conflict into symptomatology. By contrast, the symptoms feigned in malingering are consciously motivated by the *secondary gain* or environmentally recognizable advantages that they are intended to procure.

Somatic symptom disorders (somatic symptom disorder and conversion disorder) are characterized by the presence of physical symptoms that are not intentionally produced and are not under voluntary control. In *conversion disorder*, also termed *functional neurological symptom disorder*, somatic symptoms are clearly incompatible with neurological pathophysiology; and in contrast to somatic symptom disorder, the patient does not express excessive thoughts, feelings, and behaviors about the disorder; the focus is typically on the presenting symptom of loss of function. In conversion disorder (e.g., blindness and paralysis), patients are cooperative, welcome evaluations, search for answers to their symptoms, yet present with historical gaps and vagaries. In *somatic symptom disorders*, complaints suggest a medical condition but cannot be fully explained by a general medical condition, by the direct effects of a drug substance, or by another mental disorder. Most of the somatic symptoms in somatic symptom disorders cannot be demonstrated to be clearly incompatible with pathophysiology (e.g., pain and fatigue). The patient focuses on the distress caused by symptoms. In contrast to the relatively more euthymic patients with conversion disorder, patients with somatic symptom disorder present as chronic worriers. Some have speculated that psychological factors, inferred to be dynamically unconscious, have an important role in regard to the physical symptoms. In this mental disorder, the patient appears to believe that his or her symptoms are real, and distress is based on the belief of having a serious disease. Hypochondriasis may be seen in approximately 75 percent of these individuals. Complaints are in excess of what would be expected from history, physical examination, or laboratory findings, or there may be an absence of objective findings fully to explain the clinical presentation. Other mental disorders and genuine medical disorders such as depressive and anxiety disorders may coexist with somatic symptom disorders.

Malingering differs from somatic symptom disorders in that the malingerer's intentionally feigned presentation of physical and psychological symptoms is under voluntary control. In somatic symptom disorders, only physical symptoms are present, and they are not intentionally feigned and are not under the patient's voluntary control. Conversion disorder within the somatic symptom disorder category is susceptible to symptom relief by suggestion and hypnosis, unlike the situation in malingering. Somatic symptom disorder may also exist with

conversion disorder. If factitious disorder (does not seek obvious external reward) is diagnosed, malingering (seeks obvious external reward) ordinarily is not added as an additional condition. However, the presence of some factitious or malingered symptoms, mixed with other symptoms that are clearly nonintentional, may be found. When this mixed situation appears present, both somatic symptom disorder (referring to the nonintentional symptoms) and factitious disorder or malingering may be acknowledged.

COURSE AND PROGNOSIS

Malingering when deemed present needs to be assessed in the overall biopsychosocial context of the person's life. The presence of diagnosable mental disorders, their histories, and their response to psychotherapy and medications must be addressed. Any acute or chronic medical or surgical problems and their effect on the patient's overall functioning need to be considered. In forensic, hospital-related settings or in military cases, malingering may be context specific and change its presentation or even disappear if circumstances change. Because malingerers do not usually comply with treatment recommendations, the malingering status often remains unaffected. Research in this area is limited but shows that if reinforced, malingering may continue; if malingered goals are met, symptoms sometimes abate; if the malingerer becomes frustrated with current care, other providers are sought.

TREATMENT AND MANAGEMENT

Since malingering is regarded as a condition that may be a focus of clinical attention, no routine or standardized interventions have been devised and recommended. A range of suitable management approaches having clarification as a primary goal includes the following:

1. A clear diagnostic delineation of any medical or psychiatric symptoms and disorders, if present, in addition to malingering.
2. A mindful consideration of legal and ethical issues, especially those reflected in the medical record, because malingering is unlikely to be proved conclusively.
3. If the psychiatrist is the treater, a nonthreatening approach should include concerned neutrality and the avoidance of any abrupt confrontation or accusations of lying. Clear recognition of issues of implicit countertransference reactions (e.g., an unreasonable or exaggerated sense of annoyance at the patient) and awareness of the possible subtle effects of *projective identification* on all parties can minimize inordinate distortions. The psychological boundary violations that result when inadequate self-reflection is given to this subtle projectively induced resonance typically stir up adverse and negative

emotional responses in the examiner that are counterproductive. Such obscure information transfers have important communicational value. They often act as defensive maneuvers whose aim is aggressively to control and force the recipient to think and feel in ways that the sender wishes.

The clinician might discuss that all findings indicate that no medical basis for the patient's symptoms can be found, and hence no medical or psychiatric diagnoses can be established. Suggesting that the patient's history appears incomplete or that something very important has been left out may offer the patient a last opportunity to reveal the illness deception or devise an explanation to save face. With tactful clarification, alternative strategies that are healthier might emerge. Parsimonious, sensitive feedback might then encourage the patient to shed further light on his or her motivations. In some cases, the patient may become angry, respond defensively, and decide to go elsewhere. It is not often that a patient will openly admit malingering.

4. If the psychiatrist is a consultant, suggested recommendations and management strategies can be provided directly to the referring party for implementation.

FUTURE DIRECTIONS

The introduction of malingering into the DSM-III in 1980 caused a vociferous reaction by many nonmedical specialists in a vigorous effort to question and to examine the validity of this attempt to medicalize deceptive illness behavior. What had seemingly been an orphan condition was and is continuing to be the subject of intense debate. The medical model is in contrast to an alternate view, which emphasizes the person in terms of having agency, free will, and choice and thus being able to make consciously planned decisions. From this later perspective, malingering as illness deception is said to be better regarded within the context of individual responsibility, accountability, conscious choice, and social action. According to such a framework, the emphasis in malingering should eschew an uncritical adoption of the medical model and instead consider legal issues and questions having to do with problems of intentional social deviance. Those who advocate this position state that patterns of misconduct that are, in fact, "bad behavior" are incorrectly being redefined as bona fide illnesses.

Many have debated the extent to which malingering has clear-cut legal implications, such as by equating it with criminal fraud. It can be argued, however, that the function of a medical expert is not principally to make a judgment on an individual's honesty. The medical expert's role is to assess a clinical presentation, evaluate reported symptoms, render diagnoses based on the best evidence possible, and suggest a treatment plan. This approach to malingering clearly gives the physician a medical rather than a

legal or even a moral role.

The effort to introduce more objectivity into the field of detecting malingering has produced sophisticated test measures from the medical field. Detecting deception has shown some success by mapping brain correlates of lying. Functional magnetic resonance imaging (fMRI) has demonstrated specific neurocircuitry in brain regions under conditions of simulated, mock crime such as stealing. Truthful, as opposed to deceptive, responses, for example, were correctly identified in 90 percent of the model testing group. Similar brain imaging studies have shown that different cerebral activation processes are involved when a symptom is subjectively experienced as real compared to when it is feigned. Such objective data are intriguing but far from conclusive.

The entire area of malingering and assessing symptom validity with all that it implies is complex. Partial malingering may be common and add to indeterminate conclusions. Obvious psychological, psychiatric, monetary, social, legal, moral, and philosophical issues, to mention merely a few, are at play. Although it is laudatory to attempt to understand malingering in an objectively valid manner as many in specialized fields are attempting to do, the explosion of research, conjecture, simulation studies, and psychological testing has raised serious ethical questions. At this time, there is no generally agreed on consensus. The psychiatric approach, with its relatively straightforward and presumed nonjudgmental guidelines, values a dedicated capacity for examining multiple sources of information, often ambiguous and indeterminate. Psychiatry, striving for a nosological reserve, therefore participates as a rational, even essential, player in the field.

REFERENCES

Bass C, Halligan P. Factitious disorders and malingering: challenges for clinical assessment and management. *Lancet*. 2014;383(9926):1422–1432.

Bass C, Halligan P. Factitious disorders and malingering in relation to functional neurologic disorders. *Handbook of Clinical Neurology*. 2017;139:509–520.

Ben-Porath YS, Tellegen A. *MMPI-2-RF: Manual for Administration, Scoring, and Interpretation*. Minneapolis, MN: University of Minnesota Press; 2008/2011.

Carone DA, Bush SS. *Mild Traumatic Brain Injury: System Validity Assessment and Malingering*. New York: Springer; 2013.

Crossman AM, Lewis M. Adult's ability to detect children's lying. *Behav Sci Law*. 2006;24:703.

Daubert v. Merrell Dow Pharmaceuticals, Inc. 113 S.Ct. 2786 (1993).

*Halligan PW, Bass C, Oakley DA, eds. *Malingering and Illness Deception*. London: Oxford University Press; 2003.

Heilbronner RL. *Neuropsychology in the Courtroom: Expert Analysis of Reports and Testimony*. New York: Guilford Press; 2008.

*Iverson GL. Ethical issues associated with the assessment of exaggeration, poor effort, and

malingering. *Appl Neuropsychol*. 2006;13(2):77.

Klein M. Envy and gratitude [1957]. In: *Envy and Gratitude and Other Works*. London: Hogarth Press; 1975:176.

Kozel A, Johnson KA, Mu Q, Grenesko EL, Laken SJ. Detecting deception using functional magnetic resonance imaging. *Biol Psychiatry*. 2005;58(8):605.

Kucharski TL, Duncan S, Egan SS, Falkenbach DM. Psychopathy and malingering of psychiatric disorder in criminal defendants. *Behav Sci Law*. 2006;24:633.

Larrabee GJ. Aggregation across multiple indicators improves the detection of malingering: relationship to likelihood ratios. *Clin Neuropsychol*. 2008;22(4):666.

Libow JA. Child and adolescent illness falsification. *Pediatrics*. 2000;105(2):336.

McCann JT. *The Significance of Malingering and Deception in Adolescents: Assessing Credibility in Clinical and Forensic Settings*. Washington, DC: American Psychological Association; 1998.

*McDermott BE. Psychological testing and the assessment of malingering. *Psychiatr Clin North Am*. 2012;35(4):855.

Mittenberg W, Aguila-Puentes G, Patton C, Canyock EM, Heilbronner RL. Neuropsychological profiling of symptom exaggeration and malingering. *J Forens Neuropsychol*. 2002;3(1-2):227.

Mittenberg W, Patton C, Canyock EM, Condit DC. Base rates of malingering and symptom exaggeration. *J Clin Exp Neuropsychol*. 2002;24(8):1094.

Ninivaggi FJ. *Biomental Child Developmental: Perspectives on Psychology and Parenting*. Lanham, MD: Rowman & Littlefield Press; 2013.

Ninivaggi FJ. Emotion as a Second Language or Should They be our First?—Emotional Literacy: A Forgotten First Language to be Remembered. *Psychologytoday.com*. February 20, 2015.

*Ninivaggi FJ. *Making Sense of Emotion: Innovating Emotional Intelligence*. Lanham, MD: Rowman & Littlefield Press; 2017.

Nock MK, Stein MB, Heeringa SG, et al. Prevalence and correlates of suicidal behavior among soldiers: results from the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *JAMA Psychiatry*. 2014;71(5):514–522.

Norfolk PA, Floyd RG. Detecting parental deception using a behavior using scale during assessment of attention-deficit/hyperactivity disorder: An experimental study. *Psychology in the Schools*. 2016;53(2):156–172.

Ord JS, Greve KW, Bianchini KJ. Using the Wechsier memory Scale-III to detect malingering in mild traumatic brain injury. *Clin Neuropsychol*. 2008;22(4):689.

Resnick PJ, Knoll J. Faking it: how to detect malingered psychosis. *Curr Psychiatry*. 2005;4(11):13.

Rios J, Morey LC. Detecting feigned ADHD in later adolescence: an examination of three PAI-A negative distortion indicators. *J Pers Assess*. 2013;95(6):594–599.

*Rogers R. *Clinical Assessment of Malingering and Deception*, 3rd ed. New York: Guilford Press; 2012.

Rogers R, Granacher RP. Conceptualization and assessment of malingering. In: Drogin EY, Dattilio FM, Sadoff RL, Gutheil TG, eds. *Handbook of Forensic Assessment: Psychological and Psychiatric Perspectives*. Hoboken: Wiley; 2011:659–678.

Sanders MJ. Symptom coaching: factitious disorder by proxy with older children. *Clin Psychol*

Rev. 1995;15(5):423.

Slick DJ, Sherman EMS. Differential diagnosis of malingering. In: *Mild Traumatic Brain Injury: System Validity Assessment and Malingering*. Carone DA, Bush SS, eds. New York: Springer; 2013:57–72.

Slick DJ, Sherman EMS, Iverson GL. Diagnostic criteria for malingered neurocognitive dysfunction: proposed standards for clinical practice and research. *Clin Neuropsychol*. 1999;13:545.

Sollman MJ, Ranseen JD, Berr DT. Detection of feigned ADHD in college students. *Psychol Ass*. 2010;22(2):325.

Stout RG, Farooque RS. Claims of amnesia for criminal offenses: psychopathology, substance abuse, and malingering. *J Forensic Sci*. 2008;53(5):1218.

Suhr J, Hammers D, Dobbins-Buckland K, Zimak E, Hughes C. The relationship of malingering test failure to self-reported symptoms and neuropsychological findings in adults referred for ADHD evaluation. *Arch Clin Neuropsychol*. 2008;23(5):521.

*Young G. *Malingering, Feigning, and Response Bias in Psychiatric/Psychological Injury: Implications for Practice and Court*. New York: Springer; 2014.

*Younga S, Jacobsonc R, Einzigd S, Grayd K, Gudjonssond GH. Can we recognize malingerers? The association between malingering, personality traits and clinical impression among complainants in civil compensation cases. *Personality and Individual Differences*. 2016;98:235–238.

▲ 28.2 Adult Antisocial Behavior, Criminality, and Violence

DOROTHY OTNOW LEWIS, M.D.

The acts subsumed under the terms "adult antisocial behavior, criminality and violence" range from stealing candy from a baby to abducting, raping, and murdering the child. They cover a multitude of sins as well as a multitude of neuropsychiatric signs, symptoms, and behaviors. In spite of the diverse aberrant behaviors included within these categories, official diagnostic manuals consider "adult antisocial behavior" to be a condition "not due to a mental disorder." For decades, criminals, violent and nonviolent, were deemed simply "bad." Those committing particularly heinous crimes were "evil." They posed a religious, not a medical question. There was nothing wrong with their brains. Professional thieves, racketeers, dealers in illegal substances, thugs, even serial killers, were regarded simply as mentally healthy malefactors. They were immoral and their acts were thought to represent their conscious choice of an alternative lifestyle. Such assumptions presupposed the existence of supportive scientific evidence. There was not any, Fagan, the Godfather, and Tony Soprano to the contrary not withstanding.

People tend to divide crime or antisocial behavior into two general categories: street crime and white-collar crime. But what about organized

crime?

ORGANIZED CRIME

In the United States, organized crime, as defined by the Organized Crime Control Act of 1970, is "the unlawful activities of . . . a highly organized disciplined association. This structured criminal activity is referred to as 'racketeering.'"

Until 1957, J. Edgar Hoover denied or ignored this violent if lucrative entity. Hoover focused his attention and manpower on uncovering alleged Communists in all levels of government, while the Mob (otherwise known as the Mafia or Cosa Nostra) flourished right under his nose. Prior to 1957, Hoover dismissed The Mob as hoodlums to be dealt with by local police. However, in the fall of 1957, as a result of a series of mobster assassinations and turf wars, a meeting of mob bosses from all over the country took place at the palatial 58 acre Apalachin estate of Joe Barbara. A group of snazzy cars parked around the Barbara mansion caught the eye of the local police who gave chase. Some mobsters were caught running away through cornfields. Headlines the next day shouted, "Police Nab 67 Mafia Chiefs in Backwoods." Of them only 20 were convicted on minor charges and the court of appeals reversed all convictions. Such was the power of this theoretically nonexistent entity, The Mafia.

Although much is written about "The Mob," its structure, its rules, its alleged rituals, little, if any research on the psychology of mobsters, exists. In 2004, two Italian psychologists attempted a clinical description of the workings of the Mafia mind. Unfortunately, the vague, undocumented nature of their database made it impossible to generalize from their publication. Nonetheless, their description of Mafia family structure was consistent with other Mafia observers. There is a consensus that at its height, many young men who joined The Mafia (or Cosa Nostra), unlike their impoverished immigrant peers, did so not out of desperation but rather out of admiration. Kingpins commanded respect. The toughest, strongest, most self-assured male children were, reportedly, groomed to take over positions of authority previously held by their fathers, uncles, or other powerful male relatives.

Absent vigorous data, Mafia psychology is open to speculation. Thus, some criminologists, years ago, asserted that criminals, in general, and mobsters in particular, suffered from a lack of self-control and sought only to satisfy their immediate needs or wishes. Others hypothesized the existence of a criminal personality characterized by risk-taking, unreliability, irresponsibility, and inability to tell the truth.

The author of this chapter could find but one systematic study comparing Mafia-linked incarcerated criminals with incarcerated non-Mafiosi. It used the Hare Psychopathy Check List-Revised (PCL-R) (allegedly a measure of psychopathy). It should be noted that the diagnostic

manual describes antisocial personality disorder as a condition that was once called psychopathy. The researchers found, to their surprise, that none of the Mafia-linked criminals scored 30 or above on the Hare Scale (a generally accepted threshold for diagnosing psychopathy), and only 3 (10 percent) scored between 20 and 29. In contrast 4 (10 percent) of the comparison inmates scored above 30, and an additional 18 (46 percent) scored between 20 and 29. These results would suggest that being a run-of-the-mill criminal is quite different from being a "made man." The most fascinating finding was that the Mafiosi had significantly lower scores on the factor in the PCL-R allegedly measuring difficulties in interpersonal relatedness and affective capacity. In other words, the Mafiosi were more caring and empathic than their common criminal jailmates, at least to the extent that it could be measured by the PCL-R. Many had stable marriages as well as mistresses.

One student of the Mafia concluded that the personality traits that predict leadership success (i.e., extroversion, controlled impulsiveness, a sense of adventure, megalomania, and Narcissistic Personality Disorder) are equally suitable traits for a successful career in either business or organized crime. To the author of this chapter, many of the behaviors described above smack of hypomania. This will be discussed when the origins of the Hare Scale are looked at further.

WHITE-COLLAR CRIME

The line between legitimate business practices and racketeering often blurs or disappears entirely. One need simply scan the front page of the daily newspaper to marvel at the nature and scope of skullduggery perpetrated by heretofore upstanding businessmen who seem to be in their right minds. Clearly there are hidden cracks in many pillars of society. It cannot be said with conviction that, as a group, racketeers, embezzlers, and money launderers are free of psychopathology any more than it can be said of doctors, lawyers, stockbrokers, or accountants. Less is known about the neuropsychiatric characteristics and brain function of so-called white-collar criminals than is known about violent criminals.

The term white-collar crime was coined by Edwin Sutherland at a 1939 meeting of the American Sociological Society. Sutherland took issue with his colleagues who theorized that crime was primarily an outcome of poverty and socioeconomic inequality. Sutherland called attention to the improper business practices of individuals in positions of power, practices he felt should be regarded as criminal offenses rather than simply as civil wrongs.

The Federal Bureau of Investigation (FBI) has defined *white-collar crime* as illegal acts characterized by deceit, concealment, or violation of trust that are not dependent on the application or threat of physical force or violence. These crimes thus are treated as conscious, rational, economic

decisions.

Over the past 4 years estimates of the cost of white-collar crime have ranged from 200 to 750 billion dollars per year, whereas the cost of street crime has been estimated at somewhere between 14 and 16 billion dollars. The human consequences of embezzlement and fraud (e.g., lost homes, lost tuition, lost pensions, and mental and physical illness) is immeasurable.

To understand the apparently increasing incidence and prevalence of financial crime, one must study the vast changes in economic structure since the 1990s. For example, with very little oversight, corporate officers in publicly held companies with the stroke of a key can artificially inflate the price of a stock. Such practices have been implemented by new technologies at the expense of customers and employees. The more automated the victim (such as the digits of a person's retirement fund), the easier it is to defraud. The farther away the target, the easier it is to destroy it. It is easier to kill with drones than with knives. The less human the target, the less troubled the conscience. Wall Street moguls, who would never consider holding up their clients at gunpoint, think nothing of tapping a key and transferring their clients' life savings into questionable, if not downright illegal, investment schemes.

Such moral compromises are not limited to the financial world. Most troubling is the degree to which bottom-line economy has corrupted the medical industry.

An especially egregious example of corporate and medical malpractice was the marketing and dissemination of the Dalkon Shield, an intrauterine contraceptive device. In spite of early warnings regarding the risky design of the product (it had a multifilament string with a tendency to wick and carry bacteria-laden fluid from the vagina into the otherwise sterile uterus), corporate officials went forward with promotion, production, and sales.

Shortly thereafter, the company began to receive complaints regarding cases of pelvic inflammatory disease and of septic abortions. Nevertheless, it continued to market the device for over 3 years. By the time the corporation filed for bankruptcy, over 300,000 lawsuits had been filed.

According to the National Health Care Anti-Fraud Association (NHCAA), most health care fraud is committed by a small percentage of health care providers. However, the NHCAA estimates that the financial losses from health care fraud are in the tens of billions of dollars.

EPIDEMIOLOGY

During the first decade of the 21st century, as violent crime rates diminished, government and private lawsuits for fraud and other financial violations soared. Scholars and law enforcement personnel have speculated that the surge in high-stake white-collar crime is a result, at least in part, of technological advances such as the Internet, automated tellers, and plastic that facilitate certain kinds of offenses (e.g., hacking, identity theft, and embezzlement) and provide opportunities for new creative kinds of white-

collar crime. The scope of white-collar crime is ever widening. It has come to include criminal acts in violation of environmental and health laws. The Department of Justice has finally begun to focus on financial institution fraud, defense procurement fraud, health care fraud, computer crimes, telemarketing fraud, and international fraud.

Health care fraud is soaring as large health care programs, public and private, steal literally trillions of dollars. According to the Government Accounting Office (GAO), in 2011 Medicare and Medicaid paid a whopping $65 billion in improper payments. According to the FBI, one of the most significant trends in health care fraud is the participation of health care professionals—in other word, doctors—who in so doing risk their patients' lives, not to mention their patients' pocketbooks. The FBI has reported upsurge in unnecessary treatments, including surgeries, the prescribing of unnecessary, potentially dangerous drugs, and overall substandard care.

Of special concern to the field of psychiatry are cases in which psychiatric/psychological professionals have engaged in fraudulent practices. Individual practitioners, as well as hospital chains, have pled guilty to providing unnecessary treatment to thousands of patients to collect insurance money.

When insurance reimbursements depend in large part on procedures performed, psychiatrists are at a disadvantage. Talk is cheap. There are few "procedures" for which to charge. At the same time, thanks to technological advances, certain new treatment modalities, "procedures," are now available (e.g., transcranial magnetic stimulation and deep brain stimulation). Electroconvulsive therapy, once rare, has experienced a rebirth.

One of the most devastating instances of white-collar crime in the United States involving the medical profession is the ongoing case involving Purdue Pharma, producer of OxyContin, an opioid touted by Purdue as being less addictive than other pain medications because of its time-release formula. Prior to marketing, internal documents had been circulated within the company indicating the product's potential for abuse. Nonetheless, sales representatives, presumably under management's instructions, falsely assured doctors of the safety of their product. The U.S. Food and Drug Administration (FDA) approved the drug in 1995 and by the year 2000 parts of the United States, particularly rural areas, experienced soaring rates of addiction and of crimes related to addiction. Although the pill as marketed was long acting, the addict, by simply crushing the pill and swallowing the powder, or diluting and injecting it, could experience an immediate, intense "high." At a news conference in 2007, a United States attorney stated that the aggressive, untruthful marketing of OxyContin had resulted in rising crime rates, teenage drug addiction, and even deaths. This particular fraud was not limited to the pharmaceutical industry. Doctors reportedly were paid handsomely to promote the drug to colleagues.

WHITE-COLLAR VIOLENCE

The more closely one examines corporate crime and criminals the clearer it becomes that white-collar criminals cause extraordinary physical damage. A graphic example is the production and sale of the Ford Pinto long after

the corporation knew of the flaw in its design that caused it to explode on collision. The "bean-counters" at Ford concluded that the cost of lawsuits from injuries and deaths would be less than the cost of a large-scale recall; so Ford stalled the recall.

Another clear example of corporate violence, of obliviousness to human life, is the coal mining industry. In April of 2010, an explosion occurred at the Upper Big Branch Mine in West Virginia, killing 29 of the 31 miners working there that day. The owners of the mine had, previously, received dozens of notifications of safety violations and had not rectified them. In the course of its investigation, the West Virginia Office of Miners Health, Safety and Training (WVOMHS and T) issued 253 violations of safety standards.

According to the US Bureau of Labor Statistics (BLS), there were over 4,500 fatal work injuries in the year 2012. That amounted to 3.4 fatal injuries per 100,000 full time workers. This compares to 4.7 deliberate homicides per 100,000 population in the United States in 2011. Thus, corporate malfeasance has a mortality rate only slightly lower than the mortality rate for deliberate homicide.

ETIOLOGY

What causes white-collar crime? In short, the answer is unknown. The field is wide open to speculation. Sutherland believed that such crimes were learned behaviors resulting from association with those who approve of them and isolation from those who disapprove. Others have argued that white-collar crime, like street crime, results from a tendency of individuals to seek immediate, short-term gratification without considering the possible long-term consequences of their acts. Such individuals, they believe, are impulsive, aggressive, and motivated by the pursuit of pleasure and avoidance of pain. Comparisons of street criminals with other kinds of white-collar offenders reveal marked demographic differences (poor vs. rich). One MRI study reported hypertrophy of parts of the frontal cortex in white-collar criminals and suggested a relationship to "abstract monetary rewards." Although some have asserted that individual tendencies lead to all crime, others have argued that perceived incentives and disincentives like those described earlier are more powerful influences. Others have pointed to ethical deficits. Experiments on college students suggest that just about anyone will cheat when given the chance. Subjects from Harvard, MIT, Princeton, UCLA, and Yale all cheated on a test when given the opportunity.

Why is white-collar crime so rampant? Callahan, in his examination of this phenomenon, called attention to the trend in the United States toward an increasingly free market, laissez-faire society with a focus on money. He posited that the bottom-line ethic (if it can be called that) has infiltrated areas of culture that previously were insulated from market pressures (e.g.,

medicine, education, journalism, and sports). He hypothesized that the greater inequality in America between the haves and the have-nots has weakened the social fabric—undermining the notion that everyone is in it together and bound by the same rules. Callahan suggests four key reasons for the increased levels of cheating:

(1) Increasing pressures to succeed in a society in which job security cannot be counted on. (2) Weakened safeguards against cheating. The FBI, Securities and Exchange Commision (SEC), and Justice Department currently have insufficient resources to stop white-collar crime. (3) Bigger rewards for those who cheat, from the CEO who inflates company earnings to the baseball player who bulks up on steroids to maintain his million dollar-plus annual income. (4) Finally, what Callahan terms "trickle-down corruption" in a society in which the ordinary middle-class person perceives that the cards are stacked and the rules are unfair. At this time, this author suggests that a fifth reason should be added: (5) The ever-increasing distance between perpetrators and victims. It is easier to harm another human being at a distance than "up close and personal."

MORALITY AND THE BRAIN

For millennia morality has been assumed to be an acquired capacity, developed with maturation and upbringing. Most religions include specific moral concepts, beliefs, and standards inculcated during childhood and encouraged throughout adolescence and adulthood. Pillars of the church and pillars of the community are considered fabricated of the same ethical stuff. Morality is, in reality, among the most sophisticated features of human judgment, behavior, social cognition, decision making, and conflict resolution. Morality includes theory of mind (i.e., the ability to understand the complex emotions and mental states of others). Scholars debate the relative biological inheritance of this capacity versus environmental influences on it. It has been suggested that the infant is born with the innate capacity for implicit morality, whereas explicit morality requires ongoing instruction. Infants will cry when they hear other babies crying but they cannot tell you why.

Damasio in the 1990s and early 21st century had established an association between frontal lobe damage and immoral dyscontrolled behavior. The past decade has witnessed a plethora of brain studies, attempting to localize the areas in the brain and the pathways among them that facilitate moral thought and behavior. Greene and colleagues (2001) were also among the first to use functional magnetic resonance imaging (fMRI) to study moral judgments. Several different scenarios involving moral choices and neutral choices were presented to subjects undergoing fMRIs. One quandary placed the subject on a footbridge over railroad tracks. If the train were allowed to run onward, five workers would be

killed. If, however, the subject pushed a fat man off the bridge onto the tracks, he would be killed but the other five would live.

Investigators found that certain parts of the brain were especially active when contemplating personal moral questions (i.e., the medial frontal gyrus, the posterior cingulate gyrus, and the angular gyrus). Some of these areas have been found to be active in emotional processing compared with cognitive reasoning and they were more active when subjects contemplated the "up close and personal" choices than when considering impersonal moral choices and nonmoral choices. Decety and Porges modified Greene's study by showing subjects scenarios in which a person was harming or helping another person. Subjects were instructed to identify themselves with the active (harming or helping) protagonist. Helping activity was associated with increased activity in the ventral striatum, while harmful activity resulted in a "strong signal decrease" in the ventromedial prefrontal cortex (VPFC) and activation in the amygdala. Thus, Decety and Porges demonstrated that the perspective of the subject has a profound effect on the recruitment of different areas of the brain during moral considerations. Investigators have found that damage to the PFC interferes with the development of aversive response to harmful actions. They have argued that what some would call psychopathic behaviors result from damage to the PFC and a failure to develop aversion to immoral behaviors.

Does it matter when in life damage to the prefrontal cortex occurs? It does. Recently Taber-Thomas and colleagues compared the moral functioning of individuals who experienced VPFO damage in adulthood with those who had similar injuries in childhood. They found that when adults who early in childhood experienced damage to the VPFO were confronted with decisions of self-interest (e.g., whether or not to kill your boss for money); they were less likely to condemn the act than were adults who had received similar damage to their brains during adulthood. The researchers suggest that the subjects with late onset injury had, over time, developed affective responses that prohibited self-interested behaviors.

Morality involves numerous other areas of the brain connected with each other as well as with the prefrontal cortex. These areas include the cingulate gyrus, the amygdala, the dorsolateral prefrontal cortex, the temporal lobes, and even the cerebellum. Studies of the timing of reactivity from one area to the next exist. Morality itself has been divided into different aspects. This brief overview should tantalize the reader into exploring the findings, agreements, and disagreements, which swirl around the fascinating exploration of the neurologic foundations of morality.

VIOLENCE

What kind of animal are humans? Was Hobbes right? Is the natural state one of "continuale feare, and danger of violent death?" Without a strong government or set of rules "is the life of man, solitary, poore, nasty, brutish,

and short?" Was Samuel Butler wise to caution "Then spare the Rod and spill (sic) the Child?" Bad advice, at least according to current scientific data. Corporal punishment has been related to subsequent violence. No matter. Butler based his belief on the Biblical "He that spareth his rod hateth his son: but he that loveth him chasteneth him betimes." This is not the only time in this 21st century that science and God have found themselves at odds.

What about Wordsworth? Was he on to something when in *Intimations of Immortality* he proclaimed that the human entrance into this world is one of purity and innocence? Humans come "trailing clouds of glory" only to be corrupted by the pressures of society, of living along side other people. Rousseau agreed. Locke disagreed. According to Locke, each person enters this world a blank slate to be scrawled upon by his or her environments and experiences.

Are all men (and women) created equal, or are some people innately more aggressive than others? For millennia these kinds of questions were the provenance of religion, philosophy, literature, and politics. Now these eternal questions have fallen into the lap of science, into offices, laboratories, and positron emission tomography (PET) scanners. Think of current technologies, statistically validated neuropsychological test materials, PET scans, fMRIs, the newly devised ability to chop up yards and yards of DNA, and discover what humans are really made of—will these developing technologies enable scientific answers based more on fact and less on faith to be found? A start has been made. It is already known that each person enters the world with certain innate but not inevitable potentials for behavior. What happens to each person makes the difference. No single factor or known constellation of factors accounts for violence. Rather, as will become clear, a variety of different intrinsic and environmental influences come together, interact, and create matrices for violence.

EPIDEMIOLOGY

Just over a decade ago, Nelson Mandela stated, "The Twentieth Century will be remembered as a century marked by violence. It burdens us with its legacy of mass destruction, of violence inflicted on a scale never seen and never possible before in human history." He blamed this extraordinary violence on "new technologies in the service of ideologies of hate." Had he lived longer he would have marveled at the start the 21st century has made.

At this time, there are four ongoing major conflicts (two in Asia, two in Africa) each of which has a death toll of over 10,000 human beings each year. Between 2011 and the start of 2015, cumulatively there have been 200,000 deaths in the "Syrian Civil War." The "Iraqi insurgency" has had 35,829 fatalities since its onset in 2011; the "South Sudanese civil war" has killed over 50,000 since its onset in 2013; and the "Islamist insurgency" in

Nigeria has resulted in over 22,000 deaths. Currently, there are nine additional conflicts worldwide causing between 1,000 and 10,000 deaths per year.

PEACE AND PEOPLE

In contrast to the easy documentation of worldwide conflict and homicidal violence, the documentation of peaceful societies is more difficult. One must almost literally beat the bushes to find data on peaceful societies. Those who have studied these kinds of societies described extremely small communities that avoided contact with other neighboring communities. They reported that adult members of such groups raised their children to be hesitant and fearful about the intentions of others. All physical violence was eschewed, including competitive games. One of the so-called "peaceful societies" is that of the Piaroa, a jungle people who live along the tributaries to the Orinoco River in Venezuela. Piaroa children are taught the evils of vanity, jealousy, arrogance, dishonesty, and ferocity. They are taught to protect themselves from unseen cannibalistic beings. Another reportedly peaceful tribe, the Semai of Malaysia, "see themselves as helpless in a hostile world that is beyond their control." In truth, there is lacking evidence that idyllic, nonaggressive societies ever existed. Margaret Mead's account of the peace-loving Samoan people was based on approximately 9 months of research on the island, most of which was devoted to studying adolescent girls. Numerous accounts of Samoan life, written before and after Mead's book, indicate that Mead's generalizations about the peacefulness of Samoan life were simply wrong. It is surprising to learn that even now, the proportion of men killed by warfare in other primitive societies reportedly dwarfs the proportion of men in the United States and Europe who die in wars.

What about Buddhist societies? Tibet is a country long regarded as a model of nonviolence. The Dalai Lama is a symbol of peace. In 1912, Tibet, formerly a Chinese principality, declared its independence from China and ruled autonomously from 1912 to 1950. Then Mao's army took over China. In 1958, following a series of uprisings in opposition to Chinese domination, a Tibetan army, made up predominantly of lay people, was formed to fight the Chinese. Over the next few years, resentment of the Chinese grew among Tibetans, and in 1959 a full-blown revolt broke out causing the Dalai Lama to flee his country. In spite of their religious vows of nonviolence, Tibetan monks were active in protest movements. The extent to which even Buddhist monks actually participated in warfare remains unclear.

For years the degree to which Tibetans, lay people and monks alike, took part in training for battle and fighting remained secret. It is easy to understand why. Both monks and lay people found themselves in an ethical dilemma. War violated the central principle of Buddhism, that is, to create

no suffering for any sentient being. Even the Dalai Lama confessed his ambivalence, his personal and political quandaries. As he explained in his autobiography, "Nonviolence is not an intrinsic Tibetan trait. It is instead a philosophical component of Buddhism and a political component of state struggle." Thus, one sees in his internal struggle mankind's difficulty in adhering strictly to a course of nonviolence. The source of this all too human dilemma will become clearer when consider the complexities of brain structure and function are considered.

Based on numerous studies of primitive and advanced societies, the consensus seems to be that violent conflicts motivated by jealousy, revenge, dominance, and sexuality (i.e., rape) are universal human behaviors and are especially characteristic of men. These findings do not mean that violence alone characterizes the human species. Counterbalancing these aggressive drives are gentler instincts and a brain that, under optimal conditions, is often capable of moral reasoning and peaceful behavior. Then again, as will become clear, the relationship of moral reasoning to actual behavior is not simple.

HOMICIDE

Epidemiology

According to the United Nations' Global Study on Homicide (2013), the major forces contributing to homicide include economic crises, food insecurity, low levels of human development, and high levels of income inequality.

A study by the Geneva Declaration on Armed Violence and Development (GDAVD) estimated that in 2004 the global homicide rate was 7.6 intentional homicides per 100,000 inhabitants. For the year 2012, the United Nations Office on Drugs and Crime (UNODC) did a similar study and found a global average homicide rate of 6.2 per 100,000. So much for progress. UNODC murder rates for the most recent years available are reported as 16.3 per 100,000 in all of the Americas, 12.5 per 100,000 in Africa, 3.0 per 100,000 in Europe, and 2.9 per 100,000 in Asia. The high rate in the Americas (16.3 per 100,000), seems to reflect the extraordinary murder rates in Mexico (e.g., in Guerrero 61.6 per 100,000; Chihuahua 42.4 per 100,000; and Sindora 43.7 per 100,000).

In the United States, there were reportedly 4.7 homicides per 100,000 inhabitants in contrast to 3.0 per 100,000 in Europe and 2.9 per 100,000 in Asia. The murder rates within the United States vary from state to state. The highest rates are in parts of the South (Washington, DC, 13.9 per 100,000, Louisiana, 10.8 per 100,000, Mississippi, 7.4 per 100,000). New England seems pretty safe (e.g., Massachusetts 1.8 per 100,000; Maine 1.9 per 100,000). In most states, the rates have been declining over the past decade. According to the FBI supplemental data, 89.6 percent of offenders

were male, 53 percent were black, 44.3 percent white, and 2.3 percent "other races." Of the victims, 77.7 percent were male, 51.1 percent were black, 46.3 percent were white, and 2.6 percent "other races."

Of the homicides for which weapons data were known by the FBI, 69.4 percent involved firearms. In Britain, including Northern Ireland, only 9 percent of homicides were gun related. The United States, reportedly, has the highest gun ownership rate in the world (89 guns per every 100 citizens) compared with Britain (6 per 100) and Canada (31 per 100). In 2010, in the United States, there were 9,960 gun homicides compared with Britain (58 gun homicides that year). The British population is about one fifth of the US population.

Who gets murdered? In the United States, 30.2 percent of victims were killed by someone they knew, and 12.5 percent were killed by a family member. Of the female victims, 35.0 percent were murdered by husbands or boyfriends. Felony circumstances such as rape, robbery, and burglary accounted for only 22.5 percent of murders.

The FBI reported the following: "Law enforcement officers justifiably killed 410 felons" while "Private citizens justifiably killed 310 people during the commission of crimes." The authors cite these numbers and descriptions verbatim since the wording assumes that the citizens who were killed by law enforcement officers were felons (absent trials) and gives no hint of the severity or nature of the crimes.

The FBI opened its 2013 report on officers killed that year with the statement "According to statistics collected by the FBI, 76 law enforcement officers were killed in the line-of-duty incidents in 2013." A closer examination of that report reveals that only 27 officers were victims of "felonious deaths" that year. The other 49 "died in accidents." And the nature of these accidents? According to the National Law Enforcement Officers' Memorial Fund, the number of fatalities of US law enforcement officers in 2013 was the lowest in 54 years, and the number dying from gunshot wounds was the lowest since 1887. The primary cause of officer fatalities in 2013 was traffic-related incidents (speeding to crime scenes?) that cost the lives of 46 officers, almost twice as many deaths as those caused by firearms.

Ironically, murder rates in death penalty states were consistently higher over the years than in nondeath penalty states. In this author's experience, none of the murderers she has evaluated gave the death penalty a second thought when they shot their victims. Most murderers do not work that way. They barely think once much less twice about the consequences of their acts.

What about wars? After wars end do nations breathe a sigh of relief and turn away from killing? No way. In 1976, sociologists studied homicide rate changes after 50 "nation-wars." They found that after such wars most nations experienced substantial increases in homicide rates. These

increases occurred after large wars and smaller wars, in victorious nations as well as defeated nations, in nations with thriving economies, and those in poor financial shape. Homicide rates increased with greatest consistency in nations that had suffered large numbers of combat deaths. Most of the wars studied occurred prior to the development of the concept of posttraumatic stress disorder (PTSD), a mental illness familiar to psychiatrists practicing after the Persian Gulf, Iraq, and Afghanistan wars when mental health workers recognized numerous examples of homicides and suicides by traumatized soldiers. The extent to which these kinds of mental disorders contribute to post-war homicides in general is not well established.

MASS SHOOTINGS

In contrast with declining homicide rates by juveniles is the startling increase in mass or rampage shootings in the past several years. Between 2000 and 2006, there were an average of 6.4 shooting sprees annually in the United States. However, between 2007 and 2013, there were, on average 16.4 such shootings each year in the United States. Bondu and Beier attempted a worldwide study of school shootings between 1886 and 2012. They studied school attacks in 32 different countries. They found that the overwhelming majority of attacks occurred in the United States. Of 232 school attacks, 150 (64.7 percent) occurred in the United States. No other country came close. During that same period of time, only 11 (4.7 percent) shootings occurred in Germany, and 11 (4.7 percent) in Canada. Numbers of attacks in other countries were in the single digits. Some have suggested typologies of mass shooters (e.g., psychopathic, psychotic, and traumatized). Some have reported that adult shooters tend to have experienced setbacks in social, family, or vocational domains as well as "autogenic components" (e.g., dysthymic reactions). Others have described shooters as having been bullied, as being obsessional, as being grandiose, and as harboring persecutory beliefs. Several have minimized the role of mental illness in mass shootings.

This author has evaluated but two teenaged mass shooters, clearly too small a sample from which to generalize. Each of these young shooters (age 16 years) had long-standing histories of maladaptive school behaviors. One of these youngsters had a history of auditory hallucinations dating back to age 12 years. He also, though highly intelligent, had specific learning disabilities for which he received special educational assistance. Of note, for years he had been obsessed with war and guns. Periodically he came to school dressed in camouflage or all in black. During the year prior to the shooting, he had been seeing a psychiatrist (a gun collector himself) who diagnosed a depressive disorder and treated him with an antidepressant medication. While in treatment and while taking his medication, he did well socially and academically. However, several weeks prior to the shooting his parents decided he was doing well enough and should stop therapy and stop taking his medication. Following cessation of treatment his functioning deteriorated. His essays and presentations in class centered on war and shooting. At one point, he jumped up in class, clutched his head, and shouted, "Goddamn these voices in my head." His teacher thereupon punished him, making him write out 100 times, "I will not swear in

class," completely overlooking his psychotic behavior. During the days prior to his shooting spree, he told his friends of his fantasies/plans of shooting up the school. No one did or said anything. One afternoon, following a bitter argument with his father (a high achiever who could not understand why his son had such trouble at school), the boy returned home and, using a firearm his parents had bought him, shot his father through the back of the head. When his mother came home, before she came in the door, he shot her in the back ("so she wouldn't see what I had done"). The following day he came to school with guns and ammunition, killed 2 classmates, and wounded several others. In addition to the author, a second psychiatrist evaluated him for the defense and diagnosed schizophrenia. However, as is often the case, the psychiatrist and psychologist who evaluated him for the state found no psychosis. On the advice of his attorney, the boy pled guilty to two counts of first-degree murder and was sentenced to life in prison.

The second shooter, a 16-year-old girl, also shot up a schoolyard of children, using a gun and ammunition gifted her by the male relative with whom she lived. Shortly before the day of the shooting, at the request of her school counselor, she was evaluated by a psychiatrist who recommended that she be psychiatrically hospitalized for severe depression. The male relative, with whom she had been living refused to allow this. At that time it was suspected that she was the victim of incest (she kept knives in plain sight around her room to prevent anyone from entering). When asked about being bothered sexually, she was evasive. She did not openly reveal the fact that she had been sexually abused by the male relative until a parole hearing 10 years later. The parole board ignored the relevance of this disclosure to mitigation.

During the psychiatric evaluation, the author learned that she had suffered a severe bicycle accident and had injured the left temporoparietal area of her skull. She also had a history of episodes of impaired memory for her actions as well as a number of symptoms consistent with possible complex partial seizures. Her memory of the shooting itself was vague and distorted and she lacked all memory for things she had said during an interview by police right after the shooting. Unfortunately prior to her trial, she never had a neurological or neuropsychological evaluation, although they had been recommended.

The psychiatrist for the prosecution found no evidence of mental illness. On the advice of her lawyer, she pled guilty to two counts of first-degree murder and, to this day, remains in prison. Several months after the shooting, the male relative with whom she had been living, married one of her girlfriends who was the same age as she. The author never saw this youngster after the initial interviews. However, years later according to press releases she was being treated in prison for what the press called a rare form of epilepsy.

Prior to the school shooting, she often talked about shooting and killing. Furthermore, approximately 3 days prior to her rampage, she warned her schoolmates that she would "do something big" the following week. People who do these kinds of things often make similar declarations that unfortunately are dismissed simply as ploys for attention. In both of the cases evaluated by the author, sharing their thoughts of murder and mayhem appeared to be efforts to be stopped. In this author's clinical experience, many murderers, young and old, indicate to others their homicidal thoughts days to weeks before they actually commit murder. Nobody gets the message.

EVOLUTION AND VIOLENCE

Does it matter that *Homo sapiens* are apes? Of course it does. It makes humans look at themselves not as special beings endowed with immortal souls, but as an evolutionary phenomenon, like any other living being. Because humans have uniquely large frontal lobes, humans seem to be the only creatures with the capacity to grow their foods, find shelter from extremes of climate, create new sources of energy, gain understanding of their physiology, and invent specific cures for their ills. Humans are the only creatures with the glimmering of an understanding of the universe, of

mortality, and of some of the things that must be done to protect the globe and the species from extinction. An enormous cognitive capacity, compared to other animals, has enabled humans to take some distance from ourselves to examine human behaviors as though they were studying a different species and to recognize their capacities for jealousy, hatred, vengeance, lust, and the urge to dominate. One need only consider the wars of the 20th and 21st centuries to appreciate the advances that have been made in the ability to annihilate one another.

For centuries, if not millennia, scholars assumed that humans were the only animals capable of forming groups to make war on each other, but humans are not alone. Jane Goodall documented raiding parties of male chimpanzees that perpetrate gratuitous violence on other, more vulnerable groups of chimpanzees. Primatologists debate whether chimpanzee violence is innate or is caused by human interference (e.g., feeding the animals and changing their habitats).

In a five-decade study of 18 chimpanzee communities researchers reported that of 152 killings by chimpanzees in 16 of the communities, 92 percent of the attackers were male (as were 72 percent of the victims) and that most killings involved intercommunity attacks in which the killers greatly outnumbered their victims. The investigators in this project asserted that variations in killing rates were unrelated to the impact of human beings. The researchers interpreted their findings as supporting the explanation that lethal violence among chimpanzees is the result of evolutionary adaptive strategies to increase access to food and to mates. The primates did not catch it from humans. Other primatologists questioned these conclusions.

There is substantial proof of human biological closeness to the "other" great apes (chimpanzees, gorillas, orangutans, and bonobos). Advances in the analysis of the structure and composition of deoxyribonucleic acid (DNA) have revealed that humans are genetically more closely related to chimpanzees than to any other known species. From the perspective of chimpanzees, humans are their closest relatives. Their DNA is more similar to human DNA than it is to the DNA of other great apes. The question remains of whether the large human cerebrum, which enables humans to devise weapons of mass destruction while simultaneously enabling humans to conceptualize moral codes by which to live, will, in the end, have survival value. How will people use their heads?

HORMONES GENDER AND VIOLENCE

Are all humans equally aggressive? Clearly not. Males of the human species, like chimpanzee males, and males of most other species are more violent than females. Whether measured in childhood in terms of roughness of play or in adulthood in terms of violent crime, the human male behaves in demonstrably more aggressive ways than the female.

These differences are not peculiar to a particular society but have been observed cross-culturally in societies as different from each other as Ethiopia and Switzerland. In the United States, men are about eight times as likely to commit murder, nine times as likely to commit armed robbery, and four to five times as likely to commit aggravated assault as women. Of note, men are also many times more likely than women to commit nonviolent crimes such as burglary, car theft, and drunken driving. Women are catching up to men in the areas of fraud and embezzlement, gentler, less physical forms of criminality.

The greater inherent aggressiveness of human males is also reflected in the patriarchal nature of almost all human societies. In fact, a 1971 survey of 93 societies around the world revealed that men retained all important political positions in 88 percent of them, and that in the majority of societies men tended to hold all significant leadership roles within kinship groups as well. Most surprising is the finding that maltreatment of women by men (i.e., wife beating) is more the rule than the exception in societies the world over. Needless to say, males of the human species are many times more likely to commit rape than are females. The human male is, in short, the predominant maker and breaker of the rules in human society.

Much as many may wish to attribute most, if not all, of the differences between human male and female aggressiveness to cultural influences, overwhelming evidence exists that, long before birth, the die is cast; male and female brains develop differently in utero. The architectural and functional differences between male and female brains are, to this day, "areas" of controversy. While some investigators have noted marked differences in what are referred to as the "sexually dimorphic" areas of the hypothalamus and corpus callosum, others have highlighted similarities in human brains. However, it is agreed that human males have about 8 to 10 percent larger brains than females and, what is more, the greatest size discrepancies are in the cerebrum, that alleged repository of reason and impulse control.

Gonadal hormones play important roles in the development of aggressive behaviors. Gonadal hormones are especially critical during fetal and neonatal development. These hormones have been shown to affect sexually dimorphic areas of the brain, and their early presence has been reported to be important to the normal development of aggressive behaviors in such diverse species as fish, lizards, birds, and chimpanzees.

Androgens such as testosterone are metabolized in the brain to form estrogens. It is now well recognized that estrogens influence the sexual differentiation of the male and female brains and thus influence behavior. Estrogens have been found to influence dendrite formation at synapses. They have permanent organizing effects on the structure and function of the brain during critical prenatal periods. The transformation of testosterone to the estradiol essential to brain differentiation requires

activation of the cytochrome P450 aromatase enzyme. Of note, in vitro studies of neurons indicate that neurons from the male fetal hypothalamus have higher aromatase activity than neurons from the female fetal hypothalamus. The relationship of the masculinization of the fetal brain to the production of neurotransmitters and receptors in the brain is not yet completely understood. However, differences in the amount, turnover, and localization of catecholamines have been documented in the brains of male and female rats. The effects on behavior of gender differences in the central nervous system (CNS) are extremely complex. It has been reported that the steroid derivatives progesterone, testosterone and glucocorticoids, referred to as neuroactive steroids or neurosteroids, can also alter behavioral responses and are important in sex-related differences. The behavioral effects of neurosteroids are numerous and include sedative, anesthetic, and anxiety reducing properties.

Several MRI and fMRI studies have reported structural and functional differences between the brains of male and female Homo sapiens. Investigators have reported sex differences in different parts of the orbitofrontal gray matter. This area is of particular interest because men who come in conflict with the law tend to have less functional orbitofrontal gray matter than their law-abiding peers. An exception: Raine has reported areas of hypertrophy in parts of the frontal cortex in white-collar criminals.

As technology has advanced, brain researchers have been able to identify other, smaller and smaller brain areas in which men and women differ. Investigators are zooming in on numerous brain areas affected by fetal testosterone (FT) including the right temporoparietal junction/posterior superior temporal sulcus, the planum temporale/parietal operculum (PT/PO) and posterior lateral orbitofrontal cortex. Subregions of both the amygdala and hypothalamus were also found to be sexually dimorphic.

Some of the literature on dimorphism reads like an exercise in political correctness. Papers on the subject tend to soft-pedal any possible intellectual differences between men and women, especially if the differences seem to favor the male brain. A recent publication from the Proceedings of the National Academy of Sciences got directly to the point: “Sex differences in human behavior show adaptive complementarity. Males have better motor and spatial abilities whereas females have superior memory and social cognition skills.” Even Steven! Such efforts at equality, such denials of differences seem to have encouraged women (naturally less aggressive than men) to demand combat roles in times of war rather than sticking to their more obvious talents for less violent means of settling differences (e.g., talking and reasoning).

Males (rats and humans) have one X chromosome and one Y chromosome, whereas females carry two X chromosomes. The Y chromosome contains that particular gene (SRY) that, during gestation,

triggers the outpouring of androgens. Not only does this flood of testosterone affect the density of neurons in different parts of male and female brains, but it also results in sexually dimorphic *circuits* among different brain areas.

The study of brain formation and function has made a giant leap ahead with the new techniques of optogenetics and pharmacogenetics as well as innovations in researchers' ability to track and record in vivo brain activity during complex social behavior. At this point in time, most of these new technologies are used to study animal brain function, particularly the brain of the laboratory rodent. The assumption is that the underlying neural mechanisms that drive aggressive behavior are likely to be conserved across species.

Using the new techniques of optogenetics and pharmacogenetics as well as newer in vivo recording methods, investigators have identified areas in the ventromedial hypothalamus, ventrolateral area (VMHvl) as a critical area both for intermale aggression and male on female sexual activity. David J. Anderson and colleagues identified a ventrolateral subdivision of the mouse VMHv1 marked by estrogen receptor 1 (Esr1+) that, when optogenetically stimulated, elicits attack behavior. However, to the surprise of the researchers, they discovered that weaker optogenetic activation of these neurons promoted mounting (i.e., sexual) behavior. The finding that at intense stimulation the same hypothalamic pathways associated with aggression, at lower stimulation are associated with sexual acts may shed light on the pairing of sex and aggression in human rape. Rape is not strictly a violent behavior. It is aggressive *and* sexual and maybe now it is known why. Research is focusing more and more on the interactions of steroidal hormones and neurochemical systems. Many studies have looked at the relationship of hormones to serotonin (5-HT) function in males and the relationship of dehydroepiandrosterone (DHEA) to the modulation of gamma-aminobutyric acid (GABA) receptor function in females. There is a growing body of data suggesting that serotonin, via its 5-HT 1A and 5-HT 1B receptors, is a major influence on the regulation of aggression in species from crustaceans to humans. Increased 5-HT has repeatedly (though not always) been demonstrated to decrease aggression, whereas a lack of it is associated with increased aggression.

If particular hormones already give men a head start on violence, one can only imagine the possible consequences of their ingesting or injecting stimulants or steroids in the interest of bodybuilding and enhancing stamina. Anabolic-androgenic steroids and some of their metabolites as well as stimulants have been associated with psychosis, disinhibition, aggression, and acts of violence.

The author was asked to evaluate a heretofore upstanding law enforcement officer who stopped a young lady for an alleged traffic violation. He then dragged her into a field of high grasses

and raped her repeatedly. Oblivious to her terror, he returned the victim to her car, gave her his card, and advised her to call him the next time she wanted to get together. The officer—a man with an impeccable record—then drove on to the department store where his recently estranged wife worked. He kept his gun in his holster when he entered the store. He just wanted to talk with his wife and win her back. In the course of attempting to clear the store of customers, the officer brandished his revolver, which went off, killing a salesclerk. During the hours that followed, the officer kept hostage several women, whom he forced to commit sexual acts with each other and with him. As time wore on, the captives got hungry and thirsty, whereupon the officer raided the till, found change, and bought candies and sodas from a machine. Thus, when he was finally captured, he was charged with felony murder (i.e., murder in the course of a robbery).

A psychiatric evaluation revealed that the defendant had periodic episodes of deep depression, which contrasted with his usual hypomanic state of mind. During the psychiatric examination, he was expansive and delusional about his powers to influence the police and the court and to be set free. A detailed history obtained from close relatives and from psychiatric records revealed a strong family history of bipolar mood disorder. Like many sexually violent offenders, the defendant also had a history of early, ongoing childhood sexual abuse that could be documented because it had occurred with a neighbor and not a family member. In most cases in which male perpetrators of sexual violence have themselves been victimized, these data are hard to obtain because the abuser is usually a close relative. Family protects family.

The clinical team was nonplussed. Why had this upstanding officer of the law suddenly lost it and raped and murdered? After re-examining hundreds of pages of educational, social service, and medical records, the clinical team found a previously overlooked single allusion to the defendant's recent use of "body building vitamins." Before then the clinicians had only inquired about alcohol and illegal substance use. When it finally occurred to the psychiatrist to ask about his body building "vitamins," he readily admitted that over the past several months he had been taking an over-the-counter potion called Ripped Fuel, and an herbal concoction that at the time contained ephedrine and caffeine. He also had been consuming large amounts of a potentially mind-altering drug, chromium picolinate. In addition, to stay awake and work double shifts, he had been drinking dozens of cans of caffeinated sodas each day. He was wired and sleep deprived. Clearly, his wild sexual and aggressive explosion was tripped by a drug-induced, manic, psychotic episode in a man primed by early, ongoing sexually abusive treatment. The defendant was unaware either of his predisposition to mania or of the power of his "vitamins" to precipitate psychosis. Note: He had never ingested illegal substances. Thus, he had a solid defense of diminished capacity. An argument might even have been made for insanity because, at the time of his offenses, he could not have appreciated the nature of his acts or that they were wrong (a legitimate M'Naughton defense). Would a man who thought he had done wrong have given his calling card to one of his victims?

Unfortunately, this man's manic braggadocio and his sexual advances toward one of his female defense attorneys alienated his entire defense team. (This author has seen literally dozens of grandiose, manic murderers who manage to alienate not only their male defense attorneys but also their male psychiatrists. They seem to threaten the authority of their "professional" examiners who then dismiss them as boastful psychopaths or whatever the diagnostic manual's term for psychopath is in use at the time of the evaluation.)

After the evaluation, the legal defense team was informed of the defendant's mania and the effects of drug ingestion on his behaviors (offenses). The defense team, unable to overcome their antipathy, advised the defendant to plead guilty and throw himself on the mercy of the court (of which there would seem to be little in parts of the United States). He was thereupon found guilty and sentenced to death.

Approximately 10 years later, a newly appointed, experienced public defender whose ego was strong enough to deal with the defendant's ego, came upon the clinical team's neuropsychiatric evaluation and contacted the author. By virtue of the previous clinical findings, the new attorney was able to get a new sentencing hearing at which two psychiatrists, a psychopharmacologist and a neuropsychologist testified to the defendant's compromised state of mind at the time of the offense. The jury is still out on this case, but the new defense team is optimistic that the death sentence will be commuted to life without parole. The wheels of justice move slowly.

GENES, NEUROTRANSMITTERS, AND VIOLENCE

It has long been recognized that criminality runs in families. However, it is hard to tease out why this statement is true. In the 1950s, a study of more than 1,000 Glasgow boys who left school at age of 14 years revealed that the greater the number of convicted family members, the more likely a boy was to be convicted of a crime.

In David Farrington's Cambridge Study of Delinquent Development, he and his colleagues found that having a convicted father, mother, brother, or sister was associated with a child's convictions. Same-sex relationships were found to be more strongly associated with criminality than opposite-sex relationships, and having an older criminal sibling was more closely associated with a child's criminality than having a younger one.

Twin studies, as well as studies of adopted away children of criminal biological parents, have been interpreted as showing a genetic component in the genesis of violence. A recent study of a large sample of twin pairs, using parent-reported data, found an association between early temperament and later aggression. Emotionality and high activity, especially in boys, was associated with subsequent aggressiveness.

Early studies of adopted away children of criminal biological fathers and noncriminal biological fathers revealed that the children of criminal fathers were more likely to become antisocial than were the offspring of noncriminal fathers. However, those children with both criminal biological fathers and criminal adoptive fathers were at greatest risk of becoming antisocial, suggesting that both intrinsic and environmental factors contribute to antisocial behavior. Of particular importance is that the antisocial behaviors in these samples were, reportedly, nonviolent.

Others have compared identical twins (concordant for all genes) and fraternal twins and have found that identical twins are more similar to each other in terms of aggressive behaviors than are fraternal twins. Even identical twins separated at birth are more likely than fraternal twins or full siblings to become involved in criminal activities. These kinds of studies seem to support a genetic predisposition to antisocial behavior. The important two words: "seem to."

Here's the hooker. What researchers find is limited to what they look for or ask about. Studies these days tend to be large and impersonal. By large is meant studies of documented characteristics of hundreds, even thousands of subjects. Such studies are too large to allow for lengthy neuropsychiatric assessments of the subjects, much less of their families. Interviews in very large studies tend to be formulaic, and conducted by trained interviewers rather than experienced clinicians. They gain in uniformity what they lack in the flexibility necessary to pursue areas of special interest when they come to light. Data are gathered from checklists filled out by subjects, interviewers and observers (e.g., parents and teachers), and/or from records, usually juvenile justice records and/or adult criminal records.

Such studies allow for sophisticated statistical analyses. More often than not, they confirm that identical twins and/or fathers and sons are more likely than nonbiologically related relatives to get into hot water. The question is, why?

The answers to that kind of question depend not only on the records one locates, but also on the time one spends interviewing the subject and relatives, and the detail of the questions one asks. If one asks only whether a father, or sibling or other relative ever had trouble with alcohol or with the law, or even with mental illness, or if one looks only at arrest records, one will often find the names not only of parents but also of the defendant's "sisters and his cousins and his aunts" (Gilbert and Sullivan). However, one cannot simply ask about mental illness. The answer to that is usually "no." One must take the time to inquire about relatives who were "going and going and going" or "too loud" or "too talkative"; "too suspicious"; "too controlling"; relatives who were "performers" or "tycoons" or "big spenders" or "gamblers." Having identified such relatives one must then ask in greater detail about that person's behaviors. In that way the researcher/clinician is more likely to come upon a family history of manic or hypomanic or even psychotic signs and symptoms in relatives. After gathering such information, the next question must be "who else in the family . . . and who else?" If one takes the time to ask about people in the family with "moodiness," "depression," "suicide attempts," "periods of isolation and withdrawal," "over-suspiciousness," "episodes of ragefulness" as well as "alcohol" or "drug" problems, one is more likely to uncover a hitherto unknown family history of severe mental disorders that manifested themselves as antisocial aggressive behaviors and landed them in jail. Many of the relatives of the murderers the author has evaluated had documented family histories of severe psychiatric illness when enough time was taken to look for those kinds of data. Many had also been imprisoned. Similarly, when taking a medical history from the perpetrator one cannot settle for "Any serious accidents?" If the answer is "none" one must pursue the topic (e.g., "How about bicycle accidents?," "Car accidents?," and "Falls?"). Any positive responses must be followed by "When else?" And "When else?" Ask. Look for scars, not just on the face, head or arms but also on the back, legs, even on the genitalia.

The secret is time. It takes a lot of interviewing and record reviewing to uncover the idiosyncrasies of perpetrators and their first, second and even third-degree relatives. It takes still more time to dig up the institutional records or find the other family, friends and relatives who can document and support indications of severe psychopathology and abuse in several generations of family members. Many of these psychiatrically impaired family members have at one time or another come in conflict with the law. That does not mean that the family carries a "warrior gene" for violence or criminality. It is tempting to settle for what one can get easily and en masse

(e.g., arrest records and birth records). Large record review-based studies are a bit like buying in bulk at Walmart. Lots and lots of stuff but the quality of the merchandise is pretty uniform. Now Tiffany is a different story. You pay a lot for a unique bauble but boy oh boy do you take your time studying those tiny jewels with care; are there flaws? Do you need a magnifying glass?

CASE STUDY: LANDRIGAN V ARIZONA

Some researchers who deal primarily with large cohorts or esoteric genetic concepts try to bring their studies to life with a colorful clinical example. Such was the case of Jeffrey Landrigan. Landrigan had killed twice and was sentenced to death in Arizona. After lengthy appeals regarding the failure of his lawyer to develop mitigating evidence, Landrigan was put to death. His case was used recently in scientific publications to illustrate the genetics of violence. He was said to have been adopted at birth into a loving middle-class professional family. The researchers stated that he was nevertheless a particularly troublesome child from the beginning. They said that even within this lovely family his behavior progressed from temper tantrums at 2 years of age, abusing alcohol at 10 years of age, being arrested for burglary at 11 years of age, abusing drugs as a teenager, to killing his first victim at age 20 years. After escaping from prison, he killed again and was sentenced to death.

By chance, another inmate noticed that Landrigan bore a striking resemblance to another career criminal who also abused drugs and also killed twice. This man turned out to be Landrigan's biological father. With a bit more inquiry, the researchers learned that Landrigan's grandfather had been an "institutionalized criminal," but the nature of the institution remained a mystery. Alas, upon discharge from the "institution," grandpa was "shot by the police." According to the researchers, Landrigan's great grandfather was "a notorious bootlegger." Now from bootlegging to murder is a long stretch (literally and figuratively). Nevertheless the original researchers stated that Landrigan's case documented not just the intergenerational transmission of violence but also how "the adoption design separates the genetic influences of the biological parents from the environmental influences of the rearing home." It did not.

At the beginning of Landrigan's trial, his lawyer had done little if any work investigating either his birth history, his medical history, or even his family history. In court, Landrigan was obstreperous, repeatedly interrupting his lawyer and boastfully telling the judge to go ahead and sentence him to death, which the judge obligingly did. Several accounts of Landrigan's behavior in court suggest that he was in an excited state and was unable to keep his mouth shut except to interrupt his lawyer and challenge the judge.

After the trial, Landrigan calmed down enough to realize the consequences of his court behavior and, with the help of public defenders, gathered mitigating evidence. His initial appeal for postconviction relief, based on his first lawyer's failure to present any mitigating evidence at all, was "summarily denied" by the Arizona Supreme Court. Eventually, in 2005, the United States Court of Appeals, Ninth Circuit, agreed to hear the petition of ineffective assistance. They stated, "We conclude Landrigan has raised a colorable claim that his counsel's performance fell below the objective standard of reasonableness and that he was prejudiced by their errors." Landrigan's initial trial attorney was deemed ineffective because "Counsel is not deficient for failing to find mitigating evidence if, after a reasonable investigation, nothing has put the counsel on notice of the existence of that evidence . . . In this case, however, the initial investigation did reveal potential mitigating evidence . . . in addition to the evidence described above . . . Counsel did little to prepare for the sentencing aspect of the case . . ."

A psychologist, who had interviewed Landrigan pretrial, recommended that additional testing be conducted but Landrigan's attorney failed to authorize any follow-up tests. The court also cited as potential mitigators that the original defense attorney had evidence that Landrigan's birth mother was alcoholic and drug addicted during her pregnancy with him and failed to so inform the court. Furthermore, she did not give him up at birth as the geneticists assumed, but attempted to raise him until Landrigan was 6 months old, a crucial period of development. She then abandoned him at day care. One day she just did not bother to pick him up. Thus, Landrigan was not adopted at birth. The state took custody of him and, at 8 months

of age he was adopted. However, further investigation brought to light the mythology that he was "adopted at birth into a loving middle class professional family and was nevertheless a particularly troublesome child from the beginning . . . etc." The court continued, "Landrigan began drinking at an early age because his adoptive mother was an alcoholic (at times consumed a fifth of vodka) and would walk around nude in front of him . . ." According to court records, she would frequently slap him and once even hit him with a frying pan. There was strong evidence that early on Landrigan suffered from serious psychiatric illness (i.e., Landrigan overdosed in class in eighth or ninth grade). Subsequent to his trial, according to court records, Landrigan did undergo "a thorough neuropsychological evaluation (which concluded) the concurrence of the above-mentioned factors—Landrigan's genetic makeup, in utero exposure to teratogenic substances, early maternal rejection, and his troubled interactions with his alcoholic, sexually inappropriate, abusive adoptive family—resulted in disordered behavior that was beyond the control of Mr. Landrigan . . . and left him unable to function in a society that expects individuals to operate in an organized and adaptive manner . . ." Thus, spoke the court. Subsequent to his trial, Landrigan's appeals attorney arranged for a neurological examination by Pamela Blake MD, at the time, an assistant professor of neurology at Georgetown. According to her report, Landrigan's mother "would come home from work and start drinking; she would drink until she would pass out on the couch." Dr. Blake stated "The neurological examination indicates several abnormalities consistent with impairment of the frontal lobes of the brain. The frontal lobes control higher level ("executive") function, impulses, regulation of behavior and actions. . . . Numerous studies of brain function have demonstrated a correlation of this type of brain disorder with repeated aggression." Dr. Blake documented "paratonia in all four extremities . . ." "inability to perform three step hand movements (The Luria)," and "abnormal antisaccade functions." All of the above are signs of significant frontal lobe dysfunction.

Eventually Landrigan's case reached the Supreme Court of the United States where his death sentence was confirmed by a 5 to 4 majority. Justice Stevens, wrote an eloquent dissent. Stevens said "Significant mitigating evidence that may well have explained respondent's criminal conduct and unruly behavior at his capital sentencing hearing was unknown at the time of sentencing . . . correspondent's counsel failed to conduct a constitutionally adequate investigation."

On Tuesday, October 26, 2010, Jeffrey Landrigan was put to death. During the week preceding his execution, the sentencing judge from his initial trial told a clemency board that she would have given him a lesser sentence had she been aware of the brain damage caused by his mother's heavy drinking while pregnant and had she known of his turbulent childhood. She said, "The death penalty in this case is not appropriate and never has been." The clemency board reportedly split on whether or not to advise Arizona's Governor, Jan Brewer, to commute Landrigan's sentence to life without parole. She did not.

The author of this chapter chose to include this case study for two reasons: First, it illustrates the ease with which a violent defendant can be dismissed as inherently bad. Second, it calls into question the existence of a "warrior gene" that trumps the effects of a nurturing upbringing. The author's research team has evaluated literally hundreds of extremely violent individuals, most of whom came from chaotic, violent households. More important, careful examination of first- and second-degree relatives usually revealed a family history of severe mental illness (e.g., bipolar disorder, psychoses, intrafamily violence, and physical and/or sexual abuse). Similar to the Landrigan case, many of the authors' patients, particularly the men, who had been abused by fathers and/or mothers, like Landrigan, refused to allow these data to be revealed. The author is currently working on just such a case. She has seen murderers who go to their deaths (i.e., executions) rather than reveal the brutal, often sexually abusive behaviors of their mothers and fathers. In two cases, families threatened to cease visitation if sexual abuse were revealed. One family actually cut off all communication with their son when extraordinary physical and sexual abuse was revealed. Their son's death sentence was commuted but he has not seen his parents in years. The author has been obliged to remain silent in most of these cases.

This case was also chosen because more often than not, what looks like a genetic transmission of violence is, rather, a predisposition to severe mental illness and/or a history of repeated CNS trauma that on clinical evaluation indicates frontal and/or temporal lobe/amygdaloid abnormalities. Landrigan had frontal lobe abnormalities. Violence, child

abuse, and mental illness are handed down from generation to generation and thus, on the surface, violence appears to be genetic. Identifying these constellations is the mystery to be solved. Combinations of intrinsic and extrinsic interactions perpetuate what looks to outsiders like a simple genetic trait. The authors' team has encountered family after family in whom they can trace intergenerational mental illness, violence, child abuse, and institutionalizations. State prisons and county jails appear in the histories of violent families much as Harvard and Yale appear in the histories of intellectual/upper middle class families. Certainly genetics plays some role in both kinds of pedigrees but DNA is rarely if ever the whole story. (Of note, one of the protective factors from violence is intelligence quotient [IQ].)

In the middle of the 20th century, at least in part in reaction to Lombroso's earlier theories that criminals represented a degenerate biological phenomenon, investigators looked almost exclusively at society for explanations of criminality. The past half century has seen an almost 180 degree about face in the study of violence and criminality, as investigators have turned back to biology to try to explain deviance, especially to explain violence. Advances in the so called "hard" sciences (e.g., genetics, biochemistry, cellular biology, neuroanatomy, and neurophysiology) have increased the ability to identify, measure, and even visualize biological similarities and differences among people, thus encouraging this direction of research. Studies of the relationship of such neurochemicals as serotonin, norepinephrine, dopamine, vasopressin, GABA, glutamates, and nitric oxide to aggressive behavior in animals have encouraged similar studies in humans. Initially, the finding of a relationship between low cerebrospinal fluid (CSF) 5-HIAA (a metabolite of serotonin) and impulsiveness and aggressiveness in humans seems to shed light on the question of why some individuals are more aggressive than others. Studies of men depleted of tryptophan, the amino acid precursor of serotonin, have indicated that they become more aggressive in a laboratory setting than do men not so deprived. Of note, those men with a tendency toward hostility as measured on an aggression questionnaire reportedly responded more aggressively to tryptophan depletion than did nonhostile men. Chimpanzees with high blood levels of serotonin are reportedly nicer to their peers and more likely to groom them.

However, the relationship of serotonin levels and serotonin system responsiveness to degrees of aggression is not simple. A study of young (7 to 11 years old) aggressive boys with ADHD revealed a *greater* prolactin response to fenfluramine (i.e., greater serotonergic responsivity) in aggressive than in nonaggressive boys. A study of blood serotonin levels in adolescents diagnosed with conduct disorder reported that those with early onset maladaptive behaviors had *higher* blood serotonin levels than later offenders, and that blood 5-HT correlated positively with ratings of aggressiveness.

In a broader study of the relationship of brain 5-HIAA levels in infants to subsequent behavioral characteristics, researchers included environmental factors. Clarke and colleagues followed a sample of children

from their first weeks of life to 30 months of age. They found but a modest correlation between low CSF 5-HIAA and aggressive behaviors. In contrast, they found that environmental factors such as poverty, and poor family cohesiveness were strongly associated with behavior problems at 30 months of age. Another study documented a relationship between low socioeconomic status, low prolactin response, and increased aggression. Finally, a meta-analysis of 39 studies linking 5-HIAA levels to violence in humans found that psychiatric patients as a group, whether violent or nonviolent, had similar 5-HIAA levels. The mean 5-HIAA level in both violent and nonviolent patient groups was lower than the mean level in groups of normal, healthy volunteers. Thus, serotonin levels may prove to be indicators of overall psychopathology that only at times is manifested by aggressiveness.

One of the most important and influential genetic studies, was Brunners 1993 report of a family in the Netherlands in which five generations had behavioral disorders, including arson, rape, exhibitionism, and attempted suicide. Many individuals had been aggressive. Each of the affected family members was found to carry an abnormality in the gene controlling the production of monoamine oxidase A (MAOA), the enzyme that metabolizes several of the neurotransmitters recognized as affecting cognition and behavior. In fact, this gene has been dubbed by some as the "warrior gene." Only the male children of mothers carrying the abnormal gene showed behavioral abnormalities.

Piton and colleagues reported on a French family with a similar genetic abnormality and with several family members suffering from a variety of behavioral abnormalities. Piton reported that several affected males had cognitive difficulties, a 7-year-old boy was autistic, impulsive and aggressive, and two maternal uncles had severely delayed psychomotor development and required placement in a special needs school. They too had autistic features, could not read or write, and were given to aggressive outbursts. Of note, both of these family members had been maltreated and sexually abused in early childhood, a topic to be revisited in this chapter.

Since the discovery of the MAOA genetic abnormality, some clinicians/researchers have coined the term "Brunner Syndrome," designating the MAOA abnormalities as a recessive X-linked disorder characterized by impulsive, aggressive behaviors and mild retardation. Others have taken issue with this phenotypic concept because of the wide variety of signs, symptoms and behaviors reported to exist among individuals carrying this MAOA abnormality.

Another gene, the serotonin transporter gene, has been associated with behavioral problems and with a number of different psychiatric disorders. Studies have focused on the short and long types of alleles of the serotonin transporter gene. People carrying the short allele have been reported to have lower than normal serotonin levels in their blood. Some studies have

reported a greater predisposition to aggressiveness in individuals with low cerebrospinal levels of serotonin.

Dopamine, another catecholamine neurotransmitter influences many mental functions including cognition and emotion. The dopamine receptors belong to the family of G protein-coupled receptors (GPCRs). DRD1 type receptors are associated with stimulatory function and DRD2 type receptors are generally associated with inhibitory functions. More than 200 polymorphisms have been identified in the sequencing of this gene. Variants of the DRD2 gene have been associated with alcoholism and addictive disorders, mood disorders and the schizophrenias. In addition to genes, a myriad of other biopsychosocial influences and constellations of influences contribute to a violent lineage. **The reader must keep in mind that, although each cell carries the same load of genes, when, where and how a gene will be "turned on" (i.e., activated) depends in great measure on environmental stimuli.**

One of the most complex genetic studies illustrating the above is that of Hexuan and colleagues. They examined the possible relationships between youth delinquency and violence and 403 variants of 39 genes against what they termed high/moderate levels of social control in contrast to low levels of social control (i.e., high/moderate parental attachment vs. low parental attachment). Their theoretical underpinning was that in both animal and human studies, changes in environmental conditions can influence expression of genes related to different phenotypes.

They found that estimates of the total variance in serious delinquency attributable to the 403 SNPs were, in and of themselves, nonsignificant at the 0.05 level. The researchers then examined the collective genetic contribution of the 403 SNPs under low social control (LSC) compared with high/moderate social control (HMSC). They found that genetic risk for delinquency and violence was significantly greater for subjects weakly attached to parents and school and loosely disciplined by them. The genetic risk for delinquency and violence was also increased by living in lower income neighborhoods. In short, their analysis showed "significant interactions of aggression related genetic variants and socioenvironmental variables . . . favorable social conditions are associated with smaller collective genetic contributions whereas adverse social conditions are associated with greater genetic contributions to adolescent delinquency and violence." This will be amplified when the nature and interactions of child abuse with neuropsychiatric variables are considered.

PSYCHOPATHY

The concept and subsequent reification of the diagnosis "psychopathy" has, to this author's mind, hampered the understanding of criminality and violence. The practice of using the term "psychopath" was spearheaded by Robert Hare, who in 1970 created a checklist of behaviors (based in part on

Hervey Cleckley's book, *The Mask of Sanity*) combinations of which Hare believed characterized a psychopath. He subsequently revised and updated the list. The list includes such items as "grandiose sense of self worth," "need for stimulation," "promiscuous sexual behavior," "impulsivity" and "many short-term marital relations" as well as "early behavior problems," "shallow affect," "juvenile delinquency," "revocation of custodial release," "lack of realistic long-term goals," and "criminal versatility." Grandiosity and shallow affect just do not fit together. Shallow grandiosity is an oxymoron. Then come a series of clinically subjective criteria (e.g., "callous lack of empathy," "cunning manipulation," "failure to accept responsibility for one's own actions," "lack of remorse or guilt," and "irresponsibility") some of which come under the rubric of "pleading not guilty." How does one define much less ascertain "cunning?" How about the adjective "callous?" Then there is the group of characteristics that is recognized in so many offenders with frontal lobe dysfunction, "poor behavioral controls," "impulsivity," "irresponsibility," and "shallow affect."

Clearly the items above, and their combinations, cover a multiplicity of different disorders. According to Hare, in many cases one need not even meet the patient. Just rummage through his records to determine what items seemed to fit. Nonsense.

To this writer's mind, psychopathy and its synonyms (e.g., sociopathy and antisocial personality) are lazy diagnoses. Over the years the authors' team has seen scores of offenders who, prior to evaluation by the authors, were dismissed as psychopaths or the like. Detailed, comprehensive psychiatric, neurological, and neuropsychological evaluations have uncovered a multitude of signs, symptoms, and behaviors indicative of such disorders as bipolar mood disorder, schizophrenia spectrum disorders, complex partial seizures, dissociative identity disorder, parasomnia, and, of course, brain damage/dysfunction.

Almost 20 years ago, this author evaluated a jittery, jumpy prisoner who had killed a policeman and whose psychiatric, neuropsychological, and neurological examinations were consistent with frontal lobe dysfunction. Nonetheless, he was advised to plead guilty and was sentenced to death. He was dismissed as a psychopath. Almost 20 years later, he was lucky enough to have a very senior, experienced appeals attorney assigned to him, a man who read our old reports and wondered why they had not been introduced at trial, at sentencing or at least on previous appeals. We evaluated him again and the signs of CNS dysfunction were even more severe than when we first saw him. This time, however, we consulted with Ruben Gur PhD, and we were able to obtain an fMRI. As can be seen in Figure 28.2–1, this man had diffuse cortical atrophy, the most serious of which was orbitofrontal thinning (see Figs. 28.2–1 and 28.2–2).

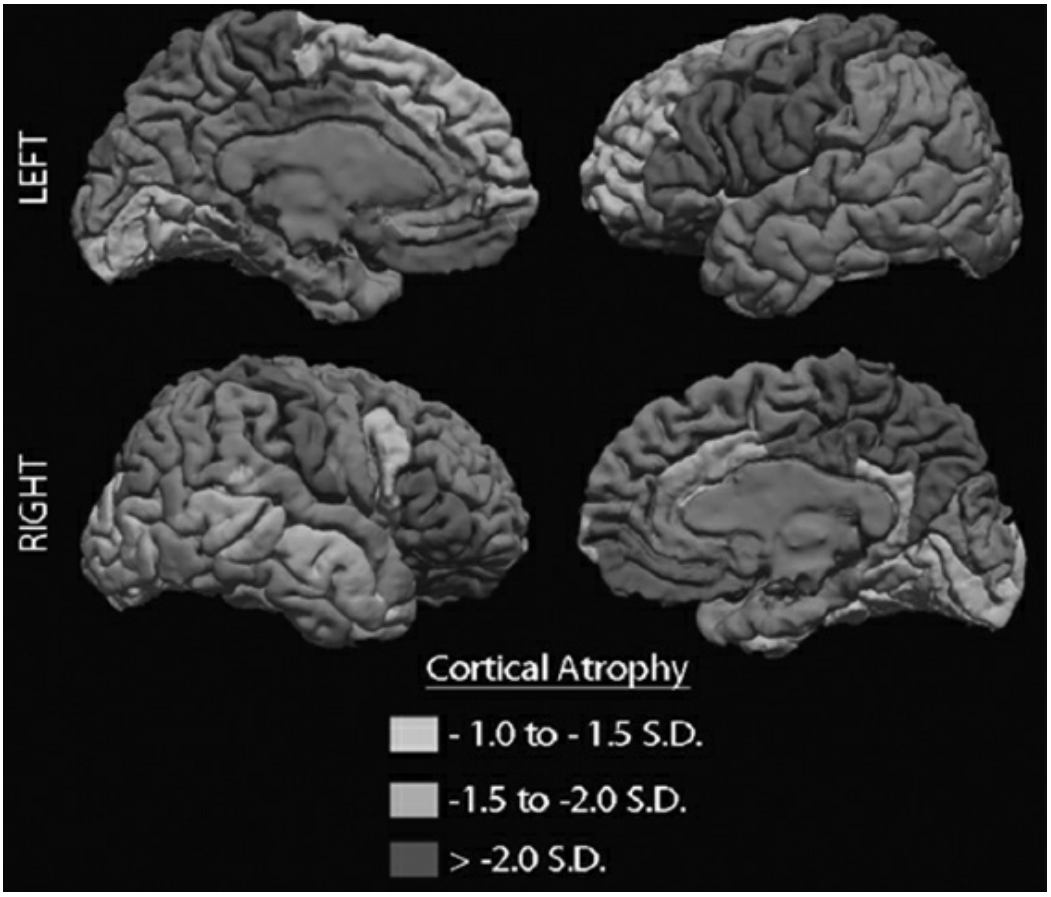

FIGURE 28.2–1. Analyzed by Ruben Gur, Ph.D. (Courtesy of Ruben C. Gur, University of Pennsylvania, Philadelphia, PA.)

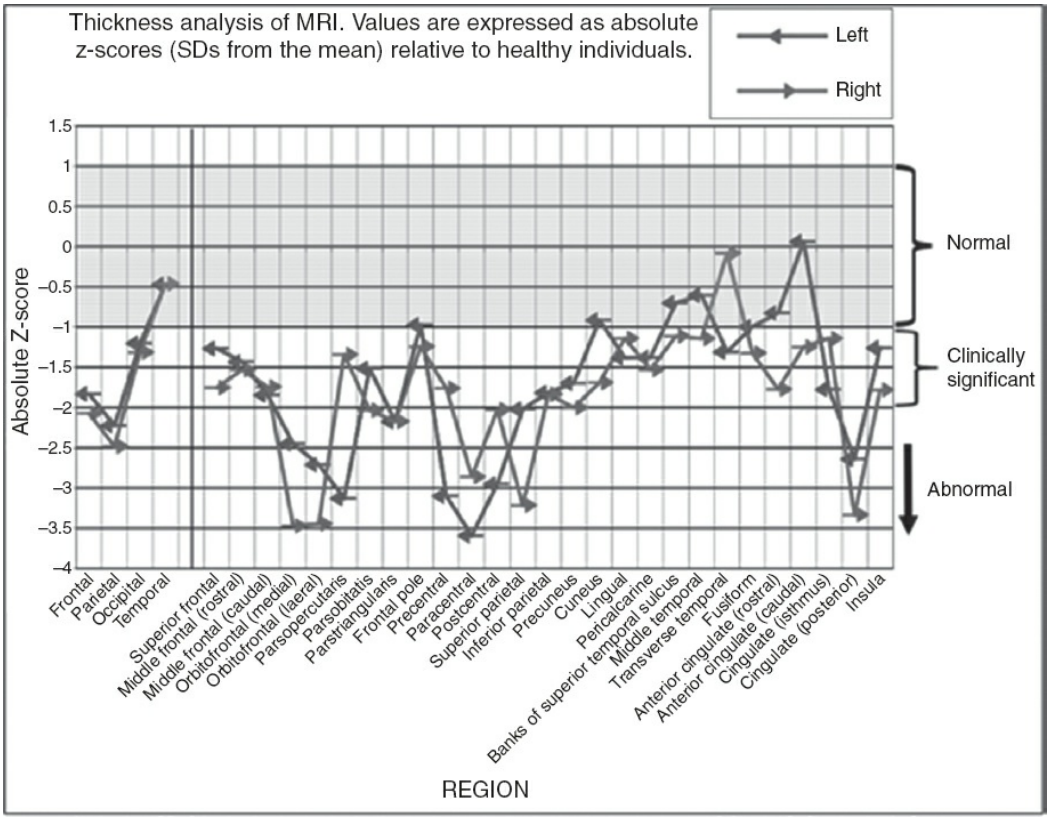

FIGURE 28.2–2. Analyzed by Ruben Gur, Ph.D. (Courtesy of Ruben C. Gur, University of Pennsylvania, Philadelphia, PA.)

The mish mash of different characteristics listed in the psychopath checklist has encouraged otherwise impeccable researchers, using sophisticated technologies (e.g., MRI, PET, and fMRI) to lump together criminals under the label "psychopaths" and compare them to noncriminals. Some have reported at least two totally different types of brain function or dysfunction in their psychopathic samples. The concept of psychopathy as described by Cleckley and "measured" by Hare just does not work. And yet, in the experience of the authors' research team, just about every violent criminal they have ever examined has at one time or another been diagnosed psychopathic or antisocial. One of the criteria for the diagnosis psychopath is having been designated "conduct disordered" as a youngster. The term conduct disorder has been used as a diagnosis in children and adolescents who have many of the same kinds of psychopathology as adult criminals. Thus, "conduct disorder" diagnosed in children (it is one of the criteria for adult "psychopathy") smoothes the way to the diagnosis psychopathy in adulthood. Both of these terms are used indiscriminately to designate individuals young and old whose behaviors are unpleasant or dangerous and whom examiners would prefer to spend as little time as possible interviewing.

ENVIRONMENTAL INFLUENCES ON AGGRESSION

Intrauterine and Perinatal Factors

The very position in utero of animals within a litter has been found to influence behavior. Female mice that develop in utero between two males are more aggressive after birth than those positioned between other females. Thus, animals from the same litter who have shared the same uterus at the same time have not had identical prenatal environments. Maternal stress during pregnancy has also been shown to affect the behavior of offspring long after delivery.

In humans, there is strong evidence that different kinds of noxious prenatal influences ranging from minor viral infections to maternal anxiety and psychological stress have been associated with childhood maladaptation. The adverse effects of maternal malnutrition, alcoholism, and other noxious substances on fetal development and on subsequent postnatal social and intellectual functioning are well documented. The authors' studies and those of others have documented the higher prevalence rate of histories of perinatal problems in the delinquent versus nondelinquent populations. Given the vicissitudes of life in utero and at birth, one should not conclude that even temperamental styles such as "emotionality" and early activity levels necessarily reflect genetic

endowment. The pregnancies of mothers who release their children for adoption cannot have been easy. There is now evidence that even heightened maternal cortisol as a result of stress affects the fetus and increases aggressiveness postnatally. Thus, the adopted away children of criminal parents in all likelihood had rocky in utero experiences. Moreover, the genes they inherited may well have predisposed them to a variety of different kinds of mental disorders other than psychopathy or antisocial personality disorder. In other words, behaviors that at first appear to be genetic may not be. "Things are seldom what they seem" (Gilbert and Sullivan).

NURTURING

Nurturing affects temperament. Animals bred to have an especially gentle nature, if cross-fostered by adult females of a violent strain, will become more aggressive than is their usual nature. Mice reared in the company of their fathers as well as their mothers grow up to be more aggressive than those raised only by mothers. Isolation during critical post-natal developmental periods can engender aggression in otherwise gentle animals. Thus, there is evidence from animal studies that innate temperament can be modified and that the quality of parenting and environment influence adult behaviors.

Very early experiences have especially powerful and enduring effects on physiological and psychological development and, therefore, on behavior. In animals, and humans qualities of parenting during the very first days and weeks of life are crucial for subsequent social adaptation. Although many reptiles are equipped to defend themselves shortly after birth, most mammals require parental protection for longer periods of time. And humans? Even longer.

In humans, the first months and years of life are periods of extreme vulnerability, emotionally, and biologically. Although the auspicious neuropsychiatric and emotional cards a person is dealt at birth may seem to predict a successful game, experience can reverse what under other circumstances should have been a winning hand. Research on attachment has shown that the ways that mothers and infants interact during the earliest weeks and months of life affect subsequent aggression in the child. Remember Landrigan. Adequate mothers sense their children's needs and respond appropriately. Mothers who lack tenderness and harbor suppressed anger toward their babies seem to engender insecure, avoidant attachments. Avoidant attachment has, in turn, been associated with subsequent aggressive behavior in boys.

Poor mothering takes its toll intellectually as well as behaviorally. Infants with disorganized attachments lag in their mental development irrespective of maternal IQ. Lyons-Ruth and colleagues found that the combination of early disorganized attachment and lags in cognitive

development were associated with aggressive behaviors at 7 years of age. Thus, it would appear that poor mothering affects brain function itself.

ABUSE

A copious literature attests to the association of maltreatment in childhood and aggression in adolescence and adulthood. The extent to which the category of "abused" children overlaps that of children who have suffered attachment disorders and the independent influences of each are difficult if not impossible to measure. Whereas complete emotional deprivation of infants leads to severe depression and even to death, lesser degrees of neglect have been associated with poor peer relationships and aggressive behaviors. Some research suggests that neglect can be as powerful an influence on the development of aggression as physical abuse. Research has shown again and again that severe ongoing physical and/or sexual abuse is associated with subsequent extremely violent behaviors.

Early repeated sexual abuse seems to have especially devastating consequences. Grotesque sexual crimes have been committed by men who, as children, experienced extraordinary sexual abuse. Many in fact have been so traumatized that they do not remember the torture.

How does severe abuse and neglect engender aggression? The answer is not simple and must be gleaned from a variety of different kinds of studies and observations. What is clear is that the human organism is so constructed that maltreatment affects every aspect of functioning—cognitive, emotional, and physiological.

A common result of abuse and/or neglect is direct injury to the CNS. Almost any CNS trauma may be associated with hyperactivity, emotional lability, diminished intelligence, poor judgment, and impulsiveness. Children with these kinds of deficits and vulnerabilities are less able to think ahead and understand the consequences of their behavior. In the authors' studies of violent juveniles and adults, medical histories were so replete with evidence of CNS insults that it was usually impossible to ascertain the exact cause of brain dysfunction. It is known that injuries to the CNS, even apparently mild concussions, can have a cumulative effect neurologically, cognitively, and behaviorally. Extremely violent juveniles have often experienced multiple accidents and injuries, many of them inflicted by those entrusted with their care. Of note, a recent study of perpetrators of domestic violence reported neuropsychological impairments of their executive functions, memory, attention, IQ, and empathy. In the authors' studies, they could not distinguish between children exposed to domestic violence and physically abused children since in most cases they went hand in hand. An examination of the heads, bodies, limbs, and even genitalia will often reveal traumas that the subject/patient cannot recall.

PSYCHOLOGICAL AND COGNITIVE EFFECTS OF ABUSE

Maltreatment has psychological consequences. Abuse changes the way in which the environment is perceived. Studies of animals as well as children have shown that one of the consequences of maltreatment is the development of hypervigilance. Defeat increases defensiveness in laboratory mice and conditioned defensiveness is generalized to other situations and other opponents. Likewise, abused children seem to become hypervigilant, repeatedly misinterpreting their surroundings, and perceiving ambiguous stimuli as threatening. Once a response has been established or conditioned (e.g., a hormonal or neurophysiological response to fear-inducing stimuli), it is easily re-evoked by exposure to similar stressors. Such recurrent, maladaptive, physiological responses seem to play important roles in the hypervigilance and retaliatory aggression of many severely abused, violent children and adults.

Child abuse also contributes to violence by diminishing a child's expressive skills. Severely abused children manifest specific kinds of deficits in verbalization. For example, toddlers who have been abused are less able than normal children to use words that reflect their inner states or feelings. The tendency of abused children to repress or deny their own painful feelings as well as the pain of others, their inability to articulate emotions, their emotional lability and impulsiveness, their tendency to distort reality and perceive threats where they do not exist, and their resultant tendency to put angry feelings into actions rather than words can become a relatively enduring adaptational style.

MODELING OF ABUSIVE BEHAVIORS

Aggression is also learned. There is sound experimental evidence that modeling plays an important role in the development of aggressive behaviors in animals as well as children. Aggressive behaviors (like more adaptive behaviors) can be learned through reinforcement. One of the most important contributions to the current understanding of the role of reinforcement in the genesis of aggression is Patterson's observation that when children's aggressive behaviors are punished severely by parents, they tend to continue. Farrington also found severe physical punishment to be a major antecedent of aggressive delinquency. Murray Strauss has documented the aggressive behaviors induced by corporal punishment. Probably the most powerful generator of aggression in animals and possibly in humans is the repeated infliction of pain. It is, therefore, no surprise that children who have been painfully physically abused tend to behave in more aggressive ways than their nonabused peers.

To summarize, child abuse contributes to violence in several ways: (1) It provides a model of behavior; (2) It teaches aggression through reinforcement; (3) It inflicts pain, a known powerful stimulus to violence;

(4) It creates neurophysiologic responses in the brain that both damage the brain and affect its functioning; (5) It damages parts of the brain related to impulses and their control; (6) It increases paranoia; (7) It impairs cognition and the ability to make judgments and foresee consequences; (8) It diminishes the ability to verbalize feelings; (9) It interferes with the child's ability to empathize with the feelings of others; (10) In the extreme, it leads to dissociation.

In short, early ongoing child abuse is the most powerful known generator of child, adolescent, and adult violence.

SOCIAL FACTORS AND VIOLENCE

Courtwright in his trenchant analysis of the sociocultural characteristics of the United States that contribute to its excessive violence, observes, ". . . other things being equal, the total amount of violence and disorder in society is negatively related to the percentage of males in intact families of origin or procreation . . ." He goes on to say, "Families rather than governments are the first best defense against violent and disordered behavior." He points out that the particular mixture of demographic, cultural, and social characteristics of the American frontier, its surplus of young, unparented, unmarried men, its code of honor, its racist attitudes, its heavy alcohol consumption, its lack of religious convictions or practices, its devotion to guns, and its lack of an effective system of law enforcement formed a matrix for violence. Moreover, the closing of the frontier did not eliminate the "national self-image" in that, to this day, Americans continue "to think a manly man is someone with a gun and an attitude." Courtwright states, "Indeed, it is possible to think of the urban ghettos as artificial, unusually violent frontier societies—vice-ridden combat zones in which groups of armed, unparented, and reputation-conscious young bachelors, high on alcohol, cocaine, and other drugs, menace one another and the local citizenry, undeterred." At this time, the United States is far more violent than most Western European and Asian countries.

In addition to the availability of guns, what else promotes violence? Material poverty and extreme economic inequality are associated with homicide rates; this is true if one compares violence in states, cities, or neighborhoods within cities. Whether the major contributing factor is socioeconomic inequality or racial subordination is academic. Violence, poverty, and racial subordination cluster together in the same populations. Being somewhat low on the socioeconomic totem pole is not so much associated with violence as is living in *grinding* poverty. But poverty alone does not seem to account for violence. Sociologists believe that a strong influence on criminality springs from "social disorganization," the lack of extended families to support youngsters whose families lack the stability to nurture them properly.

Numerous studies from various parts of the world have documented an

association between low or low-normal IQ and violence. It would be important to stress that neither impulsivity nor low IQ is necessarily inherited. The wealth of data cited above attest to the potentially adverse effects on intelligence, temperament, and behavior, of impaired attachment, child abuse, neglect, and neurologic insults. In summary, the kinds of external stressors and intrinsic vulnerabilities that play roles in individual violence (e.g., physical discomfort, lack of nurturance and stability, CNS dysfunction, increased rates of mental illness, and child abuse) are more pervasive in disadvantaged communities and contribute significantly to the disproportional rate of violence in the most impoverished sectors of society.

BRAIN DYSFUNCTION AND VIOLENCE

The part of the human brain thought to be involved in sexual and aggressive drives lies deep in the areas often referred to as the *limbic system*. The areas of this region thought to be most closely involved in aggression are the hypothalamus, septal area, and the amygdala. Lesions, electrical stimulation, and infusion of neuroactive substances into the amygdala have shown those nuclei to be necessary for animals to experience particular emotions, especially fear. Fear, in species as diverse as crocodiles and humans, is processed in the amygdala, stimulation of which elicits defensive, aggressive responses.

Humans can usually perceive threat, assess it intellectually, and choose to act or not to act in response to the stimulus. The prefrontal cortex allows humans to exercise some control over their instinctive responses. The reflexive, violent response of some repeatedly aggressive individuals to innocuous stimuli is often the result of paranoid misperceptions and immediate defensive responses. In the late 1960s, Vernon Mark and Frank Ervin in *Violence and the Brain* explored the association between electrical activity in the limbic system and aggressive behaviors in humans. By inserting electrodes deep in the brains of recurrently aggressive subjects, they reported abnormal electrical activity corresponding to episodic rages. Whether directed aggression can occur during a complex partial seizure generated within the limbic area remains a subject of debate among neurologists. However, aggressive behavioral changes have been reported in individuals with damage to the medial temporal lobe and with developing brain tumors in the limbic system. Violence can also occur interictally in patients with epilepsy. The author's own studies have indicated a high prevalence of a variety of psychomotor symptoms and paranoid ideation in violent delinquents.

The importance of obtaining a detailed medical history, as well as a detailed account of the offense(s) in question, is illustrated in the case of the serial killer Arthur Shawcross. Shawcross had a myriad of dissociative symptoms, as well as a history of early severe ongoing sexual abuse, that could easily have accounted for the rapes and murders with which he was charged.

However, a history of several serious blows to his head, evidence of which could be documented by the scars and indentations on his head, suggested the possibility that brain dysfunction might have contributed to his episodic violence. A detailed history of the offenses (i.e., what happened before, during, and after the offenses) revealed Shawcross's confusion and impaired memory of the acts. As Shawcross recalled, before the acts, he would be having a nice conversation with his victim. Then, out of the blue, something she did or said provoked a reaction. In response he would start to sweat and everything would turn bright white. A struggle would ensue and, shortly thereafter, his hands would close round his victim's neck. Shawcross would subsequently fall into a deep sleep. Hours later, he would awaken in his car, only to find a corpse in the seat beside him. Realizing what he must have done, he would drive around, seeking a place to hide the body. At least once, on a return visit to the body, he cut the victim's genitalia and ate them. Acts of extraordinary and bizarre viciousness have been committed by individuals with temporal lobe lesions (The psychiatrist for the prosecution opined that Shawcross's cannibalism was easily explained as an effort to destroy the DNA. Surely, doctor, there must have been an easier way!) The above account—an aura (bright lights and sweating), an act for which memory was impaired, followed by a deep sleep—had the hallmarks of a psychomotor or complex partial seizure. A review of prison records showed that Shawcross had been found unconscious in his cell on numerous occasions. An MRI of his brain revealed a cyst in the area of the amygdala on his right temporal lobe and scars in the white matter of both frontal lobes. A single, routine EEG was normal. His attorneys were advised to obtain a more sophisticated neurological assessment, including a quantitative EEG with hyperventilation and photic stimulation, but they refused. They chose rather to pursue a defense of dissociation. Such a defense is a big mistake when clear evidence of brain damage exists. Juries are more likely to believe what they can see (e.g., MRI and EEG printout). They can neither see nor hear dissociative states. Shawcross was tried, found guilty, and sentenced to life in prison for 10 of the 11 murders he had committed.

Just before going to trial for the one remaining murder, Shawcross's second attorney decided to obtain a quantitative EEG. It revealed spiking, not only in one temporal lobe but also in both frontal lobes. In spite of these findings, this new confirmatory evidence of complex partial seizures and their relationship to the crime in question, Shawcross's attorney advised him to plead guilty to the last charge, which he did. It was clear from the nature of the offenses that Shawcross would be institutionalized for life either in prison or in a hospital. The tragedy was the failure of the defense to educate the judicial system and the public about the ways in which damage to specific parts of the brain, and early, ongoing sexual abuse can engender extraordinary, repetitive sexual violence.

A recently published CT and MRI study of 287 male prison inmates, comparing 162 violent offenders with 125 nonviolent offenders (none previously diagnosed mentally ill), and 52 noncriminal controls reported that offenders displayed a significantly higher rate of morphological brain abnormalities, with violent offenders having significantly more abnormalities than nonviolent offenders. Differences were statistically different for frontal/parietal cortex, medial temporal structures, third ventricle, and left but not right lateral ventricles. The authors concluded that their findings raised questions regarding the culpability of violent offenders with brain abnormalities.

PET and studies of regional glucose metabolism have revealed hypoactivation in lateral and medial zones of the prefrontal cortex in individuals given to impulsive aggressive acts. These kinds of frontal lobe deficits evident on PET scans have also been supported by findings from neuropsychological testing of patients with known prefrontal lesions.

In short, results of studies from a variety of different fields suggest that

many impulsive, aggressive individuals have abnormalities in the structure and/or function of the brain and in circuits that interpret outside stimuli and regulate emotional response.

Just as it is now known that what heretofore was called schizophrenia is a multiplicity of different psychoses with different genetic constellations and different biopsychosocial correlates so enough data have been accumulated regarding biopsychosocial differences between aggressive and nonaggressive offenders to separate them into different diagnostic categories and not label the group psychopathic. It is important to cease trying to cram them all into the same cubbyhole, psychopathy, or antisocial personality disorder. A review of the behaviors of the patients Cleckley described in *The Mask of Sanity* reveals that most were manic or hypomanic and would now be treated with mood stabilizers. Here is yet another distinct group of patients heretofore also dubbed psychopaths.

SEVERE PSYCHIATRIC DISORDERS AND VIOLENCE

The question of whether or not mentally ill individuals are more violent than people in the general population has become politicized. Thus, perfectly reputable scientists have sometimes turned statistical and intellectual somersaults to find ways to prove that severely psychiatrically disturbed individuals are no more a threat to society than the little old lady living next door. These politically correct efforts, while well intentioned, can have dire consequences for psychotic, brain-injured, impulsive and aggressive patients and their neighbors. They get discharged from mental hospitals only to find themselves frightened, alone, homeless, and confused. When made uneasy (often paranoid) by their surroundings, they often lash out to defend themselves against real and/or imagined threats and wind up in jail more often than in the therapeutic settings they need. Their aggressive, disordered behaviors in prison as well as in hospitals, often as not land them in solitary confinement, a setting known to induce psychosis in normal people and to exacerbate it in the mentally ill.

There is no question in this author's mind that the thrust of the 1998 MacArthur Foundation's Violent Risk Assessment, namely that with few exceptions, discharged patients from mental institutions are no more dangerous than other persons in the general population, is simply wrong. It is true that most individuals with mental illness are not violent. However, studies from around the world, from more developed and less developed nations, indicate that violence is more common among seriously mentally ill individuals than among mentally healthy individuals. These findings should not come as a surprise. After all, severe mental illness often distorts perceptions, increases suspiciousness, is accompanied by extremes of emotion, and impairs impulse control. When the brain is not functioning properly, humans often take umbrage easily and feel an urgency to protect themselves at all cost. It is called paranoia. And paranoia is part of almost

every serious mental illness known to man.

Studies have found that the prevalence of violence is approximately five times higher in individuals who meet the diagnostic and statistical manuals' criteria for an Axis I diagnosis than in people not meeting criteria for serious mental illness. The prevalence of violence in individuals who have schizophrenia, major depression, or mania/bipolar disorders was found to be similarly high (i.e., approximately 11 to 13 percent). Individuals with alcoholism or other substance abuse reportedly commit violent offenses (i.e., 25 and 35 percent, respectively) more often than patients with schizophrenia or major depression alone. Those with the highest prevalence of violence are young lower-class men with substance abuse disorders and/or major mental disorders. Of note, almost all of the differences in prevalence of violence between patients and nonpatients seem to be accounted for by the presence of active psychotic symptoms. Patients reportedly are no more violent than anyone else when they are not actively psychotic.

In the February 2008 issue of *Psychiatric Services*, E. Fuller Torrey and Jonathan Stanley reviewed the critical literature published since the MacArthur report and in a dialogue with two of the report's authors, challenged their major conclusions. The MacArthur study followed 951 psychiatric patients ages 18 to 40 after discharge from state and university psychiatric inpatient units.

As Torrey and Stanley reported, the prevalence of violent, sometimes murderous behaviors among the discharged patients, was high. Among the 951 discharged patients, 262 (27.5 percent) committed at least one violent act over a 41-week follow-up. This group was responsible for 608 violent acts; 1 percent of the violent acts were homicidal and there were 6 victims of these murders. The homicide rates in the United States during the period of the MacArthur study were at a peak, ranging from 8.2 per 100,000 to 9.5 per 100,000. In the most impoverished, minority populations, among young adults, the rates approached 200 per 100,000. These numbers pale in comparison with the homicide rate in the discharged patient sample that was over 600 per 100,000 population. Thus, the argument that the violence rate among discharged patients was close to that in the disadvantaged communities from which they came is not so.

There is a debate over which Axis I diagnoses are most closely associated with violence. In a study of a Danish birth cohort, hospitalized patients diagnosed with schizophrenia had higher arrest rates for violent crime than did nonpatients. A recent paper on the neurobiology of aggression in schizophrenia suggests a possible role of the catechol-*O*-methyltransferase gene that is involved in the metabolism of dopamine. Others have reported that patients with affective disorders have even higher rates of violence after hospital discharge than those diagnosed with schizophrenia. Some have found the onset of antisocial behaviors to

coincide with the onset of affective disorders in many patients. Many so-called antisocial youths have unrecognized bipolar disorders. Their wild, antisocial behaviors are often dismissed as part of normal adolescent development. A glance at Hare's criteria for the diagnosis of psychopathy reveals a plethora of manic symptoms (e.g., glibness, superficial charm, grandiose sense of self-worth, and promiscuous sexual behavior). Thus, use of the PCL-R criteria can lead to serious misdiagnosis.

Other psychiatric disorders associated with episodic violence, although usually not considered in the differential diagnosis, include PTSD and dissociative identity disorder. Violent behaviors may occur in reaction to flashbacks in cases of PTSD and in cases of dissociative identity disorder. Most patients with dissociative identity disorder develop aggressive alternate personality states early in life. In these states, they experience the actual pain inflicted and develop stoic but enraged and aggressive coping styles or personality states to deal with the pain. Subsequent stressors or threats, real or perceived, may evoke a switch into these aggressive personality states. At such times, these individuals may lash out, physically or sexually, and even become homicidal. Patients with dissociative identity disorder are often partially or completely amnesic for their behaviors and may vigorously deny responsibility for violent acts. Their stories may change frequently during interrogations, depending on their personality states and on what each state remembers of an event. These kinds of patients tend to be dismissed as pathological liars. Because forensic psychiatrists and psychologists rarely entertain dissociative disorders as part of their differential diagnosis, dissociative individuals are routinely misdiagnosed as episodically explosive, simply antisocial, or as "pathological liars." Because they do hear voices in their heads, they are sometimes diagnosed with schizophrenia. Alcoholic offenders, who, because of brain damage, do not recall parts of their behaviors, fill in the missing spaces with varying stories. They, too, tend to be dismissed as liars.

Again, it is important to restate that most mentally ill individuals are *not* violent. However, they are more likely than are mentally healthy individuals to act aggressively when stressed, especially if they have endured early, severe, ongoing maltreatment.

Off-Color Crime? Violence by the Medical Profession

How does one pick a color for psychiatrists and psychologists who participate in torture, even death of prisoners? It is known that at the time of the Second World War, American and Canadian doctors worked with the government on methods of getting prisoners to reveal what they knew. The use of mind altering drugs, the effects of sensory deprivation, and other methods were not only suggested but also implemented by respected individuals in the medical profession. Psychiatrists and psychologists at other times have consulted with the military on how to train soldiers to kill.

(It does not come naturally to healthy people.)

The current wars in the Middle East seem to have further degraded the morals and standards of some in the psychiatric/psychological professions. The role of psychologists and physicians in the Central Intelligence Agency's torture program has been likened to the four main qualities of behavior of the Nazi doctors. The report of the 2014 US Senate Select Committee on Intelligence Report on the CIA's Detention and Interrogation Program revealed that medical professionals (1) experimented on detainees; (2) perverted medical procedures, turning them into rape and torture (e.g., giving their detainees enemas of fluids and pureed food substances in the guise of feeding and hydrating them); (3) induced disease in their detainees; and (4) provided enough medical care to put their subjects into condition to be tortured again. Only two so-called professionals were named, James Mitchell and Bruce Jessen; however, they were assisted by anonymous MDs.

Although most of this chapter is devoted to the study of violence and criminality perpetrated by physically aggressive criminals and by greedy, flamboyant white-collar criminals and Mafiosi, next to nothing is known about what enables trained professionals to indulge in such brutality. Did they experience some sort of high? Some sort of sexual excitement while torturing? What were their incentives? According to government and newspaper reports, one incentive, money, may have been the draw. The sum of $81,000,000 has been bandied about. Was wealth the lure? Does that mean that for the right price anyone could be turned into a sadist or murderer?

The people who indulge in such atrocities without blinking an eye are the ones who should be studied, if only to learn how to keep everyone else from falling into that inhuman (or all too human) trap.

Differential Diagnosis

The differential diagnosis of violent behaviors is as broad as are the fields of psychiatry, psychology, and neurology. The diagnoses oppositional defiant disorder, conduct disorder, and antisocial personality disorder are hodgepodges that, as stated, cover a multitude of sins and a multitude of different kinds of signs, symptoms, and behaviors and combinations thereof. These categories if used at all should be thought of as interim designations on their way to proper diagnoses. Because violence toward others is the final common pathway of so many different kinds of neurophysiological vulnerabilities and environmental stressors, the clinician faced with the task of diagnosing and treating a dangerously violent person must work hard. Every aspect of functioning—behavioral, cognitive, emotional, social, psychiatric, neurologic, neuropsychological, and environmental—must be investigated thoroughly.

There is no single treatment or intervention for violence, because it has

a myriad of causes. What is more, because so much of violence is grotesque and abhorrent, the perpetrator often is dismissed as simply evil or bad. One need not be a clinician to make that diagnosis. Evil is not a scientific or medical term.

ETHICAL IMPLICATIONS OF WHAT IS KNOWN ABOUT VIOLENCE

The diagnostic evaluation of the violent and/or criminal person is a weighty responsibility. The diagnostic evaluation for purposes of determining guilt, innocence, or mitigating factors in murder cases is weightier still. It can mean the difference between life and death. Remember Landrigan.

Extreme violence, even in war, does not come naturally. Soldiers must be trained to kill.

Studies of delinquents and criminals have shown that they can often conceptualize moral principles but are unable to behave in accordance with them. Bioethicists and philosophers have maintained that true autonomy, that is, the ability to make reasoned choices, requires the capacity not only to identify and reflect on basic, instinctual desires and impulses, but also to control them, and thereby to do the right thing (i.e., act virtuously). Part and parcel of that precept, of course, is the capacity not to do the wrong thing. Unfortunately, not all brains have the same capacity to judge consequences and monitor and control behavior. Some people by virtue of birth trauma, emotional neglect, brain injury, genetic predispositions to mental illness, and exposure to abuse and violence—by virtue of what the bioethicist Rawls has called "the natural lottery"—are more vulnerable to emotional lability, distortions of reality, and loss of impulse control than are those who, by chance, have drawn a better hand.

This chapter asks the question posed by bioethicists Beauchamp and Childress: Do recurrently violent people, at the time of their violent acts, "have the capacity to rationally accept, identify with, or repudiate a lower order of desire or preference in a manner that [was] independent of the manipulation of desire?" Virtue, as conceptualized by Aristotle in *Nicomachean Ethics*, may spring from an innate capacity to do the right thing. However, according to Aristotle, in order for a person to make use of this potential for moral action, this capacity must be cultivated actively. It must be exercised with the same perseverance, intensity, and training necessary to perform well athletically. Aristotle recognized that virtue also required mentors, models for controlled behavior. The ethical question posed in this chapter is whether defendants at the time of their offenses could have acted virtuously, or whether the combination of psychopathology, frontal lobe dysfunction, and the ongoing exposure to violent models of behavior prevented them from thinking once, much less twice, about what they were doing? According to one of the authors' subjects, "Ain't no thinkin' anymore . . . it's a reaction."

These data call for a re-examination of the outdated 19th century

definition of insanity (i.e., ". . . if the party accused was labouring (sic) under such a defect of reason, from a disease of the mind, as not to know the nature and quality of the act . . . or, if he did know it, that he did not know . . . (it) was wrong"). Queen v M'Naughton (1843). This rule limits the defense to cognitive insanity, although it is known that cognition plays a limited role in violence. In this author's experience, some of her colleagues have taught defendants how to parrot the M'Naughton answers and thus have enabled them to go to trial and on to execution. Numerous well-trained psychiatrists have declared chronic schizophrenic and unalterably brain damaged people to have been "sane" at the time of their offenses. Did their psychoses temporarily remit? Did their damaged frontal lobes function briefly during their frenzied act? Given the biopsychosocial impairments that are now known to underlie many criminal, violent acts, has the time not come for the neuropsychiatric professions to assist the judiciary in redefining the criteria for mitigation and insanity?

References

Alvaro-Gonzalez LC. Neuroethics: moral pathways in normal brain [Review] [Spanish]. *Rev Neurol*. 2014;58(5):225–233.

Arizona v Landrigan 176 Ariz 1, 8, 859 p2d 111 (1993) Arizona Supreme Court.

Arnedo J, Savrakie DM, del Val C. Uncovering the hidden risk architecture of the schizophrenias: confirmation in three independent genome wide association studies. *Am J Psychiatry*. 2015;172(2):139–153.

Baez S, Kanske P, Matallana D, et al. Integration of intention and outcome for moral ludgment in frontotemporal dementia: brainstructural signatures. *Neurodegener Dis*. 2016;16(3–4):206–217.

Bondu R, Beier S. Two of a kind? Differences and similarities of attacks in schools and in institutes of higher education. *J Interpers Violence*. 2015;30:253–271.

Brunner H, Nelen M, Breakfield XO, Ropers HH, van Oost BA. Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science*. 1993;262:578–580.

Callahan D. *The Cheating Culture: Why More Americans are Doing Wrong to Get Ahead*. Orlando, FL: Harcourt; 2004.

Courtwright D. *The Violent Land: Single Men and Social Disorder from the Frontier to the Inner City*. Cambridge, MA: Harvard University Press; 1996.

Cristofori I, Zhongshan W, Mandoski V, et al. Brain regions influencing implicit violent attitudes: a lesion-mapping study. *J Neuroscience*. 2016;36(9):2757–2768.

Cushman F. The neural bases of morality: Not just where but when. *Brain*. 2014;137(4):974–975.

Decety J, Porges EC. Imagining being the agent of actions that carry different moral consequences: an fMRI study. *Neuropsychologia*. 2011;49(11):2994–3001.

Eidelstein A. Dissociative identity disorder theory explaining serial murder and murderers (review article). *J Trauma Treat*. 2015;S4:019.

Faulkner AL, Lin D. Recent advances in understanding the role of the hypothalamic circuit

during aggression. *Front Syst Neurosci*. 2014;8:168.

Fumagalli M, Priori A. Functional and clinical neuroanatomy of morality. *Brain*. 2012;135:2006–2021.

Glenn AL. The other allele: exploring the long allele of the serotonin transporter gene as a potential risk factor for psychopathy: a review of the parallels in findings. *Neurosci Behav Rev*. 2011;35:612–620.

Hare RD. *Psychopathy: Theory and Research*. New York, NY: Wiley; 1970.

Hart H, Rubia K. Neuroimaging of child abuse: a critical review. *Front Hum Neurosci*. 2012;6:52.

Hexuan L, Yi L, Guang C. Gene by social environmental interaction for youth delinquency and violence. *Soc Forces*. 2015;93;(3):881–903.

Hortis CA. *The Mob and The City: The Hidden History of How the Mafia Captured New York*. New York: Prometheus Books; 2014.

Hutchings B, Mednick SA. *Registered Criminality in the Adoptive and Biological Parents of Registered Male Criminal Adoptees*. Mednick SA, Schulsinger F, Higgins J, eds. Amsterdam: Elsevier/North Holland; 1974.

Ingalhaliker M, Smith A, Parker D. Sex differences in structural connectivity of the human brain. *Proc Nat Acad Sciences*. 2014;111(2):823–828.

Lee H, Kim DW, Remedios R. Scalable control of mounting and attack by Esr1+ neurons in the ventromedial hypothalamus. *Nature*. 2014;509:627–632.

Lewis DO. From abuse to violence: psychophysiological consequences of maltreatment. *J Am Acad Child and Adol Psychiatry*. 1992;31(3):383–391. (Recipient of Reiger Award of the Am Acad Child and Adol Psychiatry.)

Lewis DO. *Guilty by Reason of Insanity*. New York: Ballantine Publishing Group; 1998.

Lewis DO, Shanok S, Pincus JA, Glaser G. Violent juvenile delinquents: psychiatric, neurological, psychological and abuse factors. *J Am Acad Child Adol Psychiatry*. 1979;18:307–319. Abstracted in *Digest of Neurology and Psychiatry*. 1979;(253) June-July. Reprinted: *Annual Progress in Child Psychiatry*; 1980.

Lewis DO, Yeager CA, Swica Y, Pincus JA, Lewis M. Objective documentation of child abuse and dissociation in 12 murderers with dissociative identity disorder. *Am J Psychiatry*. 1997;154:1703–1710. Abstracted in: *J Watch for Psychiatry*. 1998;4(2). The Harvard Mental Health Newsletter; 1998.

Mark V, Ervin F. *Violence and the Brain*. New York: Harper and Row; 1970.

Miczek KA, Meyer-Lindenberg A, eds. *Neuroscience of Aggression*. Heidelberg and New York: Springer; 2014.

Moskowitz A. Dissociation and violence: a review of the literature. *Trauma Violence Abuse*. 2004;5:21–46.

Piton A, Poquet H, Redin C. 20 ans après: a second mutation in MAOA identified by targeted high-throughout sequencing in a family with altered behavior and cognition. *Eur J Hum Genet*. 2014;22:776–783.

Raine A. *The Anatomy of Violence: The Biological Roots of Crime*. New York: Pantheon Books; 2013.

Rawls JA. *Theory of Justice*. Revised Edition. Cambridge: Harvard University Press; 1999.

Schiltz K, Witzel JF, Bausch-Holterhoff J, Bogerts B. High prevalence of brain pathology in

violent prisoners: a qualitative CT and MRI scan study. *Eur Arch Psychiatry Clin Neurosci*. 2013;263:607–616.

Stuber GD, Mason AO. Integrating optogenetic and pharmacological approaches to study neural circuit function: current applications and future directions. *Pharmacol Rev*. 2013;65:156–170.

Teicher MH. Wounds that won't heal: the neurobiology of child abuse. *Cerebrum*. 2000;2:4, Fall.

Wilson ML, Boesch C, Froth B, et al. Lethal aggression in Pan is better explained by adaptive strategies than human impact. *Nature*. 2014;513:414–417.

Yoder KJ, Decety J. The good the bad and the just: justice sensitivity predicts neural response during moral evaluation of actions performed by others. *J Neuroscience*. 2014;34(12):4161–4166.

▲ 28.3 Borderline Intellectual Functioning and Academic or Educational Problem

FRANK JOHN NINIVAGGI, M.D.

This section introduces one of the most clinically oriented, if not relevant, portions of this textbook—how to recognize borderline intellectual and adaptive functioning. The psychiatric focus of this section rests on variable patient presentations—clinical expressions—that may indicate risk, and possible disability burden, due to both mental abilities and adaptive performance skills being *well below average*. When the threshold below typical and into borderline intellectual and adaptive functioning is crossed, significant slowness and risk are present. Borderline functioning, however, is neither currently viewed as a formal disability nor as representing clear-cut deficits in intellectual and adaptive domains; thus "borderline intellectual functioning" (BIF) is not a formal DSM-5 disorder. Yet, *inability to function in a typically average range may indicate probability of risk resulting in potential harm and loss in the future that can be prevented or ameliorated in the present, if identified and managed*. Mindful awareness of BIF enhances higher reliability of both remedial anticipation and building resilience. Such prevention enhances cognitive reserve and disengages the trajectory for opportunities for cumulative impairments to coalesce along the developmental line. Screening and assessment for risk factors, especially early adversity, predispositions to vulnerability, and to stressful negative life outcomes strongly indicative of developing disorders, constitutes good clinical care. It reduces their often-underestimated residual burden. These considerations about risk, risk reduction, and intelligent performance (not identical to IQ scores) will be discussed, since considerations of risk and risk reduction have a protective effect on many levels and maintain patient-centered values at the heart of quality medicine.

Protection averts loss and builds a fence around resiliency. Yet, the following question remains: While duly respecting the currently limited state of evidence-based literature regarding BIF, does identifying clinically significant risk and addressing it merit inclusion of BIF, currently a diagnostic gray zone, in official DSM-5 diagnoses? Such inclusion, in fact, would acknowledge the evolving knowledge base of neuroscience and psychiatry, as well as the limitations of diagnostic criteria.

It is suggested here that vanguard trends in primary prevention—eliminating or reducing risk—merit broadening risk assessment, lowering diagnostic thresholds, and identifying cases having incipient clinical significance. The broadened scope of this inclusion, therefore, may

accommodate the natural diversity inherent in the heterogeneity of populations and not artificially exclude groupings of heterogeneous syndromes with inferred different pathophysiological mechanisms. Recognizing that designated cutoffs may be better understood as moving targets than as tipping points may be more realistic, and thus reflect the enhanced developmental focus of DSM-5. The magnitude of the net benefit of this remains to be seen. It may include authorizing coverage and making possible needed clinical services. Identifying such needs and providing services is the essence of a prevention mode that may thwart, if not obviate, additional complications. Such harm reduction therefore is a clinically salient goal of very good medicine. Greater strides toward precision medicine remain an ultimate vision.

The condition now termed BIF should be understood to denote borderline intellectual *and adaptive* functioning, and this phrase should be its proper designation. This reflects a more realistic developmental, clinical, and functional understanding, as well as diagnostic notation of the whole child. Furthermore, it is consistent with conceptions and definitions of intellectual disabilities (IDs) held by the American Association on Intellectual and Developmental Disabilities (AAIDD) and put forth in the World Health Organization's ICD-10 and its International Classification of Functioning, Disability, and Health.

The international community has taken a striking lead in research and has proposed bold initiatives about pursuing serious investigations concerning the significance of BIF. A recent summary paper from Spain in 2013 by a 19-member panel documents the Catalan association's Nabiu (ACNabiu) CONFIL project begun in 2007. This comprehensive study funded by the Department of Health of the Catalan government draws from CONFIL 1, 2, 3, and 2010 conclusions. Using evidenced-informed clinical data, it advances BIF as a priority condition demanding an international consensus that includes setting forth the following: construct, operational criteria, screening, coding, incidence, prevalence, effective interventions, and an integrative model of care centered on the individual that includes family and extended social, educational, and occupational environmental considerations. Herein is found a wealth of potential research trajectories. This movement is spreading in Europe and internationally.

INTELLECTUAL FUNCTIONING

Whether intelligence is best regarded as a single general capacity (g) as proposed by Charles Spearman in 1904 and1927 or as a range of separate, modular capacities as first proposed by Thurstone in 1938 remains an open question. No consensus prevails. Contemporary theorists view intelligent cognition as reflecting an array of abilities, strengths, and weaknesses across multiple information processing and adaptive domains. The Cattell-Horn-Carroll (CHC) theory of intelligence, one of the dominant models in

modern psychology, reflects this view.

The very nature of intelligence and its multiple definitions lingers in an unsettlingly, intriguing flux and as the ongoing subject of intensive research. As far back as 1916, American psychologist, Lewis M. Terman, recognized the spectrum of intellectual capacities and formalized the notion of borderline intelligence to indicate those who not only functioned well below average but were also not intellectually disabled. The emergence of rapidly evolving disciplines such as cognitive science and its subspecialties devoted to an exploration of the meaning and pragmatics of all aspects of human cognition (e.g., conceptual, emotional, social, practical, and motivational) attests to the central importance of these undeniable epistemological realities. The role of powerful nonconscious emotions such as envy in eliciting comparisons of superiority and inferiority that trigger cognitive discrepancies, distinctions, and emotional dysregulation attests to this.

Intelligence—acquiring and applying knowledge—is broader than cognition—knowing—considered in isolation or merely by IQ scores. Intelligence and IQ scores, though related, are different conceptions. Intelligence first denotes the capacity for awareness and discrimination of relevant and novel environmental and subjective stimuli. This information is then used in integrated fashion in several ways:

1. To eductively make sense out of complexity through analysis, synthesis, and reproducing meaning, especially through perceptual reasoning.
2. To maintain mental and physiological infrastructure.
3. To promote regulatory expansions in adaptive survival that includes flexible and effective performance interactively with the environment.

Language and effective communication facilitate executive functions—focused attention, selection, planning, organization, implementation, and refinement of both thinking and behavioral responses. Words become used as “signifiers” that represent something, and these must be combined in the brain with the “signified” to form a meaning-imbued “sign.”

This unique human symbolic process correlates with the level of an individual’s capacity for abstract thought, which facilitates executive functioning. Emotions as well as specific and nonspecific motivational urges give dynamic meaning to the way these functions are integrated, experienced, and refined. An individual’s intelligence, moreover, does not exist in a vacuum; it is a function of interpersonal engagements on all levels throughout the lifecycle. These contribute to eliciting strengths and accommodating to challenges. Respect for the life of the mind and for the totality of the functioning adaptive human being underlies these considerations.

Intelligence or intellectual functioning is a complex construct (a measurable representation of a hypothesized abstraction) that does not

necessarily manifest itself in unitary fashion. It includes the following mental abilities: understanding relationships; adeptly processing information in an efficient manner; spontaneously applying problem-solving strategies; acting purposefully in novel, ambiguous, and complex situations; and learning from experience to further refine adaptive and problem-solving skills. Intelligence is cognitive competence, the mental ability to function and perform successfully. *Intelligence is both intelligent cognition and intelligent adaptive performance*. A range of latent inner resources and potential capacities can be developed by learning, experience, and subsequently expressed as real-time capabilities to perform a demanded task. Pragmatic and efficient skill performance denotes adaptive functioning.

From a scientific and academic perspective, measurements of intelligence reflecting an individual's position or standing relative to others traditionally have resided on two quantitative indices:

1. Measurement by psychometric ("measuring the mind") instruments yielding profile arrays with numeric ranges of component scores of cognitional abilities including a composite score—full-scale IQ.
2. A statement about the effectiveness of adaptive behavior as expressed by arrays of varied performance skills.

For both criteria, cutoff scores are approximately two standard deviations below the mean of the respective assessment instrument. These scores vary by test, subgroup, and age group while considering the standard error of measurement for the specific instrument used and its strengths and limitations. Current testing uses complex neuropsychological examinations that yield arrays of profiles that eschew previously held reductionist assessments and single scores in isolation. Such results give full weight to the nuances of multifaceted cognitive and adaptive processes. Intelligence is neither a static given nor a single number, but rather a dynamic and relatively changing confluence of phenomena affected by chronological, maturational, developmental, neurological, psychological, social, and contextual factors. In BIF, real-world reasoning and judgment are well below average. These atypical features are apparent in discernibly poor academic performance.

Taking the role of intelligence into serious consideration can be vital in assessing health and well-being. This derives, in part, from research demonstrating that intelligence plays a crucial role, along with socioeconomic influences, not only in predicting academic success but also in promoting a variety of positive life outcomes. Along with the central importance of denoting intelligence as mental ability, the inclusion of emotional intelligence (*emotional knowing*) and motivational factors in their real-life adaptational mixture is essential in all considerations of intellectual functioning. The expression of intelligence and its limitations is

multidimensional. It emerges from the dynamic and reciprocal engagement among intellectual abilities, adaptive behavior, health status, social participation, cultural context, individualized opportunities, and support availability. These considerations must temper, if not humble, any endeavors at reductionist measurements of this basic human functioning.

While the main thrust of this section is on the cognitive and performance role central to intact intellectual functioning, emotional conflict certainly affects thinking and activities of daily life. From a psychodynamic perspective, when psychic conflict escalates and mechanisms of defense are elicited, the use of *intellectualizations* may be employed to block and constrain conflictual feelings thus temporarily subduing anxiety. Avoiding and denying unruly emotional turmoil by using lofty, often specious, cognitive explanations normatively arise in adolescence as a way of managing newly emerging psychosexual impulses. If such strategies bring relief, they are carried forward and may become embedded as enduring character traits. Those with BIF, however, rarely become proponents of such abstruse, pedantic legalisms or present themselves as voluble, creamy intellectuals whose excessive verbosity mesmerizes the listener. Intellectualizations are cognitive obfuscations rarely if ever used by those with BIF. These contrasts are raised since they may serve to sharpen attention to the typical presentations of BIF, more literal and concrete, as they tend to be.

The increasingly recognized protective effect that intelligence affords, furthermore, merits its appreciation as a powerful therapeutic focus. Protective factors increase chances of resilience, healthy coping, effective action-oriented behaviors, and positive life outcomes. Such knowledge has clinical utility by providing useful information needed to convey material about the patient that is relevant regarding clinical course and comprehensive treatment planning. *Attention to IQ scores and adaptive strengths and weaknesses has provided helpers with a means of identifying persons who have developmental and other special needs and offering services to optimize social inclusion, equity, and personal empowerment. Moreover, these models influence research, policy, funding, and treatment reimbursement.* These considerations also comprise a vision aimed toward raising an individual's biopsychosocial quality of life, a life with meaning and dignity. Cognitive and emotional capabilities pragmatically expressed as satisfying life skills in the real world are thereby enhanced.

BORDERLINE INTELLECTUAL AND ADAPTIVE FUNCTIONING AS A CLINICAL ENTITY OF SIGNIFICANCE

DSM-5, as have its previous renditions, recognizes and codes BIF as a condition or focus of attention, and not as a disorder. Empirical clinical utility is thus acknowledged, especially by definition, as it may have an

impact on clinical care and treatment progress. As far back as DSM-III (1980), the precise distinction between normality and psychopathology had been seen as a gray area whose inexactness remains subject to continuous scrutiny and reexamination. The threshold for "caseness" and the "cut off area" concepts used to denote a diagnostic label of "mental disorder" are not yet deemed sufficient for BIF. While the real-world population abounds with persons having no significant difficulties, having meaningful problems due to borderline intellectual and adaptive functioning typically exists in these well below average individuals. Borderline intellectual and adaptive functioning, moreover, may exacerbate comorbid mental disorders. Yet, the psychological, educational, ID, and psychiatric literature meagerly yet valiantly strives to convincingly document this challenge. DSM-5 nonetheless considers BIF worthy of inclusion since conditions such as these draw attention to a wide scope of problems encountered in clinical practice. BIF is coded both for its own clinical importance and as it affects the diagnosis, course, prognosis, or treatment of the patient's co-occurring mental disorder(s). It is useful information that may affect patient care.

A consistent operational definition of BIF that covers all or most situations and considers, for example, symptom presentation, structural pathology, pathophysiology, and etiology, along with valid and reliable studies has been absent. This correlates directly with the absence of clear-cut measurement criteria. To what extent BIF resembles a disorder, a disability, a deficit, or aspects of each, and shows features such as distress, dyscontrol, dysfunction, malfunction, syndromal patterning, pervasive inflexibility, gross irrationality, severity level, disadvantage, social or occupational impairment, risk vulnerability, developmental atypicality (not emerging as expected for age), heterogeneity, and meaningful statistical limitations and deviations are questioned, debated, and controversial. These matters remain unsettled. For example, could disability be construed not as resulting from deficit in isolation but from neurocircuitry in conflict and competing with one another to create a chaos that results in less disorganized or "well below average" or "borderline" outcomes?

BIF is, in fact, a boundary or transitional phenomenon—at least by previously accepted criteria (DSM-IV-TR), and now by rather broad, if not somewhat ambiguous, parsimonious guidelines ("careful assessment of intellectual and adaptive functions and their discrepancies," DSM-5, p. 727). Since BIF had been viewed as the transition from the ID spectrum to the average intellectual spectrum, and since psychometric and adaptive testing shows some degree of excursion from statistical norms in one or more important areas of functioning, a "disability model" is suggested. One could also argue that the individual must have some internal "disorder" that is elicited or brought forward when interacting with an environment (animate and inanimate) not quite capable of meeting atypical needs and thus causing distress. A last consideration that shall not be elaborated here

is that regarding the term "mental." Mental disorders assume the implicit split in an individual between brain and mind. This dualism is not merely traceable to 17th century Western Cartesian dualism but is recorded as far back as the extant Sanskrit literature of India and some Eastern philosophical systems. Contemporary Western science now understands individuals from more unified, less anachronistic perspectives. "Mental" disorders, therefore, are now regarded as brain-mind-behavior disorders —*biomental*—and this biomental quality also describes the nature of BIF.

BIF is a developmental limitation, not a psychiatric disorder. Not all persons with the condition of BIF or an IQ score of about 71 to 84 are impaired or dysfunctional. Lower-than-average (average IQ score range: 85 to 115; mean = 100) cognitive (analytical-conceptual-critical reasoning, verbal, and spatial) ability must be present concurrently with significant limitations in adaptive behaviors and functionally restricted practical performance skills across several domains of everyday living. Clinical experience is consistent with BIF occurring in the developmental period before age eighteen. Integrating assessment of IQ scores with an individual's everyday functioning supports ecological validity in interpretations.

Adaptive behavior domains include the following collection of conceptual, social, and practical dimensions of performance behaviors:

- Communication,
- Self-care,
- Social and interpersonal skills,
- Home living,
- Use of community resources,
- Self-direction,
- Health and safety,
- Functional academics,
- Leisure, and
- Work.

Evaluation and treatment planning warrant serious consideration of adaptive performance criteria with particular attention to an individual's physical assets, strengths, preferences, and productive habits (e.g., exercise, sports, and other physical activities). Adaptive behaviors include a range of both self-care and instrumental activities of daily living, notably skills needed to be reasonably successful in the community. Uneven performance in adaptive skills, as also seen in cognitional abilities, is common. Many distinguished contemporary psychologists whose work on intelligence testing is held in high esteem repeatedly emphasize the difficulties in adequately describing ability levels of individuals whose IQ scores reside in the borderline category.

BIF, therefore, is a heterogeneous condition that, when dysfunctional, denotes the co-occurrence of clinically, significantly limited intellectual and adaptive functioning. Together, these deficits result in uneven and often marginal scholastic, practical, and social competence. In such situations, BIF acts as a functional neurodevelopmental disability placing the individual at risk for failure and subsequent lowered self-esteem, dysphoria, and distress. As these children grow, the gap between their capabilities and demands from school, the workforce, and society grow larger and more difficult to fulfill successfully; hence failure and blame ensue. Such academically fragile persons have been referred to as "slow learners" and "hard-to-teach" children, especially since they typically have difficulty understanding and learning average amounts of material. Dysfunctional BIF denotes a general life skill's coping disadvantage and a heightened risk for maladaptive outcomes. Yet, it should not be regarded as a static condition. Over-reliance on IQ scores alone incorrectly de--emphasizes the critical importance of viewing an individual's functioning as the complex interaction of strengths, limitations, opportunities, and appropriate supportive interventions. The entire conception of supports—a contemporary and novel construct—includes resources, strategies, and accommodations that enable people to live successfully in regular community settings.

The essence of borderline intellectual and adaptive functioning, therefore, descriptively indicates *ranges of abilities that are well below average*. Clinically, those with BIF are very slow to understand and to problem solve, across areas. Executive functioning reflected in planning, anticipating, performing, and self-correcting are well below average; this may manifest in the face of apparently adequate language skills. BIF in effect can become a meaningful developmental limitation in learning and real-life adaptation. It, therefore, can contribute to marginal abilities or inability to succeed in major life functions such as self-care, academic achievement, and social-occupational tasks. Limited ability to perform may be due to unavailable inner resources or inability to access and make efficient use of what is already psychologically present but undeveloped. Sensitive environments can respond accordingly and enhance unsteady performance.

Last, as discussed in this section, the condition of academic or educational problem is coded, usually in addition to other diagnoses, when scholastic performance is severely impaired and becomes an outstanding feature of a patient's overall clinical presentation.

CULTURAL SENSITIVITY: A FRESH PERSPECTIVE AND AN ORIGINAL VOICE

The approach advocated here to recognizing, screening, identifying, and therapeutically working with persons with borderline intellectual and adaptive functioning highlights direct clinical focus on the individual. This

acuity is an astute individualized line of engagement. It rests, in part, on the premise that the array and range of factors (conceptual, social, and practical performance skills, IQ scores, and details of an individual's culture) are multidimensional and always varied. Effective clinical work requires a center of gravity in clinical details with periodic reflection on theoretical generalizations—sound clinical judgment. Acute clinical awareness and a high index of suspicion are needed to detect borderline functioning, which may indicate clusters of additional risk factors predisposing to limited and even impaired performance.

Recognition of cultural sensitivity in its broadest dimensions (language; styles of communication, movement, gesture, behavior; and the community environment typical of the individual's peers and cultural milieu) has been axiomatic of good clinical care. DSM-5 has updated its "Outline for Cultural Formulation Interviews" with field tests for diagnostic usefulness in many areas to obviate difficulties in communication that may bias diagnosis and treatment. In this spirit, clinicians who encounter those with BIF and adaptive performance need to develop a higher level of sophisticated sensitivity and amplified empathy. This more intensive recognition will support a cultural formulation producing the following:

- Increased patient engagement,
- A contoured comprehensive treatment plan,
- Higher powered effectiveness, and
- Improvements in the patient's quality of life.

Toward this vision, an acute, forwardly aware recognition and understanding of these cultural dimensions in the clinician and how they have an impact on those served including their families, schools, and allied agencies is essential. This partially self-reflective technique helps bring the expression of BIF into sharper focus because BIF is typically "obscure." This instrumental perspective is the clinician's adding such a self-observing awareness as a primary tool to the repertoire of all assessments. Evaluators and treaters undoubtedly experience some differences in perception, conception, and emotion—not themselves having a personal experience of BIF. Sharper self-awareness provides the opportunity to recognize the transitional areas in functioning and adaptive skills that approach theoretical ideals or "norms." This helps clinicians capture subtle impairments that may be severe enough to produce functional yet under-recognized disabilities. Effective clinical understanding may become more pragmatic in that it traverses this transitional boundary between divides of helper-helpee and typical versus slower or uneven mental abilities. The clinician's more intimate personal self-reflection and recognition of the individual with BIF and his or her distinctive profile of mental abilities, adaptive behaviors, and personality traits thus may be heard by the patient as an "original voice" that is fresh and empowering. The patient

experiences the clinician not as an outside stranger, but more as an inside ally, mentor, guide, and counselor. Such a motivated engagement magnifies the experience of care and may mitigate iatrogenic biases, enhance adherence, reinforce support, and bolster greater chances of success.

BORDERLINE INTELLECTUAL FUNCTIONING

Definition

The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) describes BIF in Section II "Other Conditions that may be a Focus of Clinical Attention" as follows:

> This category can be used when an individual's borderline intellectual functioning is the focus of clinical attention or has an impact on the individual's treatment or progress. Differentiating borderline intellectual functioning and mild intellectual disability (intellectual developmental disorder) requires careful assessment of intellectual and adaptive functions and their discrepancies, particularly in the presence of co-occurring mental disorders that may affect patient compliance with standardized testing procedures (e.g., schizophrenia or ADHD with severe impulsivity).

DSM-5 (ICD-10) coding: R 41.83 has significantly revised its definition of BIF from that given in DSM-IV-TR in 2000. DSM-5 begins with acknowledging that BIF may be a focus of clinical attention or have an impact on treatment or progress. No numerical IQ scores (i.e., IQ 71 to 84) are now explicitly given as definitional criteria, as they had in DSM-IV-TR. Metrics for identification are suggested by indicating careful assessment of intellectual and adaptive functions and their discrepancies. Adaptive functions are now mentioned in the DSM-5 entry. DSM-IV-TR had contrasted BIF with mental retardation (MR)/ID. DSM-5 now contrasts BIF with mild ID only. Co-occurring mental disorders now include ADHD with severe impulsivity and mentions schizophrenia that may affect patient compliance with standardized testing procedures. DSM-5 does not give highly detailed coding instructions.

Table 28.3–1.
Clinical Features Associated with Borderline Intellectual Functioning

1. Mild, impairing global (not specific) learning problems
2. Well below average, limited ability
3. Uneven academic achievement and adaptive performance
4. Concrete mindedness
5. Well below average comprehension of functional relationships
6. Slowness in understanding, broad comprehension, and response
7. Poor attention, concentration, and organizational skills
8. Impaired ability to generalize and transfer learning
9. Low frustration tolerance
10. Intermittent dysphoria from perceived inadequacy
11. Low self-esteem
12. Variable social and adaptive behavior impairment

13. Task performance overcompensation (physical tasks, sports)
14. Compensatory verbal fluency
15. Naiveté
16. Need for more-than-average direction

Since DSM-5 indicates differentiating BIF from mild ID, a brief discussion of this follows. DSM-5 has shifted emphasis in its perspective on IDs. Levels of severity are determined by the assessment of adaptive functions rather than by IQ scores. Of great significance in the recent ID literature is the serious consideration that features characteristic of mild ID, in more muted and less intense forms, may at times characterize BIF. DSM-5's section on ID criteria, Table 28.3–1, discusses severity levels in the three adaptive domains: conceptual, social, and practical. The reader is referred to the DSM-5 discussion, particularly, Severity Level Mild, p. 34, for an elaboration of individual support needs, intensity, and systems of support. Detailed elaborations of these comprehensive systems also can be found in AAIDD texts. AAIDD has made and continues to make contributions of inestimable importance both theoretical and practical. The enormous significance of supports is reflected in the AAIDD User Guide statement: "The implication is that whereas the life functioning of the individual with ID can improve with personalized and sustained supports, a valid diagnosis will not change." (p. 2).

History and Comparative Nosology

The history and development of definitions of BIF directly reflect evolving conceptualizations of mental retardation (MR), now termed ID. Efforts toward reducing stigmatization and unfair segregation resulted in a federal statute in the United States (Public law 111–256, Rosa's Law, 2010) replacing the term "mental retardation" with "intellectual disability." Whereas BIF is a borderline cognitive and adaptive limitation, ID represents an array of marked, severe, serious, and substantial developmental disabilities that include both intellectual and adaptive functioning deficits in conceptual, social, and practical domains. The DSM-5 (2013) definition correlates with that of the AAIDD 2010 definition except that the DSM-5 categorical definition includes the terms "disorder" and "deficit." AAIDD instead uses the phrase "significant limitations" to denote and describe features of intellectual "disability." Of note is that ICD-11, in development, uses the phrase "intellectual developmental disorder" to indicate ID as a disorder that involves impaired brain functioning early in life. Since about 1973 in addition to an IQ score of between 70 and 75 up to 85 and an onset before 18 years of age, classification schemes have placed significantly greater emphasis on functional adaptive skills, the elimination of the notion of borderline mental retardation, and the inclusion of the construct of BIF.

The following is a brief historical review of the coding and terminological vicissitudes of BIF in the DSM. DSM-I (1952–1967), which correlated with ICD-6 of 1948, coding in Section V reads: "Code no. 60.0, mental deficiency (familial or hereditary), mild, standard nomenclature: 000-x901, international list: 325.3." DSM-II (1968), which correlated with ICD-8, 1965, used the code and phrase "310. borderline mental retardation." The phrase "borderline mental retardation" was eliminated in 1973 (official reclassification by the American Association on Mental Deficiency, now AAIDD). It was changed to "borderline intellectual functioning" (BIF). DSM-III (1980), which followed ICD-9, 1975–1978 [ICD-9 clinical modification to increase resolution of diagnoses and procedures, an adaption created by the U.S. National Center for Health Statistics, became official in January 1979 as "ICD-9-CM"]. DSM-III first introduced the five-axis system and identified BIF on Axis I coded as "v 62.89." DSM-III-R (1987), still following ICD-9-CM, now shifted BIF from Axis I to Axis II and coded it as v40.00. DSM-IV (1994), working closely with ICD-10, 1992, maintained BIF on Axis II, and coded it as v62.89. DSM-IV-TR (2000) maintained previous definitions, classifications, and coding. DSM-5 (2013), now intimately aligned with ICD-10, revised its description of BIF, eliminated all axes systems, and codes BIF as "R 41.83."

Rapidly changing paradigms within the ID field have adopted an interaction-based ecological perspective with an evolving vocabulary. For example, with the AAIDD's 2010 publication of its 11th edition, *Intellectual Disability*, the nature and amount of support (frequency, daily support time, and type of support) that a person needs to participate fully in life is defined. As of October 1, 2014, DSM-5 codes ID based on both intellectual functioning deficits and adaptive competencies, but specifiers will be based on levels of severity defined by adaptive functioning parameters. Although DSM-5 criteria remain broad, identifying BIF requires "careful assessment of intellectual and adaptive functions and their discrepancies."

Etiology and Epidemiology

As with other conditions that reflect intellectual and adaptive limitations, genetic, hereditary, congenital, health, environmental, interpersonal, sociocultural, and educational opportunity have been implicated as etiological factors. For BIF, however, no factor in isolation has been shown to be definitively causative. Prenatal, perinatal, and postnatal medical conditions must be identified, since they may provide clearer explanatory and remedial information. With reference to a normal distribution or bell curve of psychometric intelligence on the general population, IQ scores in the borderline range (70 to 85) occur from between one and two standard deviations below the mean or average of 100, that is, approximately 13.6 percent of the time. The clinical condition of BIF is strongly suggested only when significant limitations or impairments in adaptive functioning

accompany this finding.

Identification and Clinical Features

Contemporary views stress *well below average* measured intelligence, limited adaptive functioning, and support requirements as the broadest dimensions within which dysfunction and needs arise. *Limitations and well below average performance may not be global and pervasive; areas of strength may be interspersed with marginal functioning*. Identification integrates clinical, academic, and common sense factors. There has been a shift from reliance on composite or full-scale IQ scores to an analysis of multiple subtests of cognitive functions together with adaptive performance. Specific cognitive abilities predict academic achievement and can be one part of evidence-based data for prescriptive, comprehensive interventions, and individualized remediation. Because information regarding academic performance and testing, including psychometric testing, is typically available for persons younger than 18 years of age, BIF may be recognized more easily in the school-age population. Initially, assessing adaptive functioning is typically done through clinical screening and thorough history taking. If needed, formal specialized testing may be performed. Intelligence testing and the conceptual, social, and practical life skills aspects of adaptive testing should be performed in a nonbiased manner with due attention to cultural contexts. Fairness in adequately identifying and not misconstruing the presence of BIF remains important in decision making. Motivational factors, effort, and fatigue not only affect test taking but also affect real-time performance in life skills. Effects of psychotropic medications on performance must also be considered. Information from testing is correlated with all aspects of the individual's life functioning for conclusions to be rendered reasonably and to reflect diligent clinical judgment. *For example, those with BIF are recognized implicitly as requiring more coaching early on in grade school. When this is successful, verbal skills are raised; hence, they may test at near average verbal levels both empirically and on test scores. In such cases, verbal comprehension may eclipse well below average capacities in perceptual reasoning, working memory, and processing speed deficits. A primary consideration of BIF thus may be missed.* Practical performance may be masked by a veneer of language. Table 28.3–1 summarizes some important clinical features associated with coding the condition of BIF. It must be emphasized that familiarity with the basics of child and lifecycle development, easily gleaned from readily available texts on child development, offers inestimable pragmatic, clinical assistance in orienting the evaluator concerning presentations reflective of typical versus *well below average performance*.

Clinical assessments and the psychiatric interview are based on first hand observation. Seeing not only what a person can do, but how he or she

does it is crucial, for adaptive performance. Direct and detailed observation —*the patient's clinical expression*—is indispensable to identification. In general, the attentive clinician experiences that the individual's cognitional process—receptive and expressive—is slow moving and lumbering. These variable expressions may demonstrate some of the following: mild but impairing global (not specific or isolated) learning problems; learning how to learn—the brain's "search engine"—appears limited; understanding is slow moving; slow working memory; difficulty grasping functional relationships; once learned, materially performing an activity or task is intact, and may even be highly skilled; uneven academic and other skills performance; tendency to concrete-mindedness; perceptual reasoning may be impaired; some slowness in response time; difficulty with simultaneous management of multiple tasks; notable difficulties in understanding and learning, which may be displayed as task avoidance, poor attention, concentration, and organizational skills; difficulties integrating new knowledge with previously learned information and applying it to novel problem solving; impaired capacity to generalize learned information and skills; low frustration tolerance; intermittent irritability, dysphoria, and distress, especially from poor achievement and perceived personal inadequacies; low self-esteem and low motivation; social awkwardness; mild-to-moderate social skills problems although social interaction is sought (not apathetic); neither hypokinetic nor hyperkinetic motor activity; poor common sense; naiveté; potential for being bullied and for victimization; and a need for more-than-average direction and supervision in general, given chronological and developmental status. Procedural memory and habitual task performance from past practice and repetition are typically intact. Uneven performance within each and between intellectual and adaptive domains may occur for a number of reasons, one of which is the individual's own compensatory processes. Verbal skills testing as average or better may compensate for well below average perceptual reasoning and problem-solving abilities. Clinically, this may mask the real presence of BIF.

When a number of the aforementioned clinical features are moderately pervasive, dysfunction is inevitable. All these factors then contribute to a less-than-average rate of learning and understanding, frustration in an academic environment that expects average or greater-than-average performance and compliance, academic and social failure, some difficulties with activities of daily living, and a greater chance of emotional distress and behavioral dysfunction.

Other sources that support this diagnosis come from standardized intelligence tests (e.g., Wechsler Intelligence Scale for Children, Stanford–Binet test, and Kaufman Assessment Battery for Children), adaptive behavior scales, and school psychological and academic testing and achievement reports. When these clinical features, integrated psychometric

profiles of multiple measures of cognitive functions and capacity, and adaptive and support needs are associated with an individual's biopsychosocial and academic dysfunction, BIF may be coded. If it is clinically deemed the principal or a central cause of the individual's problems or visit, the terms "principal" or "reason for visit" may be indicated in parentheses. If not enough information is available although a strong presumption is present, such uncertainty can be indicated by recording "(provisional)" after the code.

Differential Diagnosis

One must first rule out substantially significant forms of ID, which can easily be detected by standardized intelligence tests, adaptive skills profiles, history of onset in the developmental period, and overall performance in a variety of contexts. Table 28.3–2 summarizes the differential diagnosis of the clinical condition of BIF. A pivotal diagnostic question with significant treatment implications is whether the typically diagnosed disorders that often mask or are mistaken for BIF, such as ADHD, attention-deficit disorder (ADD), and learning disorders (LDs), are in fact present or absent. In addition, if comorbidity exists, it is essential to designate whether BIF is primary and comorbid, or comorbid and secondary to the individual's presenting and fundamental borderline intellectual problems. Careful and diligent history from family and collateral sources such as school reports including psychological testing are indispensable.

BIF is typically under-reported. The most frequently diagnosed disorders that mask, overshadow, or may be mistaken for BIF are *ADHD, ADD*, and *LDs*. Common features of all these are focusing attention and sustaining concentration that prevent adequate learning from occurring. Although there may be diagnostic overlap, as well as definitional variations according to variable criteria used, ADHD, ADD, and LDs are often, though not necessarily, found in those with average or higher IQ scores (above 85).

In ADHD and ADD, impulsivity and behavioral inhibition impairments are primary along with executive functioning impairments. Many stress that difficulties reside in performing academic skills, not in learning them. In ADHD, impulsive behaviors typically pose significantly disruptive problems at home, school, and in other contexts. In ADD and ADHD academic functioning is notably affected by an impulsive need to get through material rapidly, which is a deficit intimately linked to poor behavioral inhibition. Impulsivity leading to immediate action also prevents sufficient encoding into working memory. In BIF, academic learning is primarily limited since overall cognitive processing is well below average and concrete oriented with an emphasis on physicality; and adaptive performance may match these limitations including limited communication, social participation, or occupational functioning. In ADHD, organizing and skillfully executing a task is a clear-cut-weakness; in

BIF, once an activity is learned, it is typically performed soundly; execution may be slow yet persistent and organized. Some studies indicate that on the WISC-IV Factor Index, Wechsler Perceptual Reasoning (Gv-Gf), those with ADHD attain average or better scores, unlike those with BIF who score lower on solving novel nonverbal problems and fluid reasoning tasks.

Table 28.3–2.
Differential Diagnosis of the Clinical Condition of Borderline Intellectual Functioning

1. Attention-deficit/hyperactivity disorder; attention-deficit disorder
2. Learning disorders
3. Autistic spectrum disorder
4. Nonverbal learning disabilities
5. Conduct disorder
6. Antisocial personality disorder
7. Mental retardation or intellectual disability

Specific LD (present in up to 15 percent of the United States population) may manifest, not globally, but in specific fashion (reading, writing, mathematics, or combinations of these) and in those school-age children whose learning and using academic skills are compromised by more modular, primarily cognitive, information processing irregularities. Samuel A. Kirk, an American psychologist, was the first to identify the construct, and coin the term "learning disabilities" in 1963. DSM- 5 implies IQ scores to be greater than about 70, thereby ruling out ID to better account for the disorder. These LDs may also occur in "gifted" individuals, that is, those with IQ scores roughly above 130. DSM-5 also terms the persistent difficulties defined by LD as "disorders" rather than "disabilities." They are localized to academic skills that are substantially and quantifiably below those expected for age; and they cause significant interference with academic or occupational performance, or with activities of daily living. Details in memory function especially encoding and retrieval, play a significant role. The entire construct of LDs has undergone radical revisions in the last several years, and its determination is dependent on multiple complex factors. A comprehensive process that includes developmental history, medical history, psychological and achievement testing, school reports, language and communication assessments, and especially response to intervention (RTI) data are used to determine eligibility. LDs (commonly referred to as "learning disabilities" in educational contexts) qualify an individual for special education services and an individual education plan (IEP). In addition, whereas ADHD and ADD may require psychotropic medication as one part of a comprehensive management plan that includes psychotherapies, behavior modification, and specialized academic programming, LDs and BIF do not include a

medication component as part of their comprehensive treatment plan. In specific LDs, performance of *specific* academic skills is well below average for age (low achievement scores ranging from 1.0 to 2.5 SD below the population mean for age). Working memory and processing speed tend to be specifically impaired. In BIF, well below intellectual and adaptive functioning has an impact on functioning as a whole, not merely on specific or isolated academic skills, and so affects interventions and overall progress.

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by persistent deficits in social communication and social interaction across multiple contexts, and restricted, repetitive patterns of behavior, interests, or activities. Reciprocity is severely constrained; pleasurable preoccupations rather than distressing obsessions are typical. DSM-5 codes three levels of severity. In higher functioning ASD (Level 1 severity), IQ scores are average or higher and visuospatial competence may be high, unlike in BIF. Academic performance usually is impaired due chiefly to severe social skills deficits interference with learning, although social approach—awkward as it is—is attempted. Although eye contact is typically impaired, there is no social or emotional indifference. Emotions are strong. Anxiety and dysphoria may emerge from social rejection and isolation. Those with ASD may appear dysmorphic and "quirky" due to difficulties modulating affect in response to social cues and to irregularities in prosody and one-sided, egocentric, long-winded conversational style. Children with BIF, in contrast, appear typical, often quiet, and exhibit modulated ranges of affect unlike the more constricted facial expression reflecting impaired socioemotional reciprocity and perspective-taking exhibited by those with ASD. Those with BIF generally demonstrate socially meaningful pragmatic language and show developing empathetic perspectives, unlike those with ASD.

Nonverbal learning disabilities (NVLDs), a neuropsychological and academic classification and not a DSM disorder, are similar in presentation to higher functioning ASD, and both ASD and NVLD overlap in their clinical features. NVLD has also been called "right hemisphere deficit" and "developmental socioemotional processing disorder." Prefrontal and frontal lobe executive functioning impairments manifest as cognitive, behavioral, and emotional dysregulation. Individuals with NVLD are characterized by less severe social skills deficits ("dyssemia") than those found in ASD. Academic problems—for example, in mathematics and reading comprehension—take center stage. The prominent features of NVLD include overattention to details and missing main themes; intact linear cognitive processing but impaired gestalt comprehension; visuospatial perceptual deficits, especially left-side inattention and neglect; impaired pragmatic language use; and speech-related rhythm, volume, and prosody irregularities. Attentional style is desultory and wandering, in

contrast to distractingly impulsive as found in ADHD and ADD. Neuropsychological testing confirms clinical impressions and may show performance IQ scores, especially low scores on coding and block design subtests (poor visuospatial functioning), to be below verbal IQ scores, which are typically average. The capacity for abstract thinking is dysfunctional regardless of average or higher IQ scores. Children with NVLD appear more motorically passive and not overactive. Anxiety and depression rather than disruptive behaviors are more typical. Individuals with NVLD typically qualify for special education services more easily than those with BIF.

Disruptive, impulse-control, and conduct disorders (CDs) typically draw robust clinical attention and require treatment. CD is characterized by aggression and bullying to people and animals, destruction of property, deceitfulness or theft, and serious violation of rules. These behaviors in children and adolescents are often associated with school failure and an apparent impairment of comprehension of academic and social standards. BIF may be comorbid and contribute to the impaired ability to understand and comply with expected norms in youth with CD. Antisocial personality disorder (APD) is a pervasive pattern of disregard for, and violation of, the rights of others that begins in childhood or early adolescence and continues into adulthood. As in CD in children, BIF can be associated with APD and impaired cognitive skills in adults, and its comorbidity, if present, needs to be identified. Any underlying psychotic disorder whether endogenous or exogenously produced must be ruled out.

In addition, it must be recognized that BIF, like other conditions found in the juvenile population, may be, and usually is, comorbid with (although usually not mistaken for) several other psychiatric disorders. PTSD associated with neglect, maltreatment, trauma, or abuse is now commonly encountered in children whose poor academic performance is primarily affected by anxiety, agitation, and preoccupation with past traumas. Primary mood instability and cycling along with a driven, distracting, and hyperkinetic motor component (cognitive, speech, and physical activity) characterize mood disorder or bipolar disorder in children; disruptive mood dysregulation disorder with severe and recurrent verbal or behavioral temper outbursts and persistently (not episodic)) irritable, peevish, and habitual angry mood may be seen in seven to eighteen year olds. Substance-related disorders such as cannabis use and other drugs that impair learning and memory must also be considered. Clear differentiation and targeted treatments are needed. In addition, BIF has no correlation with borderline personality disorder.

Sluggish cognitive tempo (SCT) is a theoretical construct being investigated as a potential presentation of attentional cognitive impairment. It is not a formal diagnostic entity and has no standardized criteria. Most researchers, however, identify its most salient symptoms as

including the following: daydreaming, trouble staying awake and alert, mentally foggy and easily confused, staring a lot, spacey with mind seeming elsewhere, lethargic, underactive, slow moving and sluggish, inability to process questions or explanations accurately, drowsy and sleepy appearance, apathetic and withdrawn, lost in thoughts, slow to complete tasks, and lacks initiative with fading effort. Although correlated with ADD and ADHD, it is thought to both be comorbid and possibly a distinct condition. While ADHD is considered to be a disorder of executive functions and have externalizing correlates (disruptive behaviors), SCT is not. SCT is correlated with internalizing features such as anxiety, depression, and with social withdrawal. Some have suggested renaming the condition "concentration deficit disorder," "developmental concentration disorder," or "focused attention disorder." Although problems in adaptive performance are evident, there are no data available regarding IQ scores, and, therefore, it would be premature to make correlations between SCT and BIF. In addition, the clinician must carefully differentiate those with temperamental styles that are avoidant, reticent, and retiring (typically discernable by 12 months in infancy and onward), as well as schizoid personality styles from presentations having cognitive sluggishness as an isolated or primary feature.

Course and Prognosis

Although there are no specific studies that show the longitudinal course and prognosis of those with BIF, allied literature describing those with mild cognitive limitations reveals consistent trends. In general, studies show that overall quality of life may be affected in multiple negative ways when the condition is unrecognized and no supports given. Quality of life is reflected in the following domains of human functioning: social relationships, family relations, employment, autonomy, recreation and leisure, personal competence, independent living, community integration, social status, health, finances, future parenting, and subjective well-being. With appropriate habilitation, remediation, guidance, and supports, risk reduction may be achieved and a good quality of life realized. Early identification and treatment are essential. In many instances, a naturally facilitating environment that elicits strengths is found in a good family, a mentor, and supportive community that reinforce positive motivation, and a strong work ethic—all of which continue to encourage success.

A 13-year-old boy was referred to a partial hospital program because of anger outbursts, temper tantrums, and moodiness in the home. He was in a regular classroom setting, and his grades were marginal. He appeared passive, but, when given adult attention and guidance in school and at home, his overall demeanor brightened and performance improved. Along with the psychiatric assessment, evaluation included a pediatric medical examination, blood work, and urine screening for illicit drugs. School psychological testing was reviewed. Results were generally unremarkable. There were no medical problems. However, verbal IQ was 80,

performance IQ was 75, and adaptive skills proficiency was several years below that expected for age. A primary diagnosis of BIF was made. Treatment included intensive emphasis on special help with resource room and accommodations in school, psychoeducation for the family, and brief individual, group, and behavior modification for the patient. Mood and behavior at home significantly improved, self-esteem was enhanced, and school performance—academically and socially—strengthened.

An 11-year-old boy was admitted to a residential facility for further evaluation, treatment, and stabilization of a long history of ADHD and intermittent explosive disorder. He had been on dextroamphetamine (Adderall), 20 mg in the morning and 20 mg at noon, and lithium (Eskalith, Lithobid) 300 mg three times daily. His clinical presentation showed a significantly small-for-age, underweight boy with a moderate degree of facial tics and grimaces, stereotyped behaviors (picking at the skin of the hands, arms, and legs), agitation, and constricted affect. Medical examination and blood work were normal, although tics were noted. Comprehensive assessment revealed a 3-year history of current behaviors, no previous special education resources although grades had been consistently poor to failing, and chronic family strife due to the patient's intermittent irritability, oppositionality, and noncompliance at home, especially regarding homework, bedtime, and getting up in the morning. Evaluation and treatment included a therapeutic trial of the gradual tapering to zero of all psychotropic medications in the context of a structured therapeutic milieu. After 3 months, the patient gained 10 lb, had considerably fewer facial tics, and had few to no explosive episodes. Sleep patterns began to normalize.

Stereotyped behaviors and moderate impulsivity remained, although school performance (behavior and academics) improved. A trial of 0.5 mg of risperidone (Risperdal) twice daily was begun to target residual agitation, impulsivity, stereotypies, and tics. After 2 months, these signs and symptoms resolved, leaving only some mild impulsivity, which was easily redirected in the structured environment. During this period of stabilization, psychological testing was done and showed a verbal IQ of 78, a performance IQ of 76, and an adaptive skills profile well below chronological age expectations. A primary diagnosis of BIF was made. Because no specific learning disabilities were noted, his global school learning problems were addressed in a classroom with general special education accommodations. Furthermore, family therapy continued to help resolve much of the chronic strain in and between parents and patient that had resulted from years of mutual frustration and apparent failure.

Treatment

BIF is a clinical condition, indicating significant risk vulnerability. It requires strong clinical and educational interventions. Although universally agreed upon standardized techniques that target those with borderline cognitive and adaptive functioning are not yet established, a *multimodal management approach* centering on the individual, the family, and the school environment yields best results. Psychosocial therapies, as one part of a comprehensive treatment plan, fostering aptitudes that elicit innate resourcefulness—potentials for intelligent performance—and improved emotional resiliency are essential. Psychotherapy, suitable to the developmental level and individual's receptivity, often helps self-image since so many of those with borderline intellectual and adaptive functioning, feel they "do not fit in" in a "regular" way with others. Since children, adolescents, and adults present with multiple problems, often with multiple psychotropic medications, careful medication assessments are crucial. Psychotropic medications have indications for psychiatric

disorders or features of disorders such as anxiety, agitation, irritability, and mood instability. There is no explicit indication for BIF. Clinical presentations may at times be confounded with iatrogenic medication side effects such as fatigue, sluggishness, cognitive dulling, lethargy, irritability, mood lability, anxiety, poor appetite, insomnia, tics, skin picking, and trichotillomania, to name a few. These potential culprits pose distractions that when eliminated make clearer the child's actual profile of signs, symptoms, and cognitive baseline; testing results are less adversely affected. Judicious use of psychotropic medications in the juvenile population is part of good clinical practice.

Therapeutic diversification—thoughtfully constructed, monitored, and refined over time—nuanced to identify and address individual needs is very good patient care. A restructured individualized academic program is the framework of all interventions. The condition of BIF typically does not qualify an individual for special education services or an individual education plan. Such plans outline specific and measurable educational and behavioral goals that may have quarterly and annually addressed objectives. Instead, a strong, cogent, and reasonable argument for an appropriate level of individualized academic intervention such as Section 504 remedial assistance can be requested to school boards. Acceptance may be successful.

After careful delineation of cognitive, emotional, social, academic, and adaptive strengths and limitations, comprehensive strategies should address not only the individual's specific needs but also the educational supports needed. The entire treatment plan should be user friendly and not aversive in any way to the child, the family, or the academic environment. The treatment-determining context may be school based, within the family, or in a residential, independent living, or group home setting. Particular attention to significant developmental and other life transitions should be highlighted. This includes early considerations of pragmatic vocational planning. Students graduating from high school, for example, might enter a community college transitional program specifically devoted to addressing the adaptive skills needed in adult life. Vocational training for adulthood is part of transition planning. Although interventions have emphasized attention to the intellectual domain, using the adaptive domains—including the physical body's role in kinesthetic activities such as sports and manual labor—should not be underestimated. Procedural memory and forms of implicit learning typically remain potential strengths. Interpersonal communication skills, time management, organizational skills, career exploration, and workplace readiness are taught in such programs. For older adolescents and younger adults, considerations of referrals to bureaus of rehabilitation and vocational services, which may be associated with local and state Departments of Social Services, are appropriate vocational considerations. Treatment planning of resources and strategies

must be reviewed on a regular basis because needs and circumstances will change over time.

Remediation efforts in the developmental period up to age 18 years old are usually school based. These core interventions identify specific areas of learning abilities, weaknesses, and capabilities, as well as adaptive skills, strengths, and limitations. Remediation first addresses increased overall academic and social participation. Moreover, *accommodations*, which are adjustments to the school program that do not change the curriculum level or performance criteria, but that permit the student to use it more efficiently, may be needed. Various forms of assistive technology are commonly used. *Modifications*, by contrast, change curriculum goals and performance criteria. Teaching adaptations must be individualized to afford the student reasonable participation and chances of success, especially since individuals perceive, interpret, and respond differently to new information. Public school programs in charter school settings have been launched and include modified curriculums considering the clinical features associated with BIF, customized educational course plans, daily assessments with ongoing plan revisions, and strong parental involvement. A treatment program with appropriate educational, behavioral, psychotherapeutic, counseling, and social interventions is then constructed. Since novelty and ambiguity may be experienced as stressors, learning is significantly enhanced from teaching styles that are more explicit and hands-on rather than those relying on the individual's gathering, inferring, and generalizing information alone. All students benefit greatly from having a dedicated "go to" person who can act as personal mentor, guidance counselor, and mediator for child and others. All these act as "stress inoculation," which is preparation by exposure that elicits above comfort zone reactions that are learned to be controllable.

Psychosocial therapies focusing on the individual are essential. Various forms of psychotherapy in the broadest sense including counseling, guidance, mentoring, coaching, and support may enhance greater social competence by strengthening social skills, especially communication, comprehension, and self-advocacy. Strengthening executive functioning and diminishing tendencies for impulsive choice—that is, decisions that result in small, immediate rewards despite long-term negative consequences—can be bolstered by psychotherapies. Recognizing boredom and redirecting energies are skills taught through self-management techniques that especially target motivation, initial encoding competence, and sequencing. This approach also imparts strategies that model learning in a step-by-step fashion and the expression of originally learned information in as many different ways as possible to enhance generalization aptitudes.

The decisive role of helping children make sense of emotions and innovate their emotional intelligence has been underestimated in the past.

We now know it can be instrumental in advancing their global performance skills. Precision medicine and, by extension, effective interventions rest on this. Attention to the important role of social competence as allied with cognition and enhanced symbol formation in effective overall life functioning should not be underestimated. Its roots, in part, can be seen in the pioneering work of such distinguished psychoanalysts as Melanie Klein, who, as early as the first half of the 20th century, advanced important theoretical directions that underscored the significant interplay of cognition, symbol formation, social adaptation, and the reciprocal influences within the parent–child relationship in clinical applications. Contemporary psychotherapies tend to be pragmatic and solution focused. Therapists use a variety of theoretical models and tailor techniques to real-time needs and interests, as well as to the responsiveness of patients and families. As with other conditions that affect learning and successful interpersonal and social adaptation, effective contemporary methods built on precedents from the past must now also use multidisciplinary team approaches.

Behavior modification, especially useful in group and academic settings, has an important place in comprehensive interventions. Whether the psychiatrist decides to work with the patient directly or work in a team with other clinicians, the psychotherapeutic and behavioral component can act as a powerful adjunct in promoting change, especially in raising the quality of life for the child and those in the child's extended environment.

An important part of team treatment is the family: identifying its specific needs and offering pragmatic supports. The importance of enlisting the family as an active collaborator in psychotherapies, social skills, school programs, and allied interventions cannot be overstated. Recent meta-analyses show that family behavioral therapy and parent management skills training have significant positive outcomes. In addition, the child or adolescent as a family member should be offered an explicit collaborating role. A specified member of the treatment team, for example, should construct setting up benchmarks for future success, in joint fashion with the child.

Psychoeducation, which has a protective effect on many levels, denotes explaining the meaning and significance of the diagnosis and treatments of borderline intellectual and adaptive functioning to families, therapists, and educational providers. Stressing to parents that their direct, *empathetic hands-on attentive care and guidance* is fundamental to their child's needs over time cannot be overemphasized. Clearly communicating this explanatory, attributional model is an important task for the psychiatrist. Studies have shown that parenting skills that emphasize genuine warmth and responsiveness act to augment adaptive outcomes. Because children with BIF are usually emotionally sensitive and responsive, especially to nonverbal communications, positive perceptions of genuine helpfulness

from all caregivers are exceedingly effective in facilitating cooperation and change. Highlighting and reinforcing all indications of effort bolsters motivation and self-esteem. All comorbid conditions must also be addressed in the comprehensive treatment plan. Behavior modification strategies (coordinated with techniques used in school) are important in home settings. Meta-analyses clearly demonstrate that family behavior therapy and management skills training as well as cognitive-behavioral therapy for school age children is both clinically and cost-effective.

Special mention of the current widespread use of interventions derived from complementary, alternative, and integrative medicine may have a place in overall treatment because motivated families often first initiate these aids with enthusiasm not only to promote but also to enhance well-being. Healthy food choices, exercise, stress management, and even prenatal health measures (primary prevention) can be important components of a successful treatment plan. When used adjunctively, rationally, and collaboratively, they may help and sustain a positively engaged therapeutic trajectory.

Proper recognition together with remediation enhances both cognitive and adaptive reserve. This optimizes potentials that extend benefits beyond the immediate needs of the identified individual with BIF. *Risk reduction may reach across generations*. The young children of such individuals are thereby given a significantly increased chance of success in all measures of quality of life. In addition, the adult children of persons with borderline intelligence can be given a rational explanation of insights into parents' long-standing intellectual limitations. This may afford them a capacity for a forgiving attitude toward what might have erroneously been perceived as intentionally motivated, less-than-optimal parenting.

A 10-year-old boy was referred for consultation because of failing grades since first grade and disruptive behaviors in the classroom, which included expressing aggression toward peers, breaking things, and having temper tantrums. At home, he would purposefully break any toys or gifts he was given; at times, he would tear his clothes. He was the only child of a single mother who reported a normal gestation, delivery, and no medical problems. On interview, the boy was mildly overweight, had some facial dysmorphic features, but appeared otherwise healthy. His mood was sullen. In the past, his mother had been given a variety of diagnoses for the child by school and pediatrician: ADHD, bipolar disorder, ASD, and CD. No psychotropic medications had been helpful. A comprehensive assessment included multiple interviews with the child and the mother, school consultation with intelligence, adaptive, and scholastic testing, and routine blood work. Verbal and performance IQ scores were in the low 70s, adaptive skills were marginal, and grades were 1 to 2 years behind. Laboratory values were within normal limits. A primary diagnosis of BIF with mood disorder, unspecified, was made. Intensive treatment was three pronged: child, mother, and school. School consultation provided psychoeducation about diagnoses and the need for special education, which was approved. Family therapy also included psychoeducation about borderline intelligence, the child's strengths and limitations, and his need for a structured, clear, and routine environment of expectations and behavioral performance. Individual therapy for mother also began. His mother was eager to use a range of health-oriented dietary and exercise guidelines. Intensive individual psychotherapy included aspects of cognitive, behavioral, and psychodynamic work. The child's sense of self, self-esteem, peer relations, and the effects of his behaviors on others

were reviewed. During play therapy, several conflictual problems were uncovered. The boy revealed his envy of other boys on multiple levels. He envied their having a father, his perceived sense of their being "smarter and not retarded," their physical competence in sports, and their acceptance and facile participation in peer group activities. His envy created hatred of, and a wish to destroy, both them and, moreover, aspects of himself, as illustrated in his either spoiling his clothing or breaking gifts he received. He often stated: "I'm angry because I just don't fit in with everybody else." Even though he was cognitively limited, he showed an interest in and some aptitude for understanding the complex envy dynamic of excessive idealization, spoiling, destructiveness, and resentment—when put into simple language and delivered flexibly at his level of readiness. Play therapy and interpretive discourse were sensitive to engage him on the specific and changing levels of cognition and affect that he presented at each session. Changes were slow but significant. This developing level of emotional insight appeared to enhance his cognition because it diminished a great deal of anxiety and preoccupying resentment. The acquisition of "envy management skills" over time helped him to begin to understand cause and effect sequences and so constrain destructive behaviors. Once-weekly psychotherapy was supplemented by small amounts of quetiapine (Seroquel). After approximately 6 months, his behavior stabilized, his aggressive outbursts were significantly eliminated, and his academic performance began to approach chronological and developmental expectations. His mother remarked that his demeanor appeared more relaxed and that his affect brightened.

ACADEMIC OR EDUCATIONAL PROBLEM

Definition

DSM-5 describes *academic or educational problem* as follows:

> This category should be used when an academic or educational problem is the focus of clinical attention or has an impact on the individual's diagnosis, treatment, or prognosis. Problems to be considered include illiteracy or low-level illiteracy; lack of access to schooling owing to unavailability or unattainability; problems with academic performance (e.g., failing school examinations, receiving failing marks or grades) or underachievement (below what would be expected given the individuals intellectual capacity); discord with teachers, school staff, or other students; and any other problems related to education and/or literacy.

The title of this entry now includes the added phrase "...or Educational Problem." Unlike the DSM-IV-TR description, no mention is made of this condition not being "...due to a mental disorder or, if due to a mental disorder, is sufficiently severe to warrant independent clinical attention." About seven additional problems are now mentioned: illiteracy, low-level illiteracy, lack of access to schooling, discord with teachers, school staff, students, and other problems related to education or literacy besides the previous two in DSM-IV-TR: failing grades and significant underachievement. No mention is made of requiring the absence of a specific learning or communication disorder or any other mental disorder that would account for the problem, as DSM-IV-TR had required. The term "illiteracy" is not defined in specific fashion; inability to read and to write, therefore, is implied. Some of the associated factors causing academic or educational problems are remediable extrinsic agents (e.g., access, discord, implied bullying, and victimization) that, if addressed, would assuage or eliminate the designated academic problem. DSM-5 coding is Z55.9.

Etiology

The presentation of significant academic or educational problems denotes this condition. DSM-5 clearly lists several causative and clear-cut associative factors, leading to academic and educational problems of this nature. Some of these may have intrinsic causes such as illiteracy associated with dyslexia or LDs. Most of the associated causative agents are apparent environmental difficulties such as lack of access, interpersonal conflict, and perhaps behavioral dissonance, as delineated in the DSM-5 description. Bullying and victimization must be a primary consideration.

Identification and Clinical Features

Scholastic problems, acute or chronic, that are associated with the very specific factors clearly mentioned in this code's description best characterize the notation of academic or educational problem. Specific medical conditions, disabilities, or primarily overshadowing psychiatric disorders such as psychosis, schizophrenia, bipolar, depression, ASD, substance-related disorders, or PTSD may be comorbid but not sufficient to account for this coding. Clinical judgment is required.

Differential Diagnosis

Adequate psychological and intellectual testing, as well as profiles of academic performance derived from standardized scholastic testing, needs to rule out forms of ID, BIF, specific LDs, and communications disorders. All medical conditions and disabilities that might contribute to poor school performance must also be explored. Major psychiatric disorders, along with severe oppositional defiant disorder, must be considered. Substance abuse or dependence and CD must be ruled out.

One must also rule out clinical and subclinical states of anxiety and depression, as well as environmental stressors, either hidden or explicit, which significantly affect the child in the family (e.g., marital discord, divorce, abuse, and mental illness in parent), the socioeconomic environment (poverty, malnutrition), and the suitability of the school setting (accessibility, presence of bullying, and so forth). Significant medical conditions (anemia, seizures, hypothyroidism, chronic infections, and genetic disorders) must be considered. Adolescent developmental concerns, such as preoccupation with burgeoning sexuality, identity issues, and peer relations, as factors that contribute to academic problems should not be underestimated. Typically, such preoccupations are highly significant and may contribute to academic problems. In addition, motivational issues and associated family and socially influenced contextual values may directly affect scholastic performance.

A 15-year-old boy was admitted to an intensive outpatient program for the evaluation of

insidiously declining grades at school, isolative behavior, and anhedonia occurring for the first time over the course of the last semester at his second year in high school. The patient had no previous psychiatric history and was in good health. School reports showed full-scale IQ score to be 100 and no previous behavior or academic problems. Interviews with family and patient revealed that the boy had an intact family with no apparent or unusual conflicts noted. His mother found evidence of drug paraphernalia—cigarette papers, small pipe, matches, and a suspicious-looking dried substance—in the patient's bedroom. When confronted, the patient revealed a 6-month history of regular use of marijuana associated with his statement of being bullied in school. A primary diagnosis of cannabis abuse along with academic problem was made. Referral to a therapeutic drug and alcohol program was immediately implemented. A comprehensive school meeting was called to investigate and address the allegation of bullying.

A 10-year-old girl was admitted to an intensive outpatient program due to a variety of disruptive, impulsive, hyperkinetic behaviors, and intermittent rage reactions only in the home environment. The patient was well behaved in school and considered bright. Her grades, for the first time over the course of the last year, were failing. Onset began about 2 years previously when her parents separated and her mother's new male companion moved in. The child's pediatrician made a diagnosis of ADHD and started methylphenidate (Ritalin). The child's condition worsened, and was marked by periodic temper outbursts, now also in school. Medical assessment showed the patient to be in good health, and the child was referred for further assessment. Psychiatric interview showed a guarded, hypervigilant child with a constricted affect. No other data were forthcoming. The patient's mother was reluctant to participate in treatment. The initial diagnostic impression was of disruptive behavior disorder along with academic problem. Over the course of several weeks and months, the patient revealed that she was being sexually abused. Further investigation and involvement with the Department of Child and Family Services revealed that the adult man in the home was the perpetrator. He was removed. Family and individual work began and continued for several months. All psychostimulant medication was discontinued. A primary diagnosis of PTSD was made, along with abuse and secondary academic problem. The child, over time, became less irritable, rages stopped, and grades in school improved.

Treatment

Identifying and addressing environmental—family, school, and socioeconomic—stressors are important. Substance abuse must be ruled out. Any concurrent major psychiatric disorders require diagnosis and treatment. Since referred children and adolescents are typically prescribed multiple psychotropic medications, careful medication assessment is crucial. Side effects, adverse drug effects, behavioral toxicity, and iatrogenic effects (e.g., hypothyroidism from lithium, or overactivation, irritability, mood instability, underweight, tics, stereotypies, and insomnia from psychostimulants) must be ruled out. When clinically indicated, taper and discontinuation of medications may be beneficial. With informed consent, school consultation can provide teachers with insight into the student's difficulties and suggest enhanced teaching strategies, especially conflict resolution techniques. Matters concerning bullying are now common and require careful consideration and treatment team coordinated efforts. Child, family, school, and professional collaboration are essential. Illiteracy, underachievement, and failing grades need to be evaluated and appropriately addressed. If academic problems are severe enough, accommodations for individual educational programming,

psychotherapeutic interventions such as counseling and guidance, or residential schooling should be recommended even though this condition may not meet current stringent criteria for formal special education eligibility.

FUTURE DIRECTIONS

Intellectual *abilities*, social and emotional cognition, motivation, and the successful implementation of these as adaptive performance skills through thought processes and behavior are functions of the developing, dynamic, and gradually unifying personality. Innovating emotional intelligence magnifies what I conceptualize as and term *emotion performance utilization* (EPU). EPU is not a phase but the target of emotional intelligence. Viewing children and adolescents in this integral fashion also implies assessments and interventions that are comprehensive, multimodal, and flexible in accord with the biopsychosocial paradigm.

In the introductory comments to this section, recent and ongoing international efforts promoting evidenced-based research and implementation of recognition and intervention into borderline intellectual and adaptive functioning are described as an area of active pursuit. Perhaps, a case can be made for translating clinical practice that is patient-centered into universally applicable research. In the United States, the AAIDD (2010) has proposed a rationale for formalizing a continuum of ID with a support framework that spans IQ limitations on either side of the definitional dividing line. AAIDD urges that individuals with BIF or persons with mild cognitive limitations—who fall outside the "cutoff" for ID—should be considered eligible for services already in place for those formally diagnosed with ID. This admonition is consonant with this section's call for integration of BIF into the mainstream of official psychiatric nosology and reflects the art guiding scientific psychiatry. The Journal of the American Medical Association (August, 2014) has advanced the viewpoint that investing in prevention should be a strategic national priority to help improve the lagging population health of the United States compared with peer countries.

A new paradigm, the *biomental* perspective, has recently been introduced into psychiatric theory and been applied to child development and parenting. The integral confluence of brain, mind, and behavior is advanced with reference to neuroscience, developmental psychology, and pragmatic parenting strategies. Adaptive intelligence, a hypothetical construct, is presumed in this model. It denotes the central plexus subtending the wide range of conceptual, emotional, and motivational actual abilities underlying ease of effective responsiveness to change. This hub includes a tacit reserve of potential capabilities whose platform consists of an individual's innate resilience. This intelligence potential can be made active by therapeutic techniques that foster "instinctive

resourcefulness," the repository of silent aptitudes, which are learning-sensitive mediators between mental and emotional resources and their timely, spontaneous, and successful implementation as practical problem-solving performance skills. This innate capacity of resourcefulness—possibly distinct from abilities currently measured by and solely dependent on IQ scores—is considered a fundamental life tool. Although its functioning may be significantly limited in those with borderline cognition and adaptive skills, therapeutic initiatives as described in this section may activate its enhanced operation.

Current efforts to integrate descriptive and explanatory insights from genetics, neuroscience, imaging of neurocircuitry, behavioral, and clinical studies are aimed toward constructing valid and reliable phenotypes that span the heterogeneous biomental range of typical to atypical, including what may be considered "borderline." Incorporating such contemporary perspectives eschews considering only single models of categorical reductionism and lauds the reality of the inclusive fact of heterogeneity in the face of complex psychiatric presentations. Inclusive treatments for BIF may also include parental guidelines that recognize *envy management skills*, which may promote understanding, reduce cognitive and emotional dissonance, and foster greater cognitive clarity and emotional stability, thereby optimizing both intellectual and adaptive functioning and proficiency. Envy management skills may be adapted to the individualized support resources of those with BIF.

The environmental embeddedness—*eco-corporeality*—of each individual is assumed in these approaches. The integrated use of such inclusive interactional paradigms is needed to achieve a sustainable, successful, and finer quality of life. This denotes a life imbued with meaning and dignity across multiple eco-contexts, including the personal, interpersonal, social, scholastic, familial, occupational, cultural, and spiritual.

References

*Baglio G, Blasi V, Sangiuliano Intra F. Social Competence in Children with Borderline Intellectual Functioning: Delayed Development of Theory of Mind Across All Complexity Levels. *Front Psychol*. 2016;7:1604.

*Barkley RA. Sluggish cognitive tempo: a (misnamed) second attention disorder? *J Am Acad of Child and Adoles Psychiatry*. 2016;55(3):157–158.

*Becker S, Marshall S, McBurnett K. Sluggish cognitive tempo in abnormal child psychology: an historical overview and introduction and introduction to the special section. *J Abnorm Child Psychol*. 2014;42(1):1–6.

Cesare C, Goofier D, Orsini A, Peseta L. Differences in the intellectual profile of children with intellectual vs. Learning disability. *Res Dev Disabil*. 2014;35(9):2224.

Edgerton R. *The Cloak of Competence: Stigma in the Lives of the Mentally Retarded*. Rev. ed.

Berkeley, CA: University of California Press; 1992.

Emerson E, Einfeld S, Stancliffe RJ. The mental health of young children with intellectual disabilities or borderline intellectual functioning. *Soc Psychiatry Psychiatr Epidemiol*. 2010;45:579.

Fenning RM, Baker, JK, Baker BL, Crnic KA. Parent-child interaction over time in families of young children with borderline intellectual functioning. *J Fam Psychol*. 2014;28(3):326–335.

Fernell E, Ek U. Borderline intellectual functioning in children and adolescents—insufficiently recognized difficulties. *Acta Paediatr*. 2010;99(5):748–753.

Ferrari M. Borderline Intellectual; functioning and the intellectual disability construct. *Intellect Dev Disabil*. 2009;47(5):386.

Fletcher R, Loschen E, Stavrakaki C, First M. *Diagnostic Manual—Intellectual Disability (DM-ID): A Textbook of Diagnosis of Mental Disorders in Persons with Intellectual Disability*. New York: NADD Press; 2007.

Hassiotis A. Borderline intellectual functioning and neurodevelopmental disorders: prevalence, comorbidities and treatment approaches. *Advances in Mental Health and Intellectual Disabilities*. 2015;9(5):275–283.

Hassiotis A, Strydom A, Hall I, et al. Psychiatric comorbidity and social functioning among adults with borderline intelligence living in private households. *J Intellect Disabil Research*. 2008;52(2):95.

Kaufman AS. *IQ Testing 101*. New York: Springer Press; 2009.

Klein M. Our adult world and its roots in infancy [1959]. In: *Envy and Gratitude and Other Works*. London: Hogarth Press; 1975:247.

Mascolo JT, Flanagan DP, Alfonso VC. *Essentials of Planning, Selecting, and Tailoring Interventions for Unique Learners (Essentials of Psychological Assessment)*. New Jersey: Wiley; 2014.

Miller DC. *Essentials of School Neuropsychological Assessment (Essentials of Psychological Assessment)*. 2nd ed. New Jersey: Wiley; 2013.

*Ninivaggi FJ. Attention-deficit/hyperactivity disorder in children and adolescents: rethinking diagnosis and treatment implications for complicated cases. *Connecticut Med*. 1999;63:515.

*Ninivaggi FJ. Borderline intellectual functioning in children and adolescents: reexamining an underrecognized yet prevalent clinical comorbidity. *Connecticut Med*. 2001;65:7.

Ninivaggi FJ. DM-ID, Diagnostic Manual-Intellectual Disability: A Textbook of Diagnosis of Mental Disorders in Persons with Intellectual Disability, (book review). *J Autism Dev Disord*. 2008;38:204.

Ninivaggi FJ. *Ayurveda: A Comprehensive Guide to Traditional Indian Medicine for the West*. Lanham, MD: Rowman & Littlefield Press; 2010.

Ninivaggi FJ. *Envy Theory: Perspectives on the Psychology of Envy*. Lanham, MD: Rowman & Littlefield Press; 2010.

Ninivaggi FJ. *Biomental Child Developmental: Perspectives on Psychology and Parenting*. Lanham, MD: Rowman & Littlefield Press; 2013.

Ninivaggi FJ. Emotion as a second language or should they be our First?—emotional literacy: a forgotten first language to be remembered. *Psychologytoday.com*. February 20, 2015.

*Ninivaggi FJ. *Making Sense of Emotion: Innovating Emotional Intelligence*. Lanham, MD:

Rowman & Littlefield Press; 2017.

Nouwens PJG, Lucas R, Petri JCM. Embregts, van Nieuwenhuizen C. In plain sight but still invisible: a structured case analysis of people with mild intellectual disability or borderline intellectual functioning. *JIDD*. 2016;1–9. Published online: 04 May 2016. http://dx.doi.org/10.3109/13668250.2016.1178220

*Salvador-Carulla L, Garcia-Gutierrez, Gutierrez-Colosia, MR, et al. Funcionamiento intellectual limite: guia de consenso y buenas practices. *Rev Psiquiatr Salud Ment (Bar.)*. 2013;6:109–120.

*Schalock RL. *Adaptive Behavior and Its Measurement*. Washington, DC: American Association on Mental Retardation; 1999.

*Shalock RL, Borthwick-Duffy S, Bradley MA, et al. *Intellectual Disability: Definition, Classification, and Systems of Supports*. 11th ed. Washington, DC: American Association on Intellectual and Developmental Disabilities; 2010.

Shalock RL, Luckasson R, Bradley V, et al. *User's Guide: Intellectual Disability: Definition, Classification, and Systems of Support*. Washington, DC: American Association on Intellectual and Developmental Disabilities; 2012.

Shaw S, Grimes D, Bulman J. Educating slow learners: are charter schools the last, best hope for their educational success? *Charter Schools Resource J*. 2005;1:10.

Snell ME, Luckasson R, Borthwick-Duffy WS, et al. What is a mental/psychiatric disorder? From DSM-IV to DSM-V. *Psychol Med*. 2010;40:1759.

*Thompson JR, Bryant BR, Campbell EM, Craig EM, Huhes CM. *Supports Intensity Scale: Users Manual*. Washington, DC: American Association on Mental Retardation; 2004.

*Thompson JR, Bradley V, Buntinx WH E, et al. Conceptualizing supports and the support needs of people with intellectual disability. *Intellect Dev Disabil*. 2009;47(2):135.

Tymchuk AJ, Lakin KC, Luckasson R. *The Forgotten Generation: The Status and Challenges of Adults with Mild Cognitive Limitations*. Baltimore, MD: Paul Brooks; 2001.

*Wieland J, Zitman FG. It is time to bring borderline intellectual functioning back into the main fold of classification systems. *B J Psych Bull*. 2016;40(4):204–206.

29

Cultural Concepts of Distress and Assessment

▲ 29.1 Cultural Concepts of Distress

ROBERTO LEWIS-FERNÁNDEZ, M.D., LAURENCE J. KIRMAYER, M.D., PETER J. GUARNACCIA, PH.D., AND PEDRO RUIZ, M.D.

INTRODUCTION

Cultural psychiatry has documented wide variations in the experience and expression of mental distress around the world. Historically, these variations were often characterized in terms of "culture-bound syndromes"—distinctive patterns of co-occurring symptoms and signs that were assumed to be rooted in or "bound" to local features of culture and context. With advances in research, this view has been challenged and a more refined set of terms have been proposed to capture the ways in which symptoms, explanations, and expressions of distress vary across cultures. This chapter reviews these newer concepts and approaches, which have been incorporated into DSM-5, and illustrate how they can be used to better understand the causes, course, and consequences of mental health problems in the community and the clinic.

Contemporary cultural psychiatry and psychology recognize that all classification schemes are rooted in particular cultural and historical frames. In turn, the diagnostic language of psychiatry influences popular culture and everyday understandings of mental health problems, ensuring that professionally defined categories help pattern social experience. Hence, Western psychiatric nosologies need to be subjected to the same kind of social, historical, and contextual analysis applied to diverse local nosologies.

Despite local variations, however, similar forms of distress have been documented in many regions of the globe, resulting in clusters of symptoms or syndromes, such as depression or schizophrenia that can be described at a level of abstraction that allows broad cross-cultural application. Local cultural variations in symptom interpretation and

experience will nonetheless influence the recognition, response, and outcome of these disorders. Moreover, some local variations may depend on specific culturally based attributions, behavioral responses, and patterns of interpersonal interaction so that they only occur where those modes of explanation and interaction are present. These indigenous illness forms may occur in Western as well as non-Western contexts. The study of cultural variations in the presentation of illness and distress yields a valuable critique of Western classification schemes by highlighting the assumptions of existing nosology and identifying alternate, indigenous categories that may better fit the experience of patients from diverse backgrounds.

Terminology and Conceptualization

Cross-cultural mental health researchers have introduced several terms to refer to and describe culture-specific forms of expressing and diagnosing emotional distress, including cultural syndromes, popular or folk illness categories, idioms of distress, and illness explanations or attributions. The need for multiple terms reflects the complexity of the ways in which people experience, understand, and communicate distress.

The term "culture-bound syndrome" was used in the past to describe patterned behaviors of distress and/or illness that appeared to be distinct from existing psychiatric categories and were considered unique to particular cultural settings. The clear implication was that Western psychiatric categories were not culture-bound, but rather universal, and that proper characterization would disclose a simple mapping of non-Western syndromes onto the categories of international nosology.

The term "syndrome" generally refers to (1) a collection of co-occurring symptoms and signs with (2) a distinctive temporal course, which may include an immediate precipitating cause or etiological context and response to treatment. Most of the entities described in the DSM as disorders are syndromes, in that their identification depends on symptoms and signs since their underlying causes and mechanisms remain uncertain and there are no reliable biomarkers. The categories are clinically useful largely because they have implications for treatment choice and prognosis. As underlying mechanisms become clear, new distinctions can be made and specific symptoms, signs, and behaviors may be distinguished and given more diagnostic weight. However, the symptom cluster itself may remain important as a central feature of patients' experience.

The notion of a syndrome being bound by culture is more problematic. The attribution of "boundedness" emerged early in the history of cultural psychiatry in colonialist contexts in which observers had limited understanding of local cultures. Hence, there was a strong tendency to exaggerate difference, generate stereotypes, and "exoticize" the Other. Reports that focused on a particular cultural group in a community noted

patterns of problematic behavior, which appeared distinctive or unique to the (usually) Western observers, whether they were missionaries, anthropologists, or clinicians. In some cases, the behavior pattern had a specific popular label and treatment within an indigenous medical system. The foreign observers often believed that the communities in question had been socially isolated and, therefore, saw the patterns of distress as unique and locality bound. Considering the massive migrations and cultural diffusions that have occurred throughout human history, this assumption of boundedness has rarely been tenable. For example, in the late 1950s, Raymond Prince described a Nigerian culture-bound syndrome of "brain fag" characterized by feelings of mental fatigue, difficulty concentrating, heat in the head, and myriad other somatic symptoms. Brain fag was found among students who were the first in their families or communities to pursue higher education and was linked to local ethnophysiological ideas about the importance of central heat and the head in African psychology. However, recent historical evidence shows that "brain fag" was commonly described in England in the late 1800s and, presumably, traveled to Nigeria as part of the colonial system to be adopted as a local mode of expressing distress. Hence, while there may indeed be an important fit between the syndrome of brain fag and local cultural concepts, it must also be understood in relation to these colonial cultural influences.

It would be more accurate therefore to term these conditions *cultural syndromes*, de-emphasizing their boundedness because, while they originate in particular cultural areas, they spread with migration and communication and are incorporated into local concepts of illness and expressions of distress. In the contemporary global market of intensified cultural exchanges, exportable syndromes readily become detached from geographical areas. Whatever boundedness they have inheres not in specific locales but in their adoption by ethnic or linguistic groups and links to the sociocultural patterns of which they form part, whether in their countries of origin or in new homes.

The term "syndrome" implies a relatively fixed pattern of symptoms. Only a subset of culture-specific expressions of distress around the world shows this level of organization. Many are typified instead by a more diffuse set of complaints, symptoms, and predicaments that vary across situations and settings, age cohorts, and social subgroups. Rather than representing separate illnesses, they constitute a general category of distress, part of an indigenous system of classifying suffering, sometimes encompassing not only illness but also other forms of misfortune or adversity. Reports of *nervios* (nerves) in Latin America, for example, comprise a continuum of conditions, labeled with different but related terms (e.g., "being nervous since childhood," "suffering from nerves," "being ill with nerves," and "attacked by my nerves") all denoting different gradations of vulnerability and of the effects of adverse events on the

nervous system, including symptoms, signs, and causal attributions. As another example, Haitians may distinguish broadly between "*maladi moun*" and "*maladi bon Dieu*"—the first referring to a set of illnesses ("maladi") triggered by humans ("*moun*," literally "man"), usually through spiritual means, and the latter referring to diseases of the natural order, sent by God ("*bon Dieu*"). A great diversity of human afflictions fall under these general headings. Anthropologists use the term *folk illness* to refer to local illness categories like these that are based on particular systems of medicine or broader cultural ontologies, independently of whether they are organized as discrete syndromes or not. A related term is *popular illness category*, which simply reflects the fact that these are labels or constructs commonly used by lay people in the general population.

One drawback to the term "folk illness" is that the descriptor "folk" tends to connote less affluent and formally educated sectors of society and ignores the fact that these categories are used in complex medical systems embraced by large segments of the population in modern, urban, industrialized societies. Thus, Traditional Chinese Medicine or Indian Ayurveda has many categories of illness defined in terms of the imbalance of energies or vital essences, respectively. Another disadvantage of the term is that many local illness labels are used to describe afflictions and expressions of distress that are on a continuum with folk illnesses but are not strictly illnesses themselves (much as the term "depression" refers to a psychiatric disorder as well as to lay descriptions of much milder reactions to adversity). DSM-5 uses the more general term *cultural concepts of distress* to encompass folk illnesses, cultural syndromes, and two other types of cultural expressions of distress described below.

One of these is the notion of a *cultural idiom of distress*, a key concept in recent medical anthropology. Introduced by anthropologist Mark Nichter in 1981, the term "idiom" emphasizes the idea of a local linguistic or nonverbal mode of communication. It highlights the communicative function of the cultural concepts of distress, the notion that suffering is expressed in particular local "languages," which are shaped by cultural understandings of what are appropriate and meaningful ways of denoting affliction. "Idiom" also does not presuppose pathology. Many idioms—like "thinking too much," "stressed out," or lay notions of "depression"—can be used to express everyday concerns or levels of distress that may be subclinical. However, idioms of distress also provide a way to express more severe suffering that leads to help seeking. Hence, the clinician must be able to decode the significance of such expressions to identify dimensions of the problem that require intervention.

A final cultural category encompassed under the term "cultural concepts of distress" involves the many expressions of distress pointing to *causal explanations* that are rooted in local systems of knowledge. These in turn may reference particular ontologies involving ethnophysiological ideas

about humors, energies or other processes, ethnopsychological theories about the mind, sociomoral explanations in terms of problems in the social order, or spiritual or other religious and cosmological systems. As etiologies, these causes are not necessarily linked to particular signs and symptoms, but rather denote explanations for the source of the problem that usually guide help seeking. Table 29.1–1 summarizes the key elements of the constructs used in this chapter to describe the cultural concepts of distress.

Cultural syndromes, folk or popular illness categories, idioms of distress, and explanations commonly coexist and the same local term, or closely related terms, may refer to more than one of these types of concepts. For example, *nervios* terms in Latin America can be used to describe a syndrome (e.g., as the subcategory *ataque de nervios*, a stereotyped sequence of precipitants, behaviors, and sequelae), a folk illness category (e.g., when denoting an illness "affecting the nervous system," as in "she is ill with nerves"), an idiom (e.g., when communicating the nature of the affliction, such as, "my nerves are killing me"), or an explanatory cause (e.g., "he was born with emotional problems because his mother suffered from nerves during her pregnancy"). Moreover, usage may change over time, reflecting cultural changes and the influence of the health-care system itself. Despite this complexity, the concepts provide a window into the ways that patients and their families understand their problems and are essential for accurate epidemiological research, clinical diagnosis, and assessment, as well as designing appropriate and effective interventions. The following sections further trace the origins and use of these terms in psychiatric theory, research, and practice.

Table 29.1–1.
Key Elements of the Constructs Used to Describe the Cultural Concepts of Distress

Folk Illness Categories	Cultural Syndromes	Idioms of Distress	Causal Explanations
Encompass cultural syndromes, idioms of distress, and causal explanations May include afflictions and predicaments not considered illnesses in the local culture	Organized as a collection of co-occurring symptoms and signs Have distinctive temporal courses, precipitants, causal attributions, and treatment responses	Constitute a linguistic or nonverbal mode of communication Represent a local "language" for expressing distress Do not presuppose pathology	Reference etiologies rooted in local systems of knowledge Are not necessarily linked to particular signs and symptoms Guide help seeking

HISTORICAL BACKGROUND

Pioneering Work

In 1962, Pow Meng Yap, an early exponent of cultural psychiatry, introduced the term *atypical, culture-bound psychogenic psychoses* to denote what he later called ". . . forms of psychopathology produced by certain systems of implicit values, social structure and obviously shared beliefs indigenous to certain areas." In response to changing nosological paradigms, this term evolved over time into "atypical, culture-bound reactive syndromes," from which the label of "culture-bound syndromes" emerged. Yap's selection of the term "culture-bound" was based on his view that aspects of culture (i.e., implicit and explicit values and notions, social structure, and folk practices) pattern the expression of psychopathology to such a degree that particular forms of illness may be confined to the geographical boundaries of the cultural group in question. He expected that, once analyzed further, these apparently unique forms of mental illness would prove to be atypical variations in generally distributed psychogenic disorders. Basic to Yap's formulation was an acceptance of the universality of psychoanalytic constructs, on which culture could exert pathoplastic effects. This implied it should be possible to locate all of these "exotic" syndromes within the classificatory systems of clinical psychiatry and Yap proposed a formal ordering of these syndromes within a universal nosological schema for clinical, didactic, forensic, and administrative purposes. Yap also suggested that attempts should be made to distinguish "essential" from "accessory" symptoms as well as to establish the "natural history" of each disorder in its "elementary psychogenic form" in order to accurately parcel the syndromes among psychiatric and somatic illness categories. His expectation was that such studies would lead first to a consistent definition of each disorder in the form of modal symptoms and then to universally valid mechanisms of psychopathology, providing a firm basis for cross-cultural clinical research.

In the 1980s, Raymond Prince and co-workers expanded on Yap's writings. They proposed a definition of these phenomena as a "collection of signs and symptoms (excluding notions of cause), which is restricted to a limited number of cultures primarily by reason of certain of their psychosocial features." This definition attempted to narrow Yap's "culture-bound syndromes" exclusively to the medical denotation of "syndrome" as a collection of descriptively defined symptoms that are "a reasonably complete descriptive picture of nature, remain fairly constant over time and are verifiable by all investigators." This kind of nosological entity was distinguished from "labels" and "causal notions" superimposed idiosyncratically on universal psychopathological entities. In this view, these terms conveyed much less scientific accuracy and were prone to nonverifiable misconceptions. Prince attributed some of the controversy surrounding the "culture-bound syndromes" to the lack of focus on

descriptive categories defined by recurrent symptom patterns. "Culture-bound syndromes" were portrayed as useful to medical anthropologists and cross-cultural psychiatrists as well as to epidemiologists. The former can use them to explain the relationships between symptom patterns and cultural processes, while the latter need to take them into account in order to accurately and completely measure the full range of psychopathology in a given population.

Following on the methods of Yap and Prince, other authors continued to categorize "culture-bound syndromes" according to contemporary professional paradigms, usually adhering to Prince's formula of stressing their descriptive aspects. For instance, Ronald Simons emphasized the phenomenological similarity of groupings of these conditions across historically unrelated and culturally dissimilar times and places, despite differences in the meaning assigned to the experiences and behaviors. He contended that it is feasible to insert the "culture-bound syndromes" into contemporary diagnostic systems; particularly once they are grouped as a cluster of related syndromes. The term *taxon*, borrowed from biological classification, was employed to denote a grouping of similar items, but without specifying the level of abstraction of the similarity or suggesting that the grouping as presently elaborated is in any way exhaustive or final. In their 1985 text, *The Culture Bound Syndromes*, Simons and his colleague Charles Hughes used the taxonomic principle to group the cultural concepts based on their phenomenological similarity across diverse cultural settings. For example, under a "Startle Matching Taxon" they listed *latah* (from Malayo-Indonesia) and *imu* (from Japan), defined in purely behavioral terms, roughly as follows: someone is startled, after which he or she says things that are normally socially taboo, talks in a manner locally considered nonsensical and amusing, matches the movements of others, and obeys or defies commands in an exaggerated and stereotyped way. Like Yap and Prince, this kind of classification method focused on a verifiable cluster of behaviors and experiences as the defining elements of the cultural concepts of distress, and de-emphasized other aspects of the illness experience, such as causal attributions or social meanings that are dependent on the context in which the cultural concepts arise.

Subsequent Work in Cultural Psychiatry

Ruth Levine and Albert Gaw distinguished several kinds of cultural concepts of distress. "True" "culture-bound syndromes" were organized as recurrent collections of symptoms, which should be classified according to their "predominant symptom or symptoms" and listed as culturally patterned versions of the most relevant DSM category (e.g., *amok* [from Malayo-Indonesia] under the dissociation phenomena and *boufée delirante* [from Western Africa and other settings] under the psychoses). By contrast,

"illnesses of attribution" were organized around indigenous etiological notions (e.g., *susto* [from Mexico and other Latin American regions] in which sudden fright will cause the soul to leave the body, making the individual vulnerable to a variety of ills) rather than discrete symptom clusters. And "idioms of distress" are more general terms that do not have specific correspondences to psychiatric categories (e.g., *nervios* [from Latin America], a general term used as an idiom for stress-related illness that can applied to disparate psychiatric disorders, from adjustment problems to psychosis). The authors themselves acknowledged some of the arbitrariness of this approach, however, as well as the technical difficulties involved in parceling out indigenous expressions of psychopathology into their three categories. For instance, in *hwa-byung*, suppression of anger is seen as a causative factor in the development of a somatic symptom cluster (indigestion, dyspnea, fatigue, and the sensation of a mass in the stomach) associated with emotional symptoms (anxiety and depression). The fact that this folk condition is organized around a Korean etiological notion that ascribes to anger (*hwa*) the capacity to produce illness (*byung*) led these authors to categorize *hwa-byung* as an illness of attribution, whereas it has been described as a cultural syndrome by others due to its usual presentation as a recurrent symptom cluster. In fact, most cultural syndromes contain embedded notions of attribution as organizing elements for the symptom clusters.

Other authors have taken a more mixed approach, combining universalistic views with the notion that each social group creates unique forms of psychopathology that are fundamentally inseparable from the cultural preoccupations that gave rise to them. Notable among these is John Carr, who hypothesized that all forms of psychopathology, including the folk illnesses, represent *culturally authorized final common behavioral pathways*. These are behavioral responses, part of a limited number of behavioral repertoires that individuals use in response to stressful conditions that are common to humanity. Although significantly shaped by culture, these pathways are mediated by universal mechanisms of learning and cognition. To elucidate the relationship between *amok* and depression, for example, Carr and his colleague Peter Vitaliano contended that these alternate expressions of distress are culturally determined variants of response to comparable life stress conditions. The critical question, in their view, is why some people respond with aggression, some with depression, and still others cope adaptively to similar aversive events. Culture-specific and idiosyncratic responses to stress result from complex interactions between physiological processes, predisposing personality styles, cognitive processes, and learned problem-solving skills.

Perspectives from Clinically Applied Medical Anthropology

Ethnographic studies of the folk illnesses have approached them as

indigenous expressions of distress found within specific cultural contexts. From the anthropological perspective, these expressions represent local patterns of distress that reference common preoccupations or concerns of particular cultural groups. However, according to this view, all expressions of distress, including those incorporated into the DSM and ICD systems as diagnostic entities, like major depression or schizophrenia, are culturally shaped, and are thus equally "culture-bound." This specificity does not negate the role of biology or personal biography in the formation of any of these concepts, much as universal biological parameters constrain the features of all languages, yet cultures produce multiple variants of human speech.

Medical anthropologists and sociologists, such as Byron Good, Mary Jo Delvecchio Good, and Margaret Lock, have noted that cultures may organize illness expressions into nosologies not only on the basis of descriptive, symptom-based approaches, but also according to diverse cultural logics that are rooted in distinct ontologies (e.g., the reality of spirit-mediated illness). Many cultures use symbolic methods of classification, including the literate medical traditions of the ancient civilizations, in which illnesses are often categorized by analogy to various aspects of the cosmos, social relations, objects in nature, and other contextual factors. The Goods noted that current Western nosological thinking is predicated on an empiricist theory of language, which posits the existence of a one-to-one correlation between *words* and *things in the world*, with the notion that the latter will remain fixed and unchanging despite evolving social situations. Other cultures and times have considered instead that things in the world are mutable and depend for their meaning on their associations with the social, environmental, and even supernatural context, and that human categorizations should reflect this fundamental embeddedness.

Ethnographic work reveals that cultural concepts of distress identify socially constructed patterns of interpretation of suffering that serve pragmatic communicative functions. Clarifying these functions requires what Good and Delvecchio Good called a "meaning-centered medical anthropology" that investigates illness forms from the perspective of *semantic networks*, the clusters of experiences, words, and interpretations that give rise to particular illness forms as modes of experience and expression. These meaning networks are called "semantic" because they reveal the role of interpretation and communication in shaping reality and may include "personal trauma, life stresses, fears and expectations about the illness . . . the metaphors associated with a disease, the ethnomedical theories, the basic values and conceptual forms, and the care patterns that shape the experience of the illness and the social reactions to the sufferer in a given society." By taking this contextual approach, the basic organizational principle underlying a particular illness form can be

established in order to assess the degree to which it is descriptive, etiological, or expressive of particular concerns.

The anthropologist and psychiatrist Arthur Kleinman approached the meaning-centered approach through the concept of *explanatory model*. This concept refers to the more-or-less systematic explanations of an illness among different actors in any given illness episode (e.g., patient, relative, and caregiver). Kleinman showed how the explanatory model approach could be incorporated into clinical assessment through the use of practical questions about basic aspects of illness experience, addressing perceived etiology, time and mode of onset of symptoms, pathophysiology, course of sickness, and treatment that can be used by clinicians and investigators to clarify the interaction of cultural, physiological, and psychological elements in an illness episode. The Explanatory Model approach was operationalized by Mitchell G. Weiss in the form of a semistructured instrument for clinical ethnographic research, called the Explanatory Model Interview Catalogue, that can be adapted for research on any illness form.

In later work Kleinman deepened this approach to explore the ways in which the intersubjective processes of interpersonal agency—communication, engagement, transaction, and negotiation—create the flow of *social experience* that links the body-self with collective meanings and social institutions. Individuals' transact their lives, including basic processes of health and illness, within relational networks of social experience, which provides "the medium through which moral commitments and negotiations join the corporeal with the collective." For Kleinman, all of life is lived in "local worlds" that constitute the moral contexts of illness experience. The importance of local categories of suffering then lies, not primarily in identifying discrete syndromes, but in understanding what is "at stake" for individual patients, families and communities in terms of power, prestige, material resources, ethnic identity, social order, or survival. This intersection of vital interests and values was well illustrated in Kleinman's studies of neurasthenia in China, which revealed the importance of relational networks at the levels of supernatural and natural worlds, human interactions, bodily processes, and politics. In Chinese thought, each level affects the functioning of all the other levels. This fundamental interdependence can be seen in the existence of illnesses whose pathology is at the network level, such as the Chinese concept of *shenjing shuairuo,* "weakness of the nervous system." In this cultural syndrome, the internal nervous system (*shenjing*) has become devitalized (weak or run down, *shuairuo*) in correspondence with a demoralized network of interpersonal and societal relations. The symptoms of *shenjing shuairuo* can thus be understood as embodied signs of a devitalized world, both internally and externally, which are linked by contextually located emotions.

Ethnographic fieldwork by Allan Young, Laurence Kirmayer, and their

colleagues has shown the complexity of types of knowledge and modes of narration embedded in everyday explanatory accounts of illness. Individuals may link symptoms to explanations based on *chain complexes* because they were closely connected in time and space (contiguity) or as a result of logical shortcuts to denote complex experiences (e.g., metonymy), such as using the term "nerves" to refer to the anatomical and functional aspects of the nervous system, including the feelings and sensations it mediates, much as journalists use the term "Washington" to refer to the entire US government. People also commonly use *illness prototypes*—illness episodes, experienced previously by the person or by others, that provide salient examples used to reason analogically about one's own experience. In effect, patients' illness narratives may combine models common to a cultural group, idiosyncratic assemblies of possibly disconnected events, and assumptions based on past experiences. These approaches were operationalized by Danielle Groleau in collaboration with Young and Kirmayer to develop the McGill Illness Narrative Interview, a semistructured interview designed to obtain illness narratives based on (a) contiguity of events, (b) cultural prototypes of the illness for the patient, and (c) explanatory models of illness.

Atwood Gaines's concept of *sickness history* was another important contribution to the anthropological understanding of concepts of distress. Gaines's constructivist argument held that folk and professional notions of sickness evolve historically, in association with other aspects of culture, and that individuals' understanding and experience of illness are patterned on this accumulated social knowledge. In this way, cultural concepts of distress may reflect distinct cultural histories of societies. Gaines presented the examples of *spasmophilie* (spasmophilia), *fatigué* (fatigued), and *triste/fatigué tout le temps* (sad/fatigued all the time) as sicknesses that are considered physical diseases in France but are unknown in the United States, and would actually be considered nonpathological forms of dysphoria if they came to the attention of American nosologists. Gaines showed how these concepts emerged from a common sickness history that evolved from medieval understandings of dejection as a sin (acedia) to a positive valuation of existential world-weariness as a sign of maturity and a possible and accepted etiology of physical disease characterized by various states of fatigue, bodily aches, and irritability, among other symptoms.

Much recent work has pursued similar constructivist and historical approaches to both local constructs and the diagnostic entities of the DSM. Combining anthropological and clinical methods, Roland Littlewood and Maurice Lipsedge traced the historical contexts of constructs like hysteria, medication overdosing, agoraphobia, anorexia nervosa, and shoplifting and linked their phenomenologies and demographic distribution to culturally patterned gender and class constraints prevalent in the United States and United Kingdom in the 19th and 20th centuries. They described these

syndromes or idioms as deeply embodied expressions of societal conflicts about which the participants were frequently only marginally aware.

A constructivist approach to the ethnopsychologies underlying culturally specific expressions of illness and distress can be elaborated in terms of specific pragmatic, linguistic, communicative, and interactional process, including (a) the handling of the gap between experience and expression; (b) the labeling of deviant behavior as voluntary or accidental; and (c) the interpretation of symptoms as symbols or meaningless events. This conceptual grid helps explain why in North America it is possible to privilege "psychologized" understandings of illness, since these arise out of local commitments to individual autonomy, expressiveness, and personal goals. Commitments to other values in other cultural settings may produce different ethnopsychologies.

RELATIONSHIP TO PSYCHIATRIC CLASSIFICATION

At present, most investigators consider cultural concepts of distress to be *orthogonal* to psychiatric diagnoses; that is, the cultural concepts comprise alternate systems of classification that describe in distinct ways the same realm of suffering as professional nosologies. This means that, at the level of global categories, the cultural concepts cannot be easily subsumed into existing psychiatric disorders. Instead, each cultural concept exhibits a *degree of overlap* with a group of Western diagnoses, which can be represented by percentages or by a series of correlations, as with any set of comorbid relationships within professional categories.

Rather than one-to-one, the correspondence between diagnoses and cultural concepts tends to be one-to-many in either direction. Symptoms or behaviors that might be sorted by nosologies like ICD or DSM into several disorders—including across distinct diagnostic clusters—may be included in a single cultural concept, and diverse presentations that might be classified by Western nosologies as variants of a single disorder may be sorted into several distinct concepts by an indigenous diagnostic system. Therefore, cases identified by a single indigenous label often cut across diagnostic clusters. For example, the symptoms of *amok* in Malayo-Indonesia exhibit diverse combinations of anxiety, depressive, somatic, dissociative, psychotic, and characterological symptoms. The obverse is also true; homogeneous psychiatric cohorts will appear locally diverse. In Latin America, for example, different cases of major depression may be labeled as distinct conditions according to locally salient features of phenomenology, causation, context, or the healing ideology of the practitioner. These include cases of *padecer de los nervios* (suffering from nerves), *susto, fallo mental* (mental failure), *cólera* (rage), *facultades* (faculties, for *espiritistas*), and demonic influence (for Pentecostals), among other categories. Cross-cultural variation in the assembling of the "meta-structure" of psychiatric disorders—the fact that different nosologies

distribute individual illness categories into diagnostic groupings in distinct ways—as well as the changes over time in this classification process (e.g., the reassigning of posttraumatic stress disorder (PTSD) away from the anxiety disorders and into a new category of trauma and stressor-related disorders) reveals the cultural underpinnings and social evolution of any classification system.

The lack of one-to-one correspondence between diagnoses and cultural concepts is intensified by the substantial phenomenological variation that exists even across individuals who endorse the same cultural concept of distress. A given instance of *ataque de nervios*, for example, may present as an acute dissociative episode, with seizure-like symptoms in a state akin to trance, for which amnesia is later reported. Other episodes may resemble cases of panic disorder, albeit with more suicidal and aggressive features, such as breaking things and verbally or physically attacking those around them. Some episodes may present as uncontrollable fits of nervous laughter, which evolve into bouts of crying and screaming and subsequent depression. Others may constitute episodes of acute grief at the funeral of a loved one without any pathological sequelae. What links all four instances in the popular classification system in the Hispanic Caribbean is their presentation as acute episodes of uncontrolled emotionality, precipitated by interpersonal stressors that overwhelm the person's coping ability, usually in a context of insufficient social and familial support. Yet from the psychiatric perspective, these presentations would be considered phenomenologically distinct. Clinical translation between the cultural concepts and psychiatric diagnoses must therefore be conducted at the level of the *individual*, describing for that particular person the overlap between both systems of classification.

Two other differences ensure that cultural concepts of distress and contemporary psychiatric nosologies overlap only partially as global systems. First, Western nosologies are based at present nearly exclusively on formal, descriptive definitions of psychopathology. Most indigenous nosologies, however, distinguish pathology from normality at least as much on the basis of contextual characteristics and causal theories as of descriptive ones. These include assessments regarding the appropriateness of the symptom in the specific setting and time in question, the relative sufficiency of the precipitating stressors, and the nature and quality of the human, and sometimes spiritual, relationships of the sufferer.

Second, while psychiatric nosologies contain only pathological conditions, cultural concepts of distress may apply to a wide range of severity, including presentations that do not meet DSM criteria for any mental disorder. As described in DSM-5 (p. 758), "an individual with acute grief or a social predicament may use the same idiom of distress or display the same cultural syndrome as another individual with more severe psychopathology." The indigenous classification systems thus tend to

describe broader realms of behavior and adaptation than the narrower purview of psychiatric classification.

Despite the differences between local and professional nosologies, information on cultural concepts of distress is vital to clinical care for at least four reasons, which are outlined in DSM-5. First, clinician awareness of cultural concepts may prevent misdiagnosis, as cultural variation in symptoms and explanatory models associated with these concepts may lead clinicians to misjudge the severity or a problem or assign the wrong diagnosis (e.g., unfamiliar spiritual explanations may be misunderstood as psychosis). Second, cultural variations in symptoms and attributions related to the cultural concepts can be associated with particular features of risk, resilience, and outcome. For example, the interpretation of trauma-related symptoms as spiritual afflictions may prompt the use of effective forms of healing and self-coping, such as church-based social support or meditation. Conversely, the presence of a cultural syndrome may be associated with specific types of risk and morbidity—for instance, endorsement of *ataque de nervios* is related to higher suicidal ideation after adjusting for psychiatric diagnosis, demographic characteristics, and trauma exposure. Third, understanding the messages being communicated by the patient in the idiom of the cultural concept (e.g., social isolation, being overwhelmed) can improve clinical rapport and engagement and help patients and clinicians select a mutually acceptable treatment plan. Some patients come to care expecting to be treated for the cultural concept of distress and may become disengaged if their understanding is not negotiated effectively with that of the clinician. Fourth, cultural concepts can illustrate key psychological mechanisms, which need to be understood and addressed to improve clinical efficacy. For example, catastrophic cognitions can contribute to symptom escalation into panic attacks in culturally specific ways, such as through the cultural salience of certain symptom-outcome pairings. The popular linkage made in the United States between chest pain and heart attacks can precipitate panic episodes when chest pain occurs as part of an anxiety reaction but is interpreted as a harbinger of a myocardial infarction. In Cambodia, by contrast, the association is more commonly between neck soreness and stroke, due to the interpretation of panic attacks as signs of disordered *khyâl* circulation, a wind-like substance important in local ethnophysiology. Being aware of these linkages can help clinicians identify how cultural interpretations can constitute the actual "cogs" of the disorder.

Attending to these aspects of the relationship between cultural concepts and psychiatric diagnoses can help clinicians meet patients' treatment expectations. Moreover, inclusion of the cultural concepts in research can reveal patterns of comorbidity—between the diagnoses linked by the idiom or cultural syndrome—and clarify the epidemiological distribution of cultural features of illness. Information on the cultural concept of distress

can be incorporated into the process of diagnostic assessment and treatment planning by eliciting an account of the person's definitions of pathology, idioms of distress, unique symptoms, contextual embeddedness, and other associated factors that are contained in the experience of the cultural concept. A systematic approach for eliciting and documenting this information is the Cultural Formulation Interview, which was included in DSM-5 and will be described later in this chapter.

EXAMPLES OF CULTURAL CONCEPTS OF DISTRESS

Only a few of the thousands of cultural concepts of distress have received sustained research attention integrating cultural and psychiatric research methods. This chapter focuses on three concepts from diverse cultural regions that have been the object of intensive research, including studies of their relationship with psychiatric diagnoses: *ataques de nervios*, possession syndrome, and *shenjing shuairuo*. Other examples of cultural concepts of distress are included in the DSM-5 Glossary. These three examples and the few cultural concepts listed in the DSM-5 Glossary are meant to be illustrative of the wide gamut of folk illness categories relevant in particular populations or contexts.

Ataque de Nervios

Ataque de nervios ("attack of nerves" in Spanish) is a syndrome indigenous to various Latin American cultures, notably those of the Hispanic Caribbean (Puerto Rico, Cuba, and the Dominican Republic). It has received considerable attention in the psychiatric and anthropological literature since the mid-1950s, mostly in Puerto Rican communities in the Island and in the United States. Early reports of presentations in male recruits to the US armed forces by military psychiatrists in Puerto Rico mislabeled *ataque* as "the Puerto Rican Syndrome." In keeping with psychoanalytic theory at the time, these investigators attributed the *ataque* to specific indigenous child-rearing practices allegedly predisposing to the hysterical discharge of aggressive drives. The power relationships embodied in the military, and the colonial situation represented by the presence of the US armed forces in Puerto Rico, were usually left unanalyzed in these discussions.

Subsequent research focused on the role of *ataque* as a culturally recognized and sanctioned expression of emotion, mostly by women, that should be understood as a form of communication about family relationships in contexts of social and economic powerlessness. These investigations located the psychosocial etiology of *ataques* in experiences of acute family discord and individual suffering—such as grief, fear, and anger—in the context of collective distress—such as migration, poverty, and discrimination. This research emphasizes the role of *ataque* as an idiom of

distress in Latino communities and illustrates how the same folk term may represent different kinds of cultural concepts—that is, a folk category such as *ataque* is a cultural syndrome that functions as an idiom of distress.

Contemporary clinical–anthropological studies have further clarified the phenomenology, cultural meanings, epidemiology, and relationship of *ataque* to psychiatric diagnoses. The work of Peter Guarnaccia and Roberto Lewis-Fernández has identified loss of control of emotions, behaviors, and a social world out of control as the core experiences of *ataques de nervios*. Their work suggests a popular nosology based on various categories of *nervios* (nerves), several of which require more systematic research.

Phenomenology. Guarnaccia, Glorisa Canino, and colleagues surveyed a representative community sample of Puerto Rico residents (N = 912) for the presence of *ataque de nervios*. Qualitative coding of general descriptions of the syndrome among a subsample of 145 respondents in this larger study—of whom 77 were *ataque* sufferers—yielded four domains of *ataque* experience. These were emotional expressions, bodily sensations, action dimensions, and alterations in consciousness. On the basis of these data, a prototypical *ataque de nervios* can be described as composed of the elements described in Table 29.1–2.

Epidemiology. A 1987 survey by Guarnaccia and Canino in a probability sample of the population of Puerto Rico (N = 912), using the Diagnostic Interview Schedule (DIS), found a weighted lifetime prevalence of *ataque* of 13.8 percent, making it one of the most frequently reported syndromes in the Island. Twelve percent of respondents with *ataque* met DIS criteria for a positive psychological syndrome—that is, the *ataque* met severity criteria (consulting a physician or other professional, taking medications, or reporting some functional impairment) and could not be explained as resulting from physical illness or substance use. The Latino moiety of the 2002 to 2003 National Latino and Asian American Survey, led by Margarita Alegría, also obtained lifetime *ataque* prevalence in a nationally representative sample of US Latinos using a diagnostic module that followed the format of the Composite International Diagnostic Interview. *Ataque* prevalence was 10.9 percent in Puerto Ricans, 6.2 percent in Cubans, 6.0 percent in Mexicans, and 5.4 percent in Other Latinos, suggesting both that *ataque* is common across Latino subgroups and that it is more prevalent among Puerto Ricans.

Table 29.1–2.
Phenomenology of *Ataque de Nervios* (Attack of Nerves)

a. Exposure to a frequently sudden, stressful stimulus, typically eliciting feelings of fear, grief, or anger, and involving a person close to the subject, such as a spouse, child, family member, or friend. Severity of the trigger ranges from mild-to-moderate (i.e., marital argument and disclosure of

migration plans) to extreme (i.e., physical or sexual abuse and acute bereavement).

b. Initiation of the episode immediately upon exposure to the stimulus or after a period of brooding or emotional "shock."
c. Once the acute attack begins, rapid evolution of an intense affective storm characterized by a primary affect usually congruent with the stimulus (e.g., anger, fear, and grief) and a sense of loss of control (*emotional expressions*).
d. These are accompanied by all or some of the following:
    i. *Bodily sensations*: trembling, chest tightness, headache, difficulty breathing, heart palpitations, heat in the chest, paresthesias of diverse location, difficulty moving limbs, faintness, blurred vision, or dizziness (*mareos*).
    ii. Behaviors (*action dimension*): shouting, crying, swearing, moaning, breaking objects, striking out at others or at self, attempting to harm self with nearest implement, falling to the ground, shaking with convulsive movements, or lying "as if dead."
e. Cessation may be abrupt or gradual, but it is usually rapid, and often results from the ministration of others, involving expressions of concern, prayers, or use of rubbing alcohol (*alcoholado*). There is return of ordinary consciousness and reported exhaustion.
f. The attack is frequently followed by partial or total amnesia for the events of the episode, and descriptions of the following for the acute attack: loss of consciousness, depersonalization, mind going blank, and/or general unawareness of surroundings (*alterations in consciousness*). However, some *ataques* appear to exhibit no alterations in consciousness.

Guarnaccia and Canino found a prevalence of *ataques de nervios* of 8.9 percent in a representative sample of children in Puerto Rico ages 4 to 17 (N = 1,897). A second, community-based probabilistic sample of Puerto Rican children ages 5 to 13 in San Juan (N = 1,353) and the South Bronx (N = 1,138), led by Héctor Bird and Canino, found a lifetime prevalence of *ataque* of 5.4 and 4.3 percent, respectively.

Risk factors for *ataque* span a range of social and demographic characteristics. Adults reporting an *ataque de nervios* in Puerto Rico are more likely to be female, over the age of 45, with less than a high school education, formerly married (i.e., divorced, widowed, or separated), and out of the labor force. However, 10 percent of the men in the sample reported an *ataque de nervios*, which indicates that some men express their distress through this cultural syndrome. Among Latinos in the United States—and like other psychiatric disorders—*ataque* is positively correlated with higher acculturation (e.g., US birth and English proficiency), suggesting higher psychiatric vulnerability in this post-migration population. Risk factors among Puerto Rican youth differ somewhat by site (New York vs. Puerto Rico) and by study; they include older age (especially adolescence in girls), single-parent household status, exposure to violence, stressful life events in the past year, and higher perceived poverty. In one of the community-based studies in Puerto Rico, multiple regression analysis also revealed an independent positive association between *ataque* status in the child and a family history of the syndrome, suggesting the possible influence of intergenerational behavioral modeling.

Lewis-Fernández and colleagues estimated the lifetime prevalence of *ataque de nervios* in two clinical settings serving adult Puerto Rican

psychiatric outpatients. The prevalence rate of *ataque* was 55.1 percent among immigrant Puerto Ricans (*N* = 89) in an outpatient Latino mental health clinic in New England and 51.5 percent among outpatient referrals to psychiatry (*N* = 97) in three rural Puerto Rico primary care clinics. In both sites, the gender distribution of *ataque* sufferers was nearly equivalent: 56.3 percent female in New England and 54.8 percent female in Puerto Rico. However, women were twice as likely to seek mental health care in both settings. Finally, in these rural clinics for low-income residents in Puerto Rico, *ataque* rates also did not differ by age cohort or level of education. All social groupings among this disenfranchised sector of the population appeared equally at risk for *ataque* episodes. A separate opportunity sample of 303 Puerto Rican primary care patients age 50 and older in the northeastern United States found a lifetime *ataque* prevalence of 26 percent, confirming the elevated prevalence of the syndrome in clinical populations. An elevated clinical prevalence was also observed among children: lifetime prevalence of *ataque* was 25.5 percent in a representative sample of children utilizing the mental health system in Puerto Rico (*N* = 767).

Precipitants/Etiology. Prototypically, *ataque de nervios* was linked by sufferers to an acute precipitating event or to the summation of many life episodes of suffering brought to a head by a trigger that overwhelmed the person's coping ability. In the Puerto Rico community sample studied by Guarnaccia and Canino, the first experience of an *ataque* was closely tied to a precipitating event. In 92 percent of cases, the *ataque* was directly provoked by a distressing situation, and 73 percent of the time it began within minutes or hours of the event. A majority of first *ataques* (81 percent) occurred in the presence of others, as opposed to when the sufferer was alone, and led to the person receiving help (67 percent). Finally, unlike the typical experience of persons with panic disorder, most respondents reported feeling better (71 percent) or feeling relieved (81 percent) after their first *ataque*. These findings suggest that first episodes of *ataque de nervios* are closely tied to the interpersonal world of the sufferer and that they result in an unburdening (*desahogarse*) of one's life problems, at least temporarily. One function of the *ataque* may be to rally the person's social network, including reversing the current interpersonal arrangement in favor of the sufferer.

Specific precipitants were varied, but typically consisted of acutely stressful experiences resulting in feelings of loss, grief, anger, or fear. Generally, the most frequent triggers identified by the Puerto Rico epidemiological study included arguments with a spouse or other close relatives, drunkenness of a family member, an accident that threatened the life of a person close to the subject, death of a close family member, natural disaster, or the passing of a fright (*susto*) through a person. Certain

precipitants were obviously traumatic and were correlated with globally elevated prevalence rates in exposed subpopulations. For example, this survey showed a higher prevalence of *ataques* among those severely affected by a recent natural disaster (19 percent) than among nonexposed individuals (14 percent), and intermediate rates among subjects only moderately affected (16 percent). However, *ataques* have been reported after apparently minor stressors, such as family squabbles, adolescent romantic disappointments, or peer razzing. In these cases, the presence of underlying interpersonal or sociocultural conflicts of longer duration—such as marital argument or abuse, migration, poverty, discrimination, precarious employment/source of income, or family discord—has been invoked by investigators to explain the enhanced *ataque*-producing power of superficially mild triggering experiences.

Study in psychiatric outpatient samples of the relationship between childhood trauma exposure and presence of *ataques* has yielded conflicting results. Daniel Schechter and colleagues found a positive association, whereas Lewis-Fernández and colleagues did not. In this latter study, childhood traumatic exposure was high across Puerto Rican psychiatric outpatients regardless of *ataque* status, suggesting that, in a clinical sample, factors in addition to childhood trauma accounted for this condition. In a different psychiatric outpatient sample of diverse Latino subethnicities with a history of traumatic exposure (N = 202), greater trauma severity was not associated with higher frequency of *ataques*. More research is needed in samples with and without traumatic exposure to clarify the association between this risk factor and *ataque*.

Relationship to Psychiatric Diagnoses. Psychiatric evaluations of *ataque* sufferers have usually revealed the presence of some form of psychopathology. In the Puerto Rico adult epidemiological survey, 63 percent of respondents who reported an *ataque de nervios* met diagnostic criteria for one or more psychiatric disorders measured in the DIS. This suggests that rather than being an expression of nonpathological distress, most *ataques* are associated with a psychiatric disorder. Overall, *ataque* sufferers had 4.35 times the odds of meeting criteria for a psychiatric diagnosis than those who did not report the cultural syndrome. At the same time, 37 percent of *ataque* sufferers did not fulfill DIS criteria. This suggests that a substantial minority of respondents with this syndrome are describing milder levels of distress using the idiom of the cultural syndrome; alternatively, they may be reporting psychiatric disorders not assessed by the study instrument.

It is important to note that *ataques* do not correlate exclusively with one psychiatric diagnosis and cannot be treated as simply a culturally shaped version of a specific psychiatric disorder. The combined research findings support a strong relationship between *ataques de nervios* and

mood and anxiety disorders in children and adults, as well as disruptive disorders in children and dissociative disorders in adults. Of particular concern is the strong relationship between *ataque* and suicidal ideation (odds ratio [OR] = 6.2) and attempts (OR = 8.1). In a US nationally representative community sample of diverse Latino subethnicities (*N* = 2,554), lifetime history of *ataque de nervios* was independently associated with suicidal ideation (OR = 2.4) after adjusting for multiple covariates, suggesting that the cultural syndrome conveys additional risk beyond recognized risk factors such as psychiatric diagnosis and traumatic exposure. During *ataque* episodes, sufferers may make goal-directed or ill-formed attempts at self-harm, increasing the morbidity (and occasional mortality) of the syndrome.

While early speculation by psychiatrists focused on whether *ataques* constitute a culturally shaped version of a panic attack, Lewis-Fernández and colleagues systematically demonstrated that only 36 percent of best-remembered *ataques* met DSM-IV-TR criteria for panic attacks and 17 percent met criteria for panic disorder. What mainly distinguished *ataques* that did not fulfill panic disorder criteria from those that did included the cuing of episode by precipitants, symptom crescendo longer than 10 minutes, and feeling relieved rather than frightened after an episode. However, features that were common to both *ataque* and panic disorder included recurrence of attacks, the experience of fear or nervousness during an episode, and anticipatory or behavioral sequelae afterwards. Patients who described feeling relieved after their *ataque* also reported significantly fewer panic symptoms than those without this feeling.

Specific characteristics of *ataque* profiles may be associated with distinct psychiatric diagnoses. A study by Ester Salmán and Michael Liebowitz of a sample of Puerto Ricans and Dominicans (*N* = 156) attending a Manhattan research clinic that was enriched with individuals suffering from *ataque de nervios* (*N* = 109) found a link between *ataque* features and specific psychiatric disorders. Among those participants in the study whose *ataques de nervios* were characterized by intense fearfulness, and feelings of asphyxia and chest tightness, diagnoses of panic disorder were more common. Those whose *ataques* were more dominated by the emotion of anger and aggressive behavior, such as breaking things, were more likely to meet criteria for co-occurring mood disorders.

Ataque is strongly associated with dissociative experiences. Severity of *ataques*, assessed as lifetime number of episodes, is positively and independently related to self-reported dissociative symptoms, measured with the Dissociative Experiences Scale (DES), and to clinician-diagnosed dissociative disorders, assessed with the Structured Clinical Interview for DSM-IV. Further evidence of the relationship between *ataque de nervios* and dissociation comes from the significant correlation between *ataque* and other Latino cultural concepts that are also related to higher

dissociative capacity on the DES. One of these idioms is characterized by "pseudo-hallucinations" in several sensory modalities—hearing voices or noises, seeing shadows (*celajes*), and feeling presences when alone—that are popularly attributed to spiritual influences. In an outpatient US Latino sample with high traumatic exposure (N = 202), *ataque de nervios*, pseudo-hallucinations, and a third Latino folk illness, "being currently ill with nerves," were all significantly associated with each other and with elevated DES scores, forming a nexus of "nerves"-related cultural concepts linked to dissociative capacity.

In children, *ataques de nervios* are associated with research diagnoses of mood, anxiety, and disruptive disorders. In a Puerto Rico community-based survey, children ages 4 to 17 with *ataques* had ORs of 7.6 for any mood disorder, 5.0 for any anxiety disorder, and 3.1 for any disruptive disorder, compared to children who did not report an *ataque*. In the clinical sample of the same study, the equivalent ORs were 5.8, 3.5, and 2.0, respectively. Children with *ataques de nervios* had 4.3 times the odds of meeting criteria for any psychiatric diagnosis in the community and 2.3 times the odds of doing so in the clinical setting. Moreover, impairment was likewise elevated among children with *ataques* compared to those without, with an OR = 5.6 in the community and OR = 2.1 in the clinical sample. In another study, after adjusting for demographic characteristics and the level of psychosocial stressors, children ages 5 to 13 with *ataques* in San Juan and the South Bronx had nearly 3 times higher odds of meeting criteria for an anxiety or a depressive disorder, nearly 2 times the odds of a disruptive disorder, and 1.4 times higher odds of global impairment.

Additional Clinical Features. Assessment of the phenomenological features of *ataques de nervios* experienced by the 77 respondents in the Puerto Rico epidemiological study with at least one lifetime *ataque* who were reinterviewed by Guarnaccia revealed that the largest number (28 percent) had suffered only one such episode over the course of their life. Thirteen percent had 2 *ataques*, 8 percent experienced 3, 9 percent had 4, and 4 percent had 5. The second largest number (23 percent) comprised those who had more than 5 lifetime *ataques*, while 15 percent were unsure about the total. The bimodal distribution of *ataque* frequency (one vs. six or more) suggests that persons with frequent episodes constitute a minority of *ataque* sufferers. This raises the question whether having an isolated *ataque* in response to a markedly stressful event—as opposed to recurrent *ataques*—represents a relatively normal expression of distress in Latino communities. However, the lifetime presence of one or more *ataques* is associated with disability and outpatient mental health service use among US Latinos, even after adjusting for demographic characteristics, traumatic exposure, psychiatric diagnoses, and chronic health conditions. This suggests that experiencing an overwhelming reaction such as *ataque* may

not be without clinical consequences.

The association between *ataque* and a sense of loss of control and of being overwhelmed highlights the importance of the relationship between the cultural syndrome and other behaviors associated with acute emotional dysregulation. Most concerning among these is the strong association noted earlier between *ataques* and suicidal ideation and attempts. Other behaviors related to this acute dysregulation include loss of aggression control, expressed as attacks on people or property, and dissociative experiences, both of which are related to the acute *ataque* experience.

Specific Cultural Factors. The complex relationship between *ataque de nervios* and psychiatric diagnosis may be further clarified by placing *ataque* within a broader Latin American popular nosology of *nervios* (nerves) of which it is part. Experiences of adversity are linked in this nosology to ensuing "alterations" of the nervous system, which result in its impaired functioning, including the peripheral nerves. This quasi-anatomical damage is evidenced in emotional symptoms, including interpersonal susceptibility, anxiety, and irritability, as well as in physically mediated symptoms, such as trembling, palpitations, and decreased concentration. Certain people are considered at greater risk for *nervios* when confronted with adversity, such as those with an inherited predisposition or with stressful perinatal experiences (e.g., mother's alcoholism). But in all cases, the possibility that sufficient adversity will result in an alteration of the nervous system leads to the popular view that persons faced with experiences of suffering should "control" their reactions to avoid nervous system damage. This usually takes the form of dampening the emotional impact of the experience, so that the nerves will not be substantially "altered."

Clinical and popular discussions about *nervios* and *ataques* in Latino communities thus typically reference cultural views about the importance of remaining in control, and of facing adversity with *tranquilidad* (equanimity) regardless of the severity of a stressor. *Ataque de nervios* can therefore be understood as a behavioral syndrome stemming from general notions and practices about how to behave, rather than a peculiar cultural expression of a particular diagnosis. As a marker of being overwhelmed, *ataques* are associated with various forms of psychopathology. But the *ataque* itself cannot be reduced to just the psychiatric disorder; rather it is a marker of the severity of the stressor, of the person's reaction style, of his/her level of support, and of the resulting distress and pathology.

Aspects of Caribbean Latino culture may promote the observed association between *ataque de nervios* and dissociation, since the appearance of dissociation during the *ataque* may be enhanced by the cultural perception that a mature social actor needs to be always in control. Dissociation may be particularly promoted by the popular view that when

out of control the person is no longer "him or herself" ("*ése no era yo* [that was not me]"). The uncontrolled behavior is disavowed as not being part of the self. This may facilitate the emergence of dissociative splits that allows the person to distance him/herself from the moral consequences of the *ataque* episode. Similarly, prevalent cultural concerns about the psychological and social dangerousness of strong emotions may reinforce the role of anxiety sensitivity in *ataque*.

Community rates of *ataque* in adults are significantly higher in women than in men. The female-to-male ratio is 2:1 in Puerto Rico and 3:2 among US Latinos. The gender difference may appear during adolescence, since it is absent in both South Bronx and San Juan-based representative samples of Puerto Rican children ages 5 to 13 but is found in a Puerto Rico-wide representative sample of youth ages 4 to 17. The gendered nature of *ataque* in adults has been attributed in part to traditional gender roles pressuring Latino women to restrict some emotions in social situations, especially around sexual, aggressive, or assertive needs. The gender-specific cultural directive to remain "in control" may predispose women to express suppressed feelings in situations of interpersonal conflict as acute paroxysms of emotionality and uncontrolled behavior, which can later be disavowed as unwilled or even ego-alien.

Treatment. No treatment studies of *ataque de nervios* have been conducted. Typical treatment in community settings involves, first, ensuring the safety of the person and those around him/her, given the association between *ataque*, suicidality, and uncontrolled aggressivity. "Talking the person down" is usually helpful, accompanied by expressions of support from relatives and other loved ones; the use of rubbing alcohol (*alcoholado*) to help calm the person is a culturally prescribed way of expressing this support.

"Telling the story" of what led to the *ataque* usually constitutes the principal therapeutic approach in subsequent stages of treatment. Since one of the main functions of the attack is to communicate a feeling of being overwhelmed, indicating receipt of the message and the desire to offer support are usually perceived as therapeutic. The person should be allowed to set the pace of disclosure and to give enough details and circumstances to feel "unburdened" (*desahogado[a]*).

In the case of single or occasional *ataques* in the absence of a psychiatric diagnosis, brief follow-up is usually sufficient. This may be discussed with the patient and the family as a way of ensuring a full return to the previous healthy state. For recurrent *ataques*, treatment depends on the associated psychopathology, the nature of the precipitants—including traumatic exposure—the degree of family conflict or support, the social context, the previous treatment experiences, and the patients' and family's expectations, among other factors.

Psychotherapy is typically the mainstay of treatment, usually with a family focus, given the usual source of the overwhelmed behavior in the interpersonal milieu. A social activism stance by the therapist that acknowledges the origins of adversity among low-income Latinos in socioeconomic disenfranchisement and ethnic/racial discrimination is helpful in addressing the sociocultural roots of *ataque de nervios*. Pharmacotherapy may also be useful in the treatment of *ataque*-related psychopathology; primary emphasis should be placed on treating the underlying disorder. Given the slow crescendo of many *ataques*, judicious use of short-acting benzodiazepines also has a role in helping abort an impending episode. However, this should not be the main form of treatment for recurrent *ataques*, since it only forestalls the principal function of the syndrome as a mode of communication.

Possession Syndrome

Involuntary possession trance states are very common presentations of emotional distress around the world. Cognate experiences have been reported in extremely diverse cultural settings, including India, Sri Lanka, China, Japan, Malaysia, Niger, Uganda, Southern Africa, Turkey, the United States, Canada, Haiti, Puerto Rico, and Brazil, among others. *Possession syndrome* is an umbrella English-language term used in the clinical and anthropological literature on South Asian presentations of involuntary possession trance that encompasses multiple names in regional languages and dialects of India and Sri Lanka. These presentations are seen as a severe affliction by the person's cultural group because they are involuntary, cause distress, and do not occur as a normal part of a broadly accepted collective cultural or religious ritual or performance.

DSM-IV-TR did not list possession syndrome or any other pathological possession trance condition as a "culture-bound syndrome." Instead, this class of presentations was subsumed in Appendix B of DSM-IV and DSM-IV-TR under the proposed category of Dissociative Trance Disorder (DTD), one of the "Criteria Sets and Axes Provided for Further Study." Categories in this section of the manual just fell short of full inclusion in DSM-IV as specified diagnoses but were included in the appendix to stimulate research for possible incorporation in DSM-5. Beyond its global prevalence, another reason why involuntary possession trance was given a special status in DSM-IV was its inclusion in ICD-10 as a subtype of Trance and Possession Disorders (F44.3), a specified diagnosis under the Dissociative (Conversion) Disorders in the section on Neurotic, Stress-Related and Somatoform Disorders. The goal of making the ICD and DSM nosologies more compatible helped to promote consideration of involuntary possession trance in DSM-IV and DSM-5. ICD-10 describes Trance and Possession Disorders as follows:

Disorders in which there is a temporary loss of both the sense of personal identity and full awareness of the surroundings; in some instances, the individual acts as if taken over by another personality, spirit, deity, or "force." Attention and awareness may be limited to or concentrated upon only one or two aspects of the immediate environment, and there is often a limited but repeated set of movements, postures, and utterances. Only trance disorders that are involuntary or unwanted, and that intrude into ordinary activities by occurring outside (or being a prolongation of) religious or other culturally accepted situations should be included here.

In DSM-5, pathological possession trance was included under Dissociative Identity Disorder, leading to its redefinition as two variants: possession-form and non-possession-form. This incorporation was due largely to the phenomenological similarity between pathological possession trance and Dissociative Identity Disorder. Both are disorders of identity alteration that occur during a dissociative state of consciousness and are characterized by involuntary stereotyped and culturally determined behaviors or movements that are experienced as being under the control of an alternate identity. Among 119 possessed patients who visited traditional healers in Uganda, for example, Marjolein Van Duijl and colleagues observed involuntary shaking of the head or parts of the body (attributed to spirits), talking in a different voice—which others recognized as the voice of a specific spirit—and feeling influenced by unidentified forces that cause unusual behavior. Nearly half of the possessed patients mentioned something holding the body so they were unable to move or speak. Among traditional healers and local health workers in these villages, case vignettes of Dissociative Identity Disorder were always attributed to spirit possession.

In both possession and non-possession types of pathological identity alteration, sufferers can report full or partial amnesia for the events of the dissociative state. Moreover, elevated exposure to traumatic experiences is typical of both conditions. In the Uganda study, for example, individuals with pathological possession scored higher on the Harvard Trauma Questionnaire and the Traumatic Experiences Checklist than 71 randomly selected healthy inhabitants of the same villages, and scores on both trauma scales were strongly correlated with symptoms of dissociation (r = 0.64 – 0.69) and pathological possession (r = 0.51 – 0.64). To rule out the misdiagnosis of nonpathological possession states, Dissociative Identity Disorder requires the presence of amnesia, as well as the fact that the identity alteration states a) be associated with clinically significant distress or impairment (criterion C) and b) not be a normal part of a broadly accepted cultural or religious practice (criterion D). In addition, the text states that the pathological possession episodes "present recurrently [and] are unwanted and involuntary."

The possession-form and non-possession-form variants of Dissociative Identity Disorder are not mutually exclusive: individuals in one cultural setting may experience either form, and a single individual may experience both forms. Typically, in non-possession-form Dissociative Identity

Disorder, the alternate identities are experienced as *internal* aspects of the person, as other personalities or identity states. In possession-form Dissociative Identity Disorder, by contrast, the alternate identities are typically experienced as *external* possessing agents, usually of supernatural or spiritual origin (e.g., demons and spirits). The main contrast between these two forms of Dissociative Identity Disorder appears to be shaped by the patient's sociocultural milieu. That is, possession-form Dissociative Identity Disorder arises in contexts where possessing agents are considered ontologically "real" (e.g., some fundamentalist religions in the United States or traditional South Asian cultures), as opposed to non-possession-form Dissociative Identity Disorder, which draws its alternate identities from individualistically focused experiences, such as stages of life (e.g., childhood) or biographical roles in the patient's life (e.g., protector and perpetrator).

The decision to subsume pathological possession trance states under Dissociative Identity Disorder is not without controversy, however. It is an open empirical question, for example, what proportion of pathological possession states meets Dissociative Identity Disorder criteria, which include substantial amnestic gaps in addition to identity alteration. The non-possession-form of the disorder is typically a chronic condition, arising in childhood in the wake of prolonged and severe trauma. Some pathological possession trance states follow this pattern, but others are triggered by nontraumatic stressors, such as romantic disappointments and work or religious conflicts, and present as single or recurrent episodes, rather than persistent conditions. However, the absence of systematic assessment of childhood trauma precludes excluding this exposure from the genesis of dissociative disorders in these populations. With accumulating evidence, the placement of pathological possession in psychiatric nosology will be further clarified. Dissociative Identity Disorder may prove to be a heterogeneous condition, encompassing a gamut of identity alteration symptoms, courses, and precipitants across cultures. Alternatively, the evidence may favor the inclusion of pathological possession trance as a separate diagnostic entity in DSM, as in ICD-10.

Pathological possession trance states thus constitute local presentations that are farther along the process of incorporation into Western psychiatric classification systems than the usual cultural concepts of distress. Their truly global prevalence makes them candidates for "universalistic" disorders, even though they are considered uncommon in modern Euro-American and European societies. Their status is still unresolved, however, given their inclusion as a variant of Dissociative Identity Disorder in DSM-5 and as a distinct disorder in ICD-10. Moreover, since local presentations of pathological possession trance vary substantially across cultural settings, possession syndrome is discussed here as a cultural concept of distress, focusing on its specific characteristics in South Asia.

Phenomenology. It is important to distinguish at the outset between possession syndrome, as an instance of possession trance, and the broader category of possession. The latter term refers to a general ideology describing the full range of direct spirit influences on human affairs including effects on physical, psychological, spiritual, social, and ecological realms. By contrast, as a subset of general possession experience, "possession trance" refers to specific alterations in consciousness, memory, behavior, and identity attributed to direct spirit influence. In addition to pathological possession trance states, South Asian cultures authorize multiple examples of normal possession and possession trance. When voluntary and normative, these states are typically seen as instances of religious devotion, mystical ecstasy, social commentary, asceticism, interpersonal relations, existential reflection, and the study of consciousness. This chapter discusses possession syndrome as a pathological entity with an established phenomenology, that is, as a special case in the general continuum of etiological ideas regarding possession illnesses. A prototypical episode is composed of the elements in Table 29.1–3.

Data on the epidemiology, precipitants, and associated psychopathology of subjects with possession syndrome in South Asia is limited by methodological considerations. These include lack of representative community samples and nonsystematic definitions of the syndrome, which shows considerable regional variation.

Table 29.1–3.
Phenomenology of Possession Syndrome

a. Onset occurs typically due to subacute conflict or stress and shows considerable variation. It may be gradual and nonspecific—e.g., diverse somatic complaints, such as dizziness, headaches, abdominal discomfort, hot-cold flashes, listlessness, or difficulty breathing—or sudden and specific, in the form of an abrupt transition to an altered state of consciousness. Onset may also vary by geographical region, presenting in certain areas of the country (e.g., parts of Uttar Pradesh) as a mounting series of rage attacks. If presentation is gradual and nonspecific, transition to the altered state may occur spontaneously or after contact with an indigenous healer specializing in the treatment of the syndrome.
b. Behavior during the altered state consists of some or all of the following:
    i. Dramatic, semipurposeful movements—such as head bobbing, bodily shaking, thrashing, gyrating, or falling to the ground—accompanied by guttural, incoherent verbalizations such as mumbling, moaning, or shrieking.
    ii. Aggressive or violent actions directed at self or at others, including spitting, striking, and impulsive suicidal or homicidal gestures. Verbalizations may be coherent and consist of derogatory comments or threats of violence directed against significant others or against the subject (in the third person) and typically considered by observers to be uncharacteristic of the subject's usual behavior.
    iii. Specific gestures, comments, or requests denoting the appearance of a known possessing personality by (1) reference to standard attributes of culturally recognizable figures or (2) the name and degree of relation of deceased family members or acquaintances.
c. The altered state is marked by the emergence of one or several secondary personalities distinct

from that of the subject. Their specific identities, which may remain undisclosed for some time, adhere to cultural norms regulating permissible agents of possession, which vary by religion, region, and caste. Acceptable agents include spirits of deceased family members, in-law relations, or known village acquaintances who died under specific conditions of distress and minor supernatural figures of the Hindu pantheon (rarely major deities) and the Islamic spiritual world. They neither include the personalities of alternate, living persons, nor apparently alternation of coconscious multiple personalities, as in non-possession-form dissociative identity disorder, though consecutive possession by two or three agents is not rare, and one case of 21 successive possessions has been reported. On occasion, a single agent may possess several members of a family over time or in succeeding generations.

d. Possession by the secondary personality(ies) is episodic, resulting in alternation between the usual personality of the subject and the altered state. The subject in his/her usual identity appears in a daze, exhausted, distressed, or confused about the situation and may report visual or auditory perceptual disturbances regarding the possessing agent, as well as "unconsciousness" and partial or total amnesia for the altered state. States of consciousness may alternate continuously within one episode over hours to weeks, with each state lasting minutes to hours. This pattern may persist intermittently over months or years.
e. Frequently, the specific identities of possessing personalities remain undisclosed for some time, requiring the active ministrations of family members and the intervention of specialized indigenous practitioners. The process of disclosure is conceived as a struggle between the family members and the beneficent agents possessing the healer on the one side, and the troublesome possessing personalities on the other. It is characterized by remarkable reactivity on the part of the subject to environmental cues, including direct questioning, strategic neglect, and aggressive manipulation. Often, this kind of ritualized intervention elicits the withheld disclosure and leads to a process of negotiation between the family, the healer, and the possessing agent regarding the propitiatory treatment necessary for the resolution of the episode. Examples of this treatment typically consist of ritual offerings, admission to a possession cult, resolution of family conflicts, cessation of abuse, and the postponement or precipitation of marriage.
f. Outcome is variable. Total recovery is often reported at the cessation of a single acute episode, which may be of several weeks' duration. Alternatively, prolonged morbidity may result, or even, rarely, death. The subject may seek exorcism treatment with multiple healers, or undertake a pilgrimage to a regional healing temple. At times, he/she is prescribed entrance into a ritual devotion cult offering programmed periodic possessions.

Epidemiology. Period-prevalence rates among psychiatric cohorts in India conducted mostly via chart review have been variously reported as 0.2 percent over 1 year, 0.06 percent over 7 years, and 0.93 percent over 12 years. Patients with dissociative disorder accounted for 0.15 to 1.5 percent of outpatients and 0.15 to 1.2 percent of inpatients at a major psychiatric facility in India between 1999 and 2008; of these, 11.5 percent of outpatients and 5.4 percent of inpatients met ICD-10 criteria for Trance and Possession Disorder.

Community-based surveys have yielded higher rates. In 1972, Morris Carstairs and Ravi Kapur found a 6-month period prevalence of 0.97 percent for involuntary possession syndrome in a South Indian village, using a symptom checklist in a 50 percent sample of all village adults (N = 1,233). They also reported a 6-month period prevalence of an additional 1.8 percent for voluntary possession states. V. Venkataramaiah and colleagues surveyed the total population (N = 1,158) of a school catchment area in another South Indian village during an epidemic of possession syndrome in

1979. They found a 1-year period prevalence of 3.5 percent for involuntary attacks, as well as a prevalence of 0.2 percent for voluntary possession trance episodes.

Anthropological accounts of widespread occurrence and cultural barriers to discussing the supernatural features of the syndrome have suggested that the real national prevalence is likely higher than reported in public surveys. The caseloads of indigenous healers probably represent the best source of global data. For example, 17 percent of those patients seen by one traditional healer during a 4-month period who met research criteria for psychiatric ailments ($N = 30$) received an indigenous label of possession syndrome. Finally, the periodic occurrence of syndrome epidemics complicates the epidemiological picture.

Possession syndrome is more common in women, with a female-to-male ratio of 2 to 3:1 in most community and psychiatric cohorts. The 1999 to 2008 cohort showed much higher female-to-male ratios among outpatients (4.7:1) than inpatients (1.3:1), suggesting that women with Trance and Possession Disorder were significantly less likely to be hospitalized. Sex distribution may also vary historically, according to evolving cultural gender attributions, and show geographical or possibly idiosyncratic variation. Various studies conducted in India suggest that women experience considerable social disadvantage related to the typically patriarchal social system, including early marriage, childbearing as the primary social function, family and spousal violence, lack of social integration, poor social support from family members, and limited autonomy. These factors may contribute to the higher prevalence of possession syndrome among women. Age of onset is usually between 15 and 35, but many cases reportedly begin in childhood. Attacks may persist well into middle age, and geriatric cases have also been reported. The syndrome has been identified in all Hindu castes, as well as in Muslims, Christians, and Tribal peoples; however, it is more commonly described among Hindus.

Caste distribution may fluctuate historically and respond to sociocultural determinants of the appropriateness and desirability of possession syndrome as an idiom of distress. It has been reported more frequently among disadvantaged castes, of low income and education, and rural residence. Nevertheless, some case studies revealed the prevalence of the syndrome among more advantaged groups, such as Brahmins, despite the cultural proscriptions against it. Case reports suggest higher rates among widowed persons, but a systematic study did not confirm this association. The presence of possession syndrome in other family members (or in ancestors) may be a risk factor, but this relationship has not been investigated systematically.

Precipitants/Etiology. Precipitants of possession syndrome were varied

but typically consisted of marked social or family conflicts, or stressful life transitions, of subacute duration, eliciting strong feelings of vulnerability in persons without firm emotional support. Examples encountered in the literature included marital conflict, abuse and neglect, at times associated with alcoholism; arrival of a new bride to the home of her husband's family; delay in arranging marriage, or in consummating it; forced marriage; widowhood; postpartum status; loss of family social standing; death of a family member; difficulty finding employment and financial difficulties; alienation from family support; and subordination to other family members and in-laws. These precipitants were identified in 67 and 73 percent of separate mental health cohorts.

Several clinical investigators stressed the pathogenic role of hysterical, histrionic, and immature ego defenses in the escalation of these stressors into triggers for a possession syndrome response. However, some of the precipitants elicited by researchers were obviously traumatic—deaths of children, murder of husband, sexual abuse of a close friend, measles epidemic, physical abuse, revelation of husband's incest with daughter. Social scientists such as Richard Castillo have emphasized the underlying context of violence against women in India—including selective abortion, infanticide, murder due to discovered premarital relations, rape, sale into prostitution, "dowry death" (new bride is killed if dowry is considered insufficient), marital abuse, witch killings, and *sati* (widow burning). This data suggests that many instances of possession syndrome may arise not from intrapsychic vulnerability, but from widespread undisclosed early trauma, an exposure that has been insufficiently investigated. Some researchers agreed on the etiological role of childhood stress as a risk factor for possession syndrome, but differed on the specific contributions of physical and sexual abuse as opposed to emotional or physical neglect. Alternatively, anthropological investigators have argued that individual biographical data was insufficient to explain the specific association of spirit possession with marginal persons and transitional life phases, such as birth, menarche, marriage, pregnancy, widowhood, and death. They located the etiology of attacks in caste-specific cultural expectations linking social and spiritual vulnerability. Additional collectively defined precipitants of possession syndrome include certain acts and settings, such as crossing physical boundaries (e.g., the village limit, an empty field at night), which invite the attention of possessing spirits.

Relationship to Psychiatric Diagnoses. Psychoanalytically based evaluations of subjects with possession syndrome often reported evidence of hysterical traits predisposing to dissociative psychopathology in otherwise normal subjects. Others distinguished the intrapsychic origin of the acute behavioral syndrome—which they attributed to hysterical processes—from the underlying personality formation and psychic

vulnerability. These authors argued that attacks could arise in subjects with diverse personality structures and biological propensities. Some research reported a substantial proportion of attacks among patients with a primary diagnosis of schizophrenia (40 to 59 percent) and manic depressive illness (11 to 13 percent), in whom the possession symptoms took longer to resolve, compared to patients without these disorders.

As noted earlier, DSM-5 incorporated most cases of possession syndrome into Dissociative Identity Disorder. However, cases that do not fulfill the amnesia criterion should receive other DSM diagnoses, including Other Dissociative Disorder. As with all cultural concepts of distress, there is no one-to-one overlap between the cultural label and any DSM disorder; psychiatric evaluation must proceed at the level of the individual. In ICD-10, possession syndrome cases would fulfill criteria for Trance and Possession Disorders. Prior to the inclusion of possession phenomenology in Dissociative Identity Disorder, possession syndrome did not have a diagnostic home in DSM: a study conducted in 1987 in a New Delhi outpatient psychiatric clinic (N = 1,517) found that 10 percent of cases with dissociative pathology received an ICD-10 diagnosis of Trance and Possession Disorders, whereas none met DSM-III-R criteria for Multiple Personality Disorder.

Additional Clinical Features. Rates of relapse are unknown, although anecdotal reports suggest it is not uncommon. Relapse within 15 years was not infrequent and it was also recorded after two to three decades. Four of 15 subjects in one study were found to have suffered previous episodes, and 1 subject in another series reported three consecutive postpartum attacks. Entry into a ritual possession cult may reduce or eliminate relapse.

Specific Cultural Factors. Possession syndrome constitutes a folk illness category throughout India and Sri Lanka. It may present initially in a variety of forms, linked by the attribution of spirit etiology. When it presents in a nonspecific fashion, indigenous diagnosis is confirmed by the appearance of the altered state during the therapeutic ritual. It is considered an affliction by its painful, involuntary nature, and attributed to the intervention of specific spiritual agencies acting independently or at the behest of a witch. Certain castes and persons in transitional states (e.g., puerperium) are considered most vulnerable to spirit attack, especially when deprived of emotional and material support. The degree of pathology ascribed to the sufferer, the therapy chosen, and the outcome attained depend on the interaction of several factors: psychological (e.g., traumatized ego), biological (e.g., psychosis), sociocultural (e.g., caste status), and situational (e.g., family dynamics). For instance, when it presents in some Moger women of South India, possession syndrome may be seen as an affliction opening the door to a sacred practice in a

possession cult. In some Havik Brahmin women, however, it may constitute instead the last step in a series of symptomatic expressions of acute distress, leading directly to exorcism and incomplete conflict resolution.

Treatment. Specialized indigenous practitioners and ritual therapies are generally available and widely utilized, but psychiatric treatment is typically avoided. Indigenous treatments include neutralization of the conflict or stress via the communal rituals involved in exorcism, as well as the reformulation of the suffering into beneficent individual and communal practice via initiation into a spirit devotion cult—such as the Siri cult of South India—or education into the roles of oracle (diviner), exorcist, or, rarely, avatar (divine incarnation).

Shenjing Shuairuo

Shenjing shuairuo ("weakness of the nervous system" in Mandarin Chinese) is a translation and cultural adaptation of the term "neurasthenia," which was transmitted into China from the West and from Japan in the 1920s to 1930s. Revived in its modern form by the American neurologist George Beard since 1868, his formulation of neurasthenia (Greek for "lack of nerve strength") originally denoted a heterogeneous syndrome of lassitude, pain, poor concentration, headache, irritability, dizziness, insomnia, and over 50 other symptoms. It was considered at first an "American disease" resulting from the "pressures" of a rapidly modernizing society, but was later adopted by European diagnosticians. Its pathophysiology was thought to derive from a lowering of nervous system function on a physical rather than emotional basis, due to excessive demand on its use, especially among the educated and wealthier classes. In Soviet psychiatry, buttressed by Pavlovian research, neurasthenia was a central component of mental health nosology, which became especially influential in Chinese psychiatry after the communist revolution of 1949.

While neurasthenia declined in importance in Western classification systems during the 20th century, *shenjing shuairuo* underwent marked popular and professional development in mainland China, Taiwan, Hong Kong, in Chinese migrant communities, and in Japan, where a similar syndrome is labeled *shinkei suijaku*. From a peak in about 1980, when it may have constituted up to 80 percent of all "neurotic" diagnoses in Chinese societies, *shenjing shuairuo* has undergone intense psychiatric and anthropological re-examination over the last 35 years. It was featured in the *Chinese Classification of Mental Disorders, Second Edition, Revised* (CCMD-2-R; 1993), under the section on "Other Neuroses." The CCMD-2-R diagnosis required three symptoms out of five nonhierarchical symptom clusters, organized as weakness, emotional, excitement, and nervous symptoms, as well as a fifth category of sleep disturbances. Like other

neurotic disorders in the Chinese manual, the condition had to last at least 3 months and (a) lower the efficiency of work, study, or social function; (b) cause mental distress; or (c) lead to treatment-seeking. Depressive and anxiety symptoms, which used to be central in previous professional definitions of *shenjing shuairuo*, became exclusionary in CCMD-2-R, probably as a result of the growing influence of Western nosologies. Physical weakness and mental fatigue, which are prominent elements of some neurasthenic presentations in other parts of the world, are not prioritized in the Chinese classification. The third edition of the CCMD (2001) retained *shenjing shuairuo* in the section on the Neuroses as a diagnosis of exclusion. This decision was based on multicenter field trials in China on the discriminant validity and clinical utility of the Neurosis subtypes.

A general category of "neurasthenia" is also included in ICD-10 (F48.0), as a result of what this manual references as its widespread and persistent use around the world. This category emphasizes the components of weakness and fatigue and relegates other possible symptoms to secondary characteristics:

> Considerable cultural variations occur in the presentation of this disorder, and two main types occur, with substantial overlap. In one type, the main feature is a complaint of increased fatigue after mental effort, often associated with some decrease in occupational performance or coping efficiency in daily tasks. The mental fatigability is typically described as an unpleasant intrusion of distracting associations or recollections, difficulty in concentrating, and generally inefficient thinking. In the other type, the emphasis is on feelings of bodily or physical weakness and exhaustion after only minimal effort, accompanied by a feeling of muscular aches and pains and inability to relax. In both types, a variety of other unpleasant physical feelings are common, such as dizziness, tension headaches, and feelings of general instability. Worry about decreasing mental and bodily well-being, irritability, anhedonia, and varying minor degrees of both depression and anxiety are all common. Sleep is often disturbed in its initial and middle phases but hypersomnia may also be prominent.

Substantial controversy remains over the proper status of weakness and fatigue in the international definition of neurasthenia, as well as whether these symptoms are in fact more prominent in *shenjing shuairuo* than accounted for by spontaneous patient reports and thus as codified in the CCMD-3. Some investigators consider the centrality of fatigue in the ICD-10 definition of neurasthenia a tendentious alignment of the international classification with the Western diagnoses of somatoform disorders and Chronic Fatigue Syndrome. For example, Sing Lee estimated that the ICD-10 definition excludes over 80 percent of Chinese patients with *shenjing shuairuo* who report prominent symptoms other than fatigue. The DSM-5 Glossary of Cultural Concepts of Distress includes an entry for *shenjing shuairuo*, which follows the CCMD-3 definition and does not prioritize the fatigue component.

Phenomenology. Given the evolution of diagnostic practice regarding *shenjing shuairuo* in Chinese societies over the last decades—which may be

labeled the professional approximation of the condition, or its "disease" aspect—phenomenological description in this chapter will be based instead on clinical histories of self-identified sufferers—or the "illness" aspect of the syndrome. The elements in Table 29.1–4 are prototypical.

Epidemiology. The changing professional definitions of ICD neurasthenia and CCMD *shenjing shuairuo* complicate interpretation of the epidemiological reports, which are in any case limited in European languages. Most studies used ICD definitions, which in turn have evolved over time or refer to rates of "neurasthenia" without distinguishing between ICD and CCMD diagnoses. Therefore, the relationship between these reports of neurasthenia prevalence and the folk presentation of *shenjing shuairuo* remains unclear.

Table 29.1–4.
Phenomenology of *Shenjing Shuairuo* (Weakness of the Nervous System)

a. Onset is usually gradual, sometimes spanning several years, and typically emerges out of a conflictive, frustrating, or worrying situation that involves work, family, and other social settings, or their combination. A sense of powerlessness in changing the precipitating situation appears central to most illness accounts of the syndrome.
b. Symptoms show substantial individual variation, but usually involve at least some of the following spontaneous complaints: insomnia, affective dysphoria, headache, bodily pains, and distortions (e.g., "swelling" of the head), dizziness, difficulty concentrating, tension and anxiety, worry, fatigue, weakness, gastrointestinal problems, and "troubled vexation" (*fan nao*). This last emotion has been described as a form of irritability mixed with worry and distress over "conflicting thoughts and unfulfilled desires," that may be partially concealed in Chinese communities for the sake of preserving social harmony.
c. The sufferer frequently seeks the sick role, attributing his/her difficulties in meeting work, school, or other social expectations to the syndrome. Sources of treatment vary substantially across Chinese communities, depending on the availability of formal and traditional service sectors. Typically, physical health services, such as general health or neurology clinics, or traditional Chinese medicine providers are preferred to psychiatric services, though this may be less pronounced among school-related presentations, especially in more Westernized settings such as Hong Kong.
d. Course is variable, and may respond closely to changing interpersonal and social circumstances. Amelioration of the precipitating stressor(s) typically brings about substantial improvement, although residual symptoms appear common.
e. Response to treatment may be strongly mediated by the illness role and its relationship to the intractability of precipitating stressors.

Several community studies have been conducted, but most are available only in Chinese. One community-based study conducted in 1959 in Chendu City, mainland China (N = 10,830), found a point prevalence of 5.9 percent for unspecified "neurasthenia" of greater than 3 months' duration. Prevalence rates from other community surveys prior to 1980 have been reported as between 6 and 10 percent. Several investigators noted a rise in presentations of *shenjing shuairuo* to medical services in the aftermath of the Cultural Revolution (1966 to 1976), when social space was opened for a

public accounting of the traumas associated with this period. However, epidemiological documentation of this likely upswing in community and clinical prevalence are limited. One nationwide community-based study from 1986 found a prevalence rate of 1.3 percent in the general population. Prevalence rates among students have typically been high, given the association of school-based stressors with *shenjing shuairuo*. In 1993, 5.1 percent of Chinese college students in Anhui, mainland China, fulfilled CCMD-2 criteria for the syndrome. Lee and colleagues reported a 6.3 percent prevalence of moderate or severe *shenjing shuairuo* among Hong Kong undergraduates in 1995.

Clinical epidemiology studies have varied markedly over time. By numerous anecdotal reports, between about 1950 and 1980, neurasthenia, either in the ICD or the CCMD formulation, constituted probably the most common psychiatric diagnosis in China. The narrowing of the psychiatric definition of neurasthenia since 1980—such as by the exclusion of primary affective presentations—and its reduced application in diagnostic practice likely decreased prevalence rates in clinical epidemiology. Yet the prevalence of *shenjing shuairuo* in medical clinics has remained high. A 1983 clinical epidemiological study in a county general hospital in Sichuan province (N = 1,401) found a prevalence of 3.9 percent for ICD-9-defined neurasthenia. In 1997 to 1998, 5.3 percent of adult outpatients using internal medicine, neurology, and traditional Chinese medicine clinics in a referral hospital in Hunan province met CCMD-2-R criteria for *shenjing shuairuo*. Rates of up to 6.3 percent have been reported in other general clinical settings. Psychiatric inpatients appear to have much lower rates, probably reflecting the outpatient nature of most neurasthenia care. In 1990, the prevalence of neurasthenia among 2,275 inpatients discharged from the Shanghai Institute of Mental Health was only 0.26.

The 12-month prevalence of ICD-10 neurasthenia among Chinese-Americans in Los Angeles County was studied by Yan-Ping Zheng and colleagues. The prevalence of all neurasthenia diagnoses, regardless of diagnostic exclusions, was 6.4 percent. The rate of "pure" neurasthenia—those without any lifetime comorbid DSM-III-R diagnoses—was 3.6 percent, higher than any other DSM-III-R disorder. These "pure" neurasthenia sufferers did not differ on most sociodemographic characteristics from respondents without psychiatric diagnosis, with "pure" depression, or with "pure" anxiety disorder. Further analysis of the study subsample (95 percent) who were foreign born revealed that those with higher neurasthenic symptom counts were older, had less education, and were more likely to be formerly married (separated, widowed, or divorced). Lower acculturation increased the risk of neurasthenia, but only in the presence of financial stress (as an interaction term). Participants with lower acculturation were more likely than those with higher acculturation to express the syndrome when confronted with financial stress. The

authors concluded that "neurasthenia is a unique form of expressing distress among low acculturated groups of Chinese Americans."

Precipitants/Etiology. Empirical assessment of the precipitants of *shenjing shuairuo* has found high rates of work-related stressors, which were made more intractable by the centrally directed nature of mainland Chinese society. These included unwelcome work assignments, job postings that caused family separations, harsh criticism at work, excessive workloads, monotonous tasks, and feelings of inadequacy or incompatibility of skills and responsibilities. Students usually described somewhat less severe study-related precipitants, particularly school failure or anxiety over the mismatch between personal or family aspirations and performance. Other interpersonal and family related stressors included romantic disappointments, marital conflict, and the death of a spouse or other relative. The work of Arthur Kleinman has emphasized the political nature of many precipitants of *shenjing shuairuo*, especially during the 1980s and 1990s as a result of the vicissitudes of the Cultural Revolution. This societal upheaval involved traumatic changes in social circumstances, family separations, acute humiliations, social ostracism, demoralization, and exposure to violence and deprivation.

A common theme to most precipitants of *shenjing shuairuo* involved frustration of a keenly held aspiration, glossed as "troubled vexation," to which social circumstances allowed no or insufficient solution. Loss of face (*mianzi, lianzi*), a key social value in Chinese communities, was also frequently noted as a trigger of *shenjing shuairuo*.

Chinese etiological understandings of the syndrome commonly invert the Western view of "psychosomatic" presentations, whereby social–interpersonal precipitants cause psychological distress that is displaced onto bodily experience. In what Kleinman described as "somatopsychic" interpretations, many sufferers of *shenjing shuairuo* located the pathological impact of external stressors directly onto physical functioning, producing somatic symptoms that only secondarily caused psychological distress. This schema permits sufferers to acknowledge a psychological contribution to their illness, while yet locating the nexus of pathology on physical processes.

Relationship to Psychiatric Diagnoses. Several studies in Chinese communities found a strong relationship between *shenjing shuairuo* and the mood, anxiety, and somatoform disorders. Kleinman found considerable overlap between *shenjing shuairuo* and these disorder clusters in his studies in both mainland China and Taiwan. In the People's Republic, his study of 100 Chinese patients diagnosed with *shenjing shuairuo* found that 93 also had a depression diagnosis, with 87 meeting criteria for major depressive disorder. Sixty-nine patients were also

diagnosed with an anxiety disorder and 25 with a somatoform disorder. In Taiwan, 22 of 51 patients with *shenjing shuairuo* also met criteria for major depressive disorder, which was the most common psychiatric diagnosis in this group of patients.

Subsequent investigations confirmed these associations, with rates of depressive disorders ranging from 30 to 70 percent in cases of *shenjing shuairuo*. The association with somatoform and anxiety disorders was also found repeatedly. The exact correlations were particularly dependent on whether the subjects came from community or clinical settings, and if clinical studies, whether they were based in general medicine or psychiatric clinics. Outpatient primary care settings tended to yield associations between *shenjing shuairuo* and the somatoform disorders. One study by Sing Lee and colleagues with 100 primary care patients in Hong Kong with unexplained fatigue of at least 6 months duration that excluded patients already in treatment for depressive disorder found that 72 percent endorsed a self-label of *shenjing shuairuo*. In the full sample, the most common DSM-III-R diagnoses on SCID interview were somatoform pain disorder (47 percent) and undifferentiated somatoform disorder (35 percent). Fourteen patients met criteria for depressive disorders, with the other diagnoses mostly in the anxiety spectrum.

A study conducted in 1997 to 1998 by Doris Chang and colleagues in three general medicine clinics in Hunan also found similar correlations. The study included 139 patients, mostly of rural origin, who presented with unexplained somatic complaints. Only 55 percent received any DSM-IV diagnosis on SCID interview, all in the somatoform disorder spectrum. Undifferentiated somatoform disorder was the closest counterpart to *shenjing shuairuo* in DSM-IV, with a diagnostic concordance of kappa = 0.24 ($p < 0.001$). The fact that 45 percent of subjects with the CCMD-2 syndrome did not meet criteria for any DSM-IV disorder, however, suggested to the authors that *shenjing shuairuo* is a distinct clinical condition with only partial comorbidity with DSM diagnoses.

A similar conclusion was reached by Zheng and colleagues in their 1997 community-based epidemiological study among Chinese Americans in Los Angeles County (N = 1,747). This study did not use the CCMD definition of *shenjing shuairuo*, however, relying instead on the ICD-10 criteria for neurasthenia, which has only a partial overlap with CCMD-2-R. Nevertheless, they identified 112 (6.4 percent) subjects with ICD-10 neurasthenia in the community. Over half (N = 63; 56 percent) did not meet criteria for a DSM-III-R diagnosis. Compared with those with depression and anxiety diagnoses, these respondents reported fewer psychological symptoms on the SCL-90-R, but similar levels of somatization symptoms.

The overlap with ICD definitions of neurasthenia has been assessed in a few studies. Chang and colleagues found an overlap of 53 percent between

the CCMD-2 *shenjing shuairuo* category and ICD-10 neurasthenia. An even more limited overlap with the international classification was found in a 1989 study conducted at the Shanghai Psychiatric Hospital. Clinical diagnoses of *shenjing shuairuo* obtained by two senior Chinese psychiatrists yielded only a 22.5 percent overlap with the ICD-9 neurasthenia category.

Additional Clinical Features. Clinical course of the syndrome may depend on the associated psychiatric comorbidity and on the degree of persistence of the precipitating stressors. One longitudinal study found complete resolution of *shenjing shuairuo* symptoms and good social adjustment 20 years after the index diagnosis in 83 out of 89 cases. Only one case was receiving continued treatment, and no subjects reported the onset of depressive disorder subsequent to the *shenjing shuairuo* diagnosis.

Chinese psychiatrists have carried out numerous studies of neurophysiological and cognitive function in *shenjing shuairuo* patients since the 1950s. Most have reported abnormalities compared to normal controls, including in tests of polysomnography, EEG, psychogalvanic reflex, gastric function, and memory function. These findings need to be replicated with well-controlled samples using contemporary diagnostic instruments.

Specific Cultural Factors. The evolving definitions of *shenjing shuairuo* have emerged out of a tradition of syncretism in Chinese medicine between indigenous illness understandings and international contributions. Nineteenth-century Western notions of a weakened nervous system due to overuse (neurasthenia) found an ancient cognate expression in Chinese concepts of bodily meridians or channels (*jing*) binding vital organs in balanced networks along which forces (e.g., *qi*, vital energy, in *yin* and *yang* forms) could be disrupted from their normal harmonious flow. This gave rise to *shenjing shuairuo*, an illness whereby the *jing* that carry *shen*—spirit or vitality, the capacity of the mind to form ideas and the desire of the personality to live life—have become *shuai* (degenerate) and *ruo* (weak) following undue nervous excitement.

As noted by Kleinman, this Chinese conception is much more relational and contextual than the typical Western medical view. Bodies are seen as acting in the midst of specific contexts and social relationships, rather than abstractly or individually. Physical networks are thus metaphorical representations of interpersonal and societal connections, with emotions acting as mediating channels. These social networks, just like bodies, "can be vitalized or devitalized, demoralized or remoralized. Indeed bodies are networked, networks embodied."

The result is an illness form, *shenjing shuairuo*, that tightly links social

and interpersonal situations with bodily afflictions, such as the metaphorical linkage described earlier between dizziness and social dislocation due to revolutionary upheavals. Other contributors to the persistence of this illness form include elevated stigma against mental illness in Chinese communities, which is forestalled by the primarily physical nature of *shenjing shuairuo*; the contributions of Pavlovian psychiatry and Japanese research to Chinese nosology, which emphasized neurasthenia beyond its role in Western psychiatry; the prominent use of the syndrome by the communist authorities as a health-based, treatable cause of reduced productivity during the 1950s and 1960s; and its popular acceptability as a veiled complaint against social deprivation and oppression at a time when direct confrontation would be harshly suppressed.

Treatment. Kleinman noted a discrepancy in treatment patterns between Taiwanese and Chinese *shenjing shuairuo* patients, with those in Taiwan more frequently resorting initially to self-help remedies, which included special foods, tonics, and dietary changes. He attributed this discrepancy to the availability of inexpensive care in the People's Republic.

When accessing formal sectors of care, most patients used both Western-trained physicians and traditional Chinese doctors. Nonpsychiatric medical settings were preferred, including neurology and general medicine clinics, in concert with cultural understandings of the somatopsychic etiology of *shenjing shuairuo*, which emphasize its physical mediation. The modality of treatment was usually traditional Chinese medicines, which were prescribed by both Western-trained and Chinese-style doctors. This conformed to the balanced status ascribed to both types of training among Chinese physicians. Polypharmacy was common, combining sedatives, traditional herbs, antianxiety agents, vitamins, and other tonics. In spite of active suppression of religious healing in China, almost a quarter of patients were also engaged in such treatment in the early 1980s.

ONGOING WORK ON THE RELATIONSHIP BETWEEN CULTURAL CONCEPTS OF DISTRESS AND PSYCHIATRIC NOSOLOGY

Research

In 1969, Yap challenged the field to learn more about "culture-bound syndromes" so that their relationship to psychiatric disorder could be resolved, yet much research still needs to be done. Without a more systematic and wide-ranging program of research on a broader range of cultural concepts of distress, current approaches that only focus on classificatory exercises are unlikely to further the current understanding. The strategy of trying to find the right classificatory scheme by basing it on

similarity between one or two symptoms of a cultural syndrome and of the psychiatric disorder is limited. It is particularly problematic when researchers privilege the DSM categories as the main organizing structure of relevance to the cultural concepts of distress. This approach is unlikely to produce new insights into these concepts on their own terms or in relation to psychiatric diagnosis. It ignores, for example, that many cultural concepts are not organized syndromically, so that classifying them on the basis of prototypical symptoms is incompatible with their organizational structure. Comparing etiological categories with symptom-based diagnoses on the nature of their symptoms is not only useless but counter-productive.

A simple, symptom-based comparison further ignores the function of cultural concepts as idioms, or modes of expressing distress. Comparison of cultural concepts and psychiatric disorders as idioms should focus on the mutual compatibility of their intended communications (do they deliver the same message?) and the cultural fit between the idiom—whether indigenous or professional—and the situation in which it arises (is the message equally appropriate?). An idiom of distress intended as a sign of anger that is misinterpreted by the diagnostician as a sign of severe mental illness (e.g., psychosis) is not being accurately translated, from the idiomatic perspective. Similarly, an idiom that is considered culturally appropriate to the situation (e.g., anger as a comprehensible reaction) would be well matched by a diagnosis that idiomatically conveys a similar sense of appropriateness.

In order to resolve the cognitive dissonance created by the complex relationship between the cultural concepts and psychiatric disorders, cross-cultural mental health researchers need to carry out intensive research that illuminates these concepts rather than simply attempting to subsume them into psychiatric categories and obliterating them. Peter Guarnaccia and his colleague Lloyd Rogler have proposed a research program on cultural concepts of distress that would bring their study into the mainstream of mental health research. This framework is faithful to the holistic nature of the cultural concepts, while at the same time applying current research approaches from a number of fields. It comprises a series of key questions that need to be answered in order to understand these concepts from the cultural as well as the psychiatric perspectives. The questions are organized into the following five broad dimensions:

- **Nature of the Phenomenon:** How is the folk category characterized within its cultural context? What are the defining features of the phenomenon? Is the folk illness organized syndromically? Does it refer primarily to an etiological construct? What is the message it attempts to convey as an idiom of distress? How salient is the concept, defined as the rapidity and extent of its recognition within a cultural group?
- **Location in the Social Context:** Who are the people who report the

folk category and what is their location in the social structure? What are the situational factors that provoke these conditions?

- **Relationship to Psychiatric Disorder:** How is the folk illness empirically related to psychiatric disorder(s)? Does it cut across diagnostic groupings? What is the range of its overlap with psychiatric categories?
- **Social/Psychiatric History of the Cultural Concept:** When the onset and offset of the cultural concept and the associated psychiatric disorder(s) do not coincide, what is the sequence of onset? How does the life history of the sufferer, particularly the experience of traumatic events, affect the sequence? How many episodes of the folk illness has the person experienced? How severe have they been?
- **Treatments and Outcomes:** Holistically, how should be person's suffering be addressed? If the folk illness is effectively treated, does the sufferer also experience relief of the associated psychiatric disorder(s)? And vice-versa, if the psychiatric disorder is resolved with psychotherapy or pharmacotherapy, does the folk illness resolve as well?

Clinical Application

Medical anthropology and cultural psychiatry and psychology have developed tools and perspectives that are useful in assessing cultural concepts of distress in clinical work and in negotiating between the concepts and psychiatric diagnoses. One such model is the Explanatory Model Questions developed by Kleinman. Another model is the DSM-IV Outline for Cultural Formulation, which was revised for DSM-5 and operationalized as a set of semistructured questionnaires known collectively as the Cultural Formulation Interview. These frameworks are designed to help clinicians inquire about and better understand the cultural dimensions of their patients' presentations of their mental health problems.

The Explanatory Model Questions were proposed by Kleinman in 1980 as a tool to aid clinicians in interviewing patients about their cultural understandings of their illnesses. This deceptively simple set of eight groups of questions provides a framework for a brief clinical interview that can identify the presence and meaning of a folk illness category. In particular, the first five groups of questions are useful in this regard and include:

- What do you call your problem? What name does it have?
- What do you think has caused your problem?
- Why do you think it started when it did?
- What does your sickness do to you? How does it work?
- How severe is the sickness? Will it have a short or long course?

Anne Fadiman's *The Spirit Catches You and You Fall Down* is a particularly accessible and informative example of the use of Kleinman's Explanatory Model Questions to elucidate a folk illness. This book concerns a young Hmong girl with *quag da peg*, which are the Hmong words for the title of the book, and which is sometimes equivalent to the medical diagnosis of epilepsy. After detailing the miscommunications, misunderstandings, and conflicts between the family and their medical providers, Fadiman used the Explanatory Model Questions to highlight her understanding of spirit possession and soul loss for the Hmong family she had come to know. She realized that had the doctors been able to ask these questions and understand the answers with the use of a skilled interpreter, they would have had a much better understanding of the family's perspective on their daughter's illness. More research-oriented examples related to depression and other psychiatric disorders can be found in Kleinman's books *Social Origins of Distress and Disease* (1986) and *The Illness Narratives* (1988).

A more specific description of how to assess patients' cultural understanding of their condition was included in DSM-IV as the Outline for Cultural Formulation. Since it was abbreviated for inclusion with no examples, and was juxtaposed to the "Glossary of Culture-Bound Syndromes" in the last substantive appendix, this may have led to the misleading impression that the formulation was only for individuals presenting with folk illnesses. For the purposes of this chapter, this linking is useful to clinicians, since they need guidance on how to inquire about and understand culture-specific presentations of mental illness. However, the cultural formulation is clearly relevant to all clinical encounters.

The key area of the Outline for Cultural Formulation in this regard is the section on cultural explanations of the individual's illness. The major elements of inquiry include the predominant idioms of distress through which symptoms are communicated; the meaning and perceived severity of the individual's symptoms in relation to the norms of the cultural reference group; local illness categories used by the individual's family and community to identify the condition; the perceived causes or explanatory models that the individual and reference group use to explain the illness; and past and current experiences with popular and professional sources of care. These areas can serve as a support for clinicians who seek to better understand their patients' perspectives on the symptoms, expressions, and difficulties they present.

In DSM-5, the Outline was rendered more operational for routine clinical practice by standardizing its implementation as the Cultural Formulation Interview. This is composed of three types of semistructured interviews that offer clinicians concrete ways to carry out the cultural assessment of patients and their entourages. The three components of the Cultural Formulation Interview are a core 16-item questionnaire, an

informant version of the core interview to obtain information from caregivers, and 12 supplementary modules that expand on these basic assessments. If used routinely to frame the intake evaluation, the Cultural Formulation Interview can help establish a foundation of person-centered information on the patient's context and perspective, and a working alliance, on which to build the rest of the diagnostic interview and treatment negotiation process.

There are several resources that clinicians can turn to for further guidance and examples using this formulation technique. The cultural formulation has been elaborated by the GAP Committee on Cultural Psychiatry in their 2002 publication *Cultural Assessment in Clinical Psychiatry*. Illustrative cases using the cultural formulation and discussing the implications of cultural understanding of mental health problems have been available in select issues of *Culture, Medicine and Psychiatry* starting in 1996. Key writings on the cultural formulation approach were assembled in *Cultural Formulation: A Reader for Psychiatric Diagnosis* (2008). The McGill Cultural Consultation Service published *Cultural Consultation: Encountering the Other in Mental Health Care* in 2014, describing the cultural assessment of over 400 cases with an expanded cultural formulation approach. Finally, the *DSM-5 Handbook on the Cultural Formulation Interview* (2016) describes in detail the components of the Cultural Formulation Interview, presents multiple clinical vignettes to illustrate its use, and suggests strategies for incorporating the interview into clinical training and routine practice.

Once the information is elicited, the next key step is for clinicians to lead a negotiation process whereby they and their patients come to understand each other's perspectives on the mental health problem and come to agreement on how to proceed. Kleinman described this approach in *The Illness Narratives*. The process that Kleinman developed is for the provider to first elicit and understand the patient's explanatory model and then to make clear the provider's explanatory model. The process then focuses on developing a treatment plan based on negotiation between the patient and provider models. It may be useful to include the patient's cultural terms and explanations in case formulations and treatment plans, as they may help clarify symptoms and etiological attributions that could otherwise be confusing. This approach has been further elaborated in a variety of cultural competence training programs for physicians to aid them in this negotiation process. The impact on quality of mental health care from these approaches for assessing cultural understandings of cultural concepts of distress and for clinician-patient negotiation are pressing areas for research.

REFERENCES

Alarcón RD, ed. *Cultural Psychiatry*. Basel: Karger; 2013.

Bhui K, Bhugra D, eds. *Culture and Mental Health*. London: Hodder Arnold; 2007.

Chang DF, Myers HF, Yeung A, Zhang Y, Zhao J, Yu S. Shenjing Shuairuo and the DSM-IV: diagnosis, distress, and disability in a Chinese primary care setting. *Transcult Psych*. 2005;42:204.

Good B, Delvecchio Good MJ. Toward a meaning-centered analysis of popular illness categories: "fright illness" and "heart distress" in Iran. In: Marsella AJ, White GM, eds. *Cultural Conceptions of Mental Health and Therapy*. Dordrecht, the Netherlands: D. Reidel Publishing Company; 1982:141.

Groleau D, Young A, Kirmayer LJ. The McGill Illness Narrative Interview (MINI): an interview schedule to elicit meanings and modes of reasoning related to illness experience. *Transcult Psychiatry*. 2006;43:671.

Group for the Advancement of Psychiatry. *Cultural Assessment in Clinical Psychiatry*. Washington, DC: American Psychiatric Publishing; 2002.

Guarnaccia PJ, Lewis-Fernández R, Martínez I, et al. Ataques de nervios as a marker of social and psychiatric vulnerability: Results from the NLAAS. *Int J Soc Psychiatry*. 2010;56:289.

Guarnaccia PJ, Rivera M, Franco F, Neighbors C. The experiences of ataques de nervios: towards an anthropology of emotion in Puerto Rico. *Cult Med Psychiatry*. 1996;20:343.

Hinton DE, Good BJ, eds. *Culture and Panic Disorder*. Stanford, CA: Stanford University Press; 2009.

Hinton DE, Lewis-Fernández R. The cross-cultural validity of posttraumatic stress disorder: implications for DSM-5. *Depr Anx*. 2011;28:783.

Hinton DE, Barlow DH, Reis R, de Jong J. A transcultural model of the centrality of "Thinking a lot" in psychopathologies across the globe and the process of localization: a Cambodian refugee example. *Cult Med Psychiatry*. 2016;40(4):570–619.

Kaiser BN, McLean KE, Kohrt BA, et al. Reflechi twòp—thinking too much: Description of a cultural syndrome in Haiti's Central Plateau. *Cult Med Psychiatry*. 2014;38(3):448.

*Kirmayer LJ, Guzder J, Rousseau C, eds. *Cultural Consultation: Encountering the Other in Mental Health Care*. New York: Springer; 2014.

*Kirmayer LJ, Lemelson R, Cummings CA, eds. *Re-Visioning Psychiatry: Cultural Phenomenology, Critical Neuroscience, and Global Mental Health*. New York: Cambridge University Press; 2015.

Kleinman A. *Patients and Healers in the Context of Culture*. Berkeley: University of California Press; 1980.

*Kleinman A. *Social Origins of Distress and Disease: Depression, Neurasthenia, and Pain in Modern China*. New Haven: Yale; 1986.

Kohrt BA, Rasmussen A, Kaiser BN, et al. Cultural concepts of distress and psychiatric disorders: literature review and research recommendations for global mental health epidemiology. *Int J Epidemiology*. 2014;43(2):365.

Lee S, Kleinman A. Are somatoform disorders changing with time? The case of neurasthenia in China. *Psychosomat Med*. 2007;69(9):846.

*Lewis-Fernández R, Aggarwal NK, Hinton L, Hinton DE, Kirmayer LJ, eds. *The DSM-5 Handbook on the Cultural Formulation Interview*. Washington, DC: American Psychiatric Publishing; 2016.

Lewis-Fernández R, Guarnaccia PJ, Martínez IE, Salmán E, Schmidt A, Liebowitz MR. Comparative phenomenology of ataques de nervios, panic attacks, and panic disorder. *Cult Med Psychiatry*. 2002;26:199.

Lewis-Fernández R, Aggarwal NK. Psychiatric classification beyond the DSM: an interdisciplinary approach. In: Kirmayer LJ, Lemelson RB, Cummings CA, eds. *Revisioning Psychiatry: Cultural Phenomenology, Critical Neuroscience, and Global Mental Health*. New York: Cambridge University Press; 2015:434–468.

Lim RF, ed. *Clinical Manual of Cultural Psychiatry*. 2nd ed. Washington, DC: American Psychiatric Publishing; 2015.

López SR, Guarnaccia PJ. Cultural dimensions of psychopathology: the social world's impact on mental illness. In: Maddux JE, Winstead BA, eds. *Psychopathology: Foundations for a Contemporary Understanding*. New York: Routledge; 2012:45.

Nichter M. Idioms of distress revisited. *Cult Med Psychiatry*. 2010;34:401.

*Ruiz P. Addressing culture, race and ethnicity in psychiatric practice. *Psychiatr Ann*. 2004;34(7):527.

Simons RC, Hughes CC. *The Culture-Bound Syndromes*. Dordrecht, the Netherlands: Reidel Publishing Company; 1985.

Thomas JJ, Eddy KT, Ruscio J, Ng KL, Casale KE, Becker AE, Lee S. Do recognizable lifetime eating disorder phenotypes naturally occur in a culturally Asian population? *Eur Eat Disord Rev*. 2015;23(3):199–209.

Tseng WS, Streltzer J, eds. *Cultural Competence in Clinical Psychiatry*. Washington, DC: American Psychiatric Publishing; 2004.

▲ 29.2 Cultural Assessments in Psychiatry

ROBERTO LEWIS-FERNÁNDEZ, M.D.*, AND NEIL KRISHAN AGGARWAL, M.D.*

INTRODUCTION

Cultural psychiatry has been the subspecialty most involved in developing clinical, research, and educational methods for patient cultural assessments. Cultural psychiatry—also known as cross-cultural, transcultural, or anthropological psychiatry—has typically focused on three areas: (1) the extent to which manifestations of psychopathology are similar or different across societies, (2) the provision of mental health services for minorities in multicultural societies, and (3) the cultural bases of psychiatric knowledge and practice to reveal how mental health systems reflect broader societal concerns. Cultural assessments have been integral to each area, clarifying the relative biological and cultural (e.g., genetic and environmental) contributions of symptom manifestations; patient preferences on how, where, when, and with whom to receive care; and cultural assumptions within the patient–clinician relationship such as communication and relational styles. The inclusion of the Outline for

Cultural Formulation (OCF) in *DSM-IV* and the Cultural Formulation Interview (CFI) in *DSM-5* indicates that the need for cultural assessments has increasingly gained acceptance within general psychiatry.

HISTORICAL OVERVIEW

Cultural assessments in psychiatry have evolved in tandem with changing definitions of culture throughout time, especially within anthropology. In 1904, renowned psychiatrist Emil Kraepelin—one of the first to differentiate *dementia praecox* (schizophrenia) from manic-depressive illness—sailed to Java, Indonesia, and investigated symptom manifestations in 100 European, 100 indigenous, and 25 Chinese patients at the Buitenzorg mental hospital. Kraepelin found that psychotic Javanese patients experienced fewer visual and auditory hallucinations and crystallized delusions than European patients, leading him to conclude that the Javanese mind was more primitive than the European mind. Kraepelin kept meticulous records of patient symptoms and examined their similarities and differences, assuming that the definition and expression of psychotic illness would be the same among his study populations. Although his assumptions and interpretations may strike contemporary readers as problematic, Kraepelin's findings were situated within intellectual trends in 19th-century physical anthropology positing clear biological differences and hierarchical relationships among races. Kraepelin energized the nascent field of comparative psychiatry and was among the first psychiatrists working in unfamiliar cultures to use interpreters.

Another landmark movement emerged in the 1920s and 1930s with the Culture and Personality School. Social anthropologists such as Gregory Bateson, Ruth Benedict, Margaret Mead, and Ralph Linton collaborated with psychoanalysts such as Abram Kardiner and drew extensively from Sigmund Freud's theories on personality structure and organization. Researchers were primarily interested in applying psychoanalytical theories to ethnographic data. The school contended that culture molds personality types, especially during child development, and that different cultures produce different personality types as a result of discrepant socialization practices. This school gained prominence during World War II with the American government using "National Character Studies" to understand the character and strategies of the Axis Powers, as exemplified by Benedict's classic work on the Japanese, *The Chrysanthemum and the Sword*. Post-war, the school lost influence for suggesting that all members in any given society act similarly through a prototypical personality, for assuming that culture is static rather than constructed by society's members, and for insufficiently demonstrating robust links between child development practices and adult personality traits.

Current trends in cultural assessments can be traced to the extensive collaboration among cultural mental health experts and interpretive

medical anthropologists since the 1970s. A substantial scholarship has grown from this tradition, reiterating the following principles: (1) formal psychiatric diagnoses may not be meaningful across all societies and empirical data must demonstrate their applicability in each society to avoid a "category fallacy," the assumption that constructs (such as those related to mental health and pathology) developed in one cultural setting are equally valid in other cultural settings; (2) the clinician's interview bears resemblance to the anthropologist's ethnography in attempting to reconcile personal lived experience with formal scientific theory; and (3) a clinical interview can approximate ethnography by understanding the patient's explanatory model (EM) of mental illness, such as its causes, onset, course, mechanism, and severity. This orientation acknowledges that individuals belong and contribute to their cultures without assuming the homogeneity of any individual's cultural experience. These principles have laid the foundation for four main ethnographic interviews used in patient cultural assessments: psychiatrist–anthropologist Arthur Kleinman's explanatory model framework (1980), the Explanatory Model Interview Catalogue (EMIC) (1992), the Short Explanatory Model Interview (SEMI) (1998), and the McGill Illness Narrative Interview (MINI) (2006). These cultural assessments all attempt to identify key domains that can be assessed for all patients irrespective of the patient's cultural background.

Kleinman's Explanatory Model Framework

Arthur Kleinman's work in medical anthropology and cultural psychiatry during the 1970s centered on his concept of EMs, the different notions about an episode of sickness and treatment employed by patients and practitioners. Kleinman contended that the study of interactions among patient and practitioner EMs could introduce a method to understand cultural influences in clinical communication. He identified five themes in illness EMs: (1) etiology; (2) time and mode of onset of symptoms; (3) pathophysiology; (4) course of sickness, including degree of severity; and (5) treatment. Differences in patient–practitioner EMs could be analyzed as the differences of illness and disease, the former referring to the patient's lived experience of a sickness and the latter referring to the practitioner's professional knowledge. Kleinman believed that analyzing patient EMs could illuminate the personal meanings of sickness and help-seeking preferences. His monograph *Patients and Healers in the Context of Culture* included eight questions to elicit EMs as a clinically relevant approach to patient cultural assessment. His framework was influential in cultural psychiatry, primary care, and medical anthropology as a form of readily applied ethnography in clinical settings throughout the 1980s and 1990s. Subsequent patient cultural assessment interviews have variously engaged with Kleinman's EM framework to expand in other directions.

The Explanatory Model Interview Catalogue

Psychiatrist–epidemiologist Mitchell Weiss and his colleagues developed the EMIC as a method for cross-cultural research in psychiatric epidemiology based on the EMs framework. The abbreviation EMIC also invokes the anthropological distinction between *emic* cultural knowledge rooted in local communities and *etic* theoretical knowledge of professionals outside such communities, as in the illness and disease distinction mentioned earlier. In his monograph, Kleinman warned against a mechanical use of EMs that would risk reducing a patient's culture to another piece of decontextualized clinical information such as a lab result or physical exam finding; this may have prevented him from elaborating a research methodology beyond unstructured in-depth interviews. In contrast, Weiss designed the EMIC to operationalize the EM framework with a clear methodology, enable systematic comparisons across sites, and complement traditional diagnostic epidemiology with a form of cultural epidemiology that could analyze relationships of illness understandings and practices with clinical outcomes. Based initially on research with patients suffering from mental illness and leprosy in India, Weiss also created the EMIC to quantitatively rate narrative differences among patients; for example, if a patient answers about a perceived illness cause in direct response to a question rather than in passing during a response to another question. EMIC instruments follow a general question format and set of domains rather than reproducing a single, unvarying interview; in effect, a specific EMIC instrument has to be developed for every research question. Starting from general introductory questions for each domain—which may be similar across EMICs—each iteration of the interview progresses to increasingly specific questions targeting a range of local illness experiences and interpretations. EMICs typically include assessment domains such as (1) patterns of distress, (2) perceived illness causes, (3) help seeking and treatment, (4) general illness beliefs, and (5) disease-specific queries. In addition to operationalizing Kleinman's EM framework, the EMIC introduces this final domain to assess the patient's general understandings about a specific disease. This domain is intended to ascertain the extent to which the patient's individual illness explanations resemble or differ from other explanations present in the patient's culture to which the individual is exposed.

The Short Explanatory Model Interview

Criticisms emerged that the EM framework was underspecified and that the EMIC was too detailed and time consuming. Psychiatric researchers in England therefore developed the SEMI as a standardized, valid tool to assess common mental disorders in large-scale field work and primary care settings. The SEMI covers five sections: (1) the patient's individual and

social background; (2) the nature of the presenting problem; (3) help-seeking behavior, including contact with nonmedical healers; (4) interactions with the clinician/healer to clarify roles and expectations for promoting satisfaction; and (5) illness understandings, covered through a presentation of three vignettes for common mental disorders—depression, phobia, and somatization—followed by open-ended questions to capture the respondent's attitudes to the vignettes. The interview was tested in four primary care sites in England and Zimbabwe. In addition to assessing patient EMs and general views on mental illness, the SEMI expands cultural assessment to inquire about the patient's background and expectations of care as possible avenues of informing current treatment.

The McGill Illness Narrative Interview

The final ethnographic interview comes from psychiatric researchers at McGill University in Canada who have raised concerns that the EM framework may be too unstructured and incomplete for systematic comparison and that the EMIC and SEMI may be too structured for narrative spontaneity. The MINI incorporates medical anthropologist Allan Young's critique that the EM framework assumes logical, coherent accounts of an illness experience. Young's work demonstrated that individuals use multiple representations and modes of reasoning to produce narratives of illness that can be complex, internally inconsistent, and even self-contradictory. Many patients may not be able to stipulate exact causes of an illness, preferring explanations based on illness prototypes common to people in their communities (e.g., learned through popular myths or media accounts) or based on chain complexes in which people reason by association, such as through a chain of events (e.g., that a symptom and a potential cause are related because they happened close together). The elicitation of illness prototypes recognizes broad social influences on an individual's meanings of illness, whereas chain complexes allow for reasoning based on categories salient to the patient rather than predetermined EM domains. To capture such variation, the MINI consists of open- and closed-ended interview components.

The DSM-IV Outline for Cultural Formulation

The four assessment interviews discussed above generated significant interest among mental health experts searching for ways to assess patient cultures in clinical practice. Anthropologists developed a related but independent scholarship demonstrating that the health-care system can also be analyzed as a cultural system and that the act of psychiatric diagnosis is a form of clinical interpretation that owes much to the influence of social networks and institutions. Throughout the 1980s, consensus emerged among mental health investigators and anthropologists

that patients and providers—independently and together—are continually involved in creating and refashioning meanings of health, illness, and treatment. Mental health experts and anthropologists have agreed that such acts of meaning-making merit attention to close gaps between patient and provider understandings during any given sickness. In general, mental health experts have found patient cultural assessments helpful to clarify mismatches between symptom manifestations and official diagnostic criteria or to engage patients optimally in treatment; anthropologists have concentrated on the extent to which illness narratives reflect local patient worlds coming into contact with the biomedical system.

These insights achieved a measure of institutional legitimacy during the DSM-IV revision process. The US National Institute of Mental Health (NIMH) convened a conference in April 1991 at the University of Pittsburgh among cultural experts and DSM-IV task force and work group representatives to enhance DSM-IV's cultural validity. Cultural experts included cultural psychiatrists and psychologists, medical anthropologists, and prominent members of racial and ethnic minority groups within the American Psychiatric Association (APA) leadership. Following this conference, the NIMH Group on Culture and Diagnosis developed recommendations for *DSM-IV* such as an introduction to the upcoming manual on the importance of culture, a section on cultural considerations relevant to each diagnosis, a glossary of "cultural syndromes and idioms of distress," a cultural commentary on *DSM-IV*'s multiaxial evaluation format, and guidelines for a cultural formulation. The NIMH group based the elements of its OCF on a literature review, demonstrating four key domains that were believed crucial during each patient cultural assessment: (1) cultural identity of the individual, (2) cultural explanations of the individual's illness, (3) cultural factors related to the psychosocial environment and levels of functioning, and (4) cultural elements of the patient–clinician relationship. A final domain was included to record how cultural information could influence diagnosis and care. Clinicians could use the OCF to supplement the multiaxial diagnostic assessment and to describe the patient's cultural and social reference groups in clinically relevant ways. Table 29.2–1 lists OCF domains for assessment.

The OCF's inclusion in *DSM-IV* provided a platform for mental health experts and anthropologists around the world to appraise the current state of knowledge, harmonize efforts at patient cultural assessments, and suggest future directions. In contrast to the multiaxial evaluation with its standardized codes, the OCF would permit clinicians to narratively describe the patient's account of suffering from an *emic* perspective rich with contextual details. The use of the word "formulation" in the title was an attempt to stimulate the type of narrative analysis already familiar to psychiatrists trained in psychodynamic formulations who assess details of a patient's psychological conflicts in the context of life experiences. The

formulation was also designed to standardize a methodology for cultural assessment and produce outcome studies in mental health services research. Field trials with various American ethnicities validated its utility and offered tangible clinical examples.

Table 29.2–1.
The DSM-IV Outline for Cultural Formulation

Cultural Formulation Domain	Subtopics for Assessment
Cultural identity of the individual	► Individual's ethnic or cultural reference group(s). ► Degree of involvement with the culture of origin and the host culture (for immigrants and ethnic minorities). ► Language abilities, use, and preference (including multilingualism).
Cultural explanations of the individual's illness	► Idioms of distress through which symptoms or the need for social support are communicated. ► Meaning and perceived severity of the individual's symptoms in relation to norms of the cultural reference group(s). ► Local illness categories used by the individual's family and community to identify the condition. ► Perceived causes and explanatory models that the individual and the reference group(s) use to explain the illness. ► Current preferences for and past experiences with professional and popular sources of care.
Cultural factors related to psychosocial environment and levels of functioning	► Culturally relevant interpretations of social stressors, social supports, and levels of functioning and disability. ► Stresses in the local social environment. ► Role of religion and kin networks in providing support.
Cultural elements of the relationship between the individual and the clinician	► Individual differences in culture and social status between the individual and the clinician. ► Problems that these differences may cause in diagnosis and treatment (e.g., difficulties in eliciting symptoms and understanding their cultural significance, in determining whether a behavior is normal or pathological).
Overall cultural assessment for diagnosis and care	► Discussion of how cultural considerations specifically influence comprehensive diagnosis and care.

Inclusion of the OCF in *DSM-IV*, however, was problematic. Although the NIMH group advocated for its placement at the beginning of the Manual, the OCF was placed in the ninth appendix. It was also combined with the *Glossary of Cultural Syndromes and Idioms of Distress*, renamed by the *DSM-IV* leadership the *Glossary of Culture-Bound Syndromes*, a decision that tended to exoticize the OCF, as if it should be used only with unusual cultural presentations rather than with every patient. Despite these obstacles, the OCF has been used worldwide since its publication in 1994. While much of this work has found the OCF a useful clinical tool, problems

have also been identified. Its format as a description of domains to be assessed, lack of implementation guidelines, and brevity—which required that clinicians already be familiar with cultural topics—have hindered widespread clinical use. Nonetheless, the OCF remains the most widely taught method of cultural assessment in US psychiatric residency programs. In addition, psychiatrists and psychologists in Canada, Denmark, England, the Netherlands, Sweden, and the United States have independently operationalized the OCF's implementation, resulting in services ranging from brief semistructured interviews for individual clinicians to use with every patient to cultural consultation services receiving referrals from other health professionals. A major contribution of the OCF has been to organize subsequent research along its four domains of cultural assessment, facilitating the revision from the OCF in *DSM-IV* to the CFI in *DSM-5*.

KEY DOMAINS IN PATIENT CULTURAL ASSESSMENTS

Cultural Identity of the Individual

Cultural psychiatrists contend that understanding a patient's identity is indispensable to a thorough clinical evaluation. Early OCF guidelines suggested that cultural identity assessments be conducted through an "interpersonal grid" to record potentially relevant variables such as ethnicity, race, geographical origin, language, acculturation, gender, age, sexual orientation, religious or spiritual beliefs, socioeconomic class, and education. The NIMH group raised concerns, however, of stereotyping patients or minimizing sociocultural phenomena based on a clinician's determination of the patient's identity. Others have advocated a more comprehensive approach that accounts for changes in self-reported identity based on context (i.e., family roles vs. work responsibilities) that also includes the patient's perception of the clinician's identity. Assessments of cultural identity should help the clinician discern the effects of identity, stigma, and language on symptom experience, expression, and interpretation in particular cultural contexts and community settings.

Assessments of patient cultural identities do not come easily or intuitively for many clinicians. Educational studies show that psychiatric trainees feel uncomfortable and invasive in asking patients about their identities without prior training. Nearly all interviews that have striven to operationalize the OCF include common elements to assess identity such as the cultural groups with which the patient identifies; the relative levels of importance and meaning ascribed by the patient to these cultural identifications; assessments of language fluency and preferences; and assessments of a patient's involvement with the culture of origin and the host culture in the case of immigrants, especially, including positive and negative experiences with involvement in both cultures. Aspects of identity

included only in some OCF-based interviews included evaluations of whether patients experience multiple identities simultaneously or over time and how cultural identity impacts the clinical problem. Several interviews also included detailed migration histories. All OCF-based interviews have used open-ended questions to encourage interaction and discussion, rather than a checklist of features for cultural identity. All authors also warn against an oversimplification of cultural factors in which social processes such as adverse experiences associated with migration and discrimination are mistakenly understood as inherent, representative, or relevant to the current encounter, rather than as time-limited and context-driven reactions to social predicaments.

Studies published since the OCF have highlighted unique challenges for certain clinical populations. Some refugees may be suspicious of questions on identity and many migrants may not know how to discuss their backgrounds in new societies as compared to native minorities. Children and adolescents may also have different cultural identities than their parents depending on mixed parentage, adoption, and socialization or acculturation. The cultural identity of older patients may be shaped by cultural meanings of old age based on their social roles and status. Cultural identity should be explored with reference to social context, developmental phase, and with the possibility of change over time.

Cultural Explanations of the Individual's Illness

EMs of illness have underpinned most studies on patient cultural assessments in mental health and anthropology. While Kleinman's (1980) initial description of EMs sought to obtain explicit information about causality, mechanisms, and anticipated outcomes, a growing literature has documented that ways of thinking about and narrating illness may not be captured by assessments that assume cognitive linearity. Patient explanations of illness may also be characterized by metaphorical thinking, emphasizing elements without a clear cause; sequences of events that occurred in temporal and spatial proximity (linked by contiguity or metonymy); and images or examples of illness that one encounters to reason analogically about the current illness episode. EMs can change based on context, including the interviewer and the setting in which they are elicited; interactions with close associates, the community, or the health-care system; and acculturation status for immigrants and refugees. Attempts to code illness narratives from patients with medically unexplained symptoms revealed three types of EMs: *nominal*, in which a condition was named without elaborating on causal attribution or mechanism; *causal attributions*, in which a symptom or illness was linked to a cause; and *process or mechanism explanations*, in which mechanistic or metaphorical devices linked cause and symptom.

Furthermore, culture and context influence the elicitation of EMs.

Clinical narratives reflect broader cultural norms affecting communications with mental health-care providers, family, and friends. Access to the illness representations of crucial members in the person's extended social network (either directly or as reported by the patient) provides a range of possible views influencing the person's experience. The organization of health-care services in a society may also determine patients' attitudes toward disease, diagnosis, and treatment based on familiarity with the health-care system.

Eliciting past treatment experiences and current treatment expectations have also been deemed as important components of EMs. Cultural scripts associate a patient's illness representations with specific patterns of self-coping and help seeking, including the perceived relevance and efficacy of mental health treatments and their potential dangerousness. Cultural adaptations of evidence-based therapies that take account of cultural scripts may improve patient engagement and response, highlighting the value of identifying cultural interpretations prior to treatment planning. Nonmedical forms of self-coping and help seeking are common, diverse, and merit clinical attention because they may complement or conflict with formal psychiatric care. Barriers to utilizing mental health services are also determined by cultural priorities for care, family values and interactions, and local health system and socioeconomic factors affecting access to mental health services. A sound cultural assessment strives to situate EMs against relevant social and health contexts of the patient.

Early theoretical guidelines on the OCF included the *culture-bound syndromes* in this section. *DSM-5* has since substituted the term "culture-bound syndrome" with three other terms of greater clinical utility: *cultural syndromes, cultural idioms of distress,* and *cultural explanations*. These terms denote the ways that suffering, behavioral problems, and troubling thoughts and emotions are experienced, understood, and communicated, either as symptom clusters, as ways of expressing distress that may not involve specific symptoms, and/or as explanations or attributions. Key aspects of patients' clinical presentations may be patterned by such cultural concepts.

Nearly all OCF-based interviews include questions on illness-related idioms, the meaning and severity of symptoms in relation to cultural norms, causes, expected course of illness, experiences with care, and treatment expectations. Nearly all interviews ask the patient to describe the views of close associates such as family and friends in order to situate the individual's perspective in community context. The instruments vary in levels of assessing how the patient describes the clinical problem to close associates, the most troubling aspects of the clinical problem, the contribution of illness prototypes within a given culture to the patient's own understandings, the most useful treatment received, treatment recommendations of close associates, and the patient's use of traditional healers and alternative services.

Cultural Factors Related to Psychosocial Environment and Levels of Functioning

This domain encourages clinicians to clarify the context of a patient's illness—especially cultural interpretations of the stressors and supports associated with onset and course—and fears and hopes affecting levels of functioning. OCF-based instruments have used open-ended questions to obtain a detailed picture, with some instruments asking questions directly of the patient and others instructing clinicians to identify these data based on patient responses. In all cases, the goal is to obtain a description of the life context in which illness emerges and to elicit responses from the patient, the family, and the community. The subjective, lived experience of the patient's and social network's response is crucial because culture shapes the perceived centrality, magnitude, and treatability of the problem as well as potential forms of stress and support. This approach synthesizes the person's subjective experience of risk, resilience, and protective factors and the clinician's analysis of the patient's social and cultural context. For example, the subjective–objective approach might rely on a functional assessment of key activities of daily living that includes the patient's and social network's cultural interpretations of normative functioning.

The OCF was partly developed to address the adverse experiences of historically disadvantaged racial and ethnic minorities. Since its publication, studies have elaborated additional types of social stressors. A recurrent finding has been the effect of migration status on the mental health of refugees as a potent social determinant of health independent of any premigratory trauma. The factors most associated with risk for psychopathology include low socioeconomic status, poor housing, unemployment or under-employment, lack of residency status, linguistic barriers, limited social networks, racial or ethnic discrimination, role strain, family conflict, status loss, acculturative stress, nostalgia, and bicultural tension. Acculturation issues may also affect racial/ethnic minorities who feel disempowered or marginalized, not just immigrants. Discrimination-related stress affects members of other social groups, such as religious minorities, Lesbian–Gay–Bisexual–Transgender persons, and low-socioeconomic status and disability communities whose experiences of stress and coping also merit assessment.

In clinical assessment, family, friends, and close associates can provide valuable information about cultural context. Collateral information from relevant members of the patient's social network is especially critical for children and adolescents, older adults, and those with any condition affecting cognitive functioning such as dementia or florid psychosis that could interfere with the clinical interview. OCF-based interviews have varied in the extent to which they seek collateral forms of information, with some protocols for children and adolescents explicitly identifying multiple sources of information and perspectives. Although patient cultural

assessments such as the OCF and its related interviews can accommodate use of medical interpreters and cultural brokers, there is little guidance on how to operationalize this process in actual practice.

Finally, psychiatrists have increasingly reported that patients frequently turn to religion, spirituality, and other secular moral traditions in understanding and responding to mental illness. A "religiosity gap" may exist among clinicians trained in empirical/positivistic scientific methods and patients searching for holistic treatments, especially within complementary and alternative systems. Cultural assessment should attend to the meanings of religion and other moral systems in crafting illness assumptions and beliefs, especially when a religious community can provide emotional, social, and material support. Clinicians may be able to mobilize such information with patients to support recovery through groups and organizations aligned with these traditions.

OCF-based interviews have assessed stressors, supports, and levels of functioning differently. Variations include focusing on the impact of stressors, the views of the social network, the role of the person's partner, forms of religious/spiritual or moral support, barriers to receiving help, and the views of the social network on subsequent levels of functioning.

Cultural Elements of the Relationship between the Individual and the Clinician

Many OCF case studies have illustrated the challenges confronting clinicians who provide cross-cultural care. These included tensions between biomedical forms of treatment and other forms of healing, the failure of health systems to provide interpreters, and difficulties in cross-cultural communication. Clinicians have frequently discussed their reactions to simultaneously considering the patient as an individual and a member of a social group. Despite the best efforts of even the most sincere clinicians, the medical encounter may be influenced by discrimination, racism, and subtle forms of cross-cultural bias. This domain of the OCF provides an opportunity to consider how such experiences affect patients' perceptions of care and promote clinician self-reflection on the challenges of delivering culturally sensitive health care.

Existing OCF-based interviews address this domain differently. Some provide queries to the clinician in the form of introspection such as questions on mutual perceptions of power and position of the patient and clinician, including the history of relationships between their respective cultural groups. This work operationalizes the concepts of "cultural transference" and "cultural countertransference" to analyze patient–physician relationships. Other instruments suggest questions that can be posed to the patient, such as the quality of communication during the interview, the need for interpreters, important topics missed by the interview, the perceived impact of ethnic or cultural matching with the clinician or lack thereof, and the role of the patient during the interview.

Although a sensitive topic, early, open inquiry into past negative experiences of care and potential miscommunications with clinicians may facilitate rapport, engagement, and a commitment to overcome treatment barriers.

THE CULTURAL FORMULATION INTERVIEW IN DSM-5

Nearly 20 years of experience with the OCF, its four domains, and its related interviews since 1994 led to review and revision of the OCF for *DSM-5*. The *DSM-5* Cross-Cultural Issues Subgroup (DCCIS) intended to develop a set of corresponding semistructured interviews to make the OCF more operational for clinicians. This set of interviews is known collectively as the CFI and consists of a core questionnaire of 16 items, an informant version of the core CFI to obtain collateral material from caregivers, and 12 supplementary modules that expand the basic cultural assessment. The APA has made these instruments available for free at http://www.psych.org/practice/dsm/dsm5/online-assessment-measures.

The DCCIS formed part of the larger *DSM-5* Gender and Cross-Cultural Issues Study Group and was composed of experts in culture and mental health from diverse countries and cultural groups. One explicit goal of the DCCIS was to improve the assessment of culture within psychiatric diagnosis. The subgroup was charged by the *DSM-5* task force with making recommendations on racial, ethnic, cultural, and contextual issues to the *DSM-5* work groups, such as differences in risk factors, precipitants, symptom presentations, prevalence, symptom severity, and course of illness for each diagnosis.

In addition, the DCCIS undertook a critical review of the *DSM-IV* OCF. DCCIS participants prepared literature reviews on the OCF as a whole, each of its four domains, and key implementation questions raised by use of the OCF in clinical care. For example, such topics included best practices; special populations in need of attention such as children, the elderly, or immigrants and refugees; and the format and content of previously developed interviews, protocols, and questionnaires operationalizing the Outline. DCCIS members were asked to provide recommendations on how to revise the OCF, including how to make it easier to implement through the CFI.

Based on information obtained from these literature reviews and the previous experiences of DCCIS members developing and using cultural assessments guided by the OCF, the OCF was revised slightly for *DSM-5*, mostly by making the *DSM-IV* text more explicit in places. Examples of this clarification process include the incorporation of elements of cultural identity not mentioned in *DSM-IV* such as religious affiliation and sexual orientation, among others; clinician instruction to identify stressors and supports in the person's social environment and their interpretations against the patient's cultural reference groups; and mention of the potential

impact of racism and discrimination on the clinician–patient relationship.

More substantial changes related to making the OCF more operational. An initial 14-item interview was developed largely from previous question lists, interview protocols, and full interviews that operationalized the OCF; most of the authors of the OCF-based approaches were members of the DCCIS. The content of the questions was formulated, discussed, and debated in 2010 and 2011 in biweekly conference calls, supported by the *DSM-5* development process. The resulting 14-item version of the CFI was tested in a field trial from November 2011 to November 2012 at 12 clinical sites in the US, Canada, Peru, the Netherlands, Kenya, and India. This trial utilized a standardized training format (http://www.ucl.ac.uk/ccs/specialist-services/#cfi) and proceeded according to a common research design to evaluate the CFI's acceptability, feasibility, and perceived clinical utility among patients and clinicians. Field trial results were used to finalize the 16-item version in *DSM-5*. An informant version of the CFI was prepared to elicit collateral information from parents and other close associates. Supplementary modules were also developed for more comprehensive assessment. Some modules expand various sections of the core CFI, and others were developed for specific populations, including school-age children and adolescents, elderly individuals, immigrants and refugees, and caregivers. This range of instruments enables clinicians to tailor the comprehensiveness of the assessment for particular patients and clinical circumstances.

Definition of Culture

To facilitate application of the interview, the CFI guidelines provide a definition of *culture* as (1) the values, orientations, knowledge, and practices that individuals derive from membership in diverse social groups (e.g., ethnic groups, faith communities, occupational groups, and veterans groups); (2) aspects of an individual's background, developmental experiences, and current social contexts that may affect his or her perspective, such as geographical origin, migration, language, religion, sexual orientation, or race/ethnicity; and (3) the influence of family, friends, and other community members (the individual's social network) on the individual's illness experience.

This definition emphasizes the comprehensive effect of culture on a person as well as the key role of the family and close associates in patterning the cultural environment that shapes illness experience and help-seeking expectations. In US health care practice, notions of culture are frequently paired exclusively with racialized/ethnic categories, which may lead to the unintended consequence of stereotyping patients. Other components of culture—for example, age cohort, occupation, and sexual orientation—are often not included under the purview of a cultural assessment. By contrast, the CFI guides clinicians to see the contextual

frame of each patient's experience—seeking to use information about the collective to understand an individual's perspective and to clarify how local environments impinge on the person's situation, including how sociocultural contingencies help pattern a set of events. In addition, the CFI focuses the cultural formulation on the patient's self-identified group rather than on visible group markers of identity that may inadvertently lead to stereotyping.

The Core CFI

Table 29.2–2 presents the core CFI interview. The interview has been designed like other semistructured interviews such as the Structured Clinical Interview for DSM Disorders (SCID) in which instructions to clinicians appear next to questions for patients. As emphasized in *DSM-5*, the CFI follows a person-centered approach to cultural assessment by asking the patient about his or her views and those of close associates to avoid stereotyping by race, ethnicity, or clinical expectations for a given illness. Clinicians are encouraged to voice the questions in their own clinical style and to follow up with additional probes as necessary to understand patient answers completely. The CFI can be used in the initial assessment of patients in all clinical settings, regardless of the cultural backgrounds of the patient or the clinician. Patients and clinicians who seem to share the same culture may differ in ways relevant to care such as by educational level, socioeconomic status, or role expectations. The CFI may be used in its entirety or incorporated piecemeal as needed; however, the developers recommend obtaining the person's story as an unbroken flow before proceeding to the rest of the clinical interview. The CFI may be especially helpful when there is difficulty in diagnostic assessment due to differences in the cultural, religious, or socioeconomic backgrounds of the clinician and the patient; when there is uncertainty about the fit between patient symptoms and diagnostic criteria; when it is difficult to judge illness severity or impairment; when patient and clinician disagree on the course of care; or in cases of suboptimal treatment engagement and adherence. In response to criticisms about the absence of guidelines for the OCF, the DCCIS offered instructions to clinicians to facilitate patient cultural assessments.

Table 29.2–2.
The DSM-5 Cultural Formulation Interview (CFI)

Supplementary Modules Used to Expand Each CFI Subtopic Are Noted in **Bold.**

GUIDE TO INTERVIEWER	*INSTRUCTIONS TO THE INTERVIEWER ARE ITALICIZED*
The following questions aim to clarify key aspects of the presenting clinical problem from the point of view of the patient and	*INTRODUCTION FOR THE PATIENT*: I would like to understand the problems that bring you here so that I can help you more

other members of the patient's social network (i.e., family, friends, or others involved in current problem). This includes the problem's meaning, potential sources of help, and expectations for services.	effectively. I want to know about *your* experience and ideas. I will ask some questions about what is going on and how you are dealing with it. Please remember there are no right or wrong answers.
CULTURAL DEFINITION OF THE PROBLEM	
CULTURAL DEFINITION OF THE PROBLEM ***Explanatory Model, Level of Functioning***	
Elicit the patient's view of core problems and key concerns. *Focus on the patient's own way of understanding the problem.* *Use the term, expression, or brief description elicited in question 1 to identify the problem in subsequent questions (e.g., "your conflict with your son").*	1. What brings you here today? *IF PATIENT GIVES FEW DETAILS OR ONLY MENTIONS SYMPTOMS OR A MEDICAL DIAGNOSIS, PROBE:* People often understand their problems in their own way, which may be similar to or different from how doctors describe the problem. How would *you* describe your problem?
Ask how patient frames the problem for members of the social network.	2. Sometimes people have different ways of describing their problem to their family, friends, or others in their community. How would you describe your problem to them?
Focus on the aspects of the problem that matter most to the patient.	3. What troubles you most about your problem?
CULTURAL PERCEPTIONS OF CAUSE, CONTEXT, AND SUPPORT	
CAUSES ***Explanatory Model, Social Network, Older Adults***	
This question indicates the meaning of the condition for the patient, which may be relevant for clinical care.	4. Why do you think this is happening to you? What do you think are the causes of your [PROBLEM]?
Note that patients may identify multiple causes depending on the facet of the problem they are considering.	*PROMPT FURTHER IF REQUIRED:* Some people may explain their problem as the result of bad things that happen in their life, problems with others, a physical illness, a spiritual reason, or many other causes.
Focus on the views of members of the patient's social network. These may be diverse and vary from the patient's.	5. What do others in your family, your friends, or others in your community think is causing your [PROBLEM]?
STRESSORS AND SUPPORT ***Social Network, Caregivers, Psychosocial Stressors, Religion and Spirituality, Immigrants and Refugees, Cultural Identity, Older Adults, Coping and Help Seeking***	
Elicit information on the person's life context, focusing on resources, social supports, and resilience. May also probe other supports, e.g., from co-workers, or participation in religion or spirituality.	6. Are there any kinds of support that make your [PROBLEM] better, such as support from family, friends, or others?
Focus on stressful aspects of the person's environment. Can also probe, e.g., relationship problems, difficulties at work or school, or discrimination.	7. Are there any kinds of stresses that make your [PROBLEM] worse, such as difficulties with money, or family problems?
ROLE OF CULTURAL IDENTITY	

Cultural Identity, Psychosocial Stressors, Religion and Spirituality, Immigrants and Refugees, Older Adults, Children and Adolescents

Sometimes, aspects of people's background or identity can make their [PROBLEM] better or worse. By background or identity I mean, for example, the communities you belong to, the languages you speak, where you or your family are from, your race or ethnic background, your gender or sexual orientation, and your faith or religion.

Ask the patient to reflect on the most salient elements of his/her cultural identity. Use this information to tailor questions 9–10 as needed.

8. For you, what are the most important aspects of your background or identity?

Elicit aspects of identity that make the problem better or worse.
Probe as needed (e.g., clinical worsening as a result of discrimination due to migration status, race/ethnicity, or sexual orientation).

9. Are there any aspects of your background or identity that make a difference to your [PROBLEM]?

Probe as needed (e.g., migration-related problems; conflict across generations or due to gender roles).

10. Are there any aspects of your background or identity that are causing other concerns or difficulties for you?

CULTURAL FACTORS AFFECTING SELF-COPING AND PAST HELP SEEKING

SELF-COPING

Coping and Help Seeking, Religion and Spirituality, Older Adults, Caregivers, Psychosocial Stressors

Clarify self-coping for the problem.

11. Sometimes people have various ways of dealing with problems like [PROBLEM]. What have you done on your own to cope with your [PROBLEM]?

PAST HELP SEEKING

Coping/Help Seeking, Religion/Spirituality, Older Adults, Caregivers, Psychosocial Stressors, Immigrants/Refugees, Social Network, Clinician--Patient Relationship

Elicit various sources of help (e.g., medical care, mental health treatment, support groups, work-based counselling, folk healing, religious or spiritual counselling, or other forms of traditional or alternative healing).
Probe as needed (e.g., "What other sources of help have you used?").
Clarify the patient's experience and regard for previous help.

12. Often, people look for help from many different sources, including different kinds of doctors, helpers, or healers. In the past, what kinds of treatment, help, advice, or healing have you sought for your [PROBLEM]?
PROBE IF DOES NOT DESCRIBE USEFULNESS OF HELP RECEIVED:
What types of help or treatment were most useful? Not useful?

BARRIERS

Coping and Help Seeking, Religion and Spirituality, Older Adults, Psychosocial Stressors, Immigrants and Refugees, Social Network, Clinician—Patient Relationship

Clarify the role of social barriers to help seeking, access to care, and problems engaging in previous treatment.
Probe details as needed (e.g., "What got in the way?").

13. Has anything prevented you from getting the help you need?
PROBE AS NEEDED:
For example, money, work or family commitments, stigma or discrimination, or lack of services that understand your language or background?

CULTURAL FACTORS AFFECTING CURRENT HELP SEEKING PREFERENCES	
Social Network, Caregivers, Religion and Spirituality, Older Adults, Coping and Help Seeking	
Clarify patient's current perceived needs and expectations of help, broadly defined. *Probe if only lists one source of help (e.g., "What other kinds of help would be useful to you at this time?").*	Now let's talk some more about the help you need. 14. What kinds of help do you think would be most useful to you at this time for your [PROBLEM]?
Focus on the views of the social network regarding help seeking.	15. Are there other kinds of help that your family, friends, or other people have suggested would be helpful for you now?
CLINICIAN–PATIENT RELATIONSHIP	
Clinician—Patient Relationship, Older Adults	
Elicit possible concerns about the clinic or the clinician–patient relationship, including perceived racism, language barriers, or cultural differences that may undermine goodwill, communication, or care delivery. *Probe details as needed (e.g., "In what way?").* *Address possible barriers to care or concerns about the clinic and the clinician–patient relationship raised previously.*	Sometimes doctors and patients misunderstand each other because they come from different backgrounds or have different expectations. 16. Have you been concerned about this and is there anything that we can do to provide you with the care you need?

The core CFI is divided into four main domains. The first domain is the *Cultural Definition of the Problem* and consists of the first three questions. Question #1 is meant to be a broad, open-ended inquiry about the patient's presenting concerns. The prompt invites patients to describe their problems and emphasizes an understanding of the patient's illness, even if these descriptions are "similar to or different from how doctors describe the problem." Question #2 clarifies the patient's description of the problem as would be discussed with close associates, in recognition that culture influences the sharing of information according to audience and social context. At this point in the interview, patients may be reticent to describe the problem in nonbiomedical ways that the clinician could find unusual and may therefore censor such descriptions. By framing the question in relation to the person's social network, the goal is to facilitate the patient's report of as wide a range of descriptions as are salient in the immediate social environment. Question #3 on what troubles the patient most about the problem is designed to ask about what is most at stake for the individual during the illness experience.

Questions #4 to 10 belong to the second domain of the core CFI, *Cultural Perceptions of Cause, Context, and Support*. Questions #4, 6, and 7 focus the interview on the patient's EMs for the causes, social stressors, and social supports of the illness. Question #4 includes the placeholder "[PROBLEM]" to solicit the patient's own terms and phrases for how a problem is framed in the clinical setting, known as *idioms of distress*. Once

obtained, the patient's vocabulary can be substituted for the placeholder throughout the interview to build rapport and bridge patient–clinician linguistic differences in understanding the illness. In some cases, the patient may use biomedical vocabulary such as "I might have depression," but in other cases psychosocial terms may be evident, such as "I worry about my rent, electricity, and phone bill." Terms communicating the presenting problem may vary greatly, and the CFI can reveal how patients frame issues so that clinicians can tailor treatment strategies most effectively.

Question #5 asks the patient to consider what close associates may understand as the cause of illness in response to the understanding of culture in *DSM-5* as the meanings and viewpoints transmitted within social groups. In this instance, the meanings being transmitted are the understandings of illness and the social group is the patient's network of close associates. As with Question #2 on the definition of the problem, framing the inquiry through members belonging to the social network encourages report of a fuller range of explanations that may influence the individual's EMs.

Questions #8, 9, and 10 also belong to this second domain, but orient the interview toward the patient's cultural identity. Question #8 begins with a brief introduction that clinicians can use to clarify the meaning of *cultural background or identity* as relevant to health care. The introduction provides examples of identities that may be important for the patient, such as language, ethnicity, religion, and sexual orientation. These examples are meant to be representative in order to stimulate thought, not a fixed set of identities from which the patient must choose. Question #8 then asks the patient directly about the most important aspects of background or identity. This method of allowing the patient to name his or her cultural identity departs from previous models of cultural competence that have often made assumptions about patient identities based on a group affiliation such as race or ethnicity without accounting for their importance to the individual. For example, one patient in the field trial initially named his identity first as "Cuban" but then explained that identifying as a Christian was more helpful in understanding his suffering existentially. A model of cultural competence that assumed ethnicity was most the important group affiliation—without accounting for the patient's self-ascribed identity—could have overlooked the role of religion in providing meaning to the illness experience. Question #9 asks the patient how this identity may impact the current illness experience or patient-centered definition of the clinical problem. Question #10 asks the patient to consider how identity may cause other problems in life that may not initially come to mind as related to the illness, but may be nonetheless important to understand, such as problems with migration, gender roles, or intergenerational conflict.

Questions #11, 12, and 13 form the third domain known as *Cultural Factors Affecting Self-Coping and Past Help-Seeking*. This domain encourages patients to share past forms of self-coping and help seeking in recognition that most forms of help are sought outside of the biomedical health-care system. Question #11 addresses the patient's coping practices and question #12 considers help seeking broadly to include help within and outside of the biomedical system (i.e., religion-based support and support groups). An additional prompt for question #12 clarifies the types of help that have been most and least helpful. This information may aid the clinician in developing a treatment plan for the current illness episode. Question #13 asks about past barriers to treatment to help devise the current treatment plan around available resources.

The final three questions constitute the last domain of the CFI, *Cultural Factors Affecting Current Help Seeking*. Question #14 asks the patient about current treatment preferences and question #15 explores treatment preferences that may be recommended by close associates. As with the questions on patient terms and causes for the presenting problem, this question on treatment preferences is examined at the individual and social levels to explore how patients individually fashion cultural meanings within their social groups. Finally, question #16 begins with an open-ended statement that can help the patient to anticipate how perceived differences with the clinician can adversely affect care. By asking the patient about this potential barrier, the clinician can validate patient concerns and work to resolve such differences. Even if the question is not answered during the initial visit, the clinician's openness to the topic may empower the patient to raise concerns about the patient–clinician relationship later in the treatment. The core CFI adopts an ecological approach to culture by first asking the patient about the illness, moving outward to the views of the social network and the role of social supports and stressors, and then finishing broadly with the illness in relation to the health-care system.

The Informant CFI

The goal of the informant CFI (I-CFI) is to guide clinicians in conducting a cultural assessment of the patient's clinical problem from the perspective of a key informant such as a spouse, relative, or friend who may be present during the clinical encounter. Patients are often accompanied by family or other potential informants during clinic visits, providing clinicians with the opportunity to include them during psychiatric diagnosis and treatment planning. Social relationships can have profound effects—positive and negative—on pathways to and through mental health treatment, including decisions about when or whether to seek care, participation in clinic visits, and implementation of treatment plans in the home such as adherence to medication. The expected roles of close associates and their perspectives on illness and treatment are rooted in culture.

Clinical care often requires consultation with family members, other kin, or friends of those with mental illness. Certain aspects of culture may be shared among patients and others within the social network, but there may be important and unexpected differences even within families. For example, an older immigrant man who is depressed and most comfortable speaking his native language may be accompanied by an adult child who speaks English and has adopted mainstream American cultural values. In such cases, the patient and the informant may have different EMs for the patient's depression or view of treatment. In such situations, the I-CFI can be a tool that assists clinicians in eliciting the perspectives of key informants through the lens of culture, enabling a deeper understanding of the patient's illness and the interpersonal dimensions of care. This can open up opportunities for clinical negotiation and the development of shared understandings of illness and goals for treatment.

The I-CFI can play an especially important role in several clinical situations when an informant is available. In some instances, the informant may be the primary or sole source of information to assess the role of cultural factors in psychiatric diagnosis and treatment. This may occur when the clinician assesses a patient unable to give a meaningful or coherent history because of conditions such as severe cognitive impairment (as due to a neurodegenerative brain process such as Alzheimer disease or a head injury), catatonia, severe paranoia, or agitation. Alternatively, the patient may be a young child, and it may be necessary to conduct a cultural assessment of the parent or other caregiver's perspective. In other cases, the clinician may seek clarification from the informant when the patient's responses to the CFI are ambiguous or partial (as when asking the patient about how others in his or her cultural group would view the symptoms). The clinician may also administer the I-CFI if there is an identified family member, other kin, or friend who is influential and active in the patient's clinical care and decision making. Finally, the I-CFI may be useful when treatment hits an impasse and a fuller assessment of the social context may identify underlying causes.

In choosing whether to use the I-CFI, the clinician should consider several issues. First, the clinician should discuss involvement of the informant with the patient and respect all desires for autonomy and privacy. Patients may not want a family member involved due to concerns about burdening family members, shame about certain aspects of the problem, or discomfort with discussing conflicts with the informant that are contributing to the patient's problem. In these situations, the clinician may decide not to involve the informant at all or opt to negotiate family involvement over time after the patient and clinician have established rapport. In other situations, as when the patient insists on having family members present during the clinical encounter and defers to them, the clinician may need to negotiate time during the visit to meet individually

with the patient to allow discussion of sensitive issues that may involve family members (i.e., cases of conflict or abuse). After the informant interview, the clinician may need to decide whether to interview the family member (or other informant) in the presence of the patient and vice versa. This decision should be based on patient preferences.

A related issue is that the clinician may find important differences in the views of patients and informants (i.e., in their illness EMs or the extent to which behavior is viewed as normative for the cultural group). One approach to such situations is to present the different perspectives in a nonjudgmental way, together with the biomedical view of illness, in order to try and negotiate a partially shared understanding of illness and treatment. When multiple perspectives cannot be resolved in this way, documenting the differences may assist the clinician in seeking additional information to clarify diagnostic issues or help in anticipating obstacles to implementing the treatment plan.

Table 29.2–3.
The CFI Supplementary Modules

Number	Module
1	Explanatory model
2	Level of functioning
3	Social network
4	Psychosocial stressors
5	Spirituality, religion, and moral traditions
6	Cultural identity
7	Coping and help-seeking
8	Patient–clinician relationship
9	School-age children and adolescents
10	Older adults
11	Immigrants and refugees
12	Caregivers

It is also essential to note that the clinician has flexibility in the timing and administration of the I-CFI. For example, the I-CFI could be used at the initial visit or later during treatment. The clinician may use all of the I-CFI or only parts most apparently relevant to clinical care. In situations with multiple family members or caregivers, pragmatic decisions must be made about whom to interview based on patient preferences and the informant's actual involvement in the life of the patient. Finally, the supplementary modules may be useful in conjunction with the I-CFI, particularly the caregiver module, which further explores the role of cultural factors in caregiving. The I-CFI can be accessed in *DSM-5* and at the website noted earlier.

The CFI Supplementary Modules

The final set of CFI assessments is the supplementary modules, listed in Table 29.2–3. The core CFI operationalizes a basic cultural assessment, but the need for supplementary modules may arise when certain cultural domains are not addressed or assessed too succinctly during the current clinical encounter. In other circumstances, the core CFI may not be needed as much as a specific supplementary module. Supplementary modules based on special populations such as children and adolescents, immigrants, and older adults include questions tailored for their cultural needs compared to the core CFI. Alternatively, a clinician can administer the EM module to learn more about the patient's understandings of the problem or the spirituality, religion, and moral traditions module to better grasp the meanings of suffering. These examples suggest how 12 supplementary modules can be utilized with the core CFI and the informant version or as part of a stand-alone interview.

All supplementary modules possess a common format. The modules begin by noting common questions with the core CFI. Instructions for administering the module are provided, followed by an introduction for the clinician to the patient. Most modules are organized by a specific subdomain, further assessed through several questions. For example, in the cultural identity module, the subdomains include language ability and migration history, assessed by 6 and 9 questions, respectively. Like the core CFI, the supplementary modules are mostly problem focused, attempting to elicit how the module topic relates to the presenting clinical problem. Some, like the cultural identity module, focus more on aspects of the patient's identity than the presenting problem. Other modules are focused on the presenting problem, such as the EM and coping and help-seeking modules. In some cases, questions in one module are assessed in another based on topical interrelationships.

FUTURE DIRECTIONS

This chapter has presented a historical overview on patient cultural assessments, from Kraepelin's initial attempts at a cross-cultural comparative psychiatry over a century ago to the CFI in *DSM-5*. Common themes in patient cultural assessments since the 1970s have included a focus on the patient's cultural background and identity, explanations of illness, levels of psychosocial support and functioning, and interactions with help seeking, including with the biomedical system. The CFI represents the most international attempt at standardizing a pertinent set of questions for patients and clinicians in all clinical encounters with the assumption that culture constantly and ubiquitously frames sickness and care.

Members of the DCCIS have continued to expand work on cultural

assessments beyond *DSM-5*. The CFI presents a brief list of questions designed to elicit clinically relevant cultural information within practice constraints. Research is currently underway on the comparative effectiveness of the CFI with other forms of resource-intense cultural assessments such as cultural consultation services in changing diagnosis and recommended treatment modalities. The CFI also reflects an attempt to mobilize contemporary notions of culture from anthropology and the interpretive social sciences, but it is not clear that these notions apply in all clinical settings. Ongoing partnership with collaborators from anthropology, psychology, and sociology may illustrate new domains necessary for thorough cultural assessments that can be incorporated in future revisions. Studies on CFI implementation in educational and clinical settings are needed to understand which questions obtain the most pertinent cultural information applicable to diagnostic and treatment planning and how participants view the CFI experience. Finally, outcome studies measuring the CFI's effects on patient symptoms, quality of life, and functional assessment are needed to convincingly make the case for cultural assessments.

References

Aggarwal NK. Hybridity and intersubjectivity in the clinical encounter: impact on the cultural formulation. *Transcult Psychiatry*. 2012;49:121.

Aggarwal NK, Nicasio AV, DeSilva R, Boiler M, Lewis-Fernández R. Does the Cultural Formulation Interview (CFI) for the Revision of the DSM-5 affect medical communication?: a qualitative exploratory study from the New York site. *Ethnicity Health*. 2015;20:1.

Aggarwal NK, Pieh MC, Dixon L, Guarnaccia PJ, Alegria M, Lewis-Fernández R. Clinician descriptions of communication strategies to improve treatment engagement by racial/ethnic minorities in mental health services: a systematic review. *Patient Educ Couns*. 2016;99:198–209.

Alarcón RD. Culture, cultural factors and psychiatric diagnosis: review and projections. *World Psychiatry*. 2009;8:131.

Eisenberg L. Disease and illness. Distinctions between professional and popular ideas of sickness. *Cult Med Psychiatry*. 1977;1:9.

Good B. *Medicine, Rationality and Experience: An Anthropological Perspective*. Cambridge: Cambridge University Press; 1994.

Groleau D, Young A, Kirmayer LJ. The McGill Illness Narrative Interview (MINI): an interview schedule to elicit meanings and modes of reasoning related to illness experience. *Transcult Psychiatry*. 2006;43:671.

*Kirmayer L. Cultural psychiatry in historical perspective. In: Bhugra D, Bhui K, eds. *Textbook of Cultural Psychiatry*. Cambridge: Cambridge University Press; 2011.

Kirmayer LJ, Guzder J, Rousseau C. *Cultural Consultation: Encountering the Other in Mental Health Care*. New York: Springer-Verlag; 2014.

Kleinman A. Depression, somatization and the "new cross-cultural psychiatry." *Soc Sci Med*.

1977;11:3.

Kleinman A. *Patients and Healers in the Context of Culture: An Exploration of the Borderland Between Anthropology, Medicine, and Psychiatry*. Berkeley, CA: University of California Press; 1980.

Kleinman A, Benson P. Anthropology in the clinic: the problem of cultural competency and how to fix it. *PLoS Med*. 2006;3:e294.

*Lewis-Fernández R, Aggarwal NK. Culture and psychiatric diagnosis. *Adv Psychosom Med*. 2013;33:15.

*Lewis-Fernández R, Aggarwal NK, Bäärnhielm S, et al. Culture and psychiatric evaluation: operationalizing cultural formulation for DSM-5. *Psychiatry*. 2014;77:130.

*Lewis-Fernández R, Aggarwal NK, Hinton L, Hinton DE, Kirmayer LJ. *The DSM-5 Handbook on the Cultural Formulation Interview*. Arlington; American Psychiatric Press, Inc.; 2015.

Lewis-Fernández R, Díaz N. The cultural formulation: a method for assessing cultural factors affecting the clinical encounter. *Psychiatr Q*. 2002;3:271.

Lewis-Fernández R, Aggarwal NK, Lam P, et al. Feasibility, acceptability, and clinical utility of the Cultural Formulation Interview: mixed-methods results from the DSM-5 international field trial. *B J Psych*. (in press)

Littlewood R, Dein S. *Cultural Psychiatry and Medical Anthropology: An Introduction and Reader*. London: Athlone Press; 2000.

Lloyd KR, Jacob KS, Patel V, St Louis L, Bhugra D, Mann AH. The development of the Short Explanatory Model Interview (SEMI) and its use among primary-care attenders with common mental disorders. *Psychol Med*. 1998;28:1231.

*Mezzich JE, Caracci G. *Cultural Formulation: A Reader for Psychiatric Diagnosis*. Lanham: Jason Aronson; 2008.

Mezzich JE, Caracci G, Fabrega H Jr, Kirmayer LJ. Cultural formulation guidelines. *Transcult Psychiatry*. 2009;46:383.

Paralikar VP, Sarmukaddam SB, Patil KV, Nulkar AD, Weiss MG. Clinical value of the cultural formulation interview in Pune, India. *Indian J Psychiatry*. 2015;57(1):59–67.

Weiss MG, Doongaji DR, Siddhartha S, et al. Explanatory Model Interview Catalogue (EMIC): contribution to cross-cultural research methods from a study of leprosy and mental health. *Br J Psychiatry*. 1992;160:819.

Young A. When rational men fall sick: an inquiry into some assumptions made by medical anthropologists. *Cult Med Psychiatry*. 1981;5:317.

*Both authors contributed equally to this work.

30

The Neuropsychiatry of Human Aggression

JEFF VICTOROFF, M.D.

INTRODUCTION

This chapter offers a brief introduction to the neuropsychiatry of individual, nonsanctioned human aggression. It includes an overview of the origins, typology, and biological mechanisms of such aggression. It reviews the association between aggression and psychiatric disorders as they are identified in current nosology. It summarizes what has been discovered, to date, regarding ways to tailor clinical management of aggression to address specific subtypes. Although the problem of collective aggression such as gang behavior and warfare is very deserving of study—and there may be important links between the biopsychosocial generators of individual and collective aggression—violence committed by groups of people is not addressed in this chapter.

Definitions

It is difficult to define aggression. Indeed, one encyclopedic effort identified more than 250 definitions of "aggression." The most common definitions equate aggression with *behavior intended to harm another*. However, that definition is problematic for two important reasons. First, it assumes intent. At this point in the development of neuroscience, valid and reliable techniques to determine motive are lacking. The second failing of the "intent-to-harm" definition is that harm is not necessarily the outcome. Aggressive play (or human sport) may lead to injury but tends to enhance the odds of longer survival. Aggression by alpha animals in dominance hierarchies may resolve disputes and paradoxically increase the fitness of subdominant animals. Parents often act rather aggressively in the course of disciplining youngsters. In most cases, neither the intent nor the outcome is harm.

Another challenge to agreement on a definition is the diversity of aggression. If aggression is regarded as a unitary concept, one needs to have a model and definition that accommodates everything from a playground stare to a nuclear war.

An even more contentious issue is the proper definition of *violence*. More so than aggression, the term violence implies undesirable behavior. It connotes unjustifiable aggression violating social sanctions. One popular

definition was conceived by the President's Commission on the Causes and Prevention of Violence, which stated that violence is "the threat or use of force that results, or is intended to result, in the injury or forcible restraint or intimidation of persons, or the destruction or forcible seizure of property." This definition emphasizes both the aggressor's likely intention and the victim's common response of feeling threatened. (And, in fact, accumulating evidence identifies abnormal threat responsiveness as a cause of untoward aggression.)

Respecting these difficult theoretical issues, and acknowledging that no definition fulfills all qualifiers, for the purposes of this section the following tentative definition is proposed: *Aggression* is overt behavior that involves threat or action that potentially or actually causes pain, withdrawal, or loss of resources. *Violence* is physically or psychologically harmful human aggression that involves the threat or use of force.

Typology of Aggression

There are many ways to classify aggression. One method common to the animal literature is to distinguish types of aggression based on their evolutionary purpose. Table 30–1 lists common types of aggression observed in mammals. It is apparent that these different but overlapping types of behavior evolved to deal with overlapping selective pressures. For example, predatory aggression and territorial aggression both help assure caloric intake; defensive aggression helps an individual to survive predatory attacks; sexual, territorial, maternal, and mate-guarding aggression all enhance reproductive success.

Given the extraordinary conservation of genes and brain systems among mammals, it is clear that human aggression is deeply derivative from these animal types. However, it is difficult to draw one-to-one comparisons. Predatory aggression in nonhuman animals is usually equated with food seeking because carnivores evolved stalking, pouncing, and killing of other species to provide calories. Humans exhibit similar behaviors in interacting with their own species, for instance, in premeditated homicide or warfare. Calories are not the obvious goal. Thus, it is plausible that neurobehavioral systems that originally evolved to support one type of animal aggression (*inter*specific attacks for food seeking) have been incorporated into the behavioral repertoire for human–human (*intra*specific) conflicts.

Many factors explain why human aggression is similar but not identical to aggression in other mammals. Some of these factors relate to cranial and brain evolution. One simple mechanistic difference is that the vomeronasal organ is vestigial in humans. This olfactory apparatus is critical to detecting social signals that directly activate brain programs for many aggressive behaviors in rodents. The virtual absence of this mechanism means that extrapolating from the results of rodent experiments may lead to errors. More important, humans have undergone *encephalization,* so that,

compared with most other animals, humans have a much higher brain-to-body ratio. Humans also exhibit *cortification*—a huge increase in the ratio of isocortex (late-evolving association cortex) to allocortex. This is one of several reasons that caution is advisable in equating animal with human types of aggression. Cortification means that, in spite of the conservation of many genes and behaviors, humans may deal with the same ecological challenges by using different cerebral mechanisms than other animals, weighting sensory information in more complex ways, sending that sensory information down different processing streams, and adjusting responses to challenges with much more complicated and plastic risk–benefit calculations.

Table 30–1.
Types of Mammalian Aggression

Predatory/food seeking
Defensive
Territorial
Dominance related
Sexual rivalry related
Intercourse seeking
Mate guarding
Maternal
Infanticide

Humans also exhibit peculiarities that, taken individually, might be seen in other species but that, taken together, are rare. Humans are (1) highly social, (2) territorial, (3) hierarchical, and (4) somewhat monogamous. Human forebrains allow very long-term planning, and association cortices support *secondary emotions,* that is, emotions that one might generate not in direct time-locked response to a threat but just by recollection and rumination (e.g., bringing from memory into consciousness a past social defeat, an insult, an emotional stress, or a personal loss) even decades after these events have occurred, thus causing arousal and sometimes lowering the threshold for aggression due to internal stimuli. The same cerebral peculiarity—very large and plastic association cortices and the resulting capacity for experience-dependent learning—helps explain the great interindividual variability that distinguishes humans from most other animals. Language is another human peculiarity that permits new types of aggression. Verbal aggression can be more complex and nuanced in communicating short- or long-term threats (and in providing nonphysical outlets for anger and conflict resolution short of fighting) than growling, hissing, or other related animal vocalizations. Cultural variability is yet another human peculiarity; chimpanzees are the only other species known definitely to exhibit different

behavioral cultures in geographically different groups. Among humans, this cultural variability leads to considerable differences in what is considered socially appropriate, sanctioned aggression versus what is taboo and punishable.

Another important caveat about comparing aggression in other species with human aggression is that most mammals display very stereotyped sequences of aggressive behavior with well-delineated stages of approach, threat, attack, biting, and resolution. Such reproducible sequences are called ethograms. The massively increased flexibility afforded by the human heteromodal association cortices produces aggressive behaviors of very much greater variety, often lacking such clear ethogrammatic stages. Moreover, the mouse is not an ideal species for comparison because male mice rarely attack female mice, whereas men frequently commit violence against women. Further research is necessary to produce a plausible accounting of shared versus nonshared biological underpinnings of similar behaviors in humans and in the standard animal models of aggression, including genes, proteins, structures, circuits, neurological reaction patterns, and gene–gene and gene–environment interactions.

Sanctioned versus Nonsanctioned Aggression. Aggression might also be categorized according to social acceptability. Socially sanctioned aggression, depending on culture, might include rough and tumble play, spanking, contact sports, hunting, police or intelligence service actions, capital punishment, or war. Socially prohibited aggression in most cultures includes criminal assault, rape, homicide, fratricide, parenticide, infanticide, child abuse, domestic violence, torture, civil disturbance, and terrorism. These distinctions are not absolute. Spanking, infanticide of girls, and capital punishment are accepted in some cultures and considered abusive or abhorrent in others. Stoning women or men to death for perceived violation of honor codes is regarded as desirable in some places and as homicide in others. The principal point is that societies reward some forms of aggression and punish others.

A closely related way to classify aggression in organized societies is to distinguish official from nonofficial violence. This distinction is a political-scientific one. Having formed groups adhering to social contracts that include acquiescence to authority, people often tolerate aggression in the name of collective welfare so long as recognized authorities commit the actual violence. As a general statement, even though one grows up with a clear, culture-centric picture of the distinction between permitted and prohibited aggression, (1) every society classifies aggression into its own socially acceptable and unacceptable categories, and (2) with the possible exception of parenticide (witness *Oedipus Rex*), it is difficult to identify types of aggression that are universally classified as taboo.

Hypoarousal- versus Hyperarousal-Related Aggres sion. Many individuals who exhibit psychopathic traits—including nonsanctioned aggression, lack of respect for the rights of others, callousness, cruelty, lack of remorse, and lack of empathy—have been found to have lower than usual resting heart rates and less autonomic reactivity. This has led to a classification distinguishing such hypoarousal-related aggression from the aggression seen in anxiety disorders, mania, or stimulant intoxication, which is called hyperarousal related.

Proactive versus Reactive Aggression. The modern literature on human aggression usually applies a different typology. The emphasis in the vast majority of the psychological literature is on aggression that is either *proactive,* in the sense that the actor initiates the aggression against a target without immediate provocation, or *reactive,* in the sense that the actor is responding to a threat. Different authorities address this supposed dichotomy using different terminology. Approximate synonyms for proactive aggression are instrumental (serving a practical goal), planned, premeditated, cold-blooded, or predatory aggression. Note, however, that each of these terms carries somewhat different connotations and conceivably involves different neural underpinnings. Moreover, the standard experimental model for so-called instrumental aggression is the resident–intruder model, in which a rodent is suddenly confronted with a strange rodent in his cage and attacks that intruder. Such behavior seems a poor model of proactive aggression and might better be interpreted as a defensive reaction to a territorial threat. This rhetorical idiosyncrasy may explain some of the inconsistencies in the neurobiology of aggression literature. Approximate synonyms used for reactive aggression include hostile, affective, defensive, hot-blooded, or impulsive aggression. Again, it may be presumptuous to attribute affect to an animal defending itself from attack, which complicates efforts to equate the physiology of reactive aggression in rodents with angry, hostile aggression in humans.

Although the proactive/reactive dichotomy has gained an unshakeable foothold in the literature, a simple thought experiment demonstrates its limitations. A soldier on wary patrol hunting for the enemy in an urban wasteland may suddenly be confronted by an opponent wielding a weapon. If the first soldier fires, is that proactive/premeditated/predatory aggression or reactive/defensive/affective aggression? It is obviously a mix. Anyone who has witnessed an argument between two men escalates into a fight—perhaps the most common cause of fatal non–war-related violence—will acknowledge that a proactive/reactive dichotomy fails to describe the mutuality of this process. The popular proactive/reactive dichotomy has heuristic virtues but is a flawed description of many real events. From domestic violence to gang wars, humans simultaneously engage well-integrated neural systems that originally evolved to serve discrete

functions.

Epidemiology of Nonsanctioned Human Aggression

Because of underreporting, there is no way to provide a systematic account of nonsanctioned human aggression. It is hard enough to capture accurate statistics measuring major violent crimes such as rape. How would one perform an epidemiological catchment area study that reliably assessed the endemic occurrence of less severe but unequivocally aggressive behaviors from spanking, sibling fights, and low-intensity domestic spats to corporal punishment, schoolyard bullying, and rough hits in contact sports? Because aggression is woven into the fabric of normal development and commonplace in human intercourse, one begins with the limitation that only a small subset of aggressive acts come to professional or juridical attention.

That having been said, criminologists have undertaken several major initiatives to determine the frequency of some types of nonsanctioned aggression, both nationally and internationally. In the United States, the Federal Bureau of Investigation's (FBI's) Uniform Crime Report (UCR) compiles police statistics from more than 18,000 agencies and gives both absolute numbers of crimes and crime rates expressed as frequencies per 100,000 population. The UCR divides crimes into "serious," Part I offenses —including willful killing, rape, robbery, aggravated assault, burglary, larceny, motor vehicle theft, and arson—and less serious, Part II offenses—including simple assault, family abuse, nonrape sexual offenses, and a large variety of nonviolent crimes. Note that crimes of interpersonal aggression appear in both Parts I and II. The methodology of the UCR has recently been revised to include more data, such as the age, sex, and race of the victim and the perpetrator, the type of weapon used, the relationship of the victim to the perpetrator, and the circumstances of murder. These more detailed data have allowed better examination of issues such as rates of violence by geographic region, by community size, and by demographic traits of perpetrators, as well as more dynamic questions such as whether gun control laws influence the net rate of homicides. The FBI has also recently revised its definition of rape: "Beginning with the 2013 data collection, the SRS [Summary Reporting System] definition for the violent crime of rape was modified to: 'Penetration, no matter how slight, of the vagina or anus with any body part or object, or oral penetration by a sex organ of another person, without the consent of the victim.' "

Another way to measure criminal aggression is the victimization survey. The US National Crime Victimization Survey (NCVS) periodically measures reported rape, robbery, assault, and other offenses as reported by about 160,000 individuals from about 90,000 households. These surveys provide much more detail about the victims, including age, race, sex, income, education, and relationship between victim and perpetrator. Because

victims report events on this survey that may not have been reported to police, the NCVS potentially provides a more complete picture of the effect of aggressive crime. However, the value of the NCVS depends on full and accurate reporting by victims. Some individuals probably overreport to impress the interviewer. Others underreport, often declining to reveal instances of domestic violence or sexual aggression, or fearing to report serious assaults by perpetrators who are still in the neighborhood.

A third way to estimate the frequency and distribution of nonsanctioned aggression is to collect self-report information from criminals or delinquents. Such data have uncovered the gap, in some cases quite shocking, between crime that is known to authorities and crime that is admittedly committed. For instance, in a study of delinquency, boys confessed to an average of 680 offenses, and in a study of robbery, they confessed to an average of 214 robberies—vastly more than the number for which they had been arrested. Again, such data are only as trustworthy as the subjects. Some may overreport out of bravado; others will underreport out of fear of prosecution.

Acknowledging the limitations of all three types of aggression measures, the available data help conceptualize the magnitude and trends of the problem of nonsanctioned interpersonal aggression. According to the UCR, in the year 2013, US law enforcement agencies collectively reported 13,716 murders, 62,034 rapes, 335,428 robberies, and 683,971 aggravated assaults. For perspective: the country's violent crime rate was 713.6/100K in 1994, 463.2 in 2004, and 367.9 in 2013. There is strong evidence of a three-decade-long increase, followed by a two-decade-long decrease across all categories of offenses and across urban, suburban, and rural reporting districts. This amounted to a gratifying 20.6 percent decline in the nation's violent crime rate between 2004 and 2013. This trend is displayed in Table 30–2.

Of the US murderers, 64.3 percent were male, 7.7 percent were female, and 28 percent were of unknown gender. Reportedly, 31.1 percent were white, 38.0 percent were black, and 16.0 percent were Hispanic or Latino; 18.8 percent of murderers were younger than 22 years. In fact, 26.3 percent of all murders were attributed to persons between the ages of 17 and 24 years. Firearms were frequently used in the United States to facilitate murders, rapes, robberies, and assaults. For instance, in 2013, among the 12,253 murders in which the weapon was identified, 8,454 involved firearms (69 percent). In 68 percent of those cases, the murder weapon was a handgun.

The NCVS, being a self-report project, cannot very well enumerate murders. It does, however, reveal a startlingly higher number of violent crimes than those reported by police agencies: the 2013 NCVS documented self-reports of 6,126,420 violent crimes that year in the United States. These data are perhaps more useful than the UCR data for appreciating the

actual risk. For instance, while the UCR recorded about 62,000 rapes that year, respondents self-reported more than 346,000 rapes/sexual assaults. What both methods suggest, however, is a significant decline in violence in the United States over the last two decades (see Fig. 30–1).

What caused this decline? In a 2015 report from the Brennan Center for Justice, the investigators are candid that no single cause has been identified. Although a massive increase in the rate of incarceration perhaps contributed to the crime decline in the 1990s, the continued increase in the incarceration rate has yielded negligible returns. More likely beneficial factors include aging of the population, an increased number of police, income growth, consumer confidence, and a decline in alcohol consumption.

Table 30–2.
US Rates of Type I Offenses (per 100,000 Population)

Year	Murder	Rape	Aggravated Assault
1960	5.1	9.6	86.1
1970	7.9	18.7	164.8
1980	10.2	36.8	298.5
1990	9.4	41.2	424.1
2000	5.5	32.0	324.0
2007	5.6	30.0	287.2
2013	4.5	25.2	229.1

From FBI Uniform Crime Reports.

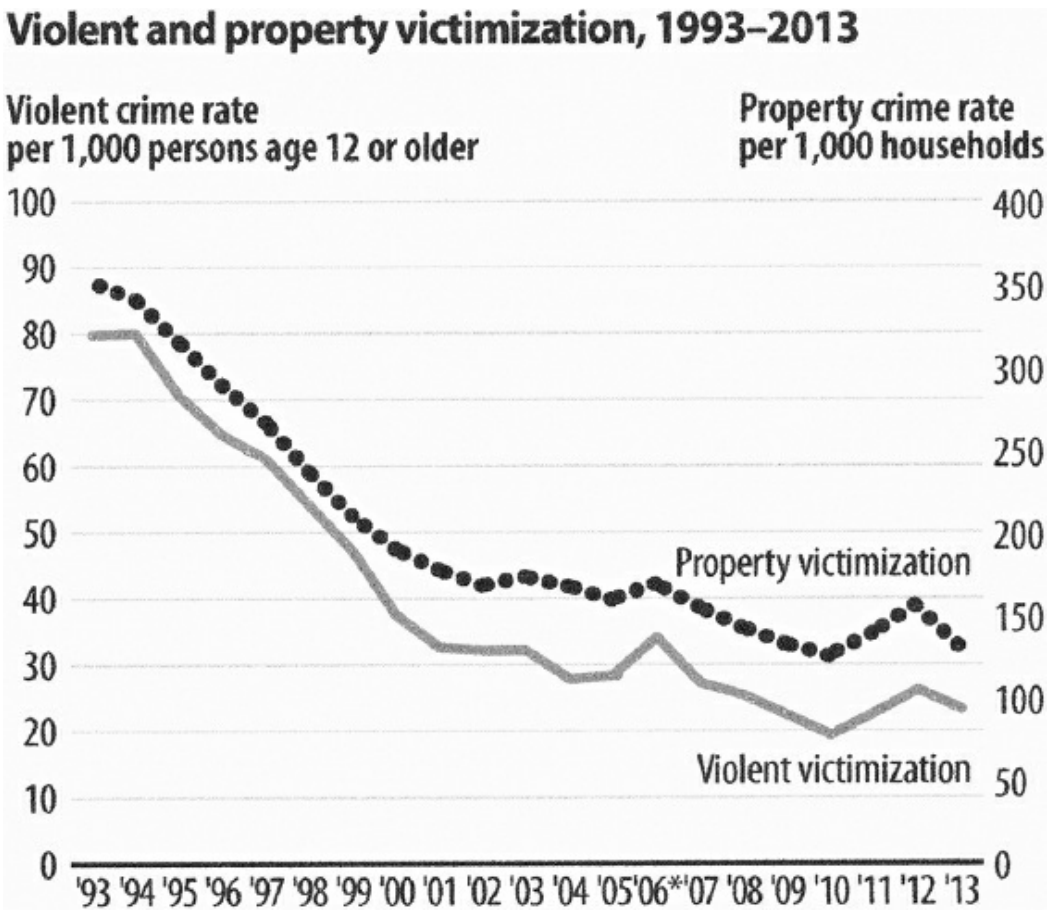

FIGURE 30–1. Violent and property victimization, 1993–2013. (From Truman JL, Langton L. Criminal victimization. U.S. Department of Justice, Office of Justice Program, Bureau of Justice Statistics. Available at http://www.bjs.gov/index.cfm?ty=pbdetail&iid=5111. Accessed June 30, 2015.)

Scholars have also looked at trends in nonsanctioned human aggression with a wider lens. Eurostat is one source of international crime statistics recorded by police, its most recent publications including data from 2002 to 2012. Violent crimes in the EU reportedly declined by about 10 percent between 2007 and 2012—a trend largely attributable to declines in England and Wales. However, rape in the EU is not surveyed because of disputes regarding the definition.

Comparative international data are also available in the surveys conducted by the United Nations Office of Drugs and Crime. Their Global Study on Homicide 2013 reveals major regional difference in rates. By far, the highest global rates are recorded in Southern Africa. In the country of South Africa, for example, the 2012 rate was ~32/100,000. The Americas also exhibit high homicide rates, yet the subregional differences are significant: In 2012, the Northern Americas reported a rate of less than 5/100,000 while South America had a rate of almost 25/100,000 and Central America had one of the highest global rates at 26.5/100,000. Even these data obscure sub-subregional variation of great interest. As illustrated in Figure 30–2, zones of elevated homicide rates are sprinkled

about the globe. Even this obscures more local effects. In the United States, for example, states in which the governments support gun possession and carriage with little restriction, such as the southern states of Louisiana and Mississippi, typically exhibit the highest rates of gun homicides.

The UN report also reports the demographics of victims, hinting at one's relative risk. Globally, 35 percent of homicide victims are males aged 15 to 29 years, 44 percent are males younger or older, 13 percent are females aged 15 to 29 years, and 8 percent are females younger or older (see Fig. 30–3).

The UN Report further specifies the "mechanism" of these homicides. As illustrated in Figure 30–4, guns are by far the main means or enablers of unsanctioned homicides in the Americas, playing a far smaller role in Europe, Africa, and Asia.

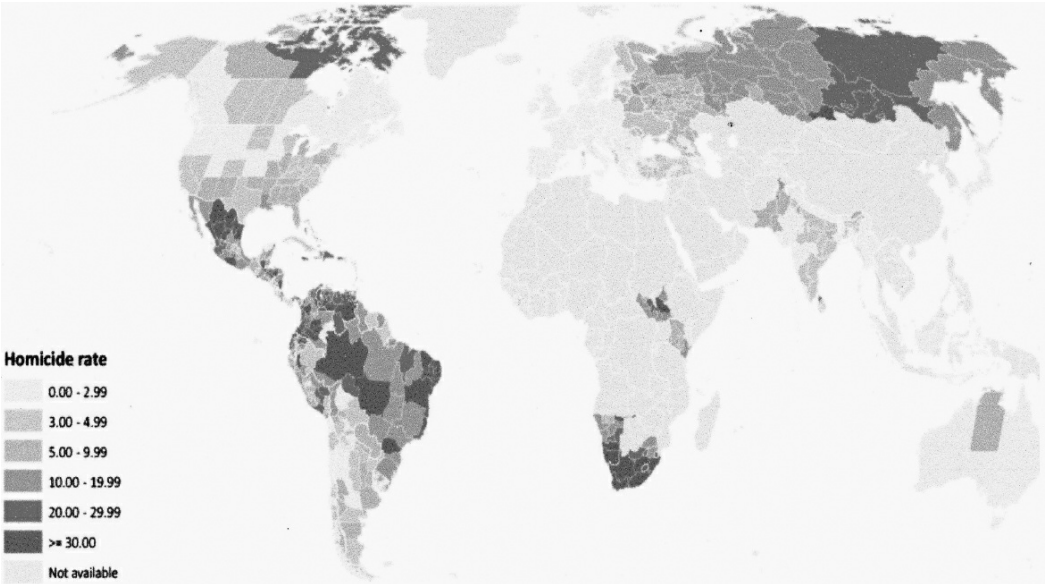

FIGURE 30–2. Homicide rates at the subnational level (2012 or latest year). (From United Nations Office on Drugs and Crime: Global Study on Homicide 2013. United Nations publication, Sales No. 14.IV.1. p. 26.)

FIGURE 30–3. Global homicide rate per 100,000, by sex and age group (2012 or latest

year). (From United Nations Office on Drugs and Crime: Global Study on Homicide 2013. United Nations publication, Sales No. 14.IV.1. p. 30.)

Demographic Correlates of Violence. The most obvious correlate of homicide is male sex. About 90 percent of those arrested for murder are men, and more than 99 percent of serial killers are men. In the United States in 2013, 88.1 percent of murdered males were killed by other males, and 90.2 percent of murdered females were killed by males. In the United States in 2013, a husband was 4.9 times as likely to murder his wife than to be murdered by her, and a girlfriend was 3.4 times as likely to be murdered by her boyfriend as he was to be murdered by her. Antisocial personality disorder (APD) is three times as common in men, and violent suicide is also about three times as common among men as among women. Warriors in almost all societies have been men. It is tempting to interpret these data as evidence of a casual theory, such as an innate male "aggressive instinct," the distribution of labor into killer male hunters and pacific female gatherers, the inevitability of male sexual rivalry, the physiological consequence of plasma testosterone, or some combination of these factors. This section cannot review the massive evidence supporting one or another of these theories. Suffice it to say that the Y chromosome (or more specifically the sex-determining region Y gene), combined with typical social imperatives affecting boys' development, constitutes a highly predictive risk factor for committing acts of interpersonal aggression. Relatively greater upper body strength, an apparently innate disposition for rough and tumble play, and increased perceptual sensitivity to dominance rivalry cues all contribute to the very strong male predominance among those who commit nonsanctioned aggression.

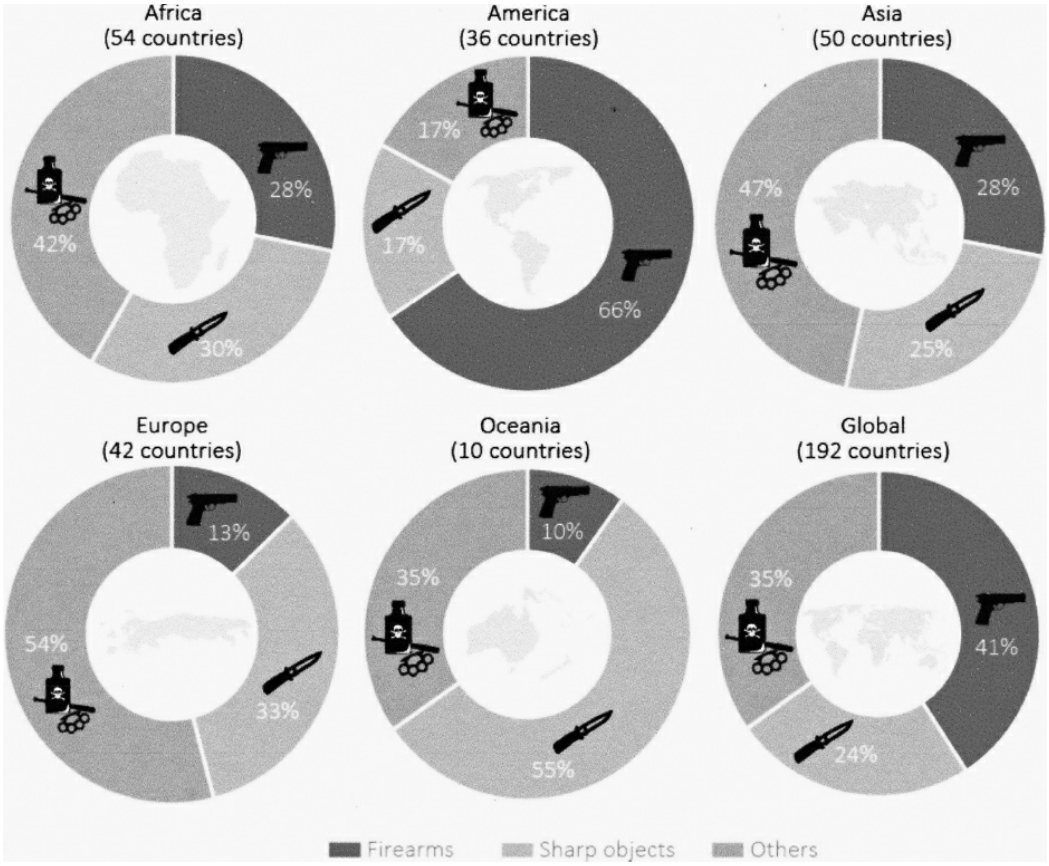

FIGURE 30–4. Global homicide rate per 100,000, by sex and age group (2012 or latest year). (From United Nations Office on Drugs and Crime: Global Study on Homicide 2013. United Nations publication, Sales No. 14.IV.1. p. 16.)

RACE. Race or ethnicity is another possible factor. The sociology of interpersonal aggression is politically sensitive because of the gross overrepresentation of minorities among arrestees, among those convicted, and among those sentenced to death. For example, although African Americans comprise 12 percent of the US population, they account for more than 50 percent of those arrested for violent crimes, including 55 percent of those arrested for murder, as well as approximately 50 percent of persons incarcerated in state and federal penitentiaries. Homicide is the leading cause of death for African Americans between the age of 15 and 34 years. An African American man has about a 1 in 30 chance of death from homicide, compared with the 1 in 180 chance of a white male. These racial differences have been studied for decades. Several large-scale studies have demonstrated that most racial differences disappear when the data are controlled for education and socioeconomic status, strongly suggesting that the risk of involvement in violence as a perpetrator or a victim has social, not racial, roots. Moreover, the overwhelming majority of murders are within-race crimes: In 2013, 90.1 percent of African American victims were murdered by African American perpetrators, and 83.5 percent of white victims were murdered by white perpetrators—data that might help dispel the myth that society's more advantaged people are suffering from the

violent behavior at the hands of its less advantaged.

POVERTY. Lower socioeconomic class is widely considered a risk factor for violent behavior. This relation held up in one review of more than 100 studies. However, there is room for reinterpretation. For example, one study found that a person's status within his class was more predictive than class membership itself. This might imply that individuals who are lowest on a proximate social ladder are subject to experiencing deprivation relative to their groups' expectations and perhaps are more at risk for low self-esteem and are at higher risk to become criminally violent.

Fear of death at the hands of a stranger predominates the consciousness of subjective risk assessment. However, in the United States in 2013, just 1,281/12,253 (10.5 percent) of murders were determined to have been perpetrated by strangers. Acquaintanceship is the most common relationship between murderer and victim, occurring in 21.7 percent of cases. The second most common relationship between murderer and victim was familial (13.6 percent).

Substance abuse is another major factor in nonsanctioned interpersonal violence. This issue is discussed in more detail in the subsection on substance abuse and aggression.

ON THE RELATIONSHIP BETWEEN GUN CONTROL, AGGRESSION, INJURY, AND DEATH. In certain countries such as Yemen and the United States, a plurality of leaders support gun possession with little or no restriction. Behavioral epidemiology strongly suggests that, as a direct result, many people die every year in those countries due to aggression by a gun possessor. In fact, the incidence of gun-related injuries and deaths is more than one order of magnitude higher in societies with weak or no gun control laws compared to similar societies with stronger gun control laws. Survey evidence suggests that the desire to possess the capacity to kill other people is widespread among humans. Historical observation reveals that confrontations provoking impulsive rage, or emotional distress promoting suicidal depression, are universal in human societies. Social psychological research has recently concluded that the problem of nonsanctioned gun-related injury and death is best accounted for by a conflict in priorities such that, among some human groups, political advantage and public health interests are in direct opposition.

NEUROANATOMY AND AGGRESSION

Many parts of the central nervous system play roles in mediating human aggression. It is easy, in a sense, to list the parts of the brain that, when stimulated or damaged, change the expression of aggression. What is very hard—and a project for the next century of research—is to sort out exactly how these parts function and interact during the years, months, hours, and

seconds before a person commits an act of either adaptive or maladaptive aggression, how these parts function during the actual moments of aggressive motor action, and what happens in the brain as the burst of motor action resolves. Progress in neuropathology, neurophysiology, neuroimaging, and molecular genetics has significantly advanced the field. Nonetheless, although convergent research has begun to clarify the neural circuitry underlying human aggression, one needs a degree of humility in attempting to catalogue the major findings.

Two neuroanatomical networks, one primarily subcortical, the other primarily cortical, with both intersecting in the amygdala—appear to account for much of the brain processing culminating in human aggression. A brief tour of the major intersections will assist in navigating this network.

Role of the Brainstem

The brainstem plays at least three distinct roles in the mediation of aggression. First, perception of a threat such as a predator provokes defensive behavior mediated, in part, by the superior and inferior colliculi and the dorsal periaqueductal gray (PAG). Reciprocal feedback among these parts of the brainstem, the medial hypothalamus, and the centromedial amygdala adjusts these responses, suggesting that arousal to threat and defensive aggression itself can be triggered by the dorsal midbrain. Second, the ventral tegmental area (VTA) projects dopamine (DA) to the limbic cortex, hypothalamus, amygdala, nucleus accumbens (NAc), and septal area (S). These dopaminergic projections support visceral "fight or flight" responses and help in the learning of rewarding or aversive outcomes. Overall, as will be discussed further, mesocorticolimbic DA innervation may be *aggression permissive,* balancing the *aggression-damping* influence of serotonin. Third, raphe projections, the major cerebral source of serotonin (5-HT), damp impulsive aggression in many species, including humans.

Role of the Hypothalamus

The conclusion from many mammalian studies is that the ventral portions of the ventromedial and lateral nuclei, when stimulated, produce offensive aggression. This region is, therefore, referred to as the *hypothalamic attack area* (HAA). The HAA receives projections from the dorsal raphe and medial amygdala and sends projections to the central gray, the mediodorsal (MD) nucleus of the thalamus, the septum, and the prefrontal cortex. It is not clear to what extent this phylogenetically ancient HAA mediates human aggression. Because serotonergic drugs affecting 5-HT_1 receptors alter hypothalamic aggression, the hypothalamus is considered to be a likely site contributing to the 5-HT aggression-damping effect. Prior experience of aggression modulates the otherwise stereotypical aggression provoked by

HAA stimulation, supporting a link between hypothalamic activity and memory.

Gonadotropin-releasing hormone (GnRH), produced in the hypothalamus, modulates vertebrate reproductive behavior via steroid sex hormones and in mammals feeds back to limbic receptors. Through variable expression of hypothalamic sex steroid hormone receptors and aromatase, the hypothalamus modulates the gender-specific aggressivity linked to reproduction (e.g., intermale fighting, sexual aggression, maternal aggression, and infanticide). The neuroendocrinology of aggression is discussed in more detail later.

Role of the Amygdala

The amygdala is a vital nexus in the neural network supporting aggression and violence. The medial temporal lobe supports both an input function—the recognition of threats and social signals—and an output function—setting the gain for the aggressive response bias. In other words, this brain region—including the amygdala—controls signals delivered both downstream to the hypothalamus to activate visceral/autonomic responses and upstream to the motor cortex, ultimately activating the muscles involved in aggressive behavior. A somewhat simplified model divides the amygdala into two major parts: A corticobasolateral amygdala (BLA) and an extended centromedial amygdala (CMA). Sensory input to the BLA mediates the conditioned fear response to threatening stimuli. In contrast, the CMA is widely connected with brainstem and hypothalamic nuclei, as well as to olfactory structures mediating social/reproductive functions in rodents.

In studies of humans and other primates, amygdaloid cells respond to perceived threat as well as to facial features, especially to emotional facial expression and to direction of gaze. Evidence suggests that definite threats activate the amygdala less than threat-related ambiguity, suggesting that this region is an assessor of the need for possible aggressive responses rather than just a releaser of stereotyped aggressive behavior. These findings together support the conclusion that the amygdala provides a center for rapid associations of sensory information relevant to food seeking, social circumstances, reproduction, and threat, and links the recognition of socially loaded stimuli with threat or reward to help predict the consequences of fight versus flight. The medial portion of the amygdala is also connected with the medial division of the orbitofrontal cortex, which is in turn connected to the anterior cingulate and dorsolateral frontal lobe. This might be considered the information conduit by which the amygdala's largely unconscious processing of threats is integrated with higher-order assessments of social circumstances.

Role of the Nucleus Accumbens, Ventral Striatopallidal System, and Septal Areas

Heath reported in 1954 that electrical stimulation of the septal area altered the comportment of a schizophrenic patient from raging to docile. Since then, many reports have confirmed a role for an anterior brain region including the septal area (S), NAc, and ventral striatum (VS) in the regulation of mammalian aggression. Lesions of the NAc are associated with striking increases in both defensive and predatory aggression in rodents, cats, and dogs; lesioned animals will attack other animals or experimenters with little provocation and even viciously attack inanimate objects. However, lesioned animals typically exhibit *decreased* social/competitive aggression, suggesting that the role of the NAc in regulating aggression depends on the social context. VS and S lesions also induce marked increases in reactive and predatory aggression, a finding that fits with Heath's early clinical observation. Increased DA or decreased serotonin in these structures is linked to increased aggressivity; thus, monoaminergic function in this region may mediate the frequently reported link between low serotonin and impulsive aggression in humans, as well as the antiaggressive benefits of some antipsychotic medications.

Role of the Thalamus

Environmental stimuli, such as a verbal challenge or an approaching fist, traverse modality-specific thalamic relay nuclei. In reptiles, motor responses are largely guided by this thalamic sensory information, with direct projections from the thalamus to the amygdala that allow for the rapid evaluation of possible threats and leading to semiautomatic fight-or-flight reactions. In mammals, a remnant of this direct thalamoamygdalar "shortcut" possibly exists by which little-processed sensory information recognized as threatening (e.g., the glimpse of a snake, the roar of a lion) is passed from thalamus to the BLA and incorporated in a rapid, if unrefined, response. In mammals with large cortices, however, the main route from thalamic relay nuclei to the BLA is the slower, indirect pathway by which information is first passed to primary and then to higher-order sensory association cortices, from which processed information is funneled to the BLA, where potential threats or predatory opportunities are detected (if not consciously perceived). The BLA, in turn, projects back to one part of the MD thalamic nucleus, whereas other parts of the nucleus receive inputs from the basal ganglia, NAc, and hypothalamus. This coalescence of projections potentially permits the MD thalamus to merge information regarding the emotional and social importance of an external stimulus with information about the body's state of preparedness. The output from the MD thalamus is widely projected to areas of the frontal lobe involved in planning motor action. Hence, the MD thalamus appears to be well positioned to adjust the gain for cortically generated aggressive motor plans according to subcortical influences.

Role of the Anterior Cingulate Gyrus

The behavioral significance of the anterior cingulate gyri (ACG) has received enormous attention in recent years because this region is activated during many functional magnetic resonance imaging (fMRI) tasks. The emerging consensus is that the ACG serves (among other functions) as a sort of conflict detector: It receives input both from the limbic system and the prefrontal cortex and is sensitive to conflicting motor plans, at which point it recruits greater cognitive control by activating the dorsolateral prefrontal cortex (DLPFC) to carefully consider the relative risks and benefits of those behavioral options. Failure of this function, therefore, might theoretically leave a person without the benefit of conscious deliberation before acting on impulsive, emotionally driven, socially inappropriate planning. Although the precise contribution of the ACG to the circuitry of aggression remains to be determined, the difference in connectivity of its divisions provides clues. For example, the motivational drive for predatory/premeditated aggression is associated with food seeking in most mammals. The connectivity of the infracallosal (inferior to the level of the corpus callosum) ACG may link the hypothalamic hunger drive with motivation and physiological preparation for stalking, killing, and eating prey. One of the great mysteries of aggression research is how human predatory aggression (e.g., serial rape or murder) is tied to the hunger-satisfying predation of other carnivores. It is conceivable that human predation uses the same circuitry but that a different kind of reward —such as the satisfaction of blood lust, sex, or dominance—substitutes for food. In contrast, the supracallosal ACG possibly plays a role in defensive and social aggression via its connections to the threat-and-face-sensitive BLA and, with the restraint provided by the lateral orbitofrontal cortex, to aid in restraining the impulse to aggress in social confrontations. These postulated separate roles of the divisions of the cingulate must be tested in future research.

Role of the Frontal Lobes

The human frontal lobes integrate sensory-limbic input with memory of the immediate past (working memory) while simultaneously monitoring social circumstances to plan, sequence, and anticipate the consequences of actions and rapidly adjust motor programs in response to ongoing environmental feedback. Because aggression calls on all of these capacities, many frontal regions contribute to conscious aggressive behavior. Yet frontal lobe injury does not uniformly increase aggression. In fact, early reports suggested that docility—accompanied by bursts of anger elicited by frustration—is the primary result. These observations must be reconciled with other findings of increased aggression (primarily but not exclusively impulsive) associated with some instances of frontocortical damage. The

explanation follows from an analysis of different frontal functions.

Role of the Ventromedial Prefrontal Cortex. The ventromedial prefrontal cortex (VMPFC) complex includes (1) the orbitofrontal cortex, (2) the VMPFC, and (3) the infra- and supracallosal anterior cingulate gyrus. Together these frontal regions help organize social behaviors and integrate them with both sensory input and visceral processing. Whereas the detailed anatomy of this region is beyond the scope of this section, it is worth noting that the orbitofrontal part of the VMPFC can be divided into medial and lateral divisions. The medial division integrates visceral drives (perhaps including anger) with the internal milieu (such as arousal and muscular readiness). The lateral division receives highly processed sensory information—for instance, the sight and sound of an approaching enemy. Both divisions are also reciprocally connected with the BLA. The VMPFC plays a role in restraining inappropriate social behaviors and reactive/impulsive aggression.

Role of the Dorsolateral Prefrontal Cortex. Intentional behavior involves the conscious will to execute a plan generated by the DLPFC. About 60 percent of the cells of the DLPFC are involved in encoding such a motor plan and about 40 percent are involved in working memory—the best candidate for a cognitive construct corresponding to human consciousness. So the DLPFC, acting under the influence of the VMPFC in concert with the conflict-alerting system of the ACC, combines new sensory data held in working memory with long-term memories to generate a plan and anticipate its consequences. Having calculated the most adaptive signal stream, the DLPFC transmits it to the supplemental motor area (SMA) and the premotor cortex (Brodmann area 6) for temporal refinement and transmission to the motor strip. All intentional aggression must use this final common pathway.

This helps explain how frontal damage might either increase or decrease aggression. Damage to the most inferior and medial region of the VMPFC may rob a person of the capacity to evaluate the social significance of events and to restraint impulsive responses. This type of damage would tend to increase aggression by disinhibition. Damage to this most conscious part of the brain, the DLPFC, might either increase or decrease aggression, depending on the precise nature of the injury (and probably the *side* of the injury, as will be discussed later). A person whose damage to the DLPFC alters the ability to integrate his or her emotions into an organized motor plan might feel upset but fail to act in feeling proportionate way. Another person with a DLPFC lesion near the corresponding area affected in the first person might be perfectly able to develop an anger-influenced motor plan but may have lost the ability to calculate the likelihood of painful consequences; the result could be imprudent aggression.

Hemispheric Specialization for Aggression?

Striking evidence exists that the processing of stimuli to which an aggressive response might be adaptive tends to occur asymmetrically in the brain of many nonhuman vertebrates. Several reports in humans are consistent with these observations. For example, one study reported that aggressive behavior was associated with left frontotemporal hypoperfusion among patients with dementia. Another study reported that adults with impulsive aggression exhibited a reduced left medial frontal response to a serotonergic stimulus. Several studies reported that left temporal lobe dysfunction is associated with greater hostility or aggression among persons with epilepsy. Evidence also exists that the right hemisphere is specialized for processing emotionally unpleasant or threatening stimuli. For example, one fMRI study reported that the right amygdala preferentially responded to fearful faces, and a positron emission tomography (PET) experiment with veterans showed that mental imagery of combat disproportionately activated the right amygdala.

One interpretation of these data is that the right brain is specialized to attend to emotionally threatening stimuli and to initiate an aggressive response, whereas the left brain normally restrains such aggression. Dysfunction on the left would thus tend to disinhibit the right brain's initiation of aggressive behavior. However, not all of the available evidence is consistent with this interpretation. Despite tantalizing results suggesting the possible conservation of an evolutionarily ancient right brain dominance for aggression, the current understanding of the lateralization of human aggression is far from conclusive.

Two Circuit Loops Supporting Human Aggression

The anatomy of aggression can be summarized as two loops of neuronal circuitry intersecting at the amygdala. As shown in Figure 30–5, a largely unconscious or subliminal loop supports communication among the amygdala, the hypothalamus, the NAc/septal region, and the brainstem. The amygdala receives sensory input via the thalamus and the cortex and assesses whether this is an aggression-provocative stimulus. Connections with the brainstem, hypothalamus, and septal area integrate arousal, serotonergic status, autonomic status, and reward feedback. The amygdalar signal for an aggressive response tendency is passed via the looping stria terminalis to the NAc and the VM hypothalamus. As shown in Figure 30–6, the amygdala is simultaneously communicating with the cortex. Highly processed sensory information about potentially aggression-provocative stimuli travels from association regions such as the inferior parietal lobule to reach both the amygdala and the VMPFC. In a partly conscious to and fro, these regions assess the threat, opportunity, social significance, and potential reward value. This assessment is crucial to developing a good

motor plan adjusted via connections with the anterior cingulate gyrus and the DLPFC. When the system is working well, a person makes rapid, appropriate, survival-critical evaluations of the circumstances, his or her brain juggling largely unconscious information with conscious awareness, and acts accordingly, for instance, by fighting or fleeing.

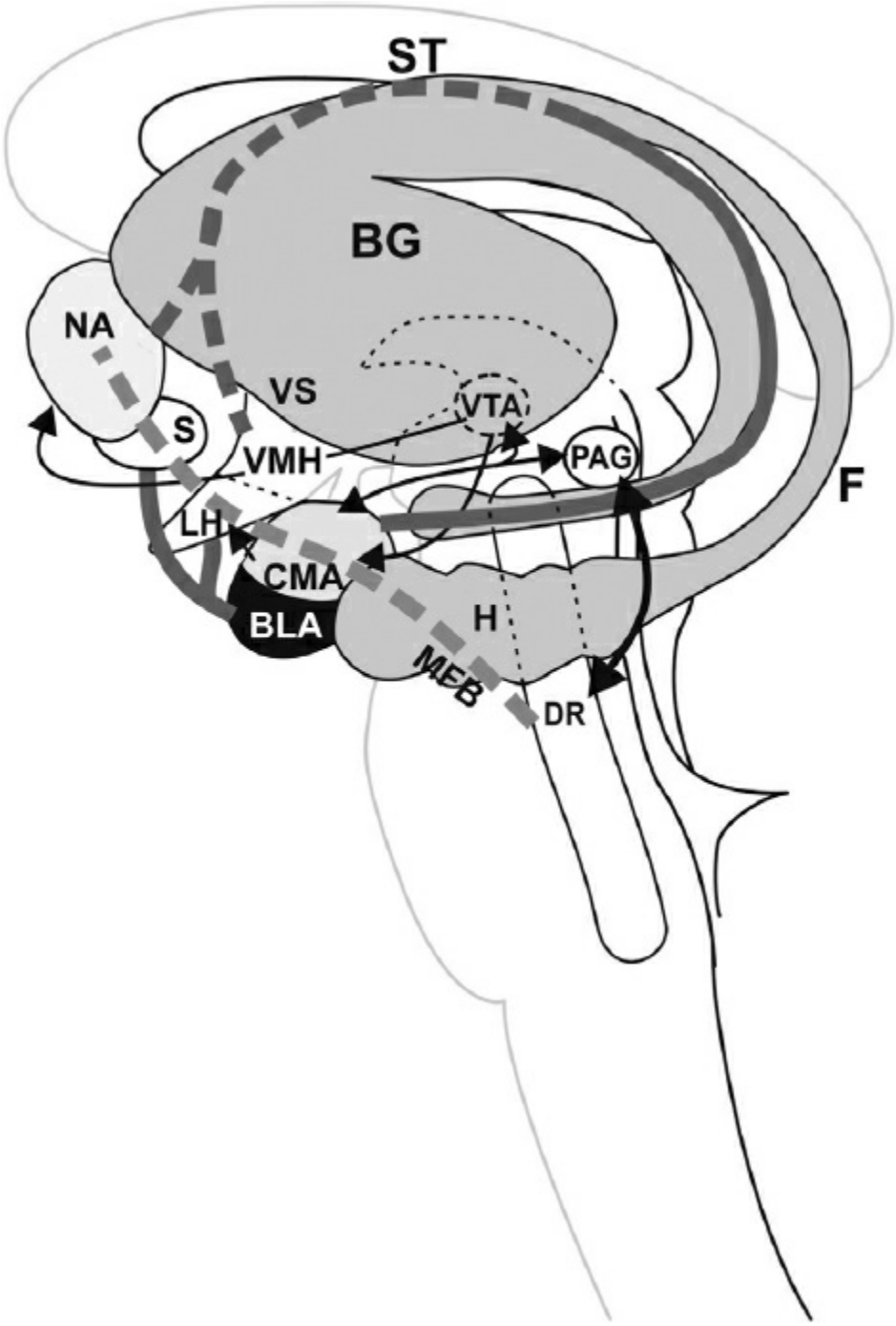

FIGURE 30–5. Subliminal processing of aggression. From semischematic of the subcortical circuit loop mediating unconscious processing of stimuli to which an aggressive response might be appropriate. BG, basal ganglia; BLA, basolateral amygdale; CMA, centromedian amygdale; DR, dorsal raphe; F, fornix; H, hippocampus; LH, lateral hypothalamus; MFB, median forebrain bundle; NA, nucleus accumbens; PAG, periaqueductal gray; S, septal region; ST, stria terminalis; VMH, ventromedial

hypothalamus; VTA, ventral tegmental area; VS, ventral striatum.

NEUROCHEMISTRY OF AGGRESSION

Virtually every amine neurotransmitter and many peptide and steroid hormones play a role in the cerebral mediation of aggression. A great deal of research, primarily with rodents, has established correlations between various measures of these chemicals and various measures of aggression. Yet there are major gaps in the current knowledge. Are ideal subtypes of aggression such as affective/impulsive versus instrumental/premeditated aggression mediated by different transmitters? What physiological processes—such as increased or decreased transmission, activation of pre- versus postsynaptic receptors, transporter function, or gene expression—account for observed correlations between levels of neurochemicals and behavior? What, if anything, do levels of chemicals measured in plasma or cerebrospinal fluid (CSF) reveal about the activation of specific brain regions or anatomical circuits? The answers to these questions are necessary for best guesses about likely treatments for subtypes of aggression and for rational drug discovery. This subsection introduces the major themes of current research in this area.

Neurotransmitters

Serotonin. A vast literature exists demonstrating correlations between serotonergic function and impulsive aggression. It was discovered in the late 1960s that aggressive mice exhibit lower serotonin (5-hydroxy tryptophan, 5-HT) and 5-hydroxy indolacetic acid (5-HIAA) content and turnover. Evidence that low 5-HT function correlates with human aggression first came to light in studies of behaviorally disturbed marines with low CSF 5-HIAA. A large number of subsequent rodent, nonhuman primate, and human studies have confirmed that indicators of low serotonergic function are more common in impulsive aggression. These findings have led to the somewhat simplistic conclusion that serotonin is an aggression damper. The details suggest a more complex physiology.

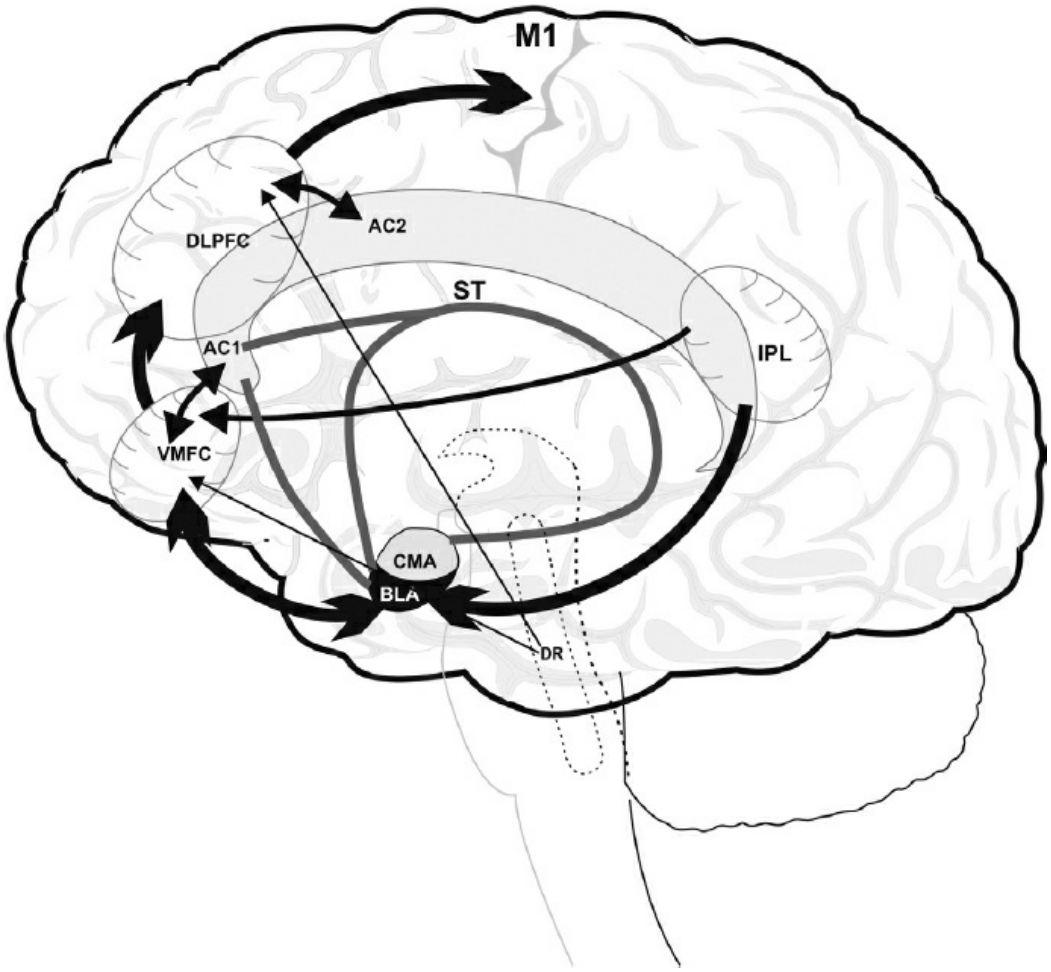

FIGURE 30–6. Liminal processing of aggression. Semischematic of the large cortical circuit loop mediating somewhat conscious processing of stimuli to which an aggressive response might be appropriate. AC1, anterior cingulated gyrus, infragenual portion; AC2, anterior cingulated gyrus, supragenual portion; BLA, basolateral amygdale; CMA, centromedian amygdale; DLPFC, dorsolateral prefrontal cortex; DR, dorsal raphe; IPL, inferior parietal lobule; M1, primary motor cortex; ST, stria terminalis; VMFC, ventromedial prefrontal cortex.

5-HT is produced primarily in the dorsal raphe nucleus toward the back of the pons. 5-HT receptors are expressed in many parts of the central nervous system that contribute to the regulation of aggression. These include the periaqueductal gray, basal ganglia, hippocampus, lateral septum, and, perhaps most important, the prefrontal cortex. Of interest, in a rodent model, studies reported no correlation between hypoarousal-related aggression and dorsal raphe activity. This is consistent with suggestions that the callous, premeditated, hypoaroused aggression of psychopathy is not mediated by the serotonergic system to the same degree as is impulsive/affective aggression.

The 5-HT_1 receptor family has been investigated most often in an effort to pin down the relevant physiology and to find promising treatments. 5-HT_{1A} agonists tend to decrease rodent aggression. Mice lacking this receptor subtype (5-HT_{1A} knockouts) are *less* aggressive, whereas 5-HT_{1B}

knockout mice are *more* aggressive. It remains unclear whether pre- or postsynaptic receptors are involved. These findings have raised hope for "serenic" drugs. Buspirone, the first 5-HT_{1A} drug used in clinical practice, has been reported in small trials to have antiaggressive effects. Experimental compounds acting on 5-HT_{1B} receptors, such as eltoprazine, anpirtoline, and zolmitriptan, reduce aggression in mice, including alcohol-intoxicated mice. However, cardiovascular side effects complicate efforts to identify a clinically useful 5-HT_{1B} serenic.

Since the introduction of the atypical antipsychotics, interest has increased in the role of the 5-HT_2 receptors. Drugs such as risperidone actually have more antagonist effect at 5-HT_{2A} than at D2 receptors. Risperidone decreases aggression in persons with several diagnoses, perhaps due to this 5-HT_{2A} effect.

Norepinephrine and Epinephrine. The catecholamines norepinephrine (NE) and epinephrine appear to play a different role in the mediation of aggression. The adrenergic system has important influences not only on central but also on peripheral functions that facilitate preparation for and response to challenging circumstances. In the periphery, adrenomedullary production of NE helps an animal face a threat such as an anticipated fight via sympathetic nervous system activation, adaptive changes in cardiovascular status, and mobilization of energy stores. One might think of this as the peripheral arousal function.

Norepinephrine receptors are also widely distributed in the brain, contributing to information processing and regulation of endocrine processes. The effects of NE on aggression may be mediated to some degree by its effect on corticotropin-releasing factor (CRF) and steroid hormone metabolism. In preclinical studies, the central NE system is activated when environmental events such as social challenges signal the possible need to aggress, as well as during both offensive and defensive aggression. Hence, binding to widely dispersed NE receptors serves as a "get ready for vigorous behavior" switch. Many adaptive processes are involved. For instance, NE enhances vigilance; it enhances olfaction, a key part of aggression motivation in many species; it supports memories of affective experiences, a very important factor regulating response to a social challenge; and it decreases somatic sensitivity via descending medullary–spinal pathways, allowing the animal to tolerate the pain of fighting. These indirect effects may account in part or in whole for the observation that NE activity is linked to aggression.

Because each subtype of adrenoceptor (α-1, α-2, β) is somewhat differently distributed in different brain regions, it is likely that regulation of aggression by pharmacological agents will depend on relative affinities for different receptors. The best evidence of serenic benefits comes from

both animal and human studies of β-blockers. Propranolol reduces predatory, defensive, shock-induced, and desipramine-induced aggression in rodents and has been particularly effective in the management of human post–brain trauma aggression. α-2 agonists such as clonidine have intriguing effects: In most rodent studies and in normal humans, clonidine increases irritability and aggression, but some evidence suggests that hyperactive or autistic children become *less* aggressive when treated with this drug. The clinical implication of these data may be that a balance between α-2 and β activity maintains healthful, socially appropriate aggressive responsiveness.

Dopamine. DA is primarily produced in the substantia nigra and the ventral tegmental nuclei of the midbrain. Like 5-HT, DA receptors are expressed on neurons in both the limbic system (especially the reward-related NAc) and the cortex. Based on elegant experiments using in vivo microdialysis, it appears that the mesocorticolimbic arm of the DA system, not the nigrostriatal arm, is involved in regulation of aggression. Unlike 5-HT, which seems directly to affect aggressive behavior, DA activity appears to play an indirect, permissive role. One theory is that aggression is often intrinsically rewarding. In the absence of DA stimulation of reward centers, motivation for aggression (and for many other behaviors) is decreased.

The pharmacology of DA and aggression is puzzling. D1 agonists and antagonists binding to D1, D2, and D3 receptors all decrease aggression. This is another hint that DA activity, in itself, does not initiate aggression. Both clozapine and risperidone sometimes reduce aggression in persons with schizophrenia. Despite the success of atypical antipsychotics for aggression management, concerns have been raised that (1) this effect is mediated not by DA but by 5-HT receptors, and (2) whatever role DA blockade plays in reducing aggression may be due to nonspecific motor effects.

γ-Aminobutyric Acid. γ-Aminobutyric acid (GABA) is a major inhibitory neurotransmitter with receptors widely distributed in the central nervous system. In animal studies, GABA agonists such as muscimol decreased aggressive behavior. Yet other pharmacological manipulations that would be expected to enhance GABAergic transmission paradoxically increased aggression. This finding may be related to the curious observation that benzodiazepines—which positively affect $GABA_A$ reception—sometimes make humans angry and aggressive. Animal studies of this phenomenon show that low doses of benzodiazepines increase whereas high doses decrease aggression. Yet dose alone may not explain the seeming paradox that some people become paradoxically aggressive when treated with drugs such as diazepam. Rodents that previously experienced social defeat are more likely to exhibited atypical responses to

benzodiazepines, and people with histories of poor impulse control seem more vulnerable to paradoxical benzodiazepine-induced anger. These findings suggest that social history may modify the relationship between GABA function and aggression.

Hormones

Testosterone. The relationship between the endocrine products of the testes and aggression has probably been apparent since the earliest observations of scrotal trauma in the distant past and the practical application of that knowledge to help domesticate large mammals. It is plausible that prehuman hominids had the cognitive capacity to make this connection. However, the association was first demonstrated scientifically in the nineteenth century. Arnold Adolphe Bertholdt observed in 1849 that castrated roosters lost their male secondary sexual characteristics (comb and wattle), which were then restored after injection of testicular extracts. Brown-Sequard created a new medical industry when he reported, in 1889, that his self-injection of testicular extracts from dogs and guinea pigs restored his manly vigor. One historian estimates that, within a year of Brown-Sequard's reports, 12,000 physicians were doing business injecting his "elixir of life." In 1920, Eugen Steinach offered a procedure to increase male vigor by purportedly increasing retention of the male essence by unilateral ligation of the ductus deferens—a surgery *to which* William Butler Yeats and Sigmund Freud *reportedly submitted*. Since the isolation of testosterone by Karoly G. David in 1935, a popular understanding has arisen that more testosterone means more aggression—perhaps the most widely held folk belief of psychoneuroendocrinology. The basic science is indisputable: Postpubertal males, who produce about 7 mg per day of the C_{19} steroid hormone testosterone, are much more physically aggressive, on average, than women, who produce about 0.3 mg per day. The actual relationship between steroid hormones and aggressive behavior is more complex.

First, rodent studies demonstrate that aggressivity of males depends on two stages of exposure to testosterone. A rodent will become an aggressive adult only if it has both been exposed to testosterone during a critical period of fetal development and exposed again in puberty. This implies that the antenatal development of a yet to be fully identified *male pattern* of central nervous system organization, perhaps related to the distribution of receptor-laden neurons, which predisposes the animal to a sex-specific behavioral predispositions. Second, testosterone does not act alone to produce its characteristic effects. The effect of testosterone in adult animals apparently depends not only on early organizational influence of this and other hormones but also on levels of estrogen, cortisol, and several amine transmitters. Third, at least in rats, testosterone seems to drive not just

aggressivity but also information processing. In experiments comparing male with female responses to features of novel environments, postpubertal males were much more attentive to the sensory cues one might use to detect the presence of other males and assess the dominance hierarchy. This was seen to enhance male-specific rapid matching of social sensations to motor responses. Moreover, an animal's past experience in competitive encounters alters its testosterone response to current challenges. Fourth, the research testing the hypothesis that more testosterone in human males means more aggression has produced mixed results. In a meta-analysis of 45 studies, a low but positive correlation of 0.14 was found between measures of testosterone and aggression measures. Plasma testosterone levels in teenaged boys are inconsistently associated with self-reported aggression, although they are perhaps better correlated with other-reported aggression. Plasma, saliva, and CSF testosterone levels in adult men are somewhat more consistently associated with aggression measures. Among prisoners, aggressive men generally have higher testosterone levels. However, a clue to an entirely different role of this hormone is that socially dominant but nonaggressive prisoners also exhibit higher testosterone levels. Testosterone is probably better regarded as an urge-to-dominate than an urge-to-harm hormone.

Testosterone levels are not constant. They are very responsive to both long- and short-term changes in social circumstances, such as position within a dominance hierarchy or success in fights. Among human prisoners, losers of fights have measurably diminished testosterone levels. Among nonprisoners, elevations for winners and decrements for losers have been reported in tennis, wrestling, and even chess matches. Several studies report that fans of losing soccer teams also experience drops in testosterone levels.

Important interactions occur between testosterone and amine transmitters. For instance, CSF 5-HIAA levels affect the degree to which testosterone activates sexual and other behaviors in animals, and CSF testosterone alters the likelihood of impulsive aggression in those with low CSF 5-HIAA. It is hypothesized that testosterone modulates the sensitivity of 5-HT receptors to alter serotonergic effects on impulsive aggression. Put simply, low 5-HT increases impulsive aggression and high testosterone increases competitive aggression. This serotonin–testosterone link may help explain why measures of testosterone alone do not correlate well with pathological human aggression: It may be that only those males who have the *combination* of high testosterone and low 5-HT are more likely to become violent.

These data together support several conclusions. First, the influence of testosterone on overt aggression depends both on fetal and pubertal brain exposure. Second, testosterone level can be conceived of both as a trait—the trait of high or low resting testosterone that somewhat predicts aggressive

reaction patterns—and state—the moment by moment and day by day fluctuations strongly linked to social events. Furthermore, whereas rough correlations are found between testosterone levels and aggression, high testosterone is probably more predictive of dominance seeking and dominance winning than of violence. Finally, testosterone hardly acts in isolation. Neurochemical interactions that help explain the role of this hormone in inappropriate aggression are just beginning to be uncovered.

Cortisol. Chronically low salivary cortisol levels are associated with disruptive, aggressive behavior in boys. Decreased cortisol levels have also been reported in adolescent girls with conduct disorder (CD). Yet not all findings are consistent with this low-cortisol–aggression association. In most studies of acute stress, cortisol rises. In studies of chronic stress, some investigators report persistent elevation and others report persistent depression of cortisol. Low cortisol may be a state that reflects long-term anxiety, or a trait—like the low heart rate reactivity often seen in antisociality—that indicates low arousal to threat. Further research is needed to determine whether the main connection between this hormonal condition and aggression is that individuals who are fearless, rule-breaking aggressors have low hypothalamic–pituitary–adrenocortical (HPA) activation or, instead, that individuals who have been chronically distressed will aggress due to hypervigilance and anxious escalation.

Peptide Hormones. More than 20 peptide hormones have been implicated in the regulation of mammalian aggression. Because of the limited confirmatory evidence in humans, just two of these will be mentioned.

The *nonapeptides* (nine amino acids) arginine vasopressin (AVP) and oxytocin have been studied for decades as moderators of mammalian social behavior. In one of the most dramatic demonstrations of this effect, male *prairie* voles have been shown to have differential genomic expression of and a much higher limbic density of AVP receptors than do *montane* voles. AVP–rich prairie voles are monogamous and exhibit extraordinary postcoital mate-guarding aggression. AVP–poor montane voles do not behave in this way. Limbic microinjections of AVP provoke aggression, an effect that is reduced by fluoxetine (Prozac) treatment, suggesting interactions between AVP and the 5-HT system. The effects of AVP in the rodent limbic system are also testosterone dependent. AVP receptors are present in the human amygdala and septal region. In one study, CSF AVP levels correlated with aggression among personality-disordered males. Research has yet to clarify whether central AVP plays a role in human males parallel to that in rodents and perhaps contributes to the high rate of jealous, mate-guarding violence.

Corticotropin-releasing factor is a 41–amino acid peptide that also helps

regulate aggression in rodents. The CRF receptors are densely distributed in the limbic systems, as well as in parts of the brainstem and the cortex. Injection of CRF into the amygdala accentuates the aggressive response to a strange male. Moreover, AVP modifies the aggression-inducing effects of CRF, suggesting yet another neurohumoral interaction.

Metabolism and Aggression. A novel field of inquiry is addressing the possibility that metabolic factors are associated with hostility and aggression, especially in men. In one study of 933 young men, controlling for testosterone levels, the most hostile and aggressive men had a cluster of risk factors for diabetes and heart disease, including being overweight and having low levels of high-density lipoprotein cholesterol. It is conceivable that this finding relates to the observation of increased aggression following the initiation of cholesterol-lowering treatments in some men. Many aspects of metabolism and endocrine function have yet to be investigated for possible links to aggression.

GENETICS AND AGGRESSION

Commentaries on the relationship between genetics and human aggression sometimes begin with the boldfaced disclaimer, "there is no 'aggression gene.'" Yet aggression is such a vital feature of human nature and variation in aggressivity is so patently apparent that it is very likely that one or more genes, acting alone or together—whose expression may depend on environmental factors—account for a significant proportion of the observed variance in trait aggressivity.

Because aggression is not a simple phenotype, studies aiming to identify relevant genetic factors must select some operational description of the subtype of aggressive behavior that they aim to explain. Elevated scores on self-rating aggression scales, measures of impulsivity, psychopathy checklist scores, laboratory measures of aggression, and criminal records may all serve as proxies for the trait of greatest interest: A tendency to commit socially nonsanctioned acts of violence.

Twin studies, adoption studies, studies of twins reared apart, and family and molecular genetic studies all suggest that genes are associated with aggression, psychopathy, or both. Influential Scandinavian studies reported that criminality among adopted away boys was much more strongly influenced by the criminality of their biological than of their adoptive fathers—early evidence of the potency of genetic influence. Twin studies show even stronger evidence of the heritability of antisociality and aggression than do adoption studies. Because environment is more likely to influence heritability in twin than adoption studies, this difference supports what neuropsychiatry has come to expect: gene–environment interactions.

More recently, molecular genetic studies have reported many specific associations between commonplace allelic polymorphisms and the risk of

aggressive behavior. The apparent damping effect of 5-HT on impulsive aggression and the permissive role of DA have both suggested promising targets for such inquiries. A subset of the notable findings is presented in Table 30–3.

Monoamine Oxidase A Gene Variants

The strongest link between a genetic variant and overt, pathological physical aggression is the reported association between a missense mutation in the gene for monoamine oxidase A (MAOA) and multigenerational violence among males in a Dutch kindred. Although this very rare mutation hardly accounts for a meaningful amount of human violence, this discovery opened the door to behavioral genetic studies of MAOA expression. The MAOA gene on the X chromosome contains a variable number tandem repeat region with differing numbers of copies of a 30–base pair (bp) repeat. Alleles 2 and 3 contain either 3.5 or 4 copies and are called the *high*-functioning variants; alleles 1 and 4 contain either 2, 3, or 5 copies and are the *low*-functioning variants. Individuals carrying the high-functioning variants express two- to tenfold more MAOA than those with the low-functioning variants. Many studies have reported that individuals with low MAOA expression exhibit aggressive traits and an increased frequency of violent offending. Neuroimaging studies help explain this phenomenon: Low-expression individuals have relatively decreased limbic volumes and, on functional imaging, exhibit increased amygdalar reactivity to aversive memories, increased amygdalar arousal to threat, and decreased prefrontal reactivity. This combination yields a volcanic potential, such that a neutral stimulus may be perceived as a threat, and the normal social restraints are loosened, significantly increasing the chances of an aggressive response.

Table 30–3.
Some Allelic Polymorphisms Associated with Human Aggression

Gene Product	Polymorphism	Behavior
MAOA	Allelic VNTR variants upstream of exon 1 on the X chromosome yielding low-enzyme activity	Conduct disorder, violent offending, antisocial behavior, impulsivity in males
TPH	TPH 779A or "U" variant in intron 7	Aggression, unprovoked anger, violent suicide
TPH	A polymorphism in the 3' region distal to exon 11	Violent suicide and impulsive aggression
5-HT transporter promoter (5-HTTLRP)	Short (S) allele containing a 44–base pair deletion	Conduct disorder, impulsivity, aggression, alcoholism, violent suicidality
5-HT transporter promoter (5-HTTLRP)	The 10-repeat variant of a 17–base pair VNTR in the second intron	Lower frequency of aggression in children

5-HT1D β	G861C variant	Increased aggression in children
COMT	Valine substitution for methionine at codon 158	Inconsistent association with aggression in different populations
TaqI DRD2	A1 allele	Antisocial traits, impulsivity, stress responsiveness, alcoholism
ABCG1 transporter (involved in sterol synthesis)	rs225374 G allele; rs914189 G allele	Increased self-rated anger and aggression
Cyclic adenosine monophosphate (cAMP) response element binding protein (CREB)	CREB 1 polymorphism, rs4675690	Anger expression in major depression, especially among males
DRD1	B2 allele 1	Psychosis and aggression in Alzheimer disease
TPH	A779C polymorphism	Aggressive hostility
TPH	A218C polymorphism	Aggression and anger in individuals without dementia and men with Alzheimer disease

COMT, catechol-*O*-methyl transferase; MAOA, monoamine oxidase A allele; TPH, tryptophan hydroxylase; VNTR, variable number tandem repeat.

Serotonergic Gene Variants

The gene for tryptophan hydroxylase occurs in two polymorphic variants: The more common L allele and the less common U allele (present in ~40 percent of whites). Several studies report that carriers of the U allele have higher scores on measures of aggression and impulsivity. Receiving much more attention are several allelic variants of the gene for the 5-HT transporter. This reuptake molecule regulates the magnitude and duration of 5-HT transmission and is highly relevant to clinical psychiatry as an initial target for antidepressant action. Short (S) and long (L) versions of the promoter region 5-HTTLPR occur, containing either a 44-bp deletion or insertion at the 5′-promoter region. About 40 percent of white carry one or more S alleles for the 5-HTTLPR, which confers less transcription efficiency, leading to less 5-HT expression and uptake. Many behavioral traits have been reported to be linked with the S allele, including anger-hostility measures, impulsiveness, CD, attention-deficit/hyperactivity disorder (ADHD), severe alcoholism, and violent suicidality.

Replication failures of studies linking 5-HT–related polymorphisms to aggression have inspired a search for both genetic and environmental interactions that modify the effect of a given allele. For instance, the association between the S allele of the 5-HTTLPR and aggression may be modified by the presence of genetic diathesis for alcoholism. An unexpected association between LL individuals and aggression may be related to antisociality in the biological parents. A great deal of research

examining such interactions would be needed to determine the size and direction of associations between any given allele and the risk for pathological aggression.

Catechol-*O*-Methyltransferase Gene Variants

A common single-nucleotide polymorphism exists at codon 158 in the gene for catechol-*O*-methyltransferase (COMT) in which valine (Val158) is substituted for methionine (Met158). The *val* variant has three- to fourfold higher enzyme activity than the *met* variant. In a somewhat puzzling series of reports, individuals with schizophrenia carrying the low-activity *met* allele were reported in one study to have an increased risk of homicidality and in another to have a decreased risk of aggression. Subsequent studies suggested that the direction of the association between *val* versus *met* allele and aggression may depend on (1) gender, (2) comorbid alcoholism, (3) comorbid schizophrenia, (4) comorbid bipolar disorder, (5) the aggression measurement method, and (6) whether self- or other-directed aggression is assessed. Over the last decade, a weak consensus seems to have emerged: Among male adults without schizophrenia or alcoholism, the high-activity *val* allele is associated with a cognitive style that involves leaping to decisions and perhaps impulsive behavior associated with relatively lower prefrontal function. These valiant *val* bearers may be, on average, more apt to address problems with action than deliberation. The *met* bearers exhibit superior prefrontal control and are perhaps less likely to imprudently aggress, although they may be more prone to anxiety and depression. A common trope is to refer to these types as "warriors and worriers." Theoretically, a cooperative society that distributes labor needs both types of behaviors.

Other Gene Variants

Several other genetic variations have been reported to contribute to aggression or related behaviors. The A1 allele of the gene for the DA DRD2 receptor is associated with reduced dopaminergic activity, increased stress responsiveness, impulsivity, and antisocial traits. The ABCG1 transporter, which plays a role in cholesterol metabolism, exhibits several single-nucleotide polymorphisms; carriers of at least one of these variants score higher on self-rating scales for anger and aggression. A single-nucleotide polymorphism for the memory-related transcription factor cyclic adenosine monophosphate (cAMP) response element-binding protein (CREB) has been reported to correlate with anger measures in depressed adults, especially among men.

Gene–Environment and Gene–Gene Interactions

An early study of male Swedish adoptees provided the first strong evidence

that genes interact with environment to influence *criminality* (note that this is not equivalent to aggression). Among 862 subject, those with neither biological nor adoptive family histories of crime had a 2.9 percent chance of becoming criminal; those who had only criminal adoptive families had a 6.7 percent chance; those with only biological criminal families had a 12.1 percent chance; and those with both biological and adoptive criminal influences had an astonishing 40 percent risk of becoming criminals. If biology and upbringing acted independently, one would expect about an 18.8 percent (6.7 + 12.1) chance of criminality for boys with both risk factors. The finding of a much higher-than-predicted rate supports the theory of synergistic gene–environment interaction. Several other adoption studies reported similar results specifically for antisocial traits.

Subsequent studies identified molecular genetic traits that may interact with adverse developmental environment specifically to yield an elevated risk of adult aggression. In the most cited (and deservedly historic) study, published in 2002, a large cohort of males was followed from birth to adulthood. As expected, those who suffered maltreatment in childhood were more likely to become antisocial adults. However, those who carried the high-activity MAOA variant *and* who suffered maltreatment seemed to be protected from this effect, implying that it is the *interaction* of a relatively low MAOA function with adversity that makes a boy prone to become an aggressive antisocial adult. Some subsequent studies confirmed this finding; others did not. One study reported that the protective effect of high MAOA enzyme activity was seen among white but not among nonwhite boys. Another recent study found that aggression among criminal alcoholic males was associated with the combination of higher CSF testosterone levels with the low-activity MAOA variant. Given that testosterone influences MAOA gene transcription, the net proclivity for aggression might depend on both factors. It is possible that this enzyme–hormone interaction helps explain not only the inconsistencies in studies of testosterone and aggression but also inconsistencies in studies of the MAOA protective effect.

Another facet of gene–environment interaction is the evidence of *environment-evocative* genetic traits. For instance, an association has been reported between antisocial traits in biological parents and harsh parenting by the adoptive parents. The thought is that children bearing antisocial traits evoke harshness in the rearing parents. This may in turn provide the adverse environment in which an innate tendency to agree will flourish.

A seven-repeat polymorphism for the DA receptor DRD4 has been repeatedly associated with the personality trait of novelty seeking. Because this trait seems to correlate with other traits associated with aggression, such as low arousal and extraversion, it is reasonable to surmise that genetic polymorphisms associated with novelty seeking will also predict aggression. This has not been demonstrated. However, one recent study

reported a gene–gene interaction: Males carrying *both* the A-1 variant of DRD2 *and* the seven-repeat variant of DRD4 were more likely to exhibit CD as adolescents or antisociality as adults.

DEVELOPMENT AND AGGRESSION

Because humans evolved to survive threatening environments in which aggressive skills were survival skills, it is readily apparent why aggression is woven into the fabric of development. The age of peak human aggression is between 2.5 and 3.5 years. Boys are more physically aggressive than girls. Some degree of antisocial, aggressive behavior is in fact *normal* in adolescent boys; the largest, longest, best-designed longitudinal studies, such as the Dunedin Study, have demonstrated that it is the rare boy who fails to exhibit such behavior. However, most people of both sexes moderate their aggression later in development, limiting conflict tactics to those generally regarded as socially sanctioned. The question addressed here is—genes aside—how nonsanctioned aggression develops. This question can be divided into several domains: (1) the question of whether aggressivity should be regarded as a trait or a disorder, (2) pre- and perinatal psychosocial and environmental risk factors, (3) postnatal psychosocial and environmental risk factors, (4) possible predictive markers, (5) the interplay of biological markers with psychosocial factors, and (6) the controversial issue of whether individuals fall into distinct categories of developmental trajectory.

Before embarking on this review, it is important to acknowledge a limitation: Regrettably, the vast majority of research has been focused on boys. Their indisputably higher rates of physical aggression emerge by preschool years, and the gender gap in physical aggression increases through adolescence. A direct/indirect aggression dichotomy has been proposed in which boys' externalizing aggression is far more likely to take the form of fighting, whereas that of girls is more likely to take the forms of teasing or reputation damaging. Some evidence suggests that oppositional and aggressive girls do not usually go on to exhibit nonsanctioned physical aggression but are at high risk to become anxious or depressed women. In fact, insufficient attention has been paid to studying aggression in girls and the overlap between the types of aggression in the two sexes.

Trait or Disorder

Unusually aggressive children represent a tremendous stress on the family and are at elevated risk for later social problems. However, a controversy exists regarding the best way to classify such behavior. From the viewpoint of evolutionary psychology, behaviors regarded as pathological in modern society might have been highly adaptive in earlier times. To consider CD a disorder, for example, may reflect modern values applied to aboriginal

behaviors. The "disruptive" boy might be closer to exhibiting what was, for 75,000 years, a preferable human trait. The child who sits quietly for hours in a chair at the age of 6 years might be the one exhibiting traits that, for the majority of the human career, were maladaptive. Put another way, there may be natural variation in human life-history strategies. The live-fast, die-young strategy may be an entirely normal trait.

For this reason, is often difficult to determine whether a disruptive, aggressive child is expressing a common variant of human nature or a psychobiological disorder. Although the clinician can determine whether the behaviors match a published algorithm of diagnostic criteria, it remains very uncertain whether those phenomenological patterns reliably identify pathological conditions. The classification systems of both the text revision of the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR) and tenth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) may be inappropriately categorical. They claim distinctions among three diagnoses associated with aggression: CD, oppositional defiant disorder (ODD), and ADHD. These three supposedly separable conditions have not been shown to be pathophysiologically discrete. One can appreciate the historical rationale for the DSM stopgap approach of operational criteria and yet be aware that plugging children's observed behaviors into a menu of observational criteria to generate a diagnosis is scientifically simplistic. Advances in neuropsychiatry are likely to mandate a reconceptualization of the origins and types of childhood aggression.

Pre- and Perinatal Risk Factors

Many studies implicate prenatal or perinatal events in later antisocial and/or violent behavior. Among the psychosocial risk factors implicated are poverty, maternal teenaged first pregnancy, single parenthood, marital discord, and low maternal intelligence quotient (IQ). Certain biotoxicities also clearly influence the rate of subsequent aggression. Fetal exposure to alcohol increases the risk of childhood CD, although, perhaps surprisingly, it remains unclear whether such exposure increases the rate of adolescent and adult violence. The alcohol effect is independent of the formal diagnosis of fetal alcohol syndrome, but it is not clear at what threshold of ethanol exposure the infant is at risk. In a study of children with fetal alcohol syndrome, a stable home environment reduced the risk of later antisocial behavior, suggesting an important environment–environment interaction.

Fetal exposure to maternal smoking also increases the risk of CD and adult violent criminality. Carbon monoxide and nicotine may be the responsible neurotoxins. The effect of fetal smoking exposure is exacerbated by other risk factors such as obstetric complications or teenaged pregnancy. In one study, for example, the risk of violent offending

was five times as greater when exposure to smoking was combined with delivery complications. Nutrition is another risk factor. In one study, Dutch men whose mothers survived a wartime famine during pregnancy went on to exhibit increased rates of antisocial behavior. Fetal cocaine exposure has recently been shown to increase aggression in boys at the age of 5 years. Prenatal lead exposure is strongly associated with childhood aggression and less strongly with adult aggression. Some evidence suggests a similar effect of mercury. Polychlorinated biphenyls, which may alter steroid hormone function, have been associated with an apparent switch in gender roles in which exposed boys exhibited less aggressive and girls exhibited more aggressive play patterns. One major gap in knowledge is the association between prenatal exposure to other potential environmental toxins and aggression. Approximately 3,000 chemicals are produced at the high volume of more than 1 million pounds per year. Of these, only 12 have been studied for their effects on the developing nervous systems. A better picture of the realistic risks represented by these agents will not be achieved for decades.

Delivery complications are also a risk factor for antisocial and violent behavior. When birth complications are combined with an "adverse" family environment—such as maternal separation or marital discord or poverty or familial criminality—the risk of violent criminality is even higher. Infants born with minor physical anomalies (e.g., low-set ears, cleft lip and palate) also have an increased risk of later criminality. Such anomalies often reflect maldevelopment in the first trimester, affecting ectodermal germ layers that become integumentary and neural. Again, when such birth anomalies are combined with problems in family life, violent offending is especially likely to occur.

Postnatal Factors

Many postnatal factors have also been identified, with various degrees of certainty, as risk factors for aggression. As in the case of prenatal factors, it is challenging to determine with confidence the degree to which any one factor accounts for the human variation in atypical aggressiveness. Because many aggressive individuals have experienced several potentially disadvantageous postnatal factors, it is difficult to isolate the effect size of, for instance, poverty versus unstable food supply, or marital discord versus inadequate paternal engagement, or abusive discipline versus unreported traumatic brain injury (TBI). Postnatal risk factors for aggression are typically divided into psychosocial versus biological, but this is a false dichotomy because (1) a social factor such as poverty might be associated with a host of biological variables, from prenatal health care to missed vaccinations to ongoing nutritional problems to lead exposure, and (2) ultimately all psychosocial factors act by changing the brain (e.g., acute stress affects the hypothalamic pituitary axis in ways that can damage the

hippocampus).

Several specific postnatal toxicities are associated with aggression. Among the best established is the strong relationship between lead and aggressive delinquency. The major source of poisoning is deteriorating lead-based paint. Approximately 25 million homes in the United States have such paint, a risk factor that will decline only with the eventual replacement of the housing stock. A linear relationship has been reported between serum lead levels in childhood and aggression in adolescence. Volatile organic compounds may also increase the risk for aggression. Styrene exposure, in particular, has been implicated.

Regarding nominally social factors, postnatal exposure to marital discord, single parenthood, and domestic violence all increase the risk of childhood and adolescent aggression. Many studies suggest an increased likelihood of aggressive/disruptive behaviors in childhood, adolescence, and adulthood if children are either abused or neglected. It has been proposed that these stressors lead to aggression by somewhat different mechanisms. Abuse may teach hypervigilance, a hostile attribution bias, and the use of violence to achieve goals. Neglect, in contrast, may lead to impaired emotional regulation due to deficient interactions. Statistical data link both problems to later violence. Maternal neglect or rejection in infancy, in particular, has been shown to interact with genes to elevate the risk of aggression. As mentioned earlier, the combination of genetically mediated low MAOA activity and early maltreatment is a strong predictor of adult antisociality.

Some studies report that early exposure to either abuse or neglect is associated with specific neurological changes, including epileptiform discharges, focal brain volume reductions, and decreased synaptic density of the hippocampus. A convergence of very strong preclinical and human clinical evidence supports the hypothesis that medial temporal lobe integrity is at risk in cases of psychic trauma. It has been proposed that early stress might alter the role of the amygdala in interpreting threat, whereas later stress might compromise the prefrontal lobe's ability to restrain impulses. It is entirely possible that psychological stressors increase the risk for other focal neurological problems, and it is virtually certain that advances will identify genetic risk or protective factors that alter the effect of such social stress on brain development. At present, it is uncertain exactly how, and how often, abuse and neglect produce neurobiological changes underlying aggression.

Parents must discipline their children, although the style of such discipline varies. Harsh, power-assertive, physically controlling parenting apparently increases the risk of children becoming aggressive. However, this finding seems to be modified by the overall quality of the relationship. Parents who are harsh disciplinarians but are nonetheless sensitive and communicative at other times might not rear aggressive children. Several

authorities emphasize that *inconsistent* discipline may be even more damaging than *harsh* discipline. The empirical evidence for this claim is limited.

Peers

By late middle childhood, roughly after the age of 10 years, peers loom much larger both as role models and sources of psychically salient rewards and punishments. Peer acceptance or rejection begins displacing parental feedback as a determinant of mood, and peers begin displacing parents as the source of definitive practical knowledge. This is the age at which peers begin to recognize the disturbing qualities of the aggressive subgroup and the age at which peer influences more deeply affect aggressiveness.

Two potentially vicious feedback loops can occur. First, a child who is more aggressive may experience peer rejection, which leads to hostility and increased aggressiveness, which leads to further rejection. Second, aggressive boys will begin to find each other and associate, which reinforces their initial aggression, especially among those more prone to peer pressure. In such groups, aggressive displays become competitive and rewards accrue to the most violent surviving risk takers. In the absence of potent alternative paths to self-respect, the bad company effect may propel otherwise moderately aggressive boys toward violence. Gangs, for some, provide an irresistible promise of affiliation, identity, and warrior bonds. The negative feedback for aggression that might otherwise come from nonviolent peers and parents is replaced by positive feedback for violent behavior until, in many cases, the cycle is interrupted by maturity or death.

Media and Aggression

Television viewing occupies more hours for adults in the developed world than any activity apart from work and sleep. There are more televisions in the average US household (2.7) the people (2.6). An average hour of children's television programming contains 25 to 30 instances of violence. Yet the precise effect of television watching on aggression remains in dispute. A very large number of studies summarized in two meta-analyses reported that childhood television watching is associated with higher rates of aggression. However, most were very short-term studies, and, in one of the only studies to follow children through to adulthood, the viewing/aggression association could be explained by the observation that aggressive children chose to watch more television. In perhaps the best project addressing this question, the Children in the Community Study followed children from a mean age of 5.8 years in 1976 to the age of 30 years in 2000. Television viewing in hours per day at the age of 14 years was strongly associated with aggressive acts against another person at the age of 16 and 22 years. Note that the content viewed was not a factor; the

time spent viewing (less or more than 1 hour per day) alone predicted aggression. Neglect, growing up in an unsafe neighborhood, income below the US poverty level, low parental education, and psychiatric disorders were all associated both with more time spent watching television and with aggression. Much less is known about the effect of playing video games or viewing violent movies. No studies have examined how biological variables might interact with this essentially ubiquitous environmental exposure to modify its apparent negative effect.

Proposed Predictive Markers

Some authorities hypothesize that patterns of infantile attachment behaviors predict later aggression. *Disorganized attachment,* meaning inconsistent responses to separation stress, is reportedly more predictive of aggression than is either *secure attachment* or even *avoidant attachment* (lack of protest on separation). Large body size, stimulus seeking, and fearlessness at the age of 3 years predicted aggression at the age of 11 years in one study. Impulsivity and inattention in preschool years predict later aggression in many studies, although only a minority of children with these traits ever commit serious violence. Behavioral inhibition—the tendency to react fearfully to strange persons or novel situations—is claimed to *reduce* the likelihood that disruptive boys will become delinquent adolescents compared with similar, extroverted boys. However, children who are described as inhibited might also be described as overcontrolled, a trait that is sometimes associated with adult-onset violence in both sexes. Several other childhood traits seem to predict later aggression. For instance, in one study, boys who were highly oppositional and hyperactive in kindergarten exhibited three times the risk of persistent aggression. However, boys who were *overtly* physically aggressive in kindergarten only remained so when mothers had both low education and teen onset of childbearing. Cognitive markers for later aggression include (1) *hostile attribution bias,* in which the child misinterprets neutral stimuli as threat, and (2) lower-than-age-appropriate problem-solving skills in which nonaggressive options fail to be considered.

Physiological traits have also been associated with increased rates of psychopathy and aggression. In particular, a syndrome of autonomic hypoarousal, with lower resting heart rate, decreased heart rate reactivity, and less skin conductance reactivity to stress, has been strongly correlated with CD and APD. This pattern is hardly determinative; aerobic fitness also lowers resting heart rate. Nonetheless, hypoarousal, more sensation seeking, and limited empathy appear to cluster and perhaps share a genetic basis. The callousness of violent persons with APD has often been compared with predatory animal behavior, and, in fact, carnivores exhibit no rise in cortisol and little autonomic arousal when stalking and killing. Of interest, the relationship between hypoarousal and antisocial behavior is

stronger in individuals from upper socioeconomic groups and intact families. In other words, the effect of autonomic nervous system function on psychopathy might wash out among those who already suffer social disadvantages that increase their risk; the contribution of hypoarousal may be easier to detect when other confounding risk factors are absent.

Other potential biomarkers for aggression have been investigated, especially markers of central 5-HT function. In several studies, the prolactin response to fenfluramine, a marker of central 5-HT activity, was reduced in children and adults with aggression. This marker may also be environment dependent: In one study, the association between low central 5-HT and parent-rated aggression was conditioned by socioeconomic status, parental antisocial traits, and number of children in the home. This finding potentially supports a stress–diathesis model in which a presumably innate biological trait either protects from or synergizes with adversity, the latter case yielding aggression. It remains to be determined to what degree central 5-HT function is an environmentally modifiable genetic trait.

Socialization to Violence?

In attempting to explain why some communities experience elevated rates of violence, contemporary criminology overlaps with the sociology of conflict. Just as urban ghettos seem to beget cultures of violence, so do children growing up amid civil or other war, or refugees from such wars, exhibit elevated rates of aggression. Stress can be posited as the common denominator. Structural theories have yet to identify the precise combination of elements of stress most likely to promote violent communities. Racialization, poverty, low parental education, and female family leadership are four factors that appear to restrict urban residential choices, but these factors in themselves do not account for the emergence of violent crime. Posttraumatic stress disorder (PTSD), family dissolution, and displacement among victims or refugees of war are well-known stressors associated with psychiatric disorder, but no specific associations have been reported between these factors and community violence rates. It is plausible that a geographic concentration of individuals who experience innate and/or developmental risk factors leads to a snowball effect, such that aggressive behavior is not only unusually necessary but also well rewarded, nurturing incipient proclivities for violence into full-blown homicidal behaviors. Similar forces may be at work in Detroit, Colombia's Medellin, and Kosovo's Mitrovica. Insufficient research has been conducted to regard this hypothesis as more than speculation. However, the concept of "socialization to violence," based on social learning and social identity theory, is not adequate to account for very violent places, given the tremendous individual variation in aggression observed among those reared in adverse environments.

Intimate Partner Violence

Domestic or intimate partner violence is difficult to classify. Although verbal and physical aggression is very common in romantic relationships, it is unclear whether to regard this phenomenon as a unique category of aggression, perhaps related to evolutionary imperatives to control access to a mating resource, a developmental stage–related phenomenon (humans tend to be especially controlling and jealous during early adulthood), a disorder of attachment, or simply the inevitable result of the fact that individuals who are pathologically aggressive will tend to aggress wherever they spend time—and many spend the majority of their time in the company of intimate partners. All of these theories probably apply. Domestic and intimate partner violence is an extremely common worldwide problem. Because of the availability of more rigorous research findings, this discussion is confined to the US experience, but intimate partner and gender-based violence may be significantly more or less prevalent in other countries.

Sources vary in estimating the prevalence of minor and major physical aggression and sexual abuse. Nearly 25 percent of US women report having been assaulted sexually or physically by an intimate partner in their lifetime. Minor incidents of pushing, shoving, slapping, or throwing objects occur in about 15 to 18 percent of US marriages. About 1 percent of men report more severely abusing their spouses (hitting with a fist, kicking, attacking with objects), whereas 5 percent of women report having been severely abused, of whom 7 percent report requiring medical attention. An estimated 85 percent of the victims of intimate partner violence are female and 15 percent are male, although reporting bias may account for some of this imbalance, since women are about 14 times as likely to report suffering physical assault from an intimate partner. Approximately one-third of all homicides of women are committed by an intimate partner, compared with less than 4 percent of homicides of men. Dating violence is getting long overdue attention. In one survey, 13 percent of college women reported having been forced to participate in sex in a dating situation. An unknown proportion of date victims also suffer other injuries.

Certain traits and diagnoses are associated with men who assault their intimate partners. On average, such men are more hostile, impulsive, and power-oriented and tend to hold stereotyped views of gender roles. In some Minnesota Multiphasic Personality Inventory (MMPI) studies, male perpetrators scored higher on the Psychopathic Deviant scale. Rates of borderline personality disorder (BPD), APD, anxiety, and depression are higher among spousal batterers. A majority of male spouse abusers were themselves abused in childhood. Unemployment is a risk factor. Although domestic violence is often associated with alcohol intoxication, at least 76 percent of perpetrators have been violent when sober. Recently, the

apparent increase in the rate of domestic violence among veterans has prompted research suggesting that combat-related PTSD is an important cause of intimate partner abuse and even homicide. Risk factors specifically for spousal homicide, in addition to the foregoing, include cohabitation, a history of stalking, sexual jealousy, and threatened separation.

It is likely that a subset of those who recurrently perpetrate domestic violence—men in particular—are neurobiologically different from those who rarely or never commit such acts. Research has identified five threads of evidence in this regard. First, perpetrators of domestic violence may have autonomic dysregulation. However, it is unclear whether hypoarousal/emotional callousness of the psychopathic type is more common than is hyperreactivity to stress. It is entirely possible that two offender populations exist with different types of dysregulation. Second, serotonergic function may be depressed in perpetrators. CSF 5-HIAA concentrations are reportedly lower in this group. Third, testosterone levels tend to be higher among men who commit domestic violence, although some evidence suggests that the association between testosterone levels and aggression obtains only in alcohol-dependent men. Fourth, perpetrators of domestic violence may have abnormal threat reactivity. This theory is supported by the observation that sodium lactate infusion—a precipitant of panic—induces rage at a higher rate in perpetrators than in control subjects. Fifth, a history of substance dependence, perhaps suggesting central reward deficiency, is very common. In one study of alcohol-dependent male perpetrators, lower cerebral glucose metabolism was reported in the right hypothalamus, and a decreased correlation was found between cortical and subcortical metabolism. The significance of these last findings is unknown. One speculation is that these men have irritability related to poorly integrated attachment processing. Even if a subset of individuals who commit domestic violence are biologically atypical, and even if an evolutionary imperative lurks in the chain of causality, intimate partner violence is not a homogeneous neuropsychiatric disorder but a final common pathway for multiple biopsychological problems.

Life Courses of Aggression

A debate exists about the typical life course or alternative trajectories of aggression. Based on the landmark Dunedin study, it was proposed that two major patterns were distinguishable: A dangerous life-persistent pattern of aggression starting in early childhood versus a much more benign adolescent-limited pattern. (Very few subjects reported no lifetime aggression, again suggesting the normalcy of much aggressive behavior.) Evidence from other sources, however, suggests that there are at least three major life trajectories, two of which can be subdivided: (1a) early childhood onset, limited duration; (1b) adolescent onset, limited duration; (2a) early

childhood onset, life course persistent; (2b) adolescent onset, life course persistent; and (3) adult-onset aggression.

It is indisputable that boys' expected physical aggressiveness occurs in two waves: The first during the toddler years prior to gaining self-control and the second during mid to late adolescence. Thus, the most common type of aggression is type 1a, the early childhood onset, limited duration type, given that the peak of human interpersonal aggression occurs at about the age of 2.5 to 3.5 years. This pattern is essentially normative. The most common pattern of clinically referred aggression is type 1b, the adolescent-onset, limited-duration type. Adolescent aggressiveness is certainly related to hormonal factors, although several developmental theories add depth to the understanding of why adolescent boys aggress. First, the role of parent as disciplinarian shifts from control toward monitoring, and the child resists that monitoring in his quest for autonomy. Second, the adolescent period forces the child to deal with increased social and cognitive demands with an incompletely developed central nervous system. The relatively slower myelination of the DLPFC leads to a mismatch in which the child is virtually ready for independence but lacks the judgment and perspective to optimally assess risks and benefits of action alternatives. Although it is very distressing and potentially dangerous, parents can be reassured that this type of aggression is common and usually time limited.

The early-onset, life-persistent type, type 2a, is the best candidate for reflecting an innate and abnormal biological condition, the expression of which is perhaps triggered by environmental factors. An example might be a genetic risk factor combined with parental neglect. Those with the adolescent-onset, persistent type of aggression, type 2b, may be perhaps especially vulnerable to peer influences or to steroid hormone changes, or to share with persons with schizophrenia or other major mental disorders the poorly understood mechanism by which behavioral atypicality is not apparent until the end of development. Both of the persistent types exhibit considerable risk of serious adolescent violence, including gang violence and adult offending. The least-understood pattern is adult-onset aggression, type 3. A commonplace is the statement that although not all aggressive children become aggressive adults, every violent adult was an aggressive child. The occasional occurrence of the serious aggression with its first appearance in adulthood belies this claim. One theory is that individuals who are overcontrolled but simmering during their developmental period may suddenly—and inexplicably to all appearances—explode with acts of serious adult violence. It is unknown how often such events are triggered by a psychologically meaningful provocation alone or by the co-occurrence of such a provocation with a second factor such as TBI, covert epilepsy, substance intoxication, cerebrovascular disease, or other organic brain condition.

NEUROPSYCHIATRIC DISORDERS AND AGGRESSION

Childhood Neuropsychiatric Disorders and Aggression

Aggression is among the most common reasons for psychiatric referral of children and adolescents. Not only are up to 60 percent of referrals to outpatient child psychiatric clinics explicitly for evaluation and treatment of aggression but serious aggression is a common occurrence in the 6 months prior to referral *regardless* of the diagnosis. Several distinct disorders frequently associated with aggression are recognized by the DSM classification system. These include Mental Retardation (MR), Autistic Disorder, Pervasive Developmental Disorder, ADHD, ODD, and especially CD. In addition, mood disorder, anxiety disorder, Tourette disorder, and disorders associated with psychosis are associated with an elevated frequency of nonsanctioned aggression. Some clinicians encountering aggressive children assign more than one label, such as "ADHD with comorbid CD" or "autism with comorbid ADHD"; as many as half of those with ODD or CCD are estimated to also have ADHD. However, although a child can clearly meet DSM operational criteria for more than one diagnosis, it is not clear that disruptive behaviors are pathophysiologically discrete or that a complex clinical picture represents two different conditions. In the case of autism (and to a lesser extent, ADHD), research seems to be honing in on truly distinctive neurobiological features. This heralds the coming era of prevention and treatment targeting the essential psychobiological problem. In contrast, very different underlying conditions can share the indistinct labels *mental retardation, oppositional defiant disorder,* or *conduct disorder*. This complicates diagnosis, treatment trials, and rational drug discovery.

Suffice it to say that one of neuropsychiatry's biggest concerns is the lack of definitive markers that would allow the determination of the most likely underlying causes of pathological aggression, whatever the diagnosis. For instance, physical aggression observed in a child with MR due to birth hypoxia may be phenomenologically identical to and biologically related to similar behavior observed in a child with lead exposure, TBI, or bipolar psychosis. Moreover, two individuals both meeting criteria for the same diagnosis may exhibit profoundly different types of aggression. For instance, a 13-year-old boy who has conned others, forged a check, stayed out late, and run away from home but who does not commit interpersonal physical aggression shares the diagnosis of CD with a 17-year-old serial rapist. A large volume of literature avoids the diagnostic issue altogether and simply refers to *disruptive behavior disorders of childhood*. For all these reasons, it is sometimes more helpful to focus on the type of aggression than on the DSM diagnosis.

That having been said, a brief comment regarding the spectrum of aggression observed in children fitting currently popular diagnostic

categories may be helpful. To provide a sense of proportion, in one study, childhood disruptive disorders were diagnosed in 32 to 48 percent of children in the juvenile justice system. From 1.3 to 12.7 percent of these met criteria for ADHD, 3.1 to 14.7 percent met criteria for ODD, and 29.9 to 40 percent met criteria for CD. Potential interventions for each are discussed in the treatment subsection.

Aggression and MR

The term MR represents a basket for many undiagnosed problems. Birth hypoxia, fragile X syndrome, inborn errors of metabolism, or undetected infantile meningitis may all result in developmental delay. Most cases of MR are never fully evaluated to determine the underlying neurobiological disorder. In 2007, the American Association on Mental Retardation changed its name to the American Association on Intellectual and Developmental Disabilities. This change reflects the fact that MR is not a specific entity but a classification of those needing certain supportive services and protections. Both the IQ measurement and behavioral features of individuals with such disabilities vary greatly. For this reason, aggression associated with MR may consist of behaviors as varied as biting, throwing feces, or premeditated murder. Childhood aggression is a common clinical compliant in MR clinics and is regarded as a clinical problem in 16 to 30 percent of developmentally delayed individuals living in residential treatment settings. The rate of criminal violence is somewhat elevated among adults with MR.

MR is often comorbid with hyperactivity, impulsiveness, depression, or substance abuse. These variations prevent generalizations regarding aggression in MR, apart from the comment that those with diminished general intellectual capacity are statistically more likely to also have central nervous system problems that (1) delay brain maturation, (2) impair the ability to learn from experience or to remember lessons, (3) impair the ability to restrain impulses, (4) deter learning complex expectations of social and legal conduct, (5) cause misinterpretation of social cues, (6) limit noncriminal opportunities for gainful employment, and (7) expose individuals to situations in which others take advantage of or exploit them. All of these factors potentially increase the risk of aggression throughout the life course.

Aggression and Autism

The autistic spectrum illnesses, including autistic disorder (AD) and Asperger disorder, appear to have increased in incidence (current estimates range from 2 per 1,000 to 1 per 200 live births) for reasons that cannot be accounted for by increased diagnostic vigilance. Aggression is commonly reported by parents of autistic children; much of that aggression is

hyperactivity related, consisting of self-directed injury or nondirected tantrums. Other-directed aggression may be a transient developmental stage that occurs at different ages, depending on the autistic child's rate of intellectual maturation. Autistic spectrum illnesses share an underlying deficit in affiliative behaviors or empathy; thus, one theoretical explanation for other-directed aggression in autism is lack of empathy. Of interest, autistic children have been reported to be empathetic with nonhuman animals. This, along with the relatively low rates of adult violence committed by patients with these diagnoses, suggests that other-directed aggression in autism is *not* primarily driven by lack of empathy so much as by low frustration tolerance, perception that others are strange and threatening, and delayed maturation of self-control. About 20 to 33 percent of autistic children develop epilepsy. One issue deserving investigation is whether comorbid epilepsy increases the risk of aggression in these children.

Aggression and ADHD

Aggression is a very common complaint among parents of children diagnosed with ADHD. Indeed, a debate exists between British psychiatry, where aggressive children with hyperactivity are often diagnosed with CD, and American psychiatry, where such children often receive the diagnosis of ADHD. Because ADHD varies in type and severity, it is likely that children who meet criteria for this diagnosis will aggress for different reasons. Estimates of the contribution of ADHD to juvenile delinquency vary; in one study, more than 42 percent of delinquent adolescents met the criteria for ADHD. Of interest, the prevalence of ADHD in female delinquent adolescents may be even higher than that among males. Impulsivity or lack of self-control and self-monitoring is the usual behavioral explanation for aggression in ADHD. However, other mechanisms, such as frustration, low self-esteem, distress at social exclusion, or stimulant side effects, may better explain aggression in many cases. Evidence exists that measures of impulsivity and lower IQ in ADHD correlate with criminality but not specifically with aggression. Therefore, although aggression in this group is classified as hyperarousal related, the mechanism by which some patients express high arousal via violence and the probable role of environmental modifiers remain inadequately investigated.

Aggression and ODD

ODD is another umbrella term weakly tethered to any consistent psychological mechanism and not shown to have any biological underpinning. Some oppositional behavior is normative in the toddler and adolescent years, frustrating efforts of supporters of this diagnosis to

demonstrate that ODD is a discrete pathological entity. It appears that coercive parenting combined with an independent and rigid childhood temperament may lead to a cycle of mutual negative reinforcement. A spectrum of oppositionality exists, but it is not evident that ODD is a mental disorder or that aggression in children who more frequently resist authority is neuropsychiatrically different from aggression in children who less frequently resist authority.

Aggression and CD

CD is one of the most frequent diagnoses in outpatient psychiatric clinics and, by far, the most common diagnosis in disruptive children. This high frequency reflects the fact that CD is an umbrella term applied to a very wide variety of behaviors of greatly varying severity. Much further research is needed to determine whether a core entity exists, meaningfully different from one part of the spectrum of human normalcy, justifying this diagnostic term. Given that CD is strongly associated with male sex, that rates of CD are elevated among adopted-away children of antisocial parents, that some children with CD exhibit autonomic hypoarousal, and that roughly 40 percent of those meeting the criteria for CD will subsequently exhibit life-persistent antisociality, at least some children with this diagnosis may be exhibiting genetically influenced central nervous system atypicality. Social rule violation is necessary for the diagnosis of CD but aggression is not. Thus, only a subset of children meeting the diagnostic criteria will display fighting, bullying, cruelty toward persons or animals, confrontational robbery, or forcible sexual aggression. One theory is that boys who express higher than the usual level of adrenal androgens or experience early-onset puberty with testicular expression of androgens are more prone to CD. Another theory points to the evidence of autonomic hypoarousal and reward insensitivity, which perhaps drive sensation seeking. It seems likely that CD will eventually be disaggregated into (1) genetically influenced, hypoarousal-related disruptive behaviors, (2) genetically influenced, hyperarousal- and impulse control–related disruptive behaviors, (3) mood disorders and related emotional dysregulation syndromes, (4) behavior associated with acquired brain injuries and toxicities, and (5) the behavior of normal children at one end of the spectrum of autonomy seeking. Pending a better understanding of so-called CD, it seems more important to carefully characterize a child's disruptive behaviors than to make the diagnosis and assume that this captures the problem or points to the solution.

Aggression Associated with Primarily Adult Neuropsychiatric Disorders

The notion that mental illness causes adult violent behavior is at least as ancient as the dialogues of Plato, and many people today believe that there

is a causal association between mental illness and aggression. Monahan et al., for instance, reported that 61 percent of US survey respondents believed that people with schizophrenia were either very or somewhat likely to commit acts of violence. Indeed, part of the stigma of mental disorder derives from the fear that individuals with such disorders may hurt others. Given the prominence of media reports linking these issues—for example, Unabomber Theodore Kaczynski's probable schizophrenia, or Virginia Tech mass murderer Seung-Hui Cho's reputed mental illness—social expectations of the mentally ill are colored with a hard-to-shake image of threat. The data are more complex than the simple story suggested by popular mythology.

Several types of studies provide relevant information: (1) psychometric measures of mental disorder among offenders, (2) measures of the prevalence of diagnosed psychiatric disorder among offenders, (3) measures of the frequency of aggressive acts committed by persons diagnosed with psychiatric disorders, and (4) longitudinal cohort studies.

In the first category, many studies, most using the MMPI in prison populations, report elevated prevalence of abnormal results on certain scales. Edwin I. Megargee et al., for example, tested 6,350 inmates of federal prisons. They found elevation above the critical 97 percent percentile level on Depression scores in 20 percent, Mania in 37 percent, Schizophrenia in 24 percent, Paranoia in 22 percent, and Psychopathic deviation in 58 percent. This approach is problematic for several reasons. First, scale elevations are not psychiatric diagnoses. Second, prisoners may feign or exaggerate mental illness, and standard MMPI validity scores have been shown to fail to detect such malingering. Most important, in the absence of a comparison group, it is not possible to interpret these results as suggesting psychological abnormality that is higher or lower than population base rates.

In the second category, many studies report remarkably high rates of clinically diagnosed mental illness among incarcerated violent felons. Descriptive studies of convicted murderers, *excluding* those acquitted by reason of insanity, report a 60 to 100 percent prevalence of diagnosable psychiatric disorders. However, in one large comparative study, the rates of mental illness were no different among homicide offenders, nonviolent offenders, and community dwellers. The question is also raised to what degree prisoners first exhibit mental illness postconviction, suggesting that either the stress of incarceration or malingering elevates prevalence rates above those that one would expect among preconviction offenders. One Canadian study of 4,156 persons convicted of serious offenses helps address this issue. The rate of preoffense mental health treatment for the 162 murderers in this sample was a strikingly high 36.9 percent. In the same study, the odds ratio of the association between prior diagnoses and conviction for a violent offense (n = 1,998) was 4.4 for schizophrenia, 4.1

for affective disorders, 9.5 for substance misuse, and 18.7 for personality disorders. These findings suggest that personality disorders and substance disorders account for a large part of the relationship between known prior mental health treatment and conviction for violence.

The third group of studies, those attending to criminality among the mentally ill, is the largest and most cited category of research. Henrik Belfrage, for example, reviewed criminality among 893 Swedish psychiatric patients and found a 28 percent rate of criminality within 10 years of hospital discharge. In this sample, only patients with schizophrenia had committed serious violent offenses. A later massive study compared criminal and psychiatric records in a Danish birth cohort of more than 350,000 and found elevated rates of offending among those with mental illness, but it did not compare rates of violence between psychiatric and nonpsychiatric samples.

In one of the most rigorous of the studies in the third category—the MacArthur Study of Mental Disorder and Violence—951 English-speaking patients between the ages of 18 and 44 years were followed for 1 year after their psychiatric admissions. The follow-up group was divided into three groups: 397 with "major mental disorders" (Axis I) without substance abuse, 392 with major mental disorders combined with substance abuse, and 142 with "other" mental disorders, mostly personality disorders or adjustment disorders, combined with substance abuse. At 1 year, the major disorders group had a 17.9 percent rate of violence, the major disorders plus substance group had a 31.1 percent rate, and the "other" plus substance group had an extraordinary violence rate of 43.0 percent. The study disaggregated the major disorders group, finding a 1-year prevalence of violence of 14.8 percent for those with schizophrenia, 28.5 percent for those with depression, and 22.0 percent for those with bipolar disorder. Substance abuse or dependence added to the risk, increasing the average 1-year violence rate for these three mental disorders from 18.1 percent to 31.3 percent. These findings and others suggest that a subset of those with serious mental disorders, especially those with the dual diagnosis of mental disorder and substance use disorder, are indeed more likely than mentally healthy persons to become violent. Other studies have focused on the Cluster BV personality disorders, including Antisocial Personality Disorder, Borderline Personality Disorder, Narcissistic Personality Disorder, and, to a much lesser degree, Histrionic Personality Disorder. According to some estimates, the majority of individuals incarcerated for homicide fulfill criteria for one of these disorders.

A major problem with all three types of research is selection bias. Persons who are institutionalized, either in a psychiatric hospital or in a prison, have by definition displayed problematic behavior. Less biased data may come from large-scale studies of unselected populations. This fourth type of research—the longitudinal cohort study—is represented by the

report by Jari Tiihonen et al. regarding an unselected birth cohort of 12,058 Finns born in 1966. After 26 years, they found 503 male offenders, of whom 23 percent had psychiatric diagnoses (insufficient numbers of female offenders prevented further analysis of this subgroup). Compared with mentally healthy reference group, the odds ratio for being an offender was 3.1 among males with schizophrenia, 6.3 among those with paranoia and other psychoses, 5.1 among those with mood disorders with psychotic features, and an extraordinary 42.2 among those with "organic mental disorders." Given that substance-induced alterations in mental status were included in the organic category, the very high rate of criminality in this subgroup may be primarily attributed to the Axis I substance-related disorders. Of the 165 violent male offenders, 7 percent were diagnosed as psychotic; the highest risk for violence was identified among subjects with alcohol-induced psychosis and schizophrenia associated with substance abuse. In perhaps the most cited community-based study, data on violence among 7,000 individuals were collected as part of the US Epidemiological Catchment Area study. The lifetime prevalence rate for serious violent behavior was 7 percent among those without mental illness versus 16 percent for those with mental illness. Again, substance abuse, alone or in combination with other mental disorders, was the largest risk factor: Lifetime prevalence by diagnosis was 16 percent for schizophrenia or major affective disorder, 35 percent for substance abuse or dependence, and a remarkable 43.6 percent for those with both major mental disorders and substance abuse. Moreover, in a study of 90 hospitalized patients, 83.5 percent of the delusionally violent subjects had histories of substance abuse. Although it is clear that the combination of mental disorder and substance abuse is a significant risk factor for violence, it remains uncertain to what degree actual intoxication at the time of the offense or simply the diagnosis of substance abuse accounts for most of this association.

Why Does Major Mental Illness Increase the Risk for Violence?

Setting aside direct psychopharmacological effects in substance abusers, it is important to consider whether and to what degree the motives or drivers of violence differ between seriously mentally ill persons and others. For example, how much of the violence perpetrated by patients with schizophrenia is conventionally motivated versus motivated by a departure from consensual reality? A popular but unproved theory claims that delusions account for the high rate of aggression among those with schizophrenia. However, the findings from retrospective and prospective studies are not consistent.

Several retrospective studies reported strong associations between delusions that people seek to harm the patient ("threat") or that outside forces control them ("control override"). At least two such studies found

that delusion content seemed to influence the violent act. Other retrospective studies supported the concern that some patients will follow command hallucinations to harm others. In contrast, the large, prospective MacArthur study found virtually no association between either the presence or the content of delusions and subsequent commission of violence. It seems likely that persons with serious mental illness commit violence with conventional motives, psychosis-driven motives, or both. The lack of clarity on this point may explain the inconsistency of results from treatment trials. For instance, atypical antipsychotic agents may be used to treat a population of persons with schizophrenia in which only a subset aggresses due to psychosis. In some cases, medication benefit may result from damping psychosis, whereas in other cases, benefits may result from antidepressant effects or nonspecific sedation. The net effects may be positive, but the real reason for treatment successes or failures may go undetected.

One possibility is that persons with schizophrenia (as well as some with mood disorders) suffer from impaired threat assessment. Even without delusions, hallucinations, or illusions, it is possible that sensory stimuli that have neutral affective connotations to others may be perceived more often as threats by those with schizophrenia. Threat detection is strongly linked to the functions of the amygdala, the dorsal anterior cingulate gyrus, and both the dorsolateral and VMPFC. These parts of the brain help interpret facial expressions and other socially important signals as indicators of possible danger. Evidence exists that persons with schizophrenia exhibit defective processing in these brain regions. Moreover, functional neuroimaging studies have shown that patients with schizophrenia display increased amygdalar activation to ambiguous stimuli. These results are consistent with the hypothesis that, although outright delusions may not account for a large proportion of the variance, at least one factor in the observed high rate of violence among patients with psychosis may be their impaired frontolimbic perceptual processing of social stimuli.

Psychopathy and Violence

Another question is the relationship between psychopathy and human aggression. The construct of psychopathy is under debate, as is the relationship between four proposed entities: (1) psychopathy, (2) sociopathy, (3) DSM-IV-TR Antisocial Personality Disorder, and (4) ICD-10 Dissocial Personality Disorder (DPD). Psychopathy should be distinguished from sociopathy, a term usually used to describe individuals, such as gang members, expressing a group value favoring criminal behavior. Although a person may fulfill the definition for both of these two terms, many psychopaths are not group members in tune with their community's values. Psychopathy is characterized by a combination of

interpersonal, affective, and behavioral traits, including arrogance, callousness, superficiality, and manipulativeness in people who are emotionally irritable but lacking in empathy and who routinely engage in irresponsible, impulsive behavior regardless of the effect on others. APD was intended to capture much of this constellation of traits but—in the spirit of nontheoretical diagnosis—focuses more on overt behaviors than on underlying traits. DPD comes closer to the traditional description of psychopathy, including callousness, inability to form meaningful social bonds, and a proneness to rationalize. Robert Hare discriminated two factors in psychopathy: (1) an emotionally detached interpersonal/affective core of selfish, callous, and remorseless behavior and (2) a social deviance core of impulsivity, irresponsibility, and antisocial behavior. Other authorities have proposed a slightly different three-factor model, including (1) arrogant and deceitful interpersonal style, (2) deficient affective experience, and (3) impulsive/irresponsible behavior. Adoption studies demonstrating some heritability for APD and findings of autonomic nervous system hyporeactivity also hint at an innate neurobiological difference. It remains to be shown which model best explains the nucleus of behaviors called psychopathy, or whether a unitary entity truly exists.

Be that as it may, many studies have demonstrated that approximately 1 percent of the general population meets criteria for the diagnosis of psychopathy and that a very strong relationship exists between measures of psychopathy (such as the Hare Psychopathy Checklist-Revised) and lifetime prevalence of violence. In forensic psychiatric hospitals, psychopathy scores are predictive of aggression and violence. In fact, arrests for violent crimes are 10 times as high among those with scores above the median in forensic populations. Fifteen percent of male and 7.5 percent of female North American prisoners are thought to be psychopathic. Unlike most criminals, psychopaths devote most of their adolescent and adult lives to persistent criminality (although half of criminal psychopaths seem to "burn out," committing many fewer crimes after the age of 40 years). Compared with other criminals, on release from prison, psychopaths are four times more likely to reoffend violently. Moreover, persons with schizophrenia who also exhibit elevated psychopathy scores are much more likely to be violent recidivists. Although only 10 to 15 percent of child molesters are psychopaths, 40 to 50 percent of rapists are psychopaths. This group tends to commit more violent and sadistic rapes and is notoriously resistant to any form of therapy. The combination of psychopathy and phallometric evidence of deviant sexual arousal is particularly strongly predictive of sexual violence recidivism.

One theoretical conundrum is the fact that although psychopaths are notoriously calculating, accounting for a much higher rate of cold-blooded instrumental homicides than others, they are also impulsive. It appears that, unlike purely impulsive individuals who may be overtaken by an

irresistible and irrational impulse to aggress, psychopaths may become excited opportunistically and will exercise a rapid risk–benefit calculation before aggressing. Thus, although functional neuroimaging demonstrates reduced frontotemporal lobar perfusion in those with psychopathic personalities, hypofrontality may not be the main problem. In one study of violent antisocial offenders, reductions in the size of the posterior hippocampus correlated with psychopathy. This result is consistent with reports that dorsal hippocampal lesions impair the acquisition of conditioned fear and suggests that, whether choosing an immediate response or planning for the future, psychopathic individuals are simply less deterred by potential adverse consequences.

Dissociation and Violence

As many as 73 percent of murderers claim amnesia for the homicide. Although it is important to be wary of duplicity, it is likely that some individuals commit acts of violence while incompletely conscious and unable to consolidate accurate memories. Evidence suggests that, under sufficiently stressful physiological or psychic circumstances, many people experience loss of full contact with their usual experience of self in the world. Time and space may become distorted. They may witness events but feel walled off from them. They may see themselves acting but feel little direct control. They may act but not remember. The question is, what actual pathophysiological process (or processes) accounts for such episodes?

One potential explanation for such an event is dissociation. Dissociation is a commonly reported symptom in populations of individuals incarcerated for interpersonal aggression. A review of 12 studies found rates of pathological dissociation from 6.3 to 76.9 percent in various offender populations; the two best-designed studies found dissociative disorders in 14.3 and 28.3 percent of adolescent offenders, respectively. Others studies have reported rates of past incarceration of 28.6 to 47 percent among men diagnosed with Dissociative Identity Disorder. In the latter study, 40 percent of the men and 20 percent of the women had been convicted of homicide. Apart from the nominal DSM-IV dissociative disorders (Dissociative Amnesia, Dissociative Fugue, Dissociative Identity Disorder, and Depersonalization Disorder), dissociative symptoms are associated with many other conditions, including schizophrenia, schizotypal personality disorder, PTSD, panic disorder, mood disorders, BPD, OCD, and substance abuse disorders, as well as a host of medical conditions that alter mental function—ranging, for example, from epilepsy, medication toxicity, fever, and neuroinfection to uremic or hepatic encephalopathy.

Dissociation is a problematic phenomenon for neuropsychiatry because the biological underpinnings of dissociative symptoms are just beginning to

be investigated and may be heterogeneous. At least two categories of dissociation should be distinguished: Dissociative *traits,* as when an individual persistently or recurrently feels disconnected from his or her environment, and dissociative *states,* when a transient disruption occurs in the integration of consciousness with environmental perception and memory. *Trait* dissociation may be associated with aggression when a person is chronically irritable, withdrawn, detached, and exhibits a flat or inappropriate affect. In psychiatric populations, elevated trait scores on the Dissociative Experiences Scale (DES) correlate both with self- and other-directed aggression, including assaultive behaviors. In a study of 241 persons with Dissociative Identity disorder, 16 percent self-reported causing their children serious bodily injuries. Individuals with this trait may have little of the empathy that helps restrain interpersonal violence. For example, one person committed multiple homicides while suffering from acute medical illness combined with an exacerbation of intractable epilepsy during which he reported perceiving no real difference between objects and people, "walls and flesh." Patients with BPD may be chronically dissociative and may self-mutilate or attack others, perhaps due to this brain state. Survivors of severe traumas such as childhood abuse may chronically be detached from the normal experience of self and may also exhibit aggression, although it is difficult to determine to what degree such aggression is explicitly dissociative or simply social learning of violence.

State, or episodic, dissociation is also associated with aggressive behavior. Again, episodes of dissociation, typically of depersonalization, occur in many mental disorders and medical conditions. Every human has probably experienced some moments of dissociation—for instance, the phenomenon of *déjà vu.* There seem to be at least three types of dissociative state violence. First, it is thought that a subset of violent attacks committed by persons with PTSD are dissociative flashback-driven events. Individuals have reported reexperiencing sexual abuse or combat and misinterpreting the identity of neutral persons as past attackers. Second, transient depersonalization or derealization may occur during or immediately after a person without PTSD commits an act of violence. One theory is that perpetrators create their own trauma and then truly cannot recall due to self-induced dissociative response. This is consistent with the finding that many officers in police-involved shootings report poor recall and dissociative symptoms. Third, in some cases, an episode of dissociation may precipitate violent behavior in person with an otherwise inapparent organic brain dysfunction—for instance, due to mild TBI, neurotoxicity, or undiagnosed epilepsy—when a stressor lowers the threshold for the expression of abnormal brain physiology.

Dissociation has been studied from the point of view of psychiatry and neurology. The psychiatric viewpoint emphasizes psychic defense against unbearable traumatic memories. From the neurological viewpoint,

dissociation is thought to be a brain state in which the stream of consciousness in the DLPFC is partially disrupted, albeit by an unknown mechanism. Dissociative experiences are very common among patients with epilepsy, especially those with complex partial seizures originating in the mesial temporal lobe. Dissociative auras among such patients are associated with anxiety. This observation has led to the speculation that stress-related abnormal brain electrical activity disrupting prefrontal–limbic integration of consciousness (but short of producing a clinical seizure) often underlies dissociation. Functional MRI studies have reported that patients with PTSD in a dissociative state display more traumatic-script–driven activation of prefrontal and limbic structures. Although abnormal brain frontolimbic electrical activity similar to a seizure is a reasonable candidate for the usual mechanism of dissociative experiences, it remains to be determined what precise neurobiological phenomenon accounts for these events.

Aggression and Specific Disorders

The conclusions, at this point, include the fact that schizophrenia, mood disorders, Cluster B personality disorders, dissociative episodes, and organic brain disorders are risk factors for violence. Psychosis probably elevates the risk, independent of diagnosis. In addition, comorbid substance abuse greatly increases the risk. Yet, substance abuse is a major mental disorder in its own right. What follows is a summary of the relationship between several specific neuropsychiatric disorders and nonsanctioned aggression.

Substance Abuse or Medication Effects and Aggression. Most of the research on the association between substance abuse and violence has focused in the role of alcohol intoxication. Fewer data are available on other notorious drugs of abuse, including stimulants, anabolic steroids, and phencyclidine. It is important to note that, despite the focus of attention on these agents, substance-related violence has been recorded in association with virtually all categories of drugs of abuse, including hallucinogens and sedative–hypnotics such as heroin. The main problem with interpreting much of the literature is that most studies either (1) examine past violent events and measure the rate of intoxication, or (2) examine substance use and violent events, comparing rates of intoxication during violent and nonviolent events. The first approach cannot determine whether intoxication played a role because there is no comparison group. The second approach is weak because knowing, for example, that 50 percent of violent offenders were drinking whereas only 25 percent of property offenders were drinking does not reveal whether drinking caused an escalation from property crime to interpersonal aggression or whether, instead, those predisposed to violence tend to be drinkers. This raises the

possibility that aggression and proneness to substance abuse are both related to a third innate neurobiological or temperamental variable. In this case, it might be incorrect to regard substance abuse as an independent (causal) variable and violence as a dependent variable.

ALCOHOL. Several lines of evidence suggest that alcohol intoxication lowers the threshold for aggression. Meta-analyses of experiment studies with college students show an increase of approximately one standard deviation in aggressiveness, as measured by willingness to shock another person, when students have one or two drinks. Although such experiments persuasively argue for a pharmacological effect, they describe events so dissimilar from community circumstances that it is difficult to assume that these results predict the influence of alcohol on serious community violence. Another approach is to examine correlations between self-reported intoxication and serious offenses. In one report, the Bureau of Justice Statistics indicated that 41 percent of convicted homicide offenders in state prisons reported drinking alcohol at the time of the offense. Other reports document that offenders had consumed alcohol before 28 to 86 percent of homicides. An 11-nation international summary of 26 studies found that alcohol intoxication was associated with 61.5 percent of violent crimes. In a meta-analysis of correctional studies, individuals who chronically drank heavily were twice as likely as light drinkers or nondrinkers to engage in domestic violence and 70 percent more likely than low-level drinkers to engage in criminal aggression. Acutely intoxicated persons are 40 percent more likely than nonintoxicated persons to commit isolated acts of violence. Whereas acute intoxication is a risk factor for violence, perhaps the strongest correlation is found between the diagnosis of chronic alcohol abuse or dependence and intimate partner aggression. It is important to emphasize that different individuals seem to exhibit different degrees of precipitation of aggression both in experimental and observation research. An especially elevated risk may occur in individuals who fail to metabolize alcohol normally, in those with psychopathic traits, in novelty seekers, and in dominant individuals. Evidence also exists that early-onset abuse is more often linked to adult violence.

The mechanism of alcohol-related aggression is unclear. In rodent studies, alcohol potentiates both offensive and defensive attacks and reduces the avoidance of dangerous situations. The argument is that approach occurs when avoidance would be prudent. There exists an inverted U-shaped curve such that the peak of aggression occurs at intermediate doses, whereas high doses sedate. (This finding may not apply well to humans who have access to weapons; these objects mitigate the negative effect of severe intoxication on fighting ability. Indeed, one study that compared the consequences of conflicts between men with and

without concomitant intoxication found that the higher rate of lethality could be explained by the intoxicated men being more likely to use guns.) Some authorities argue for an alcohol expectancy phenomenon, such that the popular image of the disinhibited drunk drives misbehavior as people feel themselves becoming intoxicated. However, direct pharmacological effects are perhaps more important in explaining the link between alcohol and violence. Alcohol reportedly briefly elevates and then depletes brain serotonin—a physiological condition associated with a lowered threshold for impulsive aggression. One study also showed that habitually violent alcoholics exhibited elevated striatal DA transporter densities compared with control or nonviolent alcoholics, suggesting altered limbic striatal affective regulation. Given that DA transmission seems to regulate both risk taking and reward, it is possible that intoxication makes risky behavior pleasurably exciting, especially among people with sensation-seeking temperaments. Some evidence supports the hypothesis that alcohol depletes brain glucose, which in turn deprives the brain of a critical energy resource for self-control. However, despite the fact that alcohol has probably been associated with aggression for several millennia, the neurobiological basis of this association—and for the marked interindividual variability in this phenomenon—remains to be determined.

COCAINE. During the 1980s in particular, great concern was expressed about a crack cocaine epidemic associated with high rates of violent behavior. In one 1989 study, among 114 Los Angeles County coroner's cases with cocaine-positive toxicology, 61 percent died violently, more than 68 percent by stabbings or shootings. However, the lifestyles of users, such as possible affiliation with competitive gangs or homelessness, also created "economic-compulsive" or "systemic" risks for violence, making it difficult to determine how many cocaine-related homicides were truly caused by the psychopharmacological effects of cocaine on the brain. Evidence also exists that alcohol increases the likelihood that cocaine-intoxicated individuals will become violent, and it is difficult to determine how much of the observed relationship between cocaine and aggression can be explained by double intoxication. It is known that cocaine intoxication produces a state of excitement, irritability, euphoria, high energy, and sometimes paranoid psychosis. Some evidence suggests that inhalation of free-base cocaine is more likely to be associated with anger than is snorting of powder cocaine. Two types of pharmacologically induced violence are most common: (1) excited delirium or acute intoxication–related psychosis and (2) cocaine withdrawal rage. Excited delirium is relatively infrequent but striking. Patients exhibit combativeness, wild running, screaming, and fighting. Such outbursts are probably less often related to paranoid misinterpretations of stimuli and more often represent spontaneous frenzies. Emergency room experience confirms that individuals who arrive

unconscious due to cocaine intoxication virtually always become combative on awakening, which is consistent with studies showing high levels of self-reported hostility during cocaine withdrawal and even during periods of abstinence in chronic users.

Some evidence suggests that chronic cocaine abuse leads to elevated numbers of striatal DA receptors in the face of depressed DA terminals, a state that might be called neuronal yearning for binding. This condition may amount to a reward deficiency that causes irritability even during periods of abstinence. It has also been proposed that cocaine produces emotional dysregulation due to electrophysiological changes in the limbic system akin to the "kindling" seen in rodent models of epilepsy. Rather than attributing violent events to this isolated effect on the brain, it seems likely that many instances of cocaine-related violence combine a pharmacological effect with a violence-supporting developmental/social milieu and instrumental pressures to aggress. Again, insufficient evidence precludes firm conclusions regarding the degree to which cocaine-associated violence is a neurobiological phenomenon.

METHAMPHETAMINE. During the 1990s and thereafter, methamphetamine became the stimulant of choice for many individuals. Methamphetamine-associated aggression has been widely reported. In several studies, from one-third to two-thirds of interviewed patients reported violence as an outcome of their use of methamphetamine. In a study of 627 California state prison parolees, a significant positive association was found between self-reported violence in the last 30 days and methamphetamine abuse. The finding that chronic abusers of methamphetamine exhibit both deficient decision making and depletion of serotonin in the orbital prefrontal cortex suggests a possible pharmacoanatomical basis for the poor judgment and disinhibition that these patients exercise, sometime associated with impulsive aggression.

ANABOLIC STEROIDS. Steroid abuse by athletes to bolter strength has become a major focus of public attention. Congressional hearings have been held, indicating both the penetration of awareness about this medical problem and its effect on the popular expectation of fairness in sports. "Roid rage" has become a term of art, expressing the widespread opinion that these agents cause aggression. Substantial evidence suggests that anabolic steroids do indeed lower the threshold for violence, including homicide. The steroids most commonly abused include the oral agents oxymetholone (Anadrol), oxandrolone (Oxandrin), methandrostenolone (Dianabol), and stanozolol (Winstrol), and the injectable agents nandrolone decanoate (Deca-Durabolin), nandrolone phenpropionate (Durabolin), testosterone cypionate (Depo-Testosterone), and boldenone undecylenate (Equipoise). It is unclear which types of steroids, forms of

administration, doses, or predisposing psychobiological factors alter the risk of aggression. It is also unclear whether steroid-associated violence is primarily related to paranoid psychosis, to dissociation, or to nonpsychotic explosive disorders. However, authorities conclude that steroids cause aggression in previously less aggressive individuals that cannot be attributed to coabuse of other drugs. In one study, 23 users reported significantly more fights, verbal aggression, and violence toward intimate partners when using steroids than when off these drugs. In a larger study, the major finding was that steroid-using strength athletes exhibited significantly more mania, hypomania, or major depression than nonusers, suggesting the possibility that mood disorder rather than psychosis accompanies the irritability that sometimes leads to violence.

PHENCYCLIDINE. The dissociative anesthetic phencyclidine (PCP) is notoriously associated with bizarre aggressive behavior and even increased physical strength, purportedly making the user especially dangerous. However, empirical data demonstrating these effects—and demonstrating that they are not due to temperament, milieu, comorbidities, or other confounding variables—are quite limited. The pharmacology of PCP is complex. Although it is best known as a noncompetitive *N*-methyl-d-aspartate (NMDA) antagonist (perhaps acting as a channel blocker in the NMDA-receptor complex) and its odd effects on mentation are usually attributed to this receptor interaction, it also has nicotinic, muscarinic, noradrenergic, and serotonergic effects. Despite its reputation, large-scale studies of violent offenders have not reported any increased prevalence of PCP use compared with other drugs. Several studies are cited as evidence of the agent's dangerousness: An epidemiological study reported that, of 104 violent deaths in St. Louis associated with PCP use between 1981 and 1986, 80 were homicides. However, this finding hardly demonstrates a causal relationship between PCP intoxication and aggression. Another study reported elevated hostility scores among men with histories of PCP use. A consensus has emerged that only a subgroup of PCP users exhibit aggression, but aggression–predictive traits of these users have not been identified, raising the question of whether the aggressive subgroup is violent for some other reason.

SELECTIVE SEROTONIN REUPTAKE INHIBITORS. The discovery that a small number of minors treated with selective serotonin reuptake inhibitors (SSRIs) developed suicidality led to a black box warning mandated by the U.S. Food and Drug Administration (FDA) in 2004. It is hypothesized that incomplete development of the central serotonergic pathways explains a significant difference between youth and adults in the effect of these medications on behavior. Equally concerning, many well-documented case reports support the conclusion that a subgroup of adults treated with SSRIs

become irritable, hostile, or homicidal. This effect apparently occurs most often in the first month of treatment. In one study, 60 of 9,219 adults (0.65 percent) treated with paroxetine (Paxil) reported hostile acts, compared with 20 of 6,455 (0.31 percent) on placebo. Thus, although the likelihood of this adverse effect is low, SSRIs perhaps double the frequency of hostile acts. Anecdotal evidence suggests that "bad responders" may display atypical anxiety or irritability in the first hours or days of treatment. Careful monitoring of response and prompt follow-up during the initial stage of treatment are warranted. Potential psychobiological predictors of this idiosyncratic adverse response are being investigated.

Borderline Personality Disorder and Aggression

BPD, along with APD, is one of a group of apparently innate temperamental conditions referred to as Cluster B personality disorders. Unlike APD, BPD seems to be more common among women than men. Self-directed aggression occurs in up to 66 percent of BPD patients, and other-directed aggression occurs in about 30 percent. Although a large body of research points to genetic, neuroimaging, and neurochemical abnormalities in these patients, fewer studies have focused on specific biological correlates of aggression in BPD. It is conceivable that different mechanisms underlie impulsivity and aggression in this condition. For example, in one PET study, impulsivity correlated with superior and middle frontal cortical metabolism, whereas hostility correlated with temporal lobe metabolism.

Many studies have supported the hypothesis that BPD is associated with diminished volume of medial temporal lobe structures, including the hippocampus and amygdala. Some evidence ties this change to early trauma. It is not clear, however, whether these volumetric changes should be regarded as congenital failure to develop versus acquired atrophy. In either case, dysfunction of this limbic tissue would be predicted to alter social reactivity and emotional regulation. Many studies also have pointed to disordered serotonergic function as measured by platelet-binding assays, CSF measures, and neurochemical challenges. For instance, BPD patients exhibit reduced metabolism in response to fenfluramine (a 5-HT releaser) in the ventromedial prefrontal lobe. A single photon emission computed tomographic (SPECT) measure of 5-HT transporter availability indicated significant differences in brainstem and hypothalamus. Of special interest, physically aggressive men with BPD were compared with men who were only verbally aggressive. Those with physical aggression had reduced frontocingulate metabolism. fMRI studies suggest that BPD patients exhibit dysfunctional amygdalar and cingulate responses to emotional faces, consistent with the theory that this condition is associated with abnormal emotional processing that could lead to rejection sensitivity and self-directed aggression or to misinterpretation of threat and other-

directed aggression. A report that fluoxetine normalized prefrontal cortical function in aggressive subjects with BPD hints at a possible avenue of treatment.

PTSD and Aggression

Many studies have reported elevated indices of self-reported aggression among veterans with PTSD, and many tragic cases have suggested that returning veterans have committed violent acts driven by the internal demons of PTSD. In one study of 17 Vietnam veterans seeking help for PTSD, subjects reported an average of 22 aggressive acts over the previous year, compared with 0.2 for non-PTSD veterans. Correlates of aggression in veterans with PTSD include severity of PTSD, lower socioeconomic group, hyperarousal symptoms, and comorbid alcohol use disorder. Nonveterans with PTSD also exhibit elevated rates of aggression. It is worth considering that PTSD may not be a unitary entity. Given that patients with this diagnosis vary greatly in the occurrence of hyperarousal, avoidance, depression, and emotional numbing, different biological stress reactions are perhaps being lumped together under an unsupportably broad umbrella. It has been suggested that elevated nervous system reactivity, as seen in one subset of those diagnosed with PTSD, may mediate the increased risk of aggression. One study reported increased cerebral blood flow in the region of the NAc in patients with PTSD and aggression. However, the specific neurobiological basis of the association between this mental disorder and violence remains a topic of speculation.

Intermittent Explosive Disorder

Intermittent Explosive Disorder (IED) is the only DSM-IV disorder that explicitly requires aggression for diagnosis. Some authorities consider it to be commonplace and underdiagnosed. Others express concern that, given the flaws in the DSM-IV criteria, any person who periodically exhibits aggression might be diagnosed with IED, whereas perhaps a smaller number truly have a distinct disorder distinguishable both from those prone to premeditated violence and those who merely have irritable temperament or poor self-control. Because of this lack of consensus, in the 2001 to 2003 National Comorbidity Study, IED was defined both broadly as three attacks in a lifetime and narrowly as three attacks in the last 12 months. According to that study, lifetime prevalence of broadly defined DSM-IV IED was 7.3 percent; the prevalence of narrowly defined IED was 5.4 percent. Both figures were much higher than is generally assumed. The average age of onset was 14 years. In the narrowly defined cases, medical attention was required to treat someone (victim or perpetrator) 233 times per 100 cases. Regardless of whether IED is a discrete, unitary entity or a behavioral phenomenon with multiple etiologies, persons meeting the

operational criteria for this behavior clearly cause a great deal of social distress and injury and some homicide.

The current diagnostic criteria are problematic. As written, they do not require that the aggression be explosive or impulsive, which permits inclusion of those with premeditated aggression. The current criteria also set no standard for frequency or multiplicity of aggressive acts and do not specify that the condition cause distress or impairment. Emil Coccaro and coworkers proposed better-circumscribed research criteria, according to which IED is described as follows:

A. Recurrent incidents of aggression manifested as either of the following:
   1. Verbal or physical aggression toward other people, animals, or property occurring twice weekly on average of 1 month; or
   2. Three episodes involving physical assault against other people or destruction of property over a 1-year period.
B. The degree of aggressiveness expressed is grossly out of proportion to the provocation or any precipitating psychosocial stressors.
C. The aggressive behavior is generally not premeditated (e.g., is impulsive) and is not committed to achieve some tangible objective (e.g., power, money, intimidation, etc.).
D. The aggressive behavior causes either marked distress in the individual or impairment in occupational or interpersonal functioning.
E. The aggressive behavior is not better accounted for by another mental disorder.

The pathophysiology of IED is unknown. Again there is controversy, with some authorities contending that IED represents epilepsy (or a subclinical seizure equivalent), and others proposing that IED represents subsyndromal bipolar disorder. Some evidence supports both hypotheses. Because impulsive affective interictal aggression is an infrequent but recognized syndrome in persons with complex partial seizures originating from the temporal lobe, it is plausible that some cases meeting IED criteria are indeed due to undiagnosed epilepsy. A *limbic–psychotic trigger reaction* has also been posited to account for some explosive aggression either in the course of subclinical seizures or between seizures in patients with abnormal brain electrical activity. On the contrary, given that some individuals with IED report distinctly manic symptoms, an atypical form of irritable bipolar disorder may account for other cases. Subnormal serotonergic function is associated with both impulsive aggression and depression, so some individuals with IED may be primarily affected by this problem. In one study, 70 percent of children with IED also met criteria for ADHD, suggesting a hyperarousal-related version of this behavior. Given that patients with IED report derealization, this syndrome may overlap with dissociative disorders. Furthermore, given that impulsivity in patients with BPD is similar to that seen in some cases of IED, these entities may

also overlap. IED symptoms have also been reported in cases of tumors impinging on the hypothalamus, leading to the speculation that such lesions interrupt limbic restraints on stereotypical hypothalamic attack behavior (although direct hormonal effects, e.g., related to AVP, cannot be excluded). Given the clinical heterogeneity of persons with intermittent explosive aggression, some theorize that IED should, at the least, be divided into hyperarousal- and hypoarousal-related types.

Several biopsychological abnormalities have been detected in different subject groups meeting the various criteria for IED. An angry, hostile baseline temperament may be a prominent distinguishing feature of one subgroup. On the Iowa Gambling Task, a standard measure of impulsive decision making, IED subjects failed to learn and persisted in making disadvantageous choices. Scalp electroencephalographic (EEG) abnormalities, more often in the left temporal lobe, have sometimes been reported. Abnormal visual-evoked potentials have been detected in a subset of children and adults with this diagnosis; the appearance of high-amplitude P100 waveforms in a subgroup is a nonspecific finding consistent with multiple central disorders but again suggests diagnostic heterogeneity. Platelet 5-HT content was lower in one study of personality-disordered subjects with IED. One study demonstrated blunted cortisol response to the 5-HT_{1A}–receptor agonist ipsapirone. This finding suggests that, at least in a subgroup, impulsive aggression relates to altered postsynaptic serotonergic receptor function. One study demonstrated a combination of amygdalar hyperactivation and orbitofrontal depression during exposure to angry faces in IED subjects, supporting the contention that these patients misperceive social signals and respond both with exaggerated emotions and impaired inhibition.

Understanding of IED is evolving slowly, and, until clinically and biological homogeneous subtypes are delineated, this diagnosis should not be taken to refer to a specific condition but to refer to a phenomenologically problematic behavioral profile due to a variety of pathophysiological conditions.

Traumatic Brain Injury and Aggression

TBI is well known to be associated with new-onset aggression. The nonspecific agitation that is a very common stage of subacute recovery should not be confused with, and perhaps bears no relation to, the less frequent occurrence of clinically significant persistent posttraumatic aggression. Because of variations in definition, the precise incidence of such persistent aggression is not clear. Some reports claim that irritability or aggression occurs in as many as 50 percent of patients, but preliminary research at the Rancho Los Amigos National Rehabilitation Center (Downey, CA) suggested a 1-year prevalence for clinically problematic

posttraumatic aggression of ~10 to 12 percent. This condition represents a major additional stress not only on the patient in his or her progress toward social reintegration but also on the caregiver: For example, wives of men with posttraumatic aggression are severely burdened and at risk for injury.

Insufficient research has explored the clinical or neurobiological correlates of post-TBI aggression. Frontal lobe injury apparently increases the incidence. Some evidence suggests that right frontal injury is especially likely to produce this complication. Other evidence suggests that posttraumatic epilepsy is a risk factor for aggression.

Cerebrovascular Disease and Aggression

Cerebrovascular disease affects the brain in many ways and degrees, from the effect of chronic hypertension on behavior, to the effect of progressive stenosis of major arteries short of stroke, to acute ischemic or hemorrhagic infarction, to periventricular white matter disease, to multiple lacunar states, to multi-infarct dementia. Each manifestation of cerebrovascular disease is associated with different behavioral features. This discussion is confined to aggression after completed cerebral vascular accident (CVA).

Episodes of aggression have been observed during the hours or days after an acute CVA. One study reported anger in 35 percent and extreme anger in 13 percent of acute stroke patients. Hemorrhagic stroke seems to produce this behavior more often than ischemic stroke. Such transient behavior can distress the family, put the patients and health care providers at risk, and compromise rehabilitation. However, the greater concern for caregivers is persistent poststroke hostility and aggression. Up to 10 percent of patients may develop this complication. Poststroke aggression has been correlated with proximity of the lesion to the frontal pole and to measures of depression. Several case reports of aggression after posterior circulation CVA have also been published; although temporal lobe or thalamic involvement may mediate this phenomenon, in one case the lesion was confined to the occipital cortex. Emotional incontinence occurs in some patients, with pathological laughter and crying consistent with VMPFC injury. However, impulsive aggression occurs only in a minority of these cases. The largest study of this problem included 145 stroke patients with inability to control anger or aggression. In these patients, anger was typically triggered by trivial or no provocation. Contrary to the theory that depression due to disability accounts for aggression, this study reported no such association. Poststroke aggression was instead correlated with motor dysfunction, dysarthria, and emotional incontinence. Frontolenticulocapsular and basis pontis lesions were the most commonly associated sites, but hemispheric laterality is not necessarily predictive.

Neurodegenerative Disease and Aggression

Agitation and aggression are among the most distressing noncognitive behavior disturbances seen in neurodegenerative disease. These problems have been best studied in Alzheimer disease (AD), but they also occur in frontotemporal lobar degeneration (FTLD), dementia with Lewy bodies (DLB), and Huntington disease (HD) and less frequently in Parkinson disease (PD). Many behaviors are labeled as aggressive in patients with dementia, from resisting bathing to committing murder. This makes it more difficult to interpret the results of the many studies that fail fully to describe the type of aggression being investigated or treated. Overt aggression scales are very helpful, but many studies use agitation scales that may obscure subtypes of aggression by conflating behaviors as diverse as wandering and physical assaultiveness. Nonetheless, a profusion of high-quality research on aggression in AD has provided something of a window into the biopsychology of human aggression.

Overt aggression is reported in 18 to 65 percent of patients with AD. This symptom is associated with a major increase in caregiver distress and early institutionalization. Aggressive resistance tends to continue until it is arrested by motor limitations or death.

Behavioral correlates of aggression in AD include (1) dementia severity (this behavior occurs most often in the middle stages), (2) measures of frontal lobe cognitive function, (3) premorbid personality disorder, and (4) concurrent psychosis. *Biological* factors reported to correlate with aggression in AD include (1) left temporal lobe atrophy, (2) right and left medial temporal lobe hypoperfusion, (3) postmortem measures of frontotemporal cholinergic function, (4) loss of noradrenergic cells in the locus coeruleus, (5) reduced 5-HT_{1A}–receptor binding in temporal cortex, (6) greater prolactin response to fenfluramine challenge (an index of central serotonergic function), (7) blunted growth hormone response to clonidine challenge (an index of central α-2 adrenergic function), (8) apolipoprotein E-e4 allele, (9) the T-427C variant of the apolipoprotein E promoter, (10) the longer variant of the 5-HT transporter promoter region, especially in homozygotes, (11) a DA DRD1 receptor polymorphism, and (12) the A218C polymorphism of the gene for tryptophan hydroxylase. One way to summarize these various findings is to conclude that aggression in AD is biologically similar to aggression in patients without dementia: Either prefrontal or serotonergic dysfunction may lead to impulsiveness and social discomportment. It remains to be seen whether the gene variants associated with aggression in AD are specific to that disorder.

Epilepsy and Aggression

British psychiatrist Henry Maudsley wrote in 1873 that “whenever a murder has been committed suddenly, without premeditation, without malice, without motive, openly and in a way quite different from the way in

which murders are commonly done, we ought to look carefully for evidence of previous epilepsy." Thus, the relationship between epilepsy and violence has garnered considerable interest for more than a century.

There are at least five ways in which epilepsy might relate to aggression:

1. *Peri-ictal aggression,* a term referring to aggression observed around the time of a seizure but in practice typically distinguished from postictal aggression and often used to refer to prodromal aggression seen in the hours or days prior to a seizure.
2. *Ictal aggression,* referring to automatic and unconscious behaviors during an epileptic seizure.
3. *Postictal aggression,* sometimes called postictal rage.
4. The *limbic psychotic trigger reaction,* an entity posited to occur in a small proportion of persons with limbic foci in which a seizure equivalent causes a psychotic state that can lead to violence.
5. *Interictal aggression,* referring to problematic irritability or violence occurring between seizures.

Most of the attention paid to aggression and epilepsy concerns complex partial seizures (CPS). Prior to the mid-1980s, this neurological disorder was often referred to as *temporal lobe epilepsy*. Since that time, recognizing that such seizures may have foci outside of the temporal lobes (most often on the frontal cortex or insula), the terminology of CPS has been used. It is important to state clearly that persons who have either lost consciousness or exhibit significantly altered consciousness very rarely commit directed violence *during* a seizure. In CPS, whether emanating from temporal or frontal lobe foci, typical behaviors are chewing, reaching, picking at clothing, pacing purposelessly, or flailing. When restrained, a person having a seizure may resist and harm the restrainer. Although somewhat complex aggressive, unequivocally ictal behaviors have occasionally been documented, such as a patient who tore a sink from a wall, another who repeatedly kicked a neighbor, or another with "wrestling seizures" who threw others to the ground during video/EEG-monitored seizure activity, the targets of the ictal aggression are virtually always targets of opportunity such as the nearest object or person. Ictal aggression, in the largest study, averaged 29 seconds and occurred in the earliest phase of a seizure. Planning and sequencing of complex motor actions involving multiple steps (such as seeking out a weapon and then a victim) are not consistent with ictal action.

Still, some patients experience prolonged episodes, perhaps lasting days, during which their seizures are so frequent that they never fully regain consciousness. This condition, called *complex partial status epilepticus,* is uncommon. During such episodes, the patient may seem to be awake and semiattentive to his or her environment and may perform more complex actions than those seen during the actual ictus. At least two

reports of confused aggressive behavior during EEG-confirmed partial complex status epilepticus have been published. It is within the realm of possibility, therefore, that a person experiencing status epilepticus could act out in violent ways, although such a patient would surely also exhibit some detectable disorientation.

Brain electrical activity is abnormal in patients with CPS for a period both before and after an ictus. An aura of fear is very common. The resulting hypervigilance and sense of impending danger may cause a patient to misinterpret environmental stimuli in a paranoid way and prepare for defensive aggression. Such prodromal aggression has been reported during the 72 hours prior to a seizure and has included irritable verbal or physical aggression toward selected targets. However, it is much more common for CPS patients to exhibit aggression in the postictal than in the preictal period. During a typical postictal period lasting from minutes to several hours, most patients are confused, withdrawn, or lethargic and often just nap. However, some patients recurrently become hostile, angry, and aggressive. Their actions may be much more complex than those observed during a seizure. Somewhat complex actions can perhaps appear to be planned and sequenced during this postictal confusional period, although, typically, a patient behaves in a simple, stereotypical way and has an incomplete memory for these events. Longer subacute postictal violence has also been described, shading into the descriptions of limbic psychotic trigger reaction or interictal aggression. One patient, for instance, about 6 to 24 hours after a flurry of seizures exhibited a psychotic state with hyperreligiosity, paranoid delusions, and violence such as attempting to stab his sister with a knife and severely assaulting an uncle. This patient had no recall of the events and was remorseful when told of them.

The *ictal psychotic trigger reaction* has been proposed to account for episodes in which an otherwise nonviolent but perhaps socially isolated schizotypal person with a flat affect experiences episodes lasting approximately 20 minutes involving delusions, hallucinations, and attacks on victims of opportunity. TBI and abnormal EEGs (but not necessarily epileptiform discharges) are common. Transient abnormal amygdalar hyperactivation is hypothesized to account for such episodes.

The most common type of epilepsy-related aggression is *interictal*. The term technically refers to periods between seizures. However, given that depth electrode recordings have demonstrated that paroxysms of abnormal brain electrical activity often occur without any obvious clinical seizure, *interictal* is better understood as "between *clinically obvious* seizures." Some patients may spend much of their lives attempting to function in society while paroxysmal depolarization shifts persistently disrupt frontolimbic brain function. Depth electrode studies also show that these abnormalities wax and wane. This may explain the observation that some

patients with interictal aggression are chronically irritable and considered to have hostile personalities, whereas others exhibit sudden unexplained bursts of aggression that might be diagnosed as IED. The literature on interictal aggression is fraught with inconsistencies, with as few as 4.8 percent and as many as 50 percent of patients reported to exhibit rages.

It is important to be cautious about concluding that a person known to have epilepsy and interictal aggression is aggressive *because* of the epilepsy. Studies have shown that aggressive individuals with epilepsy tend to come from violent families and lower socioeconomic groups, have lower IQs, and have higher rates of perinatal and other trauma. Because any of these factors may elevate rates of aggression in persons without epilepsy, it can be hard to say when epilepsy is the major causal factor. One way to determine whether epilepsy or another factor is likely to account for aggression is to look at sibling controls: If a man is the only one of four brothers who has epilepsy and the only one who is violent, it is plausible to conclude that his neurological disorder rather than his developmental environment explains his aggression.

Many Neuropsychiatric Disorders Increase the Incidence and Prevalence of Aggression

The conclusion is inescapable. Neuropsychiatric illness is a major risk factor for violent behavior. Perhaps surprisingly, affective disorders and schizophrenia both elevate the risk to about the same degree. Unsurprisingly, psychopathy, personality disorder, and substance abuse account for a large amount of the association between violence and mental illness. Although roughly 85 percent of violence is committed by offenders who are not mentally ill, mental health authorities must be cognizant that their patients may harm others. The physician is not immune; the annual rate of nonfatal job-related violence is 12.6 per 1,000 in all occupations but 68.2 per 1,000 among psychiatrists. Psychiatrists also need to know that certain diagnoses tend to be associated with higher risks of aggression and that legal requirements in many jurisdictions oblige the treating clinician to break with the conventions of confidentiality and report threats in the interest of public safety.

As will be discussed in the subsection on treatment, it is challenging to choose and administer effective therapies to aggressive individuals. A larger-scale challenge is how to deliver care to the millions of untreated mentally disordered individuals in the community who are at risk to commit violence and how to deliver meaningful psychiatric care to the hundreds of thousands of desperately underserved persons crowding jails and prisons. Cost-effective policy recommendations lie beyond the purview of this section.

TREATMENT OF AGGRESSION

The very concept of treatment for aggression gives pause. First, the overwhelming majority of aggressive behavior is biopsychologically normal or at least not due to a mental disorder. Two-year-old boys' frequent hitting, parental corporal punishment, fighting for sport, aggressive middle school behavior ranging from rumor mongering to bullying (and arguably gang violence), the dominance-seeking male's intimidation of others, manhandling of employees, servants, and intimate partners, riots, pogroms, and warfare are commonplace and typical of human societies. As much as these behaviors are indisputable causes of tremendous human distress, physicians must consider whether they are being asked to treat human nature—or part of the evolved adaptive spectrum of innate temperament—rather than a discrete pathological disorder. Indeed, a large proportion of aggressive behavior is socially sanctioned.

Second, only a small subset of those committing aggression regard their behavior as problematic and present for voluntary treatment. Clinicians often divide their energy between treatment and diplomacy involving the identified patient and a third party such as a parent, spouse, or judge. The chance of developing a therapeutic alliance is frequently depressed.

A third conundrum is the extremely wide variety of problematic aggression. Aggression varies tremendously in etiology, pathophysiology, and phenomenology, from the type seen in a hyperactive boy with ADHD to that seen in an adult serial murderer. Psychiatrists may be asked to address aggressive behavior in individuals who (1) have no psychiatric diagnosis (e.g., angry adolescents who, on assessment, are essentially normal), (2) are suspected to have such a diagnosis (e.g., persons referred in a medicolegal setting), (3) are known or discovered to have such a diagnosis (e.g., in- or outpatients with psychiatric disorders), or (4) are known or discovered to suffer from two or three psychiatric and medical diagnoses (e.g., a person suffering from schizophrenia, substance dependence, and cerebrovascular disease). Behaviors that are phenomenologically indistinguishable, such as recurrent domestic violence, may have one set of primary factors in one case (e.g., a man with alcohol-related disinhibition who witnessed a parade of father figures battering his mother) and another set in another case (e.g., a comfortably raised upper-middle-class man with a low-functioning MAOA polymorphism whose threshold to commit aggression is further lowered by the pain of frequent migraine headaches). In the text that follows, treatments are recommended on the basis of the primary diagnosis, with the recognition that current nosology tends to conflate neurobehaviorally different problems.

A fourth problem is that aggressive individuals are imbedded in social milieus. Aggression in a setting of commonplace family or community violence may indicate adaptation. The same behavior in a nurturing and safe environment may signal profound disorder. Furthermore, the possibility of effecting a beneficial change may also depend greatly on the

social environment. A youngster with ADHD living in a stable household with two supportive parents and a cooperative school system will probably respond to the same intervention with a much higher rate of success than a similar child living in an unstable setting in which physical aggression is often the first recourse in conflicts.

A fifth problem is financial. Almost no health care system funds the type of intensive, long-term, expert intervention necessary to reduce most forms of aggression meaningfully in a lasting way. It would seem that that societies have chosen to tolerate a certain amount of violence of the type that might be moderated by psychiatric treatment. Many preventable homicides and occasional massacres are inevitable.

The major dilemma, however, is the paucity of randomized, double-blind, placebo-controlled (RDBPC) research trials. Even when this gold standard of research design is followed, many studies lump together individuals with different diagnoses and report the efficacy of interventions without distinguishing differences in benefit that are virtually certain to exist among pathophysiologically unrelated conditions.

Much of the literature resorts to truisms: Get to know the patient; establish a therapeutic alliance; individualize the approach; treat the underlying neuropsychiatric disorders; determine and modify environmental triggers; reinforce positive behaviors. In this subsection, an attempt is made to offer a more evidence-based algorithm proposing specific approaches for several types of problematic aggression. Table 30–4 summarizes the most common forms of antiaggression treatment.

Table 30–4.
Antiaggression Treatments

Nonpharmacological interventions
Individual therapies
Classical psychoanalysis or other transference-focused psychotherapy
Insight-oriented psychotherapy
Cognitive-behavior therapy
Supportive psychotherapy
Behavior modification
Multisystemic therapy
Group therapies
Pharmacological interventions
Antipsychotics
Typical
Atypical
Lithium
Antiepileptic drugs
With mood stabilizing qualities
Without mood stabilizing qualities
Antidepressants
β-Blockers
α-2-Agonists

General Statement Regarding Nonpharmacological Interventions

Nonpharmacological modes of intervention were likely attempted for millennia, in families and tribes, before the rise of the concept of medicine. Systematic study of different talk methods, however, has barely begun. Although society looks to mental health professionals for guidance in this arena, one should be candid about the weakness of the empirical basis for prescribing talk-based intervention.

The most common interventions owe their theoretical genesis to psychoanalytic theory, behavioral modification, insight-oriented psychotherapy, cognitive-behavior theory (CBT), supportive psychotherapy, anger management, and group therapies with various theoretical orientations.

Psychoanalytic Therapy

Psychoanalytic therapy addresses aggression in terms of core drives and problematic development, especially unhealthy resolution of purported psychosexual developmental conflicts such as the *Oedipus complex* and the supposed defense of pathological narcissism. Treatment is aimed, in part, at uncovering deep-seated drivers of maladaptive behavior and helping patients to find value in behaviors that do not require acting out primitive responses to ego trauma. Most of the publications in this field are case reports claiming to have unearthed the root causes of the aggression and modified them through talk involving transference. No scientific evidence supports the use of classic psychoanalysis for the management of any type of aggression, although the absence of controlled trials does not in itself prove a lack of benefits. Indeed, one controlled trial supported a form of transference-focused psychotherapy to reduce aggression in BPD.

Behavioral Modification

Behavioral modification regards the aggressive individual as a black box. Neither the psychological nor the neurological mechanisms responsible for the aggression are considered or monitored. Instead, clinicians who use in this approach apply operant conditioning based on experimental results seen in animal models, using rewards (reinforcers) or punishments via *behavioral modification* to decrease the occurrence of undesirable behaviors. Some research suggests that this method is effective in aggressive young children, in those with MR, in schizophrenia, or in persons with TBI. Less evidence supports efficacy in other forms of aggression. Many studies fail to disaggregate important differences in modifiable pathophysiology, for instance, hyperarousal- versus hypoarousal-based aggression or aggression due to treatable mood

disorder. This failure may both obscure the benefit of this approach for subgroups and put some patients at risk for delayed definitive intervention.

Insight-Oriented Psychotherapy

Insight-oriented psychotherapy comprises a very broad range of talk therapies, all sharing the premise that awareness of the psychological basis of one's maladaptive patterns of interaction will ameliorate those problems. Insight aside, a therapeutic alliance with a compassionate professional probably has value for managing many aggressive individuals. In theory, this type of therapy might be especially efficacious for reducing aggression stemming from depression or anxiety; some empirical evidence supports this contention. Less evidence supports efficacy for aggression due to disorders of impulse control, personality disorders, or substance abuse. However, given that the literature reports very few controlled trials of patients exhibiting a homogeneous type of aggression with a homogeneous type of therapy (and because it is so challenging to conduct controlled trials in which nontherapeutic talk occurs in the placebo arm), it is premature to conclude which types of aggression are most amenable to this type of intervention.

Cognitive-Behavior Therapy

CBT has eclipsed most other talk therapies in current practice. This treatment, first promoted by Aaron T. Beck, is based on the theory that individuals learn to think in automatic, maladaptive ways, making irrational assumptions and holding irrational beliefs linked to painful emotions. Helping them through informed, supportive talk to cognitively reframe the dilemmas they face in more rational, accurate, and useful ways is postulated to improve their success in dealing with life's challenges. CBT is similar to Albert Ellis' *rational emotive behavior therapy* in attempting to link conscious cognitive control to unconscious affective processes. Because the drivers of human behavior are tenuously controlled by and significantly inaccessible to consciousness, it seems unlikely that patients can think themselves out of deeply ingrained (and often genetically influenced) behavior patterns or the profoundly aberrant biology of, for example, major depressive disorder. Evidence exists that the particular theoretical orientation of a psychotherapist is largely irrelevant, given that observers note that clinicians who claim to be using one form typically combine approaches from different theoretical schools. Still—by whatever mechanism—CBT sometimes catalyzes valuable change. Some empirical evidence supports the efficacy of CBT in the management of many types of aggression, including childhood school aggression, road rage, intimate partner violence, and aggression associated with major mental illness.

Supportive Psychotherapy

Supportive psychotherapy refers to regular meetings with an empathetic care provider. The author believes that this intervention—whether or not it occurs within the setting of another model of talk therapy; whether administered by a mental health profession, a religious leader, or another empathic authority figure; and whether or not it is associated with medical interventions—is an essential minimum component of any therapeutic strategy. Theoretical orientation aside, a therapeutic alliance is perhaps the sine qua non of help.

Anger Management

Anger management is an invention of the 1970s. Since that decade, innumerable programs have sprung up, claiming to provide effective—usually group-based—nonpharmacological interventions for many aggressive problems, from road rage to spousal battering. Courts and schools sometimes make completion of such programs obligatory as part of a disciplinary process. However, publicly salient apparent treatment failures such as boxer Mike Tyson and the Columbine shooters suggest suboptimal results. The premise of anger management programs is that, whereas anger is inescapable, the psychological mechanisms that allow a person to control the behavioral outcome of anger arousal are subject to reeducation. Having patience, exhibiting composure, engaging in constructive problem solving, pausing for thought, and visualizing consequences are some of the hoped-for results of a typical program. The phrase "rational-emotive psychoeducational therapy" is sometimes used to describe programs that ostensibly give clients the opportunity to practice new strategies for responding to provocations and to receive feedback in a supportive group. Anger management practitioners have stated that their methods are not intended for severe aggression or for disordered behavior but for ratcheting down the counterproductive behaviors that arise in ordinary people in day-to-day situations. Despite several positive reports, insufficient research has prevented reaching firm conclusions about the efficacy of a well-defined anger management program for long-term benefits in any subgroup of aggressive individuals.

General Statement Regarding Neuropharmacological Interventions

As noted earlier, relatively few RDBPC trials have been conducted in which a genuinely homogeneous patient population was assessed and the outcome measure was actual aggression rather than some proxy measure such as self-rated hostility. Goedhard et al. authored a relatively recent review of medication treatments for externalizing aggression in adult psychiatry. Altogether, 467 publications were screened, and 35 RDBPC trials were analyzed. "Weak evidence of efficacy" was reported for the

following:

1. Classic antipsychotics versus placebo in patients with schizophrenia or BPD (one of two studies with acceptable generalizability).
2. Atypical antipsychotics versus placebo and/or haloperidol in patients with schizophrenia, BPD, or PTSD (five of seven studies generalizable).
3. β-Adrenergic blockers versus placebo in patients with schizophrenia (four of five studies generalizable).
4. Anticonvulsants versus placebo in patients with cluster B personality disorders (two of six studies generalizable).
5. Antidepressants versus placebo in patients with schizophrenia, autism, cluster B personality disorders, PTSD, and depression (six of ten studies generalizable).

Unfortunately, this review collected studies of aggression associated with heterogeneous clinical problems using noncomparable outcome measures and did not stratify results according to specific medications. Because two medications approved for the same indication (e.g., antidepressants) may have quite different pharmacological actions, it is difficult to interpret the results of this review or to translate them into practice recommendations.

In the interests of summarizing the state of the art, such as it is, the literature has been reviewed on nonpharmacological and pharmacological interventions for aggression associated with a number of specific conditions. The results of this review are not represented as definitive; a comprehensive assessment and selective meta-analysis of published data have not been completed. The treatments suggested here represent conclusions combining a large-but-incomplete review with personal experience. In every case when a medication is suggested, a concurrent psychotherapeutic relationship is assumed to be an invaluable component of the treatment.

Treatment of Selected Aggressive Behaviors Primarily Seen in Childhood and Adolescence

Aggression in Healthy Children or in Children with ODD. For several decades, ambitious studies have attempted to reduce societal aggression through true primary prevention. In several much-cited experiments, psychosocial intervenient with children aged 0 to 9 years who were not previously identified as “at risk” were shown to decrease conduct problems, delinquency, and arrest rates in follow-up periods as long as 27 years. Some of these interventions emphasized educational services and academic skills, others parenting support via home visits, others stress reduction, and still others child care. Most effective programs included a mix of these components. This makes it difficult to state which component is critical to success—a problem that stands in the way of designing optimum

aggression-reducing public policies.

Child psychiatry has also worked hard to identify interventions to prevent aggression in "at-risk" school-aged children (typically meaning those living in poverty with single parents). Multifaceted programs have incorporated various combinations of academic help, classroom-based problem-solving training for children, home-based parental skills training, and group parental skills training. Such combined programs have demonstrated efficacy in follow-up periods as long as 6 years. Again, insufficient evidence is available to distinguish one superior strategy, frustrating the hope of fundamental policy planning to prevent aggression in large populations.

Regarding the aggression most commonly reported to pediatricians—that observed in essentially healthy children, some of whom meet criteria for the diagnosis of ODD—many studies have reported the results of supportive psychotherapy, insight-oriented psychotherapy, individual and group CBT, or "anger management" interventions. The quality of these studies varies. Some evidence supports a beneficial effect of psychotherapy, but insufficient data are available to assess the relative efficacy of different theoretical approaches. Many authorities emphasize the problem of excessive focus on the identified aggressor and that family therapy is usually necessary.

Treatment of Aggression in Mental Retardation or Developmental Disabilities. Because very different anatomical and neurochemical problems may be present, even more so than in the other categories of aggression, treating aggression in children who fall into this wide diagnostic orbit becomes a matter of trial and error.

NONPHARMACOLOGICAL INTERVENTIONS. Behavioral modification approaches to managing aggression in MR are supported by limited studies. One controlled trial demonstrated reduced anger with individual CBT. Less research supports "anger management" approaches in this group of disorders.

PHARMACOLOGICAL INTERVENTIONS. Lithium carbonate was reported in several studies (only one placebo controlled) to help manage aggression in individuals with MR. Some evidence supports the use of atypical antipsychotics, including risperidone (Risperdal), clozapine (Clozaril), olanzapine (Zyprexa), and quetiapine (Seroquel), to reduce aggression in individuals with intellectual disabilities. At least three RDBPC trials specifically demonstrated the efficacy of risperidone. Some evidence also supports the use of anticonvulsants, including carbamazepine (Tegretol) and divalproex sodium (Depakote). β-Blockers, including propranolol (Inderal), nadalol (Corgard), and pindolol (Visken), have all been reported to be effective in reducing aggression in MR in a few trials. A useful

provisional treatment recommendation is that risperidone seems to be the best-supported single choice, although the risks of weight gain, metabolic syndrome, and hypercholesterolemia mandate caution, especially in adolescents. Studies are needed to determine whether evidence of cortical irritability on EEG predicts responsiveness to antiepileptic drugs/mood stabilizers such as divalproex sodium.

Treatment of Aggression in Autism

NONPHARMACOLOGICAL INTERVENTIONS. Most of the research in this field reports behavioral modification interventions. The goal is to replace problematic behaviors such as aggression with more developmentally appropriate behaviors. *Early intensive behavioral intervention* is a protocol that promises to guide behavioral development through systematic sequencing of environmental events. The studies have been small, and some appear to display a behaviorist bias. One trial of the TEACCH residential model (treatment and education of autistic and related communication-handicapped children) reported reduction in disruptive behaviors.

PHARMACOLOGICAL INTERVENTIONS. Risperidone (0.5 to 3.5 mg per day) has been reported to reduce irritability and aggression among children with autism in a multisite RDBPC trial and has been approved for that indication by the FDA. Other atypical antipsychotic medications, such as aripiprazole, are being investigated. Levetiracetam (Keppra) reduced aggressive behavior in children with autism who had not recently been weaned from other agents thought to reduce aggression. One should be aware that some evidence implicates this agent as a cause of irritability, depression, or psychosis. Of interest, in a small open trial of divalproex sodium, all of the children with abnormal EEGs exhibited benefits, and none of the nonresponders had abnormal EEGs. It remains to be determined whether IQ or some biological marker such as EEG evidence of cortical irritability will distinguish anticonvulsant responders from nonresponders. Limited evidence supports the use of β-blockers or α-2 agonists. A useful provisional treatment recommendation is that risperidone is the best-supported choice. However, it should be noted that none of the children in the largest trial had frequent seizures. Risperidone could lower the seizure threshold in the large proportion of autistic children with epilepsy. In such children, divalproex sodium might be the first-line choice.

Treatment of Aggression in ADHD

NONPHARMACOLOGICAL INTERVENTIONS. Depending on the child's maturity and tolerance for sustained talk, CBT may be helpful. This therapy is reportedly

especially useful for those with both ADHD and depression. *Multisystemic treatment* refers to a protocol that includes family, school, and prescribing physician engagement, usually orchestrated by a psychologist or a psychiatrist. This protocol is theoretically sensible, but its efficacy for reducing aggression in ADHD has yet to be demonstrated.

PHARMACOLOGICAL INTERVENTIONS. For clinicians who have observed the devastating proaggression effect of stimulants in adults with substance abuse or dependence, it may seem surprising that stimulants are in fact the best-studied and most effective medication for reducing aggression among children with ADHD: Several RDBPC trials and one meta-analysis confirmed the efficacy of methylphenidate (Ritalin). A small RDBPC trial of risperidone concluded that this agent is modestly effective when combined with stimulants. Two trials also reported that the α-adrenergic agonist clonidine (Catapres) reduces aggression among children diagnosed with ADHD and CD. Some evidence suggests that bupropion or desipramine might be effective. Children with mood and anxiety disorders and emotional lability may benefit from antidepressants or mood stabilizers, but the empirical evidence is not strong. A useful provisional treatment recommendation is that stimulant therapy, with or without clonidine or risperidone supplementation, is the best-supported choice.

Treatment of Aggression in Conduct Disorder. Because the criteria for CD permit this diagnosis in children with heterogeneous psychobiological characteristics and very different patterns of aggression, the biggest concern in interpreting published treatment studies is to what degree the subjects were comparable. Treatments reported to be useful in CD primarily associated with hyperactivity may be harmful in CD associated with hypoarousal. Perhaps the biggest concerns in selecting an intervention are to identify the provocative or psychobiological causal factors in each case, to individualize the therapy, and—given the proliferation of unlicensed for-profit centers capitalizing on the distress of parents—to ensure the competence of the treater. The most important known predictors of treatment outcome are age and severity. The younger and less disruptive the child, the better is the chance of success, which accounts for the encouraging new emphasis on preventive interventions. Ultimately, it can be expected that CD will be shown to embrace multiple entities such as hyperactivity, psychopathy with hypoarousal, preschizophrenia, or childhood bipolar illness. Each version of CD will presumably exhibit a different profile of treatment responsiveness.

NONPHARMACOLOGICAL INTERVENTIONS. An intensive 2- to 6-month parental monitoring and child problem-solving curriculum is being investigated to prevent children at risk from developing CD. For the treatment of already aggressive children, many randomized controlled studies of varying quality

have reported that structured, short-term parent psychoeducational programs with a behaviorist orientation (sometimes called *parent management training*) produce short-term (up to 4 months) symptom reduction. There is a trend for more intensive programs that involve the child to be more effective. Individual psychotherapy for such children has reportedly reduced symptoms but has yet to be studied rigorously. *Cognitive problem-solving skills training* refers to a CBT-oriented approach that has also been less well studied. *Multisystemic therapy,* as noted earlier, refers to a comprehensive treatment program intended to involve the parents, child, clinician, school, and sometimes community in a systems-based analysis of the causes and best responses to abnormal conduct. Limited controlled studies suggest a benefit from this approach. Day hospital or fully residential group treatment is common but not proven. Military style boot camps to treat CD, often administered against the patient's will, tend to be poorly regulated and potentially dangerous operations conducted by nonprofessionals. That caveat notwithstanding, one study reported that a *voluntary* 22-week residential camp of this type reduced externalizing symptoms compared with wait-list controls.

PHARMACOLOGICAL INTERVENTIONS. An increasing number of RDBPC trials have reported the efficacy of various agents in CD. Atypical antipsychotics have been shown to reduce aggression by 25 to 50 percent in CD. Risperidone is the by far the best-researched agent, reducing disruptive, aggressive behaviors and anger attacks in at least five studies. In one report, olanzapine reduced impulsive (but not predatory) aggression in adolescents with CD. Lithium was efficacious in three studies. Divalproex sodium, carbamazepine, and clonidine have demonstrated antiaggressive effects in fewer studies. SSRIs have not been convincingly shown to reduce aggression in this population, even in the presence of depression—perhaps more evidence that these agents act differently in populations of different ages. A useful provisional treatment recommendation is that risperidone is again the first-choice agent.

Treatment of Selected Aggressive Behaviors Primarily Seen in Adulthood

Treatment of Aggression in Schizophrenia and Schizoaf fective Disorder

NONPHARMACOLOGICAL INTERVENTIONS. Behavioral modification is reportedly helpful, although the published studies have not used a single strategy and have not been appropriately controlled. One promising approach has been to combine group-based cognitive skills training with a stepwise system of rewards for target behaviors. CBT has also been shown to reduce aggression among patients with chronic psychosis.

PHARMACOLOGICAL INTERVENTIONS. Many controlled trials of neuroleptics have been conducted, making violence in the schizophrenia spectrum one of the best-studied forms of aggression. Again, it should be emphasized that the relationship between paranoid ideation and aggression is not strong, so that the reduction in psychosis may be a weak marker of the efficacy of an agent for management of aggression. Because DA, serotonin, and NE dysregulation may affect aggression in schizophrenia, it should not be assumed that all antipsychotics would have equal efficacy. In several randomized comparison studies, atypical agents such as paliperidone, clozapine, risperidone, and olanzapine have been more effective than typical agents in reducing violence. Clozapine is postulated to have a specific antiaggressive benefit, beyond the effect of other neuroleptics and perhaps beyond the effect of other atypical agents, independent of the effect on psychosis or sedation. The risks of clozapine to bone marrow may outweigh its possible superiority as an antiaggression treatment. In a recent systematic review of pharmacological management of aggression in persons with schizophrenia spectrum disorders, Victoroff et al. (2014) reported that the strongest evidence supports the efficacy of paliperidone. It is unclear whether concomitant antidepressant or mood-stabilizing therapy will boost the antiaggressive benefit of neuroleptics in schizoaffective patients. A provisional treatment recommendation is that paliperidone may be a useful first choice. If it fails, risperidone, olanzapine, or even clozapine is worth considering.

Treatment of Aggression in Intermittent Explosive Disorder. IED is not a unitary entity. As noted earlier, some cases may be physiologically more closely related to epilepsy, others to mood disorders, and others to entirely different etiologies. Treatment with a wide variety of methods has been proposed and is probably successful for subpopulations for which a given intervention addresses the major underlying problem. It can be expected that clinical and biological subtyping will eventually improve on the current trial-and-error approach to treatment.

NONPHARMACOLOGICAL INTERVENTIONS. Brief psychological interventions have been proposed as helpful in cases of court-referred drivers with presumed IED. CBT seems to be the only form of talk therapy demonstrated to be beneficial in a placebo-controlled trial. It is likely that IED patients with affective distress (and perhaps with comorbid ADHD) are more likely to improve with talk therapy than are those with underlying personality disorders, hypoarousal, or brain tumors. Medications trials are usually indicated, whatever the diagnosis.

PHARMACOLOGICAL INTERVENTIONS. Many retrospective chart reviews and open trials of medications have explored potential treatments for explosive

anger episodes. The largest amount of published evidence from RDBPC trails supports the benefit of lithium (Carbolith) and sodium valproate (Dapakene). Phenytoin (Dilantin) reduced aggression in a mixed population of aggressive, impulsive prisoners, including many who probably met IED criteria. The β-blockers propranolol and pindolol have both been effective in small RDPBC trials. Carbamazepine was popular in a large practice survey and two small RDBPC trials support its use. The α-2 agonist clonidine and the antidepressant citalopram (Celexa) have been effective in small open trials.

Although it has been proposed that patients with borderline abnormal EEGs are more likely to benefit from antiepileptic drugs, this has never been formally investigated. Some authorities suggest an algorithm that considers the presence of an affective component: When depressive symptoms are present, SSRIs or lamotrigine (Lamictal) may be tried. When manic symptoms are present, mood stabilizers, antipsychotics, or antiepileptic drugs are tried first. A useful provisional treatment recommendation is that in cases with hints of abnormal brain electrical activity by history or testing, mood-stabilizing antiepileptic drugs might be the first resort. In cases of IED with histories of TBI but absent epilepsy, β-blockers may be a reasonable first choice. Patients with elements of hyperactivity might respond to stimulants, whereas those with affective distress may require SSRIs or mood stabilizers.

Treatment of Aggression in Borderline Personality Disorder

NONPHARMACOLOGICAL INTERVENTIONS. One impressive study compared the efficacy of twice-weekly transference-focused psychotherapy, weekly individual and group dialectic behavior therapy, and weekly supportive treatment. Only transference-focused therapy reduced overt aggression at 1 year. This appears to be the sole controlled study demonstrating efficacy for this type of intervention for any aggressive condition. The fact that patients in the transference-focused arm received twice the treatments makes it impossible to conclude that the type of clinical–patient interaction was relevant to outcome. Still, these results support the contention that intensive talk can reduce aggression in BPD.

PHARMACOLOGICAL INTERVENTIONS. Several RDBPC and open trials support the efficacy of divalproex sodium to diminish aggressive behaviors in BPD. Some evidence suggests that trait impulsivity predicts such responsiveness. One might speculate that the nonepileptiform EEG abnormalities sometimes seen in these patients would predict response to an antiepileptic drug, but one study found that EEG results did not predict response, and another study found no relationship between evoked potential results and response to valproate. One RDBPC trial found that lamotrigine reduced anger in women with BPD, and another RDBPC trial reported that

topiramate (Topamax) reduced anger and aggression in men with BPD. Several trials of atypical antipsychotic agents suggest benefits. Aripiprazole (Abilify) reduced anger in an RDBPC trial in a mostly female group of patients with BPD, a benefit that was sustained for 18 months. Open-label trials also reported reduced anger outbursts in patients treated with quetiapine, clozapine, and risperidone. In one intriguing study, the omega-3 fatty acid ethyl-eicosapentaenoic acid reduced aggression among women with BPD, another suggestion of a possible link between cerebral fatty acid metabolism and aggression. A useful provisional treatment recommendation is that the chance of success probably depends on tailoring treatment to the most likely driver of aggression. Aggressive BPD patients with prominent affective disturbance may benefit from mood stabilizers. Those with prominent idiosyncratic thoughts, especially paranoia, may do better with antipsychotics.

Treatment of Aggression in Alzheimer Disease and Related Dementias

NONPHARMACOLOGICAL INTERVENTIONS. The benefits of talk or other nonpharmacological therapies depend on the level of cognitive decline. Few mildly affected patients (those who could most benefit from a therapeutic alliance) are aggressive, and, unfortunately, by the time aggression is a significant problem, most patients have lost the ability to recollect from one talk therapy session to the next, obviating carryover. Behavioral modification has been useful according to some reports. Care-resistant aggression—lashing out during attempted dressing, bathing, and toileting—is common and may be reduced by caregiver training. A review of 41 studies conducted in health care facilities determined that staff training programs and environmental modifications to reduce triggers of aggression were the most efficacious approaches; however, very few of the reviewed studies were systematic or controlled. Structured activities, music therapy, and even aromatherapy have been reported to reduce agitation and aggression in small studies.

PHARMACOLOGICAL INTERVENTIONS. Many agents have been proposed and tried for managing agitation and aggression in AD. Atypical antipsychotic agents are the most prescribed medications. A formal review based on nine RDBPC trials concluded that risperidone and olanzapine are both useful in managing aggression in dementia. Less evidence supports the benefit of β-blockers. One RDBPC trial reported that memantine (Namenda) reduced aggression among patients treated with donepezil (Aricept). Of interest, in one study, growth hormone response to clonidine—an index of central noradrenergic function—predicted the antiaggression benefit of pindolol. It can be hoped that clinicians will soon be using both psychological and biological markers of treatment responsiveness to tailor aggression management for individual patients. A useful provisional treatment

recommendation is that in demented patients with reactive/impulsive aggression (pending identification of response predictors), low-dose risperidone and olanzapine are the best-supported choices. Caregiver training, support, and protection are essential.

Treatment of Aggression in TBI

NONPHARMACOLOGICAL INTERVENTIONS. Higher-functioning patients may significantly benefit from CBT, supportive, or insight-oriented psychotherapy. Behavioral modification methods may be helpful for patients with poor cognitive outcomes who reside in institutions or with supportive caregivers. Preliminary evidence supports the use of transcranial electrical stimulation.

PHARMACOLOGICAL INTERVENTIONS. Based on RDBPC trials, the best evidence supports the use of β-blockers, including propranolol and pindolol. Little evidence favors the efficacy of valproate, carbamazepine, lamotrigine, antidepressants, buspirone (Buspar), or methylphenidate. It is notable that, despite the relatively occasional concurrence of posttraumatic seizures and aggression and the more frequent presence of sharp waves on EEGs, antiepileptic drugs have not been proven efficacious. It is possible that (1) the efficacy of β-blockers may depend on whether hyperarousal occurs, and (2) the efficacy of antiepileptic drugs may depend on electrophysiological markers of cortical irritability. A useful provisional treatment recommendation is that for reactive/impulsive aggression or irritability after TBI in the absence of depressive symptoms, propranolol in doses beginning at 10 mg twice a day is a reasonable first choice.

Treatment of Intimate Partner Violence. The major problem with evaluating the efficacy of interventions in domestic violence/intimate partner violence is the tremendous diversity of perpetrators. The presence or absence of critical traits such as major mental illness, psychopathy, controlling temperament, violence use in other settings, and substance abuse may completely alter the clinician's approach. College men live in unique environments, and their aggression may require unique responses. It has been hypothesized that the male and female perpetrators perhaps require different approaches. It is also known that completion of CBT-based programs is linked to education. Therefore, this summary of research findings must be considered as especially tentative, given that the effect of interventions certainly depends on the aforementioned factors and others.

NONPHARMACOLOGICAL INTERVENTIONS. It should be acknowledged that a fierce debate accompanies the literature on psychological interventions. Some authorities strongly encourage the *Duluth model* of treatment, which consists in structured psychoeducational programming (a curriculum often

referred to as feminist oriented, teaching men that their behavior is part of an aberrant women-controlling male behavior pattern), combined with teaching alternative conflict behaviors and tension-reducing exercises. Other authorities promote CBT or insight-oriented psychotherapy. Most interventions comprise poorly delineated batterer intervention programs (BIPs), typically containing some element of psychoeducation or CBT and administered through courts and departments of social service by service providers of varying expertise. Integrated substance abuse/domestic violence treatment programs also exist, theoretically increasing the likelihood of success by simultaneously addressing these often-related problems. The effect size of BIPs is debatable. One recent meta-analysis found that these programs reduced official reporting of violence but had no effect (or even a harmful effect) on victim-reported outcomes. There may be a dose effect, given that completers exhibited much better success rates. One study reported the promising outcome that only 14.3 percent of completers of BIPs in Chicago (vs. 34.7 percent of noncompleters) were rearrested after more than 2 years of follow-up. Conjoint therapy has been proposed to increase the rate of success by treating the relationship rather than just the perpetrator. A provisional bottom line is that treatments that engage a batterer who wishes to change in a structured psychotherapeutic milieu, whether individual or group based, whether explicitly CBT based or not, sometimes reduce this dangerous behavior. Batterers who are not motivated, have criminal histories, or have Cluster B personality disorders or substance abuse disorders are much less likely to exhibit a treatment response.

PHARMACOLOGICAL INTERVENTIONS. One recent study reported the efficacy of fluoxetine to manage domestic violence. Unfortunately, this study did not stratify the sample according to testosterone level, history of family violence, presence of violent behavior outside of the relationship, substance abuse, or 5-HT polymorphisms. Because 5-HT deficiency has especially been associated with impulsive violence in psychopathy and in mood disorder, it is possible that partners with these diagnoses are more likely to benefit from SSRI therapy. No treatment has been convincingly proven to be effective, although the combination of psychotherapy with pharmacological interventions addressing specific modifiable disorders such as depression may be fruitful. Rigorous research is urgently needed to better understand how to match treatments to individuals in an effort to contain this major social problem.

Treatment of Sex Offenders. It is widely believed, even among mental health professionals, that sex offenders are beyond treatment. This belief is probably based on the many studies showing extraordinarily high rates of recidivism among incarcerated male rapists and child molesters. Given the

evidence that sex offenders typically commit many more crimes than those that come to official attention, it is also a concern that low rates of rearrest do not mean low rates of reoffending. The belief that these individuals are not treatable has been a major factor in a wave of legislation mandating indefinite confinement, indefinite monitoring, limited community reentry, or castration. Although it is well to maintain skepticism about the lasting efficacy of treatment in many cases, it is also important to acknowledge that recent research supports important differences in types of offenders and in their treatability.

NONPHARMACOLOGICAL INTERVENTIONS. As many as five recent meta-analytic reviews reported that CBT reduces the rate of recidivism among rapists. Thus, contrary to a nihilistic tone in many discussions of treatment responsiveness, talk therapy is well documented to help. Studies are needed to identify patient characteristics predicting response to this form of therapy—such as variations in psychopathy, depression, testosterone level, or genetic markers. The observation that obstructive sleep apnea may increase hostility among sex offenders led to one study reporting that nocturnal continuous positive airway pressure treatment reduced self-reported aggression.

MEDICAL INTERVENTIONS. Orchidectomy reduces the rate of recidivism to ~3 percent. This strongly suggests that circulating steroid hormones, testosterone in particular, play a key role in sex offender behavior. Because relatively few offenders will be referred, volitionally or not, for this surgical procedure, antiandrogen therapies have long been used in an attempt to produce the same benefits. Estrogen therapy is associated with many serious side effects, including breast cancer, and has been abandoned. Instead, the most popular treatments are the synthetic progesterone derivatives medroxyprogesterone acetate (MPA) (Provera) (mean dose about 300 mg per week intramuscularly) and cyproterone acetate (CPA) (Cyprostat) or the more recently introduced GnRH agonists, such as or the luteinizing hormone–releasing hormone (LHRH) analogue triptorelin (Trelstar) or the LHRH agonist leuprolide acetate (Lupron) (7.5 mg per month intramuscularly or 22.5 mg every 3 months). This approach suppresses gonadotropin secretion, although there is usually an initial transient increase in androgenesis. The peripherally acting antiandrogen flutamide (Eulexin) (750 mg per day in divided doses for 2 to 4 weeks) antagonizes testosterone receptors and may be a useful adjunct therapy. These hormonal "castration" treatments are effective in reducing sexual aggression, but success depends on hard-to-ensure compliance with long-term treatment and probably with concurrent psychotherapy. Some evidence suggests that sexual aggression may be reduced with antiandrogen therapy, but other types of physical aggression may persist. A

useful provisional treatment recommendation is that, given the risks of side effects such as erectile dysfunction, gynecomastia, and bone density changes and the complex ethical and legal implications, treatment of this patient group is best left to subspecialists.

Several states in the United States have recently responded to the perception that there exist dangerous and untreatable sex offenders (referred to as a *sexually violent predators/sexually dangerous persons*) with very restrictive and coercive measures. These include involuntary chemical castration, indefinite incarceration, and a California law mandating lifetime Global Positioning Satellite monitoring. Some persons convicted under these laws have selected orchidectomy as the only route to release. These laws have not been formulated on the basis of evidence of the efficacy of these interventions or with a view toward psychiatrically monitored individualized treatment. This legal initiative, which may sometimes exhort an ethical deviation from the physician's expected role, has generated a serious controversy in the community of sex offense and paraphilia treatment professionals, which is unresolved at the time of this writing.

FUTURE DIRECTIONS

Research is expected to yield progress in at least three domains of inquiry in the next several decades. First, the biopsychosocial mechanisms that mediate pathological aggression will be better understood. Rather than making diagnoses fulfilling lists of operational criteria, neuropsychiatrists will be closer to distinguishing discrete problems resulting from genetic variation, epigenetic variation, and gene–gene (epistatic) and gene–environment interactions that render some individuals more prone to aggressing in ways that society cannot sanction. For instance, rather than labeling both of two similar 8-year-old boys with "CD," clinicians may soon be able to determine that one boy has a reward-deficiency syndrome related to DA receptor polymorphism causing persistent hypoarousal, whereas the other exhibits early-maltreatment–related expression of a serotonin-transporter-allele–mediated condition producing frontolimbic hyperreactivity to social threat. The optimum treatments might be dramatically different.

Second, as the genetic and developmental/environmental influences that produce pathological aggression are better understood, not only will treatments be better matched to pathophysiologies but also early detection and early intervention will decrease the net social burden of adult aggression. Regarding the genetics of aggression, neuropsychiatry has barely glimpsed the tip of the iceberg. Modern technologies are fast reducing the cost and time required to map human genomic individuality. The variations contributing to early-onset life-persistent aggression will eventually be uncovered. Rather than a one-gene, one-disease model,

psychiatry will soon be able to consider how 5 or 25 gene variants, interacting with epigenetic and environmental modifiers of gene expression, contribute to a trait, as well as how 5 or 25 traits are affected by a single gene variant. Eventually, a genome profile might prove helpful to predict a patient's risk for aggression and to optimize primary prevention. As the price of scanning a person's entire genome falls by orders of magnitude, individual access to such data will become commonplace.

Third, future research should also alter society's management of aggressive individuals. At present, the legal response to violence is largely based on an insupportable dichotomy: either the alleged perpetrator was sane or insane. Either he or she had the capacity to distinguish right from wrong at the time of the alleged offense or did not (a paid expert opinion that is little more valid than a guess). Neurolaw concepts are slowly maturing. *Degrees* of responsibility are slowly being acknowledged. For example: human groups often moderate their rewards and punishments on the basis of age. But consider a comparison: two sane men are both chronologically aged 28 years. One is neurobiologically typical for the age of 28 years. The other is neurobiologically atypical for that age, with a 28-year-old's capacity for verbal, mathematical, and memory functions; the emotional regulation capacity of an 11-year-old; the impulse control capacity of a 9-year-old; and a loss of reality testing 15 to 20 percent of the time. Are they equally responsible? One does not expect societies to embrace a sliding scale of moral responsibility and proportional administration of restraints on liberty versus treatments informed by neuroscience. As shown by the ongoing enthusiasm for the death penalty, some societies have yet to be weaned from the fantasy that retributive justice is an effective deterrent. However, the indefensible dichotomy of sane versus insane cannot be sustained. The all-or-none concept of criminal responsibility will be eventually replaced with wiser judgments of the way that human variation requires a nuanced administration of the social contract.

References

*Adams DB. Brain mechanisms of aggressive behavior: an updated review. *Neurosci Biobehav Rev*. 2006;30:304–318.

Ainsworth SE, Maner JK. Sex begets violence: mating motives, social dominance, and physical aggression in men. *J Pers Soc Psychol*. 2012;103:819–829.

Anholt RR, Mackay TF. Genetics of aggression. *Annu Rev Genet*. 2012;46:145–164.

Bortolato M, Pivac N, Seler DM, Perkovic MN, Pessia M, Di Giovanni G. The role of serotonergic system at the interface of aggression and suicide. *Neuroscience*. 2013;236:160–185.

Bureau of Justice Statistics. Data collection, National Crime Victimization Survey. Available at http://www.bjs.gov/index.cfm?ty=dcdetail&iid=245. Accessed July 11, 2015.

Coccaro EF, Schmidt CA, Samuels JF, Nestadt G. Lifetime and 1-month prevalence rates of

intermittent explosive disorder in a community sample. *J Clin Psychiatry*. 2004;65:820–824.

Craig IW. The importance of stress and genetic variation in human aggression. *Bioessays*. 2007;29:227–236.

de Boer SF, Olivier B, Veening J, Koolhaas JM. The neurobiology of offensive aggression: revealing a modular view. *Physiol Behav*. 2015;146:111–127.

Eurostat. Crime statistics. Available at http://ec.europa.eu/eurostat/statistics-explained/index.php/Crime_statistics. Accessed July 11, 2015.

Goedhard LE, Stolker JJ, Heerdink ER, Nijman HL, Olivier B. Pharmacotherapy for the treatment of aggressive behavior in general adult psychiatry: a systematic review. *J Clin Psychiatry*. 2006;67:1013–1024.

Haller J. The neurobiology of abnormal manifestations of aggression–a review of hypothalamic mechanisms in cats, rodents, and humans. *Brain Res Bull*. 2013;93:97–109.

Hassiotis AA, Hall I. Behavioural and cognitive-behavioural interventions for outwardly-directed aggressive behaviour in people with learning disabilities. *Cochrane Database Syst Rev*. 2008;3:CD003406.

*Hemenway D. *Private Guns, Public Health*. Ann Arbor, MI: University of Michigan Press; 2010.

Loeber R, Pardini D. Neurobiology and the development of violence: common assumptions and controversies. *Philos Trans R Soc Lond B Biol Sci*. 2008;363:2491–2503.

Malki K, Pain O, Du Rietz E, et al. Genes and gene networks implicated in aggression related behaviour. *Neurogenetics*. 2014;15:255–266.

Megargee El, Carbonell JL, Bohn M, Sliger GL. *Classifying Criminal Offenders with MMPI-2: The Megargee System*. Minneapolis, MN: University of Minnesota Press; 2001.

*Nelson RJ, Trainor BC. Neural mechanisms of aggression. *Nat Rev Neurosci*. 2007;8:536–546.

Parletta N, Milte CM, Meyer BJ. Nutritional modulation of cognitive function and mental health. *J Nutrition Biochem*. 2013;24:725–743.

Patrick CJ. Psychophysiological correlates of aggression and violence: an integrative review. *Philos Trans R Soc Lond B Biol Sci*. 2008;363:2543–2555.

Rampling J, Furtado V, Winsper C, et al. Non-pharmacological interventions for reducing aggression and violence in serious mental illness: a systematic review and narrative synthesis. *Eur Psychiatry*. 2016;34:17–28.

Roeder O, Eisen L-B, Bowling J. *What caused the crime decline? Brennan Center for Justice*. 2015. Available at https://www.brennancenter.org/publication/what-caused-crime-decline. Accessed July 11, 2015.

Rosell DR, Siever LJ. The neurobiology of aggression and violence. *CNS Spectr*. 2015;20:254–279.

Salzman C, Jeste DV, Meyer RE, Cohen-Mansfield J, Cummings J. Elderly patients with dementia-related symptoms of severe agitation and aggression: consensus statement on treatment options, clinical trials methodology, and policy. *J Clin Psychiatry*. 2008;69:889–898.

Sharma T, Guski LS, Freund N, Gøtzsche PC. Suicidality and aggression during antidepressant treatment: systematic review and meta-analyses based on clinical study reports. *BMJ*. 2016;352:i65.

Struber D, Luck M, Roth G. Sex, aggression and impulse control: an integrative account.

Neurocase. 2008;14:93–121.

Tiihonen J, Isohanni M, Rasanen P, Koiranen M, Moring J. Specific major mental disorders and criminality: a 26-year prospective study of the 1966 northern Finland birth cohort. *Am J Psychiatry*. 1997;154:840–845.

Truman JL, Langton L. *Criminal victimization. U.S. Department of Justice, Office of Justice Program, Bureau of Justice Statistics*. Available at http://www.bjs.gov/index.cfm?ty=pbdetail&iid=5111. Accessed June 30, 2015.

United Nations Office of Drugs and Crime: Global Study on Homicide 2013. United Nations publication, Sales No. 14.IV.1.

United States Department of Justice, Federal Bureau of Investigation. Crime in the United States, 2013. Available at https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2013/crime-in-the-u.s.-2013/violent-crime/violent-crime-topic-page/violentcrimemain_final. Published November 2014. Accessed July 10, 2015.

Victoroff J, Coburn K, Reeve A, Sampson S, Shillcutt S. Pharmacological management of persistent hostility and aggression in persons with schizophrenia spectrum disorders: a systematic review. *J Neuropsych Clin Neurosci*. 2014;26:283–312

Wintemute GJ. The epidemiology of firearm violence in the twenty-first century United States. *Annu Rev Public Health*. 2015;36:5–19.

31

Special Areas of Interest

▲ 31.1 Psychiatry and Reproductive Medicine

SARAH L. BERGA, M.D., BARBARA L. PARRY, M.D., AND EYDIE L. MOSES-KOLKO, M.D.

Reproductive events and processes have both physiological and psychological concomitants. Likewise, psychological states affect reproductive physiology and modulate reproductive events. This section examines these bidirectional relationships with the goal of introducing fundamental concepts related to the classic reproductive events, such as menarche, pregnancy, delivery, postpartum, and menopause. The fields of psychiatry and reproductive medicine continue to define the multiple mechanisms by which psyche and soma interact to determine a woman's gynecological and psychological health. For instance, premenstrual dysphoric disorder (PMDD)—the mood, cognitive, and behavioral changes that occur in association with the menstrual cycle—exemplifies a somatopsychic disorder in which biological changes occurring in the soma trigger changes in psychological state. In contrast, functional forms of hypothalamic anovulation represent psychosomatic illness that originates in the brain but alters somatic functioning. Medicine has traditionally separated the treatment of reproductive events from that of psychological functioning; however increasing awareness of a false dichotomy between mind and body is beginning to shape new systems and approaches to care that place greater attention on addressing mental health needs in medical settings.

REPRODUCTIVE DEVELOPMENT

Fetal and Neonatal Development

The sex of an embryo is determined during the fertilization process. The normal chromosomal contribution of the oocyte (egg) is 23 (including an X chromosome), and the sperm contributes 23 (including an X or a Y chromosome). Translation of gonadal sex into phenotypic sex largely depends on the type and secretory activity of the gonad formed and on the

responsiveness of target tissues in the developing fetus to gonadal products. The fetal testes begin to secrete testosterone and anti-müllerian hormone (AMH) toward the end of the first trimester. If secretion is absent or compromised or if the fetal tissues are androgen insensitive, a spectrum of incomplete masculinization ensues. In addition, excess androgen exposure at this time from maternal or fetal sources, such as occurs in fetal congenital adrenal hyperplasia, can partially masculinize a female fetus.

Sexual differentiation of the central nervous system (CNS) is largely dependent on the effect of testosterone on target areas within the developing brain. The fetal testes begin to secrete testosterone in the late first trimester in response to placental human chorionic gonadotropin (hCG). In contrast, the fetal ovary does not develop functional follicles within the fetal ovary until the late second trimester. In late pregnancy, gonadotropin secretion is restrained by placental-steroid production. When that restraint is lost at the time of birth, gonadotropin secretion rises dramatically in both sexes. In boys, the gonadotropin rise is followed by an elevation of testosterone concentrations to adult levels, with ongoing influences on brain development for a short period postnatally that does not last for more than 1 year. In girls, estrogens from fetal or placental sources do not cross the blood–brain barrier because of circulating binding proteins, and thus circulating estrogens do not have access to and cannot imprint the developing CNS. However, the gonadotropin hiatus is much later in girls than in boys, so the ovary of young girls is under less restraint postnatally than are the testes of boys. Thus, by 2 years of age, the brains and bodies of girls and boys have been exposed to dramatically different patterns of sex-steroid secretion.

The organizational effects of testosterone on the developing CNS are thought to depend primarily on in situ aromatization (the conversion of androgens to estrogens by the enzyme aromatase) of testosterone to estradiol. In addition, testosterone may bind directly (without conversion to estradiol) to androgen receptors in the CNS. Therefore, sexual dimorphism of brain structure may reflect not only androgen exposure, but also central dimorphism in key brain areas in terms of both high aromatase activity (which, thus, can convert testosterone to estradiol) and the presence of androgen receptors. In the nonhuman primate brain, androgen receptors are located in the hypothalamus, amygdala, and prefrontal visual and somatosensory cortices, whereas estrogen receptors are more highly concentrated in hippocampus in females relative to males. The degree to which gender-related behavioral asymmetries are accounted for by differences in fetal/infant-steroid hormone exposure is open to debate, but, clearly, a mechanism for inducing differences exists.

Morphometric studies of the developing and adult brain support the sex-steroid theory of CNS sexual dimorphism. Sex differences cluster in brain regions with developmentally rich sex–steroid-receptor distribution. Sexual dimorphism is prominent in the cortex, which may be a result of rapid postnatal decreases in sex–steroid-receptor concentrations. By the age of 6 years, the brain volume is 95 percent of that of an adult; however, boys and girls follow different trajectories in brain tissue changes between the age of 6 years and adulthood. By adulthood, men have 10 percent larger brain volumes relative to women, larger corpus callosum genu and overall white matter volume, and increased ventricular size. By adulthood, women have larger cortical volume relative to men, higher cortical gray matter volume, and larger volumes in Broca area, the superior temporal cortex, hippocampus, caudate, thalamus, anterior cingulate gyrus, dorsolateral prefrontal cortex, and interhemispheric white matter connections.

Finally, the coming years will reveal greater understanding of how sex-chromosome genes, independent of gonadal hormones, play a role in sexual differentiation of brain cells. This concept has already been substantiated from research on embryonic tissues studied before gonadal development or through animals engineered to be discordant for genotypic and gonadal sex. For example, sex-chromosome determination of the midbrain dopamine system has been described. New models are under development to guide further exploration of genetic contributions to CNS sexual differentiation.

Puberty

In humans, the postnatal gonadal hiatus involves the desynchronization and suppression of the hypothalamic gonadotropin-releasing hormone (GnRH) pulse generator by undetermined central mechanisms that are largely gonad independent. Puberty—the postnatal resumption of gonadal activity—is also a centrally initiated process that depends on synchronization of the pulsatile release of hypothalamic GnRH, which is likely triggered at least in part by expression of the *KiSS1* gene (a gene that encodes a family of neuropeptides called kisspeptins). GnRH neurons are endogenously pulsatile and active during fetal and neonatal life. In experimental conditions with primates, the entire pubertal process can be initiated and completed simply by providing exogenous pulses of GnRH or by stimulating the dormant GnRH neurons with excitatory amino acids. Although prenatal exposure to sex steroids has organizational effects on the developing brain, the reinitiation of gonadal activity during puberty is viewed as exerting primarily activational influences on the brain and behavior.

Gender

The phenotypic sex of an individual depends on chromosomal sex and on exposure and responsiveness of target tissues to endogenous and exogenous sex steroids. *Gender* refers to the self-image and sex-role identity of that person. Gender is determined not only by exposure and responsiveness to sex steroids, but also by expectations and behavioral patterns that are learned in early childhood from parental, familial, and societal models. Gender develops across the life cycle in response to psychological and biological life events. Notions of gender direct marital, career, and other behavioral choices and may depart from biological options. For example, a woman with *müllerian agenesis* has ovaries but no uterus and vagina. Her childbearing and sexual options are constrained biologically, yet her sense of gender-appropriate behavior may drive her to pursue a heterosexual relationship and motherhood via surgical and technological advances.

Although gender is influenced by cultural expectations, each individual understands and responds to these expectations differently. To safeguard gender development, psychiatrists and psychologists must understand the medical and technological options available to individuals with disorders that alter phenotypic sexual development, and pediatricians, internists, and gynecologists must be mindful of the psychological effects of these disorders. Honesty is important, but physicians must convey information in a sensitive fashion. The clinical objective is to maintain a flexible mindset to facilitate and support appropriate developmental choices.

Disorders of sexual development in phenotypic girls generally present at birth or at the time of expected puberty. *Adrenogenital syndrome* can be caused by congenital adrenal hyperplasia in female fetuses; the fetal adrenal secretes excess androgens, which then masculinize the external genitalia and, possibly, the brain. In female infants, the masculinized genitalia are usually recognized in the delivery room. Surgery to reduce clitoral size and create or widen the vaginal introitus may be necessary at a later date to restore the external appearance. However, the internal reproductive tract is normal, and puberty generally occurs around the expected time

if adrenal replacement therapy is adequate. Despite medical and surgical interventions, concerns over capacity for sexual function may emerge, particularly in late adolescence or young adulthood.

Turner syndrome (XO gonadal dysgenesis) may be recognized at birth because of the associated physical stigmata, whereas XX gonadal dysgenesis usually presents at puberty. From a reproductive perspective, the two conditions are similar in that the ovaries do not contain responsive oocytes due to premature atresia of oocytes or failure of germ cell migration. Donor oocytes allow the option of pregnancy.

Other disorders of sexual development include testicular feminization and its variants, in which the gonads are testes, but the fetus is phenotypically female because a defect in the androgen receptor confers androgen insensitivity. Because the testes are normal, they secrete AMH, and *müllerian regression* (regression of the anlage that would develop into the uterus and tubes) occurs. The vagina ends blindly, and the presenting complaint is usually primary amenorrhea. *Thelarche* (onset of breast development) occurs at the normal time because, at puberty, the testes secrete testosterone, which then is aromatized to estradiol and stimulates the growth of breast tissue. Orchiectomy (removal of the testes) is recommended after puberty to avoid the risk of gonadoblastoma, unless there is partial or incomplete androgen sensitivity, in which case it is performed earlier to prevent the pubertal development of hirsutism and partial masculinization.

There are many types of androgen receptor defects and a spectrum of clinical presentations. Some individuals look phenotypically like men and may present with infertility secondary to azoospermia or oligospermia. A related disorder occurs when there is an inefficiency of the enzyme 5-α-reductase, which converts testosterone to dihydrotestosterone (DHT); DHT is required for full masculinization of the external genitalia. Boys with a deficiency of that enzyme may look phenotypically female or incompletely masculinized at birth. At puberty, further masculinization, including phallic enlargement, may develop as the increased testicular secretion of testosterone partially overcomes the enzyme deficiency and more DHT is made. Fertility is preserved in this condition, so it is prudent to rear the child as a male if possible. Neonatal treatment with DHT may help to masculinize the external genitalia.

XY gonadal dysgenesis is caused by a streak gonad that fails to secrete testosterone and AMH; generally, the uterus, tubes, and vagina are present. This condition is detected when thelarche and menarche do not occur. A karyotype can confirm the XY chromosomal status. Because the gonads are inactive hormonally, they should be removed before puberty to avoid the risk of malignant degeneration. Exogenous hormone replacement is required to stimulate puberty and cause the development of secondary sexual characteristics. If donor oocytes are available, pregnancy is possible after in vitro fertilization and embryo transfer. Telling a young adolescent and her parents about the diagnosis can be difficult, but it is best to be honest, particularly because surgery will be required to remove the gonads.

A clinician who deals with patients with disorders of sexual development must have a clear understanding of the basic physiology of sexual differentiation, including that of the brain, as well as a sensitive perspective on the interactions among gender, sexuality, and reproductive capacity.

Early Sexuality

The behavior of children can be understood by acknowledging the sexual motivations that, until Sigmund Freud's time, had been assumed to be quiescent until adolescence. The fact that incest, sexual abuse, and other sexual taboos that may occur in childhood have powerful aftereffects, even in adult life, bespeaks the importance of childhood sexuality in establishing healthy relationships.

In puberty, the overriding influence on sexuality is the dramatic change in the hypothalamic–pituitary–gonadal axis that occurs in both boys and girls at that time. Major hormonal changes, combined with the issue of identity diffusion, contribute to the characteristic tension in relationships, particularly between male and female peers and between adolescents and their

parents.

In adulthood, one of the major indicators of maturity is the capacity to establish healthy sexual relationships. When it is in harmony with other components of the self and with others, sexuality is an integral component of the adult character.

Adolescence

Carol Gilligan observed that preadolescent girls exhibit a sense of self-confidence and a belief that they can influence the course of events in their lives, whereas during adolescence, that self-esteem and belief in their own efficacy appear to diminish. Before puberty, boys are more vulnerable than girls are to depressive illness. Over the course of adolescence, girls display a precipitous rise in depression rates that outpaces that of adolescent boys. By age 15 years, girls are twice as likely as boys are to have experienced a lifetime episode of major depression. This gender difference was linked to the onset of puberty rather than to chronological age and persists for the next 35 to 40 years, spanning the female reproductive years. The predominance of depression in women and the adolescent onset of this gender difference have been replicated cross-nationally and are considered to be one of the most robust findings in the field of psychiatric epidemiology.

Negative life events and chronic psychosocial stressors are known to place individuals at risk for major depressive episodes. A number of theories about the gender difference in depression point to the fact that women are more likely than men are to experience certain types of traumatic life events, such as physical or sexual abuse, or chronic life stressors, such as single parenthood, gender discrimination in the workplace, or financial adversity. Growing evidence indicates that puberty may sensitize girls to the depressogenic effects of life stress. Because women tend to develop a strong affiliative style in their social relationships, as well as a sense of self in connection with others, social stressors, such as conflicts, breaches, or losses within interpersonal relationships, contribute to their depression risk. (Men, in contrast, tend to focus on issues of individuation and assertiveness within social relationships and to develop an agenic and autonomous sense of self.)

Neurohormonal processes associated with reproduction or regulated by female reproductive hormones may contribute to the intensification of affiliative behaviors observed in girls at puberty. Recent theories have implicated the hypothalamic neurohormone oxytocin in facilitating such affiliative behaviors as maternal caregiving and adult pair-bond formation in female mammals. Shelley Taylor and coworkers proposed the existence of a uniquely female response to stress, which they termed the *tend-and-befriend response* based on this oxytocin-mediated caregiving system, which, for women, serves to decrease sympathetic and hypothalamic–pituitary–adrenal (HPA) responses to stress while simultaneously promoting affiliative responses. This stands in contrast to the prototypic model fight-or-flight human stress response. Instead of fighting or fleeing, quieting and tending to the needs of offspring when under stress and affiliating with other adults to promote the protection of oneself and one's offspring may represent a more adaptive female response to environmental threats.

Although gender role socialization may begin well before puberty, social pressures to conform to stereotypically feminine versus masculine gender roles likely intensify during the adolescent transition, as the development of secondary sex characteristics make adolescents look more like adults and they begin to engage in greater levels of cross-sex interaction and dating. Of importance, hormonal changes in the reproductive neuroendocrine axis at puberty may potentiate and interact with this socialization process to promote an adolescent intensification of gender-based behaviors. Some have speculated that the increase in testosterone in boys and the attendant aggressiveness, often directed at girls, contribute to a change in girls' behavior during

adolescence.

MENSTRUAL CYCLES

Menstrual Physiology

Menstrual cyclicity results directly from ovarian cyclicity. Each ovarian cycle starts with the development of a group or cohort of follicles, one of which becomes dominant. The follicles are composed of an oocyte surrounded by granulosa cells, which, in turn, are surrounded by theca cells.

Follicular development is initiated by the hypothalamic release of GnRH at a pulse frequency of approximately one pulse every 90 minutes. GnRH stimulates the release of the pituitary gonadotropins luteinizing hormone (LH) and follicle-stimulating hormone (FSH). In turn, LH stimulates ovarian theca cells to synthesize and secrete androgens; FSH induces granulosa cell development, including the enzyme aromatase, which converts the thecally produced androgens to estrogens. In the presence of a constant GnRH pulse frequency of one pulse each 90 minutes, the secretion of LH and FSH in the follicular phase is regulated primarily by estradiol feedback at the level of the pituitary. Rising estradiol concentrations suppress FSH, thereby limiting the number of follicles that become mature oocytes capable of ovulating. When estradiol concentrations rise exponentially to exceed a critical threshold and remain elevated for at least 36 hours, which is the pattern one fully mature follicle produces, an LH surge is triggered, and ovulation (release of the ovum from the follicle sac) ensues approximately 36 hours later. Thereafter, granulosa cells transform into progesterone-secreting luteal cells, and the ovulated follicle is then referred to as the *corpus luteum,* which secretes progesterone.

The target tissues for ovarian steroids include the endometrium, whose developmental sequence is illustrated along the bottom panel, and the hypothalamic GnRH pulse generator, whose frequency, as indicated in the top right panel, is slowed dramatically by the combination of estrogen and progesterone secreted during the postovulatory or luteal phase of the menstrual cycle. This inhibition of GnRH is followed by decreased secretion of LH and FSH so that new follicular development is prevented until the corpus luteum regresses. As progesterone concentrations decline, GnRH pulsatility increases, and gonadotropin, especially FSH, secretion rises. The phases of the menstrual cycle can be termed *follicular* and *luteal* in reference to ovarian events or *proliferative* and *secretory* in reference to endometrial events.

It is generally assumed that regular menstrual cyclicity at intervals of 25 to 40 days signals ovarian cyclicity and ovulation, but that assumption is erroneous. A normal luteal phase has a length of more than 11 days and midluteal concentrations of progesterone secretion that exceed 10 ng/mL (30 nmol/L). The woman whose hormonal levels across a 39-day menstrual cycle had a luteal phase whose adequacy, with a length of 16 days and a peak progesterone concentration of 19.1 ng/mL, cannot be doubted. In contrast, the woman whose hormonal concentrations are shown in the top two panels had an overall cycle length of 25 days. Although her estradiol levels rose to almost 300 pg/mL, the elevation in estradiol and LH was not followed by a rise in progesterone secretion. Thus, she had bleeding when her estradiol concentration fell, but the cycle was anovulatory.

As these examples demonstrate, menses and menstrual cycle interval cannot be used to determine whether a woman is or is not ovulatory or if the luteal phase was normal. The clinician who is linking a psychological state to a reproductive event must know how to determine the presence or absence of gonadal activity. Ovarian activity can be estimated by determining weekly estradiol and progesterone concentrations from the onset of one episode of menstrual bleeding to the next. Testicular activity can be estimated from one or two random determinations of testosterone. These assessments may not detect subtle compromise of gonadal function that may impair fertility, however, and further consultation with an appropriate specialist may be needed. Menstrual cycles can be expected to become more regular with advancing gynecological age (time since first menses), and, in the absence of pathology, they generally remain regular until 5 to 10 years before menopause. During the perimenopausal years, ovarian function is

characterized by higher estradiol and lower progesterone secretion. Ovarian function does not gradually decline but rather becomes erratic and unpredictable, with the potential for exposing the brain and soma to large fluctuations in hormone concentrations.

Premenstrual Syndrome

PMDD has now been included as a diagnosis in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), and is a depressive disorder with 5 percent incidence, that is associated with functional impairment and is entrained to the expected fluctuations in sex steroids associated with ovulatory menstrual cycles. The premenstrual syndrome (PMS) is a broader clinical category, without functional impairment, that affects up to 80 percent of reproductive-aged women. In PMDD, menstrual cycles are hormonally normal. Neuroimaging research has provided empirical evidence of hormonal effects on emotion processing–related activity in orbitofrontal, amygdala–hippocampal, and anterior cingulate cortices as well as brain reward circuits in monetary reward paradigms.

Functional Hypothalamic Anovulation

Defined as the absence of menstruation for more than 6 months for which there is no organic or pathological cause, *functional hypothalamic anovulation* or *amenorrhea* is a consequence of a nonorganic reduction in GnRH secretion that results in reduced pituitary secretion of gonadotropins and subsequent anovulation. Previously referred to as *idiopathic* or *psychogenic amenorrhea* in the psychiatric literature, functional hypothalamic anovulation is a diagnosis of exclusion. This implies that the ovaries are capable of responding to appropriate gonadotropin input of either endogenous or exogenous origin; that there are no structural abnormalities of the thyroid, adrenal, pituitary, or brain; that the use of drugs, including antipsychotic medications, does not account for the suppression of GnRH; and that the patient is not pregnant. Neuroimaging may be needed to establish that there are no significant anatomical lesions of the brain or pituitary. Two major questions concerning the pathogenesis of functional hypothalamic anovulation remain unresolved. First, the peripheral and central signals that disrupt GnRH pulsatility are poorly understood. Second, how do behavioral, cognitive, and personality variables activate the neural systems to disrupt GnRH secretion? Physicians and patients alike wonder what lifestyle variables provoke or contribute to this type of ovulatory dysfunction and what pharmacological and nonpharmacological treatment interventions should be considered.

The idea that psychogenic stress can induce reproductive dysfunction in women was formally introduced in 1946. The best biochemical evidence in support of the concept that stress impairs GnRH release in women with functional hypothalamic amenorrhea is the consistent demonstration that the activity of the HPA axis is increased. For instance, a prospective study found that young US women who developed transient amenorrhea while studying in Israel had

higher urinary cortisol concentrations on arrival than did those whose menses remained regular.

There appears to be a dose–response relationship between the type, severity, and number of stresses on one hand and the proportion of women who develop anovulation on the other. Biological and psychological predispositions may confer resistance or sensitivity to various stressors. Exercise, low weight and weight loss, affective and eating disorders, various personality characteristics such as perfectionism and unrealistic expectations, drug use, and a variety of external and intrapersonal stresses have been linked to the development of anovulation. Most women with functional hypothalamic anovulation, when carefully evaluated, display more than one of these traits or behaviors. Recent evidence suggests synergism between metabolic stressors that suppress the hypothalamic–pituitary–thyroidal axis, such as excessive exercise and nutritional restriction, and psychosocial challenges that activate the HPA axis. Thus, multiple chronic stressors are more likely to induce functional hypothalamic anovulation than is a single, larger-magnitude stressor.

Recovery is possible if response patterns to ongoing psychosocial demands are developed that are less likely to activate the central and metabolic processes that disrupt pulsatile GnRH release. The current standard of practice, other than observation, is to offer pharmacological interventions, such as oral contraceptives or hormonal replacement, if fertility is not desired and pharmacological ovulation induction if fertility is sought. However, pharmacological intervention alone does not lead to spontaneous recovery and cannot be expected to ameliorate stress-induced alterations in central neurotransmission and hypothalamic function or to reverse ongoing metabolic derangements secondary to exercise or weight loss. For instance, bone accretion does not proceed apace in the face of metabolic compromise, even if exogenous hormone replacement is given in supraphysiological doses. Although women with functional hypothalamic anovulation frequently do not meet DSM-IV-TR criteria for eating disorders or depressive disorders, all of these states are characterized by increased cortisol secretion, which alone can alter thyroid hormone secretion and action and alter metabolism and attendant central signals. Thus, it is not surprising to find that women with amenorrhea due to decreased GnRH drive, regardless of cause, have lower bone mineral density. Furthermore, pharmacological intervention alone may mask the recognition of psychological dysfunction and forestall the development of more effective response patterns. In addition, ovulation induction may place low-weight women with functional hypothalamic amenorrhea who conceive at risk for premature labor and intrauterine growth retardation. Fetal neuropsychological development may be compromised by the concomitant presence of hypothalamic hypothyroidism (euthyroid sick syndrome) or hypercortisolemia if ovulation is pharmacologically induced. If the parenting skills of women with functional hypothalamic amenorrhea are impaired by ongoing stress, their children may be at risk for poor psychosocial development. Clearly, treatment strategies need to consider that stress and mild psychological dysfunction can play an important role in the genesis of this form of anovulation. If an eating disorder is recognized, specialized psychiatric treatment is indicated.

Psychobiological characteristics can predispose a woman to chronic activation of central neural processes in response to commonplace events. Global hypothalamic dysfunction, including inhibition of GnRH pulsatility, can result. Although identifying the neuromodulators that mediate the development and maintenance of the disruption of GnRH is of academic interest, clinicians must make sure that all other possible causes of anovulation and amenorrhea have been excluded before recommending psychological, psychiatric, or behavioral interventions.

REPRODUCTION

Pregnancy, parturition (childbirth or delivery), and the postpartum period involve a series of tightly orchestrated hormonal events that create the potential for unique psychological states. As pregnancy advances, estrogen and progesterone concentrations rise and a host of placentally elaborated polypeptide hormones induce metabolic adaptations. Parturition is an intense biological event that eventuates in the abrupt withdrawal of estrogens, progesterone, and other placental hormones. Parturition is also

accompanied by a large-magnitude release of oxytocin and initiation of the neuroendocrine cascade that establishes lactation. The metabolic demands of pregnancy, parturition, and lactation, the need to deliver infant care across the usual circadian rest–activity cycle, and the hypoestrogenism that accompanies lactational amenorrhea present unique physiological challenges. Menses (and ovulation) generally resume in 6 to 12 weeks if the mother declines to nurse or when lactational frequency declines (usually after 4 to 6 months) if she does nurse.

SEXUALITY AND SEXUAL DYSFUNCTION

It has been said that the most important sexual organ is the brain. This truism is meant to emphasize the essential role of desire and comfort in mediating sexual awareness and expression. Brain centers mediate libido in a general way. Other brain centers influence the selection of sexual partners and the circumstances under which sexual longings are expressed. Hormones made by the ovaries and testes, such as estrogens and androgens, bind in certain brain centers and increase sexual interest. Whether and how this sexual interest gets expressed as behavior depend on a number of factors, including general health and well-being, the availability of a partner, a sense of what is appropriate, and prioritization of this interest above other interests and obligations. One of the major aims of social and cultural institutions is to channel or direct sexual interest in a manner that minimizes negative outcomes, such as sexually transmitted diseases (STDs), while ensuring that children will be born and thrive. Sexual interest can be inhibited by any number of negative emotions, including anger, fear, worry, and dislike, and by a marked decline in health or hormones. Just like other personality factors, sexual interest, or libido, has a spectrum, with some people having more and others less. *Sexuality* is the term used to refer to a person's gender identity, feelings, sexual orientation, and attitudes, and it is distinct from the expression of sexuality and sexual behavior. The sexual response cycle and the dysfunctions are covered in a separate section in this textbook. See Chapter 21.1 on Human Sexuality.

FERTILITY AND INFERTILITY

Gametogenesis

Gametogenesis, the production of oocytes and sperm, occurs as a result of puberty. Gametogenesis is the minimum physiological requirement for fertility. In women, the ovary must have oocytes that respond appropriately to gonadotropin input with the maturation of an oocyte that is ovulated. In men, the testes must be able to manufacture normal sperm capable of penetrating the layer of granulosa cells surrounding the oocyte and then fusing with the outer oocyte cell wall, the zona pellucida. The fallopian tube

must be capable of oocyte capture, and the density of sperm in the fallopian tube must be such that there is juxtaposition of sperm and oocyte. The sperm must be able to negotiate the layers of granulosa cells surrounding the oocyte (cumulus) and fuse with the outer cell membrane of the oocyte (zona pellucida) to release the paternal DNA into the oocyte cytoplasm. Thereafter, the tube must be capable of timely transfer of the fertilized oocyte or zygote to the primed endometrium of the uterus. Implantation must occur, and normal fetal development must ensue for the pregnancy to continue. In the absence of medical intervention, the customary manner in which ovulated oocytes are fertilized via coitus, or sexual intercourse.

Table 31.1–1.
Tests in the Infertility Workup

Possible Cause	Test	Comments
Anovulation	Basal body temperature chart	Patient must do each morning
	Endometrial biopsy	Office procedure in the late luteal phase
	Serum progesterone	Single or multiple blood tests
	Urinary ovulation detection kit	Home use at midcycle to time intercourse
	Ultrasound	Visualizes ovarian follicles and their rupture
Anatomic disorder	Hysterosalpingography	X-ray done in the proliferative phase, in which a radio-opaque dye is introduced through the cervix into the uterus to define intrauterine contours and tubal patency
	Sonohysterography	Transvaginal ultrasound examination with instillation of saline into the uterine cavity to define contours
	Diagnostic laparoscopy	Views external surfaces of internal structures
	Hysteroscopy	Visualizes endometrial cavity directly
Abnormal spermatogenesis	Semen analysis	Normal value >20 million/mL, 2 mL volume, 60% motility
	Postcoital test	Midcycle timing to look at sperm–cervical mucus interaction
Immunological disorder	Antisperm antibodies	Male semen test
Azoospermia	Testicular biopsy	Determines eligibility for intracytoplasmic sperm injection
	Semen analysis for fructose	Determines if the vas deferens are patent

Adapted from Beckmann CRB, Ling FW, Barzansky BM, Bates G, Herbert W. *Obstetrics and Gynecology for Medical Students*. Baltimore, MD: Williams & Wilkins; 1992.

Infertility

Infertility is estimated to affect 10 to 20 percent of all couples. With the advent of sophisticated techniques for both diagnosing and treating the causes of infertility, many couples who might have remained childless in a

previous era or adopted a child to build a family can now have a biologically related child. Common causes of infertility include endometriosis, impaired spermatogenesis, damage to fallopian tubes or the vas deferens from infections or sterilization procedures, and ovulatory dysfunction or anovulation. The basic infertility evaluation can involve a semen analysis, hysterosalpingogram or sonohysterogram, hormonal studies, endometrial biopsy, and diagnostic laparoscopy to define possible contributing or causal factors (Table 31.1–1). Ovulation and ovulatory adequacy can be monitored by obtaining serial blood samples or performing an endometrial biopsy. In contrast, spermatogenesis is difficult to assess because the process takes 70 to 90 days to complete and the morphological features of sperm that can be assessed by a standard semen analysis do not correlate well with function (i.e., ability to fertilize a human oocyte). Diagnostic uncertainty and a seemingly endless array of treatment approaches can easily provoke tension, indecision, and disagreement in the couple with infertility.

The sense that the window of opportunity is narrow can engender panic; a woman's age correlates well with fecundity, and treatment success declines sharply after age 37 years, even when ovulation occurs regularly. Although fertility also declines with age in men, the decline is gradual until age 50 years, and fertility may persist beyond age 60 years. The offspring of older mothers have an increased risk of chromosomal trisomies, particularly trisomy 21. An increase in autosomal dominant conditions has been reported in the offspring of older fathers. Of the identifiable causes of infertility, approximately 40 percent are thought to be due to male factors, 40 percent to female factors, and 20 percent to both. Anovulation is the most common cause of infertility in women. Approximately 20 percent of men with suboptimal semen parameters have microdeletions of the long arm of the Y chromosome in a region thought to play a critical role in spermatogenesis. Approximately 10 to 20 percent of infertile couples have unexplained infertility, which means that there is no identifiable cause.

The degree to which psychogenic or lifestyle factors play a role in infertility is controversial. Couples should not be told that their infertility is related to stress unless there is documented evidence of sperm dysfunction, ovulatory dysfunction, or sexual dysfunction. However, because lifestyle variables, such as tobacco exposure, alcohol consumption, drug use, excessive exercise, weight loss or gain, and psychogenic stress, can compromise reproduction, it is worth reviewing lifestyle factors with the couple and informing them of the potential adverse consequences of these factors in a supportive manner. Psychogenic variables can interfere with coital frequency and reduce the likelihood of conception and can suppress the central hypothalamic–pituitary drive to the gonad and compromise the quality and quantity of gametes produced.

Standard interventions were once limited to restorative surgery and gonadal stimulation, but they now include a variety of assisted reproductive techniques, including in vitro fertilization with embryo transfer (IVF-ET) and intracytoplasmic sperm injection (ICSI). In assisted reproductive techniques, the ovaries are generally hyperstimulated to produce at least three or four mature follicles from which the oocytes can be retrieved surgically. Sperm and oocytes are then mixed and supported in vitro to promote fertilization. If fertilization occurs, the fertilized ovum or zygote is transferred to the recipient's uterus. The chance of implantation if three or more zygotes are transferred is generally around 50 percent, but it can be lower if fewer embryos are available or if the woman is older than age 38 years. Approximately 25 percent of implantations (as evidenced by a transiently elevated serum β-hCG concentration) are not sustained. The likelihood of a successful full-term pregnancy increases dramatically if fetal cardiac activity is observed approximately 2 weeks after missed menses (i.e., 4 weeks after the gamete or embryo transfer). The main indication for assisted reproductive procedures is failure to conceive with conventional interventions. In some instances, the reproductive impairment is so severe that conventional interventions are unlikely to offer success and an assisted reproductive procedure is recommended primarily. ICSI is used when there is failed fertilization

during a previous IVF procedure or when too few sperm exist for conventional IVF.

Gamete intrafallopian transfer (GIFT) is a form of assisted reproduction with similar rates of success in which three or four oocytes are mixed with sperm and placed into the fallopian tube during a laparoscopy. Fertilization occurs in the tube and, thus, is not observed. IVF-ET has largely supplanted this procedure because GIFT requires a laparoscopy and general anesthesia and because pregnancy results are not better than with IVF-EF.

When it is not possible to correct underlying causes, substitutions are possible. Not only are donor sperm available, so too are donor oocytes and surrogate mothers who will gestate the embryo of another couple. Embryos can be cryopreserved indefinitely and potentially transferred to a biologically unrelated recipient or to the woman herself years after menopause.

Treatments for infertility are expensive and consume much time and energy. A sense of psychological and physical invasion is common. Infertile men and women may feel angry, damaged, or guilty. Studies suggesting a link between ovarian cancer and infertility further increase anxiety and psychological turmoil. The potential for husbands and wives to respond differently to the experience of infertility and its treatments, particularly any of the assisted reproductive technologies, can overwhelm customary adaptation patterns and lead to psychological decompensation. Even successful infertility treatment poses challenges. Assisted reproductive technologies carry a 25 to 30 percent chance of twins and an almost 1 percent chance of triplets. If quadruplets or more result, the couple must grapple with the option of selective reduction. A multiple pregnancy increases fetal and maternal risks, such as congenital abnormalities and preterm labor. Furthermore, multiple gestation complicates postpartum adjustment and carries a lifelong parenting challenge. There is some evidence that children born to previously infertile parents are viewed by their parents as special. These parents may have trouble setting limits or may be overly protective.

Whether or not it is treated, infertility can produce a variety of psychological reactions. Responses depend on personality attributes, including adaptability, cultural expectations, support systems, knowledge about reproductive processes, and the attitudes of the involved clinicians. Infertility can damage self-image and impair a sense of health and integrity. Some individuals feel that they are being made to suffer for past wrongs; others blame the medical profession for being inadequate or inattentive. A sense of panic, helplessness, and loss of control may compromise self-esteem. Women and men alike are generally overwhelmed by the complexity, cost, invasiveness, and uncertainty associated with medical intervention or adoption. Almost all patients who undergo infertility treatment (Table 31.1–1) experience a loss of privacy as they cope with physicians, nurses, insurance carriers, psychologists, hospitals, and laboratory personnel. The sense of intimacy that reinforces marriage and mutuality may be challenged when sex and procreation are separated by technology (Table 31.1–2). Ultimately, infertility presents the specter of a psychological death. Thus, it is not surprising that patients with infertility report as much anguish and distress as patients newly diagnosed with acquired immunodeficiency syndrome (AIDS).

Ideally, psychological support services should be available for all individuals and couples undergoing active evaluation and intervention for infertility. Unlike the traditional medical model, in which diagnostic testing precedes and directs the course of treatment, in infertility, empirical therapy may precede more-invasive diagnostic testing. Cause and effect can be difficult to identify, and magical or superstitious thought processes are common in patients and practitioners. Infertility can threaten marital adjustment, particularly if the partners react differently to the challenges. Although fertility is an attribute of a couple, blame may be prematurely or incorrectly assigned to one person in the dyad. In addition to individual or couples counseling, groups, such as Resolve, can offer support and information. When donor sperm, oocytes, or embryos or surrogate motherhood is being considered, a thorough psychological evaluation is recommended in advance of treatment to ensure that all parties agree to the treatment plan and understand the potential implications and ramifications.

Reproductive technologies have made it possible to separate biological and social parenthood. A single woman may choose parenthood. The use of donor insemination or other fertility therapies in single women is controversial. Most women who seek such services have considered the consequences of single parenthood and feel that they will be able to rise to the challenges. Nonetheless, psychological evaluation should be made available to ensure that important topics and consequences have been considered. No long-term follow-up on the mental health and development of children born to single mothers has been done.

Dysfunctional attitudes and disharmony constantly threaten the infertile person. Prophylactic evaluation is preferred in all treatment settings, but specialized services may not always be immediately available. Stress can compromise gametogenesis and libido, and infertility treatment alone can be stressful enough to activate the central mechanisms that compromise reproductive function. Psychological support is intended to lessen the likelihood of this effect and reverse it if it is already occurring.

Pharmacological agents used to treat disorders of ovulation or to hyperstimulate the ovaries may alter mood and cognition. The agents include clomiphene citrate (Clomid) and other antiestrogens, GnRH analogs, human gonadotropins, hCG, progesterone, and bromocriptine (Parlodel). Danazol (Danocrine), a synthetic androgen, may be used to arrest the growth of endometriosis as a surgical or medical adjunct to later infertility therapy. Treatment of infertility itself is anxiety provoking because of the uncertainty and sense of expectation. However, most of the agents do alter estradiol and progesterone levels, and the resultant hormonal fluctuations may trigger mood lability and depression in sensitive individuals. Agents such as clomiphene, danazol, and bromocriptine may also have direct effects on the brain. If the use of these agents is indicated, psychological support should be offered instead of anxiolytics and antidepressants. A further concern of infertile couples is whether assisted reproductive techniques increase the risk of neurodevelopmental or other abnormalities, including childhood cancers. The uncertainty as to whether there is a link between infertility treatments and poor health outcomes in the mother or child only serves to exacerbate anxiety and guilt. Infertility is a far more devastating disorder than most laypersons and physicians realize.

Table 31.1–2.
Assisted Reproduction Techniques

Method	Comments
Ovulation induction or augmentation (multiple agents, particularly recombinant or highly purified gonadotropins)	Stimulates multifollicular development and ovulation; may produce multiple births; used in anovulation, luteal phase deficiency, unexplained infertility, and assisted reproduction
Induction of spermatogenesis	Used in men with idiopathic hypothalamic hypogonadism of functional or organic nature
Artificial insemination	Donor sperm is injected into the uterine cavity or the fallopian tubes; the sperm of the male partner may be used if healthy
Gamete intrafallopian transfer	Transfer of collected oocytes and sperm into the fallopian tubes; zygote may also be transferred; used for infertility from endometriosis and unexplained infertility
In vitro fertilization and embryo transfer	Transfer of developing embryos into the uterus after extracorporeal incubation of collected sperm with oocytes retrieved by laparoscopic surgery or by ultrasound-guided transvaginal aspiration; used with occlusion of the fallopian tubes or significant sperm dysfunction; permits

	preimplantation genetic diagnosis
Intracytoplasmic sperm injection	Injection in vitro of sperm head or sperm DNA to cause fertilization and production of embryos for transfer to receptive endometrium; may be used even if sperm are barely viable; genetic causes of male infertility may be transmitted to offspring
Gamete donors	Donation of sperm or oocytes to another couple; can include oocytes only, oocyte cytoplasm only to restore reproductive competence to aged oocyte, donor sperm alone, or any combination; cytoplasmic transfer considered unethical and is illegal in some states
Surrogate mother	Surrogate mother receives donated embryo and carries the baby to term; a highly controversial technique with unclear legal ramifications

Data are taken in part from Susman V. *Pregnancy in Behavioral Sciences for Medical Students*. Baltimore, MD: Williams & Wilkins; 1993.

Pregnancy Loss

Unlike infertility, which can continue indefinitely, pregnancy loss is a defined event with an end. Acute reactions to an isolated miscarriage or perinatal loss, such as blame, guilt, anger, and denial, gradually lead to acceptance and resolution. A single miscarriage does not predict poor future reproductive performance. *Recurrent pregnancy loss,* defined as three or more losses before 28 weeks of gestation, carries an unfavorable prognosis and resembles infertility in its chronicity and sense of lost potential. Although it might be assumed that grief and emotional turmoil would be related to gestational age, this assumption is not supported by observations. Women who miscarry have more than twice the relative risk of a major depression compared with community women. More than half of women with a history of major depression experienced a recurrence after miscarriage. Because of the significant risk of depression, acute intervention is recommended for all women and couples who experience a miscarriage or perinatal loss; services should be extended beyond the acute event when indicated.

Pseudocyesis

Pseudocyesis is the development of the classic symptoms of pregnancy—amenorrhea, nausea, breast enlargement and pigmentation, abdominal distention, and labor pains—in a nonpregnant woman. Pseudocyesis demonstrates the ability of the psyche to dominate the soma, probably via central input at the level of the hypothalamus. Predisposing psychological processes are thought to include a pathological wish for and fear of pregnancy; ambivalence or conflict regarding gender, sexuality, or childbearing; and a grief reaction to lost potential after a miscarriage, tubal ligation, or hysterectomy. The patient may have a true somatic delusion that is not subject to reality testing, but, often, a negative pregnancy test result or pelvic ultrasound scan leads to resolution. Psychotherapy is recommended during or after a presentation of pseudocyesis to evaluate

and treat the underlying psychological dysfunction. A related event, couvade, in which the father of the child undergoes a simulated labor as if he were giving birth, occurs in some cultures.

Table 31.1–3.
Types of Abortion

Spontaneous	Spontaneous expulsion of the products of conception before viability (500 g or approximately 24 wks from last menses)
Recurrent	Three or more spontaneous abortions
Missed	Abnormal development of an intrauterine pregnancy; usually caused by the presence of a blighted ovum and lack of fetal development
Threatened	Uterine bleeding or cramping and positive pregnancy test; must be distinguished from an ectopic (usually tubal) pregnancy
Incomplete	Spontaneous passage of a portion of the products of conception and the retention of placental fragments that result in ongoing bleeding
Elective	Induced by medical or surgical techniques before fetal viability; techniques include dilation, evacuation, and curettage; suction curettage; injection into the amniotic sac of saline or prostaglandins; hysterotomy; prostaglandins with antiprogestins (RU-486) or methotrexate; medical indications include the detection of fetal abnormalities by ultrasound or amniocentesis

Fertility Control

Elective Abortion. Elective abortion is distinct from spontaneous abortion (i.e., miscarriage) (Table 31.1–3). Elective abortion is the planned termination of a pregnancy. The Centers for Disease Control and Prevention (CDC) statistics indicate that approximately 840,000 abortions were performed in the United States in 2004—239 abortions for every 1,000 live births. As in the past, in the United States, a disproportionately higher number of abortions were obtained by white women, unmarried women, and women younger than age 25 years. In Western nations, most women who obtain abortions are young, unmarried, and primiparous; in underdeveloped nations, abortion is most common among married women with two or more children.

CDC statistics from 2004 show that 88 percent of legally induced abortions were performed at or before 12 weeks of gestation. The remainder occurred after 13 weeks, with 1.5 percent occurring after 21 weeks. Table 31.1–4 summarizes the most common abortion techniques. In contrast to what was previously thought, recent studies show no significant untoward psychological effects of elective abortion on the mental health of women. In fact, in one study comparing the outcomes of women who sought treatment in a clinic because they thought they were pregnant, women who were pregnant and chose to abort were reported later to be doing better emotionally and in pursuing personal, educational, and work endeavors than were either those who were not pregnant or those who chose to continue the pregnancy. Another study suggested untoward consequences on the mental health and well-being of women who were denied the option to abort and, thus, were required unwillingly to carry the pregnancy to term and then raise an unwanted child. This study suggested untoward consequences on the mental health and well-being of the child as well. However, termination of a wanted pregnancy because of an abnormal karyotype or fetal anomalies can be traumatic, and supportive intervention is recommended.

Table 31.1–4. Abortion Techniques

Type	Benefits	Risks
Cervical dilation and evacuation of uterine contents by curettage or vacuum aspiration	Most commonly performed procedure for termination of pregnancy; can be done before 24 wks' gestation	Uterine perforation, adhesions, hemorrhage, infection, incomplete removal of fetus and placenta (all rare)
Menstrual aspiration (miniabortion)	Can be done within 1–3 wks of missed period	Implanted zygote not removed, uterine perforation (rare), failure to recognize ectopic pregnancy
Medical induction (cervical dilation with laminaria followed by intravaginal prostaglandin or intravenous oxytocin)	Can be used for second trimester abortion	Water intoxication, uterine rupture, infection
Intra-amniotic hyperosmotic solutions (salting out)	Can be used for second trimester abortions	Hyperosmolar crisis, heart failure, peritonitis, hemorrhage, water intoxication, myometrial necrosis
Prostaglandin (oral, intravaginal, cervical, or intra-amniotic)	Noninvasive procedure; can be used in conjunction with antiprogestins (RU-486) or methotrexate	Expulsion of live fetus, missed ectopic, hemorrhage, incomplete abortion
Antiprogestins (RU-486) with or without concomitant prostaglandin use	Nonsurgical; first trimester only	Incomplete abortion, hemorrhage
Methotrexate with or without prostaglandin use	Nonsurgical; first trimester only	Leukopenia, hemorrhage, incomplete results

Techniques for inducing abortion (Table 31.1–4) have been known for at least 5,000 years. In view of the risks of childbearing and the requirements of child rearing, induced abortion was viewed historically as a means of modulating the maternal and societal risks of unrestrained childbearing. Induced abortion affords women a measure of autonomy over a unique and psychologically powerful bodily function, but it arouses personal and public ambivalence and inspires controversy.

Historically, the concept of therapeutic abortion before *quickening*, which is generally defined as the time at which the mother perceives fetal movements, was well accepted. In the United States, legislative regulations regarding induced abortion were introduced in the early 1800s, but their intent was to protect women from sepsis and other complications associated with the procedures. Induced abortion, whether by surgical or chemical means, now entails minimal medical risk and is clearly safer than carrying the pregnancy to term.

Current legislative attention has shifted from safeguarding maternal health to the ethics of abortion and the right of women to exercise control over their bodily functions. Traditionally, ethical deliberations regarding abortion were reserved for religions. Because of the diversity of religious perspectives on the role of women and on procreation, consensus is neither likely nor necessarily desirable. Nonetheless, the current legislative debate centers on who has the authority to regulate induced abortion and to what degree. The 1992 Supreme Court ruling in *Planned Parenthood of Southeastern Pennsylvania v. Casey* upheld the basic right of women to elect pregnancy termination before fetal viability, but it also upheld the right of the state to regulate and restrict abortion and to define viability. Furthermore, in some states (e.g., Pennsylvania), pregnant women do not have the right to select whether they wish to receive advanced life support in life-threatening situations. The Freedom of Choice Act put before the US

Congress recognized the right of states to regulate the strictly medical aspects of abortion to safeguard maternal health. It restricts the freedom of a woman to terminate a pregnancy after, but not before, fetal viability. This legislative effort represented a compromise between those who argued against any medical mediation of the process and those who would grant women complete autonomy. Application of this law would hinge on the interpretation of the phrase *fetal viability,* a concept that is not easily defined. Making the option of induced abortion contingent on fetal viability reveals continuing ambivalence about allowing women the autonomy to make decisions about physiological processes that directly shape their lives. It is difficult to appreciate how a legislative body can entrust the life of a child to the mother and simultaneously limit the mother's autonomy in making reproductive decisions. Current legislative efforts have focused on restricting the techniques used to perform late-term abortions by criminalizing physicians. This debate regarding abortion is far from resolution and is likely to continue as medical advances such as mifepristone (RU-486), the antiprogestin that can chemically induce abortion in the early first trimester, increase therapeutic options.

The medical risks of either surgical or medical abortion are few, but both procedures require physician input. Medical abortion, however, requires less physician input, although it does require careful monitoring to ensure that the abortion occurs in a timely fashion. Rarely, hemorrhage occurs that requires surgical intervention or a transfusion. Medical abortion can potentially increase a woman's autonomy of her reproductive capacity. Studies indicate that women who chose to terminate their pregnancies do well in the long term in relation to physical and mental health and educational, occupational, and marital functioning compared with women who continued their pregnancies or were never pregnant. There is no substantial evidence that pregnancy termination contributes to major psychiatric illness.

Contraception. Methods of contraception fall into four main categories: Barrier methods, hormonal agents, intrauterine devices, and natural family-planning methods that require abstinence around the time of ovulation (Table 31.1–5). The perfect contraceptive would be completely reversible, free of adverse effects, and completely efficacious. Because no perfect contraceptive is available, people must select a method that best meets their needs. Frequently, the responsibility for choosing and using a contraceptive falls primarily on the woman. This can lead to resentment, lack of compliance, and marital disharmony. For those in a stable relationship, contraceptive counseling can include both partners.

The assessment of the risks and benefits of a given contraceptive strategy should not be limited to a consideration of the biological risks alone but must take into account the mindset of each party, particularly with regard to gender issues and cultural perspectives. If significant discord exists within a couple, counseling may be indicated. Counseling can also be considered for the person or couple who feels guilty about the idea of interfering with natural bodily functions for this purpose.

Clinicians must be mindful of the difference between actual and perceived risk. Concerns of current and later bodily harm from the use of contraception are common, and the risks of childbearing and child rearing frequently are underestimated. The mortality associated with

childbearing is estimated to be 10 per 100,000 live births between the ages of 20 and 24 years, 44 per 100,000 between the ages of 35 and 39 years, and 71 per 100,000 after age 40 years. However, comparable mortality in nonsmoking women on oral contraceptives is 1.2 per 100,000 live births, 3.9 per 100,000 live births, and 6.6 per 100,000 live births, respectively. The potential risks of child rearing are difficult to quantify, but they include role strain and stress, interrupted education or career advancement, depression, loss of independence, and compromise of the psychosexual relationship with the partner. This misperception of risk is particularly prevalent among adolescents, who may be uninformed and have unrealistic ideas about the rewards of pregnancy and the risks of STDs. In addition, a teenager's aptitude for listening and making decisions may be limited by ambivalence and confusion. In short, despite the availability and relative affordability of contraceptive means, their use is constrained by the discrepancy between actual and imagined risk. Because coitus is not always anticipated, there is a need for postcoital or so-called emergency contraceptives. Available methods include the insertion of an intrauterine device (IUD) and the use of high doses of levonorgestrel, high doses of estrogen–progestin combinations, and antiprogestins. The use of levonorgestrel, 0.75 mg taken twice 12 hours apart (Plan B), within 72 hours of intercourse was more effective (greater than 99 percent) and was associated with less severe adverse effects than the use of a high dose of combined oral contraceptives (Preven). In the United States, Plan B was recently approved to be sold over the counter to women who are 18 years or older.

The mental and physical health risks of contraceptive use are significantly lower than are those associated with pregnancy, parturition, and parenthood. However, contraception often gives an unwarranted illusion of control over a bodily process. No method is perfect, and unintended pregnancy can occur despite responsible use. In addition, if a couple or person intends to have children, childbearing should not be postponed indefinitely in the hope that the perfect time will arrive. It can be difficult to conceive on demand.

Sterilization. Sterilization can be accomplished by several methods of tubal ligation—by hysterectomy with or without oophorectomy in women and by vasectomy in men. Hysterectomy for sterilization alone is associated with more morbidity and mortality than any method of tubal ligation; vasectomy is safer than any of the present methods of tubal ligation. Most studies do not support the contention that tubal ligation carries an increased incidence of gynecological sequelae, but failure rates are high in the first year after tubal ligation. In general, it is safest to perform a tubal ligation by laparoscopy at least several weeks after delivery. Tubal ligations immediately after delivery are done because of convenience or because of a woman's unwillingness or inability to return for a later tubal ligation.

Table 31.1–5.
Current Methods of Contraception

Type	Efficacy[a]	Advantages	Disadvantages	Potential Complications[b]
Barrier (chemical or mechanical)				
Spermicidal agents	Moderate	Readily available and easy to use	Messy; loss of spontaneity; requires forethought	Allergic reactions

Diaphragm, cervical cap	Moderate	Inexpensive; does not interfere with menstrual cycle	User familiarity required; prescription and fitting required; may interfere with spontaneity; requires forethought	Recurrent urinary tract infections with diaphragm; allergic reactions to latex or spermicide
Male condom	Moderate	Readily available and easy to use; protects against sexually transmitted infections	May interfere with spontaneity; requires forethought	Allergic reactions to latex or spermicide
Female condom	Moderate	Inconvenient to use; protects against sexually transmitted infections	May interfere with spontaneity; requires forethought	Allergic reactions to latex or spermicide
Hormonal (suppresses ovulation and/or impairs endometrial development)				
Oral contraceptives, contraceptive patches, and intravaginal rings	High	Protect against uterine and ovarian cancer, some sexually transmitted infections	Pharmacological side effects; require daily or weekly use, regardless of frequency of sex; must be prescribed and monitored by health care professional	Depression; breast tenderness; nausea; headaches; may be contraindicated in some medical conditions; can be taken continuously to achieve greater ovarian suppression and amenorrhea and to treat medical conditions
Postcoital steroids	High	Can be used after intercourse; inexpensive	Must be initiated within 72 hrs; requires some medical supervision	Side effects, particularly nausea and headache; not universally available
Contraceptive implants (rods)	High	Implantable device that provides contraception for up to 1 yr; once in place, no forethought required	Irregular bleeding and spotting due to endometrial effects and suppression of ovarian function; must be surgically placed and removed	Bone loss, depression; other medical sequelae of associated hormonal alterations unclear
Injectable steroids	High	Injectable progestin or combination of estrogen and progestin that	Slow return of ovarian activity after last dose; not removable	Bone loss and depression greater with progestin-only injections due to induction of greater

		prevents ovulation and suppresses ovarian activity; given intramuscularly at an interval that depends on the product		hypoestrogenism; other consequences of hypoestrogenism unclear
Antiprogestins (RU-486)	High	Easy to use; does not disrupt menstrual cycle when administered in luteal phase; can be used postcoitally	Currently not available in the United States; must have predictable cycles	Impairs implantation rather than preventing conception
Sterilization				
Male sterilization (vasectomy)	High	Failure very rare; 20-min office procedure	Morbidity in 1–2% of patients includes infections, clots	Can be reversed in only 80% of cases; rare neurotic impotence reaction; used by 10.4% of men
Female sterilization	High	Almost 100% protection; no impairment of sexual function or pleasure	More complex procedure than vasectomy; reversal is complicated and difficult	Surgical morbidity; used by 13.6% of women
Other				
Intrauterine device	High	Once in place, no forethought required; impairs endometrial receptivity	May cause nonbacterial endometritis or heavy menses; requires professional insertion	Likelihood of pelvic infection may be increased and tubal damage with infertility may result; uterine perforation or spontaneous expulsion
Rhythm	Low	No cost	Imposed coital timing	None
Natural family planning	Moderate	Readily available	User must monitor cervical mucus and body temperature closely; imposed coital timing	None
Withdrawal, coitus interruptus	Low	Readily available	Difficult to implement	None

[a]Estimates of efficacy are those associated with use rather than theoretically derived. High, <5% chance of contraceptive failure during the first year of use; moderate, <20% chance; low, >20% chance.

[b]Other than failure (pregnancy).

It is difficult to determine how often reversal of a tubal ligation or

vasectomy is sought or performed, and it is even more difficult to estimate the frequency of regret that the sterilization was performed. No formula can predict who will seek reversal, but factors such as remarriage, young age at time of sterilization, death of a child, and performance of the sterilization during the postpartum period seem to play roles. Regret is high when sterilization is undertaken as a way to stabilize a marriage. The success of reversal depends on a variety of factors, including the method of sterilization, the ages of the partners, and the elapsed time since the sterilization procedure. Clinicians need to make couples and individuals aware of these considerations but should not use them as reasons to deny the planned procedure. Again, when a couple is involved, the mutuality of the decision should be discussed to avoid later resentment. Voluntary sterilization, especially vasectomy, has become the most popular form of birth control in couples married for more than 10 years.

SEXUALLY TRANSMITTED DISEASES

An STD is a contagious disease acquired as a result of a physical sexual interaction. STDs have always been a reality, but from the 1950s through 1970s the infections were considered treatable and not life threatening. AIDS, which is caused by infection with human immunodeficiency virus (HIV), is currently incurable, life threatening, and transmissible from mother to fetus. The specter of AIDS has captured the popular imagination. Although it was initially found in male homosexuals and intravenous drug abusers, HIV infection knows no boundaries. Sexual monogamy or abstinence has been advocated to ensure emotional health and a stable family arrangement for the potential outcome of sexual intercourse—children—but mutual sexual monogamy or abstinence also prevents STDs, particularly AIDS.

Another sequela of STDs such as gonorrhea and chlamydia is pelvic inflammatory disease (PID). Untreated PID can develop in bilateral tubo-ovarian abscesses and necessitate hysterectomy and bilateral salpingo-oophorectomy. Early antibiotic treatment is advocated to prevent the development of the abscesses and to reduce the likelihood of infertility, chronic pelvic pain, and ectopic pregnancy from tubal damage. These infections can also lead to obstruction of the vas deferens and chronic prostatitis in men and subsequent male infertility.

Another STD that can have serious consequences is venereal warts, or human papilloma virus (HPV) infection. Genital infections with certain subtypes of HPV can lead to premalignant changes of the penis, vulva, vagina, and cervix and are thought to cause cervical cancer. Venereal warts can be removed chemically or surgically but are difficult to eradicate completely. Women who contract HPV are encouraged to have regular gynecological examinations and Papanicolaou smears to detect premalignant lesions. Recently, a vaccine against HPV has been developed and its use is being advocated as prophylaxis, especially in young women, before the onset of sexual activity. One could make the case that both young men and women should be vaccinated, but initial attention has focused on raising awareness of the vaccine's availability among young women and their parents.

Sexual monogamy and abstinence prevent most STDs and are advocated as public health measures. However, libidinal impulses can be difficult to control and restrict. Therefore, measures such as condom use are recommended as an alternative public health measure. The ideal can be defined, but clinicians and policy makers need to remember that it cannot always be

attained. Adolescents, in particular, need to know the potential consequences of sexual activity with regard to STDs and pregnancy. Admonishing teens to remain chaste is unlikely to be completely effective and may be counterproductive. To make responsible decisions, people need to understand the risks.

Clinicians need to remember that the risks of sexual intercourse may be forgotten or seem minimal in comparison to the need for affection, contact, release, or escape. Persons who have low self-esteem or are stressed may view sex as a means of bolstering their self-image or escaping their stresses. The reinforcing properties of sex ensure that the problem of STDs will endure.

PELVIC PAIN

Pelvic pain can have many causes, including endometriosis, pelvic adhesions, ovarian or adnexal masses, hernias, and bowel or rectal disease. Pelvic pain can also be secondary to psychogenic causes, such as guilt, fertility or infertility fears, and the emotional disturbances associated with ongoing or past incest or sexual abuse. Pelvic pain should not be attributed to psychogenic causes unless thorough evaluation has excluded organic causes. In most instances, the evaluation should include a diagnostic laparoscopy. Likewise, dyspareunia—pain with intercourse—should not be assumed to have a psychogenic origin unless all anatomical causes have been excluded.

PERINATAL PSYCHIATRIC DISORDERS

Postpartum Blues

The postpartum blues or baby blues is the most common constellation of mood symptoms experienced by women (occurring in 50 to 80 percent of women) in the immediate postpartum period. The postpartum blues include transient symptoms and rapid mood shifts, including tearfulness, irritability, anxiety, insomnia, lack of energy, loss of appetite, and the general experience of feeling overwhelmed—particularly with regard to newborn caregiving tasks. By definition, the postpartum blues are transient in nature. Onset typically occurs after the third postpartum day, after the mother has left the hospital after delivery. Symptoms typically peak by day 5 and spontaneously resolve by day 10 postpartum. It has been estimated that 20 to 25 percent may of women who experience symptoms of postpartum blues will go on to experience postpartum major depression. Treatment of the postpartum blues should include psychoeducation, validation of the mother's experience, and careful monitoring for a worsening or prolongation of symptoms.

Perinatal Depression

As perinatal mental health has gained increasing public health and scientific attention, there is greater appreciation that a multitude of affective disorders (not just postpartum depression [PPD]) emerge both during pregnancy and postpartum. A change in conceptualization of PPD occurred with the revision of the time specifier in DSM-5, published 2013,

in which “postpartum onset” was changed to “perinatal onset.” “Perinatal onset” recognizes new symptom onset anytime in pregnancy through 1-month postpartum, based on data that many postpartum affective episodes either have their onset in, or predate pregnancy.

Although pregnancy was once considered a time of mental wellness, minor or major depression affects up to 25 percent of pregnant women, a similar rate as nonpregnant women. Postpartum depressive syndromes affect 15 percent of mothers, with rates that triple among inner-city and adolescent mothers. Completed suicide is a leading cause of death in new mothers, and infanticide is another devastating outcome which has brought this disorder frequent media attention. Several studies examined the course of unipolar depression in pregnancy. In a specialized perinatal psychiatry clinic the rate of depression relapse was 68 percent in women who discontinued antidepressants prior to pregnancy, compared to 26 percent in women who maintained their medication during pregnancy. In contrast, in a community sample with less severe illness, only 16 percent of the cohort stopping their medication in early pregnancy experienced a recurrence. In both studies, 50 to 75 percent of relapses occurred in the first trimester, which argues against delaying medication until second trimester. Also, both studies reported that duration of the depressive episode, a history of more recurrences, and a major depressive episode in the 6 months prior to pregnancy were associated with increased likelihood of pregnancy relapse.

While psychosocial adjustment to motherhood lasts longer than 1 month, the DSM-5 time specifier ends at 1-month postpartum to denote a syndrome with greatest biological link to the hormonal shifts of pregnancy and delivery. Heightened biological sensitivity for PPD at 1 month comes from familial studies, wherein PPD onset within 30 days was concordant among parous sisters and postpartum mood symptoms within 30 days was associated with strong genome-wide linkage signals. From linkage studies of large obstetric and psychiatric databases based upon hospitalization and outpatient contacts, there was a 2- to 3-fold increased risk of PPD in primarous women that was evenly distributed across the first 5 months postpartum, with risk remaining elevated out to 2 years postpartum in one study.

Postpartum episodes are commonly characterized by obsessive-compulsive, particularly intrusive violent thoughts, psychotic symptoms, and anxiety. Risk factors for PPD include a personal or family history of depression, as well as a personal history of previous depressions related to reproductive events (such as PMDD or previous PPDs). As with non-PPDs, women facing multiple or severe psychosocial problems, trauma, or chronic interpersonal difficulties are at increased risk of experiencing a major depressive episode during the postpartum period. One study revealed a 50 percent rate of childhood sexual abuse history in half of women hospitalized for PPD.

Perinatal Anxiety

Anxiety disorders are the most prevalent of psychiatric disorders, and women, compared to men, have greater overall lifetime prevalence of anxiety disorders, as well as increased peak prevalence during the childbearing years. Thus, it is important for both gynecologists and primary care practitioners to have a basic familiarity with the anxiety disorders to promote early detection and treatment. Postpartum anxiety disorders range in incidence from 8.5 to 13 percent with increases to 20 percent when subsyndromal forms of anxiety are included. At a prevalence of 8.2 percent,

generalized anxiety disorder (GAD) is more common in postpartum women relative to the general population. Rates of anxiety comorbidity in women with perinatal depression reach 50 percent. Medical diagnoses such as hypo/hyperthyroidism, anemia, and hypertension have a similar phenotype and should be ruled out.

Prenatal stress and pregnancy-specific stress symptoms are associated with increased glucocorticoid concentrations in the mother–fetus dyad, pregnancy complications, earlier timing of delivery, prematurity, and low birth weight. Risk factors in pregnancy include young adulthood, single motherhood, pregnancy complications, stressful life events, poor physical health, and past year victimization. Assertive screening in pregnancy is therefore warranted and affective symptoms shouldn't be assumed to be normative responses to pregnancy. Both antenatal anxiety and depression are highly predictive of PPD, which is important in framing the decision-making process for treatment choices in pregnancy.

Subsyndromal obsessive-compulsive disorder (OCD) symptoms may reach 25 percent prevalence in perinatal women, with 3 to 9 percent prevalence rates of the disorder itself, higher than general population OCD prevalence of 2 to 3 percent. Such symptoms may be an exaggeration of adaptive maternal behavior. Most common OCD manifestations in postpartum women include aggressive obsessions, contamination obsessions, and cleaning and checking compulsions. OCD and depression in postpartum women are highly comorbid (40 percent) and subsyndromal aggressive obsessions are significantly more common in postpartum compared with nonpostpartum women. Ego-dystonic aggressive obsessions toward the infant, usually regarded as highly distressing by new mothers, must be distinguished from ego-syntonic aggressive ideation which may accompany depression and psychosis (see below).

Posttraumatic experiences are reported in 20 to 33 percent of new mothers, while the prevalence of PTSD ranges from 1.7 to 6 percent. These symptoms can be categorized into context of delivery complications, reactions to miscarriage, and reactions linked to previous sexual trauma. The importance of treatment stems from difficulties in mother–infant bonding or even avoidance of contact with infant in this disorder.

Panic disorder has similar prevalence in pregnancy compared to the general population of approximately 1 to 3 percent. There may be an increased prevalence of new onset panic disorder in the first 3 months postpartum based upon a retrospective study, which awaits confirmation in a prospective study design. There is debate about the course of panic disorder symptoms in women with pre-existing illness, with no-change being the predominant finding in pregnancy with a possible increase in the postpartum period.

Perinatal Depression and Anxiety Treatment

In the absence of specialized perinatal treatment guidelines, the APA guidelines apply generally. The Center for Disease Control "Treating for Two" initiative draws attention to the fact that in some cases, avoiding or stopping medication use during pregnancy may be more harmful than taking a medication. The use of medication management is increasingly common for women with moderate to severe psychiatric disorders, for whom removal of treatment would increase the likelihood of neurovegetative symptoms, incapacitating anxiety, functional impairment, drug and alcohol abuse, suicidality, or psychosis. For new onset of depression, an SSRI may be tried initially because such agents are associated with a low risk of toxic effects in patients taking an overdose, as well as with ease of administration. However, if the patient has previously had a positive response to a specific drug from any class of antidepressants, that agent should be strongly considered. Women may be more likely to

have a response to serotonergic agents, such as the SSRIs and venlafaxine, than to nonserotonergic tricyclic antidepressants (TCAs), yet safety and efficacy data for TCAs such as nortriptyline and desipramine are generally good, given that they have been in use for a relatively long period of time.

Slow increases in the dose are helpful in managing side effects. Women who have given birth recently often are sensitive to medications side effects, possibly related to anxious, agitated features early in the postpartum course. Treatment should be initiated at half of the recommended starting dose (e.g., 25 mg of sertraline per day or 10 mg of paroxetine per day) for 4 days, and doses should be increased by small increments (e.g., 25 mg of sertraline per week or 10 mg of paroxetine per week) as tolerated until full remission is achieved. If the patient has a response to an initial trial of medication lasting 6 to 8 weeks, the same dose should be continued for a minimum of 6 months after a full remission has been achieved to prevent a relapse. If there is no improvement after 6 weeks of drug therapy or if the patient has a response but then has a relapse, a medication change (within-class or different class) is recommended. Drug doses may need to be increased in the third trimester during the period of increased hepatic metabolism and expansion of blood volume. Antidepressant trials are often initiated in primary care and OB-GYN offices, and consultation with psychiatrists can facilitate optimal dose titration to achieve remission, addressing treatment resistance, and treating psychiatric comorbidities. Long-term maintenance treatment for the prevention of recurrence should be considered for women who have had three or more episodes of severe depression.

A few placebo-controlled trials for PPD have been conducted. The SSRI fluoxetine was compared with CBT, combination treatment, and placebo. Fluoxetine and CBT were similarly effective and fluoxetine was significantly more effective than placebo. Sertraline failed to show a significant difference over placebo as an add-on treatment to brief dynamic psychotherapy. In a pragmatic trial of SSI versus supportive care listening visits, the SSRI proved superior by 4 weeks of treatment. Paroxetine was associated with greater rates of remission compared to placebo in an 8-week trial. For PPD with comorbid anxiety, both antidepressant and CBT were efficacious; however there was no added benefit of combining modalities. In small, open trials, sertraline, venlafaxine, escitalopram, paroxetine, and bupropion SR were effective for the treatment of PPD. In studies of the prevention of PPD, the TCA nortriptyline did not confer protection compared with placebo, but a pilot randomized trial of sertraline given immediately after birth for 17 weeks had greater benefit than placebo (93 percent vs. 50 percent remained healthy, respectively) in conferring preventive efficacy.

PPD treatment or prophylaxis with the reproductive hormones estrogen and progesterone has been examined as a means of compensating for the abrupt hormonal withdrawal that accompanies delivery. In an RCT of transdermal 17β-estradiol 200 mg per day versus placebo, the estradiol-treated group had a significant reduction in depression scores beginning 4 weeks after treatment initiation that was maintained at 3 months. An open trial of sublingual estradiol (mean dose of 4.8 mg/day) in women with severe depression resulted in rapid response within 1 week and an 83 percent rate of remission by 2 weeks of treatment. An RCT of the intramuscular depot-progestogen norethisterone enanthate 200 mg at 48 hours postdelivery, resulted in higher depression scores compared to placebo, perhaps mediated by lower estradiol concentrations in the progestogen group. Smaller open trials of prophylactic progesterone have been promising but merit replication with randomized, controlled study methods. Complementary treatments for PPD follow similar guidelines as for antenatal depression (see above).

Psychosocial interventions for individuals and groups, delivered in home, clinic, or by phone, compared to usual care, are associated with symptom reduction, reduced psychosocial stress, increased mental wellness, and prevention of PPD in an increasing number of studies. Promising modalities include brief interpersonal psychotherapy, cognitive-behavioral therapy, computerized resources, in-home nurse visits, and peer-based telephone support. Complementary and alternative medicine therapies suggest potential benefits or adjunctive use. Five-week morning 7,000 lx bright light therapy was associated with a significant improvement in depressive symptoms in an RCT comparing it to red, dim light. Fish oil administered in a dose of 1 g, 50 percent eicosapentaenoic acid (EPA) and 50 percent docosahexaenoic acid (DHA), is frequently recommended based upon efficacy in nonperinatal populations. Prenatal education attendance, exercise, yoga, tai chi, social/peer support, self-guided workbook with weekly

telephone support, relaxation therapies, mindfulness, acupuncture, massage in a variety of studies with variable rigor have shown promise for reducing antenatal anxiety.

There is a dearth of study of postpartum anxiety treatment studies, although given comorbidity of depression and anxiety and overlap of treatments for both, it is reasonable to offer behavioral treatments as described above for depressive syndromes. There was no added benefit of 12-group CBT to paroxetine in women with comorbid depression and anxiety. Notably, prevention of postpartum anxiety has been successful in prenatal educational formats to deliver CBT for postpartum OCD symptom reduction as well as parenting communication for reducing postpartum generalized anxiety and improving overall postpartum adjustment and marital satisfaction.

Standard guidelines for the treatment of GAD, PTSD, OCD, and panic disorder should inform the treatment of postpartum women with these disorders, using specific psychotherapeutic interventions when available. Medication management described in standard guidelines includes SSRIs as a first-line treatment; however there is also considerable evidence for the benefit of TCAs, such as clomipramine for OCD and imipramine for panic disorder. Nightmares of PTSD can be effectively treated with the alpha-1 adrenergic receptor antagonist prazosin, and could benefit this population. The use of benzodiazepines to promote sleep and/or treat residual anxiety symptoms not addressed by the primary behavioral or antidepressant medication is a common clinical practice.

Postpartum Psychosis

The most severe of the perinatal mood disorders, postpartum psychosis is rare, affecting approximately 0.1 to 0.2 percent (1 in 1,000 to 1 in 500) of postpartum women. This disorder escalates rapidly to florid psychosis within days postpartum and often results in presentation for treatment within 3 weeks postpartum. Postpartum psychosis is commonly associated with the diagnosis of bipolar disorder, unipolar psychotic depression, or a family history of affective disorder. There remains a large percentage of de novo postpartum psychosis cases for whom a brief psychotic (formerly called cycloid psychosis) clinical picture is evident, characterized by disorganized behavior and rapid shifts from excited to inhibited motoric states, from confusion to lucidity, and from anxiety/paranoia to happiness. Auditory or visual hallucinations, paranoid or grandiose delusions, elements of delirium or disorientation are common, as are extreme deficits in judgment and high levels of impulsivity. This brief psychotic phenotype is distinct from a mixed or cycling bipolar affective phenotype; however in one study that examined brief psychosis cases 12 years later, the clinical phenotype shifted to one of bipolar disorder in 16 to 20 percent of cases. During follow-up in this study, there was a striking lack of psychotic

recurrence in nearly 30 percent of women with an index episode of postpartum psychosis; furthermore these women were largely unmedicated at the time of follow-up.

Postpartum psychosis is a medical emergency because of the potential risk for infanticide or suicide and therefore imminent referral to a psychiatrist, hospitalization, or initiation of antipsychotic medication is indicated. Postpartum psychosis was associated with a 70-fold increased risk of suicide in the first 12-month postpartum with peak rates within a few days to a few weeks after discharge. Although being a mother is protective against suicide, this narrow time interval postpartum for very ill women constitutes an exception and warrants greater vigilance. Infanticide risk is also greatly increased in this patient population. In a classification of maternal filicides, 24 percent occurred within the context of acute maternal psychosis and another 56 percent due to altruistic beliefs about alleviating suffering in the infant, which was often accompanied by maternal suicide.

If recognized in its early stages and treated aggressively, postpartum psychosis can have a good prognosis. Maintenance on antipsychotic medication as per standard treatment guidelines is recommended. Women at high risk for psychosis, many of whom have a diagnosis of bipolar disorder, have been successfully treated with immediate postpartum initiation of lithium (see below). High-dose estradiol-administered postdelivery was shown to reduce the risk of recurrent postpartum psychosis in an open study that awaits replication.

Bipolar Disorder

Although an important proportion of postpartum psychosis cases will be attributable to bipolar disorder, it is critically important to recognize that the mothers with bipolar illness have a higher risk of experiencing a nonpsychotic, nonmanic episode in the perinatal period. It is estimated that 50 percent of women with bipolar I disorder had perinatal mood disorder episodes, with 25 percent constituting nonpsychotic mania and 20 percent constituting mania or psychotic depression. Episodes were notable for short postpartum latency and increased postpartum relative to pregnancy incidence. Mothers with bipolar II disorder had more widely distributed timing of mood episode, both during pregnancy and after 1-month postpartum. Clinicians should maintain a high suspicion for bipolarity given some indication that this diagnosis might be missed in as many as 50 percent of cases presenting with unipolar PPD.

Somatic treatments are the front line treatments for this disorder; however close clinical monitoring and efforts to enlist the aid of social supports to prevent sleep disruption and stabilize social rhythms in the new mother with a history of bipolar disorder may be particularly important. Illness recurrence after discontinuation of treatment was approximately 50 percent in pregnant and nonpregnant women but 3-fold higher in the postpartum period in postpartum compared to nonpostpartum women with bipolar disorder. Women who choose to maintain mood stabilizer treatment in pregnancy (vs. those who discontinue abruptly) have half the rates of recurrence, but may still have recurrence rates of 20 to 25 percent. For those women who do recur after medication discontinuation, illness recurrence occurred within the first trimester 50 percent of the time.

Lithium prophylaxis in late pregnancy or with 24 hours postbirth has been associated with a nearly 10-fold decrease in the rate of postpartum illness recurrence. Sustained remission out to 9 months postpartum was optimized with lithium maintenance therapy. For unresponsive illness, benzodiazepines, antipsychotics, and ECT also have an important role. A low-dose, mid-potency antipsychotic, such as perphenazine, can be added as a PRN agent for prodromal or breakthrough symptoms.

Nonaffective Psychosis

This category includes diagnoses of schizophrenia, delusional disorder, brief and unspecified psychosis. While first-generation antipsychotics (FGAs) are associated with diminished fertility rate via hyperprolactinemia and anovulation, second-generation antipsychotics (SGAs) are associated with normal fertility rates and high rate of unwanted pregnancies and repeated abortions in this population. These are often high-risk pregnancy due to low rates of prenatal care, high rates of poverty, homelessness, poor social supports, poor nutrition, high use of prescribed medications, alcohol, drugs, and tobacco. These women also have increased rates of sexual abuse pre-pregnancy and physical abuse in pregnancy. Often these mothers are less aware of signs of labor and delivery and are at greater risk of obstetric complications. Women with nonaffective psychosis represent the minority (15 percent) of hospitalizations due to psychosis, which are largely accounted for by affective psychosis. In women with schizophrenia, the relative risk of hospitalization in the first 30 days postpartum compared to 6 months postpartum is 5.7. Common delusional themes include ideas that the birth did not occur, that the baby is dead or defective and there can also be command hallucinations to harm the baby. Fifty percent of these mothers are separated from infants at discharge from the obstetric hospital due to incapacity to provide for child's minimal needs for physical sustenance. There is a 75 percent puerperal recurrence rate in schizophrenia and a 100 percent readmission rate 10-year post index postpartum admission due to schizophrenia relapse.

This is a diagnostic group for whom severe and persistent mental illness is convincing for medication need. Psychotic illness confers the greatest increase in risk of poor fetal outcome. Evidence from standard guidelines would be used to guide treatment with no knowledge that treatment efficacy would differ between postpartum and nonpostpartum populations. Psychosocial treatments include individual cognitive treatments, family support, and ideally rehabilitation programs would facilitate mother–infant caregiving and bonding for new mothers. Close collaboration among health professionals is necessary during the perinatal period in these women.

Screening

Western European countries and Australia have been international leaders in establishing nationwide surveillance programs for detection of postpartum mood disorders. Legislation at the national and state-level in the United States over the past decade reflects increased recognition of the seriousness of maternal mood disorders with the 2010 Melanie Blocker Stokes MOTHERS Act and the recent Bringing Postpartum Depression Out of the Shadows Act, which was passed as part of the 21st Century Cures Act. The recent postpartum bill authorizes the Secretary of Health and Human Services to provide federal grants to states for the purpose of screening, assessing and treating PPD. Increasing number of US states mandate

health care providers and facilities to screen newly delivered mothers for PPD.

Just as for other medical conditions, screening for perinatal psychiatric disorders has caveats and must be implemented responsibly so that cases identified can be referred and treated in a timely manner. Careful planning for the choice of instrument, the setting, the clinical cut-off scores, and the timing of administration must be appropriate to individual settings. Screening offers little hope for symptom improvement if referrals for treatment cannot be made. Because the assessment of perinatal disorders can be confounded by somatic complaints common to the perinatal period, such as sleep and eating disturbances, decreased sexual interest, weight gain, and fatigue, widespread use of the Edinburgh Postnatal Scale for Depression, which de-emphasizes such symptoms, has become a common instrument utilized by obstetricians, family doctors, psychiatrists, and pediatricians. Extensive experience with the ten-item Edinburgh Postnatal Depression Scale (EPDS) (http://www.postpartum.net/learn-more/tools-for-mom/) has shown it to accurately identify depression in pregnant and postpartum women cross-nationally; furthermore, this brief screening instrument has been validated in pregnant and postpartum populations, and in many languages, and may be easily incorporated for standard practice use in diverse treatment settings. Notably, the EPDS is sensitive to symptoms of depression and anxiety and does not distinguish between them. Screening for possible suicidal or infanticidal thoughts is especially important in more severe forms of the disorder. A high index of suspicion is warranted in women with severe, early-onset postpartum disorder.

Mechanisms

What constitutes the "hormonal trigger" prompting perinatal symptoms has been a topic of significant scientific curiosity, but remains poorly understood. Well appreciated are the hormonal excursions of the perinatal period, whereby late gestational increased concentrations of CRH (1,000-fold), cortisol (3-fold), estradiol (100-fold), and progesterone (10-fold), are followed by different rates of decline after delivery. Postpartum withdrawal of estradiol and progesterone, which occur rapidly within the first 2 to 7 days postpartum, when studied under simulated conditions, induced depressive relapse in women with a history of perinatal mood disorder but not in women without such history. Other experimental simulations of this kind indicate that reproductive hormone fluctuations are associated with disruptions in CNS corticolimbic connectivity that promotes healthy emotion processing. Likewise, the hormones CRH, cortisol, oxytocin, and prolactin, which fluctuate across the perinatal period and for which interindividual differences may confer differential risk for mood disorders, are implicated in stress response, emotion processing, and social cognition functions. Although there are no peripheral hormone biomarkers for perinatal mood disorders, gestationally elevated CRH and reduced oxytocin, and prolonged blunting of the postpartum HPA axis have been associated with perinatal depression in, as yet, unreplicated studies. Subclinical hyperthyroidism poses risk for pregnancy depression whereas thyroid hypofunction or increased antibodies was associated with PPD.

Functional magnetic resonance imaging of healthy mothers in the early puerperium as they attend to visual and auditory cues of their baby reveal activation of brain regions involved in reward (striatum, midbrain, orbitofrontal cortex), empathy (superior temporal sulcus), emotional appraisal (insula and amygdala), and emotion–cognition integration centers (anterior

cingulate gyrus) with transitions to the maternal role. Depressed mothers activate this circuitry to a lower extent. Genetic mechanisms research assumes a general vulnerability for mood disorders coupled with a specific hormonal sensitivity. PPD risk is increased in mothers with a personal and/or family history of depression. Postpartum mood disorder within 30 days of delivery was concordant among parous sisters and was associated with strong genome-wide linkage signals. A number of monoamine and steroid hormone gene polymorphisms and hormone-sensitive epigenetic signals have been described which require further investigation.

Psychological and sleep-related mechanisms have a critical place in the understanding of perinatal disorders. Maternal sleep deprivation is a common, if not universal, consequence of the infant's lack of a circadian rest–activity cycle, but in and of itself does not result in a major depressive disorder. Nevertheless, sleep therapies may improve mood in depressed patients. The average time for an infant to develop a recognizable nocturnal sleep pattern is 12 weeks. At this time, infants display a clear nyctohemeral pattern of melatonin secretion, indicating maturation of the master clock, the suprachiasmatic nucleus. Becoming a mother represents an unparalleled identity transition, and a disruption of life's routine including relationships, work, sleep, exercise, nutrition and finances. High social stress, marital discord, minority status, adolescence, and neuroticism are all associated with increased risk of perinatal mood disorders. Women with a history of attachment insecurity or trauma might have particular struggles when faced with an infant to protect from the abuse or neglect which she might have experienced.

Breast-Feeding

Lactation occurs because of a complex psychoneuroendocrine cascade that is triggered by the abrupt decline in estrogen and progesterone concentrations at parturition. Lactation becomes established in response to the neurological stimulation transmitted by suckling. The composition and amount of breast milk change as the infant grows. In general, babies should be fed as needed rather than by schedule. Breast-feeding has many benefits. The composition of breast milk supports timely neuronal development, confers passive immunity, and reduces food allergies. Psychological benefits accrue to both mother and infant. Every effort should be made by health professionals to convey these benefits. The workplace should include a place where women so interested can express breast milk. If for some reason, however, the mother is not able to breast-feed, reassurance that the child can be well cared for and adequately fed with current formulas should be provided to minimize the attendant guilt that may occur. Certain organizations that insist on breast-feeding for 2 years so as not to deprive the child of his or her "natural rights" may deter maternal and child well-being.

There is a theoretical argument for depression and lactation difficulties having a common neuroendocrinologic basis, perhaps mediated by disrupted oxytocin, prolactin, or stress neuroendocrine circuitries. Studies are inconsistent regarding whether lactation is protective or aversive for mental health. Mothers with anxiety or depression history should be encouraged to obtain lactational education and support early to optimize success. The fostering of skin-to-skin contact between mother and infant during lactation has important emotion-regulatory influences for both members of the dyad and also facilitates mother–infant socioemotional bonding.

Psychiatric Treatment Decision Making during Pregnancy and Postpartum

Structured Clinical Decision Making. Because of the temporal coincidence of mood disorders during women's reproductive years and

because pregnancy is unplanned in 50 percent of cases, psychiatrists must routinely ask about contraception and family planning during pharmacological management visits. The family planning period is the ideal time to have a structured preconceptual counseling discussion regarding perinatal treatment with prospective parents. Decision making in the treatment of pregnant women should be structured to include patient education regarding the natural history of her psychiatric disorder, the identification and review of likely risks and benefits associated both with the disorder and available somatic treatments, and frank discussion of available treatment options. This decision-making process must be tailored to address the severity of the patient's current illness and psychiatric history, as well as the experiences, values, and concerns of the patient. Clinical decision making must balance the known and unknown risks of medication treatment with the known and unknown risks of untreated psychiatric illness.

Open communication, provision of risk–benefit information to patients in a comprehensible format, and the ability to listen to, reflect, and respect the patient's concerns and value structure are critical to optimizing treatment decisions and obtaining fully informed consent for treatment. Such conversations take place in an iterative manner. Some women will desire a withdrawal of psychotropic medications under close supervision prior to pregnancy with iterative decisions about medication resumption over time in the event of symptom recurrence or if there is a recommendation for postpartum prophylaxis. When pregnancy is discovered during an ongoing medication treatment course, a risk–benefit analysis of medication continuation should be re-evaluated in light of altered risks associated with treatment continuation and risk of illness recurrence during the perinatal period.

Clinical decision making is complicated for pregnant and breast-feeding women because of the exponential growth in reports about psychotropic use in pregnancy that can be difficult to interpret and are often conflicting. Furthermore, clinicians in psychiatric and obstetrics may lack experience or comfort in this domain and thus the perinatal patient may lack access to a useful decision-making process. Perinatal psychiatrists are skilled in this particular area and should be consulted when needed. Updated data about perinatal risks of psychotropics can be found at a variety of free websites including the US National Library of Medicine http://www.toxnet.nlm.nih.gov/ Developmental and Reproductive Toxicology Database or Drugs and Lactation Database, the Organization of Teratology Information Specialists http://www.otispregnancy.org/, Thomas Hale's website on lactation http://www.infantrisk.com/, and the Massachusetts General website http://www.womensmentalhealth.org/. Subscribers to Micromedex Solutions can also access Reprotox, TERIS, and Shepard's databases.

Notably, the U.S. Food and Drug Administration published a new Pregnancy and Lactation Labeling Rule that went into effect June 30, 2015 to improve access to a summary of all pertinent studies that bear upon perinatal use of individual medications. This update replaces the older system of pregnancy letter categories A, B, C, D, and X and instead "requires that labeling include a summary of the risks of using a drug during pregnancy and lactation, a discussion of the data supporting that summary, and relevant information to help health care providers make prescribing decisions and counsel women about the use of drugs during pregnancy and lactation (http://federalregister.gov/a/2014–28241)." Drugs submitted to the FDA after June 30, 2015 will be required to use this new format while drugs approved previously will be gradually phased in.

Risks of Psychiatric Illness during Pregnancy

IMPACT ON MOTHER. Untreated illness can lead to clinical deterioration with concomitant psychological distress, social and occupational dysfunction, financial hardship, an inability to plan for and successfully cope with the impending life transition of motherhood, and possible suicidal ideation and self-harm. With many psychiatric disorders, functional impairment results in sleep disturbance, poor nutrition, lower use of prenatal vitamins, weight loss, as well as increased risk of cigarette, alcohol, and drug use. Poor rates of prenatal care can result in failure to detect gestational anemia, diabetes and hypothyroidism, which can have significant negative impact both on mother and fetus. Close attention must be paid to the host of behaviors which accompany psychiatric relapse in framing the risk–benefit discussion. As an example, cigarette smoking is associated with increased rates of miscarriage, intrauterine and perinatal death, preterm birth, and intrauterine growth retardation, outcomes which may be worse than those of psychotropics medications.

IMPACT ON OFFSPRING. Untreated mental illness carries significant independent risks for pregnancy outcomes. Increasingly appreciated is the impact of the prenatal milieu on fetal and infant health. The maternal HPA axis can have fetal programming effects to increase offspring risk for subsequent risk adult-onset medical and psychiatric disorders. Severe prenatal maternal stress during the first trimester of pregnancy has been associated with congenital anomalies. Furthermore, prenatal anxiety, pregnancy-specific anxiety, depression and maternal stress hormones have been associated with preterm birth, increased rates of obstetrical interventions during labor and delivery, small-for-gestational age birth weight, impaired neonatal stress regulation, infant fearful temperament, adverse cognitive outcomes, and reduced gray matter brain volumes. Novel tests of fetal cardiovascular reactivity during maternally stressful tasks are associated with altered nervous system function in utero in fetuses of depressed and anxious mothers.

Also well understood is the potential negative impact of psychiatric illness on maternal caregiving functions. Mentally unwell mothers have increased likelihood of having intrusive or withdrawn interactional styles with their infants and may lack sensitivity and attunement to infant cues. These maternal behaviors have been associated with offspring vulnerability for poor self-regulation, low self-esteem, insecure attachments, behavioral problems, and delays in motor and cognitive development. Ultimately, the complex multifactorial nature of the fetal environment, postbirth environment, and shared parental–infant genetic factors make unitary explanations of infant development overly simplistic. Yet, awareness of these distinct levels of infant developmental origins empower newly pregnant women to optimize conditions for their offspring.

Nonsomatic Treatments for Perinatal Mood Disorders. Treatment options should be presented according to a hierarchy from least to most somatically invasive. Nonpharmacological depression treatments, such as outpatient psychotherapy or mindfulness-based stress reduction, can be

advantageous for mild to moderately depressed or anxious women by providing her with emotional support and instrumental skill development needed to facilitate coping through this life transition. Maintenance of ongoing clinic visits and psychiatric monitoring, along with optimizing general health and social behaviors are all important strategies that promote primary prevention or early intervention as needed, thus conferring the best prognosis and preventing the untoward consequences and refractory course that otherwise might develop.

Chronobiological treatments were developed on the premise that depression results from a desynchronization of circadian rhythms. Treatment implications are to resynchronize the biological rhythms of sleep and activity with those of other underlying circadian rhythms, such as those of melatonin, cortisol, thyroid-stimulating hormone (TSH), or prolactin. This end may be achieved by the administration of bright light (greater than 2,500 to 10,000 lx) or wake therapy (sleep deprivation) at critical times of the day. Although still in its experimental stages, the use of bright light or wake therapy in women with pregnancy or PPD has had promising results in pilot studies. Further studies in this area would be worthwhile, given the relatively short onset of action of these treatments and the avoidance of risks associated with pharmacological interventions.

The benefits of social and material support that is not delivered in traditional clinics and which promotes the organic routine of mother caring for her children in her home, cannot be understated for new mothers. Models for treatment delivery continue to change over the years. In the United Kingdom, Australia, and several European countries, for example, home health care providers have had a major effect on improving the recognition and management and reducing the adverse consequences of these disorders. In the United States and elsewhere internet-based intervention, health coaches by phone and support groups are increasing. For women with severe mental illness, mother–baby inpatient units (in Europe and Australia) enable mothers to undergo the necessary treatment while continuing to breast-feed and develop a relationship with their baby. In the United States there are several mother–baby day hospitals and the first Perinatal Psychiatry Inpatient Unit was created in 2011 at the University of North Carolina at Chapel Hill.

Pharmacological Treatments during Pregnancy and Lactation

MAKING SENSE OF THE LITERATURE. Although ethical considerations limit performing randomized, prospective medication studies with pregnant women, improved methods for naturalistic studies (large samples, propensity score matching) and ongoing collection of observational data in registries have increased the information available to women facing perinatal treatment decisions. Increasingly sophisticated study designs and statistical methods have been important for reducing sources of bias, such as confounding by indication, which refers to the host of demographic, behavioral, and health-related characteristics associated with the illness itself (i.e., increased smoking, poor prenatal care) or use of medication (i.e., potentially more severe illness). Both medication exposure and also maternal mental illness contribute to the outcomes of interest and thus the most informative studies are those that can measure and control for each. A true teratogen is associated with a consistent birth defect across studies; therefore, inconsistent findings of disparate defects are harder to interpret, especially in view that the psychiatric disorder itself is not always

accounted for. Additionally, well-powered studies which properly control for confounding are needed to detect rare infant malformations. The multifactorial nature of infant development suggests that large studies are needed which can control for numerous contributory variables. Greater standardization of outcomes and pooling of data across research centers will facilitate advances in this domain.

Pharmacotherapy is an important modality for women with illness unresponsive to alternative treatments. The psychiatrist plays an important role in supporting the patient to make an informed decision and in providing education to and fostering communication with the obstetrician and pediatrician. A guiding principle for treating perinatal women is to minimize exposures both to the untreated illness and to psychotropics. The medication to which a mother has responded in the past is often the best choice even in the absence of substantial data given the known risk of severe, untreated illness. Most drugs cross the placenta albeit to a varied extent. Psychotropic transfer to the baby via lactation is much lower than placental transfer, due to incomplete psychotropic absorption into breast milk as well as first-pass metabolism effects in the infant. If fetal exposure to a particular psychotropic has already occurred, this agent would be reasonable to continue during lactation rather than switching to one with theoretically lower exposure. If the drug was started postpartum, observation of the infant's behavior before the mother is treated permits clinicians to avoid misinterpreting typical behavior as potentially drug related. For each psychotropic class below, the available evidence is summarized in order of reproductive domain: congenital teratogenicity, late gestation effects, neonatal withdrawal/toxicity, developmental (neurobehavioral) teratogenicity, and lactation.

SEROTONIN REUPTAKE INHIBITORS (SRI = SSRIs AND SNRIs). Across numerous studies of first trimester exposure, no SRI has reached the level of a confirmed teratogen. There was significant concern about the association of paroxetine with cardiac septal defects which resulted in an FDA label change in 2005 to category D; however; subsequent meta-analyses have had contrasting results, which challenges the conclusion that paxil is a cardiac teratogen. In several methodologically rigorous, large studies which used an untreated depressive control group as well as propensity score matching to control for depression severity and other confounders, the relative risk of cardiac malformation as a whole or limited to septal defects with SRI use was not significantly greater than with no SRI use.

The possible link between late pregnancy (after 20 weeks of gestation) exposure to SRIs and persistent pulmonary hypertension of the newborn (PPHN) also created public alarm and FDA label changes warning of risk in 2006, which was then rescinded in 2011. An increasing number of studies and meta-analyses draw contrasting conclusions and don't adequately

control for confounders such as maternal depression severity, preterm birth, and mode of delivery. Current information, therefore does not support SRI discontinuation on the basis of PPHN risk.

Spontaneous abortion rates were initially estimated to be higher in antidepressant-using women, however a significant limitation of earlier research was the lack of control for psychiatric status and the concern that therapeutic abortions were misrepresented as miscarriages due to stigma. Furthermore lack of control for prior miscarriage is a significant shortcoming of prior research. Several analyses in a large cohort found no increased risk attributable to SRIs when depression status was controlled for. Additionally, SRI use was not associated with any increased risk of stillbirth, neonatal death and postneonatal death in a second well-controlled, large study.

Antidepressant exposure in late pregnancy has been significantly associated with lower gestational age and lower APGAR scores at 1 and 5 minutes, even with proper controls for psychiatric status. SRI exposure was associated with lower infant birth weight only when maternal psychiatric status was unaccounted for. Importantly, these effects are quite low and of questionable clinical significance. A 1.2 to 2-fold increased risk of preterm delivery, defined as birth prior to 37 weeks of gestation, was associated with in utero SRI exposure in a meta-analysis that controlled for illness exposure; however other characteristics of women who maintain meds through pregnancy could confer residual confounding. The extent to which depression is associated with preterm birth parallels that from SRIs, but remains to be studied further.

A neonatal adaptation syndrome is reported to affect 30 percent of late pregnancy, SSRI-exposed newborns, and substantially more preterm infants are affected. Features of poor neonatal adaptation include respiratory difficulty, hypoglycemia, hypertonia, jitteriness, vomiting, seizures, and sleep disturbances. The syndrome appears to be time limited with greatest incidence in the first 48 hours postbirth, but potentially extending to 2 weeks postbirth. While one study provided evidence for similar rates of neonatal signs in the absence of SRI exposure in 2 weeks prior to delivery, thus implicating maternal or other neonatal factor in this syndrome, another study found increased adaptation syndrome risk even after control for maternal characteristics.

A number of prospective studies have found no impact of SRIs or TCAs on temperament, mood, arousability, activity level, distractibility, global intelligence quotient (IQ), or language development in children followed to 7 years of age. The concern that in utero SRI exposure is associated with a delay in motor skills and milestone achievement development has been lessened by evidence that development was still in the normative range, in these studies, and that no effects of in utero exposure were noted at evaluations beyond 1.5 years of age. Similarly, studies which suggest that

diagnoses of autism spectrum disorders (ASDs) and attention-deficit disorders (ADHDs) as related to in utero SRI exposure inconsistently control for confounding factors such as underlying maternal diagnosis/illness severity, other exposures such as tobacco and alcohol, as well as family history of these behavioral disorders. More recent studies which do control for these variables to the greatest extent possible find no association between SRI exposure and ASD but a possible relationship with ADHD (OR = 1.8).

Sertraline and paroxetine are the SRIs with the largest evidence-base of use in breast-feeding and the negligible concentrations detected in infants. Relative infant doses from breast milk are often less than 2 percent of adjusted maternal dose or have milk/plasma rations less than one. Citalopram, fluoxetine, and venlafaxine are more commonly associated with infant concentrations greater than 10 percent of adjusted maternal dose. In general, there are low rates of adverse effects, which are short-lived, in breast-fed infants whose mothers were treated with SRIs. With fluoxetine, breast-fed infants of fluoxetine-treated mothers gained less weight after birth, although adverse behavioral effects were not observed. Colic was reported in three infants who were breast-fed by mothers taking fluoxetine. The full-term neonatal cytochrome P450 activity is approximately one-half of that found in adults. The collective data on serum levels suggest that infants older than 10 weeks of age are at low risk for adverse effects from antidepressants. The higher lipid content of hind milk makes it likely that the second half will have a higher concentration of maternal medication than the fore milk. Taking medication immediately after breast-feeding minimizes the amount present in milk and maximizes clearance before the next feeding.

Duloxetine is a newer SNRI that has an important place in the treatment armamentarium for women given its efficacy in MDD, GAD, and fibromyalgia. Limited study to date reveals no increased rates of congenital anomalies in approximately 450 cases. Rates of spontaneous abortion, prematurity and perinatal complications are similar to historic controls and the FDA adverse events reporting system indicated that duloxetine doesn't confer any disproportionate adverse pregnancy outcome compared to other drugs or other antidepressants. The estimated infant dose through breast milk low and estimated at 0.14 percent of the maternal dose.

TRICYCLIC ANTIDEPRESSANTS. The data show no evidence of increased incidence of congenital malformations in fetuses exposed to TCAs, although evidence is less extensive than that for the SRIs, and the majority of exposures have been to clomipramine. Increased likelihood of preterm birth and low birth weight at an OR of ~1.5 mirrors the findings of SRI-exposed pregnancies, and like SRIs, there have also been reports of PPHN and cardiac septal defects related to TCA exposure. Secondary TCAs (i.e.,

nortriptyline, desipramine) may be preferred in pregnancy based upon fewer hypotensive and anticholinergic effects. A poor neonatal adaptation syndrome has been described similar to that with SRI exposure and some evidence suggests it could be more severe and include anticholinergic symptoms of urinary retention and bowel obstruction. TCAs are not typically found in measurable amounts in nursing infants. Respiratory depression and sedation, the only adverse outcomes reported with any TCA, occurred in an infant whose mother was taking doxepin. Of the drugs in this class, nortriptyline has been studied the most as a treatment for breast-feeding women. Children who were exposed to TCAs both in utero and through breast milk have been followed through preschool and compared with children who were not exposed to these drugs, and no developmental problems were found. Due to increasing preferences for SRI treatment, fewer long-term studies are available following in utero TCA exposure.

MONOAMINE OXIDASE INHIBITORS. The use of monoamine oxidase inhibitors (MAOIs) is contraindicated in pregnancy for several reasons. First, studies found growth retardation in animals receiving doses in excess of the maximum human recommended dose. Second, pregnancy-induced hypertension is a common complication of pregnancy that often has an insidious onset; further exacerbation of that condition by use of MAOIs may lead to placental hypoperfusion and have serious fetal consequences. Third, if premature labor occurs, tocolysis with β-mimetics may be obstetrically indicated but not possible because of potential interaction with the MAOI. Fourth, anesthetic management in labor may be complicated by the relative contraindication of opioids in patients taking MAOIs. If pregnancy occurs while the patient is taking MAOIs, the drug should be discontinued. Guidelines have been suggested for appropriate obstetrical anesthesia and analgesia for women treated with MAOIs.

BUPROPION. Bupropion's dopaminergic mechanism of action and role in smoking cessation give it an important place in the treatment armamentarium. The literature raises inconsistent findings about incidence of cardiac malformations attributable to first trimester bupropion exposure, with suspicion raised regarding potentially increased risk of ventricular septal defects and left outflow tract heart defects. It is a great challenge to adequately power studies to look at specific, rare, heart defects and the studies to date have not consistently controlled for important confounding effects of maternal illness. The estimated infant dose through breast milk is estimated at 5.7 percent of the maternal dose.

MIRTAZAPINE. Mirtazapine is a novel antidepressant with noradrenergic and selective serotonin effects. In the case literature it was shown effective in rapid alleviation of hyperemesis gravidarum, on the basis of serotonin 3

receptor blockade. Treatment benefits were obtained at doses of 7.5 to 45 mg/day and in some cases, mirtazapine was administered initially through intravenous fluids and treatment course was concluded after 10 days. In a study of first trimester mirtazapine exposure, there were no increased rate of malformations and similar rates of spontaneous abortion and preterm birth compared to pregnancies with other antidepressant exposures. The estimated infant dose through breast milk is estimated at 1.8 to 4.4 percent of the maternal dose.

FIRST-GENERATION ANTIPSYCHOTICS. The greatest amount of data is available for FGAs. In a meta-analysis of nearly 75,000 births in women treated with low-potency antipsychotics (chlorpromazine), there was a 0.4 percent increased risk of malformations compared to the general population. Another systematic review found that the high-potency antipsychotic haloperidol was not associated with increased risks of malformation. Furthermore, in a meta-analysis of studies of parents with schizophrenia, rate of malformations and perinatal death did not differ from the general population. Preterm birth and small-for-gestational age births may be more common with FGA relative to SGA exposure.

Use of these agents in the second and third trimester is associated with increased rates of perinatal complications, low birth weight and transient perinatal syndromes (motor restlessness, tremor, hypertonicity, abnormal movements, difficulty with feeding, and possible neonatal jaundice and functional bowel obstruction) that generally resolve within days but may last for up to 10 months after birth. The estimated infant dose through breast milk is estimated at 3 percent of the maternal dose, however there have been some reports of drowsiness in the infant. From studies conducted four decades ago, there is no evidence of long-term effects on behavioral functioning or IQ up to age 4 years in children. One study found that in utero exposure was associated with a nearly 4-fold increased rate of abnormal psychomotor development at 7 to 10 months of age which concurs with some data in SGAs and may be also related to maternal illness effects.

Given the high risk to the mother and fetus or infant during a psychosis recurrence and the minimal risks described, continuation of treatment is recommended. Use of such agents as augmentation or adjunctive agents (in combination with antidepressants or mood stabilizers) in the treatment of mood and anxiety disorders must be weighed against the severity of illness. Lower-potency FGAs have greater risk of hypotension while higher-potency agents confer risk for extrapyramidal symptoms. Intermediate-potency agents such as perphenazine constitute a useful compromise in this drug category.

SECOND-GENERATION ANTIPSYCHOTICS. While data are slowly accumulating for

SGAs in pregnancy, currently samples sizes remain fairly small. The results of three published studies examining the SGA as a class found conflicting results with no increase in major malformations in one but 2.1 to 2.7× rate of malformation in the others. Data from the first 195 analyzed registry participants in the National Pregnancy Registry for Atypical Antipsychotics cohort found no difference in malformation rates in SGA-exposed compared to controls. In a study of placental passage, the order of greatest to lowest passage was olanzapine, risperidone, and quetiapine. Low birth weight and higher NICU admissions were present for olanzapine-exposed infants. Via SGA-related increases in maternal weight gain, there is increased risk of neural tube defects in exposed infants, as well as increased risk of gestational diabetes, large-for-gestational age infants, and increased risk for c-section delivery. Because clozapine may induce fatal agranulocytosis in adults caution should be demonstrated during use in pregnancy. SGAs confer less risk for extrapyramidal side effects in the mother. In a study of 142 pregnancies of largely SGAs, antipsychotic dose was inversely related to gestational age at birth and positively related to the likelihood of NICU admission. Higher rates of respiratory distress were reported when SGAs were combined with mood stabilizers. Intake of SGA through milk leads to 0.1 to 10 percent maternal dose.

In a small group on antipsychotic-exposed infants (n = 22) at 6 months of age, of whom approximately 50 percent were SGA-exposed, lower psychomotor scores on posture, tone reflexes, and motor skills were detected which were also associated with increased severity of maternal psychiatric history. Larger studies are needed to replicate this finding and determine the longevity of the effect. There are no data yet available on long-term neurodevelopment of infants with in utero SGA exposure. The National Pregnancy Registry for Atypical Antipsychotics continues to accept eligible subjects to examine the above outcomes in greater detail (http://womensmentalhealth.org/clinical-and-research-programs/pregnancyregistry/).

LITHIUM. The scientific community has progressively reassessed the risk of Epstein cardiac anomaly following first trimester lithium exposure. While the risk was originally cited as 400-fold, this risk was later revised to 10 to 20-fold (0.5–1 per 1,000), and some argue the risk may be still lower due to methodological limitations in extant studies. Fetal echocardiography is recommended to rule out the presence of cardiac malformations as a precaution. There is a lack of consensus regarding gestational size increases and preterm delivery. Neonatal signs of lithium toxicity have been described to include muscle flaccidity, inhibition of normal neonatal reflexes, lethargy, cyanosis, and cardiovascular effects, including atrial flutter, tricuspid regurgitation, and congestive heart failure. Careful monitoring of maternal lithium concentration in pregnancy is required

because of the dramatic shifts in fluid volume during pregnancy. A higher glomerular filtration rate, coupled with the increased plasma volume, often leads to a requirement for higher doses in the gravid woman to achieve comparable serum lithium concentrations. Lithium crosses the placenta freely, and maternal and fetal plasma concentrations are similar. Infants with higher lithium concentration at delivery had low APGAR scores, longer hospital stays, more neuromuscular signs; therefore, withholding maternal lithium therapy for 24 to 48 hours prior to delivery is a suggested strategy to lower these adverse outcomes. Because of possible dehydration during labor, rapid loss of fluid volume after delivery, and potentially elevated gestational lithium doses, reinstatement of lithium therapy in the recently delivered mother should be gradual to approximate pregravid doses. Early postpartum measurement of lithium concentrations and maternal side effects is required.

New evidence of successful breast-feeding and healthy infants in well-educated and clinically stable mothers treated with lithium has started to change perceptions about lithium during lactation. The mean concentration of lithium in breast milk is approximately half that of the maternal serum concentration, with a range from 25 to 77 percent. Lithium concentrations in the serum of breast-feeding infants are estimated to be one-fifth those in maternal serum. Laboratory monitoring of infant serum lithium, TSH, blood urea nitrogen, and creatinine immediately postpartum and up to 6 weeks of age is recommended with ongoing monitoring every 2 to 3 months thereafter during breast-feeding. Long-term neurobehavioral assessments have found no difference in exposed relative to nonexposed siblings in two studies; however one of these reported mild developmental delay in both groups.

ANTIEPILEPTIC DRUGS (AED). The neurology literature provides the majority of data regarding safety and risks of the AEDs in perinatal women. Convergent studies from numerous antiepileptic drug registries associate first trimester use of carbamazepine and depakote with neural tube defects in 0.5 to 5 percent and major malformation in up to 16 percent of exposed offspring. Valproate mono- or polypharmacy conferred the greatest risk. Valproate was also associated with greater rates of fetal death compared to other AEDs. The recommendation is for women taking AEDs to take high-dose folate (4 mg/day), which may mitigate against the risk of neural tube defects. Reported malformation other than neural tube defects includes heart defects, urogenital defects, and oral clefts. Lamotrigine has gained particular attention as a therapy for bipolar depression in women with bipolar disorder because of recent registry data finding no difference in rates of malformations compared to the general population. There is not a consensus about this result and earlier registry reports raised caution for potentially increase risk of oral clefts.

During breast-feeding, valproate concentration in infants can reach up to 40 percent of those in the mother with generally lower concentrations for carbamazepine-exposed infants. Lamotrigine infant exposure can reach up to 21 percent of maternal dose. Higher range concentrations are related to the combination of in utero and breast milk exposure. Thrombocytopenia was reported in one case of a breast-fed child with valproate, transient mild hepatic toxicity, poor feeding and sedation with carbamazepine and increased platelets and rash with lamotrigine.

When AED exposure is limited to breast milk, they confer no adverse impact on IQ, verbal and nonverbal memory or executive function by age 6; however, infants exposed to valproate in utero have been reported to have dose-related, 3-fold higher rates of developmental difficulties and a greater incidence of IQ under 70, which was lower than those infants exposed to carbamazepine or lamotrigine.

REPRODUCTIVE SENESCENCE

Menopause

Menopause—the cessation of ovulation—is due to the depletion of responsive ovarian oocytes and generally occurs between the ages of 47 and 53 years. Menopause is generally heralded by 5 to 10 years of erratic ovarian function as revealed by irregular uterine bleeding, followed by sustained amenorrhea. The accepted definition is the cessation of menses for 1 year (not due to other causes), but this narrow definition belies the complexity of the reproductive and psychological transition that the term *menopause* connotes.

Study of the relationship between the menopausal transition and psychiatric symptoms has been greatly assisted by the operationalization of the stages of reproductive aging. The Stages of Reproductive Aging Workshop (STRAW) staging system is a seven-stage system that is anchored by the final menstrual period. Stages have interindividual variability in length and age of onset. Menstrual cyclicity and FSH measurement are indices that are used to determine reproductive stage. Normal menstrual cycle length is 21 to 35 days. Individual-specific cycle length changes of greater than 7 days can be an early sign of the menopausal transition. A single FSH level that is drawn at approximately day 3 of the menstrual cycle and is 10 mIU/mL or greater can stage a woman as being in her late reproductive years (stage −3). The perimenopause is marked by variable cycle length, skipped cycles, and FSH concentrations of 10 mIU/mL or greater (without meaningful cutoffs when concentrations are greater than 10 mIU/mL). The postmenopausal stage is characterized by an increase in FSH (greater than 30 mIU/mL) and LH and a reduction in estradiol (less than 30 pg/mL), progesterone, and inhibin B.

The erratic and fluctuating levels of estradiol and progesterone observed during the

menopausal transition as early as stage −3 are associated with hot flashes, night sweats, sleep disturbances, irritability insomnia, breast tenderness, migraines, joint pain, and cognitive and affective disturbances. In stages +1 and beyond, vaginal atrophy, urinary incontinence, and sexual dysfunction become manifest. Eumetabolic hypoestrogenism predisposes to osteoporosis, dementia, and cardiovascular disease. Before the 1900s, menopause tended to occur just before death. Previous negative associations partly reflect the nearly coincident timing of menopause and death. This relationship no longer holds, and it is commonplace for women to live 30 or more years after menopause. The challenge now is to avoid the acute and chronic disabilities that are related, in full or in part, to the uniquely female loss of gonadal function. Given that there may be as many as 30 to 50 years of profound hypoestrogenism, this is no small challenge. Furthermore, today's and tomorrow's women must adjust to the challenge of life after the loss of endogenous reproductive capacity. Given women's longevity and short reproductive window, it makes little sense to define them solely by their individual or collective reproductive competencies or events.

Surgical menopause occurs when the ovaries are removed. This removal is often done in women older than age 35 years who require a hysterectomy in an effort to reduce the probability of ovarian carcinoma. Prophylactic oophorectomy is a controversial procedure, however, and the symptoms that follow surgical castration before the age of natural menopause generally are more pronounced than are those that follow natural menopause. In natural menopause, the ovaries remain intact and continue to secrete androgens, including testosterone and androstenedione, which can be converted in many tissues to estrone, a weak estrogen. Surgical menopause results in an abrupt and complete loss of the ovarian secretion of androgens, estrogens, and progesterone, often well before the age of menopause. Estrogen replacement therapy is generally indicated when bilateral oophorectomy is performed before natural menopause, but compliance may be compromised by ambivalence or a sense of premature loss if the decision to undergo oophorectomy was made without full consideration of all the risks and benefits. Recent longitudinal studies have demonstrated that oophorectomy before the age of natural menopause is associated with an increased risk of all-cause dementia and Parkinson disease. Androgen therapy is more controversial, in part, because a tested, safe, physiological replacement product is not yet commercially available. The adrenal hormone dehydroepiandrosterone (DHEA) is an androgen being promoted as a dietary supplement to restore a sense of well-being, but few data are available on its long-term safety and effects. The same holds for the androgen androstenedione, which is often promoted to enhance athletic performance or support libido. There are no commercially available testosterone products, but one hormone product combines estrogen and methyltestosterone. Its use is recommended to enhance libido. It remains to be demonstrated that methyltestosterone crosses the blood–brain barrier and gains access to the key brain centers that drive libido.

Hormone Replacement Therapy. The pros and cons of hormone replacement therapy (HRT) have not been fully clarified. There is increasing awareness that not all women are the same, that not all estrogens are the same, and that not all progestins are the same. This complexity lends further confusion to the discussion and interpretation of the data. Probably the most important variable in gauging hormone effect on long-term health is the age of the woman when hormones are initiated. Thus the current concept is that recently menopausal women may benefit from extended exposure to an estrogen, whereas initiation many years remote from the time of menopause is unlikely to help and may harm. The question of whether to initiate or continue this therapy poses a medical, financial, and philosophical challenge. Until recently, most studies supported the conclusion that estrogen replacement increased longevity because of its cardioprotective properties. The results of the Women's Health Initiative (WHI) and its secondary analysis suggested that

cardiovascular risks and benefits conferred by a standard combination of an oral estrogen plus progestin (conjugated equine estrogens plus medroxyprogesterone) or a standard dose of equine estrogen preparation are dictated by the timing of their administration. Oral estrogen or estrogen–progestin protected women from cardiovascular disease if a woman was within 10 years of menopause but increased risk in women 20 or more years postmenopause. Stroke risk was increased 1.3-fold in women randomized to HRT relative to placebo and did not vary on the basis of age at HRT use. This may be because oral estrogen promotes venous thromboembolism regardless of age. Benefits conferred by HRT in the WHI included protection from bone fracture and reduced colorectal cancer risk.

However, because not all estrogens have the same molecular properties, the results of the Women's Health Initiative are not generalizable to other hormone products. In particular, evidence suggests that more-physiological preparations, such as transdermal estradiol patches, may minimize risks while retaining benefits. A negative effect of synthetic progestins as compared to progesterone has also been advanced as an explanation for some of the findings of the Women's Health Initiative and related epidemiological studies. Although concern exists that combined oral hormone use after menopause increases the risk of breast cancer and affords minimal, if any, cardioprotection, there is also evidence of benefits. Hormone use retards bone loss and reduces the risk of osteoporosis and fracture, maintains urogenital integrity, stabilizes mood and other critical brain outputs, including motor abilities, and maintains body habitus and composition. Whether hormone therapy protects against dementia also remains controversial. Some experts have suggested that the increased risk of venous thromboembolism and stroke with orally delivered estrogens may have counteracted the trophic effects of estrogens on neurons and glia. It is thought that nonoral routes of delivery of estradiol may prove neuroprotective because there is not an increased risk of venous thromboembolism with nonoral formulations. Although women at midlife tend to gain weight, studies have demonstrated that women who use HRT gain less weight and that the weight gained is preferentially deposited in a gynecoid (hips and thighs) rather than android (abdominal) pattern.

Of note, recent, prospectively collected population data showed that rates of dementia increased in both men and women after age 75 years but that the increase was much greater in women than in men (relative risk, 2.1). However, rates of dementia in women were reduced in proportion to years of hormone use after menopause. Ongoing estrogen use for longer than 10 years, especially when begun contemporaneously with the onset of menopause, reduced the risk of dementia in women to that of men. It is not widely appreciated that the brain converts testosterone to estradiol and that men typically have, at all times in their lives, more estradiol bathing their neurons and glia than women do except when women are pregnant. Because men do not experience gonadal cessation at midlife as women do, their brains continue to be protected by estradiol indefinitely.

Increasing basic research substantiates a role for estrogens in neurotrophic and neuroprotective functions. These are alternatively mediated via classical genomic functions of

estrogen receptors (ERs), which, when ligand activated, travel to the cell nucleus, where they act as transcription factors. In addition, cell membrane ERs are associated with activation of second messenger cascades and highly amplified gene transcription. Selective ER modulators (SERMs) target the same receptors as estrogen—ER subtypes ERα and ERβ—but differ in transcriptional ability based on unique ligand-induced conformational change of the ERs to facilitate interaction of ERs with coactivators and consequent agonism or antagonism of ER-mediated DNA transcription.

The risks and benefits of the use of existing SERMS by postmenopausal women are being defined. Antagonism of ERβ may pose risk for heightened depression and anxiety, given the demonstration of ERβ facilitatory effects on serotonergic activity, social interaction, and positive behaviors in rodent models of depression. Both ER subtypes are involved in neuroprotective functions. Relative density of each receptor subtype in a particular brain region likely underlies the ultimate action of any ER ligand. Although it was postulated the mixed agonist/antagonist raloxifene would have adverse effects on mood and memory because it is associated with hot flashes, there are no long-term studies that have addressed this in a group of women older than age 70 years when the risk of dementia is sufficient to be able to discern effect. The majority of studies find no increased risk of depressive symptoms in SERM-treated women, and some find SERM-associated amelioration of depressive symptoms. Likewise, raloxifene in doses of 60 or 120 mg was not associated with memory impairment and was recently reported to reduce the risk of mild cognitive impairment in postmenopausal women. In contrast, tamoxifen was associated with memory impairment when added to standard chemotherapy in breast cancer survivors. Aromatase inhibitors have been shown to have the most deleterious effect on cognitive functioning, which argues for an important role for estrogen in safeguarding neural health. The effect of long-term specific SERM use on cognitive function is an area in need of further study.

Current evidence does not reveal a strong association between estrogen use and breast cancer. In particular, oral contraceptive use, even in older women, is not associated with an increased risk of breast cancer. In terms of quality of life, hypoestrogenism can impair sexual enjoyment and libido and can cause hot flashes, night sweats, and profound fatigue. These sequelae can generally be ameliorated or reversed with estrogen use. The trend is to use the smallest dose that relieves these symptoms.

Progestin use is generally recommended for women with a uterus to guard against the development of endometrial hyperplasia and carcinoma; whether or to what degree progestin use compromises the beneficial effects of estrogen use requires clarification. The availability of a progestin-containing IUD may minimize or obviate altogether the risks of systemic progestin exposure. The adverse effects that most limit progestin use are depression and bloating. In animals, including monkeys, progestins increase irritability and diminish sexual receptivity, and progesterone has fewer negative side effects than do synthetic progestins such as medroxyprogesterone acetate (Provera). Furthermore, concern has been raised that the use of progestins other than progesterone will reverse cardioprotection by promoting vasoconstriction. Progestins may also increase the risk of breast cancer, particularly when used continuously. Women who previously experienced depression or PMS are most prone to mood disturbances with progestin use. To avoid or minimize the effects of progestin exposure, it is prudent to prescribe the smallest amount of progestin necessary to protect the endometrium from hyperplasia or use a progestin-containing IUD. Recent data support the notion that antidepressant and antidementia agents are more efficacious in women when they are estrogen replete. To achieve the best risk–benefit profile, it is worth titrating and individualizing route, dose, type, and regimen with the goal of maximizing benefits and minimizing adverse effects.

The recent introduction of a wide range of estrogen and progestin products, including so-called designer estrogens or SERMs, has expanded options. There is even less information about the long-term risks and benefits of partial estrogen agonists or SERMs than about more conventional estrogen preparations. The same holds for phytoestrogens derived from soy and flaxseed. Direct-to-consumer marketing of hormone preparations means that patients have more questions, concerns, and ambivalence. Although women want more information and guidance, physicians are poorly reimbursed for this type of counseling and often lack sufficient information.

Both men and women age and experience an age-related decline in reproductive capacity, but only women experience complete gonadal cessation. Loss of reproductive capacity may present a psychological challenge to those who are not reconciled to the loss of fertility. However, even

with gonadal failure, the availability of donor oocytes and sperm means that pregnancy can be initiated in a menopausal woman with an intact uterus who elects to pursue that option. Some women find these multiple options overwhelming. Menopause is not only a marker of life stage, it also presents biological and psychological challenges unique to women.

Menopausal Status and Depression

Depression at menopause was previously attributable to the "empty-nest syndrome," or what was termed "involutional melancholia." Nevertheless, many menopausal women report an enhanced sense of well-being and enjoy the opportunity to pursue postponed goals. Recent acknowledgment that women's roles may encompass more than child rearing has helped to facilitate this attitude. In recent years, however, it has become clear that the transition to menopause confers a risk for depression onset. New-onset depression affects 20 to 30 percent of women during the menopausal transition. The perimenopausal transition is associated with a 2- to 4-fold increased risk of depressive symptoms and a 2-fold increased risk of a depressive disorder in medically and reproductively healthy women *without* prior history of major depressive disorder relative to the time of healthy reproductive function. These increased risks remained after correction for numerous factors, including hot flashes, poor sleep, and adverse life events, factors that have challenged the interpretation of earlier studies. Although there was no association between mean hormone levels and mood, higher variability of between-visit estradiol concentrations in women relative to their nondepressed, premenopausal baseline was the strongest risk factor for the onset of a depressive disorder. This strongly argues for a mechanistic role of estradiol fluctuations in the onset of perimenopausal depression. Hot flashes were 1.3- to 2.2-fold more likely in women with perimenopausal-onset depressive symptoms and depressive disorder, a factor that argues for a shared neurophysiological mechanism of perimenopausal depression and vasomotor symptoms. Other risk factors for perimenopausal depression include personal history of mood disorder, previous reproductive-related mood disturbance (such as postpartum and/or premenstrual mood disorder), and psychosocial stress.

Management of Perimenopausal Vasomotor and Depressive Symptoms

There has been an interesting convergence of hormonal, antidepressant, and insomnia pharmacotherapies for premenopausal symptoms over the last decade. The replacement of perimenopausally declining estradiol and progesterone with exogenous hormones, previously the mainstay of perimenopausal symptom management, has been challenged by increased evidence of potential thromboembolic, cardiovascular, and cancer risks conferred by hormonal regimens (as found in the WHI). Although these risks are understood to be reduced based on patient age and HRT formulation, alternatives to HRT have nonetheless been widely promoted.

Antidepressant drugs are increasingly prescribed for hot flashes based on clinical trials that show benefit with low-dose serotonin/serotonin-norepinephrine reuptake inhibitors (escitalopram, venlafaxine, and duloxetine, but not sertraline). It is proposed that antidepressants may improve distress tolerance of hot flashes via symptomatic improvement of coexisting mild insomnia, anxiety, dysphoria, and/or irritability. Sleep disturbance affects nearly 50 percent of perimenopausal women and significantly contributes to reduced quality of life. Perimenopausal insomnia is variably attributed to vasomotor symptoms, night sweats, and mood disturbance. Targeting the sleep symptoms with sleep hygiene interventions may result in sleep and quality-of-life improvement of similar magnitude to that provided by HRT.

Conversely, estradiol has proven to be useful in the treatment of minor and major depression. Women with endocrinologically confirmed perimenopause and with structured interview–based diagnosis of depression experienced a 70 to 80 percent rate of remission during treatment with 50 to 100 μg per day of 17β-estradiol, compared to a 22 percent rate of response in placebo-treated subjects. Estradiol was equivalent to placebo in a postmenopausal trial, which may be a function of a nonresponsive age group.

Aging

Further investigations are required to distinguish the psychological and biological effects of aging from those of menopause. Aging is associated with diminished resiliency and circadian rhythmicity in both men and women. Desynchronized or reduced biological rhythms may compromise sleep, cognition, and mood. The brain is a target organ for gonadal steroids. Only women experience complete gonadal cessation. The loss of ovarian estrogen and progesterone exposure can precipitate hot flashes, reduced libido, sleep disturbances, affective complaints, and decreased memory. However, both men and women experience a gradual decline in adrenal and gonadal androgen production. Androgens help to maintain libido, muscle mass, strength, and energy. Androgen therapy has been advocated as a way to retard the aging process, but a tested, safe androgen replacement preparation is not commercially available for women. Many other chemotherapeutic agents have been advocated to reduce the burdens of aging. These include antioxidants, such as vitamin E, phytoestrogens, vitamin C, aspirin, alcohol, folate, vitamin D, and calcium; herbal supplements, such as *Gingko biloba;* and hormones, such as DHEA, androstenedione, growth hormone or analogs, and melatonin. Menopausal women often expect that physicians who offer hormonal products at menopause to increase long-term health also know whether other chemotherapeutic strategies have merit.

Psychological factors play a role in age-related disability. Disease and disability present psychological challenges and engender the need for changes in lifestyle and employment. Furthermore, stress may accelerate certain aging processes. In a longitudinal population study of the elderly, women whose cortisol levels remained elevated across the observation interval of 2.5 years displayed progressive cognitive impairment, whereas cognition improved in those whose cortisol levels returned to the normal range. The most common stress was social isolation. Depression in the elderly remains underrecognized and, therefore, undertreated, in part because the presenting signs and symptoms may be confused with dementia or other concomitants of aging. Successful aging demands adaptation to unique medical, biological, and psychosocial challenges.

FUTURE DIRECTIONS

Molecular medicine is poised to exponentially expand the understanding of the mechanisms mediating sexual differentiation and reproductive physiology. As it becomes ever more possible to translate understanding into clinical therapeutics, therapeutic promise will be greatly expanded. Concomitantly, given the importance of reproduction to the sense of mortality and immortality, it is likely that even greater anxiety will be aroused by the development of ever-more-clever ways to tinker with reproductive options. The idea that sex steroids imprint the brain in fundamental ways that determine sexual identities, longevity, and aptitudes will make it difficult to think in a polarized manner about gender, families, and aging. The psyche, both individual and collective, will continue to be ever more troubled by the implications for society and self.

Reproductive behavior is extremely personal, but governments and religions often feel compelled to interfere with personal decision-making and to regulate what is and is not acceptable. A sense of invasion is unavoidable. Offspring may or may not applaud their biological and social parents' choices and actions. Indeed, concepts of belonging and identity will likely be ever more challenged and challenging as more and more reproductive choices or offered to individuals and couples. As the choices burgeon, individuals may feel ever more compelled to seek guidance from authorities and those in the helping fields. This means that the knowledge and scientific expertise needed for those who write laws, interpret supposedly normative teachings, and provide guidance will increase. The medical insurance system is ill equipped to deal with the reproductive options that are currently being offered to prospective patients. It is difficult to see how an insurance company will be have a nuanced appreciation of what constitutes medical necessity or coverage eligibility when reproductive options are seen to be departures from convention. Controlled reproduction has already refined the understanding of reproductive wellness and reproductive rights, but access has not kept pace with the options. It will be interesting to see how various cultures grapple with ideas about the role of women in reproduction and family building in the face of population expansion in some parts of the world and contracture in others. Similarly, scientists, physicians, and health care

workers will need a deeper and more sensitive appreciation of the cultural effect and implications emerging from the promise and peril of expanding reproductive options. The issue of stem cell research has revealed clearly the potential for polarization, fear, and the imposition of prohibitions on research. It will take true societal wisdom to harness the promise of molecular medicine for reproductive and women's health in a manner that protects the vulnerable and satisfies critics and regulatory authorities. Science literacy will continue to be critically necessary, especially for those involved in policy development and application.

REFERENCES

Adler NE, David HP, Major BN, Roth SH, Russo NF. Psychological responses after abortion. *Science*. 1990;248:41–44.

Arnold AP, Xu J, Grisham W, Chen X, Kim YH. Minireview: sex chromosomes and brain sexual differentiation. *Endocrinology*. 2004;145:1057–1062.

Berga SL. Disorders of gonadotropin secretion. In: Shalet S, Wass J, eds. *Oxford Textbook of Endocrinology*. Oxford: Oxford University Press; 2002:1119.

*Berga SL, Marcus MD, Loucks TL, Hlastala S, Ringham R. Recovery of ovarian activity in women with functional hypothalamic amenorrhea who were treated with cognitive behavior therapy. *Fertil Steril*. 2003;80:976–981.

Byatt N, Deligiannidis KM, Freeman MP. Antidepressant use in pregnancy: a critical review focused on risks and controversies. *Acta Psychiatr Scand*. 2013;127(2):94–114.

Cohen LS, Altshuler LL, Harlow BL, et al. Relapse of major depression during pregnancy in women who maintain or discontinue antidepressant treatment. *JAMA*. 2006;295:499–507.

Cohen LS, Friedman JM, Jefferson JW, Johnson M, Weiner ML. A reevaluation of risk of in utero exposure to lithium. *JAMA*. 1994;271:146–150.

Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. *Br J Psychiatry*. 1987;150:782–786.

Cyranowski JM, Frank E, Young E, Shear MK. Adolescent onset of the gender difference in lifetime rates of major depression: a theoretical model. *Arch Gen Psychiatry*. 2000;57:21–27.

Finkelstein JW, Susman EJ, Chinchilli VM, et al. Estrogen or testosterone increases self-reported aggressive behaviors in hypogonadal adolescents. *J Clin Endocrinol Metab*. 1997;82:2433–2438.

Gaynes BN, Gavin N, Meltzer-Brody S, et al. Perinatal depression: prevalence, screening accuracy, and screening outcomes. *Evid Rep Technol Assess (Summ)*. 2005;119:1–8.

Goldstein JM, Seidman LJ, Horton NJ, et al. Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cerebral Cortex*. 2001;11:490–497.

Heldring N, Pike A, Andersson S, et al. Estrogen receptors: how do they signal and what are their targets. *Physiol Rev*. 2007;87:905–931.

Hollander LE, Freeman EW, Sammel MD, Berlin JA, Grisso JA, Battistini M. Sleep quality, estradiol levels, and behavioral factors in late reproductive age women. *Obstet Gynecol*. 2001;98:391–397.

Huybrechts KF, Palmsten K, Avorn J, et al. Antidepressant use in pregnancy and the risk of

cardiac defects. *N Engl J Med*. 2014;370(25):2397–2407.

Kapfhammer HP, Reininghaus EZ, Fitz W, Lange P. Clinical course of illness in women with early onset puerperal psychosis: a 12-year follow-up study. *J Clin Psychiatry*. 2014;75(10):1096–1104.

LeBlanc ES, Janowsky J, Chan BK, Nelson HD. Hormone replacement therapy and cognition: systematic review and meta-analysis. *JAMA*. 2001;285:1489–1499.

Marchbanks PA, McDonald JA, Wilson HG, et al. Oral contraceptives and the risk of breast cancer. *N Engl J Med*. 2002;346:2025–2032.

McEwen BS. Clinical Review 108: the molecular and neuroanatomical basis for estrogen effects in the central nervous system. *J Clin Endocrinol Metab*. 1999;84:1790–1797.

Meador KJ, Baker GA, Finnell RH, et al; NEAD Study Group. In utero antiepileptic drug exposure: fetal death and malformations. *Neurology*. 2006;67:407–412.

*Munk-Olsen T, Laursen TM, Pedersen CB, Mors O, Mortensen PB. New parents and mental disorders: a population-based register study. *JAMA*. 2006;296:2582–2589.

Neugebauer R, Kline J, Shrout P, et al. Major depressive disorder in the 6 months after miscarriage. *JAMA*. 1997;277:383–388.

Parry BL. Perimenopausal depression. *Am J Psychiatry*. 2008;165:23–27.

Person ES. *Dreams of Love and Fateful Encounters: The Power of Romantic Passion*. New York: Norton; 1988.

Ross LE, Grigoriadis S, Mamisashvili L, et al. Selected pregnancy and delivery outcomes after exposure to antidepressant medication: a systematic review and meta-analysis. *JAMA Psychiatry*. 2013;70(4):436–443.

Sandman CA, Davis EP, Buss C, Glynn LM. Exposure to prenatal psychobiological stress exerts programming influences on the mother and her fetus. *Neuroendocrinology*. 2012;95(1):7–21.

Schmidt PJ, Haq N, Rubinow DR. A longitudinal evaluation of the relationship between reproductive status and mood in perimenopausal women. *Am J Psychiatry*. 2004;161:2238–2244.

*Schmidt PJ, Nieman LK, Danaceau MA, Adams LF, Rubinow DR. Differential behavioral effects of gonadal steroids in women with and in those without premenstrual syndrome. *N Engl J Med*. 1998;338:209–216.

Seeman TE, McEwen BS, Singer BH, Albert MS, Rowe JW. Increase in urinary cortisol excretion and memory declines: MacArthur studies of successful aging. *J Clin Endocrinol Metab*. 1997;82:2458–2465.

Taylor SE, Klein LC, Lewis BP, Gruenewald TL, Gurung RA. Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. *Psychol Rev*. 2002;109(4):745–750; discussion 751–753.

Weissman MM, Wickramaratne P, Pilowsky DJ, et al. Treatment of maternal depression in a medication clinical trial and its effect on children. *Am J Psychiatry*. 2015;172(5):450–459.

Wisner KL, Sit DKY, Moses-Kolko EL. Antipsychotic treatment during pregnancy and lactation. *Adv Schizophr Clin Psychiatr*. 2007;3(2):48–55.

Yaffe K, Krueger K, Cummings SR, et al. Effect of raloxifene on prevention of dementia and cognitive impairment in older women: The multiple outcomes of raloxifene evaluation (more) randomized trial. *Am J Psychiatry*. 2005;162:683–690.

*Zandi PP, Carlson MC, Plassman BL, et al; Cache County Memory Study Investigators. Hormone replacement therapy and incidence of Alzheimer's disease in older women: the cache county study. *JAMA*. 2002;288:2123–2129.

▲ 31.2 Genetic Counseling for Psychiatric Conditions

HOLLY L. PEAY, PH.D., BARBARA BOWLES BIESECKER, PH.D., AND JEHANNINE C. AUSTIN, PH.D.

Patients with psychiatric conditions and their relatives may seek genetic education and counseling to understand the origins of the condition(s) in their family and the chances of recurrence. Genetic counseling about psychiatric conditions has been provided by professional genetic counselors for several decades, and interest in the service is widening. Increased interest may be attributed to refinement in diagnosis, enhanced understanding of genetic and environmental origins, improved treatment options, and enhanced focus on risk modification and management. Increasingly, requests for genetic counseling may be made directly to psychiatrists.

Many psychiatrists are well placed to initiate and respond to requests for genetics education and counseling because they have knowledge of their patients' needs and family histories within ongoing therapeutic relationships. An ideal model for psychiatric genetic counseling is through a multidisciplinary team, with collaboration between genetics and mental health professionals. Genetics professionals are available for consultation with psychiatry regarding advanced risk assessment, collection and construction of complicated family medical histories (FMHs), and the availability and limitations of genomic testing. Precision medicine is evolving to include use of panels of genetic variants and genome sequencing to assess origins of conditions in families where multiple individuals are affected or psychiatric symptoms accompany signs of rare conditions.

CASE STUDIES

AR is a 33-year-old client with bipolar disorder. She discusses how much she would like to be a mother, but she is fearful about her mental health during and after a pregnancy. She is also haunted by the idea that her child would develop bipolar disorder. She asks for expert advice about the advisability of having children.

BN is a 57-year-old client with recurrent major depression. He has two teenaged children and is concerned about their mental health. He asks for advice about how to distinguish between normal teenage moods and a developing mood disorder. He wonders what symptoms to look for and what the appropriate response would be to any emerging symptoms. BN expresses his belief that at least one of his two children is bound to become ill because the illness is so common in his family.

CASE DISCUSSION

Although the situations and queries appear quite different, the underlying issues are similar. Their client's concerns are based on the etiology of psychiatric illness, symptom identification, and risk for recurrence in family members. Mental health professionals likely have clients who express such concerns regularly. These questions can lead to discussions about the cause of psychiatric illness, symptom effects during and after pregnancy, teratogen concerns with psychiatric medications, recurrence risks for asymptomatic family members, and early

intervention and identification efforts—all topics that fall under the umbrella of psychiatric genetic counseling.

Growing use of genomic technologies may also increase the rate of inquiries to psychiatric professionals based on concerns about future psychiatric health. Healthy individuals may use these technologies to learn about genetic predisposition to a range of outcomes, including both physical and mental health. Individuals who use genomic technologies seeking diagnosis for a set of manifesting symptoms may learn of increased risk for psychiatric symptoms as part of a genetic diagnosis. In the future, secondary findings (i.e., a reported result that is not the reason for testing) associated with neuropsychiatric outcomes may be reported to research participants or patients based on whole exome or whole genome sequencing (WES/WGS). These cases represent a new type of client with different education and counseling needs. Clients who are well may perceive the importance of genomic results differently than those who are affected, while those who have a history of mild symptoms may reevaluate past symptoms on the basis of the results. In most cases, genomic results are associated with only a modest increase in risk that should be considered in light of personal and family history. The clinical utility of tests for common psychiatric disorders with (typically) low associated odds ratios and a range of etiological risk factors (see sections that follow) is still in question.

WHAT IS GENETIC COUNSELING?

Genetic counseling is the process of helping people to understand and adapt to the medical, psychological, and familial implications of genetic contributions to disease. The National Society of Genetic Counseling describes the process as integrating the following factors:

- Interpretation of family and medical histories to assess the chance of disease occurrence or recurrence.
- Education about inheritance, testing, management, strategies for mitigating risk, resources, and research.
- Counseling to promote informed choices and adaptation to the risk or condition.

The process aims to minimize distress and facilitate adaptation, to increase one's feeling of personal control, and to facilitate informed decision making and life planning.

Although the term *genetic counseling* suggests a primary focus on the genetic contributions to disease, it is not limited to that domain. Genetic counseling entails risk assessment, education, and counseling for common, complex diseases and necessarily considers both *genetic* and *environmental* components of the presenting illness. The growing

awareness of the complex interplay between genes and environment dictates a comprehensive approach to care.

To facilitate comprehension and application of the material in this section, Figure 31.2–1 presents standard symbols used in pedigree construction and Figure 31.2–2 illustrates a complex FMH presented in the form of a pedigree.

GENETIC COUNSELING FOR PSYCHIATRIC DISORDERS

Many providers consider genetic counseling as a resource when genetic testing may be an option for a patient. Yet, genetic counseling has long been provided prior to development of genetic tests, and current genetic counseling aims to provide an assessment that includes consideration of available testing. Outcomes of psychiatric genetic counseling include a better understanding of etiology and risk in the family, psychological support related to familial risk, a discussion of the limited role of parenting in causing or preventing psychiatric disorders (reflecting current scientific knowledge), early symptom identification and early intervention, and facilitated meaning making and coping.

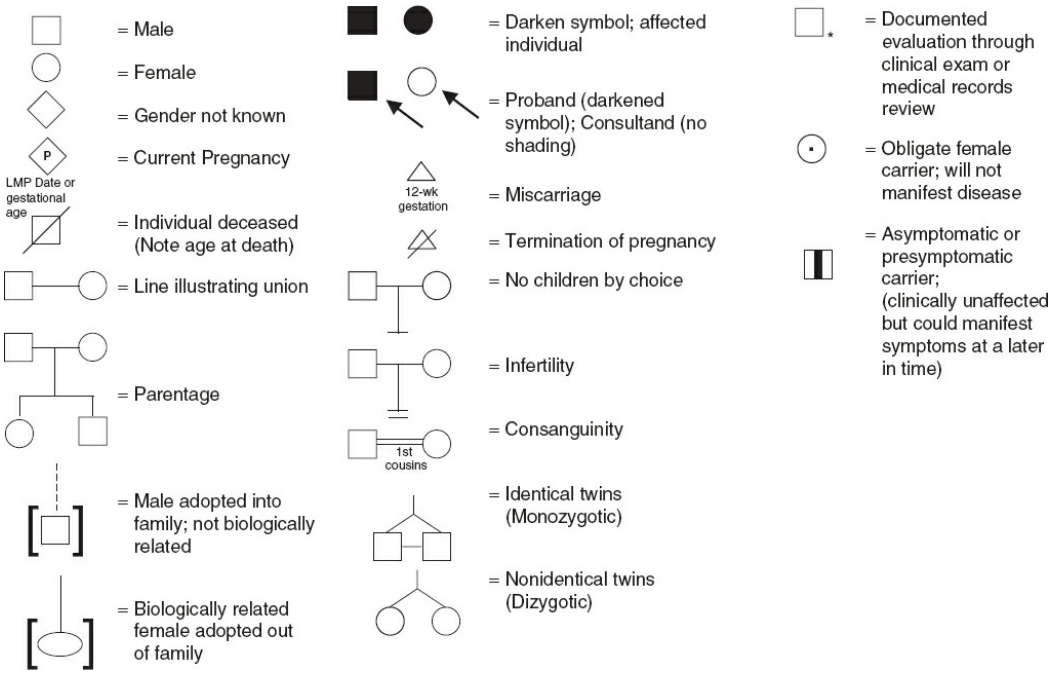

FIGURE 31.2–1. Standard symbols used in pedigree construction. LMP, last menstrual period.

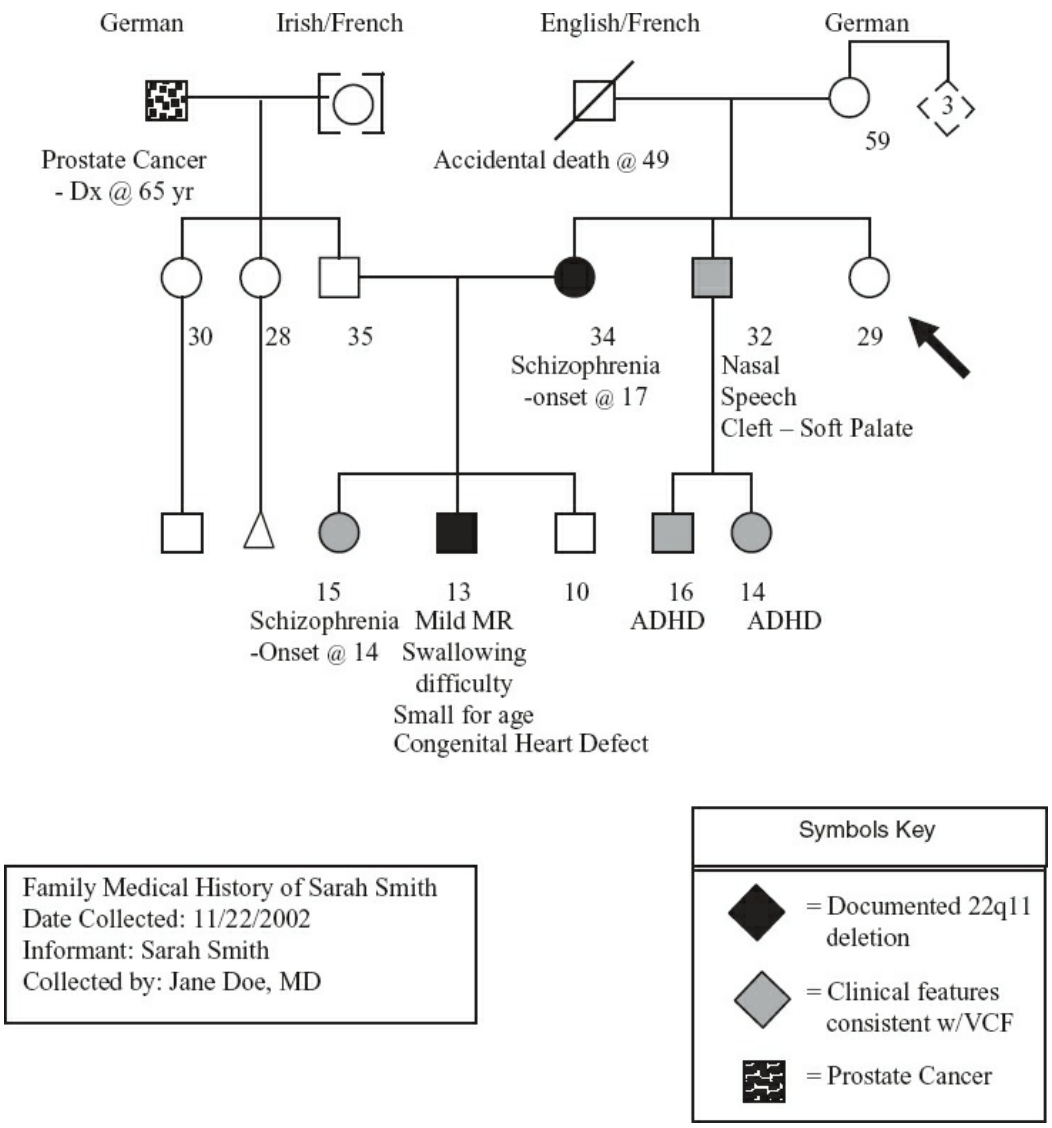

FIGURE 31.2–2. Pedigree of a family with velocardiofacial (VCF) syndrome. ADHD, attention-deficit/hyperactivity disorder; Dx, diagnosed; MR, mental retardation.

Individuals with psychiatric diagnoses and their close relatives often express interest in genetic counseling services, even if they do not have access to highly predictive genomic testing. Clients seek psychiatric genetic counseling for a spectrum of indications. These include decision making around reproduction or adoption, teratogenic concerns about psychiatric medication use, concerns about existing young children, and at-risk individuals' concerns about their mental health. Individuals with extensive family histories of illness or relatives with a severe course of illness may be more motivated to seek counseling, which is appropriate because those families are more likely to be at higher genetic risk. Studies have reported a strong interest in genomic testing, indicating the potential for significant uptake as approaches to clinically and personally useful testing emerge.

Interest in genetic counseling is generally driven by the client's appraisal of psychiatric illness within the family. Several factors that play

into a client's evaluation of risk include the perceived disease burden, beliefs about what caused the disease within the family, existing perceptions of risk to family members, and the ability to control risk. The following situations are specific to psychiatric disorders:

- Individuals affected with bipolar disorder and schizophrenia often perceive moderate to high illness burden, and relatives often perceive high burden.
- Affected individuals and family members often describe the disorders as caused by genetic/biological and environmental factors.
- When considering risk to family members, many close relatives and individuals with bipolar disorder and schizophrenia overestimate the risk for family recurrence.
- There is some evidence that a subset of individuals with mood and psychotic disorders, and a subset of their relatives, are less willing or unwilling to have children because of fear for illness in the children.

During a psychiatric genetic counseling session, the provider should query the client about these factors because they will likely have a strong effect on how clients perceive the information provided and the resulting decisions and life plans that they make.

Knowledge about the etiology of psychiatric disorders is improving steadily. These improvements allow more accurate and specific answers about why a disorder has occurred in a particular individual and family. Likewise, it also increases the provider's ability to offer more accurate assessments of risk and counseling about environmental modifications. These advances in psychiatric genetics are expected to lead to a parallel increase in interest in genetic counseling by individuals with major psychiatric disorders, close family members, and others affected by familial risk. Having an explanation regarding why one developed an illness is critical to the process of adaptation. In the absence of being provided with a comprehensible explanation for the cause of illness, affected individuals will develop their own explanatory model that is important in determining how one behaviorally responds to that illness, including treatment adherence. Fundamentally, studies show that people are more likely to actively participate in treatment if these are seen to directly address what the individual perceives to be the cause of his or her illness (see Inglis et al., 2014).

THE ROLES OF THE GENOME AND ENVIRONMENT IN MENTAL HEALTH

It is widely accepted that psychiatric illnesses tend to run in families. Disorders can recur in families for several reasons, including the functioning of the genome (rarely, due to one genomic variation with a large effect on susceptibility, or more commonly due to multiple variations that each have small effect on susceptibility) and/or shared environmental

exposures. Thus, as described previously, genetic counseling for a complex disorder entails discussions around *both* genetic and environmental risk factors. For example, consider a case in which a father with obsessive–compulsive disorder (OCD) has concerns about symptoms developing in his children. Genetic counseling may focus on the following:

- The family history, indicative of both genetic and shared environmental risk in the family.
- Risk factors for the children, stemming from those genetic and environmental factors.
- Early symptom identification and early intervention (aiming to modify the environment and thus affect genetically mediated disease progression).
- A discussion of the benefits and limitations of environmental modifications aimed at reducing risk or minimizing emerging symptoms.

The concept of complex disease (which has replaced *multifactorial* as an etiological category for common disorders) not only takes into account the interplay of the genome and the environment but also includes the broader biological and physiological mechanisms of evolution and development that operate within societies and cultures. The onset of most cases of psychiatric disease, like the onset of other common disorders, is thought to arise when an individual has several risk-increasing genomic variants and has experienced one or more environmental triggers; overall risk is influenced by development, growth, and maturation. The modern perception that psychiatric symptomatology causes disadvantageous "disease" is related to the effects of the symptoms on the ability of the individual to interact successfully with his or her social and cultural environment.

The complicated nature of the etiology of psychiatric disorders is best understood by considering an illustrative family history, such as shown in Figure 31.2–3. This pedigree illustrates many different aspects of complex disease, including the following:

1. Although the variants predisposing to disorders A and B segregate through the family, the disorders in question aggregate in the family.
   a. Individuals may have all of the necessary genomic susceptibility and never experience any illness, because they lack the necessary environmental exposures.
   b. Individuals may have all of the necessary genomic and environmental susceptibility and never experience any illness due to protective factors (genomic and/or environmental).
2. A different number or complement of susceptibility variants and/or different environmental exposures leads to the existence of two etiologically related disorders in the family, A and B.

3. The typical ages at onset of disorders A and B range from early adolescence through middle adulthood, making it difficult to determine who in the family is truly unaffected.
4. Genomic variants that increase risk are common in the general (unaffected) population.

The first evidence that genetic factors were operating in psychiatric disorders came from analysis of families with multiple affected members, twin studies, and adoption studies. Family studies clearly indicated familial etiology, although the relative contributions of genes and environmental factors were unclear. Twin and adoption studies provided insights into the relative contributions of etiological risk factors, and for all major psychiatric disorders consistently indicated the importance of genes and environment. Family, twin, and adoption studies also provided insight into the genetic heterogeneity and phenotypic variability of mental disorders. Table 31.2–1 provides a summary of family, twin, and adoption data for schizophrenia.

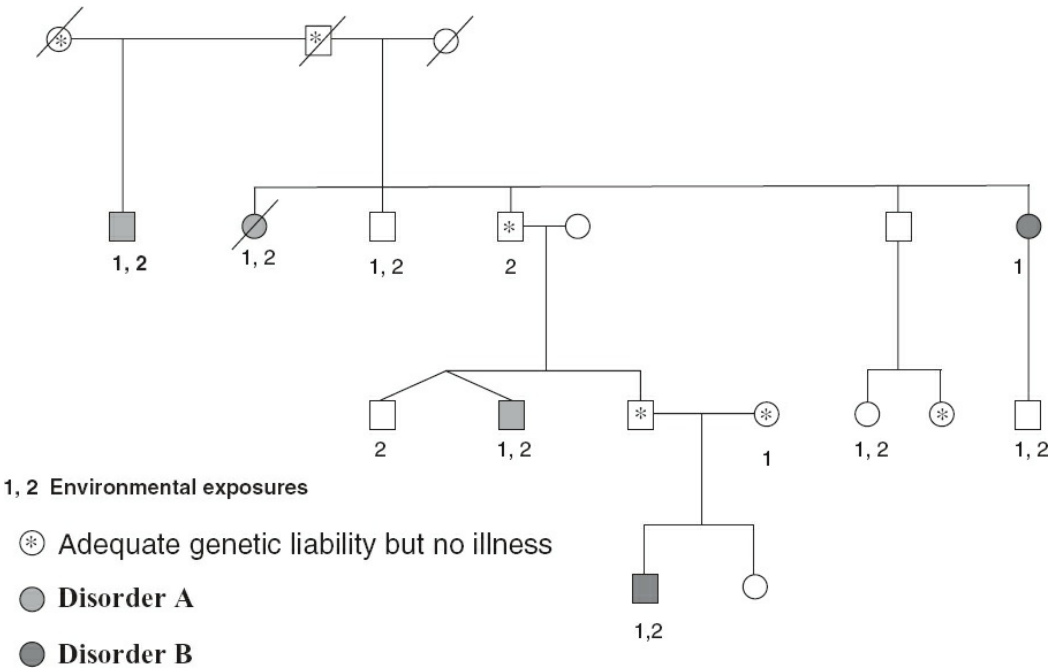

FIGURE 31.2–3. Simulated pedigree displaying the aspects of complex inheritance.

CASE STUDY

Cassie and her partner, John, sought genetic counseling as a couple to understand more about causes of mental illness. John said that he wanted to better understand—as he put it: "why is my sister-in-law's OCD so much more severe that Cassie's?" In discussion, it emerged that he had doubt that Cassie should ever have even been diagnosed with OCD at all. In exploring the couple's existing explanations for cause of illness, it emerged that John perceived environmental factors to be more important in the etiology of OCD. Cassie's sister had a traumatic event before her symptoms became noticeable to the family, while Cassie had experienced nothing similar. After John repeated several times that he felt that Cassie's OCD was not as severe as her sister's, she broke down in tears saying that she felt John minimized/dismissed her OCD, and that she felt that her symptoms were just as severe though

not as overt as her sister's. The clinician was able to help John identify that he had been terrified of thinking that genetics might be involved because of his fears around what this might mean for the children that he and Cassie planned to have together, and it was this that had caused him to deny Cassie's experience. By the end of the session, John and Cassie had come to a better understanding of each other and were developing a plan to work on protective factors together.

CASE STUDY

After Anna (12) received a diagnosis of an eating disorder, she and her mother requested a session to discuss the causes of eating disorders. Both Anna and her mom were very distressed about the diagnosis, and both seemed to attribute the disorder entirely to genetic factors. Anna said: "Its genetic, right? I'm just different … and weird." Anna's mom was feeling helpless, watching Anna as she withdrew from her friends. It emerged that Anna's mom had experienced anorexia as an adolescent and felt profoundly guilty for having "passed on" a condition that she had experienced as life-threatening at a similar age. As the clinician discussed with them the roles of genes and environment in the etiology of eating disorders, relating this back to Anna's life experiences, both mom and daughter expressed feelings of relief. For the mom, the relief was about feeling that perhaps it was not "all my fault," and for Anna, it was about feeling less fundamentally different (flawed) than she had. As the conversation turned toward protective factors, Anna's mom began sharing strategies with her daughter that she had found to be helpful in her own recovery. This provided her with a feeling that she was not entirely helpless and seemed to bring the pair closer together.

Table 31.2–1.
Data from Family, Twin, and Adoption Studies for Schizophrenia

Family Studies	Twin Studies	Adoption Studies
Empirical risk for first-degree relatives: ~5–15% Empirical risk for second-degree relatives: ~2–8% Also increased risk for spectrum disorders (schizoaffective disorder, schizotypal personality disorder, and paranoid personality disorder)	Monozygotic twin concordance: 40–60% Dizygotic twin concordance: 10–16% Heritability: ~80%	Support high heritability; unique environmental exposures play a significant role, whereas the influence of shared environmental exposures is more limited

Table 31.2–2.
Examples of Psychiatric Disorders Recognized as Having a Genetic Component to Their Etiology

Psychotic disorders: Schizophrenia, schizoaffective disorder
Mood disorders: Bipolar disorder, recurrent unipolar depression
Personality disorders: Antisocial personality disorder, schizotypal disorder
Anxiety disorders: Generalized anxiety disorder, obsessive–compulsive disorder, panic disorder, phobia
Substance-related disorders: Substance dependence and abuse
Eating disorders: Anorexia nervosa, bulimia
Childhood disorders: Attention-deficit/hyperactivity disorder, autism, chronic tic disorders, including Tourette syndrome
Memory disorders: Alzheimer disease

GENETIC CONTRIBUTIONS TO PSYCHIATRIC DISORDERS

Thousands of candidate gene, linkage, and genome-wide association studies have provided insight about genetic variants that increase susceptibility to psychiatric illness (Table 31.2–2). New research approaches and a growing understanding of the function of the human genome have opened promising areas of research, including studies of epigenetic phenomenon (heritable changes that do not involve alteration to the DNA sequence) and the evaluation of copy number variants (variation found between two or more genomes in the number of DNA copies at a particular location). A detailed understanding of the causal pathway from any of the identified genomic contributions to disorder susceptibility is not yet clear. It is apparent, however, that changes in the genome do not unequivocally result in major psychiatric disorders. Rather, current knowledge suggests that particular variants may increase the risk for an adverse response to an environmental stimulus or may more directly affect a biological process in a way that facilitates the onset of particular symptoms. As described by the Schizophrenia Working Group of the Psychiatric Genetics Consortium (2014), findings from schizophrenia research suggest that most genomic variants exert an effect through altering gene expression rather than through altering protein structures.

The genomic variants that increase the risk for psychiatric disorders are believed to regulate a range of traits that demonstrate phenotypic variability. Phenotypic variability can be illustrated by considering spectrum disorders, such as the bipolar disorder spectrum, which is a widening of the phenotype to include etiologically related disorders more frequent in relatives of individuals with bipolar disorder. Because of varying interactions among genome variants and environmental factors, individuals with a biological family member with bipolar disorder are more likely to have bipolar disorder, major unipolar depression, or cyclothymia.

Susceptibility variants identified to date provide support for the prediction that many risk variants will be of small effect and may be common in the unaffected general population. For example, SNPs associated with bipolar disorder and schizophrenia have odds ratios in the range of 1.1 to 1.3. Less common structural mutations, de novo copy number variants, are thought to be present in about 5 percent of individuals with bipolar, schizophrenia, and autistic spectrum disorders; they are estimated to be associated with an absolute risk of ~15 percent (see Gershon & Alliey-Rodriguez, 2013). It is important to keep in mind that many individuals with genomic susceptibility will never develop psychiatric illness. Thus, one can appreciate the challenges in using variants to predict onset of disease.

ENVIRONMENTAL CONTRIBUTIONS TO PSYCHIATRIC DISORDERS

In disorders with a complex etiology, including most instances of psychiatric disorders, environmental factors are necessary for disease onset. In this context, "environment" is a catchall phrase meaning anything outside of the genome that can affect the expression of a gene, including the in utero environment during development, shared and nonshared family environment, exposures to harmful substances, and medical history.

Adoption studies have indicated that the strongest environmental risk factors are unique factors, meaning factors that make siblings less alike. Many misinterpret this to mean that home environment is not important in the development of psychiatric illness because shared home environment should make siblings more alike. But home environment and parenting style are not equal for all siblings and may not be stable over time. Outside effects on the home environment, sibling order, sibling relationships, and differences in sibling personality traits (themselves genetically mediated) have effects on the home environment. Even if the home environment is identical, siblings likely perceive differences in their shared environment on the basis of their individual developmental history and personality traits, both of which are genetically mediated. In some cases, the perception of differences may be as important as actual differences. Thus, the identification of unique factors as the most important environmental contributors does not exclude the relevance of home environment in the etiology of psychiatric illness. It is likely that, as research capabilities increase, providers will be able to provide clients with a more nuanced account of the role of child rearing and home environment in the development of psychiatric illness. However, it is important for parents to know that in most cases, home environment is unlikely to be sufficient to either cause or prevent psychiatric illness in children.

Unique environmental risk factors are unequivocally important in the development of psychiatric illness. Factors that may or are likely to play a role in schizophrenia and bipolar disorder include the following:

- Obstetrical complications
- Substance use, specifically cannabis and methamphetamine
- Childhood trauma
- Being reared in an urban environment
- Advanced paternal age
- Immigration
- Maternal prenatal infections
- Season of birth
- Head injury
- Stressful life events
- Viral infection

Unfortunately, many of these factors are pervasive and largely outside

of an individual's control.

The interplay between genomic and environmental risk factors is complicated; for example, depression is expected to be caused through an interaction of genomics and environment, such that susceptibility variants are expressed in an adverse environment but not a supportive one. Although the data are limited, there is evidence to support what is likely a widely held perception of health care providers: Persons at highest risk are likely to be those who have both genetic liability and vulnerability in terms of their environment.

Environmental Triggers

In the underlying disease etiology (i.e., factors combining with genetic susceptibility to influence onset), environmental risk factors are considered separately from environmental triggers that have been shown to exacerbate symptoms in someone who is already affected. Although there may be overlapping environmental risk factors and symptom triggers, the contrast between the two is important due to issues of timing and behavioral control.

FAMILIES AT HIGH RISK (EXCEPTIONS)

Matters are further complicated by the presence of families in which a combination of etiological factors for psychiatric illness are skewed in a fashion that mimics a Mendelian pattern of inheritance (autosomal dominant or recessive), with multiple affected individuals in one or more generation (Fig. 31.2–4). When evaluating such "loaded" families, it is important for the provider to consider the potential causes. Besides presenting unique psychological needs, highly loaded families have higher risks for recurrence, and genomic testing may be considered. Psychiatrists who are not highly experienced in genetics should refer such individuals and families to genetics professionals for evaluation and counseling.

The skewed presentation of psychiatric illness in selected families may be caused by a number of colliding factors that mimic single-gene inheritance. Some examples follow.

Complex Etiology with Unusual Accumulation of Risk Factors

It is possible that families with numerous affected individuals might have an unusual number of susceptibility variants segregating within the family, such that many family members have inherited multiple susceptibility-causing variants. One mechanism for this is assortative mating, in which individuals with psychiatric illness have children with other individuals with psychiatric illness. Such families may also experience recurrent shared environmental exposures that trigger onset of the disorder. In these cases, the risk to other family members might be high, although not currently

reliably quantifiable.

Genetic Variants of Large Effect

Whereas the majority of individuals with psychiatric disorders are expected to have a complex etiology, some families may segregate genetic variants of large effect, which cause an increased risk for psychiatric illness in the absence of any other phenotype (while the presence of any other phenotype would suggest a possible syndrome). Although these variants would likely exhibit reduced penetrance and variable expression, the risk to others in the family would likely be greatly increased over empirical risks but less than 50 percent.

Syndromic

Although overall they are rare, there are a number of genetic syndromes that have significant psychiatric manifestations, along with other features. Examples include Huntington disease, velocardiofacial (VCF) syndrome (also called 22q deletion syndrome), Wilson disease, Wolfram syndrome, late juvenile- or adult-onset metachromatic leukodystrophy, homocystinuria, acute intermittent porphyria, coproporphyria, frontotemporal dementias, and early-onset familial Alzheimer disease. These conditions provide examples in which the presence of a genetic variation responsible for the disorder increases the risk (although typically less than 100 percent) for the development of psychiatric disorders compared to the general population.

22q deletion syndrome is an example of an inherited disorder with significant variability and increased risks for psychiatric disorders. It is an autosomal dominant microdeletion syndrome localized to the long arm of chromosome 22 (22q11). The clinical spectrum is variable and may include clefting of the soft palate, heart defects (most commonly tetralogy of Fallot, ventricular septal defect or right aortic arch), velopharyngeal incompetence, immune dysfunction, mild intellectual impairment, and psychiatric disorders, although rarely does an individual have all of the symptoms. Approximately 30 percent of individuals with 22q deletion experience psychotic disorders; risk is also increased for mood disorders. Onset of psychotic symptoms may be earlier than typical. Attention-deficit/hyperactivity disorder (ADHD) is also a relatively common characteristic in affected individuals. An FMH from the following case study illustrating the variability of expression of 22q deletion syndrome is shown in Figure 31.2–2.

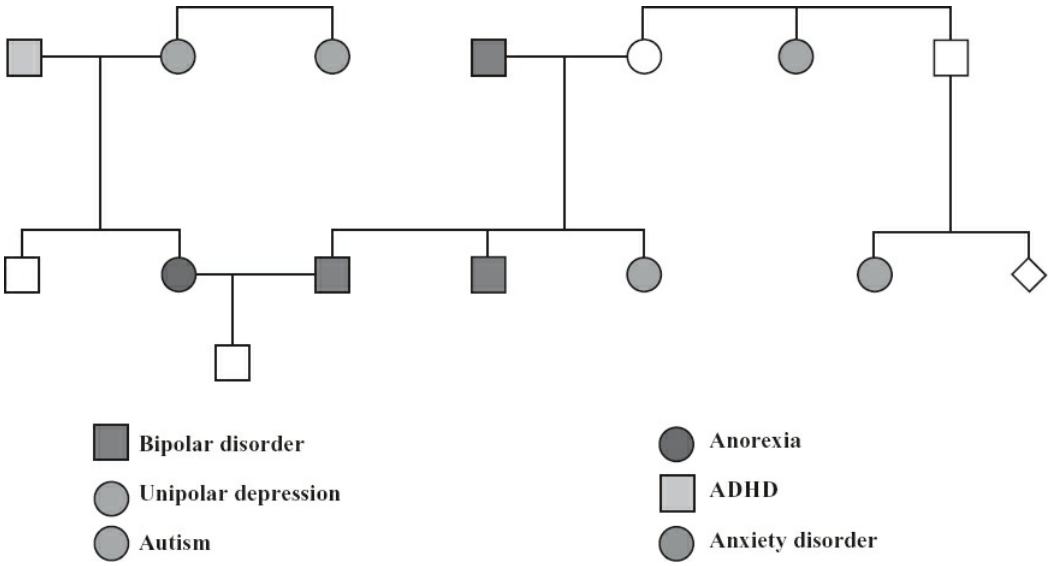

FIGURE 31.2–4. High-risk family. ADHD, attention-deficit/hyperactivity disorder.

CASE EXAMPLE

A 29-year-old woman contacted a local psychiatrist for genetic counseling regarding her family history of schizophrenia and cognitive impairment. She was concerned that her future children might experience problems similar to those of her sister, nephew, and niece. The psychiatrist collected the FMH and constructed the pedigree. To confirm the diagnoses and to gain further information, the psychiatrist discussed the potential benefits and risks of obtaining medical records. The client decided to pursue this option and obtained permission from her affected sister. The psychiatrist received records on the affected sister and her 13-year-old son, who had mild intellectual disability. The records on the boy indicated that he had undergone genetics evaluation and testing and been identified as having a microdeletion on the long arm of chromosome 22, consistent with 22q deletion (VCF) syndrome. The records indicated that the client's sister also carried the deletion, although this information had not yet been shared with the family.

The identification of 22q deletion syndrome in this family offered the option of genetic testing for others in the family. The 22q microdeletion is estimated to arise spontaneously in 80 to 90 percent of cases. However, this pedigree was suggestive of a hereditary form. Counseling in such situations is difficult owing to the variability of the disorder and the inability to predict the phenotype (e.g., the client may have the deletion and have a very subtle presentation). Family studies are warranted to attempt to clarify recurrence risks; however, issues regarding confidentiality of the acquired information require attention.

PRIORITIES OF PSYCHIATRIC GENETIC COUNSELING

Requests for genetic counseling are often initiated on the basis of the patient's or relatives' questions about the condition present in the family that are posed to the treating psychiatrist. The provision of information on the potential genomic origins of the condition and the options for genome testing or assessment of the family can be conceptualized as shared decision making. Risk assessment and genome testing are dictated by patient preference. How much the patient prefers to pursue and for what purpose can be weighed against the estimated likelihood that useful information will come from testing. The basic components of genetic counseling are outlined in Table 31.2–3.

Genetic counseling can be facilitated by the psychiatrist, by a genetics professional, or both. Ascertainment and clarification of the client's specific questions and expectations of the session promote understanding of the patient's concerns. A two-way exchange of information with respect to the client's understanding of the problem and diagnosis, expectations for the session, and participation in the choice and discussion regarding future plans or options is optimal (Table 31.2–4).

Educating Clients about Genetics

By its very nature, genetic information is complex and often technical, with a vocabulary that is not familiar to clients. Traditionally, the professional selects the initial core of information presented to the client. Significant effort should be made to identify and to include the information considered most relevant by the client and to identify and address the emotional and psychological issues that discussing this information can raise. The topics that are most often included are a description of the disorder and its natural history, explanation of the possible modes of inheritance and possible implications for family risk, discussion of environmental risk factors, options for early intervention, medical management, prevention or treatment, and (in the cases in which it is applicable) the availability of genomic testing and the associated risks, benefits, and limitations. Owing to the complexity of the information and discrepancies in client learning styles, providers may use a variety of means to communicate the information, including written material; graphics that visually present genetic terms, concepts, and modes of inheritance; and a detailed letter summarizing the clinical encounter; as well as other available resources (support groups and Internet resources). Discussions about etiology can raise emotional and psychological issues related to guilt, shame, and stigma. Thus, the provision of understandable and meaningful information should be coupled with empathy, compassion, and sensitivity to provide a multifaceted genetic counseling encounter.

Table 31.2–3.
Steps and Process of Genetic Counseling

Solicit and clarify the client's presenting questions and goals for genetic counseling.
Explore client's existing explanation for cause of illness.
Collect and review medical and family medical history.
Identify support systems.
Verify diagnoses in proband and other affected family members when possible.
Address issues and concerns identified through the genetic counseling process.
Assess the client's emotional and intellectual capacity before proceeding to determine the approach to the provision of education and counseling.
Provide information at cognitively appropriate levels.

Note that the processing of emotional reactions is intertwined with the provision of information.

Assess the personal meaning of the information and the client's willingness to negotiate various risks and burdens.

When applicable, assist the client in arriving at a decision by discussing available options; discuss benefits and limitations of each alternative.

Assist the client with adapting to the risk status in the family.

Assist in formulating a plan that the client is able to carry out.

Provide follow-up counseling and support. Continue to assess the client's and family's understanding of the information and the effect of the information, risks, or decisions.

Periodic follow-up provides opportunities to address issues as they present to the client. The extended nature of such contact provides optimal opportunity to assist the client or the family member in receiving information when it is most relevant, as well as facilitating their use of the information and adaptation to its implications.

Table 31.2–4.
Goals of a Genetic Counseling Interaction: Concerns about Recurrence in Children

Client's Goals	Psychiatrist's Goals
Reassurance that children are well	Evaluate children's development
Guidance about changes to the home environment that might reduce risk	Discussion of limited ability to reduce risk, and resulting psychological effects
Guidance about the likely magnitude of risk to children	Evaluation of likely magnitude of risk to children, understandable discussion of the risk, and psychological support
Guidance about early warning signs	Education about early symptoms and appropriate indications for a psychiatric evaluation

In some cases, genomic testing may inform risk, especially in cases with unusual presentation (i.e., highly loaded families; psychiatric symptoms manifesting along with dysmorphic features, birth defects, and/or learning difficulties that are not associated with the psychiatric phenotype). Consideration of genomic testing should include a careful review of potential benefits, limitations, and risks. A genetic counselor with expertise in psychiatric conditions should be consulted to confirm the appropriateness of genomic tests that may be pursued.

Obtaining the Data Needed for Risk Assessment

The FMH is used to assess the likely etiology of the disorder and estimate risks for recurrence in the family. Confirmation or clarification of the diagnosis or diagnoses in the family is essential to the provision of valid information.

Topics that are specific to psychiatric disorders and are important to include in the FMH are listed below.

- Any psychiatric diagnosis, whether childhood, adult, postpartum, or geriatric
- Age at symptom onset and at diagnosis
- Subjective assessment of illness severity and treatment response
- Potentially important environmental exposures (e.g., obstetrical complications, cannabis use, head injury)
- Symptomatic, undiagnosed individuals and/or undiagnosed individuals treated with a psychiatric medication
- Developmental history (e.g., has the individual achieved normal milestones for age, such as independent living and employment for adults?)
- Social history
- Substance abuse
- Suicide and self-harm
- Birth defects, intellectual disability or learning disabilities, and unusual medical conditions
- The age and sex of at-risk family members

As mental health providers will be aware, challenges to collecting the psychiatric FMH may include guilt and stigma that limit knowledge of diagnoses among family members; the frequency of undiagnosed psychiatric illness; and the dynamic nature of emerging psychiatric symptoms that may result in diagnostic change over time.

Risk Assessment Based on Family and Medical History

In most cases, risk assessment is based on FMH in the absence of meaningful genomic test results. FMH data are used together with empiric risks. It is vital to understand that empiric risks are not specific to any one individual or family but rather are risks obtained from studies that aggregate recurrence risk among large numbers of families. Factors specific to the proband, the family, and at-risk individuals allow for tailoring of empirical risks. At its most basic, such tailoring requires considering the following series of questions.

Questions about the affected individual(s) in the family:

1. Who will be considered to be affected in the family?
   - Many individuals who meet criteria for a major psychiatric disorder are not diagnosed or treated, many are incorrectly diagnosed, or the diagnosis may not be known by the client.
   - The typical late-teenage/young-adult onset of many major psychiatric disorders confounds pedigree assessment because it is possible that individuals who are genetically susceptible and will become symptomatic have not yet experienced the onset of illness.

2. Are illnesses in the family etiologically related?
   - ► Heterogeneity within diagnostic groups and the occurrence of comorbidity make it more difficult to evaluate which disorders might be etiologically related in a family. The high prevalence of psychiatric disorders confounds the issue because some illnesses might be unrelated to other illnesses in the family.
   - ► Generally, disorders that share a spectrum and disorders that are frequently comorbid are considered to be etiologically related. It is difficult to evaluate the relationship of other groups of disorders, although family studies have indicated that, on a research population level, no psychiatric diagnosis occurs in a vacuum, that is, if a relative has one psychiatric illness, family members are at empirically increased risk for a range of other disorders as well.
   - ► Assortative mating, whether between a couple with concordant or discordant diagnoses, complicates the assessment of etiological relationship.
3. What factors in the family history may raise or decrease risk?
   - ► Families in which there are multiple affected individuals over several generations, or an unusual clustering of illness in one sibship, likely have the highest chance for recurrence.
4. What factors in the affected individual's psychiatric history may raise or decrease risk?
   - ► Age at onset: If age at onset is younger than expected, the risk to relatives may be increased. Relatives likely have an increased risk of experiencing early onset as well.
   - ► Severity: More severe illness, and illness that is treatment resistant, may be associated with increased risk to relatives.
   - ► Sex: Depending on the psychiatric disorder in question, the sex of the affected individual may increase or reduce the risk to relatives, although the data are mixed. For example, siblings of females with schizophrenia have been shown in some studies to have higher risk than do siblings of males with schizophrenia.
   - ► Environmental exposures: Relatives of affected individuals with known significant environmental exposures may have a reduced risk of the disease if they do not share the same environmental exposure. For example, relatives of an individual with psychosis who has a history of crystal methamphetamine use may have a lower risk of illness (assuming that they are not exposed to the drug). However, consideration of environmental factors is difficult to apply when providing quantitative risk assessment.
   - ► Birth defects or a developmental/medical history suggesting an

underlying genetic syndrome potentially increases the risk to family members.

5. What etiology best explains the illness in the family? What are alternate explanations?

   ► Although research suggests that the vast majority of psychiatric illness is complex in etiology, at this time it is not possible definitively to rule out Mendelian or nontraditional modes of inheritance in any particular family. When family patterns suggest the possible involvement of a genetic variant associated with major effect, risks often exceed the empirical risk range. In many cases, such "loaded" families do not actually reflect the actions of a genetic variant of major effect, but regardless of cause, such a pattern in the family is indicative of increased risk.

Questions about the at-risk individuals in the family:

1. For what disorder(s) might the individual(s) be at risk?

   ► In the case of psychiatric spectrum disorders, it may be difficult to evaluate *for what specific disorder* family members are at risk. Providers should familiarize themselves with appropriate spectrum disorders based on the at-risk individual's family history. Family members may be at increased risk for comorbid disorders as well.

2. What factors about the individual may raise or decrease risk?

   ► Current age: At-risk individuals who are young have all of their years of greatest disease risk ahead of them, whereas older individuals have "lived through" a portion of their risk. Empirical risks generally reflect lifetime risk and can be qualitatively modified to reflect the at-risk individual's age.

   ► Past psychiatric and developmental history: Individuals with existing psychiatric symptoms, another psychiatric disorder, or evidence of emerging psychiatric symptoms are at greater risk. Individuals with usual developmental histories (e.g., unusual social relationships; prolonged lack of gainful employment) may have a greater risk.

   ► Environmental exposures: Exposure to known environmental triggers may increase risk, although such exposures are currently difficult to identify and evaluate.

   ► When evaluating any family history, providers should be aware that the presence of any of the following "red flags" may significantly increase risk. The identification of any of these factors should call into question the use of a standard empirical risk range.

   ► Learning disabilities or intellectual disability (preceding/unrelated to cognitive deficits due to psychiatric phenotype)

- Birth defects in the client or family members, most notably heart defects or oral clefts
- Dysmorphic facial features
- Unexplained major medical problems
- Neurological impairment not explained by the psychiatric diagnosis or medication side effects
- Very early onset of psychiatric symptoms (e.g., schizophrenia during childhood)
- Highly loaded family history (e.g., multiple individuals in several generations affected with psychiatric disorders; diagnoses may vary)

Figure 31.2–5 provides an overview of risk assessment for a family at moderate risk.

There was no relevant family history on the maternal side of the family. Ages at onset and illness severity were typical in the family, and no obvious environmental risk factors were reported. The empirical risk for bipolar disorder for a child born to this couple (reflecting one affected parent) is between 5 and 20 percent. However, this risk does not incorporate the distantly related paternal cousin who also carries the diagnosis. Therefore, risks at the lower end of the range provided may underestimate the risk for offspring of this couple. The empirical risk for this child to develop any major affective disorder is between 20 and 30 percent (again, the family history suggests that the lower end of the range may underestimate risk). The history of depression in the paternal great-aunt is difficult to evaluate because depression may be part of the bipolar disorder spectrum but is very common in females without family liability to bipolar disorder.

Risk Assessment Based on Genomic Testing

Increasingly, results from genomic testing can be incorporated into the risk assessment. Because most known changes to the genome result in small odds ratios for psychiatric disorders, research teams are exploring use of risk scores based on aggregated data about the individual's genome. Researchers are also assessing the addition of data from family and environmental history to boost risk prediction from the genomic risk score. In rare cases, variants in one gene may increase risk considerably (as in the case of 22q deletion syndrome). Given the complexity and uncertainty that come with genomic testing, the pretest counseling, ordering, interpretation, and communication should be done by experienced researchers and clinicians.

Risk Communication

Different clients seek different types of risk counseling, and they should be queried about their preference before risk communication begins. Some

may seek broad information about the underpinnings of risk for psychiatric disorders, while others prefer a qualitative discussion of the risk magnitude. These individuals may not welcome or benefit from receiving numerical risk information. For others who prefer a detailed numerical estimation of risk for recurrence, risks based on FMH should be presented as a range to reflect the ambiguity of the estimates. In some cases, it may be possible to suggest that one end of the range is more appropriate for a specific client than midrange.

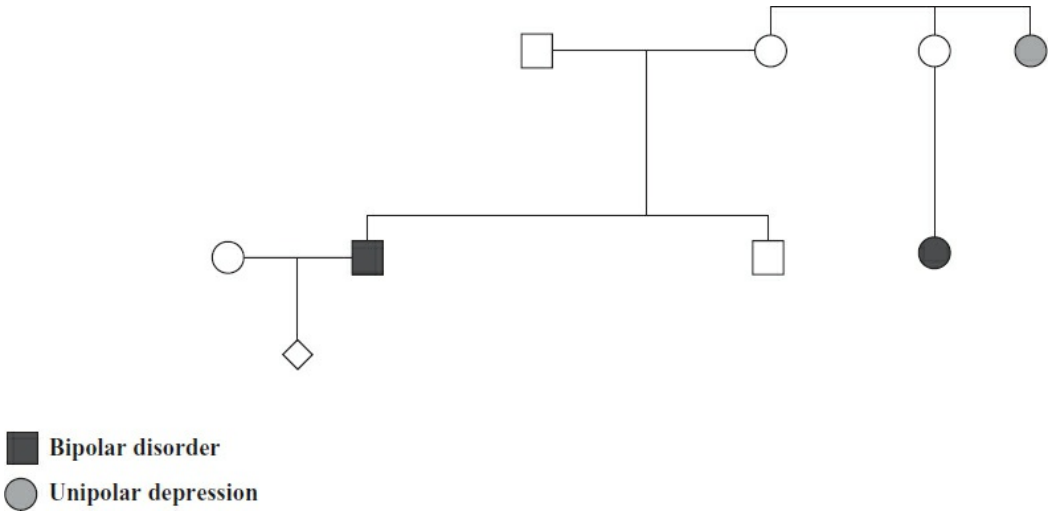

FIGURE 31.2–5. Pedigree of a family at moderate risk for affective disorder.

Perceptions of high illness burden and limited control over onset are likely to result in less tolerance for familial risk. Clients often have preconceived ideas about his or her (or other family members') risk or affected status that may be influenced by their perceptions, such as, similarities to affected relatives or family and cultural beliefs. As such, risks as provided may differ from what is expected. Individuals vary in their capabilities in understanding risk. They frequently perceive risks from a dichotomous, all-or-nothing perspective (it is or it is not going to happen); therefore, the presentation of risks in a numerical fashion (1 chance in 10, or 10 percent) may be difficult for many clients to process. The baseline for risk provision should be absolute risks. Relative risks may be interpreted incorrectly, especially when the disorder is rare (such as autism). Once clients understand the absolute risk, in some cases, it may be useful to clarify understanding by discussing relative risks.

Some individuals do not wish to be given a risk number but instead prefer a qualitative assessment of risk. The majority of families have only one affected family member, and thus empirical risks often are in the range of 10 to 15 percent or lower. For those families, qualitative assessment might be posed as follows: "You are right to think that your children have an increased chance of having [the disorder(s) in question]. However, your children have a better chance of *not* having [the disorder(s) in questions]

than having it." The meaning of descriptive risks is highly subjective and patients should be consulted about their understanding.

Finally, it is important to remember that it is the client's perception of risks (perceived or presented), rather than the actual risk, that is more likely to affect future decisions and their outcomes. A client's perceived estimate of the likelihood that an event will occur is affected by his or her ability to recall or visualize examples of the occurrence and by his or her perceptions of the nature of the event (i.e., severity).

As described previously, when risk communication includes the return of genomic test results, the authors suggest referral to a genetics professional to integrate FMH and results and communicate overall risks.

PERSON-CENTERED COUNSELING

Genetic counseling integrates psychological and social issues that result from living with a psychiatric condition in the family with FMH and increasingly will also include genomic risk information. Genetic counseling has historically been described as nondirective, which is deemed particularly appropriate within the context of reproductive decisions and presymptomatic testing for untreatable disorders, such as Huntington disease. However, nondirectiveness is difficult to practice and has been replaced with a person-centered approach as health benefits are realized by improved treatments and risk behaviors.

The following case exemplifies personal experiences and meaning of psychiatric disorders:

A couple in their mid-30s with a 10-year history of infertility had been trying to adopt a child for a number of years. Recently, the adoption agency they were working with told them of a baby who was being placed for adoption because the biological mother was affected with bipolar disorder and did not feel that she could provide adequate care for the baby. The FMH collected on the newborn baby did not identify others in his family with psychiatric disorders. The recurrence risk for bipolar disorder to the newborn was, therefore, estimated to be between 5 and 20 percent, with additional risks for other psychiatric disorders (Table 31.2–5). The couple individually reacted quite disparately to the estimated risks. In attempting to help them clarify the factors contributing to their feelings regarding the risks, the husband shared his experience with a childhood neighbor who had "some kind of mental illness" and detailed the "torment and agony" that the child brought to the family. Retorting, the woman shared the fact that her coworker also had bipolar disorder and did "just fine" at work with the help of medication. She, therefore, did not feel that the risks for psychiatric disorders were of concern. The psychiatrist facilitated the couple's discussion of the spectrum and meaning of mental illness, along with recurrence risks in the context of a genetic education and counseling session. Although the couple did not come to agreement at that meeting over the potential for adopting the child, they did feel that the information and sharing of experiences and perspectives about psychiatric disorders were beneficial. They agreed to return in 1 week after further considering the issues in an effort to reach a decision regarding the adoption.

Individuals, couples, and families seeking genetic counseling pertaining to psychiatric illness who experience greater feelings of stigma, higher perceived burden, and higher perceived risk for recurrence in offspring or

siblings are likely to require more intensive counseling around family vulnerability. Whereas issues of stigma associated with mental illness are commonly encountered by the client and addressed by the practitioner within the psychiatric setting, additional efforts to assess stigma pertaining to a genomic etiology seem prudent. Likewise, some clients may experience a higher perceived burden due to the recognition of a genomic contribution and potentially feel that they have less control in preventing or treating disease. Communication of information to interested relatives is appropriate and encouraged within the process of genetic counseling. Invitations to family members to receive information related to the disease in question, the etiological underpinnings, recurrence risks, and counseling to address the related psychological or emotional issues are encouraged.

Counseling about Risk Reduction in Children

Clients with children or those considering having children often benefit from counseling specifically focused on management of uncertainty and the potential for risk reduction. Clients are often keenly interested in information about environmental risk because it may facilitate the placement of blame away from parents while at the same time providing potential interventions that may reduce the risk for their children. Given the nature of risk factors, this counseling often focuses on their limited ability to reduce risk of onset of illness in at-risk individuals. However, there are potential targets for intervention. At-risk individuals have been found to be more vulnerable to certain street drugs (most notably cannabis and methamphetamine) and should be educated about the possibility that such substances may trigger episodes of psychiatric illness. Thinking about environment more broadly, it may be useful to suggest cognitive–behavioral therapy and the development of strategies for effectively coping with and managing stress. For affected parents with young children, parenting programs, social support, and day care may be helpful. Providers should discuss the importance of early detection of psychiatric illness because early and appropriate interventions improve prognosis.

Individuals with psychiatric conditions are often vigilant about monitoring for symptoms in children, though they may have difficulty determining normal behavior from abnormal. Psychiatrists should assist such families in achieving a balance between appropriate vigilance and inappropriate boundary setting or anxiety. Although it may be worrisome to have risk in the family, a family with affected individual(s) is well placed to recognize early manifestation of symptoms. Encouraging and educating clients about their abilities in this capacity can increase their sense of control.

Table 31.2–5.

Empirical Risks for Selected Mental Disorders

	Schizophrenia	Bipolar Disorder	Unipolar Major Depressive Disorder		Schizoaffective Disorder	C C D
			Men	Women		
Affected relative						
General population	1%[a]	0.8–1.6%[b,c,d,e]	1–15%[b,c]	2–23%[b,c]	0.5–<1%[b,c,d]	1
First degree (pooled)	9%[a]	5–20%[b,c,e]	9%[a]	18%[a]	1–10%[b,c,d]	1
Siblings	9–16%[d,j,k,l]	5–20%[b,c,e,m,j]	5–30%[b,c,m,j]	5–30%[b,c,m,j]	—	2
Parent	5–13% [o,j,k,l,p]	15%[c,d,e,m,q]	7–19%[m,j]	7–19%[m,j]	—	2
Both parents	45%[m,j]	50–75% for affective disorder[b,d,m,j]	—	—	—	-
Second degree (pooled)	2–6%[m,j]	5%[m]	—	—	—	-
Uncle/aunt	1–4%[j,k]	—	—	—	—	-
Nephew/niece	2–4%[j,k,p]	—	—	—	—	-
Grandparent	2–8%[j,k]	—	—	—	—	-
Half-sibling	4%[j,k]	—	—	—	—	-
Third degree	—	—	—	—	—	-
First cousin	2–6%[j,k]	—	—	—	—	-
Risks for additional mental disorders	Spectrum disorders (schizoaffective disorder, schizotypal personality disorder, paranoid personality disorder); unipolar and major depression	Unipolar disorder, substance dependence, schizoaffective disorder, cyclothymia, anxiety disorder, FDR; RR any affective disorder: ~20–30%	Anxiety disorders, alcohol dependence, dysthymia, ADHD		Schizophrenia, other psychotic disorders, bipolar disorder, unipolar depression, FDR; RR mood or psychotic disorder ~30%	T

ADHD, attention-deficit/hyperactivity disorder; FDR, first-degree relative; RR, Recurrence risk.

[a]Moldin SO. Psychiatric genetic counseling. In: Guze SB, ed. *Washington University Adult Psychiatry*. Mosby-Year Book; 1997.

[b]Duffy A, Grof P, Robertson C, Alda M. The implications of genetic studies of major mood disorders for clinical practice. *J Clin Psychiatry.* 2000;61:630–637.

[c]Gershon ES, Hamovit J, Guroff JJ, et al. A family study of schizoaffective, bipolar I, bipolar II, unipolar and normal control probands. *Arch Gen Psychiatry.* 1982;39:1157–1167.

[d]Gershon ES, DeLisi LE, Hamovit J, et al. A controlled family study of chronic psychosis. Schizophrenia and schizoaffective disorder. *Arch Gen Psychiatry.* 1988;45:328–336.

[e]Potash JB, DePaulo JR Jr. Searching high and low: a review of the genetics of bipolar disorder. *Bipolar*

Disord. 2000;2:8–26.

[f]Swedo SE, Rapoport IL, Leonard H, Lenane M, Cheslow D. Obsessive-compulsive disorder children and adolescence. Clinical phenomenology of 70 consecutive cases. *Arch Gen Psychiatry.* 1989;46:335–341.

[g]Crowe RR, Noyes R, Pauls DL, Slymen D. A family study of panic disorder. *Arch Gen Psychiatry.* 1983;40:1065–1069.

[h]Barkley RA. Attention deficit hyperactivity disorder. *Sci Am.* 1998;9:66.

[i]National Society of Genetic Counselors Psychiatric Special Interest Group: At a glance empiric risk data. http://www.nsgc.org/members_only/sig/sig_psyc_empiric.cfm.

[j]Nurnberger J Jr, Berrettini W. *Psychiatric Genetics*. 1st ed. London: Chapman & Hall; 1998:164.

[k]Hodgkinson KA, Murphy J, O'Neill S, Brzustowicz L, Bassett AS. Genetic counseling for schizophrenia in the era of molecular genetics. *Can J Psychiatry.* 2001;46:123–130.

[l]Kendler KS, McGuire M, Gruenberg AM, Walsh D. An epidemiologic, clinical and family study of simple schizophrenia in County Roscommon, Ireland. *Am J Psychiatry.* 1994;151:27–34.

[m]Harper PS. *Practical Genetic Counseling*. 4th ed. Oxford: Butterworth-Heinemann; 1994:348.

[n]Rasmussen SA, Tsuang MT. The epidemiology of obsessive compulsive disorder. *J Clin Psychiatry.* 1984;45:450–457.

[o]Biederman J, Faraone SV, Keenan K, et al. Further evidence for family-genetic risk factors in attention deficit hyperactivity disorder. Patterns of comorbidity in probands and relatives psychiatrically and pediatrically referred samples. *Arch Gen Psychiatry.* 1992;49:728–738.

[p]Gottesman II, Shields J. Schizophrenia: *The Epigenetic Puzzle* New York, NY: Cambridge University Press; 1982.

[q]Goodwin FK, Jamison KR. *Manic Depressive Illness* New York, NY: Oxford University Press; 1990:938.

[r]Asarnow RF, Nuechterlein KH, Fogelson D, et al. Schizophrenia and schizophrenia spectrum personality disorders in the first-degree relatives of children with schizophrenia: the UCLA family study. *Arch Gen Psychiatry.* 2001;58:581–588.

[s]Kendler KS, Gardner CO. The risk for psychiatric disorders in relatives of schizophrenic and control probands: a comparison of three independent studies. *Psychol Med.* 1997;27:411–419.

[t]Kendler KS, Walsh D. Schizophreniform disorder, delusional disorder and psychotic disorder not otherwise specified: clinical features, outcome and familial psychopathology. *Acta Psychiatr Scand.* 1995;91:370–378.

[u]McGuffin P, Rijsdijk F, Andrew M, Sham P, Katz R, Cardno A. The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch Gen Psychiatry.* 2003;60:497–502.

[v]National Institutes of Mental Health. Genetics and mental disorders. 1999. Available at http://www.nimh.nih.gov/research/genetics.htm.

CHALLENGES POSED BY PRESYMPTOMATIC AND SUSCEPTIBILITY GENOMIC TESTING

As genomic technologies continue to improve and knowledge about variants that can increase vulnerability for psychiatric illness grows, identification of individuals who are at higher risk for developing psychiatric illness based on (for example) the number of illness-associated single nucleotide polymorphisms they carry will become more common. Appropriate applications of such developments clearly need careful

consideration from both economic and ethical perspectives but may have important clinical uses. For example, at present, it can be very difficult to distinguish between emerging (prodromal) psychosis and normal adolescent behavior in youth who are at higher risk for psychosis (e.g., as a result of close family members who have experienced psychosis). However, some studies of prophylactic strategies to reduce risk among high-risk youth show promise, and for those in whom psychosis does emerge, early intervention is critical to allow for the best possible outcomes. Therefore, genomic technologies could be useful in determining which individuals might benefit most from prophylactic therapeutic approaches.

The clinical implementation of genomic testing for psychiatric illness is, and will continue to be, the subject of intense debate. Decisions about the initiation of clinical testing will likely hinge on three issues: (1) the demonstration of adequate proof that the identified variants unequivocally increase risk for psychiatric illness; (2) affirmation that the information gained from the testing has clinical utility (e.g., risk management strategies); and/or (3) demonstration of personal utility, such as testing that aids in coping and adaptation. Conflicting opinions notwithstanding, genomic testing for susceptibility to psychiatric illness is likely to become more commonplace. Psychiatrists will be on the front line for receiving requests for genetic counseling and testing because of their established relationship with clients and families with psychiatric disorders.

The clinical identification of genomic risks will most likely occur before most individuals have access to preventative options. The option of knowing risks without having access to preventative options raises concerns regarding the effect of such knowledge on the individual's mood, anxiety, distress, self-image, reproductive decisions, and decisions regarding career, family relationships, insurability, employment, and, potentially, other areas.

Across many disease indications, the majority of persons undergoing genomic testing and receiving their identified risk status do not experience long-term adverse effects. However, it should be anticipated that those receiving results indicating increased disease risk may experience elevated levels of disease-related anxiety and depression soon after receiving results and a small number have prolonged difficulty in adjusting to their results. Supportive counseling should be available to facilitate adjustment.

The pursuit of genomic counseling and testing by an individual has clear and profound implications for other family members. Consideration should be given to helping the client understand the potential reactions and effect that his or her decision to pursue genomic testing could have on close family members. The health professional can play a role in assisting the client in considering ways to share the information with selected significant others. Similarly, consideration should be given to the effect of such information on the spouse or significant other.

Individuals who have psychiatric susceptibility are arguably more likely to enter genetic counseling with higher baseline levels of mood disturbance and may be more likely to respond negatively to predictive testing information. The important role of stress in the onset and recurrence of psychiatric symptoms must also be considered; by providing an at-risk individual with stressful, anxiety-provoking risk information, might providers inadvertently hasten the onset of psychiatric symptomatology? More research in this area is vital to ensuring that clinicians do not take too paternalistic a stance, refusing wanted information to individuals who might benefit psychologically and medically (as well, refusing to provide wanted information may be a stressful, anxiety provoking experience for a client), while on the contrary, considering the inherent vulnerability in a population at risk for psychiatric illness. It is likely that the manner in which these delicate conversations unfold is important to client outcomes, and that a focus on counseling around these issues is key.

ETHICAL, LEGAL, AND SOCIAL CONSIDERATIONS

Although there are many benefits of genomics knowledge, scientists and the lay public have worried, from the beginning of genetic research, about the potential for harmful consequences of knowing one's genomic makeup. How will family members, neighbors, communities, employers, health care providers, and insurance companies react to such information? Might health care providers request that persons at increased genomic risk take psychiatric medications prophylactically? Might employers limit job opportunities for those with increased risk for psychiatric illness, keeping them in positions with lower levels of stress? Will persons at increased risk be pressured to not have children? Might providing such risk information induce significant enough anxiety/stress to trigger the onset of illness? And finally, how will families and the public be educated about the availability of susceptibility testing for psychiatric illness? Public messages about susceptibility testing should be crafted in a way that does not increase the already heavy burden of stigma attached to individuals with psychiatric disease and does not add to "courtesy stigma" attached to at-risk relatives. Increased public awareness of the hereditary component to psychiatric illness may heighten stigmatization. Conversely, having an identified, biological basis may supplant current public perceptions that mental illness is somehow a personal or family failure in moral, spiritual, or attitudinal perspectives.

The ability to provide susceptibility testing is likely to make psychiatric practice more complicated and will challenge the process of education and counseling (informed consent). Psychiatrists are accustomed to communicating and receiving information that requires special consideration owing to its sensitive nature; handling genomic information from clients will be yet another sensitive topic to be considered from social

and ethical perspectives. The diagnosis of a genomically influenced disorder may create risks not only for the patient but also for others within the family. Some clients may not choose to share genomic information with relatives, even though the information could be relevant to reproductive decisions, relationships, or even medical care. Coercion of affected relatives to undergo testing may become an issue, especially in the case in which at-risk relatives are pursuing genetic counseling to help guide their decisions. Parents may desire susceptibility testing of children even in the absence of preventive interventions and may risk stigmatizing the children and forcing information on them without informed consent. These situations call for careful attention to the confidential nature of the information, as well as the need to encourage family communication to discuss the potential for the associated personal, family, and social risks.

Historically, some attribution of psychiatric illness has been focused on child rearing and the home environment. The current focus on the biological nature of psychiatric illness has been important for reducing guilt and feelings of blame in many families. It is unclear how a more nuanced understanding of environmental risk factors, potentially including home environment and parenting style, might affect perceptions of guilt. It is important that messages about risk reduction be crafted in such a way as to not lend credence to outdated notions of parental causation.

It can be anticipated that many individuals with clear-cut psychiatric diagnoses will not be found to have identified susceptibility variants. Test results with no identified variants are likely to cause some individuals to question the validity of their psychiatric diagnoses and may have a negative effect on treatment compliance. In addition, the absence of an identified genomic susceptibility might increase parental guilt and "force" responsibility for their child's illness on their parenting efforts. A key aspect of pretest genetic counseling will be discussions about the meaning of a "negative" result.

Conversely, based on the very nature of susceptibility testing, many individuals identified as having susceptibility variants will never develop a psychiatric disorder. Counseling around the notion of susceptibility is complex, as is the necessary support for individuals who are at risk. Although the potential harm for such individuals is apparent, many may perceive benefits (perhaps related to perceived control and coping with vulnerability) that outweigh risks. If benign environmental modifications that reduce risk are identified, the concern around individuals with "positive" test results will decrease.

Finally, it is not yet clear how the public will react to genomic tests that indicate levels of risk that are only moderately increased over the population risk. It may be that individuals will interpret the resulting risk as higher than appropriate, or that the predictive limitations will lead them to discount the results. Such appraisals are likely to vary greatly among

individuals on the basis of their history, personality types, and coping styles.

As individuals within society become more aware of the identification of genomic factors within the complex etiology of psychiatric disorders, it is anticipated that requests for psychiatric genetic counseling will increase. Although genomic testing may provide important psychological benefits for many individuals, the uncertainties inherent in the current level of knowledge pose challenges in the education and counseling of clients. Psychiatrists are appropriately positioned to provide genetic counseling (or referrals for genetic counseling) and ongoing therapy to families experiencing psychiatric disorders. The greatest opportunity for benefit to clients will be provided through a multidisciplinary team comprising mental health and genetics professionals who adhere to the basic principles of genetic counseling and maintain an awareness of the latest research findings relevant to psychiatric genetics. Genetic counselors within the geographical area of the psychiatrist can be identified through the National Society of Genetic Counselors (http://www.nsgc.org).

ACKNOWLEDGMENT

Donald W. Hadley, M.S. contributed substantially to prior versions of this chapter.

REFERENCES

*Austin JC, Honer W. The genomic era and serious mental illness: a potential application for psychiatric genetic counseling. *Psychiatr Serv*. 2007;58:254–261.

*Austin JC, Palmer C, Rosen-Sheidley B, McCarthy-Veach P, Gettig B, Peay HL. Psychiatric disorders in clinical genetics II: individualizing recurrence risks. *J Genet Couns*. 2007;17:18–29.

*Austin JC, Peay HL. Applications and limitations of empiric data in provision of recurrence risks for schizophrenia: a practical review for healthcare professionals providing clinical psychiatric genetics consultations. *Clin Genet*. 2006;70:177–187.

Baker K, Costain G, Fung WL, Bassett AS. Chromosomal microarray analysis—a routine clinical genetic test for patients with schizophrenia. *Lancet Psychiatry*. 2014;1:329–331.

Baker DL, Schuette JL, Uhlmann WR, eds. *A Guide to Genetic Counseling*. New York, NY: Wiley-Liss; 1998.

Biesecker BB, Peay HL. Ethical issues in psychiatric genetics research: points to consider. *Psychopharmacology (Berl)*. 2003;171:27–35.

Clatworthy J, Bowskill R, Rank T, Parham R, Horne R. Adherence to medication in bipolar disorder: a qualitative study exploring the role of patients' beliefs about the condition and its treatment. *Bipolar Disord*. 2007;9(6):656–664.

Costain G, Esplen MJ, Toner B, et al. Evaluating genetic counseling for individuals with schizophrenia in the molecular age. *Schizophr Bull*. 2014;40(1):78–87.

Duffy A, Grof P, Robertson C, Alda M. The implications of genetics studies of major mood

disorders for clinical practice. *J Clin Psychiatry*. 2000;61:630–637.

Faraone SV, Tsuang MT, Tsuang DW. *Genetics of Mental Disorders*. New York, NY: Guilford Press; 1999.

*Finn CT, Smoller JW. Genetic counseling in psychiatry. *Harv Rev Psychiatry*. 2006;14:109–121.

Gamm JL, Nussbaum RL, Biesecker BB. Genetics and alcoholism among at-risk relatives I: perceptions of cause, risk, and control. *Am J Med Genet A*. 2004;128:144–150.

Gershon ES, Alliey-Rodriguez N. New ethical issues for genetic counseling in common mental disorders. *Am J Psychiatry*. 2013;170(9):968–976.

Hippman C, Lohn Z, Ringrose A, Inglis A, Cheek J, Austin JC. "Nothing is absolute in life": understanding uncertainty in the context of psychiatric genetic counseling from the perspective of those with serious mental illness. *J Genet Couns*. 2013;22(5):625–632.

Hodgkinson KA, Murphy J, O'Neill S, Brzustowicz L, Bassett AS. Genetic counseling for schizophrenia in the era of molecular genetics. *Can J Psychiatry*. 2001;46:123–130.

Inglis A, Koehn D, McGillivray B, Stewart SE, Austin J. Evaluating a unique, specialist psychiatric genetic counseling clinic: uptake and impact. *Clin Genet*. 2015;87(3):218–224.

Iyegbe C, Campbell D, Butler A, Ajnakina O, Sham P. The emerging molecular architecture of schizophrenia, polygenic risk scores and the clinical implications for GxE research. *Soc Psychiatry Psychiatr Epidemiol*. 2014;49(2):169–182.

Jolin EM, Weller EB, Weller RA. Velocardiofacial syndrome: is there a neuropsychiatric phenotype? *Curr Psychiatry Rep*. 2006;8:96–101.

Jones I, Scourfield J, McCandless F, Craddock N. Attitudes towards future testing for bipolar disorder susceptibility genes: a preliminary investigation. *J Affect Disord*. 2002;71:189–193.

McMahon W, Baty B, Botkin J. Genetic counseling and ethical issues for autism. *Am J Med Genet C Semin Med Genet*. 2006;142:52–57.

Meiser B, Mitchell P, Kasparian N, et al. Attitudes towards childbearing, causal attributions for bipolar disorder and psychological distress: a study of families with multiple cases of bipolar disorder. *Psychol Med*. 2007;37:1601–1611.

Meiser B, Mitchell P, McGirr H, Van Herten M, Schofield P. Implications of genetic risk information in families with a high density of bipolar disorder: an exploratory study. *Soc Sci Med*. 2005;60:109–118.

Miller SM, McDaniel SH, Rolland JS, Feetham SL. *Individuals, Families, and the New Era of Genetics: Biopsychosocial Perspectives*. New York, NY: WW Norton; 2006.

*Peay HL, Austin J. *How to Talk to Families about Psychiatric Illness*. New York, NY: WW Norton; 2011.

Peay HL, Rosenstein DL, Biesecker BB. Adaptation to bipolar disorder and perceived risk to children: a survey of parents with bipolar disorder. *BMC Psychiatry*. 2013;13:327.

Peay HL, Rosenstein DL, Biesecker BB. Parenting with bipolar disorder: coping with risk of mood disorders to children. *Soc Sci Med*. 2014;104:194–200.

Peay HL, Veach PM, Palmer CG, Rosen-Sheidley B, Gettig E, Austin JC. Psychiatric disorders in clinical genetics I: addressing family histories of psychiatric illness. *J Genet Couns*. 2007;17(1):6–17.

Quaid KA, Aschen SR, Smiley CL, Nurnberger JI. Perceived genetic risks for bipolar disorder in a patient population: an exploratory study. *J Genet Couns*. 2001;10:11.

Roberts JS, Cupples LA, Relkin NR, et al. Genetic risk assessment for adult children of people with Alzheimer's disease: the Risk Evaluation and Education for Alzheimer's Disease (REVEAL) study. *J Geriatr Psychiatry Neurol*. 2005;18:250–255.

Rutter M, Moffitt TE, Caspi A. Gene–environment interplay and psychopathology: multiple varieties but real effects. *J Child Psychol Psychiatry*. 2006;47:226–261.

Rutter M. The interplay of nature, nurture, and developmental influences. The challenge ahead for mental health. *Arch Gen Psychiatry*. 2002;59:996–1000.

Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511(7510):421–427.

Smith L, Sapers B, Reus VI, Freimer N. Attitudes towards bipolar disorder and predictive genetic testing among patients and providers. *J Med Genet*. 1996;33:544–549.

Smoller JW, Sheidley BR, Tsuang MT, eds. *Psychiatric Genetics: Applications in Clinical Practice*. Virginia: American Psychiatric Publishing; 2008.

Trippitelli C, Jamison K, Folstein M, Bartko J, DePaulo R. Pilot study on patients' and spouses' attitudes toward potential genetic testing for bipolar disorder. *Am J Psychiatry*. 1998;155:899–904.

▲ 31.3 Physical and Sexual Abuse of Adults

SHANNON F. STROMBERG, M.D.

One of the biggest threats to public health in the United States is interpersonal violence. It comes in many forms to include assault, rape, domestic violence, torture, and homicide. Although there are a plethora of types of violence, there is an unfortunate common denominator for them all. Namely, that every act of violence must have a victim. These victims span all races and genders. As an example, the leading cause of death in male African Americans continues to be homicide. Women continue to be at highest risk for lifetime sexual assault. Those in law enforcement may face threats daily just by going to work. In fact, statistics show that the majority of Americans will fall victim to a violent crime during their lifetime.

The most egregious result of violence is undoubtedly death, but the greater costs of violence may lie well beyond mortality. Morbidity for an individual exposed to violence is seen in decreased productivity, medical disability, and psychiatric sequelae. Morbidity from a societal standpoint occurs when people let violence impact where they live, where they raise and school their children, how late they stay out, or even where they walk or drive.

This threat has not been ignored though as evidenced by increased public health awareness. The last 20 years have seen an increase in emphasis on education and prevention. Statistics show an overall decrease in the total number of cases and the rates of violence in general during that

time. Identification of those who are at risk and of those who have been victimized continues to improve. Finally, the treatment modalities available to victims have expanded greatly and have led to decreased morbidity in survivors. That being said, there exists a need to continue to treat and attempt to prevent violence in all forms.

TYPES OF ASSAULT, PREVALENCE, AND RISK FACTORS

Assault can be viewed in the context of two variables. The first involves who is being assaulted, and the second is where the assault occurs. With these parameters, assault can be classified into several categories with some overlap among them. Violent crime, rape, domestic violence, workplace violence, and torture are the most commonly used categories to classify assault. The prevalence of these different types of assaults is most often reported via two different data collection systems. The first is the Federal Bureau of Investigation's Uniform Crime Report (UCR), which collects reported crime information from local law enforcement agencies. The second is the Bureau of Justice Statistics' National Crime Victimization Survey (NCVS). This survey generates estimates of the likelihood of victimization from different types of assaults. For each type of assault, there have also been attempts to develop risk factor assessments. Gender, age, race, history of prior assault, past psychiatric history, substance abuse, and income have been found to have some relevance to assault risk for different categories of assault. The last component to each type of assault is the makeup of the perpetrator. This varies widely both within and between the different categories of assault.

VIOLENT CRIME (AGGRAVATED AND SIMPLE ASSAULTS, ROBBERY)

Violent crime is defined by the UCR as murder and nonnegligent manslaughter, forcible rape, robbery, and aggravated assault. These categories exclude simple assault, which in the UCR is defined as an assault or attempted assault in which no weapon was used or no serious injury occurred to the victim. Included within this definition are stalking, intimidating, coercing, and hazing.

Prevalence

In 2013, an estimated 6,126,420 violent crimes occurred in the United States excluding homicide. Aggravated assaults accounted for an estimated 994,220 of this total, simple assaults were estimated at 4,186,390, and robberies were estimated to be 645,650. Ten-year trends show a decrease of 29.2 and 5.6 percent since 2004 for the estimated number of aggravated and simple assaults, respectively. Over the same 10 years, there was an increase of 4.7 percent for the estimated number of robberies. The rates of all types of violent crimes in 2013 were lower than the rates observed in

2004.

Risk Factors

Gender carries a small risk with males more likely to experience violent crime than females. Age plays a larger role with those younger than 25 years being at highest risk. Racial and ethnic groups also appear to have different rates of violent crime victimization. The groups most at risk for experiencing violent crime are persons of two or more races, American Indians, and Alaska Natives. African Americans and Hispanics have higher rates than whites, while Asians have the lowest rate of any group. Another risk factor is household income with lower income households being more likely to be burglarized. Homelessness has also been shown to be a factor in increased physical assaults. Finally, substance abuse has been shown in multiple studies to increase the risk of victimization.

Perpetrators

Both the UCR and NCVS suffer from limitations with regard to reporting on offender characteristics. One problem is the heavy reliance on arrest data, which exclude any data from assaults and robbery not resulting in arrest. Another problem is that data are largely generated on the basis of the victim's report of the incident. Although these incident data can be used to define the victim's characteristics, they cannot be used accurately to define the offender. By way of example, the NCVS reports on victimization as reported by the victims, but a tool as well developed as the NCVS does not yet exist to collect information on offending as reported by the offender. Furthermore, incident data for victims cannot control for a situation in which one offender has assaulted multiple victims.

RAPE

The legal definition of rape varies from state to state. Some states strictly define rape, whereas other states describe any sex crime as varying degrees of sexual misconduct or sexual assault. The UCR defines rape as follows: "Penetration, no matter how slight, of the vagina or anus with any body part or object, or oral penetration by a sex organ of another person, without the consent of the victim." In this section, rape refers to forced oral, vaginal, or anal penetration, and sexual assault refers to any forced sexual contact.

Like other violent crimes, sexual assault is on the decrease; however, every 107 seconds someone in the United States is sexually assaulted. The stereotype of the perpetrator being a stranger has been proven inaccurate by research, with nearly two-thirds of all sexual assaults being committed by someone known to the victim.

Prevalence

Unfortunately, accurate statistics are difficult to obtain because of underreporting and unacknowledgment. The Rape, Abuse and Incest National Network (RAINN) and NCVS estimate that nearly 70 percent of sexual assaults go unreported. Although in the United States rapes have fallen in number since 1993, in 2013, there were more than 300,000 victims of rape, attempted rape, or sexual assault. RAINN estimates that one in six US women and 1 in 33 US men are victims of sexual assault.

In addition to being underreported, there are many rapes that go unacknowledged by the victim. They often refer to the assault in more benign terms such as a misunderstanding, despite the fact that it meets the legal definition of rape. Research has reported that this percentage is substantial; more than 50 percent of sexual assaults may go unacknowledged and, thus, unreported. Research has further revealed that victims who fail to acknowledge having been raped usually believe that rape involves two strangers and greater force, as opposed to the views of those who acknowledge their having been sexually assaulted.

Risk Factors

Although females are usually the victim of rape and sexual assault, greater than 10 percent of victims are estimated to be male. Furthermore, most experts believe that males underreport more than females. Despite growing awareness, research on the male sexual assault victim remains sparse. Nevertheless, young females have four times the risk of any other group of becoming a victim of sexual assault. Eighty percent of all rape victims are younger than 30 years, and Bureau of Justice statistics indicates that women 16 to 24 years of age are at the greatest risk for rape in the United States. Having been a victim of childhood abuse or previous assault increases the likelihood of further assault of all types. Many studies have also showed substance use and abuse to be a risk factor for sexual assault.

Perpetrators

In general, rape is considered a crime of power and aggression, not one of sexuality. However, some controversy remains about whether the sexual predator's motivation is anger, power, sexual gratification, or control. What has been shown in multiple research studies is that the majority of perpetrators are known by their victims. Approximately 73 percent of all sexual assaults and 66 percent of rapes are being committed by nonstrangers.

The Centers for Disease Control and Prevention (CDC) conducted a review looking at the risk factors for perpetration and found a multitude of individual and relationship components. Individual risk factors included prior sexual victimization or perpetration, substance use, difficulty with

empathy, aggressiveness and acceptance of violence, different sexual issues (early initiation, coercive fantasies, and impersonal sex), and hostility toward females. Relationship risk factors for perpetration included family environments characterized by physical violence or that were emotionally unsupportive, association with sexually aggressive peers, and poor parent–child relations.

DOMESTIC VIOLENCE AND INTIMATE PARTNER VIOLENCE

Domestic violence is defined as a pattern of abusive behavior in any relationship with the intent to gain power or control. It can be in the form of physical, sexual, emotional, economic, or psychological abuse. This abuse can be directed toward children, elders, or partners. Domestic violence directed at a partner by a current or former partner or spouse is referred to as intimate partner violence (IPV). Domestic violence is often broken down into two categories: "high-severity abuse," which includes being threatened or hurt with a weapon, burned, choked, hit or kicked, and resulting in broken bones and head or internal injuries; and "low-severity abuse," which includes being slapped, hit, or kicked without injury, but could also include bruising, minor cuts, and sprains. The CDC has defined four subtypes of IPV: physical violence, sexual violence, stalking, and psychological aggression.

Elder abuse is a variant of assault either when the abuser is an unrelated caretaker or as a result of domestic violence. Elder abuse can include physical, sexual, emotional, and/or financial abuse.

Prevalence

Estimates of the prevalence of domestic violence in the United States vary widely, and many studies include psychological or emotional abuse. One of the more consistent statistics is that nearly 30 percent of women in the United States have experienced some form of physical or sexual abuse from an intimate partner. The NCVS reported more than 1,116,090 incidents of domestic violence in 2013 and 748,800 cases of IPV. The prevalence rates for both are unchanged over the previous 10 years. The National Intimate Partner and Sexual Violence Survey (NISVS) from 2011 found that 27.3 percent of women and 11.5 percent of men experienced sexual violence, physical violence, or stalking by an intimate partner during their lifetimes.

Risk Factors

Studies have found that certain individual, relationship, community, and societal risk factors are associated with domestic violence and IPV. Individual risk factors may include younger age, low income, substance use, exposure to violence and abuse as a child, low level of education, and having traits of Borderline or Antisocial personality disorder. Relationship

risk factors include marital discord and instability, economic stressors, the man having multiple partners, and male dominance in the family. Finally, multiple studies have shown that community and societal risk factors such as poverty, overcrowding, acceptance of IPV or lack of sanctions against it, and traditional gender roles or norms play roles in increased rates of domestic violence and IPV.

Perpetrators

The perpetrators of domestic violence come from all races, cultures, and ethnicities. They also span all genders, religions, educational backgrounds, and socioeconomic levels. Many studies have shown that having been a victim of abuse or witnessing abuse in the home increases the risk of someone becoming an abuser. Substance abuse is often a risk factor for episodes of perpetration. Various authors have described typologies to describe perpetrators of domestic violence. Generally, these can be loosely grouped into two types of offenders. The first is described as being "emotionally dependent," with a high level of jealousy, a high level of interpersonal dependency, a high level of anger, and low self-esteem. These abusers are prone to interpret their partner's behavior as betrayal, and their anger can explode into violence. The second type is the "antisocial/narcissistic" person, who tends to have hostile attitudes toward women and low empathy. These abusers become less physiologically aroused as violence escalates. They also have the highest rates of alcohol dependence.

Several traits can be associated with domestic violence and IPV, including exhibiting verbal abuse, projecting blame, objectifying the victims, and isolating them from social support. Finally, abusers are noted to be charming and have manipulative personalities in public while being cruel and hostile to their intimates.

WORKPLACE VIOLENCE

Acts of violence at work are defined as simple assault, aggravated assault, robbery, rape/sexual assault, and homicide.

Prevalence

Between the years of 2002 and 2009 in the United States, the rate of nonfatal workplace violence declined by 33 percent according to the NCVS. In 2009, nonfatal workplace violence accounted for about 15 percent of all nonfatal violent crimes. During this same period of time, about 3,800 workplace homicides occurred of which 80 percent involved firearms. With some variation, total violent crime in the workplace declined 35 percent between the years 2002 and 2009. This correlates with a similar 37 percent decrease in all violent crime (workplace and nonworkplace) during the

same period.

Risk Factors

Unlike other violent crimes, workplace crime rates are higher among whites than other races. An important component of workplace violence is the relationship between the risk of violence and the type of occupation. The NCVS found that people working in law enforcement are at the highest risk for victimization (47.7 per 1,000) and constituted 19 percent of all workplace assaults. This included law enforcement officers, security guards, and corrections officers. Other high-risk occupations include bartenders (79.9 per 1,000), the mental health fields (20.5 per 1,000), transportation fields (12.2 per 1,000), and high school teachers (13.5 per 1,000). Preschool and college/university teachers are at the lowest risk for victimization (0.9 and 1.9 per 1,000, respectively).

Perpetrators

There are several distinct characteristics among offenders, of which gender is the most pronounced. As reported by the victims, four-fifths of all violent workplace crime are committed by men, and this is regardless of the victim's gender. The race of the offender is most often white (55 percent), followed by African American (30 percent). Substance use in the offender is reported more often by victims in the law enforcement and retail sales occupations as opposed to the mental health and teaching fields. Unlike domestic violence, workplace violence is more often perpetrated by strangers (53 percent) than by well-known acquaintances (12 percent) or known intimates (1 percent). Mental health and teaching are the only fields in which attacks occur more often at the hands of a known perpetrator.

SEQUELAE OF VIOLENCE AND ASSAULT

Survivors of violence have varied reactions, but they are similar to those exposed to other types of trauma. Furthermore, the severity of the sequelae varies with the individual. However, multiple studies have shown that many people who experience violence have decreased physical and mental health, resulting in higher utilization of health care services.

The most common reported sequelae after sexual assault in women are posttraumatic stress disorder (PTSD), mood disorders, substance abuse, eating disorders, and sexual disorders. Although PTSD rates vary in the literature depending on methodology and sample, one study in women instituting legal proceedings found a PTSD rate of 95 percent versus a control group of women who had experienced a life-threatening trauma unrelated to any form of sexual abuse, who had a rate of PTSD of 47 percent. Perceived life threat and physical injury are also linked to a diagnosis of PTSD. One of the most protective factors in ameliorating the

development of PTSD is social support. Furthermore, the lack of social support and/or the perception of being treated differently can be quite detrimental to the survivor, causing an increase in PTSD symptoms.

Domestic violence has been associated with depression, anxiety, low self-esteem, substance abuse, sexual dysfunction, functional gastrointestinal disorders, headaches, chronic pain, and multiple somatic symptoms. Research has also shown a dose effect of escalating violence and increasing physical complaints. Physical signs and symptoms associated with current abuse include loss of appetite; frequent or serious bruises; nightmares; vaginal discharge; eating binges or self-induced vomiting; diarrhea; broken bones, sprains, or serious cuts; pain in the pelvic or genital area; fainting; abdominal pain; breast pain; frequent or severe headaches; difficulty in passing urine; chest pain; problems with sleeping; shortness of breath; and constipation. Many studies have shown that survivors of domestic violence attempt suicide more frequently than those who have not experienced violence. Substance abuse appears to be a significant factor for both the survivor and the batterer. It is both a risk factor for abuse and a consequence of it.

The relationship between childhood abuse and adult abuse is complex. Childhood abuse is a risk factor for abuse as an adult and it increases both physical and mental health sequelae of adult abuse. As expected, survivors of sexual assault or domestic violence who are older and without a history of childhood abuse have fewer symptoms.

TREATMENT ISSUES

In the evaluation of assault victims, multiple treatment issues need to be addressed. Assessing safety, treating acute psychiatric and medical issues, and evaluating the need for hospitalization are primary foci of the initial assault workup. Another important issue to be dealt with during the initial evaluation is the mandatory reporting requirements for abuse victims. After these treatment issues have been addressed, the clinician then needs to assess whether the patient may benefit from psychotherapeutic and/or pharmacological intervention.

Initial Evaluation

Many victims of assault initially present to a medical doctor for treatment of their injuries. In emergency departments, patients may often be treated for their injuries without the assault being recognized or addressed. Consequently, it is important to consider any concerning injuries as being potential stigmata of an assault. The patient may initially be avoidant when asked about the cause of injuries that appear related to assault. This may be further complicated by the need to complete specific tasks (rape kit, photo documentation, reporting of legal concerns) that require the patient to

describe and relive the recent assault. Consequently, it is paramount to build a rapport with the patient while continuing to complete a thorough evaluation. To facilitate this process, it is important to reduce the victim's stress and anxiety regarding discussing the event. The initial evaluation process should be explained to the patient before beginning. Allowing the patient some say in dictating the pace and content of the interview is preferable to having a patient feel a lack of control over the process. The crime may represent a time when they lacked control, and similar feelings in the initial interview can lead to a triggering of traumatic or anxious responses. Finally, the physician should remain attuned to nonverbal patient responses that signify discomfort and conduct the interview accordingly. This early therapeutic contact with assault victims may help influence their future prognosis positively. Furthermore, when a patient discloses domestic violence, this should be documented in the chart for future follow-up and possible future legal documentation.

Safety

After an assault, it is imperative that a safety assessment be performed on the victim. The patient needs to be evaluated for suicidal and homicidal ideation as it pertains to the recent assault. The patient's safety from further assault also needs to be addressed, especially in a setting in which the perpetrator was an intimate partner or known acquaintance. Finally, the patient needs to be screened for severe psychological symptoms that would cause difficulty in self-care. These would include acute mood deterioration or affective instability, self-destructive behaviors, dissociative symptoms, and/or frank psychosis. If the evaluation shows that any of these areas are inadequate to ensure patient safety, then a plan needs to be developed. This plan should provide for the patient's physical safety and give the patient a place to sleep, recuperate, and eat.

Hospitalization

In the event that the patient's safety cannot be guaranteed, hospitalization of the assault victim may be necessary. Common indications for hospitalization include (1) severe medical injuries, (2) suicidality and/or homicidality, (3) dissociative or psychotic symptoms, (4) mood instability or affective dysregulation, (5) self-destructive behaviors, and (6) a continued serious threat to the patient's life or well-being. If the victim is admitted to the hospital, then an individualized multidisciplinary treatment plan should be developed. Safety, milieu therapy, mood stabilization, and medication evaluation are the primary treatment modalities provided in the hospital and reflect the short-stay nature of most psychiatric facilities. Because the assault victim will likely need longer treatment than what can be provided in the hospital, it is important for

there to be a coordinated inpatient-to-outpatient transition. This, like the treatment plan, should be a multidisciplinary transition, with psychiatric, medical, and social work issues all being addressed prior to discharge.

Legal Issues

A physician evaluating an assault victim needs to adhere to mandatory reporting requirements for the state in which he or she is practicing. Mandated reporting of child abuse and abuse of developmentally disabled children exists in all 50 states. Mandatory reporting laws also exist in all 50 states for confirmed elder abuse. In cases of assault secondary to domestic violence, states do not usually require mandatory reporting. However, many states do require physicians to report serious injuries due to acts of criminal violence. Thus, in the setting of injury secondary to domestic violence, this can sometimes be considered a de facto mandatory reporting situation.

Continuing the Assessment

Initially, the victims of assault may report a variety of different psychiatric symptoms and psychosocial issues. They may endorse feelings of depression, anxiety, aggressiveness, guilt, shame, and low self-esteem. Interpersonal relationships may suffer, and substance abuse may occur. These symptoms and issues should be addressed early because of the risk to the patient of future psychological sequelae. It is well documented that victims of assault are at increased risk for PTSD, major depression, agoraphobia, OCD, social phobia, and simple phobia. The likelihood of developing these illnesses after an assault varies according to the type of assault and risk factors related to being assaulted. Therefore, a treatment plan must be developed that (1) assists the patient in dealing with psychiatric symptoms that he or she is currently experiencing because of the assault and (2) assists the patient with psychiatric symptoms that may arise later.

Postassault Psychotherapy Treatment

After the patient's safety has been ensured and the initial evaluation is complete, there are a variety of psychological interventions that can be initiated. Cognitive-behavioral approaches are the most researched techniques demonstrating efficacy. Exposure therapy is a variant of CBT that has been shown to help victims emotionally process the assault by decreasing their fear to memories or cues of the event. Recent literature has shown that brief trauma-focused CBT may accelerate recovery in victims manifesting acute PTSD and is superior to supportive therapy. These individual psychotherapies may be augmented with group psychotherapy, art therapy, dance and movement therapy, music therapy, and body-

oriented approaches if these might prove beneficial to the patient.

Postassault Psychopharmacological Treatment

Although medications are not recommended in the acute treatment of all assault victims, they may be beneficial in certain circumstances. The clinician may decide to medicate patients with incapacitating anxiety, extreme aggression toward themselves or others, and/or dissociation or psychoses immediately after an assault. The patient's safety and the safety of those around the patient will help decide the need for pharmacological intervention. Most medication treatment will be initiated much later after the assault as the patient develops symptoms of PTSD, depression, anxiety, OCD, and/or psychosis. Although medications may be useful in symptom management, they should not be viewed as a replacement for psychotherapy aimed at trauma symptom resolution.

References

*Acierno R, Resnick HS, Kilpatrick DG. Health impact of interpersonal violence. 1: prevalence rates, case identification and risk factors for sexual assault, physical assault and domestic violence in men and women. *Behav Med*. 1997;23(2):53–64.

*Breiding MJ, Smith SG, Basile KC, Walters ML, Chen J, Merrick MT. Prevalence and characteristics of sexual violence, stalking, and intimate partner violence victimization—National Intimate Partner and Sexual Violence Survey, United States, 2011. *MMWR Surveill Summ*. 2014;63(8):1–18.

Bureau of Justice Statistics. *National Crime Victimization Survey, Violence in the Workplace, 1993–2009*. Washington, DC: U.S. Department of Justice; 2011.

Bureau of Justice Statistics. *National Crime Victimization Survey*. Washington, DC: U.S. Department of Justice; 2014.

Carlson EB, Dutton MA. Assessing experiences and responses of crime victims. *J Trauma Stress*. 2003;16(2):133–148.

Faravelli C, Giugni H, Salvatori S, Ricca V. Psychopathology after rape. *Am J Psychiatry*. 2004;161(8):1483–1485.

Federal Bureau of Investigation. *Criminal Justice Information Services Division, Reporting Rape in 2013*. Washington, DC: U.S. Department of Justice; 2014.

Federal Bureau of Investigation. *Uniform Crime Reports, Crime in the United States, 2013*. Washington, DC: U.S. Department of Justice; 2014.

Kilpatrick DG, Acierno R. Mental health needs of crime victims: epidemiology and outcomes. *J Trauma Stress*. 2003;16(2):119–132.

Kluft RP, Bloom SL, Kinzie JD. Treating traumatized patients and victims of violence. *New Dir Ment Health Serv*. 2000;(86):79–102.

Leserman J. Sexual abuse history: prevalence, health effects, mediators, and psychological treatment. *Psychosom Med*. 2005;67(6):906–915.

McCauley J, Kern DE, Kolodner K, Derogatis LR, Bass E. Relation of low-severity violence to women's health. *J Gen Intern Med*. 1998;13:687–691.

McCauley J, Kern DE, Kolodner K, et al. The “battering syndrome”: prevalence and clinical characteristics of domestic violence in primary care internal medicine practices. *Ann Intern Med*. 1995;123(10):737–746.

Palmero GB. The mind of the sexual predator. *Curr Opin Psychiatry*. 2007;20:497–500.

*Resnick PJ, Scott CL. Legal issues in treating perpetrators and victims of violence. *Psychiatr Clin North Am*. 1997;20(2):473–487.

*Roberts NP, Kitchiner NJ, Kenardy J, Bisson J. Early psychological interventions to treat acute traumatic stress symptoms. *Cochrane Database Syst Rev*. 2010;(3):CD007944.

Sijbrandij M, Olff M, Reitsma JB, Carlier IV, de Vries MH, Gersons BP. Treatment of acute posttraumatic stress disorder with brief cognitive behavioral therapy: a randomized controlled trial. *Am J Psychiatry*. 2007;164(1):82–90.

*Tharp AT, DeGue S, Valle LA, Brookmeyer KA, Massetti GM, Matjasko JL. A systematic qualitative review of risk and protective factors for sexual violence perpetration. *Trauma Violence Abuse*. 2012;14(2):133–167.

Ullman SE, Filipas HH. Predictions of PTSD symptom severity and social reactions in sexual assault victims. *J Trauma Stress*. 2001;14(2):369–389.

▲ 31.4 Complementary, Alternative, and Integrative Approaches in Mental Health Care

JAMES H. LAKE, M.D.

DEFINITIONS AND HISTORICAL BACKGROUND

The National Center for Complementary and Integrative Health (NCCIH) of the National Institutes of Health defines complementary and alternative medical (CAM) practices as "those healthcare and medical practices that are not currently an integral part of conventional medicine." In the United States and other Western countries where biomedicine is the dominant healing tradition, *alternative* treatments are approaches for which a mechanism of action is not supported by *current* biomedical theory. Acupuncture is regarded as an alternative form of treatment in Western countries because Chinese medicine postulates the existence of an energetic principle called "*qi*" that is not congruent with the assumptions of current science. In contrast, complementary therapies are based on mechanisms that are empirically verifiable by current science and consistent with biomedical theory but remain controversial or are not endorsed by biomedicine for ideological or other reasons. St. John wort and *S*-adenosyl methionine (SAMe) are presently regarded as complementary treatments in the United States. Although the postulated mechanisms of action of both naturally occurring substances are consistent with biomedical theory and well-designed studies have demonstrated beneficial effects in depressed mood, however, there is ongoing debate in biomedical psychiatry over the appropriate uses of St. John wort and SAMe in the treatment of depressed mood. The majority of traditional healing practices have not been investigated in Western-style research studies and their mechanisms remain poorly understood. Integrative medicine combines biomedicine with alternative and complementary approaches aimed at maintaining wellness or treating illness.

In the North America, Europe, and other industrialized world regions, an increasing percentage of the population is using complementary, alternative, and integrative approaches to treat or *self-treat* medical and mental health problems. A large patient survey (N = 7,503) found that females, college graduates, and persons who believed that they received poor health care were more likely to use CAM, and only one-half of individuals who used CAM notified their health care providers. The majority of CAM users are well educated, committed to personal growth, satisfied with the conventional care provided by their physician or other health care provider, and use both prescription medications and CAM approaches for the same problem. An increasing number of medical

schools, nursing schools, and psychology graduate programs offer courses on CAM. Symposia on CAM modalities are included in the annual meetings of the American Medical Association (AMA), the American Psychiatric Association (APA), and other professional medical associations. Approximately half of US physicians—and the majority of European physicians—believe that acupuncture, chiropractic, and homeopathy are valid therapeutic modalities and refer patients to practitioners of these therapies. Increasing numbers of primary care physicians are becoming certified to practice massage, acupuncture, herbal medicine, homeopathy, and other nonallopathic modalities.

Notwithstanding growing interest in CAM therapies among conventionally trained health care providers, the efficacy and safety of most CAM approaches have not been established. Emerging evidence for select CAM treatments and safety or efficacy concerns about pharmaceuticals have led to the use of CAM by increasing numbers of patients. Limited research on the majority of CAM treatments is complicated by inconsistent research findings that make it difficult to distinguish placebo effects from outcomes related to a discrete mechanism. For example, many studies on St. John wort (*Hypericum perforatum*) in depressed mood have been criticized because of research design flaws resulting in controversy over the significance of findings. Beneficial effects of yoga, *qigong,* Healing Touch, and Reiki on general well-being or a particular psychiatric disorder are even more difficult to interpret because of the absence of a falsifiable hypothesis describing a discrete mechanism of action and problems designing controls. Problems defining a discrete mechanism of action for the beneficial effects of mind–body or "energy" therapies limit the ability of science to verify correlations between clinical outcomes and the *effect* of a particular CAM approach.

Persons diagnosed with a major psychiatric disorder are significantly more likely to use CAM treatments than the general population, and the majority of people who use CAM to self-treat a mental health problem take prescription medications concurrently. Almost one-half of persons diagnosed with major depressive disorder or panic disorder use at least one CAM treatment, compared to less than one-third of the general adult population. Roughly two-thirds of severely depressed or anxious persons who use CAM therapies consult with a mental health professional and 90 percent of the time this is a psychiatrist. The majority of persons who use CAM therapies for a mental health problem do not disclose this to their mental health care provider. This trend is alarming in view of potentially serious safety problems that can result when combining certain herbals or other natural products with pharmaceuticals. More than one-half of persons who self-treat severe depression or anxiety using CAM while concurrently taking a pharmaceutical believe that CAM treatments and conventional medications are equally efficacious.

THE EVOLVING CONCEPTUAL FOUNDATIONS OF BIOMEDICINE

Experiences of mental illness are shaped by cultural factors, including beliefs about the causes or *meanings* of symptoms and culturally determined understandings of the way treatments work to alleviate suffering. In contrast to the reductionist perspective of biomedicine, most nonbiomedical healing traditions reject the materialist explanations of current science as incomplete or arbitrary. Chinese medicine, Ayurveda, and Tibetan medicine, for example, are based on fundamentally different *nonmaterialist* assumptions about the nature of reality—including the microcosm of the human body—and the subtle influences of energetic and spiritual factors on illness. The metaphysical assumptions of nonbiomedical healing traditions have resulted in disparate explanatory models of illness and, by extension, disparate diagnostic and treatment approaches addressing the causes *or meanings* of illness within the patient's cultural and spiritual milieu.

LEGAL AND ETHICAL ISSUES

CAM use in mental health care raises important legal and ethical issues for psychiatrists and allied mental health professionals. The increasing use of CAM in mental health care is resulting in competing demands of patient expectations, clinical judgment, and liability considerations. Clinicians who recommend or prescribe CAM often do so in the absence of institutional oversight or professional guidance. This is complicated by the absence of treatment guidelines, policies, and procedures to help clinicians deal with liability issues. The ethical and legal relationship between a psychiatrist or other heath professional and his or her patient may consist of a limited one-time consultation or ongoing treatment using medications, psychotherapy, or CAM therapies. The choice of treatment defines the purpose of a contract between a patient and a health care provider.

Scope of practice refers to legally authorized boundaries of care for a particular conventional medical specialty or CAM practice. There is no clearly established legal–ethical framework that defines the scope of practice for physicians or CAM practitioners who treat patients using CAM modalities. Furthermore, substantial differences may exist between state jurisdictions for the kinds of medical practices that are legally sanctioned for a particular CAM specialty. For example, chiropractors are authorized to give nutritional advice in some states but not others, and massage therapists are prohibited from engaging in mental health counseling in most states but not all. All health care providers are ethically obligated to provide competent professional care. This duty includes the demonstration of professional competence when treating patients and the exercise of sound judgment when referring a patient for consultation with other conventionally trained medical specialists or CAM practitioners. The

patient assumes responsibility for complying with prescribed treatment and notifying his or her health care provider of adverse effects or other problems that emerge during treatment. Medical practitioners who use interventions that are not recognized as falling within the scope of their professional competence may invite disciplinary action by the state medical board or other regulatory agencies if patients or professional peers make formal complaints to those bodies. Disciplinary action may result in probation, suspension, or revocation of a professional license to practice.

It is advisable for health care providers who recommend or prescribe CAM approaches to first learn about all applicable restrictions imposed by their state medical boards on the scope of medical practices in their respective states. When a patient's medical or mental health problem is outside of the scope of the provider's expertise and experience, he or she is ethically obligated to refer the patient to an appropriate and competent provider. A provider who refers a patient to a CAM practitioner assumes liability for negative outcomes resulting from the referral, including harmful effects that may result from treatment. Referrals should be made only after confirming that a particular CAM practitioner is qualified to practice in his or her specialty. Malpractice refers to negligence, which is defined as failure to use due care in treating a patient, resulting in injury to the patient. Courts make rulings on malpractice using different standards of care for physicians and diverse alternative medical practitioners. The issue of malpractice insurance coverage pertaining to the practice of CAM is poorly defined. States vary significantly in their requirements for malpractice coverage, and individual insurers vary in their coverage of CAM therapies. Statutes pertaining to fraud in health care are intended to prevent intentional deception of patients by physicians and other providers. In cases in which a clinician or an institution submits a reimbursement claim for a treatment that the clinician knew or *should have known* was medically unnecessary, this might be grounds for a finding of fraud.

SAFETY

Evaluating the safety of CAM and integrative approaches poses complex challenges. Many herbal treatments are widely regarded as safe because of their long-standing historical use in traditional healing practices. There are limited safety data on the majority of herbals and other natural products used to treat medical or mental health conditions. The safety of imported herbal medicines and other natural products used for medicinal purposes from Asia is more difficult to assess than Western botanicals because of the absence of strong import controls. Herbals imported from China and India are sometimes adulterated with steroids, barbiturates, acetaminophen, and other drugs that may result in serious adverse effects and toxic interactions when combined with certain herbs or prescription medications.

Herbals and other natural products may cause adverse effects and potentially interact with other prescription medications with harmful consequences. In contrast to herbals and nonherbal natural products, few safety risks are associated with somatic therapies, mind–body practices, and so-called "energy" healing. Most safety information on CAM comes from case reports of limited reliability, and published reviews of adverse effects often cite conflicting findings. St. John wort has the greatest number of reported cases of adverse effects because it is one of the most widely used CAM treatments of depressed mood. In contrast, relatively little is known about adverse effects and toxicities associated with the majority of natural products. Safety information on natural products is especially incomplete with regard to young children, pregnant and nursing women, and patients with significant liver or kidney disease. The paucity of reliable safety data is complicated by the absence of established standards of professional training or credentialing for many CAM practitioners. Furthermore, persons who use prescription medications frequently self-treat using natural products without disclosing this information to their prescribing physician, resulting in unknown rates of toxic interactions and underreporting of potentially serious interactions between medications and natural products.

Table 31.4–1.
Safety Issues Associated with Botanicals and Other Natural Products

Natural Product	Adverse Effects	Potentially Toxic Interactions and Contraindications
Ginkgo biloba	May increase bleeding risk if used before surgery or labor/delivery Gastrointestinal discomfort, skin allergy, dizziness, headache	**Contraindicated** in combination with warfarin due to interference with platelet activation factor (PAF)
Kava (*Piper methysticum*)	Gastrointestinal discomfort (generally mild), allergic skin rashes, mild headache Case reports of severe hepatoxicity (always recommend routine liver function tests for all patients taking kava products)	Warn against combined use with benzodiazepines Case reports of lethargy and disorientation when used with alprazolam, other benzodiazepines
St. John wort (*Hypericum perforatum*)	Uncommon adverse effects include mild stomach discomfort, allergic skin rashes, tiredness, restlessness Rare case reports of mania	**Contraindicated** in combination with cyclosporine, other immunosuppressive agents, oral contraceptives, warfarin, protease inhibitors, and other antiretroviral drugs (may induce cytochrome P450 hepatic enzymes, causing increased metabolism and decreased serum levels of these

	Rare case reports of mild serotonin syndrome	drugs)
Valerian (*Valeriana officinalis*)	Uncommon reports of morning grogginess; headache Caution against driving for up to a few hours after using	Warn against combined use with benzodiazepines or barbiturates (may potentiate central nervous system sedating effects)
L-Tryptophan	Uncommon adverse effects include drowsiness, dry mouth, and blurred vision Fatal cases of eosinophilia-myalgia syndrome (EMS) attributed to a single contaminated batch of L-tryptophan in the late 1980s L-tryptophan remains in widespread use and no subsequent cases of EMS have been reported	Rare case reports of possible serotonin syndrome when combined with SSRIs and other antidepressants
5-Hydroxy-tryptophan (5-HTP)	Case reports of mild nausea, dry mouth, stomach irritation, and drowsiness at commonly used doses (50–100 mg thrice daily)	No case reports of serotonin syndrome when used in combination with SSRIs, other antidepressants
S-Adenosyl-methionine (SAMe)	Well tolerated Dose-related gastrointestinal distress is most common side effect (can range from mild nausea to vomiting and diarrhea) Case reports of mild nervousness and insomnia	Case reports of serotonin syndrome when used in combination with SSRIs, other antidepressants Warning: Risk of mania induction in bipolar patients; should be avoided in this population
Omega-3 fatty acids	Well tolerated at low doses, but there are reports of nausea and loose stools at higher doses; gastrointestinal side effects are reduced by taking with meals	Mild blood-thinning activity but no bleeding problems caused when used alone or with aspirin Warning: Increased risk of hypervitaminosis A and D with chronic high doses of cod liver oil because of high concentrations of vitamins A and D Caution: Pharmaceutical-grade products are recommended to minimize the risk of contamination with heavy metals, polychlorinated biphenyls, organochlorines

SSRI, selective serotonin reuptake inhibitor.

Patients should be advised about adverse effects or toxic interactions before starting *any* natural product alone or concurrently with *any* prescription medication. Table 31.4–1 shows adverse effects, potential interactions, and contraindications of natural products widely used in

North America, Western Europe, and Australia to treat or self-treat psychiatric disorders.

Rare cases of injury or adverse effects have been reported with CAM therapies that are not based on a described biological mechanism of action, including, for example, therapeutic massage, acupuncture, yoga, *qigong,* and Reiki. Where applicable, these are reviewed in the text.

Although most alternative practitioners are thoroughly trained in their respective CAM disciplines and dedicated to patient care, some lack the skills needed to manage serious psychiatric disorders. CAM practitioners sometimes give patients advice about medications resulting in potentially serious safety issues when medication doses are inappropriately changed or abruptly discontinued, leading to increased risk of relapse and unnecessary treatment delays. In view of these liability issues, patients should be referred to CAM practitioners *only after verifying their credentials, training, and reputation as competent clinicians*. Directories of professional societies for many complementary and alternative therapies are available through NCCAM.

OVERVIEW OF CAM ASSESSMENT APPROACHES IN MENTAL HEALTH CARE

Assessment in conventional biomedicine is based on the assumption that reasonable inferences about the psychological, social, or biological *causes* of symptoms can be made when core symptoms have been identified. Recent advances in neuroscience suggest that the causes of mental illness may be both more *subtle* and more *complex* than the prevailing biopsychosocial model is capable of elucidating. Models of consciousness based on complexity theory, quantum mechanics, and other emerging theories may be consistent with the assumptions of "energetic" assessment methods used in traditional healing. CAM assessment approaches may eventually permit a more *complete* "picture" of the psychological, biological, and *energetic* causes of symptoms. CAM approaches used to assess the nature and causes of mental illness include assays of biological structure or function, measurements of scientifically validated forms of energy or information, and measurements of postulated *subtle* forms of energy or information that cannot be explained by current science. Quantitative electroencephalography (QEEG) and heart rate variability (HRV) analysis are examples of scientifically validated assessment methods that provide valuable clinical information about functional dysregulations of electromagnetic activity in the brain and heart related to symptoms of mental illness. Assessment approaches that purport to detect *subtle* energetic phenomena that may underlie physical and mental illness include pulse diagnosis in Chinese, Tibetan, and Ayurvedic medicine and analysis of electrodermal potentials following acupuncture or medical *qigong*.

Functional medicine is an interdisciplinary field that uses established quantitative analysis methods to assess relationships between nutritional

status, neurotransmitters, endocrine and immune function, and psychological symptoms. Functional tests used to evaluate mental health problems include urinary assays of neurotransmitters and their metabolites and serum assays of cholesterol, triglycerides, vitamins, minerals, amino acids and their metabolites, hormones, fatty acids, proinflammatory cytokines (e.g., interleukin-6 [IL-6], IL-8, and IL-1b), and immunologic factors. Specific biomarkers of particular psychiatric disorders have not been identified, and functional assessment findings generally do not translate into specific treatments that yield improved outcomes. For example, although low serum folate has been causatively linked to depressed mood, research findings are inconsistent, and this *apparent* relationship may be due to chance, bias, confounding, or reverse causality. Preliminary findings suggest that total cholesterol and triglyceride levels may be inversely correlated with the severity of depressed mood. Depressed patients who attempt suicide frequently have abnormal low serum total cholesterol and triglyceride levels, which may provide useful clinical markers for suicide risk. Nonsuicidal depressives tend to have cholesterol levels in the vicinity of 180 mg/dL while severely depressed suicidal patients tend to have significantly lower serum total cholesterol levels in the vicinity of 150 mg/dL. On the basis of these findings, it is always prudent to check cholesterol levels of severely depressed patients, make appropriate dietary recommendations, and address issues related to the use of cholesterol-lowering medications. In contrast to abnormal low cholesterol in severe depressed patients, *elevated* serum cholesterol may be a sensitive marker for various anxiety disorders, including generalized anxiety disorder, panic disorder, and possibly also OCD. Amelioration of anxiety symptoms in response to prescription anxiolytics is often associated with a reduction of total serum cholesterol to normal levels. The findings of a large observational study suggest that abnormal low serum levels of certain B vitamins (thiamin, B_6, and niacin) and vitamins C and E are associated with an increased risk of generalized anxiety. Assessment of vitamin deficiency states is important when working with patients who do not take supplements or who are at risk of one or more vitamin deficiency syndromes because of chronic malnutrition or a medical problem that interferes with normal absorption or metabolism.

The risk of developing Alzheimer disease may be associated with chronic low blood levels of zinc and magnesium, but research findings are inconsistent. High levels of zinc are neurotoxic and are believed to promote formation of amyloid plaques, whereas histopathology studies of the brains of patients with Alzheimer disease reveal deficient zinc in the brain. Low serum zinc levels result in impaired zinc transport to the brain and have been cited as a possible factor in the formation of amyloid fibrils. Consistently reduced red blood cell membrane levels of two fatty acids—arachidonic acid (AA) and DHA—have been reported in nonmedicated

schizophrenics compared to medicated schizophrenics and normal controls. Low brain levels of fatty acids and cholesterol may indirectly affect the number and activity of neurotransmitter receptors by modifying their three-dimensional structure in the lipid bilayer of nerve cell membranes. Assessing flushing response following niacin challenge may provide useful information when evaluating psychotic patients. Research findings suggest that normal flushing response to niacin is attenuated in schizophrenia and other psychotic disorders. Approximately 80 percent of patients with schizophrenia do not flush when challenged with niacin compared to healthy controls. The mechanism of action probably involves dysregulation of phospholipid-dependent signal transduction and increased phospholipase A2 activity. In one series, 43 percent of patients with schizophrenia experienced abnormal low vasodilation in response to a 200-mg challenge dose of niacin, compared to only 6 percent of individuals diagnosed with bipolar disorder.

There is preliminary evidence for the clinical utility of HRV monitoring and QEEG in psychiatric assessment; however, these approaches are still mainly used for research. Abnormal electroencephalography (EEG) findings occur in up to 40 percent of depressed patients, and "small sharp spikes" are often present in severely depressed suicidal patients. QEEG analysis of unipolar depressed patients typically reveals increased alpha or theta power and decreased interhemispheric coherence. In contrast, bipolar depressed patients often have reduced alpha activity and increased beta power. Reduced alpha activity and increased beta power are more likely in bipolar disorder. EEG changes frequently associated with anxiety include decreased alpha activity in generalized anxiety, increased theta activity in OCD, and paroxysmal activity in patients who experience panic attacks. QEEG findings may predict differential response rates of patients with OCD to conventional medications. In one series, approximately 80 percent of OCD patients who exhibited increased alpha power responded to SSRIs, compared to 80 percent of OCD patients with increased theta activity who failed to respond to SSRIs. Abnormal QEEG findings associated with other anxiety disorders are highly inconsistent. Neurometric brain mapping is a specialized QEEG approach that compares EEG characteristics of the individual being evaluated with normative databases for the same age. Neurometric mapping helps clarify functional brain correlates of cognitive impairment and yields information that is useful for planning EEG biofeedback protocols addressing specific kinds of dysfunction (see later discussion of EEG biofeedback). Neurometric brain mapping is increasingly being used to differentiate cognitive impairments due to head injuries, medical disorders, progressive dementia, alcohol or substance abuse, depressed mood, and learning disorders.

Relatively greater auditory-evoked potentials (AER) correspond to lower activity of brain serotonin and predict improved response of unipolar

depressed patients to SSRIs. This finding has been corroborated by prospective studies examining the relationship between *cordance* and response to antidepressants that increase brain serotonin levels. Cordance is a measure of localized electrical brain activity relative to averaged brain EEG patterns. More than one-half of patients with severe depressed mood who subsequently had the highest response rates to SSRIs showed significant decreases in prefrontal cordance during the first 48 hours of therapy, suggesting that improvement was due to normalization of low brain serotonin levels. In contrast, nonresponders to SSRIs or placebo did not show negative prefrontal cordance following treatment, suggesting that low serotonin was not the primary cause of depressed mood in these cases. Differences in prefrontal EEG cordance may also predict response differences of depressed patients to homeopathic remedies.

Intuitive assessment approaches used by "energy healers" are based on the belief that *imbalances* in postulated subtle "energies" manifest as psychological symptoms. Analysis of the pulse is used to assess energetic imbalances associated with both physical and mental symptoms in Chinese medicine, Ayurveda, and Tibetan medicine. A treatment plan consisting of psychological, physical, and energy "balancing" therapies is constructed with the goal of normalizing energetic imbalances and improving health. Anecdotal reports suggest that subtle characteristics of the pulse described in Chinese, Ayurvedic, or Tibetan medical theory correspond to specific energetic *imbalances* that manifest as physical or psychiatric symptoms. For example, mild to moderate symptoms of anxiety and depressed mood are often associated with a "tight" or rapid rate when the pulse is palpated at the pericardium position or as a subtle sensation of "smooth vibration" over the entire pulse at all depths. More severe symptoms of mental illness are often associated with a "rough vibration" over the entire pulse or a sensation of "slipperiness" in some aspects of the pulse. The most severe psychiatric symptoms, including psychosis and acute mania, are reportedly associated with the most abnormal energetic pattern of the pulse.

The measurement of electrodermal potentials at specific acupoints may provide useful information when assessing clinical responses to acupuncture or *qigong*. Electrodermal skin testing (EDST) may also help determine whether conventional or subtle energy treatments of anxiety are likely to be effective for a particular patient. Healing Touch practitioners conducted six trials comparing the anxiety-reducing efficacy of conventional relaxation techniques with Healing Touch and concluded that EDST reliably measured changes in "bioenergy" in meridians following both conventional and subtle energy treatments of anxiety.

COMPLEMENTARY, ALTERNATIVE, AND INTEGRATIVE APPROACHES USED TO TREAT MENTAL ILLNESS

Overview

Advances in the understanding of the pathogenesis of mental illness have important implications for therapeutic uses of natural products in the treatment of depressed mood, bipolar disorder, schizophrenia, and other disorders. Omega-3 fatty acids have beneficial effects at many levels in the body and brain and are now commonly recommended as adjuvants to antidepressants, mood stabilizers, and other psychotropic medications. For example, mitochondrial psychiatry postulates that genetic or acquired factors that interfere with mitochondrial energy metabolism predispose certain individuals to developing bipolar disorder or other severe chronic psychiatric disorders. *N*-acetyl-cysteine (NAC) is an amino acid that is being investigated for its potential role as a mitochondrial modulator that may mitigate symptoms of bipolar disorder.

While some CAM modalities are substantiated by consistent positive findings from well-designed studies, the majority of CAM therapies reviewed in this section are used in the *absence* of compelling evidence. Functional medicine is an interdisciplinary discipline that conceptualizes the causes of illness in relationship to *informational changes* in intercellular communication processes that cause pathology. Functional medicine approaches used in mental health care include dietary modification, amino acid therapy, and supplementation with high doses of specific vitamins and minerals. The postulated mechanisms of action of so-called "energy" therapies, including Reiki, *qigong,* Healing Touch, and energy psychology, pose complex challenges to contemporary Western-style research methods. A more complete understanding of clinical benefits associated with light, electricity, and sound, as well as so-called *"energy"* therapies, may require conceptual advances in the foundations of science and greater openness to emerging models of physics including complexity theory, quantum mechanics, and quantum information theory.

Table 31.4–2.
Nonconventional Modalities Reviewed in This Section

Category	Modality
Biological treatments	Dietary modification
	Vitamins
	Minerals
	Amino acids
	Essential fatty acids
	Nutrient formulas
	Botanicals
	Essential oils
	Hormones and prohormones
Somatic therapies	Exercise
	Expressive movement therapy
	Animal-facilitated psychotherapy
	Snoezelen
	Massage

Mindfulness training and mind–body practices	Applied relaxation Yoga Meditation
Energy–information modalities validated by Western science	Music and sound Light High-density negative ions Electroencephalography (EEG) biofeedback Heart rate variability (HRV) biofeedback Virtual reality graded exposure therapy Microcurrent electrical stimulation
Energy–information modalities not validated by Western science	Acupuncture Homeopathy Healing touch and therapeutic touch *Qigong* Reiki Energy psychology Craniosacral therapy and somatoemotional release Shamanic healing

Approaches used as therapeutic modalities in biomedicine and CAM can be conceptualized as belonging to one of five general categories corresponding to broad differences in postulated mechanisms of action (Table 31.4–2):

- *Biological treatments* operate through a presumed discrete biological mechanism of action at the molecular or cellular level (examples include conventional medications, vitamins, minerals, medicinal herbs, amino acids, and omega-3 fatty acids).
- *Somatic therapies* improve general well-being and affect the body as a whole but do not operate through a discrete (i.e., molecular or cellular) biological mechanism of action (examples include exercise, massage, and dance therapy).
- *Mind–body and mindfulness practices* use a particular mental technique or combine a mental technique with physical training with the goal of optimizing mental or mind–body functioning and alleviating symptoms (examples include yoga and mindfulness-based stress reduction [MBSR]).
- *Scientifically validated forms of energy or information* have been validated by current science. Examples include electromagnetic fields, bright light, dim light, microcurrent electrical stimulation, high-density negative ions, music, binaural sound, and biofeedback. Specific therapeutic applications use different durations, frequencies, and power characteristics of these established forms of energy to achieve beneficial effects.
- *Postulated forms of energy or information* have not been validated by current science but are theorized to be the mechanisms underlying "energy" healing approaches including Reiki, *qigong,* prayer, Healing

Touch, Therapeutic Touch, and energy psychology.

Identifying Appropriate CAM Modalities

Deciding on which CAM or integrative approach to recommend to a patient poses many complex issues. When formulating clinical treatment recommendations, five basic issues should be kept in mind:

- The symptom or disorder that is the focus of clinical attention.
- The patient's history of response to previous treatments for the same symptom(s) or disorder.
- The relative strength of evidence for various biological, somatic, mind–body, and energy therapies.
- Practical and personal considerations, including safety, cost, availability of treatment, and cultural and personal preferences that will lead to realistic and acceptable treatment recommendations.
- Appropriate methods for assessing treatment response.

The balance of this section is devoted to a review of evidence for CAM and integrative modalities currently used in developed world regions to treat or *self-treat* common mental health problems. Because of space constraints, only modalities that have been validated by research findings are discussed in detail, whereas CAM and integrative approaches supported by limited evidence are briefly reviewed. The reader should keep in mind that future studies will probably validate the efficacy claims of certain CAM and integrative approaches not currently supported by strong evidence and refute other approaches that are now in common use.

Reviews of Complementary, Alternative, and Integrative Approaches

Biological Treatments

DIETARY MODIFICATION. The general health of the digestive system modulates immune functioning and brain activity through the *microbiome-gut-brain* axis. Recent findings point to a possible link between imbalances in bowel microflora, increased inflammation of the mucosal lining of the intestines, and systemic immune dysregulation resulting in increased risk of depressed mood. Preliminary findings from animal studies and early human clinical trials suggest that probiotics have therapeutic effects in autism, depressed mood, and anxiety. Research findings on the relationship between nutrition and mental health are inconsistent; however, reducing or eliminating refined sugar and caffeine from the diet significantly improves mood in some depressed patients. Foods rich in folate, pyridoxine (B_6), and methyl-cobalamin (B_{12}) are therapeutic because these vitamins are essential cofactors in the synthesis of serotonin, dopamine, and norepinephrine. Increased consumption of fish high in

omega-3 essential fatty acids may provide a protective effect against depressed mood. Important dietary sources of omega-3s include salmon, halibut, other deep-sea fish, and flaxseed oil. Relatively greater intake of omega-3 fatty acids (i.e., compared to omega-6 fatty acids) is correlated with lower C-reactive protein levels and reduced risk of inflammatory-mediated disorders, including possibly depression and Alzheimer disease. Prevalence rates of severe depression and suicide attempts are lower in countries where fish is an important part of the average diet. In Japan, where fish consumption is very high, the annual incidence of depressed mood is only 0.12 percent, compared to 6 percent in New Zealand, where fish consumption is relatively low. In a 6-month study, 452 men with histories of cardiovascular disease were randomized to a diet high in fatty fish versus no fish. No significant group differences were found in new cases of anxious or depressed mood at the end of the study period, and more patients in the high-fish group reported depressed mood or anxiety. Caffeine consumption increases serum epinephrine, norepinephrine, and cortisol levels, resulting in general feelings of "nervousness" and increased risk of panic attacks in predisposed individuals.

A large epidemiological study concluded that fish consumption two to three times weekly significantly reduced the risk of cognitive decline in the elderly. High fish consumption was inversely correlated with cognitive impairment. In contrast, individuals who preferred foods rich in linoleic acid (an omega-6 fatty acid) exhibited significantly higher rates of cognitive decline.

Persons who consume a high-fat, high-calorie diet are at significantly greater risk of developing Alzheimer disease than persons who have moderate fat intake and restrict total calories. Excess caloric intake and high fat intake promote the formation of damaging free radicals that cause diffuse neuropathological changes in the brain. A meta-analysis of findings from 18 community-wide studies concluded that the risk of Alzheimer disease increased linearly at a rate of 0.3 percent with every 100-calorie increase in daily intake. Average daily fat consumption was strongly correlated with increased risk of developing dementia. Moderate but not heavy consumption of wine (two to four glasses per day) is associated with a reduced risk of Alzheimer disease. In a large 5-year cohort study, 19,000 women aged 70 to 80 years who drank 14 g of alcohol daily (the equivalent of one and a half glasses of wine or one 12-oz beer) were 20 percent less likely to experience significant cognitive impairment compared to nondrinkers. Lifestyle factors—including exercise and alcohol consumption—and educational level are associated with dietary preferences, and thus a relatively lower risk of developing Alzheimer disease cannot be ascribed to fish consumption alone.

Chronic alcohol abuse is often associated with malnutrition due to malabsorption of essential nutrients through the mucosa of the stomach

and small intestines resulting in significantly reduced blood levels of thiamin, folate, B_6, and important trace minerals. Hypoglycemia can result from toxic effects of alcohol on the liver manifesting as confusion, anxiety, and impaired cognitive functioning. Alcoholics who improve their general nutrition probably have a better chance of maintaining sobriety than those who do not. Eighty-one percent of alcoholics who stayed on a caffeine-free, low–refined sugar diet high in wheat germ and fruit remained abstinent at 6 months.

VITAMINS. Depressed individuals with low serum folate and B_{12} levels and elevated serum homocysteine levels often fail to respond to conventional antidepressants. Elevated serum homocysteine is a marker for folate, B_6, and B_{12} deficiencies and is a risk factor for Alzheimer disease and other neurodegenerative diseases. The Framingham community-based study confirmed that elevated plasma homocysteine levels (greater than 12 μmol/L) double the risk of developing Alzheimer disease and non–Alzheimer dementias.

B-Family Vitamins. Folate is required for the synthesis of SAMe from homocysteine. Folate deficiency results in increased serum homocysteine levels and increased risk of depressed mood. Folate in the form of l-methylfolinic acid does not require methylenetetrahydrofolate (MTHF) reductase to become activated, more readily crosses the blood–brain barrier, and is a more effective adjuvant than regular folate. Clinical improvement in depressed patients treated with an SSRI and l-methylfolate (0.5 to 1 mg) was 30 percent greater than that in matched patients treated with an SSRI only. A systematic review of controlled studies (total $N = 247$) concluded that folate augmentation (1 to 15 mg per day) enhanced the efficacy of conventional antidepressants. The findings of a small double-blind study ($N = 75$) suggested that folic acid at 200 to 400 μg per day might enhance the therapeutic benefits of lithium carbonate in acutely manic patients. Preliminary findings suggest that biomarkers associated with inflammation may be useful for identifying nonresponders to SSRIs' major depressive disorder who are more likely to respond to adjunctive L-methylfolate.

Widespread folate deficiency in patients with schizophrenia is probably caused by chronic malnutrition and effects of conventional antipsychotic medications on absorption. Abnormal low brain folate levels in patients with schizophrenia may also result from a heritable deficiency in the enzyme glutamate carboxypeptidase II (GCPII), which is required for folate absorption through the gut. Patients with schizophrenia who take daily folic acid in the form of methylfolate 15 mg together with a conventional antipsychotic may have fewer positive and negative psychotic symptoms and may respond more rapidly than patients taking antipsychotics alone.

Low serum levels of folic acid, niacin, and thiamin are associated with cognitive impairment in general. However, the evidence for vitamin B supplementation as a treatment of cognitive impairment including dementia is inconclusive. Individuals with dementia frequently have normal red blood cell folate levels and fail to improve with supplementation. In a prospective study (*N* = 370), deficiency of folate or B_{12} doubled the risk of developing Alzheimer disease. However, a systematic review of four studies did not find strong evidence supporting the use of folic acid with or without B_{12} as a treatment of dementia or other forms of severe cognitive impairment. In a double-blind, placebo-controlled study, after 4 weeks of treatment with 50 mg per day of methylfolate, depressed patients with dementia experienced significant improvements in both mood and memory. Folate supplementation can mask the macrocytic anemia caused by B_{12} deficiency; thus, patients should be checked for B_{12} deficiency before starting folate. However, folate supplementation in the form of L-methyl-folinic acid does not mask B_{12} deficiency and can be recommended without first checking B_{12} levels.

Vitamin B_{12} is often recommended for elderly nondemented individuals who complain of impaired cognition; however, research findings are inconsistent. In a pilot study (*N* = 22), 18 elderly patients with abnormal low serum B_{12} levels and cognitive impairment were treated with B_{12} (1 mg per day intramuscularly) for 1 week, followed by weekly 1-mg injections for 1 month, then monthly 1-mg injections for 6 months. Patients who had been cognitively impaired for less than 1 year experienced significant improvement, whereas patients who had been impaired for longer than 1 year became worse, suggesting that supplementation may ameliorate symptoms related to B_{12} deficiency only in the early stages of cognitive impairment. Chronic alcoholics are typically deficient in thiamin, folate, B_6, and B_{12} because of toxic effects of alcohol on the mucosal lining of the stomach and small intestine. For individuals who are unable to remain abstinent, taking 2 g of vitamin C close to the time of alcohol consumption may reduce or prevent hangover symptoms by neutralizing metabolites of alcohol that cause oxidative damage to the body and brain. Thiamin is essential for the metabolism of acetylcholine and its release from presynaptic neurons. Depressed patients report improved mood and energy with daily thiamin supplementation of 50 mg. When taken together with a conventional antipsychotic medication, 500 mg thrice daily of thiamin (vitamin B_1) may lessen the severity of psychotic symptoms in patients with chronic schizophrenia. Thiamin-dependent enzymes may be abnormal in Alzheimer disease. Findings on high-dose thiamin (3 to 8 g per day) in Alzheimer disease are inconsistent. Animal studies suggest that low serum thiamin levels are associated with increased alcohol craving. Niacin in the

form of 1.25 g per day of nicotinamide taken with a meal before drinking may protect the liver against the acute toxic effects of alcohol in individuals who have relapsed. Early double-blind, RCTs reported that large doses of niacin (up to 3 to 8 g per day) ameliorated symptoms of schizophrenia. A combined regimen of high-dose niacin and a conventional antipsychotic medication reportedly resulted in differentially greater improvement in positive psychotic symptoms compared to schizophrenic patients treated with conventional antipsychotics alone. A subsequent review of 53 trials on niacin and other vitamins in the treatment of schizophrenia and other chronic psychiatric disorders identified methodological flaws in early studies and concluded that findings reached only marginal statistical significance. Nevertheless, high-dose niacin has become a standard therapy for schizophrenia in orthomolecular psychiatry, and case reports of dramatic outcomes continue to be published. A retrospective analysis of negative findings of controlled trials suggested that nonmedicated patients with schizophrenia in the early acute phase of illness may experience significant and sustained benefits from high doses of niacin, in contrast to medicated patients with schizophrenia or individuals in the chronic phase of the illness. Supplementing individuals with folate can mask the macrocytic anemia associated with B_{12} deficiency and exacerbate depressed mood and cognitive symptoms. Serum B_{12} levels should be routinely checked before recommending folate supplementation. Folic acid interferes with phenytoin metabolism, lowering serum levels and placing individuals with seizure disorders at increased risk.

Inositol is regarded as a member of the B family of vitamins and is synthesized in the body in the form of myo-inositol. Inositol is the precursor of phosphatidyl inositol, a second messenger required for normal neuronal function. Inositol in doses up to 20 g per day may improve depressed mood; however, findings are inconsistent. A systematic review and meta-analysis of seven RCTs (N = 242) on inositol in mood and anxiety disorders found nonsignificant benefits for inositol over placebo for depressed mood, OCD, PTSD, and generalized anxiety, and a trend toward superior response over placebo for premenstrual dysphoric disorder. In a 4-week, double-blind, crossover study (N = 21), inositol at 12 g per day and imipramine were equally effective in reducing the frequency and severity of panic attacks in individuals diagnosed with panic disorder with and without agoraphobia. In a small crossover study (N = 20), inositol (up to 18 g per day) and fluvoxamine (150 mg per day) had comparable efficacy in reducing the frequency of panic attacks. Larger studies are needed to evaluate possible synergistic effects of inositol when used in combination with conventional antidepressants or mood stabilizers. Some patients taking inositol report gastrointestinal distress.

Case reports suggest that choline (a vitamin in the B complex) reduces the severity of mania. Choline is required for synthesis of acetylcholine

(Ach). In an open trial (*N* = 6) of treatment-refractory, rapid-cycling bipolar disorder patients on maintenance lithium therapy, four patients experienced reductions in the severity of manic symptoms, with 2 to 7.2 g per day of free choline. Open trials and one small controlled trial suggested that supplementation with phosphatidylcholine (PC) at 15 to 30 g per day might reduce the severity of both mania and depressed mood in bipolar patients and that symptoms worsened when PC was discontinued. Cytidine 5-diphosphocholine (CDP-choline) is a precursor of PC and other important structural phospholipids in neuronal cell membranes. CDP-choline increases mitochondrial energy production, increases central nervous system (CNS) noradrenaline and dopamine levels, and has general neuroprotective effects following hypoxic or ischemic brain injury. Preliminary findings suggest that CDP-choline (1,000 mg per day) may improve cognitive performance in the early stages of Alzheimer disease. Two systematic reviews concluded that CDP-choline has consistent positive effects on recovery rate following stroke and mitigates the severity of cognitive impairment related to cerebrovascular disease in elderly individuals.

MINERALS. Magnesium deficiency is widespread in Western countries and may be associated with cardiovascular disease, depressed immune functioning, depression, diabetes, insomnia, and other disorders. Case histories suggest that oral magnesium in the form of magnesium glycinate or taurate (125 to 300 mg per day) may result in rapid recovery from severe depressed mood and may ameliorate symptoms of anxiety, irritability, insomnia, postpartum depression, and short-term memory. A study on elderly depressed persons with diabetes suggested that hypomagnesemia, defined as serum magnesium levels less than 0.74 mmol/L, might be an independent risk factor for depression in this population.

Magnesium supplementation at 500 to 1,500 mg per day may improve neuropsychological deficits associated with chronic alcohol abuse by improving cerebral blood flow, which is often diminished in this population. Animal studies and preliminary human findings suggest that abnormally high copper levels may be a causative factor in the pathogenesis of Alzheimer disease by damaging the molecule required to eliminate amyloid beta plaques soon after they are formed. Large cohort studies are needed to confirm this hypothesis. Deficiencies in zinc, manganese, and iron are common in alcoholics and worsen with chronic heavy use.

Zinc is an essential trace element for normal brain function and may mediate anti-inflammatory effects via several cytokines. Zinc has general neuroprotective effects, is involved in hippocampal neurogenesis, and modulates the activity of the hypothalamic-pituitary axis. Increased risk of depressed mood associated with abnormal low zinc serum levels may be mediated by dysregulation of the immune system. In a cross-sectional

study of the relationship between dietary zinc intake and depressed mood, low serum zinc levels were strongly correlated with increased risk for depressed mood. Low brain zinc levels may cause diffuse neuronal damage through increased free radical formation and exacerbate *N*-methyl-D-aspartate (NMDA) excitotoxicity, resulting in increased susceptibility to alcohol withdrawal seizures. Preliminary evidence from randomized controlled studies supports the antidepressant efficacy of zinc supplementation used as an adjuvant to prescription antidepressants.

***S*-ADENOSYL METHIONINE.** SAMe is a methylation product of adenosine triphosphate and the amino acid methionine. SAMe is essential for many important biosynthetic pathways, including transmethylation reactions required for the synthesis of serotonin, dopamine, and norepinephrine. After donating a methyl group, SAMe is converted to *S*-adenosyl-homocysteine, which has been implicated in the pathogenesis of Alzheimer disease, atherosclerosis, and other inflammatory diseases. SAMe is a widely used treatment of depression, arthritis, and fibromyalgia. The compound is available by prescription in Europe in oral, injectable, or intravenous forms, but in the United States, it can be obtained only as a nonstandardized, over-the-counter oral supplement. Folic acid and B_{12} are required cofactors for remethylation of homocysteine back to SAMe and are routinely recommended to patients taking SAMe. The research literature on SAMe in depressed mood is extensive including studies on both parenteral and oral preparations. To date, more than 45 RCTs have been conducted with consistent findings of antidepressant efficacy comparable to conventional antidepressants when used as a monotherapy or adjunctive agent. SAMe is generally well tolerated with infrequent reports of nausea, dry mouth, dizziness, and restlessness. Headaches, irritability, and sedation have been reported when SAMe is used adjunctively with antidepressants; however, no cases of serotonin syndrome have been reported. SAMe may unmask mania or hypomania and should be avoided in patients diagnosed with bipolar disorder.

AMINO ACIDS. Amino acids used to treat or self-treat psychiatric symptoms include L-tryptophan, 5-hydroxytryptophan, L-taurine, tyrosine, acetyl-L-carnitine, and glycine.

5-Hydroxytryptophan, l-Tryptophan and Branch Chained Amino Acids. The amino acids 5-hydroxytryptophan (5-HTP) and L-tryptophan are precursors of serotonin. 5-HTP is generally recommended over L-tryptophan because it crosses the blood–brain barrier at a higher rate, is converted into serotonin more efficiently than L-tryptophan, and has a more pronounced antidepressant effect. A systematic review of RCTs on 5-HTP or L-tryptophan in depressed mood identified 108 studies, but design problems and small study size

limited analysis of findings to only two studies (*N* = 64) that met inclusion criteria. On the basis of those limited findings, the reviewers concluded that 5-HTP is *probably more effective* than placebo at doses of 100 to 300 mg per day for moderate depressed mood. Sixty-three depressed patients randomized to fluvoxamine 150 mg per day or 5-HTP 300 mg per day reported equivalent improvements in mood. In an open study, almost half of treatment-refractory depressed patients (*N* = 100) responded to 5-HTP (up to 600 mg per day) in combination with 150 mg per day of carbidopa after several weeks of treatment. Case reports show that some treatment-refractory patients improve when 300 mg per day of 5-HTP is combined with carbidopa (a peripheral MAOI inhibitor), tricyclic antidepressants, MAOIs, or SSRIs. l-Tryptophan is typically used at night because of its sedating properties and is dosed between 1.5 and 5.0 g, depending on therapeutic response. 5-HTP is moderately sedating and doses greater than 100 mg should be taken at bedtime. Findings of RCTs show consistent beneficial effects of 5-HTP for generalized anxiety and panic disorder. Patients diagnosed with generalized anxiety disorder randomized to 5-HTP with carbidopa versus clomipramine reported equivalent and significant reductions in generalized anxiety. 5-HTP at 50 to 100 mg thrice daily is well tolerated when used at recommended doses alone or in combination with conventional antianxiety agents. Preliminary findings support that combining 1 to 3 g per day of L-tryptophan and early-morning bright light exposure is more efficacious than either treatment alone in patients diagnosed with seasonal affective disorder. In a double-blind study, 58 percent of generally anxious patients (*N* = 79) randomized to 3 g per day of L-tryptophan reported significantly greater reductions in baseline anxiety compared to placebo. L-Tryptophan, 1 g, at bedtime reduces time to sleep onset in mild situational insomnia, and doses up to 15 g at bedtime may be necessary for severe insomnia. Case reports suggest that nighttime use of 5-HTP at 300 to 600 mg may improve mild to moderate insomnia and lessen sleep disturbances related to obstructive sleep apnea and narcolepsy. Combining 2 g of L-tryptophan with 20 mg per day of fluoxetine resulted in more rapid response and improved sleep quality in depressed patients complaining of chronic insomnia. No cases of serotonin syndrome or other serious adverse effects were reported with these doses. Transient mild adverse effects reported with L-tryptophan and 5-HTP include nausea, constipation, dry mouth, blurred vision, drowsiness, and decreased libido. Cases of serious serotonin syndrome have been reported when MAOIs are combined with L-tryptophan, including orthostatic hypotension, hyperreflexia, diaphoresis, and delirium. Fifteen hundred cases of eosinophilia-myalgia syndrome (EMS) and 37 deaths in patients taking L-tryptophan were reported in the late 1980s and early 1990s. All cases were traced to contaminants in a single batch of one over-the-counter brand of

L-tryptophan. The manufacturing problem that resulted in the contaminated batch was identified and corrected, and there have been no subsequent reports of EMS associated with L-tryptophan.

Patients with bipolar disorder may be genetically susceptible to mood swings when L-tryptophan, tyrosine, or phenylalanine is restricted or excluded from the diet. Patients with bipolar disorder who were stable on lithium and subsequently subjected to acute dietary tryptophan depletion were found to have reduced brain serotonin levels. A mixture of bioavailable amino acids excluding tyrosine and phenylalanine (the precursor of tyrosine) may reduce CNS dopamine levels in patients with bipolar disorder, ameliorating manic symptoms and improving cognitive functioning. In a small (N = 25), 1-week RCT, patients with bipolar disorder randomized to 60 g per day of a special tyrosine-free amino acid drink versus placebo reported significant reductions in symptom severity within 6 hours following treatment. Improvements were sustained with repeated administration of the amino acid drink and persisted for 1 week following the end of the study. These findings suggest that the ingestion of an amino acid drink that results in acute tyrosine depletion may be an effective treatment of acute mania. These findings should be regarded as preliminary pending confirmation by larger controlled studies.

Acetyl-L-Carnitine. Acetyl-l-carnitine occurs naturally in the brain and the liver. RCTs have evaluated the efficacy of ALC supplementation in severely depressed patients, elderly depressed patients, and depressed patients with dementia. Neuroprotective effects of ALC including enhanced mitochondrial energy production may improve depressed mood and reduce the severity of cognitive impairments in normal aging, dementia, or traumatic brain injury. A review of RCTs on ALC in depressed mood showed superiority over placebo for moderate depressed mood and equivalent efficacy and greater tolerability compared to the antidepressants fluoxetine and amisulpride. A typical antidepressant dose is 500 mg four times daily.

ALC is widely used to treat cognitive impairments related to dementia and other neurodegenerative disorders; however, findings are inconsistent. In a 2-month RCT, depressed patients with dementia randomized to 3 g per day of ALC experienced significantly greater improvements in mood and global cognitive functioning than the placebo group. RCTs show that 1,500 to 3,000 mg per day of ALC improves overall performance on tests of reaction time, memory, and cognitive performance in patients with dementia and may slow the rate of progression of cognitive impairment. A systematic review of 11 RCTs of ALC in dementia confirmed significant positive effects at weeks 12 and 24. However, improvements were not sustained at 1 year with continued treatment. Younger individuals experiencing presenile dementia may benefit more than older individuals,

and for unclear reasons, 62 years may be the "optimal age" beyond which ALC loses much of its cognitive-enhancing effects. Findings of RCTs suggest that ALC at 1.5 to 3 g per day improves age-related symptoms of cognitive impairment in healthy elderly nondemented individuals and depressed elderly patients. There is preliminary evidence that ALC at 2 g per day improves memory and word recall in cognitively impaired abstinent alcoholics. ALC is generally well tolerated and may be safely used in combination with conventional antidepressants and cholinesterase inhibitors.

l-Theanine. L-Theanine (*N*-ethyl-L-glutamine) is found in green tea and is widely used in China, Japan, and other Asian countries to treat anxiety and depressed mood. Animal studies confirm that L-theanine increases brain serotonin, dopamine, and γ-aminobutyric acid (GABA) levels and may have general neuroprotective effects. Anxiety-reducing effects of L-theanine are believed to be mediated by enhanced alpha brain wave activity in the occipital and parietal regions that are dose dependent and reportedly similar to EEG changes observed during meditation. Calming effects may last 8 to 10 hours and are usually experienced within 30 minutes following ingestion of L-theanine at doses between 50 and 200 mg. Findings of clinical studies on L-theanine are inconsistent. In one small RCT (N = 16), healthy adult volunteers randomized to alprazolam (1 mg) versus L-theanine (200 mg) or placebo experienced equivalent and nonsignificant anxiolytic effects during an experimentally induced anxiety state. In another small RCT (N = 12), individuals taking L-theanine experienced relatively greater reduction in acute stress response as measured by changes in HRV and salivary immunoglobulins (s-IgA) compared to the placebo group. These findings suggest that stress-reducing effects of L-theanine may be mediated by inhibition of cortical neuron excitation. In a crossover study, healthy adults were randomized to L-theanine (250 mg) alone or in combination with caffeine (150 mg). Compared to the L-theanine–only group, the combination group experienced improved visual information processing, reduced mental fatigue, faster reaction time and faster working memory, generally increased alertness, and fewer headaches. Moderately severe anxiety symptoms may respond to a regimen of 200 mg twice daily, and more severe anxiety may require doses up to 600 to 800 mg per day divided into 200-mg increments every 3 to 4 hours. L-Theanine is generally well tolerated, and there are no published reports of serious adverse side effects or interactions with other natural products or conventional medications.

N-Acetyl-Cysteine. NAC is being investigated for its role as a mitochondrial modulator and may reduce the severity of depressive symptoms in bipolar disorder. In a 2-month open label phase of a randomized placebo

controlled study, 149 moderately depressed patients diagnosed with bipolar disorder were treated with NAC (1 g BID) as a maintenance therapy. By study end symptoms of depressed mood had decreased significantly and indices of functioning and life quality improved. In a subsequent double-blind placebo-controlled study, 149 stable bipolar patients randomized to NAC (2 g per day) versus placebo while continuing on treatment as usual did not experience significant differences in recurrence rates or outcome measures of functioning or quality of life. At study, end point symptom scores remained low and there were minimal changes in outcome measures in the NAC and placebo groups.

Branch Chain Amino Acids and Excluding Tyrosine. Preliminary findings support that oral administration of certain branch-chain amino acids may rapidly improve acute manic symptoms by interfering with the synthesis of the catecholamine neurotransmitters norepinephrine and dopamine. A mixture of bioavailable amino acids excluding tyrosine and phenylalanine (the precursor of tyrosine) may reduce brain dopamine in bipolar patients, reducing the severity of mania and improving global cognitive functioning. In a 7-day double-blind placebo-controlled study, bipolar patients (N = 25) randomized to a special tyrosine-free amino acid drink 60 g per day versus placebo reported significant reductions in mania severity within 6 hours of treatment. Improvements were sustained with repeated administration, and the effect persisted 1 week following the study end point.

Glycine. The amino acid glycine functions as an inhibitory neurotransmitter. Glycine and glutamate are required coagonists for the activation of NMDA receptors. Case reports and small RCTs suggest that glycine and other naturally occurring NMDA agonists help correct NMDA receptor dysfunction involved in the pathogenesis of negative symptoms and cognitive impairments in schizophrenia. A small open trial found that negative psychotic symptoms improve with glycine at doses up to 60 g per day. In a 6-week crossover RCT (N = 22), patients with treatment-resistant schizophrenia randomized to glycine at 0.8 g/kg in addition to their conventional antipsychotic versus placebo plus antipsychotic experienced a 20 to 30 percent greater improvement in negative symptoms compared to drug plus placebo. In a multicenter 16-week RCT (N = 157), inpatients diagnosed with schizophrenia or schizoaffective disorder were randomized to glycine versus D-cycloserine or placebo. No significant differences in outcomes were observed in negative psychotic symptoms or cognitive impairment. In a 10-week open-label prospective study, stable patients diagnosed with schizophrenia (N = 29) randomized to glycine (up to 60 g per day) while continuing on antipsychotic medications experienced significant improvements in depressed mood and reductions in extrapyramidal symptoms. Glycine is well tolerated.

OMEGA- 3 FATTY ACIDS AND PHOSPHATIDYL SERINE

Omega-3 Fatty Acids. Certain highly polyunsaturated fatty acids (N-3 PUFA)—EPA and DHA, the so-called "omega-3" fatty acids—may ameliorate symptoms of depressed mood and other mental health problems. Inconsistent results, small study size, and design flaws render the majority of research findings inconclusive. Disparate mechanisms of action have been proposed to explain the putative antidepressant efficacy of omega-3 fatty acids, including increased CNS serotonin activity, anti-inflammatory effects, suppression of phosphatidyl–inositol second messenger activity, and increased HRV. The suppression of proinflammatory cytokines by omega-3 fatty acids may help explain the reported relationship between coronary artery disease and major depressive disorder.

EPA may augment conventional antidepressants; however, DHA is probably no more effective against depressed mood than placebo. A meta-analysis of 10 placebo-controlled, double-blind studies on omega-3 fatty acids in depressed mood lasting 4 weeks or longer concluded that EPA significantly improved depressed mood in patients diagnosed with major depressive disorder or bipolar disorder. However, there was not a consistent relationship between EPA dose and antidepressant efficacy. The authors concluded that publication bias and different research design methods precluded validation of EPA as a stand-alone treatment of depression. Three small controlled trials evaluated omega-3s in combination with conventional antidepressants. In two studies, EPA (1 or 2 g per day) was added to conventional treatment. In the third study, patients were treated with a mixture of EPA and DHA (9.6 g per day) in addition to their conventional medication. In all three studies, treatment response was significantly greater in the combined omega-3–antidepressant groups than in the antidepressant-only groups. Many patients who had been refractory to conventional antidepressants improved markedly following augmentation with omega-3s. No adverse effects were reported in the combined treatment groups or the conventional treatment groups. These findings support that EPA augments conventional antidepressants; however, it is too early to conclude whether EPA is an effective monotherapy or that a specific ratio of EPA to DHA has optimal antidepressant efficacy. A meta-analysis of controlled trials on omega-3 fatty acids in bipolar disorder identified only one study in which omega-3s used adjunctively with a mood stabilizer showed a differential beneficial effect on depressive but not manic symptoms. In a 4-month controlled trial, 121 rapid-cycling or depressed patients with bipolar disorder treated with EPA (6 g per day) in combination with a conventional mood stabilizer versus placebo achieved equivalent outcomes.

EPA augmentation may be an effective preventive strategy in individuals at high risk for developing schizophrenia or in the early

prodromal phase of psychotic illness but not as a treatment of established cases of schizophrenia. The neuroprotective role of omega-3s during the early prodromal phase of schizophrenia may be related to general effects on the brain's antioxidative intracellular defense mechanisms or direct interactions between EPA and the glutamate receptor system. EPA augmentation may be especially appropriate for younger or antipsychotic naïve individuals in the early phase of illness or cases in which metabolic or sexual adverse effects may result in medication nonadherence and relapse. Preliminary findings suggest that a combination of EPA and DHA may reduce the rate of transitioning to full-blown schizophrenia in high-risk populations including individuals who chronically abuse substances or already have impairments in neuropsychological functioning.

In contrast to positive findings of studies on EPA augmentation in the prodromal phase of schizophrenia, a meta-analysis of placebo controlled studies on EPA augmentation on symptom severity in patients with chronic schizophrenia or other psychotic disorders found no evidence for greater benefit of EPA augmentation over antipsychotic medications alone. All studies included in the meta-analysis lasted 12 weeks and used EPA doses between 2 and 3 g per day. Possible explanations of reported negative findings include small study size, absence of clinical efficacy of EPA supplementation in chronic schizophrenia, a "ceiling effect" on response to EPA augmentation associated with concurrent treatment with antipsychotics, combinations of different fatty acids may be more effective than EPA alone, and dietary differences and other sociodemographic factors that may have confounded outcomes.

A systematic review and meta-analysis of studies on the effects of omega-3 supplementation on cognitive development, function, and decline throughout the life span concluded that omega-3 supplementation—especially DHA—may significantly improve cognitive development in infants but not in children, adolescents, adults, or the elderly. Omega-3 supplementation was not associated with a reduced rate of cognitive decline in healthy elderly adults or with a slowing in the rate of progression of Alzheimer disease.

A systematic review and meta-analysis of placebo-controlled trials on omega-3s in the treatment of ADHD in children and adolescents reported inconclusive findings due to methodological problems including small sample sizes, variability of selection criteria, and type and dosage of supplements and short-term follow-ups. Increased hippocampal neurogenesis soon following trauma may result in more rapid clearance of fear memories and interfere with consolidation of immediate posttrauma memories into long-term memories, reducing the risk of developing PTSD. Animal studies support that omega-3 fatty acids increase hippocampal neurogenesis. Studies done in Japan following the 2011 tsunami investigated the effectiveness of pretreatment with omega-3s in preventing

the development of PTSD following exposure to trauma in first medical responders. In a small, 8-week controlled trial (N = 30), women diagnosed with moderately severe borderline personality disorder randomized to ethyl-EPA (1 g per day) versus placebo reported less severe symptoms of aggression and depressed mood compared to the placebo group.

Omega-3 fatty acids are generally well tolerated and pose few safety issues. Transient mild gastrointestinal distress is sometimes reported by patients who take omega-3 fatty acids. There is one case report of apparent hypomania induced by omega-3 fatty acids. Rare cases of increased bleeding times have been reported in patients who take aspirin or anticoagulants concurrently.

Phosphatidyl Serine. Phosphatidyl serine is one of the most important phospholipids in the brain and is an essential component of nerve cell membranes. Findings of RCTs suggest that phosphatidyl serine (300 mg per day) may improve global functioning and memory in the early stages of Alzheimer disease. However, findings are inconsistent. Controlled trials on older adults suggest that phosphatidyl serine (300 mg per day) may result in small but consistent improvements in age-related cognitive decline.

NUTRIENT FORMULAS. Findings of two case series suggest that a proprietary nutrient formula (EMP+) containing 36 separate constituents, including chelated minerals, vitamins, essential fatty acids, amino acids, phytonutrients, and trace elements, may significantly reduce symptoms of mania, depressed mood, and psychosis in patients with bipolar disorder when taken alone or together with mood stabilizers. The putative mechanism of action involves correction of inborn metabolic errors that manifest as cyclic mood changes when the diet is deficient in certain micronutrients. In one series, 11 patients who completed a 6-month open study were able to reduce doses of conventional mood stabilizers by half while exhibiting sustained clinical improvement. In another series, 13 of 19 patients with bipolar disorder reportedly remained stable after discontinuing conventional mood stabilizers while taking the proprietary formula. Four patients stopped taking the nutrient formula because of gastrointestinal side effects including nausea and diarrhea. Three patients resumed conventional mood stabilizers because of recurring manic symptoms while taking the nutrient formula. Potentially toxic micronutrient–medication interactions have been reported in patients taking the nutrient formula with conventional mood stabilizers.

Preliminary findings suggest that taking the proprietary nutrient formula before exposure to trauma may increase emotional resilience and reduce the severity of PTSD symptoms following exposure. Adults enrolled in a study on the same proprietary formula for ADHD at the time of a major New Zealand earthquake reported feeling significantly less anxious and stressed than matched adults who were not taking the supplement. A

subsequent study following a severe aftershock several months after the earthquake compared measures of anxiety, mood, and posttraumatic symptoms (e.g., intrusive thoughts, avoidance, and hyperarousal) in individuals in the general population taking two doses (four capsules vs. eight capsules) of an identical formula. At the end of 4 weeks, individuals taking the formula reported significant decreases in stress, anxiety, avoidance, and arousal. There were nonsignificant differences between the high-dose and low-dose groups for all outcome measures, and all individuals in the treatment group reported clinically significant reductions in symptoms compared to the control group. Measured outcomes using micronutrient formulas were comparable to those observed with conventional medications, behavioral therapy using an earthquake simulator, and eye movement desensitization and reprocessing, but with fewer adverse effects and better retention rates. The significance of findings is limited by small study size and the absence of a placebo group, blinding, and randomized protocols.

BOTANICALS

St. John Wort (H. perforatum). St. John wort (Fig. 31.4–1) is widely used to treat and *self-treat* depressed mood. Constituents of St. John wort function as serotonin reuptake inhibitors, decrease binding to benzodiazepine receptors, weakly inhibit MAO, and possibly bind to NMDA receptors. Hyperforin probably contributes more to the antidepressant efficacy of St. John wort than hypericin. However, many commercial preparations are standardized to 0.3 percent hypericin. More than 40 controlled trials on St. John wort in patients with depressed mood have been conducted with mixed methodological quality and inconsistent results. A systematic review of 27 trials (total N = 2,200), 17 of which were placebo controlled, concluded that St. John wort is significantly more effective than placebo for the treatment of mild to moderate depressed mood. However, a more recent meta-analysis using stringent inclusion criteria controlling for publication bias and small study effects found significantly lower antidepressant efficacy than reported in previous meta-analyses. Therapeutic doses of standardized St. John wort preparations for moderate depression typically range from 300 mg twice daily to 900 mg twice daily. St. John wort is generally associated with few adverse effects, including nausea, restlessness, and mild sedation. There are no published reports of serotonin syndrome or other serious side effects when St. John wort is used as a monotherapy at recommended doses, but there are several case reports of *possible serotonin syndrome* when St. John wort and a conventional antidepressant were used together. Pregnant women should not take St. John wort. There are no case reports of serious interactions when St. John wort is taken with MAOI antidepressants. However, their combined use should be avoided to minimize risk. St. John wort should not be taken with

protease inhibitors, immunosuppressive agents, Coumadin, digoxin, and oral contraceptives.

FIGURE 31.4–1. St. John wort (*Hypericum perforatum*).

Ginkgo biloba. Ginkgo (*Ginkgo biloba*) extract is used in Europe and North America to treat dementia and other neurodegenerative diseases. Most commercially available preparations are standardized to two bioactive constituents: 24 percent flavone glycosides and 6 percent terpenoids. The flavonoid constituent functions as a strong antioxidant and is believed to have a general neuroprotective effect. The terpenoid fraction antagonizes platelet-activating factor (PAF), facilitating poststroke recovery through vascular reperfusion by decreasing thrombosis and nerve cell death associated with cerebral ischemia. Constituents of ginkgo may inhibit neurotoxicity and nerve cell death caused by nitric oxide. Two RCTs found that a compound herbal product containing both *G. biloba* (160 or 320 mg) and *Panax ginseng* (400 mg) resulted in transient improvement in recall performance in healthy adults. Enhanced cognitive functioning appears to reach its peak approximately 6 hours after the herbal preparation is taken. Preliminary findings suggest that combining ginkgo with another traditional Chinese herbal, *dangshen* (*Codonopsis pilosula*), may improve cognitive performance in healthy adults more than either herbal alone. A meta-analysis of 11 clinical trials on standardized ginkgo extracts in elderly cognitively impaired individuals who did not meet full criteria for dementia suggested consistent cognitive enhancing effects. Early systematic reviews

of RCTs on ginkgo in dementia concluded that ginkgo preparations in doses between 120 and 600 mg per day taken over several weeks to 1 year result in modest improvements in memory, general cognitive functioning, and activities of daily living in mild to moderate Alzheimer dementia and multi-infarct dementia. However, a systematic review of all RCTs published up to 2006 on ginkgo in the treatment of dementia and other types of acquired cognitive impairment found inconsistent evidence, noting that early trials were generally small or methodologically flawed, including the absence of standardized preparations and the use of different dementia rating scales across studies. In a 24-week, double-blind RCT, patients with mild to moderate dementia were randomized to a standardized ginkgo extract (160 mg per day) versus the cholinesterase inhibitor donepezil (5 mg per day) or placebo. At 24 weeks, patients taking ginkgo and donepezil experienced equivalent improvements in cognitive performance and global functioning. The findings of one study suggest that the rate of overall cognitive decline is moderately slowed in severe dementia. Ginkgo is generally well tolerated. Mild transient adverse effects include upset stomach, dizziness, and headaches. Cases of spontaneous bleeding have been reported. Ginkgo should be avoided in patients taking aspirin, warfarin, or heparin. *G. biloba* preparations should be discontinued at least 2 weeks prior to surgery. *G. biloba* may cause elevation of hepatic enzymes, and there are case reports of possible serotonin syndrome when taken with SSRIs.

Kava. In traditional Polynesian cultures, kava (*Piper methysticum*) (Fig. 31.4–2) is used for ceremonial purposes and as an inebriant. Preclinical studies suggest that α-pyrones interfere with serotonin neurotransmission in the amygdala. Kava also interferes with norepinephrine reuptake and is known to have binding affinity with both GABA and NMDA receptors. Kava may also reduce anxiety by directly modulating vagal heart tone. Double-blind studies and one meta-analysis support the use of kava preparations standardized to 70 percent kava lactones at doses between 70 and 240 mg per day for the treatment of "stress" and moderate anxiety but not severe anxiety or agitation. An early systematic review of seven quality studies (total N = 377) found that kava at 300 mg per day is more effective than placebo for reducing nonpsychotic anxiety states. A systematic review of 11 controlled, double-blind studies (total N = 600) concluded that kava was superior to placebo for the short-term management of generalized anxiety. Daily use of kava at doses of 100 to 200 mg may reduce anxiety symptoms associated with menopause. Findings of a small controlled trial suggest that generally anxious patients who gradually increase their daily dose of kava (up to 300 mg) while tapering off of a benzodiazepine do not experience worsening anxiety or benzodiazepine withdrawal. A multicenter RCT (N = 129) concluded that a standardized kava preparation (LI 150) was as

effective as two commonly prescribed antianxiety agents (buspirone and opipramol) in the treatment of generalized anxiety. Kava is generally well tolerated at therapeutic doses. However, some individuals report nausea, rash, headaches, or dizziness. Kava use should be avoided during pregnancy. Kava may potentiate the sedating effects of alcohol and benzodiazepines. Rare case reports of hepatitis and fulminant liver failure have led to restrictions in the sale of kava products in many European countries and a warning by the FDA. A thorough review concluded that most reported cases of liver failure were associated with a processing mistake resulting in potentially toxic levels of alkaloids in a single batch of kava.

FIGURE 31.4–2. Kava (*Piper methysticum*).

Valerian. Valerian (*Valeriana officinalis*) has been used in many traditional systems of medicine for more than a millennium. The mechanism of action may involve both inhibition of GABA reuptake and binding to serotonin receptors by valepotriates, important bioactive constituents of the herb. In a large, multicenter, 6-week study, chronic insomniacs randomized to a standardized valerian extract at 600 mg versus oxazepam at 10 mg reported comparable improvements in sleep quality and duration. Taking valerian extract following discontinuation of benzodiazepines may improve overall sleep quality. An early systematic review of RCTs on valerian for insomnia concluded that a standardized extract (600 and 900 mg at bedtime) improves the subjective quality of sleep with few side effects. However, a more recent systematic review of 29

RCTs that met rigorous criteria for study design and methodological quality concluded that the most rigorously designed studies failed to demonstrate significant outcomes differences between valerian and placebo. Another systematic review found a statistically significant benefit of valerian extract in insomnia but noted significant methodological flaws in most studies examined and evidence of publication bias. There is no risk of addiction with valerian preparations, and there are no reports of daytime psychomotor slowing. Because of its potentiating effect on GABA, valerian use should be avoided with benzodiazepines or alcohol. There are rare reports of elevated liver enzymes. Valepotriates may be mutagenic, and this herb should be avoided in pregnant or nursing women.

Ginseng. Individuals diagnosed with schizophrenia frequently show marked deficits in working memory, which can limit functional improvements with conventional drug treatment. In a 4 week placebo-controlled trial, 64 stable individuals diagnosed with schizophrenia were randomized to a proprietary North American ginseng extract versus placebo while remaining on treatment as usual. Visual working memory was significantly improved in patients receiving the ginseng preparation but not in the placebo group. Of note, individuals taking the ginseng preparation reported significantly fewer neurological adverse effects after 4 weeks than the placebo group. Further studies are needed to confirm these findings and determine optimal dosing of ginseng as an adjuvant to antipsychotics.

Huperzine-A. Huperzine-A is an alkaloid derived from the herb *Huperzia serrata* and is an ingredient of herbal formulas used in Chinese medicine to treat Alzheimer disease and age-related cognitive decline. Animal studies show that huperzine-A reversibly inhibits acetylcholinesterase, slows age-related neurotoxicity, regulates the secretion of nerve growth factor, and protects against oxidative stress and neuronal cell death associated with β-amyloid formation. Findings from preclinical trials suggest that huperzine-A may have a longer duration of action, better penetration of the blood–brain barrier, and be a more potent and more specific inhibitor of acetylcholinesterase than commercially available conventional cholinesterase inhibitors. Controlled trials report consistent beneficial effects in age-related memory loss, Alzheimer disease, and vascular dementia at doses between 200 and 400 μg per day. A meta-analysis of eight placebo-controlled RCTs (total $N = 733$) on huperzine-A in Alzheimer disease and two RCTs of huperzine A in vascular dementia reported improved global mental functioning as measured by changes in minimental status, improvements in activities, and infrequent mild adverse effects. The significance of findings is limited by small study size and poor methodological quality of many studies. Infrequent adverse effects include transient dizziness, nausea, and diarrhea.

Golden Root. In traditional Russian society, golden root (*Rhodiola rosea*) is prepared as a tea and consumed as a tonic for promoting general health and longevity. Among herbalists, golden root is regarded as an *adaptogen* because of its antifatigue and mood-enhancing effects. Animal studies support that beneficial effects on mood and cognitive performance are mediated by increased brain levels of dopamine, serotonin, and norepinephrine. Mental health benefits are probably related to increased dopamine, serotonin, and norepinephrine, and include improved memory, increased mental stamina, and a general calming effect. Early studies were done in the former Soviet Union and have not been carefully reviewed in English language journals. Findings of recent open studies suggest that 500 mg of a standardized *R. rosea* extract has antidepressant effects, improves overall mental performance and stamina, and may accelerate the recovery of normal cognitive functioning following traumatic brain injury. So far, only one RTC has examined putative antidepressant benefits of *R. rosea* with disappointing results. *R. rosea* can be safely combined with conventional antidepressants, and there is preliminary evidence of an augmentation effect. There are no cases of serious adverse effects or drug–drug interactions. However, there are reports of possible mania induction. There are no reports of toxicities or serious drug–drug interactions; however, there are case reports of possible mania induction, and patients diagnosed with bipolar disorder should be advised against using this herb.

Saffron (Crocus sativus). A systematic review of six studies on *C. sativus* in the treatment of depressed mood concluded that doses of 30 mg per day have antidepressant efficacy comparable to conventional antidepressants for mild to moderate depressed mood. Several mechanisms may be involved including serotonergic, antioxidant, anti-inflammatory, neuroendocrine, and neuroprotective effects. In a 6-week RCT on saffron, supplementation improved some sexual adverse effects associated with fluoxetine. Preliminary findings suggest that saffron may have comparable efficacy to cholinesterase inhibitors in Alzheimer patients.

Other Botanicals. Select herbals and other natural products may be safely used adjunctively to treat symptoms that frequently co-occur with depressed mood and other psychiatric disorders or mitigate adverse effects caused by conventional psychopharmaceuticals. *G. biloba* may be safely combined with antidepressants to mitigate adverse sexual effects. The extract of Hawthorn berry (*Crataegus* spp.) may reduce symptoms of moderate generalized anxiety. Passion flower (*Passiflora incarnata*) extract may have equivalent efficacy to benzodiazepines for the management of moderate anxiety symptoms. Lavender (*Lavandula angustifolia*) in the form of a tincture may have beneficial augmentation effects when taken with conventional antidepressants. Preparations of two botanicals used in Ayurvedic medicine, *Bacopa monnieri* and *Centella*

asiatica, have been used traditionally for millennia to treat generalized anxiety. Preliminary research findings suggest that both herbals are well tolerated and diminish symptoms of moderate anxiety. A comprehensive narrative review of studies on plant-derived anxiolytics found evidence of efficacy for chronic use of several herbals in the treatment of generalized anxiety including *Matricaria recutita, G. biloba, Scutellaria lateriflora, Silybum marianum, P. incarnata, Withania somniferum, Galphimia glauca, C. asiatica, R. rosea, Echinacea spp.*, *Melissa officinalis,* and *Echium amoenum.* Acute-onset anxiolytic activity was found for *C. asiatica, Salvia* spp., *M. officinalis, P. incarnata,* and *Citrus aurantium.* Preclinical animal studies reported promising early evidence for many other herbals in which anxiolytic effects are believed to be mediated by GABA or cannabinoid receptors.

COMPOUND HERBAL FORMULAS

Kami-untan-to. *Kami-untan-to* (KUT) is a compound herbal formula consisting of 13 different herbs and is used in Japanese traditional Kampo medicine to treat symptoms of cognitive impairment and dementia. Animal studies show that KUT increases brain levels of nerve growth factor (NGF) and choline acetyltransferase (ChAT), the enzyme required for the synthesis of acetylcholine. KUT protects against memory impairment induced by thiamin deficiency in mice, suggesting a possible beneficial role in the management of delirium tremens and other syndromes of cognitive impairment related to thiamin deficiency. In a 12-month open trial (N = 20), moderately demented patients with Alzheimer disease were treated with KUT alone versus a combined regimen of vitamin E, estrogen, and a nonsteroidal anti-inflammatory drug or placebo. Both treatment groups deteriorated at a significantly slower rate compared to the placebo group. Large, controlled trials are needed to confirm possible beneficial effects of this promising herbal formula in dementia.

Brahmyadiyoga. Brahmyadiyoga is a compound herbal formula used in Ayurvedic medicine for the treatment of syndromes resembling schizophrenia. The formula contains six herbs, including *Rauwolfia serpentina,* from which reserpine was isolated in the early 20th century and shown to have significant antipsychotic efficacy. Reserpine is seldom used to treat schizophrenia in Western countries today because of concerns over reports of severe depressed mood. In an early 2-month RCT, 108 chronically psychotic patients were randomized to 8 to 12 g per day of brahmyadiyoga, 8 to 12 g per day of valerian, 200 to 300 mg per day of chlorpromazine, and placebo. Patients treated with brahmyadiyoga and chlorpromazine reported significant and comparable improvements in symptom severity. A subsequent RCT confirmed the antipsychotic efficacy of brahmyadiyoga. Findings of a systematic review of three small RCTs

(total N = 250) comparing brahmyadiyoga with a conventional antipsychotic or placebo were inconclusive because of small study sizes, high attrition rates, methodological differences, and the use of nonstandardized assessment approaches. Cases of nausea have been reported with brahmyadiyoga.

Chinese Herbal Formulas. A review of well-designed RCTs and systematic reviews of studies on Chinese herbal medicines used to treat major depressive disorder identified several herbal formulas that were more effective than placebo and as effective as commonly used antidepressants for moderate depressed mood. Chinese herbal medicines were associated with fewer adverse effects compared to conventional antidepressants and reduced reports of adverse effects when used adjunctively with antidepressants. The significance of these findings is limited by the small number of studies included in the review.

ESSENTIAL OILS. Essential oil extracts of certain botanicals can be inhaled as aromatherapy or applied topically; however, research findings are inconsistent. Some studies report general calming of agitated behavior in demented persons when the essential oils of lemon balm and lavender are applied to the face and arms. Other studies report that exposure to the same essential oils is no more effective than placebo in agitated patients with dementia. It has been suggested that essential oils are effective in this population through systemic absorption or the combined effects of massage and olfaction but *not* through olfaction alone. Infrequent adverse effects of essential oils include skin allergies, phototoxic reactions, and potentiation of sedative–hypnotics when used with lavender or other oils known to have sedating effects. Pregnant women should avoid topical use of essential oil therapies because of possible toxic effects on the fetus or the uterus caused by systemic absorption.

HORMONES

Dehydroepiandrosterone and Estrogen. Dehydroepiandrosterone (DHEA) is a precursor of testosterone and other hormones. The sulfated form of DHEA —DHEA-S—is the most abundant steroid in the body. DHEA is an important neuroactive steroid and modulates neuronal excitability by acting as an antagonist at the GABA receptor complex. Preclinical studies suggest that putative mood-enhancing effects of DHEA are mediated through androgen receptors, estrogen receptors, serotonin, GABA, NMDA, and possibly norepinephrine. DHEA's antipsychotic effects may involve increased dopamine release in the frontal cortex and increased activity of NMDA and sigma receptors.

In a 6-week study (N = 22), depressed patients were randomized to DHEA in an escalating dose (30 mg per day for 2 weeks, followed by 30 mg

twice daily for 2 weeks and 30 mg thrice daily for two weeks) versus placebo. Half of the patients in the DHEA group improved by 50 percent or more on standardized rating scales. In a 6-week RCT (*N* = 46), moderately depressed adults off antidepressants were randomized to 90 mg per day of DHEA for 3 weeks, followed by 450 mg per day of DHEA (150 mg thrice daily) for 3 weeks versus placebo. The majority taking DHEA reported a 50 percent or greater reduction in depressive symptoms and improved sexual functioning. Most patients who responded to DHEA remained asymptomatic at 12 months' follow-up. More studies are needed to replicate these findings, evaluate DHEA for severe depressed mood, and clarify the mechanism for a putative synergistic or independent antidepressant effect.

In a small RCT (*N* = 30), inpatients with schizophrenia treated with 100 mg per day of DHEA in addition to their conventional antipsychotic medications experienced significant improvements in depressed mood, anxiety, and negative psychotic symptoms. Women improved more than men, and serum cortisol levels did not change during treatment.

DHEA is used in Europe and North America to self-treat declines in cognitive functioning associated with normal aging in spite of limited evidence. Pilot studies suggest that DHEA (25 to 50 mg per day) improves memory and enhances general cognitive functioning in healthy adults. However, findings are inconsistent, and negative findings have been reported at doses less than 90 mg per day. A systematic review of studies on DHEA or DHEA-S in healthy older adults found no evidence of improved memory or cognitive functioning. Large studies of at least 1 year or longer are needed to fully explore claims of long-term beneficial effects of DHEA. DHEA may differentially improve memory in elderly patients who have low DHEA serum levels more than in healthy adults with normal serum levels. However, no correlations were found between serum DHEA-S levels and severity of cognitive impairment or cumulative mortality in a large cohort of individuals newly diagnosed with Alzheimer disease. Preliminary findings suggest that 200 mg per day of DHEA may improve symptoms of cognitive impairment in multi-infarct dementia. However, no controlled trials on DHEA in Alzheimer disease have been completed. In a 6-month RCT (*N* = 47), men with mild dementia and healthy men aged 50 years and older were randomized to receive 75 mg of testosterone (in the form of a dermal gel) versus placebo together with their usual medications. Quality-of-life measures improved in men with mild dementia and healthy men taking testosterone, and fewer men with mild dementia who received testosterone experienced declines in overall functioning and visuospatial abilities.

In a 6-week RCT (*N* = 30), inpatients with schizophrenia randomized to DHEA at 100 mg per day in addition to their regular antipsychotic medications experienced significant improvements in *negative* psychotic

symptoms, including reduced apathy and social withdrawal. However, there were no significant changes in *positive* psychotic symptoms, including auditory hallucinations and delusions. Findings of another small RCT (*N* = 30) suggested that augmentation of conventional antipsychotics with DHEA (100 mg day for 6 weeks) significantly reduces negative symptoms and may be especially effective in women. However, a systematic review of three small RCTs (*N* = 126) on DHEA or testosterone as adjunctive therapies to conventional antipsychotics found equivocal evidence for an augmentation effect on measures of positive and negative symptoms, global functioning, and quality of life. Estrogen may play a "protective" role in women with schizophrenia because serum estrogen levels influence brain dopamine activity.

Preliminary findings suggest that DHEA may be efficacious against PTSD. Increased severity of PTSD is correlated with reduced DHEA blood levels. In a small open-label study (*N* = 5), women with treatment-refractory PTSD related to early childhood abuse and who had not responded to conventional pharmacologic therapy experienced decreases in numbing, reexperiencing, hyperarousal and other core symptoms, improved sleep, and improved libido with DHEA at doses between 25 and 100 mg per day. DHEA should be avoided in men at risk of prostate cancer and women who have a history of estrogen receptor-positive breast cancer.

In a dramatic case report, acute psychosis associated with chronically low estrogen levels completely remitted with hormonal replacement therapy using estrogen and progesterone. Findings of a pilot study suggested that estrogen replacement may be an effective treatment of postpartum psychosis. Estrogen may augment the efficacy of conventional antipsychotics.

DHEA is generally well tolerated at therapeutic doses, but cases of acne, hirsutism, and rare cases of mania have been reported with chronic high doses. DHEA supplementation may promote cancer in women who have a history of estrogen receptor–positive breast cancer and may increase the risk of prostate cancer in men with early or undetected prostate cancer. Women should be cautioned about an increased risk of heart disease and breast cancer.

Melatonin. A meta-analysis of 17 RCTs (total *N* = 284) concluded that exogenous melatonin decreases the time required to fall asleep and increases total sleep duration. Double-blind studies demonstrate consistent improvements in sleep quality and duration and reduced sleep latency in chronic insomniacs who take melatonin at doses between 0.3 and 3 mg. Doses that replace normal physiological levels (0.3 mg) in insomnia patients with low serum melatonin levels improve restful sleep in the middle part of the sleep cycle. Higher pharmacological doses (3.0 mg) also improve sleep quality but may cause hypothermia, elevated daytime

melatonin levels, and residual grogginess. A systematic review of 10 RCTs concluded that melatonin in jet lag was effective at doses of 0.5 to 5.0 mg when taken close to bedtime in the country of destination. Overall efficacy was relatively less for westward flights and relatively greater when more time zones were crossed. Melatonin has been used with moderate success to treat insomnia in patients with schizophrenia and agitated patients with mania. Daytime use of melatonin should be avoided because of potentially impairing declines in cognition and job performance. Uncommon mild adverse effects include headache and itching.

Vitamin D. Vitamin D comprises a family of prohormones manufactured in the skin in response to ultraviolet energy from sunlight and can be taken as a supplement. Vitamin D is centrally involved in the regulation of calcium and phosphorous metabolism and immunomodulation. Vitamin D_3 (cholecalciferol) deficiency is common in Western populations—especially among the elderly—and may be associated with an increased risk of depression in susceptible populations. Preliminary findings suggest that daily supplementation with vitamin D_3 at doses of 400 to 800 IU may reduce the severity of depression in patients diagnosed with seasonal affective disorder when serum D_3 levels are raised to 35 ng/mL or greater. A systematic review of RCTs on vitamin D supplementation in depressed mood found that doses greater than 800 IU per day had comparable efficacy to conventional antidepressants for moderate depressed mood. Daily supplementation with high doses of vitamin D up to 100,000 IU per day may be more effective than bright light therapy in the treatment of seasonal depressed mood.

SOMATIC THERAPIES

Exercise. Epidemiological studies support that infrequent physical activity is a risk factor for developing depressed mood. Physical exercise increases the levels of brain-derived neurotrophic factor and may enhance neural plasticity and new synapse formation. Regular exercise is associated with increases in the relative size of the frontotemporal and parietal lobes, which are important centers for learning, memory, and executive functioning. Brain levels of endorphins, dopamine, norepinephrine, and serotonin are increased following sustained exercise. Two meta-analyses of controlled trials confirmed consistent positive effects of regular exercise on depressed mood. Aerobic conditioning and strengthening exercise are equally effective against depressed mood. The optimum duration and frequency of exercise for depression have not been determined but are probably related to age and fitness level. In a 16-week study (N = 156), depressed patients older than 50 years randomized to aerobic exercise three times a week versus sertraline (up to 200 mg per day) versus exercise plus sertraline experienced equivalent improvements in standardized measures of mood,

self-esteem, and negative thoughts. Patients in the antidepressant-only group initially improved faster. However, patients in the exercise-only group had a lower 6-month relapse rate. The antidepressant effects of running or fast walking may be equivalent to those of CBT and conventional antidepressants for moderate depressed mood. Depressed individuals who exercise in a brightly lit (2,500 to 4,000 lux) indoor environment reported greater improvements in mood and vitality than depressed individuals who exercised indoors in ordinary room light (400 to 600 lux).

Routine physical activity is associated with reduced risk of all categories of dementia. More than 2,000 physically nonimpaired men aged 71 to 93 years were followed with routine neurological assessments at 2-year intervals starting in 1991. At the end of the study period, men who walked less than 1/4 mile daily had an almost twofold greater probability of being diagnosed with any category of dementia than men who walked at least 2 miles daily. Findings of the Nurses' Health Study based on biannual mailed surveys over 10 years showed that women aged 70 to 81 years who engaged in regular vigorous physical activity were significantly less likely to have been diagnosed with dementia than women with more sedentary lifestyles. A small RCT found that regular daily exercise in individuals with moderate dementia receiving in-home care reduced depressed mood but did not improve cognitive functioning, suggesting that regular exercise does not slow the rate of cognitive decline once dementia has begun.

A period of 20 to 30 minutes of regular daily exercise significantly reduced symptoms of generalized anxiety. A prospective 10-week study on exercise in individuals diagnosed with panic disorder found that regular walking or jogging (4 miles three times a week) reduced the severity and frequency of panic attacks. In a small open study of alcoholics who exercised regularly while hospitalized for acute detoxification reported significant improvements in mood and general well-being. Abstinent alcoholics enrolled in outpatient recovery programs reported improved mood with regular strength training or aerobic exercise. A systematic review of studies on the relationship between exercise and sleep in the elderly concluded that exercise probably enhanced sleep quality and improved overall quality of life.

Expressive Movement Therapies. Dance therapy and other expressive movement therapies are widely used in both outpatient and inpatient settings as treatments of depression, anxiety disorders, and schizophrenia. Most studies on expressive movement therapies are qualitative in design and difficult to analyze because of methodological differences, small study sizes, and the absence of blinding. However, evidence suggests that the regular practice of dance therapy and other expressive movement therapies is associated with significant reductions in the severity of anxiety and

depressed mood.

Animal-Facilitated Psychotherapy. Animal-facilitated therapy provides a sense of social connection and renewed purpose and has beneficial effects on mood, anxiety, and global cognitive functioning in both outpatient and inpatient settings. Animal-facilitated psychotherapy has not been rigorously evaluated in controlled trials; however, observational studies suggest that regular contact with therapy dogs and other animals may reduce anxiety, improve mood, and increase attention span. Preliminary findings suggest that the use of animal companions in the context of psychotherapy probably improves general emotional well-being in depressed and anxious persons.

Massage. The anxiety-reducing effects of massage are probably mediated by decreased cortisol and increased parasympathetic tone. However, the psychological benefits of massage are difficult to quantify. Findings of open trials suggest that regular massage therapy reduces the severity of chronic moderate generalized anxiety, especially when anxiety is related to test taking, problem solving, or work stress. Massage may also reduce the severity of depressed mood. A review of massage techniques for depressed mood identified several open studies with positive outcomes.

Snoezelen. Snoezelen is a multimodal treatment of dementia that employs music, light patterns, tactile surfaces, and essential oils to stimulate the senses of sight, hearing, touch, taste, and smell. Snoezelen sessions are usually administered two to three times weekly to patients with dementia in skilled nursing homes. A systematic review of RCTs of snoezelen in dementia concluded that apathy, speech skills, and psychomotor agitation improved in demented patients after four to eight sessions.

MINDFULNESS TRAINING AND MIND–BODY PRACTICES. Mindfulness training includes many styles of meditation and guided imagery. Mind–body practices include *taijijuan, qigong,* yoga, and others. Improved capacity for focused attention and reflection has been proposed as an important nonspecific psychological benefit of mind–body practices. MBSR and mindfulness-based cognitive therapy (MBCT) are widely secular approaches widely used to treat mental health problems. Research findings support that MBCT is a beneficial adjunctive therapy for unipolar depression and both MBSR and MBCT are beneficial adjunctive interventions for anxiety disorders. The regular practice of a mind–body skill may be as effective as CBT or conventional antidepressants in patients with moderately depressed mood. Mindfulness training combined with cognitive therapy may enhance outcomes compared to cognitive therapy alone. In an open trial, 60 women with postpartum depressed mood and anxiety experienced significant relief during the first 4 weeks following childbirth using a combined relaxation-guided imagery protocol.

Depressed patients who had previously been diagnosed with somatoform disorders experienced greater improvements in mood when guided imagery was combined with antidepressants than with antidepressants alone. Findings of a small open pilot study suggest that regular use of MBCT may significantly improve executive functioning, memory, and ability to complete tasks in stable patients diagnosed with bipolar disorder. Improvements in executive functioning persisted up to 3 months after treatment ended. Twelve individuals diagnosed with bipolar disorder who participated in 12 weekly group sessions of MBCT exhibited increased mindfulness, lower residual depressive mood symptoms, increased capacity for emotional regulation abilities, and generally improved psychological well-being and psychosocial functioning. Large prospective trials are needed to confirm these findings and determine the optimal type and frequency of MBCT for cognitive symptoms of bipolar disorder.

Applied Relaxation. Relaxation techniques include sustained deep breathing, progressive muscle relaxation, guided imagery, and systematic desensitization. The mechanism of action may involve reducing muscle tension using behaviors or cognitions that diminish tension and autonomic arousal. The regular practice of progressive muscle relaxation or sustained deep breathing at bedtime reduces the time needed to fall asleep in chronic insomniacs. Individuals who have difficulty falling asleep because of somatic complaints respond better to progressive muscle relaxation whereas those who stay awake because of chronic worrying are more likely to sleep better using guided imagery. A meta-analysis of 66 studies (total N = 2,000) on somatic and mind–body treatments of insomnia concluded that all interventions resulted in "reliable and durable benefits" that were superior to placebo, including improved quality of sleep and reduced sleep latency. A separate meta-analysis concluded that nonpharmacological treatments of chronic insomnia are initially more expensive and require more time than conventional drug therapy but are reliable, yield sustained benefits, and are thus more cost-effective in the long run.

In a 5-month prospective study, individuals diagnosed with generalized anxiety disorder randomized to relaxation only versus conventional antidepressants plus relaxation experienced equivalent and significant improvements in state anxiety levels by the end of the trial. In a small RCT (N = 36), anxious adult outpatients randomized to 12 weekly sessions of applied relaxation versus conventional cognitive therapy experienced significant and comparable reductions in anxiety. Combining relaxation with guided imagery may be more effective than either approach alone. In an open trial, 60 women reporting anxiety and postpartum depressed mood experienced significant reductions in both anxiety and depressed mood using a combined relaxation-guided imagery protocol during the first 4 weeks after childbirth. Guided imagery has beneficial effects on the

immune system, physiological stress responses, and mental–emotional functioning in general. Consistent daily practice of guided imagery reduces the severity of many anxiety symptoms, including generalized anxiety, panic, and traumatic memories.

Yoga. More studies have been done on yoga than any other mind–body discipline. The calming effects of yogic breathing are probably related to modulation of parasympathetic–sympathetic tone. The regular practice of yoga postures (asanas) probably results in beneficial changes in autonomic activity, resulting in improved cardiorespiratory performance and increased feelings of general well-being. A review of recent studies on yoga in depressed persons reported significant positive findings but most studies were limited by methodological flaws. Combining a mindfulness practice with yoga may have greater mood-enhancing benefits than forms of yoga that do not include mindfulness. Moderately depressed patients improved significantly by the end of a 5-week class in Sudarshan kriya (SK) yoga. Improvements associated with SK breathing practices in severely depressed hospitalized patients were comparable to responses achieved with antidepressants and only slightly less robust than patients receiving electroconvulsive therapy. Preliminary findings support that the regular practice of forms of yoga that do not include specialized breathing exercises also improve depressed mood, including hatha yoga, Omkar meditation, and Iyengar yoga. Regular SK yoga practice may reduce the need for conventional anxiolytics in individuals diagnosed with generalized anxiety disorder. Findings of a small RCT suggested that the regular practice of a specific kundalini yoga left-nostril breathing technique may significantly reduce symptom severity in patients diagnosed with OCD. A literature review of studies on yoga and other mind–body practices used to treat PTSD found that many approaches in current use reduce some PTSD symptoms including intrusive memories, avoidance, and emotional arousal.

Meditation. Meditation has been extensively investigated as a treatment of anxiety. Beneficial physiological effects of meditation include decreased oxygen consumption, respiratory rate, and blood pressure, as well as EEG changes associated with decreased autonomic arousal. MBSR is an eclectic approach that incorporates elements of different Eastern meditation practices and Western psychology. Findings of RCTs show that the consistent practice of mindfulness meditation significantly reduces generalized anxiety. In one series, the majority of anxious individuals (N = 322) who successfully completed a 10-week MBSR program reported significantly decreased physical and emotional distress, improved quality of life, an increased sense of general well-being, greater optimism, and increased feelings of control. Patients diagnosed with irritable bowel syndrome (IBS), a frequent concomitant of generalized anxiety,

experienced significantly fewer symptoms of both IBS and anxiety when practicing two brief (15 minutes) daily sessions of mindfulness meditation.

Studies have evaluated mindfulness training, mantra reciting, and compassionate meditation (i.e., "Vipassana") for their potential beneficial effects in PTSD. A review of meditation practices aimed at preventing PTSD found more evidence supporting mindfulness meditation than mantra reciting or compassionate meditation. Preliminary findings suggest that symptoms of reexperiencing and psychic numbing may be less responsive to meditation than other PTSD symptoms. In a 12-week pilot study (N = 5), veterans from the middle east conflict diagnosed with PTSD who trained in transcendental meditation experienced significant improvements in overall quality of life and reductions in core PTSD symptoms. Further studies on the therapeutic effects of meditation in PTSD are needed to better characterize the meditation techniques being studied using validated instruments that reliably measure targeted outcomes.

THERAPIES BASED ON SCIENTIFICALLY VALIDATED FORMS OF ENERGY AND INFORMATION.

CAM approaches based on validated forms of energy or information include biofeedback, sound, full-spectrum bright light exposure, dim blue light, high-density negative ions, virtual reality graded exposure therapy (VRGET), brain–computer interface technologies, and microcurrent electrical stimulation. Research findings point to both direct and indirect effects of well-understood forms of energy and information on the psychological and neurophysiological causes of mental illness. In contrast to scientifically *validated* modalities, treatments based on postulated *subtle* forms of energy or information cannot be verified by *current* science.

Music and Sound. In a small RCT (N = 40), anxious adult patients were randomly assigned to conventional cognitive therapy versus music-assisted reframing. Patients in the music group experienced greater reductions in overall anxiety. When headphones are used to route sounds of slightly different frequencies to the right and left hemispheres, the brain stem "constructs" binaural beats based on the frequency difference between sounds processed in each hemisphere. Preliminary findings support that certain binaural beats (16 to 24 Hz) may improve depressed mood, enhance attention, and improve short-term and immediate recall in healthy adults, whereas binaural beats at lower frequencies may consistently induce a calm, relaxed state. A specific antidepressant effect of music has not been established. However, frequent listening to music probably has beneficial effects on endorphins and other neurotransmitters that mediate improvements in depressed mood.

Findings of RCTs suggest that music alone or in combination with guided imagery improves mood in depressed patients with cancer. Patients with cancer who participated in weekly sessions of combined music and guided imagery reported greater improvements in depressed mood and

overall quality of life than patients in a wait-list group. Depressed outpatients reported significant improvements in mood and beneficial changes in heart rate and blood pressure soon after listening to tranquil music. A meta-analysis of studies on music therapy in individuals with dementia concluded that regular singing, engaging in dance therapy, listening to music, and participating in musical games improve cognitive and behavioral functioning in individuals with severe dementia, who experience reduced agitation, reduced wandering, enhanced social interaction, improved mood, reduced irritability or fear, and increased cooperative behavior. Calming background music significantly reduces irritable behavior, anxiety, and depressed mood in nursing home patients with dementia. Regular music therapy reduces irritability and improves expressive language in demented individuals.

Light. Light of different intensities and colors is used in conventional biomedicine and many nonconventional systems of medicine to treat both medical and psychiatric disorders. Diverse mechanisms of action are probably involved, including regulation of melatonin and neurotransmitters, especially the monoamines. Entrainment of sleep–wake cycles by external bright light cues and the associated suppression of melatonin production by the pineal gland is the established mechanism of action underlying the therapeutic benefits of light exposure on sleep and daytime fatigue. Exposure to natural sunlight, especially in the early morning, has a significant antidepressant effect and may reduce the length of hospital stays in severely depressed inpatients. A systematic review of controlled studies confirmed an antidepressant effect of bright light (10,000 lux) exposure therapy in seasonal depressed mood but provided only limited evidence supporting bright light as a treatment of nonseasonal unipolar depressed mood. A more recent meta-analysis of controlled studies concluded that bright light exposure or dawn simulation for *seasonal* depressed mood and bright light exposure (but not dawn simulation) for *non*seasonal depression had efficacy comparable to conventional antidepressants. Findings of a small RCT suggested that bright light therapy may be an effective alternative to conventional antidepressants in pregnant depressed women. Recent studies suggested that regular exposure to dim red or blue light might be as efficacious as bright light, especially in the management of seasonal depressed mood. In a 4-week single-blind study (N = 57), patients diagnosed with seasonal affective disorder were randomized to daily bright light versus dim red light exposure. Both groups experienced an equivalent and significant (40 percent) reduction in symptoms. Preliminary findings suggest that antidepressant effects of early-morning exposure to narrow spectrum blue or green light may be equivalent to benefits obtained from full-spectrum bright light exposure.

Exposing depressed patients to low-intensity artificial light approximately 2 hours before they would naturally be exposed to early-morning daylight increased the speed of response to conventional antidepressants. Thirty depressed inpatients treated with citalopram at 40 mg per day were randomized to early-morning dim green light, 400 lux, versus a sham, non–light-emitting device during the first 2 weeks of drug treatment. Patients in the combined citalopram–light exposure group reported significantly greater and more rapid improvements in mood compared to the citalopram-only and placebo groups.

No adverse effects have been reported with regular exposure to dim light. Findings from a double-blind study suggested that naturalistic dawn simulation (i.e., exposure to dim blue light before waking), postwaking bright light exposure, and high-density negative ions may be equally effective treatments of seasonal depressed mood. An expert consensus report on light treatment for sleep disorders concluded that there is compelling evidence for the efficacy of bright light exposure in the management of circadian rhythm sleep problems but not for other types of insomnia. Most protocols recommend 30 to 40 minutes of bright light exposure daily for beneficial effects in shifting circadian rhythms and changing sleep–wake cycles. Beneficial effects of bright light exposure therapy are usually noticeable within a few days, and an appropriate treatment plan often resynchronizes the patient's sleep–wake cycle with his or her time zone in 2 to 3 weeks. Correct timing of light exposure is essential for successful treatment of circadian sleep problems, and bright light exposure between the hours of late morning and late afternoon probably has no effect on circadian rhythm phase. Preliminary findings suggest that melatonin used in conjunction with bright light exposure may achieve more rapid normalization of the sleep–wake cycle. Patients should be cautioned to avoid using sun lamps (which emit harmful ultraviolet rays) and instructed to keep the eyes open while viewing the light fixture at an oblique (approximately 45 degrees) angle to avoid excessive eye strain.

Some patients exposed to bright morning light at 10,000 lux on a regular basis reported transient side effects, including mild jitteriness or headaches (10 percent) and mild nausea (16 percent). Sporadic cases of hypomania have been reported, especially in winter depressives or bipolar patients exposed to early-morning bright light. Almost two-thirds of patients who use bright light exposure therapy in the evening report insomnia.

High-Density Negative Ions. Regular daily exposure to high-density negative ions may be an effective treatment of seasonal depressed mood and may have comparable efficacy to bright light exposure for this condition. Twenty-five depressed patients with seasonal depressed mood were randomized to high-density negative ions (2.7×10^6 ions/cm^3) versus low-

density negative ions (1×10^4 ions/cm^3) using in-home ion generators 30 minutes daily for 3 weeks. Fifty-eight percent of patients exposed to high-density negative ions experienced significant improvements in mood on standardized rating scales, compared to 15 percent of patients exposed to low-density negative ions. In a 2-week RCT (N = 158), patients with seasonal depressed mood were randomly assigned to bright light exposure (10,000 lux) versus high-density or low-density negative ions. Patients exposed to high-density negative ions or bright light experienced significant and equivalent improvements in mood. In an RCT (N = 128), bright light exposure, a pleasing auditory stimulus, and high-density negative ions resulted in equivalent and rapid improvements in mood and alertness in mildly depressed and nondepressed adults. A review of studies on the effects of air ionization on mental health found evidence for a consistent antidepressant effect of high-density negative ions (but not low density ions) on unipolar depressed mood and seasonal depressed mood and no consistent effects on anxiety, mood, relaxation, sleep, and measures of well-being. Adverse effects have not been reported with regular daily exposure to negative ions following the protocols used in published studies.

Biofeedback Using EEG and HRV. Electroencephalographic biofeedback and HRV biofeedback employ information extracted from the body's intrinsic electromagnetic rhythms as sources of feedback. EEG biofeedback (also called *neurofeedback*) translates electrical brain activity into a visual or auditory signal with the goal of modulating the *general arousal level* of the brain or correcting *dysrhythmias* in specific brain regions or frequency domains to achieve therapeutic changes in psychological and psychosomatic symptoms associated with depressed mood, bipolar disorder, generalized anxiety, PTSD, social anxiety disorder, attention-deficit disorder, and other psychiatric disorders.

Double-blind studies have evaluated the efficacy of EEG biofeedback (also called *neurofeedback*) as a treatment of anxiety, depressed mood, attention-deficit disorder, and PTSD. Sustained "shifts" in the dynamic equilibrium of electrical brain activity have been proposed as a general mechanism underlying therapeutic effects of EEG biofeedback. Case reports and controlled studies show reductions in the severity of depressed mood using photic stimulation or alpha-asymmetry EEG biofeedback protocols. Beta-frequency training in EEG biofeedback may result in improved performance on standardized intelligence tests in nonimpaired individuals. EEG biofeedback has been successfully used to treat depressed mood in chronic alcoholics and to reduce relapse rates; however, large double-blind studies are needed to confirm consistent and robust benefits of EEG biofeedback in depressed mood. Regular EEG biofeedback training and conventional antianxiety medications may have comparable efficacy for the management of generalized anxiety.

Recent research findings support that EEG biofeedback involving very low frequencies, between 0.02 and 0.2 Hz, results in rapid significant reductions in the severity of PTSD symptoms and improvements in global cognitive functioning. It has been suggested that *infra-low* frequencies (ILF) may induce therapeutic shifts in the functional connectivity of the brain's resting state networks resulting in reduced overall arousal, enhanced cognitive functioning and increased emotional resilience. In a pilot study, seven Vietnam era war veterans with chronic treatment-refractory PTSD who trained with an ILF neurofeedback protocol reported significant reductions in symptom severity after 20 sessions; however, wait-listed controls subsequently reported similar improvements. At one large military base, more than 500 active duty combatants diagnosed with PTSD were trained in ILF EEG biofeedback. Training was done 1 to 3 times per week and was administered by licensed psychotherapists certified in EEG biofeedback therapy. Symptom severity was evaluated weekly using the PCL-5—the military version of the PTSD Checklist. A cohort analysis of 300 of the 500 active duty Marines in the original group found that 75 percent of persons with moderate to severe PTSD symptoms experienced significant clinical improvement including reductions in psychological, cognitive, psychophysiological, and physiological symptoms. Twenty-five percent of subjects in the cohort reported that all symptoms had resolved completely with fewer than 20 EEG biofeedback sessions; another 50 percent experienced significant reductions in symptom severity after 40 sessions. The remaining subjects took much longer to respond to treatment, continued to report clinically significant symptoms, discontinued training prematurely, or were nonresponsive to the neurofeedback training protocol.

In a review of nonpharmacologic treatments of insomnia, EEG biofeedback was found to be more effective than progressive muscle relaxation. The American Academy of Sleep Medicine recommended EEG biofeedback for the treatment of chronic insomnia. Several EEG biofeedback protocols are beneficial in insomnia, but a protocol that modulates alpha and theta activity along the primary sensory and motor regions in the frontal and prefrontal areas of the cortex may yield the most consistently beneficial outcomes in improving overall quality of sleep and reducing the frequency of middle awakenings.

Biofeedback using information obtained from measures of HRV may have therapeutic benefits in anxiety, depressed mood, and PTSD. HRV biofeedback is based on the theory that positive emotional states result in increased *coherence* of several physiological aspects of heart rhythm, and that feed-forward and feedback loops between the heart, brain stem, and neocortex result in complex relationships among specific emotional or mental states and specific patterns of HRV. Abnormal low HRV is associated with deficits in attention and short-term memory in combat

veterans diagnosed with PTSD. Case reports support that HRV biofeedback may improve depressed mood and anxiety resulting from chronic stress and significantly reduces arousal and improves general emotional well-being in individuals who are subjected to acute job-related stress. Beneficial changes in baseline anxiety achieved with HRV biofeedback are associated with decreased serum cortisol levels and increased serum DHEA levels. In a 4-month RCT, the majority of 29 police officers trained in HRV biofeedback reported significant improvements in baseline anxiety compared to 36 officers assigned to a wait-list group who did not improve. Individuals diagnosed with generalized anxiety disorder reported significant improvements in general emotional well-being and reduced baseline anxiety with regular HRV biofeedback training. Preliminary findings support that increased cardiac coherence may lessen the severity of PTSD; however, results are inconsistent. In a small 3-week open exploratory study, a group of active duty service members diagnosed with PTSD or depressed mood who received HRV biofeedback plus treatment as usual did not report greater reductions in symptom severity compared to a group receiving treatment as usual only. Correlations between electrocardiogram (EKG) and EEG activity provide evidence for complex feed-forward and feedback loops between the heart and brain that regulate emotion. This integrative biofeedback approach is called "energy cardiology" and is only beginning to be examined in controlled trials.

Mary could not recall a time when she was able to focus or sit still in class. She had been maintained on Ritalin starting at the age of 10 years and had responded immediately. There had been no side effects, and by the age of 16 years, Mary was an honors student and felt like she was in control of her life. One day while watching the evening news, she was intrigued by a report about EEG biofeedback. A psychiatrist in the community was using a special machine to "train" brain waves and lessen symptoms of ADHD, anxiety, and other mental health problems. Mary decided to find out how this approach worked and made an appointment. The psychiatrist explained that EEG biofeedback training is like "playing a computer game" using only your brain. Mary had her first 30-minute session the same day. The game was fun, and she was surprised at how clear and focused she felt during and after the session. After 10 sessions, the psychiatrist suggested lowering Mary's Ritalin dose from 20 mg twice daily to 15 mg twice daily. With her parents' consent, Mary tried the lower dose, and her ability to focus did not diminish. After 20 sessions, Mary was able to lower her dose to 10 mg twice daily and maintain good concentration in all of her classes. She was less fatigued in the evenings and her appetite improved. On her psychiatrist's advice, Mary purchased software for daily brain training at home and continued to take Ritalin at the lower dose.

Virtual Reality Graded Exposure Therapy. VRGET is a rapidly evolving therapy with important implications for the management of PTSD and severe phobias that are difficult to treat using conventional therapies. VRGET combines advanced computer graphics, three-dimensional visual displays, and body-tracking technologies to create realistic virtual environments with the goal of simulating feared situations or objects. Virtual environments have been designed to provide visual, auditory, tactile, vibratory, vestibular, and

olfactory stimuli to patients in highly controlled settings. During a virtual exposure session, the therapist closely tracks the patient's state of arousal by monitoring physiological indicators of stress, including heart rate and respirations. Patients are ready to practice in vivo exposure to the feared object or situation after they have been desensitized to a virtual environment.

VRGET is more effective than conventional imaginal exposure therapy and has comparable efficacy to in vivo exposure therapy for the treatment of specific phobias, agoraphobia, panic disorder, and PTSD. Like in vivo and imaginal exposure therapy, VRGET desensitizes the patient to a situation or object that would normally cause anxiety or panic. In an RCT, VRGET and conventional CBT were equally effective treatments of panic disorder with agoraphobia, and patients who underwent VRGET required 33 percent fewer sessions to achieve similar results. RCTs confirm that VRGET is an effective treatment of fear of flying, fear of heights, fear of small animals, fear of driving, and other phobias. VRGET is as effective as and more cost-effective than conventional exposure therapy for fear of flying because patient and therapist avoid significant time commitments and the need to use airplanes.

A virtual environment simulating the devastation of the September 11, 2001, attacks on the World Trade Towers has been successfully used to treat individuals diagnosed with severe PTSD following the attacks. Findings of a study on combined multisensory exposure and VRGET reported significant reductions in severity of PTSD symptoms in active duty combatants who had failed to respond to other forms of exposure therapy. Several subjects reported significant improvement following only five VRGET sessions; however, there was considerable variability in the number of VRGET sessions needed to reduce symptom severity to the same level. Combining VRGET with D-cycloserine, a partial NMDA agonist, may yield differentially greater improvement in symptoms of acrophobia compared to VRGET alone.

Considerable research is ongoing to develop virtual reality tools for assessing and preventing combat-related PTSD. STRIVE (Stress resilience in virtual environments) is a highly integrative "stress resilience training" aimed at enhancing emotional coping strategies prior to active deployment. STRIVE employs an immersive VR environment to simulate combat situations that includes a "virtual mentor" who guides the combatant through a virtual experience while coaching him or her in relaxation and emotion self-regulation skills. The intensity of the virtual stimulus used is determined by the individual's habituation based on HRV and other measures of autonomic arousal. Physiological biomarkers of stress response are measured before and after VRGET sessions. The STRIVE system permits users to be immersed in stressful combat scenarios and interact with virtual characters for training in a variety of coping strategies

that may enhance resilience in the face of extreme stress. The STRIVE protocol may provide a useful tool for predicting the risk of developing PTSD or other psychiatric disorders in new recruits prior to actual combat exposure. Recruits who display high resilience and thus presumably at relatively lower risk of developing PTSD might be more suitable for direct combat roles while individuals who display low resilience might preferentially be assigned to noncombat roles.

Future integrative approaches to phobias, panic attacks, and other severe anxiety disorders will combine VRGET with biofeedback in outpatient settings or in the patient's home via broadband Internet connections, with CBT, relaxation, mind–body practices, and appropriate medications. Combining VR environments with real-time feedback based on neurophysiological responses to stress may permit each unique patient to optimize the level and type of VR exposure to enhance resiliency training and speed the rate of recovery from PTSD. Human–computer interface (HCI) systems based on CBT and biofeedback are being developed for resilience training in individuals at risk of developing PTSD following exposure to trauma. Larger studies on patient populations diagnosed with PTSD using head-mounted displays and other technologies that create more immersive virtual environments are needed to determine whether combining VRET and EEG biofeedback is practical in clinical settings and yields superior outcomes compared to either approach alone.

Infrequent cases of disorientation, nausea, dizziness, headache, and blurred vision have been reported following VRGET. Intense sensory stimulation during VRGET can trigger migraine headaches, seizures, or gait abnormalities in individuals who have these disorders. Patients with schizophrenia should not use VRGET because immersion in a virtual environment can exacerbate delusions.

Cranioelectrotherapy Stimulation. Cranioelectrotherapy stimulation (CES), also called microcurrent electrical stimulation, is based on the application of extremely weak electrical current to the head and neck to treat anxiety, insomnia, and substance abuse. The mechanism of action may involve stimulation of endorphin release and induction of EEG frequency changes that reduce the severity of anxiety. A meta-analysis of controlled trials comparing CES with a sham treatment found that generalized anxiety improved in seven of eight studies, and the magnitude of improvement reached statistical significance in four of these. A review of 34 sham-controlled trials concluded that regular CES treatments resulted in short-term symptomatic relief of generalized anxiety through direct effects on autonomic brain centers. Patients diagnosed with one or more phobias reported significant reductions in state anxiety when exposure to the phobic stimulus was followed by 30 minutes of CES treatment. Daily application of weak electrical current to the neck or the head may stimulate

global brain activity, resulting in beneficial changes in the activity of neurotransmitters implicated in dementia. A meta-analysis of three studies on CES in patients with dementia found evidence of significant but transient improvements in word recall, face recognition, and motivation immediately following treatment that were not sustained at 6-week follow-up.

CES is used to lessen symptoms of alcohol and drug withdrawal. The postulated mechanism of action involves stimulation of release of endorphins, enkephalins, and other endogenous opioid peptides and may prove to be similar to the mechanism underlying electroacupuncture. In a 4-week RCT (N = 20), depressed alcoholics were randomized to receive 20 CES treatments at 70 to 80 Hz, 4 to 7 mA, versus sham treatments. Patients who received CES treatments experienced significantly reduced anxiety by the end of the study. Findings of a sham-controlled study (N = 60) on inpatient alcohol or polysubstance abusers suggested that daily 30-minute CES treatments significantly improved cognitive functioning and reduced measures of stress and anxiety during the acute phases of withdrawal in this population. In a 7-year prospective study of CES in the treatment of alcohol, drug, and nicotine addiction, acute and chronic withdrawal symptoms were diminished, normal sleep patterns were restored more rapidly, and more patients remained addiction-free following regular CES treatments compared to conventional psychopharmacological management. The discrepancy between largely positive outcomes from studies on CES and frequent negative findings on electroacupuncture (see later discussion) in detoxification may be partly attributable to suboptimal current or frequency settings in electroacupuncture protocols investigated for this clinical application.

CES is widely used to treat insomnia. However, findings of sham-controlled studies are inconsistent. Whereas some studies showed no beneficial effects of CES in insomnia, others reported sustained improvements in the timing of sleep onset and total sleep duration. Disparate findings on the therapeutic benefits of CES in insomnia may be related to differences in equipment, protocols, and duration and timing of treatment. Serious adverse effects of microcurrent electrical brain stimulation have not been reported.

THERAPIES BASED ON POSTULATED FORMS OF ENERGY NOT VALIDATED BY CURRENT SCIENCE.

CAM approaches based on hypothesized *subtle* forms of energy include acupuncture, homeopathic remedies, Healing Touch, *qigong,* Reiki, and energy psychology. Beneficial effects of these modalities rest on the postulated role of poorly characterized energetic *fields* or directed human intention that are difficult to examine using contemporary research methods and may not be *explainable* by current science. Chinese medicine, Ayurveda, and Tibetan medicine posit the existence of a subtle “energy”

body and a gross physical body. All Asian healing traditions use physical, psychological, and spiritual approaches to treat postulated *energetic* imbalances and restore optimal health. In Chinese medicine, "*qi*" is regarded as a fundamental kind of energy that cannot be explained by contemporary science but may have properties consistent with quantum mechanics. Ayurveda posits an analogous energetic principal called "*prana*." In both systems of medicine, a central goal of treatment is to restore "balance" in the energetic and physical body, thus enhancing the body's capacity for self-healing. Like Asian healing traditions, homeopathy posits that physical and emotional symptoms result from imbalances in a fundamental *vital force*. Current science cannot explain the postulated curative mechanism of homeopathic remedies, which are typically diluted beyond the limit of possible biological action. Reiki holds that a fundamental energetic principle, "*ki,*" promotes healing when guided by a skilled practitioner. Physical contact between the practitioner and the patient is not believed to be necessary for beneficial effects to take place. Healing Touch (HT) is an "energy healing" method used to treat both physical and mental illness. Specific energetic techniques include "charka spreading," "magnetic unruffling," "mind clearing," and "stopping." Healing Touch practitioners claim that "energetic" contact between the practitioner and the patient results in alleviation of symptoms. In Healing Touch, the practitioner positions the hands over specific areas of the body but does not physically *touch* the patient's body. Therapeutic Touch (TT) is similar to Healing Touch but uses gentle physical contact to promote healing. In contrast to Healing Touch and Therapeutic Touch, the Reiki practitioner does not consciously direct healing intention to specific regions of the body but serves as a "conduit" for *ki* to *flow into* the patient to restore healthy energetic functioning. "Somatoemotional release" is an energy therapy in which the practitioner "releases" pathological energetic states that accumulate in the body following physical injury. The technique involves gentle touch while guiding the patient to assume positions that stimulate body memories of past trauma. "Energy psychology" is an eclectic combination of Western psychological theory and Chinese medical theory that posits that energetic imbalances in the meridians (i.e., pathways through which *qi* flows) are associated with different physical, emotional, and mental symptoms. "Thought field therapy" (TFT) and "emotional freedom technique" (EFT) are specific "energy psychology" approaches used to treat anxiety and depression. In TFT, the patient is asked to invoke a "thought field" associated with depressed mood, a traumatic memory, or other principal symptom pattern. The TFT practitioner then *reattunes* energetic imbalances associated with the target symptom by tapping on specific acupuncture points in a specified manner in an effort to alleviate symptoms. EFT is a simplified version of TFT that involves only one routine for stimulating acupuncture points. Twelve of 18 RCTs included in

a review of noncontact *biofield* therapies including Reiki, Healing Touch, Johrei and Therapeutic Touch reported statistically significant positive outcomes; however, findings were limited by the small sample size of most studies and methodological problems.

Acupuncture. Evaluating the clinical efficacy of acupuncture in depressed mood and other mental health problems poses many methodological challenges because of differences in the severity and comorbidity of mental, emotional, and physical symptoms in study populations; concurrent use of other conventional or CAM treatments; conceptual differences between Chinese medicine and biomedical diagnoses; and the use of individualized acupuncture treatment protocols reflecting the patient's "energetic" balance. Sham-controlled studies suggest that conventional needle acupuncture, electroacupuncture, and computer-controlled electroacupuncture (CCEA) have consistent beneficial effects on depressed mood. Findings of a double-blind, sham-controlled study support that traditional manual acupuncture (i.e., needles in the absence of electrical current) is an effective treatment of severe depressed mood. By the end of the 8-week study, 68 percent of 33 female outpatients being treated with a specialized acupuncture protocol directed at depressed mood had achieved full remission. However, matched depressed women in a wait-list group showed equivalent improvement. In a large 6-week, multicenter study (N = 241), depressed inpatients randomized to receive electroacupuncture plus placebo versus electroacupuncture plus amitriptyline experienced equivalent improvements in mood. Electroacupuncture was superior to amitriptyline in patients who did not report comorbid anxiety. Patients treated with electroacupuncture had significantly elevated plasma norepinephrine concentrations following a 6-week course of treatment, suggesting that the mechanism of action of electroacupuncture involves stimulation of norepinephrine release. CCEA uses computer-guided modulation of the frequency and waveform of electrical current delivered through acupuncture needles. Preliminary findings suggest that high frequencies (up to 1,000 Hz) yield responses in depressed patients that may be superior to both conventional acupuncture and electroacupuncture. A narrative review of controlled studies, outcomes studies, and published case reports on acupuncture as a treatment of anxiety and depressed mood supported that sham-controlled studies yielded consistent improvements in anxiety using both regular acupuncture and electroacupuncture. Research findings on acupuncture in depressed mood and anxiety should be regarded as inconclusive because of methodological flaws including the absence of standardized symptom rating scales in most studies, limited follow-up, and poorly defined differences between protocols used in different studies.

A review of prospective trials on acupuncture as a treatment of PTSD

identified four quality sham-controlled studies and two quality uncontrolled trials that met inclusion criteria. One high-quality trial included in the review showed statistically significant differences between the acupuncture and wait-list group but nonsignificant differences between the acupuncture and CBT groups. Patients receiving acupuncture or CBT continued to report clinical improvements in PTSD symptoms 3 months after study endpoint. These findings are limited by the small number of trials that met inclusion criteria (only one study reviewed was included in the analysis), the absence of sham-controlled studies, the use of different study designs across trials examined, and poor methodological quality of many studies.

Research findings on acupuncture for relapse prevention in alcohol abuse are inconsistent, reflecting differences in the selection of acupuncture points, the treatment protocol used (i.e., conventional vs. electroacupuncture), frequency of treatments, the duration of total treatment, and the relative skill or specialized training of practitioners. Positive findings of two controlled trials supported the hypothesis that specific acupuncture protocols significantly reduced alcohol craving and reduced relapse rate in recovering alcoholics. However, a subsequent RCT found no significant differences in craving or relapse rates between an acupuncture protocol traditionally used to treat addiction, sham transdermal stimulation on random points, and a wait-listed group.

Findings of most controlled trials on acupuncture for smoking cessation have been negative or equivocal. Three meta-analyses of sham-controlled studies on the efficacy of acupuncture for smoking cessation (total N >2,000) concluded that therapeutic acupuncture protocols and sham acupuncture have equivalent efficacy, suggesting a significant placebo effect. Longer sham-controlled studies are needed to determine whether more frequent acupuncture treatment following a specific protocol or a greater number of total treatments is effective for smoking cessation. Two systematic reviews of sham-controlled trials concluded that both conventional acupuncture and electroacupuncture are equally ineffective in reducing symptoms of nicotine withdrawal and controlling cocaine addiction. A large study examining three auricular acupuncture protocols for relapse prevention in narcotics abusers concluded that all three protocols were associated with reductions in drug use over time. An 8-week RCT (N = 32) compared acupuncture with two conventional drugs versus placebo in cocaine addicts on methadone maintenance therapy. Half of the subjects dropped out. However, almost 90 percent of those who completed the study achieved abstinence by the end of the study. Patients who achieved abstinence reported diminished craving and significantly improved mood.

The absence of standardized acupuncture treatment protocols for insomnia has interfered with efforts to design rigorously controlled double-

blind studies. A systematic review of all studies on acupuncture for insomnia published in the English language through 2002 identified few studies for analysis because most studies had not employed randomized, double-blind procedures. Acupuncture may be beneficial for insomnia associated with generalized anxiety and schizophrenia. In a large-case series (*N* = 500), schizophrenics reported significant improvements in sleep following acupuncture treatments. In a small RCT (*N* = 40), patients complaining of insomnia were diagnosed using Chinese medical assessment methods and randomized to an acupuncture protocol addressing the specific energetic "imbalance" underlying their sleep disturbance versus a sham protocol. Significant and sustained improvements in the subjective quality of sleep and changes on EEG polysomnography reflecting normalization of sleep were found in the treatment group but not in the sham group.

Uncommon transient adverse effects associated with acupuncture include bruising, fatigue, and nausea. Infrequent cases of infection with human immunodeficiency virus, hepatitis B, and hepatitis C have been reported when nonsterilized needles are used. Rare cases of pneumothorax and cardiac tamponade have been reported as a result of accidental puncturing of the lungs or the pericardium.

Homeopathy. Preliminary findings from open studies suggest that certain homeopathic remedies may improve depressed mood. Short-term and long-term constitutional remedies are used, depending on symptom severity, personality factors, and the "energetic pattern" associated with depressed mood. Homeopathic preparations are widely used to self-treat generalized anxiety and other anxiety symptoms. However, research findings are inconsistent and based largely on case reports. Case reports of symptomatic improvement following the administration of specific homeopathic remedies have been published for PTSD, social phobia, panic disorder, and OCD. However, findings of two placebo-controlled studies were negative or equivocal. Students with test-taking anxiety (*N* = 62) randomized to a specific homeopathic preparation (*Argentum nitricum* 12X) versus placebo twice daily for 4 days reported equivalent minor reductions in the severity of anxiety. In a 10-week prospective study (*N* = 44), patients diagnosed with generalized anxiety disorder were randomized to an energetically appropriate homeopathic treatment versus placebo. All subjects were medication-free at least 1 month prior to the start of the study. All subjects were individually diagnosed by a classically trained homeopathic physician and treated with a homeopathic remedy addressing identified "energetic imbalances." Significant and *equivalent* improvements in baseline anxiety were found in the placebo and treatment groups at week 5 and at the end of the study. Homeopathic remedies are generally too dilute to potentially cause adverse effects. However, rare

cases of toxicity have been reported when low dilutions of arsenic, cadmium, or other heavy metals are used. Concerns about possible carcinogenic effects of remedies containing *Aristolochia* have been raised, and there are rare reports of allergic reactions.

Healing Touch and Therapeutic Touch. Healing Touch practitioners claim to be able to intuitively assess energetic *imbalances* manifesting as physical or psychological symptoms and "*direct*" healing energy to specific bodily regions to correct purported imbalances, resulting in improved functioning or general feelings of well-being. In a 3-week open study (N = 30), moderately to severely depressed patients were randomized to receive HT versus a wait list. Individuals in the treatment group received two HT sessions weekly. Control patients received no treatment, including no exposure to a sham HT protocol. By the end of the study, patients in the treatment group experienced significantly improved mood that was sustained at 1-month follow-up. These findings should be regarded as inconclusive because of the absence of blinding, small study size, and the absence of a sham treatment group. Open studies, case reports, and one double-blind trial suggest that HT and TT may reduce the severity of agitation in dementia. In one small open study, measures of agitation improved significantly in 14 individuals with dementia in a residential setting who received three HT treatments weekly over 4 weeks. Three individuals in the study needed fewer sedating medications during the active treatment phase, and two residents required dose increases in the first 2 weeks after HT treatments were stopped. In a small 5-week, sham-controlled study, two weekly 10- to 20-minute HT treatments were administered to patients with Alzheimer disease. Patients who received regular HT treatments displayed consistent reductions in disruptive behaviors, globally improved emotional and cognitive functioning, enhanced socialization, a more regular sleep schedule, improved compliance with nursing home routines, greater emotional stability, and improved communication with staff. Individuals who received regular HT treatments complained of physical distress or discomfort significantly less often than the control group. Measures of agitation improved significantly in the HT group but did not change in the control group. In a double-blind study (N = 57) that included mock-TT in the control arm, agitated patients with dementia who received two brief TT treatments daily for 3 days exhibited significantly fewer behavioral symptoms of dementia, including reduced restlessness and fewer disruptive vocalizations, compared to patients who received mock-TT. In a 3-week RCT (N = 123), active duty military diagnosed with PTSD randomized to semi-weekly sessions of HT plus guided imagery experienced greater improvement in PTSD symptom severity and quality of life compared to those receiving treatment as usual. The significance of research findings on TT and HT is limited by study

design flaws, including lack of blinding, the confounding influence of frequent human contact, small numbers of enrolled subjects, the lack of standardized methods, and the absence of control groups receiving sham HT or TT treatments in most studies.

Qigong. *Qigong* is an ancient energy healing method used in Chinese medicine that has its origins before written language. Most published studies have examined the effectiveness of regular *qigong* practice on general indicators of physical or emotional health. Individuals can engage in a regular *qigong* practice or receive "emission *qigong*" treatments for a particular medical or psychological problem from a skilled *qigong* master. Findings of open trials on *qigong* as a treatment of "neuroticism" suggest that anxiety, depressed mood, and quality of sleep were significantly better in subjects who had practiced *qigong* consistently for at least 5 years than in those who had never practiced *qigong*. An open study reported that depressed individuals who practiced *qigong* consistently for at least 2 years experienced improved mood and general improvement in their baseline emotional state. In a 2-month open pilot study (N = 39), patients with a diagnosis of major depressive disorder, dysthymic disorder, or bipolar disorder based on criteria of the text revision of the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR) received *qi* "emission" treatments administered by *qigong* masters. Half of subjects engaged in the regular daily practice of *qigong*. By the end of the study, standardized ratings of depressed mood had improved for all subjects, and there were no significant differences in response between individuals who practiced *qigong* while receiving *qi* "emission" therapy and subjects who received *qi* "emission" therapy only. Larger studies that include a sham *qigong* arm are needed to rule out significant placebo effects. A systematic review and meta-analysis of studies on qigong "exercise" for stress and anxiety in adults identified seven RCTs that met the inclusion criteria. Four RCTs reported positive findings of qigong for generalized anxiety and three studies reported beneficial effects on stress; however, the significance of findings was limited by methodological flaws and small study size. A systematic review and meta-analysis of 15 studies on qigong for anxiety, depressed mood and psychological well-being found no conclusive evidence for qigong as a treatment of a specific mental health condition; however, the findings suggested that regular qigong practice may have positive effects on general psychological well-being in patients with chronic illnesses.

The *unskillful* or excessive practice of *qigong* can result in serious psychological or psychosomatic symptoms. Psychotic symptoms that have been reported following *erroneous qigong* are described as a culture-bound syndrome, "*qigong* psychotic reaction," in the DSM. Schizophrenics and individuals diagnosed with severe personality disorders should be

cautioned to practice *qigong* moderately and only under the guidance of a skilled *qigong* instructor.

Reiki. Findings of a 6-week, single-blind, sham-controlled trial of Reiki on depressed mood (N = 73) suggested that regular Reiki treatments may increase feelings of general well-being and reduce stress. The study population consisted of volunteers who complained of depressed mood or "stress" but did not meet diagnostic criteria for a major psychiatric disorder. Subjects were randomized to weekly 1½-hour, hands-on Reiki versus distant Reiki or distant sham Reiki treatments. Patients who received hands-on or distant Reiki treatments experienced significant and equivalent improvements in depressed mood and stress, which were sustained 1 year after treatment was terminated. No one in the sham Reiki group improved during or following the study. Reiki treatments may improve state anxiety in chronic pain patients. In a small sham-controlled study (N = 12), chronic pain patients were randomized to receive Reiki versus sham Reiki, progressive muscle relaxation, or no treatment. Improvements in pain and state anxiety were significantly greater among patients receiving Reiki than sham-Reiki and relaxation training. The significance of these findings is limited by the absence of controls on anxiolytic medication use during the study. Although these preliminary findings are suggestive of therapeutic effects of Reiki, larger sham-controlled studies enrolling subjects diagnosed with mood or anxiety disorders are needed to evaluate putative mood-enhancing and stress-reducing effects of hands-on and distant Reiki.

Energy Psychology. EFT, TFT, and body psychotherapy are "energy psychology" techniques used to treat mood disorders, phobias, and other anxiety disorders. To date, few controlled studies have been done, most studies are methodologically flawed, and findings are highly inconsistent. EFT treatments can be self-administered using a manual following an initial training session, and no technical expertise is required. In a double-blind, prospective trial (N = 35), patients who met DSM-IV-TR criteria for a specific phobia were randomized to receive one EFT treatment versus one session of diaphragmatic breathing (DB) while exposed to a feared small animal. EFT-treated patients experienced greater improvement than DB-treated patients in four of five measures of anxiety, and improvements were sustained at 3 months in the absence of further EFT or DB treatments. These findings were replicated in a subsequent study. A study comparing EFT and sham-EFT in anxious individuals who had not been formally diagnosed found no differences in anxiety reduction between a defined EFT protocol and a sham protocol. These findings suggested that therapeutic benefits of EFT may be attributable to desensitization and distraction. A sham-controlled study on EFT in depression found equivalent nonsignificant changes in mood following EFT and sham treatment.

Somatoemotional Release And Craniosacral Therapy. Practitioners of craniosacral therapy gently guide the spinal and cranial bones with the goal of bringing the CNS into greater harmony. The movement of the cerebrospinal fluid (CSF) is believed to be restricted by bodily trauma or chronic stress. Gentle manipulation of the spine, the skull, and fascia is believed to ease restrictions on the openings through which nerves pass while optimizing flow of CSF through the spinal cord. Vietnam veterans (N = 22) diagnosed with PTSD completed a 2-week intensive program in which they were treated with both somatoemotional release and craniosacral therapy. The majority of veterans reported significant reductions in pain and physical distress and improvements in depressed mood, anxiety, suspiciousness, guardedness, and behavioral isolation. The significance of these findings is limited by the small study size, the absence of a sham treatment arm, lack of follow-up, and the use of nonstandardized symptom rating scales.

Shamanic Healing. A shaman is a person who is regarded by his or her tribe or sociocultural group as having the power to heal the sick, to predict the future, and to communicate with the supernatural world, including primitive natural forces and spirits of the dead. The term *shaman* has ancient roots in Siberia. However, similar traditions of spiritual healing can be found in all world regions. Shamanic healing ceremonies often focus on retrieving the patient's *lost soul*, which is believed to be a root cause of many physical and psychological illnesses. The therapeutic goals of ritual acts, stories, and chants used in shamanic healing vary widely depending on the parent culture. Research on shamanic healing is generally descriptive, and there are few published studies using clinical research methodology. One study reported on shamanistic contributions to the treatment of schizophrenia and other mental illnesses in Japan. Another study examined the outcomes and satisfaction of 188 patients treated by shamans in Taiwan with a matched sample of patients treated by physicians. Five distinctive sickness types were identified, including mental health problems. Patient and experimenter reports suggested that three-fourths of patients improved. However, shamanic treatment was more likely to be described as difficult or uncomfortable than conventional medical treatment. Findings of a more recent study suggested that shamanic healing approaches contribute to the psychiatric stability and general emotional well-being of individuals diagnosed with bipolar disorder.

Outcomes research on shamanic healing is at a very preliminary stage. Applying standard biomedical research designs to indigenous healing practices poses many complex challenges, resulting in studies that are primarily naturalistic and observational in design. However, increasing interest in indigenous healing practices in the context of promising early findings on shamanic healing and the importance of fostering cultural

belief systems in community mental health care will probably lead to growing acceptance of shamanic healing in North America, western Europe, and other developed world regions.

References

Ashford JW, Mahoney L, Burkett T. A role for complementary and integrative medicine in Alzheimer's disease prevention. *J Alzheimers Dis*. 2015;48(1):13–14.

Aung SK, Fay H, Hobbs RF III. Traditional Chinese medicine as a basis for treating psychiatric disorders: a review of theory with illustrative cases. *Med Acupunct*. 2013;25(6):398–406.

*Borre YE, Moloney RD, Clarke G, Dinan TG, Cryan JF. The impact of microbiota on brain and behavior: mechanisms & therapeutic potential. *Adv Exp Med Biol*. 2014;817:373–403.

De Oliveira AM, Radanovic M, De Mello PC, et al. Nonpharmacological interventions to reduce behavioral and psychological symptoms of dementia: a systematic review. *Biomed Res Int*. 2015;2015:218980.

Edenfield TM, Saeed SA. An update on mindfulness meditation as a self-help treatment for anxiety and depression. *Psychol Res Behav Manag*. 2012;5:131–141.

Freeman MP, Fava M, Lake J, Trivedi MH, Wisner KL, Mischoulon D. Complementary and alternative medicine in major depressive disorder: the American Psychiatric Association Task Force report. *J Clin Psychiatry*. 2010;71(6):669–681.

Gevensleben H, Rothenberger A, Moll GH, Heinrich H. Neurofeedback in children with ADHD: validation and challenges. *Expert Rev Neurother*. 2012;12(4):447–460.

Gillies D, Sinn JKh, Lad SS, Leach MJ, Ross MJ. Polyunsaturated fatty acids (PUFA) for attention deficit hyperactivity disorder (ADHD) in children and adolescents. *Cochrane Database Syst Rev*. 2012;(7):CD007986.

Hammerschlag R, Marx BL, Aickin M. Nontouch biofield therapy: a systematic review of human randomized controlled trials reporting use of only nonphysical contact treatment. *J Altern Complement Med*. 2014;20(12):881–892.

Herrmann W, Lorenzl S, Obeid R. Review of the role of hyperhomocysteinemia and B-vitamin deficiency in neurological and psychiatric disorders—Current evidence and preliminary recommendations. *Fortschr Neurol Psychiatr*. 2007;75(9):515–527.

Jiao J, Li Q, Chu J, Zeng W, Yang M, Zhu S. Effect of n-3 PUFA supplementation on cognitive function throughout the life span from infancy to old age: a systematic review and meta-analysis of randomized controlled trials. *Am J Clin Nutr*. 2014;100(6):1422–1436.

Kasper S. Phytopharmaceutical treatment of anxiety, depression, and dementia in the elderly: evidence from randomized, controlled clinical trials. *Wien Med Wochenschr*. 2015;165(11–12):217–228.

Lake JH. *Textbook of Integrative Mental Health Care*. New York, NY: Thieme Medical; 2006.

Lake JH, Spiegel D, eds. *Complementary and Alternative Treatments in Mental Health Care*. Washington, DC: American Psychiatric Publishing; 2007.

Lopresti AL, Drummond PD. Saffron (Crocus sativus) for depression: a systematic review of clinical studies and examination of underlying antidepressant mechanisms of action. *Hum Psychopharmacol*. 2014;29(6):517–527.

Louie L. The effectiveness of yoga for depression: a critical literature review. *Issues Ment Health*

Nurs. 2014;35(4):265–276.

*Marchand WR. Mindfulness-based stress reduction, mindfulness-based cognitive therapy, and Zen meditation for depression, anxiety, pain, and psychological distress. *J Psychiatr Pract*. 2012;18(4):233–252.

*Moshiri M, Vahabzadeh M, Hosseinzadeh H. Clinical applications of saffron (Crocus sativus) and its constituents: a review. *Drug Res (Stuttg)*. 2015;65(6):287–295.

Papakostas GI, Shelton RC, Zajecka JM, et al. Effect of adjunctive L-methylfolate 15 mg among inadequate responders to SSRIs in depressed patients who were stratified by biomarker levels and genotype: results from a randomized clinical trial. *J Clin Psychiatry*. 2014;75(8):855–863.

Qureshi NA, Al-Bedah AM. Mood disorders and complementary and alternative medicine: a literature review. *Neuropsychiatr Dis Treat*. 2013;9:639–658.

Reiter K, Anderson SB, Carlsson J. Neurofeedback treatment and posttraumatic stress disorder: effectiveness of neurofeedback on posttraumatic stress disorder and the optimal choice of protocol. *J Nerv Ment Dis*. 2016;204(2):69–77.

Sarris J. Herbal medicines in the treatment of psychiatric disorders: a systematic review. *Phytother Res*. 2007;21(8):703–716.

Sarris J, Lake J, Hoenders R. Bipolar disorder and complementary medicine: current evidence, safety issues, and clinical considerations. *J Altern Complement Med*. 2011;17(10):881–890.

Sarris J, McIntyre E, Camfield DA. Plant-based medicines for anxiety disorders, part 1: a review of preclinical studies. *CNS Drugs*. 2013;27(3):207–219.

Sarris J, McIntyre E, Camfield DA. Plant-based medicines for anxiety disorders, part 2: a review of clinical studies with supporting preclinical evidence. *CNS Drugs*. 2013;27(4):301–319.

Sarris J, Murphy J, Mischoulon D, et al. Adjunctive nutraceuticals for depression: a systematic review and meta-analyses. *Am J Psychiatry*. 2016;173(6):575–587.

Slyepchenko A, Carvalho AF, Cha DS, Kasper S, McIntyre RS. Gut emotions—mechanisms of action of probiotics as novel therapeutic targets for depression and anxiety disorders. *CNS Neurol Disord Drug Targets*. 2014;13(10):1770–1786.

*Spedding S. Vitamin D and depression: a systematic review and meta-analysis comparing studies with and without biological flaws. *Nutrients*. 2014;6(4):1501–1518.

Wang SM, Han C, Lee SJ, Patkar AA, Masand PS, Pae CU. A review of current evidence for acetyl-l-carnitine in the treatment of depression. *J Psychiatr Res*. 2014;53:30–37.

▲ 31.5 Disaster Psychiatry: Disasters, Terrorism, and War

DAVID M. BENEDEK, M.D., JOSHUA CHAIM MORGANSTEIN, M.D., HARRY C. HOLLOWAY, M.D., AND ROBERT J. URSANO, M.D.

Disasters impact hundreds of communities and affect the lives of millions of persons around the globe every year. They may result from the destructive forces of nature (as in hurricanes, earthquakes, tsunamis, or infectious disease outbreaks), from human negligence (as in forest fires sparked by carelessness or a plane crash), or from deliberate acts (as in large-scale terrorist attacks or military combat operations). Regardless of origin, a disaster's massive destruction creates an atmosphere of chaos that compels individuals to face the consequences of injury, loss, and death. In any disaster, psychological injury may occur in conjunction with physical injury, exposure to the injury, and death of others, as well as from the disruption of the physical environment or as a consequence of the terror and helplessness that these events evoke.

Disaster psychiatry includes anticipating the scope of behavioral, emotional, and distress-related responses of individuals and communities, planning the response, and ultimately providing care for those exposed to the trauma of human-caused and natural disasters. War and terrorism are particular subtypes of human-caused disasters and typically cause greater adverse psychological impact than natural disasters. Because war (and also military peacekeeping and humanitarian relief operations) places military health care providers in direct and prolonged contact with disaster-affected populations, the knowledge base, skills, and professional attitudes required for psychiatric practice in military operations are closely aligned with those necessary for other disasters. In any disaster (including war and terrorism), psychological injury must be addressed in a timely manner. Mental health care must be provided, not only in hospitals or clinics but also in environments near sites of chaos and destruction, where traditional medical resources may be scarce. Care providers must use public health models of health surveillance, outreach, and education to overcome barriers to care, including those safety and survival interests coexisting with mental health needs, limited physical and psychiatric care resources, and stigma associated with requesting or receiving mental health assistance. The disciplines of military and disaster psychiatry address these care demands in nontraditional environments and in mass casualty situations in which resources are overwhelmed. As the reality of terrorism and the potential for mass casualties within the United States becomes clear, disaster psychiatry highlights the importance of integrating mental health and public health for mass disasters. Care in these environments

relies on not only psychiatrists but also other physicians, social scientists, epidemiologists, nurses, and emergency responders such as police and firefighters. The coordination of these diverse care providers is necessary to ensure that needs are met and that no subgroup of victims is overlooked. This approach to population health care delivery and management, including primary (prevention), secondary (early treatment), and tertiary (treating established disease and rehabilitation) care, is a critical aspect of disaster psychiatry in military operations and other disaster environments.

In the United States, the terrorist attacks on the Pentagon and the World Trade Center and the distribution of anthrax-tainted mail in the weeks that followed resulted in widespread fear and uncertainty in the American public. Subsequently, around the globe the issues of severe acute respiratory syndrome (SARS), tube (subway) bombings in London, train bombings in Spain, suicide attacks in Pakistan, and polonium poisonings in London further focused disaster care to individuals and groups across explosive, radiological, and biological agents. These events also resulted in neuropsychiatric syndromes in many survivors of the attacks and their relatives. Ongoing response to these events involved numerous government public health and law enforcement agencies, civilian disaster relief organizations, and military medical, mental health, and law enforcement personnel. The response to the events in the United States (including the continued war on terrorism), as well as military participation in humanitarian relief efforts after the Southeast Asian tsunami of December 2004, Hurricanes Katrina and Rita of 2005, Haiti earthquakes of 2010, the Philippines' Typhoon Haiyan of 2013, and the West Africa Ebola outbreak of 2014 illustrate the intertwined nature of military and disaster psychiatry.

ENVIRONMENTS OF CARE IN DISASTER AND MILITARY OPERATIONS

The term *disaster* has several definitions or connotations. The word is derived from the Latin words *dis* (meaning *against*) and *astrum* (meaning *stars*)—*the stars are against us*. Ancient societies may have conceived of disasters, such as floods or earthquakes, as acts of the gods. However, the distinction between *natural* disasters (e.g., earthquakes) and *human-caused* or technological disasters, such as explosions or plane crashes, is increasingly difficult to make. Much of the death and destruction from an earthquake may result from damage to poorly constructed housing in urban areas—thus, there is a human-caused aspect even to natural disasters. Certainly, human-caused disasters can result from accidents (e.g., factory explosions), but human-caused disasters also result from intentional acts in the cases of war or terrorism. The degree to which a disaster may be anticipated and the characteristics of the affected population (soldiers or innocent civilians) may be more critical distinctions than whether the disaster was natural or human caused.

War is an all too common human-caused disaster. At the time of

writing, there are more than 46 ongoing armed conflicts involving 41 countries occurring around the globe. War has been defined as a political act (generally involving violence) to achieve national objectives or to protect natural interests. In these actions, militaries create their own and civilian casualties in armed conflicts of all kinds. During recent years, militaries around the world have also been prominently engaged in humanitarian disaster relief, as for hurricanes in the United States, earthquakes in Turkey, and the recent tsunami in Southeast Asia, to name only a few. Indeed, for most countries around the world, their military forces typically represent the most significant portion of disaster-response personnel resources. In addition, militaries are involved in government stabilization and peacekeeping efforts (military operations other than war). Terrorism, like war, involves violence to achieve political objectives. Here, however, the creation of fear or terror to undermine society's confidence in its government and leaders is the objective. As such, fear, anxiety, and a sense of chaos are the goals of terrorism rather than the (albeit predictable) aftereffects observed in wars and other human-caused or natural disasters. Therefore, mitigating the emotional and psychological effects of terrorism is crucial to sustaining and protecting a nation from the impact of terrorist attacks.

The causes of emotional symptoms in response to disasters including war have been historically attributed to sources ranging from the gods, the concussive force of a passing cannon ball, and various natural or supernatural sources of contagion. Emotional consequences are alluded to frequently in the descriptions of the terror induced by the attack of the hero in Homer's *Iliad*. Ancient Greeks attributed epidemic illness to Apollo's wrath after the desecration of his temple. The French military surgeon Dominique-Jean Larrey commented clearly on the ill effects of war on the health of Napoleon's soldiers. Combat-related pathological behavior similar to the symptoms of posttraumatic stress disorder (PTSD), defined in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), is described in Civil War medical treatises. The science of neurology entered military medicine with Weir Mitchell's work in this era. During the remainder of the 19th and 20th centuries, concepts of neurasthenia, dissociation, hysteria, and psychological suggestion were developed to define psychoneurological states for which no demonstrable neuroanatomical pathology could be identified. Military physicians in the Russian–Japanese War made similar observations. Recognition that nontraumatic injuries followed train derailments occurred even as World War I and II physicians initiated treatment models for conditions such as *shell shock, gas neurosis,* and *conversion paralysis*. Civilian physicians have also long recognized the concept that war results in physical and emotional human suffering. In 1859, Henri Dumont arranged for civilian medical services for the injured after observing soldiers die from lack of

medical attention during the Battle of Solferino. His efforts led to the establishment of the International Red Cross and to international guidelines for humane care of the sick and wounded during wartime. Beginning in the last quarter of the 20th century, the Red Cross and other international medical and relief agencies, such as Doctors without Borders, have increasingly provided mental health–related consultation and direct care in the aftermath of war and natural and human-caused disasters. Observations during these military and civilian medical efforts formed the basis for today's understanding that aspects of the environment and the unwanted environmental changes brought about by war or disaster result in emotional trauma.

Regardless of the source, potential stressors in disaster and military environments include exposure to the dead and grotesque, immediate and pending threat to life, loss of loved ones, loss of personal property, and physical injury with associated disability and pain. Such stressors may occur transiently but more often must be endured for months or years. Although the military brings supplies and logistical infrastructure, including food, shelter, and specialized equipment and clothing to protect soldiers at war, civilians exposed to combat in modern times may be subject to large-scale destruction within their communities. They may become refugees and may experience shortages of food or shelter that threaten life. These noncombatant victims may be fortunate enough to receive basic medical attention for injuries, but often they do not receive treatment for psychiatric symptoms that may emerge as a consequence of exposure to bombings, battle, rape, torture, and unrestrained murder. Although an earthquake may last only seconds or minutes, the traumatic experiences resulting from the destruction of home, geographic displacement from familiar surroundings, a loss of employment or because of injury or loss of loved ones may continue for weeks, months, or, possibly, years. Combat environments, ongoing terrorist campaigns (e.g., the suicide bombings in Israel), and cases of natural disaster, such as earthquake exposure, over time may include vaguely anticipated or entirely unexpected life-threatening experiences, followed by daily life in an austere or disrupted environment. This backdrop of chronic loss, depletion of resources, and exposure to reminders of recent tragedies shapes and defines the environment in which emotional and behavioral responses to disaster and military operations may evolve.

TRAUMATIC STRESS RESPONSE IN POPULATIONS AND INDIVIDUALS

The psychological and behavioral responses after disaster occur in six phases. An understanding of these phases allows those charged with disaster management to more effectively plan and respond to the mental health needs of individuals and communities following a disaster event. A Pre-Disaster phase occurs when an event is anticipated or advanced

warning is given. This phase can result in feelings of vulnerability and worry about safety. Individual who do not heed advanced warnings to take action, such as evacuation, may experience significant remorse and feelings of responsibility for subsequent injury to loved ones or damage to property.

The Impact phase occurs immediately after an acute event and consists of strong emotions, including feelings of disbelief, numbness, fear, confusion, and panic behavior. These are often normal emotional responses to an extraordinary event. During this time, if the scope of a disaster broadens, the psychological effects are typically increased.

This is followed by the brief Heroic phase, which usually lasts days to a few weeks. Adrenaline-induced rescue and survival behaviors predominate, though disorientation from the acute event may result in activity of low productivity. Injury of family members or separation of the family unit can produce increased anxiety and worry that detracts from the attention necessary for immediate problem solving. This phase is frequently accompanied by the initial appearance of assistance from outside agencies and communities and involves efforts at adaptation to the new environment. Intrusive symptoms (unbidden recollections of the attack, explosion, fire, building collapse, and others, in the form of flashbacks or nightmares), as well as hyperarousal, emerge during this phase. Somatic symptoms, such as fatigue, dizziness, headaches, and nausea, along with anger, irritability, and social withdrawal, may also emerge.

The Honeymoon phase coincides with more extensive availability of government and volunteer assistance. Community bonding often occurs as a result of sharing the catastrophic experience as well as the giving and receiving of assistance. Survivors are often more hopeful during this phase and experience an optimism that the help they will receive will make them whole again. Disaster behavioral health workers are most likely to be perceived as helpful during this phase, be readily accepted by community members, and develop a foundation from which to provide assistance in the difficult phases ahead.

The Disillusionment phase is marked by feelings of disappointment and resentment, as disaster assistance agencies and volunteer groups begin to pull out of the community and hopes for aid and restoration of the predisaster emotional and physical environment are not met. The sense of community is weakened as individuals focus on their personal needs or the extent to which these needs are still unmet. Ill will and resentment may surface as survivors receive unequal monetary compensation for what they perceive to be equal or similar damage. In addition, neighboring communities less impacted by the disaster have often returned to life as usual which can discourage and alienate those who were more severely impacted. During this phase, survivors may become physically exhausted because of enormous multiple demands, including financial pressures as well as relocation or living in a damaged home, family discord, bureaucratic

hassles, and a lack of free time for recreation or self-care. Health problems and exacerbation of preexisting conditions emerge because of ongoing stress and fatigue. The disaster "Anniversary" experience can occur during this phase and can be a critical opportunity for leaders to support the psychological well-being of disaster victims through memorializing, making meaning and focusing on "building back better." Failure to effectively address a disaster anniversary experience can further demoralize survivors and exacerbate underlying psychological distress and feelings of disillusionment.

The final, or Reconstruction, phase may last for years. Here, survivors attempt to rebuild their lives and social and occupational identities by returning to old jobs or finding new work, rebuilding homes, and resuming or establishing new social ties and emotional support systems. For some survivors, this phase is marked by an acceptance of new circumstances, including the changes and losses that have occurred. Individuals who are able to find meaning may experience posttraumatic growth, ultimately emerging from the disaster event and its aftermath with an increased sense of personal strength. Because individuals may progress through these phases at variable rates, persons involved in planning and delivering care to victims of disasters must recognize that persons may manifest emotional symptoms over different timelines in response to the same event. Moreover, depending on the severity of the trauma, the resources available during and after the event and individual coping skills, varying numbers of individuals develop persistent symptoms requiring prolonged treatment. Anger may be directed at caregivers and community leaders if these factors are not sufficiently accounted for in medical and psychological response plans.

Symptoms versus Functioning

Disaster psychiatry addresses the clinical concerns of identified patients and supports interventions to prevent potentially incapacitating morbidity in entire populations. Distress-related symptoms are universally experienced during an evolving disaster, as well as in the midst of combat. Symptoms occurring in persons who are not functionally impaired are, by necessity, a secondary concern.

The label given to clusters of symptoms has political, economic, and research-related significance. However, in the context of ongoing battle or relief efforts, the perception that one is ill can itself lead to impaired functioning, which may compromise these efforts and may increase long-term disability. Indeed, transient and nearly universal symptoms that represent normal responses can become medicalized if health care personnel reinforce a view that these symptoms constitute a disease. Therefore, initial mental health interventions must focus on mobilizing effective functioning to allow rescue and relief efforts or battle to continue.

Psychological and Behavioral Syndromes in the Aftermath of Disasters, Terrorism, and War

A considerable body of literature exists on the effects of traumatic events on individuals and communities in response to disasters and war. Many individuals emerge with limited or no adverse effects, effectively resuming their social and occupational roles (resilience). Some may even experience an increased sense of competence, self-efficacy and belief in their ability to manage future stressors (sometimes termed "posttraumatic growth"). However, a sizable minority will experience a range of adverse psychological and behavioral effects, including distress reactions, health risk behaviors, and psychiatric disorders.

In the immediate aftermath of a disasters, distress reactions predominate. Individuals feel a sense of vulnerability and often engage in blaming, scapegoating and expressions of anger at organizations and leaders perceived as responsible. Demoralization and a loss of faith may also occur. Many individuals experience insomnia, irritability and feelings of distractibility. Some present to healthcare settings with physical symptoms as a manifestation of psychological distress. Symptoms such as headache, dizziness, nausea, fatigue and weakness are common in the wake of a disaster, even when an identifiable physical disorder cannot be found. These are normal reactions to an extraordinary event. Planning for these distress reactions requires ensuring adequate resources to respond to individuals with distress symptoms in a timely and supportive manner and triage at emergency and other healthcare settings to enable management of other physical and mental disorders.

Much of disaster psychiatry focuses on the neurobehavioral responses to stressful emotional events. Well-defined psychiatric illnesses certainly result from such exposures. The DSM-5 definition of PTSD necessitates direct or indirect exposure to death, threat to the physical integrity of self or others, or aversive details about an event. PTSD includes (1) intrusion symptoms, such as involuntary memories, traumatic nightmares, or dissociative reactions; (2) avoidance manifested by persistent efforts to avoid trauma-related thoughts, feelings, or external reminders such as people, places, or activities; (3) negative alternations in cognition and mood, including inability to recall key features of the traumatic event, persistent negative beliefs about oneself and the world, distorted beliefs, diminished interest in activities and constricted affect; and (4) alterations in arousal and reactivity demonstrated by irritability or aggression, self-destructive behavior, hypervigilance, concentration problems, and sleep disturbance. When such symptoms result in functional impairment and persist for more than 1 month, PTSD is diagnosed. Transient symptoms occurring during the first month and resulting in dysfunction are diagnosed as *Trauma and Stressor-Related Disorder Not Otherwise*

Specified or Adjustment Disorder. War, disaster, or a terrorist attack often results in exposure to stressors meeting the criterion for PTSD. PTSD is conceptualized as a response to catastrophic trauma or stressor, has been as carefully defined as many other psychiatric disorders, and has target symptoms that military and disaster psychiatric preventive and evidence-based treatments seek to mitigate.

Table 31.5–1.
Psychological Responses to Disaster and War: Associated Mental Disorders and Symptoms

Mental disorders
Acute stress disorder
Posttraumatic stress disorder
Generalized anxiety disorder
Panic disorder
Major depressive disorder
Brief psychotic episode
Adjustment disorders
Alcohol abuse and dependence
Substance abuse and dependence
Symptoms
Anger and irritability
Fear
Restlessness
Concentration and attention difficulties
Sadness
Insomnia (with or without nightmares)
Somatic complaints
Headaches, gastrointestinal distress, musculoskeletal pain
Exacerbation of existing medical illness
Increased alcohol and tobacco use
Diminished interest in activities

The constellation of symptoms, severity and refractoriness of symptoms to conservative measures distinguishes between subsyndromal responses and psychiatric disorders.

In addition to PTSD, other disabling responses to trauma may occur (Table 31.5–1). Adjustment disorders (characterized by anxiety, depression, and changes in behavior) can occur in response to the stressors associated with war, terrorism, or disaster. Conversion disorders are characterized by unexplained deficits in sensory or motor function (e.g., sudden paralysis of an extremity) without demonstrable neuroanatomical injury. Bereavement may occur in response to the death of a loved one as a result of war, disaster, or other catastrophe. Other observed distress-related responses are also seen, including anxiety and depressive syndromes and antisocial behavior (including criminal behavior, military misconduct, and war-related atrocities). In addition, alterations in health-related behavior (e.g., increased use of alcohol, tobacco, or drugs; decreased exercise;

deterioration in eating habits) are not uncommon after traumatic events.

Combat and Operational Stress and Medically Unexplained Physical Symptoms

Besides DSM-5 conditions, in the practice of military and disaster psychiatry, *combat and operational stress* (previously termed 'battle fatigue') is also significant. These terms encompass a variety of somatic and emotional responses to disaster and military operations that do not fully match the specific constellations of symptoms seen in major depressive disorder, adjustment disorder, PTSD, or conversion disorder. Symptoms such as gastrointestinal (GI) distress, tremulousness, and perceptual disturbance, including depersonalization and derealization, occur transiently in response to the traumatic exposure, sleep deprivation, loss of social supports, or any combination of these stressful factors associated with military operations. Minor injury, parasitic infection, starvation, heat exhaustion, or cold injury may also contribute to a decreased ability to cope and, therefore, to operational stress symptoms.

Finally, war historians and military and civilian physicians have observed that unexplained physical syndromes are a common sequelae of combat. Syndromes such as *soldier's heart* and a variety of physical symptoms (arthralgias, rashes, and dysesthesias) have been attributed variously to undetected neurological or chemical attacks, Agent Orange, depleted uranium, or other toxic factors on the battlefield. Although debates among scientists, clinicians, veterans, and journalists persist, the consistent appearance of these syndromes—affecting various persons in various ways—after wars or other military operations points to the likelihood that psychosocial factors contribute to etiology.

A 28-year-old Army sergeant was involved in US military operations in Afghanistan. His commander escorted him to the troop medical clinic because "he just isn't acting right." On interview, he reported that he had been involved in close combat with members of the Taliban for a period of 5 days in the nearby caves. His temperature was 100°F and his pulse was 90 beats per minute; his blood pressure and respiratory rate were normal. His physical examination was unremarkable. His mental status examination was notable for poor grooming; he was unshaven and smelled of perspiration. He was oriented to month and year but unsure of the day and date. He displayed poor eye contact and psychomotor retardation. His thought processes were logical, but he asked for several questions to be repeated. There was no evidence of psychosis, but he seemed focused on his perceived need to "get back to my men." He noted that he and others had been under machine gun fire for "a few days, I guess," and that members of his unit were wounded, and two were killed as a result of a grenade blast. He reported these events with little display of emotion. He denied any history of psychiatric illness or contact with mental health professionals and denied any such history in his family.

DISCUSSION

This service member's mental status examination indicated the possibility of depression. His recent life circumstances suggested an adjustment disorder or acute stress disorder. However, the presentation was also consistent with battle fatigue. His vital signs (as well as recent circumstances) indicated that dehydration may have been contributing to his presentation. Infectious diseases endemic to Afghanistan could also have accounted for these observations. His report of proximity to a hand-grenade explosion suggested the possibility of blast-related

head injury as well. The information presented was not sufficient to render a psychiatric diagnosis but suggested avenues that must be addressed as the service member received psychological first aid, including rest, rehydration, and respite from the ongoing battle. Further history from other members of his unit might have helped clarify the extent to which the service member had been involved in previous combat and his degree of resiliency to these exposures. Although a negative workup for infectious disease, head injury, or other medical conditions and a rapid return to baseline function with conservative measures (respite, nourishment) would suggest battle fatigue and would permit return to duty, the possibility of future emotional sequelae from this event necessitated continued surveillance.

ETIOLOGY AND EPIDEMIOLOGY

The mental health effects of disaster have been historically well recognized. Epidemiological data derived from methodologically sound studies in the aftermath of disaster have been limited—understandably, given the research challenges imposed by the chaotic environment of disasters. However, studies conducted in Manhattan and Washington, D.C., in the months after the September 11, 2001, terrorist attacks demonstrated a substantial burden of PTSD and depression, as well as increased use of alcohol, tobacco, and marijuana. Rates of PTSD ranged from 7 percent to more than 17 percent, and rates of depression ranged from 9 to 15 percent. In one survey, nearly 29 percent of persons reported an increased use of alcohol, cigarettes, or marijuana. Studies of the terrorist bombing of the Oklahoma City Federal Building demonstrated that nearly 35 percent of the directly exposed population developed PTSD within 6 months of the event. Of importance, nearly 40 percent of those with PTSD had no history of a psychiatric disorder. The National Comorbidity Replication Study Survey established a prehurricane prevalence rate for mental disorders of 15 percent in the region affected by Hurricane Katrina, whereas a posthurricane survey of the same population effectively demonstrated a doubling of overall rates of mental disorder to 30 percent at 6 months after the hurricane. A systematic review of suicidality following a wide range of international and US natural disaster events revealed a trend toward decreased suicidality in the first few months following a disaster, during the "Honeymoon" phase. However, more notable were numerous studies that observed increased rates of suicidal thinking and behavior, which commonly began during the 6- to 12-month period following a disaster event, coinciding with the "Disillusionment" phase. Consistent with these finding, when compared to predisaster data, surveys given to greater than 1,000 affected residents 9 to 18 months following Hurricane Katrina found an increase in suicidal ideation (6.4 vs. 2.8 percent) and suicide plans (2.5 vs. 1.0).

Historically, the suicide rate in the US military has been below that in the civilian population, but it has climbed steadily since the beginning of the Iraq and Afghanistan conflicts to the point where suicide is presently the second leading cause of death behind combat deaths (Armed Forces Health Surveillance Center, 2012) and has exceeded demographically

matched civilian rates since 2008. The rise in US Army suicides has persisted despite substantial efforts to publicize and encourage use of mental health services. The Department of the Army has responded to these trends by entering into an agreement with the National Institute of Mental Health (NIMH) to fund the Army Study to Assess Risk and Resilience in Service members (Army STARRS; http://www.armystarrs.org), a multicomponent epidemiological and neurobiological study of risk and resilience factors for suicidality and its psychopathological correlates among Army personnel. Army STARRS is the largest study ever to address risk and protective factors for suicide in the US Military. Initial analyses have shown that the suicide death rate increased for those presently deployed, those never deployed, and those previously deployed. Also, those presently deployed or previously deployed were at greater suicide risk and being demoted in the past 2 years was associated with increased risk. The study is an example of and model for large team science to address critical behavioral health issues.

A 2002 study of 2,583 veterans aged 18 to 54 years who received standardized psychiatric interviews in the National Comorbidity Survey demonstrated that combat exposure resulted in high rates of psychiatric diagnoses and psychosocial difficulties: 28 percent had PTSD; 21 percent engaged in spousal or partner abuse; 12 percent experienced loss of employment (9 percent unemployed at the time of the survey); 8 percent reported substance abuse problems within 1 year after combat; and 7 percent reported symptoms consistent with major depressive disorder. A 2004 anonymous survey administered to US servicemen either before or 2 to 4 months after deployment to Iraq (n = 2,530) or Afghanistan (n = 3,671) suggested that a significantly higher percentage met screening criteria for major depression, generalized anxiety, or PTSD (15.6 to 17.1 percent) after deployment to Iraq than they did after deployment in Afghanistan (11.2 percent). The study also showed that more than one-third of Iraq War veterans accessed mental health services in the year after returning home and suggested that concerns over stigma and potential career implications prevented others from doing so. The incidence and prevalence of intermediate and long-term effects of combat are being clarified, and the neurobiological and psychosocial explanations for observed psychopathology are also being identified.

Neurobiological Factors

The clinical, emotional, and behavioral responses to trauma are rooted in a combination of social, autonomic, and voluntary mechanisms increasingly being understood at the molecular level. In the immediate phase of the stress response, the release of corticotropin-releasing factor (CRF), the secretion of adrenocorticotropic hormone (ACTH), the surge of peripheral catecholamines, and the activation of cortical brain areas related to

perception of threat accompany the extremes of environmental stress. Changes in behavior and cognition correlate with these noradrenergic phenomena. The immediate effect of acute stress under most conditions is improved performance. However, as the capacity to act becomes inadequate to meet the continued demands, the risk for cognitive dysfunction increases, and behavior often becomes excessively narrowly focused. An aroused but focused state results in difficulty scanning for alternatives, shifting sets, or changing plans of action (e.g., adapting). If extreme distress disrupts cognition to the point of creating chaotic thinking, the overfocused fight-flight-freeze response may result in immobility.

The immediate alarm response is followed by a cascade of neuronal and intracellular events, leading to elevated levels of CRH, increased synthesis of cortisol-related receptors, and activation of protein synthesis in subcortical nuclei of the amygdala that are responsible for the encoding of emotionally laden memories and the development of conditioned responses or habits. Hypersecretion of epinephrine seems to exaggerate and consolidate fear-related memories; NMDA receptor functional activity within the amygdala also plays a critical role in this consolidation, as well as the extinction of conditioned fear responses necessary for learning to forget conditioned responses and critical to recovery. Although increased neuronal synthetic activity (neuroplasticity) may serve an adaptive role in preparing for future threat by establishing a memory of trauma, it may also describe the molecular foundation of the intrusive thoughts, exaggerated startle response, and general state of hyperarousal observed as clinically significant pathology in PTSD.

It is less clear how this cascade of neuronal events may relate to the social withdrawal or depressive symptoms seen in individuals exposed to trauma (in conjunction with PTSD or acute stress disorder or independent of these syndromes). Social withdrawal is a major contributor to the pathological consequences of traumatic experience; perhaps the threat avoidance response becomes overgeneralized in withdrawn individuals. One factor that seems to mitigate the psychopathological effects of trauma is the availability of social supports. The increased neuroplasticity immediately following the recognition of a threat may provide a rationale for rapid intervention in the form of social support. However, empirical studies are needed to demonstrate the efficacy of psychosocial interventions in subsyndromal or asymptomatic populations.

Predisposing Factors

PTSD, other anxiety and depressive disorders, and physical syndromes in response to a given stressor are more frequently diagnosed among women than men. Explanations for this involve neurobiological and psychosocial factors, including the greater rate at which women seek treatment for

stress-related symptoms and the fact that the duration of illness may be longer for women and, therefore, may be more likely reach clinical attention. In the aftermath of a disaster or during war, women are at higher risk for victimization through intimate partner violence as well as physical and sexual assault outside the home. Women may also be at increased risk as a result of having greater responsibility for managing family emotions and distress. Men are at greater risk for posttraumatic problems with substance abuse and antisocial and violent behavior. Gender-specific cultural factors and as-yet-unspecified neurophysiological factors are likewise implicated in these differences.

The level of functional impairment in the aftermath of combat or disaster is also related to pretrauma functioning. Individuals who function marginally in social or occupational roles before combat or disaster exposure are at increased risk of poor function after trauma exposure. A history of childhood abuse or significant traumatic exposure generally appears to be a risk factor for the development of PTSD after exposure in adulthood. If persons have successfully negotiated disaster or combat situations in the past, they may be less traumatized by subsequent exposures (even if the circumstance of the trauma changes from combat to disaster or vice versa). On the contrary, if past traumatic experiences produced PTSD or other psychiatric syndromes, further traumatic exposures may trigger exacerbations or the emergence of new psychiatric disorders. Like depression, PTSD appears to be related to serotonergic dysfunction, and genetic factors (i.e., the presence of the short/short genotypes of serotonin transporter protein) may determine the degree to which an environmental exposure results in pathology or resilience. Finally, specific subpopulations, such as children and those suffering from chronic mental illness may be particularly vulnerable to the emotional consequences of traumatic experience. Elderly persons with cognitively impaired or limited mobility are less able to respond to the requirements of a disaster environment, including movement, evacuation, and seeking medical care, and are generally at increased risk for psychological injury. The loss of social supports after disasters and in war-torn societies also results in the need for assistance in others who may have been previously able to cope (Figs. 31.5–1 and 31.5–2). The loss of access to medical care (i.e., medications), as well as health care providers, is an indirect but substantial cause of casualties after disasters and terrorism.

Precipitating Factors

Precipitating factors are the proximate circumstances that initiate the various sequelae of trauma. For disaster responders and military populations, disaster and combat deployments and peacekeeping missions disrupt families and are often poorly timed with regard to other life events. The high intensity and long duration of disaster or combat exposure

increase the likelihood of psychiatric casualties. Specific experiences, such as physical injury or witnessing grotesque deaths, torture, or other atrocities, place individuals at increased risk for adverse mental health consequences. Victimization in the form of rape, harassment, or assault can precipitate significant distress or illness in those victimized. Sexual assault is a potent precipitant of adverse neurobehavioral changes.

FIGURE 31.5–1. Children are particularly vulnerable to the emotional consequences of disaster.

Mitigating or Perpetuating Factors

Ongoing factors, including the security and safety of recovery environments, the extent of secondary traumatization, and—in first responder and military populations—rotation schedules, the degree of recognition or compensation for efforts, and one's belief in the mission all affect the rate and severity of psychiatric illness and distress symptoms. Personal attributes or behaviors, such as overdedication to the task at hand, may further contribute to the development of dysfunction. Symptoms in victims of disaster and war may be mitigated or exacerbated by perceptions of community leadership's preparedness for disaster, response to crisis, recognition of heroes, and the distribution of resources such as the provision of medical, financial, or emotional assistance immediately after the crisis and over time.

FIGURE 31.5–2. Separation from social support networks contributes to stress reactions in the context of military operations.

Nongovernmental organizations (NGOs), such as the American Red Cross and the Salvation Army, help minimize the stress after a disaster. By attending to basic human needs, such as food, clothing, and shelter, they reduce the psychological and physiological effects of the event. In recent years, the American Red Cross has developed training for volunteer health care workers to recognize, minimize, and treat stress responses in disaster workers and victims of disaster on the basis of the principles of psychological first aid (PFA).

Protective Factors

In groups such as military units, police or firefighters, and paramedic

rescue units, protective factors may diminish the potential sequelae of trauma. Loyalty and interpersonal cohesion among members of such organizations may provide immediate access to social supports during the acute experience of a disaster and its aftermath. When such organizations possess strong leaders, these persons create and foster an environment of strong loyalty, mutual trust, and respect. Such an environment may reduce the extent to which one member of the group becomes excessively critical of another's performance during an emergency situation, may contribute to voluntary participation in high-risk rescue-and-recovery or combat missions, and may facilitate the synchronization of individual efforts that is necessary for success in these types of operations.

An important factor in the development of such trust and synchronization is the number of demanding, specific, and repetitive overtraining exercises in which elite police, rescue-and-recovery, and combat units participate. If the level of training is high, individuals in such units rely on ingrained principles aimed at mutual support in a quest for mission success. Recently, increasing numbers of disaster responders (firefighters, police, physicians, and civic leaders) have assembled to train for responding to terrorist attacks, as well as natural disasters. Government emergency preparedness agencies, such as the US Federal Emergency Management Agency, are increasingly coordinating such training. In addition, the National Center for Disaster Medicine and Public Health published disaster behavioral health curriculum recommendations that provide a comprehensive overarching framework for the development of training in disaster mental health preparation and response.

The nature and extent of the fit between persons, equipment—such as chemical protective suits and masks, remote-controlled bomb detonators, or mine detectors—and the environmental extremes that may be experienced during disaster (e.g., heat, cold, and sleep deprivation) modify stress responses in individuals and communities. The fit between pilot and aircraft, as well as pilot, aircraft, and other members of the aircrew, may be improved by specific training as well. The extent to which subtler factors within the environment may modify stress response has been noted in those forced to exist in close quarters with limited contact with the outside world, such as those aboard submarines or ships. Although these factors cannot be eliminated from the practice of military operations or disaster response efforts, the extent to which these additional stressors exacerbate responses to other primary stressors may be reduced through training.

MANAGEMENT AND CARE DELIVERY

General Principles

Disasters, well-orchestrated terrorist attacks, and military conflicts shatter the expectation of a just and safe world within populations in which such

notions of basic safety and justice are cultural norms. Because a lack of emotional preparedness makes chaos and disorder more likely after such events, primary prevention must begin with a recognition of the need for rapid dissemination of information regarding the threat (e.g., early warnings for tornados and hurricanes or reporting of anticipated terrorist attacks gleaned from reliable intelligence sources), followed by the development of a planned response, the practice of that response, and the provision of funding and facilities in support of the plan. All of these steps mitigate a community's sense of helplessness before such an event occurs and increase a community's capacity to provide social support to victims.

Secondary prevention (i.e., early intervention) begins with the initial triage and treatment of victims. First responders, medical care personnel, and mental health care providers must be prepared to provide psychoeducation and symptomatic relief for individuals with acute (and frequently transient) emotional or behavioral disorganization or other symptoms. The creation of a *holding environment*—a location or locations in which such symptoms can be observed and monitored—is important. This area should be sufficiently removed so as not to disrupt the ongoing triage and stabilization of life-threatening physical injuries but should be close enough to allow reevaluation and further medical intervention should symptoms worsen. A holding environment may also prove to be a location in which affected individuals attend to such basic needs as food, shelter, and personal hygiene, which are critical to restoring a sense of normalcy after a disaster or traumatic event (Table 31.5–2). For military personnel, maintaining individuals in close proximity to other unit members can aid in symptom resolution and return to normal function.

Initially, establishing a safe environment and managing life-threatening physical injury and the myriad disease possibilities that result from lack of food, potable water, or infection are the most important psychiatric interventions. Subsequently, identifying high-risk populations, such as disaster workers, persons at impact zones, children, and the chronically ill, can focus further interventions. Outreach and public education strategies are critical because disaster and trauma victims rarely specifically seek mental health care. It is important to educate medical and community groups about the normal and predictable responses to traumatic events of disaster and war, as well as when to refer for additional mental health treatment. Informing community and governmental leaders about the behavioral and emotional sequelae associated with disasters ensures the allocation of funds and resources for these phenomena in the disaster recovery plan.

Table 31.5–2.
Disaster Mental Health Response: Management Principles

Primary interventions
- Early warning systems
- Development, rehearsal, and resource allocation for event response, including identification and roles of responders, methodology for information dissemination, potential locations for triage, initial treatment, and holding environments

Secondary interventions
- Initial triage and identification of at-risk individuals or subpopulations
- Creation of holding environment, ensuring that needs for food, shelter, and hygiene are met and allowing for close observation and medication management of acutely agitated individuals
- Brief support and psychoeducation, application/implementation of the principles of psychological first aid

Tertiary interventions
- Medication management
- Cognitive-behavioral therapy
- Population surveillance for further development of symptoms or disorders

The responsibility for preventive measures and the initial recognition and treatment of the psychological consequences of disaster cannot be limited to available psychiatrists. General physicians, psychologists, and other social scientists must use their diverse skills to care for the victims of disasters and wars. It may be necessary for well-trained allied professionals to make preliminary diagnoses and to provide initial psychoeducation and consultation to medical and surgical colleagues and other first responders. Negative findings regarding the effectiveness of critical incident stress debriefing (CISD) have prompted a renewed search for acute public health interventions in the aftermath of disaster. PFA has been developed as a concept to organize evidence-based and informed interventions in the acute phase of trauma and disaster exposure. A manual for PFA was developed by the National Center for PTSD and the National Child Traumatic Stress Network for use by mental health professionals. However, since its Internet release in the aftermath of Hurricane Katrina, other organizations have developed adaptations of the manual and its principles for use by a variety of first responders, including school crisis teams, clergy, emergency medical personnel, and military rapid deployment teams. PFA is an evidence-informed process that operationalizes the aspects of social support, psychoeducation, stress management, and cognitive restructuring across the domain of short-term and intermediate mental health response supported by empirical evidence. PFA aims to establish and enhance the sense of safety, calmness, and hope (or optimism) and also to foster community self-efficacy and social connectedness in the aftermath of traumatic exposure. These “essential elements” are associated with improved psychosocial function after mass violence.

Psychiatrists necessarily devote attention to those with more severe initial presentations and to those who develop more chronic conditions despite preventive interventions.

Therapeutic Interventions

Symptoms ranging from lethargy and depression to disorientation, dissociation, hallucinations, paranoia, and cognitive dysfunction may be observed in the immediate aftermath of a disaster. The diversity of response is complicated by the fact that symptoms may result purely as an emotional response, as a consequence of a neurological injury from a blast or other blunt trauma, or, in the case of terrorist attack, from the neurotoxic effects associated with specific chemical or biological agents (Tables 31.5–3 and 31.5–4). Rapid identification of the underlying infectious, chemical, or physical insult to the brain directs the treatment strategy. Regardless of etiology, delirium must be addressed rapidly because mortality from all underlying causes increases with untreated delirium. Moreover, acutely agitated patients place exceptional demands on treatment staff that would be otherwise available to deliver group-related interventions and education.

Antipsychotic and anxiolytic medications used in the acute management of psychosis are effective in delirium due to infectious causes. However, their metabolism by way of the hepatic cytochrome P450 system may alter levels of antimicrobial agents and acetylcholinesterase inhibitors used in the treatment of biological or chemical terrorist attack victims. The potential for drug–drug interactions bears on the selection of antibiotics, chemical antidotes, and antipsychotic agents in specific cases. Dose-related side effects of antipsychotic agents (akathisia, somnolence) may also be mistaken for primary symptoms of an infectious, postphysical trauma or chemically mediated encephalopathy. The importance of keeping treatment as simple as possible under these confusing circumstances cannot be overemphasized. Therefore, a conservative approach that minimizes the initial use of psychotropic medications is warranted. Although antipsychotics or anxiolytics may be necessary to control acutely disruptive or dangerous behavior, the pharmacological suppression of acute anxiety and arousal symptoms does prevent the development of PTSD. Short-acting hypnotics may be used to restore normal sleep patterns in cases in which insomnia appears as the primary cause of dysfunction. Although supplies of psychotropic medications are less likely to be depleted than stores of vaccines or antibiotics, triage to an environment that favors observation and education may allow for the most judicious distribution of potentially scarce medications. If the diagnosis of PTSD is clear and follow-up plans can be coordinated, selective serotonin reuptake inhibitors are an important element in the treatment of PTSD. Other antidepressants and anxiolytics, as well as atypical neuroleptics, are regarded as second-line or augmentation strategies. Although not approved by the FDA for PTSD, a number or studies have demonstrated the efficacy of prazosin, an α1-antagonist antihypertensive agent, in alleviating symptoms of PTSD, particularly sleep disturbance and trauma-related nightmares.

Table 31.5–3.
Neuropsychiatric Syndromes or Symptoms in Selected Biological Agents

Biological Agent	Syndrome or Symptom(s)	Comment
Anthrax	Meningitis	May be rapidly progressive
Brucellosis	Depression, irritability, headaches	Fatalities associated with central nervous system involvement
Q fever	Malaise, fatigue	In one-third of patients
	Encephalitis, hallucinations	In advanced cases
Botulinum toxin	Depression	Due to lengthy recovery
Viral encephalitides	Depression, cognitive impairment	Other mood changes are also possible
All biological agents	Delirium	Acutely impaired attention and memory, and perceptual disturbances

Table 31.5–4.
Neuropsychiatric Effects of Chemical Agents

Organophosphates (nerve agents such as sarin, soman, VX, and tabun)	Depression, sleep disturbance, and impaired cognition; poor judgment, attention, and short-term memory; poor concentration, confusion, drowsiness, delirium, irritability, insomnia with nightmares, excessive dreaming, and psychosis
Atropine (treatment for organophosphates)	Blurred vision, tachycardia, dry mouth, sweat suppression, urinary retention, cognitive impairment, psychosis, and delirium (disorientation, hallucinations, and delusions)
Cyanide	Anxiety, confusion, giddiness, and hyperventilation
Blistering agents, mustard gas, and phosgene	No direct acute effects; inactivation of charcoal may result in sensations of suffocation or anxiety; psychic distress over disfigurement
Mycotoxins (fungal toxins)	Psychosis and somatic complaints, involuntary movements, anxiety, and agitation

Beyond pharmacological intervention, a number of psychosocial interventions have been described as helping in the prevention of PTSD or other neurobehavioral responses to disaster, often, however, without evidence-based support. Group-debriefing techniques and CISDs have been used in the aftermath of school shootings, natural disasters, and terrorist events. However, there is no convincing evidence that debriefings reduce the incidence of PTSD, and at least one study suggested that these interventions may be harmful. In addition, group meetings of individuals with greatly different exposures may actually increase the exposure of some and, therefore, increase symptoms. Much like an analgesic medication, debriefing may have effects to relieve pain, restore function, and limit disability but not to prevent disease or disorder. On the contrary, for response teams, ongoing and frank discussions among care providers and emergency responders can open lines of communication to coordinate and to evaluate the efficacy of response and foster cohesion and group

understanding of the unfolding event. This may serve to sustain the performance of persons critical to the management of the disaster, decrease individual isolation, and help identify team members who may require further psychiatric or other mental health attention.

Clinical trials indicate that cognitive-behavioral psychotherapeutic approaches to anxiety, social withdrawal, depressive symptomatology, and hyperarousal are effective. Clinical trials suggest that even brief cognitive and behavioral psychotherapy interventions may reduce immediate symptoms of depression, anxiety, and PTSD and may diminish the development of long-term morbidity. Exposure-based cognitive therapies (including cognitive processing therapy [CPT] and virtual reality therapy [VRT]) are under active investigation in disaster military populations. CPT uses a patient-created written narrative of his or her traumatic experience as the basis for reorganizing cognitions about the meaning of the event. VRT augments the patient's narrative or imagined reexposure to traumatic environments with video game technology. The efficacy of these approaches as preventive measures has not been established. Because medical resources may be quickly overwhelmed in a large-scale disaster or combat situation, nonphysicians, provided that they are appropriately trained in the delivery of these therapies, allow for more effective delivery of care. Other treatment approaches and medications become more important as psychiatric disorders become more clearly defined, chronic, and persistently impact the life events of individual victims. Some small-scale studies indicate that Internet-based CBT, with human contact and coordination, may be an effective intervention for large-scale events and populations with a high need for services for PTSD.

Three weeks after a train derailment, a 42-year-old budget analyst presented to the mental health clinic. He noted that he was embarrassed to seek care, as he was previously a firefighter, but he felt he that he needed "some reassurance that what I'm experiencing is normal." He reported that, since the wreck, he had been feeling nervous and on edge. He experienced some difficulty focusing his attention at work, and he had occasional intrusive recollections of "the way the ground just shook; the tremendous 'bang' and then the screaming when the train rolled over." He noted that he has spoken with five business colleagues who were also on the train, and three acknowledged similar symptoms. However, they said that they were improving. He was more concerned about the frequency of tearful episodes, sometimes brought on by hearing the name of a severely injured friend, but, at other times, occurring "for no particular reason." In addition, he noted that, when he evacuated the train, rescue workers gave him explicit directions about where to report, and, although he complied, he now felt extremely guilty about not returning to the train to assist in the rescue of others. He reported a modest decrease in appetite and denied weight loss but noted that he had stopped jogging during his lunch break. He had difficulty initiating sleep, so he had begun consuming a "glass or two" of wine before bed to help with this. He did not feel rested on awakening. He denied suicidal ideation or any psychotic symptoms. His sister took an antidepressant several years ago, but he did not desire medication. He feared that side effects could further diminish his ability to function at the workplace and could cause him to gain weight.

DISCUSSION

Human-caused disasters, like natural disasters or war, may precipitate mental illness.

Although this individual reported reexperiencing phenomena and anxiety, which often follow a traumatic experience and that acute stress disorder, he also reported neurovegetative symptoms and guilt, which are more consistent with major depressive disorder. His report also suggested a family history of depression that predated the trauma. Given the patient's reticence to take medication, brief supportive and cognitive-behavioral therapeutic approaches are appropriate. Once rapport and trust are established, it may be possible to advise potential benefits (and favorable side-effect profiles) of antidepressant therapy, if somatic therapy seems necessary.

Military Mental Health Care in Peacetime

When not directly involved in combat operations, peacekeeping, or humanitarian relief, military psychiatric practice is, in many ways, similar to civilian practice except for the fact that military psychiatrists go to work in military uniforms. The US Department of Defense has tertiary care hospitals throughout the United States, Europe, and Asia, and these institutions all have outpatient psychiatry or interdisciplinary behavioral health clinics or both, and inpatient wards for severely mentally ill military members, as well as their family members. Some military psychiatrists receive their undergraduate medical and residency training at civilian institutions. However, the US Congress in 1975 authorized the Uniformed Services University of the Health Sciences (USUHS) to provide medical education and produce physicians for military service. USUHS provides a 4-year medical degree program and a number of graduate degree programs in the basic and clinical sciences. The USUHS Center for the Study of Traumatic Stress conducts research and consults communities and federal and international agencies on matters surrounding individual and community responses to trauma, disaster, and war. Japan, Great Britain, and Russia are among nations with institutions that teach military-specific curriculum to military medical care providers. As in other nations, these countries also call to national duty physicians not specifically trained in military institutions during times of war or crisis.

The US Department of Defense has residency programs in psychiatry at tertiary care Army, Air Force, and Navy Medical Centers within the United States. These institutions are mandated by Congress to provide military-unique curriculum to ensure that the physician in training will be able to provide care in combat or disaster environments. This type of training is not necessarily provided in civilian programs. Military training programs also introduce physicians to aspects of the military lifestyle (e.g., frequent moves and separations from family, living and working overseas, special occupational duties, and requirements for maintaining security over classified information) that affect military members and families, even when service members are not deployed. In addition, the USUHS offers the nation's only Disaster Psychiatry Fellowship in collaboration with the Center for the Study of Traumatic Stress.

The US military has attempted to decrease the incidence and severity of

combat and operationally induced psychiatric disorders through a number of mechanisms. Mental health teams are now routinely assigned to US forces in combat and are deployed in operations other than war. Each branch of the US military services has specialized rapid intervention teams to provide consultative assistance and acute treatment as necessary to units that have experienced traumatic events. These teams educate unit commanders on likely behavioral responses to stressful events and recommend leadership interventions that may reduce negative stress responses. Military mental health units have deployed in conjunction with peacekeeping forces to the Balkans and to Kosovo and in the aftermath of the terrorist attacks on the Pentagon and the World Trade Center. Although debriefings have been offered to exposed units requesting such intervention, increasingly, the focus has been on psychoeducation, PFA, and buddy care. In addition, through a highly visible presence, attention has focused on improved access to care for those who might not otherwise request assistance and destigmatization of those who seek mental health care. Military leadership has strongly supported the increased use of these units in an effort not only to establish immediate access to care but also to study the potential efficacy of such interventions in reducing the incidence of PTSD and somatic symptom syndromes that may be tied to psychological responses to combat. These teams have also provided consultation and intervention in military humanitarian relief efforts after hurricanes, floods, infectious disease outbreaks, and other natural disasters within the United States and abroad.

Different mission types, patterns of deployment, and medical support systems among the various branches of the US military pose major difficulties for the development of a unified tri-service approach to the management of operational stress. Armies typically deploy large units for extensive periods of time and allocate large amounts of medical assets to support these units. This medical support includes specialty services. The Navy and Marine Corps deploy smaller units at sea and ashore. General medical officers and nonphysician providers furnish medical support, and specialty care is not routinely available in the operational theater. The Air Force has short- and long-range missions. In addition to direct patient care, military psychiatrists provide consultation to community leaders and civilian physicians not accustomed to responding to large-scale physical and emotional traumas.

In the United States, definitive treatment of psychiatric illness is often provided in the military's system of hospitals. Medical care is provided to active duty personnel and their families. Other mental health specialists, nurses, social workers, and psychologists augment this care. Military members who develop psychiatric disorders while on active duty are eligible for medical retirement disability pay and continued treatment through a system of Veterans Administration hospitals. Under present

regulations, some individuals may be separated from service because of personality disorders that existed prior to service entry without disability payment or ongoing medical care from the military.

Other nations with recent wartime experience, such as Israel and Croatia, have developed programs to evaluate and to treat soldiers and civilians exposed to combat. Their experiences are somewhat different from experiences in the United States because they rely much more heavily on reserve forces. These nations also have a more inclusive social medical infrastructure; therefore, treatment programs are less reliant on the military medical system. Other nations are increasingly confronted with the management of operational stress in peacekeeping and humanitarian missions. Asian nations that have recently experienced natural disasters and terrorist events are also studying approaches to evaluating and treating individuals exposed to trauma.

FUTURE CHALLENGES AND EVOLVING ISSUES

As political, social, scientific, and technological factors evolve, societies will change their responses to the consequences of disasters and wars. Psychiatric practice associated with wars and disaster has changed with the evolution of scientific understanding of illness. In the future, the resources to deal with the consequences of disaster or war and the relative importance assigned to dealing with the resultant injuries and disabilities are likely to be influenced by political and sociocultural values, as well as advances in knowledge of disease processes, treatment paradigms, and technology. Disaster and military psychiatry must anticipate these changes to ensure that current trainees are prepared to deliver care in a changing environment.

Emerging Threats

Reports of animal-to-human transmission of avian influenza have sparked renewed interest in the eventuality of pandemic influenza. Such an occurrence would raise issues related to the availability of limited quantities of medications for prevention and acute treatment and the likelihood of widespread morbidity, death, chaos, disruption, and distress. The necessary implementation of social distancing measures (e.g., quarantine, isolation, closing of schools and workplaces) to control infection will establish significant obstacles to the creation of a holding environment and the fostering of social connectedness critical to public health response in disaster. Along with planning for unprecedented surge in health care demand, alternative mechanisms for the delivery of health care, health risk communication, PFA, and the preservation of workplace function must be identified, resourced, and rehearsed.

Advanced technology has implications for those responding to human-

caused disasters, such as terrorist attacks involving chemical, biological, radiological, and nuclear agents (CBRN). In the United States, the terrorist attacks on the Pentagon and the World Trade Center as well the fatal distribution of anthrax-tainted mail in the fall of 2001 precipitated an ongoing "war" on terrorism. These events intensified efforts to increase coordination among civilian, nonmilitary, governmental, and military agencies for response to future terrorist events within the United States and abroad, as the scope of terrorist activity has become more globalized and technologically advanced. Though post-9/11 concerns about escalating CBRN terrorist attacks went largely unrealized, the use of coordinated multi-site attacks has become an increasing terror threat. Recent such events around the world have resulted in unprecedented levels of community fear and disruption. Clarifying the roles of military and civilian responders and integrating mental health and public health activities in terms of triage, treatment, consultation, and education in any joint responses to crisis are a challenge for military and disaster psychiatrists.

New technologies in combat will also modify the means of sorting and treating persons with medical and psychiatric injuries. Future militaries in technologically advanced nations are likely to become much smaller, move rapidly across the battlefield, use advanced sensors, and direct intense fire across a considerable distance. These capabilities, coupled with the use of weapons of mass destruction, will likely make the battlefield more chaotic and inhospitable to human life. The increasing use of unmanned aerial vehicles in combat operations and disaster reconnaissance may reduce exposure to physical injury but, paradoxically, create a new avenue for psychological and emotional distress as those who remotely operate this technology will have both a close-up and personal experience of human suffering and death, possibly at their own hands. Individuals who gather intelligence information from unmanned aircraft engaged in combat operations are effectively "deployed in place," creating a number of potential challenges that result from being persistently exposed to the stimulus of war without the geographic or psychological distance afforded to traditional combat units. Emergency care and evacuation of those with disease and injury may become increasingly difficult. The inability to maintain contact with rapidly moving units may preclude returning individuals to their original units. Future military casualties may increasingly rely on care by unit buddies, medics, and front-line leaders rather than specialized medical units or specialists at hospitals in the rear. Underdeveloped nations may have limited access to advanced technologies, so that more traditional ways of organizing medical and psychiatric practice may continue to be relevant in certain circumstances.

The evolution of highly mobile units on widely dispersed battlefields will decrease the opportunity for exchanging rested troops from the rear area for those exhausted by front-line combat. Provision of brief respite for

exhausted troops—a hallmark of management of battle fatigue—may become impossible because each individual may be performing a critical specialized task. Small medical units operating within the area of combat are likely to be eliminated from this technology-intensive battlefield. Although treatment may, by necessity, move to the battlefront, medical specialists at the rear may render triage decisions and diagnoses through the use of telemedicine communication technology. Experience has shown that front-line mental health providers take a pragmatic view of acute psychiatric symptoms and tend not to make hasty formal diagnoses on overstressed troops. Rear-echelon providers, by contrast, tend to assign formal psychiatric labels that may be inaccurate and that may stigmatize troops without contributing to treatment. Rear-echelon mental health specialists in future battles must address the challenge of providing useful therapeutic advice from afar while avoiding unnecessary diagnostic stigmatization.

Cultural Issues

Social scientists note that responses to trauma may be considered either normal or pathological, depending on the cultural context. Many have expressed fears that mental health practitioners, motivated by profit, will try to convince individuals experiencing normal uncomfortable responses that they need treatment. Because of the persistent stigma associated with mental health care, a future challenge to military efforts at providing assistance to other nations in crisis may be overcoming the belief that accepting assistance signals weakness.

Most individuals exposed to traumatic events deserve to be reassured that, with return to work, community, and family, they will recover. However, some individuals (and, perhaps, some cultures) will experience greater psychopathological responses and more prolonged symptoms after trauma. The nature of the behaviors and symptoms associated with trauma response across cultures is uncertain. Psychiatrists providing care during combat and humanitarian environments in a cross-cultural environment must do so with an awareness of the unique aspects of how culture influences the experience of mental illness. The extent to which social supports, biological–genetic predisposition, concurrent illnesses, and other political and economic parameters contribute to variability across cultures is not known. Although it appears clear that the severity of the trauma is important, reliable measures of severity remain to be determined. Trauma may, under some conditions, create the opportunity for personal growth as well, and further understanding of this potential must also be exploited to reduce morbidity.

Ethical Challenges

Providing care in the mass casualty situation raises ethical questions about the equitable distribution of resources and the application of moral values. The hypersuggestibility of recently traumatized individuals has, at times, led to the exercise of political influence and the manipulation of loyalties. Governments in trouble have withheld treatment to minority racial or political groups—clearly, an ethical breach. Because governmental terrorism is a common form of terrorism, care providers and leaders must be sensitive to the possibility that disasters can afford unscrupulous leaders an opportunity to manipulate citizens for their own purposes.

Military and combat operations raise challenging ethical dilemmas. To facilitate command assessment of troop health status, militaries have denied members' confidentiality in medical communication. Mental health care providers must strike a balance between a promise of privacy that encourages persons to seek care and responsible reporting to higher command regarding situations that pose danger to larger groups. Thus, dual allegiance to individuals and the larger community presents an ethical challenge that must be negotiated by the military care provider. Persons in extreme circumstances may behave in ways that they later view as shameful. This self-view may contribute to symptom development and at the same time discourage care seeking.

An innate aspect of war involves the collection of military intelligence to enhance the effectiveness of combat operations. Prisoners of war or other detainees are often a source of intelligence information. Commanders may seek involvement of mental health care personnel in the process of managing detainees. Providers will be asked to deliver mental health care to alleviate detainee suffering. They may also be sought out to provide guidance on the use of behavioral interventions to facilitate gathering of intelligence information. Physicians must be aware of ethical prohibitions against participating in cruel, inhumane or coercive practices with prisoners of war or other detainees and of the challenges that emerge when their involvement is intended for purposes other than the direct provision of care. In additional to personal ethical dilemmas that arise, medical and allied professional organizations have provided guidance regarding the appropriateness of these roles.

Disaster triage is frequently carried out in large open areas that allow everyone present to hear what patients say to caregivers. Given the social stigma assigned to the manifestations of psychiatric illness, it is easy to understand patients' reluctance to communicate and doctors' reluctance to inquire. Research on disaster- and war-exposed populations also raises significant ethical issues surrounding balancing requirements to develop knowledge and to help others with the potential exploitation of those in need. Perhaps re-educating the population can reduce ethical and therapeutic problems associated with stigma. However, altering deeply ingrained cultural expectations is just as challenging as providing privacy

in chaotic triage environments.

CONCLUSION

Natural and human-caused disasters, including terrorism and war, result in the disruption of societal function. Wars and acts of terrorism, with their attendant large-scale death, injury, and destruction, affect populations in much the same way as massive natural disasters. However, the intentional, human-caused nature of these events increases their mental health consequences. Disasters and war leave in their wake populations experiencing psychological disturbances that have been described by social scientists, civil leaders, physicians, and other care providers throughout the ages.

The consequences of exposure to disaster and war may take the form of psychological disorders, such as PTSD, or may manifest as various (and sometimes more subtle) forms of behavioral change, anxiety, or depression. Symptoms may present at different times during and after traumatic exposure. Many factors complicate the evaluation and treatment of neuropsychological syndromes in the aftermath of war or disaster. Resources are overwhelmed, life-threatening illnesses require immediate treatment, and psychological casualties are often reluctant to seek assistance.

Progress has been made in identifying the nature of trauma-related psychological responses. Predisposing, exacerbating, and mitigating factors have been identified, as have traumatic, stress-related, neuronal, and intercellular events. The value of multidisciplinary preparation and training for disaster management and the need for outreach programs have also been demonstrated. The US experience in the aftermath of September 11, 2001, the ongoing war on terrorism, and increased frequency of large-scale natural disasters have clarified the need for greater integration among military, government, and civilian organizations charged with the responsibility of mental health response to these events. However, further study must clarify the manner in which psychotropic medications and various forms of psychosocial support and psychotherapy in the treatment of war- and disaster-related morbidity can be delivered by cooperating agencies in times of national need. With technological advances and global economic shifts, the nature of war and other human-caused disasters has changed. Disaster and military mental health care delivery must anticipate future changes to develop improved methods of prevention, evaluation, and care for individuals and groups devastated by war or disaster.

REFERENCES

Assistant Secretary of Preparedness and Response. *Health and Human Services Behavioral Health Concept of Operations*. Available at http://www.phe.gov/abc. Published February 2014.

*Benedek DM, Fullerton C, Ursano RJ. First responders: mental health consequences of natural and human-made disasters for public health and public safety workers. *Annu Rev Public Health*. 2007;28:55–68.

Davidson TM, Price M, McCauley JL, Ruggiero KJ. Disaster impact across cultural groups: comparison of Whites, African Americans, and Latinos. *Am J Community Psychol*. 2013;52(1–2):97–105.

Fergusson DM, Horwood LJ, Boden JM, Mulder RT. Impact of a major disaster on the mental health of a well-studied cohort. *JAMA Psychiatry*. 2014;71(9):1025–1031.

Flynn BW, Morganstein JC. Curriculum recommendations for disaster health professionals: disaster behavioral health. Available at ncdmph.usuhs.edu/Documents/BehavioralHealthRecommendations-201401.pdf. Published January 2014.

Flynn BW, Speier AH. Disaster behavioral health: legal and ethical considerations in a rapidly changing field. *Curr Psychiatry Rep*. 2014;16(8):457.

Fullerton CS, Herberman Mash HB, Benevides KN, Morganstein JC, Ursano RJ. Distress of routine activities and perceived safety associated with post-traumatic stress, depression, and alcohol use: 2002 Washington, DC, sniper attacks. *Disaster Medicine and Public Health Preparedness*. 2015;9(5):509–515. Available at http://doi.org/10.1017/dmp.2015.67

Fullerton CS, Ursano RJ, Liu X, McKibben J, Wang L. Depressive symptom severity and community collective efficacy following the 2004 Florida hurricanes. *PLoS ONE*. 2015. Available at http://doi.org/10.1371/journal.pone.0130863.t002

*Goldman E, Galea S. Mental health consequences of disasters. *Annu Rev Public Health*. 2014;35:169–183.

Hobfoll SE, Watson P, Bell CC, et al. Five essential elements of immediate and mid-term mass trauma intervention: empirical evidence. *Psychiatry*. 2007;70(4):283–315; discussion 316.

Hoge CW, Castro CA, Messer SC, et al. Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *N Engl J Med*. 2004;351(1):13–22.

Kessler RC, Galea S, Gruger NM, Sampson NA, Ursano RJ, Wessely S. Trends in mental illness and suicidality after Hurricane Katina. *Mol Psychiatry*. 2008;13(4):374–384.

Kessler RC, McLaughlin KA, Koenen KC, Petukhova M, Hill ED; WHO Mental Health Survey Consortium. The importance of secondary trauma exposure for post-disaster mental disorder. *Epidemiol Psychiatr Sci*. 2012;21(1):35–45.

Kolves K, Kolves KE, De Leo D. Natural disasters and suicidal behaviors: a systematic literature review. *J Affect Disord*. 2013;146(1):1–14.

La Greca AM, Lai BS, Llabre MM, Silverman WK, Vernberg EM, Prinstein MJ. Children's postdisaster trajectories of PTSD symptoms: predicting chronic disease. *Child Youth Care Forum*. 2013;42(4):351–369.

Morganstein JC, West JC, Huff LA, et al. Psychosocial responses to disaster and exposures: distress reactions, health risk behavior, and mental disorders. In: Shigemura J, Chhem RK, eds. *Mental Health and Social Issues Following a Nuclear Accident*. Tokyo: Springer Japan; 2015:99–117. Available at http://doi.org/10.1007/978-4-431-55699-2_8

*Norris FH, Friedman MJ, Watson PJ, Byrne CM, Diaz E, Kaniasty K. 60,000 disaster victims speak: part I. An empirical review of the empirical literature, 1981–2001. *Psychiatry*.

2002;65(3):207–239.

*North CS, Pfefferbaum B. Mental health response to community disasters: a systematic review. *JAMA*. 2013;310(5):507–518.

North CS, Ringwalt CL, Downs D, Derzon J, Galvin D. Postdisaster course of alcohol use disorders in systematically studied survivors of 10 disasters. *Arch Gen Psychiatry*. 2011;68(2):173–180.

*Pfefferbaum B, Flynn BW, Schonfeld D, et al. The integration of mental and behavioral health into disaster preparedness, response, and recovery. *Disaster Med Public Health Prep*. 2012;6(1):60–66.

Pietrzak RH, Tracy M, Galea S, et al. Resilience in the face of disaster: prevalence and longitudinal course of mental disorders following hurricane Ike. *PLoS One*. 2012;7(6):e38964.

Shuster MA, Stein BD, Jaycox LH, et al. A national survey of stress reactions after the September 11, 2001, terrorist attack. *N Engl J Med*. 2001;345:1507–1512.

Stene LE, Dyb G. Health service utilization after terrorism: a longitudinal study of survivors of the 2011 Utøya attack in Norway. *BMC Health Services Research*. 2015;15:158. Available at http://doi.org/10.1186/s12913-015-0811-6

Stoddard FJ Jr, Gold J, Henderson SW, et al. Psychiatry and terrorism. *J Nerv Ment Dis*. 2011;199(8):537–543.

Tucker P, Pfefferbaum B, Nitiéma P, Wendling TL, Brown S. Intensely Exposed Oklahoma City Terrorism Survivors: Long-term Mental Health and Health Needs and Posttraumatic Growth. *The Journal of Nervous and Mental Disease*. 2016;204(3):203–209. Available at http://doi.org/10.1097/NMD.0000000000000456

*Ursano RJ, Fullerton CS, Raphael B, Weisaeth L, eds. *Textbook of Disaster Psychiatry*. Cambridge: Cambridge University Press; 2007.

Van der Velden PG, Wong A, Boshuizen HC, Grievink L. Persistent mental health disturbances during the 10 years after a disaster: four-wave longitudinal comparative study. *Psychiatry Clin Neurosci*. 2013;67(2):110–118.

Vineburgh N, Ursano R, Hamaoka D, Fullerton C. Public health communication for disaster planning and response. *Int J Pub Policy*. 2008;3(5/6):292.

Vlahov D, Galea S, Resnick H, et al. Increased use of cigarettes, alcohol, and marijuana among Manhattan, New York, residents after the September 11th terrorist attacks. *Am J Epidemiol*. 2002;155:988–996.

West JC, Morganstein JC. Psychological effects on military personnel assigned to humanitarian assistance and disaster response missions. In: Christodoulou GN, Mezzich JE, Christodoulou NG, Lecic-Tosevski D, eds. *Disasters Mental Health Context and Responses*. Cambridge Scholars; 2016:69–82.

▲ 31.6 Psychiatry and Spirituality

ARMANDO R. FAVAZZA, M.D.

Historians have traced the slow but progressive decline of religion in Western culture from the 16th century when the Protestant Reformation

led to the fragmentation of Christianity. Among the key events contributing to secularization are the scientific revolution started by Copernicus, the rise of mathematics and empirical investigation, the Age of Enlightenment in which intellectuals opposed religion as an institution and as a force in higher education, the French Revolution that expunged religious influence from public life, and 19th century theories of thought that dispensed with God by stressing the importance of economics (Karl Marx), biological evolution (Charles Darwin), and sexual urges (Sigmund Freud).

Secularization in the West is epitomized by empty European churches: 72 percent of the population of Britain and Germany claim to be Christians but less than 2 percent attend church regularly. Although the United States and Ireland are relatively pious nations, every indicator of traditional religiosity is either stable or declining, for example, lessened confidence in religious leaders and a decline in those choosing religious vocations. Religion no longer guides everyday life activities and one can engage in cultural life, for example, attend school, marry, watch television, and listen to music, without directly confronting God. The tide of history, especially in regard to sexuality, is turning away from traditional religious beliefs such as the prohibition of birth control and the sinfulness of homosexuality and same sex marriage.

Spirituality, described by one scholar as "a fuzzy concept that embraces obscurity with passion," is stepping into the space vacated by religion. The word itself has biblical origins and derives from the Latin translation of the Hebrew word variously meaning breeze, wind, breath, and by extension the life principle or breath of life. Traditionally, it has been linked with religion but in polls many people, especially those born between 1980 and 2000 [Generation Y] self-identify as "spiritual but not religious." They seem to be put off by religious dogmatism and intolerance. While they still may profess a belief in God, they are more likely to encounter Him/Her while walking in the woods at sunset rather than a church.

There is no generally accepted definition of spirituality, but studies and descriptions of it almost always mention transcendence. The concept often refers to short-lived moments during which the personal self blissfully merges with God, the Void, or the universe. This merging is accompanied by an unutterable understanding of external truths, insights, and ultimate reality. This epiphany is also called a mystical experience. Some have proposed that transcendence may occur on a continuum ranging from a peak experience or losing the sense of self during a special connection with nature, with art or music, or becoming one with a beloved in romantic love.

George Vaillant describes spirituality as an amalgam of the positive emotions—compassion, forgiveness, love, hope, joy, faith/trust, awe, gratitude. The spirituality that arises from these emotions "Binds humans to one another to experience 'God' as we may understand Him/Her." The positive emotions have an evolutionary basis and architecture. Through

genetic changes, reptiles evolved into loving mammals through the development of the brain's limbic system that underlies spirituality. The human neocortex has evolved over 200 million years to generate science, culture, ideas, behaviors, and religion, and has brought the limbic positive emotions into consciousness. Cultural evolution, working over a shorter time frame, was facilitated by the development of language that allowed for mastery of past and future tenses. A transformative cultural evolution occurred during the Axial Age, variously described as ranging from about 800 BCE to 100 CE, when in many areas of the world major religions were formed along with the development of a moral conscience. Rationalism in political and ethical behavior became increasingly important. These changes pushed forward the concepts of self-consciousness and a transcendent element in life. The third evolution in the development of spirituality is that which occurs when the positive emotions are learned and grow during the maturation process from infancy to old age.

Love, for example, is Vaillant's shortest definition of spirituality and "lives in the anterior cingulate gyrus." It is catalyzed by genes and by neurochemistry, for example, oxytocin in limbic dopaminergic centers is linked with lasting human attachment. It is learned both through cultural examples—poetry, songs, stories—and through identification with persons who exemplify it by their behaviors. All the positive emotions are grounded in a threefold evolutionary heritage and while they allow for transcendent moments they represent "religion without the side effects." Religion with its sacred texts and dogmas arises solely through cultural evolution where it has played an uneven role. It has produced intolerant and selfish behaviors and also is an opening through which the positive emotions are brought to conscious attention and reflection. Religion involves beliefs, is authoritarian, binds believers to their own community, and is imposed from without. Spirituality involves emotions and trust, accepts the experiences of others as valuable, is democratic, and is imposed from within. The ultimate promise of many of the world's religions is a happy eternal afterlife in some sort of heaven but the promise of spirituality is living an ethical, moral, and compassionate life guided by the positive emotions and sometimes punctuated by transcendent moments.

A biological basis for religious experiences has long been evident in cases of persons with temporal lobe epilepsy and some psychiatric disorder such as religious delusions and obsessions in schizophrenia and OCDs. The past few decades have seen growing numbers of studies aimed at understanding the neurological basis of spirituality. Neurological studies have generally revealed a frontal parietal circuit in which increased frontal lobe activity is associated with advanced meditative/prayer states and with an enhanced propensity for the representation of religious concepts. A more robust finding is that decreased activity in the right parietal lobe (RPL) is associated with a decreased focus on the self and is the primary

neuropsychological foundation for spiritual transcendence.

Brick Johnstone has been a pioneer in developing a functional geographic model based in part on the studies of persons with traumatic brain injury where it is possible to relate specific cerebral structures and neurophysiological abilities to specific spiritual traits and abilities. After administering comprehensive neuropsychological evaluations and various measures of spirituality and religion, he determined that decreased activation of the RPL, due to brain injury, surgery, transcranial magnetic simulation, or meditation/prayer practices, provides the foundation for transcendent experiences. The decreased activation results in the experience of selflessness and a diminished personal reference to space, time, or relationships. This selflessness "allows individuals to focus on and emotionally connect with things beyond the self, i.e., the very definition of self-transcendence." The reason that this selflessness is experienced as spiritual transcendence rather than typical symptoms of RPL dysfunction such as anasognosia or left-sided neglect "appears to be partially related to the context in which the selflessness is experienced, e.g., in an artistic/musical context, personal relationships, religious/spiritual setting etc." Johnstone notes that neurologically based spiritual transcendence can be perceived as a numinous connection with God or a high power or as a mystical connection with the universe based on the environmental context and a person's religious and cultural upbringing, for example, numinous if reared in a monotheistic/polytheistic culture or mystical if reared in a Buddhist culture. Differences in neurological processing may reflect differing cultural values such as the Western independent self in contrast with the East Asian collective self. The left temporal lobe with its connection to the limbic system may serve as a brain center for universal experience of common religious archetypes (God, demonic figures, prophets, etc.) that can be a basis for differing theological, ethical, and moral constructs. Increased frontal lobe activity is associated with increased frequency of prayer attendance but its association with transcendence is unclear. It may have a role in focusing attention on spiritual thoughts, beliefs, and practices.

In addition to rational understandings of spirituality, there is also an esoteric tradition as exemplified by the Greek Eleusinian and Northern European Odinic Mysteries. New Age spirituality emerged in the 1950s. It emphasizes irrational personal experiences and notions such as spirituality-laden force fields, extraterrestrial and earthly vibrations, and topsy-turvy interpretations of quantum mechanics. It is amazing that so many people worldwide hold such irrational, albeit usually harmless, beliefs.

FORMS AND METHODS OF ACHIEVING SPIRITUALITY

Because of the exceedingly broad definitions that people apply to the

notion of spirituality, there are myriad ways to achieve a spiritual moment and to lead a spiritual life. Being a vegetarian, for example, may be considered by some people to be spiritual because it aims at sparing animals from being slaughtered, thus preserving the sanctity of living creatures. This section describes some of the many ways that people induce or enhance spirituality. Many use altered states of consciousness but others do not. Some rely on reducing sensory stimuli, whereas others rely on increasing stimuli. Some persons find spirituality deep within themselves, whereas others find it in the world around them.

Consciousness

Consciousness is difficult to define or explain except to state that it is the context in which one experiences the physical world and his or her mental world, including a sense of self. How it emerges is unknown. In addition to awareness, it has an unconscious component—dream states and the memories and feelings of past experience that are not readily available to conscious attention. It is scientifically possible to parse the various contents of consciousness, but attempts to explain any findings and to understand why they are experienced lead people in circles because any explanations and understandings arise from the totality of consciousness itself.

Consciousness can be altered and the sense of self changed but the concept of self is multilayered. I can cut myself, for example, but I cannot cut my Self. Be that as it may, one perspective on spirituality is that it is a meaningfully interpreted alteration of consciousness, especially the sense of self.

People spend much of their lives in an altered state of conscious sleep, during which they have dreams. For millennia, dreams have been associated with spirituality. In ancient Greece, asclepeions were healing temples sacred to the god Asclepius. Sick pilgrims slept overnight in the temple and reported their dreams to a priest who then prescribed a cure. If barren women dreamed of snakes, they believed this to portend pregnancy. Throughout the Bible, God interacts with people through dreams. It is likely that many religious visions were actually remembered dreams.

When people are awake, altered states of consciousness are brief (from a moment to an hour, except in cases of pathological dissociation, drug intoxications, or sickness such as coma or delirium), typically pleasurable, unusual experiences that cause a break with their normal sense of reality. This detachment is a form of dissociation and a gateway to the world of spirituality. Depending on the social context, the altered state of consciousness may be interpreted as having either a religious or a naturalistic meaning. The psychologist Abraham Maslow, for example, depicted a hierarchical pyramid ranging from basic physiological needs upward to safety, love, esteem, and self-actualization needs. At the top of

the triangle is the “peak experience,” which consists of a sudden moment of well-being, wonder, awe, and at times an awareness of a higher truth and the unity of everything in the world. Triggers for these peak experiences can come from exposure to beautiful scenes of nature or to great art and music, intense feelings of love, and a confrontation with severe illness or death. Maslow believed that these moments were therapeutic because they increased empathy, creativity, self-determination, and personal growth. He claimed that everyone was capable of experiencing peak moments, and that they should be cultivated.

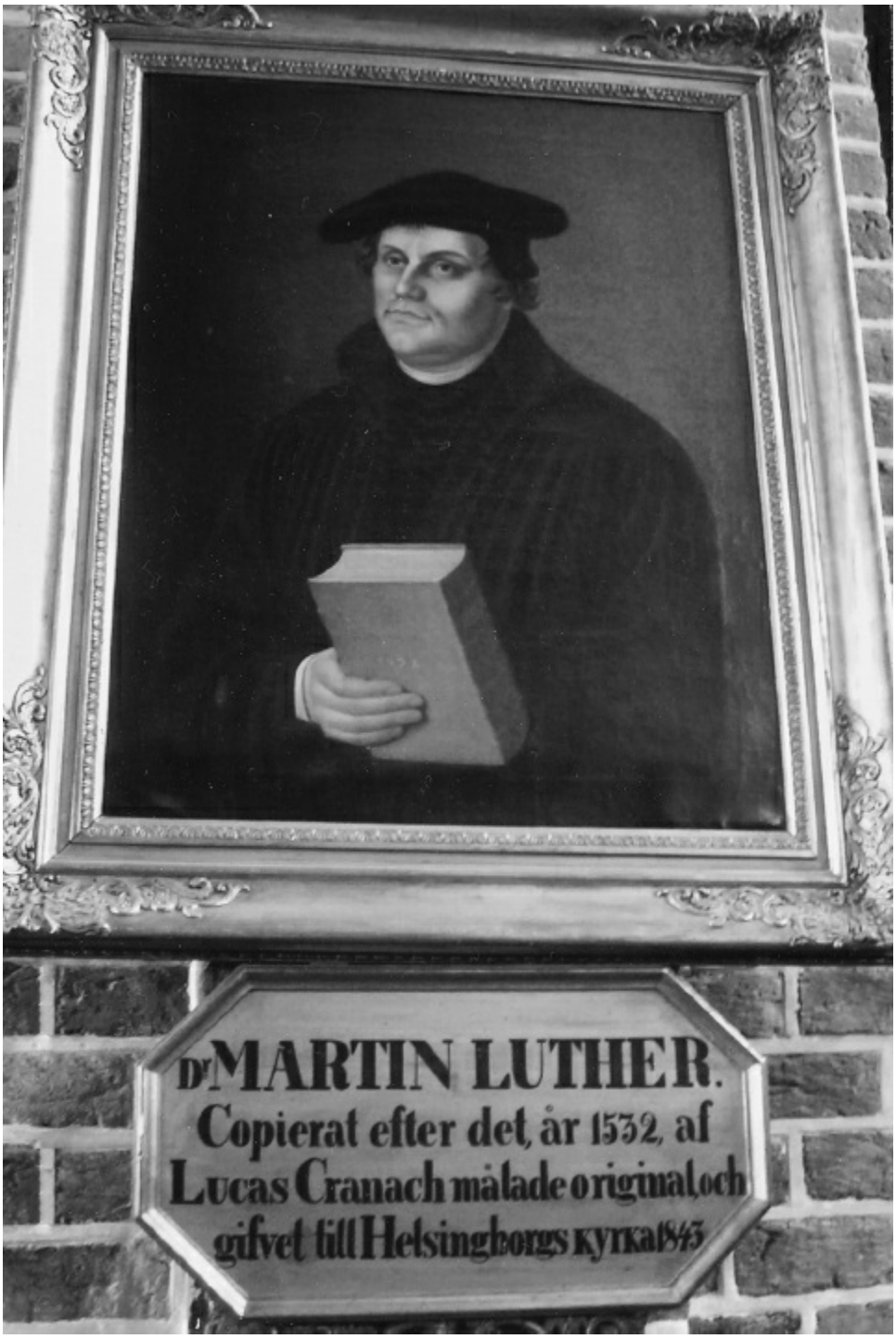

FIGURE 31.6–1. Martin Luther (1483–1547) led the Protestant Reformation and spurred the movement to translate the Bible into local languages. He believed that salvation could not be achieved by good deeds but rather by faith alone. (Courtesy of Ryan C. W. Hall.)

It is not uncommon for people to experience inexplicable fleeting insights that are less than peak moments and not quite altered states of consciousness but that have a spiritual dimension and may become incorporated into a person's spirituality. Martin Luther (Fig. 31.6–1), for example, wrote that when he heard a fellow monk recite the words from the Apostle's Creed, "I believe in the forgiveness of sins," he saw Holy Scripture in an entirely new light, felt that he was born anew, and had found the door of paradise thrown wide open. A single word or phrase, a line in a poem or a song, particular sounds from a symphony or an opera, even the play of sunlight on rocks or water can strike a chord of seemingly great importance. The taste of a madeleine cake sets off a string of childhood memories in the narrator of Marcel Proust's seven-volume masterpiece, *In Search of Lost Time*. Some persons under the influence of a psychedelic drug or a psychotic state may experience what they claim to be brief epiphanies of cosmic insight.

Meditation

Meditation is an activity usually practiced in a tranquil setting that attempts to self-regulate the mind by freeing it of wandering thoughts and emotions, especially negative ones, by promoting psychophysiological relaxation. There are many types of meditation practices that are best understood by relating them to the specific sociocultural traditions from which they arose. Generally, Western meditation practices focus on increasing a deeper love of God and a sustained focus on religious symbols and on the revelations found in sacred texts. Spirituality is achieved when the mind is filled with thoughts of God and devotional images. Transcendent states may be welcomed but play a minor role. Eastern meditation practices, especially Buddhist ones, focus on the primacy of the mind without reference to religion or faith. Spiritualist is achieved when the foundations of the self are transcended by comprehending that the self is an illusion and by elimination of all thoughts. In both traditions when meditation practices are mastered—no easy task—they can cause decreased activation of the RPL that may result in transcendent moments.

Mysticism

The quintessential example of pure spirituality is the intense altered state of consciousness known as a mystic experience, which has traditionally been described within a religious context. The characteristics of these experiences are many. Individuals who have these experiences claim that

they are so uniquely personal that their quality and worth cannot be fully communicated to others. They are a totally absorbing combination of feelings and of revelatory insights felt to be of such enormous significance that they transcend logic and intellectual comprehension. During the experience, mystics lose their sense of will and feel totally absorbed by the awe of the unity of the universe and of their immersion in the infinite ocean of God or of nature or of nothingness. They feel an elation and joyfulness unlike any other experience. The ultimate phase of the intense mystic experiences, at least in the Christian tradition, but probably in others as well, is often preceded by what St. John of the Cross called "the dark night of the soul." It is a period of helplessness that may be brief or last for decades before the self disintegrates, surrenders itself to the divine will, and eventually merges in rapture with the only true reality God. As normal consciousness returns, mystics who have had this experience return to normal life but have been transformed into intercessors between God and humankind. Although today such claims are met with skepticism, it was not uncommon in the past for mystics to claim ecstatic raptures in which they dueled with devils and interacted with Jesus and the angels. One of the best known mystics is St. Teresa of Avila (1515–1582). In one of her most famous raptures, she described an angel who pierced her heart into her entrails with a long dart with a fiery gold tip. As he withdrew it, she burned with a great love of God: "So great was the pain that it made me give those moans; and so utter the sweetness that this sharpest of pains gave me, that there was no wanting to stop, nor is there any contenting of the soul with less than God." This orgasmic rapture is immortalized in a statue in the Roman Church of Santa Maria della Vittoria by Gian Lorenzo Bernini.

Dhyana. *Dhyana* is a form of concentrated meditation taught by the Buddha. It is a practice of mind control to achieve an altered state of consciousness during which all thinking is abolished and truth is realized. Practitioners of *dhyana* must repent their sins, own only one appropriate garment, eat simple foods in small quantities, live in a quiet, humble abode, and be disentangled from worldly affairs. They must censor the desires that arise from the five senses, depriving themselves of gratifications and their ensuing cravings. They must clear their minds of thoughts that may result in greed, anger, laziness, sleepiness, recklessness, remorse, and doubt. They must sit quietly in a certain position, cross-legged with the spine erect, and breathe silently for 1 to 6 hours. Their minds must be tranquil and in harmony with their bodies and their breathing, and they should be totally aware of the emptiness and deceptive nature of the world with all its fraud and suffering. They may have visions of swollen, dead bodies with pus oozing out and maggots fattening on them so as to lose all pride and confidence in the ego self and in the selfhood of others and thus develop a dissolution of all attachments, as well as compassion for all sentient beings.

They then realize that neither the phenomena of the world nor the ego self are real. As they focus on the merits and purity of all the Buddhas, they experience a pervasive sense of bliss and safety. They are desireless and develop insight into both the emptiness of all phenomena and the potentiality within their emptiness, an insight that allows them to follow a "middle way" to attain supreme, perfect wisdom and Nirvana, which ends the cycle of reincarnation.

Samadhi. In Hinduism, *Samadhi* is a process of meditation that also leads to awareness of one's existence without the distractions of normal consciousness. It often involves an initial trance state that evokes a joyous sense of well-being and then, with the help of spiritual masters known as gurus or yogis, grows into detachment from one's body and eventually from one's ego. The meditator may attain the deepest moment of self-actualization during which the heart supposedly stops beating, breathing ceases, and there is a dissolution into the "divine ocean of love" and a realization of the unity of the universe.

Zen. The core practice of Zen Buddhism is *zazen,* the Japanese word for seated meditation in which a spiritual master instructs his students, some of whom will go on to become masters themselves, thus establishing a lineage of authority. Zen meditation uses *koans,* paradoxical or seemingly meaningless dialogues or questions, to test the progress of students. "What is the sound of one hand clapping?" is an example of a *koan* and is akin to the Christian theological question, "How many angels can dance on the head of a pin?" The goal of *zazen* is an epiphany (*satori* in Japanese) or lasting enlightenment about the true nature of creation.

Meditation in the Christian tradition is primarily an intellectual, sustained focus on spiritual themes such as Christ's suffering on the Holy Cross and his resurrection. Meditation on verses from the Bible is a common practice and derives from the book of Joshua 1:8, which states, "Do not let this Book of the Law depart from your mouth; meditate on it day and night, so that you may be careful to do everything written in it, then you will be prosperous and successful." Adoration of the Eucharist, which demands focused concentration sometimes for hours at a time, is a particularly Catholic practice.

Transcendental Meditation. Transcendental meditation, derived from Hindu meditative practices, is the most popular form of New Age meditation. It was trademarked by its founder, Maharishi Mahesh Yogi, in 1958. Although it is "officially" nonreligious, its goals are to attain "God Consciousness" and "Unity Consciousness," which are described as pure states of "being" in which a person sees the Self as integrated into all creation. This form of meditation uses physical exercises with softly spoken

suggestions that result in a mild trance that is facilitated by a group setting with teachers. Repetitive suggestions of what is and what will happen lead students to believe what they are told. Maharishi refers to Transcendental Meditation as a pathway to God. Some clergy see spiritual value in it, whereas some critics see it as bogus, as in the false claim that people actually levitate during Transcendental Meditation, whereas all that happens is that meditators sit cross-legged on rubber mats and hop forward on their derrieres while being told and believing that they are levitating.

Mindfulness Meditation. The concept of mindfulness meditation and its practices derives from a Buddhist tradition that focuses on the true nature of reality. In 1979, Jon Kabat-Zinn developed a secular Mindfulness Based Stress Reduction program for the treatment of chronically ill patients. It is now also used to promote better mental and physical health for persons in the community and in schools, businesses, government agencies, prisons, and veterans centers. The book *Waking Up: A Guide to Spirituality Without Religion* by Sam Harris is a thoughtful presentation of mindfulness.

A simplistic view is that it is the nonjudgmental, undistracted acceptance of all person's thoughts, emotions, and sensations in the present moment For Harris, a person's experience "is always now" and what is subjectively real in every moment. In advanced mindfulness meditation, the uncluttered mind aims not at being lost in thought or thinking more clearly about experiences but rather at experiencing consciousness itself more clearly. Thinking is indispensable to life but a person's identification with thoughts and failure to recognize thoughts as experiences in consciousness are a primary source of human suffering and of the illusion that a separate sense of self lives inside one's head. The self is an illusion even though people feel that it exists as their body and brain. Mindfulness breaks the spell of thought and of an illusory self. It allows a person to what pure consciousness is between one thought and the arising of the next one and reveals that consciousness is intrinsically selfless. "The conventional self is a transitory appearance, and it vanishes when it is looked for." Enlightenment, the direct experience of pure consciousness, reveals a Void that is filled with everything and is the basis of a spiritual life. Selflessness gives rise to spiritual/mystical experiences and gives one's life a moral dimension. Despite Harris's belief that it is conductive both to physical and mental well-being, specific health claims have probably been exaggerated.

Prayer

For the majority of humankind, prayer is the major entry into the world of spirituality. Prayer is a method of communicating with what are believed to

be a supreme spiritual being, other spiritual beings (angels, saints, the Virgin Mary, totemic animals, etc.), or ancestors. The communication may be verbal or silent, private or public; sung, danced, or acted out behaviorally; spontaneous or planned as part of a religious service, spinning prayer wheels, reciting the rosary, praying at certain times of the day, assuming a particular physical position, or facing a specific direction. Prayer may be used to worship or to placate spiritual beings; to confess sins; to petition for help in a variety of cases such as illness or poverty; to seek guidance, justice, or mercy; to enhance spiritual growth; to curse enemies; and to exorcise demons (Fig. 31.6–2).

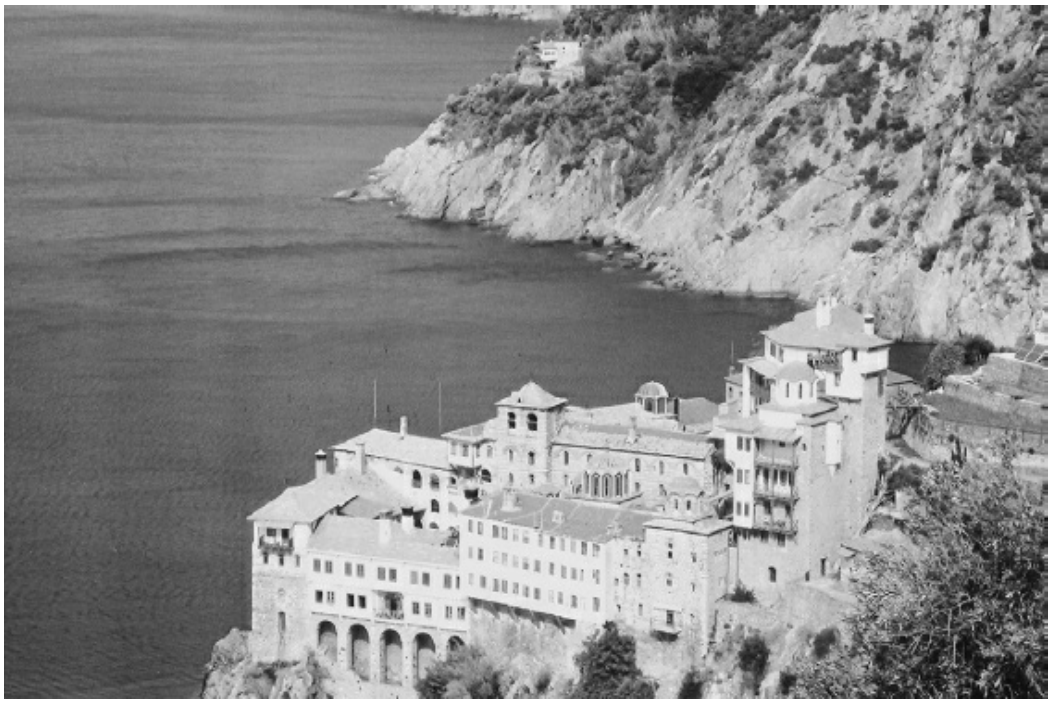

FIGURE 31.6–2. The monastery of Grigoriou on Mount Athos, Greece, the major center of institutions devoted to Christian Greek orthodox prayer. (Courtesy of Scott Cairns.)

FIGURE 31.6–3. A Lao Hmong shaman who is a specialist in curing opium addiction prays as he ties a group of addicts to a sacrificial pig that will travel as a messenger to the opium divinities with "spirit money." The therapeutic ritual is meant to free the addicts' souls from dependence on opium. (Courtesy of Wolfgang Jilek.)

People who pray assume that their prayers are received by an appreciative spiritual being who may or may not acknowledge or act on them. Prayer as defined traditionally posits a spiritual universe that can become operational in the phenomenal world—that is, prayer activates a spiritual energy that becomes manifest in everyday reality. Belief in a deity who has a personal interest in the life of all persons heightens the intensity of the prayer experience, but the experience may also be salient for those who believe in an impersonal deity that is not a being but rather is *being* itself, the god above god, the force of ultimate reality from which grace flows and on whom everything is dependent (Fig. 31.6–3).

Spirit Possession

Belief that a supernatural entity may enter into and even control a person's body and mind is an ancient and still-current spiritual belief found in most cultures. It is a belief that underlies such comments as "What's gotten into you?" or "What possessed you to do that?" and serves as the original prototype for Axis I mental disorders that people may have but from which they may be dispossessed with treatment. The possessing spirits may be malevolent or benevolent, and the possessed person may be perceived to be disturbed and sick or to be a healer or a prophet. A worldview that allows for spirit possession can be understood fully only on its own terms and not scientifically. Within a given culture, spirit possession may be quite logical,

and the pantheon of consensually validated spirits is unlike a projected paranoid pseudocommunity.

An entire chapter (Surah 72) of the Qur'an is devoted to *jinn,* a race of supernatural creatures created by Allah and sharing many characteristics of human beings although they are invisible. When angered or insulted, *jinn* may take possession of a person, causing all sorts of illnesses ranging from depression and psychosis to blindness, seizures, and paralysis. Although it is forbidden in Islam to summon *jinn* for evil purposes such as cursing a person or causing marital discord, some magicians are paid to do so.

In Egypt, Sudan, Ethiopia, and the Arabian peninsula, *zar* spirits, somewhat akin to *jinn,* are thought to possess women and cause headaches, anxiety, nightmares, and similar symptoms. The diagnosis of *zar* possession is made during a ceremony. A possessed woman will twitch and dance when a song that becomes special to her is played by musicians. She then goes into trance, and her possessing *zar* will demand gifts such as jewelry or clothing as an appeasement price for it to depart. The woman's husband and family give her the gifts, which she can keep for a while before returning them as her symptoms disappear. Thus, the *zar* interaction results in the woman's gaining the attention of her husband and asserting her position as a wife whose wishes should be respected. Many communities have weekly *zar* ceremonies during which women dance when their special song is played.

The utility of spirit possession is demonstrated in many shamanic rituals. A shaman is a generic term for a person who follows a divine call to healing, has received instructions directly from spiritual contacts, and may be a medium who communicates those spiritual instructions. Different types of shamans can be found in almost all cultural groups, such as Haitian *hungans,* West Indians *obeah* men, Puerto Rican spiritists, and Christian faith healers. Psychosocial profiles of shamans show that some are imposters, some are mentally disturbed, and some are mentally healthy, mature persons with a special cognitive style. The last may be considered true shamans who possess an ability to deal with and organize unconscious needs and primary process thinking. Often, through the mechanism of trance and spirit possession, the shaman acts as an agent of a powerful spirit and is thus able to perform psychological cures that may be temporary or permanent; for example, the spirit may declare that a patient's lost soul has been retrieved, as in cases of *susto* (spirit loss) in Hispanic cultures (Fig. 31.6–4).

Shamanic rituals are usually public affairs, and the persons in attendance include not only identified patients or troubled persons but also family, friends, community members, and cult devotees. This openness helps dissipate the stigma of illness, clearly expresses societal approval for attempts at remediation, and demonstrates recognition of the connection between the private health of individuals and the public health of society.

The healing experience often involves altered states of consciousness in both the shaman and the patient. Persons who enter a trance state as part of a healing ritual are relieved from personal responsibility for what they say or do because it is the possessing spirit who is "really" talking or acting. Such a mechanism allows for the expression of thoughts or emotions that otherwise might remain suppressed or repressed. It can bring out into the open any interpersonal and social conflicts that are troubling individuals, families, or the entire social group. Such conflicts have a better chance of resolution when they are brought out in the context of a healing ritual rather than in other forms or private confrontations.

In Puerto Rican *espiritismo,* there are several procedures for removing pathological spiritual influences. In *despojo* (exorcism), the mediums may fumigate patients with cigar smoke and stroke their necks. In addition to attracting benign spirits to the patients, the mediums may also attempt to draw off the evil spirits into themselves. *Trabajando la causa* (working the cause) is a technique used when special pathogenic spirits have been identified; it is the treatment of choice for sorcery. Possessed mediums interrogate the patients' spirits and try to convince them to depart. *Dando la luz el spirito* (giving light to the spirit) is a technique in which a patient follows a set of rituals to remove a spirit's harmful influences, such as purchase flowers or candles, perform specific therapeutic tasks, and attend meetings at the medium's *centro*. *Desarollo* (spiritual development), which is indicated for persons with dissociative symptoms, is a process by which patients are taught to display their symptoms at culturally and ritually appropriate times and places.

Many religions allow for possession by both benign and evil spirits. In Christianity, for example, possession by the Holy Spirit is a highly esteemed experience, especially among Pentecostals, in whom it may become manifest in the gifts of speaking in tongues and prophecy. The exorcism of harmful possessing demons is prominent in the New Testament. Nearly 20 percent of the Gospels are devoted to Jesus' healings, with the commonest ailment being illness caused by demon possession. According to Christian teaching, he passed on this ability to all baptized Christian believers: "And these signs will follow all who believe; in my name they will cast out demons; they will speak in tongues; they will take up serpents; and if they drink anything deadly, it will not hurt them; they will lay hands on the sick, and they will recover" (Mark 16:16–18).

FIGURE 31.6–4. Horse possession dance (*Kuda lumping*) in Java, a ceremony performed for spiritual and health benefits. Participants are possessed by the spirits of horses of historical heroes. In trance states, they act like horses and dance-jump to exhaustion. (Courtesy of Wolfgang Jilek.)

Almost all religions have rituals for exorcising demons. Exorcisms were common in the early centuries of Christianity and then fell to a low level until a renewed interest in healing emerged at the turn of the 20th century, especially in the US Midwest. The Catholic Church retains its formal, infrequently used rite of exorcism, although one or more simple exorcisms are routinely pronounced in the sacrament of baptism. Fundamentalist ministers perform exorcisms when they perceive evidence of often-dramatic demon possession. Televangelists have become more sophisticated in their approach; they are more apt nowadays to speak of demonic *oppression,* especially in persons with impulse control problems such as alcohol and substance abuse, self-injury, gambling, bulimia, and promiscuity.

Drugs

Drugs have been used throughout the world for millennia as a portal into the world of spirituality. Drugs such as LSD (lysergic acid diethylamide), DMT (*N,N*-dimethyltryptamine), peyote, ayahuasca, and psilocybin are called hallucinogens, although they rarely cause hallucinations. They are more properly called psychedelics ("mind manifesting") in the lay press. The new name for these drugs, especially among those who use them, is entheogens (derived from the Greek word for "God within").

The psychedelic experience can range from mild to very intense,

depending on the dose of the drug and, perhaps more important, on the psychological expectations of the user and the social setting of the drug use. The experience, or "trip," involves visual, auditory, and time distortions as well as synesthesia. At its peak, the drug users have a loss of ego with feelings of merging with space and objects or the universe. They may feel a sense of reaching to the beginning or end of space and time. The psychiatrist Daniel X. Freedman, a pioneer in psychedelic research, described the portentousness of the LSD experience: "The capacity of the mind to see more than it can tell, to experience more than it can explicate, to believe in and to be impressed with more than it can rationally justify, to experience boundlessness and boundary-less events, from the banal to the profound." The use of psychedelics may alter peoples' concept of reality, change their views on life or death, and challenge or promote their most passionately held beliefs. Like spiritual transcendence or ecstatic religious rapture, the peak LSD experience cannot be fully described verbally. However, when "trips" go sour, the experience can be devastating, nihilistic, fearful, and filled with demonic visions.

In a recent blinded study, 36 volunteers with spiritual or religious learning who had never taken a psychedelic drug were administered alternate doses of psilocybin and methylphenidate for an 8-hour session in a comfortable room. Based on responses to a Mystical Experience Questionnaire, 60 percent had a mystical experience, and 70 percent rated the experience among the top five in their lives, on a par with the birth of a child or the death of a parent. The effects persisted for 16 months and were confirmed by persons who knew the subjects intimately. Unpleasant sessions manifested by short-term anxiety and troublesome personal struggles were reported in 22 percent of subjects, some of whom nonetheless found the experience meaningful.

Cannabis has been used to achieve spirituality since ancient time. It is consumed as a drink called *bhang* by many Hindus during the festival of Holi, and it is smoked especially by devotees of the god Shiva. Although the Qur'an forbids wine, most Muslims interpret this to include all drugs that affect consciousness. Nonetheless, cannabis use has been present in Muslim society, especially by Sufis, who use it for spiritual reasons. The Rastafari religion, which emerged in Jamaica in the 1930s, equates cannabis with the biblical Tree of Life and encourages its use sacramentally as a means of discovering the truth and getting close to JAH (God). Rastafari claim that the drug must be used regularly in order to achieve cosmic consciousness and to rid their hearts of corruption. Its use is often associated with Bible study and with truth-seeking group discussions. The Ethiopian Zion Coptic Church, a small religious movement that was incorporated in Florida in 1975, equates cannabis with the Christian Eucharist (consecrated communion wafer).

Alcohol has long had spiritual connotations. Many Greeks worshipped

the god Dionysus, who supposedly invented wine. At sacred rites better known by their Roman name, the bacchanalia, his religion was characterized by the drinking of wine, which then released almost bestial emotions. Women were especially attracted to the chaste orgies, where they were free to forget their domesticity, dance ecstatically in the darkness of the mountains, and transcend human cares and fears. With alcohol-driven superhuman agility and strength, they ripped apart wild animals and devoured them sacramentally, believing that each bloody morsel contained a bit of Dionysus himself. The wine cult spread rapidly to areas as far apart as India and Britain. It also took hold in Palestine and Asia Minor, where, according to some scholars, the Dionysian orgy evolved into the Canaanite Feast of Tabernacles.

The Bible refers to wine as "the blood of the grape," and in its most revered transformation it became the very blood of Christ. At the Last Supper, Christ instituted the Eucharistic sacrifice, in which he offered to his disciples bread and wine that he declared to be his body and blood. By eating the bread–flesh and drinking the wine–blood, the disciples ingested their God. He became part of them and his immortality became theirs, an immortality available to all believers who partake of this sacrament.

The psychologist William James wrote of alcohol's intoxicating power to stimulate the mystical qualities of human nature and to provide a core religious feeling of transcendence and a sense of meaning in life for alcoholics. Alcohol's physiological power to transform consciousness can, depending on the context, be interpreted as a transcendence of everyday life, an entry into the cosmic world of spirituality. Alcohol fosters an experience of trust, relatedness, and unity through an illusory closeness to others. It also causes a psychological regression so that, just as infants expect mother to care for them, alcoholics feel entitled to God's protection. Drunkenness may be perceived as a shortcut to a higher life and the achievement of a higher state without any emotional or intellectual effort. In some ways, alcoholism is religion gone awry, a metabolic error of the religious impulse. Over time, as alcohol dominates the lives of alcoholics, it replaces a higher spiritual power with itself and alcoholics turn to it for redemption and the alleviation of guilt and of insecurity.

Body Modification and Mortification

Nonmedical alteration and mortification of body tissue, such as tattooing, flagellation, and piercing, for spiritual reasons have a lengthy history. Pope Benedict XIV (1675–1758) wrote: "With the exception of the martyrs, the Church venerates and gives the sanction of her authority to the sanctity of those only whom she finds to have been zealous in the modification of the flesh and senses." One such person was the blessed mystic Heinrich Suso (1300–1366) who, naked, beat himself with a spiked scourge until blood poured off his arms, legs, and torso, and bits of his flesh were ripped from

his body. Gazing at his beaten body, he was reminded of the beatings endured by Christ. Another was St. Catharine of Siena (1347–1380), who ate very sparingly, whipped herself with chains for 90 minutes three times a day, and deliberately scalded her body. After drinking the pus that oozed from the breasts of a cancerous woman, she said that she was visited by Christ, who invited her to drink the blood from his wounds.

Flagellation was a common Christian practice. In the 14th century, for example, itinerant bands of self-flagellators were received as heroes by a public who prayed that the bloody penitents would protect them from the plague. Mild flagellation is today practiced during prayer by some strict monastic orders and, along with brief crucifixions, by lay penitents in the Philippines and Latin America. In Islam, the death of the martyr Husain ibn Ali, Muhammed's grandson, is memorialized yearly by Shiites. On this day, some mourners beat themselves, often severely, with chains and spiked clubs while recalling his words, "Trial, afflictions, and pains, the thicker they fall on men, the better do they prepare him for his journey to heaven."

Within the Eastern Orthodox tradition, the sect known as the Skoptsi (eunuchs) included castration as a central religious practice. Its founder believed a medieval legend that Adam and Eve were born sexless. After they ate the forbidden fruit in the Garden of Eden, God supposedly grafted the halves of the fruit onto Adam as testicles and onto Eve as breasts. Only through castration and through removal of the nipple or breasts in woman could believers return to the pristine state of purity that existed before original sin. Despite severe repressive measures by the Russian government, the sect persisted for several hundred years until the turn of the 20th century.

Among Hindus, the yearly festival of Thaipusam commemorates both the birthday of the deity Lord Murugan and the day he was given a *vel* (lance) to conquer an evil demon. The highlight of the festival is a brief pilgrimage to a holy site before which devotees, after much preparation, enter a trance state and then pierce their skin, tongue, and cheeks with a *vel.* Because Lord Murugan's conveyance is a peacock, some persons carry a *kavadi* (somewhat of a portable altar) that may weigh 100 lb on their shoulders. The *kavadi* is festooned with peacock feathers and is attached to the carriers by hooks in their backs and chests. The festival, a joyous event held in several sites, is growing in popularity. Those who pierce themselves not only give thanks to the deity but also accrue spiritual merits. In Islam, some Sufis practice potentially dangerous forms of body piercing to achieve a spiritual state.

FIGURE 31.6–5. Fakir Musafar, one of the founders of the Modern Primitives Movement, re-enacts the Sun Dance of the Native American Plains Indians. (Courtesy of Charles Gatewood.)

Fakir Musafar, the adopted name of a Caucasian South Dakotan, is a primary founder of the Modern Primitive movement. He has devoted his life to understanding the spirituality of body modification in rituals throughout the world and putting this knowledge into practice. He is a shaman of sorts who has established a licensed school for body piercing and modification in California in which he indoctrinates students into the

spiritual underpinnings of his craft (Fig. 31.6–5).

Confession, Penance, and Reconciliation

In many religions, sins are offenses against the deities, ancestral spirits, and certain taboos. Sins, especially grave ones such as blasphemy and murder, separate people from God and destroy the charity and love of God in their hearts. Only by confession and reconciliation with God can sinners avoid divine punishment, sometimes in this life or certain eternal damnation in the afterlife, restore a right relationship with God, and regain spiritual purity. In many religions, atonement is achieved by participation in holy rituals. For example, Christians believe that atonement comes from the sacrificial crucifixion of Jesus Christ, which washed away the sins of the world and allowed humankind to be reconciled with God.

Visiting Spiritual Sites

Every religion and culture has designated sites considered to be sacred and whose spiritual forces may affect the lives of visitors. The sites may be churches, temples, or natural landmarks such as mountains, rivers, or rock formations. The Temple Mount in Jerusalem is one of the holiest places on earth. The Church of the Holy Sepulcher has been a pilgrimage destination for Christians since the fourth century, who believe that it contains the sites of both Jesus' crucifixion and burial. Also on the Mount is the Islamic shrine the Dome of the Rock, which Muslims believe is the place from which Muhammad made his Night Journey to visit Allah in heaven. This site was built over a Roman temple devoted to Jupiter and, before that, the Jewish Temple. The rock or Foundation Stone in the center of the Dome is thought by Jews to be the place where the prophet Abraham was tested by God to see whether he was willing to sacrifice his son, Isaac.

Charity and Altruism

In all the major religions, charity is considered to be the greatest or among the greatest of all spiritual values. Its fullest explication in Christianity is found in Paul's First Letter to the Corinthians. Paul used the Greek word *agape,* which denotes the purest and most absolute level of unconditional love. The word also refers to the theological virtue by which people love God above all things for his own sake and love themselves and their neighbors for the love of God.

The love of one's neighbors, a highly spiritual virtue, can be manifested physically by engaging in acts of kindness and providing both comfort and material things to the less fortunate. The Salvation Army, for example, aims to save people's souls but provides shelter, food, and clothing to the poor. Giving alms is one of the Pillars of Islam. Hindus give food to sadhus and Buddhists to monks. Such giving enhances the spirituality of the giver.

Buddhist monks may demonstrate against unjust political rulers by walking with their begging bowls face down, thus denying their oppressors the opportunity to give them food.

Altruism is a secular term referring to a selfless obligation of benevolence, respect, and caring for humankind, all living creatures, and the environment. Philanthropists and those who selflessly feed the hungry, clothe the naked, preserve and improve the environment, shelter the homeless, care for the sick, protect the defenseless, assist the weak, and enlighten the uneducated rightfully feel that they have accrued spiritual merits for their good deeds and attain a sense of meaningfulness in their lives.

New Age Spirituality

The term New Age refers to a cultural movement begun in the late 20th century with multiple roots, beliefs, and concerns that vary widely and include aspects of the following: mainstream religious traditions, Sufism, mesmerism, shamanism, Kabbalah, occultism, neopaganism, theosophy, magic, karma, unidentified flying objects (UFOs), meditation, psi energy, astrology, synchronicity, numerology, fortune-telling, parapsychology, ecological responsibility, guardian angels, cosmic forces, yoga, fasting, contact with aliens, near-death experiences, levitation, dream interpretation, and Wicca. This list is hardly inclusive but does give a sense that New Age spirituality has no sacred texts, central organization, membership, clergy, central headquarters, dogma, or creed.

New Age believers are generally optimistic and eager to enhance human potential by tapping into so-called spiritual vibrations and powers that emanate from the cosmos, crystals, pyramidal objects, and geographic locations such as Sedona, Arizona, where the Center for the New Age employs psychic readers and energy healers specializing in spiritual quests and inner transformation.

A major result of the New Age movement is the development of a thriving spiritual tourism industry. Using a computer in the comfort of one's home, it is possible to book spiritual pilgrim tours anywhere in the world. Thus, one can seek spirituality among the Kung Bushmen of the Kalahari Desert, Australian aborigines, or Amazonian shamans or at exclusive spas in Bali or ashrams in India. Such tours often may dramatically affect the lives of participants.

SPIRITUALITY AND RELIGION IN CLINICAL PRACTICE

Spirituality can be a positive force in peoples' lives. In 1984, the California Psychological Association's Task Force on Spirituality and Psychotherapy provided a new definition of spirituality as "The courage to look within and to trust what is seen and what is trusted appears to be a deep sense of

belonging, of wholeness, of connectedness, and openness to the infinite." The psychologist David Elkins outlined the values of spiritual persons. These include beliefs in "a greater self" or a personal God; a sense of purpose in life and a quest for meaning; acceptance of the sacredness of nature and of all human experience; knowledge that ultimate fulfillment is found in spirituality and not in material things; altruism; idealism; awareness of suffering and death; and leading a life that has a positive effect on people, nature, and their relationship with whatever they consider to be the ultimate and transcendent reality. This is an impressive list, but it is a challenging task, for example, to reconcile spirituality in the sacredness of nature and the occurrence of devastating tsunamis, tornados, volcanic eruptions, earthquakes, congenital malformations, and stillbirths unless they are considered as part of God's great plan for humankind, which includes all the wonderful and horrible events of the past and present or as tough-love warnings by Mother Nature to tread lightly on the Earth.

The list of spiritual values would also seem to apply to religious values, perhaps the closest being those of Buddhism. However, in the case of Christianity, for example, the list would need to contain specific beliefs about Jesus Christ as the savior of humankind, personal salvation as the ultimate goal in life, participation in church rituals, and acceptance of the Bible as the only true sacred text of the only true religion. Consideration would then have to be given to sectarian beliefs such as obedience in spiritual matters to a church hierarchy, reservation of the priesthood for men, a wife's duty to be obedient to her husband, and denial of homosexuality as an acceptable form of human personhood.

Religion, Religiosity, and Spirituality

Religion, religiosity, and spirituality may have some overlapping dimensions, but the concepts are not interchangeable. Spiritual persons may neither participate in a religion nor possess religiosity, or they may participate in a religion but lack religiosity, or they may participate and have religiosity. Other permutations are possible; for example, people who attend church may be spiritual but lack religiosity, or people may possess religiosity but not attend church, and so on. Great confusion results when these distinctions are not made. Thus, in many studies, "religion" is measured by participation in religious services at a church, mosque, synagogue, or temple. However, some people attend only a few special services each year, whereas others may call themselves a Christian, Jew, and so on, and consider themselves to be a member of a religion even though they have not attended services for years.

Religiosity refers to peoples' devotion to their religion; Allport and Ross's Religious Orientation Scale, for example, contains the statement, "My whole approach to life is based on my religion." An affirmative answer ideally implies a sincere commitment to one's religious beliefs, which are

internalized and serve as the guiding motivation of behavior. In reality, religiosity is not an all-or-none phenomenon. More than 20 variables have been identified as components of religiosity, and people may score high or low on different variables. Religiosity usually does conform with many spiritual values, but it may be spiritually destructive, for example, in the case of cults who prey on children and vulnerable persons who are disillusioned with life, frustrated by failed relationships, lonely, or experiencing a sense of disorientation that may be apparent, for example, in newly arrived college students, foreigners, and runaways. Destructive cults use deceitful recruiting practices and masterful indoctrination programs that, often in the course of several days, can cause persons to abandon their autonomy and to conform to the groups' behavioral norms in exchange for a false sense of well-being and exclusivity. Once indoctrinated, cult members lose their cognitive flexibility and adaptability; demonstrate a narrowing, blunting, or artificiality of affect; and develop a childish dependence on cult leaders to make decisions for them. In the worst scenarios, cults may use perverse religiosity to justify their sexual abuse of children, as was the case with the Branch Davidian cult in Waco, Texas, whose compound was destroyed and many of its members (including 21 children) killed during a raid by federal agents in 1993.

In an effort, perhaps, to counterbalance the overwhelming focus of medicine on biology and biotechnology, it is not surprising that a host of the major health care, medical, graduate *medical education,* nursing, and social work organizations have jumped onto the spirituality bandwagon. Most medical schools have courses on spirituality and health, and residents in psychiatry, family medicine, and internal medicine are required to learn about spirituality. The movement has gained traction by numerous correlational and cross-sectional studies that make unwarranted conclusions about the effects of religion and prayers on medical and psychiatric illnesses.

The lay press has publicized widely exaggerated claims about the "God spot," an area in the brain that supposedly demonstrates the biological basis of spirituality. Neurologists have long known that patients with temporal lobe epilepsy may develop a newfound religiosity (often obsessive in nature), have transcendent experiences, and experience religious auditory and visual hallucinations. Some have speculated that St. Paul's "thorn in the flesh" was temporal lobe epilepsy, which facilitated his conversion to Christianity on the road to Damascus. In fact, before Hippocrates, epilepsy was known as "the sacred disease." The controversial psychologist and countercultural icon of the 1960s Timothy Leary championed the use of LSD for its therapeutic and spiritual benefits; he wrote a book titled *Your Brain Is God*. Recent functional magnetic imaging studies have shown that a large number of neural circuits involving the frontal, prefrontal, temporal and cingulate cortices; the caudate nucleus;

and the parietal lobes "flickered into life" when 15 Carmelite nuns reported feeling the presence of God. The psychiatrist Herbert Benson stated, "To put it simply, we're wired to God." Albert Ellis, the psychologist who developed rational emotive therapy, retorted, "The whole field is off its rocker." Benson reasoned that because faith in a medical treatment can be therapeutic in up to 90 percent of the most common medical problems, then "faith in an invincible and infallible source carries even more healing power." In other words, belief in the healing power of God is the greatest placebo of all. Benson is probably correct, in that this belief engenders hope, which is no small thing for people who feel hopeless.

At a conference sponsored by the Association of American Medical Colleges in 1999, a consensus of clinicians, medical educators, and chaplains agreed that spirituality involves a person's search for ultimate meaning in life through participation in religion, as well as through relationships with God or family, and is expressed in nature, rationalism, humanism, and the arts. This declaration is somewhat idealistic in that very few persons lead an examined life, much less search for ultimate meaning. Sages from the West, wise men from the East, prophets, philosophers, poets, priests, psychiatrists, and Ph.D. candidates have all struggled to find the ultimate meaning of life. Multiple answers have been proposed—to have offspring and occupy and subdue the Earth; to act justly, love mercy, and walk humbly with God; to glorify and worship God; to love one's neighbor; to defeat one's enemies; to experience oneness with the universe; to be content with cultivating one's garden; to experience pleasures; to seek truth and be wise; and so on—but no single answer suffices for everybody.

From a global perspective, most people lead simple, unquestioning lives that go from paycheck to paycheck. Sports, sex, movies, children, and television shows are more apt to be topics of concern and conversation than transcendence or the meaning of life. The consensus definition of spirituality is somewhat high-minded, although it states that spirituality "may be expressed" by listening to music or painting, for example. Does listening to rap music qualify, or must it be mostly Mozart? Does painting-by-the-numbers count? What about relationships with pets?

Because the word "spirituality" is used in many different contexts, a distinction should be made between *Spirituality* with a capital S and *spirituality*. The former refers to a unique and very special experience as defined in the following discussion, whereas lower-case spirituality is a more general term that refers to meritorious, estimable, and honorable behaviors and feelings that benefit both individuals and the common good.

Based on this distinction, Spirituality refers to the following:

Spirituality is a feeling of a special connection with a person, a group, any living thing, nature, a cause, sacred or culturally significant objects or activities, the universe, or a supernatural being such as a deity. This feeling is special because it penetrates to the perceived inner core of a

person's being and may evoke deep passion, intense love, joyful bewilderment, visionary experiences, or a sense of profound understanding that cannot be expressed adequately in words. The settings in which this feeling occurs are highly variable, as are their degrees of intensity. Traditionally, the most exalted and intense feelings are those of mystical transcendence and of understanding the true nature of reality and the ultimate meaning of life. However, equally valid and more commonly experienced are the mundane feelings of connections such as wonderment in the beauty of nature and reverence for sacred landmarks; sexual merging with a loved one; caring for children, grandchildren, or a faithful pet; working for charitable causes; cheering for a favorite sports team; and maintaining good relationships with friends, family, and neighbors in good times and bad. A person may seek after and prepare for this feeling, but it almost always comes spontaneously and lasts for a moment or, in exceptional cases of mystical transcendence, for up to 30 minutes. Although this feeling is transitory, it may have enduring effects. It can be a positive force in coping with the travails of daily life and a sustaining force, particularly in times of physical and emotional crisis and in old age, especially when the specter of death is on the horizon.

The adjective "spiritual" in everyday usage usually refers to spirituality rather than to the specific feeling of Spirituality. Many persons may consider themselves to be spiritual or to lead spiritual lives because they follow the Ten Commandments and practice the behaviors specified by Jesus in his Sermon on the Mount; attend religious services regularly; pray often; fulfill the requirement of the Five Pillars of Islam; follow the Noble Path; live in a cloistered environment; evangelize for their faith; make pilgrimages; participate in charitable, patriotic, and ecological causes; remain faithful to their marriage vows; provide for the well-being of their children, parents, and pets; accept people for the content of their character; forgive those who have harmed them; appreciate the beauty of nature; serve as positive role models; and in certain contexts attempt to influence otherworldly beings on behalf of humankind. This list is not exhaustive but does provide a sense of the attributes of spiritual persons and of spiritual lives. Spiritual persons who lead spiritual lives may never experience Spirituality as defined here but may approximate and even share some portion of it. If the Virgin Mary's love for her child Jesus was spiritual, then there is some measure of Spirituality in the love of every mother for her children. If the Native American spiritual belief that all things are connected is true, then there is a potential spirituality in all the thoughts, emotions, and deeds of people's lives. Evangelical, Fundamentalist, and Pentecostal Christians claim to have a personal conversion encounter with God in which they are born again. The Biblical book of 2 Corinthians states, "Therefore, if anyone is in Christ, he is a new creation; old things have passed away; behold, all things have become new." When this conversion moment is intense and occurs suddenly, the experience may be one of true Spirituality or, at the least, a close approximation.

Unfortunately, some people who consider themselves to be spiritual can be misguided and create mischief and even havoc when they try to force their beliefs and values on others in the name of a cause or a religion. Much suffering can be caused by those who believe in their own righteousness. The austere Puritans, for example, escaped religious persecution by fleeing

to America from England in the mid-16th century. Historians have characterized them as persons who were distressed by the thought that some people, somewhere, were enjoying themselves.

Psychiatrists' Perceptions about Religion and Spirituality

Surveys of physicians in general and psychiatrists in particular have always reported that they have both a lower level of belief in God and participation in religious activities than the general population. A recent survey, although involving only 100 psychiatrists, also showed a lower level of positive attitudes toward Spirituality, which was defined in terms of religion and feelings of connectedness with an entity greater than oneself that give meaning to life. This is certainly not anything new: A medieval aphorism was, "Three doctors, two atheists." These attitudes have probably resulted from the experiences of physicians in watching so many good people and innocent children die. It takes great effort to reconcile these experiences with the will of a just and merciful deity or to find in them any sort of transcendence. Sigmund Freud contended that religion was an infantilizing, neurosis-producing, tyrannical illusion, and most psychoanalysts followed his critique, with the prominent exception of Carl Jung, who, for example, believed that every person's mind contained unconscious prehistoric memories of the real Father–God who is perceived as one's biological father, contrary to Freud's belief that it was the other way around. When psychoanalysts had something positive to say about religion, they often managed to find a rankling counterbalance. Ernest Jones, for example, wrote about "the enormous civilizing influence of Christianity" and "sublimated homosexuality" in the same sentence. Even Henry Maudsley, the most influential British psychiatrist of his time and certainly not a Freudian, wrote in 1918 that, "the corporeal or the material is the fundamental fact—the mental or spiritual only its effect." In 1975, the outspoken psychologist Albert Ellis declared that religion itself is a self-depreciating, dehumanizing mental illness. These attitudes have probably attracted some persons to gravitate toward psychiatry as a career.

The change among psychiatrists to a more positive view of religion began in the 1960s and 1970s. The psychiatrists William Meissner and Ana-Marie Rizzuto each proposed a nontheological and easier-to-accept understanding of the human experience of God through the life cycle. A child's initial image of God is based primarily on the image of his or her parents. The child's God may be loving and protective, as well as fearsome, cruel, and not always available to gratify the child's wishes. A positive mother–child experience at this stage results in a basic sense of trust that, in later years, may lead to a trusting faith in God, whereas a negative, insecure experience may distort the experience of God. As the child begins the process of separation and of becoming an individual with some independence, it hears about something called God but lacks the capacity

to understand what a spirit or transcendent force is supposed to be. The child hears that God has made the world and, thus, likens God to the most powerful person that he or she knows, namely, both parents, but especially father, who tends to be the more forceful or aggressive parent. Over the next few years, as the child comes to recognize that his or her parents neither know everything nor are all-powerful, a romantic view of a perfect, heavenly father may emerge, as well as a distorted view of the family. God and parents may be idealized as totally protecting, or there may be a split into a good God and mean parents or into good parents and a mean God. Between 6 and 11 years of age, the child begins to appreciate symbols. The concept of God shifts from that of a flesh-and-blood person to a still imperfectly understood spirit whose power must be revered and feared because God exacts punishment for every misdeed. In this stage, an immature obsessional religiosity may emerge in conjunction with ritualistic behaviors.

During adolescence, God becomes more personalized as "my Savior" or "my Father" and is more connected to emotions and to subjective attitudes, such as love, obedience, trust, and fear. God is experienced as the ideal leader and advisor who patiently and kindly listens to one's innermost problems and desires. When not tempered by some maturity of judgment, this realization of God may lead to fanatical tendencies. As adolescents become moralists, the burden of sinfulness emerges, and they may engage in the extremes of loose or puritanical behavior. Rebellion against parental authority may carry over to the parent's religion, as well as to the God whom the parents symbolically represent.

As adolescents step into adulthood, they must decide to follow the religion of their parents, to seek a different religion, or to avoid religion altogether. A true religious faith that is based less on conflict and more on mature understanding probably is not acquired until about 30 years of age. For some, the inner representation of God remains solid; for others, it may change; and for still others, it may be rejected or replaced. Thoughts about the memory of God usually return in old age, especially when death nears. A person may decide to embrace God gracefully or grudgingly, for reasons that include mature acceptance of the cycle of life, hope of entering a heavenly afterlife, avoidance of eternal damnation, and neurotic fear, or to spurn God's ultimate encroachment; a person's last breath may be an inspiration or an expiration.

Freud held that belief in God was an illusion that softens fears about the dangers of life; that the moral world order that God promises instills hope that justice will be served; and that belief in an afterlife creates an ultimate setting for the fulfillment of wishes. For Freud, this illusion was noxious because it distorted reality and promoted a state of psychic infantilism. However, in some modern psychiatric thought, God is a benevolent illusion necessary for the growth and nourishment of mental and spiritual health.

God is like the baby blanket or teddy bear that is special to the infant who endows it with a unique quality, cuddles it lovingly, and sometimes hates it. As the child matures, these transitional objects are neither mourned nor forgotten but rather simply lose meaning. They are the stuff of illusion and exist mentally somewhere between subjective and objective reality. In this formulation, God is a special transitional object that does not lose meaning totally and is always available through religion. Rizzuto concluded that a person's humanity depends on illusion; the type of illusion selected, be it science, religion, or something else, "reveals our personal history and the transitional space each of us has created between his objects and himself to find a resting place to live in."

Another factor that began to emerge in the 1960s and 1970s was that psychoanalysis was giving way to the community mental health movement, which decreased the traditional role of psychiatrists as therapists. Nowadays, the majority of psychiatrist–patient interactions involve 15-minute medication checks. In such an environment, psychiatrists usually have little trouble calling on the clergy and religious institutions to help with patient care, for example, by providing relief from social isolation through attending religious activities, by advising about life issues as well as spiritual ones, and by giving emotional and material support. Some clergy and religious institutions are quite amenable to including persons with mental illness and mild mental retardation in their congregations, whereas others may be less inclined to do so. Psychiatrists need to be aware that the clergy, a group with diminishing members, have little to no training in providing mental health services and that their major priorities are saving souls and turning sinners toward holiness. Certainly, there are pastoral counselors, but their numbers are few, and they are most likely to be called on in hospital settings in times of impending death or major medical crises.

Undoubtedly, some psychiatrists harbor conscious and unconscious negativity toward religion and religious patients. This disdain may stem from multiple sources, such as contact with or reports of pederastic and philandering clergy; exploitative cults; greedy televangelists; anger toward God over the premature death of a loved one; personal guilt (the Semitic religions are quite effective in stirring this emotion); the belief that religion is an irrational and coercive force in peoples' lives; negative experiences while attending a religious grade school; and so on. There also undoubtedly are psychiatrists who overvalue the role of religion and religiosity in the treatment of the mentally ill. Such under- or overvaluation distorts the clinical encounter and compromises the care of religious patients, who may detect this bias and skew the information they provide, which then may be interpreted as denial, resistance, paranoia, or characterological psychopathology. It is imperative that psychiatrists be aware of their feelings and beliefs about religion and religiosity to provide appropriate

clinical care. The same caveat holds true for encounters with patients whose spirituality is manifested in secular rather than religious beliefs and practices, although it would be rare for psychiatrists to harbor disdain in dealing with patients who love their family and their pets, engage in charitable and altruistic activities, and find peace in natural beauty. Psychiatrists may consider some of the practices of New Age spirituality to be odd, but, with the exception of protracted use of psychedelic drugs, these practices do not carry the same baggage as religious ones.

Spiritual Assessment

For a thorough clinical assessment of a patient, the biopsychosocial paradigm is insufficient, and the adjective "cultural" should be added to include consideration of such factors as ethnicity and spirituality. A number of spirituality assessment tools, such as the (British) Royal Free Interview for Religious and Spiritual Beliefs and the University of Minnesota Module on Taking a Spiritual History, have been published and share many commonalities.

The core initial question is "Do you consider yourself to be a spiritual person?" If the answer is "No" or "What do you mean by spiritual?" Some appropriate responses include "Are there any things that you feel passionate about or that give a special meaning to your life?" and "What are the things in your life that give you the strength to go on living?" Sometimes, a more dramatic question may be posed, such as "What is it that keeps you from ending your life?" The most common answers, in no particular order, are linked to pets, grandchildren, and God and/or religion.

It should be no surprise that pets are ranked so high; animals are significant in religious history and imagery. Among the ancient Israelites, the Leah tribes were linked to antelopes, the Rachel tribes to lambs, the Manasseh tribe to asses, and so on. Australian aborigines have long revered the totemic kangaroo. The Sacred Jaguar was central to the Aztec and Olmec religions and remains important to Amazonian tribes. Native Americans have animal spirit guides. Ancient Egyptian religion was polytheistic and often zoomorphic; for example, the jackal god Anubis watched over the dead; the falcon-headed Horus was god of the sky; Menhit was the lion goddess of war; and Wadjet the cobra goddess was the protector of Egypt, and her image always appeared on the crown of the pharaohs. Cows are sacred to many Hindus. In Christianity, Jesus is often depicted as the Lamb of God and the Holy Spirit as a dove.

In Western cultures, dogs and cats are the most common pets, and not infrequently patients state that they could not commit suicide because of the disastrous effect it would have on their pets. Popular fictional animals such as Snoopy the dog, Garfield the cat, and Bambi the deer demonstrate peoples' willingness to accept, to be amused by, and to cherish animals.

Dogs, perhaps because of their fidelity, transparent emotionality, and the seemingly unlimited amount of unconditional "love" that they offer, often stir the passions of their owners: "I would kill anyone who tried to hurt my dog"; "No one understands me like my dog." Pets certainly have a calming influence, and taking care of them does provide some meaning to many peoples' lives. Asking about the status of a pet, even in a brief clinical encounter, often helps with treatment compliance because patients feel that the psychiatrist is interested in more than their symptoms. Encouraging depressed patients to take walks with their dog exposes them to the beneficial effects of exercise and sunlight.

The birth of a grandchild often changes the relationships between grandparents and their children, especially when these relationships are strained. Grandparents may experience both delight and a renewed meaning to their lives in caring for grandchildren, passing on their wisdom, and providing some respite for weary parents. Intergenerational bonding and continuity of the family partake of the spirituality that exists in daily life.

When patients say, for example, that God helps them to persevere and gives meaning to their lives, it is important for psychiatrists to gather more information by responding, for example, as follows: "Tell me about your relationship to God." "Do you believe that God is somehow playing a role in your illness?" "Has God ever made his presence felt in your life?" "Has God ever disappointed you?" "Have you ever disappointed God?" "Do you believe that God will be a factor in recovering from your illness?" "How do you stand with God now?" "How do you think that God will judge you?"

It is quite rare for patients to describe themselves to psychiatrists as sinners, a trend that has grown over the last century. Karl Menninger's popular 1973 book *Whatever Became of Sin?* noted that sin no longer is the topic of daily discourse but rather has been transmuted into mental illness and criminality. In addition, in recent decades, the concepts of free will and of personal responsibility, necessary components of sin, have been greatly diminished by explanations of behavior that stress genetics, childrearing practices, childhood experiences, chemical unbalances in the brain, defects in neural circuits, exposure to traumatic events, peer pressure, and the influence of advertising. Only in extraordinary circumstances is it appropriate for psychiatrists to ask patients about sin. Patients who regard sin as one factor in their mental illness should be referred to a pastoral counselor to deal with this issue but still may need to receive treatment by a psychiatrist.

When patients state that religion provides meaning to their lives, then this needs to be explored. Some appropriate questions are as follows:

What religion do you belong to?
Has religion been important to you throughout your life?

Do you attend church (synagogue, mosque, temple) services? How often? When was the last time that you attended?
Are you satisfied with your church?
Do you feel that you are welcomed by the congregation?
Tell me about the congregation.
Were you exposed to religion as a child?
Is your current religion the same one you grew up with?
Do other family members attend religious services with you?
Do you discuss religion at home?
What is it about your religion that you find most gratifying?
Does your religion or your religious practices help you to cope with stress and difficult times in your life?
Have you ever had any negative religious experiences?
Are there any aspects of your religion that you disagree with?
Do you read Holy Scripture or religious materials? Does this reading influence the way you feel or how you lead your life?
If you were to die right now, would you go to heaven or hell?

Most patients tend to elaborate their responses, but if they do not, then more detailed questioning is necessary.

It certainly should not be surprising that not all congregations or church leaders are alike even though they are members of the same religion. Each church, synagogue, mosque, and temple has its unique dynamics, personalities, and ordering of priorities. In some Sikh temples, for example, major confrontations may occur over the practice of eating at a table or on the floor. In highly hierarchical religions, such as the Catholic Church, individual members or a majority of a congregation may choose to disregard pronouncements by the pope or local cardinals or bishops over such issues as birth control and not allowing divorced persons or those who are tolerant of abortion to receive communion. Mosques that are attended by both Shia and Sunni Muslims may be troubled by internal politics. Congregations may be politically liberal or conservative. Some may be somber and others joyful in their religious practices. Some may be more tolerant of the mentally ill than others. A thorough spiritual assessment should take into account the characteristics of the specific church, synagogue, mosque, or temple that a patient attends.

Spiritual and Religious Problems in Clinical Practice

Asking all of the questions involved with a spiritual assessment in a respectful, nonbiased manner may set a positive tone for the doctor–patient relationship, especially with patients for whom spirituality and religion are important. Even patients who profess no interest in religion, however, may be intrigued by the idea that spirituality does not

automatically imply religion and that even atheists can lead spiritual lives.

If a thorough biopsychosociocultural assessment demonstrates that a patient is not mentally ill but rather has a spiritual or religious problem such as issues about a loss of faith or conversion to a different religion, *Diagnostic and Statistical Manual of Mental Disorders* code V62.89 may be the appropriate diagnosis. Such problems may arise during the course of treatment or may be the presenting chief complaint of new patients who may think that they are mentally ill or who are seeking reassurance that they are not mentally ill. Once a diagnosis is made, referral to a pastoral counselor should be made, even if the patient remains in treatment for an Axis I or II disorder.

Psychosis, Depression, and Anxiety. Some of the feelings associated with Spirituality as defined in this section are similar to those experienced during a psychotic episode—for example, visions, fleeting profound insights, closeness to God. There is an eeriness to a psychotic episode, however, that may be manifested by turmoil in the patient's life, by disorders of thought (loose associations, thought blocking, etc.) and often of emotion (mania, depression, etc.), and general appearance (odd clothing, mannerisms, and comportment). When one is unsure whether a patient is psychotic because of religious claims, the patient's peers—such as church members—should be consulted. They commonly will be able to identify the patient's claims as either abnormal or within the group's concept of normalcy. Psychotic persons demonstrating high religiosity and concrete thinking may interpret scriptural passages literally and are in peril of harming themselves (eye enucleation and self-castration are mentioned in the Gospels) and others (killing children during a postpartum depression to "save" them). Claims made by a person who has gone through a rapid conversion from a traditional religion to a cultic or somewhat "exotic" religion, such as Hare Krishna or Druidism, are highly suspect because such conversions are usually associated with emotional tumult and sometimes with psychotic episodes.

It is possible, however, for persons with severe, chronic mental illness who are not currently psychotic to experience Spirituality or to lead spiritual lives. In a Swiss study of 100 outpatients with severe, chronic mental illnesses, some subjects reported that religion instilled hope, purpose, and meaning in their lives (71 percent), although it induced spiritual despair in others (14 percent). Subjects also reported that religion lessened (54 percent) or increased (10 percent) psychotic symptoms, increased social integration (28 percent) or social isolation (3 percent), reduced (33 percent) or increased (10 percent) the risk of suicide attempts, reduced (14 percent) or increased (3 percent) substance abuse, and fostered adherence (16 percent) or opposition (15 percent) to psychiatric treatment. Religious delusions worsen the prognosis of patients and are associated

with a higher severity of hallucinations, bizarre beliefs, lower levels of functioning, longer durations of illness, and the need for higher amounts of antipsychotic medication.

The relationships among depression, religion, and the spiritual life are complex. In almost all cultures, severe major depression replete with suicidality, deep regression, marked agitation or immobilization, and total hopelessness has been recognized as an illness. The concepts of moderate depression and dysthymia, however, have been considered to be pathological conditions only in relatively recent times and have spread from Western cultures to the rest of the world. The number of adjectives that have referents to depression but are used to describe aspects of the human experience are truly vast: For example, dejected, sad, downtrodden, disheartened, downcast, despondent, disconsolate, gloomy, pessimistic, melancholic, despairing, grave, lugubrious, funeral, lamentable, sullen, grim, and so on.

All persons experience both sadness and happiness in their lives, but the former tends to be more persistent and is usually interpreted as more meaningful; thus, Oscar Wilde's epigram, "Where there is sorrow, there is holy ground." The most iconic image of Western civilization is the image of Jesus on the cross, bloody and beaten. The New Testament promises Christians suffering and persecution and urges them to follow Christ's example by sharing in his afflictions. Unlike gladiators, who glistened with oil and fought for their lives, martyrs aggressively sought a passive death—the bloodier, the better. In an impassioned tract written during the worst days of Roman persecution, the theologian Tertullian boasted,

> Some Christians have been tested by the fire, some by the sword, some by the beasts; some, lashed and torn with hooks have just tasted martyrdom, and lie hungering for it in prison .... Nothing whatever is achieved by each more exquisite cruelty you invent; on the contrary, it wins men for our cause. We are made more often as you mow us down. The blood of Christians is seed (*semen est sanguis Christianorum*).

Over the centuries, many Christians willingly accepted suffering for Christ's sake. Perhaps the greatest devotional manual of Christianity, *The Imitation of Christ,* written in the 15th century by Thomas à Kempis, is based on the principle of self-renunciation. Many of the great mystics have recounted long periods of despair, dryness of spirit, terrible darkness, and questioning of their faith in their quest for holiness.

The ancient notion that suffering is a just punishment for sins is found in the Hebrew scriptures, but some of its writers were aware that this is not invariably the case. Thus, Jeremiah (12:1) pleaded with God to answer "Why is it that the wicked live so prosperously? Why do scoundrels enjoy peace?" The answer that emerged was described by the theologian John Bowker as "perhaps the supreme contribution of Israel to a human response to suffering, that suffering can be made redemptive, that it can become the foundation of better things, collectively, if not individually."

In Hinduism, the deities often represent opposite and complementary qualities. Siva personifies the forces of both creation and destruction in the universe and must be propitiated on the basis of the relationships of devourer, devoured, and devouring. Kali, the Dark Goddess, is a bloodthirsty aspect of Siva. Her teeth are fangs, and she wears a necklace of skulls and a belt of amputated limbs. Her stomach is a void that can never get enough blood from beheaded offerings in her temples. Human sacrifices to Kali were abundant until forbidden by law in 1835. However, because Siva also represents the force of creation, suffering must be put into perspective as the connection between the two conflicting universal principles, the urge to life and the urge to death. Suffering is inherent to the tensions of the cosmos and is a significant part of the human experience. It is the process of being devoured, the better to be free from worldly attachments in one's search for transcendent spirituality. Sadhus are so revered in the towns and villages of India because they allow themselves to be devoured by living a life of asceticism and of twisting their wasted bodies into odd positions, sometimes for years at a time. It is this suffering that creates true liberation from the bondage of the flesh.

Buddhism is the belief system that focuses most directly on suffering. Indeed, the First Sermon of the Buddha at Benares contained the essence of his wisdom. In this sermon, Buddha disavowed both devotion to bodily pleasures and self-mortification. Instead, he preached a middle path that gives knowledge and leads to calmness, insight, enlightenment, and Nirvana. To reach this path, a person must comprehend the Noble Truth of suffering: Aging is suffering; sickness is suffering; death is suffering; sorrow, lamentation, pain, grief, and despair are suffering; association with the unpleasant is suffering; disassociation from the pleasant is suffering; not to get what one wants is suffering; all types of attachment are suffering. The origins of suffering are cravings for bodily pleasure, for existence and nonexistence, for becoming, and for annihilation. Only renunciation and detachment from these cravings will allow a person to travel the Noble Path leading to the cessation of suffering: Right view; right thought; right speech; right action; right livelihood; right effort; right mindfulness; and right concentration. In areas where these Buddhist beliefs are prevalent, the thoughts, acts, and emotions associated with the Western concepts of moderate depression and dysthymia have been perceived traditionally as within the normal range of human experiences. It has taken a large public health campaign along with much advertising by pharmaceutical companies, for example, to convince people in Japan that they are truly depressed.

Patients' declarations that they are sinners, evil persons, and, to a lesser extent, guilt ridden are encountered less frequently by psychiatrists than in the past. Some religious patients may try to extract meaning from their depression or psychosis or feel that God is testing them. A simple and

truthful approach should be used. All of the patient's symptoms should be enumerated and the diagnosis explained. They should be told that in addition to psychosocial factors that are contributing to their illness, modern science has shown that disturbances in brain chemistry and neural circuitry play a significant role and have no special meaning. Besides, it is not appropriate to seek any meaning during the course of an illness because cognitive and emotional distortions compromise their ability to reason clearly. Any quest for meaning should occur after the illness has either passed or come under good control. This approach does not discount the use of spiritual or religious resources during the course of treatment.

Anxiety is an experience of everyday life, and everyone, from the born-again Christian to nasty sociopath, has anxious moments—for example, when the telephone rings in the middle of the night, when one is crossing a busy street with six lanes of traffic, when a laboratory report indicates a possibly fatal illness. Persistent anxiety can be terribly corrosive both by itself and in panic, social phobia, posttraumatic stress, and OCDs.

The literature on anxiety, religion, and spirituality is sparse and consists mostly of cross-sectional studies. The only clear findings are that the presence of religious obsession and compulsions offer a poor prognosis, and that the collapse or weakening of faith in posttraumatic stress disorder patients often results in unresponsiveness to therapy. An Internet survey of 1,402 adults examined the association between mental disorders (using a subscale of the Symptom Assessment-45 Questionnaire) and religious characteristics. The study found that frequency of prayer was significantly related to higher phobic scores, whereas belief in an afterlife was related significantly to lower phobic scores. No association was found between phobia and religious fundamentalism or religious attendance. A 2002 study of 37,000 representative Canadians is the largest on the interactions among religiousness (defined as how often a person worships in an organized religious setting), spirituality (defined by how important a search for a meaning for life was for a person), and mental illness. The study found that religiousness and spirituality were linked positively with older age, higher income, and greater social support. Separated, divorced, and widowed persons were the most spiritual. With regard to depression, manic episodes, panic disorder, and social phobia, spirituality was significantly associated with higher lifetime odds, whereas religiousness had the opposite relation. The study was unable to determine whether spirituality helped set the stage for the genesis of the disorders, was a response to the disorders, or was simply a correlative finding.

Existential anxiety was a topic of psychiatric concern, especially in the mid-20th century, when it was heralded as the root cause of all mental illness. This anxiety supposedly stemmed from the experiences of being thrown from the womb into a meaningless world and then spending one's life confronting the fearful freedom that comes with meaninglessness and

eventual death, the trauma of nonbeing. People were said to cope with this anxiety by rationalizations and leaps of faith that attempted to impose some order in the world; for example, religions redefine death as the end of physical existence while allowing for the continued existence of the soul, which endures eternally. Although existentialism has a rich and provoking literary heritage, it is rarely even mentioned in modern psychiatric textbooks.

Anomalous Experiences. Psychiatrists may encounter nonpsychotic patients who report anomalous experiences that defy rational explanations, such as so-called near-death experiences, alien abductions, being touched by an angel, telepathy, and clairvoyance. Although these phenomena are odd, persons who have experienced them and have been studied do not appear to be mentally ill but rather go on to develop increased empathy with others and involvement with family members, enhanced interest in spirituality, appreciation of the purpose of life and of mundane occurrences, and a stronger belief in an afterlife.

Of interest, the literature on this topic includes mystical experiences—defined in this section as true Spirituality—as anomalous events. In various surveys, up to 40 percent of subjects reported a mystical experience at least once in their lives. Undoubtedly, some of these represent Spirituality but it is probable that many are out-of-the-ordinary emotional New Age experiences and may even include dissociative or brief psychotic episodes.

NEAR-DEATH EXPERIENCES. Near-death experiences or something like them have a lengthy and widespread tradition. In *The Republic,* the Greek philosopher Plato described the hero Er, who was slain in battle and carried home for burial. Twelve days later, his body, surprisingly unaffected by decay, was placed on a funeral pyre. Er then came to life and reported that his soul had left his body and gone on a journey with a great company to a mysterious place where the souls of the dead resided. He saw a rainbow-colored column of light extending through heaven and earth and scenes of judgment that were both terrifying and joyful. Descriptions have been found in Buddhist and Islamic texts and in Chinese, Siberian, Finnish, and North and South American accounts of out-of-body experiences in which a person's spiritual body experiences a bright light accompanied by joyful peace and meets with ancestors and departed friends after crossing the border between the living and the dead.

The psychologist Raymond Moody's book *Life After Death* contained 150 interviews with persons who reported near-death experiences and brought attention to this phenomenon in 1975. These experiences occur when persons are in a state of what appears to be death (they may even be pronounced dead by a clinician) but then are revived. Many report that during the period of not-quite-life and not-quite-death, they were out of

their bodies looking down at themselves and felt themselves moving through a dark tunnel peacefully and quietly. Their lives were reviewed as in a panorama. They saw a glowing light with a human shape, cities of light, or a border from which there was no return. They meet deceased relatives and friends. They heard doctors or spectators who pronounced them dead, but they were rescued from death by a spirit.

Remarkable similarities exist between the reports of near-death survivors and by users of hallucinogenic and other drugs. Patients recovering from anesthesia, for example, may recall words spoken by the operating team during surgery. Bright lights and tunnel imagery can occur when drugs that mimic the effects of light on the retina stimulate the central nervous system or when the threshold for light perception is lowered so that bright lights are seen surrounded by darkness. Published statements from hallucinogenic drug users describe

> bright colors not like here on earth but just indescribable .... I could see a city, a city of light .... There are tall structures all around me in all colors .... It seems like I'm getting closer and closer to the sun and there's a geometric network or lattice in the distance.

Even the panoramic, rapid review of a person's life has been described by users of drugs that excite the central nervous system. Such excitation may also be induced by fever, exhaustion, and psychological processes associated with dying; decreased sensory input from failing body organs; a turning inward of attention and a blocking out of the environment; reminiscence; and a fear of death that may trigger both comforting fantasies and dissociative mental states. A fanciful theory is that the near-death experience is so traumatic that it kindles the memory of birth trauma when the infant travels through the vaginal tunnel toward light. Yet another theory is that all memories are stored but most are suppressed by new experiences that create more memories. In situations such as the near-death experience, new input is markedly decreased, thus allowing the old memories to be released and to be psychodynamically experienced as hallucinations, dreams, or fantasies. When faced with imminent death, a person's thoughts often turn to God, heaven, and culturally accepted religious images.

It is likely that the imagery and sensations of the near-death experience are caused by biological processes and mediated by cultural expectations. Another example of a biological process that results in a strange experience occurs when a person whose painful limb had been amputated feels pain as if the leg was still present. This phantom limb pain starts at the limb stump and persists in the sensory centers of the brain. A different type of experience, the autoscopic phenomenon, is associated with epilepsy and migraine headaches. It consists in the sudden brief perception of a person's face or body as a colorless, transparent phantom that imitates the person's actions. The German writer Johann Wolfgang von Goethe described a likely

autoscopic experience when, while horseback riding, he saw an approaching horseman who was his double and who then rapidly disappeared. The biological process of sleep paralysis is associated throughout much of the world with culturally elaborated beliefs and visions of a ghost, the spirit of a dead person, or an "old hag" who presses on the sleeping person's chest. The Hmong of Cambodia claim that they see a tiny figure, or demon, who crushes their chests. In fact, so many young Hmong men die in their sleep that a sudden unexpected nocturnal death syndrome is now medically recognized. It is a genetic disorder that causes a fatal ventricular arrhythmia. Survivors describe intense fear, paralysis, and a pressure on their chests caused by an alien being.

Most persons endow their near-death experience with a favorable spiritual explanation, such as having been "touched by the light," which supposedly has some association with heaven and with God. Many religious traditions associate darkness with hell and the forces of evil, whereas light epitomizes heaven, goodness, and God. In *The Divine Comedy,* one of the greatest literary evocations of God in Western literature, the poet Dante saw in Paradise both a cross that brilliantly "flames forth Christ" and God, "the eternal light." Unfortunately, some near-death experiences are not so splendid. In various studies, up to 22 percent have been described as hellish and terrifying, with the person having a sense of being condemned to a barren void for eternity or falling into a demon-filled pit.

ALIEN ABDUCTIONS. Some persons claim that they have been abducted by extraterrestrial aliens. In many reports, these abductions take place at night in a rural setting while the victim is driving a car and spots a UFO. The victim may stop to investigate, or the car will stop because of mechanical failure presumably induced by the UFO. The driver and passengers, if any, get out of the car, become amnesic for several hours, and then regain awareness. Usually, they are mystified by the lost time, but in subsequent dreams, hypnotic session, or regression therapy to recover memories, they remember being paralyzed, transported in a beam of light, and abducted by gray-skinned humanoids with large heads and very large, dark eyes. During this terrifying experience, the victims report being surgically probed, having devices implanted in their body, and being sexually abused or impregnated by the aliens. A few victims may have some sort of mystical experience, but this is soon followed by medical and psychological problems.

This scenario seems to be a North American phenomenon that first started to receive public attention in the 1960s. Well-respected psychiatrists and psychologists who have interviewed persons claiming to be abductees have found them to be rather ordinary with a level of mental health reflecting that of the general population. Many have created support groups and a subculture devoted to explaining the abductions, including

the belief that they represent alien attempts to better understand human beings or to repopulate the Earth with human–alien hybrids, secret governmental mind control experiments, and so on. Similar reports of abductions from around the world tend to be more benign, with the abductors being demons or spiritual entities.

The physical reality of the abduction phenomenon is scientifically absurd, but there is no doubt that it might be a culturally elaborated psychospiritual reality. Because reports of abduction rely so often on the recovery of lost memories, one theory is that the basic repressed memory is of childbirth: The fetus is forcibly plucked from the safety of the womb and placed in an alien, cold environment; bright lights are shined on it, and it sees the eyes of masked strangers who probe its body. Because newborns cannot distinguish colors, the strangers appear to be gray, which accounts for the perception of abductors as "grays."

The historical, religious, and anthropological literature of the world contains accounts of many visionary encounters with deities, demons, and ancestors. The Hebrew Bible, for example, reports encounters with God and his surrogates. In 1149, an Irish monk wrote *The Vision of Tundal,* a highly successful book that was translated into most of the languages of Europe. Tundal saw Satan in hell viciously torturing sinners. At that time, no one doubted the veracity of Tondal's accounts or those of many other writers. The slow demise of visionary encounters began with the publication of Dante's *The Divine Comedy* in the 14th century. Dante provided a compelling and profound first-person account of his voyage to hell, purgatory, and heaven but admitted that it was a poem derived from his imagination. In Western culture, skepticism about visions continued through the 17th and 18th centuries up to modern times, in which visions have usually been associated with psychosis, although in folklore and in shamanic and some tribal cultures, visions are accepted as reality.

In today's world, only a few persons believe in the physical reality of the abduction phenomenon, a surprisingly larger number would endorse its possibility, and the vast majority reject the idea. John Mack, the most serious psychiatric investigator of the phenomenon, concluded that it is a modern type of visionary experience but overoptimistically felt that it might portend a dramatic change in human consciousness and a blow to the world's collective egoism in the tradition of Copernicus, Darwin, and Freud. Although it may be futile to try to convince abductees that their experiences were delusional and not physically real, Mack determined that many eventually come to regard the phenomenon more positively in that it resulted in a broadening of their consciousness.

TOUCHED BY ANGELS. Belief in guardian angels is a staple of both New Age spirituality and most persons who consider themselves to be religious.

Computer search engines typically show more than 2 million "guardian

angel" sites ranging from offerings of guardian angel pins, wallpaper, and poems to tattoos, newsletters, and dog rescue teams that are specific to certain breeds. Large bookstores have sections devoted to angels, and in the film *It's a Wonderful Life,* one of the most watched movies of all time, a guardian angel saves the day. There have been literally millions of testimonials throughout history, especially in the last half-century, attesting to the power of a person's guardian angel miraculously to heal relationships and illness, provide expert guidance in life decisions, expand consciousness and spiritual evolution, and render material support in hard times. In his 1975 best-seller *Angels: God's Secret Agents,* the evangelist Billy Graham's language was probably influenced by the recent US experience in Vietnam when he described angels as

> God's dynamite and added that "Satan's BB guns are no match for God's heavy artillery .... If your valley is full of foes, raise your sights to the hills and see the holy angels of God arranged for battle on your behalf."

An issue of the *Angel Watch* newsletter noted that "Each person has a guardian angel. They're nonthreatening, wise, and loving beings. They offer help whether we ask for it or not. But mostly we ignore them." One recent book even described guardian angels as "social workers of the universe" and offered a 12-step program to assist readers in linking up with their angels.

All of the major religious traditions profess a belief in angels or similar spiritual beings, who usually are considered to be God's messengers and bringers of good news—the archangel Gabriel told the Virgin Mary that she would give birth to the son of God—and look after people, although sometimes they can bring destruction. Christians consider Satan a fallen angel, but Muslims consider him to be a *jinn*. Because they are spiritual beings, angels are sexless, although in Christian iconography, they are depicted as men who often have feminine features. Buddhist scriptures mentions angels who, in Japanese Buddhism, are called *tennin* and are depicted as exceedingly beautiful and colorfully dressed female companions of the Buddhas and Bodhisattvas. The heyday for angels was 14th century Europe; some authorities calculated their exact number as 301,655,722. The situation was out of control, with religious scholars vying to identify and name as many angels as possible. Sammuel, for example, was variously described as the angel of death and the disguised-as-a-serpent seducer of Eve. Among his mates were the four angels of prostitution, Eisheth Zenumin, Naamah, Agrat bat Mahlat, and Lilith, who was notorious also as Adam's first wife and who gave birth to 100 children daily. Catholics, accustomed to the notion of intercessors with God, used images of angels in elaborate church decorations. Protestant reformers, however, had little tolerance for the presence of gilded statues in churches and taught that intercessors were not necessary because each person could

communicate directly with God.

Modern belief in guardian angels is typically superficial. Psychiatric patients, some of whom hold a deep conviction that a personal angel watches over them, should be treated with respect. It is a harmless belief that often provides hope.

Treatment Issues. There is much confusion in both the literature and in the minds of most people, including psychiatrists and patients, because of the conflation of the terms Spirituality, spirituality, spiritual, religion, and religiosity (or religiousness). The word Spirituality should be used to refer to the extraordinary moment or brief period of time during which a person feels a special connectedness to living beings, other forms of life, nature, causes, sacred or culturally significant artifacts, the universe, supernatural entities, or even ancestors. This feeling is special because it strikes deep to the perceived inner core of a person as a revelation or epiphany that cannot be fully described in words. Spirituality appears suddenly and it is possible for some persons to cultivate its reappearance. Some religious mystics, St. Teresa of Avila, for example, described multiple episodes of a transcendent connectedness with God. In a secular context, the polymath S. Giora Shoham described "experiences of blissful temporary dissolution of the self's cognitive partitions" while watching the sun set behind some ruins or the Mediterranean shore while the waves, the darkening sky, and his self fused into an opaque unison. He developed a capacity to recapture the experience; for example, "By concentrating on the yellow petals of a wild chrysanthemum while the petals extended and widened, my whole self was engulfed into a vision of a single omnipresence." Some patients report having had such a moment of Spirituality (not to be confused with moments of therapeutic clarity or insight) that influenced them to lead spiritual lives. The experience of Spirituality is a highly prized phenomenon but cannot be induced by psychiatric expertise, which can only enhance the probability that it might occur.

The Spiritual Life. Patients seek psychiatric treatment because they feel ill and beset by stress and suffering. Increasingly, medication has become the most common treatment modality, but supportive psychotherapy should be an integral component of every clinical encounter, even brief ones. As opposed to more traditional psychotherapy that occurs over an extended period of time and to symptom-focused, shorter-term cognitive and behavioral forms of therapy, supportive therapy is the most difficult. It presupposes that the psychiatrist understands the basic nature and course of an illness, the individual psychological and social dynamics of the case, and the patient's cultural beliefs and practices. The psychiatrist then uses this information to ensure treatment compliance through advice,

exhortations, admonitions, and even threats, such as commitment to a hospital. The psychiatrist's duty should be not only to reduce or eliminate symptoms but also to enhance a patient's quality of life.

Biological interventions rarely cure mental disorders, but they do make it possible for patients to deal with the issues that cause them stress and suffering. Through advice, exhortations, and so on, the psychiatrist's role is to point patients in the right direction on life's path. Some advice is immediately practical—for example, get a payee to help with finances or move to a different apartment to avoid the drug dealers next door. Much of the advice, however, is geared toward having patients lead spiritual lives, even if the adjective *spiritual* is not articulated by the psychiatrist or the patient.

Some of the things that constitute a good spiritual life are based on mundane experiences such as doing charitable deeds; heeding one's marriage vows; appreciating the beauty of nature; accepting the responsibilities of parenthood; caring for pets; protecting the environment; visiting the ill; forgiving those who have harmed one; maintaining good relationships with parents, relatives, and neighbors; caring for the integrity of one's body; cheering for a favorite sports team; supporting the troops, in whatever fashion one envisages; engaging in civic affairs; developing hobbies; maximizing one's talents and special abilities; refraining from illegal activities; not drinking to excess; and so on. No one is perfect, however, and a person does not have to be a saint to lead a mostly spiritual life. Very few people can follow all of these precepts all of the time. There will be occasions when laziness, envy, affairs, divorces, drunkenness, and other problems will appear on the scene. When this occurs, it is the psychiatrist's role to minimize their deleterious effects and provide advice to help a patient get back on track.

THE RELIGIOUS LIFE. For some patients, a spiritual life includes religion, a facet of life that provides for both much that is wonderful and much that is horrible in human experience. In his 1902 masterpiece *The Varieties of Religious Experience,* the psychologist William James went so far as to declare that religion is an essential organ of a person's life and that it performs a function that no other portion of human nature can fulfill. He wrote that during a special state of mind known only to religious persons,

> What we most dreaded has become the habitation of our safety, and the hour of our mortal death has turned into our spiritual birthday. The time for tension in our soul is over, and that of happy relaxation, of calm, deep breathing, of an eternal present, with no discordant future, has arrived. Fear ... is positively expunged and wasted away.

A contrary view is articulated in Sam Harris's 2005 book *The End of Faith: Religion, Terror, and the Age of Reason,* in which he provides voluminous accounts of the historical and current episodes of bloodshed and terror spawned by religion, such as the killing of heretics and the

suicide bombing of innocent civilians.

When psychiatrists are sure that religion is not a factor in patients' psychopathology and that their participation in religious activities fosters leading a spiritual life, then patients should be encouraged to follow their religious proclivities. Many patients find much solace simply in reading holy scriptures, even though most sacred books are difficult to understand, contain conflicting messages, and may even lead to invidious and harmful acts if interpreted concretely or wrongly.

There is much to be said for the value of participating in organized religious activities and holding religious beliefs. Certainly, each religion imposes its particular order on life and provides rules for living that, if followed, usually result in the belief that an externally felicitous afterlife will be achieved. It gives hope for all impoverished, unfortunate, disfigured, badly treated, sick, lonely, and dying people that they might be happy and blessed in the next world, be it heaven, Nirvana, or the happy hunting grounds. Questions about the meaning and purpose of life arise particularly in old age and during palliative treatment. In such situations, believers want to hear religious explanations rather than medical clichés. Questions about faith also are common in traumatized patients who feel abandoned by God and who experience survivor guilt. Wounded soldiers may feel that they deserve to be tormented as a punishment for their wartime activities. A traumatic event may become so central to victims' lives that, especially in times of crisis or severe stress, they re-enact it in their minds or, more tragically, in their behaviors as a pathological form of worship. A religious approach to their therapy is often needed to provide a more benign object of worship.

By belonging to a caring religious community, persons can count on emotional and material support in times of crisis, but the largest study on religion and coping with crises found that, "On the face of it, people appear to be able to cope just as effectively without religion as with it." Furthermore, religious rituals in times of crisis were helpful in 40 percent of cases and harmful in 23 percent. It would seem that even for the true believers, the faith, hope, and compassion offered by religion might be necessary but not totally sufficient for coping with crises or dealing with illness.

There is a growing literature demonstrating some health advantages for persons who attend religious services. One explanation is that better physical health leads to an increased ability to attend services. It may reflect a placebo effect or simply a correlation. There is little information comparing the relative health advantages among religions or branches of religions: Do Muslims have greater health advantages than Christians, Jews, or Hindus? How do Baptists compare to Catholics or Pentecostals? At any rate, depending on the results of a spiritual assessment, psychiatrists should encourage church-going patients to continue their

religious practices. Lonely, shy, or morally lost persons who have stopped attending religious services for any number of reasons (panic attacks, getting caught up in other activities, etc.) but for whom religion was once important should be gently encouraged to at least consider searching for a church and a congregation where they feel comfortable.

Prayer is the most common form of religious worship, and there is no reason for psychiatrists to dissuade patients from praying because it does no harm but rather offers hope and consolation. Hands-on prayer often provides comfort and temporary relief from some illnesses and can be explained by psychological mechanisms. Studies of prayer-at-a-distance in which prayed-for sick persons are unaware that they are being prayed for implicitly test a supposed supernatural mechanism. Such studies have not demonstrated any advantages to such subjects. A major study of 748 patients with coronary artery disease who were undergoing certain cardiac procedures showed that neither prayer-at-a-distance nor bedside MIT (music, imagery, and touch) therapy had a significant effect on the primary clinical outcome measure. A confounding variable in these studies is that non–prayed-for control group subjects may be offering their own prayers or know that relatives, friends, or even hospital staff are praying for them.

For Christians, the New Testament contains seemingly difficult-to-understand comments on prayer. In three instances, the Gospels promise that believers who pray to God will have their prayers answered. When prayers are not answered, however, the explanation offered is that even Jesus was unable to convince God the Father to spare him from his crucifixion, recognizing that God's will must take precedence. In other words, God only answers prayers that conform to his will, and there is no way to convince him otherwise. Whatever happens *is* God's will. Nonetheless, Christians continue to offer intercessory prayers for themselves and for others. Pastoral and Christian counselors often pray with patients, as do some psychiatrists. As a general rule, however, when patients ask psychiatrists to pray with them, the most judicious response is for psychiatrists to say that they will witness the patients' prayers in respectful silence (Fig. 31.6–6).

Domesticity, particularly in Christian families, may be threatened when a family member is discovered to be a homosexual. Either the homosexual person, who may be depressed, or his angry family members may seek psychiatric treatment because in their tradition, homosexuality is reviled as an unnatural, repulsive condition. For Christians, the basis for this attitude is biblical. The book of Leviticus contains laws that God presented to Moses; verse 20:13 reads, "If a man lies with a man as he lies with a woman, both of them have committed an abomination. They shall surely be put to death. Their blood shall be upon them." Death sentences also are proscribed for adulterers, men who sleep with their father's wife or their daughter-in-law, people who mate with animals, and fortune-tellers. The

hundreds of Levitical laws condemn many behaviors, such as trimming one's beard; sitting on a seat that a menstruating woman has sat on; letting defective people such as those who are blind, dwarfs, have marred faces, very long limbs, or eczema defile the Temple sanctuary by their presence; touching a mouse; eating shellfish, rabbits, or pork; and wearing garments made of a mixture of linen and wool.

FIGURE 31.6–6. St. John writing his Gospel, as portrayed in a painting in St. Isaac's Cathedral in Saint Petersburg, Russia. The eagle was his representative animal and indicates the spreading of John's Gospel throughout the world. (Courtesy of Ryan C. W. Hall.)

Psychiatrists should at least be aware that the concept of homosexuality as currently understood did not begin to be formed until the 11th century, when St. Peter Damian coined the word "sodomy." Before then, persons who were intimate with members of their own sex were simply regarded as persons with special sexual preferences. Other biblical references to "homosexuality" are ambiguous and can be interpreted in differing ways. For example, for at least two millennia, the destruction of the city of Sodom was attributed to the lack of civility and hospitality to God's angels rather than to homosexual activity. John Boswell's *Christianity, Social Tolerance and Homosexuality* and Armando Favazza's *PsychoBible* are books that might be consulted by psychiatrists and patients for more details on this issue.

Psychiatrists should also be aware that, according to Christian theologians, Jesus overturned the Levitical laws and that Christians are neither expected nor obligated to follow them. Some theologians, however, hold that only the ceremonial laws were overturned. Even if this were the

case, Christians have no problem with men being in close presence to menstruating women, men trimming their beards, and people eating pork, and they certainly do not condemn adulterers or fortune-tellers to death. The persistent Christian fixation on homosexual behavior as unnatural and on homosexuality as a deliberately chosen preference is irrational from the perspective of psychiatry, which dropped homosexuality from its list of mental disorders in 1973. Popular Western culture is showing more and more an acceptance of homosexuality, along with the realization that both homosexual and heterosexual persons may lead spiritual lives.

Psychiatrists are often called on to treat abused woman and should be alert to the possible misuse of religion in such situations. The Semitic religions originated in patriarchal cultures, and many of their beliefs and practices reflect a belief in the superiority of men over women. The book of Ecclesiasticus, found in Catholic and Greek Orthodox Bibles, states, "From a woman (Eve) sin had its beginning and because of her we all die." In the book of Numbers, women are classified with chattel property. In the New Testament, Paul notes that "the head of every woman is man," and this view is repeated in various writings: For example, "Wives, submit to your own husbands, as to the Lord" (Ephesians 5:22), and "Wives, be submissive to your own husbands" (1 Peter 2:3). In the nonbiblical Gospel of Thomas, Peter the disciple says to Jesus, "Let Mary [Magdalene] leave us, for women are not worthy of life." Jesus replies, "I myself shall lead her in order to make her male, so that she also may become a living spirit like you males." St. Augustine, four centuries later, settled things by declaring that man, either alone or when joined by his wife, is the image of God, whereas women alone are not.

Some Christian, Jewish, or Muslim men may misuse religion by citing sacred texts to rationalize the humiliation and abuse of their wives. By using passionately pious appeals, abusers may gain the support of their in-laws and congregations. Psychiatrists who treat the victims of such persons may be portrayed as evildoers who subvert the will of God unless they endorse the status quo. However, if therapeutic efforts to ameliorate the abusive situation fail, then divorce must be considered, even though the patient may find herself estranged from her family, her congregation, and maybe even her religion. Escaping from an abusive relationship is a spiritually wise decision.

Specific Spiritual Therapies. Alcoholics Anonymous (AA) exemplifies the clinical efficacy of treating a specific illness by leading a spiritual life. The movement started in 1931 when a wealthy American alcoholic was unsuccessfully treated for a year by Carl Jung. Jung advised him to place his hope in a spiritual experience, which he eventually found in the Oxford Group, an interdenominational organization devoted to reexperiencing the dynamism of Christianity's early days through inspirational talks, group

discussion, and mutual support. His religious conversion and the process through which he achieved it was a founding moment of AA.

One of his alcoholic friends also experienced a conversion and in 1934 tried to convince an alcoholic stockbroker named Bill Wilson that religion was the answer. Wilson responded by going on a drinking binge, was hospitalized by a psychiatrist, and cried out for God to show himself. He then experienced a moment of Spirituality: "Suddenly the room lit up with a great white light. I was caught up in an ecstasy where there are no words to describe." He pictured himself on a mountain in a Spiritual windstorm. "And then it burst on me that I am a free man." As the ecstasy subsided, he experienced,

> A new world of consciousness. All about me and through me there was a wonderful feeling of Presence and I thought to myself, "So this is the God of the preacher." A great peace came over me and I thought that no matter how wrong things seem to be, they are all right.

AA is not a religion, although it comes close, and some critics have called it a religion in denial. It is not a Christian organization either, but its roots are in Christian teachings. Each of its famous 12 steps contains a spiritual message, such as the necessity for alcoholics to turn their will and lives over to the care of God as they understand him; to make a searching and fearless personal inventory; to make direct amends to persons that they have harmed; and to carry the AA message to other alcoholics and to practice these principles in their lives.

According to AA, alcoholics' acceptance of self as human is founded in their rejection of any claim to be more than human. AA members must accept the limitation of their being: An alcoholic cannot drink alcohol safely. Formal religions take personal limitations as a problem that can be overcome by participation and belief in the mechanisms of salvation offered by a particular church. Religion aspires to perfection and contains what it considers to be absolute truths. However, organized religion promised too much for the founders of AA. When asked why AA separated from the Oxford Group, Bill Wilson replied, "The Oxford Group wanted to save the world, and I only wanted to save the drunks."

Another example of a psychiatric therapy that utilizes guidelines for achieving a spiritual life is found in a series of DVD modules, *The Happy Life—Voyages to Well Being,* developed by the psychiatrist C. Robert Cloninger. The modules attempt to enhance character traits associated with frequent positive emotions and to enhance a sense of hope and mastery for self-directedness, kindness, forgiveness, cooperativeness, mindfulness, and finding spiritual meaning. The therapy begins with a focus on what makes patients happy or unhappy, the experience of well-being, and awakening their physical and spiritual senses. Patients then observe and strive to elevate their thoughts and human relationships, chart their maturity, and contemplate their being. The therapy ends with modules on stress

reduction, quieting fears, observing the power seekers in their life, contemplation of mysteries, and constant awareness. Although a happy life for patients is supposed to be a result of the therapy, there is a branch of psychology devoted to searching for what actually constitutes happiness. The theologian–physician Albert Schweitzer, who became depressed after concluding that Jesus had not really been resurrected, noted that "Happiness is nothing more than good health and a bad memory." This new form of therapy is based on some solid research findings; for example, low self-directedness scores on the Temperament and Character Inventory indicate a vulnerability to major depression, whereas high scores predict a rapid and stable response to both antidepressant medication and cognitive–behavioral therapy. Whether this therapy will be widely used or prove to be effective is still unknown.

Evidence from psychological studies of personality points to a biological consideration for a partial understanding of spirituality. A widely accepted model of fairly stable core personality traits rooted in biology identifies five dimensions: Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. Each dimension has associated descriptors or facets such as degree of emotional stability, compassion, self-discipline, and so forth. Characteristics of the five dimensions arise because of environmental forces and result in the development of beliefs, attitudes, goals, and a sense of self. Adaptations come into play during a person's lifetime experiences and are translated into behaviors and emotional reactions. In research findings, the openness to Experience, Extraversion, and Agreeableness correlate highest with spirituality. A recent study among persons from differing faith traditions suggests that spirituality may be a sixth core personality dimension. In this study, the core personality traits had an impact on health outcomes but spirituality and religious subscales did not. Only the forgiveness subscale predicted better mental health outcomes. Forgiveness is one of Vaillant's positive emotions that comprise spirituality, but it is possible to regard forgiveness as a coping strategy that is related to but distinct from spirituality.

The "hard wired" for God/spirituality concept is popular, especially in the lay press, and has given rise to the pseudoscientific field of neurotheology with notions of God genes, neurons, and neurotransmitters. Monoamine neurotransmitters have been identified in a popular book as the molecules partly responsible for self-transcendence and spirituality. A variation in the vascular monoamines transporter 2 gene in some persons supposedly enhances a tendency for spirituality. However, the book ends with a strange disclaimer: Spirituality may be partly genetic but genes are like verbal cooking recipes, and what you end up with depends upon how you interpret and execute the directions.

Most people believe that almost any type of meditation does no harm and probably does some good just by reducing stress. But only some types

that are readily accessible, such as mindfulness, hold the promise of achieving spirituality. In most instances, mindfulness is practiced in group workshops, although it may be utilized in conjunction with cognitive behavioral therapy. The attainment of mindfulness spirituality is a gradual process and may take years of hard work. In 2014, the US Agency for Healthcare Research and Quality issued a review of 41 meditation trials with matched control groups. The review revealed that mantra meditation had very little, if any, effect on psychological stress and well-being outcomes while mindfulness meditation had a moderate positive effect on anxiety, depression, and pain. The conclusion was that meditation programs were not more effective than exercise, progressive muscle relaxation, cognitive behavioral therapy, or drugs in changing health outcomes. The review did not address the issue of spirituality.

REFERENCES

Cardena E, Lynn SJ, Krippner S. *Varieties of Anomalous Experience*. 2nd ed. Washington, DC: American Psychological Association; 2014.

Cook C, Powell A, Sims A, eds. *Psychiatry and Spirituality*. London: Royal College Psychiatry Publications; 2009.

*Crescenti C, Di Bucchianico M, Fabbro F, et al. Excitatory stimulation of the right parietal cortex lessens implicit religiousness/spirituality. *Neuropsychologia*. 2015;70:71–79.

de Souza M, Francis CJ, O'Higgins-Norman J, Scott D. *International Handbook of Education for Spirituality, Care and Wellbeing*. New York: Springer; 2009.

*Favazza A. *PsychoBible: Behavior, Religion, and the Holy Bible*. Charlottesville, VA: Pitchstone; 2004.

Galanter M. *Spirituality and the Healthy Mind*. New York: Oxford University Press; 2005.

*Harris S. *Waking Up: A Guide to Spirituality Without Religion*. New York: Norton; 2014.

*Johnstone B, Bodling A, Cohen D, Christ S, Wegrzyn A. Right parietal lobe-related "selflessness" as the neurophysiological basis of spiritual transcendence. *Int J Psychol Religion*. 2012;22(4):267–284.

Johnstone B, Gloss B. Support for a neuropsychological model of spirituality in persons with traumatic brain injury. *Zygon*. 2008;43(4):861–874.

*Johnstone B, Cohen D, Konopaki K, Ghan C. Selflessness as a foundation of spiritual transcendence. *Int J Psychol Religion*. 2016;26:287–303.

Kabat-Zinn J. *Coming to Our Senses*. New York: Hyperion; 2006.

Koenig H, Huguelet P. *Religion and Spirituality in Psychiatry*. Cambridge: Cambridge University Press; 2009.

Oviedo L. Religious attitudes and prosocial behavior. *Religion Brain Behavior*. 2016; 6:169–184.

*Pargament K. *Handbook on Psychology, Religion and Spirituality*. Washington, DC: American Psychological Association; 2013.

Peteet JR, Lu F, Narrow W. *Religious and Spiritual Issues in Psychiatric Diagnosis*. Washington, DC: American Psychiatric Publishing; 2009.

Purzycki BG, Willard AK. MCI theory: a critical discussion. *Religion Brain Behavior*.

2016;6:207–274.
*Vaillant G. *Spiritual Evolution*. New York: Broadway Books; 2008.
Van Cappellen P. Effects of oxytocin on spirituality and emotional responses to medication. *Soc Cog Affective Neuroscience*. 2016;10:1579–1587.

▲ 31.7 Physician and Medical Student Mental Health

KHLEBER CHAPMAN ATTWELL, M.D., AND GRACIANA LAPETINA, M.D.

INTRODUCTION

The Hippocratic oath charges the physician with an ambitious set of ideal character traits, not limited to generous teacher, sound diagnostician, careful pharmacologist, doer of no harm, and one who values the patient's good above all other. It also highlights humanistic qualities of compassion and humanism. The imperfections of the profession—behaviors of the burned-out, depressed, fraudulent, drug-dependent, or boundary-crossing physicians do not receive formal recognition in these ideals—even if the very medical training and practice of medicine jeopardize the physician's health. Given the tension between these ideals and these actualities, and given the inevitable unsustainability of any ideal, where is the professional embrace of the physician to become the patient? Or, especially—the mental patient?

DISCIPLINES INVOLVED

Medical student and physician mental health care is intrinsically multidisciplinary. Clinical medicine and psychiatry often provide the front line of care through private or employee health sectors. Psychoanalysis provides unique windows into the understanding of the developmental histories and mental conflicts of physicians, as well as into effective treatments of disorders. Psychology offers reliable, quantifiable measures for examining the particular cognitive and emotional challenges unique to the physician suffering from a learning disability or other set of mental barriers to the acquisition of knowledge. National physician organizations such as the AMA and the APA provide critical epidemiological data with accompanying position statements and ethical guidelines. Medical schools assume the responsibility to train and educate their students to become both knowledgeable and ethical, professional doctors; disciplines of professionalism and doctor–patient communication begin their impact at this stage of training. Governmental agencies at the county, state, and national levels interface with potentially impaired physicians via state licensing boards, civil or criminal courts of law, and treatment programs. Professionals from disciplines as wide ranging as fitness and nutrition to sleep hygiene or the chaplaincy can prove essential to the health maintenance of the physician. In the broadest sense, public health efforts to restore medical student and physician health seek to rehabilitate doctors who care for society while at the same time protecting patients from doctors who may do harm.

Clinical versus Administrative Psychiatrist

The distinction between clinical and administrative psychiatrist is critical. If an ounce of prevention is worth a pound of cure, then seeking clinical attention when symptoms of a problem first present makes treatment a private matter that can restore health and its optimal functioning. Not infrequently, physicians speak with their peers to make such referrals possible when first observing unusual behavior in a colleague. Many physicians will not seek treatment voluntarily, however, prompting their residency training director, department chair, or hospital administration to make a referral for confidential, clinical care. If this intervention does not prove successful, use of an administrative psychiatrist may provide the only successful avenue for treatment. Often a forensic psychiatrist, the administrative psychiatrist is not bound by the usual parameters of confidentiality and often speaks to spouses, fellow colleagues, or relevant academic deans or hospital staff to perform a fitness-for-duty evaluation. These reports, in becoming part of one's permanent record, prove critical to halting dangerous professional practices and leverage the necessary treatment interventions. They also document the effective recovery and treatment plan linked to a safe return to the practice of medicine. One parameter that can be used to measure the efficacy of mental health services in a given community is the frequency of resort to administrative psychiatric services; if the clinical services and their referral outreach provide adequate care, the use of the administrative arm should be much less.

Medical School Administrative Psychiatrist–Clinical Psychiatrist Interface. Years of experience underscore the value of an open, well-maintained relationship with the dean of students and the administrative psychiatrist. The administrative psychiatrist becomes involved in any case in which a student's academic performance becomes linked to his or her psychiatric condition. In practice, if a medical student seeks services in a timely manner and the student mental health service functions optimally, that student need may never see either the administrative psychiatrist or the dean of students. However, the reciprocal is seldom the case: invariably, any student who sees the dean of students or administrative psychiatrist for any reason deemed to be potentially psychiatric is referred for consultation. That referral could be as benign as a learning style evaluation after failing a particular examination or as dangerous as finding a student unconscious in his dorm room unconscious from a drug overdose. In practice, students will utilize clinical services only if they know that their confidentiality is firewalled from the dean of students—the standard of care unless significant risk to the student or other is at hand (Fig. 31.7–1).

In more extreme cases, confidentiality becomes an afterthought and the administrative psychiatrist, the dean of students, and student mental

health service must collaborate. If a student is arrested while intoxicated on drugs, draws legal attention to herself while on rotation in a school capacity, is admitted to a psychiatric hospital for dangerousness to self, or has been found to be dangerously at risk from an eating disorder, more immediate attention is required. Not surprisingly, had any of these students been more open to seeking mental health evaluation and treatment earlier in the course of his or her illness, a less intrusive and more longitudinal treatment plan might have been effective. From the safety of a hospital unit or inpatient treatment facility, the return to medical school can become a fulcrum of leverage to motivate a student for treatment and to create a subsequent frame of accountability to which the student must adhere for continued matriculation. Invariably, the philosophy of this treatment shifts from that of 1:1 doctor–patient to that of treatment team where all circles of the Venn diagram become essential. If a medical school invests enough in a student to accept him or her for training, it follows that the philosophy of care is to do anything reasonable and rehabilitative to allow that student to continue training and to accomplish the original intention of working as a physician until proven unsustainable. The ethical balance of this stance swings between the rights of any student under the Americans with Disabilities Act (ADA) and the obligation to protect patients from a dangerous health care professional—just as an airline has an obligation to keep its passengers safe from a dangerous pilot.

Case Vignette: Ms. E is a 25-year-old medical student who was referred by the dean of students to the medical student mental health service for evaluation and treatment after being found by her roommates vomiting in their shared bathroom. These same roommates were worried about the drug paraphernalia and empty alcohol bottles in their shared kitchenette. Upon presentation, she weighed 89#, well below the 0.9 percent body mass index needed to diagnose her with anorexia nervosa. The student health team met with her and her family to explain the necessity of inpatient treatment were she to continue her course of study. At the eating disorder treatment facility, she was also diagnosed with mixed alcohol and cocaine dependence and began a series of overeaters anonymous, narcotics anonymous, and alcoholics anonymous 12-step groups in addition to individual treatment and pharmacotherapy to manage the compulsive ways in which she would use drugs, alcohol, food, and at sex (at times) to regulate her painful feelings of rage and grief when she had to sit still or with herself. Issues of anxiety and PTSD also emerged as she came to grips with the impact and origins of her suicidal behaviors; she was sexually abused by a trusted mentor as a young teenager, which prompted a cascade of thought in response to the deep shame she harbored from this betrayal. She came to believe that were she not "perfect" in every way, she would not be loveable. Concurrent family therapy allowed for effective confrontation with this mentor and repair within her nuclear family of damaged intergenerational rifts.

Though Ms. E re-entered the student mental health service originally in an intense manner and harboring the conviction that we only existed to make her life miserable in the name of "pretending to care," she reintegrated well into her new class after a year of intensive treatment away. Several months before her rematriculation, she moved back to New York City to secure providers and a local treatment team—including regular check-ins with the administrative psychiatrist and the medical student mental health service. Tight, contracted obligations were established to insure the return to a higher level of care was such to be necessary. To her surprise, Ms. E responded positively to the structure and intervention at hand. Through much

trial and error and countless 12-step meetings and individual treatment sessions, she used her experiences as a doctor in training to connect empathically with her patients and to provide superlative care. She went on to win the Humanism award from her class and matched in a competitive surgery residency program with plans for the subspecialty of reconstruction.

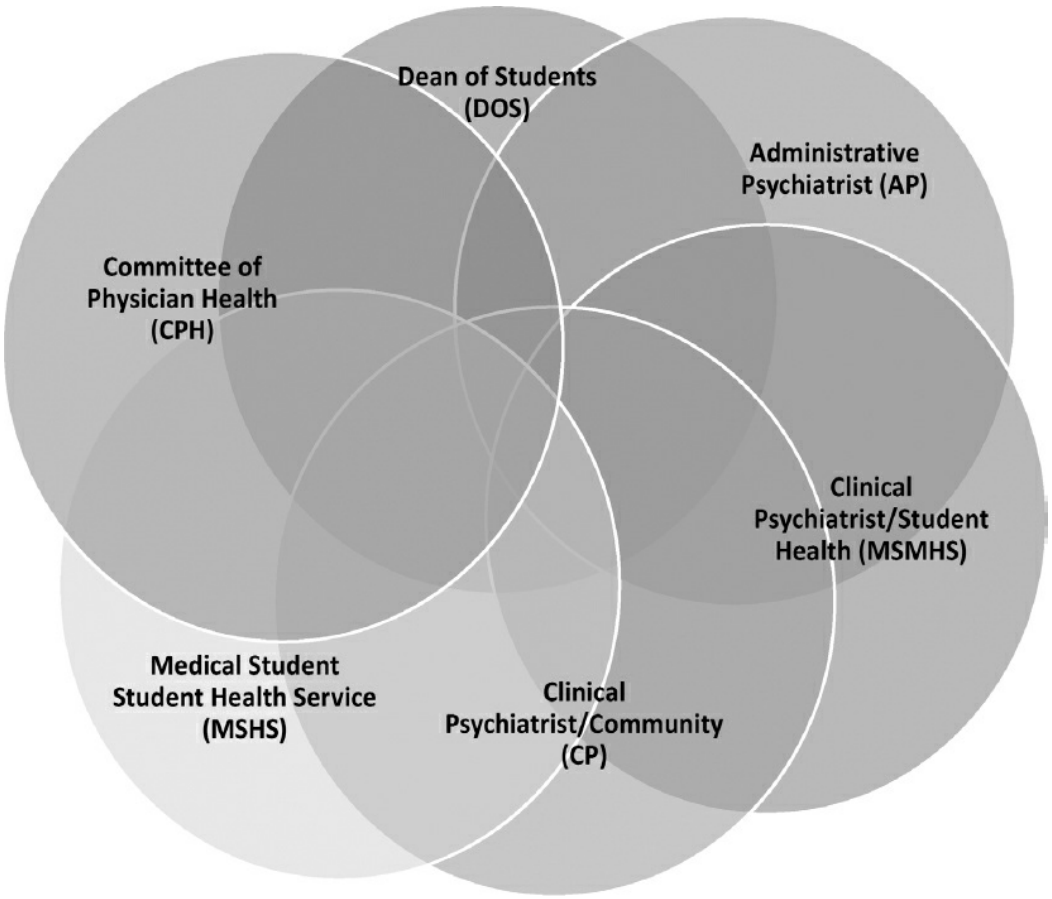

FIGURE 31.7–1. Venn diagram. Intersection of major medical student mental health parties.

Clinical Vignette: Mr. F is a 27-year-old third-year medical student referred to the medical student mental health service for vandalizing school property. He sheepishly presented for evaluation, only because he had been required to do so. He denied all overt psychiatric history of treatment, although he made references to strained family relationships, vague gaps of time or leaves of absence in his past history, to a medical condition now treated by report, and to a family history of incarceration. His peer relationships drew largely from outside of the medical center community. Shortly after his first presentation, he was found intoxicated in the laundry room of his dorm. After 911 was called and he was taken to the very emergency room staffed by his peers, he reported that he was unable to stop drinking alcohol. He was placed on a medical leave and offered multiple avenues of care in his home state. Mr. F was allowed to matriculate after having obtained documentation for his treatment back home, even if less than optimum. Shortly upon his return, he was charged with sexual harassment of a classmate, placing her in a situation deemed less than compatible with learning in a respectful environment. Though he protested his case at all levels of the school administration, the administrative psychiatrist, the dean of students, and the medical student mental health service all concurred that his psychiatric condition of alcohol dependence and documented antisocial behavior was incompatible with good academic standing. Mr. F was dismissed permanently from school and from all possible consideration of retraining at our institution.

HISTORY OF THE DISCIPLINE

The expression "Physician, heal thyself" likely stems from ancient times. More recently in the United States, formal avenues to address the issue of the impaired physician burgeoned beginning in 1969 (Florida) and 1971 (Texas) with the passage of "sick doctor statutes," which allowed for reporting of and referral for care of doctors with probable cause of distress without the actual legal allegation by a patient of harm. The number of programs providing treatment of physicians abusing substances and the accompanying statewide committees on physician health mushroomed, and these eventually became part of every state's laws. Initially focused on physicians' abuse of drugs, alcohol, or prescription narcotics and benzodiazepines, physician mental health programs have grown in scope to acknowledge the impact of untreated mood and bipolar disorders, more subtle disorders of personality that may drive the sexual exploitation of patients or the abuse of colleagues, and impairment from comorbid medical conditions. Finally, the field now fully recognizes the impact of wellness and has campaigns to prevent medical illness, psychiatric disorder, or practice burnout. It recognizes the value of education of the profession at all levels of training, beginning in medical school and continuing throughout with a constant provision of confidential access to care.

RISK FACTORS UNIQUE TO THE PROFESSION OF MEDICINE

The demands of clinical care of patients can prove exceptionally challenging and worthy of the respect bestowed on the healing professions by society. The regular grief involved with exposure to death and dying of longstanding patients; feelings of helplessness at not knowing which path to take with a patient and of limitation in the face of human nature; exhaustion from working long or emotionally taxing hours with the accompanying deprivation of sleep; and invitations to allow responsibility for clinical care to assume the weight of guilt when patients' conditions decline precipitously all contribute to the occupational hazards of the profession. Beyond the immediate clinical level, complex communications with perfection-seeking colleagues, demanding patients, or cultural and language boundaries unfamiliar to the physician add to the background stress of health care delivery. A professional culture unwelcoming of medical error, high levels of work–home interference, limited control over call schedules, and an increasing educational debt burden further magnify the impact of any given stress. Just as each of these factors creates stress, each also illuminates directions for better physician self-care and for change within the culture of medicine.

PHYSICIAN WELLNESS AND PREVENTIVE HEALTH

The physician's healthy mind, body, and sustaining relationships with others are the most essential instruments for the sound practice of medicine. Not only do healthy physicians practice more preventative medicine with their patients; decreased physician wellness has also been linked to adverse outcomes for patients, poor prescription habits, ineffective ordering of tests, and decreased patient adherence and/or satisfaction. These trends make sense: A doctor who smokes or is overweight will have a harder time using his personality to sell the patient on the idea of quitting smoking or eating more in moderation; a stressed, anxious physician may rely more on laboratory tests and prescriptions to intervene with a patient when simply engaging the patient in a deeper dialogue might get to the heart of the matter.

On one hand, specific data provide a meaningful comparison of physicians with the rest of society and convey that, in essence, physicians take relatively good physical care of themselves with respect to diet, exercise, and smoking. In parallel, mortality trends reflecting better health than population averages, although perhaps not as good as those with a high socioeconomic status. According to the Women Physicians' Health Study (WPHS), data on female physicians show them to be superior in health to their male counterparts. Only 4 percent of physicians smoke, compared to 25 percent of the general population; otherwise, compared to the general population, most physicians drink more frequently but more moderately; eat more fruits and vegetables and consume less fat; exercise more regularly but still less than the recommended 30 minutes three times per week; and engage in health screening tests such as cholesterol measurement, fecal occult blood testing, sigmoidoscopy, or mammography. Male physicians die later than do all male peers, male professionals, and male lawyers. Rates on female physicians' mortality are invalid due to the ninefold increase in women in medicine from 1950 to 1990. Either the healthy individual chooses medicine as a career or medicine as a career helps the individual to make healthier lifestyle choices. Naturally, each notion could be true and mutually reinforcing of the other.

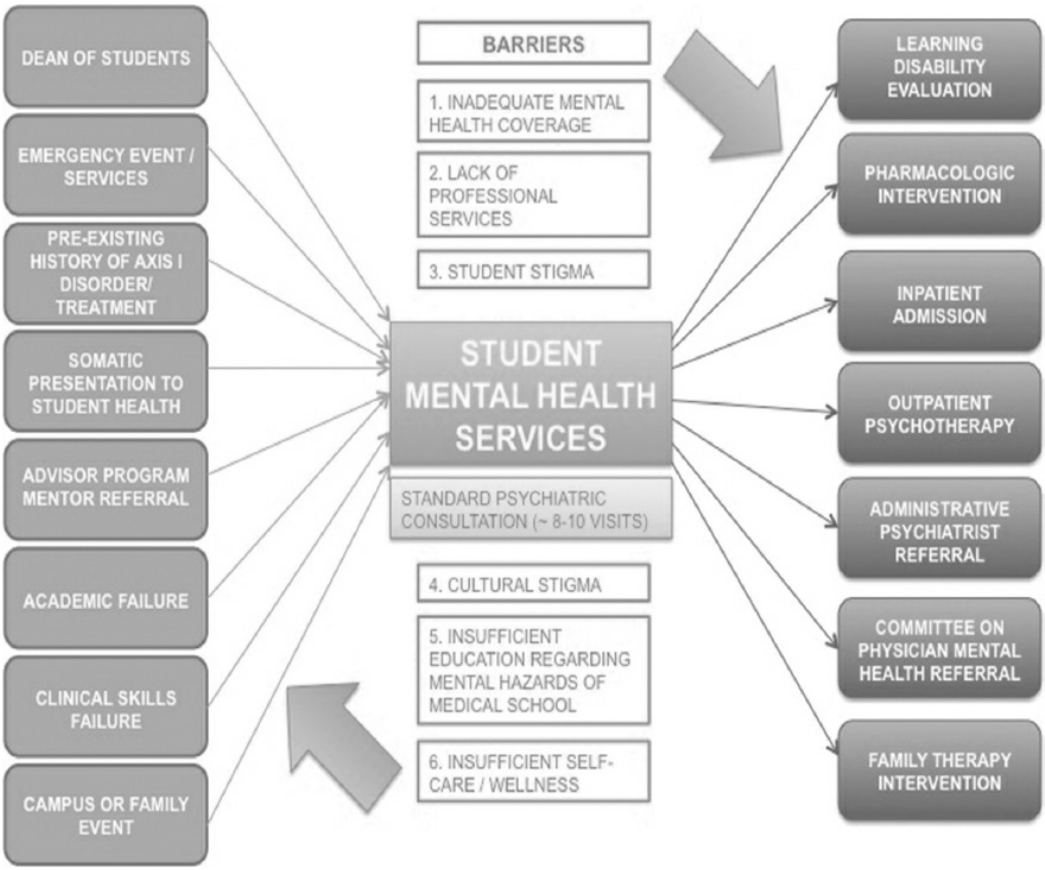

FIGURE 31.7–2. Flow diagram of various intervention avenues and outcomes.

On the other hand, numerous examples stand out to show how physician physical health can be jeopardized. Sleep deprivation and regular 80-hour work weeks characterize many fields of medicine. Physicians are more likely to die from suicide or accidents than their peers, and multiple studies demonstrate that a majority of physicians treat themselves as patients in some capacity, from self-prescription to self-chemotherapy. Physicians only report having a health care provider one-third to two-thirds of the time, and even then, that provider might be their spouse or themselves.

Beyond the meaning of any statistic measuring physician use of substances or mortality, regular practices have proven timeless in preserving physician mental availability to be present with patients in their care. Learning to deliver bad news effectively; allowing personal time to grieve losses sustained at work; limiting workload; and meeting with a safe, confidential peer group to discuss the vicissitudes of patient care and its outcomes all protect the mind and life of the practitioner. Developing religious, spiritual, or other lifestyle interventions that promote balance likewise provide avenues for the psychic metabolism of stress encountered in the workplace. It has even been suggested that the most useful intervention a physician can implement to protect himself or herself from any kind of boundary crossing, boundary violation, or other inappropriate

use of a patient for the physician's needs is the maintenance of a mutually satisfying marriage or other meaningfully intimate relationship (Fig. 31.7–2).

PERSONALITY TYPES, PHYSICIAN TEMPERAMENT, AND RISK FACTORS

Physicians choose the career of medicine for myriad reasons, but one immediate root of this career choice that can become symptomatic is prior trauma. It has been said that only the wounded can heal, and it seems beyond coincidence that psychiatrists often have mentally ill family members; that cardiologists may have family histories of cardiac disease; that pediatricians may have suffered in some ways as children, including from medical illness; and that surgeons may enjoy the sublimation of violent urges into a sanctioned, socially esteemed form of cutting into another's body. Becoming a physician to master a prior trauma can be adaptive, and the list of possible traumas is almost as long as the number of physicians. But some particular ones appear repeatedly: Competition with a physician parent; childhood illness, hospitalizations, or surgeries; traumatic immigration; childhood loss of a sibling or parent; coping with a learning or physical disability; fear of passive submission to a sexual, physical, or emotional abuser; emotional neglect in the household of cold or alcoholic parents; and strong demands on behalf of the child's parents to become a physician to meet sociocultural expectations.

Two particular stories capture this mobilization of drive.

Alfredo Quinones-Hinojosa described his amazing transition from migrant farmer and illegal alien to professor of neurosurgery. Determined to fight his way out of poverty, he learned English, worked as a welder, and became a college student to learn English. Along the way, he fell into a tank of liquefied petroleum gas and woke up in a hospital, seeing a person dressed all in white and feeling secure and protected in the hands of the hospital's doctor. In time, against many odds, and with the encouragement of a string of mentors, he worked his way through neuroscience, medical school, and neurosurgery. He cited the influence of his first role model, his grandmother (and medicine woman of his Mexican village) in allowing him to keep a hold on his dreams and actualizing them.

Dr. Diana Wood, now an anesthesiologist, came to medicine via the murder of her then–pediatric resident husband. She used medical training as a way to manage the feelings of such inestimable loss.

These stories speak to the power of a secure mind to adapt creatively and constructively to overwhelming trauma, the very force that can be so powerful when mobilized within the context of treating a medical student or physician in distress.

Although these factors may lead one to the career of medicine for conscious or unconscious reasons, challenges can arise when the psychology of the doctor becomes the psychology of the patient. It is in these moments when listening for the patient's identification with a close

physician parent, mentor, or childhood personal physician can help navigate the consultation. In particular, physicians struggle when faced with feelings of dependency, helplessness, or lack of control—feelings so common to the role of the patient—preferring instead to seem objective, problem solving, or emotionally distant.

In his writings about medical marriages, Glen Gabbard detailed the psychology of the physician. He documented the emotional distance, perfectionism, compulsiveness, need to control others, unfulfilled dependency yearnings, doubt, guilty feelings, and exaggerated sense of responsibility characteristic of physicians. He characterized their search for love and approval via their patients, colleagues, and the world at large, often at the expense of what a spouse or child might provide or might need. Medicine is a jealous mistress. In fact, many physicians detach sexual relations from emotional intimacy, approaching sex as a purely physical event. If adult physicians can engage a heroic rescue fantasy by "saving a patient" and thus involve themselves with the heady matters of life and death, then they can attempt to feel that they are valuable in the eyes of a longed-for parent. Karl Menninger speculated that the physician's constant struggle to check for laboratory errors and other potentially lethal mistakes provides a concealment of aggressive, sadistic wishes. George Vaillant hypothesized that physicians use the role of caregiver to provide for the care that they longed to receive as children. He documented overly chaotic childhoods and adult emotional instability in the lives of internists, psychiatrists, obstetrician/gynecologists, and pediatricians on the front lines of primary care. Vaillant also highlighted the frequent physician use of altruism, masochism, reaction formation, and hypochondriasis as defense mechanisms. Lane Gerber and John Rhoads focused on giving to others as an avenue to conceal hostility, conscious or otherwise, that physicians may feel toward their parents and on workaholism as a means of conquering the guilt felt at harboring such anger.

WOMEN IN MEDICINE

Malkah Notman has written about the differences experienced by women in medical training and practice, which are becoming increasingly significant as the number of women becomes equal to the number of men in the profession while the expectations of training still prioritize work as the dominant pursuit of the trainee's life. The most obvious issue is the conflict over pregnancy and raising a family while practicing medicine; women can still feel guilty or less than committed to their field if they choose to have children. Sexism, when it exists, can induce further risks to self-esteem and subsequent depression or highlight awkward feelings about prioritizing one's family and one's career. Male physicians are naturally allowed to blend the aggressive nature of achievement with the more feminine aspect of nurturing and caretaking; female physicians may be

seen as good caretakers while being not allowed the same avenues for focus on competitive ambition. Women physicians may struggle in unique ways as they manage the increased value of relationships and attachments compared to men; their greater degree of expressiveness about their emotions; and their tending more toward conformity. Conforming to masculine standards of not seeming vulnerable to avoid being seen as weak can prompt women physicians to avoid the proper attention for any mental health problem. Feeling conflicted about caring for patients and children can invite feelings of guilt from both angles. Moreover, should a woman physician choose to start a family, numerous specific issues invariably must be addressed. Pregnancy stirs up multiple issues for the doctor's patients, coworkers, and herself. Interruptions of call schedules can bring resentment and envy from coworkers. Balancing child care and patient care usually prompts mixed feelings and practical difficulties. Integrating a doctor's schedule into that of a family can trigger spousal and filial resentment. For the woman physician with a family who chooses to work part-time, the natural delay in career progression may bring less income to pay burdensome student loans in real terms and feelings of exclusion from a peer group in emotional terms.

MINORITY PHYSICIANS

It is not surprising that the data on minority physicians are scant. Much as minorities struggle to find a solid identity within the majority world of their profession and its culture, the profession must now struggle to study the specific attributes unique to any individual minority subgroup of medical students and physicians. Michael Myers and Glen Gabbard began ground-breaking work in this field and offered a framework for the assessment of cultural background in a psychiatric presentation. They focused specifically on the plight of African American; Hispanic; Asian; Asian Indian; lesbian, gay, bisexual, and transgendered (LGBT); and international medical graduate (IMG) physicians. They illustrated via case examples the complexity of the struggle to forge one's sense of identity when faced with obvious external and life history differences. They illustrated how what makes one unique on one hand can feel so isolating and burdensome on the other. Myers and Gabbard also call on the evaluating psychiatrist to be aware of the powerful roots of shame in its connection to depression, anxiety, substance abuse, and suicide. When the majority culture does not accept difference (be it of race, attitude toward mental illness, or homosexuality, as a few of many examples), the individual's proclivity to internalize such shame as the symptoms of guilt, worthlessness, or depression can be overwhelming. They also highlight the unique types of countertransference that can emerge when evaluating a medical professional of difference. Although some features may seem more obvious (unconscious racism, for example), others may be more subtle. The

profession of psychiatry can tend to be overwhelmingly accepting and celebratory of difference, which, paradoxically, can make it harder to appreciate the patient's internalized hatred of his or her difference. These examples are few but require the field to be more mindful and proactive with regard to colleagues who share an overtly or covertly different cultural background, identity, set of assumptions of normality, or family or ethnic history.

PHYSICIAN MENTAL HEALTH

Although physicians take good care of themselves physically, they may struggle to do so emotionally. Each of the syndromes outlined in what follows carries a particular profile and each of the vignettes is unique, but few of the features occur in isolation. For instance, the burnout syndrome often leads to depression, which often leads to suicidal thinking or substance abuse. Being reported to a state medical misconduct administration may trigger more anxiety and fear than relief at the outset. An untreated mood disorder may seem like a disorder of personality until properly addressed. Like any other patient, a physician patient may live a life much like his or her peers until one day a profound life event or professional upset occurs.

Stress, Overwork, and Burnout

The concept of overwork is simple: Exceeding one's endurance and/or recuperative capacity to work. Much as a spring has a particular stretch coefficient that, once crossed, leaves the spring permanently damaged, human bodies have limits. The easiest definition of burnout is that of all of the symptoms of depression in some permutation but in a manner that is time limited to the hours spent on the job. Relief of symptoms as one leaves the hospital is the *sine qua non* of burnout. Although many of the nation's employees are protected from overwork by union regulations or federal laws (such as autoworkers or airline pilots), physicians who are self-employed have to create their own set of limits. More specifically, patients with burnout experience fatigue, irritability, sleep disturbances, difficulty concentrating, confusion, depressive affect, and any number of traditional psychosomatic symptoms such as headaches, back pain, diarrhea, constipation, or unusual neurological sensations such as dizziness, tremor, or shortness of breath. Often affecting the most driven and successful of individuals, burnout begins with a characteristic event that prompts the person to strive even harder to achieve some particular goal.

Whether driven by religious fervor, an overly punitive superego, unresolved oedipal competition with one's father, residual sibling rivalry, sociocultural expectations, identifications with deceased family members who were workaholics when alive, a narcissistic sense of self-believing that

one is irreplaceable at one's work, an unassuaged force of unconscious guilt, or any other number of dynamics, as the physician strives to work harder and harder, he or she neglects basic needs such as sleeping, eating, or socializing. The realization that something is wrong becomes linked in a distortion to believing that more work is needed to achieve one's goal. The spiral continues to cycle between denial of one's needs, progressive isolation, depersonalization, a deeper sense of inner emptiness, and finally a full-blown depression. Although burnout can strike any hard-working physician across the life cycle, residency training is famous for eliciting scores of burned-out doctors. In addition to the long hours and intense patient care demands, the vulnerability of attempting to master one's craft, often as one attempts to master the beginnings of family life, can make simple rest and relaxation insufficient to mend the rifts of fray so essential to full recovery.

Liselotte Dyrbye and colleagues have recently studied the impact of burnout on suicidal ideation and attempts in the US medical student population: their work is compelling and poignant. While more is known about the correlations between untreated depression, manic depression, and/or substance abuse with suicidal ideation and attempt, little is known about the relationships between medical student and physician burnout and self-injury. Their results indicate a high prevalence of suicidal ideation among US medical students (10 percent) and of burnout (50 percent). Furthermore, they divide the syndrome into its multiple facets of emotional exhaustion, depersonalization, and low personal accomplishment and correlate each with risk of suicidal ideation. They illustrate that the reduction in burnout symptoms correlates with less suicidal ideation. The implications of this work are profound in their clear, comprehensive assessment of the extent of depressive symptoms and burnout in medical student populations. If indeed burnout is a major risk factor for both depression and suicidal ideation, as this study suggests, then US medical schools have all the necessary data to substantiate any intervention that targets student well-being and provision of mental health services and the study of its efficacy.

Depression and Manic Depression

Mood disorders are among the most frequent mental health presentations, and physicians take on depressive symptoms in unique ways that make them more camouflaged than other patients. The obvious symptoms of depression (loss of appetite, sleep disturbances, sadness of mood, lack of pleasure, crying spells, morbid preoccupation, suicidal rumination and planning, difficulty concentrating) and of mania (sleeplessness, distractibility, racing thoughts, grandiosity, agitation, irritability, use of nonprescription substances for regulation of sleep and mood) all apply as readily to the physician as to anyone suffering from a mood disorder.

Perhaps the greatest barrier to detecting depression in medical students and physicians is their shell of pride. Anxiety, fear of stigma, and reluctance to accept depressive or manic symptoms as evidence of a mood disorder layer the complexity of genuine reflection on the consequences of one's behavior. Hypochondriacal physicians can encourage fruitless and expensive medical workups to no avail. Depressed doctors can spend hours at home doubled up in a fetal position crying for reasons they have a hard time even identifying, other than "something is wrong with me" but show up for rounds the next morning as if the episode had occurred in isolation weeks ago. Suicidal preoccupation can certainly be more complex or detailed in medical professionals, given their education and familiarity with the body and its anatomy and physiology. The work schedule of medicine can be used to compartmentalize uncomfortable symptoms, whereas manic energy can fuel numerous forays into patient care, teaching, or research. It is commonly reported that as long as a physician patient can be busy at work, he or she does not have to dwell on any troubling personal feelings or relationship issues. Psychologically, this characteristic wish to avoid psychic pain, when combined with an intense drive to succeed and a parallel intense sense of disappointment with one's self when the perception of failure is present, facilitates the growth factors for depressive illness and manic defensiveness or behavior.

Physician Suicide and Psychiatric Emergency

Naturally, the most morbid and most disturbing symptom of psychiatric illness, often stemming from untreated depression or manic depression, is suicide. Few acts leave as many unanswered questions and as much unresolved grief as suicide, and when a doctor kills himself or herself (and women physicians are at higher risk), spouses, children, colleagues, and scores of patients are left to pick up the pieces. The demographics of physician suicide are replete with unanswered questions and methodological issues, but several basic trends seem to hold. Roughly 300 to 400 physicians per year take their lives by their own hands, and women physicians do so significantly more than their age-matched controls. Risk factors are as follows: Being older (45 years of age and older); being white; being divorced, single, separated, or with a disruption of relationship status; having risk-taking, substance-abusing, workaholic, or impulsive personality traits; having a history of psychiatric or medical illness; experiencing a recent change in role status (through a lawsuit, bankruptcy, loss of spouse or child, loss of professional esteem); and having access to lethal medications or firearms. Psychological autopsies of suicide victims are in their infancy, and clinical experience reveals that often the act may be concurrent with an overwhelming intrapsychic stressor that feels unbearable for myriad highly personal reasons.

Whatever the final impetus to cross the bridge into the proverbial dark

night, experience also shows that prompt, thoughtful, attentive, and professional responses to suicidal feelings or behavior in medical students and physicians can save lives. The treating clinician must conquer his or her wish to deny that someone who spends so much of the day attempting to master illness and death might just as readily be able to inflict it upon him or herself. Clinically, such a stance means recommending hospitalization and accompanying the student or doctor–patient to the emergency room if necessary, contacting immediate family members to mobilize support, and educating the physicians on the regular and natural occurrence of suicidal symptoms in mood disorders, as well as on their amenability to treatment. By definition, if a patient speaks of his or her wish to commit suicide, there is a good chance that the clinician can harness the caring side of the ambivalence to view that suicide is a one-way valve with no probability of working problems afterward, a challenge physicians like to tackle. Restoring a sense of hope and a sense of perspective that has been lost when someone has the monomaniacal delusion that suicide will solve the immediate problem at hand can literally save a life that, later, when treated, might go on to care for scores of patients in compassionate ways.

The same dynamics of treating suicidal medical students and physicians applies to all cases of psychiatric emergency. Dangerous eating disorders, raging substance abuse and dependence requiring immediate detoxification or rehabilitation, flaring comorbid medical conditions, or a transient meltdown of one's ability to test reality all may require the same deftness of conviction and intervention that a greater level of care must be used to provide safe relief from symptoms.

Michael Myers and Carla Fine wrote on the aftermath of physician suicide in *Touched by Suicide: Hope and Healing after Loss*. Fine's husband, a urologist at NYU School of Medicine, killed himself at the early peak of what promised to be a bright career. In particular, Fine wrote about the isolation and ostracization she faced from her husband's medical colleagues after his death. Other accounts of physician suicide, both attempts and completions, anchor this life-altering event in the complexity of its biological, personal, professional, social, and family webs. These accounts drive home the magnitude of such loss, as well as the urgency of preventative measures, education, and destigmatization of mental illness within the house of medicine. They also underscore the value of outstanding and empathic psychiatric care, not only from the perspective of the patient in treatment but also from that of the wounded healer evolving as a more sophisticated physician.

Postvention: Support Services in the Aftermath of a Tragedy

Unfortunately for medical students, suicide is the second most common cause of death after accidents. Given the role of physicians in their

communities and the close-knit nature of medical communities, the effects of a physician suicide are profound, far-reaching, and long-lasting. It is critical for medical schools and centers to have suicide postvention plans in place in order to respond quickly and effectively in the event of a suicide. Whereas suicide prevention is aimed at decreasing the incidence of suicides, suicide postvention is geared toward providing support to the community in the aftermath of a suicide, helping to return to normal routines, assisting those who are exhibiting unhealthy behaviors, identifying those at risk for psychiatric illness, and minimizing the likelihood of additional suicides through contagion. Crises, whether natural disasters such as Hurricane Katrina, school shootings, or a campus suicide, demand immediate response to provide safety to the community and to restore normalcy as soon as possible. Ideally, the institution will have a postvention team already in place that can be activated in the event of a suicide. The first step following a crisis is to perform a risk assessment. In the case of a suicide of a medical student, it is of utmost importance to identify whether there are other at-risk students and to ensure their safety. In addition, it is necessary to reach out to individuals who are known to be close to the deceased person, such as roommates, boyfriends or girlfriends, close friends, and academic mentors. Psychological first aid is an approach used to help people in the immediate wake of a disaster to decrease their acute level of distress and to encourage short- and long-term adaptive functioning. It is important to provide psychoeducation about typical reactions to trauma, including disturbances in sleep, appetite, energy, and interest, as well as to emphasize the importance of self-care in returning to precrisis levels of well-being.

Case example: Mr. A was a 24-year-old Korean–American who had just started his third year of medical school. He was a popular and well-liked student who had been active in the Christian fellowship. He was an excellent athlete and often played pick-up basketball games with his peers. His classmates saw him as affable, kind, and funny. However, behind closed doors, Mr. A struggled with severe self-doubt and insecurities. Despite his high grades, he thought that his peers knew more than he did, and he questioned whether he could get into a competitive plastics surgery residency. Two weeks before the first step of his national board examinations, his girlfriend of 1 year broke up with him and started dating one of his close friends. Mr. A was devastated. The breakup added to his feeling that he was inferior to his peers. Mr. A had come from a high-achieving family, and he was expected to do the same. Despite evidence of significant mental illness in his family, it was never acknowledged or discussed. The day after Mr. A got his board scores back, he was found dead in his dorm by his roommate. He had hanged himself from his door frame. His roommate notified campus security, leading to the activation of the postvention response plan. The family was contacted immediately, and they expressly indicated that they did not want the cause of his death to be disclosed, which was honored. After the family was notified, the public affairs office of the medical center sent out an e-mail to the entire medical community, alerting them to Mr. A's untimely death. In addition, the dean of student affairs e-mailed the medical student community, including medical school faculty, confirming that a medical student had died and that staff were available to speak with any student wishing to do so. The same day, the postvention team set up shifts at the medical student dorm from 9 in the morning until 9 at night to provide psychological first aid for any students or staff. The counseling center

contacted medical students who were known to have had suicidal thoughts in the past or who were being treated for severe depression to offer additional support. In addition, the counseling center posted fliers in the dorm announcing a support group for those students who were especially close to or distressed by Mr. A's death. The leader of the Christian fellowship also held a special session to help its members talk about the recent loss of Mr. A. The dean of students also met with medical school faculty who were involved in teaching Mr. A both to provide support and to review common reactions to a crisis so that they could be on the lookout for themselves and students exhibiting any symptoms. In addition, the dean stressed to faculty that some students may need extensions on examinations or other accommodations. Meanwhile, the postvention team met on a daily basis to discuss the effectiveness of their efforts, to determine whether any adjustments were necessary, and to provide necessary support to each other. Over the subsequent weeks and months, the medical school community resumed its usual functioning. However, the postvention team continued to meet to discuss improvements for future events.

Anxiety Disorders and Addictions

Anxiety—the concern that something bad will happen—is distinct from fear—the body's flight-or-fight response to an actual dangerous situation. Anxiety in its best light (also known as eustress) can be motivating because it prompts the mind and the body to prepare for future situations, such as giving grand rounds or one's first experience with physical examination and patient interviews. In its excess (distress), anxiety can cripple the most highly functioning member of society. Formal diagnoses of anxiety, ranging from the diffuse generalized anxiety disorder to the specific obsessive–compulsive disorder to the context-specific post-traumatic stress disorder, are frequent presentations for medical students and physicians. When overly anxious, self-soothing behaviors attempt to modulate one's subjective sense of dread. When left unattended, anxious symptoms can lead to addictive behavior, chemical or nonchemical, in an effort to contain overwhelming affects. By definition, all addictions are intended to relieve a subjective sense of distress. Hence, the disorders of addiction are inseparable from the disorders of anxiety.

Numerous situations unique to medical training and the practice of medicine provoke regular anxiety. Self-soothing behaviors range on a spectrum from healthy and adaptive to pathological and self-destructive. Locating the degree of consequences stemming from one's self-soothing behaviors assists in the recognition of what might be worrisome. As with all mental distress and psychiatric illness, early intervention is of the highest yield.

Personality Disorders

Many physicians come to the attention of department chairmen, partners, or other colleagues for behavior that might pass on the surface as acceptable—as long as those involved choose to not pay too close attention or ascribe it to some other seemingly rational explanation—but that on closer examination is frankly disturbing or abusive to those more directly

involved. These behaviors, ranging from screaming at a floor nurse to assaulting an operating room scrub nurse, from asking a patient to edit a manuscript to asking a patient to have sex with the physician, from any moderate boundary crossing that is kept secret to a more frank boundary violation, are not necessarily the result of one of the formal mental illnesses described earlier, although affective illness or substance abuse may well complicate the picture. Regardless of diagnosis per se (e.g., narcissistic personality disorder, borderline personality disorder, antisocial personality disorder), the behavior that crosses the line, with or without the patient or fellow colleague's consent, is the behavior that abuses the power differential between doctor and patient or ethical respect for a fellow colleague. Personality disorders may present behavior that seems more subtle or easier to write off as "just who someone is" because it may depart much less severely from one's baseline than behavior that stands out as a sudden change of character and part of a new-onset medical, psychiatric, or substance abuse problem. By definition, most medical students and physicians have achieved numerous accomplishments so as to succeed as much as they have and have used considerable strengths in order to do so. Attempting to integrate an example of egregious behavior with a composite picture that also includes the physician's demonstrated giftedness recognized by one's community can prove confusing.

The behaviors in question likely attributable at least to elements of disturbance of personality include the basics of lying, cheating, or stealing —all versions of crossing or violating boundaries. Reading the newspaper or professional journals or simply keeping one's eyes and ears open at the hospital provides scores of examples. A surgeon's carving his initials into his patient's abdomen as a signature of his pride in work well done; a psychiatrist sleeping with his suicidal patient in a desperate attempt to bring her back to life; a gifted researcher convicted for double billing Medicaid and his federal grant for research patients' clinical services; an internist lying to his morning sign-out team about laboratories checked and services performed, which jeopardize patient care; all of these add to the scores of examples dealt with in physician mental health programs and clinics across the country. Although a physician impaired as a chronic function of certain personality traits may not jeopardize the number of patients that an impaired airline pilot might, to those affected directly by his or her mismanagement, the consequences may be just as dire.

It seems relevant to offer a cautionary message about the risks involved with the assessment of a colleague's abusive behavior stemming from personality dysfunction. When one's life is going well and one has plenty of loving family, health, and money, a physician may forget the depths of despair experienced by someone attempting to care for sick patients in the wake of a spouse's death, a child's illness, or any other family tragedy. Colleagues commonly report when hearing of such news or in studying case

reports in a retrospective manner that such behavior would never invite them to frankly cross that line with a patient or colleague in that way. If that were true, physician mental health programs would not exist in the capacity they do; experience shows time and again across all medical specialties that even the most gifted, inspirational, and leading physicians are susceptible to breakdown under the proper "perfect storms" of loss, illness, or blows. Experience also shows that appropriate windows for intervention also exist, namely, before the boundary-crossing behavior has begun and while multiple avenues for proper, confidential treatment exist. As abusive of patients, colleagues, or institutions boundary-crossing behavior can be to all parties, ultimately they are also entirely self-destructive to the physician in question. The task at hand is to attempt to suggest or provide treatment before such self-destructiveness mushrooms into damage that jeopardizes license, career, relationships, and family. That actions begin by bringing these behaviors to the attention of a superior or chairperson sooner rather than later.

Relationship Problems

All intimate relationships form their foundations on the cornerstones of time spent together and emotional availability. The practice of medicine demands time and emotional availability for patient care, often leaving its practitioners depleted in reserves of empathy and freshness. Whether the couple is newlywed or seasoned, two physician or traditional, gay or straight, rearing children or with an empty nest, or financially stable or vulnerable, the practice of medicine provides constant challenges to the stability and structure of the intimate relationships of medical students and physicians. Michael Myers and Glen Gabbard have written extensively on the infinite variety of struggle that couples within the house of medicine experience and on the utility of couples therapy to assist with critical relationship mending.

Medical Illness/Disability

Eventually, every medical student and physician becomes a patient as well. Whenever the convergence of biological dysregulation, psychological vulnerability, and life circumstance leads to any medical manifestation—from new-onset mild symptom (rash or fatigue) to moderate symptom (chest pain, incapacitating headache or pain) to undeniable event (collapsing on a patient during rounds)—the medical profession's privileged status of educated and powerful keeper of knowledge and cure blends with the patient's sense of helplessness and uncertainty. Whether acute or chronic, lethal or benign, the role of patient challenges the imagined security and sense of invulnerability fostered by, or even psychologically necessary for, constant exposure to the wounds of patients.

Whereas medical illness by definition compromises some aspect of physical health, the psychological reactions to medical illness and disability may be just as if not more profound for the physician. Many may prefer to avoid treatment at all costs, hoping instead to self-diagnose or self-treat; maintain total confidentiality and control sharing of potentially dangerous news to colleagues, family members, or patients; avert any type of procedure they might routinely recommend for patients; and avoid the potential meanings of the illness and its loss of control. Not only does medical illness bring loss of bodily control, pain, or professional impairment, it also raises the deepest of questions, ranging from intensely practical (how to feed one's family?) to emotional (how to navigate these chapters of loss and sadness?) to more spiritual and philosophical (why now? and how does this event fit into the larger narrative of life?). Multiple family reactions to the health care professional's illness understandably arise, naturally ranging from mild and contained for a time-limited episode with mild dysfunction to overwhelming grief for a mortal illness. Prior family episodes of dealing with illness or sick family members; pre-existing family conflicts and unresolved resentments; family cohesiveness versus divisiveness; power dynamics among and between grandparents, parents, spouses, and siblings; and the meaning of the particular illness to any individual in the family can all play vital roles. Robert Klitzman studied some of the key questions that arise as doctors become patients via comprehensive interviews with 48 doctors coping with a range of illness. He critically evaluated the nature of physician identity, the role of denial in medical professionals' lives, the intersection of being a doctor while being a patient, the poignant lessons of patient care that can be learned, the role of magic and spiritual considerations in the experience of illness and its treatment, and the nature of empathy as these doctors shuttle back and forth between the role of patient and provider. Within this matrix, the need for competent psychiatric care is indispensable. Careful psychiatric attention to the biopsychosocial model can assist greatly with the evaluation of impairment, further medical diagnostic considerations, and the often long overdue handling of emotional issues in life and relationships, as well as assistance in navigating the psychological vulnerability of the losses of identity, control, and power. It can also provide the critical advocacy that health care professionals may be unable to provide for themselves with their families, patients, and colleagues.

Physician Impairment, Administrative Response, and the Role of Professionalism

In the course of evaluating any physician for psychiatric cause, the question of impairment always rests quietly in the background. At some point, certain syndromes or behaviors cross a line and compromise the reasonable skill and safety of the practice of medicine, whether from medical or psychiatric reason. Common medical causes include delirium or

dementia, whether from acute toxicity from medication or cerebrovascular events or from chronic degenerative illness. Psychiatric causes span the range from more purely psychiatric (mood or anxiety disorders) to substance abuse to more characterological in nature (antisocial personality). The range of behaviors that receive disciplinary action at the level of the licensing board include physically or sexually abusing a patient; prescribing drugs in excessive amounts or without legitimate reason; exhibiting an impaired ability to practice because of addiction, medical illness, or mental illness; being dishonest; having a criminal conviction; delegating the practice of medicine to an unlicensed individual; not recognizing or acting on common symptoms; and failing to meet continuing medical education requirements. Each individual case of these issues may represent an event in isolation or the tip of the iceberg, highlighting the need for careful attention to detail on behalf of the medical board evaluation committee. Some doctors come to the attention of the board via a reported medical mistake or patient complaint; others enter as referrals for substance abuse or mental illness.

The Joint Commission for the Accreditation of Health (JCAHO) requires that medical centers have a standing committee for physician health and that physician health matters be dealt with separately from disciplinary matters. States differ in terms of criteria of evidence required for taking action against doctors and type of intervention. New York State, for example, requires only a "preponderance of evidence" in lieu of "clear and convincing evidence" and maintains a confidential, treatment-focused, rehabilitative philosophy. Any successful program that diverts these complaints to treatment committees also carries consequences, depending on the severity of the offending behavior. For example, a physician may need to have a chaperone present for all examinations, or perform a procedure with a fellow specialist present, or be subject to random drug testing or office visits to assess record keeping. Infractions of these treatment programs can lead to the suspension of a medical license if the behavior continues to place patients at risk or be otherwise unamenable to treatment. In New York State, doctors are usually monitored for 3 years for a professional conduct problem and for 5 years if they were found to be impaired from substance abuse, mental illness, or other behavioral problem. The focus remains on prevention to facilitate safe, regular, and reliable participation in the program and the successful modeling of the best of medical ethical care. Each state's society provides the particular details of its program.

SPECIAL POPULATIONS: MEDICAL STUDENTS, HOUSE STAFF, MIDCAREER PHYSICIAN, AGING PHYSICIAN

Although it may seem easy from a distance to lump medical students and physicians into one category for the obvious reasons of shared training and

clinical values, such a stance misses the specific developmental ramifications and struggles unique to each phase. The resident on the ward and the president of the board may share the same M.D. degree but face markedly different work tasks and life cycle tasks. The following sections with included vignettes address common mental health presentations within the medical profession and cast them with sensitivity to the unique phase-specific issues of being a medical student, house staff officer, attending physician, or aging professional facing retirement or the ending of one's career.

Medical School and Provision of Care to Medical Students Seeking Mental Health Treatment

Access to Treatment: Access to high quality and timely mental health treatment is paramount for all medical schools to ensure that medical students receive the care they need. There are various ways by which medical students connect to mental health services. Some medical students have pre-existing psychiatric disorders and know that they will need ongoing mental health treatment upon their arrival to medical school. Those students either seek out treatment on their own or request a referral from the health services of their medical school. Others may have had an episode of anxiety and depression earlier in their life, such as during college or high school, and then experience an exacerbation at some point during medical school. These students may also call for a consultation at student health services or may seek a consultation with a mental health professional in the community. At times, a medical student may not be aware of having a mental health problem. They may be presenting frequently to student health services for various somatic complaints that are either minor or for which the medical workup is negative. For example, the student may present with cardiac concerns that are determined to be panic attacks. Through the course of these visits, the medical provider may recommend that the student schedule a consultation with the psychiatrist. At times, a student may present for a mental health consultation in context of academic difficulties. When severe depression or anxiety is interfering with a student's academic functioning such as a failing grade, the student's advisor or dean may suggest to the student that he or she seek psychiatric help. Finally, on rare instances, an emergency, such as a suicide attempt, a manic episode, or alcohol poisoning, may compel a student into psychiatric treatment.

Medical training corresponds elegantly with a number of developmental tasks of extended adolescence. Much as childhood fantasies are no longer appropriate actions for the adolescent who is still not yet an adult, the medical student, although actually an adult, is not yet equipped with the intellectual knowledge or professional mentality necessary for patient care.

Medical students deal with predictable challenges across their 4 years of

training. Understanding the actual stressors as well as the psychological tasks at hand can facilitate a connection with the medical student presenting for treatment. Eric Marcus provided a stimulating blueprint of medical student and physician development via the study of a database of dreams. Through research collection and detailed analysis of hundreds of medical student dreams, he documented a progressive unconscious developmental process as medical students master skills and tolerate responsibility for human life. Marcus traced the evolution of the grandiose, empathic, patient-centered, and completely overwhelmed first-year medical student to the more reality based but pessimistic hero-healer of the graduating fourth year. His work adds strength to the concept of a physician life cycle.

The first year of medical school involves multiple new anxieties. Many students move from other parts of the country, leaving behind significant family and/or romantic relationships. Working with a cadaver, many come face to face with death for the first time. The academic tasks at hand can be as voluminously overwhelming as they are exciting; seeing a patient in a group setting for the first time as a professional can evoke passion, fear, empathy, awe, and confusion over boundaries between the patient's life and the medical student's. Blending with a new peer group—often one's first professional peer group—requires students to ask whether they will be able to fit in and find a place of their own.

Mr. A is a 24-year-old first-year medical student. He grew up in Boston and attended a liberal arts college on the East Coast where he studied history. He dated the same woman throughout college, and they were both very interested in humanitarian causes. They spent 2 years together in Kenya serving in the Peace Corp. Upon returning, Mr. A moved to California to attend medical school in California, while his girlfriend moved to New York City to attend business school after which she would join him in California. When Mr. A arrived to medical school, he struggled with the separation from his girlfriend, missing her immensely. He spent increasing amounts of time in his room, talking or e-mailing his girlfriend, which interfered with his making new friends in medical school. In addition, he had a hard time adjusting to the volume of schoolwork and to the frequent examination schedule. He had always been a procrastinator in college, finishing in his papers minutes before they were due. However, he had always earned top grades. After failing his biochemistry final, Mr. A was encouraged to meet with the school psychiatrist. Mr. A explained that he was unable to keep up with his schoolwork. He spoke about his long history of procrastination. In high school, his teachers often forgave his lateness with assignments because his work was outstanding. Because he was so bright, he rarely needed to study for examinations. In college, he would go to review sessions before the examinations and also would attend study groups with his peers. He managed to absorb enough of the material to do well on the examinations. However, his previous study habits were no longer working in medical school. Given the volume of material and the fast pace, he quickly got behind. He went on to explain that he had a hard time remembering where he needed to be and when. He described that his room was a mess, and he would often spend at least half an hour searching for his keys and ID before leaving his room. In addition to his mild symptoms of depression resulting from his being apart from his girlfriend, Mr. A was describing many signs and symptoms suggestive of attention deficit disorder, which was confirmed on neuropsychiatric testing. Mr. A began a trial of stimulant medication and reported a dramatic difference in his ability to concentrate and organize. He also had a brief psychotherapy to help him adjust to medical school.

The second year of medical school introduces students to illness and to patients' bodies, whether via coursework, history-taking interviews, physical examination, or hospital rounds. Contact with illness can trigger painful memories of loved ones who have died from disease. Touching patients' bodies can prove invigorating and terrifying, especially if one's history of touching or being touched is complicated by a history of physical or sexual abuse. Sitting with a sick or dying patient in detailed interviews while also obtaining detailed medical history can feel disintegrating as one all the while attempts to prepare for what feels like the torture of medical boards at the end of second year. The intensity of commitment to the process heightens, and important personal areas of interest often are left by the wayside in attempts to survive the process.

Ms. B is a second-year medical student who presented to student health services when she started studying for Step 1 of her national board examinations. Ms. B had done well academically during the preclinical years, passing all of her subjects. However, as she started to prepare for her board examination, Ms. B became consumed with worry. She was unsure what area of medicine she wanted to pursue for residency and wanted to "keep all her options open." She started to have trouble sleeping at night, especially difficulty falling asleep. She would lay awake in bed thinking about what she had studied that day and whether it was enough. She avoided going to the library because she felt her peers were ahead of her in their studies, causing her to feel even more anxious. She spent hours studying in her room but often got distracted with worry thoughts about the examination. After having a panic attack during a practice examination, Ms. B sought help. During the consultation, Ms. B indicated that she had had similar difficulties when studying for the MCAT. In fact, she had had a panic attack in the middle of taking the MCAT and had to repeat the examination. Her recent panic attack reignited her fear of having this problem during Step 1 examination. Ms. B was started on a low dose of a benzodiazepine to help her as needed with sleep and intense daytime anxiety. In addition, she was referred for cognitive behavioral therapy to help her learn new coping skills to manage her anxiety as well as relaxation techniques to help her during her panic attacks. Ms. B responded well to treatment and passed her board examination. She was pleased with her final score.

The third year of medical school involves the impossible task of synthesizing basic science information with actual life histories of real patients in hospital and clinic settings. Students must deal hands-on with medical illness and its aftermath, bodily trauma, patient self-destructiveness, and the grief of end-of-life decisions. Marcus documented this year as the time when sensitive students become overwhelmed at the impossibility of their tasks and the potential for their being exposed as inadequate. Many students transform from thoughtful and caring to cold, hardened, and aggressive but numb to the feelings of worthlessness and vulnerability they carry inside. Undergoing this transformation can prove scary for students used to channeling their negative feelings toward themselves primarily. It can make students question the integrity of their decision to pursue medicine as a career, often prompting a consultation to explore other career options and the depression of wondering whether one has made a profound mistake.

Mr. C was eager to start his third year of clerkships. While he enjoyed his preclinical years, he was excited to experience daily patient contact and to work in the hospital setting alongside his peers. However, a few weeks after starting his clerkships, Mr. C started to feel increasingly down. He started to feel insecure about his knowledge base, constantly feeling as though his peers knew more than he did. He also felt more isolated. Having grown accustomed to seeing his friends on a daily basis during the preclinical years, he now felt lucky if he saw them once a week. In addition, he was increasingly disillusioned with his peers, who seemed like they were constantly trying to "brown-nose" the attending physicians and were generally acting like "gunners." In addition, he thought that the residents seemed jaded from overwork. He worried about losing his enthusiasm for medicine and becoming a cynical and sarcastic doctor. He started to question whether he should have gone to medical school at all. He started withdrawing from his friends on weekends and spending more time alone. He was calling home more frequently, and his parents urged him to seek help. After a consultation with the school psychiatrist, Mr. C began a course of dynamic psychotherapy to address his feelings of inadequacy and also to help him work through his ambivalent feelings about medical school. After a few weeks, Mr. C began to feel better but he continued in an exploratory treatment to address these emerging feelings of inadequacy that had been long-standing. The son of a cardiothoracic surgeon, Mr. C had felt that he needed to pursue a surgical subspecialty like his father. Over time, he was able to be more comfortable with his interest in pediatrics. At the time of graduation from medical school, Mr. C was excited again about clinical medicine and matched at his residency of choice in pediatrics.

The fourth year of medical school presents the excitement of putting together in a new way all that one has learned. Many students savor the fantasy of at last being able to do what they might ultimately want to and even get paid for it. Marcus detailed a fantasy of being able to do over in postgraduate training what one failed to do well in medical school, often focused on a feeling of resurrecting a sense of empathy for patients and idealism about the profession of medicine. Graduation from medical school is often coupled with marriage or other kinds of formal separation from one's family of origin, an anxiety-provoking task that invites regressing back to old patterns of behavior.

Ms. D decided to consult the school psychiatrist during her fourth year of medical school. She had been in a two and a half year relationship. Although she was happy in her relationship, she was unsure whether she wanted to marry this person. She and her boyfriend had been discussing whether to couples match. Her boyfriend was very eager to do so and had talked about his desire to propose to her. Ms. D felt a tremendous amount of pressure and had feelings of guilt about her indecisiveness. When Ms. D consulted the psychiatrist, she explained that her parents were divorced and that she did not want the same for herself. She worried that going through the couples match together was tantamount to marriage. She went on to say that she had been in two serious relationships prior to her current one. Ms. D continued to meet with the school psychiatrist over the subsequent weeks. What emerged was that Ms. D was very much in love with her boyfriend, but that she was afraid of increased closeness and commitment. After being able to talk openly about her concerns, her fears gradually subsided. Ms. D realized that she wanted to couples match with her boyfriend and was also eager for him to propose. Ms. D contacted the school psychiatrist a year later to notify her of her recent marriage.

Postgraduate medical training offers the freshly minted M.D. the chance to invest deeply with a group of colleagues, set of patients, and field of knowledge. These goals foster the formalization of the identity of a

physician after graduation from medical school and occur with the payment of a salary. Residents and fellows are invited perpetually to join in the exhaustion of frequent overnight calls, the stress of tending to high volumes of patient loads, and the heroic task of integrating formal medical knowledge with the demands of patient care. Although residents may feel impotent and inadequate all over again, they also have the chance to grow in ways previously scarcely imaginable. They can carry their own patient loads, work more independently, develop lasting friendships with like-minded colleagues, cultivate transformative mentorships, and become a formal member of a professional group. Residents can become the integrated, longed-for, competent, and optimistic individuals that they dreamed of being before they started medical school, a fantasy strengthened by increased competence and the actuality of starting one's own practice or gaining one's first faculty position. Naturally, many residents or fellows finally feel that they that can make room for more intimate, committed, loving relationships and choose to marry or partner. Residents become more capable of displaying love for what they do and for those with whom they do it.

Dr. E presented for consultation in his third year of his psychiatry residency for complaints of depression and burnout. He lamented that his passion for learning and zest for life had each atrophied and dried up since the beginning of his PGY-II year. He could not say for sure why he thought this descent had occurred, although he readily identified a couple of immediate possibilities: failure in his love life relative to his high expectations and experiencing himself as a cog within the greater wheel called profit-driven medical care, each leaving little protected time for his formerly passionate thirst for knowledge. He also owed more money than he could calculate to credit cards and student loan agencies—when would a day come that he could ever pay this debt or live in the expensive city of his training on such a thinly stretched salary? While he enjoyed basic invitations to know more about the human condition in sickness and in health and about the ubiquity of human suffering, he found himself directionless and listless in a bigger picture of his life. His prior experience in psychotherapy had soured, which left him in a hypocritical stance toward his patients and his work: how could he recommend and/or perform a procedure that had so wounded him?

Consultation over several sessions focused on several core concepts: the establishment of a supportive, mending therapeutic alliance; the restoration of hope; the normalization of Dr. E's plight relative to his stage of training and deep capacity for empathy; and exploration of pharmacologic options to manage his mood and anxiety. Numerous routes of treatment were considered, including participation in a group therapy program for physicians. Over the course of the year, Dr. E made steady gains in his life and his mood. He broke up with his on-again, off-again lover with whom he had spent years in a passionate bad fit and worked on some searing emotional pain that stemmed from his family of origin, identifying his mother's history of alcoholism and his fathers' covert history of PTSD from Vietnam as significant drivers of his wish to become a psychiatrist. He reconnected with his brother in new and rewarding ways, fell in love with a fellow resident from the emergency medicine program, and refined a professional path focused on intensive inpatient hospital work with physicians in crisis.

Life Beyond Training: Marriage, Children, Midcareer, Midlife

Midlife coincides with a time when the physician can display concern for the people, training, and institutions in which she works or begin new

phases of creative thinking and personal involvement with important relationships. Moving comfortably into this expansion means mourning the loss of one's prior role as a student and trainee. It means renouncing one's longed-for cravings in lieu of caring for the needs of those dependent on the physician. It also allows for the gratification of seniority within the tribe. A physician commonly has to ask the serious questions of time and family, reformulating who he or she wants to be and with whom he or she wants to be that person. Such editing can yield a new set of dreams and a more limited structure or scope than prior phases of life offered. Investing in the next generation can take many shapes, expose fears of one's mortality, and offer a chance to let go of a sense of control over life.

Dr. F presented for treatment in the context of her midcareer professional crisis. Because of a confluence of native gifts, affinity for patient care, and unusual capacity for three-dimensional visualization of anatomy, Dr. F excelled in medicine beyond the self-perception of stupidity that she carried from her years growing up as both bright and learning disabled—somewhere between the ADHD–dyslexia boundary. Using her passion for life and innovation, she pioneered a technique within the field of outpatient surgery that could save lives, minimize invasive procedures, and prove quite lucrative. Her brilliance and level of actualization of accomplishment set her up to be simultaneously a coveted and envied object in her department. In particular, her department chair dealt with her primarily through the use of neglect. While benign enough on the one hand to allow her to continue to develop her craft, it was also malignant enough to stunt her professional creativity and professional reach. This neglect cost her hospital bundles of lost income and missed opportunities to help scores of patients.

In response to the consultation, Dr. F began to trust her instincts more sharply and began to seek positions elsewhere. She could see quickly that she had unwittingly repeated a pattern from her childhood. She grew up with a father addicted to gambling, a profession that left the family intermittently wealth and destitute. Dr. F developed a reluctance to trust any one of significance too closely while she still found herself drawn to risk-taking and destructive men. In response to her family's outright neglect of feminine sexuality, she found herself passionate about all reproductive anatomy and biology and in equal measure ashamed of this passion. The basic question emerged: sweat bullets, work 80 hr/wk at minimum compensation without protected academic time and without the kind of downtime needed for a creative life; or take a position at a rival institution that would hire her at a time of her choosing at a lucrative salary and lots of flextime for her research and child rearing—all with oodles of ancillary staff and research associates?

While the question at the level of logic was effortless to answer, Dr. F was surprised at just how hard it proved to make the shift to a better fit. Using a metaphor of a "passionate bad fit," the authors framed her "truest" passion in maintaining the long-suffering, understated genius who knew all the family secrets and was still ridiculed and scapegoated for being "the baby" of the family. Once the authors were able to uncover—and Dr. F was able to own—her hidden pleasure in suffering professionally in a manner parallel to her identity as a girl in her family structure, lights started to go off as she engineered a passionate better fit for herself. She decided to join a therapy group for physicians that allowed her to play out the back and forth between old bad fit and the created better fit so that she did not have to endure these frustrations alone and could harness the power of the group to galvanize her own motivation for change.

The Aging Physician

Breaking up is hard to do, and doctors leaving their careers, practices, and patients behind are no exception to this rule. Feelings of loss can carry the

day at the expense of healthy consolidation and recognition of services performed, feats accomplished, and relationships cultivated. As if dealing with progressive loss of sight, hearing, physical acuity, and bodily integrity were not enough, the aging physician must cope with new configurations in practice and family life, making way throughout for new blood and younger generations. He or she must let go of the power and esteem granted as a regular function of the previously inhabited role. Financial questions that were ignored during one's busy practice can come to a head, whether via the negotiation of the sale of one's practice or the management of one's savings and investments. The same choices of any prior phase of life—to choose life and its positive integrity or to choose death and its despair—beckon in an even more pronounced way, often as a function of having more free time and an empty nest on one's hands. Seeking treatment—medical and psychiatric—and staying active in regular routines that connect one to a passion or joy can prove life affirming.

Dr. G was an 80-year old rheumatologist who presented at the time of consultation. He had suffered from a series of mild yet cumulative TIAs and was bereft to face the loss of his practice, his career, his long-standing teaching role in his department, his pleasure as a husband, father, and grandfather, and his identity as a competent, much valued healer and leader. If he was to be honest with himself, it was time to acknowledge his pending limitations to himself, his practice partners, his department, and his patients. He had taken part in exploratory psychotherapy for more than five decades ago as a resident and fellow—a process that he had enjoyed immeasurably and one that he credited with allowing him to marry his wife and to be a warm, loving, and available father in spite of the unconscious wish to join the model of his rage-filled, alcoholic father. Dr. G—ever the careful physician—was open to starting an antidepressant to minimize the impact of post-CVA depression and was open to the idea of a 10-week consultation to perform a "life review" to take an inventory on what business he most wanted to finish in whatever days of sickness or health he may have ahead. Using the best of his internist mind, he created spreadsheets of categories with all of the important branches of his life activities, using leaves to illustrate the identities of all the people who had so enriched his life. He made a point of keeping close track of the progress he made in connecting and reconnecting with unfinished business of recent or former times in his life. Although he could still practice effectively, he chose to engage a succession plan in his practice and maintain a voluntary teaching role at the hospital. He made a point of taking trips with his wife; scheduling frequent time with his children and grandchildren; and writing daily about his ups and downs in the transition. Because of the depth of the grief, and because of the benefit he perceived obtaining during the consultation, he elected to pursue individual treatment again in this stage of his life and wanted to include periodic joint meetings with his wife, children, and in-laws to maximize the probability of all members of his tribe having a chance to come to terms with his decay and to build useful structures of relationships for the grieving and transitions as he let go of his title of patriarch.

Treatment Approaches and Alternatives

Psychiatric consultation: A thorough psychiatric consultation is critical in determining the particular needs of the student and establishing a treatment plan. This consultation is often done over the course of several visits. There can be several different outcomes of a psychiatric consultation. For students where there is imminent danger or first break of psychosis,

inpatient hospitalization may be necessary for the safety of the student. For less acute situations, a referral for outpatient treatment for psychotherapy, psychopharmacology or both may be the treatment of choice. At times, a previously undiagnosed learning disability or attention-deficit disorder may become manifest during medical school. In those situations, neuropsychological evaluations may be critical for determining the area of deficit. When a student needs to take a medical leave of absence due to his or her psychiatric illness, the student must meet with the school's administrative psychiatrist for both an exit and reentry interview to ensure that the student receive proper psychiatric care while on leave and to determine when the student is ready to return to the rigors of medical school. In situations in which a substance use disorder is interfering with the student's functioning, a referral to the committee on physician mental health can be appropriate. This committee oversees the care of the student and administers random urine drug screening to ensure that the student remains abstinent. Finally, there may be times when family therapy is indicated and may be recommended.

Benefits of Treatment. Approaching medical students and physicians through the biopsychosocial model proves most useful. Making an appropriate diagnosis when indicated; providing psychoeducation about the disorder; obtaining neuropsychological testing; providing pharmacologic interventions; finding special tutoring; creating an administrative psychiatric interface separate from one's confidential treatment to manage contact with a dean's office or medical board; and providing or making the necessary referrals for cognitive–behavioral strategies, longer-term psychodynamic psychotherapy, psychoanalysis, couples therapy, family therapy, 12-step groups, and/or group therapy all may provide useful interventions.

The initial psychiatric interview can open the door to a lifetime of change. The same sort of trauma that motivates physicians to provide outstanding care to patients can change into intense self-criticism and uncertainty. Physicians and medical students commonly experience profound shame in seeking mental health care, viewing it as an absolute last resort to be tried only after every imaginable self-coping strategy has been used unsuccessfully. In the eyes of the man or the woman who always wants to be right and always wants to be in control of difficult situations, settling into a psychiatric interview awakens tidal waves of humiliation. Paradoxically, this very fear of stigmatization allows for the perpetuation of the prejudice against psychiatry within the medical field. Think of the number of jokes physicians of other subspecialties make about psychiatrists and psychiatric patients, even though in private they reveal how acute the anxiety they felt on their psychiatry rotation may have been in response to working with patients suffering the agony of mental pain.

Keeping all of this in mind and engaging the physician patient empathically about his or her struggles can allow that patient to feel cared for, a feeling that is very likely lacking in his or her life if he or she is seeking professional help. Disproving the physician or medical student's internal theory of mental distress as being equivalent to weakness may allow the patient to open up about the genuine nature of his or her struggle. One of the more gratifying aspects of treating health care professionals is witnessing the improved care these medical student and physician patients provide their patients in response to the therapeutic alliance with a competent practitioner.

Taking a careful history sheds light on the degree to which a health care professional's presenting emotional turmoil stems from a present stressor of medical practice or training and/or how much that turmoil stems from a mental reactivation of painful memories from prior life experiences. Commonly, the past and present sources coincide in a way that makes the medical student or doctor symptomatic in a newly debilitating way. Inquiring about the person's motivation to go into medicine reveals significant family relationships, mentors, or personal experiences that motivated him or her to make all of the sacrifice inherent to medical training. It also opens the door to learning which relationship patterns serve as healthy foundational structures on which to build in treatment.

Barriers, Resistance to Treatment, and the Stigmatization of Mental Distress

By the time a medical student or physician has created a situation in which reporting has occurred, whether for making a medical error, violating a boundary with a patient, having alcohol on the breath at work, or experiencing intense new-onset suicidal thinking, many opportunities for prior intervention have been missed. It may be that the anxiety about seeing a problem as existing within one's self poses such a threat in the highly rigorous, scrutinized, public profession that many physicians prefer to resort to automatic denial of the problem until it explodes. Public health efforts to mold the culture gradually into a more accepting one are infinitely worthy. Any chance to treat a medical student or physician patient with empathy, curiosity, and specific interest in the history of his or her life can prove worthwhile. It might provide access to a central, painful prior life conflict that may become reactivated in the context of the stress of medical training, or it may help process a personal loss that has been summarily suppressed until it overwhelms the patient's professional capacities. Appropriate treatment at these moments can prevent many of what later become problems for a physician's practice, patients, medical community, and family. With that in mind, multiple levels of resistance serve as rate-limiting factors in the implementation of healthy, sustainable, integrated interventions.

Perceived Lack of Confidentiality to Medical Students and Residents in Training. When trainees perceive that what they say may not remain confidential, they become reluctant to utilize the services provided. This perception—much life stigma—can take many forms; it also perpetuates the very stigma of mental suffering during the course of training, perpetuating the notion that talking about one's mental stress is both bad and dangerous when, in fact, talking to another is the first step toward getting relief. Not talking is the fulcrum of danger toward emotional isolation, numbness, and danger. If the dean of a medical school does not fully welcome taking care of one's own mental health as a priority in the provision of first-class medical care, then none of his or her department chairs can be expected to follow suit. In parallel, that same attitude will trickle down to all residency directors. As in the military, leadership begins at the top. If a student's or resident's mental health care is not fully confidential (with the exception of dangerousness to self or other), then the trainee is stuck in a genuine bind: seek care and risk the compromise of confidentiality (perceived or actual); "white-knuckle" the natural history of one's condition; or to seek care outside of the system, an avenue often prone to backfire for financial or substandard clinical reasons. Traditionally, the lower the level of training, the fewer financial resources are made available. In parallel, the less psychological sophistication available to the trainee to locate and secure his or her independent treatment arrangements outside of those offered by the system at hand.

Hazards of the Electronic Medical Record. One hazardous venue in which this stigma may play itself out is within the evolving electronic medical record (EMR). While many reasons are cited as advantages of the EMR, these benefits neglect the sensitive nature of mental health treatments of those in training within a given institution. Some trainees may end up in significant positions of influence in this field in the future—as chiefs of service, department chairs, congressmen, or other physician positions of public stature and scope. Any number of readily available public examples of breach in confidentiality—from the availability of Farrah Fawcet's medical records for breast cancer treatment to a live illegal access of protected health records on the Today Show. The 2001 Supreme Court ruling of *Bartnicki v. Vopper,* which holds that the media has a First Amendment right to publish information about a person in public life, even if the information was obtained unlawfully—as long as the media outlet did not take part in the unlawful acquisition of data. This ruling provides a mechanism whereby countries such as North Korea or terrorist organizations such as ISIS might hack confidential health information as part of a broad sweep and then unwittingly allow that information to be open to legal publication.

While early in the process of possible long-term impact on future

careers, at the immediate training level hazards remain more at the level of stigmatization. Although accessing a colleague or trainee's protected health information is illegal, EMR systems routinely receive thousands of requests per day to "break the glass" of a patient record, making the practical policing of such incidents particular to trainees difficult at best. Furthermore, the face sheet of such records can be accessed by someone as casual as the clerk at an ultrasound facility, making basic information such as medications and problem lists available to many a crafty snooper. The bind is further magnified when one learns that the electronic prescribing of a medication will—by definition—place that prescription on the first face sheet. Some medical schools have become sufficiently worried about the impact of such unintended yet dangerous access to protected health information that they have outsourced the EMR of students to a firm separate from the University's main provider. This conflict places the students in a real bind: get the help they need immediately and take the risk of all that documentation may bring, from stigmatization by an unfriendly chief resident or attending physician to sharing of these records with a state licensing board from a state that still asks whether the trainee has ever consulted with a mental health professional in lieu of asking whether the trainee has ever suffered from a mental disorder that has prevented the individual from practicing medicine safely.

Most importantly, medical students and residents in training—when in the grips of acute stress, depression, or intoxication—do not think in a linear manner consistent with a larger picture of the order of their suffering. Many will choose to prioritize the power of stigma over their own care. Many can and will after sufficient treatment that they no longer worry about the implications of any disclosure of their treatment upon the trajectory of their careers; they just as readily say that they could not have known that at the beginning of the therapeutic adventure.

Financial Barriers. While the LCME (Liaison Committee on Medical Education) and ACGME (American College of Graduate Medical Education) require that all students and trainees have similar access to affordable mental health services, in actuality, the picture can look far more complicated up close. For example, a medical school may require that all students provide documentation of proof of mental health insurance in order to opt out of the school's policy, which only a third will purchase. Imagine then that a student presents in mental health crisis with no family financial resources, restricted student loan funds, and in-network coverage only from a state across the country that covers one psychiatrist 10 miles away, one specializing in psychopharmacology. Or a resident learns that her private hospital pay line's health insurance will cover about 10 percent of the out-of-network costs after meeting the deductible while her colleague at the public hospital receives benefits that pay 90 percent of outpatient

visits with unlimited frequency. All before securing a provider who is willing to work at 6 or 7 am, 8 or 9 pm, or irregularly while postcall to accommodate the call schedule of the service. All of these imbalances and this incoherence create disruptions in the consultation service and compromise the availability of the clinical staff; increase the possibilities for burnout; restrict the time available for administrative matters and future service-building plans; restrict the options for self-study of the service; and challenge the primary care providers of the service who invariably end up seeing the overflow of patients for less time and with a less developed skill set.

Vignette: Dr. M is a 30-year-old plastic surgery resident who presented for help with his anxiety in response to developing a tremor in the operating room at the time of his father's unexpected illness and demise. Dr. M had started drinking alcohol daily in the evening to attempt some form of self-soothing. He had heard positive reviews about the mental health services for trainees at his institution while also harboring a profound terror that whatever he said privately in treatment might be used against him in his application for a distinguished surgical fellowship. He postponed consultation until unwittingly sticking one of his peers with a needle driver while operating on an HIV+, hepatitis C+ patient. Only after a frank talk with his department chair—a lovely man with an extensive history of depression and useful history of successful treatment for the depression—did Dr. M present for care. His mental health benefits were not on par with his peers because his pay-line was from the private hospital, which had a long history of providing slim benefits for mental health coverage. Dr. M was referred to a psychiatrist with interest in working with physicians regardless of ability to pay and found himself using his innate intelligence and problem-solving skills to reconstruct his own history of burnout in new ways; he began attending Caduceus 12-step groups for physicians with addiction and integrated sobriety into his own healthy lifestyle. In part from his gratitude to the physicians who helped him to treat him at the low moments of his trajectory, Dr. M donated a regular portion of his time to the education of physicians with substance abuse and burnout and to tackling the impact of stigma against mental illness in the house of medicine.

INABILITY TO CARVE OUT FREE TIME FROM THE DEMANDS OF PATIENT CARE. The most frequently cited barrier to treatment in the health care professional is a busy work schedule.

NARCISSISTIC PSYCHOLOGY. Physicians take the adage to "heal thyself" too literally and cut dangerous corners in prescribing for and treating one another without an appropriately framed doctor–patient relationship. This attitude also perpetuates the idea that physicians can manage their own in distress, even though intellectually they know that treating their own Parkinson disease, for example, would be fraught with risk.

FEAR OF STIGMA OR DIAGNOSIS. Most medical students and physicians (as well as those in other professions) fear repercussion from their diagnosis. Will the dean's office or medical board find out? Will they have to ask for extra time while taking examinations? Will they need to report their diagnosis to the residency programs to which they apply? Students and physicians report extensive perceptions of stigma associated with mental

illness, seeing it as a sign of weakness to be avoided. Thus, they move away from any help-seeking interventions. When students and physicians do seek mental health treatment, invariably they wish to know the specifics of documentation and confidentiality for fear that the information gathered in the consultation might later be used against them. Few see talking freely about mental distress as the sign of mental strength, trust, and flexibility that it is, and thus most students are loathe to let any official at the school or hospital learn of their struggle and treatment, fearing subtle discrimination. Paradoxically, this fear also keeps students from obtaining the treatment that will allow them to avoid the dean's office or medical board. Physicians fear that leaving the hospital during daytime hours for a doctor's appointment will lead to poor evaluations and suspicion, imagining that they must explain all details of a physician visit to their chief residents, attending physicians, program directors, or chairpersons. Unconsciously, many patients fear an overwhelming loss of control of themselves if they are labeled with a psychiatric diagnosis. Avoiding psychiatric jargon and focusing instead on the emotional crux of the presentation can facilitate the medical student or physician patient's moving out of the medical model of "diagnosis and prognosis" that he or she so fears but cannot resist. If medications become a part of the treatment recommendation, then focusing on the use of medication to alleviate symptoms rather than treat a "diagnosis" can make the prospect of medication seem less intimidating. It can also allow deeper history to unfold that can be used to formulate a meaningful plan of treatment.

FEAR OF MEDICATION. Medical students and physicians commonly fear taking medication, regardless of their willingness to prescribe it for others. They become anxious about side effects of stimulants, antidepressants, or anxiolytics, often wishing to pursue a more "natural course." Perhaps afraid they are taking "brain steroids," or perhaps scared to find out that they might actually perform better if properly medicated for an underlying instability of attention, mood, or anxiety, medical students and physicians worry about the meaning of taking these medications. They worry how medication might alter their sense of self, commonly asking, "If I do so much better on a medication, then who was I before I took this pill?" Conversely, performing better with proper treatment can lead students to feel less stupid or insecure, knowing perhaps that some subtle learning or emotional difficulty always went along with the hard work necessary to achieve their earlier goals. Obviously gifted students have often said, "I thought I was just dumb or defective in some way."

WISH FOR DIAGNOSIS TO BE STRICTLY BIOCHEMICAL. A consensus on the accurate diagnosis of mental illness does not exist. However, multiple models compete for explanatory primacy. The pharmaceutical industry and some

journals implicate a host of neurobiological and genetic factors. Psychoanalytic models wrestle with the impact of any given constitutional issue and its impact on the developing mind or seek the impact of early relationships and environment with the correlation of disturbance. Regardless of the relative effects of nature and nurture, it is known that students and physicians do better when treated appropriately. Invariably, good treatment means looking at more than the immediate effects of medication. It means exploring feelings of low self-esteem, fear of one's aggression, anxiety at leaving families behind through successful maturation, longings to be taken care of, and any other number of developmental themes so common to the medical student and maturing physician experience.

Countertransference Issues

Working with physicians opens the door to a particular set of countertransference issues that can impede the provision of a high standard of care. For instance, the treating physician might fear that the physician patient is smarter—which may well be the case—thus minimizing a contribution to the careful management of the patient, which is often more emotional in nature than an issue of "book-smarts." The treating physician can feel compelled to join the physician patient in thinking about the clinical struggle as purely an intellectual one, thus depriving the patient of crucial emotional clinical data. Most significantly, the treating physician can identify with his or her health care professional colleagues and normalize their distress as secondary to the trials of medical training, thus stretching the realm of "normal" to miss significant avoidance of psychic pain in need of mending. The treating physician might forget to think of suicide as a central concern, given the mutual shared value of life in the profession; wish not to see a colleague as impaired; mistake the patient's high intellect as an absence of serious disturbance; or wish to avoid destructive feelings as the provider. All of these factors can lead the treating physician to buy into the medical student or physician patient's wish to avoid treatment and the wealth of uncomfortable feelings likely experienced, even if with a sensitive, compassionate listener. Buying into a medical student or physician's wish to avoid treatment when it is indicated will only defer the struggle until a later time, when it may be much more difficult to handle, if only because many more people are then affected. It also postpones the very relief that invariably is longed for if any potential physician or medical student patient musters the courage to pick up the phone and seek a consultation for any issue that poses as the chief complaint on interview.

Future Directions

Much as the military has had to confront an epidemic of suicides since September 11, 2001, the medical profession is at its own tipping point. A window of opportunity exists now to expand upon the work of so many individuals, medical student mental health services, families of medical students or physicians who have died by suicide and are still grieving in isolation, and the pioneering work of the American Foundation for Suicide Prevention. It is notable that while there is a journal for College Mental Health, there is no equivalent for Medical Student Mental Health. If a movement were to take root, likely it will stem from action by the very medical students and residents who are most at risk. Technology and the World Wide Web offer unforeseen and unprecedented avenues to reach large numbers of people. If much of the mental duress that occurs during training stems from the emphasis of science over intuition, of rational thought over feelings, and of the prioritization of taking care of one's patients over taking care of one's self, then it makes sense that much of the risk of training comes from impossible expectations of omniscience, suppression of personal emotion in the name of patient care, and an extinguishing of one's capacity to tend to physical, mental, and spiritual health.

A Prioritization of Prevention. Conceptualizing mental health, wellness, and suicide prevention interventions along a spectrum provides multiple levels of future interventions. Primary prevention might include weekly well-led process groups for all medical students or residents that implicitly enhance empathic sensitivity, promote professionalism, foster positive attachments within peer group, and strengthen mentalization skills. Secondary prevention might entail education of all levels of staff on suicide risk, identification, and intervention. Requiring in-services on the nature of physician impairment and responsibility of peer-reporting or hosting screenings of the American Foundation for Suicide Prevention documentaries on physician suicide can be highly educational. Tertiary prevention focuses on the provision of affordable, local, high-quality mental health services at all levels of training. Quaternary prevention builds an administrative shift in attitude toward stigma. Sensitivity to electronic documentation of mental health concerns as a barrier to access of services in the context of a doctor becoming depressed and a shift in licensure reviews of mental health services from "have you ever sought mental health treatment" to "have you ever been impaired as a function of mental illness, addiction, etc." mirror this shift. Finally, the presence of a postvention team of mental health professionals trained in psychological first aid to provide individual and community intervention in the peer environment in the case of student, resident, or staff physician brings to life the actuality that postvention is the best prevention at a time of crisis.

References

Beyond Blue. National mental health survey of doctors and medical students. Available at www.beyondblue.org.au. Published 2013.

Branch W, Pels RJ, Lawrence R, Arky R. Becoming a doctor—critical-incident reports from third-year medical students. *N Engl J Med*. 1993;329:1130–1132.

Chew-Graham CA, Rogers A, Yassin N. I wouldn't want it on my CV or their records: medical students' experiences of help-seeking for mental health problems. *Med Educ*. 2003;37:873–880.

Cochran, J, Kenney, C. *The Doctor Crisis: How Physicians Can, and Must, Lead the Way to Better Health Care*. New York: Public Affairs/Perseus Books; 2014.

Dyrbye LN, Thomas MR, Massie S, et al. Burnout and suicidal ideation among U.S. medical students. *Ann Intern Med*. 2008;149:334–341.

*Gabbard G, Menninger R. *Medical Marriages*. Washington, DC: American Psychiatric Press; 1988.

Gerber LA. *Married to Their Careers: Career and Family Dilemmas in Doctors' Lives*. New York: Tavistock; 1983.

*Goldman LS, Myers M, Dickstein L. *Handbook of Physician Health: The Essential Guide to Understanding the Health Care Needs of Physicians*. Chicago, IL: American Medical Association; 2000.

Higher Education Mental Health Alliance (HEMHA). *Postvention: A Guide to Suicide on College Campuses*. 2014.

Klitzman R. *When Doctors Become Patients*. Oxford: Oxford University Press; 2008.

*Marcus E. Medical student dreams about medical school: the unconscious developmental process of becoming a physician. *Int J Psychoanal*. 2003;84:367–386.

*Myers MF. *Doctors' Marriages: A Look at the Problems and Their Solutions*. New York: Plenum; 1998.

Myers MF, Fine C. *Touched by Suicide: Hope and Healing after Loss*. New York: Gotham/Penguin Books; 2006.

*Myers MF, Gabbard GO. *The Physician as Patient: A Clinical Handbook for Mental Health Professionals*. Washington, DC: American Psychiatric Publishing; 2008.

*Newton DS, Newton PM, Erikson HE. In: Sadock BJ, Sadock VA, eds. *Kaplan and Sadock's Comprehensive Textbook of Psychiatry*. 7th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2000:607.

Ofri D. *What Doctors Feel: How Emotions Affect the Practice of Medicine*. Boston, MA: Beacon Press; 2013.

Quinones-Hinojosa A. Terra firma—a journey from migrant farm labor to neurosurgery. *N Engl J Med*. 2007;357:529–531.

Rhoads J. Overwork. *JAMA*.1977;237:2615–2618.

Shanafelt T, Adjel A, Meyskens FL. When your favorite patient relapses: physician grief and well-being in the practice of oncology. *J Clin Oncol*. 2003;21:2616–2619.

Shanafelt TD, Bradley KA, Wipf JE, Back AL. Burnout and self-reported patient care in an internal medicine residency program. *Ann Intern Med*. 2002;136:358–367.

Taub S, Morine K, Goldrich M, Ray P, Benjamin R; The Council on Ethical and Judicial Affairs of the American Medical Association. Physician health and wellness. *Occup Med*. 2006;56:77–82.

Vaillant GE, Sobowale NC, McArthur C. Some psychologic vulnerabilities of physicians. *N Engl J Med*. 1972;287:372–375.

Wood D. After years of suffering, a widow finds new hope. *People Weekly*. April 3, 1995:43.

32

Psychiatric Emergencies

▲ 32.1 Suicide: Overview and Epidemiology

DANUTA WASSERMAN, M.D.

INTRODUCTION

The problem of suicide cannot be limited to just one discipline, and the complexity of self-destruction transcends the psychiatric field. Suicide can be viewed from many perspectives like religious, philosophical, anthropological, sociological, ethical, psychological, and psychiatric/biological.

In suicide prevention, strategies can be directed toward health care services or at the general population. The health care strategy aims at identifying risk groups, improving diagnostics, treatment, and offering better rehabilitation for suicidal patients. Health care interventions can be selective and target subgroups displaying risk factors for suicide, or indicated, that is, targeting individuals with identified psychiatric disorders, who have severe suicidal ideation or have attempted suicide.

The public health approach comprises strategies that aim to increase awareness about the role of protective factors for mental health, attitudes toward the mentally ill or suicidal persons, and promote measures to eliminate fears and misunderstandings that surround suicide.

Protective factors for suicide include cognitive flexibility, active coping strategies in difficult life situations, healthy lifestyles, active social networks, confidence and the sense of personal value, and the ability to seek advice from others and help from the health care system for subsequent treatment. It comprises also strategies for promoting mental health, social welfare, and education policies.

This chapter focuses on the broad spectrum of psychopathology in suicidality, and on treatment and preventive interventions used in the health care and public health sectors.

Universal public health interventions like restricting lethal means to suicide and school-based interventions are also shortly described.

EPIDEMIOLOGY

Suicide

The WHO report "Preventing suicide: a global imperative" published in 2014 estimates that over 800,000 people die by suicide, and more than 20 million attempt suicide each year. This implies that every 40 seconds, a person dies by suicide somewhere on the globe, and every 1.5 seconds, someone will attempt to take his/her own life. However, those numbers are underreported, as not all countries in the world report suicide mortality to the WHO. Only 59 out of 194 Member States have good quality mortality registration data. Furthermore, in countries with good mortality statistics, suicides can be misclassified as accidents, homicides, or unknown causes of death.

Annual global suicide rates are 15 for males, 8 for females, and 11.4 per 100,000 population.

Globally, suicides account for 52 percent of all violent deaths in men and 71 percent of all violent deaths in women. In high-income countries, 79 percent of violent deaths in both males and females are caused by suicide.

Suicide occurs in all regions of the world and throughout the life span, and it accounts for 1.4 percent of all deaths worldwide, by that, ranking as the 15th leading cause of death.

Among young people 15 to 29 years of age, suicide is the second leading cause of death globally after traffic accidents and accounts for 8.5 percent of all deaths.

Tables 32.1–1 and 32.1–2 show the number and suicide rates by age group in all WHO regions for males and females, respectively.

Suicide rates for males are highest in Europe, followed by Southeast Asia and Western Pacific. Suicide rates for females are highest in Southeast Asia and Western Pacific.

In Table 32.1–3, the absolute numbers, crude, and age-adjusted suicide rates per 100,000 persons in each state in the United States are shown. The age-adjusted suicide rate for the United States is 12.6 per 100,000 population. The highest age-adjusted suicide rates are in Montana, Alaska, Wyoming, Utah, New Mexico, Idaho, Colorado, Nevada, South Dakota, while the lowest are in the District of Columbia, New Jersey, New York, Massachusetts, and Connecticut.

Within the same region, suicide rates can vary from one district to another. Therefore, it is important for tailored suicide prevention to survey local data from both population and clinical settings, as risk groups and risk factors can vary considerably at the local level.

Socio-Demographic Factors

In low- and middle-income countries, the male-to-female ratio of age-standardized suicide rates is approximately 1:6 in comparison with high-

income countries, where the ratio is 3:5. Differences between genders in coping with stress and conflicts, patterns of alcohol consumption and help-seeking behaviors for mental disorders, as well as access to lethal methods to suicide can explain regional and gender differences.

Suicide rates are highest both for males and females in persons aged 70 years and older in nearly all regions of the world. Young adults and elderly women have, however, much higher suicide rates in low- and middle-income countries than in high-income countries. The opposite has been found among middle-aged men in high-income countries, who show a much higher suicide rate than middle-aged men in low- to middle-income countries.

Status of being unmarried, divorced, widowed, elderly, unemployed, immigrant, and having different sexual orientation are risk factors increasing the propensity for suicide. Moreover, individuals in the armed forces, veterans, prisoners, and indigenous people have an increased risk of suicide. Since Durkheim, religion, marriage, and parental responsibilities are known key factors that play a crucial role in controlling suicide impulses. Moreover, the impact of values, social political trends, societal patterns of drug and alcohol use, and violence is also of great relevance.

Table 32.1–1.
Estimated Numbers and Rates of Suicide by Age Group for Men in WHO Regions[a]

WHO Region	0–4	5–14	15–29	30–44	45–59	60–69	70–79	80+	Total
Numbers									
World	0	5,961	136,532	128,196	117,916	52,222	43,984	21,676	506,487
Africa	0	724	11,422	9,603	6,939	3,580	2,251	903	35,423
Americas	0	571	15,171	15,127	13,936	5,309	3,478	2,635	56,227
Eastern Medit.	0	548	6,753	4,585	2,787	981	896	681	17,230
Europe	0	509	19,635	25,842	28,615	10,480	9,329	5,070	99,480
SE Asia	0	2,502	64,548	49,869	33,048	9,965	6,776	2,052	168,761
Western Pacific	0	1,107	19,002	23,170	32,590	21,907	21,254	10,335	129,366
Rate per 100,000									
World	0	0.9	15.3	17.8	23.4	28.2	42.2	60.1	14.9
Africa	0	0.7	10	15.1	20.9	30.7	43	77.5	8.8
Americas	0	0.7	13.4	15.9	18.9	18.9	22.2	34.1	12.4
Eastern Medit.	0	0.8	7.6	8.1	8.8	9.5	17	44.7	5.8
Europe	0	1	19.9	26.8	33.6	29.2	37.9	53.2	23.1
SE Asia	0	1.4	25.6	26.8	27.5	25.6	35.8	40.4	18.8
Western Pacific	0	0.8	8.5	10.4	20.4	36.5	61.6	93.5	14.1

[a]Värnik P. Suicide in the world. *Int J Environ Res Public Health*. 2012;9(3):760–771.

Methods of Suicide

National level data on suicide methods are limited, as many countries do not collect this information. Only 76 out of the 194 WHO Member States reported data on methods of suicide in 2012. The coverage rate is much better in high-income countries than in low- and middle-income countries.

Hanging accounts for 50 percent of all suicides in high-income countries, while firearms account for 18 percent. Suicide methods and rates, however, vary between regions. In the Americas, for example, firearms account for 46 percent of all suicides, in comparison with other high-income countries where firearms account for only 4.5 percent of all suicides.

It is estimated that self-poisonings with pesticides are responsible for approximately 30 percent of global suicides in low- and middle-income countries. In some countries in Asia, jumping from high buildings is a common method of suicide. Since 1998, the use of barbeque charcoal, which produces the highly toxic carbon monoxide gas, as a means for suicide has spread rapidly in many Asian countries.

Suicide Trends

The global age-standardized suicide rate has fallen 23 percent in males and 32 percent in females from 2000 to 2012. The reasons for such positive change can be many. One likely explanation could be the improvement of global health due to the awareness of the problem and widespread use of new treatment methods.

Attempted Suicide

There is a lack of systematic registration of attempted suicides across national levels. Data on suicide attempts typically derive from self-reported surveys in the general population, or from hospital-based data of medically treated suicide attempts.

Table 32.1–2.
Estimated Numbers and Rates of Suicide by Age Group for Women in WHO Regions[a]

WHO Region	0–4	5–14	15–29	30–44	45–59	60–69	70–79	80+	Total
Numbers									
World	0	5,764	94,959	60,378	48,413	24,791	24,077	17,145	275,527
Africa	0	354	6,303	2,273	2,775	1,974	1,452	778	15,909
Americas	0	409	4,311	4,124	4,517	1,378	638	460	15,837
Eastern Medit.	0	555	8,124	3,443	1,776	410	444	353	15,104
Europe	0	203	4,001	5,228	7,017	3,464	3,559	2,924	26,395
SE Asia	0	3,413	59,109	23,793	10,378	5,284	2,635	1,070	105,683
Western Pacific	0	830	13,112	21,517	21,950	12,282	15,349	11,560	96,599

Rate per 100,000

World	0	1	11.2	8.6	9.5	12.4	18.7	27.8	8.2
Africa	0	0.3	5.5	3.6	7.7	14.7	22.3	44.5	3.9
Americas	0	0.5	3.9	4.3	5.8	4.4	3.3	3.5	3.4
Eastern Medit.	0	0.9	9.6	6.6	6	3.9	8	21.1	5.3
Europe	0	0.4	4.2	5.4	7.7	8.1	9.9	14	5.8
SE Asia	0	2	25	13.4	8.9	12.6	11.7	16.6	12.3
Western Pacific	0	0.7	6.3	10.1	14.1	20.5	39.5	64.7	11.1

[a]Värnik P. Suicide in the world. *Int J Environ Res Public Health*. 2012;9(3):760–771.

Table 32.1–3.
Number, Crude, and Age-Adjusted Suicide Rates per 100,000 of Population in Each State of USA for All Ages and Both Genders, 2013 (Ranked by the Age-Adjusted Suicide Rate)

State	Number	Crude Rate	Age-Adjusted Rate	Rank
United States	41,149	13.0	12.6	
Montana	243	23.9	23.7	1
Alaska	171	23.3	23.2	2
Wyoming	129	22.1	21.5	3
Utah	579	20.0	21.4	4
New Mexico	431	20.7	20.3	5
Idaho	308	19.1	19.2	6
Colorado	1,007	19.1	18.6	7
Nevada	541	19.4	18.6	8
South Dakota	147	17.4	18.0	9
Arizona	1,163	17.6	17.5	10
Maine	245	18.4	17.4	11
Arkansas	516	17.4	17.3	12
North Dakota	128	17.7	17.3	13
Oklahoma	665	17.3	17.2	14
Oregon	698	17.8	16.8	15
Vermont	112	17.9	16.8	16
West Virginia	323	17.4	16.4	17
Missouri	960	15.9	15.6	18
Kentucky	701	15.9	15.5	19
Tennessee	1,030	15.9	15.4	20
Kansas	425	14.7	14.7	21
Alabama				

	721	14.9	14.4	22
Iowa	447	14.5	14.4	23
Wisconsin	850	14.8	14.4	24
Indiana	944	14.4	14.2	25
South Carolina	696	14.6	14.0	26
Washington	1,027	14.7	14.0	27
Florida	2,928	15.0	13.8	28
Pennsylvania	1,788	14.0	13.4	29
Mississippi	388	13.0	13.0	30
Michigan	1,295	13.1	12.9	31
Ohio	1,526	13.2	12.9	32
New Hampshire	185	14.0	12.8	33
North Carolina	1,284	13.0	12.6	34
Delaware	122	13.2	12.5	35
Virginia	1,072	13.0	12.5	36
Louisiana	583	12.6	12.4	37
Rhode Island	132	12.6	12.2	38
Minnesota	678	12.5	12.1	39
Georgia	1,212	12.1	12.0	40
Hawaii	171	12.2	11.8	41
Texas	3,059	11.6	11.7	42
Nebraska	220	11.8	11.6	43
California	4,025	10.5	10.2	44
Illinois	1,321	10.3	9.9	45
Maryland	569	9.6	9.2	46
Connecticut	330	9.2	8.7	47
Massachusetts	572	8.5	8.2	48
New York	1,687	8.6	8.1	49
New Jersey	757	8.5	8.0	50
District of Columbia	38	5.9	5.7	51

Murphy SL, Xu J, Kochanek KD. *Deaths: Final Data for 2010*. Division of Vital Statistics, Centers for Disease Control and Prevention. USA; 2013.

Based on national surveys from 10 high-income, 6 middle-income, and 5 low-income countries, the 1-year prevalence of having made one or more suicide attempts in the year prior is estimated to be 4.1 per 1,000 adults. In relation to the estimated 2012 global suicide rate of 15.4 per 100,000 adults aged 18 years and over, the global estimate in the WHO 2014 report is that for each adult who committed suicide there were 27 persons who made one or more suicide attempts. There is a wide variation in the attempt-to-suicide death ratio with respect to age, sex, region, and methods used.

Suicide attempts are known to be the single most important risk factor

for completed suicide. In spite of this, there are very few countries that contain registries of medically treated suicide attempters.

In clinical work, the presumed continuum from suicidal ideation through suicide attempts to suicide completions can be a helpful model in understanding the suicidal process, risk detection, and treatment choices, even if this model cannot be applied to all suicides.

PSYCHIATRIC DISORDERS AND SUICIDE

Risk factors for suicide are both individual and familial. Suicidal behaviors aggregate in families, and family history of suicidal behaviors is an independent risk factor for suicide attempts and completed suicides. The presence of current and lifetime psychiatric diagnoses like major depressive disorder (MDD), bipolar disorders, anxiety disorders, alcohol and substance misuse, schizophrenia, eating disorders, personality disorders, different types of trauma, chronic somatic disorders, and current stressful life events are significant risk factors for suicidal behaviors.

In the context of suicide, there is a growing body of evidence showing that exposure to early-life maltreatment can affect molecular mechanisms involved in the regulation of behavior through methylation and histone modification, supposed to induce behavioral deviations during the early development, and possibly later in life, affect genes involved in crucial neural processes. This mechanism is called epigenetics. Childhood abuse and other detrimental environmental factors seem to target the epigenetic regulation of genes involved in the synthesis of neurotrophic factors and neurotransmission.

MDD is strongly linked with suicide, especially if long-term comorbidity and acute negative life events are present.

The clinical picture is characterized by symptoms like weight loss or gain, sleep disturbances, fatigue, concentration difficulties, changes in psychomotor capabilities, feelings of worthlessness, guilt, and recurrent thoughts of severe suicidal ideation with suicidal plans. In the melancholic type of MDD, despondency, despair, irrational guilt, and emptiness are profound symptoms.

Patients diagnosed with dysthymia (persistent depressive disorder) complain of irrational patterns of negative thinking and chronic dysphoria.

Depressed suicidal patients have similar risk factors for suicide as patients with other psychiatric diagnosis, namely, living alone, being unmarried, unemployed, and a history of previous suicide attempts. Recurrent MDDs are robust precipitants of suicidal behaviors. Physical illnesses also increase the risk of suicide and attempted suicide in depressed patients, especially in the elderly population.

The most common comorbidity for affective disorders in suicidal behavior is alcohol and other substance use, multiple physical impairments, and personality disorders. Independently of the direction

concerning the link between different forms of depression and comorbidity, it is important from the clinical point of view to treat all disorders.

The risk of suicide varies between the different subtypes of depression. Affective temperaments like cyclothymic disorder and anxiousness are associated with both suicide attempts and suicide. Such temperaments, along with irritability and rapid mood switches, are important contributors in triggering suicidal acts.

Bipolar Disorders

Bipolar disorders are separated from other depressive disorders in the DSM-5 classification and are placed between the diagnostic class on depressive disorders and schizophrenia/other psychotic disorders, in order to denote a bridge between those two diagnostic classes with respect to symptoms, family history of psychiatric disorders, and molecular genetic findings. Bipolar disorders constitute a high-risk group for suicide and attempted suicide. The majority of persons with bipolar disorders commit suicide when they are in a major depressive episode or in a mixed depressive state. Suicide during the manic phase is rare. The prevalence of suicide attempts is similar in both type I and type II bipolar disorders. Comorbidity of substance use disorders, depression, and anxiety is almost always present in persons who committed suicide.

Effective pharmacological treatment supported by psychological techniques is the foremost strategy to prevent completed suicide.

Alcohol and Substance Use Disorders

Excessive alcohol and other substance misuse leading to significant impairment and distress are well-recognized conditions for an increased risk of suicidal behaviors. Comorbidity with personality disorders magnifies the risk of suicide. Suicide mortality is highest among drug users and lower, but still high, among persons with alcohol use disorders, as well as more prevalent among males compared to females.

Alcohol and other substance misuse increase aggressivity, impulsivity, and cause deterioration in cognitive capacity and flexibility to find constructive coping strategies. Suicidal alcohol misusers, however, have a fairly good psychosocial coping ability, which could make their suicide appear more astounding. Good psychosocial functioning could potentially explain why they did not seek treatment or receive attention from significant others to motivate them for treatment.

Triggers for suicidal behaviors in persons with alcohol and substance use disorders are losses of important relationships, work, economical security, and self-esteem. The more dependence the substance users' experience in their relationships, the greater the risk that separation may push them into self-destructive acts.

Guidelines for the treatment of alcohol and substance misuse are published by the American Psychiatric Association (APA) and by the National Institute for Health and Care Excellence (NICE). The duration of treatment may vary, but it is important to use both pharmacological and psychological treatments.

Decreasing the accessibility to alcohol both on a societal and individual level, as well as moulding attitudes toward alcohol intake show significant results in diminishing suicide, as described during the Perestroika period in the former USSR.

Anxiety Disorders

Anxiety as a risk factor for suicide has been neglected for a long time, that is, until studies by Fawcett et al. put forward the role of severe anxiety in precipitation of suicidal behaviors. There are many ways to assess the severity of anxiety, and it is important that suicidal psychiatric patients, especially those with mood disorders, substance use disorders, and in psychotic states are assessed for the presence of severe anxiety and treated for it.

Patients who suffer from severe anxiety may sometimes deny suicidal thoughts or suicidal intent and refuse hospitalization. It is recommended to document in the medical journal, the patient's decision and secure that outpatient treatment includes frequent follow-up and, whenever possible, involve the family.

Schizophrenia and Other Psychotic Disorders

In the schizophrenic population, estimates show that approximately 5 to 10 percent die due to suicide. Already Bleuler drew clinicians' attention that the most serious of the schizophrenic symptoms is the suicidal drive. Moderate-to-severe depression is one of the most frequent features of schizophrenic patients who commit suicide.

Schizophrenics who commit suicide usually have poorer treatment compliance, not seldom due to the side effects of antipsychotic medication like akathisia. Suicides occur after abrupt discontinuation of medication. Negative attitudes toward medication and treatment are generally high in both schizophrenic and suicidal patients with other psychiatric diagnosis. Poor treatment compliance, social isolation, and increased expectation of good performance from others and from patients themselves, are risk factors for suicide in schizophrenics.

For a long time, there has been controversy surrounding the question whether schizophrenics commit suicide during the intense and frightening psychotic activity or during periods of remission. Studies found that command hallucinations are rare among completed suicides.

On the contrary, a good premorbid functioning and higher level of

education may predispose to suicide in younger schizophrenic patients, as they experience more disruption of performance and may have more difficulty to accept chronic illness and prospects of mental deterioration than older schizophrenic patients. When giving information about the diagnosis, course of illness, and treatment, one should be aware that it is a risk situation for suicide.

Eating and Adjustment Disorders

Patients with anorexia nervosa have an increased suicide risk, thus suicide risk assessment should be included in a comprehensive clinical evaluation. Patients with bulimia nervosa and binge eating disorders also have an increased risk of suicidal behaviors. A high comorbidity of mental illnesses like MDD, bipolar disorder, anxiety disorders, and borderline personality disorders (BPDs) among persons who are underweight as well as overweight is to a great extent responsible for the high suicide risk in all eating disorders. Treatment is often complicated and requires behavioral therapy combined with medication and supportive measures. Low body mass index (BMI) and low serum cholesterol have been shown to be associated with a higher risk of attempted and completed suicide.

Prevention

Organizing home visits, case management, and regular telephone contacts with somatically ill and vulnerable elderly persons are effective preventive methods, as it diminishes isolation and provides the opportunity to early detect risk factors and risk situations for suicide.

PERSONALITY DISORDERS AND SUICIDE

Personality disorders have been previously understudied in clinical and population studies on suicidal behaviors. In 894 cases of suicide among young people, 11 percent received personality diagnosis, mainly antisocial and borderline. The remaining diagnoses were paranoid, narcissistic, anxious, and histrionic personality disorders. This low percentage depends on, among other factors, the difficulties of retrospective diagnostics of personality disorders.

Persons diagnosed with Cluster B comprising borderline, antisocial, histrionic, and narcissistic personality disorders are characterized by poor impulse control, aggressivity, unstable identity, poor self-image, affective instability with proneness to experience and strongly react to real or imagined environmental stresses and losses. Difficulties with interpersonal relationships, problem solving, and inability to plan and think about the future in positive terms increase the propensity for triggering suicidal behaviors.

Cluster C personality disorders, which include avoidant and obsessive

compulsive disorders, are not strongly associated with suicide. Only dependent personality disorder in the Cluster C is significantly associated with suicide attempt, but this association is likely due to the comorbidity of lifetime and concurrent depressive disorders.

Among Cluster A personality disorders, mainly schizoid personality disorder has been shown to increase the risk of suicide. Schizoid personality disorder is quite rare, both in the general and clinical populations, and therefore, it is difficult to estimate a true prevalence and association with suicidal behaviors.

Personality disorders contribute even more to the risk of suicidality when comorbid with other psychiatric diagnosis, and in the presence of negative life events, especially those involving interpersonal distress and loss. Concurrence of psychiatric disorders with personality disorders is now recognized as a major factor in suicide and in attempted suicide.

SOMATIC DISORDERS AND SUICIDE

Studies performed in the United States, Australia, and Europe show an association between suicidal behaviors and chronic somatic disorders.

Somatic disorders involve physical, psychological, and social implications. They imply stress, pain, sometimes handicaps, limited social performance, decrease in the capability to work, and the increased need for help from others. The comorbidity of somatic disorders with psychiatric disorders, especially with MDD and personality disorders substantially increases the risk for suicide. Age and separation from loved ones, loneliness, hopelessness, helplessness, and social isolation are parameters of importance for suicide risk.

The suicidal situation and the suicidal propensity vary during the course of the somatic disorder and depend on the treatment outcomes and pain control, as well as on psychosocial comfort.

Cancer, HIV infection and AIDS, stroke, diabetes mellitus, epilepsy, Parkinson disease, trauma with subsequent brain damage, spinal cord injury, multiple sclerosis, Huntington disease (HD), and amyotrophic lateral sclerosis are associated with an elevated risk of suicide.

Elderly patients with chronic or incurable diseases need to have an adequate somatic and psychiatric treatment, as well as good psychosocial care in times of shrinking economical resources. Different opinions on the distribution of economical resources can sometimes be associated with advocacy concerning euthanasia and assisted suicide.

SPECIAL POPULATIONS

Children and Adolescents

Poor Mental Health. Mental ill-health is the leading cause of disability in young persons aged 10 to 24 years and is responsible for 45 percent of

the overall burden of disease in this age group. The high prevalence of mental health disorders was shown in the Great Smoky Mountains prospective cohort study in the United States, as well as in the National Comorbidity Survey, also in the United States.

The results of the Saving and Empowering Young Lives in Europe (SEYLE) study showed that in a usual school class of teenagers approximately 10 percent of young people display high-risk behaviors like excessive alcohol and illegal drug use, heavy smoking, truancy, etc. Thirty percent display unhealthy lifestyles like poor sleep and diet, physical inactivity, excessive Internet/media use. Both of those groups have high levels of depression, anxiety, and suicidal behaviors. Approximately 12.5 percent of adolescents required qualified psychiatric and psychological help. However, experiences from the SEYLE study in Europe and the Columbia University TeenScreen program in the United States show that there is a fear of stigma when screening for mental health problems in schools, and both parents and adolescents show little trust in the mental health care systems. Therefore, the process of destigmatization of screening for mental disorders and increasing help-seeking behaviors is an important public health issue. SEYLE results show that younger adolescents are more likely to adhere to rules, procedures, and recommendations with higher attendance rates to health care system for girls and for pupils victimized by peers. Suicidal behaviors also appear to be an important predictor for help seeking and a predictor of referral to mental health care.

Risk for Suicide. For most adults, it is difficult to acknowledge the child's despair and suicidality. Caretakers as well as parents can deny serious childhood diagnosis due to guilt feelings and frustrations when they have difficulties to handle a suicidal, depressed child or a suicidal child with serious conduct difficulties. Suicidal behavior in children and adolescents occurs in the context of stressful, chaotic, and often unpredictable family events. Suicidal children and adolescents have poor self-esteem and poor personal identify, are often truant, not seldom bullied, and have poor school grades, which leads to a sense of inadequacy. Issues of gender identity are well recognized as risk factors for adolescent suicides. The families of these children and adolescents show a high incidence of affective disorders, alcohol misuse, and other psychiatric diagnosis. The major risk factors for suicide in young people are the presence of a psychiatric disorder, especially affective disorder, substance misuse, and BPD. Attempted suicide is an important risk factor for future suicide.

The four comorbid clinical constellations described by Apter et al. having a special significance for young suicides are; the combination of schizophrenia, depression, and substance misuse; the combination of substance misuse, conduct disorder, and depression; the combination of

affective disorder, eating disorder, and anxiety disorders; the combination of affective disorder, personality disorder of paranoid and schizoid type, and dissociative disorders characterized by disruption in integrated functions of consciousness, memory, identity, or perception of the environment.

Treatment. The treatment of suicidal children and adolescents should include the individual child and the entire family. The aim is to improve coping styles and communication between parent and child. Psychosocial rehabilitation measures including school consultation and academic remediation are important.

The treatment of underlying psychiatric disorders requires psychotherapeutic and pharmacological treatment when deemed necessary. When medication is used in the management of suicidal young persons, family can get a false sense of security, as well as the fantasy that the problem has been solved, which lowers motivation to work through the problems in family and school situation.

If child or adolescent manifests psychosis, then hospitalization is required. Hospitalization is also necessary in case of alcohol and drug misuse in the context of a disruptive home situation. If a young person perceives rejection in the family, a replacement home can be a solution during the period of treatment and rehabilitation.

Elderly

Risk Factors. Elderly persons have the highest suicide rate compared to any other age group. Decreased economical resources after retirement, placement into long-term care or residential care can increase the risk of suicide. The elderly suicides are characterized by high lethality methods, social isolation, and the presence of physical illnesses. The ratio of suicide attempt to suicide is very low among the elderly, and therefore each suicide attempt in older persons should be taken seriously, being an important harbinger of completed suicide.

Diagnoses of MDD are the most common in elderly suicides, followed by diagnosis of alcohol misuse and organic brain syndrome. Changes of sleep pattern, appetite changes, as well as somatic complaints or exacerbation of chronic concerns, which stop an elderly person from going out can be a sign of depression. Symptoms of guilt, sadness, or anhedonia can be masked by somatic expressions. Unconscious negative attitudes toward the elderly may create a risk of not taking suicidal risk in older persons seriously.

Dementia

It is not seldom that cognitive functions are severely impaired and

depression can sometimes be misdiagnosed as dementia. Evidence shows that suicide risk in people with dementia is low. However, in the period immediately after diagnosis of dementia, especially in those of younger age, the suicide risk is high.

Treatment

The treatment of suicide risk in the elderly does not differ from suicide treatment in other age groups. In cases of severe depression, when medication does not help, electroconvulsive therapy (ECT) can be recommended. In the treatment of the elderly, drug-to-drug interactions are important to observe, as the elderly often have multiple medications.

In talkative therapy, it can be of value to focus on what the person has accomplished in life, rather than to talk about what may or may not lie ahead. To discuss suicidal ideation can be part of a process that puts one's life into perspective, and by acknowledging their suicidal thoughts, the clinician can help the patient move away from the wish to die. Inclusion of family and other resources in the community should be encouraged in order to diminish loneliness and social isolation.

SUICIDE RISK ASSESSMENT

Decisions regarding the level of suicide risk are made every day in all clinical specialties, even if suicide is regarded as a psychiatric domain. There are many obstacles when performing the suicide risk assessment due to the uniqueness of each individual possessing a unique set of risk factors. Uncontrollable environmental influences, which contribute to the outcome, are another obstacle.

Clinical Interview

In suicide risk assessment, the most important instrument is clinical inquiry. This inquiry has an impact on the suicidal patients' communication which should be clear, empathetic, free from prior criticism, and focus on facts and patients' emotional and communication style.

In a systematic clinical assessment of suicide risk, the presence of severe anxiety, MDD, use of alcohol or illicit substances, and previous suicide attempts should be covered. The presence of suicidal communication and personality type should be scrutinized. Previous psychiatric disorders, suicide or suicidal behaviors in the family, as well as negative life events along with feelings of helplessness, hopelessness, sleep problems, and impulsive behaviors should be examined.

Repeated assessments are recommended to understand the mechanisms that generate the suicidal urge or impulse. Interview with the family or significant others is an important source of information.

However, permission by the patient must be given and it can sometimes take time to motivate the suicidal person to involve the family.

The evaluation of the social network and quality of family support should always be included when suicide risk assessment is done before temporary or final discharge from the hospital.

Psychometric Scales

A set of scales can be used to compliment clinical inquiry. The Suicide Intent Scale (SIS), Scale for Suicidal Ideation (SSI-C), the Beck Hopelessness Scale, the SAD PERSONS SCALE, and the Columbia Suicide Severity Rating Scale (C-SSRS) are used for assessing the risk of suicide.

Some biological and genetic markers have been described, but none of them has a practical application for clinical suicide risk assessment at the present moment.

The final decision in the suicide risk assessment is always based on an intuitive judgment and therefore the clinician should be aware of feelings toward a suicidal patient, which can contribute to a denial of suicide risk, a rigid approach to the patient, or even negative emotional states like ambivalence or hostility. Those feelings and behaviors are called countertransference.

MANAGEMENT OF THE SUICIDAL PATIENT

Hospitalization or Not

Following the assessment of suicide risk, the first important decision to be taken is the choice between admission of the patient to a psychiatric ward or treatment in outpatient care. The decision does not depend only on the status of the patient, but also on the quality of the available family support and the accessibility to qualified outpatient treatment.

Hospitalization is usually not the first option for most suicidal patients, but in severe cases it provides security, if safety measures at the psychiatric ward are followed. In the management of the suicidal patient in the hospital, precautions are important in restricting access to means of suicide in the ward and establishing different degrees of freedom while maintaining an empathetic encounter.

Suicide in the Ward

When a suicide or a suicide attempt takes place in the psychiatric ward, it is important to give support to all patients. The regular suicide risk assessment should be carried out for patients judged to be at risk. In addition to the usual treatment, an intensive psychosocial support should be given, to work through the emotions and the possible identifications with the patient who committed suicide at the same or neighboring unit. Also, the personnel at the ward should be given psychological support.

Treatment of Acute Suicidal Crisis

In the treatment of the acute suicidal crisis, it is important to focus on reducing anxiety, insomnia, depression, and psychotic symptoms. Collaboration with the family and significant others must be secured. Safety measures should also be considered in the home environment and family members should be alerted of the increased risk of suicide after discharge from the hospital. The plan for follow-up treatment and establishing collaboration with rehabilitation units and community services for chronically suicidal patients, especially those with the diagnosis of substance misuse, schizophrenia, or other psychotic disorders is recommended.

Finally, it is important to have a proper documentation about communication between different departments, family, social services, and have a clear follow-up plan.

The patient and the family should also be informed that they need to seek psychiatric emergency services as soon as a new suicidal crisis arises.

PHARMACOTHERAPY

Mood-Stabilizing Treatments

Lithium. Meta-analyses, reviews, and original studies show that long-term lithium treatment is associated with a substantial reduction of the risk for suicide and attempted suicide in patients with bipolar spectrum disorders. The analyses by the US Center for Disease Control and Prevention (CDC) show that mortality by overdoses of lithium is similar to the mortality from overdoses of new antidepressants and antipsychotic agents.

Ecological studies also show that lower rates of suicide as well as other violent behavior are found in populations exposed to drinking water with relatively high concentration of lithium salts. There is no plausible explanation for this finding as the daily intake of lithium in the drinking water is much lower than dosages required for clinically effective treatment of mood disorders.

Other Mood Stabilizers. In bipolar depression, other mood stabilizers, such as anticonvulsants or second-generation antipsychotics have beneficial effects on suicidal behavior. However, lithium is noted to be significantly superior in reducing suicidal behaviors.

Antidepressants. Selective serotonin reuptake inhibitors (SSRIs) are nowadays widely used in the treatment of suicidal patients with MDD and related conditions. When choosing an antidepressant for a suicidal depressive patient, compounds with a sedative profile are preferable. Medications that increase drive may increase the risk of suicide.

Adverse outcomes in some patients during treatment with SSRI antidepressants like agitation, restlessness, irritability, dysphoria, anger and insomnia can worsen the suicide risk in these patients. Due to those responses to serotonergic antidepressants, especially in some children and young adults below the age of 25 years, the Food and Drug Administration (FDA) and the European Medicine Agency (EMA) recommend continuous and competent monitoring of side effects. Antidepressants may also increase suicide risk in bipolar patients.

Electroconvulsive Therapy. Cases of depression with suicidality, which are difficult to treat by other means, can be treated by ECT, which has a rapid onset of action and relief of symptoms.

Antipsychotics. Suicidal symptoms in schizophrenic patients require, in addition to the standard treatment of the schizophrenia with antipsychotics, an additional medication to control anxiety or agitation. Sedating antipsychotics are mostly used in these cases. When high dosages are required, attention must be paid to the risk of acute hypotension with tendency to collapse when standing up. In some cases, to better control anxiety, benzodiazepines can be administered for a short period of time.

In the case of post-psychotic depression, an additional treatment with antidepressants is recommended. If the suicidal symptoms are part of a depressive syndrome due to the side effect of neuroleptic treatment, the dosage should be reduced, if possible.

Anxiety Control. In controlling anxiety as a comorbid symptom to depression and in anxiety disorders, antidepressants are the preferred choice for acute and maintenance treatment. In cases where treatment response is not sufficient, adding short-term administration of benzodiazepines may be necessary.

PSYCHOTHERAPY

Treatment in prevention of suicide requires a complex approach of psychosocial, psychotherapeutic, and psychopharmacological interventions. Suicide cuts across the whole spectrum of psychiatric illnesses and therefore therapists should be familiar with several therapeutic schools and strategies as they give valuable contributions to the treatment of suicidal persons. The choice of treatment depends on the condition of the patient. The combination of psychopharmacological treatment with psychotherapy should always be taken into consideration as part of a complex treatment strategy.

Cry for Help

Suicidal communication both verbal and nonverbal must be taken

seriously, independent of whether the clinician considers that the patient has a genuine wish to kill him/herself or if the purpose is to get attention. Suicidal communication always conveys an underlying sense of helplessness and hopelessness and a cry for help.

Risk of Suicide during Treatment

Treatment of suicidal patients is challenging, as the possibility of death by suicide is ever present, independent of whether the treatment is pharmacological or psychotherapeutic. Establishing an empathetic relationship with the patient when the threat that the patient might commit suicide is present, can be difficult. This can be a discouraging condition to work under.

Treatment of a suicidal person requires self-involvement and confrontation with one's own attitudes toward life and death.

In treating suicidal patients, one must allow the patient to feel suicidal and not to minimize, reject, or deny the seriousness of their suicidal intent. Suicidal patients must be allowed to feel that their pain can be shared with another person. The challenge is to follow the patient and to understand, as well as to utilize psychotherapeutic methods without denying suicidality.

Transference and Countertransference

Suicidal patients have many different feelings toward their therapists and doctors, which in the psychotherapeutic literature are called transference. Therapists and doctors also have feelings toward their patients, which are called countertransference. Countertransference feelings toward suicidal persons can be empathetic but ambivalent, aggressive, and even hostile. It implies understanding of one's own attitudes and values, as well as origins of suicidal patient positive and negative feelings toward the clinician. A therapeutic alliance, which is important to establish, comprises elements of a real relationship between the therapist/doctor and the patient, as well as aspects derived from patient's experiences of earlier relationships, which are transferred to the present treatment situation.

Long-Term Psychological Treatments

Two long-term–based treatments, cognitive behavioral therapy (CBT) and dialectical behavioral therapy (DBT), show the strongest evidence in preventing suicidal behaviors.

There are also a sufficient number of studies of good quality showing promising results in the treatment of suicidal patients with family psychotherapy, developmental group therapy, psychodynamic therapy, and in-home interpersonal psychotherapy. Interpersonal psychotherapy (IPT) and psychodynamic approaches like transference-focused psychotherapy, mentalization-based treatment, and sequential Brief–Adlerian

psychodynamic psychotherapy are supported by randomized controlled trials in which the psychodynamic approach was found to be effective in reducing suicidal behavior among patients with BPD. Other therapies like Collaborative Assessment and Management of Suicidality (CAMS) are using a combination of therapeutic techniques from multiple previously studied interventions.

Brief Psychological Treatments

Brief psychosocial suicide preventive interventions with some empirical support are CBT-based interventions, contact interventions, combination of those two approaches, and manual assisted cognitive behavioral therapy (MACT).

Psychosocial support and involvement of family, friends, and significant others after obtaining consent from the suicidal person is also recommended.

Supportive Telephone Calls and Letters

Treatments consisting of supportive telephone calls, SMS, or written contacts after discharge of the suicidal patient from the emergency department or hospital showed preventive effects. Involving community resources with the goal that the social intervention should give support to the patient in breaking through loneliness and finding social networks, and in this way, enhancing a sense of meaning in their lives is essential.

SURVIVORS

The term suicide survivors refers to those who have lost a significant other in suicide. When suicide occurs, the family and the surrounding people experience severe trauma and a range of reactions from shock, sense of unreality, sadness and grief reactions, anger, and feelings of abandonment and rejection. Suicide survivors experience a devastating event and desperately seek an explanation of what has happened. Not seldom they take contact with the family doctor or doctor who treated the patient and who also is a survivor of the patient's suicide. It can be wise to ask an experienced colleague to help with advice on how to manage this situation. The doctor needs not only to work with their grief, but also to assess the potential suicide risk in survivors, who have a high risk for suicide.

The development of support groups can be considered as an important step forward to work through the aftermath of suicide in a compassionate and supportive way, educate about grief, demystifying suicide, eliminating excessive guilt, shame, and blaming. The WHO issued guidelines on how to form support groups.

PUBLIC HEALTH PERSPECTIVE AND PREVENTION

Controlling the Environment

The fatality of suicide methods varies and the rationale behind WHO recommendations to diminish access to suicide methods is to curb a sudden and strong impulse to end one's life. Several studies on the restriction of access to firearms, pesticides, carbon monoxide and vehicle exhaust, charcoal burning, placing barriers at hot spots used for jumping, reducing access to medications, and controlling access to railroads and subways, is an effective approach in suicide prevention and should be incorporated not only at the public health level but also at the clinical level.

One of history's most effective suicide prevention program was observed in the second half of the 1980s in the former USSR during *Gorbachev's Perestroika* (restructuring), characterized not only by openness and freedom, but also by the strict limitations on the access to alcohol by cuts in the production and sale, primarily with the aim to improve the health of the population. Suicide for men decreased by 40 percent in the years 1984 to 1986 in comparison with 3 percent in 22 European countries during the same period. This decline occurred in all 15 republics of the former USSR. The sharp decrease in suicide applied to women as well as men. The largest decreases were observed among men in the workforce aged 25 to 54 years. No corresponding decline for this age group was noted in any country during the 20th century.

The Role of the Media in Suicide Prevention

Sensational and irresponsible reporting by different types of media may precipitate or induce suicidal acts through imitation or identification mechanisms in suicidal persons. The WHO issued guidelines on media coverage, describing how the press and broadcasting media should report on suicide in order to avoid copycat effects. It is essential to avoid the description of suicide as courageous or desirable.

School-Based Suicide Prevention

Children and young people spend many hours each day at school and the school is a good arena for comprehensive suicide preventive work involving a continuum of activities from mental health promotion, prevention, intervention (treatment), and postvention when suicide occurs in school.

Mental health intervention in school gives the opportunity to overcome mental health problems and inequalities while adolescents' brains are still developing.

Results from randomized controlled trials of classroom-based intervention programs in the United States, like Signs of Suicide (SOS) and Good Behavior Game (GBG), along with the SEYLE study show that suicidality can be prevented.

The Youth Aware of Mental Health (YAM) suicide preventive program

used in SEYLE aims at changing adolescents' negative perceptions and improving coping skills in the management of negative life events and stressors causing suicidal behaviors. The YAM program has reduced severe suicidal ideation with suicide plans and suicide attempts by 50 percent. This effect is higher than effects noted in other successful public health interventions like bullying and bullying victimization (7 to 23 percent), and school-based interventions addressing smoking cessation (14 percent). The YAM program is perceived as stimulating by young people and adult instructors. However, in each classroom, there are a small number of children and adolescents in need of qualified psychiatric care.

To secure both preventive aspects and treatment aspects for young people in schools, stakeholders must build continuously, models for collaboration across the education and health care systems interface.

References

American Psychiatric Association Steering Committee on Practice Guidelines. *Treatment of Patients with Substance use Disorders*. 2nd ed.; 2010. Available at https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/substanc

Apter A, Gvion Y. Adolescent suicide and attempted suicide. In: Wasserman D, ed. *Suicide: An Unnecessary Death*. 2nd ed. London: Oxford University Press; 2016.

Brent DA, Melhem N. Familial transmission of suicidal behavior. *Psychiatr Clin North Am*. 2008;31(2):157–177.

Carli V, Hoven CW, Wasserman C, et al. A newly identified group of adolescents at "invisible" risk for psychopathology and suicidal behavior: findings from the SEYLE study. *World Psychiatry*. 2014;13(1):78–86.

Chesin M, Yuryev A, Stanley B. Psychological treatments for suicidal individuals. In: Wasserman D, ed. *Suicide: An Unnecessary Death*. 2nd ed. London: Oxford University Press; 2015.

Dunne-Maxim K, Dunne EJ. Family involvement in suicide prevention and postvention: a psychoeducational perspective. In: Wasserman D, ed. *Suicide: An Unnecessary Death*. London: Martin Dunitz; 2001.

Fawcett J, Kravitz HM. Anxiety syndromes and their relationship to depressive illness. *J Clin Psychiatry*. 1983;44(8 Pt 2):8–11.

Fleischmann A, Bertolote JM, Belfer M, Beautrais A. Completed suicide and psychiatric diagnoses in young people: a critical examination of the evidence. *Am J Orthopsychiatry*. 2005;75(4):676–683.

Leadholm AK, Rothschild AJ, Nielsen J, Bech P, Ostergaard SD. Risk factors for suicide among 34,671 patients with psychotic and non-psychotic severe depression. *J Affect Disord*. 2014;156:119–125.

Lönnqvist J. Major psychiatric disorders in suicide and suicide attempters. In: Wasserman D, Wasserman C, eds. *Oxford Textbook of Suicidology and Suicide Prevention: A Global Perspective*. London: Oxford University Press; 2009:275.

Lönnqvist J. Physical illnesses and suicide. In: Wasserman D, ed. *Suicide: An Unnecessary Death*. 2nd ed. London: Oxford University Press; 2015.

McGorry PD, Goldstone SD, Parker AG, Rickwood DJ, Hickie IB. Cultures for mental health care of young people: an Australian blueprint for reform. *Lancet Psychiatry*. 2014;1(7):559–568.

Murphy SL, Xu J, Kochanek KD. *Deaths: Final Data for 2010*. Division of Vital Statistics, Centers for Disease Control and Prevention. USA; 2013.

Nordentoft M, Wahlbeck K, Hallgren J, et al. Excess mortality, causes of death and life expectancy in 270,770 patients with recent onset of mental disorders in Denmark, Finland and Sweden. *PLoS One*. 2013;8(1):e55176.

Novick DM, Swartz HA, Frank E. Suicide attempts in bipolar I and bipolar II disorder: a review and meta-analysis of the evidence. *Bipolar Disord*. 2010;12(1):1–9.

O'Dwyer ST, Leo DD. The elderly and suicide. In: Wasserman D, ed. *Suicide: An Unnecessary Death*. 2nd ed. London: Oxford University Press; 2015.

Oquendo MA, Perez-Rodriguez MM, Poh E, et al. Life events: a complex role in the timing of suicidal behavior among depressed patients. *Mol Psychiatry*. 2014;19(8):902–909.

Pilling S, Yesufu-Udechuku A, Taylor C, Drummond C; Guideline Development Group. Diagnosis, assessment, and management of harmful drinking and alcohol dependence: summary of NICE guidance. *BMJ*. 2011;342.

Rihmer Z, Gonda X, Torzsa P, Kalabay L, Akiskal HS, Eory A. Affective temperament, history of suicide attempt and family history of suicide in general practice patients. *J Affect Disord*. 2013;149(1–3):350–354.

Scott MA, Wilcox HC, Schonfeld IS, et al. School-based screening to identify at-risk students not already known to school professionals: the Columbia suicide screen. *Am J Public Health*. 2009;99(2):334–339.

Sokolowski M, Wasserman J, Wasserman D. An overview of the neurobiology of suicidal behaviors as one meta-system. *Mol Psychiatry*. 2015;20(1):56–71.

Sörberg A, Gunnell D, Falkstedt D, Allebeck P, Åberg M, Hemmingsson T. Body mass index in young adulthood and suicidal behavior up to age 59 in a cohort of Swedish men. *PLoS one*. 2014;9(7):e101213.

Stenager E, Stenager E. Somatic diseases and suicidal behaviour. In: Wasserman D, Wasserman C, eds. *Oxford Textbook of Suicidology and Suicide Prevention: A Global Perspective*. London: Oxford University Press; 2009:293–299.

Turecki G. Epigenetics and suicidal behavior research pathways. *Am J Prev Med*. 2014;47(3 suppl 2):S144–S151.

Värnik P. Suicide in the world. *Int J Environ Res Public Health*. 2012;9(3):760–771.

Wasserman D, Hoven CW, Wasserman C, et al. School-based suicide prevention programmes: the SEYLE cluster-randomised, controlled trial. *Lancet*. 2015;385:1536–1544.

Wasserman D, Rihmer Z, Rujescu D, et al. The European Psychiatric Association (EPA) guidance on suicide treatment and prevention. *Eur Psychiatry*. 2012;27(2):129–141.

Wasserman D, Värnik A. Suicide-preventive effects of perestroika in the former USSR: the role of alcohol restriction. *Acta Psychiatr Scand Suppl*. 1998;394:1–4.

Wasserman D. *Suicide: An Unnecessary Death*. 2nd ed. London: Oxford University Press; 2015.

Wasserman D, Wasserman C. *Oxford Textbook of Suicidology and Suicide Prevention: A Global Perspective*. Oxford: Oxford University Press; 2009.

World Health Organization. *Preventing Suicide: A Global Imperative*. Geneva, Switzerland:

World Health Organization; 2014.

▲ 32.2 Suicide Treatment

HOWARD S. SUDAK, M.D.

INTRODUCTION

This chapter focuses on ways to identify individuals and groups at particular risk for suicide and how to prevent them from acting on such impulses (Table 32.2–1). Although the theoretical goal as clinicians may be to prevent all suicide attempts and completions, it is clearly not possible—however, reducing attempts and completions is both possible and feasible. Furthermore, "suicide treatment" encompasses a broad spectrum of "patients." Whom are we treating? The attempter, the individual considering it, the family of the attempter or completer, or even the public when the victim is prominent or the act so dramatic? A discussion of working with each of these groups follows.

The Attempter. A patient who indicates that he or she is suicidal or one who has just attempted suicide represents a psychiatric emergency. And, since patients who express intent to kill others are extremely rare, attempters and those planning an attempt comprise the majority of the psychiatric emergencies managed by clinicians, aside from pharmacological and general medical emergency situations. Since the condition of having serious thoughts of suicide may be a "symptom" of a disease, not a disease per se, conventional wisdom has dictated that it is the disease which needs to be treated, not one of its symptoms. One treats a depression with antidepressants, a psychosis with antipsychotics, a mania with mood stabilizers, etc. But this is not universally true. Clinicians treat other symptoms (e.g., fever, headaches, and most viral infections symptomatically with aspirin); insomnia with hypnotics; constipation with laxatives, etc. and, to some degree, this holds true for suicide as well. Dialectical behavior therapy (DBT), a treatment specifically designed for patients with BPD also appears useful in some other impulsive psychiatric conditions such as nonsuicidal self-cutting, as well. Similarly, lithium, a treatment for stabilizing bipolar patients, appears to be useful in other conditions. Antipsychotics are not uncommonly used for nonpsychotic agitation, and so forth. The most common condition associated with suicide is depression, and the psychotherapies and pharmacotherapies which address this are all part of the arsenal employed to decrease the risk of a suicide "attempt" or "completion." (Note that the term "completion" is preferable to "successful" suicide with regard to treatment, since an attempt or completion is hardly a success.) Unfortunately, "suicide

attempter" has become almost a pejorative label, and such individuals too often are treated dismissively and peremptorily by emergency personnel, as though the attempt is merely a way to control or manipulate others. Certainly, most emergency rooms (ERs) are understaffed with stressed-out personnel who too often view attempters as unnecessary burdens which may stem from staff's own discomfort, anxiety, and lack of knowledge. A recent follow-up study revealed that from a total of 65,784 patients seen in emergency rooms in Canada, during a median follow-up period of 5.3 years, 4,176 had died, including 976 (23.4 percent) by suicide. This is hardly a low-risk subpopulation.

Table 32.2–1.
Risk Factors for Suicide

Primary Diagnosis	Demographic and Miscellaneous Factors	Personality Factors	Comorbidities	Social Factors	Other Factors
Bipolar	Male	Borderline	Substance abuse	Divorced	Means available
Schizophrenia	Older age	Narcissistic	Panic disorder	Widower	History of child abuse
Major depressive episode	White race	Antisocial	Anxiety	Lives alone	Few reasons to live
Dysthymia	Homosexuality	Conduct disorder	Axis III diagnoses	Isolated	Lots of adverse events
Adjustment disorder with depression	History of attempt	Impulsive		Money worries	Change grades
Conduct disorder	Family history			Other losses	Change friends
Psychosis	Suicidal ideas			No religion	Giving things away
	Hopeless Helpless				Guns in the home

Helping Significant Others When a Suicide Occurs. Foremost among "others" is the decedent's family. Assuming that they desire help, they may be referred to a local "SOS" (Survivors of Suicide) group near them. Both the American Foundation for Suicide Prevention (AFSP—headquartered in New York) and the American Psychological Foundation can provide a geographic directory for such groups. SOS groups are *not* designed for those who have attempted suicide but failed to die (although such persons are clearly in need of help, since SOS groups exist to support individuals bereaved by suicide). Some SOS groups are reluctant to have family members attend meetings immediately after their loved one's death

because they have a difficult time managing in groups, not letting others speak, since they are too aggrieved to contain themselves. Meetings are generally conducted by other survivors, rather than mental health professionals, on the principle that they "know" and empathize with what the families are suffering better than do the professionals. This certainly does not mean that the treating professional's role ceases following a patient death. Attending the funeral, meeting with the family to offer condolences and to discuss with them whatever is possible within the bounds of confidentiality is generally perceived by grieving families as helpful. Following a patient's suicide, many mental health professionals are reluctant to have any contact with the decedent's family—often because they feel so guilty or are so frightened of litigation that they wish to terminate all contact. However, such contact may actually diminish such an outcome and be comforting to the family. The mental health provider is a survivor, as well. In the article, "Client Suicide—What now" the authors note that of 30,000 suicides per year, 10,000 were in treatment during the preceding year and 6,000 were in treatment in the preceding month.

Because having one's patient suicide is so traumatic for therapists, they often seek help for themselves. Institutional review boards which routinely conduct suicide reviews, postmortems, or "root-cause" reviews are generally designed with the institution's needs in mind, not necessarily the therapist's, so that therapy and support for the involved provider can be extremely helpful.

"Significant others" also include classmates, coworkers, and friends of the victim and mental health professionals may become involved in assisting and supporting them, as well. There are a number of programs to assist schools in handling some of the emotional distress produced by a schoolmate's suicide. The AFSP has printed guidelines available to schools at no cost, as have many other organizations. These suggestions include the recommendation that schools should avoid permanent memorials to decedents, for instance, and other similar steps to avoid "contagion effects" resulting in additional suicides. Similarly, other organizations, including the University of Pennsylvania's School of Journalism, and others have published guidelines for the media to help them present suicide information in ways which do not engender "copycat" suicides.

The Neurobiology of Suicide

There is a large and exciting literature on underlying neurobiological aspects of suicide. For instance, Oquendo et al.'s 2014 article "Toward a Biosignature for Suicide," includes a review and synthesis of the research literature of the biomarkers of suicide in order to develop a coherent model for the biological underpinnings for suicidal behavior. The authors note that several lines of evidence indicate that stress response systems in suicidal individuals become dysregulated, especially the HPA (cortical–

hypothalamic–pituitary axis), constituting a diathesis for suicide. Additional findings, such as neuroinflammatory indices, glutamatergic function, and neuronal plasticity may be a reflection of the downstream markers of such dysregulation. They suggest that studying these systems in the same populations may reveal the role and interplay of each. This could lead to new treatment targets and biological predictors. Guintivano's group has identified a combined genetic and epigenetic biomarker, SKA2, in the prefrontal cortical tissue obtained from the Stanley Medical Research Institute and the Harvard Brain Bank at McLean Hospital. This group subsequently replicated their results with peripheral blood from three living groups. They found that SKA2 gene expression was significantly lower in suicide decedents and was significantly associated with genetic and epigenetic variation in the rs7208505 region of the SKA2 gene. They suggest that the SKA2 genetic and epigenetic variation represents an underlying state which increases suicide risk in the presence of a stressor. Although these preliminary findings need to be replicated by others, they represent at least one aspect of the promise of neuroscientific inquiry.

The Suicide Inquiry

It cannot be stressed enough that every psychiatric evaluation must include a documented suicide inquiry. If there is no indication of a prior suicide ideation or attempt or depression this may be brief. Be certain you are not inadvertently signaling that you're looking for a negative answer, as in "you're not suicidal, are you?" Better to start with "have you ever wished you were dead?" or "ever thought seriously about suicide?" The inquiry proceeds stepwise until the physician feels reassured that the patient is in no need for more restrictive treatment. If the patient is psychotic and/or depressed, one needs to expand the inquiry. "How depressed are you" or "does it ever get so bad that you wished you were dead?" or "that you ever thought about suicide?" "Have you ever felt suicidal at another time in your life?" "Have you ever attempted it?" Any complete assessment must include "any family history of attempts or completions?" If a patient responds affirmatively to "have you now or ever felt you wished you were dead" it is often helpful to ask "are you concerned that you might harm yourself?" If the patient says "no" you then need to wonder if he or she is only saying "no," because he thinks that's what you want to hear. If he says "yes, I am worried," you should be quite concerned. Asking other informants, such as the patient's spouse or other family members accompanying the patient to your office or telephoning the patient's spouse should you be suspicious that the patient may not have been forthcoming with you, is also helpful (assuming that the patient agrees). Note that if you are very concerned that the patient is minimizing his suicidality, you may tell the patient that you need to speak with his or her spouse even without consent because you are so concerned. Your assessment about the acuity of the situation determines

what you do next. If you feel that the patient needs to be admitted directly from your office, you or a responsible adult (e.g., the patient's spouse, a nurse or aid, a parent, or other responsible adult) needs to be constantly with the patient until he or she is transported (and presumably admitted) to a psychiatric facility which is capable of caring for such patients. "Constantly" includes use of the bathroom, a common site for attempting or completing suicide. The most privacy one should allow is that the bathroom door be left ajar. Now, in an era of hospitalists generally being the ones to care for your private patients, your decision to seek hospitalization may not necessarily suffice. What if the hospital staff feels that admission isn't required and that the least-restrictive environment rule applies and says your patient may leave? If you agree, fine. If you remain convinced that this is too unsafe, you can try to convince the staff of the need. If that fails, you may try to arrange admission elsewhere, or seek a commitment, assuming the family agrees with you. These may be painful decisions which complicate your busy schedule but life, injury, or death may hang in the balance and "you're it" and have to do what the prudent physician would do—meaning, doing your best to insure the patient's safety. For those physicians who believe that such efforts on behalf of their patients is infantilizing and paternalistic and, therefore, prefer to be laissez-faire in such matters, if the patient goes on to harm or kill himself, you may be held responsible, nonetheless, should litigation arise (see section on Suicide and Malpractice).

There are numerous "suicide risk scales" in use. One of the best of these is the Columbia Suicide Severity Risk Scale (C-SSRS), which was determined to be reliable and valid in three multisite studies. An excellent recent review article is "Training for Suicide Risk Assessment and Suicide Risk Formulation" in which the authors describe two components of the evaluation: assessing the risk factors and then formulating the level of risk. Despite the general belief that active suicide ideation is more predictive of an event occurring than passive, they do not believe that this is so, nor do they see that "active" is more predictive of death by suicide than "passive" ideation. They agree with A.T. Beck's contention that learning about suicide ideation at its worst point is a better predictor of death by suicide than current suicide ideation or hopelessness. They feel that "warning signs" primarily relate to suicide over the next 12 months, not necessarily "near" events, and cite other work showing that, in their last contact with suicidal patients, therapists estimated "no risk" in 30 percent and "low risk" in 54 percent of patients who killed themselves before their next visit. Further, they cite Spangler's study of 4,600+ clinicians and conclude that education and clinical experience results in only 13 percent accuracy of their clinical judgments! "Competence can only develop with extensive deliberate practice and corrective feedback." They (wryly) note that all clinicians feel they are competent judges, nonetheless.

Suicide and Malpractice

What standard of care applies to patient suicide? The usual cliché is that "the psychiatrist isn't expected to predict all suicides," only to properly appraise the "risk" of it and to determine how "foreseeable" the risk is. Although correct in theory, it is more accurate to conclude, "it all depends on the judge and the jury." Most psychiatrists will have at least one suicide in their careers. The great majority of these do not proceed to litigation. Psychiatry residents, unfortunately, not uncommonly have patients who suicide. Then, if litigation ensues, it is generally against the institution and the attending physician as well as the resident. The resident is often called as a witness but not held financially accountable. This is of small comfort to the resident because he has less capacity for perspective to fall back on and less confidence—hence, the critical need for support programs for such residents, lawsuit or not. How does the prudent psychiatrist keep his risk at a minimum? The medical record will generally be decisive in court proceedings so it behooves clinicians to absolutely ensure that the record accurately reflects the evaluation, diagnosis, the suicide inquiry, estimation of risk, treatment plan, course, results, etc. Any contact with the patients' family to corroborate the diagnosis and explain the treatment should be documented. The inquiry about available means for suicide, for example, guns in the home, potentially lethal medications, etc., should be detailed. If there is reason to believe that the patient might be suicidal, was he asked about it at every visit and was this documented in the chart? Did the therapist ask about past history and family history of suicide? Did he ensure that the doses and types of medicine prescribed were reasonably safe, effective, etc.? Were the number of pills prescribed and their doses excessive or insufficient for the situation? A good relationship between the provider and the patient's family may be the most critical factor in terms of avoiding lawsuits after a suicidal death. Importantly, even the best care may not protect against litigation ensuing and then it will all be up to the defense attorney, the plaintiff's attorney, credibility of the plaintiff's and defendant's "expert witnesses," and, ultimately, the judge and jury. Note, too, that merely because a judgment for malpractice has been obtained does not mean one is an inadequate psychiatrist!

Therapy

Introduction. The most effective treatment for preventing suicide is undoubtedly combined treatment—that is, both psychotherapy and pharmacotherapy targeting the underlying psychiatric disorder while working specifically on the suicidal behavior. As stated in the Introduction section, however, there are some approaches, both psychological and pharmacological, which may transcend the "underlying disorder concept," for example, CBT for a wide array of apparently discrete mental disorders;

major and minor tranquilizers for anxiety and agitation across diverse disorders, etc. Nonetheless, for clarity and simplification, medications and psychotherapy are discussed separately here. A major, critical, problem is to find effective treatments for suicide outcome studies is that "suicidal" patients are generally avoided as research subjects because of ethical and medicolegal concerns re litigation should they die or seriously injure themselves while so enrolled.

Role of the Hospital. When the patient appears to be so suicidal, so unreliable about his intent or untrustworthy regarding his intent, or so out of control that the therapist feels hospitalization is necessary to protect him, then hospitalization is required. Although this sounds simple and straightforward, as discussed earlier, this is a complex situation, even when the patient and his significant other agree. Since the therapist generally is not the decider of who gets admitted, the therapist will need to provide a convincing reason for "why now in?" Assuming that hospital staff agree to keep the patient, they may or may not want your input in helping decide what level of care is required: closed or open unit, single or double room, frequency of observation, etc.?

Presumably, you (the therapist) and the hospitalist will work together on this. Before your patient can be safely discharged, however, virtually the same question must be asked by the attending, namely "why now out?" What has changed to indicate that he is now safe? The mere fact that he has been in the hospital for a few days, in itself, is insufficient evidence. How has his improved outlook been documented? Just because he's taking his medicines and attending groups is not sufficient evidence by itself. What do staff and his family report? Such evidence must inform the discharge plan.

Somatic Treatment. Obviously, somatic treatment has to be tailored to the underlying condition, for example, depression, bipolar disorder (both manic and depressed phases), schizophrenia, substance abuse, etc. However, when selecting an agent for a patient deemed "suicidal," it is preferable to select one with known antisuicide effects such as lithium for bipolar patients or clozapine for psychotic or bipolar patients, assuming that there are no contraindications for such choices. In BPD, the best approach is psychological (e.g., transference-focused therapy, "mentalization-based" behavior therapy, or DBT) rather than pharmacological (unless the BPD is accompanied by a comorbid disorder which responds to medication), since no medications have yet been proven effective for this disorder.

Psychotherapy. There is recent evidence that psychotherapies may specifically target suicide. The best example of this may be DBT for BPD. Although DBT has come to be regarded as the gold standard for treating

BPD, there are a few studies showing that other therapies may be equally valuable in reducing suicidal behaviors. An interesting comparison study was carried out by Clarkin et al., comparing 90 BPD patients treated by experienced clinicians using three modalities: transference-focused psychotherapy, DBT, or supportive treatment (along with medication, as indicated) for 1 year. Patients in all three groups showed a significant positive change in depression, anxiety, global functioning, and social adjustment. Both DBT and transference-focused therapies were significantly associated with improvement in suicidality. Only transference-focused and supportive therapies were associated with improvement in anger and impulsivity. Only transference-focused was significantly predictive of change in irritability, and verbal and direct assault. CBT is another highly valued treatment for suicidal behaviors. A 2012 article in *The Behavior Therapist,* indicates that suicidal behavior should be viewed as the primary problem rather than merely viewing this behavior as a symptom of a psychiatric disorder. The authors target specific cognitions and behaviors which fuel suicidal behavior. They detail an approximately 10-session program and note that, in an 18-month follow-up, 24 percent of the CBT-treated patients versus 42 percent of the treatment as usual (TAU) ones made at least one additional attempt.

Daniel and Goldston, guest editors for a special series, "Working with a Suicidal Client—Not Business as Usual," noted that "until recently, suicide literature focused on risk factors and assessment, and treatment focused on the 'underlying DSM-5 disease' in which suicide is seen as a symptom of another disease such as depression, psychosis, etc., and on prevention." In this same "special series," Linehan, Comtois, and Ward-Cieleski note that the paucity of random controlled trials is a major reason that suicide rates have barely budged in the past 100 years.

In addition to DBT and CBT, there are other therapeutic approaches which appear to be helpful for suicidal patients. One of them is mindfulness. As Luoma points out, experiential avoidance, the tendency to avoid or escape from unwanted psychological experiences, may produce harmful consequences like suicidal behaviors. Mindfulness and psychological acceptance may serve as a means to target experiential avoidance, thereby decreasing the risk of suicide. This approach directs treatment of suicide not solely as a symptom of disease but as the primary target of therapeutic efforts. Mindfulness is a key feature which overlaps the four psychological processes in acceptance and commitment therapy (ACT).

SPECIAL POPULATIONS

Military. For both active duty and, to a lesser extent retired, military personnel have increased rates of attempts and completions. A recent study by Rudd et al., found that eight participants in a brief CBT treatment (13.8

percent) versus 18 subjects (40.2 percent) treated with TAU made at least one suicide attempt. Thus, CBT subjects were 60 percent less likely to attempt than TAU ones. Subjects were drawn from 76 subjects who had either attempted suicide or had suicide ideation with intent to act and were treated with TAU versus 76 attempters or serious ideators treated with CBT + TAU.

Adolescents. Depression in adolescents is less likely to manifest the same symptoms as in adults. Verbal or even physical aggression is just as apt to be seen as sadness, tearfulness, etc. Giving prized possessions away (books, clothes, CDs, etc.) or changes in behavior (becoming unruly, uncooperative, abruptly starting to use drugs or alcohol) may be all that the parents note as changed. Similarly, a drop in grades in a student who has always done well may be all that is observed by the parents. More "moodiness" or more sleep irregularity are common symptoms but not as common in adolescence as to be poor markers for depression. School-wide suicide screening examinations, such as the Columbia Teen Screen are now widely used in junior and senior high schools in this country to help identify those students at increased risk for depression and suicide.

The Elderly. The geriatric population has the highest ratio of completion to attempts. That is, when they attempt suicide, it is more frequently a lethal attempt than the methods seen in younger populations. They are also harder to predict, partly because their depressions are more likely to be misperceived as "normal" ("of course he's depressed, he's old, feeling useless and tired, so it's just natural"). The suicide rates for the elderly have been the highest of any age group for approximately the past 250 years but, beginning around the year 2000, rates in 40 to 65 year olds rose enough to surpass the 65+ year olds.

The Publicly Shamed. Defrocked clergy, disbarred attorneys, de-licensed physicians, and other "prominent and respected" individuals, who have became publicly shamed due to one scandal or another (e.g., undue familiarity; child molestation; financial "mischief," etc.) are anecdotally reported to have suicide rates within the year following their "outing" higher than any other subgroup.

Victim-Precipitated Suicide. Also known as suicide-by-cop, is a particularly cruel method of killing oneself since the pain it engenders is not only to the decedent's friends and family but also to a totally innocent person. The unfortunate policeman who is confronted by a stranger who appears to be homicidal and who shoots and kills that stranger thinking his own life or some third party is in danger, ends up becoming a victim himself. The legacy of suicide is terrible enough without making someone else responsible. Jumping to one's death in front of a passing train or car leaves a similar painful legacy for innocent second parties.

Physician-Assisted Suicide. Already first legal in Oregon and now a few other states, the right for physicians to prescribe (and administer) lethal doses of medications to terminally ill and suffering patients, will undoubtedly expand. Although this appears humane, the potential for harm is real and needs to be guarded against, such as families pressuring infirm and older relatives to request an end-of-life to serve their own purposes rather than the patient's, as Hendin warned against in Seduced by Death. An additional related concern is that depression is common in chronically medically ill patients and may result in their pressing of their families to help them die, despite the fact that most such depressions are treatable.

Future Directions. What can one hope for over the next 25 to 50 years? This writer's wish list vis-à-vis suicide follows. Suicide attempt and completion rates will decrease in concert with diminished gun ownership. The likelihood of that happening is none to minimal, however. What is more probable is a vastly improved array of medications to rapidly, effectively, and safely treat (a) depressive and bipolar disorders, (b) substance abuse, (c) schizophrenia, and (d) personality disorders, especially BPDs. An apparent taste of this was provided when near-immediate lifting of depression appeared to arise when Ketamine was administered and was reported to be a wonder drug. It wasn't but that hardly rules out another drug with super-rapid action coming along. Better pharmacological therapies would result in basic science discoveries re depression, mania, and psychoses. Also, improved diagnostic tools will help ensure optimal and prompt diagnoses. In addition to a vastly improved pharmacology, would go a comparably improved psychotherapy. CBT, DBT, and psychoanalytically oriented psychotherapy, as well as other psychotherapies, need to be mastered by all therapists and shortened by discerning what parts of these are most relevant to suicidal individuals.

Acknowledgment

The author wishes to express his appreciation to Dr. Donna Sudak for her invaluable bibliographic and editorial assistance with this chapter.

References

Beck AT, Brown GK, Steer RA, Dahlsgaard KK, Grisham JR. Suicide ideation at its worst point: a predictor of eventual suicide in psychiatric outpatients. *Suicide Life Threat Behav*. 1999;29:1–9.

*Brent D. In search of endophenotypes for suicidal behavior. *Am J Psychiatry*. 2009;166(10):1087–1089.

Brent D. Nonsuicidal self-injury as a predictor of suicidal behavior in depressed adolescents. *Am J Psychiatry*. 2011;168(5):452–454.

Brent DA, Emslie GJ, Clarke GN, et al. Predictors of spontaneous and systematically assessed

suicidal adverse events in the treatment of SSRI-resistant depression in adolescents (TORDIA) study. *Am J Psychiatry*. 2009;166:418–426.

Brent DA, Melhem NM, Oquendo M, et al. Familial pathways to early-onset suicide attempt: a 5.6-year prospective study. *JAMA Psychiatry*. 2015;72(2):160–168.

*Clarkin JF, Levy KN, Lenzenweger MF, Kernberg OF. Evaluating three treatments for borderline personality disorder: a multiwave study. *Am J Psychiatry*. 2007;164(6):922–928.

Daniel SS, Goldston DB. Hopelessness and lack of connectedness to others as risk factors for suicidal behavior across the lifespan: implications for cognitive-behavioral treatment. *Cogn Behav Pract*. 2012;19:288–300.

Ellis TE, Patel AB. Client suicide: what now? *Cognitive and Behavioral Practice*. 2012;19:277–287.

Guintivano J, Brown T, Newcomer A, et al. Identification and replication of a combined epigenetic and genetic biomarker predicting suicide and suicidal behaviors. *Am J Psychiatry*. 2014;171:1287–1296.

Kelleher I, Lynch F, Harley M, et al. Psychotic symptoms in adolescence index risk for suicidal behavior: findings from 2 population-based case-control clinical interview studies. *Arch Gen Psychiatry*. 2012;69:1277–1283.

Linehan MM, Comtois KA, Ward-Ciesielski EF. Assessing and managing risk with suicidal individuals. *Cognitive and Behavioral Practice*. 2012;19:218–232.

Luoma JB, Villatte JL. Mindfulness in the treatment of suicidal individuals. *Cogn Behav Pract*. 2012;19:265–276.

*Oquendo MA, Sullivan GM, Sudol K, et al. Toward a biosignature for suicide. *Am J Psychiatry*. 2014;171:1259–1277.

Posner K, Brown GK, Stanley B, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry*. 2011;168:1266–1277.

*Silverman MM, Berman AL. Training for suicide risk assessment and suicide risk formulation. *Acad Psychiatry*. 2014;38:526–537.

Simonetti JA, Mackelprang JL, Rowhani-Rahbar A, Zatzick D, Rivara FP. Psychiatric comorbidity, suicidality, and in-home firearm access among a nationally representative sample of adolescents. *JAMA Psychiatry*. 2015;72(2):152–159.

Stanley B, Brown GK. Safety planning intervention: a brief intervention to mitigate suicide risk. *Cognitive and Behavioral Practice*. 2012;19:256–264.

Wenzel A, Berchick ER, Tenhave T, Halberstadt S, Brown GK, Beck AT. Predictors of suicide relative to other deaths in patients with suicide attempts and suicide ideation: a 30-year prospective study. *J Affect Disord*. 2011;132:375–382.

Wenzel A, Jager-Hyman S. Cognitive therapy for suicidal patients: current status. *Behav Ther (N Y N Y)*. 2012;35(7):121–130.

▲ 32.3 Other Psychiatric Emergencies

DAVID A. BARON, D.O., ROBERT T. COBB, M.D., AND GRACE MITU JUAREZ, M.D.

INTRODUCTION

The previous edition of this chapter highlighted the important role of the Community Mental Health Centers Act of 1963. It established Emergency Psychiatry as one of the five essential services mandated for all federally funded community mental health centers nationwide. This legislation ultimately helped establish Emergency Psychiatry as a subspecialty, through the increased recognition of its importance to the health care delivery system. More recent legislation has also had an impact on the diagnosis and treatment of psychiatric emergencies. The Affordable Health Care Act has enabled millions of previously uninsured Americans to obtain medical insurance benefits. The effect this will have on utilization of psychiatric emergency services (PES) is yet to be determined. With better benefits, will patients require fewer emergency visits resulting from better routine care and earlier initial assessment? Or, will having insurance benefits with expanded options for care delivery increase patient volumes in psychiatric emergencies centers? These are questions future versions of this chapter will need to address. This chapter will provide a historical background of the field and an update changes with DSM-5 as it pertains to PES. Recent advances in pharmacotherapy and an update on contemporary drugs of abuse in 2015 seen in the PES also are highlighted. This edition of the chapter includes updated references and new subject material, in addition to the most relevant information from the previous version.

HISTORY

Emergency psychiatry originated in the Russo-Japanese War (1904–1905). Psychiatrists in the Imperial Russian Army championed treating psychiatric emergencies at the front lines. Healthy components of personality were strengthened, esprit de corps was encouraged, and soldiers returned to action as soon as they were stabilized and reconstituted. This countered effects of stigmatization, secondary gain, and avoidable iatrogenic effects of longer-term military therapies that entail removal of soldiers from the front lines. With this approach, the Italians found that 60 percent of posttraumatic stress disorders remitted. American military psychiatrists did not at first adopt this technique; only 18 percent of US psychiatric casualties returned to action in 1943. A year later, when crisis-oriented therapy was introduced using abreaction, ventilation, and social support, as many as 85 percent returned to fighting. It is apparent that brief crisis-oriented approaches at times of emotional crisis can not only restore functioning but also reinforce and develop healthy coping styles that protect against future stresses. Crisis intervention approaches have been incorporated into psychiatric practice and extended to the civilian populations over the years and serve as an important arm of mental health care services available to patient populations.

Erich Lindemann studied patterns of adaptation to the catastrophic deaths of more than 500 people in the Coconut Grove nightclub fire in Boston in 1941 and found that mourning is necessary to adjust to loss. Lindemann's studies of grief reactions in the 1940s were followed by other clinicians in the early 1950s who conceptualized a crisis as a time-limited event that disrupted a patient's psychological equilibrium. Rapid psychotherapeutic intervention could facilitate a return to a psychological homeostasis and help the individual to make changes in his or her life and reestablish a psychological sense of well-being. Rapid intervention strategies were termed crisis intervention and referred to a variety of psychotherapeutic techniques for treating patients who are confronted with stressful or overwhelming psychological events.

Several major forces converged in the 1950s and 1960s that gave rise to the field of emergency psychiatry as it is known today. With the beginning of deinstitutionalization in the mid-1950s, and the emergence of emergency medicine as a medical specialty, hospital emergency departments began to see an increase in the number of chronically ill psychiatric patients who were being admitted for assessment and treatment. These patients were generally seen in the medical emergency department by a psychiatrist who acted as a consultant and with recommendation for either referral to an outpatient clinic or a psychiatric hospitalization. With the passage of the Community Mental Health Centers Act of 1963, which mandated 24-hour emergency services, the PES was created as a specialized program with its own staff and designated clinical area. The PES evaluated and treated patients suffering from acute psychopathology with the goal of providing treatment at the least restrictive level of care.

A psychiatrist working in a PES must have an in-depth knowledge of psychiatry, substance-related illnesses, medicine, neurology, and the available community resources. Many clinicians consider emergency psychiatry a subspecialty of psychiatry. The American Association for Emergency Psychiatry was founded in the 1980s for psychiatrists with a special interest in emergency psychiatry. Other programs for providing emergency psychiatric care have since evolved, including crisis hospitalization programs, crisis residences, and mobile crisis teams.

The priorities and goals of an emergency psychiatric assessment are as follows:

- To control aggressive behavior in order to protect other patients, the PES staff, and the patient from him- or herself.
- To rule out any etiology for a patient's behavior that might be life threatening or may increase medical morbidity.
- To facilitate a thorough clinical evaluation and an appropriate longitudinal plan.

First Priority

The first priority is to ensure that the PES environment is safe for the clinical evaluation. Clinical personnel should be trained in the management of aggressive behavior and have a working understanding of their clinical responsibilities and institutional policies and procedures. Various formal training programs have been designed to ensure that providers are vetted on management strategies. These programs serve to mitigate risk of injury to patients and clinical personnel and are essential to providing a secure and safe environment for patient care.

Certain environmental variables should be modified to decrease the potential for violence. These include, but are not limited to, shortening waiting times, offering the patient drink or food, and decreasing environmental stimuli. Providing the patient a comfortable chair in a quiet area or the opportunity to lie down is also effective in mitigating risk of worsening agitation. Early identification of potential dangerous pathology and early intervention can markedly decrease negative outcomes due to behavioral escalation, and a skilled practitioner should be able to intervene safely and appropriately.

It is important also to understand that the physical structure of the PES is also crucial to facilitating a safe environment for patient care. Having shatterproof windows in the doors of interview rooms provides privacy but allows an interview to be observed. In addition, security cameras can allow for observation of clinical providers and ensure another layer of safety. Cameras can also provide the patient with a sense that his or her behavior is being monitored. Posting rules visible to patients within the PES can be useful in making it clear that violence will not be tolerated and has consequences is another useful strategy.

Safety and security can be further enhanced by

1. weapons screening prior to entry of the secured area;
2. rooms in which the examiner cannot easily be trapped;
3. a choice of open or enclosed interviewing areas;
4. an alarm system to call for help, preferably a panic button in every room; and
5. adequate personnel to respond if help is needed, including trained security personnel.

If a person is brought in handcuffed by the police with a history of a very recent assault, then it may be necessary to examine the patient in handcuffs or possibly in physical restraints if the agitation or threats of violence persist. If the patient is acutely agitated and threatening, then it may be necessary to administer medication to help him or her regain control before beginning an interview. Interventions for calming the agitated and aggressive patient are discussed later in this section.

Second Priority

The second priority in managing emergent psychiatric pathology is to rule out a medical etiology for the patient's behavior. The incidence of patients presenting with psychiatric illness who have a medical etiology for their symptoms varies from 15 to 90 percent. Medical examinations of psychiatric patients in the emergency department are often limited in scope, and psychiatrists and emergency department physicians may have different viewpoints on the utility of laboratory screening. Therefore, psychiatrists must maintain a high index of suspicion for underlying medical problems and may need to initiate further studies than those that are recommended by emergency medicine guidelines. Clinical history, signs, and symptoms that are suggestive of a medical etiology include

- patients older than 40 years or younger than 12 years with no previous psychiatric history;
- acute onset (hours to weeks);
- fluctuating course;
- clouding of consciousness;
- visual or olfactory hallucinations;
- abnormal vital signs;
- disorientation;
- known medical illness or neurological symptoms;
- memory impairment;
- complex medication regimen;
- a history of alcohol or drug use.

Patients with dementia are typically brought to the PES as a result of wandering away from home, poor nutrition, psychosis, violent outbursts, or lack of attention to everyday safety hazards. In such patients, it is imperative that the clinician rule out an underlying medical or neurological etiology for the behavior. Elderly patients presenting to psychiatric emergency departments, especially those with diagnosed dementia, are at high risk for behavioral disturbances. They often have complex medication regimens, are high risk for urinary tract infections, electrolyte disturbances, and sequela from cerebrovascular insult (e.g., stroke).

Acquired immune deficiency syndrome (AIDS)-related psychiatric emergencies have seen a recent decrease in incidence. This is in large part due to more effective management of human immunodeficiency virus (HIV) seropositive patients with highly active anti-retroviral treatment strategies. However, the associated medical complications, a comorbid psychiatric disorder, or HIV/AIDS-related neuropsychiatric symptoms may present as a primary psychiatric complaint. Psychiatric disorders can often be the first manifestation of HIV infection, and many patients with HIV have psychiatric and/or substance use disorders and are labeled as "triply

diagnosed patients." Obtaining a thorough history and/or appropriate serology testing is important in patients at risk for HIV.

When medical conditions are suspected as the cause of the patient's symptoms, a physical examination and an appropriate medical workup should be completed, if possible. Initial diagnostic studies should include a complete blood count with differential, a complete metabolic panel (including electrolytes, blood urea nitrogen, creatinine, liver function studies, glucose), an electrocardiogram, urine drug screen, urinalysis, blood alcohol level, and neuroimaging when indicated. Medication blood levels should be obtained if applicable, for example, valproic acid, lithium, or carbamazepine. An ammonia level should be obtained for patients taking valproic acid. If the initial diagnostic studies are negative, but the clinician still suspects an underlying medical etiology, then additional tests can be completed once the patient is admitted to the hospital. Various medical illnesses that present with psychiatric signs and symptoms are listed in Tables 32.3–1 through 32.3–4.

Third Priority

Facilitating appropriate treatment and disposition through a comprehensive psychiatric evaluation is the third priority. The patient interview should always be regarded as a therapeutic opportunity. Patients in the PES are frequently in distress and should be given the opportunity to express their concerns and ventilate their affect. An open-ended question will usually elicit the primary reason for the visit and potentially aid in rapport development. Even if the staff has already obtained an initial history, it is still important to let the patient talk for a few minutes without interruption. Closed-ended questions should be reserved for the end of the interview or for those patients who are unable to adequately provide a meaningful and coherent history. Reassurance should be given frequently, and long silences should be avoided during the interview.

Table 32.3–1.
Differential Diagnosis of Anxiety

Medical
AIDS/HIV
Cerebral arteriosclerosis
Encephalitis
Epilepsy
Essential hypertension
Hyperthyroidism
Hyperventilation syndrome
Hypocalcemia
Hypoglycemia
Hypokalemia
Impending myocardial infarction

- Internal hemorrhage
- Mitral valve prolapse
- Paroxysmal atrial tachycardia and other cardiac arrhythmias
- Pheochromocytoma
- Postconcussion syndrome
- Pulmonary embolism
- Subacute bacterial endocarditis
- Temporal lobe disorder

Substance induced
- Alcohol delirium and withdrawal
- Stimulant withdrawal disorders
- Caffeine intoxication
- Stimulant intoxication
- Sedative, hypnotic, or anxiolytic withdrawal and withdrawal delirium

Psychiatric
- Anxiety disorders
- Bipolar disorder
- Borderline personality disorder
- Chronic schizophrenia
- Major depressive disorder
- Generalized anxiety disorder
- Panic disorder
- Phobias
- Sexual disorders

Table 32.3–2. Differential Diagnosis of Depression

Medical
- AIDS/HIV
- Antihypertensive toxicity
- Cerebral neoplasia
- Cirrhosis of the liver
- Dementia of the Alzheimer type
- Hepatitis
- Hypokalemia
- Hyperthyroidism
- Hypothyroidism
- Infectious mononucleosis
- Occult malignancy
- Pancreatic carcinoma
- Postviral infection syndrome
- Steroid psychosis
- Vascular dementia

Substance induced
- Stimulant withdrawal

Psychiatric
- Adjustment disorder with depressed mood
- Bipolar disorder
- Borderline personality disorder
- Brief psychotic disorder
- Chronic schizophrenia
- Cyclothymic disorder
- Persistent depressive disorder

Major depression
Schizoaffective disorder
Schizotypal personality disorder

Table 32.3–3.
Differential Diagnosis of Mania

Medical
AIDS/HIV
Antidepressant-induced mania
Amphetamine-induced mania
Bronchodilator-induced mania
Decongestant-induced mania
Delirium
Hyperthyroidism
L-Dopa–induced mania
Postencephalitic syndrome
Steroid-induced mania
Substance induced
Alcohol intoxication
Amphetamine-induced mania
Cocaine-induced mania
Phencyclidine-induced mania
Psychiatric
Atypical psychosis
Bipolar disorder
Catatonic schizophrenia
Schizoaffective disorder

The emergency evaluation varies in length and may on occasion exceed several hours. Ideally, the evaluation should include

- identifying information: age, sex, ethnicity, marital status, insurance status, and source of referral;
- chief complaint;
- history of present illness;
- past psychiatric history;
- substance use history;
- medical history;
- social history;
- educational, occupational, and military history;
- legal history;
- family history;
- review of systems, if indicated;
- physical examination, if indicated (breast, genital, pelvic, or rectal examinations should not be done); and
- mental status examination including a cognitive screen.

Table 32.3–4.
Differential Diagnosis of Thought Disorder

Medical
AIDS/HIV
Dementia due to Pick disease
Dementia of the Alzheimer type
Endocrine disease
Frontal lobe neoplasm
Migraine equivalent
Pernicious anemia
Steroid psychoses
Syphilis
Temporal lobe epilepsy
Vascular dementia
Substance induced
Alcoholic hallucinosis
Alcohol intoxication
Substance-induced psychotic disorder (e.g., PCP or amphetamines)
Psychiatric
Adjustment disorders
Atypical psychosis
Bipolar disorder
Brief reactive psychosis
Delusional disorder
Dissociative disorders
Major depression
Schizophrenia
Schizophreniform disorder

It is essential that suicidal and homicidal ideations are assessed thoroughly. A structured suicide and violence risk assessment should be performed and documented during the course of examination. The most common liability in the PES is failure to hospitalize a patient who requires admission for risk of violence to self and/or others. If a decision is made to allow for discharge from the PES and not hospitalize, a thorough record of the clinical determination should be documented. This includes protective factors, risk factors, and the overall risk stratification of the patient for potential for harm to self or others. While there is not a “gold standard” for documenting the clinical decision, it is important to have documentation reflect one’s decision should an unfortunate adverse event occur.

Obtaining a comprehensive history in a PES can be challenging as patients who are psychotic or agitated may provide unreliable or inconsistent histories. This is especially true, if the patient is mute, cognitively impaired, or agitated. In these situations, it becomes even more important to utilize collateral sources of information including significant others, family, therapists, neighbors, police, or paramedics. Managed care companies may have significant clinical information and provided a complete history of the patient. Previous hospital records, if available,

should be reviewed and important information should be summarized in the final assessment.

PRINCIPLES OF TREATMENT

Overview

Treatment in the PES usually focuses on behavioral and symptom management. Patients often have multiple confounding clinical problems. These include psychosis, depression, substance intoxication or withdrawal, extreme psychosocial stresses, and overlapping character pathology. Furthermore, psychosis and depression have an extensive differential (Tables 32.3–2 through 32.3–4). Given the complexity of patients presenting for care, making a definitive diagnosis may be difficult in the limited amount of time available to the PES clinician. Beginning diagnosis-specific treatment may be problematic unless continuity of care can be coordinated and assured with the next provider. The focus of the PES intervention should be acute stabilization and safety with subsequent longitudinal planning based on outcomes of initial interventions.

Psychotherapeutic Management of Psychiatric Emergencies

Treatment should ideally be interwoven with the evaluation process. The initial interview not only serves to elicit important diagnostic information but is also therapeutic if approached correctly. The goal of the interview is to gather information about the present illness, past psychiatric and medical history, family history of mental illness, social history, and drug or alcohol use but can serve as an important vehicle in establishing rapport and ensuring a patient is aware of his or her specific treatment plan.

Ideally, the interview should take place in a quiet, comfortable setting with both the patient and the clinician seated. However, often the intervention may take place in a less than ideal setting. In these instances, at minimum the clinician should have an unimpeded exit in case of patient behavioral escalation. If the patient is seen after the medical evaluation, then it is important to remove unnecessary intravenous needles and ask the patient to sit up on the stretcher. This helps restore the patient's autonomy and improve the physician–patient alliance. The physician should always introduce him- or herself and address the patient as Mr. or Ms. X as this helps restore the patient's sense of dignity.

The interview should begin with nonspecific, less intrusive questions. Only after the patient begins to show some comfort with the interviewer should the physician start questioning him or her about the specific details of the present illness. Questions should be open-ended. The physician should be flexible in his or her approach to the interview as each patient presents a unique set of circumstances, behavioral triggers, stressors, and life experiences that will influence the diagnostic evaluation.

When interviewing an acutely disturbed patient, it is frequently necessary to structure the interview by asking straightforward questions. If the patient begins to ramble, then the interviewer should indicate that he or she understands what the patient is saying, while helping him or her to be redirected in providing important content. When patients have disordered perceptions, it is important not to attempt to correct the misperceptions but to clarify how the patient experiences them. The clinician should avoid the temptation to use logic to convince the patient that he or she is wrong. Such an approach makes the patient more defensive and can lead to escalation with potential for dangerousness. The interview should be geared toward fact gathering and establishing a baseline understanding of psychopathologic-related impairments.

Establishing emotional contact with the patient to help identify the patient's feelings will help build the treatment alliance. Even a psychotic patient's feelings are not unique, and the interviewer should be capable of identifying and empathizing with them. A patient's need to protect his or her sense of integrity and self-control may initially make him or her reluctant to engage in the interview. If patients are resistant to the idea of talking, then it is useful to elicit their feelings about the situation and try to understand the patient's discomfort with being in a PES and not knowing what may happen to him or her. Sometimes it may be better to drop a particular subject and come back to it later in the interview, when the patient is more at ease. It is of great importance to understand that the patient is experiencing a significant loss of autonomy in requiring an assessment in the psychiatric emergency room setting. Individuals react differently to their circumstances on the basis of the level of impairment and previous life experiences. Impulse control is often compromised and is often a reflection of the normative human response to stress. The innate "fight or flight" response based in sympathetic nervous system drive can be elevated in PES patients, especially those who do not want assistance with their underlying psychopathology. The goal of the interview should be to develop rapport with the patient and aid in his or her understanding of the unique set of circumstances that have led to the clinical intervention. It is also important to understand that the clinician may be in a position of perceived power by the patient. As a result, the goal should be to have the patient understand that you are an advocate, and there to help aid in the recovery process.

When psychotic patients do not respond to questioning, the interviewer should use whatever data are available to make a meaningful connection with the patient. This may include the patient's words, expressions, appearance, or behavior, as well as a physician's feelings. Comments should be as specific as possible in order to provide some structure to an interview that has little to no momentum—"I don't know what Jesus says to us" or "I see you are in a bathrobe. I gather you were brought to the

hospital unexpectedly." By attempting to develop rapport in understanding the patient's experiences, a clinician can have significant impacts in developing a therapeutic alliance and improve patient outcomes.

Of all patients, the most difficult to evaluate is the hostile, paranoid patient. If a patient is frightening, then the physician should not ignore his or her own fear. Even the most impaired patients can quickly sense a physician's discomfort and may become frightened, resulting in a reactive behavior. This may in turn lead to a violent episode. At times, however, the physician may have to acknowledge his or her fears to the patient, "The way you are looking at me is scary, like you are on the verge of striking out. I won't be able to give you my full attention if I am afraid." If the physician remains professionally confident and does not avert his or her eyes, then the patient is reassured. The patient's anxiety can be further ameliorated if the interview is conducted with the door open, thus allowing for a sense of safety for the patient. If the patient engenders a level of anxiety in the clinician such that it interferes with the evaluation, then the clinician should request another clinician to interview the patient or interview the patient with another staff member present in the interview room. It is often difficult for clinicians to understand their limitations in engaging with patients, but ensuring your safety and the safety of others is paramount in the PES. Termination of a clinical assessment may be necessary and is not usually a function of poor training.

It is important to be extremely tactful with paranoid patients because they are easily humiliated and made to feel guilty for their actions and thoughts. The clinician should never become involved in how the patient's beliefs are distorted. If the physician tries to determine what is real and what is not, then he or she will be experienced as a prosecutor, and the patient may become more argumentative and defensive. By identifying a patient's thought distortions in his or her delusional belief system, especially those that are paranoid in nature, a clinician may become incorporated into the delusional construct and disrupt rapport development.

Frequently, paranoid patients who are angry begin with a tirade of accusations about being mistreated. The physician may have to interrupt these tirades and ask the patient, "How can I help you?" in an attempt to develop a supportive therapeutic alliance. With the hostile, paranoid patient, the physician should always keep a professional distance by not becoming too friendly, avoiding jokes, and avoiding parental reassurance as patients can quickly escalate on the basis of external stimuli or a misperception of circumstances. Occasionally, highly paranoid patients can make the interviewer feel defensive and foolish by twisting the meaning of the interviewer's words, making it impossible to sustain any direction in the interview. Under these circumstances, it is advisable to alter the course of the interview and attempt to process with the patient his or her conduct.

If meaningful contact cannot be made, then the interview should be terminated, reassuring the patient of concern for his or her safety and ensuring that he or she is aware of the clinician's role as advocate and care provider.

When patients are unable or unwilling to provide clinical information, significant others should be contacted (e.g., spouse, relatives, immediate family, or previous treating physicians) to obtain collateral information. Whenever possible, consent to contact others should be obtained. Written consent is preferable to verbal consent. In potential life-threatening situations such as suicidal ideation or gestures, homicidal ideation or threats, or an inability to communicate, the emergency takes precedence over confidentiality, and significant others should be contacted even without consent. In these cases, the clinician should carefully document in the medical record the reasons for breaking confidentiality.

An important aspect of the interview is its termination. Although there is considerable pressure on the clinician to work rapidly, the interview should be structured to make the patient feel that he or she has been given sufficient time to present his or her story. Allowing a patient an opportunity to present information independent of the structured interview can often deescalate a tense interaction and the clinician should validate the experience of the patient. The clinician should give a summary of his or her understanding of the patient's problem, followed by recommendations for further care.

Difficult Patients

Difficult patients encompass a wide category of diagnoses. A common theme with these patients is their ability to induce extremely negative reactions in PES staff and clinicians. Most often, patients with Personality Disorder—Trait Specified (PD-TS) or formal Personality Disorder (particularly Borderline Personality Disorder), those who present repeatedly ("repeaters," "frequent flyers," or "regular customers"), and those with addiction-related behaviors who might be demanding, manipulative, and entitled may fit this category. Difficult patient populations often demand immediate attention, have a poor frustration tolerance, and relate to others in a negativistic manner. As a result, clinicians may feel anxiety, ambivalent anger and rage, and even fear. It is essential to remain objective and ensure appropriate professional responses to the patient. The clinician's responses to challenging patient populations can have a significant impact on treatment and outcome.

Because of the wide diagnostic variability with difficult patients, similarly diverse clinical presentations may be seen in the PES. Patients may come voluntarily or may be brought in by police for disruptive or violent behavior. Those with borderline traits or personalities may present with suicidal ideation, suicidal gestures, suicide attempts, homicidal

ideation or gestures (and rarely attempts), drug abuse, brief psychotic episodes, and impulsive behavior, often precipitated by interpersonal conflicts. Patients with antisocial personality disorders often use the emergency service for the purpose of secondary gain, such as obtaining medication, avoiding legal issues, or using the prospect of hospitalization to escape from some psychosocial predicament.

The management of these challenging populations is similar, beginning with a clear understanding of acceptable behavior and firm but supportive limit setting by the clinician. An effective strategy to accomplish this is clearly stating that certain behavior is unacceptable and the reasons why it is unacceptable. The patient should be offered two therapeutic alternatives. For example, *you cannot keep punching the walls, because your behavior is scaring other patients and you could injure yourself. Therefore, you can take some medication or lay down in the quiet room until you feel more in control.* It is important to understand, especially in dealing with patients who are involuntarily being observed, that a loss of autonomy can precipitate untoward behaviors. In providing a patient with alternatives to maladaptive behaviors and allowing him or her to make decisions about his or her own care, a clinician can potentially de-escalate a tense situation. Threatening the patient with punitive consequences will exacerbate the negative behavior and may lead to further acting out behaviors. Medication such as benzodiazepines or antipsychotics may be useful in reducing the patient's agitation. Choice of medication should be determined by the presumed etiology of the patient's behavior. Effective communication between staff members, clinicians, and the patient is essential, as these patients may be manipulative or try to divide the PES staff. The strategies discussed earlier in dealing with agitated and aggressive patients are effective in dealing with difficult patients. These patients may require more time for an evaluation in order to discharge their affect, regain control of their behavior, and realistically discuss their options. An effective interpersonal intervention may avert hospitalization. A primary diagnosis of adjustment disorder, mood disorder, or brief psychosis will show that the decision to involuntarily hospitalize a patient is based on an illness rather than a pattern of maladaptive behavior. Involuntary commitment may be necessary for patients with unremitting homicidal or suicidal ideation to ensure safety.

Pharmacological Management of Psychiatric Emergencies

In the PES, a high priority is given to the pharmacological treatment of agitated patients to reduce the incidence of patient and staff injuries and to reduce the patient's psychological discomfort. Management of agitation and aggression is complex because these nonspecific symptoms can occur in a wide variety of clinical conditions. These may include delirium, dementia, alcohol and drug intoxication or withdrawal, personality

disorders, and psychosis secondary to primary psychotic illnesses. At times, patients are so agitated that they are unable to cooperate with a psychiatric or medical evaluation and are incapable of providing any relevant information.

Pharmacological management may initially serve as the primary therapy or as an adjunct to a verbal intervention. Clinically significant agitation may be defined as abnormal and excessive verbal or physical aggression, purposeless motor behaviors, heightened arousal, and significant disruption of patient's functioning. Typical behaviors of clinically significant agitation include

- explosive and/or unpredictable anger;
- intimidating behavior;
- restlessness, pacing, or excessive movement;
- physical and/or verbal self-abusiveness;
- demeaning or hostile verbal behavior;
- uncooperative or demanding behavior or resistance to care;
- impulsive or impatient behavior; and
- low tolerance for pain or frustration.

The goal of pharmacological intervention is to calm the patient without sedation so that he or she can participate in the evaluation and treatment plan. Whenever possible, the patient should be given the option of the route of administration as this can facilitate his or her sense of having some measure of control. The use of oral liquid or dissolving tablets is the least invasive pharmacological intervention and allows the patient to have a feeling of control and participation in treatment.

The most frequently used medication strategies consist of benzodiazepines, second-generation antipsychotic medication alone or in combination with a benzodiazepine, and haloperidol alone or in combination with a benzodiazepine. The most commonly used benzodiazepine is lorazepam.

Studies of the pharmacological treatment of agitation in a busy PES are methodologically challenging. A treatment algorithm for the management of acute agitation or aggression based on the extant literature is as follows:

- Orally disintegrating or liquid risperidone (2 mg) combined with oral lorazepam (2 mg) or orally disintegrating olanzapine (5 to 10 mg) is the first-line intervention for agitation.
- When oral medication is not appropriate because of the severity of the agitation, intramuscular (IM) lorazepam (2 mg) is recommended for delirium, substance withdrawal, and unknown causes or conditions not associated with psychosis.
- For severe agitation secondary to psychosis, IM ziprasidone (20 mg) that can be supplemented with IM lorazepam (2 mg) is first line to limit

potential for side effects.

- Second-line recommendations for severe agitation secondary to psychosis are IM haloperidol (5 mg) and IM lorazepam (2 mg) or IM olanzapine (5 to 10 mg). Lorazepam should not be used in combination with olanzapine because of the risk of cardiorespiratory depression.

Medication may be given every hour, up to three or four doses every 24 hours, though most patients will respond to a single dose of medication. The most common side effects with antipsychotic medications are dystonic reactions that can be effectively treated with IM benztropine (2 mg). Akathisia may develop in the use of highly potent agents and should be monitored. Management is with PO beta-blockers such as propranolol, and patient management can be optimized with concomitant use of benzodiazepines. The most common side effects with benzodiazepines are sedation and ataxia.

Pharmacological management of agitation in children and adolescents in the PES setting is a complex and difficult task for the psychiatric provider. This is especially concerning as the Food and Drug Administration (FDA) has generally not approved the use of atypical antipsychotics in children and adolescents with the exception of risperidone and aripiprazole in certain diagnoses. However, in emergency situations, a consensus of PES medical directors responded that risperidone or lorazepam is their preferred medications. There are also no outcome studies of or recommendations for treating the agitated pregnant patient. In the absence of safety data, clinicians should use the minimal amount of medication necessary to reduce agitation and aggression in these two groups of patients and work collaboratively with specialty services to ensure consensus care as a treatment team.

Management of Agitation and Aggression

Agitation is a nonspecific symptom that occurs across all medical and psychiatric diagnoses. Untreated agitation can escalate to violence. Managing agitation in the emergency setting is complicated by the need to make clinical decisions on the basis of relatively incomplete psychiatric and medical assessments and histories. Severe agitation may make it impossible to obtain an initial medical history, complete a medical examination, or collect screening laboratory studies. Once agitation subsides and the patient is more cooperative, a medical history and evaluation should be considered as medical illness can precipitate violent and aggressive behavior. The differential diagnosis for violence is summarized in Table 32.3–5.

Interventions for agitation should be conceptualized as hierarchal, beginning with the least restrictive and least intrusive interventions and only using the most restrictive interventions when all other treatment

approaches have failed. The core clinical strategy for managing agitation is the use of interpersonal strategies that emphasize verbal de-escalation interventional techniques. It is important for the clinician to reduce the patient's anxiety and fear by maintaining a humane and respectful disposition. A patient who is treated with honesty, dignity, and respect is likely to respond with reciprocity if there is retained reality testing. Affect management is central to any effective aggression management technique. It involves acknowledging the patient's affect, validating it when appropriate, and encouraging the patient to talk about his or her feelings. Attempting to use logic and rational thought with a patient who is flooded with affect may lead to a further increase in agitation. An emotionally distraught patient requires an active response from the clinician. Use of active listening techniques, such as paraphrasing to the patient a brief encapsulated form of his or her statements, helps convey that the clinician is attempting to understand what the patient is experiencing. It is important to be honest and unambiguous. Verbal de-escalation techniques require time and are often sufficient to reduce agitation and eliminate the risk of violence. Offering food or drink may also help facilitate a therapeutic alliance and reduce a patient's agitation. Management of aggressive patients may be compromised when the clinician inadvertently evokes feelings of fear, anger, or rejection. This can generate thoughts of retaliation aimed toward the clinician. If identified, the clinician should stop the interview and ask another staff member to continue with the therapeutic intervention.

Table 32.3–5.
Differential Diagnosis of Violent Behavior

Medical
Cerebral infection
Cerebral neoplasm
Electrolyte Imbalance
Hepatic disease
Hypoglycemia
Hypoxia
Infection
Renal disease
Temporal lobe epilepsy
Vitamin deficiency
Substance induced
Alcoholic intoxication
Alcohol withdrawal
Stimulant intoxication
Delirium tremens
Inhalant Intoxication
Phencyclidine (PCP) intoxication
Sedative/hypnotic withdrawal
Psychiatric

Antisocial personality disorder
Bipolar disorder
Borderline personality disorder
Schizophrenia
Decompensating obsessive-compulsive personality disorder
Delusional disorder
Dissociative disorder
Impulse control disorder
Schizophrenia
Social maladjustment without psychiatric disorders
Uncontrollable violence secondary to interpersonal stress

Assaultive patients, due to their potential for unprovoked and unpredictable actions, should be provided a significant amount of space by all clinical providers. A general rule of thumb is to keep at minimum two quick steps or at least a legs length distance from the patient. Personal space can be visualized as an oval zone extending 4 to 6 ft all around. If the patient is standing, then the clinician should stand. If the patient is sitting, then the clinician should also sit down, preferably at the same level as the patient. The clinician should never stand over the patient during the interview. If the patient is pacing, then the clinician can walk with the patient but at a much slower pace. The clinician should take a posture that makes him or her appear small and thus less threatening. This can be done by holding the hands at waist level with palms up and open. This posture ensures the patient, especially if paranoid, does not misperceive the intentions of the clinician and ensures that there are no syringes or other devices of perceived harm in the practitioner's hands. Assuming the "Thinker" stance (one forearm crosses the chest, and the opposite elbow rests on it with the index finger on the cheek or chin) is also an effective technique. In this position the patient's content can be validated by affect but also ensure the practitioner is ready for any unexpected behaviors by the patient. Some patients may not respond or only partially respond to verbal intervention and will require psychotropic medication.

When both interpersonal and pharmacological interventions fail to reduce agitation and the patient remains uncooperative with treatment efforts, a show of force with five or more staff and/or security guards may encourage the patient to cooperate with treatment. The presence of additional staff supporting the clinician conveys to the patient that his or her aggressive impulses are not appropriate in the given circumstances and will need to be controlled. A show of force should not be implemented in a way that threatens or humiliates the patient but serve as a vehicle for ensuring safety of the patient, patient peers, and staff.

When less restrictive interventions fail to reduce a patient's agitation, physical restraints may be necessary to prevent imminent harm to the patient, patient peers, or staff. Prevention of serious disruption of the treatment setting, prevention of bodily harm, and/or prevention of

significant damage to property is the goal and restraints should be considered a last resort in the PES setting. Recent evidence suggests that the acute symptoms and presentation characteristics in a PES clinical encounter are more likely to be associated with restraint/seclusion than patient's underlying disease process or demographics. Careful approach to every patient is warranted especially if there is a history of impaired impulse control or assaultive behaviors, and PES providers should be prepared for behavioral escalation. Once the decision is made to restrain a patient, the restraint process should be implemented immediately and without negotiation. Rigorous attention to the patient's safety is paramount as restraints are contraindicated in the patient with an unstable medical condition such as infection, cardiac illness, body temperature instability, metabolic illness, or orthopedic problems. In addition, patients with delirium or dementia may experience a worsening of symptoms secondary to the sensory isolation induced by restraints and seclusion. In all restraint episodes, documentation should clearly outline the behavior requiring restraints, the interventions that were made to reduce the patient's agitation prior to restraints, and all interventional efforts to remove the patient from restraints. While a patient is in restraints, continuous monitoring of the patient should occur to prevent injury with 15-minute checks of extremities to insure adequate circulation, adequate hydration, and limb exercise. All clinical efforts should focus on removing the patient from restraints as quickly as is clinically possible. All clinicians should review and thoroughly understand the restraint guidelines, polices, and procedures of their respective institution.

PES should have written policies and procedures that are understood and practiced by all clinical staff. Patients should be evaluated and treated in areas that have adequate space, appropriate lighting and alarm systems, and the ready availability of security officers. Interviewing rooms should have multiple exits and a minimum of materials and furniture that could be used as weapons. As a clinician, it is important to understand your physical and psychological limitations in dealing with the unpredictability of a psychiatric emergency setting. Ensuring safety of patient and staff is the first priority and will mitigate the risk of a potential secondary emergency. Be aware of your surroundings, communicate effectively, and utilize resources available on site.

PSYCHIATRIC EMERGENCIES RESULTING FROM SUBSTANCE INTOXICATION

Stimulant Intoxication

Cocaine is short acting and produces rapid effects when snorted, smoked, or used intravenously (IV). Symptoms of acute stimulant intoxication include a heightened sense of well-being, increased sense of alertness, decreased anxiety, reduced social inhibitions, increased talkativeness, and

poor judgment manifested by loss of money, sexual indiscretion, and potential illegal activities. Patient symptoms usually resolve within 48 hours. When stimulant intoxication progresses to delirium, the patient exhibits grandiosity, psychomotor agitation, paranoid ideation, formication (sensation of insects crawling on the skin), bizarre behavior (such as sorting objects into various groups), stereotyped movements of the mouth and the tongue (such as teeth grinding, lip biting, and skin picking), tactile or visual hallucinations, and distorted perceptions. Delusional states, especially paranoia can also occur.

Physical findings include tachycardia or bradycardia, hypertension or hypotension, hyperthermia, dilated pupils, perspiration or chills, hyperreflexia, nausea, and vomiting. Patients may have evidence of weight loss, muscular weakness, respiratory depression, chest pain, or cardiac arrhythmias. Intoxication can lead to confusion, seizures, or coma. Patients who abuse stimulants IV are at risk for numerous medical conditions including endocarditis, tetanus, hepatitis, abscesses (MRSA), and AIDS. Urine and serum drug screens provide corroboration of the diagnosis and should be part of the PES intervention if suspected, or if there is a need for determination of diagnosis based on readily available information.

Treatment should include placing the patient in a quiet room away from the general traffic of the PES as external stimuli may exacerbate the physiologic effects of sympathomimetics. Explain the treatment in simple language. Reassure the patient that the effects of the stimulant will subside. Because the patient should not be left alone, family members or friends should be included in the treatment to provide further reassurance and support if safe to do so. Strategies discussed earlier in this section for managing aggressive patients should be utilized. If interpersonal interventions fail to relieve the patient's agitation, then a benzodiazepine can be administered. If the patient does not respond or is delusional, extremely agitated, or combative, then an antipsychotic medication with or without a benzodiazepine can be given as per the medication algorithm discussed earlier. The patient should be discharged to the care of family or friends with a follow-up outpatient appointment for the next day in a drug treatment program.

Hospitalization should be considered if

- the psychosis does not remit and effects ability to negotiate his or her needs;
- the patient has a polysubstance abuse problem and requires detoxification;
- the patient has no family or other support system and appropriate housing or shelter cannot be arranged;
- the patient has protracted cognitive impairment; and
- the patient has serious associated medical problems.

Alcohol Intoxication

Patients who are intoxicated with alcohol in the PES may be self-referred or brought in by police, friends, or relatives because of self-destructive, disruptive, or violent behavior. People under the influence of alcohol commit one-quarter of suicides. And more than one-half of murderers and their victims are intoxicated at the time of the act.

Blood alcohol levels usually correlate with the clinical presentation of the alcohol-intoxicated patient:

- 30 mg/dL: Attention difficulties and/or mild euphoria
- 50 mg/dL: Coordination problems and/or unsteadiness
- 100 mg/dL: Impaired memory and/or ataxia
- 200 mg/dL: Confusion and/or stupor
- >400 mg/dL: Seizure, coma, respiratory arrest, and/or death

Chronic alcohol use disorder patients will have a higher tolerance to the effects of alcohol so that blood levels of alcohol will not always correlate with the clinical presentation. Blood alcohol levels should always be obtained in the assumed intoxicated individual. Alcohol is metabolized at the rate of about 1 oz per hour or 10 to 30 mg per hour in patients with healthy hepatic functioning.

Symptoms of alcohol intoxication include aggressive or inappropriate behavior, mood lability, impaired judgment, slow or slurred speech, incoordination, unsteady gait, and nystagmus. Objective physical findings of intoxication may include autonomic instability; tremulousness; nausea and vomiting; transient visual, tactile, or auditory hallucinations or illusions; psychomotor agitation; and anxiety. Progression to withdrawal may be rapid in chronic and heavy users of alcohol, and symptoms should be monitored closely in patients with complicated use histories. At minimum, vital signs and neurological function checks should occur every 4 hours for patients at high risk for withdrawal.

Medical complications are common in chronic use of alcohol. These include but are not limited to gastrointestinal bleeding, gastritis, stomach or duodenal ulcers, cirrhosis, pancreatitis, vitamin deficiencies, cardiomyopathy, peripheral neuropathy, and malnutrition. In acute presentation, patients may have sunburn exposure and subdural hematoma from head trauma, thus making an exploration of patients' recollection of events prior to arrival imperative. Exploration of duration of use, quantity used, and timeline of the most recent exposure must be explored. Other drugs may also have been ingested, thus further complicating the clinical picture. Death may occur from aspiration of vomitus or respiratory depression if not clinically identified.

The first treatment priority is to protect the acutely intoxicated patient from harming him- or herself or others and to identify medical problems

that require immediate attention. A patient may be hyper- or hypoglycemic, bleeding internally, hypertensive, have end-stage cirrhosis, a cardiac arrhythmia, pneumonitis, hematoma(s), or one of the alcohol-related syndromes such as alcohol-induced persisting amnestic disorder or Wernicke encephalopathy. A person may appear obtunded by alcohol but have clouding of consciousness from hyper/hypoglycemia, blood loss, or a subdural hematoma. Clouded consciousness in the clinical presentation of a negligible serum alcohol level is concerning for medical pathology warranting intervention. Patient's vital signs should be obtained every 15 minutes if possible, and depending on the patient's medical history and physical examination, appropriate laboratory studies should be obtained. If control of agitation is required, then the use of antipsychotics and/or benzodiazepines should be used as described earlier in this section. Patients may require a higher level of care than may be offered in the PES setting, and emergency medicine personnel should be involved in care as acute intoxication can be a medical emergency.

Intoxicated patients tend to evoke negative responses from hospital staff, which in turn may further agitate an intoxicated patient. As with all difficult patients, personal feelings should not be allowed to influence professional judgment and clinical care. Utilization of all available resources should be the standard of care to ensure safety, as intoxicated patients can be unpredictable. Security forces, though seldom needed, should be alerted before interviewing an intoxicated patient. If possible, the interview should be conducted in a large comfortable room. The clinician should assume a tolerant, nonthreatening manner and be prepared to tolerate insults and threats. Offering food, coffee, or juice may also be helpful as well as other strategies for intervention with aggressive patients.

The Patient Placement Criteria of the American Society of Addiction Medicine can be very helpful for determining the level of aftercare. These criteria include

- level of intoxication or potential for withdrawal;
- medical comorbidity;
- emotional and psychiatric comorbidity;
- patient's acceptance or rejection of treatment;
- relapse potential; and
- availability of supportive environment.

Disposition should be to home with family, if available, or other community placement, if the patient presents no immediate danger and has no acute medical problems. The patient should be given a follow-up appointment to an outpatient program and referral to alcoholics anonymous (AA). Many facilities offer on-site substance use counselors, which may be beneficial in aiding in longitudinal planning as well.

Benzodiazepine, Sedative, and Hypnotic Intoxication

Benzodiazepine, sedative, or hypnotic intoxication is marked by impaired judgment, loquacity, mood lability, dreamlike feelings, and a disinhibiting of sexual and aggressive impulses. Slurred speech, incoordination, unsteady gait, lateral nystagmus, areflexia, muscle hypotonicity, hypotension, and hypothermia are common physical findings.

Management of benzodiazepine, sedative, or hypnotic intoxication is a complex medical intervention best left to clinicians who have experience in treating such cases. These patients should be treated in the medical emergency department. Barbiturate overdose, in particular, is a medical emergency. Most patients require admission to a medical intensive care unit for continued evaluation and treatment. Patients may require management of a comorbid psychiatric illness but intervention should be deferred until medical stability is ensured.

Opioid Intoxication

Symptoms of opioid intoxication include drowsiness, slurred speech, and potentially coma. Physical findings include decreased respiration, blue lips, blue or pale skin, pinpoint pupils (if brain damage has occurred, usually secondary to anoxia, pupils may be dilated or meperidine can cause mydriasis), hyperemic nasal mucosa, recent needle marks, pulmonary edema, cardiac arrhythmias, or seizure.

Death is secondary to pulmonary or cerebral edema.

If a patient is unstable due to the direct physiological effects of opioids, then the signs and symptoms are rapidly reversed by administration IV of a narcotic antagonist, for example, naloxone. Symptoms of opioid intoxication may be complicated by varying levels of intoxication or withdrawal from other drugs. Therefore, patients should have urine and serum drugs screens. Because of the life-threatening risk or its potential medical complications, opioid intoxication or overdose should be treated in the medical emergency department. Patients may require management of a comorbid psychiatric illness, but intervention should be deferred until medical stability is ensured.

Cannabis Intoxication

Cannabis includes hashish, bhang, marijuana, and purified Δ9-tetrahydrocannabinol (THC). Abusers commonly use cannabis with other substances, particularly alcohol and cocaine. The most common methods of abuse are smoking and eating. Drug effects develop within 20 to 30 minutes. There has been an increasing prevalence of use of cannabinoids due to recent legalization in various jurisdictions. Providers should always inquire about use independent of legality as cannabinoids may influence directly the psychiatric phenotype of the patient in the PES.

Clinical findings include impaired motor coordination, euphoria, anxiety, sensation of slowed time, impaired judgment, and social withdrawal. Within 2 hours of cannabis use, the patient may develop the following:

- Conjunctival injection
- Increased appetite
- Dry mouth
- Tachycardia

Marijuana intoxication usually will not be the primary reason for evaluation in the PES unless an adverse reaction to the drug occurs. Treatment involves reassurance and a calm environment. Symptoms will usually clear within 4 to 8 hours. If the patient's anxiety or panic becomes severe, then a benzodiazepine can be given. Paranoid delusions are rare but may require an antipsychotic medication. If a psychotic reaction lasts more than a few days, then the patient should be evaluated for a major psychiatric illness.

It is important to note that patients may have exposure to synthetic cannabinoid products in presenting to the PES. Patients may endorse use of "marijuana" use only, but synthetic cannabinoids have seen recent increases in incidence rates in the emergency setting. Patients may present with symptoms similar to natural occurring cannabinoids but can develop severe tachycardia, hypertension, marked restlessness, agitation, anxiousness, nausea with emesis, mydriasis, somnolence, hallucinations, elevation of blood glucose, hypokalemia, and even seizures. Patients may require supportive medical interventions to ensure stability, and physicians may need to treat symptomatically underlying psychopathology, especially when agitation and psychosis are salient features. If the presentation of a patient is inconsistent with use of cannabinoids alone, clinicians should ensure thorough screening for other substances, including synthetic cannabinoids. Patients should be monitored closely and if psychotic reaction lasts more than a few days, the patient should be evaluated for a major psychiatric illness.

Hallucinogen Intoxication

Hallucinogens include the following:

- Lysergic acid diethylamide (LSD)
- Psilocybin ("magic mushrooms")
- Dimethyltryptamine (DMT)
- Peyote
- Dimethoxymethylamphetamine (DOM or "STP")
- Mescaline

- Phencyclidine (PCP)
- Ketamine
- Methylenedioxymethamphetamine (MDMA, "Ecstasy," or "X")

Hallucinogens cause no known withdrawal symptoms. Drugs are mixed, adulterated, and sold under different name and, therefore, clinicians can never be certain what hallucinogen was ingested on the basis of a patient's provided history. Therefore, serum and urine screens must be obtained. Pre-existing psychiatric or medical illness or premorbid personality disorders may intensify hallucinogenic symptoms. The symptom complex is a result of drug–drug interactions, dose ingested, duration of action, effects of other drugs or concurrent illness, and rate of onset of symptoms. In addition, a user's premorbid personality, expectations of the user, and social setting can directly affect the behavioral phenotype and described effects. Hallucinosis typically lasts 16 hours to 2 or 3 days.

Symptoms include perceptual, emotional, and cognitive changes. Patients may see colorful mobile displays of geometric designs or complex human or animal forms. Colors appear more intense and fixed objects undulate and flow. The patient has a greater sensitivity to touch and an altered sense of taste and smell. Synesthesia may occur (an overflow of one sensory modality into another; for instance, the patient "hears" colors or "sees" music). Other perceptual changes include depersonalization, derealization, and body image distortions in which body parts seem to grow, shrink, or disappear.

Emotional presentations range from euphoria, elation, and bliss to tension and anxiety and may culminate in a panic attack. The patient may laugh or cry for prolonged periods of time. Cognitive changes include loosening of associations with unusual content, illogical and fantasy-laden thought sequences, paranoid grandiosity, and, less commonly, persecutory delusions. Physical findings include hypertension, hypothermia, dilated pupils, tachycardia, sweating, palpitations, tremors, blurred vision, incoordination, and hyperreflexia.

Hallucinogen intoxication usually will not present to the PES unless an adverse reaction to the drug occurs or a patient places him- or herself in imminent danger due to impaired judgment and is brought in by a concerned family member or agency. The primary intervention involves reassuring the patient that the drug's effects will subside within several hours. It is helpful to remind the patient that the hallucinogen has produced his or her extraordinary ideas and feelings and that the drug effects will soon dissipate. When possible, friends and family members can help alleviate the patient's anxiety by maintaining a continuous dialogue in a supportive, nonthreatening environment. Because closing the eyes can intensify the hallucinogenic state, the patient should sit up or walk. Clinicians should try to avoid pharmacological intervention. If the patient's

anxiety becomes overwhelming, then a benzodiazepine can be given. Severely agitated patients may require an antipsychotic agent.

Phencyclidine Intoxication

Originally marketed as an anesthetic agent, PCP was taken off the market because of the psychotic reaction experienced by patients postoperatively. PCP commonly called "angel dust" is an illegal drug often sold to consumers under many names, including crack or LSD (it is neither). The drug may be taken orally, inhaled, snorted, or injected IV. PCP presents a confusing clinical picture with diverse psychological, physiological, and behavioral effects and may mimic schizophrenia. PCP highlights the importance and utility of urine and serum drug screening in the complicated patient presentation.

Patients can become extremely violent when intoxicated with this illicit substance. It can create a significant paranoid psychosis resulting in harm to others and, when taken in sufficient amounts, can produce death from respiratory failure, cardiac arrhythmias, and adrenergic crisis. Clinical findings are diverse and can include calmness to the point of sleep, agitation, hostility, disorientation, emotional lability, a sense of slowed of time, out of body feelings, assaultiveness, unpredictability, impaired judgment, hallucinations, and paranoid ideation. Because of the amnestic quality of PCP, patients will often not remember how they behaved while intoxicated.

Clinical signs and symptoms tend to correlate with PCP blood levels and include vertical and horizontal nystagmus, ataxia, loss of pain response, drooling, nausea, and mild pulse and blood pressure elevation at low doses. Higher blood levels include catatonia, rigidity, myoclonus, fever, elevated pulse and blood pressure, hyperreflexia, and elevated creatinine phosphokinase secondary to muscle trauma. PCP intoxication can proceed to death as the dose increases.

Patients with mild to moderate intoxication should improve with supportive care. The patient will require supervision in a nonstimulating environment that should include protection from self-injury. If symptoms remit and there is no cognitive impairment after 12 hours, then the patient can be discharged from the PES. Moderate to severe intoxication will require supportive care in an emergency department or medical inpatient unit with special attention to respiratory and cardiac function. It is often important to obtain serum creatinine kinase levels. Patients may require IV fluids to ensure that renal damage does not ensure. Vital signs need to be monitored for laryngeal stridor or respiratory depression. Restraints should be avoided to avoid the increased risk of rhabdomyolysis and resultant acute renal failure.

The preferred method of physical restraint for combative patients is total body immobilization by rolling the patient in a sheet. Sedation with a

benzodiazepine may be required for anxiety, agitation, or harmful behavior. An antipsychotic may be given for psychosis. Patients who take PCP orally may benefit from gastric lavage. Urine acidification is controversial, and this decision is best left to emergency medicine physicians. PCP is highly lipophilic and absorbs quickly into adipose tissue. Acidifying the urine to a pH below 5.5 may help flush PCP from adipose.

Inhalant Use

Inhalants can cause irreversible damage. They can cause death by suffocation, aspiration vomitus, or myocardial infarction. Patients who abuse inhalants sniff, aerosols, cleaning solutions, nail polish remover, lighter fluids, paint thinner and other petroleum products, or glue containing toluene. Intoxication occurs within 5 minutes and ceases in 1 to 2 hours.

Clinical findings include belligerence, assaultiveness, apathy, impaired judgment, and impaired social and occupational functioning. Physical findings include dizziness, nystagmus, slurred speech, unsteady gait, depressed reflexes, tremors, blurred vision, diplopia, stupor, coma, rash around the nose and mouth, breath odors, residue on face, hands, and clothing, redness and swelling of the eyes, and irritation of the throat, lungs, and nose. Major medical problems associated with inhalant abuse include cardiac arrhythmias, hypoxia, hepatitis, renal failure, bone marrow suppression, impaired pulmonary functioning, and skeletal muscle weakness. Death can result from central nervous system depression, cardiac arrhythmias, bronchospasm, and ventricular fibrillation.

Care is mainly supportive. Patients should be kept in an environment with minimal stimulation. Clothes smelling of solvents should be removed, and solvents on skin or hair should be washed away with a mild soap and water. Marked agitation should be controlled with antipsychotic medication as benzodiazepines may potentiate the effects of inhalants. Hospitalization may be warranted. Chronic, heavy use of inhalants has been reported to precipitate withdrawal symptoms within a few hours to several days after cessation of use. These symptoms include tremulousness, anxiety, inattention, diaphoresis, diarrhea, restlessness, and insomnia. Withdrawal symptoms can be treated with benzodiazepines. Whether there is an actual withdrawal state from inhalants is controversial. However, chronic inhalant users often have polysubstance involvement and clinicians should carefully screen for other substances and treat accordingly.

PSYCHIATRIC EMERGENCIES RESULTING FROM SUBSTANCE WITHDRAWAL

Withdrawal from Alcohol

Alcohol withdrawal is a medical emergency. If left untreated, it may lead to death. Clinical signs of alcohol withdrawal include disorientation,

fluctuating attentiveness, and intermittent agitation, commonly violent in nature. Physical findings include tachycardia, sweating, hypertension, nausea, vomiting, tremulousness, hyperreflexia, orthostatic hypotension, and, occasionally, generalized seizures. Patients in withdrawal are usually depressed, irritable, and anxious and may experience illusions or transient visual, tactile, olfactory, and auditory hallucinations. Hallucinations are usually frightening—the patient sees and feels mice or lice crawling on the skin or sees animals, especially snakes. Alcohol withdrawal can be complicated by seizures and delirium tremens.

Death occurs in 4 to 20 percent of cases from complications such as hyperthermia, aspiration, or vascular collapse. Onset of complicated withdrawal is usually 2 to 10 days after cessation or decreases in alcohol use and may last hours to days. Symptoms of uncomplicated alcohol withdrawal peak 24 to 48 hours after the last drink and subside within 5 to 7 days, even without treatment. Irritability and insomnia may last for 10 days or more. Withdrawal is more severe in patients with previous episodes of withdrawal or medical illnesses.

Alcohol withdrawal symptoms are generally effectively treated with benzodiazepines. Chlordiazepoxide (25 to 100 mg every 4 to 6 hours and as needed) provides for a smoother detoxification given its long half-life. Lorazepam (1 to 2 mg orally or intramuscularly every 6 hours and as needed) is preferred when IM medication is necessary or patients have markedly impaired liver function. Lorazepam is safer for patients with hepatic disease or brain damage due to renal excretion and shorter half-life. An antipsychotic as recommended earlier in this section could be used for patients with hallucinations. Pulse, blood pressure, and temperature must be monitored every 30 minutes in the PES. One hundred milligrams of thiamine is given IM to address expected vitamin deficiency. Delirium tremens or a seizure is a medical emergency and must be treated in the medical emergency department. If a patient has a documented or verbalized history of alcohol withdrawal related pathology, it is considered a medical emergency and patients will require monitoring and potential hospital admission, independent of any comorbid primary psychiatric pathology.

Withdrawal from Benzodiazepines, Sedatives, and Hypnotics

Withdrawal from benzodiazepines, sedatives, and hypnotics is similar to alcohol withdrawal. Physical findings include tremors, increased pulse and respiration, hyperthermia, loss of appetite, nausea and vomiting, postural hypotension, hypertension, seizures, and delirium. Patients who abuse sedatives are at risk for infections and hepatitis.

Withdrawal from short-acting agents will have an onset 12 to 24 hours after the last dose, peaking between 24 and 72 hours. Withdrawal symptoms with longer-acting agents peak 5 to 8 days after the last dose.

The elderly or patients with hepatic disease may develop symptoms more slowly because of decreased drug metabolism.

There are several pharmacological options for withdrawing a patient from benzodiazepines, sedative, and hypnotic drugs.

- One method is to gradually withdraw the agent of dependency.
- Substitute the dependent agent with phenobarbital or a longer-acting benzodiazepine and taper. For short half-life benzodiazepines, treatment would be with a longer-acting drug such as clonazepam or chlordiazepoxide.
- In cases of barbiturate or other mixed sedative/hypnotics, a phenobarbital taper can be used, as this will allow for a safer detoxification because it produces few changes in blood levels between doses, it has observable signs of toxicity, and intoxication does not produce euphoria. A phenobarbital taper is done as follows:

  1. Administer a 200-mg test dose of pentobarbital orally and assess neurological status after 1 hour. Pentobarbital can be converted to phenobarbital, which has a longer half-life with fewer variations in blood level (100 mg of pentobarbital is equivalent to 30 mg of phenobarbital).
  2. If the patient is asleep but arousable, then no treatment is required. If the patient is drowsy with slurred speech, coarse nystagmus, or ataxia, then the patient's estimated pentobarbital requirement is 400 to 600 mg per 24 hours.
  3. If the patient is comfortable with only fine lateral nystagmus, then the 24-hour pentobarbital dose is 800 mg.
  4. If the pentobarbital has no effect, then the 24-hour pentobarbital dose is estimated at 1,200 mg or more.

Patients who are withdrawn from sedative hypnotics should be admitted to the hospital for the withdrawal protocol because this is a life-threatening condition. Clinicians should also familiarize themselves with the dose equivalents of benzodiazepines and other sedative drugs. Some common benzodiazepine dose equivalents are

- Alprazolam (Xanax) 0.5 mg
- Chlordiazepoxide (Librium) 10 mg
- Clonazepam (Klonopin) 0.25 mg
- Diazepam (Valium) 5 mg
- Lorazepam (Ativan) 1 mg
- Oxazepam (Serax) 15 mg

Withdrawal from Opiates

Patients' opioid use disorders may present to the PES requesting

detoxification or, more commonly, present in withdrawal requesting relief. In addition, patients may have significant mood symptoms that may potentiate the untoward effects of withdrawal. Cessation of opiates, while uncomfortable, is not life threatening. Opioid withdrawal is precipitated by the abrupt cessation of opioid administration. When heroin or morphine use is discontinued, withdrawal symptoms appear within 6 to 8 hours, peak within 36 to 72 hours, and last for 7 to 10 days. Meperidine withdrawal may not begin until 1 to 3 days after the last dose with symptoms dissipating after 10 to 14 days.

Withdrawal symptoms include nausea, vomiting, diarrhea, dilated pupils, muscle aches and pains, insomnia, lacrimation (tearing eyes), flu-like weakness, rhinorrhea (running nose), piloerection, sweating, flushing, abdominal cramps, increased temperature, yawning, and semen ejaculation. Patients who are dehydrated or debilitated can have life-threatening complications. Seizures and delirium tremens do not occur.

A detoxification regimen should be started only when physical signs of withdrawal appear. The most common approaches to opioid detoxification are methadone substitution and withdrawal and clonidine for symptom blockade. Symptom management may include trimethobenzamide for nausea, dicyclomine for gastrointestinal cramping, and ibuprofen for body aches. As opioid withdrawal is not life threatening, clinicians should not start methadone in the emergency setting. Patients should be referred to a methadone maintenance or drug rehabilitation program for further evaluation and treatment. However, if a patient is in established treatment for opioids with methadone, efforts should be made to continue care. This includes contacting current providers and ensuring a prevention of lapse in care if possible.

Withdrawal from Cocaine and Amphetamines

Withdrawal from amphetamines and cocaine produces dysphoria, hypersomnia or insomnia, increased appetite, apathy and lack of energy, muscle aches, chills, depressed mood, anxiety, irritability, and vivid and unpleasant dreams. Patients may be agitated or have psychomotor retardation. Drug cravings may be intense. Suicidal thoughts can occur, and in some cases, patients may experience paranoia or hallucinations. Severity of the symptoms will depend on the potency of the stimulant. With amphetamines, these symptoms commonly occur 3 or 4 days after a decrease or cessation of heavy amphetamine use. The depression can last for months, although symptoms may last only 8 to 48 hours with milder symptoms persisting for several weeks.

There is no detoxification regimen for stimulants. Usually after several nights of sleep, the dysphoria and suicidal ideation resolve. A calming environment can decrease agitation, and clinicians should avoid a confrontational approach. Quiet time, food, and sleep are the best

interventions. However, in the PES, the patient has to be carefully evaluated for suicide risk and may require inpatient hospitalization. Many of these patients do not regularly use the health care system, and clinicians should always maintain a high index of suspicion for undiagnosed and/or untreated medical disease. Because these patients are at a high risk for relapse, an inpatient rehabilitation program should be considered. Given that antidepressants have a delayed onset of action, their use is of no value in substance-induced mood disorders in the PES. If the patient is discharged home, then it is preferable to have family or another supportive individual accompany the patient home from the PES.

OTHER PES-RELATED ISSUES

Domestic Abuse

Domestic abuse refers to a behavior pattern in which a person repeatedly inflicts physical injury, pain, fear, or mental anguish on another member of the household. The abuse may be imposed through physical, psychological, sexual, or economic means. Neglect occurs when a care provider fails to provide the basic necessities for living to a dependent individual. Given the marked prevalence of the problem and the difficulty of identifying domestic abuse and neglect, clinicians should maintain a high index of suspicion. Knowing the signs that suggest domestic violence can help the clinician to determine the necessary intervention. If the injuries form a pattern of abuse it may inform whether the abused or neglected patient should be separated from the abuser, or whether long-term treatment is necessary. Although men are less likely to have a history of abuse than women, they are more likely to be attacked by their partners with knives, guns, or household objects and suffer serious injury or death.

Treating a patient's physical injuries is the first priority. It must then be determined whether the patient's significant other caused the injuries. The clinician should obtain a history of the current relationship as well as a history of the previous types of abuse. Victims of abuse often will give inconsistent explanations for their injury. Frequently, the abused adult patient may minimize the pain suffered from the attack and impede investigating the abuse or claim that the abuser did not really mean any harm. Issues such as whether the patient should report the abuse and whether the patient should seek placement other than his or her home are complicated and require a great deal of patience and skill on behalf of the clinician. Whether or not to seek further social services and legal aid also factor into the intervention. Treatment options may include psychotherapy, special abuse programs, and self-help and advocacy groups. Knowledge of available community resources is important for PES staff.

Elder Abuse

Elderly patients can be victims of passive neglect or physical abuse. The most common form of abuse is neglect, which accounts for 55 percent of reported cases, although physical abuse (14.6 percent) and financial exploitation (12.3 percent) are also common. The clinician's high index of suspicion is a critical component to making the diagnosis. The first priority is always to treat the physical consequences of the abuse or neglect. An elderly patient may minimize the problem, appear afraid of the caregiver, and express fears about returning home. Patients should be asked about being hit, being physically harmed in the past, or being deprived of care or necessities such as food and clothing. Signs of neglect in an elder person may include malnourishment, dehydration, poor hygiene, or pressure sores. The misuse of drugs, including overdosing, underdosing, or inability to obtain prescribed medications may reflect abuse. Sometimes additional evidence of neglect can be obtained from the person that brought the patient to the emergency department. This person may have seen the inside of the patient's home and can comment on the level of cleanliness, the functioning of utilities, and the adequacy of appliances such as walkers and commodes.

Perpetrator risk factors include substance abuse, psychiatric illness, psychosocial stressors, emotional and financial dependency, history of abuse, social isolation, inexperience in care giving, and misinformation as to how to provide care. The single most predictive factor for elder abuse is alcohol abuse by the caregiver.

Interventions include reducing the stress on the family by expressing empathy and providing support to the caregiver. Identifying interpersonal conflicts that underlie the abuse may also help alleviate tension in the family. Treatment for a patient, the abuser, and another family member must continue outside the hospital. The clinician must inform the family of appropriate community resources and make appropriate referrals. If elder abuse is suspected, mandated reporting laws require a concerned clinician to report findings to the appropriate authorities. Clinicians should be familiar with the local statutes concerning this issue and the available mechanisms for reporting suspected abuse.

Child Abuse

Of the approximately 3.4 million cases of child abuse reported annually, the majority of reports are cases of neglect, with smaller proportions related to physical abuse, sexual abuse, and emotional mistreatment. Because child abuse is underreported, the clinician should always maintain a high index of suspicion. Risk factors associated with child abuse include a single-parent household, especially if the father is the sole caretaker, children between 3 and 12 years in age, behavioral problems, large family size, and poverty. Of all of the risk factors, poverty is most strongly associated with abuse and neglect. Caretaker factors that have been associated with or have

contributed to child abuse and neglect include younger age, poverty, personal history of abuse, large family size, unplanned pregnancies, illicit drug abuse and alcohol use, inadequate parenting skills, low-birth-weight babies, and social isolation.

The interview of a child who may be a victim of abuse requires a compassionate and sensitive approach, which assures the child that he or she is in no way at fault. The most successful interviews occur when the child is interviewed alone. The clinician should sit at the child's eye level, explain the purpose of the interview, use open-ended questions when possible, and not display shock or horror at the child's descriptions. Depending on the outcome of the interview, the clinician has numerous options for treatment including hospitalization, outpatient treatment and family therapy, contacting the state or city social welfare agencies, and, depending on the state legal statutes, contacting legal authorities. Medication may be considered to treat underlying psychopathology but not to treat reactive emotional states. If underlying mental illness is discovered, use of child and adolescent psychiatrists is optimal to ensure appropriate longitudinal care by specialized providers.

Rape

Rape is a major psychiatric emergency. A victim who is not managed properly may develop enduring patterns of psychological and sexual dysfunction. Because of social stigma, taboo, or fear of retaliation by the offender, the victim will often not report rape. There is often an accompanying fear of violence and death and of contracting sexually transmitted diseases. For women, there is the additional fear of becoming pregnant. The management of the rape victim involves medical, mental health, and forensic interventions that may be initiated in the PES. However, the primary goal of the assessment for a victim of rape is to restore psychological safety by providing information, correcting misattributions, restoring and supporting effective coping, and ensuring social support. A respectful, compassionate, and sensitive approach is necessary to forge a therapeutic relationship with victims.

Helping the patient recognize that he or she is in the hospital and is safe can restore psychological safety. Because patients may be sensitive to power and control issues, a common respectful approach will help develop this sense of safety. Patients who react to the trauma with dissociative symptoms may appear calm and unaffected by their ordeal yet have a high risk of developing posttraumatic stress disorder. Once a sense of safety is achieved, it is essential to provide information to the rape victim about his or her current medical status. Addressing fears of future health problems, including pregnancy, sexually transmitted diseases, and HIV infection is paramount and will aid in reducing anxiety. Information must also be given about the legal system, the responsibility of the police, and the need for

evidence collection. Clinicians should acknowledge the patient's fears of being touched and undergoing vaginal and rectal examinations and discuss the potential for these examinations to trigger flashbacks or intrusive thoughts of the rape or sexual assault. The presence of a trusted family member or friend may help the patient to cope with the examination. Victims of rape or sexual assault commonly blame themselves for what happened and may have persistent fears of being out in the dark or a belief that going out alone is dangerous. Helping the patient to recognize the impairing nature of his or her postrape beliefs may decrease anxiety, guilt, and anger. Education about common reactions to rape such as irritability, hypervigilance, and problems with sexual intimacy is important to share with the patient once a therapeutic relationship has been established. Discussing sleep problems, intrusive thoughts, nightmares, avoidance, and numbing may also aid restoring psychological stability and allay fears that he or she is crazy or has lost control. Ongoing counseling and psychotherapy should be recommended for all sexual abuse victims. Various institutionally based services may be available and it is important for the PES provider to be familiar with all potential resources to help patients of sexual trauma.

PSYCHOSOCIAL ISSUES IN THE PES

Homelessness

Many of the urban homeless are chronically mentally ill. PES and social services personnel are frequently called on to work with homeless persons after their immediate physical problems (thermal injuries, infections, congestive heart failure, and trauma) are treated. Although the typical "street person" is difficult to characterize, these patient populations may suffer from one or more psychiatric illnesses, including schizophrenia, substance use disorders, neurocognitive disorders, and personality disorders. Because of the extensive biopsychosocial pathology seen in the homeless, numerous resources must be devoted to their evaluation and treatment.

It is often difficult to engage homeless populations. Homeless individuals often do not trust or seek to avoid contact with the medical or social service systems. If the clinician cannot develop a rapport with the patient, it is not generally an effect of poor training. In these instances, other PES staff members may be utilized to try to establish a relationship with the patient in order to assess his or her needs. The clinician's initial task is to determine what the patient wants (shelter, medical attention, or antipsychotic medication) when self-presenting. However, police or other agencies may bring in homeless patients for care, and the chief complaint may be obscure or often nonexistent.

Helping the patient to clean up and offering food and drink may be the

best ways to engage the patient and learn about his or her psychiatric and social needs. As a general rule, all homeless persons should be screened for medical illness. Some people may be demented as a result of head trauma or nutritional deficiencies. If possible, these patients should be hospitalized for further evaluation.

While restarting medication is an option, it is unlikely that the patient will continue to take medication following discharge from the PES unless a more structured living situation can be arranged. The mentally ill homeless person is commonly resistant to a meaningful change in lifestyle. A caring, nonthreatening clinician may be able to place the patient into the mental health care system. Linking a patient to a drop-in center or day hospital may be worthwhile, if possible, and many urban facilities may offer specialized programs in providing assistance. Working with the patient and targeting specific identified goals can potentially develop a positive therapeutic alliance and allow a homeless patient to receive necessary services longitudinally.

High Utilizers of Service

Patients may utilize psychiatric emergency care services for inappropriate secondary gain. Beginning with symptoms, patients may complain of urgent needs that are ripe with unrealistic expectations and entitled wishes for immediate relief. Patients may relate in a generally negativistic manner and are often regarded by staff as demanding and manipulative. These patients' desires and demands are usually quite different from their clinical needs on assessment, and can create intense frustration on the part of both the patient and the treatment team. High utilizers of resources have a paucity of supportive interpersonal networks as they tend to interact within their social environments in much the same way they do in the PES, leaving available resources exhausted and depleted.

Managing these patients requires a flexible approach. The patient's autonomy should be supported, but each emergency visit should be viewed as an opportunity to encourage both the patient's engagement and participation in treatment. In addition, every effort should be made to enhance the patient's understanding of what the PES provides and the limitations of service. Some of these patients may ultimately develop a positive transference to the PES that will help them maintain a relatively consistent level of functioning. While the optimal goal would be to have them engaged in an outpatient program, their relationship with the PES may keep some patients under the protective mantle of a single health care facility and ensure appropriate and safe longitudinal care. A partial hospital or intensive outpatient program is an effective, relatively stabilizing disposition and should be encouraged by PES staff. However, the availability of such resources may be limited and the care in many instances rests on PES providers.

Perhaps the most difficult aspect in managing any patient whose expectations are unmet is the clinician's own discomfort. Because these patients may be manipulative, uncooperative, and ambivalently help-rejecting, the clinician must remind him- or herself that the patient's style is likely a reflection of interpersonal deficits. The deficit usually being the result of early developmental problems having little to do with the clinician. Such an understanding should help the provider resist being drawn into the transference struggle and instead should help him or her to respond empathically despite provocation. In addition, a supportive approach combined with operant shaping strategies that set limits on demanding or inappropriate behavior can be effective when interacting with this population.

LEGAL ISSUES IN EMERGENCY PSYCHIATRY

Relevant legal issues in the PES include confidentiality and release of information, duty to warn and/or protect, informed consent and right to refuse treatment, competency, and involuntary commitment. An overarching concern for all of these issues is the need for adequate documentation in the medical record. The chart is a legal document; therefore, a court can subpoena a medical record. This is the most important piece of evidence in most malpractice cases. Charting should include an accurate assessment of the patient and all procedures, interventions, treatments, and plans for continued care. All chart entries should be timed, dated, written legibly, and signed. Fraudulent documentation or "doctoring" the chart is a very serious legal offense.

Confidentiality

Since the time of Hippocrates, information shared between a doctor and a patient has been considered confidential. Accordingly, the American Psychiatric Association has established ethical codes addressing confidentiality. A physician is ethically and legally obligated to maintain the privacy of all information shared by the patient in a clinical doctor–patient setting. This "privileged" information belongs to the patient and can be shared with others only if the patient gives his or her permission to release the information or the court system demands it. There are, however, several clinical situations in which confidentiality is waived such as the court system deem more harm than good will result if patient confidentiality is maintained. These include situations of child/dependent/elder abuse or neglect, competency hearings, court-ordered psychiatric assessments, danger to self, danger to others, expressed intent to commit a crime, and communication with other treatment providers involved in the care of the patient.

Duty to Warn

For the PES physician, the duty to warn is of particular significance. If a physician determines that there is a substantial risk of harm to another person if no intervention is made, then he or she is legally obligated to warn the person at risk. It is imperative for the psychiatrist to determine whether information provided is a credible threat. Statements such as "my spouse makes me so angry sometimes I'd like to kill him" should not be taken at face value as a legitimate threat. Further questioning and clarification, however, may be necessary in order to determine any real lethal intent. Obtaining information from family, friends, and other treatment providers will often help clarify the credibility of a threat. When available, medical records of prior psychiatric treatments or the patient's managed care company's clinical database can provide pertinent clinical information. The medical record should detail the patient's statements and all follow-up information obtained, including the physician's opinion on the credibility, or lack thereof, of the threat. The record should justify the clinical action taken by the doctor and should include any specific plans expressed by the patient, capacity to carry out the plan, prior history of violent behavior or aggression, access to weapons, current mental status examination, and a current psychiatric differential diagnosis. The clinician should also document any consultation with colleagues, supervisors, administrators, or other authorities. If the threat is deemed credible or the clinician is unsure, then the patient should be admitted to a psychiatric unit for further evaluation as potential imminent dangerousness to others is a viable reason for inpatient care.

Evaluation of Capacity and Competency

Clinical capacity is broadly defined as the ability of a patient to make informed decisions and consent to receive medical treatment as assessed by any physician. Similarly, competency is a patient's legal capacity to engage in the decision-making process regarding his or her medical care. A patient's incompetence, however, may be determined only by a judge who is often supplemented with a physician's capacity evaluation. By law, all adults are presumed competent unless otherwise determined by a court of law. In a psychiatric emergency, waiting for a judgment from court is rarely practical. In assessing a patient's decision-making capacity, a patient to must be able to

1. understand the basic information necessary to make a decision;
2. understand alternative treatment strategies available;
3. appreciate the consequences of treatment and of no treatment; and
4. integrate all of the relevant information to make a consistent decision on their own behalf.

Aside from these four criteria, the standard for capacity may also vary on the risks to benefit ratio of the proposed treatment. A high benefit with

low-risk procedure may carry a higher standard of capacity to refuse and a lower standard of capacity to accept the treatment.

Competency is situation specific, but a patient may also be assessed for global incompetence. A patient may be competent to consent for treatment and still suffer from major mental illness. A patient who is delirious, floridly psychotic, demented, or who is acutely intoxicated or withdrawing from drugs is likely not competent to make decisions concerning his or her treatment. In these cases, a surrogate decision maker, usually a family member acting in the best interest of the patient, may decide on the patient's behalf.

Competency is a medicolegal determination and should not be used to coerce a patient into treatment, even if the physician feels it is in the best interest of the patient. The physician may not agree with the decision, but the patient, if found competent for decision-making capacity, has the legal right to determine his or her or her own treatment. Accurate and timely documentation, including a mental status examination in the medical record is mandatory in all cases regarding patient competency, especially when the patient refuses the recommended treatment.

Informed Consent

Physicians often equate informed consent with a person giving permission to participate in a clinical research trial. The concept of informed consent, however, is relevant to all clinical settings, especially the PES. The doctrine of informed consent mandates that a patient must voluntarily give permission to receive or deny medical or psychiatric treatment. However, in order to provide consent, the patient must be deemed competent to make medical decisions on his or her own behalf (see "Evaluation of Capacity and Competency" previously) and be free of any coercion. A key component of informed consent is that adequate information must be provided to the patient or surrogate decision maker. In some states, a materiality standard or the amount of information needed for an average individual to make a decision on a proposed treatment is required versus, in other states, where the amount of information required depends on the needs of the specific individual. Typically, the information for informed consent includes the patient's diagnosis, nature of the treatment proposed, treatment risks, benefits, and alternatives, if available, likely outcome with or without the proposed treatment, and the unpredictability of some treatment outcomes.

The medical record should document informed consent granted for treatment given or why treatment was necessary despite a patient's refusal. In some cases with patients who have chronic psychiatric conditions, a court-appointed delegate might be designated to provide consent on the patient's behalf. This is a legal determination and not one made by a psychiatrist. In children, parents generally have the right to treatment

decisions except in the case of emancipated minors. The age and nature of treatment where parental authorization is required vary from state to state.

There are four circumstances in which a psychiatrist may treat a patient without first obtaining informed consent from the patient.

1. The patient is incompetent (a legal determination).
2. Emergency care is necessary to prevent serious risk to the patient or others.
3. There is a therapeutic waiver, where a competent patient designates the physician to make treatment choices for him or her.
4. Therapeutic privilege, which allows the physician to withhold information from the patient if it poses a threat to the patient or would impair his or her ability to make a competent decision.

Whenever informed consent is not obtained, precise and thorough documentation of the clinical circumstances should be placed in the medical record in the form of a legible, dated, and timed progress note.

Involuntary Commitment

Psychiatrists are in a unique position in that they have an ethical and legal responsibility to care for patients against their will if they meet specific criteria. Depriving an individual of his or her civil liberties is often a difficult decision as patients, due to an underlying mental illness, may not have the insight into their actions and subsequent repercussions of behaviors. Although each state has its own involuntary commitment procedures, virtually all include dangerousness to self, dangerousness to others, or grave impairment leading to an inability to care for oneself as a result of mental illness. A patient may be severely mentally ill, but if he or she is not at imminent risk of harm to him- or herself or others, then he or she generally cannot be placed into involuntary care. Family members, friends, police, or physicians can petition for the patient to be evaluated by a psychiatrist, at which point the psychiatrist determines the dangerousness of the patient's behavior. If deemed dangerous, then the patient may then be hospitalized for a few days to a week, depending on the state mental health laws. Within a designated time period, the patient must have his or her specific case reviewed by a mental health court to determine whether he or she may be released or committed for longer treatment. It is important to emphasize that mental illness is necessary but not sufficient for involuntary psychiatric hospitalization. Homelessness, for example, does not justify forced hospitalization.

CONCLUSION

This section has reviewed some of the unique challenges that present to clinicians working in a PES. As is the case in all emergency department settings, the need for a focused, thoughtful initial evaluation is the key to

clinical success. Unlike an outpatient service, time pressure is an important factor. The PES clinician must be able to complete a biopsychosocial assessment, formulate a differential diagnosis, and make treatment or placement decisions, often in a few hours or less. The primary focus is on patient and staff safety and initiating appropriate treatment. The importance of concise, accurate charting cannot be overstated. It is unlikely that the PES clinician will be following the patient long term. The medical record will be the enduring material supporting placement decisions and must contain the supporting clinical information. Accurate dating and timing of all notes is particularly important in the PES, as this may impact on legal issues (such as involuntary commitment) and clinical issues (such as medication dosing). Maintaining a high index of suspicion for medical illness will minimize potential mortality and morbidity of PES patients. Managing violence is of particular importance. All patients presenting to the PES should be considered potentially violent. This "universal precaution" mentality will help avoid dangerous situations from developing. PES staff must be familiar with the array of social resources available in the community, including shelters, drug and alcohol treatment services, and community mental health services. These are vital to running an effective PES. PES staff must also be familiar with the legal aspects in providing the standard of care in a psychiatric emergency setting. Although emergency psychiatry is challenging for clinicians, it is also very rewarding and a critical component of the mental health delivery system.

References

Allen MH, Currier GW, Carpenter D, Ross R, Docherty JP; Expert Consensus Panel for Behavioral Emergencies 2005. The expert consensus guideline series. Treatment of behavioral emergencies 2005. *J Psychiatr Pract*. 2005;11(suppl 1):5–108; quiz 110–112.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

Berg AZ, Bell CC, Tupin J. Clinician safety: assessing and managing the violent patient. *New Dir Ment Health Serv*. 2000;(86):9.

Buckley PF, Noffsinger SG, Smith DA, Hrouda DR, Knoll JL IV. Treatment of the psychotic patient who is violent. *Psychiatr Clin North Am*. 2003;26:231–272.

Cameron CL. Hospital based psychiatric emergency services: a guide to service delivery and clinical best practices. *Int J Emerg Ment Health*. 2006 Winter;8:23–33.

Crowley TJ, Sakai JT. Inhalants. In: Galanter M, Kleber HD, eds. *Textbook of Substance Abuse Treatment*. 3rd ed. Washington, DC: American Psychiatric Publishing; 2004.

Currier GW, Allen MH. Emergency psychiatry: physical and chemical restraint in the psychiatric emergency service. *Psychiatr Serv*. 2000;51:717–719.

Currier GW, Medori R. Orally versus intramuscularly administered antipsychotic drugs in psychiatric emergencies. *J Psychiatr Pract*. 2006;12:30–40.

D'Orio BM, Purselle D, Stevens D, Garlow SJ. Reduction of episodes of seclusion and restraint in

a psychiatric emergency service. *Psychiatr Serv*. 2004;55:581–583.

Dubin WR, Weiss KJ. *Handbook of Psychiatric Emergencies*. Torrance, CA: Homestead Schools; 1997.

Hermanns-Clausen M, Kneisel S, Szazbo B, Auwarter V. Acute toxicity due to the confirmed consumption of synthetic cannabinoids: clinical and laboratory findings. *Addiction*. 2013;108(3):534–544.

Hillard R, Zitek B. *Emergency Psychiatry*. New York, NY: McGraw-Hill Companies; 2004.

Kahn MW. Emergency psychiatry: tools of engagement: avoiding pitfalls in collaborating with patients. *Psychiatr Serv*. 2001;52:1571–1572.

Korbin JE, Krugman RD. *Handbook of Child Maltreatment*. Netherlands: Springerlink; 2014.

Kosten TR, O'Connor PG. Management of drug and alcohol withdrawal. *N Engl J Med*. 2003;348:1786–1795.

Ladavac AS, Dubin WR, Ning A, Stuckeman PA. Emergency management of agitation in pregnancy. *Gen Hosp Psychiatry*. 2007;29:39–41.

McKinney C, Renk C. Atypical antipsychotic medications in the management of disruptive behaviors in children: safety guidelines and recommendations. *Clin Psychol Rev*. 2011;31(3):465–471.

Osterman JE, Barbiaz J, Johnson P. Emergency psychiatry: emergency interventions for rape victims. *Psychiatr Serv*. 2001;52:733–734, 740.

Pechnick RN, Ungerleider JT. Hallucinogens. In: Galanter M, Kleber HD, eds. *Textbook of Substance Abuse Treatment*. 3rd ed. Washington, DC: American Psychiatric Publishing; 2004.

Petit JR. *Handbook of Emergency Psychiatry*. Philadelphia, PA: Lippincott Williams & Wilkins; 2004.

Petit JR. Management of the acutely violent patient. *Psychiatr Clin North Am*. 2005;28:701–711.

Rudolph MN, Hughes DH. Emergency psychiatry: emergency assessments of domestic violence, sexual dangerousness, and elder and child abuse. *Psychiatr Serv*. 2001;52:281–282.

Schnoll SH, Weaver MF. Phencyclidine and ketamine. In: Galanter M, Kleber HD, eds. *Textbook of Substance Abuse Treatment*. 3rd ed. Washington, DC: American Psychiatric Publishing; 2004.

Simpson SA, Joesch JM, West II, Pasic J. Risk for physical restraint or seclusion in the psychiatric emergency service (PES). *Gen Hosp Psychiatry*. 2014;36(1):113–118.

Slaby AE. Crisis-oriented therapy. In: Lipton FR, Goldfinger SM, eds. *Emergency Psychiatry at the Crossroads: New Directions for Mental Health Services*. San Francisco: Jossey-Bass; 1985.

Slaby AE. Definitions and conceptual framework of psychiatric emergencies. *Emerg Health Serv Rev*. 1985 Winter-1986 Spring;3:9–24.

Stone AL, Becker LG, Huber AM, Catalano RF. Review of risk and protective factors of substance use and problem use in emerging adulthood. *Addict Behav*. 2012;37(7):747–775.

Tardiff K. *Medical Management of the Violent Patient: Clinical Assessment and Therapy*. New York: Marcel Dekker; 1999.

Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet*. 2013;382(9904):1575–1586.

Wilson MP, Pepper D, Currier GW, Holloman GH Jr, Feifel D. The psychopharmacology of

agitation: consensus statement of the American association for emergency psychiatry project BETA Psychopharmacology Workgroup. *West J Emerg Med*. 2012;13(1):26–34.

Work Group on Psychiatric Evaluation, American Psychiatric Association Steering Committee on Practice Guidelines. Practice guideline for the psychiatric evaluation of adults. Second edition. American Psychiatric Association. *Am J Psychiatry*. 2006;163(suppl 6):3.

Zun LS. Evidence-based evaluation of psychiatric patients. *J Emerg Med*. 2005;28:35–39.

33 Psychotherapies

▲ 33.1 Psychoanalysis and Psychoanalytic Psychotherapy

SYLVIA R. KARASU, M.D.

> For one who'd lived among enemies so long;
> If often he was wrong and, at times, absurd, to us he is no more a person now but a whole climate of opinion
>
> —W.H. Auden, *In Memory of Sigmund Freud*

"Psychoanalysis is really a method of treatment like others. It has its triumphs and its defeats, its difficulties, its indications . . . its therapeutic successes give grounds for neither boasting nor being ashamed."—Freud, 1933, *Introductory Lectures*

INTRODUCTION AND DEFINITION

"Psychotherapy," wrote Freud, in his 1905 paper of that name, "is in no way a modern method of treatment. On the contrary, it is the most ancient form of therapy in medicine." He went on to explain that there is an old saying that certain diseases are cured not by medication, but by the "mental influence" of the "personality of the physician." No one, of course, can really trace the origins of psychotherapy, but they probably extend back to primitive medicine men, priests, and soothsayers. According to the *Oxford English Dictionary,* the earliest use of the term "psychotherapeia," meaning "remedial influence of the mind," appeared in the 1850s. By 1897, in a textbook of mental disease, the term "psychotherapy" was defined as "every means and every possible agency which primarily affects the psychical rather than the physical organization of the patient in a curative direction."

Psychoanalysis, of course, very clearly originated with Sigmund Freud, although Freud gave credit to his medical colleague, Josef Breuer, in connection with the first published appearance of the word "psycho-analysis" in 1896: Freud spoke of a "new method of psycho-analysis, Josef Breuer's exploratory procedure" that he was now employing to trace hysterical symptoms back to their origins. James Strachey, Freud's official translator for the so-called *Standard Edition* of Freud's writings, explains that Freud actually used the terms "analysis," "psychical analysis,"

"psychological analysis," and "hypnotic analysis" in an earlier paper, "The Neuro-Psychoses of Defense" in 1894. Freud chose the word "analysis" to suggest an analogy to work done by chemists, who isolate fundamental substances, that is, chemical elements, in their laboratories. Freud saw his work with patients as "analyzing," in that he separated out and isolated a patient's mental processes into their elementary constituents. To Freud, a patient's symptoms were "highly complicated mental formations" that required analysis.

What has been confusing is that Freud used the words "psychoanalysis" and "psychotherapy" interchangeably for many years, and it was only considerably later that he and others sought to distinguish psychoanalysis from other psychotherapies. In an early paper, *On Psychotherapy,* for example, Freud wrote that psychotherapy seems "positively unscientific and unworthy of a serious investigator" but asked his reader to allow him to defend it: "There are many ways and means of practicing psychotherapy, to the method Breuer called 'cathartic,' but which I prefer to call 'analytic.'" By 1933, however, Freud, in a series of introductory lectures, wrote ". . . psycho-analysis possesses a doctor either entirely or not at all. Those psychotherapists who make use of analysis among other methods, occasionally, do not to my knowledge stand on firm analytic ground; they have not accepted the whole of analysis but have watered it down." Ironically, over time, with changing patient populations and the realities of 21st century life, psychoanalysis and psychoanalytic psychotherapy have sometimes become harder to distinguish.

As defined, psychoanalysis traditionally has three distinct meanings. Freud was trained as a physician. His laboratory was his consulting room, and his theories came directly from his observations of his patients. So its first meaning is that it is a method of investigating the mind. It is also a theory that explains both normal and abnormal human behavior. Third, it is a specific treatment method designed initially to treat neuroses, such as hysteria and obsessive-compulsive symptoms, but later used for many other disorders.

Significantly, Freud's therapeutic goals were limited. For example, in an early paper on hysteria, he wrote "much will be gained if we succeed in transforming your hysterical misery into common unhappiness." Near the end of his life, in his classic 1937 paper, *Analysis Terminable and Interminable,* Freud described psychoanalytic therapy as a "time-consuming business" whereby "even successful analytic treatment does not protect the patient, who at the time had been cured, from falling ill later on of another neurosis" or even from a recurrence of the original neurosis. It was in that paper that Freud spoke of analysis as being one of three (the others being government and education) "impossible professions in which one can be sure beforehand of achieving unsatisfying results." In fact, to Freud, analysis "is always right in theory but not always right in practice."

Freud, though, wrote in his 1927 paper, *The Question of Lay Analysis,* that he did not want psychoanalysis to be a "handmaiden to psychiatry": "We do not consider it at all desirable for psychoanalysis to be swallowed up by medicine and to find its last resting place in a textbook of psychiatry under the heading 'Methods of Treatment.'"

The current practice of psychoanalysis and psychoanalytic psychotherapy represents a century of "revolution and evolution" of radical revision and gradual refinement by both its creator, Sigmund Freud, and by his followers and detractors. A wealth and diversity of psychoanalytic thinking, both theoretical and technical, have produced a variety of clinical modifications that are increasingly applicable to a broad spectrum of patients and problems. The psychoanalyst Leo Rangell called it "kaleidoscopic diversity." Technical parameters now range from the vestiges of orthodox long-term psychoanalysis to an array of shorter-term, neo-Freudian and non-Freudian psychodynamic treatments, some with a fixed time limit. The problems currently treated extend from classical symptomatic neuroses to character pathology, especially borderline and narcissistic disorders of the self. This in turn means addressing very early developmental deficits, as well as subsequent intrapsychic conflicts that stem from childhood. Similarly, since the 1980s, the predominant theoretical model of (one-person) ego psychology has been supplanted, and at times replaced, by object relations and self-psychology, models primarily developed by British analysts, and ultimately introduced into the United States by Otto Kernberg and Heinz Kohut, respectively. These and other new directions in (two-person) relational and intersubjective perspectives have major implications for both theory and practice. Today, as a consequence of the many changes since Freud, there is greater diversity, as well as increased integration of multiple perspectives and techniques—exploratory (also called expressive) and supportive, brief and long-term, intrapsychic and intersubjective—and the incorporation of old and new therapeutic elements from the varied fields of linguistics, hermeneutics, philosophy, and religion. As MacKinnon, Michels, and Buckley emphasized, however, as long as the interview is the major tool in psychiatry, a psychoanalytic (i.e., psychodynamic) understanding of a patient will remain essential. One of the early Freudian theoreticians, Otto Fenichel, wrote, "There are many ways to treat neuroses, but there is only one way to understand them."

HISTORY

Pre-Freud: 18th and Early 19th Centuries

Eighteenth-century mesmerism, originated by Franz Anton Mesmer (1734 to 1815), sowed the earliest seed of psychoanalytic practice, even though neither Mesmer's theoretical premises nor his actual practices are accepted

by psychoanalysts today. Using a procedure referred to as the magnetic pass, Mesmer and his patient would sit face to face with knees touching, and the charismatic Mesmer would cast a spell by making sweeping movements across the patient's body. In successful cases, a "magnetic crisis" ensued: After a hysterical seizure, the patient's symptoms would disappear.

Nineteenth-century followers of Mesmer remained within the same basic framework, retaining both the general theory of magnetic fluid and the patient's acting as a passive recipient of the therapist's active maneuvers. Changes in theory were minor until James Braid originated the name "hypnotism" to describe the earlier process, attributing its effectiveness not to physiological factors (i.e., a particular "fluid"), but to psychological ones (i.e., what the therapist expected the patient to experience). Braid thus laid the foundation for the modern concept of suggestion, on which Sigmund Freud based his initial practices, though he would ultimately reject them.

Freud: Late 19th and Early 20th Centuries

Hypnosis. The emergence of psychoanalysis as created by Freud in the late 1800s was a direct outgrowth of the treatment of neurotic patients at that time by physicians in general practice—the use of hypnosis for hysteria. Freud's friend and older colleague, the Viennese internist Josef Breuer, set the stage in 1881 with his serendipitous discovery that a patient's symptoms disappeared as a consequence of their verbal expression under a hypnotic trance. This was the famous case of Fraulein Anna O., who resides in the annals of psychoanalysis for having given it its impetus as well as its lasting name, the "talking cure," or, as she jokingly referred to it, "chimney sweeping." When Freud heard of the success of the technique in 1882, he was just completing his medical training with an assignment at Vienna's Allgemeine Krankenhaus, under the direction of Theodor Meynert. Struck by Breuer's compelling account, he spent the year from 1885 to 1886 studying hypnosis at the Salpêtrière in Paris with the French neurologist Jean-Martin Charcot. The latter's direct demonstration that with hypnosis he could both elicit and abolish hysterical symptoms so captured Freud's imagination that it became a major turning point in his thinking. In his practice in the late 1880s, Freud used hypnosis to enable patients to suppress or deny their symptoms while under its influence. More important to Freud's subsequent thinking, however, was the use of hypnosis in retracing the history of symptoms.

Cathartic Method. When Freud returned to Vienna and subsequently established his medical practice, he had no success in treating hysterical patients with the more typical methods of the day, such as hydrotherapy, electrotherapy, and rest cures. He then decided to draw on the profound

influences of Breuer and Charcot. The hypnotic method at that time was used strictly for psychic catharsis, attempting to gain access to patients' reminiscences that they could not remember under any other conditions. The particular method was a form of cathartic therapy, borrowed from Breuer.

In their collaborative work on hysteria in the mid-1890s (*Studies in Hysteria*), Breuer and Freud discussed their theory that the symptoms represented repressed ideas that had not reached consciousness and were a transformation or "conversion." They suggested that hypnotic treatment was effective because of the discharge of affect previously attached to the repressed mental acts, called "abreaction." A major requisite for the cathartic method, which was contingent on that theoretical position, was that the patient could be hypnotized and thus be enabled to revert to the psychological state in which his or her morbid symptoms had appeared for the first time. Their clinical experience and observations suggested to them that the patient's memories, thoughts, and impulses that were no longer conscious could enter the mind under hypnosis, and that when recounted to the physician, especially under circumstances of intense emotion, the symptoms would disappear automatically. It was consistent with the theory that hysteria was the result of traumatic experiences, usually sexual in nature, whereby large quantities of affect that had been barred from consciousness appeared in disguised form as symptoms. Freud poignantly noted the "hysterics suffer from reminiscences."

Following Breuer's precedent with Anna O., Freud used the cathartic method (in conjunction with hypnosis) in 1889 to locate the traumatic origin of the hysterical symptomatology of Frau Emmy von N. In describing the case, Freud anticipated his later thesis, which speculated that inhibited sexuality played a role in the etiology of her illness. Freud grew dissatisfied with hypnosis because he was not able to hypnotize all his patients who suffered from hysteria. He was also frustrated by the transitory effects of hypnosis that seemed to last only as long as the patient's contact with the physician. He began to suspect that the therapeutic effects depended more on the relationship between patient and doctor than on the technique per se. His suspicions were prompted by a striking incident in which one of his female patients threw her arms around him on waking from a hypnotic trance. (Breuer had earlier reported a similar incident.) As a result, Freud turned his thinking in the direction of the physician–patient relationship itself and, with that, the origins of his eventual discovery of transference.

Concentration Method. It was not until his use of hypnosis failed in the case of Fräulein von R. in 1892 that Freud felt obliged to experiment with modifications of the original hypnotic–cathartic method. The physician Hippolyte Bernheim of the Nancy School, an adversary of Charcot, influenced the technical direction that Freud took. Freud's decision to give

up the hypnotic–cathartic method was largely based on Bernheim's observation that the events recalled during hypnosis (in a hypnoid state) could be re-evoked by patients in states of consciousness when prompted by leading questions and urgings of the physician. Basing his concentration method on that premise, Freud developed a strategy in which he requested the patient to lie on a couch with eyes closed, then to concentrate on a particular symptom and attempt to recall past memories; Freud would place his hands on the patient's forehead as a physical way of facilitating the memory process.

The Early Psychoanalytic Method

FREE ASSOCIATION. Through his work with the same recalcitrant patient Fräulein von R., Freud discovered the method that is identified with the true creation of psychoanalysis—free association. Freud soon realized that his interruptions of the patient and his constant prodding for additional details hindered the spontaneous flow of the patient's thought. A simple remark by Fräulein von R., namely that she did not know what to say because she was not sure what Freud wanted to hear, sparked his realization of the undesirable effects of censorship and direction by the analyst. He began to instruct his patients to express even the most irrelevant, unpleasant, or trivial material. The method of free association is thus the gradual consequence of experimentation with technique—from hypnosis and the cathartic method, to a concentration method, to an approach almost antithetical to the original suggestive methods. As a new technique, free association was (1) unlike hypnosis, which required patient suggestibility; (2) unlike catharsis, which focused on the release of affects only; and (3) unlike the concentration method, which prodded, directed, and interrupted the patient.

TRANSFERENCE AND RESISTANCE. In altering his therapeutic technique, Freud realized that a "different though not contradictory conception of the therapeutic process" was warranted. With the new emphasis on the removal of amnesia and the recovery of repressed memories, he observed that the same instinctive forces that brought about repression to obliterate pathogenic material from consciousness continued to exert force—resistance—against full disclosure through free association. Moreover, unexpected feelings toward the analyst—transference—provided additional resistance to treatment. The persistence of repression with free association brought to the fore those two concepts, which remain the critical constructs of modern psychoanalysis.

Freud's first reference to transference in the psychoanalytic sense came in the mid-1890s in *Studies on Hysteria*. He noted that some patients consistently tended to bring their intense personal feelings toward the analyst into treatment, transferring to the physician disturbing ideas that

arose from their past and that recapitulated earlier ties with significant persons from their childhood, thereby making what he called "a false connection" onto the analyst. In the subsequent case of Dora, in 1905, hostile feelings resulted in her breaking with Freud. Instead of continuing to analyze her feelings, Dora chose to leave her analysis precipitously, that is, act out her feelings by her action of leaving. Disturbed by her premature termination, Freud recognized the need to analyze resistance and transference manifestations.

Just as Fräulein von R. had been the pivotal failure with hypnosis that led to Freud's major shift away from its use, so the case of Dora became the turning point in psychoanalytic technique by highlighting for Freud the necessity to interpret feelings transferred onto the analyst. Again it was a therapeutic failure that instigated the change in approach.

Freud's Psychoanalytic Technique and "Classical" Psychoanalysis. In their book *Unorthodox Freud,* Lohser and Newton note that Freud wrote comparatively few papers about the technique of psychoanalysis: of 300 scientific publications, only 20 papers from 1895 to 1937 focused on technique. In his paper *Wild Analysis* (1910), Freud emphasizes the importance of undergoing psychoanalysis rather than just learning about it, "If knowledge about the unconscious were as important . . . listening to lectures or reading books would be enough to cure him. Such measures have as much influence on symptoms of nervous illness as a distribution of menu-cards in time of famine has upon hunger." In his paper, *On Beginning Treatment* (1913), Freud suggests, in explaining free association, that his analysand think of himself "as a traveler on a train sitting at a window" who describes everything he sees outside to his fellow travelers who do not share the same view.

Lohser and Newton speculate that Freud did not want his patients to know too much about his method, did not want his opponents to have "further ammunition" against psychoanalysis, or "didn't want to give his followers guidelines that would be incomplete and imprecise." Freud never wrote a treatment manual per se, and much of what he actually did with patients comes from his own case descriptions as well as recollections from some of his former patients who chose to write about and share their own analytic experiences. Samuel D. Lipton notes that even in Freud's extensive case history of the "Rat Man," Freud did not seem to write the case for the purposes of elucidating technique and his "occasional comments on technique seem only incidental." Freud's technique, as described either by him or others, was hardly the classic or orthodox technique that evolved over time, particularly in the United States (and most evident in the technique of analysts Kurt Eissler or Charles Brenner). For example, Freud reported that when the Rat Man was hungry, he fed him herring, something a "classic" analyst would find inappropriate. Freud told the

“Wolf Man” at one point that he had “first class intelligence” and suggested to him that he give Freud a gift at the end of his analysis. Reportedly the “Wolf Man” was pleased to see years later that the antique statuette he had given Freud was photographed among Freud’s collection.

Freud’s technique as described in the “Rat Man” case, says Lipton, was not criticized until about 40 years after Freud had published it. Says Lipton, “The criticisms of Freud’s technique show that modern technique places great weight on minor matters—trivial matters like a meal (or sending a) postcard . . . ” In other words, it had moved from judging an analyst’s technique on its purpose to one where his or her technique is judged according to the behavior of the analyst. Lohser and Newton note that what has come to be called “classical” or “orthodox” analytic technique and attributed to Freud was “actually a post-World War II invention.” Lipton speculates that modern technique may have developed as a reaction against Franz Alexander’s technique of the corrective emotional experience (i.e., emphasizing experience rather than insight and “manipulating” the transference rather than analyzing it)—all principles that classical analysts saw undermined psychoanalytic theory. As a result, though, so-called classical or orthodox technique resulted in “a panoply of behavioral rules that tend to be patronizing to the analyst and infantilizing to the patient,” says Lipton. Even Eissler’s classic paper on parameters had “important though unintended effects” because “parameter” developed “a pejorative connotation” and a “myth” arose that a classical analyst should do nothing other than interpret.

Many analysts have documented major discrepancies between Freud’s actual technique and his recommendations to analysts. Some examples from the recollections of patients include Smiley Blanton’s memory that after he had told Freud he was saving up his money to buy Freud’s works, the next day Freud presented him with a gift of his collected papers. Apparently afterwards, Blanton was unable to associate to his dreams and Freud acknowledged that giving gifts can be problematic. Hilda Doolittle’s recollected that Freud took her to his alcove where he housed many of his antiquity collection and showed her certain pieces, an act she later described as “extremely gratifying,” as were their “animated conversations” about ancient civilizations. Freud also seemed to be quite arbitrary regarding ending analysis. For example, analyst Abram Kardiner described being “shocked” when Freud suddenly announced to him it was time to prepare for termination after only 5 months of treatment. Freud also invited Kardiner to visit his son-in-law in Hamburg on his way to return to the United States and even signed a photograph for Kardiner in their last session. Joseph Wortis was told by Freud that Wortis was “too bourgeois” and told to give up his “narcissistic attitude” and be more receptive to what Freud said. Alix Strachey describes how Freud lit a cigar and said “Such insights need celebrating” after he had made an interpretation to a

particularly obscure dream. There are many other examples of extraordinary differences from classical technique, but clearly Freud's clinical method was evolving with experience (Fig. 33.1–1).

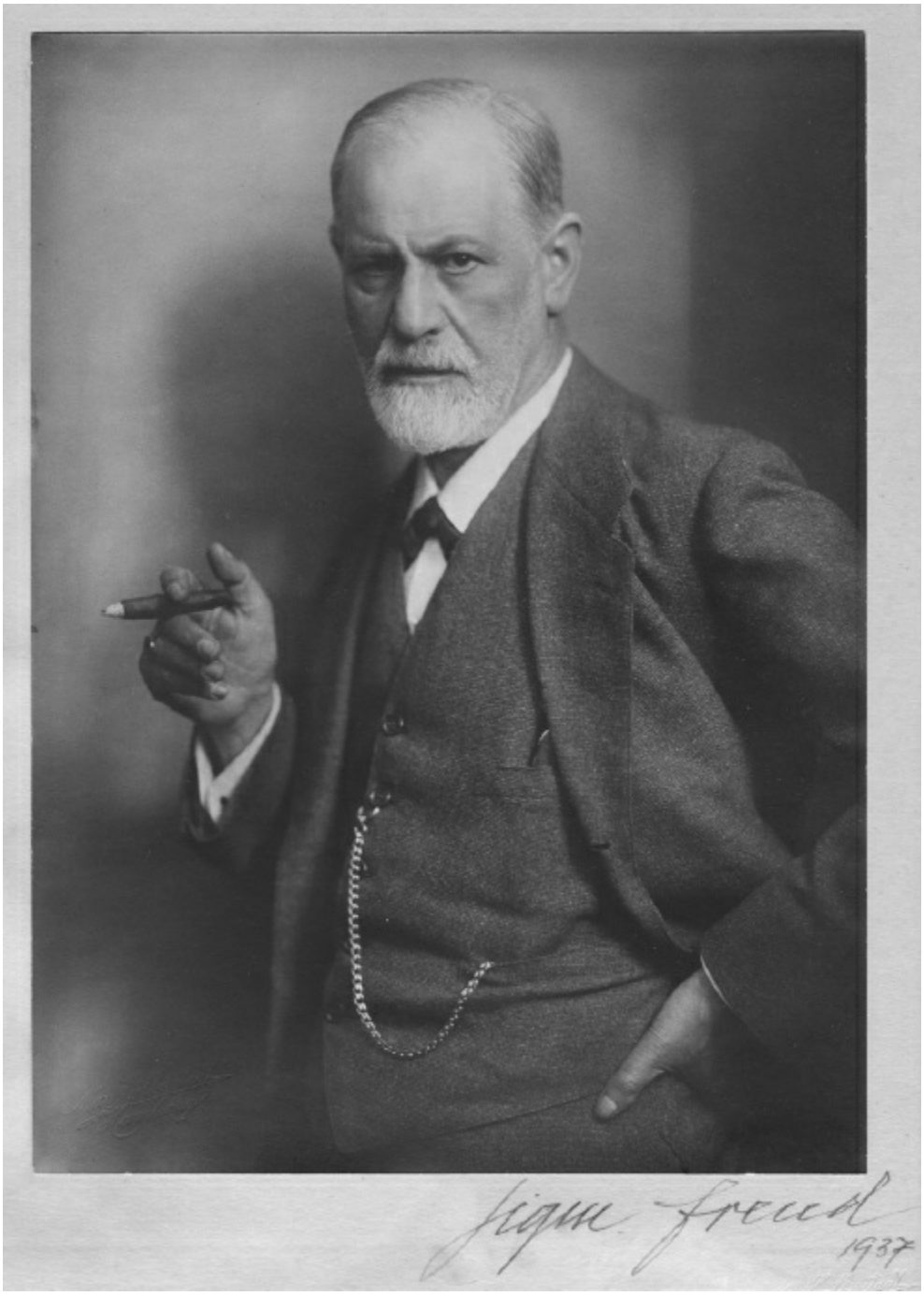

FIGURE 33.1–1. Photo of Sigmund Freud, 1937. (Courtesy of the Oskar Diethelm Library, DeWitt Wallace Institute for the History of Psychiatry, Weill Cornell Medicine, New York.)

Freud and Post-Freud: Early 20th Century

The history and growth of the psychoanalytic movement have been marked by repeated revisions by Freud, as well as by reappraisals and rebellions by

neo-Freudian and non-Freudian contemporaries and descendants, as chronically manifest in internecine conflicts over diverse and intricate points of theory and practice (Table 33.1–1). From the beginning, some analysts have argued against nearly all of Freud's basic conceptual premises, from the strictly sexual etiology of the neuroses to his views on female psychology. Others have reacted more directly to their clinical concerns that too few patients are amenable to the rigorous requirements associated with orthodox psychoanalysis, and they have attempted to make the treatment more widely applicable, affordable, and terminable.

George Makari, in his 2008 book, *Revolution in Mind,* emphasized that Freud was "reasonably charitable in the face of dissenting views . . ." and was not rigidly dogmatic. Throughout the years, Freud often modified his theories and incorporated the ideas of others. But Makari also noted that "one could not—it seems—question Freud's clinical authority, his veracity as an empirical observer, or his therapeutic claims." For example, when Freud disagreed with the theorist Alfred Adler, ". . . Freud did not prove Adler was wrong; rather he declared Adler was not a psychoanalyst." Freud's later papers were preoccupied with the dynamics of transference and resistance, their positive and negative aspects, and warnings to analysts about how to behave in order to manage their inherent difficulties. Those papers, which spanned the years 1912 to 1917, addressed the foundation of psychoanalytic technique as practiced to the present time.

In the 1920s and 1930s such analysts as Sandor Ferenczi, Otto Rank, and Wilhelm Stekel tried to broaden the applicability of psychoanalysis to a larger clinical spectrum by shortening treatment time and emphasizing a more active, affective, caretaking approach. Stekel's Institut für Aktive Psychoanalyse was the first organized attempt at psychoanalytically-based brief psychotherapy.

Blurred Boundaries and Boundary Violations among the Early Psychoanalysts. Glen Gabbard notes that cases of blurred or overt boundary violations among the early proponents of psychoanalysis demonstrate how the concepts of transference and countertransference evolved. He believes these cases should be read and studied "to understand the fundamental vulnerabilities of the psychoanalytic situation." Freud demonstrated many instances of blurred boundaries. For example, he suggested that one analyst marry a wealthy heiress in order to arrange a large financial donation for psychoanalysis. Peter Gay, in his biography of Freud, wrote that Freud also apparently used the connections of a patient to procure his professorship that had eluded him for years. Freud also analyzed his own youngest child (of his six children), daughter Anna Freud, who later became a child psychoanalyst.

One of the most egregious violations occurred between Carl Jung and his patient Sabina Spielrein, his first analytic case. Gabbard notes that the

relationship between Jung and Spielrein had become "boundary-less," and Jung consulted with Freud to help him extricate himself from his tie to Spielrein. Another early follower, Sandor Ferenczi, also consulted with Freud when he found himself in a difficult and compromising situation when treating a mother and her daughter.

Salvador Dalí and Freud. Salvador Dalí (1904–1989) was one of the greatest painters of the 20th century and the painter most associated with Surrealism, a "revolutionary" art movement that grew out of political and social disillusionment of the World War I. Surrealistic artists, including Dalí, embraced Freud's theories and expressed representations of them in their art.

Table 33.1–1.
Historical Development of Psychoanalytic Psychotherapy

Theorist	Major Contributions to Psychoanalytic Psychotherapy
1920s–1940s	
Rank, Ferenczi	"Active technique"; nurturing approach with emphasis on the affective aspect of the analytic experience; "end-setting" time limit
Stekel	Institut für Activ Psychoanalyse; the first attempt at psychoanalytically based brief psychotherapy
Reich, Fenichel	Expanded analysis of resistances to pathological character traits and defenses, e.g., analysis of character "armor" (Reich)
Alexander, French	Established an innovative school; the role of the analyst is to provide "corrective emotional experience" by being different from parents and offering an alternative to early developmental experiences; technical modifications of short-term duration, emphasis on face-to-face interviews
Horney	Rejection of Freud's theory of neurosis with an interpersonal approach led to an extension of transference to current context, closed the distinction between analysis and therapy, a belief that free association and the couch were not essential, an active therapist role, and short-term goals
1950s–1960s	
Sullivan	Pioneering work with schizophrenics; developed the face-to-face "psychiatric interview" with the therapist as "participant-observer"; interpersonal extension of transference beyond parents to teachers and peers based on the notion of "parataxic distortion"
Eissler, Bibring, Stone, Gill	Expanded definition, indications, and scope ("parameters") of psycho-analysis (Eissler, Bibring, Stone), including a broadened analysis of transference to the "here-and-now" (Gill)
M. Klein, Winnicott, Bion, Balint	Influence of British object relations school on psychoanalytic theory and technique (M. Klein); concepts of a "holding" environment, "good-enough mothering (Winnicott), the therapist as "container" (Bion), and "healing the basic fault" (Balint) addressed patients without adequate mothering in the early months of life
Zetzel, Greenson, Boss	Extension of therapeutic relationship to nontransferential aspects; concepts of "therapeutic alliance" (Zetzel) and "working alliance" (Greenson) advanced the idea that analysis incorporate aspects of reality or the "real" relationship into treatment, using both observing and experiencing ego of the patient; existential emphasis on the genuine human encounter between analyst and patient (Boss)
Hartmann,	Advanced ego psychology in adaptation; postulated "conflict-free sphere" of ego and

Kris	defined a concept of the self as a separate structure (Hartmann); concept of "regression in the service of the ego" emphasized the ego's participation in the analytic process (Kris)
A. Freud	Application of ego psychology to psychoanalytic treatment and child analysis, with emphasis on the adaptive function of defense mechanisms
Jacobson	Developmental model as a basis of a comprehensive psychoanalytic theory, integrating ego psychology, object relations, and drive theories
Mahler	Direct observation and research on infants and mothers; delineated "separation–individuation" subphases of child development, with effect on adult personality and pathology and implications for analytic treatment process
Lacan	Joint function of therapist and patient "is not to inform but to evoke"; therapist's unconscious is able to reconstruct/interpret unconscious of the patient; therapist remains quiet so that "Other" can speak; varying the length of the sessions—"the cutting up of the timing"
1970s–1980s	
Sifneos, Davanloo, Malan, Mann	Development of different types of brief psychoanalytically oriented psychotherapy; use of active, confrontative; and supportive techniques, and reduced or set time limits
Kohut	Development of self-theory in analytic treatment of narcissistic disorders; delineation of "self-object" (mirroring and idealizing) transferences; emphasis on empathic atmosphere to facilitate insight and to crystallize the self
Kernberg	Extension of object relations theory to psychoanalytic psychotherapy techniques in the treatment of borderline disorders; delineation of "expressive techniques alone" for analyzing primitive transferences and defenses
Strupp and Binder, Luborsky	Operational manuals for the clinical application of dynamic psycho-therapy (Strupp and Binder) and psychoanalytic supportive-expressive treatment (Luborsky)
1990s–2000	
Reformulation and integration of psychoanalytic theories and unifying conceptual schemas for treating patients across a broad diagnostic (neuroses, borderline, narcissistic disorders) and developmental (oedipal, preoedipal) spectrum	
Schafer, Chessick, Yalom, Karasu	Application and blend of philosophy, religion, and the humanities, such as existentialism (Yalom), hermeneutics (Chessick), language (Schafer), and spirituality and "the psychotherapist as healer" (Karasu), to a new understanding of patients in psychoanalysis and psychotherapy
Stolorow, Mitchell, Ogden, G. Klein, Sandler	Focus on intersubjective perspective (Stolorow) and relationality (Mitchell) in the interaction between analysand and analyst; both psychopathology and treatment viewed as a two-person event; greater attention to countertransferential phenomena (Ogden, G. Klein, Sandler)
Gill, Hoffman	Emphasis on analyst's actual behavior as continually affecting the patient's transferential feelings and fantasies; social constructionism (Hoffman): analyst and patient construct a mutually negotiated reality
Pine, Karasu, Stern	Developmental movement away from a single theory of the mind toward multiple perspectives of clinical phenomena, including drive, ego, object relations, and self-experience (Pine); generic conceptualization of practice based on dyadic and triadic deficits and conflicts (Karasu); psychoanalysis and developmental psychology combined in research studies of the earliest inner life of the infant and its effect on the mother–child relationship (Stern)
2000s and Beyond	
Rangell,	Analyst as "humanistic scientist" and "scientific humanist" transference-focused

Kernberg, and Clarkin	psychotherapy; focus on predominant affect-laden themes that develop between patient and therapist in the course of therapy
Milrod	Evaluating efficacy of psychodynamic psychotherapy with specific focus (e.g., panic disorder)
Kandel	Psychotherapy as scientific endeavor that produces actual changes in brain; the "biology of psychotherapy"
Fonagy, Bateman	Importance of mentalizing—"having the patient's mind in mind"; interpreting behavior as being motivated by mental states, both in themselves and others; based on attachment theory
Per Høglend	Research evaluation of transference
Young	Schema therapy: integrative psychotherapy combining cognitive, experiential, behavior, psychodynamic, and interpersonal techniques to achieve structural change; maladaptive schema result of individual's core needs not met in childhood

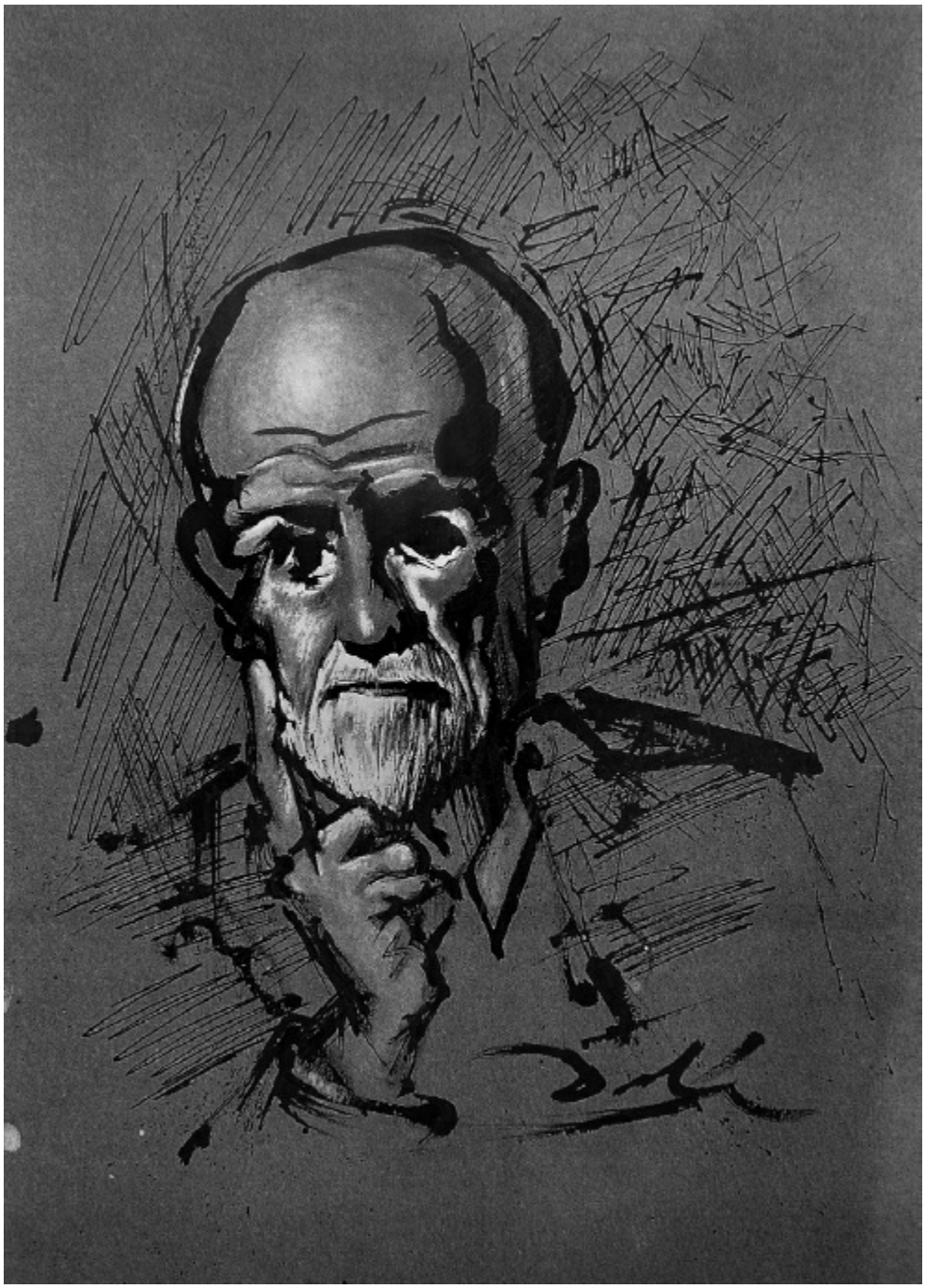

FIGURE 33.1–2. Salvador Dalí Portrait of Freud, 1938. (Courtesy of Copyright Salvador Dalí, Fundacio Gala-Salvador Dalí, Artists Rights Society (ARS), New York, 2014.)

According to sources, Dalí had read Freud's *Die Traumdeutung* (*The Interpretation of Dreams*) and was considerably influenced by it. Many of Dalí's paintings have titles evocative of psychoanalytic themes, such as *The Persistence of Memory, Le Grand Masturbateur, Freud's Perverse Polymorph (Bulgarian Child Eating a Rat)* and *The Enigma of Desire: My Mother*.

Dalí tried repeatedly to meet Freud and ultimately was able to arrange a meeting in July 1938, when 82-year-old Freud, then quite ill with the

cancer that would kill him the following year, had relocated to London. Reportedly, Freud had been somewhat contemptuous of the Surrealists. After the meeting, though, Freud wrote to one of his colleagues, "Up to now I have been inclined to consider Surrealists, who seem to have chosen me as their patron saint, as incurable nutcases. The young Spaniard, however, with his candid, fanatical eyes and unquestionable technical skill has made me reconsider my opinion . . ." Dalí later described the meeting, "Contrary to my hopes we spoke little, but we devoured each other with our eyes . . ." Dalí created several portraits of Freud in the late 1930s (see Fig. 33.1–2), and he made several etchings of Freud, including one of Freud (see Fig. 33.1–3) as part of his "Famous Men" series (from the late 1960s) that, among others, also featured John F. Kennedy and Albert Einstein.

Post-Freud: Mid-20th Century

Freud fled Vienna and the Nazis in 1938, reportedly only after his daughter Anna Freud had been questioned by the Gestapo. He died in exile in England the following year of cancer of the jaw that had plagued him for many years. Many other European analysts also fled the Continent and landed primarily in England and the United States. There were technical innovations in the late 1940s by the school of Franz Alexander and Thomas French, the first group to experiment directly with influencing the transference by varying the frequency of sessions and emphasizing face-to-face interviews. Alexander and French defined the major role of the analyst as providing a corrective emotional experience to compensate for pathogenic early developmental experiences by being different from the patient's parents. However, those therapies soon sparked a major controversy over whether the therapists could still be considered psychoanalysts. Threatened by the widening scope of psychoanalytic technique, orthodox therapists claimed that any departure from a strictly neutral-interpretive stance would compromise treatment.

FIGURE 33.1–3. Salvador Dalí Etching of Freud, 1968, part of "Famous Men" Series. (Courtesy of Copyright Salvador Dalí, Fundació Gala-Salvador Dalí, Artists Rights Society (ARS), New York, 2014.)

Although Freud, who had visited the United States only once, in 1909, thought of America as "an experiment that failed," the United States became the major center for psychoanalysis after World War II. There was widespread dissemination of psychodynamic concepts and applications, and the 1950s were the heyday of psychoanalysis in the United States. As the practice of psychoanalysis broadened to include a great variety of

clinical conditions, especially with the extension of classical analysis to psychoses and character disorders, analysts were obliged to alter both their strategies and the theories that they used to explain the changes in perspective. Initially, especially in the earlier era of postwar psychiatry, psychoanalysis was applied almost indiscriminately to a wide range of mental disorders reflecting every degree of severity of illness; during that experimental period analysis began more actively to recognize the limitations of standard methods.

Those critical years saw the expansion in interpersonal, cultural, and phenomenological directions of the psychoanalytic movement. European and American analytic theorists—including Alfred Adler, Otto Rank, Erich Fromm, Karen Horney, Harry Stack Sullivan, and Medard Boss—separately seceded from Freud (some to form their own schools) and sought sufficient theoretical changes to challenge and extend the early boundaries of psychoanalysis.

Horney's rejection of Freud's libido theory of neurosis in favor of a more social orientation led to an interpersonal approach in the analysis of neurotic patients that drew less distinction between analysis and psychotherapy, gave the analyst a more active role, and dispensed with free association and the couch. Sullivan's pioneering work with a population of schizophrenic patients drew attention to their distorted parataxic thinking (illogical causal connections between concurrent events) in adult interpersonal relationships, and he influenced treatment techniques by defining the face-to-face psychiatric interview that fostered the role of the analyst as participant-observer. Within the framework of existential analysis (e.g., Boss' *Psychoanalysis and Daseinsanalysis*), the analyst emphasized the real, here-and-now encounter. Others, who were engaged in the treatment of patients with character problems, began to acknowledge the increasing need to enlarge the scope of treatment by introducing so-called "parameters" that would enable such patients to be treated by analysis. "Parameter" was the word introduced by analyst Kurt Eissler in his classic 1953 paper. Parameters were deviations from what came to be called "classic" or "orthodox" technique, which focused almost exclusively on interpretation. Parameters, which could include advising a patient to perform a certain action or even "forcing him indirectly to do it" (e.g., advising a phobic patient to confront his phobia directly), were meant to facilitate an otherwise psychoanalytic process and enable the patient to gain insight. Eissler noted that the goal of interpretation is insight, and the "problem" for the analyst is "only when and what to interpret." In the "ideal case, the analyst's activity is limited to interpretation," says Eissler. He did acknowledge that the analyst did have "one tool which is indispensable": the use of questions, a tool that had "unfortunately" been "taken for granted." Eissler acknowledged, ". . . the proper use of interpretation is difficult and complicated." Introducing parameters into analytic treatment,

though, “contains dangers which must not be overlooked. Each parameter increases the possibility that the therapeutic process may be falsified . . .” He believed that the introduction of parameters affected the transference relationship to the analyst and that any parameter must ultimately be subject to interpretation. Eissler also cautioned, “There is a great temptation to cover up, by the introduction of parameters, one’s own inability to use properly the interpretative technique.”

In the 1950s and 1960s, new technical directions were influenced by the growing effects of ego psychology on the understanding of ego defects and associated problems. Wilhelm Reich’s analysis of character armor, as well as Otto Fenichel’s investigations of problems of psychoanalytic technique, expanded the analysis of resistances in determining pathological character traits. Analysts such as Merton Gill, Leo Stone, and Edward Bibring attempted to define and expand the horizons of psychoanalysis. Heinz Hartmann and his colleagues explored the technical implications of ego psychology. They postulated the existence of a conflict-free sphere of the ego and further defined the concept of the self as a separate structure within the ego. Anna Freud had set the stage for a focus on the ego by publishing her classic work, *The Ego and Mechanisms of Defense,* in 1936 as an 80th birthday present for her father. Leo Stone spoke of the “widening scope of psychoanalysis.” Edith Jacobson’s developmental model became a foundation of comprehensive psychoanalytic theory, which integrated drive theory with ego psychology and object relations theory. This model was supported by Margaret Mahler’s direct observational research in the 1950s and 1960s on infants and their mothers during the separation–individuation stages of child development. Mahler and her colleagues came to believe that biological birth is not concurrent with the psychological birth of an infant.

Mahler delineated the following stages as part of normal development: *Normal autistic* (birth to 4 weeks): A stage in which the infant is half-awake, half-asleep and whose main task is maintaining homeostatic equilibrium; *normal symbiotic* (4 weeks to 4 to 5 months): Dim awareness of the caretaker, but still a state of “undifferentiation”; and the *subphases of separation–individuation* (5 months to 36 months and beyond). There are four subphases, according to Mahler: (1) A differentiation subphase (5 to 10 months), which Mahler called “hatching from the autistic shell.” The characteristic anxiety here is stranger anxiety, which involves both curiosity and fear of strangers. (2) A practicing subphase (10 to 16 months), marked by upright locomotion and a new perspective on the world, with a mood of elation. The characteristic anxiety is separation anxiety. (3) A rapprochement subphase (16 to 24 months), in which the infant has become a toddler and more aware of separateness from mother. The mood of elation is dampened, and temper tantrums are common as the toddler wants to be soothed by mother and yet wants to be independent. (4) Object

constancy (24 to 36 months and beyond), a phase in which the child is able to cope with mother's absences and engage substitute caretakers (e.g., nursery school teachers). There is a gradual internalization of image of mother as reliable and stable.

Concurrently with Jacobson and Mahler, object relations theory came to the fore in Great Britain, in which internalized relationships and interactions between the infant and his or her caretakers replaced biological instincts as having a primary motivating power in the life of the child and later adult. The influence of the British object relations school and the theories of Melanie Klein, Wilfred Bion, Donald Winnicott, and Michael Balint led to modifications in technique to accommodate patients without adequate mothering in the early months of life, embodied in the concepts of therapy as a "holding environment," the therapist as container, and healing of the "basic fault."

Elizabeth Zetzel and Ralph Greenson wrote of the "real" relationship between analyst and patient by developing the notion of the therapeutic or working alliance, which emphasized the need to recognize nontransferential elements of the analyst–analysand relationship.

An evolution in psychoanalytic techniques paralleled the change in analytic focus from inhibitory traits to character pathology and then to a direct interest in the vicissitudes of defective egos. The analysis of character armor in psychoanalytic practice, an early effort of Reich, has become the basis for a major thrust in psychoanalytic psychotherapy—its application to characterological disorders.

The Translation of Freud's Collected Works. Freud had written primarily in his Viennese German, for which he won the Goethe Prize for Literature in 1930. Translations of Freud's work into other languages, including English, began during Freud's lifetime. Freud was aware of difficulties of rendering his new concepts into plain language. Near the end of his life, for example, in *Beyond the Pleasure Principle,* in discussing the concept of the death instinct, Freud wrote, "The deficiencies in our description would probably vanish if we were already in a position to replace the psychological terms by physiological and chemical ones."

Nevertheless, it was apparent to his disciples of the importance of a translation in spreading Freud's ideas, especially to England and America. Englishman James Strachey, a member of the celebrated Bloomsbury group, and his wife Alix had been patients of Freud when Freud asked (another example of a boundary violation) Strachey to translate one of his papers. So began in the 1920s James Strachey's life work as general editor and translator of all of Freud's collected works, with the help of Anna Freud, Alix Strachey, and Alan Tyson. The official invitation had come from English analyst (and future biographer of Freud) Ernest Jones, who wanted "to secure a definitive edition for generations to come; if it is done after our

time, it can never be done so well." This project grew into the 24 volumes of the *Standard Edition,* which became the Bible to generations of English-speaking analysts. It was published by Leonard Woolf (husband to Virginia Woolf) at the Hogarth Press. Ironically, though, Woolf never consulted Freud about his wife Virginia's severe depression, a depression that ultimately resulted in her suicide by drowning.

As English analyst Clifford Yorke said, ". . . it would be quite impossible to undertake any translation of Freud in English without attracting criticism," but despite that, Yorke added, ". . . the Standard Edition is one of the great publishing events of the century." Strachey's primary aim was to render Freud's meaning "with the greatest possible accuracy" but in the mold of "an English man of science." As a result, instead of the highly metaphorical (and sometimes tentative) language of Freud, Strachey used scientific-sounding, formal, or technical words derived from Greek and Latin. Most controversial was Strachey's acceptance of the words "ego," "id," and "superego" for Freud's simple "ich," "es," and "super-ich" that had been insisted upon by analysts Ernest Jones and Joan Riviere, who had earlier translated some of Freud's papers. There are, though, many other examples of controversial choices, such as use of the previously unknown word "cathexis," a quantum of psychic energy or affect, when Freud's original word ("besetzung") suggested an "engagement" or even "an occupation" such as might be done by an army, or for Freud's "trieb," literally "drive," Strachey inserted "instinct," a very different concept.

Criticisms of the Strachey's translation began almost immediately but gained momentum in the 1980s, after the death of Anna Freud. It was not until the early 1990s that there was a decision by the Institute of Psychoanalysis in London to revise the *Standard Edition* and analyst Mark Solms became its editor. Solms described the "legal minefield" regarding copyrights because of Freud's own "laissez-faire attitude" about copyrights and the political battles that followed. As a result, this project has taken more than 20 years and, as of this writing, has yet to be released. Solms notes that the revision will correct Strachey's errors and add extensive commentaries. Surprisingly, though, Solms has elected not to revise Strachey's translation "in the vain illusion of creating a better language than he did" and "to retain Strachey's vocabulary" though he has added "lengthy discussion of the technical terms."

Solms is not the only one who has now attempted another translation of Freud into English. Psychoanalyst Adam Phillips was approached by Penguin Publishing. His view of translation was completely different from both Strachey and Solms. Phillips thought that each book of his series should be translated by a different person, and there should not be any consensus regarding the use of technical terms. In the end, there were 16 volumes and 10 different translators, and each author could include a preface that explained the choices made, and the general editor "should be

available for consultation." Furthermore, Phillips' goal was not to replace Strachey but provide alternative—"coexistent"—versions that would not be necessarily comprehensive and not even include an index. Rather, these translations were meant to appeal to a "common reader."

The Sigmund Freud Archives. The Sigmund Freud Archives, Inc., according to its website, http://www.freudarchives.org, is an "entirely independent organization" that was founded in 1951 by psychoanalyst Kurt Eissler for the purpose of "collecting, conserving, collating, and making available within a single archive for scholarly use" the professional and personal papers, including correspondence and photos, of Freud. Most of the collection came from Anna Freud and other donors. Between 1952 and 2005, this collection of thousands of documents was given to the U.S. Library of Congress in Washington, D.C. where it remains today. Initially the Archives put restrictions on release of many of the papers until the year 2000 in order to protect the privacy of its subjects. More recently, the Archives has a "policy of de-restriction, except for ethical and legal constraints" (e.g., to preserve patient confidentiality), and about 98 percent of the collection is now available to scholars. Recent examples of "derestricted and released" documents include the Freud/Anna Freud Letters, the Freud/Minna Bernays Letters, and the complete Freud/Abraham Letters. Some documents continue to have stipulated time-release restrictions, as for example, the papers of psychoanalyst Princess Marie Bonaparte are not to be released until 2020, and interviews and recollections of analysts Felix and Helene Deutsch are not to be released until the year 2050.

Years ago, the Archives was the subject of controversy when Jeffrey M. Masson, originally appointed by Eissler as his successor, delivered a paper highly critical of Freud, and was dismissed as its director. Masson ultimately wrote a book, *The Assault on Truth: Freud's Suppression of the Seduction Theory*. Over the years, the organization has been run by its Board of Directors, including an Executive Director.

Late 20th Century

Several competing and often overlapping theoretical domains characterize psychoanalytic theory in the latter part of the 20th century, with important implications for revisions in analytic techniques and applications to patient populations other than those with neuroses. Extending beyond ego psychology, they include the major contributions of object relations theory, self-psychology, and the newest arrivals, the intersubjective and relational schools of thought.

One of the most controversial figures in psychoanalysis was French analyst Jacques Lacan, whose seminal work *Ecrits* was first published in the 1960s. One of his major contributions was Lacan's focus on language

—"the function of language is not to inform but to evoke," though according to Auchincloss and Samberg, Lacan incorporated philosophy, mathematics, and aesthetics into his theories. He believed the analyst's function was to reconstruct (i.e., interpret) the unconscious, which was structured much like a language, from what was communicated to him by the patient, rather than provide the patient with empathy or a substitute object. He emphasized that the analyst must remain quiet so that what he called the "Other" can speak. For Lacan, the "Other" was the "very foundation of intersubjectivity" and was what the patient considered strange or different from himself. Most controversial, though, was Lacan's technique of varying the length of his sessions and even shortening them to a few minutes—what he called the "cutting up of the timing."

Since the 1970s the changing analytic direction has been greatly influenced by the direct clinical study of narcissistic and borderline personality disorders, which provides the pivotal foundations for both the theory and practice of psychoanalytic psychotherapy today. The attention to turbulent transferences, disturbances in self-image and object relations, and such primitive defenses as projective identification and splitting has inspired a therapeutic concern with developing theories to understand the etiology of such disorders and techniques for managing them. Otto Kernberg's application of psychoanalytic object relations theory to the development of psychoanalytic techniques in the treatment of borderline personality-disordered patients and Heinz Kohut's psychology of the self in the understanding of narcissism remain at the center of current analytic practice, the latter sparking a virtual revolution in clinical thinking. Kohut's classic paper "The Two Analyses of Mr. Z" demonstrated Kohut's dramatic change of clinical perspective from that of the classical Freudian analyst, whose focus is primarily on the oedipal phase, to his later focus on preoedipal factors and the concept of the self. Kohut's subsequent delineation of "mirror" and "idealizing" transferences, which are pathognomonic for narcissistic personality disorders, enabled him to perceive meanings and their significance differently from his original understanding of Mr. Z. The major mutative change agents of psychoanalytic psychotherapy became less emphasis on insight per se and more on creating an empathic atmosphere.

The techniques of Kohut and Kernberg have often been contrasted in ways that support their respective theoretical tenets and have implications for the therapist. Thus, Kernberg's predominant technique of confronting the defense of the patient's grandiose self (as it presumably protects against hatred and envy of the therapist) means that the clinician may be the recipient of the oral aggression that is conceptually viewed as the core of severe personality disorder. In comparison, Kohut's major strategy of allowing the transference to unfold and to have the analyst become the empathic self-object that was presumably missing in the patient's past (so

that he can gradually explore the meaning of the patient's disappointment in the parent/therapist) supports his more benign theoretical contention that aggression is secondary to narcissistic psychopathology.

CONTEMPORARY TRENDS

Intersubjectivity and Relational Psychoanalysis

Post-Kohutian self-psychology has been marked by the rise of intersubjectivity, as posited by Stolorow and his associates. It suggests that psychoanalysis is the science of the intersubjectivity and focuses on the interplay of two differently organized subjective worlds of patient and analyst in which pathology is understood as a two-person event. How the patient's self-object needs are met within an "empathic-introspective" intersubjective field of treatment is thus believed to influence both the nature of the pathology and the therapeutic course and outcome. Whatever goes on in analysis is always viewed as a shared construction, based on the critical thesis that the fundamental operation of the psyche is based on its striving not for gratification of instinctual drives but for relational connections and communication.

Another proponent of the importance of intersubjectivity is Daniel Stern, whose direct observations of infants convinced him that from birth one's feelings, intentions, and thoughts are all shaped by those of others. Stern believed that "our mental life is co-created" by continuous dialogues with the minds of others, and that is what he meant by an "intersubjective matrix." He believed that a one-person psychology, with its focus on the intrapsychic, was no longer tenable.

Stern also emphasized that even though symbolic functioning (e.g., language) does not become available until the middle of the second year of life, this does not mean that the infant is not thinking. There is, in fact, a whole system of implicit thinking, which is not symbolic and not verbal, that operates in the infant before 18 months of age. Stern called this domain the "nonconscious," as distinguished from the dynamic unconscious or preconscious.

Furthermore, Stern found what he called *affect attunement* between mother and infant, which he defined as a form of selective imitation (and "the path to sharing inner feelings states") and which serves as a forecast for lifelong patterns of behavior. Significantly, Stern called into question Mahler's normal autistic and normal symbiotic stages and even the need for a separation–individual phase. He argued that infants are capable of self-other differentiation from birth, as evidenced by the fact, for example, that they imitate adults (e.g., stick out their tongues) and that they can infer intentions of others as well. Stern even questioned Freud's stages of infantile sexual development (i.e., oral, anal, and phallic-oedipal) and noted that 100 years after Freud's formulations, there is no confirming

evidence of these stages from systematic observations of infants. Stern reported experiments by Meltzoff and colleagues showing that when infants want to learn about inanimate objects, they mouth them, and when they want to learn about people, they imitate them.

Relational psychoanalysis emphasizes the intersubjective by suggesting that the clinical process in psychoanalysis has always been, first and foremost, relational, or, in Stephen Mitchell's words, "our relational embeddedness with others." Incorporating aspects of attachment theory, he further suggested that the analyst's interpretations of the patient's verbalizations are a form of interpersonal attachment experience. Ultimately, he saw relational psychoanalysis as a broad, multidimensional vision of human intersubjectivity.

Irwin Hoffman wrote, along the lines of a relational approach, of social constructionism, that is, how the analyst's actual behavior constantly affects the patient's transferential reactions. He and Merton Gill note that, unlike those who emphasize "distortion of reality as the hallmark of transference," they see "perceptions of the therapist's 'actual' behavior are implicated in the transference at every point." In other words, there are always aspects of the therapist's behavior that become incorporated into the patient's transference. Gabbard notes that the constructionist view is borrowed from philosophy, which emphasizes that reality, rather than just an objective set of facts "out there" is a construction influenced by the observer's own subjective reactions. In the context of a psychoanalytic office, then, analyst and patient construct "a mutually negotiated reality."

Conflict/Deficit Duality

Competing cross-currents of psychodynamic thought have given rise to two predominant paradigms of theory and therapy. There remains an unresolved debate over conflict and deficit models that represent the culmination of nearly a century of internecine questions over the etiology and evolution of individual development. In brief, the original *conflict model* of ego psychology primarily pivots on intrapsychic drives and ego defenses based on repressed memories and distortions of the past, whereas the *deficit model* places its crucial focus on the creation and destruction of the self-structure based on real experiences and traumas in primary relationships. The former theorizes intrapsychic clashes among unconscious forces of the mind, especially repressed sexual drives and the defensive and adaptive mechanisms that guard against them (i.e., Freud's formation of the neuroses). This conflict model proposes a primitive splitting between object love and hate in order to deal with unconscious aggression, oral envy, and rage (i.e., Kernberg's formation of borderline personality disorder). When treating severe character pathology with this model, analysts maintain that unconscious intrapsychic conflicts are not simply struggles between impulses and defenses but denote two opposing

units or sets of internalized object relations. Such psychopathology is thus composed of a defensive constellation of self and object representations directed against an opposite (and often dreaded) repressed self and object representation.

These conflict-based conceptualizations have been contrasted with a deficit model, which postulates a defective development of the self due to early environmental events, especially lack of affective attunement between the original caregiver and the infant. With a deficit model, the focus is no longer on frustrated and *repressed* libidinal fantasies of oedipal origin within a triangular struggle (i.e., father–mother–child triadic structure). Rather, the focus is now on traumatic failures of empathic responsiveness that occurred during the preoedipal period (i.e., mother–child dyadic structure). The *deprived* patient yearns for unmet maternal mirroring and merger. The conflict-deficit dichotomy has also been framed in yet other terms, representing a pervasive shift from an *intrapsychic* to an *intersubjective* perspective in theory and treatment.

These developmental distinctions have also had direct implications for psychotherapy practices as well as for the populations treated. They broadly coincide with the movement from "expressive" (i.e., exploratory) to "supportive" strategies, that is, from a therapist who interprets and confronts a more developmentally advanced erotic transference to a therapist who is responsive to a more primitive (i.e., narcissistic) transference. With this change, there is a shift from a belief in the curative power of "insight" to that of "empathy."

These pivotal positions have also been viewed as a means for depicting basic human motivation and the resultant ills that befall humanity. Two predominant types have been proposed, each evolving out of a different historical period within the temper of the respective times. According to Kohut, an earlier prototype, termed *guilty man,* embodies the search for the gratification of hidden and forbidden sexual longings, whereas a later prototype, termed *tragic man,* expresses the basic striving for cohesion of a fragmented or missing self. Erikson put it more pragmatically in clinical terms when he observed that the patient of today suffers most under the problem of what he may be or become, whereas the patient of earlier days suffered most under inhibitions that prevented him from being who he thought he knew he was.

Short-Term (Brief) Therapy

The emergence of a variety of short-term (brief) psychodynamic psychotherapies was also a hallmark of the 1970s and 1980s. They are now primary, preferred modalities in their own right. Like other forms of psychoanalytic psychotherapy, they meet several contemporary needs: Socioeconomic constraints that make protracted and costlier treatment less viable, the wish to serve larger numbers of patients, and the pressures for

accountability (which calls for systematic studies of psychotherapy outcomes based on specific formats and circumscribed time frames). As a further attraction, the earlier fears of symptom substitution as an inevitable consequence of short-term treatment have essentially been quelled, allowing the enduring strengths of the Freudian legacy to be retained as the foundation for newer, more expedient techniques.

Peter Sifneos' anxiety-provoking psychotherapy, David Malan's intensive brief psychotherapy, Habib Davanloo's broad-focus, short-term dynamic psychotherapy, James Mann's time-limited psychotherapy, and Lewis Wolberg's eclectic, short-term psychotherapy are examples of the range of brief treatments that derive from the psychoanalytic approach. Their combined interpretive and noninterpretive techniques, active approaches, and abbreviated time requirements are all accepted features of current psychoanalytic psychotherapy. Operational guides for psychoanalytic practice have been developed. Hans Strupp and Jeffrey Binder's time-limited dynamic psychotherapy (TLDP) and Lester Luborsky's principles of psychoanalytic psychotherapy, a manual for supportive-expressive treatment, are two examples of the systematic approach to psychodynamic treatment.

Toward an Integration of Psychoanalytic Models

Although some analysts have sharpened the schism between conflict and deficit (defect or deficiency) models, others have addressed the coexistence and potential integration of multiple perspectives (i.e., "pluralism"). For example, in considering structural versus object relations conflict, Theodore Dorpat delineates two models of the mind: A "higher" tripartite model of id, ego, and superego, and a "lower" object relations model that occurs prior to it; he regards them as two successive stages of individual development, which together comprise the psyche's internal architecture.

Others have suggested that looking at symptomatology alone may not reveal whether intrapsychic conflict (neurosis) or developmental arrest (self-object relationship disorder) is responsible. Stolorow and Lachmann, who began with conflict/deficit as a dual line of thinking, have elaborated on a complex confluence of subtle and often masked manifestations. Their clinical examination of developmental arrests has elucidated a crucial distinction between two types of mental activity: defenses that primarily function to ward off conflicts (within a well-established psychic structure) as compared to those manifestations that predate them (at a pre-stage of defensive development), which are vestiges of structural voids and defects. For example, detachment can be a result of either an early developmental failure of attachment or a later instance of defensive avoidance of interpersonal contact. Their natural overlap or interface is demonstrated by their simultaneous occurrence in actual psychic life. Trauma due to empathic failure by a parental figure does not exempt the child from

conflict-based ambivalence toward that significant other. Eagle suggests that the two sources of psychopathology are intertwined, insofar as developmental deficiencies may influence the capacity to deal with conflict and unresolved conflicts may trigger developmental regressions or arrests.

On a broader level, Fred Pine attempted a synthesis of clinical phenomena derived from four major sources: drive theory, ego psychology, object relations theory, and self-psychology. Pine offers a multimodal perspective, based on studies of child development and its effect on adult psychopathology. Each theory has something to offer the clinician. Drive psychology emphasizes the taming and gratifying of the drives; ego psychology emphasizes defenses, adaptation, and reality testing; object relations theory focuses on all relationships throughout life from infancy on; and self-psychology focuses on the task of forming a differentiated sense of self. Ultimately, Pine believes that all four psychologies share the assumptions of psychic determinism (i.e., all present events are shaped by the past), unconscious mental functioning, and primary process thinking (i.e., irrational and wishful thinking).

Similarly addressing the need to bridge the theoretical and the clinical, T. Byram Karasu integrated various constructs of past and present analytic thinking, especially the schism between deficit and conflict viewpoints. Karasu differentiates among defects, deficits, and conflicts: for example, if a person has a problem with his eyesight, he has a defect that might be correctable by eyeglasses but that no amount of driving lessons will alleviate; if the person, though, has good eyesight but does not know how to drive, this is a deficit that can be corrected by driving lessons. If, however, that person has good eyesight and knows how to drive, but has a fear of driving or is ambivalent and cannot make a decision whether to drive to visit his mother, he has a conflict that neither glasses nor any number of driving lessons will ameliorate. Psychotherapy, according to Karasu, deals with deficits and conflicts but cannot specifically correct defects.

Karasu's theory is a trans-theoretical model of psychotherapy in which patients are viewed not in terms of traditional classifications or standard diagnostic nomenclature, but rather according to major maturational configurations. Both deficits and conflicts having origins in different phases of development are represented, including the combined effect of the very early mother–child relationship and the subsequent father–mother–child relationship. He establishes a fourfold matrix in which patients typically fall primarily into a dyadic deficit, dyadic conflict, triadic deficit, or triadic conflict classification.

The *dyadic deficit* patient has experienced inadequate, absent, or inconsistent mothering and exhibits a helpless dependency. This patient may present with an inner deadness and chronic depression. He or she fears abandonment but sees therapists as interchangeable and may tend to deny the therapist's existence as a person. As a result, therapists may feel

disconnected or discouraged in working with these patients.

The *dyadic conflict* patient has experienced an intrusive, demanding, or controlling mother and has difficulties with autonomy and separation–individuation. This patient may be chronically angry, even enraged at an "unrewarding object world" and devaluing of the therapist and may precipitate the rejection that he or she fears. The therapist, in turn, may feel unappreciated and unjustly attacked but feels worse if this patient does drop out of therapy.

The *triadic deficit* patient has experienced absent, abusive, or inadequate fathering. As a result, this patient has defective superego formation, a kind of "ethical poverty" that leads to sociopathic behavior. The patient fears being punished and is less likely to experience guilt from his wrongdoing. He often resists the "patient role" and sees the therapist as a "dispensable expert" from whom he expects special treatment. There is often a hidden agenda in seeking therapy. The therapist may feel compelled to "perform" for these patients by displaying wit or knowledge or even turning therapy into an informal conversation.

Triadic conflict patients experience strong erotic, competitive, and aggressive wishes and/or fears. These patients easily engage the therapist and can be seductive, jealous, envious, and competitive, but they can be highly motivated and accept the patient role readily. The major danger for the therapist, however, is being responsive to this patient's seductive and aggressive play.

In other attempts to find a unifying conceptual schema, some theorists have used a human science framework that extends beyond metapsychology to philosophy, hermeneutics, language, and religion. Roy Schafer, for example, has reconceptualized psychoanalysis in terms of an "action language." He continues to search for the best terms in which to render the findings of psychoanalytic practice and to systematize them. In his post-Freudian conceptualization, he has developed a clinically useful alternative to traditional descriptions of the apparatus of the mind and its intrapsychic processes.

In a different vein, Richard Chessick also goes beyond psychology per se in questioning the implicit assumptions on which the understanding of a patient rests. By delving into the phenomenological and philosophical thinking of Jean-Paul Sartre, Jacques Lacan, Jean Bernard León Foucault, Martin Heidegger, Erich Fromm, and numerous others as their ideas affect psychotherapy and psychoanalysis, he emphasizes multideterminants in the formation of psychic structure. Chessick also considers what actually brings about cure in psychoanalytically-oriented treatment. He wrote of his experience with a patient who he had helped considerably over the years. Near the end of therapy, he asked this woman what interpretations had the most significant effect on her improvement. Her unexpected response, "You have kind eyes," led to his appreciation that the therapist's

interpretations are only partly responsible for therapeutic efficacy. The person of the therapist can have a substantial effect on a patient.

In another exploration of what is common to all patients, Irvin Yalom examined the relevance of an existential dilemma for both patients and psychotherapists. He stresses that all human beings look for meaning and certainty in a universe that has neither. He also emphasizes the participant role of the therapist in the lives of his or her patients—both clinician and patient are affected and changed by the mutual encounter. T. Byram Karasu, who emphasizes "the psychotherapist as healer," shares this existential domain. His 2013 book, *Life Witness: Evolution of the Psychotherapist,* goes even further in suggesting a "master therapist witnesses and experiences a patient's dilemma . . . and transcends theoretical dogma as much as theoretical atheism." Psychotherapy becomes a way of being. Says Karasu, "The tranquility of the therapist largely depends on the ability to live with optimum disillusionment and still maintain profound faith in all persons . . ." Karasu believes that psychotherapeutic theory is useful for enhancing a therapist's professional identity and bonding with his or her peers, but is ultimately no longer needed for a master therapist's self-congruency and self-grounding. Experienced therapists are both "more disciplined as well as more improvisational in their therapeutic repertoires." For him, psychotherapeutic change does not depend on the elucidation of "historical antecedents" but rather on the patient–therapist relationship. In other words, collecting biographical information on a patient can be a "form of objectification" of that patient and a type of intrusion—"psycho-peeping"—without regard to any consequences and ramifications for the patient. For Karasu, psychotherapy is not illness-specific but rather person-specific, and a therapist must seek for what is both "universal (shared) about a patient and what is unique (selective)." Diagnostic labels can become "procedural, bureaucratic, or even punitive."

Arnold Cooper described a "multiplicity of authoritarian orthodoxies, each derived from a particular thinker, rather than a scientific discourse." Whereas the boundaries of psychoanalysis were once sharp, they have become blurred, but "pluralism alone is chaos." Cooper delineates the fundamental controversies among psychoanalytic thinkers: relative importance of nature versus nurture; origins of neurosis (i.e., whether oedipal or preoedipal); the nature of man (i.e., whether tragic or guilty); the nature of inquiry (i.e., whether scientific or hermeneutic); the nature of analytic language (i.e., whether metapsychological or clinical); the origins of motivation; the intrinsic nature of the mind (i.e., whether unified or conflicted); the nature and source of therapeutic effectiveness and the limits of analytic technique. Cooper believes that it is crucial for the survival of psychoanalysis to abandon "our old isolation from each other," and though techniques change over time, "the essence of psychoanalysis

lies in the nature of its inquiry and its views of man, not in its technical procedures."

Another clinician who strives for unity among theoretical plurality is psychoanalyst Leo Rangell. For Rangell, the complete psychoanalyst is both "a humanistic scientist" and "a scientific humanist." Rangell raises the question of whether it is acceptable that a patient has an oedipal conflict when treated by one therapist and a problem with self-cohesion when treated by another. Rangell's encompassing theory would have room for drives and objects, the oedipal and preoedipal, constitutional as well as environmental inputs, the cognitive and affective, conflict and deficit, the intrapsychic and the interpersonal, actual traumas and fantasies, as well as both historical truth and narrative truth and the entire developmental cycle.

PSYCHOANALYSIS

Theoretical Constructs in Psychoanalysis: Metapsychology

Freud first used the word *metapsychology* in the late 1890s in letters to his friend Wilhelm Fliess, but did not elaborate on its meaning until his 1915 papers. Auchincloss and Samberg note that metapsychology has been defined as "the highest level of abstraction in a hierarchy of psychoanalytic conceptualization" wherein clinical observation is the most basic. Freud originally had three ways of conceptualizing mental functioning: topographic, dynamic, and economic. Later theorists added the genetic and structural "points of view." These theoretical concepts form a major foundation for psychoanalytic theory and practice and have been used to describe a patient's mental functioning. Over the years, some analysts, though, question the necessity of using this level of abstraction to understand a patient rather than relying on clinical observation.

Topographical. The topographical perspective refers to the premise that mental phenomena reveal themselves at different levels of manifestation, from the deepest recesses of the mind (unconscious), to the border of awareness (preconscious), and finally to the surface (conscious). That model laid the groundwork for the etiological phenomenon of repression, which describes the processes by which forbidden ideas and impulses (initially believed to be largely sexual in the narrower and wider senses of the word) are denied access to the conscious mind and whose reversal removes the original source of illness. Its general orientation recognizes the human being's pervasive avoidance of painful feelings or experiences by keeping unpleasant thoughts, wishes, and affect from awareness. It also recognizes the persistence, resilience, and inaccessibility of underlying conflicts that remain alive and active but may appear in diverse and disguised forms that are often unrecognized by their recipient.

Freud, before his later structural model that introduced the id, ego, and superego, believed the goal of analysis was to make conscious the unconscious.

Dynamic. The dynamic perspective reflects the notion that mental phenomena are the result of a continual interaction of forces that oppose one another, that is, with the conflict between instinctual drives and forces that restrain them. It implies that human behavior and motivation are active and changing at all times. It is the basis for such fundamental concepts as conflict, compromise formation, and resistance, which relate respectively to the foundations of mental illness as the consequence of competing demands and wishes, and the largely unwitting ways in which the patient fights getting well.

Economic. The economic perspective relates to how psychic energy is distributed, discharged, and transformed into symptoms, or displaced onto someone or something else. This perspective, according to Auchincloss and Samberg, is the most controversial. It proposes that mental events "are fueled by psychic energy, used to account for the relative intensities and peremptory force of wishes, feelings, and ideas, as well as for the power of defensive operations." It has implications for how ideas and affect are expressed (e.g., verbally or somatically).

Structural. The structural perspective refers to the idea that the mental apparatus is organized into functional units of a tripartite nature—id (instincts), ego (external reality), and superego (moral and ethical precepts). That basic personality organization forms the theoretical structure of intrapsychic conflict among three harsh masters. The ego serves the pivotal adaptive and executive function in mediating between the person's primitive instinctual drives and his or her internalized parental and social prohibitions. Once Freud had elaborated this structural model in 1926, he believed the major goal of analysis was "where id was, there shall ego be." Ironically, although Freud was originally trained as a neuroanatomist, he wrote in 1915, in his paper *The Unconscious,* "Every attempt to discover a localization of mental processes . . . has miscarried completely . . . Our psychical topography has for the present nothing to do with anatomy." Freud apparently never attempted to find neuroanatomical correlates for his ego, id, and superego within the brain, nor for his conscious, preconscious, or unconscious.

Significantly, in recent years, psychoanalyst Charles Brenner has proposed that there is no longer any need to divide the "mental apparatus"—Freud's term for the mind first introduced in his *Interpretation of Dreams*—into the structural components Freud eventually called *ego, id,* and *superego*. Brenner noted that even when

Freud conceptualized and focused on the structural theory rather than the topographical model of conscious, unconscious, and preconscious, he never abandoned the notion that the mind consists of separable systems or structures. He also notes that Freud never used the terms "topographical" or "structural." Rather than divide the mind into structures, though, Brenner believed the mind is better understood in terms of conflicts related to the sexual and aggressive wishes of childhood and the compromise formations that result between these conflicts and defenses against them. Brenner believed that all humans want to gain as much pleasure as possible by gratification of these sexual and aggressive wishes of childhood and simultaneously want to avoid as much unpleasure as possible. When a pleasure-seeking wish is associated with unpleasure, the mind, said Brenner, is in conflict, and this results in compromise formations in both normal and pathological conditions. What determines the difference is degree, namely, that a compromise formation that results in too much unpleasure—that is, too much anxiety or depression, or too much inhibition or self-destructive behaviors—is considered pathological.

Genetic. The genetic perspective focuses on ascertaining the historical roots of a patient's illness in childhood and has emphasized the importance of the oedipal stage, the source of the tabooed erotic bond between the child and his or her opposite-sexed parent. In more recent years, analysts have also recognized the importance of the pre-oedipal period, particularly in determining aspects of personality and personality disorders. The use of the word "genetic" is an unfortunate word choice as "genetic" in this context has nothing to do with genetics as it is used in common parlance today to indicate the study of heredity.

Psychoanalysis and Its Relationship to Hermeneutics

As a physician initially trained in neurology, Freud very clearly thought of himself as part of a scientific tradition. One of his first papers had been on aphasia, and in 1895 he wrote a paper, though published posthumously that he described as "a psychology for neurologists" but which was ultimately called *Project for a Scientific Psychology*. In a *Postscript* to *The Question of Lay Analysis* (1927), Freud writes, "In psychoanalysis there existed from the very first an inseparable bond between cure and research . . . It is impossible to treat a patient without learning something new . . . This prospect of scientific gain has been the proudest and happiest feature of analytic work."

Whether psychoanalysis is a science or a branch of the humanities, though, has been debated for years, both by analysts themselves and their critics. One particularly vociferous critic was the Austrian born British philosopher Karl Popper, who believed in the importance for a science of the doctrine of falsifiability—the use of empirical evidence to refute and

criticize, rather than look for certainty. He noted that those who followed Freud saw confirmation of their theories everywhere and this apparent strength was, in fact, a major weakness in the theory. Furthermore, Popper compared analysts' clinical observations to the pronouncements of astrologers and Freud's structural concept of the mind as no better than Homer's mythology.

Charles Brenner, on the other hand, strongly believed that psychoanalysis is a natural science based on clinical observation. He emphasized that it was the analyst's approach to clinical data, rather than any particular subject matter, that made it distinctive and scientific.

Between these two poles is Robert Wallerstein, who believes that psychoanalysis is a psychological science and not in the same category of the science of neurophysiology. Wallerstein distinguishes "search" from "research," which entails systematic, observable, and replicable scrutiny. Wallerstein noted that despite the fact that Freud thought of psychoanalysis as a science, he never required any specific formal systematic research that science requires. Arnold Cooper expressed the concern it seemed "an extraordinary admission of mistrust concerning the scientific power of his theories and the identity of our profession" when generations of analysts were unwilling to overturn any of Freud's ideas.

Those who question its scientific underpinnings view psychoanalysis as part of the humanities, that is, a so-called "hermeneutics discipline." Hermeneutics is the study of interpretation, from the Greek word "to interpret" and was employed to study (i.e., interpret) texts such as Homer or the Bible. An interpretation of the entire text depended on the interpretation of its parts. In that sense, the analyst's interpretations to a patient during a session are seen in a wider context of the patient's life as a whole and his or her cultural context, just as the interpretation of a passage of literature may take on a different meaning in the wider context of the entire text. Within psychoanalysis, hermeneutics signifies understanding a patient based on his or her "coherent narrative," rather than establishing specific scientific truth. Lawrence Friedman notes, "In every other science the object of study is selected by the investigator. In psychoanalysis, a science which is also a therapeutic discipline, the object of study—the entities within the mind—is selected by the subject." He also thought that a hermeneutic philosophy enables an analyst to "accept the impossibility of capturing the patient's unique meaning: understanding does not come from fitting the patient's productions into a stock formula . . ." Merton Gill believed the analytic process should be considered hermeneutic but the process itself could be studied scientifically, and as a result, he called psychoanalysis, "a hermeneutic science."

Fundamental Concepts of Treatment

Psychoanalytic practice is distinguished by the recognition of two integrally

related events of treatment: transference and resistance. Considered opposite sides of the same coin, they are inextricably linked insofar as transference can at any time be transformed into resistance and threaten to interrupt treatment. Composing the essential dynamics of the analytic situation, these dual forces of shifting valence are operating at all times as the patient continually struggles to be cured of his or her illness.

Freud initially believed that the primary purpose of psychoanalysis was to analyze transference and to circumvent resistance, but he later advocated the uncovering and repeated analysis of resistances in a systematic way. Rather than to address the actual content of the patient's verbalizations, the analyst has the primary objective to confront and understand the defensive maneuvers by which the patient delayed getting well. Only then could the deeper levels of neurotic conflict be exposed and ultimately resolved.

Transference. Transference is unequivocally the heart of psychoanalysis, and the analysis and specifically, the interpretation of transference is what distinguishes psychoanalysis from all other forms of psychotherapy. *Transference* is the displacement and reactivation of feelings originally experienced toward significant people (usually parents) in one's childhood onto the person of the analyst. Gabbard notes that patterns of transference appear automatically and unconsciously within the analytic relationship.

Lawrence Friedman speaks of the "paradox of transference"—it both helps and hinders the analyst. Since Freud's original serendipitous discovery of the "strange phenomenon" of transference in the psychoanalytic treatment of every neurotic patient, transference has been conceptualized in various related ways: (1) as a distinct type of therapeutic relationship, which has since been distinguished from nontransferential therapist–patient bonds (e.g., therapeutic alliance or real relationship); (2) as substitute pathology, expressed in the formation and resolution of a so-called transference neurosis; and (3) as a general phenomenon that transcends the boundaries of analysis and extends into all human relationships.

TRANSFERENCE AS A THERAPEUTIC RELATIONSHIP. The typical transference relationship is one in which the patient perceives the analyst in a specific although not implausible way because of a predisposition to view the analyst through the prism of earlier relationships. The feelings engendered may include an unusual degree of attachment and affection and even hostility and anger that can be traced to wishful unconscious fantasies from childhood. The analyst promotes their appearance through special methods and analytic rules (e.g., a state of abstinence and lack of directly gratifying the patient) that intensify reactivation. Transference as a unique unreal

interpersonal bond is a function both of the inherent nature of neuroses and of the design of the analytic situation that facilitates regression and may lead to a substitute neurosis, that is, the "transference neurosis."

The peculiarity of the transference relationship with the analyst lies in its excess, in both character and degree, over what is rational and justifiable. Its major manifestation is an over-endowment of the analyst as an idealized image, often including overestimation of the analyst's qualities, adoption of similar interests, and jealousy of other persons in the analyst's life. The analyst, becoming the predominant object of the patient's unresolved sexual (and aggressive) desires, arouses intense erotic feelings that, in the interests of cure, must remain unrequited. Patients may genuinely believe they are passionately in love with their analyst (e.g., the so-called *erotic transference*). Because the transference relationship is based on displacement and fantasy on the part of the patient, and because the analyst neither responds to cravings of the patient nor reveals himself or herself in a reciprocal or personal fashion (at least as practiced in classical analysis), it has been characterized (and often criticized) as being an artificial and asymmetrical bond.

Often what distinguishes a transference from a nontransference reaction is not its content per se but a group of qualities that tend to characterize transference responses and that may be used as signals to the analyst to denote their occurrence: Inappropriateness (the largely irrational character of the transference response); intensity (the unusual strength of reaction elicited in the patient); ambivalence (the shift or split in affect toward the analyst that occurs simultaneously or alternately); tenacity (the resilience with which such feelings tend to persist despite the analyst's actual behavior); and capriciousness (the erratic, and sometimes trivial or frivolous, events that evoke the feelings).

As transference distortions develop, their manifestations can be either positive or negative, paralleling the ambivalence that underlies all feelings that are in part unconscious. *Positive transference* refers to the expression of good feelings toward the analyst, of love and its many variations—as manifest in (albeit excessive) interest, trust, admiration, respect, sympathy, and so on—that can predominate as the motive force behind the wish to change and to receive the analyst's approval. In the state of positive transference, the patient overvalues and endows the analyst with some of the same magical powers attributed during infancy to the patient's parents. Such feelings may be the basis for the benign dependency used by the analyst to gain the patient's trust and establish rapport insofar as the patient is well-motivated and receptive to the analyst's influence. Freud wrote of the *unobjectionable positive transference* whereby the analyst is seen as clinician who is helpful and works collaboratively with his patient. Freud did not believe this transference attitude needed to be analyzed because it did not interfere with the work of analysis, that is, with the

patient's free associations. Subsequent analysts have been critical of Freud's view and felt even these benign positive feeling may conceal negative ones. At the other end of the affective spectrum, *negative transference* refers to equally intense bad feelings toward the analyst—including anger, hostility, mistrust, and rebelliousness—in which the patient undervalues the analyst in ways that also repeat comparable feelings toward parents or parent-substitutes of the past. Both types of transference are inevitable aspects of psychoanalysis (and often occur simultaneously or in close proximity to each other) and must be interpreted. Some analysts consider positive transference to be libidinal, based on sexual drives, and negative transference as predominantly a function of unresolved aggressive strivings.

It is usually negative transference that becomes most problematic and requires immediate analysis if treatment is to proceed; that happens because it becomes manifest in ways that interrupt treatment progress, either through direct attacks on the analyst or by the acting out of negative feelings instead of exploring them. Nevertheless, very intense positive transference, often expressed in excessive passionate demands on the analyst, can be misleading because it may also be a manifestation of resistance as the patient defends against further probing into unresolved conflicts. As traditionally understood, transference refers primarily to unrealistic distortions from the past, whether positive or negative; it does not pertain to reactions to present reality, as when the patient becomes legitimately angry. However, transference responses are increasingly recognized as having both objective and subjective components, relating both to significant figures of the past and to real responses to the analyst. Fred Pine summarized the changing views of transference: analysts have gone from Freud's explaining to patients, "You think you are talking about me, but it is really someone else in the past," to "You think you are talking about someone else in the past, but you are really talking about me," to the current view, "It doesn't matter what you are talking about, for there is an analytic drama happening between us, right here in this room." Arnold Cooper has described the present view of transference as "an interpersonal interaction, in which analyst and patient are powerfully influencing and affecting each other . . ."

TRANSFERENCE AS SUBSTITUTE PATHOLOGY: TRANSFERENCE NEUROSIS. The most vivid expression of transference is the formation of a transference neurosis, a substitute pathology: the original neurosis is displaced onto the relationship with the analyst, and a new, artificially acquired neurosis replaces the earlier one. All of the patient's symptoms lose their original significance and adapt themselves to a new meaning. Repetitions of childhood are fostered as the patient becomes continually frustrated in the search for gratification from the analyst; this is an approach based on the

assumption that a full-blown transference neurosis should eventually take place in which inappropriate infantile feelings are revived and re-experienced in their original intensity.

Manifestations of the transference neurosis do not occur immediately but emerge in the so-called middle phase of analysis, when the patient is most subject to the regressive forces induced by the analytic situation and to the emergence of infantile needs for gratification. Its appearance may be episodic, or it may never truly appear, although much of the effort of the middle phase is exerted to remove resistances and allow the transference neurosis to surface. The transference neurosis was originally regarded as a serious obstacle to analytic work, even though it also allows the analyst to observe directly the recapitulation of the patient's childhood responses. The resolution of the transference neurosis coincides with the cure of illness.

The depth of the transference neurosis has several determinants. The treatment situation influences the nature of transference dependency, including the degree of frustration (intentional withholding of wish fulfillment and of any information about the analyst in order to increase drive tension), the absence of reality information (the paucity of cues and feedback to validate the patient's perceptions), and the analyst's stance and activity (the greater the neutrality and silence, the deeper is the transference regression; the more extensive the use of analytic techniques, the more the transference is encouraged). Each interpretation may be seen as a form of frustration or deprivation that induces transference formation.

Apart from individual variations in the timing and extent of the transference neurosis, there is some controversy as to its content. For some it consists only of the new editions of old conflicts that Freud proposed, whereas for others it pertains not only to the original oedipal transference as displaced onto the analyst, but also to the defenses and resistances that are built up against it. Most investigators agree that the only way to resolve the transference neurosis is to analyze it. As the unearthing of early memories and the formation of a substitute neurosis is a gradual process that is not easily attained and that may not be feasible for some patients, achieving a complete transference neurosis and working through to its resolution may best be regarded as a final theoretical goal; it constitutes an ideal that is sought but rarely fully accomplished.

In more recent years, the focus on the importance of the transference neurosis per se has lessened. Although all analysts recognize the *sine qua non* importance of transference, some analysts, such as Arnold Cooper, no longer consider the establishment of the transference neurosis of significant importance in their work with patients.

TRANSFERENCES IN PATIENTS WITH NARCISSISTIC AND BORDERLINE PATHOLOGY. As noted, the term "transference neurosis" is generally used to describe the intense constellation of events that gradually places the analyst in a position of

unsurpassed dominance in the patient's life. Freud also used the term diagnostically to separate patients who were able to develop such an irrational attachment from those who could not. He characterized the former as having a transference neurosis and the latter as having a narcissistic neurosis. In his *Introductory Lectures* (1916 to 1917), Freud wrote that with these much sicker patients (e.g., those with dementia praecox or forms of paranoia) "after proceeding for a short distance, we come up against a wall . . . the resistance is unconquerable." He further noted that these patients "reject the doctor, not with hostility, but with indifference . . ." As a result, since they manifest no transference, they are "inaccessible to our efforts and cannot be cured by us."

Though Freud was writing about psychotic patients, his division contains the seeds of later work by Kohut, Kernberg, and others who developed new analytic theories and techniques to accommodate differences among patients with more severe psychopathology.

The increasing appearance in treatment of severe character pathology, especially narcissistic and borderline personality disorders, has confirmed that the intrapsychic structural organization of those patients is very different from that of their psychoneurotic counterparts. Instead of simply excluding such patients from psychoanalysis, analysts have adopted both theoretical and technical modifications to deal with the particular ego disturbances of self-image, faulty self-esteem, and split object relations that dominate their clinical picture.

The Freudian premises, which primarily focus on libidinal and aggressive impulses, are considered inadequate for understanding self-disorders. Not only is the actual content of the transference neurosis (i.e., incestuous sexual strivings) less pronounced, but the ideal of oedipal origin (which presumes that unconscious conflicts derive from a developmental stage after stable differentiation between self and others has occurred) also needs modification. In psychoanalysis the patient is assumed already to possess a cohesive sense of self and to have achieved a stage of developmental maturity beyond that attained by patients with serious ego disorders. Much of the concern about altering analysis relates to the need to manage the particular archaic preoedipal transference configurations, partial object relations, and primitive defense mechanisms of splitting and dissociation that are exhibited in patients with self-pathology toward the primary goal of restoration of the damaged or deficient self.

Kohut observed that patients with narcissistic personality disorders reactivate the specific mirroring and idealizing needs to which significant figures (self-objects) of early infancy failed to respond, and thus they do not develop typical transference neuroses: the central manifestations involve ways in which the patient deals with unfulfilled narcissistic needs. Kohut defined two types of narcissistic (or self-object) transferences that are considered pathognomonic for self-disorders and are direct repetitions in

treatment of earlier preoedipal interaction: (1) mirror transference, in which an unfulfilled need for accepting-confirming mirroring is revived, and (2) idealizing transference, which recreates the patient's need to merge with a source of idealized strength. In a similar vein, Kernberg observed a basic personality constellation in borderline personality disorder (BPD) patients, who also produce atypical transference manifestations (i.e., chaotic, contradictory, turbulent, and impulsive).

TRANSFERENCE AS A GENERAL PHENOMENON. Beyond the boundaries of the analytic situation, in which transference feelings are particularly intense, transference may also refer to the ubiquitous (and usually less exaggerated) tendency to displace the emotions and forgotten memories of one's past, particularly from the earliest, most affect-laden years of life, onto another person in one's present day life. Most notably, they are seen in other treatment situations, such as a hospital environment (called "institutional transference") or between the patient and authority or establishment figures (e.g., teachers and physicians). Significantly, transference to any physician, including an analyst or therapist, can typically begin even before the patient sees the doctor. Just being given a doctor's name, address, and telephone number may engender many fantasies and "false connections" related to a patient's past and not specifically related to the person of the doctor.

Resistance. Gabbard notes that one of Freud's greatest insights was that most people, no matter how pervasive their symptoms or their suffering, are ambivalent about changing and hence getting well. Resistance, as defined in the analytic situation, refers to all those powerful internal forces within the patient that obstruct the progress of analysis and oppose the analyst's efforts to help the patient achieve insight. Freud wrote, "The resistance accompanies the treatment step by step." Comparable to transference as a major pivot of psychoanalysis, analyzing or managing resistances is at the heart of analytic work. Gabbard called resistances the "analyst's daily bread-and-butter work." Resistances function in counterpoint to transference in two ways: (1) as resistance to the transference, which means that the patient fights against the development of a transference neurosis and thus prevents the analyst from tapping the source of intrapsychic conflict, and (2) as transference resistance, which means that the transference is used as a resistance by stubbornly adhering to irrational transference manifestations instead of as a path to understanding earlier experiences and memories.

Freud introduced the term *resistance* in the 1890s. He recognized that resistances were obstacles to treatment and had to be removed. By 1900 Freud had made several references to resistance in relation to dreams. His 1905 published case of Dora had demonstrated how the transference

relationship became a form of resistance in itself, an idea he subsequently elaborated years later in his papers on technique. The intensity of Dora's transference to Freud and Freud's failure to interpret it, as noted, led to Dora's premature and sudden termination of her therapy. The notion of transference as the most powerful and frequent form of resistance is still accepted.

Resistance can be conscious or unconscious, and can be produced by the ego, the id, or the superego. Conscious resistance is the deliberate withholding of information from the analyst, or the like. Such resistance is usually transient and usually easily rectified by pointing it out to the patient. Unconscious resistance, however, is a more significant and resilient phenomenon; it arises as a defense against uncovering repressed material that the ego wishes to avoid at all costs. Instead of appearing early in treatment or very intermittently, unconscious resistance occurs throughout analysis as a constant oppositional force whenever significant data begin to surface. Every association and every act of the person undergoing treatment must confront the resistance, which represents a compromise between the forces that are striving toward recovery and the opposing ones that need to keep the neurosis intact.

The clinical signs and manifestations of resistance are manifold. Any persistent, stereotyped, or inappropriate interruption of the treatment process may be a clue to resistance. Common examples include the silence of a patient who impedes progress through failure to verbalize and, at the other end of the spectrum, the compulsive talking of the patient who is ostensibly obeying the fundamental requirement to say whatever comes to mind but whose verbal productions are unconscious barriers to insight. Specific variations of resistant behavior include undue focus on the past that diverts attention away from the present and incessant inclusion of trivia or external events to avoid painful or emotion-laden topics. Typical forms of resistance also include lateness, missed hours, or delaying ("forgetting") to pay one's bill. Even nonverbal cues, such as physical posture of the patient, can indicate resistance.

Patients' resistances are manifestations of their defense mechanisms, that is, their typical ways they protect themselves from unpleasure. Gabbard notes that defenses, which operate intrapsychically, "are translated into resistances, which operate interpersonally" in the context of the patient's relationship to the analyst. Managing resistances in psychoanalysis means that the analyst must first address the defensive maneuvers of the patient before the avoided material can be approached. Because psychoanalysis is characterized by a systematic and thorough analysis of resistances, the analyst must discover how the patient resists, what is being resisted, and why. The immediate cause of resistance (e.g., anxiety, guilt, or shame) may be a superficial or surface indication of what is going on in the patient. Repeated uncovering and confronting of

resistances should reveal the affects underlying those behaviors. An important clinical distinction is made between those resistances that are ego-syntonic (viewed by the person as acceptable or consistent with his or her total personality) and those that are ego-alien (unacceptable or inconsistent). Only when the patient feels that a resistance is ego-alien will he or she be willing to give it up. The patient will generally deny the existence of, belittle, or rationalize ego-syntonic resistances. Gabbard has noted that resistance tends to change over the course of the analysis. For example, in the opening phase, a patient may avoid forming an attachment to his analyst; in the middle phase, it may involve extremely intense and particularly irrational feelings about the analyst; in the termination phase, it may involve a return of the original symptoms. Furthermore, says Gabbard, analysts' understanding of resistance has changed since Freud: it is now believed it is the analyst who is resisted rather than the recovery of unconscious memories.

Countertransference. Another potential obstacle to analytic progress may be the analyst's strong reactions to the patient, or *countertransference,* analogous to the concept of the patient's transference. These reactions of the analyst can inappropriately enter the treatment if the analyst is not sufficiently aware of personal unresolved feelings. It generally refers to the analyst's unconscious emotional needs, wishes, and conflicts that are evoked by the patient, brought into the analytic situation, and influence (usually negatively) the analyst's objective judgment and reason.

The term "countertransference" was coined by Freud in 1910. Freud wrote, ". . . no psychoanalyst goes farther than his own complexes and resistances permit, and we consequently require that he shall begin his activity with a self-analysis . . ." Freud, though, wrote little more about it. Initially, countertransference was seen more as problematic and something that interferes with the work of analysis. More recently, countertransference is seen quite differently. In fact, Theodore Jacobs maintains that it is a "concept whose time has come" and notes that "communication between the unconscious of the patient and analyst is a central feature of the analytic situation" and may reflect "the interplay between the minds of patient and analyst . . ." The notion of countertransference, though, reflects a related issue—the need not only for self-analysis, as Freud had mentioned in 1910, but for prospective analysts, as well, to undergo a training analysis as a prerequisite for working with patients. Arnold Cooper notes that Freud considered the most important achievement of a training analysis to be "a depth of conviction concerning the nature of the unconscious." In one of his last papers, *Analysis Terminable and Interminable,* published in 1937, Freud spoke of the need for an analyst to return for reanalysis periodically.

Countertransference becomes manifest in many ways. It is commonly acute, temporary, superficial, and easily recognized and managed, but it can also be chronic, permanent, deeply rooted, largely unconscious, and out of the analyst's control. Acute countertransference may occur in response to very specific content that arises or in identification with some concrete aspect of the patient's personality. Chronic countertransference involves more general and ingrained patterns of behavior, often pathological, that pervade the analysis in a way that is not therapeutic and to which the analyst remains blind without external intervention.

The former type of countertransference is more typical and occurs in every analysis. Like the patient, the analyst brings into the analysis aspects that do not belong there. Classic countertransferences may become manifest in special consideration of an attractive patient, as in eagerly making an unavailable hour available, or in failing to remember the changed appointment hour of an uninteresting patient. Common warning signs of countertransference in analysis include experiencing uneasy feelings during or after sessions with certain patients; persistently feeling drowsy or actually falling asleep during sessions; altering sessions or showing carelessness about scheduling (e.g., extending hours or forgetting appointments); making special financial arrangements (e.g., being overly strict with some patients and less assiduous with others); wishing to help the patient outside the session; dreaming about one's patients or being preoccupied with them in one's leisure time; using the patient as an example to impress a colleague or having the urge to lecture or write about a particular patient; reacting strongly to what the patient thinks of the analyst (i.e., needing a particular patient's approval); wanting, or not wanting, the patient to terminate; finding oneself unable to explore certain material or to understand what is going on with the patient; and evincing sudden or excessive feelings, such as anxiety, depression, or boredom during sessions.

Countertransference is presumed to relate primarily to unresolved and irrational responses, yet it may also refer to reasonable reactions to a patient's behavior, as when feeling aroused by a seductive patient, paternal toward a deprived patient, frightened by an aggressive patient, burdened by a demanding patient, or envy of a successful patient. Countertransference feelings are an inevitable part of any treatment. However, when such feelings are not situation-specific and evoke strong reactions that belong to former events or persons in the analyst's life, there is a chance that the analyst will bring them into the analysis in the form of unnecessary and even nontherapeutic behaviors.

Forms of countertransference may also reflect chronic residual problems from the analyst's analysis. These forms are obviously more serious, especially if the analyst remains unaware of them. Examples are the analyst with an underlying masochism who accepts abuse from patients

without adequately analyzing its reasons; the grandiose analyst who takes on the most difficult patients with promises of cure without recognizing when he or she needs help if the treatment is not going well; the so-called "macho male" analyst who allows the seductive patient to act out, reciprocating her sexual advances instead of examining her wish to arouse him and acknowledging his need for sexual conquest; or the lonely analyst who encourages the patient's dependency and will not terminate treatment for fear of abandonment. When the analysis becomes a source of narcissistic gratification for the analyst, who encourages the love or idolatry of the patient without introducing a more realistic appraisal or prematurely terminates patients who do not improve sufficiently, the analyst's return to analysis may become necessary.

Recent changes in theoretical and technical emphases on relationality and intersubjectivity have resulted in greater awareness of the complexity, multiplicity, and subtlety of forms of countertransference. These encompass such contemporary post-Freudian phenomena as projective identification (as used by Otto Kernberg, Thomas Ogden, and Heinrich Racker), role responsiveness (as originated by Joseph Sandler), and countertransference enactment (as developed by Theodore Jacobs, James McLaughlin, Judith Chused, and others). Some analysts, such as Owen Renik, believe that it is even helpful to share some of their countertransference feelings with a patient (i.e., self-disclosure). Renik believes this leads to a more collaborative relationship between analyst and patient. Other analysts are critical of this kind of self-disclosure, which goes against the notion of an analyst's neutrality or "blank screen" persona.

Ogden, a major thinker on the concept of countertransference, focused on projective identification, a term described in the 1940s by Melanie Klein. He described a series of three steps or aspects (an interpersonal dialectic rather than linear process), in which (1) a disavowed part of the patient's self is unconsciously placed into the analyst, (2) the analyst is in effect coerced by the patient to identify unconsciously with the particular projection, and (3) the analyst reprocesses the projection, such that a modified form is reintrojected by the patient. This triple relationship process simultaneously serves to go beyond two individual subjectivities of analyst and patient, creating what Ogden refers to as an "analytic third" (i.e., the analyst as a joint creation of the two subjectivities). He goes further to suggest that understanding these very early interpersonal processes helps to clarify the patient's relationship to the analyst in the present. The complexity of this triad of intersubjective experiences is further compounded by pre-existing feelings of the analyst and those projected by, or evoked by, the patient. The subtlety of these manifestations is also defined by Racker's distinction between complementary and concordant countertransference reactions. In the former, the analyst identifies with a self-representation of the patient, whereas in the latter, the

analyst identifies with a projected internal object representation. The presumption here is that the infant's early ego is able to perceive and process such self and object representations.

Sandler's term *role responsiveness* is another variation on this theme, emphasizing the "reflexive acceptance of the role which the patient is forcing on him." It is reminiscent of Bion's explicit emphasis on the coercion the analyst feels to play a role in the patient's fantasy. Although the analyst may totally attribute an irrational response on his or her part as a personal blind spot, in fact it may be regarded as a compromise formation between his or her tendencies and the coerced acceptance of an open-ended role. On the other hand, the analyst should not too quickly absolve himself or herself of responsibility for intense emotional reactions to a patient's words or deeds by attributing them to affects solely induced by the patient.

Theodore Jacobs wrote of *countertransference enactment* to refer to subtle interlocking transference–countertransference phenomena, such as a change in body posture as an unconscious enactment on the part of either party, insofar as both are continually reacting to one another. Chused defined it as the actualization of transference fantasy that evokes a countertransference reaction. As a special interpersonal phenomenon, it too can provide significant information about the patient's internal world. In fact, one implication of work in this area is the belief that insight into one's countertransference may only occur after some countertransference enactment has been recognized and understood. Other members of the relational school of thought (e.g., Stephen Mitchell and Lewis Aron) also insist that analysts always unconsciously play assigned roles in an open-ended process, no matter how consciously they try to do otherwise.

Sandor Ferenczi, back in the 1920s, first advocated the importance of self-disclosure and even experimented during the analyses he conducted whereby he would share his feelings and exchange roles with the patient, taking turns analyzing each other. His technique was initially severely criticized by so-called classical analysts. Theodore Jacobs, in his 2013 book *The Possible Profession,* wrote that more recently some analysts interested in the intersubjective connection between analyst and patient have reconsidered aspects of self-disclosure. Jacobs, though, notes the need for "caution when approaching the question of self-disclosure," particularly because what the analyst says carries considerable weight for the patient and the patient may hear the analyst's interventions "in ways that are quite different from what the analyst intended." He also distinguished intentional self-disclosure from that which is unintentional. Jacobs, for example, noted how his patient became much less critical of him after she had observed his alumni yearbook he had inadvertently left on his desk. Ultimately, self-disclosure is a clinical decision that depends on the specific patient, the phase of the analysis and the dynamics of the transference-

countertransference relationship. All would agree, however, that any self-disclosure should be for the patient's benefit rather than for the analyst's benefit.

Concepts of Nontransference. Before the 1950s whatever was not transference in analytic practice was regarded as "subversive" (Anna Freud's term). It was not until the 1950s, however, that analysts began to explore directly the potentially different relationships that occur in the therapeutic situation. As noted, Zetzel made the first broad differentiation in analytic treatment between its transferential and its realistic aspects by introducing the term "therapeutic alliance." Greenson later revised the term to "working alliance" and further delineated that nontransferential aspect of the analytic relationship from both transference and another form of nontransference—the real relationship.

THERAPEUTIC ALLIANCE. The therapeutic (or working) alliance is the task-oriented collaboration between therapist and patient in which the therapist's aims are to form an alliance with the conscious adult ego of the patient and to encourage the patient to be a scientific partner in the exploration of his or her problems. Non-neurotic rapport is the basis for mutual, purposeful, and realistic interaction, despite transference (and countertransference) impulses that are simultaneously at play.

No analysis can proceed without the formation of a rational, trusting therapeutic alliance; its establishment is the first task of treatment before a deeper transference neurosis can be facilitated. The patient's sense of suffering and the prospects of help from the analyst impel the patient to want to form a cooperative relationship. In addition to the therapist's consistent concern, care, and respect, which contribute to the formation of a rational working alliance, the therapist helps the patient to distinguish between what is realistic, healthy, and appropriate and what is distorted, neurotic, and inappropriate—to separate the fantasy figure from the real one. To the extent that a nontransferential, nonregressive interaction is sustained, the therapist will be able to monitor the patient's regression and irrationality and to enable the patient to receive and comprehend the therapist's communications rationally, to review and assess interpretations reasonably, to participate cooperatively and responsibly, and, ultimately, to integrate the insights that are gathered in treatment.

The therapeutic alliance does not arise or sustain itself automatically. A trial period may be necessary to assess the extent to which a patient can form and maintain an ongoing collaboration. For some fragile patients, the connection to another human being must be consolidated before therapeutic work can be done. For others, the empathic regard and professional concern of the therapist are more easily transmitted and serve as foundation for an already developed sense of security. During treatment

the inevitable intrusion of transferential elements requires the therapist repeatedly to restore the therapeutic alliance so that irrationality and regression do not disorganize the patient and threaten treatment.

The psychoanalyst Charles Brenner has questioned the concepts of the working or therapeutic alliance and feels that they are neither valid nor useful. According to Brenner, when a patient is in analysis, "it is best to approach *everything* analytically," that is, in the context of the transference. He believes that a discussion of these concepts relates to the issue of frustration or gratification of a patient within the analytic situation. When an analyst decides to gratify a patient's wishes, such as by answering a question or giving information, and so on, Brenner says that the decision must rest firmly on the analyst's understanding of the patient at that particular time. Brenner takes the hard line that an analyst should not even express condolences to a patient when a loved one has died because that may make it harder for the patient to express ambivalence toward the deceased. Many other analysts would take issue with that approach, believing that expressing condolences does not preclude exploring a patient's ambivalence. Years ago, psychoanalyst Ralph Greenson described his reaction of disbelief when his self-righteous supervisee, in the name of analytic technique, did not inquire about the condition of a patient's baby who had had convulsions and fever the day previously. Said Greenson, "It is not only natural, but imperative to indicate some compassion." Further, it was not unexpected to Greenson that the patient, who said nothing that following session but cried quietly throughout, announced subsequently that she had decided to terminate her treatment with this insensitive analyst.

REAL RELATIONSHIP. The real relationship is the realistic (reality-oriented and undistorted) authentic and true aspect of the therapeutic relationship. It is contrasted with the transference, which is genuine but unrealistic, and the working alliance, which is realistic but not genuine (i.e., an artifact of the analytic situation).

Most current psychoanalysis recognizes the distinction between transference (which is relived as if it were real) and that part of the therapist–patient interaction that is real, constituting an authentic human relationship; the latter is the medium through which transference reactions occur, but it also has an independent existence. Recent psychoanalytic work with patients with severe mental disorders has placed greater emphasis on the real reparative object relationship. In contrast to the treatment of neurotic patients, the treatment of psychotic patients has traditionally centered on the real relationship (as represented by such analysts as Frieda Fromm-Reichmann, Paul Federn, and Harold Searles, as well as Sullivan and Winnicott).

Especially in the initial contact, the therapist presents to the patient as a

real person who sets up the structure of reality through mutual negotiations regarding such issues as scheduling, fees, and vacations. Psychoanalytic psychotherapy not only introduces greater reality into treatment, but it accepts more readily that the patient's perception of the therapist may be the result of objective, or real, responses, not just distorted projections. Exclusive attention to transference may produce biased or erroneous interpretations, and in some therapists it reflects a defensive flight to avoid their discomfort with direct interpersonal engagement. Existential analysis, the most ardent expression of that position, is based on the notion that transference is always a genuine relationship *sui generis* between the analyst and analysand and places its entire emphasis on the real immediate encounter between the two partners.

Interpretive Process. Psychoanalysis is generally characterized by the role of interpretation as its fundamental, defining strategy—in effect, to analyze means to interpret. As a process it is preceded by the analytic techniques of confrontation and clarification, each of which plays an integral part in paving the therapeutic path to self-understanding. The goal of interpretation is the acquisition of insight.

Confrontation refers to pointing out or highlighting a particular mental event that the patient recognizes consciously, must face and has been avoiding. It may identify contradictory ideas. Auchincloss and Samberg note that confrontation as used in the analytic context "is presented tactfully and without aggression," that is, without the typical connotations outside of analysis. Confrontation can be preparatory to interpretation.

Clarification places a vague or confusing conscious event in clearer focus by separating important aspects from unimportant ones and elaborating on them. Clarification, according to Auchincloss and Samberg, just as confrontation, can be preparatory to the process of interpretation and can take the form of asking the patient questions or observing and commenting on a patient's behavior.

Interpretation goes beyond the material that has been discerned and elucidated by making the patient aware of previously unconscious psychological content and conflict that has been warded off and by attributing an underlying meaning or cause to the event in question, that is, it is the sine qua non of psychoanalytic technique. According to Auchincloss and Samberg, "effective interpretations" usually involve more than just an intellectual understanding; they should "resonate with some aspect of the patient's emotional experience."

Working through, most often referred to as an ongoing process rather than as a technique, is the culminating procedure that combines all three techniques. It entails the repetitive, progressive, and elaborated exploration of interpretations made until the material has become fully integrated into

the patient's ego. Auchincloss and Samberg note that for Freud, working through was the distinguishing feature of analysis because he believed other treatments were more apt to utilize suggestion.

Interpretation involves the special skill (or art) by which the analyst deciphers the latent meanings of the patient's verbalizations and translates them in a way that can be comprehended and accepted. It includes hidden historical connections between present and previous behaviors, between feelings toward the analyst and those meant for significant others, between instinctual impulses and defenses built against them, and, ultimately, between conscious thoughts and unconscious wishes traced to their infantile roots. The interpretive process should form an intelligible sequence that links the neurotic adult with the repressed child who lurks beneath the surface. By extrapolations from the manifest material of free association, dreams, and parapraxes (i.e., the so-called "Freudian slips," such as slips of the tongue, mislaying of objects, and so on) the analyst decodes the disguised symbolic language of the unconscious into meaningful terms that are relevant to the patient. Interpretations have both direct and indirect functions. The major purpose is to provide the patient with an understanding of himself or herself and the source of the problem. They also serve the technical purpose of frustration rather than gratification because each introduced interpretation takes the patient away from the analyst as the fulfiller of wishes. T. Byram Karasu believes, "All interpretations are deprivations: good ones bring disappointment and bad ones cause disengagement." As the primary communication link between analyst and patient, the act of interpreting, though, may also serve to transmit empathy, concern, and care along with the painful content of the interpretation.

Psychoanalysis is often distinguished by the use of interpretations that are dynamic and genetic in nature. The dynamic interpretation demonstrates how intrapsychic forces operate to produce a particular effect on the patient's psychological life (e.g., the superego's role in feelings of guilt). The genetic interpretation traces patient behaviors and feelings back to their early origins in infancy and childhood. Those reconstructions maintain an essentially cause–effect framework, seeking the meaning of events in old, disguised conflicts or traumas (e.g., tracing patient affect to early sexual or rivalrous feelings as expressions of unresolved oedipal wishes). Sometimes, deep genetic interpretations are too threatening or premature. There are also distinctions between transference interpretations, which expressly focus on the analyst as a substitute or fantasy figure from the past, versus nontransference interpretations (also called dynamic interpretations), which essentially view the patient's responses as reactions to current events and persons in the "here-and-now."

GUIDELINES FOR INTERPRETATION. Ideally interpretation is an individualized, patient-centered strategy, with both content and timing based on the unique psychic activity and receptivity of the analysand, not on an analyst's preconceptions or preferences. A careful interpretation meets four criteria: optimum timing, minimum dosage, concrete detail, and individual focus.

The further the analyst departs from the spontaneous data provided, the greater is the chance of erroneous, premature, or "inexact" interpretations. Other general guidelines pertain to the use of a topographical sequence in which the analyst starts from manifest content or from the patient's defense before attempting to deal with the latent content that underlies it and the viewing of conflicts in increasingly historical perspective by providing interpretive links to early emotion-laden past events only when more recent material has been integrated. Titrating the timing and dose of interpretation is a clinical skill whereby the analyst may gauge the areas of strongest or most inappropriate affect or those often-repeated themes that indicate focal conflicts and, at the same time, assess the patient's capacity to accept the particular interpretation at that moment. Interpretation in psychoanalysis is traditionally characterized by a centripetal orientation toward the analyst–patient axis in which analysis of transference is regarded as a relatively closed or circumscribed intrasession activity as opposed to a more centrifugal orientation of psychotherapy away from that axis in which greater attention is paid to extratherapeutic events.

Dream Interpretation. Freud's classic work *The Interpretation of Dreams* and subsequent papers on the mental phenomena of everyday sleeping and waking seek to demonstrate that dreams, parapraxes, jokes, puns, and metaphors are not chance or senseless events but are meaningful acts subject to the basic laws governing unconscious processes. Freud believed early on that the analysis of dreams, such as he had conducted on himself, was sufficient for a full self-analysis. As the fastest and surest route to the origins of psychic conflicts—"the royal road to the unconscious"—dreams remain for many analysts the perfect paradigm of the interpretive process.

Dreamwork is a distorting process by which latent dream thoughts are condensed, displaced, and converted through symbolic substitution into manifest dream material, of which the recollection, on wakening, is only an unintelligible facade or disguise. The dream is generally looked on as a microcosm of the major concerns of the patient, especially as a commentary on the events of the preceding day (day residue). For the dreamer it serves a largely wish-fulfilling function behind which resides the original unsatisfied instinctual impulse. Current views also suggest a prospective function of dreams, as a direct attempt to adapt to present stresses or to solve anticipated problems. Despite its salient role in analysis, however, dream interpretation should not be pursued for its own sake but

should be subject to the same rules that govern all analysis. Excessive attention to dreams can become, in fact, a form of resistance by drawing the patient away from more spontaneous concerns. Sometimes, dreams can function as a "gift" a patient brings to the analyst, especially if the patient believes that the analyst has a special interest in dreams. Although Freud thought of dreams as special and as often supplying a source of information that cannot be derived in any other way, many analysts, such as Charles Brenner, believe that there is nothing particularly unique or special about the dream, and that anything a patient discusses—a symptom, a joke, a movie, a book—can be the royal road to unconscious mental processes.

In the first days of her analysis, a psychologically minded 19-year old college student reported the following dream image (*manifest content*):

An old man, somewhat stooped, was putting large corrugated cardboard boxes in my way on the road.

She quickly made the connection that the old man stood for her analyst. Her *free associations* led her to think of the considerable difference in age between her and that of her analyst. The day before (*day residue*), she had also noticed an attractive painting in the analyst's office of an old, bent-over man, very similar to the image in the dream that confirmed for her the connection of the images to her analyst and his office. The dream seemed to represent her beginning her course of analysis symbolized by the image of the road (i.e., *displacement* where one image/idea is substituted for another). In the dream, she remembered, she had actually been in her car. The image of the car reminded her that she had now taken on a major time commitment in beginning analysis 5 days a week because it entailed a long car commute between her school and her analyst's office. Further, she knew that the analyst was going to confront her along the way, with enormous challenges, symbolically represented by the large boxes. The analyst asked for her associations specifically to the corrugated cardboard boxes. A Latin scholar, she knew that the derivation of the word was from the Latin "ruga," meaning "wrinkles," but it also made her think of the "rugae" of the vagina. She knew that sexual material would be part of the challenges she would face in analysis. These associations, though, made the analyst believe that the image of the corrugated boxes was an example of *condensation* (i.e., one image standing for many thoughts). He speculated that it could be reflective of the beginning of an (as yet unconscious to the young woman) erotic transference (i.e., her disguised, *latent wish* to share her vaginal "rugae" with him) that was to develop over the course of the analysis. The analyst did not interpret (i.e., make conscious) this notion to the patient at this initial stage of her analysis.

Psychoanalytic Process

The psychoanalytic process involves bringing to the surface repressed memories and feelings by means of a scrupulous unraveling of hidden meanings of verbalized material and of the unwitting ways in which the patient wards off underlying conflicts through defensive forgetting and repetition of the past. Psychoanalyst Lawrence Friedman, though, emphasizes the need for a therapist to have a theory. His two principles are "act like therapist" and "be curious," or in other words, "look for unsolved puzzles." What makes therapy (and particularly psychoanalysis) different from other encounters are the responses of the therapist.

The overall process of analysis is one in which unconscious neurotic conflicts are recovered from memory and verbally expressed, re-

experienced in the transference, reconstructed by the analyst, and ultimately resolved through understanding. Freud referred to those processes as "recollection, repetition, and working through," which make up the totality of remembering, reliving, and gaining insight. Recollection entails the extension of memory back to early childhood events, a time in the distant past when the core of neurosis was formed. The actual reconstruction of those events comes through reminiscence, associations, and autobiographical linking of developmental events. Repetition involves more than mere mental recall; it is an emotional replay of former interactions with significant persons in the patient's life. The replay occurs within the special context of the analyst as projected parent, a fantasied object from the patient's past with whom the latter unwittingly reproduces forgotten, unresolved feelings and experiences from childhood. Finally, working through is both an affective and a cognitive integration of previously repressed memories that have been brought into consciousness and through which the patient is gradually set free (i.e., cured of neurosis).

Narrative versus Historical Truth. Many analysts today would take issue with the idea that events are actually reconstructed in analysis. Donald Spence is a proponent of the notion that a patient's "narrative truth" is different from his "historical truth." For Spence, historical truth is "time-bound and its aim is to come as close as possible to what 'really happened.'" But, says Spence, what really happened does not necessarily correspond to what the patient remembers and reports in analysis. Spence believes that Freud's use of the metaphor of archeology (and "a faith in the idea that psychoanalysis is an archeology of memories") may be "a substantially misleading analogy" when comparing it to the work of uncovering in analysis. He believes that an interpretation is more often a "construction—a creative proposition, rather than a reconstruction that is supposed to correspond to something in the past." Says Spence, "In making an interpretation, we exchange one kind of truth—historical truth—for the truth of being coherent and sayable—narrative truth" and ". . . the construction not only shapes the past—it becomes the past." Interpretation, then, is possibly less about "a referent in the past" and more a "pragmatic statement." He believed that Freud was too willing to generalize from one case. Says Spence, "We will always be hampered by incomplete data."

Stages of Psychoanalysis. The classical analytic course may be subdivided into three major stages as follows:

STAGE ONE. During the initial phase, or beginning of treatment, the patient becomes familiar with the methods, routines, and requirements of analysis and a realistic therapeutic alliance is formed between patient and analyst. Basic rules are established, and the patient describes his or her

problems, gives some review of history, and gains initial relief through catharsis and a sense of security before delving more deeply into the source of the illness. The patient is primarily motivated by the wish to get well.

STAGE TWO. Freud compared psychoanalysis to a chess game in which the initial and end stages are subject to uniform technical rules that are virtually identical for all players but whose long middle phase represents a spontaneous unfolding in which each protagonist's part is played out differently; it is thus often thought to be the most creative stage. The analogy still broadly holds, insofar as each analysis is highly individual, its course can never be precisely predicted, and the participants must carefully receive their cues from each other at any particular moment.

During the long middle phase, a transference neurosis may develop. This may substitute for the actual neurosis of the patient and in which the wish for health comes into direct conflict with the simultaneous wish to receive emotional gratification from the analyst. There is a gradual surfacing of unconscious conflicts; an increased irrational attachment to the analyst, with regressive and dependent concomitants of that bond; a developmental return to earlier forms of relating (sometimes compared to that of mother and infant); and a repetition of childhood patterns and recall of traumatic memories through transfer to the analyst of unresolved libidinal and aggressive wishes.

In essence the middle stage is characterized by a continual battle within the psyche of the patient and between patient and analyst wherein infantile problems of childhood are played out with the analyst as the object of the patient's repressed strivings. It is also the most protracted and painstaking phase, in which recall, repetition, and synthesis of early interpersonal patterns are relived and reconstructed over and over until the patient develops new modes of reacting and interacting that are more mature and realistic. Through the process of working through, the patient is confronted with the same material, or variations of it, many times over until the warded-off material can be sufficiently neutralized and permanently integrated into the psyche. The length of the middle phase (sometimes interminable) is a function of the resilience of infantile behavior patterns and a manifestation of the fact that it takes numerous, repeated confrontations, clarifications, and interpretations to alter the psychic structure to effect permanent change.

Sometimes, however, there may develop what is called a "negative therapeutic reaction." Here, paradoxically, the patient seems to get worse in the course of treatment, often seen after the understanding of an important insight or some sense of progress. This negative reaction may have multiple determinants. For example, patients may be masochistic and have strong feelings of guilt or wishes for punishment and suffering. Sometimes, this reaction occurs in patients who are competitive with and

envious of their analyst, with strong wishes to defeat the analyst and are determined unconsciously to have treatment fail. Treatment may, in fact, end in a stalemate if these unconscious wishes are not interpreted. Sometimes, an analyst overzealous and over-eager to cure a patient—the so-called "rescuers"—may arouse this clinical situation in patients who are psychologically predisposed to it.

STAGE THREE. The termination phase is marked by the dissolution of the analytic bond as the patient prepares for leave taking. The irrational attachment to the analyst in the transference neurosis has subsided because it has been worked through, and more rational aspects of the psyche preside, providing greater mastery and more mature adaptation to the patient's problems. However, it is not an unambivalent experience because it often revives separation anxiety and old issues of dependence on versus independence from significant persons of the past (as well as transient recurrences of original symptoms). It has been likened to a mourning process in which the patient grieves over the impending loss of the analyst and the loss of his or her former self. However, by incorporating the insights received the patient can better use his or her self-knowledge to separate the reality of the analyst from the fantasy and the past from the present. Termination is not a hard-and-fast event, and the patient invariably has to continue to work through any problems outside of the therapy situation without the analyst or may need intermittent assistance after analysis has technically terminated. The patient is expected to continue "self-analysis" after treatment ends.

Indications and Contraindications

In general, chronic cases of psychoneuroses (and relatively mild psychopathology), including all forms of symptomatic anxiety and depression, as well as obsessional, and hysterical neurotic manifestations, are considered the most amenable to psychoanalysis. Psychoanalysis is precluded when dangerous behavioral or physical symptoms require more immediate attention or for virtually the entire realm of the psychoses, in which severe ego deficits and tenuous reality are apparent. Merton Gill described the indications for psychoanalysis: "the ego is sufficiently damaged that extensive repair is necessary, but sufficiently strong to withstand pressure."

However, psychoanalysis has been applied to a wide range of patients well beyond those with symptomatic neuroses, in part because the kinds of patients who appear for treatment today differ diagnostically from those in the past. Patients with hysteria and other neuroses have been largely replaced by patients with character disorders, such as those with narcissistic personality disorders or borderline personality disorders. These patients can be treated with psychoanalytic methods, but results tend to be

favorable to the extent that neurotic traits (i.e., areas of conflict) are evident along with narcissistic ones. The same appears to be true of sexual disorders: Some patients with deviant behaviors seem to be intractable to psychoanalysis, but when those behaviors are accompanied by neurotic difficulties, they are considered amenable to psychoanalytic work. A small percentage of patients with psychosomatic illness may also need special modifications of psychoanalysis, in light of a cluster of characteristics often observed in such patients—called alexithymic (an inability to verbalize feelings, poverty of fantasy life, and denial of psychic conflict)—which makes them poor candidates for traditional methods. Persons with schizoid or paranoid personality disorders are also unsuitable, as are those with more serious psychosocial disorders, including serious substance use disorders (narcotic addictions and alcoholism) and antisocial personality disorder. Infantile demands, poor impulse control, an inability to tolerate frustration, impaired social judgment, and the physical concomitants of substance abuse severely limit the benefits of psychoanalysis.

Concomitant Use of Medication with Psychoanalysis

The concomitant use of medication in the context of analysis "is one of the most prevalent changes in psychoanalytic practice over the past thirty years," according to psychoanalysts Robert Glick and Steven Roose. Years ago, psychoanalysts discouraged any other concurrent therapies, such as use of psychotropic medication, because they felt that medication would obscure and attenuate symptoms, particularly anxiety and depression, interfere with a patient's motivation, and lead to a "deleterious effect." There is now "growing evidence" that medication can work synergistically with analysis. Use of medication, particularly for depression, is much more common now. These authors quote studies that document that about 30 percent of patients in analysis are on psychotropic medication, and rather than having a deleterious effect on the analysis, this may actually lead toward more psychological-mindedness in patients and a deepening of the analytic process. Glick and Roose, however, caution that analysts may develop countertransference reactions. For example, these analysts may be more concrete, take symptoms at face value, and be less psychoanalytic ("analytically tone deaf") when discussing medication and its side effects. They note that some patients will claim to be "sensitive to all medications" "an intrapsychic reality" but a "physical impossibility." They found that with some patients and analysts, when medications are introduced into the treatment setting, an idealization of both the analyst and analysis is threatened—that is, "the very fact of a prescription means the analysis alone is insufficient." For analysts, there may be guilt and shame and feelings that they are committing "an act of treason" by writing a prescription. Alternatively, prescribing medication can feel like it is a magic potion. It can also seem like a boundary violation in that "analysis is a non-

contact sport" and medication is "crossing the body boundary." Glick and Roose emphasize that an analyst must be "bilingual," that is, keep in mind the different and shifting perspectives involved when medications are part of the analysis. All associations must be subject to analytic scrutiny.

There is also a growing literature on the importance of the psychosocial aspects of psychopharmacology. These "nonspecific elements" may be just as potent as a medication's active ingredients. In other words, *how* a therapist and physician in general prescribe medication may be just as important or even more important than which medication is chosen. Furthermore, there is often no way of ascertaining whether when a patient benefits, it is from the prescribed medication, the patient's alliance to the physician, the patient's expectations, or even a placebo response. One patient described feeling so much better just having the bottle of medication with the therapist's name on the bottle. The prescribed medication seemed like a gift from the therapist.

Patient Requisites

The most important patient requisites for psychoanalysis are the following:

1. *High motivation*. The patient needs a strong motivation to persevere in light of the rigors of intense and lengthy treatment. The desire for health and self-understanding must surpass the neurotic need for unhappiness. The patient must be willing to invest considerable time and money and to endure the pain and frustration associated with sacrificing rapid relief in favor of future cure.
2. *Ability to form a relationship*. The capacity to form and maintain, as well as to detach from, a trusting object relationship is essential. Psychoanalysis cannot begin, be sustained, or terminate properly without such a basis. The patient also has to withstand a frustrating and regressive transference without decompensating or becoming excessively attached. Patients with a history of impaired or transient interpersonal relations and who cannot establish a viable connection to another person make poor psychoanalytic patients.
3. *Psychological mindedness and capacity for insight*. As an introspective process psychoanalysis requires curiosity about how the mind works as well as curiosity about oneself and the capacity for self-scrutiny. Those who are unable to articulate and comprehend their thoughts and feelings cannot negotiate with the fundamental analytic coin—words and their meanings. The inability to examine one's motivations and behaviors precludes benefits from the analytic method. Psychological mindedness is considered a process while insight is the product of that process. Whether a patient can be trained to become psychologically minded is open to question. Many analysts believe it may be like musical proficiency that is dependent on a basic musical ability, but if dynamic

conflicts are interfering, resolution of those conflicts can increase a person's ability and capacity for psychological mindedness.

4. *Ego strength*. Ego strength is the integrative capacity to oscillate appropriately between two antithetical types of self-regulation: On one hand, the patient must have the conscious cognitive ability to reflect, to relinquish reality for fantasy, and to be dependent and passive. On the other hand, the patient has to be able accept analytic rules, to integrate interpretations, to defer important decisions, to shift perspectives so as to become an observer of his or her intrapsychic processes, and to function in a sustained interpersonal relationship as a responsible adult.

In addition, Freud felt that people "near or above the age of 50" were too old and inflexible for psychoanalytic treatment—that "the elasticity of the mental processes, on which the treatment depends, is as a rule lacking—old people are no longer educable—and the mass of material to be dealt with would prolong the treatment indefinitely." Few analysts today would accept that limitation.

Goals

The aim of psychoanalysis, in its most parsimonious form, is expressed in the now-famous Freudian dictum, "Where id was, so shall ego be." Ideally, there is a structural intrapsychic reorganization such that there is a reduction in the patient's punitive superego, as expressed in less paralyzing guilt and a greater capacity for pleasure. The patient is better able to accomplish the two life-goals (attributed to, but probably never explicitly stated by Freud), *arbeiten und lieben*—to work and to love—as those basic activities, along with play, of living are released for fuller gratification because the underlying inhibitions have been removed.

Stated in developmental terms, psychoanalysis aims at the gradual removal of amnesias rooted in early childhood, based on the assumption that when all gaps in memory have been filled, the morbid condition will cease because the patient no longer needs to repeat, or remain fixated to, the past. Giving up infantile wishes and not having to recapitulate early inappropriate behaviors is an expression of the patient's successful negotiation of successive stages of psychosexual development, culminating in the genital phase of mature sexuality and intimacy. The patient should be better able to relinquish former regressive patterns and to develop new, more adaptive ones, particularly as he or she learns the reasons for his or her behavior. A related goal of psychoanalysis is for the patient to achieve some measure of self-understanding, or insight and have a relief of symptoms.

Psychoanalytic goals are often considered formidable (e.g., a total personality change), involving the radical reorganization of old developmental patterns based on earlier affects and the entrenched

defenses built up against them. Goals also may be elusive, framed as they are in theoretical intrapsychic terms (e.g., greater ego strength) or conceptually ambiguous ones (resolution of the transference neurosis). Criteria for successful psychoanalysis may be largely intangible and subjective and are best regarded as conceptual endpoints of treatment that must be translated into more realistic and practical terms.

In practice, the goals of psychoanalysis for any patient naturally vary, as do the many manifestations of neuroses. The form that the neurosis takes—unsatisfactory sexual or object relationships, inability to enjoy life, underachievement, and fear of work or academic success, or excessive anxiety, guilt, or depressive ideation—determines the focus of attention and the general direction of treatment as well as the specific goals. Such goals may change at any time during the course of analysis, especially given that many years of treatment may be involved. Psychoanalyst Ernst Ticho has divided goals into *treatment goals* and *life goals*. Life goals, including both personal and professional goals, are those that a patient would hope to reach if he can put his "potentialities to use" whereas treatment goals are those that are involved with the "removal of obstacles to the patient's recovery of what his potentialities are." For Ticho, treatment goals should be discussed with the patient during the very first few hours of analysis but should "unfold and develop" as treatment proceeds. Treatment goals include establishment of mature relationships and a decrease in "narcissistic self-centeredness," says Ticho, but may also include increased frustration tolerance, freer access to the patient's unconscious, and a "more tolerant, integrated and loving superego." The analyst, though, must separate his own life and treatment goals from those of his patient.

The Main Approach

Structurally, psychoanalysis usually refers to individual (dyadic) treatment that is frequent (four or five times per week) and long term (several years). Freud initially saw patients 6 days a week for 1 hour each day. Abraham Kardiner, a psychoanalyst and patient of Freud, told how Freud, taking the suggestion of his daughter, Anna, cut back on 1 hour a week of every patient to make room for an additional patient. That is how 5 hours a week evolved, a routine now sometimes reduced to four sessions of the classic 50-minute (and now sometimes 45-minute) hour. The International Psychoanalytic Association recently approved three sessions a week as acceptable. Long intervals between sessions are avoided so that the momentum gained in uncovering conflictual material is not lost and confronted defenses do not have time to restrengthen. Over time, psychoanalyses have gotten much longer. Freud saw many of his patients for a few months. Kardiner, who saw Freud for 6 months during 1921 to 1922 (and was charged $10 a session by Freud), once asked Freud what he thought of himself as an analyst. Freud replied, "I am much too impatient

now. . . . I am much too occupied with theoretical problems all the time . . . rather than paying attention to the therapeutic problems. . . . I have no patience in keeping people for a long time. I tire of them, and I want to spread my influence."

Psychoanalytic Setting. Like other forms of psychotherapy, psychoanalysis takes place in a professional setting, apart from the realities of everyday life, where the patient is offered temporary sanctuary in which to ease psychological pain and reveal intimate thoughts to an accepting expert. The psychoanalytic environment is designed to promote relaxation and regression. Significantly, Freud's office was anything but neutral: His couch was piled high with Oriental rugs, and he kept his extensive collection of archeological antiquities in full view on his desk. He even had his dog Jofi, Freud's favorite chow (among the several dogs he had over the years), in the office during sessions.

USE OF THE COUCH. The traditional setting is a formalized part of the psychoanalytic heritage, still modeled on the physical arrangement of Freud's medical office at Bergasse 19 in Vienna. The analyst's couch is the most enduring structural requirement. The couch allows the patient to lie down within view of the analyst, who remains seated in a chair outside of the patient's field of vision.

The couch is a convention that preceded psychoanalysis, a carryover from the days when Freud used hypnosis. It has persisted as a primary feature of psychoanalysis, designed to facilitate patient communication by way of free association. Its original rationale was neither comfort nor convenience, but it involved a personal motive for Freud, namely he admitted that he could not bear being looked at for 8 hours a day and believed that other analysts shared that distaste. Since Freud, many analysts have commented on the relevance of the couch. Both Jung and Fairbain questioned its use. Brenner thought it had "real and concrete advantages" and was not an "atavistic anachronism." Merton Gill, who distinguished between intrinsic criteria (e.g., theory behind analytic technique and the analyst's neutrality) and extrinsic criteria of analysis, considered frequency of sessions and the use of the couch as extrinsic criteria. Gill, in his 1984 discussion of the differences between psychoanalysis and psychotherapy, wrote, "The argument that psychoanalytic technique cannot be used in the face-to-face position because of the reality cues afforded the patient is a variant of the mistaken idea that it is the external stimulus rather than the interpretation of that stimulus that matters." The couch is often considered the "border" between psychoanalysis and psychotherapy.

Nevertheless, the couch has several clinical advantages that are both real and symbolic: (1) The reclining position is relaxing and eases the

patient's conscious control of thoughts; (2) it minimizes the intrusive influence of the analyst, thus curbing unnecessary cues; (3) it permits the analyst to make observations of the patient without interruption; and (4) it holds symbolic value for both parties, a tangible reminder of the Freudian legacy that gives credibility to the analyst's professional identity, allegiance, and expertise. For some patients, though, the couch can be seen as threatening, with personal meanings of submission to an authority figure; for others, a relief from confrontation by the analyst (e.g., fear of use of the couch and over-eagerness to lie down, or "couch-diving," may reflect resistance and thus need to be analyzed). Although the use of the couch has been considered a requisite to analytic technique, it does not have to be applied automatically; it can be introduced gradually, and may be suspended whenever additional regression is unnecessary or countertherapeutic.

FUNDAMENTAL RULE. The fundamental rule of free association requires patients to tell the analyst everything that comes into their minds, however disagreeable, unimportant, or nonsensical. It differs decidedly from ordinary conversation—instead of connecting one's remarks with a rational thread, the patient is asked to reveal those very thoughts and events even though they are objectionable and regardless of how irrational they seem.

That directive represents an ideal because free association does not arise freely but is guided and inhibited by a variety of conscious and unconscious forces. On the most overt level, to renounce the natural tendency toward deliberate censorship, especially with a relative stranger, requires a firm basis in trust. On a more covert level, there is an interplay of motives that direct the ambivalent course of free associations, alternately fostering and threatening the progress of analysis. The analyst not only must encourage free association through the physical setting and a nonjudgmental attitude toward the patient's verbalizations, but he or she also must examine those *very* instances when the flow of associations is diminished or comes to a halt—that is, both the content and the process of the associations. The analyst should also be alert to how individual patients use or misuse the fundamental rule. Some obsessive-compulsive patients, for example, may apply the rule so literally that they appear to be fastidiously complying with it while actually sabotaging treatment by using it to avoid self-observation. In other patients, the inability to stop primary process associations because of a breakdown of ego functioning may signal a latent psychosis.

Aside from its primary purpose of eliciting recall of deeply hidden early memories, the fundamental rule reflects the analytic priority placed on verbalization, which translates the patient's thoughts into words so that they are not channeled physically or behaviorally. As a direct concomitant of the fundamental rule, which prohibits action in favor of verbal

expression, patients originally were expected to postpone making major alterations in their lives, such as marrying or changing career, until they discussed and analyzed them within the context of treatment. As analyses have gotten considerably longer, this recommendation is no longer practical.

Closely tied to the fundamental rule is the concept of *psychic determinism,* namely that mental events (i.e., associations) are causally linked to each other. In other words, all thoughts, associations, and behavior are not random, arbitrary, or haphazard and are motivated, either by conscious or unconscious forces. Another related concept is that of *multiple function* (Robert Waelder's term) or *overdetermination*. Here behaviors, symptoms, or even images in a dream must have more than one motive, meaning, or even determining factor, and it is incumbent on the analyst to look for these layers of meaning when listening to a patient.

PRINCIPLE OF EVENLY SUSPENDED ATTENTION. As a reciprocal corollary to the rule that patients communicate everything that occurs to them without criticism or selection, the principle of evenly suspended attention requires the analyst to suspend judgment and to give impartial attention to every detail equally. The method consists simply in making no effort to concentrate on anything specific but to maintain a neutral, quiet openness to all that is said. Freud described the analyst's attitude as the surrender to unconscious mental activity to avoid as much as possible any reflection or conscious expectations—to "catch the drift of the patient's unconscious with his own unconscious." That special attitude, or way of listening, was originally recommended out of Freud's major concern that the analyst needs to keep in mind innumerable dates, names, detailed reminiscences, and associations without using any expedient, such as note taking, as an assist to memory. Often misunderstood as an intellectual or objective orientation that distances the analyst from the patient, free-floating attention means forming a connection to the patient while giving free play to ideas aroused in oneself. Lawrence Friedman notes that "free-floating attention" is a relative term: "it is not absentmindedness, reverie, transcendental meditation, or untetheredness." Friedman also notes that the "paradigm" of psychoanalysis, unlike other talking therapies, is that what comes across from the analyst to the patient is "basically the patient's own voice," and the analyst "should not contaminate the patient's voice."

The notion of equidistance, that is, so-called technical neutrality, suggests that the analyst attempt, in effect, to keep equidistant from all of the psychic agencies (id, ego, and superego) and not ally too closely with any of them. The analyst thus allows each area of mental functioning to be fully expressed without influencing its strength by excessive attention or collusion.

ANALYST AS "MIRROR." In his 1912 paper, *Recommendations to Physicians Practicing Psycho-analysis,* Freud suggested that an analyst "should be opaque to his patients, and like a mirror, should show them nothing but what is shown to him." That means that analysts are not to bring their values or attitudes into the discussion or to share personal reactions or mutual conflicts with their patients, even though they may sometimes be tempted to do so.

The bringing in of reality and external influences can interrupt or bias the patient's unconscious projections. In effect, the analyst accepts without censure all the forbidden or objectionable responses of the patient. However, the mirroring role is never entirely neutral or passive, as it is sometimes characterized. The analyst as totally value-free is a myth because it is impossible not to impart one's social, cultural, or personal standards through unconscious cues or behaviors. In practice, though, the analyst as mirror can misrepresent the human part the analyst implicitly plays. The analyst's restraint from direct suggestion and affective interaction should not preclude transmitting respect, concern, and nonjudgmental caring.

RULE OF ABSTINENCE. Freud advised that the analyst carry through the analytic treatment in a state of renunciation. The analyst must deny the patient who is longing for love the satisfaction that the patient craves. This not merely means corporal or sexual abstinence, but also encompasses the frustration of emotional needs and wishes. The fundamental rule of abstinence allows the patient's longings to persist and to serve as driving forces for analytic work and motivation to change. The analyst must curb the temptation to take on the role of a nurturing figure (parent, lover, or friend) who satisfies the patient's infantile libidinal wishes or to interact behaviorally instead of analyzing the patient's associations. Like any other instrument, the rule of abstinence must be applied with professional judgment to achieve the optimal balance of frustration and gratification (in light of the patient's tolerance for being denied what he or she craves). It is a misunderstanding of analytic intent, not to mention poor practice, to assume that the analyst must be totally sterile and withholding or that the patient should be expected to endure such treatment without getting worse or leaving. Along those lines, Leo Stone has suggested that much of the "prevailing psychoanalytic practice" constitutes the "overwrought and indiscriminate application of the principle of abstinence."

Question of Lay Analysis. Freud was trained as a physician, but many of the early psychoanalysts, such as Erich Fromm and Melanie Klein, and even Freud's youngest daughter Anna, were not physicians. They were called "lay analysts." Freud never thought that a medical degree was necessary for psychoanalytic training; to him, an analyst's personal analysis

was most important, and although he thought of psychoanalysis as a science, he thought of it as part of psychology. He even defended the nonmedical analyst Theodore Reik when he was prosecuted for "quackery" by the Austrian government.

When psychoanalysis was brought to the United States in the early 20th century, however, it was the time of medical reform and the Abraham Flexner Report that standardized medical education. Psychoanalytic training, at least under the auspices of the American Psychoanalytic Association, was exclusively limited to physicians with or in psychiatric training. All of that changed when nonmedical analysts sued the American Psychoanalytic Association in the 1980s and won. This case led to the "Gaskill Amendment" that stated psychoanalytic institutes could not ban nonmedical graduates from their programs because of a lack of a medical degree. Psychoanalyst Arnold M. Cooper, former President of the American Psychoanalytic Association, believes the passage of the Gaskill Amendment "effectively ended the separate identity of medical psychoanalysis in America."

Limitations. Critics of psychoanalysis as a theory and therapy have questioned the universality and validity of major psychoanalytic constructs (from the Oedipus complex to the tripartite mental structure of id, ego, and superego) and the reliability of the case study method, which is considered the most subjective form of research observation in establishing effectiveness. More specific is the question of the validation of interpretations, on which rests the integrity of psychoanalysis as a clinical method.

The predominant treatment constraints are often economic, relating to the high cost in time and money, both for patients and in the training of future practitioners. In addition, because clinical requirements emphasize such requisites as psychological-mindedness, verbal and cognitive ability, and stable life situation, psychoanalysis may be unduly restricted to a diagnostically, socioeconomically, or intellectually advantaged patient population. Other intrinsic issues pertain to the use and misuse of its stringent rules, whereby overemphasis on technique may interfere with an authentic human encounter between analyst and patient and to the major long-term risk of interminability, by which protracted treatment may become a substitute for life.

Ms. M, a 29-year-old, unmarried woman who worked in a low-level capacity for a magazine, presented for consultation with the chief complaints of considerable sadness and distress over her parent's reaction when they found that she had had a lesbian relationship. She also realized that she had been working far below her potential. She had never sought any treatment before. She was clearly intelligent, sensitive, self-reflective, and insightful. When the possibility of psychoanalysis was presented to her, she worried that meant she was "sicker." Ms. M, however, began reading Freud, realized that analysis was actually recommended for those who are higher functioning, and became intrigued by the idea. She agreed to come 4 days a week for 50-

minute sessions.

She was the oldest of three children and the only girl. Ms. M's father, a successful professional, was described as very demanding and intrusive, someone who never thought anything was good enough. He had always expected his children to do the "extra credit" assignments as part of their regular work. Ms. M, however, was very proud of her father's accomplishments. She spoke of her mother in conflicting terms as well: She was a homemaker, weak, and sometimes acquiescent to the powerful father but also a woman in her own right who was involved in community volunteer work and could be a powerful public speaker.

Just prior to beginning her analysis with a female analyst, Ms. M had had her wallet stolen. In her first analytic session, she spoke of losing all of her identification cards, and to her it seemed as if she were starting analysis "with a completely new identity." Initially, she was somewhat hesitant to use the couch because she wanted to see her analyst's reactions, but she quickly appreciated that she could associate more easily without seeing the analyst.

As her analysis proceeded, through dreams and free associations, Ms. M became quite focused on the analyst. She became extremely curious about the analyst's life. What emerged from her associations to seeing the analyst's appointment book on the desk was that she felt "slotted in." Whenever Ms. M saw other patients, she felt the office was "like an assembly line." Further associations led to her feeling slotted in by her parents as they ran from one activity to another. Her resistance manifested itself in Ms. M's often coming as much or more than 15 minutes late to her sessions. Her associations led to her admitting that she did not want her analyst to think that she was "too eager." Ms. M was able to see that she needed to devalue her analyst and her importance to Ms. M as a defense against an overwhelming positive and even erotic transference toward her. For example, Ms. M wanted to improve her appearance so that her analyst, who was a "role model" for Ms. M, would find her more attractive. Her negative transference, however, was never far from the surface, and she denigrated the analyst by wondering if the analyst were a "clotheshorse" who was financing her wardrobe by the patient's payments.

Her conflicts about her sexual orientation were a central preoccupation in the course of her analysis, particularly because her father was so homophobic. Early on, Ms. M felt awkward and uncomfortable when she went to a lesbian bar, and when asked if she qualified for the "lesbian discount," she said she did not. At one point, she began seeing several men, including a male psychologist. The analyst made the transference interpretation, which Ms. M accepted, that a date with this psychologist seemed as if it were a date with her analyst and sleeping with him would somehow be a stand-in for sleeping with her analyst. Ms. M was also able to see that her transient choice of dating a male therapist was a defensive compromise. Although her homosexual object choice was multidetermined, Ms. M came to appreciate, through her work in analysis that at least a part of her conflicts about homosexuality stemmed from her relationship toward her father. It was a means of securing his attention as well as infuriating him.

Over the course of 4 years, Ms. M began to do considerably better at work and was promoted to a job commensurate with her potential. She was also able to deal better with both her parents, and particularly her father, regarding her sexual orientation. She became much more comfortable with her "new identity" and became involved in a relationship with a professional woman. At the end of therapy, Ms. M and this woman were committed to each other and were thinking of adopting a child. Years later, she wrote to her analyst to say that she and her partner had decided to marry in the state where they both lived.

PSYCHOANALYTIC PSYCHOTHERAPY

Freud had used the terms "psychoanalysis" and "psychotherapy" interchangeably at various times throughout his writings, as mentioned, and many of his so-called "analytic" cases would be considered psychoanalytic psychotherapy by virtue of their length by today's standards and their failure to meet the standards of "classical analysis" as defined by generations after Freud. Nevertheless, Freud had a very clear idea of differentiating psychoanalysis from all other techniques that incorporated suggestion. In fact, he wrote, "There was the greatest possible antithesis between suggestive and analytic technique." In 1919, he had written his famous dictum, "It is very probable . . . that the large-scale application of our therapy will compel us to alloy the pure gold of analysis freely with the copper of direct suggestion."

Yet, although psychoanalytic psychotherapy does go beyond the boundaries of classic practice by permitting analytic techniques to be modified and combined with nonanalytic techniques, it is no longer necessarily considered an uncontrolled or contaminated treatment based solely on suggestion. Rather, modifications offer acceptable and appropriate applications of the analytic armamentarium in keeping with the socio-cultural climate and temper of the times. The early-20th-century Victorian *Weltanschauung* expressed in an age of sexual repression that made hysterical neurosis the consummate disorder of its day has been superseded by the current culture of narcissism, with its more modern manifestations of mental illness, such as the narcissistic and borderline personality disorders and other ego disturbances. At the same time, current economic constraints and the effect of public policy on the nature of psychiatric treatment have enforced practical considerations in the design of revised analytic techniques to meet a broader spectrum of patients and problems. The result has been an array of shorter-term psychodynamic strategies in addition to long-term psychoanalysis.

Psychoanalytic, also called "dynamic," "exploratory," or "expressive" psychotherapy, however, was always grounded in the principles of psychoanalysis. The psychoanalyst Robert Wallerstein reported that dynamic psychotherapy, as it developed in the post-World War II period, was the "one distinctly American contribution" to modern-day psychiatry.

Distinguishing Psychoanalysis from Psychoanalytic Psychotherapy

As derivatives of classic psychoanalysis, the techniques of psychoanalytic psychotherapy represent modifications that both resemble and depart from it. As broadly practiced today, psychoanalytic treatment encompasses a wide range of uncovering strategies used in varied degrees and blends. Despite the inevitable blurring of boundaries in actual application, the

original modality of classic psychoanalysis and major modes of psychoanalytic psychotherapy (i.e., expressive or exploratory or dynamic, and supportive) are delineated separately here (Table 33.1–2). Analytic practice in all of its complexity resides on a continuum. Individual technique is always a matter of emphasis as the therapist titrates the treatment according to the needs and capacities of the patient at every moment. Psychoanalytic psychotherapy is based on fundamental dynamic formulations and techniques that derive from psychoanalysis and is designed to broaden its scope. Arnold Cooper notes that one of the major distinguishing features of psychoanalysis from other analytically-oriented therapies is the degree of internalization of the therapist and the process of therapy after a patient has terminated. For example, he finds that after therapy, a patient will remember it vividly as a powerful and important experience, whereas after an analysis, a patient tends to remember it poorly, reflecting the analytic experience has become an integrated part of the patient, rather than sequestered as a separate internal representation.

Gabbard notes that the strategies of psychoanalytic dynamic psychotherapy range from exploratory or expressive (e.g., insight-oriented clarification, confrontation, and interpretation) techniques to supportive (e.g., relationship-oriented suggestion, empathic validation, advice and praise, and affirmation) techniques. Although those two types of methods are sometimes regarded as antithetical, their precise definitions and the distinctions between them are by no means absolute.

The primary distinction between psychoanalysis and other forms of therapy, though, is in the handling of the transference. It follows that the pivotal clinical issue for psychoanalytic psychotherapy is how the transference is used or managed, to what extent it is interpreted or left uninterpreted, and to what extent it is fostered by the therapist. In psychoanalysis, all transference manifestations are analyzed. Robert Michels has noted that in psychotherapy, as distinct from psychoanalysis, the transference is analyzed primarily when it becomes a source of resistance, that is, when it interferes with the progress of the therapy.

Psychoanalyst Lawrence Friedman, whose classic work, *The Anatomy of Psychotherapy,* believes that the defining role of the therapist is his "commitment to ambiguity." But he adds, "If there were no theory behind it, the psychotherapy session would be a very strange meeting." At least one source noted that all the controversies among analysts about what constitutes analysis (e.g., how many times a week, whether to use the couch, what theory to use) is analogous to being concerned about reshuffling the deck chairs on the Titanic when the ship is going under. In other words, maybe the focus has been misdirected. Cooper maintained that there is often not a clear boundary between psychoanalytic psychotherapy and psychoanalysis and said, "I had the not infrequent experience that some of my sitting up so-called psychotherapy patients

were experiencing an analytic process and that some of my lying down five times a week psychoanalytic patients were at best having a psychotherapy." He added, "The heart of analysis cannot be a matter of how many times a week we see the patient and the patient's posture."

In support of that idea, Robert Wallerstein had demonstrated, years ago in the Psychotherapy Research Project of the Menninger Foundation, that even in psychoanalytic treatment, much of what is called structural change came from supportive, noninterpretative means. His conclusion was that the psychotherapies accomplished much more than initially expected and "real treatments in actual practice are inextricably intermingled blends of more or less expressive and more or less supportive elements." Despite the potential for intermingling blends, says T. Byram Karasu, "Psychotherapeutic communication begins only when ordinary communication ends."

Thirty years after delineating four intrinsic criteria of psychoanalysis (centrality of transference, neutral analyst, regression to transference neurosis, and interpretation as sole instrument), Merton Gill revised them in 1984 to provide the backbone of modern psychoanalytic psychotherapy: (1) The centrality of transference is broadened to encompass nontransference manifestations; (2) the guiding principle of neutrality is loosened to acknowledge the real personality and attitudes of the analyst; (3) the regressive transference neurosis is supplanted or replaced by less frustration, fantasy, and focus on the past (and greater gratification, reality, and focus on the present); and (4) interpretation is extended to include more nonexploratory methods (with genetic interpretations expanded by here-and-now interpretations). Those intrinsic changes accompanied three major extrinsic changes: reduced session frequency, reduced or more flexible treatment duration, and abandonment of the couch.

The duration of psychoanalytic psychotherapy is generally shorter and more variable than in psychoanalysis. Treatment may range from brief, even with an initially agreed-on or fixed time limit (e.g., Mann's time-limited psychotherapy), to a less definite number of months or years. Brief treatment is chiefly used for selected problems or highly focused conflict, whereas longer treatment may be applied in more chronic conditions or for intermittent episodes that require ongoing attention to deal with pervasive conflict or recurrent decompensation. In either event, when a time frame is established at the beginning, it can often be altered in the course of therapy as patient needs and goals change. Unlike psychoanalysis, psychoanalytic psychotherapy rarely uses the couch; instead, patient and therapist sit face-to-face. That posture helps to prevent regression because it encourages the patient to look on the therapist as a real person from whom to receive direct cues, even though transference and fantasy will continue. The couch is considered unnecessary because the free association method is rarely used except when the therapist wishes to gain access to fantasy material or

dreams to enlighten a particular issue. In current practice the use of the couch may not even be advisable in many instances (especially in the supportive type) because extensive unconscious data are not sought or because the primitive thinking or behavior that may accompany severe illness makes further regression undesirable.

Table 33.1–2.
Scope of Psychoanalytic Practice: A Clinical Continuum

Features	"Classic" Analytic	Expressive Mode (Exploratory)	Supportive Mode
Frequency	Regular 4–5 times/wk; "50-minute hour"	Regular 1–3 times/wk; $^1/_2$ to full hour	Flexible 1 time/wk or less; or as needed $^1/_2$ to full hour
Duration	Long term; usually 3–5+ years; maybe considerably longer	Short or long term; several sessions to months or years	Short or intermittent long term; single session to lifetime
Setting	Patient primarily on couch with analyst out of view	Patient and therapist face to face; occasional use of couch	Patient and therapist face-to-face; couch contraindicated
Modus operandi	Systematic analysis of all positive and negative transference and resistance; primary focus on analyst and intrasession events; transference neurosis facilitated; regression encouraged	Partial analysis of dynamics and defenses; focus on current interpersonal events and transference to others outside of sessions; analysis of negative transference; positive transference left unexplored unless it impedes progress; limited regression encouraged	Formation of therapeutic alliance and real object relationship; analysis of transference contraindicated with rare exceptions; focus on conscious external events; regression discouraged
Analyst/therapist role	Absolute neutrality; frustration of patient; reflector/mirror role	Modified neutrality; implicit gratification of patient and greater activity of therapist	Neutrality suspended; limited explicit gratification, direction, and disclosure
Mutative change agents	Insight predominates within relatively deprived environment	Insight within more empathic environment; identification with benevolent object	Auxiliary or surrogate ego as a temporary substitute; holding environment; insight to the degree possible
Patient population	Neuroses; mild character psychopathology	Neuroses; mild to moderate character psychopathology,	Severe character disorders, latent or manifest psychoses, acute crises,

		especially narcissistic and borderline disorders	physical illness
Patient requisites	High motivation, psychological mindedness, good previous object relationships, ability to maintain transference neuroses, good frustration tolerance	High moderate motivation and psychological-mindedness, ability to form therapeutic alliance, some frustration tolerance	Some degree of motivation and ability to form therapeutic alliance
Basic goals	Structural reorganization of personality; resolution of unconscious conflicts; insight into intrapsychic events; symptom relief an indirect result	Partial reorganization of personality and defenses; resolution of preconscious and conscious derivatives of conflicts; insight into current interpersonal events; improved object relations; symptom relief a goal or prelude to further exploration	Reintegration of self and ability to cope; stabilization or restoration of pre-existing equilibrium; strengthening of defenses; better adjustment or acceptance of pathology; symptom relief and environmental restructuring as primary goals
Major techniques	Free association method predominates; full dynamic interpretation (including confrontation, clarification, and working through), with emphasis on genetic reconstruction	Limited free association; confrontation, clarification, and partial interpretation predominate, with emphasis on here-and-now interpretation and limited genetic interpretation	Free association method contraindicated; suggestion (advice) predominates; abreaction useful; confrontation, clarification, and interpretation in the here-and-now are secondary; genetic interpretation is contraindicated
Adjunct treatment	Primarily avoided; if applied, all negative and positive meanings and implications thoroughly analyzed more recently, medication adjunct	May be necessary, e.g., psychotropic drugs as temporary measure; if applied, negative implications are explored and diffused	Often necessary, e.g., psychotropic drugs, family rehabilitative therapy, or hospitalization; if applied, positive implications are emphasized

This division is not categorical; all practice resides on a clinical continuum.

A 32-year-old woman had been in twice-weekly psychotherapy for about a year when her therapist presented the possibility of her coming into analysis with him. It was just around the

time that the therapist had gotten a large ottoman for the chair his patients used. This patient, who became quite frightened and confused by the idea of switching her therapy to analysis, felt this new furniture represented a "neither here nor there"—that is, "neither a couch nor a chair." She chose not to enter analysis but did continue her therapy. She was never was able to use the ottoman.

Mutative Change Agents

Insight. *Insight* has been defined as the process by which the meaning, significance, pattern, or use of an experience becomes clear or as the understanding that results from that process. Theoretically, it has been portrayed as occurring in four successive stages, which can be applied to the psychoanalytic process as follows: (1) preparation, characterized by frustration, anxiety, feelings of ineptness, and despair, often accompanied by trial-and-error activity and falling into habitual patterns (repetition compulsion); (2) renunciation, in which one desires to escape from the problem or is unmotivated to make insightful efforts (resistance, negative transference); (3) inspiration or illumination, in which the problem is grasped and solutions suggest themselves (beginning of discovery based on interpretive process); and (4) elaboration and evaluation, in which the validity of the insight is checked and confirmed against external reality (working through).

Insight is not fully understood, and it is frequently misidentified as a dramatic "eureka" phenomenon. Such sudden enlightenment rarely occurs in analysis, and if it does, it is usually short lived but may mislead the analyst or the patient into wishfully believing that profound understanding has occurred. It is much more common for the patient to achieve insight in a slow and subtle manner, in small ripples rather than in sudden tidal waves. Insights also tend to be circumscribed and specific to certain problem areas rather than whole-truth revelations that are associated with mystical experiences or religious conversions.

Despite its presumed potency as a therapeutic agent, insight in psychoanalysis is not definitive in its role or its function. There is no proof of a necessary relation between the truth (or falseness) of insight and therapeutic results, although that finding may be clinically obviated by the role of mutual belief as a nonspecific healing agent. What is clinically known, however, is that intellectual insight alone is of minimal value; insight must be emotionally integrated (even though the distinction may be difficult to make). Moreover, not all change in psychoanalysis is attributable to insight, and not all insight leads to behavioral change. Perhaps the major controversy that characterizes contemporary thinking on the subject is the mutative role of insight versus the role of the therapeutic atmosphere or the empathic relationship within which the interpretation occurs. In his 1953 classic paper that introduced the term "parameter," Kurt Eissler wrote, "It is still a riddle why a human being

should refuse to make maximal use of insight which is conveyed to him."

Identification and Empathy. Insight is still regarded as the sine qua non of analytic cure, and the interpretation of transference remains the primary vehicle through which mutative change occurs, but greater attention in recent years has been paid to the role of the therapeutic atmosphere in facilitating insight and to the therapeutic relationship itself. The traditional concern has been more with recreating the original infantile neurosis than with the ameliorating qualities of the analyst as an ideal or real person who is different from desired past objects.

Identification is an automatic, unconscious process that occurs when an individual takes on the qualities of a substitute object or model. This phenomenon—an early expression of an emotional tie with another person—is considered a consequence of the desire to believe in the benevolent power of another, the tendency both consciously and unconsciously to want to be like the loved or admired object. As the therapist's values are presumably more realistic and constructive than those of the patient, repeated identification with the image of the therapist serves not only to lessen the harsh quality of the patient's superego, but it can also modify the patient's ego with the therapist as role model, increasing self-esteem through identification with the therapist. The therapist's role as an identification model is facilitated, at least in part, by his or her elevated transferential image as an idealized authority.

In recent years, there has been an emphasis on the important role of the therapeutic atmosphere, as expressed in the concept of empathy, defined by Auchincloss and Samberg as "a complex affective and cognitive process of feeling, imagining, thinking, and somatically sensing one's way into the experience of another person." Kohut had focused on the importance of empathy in his theory of self-psychology and defined it as "vicarious introspection." For Kohut, major failures of empathy by parental figures during childhood interfered with the crystalization of the self. As the therapist communicates an understanding of the patient's problems and psychological pain (i.e., an empathic resonance), the patient is more likely to gain insight. Current research suggests that even the most accurate and well-timed interpretation will not be effective if the patient does not experience such an empathic therapeutic atmosphere.

The complex question of mutative change agents in psychotherapy is unresolved. Investigations by Jerome Frank, Hans Strupp, T. Byram Karasu, and others suggest that the critical ingredients of psychotherapeutic effectiveness may transcend specific strategies or types of treatment and to a large extent reside in nonspecific ingredients that all therapies share, such as the personality, commitment, and availability of the therapist.

Adjunctive Therapy and Concomitant Use of Medication in Psychotherapy

Another issue is the use of adjunct treatment during psychoanalytic psychotherapy. Patients may present with more severe pathology in psychoanalytic psychotherapy and may require concurrent treatment with psychotropic drugs to relieve acute symptoms of anxiety or depression before any dynamic exploration can ensue. Just as in the concomitant use of medication with psychoanalysis, there may be similar complex transferential and countertransferential issues in the context of the analytically-oriented psychotherapies. Because these treatments lack the depth of psychoanalytic exploration, therapists need to be particularly aware of their own, as well as their patients' fantasies, regarding the use of medication during the course of therapy. A 2014 article by Huhn et al. conducted a systematic overview of the efficacy of pharmacotherapy and psychotherapy, with 61 meta-analyses and over 137,000 participants. Because patients often benefit from both modalities, the authors recommended that future research should focus on how both modalities can be combined "to maximize synergy" of the two rather than debate which treatment is more efficacious. The authors found that direct comparisons of medication and psychotherapy did not demonstrate consistent differences but their combination was superior to either alone. A survey of psychotherapists (61 psychiatrists, and over 1900 patients in treatment) in the New York/New Jersey areas found that so-called "split care"—care by two professionals, one of whom does the medication management and the other does the psychotherapy—was common. Regrettably, though, there was often poor, irregular, or infrequent communication about the patient between the professionals involved.

With the analytically-oriented therapies, depending on the nature of the patient's pathology and support system, other treatments may also be necessary, including hospitalization, rehabilitative therapies, couple or family therapy.

Exploratory (Expressive) Psychotherapy

Indications and Contraindications. Exploratory psychotherapy has become a desirable and widely applicable form of treatment under the following circumstances: (1) Psychoanalysis is diagnostically or clinically contraindicated because the patient lacks sufficient ego strength to tolerate the extent of regression, frustration, or suspension of reality that is required; the patient is deficient in the cognitive resources necessary for the achievement of deep insight; the patient is not sufficiently motivated because of competing family, social, or cultural influences; or the problem is so pressing that lengthy treatment is unsuitable. (2) Practical or logistical considerations prevail whereby patients otherwise suited for a more lengthy psychoanalysis are unavailable for long-term treatment for reasons

of time, money, or geography. For that group the techniques used are closer to those of psychoanalysis proper because extrinsic rather than intrinsic factors are largely responsible for the change of approach. Nonetheless, the category of healthier patients constitutes a sizable population of recipients of current psychoanalytic psychotherapy. (The inclusion of both categories here in part accounts for the blurred boundaries and confusion in the application of expressive psychotherapy.)

Diagnostically, psychoanalytic psychotherapy in its mode is suited to a range of psychopathology with mild to moderate ego weakening, including neurotic conflicts, symptom complexes, reactive conditions, and the whole realm of nonpsychotic character disorders. The last includes those disorders of the self that are among the more transient and less profound on the severity-of-illness spectrum, such as narcissistic personality disorders. It is also the treatment controversially recommended for patients with borderline personality disorders, although special variations may be required to deal with the associated turbulent personality characteristics, primitive defense mechanisms, tendencies toward regressive episodes, and irrational attachments to the analyst. Depending on the level of pathology, however, narcissistic and borderline personalities may be amenable, as mentioned, for psychoanalysis. Sometimes a trial of therapy is indicated to assess a patient's suitability. There was considerable discussion among many psychoanalysts and psychotherapists about one of the very last episodes of the popular television program, *The Sopranos*. In this particular episode, Dr. Melfi, the long-time therapist of mobster Tony Soprano, summarily dismisses him from his psychoanalytic psychotherapy after hearing about a study that indicated sociopaths did not benefit from psychotherapy. Glen Gabbard, who has used the program to teach psychotherapy, has written extensively on *The Sopranos* and discussed the complex issues involved in treating "a man who is beyond redemption" in his book, *The Psychology of the Sopranos*.

The persons best suited for the exploratory psychotherapy approach have fairly well integrated egos and the capacity both to sustain and to detach from a bond of dependency and trust. They are, to some degree, psychologically minded and self-motivated, and they are generally able, at least temporarily, to tolerate doses of frustration without decompensating. Patients must have some capacity for introspection and impulse control, and they should be able to recognize the cognitive distinction between fantasy and reality. Often, they may appear similar to those suitable for psychoanalysis. Sometimes, logistical issues, such as time constraints and ability to commit to a lengthy, intense process, determine the method of treatment.

Goals. The overall goals of exploratory psychotherapy are to increase the patient's self-awareness and to improve relationships with people in the

patient's life through exploration of current interpersonal events and perceptions. In contrast to psychoanalysis, major structural changes in ego function and defenses are modified in light of patient limitations. Instead of systematically resolving the unconscious nuclear conflict, the therapist may opt to resolve some conflict areas and undo specific resistances, intentionally overlooking or reinforcing others. The aim is a more limited, and thus select and focused, understanding of one's problems. Rather than uncovering deeply hidden and past motives and tracing them to their origins in infancy, the major thrust is to deal with preconscious or conscious derivatives of conflicts as they became manifest in present interactions. Although insight is sought, it is less extensive; instead of delving to a genetic level, there is a greater emphasis on clarifying recent dynamic patterns and maladaptive behaviors in the present. Symptom relief may be an acceptable aim rather than simply a concomitant of the better resolution of conflicts. In more integrated patients who have the capacity for greater insight, it may be a prelude to further analytic work. Much like with analysis, though, both patient and analyst can divide goals, as Ernst Ticho has noted, into life goals and treatment goals.

The Main Approach. Facilitating a full-blown, regressive transference neurosis is neither necessary nor appropriate. Rather, the major modus operandi involves establishment of a therapeutic alliance and early recognition and interpretation of negative transference. Only limited or controlled regression is encouraged, and positive transference manifestations are generally left unexplored unless they are impeding therapeutic progress; even here, the emphasis is on shedding light on current dynamic patterns and defenses.

Although technical neutrality is largely maintained, the therapist is less strictly anonymous and silent and is more active and responsive in his or her overall stance and in the nature of specific interventions. The free association method is no longer used as the major patient tool, and primary process material is elicited only as a selective way of accessing specific conflictual data. Clarification and interpretation are still used, but they are altered both qualitatively and quantitatively. Partial interpretations address some conflict areas as others are left untouched; the depth of interpretation is less penetrating, with emphasis on the present.

Limitations. Some therapists fail to accept the limitations of a modified insight-oriented approach and so apply it inappropriately to modulate the techniques and goals of psychoanalysis. Overemphasis on dreams and fantasies, zealous efforts to use the couch, indiscriminate deep interpretations, and continual focus on the analysis of transference may have less to do with the patient's needs than with those of a therapist who is unwilling or unable to be flexible.

Ms. S was an attractive 30-year-old, unmarried woman working as a secretary when she presented for consultation. Her chief complaints at the time were feeling "only anger and tension" and an inability to apply herself to studying voice, "which is one of the most important things to me."

In obtaining a history, the therapist noted that Ms. S had never completed anything: She had dropped out of college; never pursued a music degree; and switched from job to job, and even city to city. What initially seemed like a woman with diverse interests (e.g., jobs as a research assistant, freelance copyeditor, part-time radio announcer; manager of data entry for a software company; and, most recently, secretary) really reflected a woman with a chaotic lifestyle and serious difficulties committing to anyone or anything. Although obviously intelligent, Ms. S presented with unrealistic expectations regarding her consultation. For example, after the first consultative session, Ms. S said she felt good afterward but felt there were "no revelations yet." Because of Ms. S's inability to commit and her somewhat disorganized life, the therapist recommended a course of psychotherapy, beginning twice a week, rather than something more intense like psychoanalysis. The therapist also realized over the course of the consultation that Ms. S would have difficulty with free association without getting disorganized. The therapist thought that Ms. S might regress unproductively on the couch without visual contact with the therapist.

Ms. S was the second oldest of four children—two brothers and a younger sister, with whom she was most competitive and who clearly seemed the mother's favorite. She described her mother as a successful professor who was demanding and critical, as if she had a "raised eyebrow" in disapproval. For example, much to her mother's chagrin, Ms. S had once wanted a sandwich "with everything on it." Ms. S was also disappointed when she was given one piece of luggage rather than a complete set for a Christmas gift. She was able to accept the therapist's interpretation that she felt "part of a set" by being one of four siblings. Ms. S initially idealized her father, who was active in the family church, but eventually saw him as disappointing and rejecting.

Ms. S's ideal therapist would be "flexible," by which she meant a therapist who might do hypnosis one session, psychotherapy the next, and, maybe, analysis another session. In fact, within the first week of beginning therapy, Ms. S had simultaneously consulted a hypnotherapist, which she mentioned in passing only weeks later, for her neck pain and tension. Although she did not pursue hypnosis, she did maintain a chiropractor for most of her therapy, also something she mentioned in passing many months after beginning therapy. She did speak of wanting to be "on best behavior" and "follow the rules." Her tremendous sense of entitlement, however, was evident: She had an expectation of getting "cut-rate prices" on everything from haircuts and car repairs to doctors' visits. Her initial fee was a much-reduced one, which she paid late and begrudgingly throughout most of her therapy. During one session, for example, she held up her hands admiringly, delightedly noting the bargain she had gotten for her $7.00 manicure. She was oblivious to the transference connection, which the therapist pointed out, that she was paying the same price for each therapy session.

Although she was seen only twice a week, Ms. S developed intense feelings for her therapist. Mostly she experienced rage when she saw evidence of the therapist's other patients, such as footprints on the waiting room floor after a snowstorm or a coat hanger turned around. She expressed the wish to keep some of her things, like bobby pins and hairspray, in the therapist's bathroom. She vacillated between feelings she wanted to move in and feelings that the therapist did not exist. For example, before she took a plane flight, she wondered who would tell her therapist if something happened to her. She had never given the therapist's name to anyone, nor did she have her name in her weekly appointment book. The therapist interpreted that she had a wish to devalue her and not to share her with anyone else. Associations to a dream with an image of a string of Baroque pearls led to thoughts that these pearls—irregular and imperfect—defective and even lopsided, represented how she viewed herself.

Over the course of the next few years, Ms. S was able to commit to coming regularly to therapy, although the course was somewhat tumultuous, with many threats of quitting and much withholding of information. At one point, she even tried to provoke the therapist by seeking a consultation with another therapist in order to "tattle" on her, just as she had tattled on her siblings. Her therapist remained unprovoked and continued to provide a safe environment for Ms. S to explore her ambivalence to the therapist and the therapeutic

situation. The therapist was also able to contain Ms. S's tendency to regress, particularly with separations, by providing her with the therapist's telephone number.

She had actually entered therapy with a wish to become a world-famous musician who would win her mother's approval and praise. Her narcissism and sense of entitlement made it difficult for her to give up on that fantasy despite repeated evidence that she did not have sufficient talent. She was finally able to settle on a compromise: She began to work diligently and closely as a research assistant to her mother, who was writing a book, and as Ms. S became more focused and organized over time, she even thought she might write a book about the Church.

Supportive Psychotherapy

Supportive psychotherapy aims at the creation of a therapeutic relationship as a buttress or bridge for a patient. It has roots in virtually every therapy that recognizes the ameliorative effects of emotional support and a stable, caring atmosphere in the management of patients. As a nonspecific attitude toward mental illness, it predates scientific psychiatry, with foundations in 18th-century moral treatment, when, for the first time, patients were treated with understanding and kindness in a humane interpersonal environment free from mechanical restraints.

That perspective underlies such diverse developments as milieu therapy for rehabilitating chronic hospitalized patients; crisis intervention to assist otherwise functioning persons through sudden periods of great turmoil or stress; and guidance or counseling practices for children, former patients, and nonpatients who need interim help in social, academic, or vocational areas but do not warrant long-term or in-depth exploration. Supportive psychotherapy has been the chief form used in the general practice of medicine and rehabilitation, frequently to augment other therapeutic measures such as prescriptions of medication to suppress symptoms or for hospitalization to provide a structured therapeutic environment, protection, and control of the patient. It may be applied as primary or ancillary treatment. The global perspective of supportive psychotherapy (often part of a combined-treatment approach) places major etiological emphasis on external rather than intrapsychic events, particularly on stressful environmental and interpersonal influences on a severely damaged self.

As a viable modality within the psychoanalytic movement, supportive psychotherapy has been described as the most ill-defined and nebulous of all psychotherapies. It is the treatment to which very difficult, characterologically or intellectually limited patients were referred when no other modality, particularly insight-oriented psychotherapy, seemed suitable. In the late 1940s, Alexander and French noted the persistent tendency to differentiate between two main categories of treatment—primarily supportive methods and primarily uncovering methods.

Relationship to Psychoanalysis. With the widening scope of

psychoanalysis, supportive psychotherapy has emerged as a specific body of techniques. The contributions of modern ego psychology and object relations theory, which have broadened the focus and goals of psychoanalytic psychotherapy beyond classic psychoanalysis, have generated renewed interest in defining the role of supportive psychotherapy in analytic work. Supportive strategies include an emphasis on suggestion, persuasion, encouragement, reassurance, advice-giving, reality testing, or environmental manipulation, in contrast to or in addition to insight-oriented strategies. Caretaking in the form of a positive relationship provides the basic context or milieu for listening to and understanding patients (no matter what specific strategies are used). Not only the actual methods, but the entire atmosphere and nature of communication are characterized by support and safety. This method provides the patient a relationship where the therapist is an auxiliary or substitute ego that offers the patient a secure and nurturing interpersonal relationship, as well as specific control, direction, and counsel.

In the psychoanalytic spectrum, supportive psychotherapy represents not so much a separate entity as an important emphasis within its exploratory-supportive boundaries. Instead of being negatively viewed as a compromised approach for patients unsuitable for analysis, it is positively viewed as effective treatment for a broad clinical range of patients. Concern and empathy are integral to all therapeutic endeavors, but the focus here is on a stance or interpersonal bond that is affectively responsive (rather than neutral) and oriented to present reality (rather than a focus on the past) and in which limited interpretive work in the broadest sense can more readily occur. Many therapists now explicitly recognize the complementarity of exploratory and supportive modes of treatment, which can be combined, modified, and individualized according to the needs of the patient. The therapeutic focus is more directly on the caring and nurturing aspect of the doctor–patient relationship (without analyzing it as a past transference manifestation).

Indications and Contraindications. Supportive psychotherapy is generally indicated for those patients for whom classic psychoanalysis or insight-oriented psychoanalytic psychotherapy is typically contraindicated —those who have poor ego strength and whose potential for decompensation is high. Amenable patients fall into the following major areas: (1) persons in acute crisis or a temporary state of disorganization and inability to cope (including those who might otherwise be well-functioning), whose intolerable life circumstances have produced extreme anxiety or sudden turmoil (e.g., persons going through grief reactions, illness, divorce, or job loss or who were victims of crime, abuse, natural disaster, or accident); (2) patients with chronic severe pathology, with fragile or deficient ego functioning (e.g., those with latent psychosis,

impulse disorder, or severe character disturbance); (3) patients whose cognitive deficits and physical symptoms make them particularly vulnerable and so unsuitable for an insight-oriented approach (e.g., certain psychosomatic or medically ill persons); and (4) persons who are psychologically unmotivated although not necessarily characterologically resistant to a depth approach (e.g., patients who come to treatment in response to family or agency pressure and are interested only in immediate relief or those who need assistance in some very specific problem areas of social adjustment as a possible prelude to more exploratory work).

Supportive psychotherapy is usually used when other complications enter the clinical picture, such as when the primary diagnosis is physical illness but the patient needs help in dealing with disability, or for long-term mental illness, in which intermittent support and guidance in everyday living augment the control of symptoms through pharmacotherapy and provide out-of-hospital management. Regardless of the diagnosis, a supportive approach may also be used in the early phases of virtually all treatments as preparation for the establishment of a firm therapeutic alliance (even when expressive treatment is the predominant mode) or intermittently whenever the patient is in danger of excessive regression.

Because support forms a tacit part of every therapeutic modality, it is rarely contraindicated as such. The typical attitude regards better-functioning patients as unsuitable, not because they will be harmed by a supportive approach, but because they will not be sufficiently benefited by it. In aiming to maximize the patient's potential for further growth and change, supportive therapy tends to be regarded as relatively restricted and superficial, and thus it is not recommended as the treatment of choice if the patient is available for and capable of a more in-depth approach.

Goals. Supportive psychotherapy has more limited objectives than does insight-oriented psychotherapy because the patient's psychopathology and diminished psychological resources restrict the potential for major intrapsychic change or growth. Because the patient may begin supportive therapy in a decompensated state, the therapist (at least initially) aims at reconstituting and stabilizing the patient's function. Although the individual aims of supportive treatment vary, the general thrust is toward the amelioration or relief of symptoms through behavioral or environmental restructuring within the patient's existing psychological framework. That often means helping the patient to adapt better to problems and to live more comfortably with his or her psychopathology. To restore the disorganized, fragile, or decompensated patient to a state of relative equilibrium, the therapist has the major treatment goal of suppressing or controlling symptomatology and of stabilizing the patient in a protective and reassuring benign atmosphere that militates against overwhelming external and internal pressures. The ultimate treatment goal

is to maximize the integrative or adaptive capacities so that the patient increases the ability to cope while decreasing vulnerability by reinforcing assets and strengthening defenses.

The Main Approach. The techniques of supportive psychotherapy are designed to restore or enhance the patient's ego strength by helping the patient to control impulses through direct limit-setting, to gain a more accurate picture of reality through the clarification and testing of perceptions, to sustain the adaptive structure by accepting (not analyzing or confronting) defensive maneuvers, and to develop better coping skills through direct teaching and practical advice and through the use of the therapist as a role model and as a constantly reassuring figure on whom the patient can rely. The active therapist serves as a substitute or auxiliary ego until the patient is able to incorporate those assets in which he or she is largely deficient, either acutely or characterologically. The safety and security of a controlled and consistent therapeutic atmosphere also act as a holding environment to contain the patient's aggression and dangerous impulses as he or she learns better to repress or sublimate the tendencies to act out feelings rather than to verbalize them.

The traditional analytic techniques of clarification and interpretation are also used to the extent that they focus on the here-and-now of the patient's problems, but genetic interpretations are contraindicated and may cause decompensation in an already fragile patient.

Transference management also differs from that of classical psychoanalysis and exploratory psychoanalytic psychotherapy. Supportive psychotherapy focuses on the fostering and maintenance of a positive transference at all times and on vitiating the effects of negative transference should they arise. A complete or partially controlled transference regression is contraindicated because intense transferences that are allowed to develop in ego-impaired patients can produce sudden, turbulent reactions that have disruptive effects on both the patient and treatment. A positive transference is fostered and rarely analyzed as such. Exceptions are those instances when positive transference may lead to acting out, which can be forestalled by drawing attention to the patient's distorted perceptions and projections. The development of negative transference is minimized by the use of highly structured interventions, with the therapist assuming a much more direct and active approach than in expressive psychotherapy. That includes discouraging free associations and fantasies and continually bringing the patient back to reality and to the immediate situation.

Techniques are thus designed to focus on conscious external events and on the therapist as a largely nontransferential figure. To accomplish such aims, the therapist suspends therapeutic neutrality, with much greater direction, disclosure, and gratification offered than would be appropriate in

more uncovering approaches. The therapist as a real object serves to validate the patient's reality testing, diminishes the breakthrough of regressive negative transference, and provides a reliable companion to sustain the patient until he or she can function independently.

Limitations. To the extent that much supportive therapy is spent on practical, everyday realities and on dealing with the external environment of the patient, it may be viewed as more mundane and superficial than depth approaches. These patients are seen intermittently and less frequently, and the interpersonal commitment may not be as compelling on either the patient's or the therapist's part. Greater severity of illness (and possible psychoses) also makes such treatment potentially more erratic, demanding, and frustrating. The need for the therapist to deal with other family members, caretakers, or agencies (e.g., auxiliary treatment and hospitalization) can become an additional complication as the therapist comes to serve as an ombudsman to negotiate with the outside world of the patient and with other professional peers. Finally, the supportive therapist needs to be able to accept personal limitations and the patient's limited psychological resources.

Mr. W was a 42-year-old widowed businessman who was referred by his internist because of the sudden death of his wife, who had had an intracranial hemorrhage, about 2 months earlier. Mr. W had two children, a boy and a girl, ages 10 and 8 years, respectively.

Mr. W had never been to a psychiatrist before, and when he arrived he admitted he was not certain what a psychiatrist could do for him. He just had to get over his wife's death. He was not sure how talking about anything could really help. He had been married for 15 years. He admitted to having difficulty sleeping, particularly awakening in the middle of the night with considerable anxiety about the future. One of his relatives had given him some of her own Klonopin for his anxiety, which helped tremendously, but he feared getting dependent on it. He was also drinking more than he thought he should. He was most concerned about raising his children alone and felt somewhat overwhelmed by the responsibility. He was beginning to appreciate just how wonderful a mother his wife had been and now saw how critical he had been of her for spending so much time with the children. "It really does take a lot of effort," he said.

Mr. W did admit to feelings of guilt. For one thing, he admitted to some sense that he could now start over. He had been somewhat restless in the marriage recently before his wife's death and had actually been unfaithful for a brief period early in the marriage. He also felt some guilt that had he been awake the night of his wife's hemorrhage—maybe he could have saved his wife. In reality, there was nothing he could have done.

Mr. W agreed to come for a few sessions to talk about his wife. At this point, only 2 months after her death, he seemed to have an uncomplicated mourning reaction. Although he talked easily in session, he was clearly worried that he might like "being here too much." The therapist chose not to interpret his dependency conflicts. Mr. W seemed to have good coping skills and used humor as a high-functioning defense. For example, in giving a eulogy for his wife (who had been a very popular member of the congregation), he looked around at the enormous crowd of people at the church service and said he had never seen so many people attending church before, adding, "Sorry, Reverend."

After about four sessions, Mr. W said he that felt better and no longer saw the need for further sessions. He was sleeping better and had stopped drinking excessively. The therapist suggested that he might want to continue to talk more about his guilt and his life as he went forward without his wife. The therapist was reassuring that there seemed to be nothing else

Mr. W could have done to save his wife. He also encouraged the patient to begin dating when he felt ready, something that Mr. W's in-laws were clearly not encouraging. For now, however, Mr. W was not interested in any further therapy. He was appreciative of the therapist and felt that talking about his wife's death had been helpful. The therapist accepted his wish to discontinue their sessions but encouraged Mr. W to keep in touch to let him know how he was doing.

Recent Developments/Variants on Psychoanalytically Oriented Psychotherapy

With the widened spectrum of more severely ill patients, therapists have decreased the frequency of sessions to influence the nature of the transference and control excessive regression and dependency. Otto Kernberg and his colleagues have devised an analytically oriented treatment for borderline and narcissistic patients who manifest severe breakdown in their intimate, social, and work relationships and who have a lack of a coherent self (i.e., identity diffusion). Their model, *transference-focused psychotherapy,* is based on a psychoanalytic object relations theory, originally derived from the work of British analyst Wilfred Bion, as well as from their work done years before at the Menninger Foundation. In transference-focused therapy, patients are seen sitting up at a frequency of two-to-three times a week for many years and encouraged to free-associate. The major techniques are interpretation, an emphasis on the transference as it relates to the patient's current problems in external reality, and therapist neutrality. The therapeutic focus is on the here-and-now, rather than an interpretation of the past. Though it is a manualized technique, Kernberg emphasizes it cannot be learned from a book. He recommends ongoing supervision, even for senior clinicians, for these particularly challenging patients, who can treat their therapists with disdain, disregard, and distrust. Some of these patients are even "serial killers," as it were, of therapists, in that they blame the therapist for their difficulties and fire one therapist after another, with an "unconscious sense of triumph at having killed off another therapist." Limit setting and involvement of the family become crucial when working with these patients as well.

Kernberg has found his approach works better than supportive therapy for patients with severe character pathology such as borderline personality disorders for several reasons: ignoring these patients' impulses only increases their fear of their impulses and encourages acting out, failing to analyze their severe negative transferences and paranoid dispositions has disruptive effects on the therapeutic relationship and pre-empts any efforts to provide a stable, reliable, and empathic atmosphere, and gratifying their transference needs significantly distorts those patients' perception of the therapeutic situation (i.e., they identify with an overly idealized version of their analyst that they feel they are not capable of living up to). Kernberg's approach is controversial and many therapists still do advocate a supportive approach for those patients.

Fonagy and colleagues have devised a form of psychodynamically-

oriented psychotherapy *mentalization-based therapy,* initially for borderline patients but now expanded to treat other personality disorders. Mentalization is the ability to understand one's own mind and the mind of others. Mentalizing is a developmental process that flourishes in the context of trusted caretakers who are sensitive to the needs of a growing child and hence foster a secure sense of attachment in the child. In this therapy, patients are seen twice a week and sitting up. The therapist encourages patients who are defective in their ability to mentalize to reflect on their own behavior as motivated by their mental states and those of others. Mentalizing, then, is a means of acquiring knowledge about social situations and being able to use this knowledge in different contexts. It establishes the individual as an agent rather than a passive victim. A key factor in this therapy is that the patient begins to experience that another person has the patient's "mind in mind."

Jeffrey E. Young and colleagues have devised *schema therapy* in recent years. An outgrowth of cognitive behavioral therapy, this treatment is an integrative approach for treating patients with personality disorders. It combines behavioral techniques with experiential, interpersonal, and psychoanalytic object relations approaches. The theory is based on a belief that patients have developed maladaptive cognitive perceptions (i.e., maladaptive schema) of themselves, others, and the environment as a result of their core emotional needs not being met during their formative years. The technique is similar to dynamic therapy in that childhood experiences are seen as the root of a patient's difficulties, and there is a focus on the therapy relationship itself but the therapist is much more directive (rather than neutral), and may make more use of self-disclosure. The therapist will try a limited "re-parenting" of patients by providing a supportive environment in which patients are nurtured, as for example, by being praised and given extra time. Patients are seen once or twice a week for 1-to-4 years.

ETHICAL ISSUES

Value and the Value-Free Fallacy

A neutral and accepting stance is often believed to be the backbone of the psychoanalytic endeavor, whereby the clinician suspends judgment by attentively listening and being objective toward the patient's problems. Theoretically at least, this means exercising professional restraint and not imposing personal views or values onto a patient. Freud was very clear in recommending that an analyst refrain from trying to change a patient into his or her image. He felt that an analyst should not "turn a patient into private property," decide a patient's fate, or force his or her ideas onto a patient, and he also felt that an analyst could treat people with whom he or she had nothing in common. In reality, however, the notion of the analyst

as value-free is a fallacy, and the current understanding is that all therapists embody particular social standards, value orientations, or personal preferences—unconsciously, if not consciously. Even diagnosis can be a form of value judgment (e.g., diagnosing personality disorders such as borderline or narcissism, or considering homosexuality a psychiatric disorder as it once was).

Other social, cultural, and religious values are expressed, directly or indirectly, every time the clinician may encourage (or even question) a married man to cease having affairs, an unmarried woman to have an abortion, a troubled couple to divorce (or to stay together), a failing student to remain in school, a traumatized child to leave home, or a sexually abused daughter to take legal action against her father. Long-term psychoanalytic psychotherapy also holds the ethical danger of keeping patients beyond necessity for the benefit of the therapist rather than the patient. Yet some ardent analysts may deceive themselves, and their patients, by failing to permit timely termination or by not mentioning alternative approaches (e.g., drug treatment or brief therapies) that might be more therapeutically efficacious.

Boundary Violations and Sexual Transgressions

As noted earlier, boundary violations, including overt sexual transgressions, were seen among the early proponents of psychoanalysis. Blurred or broken professional boundaries remain significant problems now over 100 years later in both psychoanalysis and psychoanalytic psychotherapy, and according to a survey of psychoanalysts, most therapists have never taken an ethics course. The intense psychoanalytic/psychotherapeutic relationship, which differs from all other professional or fiduciary relationships, is therapy's greatest strength but potentially its greatest liability. The transference implicitly places the therapist in an idealized position of power and even persuasion. It sets the stage for potential boundary violations because of the expected establishment of transference "love" and the reciprocal countertransferential responses that it might arouse in the therapist. The intensity of the bond can easily activate sexual feelings and fantasies while potentially weakening the objectivity needed to retain ethical boundaries. Graphic sexual content, which is traditionally explored in psychoanalysis and dynamic psychotherapy, also reinforces such thoughts and emotions. Even the use of the couch can lead to boundary violations. Andrea Celenza, in her paper "When is a Couch Just a Couch?" notes that the "props" of the psychoanalytic situation, for example, the couch, can be fetishized, objectified, and degraded by either patient or analyst. In this situation, the couch becomes "mere furniture." "In all cases of sexual boundary violations," she says, "it is common to find a vast reservoir of unresolved anger toward authority," and sexual exploitation can reflect "a rebellion

against the authority of the profession and an underlying desire to break the rules."

Boundary violations may range along a continuum from minor ones such as treating a patient in a special way by giving extra time or phone availability, to exchanging gifts, to the cardinal sin of therapy, committing a sexual transgression. Gabbard noted that sexual boundary violations are sometimes more likely in the course of psychoanalytic psychotherapy or even psychoanalysis when a therapist is working with very disturbed, suicidal patients. He emphasizes, however, that regardless of what the pathology is in a particular patient, that *patient is never to blame* for the boundary violations or any transgressions on the part of the analyst. He and Celenza also emphasize that any therapist can be vulnerable, particularly when the therapist is in the midst of his (overwhelmingly, male therapists are involved) life crisis, such as a divorce, a "lifeless" marriage, an illness or death of a family member, a malpractice suit, or even bankruptcy. Though most cases involve a male therapist with a female patient, about 15 to 20 percent of cases involve a female analyst who is often "intensely involved" in heroic efforts to rescue a difficult patient.

Celenza and Gabbard note there is no consistent psychopathology in those who commit the extreme violation of sexual transgression. In fact, in the vast majority of cases studied, the therapist is a one-time transgressor and not a psychopath or sexual predator. Typically, these therapists are often masochistic or even "lovesick." In general, Celenza and Gabbard were optimistic that many of these therapists can be rehabilitated over years, but the therapists must be genuinely remorseful about their behavior (and recognize that it was clearly wrong), agree to be supervised, and be in therapy. Their view is that an evaluation should include a discussion with the therapist to ascertain his or her understanding of why the behavior is unethical. Its purpose, though, "is not to 'find the facts' but rather to assess the psychopathology in the analyst . . . " and to assess his attitude and sense of responsibility toward the transgression as a means of assessing the possibility of rehabilitation. The rehabilitation process can involve limitations on the therapist's practice (e.g., practice limited to seeing only male patients), supervision, or tutorials. The rehabilitation, however, is not in lieu of any legal or ethical sanctions the therapist's behavior might engender. Celenza and Gabbard found that there is a "temptation to master our confusion by simplifying a complex problem," namely by assuming the transgressors are "not only fallible but fatally flawed" and "different from you and me—an 'us vs. them' mentality." Gabbard and Hobday found their experience with over 300 physicians who were referred by licensing boards or ethics committees has convinced them that it is "virtually impossible to predict prospectively which physicians will commit ethics transgressions and which will not . . ." a situation the authors find a "sobering observation" and "highly disconcerting."

There are, of course, some therapists who are not amenable to rehabilitation, particularly when they have serious character pathology and a complete lack of remorse.

Ethical and Legal Options in Sexual Transgressions. Therapists who violate sexual boundaries may face censure, possible dismissal from professional organizations through the action of their ethics committees, the risk of malpractice litigation, and, in some US states, criminal prosecution. The controversy over criminal sanction may be viewed from several standpoints, including philosophical, clinical, legal, and empirical views: (1) philosophically, the definition of consent (vs. coercion) resides at the center of the criminal charge. The issue is what constitutes mutual consent within the therapeutic relationship—how much is genuine, rational, and autonomous as opposed to pathological, irrational, and coercive. In light of the fiduciary relationship and unconscious forces at play, the patient is the nonconsenting party and the therapist is seen as the perpetrator. (2) Some suggest that the therapeutic relationship is *never* over; transference distortions remain long after the treatment has terminated. (3) Legally, the right to privacy interfaces with the nature of consent, given that mutual consent has been used as a defense in these cases; it has even been suggested, though wrongly, that it was the *therapist* who had not consented. (4) From an empirical perspective, there is the question of the effects of criminal actions—are they really a deterrent? And if the therapist is deemed a criminal, should victimized patients be compensated?

Because of the timelessness of the unconscious, what is clear, says Gabbard, is that "once a patient, always a patient," because this idea "represents the most ethical and responsible professional position." This is also the position of the American Psychiatric Association and is in keeping with its *Principles of Medical Ethics,* which explicitly states that, because of the "inherent inequality in the doctor–patient relationship, sexual activity with a current or former patient is unethical." Gabbard notes that patients often terminate with a "sense of unfinished business." In a study in the 1980s, of 71 successfully analyzed patients, two-thirds had contacted their analysts within 3 years of termination. T. Byram Karasu notes that transferences and countertransferences never complete dissolve and "what remains after the therapy has terminated is future potential therapy."

Confidentiality

Confidentiality and Privilege: The Protection of Privacy. Because psychological problems pertain to some of the most secret and troubling thoughts and feelings about a person's private life, their revelations are frequently fraught with fear, shame, and guilt about socially unacceptable events and behaviors. In fact, many countries grant the privacy between

psychiatrist and patient the same absolute protection accorded to husband and wife or lawyer and client. Whereas *confidentiality* is concerned with the protection of patient's disclosures in a therapeutic setting (and has a long ethical history), *privilege* extends that protection to decisions about specific material to be potentially released, usually for judicial or quasi-judicial purposes (and is of relatively recent legal origin).

Because confidentiality requires utmost care and impeccable therapist judgment regarding all formal and informal verbal exchanges, as well as written reports and case records, potential ethical dilemmas for the psychotherapist can arise in relationship to treatment, training, or administrative responsibilities (i.e., in contact with family members, in consultation with other professionals or social agencies, in bureaucratic accessibility to clinical notes and hospital records, in training use of patient presentations and teaching aids). It may also carry over to more personal transgressions, such as the use of case material for the therapist's writings or the more insidious instance of gossip among colleagues and friends. There are two major circumstances under which the patient's right to confidentiality is waived: When the patient chooses to do so (i.e., *patient-litigant exception to testimonial privilege*) and when the therapist must do so (i.e., *protection of public or private peril*).

PATIENT-LITIGANT EXCEPTION TO TESTIMONIAL PRIVILEGE. Despite the patient's general expectation and right to confidentiality, exception is often made in instances of legal action, as when the patient wants information released because he or she is a plaintiff in an emotional damages lawsuit. An ethical dilemma often arises here, not between clinician and court, prosecuted party, or a third party (e.g., worker's compensation board), but between therapist and patient. This is due to the therapist's erroneous belief that he should have the last word on the nature of privilege. Rather, patient consent is always needed (alone or, only if the patient desires, with professional assistance). The therapist has no choice; he or she must release case records or provide professional testimony. Nonetheless, at least three situations constrain the patient's right; there is no privilege if (1) the psychotherapist has determined that the patient is in need of hospitalization; (2) a court orders an examination of the patient's physical, mental, or emotional state; and (3) a communication of the patient's condition is essential to a legal claim or defense.

PROTECTION OF PUBLIC PERIL: THE TARASOFF DECISIONS. This groundbreaking case involved a male graduate student at the University of California Berkeley who was rebuffed by his then girlfriend Tatiana Tarasoff. The student confided to his therapist a fantasy that he wanted to harm this woman and apparently, through a third party, the therapist learned that the student had been thinking of purchasing a gun. The therapist became concerned

about the possibility of violence and contacted the campus police, who went to the student's apartment. Rather than bringing him to the hospital as was planned, the police found him "rational" and just warned him to avoid the woman. The patient never returned for further treatment. Two months later, he stabbed this young woman to death. The woman's parents sued the campus police and the clinicians. The initial Tarasoff Decision (1974) in California found those involved were negligent and had had a "duty to warn" the young woman. The American Psychiatric Association filed a brief supporting the clinicians in which it argued that clinicians are not necessarily able to predict violence. The California Court agreed to rehear the case and issued what has been called "Tarasoff II" (1976) in which the Court found that therapists not only do have a legal duty *to warn* but also noted they have a broader responsibility of a "duty to protect" the endangered person. This may include warning the potential victim or others, notifying the police, "or take whatever steps are reasonably necessary under the circumstances." According to a review by Stuart Anfang and Paul Appelbaum 20 years after the decisions, 20 states had statues that clearly defined the duty to protect.

One prominent example was that the then Press Secretary James Brady and others affected by the attempted assassination on President Ronald Reagan sued John Hinckley's psychiatrist. The case was dismissed because there was no evidence of a prior history of violence in Hinckley. The court ruled "Human behavior is simply too unpredictable and the field of psychotherapy presently too inexact, to so greatly expand the scope of therapists' liability." Over time, courts reportedly have become more reasonable and have accepted that it is an "impossible burden on clinicians" to predict with complete accuracy the possibility of violence in a patient. It is even difficult to assess whether a threat is just a fantasy. A survey of cases published in 2006 found that where courts had found negligence in the time shortly after Tarasoff, they were now more inclined toward "increased judicial sympathy" for clinicians faced with patients who threaten violence.

The Tarasoff decisions involve issues of confidentiality as well. The primary exception to the therapist's protection of patient confidentiality, which occurs irrespective of the patient's wishes, arises in the context of prospective danger to others. The Tarasoff decisions enunciate the maxim that "Protective privilege ends where public peril begins." The case highlights the complex ethical dilemma of weighing the right of confidentiality of individuals versus protection of society from danger. The therapist is in an extremely difficult position—he or she is liable not only for failing to warn the prospective parties but also for invasion of privacy (and even defamation of character) if the potential threat fails to materialize; at the same time, the fidelity of the treatment process is at risk, and the therapist may lose the betrayed patient just at a time when therapy

is most needed, as was the case with the Berkeley student who killed Tarasoff.

Confidentiality and the Future. The protection of confidentiality of the traditional therapist–patient relationship has been under unprecedented pressure in recent years as third-party insurance and peer-review organizations require specific information about treatment. Three forces have been identified in the escalating conflict between the right to protect confidentiality and the right to information: (1) increasing involvement of the government in areas that were previously considered as private, such as business regulation and health care; (2) the technological revolution in data collection, storage, and retrieval (computerized records, videotapes, wiretapping) in response to the greater threat to privacy; and (3) increased awareness of consumer rights and actions. In light of these factors, major ethical questions arise: Whose agent is the psychotherapist—the patient's, the family's, society's, the law's? What are the goals and motives for sharing a confidence—better evaluation and treatment, professional support and validation, countertransferential need (e.g., self-aggrandizement and peer approval)? What are the risks and ramifications of divulging information—jeopardy of patient trust, termination of treatment, or a possible lawsuit?

THE JAFFEE VERSUS REDMOND DECISION. Throughout the years, most states have recognized some form of privilege for psychotherapeutic communications, as contrasted with records for medical treatment. It was not until the Supreme Court 1996 decision of *Jaffee v Redmond*, however, that absolute (i.e., not case by case) privilege was established as "reliable and unequivocal" as possible. The Jaffee case involved a police officer Mary Lu Redmond in Illinois in the early 90s who was accused of using excessive force, while on the job, when she killed a man she perceived to be a public threat. Redmond was sued by the estate of the man she had killed. During discovery, it was found that Redmond had sought counseling from a social worker, and Redmond, claiming the social worker's notes were protected by psychotherapist–patient privilege, opposed the request to supply these records. Ultimately the case went to the Supreme Court. The majority opinion sided with Redmond: "effective psychotherapy depends on an atmosphere of confidence and trust . . . disclosure of information may cause embarrassment or disgrace." The Jaffee decision, though, does not include disclosures once a patient has given his or her consent, including to third-party payers or for applications for disability, or even for security clearance.

Of note is that Supreme Court Justice Antonin Scalia, who wrote the dissenting opinion, disputed the relative importance of psychotherapy and even the whether its contribution is "so distinctive." Scalia believed there was "a lack of sufficient basis to distinguish" psychotherapists from others

in society in which people place confidences . . . (e.g., bartenders!) and even wondered when psychotherapists "came to play such an indispensible role in the maintenance of the citizenry's mental health."

HEALTH INSURANCE PORTABILITY AND ACCOUNTABILITY ACT. In 1996, the U.S. Department of Health and Human Services, by passing the Health Insurance Portability and Accountability Act (HIPAA), established stronger national confidentiality protections for personal health information (PHI) in medical records in general and for psychotherapy notes, in particular, because of the Jaffee decision. Compliance went into effect in 2003. Its provisions apply to all electronic, written, or oral communications involving a patient. Psychotherapy notes (sometimes referred to as "process notes") are not specifically required but if kept by a therapist, they must be kept separate from the medical record and are not to be entered into the electronic record if therapists want to keep them private. They include contents of conversations during individual, family, joint, or group therapy sessions. Sources do acknowledge that the term "process notes" is imprecise and without any universally accepted meaning. Psychotherapy notes, though, do not include medication prescriptions and monitoring, stop and start times of sessions, modality and frequency of sessions, results of clinical tests, treatment plan, functional status, diagnosis, prognosis, and clinical progress.

HIPAA does allow a patient to release information to third parties. HIPAA also has an "opt-out" system whereby health protected information can be disclosed for treatment, payment, and health care operations once a person enters into medical care. Permission to disclose information is called "consent" or "authorization." In the context of privilege, permission to disclose is called a "waiver of the privilege." HIPAA, though, does not protect psychotherapy notes when a health care professional is defending a malpractice suit that is brought by a patient, though, in general, a patient does not have the right to see these notes. A patient does have the right to see his or her electronic medical record and even the right to correct the information. Sources have raised serious question about the privacy of medical records. One survey found that in large institutions, hundreds of employees have access to patient records during a hospital stay. There have, unfortunately, been several, both inadvertent and purposeful, major security breaches. One instance, reported in 2008 in the *New York Times,* found that 40,000 medical records had been stolen by an employee. A general recommendation is that only information that is the "minimal necessary for treatment" should be included when it will be widely accessible to multiple health care professionals.

WRITING ABOUT PATIENTS. In writing about his failure with his case of Dora, Freud said, "It becomes the physician's duty to publish what he believes he

knows of the causes . . . and it becomes a disgraceful piece of cowardice on his part to neglect doing so, as long as he can avoid causing direct personal injury to the single patient concerned." Significantly, though, over the years, even though Freud used pseudonyms for his patients, the identities of many of these patients have been revealed. One of the major concerns of the Freud Archives, as it releases more of its documents, is to continue to protect the identity of these patients.

There are several ways of writing about patients, including obtaining informed consent from a patient, disguising the patient's identity sufficiently without sacrificing the authenticity of the report, and creating case composites. Sources question whether "informed consent" is really possible from a patient: the patient may feel compelled to agree. Furthermore, the request will "indisputably" affect the therapeutic process. For example, a therapist may feel indebted to a patient and be more "lenient" in terms of fee or schedule if the patient agrees whereas the therapist may feel angry and resentful, either consciously or unconsciously, if the patient refuses. There is no good time to ask for consent, whether before, during, or even after therapy has ended, and therapists should consider how likely the patient will come across the document once it is published. In his study of his 42 lives in treatment, Wallerstein noted that it was quite unlikely, for example, that his patients would ever read their case histories.

TELEPHONE AND EMAIL: PSYCHOTHERAPY IN THE 21ST CENTURY (TELETHERAPY)

It is a rare psychotherapist in today's world who has not been confronted by a patient who requests communication (and even entire sessions) via telephone or correspondence by email or text messaging. Men and women today lead vastly different lives from the days of Freud in old Vienna: More and more women are in the workplace, and both men and women are more likely to travel as part of their customary work responsibilities. Furthermore, as psychoanalysis continues to spread worldwide, telecommunications, including providing both treatment and supervision to colleagues, have been used in places such as China where there are few analysts. Teletherapy is the term for therapy conducted over the telephone, over the internet, and by video-conferencing. Patients should be alerted that there is no guarantee of privacy when using these devices, at least at this time. It can also be quite jarring when patients are in vastly different time zones and a "good morning" is really a "good night" to the other.

Some (usually older, more traditional) therapists reject these requests under any circumstances. Many others will find themselves acquiescing, (lest they be considered inflexible and old-fashioned), at least to some of these requests, with feelings ranging from enthusiasm to resentment. What is interpreted to the patient depends on the type of therapy.

Psychoanalysis, exploratory psychoanalytic psychotherapy, and supportive psychotherapy all can be conducted on the telephone or even involve the use of email. These devices are most effective, though, when they provide a temporary bridge to maintain a connection between therapist and patient when the patient is unable to attend a session (e.g., due to illness, weather, and business out of town).

Telephone sessions, at least for an analytic patient, may seem not unlike an analytic session. The analyst is always out of view anyway. But a telephone session is always second best to the subtleties and nuances a therapist may pick up in all the nonverbal communication the patient reveals. Even the way a patient comes dressed to a session may have significance that is lost by the telephone. And silences may be more difficult to manage on the telephone. MacKinnon, Michels, and Buckley believe a therapist "is always at a disadvantage" when treating a patient by telephone. Some patients, obviously, are better suited to use telephone for sessions than others. Patients in supportive psychotherapy, for example, who are particularly fragile, may find comfort in just hearing the therapist's voice, whereas others may need the "face-to-face" contact the office setting provides. What works best is that a patient (and therapist) should create his or her own private and quiet place for the entirety of the session. Just as with sessions in person, the more consistency, the better. Patients sometimes fail (or are unable to do so), such as when a patient chose to call from a hotel lobby or from a busy street (both of which were actual occurrences for one therapist). And patients can abuse the privilege, such as when they oversleep and request a telephone session when the therapist's office is just blocks away from where the patient lives. Other examples of resistances can include forgetting to make the phone call, calling late, and speaking too softly. Further, a request for a telephone session may have many meanings, both adaptive and maladaptive, conscious, and unconscious, that should be analyzed.

A more recent innovation is the use of Skype in which patient and therapist are connected and are able to see each other on their computers or cell phones via an Internet exchange. Jill Savege Scharff, in her edited 2013 book, *Psychoanalysis Online,* details how therapists have described their own, as well as their patients' experiences of helplessness, powerlessness, and vulnerability when the connection fails as it invariably does (at least with the current technology available) during a session. There can be extraneous noises on the line, fuzzy images, and a weak, pixilating or frozen image. It can also lead to a patient's disillusionment with the therapist with whom he or she has endowed omnipotence when the Skype connection is disrupted or malfunctions. Furthermore, when there is silence, either cannot necessarily tell whether the silence is due to a technical problem or is deliberate by either patient or therapist. Skype is considered "a free social platform" and not secure, and as such, "does not

claim to be HIPAA compliant."

Furthermore, both patients and therapists should refrain from multitasking simultaneously—that is, checking their email or doing other computer-related activities, during the telephone call. One patient, for example, was unable to come for regular sessions as she had two young children and had moved away from the therapist but wanted to continue her therapy. She admitted, somewhat sheepishly, to her therapist that she found herself picking up her children's toys, making them lunch, etc., while having her telephone session on a mobile phone. Though she was able to maintain the contact she so desperately needed, both therapist and patient realized these activities "watered-down" the effectiveness of her psychoanalytic psychotherapy and, as the expression goes, "was literally phoning it in."

Email contact most commonly occurs for administrative or logistical reasons, such as for scheduling changes, but any dynamically oriented therapist appreciates even this use should be analyzed, that nothing is ever "just administrative." In other words, there lies an unconscious meaning behind any behavior, including whether a therapist chooses to respond to a patient's email.

Sources describe how email communications may interfere with time and place boundaries for the therapist and may lead to countertransference reactions when the therapist feels unable to "maintain a mental separateness" between his personal life and his professional responsibilities. Email may also facilitate transference fantasies in a patient regarding what a therapist does outside of the office, and it may seem like a gift or even as a transitional object to a patient. Sources also distinguish a telephone call, which usually occurs in a "shared 'real' time" from an email, which has no time boundaries (though instant messaging can have them).

Further, email communication is not always confidential. A therapist must be sensitive to conducting any kind of dialogue with a patient when wireless networks provide less security and patient protection.

Email can, though, lead to its own complications. Gabbard describes how cyber communication by email was disinhibiting to one patient and led the patient to reveal highly erotic, passionate fantasies about him that she had been reluctant to reveal in session. With the use of passwords and an online identity as well, an email communication can take on an "illusion of a secret forbidden relationship." Gabbard describes poignantly his own discomfort ("part of me felt I needed to close the door while I read her messages") at reading what, at times, seemed like embarrassing "pornographic communications" from this patient.

When using telecommunications, Scharf and her authors recommend that a therapist check to see if he or she is covered by malpractice insurance, as well as consider tax liabilities if a patient is in a different state.

There are some contraindications to telecommunications: patients unable to maintain an alliance under these conditions or are not willing to share responsibility for managing the session; patients who have suicidal tendencies or substance abuse; or those who require medication that cannot be monitored locally.

RESEARCH AND EVALUATION

Complexities in Psychoanalytic and Psychotherapeutic Research

Freud's research lab was his consulting room, and his method was the single case study method of observation, as previously noted. He fully appreciated that psychoanalysis as a therapy had its limitations. In one of his early papers, he said, "I consider it quite justifiable to resort to more convenient methods of treatment as long as there is any prospect of achieving anything by their means." Freud was also firmly convinced that physiology and chemistry would eventually provide a greater understanding of mental processes than he could offer, and he predicted, in his 1914 paper on narcissism, that psychology would "presumably one day be based on an organic substructure."

Throughout the years, research in psychoanalysis and psychoanalytic psychotherapy has proved challenging, particularly because of the anecdotal nature of individual case reports, the difficulty in assessing change, and the complexity of therapist, patient, and relationship variables. Back in the 1980s, T. Byram Karasu, then Chairman of the American Psychiatric Commission on Psychotherapies, and his committee were some of the first psychiatrists to review the methodological and efficacy issues inherent in psychotherapy research. Psychotherapy research can be outcome-oriented (e.g., the effect of therapy once therapy has terminated) or process-oriented (e.g., the study of "transactions" between therapists and patients during the course of therapy). Process-oriented research can be further divided into that conducted in a naturalistic setting of a therapy office setting or else conducted under controlled situations such as in a laboratory setting. There are innumerable methodological problems inherent in this kind of research, including finding a suitable control group, high dropout rates, and the multiplicity of uncontrolled life events that occur outside the therapeutic situation. Even defining what constitutes a "therapist" is an issue as therapists differ in psychological sophistication, training, age, gender, experience, and even their belief in the efficacy of a particular psychotherapeutic treatment. The Commission found that it was often difficult to separate the treatment from the therapist who provides the treatment. There are also ethical issues involved in psychotherapeutic research, such as the ethics of withholding potentially beneficial treatment for a waiting list or control group or even the nature of *informed consent* that implies consent must be voluntary, risks must be specified, and

subjects have the ability to withdraw at any time.

More recently, Scott Lilienfeld and his colleagues, who addressed the multitude of errors that occur when psychotherapy researchers rely on "informal clinical observations" for assessing the effectiveness of treatment, asked the provocative question, "When a patient becomes free of symptoms, was the improvement due to the therapy?" The researchers found four cognitive biases that lead to potentially inaccurate assessment: *naïve realism* (i.e., tendency to believe change following therapy necessarily means change because of therapy); *confirmation bias* (i.e., seeking out evidence that is consistent with an initial hypothesis and discarding evidence to the contrary); *illusory causation* (i.e., tendency to perceive cause when there is none); and the *illusion of control* (i.e., the proclivity to overestimate one's own ability to influence events around them). Lilienfeld and colleagues further delineated 26 "causes of spurious therapeutic effectiveness," including confusing insight with improvement, retrospective rewriting of a patient's pretreatment functioning; not appreciating the tendency of some psychological conditions to spontaneous remit on their own or have a fluctuating course; and failing to understand that improvement can result from a therapist's enthusiasm or the patient's expectation (i.e., the placebo effect). Previous researchers had also found that the researchers' own treatment allegiances could affect therapeutic outcomes. These researchers had become curious to explain why research comparisons of the same treatments could yield different results. Lilienfeld et al. also noted that the longer patients remain in therapy, the more likely so-called "extra-therapeutic" effects such as social support or positive life experiences may contribute to improvement. They quote Karen Horney, ". . . life itself remains a very effective psychotherapist." They call for evidenced-based practice, with more systematic designs and randomized controlled trials.

Early Research: The Psychotherapy Research Project

Back in the 1950s, Robert Wallerstein and his colleagues at the Menninger Foundation in Topeka, Kansas, began a "naturalistic" research project that would eventually span over 30 years with the aim of learning what changes actually take place in psychoanalysis and psychoanalytically-oriented psychotherapy and how those changes come about. The researchers established three treatment groups of patients, all treated by the therapeutic modality that was most clinically indicated—those in psychoanalysis, expressive psychotherapy, and supportive psychotherapy. Altogether, 42 patients were followed; over a third of the patients were followed from 12 to 24 years after their therapies had terminated.

Significantly, such a study could never be conducted today as ethical guidelines have changed considerably over the years. The design of this project raises serious ethical issues: neither patients nor therapists knew

they were part of a research study for the duration of a patient's treatment. In other words, there was no informed consent and both patients and therapists had been unwittingly "drafted into their participating roles" as Wallerstein described in his book *42 Lives in Treatment*. Furthermore, there was no randomization of the patient population and no traditional controls. It was only after the therapy terminated that therapists and patients learned they had been part of a long-term study. The population consisted of an equal number of men and women, ages 17 to 50 years. One of the major findings was that not a single patient of the 22 patients in psychoanalysis had a so-called "classical" analysis. There were many kinds of modifications, including extra contact by telephone, episodic or prolonged sitting up (rather than consistent use of the couch), and even hospitalizations required for some patients. Wallerstein notes, though, the Menninger population outpatient tended to be a sicker population than the customary psychoanalytic outpatient setting. Furthermore, contrary to expectation, the researchers found it was also difficult to distinguish—"a real blurring of boundaries"—supportive therapy from expressive therapy, including psychoanalysis. More supportive therapies, surprisingly, had changes "just as stable and enduring" as structural changes (e.g., changes in defensive pattern, thought processes, anxiety tolerance, and ego strength) seen more typically after a classical analysis. Wallerstein summarized his findings: an analyst/therapist should "be as expressive as you can be and as supportive as you have to be." Rather than widening the scope of psychoanalysis as had been suggested over the years, Wallerstein urged a "repositioning of the pendulum" and more "in the direction of narrowing indications" for psychoanalysis. He urged clinicians to be more respectful of supportive elements in all therapies.

Otto Kernberg, who had been part of the Menninger Project, ultimately reanalyzed Wallerstein's data. He and his associates found that patients with high initial ego strength improved regardless of the treatment modality or even the therapist's skill, but with these patients, psychoanalysis brought the greatest degree of improvement and supportive treatment brought the least. In other words, psychoanalysis was found to be the ideal treatment for those who needed it the least. These researchers found those with lower levels of ego strength required a consistent focus on the transference in the "here-and-now." From his review of the data, Kernberg found it was an ominous prognostic sign for all types of treatment when patients initially had a poor quality of interpersonal relationships, and a purely supportive approach without a focus on the transference was contraindicated.

Later Research Studies

Investigations at major psychoanalytic institutions, including the Menninger Foundation, the New York Psychoanalytic Institute, with

studies by Joan Erle and Daniel Goldberg, Columbia University, with studies by Henry Bachrach, and the Boston Psychoanalytic Institute, where Judith Kantrowitz conducted her studies, have all demonstrated that many patients do derive substantial benefit from psychoanalysis or psychoanalytic psychotherapy. There was, however, no uniformity of design among the studies, and researchers emphasized different measures. Kantrowitz, for example, highlighted the importance of the patient–therapist match, "an interactional concept . . . a spectrum of compatibility." Others explored such issues as whether tape recording of sessions had an effect on the therapy.

Despite specific claims and counterclaims on behalf of one or another modality, overall research suggests that all psychotherapeutic approaches may be efficacious. This may be largely due to nonspecific factors, or common elements, that all therapeutic methods are presumed to share: a healing setting, a mutual belief system between therapist and patient, an offering of hope for help, an emotionally charged, trusting relationship, and enhancement of a sense of mastery. T. Byram Karasu, for example, delineated three aspects of all therapies, including those that are psychoanalytically-oriented: affective experiencing, cognitive mastery, and behavioral regulation, but noted that contributions of both specific and nonspecific factors to therapeutic efficacy may not be mutually exclusive. In other words, "both ideologically and clinically we could be dealing with a false dichotomy." He defined "technique" as synonymous with specificity and the relationship with the therapist as synonymous with nonspecificity, but the relationship with the therapist is the therapeutic influence that "transcends" specific techniques.

In one study that examined four reputedly different therapeutic approaches (three of them psychoanalytic)—Freudian, Kleinian, Jungian, and Gestalt—to the same patient, descriptive ratings by objective observers failed to differentiate the respective schools of thought. This finding came as a surprise not only to the investigators but also to the therapists. A review of literature based on several research investigations subsequently concluded that effective therapists, irrespective of theoretical orientation, behave similarly—they appear confident, express concern, communicate clearly, and are empathic. Significant differences had more to do with general level of experience than with specific avowed theoretical allegiances.

Other researchers have reviewed the so-called common factors in therapies and concluded that there was not empirical support for a dichotomy between common versus specific factors: both factors are important and not mutually exclusive. The most relevant common elements were found to be patient engagement, affective experiencing, and therapeutic alliance.

Diagnosis has always been a significant differentiating factor in

therapeutic results with psychodynamic treatment (i.e., favoring less severe pathology). Leichsenring and colleagues found 47 randomized controlled studies (from 1970 through 2013) of psychodynamic therapy that supported that this therapy is "efficacious" for a wide range of disorders, including anxiety, depression, personality disorders, pathological grief reactions, post-traumatic stress, and even substance abuse. They suggest that future studies should clarify whether psychodynamic therapy leads to "specific gains" or what they call "added value" only seen with this therapy, as opposed to other therapies with which it is compared. They also found that there is a dose–response relationship in that longer-term dynamic therapy seemed to be more efficacious than shorter term or less intense, though they acknowledged that there is no "generally accepted 'standard duration.'" One problem, though, acknowledged by the researchers is that treatment adherence is difficult to assess in dynamic therapy because there is no specific treatment manual, unlike in cognitive behavioral therapy.

Studies by John Clarkin on borderline patients compared Otto Kernberg's transference-focused psychotherapy with both dialectical behavior therapy and supportive therapy. All treatment groups showed significant positive change in depression, anxiety, and social adjustment, but transference-focused therapy seemed to produce more positive results with symptoms of irritability and behaviors of verbal and physical assaults. A randomized, controlled clinical trial with panic disorder patients by Barbara Milrod and her colleagues demonstrated the efficacy of psychodynamic psychotherapy with these patients.

The Stockholm Outcome Study of Psychoanalysis and Psychotherapy (STOPP) began in 1992 when Sweden launched a program to subsidize both analysis and long-term therapy with private therapists. From a group of 156 patients who completed therapy (24 in analysis, 129 in dynamic therapy, and 3 in "low-dose" supportive therapy), ultimately two subgroups from analysis and dynamic therapy were randomly selected. The researchers found that those in both groups used self-analysis after termination, but those who received analysis, contrary to expectation, were more likely to use "self-supporting strategies" as well, that included thinking of the analyst's voice or office atmosphere or even recalling a meaningful insight as a source of inner support. The researchers are cautious in their conclusions: they are not certain if the patients chosen were representative of the group as a whole (since those who had more successful treatments were more apt to agree to be interviewed) and cannot be certain of causality, that is, whether these differences noted had existed prior to the patients' therapy experiences.

Per Høglend has reviewed the extensive literature on transference and found there was not even a consensus on a definition of transference or even what a transference interpretation "sounds like" and when this kind of interpretation can be effective. Studies, as well, were "too heterogeneous"

to be evaluated by meta-analysis. He has adopted a broad definition of transference: "the patient's patterns of feelings, thoughts, perceptions, and behavior that emerge within the therapeutic relationship and reflect aspects of the patient's personality functioning . . ." He and his colleagues studied transference interpretations in psychodynamic (i.e., exploratory therapy) in contrast to supportive therapy. In their studies, they used treatment manuals that were "manuals of principles, not step-by-step procedures." They found results contrary to their expectations: patients with more severe and chronic difficulties did better with transference interpretations than those who were higher functioning. The researchers speculated that for those patients whose relationships are more mature a focus on the transference and the relationship to the therapist seems "jarring and strange" and "needlessly focusing the patient's attention on the therapist." They also noted that the timing of a therapist's interventions, the chronicity of a patient's disorder, and how much insight the patient gained may all factor into eventual treatment outcomes, but were not measured in their studies.

Studies, under the aegis of the American Psychoanalytic Association, are currently underway to assess whether psychoanalysis "changes people in profound, meaningful, and enduring ways" that only its in-depth approach can deliver. The researchers note that there is no treatment manual for psychoanalysis and no controlled outcome studies to date for psychoanalysis. They have addressed issues such as the use of rating scales to measure outcome, the question of randomization rather than patient self-selection, the concomitant use of medication, the expertise of the individual analysts, and how to measure adherence to analysis specifically so that the comparison treatments are not just different "in name only."

A wealth of evidence suggests that psychotherapy is indeed efficacious. The jury is still out on whether psychoanalysis and psychodynamic psychotherapy are significantly different in efficacy from each other, as well as from cognitive–behavior therapy, interpersonal therapy, or other relatively standardized modalities.

Future Research Studies

Neuroimaging and the "Biology of Psychotherapy." On a totally different plane has been the recent provocative thesis that psychotherapy actually produces changes in the brain—what Nobel Prize winner Eric Kandel and his colleagues call the "biology of psychotherapy." These researchers are optimistic that neuroimaging, which can simultaneously assess activity in specific brain areas, will eventually be able to predict which treatment (e.g., medication or psychotherapy) is most efficacious, as well as be able to track a patient's progress (through repeated scans over time) and even be able to predict a patient's relapse. Kandel et al. speculate that predicting the efficacy of a treatment may depend more on the "functional characteristics"

of a patient's brain rather than on a specific diagnosis. Neuroimaging, though, in this capacity, is in its infancy: there are many complex variables, including issues of design (e.g., different variables from patient, therapist, and investigator) and translation (e.g., translating different psychological constructs, such as empathy, "from the person to the neuron").

A functional magnetic resonance imaging (fMRI) is a noninvasive technique that measures blood flow in the brain by measuring subtle changes in the brain's use of oxygen when the patient is simultaneously given tasks such as problem solving. In other words, an fMRI measures the "metabolic sequelae" of neuronal activity, but many things can lead to "error variability" or "noise," including the properties of the machine, head motion, the patient's heart and respiration, or even his or her attention. There are many different tasks used and which specific ones are selected has been compared to choosing which psychological tests are to be used for a patient's clinical assessment. Researchers suggest that the fMRI should be part of a full evaluation, including behavioral assessments and patient self-reports.

Mirror Neurons. Mirror neurons, originally discovered in the frontal cortex of macaque monkeys, were found to fire not only when a monkey performed an action but also when that monkey passively watched the action performed by one of the researches. The mirror neuron concept captured the imagination not only of the scientific community but also of the media. These neurons were called in one newspaper report "cells that can read minds." Thought to mediate "action understanding," that is, understanding the intention of another, they were implicated in a vast array of conditions, such as autism ("broken neuron" theory), schizophrenia, empathy ("neurobiological empathy"), and the complex transferential–countertransferential relationship between therapist and patient, even though there was initially only indirect evidence that they existed in humans. Researchers have now debunked much of the hype surrounding the application of mirror neuron theory to human nature. Among other things, humans can understand the motions involved in activities, such as playing a musical instrument, even though they do not know how to play themselves. Future research will determine whether mirror neurons have any relevance to concepts such as empathy or even any specific relevance for psychoanalysis and psychoanalytical psychotherapy.

REFERENCES

*Auchincloss EL, ed. *The Quiet Revolution in American Psychoanalysis: Selected Papers of Arnold M.* Cooper. New York: Brunner-Routledge; 2005.

Auchincloss EL. *The Psychoanalytic Model of the Mind.* Washington, DC: American Psychiatric Publishing, 2015.

*Auchinloss EL, Samberg E, eds. *Psychoanalytic Terms and Concepts*. (American Psychoanalytic Association). New Haven: Yale University Press; 2012.

Cooper AM. American psychoanalysis today: a plurality of orthodoxies. *J Am Acad Psychoanalysis*. 2008;36:235–253.

Dahl HJ, Høglend P, Ulberg R, et al. Does therapists' disengaged feelings influence the effect of transference work? A study on countertransference. *Clin Psychol Psychother*. 2016. doi: 10.1002/cpp.2015.

Etkin A, Phil M, Pittenger C, Polan HJ, Kandel ER. Toward a neurobiology of psychotherapy: basic science and clinical applications. *J Neuropsychiatry Clin Neurosci*. Spring 2005;17(2):145–158.

Fenichel O. *The Psychoanalytic Theory of Neurosis*. New York: W.W. Norton & Company, Inc., 1945.

Fonagy P, Allison E. The role of mentalizing and epistemic trust in the therapeutic relationship. *Psychotherapy*. 2014;51(3):372–380.

*Freud S. *The Standard Edition of the Complete Psychological Works of Sigmund Freud*. London: Hogarth Press; 1953–1964.

*Friedman L. *The Anatomy of Psychotherapy*. Hillsdale, NJ: The Analytic Press; 1988.

Gabbard GO. The early history of boundary violations in psychoanalysis. *J Amer Psychoanalytic Association*. 1995;43:1115–1136.

Gabbard GO, Hobday GS. A psychoanalytic perspective on ethics, self-deception and the corrupt physician. *Brit J. Psychotherapy*. 2012;28:235–248.

Gabbard GO. *Psychodynamic Psychiatry in Clinical Practice*. 5th ed. Washington, DC: American Psychiatric Publishing Inc.; 2014.

Glick RA, Roose SP. Talking about medication. *J Amer Psa*. 2006;54:745–762.

Greenson R. *The Technique and Practice of Psychoanalysis,* Vol. 1. New York: International Universities Press; 1967.

Høglend P. Exploration of the patient-therapist relationship in psychotherapy. *Am J Psychiatry*. 2014;171:1056–1066.

Jacobs TJ. *The Possible Profession: The Analytic Process of Change*. New York: Routledge Taylor and Francis Group; 2013.

Karasu TB, ed. *Psychotherapy Research: Methodological and Efficacy Issues*. Washington, DC: American Psychiatric Association; 1982.

*Karasu TB. *Life Witness: Evolution of the Psychotherapist*. Lanham: Jason Aronson; 2013.

Kernberg OF. New developments in transference focused psychotherapy. *Int J Psychoanalysis*. 2016;97(2):385–407.

Kohut H. The two analyses of Mr. Z. *Int J Psychoanalysis*. 1979;60:3–27.

Leichsenring F, Klein S. Evidence for psychodynamic psychotherapy in specific mental disorders: a systematic review. *Psychoanalytic Psychotherapy*. 2014;28:4–32.

Lilienfeld SO, Ritschel LA, Lynn SJ, Cautin RL, Latzman RD. Why ineffective psychotherapies appear to work: a taxonomy of causes of spurious therapeutic effectiveness. *Perspectives on Psychological Science*. 2014;9(4):355–387.

Lipton SD. Further Observations on the advantages of Freud's technique. *Annual of Psychoanalysis*. 1988;16:19–32.

Lohser B, Newton PM. *Unorthodox Freud: The View from the Couch*. New York: The Guilford Press; 1996.

MacKinnon RA, Michels R, Buckley PJ. *The Psychiatric Interview in Clinical Practice*. 3rd ed. Arlington, Virginia: American Psychiatric Association Publishing; 2016.

Makari G. *Revolution in Mind: The Creation of Psychoanalysis*. New York: Harper Collins Publishers; 2008.

Milrod B, Chambless DL, Gallop R, et al. Psychotherapies for panic disorder: a tale of two sites. *J Clin Psychiatry*. 2016;77(7):927–935.

Scharff JS, ed. *Psychoanalysis Online: Mental Health, Teletherapy, and Training*. London: Karnac Books Ltd.; 2013.

Solms M. Notes on the revised Standard Edition. *Psychoanalytic Review*. 2013;100:201–210.

Spence DP. *Narrative Truth and Historical Truth: Meaning and Interpretation in Psychoanalysis*. New York: W.W. Norton & Company; 1982.

Wallerstein RS. Psychotherapy research and its implications for a theory of therapeutic change: a forty-year overview. *Psychoanalytic Study of the Child*. 1994;49:120–141.

▲ 33.2 Psychoanalytic Treatment of Anxiety Disorders, Obsessive-Compulsive and Trauma-Related Disorders

ERIC M. PLAKUN, M.D.

INTRODUCTION

Since the history of psychoanalysis extends back over 100 years, its conceptualization of psychiatric disorders has evolved over time, extending and correcting but remaining consistent with Sigmund Freud's earliest formulations of the importance of the unconscious, mental representations, transference, countertransference, resistance, and defenses. The reader is referred elsewhere in this volume for a description of the history and theory of psychoanalysis. This clinically focused section endeavors to make accessible a psychoanalytic way of looking at and treating patients with anxiety disorders and several related disorders and to put theory and clinical practice into ordinary language.

Although psychoanalysis is a theory of the mind and of human development, as well as a specific form of treatment developed by Freud for treating "psychoneuroses," the adjective "psychoanalytic" refers to more than these. Today its more usual form in psychiatry is as the word "psychodynamic." This term refers both to a less intensive form of individual therapy based on psychoanalytic theory and to an entire way of looking at things that happen in the minds of (and between) patients and clinicians, but also within (and between) families, groups, and systems, and to a way of understanding, formulating, integrating, and responding to them. Psychodynamic psychiatry is the interface between psychoanalytic thinking and psychiatry. A psychodynamic approach to anxiety and related disorders includes both a specific form of psychotherapy used to treat patients and a way of integrating the overall treatment plan of patients with these disorders. For example, and as illustrated later, this might include a specific way of conducting the psychopharmacologic treatment of anxiety and other disorders that has been called "psychodynamic psychopharmacology."

THEORETICAL ISSUES

Two Types of Anxiety

When Freud initially studied the affective state known as anxiety, he concluded that it was a physiological phenomenon resulting from blocked discharge of libido. Thirty years later, in "Inhibitions, Symptoms and Anxiety," Freud modified this view, as he had others, to differentiate between *primary anxiety* and *signal anxiety*. Primary anxiety is associated

with an overwhelming, inchoate experience of total disintegration and is thought to have origins in the earliest threats to the cohesiveness of an individual at or shortly after birth. It implies no capacity to think about or respond to whatever the overwhelming stimulus may be. Signal anxiety is a warning of potential danger or of a potential threat to equilibrium. It warns that primary anxiety may not be far off and leads an individual toward thought and action. In psychodynamic thinking, signal anxiety is seen as a symptom of a deeper problem rather than necessarily as a specific focus of clinical attention, just as the dashboard "check engine" warning light in a motor vehicle is a warning to deal with some hidden but discoverable danger, not a warning to turn off the light. In both cases, there is an underlying problem that has led to the signal, and the solution lies in the direction of dealing with the underlying problem.

In psychoanalytic structural theory, structures like the ego, id, and superego are used to understand internal mental processes. Viewed through the lens of structural theory, signal anxiety is the ego's response to sexual and aggressive drives originating in the id that are in conflict with the moral standards of the superego. Humans are never far from a conflict between their wish to gratify id impulses immediately and the superego's demands for adherence to standards of behavior that require delay in acting on impulses. Signal anxiety is ubiquitous and part of ordinary human functioning, leading to Freud's observation that anxiety is the price humans pay for civilization. However, some individuals suffer from marked signal anxiety in ways distressing enough to come to clinical attention.

Psychoanalytic theory includes recognition of a predictable progression of developmental issues that emerge in the course of human life. Signal anxiety may occur in response to conflicts that emerge as the developmental sequence unfolds, including but not limited to conflicts about loss of the self, separation from an attachment figure who is loved and needed (separation anxiety), loss of the love of an attachment figure who is loved and needed, fear of massive retaliation by a powerful rival for the love of an attachment figure who is claimed by the powerful rival (castration anxiety within the Oedipus complex), loss of control, dependency, aggression, fears of sexual impulses, and others. Signal (or even primary) anxiety may also emerge from severe early adverse experiences of loss, neglect, deprivation, trauma, or abuse.

Nosological Issues

DSM-5 anxiety disorders are listed according to age of onset, and include separation anxiety disorder, selective mutism, specific phobia, social anxiety disorder, panic disorder, agoraphobia, generalized anxiety disorder (GAD), substance/medication-induced anxiety disorder, anxiety disorder due to another medical condition, other specific anxiety disorder, and unspecified anxiety disorder. This chapter will focus on the psychoanalytic

psychotherapy of three of these: specific phobia, panic disorder, and GAD. In the move from DSM-IV to DSM-5, two anxiety disorders were relocated to other chapters—obsessive-compulsive and related disorders like body dysmorphic disorder, and trauma and stressor related disorders. This chapter on psychoanalytic treatment of DSM-5 anxiety disorders will also address obsessive-compulsive disorder (OCD) and post traumatic stress disorder (PTSD). Although disorders will be discussed as separate entities, there is often comorbidity among anxiety disorders and between anxiety and other disorders. Except for research purposes, psychoanalytic therapists do not generally join what is arguably a false assumption of psychiatry that most or even many patients will present with a single disorder that will respond to a single evidence-based treatment of that disorder. Studies like the STAR*D for depression have shown that the vast majority of patients present with multiple comorbid conditions despite the "one disease-one treatment" models of psychopathology that predominate in the field. Indeed, one of the most robust indications for psychoanalytic treatment may well be the presence of complex comorbid conditions.

While psychiatry groups a series of disorders together as anxiety disorders, anxiety is often part of these disorders in different ways that are meaningful to a psychodynamic therapist. Although some direct experience of a form of anxiety is part of GAD and panic disorder, other anxiety disorders may come to clinical attention because of efforts to avoid anxiety, as in agoraphobia to avoid the risk of exposure to sources of anxiety, other phobias in which anxiety is limited by limiting the risk of exposure to the phobic stimulus to a degree that may compromise functioning, or when obsessions and compulsions in OCD serve to ward off the experience of anxiety, or because of symptoms beyond anxiety (as in avoidance, re-experiencing, negative alterations in cognition, or hyperarousal phenomena in PTSD). These distinctions are meaningful to a psychodynamic therapist. Signal anxiety is an affect, whereas phobic avoidance, obsessions and compulsions, and re-experiencing phenomena are symptoms that may be associated with anxiety (e.g., a flashback in PTSD) or with its relief (e.g., counting rituals that ward off anxiety). Psychodynamic therapists are generally less interested in eliminating a symptom or affect than in helping a patient to discover and then take charge of what lies hidden beneath. What is the conflict hidden beneath the affect of generalized (signal) anxiety? What is the meaning of the symptom of phobic avoidance or of a counting ritual? Psychodynamic treatments view symptoms and affects like signal anxiety as the visible phenomenological tip of an iceberg that really goes much deeper than a surface view suggests.

The Unconscious and Psychotherapeutic Technique

Elsewhere in this volume the evolution of psychoanalysis is addressed, including its move from a topographical model to a structural model, to an

object relations-ego psychological model, and more recently to an intersubjective-relational model. For the current chapter, the model used will be based on the notion that psychoanalytic or psychodynamic psychotherapy of any disorder presupposes that there is a relationship between the disorder and the functioning of the dynamic unconscious. This point of view is based on the notion that much of the behavior of any individual is governed by forces outside conscious awareness. Unconscious processes are active and make significant contributions to psychiatric disorders, including anxiety disorders. Although a patient may be aware of anxiety or of behavioral symptoms that cause distress, the conflicts or other meaning that underlie these may well remain out of awareness, although they are potentially discoverable—and discovery can lead to mastery. These conflicts and the issues associated with them are repressed or kept out of conscious awareness by defense mechanisms.

From this perspective, the task of psychodynamic therapy is to bring into conscious awareness and under conscious control these unconscious conflicts and other meanings that are signaled by anxiety or encoded as nonverbal communications in behavioral symptoms. Making the unconscious conscious is a significant challenge, but psychodynamic therapists understand that there is a range of ways to do this.

Psychodynamic therapists conceptualize three parts of the therapeutic relationship. The *real relationship* always exists, and is based on the reality elements of two individuals involved in the therapy situation, including age, gender, ethnicity, and personality features, among others. The *transference relationship* also always exists, and is based on life experiences from the patient's past that are projected onto the therapist (and the therapist's transference to the patient or countertransference based on the therapist's own life history). The third part of the therapeutic relationship, the *therapeutic alliance,* exists only if it is negotiated between therapist and patient, and this negotiation is an essential early step in treatment. Psychodynamic therapists negotiate an agreement to explore the patient's mental life, encouraging patients to *free associate,* that is, verbalize their thoughts and feelings freely as they arise in sessions, including thoughts and feelings about the therapist. Psychodynamic therapists recognize that a patient's fantasies, dreams, fleeting thoughts, slips of speech, and habitual or unusual behaviors offer opportunities for therapist and patient to discover clues about unconscious motives, fantasies, and conflicts that can illuminate and make sense of the patient's disorder and make it possible to acknowledge, bear, and put into perspective that which is discovered through therapeutic work. From a psychodynamic perspective, the therapeutic alliance includes an agreement for both parties to be committed to the task of therapy, especially in times of regression.

Psychodynamic therapists understand that, in addition to the wish to

change that brings them to treatment, patients, like other people, are fundamentally ambivalent, and thus also fear change and prefer the familiar. This is partly because a patient's symptoms are not only problems but also represent the best solution a patient has been able to find for those problems. Patients may pay a high price for their solutions, but their symptoms are solutions nevertheless. Hence, it is understood that patients will inevitably bring to the treatment a reluctance to change that is known as *resistance*. Given the request that patients free associate, resistance often takes the form of conscious censorship or unconscious repression of thoughts, fantasies, or feelings. It is important to emphasize that not all parts of the resistance are conscious or entirely under the patient's control, yet resistance is motivated and powerful, and can be discovered and engaged, with the patient. In psychodynamic therapy, the therapist and patient work to discover and bring the patient's resistances into consciousness as a necessary step before change can occur.

Psychodynamic therapists work with a focus on transference and countertransference as inevitable therapeutic phenomenon and as powerful opportunities to learn more about what a patient is struggling with, to engage relevant issues in the here and now of the therapy relationship, to deepen the work, and to help a patient to move toward change.

In psychodynamic psychotherapy, the therapist maintains a stance of *technical neutrality* rather than being directive or self-disclosing. Technical neutrality should not be misconstrued as being cold and distant. It is in fact a warm, empathic and nonjudgmental listening stance with evenly hovering attention and responsiveness. Patients who meet with a therapist operating from a stance of technical neutrality will react to the therapist in ways shaped substantially by past relationships that have influenced their experience of others. The psychodynamic therapist strives to accept whatever the patient brings to the sessions, including distorted perceptions of the therapist based on past relationships, that is, the transference, while renouncing ambition and desire about how the patient's life should be or what choices the patient should make. The psychodynamic therapist knows central experiences from the patient's life history will inevitably be recreated in the transference to the therapist. For example, it is likely a patient with a history of early sexual trauma perpetrated by a relative will come to experience the therapist at different times as like the abuser or like another relative who failed to stop the abuse or even as the victim of abuse. Psychodynamic therapists try to accept the transferences that are offered, including distasteful negative transferences, without refusing them or trying to talk the patient out of them, because they recognize that this is the way a patient's problems come to life in the therapeutic relationship.

Because therapists are also human beings, they will have a range of reactions to a patient. Some of these will be the reactions anyone would have to a patient acting in that particular way, whereas others may be more

uniquely related to the therapist's character as shaped by his or her life history. For example, a therapist who was unable to protect his or her own younger sibling from sexual abuse may respond to the aforementioned patient's transference to the therapist as an abusive relative with guilt and a wish to talk the patient out of the transference through reassurance about how professional and competent he is. This kind of situation, in which a patient's transference engages the therapist's character and life history in a way that creates a reciprocal countertransference that the therapist cannot endure, leading to action that pushes the transference away, is called an *enactment*. Enactments are ubiquitous therapeutic phenomena. In fact, the difference between beginning and seasoned therapists is better characterized by the seasoned therapist's ability to detect, analyze, and utilize enactments to deepen the work, rather than entirely eliminate them. To be sure, enactments are a slippery slope, but, as in skiing on a slippery slope, the task is not to avoid sliding down at all, but rather to develop skill in maneuvering and maintaining one's balance and control while sliding down the slippery slope. It is partly because of the inevitability of countertransference enactments that it is so valuable for psychodynamic therapists to undergo psychoanalysis or psychodynamic therapy themselves so that they become aware of their blind spots, character styles, and countertransference vulnerabilities.

SPECIFIC PHOBIA

In phobic disorders, the affect of signal anxiety is evoked in response to a particular phobic stimulus. Avoiding the stimulus often eliminates anxiety, but the price paid to be free of anxiety may be high, seriously compromising social or role functioning. Psychodynamic therapists approach phobias by exploring the unconscious meaning of the phobic stimulus to the individual as revealed by associated feelings, memories, dreams, and fantasies. As these emerge in the transference relationship, there is an opportunity to engage them and to work through the resistances, and the associated issues and conflicts.

D, a woman in her 40s, had a long history of a specific phobia about driving over bridges. Over the course of years, she had been treated for this phobia with benzodiazepines and then, with good effect, exposure and response prevention. After the sudden death of her father, D's phobia returned worse than ever, making it nearly impossible to commute to work. This time behavioral treatment did not seem to help. D was referred to a psychodynamic therapist, who worked with her to explore and elaborate the details of her phobia, which was most anxiety producing when it involved bridges with open sides. Associated with her anxiety was a sense of dread that she would be swept off the side of the bridge and fall to her death. She could sometimes force herself to cross them, but there was a horrible knot in the pit of her stomach the entire way across. Having someone else in the car helped, but only if they agreed to hold her hand while they drove across the bridge.

The therapist explored D's relationship with her father, given that his death seemed to have made the phobia worse. D had been very attached to her father, who seemed more loving and tender toward her than her aloof mother, but, despite his warmth, her father was often absent.

She had not really grieved his sudden death because of how much her widowed mother needed D to be her solid support.

At one point in a session, when the therapist asked if she had ever previously felt the kind of knot in the pit of her stomach that was associated with driving over bridges, D recalled a memory from a family vacation when she was 13 years old. With her father she had climbed a steep and exposed fire tower. Going up had been exciting and the view was exhilarating, but when she tried to climb down D became terrified by the steep pitch and small steps and felt the same horrible knot in the pit of her stomach that was now associated with crossing bridges. D asked her father to hold her hand during the descent, but he declined, saying she had to make it down on her own, then turned away from her and descended. He did not stop when she called him. Alone at the top, D was terrified, but she had no choice but to make her way down alone. She sat on the steps and slid down one at a time, all the while terrified she would slip off a step and fall to the bottom, where her family was gathered waiting and laughing at her timidity. When she got down her father laughingly commented that he knew she could get down on her own one way or another.

In the next phase of her therapy driving over bridges became somewhat easier while therapeutic work was done around the earlier experience at the tower that caused the same knot in her stomach, the same fear of being swept into space and falling, and the same sense of shame and humiliation for being laughed at for not being able to do what others could do easily. These feelings were also associated with a painful sense of abandonment and ridicule from the father she needed and loved. She had never spoken to him again about the incident. A great deal emerged in the therapy about her longing for her father and the way his absence in her life and tendency to ridicule and tease her had hurt her. She felt that there was something about her that made her less interesting to him than her older brother, with whom he spent more time. As these feelings emerged, with them came intense grief over father's death. Now she would never have him with her again, and the hope of understanding why he had acted so uncaringly on the tower, or that he would apologize for it, was gone forever. As she grieved in and outside the sessions, D's bridge phobia diminished further.

After about 6 months of weekly therapy, D felt ready to think about ending therapy, so she and her therapist set a termination date a couple of months in the future. In the termination phase of their work, D began to have more trouble commuting over bridges again. This led to an opportunity to address the ways in which the loss of her therapist was presenting D with another abandonment. Termination was another occasion she would feel all alone and in danger of being swept into the void. During the termination phase, abandonment and dependency issues were explored in the transference. D's symptoms subsided again by the time of termination. At 2-year follow-up she was free of the bridge phobia.

D's phobia about bridges was revealed to be a translation into a symptom of a humiliating memory of an experience of abandonment by her beloved but belittling father, whose death could not be fully grieved until she gained access to issues in their relationship that had been frozen in time in her symptom. In this instance, the "anxiety disorder," that is, D's phobia, and the signal anxiety associated with driving over bridges, concealed issues of unresolved grief and loss.

PANIC DISORDER

Psychodynamic therapists view panic attacks as a form of signal anxiety triggered by unconscious, unresolved past conflicts often involving aggression, anger, separation, and dependency. By following affects, by listening to the patient's wishes dreams and fantasies, and carefully exploring the patient's life history for therapeutic stories that illuminate the meaning of the panic attacks, as well as by attending carefully to the

triggering context of panic attacks, a psychodynamic therapist joins the patient in a search for unconscious past conflicts related to panic. These conflicts are brought into awareness and under the patient's conscious control by engaging these issues as they arise in the transference relationship.

A specific form of psychodynamic therapy has been developed and studied for panic disorder. A randomized, controlled trial compared a manualized, twice-weekly, 24-session therapy called panic-focused psychoanalytic psychotherapy (PFPP) to twice-weekly, manualized, applied relaxation training. PFPP was found to be efficacious in patients with panic disorder.

PFPP is conducted in three phases. Phase 1 focuses on acute panic. The therapist works with the patient to uncover unconscious meanings of panic to achieve relief from panic attacks. PFPP therapists explore the circumstances and feelings surrounding the onset of panic, the personal meaning the patient assigns to panic symptoms, and the feelings and mental content of panic attacks. Conflicts involving separation, autonomy, and suppressed anger are frequently discovered. Within phase 1 some panic relief is typical, together with a reduction in any associated agoraphobia.

Phase 2 of PFPP focuses on lessening vulnerability to panic through the understanding and alteration of core unconscious conflicts. These are addressed as they arise in the transference and are shown to be ubiquitous in the patient's life. Responses expected in phase 2 include improved relationships, less conflicted and anxious experiences of separation, anger, and sexuality, and reduced recurrences of panic.

Phase 3 of PFPP occurs during the last third of the sessions and focuses on termination. Conflicts—for example, those about separation or anger about losing the relationship with the therapist—are re-experienced directly within the transference. Although there may be a temporary recrudescence of symptoms during termination, as if the patient were saying, "If you let the treatment end and abandon me, then all my symptoms will return," this allows consolidation of gains in the termination phase through transference work on the central separation, anger, and autonomy issues associated with panic. Patients generally emerge with a new ability to manage separation and anger and to function autonomously.

GENERALIZED ANXIETY DISORDER

For a psychodynamic therapist even a GAD, in which (signal) anxiety seems to be ubiquitous and without specific content, would be approached with curiosity about and a search for whatever pieces of context may emerge in the course of exploration of a patient's fantasies, fleeting thoughts, and other clues from the unconscious that may emerge as associations to the affect of anxiety are followed and explored. Treatment is oriented toward discovering the unconscious fantasies that illuminate the conflicts

associated with anxiety, some of which may involve fears of separation from attachment figures, fears of loss, fears of castration, or fears of sexual impulses. The therapist approaches the patient with no preconceived notion about what these conflicts are, although with a knowledge of potentially conflictual issues in human development. The following case example illustrates psychodynamic therapy with a patient with a GAD.

B, a 28-year-old man with a history of a GAD, had an alcohol use disorder now in remission and regularly attended Alcoholics Anonymous (AA) meetings. Because of sexual side effects, B was unwilling to take selective serotonin reuptake inhibitor (SSRI) antidepressants, serotonin norepinephrine reuptake inhibitors and buspirone (BuSpar) had been ineffective, and gabapentin (Neurontin) was too sedating. Clonazepam (Klonopin) was effective, but B's peers in AA urged him to give up benzodiazepines. Partly because of these pressures, B sought weekly psychodynamic therapy with a psychiatrist, who suggested that they begin tapering clonazepam in light of B's history of an alcohol use disorder. B balked at this idea, worried that he would become more anxious. The therapist suggested that it might be useful to bring his anxiety to sessions if their task really was going to be to learn more about his anxiety. B agreed to the plan after some reflection.

On a decreasing dose of clonazepam B's anxiety increased. He complained that his male therapist was unempathic, making B suffer with anxiety while the therapist did nothing. As the treatment unfolded, the therapist learned B had been especially close to his mother, who, with B, had been the target of criticism from his often absent, short-tempered, mean-spirited, alcoholic father. B's mother had surgery and chemotherapy for breast cancer when B was 10 years old. It was shortly after this that B's anxiety symptoms began.

When clonazepam was discontinued B experienced an increase in anxiety and had an outburst of anger at the therapist for making B suffer so much. The therapist quietly accepted B's anger at him, empathically noting that he had asked B to endure more anxiety, while leaving him alone and on his own most of the week. When he pointed out that B seemed to be finding in the therapist someone like his absent and sadistic father, B thought this made sense, and began to trust the therapist more. Over time B stated his appreciation that the therapist could endure and understand his anger without needing to retaliate. B also realized that the therapist was sticking to a treatment plan they had agreed to from the outset. As the alliance deepened, B struggled to put words to his experience of anxiety. B spoke more of his attachment to his mother and to the way he would cling to her to support her, pressing himself against her ample bosom, while his father would rage at them both while drunk, sometimes suggesting that B's clinging to her was unnatural and inspired by lust.

B reported a dream in one session in which he watched passively, frozen with fear and guilt and unable to move, as a man murdered and dismembered a naked woman. B's associations to the dream led to painful memories of his mother's disfiguring cancer surgery and to his guilt about not having been able to stop his father from angrily criticizing her before or after the surgery. B then added there was another part of the dream he had left out because of shame; B had been sexually aroused during the dream. B suddenly reported an intrusive thought that upset him—a thought that the breast cancer had come because he had been unable to protect his mother—and because he had been aroused by her breasts. B wept for the first time in the therapy. Over time therapist and patient explored the dream and his intrusive thoughts, learning that B guiltily felt as if he were the cause of his mother's illness and disfiguring surgery. He felt she had become ill because he could not protect her from father's rages, but also as a punishment to him because of his attraction to his mother's breasts. He spoke of the way his father's drunken accusation of lust toward his mother had actually been partly right. B feared, too, that he would be disfigured because of a disease or accident, perhaps by castration, for what he had done to his mother. It was not easy for B to explore these feelings, but as he did his anxiety diminished.

By the end of B's treatment, it was evident that what had appeared initially to be a GAD without context had actually been the surface presentation of unconscious psychological conflicts around attraction to his mother, guilt about her disfiguring illness, and fear of his

father. These conflicts were revealed in the course of B's therapy, with resolution of his symptoms.

POSTTRAUMATIC STRESS DISORDER

In PTSD, a patient has experienced something horrifying that is out of the range of ordinary human experience. The experience of anxiety during the traumatic event may approach that of primary anxiety. The clinical presentation of PTSD includes more than anxiety, with a range of symptoms that carry discoverable meaning.

Freud's early theoretical formulations dealt with the horror of actual but repressed sexual trauma as the cause of neurosis as he came across patient after patient whose unfolding unconscious conflicts led to descriptions of childhood sexual abuse by adults. Freud later moved away from theoretical formulations based on the notion of actual seduction of children by adults to formulations based on the notion that such seductions were wished for and feared fantasies of patients. There is no simple equation explaining such matters. Some patients may have actually experienced early seduction, whereas many others struggle with seduction within the realm of fantasy.

Patients with PTSD have experienced actual sexual or other kinds of trauma. Psychodynamic therapists recognize that patients with PTSD are likely to struggle with difficulty trusting authority figures, given that past abusers may well have been authority figures. Psychodynamic therapists negotiate an alliance within which patients are encouraged to explore the details of their experience of abuse and the associated affects as they are able to bear them. Psychodynamic therapists also anticipate that patients with PTSD will likely mobilize in their transference relationships with therapists aspects of the past abusive relationship. Thus, psychodynamic therapists understand they may be experienced within the transference as the abuser, as the helpless victim of abuse in response to the patient's sadism, or as the passive witness or bystander, like the other parent in an abusive family who seems to the child like someone who ought to know about and put a stop to the abuse—unless he or she is sanctioning it. These are often particularly difficult transferences for clinicians to endure because therapists are motivated primarily by the wish to be healers and helpers. Such transferences may evoke powerful countertransference reactions in the therapist—and enactments in the therapy.

The following case illustrates psychodynamic therapy with a patient with PTSD.

M, a 33-year-old woman, was employed by a domineering male boss who singled her out at work as the target of his anger, publicly humiliating her over the slightest failures. Her efforts to report him to superiors were ignored and his behavior minimized. After a frightening sexual approach by him that she successfully rebuffed, she became symptomatic, with re-experiencing phenomena, negatively altered cognitions, hyperarousal, and intrusive memories of the abuse

by him. She quit her job, felt unable to follow advice from friends to take legal action against her boss, and sought outpatient treatment. M was initially treated by a psychiatrist who prescribed medications for her symptoms and urged her to file a complaint against her boss. Although M felt unable to file a complaint because of her unhelpful past experience with her boss' supervisor, her psychiatrist urged her to do so because it might protect other employees, but she continued to refuse. When she experienced significant side effects, M stopped the SSRI he had prescribed. Her psychiatrist pressured her to try another medication, but this one, too, was associated with intolerable side effects. Unhappy, M ended her treatment, but, when symptoms persisted, began work with a new psychiatrist, although with considerable trepidation that she would have another bad treatment experience.

M began psychodynamic therapy with the new psychiatrist. She quickly felt comfortable with her new psychiatrist when he helped her see that her previous treatment experience had unwittingly recreated the experience of abuse by a domineering authority figure as the well-meaning first psychiatrist had badgered her about making a formal complaint she did not feel ready to file and pressured her to take pills associated with intolerable side effects. The alliance was off to a good start, with a strong positive transference attachment.

Soon, though, M started to miss or come late to sessions. Although still symptomatic, she came to a session with a decision to end the therapy. Taken aback, the therapist pressured her to continue the therapy, and M reacted with intense anxiety and a flashback to the experience of abuse by her boss. The therapist realized he was in an enactment, noticing to himself that the abusive relationship had once again been recreated, this time between them. Now it was the new psychiatrist who was experienced as domineering and pressuring her to do something she did not want to do but had trouble resisting. Aware of the enactment, the therapist noted to the patient that they had unwittingly recreated the abusive scenario. The psychiatrist suggested that they take stock of what was happening between them that would lead to another repetition of a now-familiar pattern. The therapist noted that a choice to end the work was hers to make, but he thought that they were in the midst of important work and thought there was more to understand about her wish to flee, if she were up for the learning that might unfold. M agreed to continue the work. With difficulty M revealed that she had been troubled by somatic sensations she had avoided revealing. For some weeks, about as long as the lateness and missed sessions had been occurring, she had been having strange, confusing and sometimes arousing genital sensations before and during the sessions. She realized with fear and disgust that she was feeling attracted to him. In further work it became apparent that becoming sexual with the therapist was not just something she feared, but also something she guiltily desired. This led to the patient's revelation of an experience of date rape by an 18-year-old man when she was 14 years old. Although she had never really forgotten this incident, it was an "unthought known" in her mind, something she tried not to think about or tried to make unimportant on the rare occasions when she did acknowledge it to herself.

As the experience of date rape was explored, it emerged that it was not her first terrifying experience of being forced to do something that she did not fully want to do by a domineering man. Her father, too, had been frightening when angry, although usually loving and never sexually inappropriate. As the exploration continued, M reported discomfort and a feeling that bringing the memory of the date rape to the therapy felt like being invited to reveal more than she wanted to. When she complained that the therapist was pressing her to reveal things that made her uncomfortable, he realized that this was true and that he was again in an enactment. He apologized for pressuring her, noting he could see he was pressing her to keep going beyond the point that felt safe in a way that replicated her traumatic experiences with men. After the apology M guiltily revealed that the experience of date rape, as terrifying as it genuinely was, was also associated with an intense orgasm. M wondered if she were to blame if that was the case. Over the remainder of the therapy work was done exploring her conflicted and guilty sexuality as her PTSD symptoms subsided. When she felt strong enough, M filed a complaint against her former boss.

M presented with symptoms of PTSD in response to a recent experience at work that had important connections with an earlier instance of sexual abuse and with her relationship with her father. The first abuse experience

was recreated in her work life and then, repeatedly, in enactments in therapeutic relationships, until the underlying fears and wishes were brought to consciousness, explored in the transference, and worked through over the course of several repetitions. Both therapist and patient recognized their own roles in recreating the problem.

OBSESSIVE-COMPULSIVE DISORDER

OCD is a disorder in which signal anxiety may be present (as when intrusive obsessions are associated with anxiety), although many patients with OCD ward off anxiety through compulsive rituals, expending huge amounts of energy through their symptoms to control and contain anxiety about unacceptable impulses.

OCD is thought to have its origins in an individual's earliest struggles to achieve a capacity for autonomous functioning in the period of toilet training. During this period a child is encouraged, lovingly and otherwise, to give up a form of dependence on parents and to take charge of his or her own excretory functions. This sets in motion the potential for a complicated struggle about control in the form of complying with or defying parental expectations. Defiance is a way of refusing to grow up, whereas complying and accepting change, which is a kind of surrender, is ultimately in the service of achieving greater autonomy. A psychodynamic therapist approaches a patient with OCD by searching for the hidden meaning encoded in the obsessions and compulsions, with anxiety often emerging as the meaning of the symptoms becomes apparent.

P, a passive, emotionally vacant, exceedingly polite and quiet man in his 30s, sought psychodynamic psychotherapy for OCD because he was having difficulty functioning at work and in relationships. P had counting rituals and compulsions that took hours. Among his bedtime and cleanliness rituals were a need to keep shoes shined and properly hung on shoetrees or aligned in closets, and behaviors to verify that knives were safely locked up. In sessions, he often droned on endlessly about emotionally empty details of his work life. The therapist became drowsy at one point as he listened to P, who noticed this and, with uncharacteristic affect in his voice, asked, "Doctor, excuse me, but are you listening?" To this the therapist replied, "No, I guess not. Are you?" P seemed taken aback, and then apologized for having been boring.

This incident led to a direct discussion between them about the way P's obsessive, circumstantial, and emotionally empty recounting of details was a form of resistance, and one in which his therapist had joined him by becoming drowsy in the session. Were they going to do the work of therapy together or not?

In subsequent sessions, P made efforts to speak more about the origin of his symptoms, which began with a ritual of kissing his parents each goodnight nine times lest he be unable to sleep. On one occasion, while explaining this, P made a slip of speech about having to kiss his father nine times, instead substituting the word "kick" for "kiss." The therapist asked P if he had noticed the slip. P insisted he could not have made such a mistake, escalated his protests about this for a minute or two as he became sadder, then burst into sobs. While weeping, P accused the therapist of pretending there had been a slip of speech to make him look bad, and then, with an outburst of anger, recalled the hurt and injury he had felt when the therapist had become drowsy in the session.

Surprised at the intensity of his feelings, P recalled when his need to kiss his parents goodnight in a ritualized way had begun. It was after he had gotten and then lost a new puppy.

His controlling and intrusive father had kicked P's beloved new puppy after the latter had a series of toileting accidents in the house. P wept as he recalled his revulsion and rage at his father and the way he had later comforted his dog when alone. P recalled having repeatedly unfolded and refolded the blade of his jack knife to show his dog his weapon and swore to his dog that he would use it on father if the latter ever tried to hurt him again. Soon after the kicking incident P's father decided he had had enough of this messy puppy and sent the dog away while P was at school. P was bereft for a while, but soon settled into an affectless, timid, and passive way of being.

P responded to what he had learned about his compulsions by actively trying to suppress them and became more anxious. As his anxiety was explored, memories emerged of P's earlier struggle with his parents around toileting before the puppy came into his life, when he had received regular enemas from his father to control the frequency of bowel movements. Despite loving his father, P was also enraged at him for taking his dog away without a chance to say goodbye and for the intrusive and terrifying experience of the enemas, while he was also furious at his mother for not stopping father. P also felt humiliated that he had experienced this kind of intrusion into his body and that he had let his dog be given away by his father after swearing that he would protect the puppy from his father.

The meaning of P's specific rituals about knives and shoes became apparent in the course of his therapy as a result of his reaction to his therapist's entry into an enactment in becoming drowsy and as a result of a slip of speech that revealed aggression hidden beneath passivity and compliance. P needed to be sure that no knife blades were exposed because such blades represented the threat of a terrifying assault on his father or an equally terrifying failure to protect the puppy that he loved but lost. Similarly, P's compulsive need to put shoes properly and safely on shoetrees was linked to an effort to put away the memory of and prevent any recurrence of his beloved puppy being kicked and injured by a shod foot. After these were clarified and P tried to control his rituals, signal anxiety emerged, and with it the recovery of memories of an earlier struggle with his parents about intrusive control of his toileting behavior.

PSYCHODYNAMIC PSYCHOPHARMACOLOGY IN THE TREATMENT OF ANXIETY DISORDERS

In addition to psychodynamic therapy as a treatment of anxiety disorders, a psychodynamic perspective informs psychopharmacologic treatment of anxiety as well as other disorders. If anxiety is a signal, then a psychodynamic psychiatrist will wonder what the advantage is of medicating it away rather than bringing the relevant conflicts into consciousness. The psychodynamic psychopharmacologist will hold this point of view in mind while formulating a treatment plan and selecting drugs and dosages.

Psychodynamic psychiatrists attend carefully to the meaning of medications when they treat patients with anxiety and other disorders, recognizing that the way in which medications are prescribed and their intended and unintended effects will have important meaning for patients whether or not there is concurrent therapy and whether or not any such therapy is provided by the psychodynamic psychiatrist or another clinician.

Psychiatrists practicing psychodynamic psychopharmacology pay attention not only to what to prescribe but also to how to prescribe. An example of this was part of the case of B, who presented with a GAD. His therapist's stance about reducing benzodiazepines so that B could bring his anxiety to the sessions led to the emergence of feelings that opened up an understanding of hidden issues in B's life.

The meaning of medications may not be the same to a patient as to a therapist. Psychiatrists who practice psychodynamic psychopharmacology recognize that their wish to help patients with medications may not be perceived simply as help by a patient. The prescription of an SSRI to reduce anxiety may be perceived by a patient who develops sexual side effects as an intentional interference with his or her sexual functioning. Prescription of medications, then, becomes an area in which psychiatrists may step into countertransference enactments. This was illustrated in the case of M, the woman with PTSD, whose first psychiatrist's well-intended efforts to medicate her PTSD symptoms was experienced as a replication of abuse by an intrusive and controlling boss. Her new psychiatrist's eventual recognition of this meaning helped strengthen their alliance and created a precedent for examining and learning from other enactments as her treatment unfolded.

There is evidence of the power of meaning effects of medication. One study suggests that a patient's readiness to change is more predictive of response of panic disorder to a benzodiazepine than whether the patient is receiving active drug or a placebo. Similarly, there is evidence the doctor–patient alliance may be more important for psychopharmacologic treatment outcome than medication. A large, multicenter, National Institute of Mental Health–supported study found that in psychopharmacologic treatment, patients had the greatest reduction in symptoms when a strong alliance was paired with an active drug. Patients receiving placebo who had strong therapeutic alliances with their doctors had greater symptom reduction than patients with poor therapeutic alliances receiving the active drug. These studies support a basic premise of psychodynamic psychopharmacology that symbolic aspects of medications, that is, their meanings, are as potent as the biologically active ingredients and should not be neglected in the care of patients.

References

Beitman BD, Beck NC, Deuser WE, Carter CS, Davidson JR. Patient stage of change predicts outcome in a panic disorder medication trial. *Anxiety*. 1994;1:64.

Busch FN, Milrod BL, Singer MB, Aronson AA. *Manual of panic-focused psychodynamic psychotherapy—Extended range:* New York: Routledge; 2012.

*Freud S. On the grounds for detaching a particular syndrome from neurasthenia under the description "anxiety neurosis." In: *The Standard Edition of the Complete Works of Sigmund Freud*. Vol. 3. London: Hogarth Press; 1996.

*Freud S. Inhibitions, symptoms and anxiety. In: *The Standard Edition of the Complete Works of Sigmund Freud*. Vol. 20. London: Hogarth Press; 1996;3:87.

Kayatekin MS, Plakun EM. A view from Riggs: treatment resistance and patient authority, Paper X: from acting out to enactment in treatment resistant disorders. *J Am Acad Psychoanal Dyn Psychiatry*. 2009;37(2):365–381.

Krupnick JL, Sotsky SM, Simmens S, Moyer J, Elkin I. The role of therapeutic alliance in psychotherapy and pharmacotherapy outcome: findings in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *J Consult Clin Psychol*. 1996;64:532.

Leichsenring F, Rabung S. Effectiveness of long-term psychodynamic psychotherapy: A meta-analysis. *J Am Med Assoc*. 2008;300(13):1551–1565.

Meissner WW. The concept of the therapeutic alliance. *J Am Psychoanal Assoc*. 1992;40:1059–1087.

Meissner WW. Notes on the therapeutic role of the alliance. *Psychoanal Rev*. 1999;86:1–33.

*Milrod B, Leon AC, Busch F, Rudden M, Schwalberg M. A randomized controlled trial of psychoanalytic psychotherapy for panic disorder. *Am J Psychiatry*. 2007;164:265.

*Mintz D, Belnap B. A view from Riggs: treatment resistance and patient authority. III. What is psychodynamic psychopharmacology? An approach to pharmacologic treatment resistance. *J the Am Acad Psychoanal Dynam Psychiatry*. 2006;34:581.

Plakun EM. Making the alliance and taking the transference in work with suicidal patients. *J Psychother Pract Res*. 2001;10:269.

Plakun EM, ed. *Treatment Resistance and Patient Authority: The Austen Riggs Reader*. New York: Norton Professional Books; 2011.

*Plakun EM. Treatment resistance and psychodynamic psychiatry: Concepts psychiatry needs from psychoanalysis. *Psychodynam Psychiatry*. 2012;40(2):183–209.

Wisniewski SR, Rush AJ, Nierenberg AA, et al. Can phase III trial results of antidepressant medications be generalized to clinical practice? A STAR*D report. *Am J Psychiatry*. 2009;166:599–607.

▲ 33.3 Behavior Therapy

DEREK R. HOPKO, PH.D., CAROLINE GRAY CLARK, B.A., M.S.S.W., AND REBECCA SHORTER, M.A.

INTRODUCTION AND DEFINITION

The principles and techniques of behavior therapy are anchored in learning theories that include both operant (instrumental) and classical (respondent) conditioning models. These theories are the basis for a variety of assessment strategies and empirically validated treatments that are frequently used in clinical practice. Examples of well-known behavioral treatments include assertiveness training, parent–child interaction training, exposure and response prevention (RP) (for obsessive compulsive disorder), systematic desensitization (for phobias), stress management, token economies, and problem-solving skills training. More recently, third wave behavior therapies have become increasingly supported and efficacious, including behavioral activation (BA), acceptance and commitment therapy (ACT), functional analytic psychotherapy (FAP), mindfulness-based therapy (MBT), dialectical behavior therapy (DBT), and the cognitive-behavioral analysis system of psychotherapy (CBASP).

A behavioral approach involves thinking about clinical symptoms as learned behaviors and developing treatment programs that help patients learn new ways of behaving (and sometimes thinking) to reduce symptoms and improve quality of life. To develop an appropriate behavioral treatment plan, it is important to use behavioral assessment strategies to formulate a case conceptualization, which then guides the selection of specific treatment techniques. Behavior therapy does not always involve a simple matching of disorders and treatment techniques, although the decision to use a specific intervention often is based on the scientific literature wherein effective treatment approaches have been demonstrated using randomized clinical trials. Another key component of a behavioral therapy involves identifying outcome measures that can be assessed over time to evaluate treatment progress, outcome, and maintenance of gains over time.

HISTORY OF BEHAVIOR THERAPY

Behavioral treatment of psychiatric symptoms has existed since at least the first century, when Pliny the Elder used a type of aversion therapy to modify drinking behavior by putting spiders at the bottom of a drinker's glass. Behavioral procedures also were used by Alexander Maconochie in the early 19th century when he set up a token economy wherein prisoners were awarded points for positive behaviors in a penal colony at Norfolk Island, Australia. However, the formal application of learning principles (e.g., classical conditioning, operant conditioning, and social learning

theories) and the movement toward a behavioral approach to the treatment of abnormal behavior emerged in the 1900s.

The first reference to the concept of behavior modification occurred in 1911, in a classic book written by Edward Thorndike (*Provisional Laws of Acquired Behavior or Learning*). In 1920, the classic case of "Little Albert" was published, demonstrating the application of classical conditioning principles to the acquisition or learning of fear. In a variant of the famous Pavlovian conditioning paradigm, Little Albert was conditioned to fear a white rat when steel bars were banged together over his head at the same time that a white rat was presented to him. After 5 pairings of the loud noise and the white rat, Albert exhibited signs of distress and fear. These principles were similarly applied to the deconditioning or unlearning, of fear in another child, Peter, who was afraid of rats, rabbits, and other furry objects. To decondition the fear, a caged rabbit was gradually moved closer to Peter while he was eating a favorite food. Repeated and gradual pairings of the rabbit and the pleasant food led to a reduction in fear. Fourteen years later, another set of investigators demonstrated the use of a *bell-and-pad* technique for treating enuresis or bed wetting. This treatment also relied on principles of classical conditioning. A bell (buzzer), embedded in a pad on the child's bed, sounded whenever the pad got wet. This pairing of wetting with the sounding of a buzzer conditioned the child to awaken at the time of bladder distension (onset of urination) and contract the sphincter. Ultimately, the child learned to awaken at the sensation of a full bladder without needing the bell/buzzer.

Behavior therapy evolved rapidly in the mid-1950s in three distinct areas of the world: the United States (B. F. Skinner), South Africa (Joseph Wolpe), and the United Kingdom (Hans Eysenck, Stanley Rachman). A major catalyst for the development of behavior therapy was the need for effective short-term therapies for anxiety and depression in World War II veterans returning with significant emotional problems. The characteristic feature of this "first wave" behavior therapy movement was a focus on classical conditioning and operant learning. In addition to forming the basis for anxiety and depression interventions, research on learning principles led to the development of contingency management programs such as token economy systems, used initially in inpatient settings to motivate patients to change behavior in response to differential consequences. In these systems, tokens (e.g., points, stars, and chips) were earned for specific behaviors and then exchanged for other reinforcers such as food or television time. The work of Wolpe, Eysenck, Rachman, and Cyril Franks in the mid-1960s led to the establishment of the first scientific journal focused on behavioral procedures and treatments (*Behaviour Research and Therapy [BRAT]*) and the first national professional association for behavior therapy (*Association for the Advancement of Behavior Therapy*).

The "second wave" of behavior therapy occurred in the 1970s and was characterized by increased focus on information processing and paved the way for a more integrated cognitive behavior therapy (CBT). Much of this work derived from Albert Bandura's *social learning theory,* which posited that people learn not only through the process of classical and operant conditioning but also by observing the actions of others in their environments. The process of *observational learning* emphasized the roles of thoughts, images, and expectations in the development of psychological or psychiatric disorders and set the stage for the development of treatments that target these internal components of behavior more directly. Attending to cognitions in the process of behavior change was a significant departure from "first wave" behaviorist thinking that focused only on behaviors that were directly observable. However, the first two major cognitive behavioral treatments, Aaron Beck's CBT for depression and Albert Ellis' rational emotive therapy (RET), had a major impact on the field of behavior therapy as well as the fields of psychiatry and social work.

Although "second wave" CBT is at present the dominant contemporary system of psychotherapy worldwide, novel behavioral therapies have been developed with steadily increasing support for their efficacy across a range of psychological disorders. This "third wave" behavioral movement largely emerged from a component analysis study directed by the late Neil Jacobson where a BA intervention was as effective as the more comprehensive CBT treatment of individuals with major depression. This finding was further highlighted in a 2006 study where Sona Dimidjian and colleagues demonstrated that BA was as effective as antidepressant medication and more effective than cognitive therapy for individuals with severe depression. This "third wave" of behavior therapies is characterized by themes new to behavior interventions: metacognition, cognitive fusion, emotions, acceptance, mindfulness, dialectics, values, spirituality, and the therapeutic relationship. The techniques used in "third wave" methods are heterogeneous, although they commonly have abandoned the use of content-oriented cognitive interventions and have reverted back to "first wave" principles such as operant conditioning. Today, behavior therapy (including CBT) has become one of the most well-accepted and well-researched approaches for treating many psychiatric symptoms and disorders.

THEORETICAL ISSUES: PRINCIPLES OF BEHAVIOR THERAPY

Behavior therapy consists of many different interventions and a wide array of treatment programs, which combine behavioral (and cognitive) techniques demonstrated as efficacious for a range of psychiatric disorders. Despite the breadth of these techniques and treatments, however, all variations in behavioral treatments share certain basic assumptions or principles.

Maladaptive behaviors are acquired through learning, according to the same principles that govern the learning of adaptive behaviors. Behavior is significantly influenced by the environment through classical conditioning, operant conditioning, and observational learning. Behavior therapists use these learning models to explain how both maladaptive and adaptive behaviors (and thoughts and feelings) develop and to create strategies that help patients change behaviors and improve functioning. Although some behavior patterns clearly have significant genetic or neurobiological mechanisms (e.g., cognitive impairment that results from mental retardation, dementia, and traumatic brain injury), behavior therapists focus on the role of learning history. Contemporary behavior therapists often consider the role of learning in the context of predisposing factors (see Case Conceptualization later in this chapter), some of which may include biologically based influences such as genetics and temperament. Although early behaviorists focused only on the impact of environmental variables, behavior therapists today often look for person–situation interactions to identify how maladaptive patterns of behavior are learned. For example, certain social cues (e.g., speaking aloud in class and giving an oral book report) will be more easily conditioned to elicit anxiety in a child with an introverted temperament.

It is not necessary to identify an underlying cause or motive for maladaptive behaviors. In a behavior therapy model, maladaptive behaviors are viewed as the problem to be treated; they are not considered *signs* of an underlying disorder. Although behavior therapists now consider a range of internal processes (e.g., thoughts and physiological arousal) as part of the targets for assessment and treatment, they do not search for an underlying motivational process or disease. Contemporary behavior therapists view internal processes as behaviors that can be measured (with thought records, physiological assessments), in opposition to more traditional psychoanalytic or psychodynamic approaches that view symptoms as manifestations of a deeper, underlying disease process that is unconscious and cannot be measured. For example, a behavior therapist would view social anxiety disorder as a constellation of symptoms that includes fear/anxiety in social situations, expectations of negative evaluation (e.g., "people will think my talk is terrible"), avoidance of social situations (e.g., not attending social gatherings and not talking out loud in class), and physiological arousal (when thinking about or entering a social situation). A psychodynamic therapist might perceive social anxiety as a sign of a more complicated and unconscious disease process that rests on early relationships with caregivers (e.g., insecure attachment and early rejection experiences). This is not to say that behavior therapists ignore historical or developmental factors. As will be evident below, behavior therapists examine a patient's history for information about the

development of current symptoms, but the focus is on obtaining information about early learning history (e.g., what models were in the person's early environment? what behaviors were reinforced?) rather than toward identifying an underlying motivational or disease process.

Learning principles can be used to modify maladaptive behavior. The same principles that guide the acquisition of maladaptive behaviors can be used to change them. As will become clear when the interventions are discussed, and the associated evidence reviewed, treatment focused on eliminating maladaptive behaviors and acquiring new ones (and new ways of thinking) can improve psychiatric disorders and symptoms. Patients can learn new ways to respond to certain situations, and family members (e.g., parents of young children, spouses, and adult children of aging parents) can learn to shape someone's behavior in more adaptive directions. Of course, just because behaviors can be modified as a result of learning does not mean that those behaviors were originally acquired in the same way or that all psychiatric disorders and symptoms will respond to behavioral approaches. Contemporary behavior therapists certainly consider the role of biological factors in the modification of problematic behavior, recognizing that many psychiatric disorders and symptoms (e.g., schizophrenia, bipolar disorder, and severe depression) warrant biological treatment. However, the behavior therapist then looks to learning mechanisms for ways to improve functioning over and above the response to medication. For example, cognitive behavioral treatments have been used to help people with schizophrenia and bipolar disorder learn to improve medication adherence, decrease the presence of auditory hallucinations, identify and manage early signs of symptom exacerbation and relapse, and enhance interpersonal functioning.

The focus of treatment is on factors that maintain current behavior rather than on historical issues. As noted previously, behavior therapists often obtain information about a patient's past in order to understand how learning history relates to current difficulties (e.g., what early patterns of learning played a role in predisposing the patient to experiencing the current difficulties). However, even more important is an understanding of how environmental (and internal) processes maintain current behavior. Interventions then focus on modifying these environmental factors and teaching the patient new ways to interact with the environment and other people to reduce problematic behaviors, thoughts, and feelings. This emphasis on the present originally was in sharp contrast to more traditional psychoanalytic and psychodynamic treatment where investigation of historical variables and achievement of insight is central. More recent short-term variants of psychodynamic therapy incorporate a greater focus on the present, but historical insight is still seen as curative.

Therapists should be knowledgeable about scientific

literature relevant to the patients they treat. Behavior therapists (and cognitive behavior therapists) place high value on using assessment and treatment techniques that have been demonstrated as efficacious in the scientific literature. Although the use of these techniques may vary depending upon the particular patient, behavior therapists use the empirical literature to guide their selection of treatment tools. Most behavior therapists have been trained within a *scientist–practitioner* or *practitioner–scholar* model, both of which teach the value of integrating attention to clinical care and science. It is the responsibility of professionals trained in these types of models to maintain familiarity with the scientific literature to offer optimal treatment to their patients.

It is important to set specific, measurable treatment goals and measure outcomes. Another important component of an integrated science and practice approach to treatment involves working together with patients to set measurable goals and evaluate the outcomes of treatment. Setting goals helps to shape a treatment plan (i.e., which behaviors need to be targeted in what order) and to define whether or not treatment has been effective (i.e., have the identified behaviors, thoughts, feelings changed in the ways expected?). In this vein, most behavior therapists utilize objective outcome measures as part of their clinical practice, first to assess symptom severity and content at the start of treatment and later to evaluate whether or not changes are occurring in the expected direction (e.g., is depression severity improving? Are problematic behaviors occurring less often?). Without objective measures, it is difficult for both therapist and patient to evaluate whether substantial change has occurred given biases in recall and perception of previous symptoms. As discussed below, a range of assessment tools are available that can be administered at the beginning of treatment and then at various intervals to assess extent of change in identified areas. The patient and therapist together can track symptoms over time to evaluate progress and modify goals/interventions as needed. In this way, the behavior therapist takes an empirical approach not only to awareness of relevant scientific literature but also with regard to working with each patient.

TECHNIQUES OF BEHAVIOR THERAPY

Behavior therapy encompasses a range of assessment and treatment procedures to evaluate problematic behaviors and their relationship to the environment, develop and implement treatment plans, and assess outcomes. Although clearly not an exhaustive list, some of the most commonly used techniques are presented below, and in the context of this case illustration:

Upon entering therapy, Jenna reported she was experiencing depressive and anxious symptoms. Among her depressive symptoms was a depressed mood that had been present for 2 years starting with her breast cancer

diagnosis. She also reported significant loss of energy and impaired concentration, along with moderate feelings of worthlessness, hypersomnia, and decreased appetite. Generalized anxiety symptoms also were reported that included persistent and uncontrollable worry about a number of life areas such as work, family, finances, social issues, and personal health. Psychosomatic symptoms of anxiety included significant difficulty concentrating, becoming easily fatigued, moderate muscle tension, sleep difficulties, and mild irritability. Jenna reported the onset of anxiety symptoms about a year prior to entering therapy, precipitated by knee replacement surgery that affected her ability to work. The patient reported utilizing avoidant behaviors to try and reduce her anxious thoughts (e.g., thought suppression, watching television, and sleeping excessively). These avoidant behaviors became linked to her depressed mood, which further inhibited her from engaging in previously rewarding overt behaviors. Another complication to this patient's situation was two significant Axis III problems (i.e., a breast cancer diagnosis 2 years prior to therapy and knee replacement surgery 1 year prior) that required her to re-evaluate how current behavioral patterns and life goals might be changed as a function of these health issues. These precipitators increased anxiety about her future. Although some previously emitted positive behaviors had been discontinued due to avoidance and decreased motivation associated with depressed affect, decreased engagement in other previously reinforced behaviors (e.g., work activities and health behaviors) were more directly related to physical limitations.

Assessment and Conceptualization

A thorough assessment and case conceptualization are required before a behavioral treatment plan can be developed. The primary goals of an initial behavioral assessment are to identify and define target behaviors and the variables that maintain them and to conduct a *functional assessment*. A functional assessment should lead directly to a treatment plan as it establishes causal (or functional) links between problematic behaviors and the internal and external antecedents and consequences that serve to maintain them. When these links are identified, learning principles and treatment techniques from empirically supported interventions can be applied to modify antecedents, behaviors, and consequences such that problematic symptoms are reduced. Assessment continues as treatment is implemented to ensure that appropriate changes are occurring.

A behavior therapist would first need to understand the nature and severity of Jenna's symptoms, the variables (internal and external) that precipitate her anxious and depressive symptoms, the consequences (internal and external) of these behaviors, and the relationship between anxiety and depression and other possible symptoms (e.g., pain and alcohol use). The assessment would also include historical information about the

onset and course of symptoms and the roles of family members in the maintenance of her difficulties. Some of the techniques that might be used by the therapist in the assessment phase are described below.

Methods of Behavioral Assessment

As with many other theoretical approaches, the behavior therapist often begins the assessment with clinical interviews. A range of additional strategies helps establish links between target behaviors and maintaining factors, evaluate symptom severity against normative data, and set the stage for measuring outcomes over the course of treatment.

Unstructured and Structured Interviews. The structure of clinical interviews has tremendous variability, ranging from a primarily unstructured and completely flexible approach to structured methods that are more restrictive and goal directed. A number of positive correlates are associated with unstructured methods that include increased therapist–patient rapport, ability to assess how patients organize responses, and the potential to explore unique details of a patient's history. Most contemporary behavior therapists allow for some degree of flexibility, although most also make use of some type of intake form or checklist to facilitate the assessment process. The goals of initial interviews involve determining which other assessment strategies might be useful, gathering data on target behaviors and maintaining factors, and educating the patient about the behavioral philosophy. Behavioral approaches typically emphasize individualized assessment given that target behaviors and maintaining factors may vary greatly across patients with common presenting complaints (e.g., two patients with depression may have very different target behaviors and maintaining factors). Nevertheless, behavior therapists typically assign DSM-5–based diagnoses because they help guide selection of appropriate evidence-based techniques. Largely due to concerns about the reliability and validity of unstructured interviews, structured interviews sometimes are used. The most commonly used structured (clinician-rated) interviews are the Structured Clinical Interview for DSM-IV (SCID), for which there are both Axis I and Axis II versions; the Mini-International Neuropsychiatric Interview (MINI), a brief interview that assesses only a subset of DSM-IV categories; the Diagnostic Interview Schedule (DIS), and the Primary Care Evaluation of Mental Disorders (PRIME-MD), which also is brief and was designed for use by primary care clinicians. Other diagnostic interviews such as the Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV) and the Schedule for Affective Disorders and Schizophrenia (SADS) are available for collecting more detailed information on a limited set of diagnostic categories.

Taken together, the main goals of the behavioral assessment interview process are to provide a valid diagnosis as well as identify problematic

target behaviors and gather information about when, where (under what circumstances), and how often they occur. In behavioral assessment, it is critical to *operationalize* and explicitly define what patients mean by various terms. Because feeling anxious or depressed may mean very different things to different patients, it is important to ask patients to explain what it means to experience these feelings (e.g., what specific behaviors are associated with being depressed (how do they and others know when they are depressed), what depressive thoughts are in their minds, what situations tend to make them depressed, and what do other people do or say? Gathering this information in the context of a detailed discussion about specific episodes of the target symptoms also is helpful. For example, in the case of Jenna, a therapist might say:

Think about the last time you became really anxious. Where were you? What were you doing and who was there? What anxious thoughts were in your mind? What did you do or say when you first started to feel anxious? Did you experience physical symptoms of anxiety such as increased heart rate, sweating, shaking, or difficulty breathing? What did the person(s) you were with do or say? How did you then feel and what did you do or say next?

This level of very specific detail provides helpful information about how to define targets for change and which antecedents and consequences need to be considered to establish a behavioral treatment plan.

Lastly, the therapist can use the initial interviews to set the stage for further assessment and to educate the patient about the general nature of behavioral treatment. It is important to explain that working together may involve daily monitoring of symptoms and behavioral observations (see below), as well as home practice assignments when treatment begins. Patients need to understand the approach taken so they are well informed about the roles they will be asked to play.

Jenna's therapist first asked her to describe her symptoms. Dr. Lejuez asked specific questions about how often Jenna felt depressed, what situations she associated with her depression, the severity of depressed mood in various situations, how avoidant and passive she had become, how depression had affected her quality of life, and in what ways her parents had perhaps (even unknowingly) rewarded her depressed behavior by completing her household chores and daily responsibilities. Dr. Lejuez also asked about her thoughts—was she thinking life was hopeless, was she having suicidal thoughts, was she thinking negatively about herself, other people, the world, and her future, was she ruminating frequently and having difficulty concentrating? He learned that rumination was associated with increased sadness, anxiety, reduced motivation, insomnia, self-criticism, and hopelessness. Dr. Lejuez further queried her about the cycle of ruminating and being passive and avoidant, learning that Jenna had elevated beliefs that rumination was useful for solving problems and resolving difficult emotions. This misperception increased avoidance behavior and inhibited Jenna from engaging in activities that potentially would increase environmental reward and decrease depressed mood. Dr. Lejuez also learned that Jenna's family was overly nurturing when she communicated being depressed, and consequently rewarded her passivity by completing her responsibilities and chores. Dr. Lejuez also used questions from the ADIS-IV, determining that Jenna's symptoms met criteria for GAD and major depression. He explained that he would like to work with

Jenna for 8 weeks and begin a problem-solving therapy (PST) and BA intervention. To begin, Dr. Lejuez asked Jena to fill out some questionnaires and to keep a daily diary monitoring log of her behaviors, associated difficulty, and level of experienced pleasure—a strategy to help them prepare a treatment plan. He also mentioned other possible types of treatment that might be useful (i.e., adjunct medications) and the scientific literature that supported them.

Questionnaires and Clinician Rating Tools. Although clinicians from other traditions also use self-report questionnaires and clinician ratings, these measures are commonly used among behavior therapists to assess baseline severity of symptoms. Establishing a baseline allows comparison of symptom severity based on normative data from the literature for people with and without psychiatric diagnoses. Self-report measures are useful as screening instruments, as auxiliaries in the diagnostic process, as tools for monitoring progress across treatment sessions, as outcome measures for assessing the efficacy and effectiveness of various psychosocial and pharmacological interventions, and as a method to assess maintenance of gains at longer-term follow-up assessments. Scales have been designed to assess a tremendous range of content areas across many symptom domains, including affective, verbal-cognitive, somatic, behavioral, and social symptoms. Presented in the annually published *Mental Measurements Yearbook,* the research databases *PsycTESTS* and *Health and Psychosocial Instruments (HaPI),* and the comprehensive bibliography *Tests in Print,* there are a couple thousand measures designed to assess psychological constructs. The majority of these instruments have adequate to excellent psychometric properties and assess a wide range of psychiatric symptoms including anxiety, depression, alcohol use, interpersonal skills, eating behavior, smoking, and relationship discord. In most cases, patients complete these self-report measures on their own, often in the waiting room or at home, making them an efficient way to obtain standardized information about symptom severity. Inspection of responses to individual items can also provide information about patterns of symptoms (e.g., endorsement of suicidal ideation on a general measure of depression such as the Beck Depression Inventory). Responses to critical items can then be discussed during the early assessment-focused clinical interviews.

Clinician rating scales also can be used to obtain standardized information about symptom severity. Many such instruments exist to assess symptoms of anxiety, depression, schizophrenia, and other psychiatric disorders. For example, the clinician-rated Hamilton Rating Scales for Depression (HRSD) and Anxiety are useful toward assessing the severity of depressive and anxiety symptoms and measuring patient progress and treatment outcome. The 24-item *Brief Psychiatric Rating Scale (BPRS)* is one of the oldest and most commonly used clinician rating scale of depression, anxiety, psychotic symptoms, and other unusual behaviors. Important to note, use of such measures requires more

administration time than questionnaires as the clinician queries the patient about a standard set of symptoms and makes severity ratings based on patient responses. Clinical rating scales also require significant clinician training and assessment of *interrater reliability* to ensure ratings are made consistently across clinicians (i.e., two clinicians interviewing the same patient should come up with approximately the same score). If the measures are not used reliably, they are not useful for comparing patient scores against normative data or for evaluating a patient's own scores over time. Nevertheless, standardized clinical ratings can be an important component of a behavioral assessment. Finally, more comprehensive personality tests such as the *Minnesota Multiphasic Personality Inventory II (MMPI) and Personality Assessment inventory (PAI)* may provide useful assessment data on various psychological symptoms such as anxiety, depression, hypochondriasis, psychosis, and symptoms of personality disorders.

At the beginning of the second session, Dr. Lejuez asked Jenna some very specific questions about the frequency and severity of her depressive symptoms. These questions were part of the HRSD, a well-known standardized clinical rating scale for depression. Jenna's score on this measure was a 36, indicating severe symptoms of depression. At the end of this session, Jenna stayed in the waiting room and filled out three questionnaires that asked about symptoms of depression (Beck Depression Inventory [BDI-II]), generalized anxiety (Penn State Worry Questionnaire [PSWQ]), and alcohol use (Alcohol Use Disorders Identification Test [AUDIT]). According to these measures, Jenna had severe depression (BDI-II = 32), moderate worrying (PSWQ = 45), and a strong likelihood of harmful alcohol consumption (AUDIT = 12). Finally, elevations on the MMPI-II depression and psychasthenia scales supported these findings.

Self-Monitoring. Another important tool for behavioral assessment is self-monitoring, a process whereby a person observes, measures, and records behavior as it occurs. This "real-time" type of assessment is in contrast to interviews and questionnaires that are typically retrospective (e.g., *How depressed have you been over the past 2 weeks? How many alcoholic beverages have you had in the last 2 weeks?*). Self-monitoring procedures often produce very different information about symptoms than retrospective assessments that rely on the patient's ability to recall precisely what happened in the past. Memory for past events is often biased or otherwise inaccurate.

Self-monitoring can be used to evaluate the frequency and duration of symptoms as well as to identify sequences of events, behaviors, thoughts, and feelings that may provide information about antecedent and consequent factors that maintain problematic behaviors. Very simple self-monitoring procedures can be used to obtain baseline evaluations of symptom frequency. Patients can be asked, for example, simply to mark on a note card each time they smoke a cigarette, have a panic attack, binge eat, or perform a ritual. Similarly, as presented in Table 33.3–1, daily diary logs can be used to assess the difficulty, enjoyment, and reward value of

activities and behaviors. More complicated monitoring procedures can be used to evaluate behavioral sequences and environmental or internal antecedents and consequences. For example, patients may be asked to record for each episode of anxiety to record the precipitating situations, associated thoughts and feelings, physical symptoms, and consequent behaviors. This type of monitoring provides both frequency and contextual information. Patients also may be asked to rate the severity of anxiety across episodes to evaluate which environmental variables are associated with more or less severe symptoms.

Table 33.3–1.
Daily Diary Activity Log

Time	Activity	Difficulty (0–10)	Enjoyment (0–10)	Importance (0–10)
5–6 AM	Sleeping	3	8	8
6–7 AM	Coffee and breakfast	5	5	5
7–8 AM	Call mother	7	7	8
8–9 AM	Make bed and clean bedroom	4	2	4
9–10 AM	Watching television on the couch	1	7	2
10–11 AM	Watching television on the couch	1	7	2
11–12 AM	Make soup and eat lunch	6	3	5
12–1 PM	Watching television on the couch	2	8	2
1–2 PM	Take dog for walk	8	4	6
2–3 PM	Take a nap	3	7	2
3–4 PM	Take a nap	3	7	2
4–5 PM	Read my Bible	5	5	9
5–6 PM	Made a ham sandwich and ate dinner	7	6	5
6–7 PM	Watched television in my living room	1	7	3
7–8 PM	Spoke to Theresa on phone	8	5	7
8–9 PM	Played game on cellphone	4	7	2
9–10 PM	Sat on chair thinking	2	2	2
10–11 PM	Did laundry	5	2	7
11–12 PM	Read Bible and prayed	6	7	9
12–1 AM	Sleeping	8	4	8
1–2 AM	Sleeping	6	6	8
2–5 AM	Sleeping	3	8	8

Table 33.3–2.
Daily Monitoring of Rumination

Each day, record the amount of time spent ruminating in the morning, afternoon, and evening

	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Monday
Morning	2 hours	1 hour					
Afternoon	45 minutes	1.5 hours					
Evening	2 hours	2 hours					

Once a day, record the following details about an episode of ruminating:

Date	Time	Situation or Trigger	Feelings	Thoughts (Ruminations)	Duration of Rumination	Consequences of Ruminating
Monday	6 PM	Finished work	Afraid Scared Worried	Should not have argued with my boss? What if he thinks I am a problem? What if he fires me?	2 hours	Better for now, maybe I will not get fired.
Sunday	2 PM	Leaving Church	Worried Anxious	I did not really pay attention to the sermon. What if the minister noticed? What if God is angry with me?	1.5 hours	Anxious because I did not get a chance to pray about it later.

In addition to measuring symptom severity before treatment begins, self-monitoring can be used to assess progress during behavioral treatment. In fact, data plots of symptom severity over time can be a highly useful indicator of treatment efficacy for both patient and therapist. Self-monitoring also is often an important component of the treatment program, not just of the assessment process. Recording one's own behavior is a *reactive* process, meaning that the behavior being recorded often changes as a result of monitoring. Simply keeping records of how often one smokes, binge eats, and drinks alcohol can change the frequency of these behaviors as people become more aware of their behavior. Awareness training and self-monitoring are often the first steps in empirically supported treatment programs, although changes that occur only as a result of monitoring often are not enduring.

By the end of Jenna's second session, her therapist was beginning to get a good idea of the nature and severity of her symptoms and some of the maintaining factors. However, in order to plan the treatment in more detail and to get a better idea of how the symptoms occurred during her daily life, Dr. Lejuez asked Jenna to keep daily records over the next week using a form he had prepared for her. In particular, Dr. Lejuez desired more data on the frequency and contextual factors associated with Jenna's tendencies toward engaging in rumination. The form was used to record the amount of time she spent ruminating each morning, afternoon, and evening, as well as more details about at least one episode of rumination each day (e.g., what was happening before and after ruminating; see Table 33.3–2).

Behavioral Observation. Like self-monitoring, behavioral observation involves measuring behavior as it occurs, with attention to frequency, duration, and associated contextual variables. In this approach, however, someone other than the patient (e.g., therapist, parent, and teacher) observes and records the behavior. Before behavioral observation

techniques can be used, it is important to operationalize the target behavior (identify explicitly what will be observed) so that measurements will be reliable (consistent over time and across raters). In the case of Jenna, for example, specific behaviors that represent anxiety would need to be identified. These behaviors might include avoidance of certain situations, reassurance seeking toward others, and the frequency of smoking and alcohol use. Her therapist also might enlist the help of her spouse to assist in obtaining accurate behavioral observations. Observations of this type might involve *event recording,* where each episode of the target behavior is monitored, or *interval recording,* where the frequency of the behavior is monitored as it occurs during a restricted interval of time. In the case of children with behavioral problems, parents and teachers often are included in the behavioral observation process. These behavioral observations may be identified retrospectively through use of measures such as the *Child Behavior Checklist (CBCL),* or more directly within the context of the classroom using the *Behavior Assessment Scale for Children (BASC-II)* that involves direct observations of maladaptive and adaptive behaviors as they occur in the classroom environment using a structured *Student Observation System*.

In some cases, it is possible for the therapist to accompany a patient into the natural environment for observations. For example, a therapist might go to the grocery store with a patient with panic disorder, might accompany a patient with acrophobia to the top floor of a tall office building, or might arrange to take a brief airplane flight with a patient with aviophobia. When natural observations are not possible, *analog* or *simulated* situations can be created. A patient who is depressed and noncommunicative, for example, might be asked to attempt to have a conversation with her therapist. The therapist can then monitor the frequency with which the patient makes appropriate eye contact, smiles, asks questions, and uses "power" and "achievement" words, all of which are scientifically associated with increased depression. Patients in these natural or simulated environments also can be asked to rate the severity of their anxiety or depression using a standardized rating scale.

Role plays are a common method of simulated observation, particularly when the therapist is interested in observing interpersonal skills. In a role play, the therapist and patient take on different roles and enact a scene that provides information about interpersonal or social skills related to depression, anxiety, or other negative moods. To assess assertive behavior, for example, the therapist might pretend to be the patient's spouse while the patient attempts to communicate needs that are being unmet in the relationship. The therapist then can assess the patient's ability to speak clearly, state the problem, and ask her partner to meet needs that are specifically defined. Levels of anxiety at different points during the interaction also can be obtained. Subjective units of distress (SUDS)

measures, which ask patients to rate on a scale of 0 to 10 or 0 to 100 their levels of discomfort, are commonly used in these types of situations. More natural interactions also are used to assess interpersonal skills. A married couple, for example, might be asked during a therapy session to discuss a contentious issue while the therapist observes. The therapist then can monitor communication skills such as eye contact, the couples' ability to listen to each other without interrupting, the on-topic nature of their conversation (e.g., do they veer off to bring up other contentious issues or negative feelings), and any nonverbal behaviors that may escalate negative interactions (e.g., fist clenching and increasing volume of speech). Measures of mood (anger, sadness) can also be obtained during these interactions.

Behavioral avoidance tests (BATs) are another form of observation used commonly to assess severity of fears and phobias. In these procedures, a patient is asked to approach a feared stimulus or situation and the therapist observes behavior, measures how close the patient is able to get to the feared object or situation, and assesses levels of anxiety, usually with SUDS ratings. For example, patients with acrophobia might be asked to move as close to the window in a high office building as they can, reporting levels of anxiety as they go higher. A passive, noncommunicative depressed patient might be asked to talk for as long as possible about a particular topic while reporting confidence and anxiety levels at various time points. During BATs, the therapist also can look for subtle avoidance behaviors (e.g., walking closer to the window, but not looking outside) that may impact the development of a treatment plan.

Designing an appropriate behavioral observation strategy often requires creativity of the therapist. Deciding how best to observe the patient in a situation where the target behaviors occur can be challenging, but these types of observations often produce very useful information about patient skills and sequences of events and behaviors that cannot be derived from interviews or other assessment tools.

Jenna's therapist asked to schedule the next visit at her home. Based on a behavioral observation strategy developed by Peter Lewinsohn, the therapist arranged to observe Jenna and her family during a meal in their dining room. After asking the family to behave as typically as possible and seated in an area largely out of view, Dr. Lejuez coded behavioral interactions in 30-second intervals. All behavioral interactions were seen as having a "source," that is, an individual who emitted an "action" that was often directed toward some other individual, the "object." Actions were followed by "reactions," which were coded as being either positive (e.g., expressions of affection, approval, and interest) or negative (e.g., criticism, disapproval, and avoidance). Using this strategy, Dr. Lejuez was able to identify interpersonal behaviors he perceived as related to Jenna's depression (e.g., lack of social reinforcement). More specifically, Dr. Lejuez observed that a significant proportion of her husband's actions and reactions toward Jenna were critical and negative in quality. He generally treated her politely but coldly, frequently belittling her, not appearing to take what she said very seriously, and rarely praising her or having something pleasant to say to Jenna.

Case Conceptualization

As noted earlier, the primary goal of behavioral assessment is to develop a functional assessment of target behaviors that establishes causal links between behaviors and their antecedents and consequences. Functional assessment of target problems and investigation of the patient's history then allow for developing a *case conceptualization* or working hypothesis that connects information from the assessment and provides a strategy for developing an integrated intervention based on behavioral principles. A case conceptualization allows one to treat multiple disorders concurrently and provides guidelines for understanding and managing patient behaviors that seem to interfere with the process of treatment. Patients with coexistent depressive and anxiety disorders, for example, might be conceptualized according to *core schemas* or maladaptive appraisal systems that are the foundation for both disorders (e.g., a patient with significant depression and worry might have developed the early maladaptive schema of abandonment and rejection, resulting in persistent thoughts and interpretations of situations such as "I am unlovable," and "No one really cares about me"). These patients also might be conceptualized as having core behavioral deficits that might include deficient interpersonal skills (e.g., poor communication and lack of assertiveness skills) as well as behavioral avoidance that limits access to environmental reward and pleasant experiences—resulting in a sad and anxious mood as well as difficulty interacting with others in the environment).

The core difficulties then are examined in terms of predisposing, precipitating, and maintaining factors. *Predisposing factors* reflect both biological components (e.g., genetic influences) and those elements of patients' learning histories that put them at high risk of developing significant symptoms. For example, a patient may be predisposed to develop anxiety or depression due to a strong family history of mental health problems, being raised by authoritarian parents who provided minimal nurturance and were overly controlling and manipulative, and had experiences with parents that were perceived as cold, rejecting, and invalidating. *Precipitating factors* are those events that occur at the start of significant psychiatric difficulties. For instance, this same patient may have recently ended an unpleasant interpersonal relationship in which she frequently tried to express her thoughts and feelings that largely were ignored, minimized, and invalidated by her partner. She also might have simultaneously lost her job, resulting in decreased income and less frequent contact with her social support system. *Maintaining factors* are those antecedents and consequences that reinforce and maintain problem behaviors once they have started. The patient who is experiencing chronic worry may be receiving attention from parents or significant others who

provide constant reassurance that inadvertently reinforces the patient's fears and worry behaviors. Completing the depressed individual's chores and responsibilities may negatively reinforce passive and avoidant behaviors. Careful analysis of core difficulties, predisposing factors, precipitating events, and maintaining factors sets the stage for the development of a behaviorally based treatment that integrates techniques from empirically supported interventions and consideration of each patient's unique learning history.

Jenna's case formulation was based on a behavioral model of depression grounded in functional assessment, based on the notion that depression is associated with an interruption of a repertoire of healthy behavior that previously was positively reinforced in the social environment. In other words, the hypothesis is that Jenna experienced an overall decrease in response-contingent reinforcement for nondepressive (or healthy) behaviors, resulting in depressed mood. A functional analytic view proposes that a combination of reinforcement for depressed behavior and a lack of reinforcement or even punishment of more healthy alternative behaviors leads to depressed and anxious mood and a pattern of depressed, unhealthy behavior. In the case of Jenna, antecedents to depressive affect were her diagnosis and treatment of breast cancer as well as her difficulty regaining mobility after knee replacement surgery. These events had a significant impact on her overt behaviors, access to environmental reinforcement, self-identity, and anticipated life goals. Previously rewarded behaviors, such as working at her job and spending time hiking or exercising had become less enjoyable and sometimes physically impossible. This led to a decrease in positive reinforcement (and increased punishment when pain became debilitating) for activities she had previously found pleasurable, causing her to begin a cycle of avoidance and less healthy behaviors (e.g., passivity and oversleeping). Although Jenna found herself spending less time being active and taking part in more avoidant, passive, and unhealthy behaviors, she recognized that this behavioral routine was worsening her negative mood. At the same time, she was mildly rewarded by these maladaptive behaviors in the context of controlling her symptoms of pain, as well as generalized anxiety, noting that she could effectively avoid anxious thoughts through such behaviors as watching television and oversleeping. Importantly, Jenna had previously been a very physically active person, and the inability to behave in a manner consistent with this value and accomplish the same goals she had previously set for herself further worsened her depression and anxiety.

"First Wave" Behavior Therapy Treatment Strategies

Developing a plan for Jenna or any other patient requires knowledge of treatment strategies that are empirically supported and selected based on the theoretical case conceptualization. A range of behavioral interventions are reviewed below, which may be used alone, or in combination with other techniques depending on the patient's presenting complaints and the therapist's operating case conceptualization. Behavior therapies are presented in the context of the three "waves" (or movements) discussed in the introduction.

Progressive Muscle Relaxation. Relaxation often is used as a component of systematic desensitization therapy (presented below), as a primary intervention for stress and some medical disorders (headache, pain, postsurgical distress), and as part of a comprehensive intervention

strategy for reduction in anxiety. Relaxation therapy increases parasympathetic nervous system activity, thereby decreasing the opposing sympathetic nervous system activity and decreasing arousal. Several different approaches can be used to producing a relaxation response. The most popular is progressive deep muscle relaxation (PDMR), where patients are taught to tense and relax various muscles in their bodies, focusing on the differing physical states of tension and relaxation. PDMR begins by first teaching patients simple diaphragmatic breathing, where they learn to breathe from their diaphragms rather than their chests. Patients are asked to place one hand on the chest and one on the diaphragm to assist in this process and then to inhale slowly and deeply to the count of 4 or 5 and then to exhale to a similar count. Patients are then guided through a sequence of activities that involve tensing and relaxing each of the identified muscles or muscle groups several times. For example, the PDMR script might include tensing and relaxing both hands; both arms; the face (including forehead, cheeks and nose, jaws); neck and shoulders; stomach; both legs; and feet. Other scripts teach the procedures using fewer or even more muscle groups. In some procedures, patients are first taught to tense and relax a large number of muscle groups, and then are asked to combine them over time and with practice ultimately to achieve full body relaxation. The act of tensing before relaxation serves to help patients identify tension and facilitate relaxation. Many people are unaware of tension in their bodies, particularly when they are chronically tense. In PDMR, tensing before relaxing helps people become familiar with the sensations of tension and notice the difference between feelings of tension and relaxation. PDMR sessions typically last 20 to 30 minutes, and patients often are given a tape at the end of the session to continue the practice at home.

A few negative side effects have been associated with relaxation. These include muscle soreness or strain, depersonalization (when practiced continuously for several hours per day), and relaxation-induced panic attacks (occasionally seen among patients with panic disorder who become concerned about the normal physical sensations that accompany relaxation).

Exposure Therapies. Exposure therapies are based on the premise that fears are acquired through associative learning (classical or operative conditioning). Interventions to eliminate fear use the same conditioning principles, and elimination of maladaptive fears requires exposure (contact) with the feared object, event or situation. Exposure is a generic term used to depict a varied set of procedures, all of which are capable of reducing or eliminating fear. Exposure therapy is based on two models: The *counterconditioning model* pairs a feared object with an incompatible response. The *extinction model* is based on habituation of the feared

response. The basic concept of counterconditioning is the substitution of one response for another. The scientific basis for counterconditioning is *contiguity theory, which* states that *unlearning* results from associating a new or incompatible response with the previous stimulus. Although it is now known that the mechanism is not "unlearning" but new learning that is more potent than the prior maladaptive learning, the intervention strategy remains unchanged.

In theory, any positive stimulus that inhibits the expression of a reaction by another stimulus constitutes counterconditioning. The most common learned alternative response for adults is relaxation, and the most prominent method of counterconditioning is the "first wave" therapy known as *systematic desensitization*. The roots of systematic desensitization can be found in Joseph Wolpe's book, *Psychotherapy by Reciprocal Inhibition*. *Reciprocal inhibition* is the permanent weakening of the ability of a stimulus to evoke anxiety, if a response antagonistic to anxiety occurs in the presence of anxiety evoking stimuli so that it is accompanied by a complete or partial suppression of the anxiety response. Reciprocal inhibition is the term used to describe the opposing functions of the two parts of the autonomic nervous system, where the parasympathetic nervous system and sympathetic nervous system work together (actually in opposition) to maintain the body in balance. Wolpe proposed that the concept of reciprocal inhibition could be used to countercondition and thereby diminish fear in humans. With respect to treatment, Wolpe began by using relaxation as an anxiety inhibiting response. His view was that relaxation instigated the parasympathetic nervous system, which in turn, inhibited the sympathetic nervous system (responsible for anxious arousal) through the process of reciprocal inhibition. In short, the autonomic nervous system effects of relaxation are diametrically opposed to anxiety and could be used as a competing response.

Systematic desensitization consists of relaxation therapy (usually PDMR) and the presentation of fear-producing stimuli arranged within a hierarchy. The hierarchy consists of a series of situations (real or imagined) that represent successive approximations to the feared object, situation or event. When constructing a hierarchy, clinicians adhere to three guidelines. First, the items should be very similar to, or when possible, represent real experiences. Second, items should be complete enough, so that patient does not have to fill in the details. Third, the items should sample broadly from the domain of the situation in which the fear might operate.

The specific steps used to construct a hierarchy are: (1) Conduct a thorough assessment to determine all relevant aspects of the fear; (2) Explain the treatment rationale to the patient; (3) Generate a list of potential hierarchical items; and (4) Rate each item using a numerical fear rating scale or SUDS (see Behavioral Observation, earlier in this chapter). Obtaining SUDS ratings for each item on the hierarchy is crucial given that

the order of items may not necessarily follow a rational thought process. For example, a child who has been bitten by a beagle may have greater fear when approaching a beagle than when approaching a much larger dog, such as a German shepherd. Therefore, the therapist should be aware that the items may not necessarily follow conventional logic. Consider the following case and associated hierarchy:

Kevin is 44 years old. When he was in his early 20s, he had an extremely unpleasant conditioning event when traveling by airplane. Specifically, during a lengthy plane flight, the combination of having the flu and violent turbulence resulted in Kevin becoming violently ill. Following this experience, Kevin had not traveled by airplane, and found himself in a difficult position when he obtained a job as a sales representative that would require him to fly frequently. After doing some research on the internet, Kevin learned that behavior therapy was one of the preferred methods for treating aviophobia, or fear of flying. In one of his first few therapy sessions, Kevin and his therapist developed a fear hierarchy that would be the basis for systematic desensitization.

Kevin's Imaginal Hierarchy

Item	SUDS
Purchasing a plane ticket	8
Thinking about flying 2 weeks before trip	10
Thinking about flying 1 week before trip	15
Thinking about the trip the night before	25
Travelling to the airport	30
Checking through security at the airport	40
At the airport in the terminal	45
Boarding the airplane	50
Sitting on the airplane with passengers boarding	56
Airplane departing from gate	60
Airplane taxiing down the runway	70
Airplane taking off	80
Airplane climbing through 10,000 ft	90
Airplane reaching flying altitude	100

In systematic desensitization, PDMR is taught, and the hierarchy is created. A state of relaxation is then paired with the presentation of the feared object, event or situation in a gradual process using imagery. In an initial desensitization session, the patient first achieves a relaxed state. The lowest item on the fear hierarchy is presented first, usually for about 15 seconds (e.g., "You are standing in the airport, waiting for your friend to arrive on the next flight. This time, you are not going to take a trip yourself.). If the patient experiences anxiety, he or she is asked to stop imagining and just relax. After a short period of time, the patient again is asked to imagine the scene. When the patient reports the imagery no longer elicits anxiety, the therapist proceeds to the next item on the hierarchy, repeating the pairing of relaxation with the anxiety-producing imagery.

Typically, no more than three steps of the hierarchy are covered in a single session. Over sessions, however, the cycle is repeated until the patient is able to imagine the items at the top of the hierarchy (previously identified as the most-anxiety provoking aspects of the stimulus) without reporting any anxiety symptoms.

In vivo desensitization is based on the same principles as systematic desensitization and also is a counterconditioning treatment. The basic differences are that PDMR is not typically used in conjunction with the presentation of the stimulus since complete muscle relaxation is not possible when the patient is using muscles in real-life situations, and the stimuli are presented entirely in "real life." Rather than imagining scenes of anxiety-producing stimuli or situations, real life hierarchical steps are developed using the same rationale as for imaginal scene construction in systematic desensitization. Positive competing responses may involve partial relaxation or deep breathing, but in some instances, no competing response is used. The key is that each step represents only a small incremental increase in arousal in comparison with the preceding step. Many of the procedural issues used for systematic desensitization are appropriate for in vivo desensitization, and in many instances, systematic desensitization and in vivo desensitization are combined. The benefits of in vivo desensitization, or graduated exposure, are that neither full muscle relaxation nor imagery is required. Although most patients easily learn relaxation skills, a subset of patients has great difficulty using their imagination. Graduated exposure eliminates the need to use imagery to produce fear. Furthermore, the actual situation, rather than an imaginal representation, is used in the therapy, allowing quicker results and enhancing patient compliance as the relevance of the therapy to the patient's presenting complaint is more clearly seen. Potential limitations of graduated exposure include patients who are too anxious and therefore reluctant to even engage in low levels of actual contact with the feared stimulus and situations where it may not be possible to replicate the exact fear. For example, desensitization to fears that driving in a car will result in an automobile accident and the responsibility for the death of others cannot be replicated solely in real life.

How does a clinician decide whether to use systematic desensitization or graduated in vivo exposure? There is no single reason to choose one over the other. When there is a need to obtain quicker results and when the stimulus is readily available, graduated exposure is preferable. When anxiety and fear are extreme, when there is a complicated fear structure, or when the fear stimuli cannot easily be reproduced, systematic desensitization is preferable.

An exposure procedure that uses the extinction process to reduce anxiety is known as *flooding*. The term *extinction* describes the process of decreasing fear and maladaptive anxiety. When operant behavior that has

been previously reinforced no longer produces reinforcing consequences, the behavior gradually stops occurring. In classical conditioning, when a conditioned stimulus is presented alone, so that it no longer predicts the coming of the unconditioned stimulus, conditioned responding gradually stops. For example, if a socially anxious individual was conditioned to become physically anxious when speaking in large groups, he will eventually stop experiencing anxiety around people if he continues to speak in groups without negative consequences. So over repeated trials, the conditioned stimulus (people) ceases to be a feared stimulus to escape or avoid. In comparison to desensitization, which typically begins with a low level of fear, the objective of flooding is to expose the patient immediately to the fear producing stimuli at its full strength. Furthermore, in comparison to desensitization, where a physiological state incompatible with anxiety is used, in flooding, no such alternative state is used. Finally, there is typically no stepwise progression to build on mastery of lower level fear before confronting higher level fears.

In most cases, flooding (exposure) is accompanied by RP. RP involves discontinuing all avoidance and other anxiety-reducing behaviors. In order for exposure to produce extinction, avoidance behaviors that typically reduce fear need to be prevented (e.g., looking at the floor when talking can reduce social anxiety; during exposure, therefore, to incorporate RP, the patient would be asked to look directly at the people and not at the floor). Over time, while the patient remains in the fear-producing situation, the anxious emotional response will be reduced. *Exposure with response prevention* (ERP) actually is the treatment of choice for obsessive compulsive disorder. In ERP, patients are exposed to the feared stimuli and obsessions while rituals, which typically serve to reduce anxiety, are prevented. ERP also is used for the treatment of eating disorders, although RP is a part of almost all anxiety reduction treatment programs that utilize exposure. Exposure sessions initially last approximately 90 minutes. When a feared stimulus is first presented, anxiety typically rises significantly. Over time, with continued exposure, indicators of anxiety (SUDS levels, physiological indices like heart rate, skin temperature) should decrease. A good indication that the session can be terminated is when the patient is exhibiting significantly reduced anxiety (i.e., close to baseline) while still attending to the exposure stimuli. This reduction in anxiety is often referred to as *within-session extinction*. Across a number of sessions, the key elements for successful treatment are that the peak anxiety level is reduced and the time to return to minimal anxiety is shortened. These responses to treatment are called *between session extinction,* when the fear producing stimulus produces reduced amounts of anxiety and faster within-session extinction over sessions.

Similar to systematic desensitization, exposure can be conducted using imagery or in vivo (in real life). The decision sometimes rests on whether

the stimuli can be reproduced in vivo or must be presented to the patient imaginally. For example, if Linda wanted to overcome her fear of lice quickly, the following imaginal flooding scene might be used:

> Imagine that you are in your home and you have just returned from the movie theater. You notice a slight tingling at the base of your neck, in your hairline. You scratch and for a moment the tingling disappears. You try to continue with some chores and notice the tingling again, but this time your neck feels itchy. You also notice that the itchy feeling has extended into your hair above both ears. You begin to scratch yourself and then you feel something at the base of your neck. You scurry to the bathroom and quickly manage to pick it off. To your shock it is a louse. You look at it in the palm of your hand and for a moment are too horrified to move. You notice its pale beige color and can see its eight legs slowly moving. You quickly throw it in the sink and begin to panic. What if there are more of them? Your heart is racing, and you cannot get your breath. Maybe these disgusting creatures are all over your hair. You scratch some more and notice that several more fall onto your hand, including live white eggs that have not yet hatched. You imagine how many more are roaming through your hair and think of those attached to your scalp, feeding on your blood. What if they are also in your clothes? They are dirty, filthy creatures. You feel light-headed and dizzy as you hurry about the house and prepare to drive to get the lice treatment. What if your children also have them? You scream in horror as you shake another one from your hair. What if the treatment doesn't work? How will you sleep tonight? What if they crawl on your body while you sleep?

Both forms of presentation are effective, although in some cases patients either cannot or may be reluctant to evoke imagery powerful or detailed enough for a sufficient length of time for habituation to be effective. If patients are having difficulty with imaginal flooding, it may be necessary to first provide basic training in imagery, for example, asking the patient to imagine himself/herself in a familiar place such as the home. If patients are successful in providing descriptive details of the situation, they should be able to participate in imaginal flooding. If not, in vivo procedures may need to be used.

As indicated above, graduated procedures (i.e., desensitization) involve the use of a hierarchical approach to exposure whereas flooding involves immediately exposing the patient to the most fearful elements of the feared stimulus. There is some indication that a graduated approach may decrease patient drop-out, noncompliance, and resistance. However, intensive exposure (flooding) achieves faster results, reducing distress in a quicker period of time. The choice often depends upon the particular characteristics of the individual and the disorder. Some factors to be considered include the age of the patient (child vs. adult) and the particular disorder (phobia vs. OCD). For example, because young children, particularly those below the age of 10 years, are still maturing in terms of their basic cognitive skills, they may have difficulty using imagery for extended periods of time. Their ability to concentrate for extended periods also may be limited, particularly when the content is of a distressful nature. Second, although adults easily understand the rationale for flooding (e.g., high intensity means fewer sessions and therefore, short-term stress is worth the long-term gain), this rationale may be difficult for young children to comprehend. In addition, because young children rarely seek treatment

on their own volition, it often is easier to engage their cooperation with a gradual approach. With respect to the issue of the type of fear, Linda's fear of lice was something that could not easily be reproduced in real life. In more extreme situations, when the fear is of a horrific nature, as is the case for many people with OCD (fear of the house burning down), it is not possible to recreate that stimulus in real life and thus an imaginable approach is necessary.

Some considerations regarding the use of flooding include whether the patient can tolerate the emotional distress, whether the patient's medical condition is capable of tolerating temporarily high levels of distress (e.g., does the patient have any cardiac condition?), and whether the therapist is willing to conduct sessions of sufficient length to obtain extinction. Again, if Linda was a 65-year-old woman with a heart condition, flooding may be contraindicated, and a more gradual approach would be appropriate.

Exposure-based treatments have been used effectively to treat a range of anxiety disorders (e.g., phobias, social anxiety disorder, obsessive compulsive disorder, and panic disorder) and eating disorders, where fears of gaining weight are central. Two new strategies may hold promise for further enhancing the effectiveness of exposure therapy. First is the use of virtual reality technology. In *virtual reality exposure,* the patient wears a head-mounted display that produces computer-generated images of feared situations. Sensors on the head and arms allow the patient to "interact" with what feels like a real-life situation. Virtual reality procedures have been used effectively to treat a range of phobias (e.g., flying, heights, spiders, and enclosed places), PTSD, social anxiety disorder, and panic disorder, although more randomized clinical trials are needed. Virtual reality offers the opportunity to carry out exposure-based programs that bypass limitations of both imaginal and in vivo exposure. Second, eHealth is a relatively recent term for health care practice supported by electronic processes and communication. For a variety of mental health disorders including anxiety and depression, technology-based self-help (i.e., computer-based, smartphone applications) shows promise in supplementing more traditional behavior assessment and treatment strategies. For both anxiety and mood disorders, however, therapist-assisted technology treatments appear more effective then implementation without periodic therapist guidance. Although the most efficacious amount of therapist contact varies by disorder, computerized treatments have been shown to be a less-intensive, cost-effective way to deliver empirically validated treatments for a variety of psychological problems.

Operant Strategies for the Acquisition of New Behaviors

Behavior therapy often is directed at reducing or eliminating maladaptive behaviors. In other cases, a lack of appropriate behaviors leads to the need for intervention. For example, children with selective mutism do not speak

in public; an adult with autistic disorder does not make eye contact. Interventions based on operant strategies are powerful methods to acquire new behaviors. In order to understand these interventions, a brief review of operant conditioning is necessary.

Operant is a term used to describe a mechanism whereby a behavior operates on the environment. In this learning model, behavior is not evoked or elicited (as in classical conditioning) but independently occurs and environmental responses influence whether the behavior is continued or stopped. Changes in behavior occur as a result of consequences that reinforce or punish emitted responses. For a consequence to alter a certain behavior, the consequence must be *contingent* on the behavior. *Contingency* refers to the relationship between a behavior and the events that follow it. A consequence is *contingent* when it is delivered only after the target behavior is performed, and it is otherwise not available. *Reinforcement* is a contingent event that strengthens the response that precedes it (e.g., a child is given a candy when he or she smiles; the child then smiles again).

According to operant theory, a behavior may be positively reinforced, negatively reinforced, or punished (see Fig. 33.3–1). *Positive reinforcement* increases the frequency of a particular response by giving something favorable immediately after the response (e.g., stickers and candy). *Negative reinforcement* increases the frequency of a response by removing an aversive event immediately after the response (e.g., nagging stops after a child cleans his/her room). *Punishment* decreases the frequency of a response by either presenting an aversive event following a response (e.g., spanking; *Positive Punishment*) or the removal of a positive event following a response (e.g., allowance is withdrawn; *Negative Punishment*).

	BEHAVIOR	
Action	Increases	Decreases
Given	Positive reinforcement	Positive punishment
Taken away	Negative reinforcement	Negative punishment

Operant theory has several different schedules by which reinforcement/punishment can be applied. *Continuous reinforcement* means that the reinforcement/punishment is applied every time the behavior occurs. Initially, continuous schedules of reinforcement are more useful in establishing a new behavior or increasing a low-level response (e.g., giving a young child a cookie each time he uses the toilet will facilitate toilet-training). *Intermittent reinforcement* occurs when only a subset of target behaviors is reinforced. Intermittent reinforcement can be applied according to a schedule, either a frequency (ratio) or time (interval) schedule. The time interval or behavioral response interval can be *fixed* (every 5 minutes or every 5th time the behavior occurs), or it can be

variable (reinforcement randomly determined). Intermittent schedules tend to produce responses that are more resistant to extinction (e.g., offering a cookie only sometimes when the child uses the toilet helps to maintain appropriate toileting after the behavior has been established, as the child continues to use the toilet waiting for the reinforcer).

Several behavioral interventions exist to assist patients in the acquisition of new behaviors. One strategy, *shaping* is a process of reinforcing (rewarding) successive approximations the come increasingly close to a desired behavioral goal. Behaviors may be reinforced because they resemble the desired behavior or include a component of the desired behavior. For example, there are certain instances where BA protocols (see below) may include elements of shaping. If BA is focused on achieving longer-term goals that require more intermediary behavior, such behavior is of secondary importance and not the unit of analysis, but instead part of successive approximations toward the longer-term (life) goal. For example, the activity hierarchy of a patient who has identified that he has a desire to rock climb might include behavior such as hiking in the woods, reading a book on rock climbing, enrolling in a rock climbing class, climbing a small, fabricated gymnasium apparatus, and then moving toward climbing increasingly steeper rock faces. After beginning with easier tasks that are likely to be completed and reinforced, progressively more difficult tasks are targeted for activation, and preceding approximations of the rock-climbing behavior are extinguished. Of course, effective completion of this activation component necessitates that the patient is experiencing contingencies of reinforcement while rock climbing!

A second behavioral intervention directed at the acquisition of new behavior is *chaining*. Most daily activities require a series of small behaviors or steps that when performed in sequence, form a chain. For example, putting on a pair of pants requires first getting the pants out of the closet, opening them up, putting one leg into the pant leg, putting the second leg into the pant leg, pulling the pants up, and buttoning and/or zippering them. When simple behaviors already exist in the repertoire, the process of *chaining* can be used to form a more complex behavioral sequence. One response produces the conditions for the next response, and so forth, and at the end of the chain, the final behavior is reinforced. Chaining can be backward or forward. *Forward chaining* consists of teaching the sequence of behaviors beginning with the initial step in the chain. An example is counting from 1 to 10. You first reinforce the child to say one, then 1, 2, then 1, 2, 3, and so on. In *backward chaining,* the final response is taught first. For example, a child is taught to pull up his pants (the complex behavior). Training begins with the child being reinforced for having his hands on the waistband. Then he is reinforced for pulling his pants up from just a few inches below his waist, then pulling him pants up from his knees, until he can complete the entire chain from the beginning

(putting on his pants by first putting foot in the leg hole). Shaping and chaining have been used to teach both simple and complex behaviors and have a high rate of success across various groups, disorders, and skill categories.

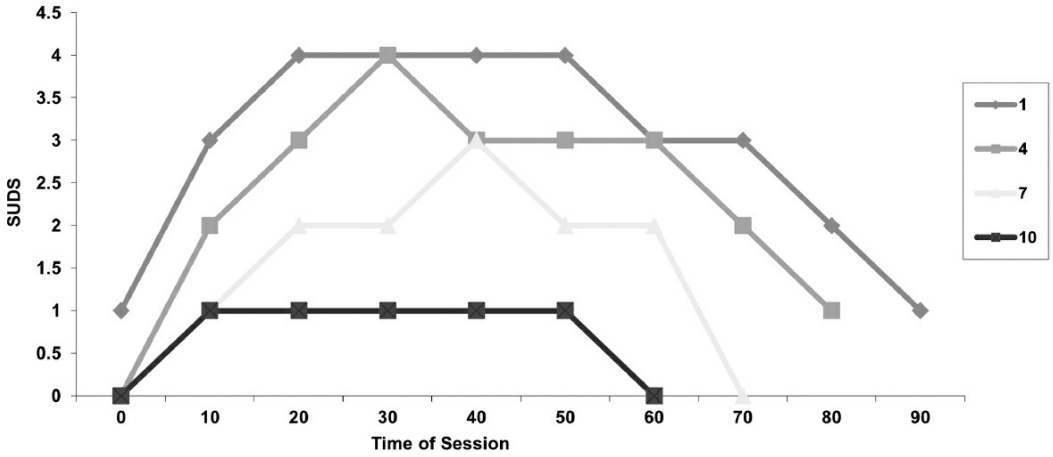

FIGURE 33.3–1. Decreased self-reported distress as a function of continued behavioral exposure.

Contingency Management

Contingency management is not one procedure but is defined as the general application of operant principles (e.g., reinforcement and punishment) in the process of behavior change. Contingency management plans adhere to the same rules of reinforcement and punishment described earlier. Most importantly, failure to apply a contingency consistently leads to the failure of the contingency management program and could inadvertently strengthen an undesired response. When used effectively, contingency management procedures can terminate negative reinforcement cycles, clarify behavioral expectations, and teach people how to bargain and compromise rather than resorting to coercive procedures and how to define and clarify what they want and expect.

Contingency management procedures are implemented using a behavioral contract, an agreement between persons desiring behavior change. Contracts are written agreements that specify the relationship between behaviors and consequences. Several assumptions underlie contingency contracting. First, behavior reinforcement must be earned. Second, good contracts are based on even exchange. Third, the value of the contract is influenced by the reinforcement received. Fourth, a contract increases freedom within the relationship for both parties to earn the reinforcer they desire. In general, effective contracts are usually part of a larger series of procedures used to shape behavior. Below is a sample contract used with a depressed patient.

Sample Contract. I, _____, will attempt to

avoid engaging in the following unhealthy behavior(s): _____; If I do engage in the above behavior(s), then _____ agrees to avoid rewarding this by _____.

Instead, I will try to engage in the following healthy behavior(s):_____.

If I succeed, then _____ agrees to reward this by _____.

Signed:

_____

Your Name

_____

Family Member or Friend

Token economy programs (mentioned earlier in this chapter; History of Behavior Therapy) are a formalized method of contingency management often used in adult and child inpatient settings, day treatment programs, and classrooms. Token economy systems clearly spell out a series of behaviors that are expected (e.g., showering, making the bed, and arriving on time for therapy sessions) and the contingencies for compliance/noncompliance with these behavioral goals. Achievement of behavioral goals leads to the acquisition of tangible rewards (snacks, poker chips that can be redeemed for prizes) or privileges (time off the ward, home visit privileges). Noncompliance does not result in privileges or rewards. Some programs include a provision called r*esponse cost* (a form of punishment), which is the removal or withdrawal of a reward following inappropriate or problem behavior (e.g., fining a child for swearing). Response cost is a very effective strategy for suppressing undesirable behaviors. Fines in a token economy system are one example of response cost, which can be used with a broad variety of problem behaviors and with a broad range of individuals. The combination of token reinforcement and response cost may be more effective than either alone. Token economies also are most effective when administered in a patient's natural environment and when reinforcers eventually are shifted from tangible rewards (e.g., food and points) to social and more natural reinforcers (e.g., praise, feeling good about oneself). Shaping procedures can be used to facilitate this transition.

Maria is hospitalized on the inpatient unit. She has been diagnosed with manic-depressive disorder, and has a history of acting aggressively. The inpatient unit has a token economy program, whereby Maria earns tokens for complying with unit rules (being out of her room during group therapy hours, maintaining her personal hygiene). Because of her history of aggression, Maria's personal program was modified to include a response cost feature—an aggressive act toward another patient or the staff would result in the deduction of all tokens earned for that day.

A number of behavioral procedures incorporate contingency management strategies as part of an overall intervention. For example, *Parent–Child Interaction Therapy* (PCIT) is a structured program that teaches parents two skill sets. In Child-Directed Interaction (CDI), parents learn to use traditional play therapy skills to enhance the parent–child relationship. The skills of CDI include praising, reflecting, imitating, describing, being enthusiastic, and providing contingent attention. In Parent-Directed Interaction (PDI), parents learn skills for improving child compliance and decreasing disruptive behaviors. PDI skills include giving good commands, praising compliance, using time-out in a chair for noncompliance, and establishing structured house rules. PCIT is a short-term intervention that generally requires about twelve 1-hour weekly sessions. There are many approaches similar to PCIT, with most including at least some of the following training modules: (a) positive reinforcement training (how to use reinforcement to change a child's behavior), (b) attending to desirable behavior to increase its frequency and ignoring undesirable behaviors to decrease its frequency (e.g., praising a child when he/she is playing quietly and ignoring the child when he/she interrupts), (c) giving commands, (d) time-out, and (e) response cost. PCIT has been used successfully with a range of behavioral problems, including oppositional behavior, aggression, eating behaviors, and homework completion.

Time-out is punishment by removal of positive reinforcement. Time-out is a complex strategy, and several variables need to be considered for it to be used effectively. When applying time-out, the therapist must consider the duration of the time-out interval, the location where time-out occurs, and what behavior the child is required to exhibit before time-out can be concluded.

Erik is 5 years old and throws temper tantrums when he does not get his way. Of particular concern is that he is beginning to hit his newborn baby brother with increasing force. The therapist, in consultation with Erik's parents, established this plan: If Erik hits or attempts to hit his brother, his parents will direct him to sit in the time-out chair. If he refuses, his parents will, in a matter of fact fashion, place him in time-out. Erik will have to sit in a chair facing the corner of the living room, and his parents will not speak to him during that 5-minute interval. If Erik attempts to leave the chair early, the 5-minute interval will be restarted. Likewise, if Erik tantrums in the chair, the time-out interval is reset. Erik cannot leave time-out until he sits quietly for a continuous 5 minutes.

Aversion-Based Approaches

Like exposure therapies, aversion-based approaches consist of different interventions; some based on operant principles and others based on counter-conditioning strategies. Time-out and response cost, discussed above, are punishment procedures that decrease the probability of target behaviors. In other forms of punishment, a noxious stimulus or event is

applied in order to decrease the likelihood of a behavior. In this scenario, punishment is defined functionally as the reduction in the future probability of a response as a result of the immediate delivery of a stimulus for that response. The stimulus is called a punishing stimulus. Punishing stimuli include loud noise, aversive tastes, mild electrical shock, or chemical agents, the latter used only with adults. Behavior therapists use punishment as a last resort, when other methods such as time-out or response cost fail to get the target behavior under control. Punishment is also used when the negative behavior is very dangerous to the person (e.g., self-mutilative behaviors such as eye gouging) and must be suppressed immediately.

Three general parameters are associated with all punishment procedures. To be effective, punishment needs to be of high intensity (but only intense enough to disrupt the behavior), immediate, and continuous (at least initially). Less aversive stimuli such as sprays of warm water to the face or a drop of lemon juice applied to the tongue are tested first, prior to the use of higher intensity stimuli. When used effectively, punishment can lead to the permanent suppression of a behavior. In fact, there is a direct relationship between the strength of the punishing stimulus and recovery of a previously suppressed behavior; the weaker the punishing stimulus, the more likely it is that the negative behavior will recur. In fact, insufficiently intense punishment can produce increased responding following termination of punishment.

Punishment may produce undesirable effects such as negative emotional reactions, escape and avoidance, and aggression. There are numerous ethical issues to consider prior to the implementation of punishment. As noted already, punishment is used only as a means of last resort. Used primarily with children, it may also be useful for adults with developmental disabilities who engage in self-injurious behaviors. Treatment plans that incorporate the use of punishing stimuli are never the decision of one person but are reviewed by various ethical and legal entities, prior to their implementation.

Other aversive procedures, such as *aversion therapy,* are based on a classical conditioning paradigm. Used to treat sexual deviations and substance abuse, an aversive physical and/or emotional state is paired with cues that elicit abusive behavior or sexually deviant behavior. The person develops an aversion to the sights, smell, tastes, sounds, and other aspects of maladaptive behavior. The result is the elimination or reduced frequency of the unwanted behavior. For example, if the person is presented with repetitive pairing of the deviant stimuli (engaging in sexual activity with a prepubescent child) and the aversive stimulus (inhaling smelling salts), the deviant behavior suppressed. Three different stimuli have been used as aversive conditioning stimuli: chemical aversion, electrical aversion (sometimes called faradic aversion), and verbal aversion (also known as

covert sensitization).

Chemical aversion may be administered via the olfactory or gustatory systems. *Olfactory aversion* is the pairing of extremely noxious but harmless odors (such as ammonia) with either sexual fantasies or sexual behaviors. Typically, the person is presented with the deviant sexual stimuli and then inhales the ammonia fumes, which cause burning and watering eyes, runny nose, and coughing. With repeated pairings, the deviant sexual behavior is suppressed, usually within a few weeks. For substance abuse, *emetic aversion* (a chemical that induces vomiting) has also been used. Similarly *faradic aversion* (electrical stimulation to the forearm) and *verbal aversion* (covert sensitization) have also been used. Covert sensitization involves the presentation of an imaginal scene involving deviant behavior that is immediately followed by a scene that produces nausea. The imaginal scenes are individually developed, as there are individual differences in what produces nausea. All three forms of aversion therapy appear to have positive effects for the treatment of sexual deviations or substance abuse, although in the case of substance abuse, aversion therapy alone is rarely sufficient to treat the abuse/dependence. Aversion therapy, in the form of rapid smoking, has been used to treat nicotine addiction; however its efficacy for this form of substance use is not yet clear.

Skills Training

Skills training approaches have a long history of use in behavioral therapy. From a behavioral perspective, lack of skill may play a role in the onset and maintenance of various types of psychopathology. For example, people with social phobia or schizophrenia report patterns of social isolation that may have existed since preadolescence. In the case of social phobia, decreasing anxiety will be ineffective if part of the anxiety is driven by a lack of knowledge and skills in social situations. Similarly, psychiatric rehabilitation efforts directed at patients with schizophrenia will be limited if they are not given the skills necessary to interact with others, to interview and maintain a job, and to become capable of living independently. Finally, patients with Axis II disorders have chronic difficulties interacting and maintaining relationships with other people. In these and other instances, *social skills training (SST*) may need to be an integral part of the treatment program.

Social Skills Training. SST is a broad-based program designed to teach a wide variety of social skills that can be fitted to the particular needs of individuals and groups. SST usually includes a focus on verbal and nonverbal behaviors (e.g., eye contact and speech latency) and can be used with individuals who have any type of psychiatric disorder and are in need of learning how to interact with others effectively. Possessing basic

conversation skills is essential for successfully completing social and performance tasks and allowing one to master social interactions and overcome maladaptive anxiety. Even for those who seem to possess some social skills, the inclusion of a skills training component into an integrated treatment plan serves as an opportunity for skill refinement and practice in a carefully controlled setting.

Conceptually, social skills training content can be divided into two broad components: *Social Environment Awareness* and *Interpersonal Skill Enhancement*. Together, these components address a specific set of common problems. Social Environment Awareness (sometimes called social perception) involves teaching the nuances of when, where, and why to initiate and terminate interpersonal interactions. Interpersonal Skill Enhancement includes teaching verbal and nonverbal mechanics of successful social encounters (e.g. initiating, maintaining, and ending conversations; establishing and maintaining friendships; and assertiveness). It also may include more complicated skills such as anger management skills, job seeking skills, and public speaking skills.

Five techniques are used for social skills training: (1) Instruction in the skill, (2) modeling the skill, (3) behavioral rehearsal (role-play practice), (4) corrective feedback, and (5) positive reinforcement. Social skills training is conducted within the context of role-play scenes, which are used to assess baseline skill level and to allow for the behavioral rehearsal of new skills. The scenes are designed to depict various types of skills (e.g., assertiveness, initiating a conversation). Dismantling studies designed to determine the critical elements of SST indicate that behavioral rehearsal is the critical element in social skill acquisition. Below is a typical curriculum used to teach social skills to adults with generalized social phobia.

Week	Topic
1	Social perception skills and greetings
2	Initiating conversations
3	Maintaining conversations
4	Listening and remembering
5	Inviting others to join activities
6	Compliments
7	Assertiveness
8	Heterosocial skills
9	Introduction to public speaking skills
10	Beginning and ending a speech
11	Body of the speech
12	Small group interactions

"Second Wave" Behavior Therapy Treatment Strategies

With increased interest in cognitive theory in the 1970s, interventions

based exclusively on operant and respondent principles, once thought adequate, were viewed as insufficient, and the absence of direct cognitive manipulations was widely regarded as a limitation of behavioral therapies. This changing zeitgeist was reflected in the increasing popularity of cognitive therapy and culminated in the inclusion of this treatment (and exclusion of behavioral therapy) in the *Treatment of Depression Collaborative Research Program* funded by the *National Institute of Mental Health*. This large study included individuals diagnosed with major depression and compared the relative efficacy of CBT, interpersonal psychotherapy (IPT), and antidepressant medication (Imipramine) with a placebo control group. Among less severely depressed individuals, there were no differences across active treatment conditions and all were superior to placebo. For more severely depressed individuals, however, Imipramine and IPT were deemed more effective then CBT. Important to point out, reanalyses of study data have since indicated that when only outcomes of patients treated by highly skilled CBT therapists were included in analyses, CBT was as effective as IPT and medication, even for severely depressed patients. Several more recent studies on both "second" and "third" wave therapies (e.g., BA) have supported the notion that CBT is effective for severely depressed individuals, equally as effective as antidepressants for this group, and potentially more beneficial than medications for individuals with less severe depression. Moving into the 21st century, the distinction among interventions for depression (and other disorders) considered purely "behavioral" of "cognitive" has become blurred because of their significant conceptual and technical overlap.

In the transition into "second wave" therapies, mental health practitioners began to empirical study how thinking (cognitions) affected emotions and behavior. This was a revolutionary challenge to traditional behavior therapy and sometimes is referred to as the "cognitive revolution." The history of the cognitive revolution in psychotherapy largely was assisted by two other developing sciences. The first was social psychology, where pioneers such as Fritz Heider, Harold Kelley, and Bernard Weiner were beginning to research and understand the complex ways in which individuals interact and think about one another's behaviors. As mentioned earlier, Albert Bandura's "social learning" model as well as "attribution theory" were the main contributions of social psychology to the development of cognitive therapy. Social attribution research explored the human tendency to think about or "attribute" causes to the actions of other people. The second boost to the cognitive revolution was the development of computer science and computer programming. Computer software and software programs were a perfect analogy for understanding the "programming rules" for human brains and behavior. As a result of such developments, many behavior therapists reconceptualized patient difficulties by considering the role of conscious thoughts and no longer

restricted themselves to the assessment and modification of only observable behavior. This combination of scientific disciplines and changing philosophy resulted in what is now referred to as "Cognitive-Behavior Therapy" (CBT).

CBT is a structured, short-term, present-oriented psychotherapy, directed toward solving problems and modifying dysfunctional, inaccurate, and irrational thinking and behavior. CBT refers to behavior therapy, cognitive therapy, and therapy that combine basic behavioral and cognitive principles. CBT acknowledges there may be behaviors that cannot be controlled through rational thought, but rather emerge from prior learning experiences and other external and internal stimuli. CBT is "problem focused" and "action oriented" in that the clinician is directive and assists patients in selecting specific strategies to address presenting problems. CBT is an effective treatment of many psychological conditions, including mood, anxiety, personality disorders, eating, substance abuse, tic, and psychotic disorders. Many CBT interventions for specific disorders have been evaluated for efficacy, resulting in their designation as evidence-based or empirically supported treatments. Several of the most common "second wave" therapies are presented here, including those without a cognitive component and arguably more aligned with "first wave" behavioral theory, but included in this section given their more recent development.

Cognitive Therapy

A founding co-father of cognitive therapy is Aaron Beck, and the fundamental premise of cognitive therapy is that abnormal behavior is the result of distorted cognitive (mental) processes. In other words, situations and events do not affect one's emotions and behaviors directly, but rather the way one thinks about and perceives those events does. For example, Beck frequently referred to the *cognitive triad,* or the tendency of depressed individuals to have three types of negative thoughts: a negative view of the self, the world, and the future. These negative beliefs are often called *cognitive distortions,* or patterns in thinking that affect individuals' mood and behavior. Although there are many types of cognitive distortions, a few of the most common are *all-or-none thinking* ("If I don't get this new job I will be a complete failure in life"), *catastrophizing* ("I did poorly on the examination and now my professor is going to hate me"), *mindreading* ("That person looked at me strangely, he must think I'm unattractive"), and *personalizing* ("It rained today and I had to cancel the family picnic, it's all my fault"). Although most frequently referenced in the context of depressive disorders, individuals with various psychological problems may consistently distort interpretations of events (even unconsciously) so that they maintain negative views of themselves, their environment, other people, and the future. In the range of approximately 20 to 25 1-hour weekly sessions, the focus of cognitive therapy is on (a) eliciting *automatic*

thoughts, or thoughts that intervene between events and the individual's emotional reaction to them; and (b) engaging in the process of *cognitive restructuring* or reframing to examine the validity of automatic thoughts and learn to modify their thinking when thought patterns are irrational or inconsistent with available evidence from their life and daily experiences. Traditional cognitive therapy is unique in emphasizing this "hypothesis testing" strategy to assess the accuracy and usefulness of thinking behavior. Largely as a result of the influential work of Aaron Beck, the Academy of Cognitive Therapy was developed in 1996 to officially standardize and certify therapists doing this form of cognitive therapy.

The other founding co-father of cognitive therapy was Albert Ellis, who developed rational emotional behavior therapy (REBT), previously called rational therapy and RET. His highly impactful work began in the mid-1950s and continued until his death in 2007. One of the fundamental premises of REBT is that the major cause of human suffering is that people are largely distressed by events and adversities because of how they construct their views of reality through their language, evaluative beliefs, meanings and philosophies about the world, themselves and others. In REBT, patients initially learn the *A-B-C*-model of psychological disturbance and change. The A-B-C model states that it normally is not merely an *A, a*dversity (or *a*ctivating event) that contributes to disturbed and dysfunctional emotional and behavioral *C*s (consequences), but also what people *B* (believe) about the *A* (activating event). The activating event can be either an external situation or internal event (e.g., a thought) and can refer to an event in the past, present, or future.

The *B*s, beliefs that are most important in the *A-B-C* model are explicit and implicit philosophical meanings and assumptions about events, personal desires, and preferences. The *B*s, *b*eliefs that are most significant are highly evaluative and consist of interrelated and integrated cognitive, emotional, and behavioral dimensions. According to REBT, if a person's evaluative belief about the activating event is rigid, absolutistic, and dysfunctional, the emotional and behavioral *c*onsequences are likely to be self-defeating and destructive. Alternatively, if a person's evaluative belief is preferential, flexible, and constructive, the emotional and behavioral *c*onsequences are likely to be self-helping and constructive. Through REBT, by understanding the role of their mediating, evaluative and philosophically based illogical, unrealistic and self-defeating meanings, interpretations and assumptions, people often can learn to identify them, begin to *D* (dispute, refute, challenge, and question them), distinguish them from healthy constructs, and subscribe to more constructive and adaptive perspectives of their world.

According to REBT, people to a large degree consciously and unconsciously construct emotional difficulties such as self-blame, self-pity, clinical anger, hurt, guilt, shame, depression and anxiety, and behaviors

and behavior tendencies like procrastination, over-compulsiveness, avoidance, addiction and withdrawal by means of their irrational and self-defeating thinking, emoting and behaving. REBT is then applied as an educational process in which the therapist often active-directively teaches the client how to identify irrational and self-defeating beliefs and philosophies which in nature are rigid, extreme, unrealistic, illogical and absolutist, and then to forcefully and actively question and dispute them and replace them with more rational cognitions. Similar to Beck's strategies, using emotive and behavioral methods and activities, the patient and therapist construct homework exercises to gain a more constructive rational way of thinking, emoting, and behaving. The Albert Ellis Institute (AEI) was founded in 1959, and Ellis continues to be one of the most influential psychotherapists in the history of clinical psychology.

Several other cognitive therapies subsequently have been developed, including Patricia Resick's *Cognitive Processing Therapy* (1993) and Jeffrey Young's *Schema Therapy* (1990), interventions demonstrated to be effective for problems such as depression, suicidal ideation, anxiety, trauma and PTSD, substance abuse, eating disorders, personality disorders, and marital problems. Although beyond the scope of the chapter to describe in detail, the important point is that all cognitive therapies to a greater or lesser degree include behavior-based assessment and intervention methods.

Problem-Solving Therapy. Psychological difficulties can result from ineffective methods by which to approach life's challenges and problems. Psychologists have identified three different types of problem-solving: (1) *impulsive/careless problem solving,* characterized by impulsive, hurried and careless attempts at problem resolution; (2) *avoidant problem solving,* defined by procrastination, passivity and over-dependence on others to provide solutions; and (3) *rational problem solving,* symbolized by a problem-solving style that is constructive, systematic, and planful. The use of these skills contributes toward the identification of an adaptive solution or coping response. When patients do not possess these behaviors, PST teaches them the skills necessary to discover effective solutions. PST is effective for a wide range of clinical populations (e.g., schizophrenia, depression, and eating disorders), psychological problems (e.g., smoking and marital discord) and stress associated with chronic medical disorders.

The specific goals of PST are to (1) increase patient understanding of the connection between current depression and anxiety symptoms with everyday problems, (2) increase the ability to clearly and accurately define current problems, (3) teach patients a specific problem-solving method to allow for a more structured skill set to address life problems, and (4) create more positive experiences through patients' improved abilities to solve problems. The basic premise of PST is that as patients learn to identify and

resolve problems, they gain an increased sense of self-efficacy, control, and confidence while becoming more active in eliciting reward from their environment. According to problem-solving and behavioral theories, such changes serve to reduce symptoms of depression, anxiety, and other negative emotions. Patients receiving PST generally receive eight weekly psychotherapy sessions (of one hour in duration). Problem-solving treatment includes seven stages: (1) understanding the PST treatment and its rationale; (2) identification, definition, and breaking down of the problem; (3) establishing achievable goals for problem resolution; (4) generating possible solutions; (5) evaluating and choosing the solution; (6) implementing the chosen solution; and (7) evaluating the outcome. During the course of PST, patients are encouraged to continue to modify their problem list as new problems arise or as other problems not initially identified are recalled. Gradually throughout treatment, patients take a more prominent role in performing the problem-solving stages, requiring less input and guidance from the therapist as they became more skilled at implementing the stages.

Shirley had been feeling depressed for a long time. It seemed like she just could not get her life together. She had not been very effective in her work when she was younger; now that she was older, she was having even more trouble motivating herself to get anything done. She was lethargic and had trouble making decisions. Since she often could not decide what to do next, she often did nothing. When she decided to see a therapist, he suggested that learning how to solve problems more effectively might help her depression. Shirley figured it was worth a try. She was able to identify a number of problems that needed to be solved and made her feel distressed—a number of appliances in her house needed to be fixed; she and her husband just did not seem interested in the same things anymore; and she had a lot of papers on her desk that needed to be sorted and filed. The therapist suggested they start with the appliances. Shirley and the therapist worked together to come up with a list of possible solutions, some of which were rather funny—but the therapist had said to name anything that came to mind, not to evaluate whether it was a good strategy or not. Here is the list they came up with:

1. Throw away all the nonworking appliances.
2. Give the appliances to a charity and buy new ones.
3. Call a friend to get the name of a good repairman.
4. Get the name of a repair company from the phone book and call them.
5. Call a friend to ask if he can try to fix the appliances.
6. Do nothing—just live with the appliances not working.

Shirley then went back and rated each of these possible solutions for its strengths and weaknesses. She was able to choose one, #3. She went home and called her friend. Shirley felt better just after getting this first step accomplished.

Self-Control Approaches

Traditionally, hypothesized constructs such as “mind” or “will power” have been viewed as operating outside of the laws of human behavior, reflecting an inner mechanism of control for human behavior. Behaviorists questioned the intrapsychic view, noting that behavior can best be

understood in terms of (1) setting characteristics, (2) learning history, and (3) transient organismic states. Although many behavioral treatments involve some elements of self-control, as patients take responsibility for changing behaviors, more direct self-control approaches are defined by the use of stimuli, behaviors, and consequences with oneself to achieve a desired outcome. Common self-control strategies include physically restraining oneself, chaining the stimuli conditions, depriving or satiating oneself, using mild aversive events, using drugs, alcohol or stimulants, administering self-reinforcement or self-punishment, and doing something else to avoid aversive consequences. Self-control procedures may be developed in conjunction with a therapist, but are usually administered by the patient to him or herself. Self-control therapies are efficacious for changing maladaptive and excessive behaviors that contribute to an unhealthy lifestyle such as overeating and smoking.

Self-control therapies generally involve three basic processes. The first stage is self-monitoring (described earlier, see Self-Monitoring), of one's own performance in the context in which it occurs. For self-regulation to occur, there must first be understanding of the behaviors that are in need of change. Consider the case of a woman who is overweight and seeks the help of a health professional to lose weight. She is assigned an eating diary, where she is instructed to monitor what she eats, how much she eats, as well as where and when she eats.

The second stage of self-control is self-evaluation. After self-monitoring, the individual compares his or her performance with some standard (internal, external, or both). After monitoring her eating behaviors for 2 weeks, she discovers the following: (a) her total caloric consumption is 2,300 calories per day; (b) in addition to eating three meals, she eats 6 snacks per day, usually when she is working at her desk or when watching television; and (c) she exercises only once per week. Most standards are developed through direct reinforcement or modeling. The patient decides to reduce her caloric intake by limiting herself to one snack per day, only eating while at the dining room table, and increasing her exercise ritual to three times per week.

Self-reinforcement is the third phase. It involves the self-communication of positive or punishing stimuli contingent on good or poor performance. Self-reinforcement operates identical to overt reinforcement. In order for self-control to work, self-reinforcement is necessary. Self-control programs are most likely to work when reinforcement is directed at a specific behavior and not effects of behavior. In the case of the woman who is seeking to lose weight, she buys some comfortable "fun" workout clothes, joins a gym where she can socialize with others after working out, and sets up a new wardrobe account, where she pays herself $1.00 every time she exercises or refuses a snack or high calorie food. It is important to note that the reinforcement is provided for engaging in positive, healthy

behaviors, not for weight loss (which should result from positive healthy behaviors).

Another type of self-control strategy is called *stimulus control.* A basic tenet of behavior therapy is that behaviors occur in the presence of specific stimuli, which serve as cues for the behavior. When these cues are present, they increase the likelihood that the behavior will occur. When stimulus control is maladaptive, behavior problems may occur. Insomnia, for example, has been successfully treated with stimulus control procedures. Even though insomnia may develop for different reasons, it is commonly maintained by poor sleep hygiene. Assessment of the behaviors associated with insomnia often indicates that people go to bed but engage in a variety of activities incompatible with sleep—eating, reading, watching television. They often take naps during the day, lessening the likelihood that they can sleep at night. Finally, when they do fall asleep, they sleep throughout the morning, shifting their sleep schedule dramatically. Behavioral interventions based on stimulus control provide patients with a sleep hygiene plan, which might include going to bed only if sleepy, getting out of bed if sleep does not occur within 10 minutes, prohibition against eating, reading or watching television in bed, elimination of naps, and getting out of bed in the morning in order to reset the sleep–wake schedule.

Habit Reversal Training

Habit reversal training (HRT) is a behaviorally based treatment approach designed to reduce repetitive behaviors (habits) such as hair pulling, stuttering, nail biting, and tics. From a behavioral perspective, these habits become associated with various situations and cues through the process of classical conditioning. They also are reinforced operantly as they serve to reduce stress. HRT as developed by Nathan Azrin and his colleagues was designed to break the conditioned habit cycles using a combination of procedures including awareness training, habit control motivation, competing response training, relaxation training; and reinforcement. *Awareness training* involves use of self-monitoring to increase patient awareness of internal and external precipitants of the repetitive behavior. Through this process, patients learn which situations put them at "high risk" of the habit to occur. They also are taught to think of the habit not as a single behavior but as a sequence of behaviors that includes precipitating motoric responses (e.g., face touching that precedes hair pulling or nail biting, shallow breathing that occurs before stuttering) and internal cues (thoughts, feelings), as well as behavioral consequences and reinforcers.

Habit control motivation is another component of the early stages of HRT. This work is akin to motivational interviewing and involves a review of the inconveniences of the habit (e.g., how does it get in the way and what are the reasons for changing it). *Competing response training* is the core of HRT. This process involves identifying a motoric behavior that is

incompatible with the habit. Fist-clenching is a commonly used response that is incompatible with what are called *body focused repetitive behaviors,* such as hair pulling, nail biting, and skin picking. Fist clenching can be done inconspicuously by making a fist, clenching a book, or firmly grasping a steering wheel. Patients learn to engage in these competing behaviors when they are in high-risk situations and when they notice precipitating behaviors or internal cues. Completing the alternative behavior for intervals of 1 to 3 minutes breaks the conditioned stimulus-habit cycle.

Susan pulled her hair most often when she was reading and watching television. Through awareness training, she noticed that her left hand went to her hair almost as soon as she sat down. Her hand moved over her head, feeling for any hairs that were of a different texture. Susan then typically pulled the coarse stubby hairs, almost without awareness. Using the fist-clenching technique, however, gave her something else to do with her hands. The therapists even suggested that she get a hand exerciser to keep her hands occupied and away from her hair when she was watching television. It was amazing how much this helped her cut back on the number of hairs she pulled.

Relaxation training, discussed earlier in this chapter, is also used to help break the conditioned habit cycles, particularly when stress or another negative mood is a precipitant for the target behavior. Finally, reinforcement is used as parents, therapists, or significant others provide rewards for reduced frequency of the targeted habit. For some patients, a token economy system or a self-control reinforcement program (both discussed earlier in this chapter) can help to facilitate and maintain treatment gains.

However, recent behavioral models of these disorders have produced more comprehensive approaches that suggest the need to incorporate a broader array of treatment techniques to fit a wider range of functional relations between the habit and its antecedents and consequences. Patients are asked to look not only for situational and motoric precursors of the repetitive behavior, but also cognitive, affective, and sensory cues that may precipitate the habit. For example, a person with repetitive hair pulling (trichotillomania) may have thoughts like, “That hair is too grey. My eyebrows are uneven,” that increase the risk of pulling. More general thoughts or even core beliefs (discussed earlier) may also precipitate pulling. In addition, a range of affective or mood states (e.g., anxiety, depression, anger, and boredom) and sensory cues (e.g., rubbing ones face and touching one’s hair) may be important precipitants. Given these broader functional relations, a wider range of coping skills is needed. For example, a patient who is reinforced for hair pulling by the sensations of biting off the hair root may instead chew on sesame seeds. A patient whose nail biting occurs primarily in response to interpersonal stress may be taught a wider range of relaxation responses, communication skills to

improve social relationships, or ways to modify thinking to decrease stress.

"Third Wave" Behavior Therapy Treatment Strategies

As mentioned earlier, during the last two decades a number of therapies, under the name of the "third wave" of CBT, have been developed: BA, ACT, MBT DBT, CBASP, and FAP. The therapies are "grounded in empiricism, principle-focused, particularly sensitive to the context and functions of psychological phenomena, not just their form, and thus tend to emphasize contextual and experiential change strategies in addition to more direct and didactic ones. These treatments seek the construction of broad, flexible and effective repertoires over an eliminative approach to narrowly defined problems, and emphasize the relevance of the issues they examine for clinicians and patients. The third wave reformulates and synthesizes previous generations of behavior and cognitive therapy and carries them forward into questions, issues, and domains previously addressed primarily by other traditions, in hopes of improving both understanding and outcomes." These are the most prominent "third wave" therapies.

Behavioral Activation

Based on behavioral theory, depression persists because: (a) reinforcement for nondepressed (healthy) behavior is low or nonexistent, (b) depressed behavior produces positive (e.g., sympathy from others) or negative reinforcement (decrease in responsibilities or avoidance of undesirable situations), (c) exposure to aversive or unpleasant life experiences is significant, or (d) some combination of these factors. Severely depressed mood leads to behavioral withdrawal, which in turn results in loss of reinforcement from pleasurable events or activities (see Fig. 33.3–2).

Developed concurrently in West Virginia (Hopko and Lejuez) and Washington (Jacobson and Martell), BA emphasizes structured attempts to increase overt behaviors that bring patients into contact with reinforcing environmental contingencies and corresponding improvements in thoughts, mood, and quality of life. Within the BA model, the process of increasing response-contingent reinforcement follows the basic behavioral principles of extinction, shaping, fading, and in vivo exposure. Initial sessions involve assessing the function of depressed behavior, establishing patient rapport, motivational exercises focusing on the pros and cons of behavioral change, depression psychoeducation, and introduction of the treatment rationale. Once efforts are made to reduce reinforcement for depressed behavior, a systematic approach for increasing healthy behavior is initiated by increasing the value of reinforcers for such behavior and devaluing reinforcers for depressed behavior.

Within this model, systematically increased activity is a necessary precursor toward the reduction in overt and covert depressed behavior.

Patients began by engaging in a self-monitoring (or daily diary) exercise to examine already occurring daily activities to provide a baseline measurement and provide some ideas with regard to identifying potential activities to target during treatment. For each behavior, they are asked to indicate their levels of reward, difficulty, and importance. Following monitoring, emphasis shifts to identifying patients' values and goals within life areas that include family, social, and intimate relationships, education, employment/career, hobbies/recreation, volunteer work/charity, physical/health issues, spirituality, and anxiety-eliciting situations. Following this exercise, an activity hierarchy is constructed where 15 activities are rated from "easiest" to "most difficult" to accomplish. Using a behavioral checkout to monitor progress, patients progressively move through the hierarchy, moving from easier behaviors to the more difficult. For each activity, the therapist and patient collaboratively determine the *final goal* in terms of the frequency and duration of activity per week. At the start of each session, weekly activation assignments are examined and discussed, with the following weekly goals established as a function of patient success or difficulty. BA can also be used to address coexistent anxiety problems by including anxiety reduction strategies into the structured hierarchy, such as PDMR, assertiveness training, and graduated exposure to anxiety-eliciting stimuli. In general, BA treatment includes between 8 and 20 sessions of approximately 1 hour in duration.

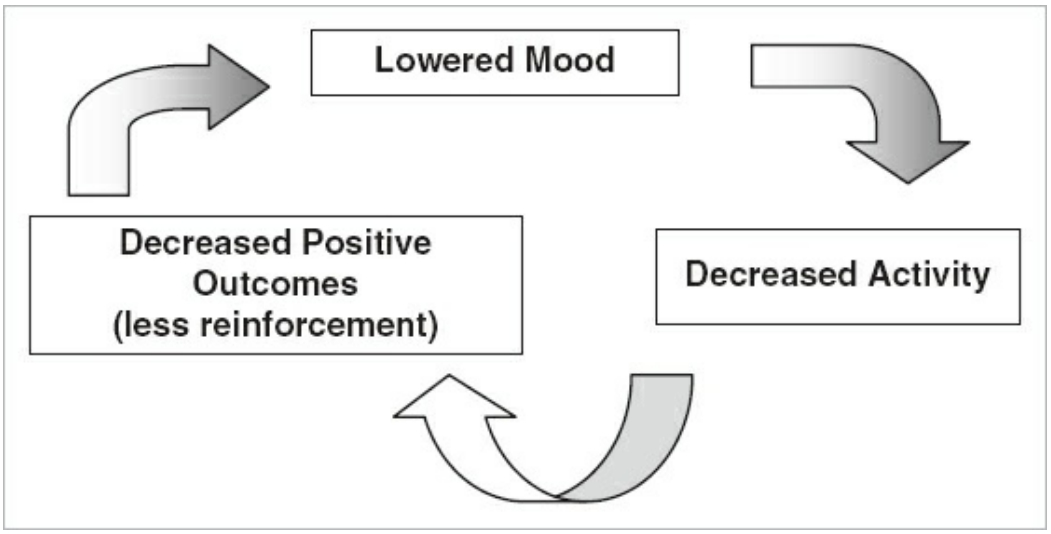

FIGURE 33.3–2. Positive cycle of increased activity, improved mood, and reduction in depressive symptoms.

BA has been effectively used with patients in community mental health centers, inpatient psychiatric facilities, college psychology clinics, and primary care settings, even when administered by previously untrained mental health nurses. In addition to major depression experienced by younger and older adults, BA has been shown to be effective for grief and

bereavement, smokers and drug users, anxiety disorders such as PTSD and GAD, schizophrenia, and for depressed patients with cancer, obesity, and HIV/AIDS. Recent data suggest BA leads to greater improvement than cognitive therapy for patients with severe depression, with equivalent outcomes to antidepressant medication.

Charles was a 70-year-old retired business executive. Throughout his life, his work consumed him. Although he married and had a family, his job was his primary focus. He went to the office early and came home late. He enjoyed what he did—it was stimulating and made him feel important and useful. But as he got older, his performance was not what it used to be and he decided it was time to retire. However, his mood was pretty low when he no longer had a job. He did not have energy to get more involved in his church or to develop other hobbies, so he sat around all day, without any social contacts. Charles; wife and best friend encouraged him to go talk to someone. The therapist suggested that they try BA. Charles was somewhat skeptic, as it seemed too simple, but he needed to do something. The therapist spent some time with Charles talking about the kinds of activities that used to make him feel good—and some of the things he used to enjoy. They then put together a list of things he might be able do—even if he did not feel much like it—just to see what would happen. The list included looking for volunteer work where he could use his job skills, spending more time with his wife in some of the activities they once had enjoyed (e.g., watching movies and taking walks), and rejuvenating an old hobby from his college days—fishing. Charles initially agreed to do some easy activities—go to one movie a week, take one walk a week, and contact his church activity leader about possible volunteer activities. He was surprised to find that even these "baby steps" helped him feel better. He had the chance to talk with other people and begin to see that even in retirement, he could find useful and fun things to do.

Acceptance and Commitment Therapy

Originally referred to as comprehensive distancing, acceptance and commitment therapy (or ACT) was developed in the 1990s by Steven Hayes and Colleagues. ACT is based on acceptance and mindfulness strategies, together with commitment and behavior change strategies, to increase psychological flexibility. Psychological flexibility means contacting the present moment fully as a conscious human being, and based on what the situation affords, changing or persisting in behavior in the service of chosen values. The core premise of ACT is that much of human (psychological) suffering is a consequence of *experiential avoidance,* or the unwillingness to experience negative feelings, thoughts, and sensations. Through metaphor, paradox, and experiential exercises clients learn how to make healthy contact with thoughts, feelings, memories, and physical sensations that have been feared and avoided. Clients gain the skills to recontextualize and accept these private events, develop greater clarity about personal values, and commit to needed behavior change. ACT views the core of many problems to be due to the concepts represented in the acronym, **FEAR: F**usion with your thoughts, **E**valuation of experience, **A**voidance of your experience, and **R**eason-giving for your behavior. According to ACT, the healthy alternative is to **ACT: A**ccept your reactions and be present focused, **C**hoose a valued direction, and **T**ake action.

Mindfulness-Based Therapy

Mindfulness-based treatments are a relatively new variation in behavioral (and cognitive) interventions that integrate various techniques based in Eastern philosophies of contemplation and meditation. MBTs, such as mindfulness-based stress reduction (MBSR) and mindfulness-based cognitive therapy (MBCT) show promise as effective interventions for anxiety, depression, and other psychological problems. The fundamental principles of MBT involve attention regulation, openness to present experience, curiosity and acceptance of the "here-and-now," and nonjudgmental awareness of thoughts, emotions, sensations, and the environment. The efficacy of MBT in treating emotional disorders is believed to be a function of an increased ability to manage life stressors, enhanced concentration and mindfulness, and reduced rumination. In MBSR, the primary strategies include learning mindful eating (e.g., the raison exercise), mindfulness meditation, body (scanning) awareness, mindful walking, urge surfing, and yoga. Originally developed in the mid-1990s as a treatment to prevent relapse into depression, MBCT integrates fundamental theories and applications of MBSR with those of more traditional CBT. Unlike CBT, however, MBCT does not emphasize changing the "content of thoughts," but emphasizes changing awareness of and relationship to thoughts. Aspects of CBT included in MBCT are primarily those designed to facilitate "decentered" views, such as "thoughts are not facts," and "I am not my thoughts." Although interventions such as MBSR and MBCT are predominantly administered over 8 (3-hour) weekly sessions followed by a daylong retreat, some data support abbreviated formats of these protocols.

Dialectical Behavior Therapy

DBT was originally developed as an intervention for chronically suicidal individuals. DBT blends behavioral and crises intervention theories with an emphasis on acceptance and tolerance based on both from Western contemplative and Eastern meditation practice. The theoretical framework emphasizes an exchange and negotiation between therapist and patient, between the rational and the emotional, and between acceptance and change (hence the term "dialectical"). Treatment targets are agreed upon, with self-harm taking priority. The learning of new skills is a core component–including mindfulness (e.g., nonjudgmental awareness of one's own feelings and actions), interpersonal effectiveness (e.g., assertiveness and social skills), coping adaptively with distress and crises, and identifying and regulating emotional reactions. DBT involves both individual and group therapy.

Cognitive Behavioral Analysis System of Psychotherapy

Developed by James McCullough in the late 1990s, CBASP is an integrative therapy for chronically depressed adults that combines components of cognitive, behavioral, interpersonal, and psychodynamic therapies. According to this model, those with chronic depression experience disconnectedness from their environment and therefore have decreased access to important feedback on problematic interpersonal patterns and relationships. The therapeutic relationship is actively used to help patients generate empathic behavior, identify and change interpersonal patterns related to depression, and heal interpersonal trauma. CBASP consists of three techniques: (1) *Situational Analysis,* a problem-solving technique designed to help the patient realize the consequences of his/her behavior on others and modify it; (2) *Interpersonal Discrimination Exercises,* examination of past traumatic experiences with others and differentiation of those from healthier relationships; and (3) *Behavioral Skill Training/Rehearsal,* such as assertiveness training, to further help depressed individuals modify maladaptive behavior. McCullough also strongly recommends that all patients beginning CBASP begin a regime of antidepressant medication.

Functional Analytic Psychotherapy

FAP is a treatment based on a behavior-analytic understanding of the therapeutic relationship as an agent of change. Although much of behavior analysis focuses on environmental manipulation, FAP presumes that significant change occurs in the context of the therapeutic relationship during the traditional weekly psychotherapy session. Because many problems that lead to distress and impaired role function are the result of interpersonal deficits, the therapy environment is made to have features much like the natural environment. The therapist and client develop a case conceptualization based on the client's description of clinical problems and the therapist's interactions with the client in the context of an intense interpersonal relationship. Change occurs as the therapist helps the client to discriminate opportunities to behave in more adaptive ways and then shapes the client's repertoire so that social interactions result in meaningful social reinforcement. In therapy, the therapist becomes an important mediator of interpersonal reinforcement. It is the contingent responding of the therapist to clinical problems and improvements that occur in the therapy session that produces change that will be maintained outside the therapy setting. In the FAP model, each client is thought to exhibit in-session behaviors that are functionally similar to the behaviors he or she emits in daily life. Client problematic and functional behaviors (referred to as clinically relevant behaviors; CRBs), as determined by a functional idiographic assessment, are contingently responded to in the session. A FAP therapist seeks to reduce (functionally punish or extinguish) the client's dysfunctional behavior (CRB1) while increasing the probability

of (functionally reinforce) more appropriate alternative behavior (CRB2) by using naturalistic, rather than arbitrary, feedback and reactions.

Third wave behavioral treatments have become increasingly studied and demonstrated as effective for a wide range of psychological problems including borderline personality disorder, anxiety, chronic pain, depression, and stress. The approaches also have been used to reduce dysfunction in patients with medical conditions (e.g., cancer and multiple sclerosis) and to increase general well-being and quality of life. Important to note, many researchers and clinicians believe that relative to traditional CBT, the evidence base of "third wave" therapies is limited in quantity, quality, and breadth of available studies, precluding any firm conclusions about their short- or longer-term equivalence. Additionally, while subtle and important differences on theoretical and procedural levels might exist, available data do not suggest differential mechanisms of action that warrants a dramatic separation from the CBT family of approaches. Accordingly, some researchers propose "new wave" treatments are consistent with the CBT approach and that the term "new wave" is potentially misleading because it is not an accurate reflection of the contemporary literature.

CLINICAL ISSUES

Indications

Behavior therapy has a wide range of applications and are useful for a wide variety of psychiatric disorders (e.g., anxiety, depression, schizophrenia, and bipolar disorder), psychological problems (e.g., marital discord, school refusal, and stress), and medical problems (e.g., adjustment to chronic illness, medication adherence, and healthy behaviors). Behavioral treatments also have been administered effectively with individual patients, couples, families, and groups, and they are useful in a range of treatment settings, including inpatient, outpatient, day treatment, and community settings. Recent work also has demonstrated the utility of behavioral treatments offered in primary care settings. Many people with psychological difficulties (in particular, anxiety, depression, and substance abuse) do not seek help in mental health settings, but rather present for care to medical clinics. In these settings, behavioral treatments offered by mental health specialists and other care providers have been effective for reducing symptoms and improving quality of life. Providing care in medical settings also increases efficiency of service provision and decreases cost. Medical patients with psychological symptoms and disorders are often high utilizers of medical care, and treating the psychological symptoms can reduce the overall cost of health care for these individuals.

Developmental Considerations. Effective behavioral treatments also

have been developed for use across the full developmental spectrum (children, adults, older adults), and these approaches are particularly valuable for patients whose verbal skills are not appropriate for more traditional therapeutic approaches. However, when using behavioral treatments across the life span, it is important to select and implement treatment techniques with careful attention to cognitive, social, and other developmental issues. Assessment and treatment procedures for children, for example, require the use of less sophisticated verbal content and more pictorially oriented written materials (e.g., rating scales with faces or other pictures to describe mood instead of words). In many instances, children have not yet developed the cognitive skills to participate effectively in interventions that have a large cognitive component. Prior to mid-childhood, few children have meta-cognitive skills (i.e., the ability to think about thinking). Therefore, interventions with children often de-emphasize cognitive strategies in favor of more behavioral interventions Treating children also typically involves work with a broader treatment team, including parents, teachers, and other caretakers. It often requires consistent contact with school personnel who are often an integral part of the treatment program. Furthermore, many intervention programs for children are designed to be offered in schools, particularly those children whose parents might not have the resources to utilize traditional outpatient settings.

Behavioral treatments also can be of particular value for older adults. Again, however, implementation requires modifications, including a generally slower pace of skills training, simpler written documents prepared with larger fonts (e.g., treatment summaries and home practice forms), and more attention to medical issues. Conducting treatment with older adults who are cognitively impaired also requires the inclusion of collaterals or caregivers who can assist with the assessment and treatment process. Older adults with dementia, depression, and other behavioral problems, for example, can be treated effectively when caregivers are taught how to use BA skills with the identified patient. Case studies also have demonstrated that older adults with dementia and anxiety are able to learn coping skills such as relaxation, self-talk, and BA with the help of a therapist and a collateral or caregiver. For these patients, different learning strategies are used to teach skills, including increased repetition and practice, use of more cuing procedures to facilitate memory (e.g., calendars, written reminders in key locations of the home), and spaced retrieval, a procedure based in classical conditioning that repeatedly pairs questions (e.g., What can you do when you feel nervous?) and answers (e.g., I can breathe deeply) over increased intervals to facilitate memory.

Cultural Considerations. Only a few studies have addressed the role of race and ethnicity in behavioral treatments, and the majority have not

found differences in outcome based on these variables. Most differences identified typically relate to the ability to access services. Minority patients tend to seek mental health services less often and often receive fewer services. Recent behavioral treatment studies, however, have begun to develop and test variations in assessment and treatment procedures designed specifically for various minority and cultural groups, such as BA with Hispanic individuals. These modifications require careful attention to language and other cultural issues that may impact treatment delivery and outcomes. Simply translating questionnaires and treatment manuals into new languages is insufficient. Materials require back translation and careful consideration by members of the minority group as well as pilot testing to evaluate feasibility and acceptability.

Although there is limited research on how cultural differences may impact the practice of CBT, there is good reason to believe the impact is there. Research shows that individuals from different cultures tend to think about different things and make different interpretations using different reasoning processes. This obviously could have important implications for CBT with its focus on addressing patients' thoughts and thought processes. Cultural norms regarding what behavior is "appropriate" in different situations can have a major impact. Also, cultural norms (both the patient's and therapist's) impact interpersonal interactions and therefore impact the therapeutic relationship that is central to psychotherapy. Furthermore, when the therapist's worldview, experience, and expectations are different from the patient's, this can complicate the process of developing a shared understanding of the client's problems and possible solutions.

Implementation Variables. Behavioral treatments typically are of relatively brief duration, although the total length of treatment depends on the nature of the target problem(s) and the maintaining variables. Some types of specific phobia, for example, can be treated in as little as a single session. Treating more complex fears and phobias (e.g., obsessive compulsive disorder and social anxiety disorder) requires a longer series of sessions. The use of homework assignments in behavioral treatments in effect extends the actual duration of treatment without requiring as much therapist time or involvement. Sometimes, sequential phases of treatment are needed to address different problems, and these phases together may require a longer duration of treatment. For example, treating substance abuse typically requires an initial intensive phase of treatment to get substance use under control followed by a longer-term relapse prevention phase that may require less frequent sessions but is important for maintaining treatment gains and addressing other core difficulties (e.g., anxiety, depression, and interpersonal skills).

A variety of self-help treatments also have derived from behavioral approaches, and many interventions have begun to incorporate different

forms of technology to assist in the administration of behavioral treatments. *Bibliotherapy,* which offers treatment primarily via written materials, is useful for the treatment of panic disorder and depression. Internet-based treatments and smart-phone applications also have been used effectively in the treatment of various anxiety disorders (GAD, obsessive compulsive disorder, PTSD), depression, alcohol use, and body-focused repetitive behaviors (e.g., trichotillomania). These types of interventions extend the availability of behavioral treatments for people who are unable or unwilling to see a therapist.

Strengths of a Behavioral Approach. The primary strengths of behavior therapy are its specificity and attention to empiricism. First, behavioral treatment requires the specific identification of target behaviors to be changed, goals to be attained, and measures to evaluate outcomes. As such, accountability is high; it should be relatively clear to both therapist and patient (and insurance companies) when treatment is having the desired effects. Second, a behavioral approach is heavily anchored in a scientific method that requires knowledge of empirical literature and an empirical approach to the therapist's work with each patient. The scientific base for behavioral treatments is larger than for any other psychosocial approach, and behavioral treatments are actually the treatment of choice for many psychiatric disorders, including depression and anxiety. A behavior therapist also should be able to identify early in treatment when a case conceptualization is off the mark given the ongoing attention to objective assessment of outcomes. When treatment is having the desired effects, a new conceptualization and associated treatment plan can be developed. Third, behavior therapy is both efficient and cost-effective. Many treatment techniques described above have an immediate effect (e.g., when a sufficiently powerful reinforcement is used to change behavior, when intensive exposure is used to treat anxiety). In addition, much of the work of treatment occurs outside the therapist's office. Finally, behavioral approaches often match well with requirements of managed care, which require ongoing documentation of identified problems and outcomes.

Goals of Treatment

The goals of behavior therapy center on identifying and modifying target behaviors. Patients are required to take an active role in the identification of target problems and goals for change. They also must be willing to work on the problems outside of the therapy session. Although the relationship between therapist and patient is important in behavior therapy, as evidenced by studies that demonstrate a strong association between the therapeutic alliance and treatment outcomes, outside of FAP, this relationship is not considered the central mechanism of change. Rather, the working relationship between therapist and patient is a necessary, but not

sufficient, feature of therapy that sets the stage for change that occurs as patients take action. This action toward change is the primary goal of therapy. Outcomes are considered to be observable and measurable.

Limitations, Complications, and Contraindications. Behavior therapy is an action-oriented approach that requires a patient's readiness to change. Behavioral treatment, therefore, may not be appropriate for patients who are not ready to change or who prefer to discuss the historical origins of their problems (e.g., "I just don't understand why I am this way.") without wanting to take action. The extent of patients' readiness to change can be identified and measured, and treatment approaches such as *motivational interviewing* have been designed to help move patients toward readiness to change such that behavioral treatments can be effective. Motivational interviewing involves a nondirective evaluation of the advantages and disadvantages of change. In many cases, reinforcement for not making change is simply more powerful than reinforcement for making change (e.g., for smoking, drinking, and other behaviors followed by powerful and immediate reinforcement). However, patients who are consistently uninterested in making real changes are not optimal candidates for behavior therapy, or for that matter, virtually any form of therapy.

At times, although behavioral treatments generally match well with managed care reimbursement practices, some forms of behavior therapy may be at odds with insurance-based practice and reimbursement availability. Intensive exposure treatments, for example, that typically require 90-minute sessions (rather than the traditional 50-minute session) and more frequent meetings may not be reimbursed at a level that covers the therapist's time. Home visits, which often are an important part of behavioral treatments, require travel time and expenses not reimbursable by insurance. Finally, in many places behavior therapists with appropriate expertise are simply not available. Particularly in more rural geographical locations, there often are not a sufficient number of behavior therapists in practice. When specialized expertise is needed (e.g., ability to work with children, older adults, or particular psychiatric problems such as obsessive compulsive disorder or trichotillomania), the ability to locate a behavior therapist with appropriate expertise often poses a significant problem.

Think back to *Jenna,* whose symptoms and problems were described earlier in this chapter. After a thorough behavioral assessment, evaluation of baseline symptom severity, and development of a case conceptualization, the therapist and Jenna embarked on a course of behavioral treatment that included BA and PST.

Jenna was treated using an adapted PST protocol for cancer patients with depression. The basic premise of PST was that as she learned to identify and resolve her problems, and she would gain an increased sense of self-efficacy, control, and confidence while becoming more active in eliciting reward from her environment. According to problem-solving and behavior

theories, such changes would improve her mood and alleviate depression symptoms. The goals of PST were identical to those highlighted above (*Problem-Solving Therapy*). To better fit the needs of Jenna, her PST protocol included psychoeducation about cancer and its relation to emotional experiences and behavioral changes. An "acceptance versus change" model was incorporated throughout treatment. This model involved the rationale that some experiences in life are controllable and changeable whereas there are other experiences that cannot be readily changed and thus must be met with acceptance (e.g., being diagnosed with cancer or experiencing periodic negative emotional states). According to this rationale, to improve affect and emotional well-being, Jenna's effort should be applied to those aspects of her life more apt to be controlled and modified, namely overt behavior.

PST was carried out based on the seven stages described earlier. The first session involved establishing a comprehensive list of current problems in Jenna's life. During each subsequent session, Jenna and her therapist collaborated on solving a different problem of her choosing using steps two through seven (step six was homework and step seven was carried out at the beginning of the subsequent session). Jenna chose each problem from the problem list generated in session one. Gradually throughout treatment, she took a more prominent role in carrying out each of the problem-solving steps, requiring less input and guidance from the therapist as she better understood how to implement steps on her own. Jenna was primarily concerned with problems related to her relationship with her husband. The therapist and patient worked to narrow the focus of general relationship complaints to specific difficulties in communicating. Utilizing PST, the therapist and patient worked to help Jenna find proactive ways of expressing her thoughts, feelings, and concerns to her husband in a way that helped her feel understood and satisfied. Her role in communication problems with her husband was emphasized, as opposed to placing all the blame on her husband. The therapist and Jenna also worked on resolving problems related to parenting, including development of a structured disciplinary plan and formulating strategies to set limits with her children.

BA also was used as a component of treatment to address her behavioral and cognitive avoidance (e.g., thought suppression, rumination, watching television, and sleeping excessively). For example, in the context of her continued work with her physical therapist to increase mobility of her knee (with compliance monitored by her clinician), a life value assessment and activity hierarchy were structured that included behaviors systematically designed to get her back to work, increase health behaviors (e.g., hiking and exercising at her gym), and increase spiritual development (e.g., attending church, reading her bible, and praying). Exposure therapy in the form of journal exercises also was incorporated into BA to target Jenna's avoidance and help her confront the negative emotions involved with living with cancer. These assignments asked her to write a personalized narrative about her experience of being diagnosed, treated, and living with cancer, culminating with the completion of a written exercise designed to help a friend cope with being diagnosed with cancer. Through each of these journal assignments, Jenna was able to approach her thoughts and emotions about her health and her diagnosis. She reported these exercises were helpful in getting her to process her painful thoughts and feelings concerning her cancer diagnosis, and by writing about this experience, she was able to better accept her illness. These journal exercises also helped her realize the tendency to avoid difficult subjects and demonstrated the benefit of approaching, confronting, and processing such material as opposed to avoiding.

Jenna's parents also were asked to stop reinforcing her (passive avoidant) depressive behaviors by discontinuing completion of her chores and household responsibilities. Her therapist also worked directly with Jenna and her husband on communication skills to facilitate more positive, respectful and validating, as well as less critical conversations, while emphasizing the need for Jenna's husband to praise her and behave more pleasantly toward Jenna. At the termination of treatment, the Jenna reported that the treatment helped her to feel more confident and more in control of her life, as well as less anxious and depressed. She felt the problem-solving and BA strategies gave her a useful set of tools to approach and resolve the various sources of stress in her life and left her feeling less overwhelmed than she had been prior to treatment. After 3 months of treatment, Jenna's score on the HRSD had dropped to 6 (depression remission), 8 on the BDI-II (minimal depression), 12 on the PSWQ (minimal worry), and 3 on the AUDIT (minimal alcohol consumption). Jenna also reported more rewarding interactions with her husband, a strong social network at her church, and that she had recently begun a part-time job.

ETHICAL ISSUES

Behavior therapy shares many of the same ethical issues that all mental health providers face, namely, providing adequate and proper care, practicing within appropriate boundaries, maintaining patient rights and confidentiality, and doing no harm. The scientific approach that serves as the basis for behavior therapy, however, minimizes some of these issues. Specifying targets for change and evaluating outcomes in an objective fashion helps ensure that treatment is appropriately administered (i.e., with regard to content and duration). Utilizing empirically supported treatment procedures also helps to ensure adequate and appropriate care. Much of the behavioral approach also involves collaboration between the patient and therapist. Treatment goals are established collaboratively, and patients are active in treatment through participation in home practice assignments. If at any point it becomes apparent that patients are not moving forward toward change, the therapist can work with the patient to modify goals and strategies.

In the early days of behavior therapy's development, there was significant concern about the use of behavioral principles to control people. The term *behavior modification* itself posed significant concerns in this realm. And there remain today specific issues that behavior therapists must consider with regard to the use of certain behavioral techniques. A number of ethical issues, for example, need to be considered in the use of aversion-based techniques. As previously noted, these procedures are used only as a last resort for very serious symptoms (particularly those that cause harm to the patient or another person) that have not responded to alternative approaches. Typically, aversion-based treatments are reviewed by a full team of providers, including physicians who can evaluate any potential medical complication or risk, and are conducted only with informed consent from the patient or an appropriate legal guardian. In many cases, ethical and legal review boards also have input. And perhaps most importantly, these procedures should be used in conjunction with alternative techniques that promote other more adaptive behaviors.

There also are ethical considerations in the use of exposure treatment. If exposure sessions are not conducted appropriately, they can create increased anxiety and cause harm. Such adverse effects rarely occur, but a behavior therapist needs to be observant of signs that exposure may not be an appropriate approach (e.g., the patient has a medical condition that contraindicates increased arousal; the patient cannot tolerate increased distress) or is not producing extinction as it is being conducted (e.g., within- and between-session reductions in arousal are not occurring). The choice to use exposure should be based on the scientific literature, a thorough case conceptualization, and a collaborative decision between therapist and patient following appropriate education of the patient and

evaluation of the cost–benefit ratio.

The ethics of utilizing a nonscientific treatment approach needs to be considered as well. In some cases, there simply are no appropriate empirically based treatments available. In these instances, the behavior therapist must rely on the use of behavioral principles that fit the behaviorally based case conceptualization. There are, however, some frequently used procedures that mimic some of the strategies of behavior therapy but have not received adequate empirical validation. Thought Field Therapy (TFT), which involves "tapping" one's fingers at certain "meridian points" on the upper body and hands, is purported to eliminate fear, anxiety depression, and other negative emotional states. However, there are few controlled trials of TFT that would meet accepted scientific standards. Eye movement desensitization and reprocessing (EMDR) combines traditional exposure therapy with repetitive eye movements, and several studies suggest it is superior to no treatment or may be equally effective to traditional behavior therapy. However, the current database is fraught with methodological difficulties, and the incremental validity of EMDR over traditional exposure therapy alone is not well established.

As is the case for all health care providers, it is essential that behavior therapists are well trained to administer treatments. If a therapist decides to use a treatment for which he or she has not had sufficient training, it is the therapist's responsibility to obtain consultation and supervision as necessary to conduct the treatment appropriately. Particularly in the case of procedures such as punishment and aversion therapy (e.g., electrical shock), careful consideration of all possible alternatives is an absolute necessity.

RESEARCH AND EVALUATION

As previously stated, research is central to behavior therapy. Although most behavior therapists do not actually conduct research, they take an empirical perspective in their clinical work, using the scientific literature to design treatment plans and evaluate patient outcomes over the course of therapy. For behavior therapists who do conduct research, there are many types of experimental designs available to examine the efficacy of behavioral treatments. Many of these are the same as those used in other types of outcome research, but a brief overview of methods is presented here. As a result of the strong research tradition in behavior therapy, there now exists a wide array of empirically validated behavioral treatments to address a range of psychiatric disorders. The definitions of empirically validated treatments and an overview of treatments that have attained this standard also will be presented.

Methods of Research and Evaluation

The simplest form of outcome evaluation is the *case study,* a detailed description of a patient (or small group of patients) and the course of treatment, typically accompanied by quantitative measures of outcome. Although case studies do not include experimental controls that allow conclusions about the cause of patient improvement over time, a series of case studies is often taken as initial evidence that a particular type of treatment may be efficacious and worthy of future study. Case studies, therefore, are useful for establishing hypotheses about the potential value of certain therapeutic approaches that can be tested in subsequent empirical studies. Case studies also are useful for studying rare clinical phenomenon for which group-based research would be nearly impossible, and they are an excellent way for clinical practitioners to conduct research and share findings with others.

Single-case experimental designs are another research and evaluation tool that can be used to test the efficacy of treatments offered to an individual patient. As opposed to the case study, however, these designs do incorporate sufficient experimental controls to allow one to make causal inferences about the impact of treatment on patient outcomes. In a single-case design, treatment with an individual patient is a complete experiment, as the patient alternately participates in both the intervention and a control (comparison) condition. The goal of the experiment is to evaluate whether the patient's behavior changes systematically as a result of participation in the intervention and comparison conditions.

In a single-case experiment, the patient first participates in a baseline assessment phase during which the frequency of the target behavior is recorded (e.g., how many tantrums a child has). Given that simple monitoring of the target behavior may actually change the behavior (see Self-Monitoring earlier in this chapter), the initial assessment phase needs to be continued long enough for the frequency of the target behavior to achieve stability. The assessment of the target behavior is then continued further as the intervention and any comparison conditions are applied and removed. The most common design for a single-case experiment is the *ABAB* or *reversal design,* where A is the baseline phase and B is the intervention. Conducting an ABAB experiment involves a baseline assessment (A) followed by application of the intervention (B), removal of the intervention with continued assessment (A), and reapplication of the intervention (B). If each "B" phase has the same effect on the target behavior (e.g., each time a token economy is instituted and the child has fewer tantrums), but the behavior returns to baseline levels when the intervention is removed (e.g., the child resumes a high level of tantrums when the token economy system is removed), one can conclude the intervention had a causal impact on the behavior (Fig. 33.3–3).

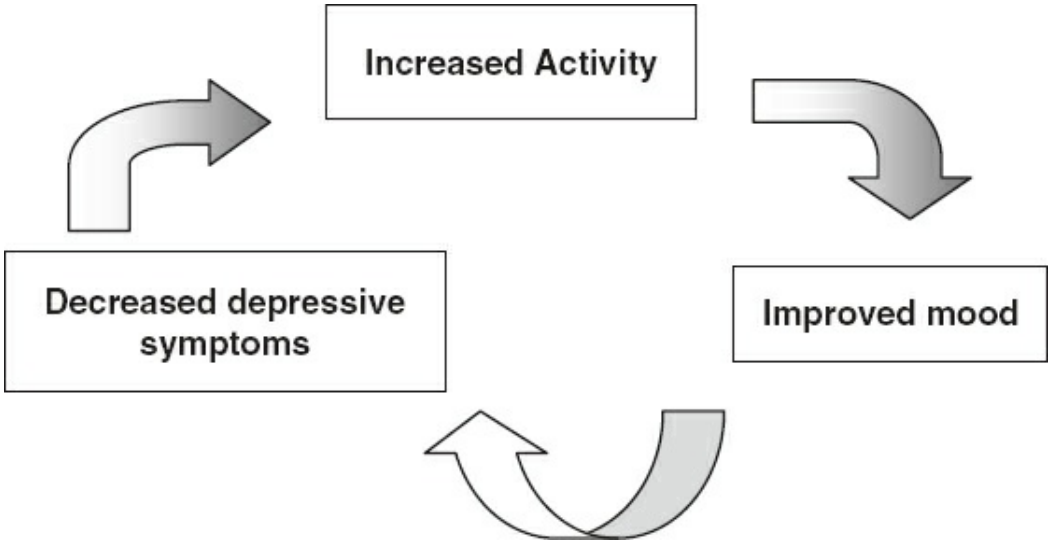

FIGURE 33.3–3. Sample of ABAB design.

In many cases, a single-case experiment is unethical or impractical. For example, removing a treatment that reduces self-injurious behavior for a child with developmental disabilities is unethical. In other cases, a treatment once applied cannot be removed. For example, after a patient is taught problem-solving skills, it is impossible to "remove" that treatment and reassess behavior. In these cases, a *multiple baseline design* may be used. In this design, only one A-B sequence is applied, but the sequence is repeated across individuals, settings, or behaviors. To repeat the design across individuals, all patients begin the assessment phase at the same time, but each patient begins the treatment at a different time point. For example, four patients with trichotillomania might begin monitoring of hair pulling at the same time and then receive HRT. However, patient 1 gets HRT after 2 weeks of baseline monitoring, patient 2 after 3 weeks of monitoring, patient 3 after 4 weeks of monitoring, and patient 4 after 5 weeks of monitoring. If hair pulling frequency changes for each patient at the time of intervention (B), one can conclude that intervention likely had a causal effect.

Single-case designs allow clinicians and researchers to use experimental strategies to determine whether treatments are effective with individual patients or small numbers of patients. They do not, however, allow one to generalize the results to large groups of patients. To draw more general conclusions, a *randomized clinical trial* is needed. The methodology of a randomized-controlled trial (RCT) is generally no different in behavior therapy research than in other types of treatment research (e.g., psychopharmacology), although some unique issues arise with regard to the nature of the control group. In an RCT, a large number of patients are assigned randomly to treatment conditions and assessed at regular intervals to determine whether patients who receive certain treatments improve more than patients in alternative treatment or control groups. The

nature of the control group is an important consideration in behavior therapy research (as it is in all psychosocial treatment research) and significantly influences the conclusions that can be drawn. In some instances, patients in the control condition receive no treatment, allowing one to conclude only how well a particular treatment works relative to nothing. It is also important to examine how well a treatment works relative to other treatments that control for what are called *nonspecific effects,* or those factors of treatment that are not specific to behavior therapy, such as amount of time with the therapist, strength of the therapeutic relationship, and patient expectations for improvement. Various attention-based control conditions (e.g., supportive psychotherapy and education–discussion groups) are often used to examine whether treatments have an impact on behavior over and above nonspecific effects.

Once a treatment has been documented as efficacious in a controlled experiment, it is then important to determine whether the treatment also has an impact when delivered in the context of routine clinical care. In this domain, the difference between *efficacy research* and *effectiveness research* is important to consider. In efficacy studies, the researcher attempts to maximize *internal validity* of the study by controlling all possible extraneous variables so that a firm conclusion can be drawn about the impact of the treatment (e.g., patients and procedures in the treatment and control groups are exactly the same with the exception only of the specific treatment procedures administered). In effectiveness studies, there is an attempt to maximize *external validity* by delivering the treatment to patients as it would occur in the "real world." These types of studies allow one to generalize findings more broadly to what happens in actual clinical practice, but they do not incorporate the same level of control over extraneous variables. For example, treatment may be conducted in a variety of settings (e.g., primary care and community-based clinics), patients may be more diverse (and therefore more representative of the population of interest), providers may be less expertly trained, and the control condition is often *usual care,* wherein patients continue to receive whatever care they would ordinarily receive in the treatment setting, making it more difficult to draw conclusions about the specific effects of the treatment under investigation. In some cases, studies incorporate elements of both efficacy and effectiveness research to balance the benefits of internal and external validity.

Empirically Validated Behavioral Treatments

The philosophy of behavior therapy and the scientist practitioner model align closely with the recent evidence-based medicine movement, which proposes that the quality of patient care is enhanced with the use of empirically supported treatments. In 1993, in fact, Division 12 (Clinical Psychology) of the American Psychological Association (APA) established a

task force to identify and promote the use of empirically supported treatments. The goals of this work were to (1) define what is meant by empirically supported treatment, (2) provide examples of interventions that met these criteria, and (3) explore the notion of disseminating this information to the professional and lay communities. The first report from this task force published in 1995 included definitions of what were called *well-established treatments, probably efficacious treatments,* and *experimental treatments*.

Well-established treatments were defined as those for which at least two RCTs or a large series of single-case design experiments demonstrated efficacy relative to a pill placebo, psychotherapy placebo, or an alternative intervention. Experiments considered supportive of these criteria had to be conducted with clearly specified patients using some type of formal treatment manual. Studies supporting efficacy also had to be produced by at least two different teams of investigators. Probably efficacious treatments were defined as those for which efficacy was demonstrated only relative to a wait-list control condition (i.e., where no treatment was provided) or where all supporting controlled trials came from the same set of investigators. Experimental treatments were those that failed to meet criteria for either of these categories.

In the 1995 report, the task force reported 25 treatments that met criteria for one of the two categories. The majority of these were behavioral (or cognitive behavioral) interventions that focused primarily on the treatment of adults. Since this time, the list has been expanded significantly, and other task forces were appointed to focus on empirically supported interventions for children and older adults. Similarly, many other professional groups (e.g., Society of Behavioral Medicine; APA Divisions 17 [Counseling Psychology] and 29 [Psychotherapy]) and individual researchers have taken on similar tasks of identifying empirically supported treatments. There is much consistency in the interventions categorized as efficacious, including many treatments (largely behavioral and cognitive) for anxiety and stress, depression, substance abuse, health problems, and marital discord.

Training programs in psychiatry and psychology are now required by credentialing bodies to teach empirically supported interventions, although surveys continue to document that many programs still do not follow these guidelines. There also have been many critiques of the production of these kinds of lists. One primary critique has to do with the concern that such reports will be misused by third-party payers and limit professional practice. Other critics have argued that the reports focus too much on brief, manualized forms therapy conducted with homogenous patient groups that fail to represent the types of patients seen in clinical practice. As stated earlier in this chapter, however, proponents of empirically supported procedures recognize that these treatments provide guidelines for

individualized patient care; they are not intended to serve as rigid rules for how to treat patients. Moreover, new lines of treatment research are beginning to broaden to an effectiveness perspective wherein interventions are tested in more real-world settings with more heterogeneous groups of patients. Other critiques of the task force report (and other reports that are similar) also are concerned that most of the treatments listed derive from behavioral and/or cognitive perspectives. However, behavior therapy is much more strongly aligned with the use of empirical studies to document treatment effects than other perspectives, and thus most experimental trials used to classify treatments as efficacious involve behavioral interventions.

In summary, although less than 100 years old, behavioral interventions now represent one of the most established and accepted interventions for individuals with a variety of disorders. Behavioral treatments are effective with children, adolescents, adults and mature adults, and they can be implemented in traditional inpatient and outpatient settings. However, they are commonly practiced in educational institutions and the community at large. Efficacious not only for psychological disorders, they also are now part of the mainstream for many medical disorders, either as a primary intervention (as in the case of hypertension or obesity,) or as a secondary intervention to assist in the management of chronic diseases, such as diabetes, asthma, or cancer. Finally, unlike some other schools of therapy, the success of behavior therapy is not limited by intellect or the need for verbal communication, having been successfully used for those with various developmental disabilities and even those who suffer from multiple difficulties such as blindness and deafness. Clearly, behavioral interventions are broad-spectrum treatments with substantial empirical support.

References

Azrin NH, Nunn RG. *Habit Control in a Day*. New York: Simon & Schuster; 1977.

Baer RA. Mindfulness training as a clinical intervention: a conceptual and empirical review. *Clin Psychol Sci Pract*. 2003;10:125–143.

Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory*. New York: Prentice-Hall; 1986.

Barlow DH. *Clinical Handbook of Psychological Disorders*. 5th ed. New York: Guilford; 2014.

Basco MR, Rush AJ. *Cognitive-Behavioral Therapy for Bipolar Disorder*. 2nd ed. New York, NY: Guilford Press; 2007.

Bernstein D, Borkovec TO. *Progressive Relaxation Training*. Champaign, IL: Research Press; 1979.

Chambless DL, Ollendick TH. Empirically supported psychological interventions: controversies and evidence. *Ann Rev Psychol*. 2004;52:685–716.

Dimidjian S, Hollon SD, Dobson KS, et al. Randomized trial of behavioral activation, cognitive

therapy, and antidepressant medication in the acute treatment of adults with major depression. *J Consult Clin Psychol*. 2006;74:658–670.

Eysenck HJ. The effects of psychotherapy: an evaluation. *J Consult Psychol*. 1952;16:319–324.

Hayes SC, Strosahl KD, Wilson KG. *An Experiential Approach to Behavior Change*. New York: Guilford; 1999.

Hofmann SG, Sawyer AT, Fang A. The empirical status of the "New Wave" of CBT. *Psychiatr Clin North Am*. 2010;33(3):701–710.

Hollon SD, Ponniah K. A review of empirically supported psychological therapies for mood disorders in adults. *Depress Anxiety*. 2010;27(10):891–932.

Hopko DR, Lejuez CW, Ruggiero KJ, Eifert GH. Contemporary behavioral activation treatments for depression: procedures, principles, progress. *Clin Psychol Rev*. 2003;23:699–717.

Jones MC. The elimination of children's fears. *J Exp Psychol*. 1924;7:383–390.

Kabat-Zinn J. An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: theoretical considerations and preliminary results. *Gen Hosp Psychiatry*. 1982;14:33–47.

Kazdin AE. *Behavior Modification Applied Settings*. 6th ed. Belmont, CA: Wadsworth/Thomson Learning; 2001.

Kohlenberg RJ, Tsai M. *Functional Analytic Psychotherapy*. New York, NY: Springer US; 1991:169–188.

Lejuez CW, Hopko DR, Hopko SD. A brief behavioral activation treatment for depression. *Behav Modifi*. 2001;25:255–286.

Lewinsohn PM. A behavioral approach to depression. In: Friedman RM, Katz MM, eds. *The Psychology of Depression: Contemporary Theory and Research*. Oxford, United Kingdom: Wiley; 1974:157–178.

Linehan MM. *Cognitive-Behavioral Treatment of Borderline Personality Disorder*. New York, NY: Guilford Press; 1993.

McCullough JP. *Treatment for chronic depression: Cognitive Behavioral Analysis System of Psychotherapy (CBASP)*. New York: Guilford; 2000.

Mowrer OH, Mowrer WM. Enuresis—a method of its study and treatment. *Am J Orthopsychiatry*. 1938;8:436–459.

Nezu AM, Nezu CM, Perri MG. *Problem-Solving Therapy for Depression: Theory, Research, and Clinical Guidelines*. New York: Wiley; 1989.

Persons JB. Empiricism, mechanism, and the practice of cognitive-behavior therapy. *Behav Therapy*. 2005;36:107–118.

Sattler JM, Hoge RD. *Assessment of Children: Behavioral, Social & Clinical Foundations*. San Diego, CA: Jerome Sattler Publications; 2006.

Spiegler MD, Guevremont DC. *Contemporary Behavior Therapy*. 4th ed. Belmont, CA: Wadsworth/Thomson Learning; 2003.

Summerfeldt LJ, Antony MM. Structured and semistructured diagnostic interviews. In: Antony MM, Barlow DH, eds. *Handbook of Assessment and Treatment Planning for Psychological Disorders*. New York: The Guilford Press; 2002.

Teasdale JD, Segal Z, Williams JM. How does cognitive therapy prevent depressive relapse and why should attentional control (mindfulness) training help? *Behav Res Therapy*. 1995;33:25–

39.
Watson JB, Rayner R. Conditioned emotional reactions. *J Exp Psychol*. 1920;3:1–14.
Wolpe J. *Psychotherapy by Reciprocal Inhibition*. Stanford: Stanford University Press; 1958.

▲ 33.4 Hypnosis

ALLAN DAVID AXELRAD, M.D., DANIEL P. BROWN, PH.D., AND HAROLD J. WAIN, PH.D.

Why and what is it about hypnosis that for centuries it has procured myths and the imagination of not only the medical community but the lay public as well? Is it a treatment, a state, a trait? Where and how can it be used effectively? Myths and misconceptions are abundant. There are a myriad of perceptions among clinicians and the lay public. Even the term hypnosis can be misleading, coming as it does from the Greek root hypnos (meaning "sleep"). Hypnosis is not sleep. It is more likely a complex process that requires alert focused and receptive attention. Many also accept the myth that the clinician projects the hypnotic trance onto the patient or has the power to influence the patient. In reality it is the patient who has the hypnotic gift, and the clinician's role is to assess the patient's capacity to capitalize on their assets and to help them to discover and use them effectively. Patient motivation, personality style, and biological predisposition may contribute to the manifestation of this talent.

The lay public and many clinicians have frequently witnessed hypnosis as a form of entertainment. Though this has kept the intrigue of hypnosis alive, it has also contributed to a possible lack of appreciation or a diminished impact of its possible clinical utility. However, for the trained clinician hypnosis can be a remarkably effective tool in helping with pain, surgery, stress habits, anxiety, somatoform, dissociative disorders as well as many medical issues and unexplained symptoms. For experienced clinicians hypnosis is not always a treatment of a syndrome but is a treatment of the individual patient as long as the clinician is comfortable and experienced using the intervention. With its effect on a myriad of medical and psychiatric conditions that range from PTSD to surgery, hypnosis can be facilitative for researchers to track both biological and psychological variables. The milieu of hypnosis can also be an exceptional research milieu to explore mind–body interactions.

HISTORY

Early Chinese and Egyptian writings, as well as discussions in the Hebrew Bible and New Testament, seem to refer to hypnotic phenomena, trance states, and spontaneous dissociation being used for medical intervention. Franz Anton Mesmer was perhaps the first to describe the significance of this phenomenon. He postulated that magnetic energy or an invisible fluid could be provided from a clinician or an object to correct imbalances and to restore health to an individual with illness. Mesmer's style was dramatic and included physical passes over the body of the patient by using his own

magnetic field in an attempt to restore equilibrium to the patient's "fluid." Many times his patients appeared to have experienced nonepileptiform seizures on their road to recovery. His theories were highly controversial and, although anecdotally successful, attracted negative attention. In 1784, the French government convened a panel of experts that included Benjamin Franklin, the chemist Antoine Laurent Lavoisier, and Joseph-Ignace Guillotin to meet in Paris. They concluded that the phenomenon was nothing but "heated imagination." As a result Mesmer's influence in France declined, although history credits him with beginning contemporary medical hypnosis.

In the 1840s, James Braid, an English physician and surgeon, described phenomena similar to those that Mesmer had noted. He reported that he could produce trance states using eye fixation and eye closure. In 1847, Braid rejected magnetic theory and created a psychological concept that he called monoideism, by which he meant mental concentration on a single dominant idea. In this state, subjects were highly suggestible and could focus their attention on specific ideas that would influence behavior.

In 1846, James Esdaile, a Scottish physician serving in India, described in his book *Mesmerism in India, and Its Practical Application in Surgery and Medicine,* the use of such techniques as the only anesthetic in more than 300 major surgical procedures. During the mid-19th century, John Elliotson, a noted professor of medicine and an editor of *Lancet,* became a strong supporter of hypnosis. Because hypnosis was still rejected by many physicians at the time, Elliotson fell out of favor with his colleagues and was forced to resign from the journal. Later many others became impressed with the effect and results of hypnosis and began studying the technique.

By 1880, the noted French neurologist Jean-Martin Charcot at the Salpêtrière in Paris considered the hypnotic state to be a neurophysiological phenomenon and a sign of mental illness. Some prominent physicians, such as Pierre Janet, supported this position about hypnosis, but others, such as Hippolyte Bernheim, believed that hypnosis was a function of the normal brain. Bernheim emphasized the role of suggestibility in hypnosis over pathological theories. Sigmund Freud studied hypnosis with Charcot, and it was central to his classic work on hysteria with Joseph Breuer. The latter facilitated a shift in hypnotic treatments by recognizing the value of abreaction under hypnosis. Of interest, Freud noted in his autobiography his setting aside the use of formal hypnosis and focusing on the concept of transference. When a patient threw her arms around his neck during a hypnotic trance he related that "I was modest enough not to attribute the event to my own irresistible personal attractions, and I felt that I had now grasped the nature of the mysterious element that was at work behind hypnotism." To concentrate on the elements of transference, he had to exclude extraneous variables, and it was therefore necessary to abandon hypnotism. Along with the rise

of behaviorism, Freud's rejection of hypnosis contributed to its decline in the early 20th century. However, later Freud spoke of the need to blend "the pure gold of analysis plentifully with the copper of direct suggestion."

After Freud discarded hypnosis, its influence continued to wax and wane. During World War I, Ernst Simmel, a German psychoanalyst, became interested in hypnosis for the treatment of "war neurosis." He developed a technique for accessing repressed material, which he called hypnoanalysis. Other psychiatric pioneers, such as Sandor Ferenczi, Paul Schilder, Lawrence S. Kubie, Sydney Margolin, Merton Gill, Margaret Brennan, John Watkins, and Erika Fromm, contributed to the early development of a psychodynamic approach to clinical hypnosis. During World War II, hypnosis played a prominent role in the treatment of pain, combat fatigue, and neurosis through the work of John Watkins and John Spiegel. Some of the first research on memory recall and the control of physiological activity with hypnosis were carried out at the U.S. Army School of Military Psychiatry.

Despite Freud's earlier perceived rejection of hypnosis, many researchers and clinicians helped to shape the modern period of scientific hypnosis. For example, Ivan Pavlov focused on a biological perspective in the understanding of hypnosis. World War I brought more interest in treating "battle fatigue" and "shell shock." Other researchers and clinicians, such as Clark Hull, Ernest and Josephine Hilgard, Andre Weitzenhoffer, Milton Erickson, Martin Orne, and Herbert and David Spiegel, all contributed to establishing the contemporary place of hypnosis. In 1955, the British Medical Society formally recognized hypnosis and recommended that it be taught in medical schools. In 1958, the American Medical Association (AMA) and the American Psychiatric Association officially recognized it as a safe and effective treatment. Three specific national societies were soon established, including the Society for Clinical and Experimental Hypnosis, the Society of Psychological Hypnosis (American Psychological Association Division 30), and the American Society of Clinical Hypnosis.

THE TRAIT OF HYPNOTIZABILITY

When Clark Hull first studied responses to direct ideomotor suggestions in the laboratory he found that responsiveness was not evenly distributed across subjects into a more suggestible and a more resistant group. In the 1950s and 1960s, Ernest R. Hilgard and Andre Weitzenhoffer developed standardized instruments to measure hypnotic responding. The Stanford Hypnotic Susceptibility Scale (SHSS) consists of a formal hypnotic induction ceremony, followed by 12 specific suggestions to alter ideomotor response or cognitive-perceptual experience, or posthypnotic behavior. Administration of these scales to large numbers of subjects led to the view that hypnotizability is bimodally distributed, with 10 to 20 percent of the

general population being highly or very highly hypnotizable and the remainder of the population being mildly or moderately hypnotizable. Figure 33.4–1 illustrates the bimodal distribution of hypnotizability.

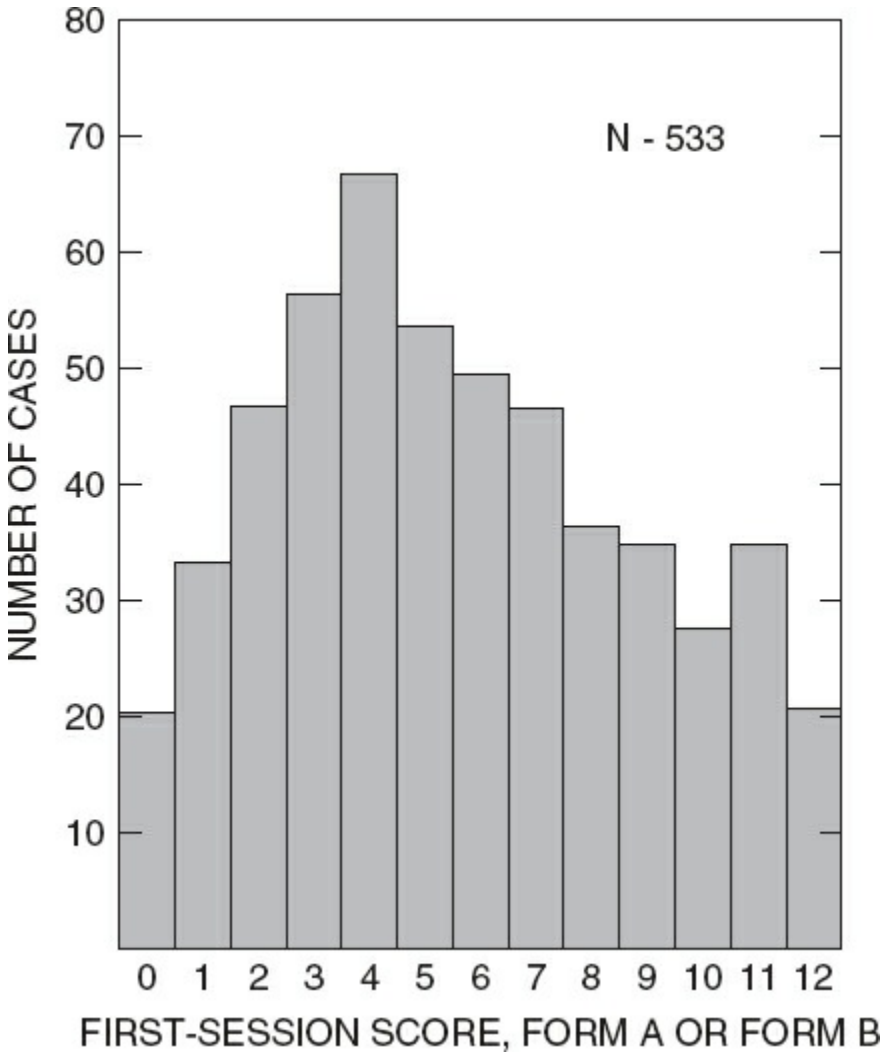

FIGURE 33.4–1. Bimodal distribution of hypnotizability resulting from administration of the Stanford Hypnotic Susceptibility Scale to a large number of subjects, reflecting that 10 to 20 percent of subjects are highly or very highly hypnotizable and the remainder of subjects are mildly or moderately hypnotizable.

Use of a variety of standardized instruments to assess hypnotizability over the last half century in large samples of subjects from different geographic areas representing various cultural and ethnic groups has consistently demonstrated individual differences in hypnotizability.

The convention has been to divide hypnotic subjects into three groups:

- High hypnotizables represent 8 to 10 percent of the general population. These are the virtuosos of hypnotic talent.
- Moderate hypnotizables make up the bulk of hypnotic subjects—from 84 to 88 percent.
- About 4 to 6 percent of the general population are low hypnotizable.

A comparison of results obtained using different instruments to measure hypnotizability shows that the cut-off points distinguishing high, medium, and low hypnotizability do not always exactly correspond.

Accurate assessment of hypnotizability in children depends on adapting the hypnotic technique to the level of cognitive development of the child. Preschool children are less hypnotizable than school-age children. Nonetheless, their degree of hypnotic responding increases with brief, open-eye hallucinatory pretend play suggestions matched to the preoperational thinking and short attention span of that age group. Early-school-age children respond best to action-oriented fantasy suggestions of longer duration, matched to the late preoperational thinking of that age group. Older school-age children (ages 9 to 12 years) respond well to closed-eye detailed fantasy suggestions and ideomotor suggestions, as well as to a wide range of suggestions for cognitive and perceptual changes, corresponding to the development of concrete operational thinking, along with a rich internal world of fantasy and affective states and stable attentional skills characteristic of that age group. Longitudinal studies of changes in hypnotizability in the same children over the course of child development consistently have shown that, irrespective of the issue of matching the style of hypnotic suggestion to the developmental age of the child, there is a progressive increase in hypnotizability from younger to older childhood, with the degree of hypnotizability reaching a peak between 9 and 12 years of age, an age roughly corresponding to the development of concrete operational thinking. Because most children from 9 to 12 years old are highly hypnotizable, this age range has been called the of peak hypnotizability. For most individuals, the degree of hypnotizability drops off by mid-adolescence, so that the majority of older adolescents and adults fall within the medium range of hypnotizability.

Test–retest correlations (0.7 to 0.8) of standardized hypnotizability scales and/or repeated experiences associated with interpersonal trust, such as psychotherapy or participation in growth groups, can lead to very modest, stable increases in hypnotizability over time, at least in some individuals. Drastic alterations in consciousness, for example, induced by hallucinogenic drugs or prolonged sensory deprivation, lead to significant but temporary changes in hypnotizability, after which the degree of hypnotizability returns to baseline for most individuals. There is little evidence that any procedure leads to significant stable changes in the degree of hypnotizability for most adults. In other words, low- and medium-hypnotizable subjects are likely to remain within that range of hypnotizability and are unlikely to become hypnotic virtuosos under any circumstances.

The trait of hypnotizability is distributed across a wide range of personalities, and as research with a wide range of personality tests has indicated, there is no simple stereotype of the type of individual who is

highly hypnotizable. With respect to the correlation of hypnotizability to single personality traits, measures of attachment style, attentional skills, absorption, dissociation, fantasy-proneness, and active imagination show a mild to modest, yet significant correlation with hypnotizability.

The correlation between hypnotizability and psychiatric diagnosis has yielded more promising results. Comparing patient groups with independently validated psychiatric diagnoses on standardized measures of hypnotizability has shown that as a group (1) patients with anxiety disorders (generalized anxiety, panic, and OCDs), somatoform disorders (e.g., conversion), and irritable bowel syndrome (IBS) are slightly but significantly more hypnotizable; (2) patients with PTSD are moderately more hypnotizable; and (3) patients with dissociative identity disorder are much more hypnotizable than the general population. These differences in the hypnotizability of various psychiatric groups have led to the belief that hypnosis may be a preferred treatment at least for certain psychiatric conditions. Again, this is not to suggest that it is only the hypnotic intervention that accounts for the change; it is the strategy one utilizes with the milieu of hypnosis that accounts for the positive outcome.

The relationship between degree of hypnotizability and clinical efficacy has not been clearly established. Some clinicians select only high hypnotizables for hypnotherapy. However, the majority of clinicians who use hypnotic procedures offer hypnosis most often to medium and high hypnotizables and some of the low hypnotizables under the assumption that anywhere from a low-to-moderate to higher level of hypnotizability is sufficient for most clinical work. Some clinicians believe that establishing certain expectancies, as well as a therapeutic alliance, facilitates therapeutic change even in patients with low hypnotizability. The prevailing view is that the clinical applications of hypnosis are not restricted to highly hypnotizable individuals. Some clinicians suggest that intervention strategies can be based on levels of hypnotizability.

Research consistently supports the view that hypnotizability predicts clinical outcome for a variety of DSM-5 diagnoses, somatic symptoms, and related disorders and/or DSM-IV Axis I psychiatric conditions and not others. Hypnotizability as a personality trait, irrespective of whether the clinical procedure is conducted in a waking or hypnotic state, predicts clinical outcome for a variety of psychiatric conditions and medical conditions, such as in the treatment of anxiety disorders, conversion disorders, posttraumatic or dissociative disorders, bulimia, pain, headache, asthma, certain skin conditions, and functional gastrointestinal disorders, but does not predict outcome for a variety of addictive behaviors, such as smoking, weight control, and alcohol/drug abuse.

The implication of this research on the distribution of hypnotizability is that hypnotic responding, at least for high hypnotizables, is a personality trait or talent. Such a view should dispel the stereotype that hypnosis

somehow exemplifies the special power of the clinician to exert undue influence. Appreciating hypnosis as a talent or gift within the patient naturally leads to a more egalitarian, less hierarchical or power-oriented approach to hypnotherapy. In permissive hypnosis, the hypnotherapist provides the relational conditions to facilitate the patient's exploration of his or her hypnotic abilities and inner resources relevant to therapeutic change.

HYPNOSIS AND SUGGESTIBILITY

Clark Hull systematically studied the effects of direct suggestions on motor response and believed that responsiveness to ideomotor suggestion constituted the primary factor of overall hypnotic responding. Andre Weitzenhoffer studied the effects of suggestion as a form of external social influence through verbal communication. Weitzenhoffer noted that some individuals significantly increased their suggestibility after a hypnotic induction ceremony was used to induce a trance state. This led to the view that trance induction caused a *condition of heightened suggestibility,* so that the individual immediately following trance induction was significantly more responsive to subsequent suggestions than to suggestions given outside of the context of trance induction. Weitzenhoffer coined the term "classic suggestion effect" to illustrate how high-hypnotizable subjects following the induction of trance accepted subsequent suggestions involuntarily and with uncritical acceptance and deep conviction, even when the suggestions violated conventional logic or normal reality testing. The *heightened suggestibility model* of hypnosis has had an important influence on hypnosis research since the last half of the 20th century, in that most of the standardized instruments to assess hypnotizability are constructed around this assumption, namely each instrument begins with a formal hypnotic induction ceremony, followed by a series of specific suggestions given in a standard format. Overall hypnotizability in these scales is assessed as a function of how many of specific suggestions are accepted by the subject following the induction ceremony. In a broader view, heightened suggestibility is best seen as one important dimension within the overall domain of hypnotic responsiveness, with a variety of state effects occurring at least partially independent from suggestion effects in certain hypnotic subjects. The state model of hypnosis assumes that a formal induction ceremony results both in a hypnotic state and an increase in responsiveness to suggestion. Subsequent suggestions serve to shape the nature of hypnotic experiences and behaviors.

THE DOMAIN OF HYPNOTIC EXPERIENCES AND BEHAVIOR

Ernest Hilgard's research on hypnotizability defined six main areas of hypnotic responding: ideomotor response, hypnotic imagery and dreaming,

effects on memory, sensory and perceptual effects, cognitive-affective changes, and posthypnotic effects.

- Ideomotor responsiveness refers both to the facilitation and the inhibition of motor response through specific suggestions, for example, to facilitate eye closure, postural sway, movement of outstretched hands together as if pulled by a magnet, and hand/arm levitation and lowering, and/or suggestions, for example, to inhibit the bending or lifting of an arm, the unlocking of clasped fingers, the opening of closed eyelids, or the verbalization of certain words. Ideomotor responsiveness accounts for the largest portion of the variance of overall hypnotic responding.
- Some hypnotized subjects experience a richer quality and greater quantity of imagery in hypnosis. Imagery as compared to cognitive information processing may play a greater role in hypnotic over waking experiences.
- The effects of hypnosis on memory are complex. Hypnotic suggestions may increase the recall (hypnotic hypermnesia) but not the recognition of personally meaningful material (but not the recall of semantic or procedure memory), but the accuracy of recalled information varies according to whether the hypnotist created realistic or unrealistic expectancies about memory retrieval, did or did not encourage guessing, and/or did or did not suggest misinformation about the recall target.
- Hypnotic analgesia is a well-established phenomenon, especially among high-hypnotizable subjects. A relatively small group of high-hypnotizable individuals are able to produce positive or negative hallucinations in response to specific hypnotic suggestions. An hypnotically suggested heat hallucination not only results in an intense subjective experience of heat that is virtually indistinguishable from actual heat in certain high hypnotizables, but may also result in burn blisters. Hypnotic suggestions can also inhibit the perception of actual experiences (negative hallucination), such as failing to see an actual light or color, failing to see a hand on a watch, or selectively becoming deaf to certain words or sounds.
- Hypnotic suggestions can also affect changes in meaning systems, affective experience, and self-representation. For example, hypnotic suggestions can produce temporary agnosias and aphasias. Direct suggestions to experience a discrete emotional state, for example fear, not only result in the subjective experience of fear but also an increase in catecholamine and cortisol output. Hypnotic suggestions can directly alter self-representation, such as the suggestion to imagine oneself completely free of pain or the suggestion to no longer identify oneself as a smoker. It can also be used to create somatic symptoms as a defense or bridge while looking for the metaphorical meaning of the presenting symptom.

- Certain suggestions given during a hypnotic state are designed to have an effect after formal termination of the hypnotic session. These are called posthypnotic effects, and include posthypnotic amnesia and other posthypnotic behavior. Posthypnotic suggestions are either cue induced, that is, the experience or behavior occurs in response to a specific cue given previously during trance, or are graded, that is, the effect occurs progressively, such as the posthypnotic suggestion that the meaning of a hypnotic dream will become increasingly more clear over time.

DIMENSIONS AND PHENOMENA OF THE HYPNOTIC EXPERIENCE

Clinicians suggest that although hypnosis is an altered state, the capacity to experience the phenomenon varies among individuals according to their biological substrata and cognitive and emotional styles. Most would agree there is a both a general capacity and a variety of specific capacities of hypnotic responding.

Ronald Shor identified eight dimensions of hypnotic experience, each of which varies according to the hypnotic subject: Fading of the individual's generalized reality orientation, nonconscious involvement, archaic involvement, drowsiness, relaxation, vividness of imagery, absorption, and access to the unconscious. Peter Sheehan and Kevin McConkey concluded from their research that the subjective experience in hypnosis depends on specific expectancies, cognitive style, the nature of imagery, perception of the hypnotist, and reality awareness. The data revealed "patterns of individual differences in reported hypnotic experiences" and that hypnotic experiences are different across subjects, depending on that subject's cognitive style. Three cognitive styles were identified: concentrative, independent, and constructive. Ronald Pekala attempted to quantify the patterns of consciousness in hypnotic trance and concluded that there were specific differences between low and high hypnotizables. Low hypnotizables scored lower on voluntary control, self-awareness, internal dialogue, rationality, and availability of memory. High hypnotizables scored significantly higher on measures of hypnosis as an altered state, absorption, and altered time sense. These data suggest that high hypnotizable's experience of an altered state of consciousness or trance is fundamentally different from the state experienced by medium or low hypnotizables.

DEFINITION OF HYPNOSIS

Until recently there did not appear to be a consensual definition of hypnosis. Earlier hypnosis was defined by any induction procedure followed by a series of suggestions. Most experts agree that hypnosis is associated with a set of unusual subject experiences and behaviors, and most agree with a limited *state definition* of hypnosis, with respect to the redistribution of attention. The central issue of ongoing debate is whether

hypnosis entails a discrete altered state of consciousness. The 2003 definition of hypnosis by the American Psychological Association (APA) failed to mention either self-hypnosis or hypnosis as an altered state of consciousness. In 2015, a new APA task force was able to reach a consensual definition. Hypnosis was defined as "A state of consciousness involving focused attention and reduced peripheral awareness characterized by an enhanced capacity for response to suggestion." An hypnotic induction was defined as "A procedure designed to induce hypnosis."

Sakari Kallio and Antti Revonsuo proposed a solution to the state/nonstate controversy. They convincingly argued that there might be two groups of hypnotizable individuals—a smaller group of highly hypnotizable individuals who have a propensity toward extreme changes in sensation, perception, and cognition, and for whom an altered state definition of hypnosis best applies, and a much larger group of mild to moderately hypnotizable individuals, whose state changes are less extreme and whose experiences are more a product of social context. Although both the state and sociocognitive theories of hypnosis partially explain overall hypnotic responding in both groups, the altered-state and trait models of hypnosis may explain a greater portion of the variance of the high-hypnotizable group, and the sociocognitive task-motivational and response-expectancy models may explain a greater portion of the variance of the medium-hypnotizable group. (A similar argument was made by Herbert and David Spiegel.) Kallio and Revonsuo's proposal resurrects the problem discussed at great length by Ernest Hilgard, namely that hypnotic responding appears to be bimodally distributed.

Kallio and Revonsuo have expanded the meaning of bimodality to include the idea that medium hypnotizables may respond more to social influence, whereas high hypnotizables may respond in a way that is at least partially independent of both social influence and hypnotic context.

HYPNOSIS AS REDISTRIBUTED ATTENTIVENESS

According to Amir Raz, "attention is an integral, defining aspect of hypnosis." Hypnosis is seen as a condition of "atypical attention." Limited state models emphasize hypnosis as condition of altered or heightened attentiveness and the effects of altered attentiveness on hypnotic experience and behavior. The hypnotic state is associated with changes in at least three of the normal attentional systems: alerting, orienting, and executive control:

- Relaxation suggestions commonly incorporated into a hypnotic induction alter waking alertfulness but do not cause sleep. David Spiegel viewed hypnosis as alert activation without arousal or vigilance, a type of alert consciousness lacking sympathetic arousal.

- Hypnotic trance is also associated with a diminished generalized orientation to external reality and inward absorption, but also with enhanced orienting to specific hypnotic suggestions while avoiding distraction.
- Hypnotic trance is also associated with changes in executive attentional processes, which results in (a) enhanced selective focus on suggested effects along with reduced focus on peripheral cues or on parallel competing attentional demands; (b) alteration in the allocation of divided attention; (c) alteration in error processing, such as the suspension of critical judgment and normal reality monitoring; and (d) alteration in response monitoring, as exemplified by the facilitation or inhibition of motor response, the shift in response criteria for reporting memory recall, or the suspension of belief with respect to suggested hallucinatory experiences.

In addition to these three frontal attention systems, the parietal lobes play an important role in levels of awareness, broadening or narrowing the field of consciousness, and shifting from local to global attention, and they may also play an important role in the hypnotic experience. The magnitude of these attentional shifts in a hypnotic state is significantly greater than in, for example, a waking absorbed state while reading a book, a deeply relaxed, nonhypnotic state, a state simulating hypnosis, or in whatever state is characteristic of low hypnotizables.

TRANCE CHARACTERISTICS

Another version of the state model uses "state" to imply a *discrete altered state of consciousness* or *trance,* uniquely different from the normal waking state and characterized by a set of unique *trance characteristics,* which include a fading of the generalized reality orientation; involuntarism; a suspension of normal critical judgment and reality monitoring (sometimes called trance logic); dissociation; and the capacity for altered cognitive-perceptual information processing. Some versions of the altered state model stress the relationship between the trance induced by a hypnotic induction and naturally occurring states outside of a hypnotic context such as self-hypnosis, daydreaming, stares of absorption, and dissociative states.

- Induction of trance is associated with a relative *reduction in generalized external reality orientation* and increased inward absorption. Hypnotized individuals are either less aware of, or aware of but less responsive to, external stimuli than in the normal waking state.
- A related component of hypnosis is *absorption*—a narrowing and intensification of attention resulting in complete focus. The shift into the hypnotic state is something like a psychological zoom lens that can shift toward highly focused attention. As attention becomes more intense and

focused, awareness of orientation in time and space decreases.

- The experience of *involuntariness* or automatic responding is considered a benchmark of high hypnotizability. For example, although it is clear that the hypnotized subject is using active imagination to experience the hands moving together in response to the suggestion that the hands are drawn together like magnets, the subject typically reports that the hands "seemed" to move together by themselves, as if top-down, voluntary executive control over the suggested response was suspended and it occurred in an automatic, nonvolitional manner. John Kihlstrom emphasized that perceived involuntariness does not necessarily imply suspension of volitional control during trance. Sociocognitive theorists have stressed that perceived involuntariness is not really involuntary, but is a function of purposeful, goal-directed cognitive activity in response to hypnotic expectancies and the nature of the hypnotic interaction. Irrespective of differences in the explanations for involuntariness, most experts concur that perceived involuntariness is a central feature of hypnotic subjective experience.
- Some high-hypnotizable subjects manifest a transient suspension of *critical judgment* during trance, which allows them more easily than in the walking state to believe in the actuality of suggested positive hallucinations, to deny the actual reality in suggested negative hallucinations, accept suggestions for reduced pain experience and hypnotic analgesia in the face of actual pain experience, accept suggestions that confuse imagination and actual perception, or accept the actuality of hypnotically suggested pseudomemories. However, the extent to which a suspension of critical judgment is a defining feature of trance remains largely untested. Few studies exist that experimentally test suggestions to improve critical judgment during trance, except for a few studies that suggest warnings about possible inaccuracies in hypnotic recall, which tend to show lower memory error rates while not eliminating memory error altogether. Trance logic is often referred to as a strong belief about suggested experiences in hypnosis that is logically incongruous from the perspective of the normal waking state, such as the presence of a hypnotically hallucinated person. Sometimes hypnosis is associated with a suspension of normal waking initiative, planning, and goal-directed behavior, so that the hypnotizable subject passively and uncritically responds to suggestions. However, some experts have emphasized hypnosis as an active cognitive, problem-solving endeavor. The common factor may be hypnotically altered action planning, either in the direction of less or more action planning, depending on the hypnotic context.
- The hypnotic state has been characterized as a *dissociative state.* However, there are two fundamentally distinct uses of the term *dissociation:* (1) dissociation as shifts in state characterized by a

narrowing or expansion of the field of conscious awareness and (2) dissociation as disintegration of the normally integrated operations of waking consciousness resulting in the dissociative compartmentalization of experience.

Dissociation as state shifts and changes in the field of conscious awareness, as exemplified by absorption, consistently accounts for a significant but small portion of the variance of overall hypnotic experience. *Dissociation as compartmentalized states* is more characteristic of high- than less-hypnotizable subjects, such as the dissociation of the "part of the mind" experiencing of hypnotic analgesia from the "hidden observer" part of the mind still registering actual pain experience, or dissociation of the emergence of conflictual mental content during trance from the normal waking defensive operations against the emergence of this material.

- High hypnotizability may represent a capacity for altered *cognitive-perceptual information processing*. The fading of the generalized reality orientation results in an overall reduction in the amount of information, and information processing is relatively restricted to the verbal suggestions supplied by the hypnotist. Temporal or sequential (serial) information processing is characteristic of the normal waking state.

Evidence from two sources implies that information processing in the trance state favors parallel, simultaneous over serial, temporal processing: (1) disordered recall after removal or suggestions for posthypnotic amnesia and (2) increased ability to generate random numbers in the hypnotic state as compared to the waking state, wherein the spontaneously generated numbers tend to be patterned. As a consequence of altered information processing, suggested or spontaneous time distortion is a characteristic feature of high hypnotizability.

Advocates of the altered-states-of-consciousness model of hypnosis such as Charles Tart view the trance state as a discrete state of consciousness with a set of special characteristics. From their perspective, the hypnotic induction ceremony constitutes a set of expectancies and specific suggestions designed to induce and ratify the special characteristics of the hypnotic state as markedly different from the normal waking state. The unique characteristics of the trance state are believed to correlate with heightened suggestibility on one hand and the clinical efficacy of hypnosis on the other.

SOCIAL INFLUENCE MODELS FOR HYPNOTIC RESPONDING

According to the *task motivational theory* of hypnosis, responsiveness to hypnotic suggestions does not require the induction of a hypnotic state but depends on the tasks set by the hypnotist, the manner in which these tasks are communicated, and the way in which the hypnotist shapes the

attitudes, motivation, and efficacy expectancies of the hypnotic subject. Attitude pertains to whether hypnosis is seen by the subject as a desirable or feared state. Motivation pertains to the willingness of the subject to be hypnotized. Efficacy expectancies refer to the hypnotic subject's belief in his or her capacity to become hypnotized.

Theodore Barber's research demonstrated that most of the unusual subjective experiences and behaviors characteristic of high hypnotizables could be produced without a hypnotic trance induction if the subject were motivated to the hypnotic task. Although these experiments led to a greater appreciation of the relative contributions of attitudes, motivation, and efficacy expectancies to overall hypnotic responding, they did not refute the contribution of a trance trait state to overall hypnotic responding, at least for certain high hypnotizables. More recently, social influence models of hypnotic responding have been revitalized in the form of a *response expectancy model*. Advocates emphasize the relative omission of expectancy effects in Andre Weitzenhoffer's suggestion model and Ernest Hilgard's state model. Response expectancies are defined by Irving Kirsch as "anticipations of automatic reactions to particular situations," such as automatic, nonvolitional behaviors or emotional states, and are especially relevant to the perception of internal states. According to sociocognitive models, response expectancies regarding hypnotic response are the consequence of culturally mediated a priori expectancies about hypnosis and/or specific information shaped by the interaction with the hypnotist. Response expectancies pertain to the prediction of probability of successfully responding to a given hypnotic situation. Within a hypnotic context, hypnotic subjects learn to produce a given hypnotic response according to their cultural understanding of hypnotic role behavior and/or the specific expectancies shaped by the hypnotist. In other words, hypnotic response is the outcome of social influence and the systematic shaping of hypnotic expectancies by the hypnotist. For example, suggestibility increased by only a small amount with an hypnotic induction when the procedure was labeled as "relaxation" but increased by a much larger amount when the procedure was labeled as "hypnosis."

Although expectancy models of hypnosis are valuable in helping in the appreciation of the contribution of expectancy effects to overall hypnotic responding, these models have several limitations. The overall contributions of expectancy effects are sometimes overstated, as if expectancy were "the final common pathway to hypnotic suggestibility." Although expectancy theory has helped in the appreciation of the relative contribution of social influence to overall hypnotic responding, it does not serve as a complete explanation for hypnotic responding, any more than placebo effects serve as a complete explanation for drug effects. Both expectancy effects and state-trait effects contribute to overall hypnotic responding.

THE STATE/NONSTATE CONTROVERSY

An ongoing debate continues between state theorists and sociocognitive theorists of hypnosis. State theorists have turned to psychophysiological and functional imaging studies of hypnotic responding to support limited- or altered-state theories of hypnosis and claim that sociocognitive theorist ignore the evidence from neuroscience. The evidence supports altered brain functioning during hypnotic responding, at least for high hypnotizables, and the hypnotic induction does activate or deactivate certain frontal attentional systems and other brain areas. A review of the neurobiological evidence follows in the next subsection.

Sociocognitive theorists argue that trance characteristics and the altered state itself are a product of social influence and the shaping of response expectancies, and that all of the experiences and behaviors characteristic of the domain of hypnosis can be produced through social influence without prior hypnotic induction or a trance state.

USING HYPNOTIC CHARACTERISTICS IN CLINICAL CONTEXT

Although the controversy over definition and specifics of theory remains, the clinical outcome of using hypnotic phenomena is undisputed. The effect of the clinical phenomena of hypnosis may rest on the biology of the patient, the presenting clinical problem, and the skills of the clinician. What follows are some of the hypnotic phenomena identified in the previous subsections that summarize strategies that can be used.

- *Catalepsy*. An involuntary tonicity of the muscles so that limbs remain where they are positioned with an inhibition of fatigue. It is used for the assessment of responsivity to hypnosis, for trance ratification, to help with obtaining radiography, for induction of deepening, and to facilitate anesthesia and analgesics.
- *Amnesia*. Inhibition or disruption of memory, which may be either spontaneous or suggested. Spontaneous amnesia suggests a deeper trance state. Duration of amnesia is unpredictable. Amnesia assists posthypnotic suggestions to occur by reducing conscious "spectatoring."
- *Dissociation*. A detachment from the immediate environment, similar to dreaming when one sees oneself doing something from outside of one's body. It is useful in promoting increased depth, in pain relief, and in facilitating "automatic writing" for purposes of unconscious hypnotic exploration.
- *Analgesia and anesthesia*. Experiencing analgesia is like turning down the dimmer switch on a light and is characterized by some milder grimace. Estimates are that 70 to 90 percent of subjects may develop at least some degree of analgesia and perhaps 20 percent can develop complete anesthesia (even during surgery). This phenomenon is useful for managing pain, reducing inflammation and facilitating more rapid

healing and providing a method of trance ratification.

- *Hyperesthesia*. Increased sensitivity to touch.
- *Ideosensory activities*. This refers to the capacity of the brain to develop responses on any of the sensory modalities: auditory, visual, kinesthetic, olfactory, or gustatory.
- *Somnambulism*. This is one of the deepest levels of hypnosis and refers to being in trance but with the eyes open. A patient may go into trance and the untrained clinician may not be aware of its significance.
- *Hallucinations*. A negative hallucination refers to not seeing or sensing a stimulus that is present. Negative hallucination for smell (anosmia) may be produced for a pungent odor in perhaps 25 to 35 percent of subjects. Having a positive hallucination is perceiving something that is not there. All five senses may be hallucinated. Taste and smell are the easiest to hallucinate, vision is the most difficult (only 3 percent were able to do this in Hilgard's research sample), and hearing is the second most difficult to hallucinate.
- *Hypermnesia*. Refers to increased capacity for memory recall.
- *Age regression*. (1) *Revivification* refers to a complete regression or reliving of a past event; voice and writing become child-like. (2) In *partial regression* to a past time or event, the patient retains some awareness of the hypnotist and the patient's adult identity and thus has the capacity to apply an adult perspective to the past experiences. This can be an instrumental means in facilitating psychotherapy, uncovering repressed or dissociative data, and working through poor object relations and attachments.
- *Ideomotor activity and ideomotor exploration*. This refers to the involuntary capacity of muscles to respond instantaneously to thoughts, feelings, and ideas—for example, the experiencing of fingers or limbs floating and seeming to move by themselves. Ideomotor exploration is a method for rapid unconscious exploration using involuntary finger signals. These signals may also be used to signal the therapist when internal processes have been completed (e.g., when numbness is present).
- *Age progression*. Pseudo-orientation in time to the future may seem like a fantasy or mental rehearsal about the future to some subjects, but to talented hypnotic subjects it can feel quite real. Purposes include covert rehearsal and confidence building; the diagnostic determination of how the patient will react to a future situation; and, in some cases, the determination of what treatment "helped you recover." In addition, for some depressive states it also facilitates allowing the patient to foresee the possible future. This can be an invaluable approach in facilitating psychotherapy and working through poor object relations and affiliations.

- *Time distortion*. There are two possible aspects: (1) Time expansion, by which 2 minutes, say, can seem like 20 minutes, with a subjective sense of time going by very slowly. An everyday example would be waiting for a ride on a snowy day or waiting in a slow checkout line. (2) Time contraction, by which 10 minutes, say, can pass in what seems like 1 minute during a pleasurable activity or by suggestion. Purposes can range from sports activities, to pain or migraine, to past review of etiological factors, to trance ratification to convince patients that they were in an altered state of consciousness. It can be used to expedite painful times while going through medical procedures, as well.
- *Depersonalization*. Through suggestion an individual can temporarily forget his or her identity and view traumatic past experiences like an objective intelligence or may take on the identity of another person and identify with him or her.
- *Induced dreams*. Hypnosis may influence dream content at night and "dreams" may be produced in a hypnotic state. Purposes include exploring and revealing causes and conflict, identifying solutions and insights, altering feelings, and recalling nocturnal dreams. In a hypnotic state, sleep ideation also may be "searched" in treating insomnia or physiologic disturbances (e.g., morning ulcer disturbance or premature labor).

NEUROBIOLOGY AND NEUROPSYCHOLOGY OF HYPNOSIS

A limitation in the physiological studies of hypnosis is the failure to distinguish between the physiological effects of the hypnotic state per se and physiological effects of specific hypnotic suggestions given within that state.

Seemingly innocuous changes in the wording of specific suggestions, such as "you're becoming more and more relaxed," "focus your attention on," "your breathing will become slower and slower," or "allow a scene to come into your mind," may produce substantial physiological changes, so that without careful control of the wording of suggestions, the overall physiological effects of hypnotic suggestions may seem to be a wash.

Electroencephalographic (EEG) studies of hypnosis have failed to reveal a clear-cut EEG signature to the hypnotic state. The hypnotic state consistently is associated with increased theta activity in high versus low hypnotizables, yet the type and location of theta activity may be more important in hypnotic responding than theta activity per se. Posterior theta activity is more a function of relaxation than hypnotic responding per se, but it might signify hypnotic responding if the overall hypnotic suggestions include engaging in progressive relaxation and passively allowing experiences to arise, whereas anterior theta may reflect the activation of attentional systems and active cognitive processing in response to specific hypnotic suggestions. Furthermore, the frequency of theta activity may

indicate the presence of imaginative as compared to realistic responding in high versus low hypnotizables and increased attentional focus. Increased alpha activity has sometimes been associated with the hypnotic state in high hypnotizables, but these findings have been disputed.

Increased high-frequency gamma activity (40 Hz) is associated with focused arousal in high as compared to low hypnotizables with and without a hypnotic induction. The eye roll sign in certain high hypnotizables is associated with left frontal and bilateral occipital activation. EEG findings clearly establish that hypnosis is not sleep.

Thus, although no definitive overall EEG signature of hypnosis exists, specific EEG patterns are associated with various components of overall hypnotic responding, such as suggestions for eye roll, hypnotic relaxation, imaginative involvement, receptive attention, and intensely focused attention.

Functional imagery studies such as positron emission tomography and fMRI studies have consistently documented significant differences in the brain functioning of high and low hypnotizables, so that high hypnotizability may constitute a unique brain state. Induction of a hypnotic sate may represent the alteration of brain functions operative in the regulation of consciousness. Helen Crawford and her associates argued that high as compared to low hypnotizables have more efficient executive control systems and supervisory attentional systems. Low hypnotizables do not activate these same brain attentional systems. Nevertheless, it is also clear that the activation and deactivation of localized brain areas during hypnosis are strongly associated with the specific hypnotic task suggested. There may be no single brain activation/deactivation profile for hypnosis, but instead a number of neurocognitive patterns, depending on the nature of the hypnotic induction and the specific suggestions given.

Various components of the hypnotic induction may also be associated with specific patterns of brain activation/deactivation as suggestions affect the alerting, orienting, and executive attentional networks accordingly. Amire Raz provided fMRI images of the three attentional networks activated in hypnosis.

Hypnotic relaxation was associated with lower levels of activation of the mesencephalic tegmentum of the brainstem, the thalamus, and the anterior cingulated cortex (ACC), comparable to changes produced by waking decreased vigilance or sleepiness, whereas hypnotic absorbed attention was associated with higher levels of activation in the brainstem, thalamus, and rostral ACC. The overall hypnotic induction resulted in lower levels of activation of the parietal areas consistent with an overall reduction in the field of conscious awareness, temporal and spatial distortion, and disorientation of self. According to John Gruzelier, the traditional induction is not a haphazard assortment of roles that the subject is encouraged to enact, but follows a sequence with neurophysiological logic

—“the engaging of anterior left-sided attentional mechanisms, which, once engaged, can be suppressed, which in turn allows selective inhibition of frontal functions permitting cardinal features of the hypnotic experience such as automaticity and involuntariness to take place.” Gruzelier added that such selective inhibition and enhancement of certain brain attentional systems also explain the dissociation component of hypnosis. Overall, the evidence strongly implicates the role of the frontal/parietal system in hypnotic responding, and particularly the role of the ACC as an explanatory system for many of the phenomenon of hypnosis. High as compared to low hypnotizables have stronger attentional filtering abilities as manifested in the capacity for focused and sustained focused attention, and also in the ability to ignore irrelevant stimuli, that is, high hypnotizability is a condition of both “extremely focused attention and disattention.”

The left dorsolateral prefrontal cortex (DLPFC) is part of the neurocircuitry of meta-cognitive monitoring. Low frequency repetitive transcranial magnetic stimulation of the left DLPFC, which presumably blocked the functioning of the DLPFC, increased hypnotizability. This finding is consistent with the view that suspension of critical judgment occurs in high hypnotizable subjects, and that deactivation of the left DLPFC may play an important role in hypnotizability.

The frontal systems are implicated in both the facilitation and inhibition of motor responding in hypnosis. Functional imaging studies both of patients with motor conversion paralysis and normal subjects given hypnotic suggestions for motor inhibition have shown that similar patterns of functional brain changes are involved in each case, namely right ACC activation and right orbitofrontal activation, along with deactivation of the right primary motor cortex. When high-hypnotizable subjects who respond positively to a range of ideomotor suggestions were given either hypnotic suggestions for limb paralysis or suggestions to feign paralysis, different patterns of brain activation were associated with hypnotically suggested and simulated paralysis.

Hypnotically suggested elimination of interference on the Stroop test is mediated by ACC activation. In the Stroop test, subjects name a color in which words are displayed. Response time is generally slower if the name is different from the color. Hypnotic suggestions to ignore competing responses during the test allow for response over-ride and quicker response to the target stimulus. However, whereas both waking and hypnotic suggestions improve Stroop performance in high versus low hypnotizables, hypnotic induction is not necessary for suggestive reduction in the Stroop effect in high hypnotizables.

High hypnotizables with the ability to respond to hypnotically suggested hallucinations as compared to subjects who could not hallucinate showed activation of the right ACC when they were actually hearing an audiotaped message and also when hypnotically hallucinating the taped message but

did not show ACC activation when instructed to imagine hearing the message. These data suggest that in terms of brain activation, hypnotic hallucinations are comparable to actual perceptual events and not simply the product of imagination.

Functional imaging studies have shown comparable ACC activation in persons with schizophrenia with auditory hallucinations. With respect to hypnotic visual hallucinations, highly hypnotizable subjects showed activation of the fusiform and lingual areas typically associated with color perception in response to a hypnotic suggestion to perceive a color whether actually shown a color or a gray-scale stimulus. In terms of brain activation, hypnotically hallucinated color perception and actual color perception are similar.

Hypnotic suggestions to increase or decrease pain experience cause changes in the activation of the ACC, insula, and parietal areas, and sometimes in the primary sensory cortex. Direct suggestions for pain alleviation activate cognitive (via prefrontal and orbitofrontal) and emotional (via insula, inferior parietal, lenticular nucleus and ACC circuitry). These neurobiological findings are consistent with the view that hypnotic analgesia suggestions may prevent higher cortical processing of nociceptive input, with inhibition being mediated by the cingulate cortex. High as compared to low hypnotizables also show a decreased amplitude in the later P100 and P300 wave components of pain information processing. The difference between high and low hypnotizables in the ability to control pain may represent not only functional but also structural brain differences, in that high hypnotizables who had the ability to control pain had significantly larger corpus callosum size than low hypnotizables.

Overall, the growing consensus is that hypnotic suggestions exert a top-down influence on frontal executive control systems characterized mainly as selective inhibition of various frontal functions.

In particular, the ACC plays a central role in the distribution of attention in the context of conflicting attentional demands, in the parallel distribution of attention, in distributing attention between cognitive and emotional processing, in meditating autonomic responses, in coordinating the intensity of attention, in motivational states and the initiation of goal-directed behaviors, in working memory, and in nociceptive experience. The selective inhibition of various frontal systems in general, and the alteration of ACC functioning in particular, in response to verbal hypnotic suggestions serves as a neurobiological explanation for such diverse hypnotic phenomenon as the fading of generalized reality orientation; intense focus and distraction resistance in hypnosis; hypnotic relaxation; hypnotic analgesia effects; the selective inhibition of competing response potentials, as exemplified by response to the Stroop test in hypnosis, or illustrated by hypnotic suggestions for selected deafness or focal blindness; the production of positive hallucinations; the selective recall of personal

memories in hypnosis, but also the diminished memory-error monitoring in certain subjects; the efficacy of hypnotic suggestion in evoking strong emotional states; and the seeming involuntary nature of hypnotic responding. It is well established that high hypnotizables have exceptional attentional filtering and sensory-perceptual gating abilities, so that high hypnotizability may constitute a selective pattern of brain activation/deactivation.

THE ASSESSMENT OF HYPNOTIZABILITY

As noted earlier, the first standardized hypnotizability scale was the SHSS, which consists of a standardized induction ceremony based on suggestions for progressive relaxation, slower breathing, and going deeper and deeper into trance.

Following the induction there are 12 specific suggestions, 9 of which are ideomotor suggestions; the other 3 include suggestions for a fly hallucination (presence of a buzzing fly), posthypnotic amnesia, and a simple posthypnotic suggestion. There are two comparable forms, Form A and Form B, both of which include the same hypnotic induction ceremony, but with somewhat different sets of the 12 specific suggestions. The SHSS is read verbatim to the hypnotic subject, and there is a standard set of scoring rules. Scoring is based on objective, behavioral response to a given suggestion. A given subject can score anywhere from 0 to 12 on the SHSS. An audiotape self-scoring version of the SHSS was also developed—the Harvard Group Scale of Hypnotic Susceptibility (HGS). It contains a similar induction and many, but not all, of the same items as the SHSS. Both the SHSS and the HGS take just about 1 hour to administer, but the HGS is time saving in that its administration does not require the presence of a hypnotist or clinician. The Stanford group also developed a graded difficulty scale, the SHSS, Form C. It contains a standard induction procedure followed by 12 specific hypnotic suggestions, each of which has been empirically shown to be somewhat more difficult than the previous item. The selection of items on Form C shows greater diversity and range than the items on the A or B scale. It contains only 4 ideomotor items, 5 hallucination suggestions, a suggested dream, a suggestion for age regression, and a suggestion for posthypnotic amnesia. There is also a self-scoring audiotape version of Form C, the Waterloo-Stanford Group Scale-Form C. London also used the Stanford scales to compose a standardized scale to assess hypnotizability in children—the Children's Hypnotic Susceptibility Scale (CHSS). The Barber Suggestibility Scale is unique, in that is measures response to 8 suggestions in a standardized manner but does not include a hypnotic induction ceremony. All of these scales meet acceptable standards of reliability and validity.

Some controversy exists about the value of using laboratory-derived hypnotizability scales in clinical settings. Some critics argue that these

scales contain a number of challenge items and create a performance-oriented ambience that is contrary to the holding environment of therapy. Others have argued that low performance on standardized scales may discourage patients from using hypnosis. However, one study showed that those given feedback of low scores as compared to high scores were *more* not less inclined to request hypnotherapy. Another criticism is that scales like the SHSS Form A scale and HGS lack range, in that 9 of the 12 items are ideomotor items. The scales do not contain items representative of the range of hypnotic phenomena directly relevant to the clinical context. The SHSS C scale has greater range and is relatively exempt from this criticism. A more practical criticism is that most of these scales require about 1 hour of face-to-face clinician–patient interaction to complete and in that sense are not time efficient. The use of audiotape self-scoring versions like the HGS and the WSGC fully addresses this problem. Having a patient listen to a tape in another office and then returning the self-scored result to the clinician at the end of 1 hour is remarkably time efficient. Moreover, the clinician is able to chart normative data of hypnotizability with little time investment. The HGS is the assessment instrument of choice in clinics seeking normative documentation on hypnotizability with little allotment of time to accomplish this.

A number of standardized clinical scales were developed to meet the needs of clinicians. The Hypnotic Induction Profile (HIP) begins with an eye roll sign, in which the patient is instructed to look upward toward the eyebrows, then toward the top of the head, and then, while continuing to look upward, to close the eyelids slowly. The eye roll sign is alleged to be a biological indicator of dissociative capacity. The eye roll is measured in terms of the visibility of the sclera between the cornea and the lower eyelid as the eyelids are closing. The eye roll is scored from 0 to 4. The rapid objectivity also fits well into the medical model. Observing the sclera is easily demonstrated on medical rounds and can facilitate debunking myths of the skeptical. Upon closing the eyelids, floating suggestions are used to induce a trance state. The patient is then told that the left arm will feel light and will float into the air until a cut-off signal is given. Then the arm is lowered prior to giving the cut-off signal, and hypnotic responsiveness is assessed in terms of the speed and number of prompts given before the arm floats up again. The great advantage of the HIP is that it only takes 5 to 8 minutes to administer. The experienced clinician is able to complete the administration in less than 5 minutes. As the first standardized clinical scale developed, the HIP has enjoyed enormous popularity among clinicians. The Spiegels have collected a normative sample of HIP scores on almost 10,000 patients and nonpatients. The disadvantage of the HIP is that it lacks range. The instrument is largely constructed around response to an arm levitation suggestion along with subjective experiences associated with that action. However, because response to ideomotor

suggestions like arm levitation are predictive of response to other dimensions of hypnosis, the choice of building a quick, standardized assessment scale around arm levitation is reasonable.

The SCSHS contains a standardized hypnotic induction followed by five specific suggestions: hands moving together, dream, age regression, posthypnotic amnesia, and posthypnotic suggestion. It is administered face to face and takes about 20 minutes. The SCSHS has a number of advantages.

It is time efficient, and it assesses mostly the range of hypnotic abilities that are directly relevant to the clinician: dream, age regression, posthypnotic amnesia, and posthypnotic suggestion. The disadvantage is that scoring allows for ready identification of high hypnotizables but does not easily discriminate between low and moderate hypnotizables.

Because the items are reasonably difficult, a small number of patients score 4 to 5, and the majority score 1 to 2. It is difficult then to know whether a score of 1 to 2 represents low or medium hypnotizability.

Overall, no single standardized hypnotizability scale perfectly meets the needs of the clinician. Clinicians wishing a time-efficient way to assess hypnotizability must decide between the HGS and the HIP, with the HGS having better normative data and greater range and the HIP being better at detecting high hypnotizables. The Stanford Clinical Scale taps the best range of clinically relevant dimensions and is relatively time efficient. Although not always practical, of course, the most empirically sound approach is to administer several standardized hypnotizability scales with the same patient, although the scales result in similar but not identical classifications of high-, medium-, and low-hypnotizable subjects.

THE HYPNOTIC INDUCTION CEREMONY

Purpose and Preparation

The hypnotic induction ceremony is a metaphorical performance that suggests a change in conventional verbal discourse. In a more traditional sense, it implies that a ritual to facilitate induction into a different mental state is going to occur. Despite the great variation in hypnotic induction ceremonies, most contain certain common elements. For example, most induction ceremonies contain specific suggestions to focus attention on some target, as well as suggestions to disattend to everything else. Most induction ceremonies strategically use language to shape the progressively developing and deepening experience: “You are becoming more and more relaxed,” “You are going deeper and deeper into this state.”

Some trance inductions, especially with highly hypnotizable subjects, explicitly contain suggestions for entering an altered state of consciousness or trance state, and they may also contain explicit suggestions to enhance or ratify certain trance characteristics such as involuntarism, dissociation,

altered time sense, and so on. Hypnotic discourse is often very repetitive.

Because repetition and monotony are associated with destabilization of ordinary waking consciousness, the hypnotic subject becomes more open to unique subjective experiences characteristic of trance. As part of the hypnotic induction ceremony, suggestions are given to increase the subject's motivation to be hypnotized and find the hypnotic experience desirable and to increase the expectation that the subject will be able to respond successfully to the given suggestions.

Many different inductions are used clinically, but the most successful one will be the one with which both the patient and the practitioner feel most comfortable. Both will have preferences that will influence the success of any given technique. It is appropriate for each clinician to develop his or her style that facilitates the incorporation of responsiveness to the patient to allow for the progressive depth of trance to occur. Although many textbooks and pamphlets describe the actual litany that should occur, they are merely templates of ideas that have been shown to work. In the clinical course of intervention, it is imperative that the practitioner be able to change course smoothly if it becomes clear that a particular technique or suggestion is not effective in order to allow the patient to remain in a state of relaxed attentiveness. A good guideline is that all hypnosis is self-hypnosis, and the clinician's role is to facilitate the patient's ability to make maximum use of the patient's trance capacity.

Before beginning, it is helpful to prepare the patient for the experience by discovering what his or her expectations are and modifying them accordingly with prehypnotic preparation. For example, popular culture's depiction of stage hypnosis performances may convey a total lack of control, and such negative expectancies can be corrected during the prehypnotic interview. Discussion of prehypnotic expectancies may also address helping the patient to become more comfortable with the idea of suggestibility and the nature of subjective experiences typically associated with hypnosis.

If the patient fails to enter what the or he believes is a trance state, then it is beneficial to try to reframe the failure as something positive, such as suggesting that the patient has unusually good control over his or her muscles, and that the discovery of this fact will be very helpful to know for planning future therapy options: The patient can use that control for increasing relaxation and reaching further depths than might not otherwise be considered possible. In most cases, patient expectations play an even greater role in success than the patient's level of suggestibility. Facilitating appropriate expectancy of change will likely enhance the therapeutic alliance and set the stage for induction.

Most induction ceremonies are relaxation and/or awareness based, the difference being whether the patient is concentrating on releasing tension or ignoring the body to concentrate on something, such as the rhythm of

the breath or walking down stairs. The most commonly used inductions are as follows.

Progressive Relaxation

The patient is asked to relax, first concentrating on slowing the breathing and then progressively releasing tension in all parts of the body, beginning with the feet and ending with the head/eyes.

Holding On and Letting Go

For some patients who are reticent to relinquish control, it is easier to first tense the body and then release the tension. Emphasis is placed on relaxing and feeling safe and secure, and holding on to only as much tension as is required for the patient to feel in control enough to avoid anxiety about submission. The patient is asked to take deep breaths, release the breaths completely, and then concentrate on how good it feels to do this. The patient is then asked to progressively tense and release each muscle group starting either from the feet or the head.

Staircase/Escalator

This is a version of progressive relaxation in which the patient is asked to imagine that he or she is walking down a staircase with 10 steps to the bottom. At the top of the steps is the patient's current state, and at the bottom is the most relaxed, safe, and comfortable state possible. The patient is instructed to descend each stair slowly on the clinician's invitation, and in each step the patient relaxes further (often suggestions such as "You will find that you are twice as relaxed on the next step as you were on this last one" are helpful), slows the breath further, and becomes more at ease. As with the previously described techniques, it can be helpful sequentially to mention relaxing muscle groups such as eyes, face, neck, shoulders, chest, abdomen, pelvis, thighs, calves, and then feet on each step.

Eye Fixation/Closure

The patient is asked to stare at a target that is preferably on the wall or ceiling that is somewhere above the normal field of vision so that some eyestrain is provoked. It is then suggested that the patient's eyes are beginning to become heavy, and that the body is becoming more and more relaxed. The patient is monitored for signs that this is working, such as blinking, watering eyes, eyelids beginning to droop, and so on. These are noted verbally to the patient as assurance that he or she is entering the hypnotic state.

If a sufficient time has passed without the patient closing the eyes completely, he or she may be invited to do so, with the statement that this

is to allow the patient to further experience and imagine the experiences that the clinician will be suggesting.

Visual Imagery Techniques

The patients are instructed to use an image to visualize a scene or scenario from the past, present, or future. Throughout the sequence, patients are instructed to relax and focus, and they are reinforced for their style.

Eye Roll Technique

During this application of this technique, patients continue typically with the screening procedure developed by Spiegel, and the technique is reinforced. Patients are asked to roll their eyes to the ceiling while taking a deep breath and holding. They are then given the suggestion for relaxation and for feeling more and more comfortable. After the induction occurs, the specific strategy begins for most patients, such as for pain, habit control, anxiety, phobia, and so on. Some clinicians describe using the same strategy for all patients, but individual strategies based on each patient's personality traits, motivation, and so on may be more successful. All of these techniques have in common paying attention, focusing attention, and decreasing peripheral awareness.

HYPNOTHERAPY TREATMENT PLANNING AND HYPNOTHERAPY TECHNIQUES

Hypnosis in and of itself is not a treatment. Hypnosis is the *medium* of treatment. Hypnosis is a state of heightened attentiveness (and sometimes an altered state of consciousness and sometimes a condition of heightened suggestibility) within which psychotherapy or suggestions are used for intervention to achieve a goal. The patient, by virtue of going into a hypnotic state, is absorbed in a task, blocks out peripheral awareness, and has rapid assimilation of data.

It is therefore imperative to use the best-matched treatment protocol for a given condition. A number of comparative outcome studies have been conducted in which the same treatment protocol was offered with and without a hypnotic induction. The overall treatment effect sizes are typically comparable in both the nonhypnotic and hypnotic conditions. The advantage is that a treatment protocol offered to a patient in a hypnotic state generally yields a faster effect, for a number of reasons:

- Discovering and learning new coping strategies is quicker in a nondistracted state.
- When more focused, the undercurrent of self-talk, schemas, affects, personal memories, and self-states is more accessible than in a nonhypnotic state, thereby giving an advantage to hypnosis in exploratory therapy.
- Hypnosis also mobilizes and intensifies the transference, aiding analytic

work.

- Treatment planning with hypnosis entails decisions.
- Assessing hypnotizability and deciding whether the patient has sufficient hypnotizability for hypnotherapy.
- Deciding whether the present problem is appropriate for hypnotherapy.
- Choosing the approach to hypnotherapy.
- Choosing the style of hypnotherapy.
- Using patient interests and personality style to facilitate induction and to maintain a level of trance.
- As in all treatment, developing a therapeutic alliance with the patient is the first necessary step for successful intervention.

A review of problems most amenable to hypnotherapy is included in the next subsection. There are a number of viable approaches to hypnotherapy. The goal of time-effective *problem-solving hypnotherapy* is symptom ameliorization, for example, in the ameliorization of pain or tinnitus. A good deal of the hypnotherapeutic interventions used in behavioral medicine focus on active coping and problem solving. *Exploratory hypnotherapy* emphasizes the exploration of psychodynamic factors operative in the development and maintenance of symptoms, for example, in the treatment of a conversion disorder.

Exploratory therapy also addresses the exploration of inner resources and active coping strategies to address symptoms. *Supportive hypnotherapy* uses suggestions to enhance the patient's sense of confidence and mastery. John Hartland called such suggestions ego-strengthening suggestions. *Developmentally informed hypnotherapy* aims at fostering normal development along developmental lines such as self-development, object relations development, and affective development. The hypnotherapist must decide which approach best meets the needs of the given patient's presenting problem.

There are three very different hypnotic styles. *Directive hypnotherapy* is derived from persuasion theory. The hypnotist serves as an authoritative source and gives direct suggestions for symptom ameliorization. This model is best suited to cultures in which doctor–patient relationships are more hierarchical than egalitarian, as was the case in the United Kingdom as represented in the older *British Journal of Medical Hypnosis* prior to 1970. *Permissive hypnotherapy* is based on the assumptions that hypnotizability is a talent within the patient, and that the patient has the inner resources to get well. In permissive hypnosis, the hypnotherapist provides just the right relational conditions in the context of an egalitarian relationship in which the patient is provided with a permissive relational context and encouraged to explore inner resources and coping strategies. Neo-Ericksonian *indirect hypnosis* is derived from the work of Milton Erickson with difficult patients—those who were not very hypnotizable

and/or who had a history of treatment failure. Indirect methods include the strategic use of metaphors and scripts, interspersal technique, and binds, double binds, and paradoxes. Although these methods are well suited for the so-called difficult patient, they are sometimes used as general hypnotherapy techniques.

The therapist must decide which hypnotic style is best matched to the given patient. In contemporary Western culture, a permissive approach is best suited to the majority of hypnotherapy patients. Directive hypnotherapy works with a small percentage of patients who fit the profile of higher hypnotizability and lower intelligence. Indirect hypnotherapy is best matched to patients with intractable symptoms of a history of treatment failure. For example, consider the introduction of hypnotic pain-coping strategies in a 12-week group of eight patients with chronic pain. One of these patients may respond quickly to direct, graded hypnotic suggestions that the pain will become less and less severe over time. The majority of the patients may discover and gradually master individualized pain coping strategies over the first two-thirds of the group using a permissive, individualized approach to group hypnotherapy. The other two patients may not respond to either of these approaches. Both may be likely to complain that the pain is becoming worse through treatment. These patients would then be given indirect hypnotic suggestions. For example, such a patient is told that it is very difficult to tolerate the uncertainty, never knowing when the pain will become worse. Therefore, as a means of gaining more control and certainty, the patient is asked to try voluntarily to increase the pain. Such paradoxical interventions, when matched to the right patient, serve to ameliorate the pain. Although such indirect methods should never be overused, they work well with a particular type of patient.

There are two main approaches to hypnotherapy: (1) a *protocol-based approach* and (2) an *individualized approach*. In this era of empirically supported treatment, there is a growing trend toward using established manualized treatments that include a set of necessary treatment foci for the treatment to be effective. A well-established treatment protocol derived from outcomes research outside of the hypnosis domain can be used with a patient in a hypnotic state.

This *hypnotic induction plus manualized treatment protocol* is becoming increasingly popular in, for example, the treatment of anxiety conditions, depression, sexual dysfunctions, and behavioral medicine. An alternative approach is to use hypnoprojective techniques (described later) to identify coping resources within the individual that helps the patient to resolve the problem or condition. For example, the hypnotized patient is asked to imagine a video about a person who had a similar problem but who discovered ways of overcoming it. A second video is about another person with the same problem who discovered a way to overcome the problem that was very different from that used by the first person. After a

series of three to five such scenes the patient is told to reflect on the scenes until the meaning of each becomes more and more clear and until the patient sees how such scenes inform him or her about the best way for overcoming the problem. Next, the patient is asked to imagine him- or herself in the scene using these coping resources effectively until he or she begins to develop a sense of mastery. The patient is given a posthypnotic suggestion that thoughts about these coping strategies will naturally come to mind in everyday life whenever the patient notices the very first signs of the symptom/problem occurring. Using this displacement technique, patients discover powerful inner coping resources through the metaphor of the "displaced other" and then learn to apply these coping strategies to their symptoms first in fantasy and then in everyday life as the symptoms just begin to intensify. The power of the displacement technique rests in the fact that all of the strategies for overcoming the symptoms come from the patient, leading to a strong sense of mastery and self-efficacy. A comparable approach is the inner wisdom method, in which the hypnotized patient is asked to consult with his or her "inner wisdom" to discover the best way of effectively dealing with a given problem.

There is a large repertoire of hypnotherapy techniques. *Hypnoprojective techniques* entail the symbolization of inner coping resources and unconscious processes. The "theatre technique," for example, entails having the patient visualize a theater with a play spontaneously appearing. To explore coping resources, suggestions can be given that "the play will somehow be about effective ways of coping with [name problem]." To explore psychodynamic processes, suggestions can be given that "the play somehow will be about whatever has contributed to the development of [name problem]." To explore transference, suggestions can be given that "the play will somehow be about the treatment relationship." The hypnotherapist can structure a play with several acts so as to gradually uncover the meaning of a given problem or gradually discover a variety of new coping resources for more effectively dealing with the problem. Other hypnoprojective techniques include the "television or video technique," the "old house technique," and the "library technique." Not all hypnoprojective techniques use imagery. *Automatic writing and drawing techniques* entail giving suggestions for dissociation of the hand, followed by suggestions that the hand is able to move involuntarily and to make writing or drawing movements on its own. Then suggestions are given that the dissociated hand is able to express inner coping resources and/or unconscious processes. A more detailed discussion of hypnoprojective techniques can be found in Daniel Brown and Erika Fromm's *Hypnotherapy and Hypnoanalysis* and John Watkins' *Hypnotherapy Techniques.*

In Calvert Stein's clenched-fist technique the patient is asked to make a fist, then scan the body, then imagine all the tension from various sources in the body gradually flowing into the arm, and then into the fist until it

gets tighter and tighter. Once all of the tension flows into the fist, the patient is told to let it go. Hypnotic suggestions can be given for soothing ("A scene will come to mind that is somehow about a deep sense of soothing or comfort"). Additional suggestions are then used to amplify to feeling of soothing or comfort. The metaphor of an affect dial is useful in regulating the intensity of affect state. The patient is asked to imagine him- or herself in a video at a time that he or she is feeling a particular emotional state that is difficult to tolerate.

Then the patient is asked to imagine holding a remote control device with a special volume dial that can turn up or down the intensity of the feeling until just the right level of feeling is found. A wide variety of affect regulatory techniques exist within the hypnotherapy literature, and many have recently been collected in Carolyn Daitch's *Affect Regulation Toolbox*.

The field of hypnotherapy contains a number of *methods for the rapid identification and resolution of conflicts*. John Watkins' *ego state therapy* consists in calling forth in trance each "part of the mind" associated with a given conflict, identifying the adaptive function of each "part," and then suggesting that other "parts of the mind" come forth giving various perspectives on resolving the conflict. For example, if a patient is not ready to commit to treatment of problem drinking, the clinician can call forth in trance the "part of the mind" that wants to stop drinking, then call forth the "part of the mind" that does not want to stop drinking, then heighten the cognitive dissonance between these two "parts," and then suggest that other "parts" come forth that have a perspective on resolving the conflict between the "parts" wanting to stop and not wanting to stop drinking.

Ultimately the dissonance is resolved *within* the patient, and the internal state shifts away from the extreme resistance to stopping drinking. Using Leslie LeCron's *ideomotor questioning technique,* the clinician establishes communication directly with the so-called "unconsciousness mind," which is told that it can communicate directly using automatic ideomotor finger signals, with one finger signifying "yes," another finger "no," another finger "don't know," and another finger "don't want to say." Then the patient's "unconscious mind" is asked a series of questions pertaining to the meaning of a given symptom, such as whether it serves an emotional need, whether it is associated with a past conflict, whether it is associated with a limiting belief, it is a way of identifying with someone, and so on. In this manner, the hypnotherapist is able rapidly and precisely to identify the conflict central to the development of and/or maintenance of a symptom. This ideomotor method has been used extensively in identifying factors involved in resistance to treatment, exploring psychosomatic problems, exploring psychological factors involved in premenstrual symptoms and infertility problems, and accurately assessing motivation and level of risk of suicidal behavior.

In general, a good suggestion for clinicians is to recognize that the

hypnotic evaluation can be made part of the overall evaluation and can offer information regarding patient processing and responding. Based on patient response to the HIP, Herbert and David Spiegel described patients as ranging along a continuum of low to high that they call the Apollonian–Odyssean–Dionysian cluster. Hypnotic and psychotherapeutic interventions are based on each patient's hypnotic capacity.

Prior to doing any hypnotic intervention, it is recommended that a biopsychosocial evaluation be performed along with a mental status exam. Recognizing and understanding the development of the behavior, symptom, or underlying problem can help to avoid later complications. Removing symptoms without understanding their development can lead to negative consequences.

HYPNOSIS'CLINICAL APPLICATIONS

As previously noted, hypnosis has been used as a clinical intervention in medical and mental health care in a variety of settings. Following the recognition of hypnosis as a form of medical treatment by the AMA in 1958, hypnosis has been used in the treatment of a variety of medical problems and a number of medical interventions. Although there is no uniform availability of hypnotic interventions, physicians, psychologists, and other mental health professionals do provide opportunities for patients to undergo hypnotherapeutic interventions for the alleviation of medical problems and as a form of medical intervention. Harold Wain, in the paper "Hypnosis on a Consultation-Liaison Service," described a variety of areas in which hypnosis has been used in the hospital setting.

In psychiatric settings, hypnosis has been employed for a number of psychiatric disorders, which are addressed later. Over the last 25 to 30 years, professional societies for hypnosis have developed beginning, intermediate, and advanced training programs for health care professionals for the application of clinical hypnosis to their patients. Furthermore, there have been a significant number of evidence-based studies that document the efficacy of hypnotic interventions for medical problems, medical procedures, psychiatric disorders, and psychological disorders.

Trauma and Hypnosis

Freud suggested that trauma is an overwhelming experience that creates havoc within the individual and puts into question one's homeostasis and psychological identity. According to David Spiegel, traumatic stress manifests as a feeling of helplessness and loss of control over one's safety.

Individuals who experience the extremes of arousal during a traumatic event are more likely to experience PTSD symptoms over time. Too high a level of anxious arousal has a disorganizing effect. Too low a level of arousal as a consequence of dissociation results in failure to integrate the traumatic

experience. In either case normal processing of the traumatic event is disrupted, and PTSD symptoms are more likely to develop over time. Dissociation is used as a means of defending oneself against what might otherwise be too overwhelming to manage.

The effects of trauma are generally classified into immediate aftermath effects, or acute stress disorder; chronic or delayed PTSD; and complex PTSD. Simple PTSD is generally associated with a single or circumscribed series of traumatic event(s).

Complex PTSD is generally a function of multiple comorbid psychiatric conditions, including a number of other psychiatric disorders (sometimes with a major dissociative disorder) and personality disorders. Chronic cumulative trauma, early childhood trauma, and/or trauma aggravated by early childhood attachment disorganization are associated with complex PTSD.

Because the development of simple PTSD after a traumatic event is correlated with higher levels of either anxiety or dissociation during the trauma, many have theorized that hypnosis might be an effective treatment due to its ability to calm anxiety and to breach the dissociative symptoms. Specifically, when hypnotic techniques are used as a way to facilitate exposure and work through the painful memories of the trauma in a controlled way, the frequency and intensity of PTSD symptoms often appear to decrease. The clinical use of hypnosis as a treatment of trauma was first formally described in the late 19th century by Pierre Janet, and during World War II clinical hypnotic techniques were highly successful in treating soldiers suffering from “combat neurosis.” Since that time, use of clinical hypnosis has continued, both alone and as an adjunct to other psychodynamic and CBTs, as a treatment of acute stress disorder, PTSD, and complex PTSD.

Randomized controlled studies of hypnosis and acute stress disorder, PTSD, and complex PTSD are much less prevalent in the research literature than are case studies. One clinical trial found hypnosis alone to be just as efficacious in the treatment of PTSD as psychodynamic psychotherapy and more efficacious than wait-list controls. With respect to simple PTSD, Daniel Brom and coworkers suggested that different treatments were efficacious for different PTSD symptoms: Whereas hypnosis and desensitization were more efficacious for intrusion symptoms, psychodynamic therapy was more useful for avoidance symptoms.

A meta-analysis by Jeffery Sherman, including the study by Brom and coworkers, as well as others, concluded that the beneficial effects of hypnosis for PTSD might be seen more at follow-up than immediately after treatment termination. In a more recent randomized, controlled trial, hypnosis was added to brief CBT in the treatment of acute stress disorder. The findings suggested that CBT combined with hypnosis was more efficacious in preventing re-experiencing symptoms immediately after

treatment than CBT alone. However, no differences were found between the CBT alone and CBT-plus-hypnosis groups in terms of prevention of PTSD at the 2-year follow-up: The two groups had a similarly high rate of prevention of PTSD compared to supportive counseling. Manualized hypnotic ego state therapy has also been used as an alternative to CBT for simple trauma. Because hypnosis has been demonstrated to be effective when used as an adjunct to many CBT and psychodynamic treatments for various disorders, many researchers contend that there should be more systematic research in the use of hypnotic techniques for the treatment of trauma. Reinforcing this recommendation, many case studies have demonstrated the effectiveness of hypnosis in treating PTSD-like symptoms such as dissociation, sleep and dream alterations, and pain, further justifying investigation into these techniques.

Aside from the psychiatric sequelae of a traumatic event, there is often a co-occurring physical trauma as well, and physical recovery and pain management are often an integral part of the overall recovery process. Although medical care is often essential after the trauma, the medical care received in hospital settings can sometimes be viewed by patients as an extension of the traumatic experience. Similar to other disorders, patients experiencing pain after a trauma are not motivated and are less likely to respond to treatments of acute stress disorder and PTSD prior to the reduction in their physical discomfort. For such patients, pain control is imperative. Complicating the picture is the possibility that their pain is a defense against the trauma. Hypnosis has been demonstrated to be a clinically useful adjunctive tool in the treatment of pain, both in trauma and nontrauma patients, and can often be used as a successful analgesic for phantom limb pain that would normally be difficult to treat. Furthermore, on recognizing that hypnosis can treat some of the physical symptoms of pain, patients may be more amenable to the use of hypnosis for related trauma residuals, including PTSD.

When used appropriately as a treatment of simple PTSD symptoms, there has been no evidence that hypnotic abreactions retraumatized patients or caused psychotic breaks. Similarly, not only did hypnotic analgesia cause no negative side effects, but it was also demonstrated to produce a multitude of benefits not exclusively related to pain reduction, including increased control over pain, increased sense of relaxation and well-being, decreased stress, and increased positive mood states. Thus, hypnosis might prove to be a better option for highly hypnotizable patients than more “traditional” treatments of pain and PTSD-like symptoms, especially if there is a high risk of medical complications and side effects from the traditional treatments.

Although levels of hypnotizability vary to a large extent among the general population, researchers have found that patients with acute stress disorder, as well as those developing subsequent PTSD, appear to be more

hypnotizable than trauma survivors who did not develop acute stress disorder or PTSD. Several studies with standardized hypnotizability scales have suggested that individuals suffering from posttraumatic symptoms tend to be more highly hypnotizable than most other clinical and/or nonclinical groups. However, other studies have suggested that the nature and duration of the trauma play a role in the level of hypnotizability, such that hypnotizability was more highly correlated with those who experienced repeated, as opposed to isolated, traumas.

These results were bolstered by similar research findings that chronic dissociation is more likely to occur in the context of repeated traumas than in the context of single traumatic events.

Military combat soldiers seem to be at a particular risk of dissociation during traumatic combat situations. Charles Marmar posited that combat soldiers often protect themselves from the full psychological effect of traumatic events by engaging in "peritraumatic dissociation," which allows them to distance themselves from an event while still staying engaged in a peripheral way. In a subsequent study, J. Douglas Bremner and Elizabeth Brett found that patients with PTSD reported higher levels of dissociation during combat-related traumas than did patients with non-PTSD. In fact, when compared to those experiencing noncombat traumas, soldiers citing combat as their worst traumatic experience were more likely to have lifetime PTSD and unresolved PTSD symptoms and were also more likely to be unemployed, fired, divorced, and physically abusive to their spouses. There has been recent significant work at Walter Reed Army Medical Center (WRAMC) and after the move to and merger with Bethesda Naval, Walter Reed National Military Medical Center (WRNMMC) using hypnosis in the treatment of trauma-related symptomatology in medically hospitalized service members returning from Operation Iraqi Freedom (OIF) and Operation Enduring Freedom (OEF). Harold Wain and his associates on the Psychiatry Consultation Liaison Service described many techniques and approaches in attempting to decrease psychiatric comorbidity in physically and emotionally traumatized soldiers. They stressed the need to develop a therapeutic alliance with the patient and then to use hypnotic techniques initially as a way of facilitating stabilization of the patient. Ego strengthening and reframing techniques are used, along with methods to enhance the patient's assets and help to support ego defenses, using a procedure called empathic exposure, in which the patient is asked to recall the trauma and then his or her assets are reinforced. Hypnotic recall through age regression or the affect bridge technique is used to help a patient to learn to reframe and master trauma-related symptoms. The affect bridge may be used as the patients are recovering and symptoms remain; they are asked to recall an earlier like feeling while under hypnosis. The affect and cognitions are processed and worked through.

If the experience becomes too overwhelming, the patient is taught to go to a safe place or to use the split-screen technique described by Herbert and David Spiegel. Use of these techniques does not preclude other psychological or pharmacological interventions.

A 42-year-old married mother of three children had been kidnapped and locked in a large packing trunk. After she had freed herself and broken out, her abductors had stabbed her multiple times, tied her up, put her back in the trunk, and thrown her down a cliff. She had eventually managed to break out and crawl to safety. Eventually she had been picked up by a passerby. She reported that others had seen her lying on the road and appeared frightened to approach her. Eventually 911 had been called and she had been transported to a hospital. Following medical stabilization, she had been discharged and found herself developing nightmares, re-experiencing avoidance, and having hyperarousal symptoms. She was referred by her internist for treatment and was initially started on 25 mg of sertraline, which was increased to 50 mg 4 days later. She was evaluated on the HIP and determined to be a mid- to high-range hypnotic subject. She was taught to go to a safe place and to use a split-screen technique. She was also given permission to describe her nightmares, re-experiences, and overwhelming anxieties and fears that she faced while being captive, as well as her feelings of abandonment while lying on the road. She was reinforced for her ingenuity in breaking out of the trunk. Her feeling of blame for her capture was reframed while she was under hypnosis. She was taught to calm herself and to reframe her negative feelings about her helplessness. Hypnotic age regression was used to help her master to her experiences and facilitate their becoming like a bad movie. Initially her startle response was used as a signal for her to go to her comfort zone. Age progression was used to help her to rehearse the future. The treatment used the milieu of hypnosis along with exposure, cognitive reframing, psychodynamic approaches, and pharmacology.

Although there is a paucity of randomized, controlled trials studying the effect of hypnotic techniques in the treatment of trauma, leading clinicians and researchers in the field have nonetheless outlined a strong rationale for its use. According to Etzel Cardena, Onno van der Hart, and David Spiegel, there are many solid reasons why hypnosis should be used as an adjunct in the treatment of simple PTSD victims. First, the highly hypnotizable nature of many patients with PTSD allows for greater effectiveness of hypnotic techniques. Similarly, because many patients with PTSD suffer from dissociation, they can recollect, reintegrate, and work through varying parts of their experiences while in a controlled setting under hypnosis.

It is known that some people who have experience a traumatic event may have distorted, fragmented memories of it, and under certain circumstances, as in cases of full dissociative amnesia, the use of hypnotic dissociation could facilitate access to previously forgotten memories or events. Finally, although hypnosis might not be a treatment in and of itself, it is a medium that can expedite treatment and can easily be added to many of the leading interventions. Research has further demonstrated that hypnosis can have a synergistic effect when added to previously established therapies.

With respect to complex PTSD, the standard of care is phase-oriented trauma treatment (POTT), reviewed extensively in Daniel Brown and coworkers' *Memory, Trauma Treatment, and the Law*. Most experts

concur that the three main phases of treatment are stabilization, memory and representational integration, and postintegrative normalization.

Because processing trauma as a sole treatment focus in complex PTSD can have a disorganization effect, most experts concur that hypnosis helps in the development of a variety of trauma-specific skills and coping strategies, as well as in facilitating the developmental repair of the sense of self, attachment representations, and affect regulatory structures *prior* to trauma processing, as part of a multifocused stabilization phase. Guidelines for the assessment and treatment of complex PTSD, as well a review of outcome studies, can be found in Chris Courtois and Julian Ford's *Complex Traumatic Stress Disorders*.

Hypnosis and DSM-5 Somatic Symptom and Related Disorders

Hypnosis in the evaluation and treatment of somatic symptoms as defined in DSM-5 is a broad category that has not yet been well studied with evidence-based trials, with the exception of conversion disorders where there are many well-documented case reports. Albeit, there are many well documented case reports. The association between conversion disorder and hypnosis was first suggested by Pierre Janet in 1907 as a form of autohypnosis. More recent clinicians have also suggested that conversion symptoms may result from spontaneous self-hypnosis involving a dissociation of sensory-motor function in reaction to traumatic events or to other nontraumatic internal conflicts. In a sized sample of patients Wain, while working on a consult liaison service, observed that patients who exhibit conversion disorder symptoms showed a significantly higher capacity for hypnosis. Several Dutch studies demonstrated similar findings. Experts on somatic symptom disorders concur that hypnosis may therefore be easily used to the clinician's advantage to both diagnose and treat the disorder. Wain often uses his modification of the HIP to ascertain hypnotizability in medical patients who have physical symptoms lacking physical finding. The eye roll measurement fits well into the medical model.

Charcot, Breuer, and Freud in the late 19th and early 20th centuries contributed greatly to the development of the concept of conversion disorder. In the first edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-I) conversion disorder was called a conversion reaction, in the second edition it was called hysterical neurosis, and in the third and fourth editions the conversion process was restricted to symptoms affecting the voluntary motor and sensory nervous systems. Motor conversion symptoms include nonepileptic seizures, contractures, paralysis, and stance and gait disturbances. Sensory-based conversion includes focal blindness, elective muteness, and conversion pain.

Symptoms for which there was some physiological understanding (generally involving the automatic nervous system) were subsumed under the psychophysiological disorders. A broadened definition of conversion

disorder was used in the DSM-III, which was characterized by symptoms involving any "loss or alteration in physical functioning suggesting a physical disorder" as long as the mechanism of conversion was evident, that is, the symptom was "an expression of a psychological conflict or need." In the DSM-IV and DSM-5, conversion disorder was again restricted to symptoms affecting the voluntary motor and sensory systems and where there was incompatibility between symptom and neurological or medical conditions.

All somatic symptom disorders share in common and are characterized by physical complaints lacking known physical basis or demonstrable physical findings where there is significant distress and impairment. These disorders are conditions characterized by physical symptoms suggestive of but not fully explained by a general medical condition or the direct effects of a substance. Symptoms are not intentionally produced and are not attributable to another mental disorder and cause distress or impairment in important areas of functioning. The disorders include illness anxiety disorder, conversion disorder, psychological factors affecting other medical conditions, factitious disorders, other specified somatic symptoms and related disorders, and unspecified somatic symptoms somatic symptoms and related disorders.

The somatic symptoms may transiently help to alleviate the internal conflict and primary gain is established. Secondary gain may also occur, allowing the patient external reinforcement by escaping from the situation, adopting a sick role, and so on.

Once a medical condition has been carefully ruled out, hypnosis can be an effective way to uncover the traumatic or nontraumatic conflict that precipitated the conversion symptom. Exploratory hypnotic techniques, such as age regression and affect bridge, are generally used in helping the clinician and patient to identify and understand the nature of the conflict.

Procedurally, one would use trance phenomena to take a patient to a safe place prior to going to the conflictual area. This establishes a safety valve for the patient and models a way of moving away from the conflict when anxiety becomes overwhelming. Thus reassured, the patient can face the conflict but transiently find safety without having a dystonic symptom. It further can expedite the attainment of insight, facilitate mastery, and allow for a more rapid working through of the conflict. Helping the patient to find and understand the conflict that perpetuated the need for the symptoms can be expedited. The use of age progression can then also help patients to see themselves facing a similar conflict in the future with the added benefit of having a rehearsal to facilitate success and a more appropriate way of responding.

While the patient is resolving the conflict, hypnotic techniques of symptom substitution can be used that are more appropriate—for example, shaking fingers versus having an entire body tremor can be extremely

useful.

Extinguishing the symptom without exploring the reason for its presence at times can enhance the level of anxiety and/or depression and lead to other potential acting-out behavior.

Hypnocognitive methods are also used to modify learned behavior. Remembering the symptom is beneficial for one's protection and its ablation can be destructive to one's homeostasis.

A 22-year-old male patient was brought to the emergency room with bilateral blindness. Following an evaluation by ophthalmology, it was determined that the blindness was psychogenic. After initial evaluation by psychiatry, a therapeutic alliance was developed, and hypnosis was used to take the patient to a safe place and then back to the time immediately prior to the blindness. After two sessions the patient was able to describe seeing his wife in an adulterous relationship. At that moment the patient vocalized a desire to harm his wife and her suitor. Immediately after this vocalization, he became amnesic for the event and blind. On describing this under hypnosis, he was given a suggestion that when he became alert "He would only remember what he felt comfortable remembering." Subsequent to the patient becoming alert, he had no idea what had occurred, and each day after the hypnotic intervention was initiated the patient's anger was reframed. When the patient felt comfortable he then confronted his wife. The patient became aware that the amnesia was being used to prevent him from acting out. Use of a psychodynamic, cognitive reframing approach with a hypnotic milieu helped this patient to gain control and understanding of his symptoms. The patient and his wife were then referred for marital counseling.

A 30-year-old male with a history of bipolar disorder who presented to neurology for a worsening motor tics which included rhythmic flexion of the left wrist and hand, eye blinking and grimacing, left sided lower extremity repetitive kicking movements, and repeated jerking of the neck and shoulder left upper extremity tremor. The patient underwent a full neurological work-up which was negative and a trial of medications was without result.

After evaluation the patient was screened for hypnotic capacity and was determined to be a midrange subject. He was taught to be sensitive to his physical cues and utilize a hypnotic intervention to curtail symptoms. Following the first session the patient had an immediate reduction in the frequency and amplitude of his tremor. He then practiced the hypnotic technique and was motivated to understand the significance and etiology and gain greater control of the symptom. With a psychodynamic approach the patient recognized the tic was his defense against the anxiety of being shot. And the tremors were his way as a child of expressing his fear responses especially when he perceived his father or other adults might hurt him. Following treatment the patient was able to control his symptoms and the initiation of the hypnosis gave him a sense of confidence to explore more of the psychological conflicts that he had faced.

A 29-year-old woman was referred for evaluation and treatment of ongoing facial pain that was not responding to traditional methods of intervention. Neurological evaluation showed no objective physical correlations. Her high mid-range performance on the HIP added support to the potential of a psychological mechanism for the pain. Initially the pain was controlled by a hypnotic intervention, but it returned 24 hours later. Her self-hypnotic technique ceased to be effective. A decision was made to explore more completely the meaning of the pain. Age regression under hypnosis was used, and the patient was regressed to a time prior to the pain.

She related that her brother had been injured by a car while he was running in the street. The patient was babysitting at the time, and her father was so angered that he hit her. Recently her friend's dog ran away, and she felt responsible. As she began to recognize her need to punish herself because of her guilt over what had occurred, she was able to understand her feelings and reframe her thoughts in a more productive manner. An "affect bridge" was also

used, and the patient was asked to go to back to a previous time when she fell guilty and was punished. She then was able to describe her feelings of being hit by her alcoholic, abusive father. She continued to gain insight and mastery over the past and was able to ablate her pain.

The foregoing vignettes demonstrate a variety of somatic disorder cases where hypnotic techniques were used to help to resolve the patients gain control over the symptom and resolve the conflicts. They further demonstrate how hypnosis can be used as adjunctive psychological therapy together with reliance on pharmacotherapy, CBT, psychodynamic therapy, and/or eye movement-desensitization-reprocessing (EMDR). Wain advocates the need for the clinician to be viewed as safe and to facilitate the development of trust and a therapeutic alliance to enhance the success of treatment. Hypnotic intervention can be introduced after these needs have been met. The clinician further needs to avoid being impatient and to recognize the consequences of immediate demands for the resolution of the conflict.

Hypnotherapy and Relational Disturbances

There are two broad categories of relational problems that patients bring to psychotherapy—problems *with* relationships, or attachment pathology, and problems *within* relationships, or focal, recurrent conflictual relational themes played out in most relationships—called core conflict relational themes (CCRTs). Hypnosis is useful for both types of relational disturbance.

Attachment disturbance is assessed in children between 12 and 20 months using the Strange Situation Paradigm, which entails direct observation of an infant and mother in an unfamiliar play room. Children are classified into four attachment types—secure, avoidant (dismissing in adults), resistant (anxious preoccupied in adults), and disorganized. Longitudinal studies have shown that these attachment types are relatively stable and predict preschool, adolescent, and adult attachment behavior.

Base rate studies consistently have shown that about 65 percent of the general population has secure attachment, and about one-third of adults in the general population in contemporary Western culture have one or another of the three forms of insecure attachment. Use of the Adult Attachment Inventory shows that anxious preoccupied attachment or disorganized attachment is strongly associated with borderline personality disorder. Disorganized attachment is associated with a range of psychopathology. Disorganized attachment combined with physical and/or sexual abuse in later childhood is associated with the development of dissociative disorders in adulthood.

Psychodynamic psychotherapy with personality-disordered patients with attachment pathology is based on the interpretation of defenses against attachment along with the provision of a stable, consistent

therapeutic attachment relationship that serves as the basis of the development of new internal working models for secure attachment. However, research with the Strange Situation Paradigm in infants shows that it is the *representation* of attachment between 12 and 20 months when symbolic or representational thinking develops, not an attachment relationship per se, that has an organizing effect on lines of development like affect regulation and self-development.

Therefore, hypnotherapeutic imagery focused on the development of positive representations of secure attachment more directly addresses the central problem of attachment pathology, namely the failure to develop positive stable internal representations of a secure base. Using an object relations model, Elgan Baker first articulated the use of hypnotic imagery to visualize the therapist as a "good object" and then to visualize interactions between the patient and therapist in which the therapist is seen a nurturing, supportive figure. Daniel Brown's hypnotherapeutic approach to treating attachment pathology follows John Bowlby's model that the establishment of a secure based is naturally followed by exploration. Structured hypnotic visualizations of ideal parent figures are used to develop new, positive internal representations of secure attachment. An outcome study is underway using this protocol with personality- and dissociative-disordered patients with the pre/post scoring of the Adult Attachment Inventory as the main measure of change in attachment status.

CCRT maps are more complex. They develop between the third and fourth year of life at a time that the child has developed the capacity for emotional ideas, wishes, complex schemas and beliefs, and psychological defenses. Whereas attachment representations develop from 12 to 20 months at an age prior to the development of narrative memory, CCRT maps develop concurrent with the development of narrative memory and are therefore subject to interpretation in dynamic psychotherapy.

The typical approach to modifying CCRT themes through dynamic psychotherapy takes the following steps:

- Formulating the CCRT accurately in terms of the underlying wish of what the given patient is looking for in each relationship, the negative reactions of others expected, and the maladaptive reactions of the self in response to not having this wish fulfilled.
- Making a test interpretation so that the patient comes to see how this pattern is operative in most relationships—past and current—and in the relationship with the therapist.
- Interpreting the manifestation of the CCRT elements in the here-and-now transference to the therapist. By adopting an emotionally corrective stance, the therapist tries to alter the structure of the CCRT. Some methods are designed to change the reactions of others, whereas others

focus on changing the reactions of the self. Outcomes studies show that reactions of others and reactions of the self-elements change in roughly 30 to 50 hours of dynamic psychotherapy that consistently focuses on CCRT themes.

Hypnotherapy can play a useful role in CCRT treatment. Hypnoprojective methods can assist in formulating the CCRT ("The scene will somehow be about the kind of problems that keep coming up in relationships"). Structured hypnotic visualizations can assist in having the patient articulate the best aspects of the central wish and then imagine ideal reactions of others and healthy adaptive reactions of the self toward the goal of developing a new, positive CCRT internal map ("The scene will be about the kind of relationship you wish for, in the best sense," "The scene will be about the ways others respond, in the best sense"). There are no outcomes studies specifically on the use of hypnotherapy for CCRT problems.

Carolyn Daitch (2014) has developed a short-term hypnotherapeutic protocol for anxiety in the context of romantic partnerships.

Hypnosis and Anxiety Disorders

Anxiety disorders can be classified in terms of whether the anxious arousal is more or less continuous, as in GAD and OCD, or discontinuous or episodic. Disorders of discontinuous anxiety can be further classified in terms of whether the source of the anxiety is external, as in specific phobia, or internal, as in panic. Unwanted thoughts are also characteristic of continuous anxiety conditions, with excessive worry about realistic situations being characteristic of GAD and misappraisal of the importance of unwanted, senseless obsessional thoughts being characteristic of OCD. Learned progressive avoidance of unwanted, senseless thoughts makes such thoughts salient and persistent and tags these thoughts with excessive anxiety, thereby justifying the progressive use of compulsive-behavioral and mental rituals to mitigate the anxious arousal. Learned progressive avoidance of discomforting bodily sensations and/or unfamiliar situations loosely associated with an increased probability of having panic attacks gradually transforms panic attacks in agoraphobia.

Furthermore, while it is well established that exposure-based treatments are effective, each type of anxiety disorder requires a specific type of exposure treatment. For example, because excessive worry generates continuous arousal in GAD, worry exposure in the best-matched treatment. Because discomforting bodily sensations generate panic, exposure to such uncomfortable sensations is critical to successful treatment of panic, as exemplified by fear-of-fear exposure, interoceptive exposure, and panic control therapy (PCT). Because fear of external objects or situations is central to the development of specific phobias, exposure to

the feared situations in fantasy and in real life through systematic desensitization is critical to the treatment of phobias.

Because compulsive rituals eradicate anxiety, behavioral exposure and response prevention (ERP) aids the patient to tolerate the anxiety and master it without resorting to the compulsion. ERP alone, however, does not directly address the obsessive thinking. Restructuring core beliefs that cause the misappraisal of senseless, unwanted obsessive thoughts is the best approach for working with the obsessive-thinking component of OCD.

Most clinicians concur that hypnosis is not a treatment; it is the medium of treatment. Hypnosis is merely a state of redistributed attention and heightened focus and, for some patients, an altered state of consciousness or trance that allows greater access to imagery, self-talk, affects, self-states, and inner coping resources than would be available in the normal waking state. However, it is still necessary to bring to hypnotherapy the protocol of treatment that is best matched to the given psychiatric condition. With respect to anxiety conditions, exposure-based cognitive-behavioral treatment protocols consistently show favorable outcomes. Hypnotherapy can serve as a useful adjunct to the cognitive-behavioral treatment of anxiety disorders. J. Christopher Clarke and Arthur Jackson devoted an entire text to hypnotherapeutic treatment of anxiety disorders. With respect to specific phobias, hypnotherapy is a useful adjunct to desensitization treatment. Sharon B. Spiegel has provided an excellent review of the current issues in the treatment of a specific phobia and has made significant recommendations for innovative applications of hypnosis in the treatment of specific phobias. She fully describes the benefits of utilizing hypnosis in the treatment of specific phobias, and she has emphasized that hypnosis does significantly manage resistance to treatment in specific phobias. She has documented that the use of hypnosis does reduce the length of treatment, and she has demonstrated that hypnosis does provide substantial reduction in anxiety, based upon the significant advances in cognitive neuroscience. The therapist and patient construct a stimulus hierarchy from least- to most-feared situations. Exposure consists of imaginal exposure to each situation on the hierarchy, from least to most, first in fantasy and then in real life. Hypnotic desensitization has advantages over nonhypnotic behavior exposure, in that hypnotic treatment entails greater focus, less distraction, richer affective imagery, a greater capacity to control the exposure imagery, enhanced sense of learning and mastery through time distortion suggestions, and greater likelihood of facilitating effective real-life exposure through posthypnotic reinforcement.

Comparative outcome studies of nonhypnotic behavioral exposure and hypnotic desensitization illustrates that both approaches have comparable treatment effect sizes but that the hypnotic treatment is typically faster for the aforementioned reasons.

With respect to panic, the first proven exposure-based treatment was interoceptive exposure. David Barlow and his associates reasoned that because the escalation of panic states starts with the misappraisal of uncomfortable bodily sensations, exposure to uncomfortable bodily sensation was a core ingredient in successful treatment. In interoception exposure, the patient is given instructions to produce analogues to the kind of bodily discomfort operative in panic states, such as twirling the head around rapidly until becoming dizzy, running in place until becoming winded, or holding the breath for longer periods until becoming uncomfortable. The rationale of interoceptive exposure is that the patient learns to sit through and master uncomfortable bodily states. Barlow more recently developed a more comprehensive treatment package for panic disorder—PCT—which includes interoceptive exposure, cognitive reframing of panic escalating cognitions, and controlled breathing. Because panic is a rapidly escalating state exacerbated by hyperventilatory breathing, learned control over breathing serves to de-escalate the panic state. Hypnotherapy can serve as a viable alternative to PCT. Interoceptive exposure to uncomfortable bodily sensations is an indirect form of exposure because the bodily discomfort generated by head twirling or breath holding is not equivalent to the bodily discomfort naturally occurring during panic states. In contrast, hypnotherapeutic exposure using a fear-of-fear hierarchy is more direct. The clinician constructs a fear-of-fear hierarchy of least to most intensity of escalating panic experience.

In hypnosis, the therapist suggests the progressive intensification of the panic state from very mild, to moderate, to full intensity and then instructs the patient to sit through the full discomfort until it diminishes and mastery over the discomfort is attained. The hypnotherapeutic fear-of-fear exposure more closely matches the sequence of cognitive and somatic events operative in the development of an actual panic state than does interoceptive exposure.

Another interesting disorder vocal cord dyskinesia (VCD) in the past has often been diagnosed as a conversion disorder rather than under the anxiety spectrum. This disorder is often known to be an asthma simulator and is characterized by upper airway obstruction and stridor without an anatomic cause. The disorder is highlighted by inappropriate inspiratory vocal cord adduction seen on fiber optic laryngoscopy. This can be distinguished from bilateral vocal cord paralysis by episodic, momentary abduction of the vocal folds during cough or patient distraction. Arterial blood gases values remain normal. These patients rarely exhibit hypoxia or hypercarbia. In addition, chest radiographs do not show hyperinflation or other immune functioning as is characteristic of asthmatics.

A 26-year-old female with a dx of VCD was referred from pulmonology for evaluation and treatment. She described an incident of sexual abuse at age 16 and had no previous symptoms

or recurring thought intrusions or nightmares. She relates her present symptoms began after running and then beginning to gasp for air. Initially the patient is diagnosed with asthma and antiasthmatics are prescribed but all treatments she was refractory to and the symptoms persist. Following laryngoscopy a diagnoses of VCD is determined development.

The patient reported a successful college and vocational experiences and initially denied any psychological concerns. During the initial evaluation a HIP screening is utilized and she is a mid-range to high subject. As the arm levitates in the post hypnotic section she then looks with amazement at the in voluntarism of the arm and she is given the interpretation that her arm is responding the way other parts of her body can respond as well.

She is then able to recognize what may be occurring with her larynx. She is taught a self-hypnotic procedure to gain control over her symptoms. Whenever she feels tightness in her chest and she is instructed to do herself hypnotic procedure. In later sessions she begins to acknowledge that she is feeling choked by the stress in her present life and recalls a similar sensation when she was choked during a past abuse episode. She begins to acknowledge the feeling of having been held down and overwhelmed. As frequently occurs she begins to discuss more of her past and present trauma and is able to reframe and gain control. An affect bridge is used to help her master her feelings of the past as well as the present. As the intervention progresses the patient slowly begins to acknowledge her difficulty with her present position and her desire to terminate her job. However, she feels compelled to remain in the position because of her contract. She gains motivation to be able to protect herself and her symptoms begin to dissipate. One year after her last treatment there still were no return of symptoms and she was able to regain control. Wain and his colleagues at WRNMMC utilize this approach with many patients with VCD who are referred for treatment.

After a pulmonary evaluation and a therapeutic alliance are developed, we initially teach them symptom control then help them understand the underlying conflict.

With respect to GAD, worry exposure directly addresses the excessive worry that generates sympathetic arousal. The patient first constructs a list of favorite categories of worry. The patient is instructed to intentionally worry about something common on the list, then to think about something worse, then something worse, and so on. Worry exposure is systematic worst-scenario thinking. After awhile the content of worst scenario thinking becomes so extreme that it fails to generate physiological arousal, and the link between excessive worry and physiological arousal is broken. There are a number of hypnotic approaches to GAD. Suggestions for relaxation, cue-induced soothing imagery, and the clenched-fist technique address the somatic arousal characteristic of GAD. Direct ego-strengthening suggestions to experience a worry-free state and/or hypnotically assisted worry exposure address the excessive worry characteristic of GAD.

With respect to the treatment of specific phobia, hypnotizability is a strong predictor of hypnotic treatment of flying phobias. Hypnosis reduces the over length of treatment, optimizes relaxation in exposure-based treatments, and builds self-regulation skills in the treatment of phobias.

With respect to OCD, hypnosis can aid ERP. Hypnotic suggestions are given to intensify and then tolerate the anxiety associated with ritual compulsive behaviors such as checking or washing until a marked reduction in the frequency of compulsions is achieved. Modification of core beliefs that drive the misappraisals of senseless, unwanted thoughts and transform them into obsessions is best accomplished by hypnotic ego state

therapy.

For example, if the patient reports a core belief of excessive responsibility, ego state therapy is used to identify the adaptive function of the overly responsible "part of the mind." Then suggestions are given to contact the part of the mind that is most opposed to the overly responsible part. By accessing that counterpart to the overly responsible part, cognitive dissonance is created. For example, a spontaneous playful part of the mind might be identified as a counterpart to the overly responsible part. Having heightened the dissonance by calling forth the spontaneous, playful part, the therapist uses conflict resolution strategies to reduce the dissonance between the overly responsible and spontaneous parts. Other parts are called on to offer a solution to the conflict between the overly responsible and spontaneous parts. To reduce the dissonance, the system typically shifts away from the extreme position of the overly responsible part. A similar approach is used for modifying any of the other core beliefs operative in maintaining the OCD symptoms until a marked reduction in unwanted obsessive thoughts is reported.

Hypnosis and Depression

Depressive disorders are among the most frequent referrals in psychiatry. The National Institute of Mental Health (NIMH) outcome study on depression showed that antidepressant medication and interpersonal therapy were equally effective, and CBT was modestly effective, in the treatment of major depressive disorder. Assen Alladin studied hypnotherapy combined with CBT as compared to CBT alone in treating depression in a 16-week randomized clinical trial. He found that although both groups significantly improved, the treatment effect size was significantly greater in the hypnosis group, and that hypnosis had an additive effect on traditional CBT treatment of depression. Miaozun Shih and associates conducted a meta-analysis of hypnotic treatment of depression showed a clinical significant effect of hypnotherapeutic treatment that significantly improved symptoms. In hypnosis, depressed patients can be given ego-strengthening suggestions for increased confidence and control, suggestions for positive affective states, and rehearsal of positive, hopeful future events to help the patient to come out of the hole in which he or she has fallen. Severe depression sometimes can interfere with concentration, thereby challenging the clinician to keep the patient focused on developing positive states and preventing a sense of failure. Hypnosis can be used as an adjunct to pharmacological approaches with depressed patients.

Hypnosis and Personality Disorders

The *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*

(DSM-5) defines a personality disorder as follows: A personality disorder is an enduring pattern of inner experience and behavior that deviates markedly from the expectations of the individual's culture, is pervasive and inflexible, has an onset in adolescence or early adulthood, is stable over time, and leads to distress or impairment.

A variety of nonhypnotic approaches to the treatment of personality disorders have evolved, ranging in diversity from behaviorally oriented DBT, cognitively oriented schema therapy, and dynamically oriented transference-focused psychotherapy (TFP) to developmentally informed psychotherapy that stresses attachment repair mentalization, self-development, and affective development. Although Marsha Linehan's DBT enjoys great popularity in the United States, outcome studies merely show that patients with personality disorders treated with DBT tend to stay in therapy and have less self-destructive tendencies over time relative to patients in conventional therapy, but no data are available to document whether DBT leads to the clinically effective result of no longer meeting a personality disorder diagnosis. Likewise, Otto Kernberg's TFP shows some promise in diminishing personality disorder criteria, but in comparative outcome studies developmentally informed treatments like Peter Fonagy's mentalization-based treatment (MBT) and cognitive-based treatments like schema therapy consistently show better outcomes than DBT or TFP in terms of numbers of patients who no longer meet sufficient criteria for a personality disorder posttreatment. More recently, outcome measures have included not only personality disorder diagnostic criteria, but also developmentally informed outcomes, such as coherence of state of mind and metacognitive capacity. Such measures are used in the attempt to demonstrate that patients with personality disorder are actually changing the organization of their state of mind in some fundamental way through treatment.

Hypnotherapy holds greatest promise as an adjunct to cognitively based approaches like Jeffrey Young's schema therapy and developmentally informed treatments. The Young Schema Questionnaire serves to identify on which of 18 maladaptive schema domains a given patient with personality disorder measures high or very high. Treatment then targets each of these areas using both traditionally cognitively focused methods, such as testing the evidence and correcting distortions, and emotion-focused methods. The main emotion-focused method entails a reparenting approach in which the therapist adopts an emotionally corrective stance. For example, if the patient characteristically expects abandonment, the patient is asked to imagine the therapist as being fully there, or if the patient expects deprivation, the patient is asked to imagine the therapist as providing the kind of care that is most needed. This emotion-focused technique is conducted in a relaxed state with structured waking imagery. A recent comparative outcome study showed that twice as many patients with

personality disorder in schema therapy improved as did matched patients with personality disorder in TFP. Young believed that a main component of the success of schema therapy was the emotion-focused reparenting imagery work.

Although Young did not use hypnosis in schema therapy, there is no reason that his emotion-focused technique could not be adapted to using hypnotic imagery. There may be several advantages using hypnosis, especially with hypnotizable patients with personality disorder. Although patients with personality disorder have a variable self-observational capacity, the hypnotic context may provide a necessary ingredient within which such patients can better resist distract, sustain focus on the emotion-focused imagery, and more quickly and efficiently alter the target maladaptive schema.

Hypnotherapy can also serve as a useful adjunct to a developmentally informed approach to the treatment of personality disorders. The work of Elgan Baker and of Daniel Brown illustrates how structured hypnotic imagery can be used to develop positive, stable internal representations of secure attachment. With respect to self-development, Janine Stevenson and Russell Meares obtained favorable results in reducing a personality disorder diagnosis using Fred Pine's approach to fostering self-development. Hypnotherapy can also be used to foster self-development. The normal line of self-development includes the development of a cohesive self-representation (self-definition), the representation of the self as having an effect on the physical environment and on unfolding events and on eliciting desired responses from others (self-agency), and the linking of positive affective states to the self-representation (self-esteem). Patients with personality disorder with deficits along the line of self-development, such as borderline patients, show deficits in self-definition and agency, and those such as narcissistic patients show deficits in self-esteem development. Repeated structured visualizations are introduced in hypnosis to foster self-definition ("the scene will be about . . . when you are most you . . . your real self . . . the unique qualities that manifest when you are being most yourself"); self-agency ("the scene will be about . . . making an impact on . . . eliciting the kind of responses from others that you need"); and self-esteem ("the scene will be about . . . some situation in which you feel especially good about yourself . . . especially like yourself . . . feel especially confident"). Hypnotic suggestions are used to generate the target scene, amplify the positive affect associated with the scene, extend the duration of experiencing the target state of mind, and then transferring that state of mind to situations in which it would normally feel difficult to hold that sense of self. The structured visualizations are repeated in treatment sessions and outside of the sessions in self-hypnosis until they foster a stable, positive internal representation of self-definition, self-agency, and/or self-esteem.

With respect to affective development, hypnotherapy can be used to link the visceral substrate of affective states to imagery and thought in patients with personality disorder who show developmental deficits in the capacity for affect experience. Hypnotically suggested soothing imagery and/or hypnotically structured exposure (hypnotically suggested progressive intensification of affective states so that the patient learns to sit through them) can be used in patients with personality disorder who show deficits in affect tolerance. There are no outcome data on developmentally informed hypnotherapeutic approaches to the treatment of personality disorders.

HYPNOSIS AND PAIN

An excellent review of controlled studies and anecdotal reports concerning research and clinical applications of hypnosis in treating clinical pain. David R. Patterson and Mark P. Jensen's excellent review of evidence-based studies of the use of hypnosis in assessment and treatment of pain disorders. Since the publication of that review, there have been other significant studies documenting that hypnosis has become an important method of intervention in the treatment of acute pain disorders, pain during medical procedures, surgical procedures, and chronic pain disorders.

Hypnosis in Acute Pain Disorders—Procedural Pain Interventions

Over the last 10 years, there has been increasing use of hypnotic interventions in the treatment of acute pain disorders and the alleviation of pain during medical and surgical procedures. There have been evidence-based studies document the efficacy of hypnotic interventions in these contexts, and these studies correspond to research findings regarding hypnotic analgesia in the hypnosis laboratory setting.

Hypnosis and Burn Care. Hypnosis has been used in the analgesic management of burns. These hypnotic interventions include suggestions with progressive muscle relaxation, followed by suggestions for direct analgesia. Imagery for cooling the inflammation and suggestions for reappraisal of feared bodily harm also have been used. In the various reports in the literature concerning the use of hypnosis in the management of burns, the hypnotic interventions appear to have been efficacious for both debridement of burns and changes of dressing involving burns. These controlled studies have also documented that the hypnotic interventions had significant effect on pain perception, anticipatory anxiety, and relaxation during and after burn care.

In addition to Theodore Barber's rapid induction analgesia method of hypnosis, other hypnotic interventions have used eye fixation associated with progressive muscle relaxation, guided imagery, and direct suggestions

for pain relief in the affected area. Overall, both clinical reports and controlled studies have demonstrated that hypnosis is associated with reduced use of opiate medications, decreased pain ratings, and reduced use of sedative–hypnotics and tranquilizers. Many of these studies included assessment of hypnotizability using the HIP and the Stanford Hypnotic Clinical Scale (SHCS).

Hypnosis in Obstetrics and Gynecology. In view of the fact that labor pain represents a significant acute pain and that women are certainly motivated for both reduction in pain and diminished anxiety, hypnosis has been used as an intervention in the acute pain associated with labor and delivery. Another benefit reported in both anecdotal reports and controlled studies involves the reduction in stress and anxiety in the labor and delivery process. This reduction in stress and anxiety provides a better maternal environment during both stage II and stage III of labor. The literature provides numerous anecdotal reports concerning the use of hypnosis in the labor and delivery process. Ralph August and Simon Chaisson described using hypnosis as a sole anesthetic for caesarean births. Karl Fuchs in Israel described the successful use of hypnotic intervention with hyperemesis gravidarum.

Hypnosis and Women's Issues

In a study by Karl Fuchs, he demonstrated that hypnotic intervention can be very successful for hyperemesis gravidarum. Out of 51 gravidas patients treated individually, 35 (68.5 percent) showed excellent results with no vomiting and no nausea, and two (3.9 percent) showed good response and had no vomiting but some nausea. Fourteen (25.5 percent) patients showed vomiting less than six times per day and had improvement in their nausea poor response to treatment, and one patient (2.3 percent) could not be hypnotized in spite of repeated trial sessions. Out of 87 gravida patients treated in groups, excellent results were obtained with 61 patients (70.1 percent), and 24 patients (27.6 percent) had no vomiting but some nausea, and 2 patients decreased their vomiting to less than six times per day but had improvement in their nausea.

Gary Elkins and his associates, as well as Cassie Kendrick and associates, have used hypnosis in the treatment of hot flashes in premenopausal women. Maitri Shah and associates found that both hypnosis and medications were equally effective in treatment dysmenorrhea.

Hypnosis in the Emergency Room

Past myths about the nature of hypnosis may have contributed to a reluctance among many to use hypnosis in the emergency department.

Some feel that hypnosis takes too long and would not be appropriate in acute care situations. The time for formalized trance induction may range from a spontaneous occurrence to several hours. However, given the focused state of the individual in an emergency situation, hypnotic induction is frequently quicker than under other conditions. Frequently, because of the discomforts of the presenting ailment, anxiety, and confusion pertaining to the extent and significance of one's symptoms, patients in the emergency department feel as if they are out of control. However, their preoccupation with acute symptoms may decrease their peripheral awareness and thus narrow their field of concentration, enabling them to gain some control. This highly focused state is like a hypnotic trance state.

The psychological trauma associated with physiological malfunctioning often precipitates an altered state of consciousness for the patient. Awareness of the existence of this altered state (trance or dissociated state) can enable the staff to use it in treatment. In this altered state the patient may be highly receptive to comments and suggestions. The staff must be very cautious with the words and statements used within the patient's hearing. The same guidelines should be followed with patients who have not been introduced to the formal hypnotic trance. Patients, by virtue of being in a focused state, are vulnerable to staff's comments and assimilate the data rapidly and at times without vigilance. This is why a mere suggestion of help or of a positive outcome can be facilitative for a patient in general.

In the emergency department the staff typically adopts a positive, calm, supportive approach to the patient. This supportive style may provide the necessary ingredients for an enhanced therapeutic alliance, which can facilitate the patient's ability to use his or her trance state more effectively. Many individuals use their trance state without being cognitively aware. The natural trance is frequently used by individuals as a protective measure, and it typically allows individuals to distance themselves perceptually from their discomfort, whether it is physical or emotional. The astute clinician can recognize a patient who is doing this and capitalize on this state. For patients who are having difficulty focusing constructively, the clinician can help to find an appropriate focus for them.

Paul Delfito described the value of hypnosis in the emergency department for treating the pain due to orthopedic injuries, burns, corneal abrasions, abdominal distress, headaches, menstrual discomfort, renal calculi, herpetic lesions, muscle spasms, and dental syndromes.

Wain, working on a Consultation Liaison Psychiatry Service, described similar results with a variety of patients in the emergency department.

A 32-year-old man presented to the emergency department with a severe headache. He was a chronic migraine sufferer and had been unable to control the pain on this occasion with his

propranolol. The emergency department recognized that he had high hypnotic capacity. The imagery of an icepack being placed on his forehead was suggested. Initially some real ice was placed on his forehead to help. The patient was able to control his pain completely with this imagery. He did not require narcotics, as he had on previous visits. On follow-up several weeks later the patient reported being able to use this strategy to control, as well as prevent, migraine attacks, and he no longer had to rely on frequent emergency department visits for pain relief.

A 55-year-old female patient suffering from coronary artery disease presented to the emergency department with severe chest pain. Her cardiac enzymes and electrocardiogram (ECG) were completely normal. The treating clinician used hypnosis, during which the patient reported that her internist had died of a myocardial infarction the previous day. When it was clear that there was no actual cardiac involvement on this occasion, the patient was allowed to reinterpret her perception of the pain. She came to the realization that she had been identifying with her internist, and the pain was her way of grieving his loss. As she began to practice the hypnotic techniques, her pain dissipated. Although pain free, she was nevertheless observed on a monitor for 24 hours. There continued to be no evidence of myocardial damage. She was discharged pain free. She returned as an outpatient and again was symptom free. She found a new internist and was monitored regularly.

Hypnosis for Alleviation of Pain in Bone Marrow Aspirations. Studies have demonstrated the efficacy of hypnosis in reducing the pain of bone marrow aspiration using such techniques as storytelling, fantasy, imagery, and deep breathing. Control studies demonstrate that hypnotic interventions with children provided lower ratings of pain than a control group.

Some studies also used specific hypnotic interventions involving posthypnotic suggestions for re-entering hypnosis following cues provided by the hypnotherapist. The hypnotic interventions used eye fixation, as well as progressive muscle relaxation and the involvement of storytelling and guided imagery, as methods of hypnotic intervention. The use of hypnosis for bone marrow aspiration in children is associated with the administration of the SHCS for Children. Ernest Hilgard and Samuel LeBaron's study of children's hypnotizability predicted analgesic response, as well as reduction in anxiety.

Hypnosis in Surgical Procedures

There have been a number of studies concerning the use of hypnosis in surgical procedures. Such studies include hypnosis used in hemorrhoidectomies, anorectal surgery, surgical removal of pedicle attachments, neurosurgical procedures, dental surgery, breast surgery, coronary artery bypass surgery, hernia repair, and angioplasty. These studies reported efficacy and positive effect on surgery, including reduced postoperative nausea and vomiting, decreased anxiety during the procedure, fewer postoperative side effects, reduced medication use, and quicker recovery time. Similar studies also reported efficacy in dental surgery, breast surgery, coronary artery bypass surgery, endocrine surgery, cervical surgery, and hernia repair. There have also been studies that

demonstrate efficacy of hypnosis in certain gynecological procedures.

Wain published a paper on the use of surgery as a sole anesthetic in septoplasty (Fig. 33.4–2). Hypnoanesthesia is not a sole anesthetic for all patients, however—probably only for hypnotic "virtuosos." There are many more patients who can use hypnosedation as an adjunct to chemical anesthesia, thus reducing reliance on the latter and decreasing risks. In his article on hypnosis for septoplasty Wain reviewed many surgeries in which hypnosis was used as a sole anesthetic or as an adjunct to traditional anesthesia.

A 34-year-old man presented with a history of nasal obstruction for many years. He had an injury to the nose at age 12 years and had a continuous obstruction to breathing. He was found to have a grossly deviated septum and a small hump on the nose, which required removal. His physical examination revealed a well-developed, well-nourished 34-year-old man in no acute distress with normal vital signs. Physical examination was unremarkable except for the presence of a grossly deviated nasal septum to the right with total obstruction and total dislocation of the septum to the right nostril obstructing the airway, with compensatory hypertrophy of the inferior turbinates on the left side. A septoplasty was performed under hypnosis. Although no chemical anesthesia was used, an anesthesiologist was in the operating room to ensure efficacy for the patient. Postoperatively the patient was able to change his packing without any pain. The strategy consisted in suggesting that he was absorbed in floating on or near a beach. He had no recollection of the surgery and reported no perception of pain. He was highly hypnotizable.

FIGURE 33.4–2. Photographs of septoplasty.

Hypnosis in Chronic Pain Disorders

In contrast to acute pain and procedural pain, chronic pain is a persistent and intractable pain disorder that persists beyond the time needed for healing to occur from an injury. Although chronic pain is sometimes associated with tissue damage, in the majority of cases tissue damage has not been substantiated.

According to the modified gate control theory of Ronald Melzack and colleagues, chronic pain is mainly a signal detection dysfunction in which a variety of haptic sensory inputs are misinterpreted as pain. Daniel Brown likened the development of chronic pain to the development of panic states, both of which entail a progressive learned misappraisal and phobic avoidance of somatic distress. Extensive phobic avoidance of pain sensations maintains the signal detection dysfunction, in that the patient rarely attends to somatic distress in a way that would allow for correction of signal detection errors. Chronic pain syndrome is also frequently associated with significant cognitive distortions and illness behaviors. According to Wain,

Nowhere in medicine, does there seem to be a stronger link between the physiological, psychological, and sociological variables, than there is in the understanding of the diagnosis and treatment of the patient with chronic pain. It is the trauma of the pain, suffering, and anguish that unites every physician, nurse, and health care specialist.

Wain further noted that a chronic pain experience is associated with significant problems in coping with associated social and environmental factors. In contrast to acute pain disorders, interventions such as prolonged rest and immobility are frequently contraindicated in the effective management of chronic pain syndromes.

Hypnosis frequently is also used to address the illness behaviors and suffering associated with chronic pain. In *Hypnosis and Behavioral Medicine,* Brown and Fromm described a phase-oriented approach to hypnosis and chronic pain. The first phase entails identification of specific pain coping strategies available to the given patient with chronic pain, followed by the application of these pain coping strategies in trance, along a hierarchy from less intense, peripheral to more intense, target pain locations.

Hypnotic techniques are used to correct negative self-talk and distorted pain beliefs to the extent to which they interfere with the successful application of pain-coping strategies. The second phase is skill development, in which the patient with chronic pain develops mastery over the pain experience within and outside of the treatment session and learns successfully to use a range of pain-coping strategies whenever pain occurs in everyday life. At this stage phobic avoidance of pain is reduced. The patient learns not only to tolerate pain adequately, but also to discriminate a wide range of sensory signals other than chronic pain. The third stage, the rehabilitation phase, addresses the issues of chronicity, for example, the

normalization of daily life, developing a healthy, pain-free self-representation and eliminating analgesic addiction.

Tiara Dillworth and Mark P. Jensen have provided an excellent review of controlled trials that have demonstrated hypnosis is an efficacious treatment of chronic pain. Their studies did review a number of controlled studies that have documented hypnosis is an efficacious treatment of chronic pain syndromes.

Multiple Convergent Pain Management. The Institute of Medicine published a landmark report in 1987 entitled "Pain and Disability: Clinical, Behavioral, and Public Policy Perspectives." The report, prepared by a committee of authorities, contained a literature review, an analysis of chronic pain and its treatment, and specific recommendations for future developments in medical research and public policy concerning chronic pain.

Although the Institute of Medicine study suggested the need for further research on the different modes of treatment and their efficacy, it stated that multimodality and multidisciplinary therapeutic interventions lessen the disability and dysfunction associated with chronic pain.

Ronald Melzack and Patrick Wall described multimodality and multidisciplinary therapy for chronic pain as a multiconvergent pain management program that includes physical, psychological (behavioral medicine), and pharmaceutical interventions and interventional anesthesia (Fig. 33.4–3). This type of pain management program uses a rehabilitation and collaborative model of intervention. The psychological interventions include several modes of treatment based on the principles of both operant conditioning and cognitive-behavior management. Hypnosis has a definite role in any treatment program that uses contingency management and cognitive-behavior management approaches to chronic pain. Certain features of the behavior of the patient with chronic pain demonstrate the value of hypnosis as a mode of treatment—specifically, demoralization, suffering, and low self-efficacy.

In addition to the aforementioned psychological problems associated with chronic pain, it is also associated with significant suffering based on the patient's fear that he or she will be unable to control the pain. The suffering component of the chronic pain experience is also amenable to being ameliorated through hypnotic interventions.

In addition to the medical suffering associated with chronic pain, chronic pain is also associated with demoralization, which is amenable to hypnotic interventions. Because of the influence of cognitive processes on the perception and integration of pain and the significant complications regarding the psychological problems associated with chronic pain, interventions that significantly modify cognition and perception would be efficacious in the management of pain. This observation provides a firm

foundation for using hypnosis in chronic pain management.

Hypnosis lends itself to both the operant conditioning (contingency management) and the cognitive-behavior modification therapies that are well established in multiple convergent pain management programs. In view of the fact that there is a nonspecific (placebo) response and a direct effect of the hypnotic experience, both the nonspecific and the direct effect of hypnosis can be used appropriately in the management of chronic pain syndromes.

There are significant anecdotal reports and some cases with controls that demonstrate that patients with higher hypnotizability scores on hypnotic assessment tests generally report greater analgesic effects of hypnosis and significantly improved cognitive functioning through the use of hypnosis. The use of hypnosis should involve multiple interventions during the course of the chronic pain management and self-hypnosis training. In their review "Hypnosis and Clinical Pain," Patterson and Jensen provided significant evidence concerning the lack of control studies in chronic pain studies. They recommended that hypnosis should be used in a more organized manner in the management of chronic pain. They also suggested that hypnosis may be efficacious in the treatment of the suffering and cognitive distortions experienced by patients with chronic pain. Their review included an important meta-analysis suggesting that more than 70 percent of the patients in multiple convergent pain management programs could have benefited from adding hypnotherapy to the treatment program. Such interventions could be directed toward both the alleviation of the pain experience and the treatment of the suffering and cognitive distortions associated with chronic pain behaviors. A more recent meta-analysis comparing hypnosis to behavioral methods for chronic pain Tononori Adachi and his associates found that hypnotic treatment of chronic (nonheadache) pain had a larger treatment effect size as compared to other psychological interventions like cognitive-behavioral treatment or guided imagery.

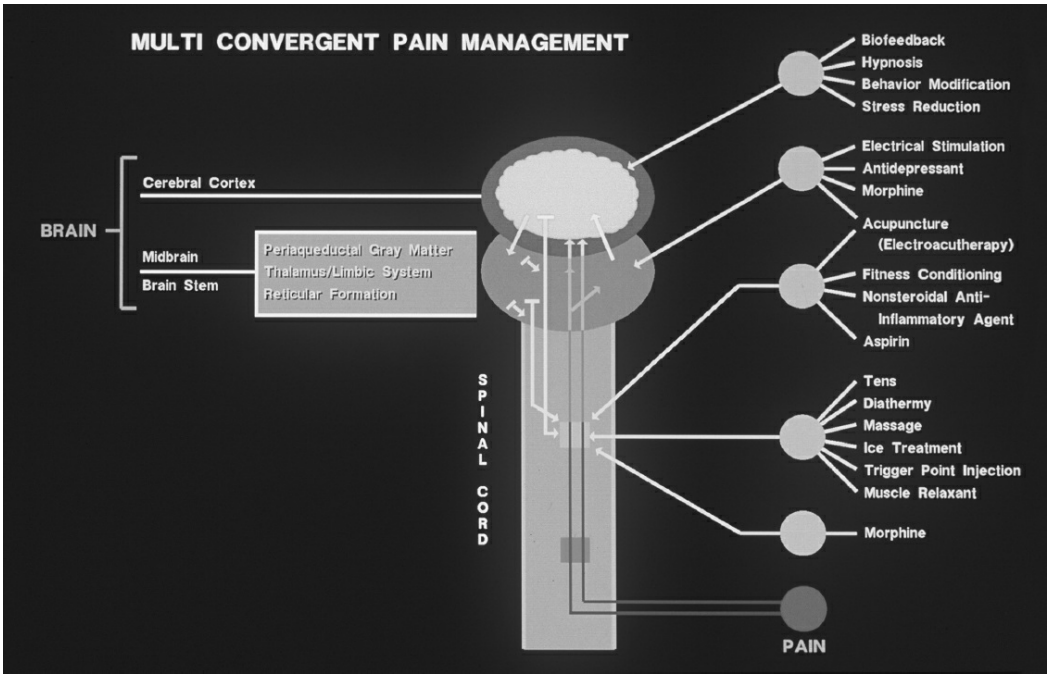

FIGURE 33.4–3. Multiconvergent pain management involves multiple interventions that act on different parts of the central nervous system and spinal cord. These interventions substantially diminish the pain, suffering, and depression produced by the chronic pain experience. Hypnosis plays an important role in significantly reducing the pain, demoralization, and suffering of chronic pain. Hypnosis also facilitates the efficacy of biofeedback, behavior modification, stress reduction, and muscle relaxation in multiconvergent pain management. TENS, transcutaneous electrical nerve stimulation.

During the initial evaluation of such patients for hypnosis and self-hypnosis training, they should be given a full explanation of the benefits of this training for the management of their chronic pain, as well as a full understanding of the misconceptions and myths about hypnosis. They should also undergo a hypnotizability assessment, such as the SHCS for Adults or the HIP. David Axelrad developed a method of induction—open neurosensory fixation—for use in the individualized hypnotherapy and self-hypnosis training of the patient with chronic pain. This method, described in the paper "The Role of Hypnosis in Multiple Convergent Management of Chronic Pain," includes engaging in eye fixation on a visual source, listening to ambient environmental sounds, experiencing kinesthetic and other pleasurable sensations in the skin, muscles, and joints, and engaging in progressive muscle relaxation (Fig. 33.4–4). This method can be used by patients with high, moderate, or low hypnotic capacity.

Patients with chronic pain should be provided with individualized hypnotherapy that incorporates suggestions for relaxation and progressive muscle relaxation, or some other form of relaxation intervention, that permits the patient to undergo relaxation responses from the hypnotic experience. The suggestions should also be directed toward direct pain reduction and pain relief.

In view of the significant association of depression and demoralization

with chronic pain, the hypnotic interventions should also include suggestions that are directed toward alleviation of depression and demoralization.

Displacement suggestions, distraction suggestions, and transformation suggestions may also be employed during the course of the hypnotherapy. Following several individualized hypnotic interventions, the patient should also be provided with training in self-hypnosis in a group setting. The self-hypnosis training experience is an integral part of a hypnotic intervention for patients with chronic pain. During the course of a multiple convergent pain management program, patients should be encouraged to undergo self-hypnosis in a variety of situations that provoke increased pain. Such experiences reinforce the "well" behavior and help the patient to establish self-efficacy and sense of personal mastery.

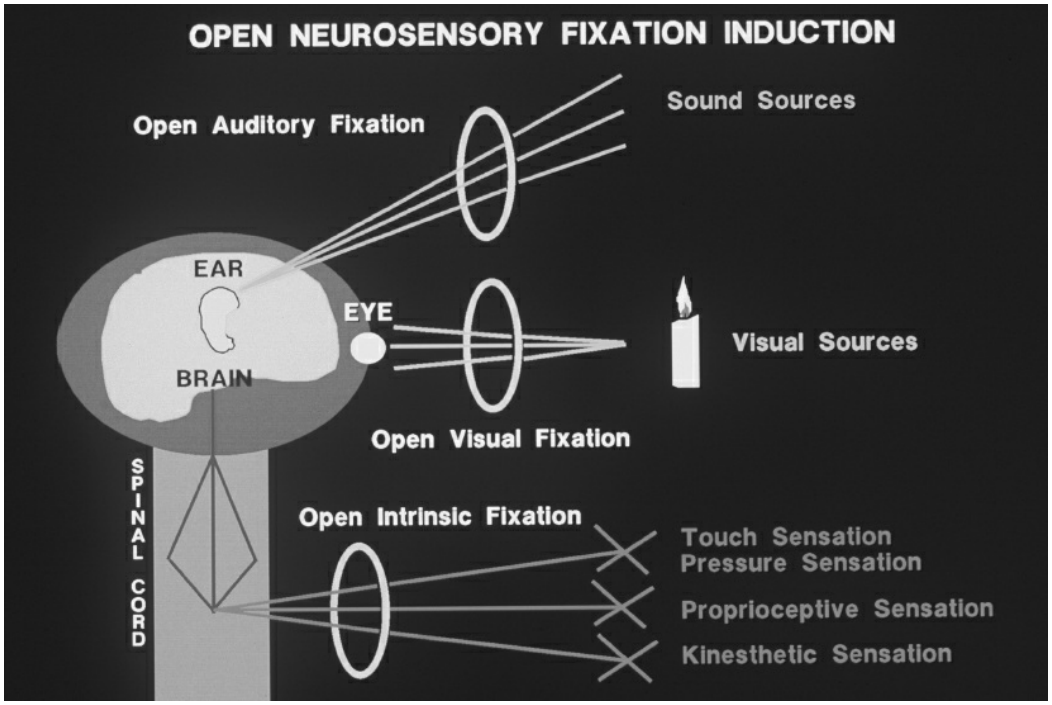

FIGURE 33.4–4. Induction of a patient, using open neurosensory fixation induction.

Ms. D was a 63-year-old white woman referred for symptoms of chronic pain and PTSD caused by the experience of being a passenger in a motor vehicle that was struck by an 18-wheeler. Ms. D had been in the front-passenger seat, restrained by a seatbelt and shoulder harness. During the accident, she had experienced severe fright and the fear that she would be killed. She had sustained injuries to her left shoulder and neck, which required an emergency room assessment and evaluation.

During the emergency room evaluation, a computed tomography (CT) scan of her neck and back reflected significant degenerative disease in her cervical spine, a disk protrusion in her cervical spine, and degenerative disease in her lumbar spine. An MRI examination also documented these injuries. Subsequent studies also revealed a negative axial CT scan and a negative MRI of the brain. The patient's chronic pain was severe. Because of the patient's significant history of insulin-dependent diabetes and other problems, her orthopedic surgeon and internist recommended conservative pain management interventions. Ms. D also experienced an acute stress disorder, followed by a PTSD, after her injuries. Her symptoms

included recurrent nightmares and re-experiencing images of the motor vehicle accident. She also exhibited hyperarousal and a significant startle reflex. She reported marked avoidance behavior, avoiding freeway driving, and experiencing increased anxiety and panic attacks whenever a motor vehicle in which she was riding was approached by an 18-wheeler. She reported significantly increased anxiety in response to both internal and external cues reminding her of the motor vehicle accident.

The patient underwent treatment of her PTSD, which included individual and group hypnotherapy and a self-hypnosis training program. She was also provided with relaxation exercises and biofeedback training. She also required multiple convergent pain medicine interventions, which included physical therapy interventions and behavioral interventions for her chronic pain syndrome. She was also provided treatment with a SSRI antidepressant.

Over the course of her treatment, she was able to use both hypnosis and self-hypnosis training to significantly reduce her anxiety, abort her panic attacks, and significantly diminish chronic pain experiences involving her neck and upper back. This self-hypnosis training program also provided her with improved control over her posttraumatic headaches. Her orientation to the hypnosis experience included the administration of the SCSHS, on which she had a score of 5, demonstrating that she had marked hypnotic capacity.

The focused open neurosensory induction program Ms. D included eye fixation on a visual object and environmental experiences (sound) and direction of her attention inward to neurosensory experiences, including kinesthetic sensations. She was instructed in diaphragmatic breathing, which is a critical component of most hypnotic inductions and relaxation or meditative experiences. The training program included suggestions calling for eye closure and progressive muscle relaxation.

During the course of the patient's training, she was provided with suggestions for diminishing her fear and anxiety associated with both internal and external cues reminding her of the accident, as well as direct suggestions for pain relief in her neck and back. She also incorporated these suggestions for reducing her headaches by imagining having a cold pack placed on her forehead and neck.

Over the course of Ms. D's treatment, she experienced a significant remission of her PTSD symptoms and significant reduction in her pain experiences, as judged by both the McGill-Melzack Pain Inventory and the Visual Analog Scale.

Mr. G was a 55-year-old African American man who experienced an occupational injury that caused a herniated disk at L4–5. He experienced intractable pain, and he failed to respond to conservative interventions. He underwent a laminectomy and fusion procedure for this injury.

Unfortunately, his pain returned, and he was unable to return to full-time employment. Over the course of his medical history, he underwent four additional surgical interventions to his low back region. He was referred to a pain center because of intractable pain associated with significant tolerance to opiate medications. At the time of his admission to the pain program, he was requiring methadone at a dose of 60 mg four times daily, augmented with hydrocodone at 2 mg two to three times daily.

Following his admission to the pain medicine program, he was administered the Stanford Hypnotic Susceptibility Scale, on which he had a score of 11, indicating marked hypnotic capacity. At the time of his entry into the program, he was seriously depressed and demoralized. He was provided with cognitive-behavioral management, self-hypnosis training, and physical therapy interventions, emphasizing aerobic conditioning and strengthening exercises. He was also treated with antidepressant medication to assist in the treatment of his depression and his chronic pain syndrome.

Mr G. was also provided with hypnosis and self-hypnosis training similar to that described in the previous case. He was able to use open neurosensory induction to significantly reduce the pain in his back, increase relaxation of his back muscles, and improve his self-image.

The hypnosis was also directed toward enabling the patient to experience significant reduction in pain without the need for opiate medications. Based on his training, the patient was able to use self-hypnosis once or twice daily to significantly reduce his pain and discomfort. Over the course of his treatment, he was successfully withdrawn from opiate medications. He was discharged free of opiate medications was successfully rehabilitated through a state-supported rehabilitation program.

Hypnosis in Headache Management

As noted in the review by Patterson and Jensen, far more studies have focused on the use of hypnosis for headache than for any other etiology of chronic pain. Their review included nine studies that involved some controls. The hypnotic interventions in the headache management studies lasted for significant periods of time, and the patients were provided with suggestions for relaxation, visual imagery techniques, verbal reinforcers, and direct suggestions for pain reduction. Patients were frequently provided with self-hypnosis training as a part of their headache management program. The headache management program was also combined with biofeedback training and other methods of autogenic relaxation. A controlled study demonstrated that patients provided hypnotic interventions experience significantly greater pain control than patients provided biofeedback training alone. The hypnotic interventions provided for patients in this study used eye fixation followed by suggestions for relaxation, analgesia, and numbness.

Many of the headache studies reviewed by Patterson and Jensen included hypnotic assessments and measures of suggestibility as a part of the hypnotic intervention program. The researchers, who adopted social influence models for hypnotic responding, also reported that patients provided with such interventions reported significant improvement of their headaches without the induction of a hypnotic state.

HYPNOSIS AND IRRITABLE BOWEL SYNDROME

Functional abdominal pain (FAP) and IBS are highly prevalent in both childhood and adulthood. IBS entails gut motility changes and recurrent pain and is accompanied by a functional dissociation of awareness of somatic events combined with excessive worry about bodily experience. Since the pioneering work of P. J. Whorwell and his associates reported in *Lancet* there have been several dozen controlled, evidence-based studies that have demonstrated that hypnotherapy for FAP and IBS is a highly effective treatment both for children and adults. These studies show hypnotherapy to be superior to standard medical care or supportive psychotherapy for the reduction in pain and stabilization of gut motility. Some of these studies also elucidated both the physiological and psychological mechanisms of the hypnotic effects. One of the mechanisms of action of hypnotherapy involved changes in both colonic motility and rectal sensitivity. Hypnotherapy also reduced excessive worry and somatic dissociation in patients with IBS. In view of these studies, hypnotherapy—individual, group, and self-hypnosis training—is recommended for the treatment of IBS in both adults and children. Vivian Miller and Peter J. Whorwell developed a similar protocol in the treatment of inflammatory

bowel disease (IBD), but studies in this area are only just beginning to appear.

Susanna Carolusson has provided an excellent review and significant applications of dynamic hypnosis in the treatment of IBS. She has provided a randomized controlled study of the use of hypnosis to the treatment of IBS.

HYPNOSIS AND ASTHMA

Asthma entails intermittent, variable airways obstruction due to chronic bronchospasm. It may also involve an ongoing inflammatory response. In the 1960s the Hypnotherapy in Asthma Subcommittee of the British Tuberculosis Association conducted a series of randomized clinical trials using direct hypnotic suggestions to treat asthma.

Whereas the results were promising, the findings were largely restricted to subjective outcome measures, with one important exception: Asthmatics treated with hypnosis as compared to a relaxation group showed a significant increase in forced expiratory volume ()over 1-year span.

In New Zealand, Timothy C. Ewer and Donald E. Stewart conducted an important randomized, controlled hypnotherapy trial with 39 adults with mild-to-moderate asthma. The hypnosis group underwent six weekly 1/2-hour hypnosis sessions plus daily self-hypnosis. The hypnosis consisted of a hypnotic induction, relaxation suggestions, ego strengthening, guided imagery, and direct suggestions for mastery over asthma symptoms. The control group intervention consisted of nurse review of daily symptom diary entries. The hypnosis treatment group showed a significant reduction in airways obstruction as measured by peak flow, a significant reduction in airways irritability as measured by a methacholine challenge test, and a significant reduction in the use of bronchodilator medication. High hypnotizability strongly predicted the positive treatment response. Marianella Castes and coworkers reported a Venezuelan study in which 35 asthmatic children used guided imagery to change their inflammatory response. The researchers found a significant decrease in forced expiratory volume, a decrease in immunoglobulin E antibodies, and a decrease in the number of $CD23^+$ cells, which meditate the inflammatory response.

Unfortunately, neither of these studies has been replicated. Nonetheless, they suggest that hypnosis is possibly efficacious in the treatment of asthma with respect to increasing airways flow, decreasing airways hyper-responsiveness, and possibly curtailing the inflammatory response and that high hypnotizability predicts a positive response. With respect to peak flow, some critics have pointed out that these seemingly positive biological changes are largely restricted to effects on the large air passages. Nevertheless, given the fact that the magnitude of airways change is clinically significant especially in high hypnotizables, hypnosis should be seriously considered for this group.

HYPNOSIS AND IMMUNE-RELATED DISORDERS

Hypnotically assisted profound relaxation and specific hypnotic suggestions for cooling localized target tissue has been shown to reduce acute inflammation in, for example, burns and also to diminish chronic inflammation associated with autoimmune conditions like rheumatoid arthritis. Stephen Black and his associates found that direct hypnotic suggestions markedly reduced the expected weal and flare in response to a passive transfer skin test to a known allergen but did not alter cellular infiltration. Despite this clear laboratory demonstration of reduced allergic response, there have been only a few clinical studies applying hypnosis to the treatment of the gross symptoms of allergic sensitivity.

The main area in which hypnosis has been used in the treatment of immune-related disorders is psycho-oncology. Hypnosis can be used to facilitate the development of active coping strategies to deal with the crisis of the cancer diagnosis, death anxiety, life disruption following diagnosis, uncertainly of course of illness, helplessness, and noxious medical procedures.

Hypnosis can be used to ameliorate psychiatric symptoms associated with a cancer diagnosis, such as depression, anxiety, sleep disturbance, and pain. Exploratory hypnosis can also be used for meaning making and benefit finding following the initial cancer diagnosis, as well as for enhancing quality of life in the face of the diagnosis. A meta-analysis of psychosocial interventions in cancer showed that the largest treatment effects are found in altering quality of life, with more modest treatment effects associated with coping enhancement and emotional adjustment. The role of psychosocial interventions in prolonging survival rate is controversial, with studies yielding both positive and negative results. Most studies find that facilitating active coping and encouraging emotional expressivity are associated with improved immune regulation, but it is less clear whether improved immunoregulation affects overall survival rate. James Hall found that high hypnotizability is a strong predictor of which individuals are more likely to alter immune response following psychosocial interventions for cancer.

HYPNOSIS AND SEXUAL DYSFUNCTION

The classification of sexual dysfunctions typically corresponds to the three phases of sexual response—desire, arousal, and climax. Conjoint sex therapy is an empirically supported approach to the treatment of disorders of arousal and climax in both men (erectile dysfunction and pre- or postmature ejaculation) and women (frigidity, anorgasmia), in which couples are given a structured series of sexual assignments designed to impart new sexual skills and identify/work through conflicts that interfere with normal sexual response. Hypnosis can be used as an adjunct to

conjoint sex therapy in couples. Daniel Araoz's book *The New Hypnosis in Sex Therapy* contains a wide range of sex hypnotherapy techniques for supporting conjoint sex therapy.

The consensus among sex therapists is that disorders of desire in men and women (inhibited sexual desire; sexual compulsivity) are more complex, in that long-standing unresolved traumatic or neurotic conflicts and maladaptive cognitive learning are important in their development. At the Masters and Johnson Sex Clinic, Mark Schwartz used a combination of traditional conjoint treatment along with dynamic exploratory hypnotherapy to identify the conflicts contributing to sexual desire disorders. Although no randomized clinical trials exist on hypnotherapy for sexual desire disorders, Schwartz's data constitute a large cohort study followed over time that offers promise for hypnotherapy in an area in which outcome success has been generally lacking.

HYPNOSIS AND DERMATOLOGY

One of the classic psychosomatic disorders is neurodermatitis. The link between emotions and skin lesions has been well reported in the literature. In particular, several clinicians have reported the successful treatment of warts with hypnotic intervention that were refractory to routine dermatological treatment. One clinician reported the successful removal of penile condyloma acuminatum. Clinicians have also reported successful treatment of ichthyosis, pruritus, alopecia areata, eczema, boils, angular cheilosis, and psoriasis. Areed Barabasz and his associates found that treatment of human papilloma virus (HPV)-related warts by hypnosis was more effective than medical tissue assault methods in long-term follow-up. In contrast to these clinical reports, some of the research has been less positive. Therefore, it may be appropriate to consider the use of hypnosis if other treatments fail. The more hypnotizable the patient, the more likely it is that physical alterations can occur. The use of hypnotic techniques as an adjunct to other treatments also appears warranted. It is safe to conclude that emotions have a strong effect on the skin and that hypnosis can have an effect on emotions.

HYPNOSIS AND ADDICTIVE BEHAVIORS

Hypnosis and Smoking Cessation

Hypnosis has been used for smoking cessation for many years. Many anecdotal studies have been published, as have several retrospective clinical studies that provide some encouragement for the use of hypnosis in smoking cessation. Unfortunately, many of these studies were limited because they did not include biological markers of smoking cessation, nor was there randomization. Furthermore, the studies also varied in the amount of hypnotherapy provided during the hypnotic intervention.

There are two *major* approaches to using hypnosis for treating smoking:

- Multimodal cognitive-behavioral treatment conducted in trance. Brown and Fromm described an integrated hypnobehavioral protocol that addresses motivation for treatment; self-monitoring of urges and smoking behavior; nicotine fading; stimulus control; self-regulation of smoking behavior; smoking-related cognitive distortions; the effect of regulatory use of smoking; the social context of smoking; identification and prevention of relapse; and smoker self-image.
- Single-session or brief hypnosis, as exemplified in the work of Herbert and David Spiegel. The Spiegel approach uses the eye roll induction method, followed by a series of brief health-benefit suggestions designed to increase commitment for smoking cessation and the use of posthypnotic suggestions coupled with self-hypnosis training. During the course of the hypnotic intervention, positive imagery may also be used.
- Others used a modified approach of the above where strategies are based on first developing the therapeutic alliance with the patient and structuring individual strategies based on their past experiences and what might motivate them to change. Those individual motivations are incorporated into the hypnotic suggestion given during the treatment session and in their self-hypnotic interventions.

At least one study using the longer hypnobehavioral protocol demonstrated that the intervention associated with an intensive hypnotherapy intervention program can be an effective means of achieving smoking cessation. In this one study, the rate of smoking cessation at 26-week follow-up was 40 percent and was confirmed by an expired carbon monoxide value of 8 ppm or less. The investigators noted that limitations in the study were the small sample size and a lack of measurement of hypnotizability.

Anecdotal reports of the single-session Spiegel approach included hypnotizability assessment using the HIP. These reports documented significant abstinence in 20 to 35 percent of subjects, with abstinence rates as high as 40 percent in subjects with higher hypnotizability scores on the HIP. The Spiegels reported that a consecutive series of 226 smokers referred for the smoking cessation program were treated with a single-session habit restructuring intervention involving self-hypnosis; they were followed for 2 years. The criterion for successful outcome was total abstinence for smoking; 52 percent of the study group achieved complete smoking abstinence 1 week after the intervention, and 23 percent maintained their abstinence for 2 years. The hypnotic interventions for smoking cessation also included self-hypnosis training in which the patient was provided with suggestions to minimize negative self-talk and to incorporate the benefits of nonsmoking and the ability to become a

nonsmoker.

Both approaches typically incorporate basic educational experiences regarding the deleterious physical effects of smoking and the significant benefits of overcoming the smoking addiction.

The suggestions provided during the hypnosis generally include increased control in mastery and the positive experiences of healthy living. The suggestions should also be provided to counterbalance the symptoms of withdrawal.

Hypnosis in Weight Management

Hypnosis has been used in a number of weight control and weight management programs. It is generally combined with exercise and comprehensive dietary management. The purpose of a brief hypnotic intervention is to provide the patient with an opportunity to restructure his or her relationship to food. During the course of the hypnotic experience, patients are provided with a number of suggestions to facilitate greater control over their eating behavior. The hypnotic interventions are combined with a number of affirmations that facilitate the patient's dietary management. Use of imagery may also assist in the patient's weight management. Patients are provided with encouragement to use self-hypnosis on a daily basis to facilitate management of their eating behavior.

Brown and Fromm described a manualized, multimodal protocol integrating hypnosis and cognitive-behavioral treatment that addresses motivation for treatment, stimulus control, self-regulation of eating behaviors, problematic eating patterns, managing urges, cognitive control over eating, relapse risk assessment and prevention, and dealing with dynamic factors associated with problematic eating. Irving Kirsch found that the addition of hypnosis to a cognitive-behavioral protocol for weight control made a significant contribution to the overall treatment effect.

Entwistle, Webb, and Abayomi have provided an excellent review of the literature concerning the use of hypnosis in weight management programs. Their review does provide an excellent understanding of the value of hypnosis in weight management programs.

Hypnosis and Eating Disorders

Beginning with Pierre Janet's classic case studies more than 100 years ago through the modern hypnodynamic treatment approach of Moshe Torem, hypnotherapy has a long history in the treatment of anorexia and bulimia. Studies consistently have shown that patients with bulimia, and certain patients with anorexia, are more hypnotizable than normal control subjects. Johan Vander Linden in Belgium developed a detailed manualized approach to hypnotherapeutic treatment of patients with eating disorder with and without a history of trauma that combines the

development of cognitive-behavioral skills in hypnosis and exploratory hypnotherapy with family therapy. The 13-step eating disorder protocol conducted over a 6-month period addresses providing psychosocial education, lowering medical risk, establishing therapeutic goals, increasing motivation, normalizing eating habits, evaluating progress, enhancing self-esteem, improving bodily experience, learning relaxation, enhancing autonomy, learning to recognize and tolerate negative and positive feelings, addressing trauma, and enhancing relapse prevention. Hypnosis is used in different ways across most of the 13 steps. Clinical outcome data show that 55 percent of the patients no longer met sufficient criteria for an eating disorder diagnosis after 1 year, with 70 percent improvement in patients with bulimia, 52 percent improvement in patients with restricting anorexia, and only 18 percent improvement in patients with mixed anorexia and bulimia.

Hypnosis and Complex Habits

Hypnosis has also been used in the treatment of complex habits such as nail biting, tongue sucking, skin picking, and hair pulling or trichotillomania. The plucking of hair may occur for some individuals whose anxieties or internal tensions reach a crescendo, but the behavior becomes so overlearned that it serves to regulate a wide variety of affective and cognitive states in the patient. Trichotillomania is defined as the recurrent failure to resist the plucking of hair. Hair, eyebrows, and moustache are the targets.

Patients are often unaware of engaging in the behavior.

HYPNOSIS IN FORENSIC PSYCHIATRY

Hypnosis has been used as a forensic investigative tool to refresh the recollection of witnesses and victims, as well as sometimes suspects, of crimes. In the 1960s and 1970s, hypnotically refreshed memory was openly admitted as testimony in most courts. The forensic use of hypnosis greatly increased its popularity in those years, especially among law enforcement agents. Hundreds of police officers were given brief training in hypnosis. One popular program, for example, consisted of only 32 training hours. Given the general lack of psychological training of the police taking the courses and the cursory training given in forensic hypnotic interviewing, it was inevitable that the quality of investigative hypnosis would deteriorate. Martin Orne first articulated the misuses of forensic hypnosis and was greatly influential in cautioning the courts about abuses of forensic hypnosis and the questionable accuracy of hypnotically refreshed testimony.

In the early 1980s, a series of landmark legal decisions changed the future of forensic hypnosis. Many states then adopted a per se *exclusion*

standard, according to which hypnotically refreshed testimony was found inadmissible. Other states found the per se rule too extreme and adopted a *guidelines rule* in which hypnotically refreshed testimony was admitted, provided that the hypnotic interview was conducted by a competent mental health professional with specialty training in hypnotic interviewing, that no one else was in the room, and that a tape recording of each session preserved the record so that it could later be reviewed and scrutinized for possible undue influence.

In 1985, the AMA published a cautionary statement specifically on the forensic, not clinical, use of hypnosis. Unfortunately, most of the landmark court decisions limiting hypnotically refreshed testimony and also the 1985 AMA report were premature, in that the findings were written before the greater bulk of research findings on hypnosis and memory had appeared in the professional literature. The AMA report only cited a small number of the scientific studies on hypnosis in print at the time. Likewise, the only meta-analytic study on hypnosis and memory included 24 studies, only 4 of which pertained to memory recall. An expanded list of 38 studies was reviewed by Matthew Erdelyi in the same year. Erdelyi included 17 studies on the recall of meaningful material, only 3 of which were included in the aforementioned meta-analysis. Reviews and updates on scientific research on hypnosis and memory that are more representative and inclusive of the available research findings can be found in *Memory Guidelines* by the American Society of Clinical Hypnosis and *Memory, Trauma Treatment, and the Law* by Daniel Brown and coworkers.

The main research findings fall into three broad areas: hypnotic hypermnesia, hypnotic pseudomemories, and confidence in hypnotically refreshed memory.

- *Hypnotic hypermnesia* pertains to the relative *completeness* of memory. Laboratory subjects are shown a series of target stimuli and are then asked to recall or recognize the original stimulus material under waking and hypnotic conditions. Hypnotic hypermnesia occurs if the subject remembers significantly more new information under hypnotic as compared to nonhypnotic conditions. Hypermnesia refers to the amount of information or the completeness of the recollection, not necessarily its accuracy.

Although the results across numerous studies are quite variable, the overall results suggest little or no advantage of hypnosis in recognition memory studies. Results across many studies show a modest increase in the amount of information recalled with hypnosis for more- but not for less-meaningful stimulus material. However, this conclusion is complicated in several respects. First, research designs do not always discriminate repeated recall and hypnotic recall; if the initial recall attempt occurs in the waking state and the second recall in hypnosis, there is no way of knowing

whether the advantage of the second recall was due to hypnosis or simply to a second recall. Second, many studies show that hypnotized subjects asked to recall a target event produce a remarkable increase in, yet a widely variable amount, of responses. Martin Orne hypothesized that hypnosis resulted in a lowering of response criteria, so that subjects used lax criteria when deciding whether to report a memory or other mental content like a thought or image as if it were a memory. However, subsequent research showed that when a shift to lax response criteria occurs, it is more a function of establishing unrealistic expectancies regarding memory retrieval and less a function of a hypnotic procedure per se. Although the *lax response-criteria hypothesis* served as the dominant explanation for hypnotic hypermnesia for several decades, a series of experiments in memory research outside of the hypnosis field by Henry L. Roediger and Matthew Erdelyi largely refuted this explanation.

The method used to control the variability in productivity in response to hypnotic hypermnesia suggestions was *forced recall,* according to which subjects were shown a fixed number of stimulus pictures, say 60, and after recalling as many as possible, say 38, were instructed to fill in the other 22 lines on the answer sheet. Unfortunately, although this method controlled for variable response criteria across subjects, it created an additional problem. The findings from many of the studies using the forced-recall procedure show an increase in both accurate and inaccurate information about the target event. This became known as the "hits and misses" view of hypnotic hypermnesia.

However, this increase in the memory error rate is likely an artifact of the forced-recall method, which essentially encourages guessing. Research comparing both forced and nonforced recall failed to show any significant differences in the number of correct or incorrect responses in the hypnotic condition.

Overall, if any pattern emerges from these studies, it is that hypnosis may facilitate a modest recall of new information about a meaningful target event, at least in certain subjects, but that the effects of hypnosis per se are not always clearly distinguishable from relaxation and repeated recall effects. Although some critics assert that hypnotic hypermnesia runs the risk of inaccurate recall, this "hits and misses" hypothesis is misleading in that it does not distinguish between the effects of hypnosis per se and the effects of poor interviewing techniques and/or making unrealistic claims about memory retrieval and encouraged guessing.

- The research on hypnotic *pseudomemories* refers to the relative *accuracy or inaccuracy* of the memory. There is an extensive scientific literature demonstrating that some normal (nonhypnotized) individuals are vulnerable to changing their memory reports for a target event when subjected to postevent suggested misinformation. This is known as the

misinformation suggestion effect. Vulnerability of misinformation suggestions occurs both in the normal waking state and in the hypnotic context. In the hypnotic context, this is known as the hypnotic pseudomemory effect. Kevin McConkey, Amanda Barnier, and Peter Sheehan defined hypnotic pseudomemory as "memory reported when a suggestion is given following a hypnotic induction that a person will remember specific information that the hypnotist knows (or believes) to be false." Extensive research on the hypnotic pseudomemory effect has compared the response to misinformation suggestions in both high- and low-hypnotizable subjects, in the waking state, and following a hypnotic induction. This extensive research generally shows that pseudomemories were reported both by high and low hypnotizables but at higher rates in high hypnotizables, and also that pseudomemories were reported across hypnotic, waking, and simulation-of-hypnosis conditions. There is a trend that some portion of but not all high hypnotizables who were also hypnotized have at best modestly higher rates of pseudomemories than high hypnotizables in the waking condition or than low hypnotizables in either condition.

High hypnotizability, not the hypnotic procedure per se, is associated with a modest increase in pseudomemory rates. However, Graham F. Wagstaff and his associates found that when high hypnotizables as compared to low hypnotizables are given instructions that imply that hypnosis leads to more, not less accurate reporting that are less likely to produce false memories. However, the consistent finding from hypnotic pseudomemory research is that a hypnotic procedure per se poses no more risk of pseudomemory production than a walking interview procedure. The primary risk factor for inaccuracy due to suggested misinformation is an interview that includes suggested misinformation, irrespective of whether this misinformation is given in the waking state or in hypnosis. The problem of pseudomemory inaccuracy lies in suggestive interviewing and setting unrealistic expectations, not hypnosis. Some critics have argued that the effects of hypnosis and misinformation suggestions are additive and that both contribute to the overall variance to the hypnotic pseudomemory effect. However, an attempt to replicate the findings with a more sophisticated signal detection design failed to replicate the earlier findings that hypnosis had a negative effect on memory accuracy, but did replicate the rather robust finding that suggestive interviewing independent of hypnosis significantly increased memory inaccuracy. If anything has been learned from the numerous pseudomemory studies, it is that hypnosis as a procedure per se does not cause pseudomemories, and that interviews conducted with and without hypnosis must scrupulously avoid supplying misinformation, that is, avoid supplying *any* content that has not been already reported by the interview subject.

- *Undue confidence*. Martin Orne and Bernard Diamond cautioned against the use of hypnosis because they believed that hypnotically refreshed memory may foster undue confidence about inaccurate memories and that such undue confidence would make such individuals unusually resistant to cross-examination in the courtroom. However, the many research studies pertaining to hypothesized increased target memory confidence following a hypnotic induction have yielded equivocal results: High hypnotizability and/or a hypnotic induction were found significantly to increase confidence in hypnotically retrieved memories in some studies and yet failed to do so in other studies. The pattern of results implies that undue confidence is not a consequence of a hypnotic procedure per se, but rather is a function of directly communicating unrealistic expectancies about hypnotic memory retrieval ("You will be able to remember absolutely everything," "Your memory is like a tape recorder and you can review all the details on a television screen in hypnosis"). Other factors contributing to inflated confidence may include high hypnotizability, less meaningful or ambiguous stimulus material, shorter retention intervals, and permission to guess. There is no compelling evidence that hypnotic procedures per se are necessarily associated with inflated confidence, and in most of the studies demonstrating inflated confidence the magnitude of the effect is not large.
- The *memory hardening position* states that undue confidence in hypnotic pseudomemories would make a witness impervious to cross-examination in the courts. Laboratory studies designed to test Orne's hypothesis have consistently failed to support the memory hardening view. These findings provide no support for the notion that the influence of hypnotic procedures on memory are so powerful that witnesses with hypnotically refreshed testimony cannot be effectively cross-examined or that hypnotically refreshed testimony should be found inadmissible because it is particularly resistant to cross-examination.
- The overall findings from extensive research on hypnosis and memory indicate that whereas hypnosis may contribute to a modest increase in new information about a target event at least in certain subjects, that there is little evidence that hypnosis per se contributes to an increase in the memory error rate, and that the greater portion of the variance of inaccuracies associated with hypnotically refreshing memory are best attributed to nonhypnotic factors—mainly to poor interviewing techniques that include fostering unrealistic expectancies about memory retrieval, encouraged guessing, and suggesting misinformation.

As an outgrowth of the enormous research literature on memory retrieval, the misinformation effect, and the effects of interview components on the completeness and accuracy of recall, there is a better

understanding of how to conduct nonhypnotic and hypnotic investigative interviews that minimize the memory error rate while maximizing the retrieval of new information about a target event. These findings for forensic hypnotic interviewing can be summarized in the following guidelines:

- *Proper informed consent*. Any forensic hypnotic interview begins with informed consent about the nature of memory, memory retrieval and hypnosis, the risks and benefits of the procedure, and possible alternatives. In the mental health field, informed consent is not seen as a single event, namely providing information at the onset of treatment, followed by signing a consent form; it is best seen as an ongoing process of open communication and adjustment of the unfolding conceptualization of treatment problems, methods, and goals, as part of an unfolding relationship. In some jurisdictions, clients need to be informed that the use of hypnosis runs the risk that any testimony associated with hypnosis, even in a clinical and not a forensic context, may be inadmissible in court. The guidelines cited earlier of the American Society of Clinical Hypnosis include a sample informed consent form.
- Certain *personality characteristics* are associated with memory distortion. The Gudjonsson Suggestibility Scale (GSS) is standardized assessment instrument that assesses the degree to which a given individual is vulnerable to accepting misinformation suggestions, confabulating memory report, and shifting memory report in response to the emotional tone of an interview. Administering the GSS at the onset of a forensic examination allows a normative assessment of risk of memory distortion prior to any forensic interview. The Inventory of Childhood Memories and Imaginings (ICMI) evaluates fantasy proneness. About 3 to 5 percent of the general population is very high in fantasy proneness, which predicts significant internal confabulation of fantasy and memory. Empirically derived instruments like the Structured Inventory of Malingered Symptomatology (SIMS), the Malingering Probability Scale (MPS), and the Structured Interview of Reported Symptoms (SIRS) also help to identify the type of individual in a forensic setting that is likely to malinger psychiatric symptoms, and the Test of Memory Malingering (TOMM) and the SIMS help to identify individuals who specifically malinger memory symptoms. These instruments, albeit not foolproof, allow for reasonable assessment of a variety of situational and personality factors that predict high distortion of memory reports in a forensic context.
- *Establishing rapport,* making the patient feel at ease with the interview context, and building a relationship are important components of successful interviewing. The interview can begin with statements like, “I

want to get to know you; tell me about yourself, something that is important to you" This section of the interview also includes the rules of interview, such as the importance of telling the truth and being accurate; the permissibility of saying "I don't know" or "I don't want to say"; the fact that if the interviewer repeats a question, it does not mean that the previous answer was incorrect; and the fact that the interviewer really does not know what happened and must rely on whatever the interviewee is able to say about the target event.

- The preamble should set forth the *expectancies externally supplied by the hypnotist* regarding hypnosis, memory, and memory retrieval. The extent to which explicitly supplied explanations about hypnosis, memory, or memory retrieval overwrite the interviewee's *a priori* expectancies is unclear. Nonetheless, the forensic examiner should make explicit both the interviewee's a *priori* and explicitly suggested response expectancies.
- The preamble should include the *establishment of realistic expectancies* about memory retrieval ("You may or may not recall more about . . .") and the limits of accuracy ("Some of the things you may recall may be more and some less accurate") and should include explicit instructions not to guess ("Report everything that comes to mind but do not guess") and explicit instructions to foster an increase in, not a suspension, of critical judgment ("Throughout the procedure you will retain your best judgment as to which things you report are likely to be more and which are likely to be less accurate, knowing that not all details you report will be accurate").
- Next the interview focuses on *establishing a waking baseline memory* for a meaningful, neutral event from the same time frame (or developmental age) as the target event. In this, the section of the interview free recall is encouraged, "Tell me everything you can about . . ."
- The interview directly addresses the *target event* without explicit cuing. The interviewer must avoid leading the interviewee or being suggestive. Open-ended questions are desirable, such as "Tell me your understanding about what you're being interviewed about today." For interviewees who fail to disclose the target event to an open-ended prompt, the interviewer can state what he or she has been told, such as "Officer X told me that you might have witnessed an accident on [date] at [location]." Then, free recall is encouraged, with questions such as "Tell me whatever you can recall about [name target event], from the very beginning to the end."
- Field research on investigative interviews by child protective service workers or law enforcement agents with children suspected of being abused consistently has shown that the failure of many children to

disclose abuse in an initial interview is an artifact of the interview situation. A similar finding comes from field research on adult suspects during police interrogations. In such situations, forensic interviewers typically make the mistake of asking too many closed-ended or option-limiting questions at the onset of the interview.

A series of seemingly innocuous, factually based questions like "What is your name?" "Where do you live?" "What grade are you in?" establishes expectancies that discourage free recall and shapes the interviewee's behavior in the direction of only answering questions posed by the interviewer. Under such circumstances, the volume of free recall is significantly reduced, including much potentially valuable information about the target event. Such field research strongly shows that the greater amount of information and the greater accuracy of that information about the target event occur in the context of *extended free recall.* There should be at least two periods of extended free recall about the target event—one prior to and one following the hypnotic induction—followed by additional periods of extended free recall in hypnosis in subsequent sessions. Free recall is encouraged by such questions as, "Tell me everything you can remember about . . ."

- Hypnotic forensic interviewing is likely to be less controversial when the hypnotic induction conveys a *heightened focus or redistribution of attention* but avoids specific suggestions about trance characteristics, for example, involuntariness, the suspension of critical judgment, or dissociative compartmentalization that might bias retrieval.
- When strong emotional arousal, like fear, has interfered with retrieval of a memory for a target event, some evidence shows that the inclusion of *hypnotic relaxation* and instructions for focused recall increases the recall of meaningful information.
- After a hypnotic induction, it is best to allow the interviewee an extended duration of free recall of the target event after hypnotic induction. Frequent closed-ended questions should be avoided.
- The Cognitive Interview was developed by R. Edward Geiselman and Ronald P. Fisher as an alternative procedure to hypnosis for refreshing memory. It is based on basic principles of memory retrieval, with extended free recall and context reinstatement serving as the main components of the interview. More than 50 laboratory studies exist on memory retrieval with the Cognitive Interview. These studies consistently show that the interview yields around 25 to 35 percent of new information about a target event, typically without a significant increase in the memory error rate. Apart from the Cognitive Interview per se, these data strongly support the value of *context reinstatement along with free recall* as a primary memory retrieval technique. In hypnotic forensic interviews, there should be an extended period of free

recall, followed by context reinstatement plus free recall. The interviewee is reoriented to the time and place of the target event, asked to place him- or herself in the scene, and then to freely recall the event.

- Throughout the hypnotic interview, the interviewer should pay careful attention to the *type of interview questions* used. Instructions that encourage extended free recall are best. After an extended period of free recall, "wh-" questions are acceptable, like "what, where, when?" Research has shown that option posing and yes/no questions are of more limited value. The best guideline is to keep interview questions as open-ended as possible and select questions that pose the greatest number of options as answers.
- Introduction of follow-up *probe questions* occurs only after an extended period of free recall and context reinstatement in hypnosis fails to yield any more new information. Probe questions must avoid supplying new information, but instead must simply reflect back to the interviewee his or her words ("Earlier you said the assailant was 'tall'; tell me more about 'tall'").
- The entire session should be *videotaped*—not just the hypnotic session, but all of the preliminary explanations and the preamble, from the point that the interviewee and interviewer meet.
- The most important principle of any forensic interviewing is completely to *avoid leading and misleading questions*. The interviewer should not suggest or alter *any* content that has not already been reported by the interviewee.

There are no hard and fast rules as to when to use a hypnotic as compared to a nonhypnotic forensic interview. Reasonably prudent forensic examiners familiarize themselves with the case law on hypnotic admissibility for the jurisdiction is question. Strong case law in a state following a per se exclusion rule would merit caution. Cases entailing the highest standards of evidence, like reasonable doubt in a murder case, also merit caution in choosing a hypnotic forensic interview over the nonhypnotic Cognitive Interview.

Most experts further concur that conducting forensic hypnotic interviews with suspects of crimes fraught with difficulties should be avoided except under special circumstances. Hypnosis is indicated in cases involving full dissociative amnesia that have not been resolved by other methods.

The most frequent, justifiable use of forensic hypnosis is in interviewing witnesses and victims of crimes, not to "solve" the crime, but to provide additional evidence. William Wester and D. Corydon Hammond have reported that the most frequent use of forensic hypnosis is interviewing witnesses and victims of crimes not to solve the crime but to provide additional evidence. All evidence derived from hypnotically refreshed

testimony, as well as evidence from other, nonhypnotic procedures to refresh memory, like the Cognitive Interview, should not stand alone without independent corroborative evidence. The most common use of forensic hypnosis is in seeking new evidence about a crime from witnesses and/or victims.

The most famous case in which forensic hypnosis provided valuable evidence was the Chowchilla school bus kidnapping in the 1970s. Three assailants kidnapped a school bus of mixed-age children, buried the bus and the children alive in a rock quarry, and demanded a large ransom. The bus driver and some of the older children dug their way out by hand. Later the bus driver was hypnotized and asked about the license plate and other identifying details. Under hypnosis, the bus driver recalled a number of but not all of letters in the license plate, and recall sequence was inaccurate. Nevertheless, the recall of some of the letters accurately along with recall of other unusual features of the assailants' vehicle led to a correct computer identification of the vehicle.

Witnesses at the scene of sniper incidents and crime scenes involving suspected serial killers or rapists are interviewed by a team whose practice includes hypnotizing witnesses for the purpose of generating a composite drawing of the assailant to be circulated in the news media. This team's protocol has been effective in successfully apprehending a number of high-profile assailants. Most federal jurisdictions do not follow the per se exclusion rule and follow the guidelines approach. Federal guidelines have been developed for the forensic use of hypnosis.

FUTURE DIRECTIONS

From the days of early Chinese writings to the passes of Mesmer, the field of hypnosis has progressed to the science of the laboratory and to the clinician's armamentarium for use in a variety of arenas. The clinical application of hypnosis has been described with emphasis on the efficacy of hypnosis in the treatment of psychiatric disorders and other medical illnesses and problems. In view of the increasing desire for alternative medical approaches and the ability for hypnosis to be a milieu in which to explore the mind–body interaction, one can anticipate increased emphasis on evidence-based studies and controlled studies in both the basic science and clinical research. Hypnosis can be expected to become an increasingly more significant efficacious intervention in the treatment of psychiatric and other medical problems and illnesses.

REFERENCES

*Alladin A. Evidence-based hypnotherapy for depression. *Int J Clin Exp Hypn*. 2010;58:165.

American Psychological Association, Division of Psychological Hypnosis. Definition and description of hypnosis. *Hypn*. 1994;11:142.

Axelrad AD. The role of hypnosis in multiple convergent management of chronic pain. *Med.* 1990;6:111.

Baker EI. An hypnotherapeutic approach to enhance object relatedness in psychotic patients. *Int J Clin Exp Hypn.* 1981;24:136.

Barabasz AF, Barabasz M. Single-session manualized ego state therapy (EST) for combat stress injury, PTSD, and ASD, Part 1: The theory. *Int J Clin Exp Hypn.* 2011;60:381.

Barber TX, Calverley DS. The relative effectiveness of task-motivating instructions and trance-induction procedure in the production of "hypnotic-like" behaviors. *Nerv Ment Dis.* 1963;137:107.

Brown D, Fromm E. *Hypnotherapy and Hypnoanalysis.* Hillsdale, NJ: Lawrence Erlbaum; 1986.

Brown D, Fromm E. *Hypnosis and Behavioral Medicine.* Hillsdale, NJ: Lawrence Erlbaum; 1987.

Carolusson S. Dynamic hypnosis, IBS, and the value of individualizing treatment: a clinical prospective. *Int J Clin Exp Hypn.* 2014;62(2):145.

Castes M, Hagel I, Paleque M, Canelones P, Corao A. Immunological changes associated with clinical improvement of asthmatic children subjected to psychological intervention. *Behav Immunol.* 1999;13:1.

Clarke JC, Jackson A. *Hypnosis and Behavior Therapy: The Treatment of Anxiety and Phobias.* New York: Springer; 1983.

Crawford HJ. Brain dynamics and hypnosis: attentional and disattentional processes. *Int J Clin Exp Hypn.* 1994;42:204.

Daitch C. Hypnotherapeutic treatment of anxiety-related relational discord: a short-term hypnotherapeutic protocol. *Am J Clin Hypn.* 2014;56:325.

Dienes Z, Hutton S. Understanding hypnosis metacognitively: rTMS applied to the left DLPFC increases hypnotic suggestibility. *Cortex.* 2013;49:386.

Dillworth T, Jensen MP. The role of suggestions in hypnosis for chronic pain: a review of the literature. *Open Pain J.* 2010;3(1):39.

*Elkins GR, Barabasz AF, Council JR, Spiegel D. Advancing research and practice: the revised APA Division 30 definition of hypnosis. *Int J Clin Exp Hypn.* 2015;63(1):1.

Entwistle PA, Webb RJ, Abayomi JC. Unconscious agendas in the etiology of refractory obesity and the role of hypnosis in their identification and resolution: a new paradigm for weight-management programs or a paradigm revisited. *Int J Clin Exp Hypn.* 2014;62(3):330.

Ewer TC, Stewart DE. Improvement in bronchial hyper-responsiveness in patients with moderate asthma after treatment with a hypnotic technique: a randomised controlled trial. *Med J.* 1986;293:1129.

Gruzelier JH. Redefining hypnosis: theory, methods and integration. *Contemp Hypn.* 2000;17:51.

Hammond DC. *Hypnotic Suggestions and Metaphors.* New York: WW Norton; 1990.

Hammond DC, Garver RB, Mutter CB, Crasilneck HB, Frishholz E. *Hypnosis and Memory: Guidelines for Clinicians and for Forensic Hypnosis.* Des Plaines, IL: American Society of Clinical Hypnosis Press; 1995.

Hilgard ER. *Hypnotic Susceptibility.* New York: Harcourt, Brace & World; 1965.

Kallio S, Revonsuo A. Hypnotic phenomena and altered states of consciousness: a multilevel framework of description and explanation. *Contemp Hypn*. 2003;20:111.

Miller V, Whorwell PJ. Hypnotherapy for functional gastrointestinal disorders: a review. *Int J. Clin Exp Hypn*. 2009;57(3):279.

Moene FC, Spinhoven P, Hoogduin KA, van Dyck R. A randomized controlled clinical trial of a hypnosis-based treatment for patients with conversion disorder, motor type. *Int J Clin Exp Hypn*. 2003;51(1):29.

Morgan AH, Hilgard ER. The Stanford Hypnotic Clinical Scale for Adults. *J Clin Hypn*. 1979;21:134.

Nussbaum F, Redoute J, Le Barsa D, et al. Chronic low-back pain modulation is enhanced by hypnotic analgesia suggestion by recruiting an emotional network: a PET imaging study. *Int J Clin Exp Hypn*. 2011;59:27.

*Patterson DR, Jensen MP. Hypnosis and clinical pain. *Bull*. 2003;129:495.

Peter B, Vogel SE, Prade T, Geiger E, Mohl JC, Piesbergen C. Hypnotizability, personality style, and attachment: an exploratory study, Part 1—General results. *Am J Clin Hypn*. 2014;57:13.

*Raz A. Attention and hypnosis: neural substrates and genetic associations of two converging processes. *Int J Clin Exp Hypn*. 2005;53:237.

Shih M, Yang YH, Koo M. A meta-analysis of hypnosis in the treatment of depressive symptoms: a brief communication. *Int J Clin Exp Hypn*. 2009;57(4):431.

Spiegel D, Maruffi B, Frischholz EJ, Spiegel, H. Hypnotic responsivity and the treatment of flying phobia. *Am J Clin Hypn*. 2014;57:156.

Spiegel H, Spiegel D. *Trance and Treatment: Clinical Uses of Hypnosis*. Washington, DC: American Psychiatric Press; 2004.

Spiegel SB. Current issues in the treatment of specific phobia: recommendations for innovative applications of hypnosis. *Am J Clin Hypn*. 2014;56:389.

VanderLinden J, VanderReychen W. *Dissociation, and Impulse Dyscontrol in Eating Disorders*. New York: Taylor Francis-Brunner/Mazel; 1997.

Wagstaff GF, Wheatcrogt JM, Jones C. Are high hypnotizables especially vulnerable to false memory effects? *Int J Clin Exp Hypn*. 2011;59:310.

Wagstaff GF, Wheatcroft JM, Caddick AM, Kirb LJ, Lamont E. Enhancing witness memory with techniques derived from hypnotic investigative interviewing: focused meditation, eye-closure, and context reinstatement. *Int J Clin Exp Hypn*. 2011;59:146.

Wain HJ. Hypnosis on a consultation-liaison service. 1979;20(10):678.

Wain HJ. Reflections on hypnotizability and its impact on successful surgical hypnosis: a sole anesthetic for septoplasty. *Int J Clin Exp Hypn*. 2004;46(4):313.

Wain HJ, Amen DG. Emergency room use of hypnosis. *Hosp Psychiatry*. 1986;8:19.

Wain H, Perito R. "Care of Combat Injured Service Members" In: Cozza SJ, Goldenberg MN, Ursano RJ, eds. *Care of Military Service Members, Veterans, and Their Families*. Washington, DC: APPI Press; 2014.

Watkins JG. *Techniques: Clinical Hypnosis*. Vol. 1. New York: Irvington; 1987.

Wester WC, Hammnod DC. Solving crimes with hypnosis. *Int J Clin Exp Hypn*. 2011;53(4):255.

*Whorwell PJ. Effective management of irritable bowel syndrome—the Manchester model. *Int J Clin Exp Hypn*. 2006;54:21.

▲ 33.5 Group Psychotherapy

HENRY I. SPITZ, M.D.

INTRODUCTION AND OVERVIEW

Group psychotherapy has always been an evolving field: sensitive to changes in the cultural climate, reactive to the lessons learned from history, cognizant of world events, aware of developments in scientific research, open to new systems of diagnosis of emotional illness, and updating theoretical orientations and clinical technique in the spirit of providing a level of care that is relevant, effective, and capable of being incorporated into new health care delivery systems.

Over the past decade, a range of issues have been integrated to expand group therapy themes. Examples include the results of trauma from sources as broad as PTSDs in soldiers returning from war and others who have experienced childhood psychological, physical and sexual abuse, gender issues including homosexuality, transgender, lesbian and bisexuality, Internet and computer themes including Internet addictions, pornography, gambling and handheld phones, leading to interpersonal problems, new discoveries in brain structure, genetics, disease, and the course and therapy of central nervous system disease; BPD has been studied and DBT has become a successful intervention.

Groups for schizophrenia and serious inpatient and outpatient psychological conditions have earned prominence with the emphasis on cognitive problems, behavioral skills, and socialization learning for its members. Relationship problems of many kinds are common as well. Heterogeneous couples have witnessed an increased divorce rate and often seek group therapy as a useful tool in exploring their marriage. New relationship forms that have emerge which include same sex couples, families that consist of a single parent and homosexual parents.

Group therapy as a core element in the treatment of substance abuse and alcoholism continues to develop. Different group approaches and treatment opinions vary from classic Twelve Step or Alcoholics Anonymous groups to midrange groups as in Recovery Incorporated and lastly a harm reduction group model where members are taken into group without present sobriety being a requirement.

The list of variety of group approaches is wide ranging. Medical illness, affective disorders, bereavement and isolation, eating disorders, psychopharmacological problems, geriatric groups, combined individual and group treatment, self-help groups, adolescent, and family groups are a brief overview to set the stage for what follows in this chapter.

HISTORY OF GROUP PSYCHOTHERAPY

The field of group psychotherapy can be conceptualized as being in a new phase of its modern history. The first wave, ushered in by World War II and the enormous amount of psychological carnage that emerged from the War, created a pressing need for treating the psychiatric casualties of wartime. Working in a group setting, as opposed to the conventional individual psychoanalytic milieu, facilitated seeing more people in less time and led to an exponential increase in the growth of psychotherapy groups predominantly those of a psychodynamic nature.

The paradigm for phase two of group therapy's development was the community mental health movement in the late 1960s and 1970s. The mandate to deinstitutionalize the chronically mentally ill patient from state hospitals back into the community presented a set of circumstances similar to those following the World War II—a surplus of those in need of psychological help and a shortage of professionally trained therapists to treat them. The practice of group therapy particularly in the areas of homogeneous groups designed to treat members who shared a common condition expanded during this period. With greater comfort and familiarity in employing group techniques came a corresponding trend toward experimentation with the application of small and large group methods to meet the needs of a broad range of patient populations and psychiatric conditions.

The 1990s was the age of health care reform. With it again came the pressure to see more people in shorter time frames, but the mandate this time had come from a different source: managed mental health care. These four emotionally charged words reflect an economically motivated model of health care delivery that places a high premium on time and cost-efficiency often at the expense of good patient care. Despite its many drawbacks, managed care allowed creative group therapists to expand the current knowledge about brief, symptom-focused treatments and the use and possible hazards of time-limited groups.

The first decade of this century has borne witness to the fourth phase in group work ushered in by the rising threat of terrorism, human and natural disasters, and consequently to the influx of group therapies designed to address issues of psychological trauma. The immediate and longer-term reactions following catastrophic events commonly include panic, emotional anesthesia, alienation, social isolation, a profound sense of vulnerability, and a host of symptoms, some of which result in true PTSD syndromes. The group setting is ideal for help in healing those whose lives have been touched by traumatic events.

DEFINITION

Group psychotherapy is a modality that employs a professionally trained

leader who selects, composes, organizes, and leads a collection of members to work together toward the maximal attainment of the goals for each individual in the group and for the group itself. Certain properties present in groups such as mutual support can be harnessed in the service of providing relief from psychological suffering and supply peer support to counter isolation experienced by many who seek psychiatric help.

Similarly, homogeneously composed small groups are ideal settings for the dissemination of accurate information about a condition shared by group members. Medical illness, substance abuse, and chronic and persistent severe psychiatric conditions including schizophrenia and major affective disorders are cases in point.

HISTORICAL ROOTS

Group therapy has a long and fascinating history spanning more than 100 years. Early efforts from varied sources are directly traceable to elements of clinical practice today.

Most historians agree that the first recognizable group intervention was conducted in 1905 by Joseph Pratt, an internist treating tuberculosis patients in Boston. The treatment of tuberculosis at the turn of the century consisted of isolating afflicted patients from the general population and housing them in long-term care facilities. Pratt conducted "classes" on the ward designed to educate patients about what was known at the time about their disease. A serendipitous observation by Pratt was that people who attended class regularly seemed to be more compliant with treatment protocols of that era and also seemed to have improvement in their state of depression.

Parallel efforts took place in 1921 when Edward W. Lazell applied the group method to schizophrenic and manic-depressive patients at St. Elizabeth's hospital in Washington, D.C., and observed similar results. Patients who participated in group sessions appeared to be less depressed and isolated as well as being more accurately informed about their psychological condition and the rationale for the nature of their therapy.

At the same time, Alfred Adler and Rudolph Dreikurs departed from the exclusively psychoanalytically based dyadic models of individual psychotherapy by involving significant others in the lives of the children they were treating at the Vienna Child Guidance Clinic. Parents, schoolteachers, and clinic staff were directly involved in small group sessions designed to integrate the therapy of the children and adolescents with whom they were connected.

L. Cody Marsh, a minister working at Kings Park State Psychiatric Hospital in New York, used a revivalist-inspirational style to help bridge the alienation he felt was at the root of many of the psychiatric problems of the time. Marsh's now famous quote "By the crowd they have been broken, by the crowd they shall be healed" captures the spirit of his orientation to

mental illness and his emphasis on the need for the social support provided by therapy groups in helping patients cope with their problems.

The 1920s saw a confluence of interest in group dynamics and early group therapies. Sigmund Freud addressed the properties of large groups and their potential for good or dangerous influence in "Group Psychology and the Analysis of the Ego." Jacob Moreno, the founder of psychodrama, was working in Vienna experimenting with the use of intense affect and physical activity in a group setting when he created his "Theater of Spontaneity." In addition to creating psychodrama, Moreno is also cited as the first person to use the term "group psychotherapy" in conjunction with his applications of the psychodramatic technique.

In 1925, Trigant Burrow applied the group method to psychiatric outpatients and is credited with creating the term "group analysis" to describe his technique. In addition, Burrow was interested in the dynamics of small groups and created some of the earliest group "laboratory" formats in which the task of the group was to examine its own internal interaction. Burrow's work represented an extension of Freud's individual focus to the larger social context in which an individual existed.

The 1930s were characterized by efforts to expand the interface between psychoanalysis and group therapy. Louis Wender, in the early 1930s, formed six- to eight-member groups of patients and met with the group two or three times per week. Paul Schilder, working at Bellevue Hospital in New York City, was also enthusiastic about outpatient psychoanalytic groups for patients with a broad range of emotional problems.

A key figure in the evolution of group psychotherapy for children and adolescents was Samuel Slavson. Slavson is considered by many to be the founder of modern group psychotherapy. Slavson focused on both the individuals and the group itself. His innovative work with activity and play group therapies was based heavily on using the therapy group as a vehicle for observing behavior as participants interacted in group sessions. He combined an analytic superstructure to his observations and utilized the analysis of transference and resistance as mainstays of his technique. Slavson founded the American Group Psychotherapy Association and was also the organization's first president.

The marriage of psychoanalysis and group psychotherapy as practiced by contemporary psychodynamically oriented clinicians is closely associated with the work of Alexander Wolf. Wolf was intrigued by the challenge of adapting the tools of psychoanalysis to the group setting. He believed the depth and availability of transference was accentuated in groups largely because the presence of other people stimulated feelings related to the families of origin of group members. He used dream interpretation, free association, and other analytic staples in innovative ways. For example, the "go-around" or "round robin" technique used in groups to include all members was originally born out of Wolf's efforts to

use free association in the group session. He began a session by asking a member to report a dream and then would ask all other group participants to share their own spontaneous associations to the elements contained in the dream.

Significant theoretical and applied group work was taking place in Europe during the late 1930s and through the 1940s following the World War II. Wilfred Bion, Henry Ezriel, S.H. Foulkes, and E. James Anthony all made seminal contributions to the understanding of the dynamics of groups. Bion focused on the "group as a whole" and delved into the unconscious life of the group using an object relations perspective based in the work of Melanie Klein. Concepts like the "basic assumption group" emanated from Bion's studies of group behavior and have led to a popular subgroup of modern group therapists that subscribe to Bion's precepts.

Ezriel agreed with many of Bion's ideas but emphasized the importance of focusing on the individuals in the group and their unique concerns as well as concentrating on the whole group level. To do so, Ezriel drew attention to interactions and the member-to-member transferences in the "here-and-now" of the group. In the same vein, Foulkes, a British psychoanalyst, coined the term the "group matrix" that reflected his view of the group as a nexus of communication shared by all members. The matrix took into account the cultural similarities and differences among members along with their styles of communication as reflected in their group interactions. Foulkes' work was significant in integrating historical material with their manifestations in the group process itself thereby bridging past and present. Foulkes along with E. James Anthony and others were instrumental in forming the Group Analytic Society in England.

In the early 1950s, another British figure, Maxwell Jones, created the "therapeutic community" concept. Jones's ideas were embraced and adapted by clinicians working with the chronic and persistent mentally ill and those dealing with problems related to substance abuse.

The history of group therapy incorporates the contributions of social scientists like Kurt Lewin. Lewin's field theory helped solidify the view that the group was greater than the sum of its individual parts and that groups were social systems, a concept that had great utility for clinicians in both group and family therapies. The term "group dynamics" was introduced by Lewin.

From the 1950s through the present, enthusiasm for experimentation with the application of group models led to further awareness of the uses and potential liabilities of experiences in a group setting. Among the former was the work of Nathan Ackerman who made noteworthy contributions at the interface between group and family therapy when he began seeing whole families as a unique small group in the therapeutic setting.

Simultaneously, the National Training Laboratories conducted community groups in Bethel, Maine. These groups formed the prototype

for the T-group or training group, encounter, and sensitivity groups that flourished through the 1960s. Important lessons about the risk factors in groups and the psychiatric sequelae that followed directly from passive or charismatic leadership.

Exponents of new "schools" of group therapy began to emerge. Eric Berne developed transactional analysis, Fritz Perls designed groups based on Gestalt therapy theory, and Robert Goulding's redecision group therapy are but a few of the burgeoning points of view leading to the broad array of theoretical points of view represented in current group therapy practice.

The appreciation of the small and large group setting as a vehicle for the communication of information about a condition or a set of goals or problem sets shared by its members is as evident in the work of these early practitioners as it is in present day group work.

THEORETICAL ORIENTATIONS

Many individually derived forms of psychotherapy have their group counterparts. A representative sample of major theoretical orientations toward group therapy can be broadly subsumed under the psychodynamic, interpersonal, and cognitive-behavioral schools of thought.

The *psychoanalytic spectrum* has evolved to include more than classical Freudian psychoanalytic theory. Exponents of the British school of object relations based on the work of Melanie Klein, Ronald Fairbairn, D.W. Winnicott, Harry Guntrip, and other self-psychologists utilize the unique properties of the therapy group to observe and alter the behavioral and ideational manifestations that originate in the inner world of the group member. Group members function as or represent part objects that parallel those that exist in the mind of an individual and are projected onto and enacted within the interactions in the group. In this way, the group provides an unparalleled setting for behavioral information about how group members perceive and misperceive their relationships to fellow group members and to the group leader. Practitioners with this orientation use the group to highlight unresolved issues involving themes of emotional attachment, acceptance and rejection, managing intense affect, and tolerance of frustration.

Ego psychologists also fall within the penumbra of modern psychoanalytic practices in group therapy. Their theoretical base builds on earlier work by Heinz Hartmann, Ernst Kris, and David Rapaport who viewed Freud's structural theory of mental function into id, ego, and superego components as incomplete. They postulated that a conflict free or adaptational portion of the ego exists as well. In clinical group work, ego psychologists rely on an appreciation of the mechanisms of defense that members use to counter their anxieties, the utilization of the peer and authority transferences in the group, and on the countertransference reactions of the group therapist.

Group therapy is an *interpersonal* form of treatment. Interpersonal theorists, most notably Irvin Yalom, advocate a group approach that centers on the interaction of group participants in the "here and now" transactions in the group session. Yalom also added universal existential dilemmas to a method that calls for active group leadership in creating an early, cohesive group framework.

Psychodrama is perhaps the most physically active form of group therapy. By employing structured dramatic activities in a group setting, the "director" or group leader guides the "protagonist" (individual with an issue to be resolved) through a dramatization that is followed by feedback from other group members.

Cognitive and behavioral approaches rely primarily on learning theory and experiential elements to reach their goals. Structured, often manual-driven protocols are utilized in a time-limited group composed of members with a shared set of problems. Cognitive approaches are the most standardized of all group interventions and are, therefore, the types of groups that are most easily measured in terms of group process and outcome.

Systems oriented group approaches have found a bona fide place in group therapy as they have in family therapy. The assumption that the therapy group is an entity larger than the sum of its parts and that each member plays a role in the success or failure of the group enterprise is a central core to the systems approach.

Lesser-known but still actively practiced orientations include *Gestalt, redecision, and transactional analysis group therapies*. The first, based on Kurt Lewin's field theory and the work of Fritz Perls and Paul Goodman enjoyed great popularity during the sensitivity and encounter group movement in the late 1960s and 1970s. A central tenet of the Gestalt approach is the emphasis on group process over group content and efforts to mobilize strong affect that stands in the way of a person actualizing his or her full potential. The group focuses is on emotional experience in the present moment and employs a variety of active, noninterpretive exercises designed to heighten self-awareness and to facilitate change.

Closely allied to the Gestalt method is redecision group therapy, a short-term model created by Robert and Mary Goulding as an offshoot of their work with Eric Berne and Fritz Perls. The format encompasses the creation of a clear behavioral contract at the outset. The group therapist aligns with the client's "authentic self" or that part of the person that is searching for new solutions to old impasses. The model eschews the use of the psychodynamic techniques such as the use of interpretation and the analysis of transference and resistance relying instead on an array of redecision interventions that aim at identifying maladaptive, rigidified, unconscious defenses, and interrupting their reflexive nature in order to substitute more functional responses and problem-solving behaviors.

Transactional analysis, commonly known as TA, is a component of redecision therapy but is also a stand-alone form of group therapy. Its founder, Eric Berne, replaced Freud's structural theory of the psyche from id, ego, and superego to what he termed "ego states" of Parent ("exteropsychic), Adult ("neopsychic"), and Child ("archaeopsychic").

Berne codified a series of dysfunctional interpersonal patterns into what he identified as "games" that people engage in repetitively without success. The goal of the therapy group was to identify transactional states among group members and to change childhood "life scripts" into more adaptive adult patterns.

Table 33.5–1 illustrates some of the similarities and differences among representative group formats.

SELF-HELP GROUPS

Leaderless or self-help group experiences usually homogeneously composed of participants with a common issue represent an additional use of the group setting as a vehicle for change. Although these groups are not strictly considered a form of psychotherapy, they do share a number of parameters and can have therapeutic benefit for participants. Perhaps the best-known examples are the Twelve Step groups of Alcoholics Anonymous (A.A.), Al-Anon, Adult Children of Alcoholics (A.C.O.A.), and other substance abuse centered groups. In addition, many self-help groups are formed around health and illness, loss or a shared life experience. Cancer support groups, groups for widows or widowers, parents without partners, and bipolar support groups are but a few well-known self-help groups.

THERAPEUTIC POTENTIAL OF GROUPS

Therapy groups contain special properties that can be utilized in the service of change. Some of these factors are unique to the group milieu and others are amplifications of more generic components of all effective psychotherapies. For many years, researchers and practitioners have tried to identify those therapeutic or "curative factors" present in groups in order to determine how they can be used to maximum effect.

Many authors have offered systematized categorizations of the natural resources available in therapy groups over time. Popular examples of these efforts can be seen in the work of Yalom and Leczsz, MacKenzie, and Sadock. Yalom and Leczsz use a 13-point model that includes universality, altruism, instillation of hope, imparting information, corrective recapitulation of the primary family experience, development of socializing techniques, imitative behavior, cohesiveness, existential factors, catharsis, interpersonal learning (both input and output), and self-understanding.

MacKenzie divided the therapeutic factors into four groups: supportive factors, learning factors, self-revelation factors, and psychological work factors. The supportive factors include universality, acceptance, altruism,

and hope as contrasted to learning factors such as modeling, vicarious learning guidance, and education. Self-disclosure and catharsis comprise the factors of self-revelation, while interpersonal learning and insight round out the psychological work factors in this schema.

Sadock has delineated a more extensive description and definition of therapeutic factors in Table 33.5–2.

Depending upon the nature of the group and its stated goals, the group leader draws upon particular therapeutic factors in order to enhance the likelihood of group members deriving maximum benefit from group therapy. As an example, in the early stages of an adult psychotherapy group, the group leader emphasizes those factors that facilitate the creation of a rapid sense of group cohesion. In practice this means relying on similarities among members, creating an emotionally safe group structure, using the support elements available to the group, avoiding premature confrontation by the therapist, and monitoring intermember hostility.

Table 33.5–1.
Comparison of Major Group Therapy Orientations

	Supportive	Psychodynamic	Self-Help	Cognitive/Behavioral
	Group therapy	Group therapy	Groups	Group therapy
Frequency	1–5 times/wk	1–2 times/wk	7 days/wk	1–3 times/wk
Individual screening	Usually	Always	None	Always
Group size	8–15 members	5–9 members	No size limits	5–10 members
Goals	Better adaptation for daily living	Reconstruction of personality dynamics	Social support	Relief of specific symptoms
Indications	Crisis situations; severe emotional	Neuroses; mild personality disorders	Shared life experiences	Phobias; anxiety disorders
Group composition	Homogeneous for level of psychopathology	Balance of similarities and differences	Homogeneous	Homogeneous for similar symptoms
Group focus	"Here and now"; family, vocational, environmental factors	Past and present; intragroup and extragroup dynamics	Education; emotional "sharing"	Training in methods that control symptoms
Use of confrontation	No	Yes	No	Rarely
Therapist activity	Actively structures and leads the group	Active around interpretation	No formal group leader	Very active in teaching skills to patient
Extragroup contacts	Encouraged	Prohibited or discouraged	Encouraged Formal peer sponsorship	Discouraged
Transference	Not analyzed	Utilized extensively	Not analyzed	Not relevant to the group work

Therapeutic factors	Cohesion; universality; reality testing instillation of hope imparting information	Cohesion; catharsis; family replay; interpersonal learning	Cohesion; universality; education; peer support	Cohesion; universality; education; reinforcement

SELECTION OF GROUP MEMBERS

The extensive application of diverse group therapies over time has resulted in the liberalization of the inclusion and exclusion criteria for group membership. The selection process is a three-dimensional process consisting of evaluating individuals to see if they are appropriate for group therapy in general; member-to-group matching, does the suitable group patient fit into the group for which he or she is being considered; and therapist-to-group matching, does the leader possess the clinical skills and emotional temperament necessary to understand and lead the group in the most effective way.

Traditionally, the group therapy literature has been more specific about exclusion criteria than inclusion criteria. Contemporary group therapy has been significantly modified from its historical roots in order to incorporate a broader range of difficult personality styles, severe psychopathological states, and interpersonal deficits. Still, there remains a nucleus of patients for whom group therapy is either contraindicated or the timing of their entry into group treatment is premature.

Among these candidates for exclusion are those inpatients who are acutely psychotic and not stabilized on psychotropic medication, people who are acutely suicidal or for whom the high level of affect in the group might exacerbate their symptoms, and those whose primary language is not the language of the group. Debate continues about whether some conditions formerly considered for exclusion from group therapy can be treated in specialized groups in venues where the leader and the system are tolerant and familiar with the condition shared by members. A person with diagnosis of antisocial personality disorder was virtually certain to be excluded from group therapy however; groups in prison settings and many substance abuse groups are examples of the possibility of having homogeneous groups of members with a diagnosis heretofore thought to be unworkable in the group milieu.

Ideally, the person well suited for inclusion in group therapy is one who presents with an interpersonal problem and possesses sufficient motivation to work on the manifestations of these issues as they play themselves out in the ongoing arena of the group. For interactive, outpatient groups, members must have the capacity to empathize with others, possess sufficient ego strength to tolerate not being the exclusive focus of therapy, have some capacity for self-awareness or psychological-mindedness, be willing to self-disclose in the group meetings, and to benefit from the reality testing and interpersonal feedback supplied by the group therapist

and group members.

In actual clinical practice, many people are referred for group therapy precisely for the very reason that they lack some or all of the aforementioned qualities. This should not be a de facto reason for exclusion from group membership. Pre-group alternatives like participation in a waiting-list group or a short-term orientation to group psychotherapy experience can give the therapist and the patient an expanded opportunity to see a prospective member in action. In so doing, it may become clear that the patient will either not be able to derive benefit from the group or will be a disruptive or undermining influence on the group in which case he or she will be deemed inappropriate for inclusion. However, pre-group participation can be invaluable in helping the group therapist design a treatment plan that is more patient-specific based on direct experience of the patient's interpersonal interactions.

Objective measures such as the Group Therapy Questionnaire and the Group Selection Questionnaire, both sections of the comprehensive revised CORE Battery, yield important data derived from patient self-reports about variables relevant to the decision about who can is likely to flourish in groups and who should be excluded from group therapies. Once the decision about inclusion has been made, the critical phase of pre-group preparation of the prospective group member follows next.

Table 33.5–2.
Therapeutic Factors

Factor	Definition
Abreaction	A process by which repressed material, particularly a painful experience or conflict, is brought back to consciousness. During the process, the person not only recalls, but also relives the material, which is accompanied by the appropriate emotional response; insight usually results from the experience.
Acceptance	The feeling of being accepted by other members of the group; differences are tolerated, and there is an absence of censure.
Altruism	The act of one member's being of help to another; putting another person's need before one's own and learning that there is value in giving to others. The term was originated by Auguste Comte (1798–1857), and Freud believed it was a major factor in establishing group cohesion and community feeling.
Catharsis	The expression of ideas, thoughts, and suppressed material that is accompanied by an emotional response that produces a state of relief in the patient.
Cohesion	The sense that the group is working together toward a common goal; also referred to as a sense of we-ness; believed to be the most important factor related to positive therapeutic effects.
Consensual validation	Confirmation of reality by comparing one's own conceptualizations with those of other group members; interpersonal distortions are thereby corrected. The term was introduced by Harry Stack Sullivan. Trigant Burrow had used the phrase consensual observation to refer to the same phenomenon.
Contagion	The process in which the expression of emotion by one member stimulates the awareness of a similar emotion in another member.

Corrective familial experience	The group recreates the family of origin for some members who can work through conflicts psychologically through group interaction (e.g., sibling rivalry, anger original toward parents).
Empathy	A capacity of a group member to put himself or herself into the psychological frame of reference of another group member and, thereby, understand his or her thinking, feeling, or behavior.
Identification	An unconscious defense mechanism in which the person incorporates the characteristics and qualities of another person or object into his or her ego system.
Imitation	The conscious emulation or modeling of one's behavior after that of another (also called role modeling); also known as "spectator therapy," as one patient learns from another.
Insight	Conscious awareness and understanding of one's own psychodynamics and symptoms of maladaptive behavior. Most therapists distinguish two types: (1) intellectual insight—knowledge and awareness without any changes in maladaptive behavior and (2) emotional insight—awareness and understanding leading to positive changes in personality and behavior.
Inspiration	The process of imparting a sense of optimism to group members; the ability to recognize that one has the capacity to overcome problems; also known as installation of hope.
Interaction	The free and open exchange of ideas and feelings among group members; effective interaction is emotionally charged.
Interpretation	The process during which the group leader formulates the meaning or significance of a patient's resistance, defenses, and symbols; the result is that the patient develops a cognitive framework within which to understand his or her behavior.
Learning	Patients acquire knowledge about new areas, such as social skills and sexual behavior; they receive advice, obtain guidance, and attempt to influence and are influenced by other group members.
Reality testing	Ability of the person to evaluate objectively the world outside the self; includes the capacity to perceive oneself and other group members accurately.
Transference	Projection of feelings, thoughts, and wishes onto the therapist, who has come to represent an object from the patient's past. Such reactions, although perhaps appropriate for the condition prevailing in the patient's earlier life, are inappropriate and anachronistic when applied to the therapist in the present. Patients in group therapy may also direct such feelings toward one another, a process called multiple transferences.
Universalization	The awareness of the patient that he or she is not alone in having problems; others share similar complaints or difficulties in learning; the patient is not unique.
Ventilation	The expression of suppressed feelings, ideas, or events to other group members; the sharing of personal secrets that ameliorates a sense of sin or guilt (also referred to as self-disclosure).

PRE-GROUP PREPARATION

All prospective group members must receive a pre-group orientation. Both the individual and group psychotherapy literature attest to the observation that the better prepared a patient is for psychotherapy, the better the outcome. In group therapy, the pre-group phase has been shown to be particularly important in reducing group drop-outs, diminishing patient anxiety about group entry, having members with more realistic expectations of group therapy, enhancing motivation for therapy, avoiding

goal incompatibility among members, and building a strong therapeutic alliance.

The use of a pre-group orientation session with each incoming group member affords the group therapist an ideal opportunity to accurately inform members about the parameters of the group they will be entering and to anticipate early resistances and behaviors that require sound clinical management in order to prevent avoidable drop-outs from group. There are many ways to prepare new group members ranging from informal discussions with patients to structured pretraining programs, the distribution of written orientation manuals, employing videotapes of simulated group sessions, group orientations involving many new group members simultaneously, and others. Regardless of the format of the orientation to group, the group leader is well advised to include an organized pre-group preparation component as a routine part of the pre-group phase of group therapy. The use of a group therapy orientation checklist is one method of addressing the central elements of pre-group preparation of group therapy participants. Fourteen items comprise the pre-group orientation checklist. Depending upon the type of group and the personality of the incoming member, group leaders have the option to focus more in-depth on those areas on the list that may be most specific to the group and the patient in question and less on the topics that are less applicable to the situation.

The pre-group session not only addresses the "nuts and bolts" of the administrative and logistical aspects of groups but it is simultaneously the start of the therapeutic process. When viewed from this perspective, the group therapist can accomplish the dual goals of preparing patients for group entry and get a preview of individual issues that will form the therapeutic agenda for a group member when he or she enters the group.

1. *Goals and general purpose of the group*. The therapist clearly states the purpose and goals of the group. Goals that are vague or too ambitious are to be avoided in favor of goals that are patient and group-specific.

2. *The composition of the group*. The therapist tries to convey a picture of the group the patient will be entering without disclosing information that violates the confidentiality of other group members. The goal is to provide the member with an understanding of who will be in the group, why specific members have been selected, and what the size of the group will be.

3. *The role of the group leader*. A key concern for new members of any group revolves around what to expect from the group leader. Members who are used to individual psychotherapy may be anxious about sharing the therapist with other group members, puzzled about how help can come from others who themselves have significant emotional problems, and wondering how much reliance they can place upon the group leader.

What the group leader actually tells members in the orientation

dialogue is aimed at addressing common patient concerns about leadership activity versus passivity; how the relationship with the leader in the group relates to the kind of help they can expect from group therapy; how structured the group will be; will the group leader be available for individual sessions outside the group; and a host of recognizable pretherapy fears that affect new patients in groups.

4. *Observing and recording group meetings.* There are settings and circumstances that may call for groups to be observed. In training programs, observation and recording of group sessions are primarily used for the benefit of teaching therapists a particular technique or method. There are also times when recording, especially videotaping of meetings, is used as part of the group therapy itself, as well as a vehicle for training, supervision, and quality control of therapeutic groups. Observation of any kind must be handled openly before the first session and with the patient's informed consent.

5. *The arrangement of the therapy setting.* When working with more seriously disturbed, suspicious, paranoid patients, or with group members who display excessive amounts of anticipatory anxiety, a simple statement about, and a short visit to, the room in which the group will be meeting helps allay many fears in the minds of these patients.

During or after the visit to the group therapy room, most people will report some sense of relief, but some will describe the experience in terms that are clinically revealing. Initial reactions like, "The room feels claustrophobic to me." or "The chairs seem so close to each other; can people actually touch you in the group meeting?" are cases in point.

6. *Time factors in the group.* The time frame of a group correlates directly with the purpose of the group. Time varies according to the group goals. As a general rule, short-term groups have a fixed time limit and aim to change maladaptive behaviors and/or to teach specific skill acquisition. Ongoing or open-ended groups are more ambitious and aspire to reach many of the same goals as their individual psychotherapeutic counterparts. Personality and attitudinal shifts as well as behavioral change are part of the long-term group model.

Considerations about the optimum time frame for each group session are also addressed in the pre-group orientation. The group parallel of the 45- or 50-minute individual session is the 90-minute group meeting. Examples of variations in the standard time frame designed to meet the unique needs and tolerance level of members include a brief cognitive/behavioral group that meets for 1 hour sessions over the course of 12 weeks and a 30- to 45-minute session group for psychiatric inpatients who may not be able to sit still or tolerate the level of affect that often accompanies a group of longer length.

7. *Loss and addition of group members.* Every group requires a sufficient number of members in order to reach its stated objectives. Some

groups, usually those of a short-term nature, start and end with the same members. If, in the course of the time, members drop out of brief therapy groups, they are not replaced. Many short-term therapy groups have a structured protocol for the conduct of each session and groups of this type do not allow for any changes in membership.

Long-term groups employ an open membership model that assumes that members will be leaving the group over the course of its existence. The two primary ways in which members leave a group are by terminating prematurely for psychological or practical reasons or through the realization of their treatment goals. In either case, gaps are left in the group.

Issues of timing are important when considering additions to the group. Long-term dynamically oriented groups do not seek immediate replacements for a departed member. Whatever the reason for a member leaving the group, it still is experienced as a loss by the group membership. The leader needs to allow time for the group to deal with the issues ushered in by the loss of a group member. Powerful emotional themes related to death, abandonment, separation, competition, family re-enactment, and loss are but a few of the feelings that accompany the exit of a member who has been in group for a long time. The dynamically inclined therapist in a long-term psychotherapy group can make extraordinarily productive use of the group members' unresolved conflicts related to loss, feelings emanating from inclusion and exclusion, and sibling relationships.

These potent emotions, most of which are historically based, get replayed in the group and are available for therapeutic exploration. Neglecting to address the issues related to loss and addition of group members invites avoidable problems in the ongoing conduct of the group.

8. *Rules about attendance.* All groups strive to establish regular attendance as a central group norm. It is a primary leadership responsibility to advocate and model regular group attendance. In so doing, the leader makes a position statement by his or her words and actions about the importance of the group and how vital each person is to the endeavor. Instructing members to call prior to a session if they are going to be absent or late is also discussed in this phase of the pre-group orientation.

Members in groups where absences and lateness are tolerated tend to devalue the group experience, lose confidence in the leader whom they regard as weak, passive, or inadequate, and feel like they are participating in a failure. All of these factors are, at the very least, demoralizing but they may actually be injurious to patients who come to group with pre-existing feelings of low self-esteem, a sense of personal failure, and a reasonable expectation that their lives will be enhanced, not made worse, by their joining a therapy group.

When working with very seriously ill psychiatric patients, rigid

restrictions on attendance may be asking patients to comply with a task that is beyond their adaptive capacity. Therapists who work with schizophrenic outpatient groups can modify this portion of the group preparation module to allow for some expected irregularity in attendance. If the patient can make an attachment to the group and the system or hospital where the group meets and remain with it over time, these patients may have poor per session attendance records but those who regard themselves as members of the group or members of the program do well over the longer life span of the group.

Higher functioning patients can be held to a higher standard of attendance. For purposes of the pre-group preparation of incoming members, the group leader decides in advance of the group what the policy on attendance will be and orient all members in a uniform way.

9. *Fees and billing procedures*. The mechanism by which therapy will be paid for is included in the orientation session. Issues of precertification for sessions, proper referral procedure, and compliance with the criteria set forth by third party payers form the essentials of this process.

The setting of fees has significant clinical implications. Group leaders must decide whether it suits the goals of the group to have a standard fee for everyone in the group or, in cases where finances are a problem for a group member, should there be fee adjustments or a sliding fee scale. Whatever choice is made, it is important for the group leader to be clear about the fee policy for each incoming member and to anticipate the dynamic issues that graft themselves onto the issue of money.

It is common in psychodynamic groups where members pay different fees for group members to feel that the therapist prefers the people who are charged the lowest fees. This is a group replay of unresolved family issues of parental preference of one sibling over another displaced onto what appear to be practical matters involving money. The alert group leader can differentiate genuine issues about money from other psychodynamic and group process issues if he/she takes care to explain to each member the policy for group fees during the orientation meeting.

10. *Coordination of other therapies*. Group therapy is frequently part of a broader treatment plan for patients. The group leader has the responsibility for coordinating the treatment of each group member. In the pre-group orientation session, the therapist goes over other concurrent or potential treatment efforts and discusses how each will be handled.

This aspect of the pre-group member preparation encompasses several important considerations. If the group member is also in another form of psychotherapy, most usually individual, couples, or family therapy, the group leader has to get permission from the patient to contact the other therapist(s) involved in the patient's care at any point the group leader deems necessary. In this way, not only can the group therapist's database of information be expanded, but some potentially troublesome clinical issues,

such as the use of “splitting” defenses by group members with borderline personality disorders, an attempt that aims to divide members of the treatment team.

The two other areas to be covered in this portion of the orientation concern medication management and the provisions for hospitalization of patients should that become necessary. When someone other than the group leader is prescribing psychotropic medication for a group member, the leader has to be knowledgeable about the drug regimen. At the same time, the clinician providing the medication can obtain valuable information referable to the patient’s response to drug therapy as demonstrated by his individual and interpersonal actions in the group.

When a group member has a history of prior psychiatric hospitalization or has a condition that becomes unstable at times, a plan for how, when, and by whom the patient will be hospitalized must be agreed upon. Patients who become overtly psychotic in group sessions may require admission to an inpatient facility directly from the group. In these instances, the understanding conveyed to the patient in the pre-group orientation is that it will be the group leader who will oversee the decision to hospitalize anyone from the group. The group therapist has to be sure that the patient understands the plan and is in accord with it in order to avoid group crisis situations for which no appropriate contingency plan has been made.

11. *Extragroup socialization.* Conventional groups, especially those with a psychodynamic point of view, actively discourage or prohibit contact among members outside of group sessions. The reasons for adopting this group policy are designed to counteract the dangers that can arise when the group leader is not present. The formation of dysfunctional subgroups that exclude other group members, business or professional affiliations that spring from group, and romantic and sexual relationships that develop between members of the same psychotherapy group are a few of the countertherapeutic results that may follow extramural socialization among group participants.

In some selected groups, the leader may encourage contact among members for some reason that is consistent with the group goals. A “buddy system” in a group of patients who are lower functioning is an example of extragroup contact that is designed to be an extension of the therapeutic group process. Group members are paired together and come to group and go home together. Since social isolation, loneliness, and coming to group regularly are problems for these patients, the creation of a group-generated plan that involves all members and excludes no one is qualitatively quite different from spontaneous or planned acting out via outside contacts among members of higher level groups. The latter is almost always an example of enactment of some issue outside the confines of the group where it would have to be openly addressed.

In the same spirit, some substance abuse groups encourage the creation

of a sobriety telephone network among members. The purpose of the networking is to give members who are at risk to relapse into drug use an alternative option whereby they can call a sober group member who will support their healthy efforts to resist using drugs.

Whatever the group therapist decides is in the best interests of the group regarding extramural contact among group members, the pre-group orientation, not the established group, is the time and place to discuss this subject in some detail. Here again, the group leader may encounter some prospective group members who are unable to agree to adhere to this portion of the therapeutic contract. These members may have to be screened out for the group on the basis of their high-risk potential for undermining the group effort.

12. *Modifications of the group contract*. This category concerns the rare but significant instance in which something in a patient's life could be at cross-purposes with his/her effective participation in the group. Many people in groups have scheduling constraints; however, it is a mistake for the leader to accept a member for group when it is clear that they will have a regular pattern of absences. Individual requests to change the meeting time or date of the group based solely on any single member's personal limitations regarding scheduling are denied. The time, date, and meeting place are non-negotiable items unless, for some reason, the therapist has to change the arrangements due to his/her professional commitments.

The problems that occur when group meeting times are changed to accommodate one member, when episodic absence is accepted, and other deviations from the standard group structure are abandoned can be of such magnitude as to threaten the viability of the entire group experience.

13. *Confidentiality*. Effective psychotherapy of any kind cannot take place in the absence of the assurance of strict confidentiality. Incoming members are notified of the gravity of the confidentiality rule and are asked whether or not they feel they can adhere to it. Breaches of confidentiality are grounds for immediate dismissal of a member from the group. Any incoming member who cannot commit to strict confidentiality in the initial interview is excluded from joining the group.

Confidentiality takes on added meaning in group therapy since members have the ability to take information about other people out of the group. In order to protect the confidentiality of all group members, the group leader must draw a clear distinction between gossip and the constructive use of information originating in the group to the member's relationships outside of group. This transfer of learning function encourages group members to apply insights and behaviors learned in the group to the significant interpersonal relationships in their everyday lives without identifying the source of information in the group. Rules about confidentiality are restated in the initial group session and again at any point in therapy where potentially sensitive material is discussed.

14. *Questions and answers*. The preparation session concludes with a portion that encourages patients to ask any questions related to the group they are joining or about any aspect of group therapy in general. It is best to leave time for questions near the end of the meeting after the other checklist items have been explained. Allowing for a question and answer period affords an opportunity for the group leader to get a sense of how the pre-group orientation meeting has gone. It also gives the group member and the leader a time to go over any areas of confusion or misunderstanding concerning the content of the session.

Since many people enter group with preconceived notions concerning what actually transpires in group therapy, the pre-group interview encourages the expression of doubts, fears, and irrational beliefs about the upcoming group. The therapist can clarify myths and misconceptions and help ease many of the avoidable anxieties attendant to the process of being a new member of a therapy group.

A sense of mistrust, fear of being attacked in group, and concerns about humiliation, public exposure, and feeling inadequate often come up in the question and answer period. These can all be clinical clues to the leader about the patient's particular difficulty with entry into the group.

Another way of viewing the fears voiced in the interview is to see them as psychodynamic harbingers of the kinds of initial resistances that person is likely to demonstrate in the interpersonal milieu of the actual group. The group therapist can use this opportunity to join with the patient, reassure him/her, and begin the process of establishing a therapeutic alliance aimed toward facilitating both group entry and retention in the group.

In a newly formed group, once all members have had the pre-group orientation, the stage is set for the first group meeting.

GROUP THERAPY ORIENTATION CHECKLIST

1. Goals and general purpose of the group
2. The composition of the group
3. The role of the group leader
4. Observing and recording group meetings
5. The arrangement of the therapy setting
6. Time factors in the group
7. Loss and addition of group members
8. Rules about attendance
9. Fees and billing procedures
10. Coordination of other therapies
11. Extragroup socialization
12. Modifications of the group contract
13. Confidentiality
14. Questions and answers

STAGES OF GROUP DEVELOPMENT

Well-designed therapeutic groups go through developmental processes over the course of their life span. The stages of group development are characterized by observable individual and group-as-a-whole phenomena. Numerous writers have identified and labeled group stages in keeping with the author's theoretical and clinical belief systems. The American Group Psychotherapy Association's *Practice Guidelines for Group Psychotherapy* integrated these major historical efforts at defining group stages with contemporary research findings resulting in a consensus about a five-stage paradigm of group development.

In this system, the stages that define the group's progress across its flight path are: forming/preaffiliation, storming/power and control, norming/intimacy, performing/differentiation, and adjourning/separation. It is essential that the group leader be aware of the stage of the group's development and intervenes with the most stage- specific techniques available. Ill-timed or poorly sequenced interventions are not only ineffective, they can be damaging to individuals and to the whole group endeavor. Confronting a fragile member before sound group cohesion and support is established is an example of this error.

A fuller description of the elements of the five group stages will serve the twofold purpose of delineating the stages and the phenomenology of each and also shed light on the group leader's role at different points along the therapeutic spectrum. Two caveats to bear in mind are that stages of development are related to the composition, ego strength and goals of the group and that they are not necessarily linear. Debate exists in the field regarding whether or not gains made in earlier group stages are retained even when a member or the group drifts back into an earlier phase posture. Similarly, the capacity and pacing of how far and how fast a group can progress in its development is limited by the level of psychopathology in the group members. With these two caveats in mind, the outpatient interactive psychotherapy group composed of midrange to high functioning members is an opportune venue in which to appreciate the clinical significance of a developmental model of group therapy.

Preaffiliation

Initial concerns around inclusion and exclusion, acceptance by the leader and other members, the search for similarities, fears of interpersonal aggression, and competitive themes typify the first stage of a psychotherapy group. The group leader's focus is on knitting the nucleus of the group into a cohesive whole.

Therapist interventions encourage identification of goal compatibility among members, modeling and eliciting supportive comments, promoting "here and now" interaction among members, discouraging excessive

dependency on the leader, establishing clear group norms, and constructing a secure therapeutic frame to contain the group as it goes forward. A strong educational function also characterizes the therapist's role in the first stage.

Power and Control

Apprehension about the emotional safety of the group, testing the group leader's strength and weaknesses, negotiating differences with others in the group and experimentation with self-disclosure and self-assertiveness are hallmarks of the second or "storming" stage. Among the leadership tasks in this phase are mediating power struggles and conflicts in the group, demonstrating firm but fair limit setting, being on the alert for issues that might lead to drop-outs from group, promoting collaboration among members, and keeping the group on task toward the attainment of its goals.

Performing/Differentiation

As the group progresses toward a more mature level, intragroup transferences become more apparent. Both vertical (member-to-leader) and horizontal (member-to member) transferences are accessible. Levels of self-disclosure are higher and deeper interpersonal bonding takes place among members. There is a gradual transition from leader-directed group interplay to members sharing more responsibility for behaving in concert with the group norms. A cooperative group climate and greater reciprocity between members typifies the intimacy stage.

The role of the group therapist during this stage can be less active that in earlier stages; however, the group therapist still plays a critical role in guiding the group process and preventing it from getting detoured into unproductive or counter-therapeutic channels. The leader aims to keep the group on track and reinforce those behaviors that promote the acquisition of insight, recognize constructive attempts at problem solving, encourage the group to keep going forward in a collaborative spirit, and to guard against the formation of subgroups.

Norming/Intimacy

Against the backdrop of the secure base of established group cohesion, group members in this stage are more willing to expose deeper feelings that might have activated real or imagined risks at an earlier point in time. This stage is an active one in which members is less afraid to show individual differences. The prevailing group climate is one of mutuality and acceptance. This atmosphere fosters greater understanding of relationships through the experiential learning that comes by virtue of being in stable relationships with others in the group over time.

Leadership focuses on letting the group be more self-sufficient and

rewards spontaneous interaction among participants. Empathic responses are valued as are sincere attempts of members to be unselfishly related to others. The leader has to guard against abdicating the leadership role by becoming too passive during this phase. Since the groundwork has been laid for members to know how to express themselves in keeping with the group norms the leader may intervene less frequently but still bears overall responsibility for monitoring the group's therapeutic trajectory. Allowing and urging the group to be proactive and trusting the group are issues for the therapist during the phase of establishing intimacy.

Adjourning/Separation

The termination stage of therapy groups is noted for the triad of denial, regression and recapitulation. In open-ended psychotherapy groups, an event often sparks the realization that termination is the last stage of group therapy. Not uncommonly, this occurs when a member announces his/her plan to stop therapy because they feel that they have realized their treatment goals. Ambivalent feelings are evoked as members contemplate the impact of the loss of a valued member. Feelings of altruism, admiration, and having contributed to a fellow group member's success are accompanied by concurrent feelings of loss, competition, envy, anger, and concerns about each member's future. Members voice anxieties about the viability of the group when a known member leaves during this stage.

The group therapist has many leadership considerations related to an outgoing member and the remaining group. First, a clinical judgment needs to be made in order to distinguish between a genuine resolution of a member's therapeutic agenda versus a flight from group. The latter is primarily fear-based and as such, calls for the patient to remain in the group to work through the unresolved issues that prompted the desire to terminate prematurely from the group.

Second, the termination phase is notorious for denial defenses, a re-emergence of earlier problems, and regressive behaviors all aimed at demonstrating to the leader that ending the group would be a damaging decision. The group therapist has to help members stay engaged with the plethora of strong emotions engendered by the prospect of losing a close relationship.

The leader conducts a review of the departing member's clinical course and the group's progress. In doing so the contributions made by fellow group members are acknowledged. This helps consolidate gains made for all in the group and helps buffer feelings of despair, loss, hopelessness, and sadness that regularly accompany the final stage of group work.

Finally, plans for life after the group are always part of a thorough termination stage. In instances where the whole group is ending, the experience of gaining closure on the group experience and saying good-bye to one another round out the tasks of the group leader as the group

disbands. An appreciation of group development and the selection of carefully planned, stage-appropriate interventions by the therapist are essential ingredients of any successful therapy group.

LEADERSHIP ISSUES

The role of the group leader is the linchpin upon which the ultimate success or failure of the group enterprise rests. The leader creates, composes, and manages the group process from the selection and orientation of members through the working phases until the group reaches termination. The needs of the whole group and its individual members pose a series of potential challenges for the leader. The group therapist's self-awareness about his/her strengths and limitations in carrying out these divers leadership demands is essential for a positive group outcome.

Generic leadership functions and desirable personal and professional qualities in the person of the group leader has been a subject of study since the inception of group psychotherapy. Brabender, Fallon, and Smolar have eloquently outlined the core elements in successful group leaders. These qualities consist of an ability to look into oneself, the capacity for empathy, having a positive belief about the value of group therapy, and possessing good interpersonal skills. Additionally, executive leadership functions that encompass the ability to set and maintain the therapeutic frame of the group experience, a capacity to appreciate multiple simultaneous levels of interaction in the group session, helping members supply accurate meaning attribution to their behavior, and maximizing the use of the role of authority figure in the group are other central leadership attributes.

In terms of specific leadership technique, the leader has to be proficient in tolerating as well as containing intense affect, managing his/her own countertransference engendered by the group, performing a teaching function for members in which the leader clarifies misperceptions and provides factual information, and uses interpretation in a timely manner. Since the leader is not immune to group pressure, care has to be exercised when using confrontational techniques and with respect to therapist self-disclosure in group meetings. As an example, when the group is in a state where anxiety in members is high and expresses itself in the form of confronting, testing, and challenging the leader, his/her competence, and/or the value of the group experience, the leader has to be able to separate personal feelings of rejection and inadequacy in favor of focusing on what and why this material is coming up at his point in time.

A helpful rule of thumb for judicious use of therapist self-disclosure is that the therapist should make sure that when self-disclosing information of a personal nature that it meets the criteria of being in his/her conscious awareness, is confined to the "here and now" of the group, and that the same goal could not be accomplished by other means. In this way, inappropriate or reflexive countertransference reactions by leaders that are

harmful to the group are avoided. It is incumbent upon the leader to examine any temptation to self-disclose that has to do with the therapist's personality as it articulates with the dynamics of the group.

Co-Therapy

In many groups, the complexity of issues presented by the group suggests that sharing the leadership responsibility with a co-therapist would be advantageous. The workload on the leader who functions alone can be formidable in certain types of groups. As an example, co-therapy is often used in groups treating drug and alcohol addiction. There are many problems of living that involve the family of the substance abuser and require attention on the part of the group leader. Extragroup issues including legal problems, debt, housing, and employment are all factors that have impact on conducting a group composed of substance abusers.

These issues coupled with the empirical observation that staff burn-out is high among those who work exclusively with the problems of substance abuse have led to a trend to use more than one leader in these groups. Team leadership is advantageous in allowing for the sharing of leadership tasks in the group. The presence of more than one leader increases the options for identification, interaction, and the exploration of intragroup relationships. As a result, the leader's therapeutic load is lightened and the pace of the group can be accelerated by the presence of two leaders as objects of psychological investment by members.

When co-therapy is utilized, leaders have to be clear about their role and task definition in group in order to function collaboratively. Co-leaders who integrate their realistic differences concerning professional discipline, gender, and theoretical orientation provide the group with an excellent role model for collaborative function in adult relationships. More opportunities for role modeling and for transference possibilities accrue when an opposite sex team leads the group. This leadership combination stimulates family feelings in members by having a new therapeutic parental pair with whom members can interact.

A specialized form of co-therapy occurs in some substance abuse groups when a recovered alcoholic or drug addict is part of the therapy team. The recovered user occupies position midway between that of a leader and a member. Their member-like qualities, including their first-hand experiences with life on drugs and alcohol, add a dimension to the group that cannot be supplied by the professional mental health staff. In order to be considered for this "elder statesperson" role in the group, former substance abusers must have a documented history of at least 2 years of sobriety. They must also consent to urine screening during the course of the group therapy. The professional staff and recovered counselors meet immediately following each group session to discuss group events and the peer counselor's reactions to them. The post group discussion or "re-hash"

is designed to plan a course for future sessions and to ensure that the dynamics of the group do not pose a threat to the sobriety of the counselors themselves.

Variations in this specific co-leadership model are also found in groups for bipolar disorder, medical illness, psychological trauma groups, and other conditions where a person has made a recovery and wants to share his/her experience with others trying to cope with an earlier stage of the problem.

CLINICAL APPLICATIONS

The versatility in the use of the group matrix can be seen in the following descriptions of contemporary innovative clinical applications of group therapies.

Combined Individual and Group Therapy

Many conditions require more than either group or individual therapies alone can supply. The simultaneous treatment of a patient in both one-to-one and group therapy is commonly referred to as "combined therapy." Therapy can be conducted by the same therapist or by two separate therapists. Different clinical issues are engendered by both types of combined therapy.

When the group and individual therapist are the same person, members only in group therapy often sees the patient in combined therapy as being in a favored position by the therapist/group leader. This often induces group themes of competition, acceptance, and rejection and unresolved emotional residue from the families of origin of group members.

In circumstances where one therapist is the dyadic therapist and another the group therapist, professional colleagues must regularly communicate with each other. This ensures coordination of a treatment plan for the patient and helps monitor whether the patient is doing an editing or splitting function by reserving some material only for individual therapy and not disclosing it in the group or vice versa.

A patient with a borderline or narcissistic personality disorder is at higher risk to play off one therapy against the other as a resistance to change. Astute use of combined therapy along with a dialectical group experience can identify this dysfunctional behavior and offer more adaptive alternatives. Specifically, the personality-disordered patient in combined therapy may use individual therapy to focus on intrapsychic issues and can apply the insights gained in therapy and experiment with new modes of behavior, thinking, and defenses in the safe and controlled interpersonal setting of the dialectic group.

At the same time, groups are concentrated on themes of appropriate interaction with peers and authority figures, the constructive expression of

intense feelings such as anger and intimacy, and the enhanced reality testing which groups provide. These attributes of the group help counter irrational beliefs and socially deviant behavior and offer more viable alternatives. Ideally, consistent communication between therapists staves off avoidable pitfalls and adds to the efficiency of the total therapeutic effort.

Bauer and McBride have created a life goals group program that combines psychoeducational, cognitive-behavioral, and self-management skills training for patients with bipolar conditions.

Groups for Mood Disorders

Affective disorders play havoc with the interpersonal relationships of those diagnosed with a mood instability. The advances in diagnosis and psychopharmacological options for treating depression have been dramatic. Group therapy has proven to be a widely accepted component of the treatment plan for patients with mood disorders. When one looks at the interpersonal sphere of the bipolar patient, their histories are riddled with interpersonal problems.

Conflicts at work, with domestic partners, original family members, and the law are common occurrences. As a corollary, depressed patients are invariably withdrawn from social contact and/or show signs of irritability as a prominent symptom.

A supportive and structured group psychotherapy milieu is ideal for addressing some of the issues that confront patients with serious mood disorders. Educating patients about their condition and the latest pharmacological options available to them, using the group to focus on learning structured problem-solving techniques while simultaneously experiencing working in a cooperative way with fellow group members and staff, plus positive use of re-framing techniques designed to counter stigmatization of mental illness are the central components of the life goals paradigm.

Groups for Personality Disorders

The model above is emblematic of many group approaches not only for affective problems but also for personality disorders where affective dysregulation is a core component of the therapeutic agenda. Exponents of group therapy know that group pressure is a powerful force for change that is unique to the group setting. When working with personality disorders in group therapy, the therapist has more leverage available to induce change in maladaptive behavior, negative attitudinal stances, and interpersonal deficits.

The ingredients in the interpersonal group (IPG) psychotherapy protocol for patients with BPD typify current thinking about the utility of

groups to achieve significant change. The therapy group brings to light the borderline patient's inner self-system with its feelings of emptiness, hunger for affect and gratification, and the inefficient interpersonal styles they have created in order to reach these ends. The IPG utilizes the dilution of the transference that sharing a group leader provides to turn down the emotional temperature of the group so that the atmosphere is more conducive to learning.

The interpersonal emphasis strives to help the patient form and experience more positive, gratifying identifications with others in the group. Using a theoretical premise that views problems with attachment as a common starting point for the interpersonal difficulties found in people with personality disorder, the group therapist uses a noninterpretive, empathic approach. In contrast to more traditional psychoanalytic approaches, the group leader avoids confrontation that invites strong affect, particularly in the early stages of the group, in favor of functioning as a "container" for the patient's confusion, anger, and dysphoria. Once a therapeutic alliance is formed, the leader works with individuals in the group to counter their self-deprecating beliefs and replace them with a positive group climate of hopefulness and the possibilities for change.

These techniques rely heavily on the group leader's subjective responses to behavior by patients in the group and use caring feedback and suggestions of alternative courses of action as the tools to implement change. At the heart of this method is the assumption that in the IGP model, "change is more due to the experience of interactions in the group and less to the acquisition of insights about the genesis of internalized conflicts."

DBT has also been adapted to work effectively in the groups as well as in the individual setting. Although DBT was originally designed for working with patients diagnosed with borderline personality disorders, it has been customized for use with patients with eating disorders, families with suicidal adolescents, and substance abusers. Still more applications of IGP- and DBT-based group therapies are expanding rapidly.

Groups for Anxiety Disorders

Anxiety is present in the vast majority of psychiatric conditions. For many people, it is the primary reason for which they seek professional help. While there are effective pharmacological agents available for treating aspects of anxiety disorders, group therapies are particularly helpful as either adjunctive or primary modalities of treatment geared to managing the symptoms of anxiety and offering a vehicle for working on the interpersonal fallout from the habits formed by people who have unsuccessfully managed the stressors in their lives.

Groups, particularly those with a cognitive-behavioral group therapy (CBGT), have been applied to the full gamut of anxiety disorders from

adjustment disorders through panic disorder up to and including PTSD and phobic states.

Anxiety, fear, and panic are all characterized by avoidance behaviors. The CBT group aims to empower patients to engage with their anxieties by supplying them with a set of anxiety-reducing skills. Universal components of individual CBTs are adapted to the group setting. The advantages of a CBT group are greater opportunities for role modeling, practicing new behaviors in vivo before testing them out in the world outside the group, getting peer feedback from others with similar anxieties, and the positive reinforcement that comes from many people, not just the therapist, which serves to reinforce new habits and thought patterns.

Examples of CBT group techniques include task assignments including personal journals and homework assignments, teaching progressive muscle relaxation, meditation, deep breathing and other stress management mechanisms, and identifying the schemas or belief systems that underlie cognitive distortions. Maladaptive thought processes such as catastrophic thinking, polarized "black or white" thinking, and over-personalization are labeled in group members and cognitive restructuring is used to substitute more reality-based responses to stressors.

CBT groups for anxiety-based problems are the fastest growing form of group therapy today. They have gained clinical popularity for several reasons: the therapeutic interventions are often time-limited; the techniques are clearly described and can be standardized thereby facilitating research on outcome by self-report, clinician administered instruments and/or cognitive and behavioral testing; precise treatment plans can be made into manuals that have a specific agenda and goals for each session; and the role of the group leader is very clearly defined.

Anxiety cuts a wide swath across generations, gender, ethnicity, race, culture, and physical illness. In CBGT, commonality of anxiety problem is the central factor around which these groups coalesce. CBGT has the expectation that change will take place outside of group as well as in formal group sessions. Out-of-session assignments, such as monitoring and recording intrusive thoughts that provoke anxiety, are regular staples in cognitively based groups. CBT groups are prototypical of most brief or short-term group psychotherapy models insofar as they have a fixed time frame, a strong educational component, a clear focus for each group session, and very active, directive leadership that strives to keep the group on task and terminates the group at a predetermined point in time.

The awareness of traumatic conditions experienced by men and women who are returning soldiers has formed a large basis for the application of effective trauma groups.

Groups for Schizophrenia and Other Psychotic Disorders

Group therapy for patients with severe psychiatric disorders has truly come

of age. Historically, inaccurate assumptions were made about the viability of therapy groups for psychotic patients. Most of these doubts came from psychoanalytic theory that proposed using techniques that were beyond the capabilities of schizophrenic patients and, in some instance, actually harmful to fragile patients.

Groups address many of the central issues involved in managing patients with severe psychiatric impairment. From crisis control to diagnostic evaluation and as a primary form of psychosocial therapy, the group milieu offers enormous flexibility by creating a foundation for a comprehensive treatment program for the chronic patient. Inpatient, transitional, and outpatient groups, all constitute the group resources available to the clinician. The choice of group options is primarily related to the patient's diagnosis and to the stage of severity of the primary psychiatric problem.

Using this paradigm, groups for chronically ill psychiatric patients can be conceptualized as acute, transitional, and maintenance or aftercare in nature. Patients with acute first episodes of a condition or with an exacerbation of a long-standing problem are primarily hospital based. Inpatient group therapies encompass quiet activity groups; art, poetry, and dance programs; occupational therapy; discussion groups, and community meetings on the inpatient service. Goals that are overly ambitious should be abandoned to suit the relatively brief time period of hospitalization. Simply stated, one of the highest priorities for inpatient groups is to get patients to be comfortable with group therapy. Having a successful group experience in the hospital increases the probability that patients will continue in after care groups and not get lost in the transfer from inpatient to outpatient status.

Transition groups refer to the transition from hospital-based treatment to outpatient and community-based programs. The start of this process begins with discharge planning in the hospital. Discharge planning groups, social skills training groups, family groups, job-focused behavioral groups, and medication groups are all representative of groups for people in the process of leaving the hospital.

Groups in an outpatient setting are designed to prevent relapse, which would require rehospitalization and to help participants develop rewarding social, familial, and vocational lives. In general, these groups aim to reduce social isolation, monitor the effect of medication as measured by the patient's level of interpersonal function, aid in reality testing, provide social support, reduce the stigma associated with being mentally ill, and transmit accurate, up-to-date information about the primary condition. Leadership is very active in setting a therapeutic structure around which the group can consolidate. Psychotropic medication is invariably an integral part of the group life whether or not the group is a medication group per se.

Most groups for chronic and persistently mentally ill patients are based

on principles of supportive psychotherapy. Exploration of unconscious material, delusional themes, and affect related to anger and hostility are to be avoided. The emphasis is on patient strengths and the accumulation of small, incremental successes started in the group and extended to life outside the group. Problem solving, advice giving, and teaching new coping strategies are all hallmarks of outpatient groups for patients with chronic and debilitating mental illness. The emphasis on improving cognitive function is also a central group goal.

Table 33.5–3 illustrates the similarities and differences between groups that have psychotic members and those which do not.

A group focus on medication compliance, education, socialization, vocation, and psychological integration is the constellation of benefits chronically ill patients can derive from ongoing participation as members of groups.

RESEARCH ISSUES

The field of group psychotherapy has departed from an antiquated, predominantly anecdotal literature to a firm footing in scientific inquiry. Data from well-designed and well-controlled studies attest to value and utility of groups as a primary form of psychiatric treatment. Significant findings related to group cohesion show that pre-group selection and orientation; higher levels of early group structure; balancing individual and intragroup factors; active group leadership, facilitating emotional expression, relationship building, and verbal feedback all result in high levels of group cohesion that, in turn, favors positive outcome. The role of a strong therapeutic alliance is also correlated with positive outcome in both brief and long-term groups.

A greater degree of both specificity of focus and research study design characterizes the cutting edge efforts to support the clinical enthusiasm for group psychotherapy. Additionally, research instruments have been created to decode the mechanisms by which groups work. Empirical observation and research focus of group work now study aspects of group process, leadership variables, and therapeutic factors that form the basis for the change resulting from group participation. The CORE Battery developed by the American Group Psychotherapy Association is the best-known example of this effort.

Table 33.5–3.
Group Therapy with Psychotic and Nonpsychotic Patients

	Psychosis	Nonpsychotic
Composition	Homogeneous for level of function	Homogeneous: Brief group therapy Heterogeneous: Long-term group therapy
Functional capacity	Poor ego strengths Interpersonal deficits	Able to learn from group feedback Capacity for empathy and altruism

Goals	Crisis intervention Stabilization Improved communication Social skills development	Improved self-perception Improved interpersonal function Resolution of personality issues
Educational elements	Central to the group model	Elective: Yes: Brief groups No: Psychodynamic group therapy
Therapeutic factors	Reality testing Identification Social support Bolster defenses	Cohesion Challenging defenses Use of transference and unconscious material
Cohesion	Desirable but hard to attain	Mainstay of the group model; easier to attain quickly
Leader role	Active around structuring the group Avoids aggressive and irrational themes	Spontaneous group interaction Leader is catalyst for group process
Confrontation	Strongly avoided	Actively encouraged
Interpretation	Only on occasion and limited to the "here and now"	Regularly used for in-group transactions and transferential issues
Medication	Almost always part of the group	May or may not be a group issue
Group focus	Present and future	"Here and now" plus historical focus
Socialization	Strongly encouraged	Generally discouraged
Time span	Brief and long-term models Open-ended "maintenance" group therapy	Brief: Symptom focused Long term: Personality and behavior change
Group participation	Involuntary and voluntary	Voluntary

FUTURE DIRECTIONS

There is a clear trend toward eclectic group therapies characterized by technical integration. Greater emphasis is being placed on those group interventions that are specific, documentable, and capable of replication when employed by other practitioners. Many group therapists draw upon a psychodynamic understanding of each group member but their modes of intervention have become liberalized and less restricted to traditional analytically based practice. Cognitive elements, educational components, structured group exercises, and the assignment of tasks based on the expectation of increased patient responsibility for change are but a few of the additional eclectic modes of intervention added to the therapeutic armamentarium.

Training modules and manuals have been created for practitioners and the public covering several critical aspects of group therapy. How to train future supervisors of group therapy, explicating the issues involved in the ethical practice of group therapy and guidelines for teaching a basic course in group psychotherapy representing three examples of a trend to educate practitioners in order to ensure quality control and to strive to continue to provide first rate group services to patients.

Times have changed and group psychotherapy has kept pace with them.

As new developments occur, tandem reflection in the addition and modification of group technique and sensitive therapeutic styles ensure the progress of group therapy as a welcome transition to modern psychiatric clinical work.

REFERENCES

*American Group Psychotherapy Association Science to Service Task Force. *Practice Guidelines for Group Psychotherapy*. New York: American Group Psychotherapy Association; 2007.

Bauer MS, McBride L. *Structured Group Psychotherapy for Bipolar Disorder: The Life Goals Program*. New York: Springer Publishing Co.; 2003.

Bernard HS, Spitz HI. *Training in Group Psychotherapy Supervision*. New York: American Group Psychotherapy Association; 2006.

Brabender VA, Fallon AE, Smolar AI. *Essentials of Group Psychotherapy*. Hoboken, NJ: Wiley; 2004.

Brook DW, Spitz HI, eds. *The Group Therapy of Substance Abuse*. Binghamton, NY: The Haworth Press; 2002.

*Buchele BJ, Spitz HI, eds. *Group Interventions for Treatment of Psychological Trauma*. New York: American Group Psychotherapy Association; 2004.

*Burlingame GM, Strauss B, Joyce AM, et al. *American Group Psychotherapy Association's CORE Battery—Revised*. New York: American Group Psychotherapy Association; 2006.

Feld B, ed. *Special Section: Couple Therapy, Group: The Journal of the Eastern Group Psychotherapy Society*. 2007;31:3. Catskill, New York.

Flores PJ. *Group Psychotherapy with Addicted Populations: An Integration of Twelve-Step and Psychodynamic Theory*. 3rd ed. Binghamton, NY: The Haworth Press; 2007.

Fuhriman A, Burlingame GM. *Handbook of Group Psychotherapy: An Empirical and Clinical Synthesis*. New York: Wiley; 1994.

Gantt SP, Cox P, eds. *Neurobiology and Building Interpersonal Systems: Groups, Couples, and Beyond,* Volume 60:4. New York: The Guilford Press; 2010.

Grundy D, ed. *The Short-Term Inpatient Group, International Journal of Group Psychotherapy,* Volume 62:2. New York: The Guilford Press; 2014.

Grundy D, ed. *Special Issue on Group Therapy Around the World, International journal of Group Psychotherapy*, Volume 65. New York: The Guilford Press; 2015.

Heimberg RG. *Cognitive-behavioral Group Psychotherapy for Social Phobia*. New York: The Guilford Press; 2002.

Jongsma AE, Paleg K. *The Group Therapy Treatment Planner*. 3rd ed. Hoboken, NJ: John Wiley & Sons; 2015.

*Kaplan HI, Sadock BJ, eds. *Comprehensive Group Psychotherapy*, 3rd ed. Baltimore: Williams & Wilkins; 1993.

Klein RH, Schermer VL. *Special Issue on Violence in America: Part I, International Journal of Group Psychotherapy*. Volume 65:1. New York: The Guilford Press; 2015.

Klein RH, Schermer VL, eds. *Special Issue On Violence in America: Part I & II, International Journal of Group Psychotherapy*. Volume 65. New York: The Guilford Press; 2015:1–109; 163–294.

Kleinberg JL. *The Wiley-Biackwell Handbook of Group Psychotherapy*. 1st ed. Hoboken, NJ: Wiley-Blackwell; 2016.

MacNair-Semands R. *Ethics in Group Psychotherapy*. New York: American Group Psychotherapy Association; 2005.

Marziali E, Munroe-Blum H. *Interpersonal Group Psychotherapy for Borderline Personality Disorder*. New York: Basic Books; 1994.

Motherwell L, Shay JJ. *Complex Dilemmas in Group Psychotherapy*. New York: Brunner-Routledge; 2004.

Roth JD. *Group Psychotherapy and Recovery From Addiction: Carrying the Message*. London: Routledge/Taylor & Francis; 2016.

Schein LA, Bernard HS, Spitz HI, Muskin PR. *Psychosocial Treatment for Medical Conditions: Principles and Techniques*. New York: Brunner-Routledge; 2003.

*Spitz HI, Spitz ST. *A Pragmatic Approach to Group Psychotherapy*. Philadelphia, PA: Brunner-Mazel/Taylor & Francis Group; 1999.

Spitz HI. *Group Psychotherapy and Managed Mental Health Care: A Clinical Guide for Providers*. New York: Brunner/Mazel; 1996.

Weber RL. *Principles of Group Psychotherapy*. New York: American Group Psychotherapy Association; 2003.

White JR, Freeman As. *Cognitive-behavioral Group Therapy for Specific Problems and Populations*. Washington, DC: American Psychological Association; 2000.

Wolf A, Schwartz EK. *Psychoanalysis in Groups*. New York: Gruneand Stratton; 1962.

*Yalom ID, Leszcz M. *The Theory and Practice of Group Psychotherapy*. 5th ed. New York: Basic Books; 2005.

Zorzella KPM, Muller RT, Classen CC. *Trauma Group Therapy: Role of Attachment and Therapeutic Alliance; International Journal of Group Psychotherapy* Volume 64:1. New York: The Guilford Press; 2014.

▲ 33.6 Family and Couples Therapy

HENRY I. SPITZ, M.D., AND SUSAN T. SPITZ, A.C.S.W.

INTRODUCTION

The past decade has been one of enormous upheaval and turmoil. There is quality of unpredictability from terrorism to climate change in the world today. The concept of who constitutes a family and a couple has dramatically evolved during this period. It is understandable therefore that the formats of family and couples therapy have gained prominence as a means of engaging people in distress.

The logic of seeing people through the prism of their family and its history, both past and present, has gained widespread acceptance. There are reverberations in the family and couple when one member suffers from chronic medical and/or psychiatric conditions like schizophrenia, bipolar disorder, eating disorders, autism and attention deficit disorder, violence, and abuse. Wars have resulted in post-traumatic stress disorders, depression, and substance abuse. Poverty, unemployment, and general economic uncertainty cause setbacks and disruption in everyday family life. Technology has advanced exponentially, altering communication, and eroding intimacy in relationships. Acceptance and rejection of religion, race and ethnicity, as well as of a broader range of sexual identification in parents and children has had repercussions in many families and couples. Divorce and single parent families are more prevalent now and these approaches are most appropriate for dealing with the sequelae of those situations.

The family is the foundation upon which most societies are built. The study of families in different cultures has been a subject of fascination and scientific interest from viewpoints as diverse as sociology, group dynamics, anthropology, genetics, the brain, evolutionary biology, and, of course, the mental health field. The confluence of information gleaned from family studies has set the backdrop against which the contemporary practice of family therapy has evolved.

Modern family and couple therapies see the family and the couple as both part of the problem but also as essential resource available for the resolution of the presenting problem. Family strengths as opposed to weaknesses are reinforced, educational elements offer families new modes of adaptation in place of rigid, well-practiced but dysfunctional interactional patterns, and the therapist is the ally of and advocate for the family and each of its members. The specific ways in which this shift in orientation has come about can be readily seen when one looks at the developments in the family therapy field in its historical context.

HISTORICAL TRENDS

Both family and couple therapists often point to the work of Alfred Adler at the Vienna Child Guidance Clinic. Adler's involvement of parents in the therapy for their children, a practice that seems axiomatic today, represents a significant departure from the dyadic psychoanalytic model prevalent at the time. It is not until the 1940s that developments relevant to the evolution of contemporary family therapy practice can be seen. During this decade, the biologist Ludwig von Bertallanffy applied concepts from mathematics and physics in an attempt to create a unifying theory for human interaction. His work emphasized the systemic nature of the human organism and the capacity for change in open human systems. He saw a system as more than the sum of its parts and capable of interacting with other systems.

Bertalanffy used the concept of equifinality to denote the capability of a system to find many routes to the same solution of a problem. A strong opponent of the dehumanized mechanistic theory prevalent at the time, he preferred a view that regarded human systems as ecological organisms with a well-defined value systems. First-generation family therapists including Murray Bowen working at the Menninger Clinic and Carl Whitaker at Emory University became interested in working with schizophrenic families. John Bowlby's work about infant–mother bonding and the perils of insecure or inadequate attachments in early life and its subsequent impact on critical adult relational issues such as trust and forming an enduring interpersonal bond caught the attention of interested family therapists.

Nathan Ackerman is credited with being the founding "father" of family therapy. Trained as a child psychiatrist and psychoanalyst, Ackerman retained his emphasis on understanding individual personalities but also had a broader view of how individuals articulated with one another in ways that clashed or meshed. The logical extension of this point of view was to conduct sessions with the whole family and to intervene in provocative and confrontational ways that would mobilize covert family conflict such that it was available for examination and change in the session.

Gregory Bateson, began working in Palo Alto, the city whose name was to become synonymous with other influential family therapists including Don Jackson, Jay Haley, and Virginia Satir (the Palo Alto group) who conducted their work at the Mental Research Institute. Bateson met Norbert Weiner, a mathematician and expert in the field of cybernetics, and became intrigued with the application of cybernetics to human systems. Cybernetics enabled the discovery of key constructs in family therapy that included the premise that change was not linear as had been previously assumed, but rather it was circular, determined by both positive and negative feedback loops within the family interaction. The "double bind" is perhaps the best-known example of a negative feedback loop in

family communication. The famous mother to son comment, “I hate you, you’re such an angel” is a case in point.

Later on, Jackson, Haley and Satir turned their attention to the process of communication in severely disturbed patients and their families. They were among the first to understand psychotic communications as having meaning and making a certain kind of sense in schizophrenic families. They began to label communication patterns and coined terms that are widely used today. Jackson studied couples and identified complementary and symmetrical relationships and the decoding of redundant relationship patterns or family and couple “rules.” Satir contributed clinical techniques that were aimed at reducing distance among family members. Her experiential approach, therapeutic use of self-disclosure, and appropriate physical interaction with family members were later to be seen in the work of many who followed her.

Jay Haley was influenced by the work of Milton Erickson, an exponent of the use of indirect hypnotic techniques, metaphor, and suggestion. Haley’s therapeutic stance was an active and pragmatic one directed toward the rapid resolution of the presenting problem in the family.

Murray Bowen was a strong theoretician as well as a clinician. His focus began when working with families with a schizophrenic member during the course of which he made observations about how family members cooperated to maintain an interpersonal space and a degree of individual differentiation from the undifferentiated family ego mass. Bowen’s awareness of triangulation, a process whereby couples and families involve a third person (extramarital affair; mother–father–child triad) or third “thing” (work, alcohol and drugs) all of which have the effect of attenuating the fears related to closeness and intimacy in primary relationships.

Three other major figures in the development of family therapy were Carl Whitaker, Lyman Wynne, and Theodore Lidz. Whitaker is perhaps best known for his emphasis on the family therapist’s need to interrupt and repattern disruptive, rigid, well-practiced family habits. His open attitude toward his own work being observed by others was the embryonic beginning of the use of the one-way mirror. The consulting team approach and the use of videotaped sessions as regular fixtures in couple and family therapies emerged from their work. Though often flamboyant in style, Whitaker used himself as an agent of change employing unpredictable, spontaneous, and often confusing techniques designed to upset the homeostasis of an ineffective family system.

Lyman Wynne was interested in repetitive, stereotypical roles taken by family members and theorized that disordered or communication deviance was the conduit through which transgenerational transmission of irrationality went from one generation to the next. He coined the term “pseudomutuality” to denote the role of a false front presented by families who fear both intimacy and separateness. Although the bulk of Wynne’s

work was with schizophrenic families, subsequent research and clinical studies have shown that the patterns he identified go beyond the schizophrenic spectrum and are readily observable in personality, mood, and anxiety disorders.

Theodore Lidz, also working with schizophrenic families, implicated the critical role played by fathers in these families. Prior to Lidz's work, the main influence on the intrafamilial environment of these families was felt to be the mother. Lidz and his colleagues defined specific destructive patterns of paternal interaction with spouses and children that led to serious adverse psychological consequences. When Lidz worked with dysfunctional couples, he also delineated maladaptive patterns that led to marital discord or disruption. Marital schism, a state wherein partners compete rather than collaborate and fail to integrate their differences in a functional manner. Marital skew, where a conspicuous power imbalance exists and active antagonism predominates are the most widely recognized examples of Wynne's terminology for dyadic relationships.

These family and couple therapy pioneers set the stage for the decades to follow that bore witness to virtual explosion of family therapy research, and practice and the emergence of defined schools of family therapy. The family and couples therapy literature expanded exponentially following the groundwork they laid. Corresponding acceptance of family therapy as a viable clinical entity now supported by empirical research studies gave greater momentum to family and couples work and the creation of several mainstream models of family therapy practice.

DEFINITION

Family therapy is a mosaic of techniques all of which have the shared goal of direct alteration of maladaptive family processes. In point of fact, it is more accurate to speak of family and couple therapy in the plural, as family and couple therapies since the array of interventions with the above goal in mind is impressive. The coalescence of discreet therapeutic orientations has resulted in the creation of several distinct tracks in actual clinical practice of family therapy. These orientations toward the treatment of family problems will be the subject of the next section.

Schools of Family Therapy

Schools of thought in family therapy are like windows in a greenhouse, each with its own orientation when observing the family through a different pane. Each has its own lexicon, goals, and focal points, but all share a common purpose in attempting to find innovative ways to provide time and cost-effective therapies for families in distress.

A representative sampling of the current state of family therapy techniques shows the clear linkage between first-generation family therapy

described above and how it manifests itself in modern clinical practice. Gurman, in the chart that follows, has simplified the apparent overlap in the different schools by outlining the major schools of family therapy, the influential practitioners of each method and the focus of each technique (Table 33.6–1).

This schema allows one to compare and contrast each family therapy model in more depth. The **cognitive-behavioral** approaches have their theoretical base in social learning theory, operant conditioning, and cognitive psychology. Cognitive and behavioral therapies require active therapist participation both in structuring the framework of therapy, finding circumscribed as opposed to global goals, and having a planned sequence of interventions and instruction in mind in order to reach the solution of the presenting problem. For the most part, cognitive/behavioral family therapy is brief, symptom-focused that places low value on revisiting historical material in sessions. Like most family and couples interventions, with the possible exception of the psychodynamic approach, it is extremely pragmatic, focused in the here-and-now, very structured, interactive and directive.

The over-riding premise is that problems are learned behaviors that are well practiced and consequently almost reflexive in nature. This model helps expand the couple or family's repertoire of functional communication and actions and, in so doing, incorporates new ways of thinking about familiar but dysfunctional family interactions. For example, the wife who feels that her husband is working too hard delivers that message from a position of caring and concern. The husband in this case "hears" the message as critical of him and feels that his partner is accusing him of being a workaholic and a neglectful father. The cognitively oriented therapist would identify this irrational belief or cognitive distortion, clarify it by supplying an alternative, more positive reading of the message, encourage the use of reality testing to keep catastrophic, or all or nothing thought patterns in check and reinforce all new, more adaptive and rewarding responses on the husband's part.

An essential ingredient in adapting individually based cognitive/behavioral techniques to the couple and family milieu is the added emphasis on social exchange theory. This adds a dimension of mobilizing the reward system in the system. Families and couples are taught to replicate those behaviors that increase the number of rewarding interchanges, decrease negatively cast feedback and acquiring collaborative problem solving skills.

The therapeutic tool kit for cognitive/behavioral therapy with families and couples contains many of the same techniques used in individually based cognitive therapy. Positive reframing, cognitive restructuring, ensuring a shared base of assumptions and expectations between partners and among all family members are all parts of the therapeutic equation.

In order to induce behavioral change, in-session exercises and educational elements are coupled with homework assignments to be completed between sessions. Cognitive/behaviorists do see whole families but feel more comfortable than practitioners from some other schools in seeing those subsystems in the family that are most relevant to the resolution of the identified problem. Therapists of this orientation are also comfortable with the concept of repetitive brief modules of treatment designed to address a specific issue each time over the course of time.

Jacobson and others have validated the benefit of the cognitive/behavioral approach. Advocates of this model have created impressive arrays of instruments for measuring parameters of change. Generally speaking, the cognitive and behavioral therapies are the most specific about technique (what they actually do in family sessions) and about the ways in which they assess the efficacy of their work.

Novel applications of these approaches have been used for parenting skills training, childhood conduct disorders, developmental disabilities, premarital counseling, and bipolar disorder. The integration of cognitive/behavioral family work with other therapies has been helpful with eating disorders, substance abuse, and anxiety disorders.

Family Systems—These formats encompass a broad range of techniques all of which are unified by their belief that the family is a human system with unique properties that contribute to the current emotional and behavioral configuration. Much like a clock with interlocking gears where one does not move without affecting the others, without therapeutic intervention dysfunctional families become synchronous and stuck over time.

The most frequently employed subdivisions within the systemic approach are the closely allied structural and strategic therapies. Salvador Minuchin and his colleagues at the Philadelphia Child Guidance center were the originators of **structural family therapy.** Their early work with families that had a symptomatic child formed the basis for the observations about dysfunctional patterns in families with a symptomatic member. In particular, strategic therapists look at boundary issues in families to see where the family lies along the spectrum of over-involvement (enmeshment) and detachment (disengagement). Similarly, those who think systemically look for mechanisms within the family that do not serve them well. The detouring of parental marital conflict by focusing on a child with psychological symptoms is a case in point.

Structural family therapy is a very leader-centered modality that uses directives geared to initiate change by stimulating affect, increasing the therapeutic alliance, and by employing both logical and paradoxical instructions that would induce change in the rigid family system. Once identified, the structural family therapist aims to reset the imbalances in the family system through a variety of therapeutic interventions.

Commonly the structural therapist tries to help the family solve the current problem by defining clear and attainable goals, establishing appropriate generational boundaries, assisting parents to take leadership positions in the family, offering alternative, often provocative viewpoints that differ from the family's skewed understanding of the problem, and restoring or helping to create function subsystems within the family. According to Nichols, structural family therapy follows seven steps: joining and accommodating—reducing initial anxieties about therapy; enactment—encouraging spontaneous interaction among family members to demonstrate the properties of the system and to allow for insertion of behavioral instructions that run counter to the family's established pattern of behavior; structural mapping—broadening the focus from the identified patient to a family-as-a-whole level; highlighting and modifying interactions—turning the volume up or down in order to illustrate a behavioral sequence; boundary making—aligning relationships between subgroups in the family; unbalancing—using joining maneuvers to induce change in an unequal or destructive dyad or triad in the family; and challenging unproductive assumptions—introducing notions that run counter to the established family belief system. The structure of the family with respect to concepts of boundaries, alliances, splits; cutoffs of relationships, complementarity, and positive and negative subsystems within the larger family system are the focal points for the structural approach.

Table 33.6–1.
Schools of Family and Couple Therapy

Therapeutic Approach (Examples)	Core Concepts	Typical Goals	Common Strategies and Techniques	Comment
Behavioral/cognitive-behavioral (James Alexander, Neil Jacobson, Donald Baucom, Gerald Patterson)	Functional analysis Social learning theory Communication and problem solving Attributional style	Resolution of presenting problem Enhanced communication and problem-solving skills Balance between change and acceptance	Communication and problem-solving training Acceptance training/reattribution techniques Parent management, emphasis on behavior/consequences Therapist as educator	Functiona not tre technic definin charact Arguably t empiric suppor family a metho
Bowen family systems (Murray Bowen, Philip Guerin, Michael Kerr, Monica McGoldrick)	Differentiation of self Triangulation Emotional cutoffs Family emotional system Sibling position	Increased differentiation Detriangulation Cutoffs resolved Improved ability to manage anxiety	Use of genogram Therapist as "coach" Education about multigenerational family processes	Often con with or patient Influence but has since B death
Experiential–humanistic (Susan	Communication styles (e.g.,	Fostering creativity	Resolving the battle for structure and the battle	Since Whi death,

Johnson and Leslie Greenberg, Virginia Satir, Carl Whitaker)	placater–blamer) Psychotherapy of the absurd Attachment theory	(comfort with "craziness") Increased family cohesion Personal growth, self-esteem Tolerance for conflict	for initiative Frequent use of co-therapists Family sculpture Use of self Family reconstruction	experie approa in visib Emotiona therap few no or psyc couple increas influen
Integrative (Alan Gurman, William Pinsof, Douglas Snyder, Froma Walsh)	Importance of and relationships among multiple levels of experience Theoretical and technical eclecticism	Improved communication and problem solving Enhanced interactional insight Presenting problems	Combinations of cognitive-behavioral and psychodynamic methods, in general systems context	Along wit behavi psycho metho most se researc Attempts interve proble family
Mental Research Institute brief therapy (Don Jackson, John Weakland, Paul Watzlawick)	Distinction between "difficulties" and "problems" Communication processes Symptom as communication First-and second-order change	Resolution of presenting problem Second-order change	Reframing Therapist maneuverability	Historicall strateg import differe "functi sympto
Feminist (Virginia Goldner, Rachel Hare-Mustin, Betty Carter)	Raising gender consciousness Underlying effect of gender	Gender equality Recognizing gender inequities	Raising consciousness Identifying and challenging culturally constructed narratives	Not so mu of inter set of is addres pervasi on fam couple
Narrative (David Epston, Michael White, Harlene Anderson)	Constructivism "Languaging" "Stories" socially created Problem-saturated descriptions	Resolution of presenting problem Enhancing undiscovered parts of self	Externalization of problems Focusing on "unique outcomes" Creation of new meaning via "re-storying" Therapeutic letters	Currently popula acknow overlap cogniti behavi
Psychodynamic–psychoanalytic (Nathan Ackerman, Henry Dicks, Clifford Sager, Jill Scharff and David Scharff)	Object relations Projective identification Splitting Scapegoating	Improved insight Genetic (historical) awareness Enhanced empathy De-emphasize presenting problem Disentangle	Transference, resistance, and countertransference analysis Creation of holding environment Interpretation Emphasis on therapeutic alliance	Probably i family t clinical than is acknow

		interlocking pathologies		
Psychoeducational (Carol Anderson, Ian Falloon, William McFarlane, David Miklowitz)	Biopsychosocial theory Stress-diathesis Expressed emotion	Enhanced family coping skills Improved communication and problem-solving skills Relapse prevention	Survival skills workshops Family management Family information and education Concurrent use of psychopharmacology	Without c most e family- approa familie membe major p disorde
Solution focused (Steve de Shazer and Insoo Berg, William O'Hanlon, Michele Weiner-Davis)	Core concepts Focus on solutions, not problems Unimportance of problem etiology Disbelief in resistance	Typical goals Resolution of presenting problem Creation of solutions	Common strategies and techniques Scaling questions Miracle questions Exception questions Client empowerment Coping questions Questions re presession change	Comment Grows ou Mental Institut philoso empha solutio than pr Has made extraor claims effectiv
Strategic (Milton Erickson, Jay Haley, Cloe Madanes, James Keim)	Power and control Family life cycle Symptom-maintaining sequences Function of problems	Resolve presenting problem Disruption of problem-maintaining sequences	Persuasion Paradoxical injections Insight downplayed Pretend and ordeal techniques	Unfortuna empha biology psychia diagno person
Structural (Harry Aponte, Salvador Minuchin, Jose Szapocznik)	Boundaries Hierarchies Coalitions and alliances Complementarity Engagement–enmeshment	Increased flexibility Adaptability to developmental change Balance between connectedness and differentiation Subsystem functioning	Unbalancing Enactment Joining	Arguably influen therap
Transgenerational (Ivan Boszormenyi-Nagy, James Framo, Laura Roberto-Forman)	Invisible loyalties Ledgers and debts Family mourning Personal authority	More universal than family-specific goals Increased trust Repair ruptured relationships	Multidirected partiality Family-of-origin sessions Use of co-therapy	Almost all /transg metho signific psycho underp

As an example of another focus under the umbrella of systemic/structural therapies is the life cycle approach favored by McGoldrick. This practice uses the in-session drawing of a genogram, similar to a genetic diagram or family tree but with much more verbal

commentary by the family. Bowen created this technique to map and flesh out a family's multigenerational history over time. The assumption that underlies this method is that transition points in normal family development can become sticking points for families who now seek therapy. The use of a life stage, transgenerational orientation helps engage resistant families by depathologizing their current struggles and relabeling it as a common and understandable function of family life.

Strategic family therapy—It has its historical base in the early work of Bateson and the Brief Therapy Group at the Mental Research Institute in Palo Alto. Don Jackson, Jay Haley, Virginia Satir, John Weakland, and Paul Watzlawick are well-known members of this school. Over time, the ranks of strategic family therapists have been joined by other prominent figures in the field of family including Maria Selvini Palazzoli and the Milan Strategic Therapy group who have incorporated their own variations on the theme of intervening strategically through the use of a team approach, circular questioning, and attempts to provide alternative understanding for dysfunctional or "stuck" family systems.

Contributions have come from Jay Haley and Cloe Madanes, building of the hypnotic techniques of Milton Erickson who emphasized a new way of thinking about and changing psychological symptoms. Madanes defined six useful parameters that were essential to understand before formulating a plan of intervention. These include assessing whether the symptomatic behavior is voluntary or involuntary, reorienting power dynamics such that the person who acts the weakest in the family is often the most powerful or controlling, viewing symptoms as part of the communication in the system and determining whether or not maintaining the symptom serves a purpose (albeit maladaptive) for the family; evaluating the hierarchical structure and corresponding role that go with them in order to approximate fairness and equality, understanding the motivational basis (benign or malignant) for problematic behavior, and deciding whether or not the troublesome symptom is a misdirected attempt to gain acceptance love and approval.

Once an appreciation of the presenting problem is attained, a plan for intervening in the system is created. Strategic therapists, like their structural colleagues, are very active in sessions and in their use of out-of-session assignments. Stemming from their origins in communication theory, strategic family therapists focus heavily on those forces that generate and reinforce resistance to change in families and work in a brief therapy format alone or with a co-therapist.

Some of the more recognizable techniques include the use of directives, circular questioning, paradoxical messages including prescribing the very symptom for which the family seeks help, assigning rituals or ordeals that disrupt the family homeostasis, reframing behavior in a less pejorative and pathological definition to one that emphasizes the protective nature of

family or individual defenses. Strategic therapists are keenly aware of the context in which the symptom arises in the family. A child's refusal to go to school but his/her willingness to go out of the home to pursue pleasurable relationships with peers can be a context dependent issue worthy of exploration in family sessions.

Strategic therapists take positions in therapy that poke and test the response surface of the family system. Depending upon the reaction evoked, a strategy is designed to alter the process. Families can be compliant, noncompliant, or pseudocompliant in response to directive or prescriptions put forth by the therapist. Noncompliant and responses that are actually resistances disguised as compliance call for astute identification and construction of tasks that tap the roadblocks to change in the family.

An often-cited criticism of strategic therapy is that it is manipulative and nonempathic. On the contrary, reputable strategic therapists are extraordinarily sensitive to the families they are seeing and are aware of the feelings evoked in them when using the strategic interventions. Additionally, strategic therapists are often called upon to see individuals or families who have been refractory to other forms of therapy and where a family may contain a member with a potentially life-threatening condition such as severe anorexia nervosa. When conventional therapies have been ineffective, the creativity, incisiveness, and pragmatism that typify good strategic therapy will often succeed where others have failed.

Experiential/humanistic family therapy is associated with the name of Susan Johnson's emotionally focused couples therapy format. The practice of family therapy has undergone dramatic change since its inception resulting in an array of technical formats that rely heavily on the "here-and-now" experience of the family in therapy. Satir retained her focus on communication patterns but added an emphasis on experiencing emotions as they would arise in a nurturing therapeutic climate. Satir was not afraid of showing compassion by touching family members at times of profound emotional experiences. This approach never encourages boundary violations but does advocate human responses to authentic suffering, struggle, despair, depression, and other intense affective states. A humanistic belief system and an appreciation of the vicissitudes of inadequate or damaging attachment experiences with its consequent feelings of isolation, deprivation, and emotional hunger form the foundation of most of the popular experiential family therapy experiences today.

A positive but not falsely reassuring tone is the backdrop against which therapy is conducted. The person of the therapist is one of the most powerful aspects of this technique. The therapist's positive belief in the family or couple's ability to change, the conviction that people have untapped resources within themselves that can be mobilized in the service

of change, contribute to an optimistic but realistic view that resolution of long-standing problems is possible. Experiential/humanistic family therapy is not only supportive. Confrontation in sensitive but direct ways also forms a regular component of the treatment philosophy.

Interventions designed to liberate submerged affect in present time can be accomplished through family sculpting, role playing, and other action-oriented methods. The overarching goal of experiential/humanistic family therapies is to validate aspects of the family and promote emotional well-being. Problem solving is a desirable but less urgent goal than it is in the structural, strategic, and behavioral approaches. Perhaps the renaissance of interest in humanistic family approaches has to do with a reaction to the underemphasis on the internal feeling state of the individual that received less attention in earlier cognitive/behavioral methods and the impact of the impersonal often insensitive. Regardless of cause, the experiential and humanistic approaches have earned a secure place in present day family therapy.

Psychoeducational Family Therapy—It has been and continues to be a mainstay in the integrated treatment of people with severe and disabling mental illness. The value of small groups for emotional support and the dissemination of accurate information about a condition shared by members have been amply demonstrated in the group psychotherapy literature. In more recent times, the scope of psychoeducational has branched out to include family problems that are child focused. These include attention deficit disorder, autism spectrum conditions, and a range of behavioral and conduct issues. Families coping with a serious medical condition are also ideal candidates for a psychoeducational experience either as a solo family or in a multifamily group format.

In principle, psychoeducation is a biopsychosocial orientation that provides a vital support function for families in crisis. Initially, psychoeducation helps stabilize a traumatized family system through the provision of a supportive and knowledgeable therapist or interdisciplinary co-therapy team. In the case of medical illness like cancer, a team is ordinarily composed of a family therapist and a co-leader with a background in oncology.

Strong emphasis is placed on teaching coping skills to better deal with the condition in question. Much of the original work in psychoeducation addressed schizophrenia and its impact on family life. When applied to psychosis, psychoeducational therapy teaches families several critical skills including how to tell when someone is showing signs of decompensation, how to arrange for rehospitalization, plans for ensuring physical safety of all involved, how to be helpful in ensuring that the schizophrenic family member complies with taking medication and keeping scheduled appointments for outpatient treatment.

Techniques to control the level of expressed emotion (EE) to a level that

is not overwhelming for the psychotic family member are also essential parts of the psychoeducational approach. Studies have clearly confirmed that lowering the levels of EE is correlated with a reduction in the relapse and rehospitalization rate in patients with schizophrenia. Similarly psychoeducational methods in general have been shown to be more effective than other nonpharmacological interventions for the prevention of decompensation leading to the need for hospital readmission.

Carol Anderson and Ian Falloon are two of the best known exponents of this method and have expanded their use of the psychoeducational model to include home visits to hold family meetings, providing intensive, day long workshops for families and coping strategy seminars for family and significant others involved in the life of a person dealing with schizophrenia. The thrust of these meetings is to help the family understand mental illness as a biologically driven entity and to avoid the temptation to blame, ridicule or chastise the affected person. Self-care for those family members who are charged with overseeing the daily life of a psychotic family member is strongly advocated in order to avoid caregiver burnout.

The psychoeducational team's stance is very supportive and, in a nonpatronizing way, synthesizes and translates the medical and psychiatric data to the family in a form that they can convert into constructive action. Mc Farlane has done just this in his list of guidelines for family and friends of schizophrenics. These guidelines orient the family along several parameters: recovery is a slow process, set realistic limits, avoid harsh confrontation or arguments, comply with the psychiatrist's recommendations regarding medication and frequency of visits, avoid drugs and alcohol, have modest goals for the patient, and expect change to occur in small steps over time. This combination of distilled scientific knowledge and applied common sense are emblematic of the psychoeducational model of family therapy.

Psychoeducational practitioners choose to use the multi-family group composed of as few as two or as many as ten families as participants. The advantages of this format is that it enhances the support elements for families, de-stigmatizes the shame in having a psychotic family member, allows for cross-family learning by seeing in other families in the group what one is unable to see in themselves, facilitates a broader exchange of information and coping strategies among peers, and counters the isolation that often evolves in families consumed by having to care for a medically or psychiatrically ill member.

Psychodynamic family therapy—The thread that runs through psychoanalysis to modern psychodynamic therapies has followed a linear path. Many early family therapists such as Ackerman, Bowen, Jackson, and Minuchin were trained psychoanalysts. Henry Dick's classic book *Marital Tensions* offers an early recognition of the differences in transference and

other analytic techniques when working dyadically as contrasted to individual psychotherapy. How to directly apply, adapt or discard time-honored elements of the analytic model transition to make the transition to family work was the first challenge these therapists faced.

Would it be relevant and appropriate to use the basic cornerstones of psychoanalysis, the analysis of transference and resistance, dream interpretation, the nondemand posture of the analyst, the quest for access to the unconscious mind, and the focus on early childhood experiences as the vehicles for treating families? The history of modern psychiatry reveals that many of the components of psychodynamic theory and practice are applicable to other therapies including group, family, and couple therapies, but that significant modifications in technique and attitudinal shifts on the part of the therapist are required in order to do so.

Freudian and earlier analytic models have given way in favor of object relations and self-psychological perspectives. Generally speaking, exponents of the modern psychodynamic school believe that viewing interpersonal as well as individual themes is essential if dynamically oriented family therapy is to succeed. Systemic concepts have been grafted onto the exclusive focus on the inner world of the individual resulting in a broader canvas on which the family therapist works. Empathy, establishing a therapeutic alliance, functioning as a container for intense affects, supplying a secure holding environment for each individual and for the whole family, and providing a tight frame within which to conduct the therapy are all parts of the modern psychodynamic method. Interpretation is used more in this model than in any other family technique. The presenting problem is important but forays into the past life of the family are actively encouraged.

The leadership style is more active than in traditional dynamic therapies but there is ample room for the development of transference. Using transference as the prototype for how psychodynamic family and couple therapists view modify and work with traditional analytically-derived concepts with families and couples, the object relations orientation as advocated by Jill and David Scharff affords an excellent example. In this model transference in family and couples exists on three levels: contextual transference, focused transference, and shared contextual transference.

Contextual transference relates to the early phase of treatment when the role of the therapist is to identify and address, in a supportive and nurturing manner, those initial anxieties that families have about unfamiliarity with therapy, the administrative aspects of treatment (fees, scheduling, physical setting), and anticipating early resistances. The therapist strives to convey a feeling of emotional safety and serves as a positive “holding environment” or container for real or imagined fears about the therapeutic process.

Focused transference is most closely allied to the transference

manifestations in conventional psychoanalytic therapy. It consists of the projections onto the therapist, distortions or misinterpretations of the therapist's words or actions, and other present day phenomena that reflect unresolved issues stemming from early childhood experience. Shared contextual transference is easiest to see when working with couples. It entails understanding a couple's response to a holding environment, a critical issue for dyadic relationships insofar as one of their major tasks is to create a secure holding environment for each other.

Additionally, recognition of regressive recreations of earlier family configurations, examining triadic transferences, and appreciating the transference issues that determine the process of partner selection in relationships are all illustrative of how transference, a core concept in psychoanalysis has been creatively incorporated into current family and couple therapies.

Narrative and solution focused therapies are other prominent postmodern constructivist theories about family life. The premise in narrative therapy is that the shared narratives constructed over time create the perception of reality embraced by the family often resulting in narrow, rigid family myths and misconceptions. The therapy endeavors to broaden the family perspective by helping the family construct new, shared stories leading to more accurate narratives that, in turn, mediate a change to new behaviors.

Solution-focused therapy is, as the name implies, a quest for positive resolutions to dilemmas with which the family is struggling. Through the use of a series of techniques that involve pinpointed questions geared to expand the cognitive and behavioral repertoire of the family, the therapist guides people toward new ways of thinking about old issues relying heavily on positive reinforcement, an optimistic attitude that looks toward solutions not problems, and a search for exceptions to the situations that the family feels mired down and powerless.

CLINICAL APPLICATIONS

General Clinical Principles

Family therapy and individual psychotherapy may be viewed as opposite poles of a spectrum of therapeutic interventions each requiring different modes of engagement. Individual therapy is a new relationship where interaction between therapist and between therapist and patient stimulates affects that form the basis for examination in dynamically oriented individual work. Initial goals are to engage the patient and create a relationship into which strong feelings, both past and present will develop.

In family and couple therapy the opposite is the case. Families and couples have relationships and histories that predate the new therapy and usually have a surplus of affect that is out of control. Coming in when a

crisis has arisen is a frequent entry posture of couples and families. These circumstances call for decelerating interventions designed to stabilize the system. Affective intensity needs to be reduced to a degree that the family system can gain a sense of equilibrium. The family therapist has to find a point of entry into the current family configuration.

With this principle in mind, it makes clinical sense to understand the strengths and limitations of family and couple therapies and for whom they are best suited and the common clinical management issues that occur that require thoughtful treatment planning and interventions.

Indications and Contraindications

Family and couple therapies are relational therapies. The clinician is primarily focused on the unresolved interactional issues that the family or couple presents at the time of the initial interview rather than on individual diagnoses. Despite this, since family and couples therapy has been demonstrated to be useful in specific psychiatric conditions, some practitioners prefer to use Diagnostic and Structural Manual V for deciding when family therapy is indicated. In these circumstances, conditions including schizophrenia, bipolar disorder, major depressive disorder, and substance abuse are the yardsticks by which a determination is made about suitability for a family or couples approach. The majority of family therapists prefer a more systemic or group-based orientation when thinking about the indications and contraindications for family or couples therapy. The following is an integrated model that incorporates both diagnosis and presenting problem as criteria for recommending a family or couple form of therapy.

Indications for Family and Couple Therapy

1. When problems exist across generational boundaries—Grandparents feel excluded from access to grandchildren, conflicts between a father and son or daughter working together in a family business, sharing the living space with more than one generation in the family.
2. At significant times in the family life cycle—Leaving home, birth of a child, a geographical move, death in the extended family.
3. Whenever a child or adolescent is symptomatic and designated as “the problem”—Adolescent acting out with drugs, alcohol, and promiscuity; cutting; children who refuse to attend school and/or have psychosomatic symptoms.
4. When a couple or family calls or comes in requesting a family therapy.
5. When serious medical or psychiatric illness is present in a family member—cancer, schizophrenia, bipolar disorder.
6. Families stressed due to racial, cultural or religious heritage, interfaith, and interracial partnerships.

7. Issues that stem from recombined or blended families—friction between step-parent and biological child; problems with role definition as parents; children who feel themselves to be in a loyalty bind between divorced parents.
8. Families in which someone intrudes upon or sabotages another member's individual psychotherapy—intrusive parents threatened by the independence of the adolescent or young adult in individual therapy; parental abuse of power by refusing to pay for treatment.
9. Families or couples where societal pressure and prejudice has significant emotional spill over into family life—gay parents with children; older couples contemplating adoption of a child; common-law partnerships.
10. When sibling issues are beyond the control of the parent(s)—physically assaulting a younger sibling; excessive competition for parental attention and affection.

For couples:

1. When the chief complaint is presented as a sexual issue—This may indeed be an indication for sex therapy but other unresolved relationship issues masquerading as a sexual problem should be ruled out first.
2. Following the discovery or disclosure of an affair—the threat of divorce or break up; restoring the power imbalance in the dyad.
3. When communication problems are the basis for seeking treatment—escalating anger; repetitive arguments without resolution; emotional detachment and break off of communication.
4. When partnerships are ending and children are involved—partners seek therapy not to resolve relationship issues but to minimize damage to children; decision-making that involves co-parenting.

Contraindications for Family and Couple Therapy

1. When a family has an emotionally unstable member or members and the risk of stirring up intense affect in session might lead to decompensation or other adverse effects—schizophrenia, major depression, impulse control disorders, addictions, dementia.
2. When there is a history of violence in the family—elder abuse; domestic violence, drug and alcohol dependence.
3. When a family is collusive, rigidly unified, and inflexible in considering that there may be a problem in the family/couple.
4. When therapy runs counter to the belief system in the family—sense of shame; religious or cultural beliefs discourage therapy.
5. When essential members of the family cannot or refuse to be included in the sessions—father feels that "real men" handle problems on their own; family member offers work schedule as a reason for being unable to participate.
6. When family therapy should not be the first line treatment—

detoxification of addicted family members should precede decision for family/couples therapy; psychotic family member who is not adequately medicated.

For couples:

7. When a pairing defense is so strong that both partners recoil at the suggestion of couple therapy; rare relationships where a folie a deux exists such that the couple has a shared delusional system and/or a rigid, ego-syntonic view that they do not require therapy.
8. When the level of motivation is unequal or deceptive—one partner has not disclosed his/her plan to leave the relationship.
9. When a member of the couple has an active relationship with someone other than the partner—extramarital affairs; homosexual relationship simultaneous with heterosexual primary partnership.
10. The threat of domestic violence or crisis creation in vulnerable partnerships—history of physical abuse.
11. In cases of documented child physical, sexual and/or emotional abuse, extreme caution should be exercised before exposing the child to the abusing parent(s) even in an ostensibly therapeutic and safe climate of family therapy.

Assessment and Evaluation

Evaluation procedures take numerous forms in family therapy. It is important to note that aside from data collecting, the initial contact with the family or couple may simultaneously be the first session of their family therapy experience. The most common approach is to see the whole family for family therapy evaluation and both members of the couple for couple therapy. The closer the therapist subscribes to a strict systemic approach the more the requirement that all members be present. Alternatively, other family therapists work with whatever portion of the family system presents itself for the initial session with the thought in mind of recruiting absent members by using the existing forces within the family.

The usual time allotment for a first family or couple session is one hour but time-extended sessions are common if, for logistical or clinical reasons, a longer meeting would allow for the most comprehensive view of the family/couple system.

Seven global areas are subject to assessment in the first meeting with a family or couple: the presenting problem (stated as the family defines it); communication patterns; styles of problem-solving; regulation of affect; family roles—alliances, splits, coalitions; family rules; impulse control; and the prevailing family beliefs, traditions, and narratives.

Pragmatically, the data for the assessment comes from the therapist's observation of the spontaneous interactional patterns and from the specific questions posed by the interviewer. Many of the techniques used in family

and couple therapies are interchangeable and applicable to both. The material that follows details of the conduct and content of an initial evaluation of a couple with particular emphasis on the role and posture of the therapist but the points outlined apply equally when modified for working with a family format.

Couples Therapy Initial Interview

1. Orient the couple to the joint format—emphasize the need for both partners to be present in order to have input and how couple therapy differs from individual psychotherapies.
2. Request that each person state his/her perception of the problem—a majority of couples experience the locus of disharmony as being in the partner rather than in themselves.
3. Try to engage both partners in the first session—balance all interventions especially praise and confrontation such that neither partner feels rejected.
4. Identify and clarify relationship patterns and structure—is the present problem ongoing since the relationship began or is it in reaction to new circumstances.
5. Formulate the problem without singling out either partner—shift the focus from the identified patient model to the identified problem model —the question is often posed in the following form, "How long has this been a problem for the two of you?"
6. State the purpose and some goals of couples therapy in nontechnical language or jargon—to resolve communication problems; to learn how to fight fairly; to address specific issues around parenting.
7. Think of ways of gaining entry into an established system—wait for a naturally occurring pause before entry or try to interrupt a pattern to see how the couple responds.
8. If couples therapy is indicated, try to define an initial treatment contract and get agreement of both partners to participate—plan for a next meeting and give some indication of what the general focus might be, that is, communication issues.
9. Anticipate early resistances—denial defenses with alcohol and drug related issues; the temptation to stay with familiar behavioral patterns even though they may not be working but are less frightening than the prospect of change through couple therapy.
10. Clarify myths and misconceptions—joint therapy is more likely to lead to dissolution of the partnership; being in joint therapy is an indication of an unworkable relationship.
11. Leave time for the couple to ask questions about anything in the session that was unclear or worrisome—common concerns are the expected duration of therapy, cost factors.
12. Assign "homework" if indicated—instruct them to arrange to spend five

to ten minutes each day or alternate days in face to face conversation on any nonconflictual subject without outside distractions that include television, children or other impediments to engagement with one another.

3. Summarize what you want the couple to take away from the first meeting —point out the themes that appear to be urgent ones; commend the couple for being open enough to explore the possibility of joint therapy.
4. Delineate what your role will be in the therapeutic process—make sure to emphasize how a collaborative treatment model works and that the therapist will be an active participant in the ongoing therapeutic process and will be unavailable for individual phone contacts or sessions that usually result in telling "secrets" to the therapist.
5. Address "housekeeping procedures"—appointment times; frequency and length of sessions; fees; what happens when one partner cannot attend a session; policy about individual sessions during the course of joint therapy; get permission that is HIPAA compliant for communication with other (individual) therapists; obtain informed consent from both partners.
6. Stop on time unless there is a crisis which cannot be handled in any other way than by extending the session—commonly a catharsis of intense feelings is opened up in a first meeting, allow time for both expression of emotion and putting a cognitive structure on it in order to gain closure on the session.

These are the general clinical tasks for practitioners of diverse theoretical orientations to cover in the first contact with a couple. In order to insure that relationship relevant themes are included in the interview, it is essential that the therapist have in mind a set of variables that will glean the information he or she needs in order to understand the partnership issues and to begin formulating a treatment strategy. Listed below is a practical set of questions and areas of observation that will yield the information most relevant to creating and directing the ongoing treatment process.

Specific Areas of Inquiry and Observation in Couple Therapy

1. Why do they seek evaluation now?
2. Opening position/posture of each person in the session
3. Locus of the problem
4. Communication styles
5. Collaborative abilities
6. Problem solving preferences—insight or behavioral preferences
7. Families of origin of both partners—three generational model starting with grandparents.
8. Overt and covert relationship "contract"

9. Methods of managing intense affect
10. Function of the presenting symptom(s) in the relationship and/or family system
11. Assessment of individual as well as joint strengths and assets
12. Power and control issues
13. Role definitions and gender based issues
14. Mechanisms used to control interpersonal closeness and distance
15. Work function of each partner
16. Sexual issues
17. Prominent defenses—individual and joint
18. Duration of the disharmony
19. Examples of historical crisis situations and their outcome
20. Motives for selecting each other initially and for staying together currently
21. Parenting issues
22. Reaction of each family of origin to the partnership
23. External stressors on the relationship—work, time, money, illness, sleep deprivation
24. Interpersonal skills outside the relationship
25. Ability to have pleasurable and restorative experiences
26. Emotional resiliency of each partner and of the relationship
27. Ethnicity, religious and cultural and gender issues
28. History of significant intimate relationships prior to the present one
29. Boundary issues—within the relationship and between the relationship and significant others especially with original families
30. Involvements outside then relationship—"Triangulation"
31. Substance use and abuse
32. Joint and individual life cycle factors
33. Questions about couples therapy

When converting principles into practice, the process is rarely without potential pitfalls that call for prompt recognition and astute clinical management. In both family and couple therapies, it behooves the therapist to anticipate that there will be resistance to change. More often than not, this is fear based and can be traversed if the therapist is alert for common roadblocks to progress. To illustrate this point, a representative sampling of the types of clinical management decisions that confront the therapist working with couples and some guidelines for handling them are highlighted next.

Common Clinical Management Issues

1. The absent or resistant partner or family member—joint sessions are not held if one partner (with families, if significant family members) is not present. Doing so runs the risk of unbalancing the therapy, being perceived

as participating in a collusive relationship with the partner who is present and risks being told "secrets" by the person seen individually. It also tacitly appears to be sanctioning the idea that that the presence of both partners is not necessary for successful couples therapy.

2. Distinguishing true emergencies from "pseudocrisis" situations—a common resistance to dyadic therapy is to try to create a crisis designed to have an individual relationship with the therapist. The form this very often takes is an urgent individual phone call between sessions. The therapist has to rapidly assess whether there is indeed a true emergency that requires immediate attention or if this is an effort to sabotage the joint therapy. In the latter case the caller needs to be informed that the content of the call will be shared with the other partner in the next couple session.

3. Communication with other current or past therapists—pivotal in formation can be obtained from prior individual, family or couple therapists referable to what has worked well and what has not. Obtain informed consent from both partners but make this a non-negotiable condition of therapy.

4. Problems with payment for sessions—delayed payment, disputes about fee arrangements and a host of money related issues not only serve as points of resistance but may also be informative that money matters are an issue for the relationship itself.

5. Management of psychotropic medication issues both pharmacologically and psychodynamically—when one member takes medication it is important to insure that that person is not seen in disadvantageous or "one down" role.

6. Management of "secrets" in relationships and in original families—the therapist should make every effort to avoid situations in which one partner tells something that is not known to the other. This presents a threat to the viability of the entire therapeutic enterprise and must be handled immediately. Not permitting individual sessions is the most common method of handling secrets. If by chance, a person tells a secret by phone or by lingering after the other partner has left at the end of the session the therapist needs to clearly communicate that it is the person's responsibility to share this information openly or else the therapist will decide how best to reveal the content to the absent partner.

7. What happens when one partner plans to or has seen a lawyer—many individuals have different motives for seeking couples therapy. In instances where one has contacted a lawyer, the underlying agenda may be one of guilt reduction by leaving the partner in the hands of the therapist rather than his/her stated desire to work on the relationship. In these instances the covert actions of the dishonest partner is a priority issue to be confronted in the next session. The obvious risk is to have one person exposed and vulnerable thinking that they are working toward problem resolution, only to be abandoned by their partner who plans to end the

relationship.

8. Videotaping of sessions—the purpose of taping whether for review by the therapist between session to show the couple aspects of problematic and effective communication or for purposes of training and supervision always need to be clearly explained. Obtaining informed consent from both partners is a precondition for audio or videotaping sessions.

9. Confidentiality—as is the case in all psychotherapies, confidentiality is a sine qua non of ethical family and couple therapy practice. Unique forms of confidentiality with couples can occur in a couples group where several couples are privy to intimate details in the lives of other group members. Breaches of confidentiality are grounds for dismissal from the group.

10. Whether or not (and when) to include children or extended family members in sessions—couples with children and those with living parents frequently raise issues in joint sessions that involve family members who are not present. One of the unique aspects of couple and family therapy is flexibility. Couple formats allow for the inclusion of other family basis or for conversion of couples therapy to larger family groupings in order to address issues in the larger system.

11. How to handle intense negative affect in sessions—rage, blaming, threats to divorce, potential physical violence are all alarming phenomena in couples therapy sessions. The therapist has to have a tolerance for intense affect, know when to intervene before a dangerous situation escalates, set a clear set of ground rules for the modes of expression of anger and help the couple find more viable mechanisms for dealing with conflict.

12. How to avoid repetitive, unproductive or destructive interchanges in sessions—a consistent property of systems is the repetitive, cyclical nature of behavioral sequences. Couples stalled in these ways tend to get demoralized and despair over the fate of their relationship. Couple therapists can supply input in the form of feedback both of a general nature and specific observations in the here and now in order to supply new information as fodder for interrupting the recycling of old arguments.

13. How to provide a different experience in the session from what the couple would be inclined to do spontaneously- asking a couple to do something in the session that they are not doing on their own is a central concept in all couple and family therapies. Asking a couple with a problem of sexual inhibition and avoidance to spend the meeting talking about sex allows them to experience that they possess the capacity to do something in the session that they feared might not be possible.

14. What constitutes appropriate timing of interventions in treatment—although couple and family therapies are brief therapies as compared to individual psychotherapy, they still go through definable early, middle, and termination phases. Therapists need to time interventions accordingly.

Building a therapeutic alliance must precede the middle or working stage of therapy otherwise interventions will be ineffective. Similarly, termination in couples may feel abrupt to patients who have been used to the ritualized, extended termination techniques of individual approaches.

15. Co-therapy considerations—when a team treats a couple many advantages accrue but potential pitfalls need to be monitored as well. Splitting the co-therapists is the most common defense against the team format. Opposite sex co-leaders are seen as parental figures and mobilize feelings related to family of origin in both partners. Ideally, the advantages of co-therapy are the increased option for modeling aspects of healthy communication, respect for differences and a working partnership between adults.

16. How to make the therapeutic experience one that is sensitive to issues of diversity—It is essential that the couple therapist make him/herself as finely attuned to the variables that fall under the heading of issues related to diversity in the partners that they are seeing. Diversity requires modification of therapeutic understanding and technique in ways that are respectful, free of bias and potentially negative preconceptions. Race, social class, culture, disability, age, gender, financial status, ethnicity, sexual preference, and religion are noteworthy examples of factors that must incorporated into any treatment plan in order to avert formulaic approaches that disrespect or ignore facts that are central to the identities of the partners. The therapist needs to be vigilant in self-monitoring that his/her own beliefs do not insinuate themselves in ways that would hinder or destroy the therapeutic alliance.

COUPLES GROUP THERAPY

Few modalities reflect the interface between family, couples, and group psychotherapies better than a group composed of couples. Couples groups also exemplify the growing trend toward more integrative approaches in family work. In spite of their popularity and utility, couples groups are poorly described in the professional literature. The prototypical couples group is an outpatient modality composed of three to six couples led alone or by a co-therapy team. The group typically meets once per week for a session of one and a half hours. Groups can be brief, time-limited cognitive/behavioral or skills acquisition training models or longer-term psychodynamically focused models with more ambitious individual as well as couples goals.

The benefits for patients including couples in groups is to counter isolation, destigmatize the idea of having relationship problems, increase the options for handling issues differently by seeing how other couples manage them, enhance the possibilities for varied feedback, experiment with new communication opportunities, and counter depressive and demoralized states that invariably accompany the frustration of

unsuccessful efforts to make progress.

Couples groups have distinct advantages for the therapists as well. Co-leadership is very desirable in terms of sharing the workload. Working with a partner allows for role modeling and demonstrating behaviors for the couples in the group that reflect alternative modes of handling individual differences, equally distributing power and control, clarifying roles of men and women as individuals and as partners and opening up issues that couples handle by avoidance. The presence of a co-leader permits the balance between therapist support and confrontation of members by having a backup system not present when one works alone. When it is necessary to confront denial defenses in a group of couples with substance abuse issues, co-therapy facilitates one leader's ability to take a stronger stance than he/she might when working solo. If necessary, the other therapist can offset any unforeseen negative fallout if it occurs. Alternating these roles between co-therapists over time demonstrates productive styles of collaboration for the group to experience.

For the psychodynamically-oriented couples therapist, seeing couples in a group milieu exponentially increases the number and variety of transference possibilities in the leadership and membership vectors. The co-therapy team is often perceived by members as having a therapeutic marriage despite the group's awareness that they might be colleagues who have chosen to work with each other. It is also wise to choose co-leaders who have different therapeutic styles but who show mutual respect and readily incorporate suggestions from their co-leader. Therapists with an object relations preference can see the inner world of members enacted in their transactions with other part objects represented by the group leader(s) and fellow members in the group.

CASE VIGNETTE

During the middle phase of a couples group composed of four couples, the theme of whether or not to have children arose. One couple had just come from a visit to the gynecologist who informed them that they were running out of time because of the wife's age. The woman in the couple did not want to have children but her husband did. His complaint about the marriage was that his wife never was demonstrative in showing her loving feelings for him. He felt her to be detached, distant, and sexually inhibited.

The prevailing sentiment among the other couples who had children was that children only added additional stress to an already stresses relationship. One other couple, however, voiced their different view by describing how their children had enriched their lives. As the talk about going forward and getting pregnant progressed, the group leader noted the nonverbal communication between the ambivalent couple. Whenever the tone of the group leaned toward having children, the wife would reach out and grasp the hand of her husband in a tender way. This invariably had the

effect of stopping him from pursuing the topic for fear of the withdrawal of the affection he hungered for. All this transpired without words.

Once identified, this repetitive nonverbal pattern was available for examination in the group and the supportive elements provided by other members and the leader encouraged a frank, direct, and open conversation between the partners who eventually chose to go forward and attempt to have a child.

Couples groups share many commonalities with multifamily groups in terms of cross-relationship learning, psychoeducational elements, strong support elements common in homogeneous groups and as a vehicle for reality testing ideas that are inhibiting progress.

Therapists who are experienced with the couples group modality frequently report that they have more therapeutic leverage once an initial base of group cohesion is established. Group cohesion acts as a supportive buffer that allows the therapist greater freedom to use appropriate anxiety-evoking interventions or apply group pressure as a force for change.

A common indication for couples groups is for couples with problems of long duration (years rather than months) where it is felt that the combination of cohesion, a strong therapeutic alliance, the multiplicity of feedback from the group and the leaders, plus the option to apply pressure when needed add up to the empirical enthusiasm expressed for this format.

RESEARCH IN FAMILY AND COUPLES THERAPY

Family therapy research presents several challenges. In the broadest sense, family research to date is a robustly evolving process resulting in a much clearer delineation of the significant variables to be taken into account in order to produce clinically relevant, scientifically informed studies. Part of the problem in conducting credible family and couples research on outcome resides with the multidimensional nature of individual, couple, and family and therapist levels present in all family work. What and who is being measured are central concerns as is the selection process of how cases are being assigned for particular therapeutic interventions.

Therapist variables such as theoretical orientation, level of experience, how issues of diversity are managed, gender similarities and differences, and the precise form and timing of technical choice all are being studied in the current research climate in family and couples therapies.

More specificity is being applied to studies in constructing instruments for measuring the outcome of family approaches, understanding what contributes to the process of change in family work, and in randomizing families under study. In particular, greater attention is being paid to clarity of intervention technique with tendency to use those methods that can be replicated under in other settings by other practitioners.

The notion of goal specificity in family and couple therapies will be followed. In cases where a family seeks therapy for help with a member

with a chronic mental illness the target goals can be reduction in caregiver burden, change in familial interaction patterns, regulating affect in the family, improved family ability for shared problem solving or all of the above. Clear and attainable goals make outcome appraisal of treatment an easier process.

Comparisons of different schools of family and couples therapy have not demonstrated superiority of one technique above another. Studies have shown documentable empirical support for behavioral and long term, supportive, psychoeducational family therapy for schizophrenia. Behaviorally based social network therapy for alcoholism and substance abuse, and a range of childhood behavioral problems that include adolescent delinquent behavior and substance abuse, phobic states in preadolescent children, and aggressive behavior in children and teenagers are emerging.

Problems associated with various aspects of parenting were demonstrated to show benefits following parental stress management training and behaviorally focused parental problem-solving education.

Other interventions that show promise in early studies but lack sufficient large-scale data at present include parent management training (PMT), functional family therapy (FFT), and multisystemic therapy (MST). Adjunctive family and/or couples therapy have a strong foothold as part of a more global treatment plan for conditions that encompass anorexia, alcohol and substance abuse, and bipolar disorder and schizophrenia.

Where great progress is being made in family and couples research is in aligning research findings with direct clinical application, thereby reducing the gap between family studies and family practice.

Couple therapy research is also showing great promise in defining those attributes that contribute to better outcome. Of the three prevailing orientations in couples therapy, psychodynamic, behavioral (BCT), and emotionally focused therapy (EFT) only the latter two have found empirical research support for their utility. Social exchange theory informed BCT while attachment theory formed the underpinnings for (EFT). BCT has added cognitive components to the model and research to date has shown appreciable advantage of the combined approach.

EFT, however, has been found to be valuable in several areas of couple functioning: emphasizing the attribution of meaning that partners incorrectly ascribe to their partner's verbal communication and behavior; teaching the exchanges that promote secure attachment; identifying, editing, and supplying alternatives to the number and forms of negative cycles of behavior; stressing the beneficial use of intense emotion; and delineating maladaptive interactional patterns that lead to dissolution of the relationship.

Couples who began therapy with lower levels of initial distress were also noted to have better outcome in couples therapy. The work on emotional

attachment themes has been a powerful addition to new approaches capitalizing on repair of childhood damage that manifests itself in bilateral interactional problems.

In EFT and psychodynamic approaches, the stronger the therapeutic alliance and the more both partners of the couple are engaged in therapy, the better the outcome in terms of problem solving, communication alignment, reduction in hostility and blame, perceived partner sensitivity, and a softening of the negative affective tone present upon entry into treatment.

Trends that are going to remain in family and couples therapies will include reducing the distance between statistical studies and their practical application. Therapy protocols that can be standardized for use by many practitioners along with suggested treatment guidelines for the condition that causes the family or couple to seek help. Ethical standards that deal with issues of confidentiality in family sessions will always be a key part of the therapy.

Experimentation with the advantages of newer techniques such as computer-based treatment plans, telepsychiatry, the use of television equipment are aimed for shortening the distance between geographically inaccessible families and couples and their therapy, creating algorithms for formalizing the sequence of treatment approaches to more difficult family and couple presentations of symptoms. Education and training will include online, computer-based courses for advanced learning of the latest techniques and developments in the field.

These trends clearly indicate that family and couple modalities of treatment will be in the vanguard of clinical practice and research.

References

Anderson CM, Reiss D, Hogarty B. *Schizophrenia and the Family*. New York: Guilford; 1986.

Atkinson BJ. *Emotional Intelligence in Couples Therapy: Advances from Neurobiology and the Science of Intimate Relationships*. New York: WW Norton; 2005.

Bigner JJ, Wechtler JL, eds. *Relationship Therapy with Same-Sex Couples*. Binghamton, NY: Haworth Press; 2004.

Clarkin JF, Carpenter D, Fertuck E. The state of family research: a positive prognosis. In: Sholevar GP, Schwoeri LD, eds. *Textbook of Family and Couples Therapy: Clinical Applications*. Washington, DC: American Psychiatric Publishing, Inc.; 2003.

Diamond GS, Diamond G, Levy S. *Attachment Based Family Therapy for Depressed Adolescents*. Washington, DC: American Psychological Press; 2013.

Fields L, Morrison TL, Beels CC. Couple and family therapy. In: Hales RE, Yusofsky SC, eds. *The American Psychiatric Publishing Textbook of Psychiatry*. 4th ed. Washington, DC: American Psychiatric Publishing, Inc.; 2003.

Framo JL. *Family-of-Origin Therapy: An Intergenerational Approach*. New York: Brunner/Mazel; 1992.

Green RJ. LGBTQ Families. *Fam Process*. 2015;54:391–575.

*Gurman AS, Jacobson N, (eds). *Clinical Handbook of Couple Therapy*. 3rd ed. New York: Guilford; 2003.

Heyman RE, Jongsma Jr AE. *The Couples Psychotherapy Treatment Planner*. 2nd ed. Hoboken, NJ: John Wiley & Sons; 2015.

Johnson SM. *The Practice of Emotionally Focused Couple Therapy*. New York: Brunner-Routledge; 2004.

Johnson SM, Zuccarini D. Integrating sex and attachment in emotionally focused couple therapy. *J Marital Farm Ther*. 2010;36(4):431–445.

McFarlane WR. *Multifamily Groups in the Treatment of Severe Psychiatric Disorders*. New York: Guilford Press; 2002.

McFarlane WR. Family interventions for schizophrenia and the psychoses: a review. *Fam Process*. 2016;55:460–482.

Megan MJ, Hecker L, (eds). *Ethics and Professional Issues in Couple and Family Therapy*. 2nd ed. London: Routledge/Taylor & Francis: 2016.

*Nichols MP, Schwartz RC. *Family Therapy: Concepts and Methods*. 7th ed. Boston, MA: Pearson Education Inc.; 2006.

Pinsof WM, ed. Special issue: marriage in the 20th century: trends, research, therapy, and perspectives. *Family Process*. 2002;41:2.

Powell B, Cooper G, Hoffman K. *The Circle of Security Intervention: Enhancing Attachment in Early Parent-Child Relationships*. New York: Guilford Press; 2013.

*Sandberg JG. Introduction to the special section on learning emotionally focused couples therapy. *J Marital Fam Ther*. 2011;37(4):377–428.

Sextan TL. *Handbook of Family Therapy*. 2nd ed. London: Routledge/Taylor & Francis; 2015.

*Sprenkle DH. Intervention research in couple and family therapy: a methodological and substantive review and an introduction to the special issue. *J Marital Fam Ther*. 2012;38(1):3–29.

Watzlawick P, Weakland JH, Fisch R. *Change: Principles of Problem Formation and Problem Resolution*. New York: Guilford Press; 2002.

*Williams L, Patterson J, Edwards TM. *Clinician's Guide to Research Methods in Family Therapy: Foundations of Evidence Based Practice*. New York: Guilford Press; 2014.

▲ 33.7 Cognitive Therapy

JUDITH S. BECK, PH.D., AND ROBERT HINDMAN, PH.D.

INTRODUCTION

Cognitive therapy, a form of Cognitive Behavior Therapy, was developed by Aaron T. Beck, M.D., in the 1960s and 1970s. He and colleagues worldwide have continued to expand and refine the underlying theory and treatment ever since and Beck's work, emphasizing the central role of cognition in psychopathology, has led to a paradigm shift in the field. It is a

comprehensive biopsychosocial system of psychotherapy, with an operationalized treatment based on an elaborate and empirically supported theory of psychopathology. It has been demonstrated in well over 1,000 randomized controlled trials to be effective for a wide range of psychiatric disorders, psychological problems, and medical conditions with psychological components. Research shows that cognitive therapy also tends to reduce relapse.

This form of psychotherapy has been adapted and studied in a variety of settings, including outpatient clinics, psychiatric hospitals, primary care offices, schools, prisons, and community mental health centers. Research has demonstrated its efficacy for young children, adolescents, adults, and older adults from a variety of cultures in countries around the world. Studies support its efficacy in individual, couples, family, and group treatment. A substantial number of studies now demonstrate its efficacy in computer-based programs.

Cognitive therapy is based on the cognitive model, which has been widely studied and has received substantial empirical support. The cognitive model posits that it is not situations that directly impact one's reaction (emotional, behavioral, and physiological). Rather it is one's automatic *perception* of the situation that is more closely associated with the reaction. A depressed individual, for example, who is reading this chapter (situation), might think, "I can't concentrate. There is something seriously wrong with me. I'll never get better," ("automatic thoughts"). As a result, his sadness and hopelessness increase (emotional reaction); he stops reading altogether (behavioral reaction); and he feels a weight in his chest (physiological reaction). On the other hand, another individual might think, "I don't want to read this. I don't want to know anything about this topic but I have to." She then feels annoyed, her body becomes somewhat tense, and she skims the pages instead of reading the article carefully. One important component of cognitive therapy involves teaching patients to identify the thoughts that lead to distress or dysfunction, to evaluate their thinking, and to effectively respond to these cognitions. Figure 33.7–1 depicts the cognitive model.

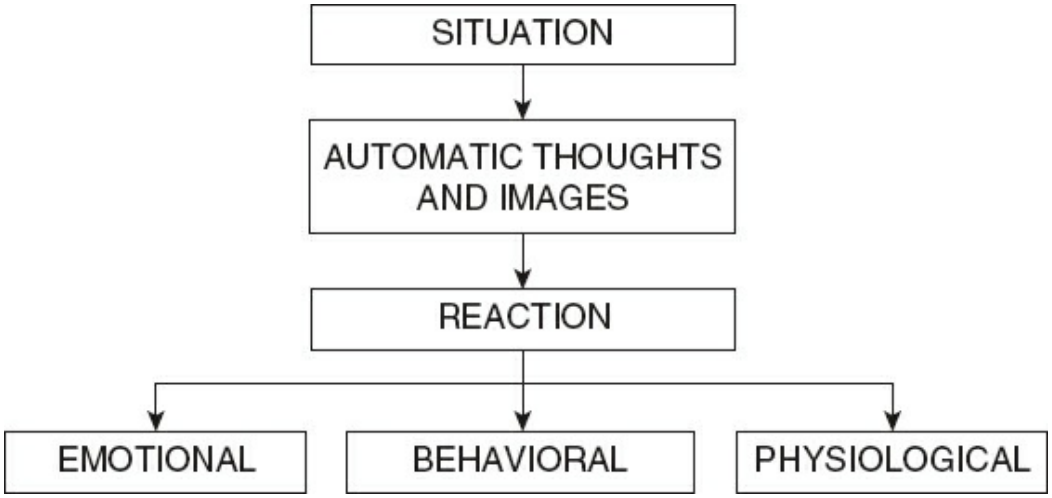

FIGURE 33.7–1. Cognitive model.

Beck and colleagues developed a manual for cognitive therapy for depression to facilitate treatment fidelity in research studies. Manuals, however, often fail to allow clinicians to vary therapy as appropriate for the individual patient and most clinicians find their treatment is significantly more effective when it is based on a cognitive formulation of the patient's disorder(s) and on an individualized cognitive conceptualization of the patient.

A strong therapeutic alliance is essential for effective treatment and the relationship is collaborative. Both patients and clinicians are active in solving problems, evaluating cognitions, and modifying behavior. Cognitive therapy is educative; clinicians teach patients how to use the techniques at home that they have employed in session. The clinician's stated goal is to help patients become their own therapists. They emphasize that the way patients get better is to apply what they have learned in session to make small changes in their thinking and behavior every day.

Originally designed as a short-term treatment for depression, cognitive therapy has evolved over the years. Patients with straight-forward cases of anxiety and depression, who were relatively psychologically healthy before the onset of their disorder, often benefit from a short course of treatment, perhaps 6 to 12 sessions. At the other end of the spectrum are patients whose disorders are chronic or complex, or those for whom standard treatment needs to be varied due to personality disorders. They may need many more sessions over a significantly longer period of time. Those with recurrent episodes or with serious mental illnesses may require intermittent treatment through their lifetime.

HISTORY

Beck, a professor of psychiatry at the University of Pennsylvania, had trained and practiced as a psychoanalyst. A scientist at heart, he conducted a series of studies starting in the late 1950s to test psychoanalytic theories

of depression. Much to his surprise, his research demonstrated that depression was not a result of retroflected hostility, unconscious anger toward others that is repressed and turned inward against the self. He found, in fact, that the dreams of depressed patients contained *less* hostility than the dreams of other patients. Their dreams did, however, contain the same themes (such as failure, loss, and rejection) that these patients expressed verbally. Another series of experiments disconfirmed the psychoanalytic theory of a need to suffer. Instead of viewing depression as a result of unconscious drives, wishes, and defenses, Beck proposed that depression was associated with highly negative, distorted, and dysfunctional beliefs and thoughts, the result of biased information processing. When he began to tailor treatment toward helping his patients identify and evaluate their biased thinking and change their behavior, they began to rapidly improve.

In his landmark monograph on depression, Beck (1967) detailed the cognitive triad of depression: negatively biased cognitions about oneself, one's world (including other people), and the future. He described the role of schemas, hypothesized mental structures that organize information, and he outlined a treatment approach that emphasized testing the accuracy of patients' cognitions. Ten years later, the first randomized controlled trial of cognitive therapy for depression was published. Patients in the cognitive therapy group responded slightly better to treatment than patients in the psychopharmacological group. In addition, a later study showed that the cognitive therapy group was found to have less relapse after treatment.

Throughout the decades, Beck and colleagues studied other disorders. They used clinical material to identify and specify key cognitions and behaviors characteristic of each of a number of psychiatric problems. They used existing rating scales or constructed and validated new instruments to measure relevant clinical variables, adapted the general cognitive model and treatment interventions for each disorder, and created treatment manuals for randomized controlled trials.

In developing cognitive therapy, Beck was influenced by Greek Stoic philosophers and theorists contemporary to the time, such as Adler, Alexander, Sullivan and Horney, as well as Kelly, Lazarus, and Arnold. Behaviorists and cognitive behaviorists such as Ellis, Meichenbaum, Mahoney, Lewinsohn, and Bandura were also influential in the development of his theories and treatment.

THEORETICAL ISSUES

Kevin is a 35-year-old male who has been married for 3 years. He presented for treatment because of depression (major depressive disorder, severe), which had developed almost a year before, after he was fired from his job in construction (for making a costly mistake while renovating a commercial building). At the beginning of therapy, it was apparent that he had many highly negative, inaccurate or dysfunctional thoughts and maladaptive behaviors,

especially significant avoidance. His depressive thinking and inactivity contributed to the onset and maintenance of his depressive episode and became major targets in treatment.

Automatic Thoughts

Automatic thoughts arise in the mind spontaneously, without conscious reflection. Most patients are generally unaware of their automatic thoughts before treatment. They are much more likely to notice their emotional, behavioral, or physiological reactions. They tend to believe that situational triggers directly cause their reactions.

When Kevin entered treatment, he believed his low mood was a result of his not working ("I don't have a job; that's why I'm so down.") After psychoeducation, he grasped that what he thought *about the situation ("It's hopeless. I'll never get a job") was more closely associated with his emotions (sadness) and behavior (avoidance of actively seeking a job) than the situation itself.*

The situations, or stimuli, that trigger automatic thoughts may be external (e.g., making a mistake in paying the bills) or internal.

Kevin interpreted his internal experiences in a highly negative way. He had thoughts about his thoughts. When the clinician helped him evaluate some of his cognitions, for example, his automatic thought was, "I can't even think right." He also had thoughts about his emotions ("Feeling depressed shows how weak I am"), about his behavior ("I can't believe I stayed in bed so late,") and about his physiological reactions ("I shouldn't be this tired. What's wrong with me?")

Even when patients are aware of their automatic thoughts, they tend to accept them uncritically, as "truths."

Kevin has hundreds of automatic thoughts throughout the day, such as, "There's no use in getting out of bed." "I'm letting Maria (his wife) down in every way." "I can't even write checks to pay the bills without screwing the whole thing up." It is as if Kevin is looking at his experience through the darkest lenses possible. Much of the time, his perceptions are unrealistically negatively biased—though he believes he is seeing reality clearly.

All-or-nothing thinking *(also called black-and-white, polarized, or dichotomous thinking)*: You view a situation in only two categories instead of on a continuum. Example: "If I'm not a total success, I'm a failure."

Catastrophizing *(also called fortune-telling)*: You predict the future negatively without considering other, more likely outcomes. Example: "I'll be so upset, I won't be able to function at all."

Disqualifying or discounting the positive: You unreasonably tell yourself that positive experiences, deeds, or qualities do not count. Example: "I did that project well, but that doesn't mean I'm competent; I just got lucky."

Emotional reasoning: You think something must be true because you "feel" (actually believe) it so strongly, ignoring or discounting evidence to the contrary. Example: "I know I do a lot of things okay at work, but I still feel like I'm a failure."

Labeling: You put a fixed, global label on yourself or others without considering that the evidence might more reasonably lead to a less disastrous conclusion. Example: "I'm a loser. He's no good."

Magnification/minimization: When you evaluate yourself, another person, or a situation, you unreasonably magnify the negative and/or minimize the positive. Example: "Getting a mediocre evaluation proves how inadequate I am. Getting high marks doesn't mean I'm smart."

Mental filter *(also called selective abstraction)*: You pay undue attention to one negative detail instead of seeing the whole picture. Example: "Because I got one low rating on my evaluation [which also contained several high ratings] it means I'm doing a lousy job."

Mind reading: You believe you know what others are thinking, failing to consider other, more likely possibilities. Example: "He's thinking that I don't know the first thing about this project."

Overgeneralization: You make a sweeping negative conclusion that goes far beyond the current situation. Example: "[Because I felt uncomfortable at the meeting] I don't have what it takes to make friends."

Personalization: You believe others are behaving negatively because of you, without considering more plausible explanations for their behavior. Example: "The repairman was curt to me because I did something wrong."

"Should" and "must" statements *(also called imperatives)*: You have a precise, fixed idea of how you or others should behave, and you overestimate how bad it is that these expectations are not met.
Example: "It's terrible that I made a mistake. I should always do my best."

Tunnel vision: You only see the negative aspects of a situation.
Example: "My son's teacher can't do anything right. He's critical and insensitive and lousy at teaching."

FIGURE 33.7–2. Cognitive distortions. (From Beck JS. *Cognitive Behavior Therapy: Basics and Beyond*. 2nd ed. 2011. Reprinted with permission of Guilford Press.)

While everyone's thinking is characterized from time to time by cognitive distortions, psychiatric patients tend to commit these thinking errors much more frequently. Figure 33.7–2 describes typical distortions.

Kevin engaged in considerable fortune-telling ("I'll never find a job), mindreading ("My wife thinks I'm a failure"), and emotional reasoning ("Since I feel like such a failure, I must be a failure").

Most automatic thoughts are verbal in nature but some are in the form of mental images.

As Kevin was thinking about looking for job opportunities on the Internet, he pictured telling a potential interviewer how long he had been out of work and he imagined the interviewer looking scornfully at him.

Core Beliefs

Individuals display characteristic, recurrent themes in their thoughts which are associated with a deeper level of cognition, their core beliefs, representing the content of cognitive schemas. Core beliefs (Table 33.7–1) are one's most central ideas about the self, other people, and the world. These beliefs are generally identified by asking patients for the meaning of their thoughts.

Core beliefs develop during childhood as one's genetic predisposition toward certain personality traits and styles of thinking affect the way they interpret their experience. People without significant psychological difficulties tend to view themselves throughout their lives as reasonably effective, able to protect themselves, likeable, and worthwhile. If they develop an acute disorder such as depression, however, negative schemas may become activated and dysfunctional beliefs, which may have been latent, become manifest. Patients begin to view themselves, their worlds, and/or other people through a negative lens, interpreting their experience in a highly negative way, and may then overgeneralize the meaning of a situation. They may see themselves as ineffective, vulnerable, inferior, unlikeable, or worthless. They may view others as potentially hurtful, rejecting, or superior. They may see their world as hostile or dangerous. During times of psychological distress, these individuals tend to regard their negative core beliefs as fixed, absolute truths, which make these deeper level ideas less amenable to change than automatic thoughts.

Table 33.7–1.
Categories of Core Beliefs

Helpless Core Beliefs	Unlovable Core Beliefs	Worthless Core Beliefs
I am helpless.	I am unlovable.	I am worthless.
I am ineffective.	I am unlikable.	I am unacceptable.
I am incompetent.	I am unattractive.	I am bad.
I am weak.	I am uncared for.	I am a waste.
I am a victim.	I am bound to be rejected.	I am dangerous.
I am trapped.	I am bound to be abandoned.	I am toxic.
I am needy.	I am bound to be alone.	I am evil.

Adapted from Beck JS. *Cognitive Therapy for Challenging Problems: What to Do When the Basics Don't Work.* 2005. Reprinted with permission of Guilford Press.

Throughout his school years, Kevin performed at a level slightly below average but more or less at the level of most of his friends. He did not value getting good grades and was more interested in developing his above average athletic skills. He felt competent enough academically and saw himself as a success athletically. His parents, however, had unrealistic expectations of him and criticized him whenever he did not perform as well as they desired (which was much of the time). Accordingly, from time to time, especially when his parents were harsh, he viewed himself as a failure. Fortunately, these episodes were generally fleeting and his view of himself as basically adequate and competent predominated. It was not until Kevin was fired from his construction job and could not quickly obtain another job that he began to view himself as a failure. The dominance of this negative core belief

contributed to the onset of his depression.

Coping Strategies

The activation of a negative core belief typically leads patients to engage in compensatory behaviors as an attempt to cope with their distress.

As Kevin began to view himself as a failure, he started to avoid tasks that he assessed were difficult. He gave up job hunting after a few weeks and spent a considerable amount of time sleeping, watching television, or playing video games. He began to withdraw from his wife, family, neighbors, and friends, believing that they all saw him as a failure, too.

Intermediate Beliefs

Intermediate beliefs are ideas about oneself, others, or the world that are at a deeper level than automatic thoughts but are more malleable than core beliefs. They consist of rules, attitudes, and assumptions, and can be a link between patients' deeper core beliefs and their coping strategies. Patients generally believe that if they engage in certain behavioral strategies, their core beliefs may not be evident; however, if they do not engage in this kind of behavior, their core beliefs will likely be proven true.

Once Kevin became depressed, several key assumptions became activated. "If I take on challenges, I'll fail but if I avoid difficult tasks, I'll be okay (at least for the moment)." He also held the attitudes "It's terrible to fail" and "Failing means I'm a failure as a person."

Schemas and Modes

Schemas are responsible for processing, storing, and retrieving different types of information, including perceptions, goals, expectations, memories, fantasies, and prior learning. They provide meaning and, when activated, affect the way individuals process their experiences. Their function is to aid the individual in selecting the most relevant information from the practically infinite amount of information in the environment. Schemas may be cognitive, behavioral, motivational, or affective. They range from adaptive to dysfunctional and vary in complexity, strength, flexibility, permeability, level of abstraction, and valence.

When Kevin is depressed and his depressive schemas are activated, he is likely to process information in a distorted manner that confirms his core beliefs. For example, he views his difficulty getting out of bed in the morning as a sign that he is a failure, even after receiving psychoeducation about the effect depression has on patients' thinking, motivation, and behavior. He fails to register the tasks he completes at home or the ways he is helpful to his wife even in the face of severe depression. He discounts his wife's positive comments, believing that she is "just trying to be nice."

The diathesis–stress model posits that schemas are triggered when individuals experience events relevant to their underlying vulnerability.

Kevin had had negative experiences from time to time throughout adulthood, such as the death of his best friend, arguments with coworkers, and rejection from his neighbors. These incidents, however, did not activate his depressive schemas as they did not match his underlying vulnerability toward interpreting a deficit in performance as failure. His depressive mode was activated only after he was fired from his job for "negligence and incompetence" and was unable to secure a new job. These events matched stimuli from his "failure" schema.

Modes are a complex organization of schemas related to individuals' self-evaluations, rules, memories, and expectancies. Primal modes are related to survival; constructive modes allow the individual to increase resources. Modes help explain the complex of symptoms that patients experience and their global, systematic bias. Each psychiatric disorder is characterized by a particular mode. One constellation of schemas, for example, characterizes a depressive mode, while a different constellation is involved in an anxious mode.

The Generic Cognitive Model

Beck recently expanded upon the cognitive model. The "generic cognitive model" describes how cognitive schemas contribute to the development and maintenance of psychopathology. Neurophysiological hyperactivity is genetically influenced and may, in turn, negatively bias individuals' attention and memories, resulting in dysfunctional cognitive schemas. Beck also described the nature of individuals' functioning along a continuum from adaptive to maladaptive, the process by which schema are activated, the nature of dual processing (through the automatic and then reflective processing systems), and protoschemas (which quickly but crudely detect and interpret critical stimuli), along with an expanded theory of modes and common processes (such as rumination, worry, attentional fixation) across psychopathological states.

Cognitive Conceptualization Diagram

The cognitive conceptualization diagram (Fig. 33.7–3) provides a map of an individual's psychopathology and helps guide treatment. The diagram describes events that contributed to the development of psychopathology (relevant life history) and helps explain the establishment and strengthening of core beliefs. It depicts the relationship among core beliefs, intermediate beliefs, coping or compensatory strategies, and current automatic thoughts. Data are collected verbatim from patients and clinicians indicate their hypotheses about patients with question marks until their inferences are confirmed by the patients themselves.

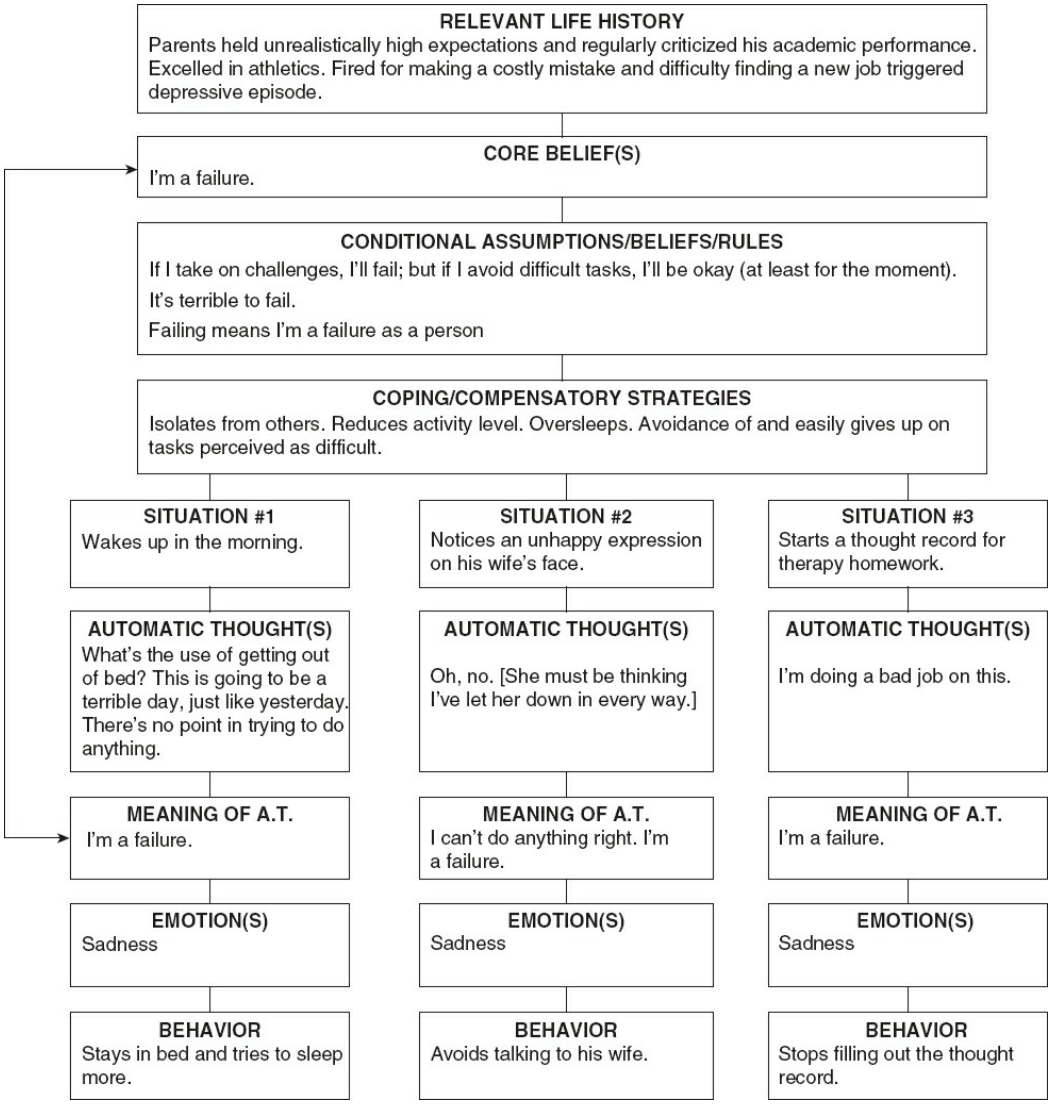

FIGURE 33.7–3. Cognitive conceptualization diagram. (Adapted from Beck JS. *Cognitive Behavior Therapy: Basics and Beyond*. 2nd ed. 2011. Reprinted with permission of Guilford Press.)

The bottom portion of the diagram contains data about an individual's automatic thoughts in specific situations. The diagram is a continually evolving document that clinicians may start to fill in as soon as they begin to see important patterns in patients' thinking and behavior. They refine the diagram after every session as new information emerges that supports, disconfirms, or refines their hypotheses. Clinicians typically do not share this diagram with patients as it is too complex. But they continually verbally summarize their conceptualization to patients and inquire whether their understanding is accurate. They usually draw a simplified version of the diagram during sessions to enhance patients' insight into how their thinking influences their reactions, how their beliefs influence their thinking, how their life experiences influenced their core beliefs, and/or how they developed dysfunctional patterns of behavior.

Applications across Specific Disorders

According to the cognitive model, psychiatric disorders consist of an interaction among triggers, cognitions, behavior, emotions, and common psychological processes. Each psychiatric disorder is characterized by a specific set of beliefs and behavioral strategies. The following is an overview of the cognitive theories for several disorders.

Depression

The maladaptive cognitions specific to depression are referred to as the cognitive triad: negative beliefs about oneself, one's life experience (and the world in general), and one's future. In a depressive mode, these negative cognitions affect (and in turn are affected by) individuals' mood, behavior, and physical symptoms. Depressed patients typically feel sad, empty, irritable, anxious and/or guilty; they often predict that the future holds nothing but unending problems and pain; they generally experience decreased energy and motivation; they may have disturbances in sleep, libido, and appetite; they tend to reduce their level of activity and to isolate themselves from others.

When depressed, individuals tend to be excessively self-focused. They are likely to disproportionately allocate attention to their negative thoughts, feelings of sadness, and physiological symptoms. Depressive rumination typically results in an exacerbation of symptoms and continued inactivity. As the depressive mode continues to be active, individuals' sense of hopelessness about the future may increase and lead to increased suicidal ideation when they believe suicide is the only solution to their problems.

Bipolar Disorder

The schemas of bipolar patients are activated when biological and psychological processes interact with external and internal triggers, leading to depression, mania, or mixed mood. In a depressed state, these patients display the same kinds of cognitive biases and dysfunctional beliefs as patients with unipolar depression, especially negative cognitions about the self, the world and other people, and the future. In a manic state, patients with bipolar disorder tend to view themselves in an inflated, expansive way. Hyperpositive core beliefs (I am powerful; I am invulnerable; I am great) predominate. Their behavior changes, consonant with the changes in their thinking. They become impulsive, minimize potential threats, and take risks. Typical conditional assumptions include: "If something seems like a fantastic idea, it must be"; "I should follow my impulses and take action immediately"; "If I'm feeling great, I don't have to take meds or be careful about my daily habits." In a euthymic mood, patients with bipolar disorder may think, "I have to drive myself hard to make up for how little I

accomplished when I was depressed." This belief may lead them to overextend and overstimulate themselves; they may then sleep poorly and ultimately become hypomanic.

Mansell et al. (2007) emphasize the role of interpretations of internal events and beliefs about affect regulation as factors that contribute to the triggering of depressive, manic, or mixed episodes. Basco and Rush (2005) detail cognitions that lead to medication nonadherence.

Generalized Anxiety Disorder

GAD is characterized by frequent worry about a variety of events or activities and by futile attempts to control the worry process. Individuals diagnosed with GAD tend to overestimate risk and underestimate their ability to cope with events. They may have developed core beliefs such as "I'm helpless and vulnerable," and "The world is a dangerous place." As a result, their automatic thoughts tend to center around potential future catastrophes. Their attention becomes fixed on worry thoughts and signs of threat, danger, and/or uncertainty.

Wells (2009) proposed a cognitive model of GAD with beliefs about worry as the central feature. Individuals with GAD hold both positive and negative beliefs about the worry process. Positive beliefs are that worry is helpful, prepares one for danger, and shows that one is being conscientious. These beliefs are initially activated when an individual interprets an event or situation as threatening. For instance, a patient notices a pain in her abdomen and thinks, "It might be cancer." She feels highly anxious and believes that worrying may help her ward off catastrophic consequences. Because most worries do not come true, the belief that worry is helpful and can prevent catastrophe is maintained.

While worry may be initiated by positive beliefs, negative beliefs may follow. Patients may see worry as harmful and uncontrollable; they may fear worry will lead them to go crazy or lose control. They may then try to suppress their worries, which only serve to increase the frequency of their anxious thoughts, reinforcing the belief that they cannot control their worry. Behaviorally, they engage in "safety behaviors," coping strategies to ward off anxiety, such as avoidance of situations in which they feel vulnerable; reassurance seeking (from other people and the Internet), and excessive checking. These strategies only serve to maintain a preoccupation with worry and prevent individuals from being exposed to situations that would have demonstrated to them that worry is harmless and at least partially controllable.

Panic Disorder

Panic disorder consists of multiple, unexpected panic attacks, which are episodes of intense fear with accompanying symptoms, such as heart

palpitations, sweating, dizziness, and shaking. Cognitive models view catastrophic misinterpretations of a specific physiological or mental symptom (or small set of symptoms) as crucial to the development and maintenance of the disorder. Individuals notice a change in their body (or mind) which might be due to activity, exercise, a strong positive or negative emotion, or a benign physiological symptom. They then make a negative attribution about this change ("There's something wrong with me."). This cognition leads to an anxious mood and physiological sensations of anxiety. Their attention becomes fixed on their physical or mental sensations and the sensations intensify. At some point, they make a catastrophic misinterpretation, believing that the intensifying sensations will lead to an imminent catastrophe ("[This rapid heartbeat and tightness in my chest mean] I'm having a heart attack." "[This feeling of unreality means] I'm going crazy." "[This pain in my head] means I'm having a stroke") and they experience a full-blown panic attack.

After a panic attack, individuals may develop a fear of the triggering sensation(s) and/or the situation in which the panic attacked occurred. They become hypervigilant for the feared sensations to try to prevent future panic attacks. This strategy often results in a greater awareness of their sensation(s) and leads to an increased likelihood of additional panic attacks. Avoidance of external situations that have triggered prior panic attacks may also occur. For instance, an individual had his first panic attack when he noticed his heart racing while in a crowded theatre; following that experience, he decided to avoid going to the movies and events in large auditoriums.

Patients with panic disorder employ safety behaviors to avoid panic attacks; these strategies also maintain the disorder. Examples of these coping strategies include carrying anxiety medication at all times, only going to feared locations with a companion, and making sure an easy exit is possible if panic arises. For example, the individual mentioned above eventually decided to return to the movie theater, but made sure that he knew where the exit was and sat in an aisle seat so he would be able to easily leave if panic symptoms began. Panic disorder is maintained because the individual does not learn that the symptoms, while uncomfortable, are actually innocuous and do not need to be avoided.

Obsessive-Compulsive Disorder

OCD consists of obsessions, compulsive rituals, or both. Obsessions are intrusive thoughts, images, or impulses that individuals attempt to suppress, neutralize, or ignore because they lead to elevated anxiety. Compulsions are repetitive behaviors or mental events that individuals feel compelled to perform in response to obsessions to reduce distress or prevent a feared event from occurring. The vast majority of the nonclinical population report experiencing intrusive thoughts; cognitive models help

explain why only some individuals go on to develop OCD.

Cognitive models propose that intrusive thoughts, images, or urges develop into obsessions when they are given excessive negative importance. For instance, a woman has an image of dropping her baby and believes, "(Having images of dropping my baby means) I'm a bad person and I will lose control and do it." These beliefs are referred to as fusion beliefs. The aforementioned intrusion is a thought-action fusion as the individual believes thinking about performing an action can cause the action to occur. Other fusion beliefs are thought-event fusion (thinking about an event means it will or has already happened) and thought-object fusion (the belief that thoughts and feelings can be transferred onto objects).

The content of intrusions usually violates patients' deeply held values and may be related to earlier experiences. For example, the woman who imagined herself dropping her baby had taken care of a younger sibling when she grew up. Her mother had frequently told her to be extremely careful when handling her young sibling. As individuals with OCD believe their intrusive thoughts, images, and urges will cause threatening events, they may take excessive responsibility for any potential negative situation. Focus of attention is allocated toward intrusive thoughts as they are assigned excessive importance.

When beliefs about intrusions are activated, individuals experience intense negative emotion and their beliefs about rituals or compulsions become activated as well. They may believe that engaging in a compulsion will both reduce the likelihood of the anticipated threat and lessen psychological distress. The urge to carry out a compulsion is also related to an inflated sense of responsibility for the potential occurrence of the feared outcome. Overt compulsions may include ordering, checking, washing, touching objects, and avoidance, and covert compulsions may consist of subvocalizations of phrases, thinking of neutralizing images, counting, and thought suppression. The woman mentioned above engaged in the covert ritual of praying for the safety of her child and overtly refused to hold the child until the intrusive thought was absent. The compensatory, compulsive rituals alleviate distress in the short term but act to maintain the disorder because the use of the rituals is negatively reinforced through distress reduction and the absence of the feared outcome. Accordingly, individuals do not learn that their beliefs about intrusive thoughts, images, and urges are not accurate.

Social Anxiety Disorder

Social anxiety disorder (SAD) involves a fear of social situations in which an individual is exposed to potential scrutiny from others. As a result, such situations are either avoided or endured with a disproportionate amount of anxiety. Individuals with SAD tend to hold beliefs that they are inadequate, defective, and/or undesirable, and that other people are inherently critical

and likely to evaluate them negatively and reject them. Cognitive models of the disorder have identified factors before, during, and after a feared social event that act to maintain these maladaptive beliefs.

Before a social event, patients with SAD engage in worry. They frequently have memories of interpersonal situations that they perceived turned out badly, and they make highly negative predictions about what will happen in an upcoming situation. During social interaction, their attention is highly self-focused. They believe others will form negative judgments of them if they do not perform at a high level or if their anxiety is noticeable. Two prominent safety behaviors are to monitor themselves excessively and to attempt to control their physiological reactions (e.g., by clasping their hands to avoid trembling). Self-monitoring leads to an inordinate focus on their symptoms (e.g., heart racing, shakiness), and rather than being able to control their symptoms, the symptoms increase and they feel increasingly more anxious. They commonly use other safety behaviors as well, to protect themselves from embarrassment or failure.

For example, a man with SAD is thinking about a job interview that is scheduled for the following week. He thinks about previous interviews in which he remembers (correctly or not) stumbling over his words. He then worries that he will perform even worse during the upcoming interview. This dysfunctional thought process elevates his anxiety and triggers compensatory behaviors such as memorizing what he intends to say. During the interview, he avoids making eye contact, fearing the interviewer will be looking at him disapprovingly and answers questions minimally.

After an event, these patients engage in postevent processing, mentally replaying their biased memory of interactions over and over, and tend to evaluate their performance in an unrealistically negative light. Patients with SAD often fail to recognize positive reactions from others or misinterpret others' reactions as negative. When their performance actually exceeds their expectations, they tend to attribute the outcomes to their use of safety behaviors instead of their innate abilities, which acts to maintain the dysfunctional beliefs.

Eating Disorders and Obesity

Anorexia nervosa and bulimia nervosa have been extensively studied and conceptualized according to the cognitive model. A cognitive model has been developed for each disorder; however, Fairburn et al. (2003) proposed that the disorders share the same underlying psychopathology and developed a transdiagnostic model. This model postulates that individuals diagnosed with eating disorders evaluate their self-worth disproportionately both in terms of their eating habits, shape or weight and in their ability to control these factors.

For instance, an individual may accurately perceive her success at work and in relationships, yet view herself as "completely worthless" as she

focuses excessively on her body image and her weight. Interestingly, individuals diagnosed with bulimia nervosa tend to have average body weight and those with anorexia nervosa are, by definition, underweight, sometimes dangerously so. Their preoccupation with how they see themselves may have been influenced by early life factors, such as parents or peers who were overly critical of their weight or shape.

The difference between the two disorders is hypothesized to be the result of the compensatory behavior used to cope with negative self-evaluations or other triggers related to eating, body image, or sense of control. Anorexic patients predominantly cope by restricting food intake and/or excessive exercise, while individuals diagnosed with bulimia use a compensatory strategy such as purging when they perceive they have overeaten.

Eating disorders are maintained through several processes. For instance, individuals feel better about themselves for successfully restricting eating and they reduce their distress through dysfunctional compensation methods after a binge eating episode, which makes those behaviors likely to continue. Additionally, these individuals hold unrealistically high standards about their weight, shape, and/or self-control. When they do not meet these standards, they may respond with further self-criticism and continued attempts to achieve "success" through food restriction or other compensatory strategies.

People who are overweight or obese and who have struggled unsuccessfully to lose weight or maintain a weight loss display characteristic ways of thinking. They have a host of dysfunctional cognitions, such as "It's okay to eat this (food I hadn't planned) because … I'm happy/I'm sad/it's free/I'll make up for it later/everyone else is eating it/ no one is watching." Having eaten this food, they think, "I cheated. I may as well eat whatever I want now and start again tomorrow." They may then consume hundreds or even more additional calories. When they get on the scale the next morning and see that they have gained weight, core beliefs of helplessness may become activated and they think, "Dieting is too hard. I just can't do it. I may as well just give up." Dieting is often the only area in their lives in which they compound an initial mistake with many more.

These individuals tend to label any desire or impulse to eat as "hunger." They fail to realize that they often want to eat when they are actually thirsty, tired, bored, or procrastinating; when they are feeling discouraged or deprived; when they see others eating or drinking freely; when they are experiencing a craving or are feeling upset for any reason. When distressed, their beliefs about emotional eating may become activated: "If I'm upset, I deserve to eat," or "If I'm upset, the only way I can calm down is through eating." Having overeaten, beliefs about being weak and out of control predispose them to continued overeating.

Substance Use Disorders

These disorders involve a problematic pattern of substance use that leads to impairment in functioning and/or significant distress. Individuals may begin to use substances for various reasons: obtaining pleasure, feeling "high," improving the experience of an activity, or sharing the experience with companions. Once substance use has started, several factors may contribute to becoming dependent on the substance, including beliefs such as, "If I use, I can endure the stress and frustration of life." "I can escape or avoid distress." These individuals also tend to have negative core beliefs. The activation of these beliefs leads to dysphoria, the individual focuses on the negative emotion and then has a permission-giving cognition which leads to the compensatory behavior of substance use to alleviate or cope with the distress. They may also use because of their belief, "I won't enjoy myself if I don't use."

As substance use becomes more frequent, external and internal triggers become associated with use. External triggers include certain people, locations, or situations, while internal factors include unpleasant emotional states such as anxiety, sadness, and boredom. These triggers also become associated with the experience of cravings and urges. Cravings are the desire for gratification or relief, and urges compel action to obtain relief or gratification. Beliefs about cravings and urges develop, such as "I need to give into urges" and "Cravings are uncontrollable."

As substance use increases, individuals are likely to encounter financial, social, and/or medical problems, about which they have negative cognitions, leading to negative mood states and subsequent substance use. Their focus of attention is excessively inward, as they monitor themselves for dysphoria. When they hold conflicting beliefs about substance use related to their adverse circumstances, they generally minimize potential consequences or develop additional permission-giving beliefs. The cycle of substance use is maintained by the dysfunctional beliefs about the self, substance use, cravings and urges, permission to use, and negative mood states.

Schizophrenia

Schizophrenia is characterized by positive symptoms (delusions, hallucinations, disorganized speech, and behavior) and negative symptoms (avolition, social withdrawal, diminished emotional expression). The development of schizophrenia is believed to be influenced by predisposing genetic, biological, and cognitive factors, which interact with stressors to bring about the disorder.

The negative symptoms of schizophrenia are mediated, at least in part, by dysfunctional beliefs: "I'm helpless," "I'm unlikeable," "There's no use in engaging with others because they'll just reject me," "If I try to be

productive, I'll just fail." Individuals with this disorder tend to isolate themselves and avoid tasks they view as challenging, even some fundamental activities of daily living.

Delusions are believed to develop, in part, due to information processing biases, cognitive errors, and impaired reality testing. Individuals experiencing delusions tend to have a self-referential bias, perceiving themselves to be the center and central focus of events, while also interpreting their subjective experiences as having an external causation. For example, an individual sees a commercial for cereal on television and thinks that it was aired solely for him because he has eaten that specific cereal in the past. He then stays up most of the night trying to figure out why the commercial was aired, feels tired, and attributes being tired to someone putting drugs in his food.

Cognitive errors common in schizophrenia include selective abstraction, extreme judgments, overgeneralization, selective focus, inferential bias (perceiving connections between irrelevant events), catastrophizing, and inadequate data gathering. Impaired reality testing, or the inability to distance oneself from beliefs and interpretations to consider whether they are accurate, is believed to be related to limited cognitive resources in these individuals. Individuals may not fully accept their delusion as true initially, but over time they selectively attend to information they believe confirms the delusion and do not consider evidence that contradicts their belief. These maladaptive cognitive processes, in addition to the use of safety behaviors, strengthen and maintain the delusion. For example, the person mentioned above engaged in the safety behaviors of covering his television with a blanket and closely monitoring his food to make sure others had not tampered with it.

Hallucinations are experienced by the vast majority of individuals diagnosed with schizophrenia. A cognitive model of hallucinations describes factors that are precursors to hallucinations, lead individuals to fixate on them, and then maintain their occurrence. One predisposing factor is a low threshold for internal auditory and visual imagining that seems real or almost real. This low threshold interacts with activated cognitive schemas that produce cognitions related to the schema and not to reality. For instance, an individual with a "failure" schema may have an intrusive, automatic thought that is perceived as an auditory hallucination saying, "You can't do anything right." The hallucination is perceived as coming from an external source due to externalization bias, cognitive errors, and impaired reality testing. Hallucinations are maintained by dysfunctional beliefs about the nature of the hallucinations, the supposed communicators, and the nature of the individual's relationship to the voices. For instance, a person may develop an idea that the voice is present to hurt him; this cognition then becomes integrated into his delusional beliefs and consequently strengthens both positive symptoms. Safety

behaviors are also performed to minimize the presence or impact of hallucinations, such as avoiding certain places or wearing earphones to muffle auditory hallucinations.

Beliefs about positive symptoms, in addition to the symptoms themselves, influence the development and maintenance of negative symptoms of schizophrenia. For instance, an individual believes that the Mafia is out to get her, which leads her to engage in the safety behavior (and negative symptom) of rarely leaving her residence. Moreover, as positive symptoms are often triggered in social situations, individuals may develop an aversion to socializing and choose to reduce social contact. Overall activity level may also decrease as individuals become excessively focused on delusions and/or hallucinations and less focused on engaging in valued activities. Additional factors that influence the development and maintenance of negative symptoms are negative beliefs about one's performance, low expectations for obtaining pleasure and success, and the perception of having limited resources for engaging in activities (e.g., too little energy).

While some clinicians focus on diminishing positive symptoms, others who engage in recovery-oriented cognitive therapy emphasize resilience, rather than focusing on symptoms. They elicit individuals' values and goals, create experiences that lead to a disconfirmation of the individuals' beliefs, and help them work toward achieving what they want from life, which usually includes increasing behavioral activation and engagement with others, and acting productively toward the fulfillment of their goals. Symptoms are addressed when they become obstacles to goal achievement and clinicians engage the patient in collaborative problem solving about them.

Personality Disorders

Personality disorders are characterized by enduring personality traits (cognitive, affective, motivational, and behavioral) that are inflexible, maladaptive, and cause significant distress and/or impairment in functioning. As with the previously discussed disorders, personality disorders develop from the interplay of genetic factors and early environmental experiences. The schemas present in individuals with personality disorders resemble those present in other syndromes, but they tend to be more rigid and remain activated on a more or less continuous basis. Dysfunctional schemas become a part of normal, everyday processing of information.

Each personality disorder is characterized by a specific set of beliefs, assumptions, and coping strategies. An individual with dependent personality disorder, for example, may have a "helpless" schema and compensate by attaching to and depending on others. An individual with avoidant personality disorder usually has core beliefs of vulnerability and

unlovability and avoids not only social interactions but also thinking about situations that could lead to emotional distress. When individuals have comorbid personality disorders, they often hold multiple sets of dysfunctional beliefs that cross diagnostic lines.

TECHNIQUES

The goal of cognitive therapy is to facilitate remission and prevent or reduce relapse by teaching patients skills they can use for their lifetime, within a highly supportive, collaborative relationship. Both patients and clinicians are active participants, jointly solving current problems, evaluating patients' cognitions in specific situations, and working toward needed behavioral change. As a result, patients' symptoms are reduced, and their mood and day-to-day functioning improves. Most patients achieve remission without exploration of childhood issues. When longstanding highly dysfunctional beliefs acquired in childhood interfere with "standard" treatment, as may be the case when working with patients with personality disorders, clinicians may help patients restructure the meaning of early experiences and apply what they have learned to help them overcome current difficulties.

This section describes how clinicians use cognitive therapy techniques in a standard 50-minute appointment with an individual outpatient. Many of the techniques, however, can be used in brief sessions (e.g., Lusk & Melnyk, 2011, Wright et al., 2010).

Clinicians may adapt techniques from almost any psychotherapeutic modality, using them in the context of a cognitive conceptualization of the patient. In recent years, cognitive therapists have drawn on techniques from the range of cognitive behavior therapies, including DBT, exposure therapy, cognitive processing therapy, and acceptance and commitment therapy. Techniques from positive psychology, psychodynamic psychotherapy, mindfulness-based therapies, and interpersonal therapy are used with some patients.

As part of the initial assessment, clinicians collect data to diagnose patients' difficulties. Treatment is guided by the cognitive formulation of patients' disorders, as described above. Clinicians begin to formulate a cognitive conceptualization of the individual patient and his/her difficulties; they refine this conceptualization at each session, as they receive additional data. They hold their hypotheses as conditional, subject to confirmation by the patient.

Clinicians use this ongoing conceptualization to plan treatment across sessions and within sessions. An accurate conceptualization allows the clinician to appropriately select techniques such as those described below. Clinicians also take into consideration the patient's goals, stage of treatment, degree of distress, and personal characteristics and preferences. They consider the strength of the therapeutic alliance, the nature of the

problem under discussion, the amount of time remaining in the session, and other factors. They suggest a technique, provide a rationale, and elicit the patient's agreement. "Is it okay if we take a look at this idea you have that you're a failure because your boss criticized you? It would be good to find out whether this thought is 100 percent true or 0 percent true or some place in the middle." Clinicians check on how the patient is feeling before and after an intervention, to assess the need for additional strategies. Two questions also help guide the selection of techniques: How can I help the patient feel better by the end of the session? How can I help him/her have a better week?

The techniques described below are somewhat artificially grouped as many techniques fall into more than one category. For example, most behavioral techniques contain both cognitive and behavioral components and vice versa.

Therapeutic Relationship Techniques

Clinicians use good counseling skills: warmth, support, accurate understanding and reflection, genuineness, empathy, realistic optimism, and compassion. They use positive reinforcement to encourage change. They show interest in their patients and may deliberately be more conversational at times, for example, when inquiring about situations in which a depressed patient felt better during the week. They also make judicious use of self-disclosure. Clinicians are highly collaborative and seek patients' feedback. During sessions, they are attuned to patients' affect and when they notice a shift, ask patients how they are feeling and what they are thinking.

Clinicians provide positive feedback when patients express negative thoughts about them or about the process of therapy. ("It's good you told me that.") Then clinicians may apologize (if they believe they have made a mistake—such as offering an inappropriate suggestion), do problem solving, and/or examine the patient's distorted conclusion. ("You're treating me like everyone else. You don't care about me.") Clinicians also routinely ask for feedback at the end of sessions: "What did you think of the session?...Was there anything that bothered you or anything you thought I didn't understand? ... Is there anything you want to do differently next time?"

Some patients, especially those with personality disorders, bring dysfunctional beliefs about other people and relationships to treatment and may react to clinicians in dysfunctional ways. Clinicians work through therapeutic relationship problems, helping patients evaluate their beliefs about the clinician. They then help clients transfer what they have learned to other relationships. It is essential for clinicians to vary their interpersonal style according to the preferences and sensitivities of patients. Some patients prefer a close, warm relationship while others react

poorly to direct expressions of empathy and prefer a more clinical approach.

Finally, research has demonstrated that a reduction in symptoms leads to a better therapeutic alliance. Therefore, it is important to spend sufficient time to develop a good working relationship with patients but clinicians need to keep as their central focus the facilitation of cognitive, emotional, and behavioral change so patients are equipped to deal more effectively with their difficulties outside of the therapy session.

Structuring Techniques

Cognitive therapy sessions tend to follow a standard structure, as it allows patients to gain the most from the appointment (Fig. 33.7–4). The structure is implemented flexibly, though, and sometimes has to be modified, especially initially, if patients have a negative reaction. At the beginning of sessions, clinicians do a mood check, ask for an update (of positive and negative events) since the previous session, review patients' "action plans" (therapy homework), and set an agenda ("What problem or problems do you most want my help in solving today?"). These problems may have arisen in the recent past or they may be difficulties patients expect may arise in the near future. Clinicians may contribute items to the agenda, too. Then they collaboratively prioritize agenda items. If, later in the session, patients bring up problems that were not on this initial agenda, the clinician notes the change of topic and they make a collaborative decision about whether to return to the initial problem under discussion or to continue discussing the newly identified topic or difficulty.

In the middle of sessions, clinicians help patients prioritize the agenda items and ask them to recall a specific recent instance in which they felt the most distressed about the first problem under discussion. They seek additional data, then summarize their perception of the patient's experience in the form of the cognitive model, insuring that they correctly understood the patient. ("Okay, so you were at work yesterday morning, and you thought, "I hate this job. But there's nothing I can do. I'm stuck. I'd never be able to find another job," and you felt very sad and hopeless. Then you just sat at your desk and stared at the computer and didn't finish your work … Did I get that right?")

Beginning of Session—usually 5–10 minutes
- Mood Check
- Bridge (Reminder of previous session)
- Agenda Setting
- Homework Review

Discussion of Agenda Items/Problems—usually 35–40 minutes
- description of the occurrence of a specific problem
- elicitation and confirmation of elements of cognitive model
- collaborative decision of how to approach problem
- rationale for and introduction of intervention(s)
- assessment of efficacy of intervention(s)
- summary by patient
- collaborative creation of "action plan" in writing.

Feedback to Therapist—usually 1–2 minutes

FIGURE 33.7–4. Session structure.

The clinician might then propose an intervention aimed toward any part of the cognitive model, for example, doing problem-solving about the most onerous aspects of the job (situation level), examining the patient's thoughts (automatic thought level) or eliciting and examining his beliefs, teaching the patient emotional regulation skills (emotional level), demonstrating relaxation or mindfulness techniques (to decrease physiological arousal), or teaching the patient needed skills (behavioral level), for example, how to recognize depressive inactivity and start on the easiest task. After such discussions, clinicians ask patients to summarize what they have learned and clinician or patient write down what the patient wants to remember and do in the coming week, creating an "action plan."

At the end of sessions, clinicians ensure that patients have a good grasp of key points of the session and are highly likely to carry out their action plans. Then they elicit and respond to feedback.

The session structure for certain problems, such as anxiety disorders, tends to be more narrowly focused. Instead of discussions of a variety of wide ranging problems, clinicians tend to set a pointed agenda, focusing on testing patients' key fears and reducing avoidant behaviors, in and out of session, that interfere with the testing of such cognitions. Focusing on key thoughts, beliefs, and/or behaviors (guided by the cognitive conceptualization) generally leads to greater improvement and can generalize to more peripheral problems. When dealing with acutely suicidal patients, clinicians emphasize developing a safety plan, responding to key

cognitions related to hopelessness, teaching patients emotional regulation skills, and doing advance problem solving about situations in which the patient might become suicidal. Recovery-oriented cognitive therapists, working with inpatients with schizophrenia who have been hospitalized for a long period of time, do not initially employ a set structure; instead, they provide experiences to engage the patient.

Goal Setting Techniques

Prior to setting goals, the clinician seeks to identify patients' major values in life. Many patients endorse being productive or achieving at work, being a good family member, having an intimate relationship, expressing their creativity, contributing to their community, working toward social justice, leading a spiritual life. Identifying values prepares patients to set specific goals in keeping with aspects of life that are most important to them. Major reasons to set goals with patients include instilling hope that life can get better and providing a concrete picture of what that would look like.

Goals are the flip side of problems (problem: I'm lonely; patient's goal: to socialize more). The clinician asks, "What are your goals for treatment? How would you like to be different as a result of treatment ... or how would you like your life to be different?" When patients provide a global response, such as "I don't want to be depressed anymore," or "I want to be happier," the clinician helps them become more specific. "If you weren't depressed, what would you be doing differently? How might your life look different?"

A number of techniques can help when patients have difficulty identifying goals. Clinicians can ask them for a detailed account of how they spend a typical day versus how they would like to spend a typical day. They can ask patients to recall what their life was like when they were feeling better; what were they doing then that they aren't doing now? They can also request that patients imagine a typical day at a specific time in the future when their disorder has remitted and they are feeling and functioning better.

When patients set goals for other people (such as "I want my partner to be nicer to me,") the clinician helps them formulate a related goal that they have control over (e.g., "Investigate ways of interacting differently with my partner"). Patients can keep adding to their goal list as time goes on. The clinician's aim is not necessarily to help patients achieve all their goals by the end of treatment, but rather to equip them with the tools they will need to continue improving their lives on their own.

What is the evidence that your thought is true? What is the evidence on the other side?

What is an alternative way of viewing this situation?

What is the worst that could happen, and how could you cope if it did? What's the best that could happen? What's the most realistic outcome of this situation?

What is the effect of believing your automatic thought, and what could be the effect of changing your thinking?

If your [friend or family member] were in this situation and had this automatic thought, what advice would you give him or her?

What should you do?

FIGURE 33.7–5. List of socratic questions. (From Beck JS. *Cognitive Behavior Therapy: Basics and Beyond*. 2nd ed; 2011. Reprinted with permission of Guilford Press.)

After setting goals, clinicians provide psychoeducation. "Most people think that once they feel better, they will start doing these things. But research actually demonstrates the opposite. To feel better, people need to start working toward their goals."

Cognitive Techniques

The clinician elicits patients' thinking when they have experienced a significant affect shift or have engaged in significantly dysfunctional behavior. (For patients with somatization disorder, pain, or some anxiety disorders, the clinician elicits their thinking just before and after a significant physiological change.) The clinician generally asks, "What was just going through your mind?" or "What were you thinking?" As mentioned above, the clinician shares his conceptualization with the patient; then they often collaboratively decide to evaluate the patient's thoughts.

Guided discovery, primarily through Socratic questioning, is the major method for helping patients assess their thinking. Typical questions involve labeling distortions, reality testing, seeking alternative explanations, decatastrophizing, examining the utility of the thought, and asking patients to consider how they would advise another person who had the same thought(s) (Fig. 33.7–5—list of Socratic questions). If the process of evaluation results in a positive affect shift, the clinician makes sure patients' new understandings are recorded, on cards (see Fig. 33.7–6), in a notebook, or on a smart phone so patients can reflect on these new ideas daily.

Automatic Thought:

"I don't have the energy for this."

Response:

Just because I'm tired doesn't mean I can't be active. There have been plenty of times in the past when I was tired, got involved in an activity, and actually felt MORE energetic afterward.

FIGURE 33.7–6. Coping card.

When the process of evaluation does not result in significant improvement in mood, the clinician may collect additional data. Has the cognitive restructuring of the automatic thought led the patient to change his/her thinking at the intellectual level but not at the "emotional" level? Does the patient have other automatic thoughts (verbal or imaginal) that need to be addressed? Was the automatic thought closely tied to a more overgeneralized underlying belief? Is the patient's thinking obsessive—if so, mindfulness techniques may be indicated.

Sometimes a realistic appraisal of patients' thinking indicates that a given thought is largely valid. In this case, the clinician engages them in problem solving, identifying and assessing the meaning they have attached to their thoughts, and/or moving toward acceptance of the situation accompanied by focusing on ways to enrich other parts of their lives.

Behavioral experiments are commonly used to help patients evaluate their predictions. Examples of techniques used in session include panic inductions (so patients can learn that no matter how intense their symptoms become, a feared catastrophe does not happen), reading a paragraph in a book (so depressed patients can learn that they have not lost the ability to concentrate), or rubbing their hands on the bottom of their shoes and refraining from washing (so patients with OCD and a fear of contamination can learn that their obsessions are inaccurate, they can withstand the anxiety, and the anxiety that arises does not increase indefinitely without resolving when they prevent themselves from engaging in a compulsion). Clinicians also help patients devise behavioral experiments to test predictions about situations that arise outside of session.

When patients are ready, clinicians may show them how to evaluate their thinking between sessions. A number of worksheets and apps facilitate this process. See Figure 33.7–7 for one such worksheet.

To identify patients' underlying beliefs (their assumptions and basic beliefs about themselves, their worlds, and others), clinicians ask patients

for the meaning of their thoughts, examine the recurrent themes in patients' thoughts, directly elicit a rule or attitude, review a questionnaire patients have completed, such as the Personality Belief Questionnaire, or provide a list of common core beliefs to see which one(s) resonate. (See Fig. 33.7–5 and Table 33.7–1.)

Clinicians use a variety of techniques to help patients assess and respond to their deeper level cognitions. They use psychoeducation to explain how the patient's information processing system automatically processes data that seems to support their beliefs but discounts or fails to register data that contradicts it. Socratic questioning, examining advantages and disadvantages of maintaining a given belief, formulating a more adaptive, realistic belief, and historical tests of the belief are frequently used techniques. Clinicians also teach patients how to continually identify and reframe evidence that seems to support the old belief and how to identify evidence that is contrary to this belief but supports the new belief (Fig. 33.7–8, Core Belief Worksheet). These techniques help patients change their ideas at an intellectual level. To modify their ideas at a deeper, "gut" level, other techniques may be necessary. Imagery, behavioral experiments, acting "as if" the patient fully accepts the new belief as true, metaphors, and experiential techniques (such as roleplaying or psychodrama to restructure the meaning of specific current or childhood experiences) can facilitate a new understanding at the "emotional" level.

Directions: When you notice your mood getting worse, ask yourself, "**What's going through my mind right now?**" and as soon as possible jot down the thought or mental image in the Automatic Thought Column.

DATE/ TIME	SITUATION	AUTOMATIC THOUGHT(S)	EMOTION(S)	ALTERNATIVE RESPONSE	OUTCOME
	1. What actual event or stream of thoughts, or daydreams, or recollection led to the unpleasant emotion? 2. What (if any) distressing physical sensations did you have?	1. What thought(s) and/or image(s) went through your mind? 2. How much did you believe each one at the time?	1. What emotion(s) (sad, anxious, angry, etc.) did you feel at the time? 2. How intense (0–100%) was the emotion?	1. (optional) What cognitive distortion did you make? (e.g., all-or-nothing thinking, mind-reading, catastrophizing.) 2. Use questions at bottom to compose a response to the automatic thought(s). 3. How much do you believe each response?	1. How much do you now believe each automatic thought? 2. What emotion(s) do you feel now? How intense (0–100%) is the emotion? 3. What will you do (or did you do)?

Questions to help compose an alternative response: (1) What is the evidence that the automatic thought is true? Not true? (2) Is there an alternative explanation? (3) What's the worst that could happen and how could I cope? What's the best that could happen? What's the most realistic outcome? (4) What 's the effect of my believing the automatic thought? What could be the effect of changing my thinking? (5) If _____ (friend's name) was in the situation and had this thought, what would I tell him/her? (6) What should I do about it?

FIGURE 33.7–7. Thought record. (From Beck JS. *Cognitive Behavior Therapy: Basics and Beyond*. 2nd ed. 2011. Reprinted with permission of Guilford Press.)

Imaginal Techniques

Many patients experience automatic thoughts in the form of spontaneous images, although most do not recognize these "pictures in their minds" without prompting from the clinician. The clinician educates patients about images and may then ask them to continue imagining the situation to its conclusion. Through this process, some patients may do an automatic reality check, recognize that an anticipated negative outcome does not occur, and feel better. Other patients may imagine a catastrophic outcome. The clinician can use a variety of techniques, for example, having patients imagine coping in the situation or jumping ahead in time to see themselves (or others) getting on in life. The clinician may take another tack: suggesting to patients that they take control of "the movie" in their head, much as a director would, and change the image to what they would like to happen.

Name: _____ Date: _____

Old Core Belief: _____

How much do you believe the old core belief **right now**: intellectually? (0–100) _ emotionally? (0–100) _____

*What's the most you've believed it **this week**: intellectually? (0–100) _ emotionally? (0–100) _____

*What's the least you've believed it **this week**: intellectually? (0–100) _ emotionally? (0–100) _____

New Belief: _____

How much do you believe the old core belief **right now**: intellectually? (0–100) _ emotionally? (0–100) _____

Evidence that contradicts old core belief and supports new belief	Evidence that seems to support old core belief with reframe

*Should situations related to an increase or decrease in the strength of the belief be topics for the agenda?

FIGURE 33.7–8. Core belief worksheet. (Adapted from Beck JS. *Cognitive Behavior Therapy: Basics and Beyond*. 2nd ed. 2011. Reprinted with permission of Guilford Press.)

The situations above describe techniques to use with spontaneous images. Clinicians may also deliberately induce images, for example, using "covert rehearsal," to have patients discover cognitions or practical problems that could arise as they contemplate doing a task on their homework list or "action plan" (or taking their medication). They may do imaginal exposure with anxiety patients, asking them to picture themselves in a feared situation, fully exposing themselves to thoughts, emotions, and physiological arousal to help them learn that anxiety rises, peaks, and then declines on its own, without intervention.

In addition, clinicians may use imagery to help patients with personality disorders restructure the meaning of early, traumatic experiences. Initially, the clinician may ask the patient to recall an incident in which an important

core belief about the self, others, or the world originated or became significantly strengthened. After restructuring the belief through intellectual methods, the clinician may ask patients to imagine they are the child and to describe the traumatic experience as if it is happening right now. The clinician may then suggest that the "child" invite his/her older self (the patient's "intellectual" side) into the image to explain to the child (the patient's "emotional" side) what has just transpired and to offer alternative explanations for the child's overgeneralized and often inaccurate beliefs. Or the clinician might suggest that the child invite someone else into the image (a trusted adult) to protect and help the child.

Behavioral Techniques

Problem solving is a key technique for most patients unless they have a circumscribed problem with anxiety (e.g., specific phobia). Some patients also need to learn effective methods for solving problems and making decisions on their own.

Behavioral activation, accomplished through activity scheduling, is very important for most depressed patients and others, especially those with entrenched avoidance. Clinicians ask patients to verbally describe or track their hour to hour activities. They discuss which activities, usually passive and solitary, are maladaptive for patients to engage in frequently, because they maintain or worsen their mood (e.g., napping, resting, watching television, surfing the web, playing video games) and which activities have the potential to bring patients a sense of mastery and pleasure. Then they make a plan in session to increase their active engagement, either creating a model schedule with specific activities for the full day or scheduling specific activities throughout the week. Clinicians often have to help patients respond to automatic thoughts that could interfere with initiating or completing activities or that could decrease the potential of these activities to increase their experience of mastery or pleasure.

Some patients, especially those who are depressed, have difficulty making decisions. Creating a four square chart can help. At the top, patients can list advantages and disadvantages of pursuing one option. At the bottom, they can list advantages and disadvantages of pursuing another option.

Graded task assignments are important for patients who are depressed, procrastinating, or overwhelmed by what they have to do. The clinician teaches them to break a large task into smaller components. A patient whose apartment is quite messy might initially just tackle the dishes in the sink or go through each room with a trash bag, looking for obvious things to throw away. Tasks can also be broken down by time: spend 10 minutes separating junk mail from important mail, spend 10 to 15 minutes paying bills.

Exposure and Response Prevention

This set of techniques, a form of behavioral experiment, is often used with patients who have OCD. In a graded fashion, the clinician exposes patients to feared situations, coaching them to refrain from engaging in compulsions so they can test their predictions and increase their tolerance of anxiety.

Motivational Techniques

Patients vary in their interest in therapy. To boost motivation to engage in treatment, work collaboratively in session, and follow through with action plans between sessions, clinicians may use a variety of techniques including the use of motivational interviewing, reminders of patients' key values and goals, an examination of advantages and disadvantages of various behavioral options, and Socratic questioning of cognitions that decrease their willingness to change.

Emotional Regulation Techniques

An important therapeutic goal is for patients to understand and accept their negative emotions and to discover that they can tolerate negative emotion without turning to maladaptive strategies such as avoidance, self-harm, aggression, or the use of substances. Psychoeducation, Socratic questioning of beliefs about the experience of emotion, and behavioral experiments to test these beliefs are often used to facilitate this understanding.

Patients also benefit from learning specific skills to better regulate their emotional arousal. They include mindfulness techniques, relaxation or breathing exercises, refocusing or distraction, reading therapy notes, positive social interaction, and self-soothing activities such as listening to music, physical contact with a pet, or walking in nature. Patients may also benefit from compassion therapy skills if their distress is related to a high degree of self-criticism.

Environmental Interventions

Some patients continually face such adverse circumstances that an environmental change is required. Patients who are emotionally or physically abused in their living environment, for example, may need help to identify a new place to live, at least for a while. Other patients might benefit from an examination of the pros and cons of staying in a hostile work environment, and, if needed, coaching on job seeking and interviewing skills.

Interpersonal Techniques

Techniques focus on solving interpersonal problems, teaching social skills, and modifying maladaptive beliefs about themselves, others (including, at times, the clinician), and relationships. Roleplaying is often used to help patients increase their facility in initiating and sustaining casual conversation, being appropriately assertive, or accurately reflecting what an intimate has said. Including a family member or intimate in session, whether or not there is friction in their relationship, can help clinicians round out their conceptualization of patients and their problems and provide an opportunity for psychoeducation and cooperation in helping patients reach their goals.

Biological Interventions and Treatment Adherence

It is important for clinicians to look broadly at patients and, for some, to encourage healthier habits—facilitating change, for example, in eating habits, caffeine intake, use of substances, sleep, and exercise. Medication adherence can be increased by considering advantages and disadvantages of taking medication, examining interfering cognitions, promoting tolerance of side effects, behavioral experiments (such as taking the first dose in the office), and checking on practical matters such as potential problems in obtaining the medication or remembering to take it.

CLINICAL ISSUES

Indications

Cognitive therapy may be used to (1) eliminate or moderate the symptoms of a psychiatric disorder or psychological distress as an independent treatment or in combination with medication; (2) modify underlying beliefs and schemas that contribute to the development and maintenance of disorders; (3) address psychosocial problems (e.g., marital discord, job stress, caregiver burnout) that may have preceded, coincided with, or been caused by the disorder; (4) reduce the likelihood of relapse or recurrence of a disorder after symptoms have resolved; (5) increase adherence to recommended medical treatment (e.g., medication, weight-loss, diabetes management); or moderate the impact of a medical condition on level of functioning (e.g., chronic pain, cancer).

Patients. Ideal cognitive therapy patients have the ability to recognize and label their emotions, automatic thoughts, behaviors, and the connection between these experiences. Nonetheless, patients who initially have difficulty recognizing and labeling these phenomena are typically able to learn these skills through instruction and practice. Accordingly, even average intelligence is not a prerequisite for treatment. Motivation for therapy is also helpful, but not required, as patients who are hopeless, unmotivated, or lethargic may become more invested in treatment as they

begin to experience improvement. New patients to psychotherapy often report appreciating the cognitive therapy approach as they see it as being pragmatic and comprehensible. Some long-term patients of other psychotherapy approaches may have an initial period of adjustment into cognitive therapy due to its structured nature. Nonetheless, most patients are able to make the transition, and structure can be flexible for those who cannot or do not wish to adjust to structured sessions.

Clinicians. The ideal cognitive therapist is attentive, empathic, and nonjudgmental, as is desired in most models of psychotherapy. To be maximally effective in an active treatment such as cognitive therapy, clinicians need to be comfortable being directive in a collaborative manner. They also must be able to ask for and receive feedback in a nondefensive manner. Although a skilled cognitive therapist knows how to sit back and listen when appropriate, they do not allow long silences as a matter of routine, and they are willing to answer most questions with a therapeutically direct answer. Cognitive therapists aim to be transparent and share their conceptualization and rationales for interventions. Overall, the goal of cognitive therapy is for patients to learn skills which they can use in treatment and also independently. Treatment frequently lasts for weeks or months, but not usually years.

As would be expected, several studies have demonstrated a high correlation between a clinician's level of competency and successful treatment outcomes (e.g., Strunk et al., 2010). Competency in cognitive therapy typically requires more than a passive knowledge of the theory and techniques. Clinicians who wish to specialize in cognitive therapy usually require at least 1 year of supervised clinical experience using the cognitive model. Ideally, supervision is performed by an expert cognitive therapist who reviews, rates, and provides feedback on recordings of therapy sessions.

Limitations

Early outcome studies on cognitive therapy indicated that depressed patients could show significant improvement in approximately 12 weeks (of biweekly and weekly sessions). Likewise, individuals with panic disorder evinced a significant reduction in panic attacks after four to eight sessions of cognitive therapy. While the short-term results of cognitive therapy are impressive, they have, at time, been interpreted too broadly, creating an inaccurate perception that all types of pathology and clinical presentations will respond as rapidly.

Research on cognitive therapy for individuals diagnosed with bipolar disorder, chronic depression, and personality disorders allow for a longer course of treatment (e.g., 1 year), and results have been promising. Moreover, even for uncomplicated cases, "booster sessions" at longer

intervals (e.g., 3 or 6 months) after termination of the course of regular treatment can help with the maintenance of treatment gains and prevent relapse.

Complications

At present, there are no known complications caused by cognitive therapy when it is used properly and in the appropriate context. The model is designed to teach patients a variety of skills to manage their presenting problems, which is meant to give them a sense of empowerment and self-efficacy. Accordingly, the cognitive therapy model does not encourage patient dependence on the clinician. Termination issues, which are relevant in any form of psychotherapy, are dealt with proactively by the clinician who explicitly educates the patient about the time-effectiveness of treatment, the importance of learning to be one's own therapist, and the use of booster sessions.

Clinical crises are handled with a combination of standard risk-reduction procedures (e.g., increased frequency of sessions, inclusion of significant others, emergency consultations, hospitalizations), along with a cognitive conceptualization of the crisis situation. Once an acute crisis subsides, clinicians explore the patient's vulnerability (the activation of his or her schemas) that was triggered by the activating event and help the patient develop alternative coping strategies for similar future situations. The goal is to prevent such instances in the future and for the patient to become less vulnerable to life stressors.

Contraindications

Cognitive therapy has been demonstrated to be efficacious as part of a treatment package for severe and persistent mental illness, although it is generally recommended that pharmacotherapy options be assessed as part of the treatment regimen for disorders such as schizophrenia and bipolar disorder. Moreover, while there is some evidence that cognitive therapy alone can be helpful for cases of bipolar disorder in which medical complications make the use of pharmacotherapy problematic; it is generally not recommended for patients with such psychiatric disorders to pursue cognitive monotherapy.

GOALS OF TREATMENT

Cognitive therapy has two major goals. The first is to reduce patients' symptomatology and suffering as quickly as possible and help them reach remission. The second is to teach them skills which they can employ long-term to reduce relapse.

ETHICAL ISSUES

The ethical issues that arise in many forms of psychotherapy also may arise in cognitive therapy: eliciting informed consent, reporting of child abuse, breaking confidentiality when the patient's life, or the life of other people are in imminent danger, maintaining appropriate boundaries, and avoidance of dual relationships. Additionally, clinicians are cognizant of not continuing regularly scheduled sessions when ongoing treatment is no longer deemed necessary. Once patients have made strides in accomplishing their goals, are no longer experiencing acute symptoms, and have demonstrated an ability to maintain treatment gains, clinicians may suggest tapering the frequency of sessions, followed by scheduling only periodic "booster" sessions.

RESEARCH AND EVALUATION

Systematically controlled studies have demonstrated the efficacy of cognitive therapy for patients with clinical problems, such as major depressive disorder, bipolar disorder, GAD, panic disorder, SAD, OCD, and somatoform disorders; anorexia, bulimia, and body dysmorphic disorder; substance use disorders; anger; chronic pain; relationship discord; personality disorders; insomnia; a variety of childhood disorders, as well as schizophrenia. A review of meta-analyses concluded that the strongest support for cognitive therapy exists in the treatment of anxiety disorders, somatoform disorders, bulimia, anger, and general stress. Moreover, cognitive therapy has been successfully adapted across various patient characteristics including; age (children through geriatric populations), gender, sexual orientation, socioeconomic status, and racial or ethnic background; in addition to treatment setting (e.g., outpatient, inpatient, schools, forensic settings, medical centers). Accordingly, studies and meta-analyses have concluded that cognitive therapy is also an effective treatment when implemented in real-world settings with diverse patient populations (e.g., Chambless & Ollendick, 2001).

A common misconception in psychological intervention research is the notion of the Dodo Bird verdict, which concludes that all psychological treatments are equal in their effects. A meta-analysis concluded that cognitive therapy was significantly more efficacious than other bona fide interventions (e.g., psychodynamic, interpersonal, and supportive therapies) for anxiety and mood disorders. Cognitive therapy has also demonstrated superiority to other treatments when comparisons are limited to studies with high methodological quality (e.g., random assignment, evaluator blind to condition). Additionally, Hofmann et al. (2012) noted that 7 of 11 meta-analyses demonstrated higher response rates for cognitive therapy than comparison conditions, while only one identified lower response rates for cognitive therapy.

Researchers have used functional neuroimaging to study the influence of psychological interventions on brain function. Neuroimaging studies

have demonstrated that individuals with psychological disorders tend to have different brain functioning than nonclinical controls. Interestingly, studies have found the brain functioning of individuals with certain psychological disorders to resemble control subjects following successful cognitive therapy. Neuroimaging research has also compared brain functioning of individuals undergoing treatment with cognitive therapy to those taking psychotropic medication. The results indicated that cognitive therapy caused similar changes in brain function to psychotropic medication in the treatment of OCD, SAD, specific phobias, and depression. Nonetheless, there are differences in how disorders present in terms of both symptoms and brain function. A more nuanced study found some differences in brain function between successful cognitive therapy and pharmacotherapy. This suggests that the type of treatment (cognitive therapy, pharmacotherapy) could be tailored to an individual based on pretreatment brain function patterns.

Cognitive therapy has been adapted to be administered in Internet (e.g., iCBT) and computer-based (e.g., cCBT) formats. A meta-analysis concluded that Internet- and computer-based cognitive therapy demonstrated large effect sizes when compared to a control group in the treatment of all disorders that were studied (depression, panic disorder, SAD, and GAD). Moreover, patient adherence was good and patients reported being satisfied with treatment.

In sum, although cognitive therapy was originally developed as a means to treat unipolar depression and anxiety disorders, it has been successfully adapted to treat various clinical populations, including patient groups who were traditionally thought to be unresponsive to psychosocial treatments (e.g., schizophrenia). A growing body of research has demonstrated the efficacy and effectiveness of cognitive therapy in addressing various disorders and patient populations. While some researchers study mechanism of change, additional studies, using mediational analyses, should be conducted to continue to establish empirical support for the cognitive model and to aid in refining treatment. Additionally, as researchers become better able to identify biological processes of psychological disorders, future research may attempt to integrate biological processes with the cognitive model, consistent with the National Institute of Mental Health Research Domain Criteria (RDoC) initiative, which was implemented to develop an alternative, more homogeneous and etiologically relevant, psychological disorder classification system to the DSM. Finally, efforts to disseminate cognitive therapy to clinicians have lagged behind the growing efficacy and effectiveness research base. Future research should place a greater emphasis on determining the most effective and efficient methods to disseminate cognitive therapy to clinicians so that patients have greater access to competently delivered evidence-based treatment.

REFERENCES

Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS ONE*. 2010;5:e13196.

Basco MR, Rush AJ. *Cognitive-Behavioral Therapy for Bipolar Disorder*. 2nd ed. New York: Guilford Press; 2005.

Beck AT. *Depression: Clinical, Experimental and Theoretical aspects*. New York: Hoeber Medical Division, Harper and Row; 1967.

Beck AT. The evolution of the cognitive model of depression and its neurobiological correlates. *Am J Psychiatry*. 2008;165:969–977.

Beck AT, Beck JS. *The Personality Belief Questionnaire. Unpublished Assessment Instrument*. Bala Cynwyd, PA: The Beck Institute for Cognitive Therapy and Research; 1991.

Beck JS. *Cognitive Therapy for Challenging Problems: What to do when the Basics Don't Work*. New York: Guilford Press; 2005.

Beck JS. *The Beck Diet Solution*. Birmingham, AL: Oxmoor House; 2007.

Beck JS. *Cognitive Behavior Therapy: Basics and Beyond*. 2nd ed. New York: Guilford Press; 2011.

Beck AT, Davis DD, Freeman A. *Cognitive Therapy of Personality Disorders*. 3rd ed. New York: Guilford Press; 2015.

Beck AT, Haigh EA. Advances in cognitive theory and therapy: the generic cognitive model. *Annu Rev Clin Psychol*. 2014;10:1–24.

Beck AT, Rector NA, Stolar N, Grant P. *Schizophrenia: Cognitive Theory, Research, and Therapy*. New York: Guilford Press; 2009.

Beck AT, Wright FD, Newman CF, Liese BS. *Cognitive Therapy of Substance Abuse*. New York: Guilford Press; 1993.

Beck JS, Liese B, Najavits LM. Cognitive therapy. In: Frances RJ, Miller SI, eds. *Clinical Textbook of Addictive Disorders*. 4th ed. New York: Guilford Press; 2016:563–587.

Chambless DL, Ollendick TH. Empirically supported psychological interventions: controversies and evidence. *Annu Rev Psychol*. 2001;52:685–716.

Clark DA, Beck AT. *Cognitive Therapy of Anxiety Disorders: Science and Practice*. New York: Guilford Press; 2010.

Clark DM, Beck AT. *Scientific Foundations of Cognitive Theory and Therapy of Depression*. New York: John Wiley & Sons; 1999.

Clark DM, Wells A. A cognitive model of social phobia. In: Heimberg RG, Liebowitz MR, Hope DA, Schneier FR, eds. *Social Phobia: Diagnosis, Assessment, and Treatment*. New York: Guilford Press; 1995:69–93.

Dobson D, Dobson KS. *Evidence-Based Practice of Cognitive-Behavioral Therapy*. New York: Guilford Press; 2017.

Fairburn CG, Cooper Z, Shafran R. Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. *Behav Res Ther*. 2003;41:509–528.

Fuchs T. Neurobiology and psychotherapy: an emerging dialogue. *Curr Opin Psychiatry*. 2004;17:479–485.

Furmark T, Tillfors M, Marteinsdottir I, et al. Common changes in cerebral blood flow in patients with social phobia treated with citalopram or cognitive-behavioral therapy. *Arch Gen Psychiatry*. 2002;59:425–433.

Goldapple K, Segal Z, Garson C, et al. Modulation of cortical-limbic pathways in major depression: treatment-specific effects of cognitive behavior therapy. *Arch Gen Psychiatry*. 2004;61:34–41.

Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. *Cognit Ther Res*. 2012;36:427–440.

Kaczkurkin AN, Foa EB. Cognitive-behavioral therapy for anxiety disorders: an update on the empirical evidence. *Dialogues Clin Neurosci*. 2015;17:337–346.

Lewis G, Lewis G. No evidence that CBT is less effective than antidepressants in moderate to severe depression. *Evid Based Mental Health*. 2016;19, 125.

Lusk P, Melnyk BM. The brief cognitive-behavioral COPE intervention for depressed adolescents: outcomes and feasibility of delivery in 30-minute outpatient visits. *J Am Psychiatr Nurses Assoc*. 2011;17:226–236.

Mansell W, Morrison AP, Reid G, Lowens I, Tai S. The interpretation of, and responses to, changes in internal states: An integrative cognitive model of mood swings and bipolar disorder. *Behav Cogn Psychother*. 2007;35:515–539.

Newman CF, Leahy RL, Beck AT, Reilly-Harrington NA, Gyulai L. *Bipolar Disorder: A cognitive Therapy Approach*. Washington, DC: American Psychological Association; 2002.

Perivoliotis D, Grant PM, Beck AT. *Recovery-Oriented Cognitive Therapy for Schizophrenia: A Comprehensive Treatment Manual*. New York: Guilford Press; in press.

Scher CD, Ingram RE, Segal ZV. Cognitive reactivity and vulnerability: Empirical evaluation of construct activation and cognitive diatheses in unipolar depression. *Clin Psychol Rev*. 2005;25:487–510.

Strunk DR, Brotman MA, DeRubeis RJ, Hollon SD. Therapist competence in cognitive therapy for depression: Predicting subsequent symptom change. *J Consult Clin Psychol*. 2010;78:429–437.

Tolin DF. Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clin Psychol Rev*. 2010;30:710–720.

Uher R, Pavlova B. Long-term effects of depression treatment. *Lancet Psychiatry*. 2016;3:95–96.

Wells A. *Metacognitive Therapy for Anxiety and Depression*. New York: Guilford Press; 2009.

Wilhelm S, Berman NC, Keshaviah A, Schwartz RA, Steketee G. Mechanisms of change in cognitive therapy for obsessive compulsive disorder: role of maladaptive beliefs and schemas. *Behav Res Ther*. 2015;65:5–10.

Wright JH, Sudak D, Turkington D, Thase ME. *High Yield Cognitive-Behavior Therapy for Brief Sessions: An Illustrated Guide*. Washington, DC: American Psychiatric Publishing; 2010.

▲ 33.8 Interpersonal Psychotherapy

ROBERT W. GUYNN, M.D.

It is more accurate to speak of the interpersonal psychotherap*ies* (plural). There have been long-term, insight-oriented, dynamically based and open-ended interpersonal psychotherapies, and there are more modern, short-term, time-limited approaches that avoid both extensive theory and insight. What all these therapies hold in common is the belief that the interpersonal relationships of a patient, now and in the past, to a greater or lesser extent, have a role in the etiology of his or her psychopathology and its treatment. It is unfortunate that all forms of the interpersonal psychotherapies bear the same name. In order to minimize confusion, for the rest of the discussion the field will be split into the insight-oriented interpersonal psychotherapies (IO-IPT) and non–insight-oriented interpersonal psychotherapies (IPT in keeping with the current usage of that abbreviation).

HISTORY

To the American psychiatrist Harry Stack Sullivan (1892–1949) goes the honor of the title of the Father of Interpersonal Psychiatry. It was from the 1920s through the 1940s that Sullivan developed his ideas as an alternative formulation to classic psychoanalytical theory. Simply put, Sullivan theorized that everything people consider human comes from their interactions with others. He postulated that there are two basic drives: (physical) satisfaction (food, warmth shelter, sex…) and (interpersonal) security. Security here is the sense that one is an adequate human being (having a healthy self-esteem, feeling confident to handle stress, being comfortable among others, being interpersonally competent, being relatively anxiety-free, etc.). Sullivan pointed out that not only do humans learn language and culture from the members of their family but they also learn how to separate fantasy from reality, develop a sense of self and self-esteem (or lack thereof), develop strategies for coping with anxiety, etc. Sullivan postulated social stages of child development (rather than psychosexual ones), the normal outcome of which being the more-or-less enduring patterns of thinking and behaving that constitute ones personality. These patterns develop not so much because of inner conflicts, according to Sullivan, but as an attempt to deal with the human environment of one's family, school, and community. To the extent that these patterns are significantly maladaptive one has a mental illness.

It is probably fair to say that, although Sullivan himself was a clinician, he is primarily remembered for his theories, through his lectures and

largely posthumous books. It was left to others to flesh out the therapeutic techniques. Some, such as Frieda Fromm-Reichmann (1889–1957), Harold Searles (1918–2015), Karen Horney (1885–1952), and Hilda Bruch (1904–1984) blended some of the Sullivan's ideas with more classic psychoanalytic thinking. They, thereby, were able adapt the traditional theory and therapy to nontraditional populations including schizophrenia. Others such as Donald Kiesler (1933–2007) proposed alternative therapeutic strategies. Of note, among all of these early writers is the recognition that a therapist is not and cannot be a neutral observer.

The general tasks of an insight-oriented interpersonal psychotherapist is to identify patterns of behavior and thinking from the patient's history, support the self-esteem of the patient and re-enforce it appropriately, address the positive aspects of the patient and his or her successes. The therapist is friendly, even an advocate, but not a friend. An important element of the therapy is the assumption that the therapist is a participant, and, therefore, cannot be entirely objective and cannot avoid being the target of the patient's stereotypical behavior. If the patient tends to interact with others by being manipulative, challenging, provocative, or suspicious, sooner or later he or she will interact with the therapist in the same way. The ability of the therapist to read his or her own reactions to the patient objectively and reflect them therapeutically is an essential part of the treatment.

Kiesler, to take but one specific example, built on Sullivan's ideas and developed a more formal therapeutic strategy. He very much emphasized that the therapy sessions are dyadic—that is, the therapist is a participant observer. He uses the term "impact messages" to designate all the feelings, thoughts, and behaviors of the therapist in response to the patient. Kiesler distinguishes several groups of impact messages: direct feelings, action tendencies, cognitive attributions, and fantasies. The therapist consciously and continuously monitors his or her reactions (impact messages) and assumes that the patient is provoking in the therapist what he or she provokes in other significant relationships. That is, the therapist, far from being aspiring to be neutral, uses his or her own reactions to the patient to understand the patient's maladaptive interpersonal style. The impact messages are not considered counter-transference, as such, but the essentially normal response that anyone might have to the repetitive interpersonal patterns that the patient has developed as a strategy for living, perhaps because it is what worked in the nuclear family. The difference is that the therapist's response and reactions will attempt to be therapeutic rather than "usual." In fact, the therapist's failure to respond to the patient's interpersonal psychopathology in the expected way begins to undermine the "usefulness" of the psychopathology.

For example, if a person treats people in his or her life as objects to be controlled, then there is no reason to think that the patient would treat the

therapist any differently. In this example the therapist might find him or herself becoming progressively irritated with the patient as the sessions go on. The therapist would then internally analyze the source of the irritation and, hopefully, ascribe it to the subtle, or not so subtle, attempts at control by the patient. This strategy makes the therapist a very active participant, monitoring his or her own feelings and reactions as a strategy to both understand and treat the patient. If the insight-oriented interpersonal psychotherapist were to choose to share his or her own reaction with the patient, in this case, the interpretation would not be to reflect back on a specific early relationship of the patient but point out to the patient what the therapist is sensing ("I feel that you are trying to control me") and wonder aloud if this is not an example of the patient's automatic and well-established behavioral pattern presumably identifiable from the patient's interpersonal history. The therapist might then discuss with the patient alternative and less stereotypic strategies.

Another example might be the shy and anxious person who treats interpersonal interactions as painful reminders and proofs of his or her inadequacy ("I never know what to say and must come across as stupid") will probably be looking for judgments from the therapist and be unusually sensitive to any reaction of the therapist that implies criticism. One might expect such a patient to provoke in the therapist feelings that the patient must not be criticized. The IO-IPT therapist might choose to comment to the patient that he or she is feeling that the patient is projecting a sense of fragility and vulnerability—is that really necessary?

In fact, Kiesler suggests for any given statement by the patient, the therapist can choose among any of the six categories of reply. He or she may simply respond to the manifest content of the statement or may ignore it and change the subject. One may also interpret the statement of the patient as an unconscious message either of the patient to himself or the patient to the therapist. Finally, the therapist may choose to share either his personal reflection on or his reaction to the patient's statement.

In this example of IO-IPT, Kiesler's therapy has four phases: (1) hooking, (2) uncovering, (3) resolving, and (4) termination. No set number of sessions is proposed. The last three phases would seem to be self-evident and do not need exposition here. The first phase ("hooking") might be worth addressing. It was an assumption that many individuals seek to come to therapy not so much to become better but to become better at their psychopathology. Such patients will quickly tire of therapy once whatever provoked their anxiety and drove them to seek help abates. The "hooking phase" is a conscious attempt by the therapist to intrigue the patient in the goal of health.

In other cases the problem may at least partially reside outside the patient's own interpersonal style. An example would be the case of the depressed woman married to a perfectionist. Without a trace of insight, the

patient pinpointed the onset of her depression to the date of her marriage. Being not as educated as her husband, she readily accepted him as an authority at all levels. His pattern was to "lovingly" make suggestions about how the patient might do things better from how she cooked to how she held the leashes when walking the dogs. After a couple of years the patient gradually had lost much of her self-esteem and had become paralyzed when trying to make any decision. During the exploration of the pattern of the interpersonal relationship between the patient and her husband, he announced that he had found another. Much to the surprise of all, the patient was not devastated but actively threw him out and never looked back. The depression resolved.

The caveat to all of the above is the fact that such approaches have not been subjected to rigorous scientific study. However logical and attractive the theories and approaches may seem, they have not been put to sufficient testing.

TIME-LIMITED INTERPERSONAL PSYCHOTHERAPY

A more modern version of interpersonal psychotherapy, IPT, is a nonpsychodynamic, time-limited psychotherapy developed in the 1970s by Gerald Klerman and colleagues for the treatment of nonbipolar, nonpsychotically depressed outpatients. This version of interpersonal psychotherapy makes no assumptions about the causes of psychiatric illness, other than that the development and maintenance of some psychiatric illnesses occur in a social and interpersonal context and that the onset, responses to treatment, and outcomes are influenced by the interpersonal relations between the patient and significant others. Klerman's approach is reminiscent of cognitive-behavioral therapy in that the focus is on here and now. There are, also, a set number of sessions, transference may be ignored or undermined, and a workbook may be used. As with all interpersonal psychotherapies, in IPT the focus is primarily on the patient's current and recent interpersonal relationships. Other differences between IO-IPT and IPT are that in the latter there is less interest in the origin of the problem and the therapist does not necessarily reflect on the relationship between the therapist and the patient. A major advantage of ITP over the other interpersonal approaches is that IPT can be researched.

Time-limited interpersonal psychotherapy moves through defined phases, each of which is associated with specific strategies and tasks for the therapist and patient. Its well-defined treatment strategies are aimed at resolving problems within four social domains: grief, interpersonal role disputes, role transitions, and interpersonal deficits. The overall goal is to reduce and/or eliminate psychiatric symptoms by improving the quality of the patient's current interpersonal relations and social functioning. Time-limited interpersonal psychotherapy has been found to be efficacious for

major depression and has also been successfully adapted to treat other types of mood and nonmood disorders. Interpersonal psychotherapy has also been successfully applied to a group format and is being evaluated in a number of trials. The extensive empirical background and theoretical foundation, as well as the strategies and techniques of interpersonal psychotherapy, are fully described in comprehensive books by Myrna Weissman and her colleagues.

Differences from Other Psychotherapies

Time-limited interpersonal psychotherapy is, of course, not the only psychological approach for the treatment of psychiatric disorders. The methods and techniques of many schools of psychotherapy share common ground, including attempts to help patients gain a sense of mastery and reduce social isolation. A major difference among the therapies lies in their conceptualization of the causes of the patient's problems and the typical length and foci of treatment. Although they use similar techniques, IPT also differs from cognitive and behavioral approaches in that maladaptive thoughts and behaviors are addressed only as they apply to problematic interpersonal relations, with the goal to change the relationship pattern rather than associated cognitions. In contrast to other therapies, IPT recognizes but does not necessarily focus on the patient's personality characteristics.

The focus of IPT is on identifying and changing the maladaptive interpersonal context in which the psychiatric problem has developed and is perpetuated. It is designed to help patients identify and address largely current interpersonal problems within four domains: Grief, identified as the problem area when the onset of the patient's symptoms is associated with the loss of a person or a relationship, either recent or in the past; interpersonal role disputes, which are conflicts with a significant other (e.g., a partner, other family member, coworker, or close friend) that emerge from differences in expectations about the relationship; role transitions, which include difficulties associated with a change in life status (e.g., graduation, leaving a job, moving, marriage/divorce, retirement, change in health status); and interpersonal deficits, which apply to those patients who are socially isolated or are in chronically unfulfilling relationships.

TECHNIQUES

Therapeutic Stance

Similar to the other therapies, IPT places importance on establishing positive therapeutic alliance between therapist and patient. Specifically, the interpersonal psychotherapeutic stance is one of warmth, support, and empathy. The interpersonal psychotherapist is active and advocates for the

patient rather than remaining neutral. Moreover the therapeutic relationship is not viewed as a classic manifestation of transference/countertransference, nor is it treated as a friendship. As noted, IPT is similar to many other therapies at the level of psychotherapeutic techniques and therapeutic stance but is distinct at the level of strategies.

Timeline of Treatment

The typical course of IPT lasts 12 to 20 sessions over a 4- to 5-month period. The initial phase is dedicated to identifying the problem area that will be the target for treatment. The intermediate phase is devoted to working on the target problem area(s). The termination phase is focused on consolidating gains made during treatment and preparing the patients for future work on their own (Table 33.8–1).

Initial Phase. Sessions one through five typically constitute the initial phase of IPT. After assessing the patient's current psychiatric symptoms and obtaining a history of these symptoms, the therapist explains the rationale of IPT, underscoring that therapy will focus on identifying and altering dysfunctional interpersonal patterns related to psychiatric symptomatology. To determine the precise focus of treatment the therapist will conduct an interpersonal inventory with the patient and will develop an interpersonal formulation based on this. In the interpersonal formulation, the therapist will link the patient's psychiatric symptomatology with the therapist's identification of the problem area(s). The patient's concurrence with the therapist's assessment of the problem and agreement to work on this area are essential before beginning the intermediate treatment phase.

Table 33.8–1.
Phases of Time-limited Interpersonal Psychotherapy

Initial phase: Sessions 1–5
Give the syndrome a name; provide information about the prevalence and characteristics of the disorder
Describe the rationale and nature of the therapy
Conduct the interpersonal inventory to identify the current interpersonal problem area(s) associated with the onset or maintenance of the psychiatric symptoms
Review significant relationships, past and present
Identify interpersonal precipitants of episodes of psychiatric symptoms
Select and reach consensus about the interpersonal problem area(s) and treatment plan with patient
Intermediate phase: Sessions 6–15
Implement strategies specific to the identified problem area(s)
Encourage and review work on goals specific to the problem area
Illuminate connections between symptoms and interpersonal events during the week
Work with the patient to identify and manage negative or painful affects associated with his or her interpersonal problem area

Termination phase: Sessions 16–20
Discuss termination explicitly
Educate the patient about the end of treatment as a potential time of grieving; encourage the patient to identify associated emotions
Review progress to foster feelings of accomplishment and competence
Outline goals for remaining work; identify areas and warning signs of anticipated future difficulty
Formulate specific plans for continued work after termination of treatment

DIAGNOSIS AND ASSIGNMENT OF THE SICK ROLE. After a thorough psychiatric interview has been conducted, the therapist gives the patient a formal diagnosis from the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) and the patient is formally diagnosed and assigned the sick role. Therapist and patient then discuss the diagnosis, as well as what might be expected from treatment. Assignment of the sick role during this phase serves the dual function of granting the patient both the permission and the responsibility to recover. The purposes of assigning the sick role are both theoretical and practical. Consistent with the theoretical model, receiving a formal diagnosis reinforces the idea that the patient has a known condition that can be treated. The giving of a diagnosis also explicitly identifies the patient as being in need of help. The sick role is not assigned to condescend to the patient but, rather, is designed temporarily to exempt the individual from the responsibilities so as to devote full attention to recovery. This is particularly important for patients who tend to set aside their needs and desires to care for and please others. If this applies in a particular case, the therapist may explicitly highlight the patient's excessive caretaking tendencies and encourage the patient to redirect this energy from others toward his or her recovery. In doing so, the therapist clarifies the rationale behind interpersonal psychotherapy—that by improving the patient's patterns of interpersonal functioning, the patient's psychiatric symptoms are expected to improve as well. Symptom relief starts with helping the patient to understand that his or her psychiatric symptoms are part of a known syndrome that responds to several treatments and has a good prognosis.

IDENTIFY CURRENT INTERPERSONAL PROBLEMS: THE INTERPERSONAL INVENTORY. The therapist conducts an interpersonal inventory, which includes a review of the patient's past and current social functioning and close relationships. A thorough interpersonal inventory is essential for adequate case formulation and development of an optimal treatment plan. The clinical importance of investing the time involved in conducting a comprehensive interpersonal inventory cannot be overemphasized; accurate identification of the patient's primary problem area can be difficult, but it is key to success in therapy.

For each individual who is important in the patient's life, the following information is assessed: frequency of contact, activities shared, satisfactory

and unsatisfactory aspects of the relationship, and ways in which the patient would like to change the relationship. The interpersonal inventory provides a structure for elucidating the social and interpersonal context of the onset and maintenance of psychiatric symptoms and delineates the focus of treatment. Of particular importance are changes in relationships (e.g., the death of a significant other, a change in jobs, an increase in marital discord, disconnection from a friend). The therapist obtains a chronological history of significant life events, fluctuations in mood and self-esteem, interpersonal relationships, and psychiatric symptoms. From this review, the therapist can make a connection between certain life experiences and psychiatric symptoms. As this interrelationship is delineated, patients understand more clearly the rationale of interpersonal psychotherapy.

ESTABLISH THE RELEVANT PROBLEM AREAS. On completion of the interpersonal inventory, the therapist develops an individualized interpersonal formulation, linking the patient's symptoms to one of the four interpersonal problem areas—grief, interpersonal deficits, interpersonal role disputes, or role transitions. Although the therapist explicitly assigns a problem area, the patient needs to concur with the salience of the problem area proposed and agree to work on it in treatment. Because the four problem areas are relatively inclusive, it is usually easy for the therapist to provide an individualized formulation. Although some patients may present for treatment with difficulties in several problem areas, the time-limited nature of the treatment necessitates a focused approach. The therapist, with the agreement of the patient, should assign one or, at most, two problem area(s). If more than one problem area is identified, the problem area that seems most likely to be responsive to treatment is addressed first. Tackling the most soluble problem first has the advantage that progress on one front often leads to progress on others. Not only are the patient's morale and overall sense of competence enhanced when progress is made on a problem, but also barriers to progress in other areas may be eroded or disappear.

COLLABORATIVELY DEVELOP TREATMENT GOALS. After the major interpersonal problem area(s) associated with the onset or maintenance of the disorder is identified, the therapist and patient work together to formulate a treatment plan with specified goals. The goals developed at this stage will be referenced at each future session and will guide the day-to-day work of the treatment. Goals should flow from the identified problem area and be designed to address it. After identification of a goal, the therapist and patient collaboratively identify specific steps the patient will take to improve relationships and socialization. To the extent possible, goals should be summarized in such a way as to include reference to specific

individuals, events, and interpersonal themes to help ensure that they are as personally meaningful to the patient as possible. Before the end of the initial phase, patients may be given a written summary of their goals, which serves as a treatment contract to guide their work for the remainder of treatment.

Intermediate Phase. The intermediate phase typically lasts from eight to ten sessions and continues the "work" of the therapy. Once the patient and therapist have agreed on the primary problem area and have set treatment goals, the intermediate phase of treatment begins. An essential task throughout the intermediate phase is to strengthen the connections the patient makes between the changes he or she is making in his or her interpersonal life and the changes in his or her psychiatric symptoms.

During the intermediate phase, the therapist implements the treatment strategies specific to the identified problem area as specified by Weissman and colleagues (Table 33.8–2). For an in-depth discussion of the interpersonal psychotherapy goals and strategies, the therapist is encouraged to read the comprehensive text by Weissman and colleagues.

PROBLEM AREAS

Grief. Grief is identified as the problem area when the onset of the patient's symptoms is associated with the death of a loved one, either recent or past. Grief, however, may not be limited to the physical death of a loved one. Grief can also result from the loss of a significant relationship or the loss of an important aspect of one's identity. The goals for treating complicated bereavement include facilitating mourning and helping the patient to find new activities and relationships to substitute for the loss. Reconstructing the relationship—both its positive and its negative aspects—is central to the assessment of not only what has been lost, but also what is needed to counter the idealization that commonly occurs. As patients become less focused on the past, they should be encouraged to consider new ways of becoming more involved with others and establishing interests.

Role Transitions. Role transition (which can overlap with grief) includes any difficulty resulting from a change in life status. Common role transitions include a career change (promotion, firing, retirement), a family change (marriage, divorce, birth of a child), the beginning or end of an important relationship, a move, graduation from school, or diagnosis of medical illness. The goals of therapy include mourning and accepting the loss of the old role, recognizing the positive and negative aspects of both the old and new roles, and restoring the patient's self-esteem by developing a sense of mastery in the new role. Key strategies in achieving these goals include thoroughly exploring the patient's feelings related to the role change and

encouraging the patient to develop new skills and adequate social support for the new role.

Interpersonal Role Disputes. Such disputes are conflicts with a significant other (e.g., a partner, other family member, coworker, or close friend) that emerge from differences in expectations about the relationship. The goals of treatment include clearly identifying the nature of the dispute and exploring options to resolve it. It is important to determine the stage of the dispute; once the stage of the dispute becomes clear, it may be important to modify the patient's expectations and remedy faulty communication to bring about adequate resolution. It may be particularly helpful to explore how nonreciprocal role expectations relate to the dispute. If resolution is impossible, the therapist assists the patient in dissolving the relationship and in mourning its loss.

Interpersonal Deficits. Interpersonal deficits refer to patients who are socially isolated or who are in chronically unfulfilling relationships. The goal is to reduce the patient's social isolation by helping him or her to enhance the quality of existing relationships and encouraging the formation of new relationships. To help these patients, it is necessary to determine why they have difficulty in forming or maintaining relationships. Conducting a review of past significant relationships will be particularly useful in making this assessment. During this review, attention should be given to both the positive and negative aspects of each relationship, as well as to an investigation of potentially recurrent patterns in these relationships. It may also be appropriate to examine the nature of the patient–therapist relationship because as this may be the patient's only close relationship, and it can be observed commented on.

Table 33.8–2.
Interpersonal Problem Areas: Description, Goals, and Strategies

Problem Area	Description	Goals	Strategies
Grief	Complicated bereavement after death of a loved one	Facilitate the mourning process Help the patient to establish interest in new activities and relationships to substitute for what has been lost	Reconstruct the patient's relationship with the deceased Explore associated feelings (negative and positive) Consider ways of becoming reinvolved with others
Interpersonal deficits	A history of social impoverishment, inadequate or unsustaining interpersonal	Reduce patient's social isolation Enhance quality of any existing relationships Encourage the formation	Review past significant relationships, including negative and positive aspects Explore repetitive patterns in

	relationships	of new relationships	relationships Note problematic interpersonal patterns in the session and relate them to similar patterns in the patient's life
Interpersonal role disputes	Conflicts with a significant other—a partner, other family member, coworker, or close friend	Identify the nature of the dispute Explore options to resolve dispute Modify expectations and faulty communication to bring about a satisfactory resolution If modification is unworkable, encourage patient to reassess the expectations for the relationship and to generate options to either resolve it or dissolve it and mourn its loss	Determine the stage of the dispute: renegotiation (calm down participants to facilitate resolution); impasse (increase disharmony to reopen negotiation); dissolution (assist mourning and adaptation) Understand how nonreciprocal role expectations relate to the dispute Identify available resources to bring about change in the relationship
Role transitions	Economic or family change—the beginning or end of a relationship or career, a move, promotion, retirement, graduation, diagnosis of a medical illness	Mourn and accept the loss of the old role Recognize the positive and negative aspects of the new role and assets and liabilities of the old role Restore self-esteem by developing a sense of mastery regarding the demands of the new role	Review positive and negative aspects of old and new roles Explore feelings about what is lost Encourage development of social support system and new skills called for in new role

Reprinted with permission from Wilfley D, Stein R, Welch R. Interpersonal psychotherapy. In: Treasure J, van Furth E, Schmidt U, eds. *Handbook of Eating Disorders*. 2nd ed. Hoboken, NJ: Wiley; 2003:253.

FACILITATE PATIENT'S WORK ON IDENTIFIED GOALS. In structuring the intermediate phase, the IPT takes a moderate position between the extremes of being highly active and merely reactive to the patient's concerns. The therapist actively guides the patient to cover material that is relevant to the treatment goals. If the patient does not bring in material, the therapist may elicit an update or more detailed information in one of the agreed-on problem areas and goals. The need for the patient to change is stressed at regular intervals. It is important to note that this constitutes a general encouragement to change rather than pressure to take a specific course of action. Meetings during the intermediate phase should not pass without reference to the interpersonal problem areas and attendant goals. Phrases such as "moving forward on your goals" and "making important changes"

are used to encourage patients to be responsible for their treatment while reminding them that altering interpersonal patterns requires attention and persistence. In sessions, unfocused conversations are redirected to central themes of treatment, and abstract and vague discussions are minimized to maintain focus. Therapists avoid asking questions that elicit general or passive responses, such as general inquiries about the patient's week. Instead, sessions open with questions such as "What would you like to work on today?" and "How have things been since we last met?" These questions keep the patients focused on recent interpersonal events and psychiatric symptoms, which the therapist attempts to link. Above all unfocused, unproductive complaining about others should be discouraged. In which case it should be pointed out that neither the patient nor the therapist is likely to change someone else in the patient's life.

A critical goal in the intermediate phase is to facilitate and help to strengthen the patient's connections between their psychiatric symptoms and the difficulties they have in their interpersonal lives. As the patient continues making these connections and developing strategies to alter the interpersonal context in which his or her symptoms occur, the patient will disrupt the cycle. The patient is encouraged to explore the problem areas and consider ways of changing; the attempts to change then become the focus of subsequent sessions. The therapist helps the patient to remain focused by ensuring that the subject matter is relevant and by providing clarification when needed. Clarification takes the form of pointing out themes and inconsistencies or highlighting points that the patient might miss. Clarification does not extend to making interpretations, in which reference is made to a theoretical view on the disorder and its treatment. Throughout, the focus remains on the present.

The behavior of the patient toward the therapist is expected sooner or later to reveal the patient's well-practiced behavioral patterns toward significant others. These may become more obvious during the intermittent phase. It may be useful for the therapist to comment on them.

By the end of the intermediate phase, patients are often acutely aware that treatment will soon be ending. The therapist begins to discuss termination explicitly and addresses any anxiety that the patient may be experiencing regarding it. The therapist also begins to prepare the patient for emotions that may arise with termination, including grief related to the ending of treatment.

Termination Phase. In the termination phase (usually four to five sessions in length), the therapist discusses termination explicitly with the patient and assists him or her in understanding that the end of the treatment is a potential time of grief. During this phase, patients are encouraged to describe specific changes in their psychiatric symptoms, especially as they relate to improvements in the identified problem area(s).

The therapist also assists the patient in evaluating and consolidating gains, detailing plans for maintaining improvements in the identified interpersonal problem area(s), and outlining remaining work for the patient to continue on his or her own. Patients are also encouraged to identify early warning signs of symptom recurrence and identify plans of action.

DISCUSS TERMINATION EXPLICITLY. The therapist systematically raises the issue of reactions to impending termination in each of the last several sessions. Introducing the idea that termination is an explicit stage will plant the seed that it is an important topic to discuss. Therapists always ask patients how they feel about the ending of treatment—not least because this provides an opportunity to emphasize what has been achieved and to stress the patient's competence at dealing with future areas of difficulty.

DISCUSS TERMINATION AS A POTENTIAL TIME OF GRIEF. Patient reactions to the conclusion of treatments are often varied. As therapy nears completion, and sometimes well before that, patients may develop anxiety about saying goodbye and going it alone. Because termination marks the end of a connection to the therapist, is has a theme of loss, an analog of grief. It is important to state this possibility explicitly because unacknowledged sad feelings may lead to fears of relapse and an increase in symptoms.

REVIEW PROGRESS. An important aspect of the termination phase is encouraging patients to talk about the progress they have made and giving feedback about the changes that have been seen. This review helps to consolidate the work that has been done. It is not uncommon for individuals to attribute changes in treatment to the therapist rather than to their own hard work. Misplaced credit may erode the patient's confidence in the ability for continued success and improvement without treatment. Therefore, the therapist needs to emphasize how the patient has begun successfully to manage his or her outside relationships.

ENCOURAGE MAINTENANCE OF THERAPEUTIC GAINS. Patients are encouraged to maintain the interpersonal behaviors that have led to changes in their psychiatric symptom profile, and therapists work to develop a plan for continued work on identified goals. However, there are always goals that are not accomplished within the time frame of interpersonal psychotherapy, and patients need to be aware of this. Indeed, future difficulties can be expected, and the therapist needs to cultivate a discussion of how to handle them. Patients should also be encouraged to view their psychiatric disorder as a vulnerability or an "Achilles heel," in the sense that it may recur at future times of difficulty. Guiding the patient in a discussion of strategies for managing future problem will bolster feelings of competence. It is vital to assist patients in thinking about

warning signs and symptoms that suggest a need for intervention.

General Therapeutic Techniques

Throughout therapy, the therapist should maintain a consistent focus on the interpersonal context of the patient's life. Symptoms should thus be analyzed in light of the interpersonal difficulties constituting the backdrop of those symptoms. Although this approach is unique to the interpersonal psychotherapies, some of the therapeutic techniques used in interpersonal psychotherapy are similar to techniques used in other therapies. These include exploratory questions, encouragement of affect, clarification, communication analysis, and the use of the therapeutic relationship.

Exploratory Questions. The use of general, open-ended questions can facilitate the free discussion of material. This is particularly true in the initial phases of a session. For example, the therapist might say, "Tell me about your relationship with your husband." Once this has generated discussion, progressively more specific questioning will ensue.

Encouraging Affect. Interpersonal psychotherapy focuses on affect evocation and exploration. Specifically, the therapist helps patients to (1) acknowledge and accept painful affects—as the feelings are expressed, it is important for the therapist to validate and help the patient to accept them; (2) use affective experiences to bring about desired interpersonal changes—the goal is to help the patient to act more constructively in interpersonal relationships, and this may involve learning when his or her needs are met by expressing affect and when they are better met by suppressing affect; and (3) experience suppressed affects—frequently patients are emotionally constricted in situations in which others would typically experience strong emotions. For example, sometimes patients will deny feeling upset when it is clear that an upsetting interaction has just occurred. In this situation, the therapist might say, "Although you said you were not upset, it appears to me that you have shut down since mentioning the situation with your husband." By explicitly noting the discrepancy, the therapist attempts to draw out the affect that has been suppressed.

Clarification. This technique is useful in calling attention to contradictions that may have occurred in the patient's presentation of material and increases the patient's awareness about what he or she has actually communicated. For example, contradictions between the patient's affect and speech may be noteworthy (e.g., "While you were telling me how upset you are about your father, you had a smile on your face. What do you think that's about?").

Communication Analysis. This technique is used to identify any

communication difficulties that the patient may be experiencing and to help the patient change ineffective communication patterns. Typically, the therapist will ask the patient to recall in great detail an important recent interaction or argument with a significant other. Together, patient and therapist work to identify any communication difficulties and find more effective communication strategies.

Use of Therapeutic Relationship. The premise behind this technique is that people have characteristic patterns of interacting with others. Use of the therapeutic relationship in IPT is less routine. If appropriate, however, the patient's thoughts, feelings, expectations, and behavior in therapeutic relationship are explored and related to the patient's characteristic way of behaving or feeling in other relationships. In IPT it is particularly relevant and useful for patients with interpersonal deficits (the patient develops a relationship with the therapist as a model for other relationships) and interpersonal role disputes (the patient receives feedback on how he or she comes across and thereby has the opportunity to understand the nature of his or her difficulties in relating with others).

INTERPERSONAL PSYCHOTHERAPY DELIVERED IN A GROUP FORMAT

A relatively recent approach in the ongoing development of IPT has been its use in a group format. Interest in this area emerged from the successful adaptation by Denise Wilfley and colleagues of two individually based treatments—cognitive-behavioral therapy and interpersonal psychotherapy—to a group format for the treatment of overweight individuals with binge-eating disorder. Subsequently, Wilfley and colleagues published a book that details the procedures for implementing interpersonal psychotherapy in a group format. Numerous investigators are using group interpersonal psychotherapy to treat other patient populations and psychiatric disorders.

Interpersonal psychotherapy delivered in a group format has many potential benefits in comparison to individual treatment. For example, a group format in which membership is based on diagnostic similarity (e.g., depression, social phobia, eating disorders) can help to alleviate a patient's concerns that he or she is the only one with a particular psychiatric disorder while offering a social environment for patients who have become isolated, withdrawn, or disconnected from others. Given the number and different types of interpersonal interactions in a group setting, the interpersonal skills that are developed may be more readily transferable to the patient's outside social life than are the relationship patterns that are addressed in a one-on-one setting. Moreover, a group modality has therapeutic features not present in individual psychotherapy (e.g., interpersonal learning). The group format also facilitates the identification of problems common to many patients and provides a cost-effective alternative to individual treatment.

Techniques

Timeline and Structure of Treatment. The typical course of group IPT lasts 20 sessions over a 5-month period. It is recommended that group size range from six to nine members, with one or two group leaders, depending on resources and training needs. In contrast to the initial phase of individual interpersonal psychotherapy, in group interpersonal psychotherapy, the assignment of the sick role, the interpersonal inventory, the development of the problem area(s), and the development of treatment goals occur in a structured, individual, 2-hour pre-group meeting. The initial phase (sessions 1 through 5) of group IPT is then devoted to the development of positive group norms as group members test out and refine their individual target problem area(s). The intermediate phase (sessions 6 through 15) of group IPT is dedicated to group members' work on their identified problem area(s). At the midpoint of the intermediate phase, a 1-hour individual meeting is held between therapist(s) and each group member. The termination phase (sessions 16 through 20) of group IPT is devoted to the consolidation of gains made by group members over the course of treatment and the preparation of the group members for termination of the group and for future work on their own. Within one week of the final session of the termination phase, a 1-hour post-group meeting is held between therapist(s) and each group member. The three individual meetings (pre-, mid-, and post-group), sequenced to correspond with critical time points in the three phases of IPT, in combination with other techniques, were designed to maintain the exclusive and strategic focus on individual patients' interpersonal problem areas—the hallmark of IPT.

Therapeutic Stance. Similar to individual IPT, the therapeutic stance of the group interpersonal psychotherapist is one of warmth, support, and empathy. The therapist is active and advocates for the group members rather than remains neutral. The open-ended group approach used in group interpersonal psychotherapy, with its emphasis on fostering direct communication between group members, is also highly compatible with the therapeutic stance, goals, and structure of individual interpersonal psychotherapy. The approach provides a framework for facilitating interaction among group members and emphasizes the establishment of a safe environment in which patients can receive feedback on how they are perceived by others and can experiment with new and risky interactional styles. Successful implementation of an open-ended group approach requires specific training for the therapist and attention to important developmental processes.

Pre-Group Meeting. The pretreatment meeting is crucial for facilitating a patient's individualized work in the first phase of group IPT. In individual

IPT, the first five sessions are devoted to a detailed examination of the patient's interpersonal history (i.e., an interpersonal inventory) and the formulation of problem areas and goals, which will guide the therapeutic work. In these critical sessions, the patient is also oriented in a more general way to the work of psychotherapy and is educated about the nature of his or her disorder and the means by which the work is used to bring about a recovery. Although it is possible to undertake individual preparation in the presence of other group members, such as an approach would simply not be an efficient or therapeutically valuable use of group time. Thus, in group IPT, the 2-hour pretreatment meeting is used to focus on identifying interpersonal problem areas, establishing an explicit treatment contract to work on problem areas, and preparing patients for group treatment. After identifying a patient's interpersonal problem(s) (interpersonal deficits, role disputes, role transitions, or grief), the therapist works collaboratively with the patient to formulate concrete prescriptions for change, in addition to the specific steps the patient will take to improve social relationships and patterns of relating. These goals of treatment are expressed in language that is as specific and personally meaningful to the patient as possible. Before the start of the group, each group member is given a written summary of his or her goals and told that these goals will guide his or her work in the group.

Another important element of the pre-group meeting involves adequately preparing patients for group treatment. That is, patients are encouraged to think of the group as an "interpersonal laboratory" in which they can experiment with new approaches to handle challenging interpersonal situations. In this regard, patients are informed about the important interpersonal skills that are learned while participating in a group (e.g., interpersonal confrontation, honest communication, expression of feelings) and are encouraged to learn from others as they see changes occur. The therapist stresses to patients the importance of keeping their work in the group focused on changing their current interpersonal situations or intensifying important existing relationships and not using the group as a substitute social network.

Initial Phase. The first five sessions of group treatment comprise the initial phase in group interpersonal psychotherapy. During this phase, the therapist works to cultivate positive group norms and group cohesion while emphasizing the commonality of symptoms among members and how they will be addressed in the group context. During this phase, group members are encouraged to review their goals with the group and begin to make some initial changes in their respective interpersonal problem areas. As members begin to experiment with the changes outlined in their goals, the therapist works collaboratively with each group member to refine and make any alterations in the target areas before the beginning of the

intermediate phase.

Intermediate Phase. During the intermediate "work" phase of group IPT (sessions 6 through 15), the therapist works to facilitate connections among members as they share the work on their goals with one another. In contrast to open-ended group approaches, the group interpersonal psychotherapist is much less likely to focus on intragroup processes and relationships unless they are specific to the work on the member's interpersonal problem area (e.g., interpersonal deficits). However, the therapist consistently and continuously encourages group members to practice newly acquired interpersonal skills both inside and, most important, outside of the group. As is the case with individual IPT, an essential task throughout the intermediate phase is to strengthen the connections that the group members make between difficulties in their interpersonal lives and their psychiatric problems. As with individual IPT during the intermediate phase of treatment, the interpersonal psychotherapist implements treatment strategies that are specific to the four problem areas as specified by Weissman and colleagues (Table 33.8–2).

MIDTREATMENT MEETING. The midtreatment meeting is held midway through the intermediate phase (usually between sessions 10 and 11). This meeting provides an opportunity to conduct a detailed review of each group member's progress on his or her individual problems and to refine interpersonal goals. The therapist(s) recontract with group members during this meeting as a means of outlining and emphasizing the work that remains, both inside and outside of the group, before the conclusion of treatment.

Termination Phase

In the termination phase (sessions 16 through 20), the therapist discusses termination explicitly with the group members and begins to help them to recognize that the end of treatment is a time of possible grief and loss. The therapist helps members to recognize their own progress and the progress made by other group members. During this phase, group members are encouraged to describe the specific changes in their psychiatric symptoms, especially as they relate to improvements in the identified problem area(s) and relationships. Although it is common for group members to want to keep meeting on their own or to have frequent reunions, group members are encouraged to use this phase of the group to formally say goodbye to one another and to the therapist(s). The therapist(s) also uses this time to encourage members to detail their plans for maintaining improvements in their identified interpersonal problem area(s) and to outline their remaining work. In the final session, a specific structure is established to

allow each group member to give feedback to other members.

POSTTREATMENT MEETING. The posttreatment meeting is scheduled within 1 week after the final group session. The therapist(s) use this final individual meeting to develop an individualized plan for each group member's continued work on his or her interpersonal problem area(s). The therapist(s) reviews the group experience and the changes the patient has made in his or her interpersonal problem area(s) and significant relationships.

Maintaining the Interpersonal Focus in Group Interpersonal Psychotherapy

Use of Group Summaries. In contrast to individual IPT, in which patients have the undivided attention of the therapist, in the early phases of group formation, patients may feel overlooked or confused about how others' interpersonal problems relate to their own. One way to address this issue is to provide weekly group summaries to group members. These four- to five-page summaries focus on the transactions that occurred during each session and their implications for each patient's recovery. Summaries include information about the group process, the common interpersonal problems of those in the group (e.g., difficulty expressing feelings, excessive caretaking of others to the detriment of meeting their own needs, conflict avoidance), and how to use the group most effectively (e.g., how to use feedback and learn from other members' work on problem areas). Wilfley and her colleagues used group summaries in their treatment study for binge-eating disorder. However, writing group summaries can be quite time intensive. Therefore, one alternative to the use of group summaries may be the use of self-help manuals. With this approach, all group members (assuming diagnostic similarity among the group membership) would receive the self-help manuals and be instructed to read them over the course of treatment. For example, Weissman developed self-help manuals for mastering depression that would be useful companions to an IPT group for depressed patients.

Use of Interventions That Are Informed by Group Stage Development Theory. As discussed earlier, maintaining the important individual focus in group interpersonal psychotherapy also requires attention to important developmental processes noted in other time-limited groups. The developmental sequence of a group merges naturally, but successful movement from one stage to another can be disrupted by an ineffectual leadership style and ill-informed intervention strategies. At the very least, a therapist unaware of group development may use interventions that lead to stagnation, and, at the worst, he or she may use strategies that are detrimental to an individual's progress. Therefore, to manage the group effectively while maintaining the interpersonal focus required in group

interpersonal psychotherapy, it is essential that therapists understand the stages of group development. Attending to these normative transitions in the group intensifies group cohesiveness, prevents premature dropouts, and assists patients in making otherwise difficult interpersonal transitions. In addition, therapists who can predict critical transition points in the group and make stage-appropriate interventions are more likely to facilitate the work of the group and its members (Table 33.8–3).

CLINICAL ISSUES

The following case example illustrates the individual treatment of a 51-year-old woman with binge-eating disorder. This vignette highlights the therapist's presentation to the patient of the interpersonal formulation linking the onset of her eating disorder to one of the four interpersonal problem areas. As discussed earlier, the therapist's ability to discern patterns in interpersonal relationships rapidly, connect events with the onset and maintenance of the disorder, and to formulate goals is crucial to the time-limited nature of interpersonal psychotherapy. In addition, this case illustrates several of the strategies and techniques used in treating the interpersonal problem area.

Indications

Indications for time-limited oriented interpersonal psychotherapy have been defined by several efficacy trials. IPT has been shown to be effective in the treatment of mood disorders and eating disorders (bulimia nervosa and binge-eating disorder). Interpersonal psychotherapy is also described by the American Psychiatric Association's (APA) *Practice Guidelines for Major Depressive Disorder in Adults* as useful for patients who face conflicts with significant others or who are having difficulty adjusting to a life transition.

Table 33.8–3.
Linking the Phases of Interpersonal Psychotherapy to the Stages of Group Development

Phases/Tasks	Group Stages	Members' Work	Therapist Interventions
Initial: Sessions 1–5; identify problem areas	Engagement: Sessions 1–2	Members look for structure as they grapple with the anxiety of being in a group and sharing their problems	Establish a structure that encourages appropriate self-disclosure; facilitate norms for effective communication
	Differentiations: Sessions 3–5	Members work to manage negative feelings over interpersonal differences as they	Help members to understand their reactions in the context of interpersonal differences in their outside social lives

		emerge in the group	
Middle: Sessions 6–15; work on goals	Work: Sessions 6–15	Members work out differences and strive toward common goals	Facilitate connections among members as they share their work with each other; encourage practice of newly acquired interpersonal skills in and outside of the group
Final: Sessions 15–20; consolidate treatment	Terminations: Sessions 16–20	Members struggle with how to manage the impending loss of connection with other group members	Help members to consolidate their work and to plan continued work; assist members in grieving the loss of the group

Reprinted with permission from Wilfley DE, MacKenzie KR, Welch RR, Ayers VE, Weissman MM. *Interpersonal Psychotherapy for Group*. New York: Basic Books; 2000:20.

Limitations

Interpersonal psychotherapy has not been shown to be effective in treating substance abuse. The literature on the treatment of anxiety with IPT is evolving.

Complications

Interpersonal psychotherapy is a very straightforward treatment that is highly acceptable to patients. In fact, very few, if any, complications from interpersonal psychotherapy have been noted. However, after a full course of interpersonal psychotherapy, treatment nonresponders should be evaluated for the possible prescription of an alternative, evidence-based treatment.

Contraindications

Interpersonal psychotherapy was not developed to treat psychotic depression or other forms of psychosis, such as schizophrenia.

Goals of Treatment

As indicated earlier, the goal of interpersonal psychotherapy is to reduce or eliminate psychiatric symptoms by improving the quality of the patient's current interpersonal relations and social functioning.

ETHICAL ISSUES

To maintain high ethical standards, therapists wishing to conduct IPT should seek training and supervision to develop proficiency. Specific IPT training guidelines are available in the comprehensive text by Weissman and colleagues, and specific group training guidelines are available in the group text by Wilfley and colleagues. In general, interpersonal psychotherapists need to be mindful of imposing their goals and values onto their patients (e.g., pushing a patient to end a relationship). As with

any psychological treatment, it is critical thoroughly to evaluate (medically and psychiatrically) all patients entering treatment to determine the best evidence-based treatment and most prudent course, including the range of treatment options.

RESEARCH AND EVALUATION

Outcome Research

IPT has been found efficacious for major depression and has also been successfully adapted to treat other types of mood and nonmood disorders.

IPT has demonstrated efficacy for the treatment of several disorders, including nonpsychotic major depression, recurrent depression, bipolar disorder, postpartum depression, dysthymic disorder, bulimia nervosa, and binge-eating disorder. In addition, a number of new research applications are under investigation for various disorders, some of which include traumatic grief, social phobia, body dysmorphic disorder, chronic somatization, borderline personality disorder, and anorexia nervosa.

IPT has been adapted for various populations, including the elderly, posttraumatic stress disorder (PTSD) patients, adolescents, and patients with comorbid medical conditions (e.g., human immunodeficiency virus [HIV] infection and cancer). Interpersonal psychotherapy also has been modified for long-term treatment, groups, couples, telephone intervention, primary care, and adjunctive self-help, as well as been translated into several languages. Together, these findings indicate that IPT is not only an effective and promising treatment for a range of psychiatric disorders, but also has the potential for wide dissemination.

Research Using Group Interpersonal Psychotherapy

Group interpersonal psychotherapy has recently become a topic of various research interests. It has been used to treat depression by K. Roy MacKenzie and colleagues, postpartum depression by Claudia Klier and colleagues, and binge-eating disorder by Wilfley and colleagues. Studies are under way to look at group interpersonal psychotherapy for depressed adolescents, depressed older adults, and individuals with eating disorders, PTSD, and social phobia.

FUTURE DIRECTIONS

There are a number of important areas within the study and practice of interpersonal psychotherapy that warrant further consideration. First, although a number of outcome studies in interpersonal psychotherapy clearly document its efficacy, little is known about the mechanisms by which interpersonal psychotherapy exerts its effects. A greater understanding of the mechanisms by which change occurs in interpersonal psychotherapy would assist in further refinements of the treatment and

yield insights about the nature of the psychiatric syndrome under investigation. Second, increased efforts to improve the effectiveness of interpersonal psychotherapy are warranted. This might entail altering the structure of the treatment itself (e.g., timing, duration, frequency of the sessions) and identifying specific therapist behaviors and patient responses associated with better outcome. Third, refining the definitions of the target populations is advisable, which would define those for whom IPT is most indicated or the treatment of choice. Fourth, IPT was developed as a research intervention and remains to be well disseminated among clinicians. Consequently, there exists a need to translate interpersonal psychotherapy efficacy data to effectiveness studies and appropriate clinical practice.

REFERENCES

Agras WS, Walsh T, Fairburn CG, Wilson GT, Kraemer HC. A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. *Arch Gen Psychiatry*. 2000;57:459–466.

Anchin JC, Kiesler DJ. *Handbook of Interpersonal Psychotherapy*. New York: Pergamon Press; 1982.

Browne G, Steiner M, Roberts J, Gafni A, Byrne C. Sertraline and/or interpersonal psychotherapy for patients with dysthymic disorder in primary care: 6-month comparison with longitudinal 2-year follow-up of effectiveness and costs. *J Affect Disord*. 2002;68:317–330.

Brunstein-Klomek A, Zalsman G, Mufson L. Interpersonal psychotherapy for depressed adolescents (IPT-A). *Isr J Psychiatry Relat Sci*. 2007;44:40–46.

De Mello MF, Myczowish LM, Menezes PR. A randomized controlled trial comparing moclobemide and moclobedemide plus interpersonal psychotherapy in the treatment of dysthymic disorder. *J Psychother Prac Res*. 2001;10:117–123.

Dew MA, Reynolds CF 3rd, Mulsant B, Frank E, Houck PR. Initial recovery patterns may predict which maintenance therapies for depression will keep older adults well. *J Affect Disord*. 2001;65:155–166.

Frank JD. *Persuasion and Healing: A Comparative Study of Psychotherapy*. Baltimore, MD: Johns Hopkins University Press; 1973.

Frank E, Spanier C. Interpersonal psychotherapy for depression: Overview, clinical efficacy and future directions. *Clin Psychol Sci Pract*. 1995;2:349–369.

Frank E, Swartz HA, Kupfer DJ. Interpersonal and social rhythm therapy: Managing the chaos of bipolar disorder. *Biol Psychiatry*. 2000;48:593–604.

Heisel MJ, Talbot NL, King DH, Tu XM, Duberstein PR. Adapting interpersonal psychotherapy for older adults at risk for suicide. *Am J Geriatric Psychiatry*. 2015;23(1):87–98.

Kiesler DJ. *Contemporary Interpersonal Theory of Research: Personality, Psychopathology, and Psychotherapy*. Hoboken, NJ: Wiley; 1996.

Kiesler D, Strupp H. *The Process of Psychotherapy: Empirical Foundations and Systems of Analysis*. Edison, NJ: Aldine Transaction; 2006.

Klerman GL, Weissman MM, Rounsaville BJ, Chevron ES. *Interpersonal Psychotherapy of*

Depression: A Brief, Focused, Specific Strategy. Lanthan, MD: Jason Aronson; 1995.

Klier CM, Muzik VI, Rosenblum KL, Lenz G. Interpersonal psychotherapy adapted for the group setting in the treatment of postpartum depression. *J Psychother Pract Res*. 2001;10:124–131.

Lenze EJ, Dew MA, Mazumdar S, Begley AE, Cornes C. Combined pharmacotherapy and psychotherapy as maintenance treatment for late-life depression: effects on social adjustment. *Am J Psychiatry*. 2002;159:466–468.

Lillie R. Getting clients to hear: applying principles and techniques of Kiesler's interpersonal communication therapy to assessment feedback. *Psychol Psychother*. 2007;80:151–163.

Lipsitz JD, Markowitz JC. Mechanisms of change in interpersonal therapy (IPT). *Clin Psychol Rev*. 2013;33:1134–1147.

MacKenzie KR, Grabovac AD. Interpersonal psychotherapy group (interpersonal psychotherapy-G) for depression. *J Psychother Pract Res*. 2001;10:46–51.

Markowitz JC. *Interpersonal Psychotherapy for Dysthymic Disorder*. Washington, DC: American Psychiatric Press; 1998.

Markowitz JC. Interpersonal psychotherapy for chronic depression. *J Clin Psychol*. 2003;59:847–858.

Miller JD, Frank E, Cornes C, Houck PR, Reynolds CF 3rd. The value of maintenance interpersonal psychotherapy (IPT) in older adults with different IPT foci. *Am J Geriatr Psychiatry*. 2003;11:97–102.

Miniati CA, Callari A, Calugi S, et al. Interpersonal psychotherapy for postpartum depression: a systematic review. *Arch Womens Ment Health*. 2014;17:257–268.

Mufson L, Sills R. Interpersonal psychotherapy for depressed adolescents (IPT-A): an overview. *Nord J Psychiatry*. 2006;60:431–437.

O'Hara MW, Stuart S, Gorman LL, Wenzel A. Efficacy of interpersonal psychotherapy for postpartum depression. *Arch Gen Psychiatry*. 2000;57:1039–1045.

Orgeta V, Qazi A, Spector AE, Orrell M. Psychological treatments for depression and anxiety in dementia and mild cognitive impairment. *Cochrane Database Syst Rev*. 2014;(1):CD009125.

Parker G, Fletcher K. Treating depression with evidence-based psychotherapies: a critique of the evidence. *Acta Psychiatr Scand*. 2007;115:352–359.

Reynolds CF 3rd, Miller MD, Pasternak RE, Frank E, Perel JM. Treatment of bereavement-related major depressive episodes later in life: a controlled study of acute and continuation treatment with nortriptyline and interpersonal psychotherapy. *Am J Psychiatry*. 1999;156:202–208.

Scocco P, Toffol E. Group interpersonal psychiatry: a review. *Sante Ment Que*. 2008;33(2):105–131.

Spinelli MG, Endicott J. Controlled clinical trial of interpersonal psychotherapy versus parenting education program for depressed pregnant women. *Am J Psychiatry*. 2003;160:555–562.

Tanofsky-Kraff M, Wilfley DE, Young JF, Mufson L, Yanovski SZ. Preventing excessive weight gain in adolescents: interpersonal psychotherapy for binge eating. *Obesity (Silver Springs)*. 2007;15(6):1345–1355.

Weissman MM. Recent non-medication trials of interpersonal psychotherapy for depression. *Int J Neuropsychopharmacol*. 2007;10:117–122.

Wilfley DE, MacKenzie KR, Welch RR, Ayers VE, Weissman MM. *Interpersonal Psychotherapy*

for Group. New York: Basic Books; 2000.

Wilfley DE, Welch RR, Stein RI, Spurrell EB, Cohen LR. A randomized comparison of group cognitive-behavioral therapy and group interpersonal psychotherapy for the treatment of overweight individuals with binge-eating disorder. *Arch Gen Psychiatry*. 2002;59:713–721.

Zlotnick C, Johnson SL, Miller IW, Pearlstein T, Howard M. Postpartum depression in women receiving public assistance: pilot study of an interpersonal-therapy-oriented group intervention. *Am J Psychiatry*. 2001;158:638–640.

▲ 33.9 Dialectical Behavior Therapy

M. ZACHARY ROSENTHAL, PH.D., AND MARCUS A. RODRIGUEZ, M.A., M.S.

INTRODUCTION AND DEFINITION

DBT is a multi-modal, cognitive-behavioral treatment originally developed to treat women who meet criteria for BPD with a history of chronic self-harm and suicide attempts. BPD is a severe, costly, and difficult-to-treat psychiatric disorder characterized by problems across interpersonal, behavioral, cognitive, and emotional domains of functioning. Most individuals diagnosed with BPD report high rates of nonsuicidal self-injury (NSSI) and multiple suicide attempts. The lifetime mortality rates by suicide are estimated to be 50 times higher among BPD patients compared to the general population. Individuals with BPD require more mental health services and evidence higher rates of treatment dropout and relapse compared to patients with other psychological disorders. Moreover, BPD is associated with therapist burnout and is one of the most stigmatized psychological disorders. According to the American Psychiatric Association there is no single efficacious pharmacological intervention for BPD. However, current practice guidelines recommend psychotherapy as the primary treatment for BPD, and, to date, DBT is the psychosocial treatment that has received the most empirical support for patients with BPD.

DBT was developed in the context of the broader rise of behavioral and CBTs. Although behavioral and cognitive interventions are fundamental, treatment strategies in DBT, this is a therapy based on the amalgamation of principles from diverse sources, including interpersonal systems, theories, and scientific research related to emotional functioning, and Eastern mindfulness and Western contemplative spiritual practices. A biosocial model featuring the reciprocally interactive effects of nature and nature, along with a dialectical world view, together influence case conceptualization and are woven into the transactional process between patients and therapists. As a result, DBT offers a comprehensive treatment approach that uses a wide range of interventions to create a compassionate and hopeful therapy for patients who have been historically considered "manipulative," virtually untreatable, or even deliberately resistant.

Structurally, there are four primary modes of treatment in DBT: group skills training, individual therapy, telephone skills consultation, and a consultation team of DBT therapists. Preliminary data suggest DBT skills training may hold promise as a stand-alone treatment for emotion dysregulation among a transdiagnostic sample. However, there is currently insufficient empirical evidence to know with certainty whether all or only some of these modes are necessary. Until further dismantling and component analysis studies have been completed, it is inappropriate to consider a treatment as DBT unless all of the four primary modes are being implemented. Accordingly, this chapter focuses on DBT as originally developed for the treatment of BPD. However, it also highlights several areas where DBT has been adapted for specific populations and mental health problems.

HISTORY

DBT was developed by Marsha Linehan at the University of Washington in the beginning of late 1970s. For years she had used standard CBT strategies to work with suicidal and actively self-injurious patients. However, Linehan found that therapists' unrelenting focus on trying to help patients change their problematic behaviors often led patients to feel invalidated. In turn, patients responded by expressing anger, ambivalence about engaging in treatment, and even threats of self-harm. Therapists working with patients diagnosed with BPD tended to report higher levels of burn out, in part, because of the sheer volume of severe and highly complex problems their patients presented during and between sessions, and in part because they could not realistically address these all of these problems in session and still have time to teach patients new skills.

Accordingly, Linehan and her team developed a multifaceted approach aimed at helping these high-risk patients who suffer tremendously from chronic and pervasive problems across many areas of their lives to create "a life worth living." They adopted acceptance and validation strategies to balance the change-focused strategies of cognitive and behavioral therapies; they implemented skills training groups and phone coaching to ensure that patients learn new skills and generalize their learning into the naturalistic environment to cope with everyday stressors; and they added consultation team meetings in order to provide therapists with support, structured problem-solving, and a context for ongoing learning and assessment of fidelity to the treatment model.

The first randomized controlled trial (RCT) for DBT was published in 1991. This study demonstrated the superiority of DBT over treatment as usual for women with BPD. Subsequently, Linehan published the DBT treatment manual and skills training manual in 1993. In her book, Linehan introduced a novel conceptualization of BPD, reframing it as a dysfunction of the emotion regulation system that can be addressed with the different

components of DBT. When Linehan and her team published positive results for a second RCT in 1999, DBT became the first psychological treatment that could claim to be an evidence-based treatment for BPD, according to the criteria outlined by the APA Division 12 Task Force.

In order to facilitate the dissemination of DBT and the training of DBT clinicians around the world, Dr. Linehan established a nonprofit organization, Behavioral Tech LLC, in 2004. Around the same time, Linehan and her team developed a coding scale for rating individual and skills group sessions for adherence to the DBT treatment manual. In order to enhance adherence and competence, the DBT-Linehan Board of Certification has recently begun to offer DBT certification for individual therapists. In 2014, the revised skill training manual was published, which features both a revision of many original DBT skills as well as the addition of select skills to be used with specific clinical problems (e.g., substance use).

As an empirically driven treatment it is expected to continue to evolve over time as a function of further research. To date, RCTs have been published from many investigators around the world examining applications of DBT for different problems (e.g., eating disorders, substance abuse problems, and depression with co-occurring personality disorders) and in different settings (e.g., outpatient private practices, inpatient psychiatric hospitals, and forensic settings). Indeed, in 2012, DBT was officially recognized by the Cochrane Review as the treatment of choice for BPD. In little over 20 years, DBT advanced from a novel behavioral therapy to a gold standard intervention for BPD.

THEORETICAL ISSUES

The basic elements of DBT are outlined here. For a comprehensive description of this treatment approach, interested readers are referred to the treatment manual and group skills training manual by Marsha Linehan. The following section includes an outline of the dialectical philosophy and biosocial theory that together bind the various treatment modes and strategies of DBT. Treatment techniques and a case example are then presented, followed by a discussion of ethical issues and empirical findings related to DBT.

Dialectics

For DBT, dialectics subsumes both a world view and a process of change. The dialectical world view refers to an ontological perspective of interrelatedness and wholeness applied to the understanding of human suffering. Accordingly, behavior is conceptualized contextually and systemically, across private (e.g., thoughts, feelings, physiological sensations) and publicly observable (e.g., overt motoric behavior)

experiences. The patient as a whole is treated rather than a discrete disease, disorder, or symptom. Dialectical thinking provides a solution to the tension between formalistic universalistic thinking (there is absolute Truth) and relativistic thinking (there are many "truths"). Dialectics adds a third way of thinking; there is both absolute truth and truth is contextual and always evolving. A core feature of dialectical thinking is understanding that for every point (a thesis) an opposite position (an antithesis) can be held while searching for a synthesis between these apparent contradictions.

Throughout the course of DBT, polarities are repeatedly identified in how the patient is thinking or behaving. For example, patients might express anxious thoughts in catastrophic but unrealistic terms or, at an overt behavioral level, might repeatedly try to reduce guilt by excessively apologizing. Through consideration of opposing or contradictory ways of thinking (e.g., the catastrophe will never happen) or behaving (e.g., perhaps no apology is necessary), natural tensions arise between two possible ways of thinking or behaving. The dialectical process of change promotes the resolution of apparently contradictory stances by yielding an integrated alternative, or a synthesis, that takes into account the fluidity of a constantly dynamic context. The catastrophe is highly unlikely to occur but could still happen; apologies are necessary in some situations but not all. As another example, a patient might insist that expressing anger in a certain way helps to improve mood in the moment, whereas the therapist might insist that this anger expression actually is the problem, resulting in a polarized stalemate. Dialectical thinking would help the therapist to create a synthesis with the patient by identifying how anger expression may be effective when done in some ways and in some contexts, and, that at the same time, anger expression in some contexts can exacerbate problems, not solve them. Working together, the therapist and the patient may look for ways to effectively express anger in some contexts and not express anger in others contexts to minimize the extent to which anger expression is an enduring problem. In other words, the synthesis would be to neither totally inhibit anger nor express anger, but inhibit or express, depending on the context. This sensitivity to context—seeking what is needed in any given moment and moving away from extreme or rigid ways of thinking or behaving—is a principle that lies within the core of interventions in DBT.

Not only does a dialectical approach reduce extreme ways of thinking or behaving, it also explicitly provides the therapist with a process designed to give up on always being right about every issue, without losing perspective. In addition, dialectics acknowledges that the patient influences the therapist while being influenced by the therapist. Dialectical strategies, because they can be implemented in a way that is not always perfectly predictable by the patient, also may keep the patient slightly off balance (e.g., lose the rigid conviction that his or her perception is the literal truth). This has advantages for patients learning new behaviors, in that the

conditions under which adaptive ways of thinking or behaving in session are broad and not restricted to a narrow topographical therapist behavioral repertoire. In other words, the dialectical process is designed to promote broad repertoires of patient behavior and cognitions. For patients with BPD, this means less rigid, extreme, or otherwise problematic behavior, emotions, and patterns of thinking. Of importance, this also means more flexible ways of thinking and behaving that are sensitive to coupling the patient's values and goals with the contingencies of the present context.

Although there are many dialectical tensions in DBT, the fundamental dialectic is that of *acceptance and change*. For the therapist, this entails balancing an acceptance of patients as they are in the moment with an explicit goal of meaningful change. This dialectic requires that patients and therapists accept the reality that painful emotions and thoughts exist (but not necessarily the literal truth of the content of thoughts) while simultaneously working to change dysfunctional and seemingly intractable problems. A perfectly balanced "middle-of-the-road" compromise between acceptance and change is not necessarily the goal. Instead, a synthesis of polarities may be more acceptance-based in one moment and more change focused in another, depending on the context and what is likely to be most effective. This notion is akin to the need, when playing golf, to factor in the speed and direction of the wind when teeing off. The goal is to get the ball to land on the fairway, not to hit it in the air over the exact middle of the fairway; thus, the ball is sometimes hit to the left or to the right, factoring in the intensity and direction of the wind in the moment before the ball is struck.

Biosocial Theory of BPD

If dialectical thinking is the philosophical heart of DBT, then the biosocial theory of BPD forms the backbone. In articulating this theoretical model, Linehan suggested that BPD is fundamentally a disorder of the emotion regulation system and results from continuous reciprocal transactions among an invalidating environment, emotional vulnerability, and emotion dysregulation (Fig. 33.9–1). Prototypical examples of environmental invalidation are childhood sexual or physical abuse. Childhood sexual abuse may be a key etiological feature of the disorder, with an estimated 40 to 70 percent of borderline individuals reporting a history of childhood sexual abuse. Childhood sexual abuse, however, does not appear to cause BPD. Indeed, meta-analytic data from 21 studies found only a moderate effect size (r = 0.28) in the association between BPD and childhood sexual abuse. As such, sexual abuse may be an important etiologic factor for some individuals with BPD, but the development and maintenance of BPD is complex and determined by a wide range of neurobiologically and socially mediated factors.

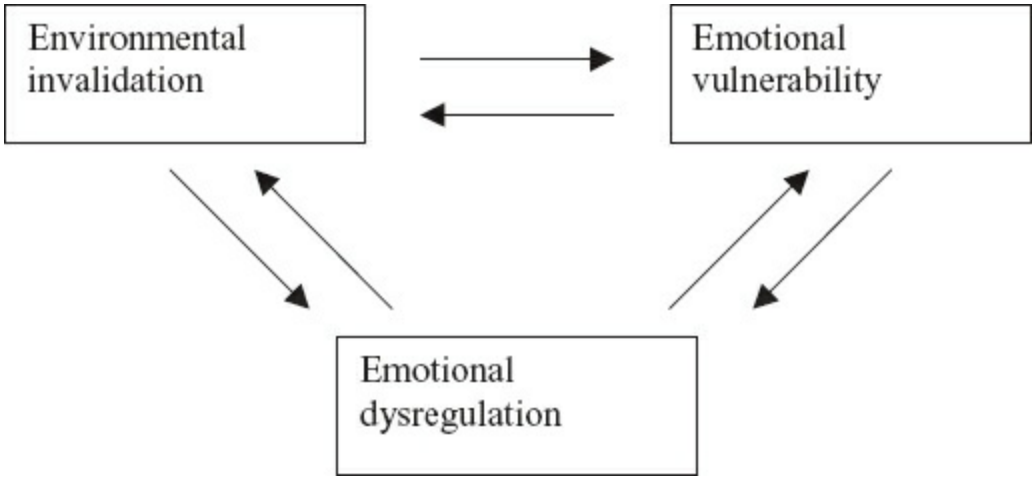

FIGURE 33.9–1. Components of the biosocial theory of borderline personality disorder.

The Invalidating Environment. According to Linehan's biosocial model, the emotion dysregulation characteristic of BPD commonly emerges among individuals with emotional vulnerabilities raised within (and often still living in) an invalidating environment. The invalidating environment is broadly characterized by pervasive criticizing, minimizing, trivializing, punishing, or erratically reinforcing communication of internal experiences (e.g., thoughts and emotions), and oversimplifying the ease of problem solving. For example, a parent may pervasively communicate "You're not that upset, you're just faking it" or "You say no, but you don't really know what you want." Verbal communication is indiscriminately rejected, and the individual may be chronically pathologized as having socially undesirable personality traits (e.g., too sensitive, paranoid, or lazy). This leads to a lack of trust in self, such that those with BPD learn that what they think, feel, and do is inappropriate, untrue, unimportant, or due to shameful personal flaws. Instead, an excessive reliance on the social environment for cues on how to think, feel, or behave develops, resulting in a hypersensitivity to interpersonal cues and often impairment with the interpretation of such cues. In addition, appropriate emotional displays are pervasively punished, and extreme emotional displays are intermittently reinforced, increasing the likelihood of emotionally escalated behavior. For example, individuals with BPD may not be taken seriously at home (e.g., dismissed, criticized, etc.) until they threaten suicide or self-injury, after which they may receive nurturance, respect, soothing, and the like, which may maintain or increase the likelihood of future self-injurious behavior in similar contexts.

Emotional Vulnerability. In addition to environmental invalidation, the biosocial model predicts that the emotion dysregulation problems characteristic of BPD are exacerbated by emotional vulnerability. Emotional vulnerability refers to a biologically mediated predisposition for

affective instability involving genetic, intrauterine, and temperamental factors. As a construct, emotional vulnerability is defined as heightened sensitivity, heightened reactivity, and slow return to a baseline level of emotional arousal. That is, individuals with BPD respond quickly to a wide range of contextual emotional stimuli, respond with a high magnitude of physiological arousal, and take a long time before arousal decreases to baseline. Once emotionally aroused, the slow return to baseline makes it increasingly likely that subsequent events will evoke extreme emotional responses. Similar to the searing physical pain a burn victim feels when the wind gusts, emotionally vulnerable individuals feel lasting emotional pain in response to seemingly ordinary daily events. Parents with emotionally vulnerable children are commonly frustrated and have difficulty managing their child's behavior. However, this may be particularly true for parents with existing psychopathology or significant deficits in parenting skills. In addition, high emotional vulnerability may make an already invalidating environment even worse, with a mismatch between child and parent emotionality contributing to additional invalidation, even if inadvertent. According to the model, in a reciprocal fashion, such increased invalidation may further exacerbate emotional vulnerability.

Emotion Dysregulation. In the context of environmental invalidation and emotional vulnerability, the biosocial model states that emotion dysregulation occurs across behavioral/motoric, physiological, and cognitive/experiential emotional systems. In a transactional, back-and-forth manner, problems with emotion dysregulation can aggravate emotional vulnerability and intensify environmental invalidation, and these problems can then further exacerbate emotion dysregulation. As shown in Table 33.9–1, the biosocial model asserts that BPD criterion behaviors stemming from emotion dysregulation can be organized across domains of functioning (i.e., emotional, behavioral, cognitive, interpersonal, and self). Processes of dysregulation thought to mediate borderline symptoms include problems with up- or down-regulation of physiological arousal, inhibition of mood-dependent behavior, reliance on avoidance strategies, attentional dyscontrol, processing of emotional information, self-soothing, and self-validation. For example, an inability to readily up- or down-regulate physiological arousal may lead to the development of extreme behavioral dyscontrol, such as self-injurious, impulsive, and aggressive behavior. Extreme emotional arousal may narrow attention, making it difficult to turn attention away from emotionally evocative stimuli, thereby biasing interpretations of events in a fashion consistent with mood state. For DBT therapists, the biosocial model provides a framework for understanding how a variety of topographically distinct problems (e.g., self-injury, substance use, anger problems, dissociation) may be actively targeted for change through

interventions aimed at similar mechanisms. Of importance, it also provides a nonpejorative, compassionate, and accepting approach sometimes difficult to adopt when working with BPD patients.

Table 33.9–1.
Diagnostic Criteria for Borderline Personality Disorder

Emotion dysregulation
Affective instability
Problems with anger
Behavioral dysregulation
Impulsive behavior
Self-injurious behavior
Cognitive dysregulation
Dissociation
Paranoia
Interpersonal dysregulation
Chaotic relationships
Fears of abandonment
Self-dysregulation
Identity disturbances
Chronic feelings of emptiness

TECHNIQUES

DBT is a principle-driven, flexible, and comprehensive treatment for multidiagnostic patients. As a behavioral therapy, it is quite change focused and driven by behavioral principles and models of classical and operant conditioning. Yet, at the same time, DBT incorporates numerous acceptance-based strategies to implement when changing behavior may not be possible or effective.

There are four primary modes of treatment in DBT: Group skill training, individual therapy, phone consultation, and consultation team. Other, ancillary treatments that are typically uncontrolled in DBT are pharmacotherapy and acute inpatient psychiatric hospitalization. To practice DBT, it is necessary to include all standard modes of treatment into a treatment program. Patients must be seen by a DBT-trained therapist, and therapists must participate in a DBT consultation team. The following section briefly reviews the function, process, and structure of DBT treatment modes.

Skill Training

Skill training occurs in the group format and is a skill class rather than a traditional interpersonal process-oriented psychotherapy group. The primary function of the group is the acquisition of new coping skills. The assumption is that these patients must learn specific behavioral, emotional, cognitive, and interpersonal skills not learned earlier in life. On the other

hand, it is also assumed that group members are not fragile, and that emotional arousal during group provides natural opportunities for skills practice.

Groups occur on a weekly basis, typically last for two hours, and usually include between four and ten patients, as well as two DBT-trained co-leaders. The first hour is usually devoted to a brief mindfulness practice and homework review from skills learned the previous week. During homework review, group members take turns discussing their successes and challenges in implementing skills learned the previous week. Group participation is strongly encouraged, but observations about others in the group and traditional "process" therapy are actively discouraged. During the second hour of group, new skills are taught didactically from the skills training manual. This workbook provides a wide variety of handouts and worksheets designed to decrease emotion dysregulation, including separate skill modules for distress tolerance, emotion regulation, and interpersonal effectiveness, as well as the core skills of mindfulness that are taught recurrently after each skills module.

Skills Modules: Mindfulness

Mindfulness is considered the core set of skills in DBT. Derived from Eastern meditative and Christian contemplative traditions, mindfulness is the practice of paying attention in a particular way: on purpose, in the present moment, and without judgment. Patients learn that their behavior is a function of current emotions (emotion mind) or logical analysis (reasonable mind). "Wise mind" knowing is emphasized as an integration of emotion and reasonable minds. The concept of wise mind relates to a synthesis between emotional and logical thinking, such that coping decisions are both effective and remain within personal values but do not need to be free of all of emotion. For instance, this can translate into suspension of short-term behaviors that have harmful consequences (e.g., drinking) for long-term gains in self-esteem or feelings of mastery, even if negative emotions are elicited. Wise mind can also be experienced as spiritual or "knowing from the gut" that a decision or action makes sense given the context.

Mindfulness skills also include the ability to observe, describe, and participate fully in one's actions in a nonjudgmental, one-mindful, and effective manner. Observing refers to noticing experiences without becoming attached, allowing thoughts or other internal experiences to flow freely with full awareness. Describing follows observing and involves labeling or putting words on experiences. Participating is somewhat different and involves behaving effectively without necessarily observing and describing internal experiences (e.g., an athlete at peak performance). Being nonjudgmental, including being aware of judgments and letting go of their literal truth, is a central skill repeatedly practiced by DBT patients and

therapists. Being one-mindful entails a sharpening of attentional focus on one thing or activity at a time. This skill involves staying in the present moment and not becoming distracted by events in the past or concerns in the future. Finally, the focus on effectiveness is a key aspect of mindfulness. Effectiveness refers to behaving in a way that is consistent with one's values and long-term goals, which often translates to not immediately reducing distress despite adverse consequences. The emphasis on effectiveness as a DBT skill illustrates how mindfulness is a behavioral, psychological, and spiritual practice—not purely a meditational pursuit.

Skills Modules: Distress Tolerance

Patients with BPD frequently experience extreme emotional distress. Many symptoms may develop as a means of reducing chronic and intense distress, albeit in maladaptive fashion. For example, self-injurious behavior may function temporarily to reduce distressing emotional states. Beyond overtly self-injurious behavior, a wide variety of topographically distinct behaviors may function abruptly and dramatically to reduce emotional distress, such as impulsivity, risk taking, anger, and dissociation. The distress tolerance module is designed to teach patients how to tolerate aversive emotional experiences without behaving maladaptively. A list of crisis management skills is taught, including strategies for effective temporary distraction, such as activities, eliciting opposite emotions, and squeezing ice or a rubber ball. Self-soothing skills are introduced, by which patients learn intentionally to soothe themselves during periods of crisis with calming visual stimuli, sounds, smells, tastes, and objects to touch. In addition, other skills such as imaginal and relaxation exercises are taught to improve the current moment in order to avert doing something that would make a crisis even worse. Other distress tolerance skills include awareness, breathing, and half-smile exercises, as well as radical acceptance of reality in the current moment. Overall, distress tolerance skills are intended to interrupt and change habitual and problematic responses to acute emotional distress, allowing the opportunity for newer, more effective behaviors to emerge.

Skills Modules: Emotion Regulation

The emotion regulation skills module is designed to help patients to better understand their emotions, reduce emotional vulnerability, and decrease emotional suffering. Specific skills taught include increasing awareness of emotions, identifying and challenging distorted ways of thinking about emotions, learning how emotions are related to problem behaviors, accurately labeling emotions, understanding the functions of emotions, reducing emotional vulnerability, increasing pleasant emotions, and *acting opposite* to behavioral urges associated with emotions. Although all of

these skills are useful, the opposite-action skill is particularly helpful for most borderline patients because it can be applied in many contexts to change a variety of problematic behaviors.

The opposite-action skill provides an algorithm for knowing when to try to change an emotional response, in order to develop a new and more effective way of responding in the future. This includes first determining whether the emotion or the intensity of the emotion is justified or effective based on the current situation. Next, it is important to know the action urge of the emotion being experienced. Each emotion has its own urging component (e.g., anger—attack, fear—run, sadness—withdraw, guilt—repair, shame—hide). When the patient is experiencing an unjustified emotion that he or she wishes to change, the skill is to go opposite to the action urge of the emotion. For example, if a patient is feeling ashamed or guilty after telling a friend that the patient feels excited about doing a new activity, opposite-action skills would teach the patient to ask first whether his or her behavior was egregious according to the patient's value system. If so, then a repair (e.g., apology) would be in order and would function to lower justified guilt. However, if the shame or guilt was unjustified (i.e., the patient did no wrong), then the patient would be instructed not to repair but to repeat the behavior (i.e., expressing enthusiasm about new activities) multiple times. This leads to less guilt over time as the patient learns that giving opinions does not always result in negative outcomes.

Skills Modules: Interpersonal Effectiveness

Because chaotic interpersonal relationships are a hallmark characteristic of BPD and because interpersonal stressors are a common trigger preceding suicide attempts, the development of interpersonal skills is crucial in the treatment of BPD. This skill module teaches group members how to identify factors interfering with interpersonal effectiveness, challenge common cognitive distortions associated with interpersonal situations, and determine the appropriate level of intensity for making requests or saying no in a given situation. Patients are taught what to say and how to say what they say depending on their priority in an interpersonal situation. In addition, group members are strongly encouraged to practice these skills in a wide variety of new situations. Guidelines for being taken seriously, attending to relationships, and preserving self-respect are introduced, and patients are instructed to practice developing new interpersonal skills based on these guidelines.

Individual Therapy

Individual therapy sessions with a DBT-trained therapist are typically 50 to 60 minutes once per week, but this can vary, depending on whatever is needed for a given patient in a given week. A primary function of individual

therapy is to improve and maintain patient motivation. A validating environment is created in which patients are treated with compassion and acceptance. Within this context, the therapist and patient relentlessly pursue behavioral change. Skills learned during group sessions are woven into skill plans in anticipation of upcoming events. In addition, episodes of emotion dysregulation from the previous week are discussed in light of skills that could have been used. Factors interfering with progress in treatment are also discussed in individual therapy sessions, preventing problems that might impede the development of new skills and helping patients to remain engaged in treatment despite urges to dropout.

Because it may take a long time for a patient's arousal to return to baseline, care must be taken to structure sessions in a way that facilitates learning and does not lead to a depletion of attentional resources. Once they are highly aroused in session, it is common for patients with BPD to become so emotionally dysregulated that new learning is not likely to occur during the session. A mindfulness or relaxation exercise may be used toward the end of session to facilitate return to baseline arousal before the patient leaves.

Diary Cards

To monitor a variety of targets, a daily diary card is used. For example, patients rate their mood, monitor the frequency of self-injurious behavior and urges, and track other relevant targets. Figure 33.9–2 shows a prototypical diary card. The diary card is reviewed at the beginning of each session, and the therapy session is organized around targets evident on the diary card. Given the plethora of treatment targets and the possibility that patients will not remember salient events from the week, the diary card is instrumental in directing therapy sessions toward highly relevant targets.

Behavioral Analysis

To change dysfunctional behaviors, DBT uses a number of problem-solving strategies. Behavioral analysis is frequently used to identify problem behaviors and understand the relevant context in which these behaviors occur. In a nutshell, behavioral analysis involves an active, directive effort by the therapist to identify specific antecedents and consequences associated with the problem behavior being analyzed. A thorough elaboration of events before, during, and after problem behaviors facilitates the selection of appropriate treatment interventions. Based on a functional-analytic approach to behavioral assessment, the goal of behavioral analysis is prediction and control of functional classes of problem behavior rather than traditional diagnostic assessment of disease entities. In other words, in DBT, borderline symptoms are conceptualized as problem *behaviors*. These may be external, publicly observable behaviors such as self-

mutilation or impulsive aggression. Alternatively, problem behaviors often are internal, publicly unobservable experiences such as self-judgmental thoughts or urges to use substances. In sum, behavior refers to anything a person does or experiences (e.g., thoughts, emotions, sensations, and overt actions). The inclusion of internal experiences as targets is consistent with other contemporary behavioral therapies and is crucial to understanding the phenomenological experiences of the patient associated with treatment targets.

Because behavioral analysis involves explicating the links in a chain of events, it is often referred to as a chain analysis. During a chain analysis, the topography, intensity, duration, and frequency of the target problem all are discussed. As links in the chain are explored, the DBT therapist considers the role of classical and operant conditioning. Classically conditioned behaviors are under the control of an antecedent stimulus, and operantly conditioned behaviors are under the control of consequent events. For example, strong urges to self-injure may be classically conditioned to occur after arguing with a partner. On the other hand, self-injurious behavior may be operantly conditioned by the consequences that follow, such as less hostility or increased attention from a partner. Knowing the controlling variables (i.e., antecedent or consequent) of a problem behavior informs the choice of interventions. Strategies for changing antecedents include behavioral exposure (e.g., rehearsing arguments) and stimulus control (e.g., avoid partner). In addition, other behavioral principles used during chain analyses are positive and negative reinforcement, punishment, extinction, and shaping. Doing a chain analysis is like being a detective solving a crime. All of the events that lead to and follow the problem behavior are uncovered, revealing critical links in the chain amenable to change.

Dialectical Behavior Therapy Skills Diary Card	Homework:	Filled out in Session? Y N (Circle)	How often did you fill out this side? ___ Daily ____2–3x ___ 4–6x ____Once	Started: Date____/_____/_____

Circle Start Day	Highest Urge To:			Highest Rating for Each Day			Drugs/Medications								Behavior				
Day Of Week	Commit Suicide	Self Harm	Use Drugs	Emotion. Misery	Physical Misery	Joy	Alcohol		Illicit Drugs		Meds.As Prescribed	PRN/Over the Counter			Mind-ful	Slee p	Skill s	Opposite Action	
	0–5	0–5	0–5	0–5	0–5	0–5	#	What?	#	What?	Y/N	#.	What.		Y/N	Hrs #	0–7	Y/N	
MON																			
TUE																			
WED																			
THUR																			
FRI																			
SAT																			
SUN																			

Chain Analysis Notes	*USED 0 = Not thought about or used 1 = Thought about, not used, didn't want to 2 =Thought about, not used, wanted to 3 = Tried but couldn't use them	SKILLS: 4 = Tried, could do them but they didn't help 5 = Tried, could use them, helped 6 = Didn't try, used them, didn't help 7 = Didn't try, used them, helped
Med Changes/Other:		

Urge to:	Coming into Session (0–5)	Ability to Control	Coming into Session (0–5)
Quit Therapy		Emotions:	
Self-Harm		Action:	
Commit Suicide		Thoughts:	

FIGURE 33.9–2. A typical diary card.

Dysfunctional links in the chain are examined and replaced with more adaptive responses during a *solution analysis*. Typically, this is guided by three questions: (1) Can the patient change the circumstance? (2) Can the patient change his or her emotional reaction? (3) Can the patient better tolerate the pain associated with the problem? Together, the patient and the therapist collaboratively develop strategies to replace problematic links and then commit to using new solutions the next time the problem behavior emerges.

Skill Enhancement

Skills learned during group sessions are enhanced in individual therapy. It is often fruitful to have patients rehearse behavior in session. Behavioral rehearsal may occur in the form of covert rehearsing of challenges to distorted cognitions or cheerleading statements said silently. Rehearsal also includes role-playing interpersonal scenarios with an emphasis on skill use. Therapists provide reinforcement and coaching, and rehearsal occurs with repetition. Whenever effective behavior occurs in session, it is the DBT therapist's job to reinforce such behavior. Alternatively, ineffective behavior evoked during session is often ignored or punished. In this way, the therapist is actively involved in *contingency management*. Positive regard is applied contingently. For example, when the patient is behaving inappropriately (e.g., threatening self-injury), the therapist remains cool and matter of fact, whereas skillful behavior (e.g., committing to not self-injure and using skills) is greeted with therapist warmth.

Another method of facilitating skill use is to use behavioral exposure

and response prevention in session. Patients will often become angry, ashamed, or fearful in session, and a range of behavioral problems may emerge in response to these emotions. Patients feeling angry may lash out, whereas patients feeling ashamed may look down or dissociate, and patients feeling fearful may abruptly terminate the session. Behavioral exposure and response prevention applied to these emotions in session would target paying attention to these emotions nonjudgmentally, observing urges to behave ineffectively, and blocking these urges by not lashing out, keeping eye contact, or remaining in the therapy room. In addition to changing the behavioral pattern, maladaptive cognitions are modified through standard cognitive therapy techniques aimed at identifying and challenging unhelpful ways of thinking.

Validation

All efficacious therapies require verification of what the patient does effectively and disconfirmation of what is ineffective. Indeed, verification of self-constructs may reduce physiological arousal, and validation as an intervention may function to enhance patient learning by modulating arousal and enhancing the therapeutic relationship.

Table 33.9–2.
Levels of Validation in Dialectical Behavior Therapy

1. Listening and observing
2. Accurately reflecting patient experiences
3. Helping patients to articulate unverbalized emotions, thoughts, and patterns of behavior
4. Communicating an understanding of behavior as valid, given past learning history or biological vulnerability
5. Communicating an understanding of behavior as valid, given a current context or what is deemed normative
6. Providing radical genuineness, treating the individual as an equal and not as a sick and fragile patient

Although validation may mean many different things, in DBT validation is operationalized on several levels. These levels of validation are shown in Table 33.9–2 and are integral to doing DBT. In addition, validation may be explicitly verbal, or it may occur more implicitly and functionally, as when the therapist offers a tissue when an emotionally inhibited patient appears on the verge of tears. Validation may be used as pure acceptance, with no directed effort toward change. DBT therapists attempt to validate what is valid and usually ignore or if needed can invalidate what is invalid. For example, a therapist may say, “It makes sense that you felt angry in that situation, because your colleague lied about you, but cussing at her in the break room is not effective.” This requires the DBT therapist to discern carefully what is valid (i.e., feeling anger) and to apply validation thoughtfully. In addition to fostering a positive therapeutic relationship

and providing a compassionate environment, validation by the therapist functions as a model for patients and may be taught directly as a set of skills.

Dialectical Strategies

As mentioned, efficacious therapy and learning requires both confirmation and disconfirmation. Dialectical strategies are fundamentally based on this notion and revolve around behavioral acceptance (e.g., validation) and change (e.g., problem solving). Dialectical reasoning is pursued with the patient, by which the therapist helps the patient to move from a polarized position of "either-or" to a dialectical synthesis of "both-and." There are a number of specific dialectical strategies used with patients (Table 33.9–3). Of importance, to be effective these strategies must be used with therapist genuineness, not as gimmicky techniques. On the other hand, the therapist must be willing to let go of the truth or rightness of any dialectical strategy and instead must continually search for other ways to help patients change maladaptive patterns.

Telephone Consultation

To enhance the generalization of skills, patients in DBT may contact their individual therapists for telephone consultation between sessions. Because DBT patients may experience unrelenting crises, it is crucial that therapists observe personal limits associated with telephone consultation. Telephone calls are intended to be brief and usually last less than 10 minutes. Patients are oriented during pretreatment to the purpose of skills calls and are told to call when they are unable to implement skills in necessary situations but before crises occur. To reduce inadvertent reinforcement of self-injurious behavior using warmth and validation over the telephone, DBT uses a *24-hour rule*: Patients will not receive skills help from their individual therapists until 24 hours after any self-injurious behavior. Instead, early in treatment a contingency plan is created for occasions when patients self-injure, and DBT therapists may refer patients to this plan in the event that they call after self-injuring. In all consultation calls, DBT therapists assess for immediate danger and provide appropriate assistance when the patient is deemed to be in imminent danger of harming himself or herself or others.

Table 33.9–3.
Dialectical Strategies Used in Dialectical Behavior Therapy

1. Providing acceptance and change-focused interventions
2. Nurturing the patient and demanding that patients help themselves
3. Being stable and persistent, as well as flexible
4. Highlighting the patient's strengths and deficits

5. Structuring the session with an agenda, and responding to in-session patient behaviors as they occur
6. Highlighting both ends of continua, and making synthesizing statements
7. Pointing out paradoxes when present (e.g., patient's behavior, therapeutic process)
8. Using metaphors
9. Playing devil's advocate
10. Extending the seriousness or implications of patient's communication
11. Helping the patient to activate "wise mind"
12. Helping to make lemonade out of lemons
13. Allowing natural changes in therapy without attempting to control all outcomes

Consultation Team

As mentioned earlier, use of a consultation team is a necessary mode of treatment. Team members commit to weekly meetings and agree to a team structure and process (Table 33.9–4). In important ways, team members treat each other with DBT by providing validation, support, and motivational enhancement strategies. This support is invaluable and can help DBT therapists with a more balanced approach toward their patients. A consultation team also provides opportunities for fresh perspectives and new solutions, helping therapists to get unstuck and become hopeful. It is not uncommon for DBT therapists periodically to become rigid in their thinking and behavior with a patient. The consultation team offers problem solving and validation for the therapist, and team members actively use a dialectical process to help find effective syntheses between polarized positions. For example, the team can help remind the therapist to continue managing contingencies in session appropriately (e.g., not being warm in response to patient hostility). If possible, it is extremely helpful to videotape therapy sessions and watch important segments during the consultation team meeting because this engenders a full appreciation for the difficulty a therapist may be having and allows the team to ensure that all members are indeed adhering to the treatment.

Table 33.9–4.
Consultation Team Agreements in Dialectical Behavior Therapy

1. Meet weekly for 1–2 hours
2. Discuss cases according to the treatment hierarchy (i.e., self-injurious/life-threatening behavior, treatment-interfering behavior, and quality-of-life–interfering behavior)
3. Accept a dialectical philosophy
4. Consult with the patient on how to interact with other therapists, and do not tell other therapists how to interact with the patient
5. Do not expect consistency of therapists with one another (even across the same patient)
6. Allow all therapists observe own limits without fear of judgmental reactions from other consultation group members
7. Search for nonpejorative, empathic interpretation of patient's behavior
8. Acknowledge that all therapists are fallible

CLINICAL ISSUES

In DBT, problem behaviors are seen as the inevitable sequelae of dysregulated emotions or as maladaptive methods of altering emotional experiences. For example, self-injury may function to regulate dysregulated emotional responding. This theoretical stance has resulted in a treatment package that largely focuses on modifying various aspects of the patient's emotional system and treatment formulations are apt to focus on reducing ineffective action tendencies linked with dysregulated emotions. In DBT, the problem behavior is targeted as a priority, in an effort to break cycles of reinforcement (Fig. 33.9–3). The present section begins with an outline of the assumptions and agreements of DBT, followed by a discussion of the levels and targets of treatment. A brief case example has been included to illustrate some of the strategies in DBT.

Assumptions and Agreements

Most therapeutic approaches have underlying assumptions about what is expected from patients and from the treatment. In DBT, these assumptions are openly discussed with patients in the first few sessions, during what may be called pretreatment (Table 33.9–5). Patients are also oriented to the philosophy, theories, levels, and modes of treatment. Patient and therapist agreements for treatment are also discussed during pretreatment sessions (Table 33.9–6). In addition, it is important for the therapist to discuss the requirement that patients commit to participating in DBT. Indeed, commitment to treatment may wax and wane over time for many psychotherapy patients. Thus, the therapist providing DBT needs continually to be alert to changes in commitment by the patient and be willing to go back and work on commitment to treatment again and again if needed.

Levels of Treatment and Treatment Targets

There are four stages of DBT. During stage I, patients often present with severe behavioral dyscontrol and life-threatening problems, including self-injurious behavior such as cutting, burning, or otherwise deliberately harming oneself physically with or without the intent to die. Patients in this stage commonly meet criteria for multiple Axis I disorders as well as BPD. In addition, unrelenting crises that may vary from week to week in content, as well as a multitude of treatment-interfering behaviors, commonly occur for these individuals. To provide a framework for therapeutic targets, a hierarchy is used to guide the therapist toward specific goals in a given session: (1) reduce life-threatening and self-injurious behaviors; (2) reduce treatment-interfering behaviors, such as arriving late to session, not attending skills training group, lying, not completing homework (e.g., diary cards), arriving to therapy intoxicated, avoiding emotions during therapy, talking excessively about topics that are not germane to the problems being

targeted, etc.; and (3) reduce quality-of-life–interfering behaviors, such as drug use, dysfunctional eating behavior, anxiety, depression, and physical health problems.

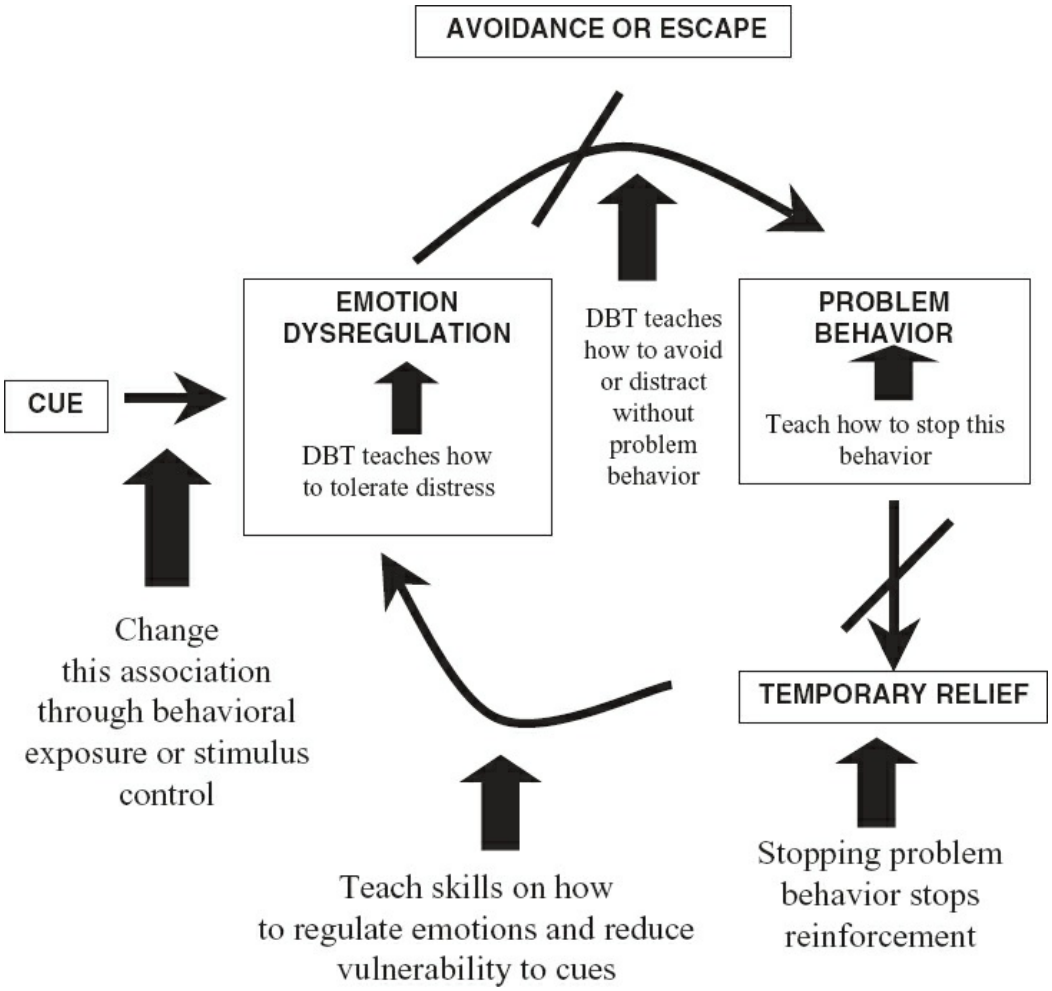

FIGURE 33.9–3. How dialectical behavior therapy (DBT) works.

Within and across therapy sessions, this hierarchy largely informs the selection of treatment targets. Often, treatment-interfering or quality-of-life–interfering problems occur in the context of life-threatening or self-injurious behavior. In these cases, it is crucial that the DBT therapist refer to the hierarchy, targeting the reduction of life-threatening/self-injurious behavior first, even if this is less desirable than talking about something else to the patient.

Table 33.9–5.
Patient and Therapy Assumptions in Dialectical Behavior Therapy (DBT)

Patients

1. Patients are doing the best they can
2. Patients want to improve
3. Patients need to do better, try harder, and be more motivated to change

4. Patients must solve their current problems, regardless of who caused these them
5. Patients are living lives that are unbearable as they are currently being lived
6. Patients must learn new behaviors in all relevant contexts

Therapy

1. Patients cannot fail in DBT, but the therapy or therapist can fail the patient
2. Helping patients work toward their ultimate goals in life is the most caring thing a therapist can do
3. DBT therapists need support
4. The therapeutic relationship is a relationship of two equals
5. Principles of behavior are universal, affecting both patients and therapists alike

Once skillful behaviors learned during treatment have replaced major life-threatening/self-injurious, treatment-interfering, and quality-of-life–interfering behaviors, patients begin stage II. During this stage of treatment, posttraumatic stress is often treated using empirically supported exposure therapy protocols while continuing to increase the patient's ability to experience aversive emotions skillfully. When posttraumatic stress is not present, longstanding difficulties with emotional experiencing and expression are targeted. For example, patients may need further to develop skills such as emotion labeling, reducing vulnerability to emotion dysregulation, emotional acceptance, or adaptive emotional change. These skills may have been learned already in stage I, but in stage II they are generalized more directly to contexts that do not involve acute behavioral dyscontrol. Patients in this stage of DBT may revert to behavioral dyscontrol characteristic of stage I. Thus, it is crucial that stage I treatment targets (e.g., severe behavioral dyscontrol) are closely monitored while patients are in stage II.

Table 33.9–6.
Patient and Therapist Agreements in Dialectical Behavior Therapy

Patient agreements

1. Stay in therapy for a specified period of time, usually 1 year
2. Attend all therapy sessions
3. Therapy will be discontinued if four consecutive sessions are missed
4. Work toward terminating self-injurious behavior and other therapeutic targets
5. Participate in skills training
6. Abide by relevant research conditions of therapy
7. Pay agreed-on fees for service

Therapist agreements

1. Maintain competence and effort
2. Provide ethical and professional treatment
3. Be available for weekly sessions and phone consultation
4. Treat patients humanely, with respect and integrity
5. Maintain confidentiality
6. Seek appropriate consultation

For those who persist in DBT beyond stage II, in stage III of DBT, the

therapist and patient target problems in living that are not debilitating but that interfere with experiencing *ordinary* happiness and unhappiness. This includes addressing unresolved problems with social and occupational functioning. Finally, stage IV targets a sense of completeness, spiritual growth, insight, enhanced awareness, sustained joy, and other fulfilling life endeavors. Although Linehan articulated these four stages of DBT, the first stage has been developed most thoroughly and is what has been tested and implemented for the most egregious and out-of-control behavioral problems among those with BPD.

In this transcript, the patient was a 28-year-old Caucasian woman presenting for psychotherapy with BPD and depression. In this prototypical example of DBT, the therapist conducted a behavioral analysis of self-injurious behavior, identified possible skills, and began to obtain a commitment to skills-based solutions.

T: So, let's think back to last week when you harmed yourself. What we're going to do is a behavioral analysis of that incident, because we really need to understand how that happened and how we can prevent it from happening again.

P: I don't really know why I cut myself last week. But I can tell you that I won't do it again this week. I really feel awful about this . . .

T: I get it. And I know it feels horrible to think and talk about this. We can do this together, and we need to, to keep us moving forward even when you are feeling down.

P: OK.

T: Now, saying that you will not cut yourself this next week is important. And, as you know, for this treatment to help you get out of the misery that you have been in for years, we have to figure out how to help you be committed with your actions throughout the week. You've heard me say it before and I'll say it again . . . commitment is as commitment does. So, let's work hard together to figure out what happened when you cut yourself last week. And then we'll figure out what to work on this week to keep you from harming yourself. We'll do our chain analysis together and see where it takes us, OK?

P: OK. So I had a bad day at work on Friday.

T: Did you cut yourself at work or was it after work, as it has been in the past?

P: After.

T: What happened at work?

P: Nothing really. Right before I left work on Friday, I tried to tell my manager about an idea I had to improve the level of comfort for the animals at the veterinary clinic where I work. But he cut me off before I finished explaining my idea, and said, "No. That's not a good idea. Why don't you just do your job and let me do mine?" It's the way he said it, and how he rolled his eyes.

T: I bet you felt angry and discouraged, and I can understand why.

P: Yeah, because he never listens to my opinions and he talks to me like I'm stupid.

T: And because I also know you really care about the animals.

P: Yeah, but the worst part is that he gets in my head, and I often wonder if he's right. Because honestly, no one really thinks I have good ideas, not even my fiancé. I think it's probably me (patient becomes visibly sad or ashamed).

T: It looks like you might be feeling a little sad right now. I'm noticing how you are looking down at the floor. Are you feeling sad right now?

P: I don't know what I feel . . . I think I just feel hopeless sometimes.

T: I wonder if you felt this way on Friday around the time you cut yourself?

P: Probably.

T: Do you think you would be willing to look up towards me as we continue?

P: Okay (looks up).

T: Excellent (leans in toward patient). Now, as we talk about when you cut yourself you feel terrible.

P: Like a loser. Like a hopeless loser.

T: So tell me, what happened after you left work?
P: I went home and told my fiancé about what happened. But instead of him comforting me and telling me what a jerk my manager is, he said my idea wasn't realistic and that I chose poor timing. Like always, he blamed me.
T: What were you thinking about in that moment?
P: My mind was blank at first, and then it started racing. I didn't know if I should scream at him or break down weeping. I felt totally overwhelmed.
T: What did you do next?
P: I held myself together at first. I gave my fiancé the silent treatment first, and then I excused myself to the bathroom. I tried to use the distress tolerance skill we talked about, the one where I put my face in cold water. But that didn't work. I still felt so empty inside.
T: What did you do right after that?
P: I tried to use a distraction skill, to play Solitaire on my phone.
T: Great. You were mindful and willing to use skills in that moment. That is a real progress!
P: Yeah, but then my cell phone didn't work for some reason. At that point I just thought forget it, you know, and got out my razors. I had enough and knew what I needed to do. So I just sat there and did it. The way I always have.
T: How did you feel after you started cutting?
P: Some pain at first, but then I felt mostly relief. My mind quieted down. I stopped thinking about everything that happened. It really distracted me. I told you, this sounds crazy, but cutting seems like, in your words "an effective skill."
T: That's one way to think of it. OK, so immediately after you stopped cutting, what happened next?
P: I did it for like three minutes then stopped. I felt ashamed, alone, and even more hopeless. I knew I couldn't call you for 24 hours, and I knew my fiancé wouldn't understand, so I ate like a gallon of ice cream and sat and binged on Netflix shows. I cut my self off from the world and just got away from everyone. I didn't want to deal with anything, you know? Plus I had to clean up the mess I made. No thanks. You know? It's just easier not to deal with any of it.
T: You are doing a great job filling in the details of this chain analysis. This is a familiar chain of events to times in the past when you have cut yourself. Let's think, what would have happened if you had taken a moment to use your skills earlier in the chain of events? For example, can you think of another skill or two you could have used before you hurt yourself?
P: I could have breathed in the car on my way home, and used my mantra.
T: Yes!
P: And I could have called you.
T: This is true too. I would love to have helped you before you cut yourself. Now, although we have the luxury of looking back in hindsight, we also have the luxury of looking to the future with our crystal ball. The good news and the bad news here is that you will have more opportunities to use these skills in similar situations in the future. And using these skills in the moment is really hard.
P: You're right. The hardest part sometimes is remembering to do it.
T: Let's talk about how we can do this effectively. We don't want to make things worse by biting off more than we can chew, but on the other hand we also want to recognize that you have the ability to hear invalidating comments without eventually harming yourself. Perhaps a good starting point would be for you to think about what you would be willing to commit to actually practicing, realistically.
P: Yes.
T: And what exactly would you practice doing in terms of your skills?
P: I don't know . . . I could practice mindfulness of my breath or checking the facts after a difficult conversation with my manager or fiancé.
T: Yes, good. But what would be most effective?
P: I guess checking the facts.
T: Outstanding! Are you willing to practice doing this right now?
P: Yes. I am.
T: Good, after we do that we can talk about the distress tolerance skills you said you used but didn't work last Friday. As we've talked about before, even if you use skills, you may want to harm yourself, since we know this is like an automatic response you have when you feel

hopeless or ashamed. So, you must, under all circumstances, commit to not harming yourself again, even if you have strong urges. If you do have strong urges, you can call me for skills help, but only before you harm yourself. If you harm yourself, you cannot call for 24 hours.

P: I know. And I am making a commitment not to harm myself again, and I will use these skills to cope with interpersonally challenging situations at work and at home.

T: Okay.

ETHICAL ISSUES

Most ethical issues faced by DBT therapists are not specific to the practice of DBT. The following section includes a discussion of ethical issues commonly faced by DBT therapists across the four modes of treatment. One such example could include a patient disclosing self-harm or suicidal ideation in the context of skills group where it is not possible to do a full risk assessment in front of other group members. These types of situations are not uncommon, which is one of the reasons why it is important to have two group leaders, so that one group leader can take the patient outside to assess risk and develop a safety plan. Furthermore, as discussed previously, the consultation team can also provide assistance to the therapist in managing these situations and documenting appropriately.

The *Four-Miss Rule,* whereby a patient is dropped out of therapy after missing four consecutive sessions of individual or group therapy, presents an ethical dilemma for DBT therapists. On the one hand, if a patient is not receiving therapy, then they are, in a matter-of-fact way of thinking about it, simply not in therapy. Similarly, irrespective of reasons for not taking a dose of medicine, if one does not take the dose, the medicine is not going to take effect. On the other hand, there may be many reasons why patients miss four consecutive weeks of individual or group therapy visits. Still, after missing four consecutive weeks, in DBT the patient is considered out of treatment. Ethically, at this point the therapist or clinic is responsible for the patient until they have been successfully referred to a provider in the community. As such, the interim period presents an ethical challenge. Therapists should therefore carefully discuss and frequently review with their patients the DBT treatment agreement and its rationale. The intent of the *Four-Miss Rule,* for instance, is to prevent patients from simply drifting out of therapy. It is also important for therapists to be proactive in communicating with and motivating patients to remain engaged in therapy if they notice a patient is approaching four missed sessions. If a patient is out of therapy due to the *Four-Miss Rule,* the treatment team needs to support the therapist in helping the patient find treatment elsewhere, or develop opportunities to overcome the problems that led to them missing treatment in the first place. After some time and evidence that the problems have been resolved, the patient can petition to the therapist, who can consult with the team, about the pros and cons of the patient being permitted back into DBT.

Another common set of ethical challenges in DBT occurs when therapists take telephone coaching calls to generalize skills in everyday contexts where the skills need to be used. Telephone coaching introduces unique concerns related to confidentiality and risk assessment. For example, sensitive personal information is at risk if a patient loses their phone or if someone overhears the content of a coaching call. Although the purpose of telephone coaching is skills generalization, therapists often need to assess the risk for self-injurious behaviors. Doing an assessment over the course of a brief conversation on the phone can cause complications. Although there are many possible solutions to this problem, ultimately the individual therapist leverages the individual sessions (e.g., by motivating patients to use skills and problem solving skills use) to prevent self-harm related crisis calls.

Another ethical issue occurs in the context of DBT team consultation, and is related to the dialectical tension between communicating with other treatment team members on behalf of a patient versus allowing the patient to speak to their provider directly. Ethical guidelines require consultation among treatment providers. Although it would often be easier to provide certain information about a patient to other members of the treatment team, one important aim of DBT is to help patients learn the skills necessary to get their needs met autonomously. Across each of the modes of DBT therapists must balance apparent conflicts between what is required ethically and what is indicated clinically.

RESEARCH AND EVALUATION

DBT has garnered sufficient empirical evidence for its efficacy in treating BPD, warranting designation as well established when utilizing criteria outlined by the Division 12 Task Force. Well-established according to these criteria means that a treatment must have demonstrated efficacy in at least two rigorous RCTs with superiority over placebo control conditions or another bona fide treatment. To date, more than 10 RCTs have been conducted across multiple independent research teams. DBT efficacy studies have been conducted in various countries, including the United States, Canada, Germany, the Netherlands, New Zealand, Spain, and Sweden. DBT has also been implemented and evaluated in inpatient, outpatient, school, forensic, military, and community settings. A comprehensive summary of DBT studies and adaptations is beyond the scope of this chapter; however, this section will review some recent research findings.

In a recent meta-analysis of 16 studies, 8 of which were RCTs across different research teams, DBT was associated most with reductions in suicidal and self-injurious behaviors. A more recent meta-analysis of five RCTs corroborated these findings. Although DBT was designed for individuals with BPD, research on standard DBT protocols has paved the

way for more studies examining adaptations for a variety of other populations. DBT has demonstrated efficacy in the treatment of substance-use disorders, eating disorders, chronic depression in older adults, and depression with co-occurring personality disorders. The use of DBT with substance use disorders leads to enhanced emotion regulation. DBT in the treatment of eating disorders can lead to decreased bingeing and purging behaviors. In depression, DBT has been found to lead to faster remission compared to medication alone. The effects of DBT may last after treatment ends. For example, at six-month follow up, in one recent study DBT was associated with reduced visits to the emergency room.

In the treatment of PTSD, studies have yielded promising results suggesting that DBT is effective in the treatment of women who are actively self-harming, victims of sexual assault or harassment in the military, and adults with a history of childhood sexual abuse. For instance, DBT was found to lead to reductions in depressive symptoms and hopelessness among female victims of domestic abuse.

Recent research suggests that DBT is also effective in treating patients characterized by dysfunctional eating behaviors. This includes adaptations of DBT used to treat binge eating (e.g., Glisenti & Strodl, 2012), as well those designed to address co-occurring BPD and anorexia or bulimia. Another adaptation of DBT, Radically Open-DBT, was associated in a pilot study with improvements in eating disorder related psychopathology symptoms among patients with anorexia nervosa-restrictive type.

DBT also has been adapted for children and adolescents. This development is important because the evidence suggests that DBT helps to reduce suicidal ideation and suicide risk behavior, and suicide is the third leading cause of death among youth ages 10 and 15. The first RCT of DBT for Adolescents (DBT-A) demonstrated greater reductions in self-injurious behaviors compared to enhanced care as usual. At the time this chapter was written, a second multi-site RCT of DBT for Adolescents was being conducted to examine the effectiveness of DBT at decreasing suicide attempts in at risk youth (R01 MH093898). Some of the ways DBT-A is different from standard DBT include the following: families are included in skills training, interventions directly target the invalidating environment (in family therapy), the length of treatment is abbreviated (from 1 year to 6 months), new skills are added (e.g., the “Middle Path”), and skills examples more relevant to teens are included in the treatment manual by Rathus and Miller (2014). In school contexts, DBT-A demonstrated promising results (e.g., reduction in number of referrals to the assistant principal, detention/suspensions, and visits to the ER) in an open trial study conducted at a high school in New York. In addition to DBT for Adolescents, DBT has also been adapted to help children (ages 8 to 11) to reduce self-injurious behaviors. Perepletchikova et al. (2011) found that the 6-week DBT skills group intervention led to significant decreases in

depressive symptoms and suicidal ideation.

Despite its strong empirical foundation and growing list of promising adaptations, a number of gaps remain in the scientific literature examining DBT. Chief among the gaps in research is the lack of research with males or minorities. In addition, there continues to be a need for research investigating what predicts treatment response via component, dismantling, or process-analytic studies, as well as large-sample effectiveness research in community settings. Finally, as recent modifications of DBT to diagnoses other than BPD have yielded promising results, future research will likely increasingly target transdiagnostic processes, such as emotion regulation, that are common in, but not specific to BPD.

REFERENCES

American Psychiatric Association. *Practice Guidelines for the Treatment of Patients with Borderline Personality Disorder*. Washington, DC: American Psychiatric Association; 2001.

Aviram RB, Brodsky BS, Stanley B. Borderline personality disorder, stigma, and treatment implications. *Harv Rev Psychiatry*. 2006;14:249–256.

Axelrod SR, Perepletchikova F, Holtzman K, Sinha R. Emotion regulation and substance use frequency in women with substance dependence and borderline personality disorder receiving dialectical behavior therapy. *Am J Drug Alcohol Abuse*. 2011;37:37–42.

Bornovalova MA, Daughters SB. How does dialectical behavior therapy facilitate treatment retention among individuals with comorbid borderline personality disorder and substance use disorders? *Clin Psychol Rev*. 2007;27:923–943.

Carmel A, Fruzzetti AE, Rose ML. Dialectical behavior therapy training to reduce clinical burnout in a public behavioral health system. *Community Ment Health J*. 2014;50:25–30.

Glisenti K, Strodl E. Cognitive behavior therapy and dialectical behavior therapy for treating obese emotional eaters. *Clinical Case Studies*. 2012;11(2):71–88.

Gunderson JG. *Borderline Personality Disorder: A Clinical Guide*. Arlington, VA: American Psychiatric Pub; 2009.

Harned MS, Jackson SC, Comtois KA, Linehan MM. Dialectical behavior therapy as a precursor to PTSD treatment for suicidal and/or self-injuring women with borderline personality disorder. *J Traumatic Stress*. 2010;23:421–429.

Hörz S, Zanarini M, Frankenburg F, Reich DB, Fitzmaurice G. Ten-year use of mental health services by patients with borderline personality disorder and with other axis II disorders. *Psychiatr Serv*. 2010;61:612–616.

Kliem S, Kröger C, Kosfelder J. Dialectical behavior therapy for borderline personality disorder: a meta-analysis using mixed-effects modeling. *J Con and Clin Psych*. 2010;78:936–951.

Koons CR, Robins CJ, Tweed JL, Lynch TR, Gonzales AM. Efficacy of dialectical behavior therapy in women veterans with borderline personality disorder. *Behav Ther*. 2001;32:371–390.

*Linehan MM. *Cognitive Behavioral Treatment of Borderline Personality Disorder*. New York: Guilford Press; 1993.

*Linehan MM. *DBT Skills Training Manual*. Guilford Press; 2014.

*Linehan MM, Armstrong HE, Suarez A, Allmon D, Heard HL. Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Arch Gen Psychiatry*. 1991;48:1060–1064.

*Linehan MM, Comtois KA, Murray AM, Brown MZ, Gallop RJ. Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs. therapy by experts for suicidal behaviors and borderline personality disorder. *Arch Gen Psychiatry*. 2006;63:757–766.

Linehan MM, Dimeff LA, Reynolds SK, Comtois KA, Welch SS. Dialectical behavior therapy versus comprehensive validation therapy plus 12-step for the treatment of opioid dependent women meeting criteria for borderline personality disorder. *Drug Alcohol Depend*. 2002;67:13–26.

Linehan MM, Heard HL, Armstrong HE. Naturalistic follow-up of a behavioral treatment for chronically parasuicidal borderline patients. *Arch Gen Psychiatry*. 1993;50:971–974.

Linehan MM, McDavid JD, Brown MZ, Sayrs JH, Gallop RJ. Olanzapine plus dialectical behavior therapy for women with high irritability who meet criteria for borderline personality disorder: a double-blind, placebo-controlled pilot study. *J Clin Psychiatry*. 2008;69:999–1005

Linehan MM, Schmidt H 3rd, Dimeff LA, Craft JC, Kanter J, Comtois KA. Dialectical behavior therapy for patients with borderline personality disorder and drug-dependence. *Am J Addict*. 1999;8:279–292.

Linehan MM, Tutek DA, Heard HL, Armstrong HE. Interpersonal outcome of cognitive behavioral treatment for chronically suicidal borderline patients. *Am J Psychiatry*. 1994;151:1771–1776.

Lynch TR, Cheavens JS, Cukrowicz KC, Thorp S, Bronner L. Treatment of older adults with co-morbid personality disorder and depression: A dialectical behavior therapy approach. *Int J Geriatr Psychiatry*. 2007;22:131–143.

Lynch TR, Gray KL, Hempel RJ, Titley M, Chen EY, O'Mahen HA. Radically open-dialectical behavior therapy for adult anorexia nervosa: feasibility and outcomes from an inpatient program. *BMC Psychiatry*. 2013;13:293.

Mehlum L, Tørmoen AJ, Ramberg M, et al. Dialectical behavior therapy for adolescents with repeated suicidal and self-harming behavior: a randomized trial. *J Am Acad Child Adolesc Psychiatry*. 2014;53:1082–1091.

Neacsiu AD, Eberle J, Kramer R, Wiesmann T, Linehan MM. Dialectical behavior therapy skills for transdiagnostic emotion dysregulation: a pilot randomized controlled trial. *Behav Res Ther*. 2010;59:40–51.

Panos PT, Jackson JW, Hasan O, Panos A. Meta-analysis and systematic review assessing the efficacy of dialectical behavior therapy (DBT). *Res Soc Work Pract*. 2014;24:213.

Paris J. *Treatment of Borderline Personality Disorder: A Guide to Evidence-based Practice*. New York: Guilford Press; 2008.

Perepletchikova F, Axelrod SR, Kaufman J, Rounsaville BJ, Douglas-Palumberi H, Miller AL. Adapting dialectical behaviour therapy for children: towards a new research agenda for paediatric suicidal and non-suicidal self-injurious behaviours. *Child Adolesc Ment Health*. 2011;16:116–121.

Rathus JH, Miller AL. Dialectical behavior therapy adapted for suicidal adolescents. *Suicide Life Threat Behav*. 2002;32:146–157.

Safer DL, Telch CF, Agras W. Dialectical behavior therapy for bulimia nervosa. *Am J Psychiatry*.

2001;158:632–634.

Skodol AE, Gunderson JG, Pfohl B, Widiger TA, Livesley WJ, Siever LJ. The borderline diagnosis I: psychopathology, comorbidity, and personality structure. *Biol Psychiatry*. 2002;51:936–950.

Stoffers JM, Völlm BA, Rücker G, Timmer A, Huband N, Lieb K. Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst Rev*. 2012;(8):CD005652.

Telch CF, Agras W, Linehan MM. Dialectical behavior therapy for binge eating disorder. *J Consult Clin Psychol*. 2001;69:1061–1065.

van den Bosch LM, Verheul R, Schippers GM, van den Brink W. Dialectical behavior therapy of borderline patients with and without substance use problems. Implementation and long-term effects. *Addict Behav*. 2002;27:911–923.

Van Dijk S, Jeffrey J, Katz MR. A randomized, controlled, pilot study of dialectical behavior therapy skills in a psychoeducational group for individuals with bipolar disorder. *J Affect Disord*. 2013;145:386–393.

Verheul R, Van Den Bosch LM, Koeter MW, De Ridder MA, Stijnen T, Van Den Brink W. Dialectical behaviour therapy for women with borderline personality disorder: 12-month, randomised clinical trial in The Netherlands. *Br J Psychiatry*. 2003;182:135–140.

Zanarini MC, Frankenburg FR, Reich DB, Fitzmaurice G, Weinberg I, Gunderson JG. The 10-year course of physically self-destructive acts reported by borderline patients and axis II comparison subjects. *Acta Psychiatr Scand*. 2008;117:177–184.

▲ 33.10 Intensive Short-Term Psychodynamic Psychotherapies

MANUEL TRUJILLO, M.D.

INTRODUCTION AND DEFINITION

The terms Short-Term Dynamic Psychotherapies (STDPs) and Intensive Short-Term Dynamic Psychotherapies (ISTDPs) denote a robust corpus of theories, clinical practices, and therapeutic techniques deriving from psychoanalysis, as discovered, developed, and formalized by Sigmund Freud at the end of the 19th century and the first third of the 20th century. Those short-term therapy systems that include the label psychodynamic in their self-definition, usually wish to signal a strong adherence to the classical psychoanalytic metapsychology and a very extensive inclusion in their therapeutic armamentarium, of essential psychoanalytic techniques such as the utilization of transference concepts and transference interpretations, and the use of the unconscious and of unconscious mental functions to understand the causal underpinnings of psychic conflict, and to frame the development of the psychopathology of common psychiatric disorders. Finally, most of the currently active systems of STDPs use the interpretation of impulses, sexual and aggressive, and feelings, loss and abandonment, shame and guilt, as key therapeutic interventions.

The growth of psychotherapy in general, and of dynamic psychotherapies derived from the psychoanalytic framework in particular, represents a landmark achievement in the history of psychiatry. Most impressive is the capacity of the STDP field to meet the challenging standards of contemporary biomedical research paradigms, by producing high-quality treatment research studies evaluating the therapy outcomes obtained in well-defined patient populations suffering from various psychiatric disorders such as, anxiety, depressive disorders, personality disorders and others that, together, represent a vast majority of the large and growing suffering and disability burden imposed by psychiatric disorders on most societies.

Probably because of these advances, and despite growing and spirited competition from many other biomedical and psychosocial interventions, more people are treated with psychodynamic approaches than ever before in history. Those who are treated achieve clinical results at least comparable to those treated with alternative psychological/psychosocial or biomedical approaches. In the cases complicated by the presence of significant personality pathology or substantive psychosocial problems, traumatic experiences, and psychosocial crises, the results achieved by psychodynamic psychotherapies are often superior to the clinical outcomes delivered through standard clinical management, including the use of

psychotropic medications.

Since the methodologically sound meta-analysis conducted in the 1970s by Mary Lee Smith, Gene V. Glass, and Thomas I. Miller which concluded that 75 percent of patients treated with psychotherapy were better off at termination and follow-up than the untreated controls, many rigorous research studies have demonstrated the efficacy and effectiveness of various models of the STDPs when treating such common disorders as anxiety and depression, selected personality disorders, certain posttraumatic stress syndromes, and selected eating and substance abuse disorders.

In 2013, the American Psychological Association's Council, after an extensive and comprehensive review of psychotherapy research studies, concluded that the general or average effects of psychotherapy are widely accepted to be significant and large and are quite constant across most diagnostic conditions, with variations being more influenced by general severity than by particular diagnoses.

The council members asserted that the results of psychotherapy tend to last longer, and be less likely to require additional treatment courses, than psychopharmacological treatments. For example, in the treatment of depression and anxiety disorders, psychotherapy clients/patients acquire a variety of life skills that are used after the treatment termination, and generally may continue to improve after the termination of treatment. Within this picture of broad effectiveness for therapies of long duration, STDPs hold a position of distinction. Systematic review and meta-analyses conducted since 2010 of the effect of STDPs on patients suffering from mood disorders, anxiety disorders, somatic conditions, and personality disorders, concluded, after reviewing studies meeting high standards of quality, that the therapies studied were clinically effective, provided respectable effect sizes, and were cost-effective compared with the control conditions. The fact that these robust outcomes were also obtained working with patients suffering from personality disorders, who are not considered very responsive to other forms of psychotherapy, highlights the potential therapeutic outcome, and the value, of these ISTDPs, designed by Habib Davanloo starting in the 1970s.

A question stemming from a public mental health perspective is, can these interventions be made readily available to the large number of patients who may benefit from them.

In parallel to the growth of research in these short-term techniques, the use of planned, psychodynamically oriented, short-term modes of treatment in clinical settings has also grown. Hanna Levenson and colleagues, in a comprehensive national survey of psychiatrists, psychologists, and clinical social workers, found that psychodynamically oriented clinicians provided a sizable proportion of all brief therapy conducted nationally. Most such clinicians acknowledged treating many

patients in brief therapy and would welcome supervisory and learning opportunities to enhance their technical expertise in STDP and ISTDP. The challenge to the field is to provide such learning opportunities to sufficient number of therapists to ensure the availability of such therapies to the large number of patients who, suffering from gravely disabling conditions, remain untreated, or undertreated in the community.

The current state of the field of STDPs can thus be described as healthy. The field has developed well-documented models of treatment and has proposed specific techniques that are often codified in research-based treatment manuals. Practitioners can count on numerous well-conducted research studies to support efficacy and effectiveness claims, and large numbers of practitioners offer short-term dynamic approaches nationwide to the growing number of patients who consult them for their psychiatric disorders and life crises.

History

The Early 1900s. The history of ISTDP traces back to Sigmund Freud, in that there is reason to believe that many of Freud's early treatments lasted from a few hours to a few months. Many such early treatments were in fact training analyses conducted with psychiatrists and other clinicians traveling from locations throughout Europe or the United States with stringent travel time limits. Freud's treatment of the conductor Bruno Walter that took place in 1906 lasted a few sessions, and the composer Gustav Mahler was treated in 1908 during a 4-hour walk through the woods of Vienna for an episode of psychogenic impotence.

As the time went on, various patient factors and technical developments may have contributed to the lengthening of the course of psychoanalysis. The increasing reliance on free association as a key technical ingredient diminished, by necessity, the analyst activity and may have diffused the therapeutic focus of the clinical situation. The use of the couch and the increasing commitment to allow for the slow development of a transference neurosis may have further contributed to the lengthening of the psychotherapeutic process.

The consolidation of the early psychoanalytic technical triad of reduced therapist activity, encouragement of free association, and the development of the transference neurosis has yielded substantial heuristic fruits, and has undoubtedly contributed to the construction of the vast edifice of psychoanalytic metapsychology and theory of psychopathology. The price, however, may have been the reduction of therapeutic potency for those patients for whom the nature and extent of their resistances (particularly what Freud termed superego resistances) proved impervious to the technique of free association—thus Freud acknowledged growing pessimism regarding the power of free association to overcome such powerful resistances.

The Middle 1900s. Among early psychoanalytic practitioners, Sandor Ferenczi and Otto Rank deserve credit for attempting to create alternative techniques to enhance the therapeutic effect of psychoanalysis. In their book *The Development of Psychoanalysis,* they delineated a therapeutic position that included enhanced activity of the therapist, the active development of a therapeutic focus, and the setting of time limits—a triad that was to become the core of the technical offerings of later innovators such as Peter Sifneos, David Malan, and Habib Davanloo. These investigators were also aware that when patients acquired just intellectual knowledge, without integrated affects, such knowledge may serve as resistance, and they recommended an active maintenance of the focus on affect through the course of therapy to overcome such as intellectualized resistance. In the late 1930s, Frank Alexander, by then a prominent psychoanalytic leader in Chicago, focused his efforts on a pioneering research project aimed at making psychoanalytic treatment more effective, to specify the recovery techniques (process variables) necessary to achieve such a goal, and to shorten the duration of the average course of treatment More than 600 patients were included in his studies, the findings of which were reported in 1946 in the book *Psychoanalytic Therapy* coauthored with Thomas French.

Consonant with the work of Ferenczi and Rank, Alexander also emphasized the importance of the therapist increased activity and the maintenance of the therapeutic focus on the patient's emotional experience —thus his emphasis on what he called the "corrective emotional experience," an experience that he considered essential to any successful therapeutic venture. Such corrective emotional experience affords the patient an opportunity to relive, in the safe and supportive presence of the therapist, the original childhood traumatic experiences that contributed to the development of the patient's neurotic symptoms and maladaptive defenses. In his 1946 book, Alexander reported therapeutic results comparable to those of long-term psychotherapy and emphasized the fact that, besides symptomatic improvement, short-term therapies can also promote significant character changes. Of note, both his belief in the therapeutic power of bringing affect into the focus of treatment, and in facilitating emotional experiences in the here and now of the therapy session (inevitably related to the transference) have been substantiated again and again in high-quality research studies.

The Contemporary Situation. Starting in the 1960s the field of STDP underwent tremendous growth. Clinicians and researchers across the United States and the United Kingdom proposed diverse short-term models anchored in the psychoanalytic tradition. Authors such as David Malan in London, Peter Sifneos in Boston, Habib Davanloo in Montreal,

and others generated models of psychoanalytically based psychodynamic psychotherapies applicable to increasingly broader populations of patients, and—with significant boldness—provided transcripts, audio recordings or videotapes of treatment sessions. These authors defined in very specific terms their treatment techniques, and initiated basic research strategies to examine the outcomes of their treatments. Additional models were developed by clinicians and investigators such as Mann, Horowitz, Luborsky, Strupp, and by the New York Beth Israel group (Winston, Trujillo, Pollack, Been, McCullough, Laikin, and others). The Beth Israel group, and other investigators such as A Abbass, have contributed seminal research studies demonstrating the efficacy and effectiveness of STDP as well as its broad applicability to patients suffering from complex psychiatric and personality disorders.

CLINICAL SPECTRUM OF STDPS

This chapter will describe in more detail the models proposed by Sifneos, Malan, and Davanloo as their models (which have much in common) can be ordered in a spectrum of complexity of techniques as well as of clinical severity of the pathology of patients who can benefit from them. Sifneos's short-term anxiety-provoking psychotherapy (STAPP) initiates the spectrum and occupies its relatively less complex position as it is mainly applicable to patients suffering from psychodynamically uncomplicated conditions, which can be successfully treated using basic technical interventions (development of a treatment focus, increased therapist activity, minimal to moderate resistance, and transference work). Increasing technical complexity, and patient applicability are provided by Malan's short-term model, which can be situated in the middle range of the theoretical spectrum. Davanloo's ISTDP provides more complex interventions to reach more complicated psychopathology.

Short-Term Anxiety-Provoking Psychotherapy—Peter Sifneos

Theoretical Issues. Developed by Peter Sifneos as a result of his 7-week treatment of a patient who presented to his clinic at the Massachusetts General Hospital complaining of severe anxiety, but needed to move from Boston after a few months, STAPP has found its place in the contemporary therapeutic armamentarium as a brief psychotherapy approach for patients who, having an oedipal conflict as the basis of their symptomatology, display high motivation for change, and suffer from few characterological complications. Typically, the oedipal dynamics lying at the core of the patient's clinical problems, have to do with the conflictual aftermath of unresolved competition with the parent of the same sex for the love and affection of the parent of the opposite sex (or his or her surrogate). The presence of significant anxiety, and of the defenses deployed by the patient

to counter such anxiety, make it difficult for the patient to resolve such conflicts through "living," leading, as time goes on to the development of overt clinical symptoms, subjective suffering and/or interpersonal difficulties. Common chief presenting complaints of STAPP patients include various experiences of anxiety, alone or in conjunction with other symptoms, grief reactions, mild-to-moderate depression, somatic symptoms, and/or interpersonal difficulties unrelated to enduring character disorders.

Techniques. As Sifneos indicates, "the technical therapeutic considerations for this kind of psychotherapy are specific and clear-cut. The therapist must encourage the development of a therapeutic alliance and must create the atmosphere where learning takes place."

Taking advantage of the prevailing positive transference feeling, staying within the specified focus, and using anxiety-provoking confrontation questions, clarifications, and interpretations, the therapist must strive to clarify with the patient the conflictual impulses and feelings underlying their psychopathology, while avoiding getting entangled in pregenital characterological issues such as passivity, dependence, sadomasochistic or narcissistic features, and acting-out tendencies, which are often used defensively by the patient and tend to prolong treatment.

Sifneos' prescription synthesizes three core technical dimensions: the use of positive transference, the active maintenance of the focus, and the avoidance of regression. Beyond this general prescription, Sifneos provides detailed descriptions of techniques covering the process of therapy and including the following domains: selection criteria, initial assessment, early and middle phases of therapy, and termination.

SELECTION CRITERIA. Sifneos summarizes his selection criteria for STAPP as follows:

(1) Presence of a circumscribed chief complaint.
(2) Evidence of a give-and-take or meaningful relationship in childhood.
(3) Capacity to relate flexibly to the evaluator and to express feelings.
(4) Above-average intelligence and psychological mindedness.
(5) Motivation for change.

INITIAL ASSESSMENT. Beyond the completion of a classical psychiatric history and mental status examination necessary for the development of a criteria-based clinical diagnosis following the fifth edition of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), the initial assessment of patients selected for STAPP treatment must provide evidence of the fulfillment of the selection criteria and must lead to the joint development of a sufficiently explanatory psychodynamic formulation, a therapeutic focus, and a therapeutic contract. Such

developments are usually the result of an active and dynamic interaction between the patient and the therapist, an interaction that optimally ends in the mutually agreed delineation of therapeutic foci that clearly and succinctly crystallize the specific psychological conflicts underlying the patient's psychological problems.

THE OPENING PHASE. The goal of the opening phase of treatment is to deepen the patient's commitment to work on the mutually agreed upon therapeutic contract as delineated in the initial interviews. To accomplish that goal, Sifneos recommended the use of techniques such as high therapeutic activity, the maintenance of the therapeutic focus, the consolidation and extension of the therapeutic alliance, the use of positive transference, and the avoidance of the defensive (and regressive) use of pregenital characterological issues.

THE MIDDLE PHASE. From the technical point of view, the middle phase of treatment has as its goal a significant working through the patient's psychological conflicts as delineated early in the treatment in the dynamic formulation and therapeutic focus. Here the facilitation of transference-based "corrective emotional" experiences is central. As depicted by Alexander, "the key to shorter therapy lies in the ability to experience and control the transference." To accomplish this task, the STAPP therapist continues to use anxiety-provoking questions, confrontations, and clarifications as he or she links the emerging transference feelings with similar experiences in the patient's past and/or in other areas of the patient's present life. In this phase of therapy these interpretative links should lead to new learning and healthy problem solving. To facilitate the consolidation of such learning, the therapist needs to support the emerging evidence of such growth and continue to use supportive, clarifying, educational, and insight-promoting techniques.

THE TERMINATION PHASE. Because STAPP is designed as short-term therapy with open-ended termination, the therapist must rely on evidence of significant clinical improvement before approaching the termination phase. To the persisting controversy in academic psychotherapy as to who should initiate the terminations of the therapy, Sifneos provides clarifying answers: The therapist, being more able to discern objectively the state of resolution of the patient's original conflicts, should initiate termination discussions. In practice, however, more than half of Sifneos' patients spontaneously mentioned the possibility of ending the therapy as they experienced better moods and enhanced self-esteem and enjoyed their new problem-solving capacities.

Regardless of who initiates the discussion, termination is "in the air" when evidence is provided that points toward substantial differences in the attitudes and behavioral patterns present at the initiation of therapy, and

new problem-solving skills are applied by the patient to the solution of their own conflicts. Abilities such as progressive independence in the pursuit of personal goals, increasing self-assurance, and improved self-esteem signal the emergence of new adaptive capacities liberated from the shadows of the unconscious conflicts that dominated the patient's life prior to the initiation of therapy.

Clinical Issues

INDICATIONS. Clinical experience with Sifneos' techniques suggests its application to the patient who fits the following DSM-5 diagnostic categories:

Adjustment disorders with depressive mood, with anxiety, and with mixed anxiety and depressed mood; generalized anxiety disorder; panic disorder, selective somatic symptom disorders, Illness anxiety disorders, psychological factors affecting other medical conditions, acute stress disorder, persistent depressive disorder DSM-IV's dysthymic disorder; and major depressive disorder.

As taught by Sifneos, beyond those diagnostic parameters, these patients should fulfill the following selection criteria: relatively circumscribed chief complaint, a history of at least one personal relationship characterized by mutuality ("give and take"), ability to interact flexibly with the therapist/evaluator, average intelligence with responsiveness to insight, and motivation for change.

LIMITATIONS. As reported by Sifneos, patients with significant character pathology, excessive dependency, multiple foci of loss, or active substance abuse may not be suitable candidates for STAPP.

COMPLICATIONS. Beyond premature abandonment of treatment, the most commonly reported complications include the development of a transference neurosis that may need additional time for resolution. Fear of termination may produce clinical regressions that can, and needs to, be handled by active interpretation within the framework of the technique. This complication is more likely to appear in those cases in which one or more early losses have not appeared in the initial psychodynamic evaluation and thus have not been incorporated in the therapeutic focus, promoting its expression through distinct behaviors in the transference, such as excessive dependency and clinging, as well as through other signs of clinical regression.

CONTRAINDICATIONS. Active dependence on alcohol or other drugs needs to be addressed independently, and a period of abstinence (30 to 90 days) should be documented prior to the initiation of STAPP. Active suicidal

ideation and overwhelming interpersonal, work, or financial adversity should prompt the use of crisis intervention techniques. Once the crisis situation is resolved, the use of STAPP could be considered after a careful evaluation of the stability of the patient's psychosocial situation and the restoration of sufficient ego strength.

Chris, a 31-year-old single man, sought help for a moderate depressive episode precipitated by the loss of his relationship with his girlfriend, Joanna. She had broken off the relationship after approximately 1 year, tired of Chris' erratic work ethic and emotional instability, and discouraged by his fear of commitment to the future of their relationship. This cycle of infatuation, increasing fear of commitment, and relationship loss had become a pattern in Chris' interpersonal life. His work life was plagued with similar problems. Chris lost several jobs because of serious conflicts, and threatening confrontations with his superiors. As conflicts arose at both work and home, Chris typically suffered increasing anxiety and episodic panic attacks. In the aftermath of each loss, Chris usually was stunned, sad and often endured prolonged depressive episodes at times accompanied by suicidal ideation.

DYNAMIC FORMULATION. The development and delineation of an accurate dynamic formulation is an essential clinical activity in STDP. The challenge is to construct, based on the findings of the clinical situation, a clinical synthesis that links the patient's subjective symptoms with the underlying original childhood foci that determine or drive such symptoms. These genetic foci may be related to repressed unconscious impulses (aggressive or sexual) and feelings (loss, grief, self-doubts) that may have first occurred at key developmental points, and remain suppressed and unresolved. The end result of the dynamic formulation can be represented by a dynamic holographic (Fig. 33.10–1) representation of the pathogenic sequence that links deep genetic impulses or feelings with broad dynamic forces (repetitive experiences such as irritability or ready to explode anger or habitual maladaptive behaviors such as automatic opposition to authority figure). The final step of the hologram, links those dynamic forces with their surface, conscious conflicts or behavioral manifestations such as sullen withdrawal and passivity or self-destructive angry acting out.

For Chris in the foregoing case, the dynamic hologram can be represented as follows:

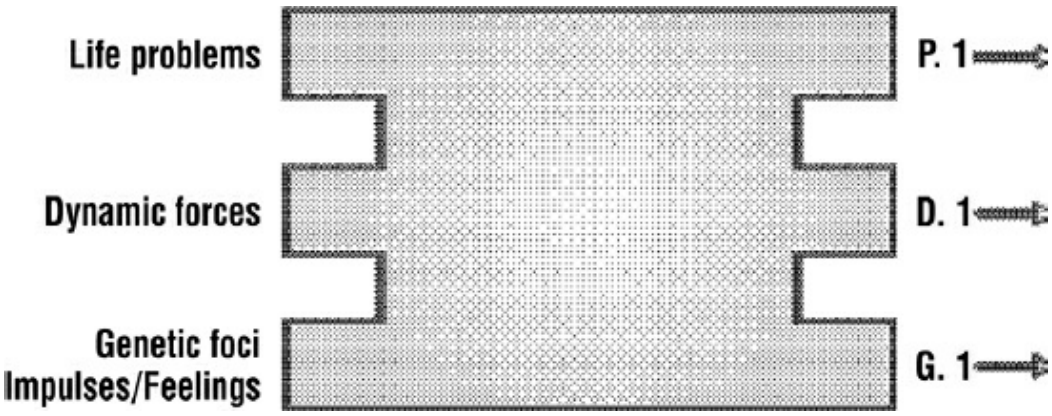

FIGURE 33.10–1. Dynamic hologram.

Life problems: Recurrent episodes of anxiety and depression; work problems; unstable interpersonal relationships; conflict with authority figures; antagonism toward, and emotional distance from, his father, brother, and other male friends; and fears of heterosexual intimacy and of commitment.

Dynamic forces: Ongoing hostility and envy toward males, authority figures, and successful people, and compulsive and possessive craving and seeking female love objects while simultaneously he had great difficulties considering fulfilling, or even recognizing their independent needs.

Genetic pathogenic foci: Unconscious loss of maternal objects precipitated by birth of brother at age 2 years; uncontrolled grief for that loss with compulsive drive to experience child-like possession of love objects; and compulsive hostility toward others perceived as rivals.

PROCESS AND GOALS OF TREATMENT. The typical clinical process in STAPP can be summarized as follows:

Objective: Resolution of Oedipal conflicts
Focus/Foci: Triangular/Oedipal
Duration: From 10 to 20 sessions, lasting from 50 to 60 minutes occurring once a week
Termination: Open

After the evaluation phase, the typical therapy process proceeds on the strength of the therapeutic alliance optimally established during the early sessions. The efforts of the therapist in the initial phase aim to enrich the dynamic formulation through an active inquiry around the foci established in the initial formulation. This active inquiry, in the case of Chris, yielded additional confirmation of the persistence of repressed sexual feelings toward his mother and the presence of hostile feelings toward all rivals for his mother's affection. A memory suffused with very visceral feelings emerged in this phase as a result of the therapist's active inquiry. In this memory, Chris saw himself in his mother's arms in a dark room. He remembered vividly the intense pleasure of the contact with the warm skin of his mother, the texture of her clothes, and the smell of her perfume. While narrating this memory to the therapist, Chris was so absorbed in the experience that he blushed intensely while describing it to the therapist. He first giggled with delight and amazement, a look of disbelief on his face, seeking the complicity of the therapist in this timeless experience of imaginary possession, and enjoyment of his mother's body. He next described the painful termination of this moment of pleasure by his father's sudden, angry forceful, and disruptive opening of the closet door, and intense pain of the flood of light that disturbed his pleasurable absorption. After this sequence, Chris felt sad and open himself to the experience of grief at this loss of the intense and exclusive bond with his mother; a loss that in his childhood he had also after his brother's birth.

Feeling safe in the therapeutic situation, he was able to experience again and again, the sadness, loneliness, sense of abandonment, and rage that had been all too familiar through his childhood and that recurred as well in his present life when his romantic attachments were threatened or actually lost. The affective link between this childhood experience and his intimacy

problems in the present became viscerally obvious to Chris, and the acceptance of this link enhanced his capacity to work through this essential component of his pathology. A parallel conflict appeared in the transference as the patient resented the "intrusion" of the inquiring therapist into the zealously guarded privacy of this primal fantasy of maternal possession.

OUTCOMES. In addition to his role as a clinical pioneer, Sifneos deserves additional credit for his commitment to make his work visible through the pioneering utilization of the rigorous outcome evaluation of his work. In two studies, Sifneos reported respectable clinical outcomes, which included moderate symptomatic improvement, significant enhancement of self-esteem, and positive resolution of the crisis situation that initially brought the patients into treatment.

Beyond this early evaluation, Sifneos' approach has been studied in a modified form by many clinical researchers. The work of Arnold Winston and colleagues represents an excellent investigational effort. In these studies, two forms of brief therapy based on procedures specified in therapy manuals were used. One of them, labeled STDP was based on principles developed by Davanloo. The second one—brief adaptive psychotherapy (BAP)—used a composite of brief therapy techniques and was heavily influenced by the work of Sifneos. The results of both of Winston's studies indicate that both treatment conditions were equally effective and obtained clinical outcomes that were significantly better than those achieved by the patients who remained on the waiting-list control for the duration of the study. Of note, the Sifneos-derived BAP technique achieved better outcomes than the STDP in measures of anxiety and phobic anxiety, whereas STDP showed significant advantage over BAP on depression subscales.

Short-Term Dynamic Psychotherapy—David Malan

Theoretical Issues. The work of David Malan flows from the interest of Michael Balint and his group at the Tavistock Clinic in London, in expanding the boundaries of psychoanalysis by training general practitioners in basic psychotherapy techniques, and in developing brief psychotherapy techniques to enrich the clinical armamentarium of experienced psychoanalytically oriented psychotherapists.

Anchored in classic psychoanalytic metapsychology and theory of technique, Malan's brilliant contributions to STDP involve both technical aspects and research insights. His contributions to technique include the focusing of the therapist's interventions on the triad of impulses, defenses, and anxiety (triangle of conflict) as it plays out in relation to the therapist and to key figures in the patient's current life and the past. In this regard, Malan advocates for a comprehensive exploration of all components of the

triangle of conflict and recommends a thorough working of the impulse component of the past conflict.

His contributions to research are many. By keeping the research enterprise very close to the clinical ground and to the way in which psychodynamic clinicians actually think and behave, he has fostered the meaningful generalizability and reliability of his findings. His attempts to make outcome predictions from baseline clinical data, place him as a pioneer in research techniques that later grew exponentially in the 1970s and 1980s. A key finding of his studies is that the higher the number of interpretations linking the current behavior of the patient with the past and the transference (TCP), the better is the outcome. This finding has stood the test of time and has been replicated in numerous other psychodynamic psychotherapy studies.

Techniques. The treatment process in Malan's STDP can be summarized in the following manner:

OBJECTIVE. To work through the patient's dynamic problems by linking the patient's impulses and feelings, defenses, and anxiety in relation to each relevant psychodynamic focus, while promoting "corrective emotional experiences" with the therapist in the transference. The typical outcome of such working through is increased tolerance of interpersonal conflict (ego strength), enhanced assertion, and substantial lessening of dependence. Negative feelings and pessimism are also reduced.

FOCI. Unresolved interpersonal conflicts and suppressed impulses and feelings, including feelings of abandonment and loss, links between traumatic experiences in the past, congruent situations in the present, and lived perceptions and emotions in the transference. A patient with unresolved abandonment and separation fears may come to understand the link between this fear (based in childhood abandonment experiences) and her present day compulsive neediness and clinginess. Supported by her positive transference experience, such patient may feel confident enough to endure fears of loss and initiate assertive independence in key areas.

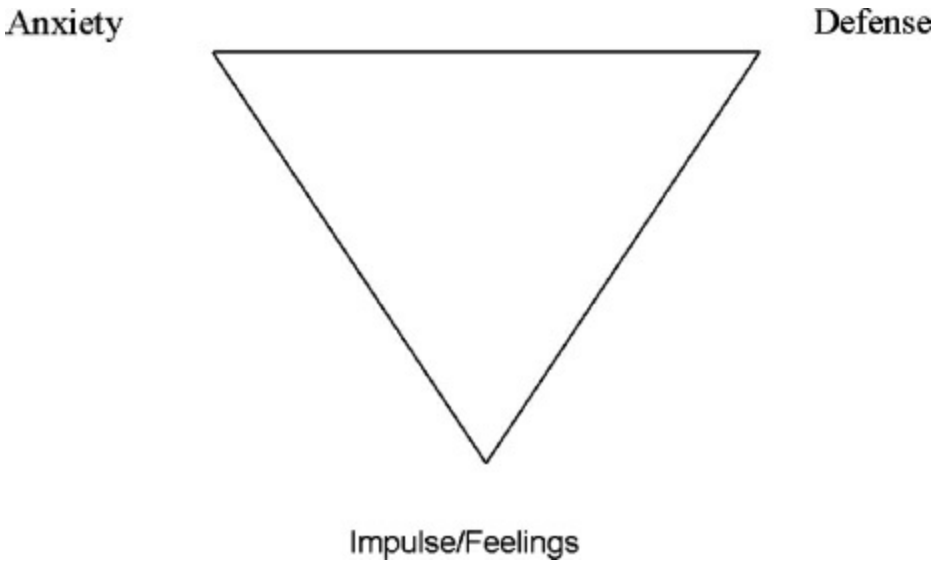

FIGURE 33.10–2. Triangle of conflict.

SELECTION CRITERIA. Motivation for psychological change, response to early tentative interpretations (trial interpretations), and ability to tolerate and work with intense feelings.

EXCLUSION CRITERIA. Inability to create emotional contact between therapist and patient, extremely rigid defenses, lack of motivation for dynamic change, and the presence of risk factors for the development of intense dependency.

Therapists working within Malan's framework used all psychoanalytic technical interventions, organized around two conceptual triangles labeled the triangle of conflict and the triangle of the person (Figs. 33.10–2 and 33.10–3). The objective of treatment is the repeated, and affect-laden, clarification of defense in the triangle of conflict, which permits access to the experience of underlying impulses and feelings when anxiety is appropriately modulated. Increased awareness of formerly repressed impulses and feelings permits their elaboration and working through as they appear in the patient's current life, the past, and the transference. As noted previously, Malan observed that the higher the number of links between experiences in the current life of the patient, the past, and the transference, the better is the outcome.

The techniques used by Malan are the ones used in standard analytically oriented psychotherapy. In addition to the clinical material spontaneously communicated by the patient or developed through the therapeutic interaction, dreams, fantasies, slips, and associations are incorporated into the clinical work.

The clinical material is kept manageable by the efforts of the therapist to keep the therapy in focus, an objective that is achieved by following Malan's recommendation to use "selective interpretation, selective

attention, and selective neglect." The therapy is terminated when there is evidence of sufficient symptomatic improvement. In most cases reported by Malan, patients also achieved substantial improvements in long-standing characteriological problems.

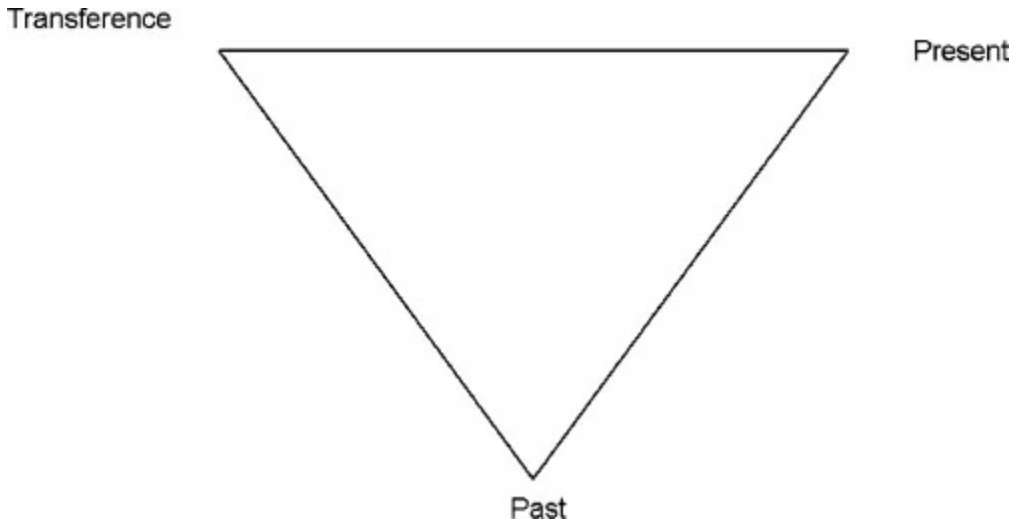

FIGURE 33.10–3. Triangle of the person.

Clinical Issues

INDICATIONS. *Adjustment disorders with depressive mood, with anxiety, and with mixed anxiety and depressed mood; generalized anxiety disorder; panic disorder, selective somatic symptom disorders, Illness anxiety disorders, psychological factors affecting other medical conditions, acute stress disorder. Selective separation anxiety disorders. Persistent depressive disorder, dysthymic disorder and major depressive disorder. Selected personality disorders such as avoidant and dependent personality disorders.*

LIMITATIONS. The range of techniques proposed by Malan can provide treatment to patients suffering from oedipal psychopathology complicated by loss and abandonment, as long as the patients have not developed too intense, character-based interpersonal defenses or express regressive trends such as diffuse pervasive anxiety, paralyzing helplessness, extensive somatization, and the like. In his series of studies, Malan and his group predicted poor outcome in patients who failed to make contact with the evaluator, remained poorly motivated after the initial interviews, or developed intense transference-based dependency.

COMPLICATIONS. The most common complication reported in Malan's studies relates to the inability to terminate the therapy within the 40-session upper limit. In Malan's study most of these patients went on to medium-length and long-term psychoanalytic psychotherapies but

ultimately achieved delayed, but respectable, clinical outcomes. Of the 30 patients treated in Malan's second study, only 1 was reported to be clinically worse at follow-up.

CONTRAINDICATIONS. As reported in Malan's studies, there were certain absolute contraindications that eliminated patients from being considered a candidate for this treatment. They included a recent history of suicide attempts, active substance abuse, history of more than one course of electroconvulsive therapy (ECT), chronic incapacitating obsessional symptoms, and gross destructive or self-destructive acting out. Whether selected samples of these originally excluded patients would be good candidates for therapy should they respond positively to trial therapy needs to be tested in additional studies.

Ana, a divorced 60-year-old woman, sought psychiatric help following a severe depressive episode lasting several months. This episode, which was one of many in her life, was especially severe in terms of loss of energy, interest, and motivation, as well as in terms of the intensity of her sadness and her wish to die. Only her profound religious convictions protected her from acting on these wishes. Ana had lost a lot of weight, had trouble sleeping, experienced many nightmares, and had difficulty with concentration. She was plagued by pervasive feelings of hatred for her mother, who was by now very old, ill, and dependent on Ana, who was unable to forgive her for having abandoned her in an orphanage when she was 5 or 6 years of age.

After an extensive assessment, the dynamic formulation of Ana's problem was represented as follows:

Life problems: Recurrent depressive episodes plagued by feelings of guilt and self-reproach; problems with men involving choosing partners who are commonly cold, distant, or otherwise unavailable; involuntary and painful emotional distance from her children, friends, and other close relationships; and unproductive and unrewarding work-life, despite considerable intellectual gifts.

Dynamics: Ambivalent relationship with her mother, whom she blames for most of the tragedies of her life; guilt and need for punishment in relation to her unrelenting hatred for her mother; and pathological grief reaction for the loss of an idealized and more optimal relationship with her mother, the one she remembers she had prior to her orphanage placement. From this focus there flows a melancholic conviction of the inevitable failure of human relationships.

Pathogenic foci: Grief and inability to mourn the loss of her mother after she was placed in the orphanage, with attendant rage and guilt; a second pathogenic focus results from her pathological grief for the loss of her father, who, because of severe alcoholism, had abandoned the family; a move that caused the mother to place her children in an orphanage in order to be able to work and ultimately recover their care. Unconsciously, Ana had always blamed her mother for the family catastrophe, thus "protecting" an idealized view of her father, to whom she was profoundly attached.

For Ana, the initial phase of treatment focused on facilitating, through quite intense resistance, the actual affective experience of her destructive anger toward her mother, the work through of which, made possible the appearance of a modicum of empathy with her mother's painful life situation around the time she placed Ana and her sisters in the orphanage. Next, the therapy focused on Ana's father. Deep feelings of idealization, disappointment, anger, and grief were experienced with increasing clarity and intensity, frequently via displaced feelings in the transference and almost always after overcoming considerable resistance. The last phase of treatment permitted the gradual development of realistic feelings of empathy and appreciation for her mother, experienced now without the cloud of anger or the numbed emotional distancing which had accompanied such feelings in the past. The reawakening within Ana of feelings of joy and hope, as well as professional ambition also needed support and

encouragement, as such "luxuries" had been blocked in her development by her feelings of guilt.

OUTCOMES. Although Malan's carefully designed and elegant research studies would not conform to today's requirements for research in STDP, they contributed rich insights to the field. Malan reported on two different naturalistic, non-controlled samples. The results of the initial study of 21 patients were reported in Malan's 1963 book, *A Study of Brief Psychotherapy*. A replication study involving 39 additional patients was reported in 1976. Both studies generated a wealth of therapeutic insights, which have guided the field since their publication. His findings confirmed that neither severity of baselines symptoms nor their relative chronicity predicted clinical outcome at termination, whereas the presence of motivation for insight was highly predictive of success. In terms of therapeutic process, use of transference and transference–past-present links was also found to be highly predictive of positive outcome.

Malan's proposals, both in direct or modified forms, were later thoroughly tested in efficacy studies. Leichsenring's meta-analyses included several studies using Malan's STDP as one of the study groups. Probably the most rigorous test of Malan-derived STDP principles and techniques is represented by a study by Svartberg and colleagues, which showed statistically significant improvements to patients with Cluster C personality disorder suffering from anxiety and depressive disorders. This study compared STDP and cognitive therapy and found changes in symptom distress after active treatment termination which were statistically significant only for the STDP group.

Intensive Short-Term Dynamic Psychotherapy—Habib Davanloo

Theoretical Issues. Davanloo developed his technique of ISTDP in a series of clinical research projects at the Department of Psychiatry of the Montreal General Hospital of McGill University. The initial studies (conducted between 1963 and 1974) established that excellent results could be achieved working largely through interpretations on a population of highly responsive and motivated patients suffering from relatively simple psychodynamic problems. A major naturalistic research project was then designed to broaden the base of eligible patients by testing the effect of techniques designed to deal with increasing resistance. This study established for Davanloo that the combination of transference focus and special handling of resistances could be used to achieve major breakthroughs into the patient's unconscious, leading all the way to major unlocking of the unconscious, a process found to have great therapeutic impact. These discoveries led to the development and consolidation of his technique of ISTDP and to a new theoretical understanding of the function of the human unconscious and of the role of transference and the

therapeutic alliance in the therapy of highly resistant patients. For Davanloo the key therapeutic force involves the mobilization of the therapeutic alliance through a highly specialized form of interviewing designed to deal with the patient's overwhelming resistance. The second major area of scientific concerns Davanloo's views of the centrality of the process of attachment for human development and of the importance of its disruptions for psychopathology. For Davanloo, it is the severe traumatization of core attachment bonds during development what causes psychic traumas, the pain of trauma, the mobilization of primitive rage and murderous impulses, and the guilt, grief, and character distortions that one reliably finds at the core of the neurotic processes of complicated psychoneurotic patients.

Techniques. The core of Davanloo's techniques flow from his research on the twin dynamics of transference and resistance. The central aim of Davanloo's system of ISTDP is the rapid reduction or removal of resistance through the careful application of a system of planned interventions that end—when successful—in direct access to the unconscious as a result of the breakdown of the major resistance and defense systems with the attendant mobilization of unconscious therapeutic alliance.

This planned sequence of steps to deal with resistance is what Davanloo called the central dynamic sequence. The eight technical steps of this sequence are as follows:

(1) Phase of inquiry into the patient's clinical and interpersonal difficulties.
(2) Pressure over the avoided or repressed feelings.
(3) Mild to intense challenge to the patient's resistance: Davanloo's "head-on-collision."
(4) Detection and challenge to the transference resistance.
(5) Direct access to the unconscious; usually in the form of viscerally intense, sudden eruption of vividly experienced primal and primitive rage, sadness, grief, and the dominant ascent of a deep therapeutic alliance.
(6) Systematic analysis of the transference.
(7) Dynamic exploration of the core unconscious configuration of conflict.
(8) Consolidation and preparation for working through of such unconscious configuration.

The end product of this process, when successful, involves increasing visceral experience and awareness of intense murderous rage and other deep feelings, which usually appear first in the transference and which then are focused on key figures in the patient's development and childhood trauma. As these feelings burst into open, they typically bring forth deep waves of sadness and guilt, both of which are experienced deeply and

viscerally in a manner that promotes the resolution of the pathogenic foci. Following the completion of the tasks of the initial interview, the therapy enters the working-through phase. Key to this phase is the ongoing affective experience of personal reactions and conflicts. The tasks of the working-through phase include the continuation of the process of resolution of previously repressed, extensive, murderous impulses and feelings.

The transformation of pathological grief reactions into acute grief processes facilitates their resolution and contributes greatly to clinical improvement.

The process of termination is initiated when there is evidence of symptomatic resolution and of character and functional changes. In complicated cases, mourning the termination of the therapeutic relationship becomes a significant focus and the key to the prevention of repressive dependency.

The working-through phase aims at the systematic resolution of the multiple foci of conflictual painful feeling accessed in the initial interview. The following principles apply to the working-through process:

The therapist should place emphasis on the patient's actual experience of his or her feelings. Feelings such as intense grief reactions, sadness, and rage followed by guilt and remorse appear repeatedly in the psychodynamic focus, facilitating their working through and resolution and promoting healthy adaptive growth.

Often, as a result of the breakdown of resistance systems, the patient experiences vivid reactivation of memories and dreams, which enrich the working-through process, as long as the therapist holds the focus on the direct affective experience of the recovered memories and dreams.

Clinical Issues

INDICATIONS. *Adjustment disorders with depressive mood, with anxiety, and with mixed anxiety and depressed mood; generalized anxiety disorder; panic disorder, selective somatic symptom disorders, Illness anxiety disorders, psychological factors affecting other medical conditions, acute stress disorder. Selective separation anxiety disorders.*

Persistent depressive disorder, dysthymic disorder and major depressive disorder. Treatment resistant anxiety and depressive disorders and chronic pain syndromes.

Major dissociative disorder (depersonalization); social anxiety disorder; obsessive-compulsive disorder; somatic symptom disorder, conversion disorder and hypochondria.

Selected personality disorders such as avoidant and dependent, obsessive compulsive and histrionic personality disorders.

Davanloo, further characterized the core psychodynamic problems presented by patients eligible for treatment through these techniques, on the basis of a descriptive sample of close to 1,000 patients. The large majority of these patients suffered from diffuse and pervasive symptoms such as anxiety, depression, and widespread somatic complaints. They displayed severe maladaptive personality and character problems, leading to significant interpersonal and work problems and to intensely self-sabotaging and self-destructive behaviors. They had clinical histories of overt (and covert) severely traumatic experiences and presented to treatment with high degrees of major resistance. The core psychopathology of these patients was dominated by the presence of highly primitive unconscious rage often directed toward multiple figures of their past who participated in their highly traumatic development or otherwise induced repeated experiences of abandonment and loss. The widespread experience of unconscious rage often led to profound feelings of guilt and grief and commonly promoted masochistic self-destructive lifestyles.

LIMITATIONS. These techniques can provide effective treatment to a large majority of patients suffering from what had been classically called severe psychoneurotic disorder and character disorders. These patients are diagnosed today with Axis I disorders, such as anxiety (including panic) and major depressive disorders, and with personality (predominantly Cluster C) disorders. They represent large cohorts from among the patients who seek treatment in outpatient settings and in private practice and present a serious public mental health burden because of the treatment-resistant nature of their problems and their high utilization of health and mental health resources.

As these techniques are designed to deal with impulse (rage)-based dynamics, they may reach their limits in the growing population of patients suffering from primary self-disorders such as narcissistic and borderline syndromes (self-disorders).

The treatment of self-disordered patients needs to go beyond the resolution of their unconscious rage and secondary guilt and grief, to reach two key components of their disorder: their severely devalued self-images and self-representations, and, more importantly, the healthy and vital self-representations that often lie dormant and repressed under the weight of more consciously experienced negative views of self as such healthy self-representations will make possible the healthy ego development and personal growth that can be induced by treatment.

Techniques for the treatment of these disorders have been proposed, and they are reviewed later in this section.

COMPLICATIONS. The most common complication of ISTDP consists of the development of a protracted or stagnant therapeutic process as a function

of the therapist's inability to break through the highly elaborate structure of resistances presented by the patient at baseline. Transference-based negative therapeutic reactions might ensue and result in a temporary worsening of the presenting symptomatic picture.

CONTRAINDICATIONS. These techniques are not indicated for the treatment of psychiatric disorders such as schizophrenia, substance abuse disorders, antisocial personality disorders, or life-threatening psychosomatic conditions.

The application of ISTDP for psychosomatic and borderline disorders requires modifications of the technique. Selected problems of patients with bipolar disorder (certain maladaptive behaviors, interpersonal problems, and pathological grief reactions) can be treated with these techniques with positive effects on the patients' overall affective functioning.

Davanloo provided numerous detailed case reports that give step-by-step descriptions of the different phases of his therapeutic approach.

The following case was treated following Davanloo's system of ISTDP.

At the time of the first interviews, John was a 35-year-old professional seeking help because of a worsening of lifelong feelings of depression, hopelessness, and impotence. He had an almost delusional conviction that he was destined to catastrophic failure and that he was helpless to change the course of such a fate. A very brilliant man, he struggled to hold back his professional ambition for fear of being perceived as arrogant. He suffered from pervasive feelings of physical and intellectual inferiority and experienced severe loneliness, with the conviction that he would never be able to marry. He believed that no worthy woman would ever show an interest in him.

John was also plagued by episodes of severe anxiety, which became intensified in conflict situations. He feared any interpersonal conflict that could lead to angry words or deeds. In his struggle to suppress any remnant of anger, he had adopted an ego-syntonic, submissive, and self-effacing character structure.

The second of two siblings raised in a middle-class family, John remembered a lifetime of submission to the capricious and sadistic acts of his older brother, who tyrannized him daily through angry words and physical violence. Neither his mother (perceived as distant, sad, and helpless) nor his authoritarian and emotionally removed father ever intervened effectively to curb his brother's relentless aggression. As early as he could remember, John's subjective experience was suffused with terror. Usually, he was afraid of triggering his brother's brutal temper. Occasionally, he feared to lose control of his intense destructive impulses toward his brother. Such impulses were deeply repressed and denied.

The dynamic formulation of John's problems can be represented as follows:

Life problems: Ongoing feelings or depression, sadness, hopelessness, and loss of motivation and interest; deep loss of self-esteem, with feelings of physical and intellectual inferiority; resignation to a life of limited relational and intellectual ambitions; fear of rejection by women and fear and avoidance of falling in love due to the subjective certainty of rejection; and subjective sense of helplessness, worthlessness, and impotence.

Dynamic forces: Struggle to suppress even minimal experience of anger in any context; suppression of assertive wishes and of ambitious hopes and aspirations, and intense guilt and self-reproaches when hoping for assertion and success.

Pathogenic foci: Powerful deep unconscious destructive feelings (including death wishes and murderous impulses) toward his older brother; and parallel destructive impulses toward John's father for his lack of emotional involvement and toward his mother for her failure to prevent the abuse perpetrated by John's older brother. These foci were sharply portrayed via a particularly violent dream that occurred in the early phase of treatment. In this dream "a serial

killer, in a murderous rampage stabbed and killed several members of the same family."

GOALS OF TREATMENT. Davanloo established ambitious goals for his system of ISTDP. He asserted that he aimed at a total resolution of the patient's core psychopathology, leading to a clinical situation in which the patient was symptom free and showed demonstrable structural, character, and behavior changes.

There should be visible and stable changes in the way in which the patient is related to others and used his or her capacity to love and work. A key component of such progress was the patient's often new ability to be in touch with and appropriately express deep feelings.

OUTCOMES AND RESEARCH. Very few studies have been conducted rigorously to assess the efficacy and effectiveness of Davanloo's STDP. In fact, in several meta-analyses of brief dynamic therapy studies conducted between 1978 and 2004, only the studies presented by Winston and colleagues purported to test the efficacy of Davanloo's STDP when compared with an alternative dynamic therapy and a wait-list control. These studies, however, have been widely disseminated, are cited by all selective meta-analytic reviews, and are often praised for their careful scientific design and clinical implementation.

In Winston's 1991 study, 32 patients with diagnoses of personality disorder were randomly assigned to two treatment conditions, and an additional 17 patients were assigned to a wait-list control group. Both therapy conditions were brief dynamic therapies, one being an STDP based on the principles of Davanloo and a second (BAP) drawing heavily on the work Sifneos. At the end of the study, the two therapy groups obtained similar outcomes (significantly better than the wait-list controls) but showed differences in outcomes in symptomatic subscales. Davanloo-based STDP obtained better outcomes on ratings of depression, whereas the Sifneos-based BAP offered better results in ratings of anxiety. Both therapy conditions demonstrated robust and clinically meaningful effect sizes (0.70 to 1.35). An expansion of the earlier study to include 81 patients reported in 1994 obtained similar results.

In a recent report by Allan Abbass and colleagues, 27 patients diagnosed with DSM-IV personality disorders (borderline, obsessive-compulsive, and avoidant personality disorders constituted 65 percent of the sample) were randomized for treatment with ISTDP or to a minimal contact, delayed treatment, control condition.

The ISTDP treatment was provided as outlined by Davanloo, and treatment adherence was carefully controlled. After an average of 27.7 (range 2 to 64) sessions, the ISTDP patients did significantly better than controls in symptomatic and functional measures, and these gains persisted in long-term follow-up. Of note, the treatment sample reached

the normal range on the individual symptomatic scales used (Brief Symptom Inventory and Inventory of Interpersonal Problems), had an 83 percent reduction in personality disorder diagnosis, and achieved large reductions in the use of medications. Just as important, these patients achieved significant reduction in functional disability.

EMERGING SHORT-TERM DYNAMIC PSYCHOTHERAPIES

Accelerated Experiential Dynamic Psychotherapy—Diana Fosha

Theoretical Issues. Building on the work of Sifneos, Malan, Davanloo, and others, Diana Fosha proposed a new model of therapy, which she called accelerated experiential dynamic psychotherapy (AEDP). This approach integrates the technical contributions of its predecessors under the aegis of an affect-centered model of relational psychodynamics; a dynamic which Fosha has later anchored in the neurobiology of attachment and affect.

Whereas the reduction of resistance by breaking through the defense of patients constituted a central tenet of Davanloo's ISTDP, Fosha proposed an alternate route to deal with such defenses: fostering the experience of deep affects and the development of an emotional bond between patient and therapist. Such bonds, by promoting feelings of safety and connectedness in the here and now, are hypothesized ultimately to reduce the need for resistance and defensiveness and to facilitate "corrective emotional" experiences of intimacy and closeness within the patient/therapist relationship.

Fosha's AEDP places the experience of affect, in the context of an emotional relationship, at the center of its understanding of normal psychic development and of the psychotherapeutic process.

Following the teachings of Donald Winnicott and Heinz Kohut, Fosha understands psychic activity as motivated by the person wishing to maximize the integrity and vitality of self-experience and the strength of meaningful attachments. In this framework, affects are, in the context of normality, deep sources of meaning and vitality and, in turn, affect frame experiences of integrity and authenticity. However, they may also become sources of suffering and anxiety when caretakers are actively destructive (through rejection, excessive criticism, or shaming) or passively destructive (by being detached, uninvolved, or absent). Affects experienced under such painful conditions may eventually overwhelm the mastery capacity of the developing ego and may make it necessary, to ensure survival, the use of increasingly complex defense mechanisms. The basic therapeutic stance of the AEDP therapist—supportive, affirming, and emotionally engaged from the initiation of therapy—is intended to create a therapeutic environment in which defenses can be momentarily withdrawn and deep feelings can become available for therapeutic reconstruction.

Techniques. The central goal of AEDP is the mobilization (and working through) of core affects and genuine experiences. According to Fosha, the key mutative agent in AEDP is the state transformation leading to the visceral experience of core affective phenomena within an emotionally engaged dyad.

AEDP aims to facilitate affective experience such as (1) previously unbearable emotional experiences (including grief, anger, sadness, fear, and joy), (2) self-experiences (positive or negative self-images), and (3) relational experiences (feelings of closeness or distance from others) and (4) affects associated with healing and with breakthrough feelings emerging as liberation from emotional traps. To achieve these aims, the AEDP therapist can use what Fosha calls (1) relational techniques (animated by empathy), (2) restructuring interventions, using the classical triangles of conflict and triangle of person to facilitate cognitive understanding, and (3) experimental/affective interventions (facilitating deep, genuine affective experiencing).

Clinical Issues

INDICATIONS. AEDP has been used widely with a variety of outpatients suffering from DSM-IV-TR Axis I disorders (anxiety disorders, dysthymia) and Axis II disorders (avoidant, dependent, and histrionic personality disorders). For patients with more severe psychopathology, such as somatoform, dissociative, and/or borderline disorders, a positive response to a "trial therapy" dose is necessary prior to the conduct of the full therapy course. Success with narcissistic personality disorders has also been reported in selected cases.

LIMITATIONS. Exclusion criteria include all psychotic disorders, bipolar disorders, major depression, impulse disorders, marked acting-out behaviors, and substance abuse disorders.

COMPLICATIONS. No complications have been reported in the AEDP literature.

Psychoanalytic Psychotherapy for Panic Disorders—Barbara Milrod and the Cornell University Group

Theoretical Issues. The psychoanalytic psychotherapy for panic disorder proposed by Barbara Milrod and the Cornell University Group shares common theoretical underpinnings and metapsychology with classical psychoanalysis but differs in its application. For these researchers, panic symptoms carry psychological meanings, and the aim of PFPP is to uncover such unconscious meanings to achieve relief. In addition, their conceptual framework assesses continuous psychodynamic conflicts which tend to be quite specific to patients suffering from panic disorders: conflicts

related to separation and the development of individuality. Anxiety is often aroused in these patients when confronting abandonment or separation threats. Conscious or unconscious anger is also a common trigger of anxiety reactions, including panic.

The theoretical underpinnings of their method (PFPP) incorporate key generally psychoanalytic precepts such as the following:

(1) The role of unconscious fantasies. Unacknowledged (unconscious) rage is considered by Milrod's group to be an important part of the mental life of these patients which—in their cases—relates to the onset of panic.
(2) The common presence of characteristic defenses such as reaction formation, undoing, and denial, which are extensively used by panic patients to minimize and deny anger, and protect their bonds with others.
(3) The central role of transference in organizing the patient's relationship to others, including the therapist. As children, these patients often lived in fearful dependence, plagued by concerns about rejection and/or abandonment. Such concerns fueled repeated experiences of anger, which in turn exacerbated the fear of separation and loss.
(4) The sudden failure of the overburdened system of defenses may lead to the onset of acute panic.

Techniques. Emerging from classic psychoanalytic metapsychology and from the common ground of psychoanalytic psychotherapy, panic-focused psychotherapy is an active, focused, interpretative technique that—at least in research projects—can be carried out in 24 sessions conducted once or twice a week. There are three phases to the therapeutic process: (1) an initial evaluation and early treatment phase normally focused on addressing the acute panic and its surrounding dynamics, (2) a middle phase focused on the underlying core conflicts, and (3) a termination phase usually focused on the patient's substantive conflicts with separation and independence.

Techniques of work vary with the phase of treatment, but in a complete course of therapy most of the classic psychoanalytic techniques are used. Clarifications, interpretations of impulses, and feelings and transference interpretations are covered. Affects (related to anger, interpersonal dependence, and fear of autonomy) are also mobilized and worked through.

Clinical Issues

INDICATIONS. The central indication of PFPP is the treatment of panic disorder, and this is the venue in which panic-focused psychotherapy has established its efficacy in randomized, controlled clinical trials. Many of the

patients included in the reported trials suffered comorbid depression and reported improvement of their depressive symptoms as well, which may suggest the ready applicability of this technique to the treatment of depression with whatever modifications are called for as a function of different psychodynamic configurations.

LIMITATIONS. The most common current limitation of this technique relates to its specific focus or panic disorder. This limitation, however, needs to be balanced by the awareness of the public health effect of the successful treatment of panic disorder, whose patients carry a high suicide risk, suffer high rates of medical and psychiatric comorbidities, and incur very high overall medical costs.

COMPLICATIONS. The reported dropout rate of PFPP is lower than that reported for psychodynamic psychotherapies in general.

Deterioration of clinical status, worsening of panic symptoms, and the development of comorbid depression are also possible complications.

CONTRAINDICATIONS. The presence of active substance abuse, high active suicide risk, bipolar disorder, and psychosis contraindicate the use of this technique. Cognitive impairment of enough magnitude to interfere with learning represents another contraindication.

GOALS OF TREATMENT. Building on the psychodynamic formulation of panic, which understands the disorder as mediated by conflicts related to separation and autonomy and with difficulties with the recognition modulation and expression of anger, the goals of treatment are designed to facilitate the resolution of those conflictual dynamics. Thus, the initial phase of treatment brings into focus detailed and specific exploration of the feelings surrounding the onset of panic and the personal meanings evoked as associated with the symptoms and the panic episode. Goals of the initial phase (phase 1) include providing relief from panic and help to overcome the functional limitations that the disorder may impose on the patient as a function of the use of avoidance mechanisms.

The goals of the middle phase (phase 2) of treatment include highlighting and working through the core dynamics of the disorder: fear of unconscious impulses (anger and sexuality) and feelings of abandonment. The goal here is to desensitize the patient to the anxiogenic effect of angry affects and to facilitate the move from dependency to autonomy while modulating fear of loss and abandonment.

Phase 3 (termination) permits the re-experiencing of conflict about reparation and dependency directly in the therapeutic situation and in relation to the therapist, thus potentially consolidating the therapeutic work of phases 1 and 2. When the treatment is successful, patients report enhanced autonomy, less fear of anger and interpersonal conflict, and

acceptance and enjoyment of independence.

OUTCOMES. PFPP has been tested in an open trial and a randomized, controlled clinical trial, both of which showed significant reduction in panic symptoms. In the second trial, twice as many experimental psychotherapy patients met response criteria at termination, than patients in the control conditions, which consisted of a cognitively enriched form of relaxation therapy. In addition, PFPP showed significantly superior reduction in the severity of a broad range of panic symptoms as measured by the Panic Disorder Severity Scale: 73 percent of PFPP patients versus 39 percent of the controls met the response criteria for panic (40 percent reduction in scale score). PFPP patients also achieved a significant reduction in functional impairment.

Milrod and her group presented a manualized form of psychodynamic psychotherapy for panic disorder, which they tested in a randomized, clinical trial. The treatment, consisting of an open-ended number of sessions in clinical practice and of 24 sessions in the clinical trial, is conceptualized as consisting of three phases:

Phase 1 involves the treatment of acute panic, phase 2 addresses the core dynamic conflicts that support the vulnerability for panic, and phase 3 is the termination phase. The goal of phase 1 is the elucidation of the conscious and unconscious meaning of panic symptom with the goal of obtaining a map of the personal architecture of unconscious conflicts surrounding the panic. Common themes highlight the role of separation and dependence and problems in the modulation of sexual and aggressive impulses. Key techniques in this phase of treatment include keeping the focus on the symptoms of panic, engaging in a detailed and specific exploration of the circumstances surrounding the panic experience, and a providing a firm supportive and reassuring stance that the problem can be treated and the symptoms ameliorated. The goal of phase 2 is the reduction of vulnerability to panic by engaging and working through (often in the transference) additional unconscious conflicts related to dependence and independence and building tolerance for instinctual experiences related to anger and sexuality.

The phase of termination, by bringing to the focus the key vulnerability of these patients (separation experienced as abandonment), provides an opportunity for consolidation of the therapeutic work. Milrod and her group addressed this problem by keeping an active focus on termination for the last one third of the duration of the therapy and by helping the patient work through such fears in the context of the therapeutic relationship.

STDP of Narcissistic and Other Self-Disorders—Manuel Trujillo

Theoretical Issues. The term narcissism has pervaded popular culture

and the sociological and psychiatric clinical literatures for many decades, depicting the increasing isolation and self-oriented individualism of affluent citizens in the United States and other developed countries.

Many features of contemporary culture mirror the narcissistic traits of individuals: the pursuit of materialistic cravings and the instrumental nature of many relationships; the seeking of success for success's sake; and the cultivation of hedonic pleasure rather than the realization of ideals. In such contexts, narcissistic personality traits may be valued, adaptive, and, therefore, preferentially fostered. Witness the success of the "Master of the Universe" meme in the late 20th century.

Freud first referred to narcissism to describe a libidinal position in which cathexes were invested in the subject's ego and not in objects. Under those conditions, true transferences could not develop, making these patients ineligible for psychoanalysis. The clinical expression of non-object libidinal cathexes included psychotic processes.

Dynamically, significant advances have been made over the last three decades in the clarification of the metapsychological dimensions of narcissistic disorders. Two authors, Otto Kernberg and Heinz Kohut, have advanced well-developed metapsychologies for these disorders, and have proposed comprehensive techniques for their treatment through appropriately modified psychoanalytically based interventions.

There are, however, fewer attempts to apply techniques of STDP to patients suffering from self-disorders, despite the fact that such techniques have demonstrated respectable efficacy in the treatment of patients suffering from complex psychoneurotic disorders, selected personality disorders, and clinical conditions such as anxiety and depressive disorders. Patients meeting criteria for these disorders have been shown to make rapid and impressive gains, as long as their dynamic problems are related to difficulties in the management of sexual and/or destructive impulses, problems in the regulation of object loss, or complex combinations of both types of pathology. Patients suffering from extensive self-pathology are typically reported to do less well, largely because the structural underpinnings of self-disorders have not been as well understood as the constructs underlying Oedipal or attachment disorders.

A significant shift in the deep psychological understanding of narcissistic disorders is represented by the work of Kohut and summarized in his book *The Analysis of Self*. Kohut advanced a complete metapsychology of self disorders, including (1) the superordinate position of the self as a whole system organizer to which other mental agencies (e.g., drives, ego, and superego) and mental capacities (e.g., motivation, vitality, goals, and values) are, to different degrees, subordinate; (2) and the relative independence—but rich interdependence—of the development of the self in relation to the development of the other components of the mental apparatus, for example, drives, id, ego, and superego.

Kohut described the relatively independent developmental line of the self and the various disordered responses to failures of environmental support, including parental mirroring. The subjective impact of these failures is registered through deficits and distortions of the perception of self (affecting self-esteem), the relative lack of integration of ambitions and capacities (creating grandiosity, self-effacement, and/or underachievement), or the incomplete development of appropriate, socially accepted values and ideals. For Kohut, repairing these deficits requires facilitating the development of a reparative transference of one of three types: (1) mirror transference proper, involving the use of the therapist as an admiring self-object and/or the resolution of perceived therapeutic failures via empathic interpretations; (2) idealizing transference, involving the acceptance by the therapist of the patient's unconscious needs to see him or her as a quasi-omnipotent object, capable of soothing and consoling and (transferentially speaking) worthy of emulation; and (3) twinship transference, which allows the patient to feel like others and develop a sense of connectedness under conditions of existential equality.

In the area of the self, and thus in narcissistic disorders, the key pathogenic phenomenon is lack of integration; that is, the inability of the mind to maintain a sense of wholeness. The signal of failure to maintain self-homeostasis is the affect of shame. Its presence indicates the inability of the self to maintain sufficient wholeness and fullness. In cases of more severe pathology, fragmentation anxiety denotes the organism's fear for its own basic integrity.

A very important theoretical and practical construct in self-psychology is the developmental process that, when successful, promotes the maintenance of a coherent, whole, optimally altruistic persona. When a "good enough" environmental fit is present, this process, fraught with complexities and opportunities for derailment, facilitates vital human functions such as (1) the slow maturation of a primal boundless, grandiose self into realistic, enduring, and resilient self-esteem, personal identity and grounded ambitions and (2) the gradual transformation of undifferentiated, global, idealizing needs into values, ideals and a moral compass that color and mediate each person's investment in, and involvement with, external social systems of beliefs, meaning, and action.

Narcissistic psychopathology occurs via two interrelated mechanisms: (1) the failure to complete key developmental steps along the pathway just delineated. Such failure may lead to a lack of maturation, integration, and realization of large segments of grandiose self (or idealized object image), may create split-off (or dissociated) self-object representations, which entail potential problems for identity formation for the maintenance of self-esteem, and for the development of an appropriate and satisfying system of values, ideals, and accomplishments; (2) pathological interrelatedness of self-development with drive-developmental processes. In this regard, the

presence of significant unresolved self-developmental pathology may color drive-developmental and interpersonal processes such as separation-individuation, the ability to cope with object losses, and the resolution (or lack thereof) of triangular Oedipal pathology. This clinical situation often leads to hyper sexuality (sex defending from abandonment), the sexualization of violence, and to other compulsive behaviors. Kohut, in fact, stated unequivocally that clinically significant Oedipus complexes occurred because the developing child's normal drives and conflicts were disturbed and intensified by unempathic parental responses.

Once consolidated, narcissistic psychopathology is expressed via: (1) painful and pervasive subjective experience of shame, emptiness, and/or fragmentation anxiety; (2) the presence of extreme intrapsychic defenses (e.g., splitting and cognitive-affective dissociation) and interpersonal defenses (e.g., distancing, feeling extreme superiority or inferiority, intense neediness, or equally unyielding indifference); and (3) symbolic or manipulative behaviors designed to restore the integrity of self by providing soothing of vicarious experiences of wholeness via symbiotic fusion. At times, such soothing is obtained through the use of addictive drugs and or compulsive sexuality.

Techniques. Therapy can be conceptualized as a process that allows the patient—via the development, experience, verbalization, and resolution of intense transferences—to reactivate primitive, dissociated, anxiety-provoking views of self and others. Once activated, a more normal developmental pathway can be resumed in which primitive self-perceptions mature and become more integrated and healthier.

INTERVENTIONS. The paradigmatic intervention for self-problems is empathic understanding coupled with active interpretations of impulses, feelings, and self-perceptions; all in a context of pervasive use of transference. In this realm empathy involves an accurate grasp, both cognitive and affective, of what others experience and its presence creates a supportive milieu. In it, the patient can either bask in the use of the therapist as an optimum self-object or, bitterly but safely, complain about the therapist's failure while benefiting from the healing effects of the possibility of resolving the damage caused by that failure.

To accomplish this task in a relatively short-term program (50 sessions or less) requires the following:

EMPATHY

(1) Accurate, empathically mediated grasp of the patient's core narcissistic vulnerabilities. For many patients, these problems are hidden behind a well-cemented wall of character defenses. Thus, these narcissistic vulnerabilities are apparent only after significant breakthroughs into

the patient's unconscious rage and/or grief. For this part of the process, the ISTDP techniques developed by Sifneos, Malan, and Davanloo are crucial.

TRANSFERENCE FOCUS

(2) Active and early attention to the development of various types of self-object transferences.

ACTIVATION OF PRIMITIVE SELF-STATES

(3) Early detection and reactivation (largely through active inquiry) of primitive, dissociated, grand, or grandiose self-images and the containment of defensive affects, such as dissociative anxiety, fragmentation anxiety, and pervasive sense of embarrassment and shame. A simple inquiry such as "and how do you see yourself as a person now?" can bring up a mélange of primitive, childlike, exaggerated or minimized view of self with a poorly differentiated mixture of affects such as shame, fear, pride, and others. Once all those feelings are "in the open," the patient can do the necessary work of pruning and shaping such personal (at times idiosyncratic) views of self in order to carve a more grounded and satisfying personal identity.

IMAGINARY RECONSTRUCTION

(4) A special form of corrective emotional experience called "imaginary reconstruction." This intervention permits the soothing and healing of narcissistic wounds via the imaginary reconstruction, following a painful shameful, often catastrophic experience or failure of self, or of a valued other. At these moments, the therapist inquiry of an imagined optimum (not perfect, but not wounding) version of the "failure narrative" often provides balsam to the narcissistic wounds. A few of such interventions are provided in the case description below.

Just as the course of an episode of ISTDP treatment can be depicted by the use of the two triangles discussed earlier, the course of an episode of treatment of a (pure or mixed) self-disorder problem can be depicted as shown in Figure 33.10–4.

The typical therapy process focuses first—and repeatedly—on the overt manifestations of defensive activities, such as character defenses buttressed by splitting and cognitive–affective dissociation. Subsequent and repeated therapeutic challenges to the defenses may prompt the emergence of anxiety or paralyzing shame, and finally lead to the reactivation of formerly repressed and distorted self-representation. Distorted negative self-images are then deactivated via this exposure because the patients can see themselves more realistically and can experience and resolve the emerging grief. Positive self-images are allowed again to flow through the barriers of

dissociation and are reactivated, owned (experienced by patients as rightfully belonging to them) and appropriately linked to hopes, dreams, skills, and values. In this way, profoundly personal projects of much meaning to the patient can re-emerge as valued blueprints for purposeful and meaningful life plans and actions.

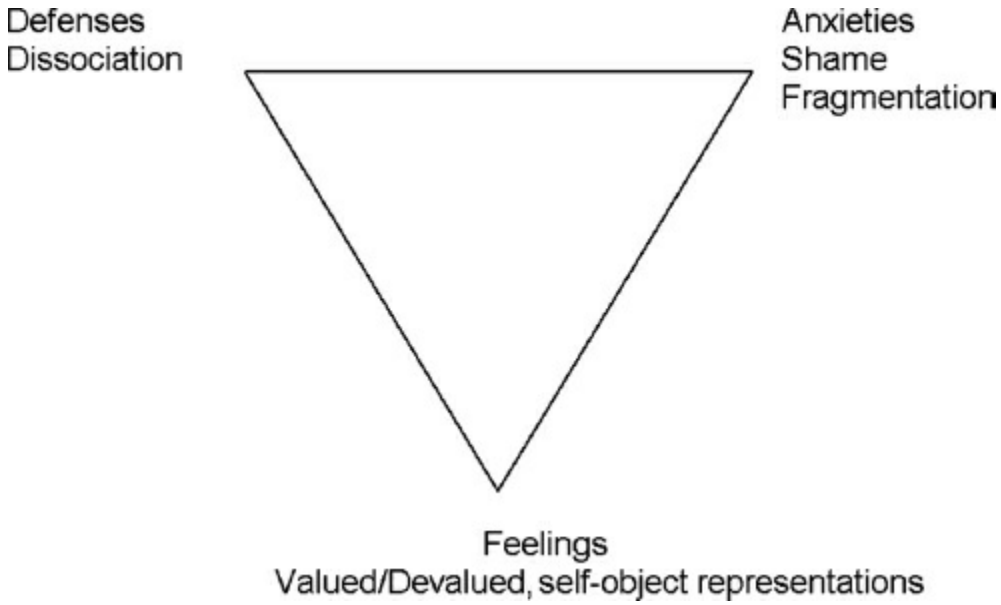

FIGURE 33.10–4. Triangle of the self.

WORKING THROUGH OF RAGE. The technique of working through narcissistic rage requires a sequence of interventions that achieve (1) keeping the focus on the rage, for example, a specific memory of rage in a patient's interaction with the patient's father, (2) reactivation of the physiological–emotional–cognitive continuum underlying the rage response to allow the patient a simultaneous visceral and cognitive experience of the rage, (3) pressure and challenges to the defenses against rage to ensure a fuller experience, and (4) pressure and challenges to the defenses against interpersonal closeness and, especially, closeness with the therapist. As the therapist acquires, via the transference, a certain standing in the patient's emotional life, the attachment to the therapist heightens the patient's ambivalent attachment to the primary object and intensifies the patient's unconscious conflict over emotional loyalty. This is yet another source of considerable resistance and active therapeutic work. Rage toward the therapist may reach murderous proportions.

REACTIVATION OF SELF

Reactivation and Integration of Self. As patients gain mastery and knowledge of their experience of rage and the pathological grief is lessened, they are often ready to reactivate grand and grandiose self-images, and to re-initiate and complete the reintegration of such self-images both with reality and

with the person's unique set of skills and personal capacities. An enhanced sense of serenity and self-regard, a capacity for confident performance of work and play activities, and increased interpersonal empathy and closeness are usual results of such processes. The work of reactivation of primal, dissociated self-images also requires clarification and correction of multiple self-distortions. For example, for GR—the patient to be presented next case—a rather large, athletic, and physically powerful man, the self-image of utter fragility as a premature, inhumanly small baby in danger of being trampled underfoot at any moment, contaminated the way he saw himself and acted physically. His hunched over posture, made him appear considerably smaller than he actually was. As the therapy advanced, he started to derive considerable pride in his body, his physical prowess, and his dress style, for which he often sought the therapist's admiration and approval. Increasingly, his dreams and fantasy life filled with images of grandness ("I feel like the captain of my own ship"), pride, and pleasure at his own physical, mental, and emotional characteristics. The resolution and integration of this sequence of work required a process akin to a recapitulation of the developmental work of adolescence. Much as in late adolescence, through newly developed capacities for introspection, one takes some ownership and responsibility for one's self and destiny, patients often experience a mini-recapitulation of this developmentally appropriate sequence. There is often a lessening of the earlier dependence on external mirroring and a sense of pride and pleasure at discovering, defining and taking responsibility for one's unique personal assets and aspects; assets that can now be used to create a life of personal purpose.

Clinical Issues

INDICATIONS. *Adjustment disorders with depressive mood, with anxiety, and with mixed anxiety and depressed mood; generalized anxiety disorder; panic disorder with agoraphobia and without agoraphobia; dysthymic disorder; major depressive disorder; depressive disorder (recurrent).*

Narcissistic personality disorder and selected borderline personality disorder as well as the self-disorder core of many other clinical conditions such as eating Disorders, certain addictions, and others.

Experienced self-psychologists recognize the following additional syndromic presentations:

Primary Self-Disorders

(1) Primary, overt, self-disorder: the narcissistic personality disorder, as defined by the DSM-5 (Table 33.10–1).

(2) Covert narcissistic personality disorder, representing a mirror image of the narcissistic personality disorder described above. Features that are overt in narcissistic personality disorder proper, are often covert and

hidden here, but can be detected either in the patient abundant fantasy life, or hiding behind considerable resistance (Table 33.10–2).

(3) Mixed self-disorders: disorders that reflect a mixture of narcissistic traits and conflict/drive problems. Narcissistic vulnerabilities and traits may fuel and compound many classic clinical conditions disorders, such as anxiety disorders, depressive disorders, and eating disorders.

(4) Disorders mixed with other personality disorders. Pronounced self-disorder traits may color the clinical presentation and outcome of other personality disorders, such as avoidant and dependent personality disorders.

Table 33.10–1.
Diagnostic Criteria for Narcissistic Personality Disorder 301.81

A pervasive pattern of grandiosity (in fantasy or behavior), need for admiration, and lack of empathy, beginning by early adulthood and present in a variety of contexts, as indicated by five (or more) of the following:

- Has a grandiose sense of self-importance (e.g., exaggerates achievements and talents, expects to be recognized as superior without commensurate achievements).
- Is preoccupied with fantasies of unlimited success, power, brilliance, beauty, or ideal love.
- Believes that he or she is "special" and unique and can only be understood by, or should associate with, other special or high-status people (or institution).
- Requires excessive admiration.
- Has a sense of entitlement, i.e., unreasonable expectations of especially favorable treatment or automatic compliance with his or her expectations.
- Is interpersonally exploitative, i.e., takes advantage of others to achieve his or her ends.
- Lacks empathy; is unwilling to recognize or identify with the feelings and needs of others.
- Is often envious of others or believes that others are envious of him or her.
- Shows arrogant, haughty behaviors or attitudes.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Table 33.10–2.
Diagnostic Presentation for Narcissistic Personality Disorder (Covert Type)

A pervasive pattern of grandiosity (in fantasy, not in overt behavior), need for admiration, excessive sensitivity to judgment and opinion of others, beginning by early adulthood and present in a variety of contexts, as indicated by:

- A covert grandiose sense of self-importance, although may compulsively minimize achievement.
- Fantasies of unlimited success, power, brilliance, beauty, or ideal love often hidden under considerable shame.
- A craving to be noted as special, but fearing automatic rejection by other high-status people or institution.
- Need for excessive admiration.
- Behavior as if driven by a sense of entitlement; i.e., easily hurt and disappointed when high expectations for favorable treatment or automatic compliance are not met.
- Compulsively seeking others as admiring objects.

Excessively focused on the behavior of others but unwilling to recognize others' independent needs.
Often envious of others or believing that others are envious of him or her.
Hidden arrogant behavior or attitudes.

Secondary Self-Disorders

(1) Disorders in which self-esteem and self-representation may be impaired through the effect of functional biologically based illnesses such as schizophrenia and bipolar disorders as they interact with the normal development of self.
(2) Disorders in which self-esteem and self-representation may be impaired through the effect of structural, biologically based disorders such as attention-deficit/hyperactivity disorder in the normal process of development of the self.

The importance of secondary self-disorders lies in the fact that, if unattended, their presence may complicate the treatment and outcome of the primary disorder and force the patient to live with unresolved unconscious, defective, and pathogenic self-images. Each clinical presentation requires different interventions, priorities, and techniques.

LIMITATIONS. In addition to narcissistic personality disorders, many other psychiatric conditions affect significant aspects of the structure and/or development of the self. Patients suffering from such disorders as schizophrenia and bipolar disorders do indeed suffer impaired self-experiences and self-representations and object representations. In these conditions, such views of self and others, develop through neurobiological alterations that may impede the structural integrity of the self. Other patients, such as those with antisocial or borderline personality disorders, may suffer from lack of ego integration or from a structural inability to develop the minimum of empathic relatedness necessary for therapy.

COMPLICATIONS. The most common complication is the development of a relatively stable transference neurosis, which requires additional time for its resolution. Negative therapeutic reactions in response to actual or perceived narcissistic injuries have been described.

GR was a 44-year-old unmarried professional man who first consulted because of unbearable anxiety, depression, and loneliness. Although periodically present throughout his life, these feelings appeared in a clinically explosive form at the end of a troubled 6-month relationship with Ms. B, a divorced professional woman who was the mother of a 10-year-old girl. Ms. B ended the relationship after a few months of sustained—if unsuccessful—efforts to please GR and cater to his many needs and demands. His escalating and often frantic demands reflected a deep and potent craving for perfectly attuned attention and responsiveness to the endless flow of his physical and emotional needs. Despite all of Ms. B's efforts (now painfully recognized by the patient), he usually treated her with a mixture of cruel devaluation and condescension, forcing her to admit, in practically every encounter, her incredible luck at being with a man with his "unique" sensitivity and passion and her pitiable incompetence at "adequately" responding to him.

Dynamically, his conscious life was dominated by the need to enact scenarios of uniqueness and specialness, in a never ending quest for positive confirmation from whomever was chosen to function as a confirming self-object. Failures in this regard met with a cascade of subjective effects such as unbearable anxiety, pervasive rage, and fury, which often led to prolonged, sadistic devaluing and attacks on the hapless offender. When such a person left GR's presence, his affect shifted to intense loneliness and profound depression, often requiring immediate medical attention.

The dynamic picture for patients like this can be represented as follows:

(1) Ongoing, somewhat unstable conscious experience of entitled grandness or frank grandiosity, with large segments of the patient's adaptive behavior devoted to seeking opportunities for confirmation of the patient's grand(iose) view of self and to searching for, and holding onto, complementing, confirming self-objects.

(2) Storms of primitive anxiety and despair often accompanying the experience of narcissistic traumas and self-object rejections. This may include painful episodes of fragmentation anxiety, depression, and self-devaluation and prolonged experiences of emptiness. When these subjective states coincide with serious professional, health, or financial setbacks the risk of suicide is high.

(3) An unconscious life dominated by negative self-images and expectations of rejection and devaluation. Such experiences color patients' conscious fears and frequently populate their dreams.

GOALS OF TREATMENT. At the end of the evaluation phase, GR's dynamic hologram appeared to have the following configuration:

Problems

Inability to achieve reasonable success despite evident intellectual assets.

History of stormy failed relationships with multiple women despite a craving to achieve stability, be married, and have a family.

Chaotic life dominated by a craving for emotionally intense professional and emotional relationships (self-object-like).

Dynamics

Ongoing and intense need for outside emotional supplies. The need to search for the perfect (i.e., all-admiring, all-giving, and nurturing) partner dominated his life to the practical exclusion of more mature relatedness.

Primitive rage and vindictiveness triggered by perceived or actual failure of his partners to meet his unexpressed needs.

Ongoing professional and personal self-sabotage.

Core conflicts

Profound repressed primitive, ambivalent feelings toward the maternal object, resulting in an insoluble tension between libidinal investment and sadistic-destructive derivatives in his actual relationships with women. Thus, his relationships with women typically started on a high of idealization and, just as often, crashed in a low of violent

recriminations.

Impulse/Feeling: Ambivalence toward paternal object complicated by unresolved pathological grief, with the consequent disruption of the development of dynamically equivalent actual adult relationships.

Self-representations: To the degree that adaptive and realistic maturation of primal grand self-representations was delayed, the patient showed incongruent behavior: The grandiose self-images were acted out in consciousness in the form of perfectionistic demands on others while the helpless self-fed enormous self-doubts and neediness.

DYNAMIC HOLOGRAM OR ROADMAP. This dynamic hologram was obtained in the initial evaluation, at a time when GR was often flooded with anger, grief, and other painful feelings. The hologram represents a roadmap for the unfolding of the therapeutic experience. For this patient, therapy will proceed through repeated runs around the triangle of impulses, person, and self. The two key pathogenic foci for him are (1) the focus of self and (2) the focus of unresolved grief for the death of his father. The pathogenic capacity of each focus reinforces the pathogenic capacity of the other. The images of helplessness and utter fragility derived from his repressed disturbed negative self-images are generated first by identification with the failed aspects of his father's life and function as a way to maintain some bond with the father. Such bond is also reinforced by his need to expiate guilt for his unconscious murderous impulses toward his father. The trigger for such destructive feelings relates to his father's failure to act as an adequate narcissistic self-object (optimum mirror), as his father, probably immersed in his own pervasive sense of failure, was unable to recognize GR's person, accomplishments and inclinations.

The working through of these problems will require revisits to the experience of anger and destruction in relationship to the people in his current life (C) or toward the therapist (T) and subsequently toward his father (P). Multiple TCP interpretations were made before a sense of peace and emotional understanding and a modicum of closure and closeness obtained for the patient in relationship to memories of his father. Simultaneously, expressed and covert demands for closeness, support, feedback, extra time, and extra resources showed up in the therapeutic situation. The slightest disappointment in the therapist's responsiveness to such needs, evoked feelings of desolation and devastation in the session.

The previously described technique of imaginary reconstruction allowed the patient to evoke emotionally satisfying moments of perceived optimum mirroring from the therapist, often following the expression of intense anger toward the therapist as a function of real or perceived failures on the therapist part. After one such episode, after the patient bitterly berated the therapist for not responding in the way he "needed" during a telephone

interchange, the following dialogue ensued:

CLINICAL VIGNETTE

Therapist: You were obviously very disappointed with my response. What would you have wished me to say?

Patient: You know—something warm, something supportive, like "Of course, G[. . .], she should have agreed with you."

Therapist: And how would you have felt if I said that?

Patient: Relaxed, accepted, or, like I have value.

Therapist: And how do you imagine I feel toward you?

Patient: Never thought of that—I think you want me to succeed, I think [with emotional conviction] you love me!

Therapist: And how do you feel about that?

Patient: Warm—Great it makes me think of my father. I think that if he would have been able to, he definitely would have loved me, how sad!

Therapist: What would he have loved especially about you?

Patient: That I am sensitive, that I am creative, that I am funny.

Therapist: Because—How do you see yourself as a person now?

Patient: Just like that: Funny, creative.

The ultimate goal of the therapy is the rapid reactivation of formerly dissociated positive self-images and their fuller reintegration with such dimensions of the self as hopes, dreams, and aspirations for the future and such means of self-realization as actual skills and capacities. This aim may be succinctly expressed by paraphrasing Freud's well-known dictum "Where id was, there ego shall be" as "where emptiness is, there self will be."

Mutually agreed termination of therapy with patients suffering from narcissistic personality disorder or substantive narcissistic traits requires the fulfillment of the following criteria: Significant reduction of initial clinical symptoms and substantive lessening of the interpersonal and work dysfunctions presented at the initial evaluation. From the dynamic viewpoint, a sense of agency and self-possession often appears de novo, partially replacing the overwhelming sensitivity of these patients to the negative behaviors of others.

OUTCOMES. The outcome of the treatment of these disorders has not been tested through appropriately designed research studies. Clinicians treating many of these patients report substantial gains in symptoms, interpersonal function, personal integrity, and object-relatedness. Characteristics of the treatment process, such as short duration, enhanced activity of the therapist, development of treatment foci, use of transference, and mobilization of intense affects, have been satisfactorily researched in

many studies of the efficacy and effectiveness of STDP in the treatment of severe psychoneurotic and personality disorders. These studies represent a good background for the design of future studies of the treatment of self-disorders.

By the 1950s, the randomized clinical trial became the gold standard of studies purporting to demonstrate the efficacy and effectiveness of medical interventions. Since then, these trials, which have guided the development of many such medical interventions, have slowly been introduced for verbal therapies as well. By the 1970s sufficient numbers of well-conducted studies permitted the meta-analysis by Smith, Glass, and Miller that demonstrated the general effectiveness of psychotherapy when applied to the treatment a broad range of conditions. More recently, the efficacy of psychotherapy has been demonstrated for various anxiety disorders, depression, certain personality disorders, including borderline personality disorder, and selected eating and addictive disorders.

Smith, Glass, and Miller studied the outcome of psychotherapy in 475 controlled studies that fulfilled preset criteria in terms of diagnostic validity, specification of the interventions provided in both the experimental and control conditions, and the systematic measurement of outcome. The main finding established that typical patients treated in the active experimental arm of the clinical trials were significantly better than 75 percent of untreated controls. The therapeutic gains reported at the conclusion of the therapy were maintained at 2-year follow-up. Both behavior and psychodynamic therapies appeared to be superior to other therapies.

No separate analysis was conducted in this review of the efficacy of STDPs. Falk Leichsenring and colleagues to assess the efficacy of STDP in treating specific psychiatric disorders conducted a more recent meta-analysis. The review included studies conducted between 1970 and 2004 and were selected for meeting stringent criteria, such as randomized, controlled trials, therapies specified in treatment manuals, verification of treatment integrity, training of therapists to preset proficiency standards, and the use of reliable and valid diagnostic and outcome measures. Core features of the STDP selected included increased activity of the therapist, focus on the core themes and conflicts defined in the initial evaluation, mobilization of affects, and active use of transference with particular emphasis on the here and now.

Table 33.10–3.
Summary of Effect of Short-Term Dynamic Psychotherapy (STDP)

Mean Effect Site	STDP Term/Follow-Up	Usual Treatment Term/Follow-Up	Wait-List Term/Follow-Up
Target problems	1.39–1.57	0.55–0.84	0.27–NA

Symptoms	0.90–0.95	0.22–0.24	0.12–NA
Social function	0.80–1.19	0.38–0.95	0.21–NA

In this meta-analysis, STDP yielded large and significant effect sizes for target problems (1.39), psychiatric symptoms such as anxiety and depression (0.90), and social functioning (0.80), effects that significantly exceeded those attained by controls consisting of wait-list and treatment as usual. Of note, patients in the active treatment arm of the studies continued to achieve clinical improvement between termination and 1- to 7-year follow-up, as reflected in Table 33.10–3.

Outcome of STDP in Specific Psychiatric Disorders

Anxiety Disorders. Although there are far fewer studies of STDP applied to the treatment of anxiety disorders than trials of CBT, those studies provide some evidence that STDP is effective in the treatment of generalized anxiety disorder, social phobia, panic disorder, and posttraumatic stress disorder. Some of these studies included comparisons between psychodynamic psychotherapy and CBT, and in most of these the two methods of treatment were reported as being equally effective.

Generalized Anxiety Disorder. Treatments developed for GAD are represented by cognitive therapy and behavioral therapies focused on anxiety management (biofeedback, relaxation). There are few studies that include well-constructed STDP therapies in the comparison arm. The results of these few studies tend to favor CBT over dynamic interventions. One well-conducted study that included 99 patients randomly assigned to CBT or analytic psychotherapy showed the effect of CBT to be superior at termination but not at 1-year follow-up.

Panic Disorder. There are few studies of brief dynamic psychotherapy for panic. Milrod's group conducted several trials of manualized PFPP. In an open study of psychoanalytic psychotherapy that included 21 patients seen twice a week for 12 weeks and followed-up at 6 months, they reported remission rates of 76 percent for the total sample and 93 percent for those who completed treatment.

The Helsinki Psychotherapy Study, a well-designed randomized study of the psychotherapeutic treatment of mood disorders, included a sizable number of patients suffering from panic disorder and generalized anxiety disorder. This study compared STDP (modeled after Malan's approach) with problem-solving, solution-focused therapy (SFT). In this study, 56 percent of patients with anxiety disorder did not meet criteria for any anxiety disorder after 7 months of STDP, compared with 42 percent of those in the SFT group. At 12 months, STDP patients had experienced a 34 percent reduction in the Hamilton Anxiety Rating Scale, compared with a

28 percent decrease achieved by SFT patients.

In a second study, a randomized, controlled trial of psychoanalytic psychotherapy was compared with applied relaxation training. Subjects treated with psychotherapy had significantly greater reduction in severity of panic symptoms and manifested improved psychosocial functioning. The effect size for symptomatic improvement was a robust 0.95 and for disability improvement was 0.74. Significant reductions in the score for depression, based on the Hamilton Scale for Depression, were also achieved, with a respectable 0.53 effect size.

DEPRESSION. A sizable number of reviews of psychodynamic psychotherapy trials for the treatment of depression is available. A review by Leichsenring in 2001 comparing STDP and CBT concluded that the clinical results achieved were not substantially different between the techniques. AA Abbass and collaborators, in a 2014 study sponsored by the Cochrane Library, reviewed all published studies in which dynamic psychotherapy lasting less than 40 hours were included.

The authors included 33 studies involving 2,173 participants and measured symptom reductions across somatic, general, anxiety and depression domains, as well as social function and interpersonal problems. Despite the studies diversity, and except for the somatic dimension, all other measures showed substantial superiority of the STDP arm of the studies over the control conditions in the short and the long term. Effect sizes increased over the long term.

PERSONALITY DISORDERS. Personality disorders represent a major challenge for clinical practice and for research due to their heterogeneity and the high level of comorbidities they share with Axis I and other Axis II disorders.

A specific meta-analysis of psychodynamic psychotherapy applied to personality disorders conducted by Leichsenring and Leibung showed robust gains in outcome domains such as target complaints, broad-based symptoms, and social functioning but did not demonstrate large outcome differences between the STDP and alternative CBT therapies. The effect size for psychodynamic psychotherapy was 1.31 versus 0.95 for CBT.

A study by Martin Svartberg and colleagues comparing STDP to CBT for Cluster C personality disorders with anxiety and depression comorbidities reported significant improvement at termination and 2-year follow-up. At the 2-year follow-up, 56 percent of patients treated in the STDP group and 42 percent of CBT patients had returned to community standards of normal functioning.

A recent survey by Abbass and colleagues evaluated the efficacy and long-term effectiveness of ISTDP in the treatment of patients with DSM-IV-TR personality disorders. The study compared Davanloo's ISTDP to a minimal-contact controlled condition. The most common personality

disorders included in the treatment sample (borderline, obsessive-compulsive, and avoidant) corresponded to reported prevalence rates in clinical samples drawn from broadly representative outpatient facilities. The sample also exhibited high levels of interpersonal problems, work disability, and use of psychotropic medications, thus qualifying as a severe sample. The ISTDP-treated group showed superior outcomes relative to controls in measures related to symptoms, social functioning, return to work from disability, and reduced use of psychotropic medications. Most participants (86 percent) no longer met criteria for any personality disorder at the completion of treatment or at the follow-up assessment.

ISTDP in Other Psychiatric Disorders. The development of manualized therapies based on the psychoanalytic psychodynamic framework has fostered their application in therapeutic trials designed to assess ISTDP's efficacy when applied to a broad range of psychiatric disorders. A study of STDP in anorexia nervosa compared it with dietary advice and found no difference in body weight at termination and 1-year follow-up, although ISTDP patients reported better social and sexual adjustment. A larger trial compared ISTDP (based on Malan's approach) with specialized family therapy, cognitive therapy, and standard psychiatric care. Within a context of modest gains, all therapy conditions performed better than the control.

Although there have been a modest number of trials of ISTDP for alcohol dependence and abuse, there is insufficient evidence to support psychodynamic psychotherapy as a robust first-line therapy for these patients. Trials with cocaine-dependent patients were also inconclusive.

A study by Woody and colleagues compared a dynamic (supportive-expressive) psychotherapy condition with CBT and with drug counseling in opiate-dependent individuals maintained on methadone. Patients in all three groups showed improvements in lessened drug use, reduced criminal behavior, and improved psychological functioning. STDP showed better results than CBT in reduction in drug use, improvement in work status, and lessening of psychological problems, whereas CBT improved the patients' ability to deal with their legal problems.

In a multicenter prospective trial, 157 breast cancer patients with comorbid depression were randomized to either individual STPP (intervention group, $N = 78$) or "treatment as usual" (control group, TAU, $N = 79$). As the primary outcome measure, the authors hypothesized a higher rate of remission defined as no diagnosis of depression (Structured Clinical Interview for DSM-IV) and reduction in depression score by at least 2 points (Hospital Anxiety and Depression Scale, HADS-D) in STPP versus TAU at treatment termination. Secondary outcomes mainly refer to quality of life (QoL).

In the intention to treat (ITT) analysis, 44 percent of the STPP group achieved highly significantly more remission than TAU (23 percent). STPP

treatment (OR = 7.64; $p < 0.001$) was the strongest predictor for remission post-treatment; time was also significant (OR = 0.96; $p < 0.05$). A high effect favoring STPP ($d = 0.82$) was observed for the HADS-D score post-treatment (secondary outcome). Regarding further secondary outcomes (QoL), analyses of covariance yielded main effects for group (favoring STPP with an effect size of at least $d = 0.5$) for global QoL, role, emotional and social functioning, pain, treatment side effects, breast symptoms, and upset by hair loss. STPP is an effective treatment of a broad range of depressive conditions in breast cancer patients improving depression and functional QoL. The interpretations of these findings are limited by the drop-out rate (~1/3) and by delayed post-treatment assessments. Future trials may consider stepped-care approaches, tailored to patients' needs and requirements in the acute treatment phase.

COMBINED ISTDP AND MEDICATIONS. Many trials combining specific psychopharmacological agents and various forms of brief dynamic psychotherapy and CBT have supported the superiority of the combined condition over each intervention provided alone. In the treatment of panic disorder, Wiborg and Dahl studied the effect of adding psychodynamic psychotherapy to treatment with the tricyclic antidepressant clomipramine. The addition of the therapy (a manualized blend of Davanloo, Malan, and Strupp and Binder's work) significantly strengthened key therapeutic outcomes. At termination, all patients in the combined therapy conditions were free of panic. Patients in the combined therapy arm suffered far fewer relapses at 18-month follow-up (20 vs. 75 percent for clomipramine alone).

Regarding depression, many studies comparing psychodynamic therapy with medication have established comparable efficacy for psychotherapy and medication and the superiority of combined therapies over the medication-alone condition.

PROCESS-OUTCOME CORRELATIONS'MECHANISM OF CHANGE. Beyond established efficacy and effectiveness for specific psychiatric disorders, therapies classifiable under the umbrella term ISTDP need to refine their knowledge regarding what specific ingredients of the proposed therapy are responsible for what aspects of the process of change that determines the outcome of the intervention. Among the ingredients mentioned in the spectrum of therapies described in this section, ingredients such as the enhanced activity of the therapist, the maintenance of a relevant focus, the mobilization of unconscious affects, and the use of the transference are most often mentioned as relevant to the therapeutic outcomes obtained.

The therapeutic efficacy of those factors has been supported by well-conducted process—outcome correlation studies and selective meta-analyses. Diener and colleagues systematically examined the relationship of emotional experience and expression and the therapeutic outcome of

dynamic psychotherapy through a meta-analysis of 10 independent studies. By calculating an overall effect size of the relationship between therapist-directed focus on affect and the outcome of psychotherapy, they concluded that in the sample analyzed, such focus on affect was associated with considerable patient improvement. The overall average weighted effect size across all outcomes measured was reported as significant (r = 0.30), suggesting a 30 percent difference in success rate for patients who received an affective therapeutic focus.

Regarding transference, transference interpretations are generally assumed to be a key treatment ingredient for psychoanalytically oriented short-term psychotherapy. Yet, despite their theoretical importance, few empirical studies have attempted to measure the long-term effects of transference interpretations in brief dynamic psychotherapy.

Glund and coworkers conducted such a study by randomly assigning 100 patients to brief dynamic psychotherapy with or without transference interpretations. Although both treatments demonstrated significant improvement after treatment termination, patients with a lifelong pattern of poor object relations benefited more from the 1-year transference-based condition than from therapy without transference interpretations, an effect that was sustained through the 4-year follow-up period. The findings of this study suggest that transference interpretations may be especially suited to improving the interpersonal functioning of patients with prior histories of poor object relationships.

Review of the growing literature on the application of brief dynamic psychotherapy to many specific psychiatric disorders has broadened the original findings of the meta-analysis of Smith, Glass, and Miller. Beyond their conclusion that the typical psychotherapy patient is better off than the untreated control at termination of the treatment, it is now possible to assert the specific efficacy of dynamic approaches when treating anxiety disorders, depression, a broad range of personality disorders and other selected eating disorders, and substance abuse conditions. Efficacy has also been established for the combination of psychodynamic therapies with medication. The challenges for future outcome research include identifying precise and specific mechanisms of change, generating additional innovative treatments for resistant conditions, and refining the technical specificity of existing psychodynamic techniques.

Psychotherapy and Neurobiology

It is to Freud's everlasting credit that psychoanalysis transformed the current understanding of mental life by contributing, for the first time in history, a scientifically grounded theory of mind. Unfortunately, his deeper ambition—to see the grand edifice of psychoanalysis anchored in the solid foundation of neurobiology—has not yet been realized. Freud was forced to abandon such ambitions, expressed in his "Project for a Scientific

Psychology," because of the immaturity of the brain science of his time, among many other reasons. In "Beyond the Pleasure Principle" Freud asserted "The deficiencies in our description would probably vanish if we were already in a position to replace the psychological terms with physiological or chemical ones. We may expect physiology and chemistry to give the most surprising information."

He also asserted in his paper "On Narcissism: An Introduction" that "we must recollect that all of our provisional ideas in psychology will presumably one day be based on an organic substructure."

The absence of a sound neurobiological base with which to link his emerging discoveries about the structure of mental life forced Freud to abandon his early ambition and to concentrate his efforts on the development of a psychoanalytic metapsychology that has stood, independent of its biological underpinnings, for more than a century and has produced abundant clinical and metapsychological fruits.

The growth of neural science in general and of cognitive neuroscience and neuropsychology in particular affords contemporary psychoanalysis a second chance to anchor classical metapsychology in a newly evolving neural science.

The following findings are highly suggestive of potential avenues for the construction of a neurobiologically anchored psychological science. Such pathways include the capacity for verbal therapies to modulate basic neurophysiological brain functions and alter serotonergic function, modify synaptic plasticity, and changing the activation of key brain circuits in areas such as the anterior cingulated cortex, various limbic structures, the prefrontal cortex, and other brain centers that play key roles in the processing of key emotions and views of self and others.

Psychotherapy and the Modulation of Brain Function. The work of Kathleen Shear in the early 1990s demonstrated that the application of CBT to patients suffering from panic disorder appeared to reverse their pretreatment vulnerability to lactate-induced panic, raising the possibility that CBT, at a minimum, modulated the biological cascade of events that trigger a panic episode in vulnerable patients following the infusion of sodium lactate and thus providing evidence of the biological effect of talk therapies.

EFFECTS OF VERBAL THERAPIES ON KEY NEUROTRANSMITTERS: SEROTONIN. More evidence for the robust biological effects of verbal therapies was subsequently produced by L. R. Baxter Jr. and his collaborators. In a study published in the Archives of General Psychiatry in 1992, Baxter investigated local cerebral metabolic rates for glucose in patients with OCDs before and after treatment with either fluoxetine or behavior therapy. After treatment, the investigators observed significant reductions

in glucose metabolic rates in the head of the right caudate nucleus, a reduction not observed in nonresponding patients or in normal controls. In considering how behavior therapy could produce brain function changes similar to those caused by neurochemically specific drugs, the authors invoked the findings of Eric Kandel, who demonstrated, in the sea slug Aplysia the changes at synapses that use serotonin to mediate learned changes in stimulus-response behavior.

It is thus possible that both medication and behavior therapy achieve their therapeutic effect via the modulation of the availability of serotonin to the synapse. Although medication is reported to act by blocking the reuptake of serotonin, behavior therapy—after all, a specialized form of learning—may achieve its therapeutic effect by enhancing the functional availability of serotonin at the synapse level.

Also in the late 1990s, Heimo Viinamäki demonstrated, via single-proton emission computed tomography, the normalization of serotonin transporter densities in the medial prefrontal cortex, midbrain, thalamus, and striatum in a 25-year-old patient suffering from depression and personality disorder who received 1 year of once-a-week psychodynamic psychotherapy.

Psychotherapy and Synaptic Plasticity. A study by Arthur L. Brody and colleagues demonstrated that subjects with major depressive disorder had regional brain metabolic abnormalities at baseline that seemed to normalize equally with the application of either of two forms of treatment: the antidepressant Paroxetine and interpersonal psychotherapy (ITP). In testing the hypothesis that in subjects with major depressive disorder, metabolism would (with treatment) increase in the dorsolateral prefrontal cortex (DLPFC), and decrease in the ventrolateral prefrontal cortex, the authors found that both paroxetine and ITP normalized the metabolic abnormalities defined at baseline. Specifically, both groups showed decreases in prefrontal cortex and left anterior cingulate gyrus metabolism, and increases in left temporal lobe metabolism. In discussing the potential mechanisms of action of ITP the authors invoked changes in synaptic plasticity through learning-induced retraining of implicit memory systems.

Because of ITP's special focus on improved socialization, areas of the brain associated with that function may undergo an attenuation of neuronal connectivity during such treatment. Just as signs of distress in socially isolated animals correlate with increased metabolic activity in the cingulate cortex, improvements in the socialization of depressed patients achieved through ITP may lead to reductions in the activity of the anterior cingulate cortex and a corresponding amelioration of depression-induced negative evaluation of self and others.

A more recent study by S. Saxena and collaborators reported findings that establish the intriguing possibility that psychotherapy (CBT in their

experiment) may share with medication some of their effects in brain physiology while producing their own specific effects in addition. In seeking to elucidate the nature of the brain response to brief intensive CBT for OCD, Saxena obtained positron emission tomography (PET) brain scans of patients suffering from OCD before and after a course of 4 weeks of daily intensive CBT supplemented by disorder-specific homework assignments, and compared the brain-scans of patients undergoing therapy, with the brain scans of untreated controls. Treated patients, as expected, showed significant decreases in thalamic metabolic activity, and, unexpectedly, they also showed a significant increase in right anterior cingulate cortex (ACC) activity, a response that the authors considered necessary to explain the patients' response to CBT, as well as the relative rapidity and robustness of such responses. In contrast to the effect of pharmacotherapy on brain function, augmentation of dorsal ACC activity maybe a primary mechanism of action of CBT for OCD.

It is known that subregions within the ACC have different functions. A critical such function, which may mediate the effect described here, involves the conscious regulation of emotion via re-evaluation of negative emotions, suppression of excessive arousal, and suppression of amygdala-mediated negative emotionality. These effects, specific for the CBT arm of the study, tended to occur rather rapidly and were correlated with treatment response at 4 weeks.

The biological relationship between negative affects and their modulation is central to disorders such as anxiety and depression, and it is a core paradigm that governs the psychotherapeutic treatment of these disorders. A central task of psychotherapy involves the repeated re-examination of painful memories, the experience and expression of the negative emotions associated with those memories, and the acquisition of a modulated soothing perspective. Such a sequence has also been studied in normal subjects, on whom the process of memory reappraisal has been found to be associated with improved mood, increased activity in prefrontal cortex subregions, and decreased activity in the amygdala and orbitofrontal cortex.

The results of studies examining the functional neuroanatomy underlying the neurobiological effects of verbal therapies are summarized in Table 33.10–4.

Brain Structures Key for Psychotherapy

The Amygdala. Although every region of the brain may participate in any mental process, selected regions may have special relevance for the health (or disorder) of key mental functions and for the treatment of such problems through psychotherapy. The amygdala is probably the brain structure most commonly understood by clinicians as essential to the experience of emotions, emotional dysregulation and, conversely,

emotional modulation through psychotherapy. It tends to show elevated activity in depression and posttraumatic stress; its dysregulation worsens in the presence of reduced prefrontal cortex activity (which supports key ego-functions) and improves through various means including verbal therapies. This dyad: amygdala, frontal cortex drives much of what is adaptive in humans and is capable of creating just as much emotional pain and dysfunction when it overwhelms human prudence and judgment.

Prefrontal Cortex. Within the frontal lobes, in addition to the already mentioned anterior cingulate cortex, the dorsomedial prefrontal cortex (DMPFC) is central to the exercise of top-down cognitive modulation of fear and other emotions, and to paying attention (self-awareness) to one's own (and other's) emotional responses facilitating mentalization and empathy. It executes many of the functions traditionally ascribed to the Ego and to the normative Superego. It also supports object relationships. It is hypoactive in depression. As long as it remains hypoactive, it limits the potential success of psychotherapy. Even without the availability of imaging technologies, the wise psychoanalytic clinicians of yore treating such patients, often insisted on affording them an initial period of therapy designed to "build up the ego," largely through being supportive, providing empathy, and supporting the patient's own efforts for tender loving self-care. Exercise however minimal in the early stages, rest, proper nutrition (protein, minerals etc.) help as well. If necessary, treatment with stimulating antidepressants rapidly refills the empty "battery." In the authors' own practice, they have been able to detect through neuroimaging (SPECT) significant under activity in the DMPFC in a sizable number of patients who became "stuck" in short-term therapy after an initial period of rapid progress, and who could resume their expected progress after the SPECT guided treatment of detected neurobiological dysfunctions.

The Dorsolateral Prefrontal Cortex. A close neighbor of the DMPFC reviewed above, the DLPFC provides, under normal conditions, the basic infrastructure of the person's executive functions.

Table 33.10–4.
Functional Neuroanatomy of Neurobiological Effects of Verbal Therapies

Disorder	Treatment Technique	Post Treatment Findings
Obsessive-compulsive disorder–social phobia (PET—Baxter, 1992; SPECT—Nakatani, 2003)	Cognitive-behavior therapy/behavior therapy	↓ Metabolism in right caudate nucleus Uncoupling of cortico- striato-thalamic activity
Pediatric OCD (fMRI—Huyser, 2010)	CBT CBT	↓ Limbic metabolism ↓ Thalamic metabolism Increased dorsolateral prefrontal and

		parietal (from baseline low)
Major depression	Cognitive-behavior therapy	Decreased ↓activation in dorsolateral prefrontal cortex Decreased ↓activation in para hippocampal gyrus ↑ Activation in limbic regions
Major depression (PET—Brody 01, SPECT—Martin 01) fMRI—Ritchey, 2011	Interpersonal psychotherapy CBT	↑ Activation in right basal ganglia ↑Activation in right posterior cingulate ↓Metabolism in prefrontal cortex, which correlates with symptom improvement ↑Ventromedial, ↑Amygdala, caudate, hippocampus
Disorder	Treatment technique	Posttreatment Findings
Panic Disorder (fMRI—Beutel, 2010)	Psychodynamic Thx	↓Amygdala and hippocampus ↑Prefrontal activity
Somatoform Disorder (fMRI—de Greck, 2011)	Psychodynamic Thx	Normalization of L posterior gyrus and R ventroposterior thalamus
Social Phobia (PET—Furmark 02) Social Phobia (fMRI—Golden and Gross, 2011)	CBT Mindfulness	↓Amygdala, in proportion to improvement ↓Amygdala

Attention, concentration, control of effort, relational processes, working memory, emotional memory, and many other functions under its control, are absolutely essential for successful adaptation to function and to people. It can also be hypoactive in depression and may contribute to the experience of loneliness, social isolation, loss of interest, and low standards of performance. In geriatric depression, the poor concentration, inability to focus, and no desire and no capacity to get much done raises the question of the differential diagnosis of dementia versus depressive pseudo dementia. Imaging studies showing hypoactive DLPFC without the widespread fronto-temporal or temporo-parietal underactivity usually present in true dementia, points toward a diagnosis of depression.

Neuroimaging and Basic Psychodynamic Principles

Imaging of Core Psychoanalytic Constructs. Of the many challenges that lie ahead as psychiatry and psychoanalysis continue their march toward becoming true science-based, biopsychosocial disciplines, two are really within the reach of currently emerging science. The first challenge consists of the need to delineate the brain systems which underlie normal psychological functions, and which may signal disorder when functioning outside of defined ranges. Validated research findings are emerging, for example, which enable one to know how the temporal lobes and thalamus act in concert to modulate mood and to restrain wild fluctuations of mood in response to every day moderately challenging stimulus. Key research

findings provide information as well, how the orbitofrontal cortex facilitates (or inhibits) impulsive activity, depending upon its functional state, and what conditions may lead to poor impulse control in a particular single event or situation, in response to a particular class of stimulus (sexual stimulation, risk taking in gambling) or as an expression of enduring personality characteristics. The second fascinating task is represented by the need to re-evaluate, and if necessary reformulate, core psychoanalytic constructs, which were formulated by Freud while he waited for a robust brain physiology that would help trace neural activity as it was converted into basic mental function. As he humbly said: "The deficiencies in our description would probably vanish if we were already in a position to replace the psychological terms with physiological or chemical ones. We may expect physiology and chemistry to give the most surprising information."

As summarized by Bradley S. Peterson and other investigators, imaging technologies can be used to develop experimental paradigms that may permit the understanding of the neural basis underpinning central theoretical constructs of psychoanalysis.

Drives and Instincts. The classical Freudian tenet that humans, as well as all living organisms, seek to maximize pleasure and minimize displeasure has been expanded by modern neuroscience into a complex theory of affects and rewards that describes numerous emotional systems served by different neural substrates. Such neural substrates, under the label pleasure, can discern and delineate and quantify various reward-systems and can help the organism to select which behaviors it enacts (or avoids) when a particular neural circuit is activated or inhibited. Thus, such circuits—and their attendant neurotransmitter(s)—may direct or even compel the organism to seek extremely risky behaviors, if driven by a severe enough "dopamine deficiency" state for example.

The nucleus accumbens and related limbic circuitry contribute to the "final reward equation" by assigning values of pleasure or displeasure to affective experiences, and the alerting/arousal system recruits the reticular formation, thalamus, amygdala, and cortex to the task of assigning personal meaning and salience to those affective experiences. Different emotional states are, in this paradigm, the result of complex and distinct interactions between the reward/pleasure systems (which promote appetitive drives) and the activation of arousal systems that assign priority to the pursuit of certain objects or activities over alternative behavioral choices. Imaging studies have shown increased activation of the nucleus accumbens and the orbitofrontal cortex while experimental subjects experience the euphorigenic effects of certain drugs, whereas the intense dysphoria experienced during abstinence correlated with metabolic underactivity of the same circuits.

As these investigations go forward, they are likely to yield a more refined theory of affects, drives, and emotions on which to create a modern psychoanalytic metapsychology and a theory of techniques built on hypotheses that can be tested in controlled studies. The role of the ego and superego as mediating agencies in the modulation of drives is as central to psychoanalysis as Freud's theory of drives.

In this area, modern neuroscience in general and neuroimaging in particular are generating seminal insights.

The relevance of frontostriatal systems is key to understanding how individuals plan, execute, and monitor goal-directed behaviors by selecting from the innumerable choices presented to them by the environment, as well as by their own perceptions, affective experiences, memories, and cognitions. Neuroimaging studies indicate that the anterior cingulate cortex and prefrontal cortex are involved in the top-down regulation of the capacity for self-regulation through their influence on broader brain regions such as temporal and parietal cortices and the basal ganglia. Neuroimaging studies of disorders of self-regulation, such as attention-deficit disorder and bipolar disorder, have established the existence of functional disturbances of cortico-striato-thalamo-cortical activity. These discoveries have direct clinical application as well; witness for example the very positive clinical impact of adding anticonvulsant mood regulators to the treatment mix of patients suffering from anxiety or depressive disorders or even attention deficit disorder, when neuroimaging determines the presence of temporal lobe over or underactivity. Under such conditions, pure antianxiety, antidepressant, or attention focusing treatments are not sufficiently effective because of the severe mood disruption precipitated by temporal disorganization.

Memory is also a core player in psychoanalytic discourse. Neuroscience research has described numerous memory systems with diverse distributions in the central nervous system. A key distinction for psychoanalysis is that between declarative and procedural memory systems. Declarative (explicit) memory is mediated by the medial temporal lobe and the hippocampus and stores conscious memory for people, places, and things. Procedural memories, on the other hand, are completely unconscious and are evident only in behavioral performance rather than conscious recall. Procedural memory involves the striatum (mediating the acquisition of skill and habits), the neocortex, the amygdala (facilitating conditioned emotional response), and the cerebellum. Many investigators have advanced the idea that a part of the unconscious (the procedural unconscious, which maps onto procedural memory) is central to key processes of psychoanalytic therapy. Early attachment relationships, for example, are internalized and encoded through procedural memory and unfold in the therapeutic process in the form of transference. Psychotherapy, in this context, represents the potential for the

development of a new primal attachment relationship that may create the conditions necessary to change attachment-related, implicit procedural memory as new interactions with the therapists are internalized by the patient in corrective self-object representations. Such maybe the deepest meaning and therapeutic value of the concept of "corrective emotional experience" advanced by Frank Alexander in the 1940s and of the "moment of meaning" developed by Sanders and Stern to describe moments in the interaction between patient and therapist that break the impasse of a transference-based pattern of relatedness and advance the therapeutic relationship to a new level.

Neuroimaging and the Prediction of Outcome

The most persuasive outcome predictions have developed out of neuroimaging studies of depression. In one such study, increased activity in the rostral anterior cingulate cortex predicted a positive response to treatment, a finding that has been replicated in other studies.

The anterior cingulate performs many functions related to conflict resolution and response to basic emotions. The central region of the cingulate (rostral subregion) may be recruited to resolve conflict between emotional stimuli or between conflictive cognitive mental contents. Thus, patients with higher pretreatment levels of activity in this area of the cingulate may be in a better position to resolve their pathogenic conflicts.

In studies related to the prediction of outcome of treatment in OCD, the orbitofrontal cortex is a key area. Neuroimaging studies have shown that lower pretreatment levels of activity predicted a good response to pharmacotherapy, whereas higher pretreatment levels or activity predicted a better response to psychotherapy.

References

Abbass A, Sheldon A, Gyra J, Kalpin A. Intensive short-term dynamic psychotherapy for DSM-IV personality disorders. *J Nerv Ment Dis*. 2008;196(3):211–216.

*Abbass AA, Kisely SR, Town JM, et al. Short-term psychodynamic psychotherapies for common mental disorders. *Cochrane Database Syst Rev*. 2014;(7):CD004687.

Alexander F, French TM. *Psychoanalytic Therapy: Principles and Application*. New York: Ronald Press; 1946.

Beutel ME, Stark R, Pan H, Silbersweig D, Dietrich S. Changes of brain activation pre- post-short-term psychodynamic inpatient psychotherapy: an fMRI study of panic disorder patients. *Psychiatry Res*. 2010;184(2):96–104.

Busch FN, Milrod BL, Sandberg LS. A study demonstrating efficacy of a psychoanalytic psychotherapy for panic disorder: implications for psychoanalytic research, theory, and practice. *J Am Psychoanal Assoc*. 2009;57(1):131–148.

*Davanloo H. *Intensive Short-Term Dynamic Psychotherapy*. Chichester, England: Wiley; 2000.

Davanloo H. Intensive short-term dynamic psychotherapy. In: Sadock BJ, Sadock VA, eds.

Comprehensive Textbook of Psychiatry. 8th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2004.

Driessen E, Cuijpers P, de Maat SC, Abbass AA, de Jonghe F, Dekker JJ. The efficacy of short-term psychodynamic psychotherapy for depression: a meta-analysis. *Clin Psychol Rev*. 2010;30(1):25–36.

*Etkin A, Pittenger C, Polan HJ, Kandel ER. Toward a neurobiology of psychotherapy: basic science and clinical applications. *J Neuropsychiatry Clin Neurosci*. 2005;17(2):145–158.

Forbes EE, Olino TM, Ryan ND, et al. Reward-related brain function as a predictor of treatment response in adolescents with major depressive disorder. *Cogn Affect Behav Neurosci*. 2010;10(1):107–118.

Fosha D, Paivio SC, Gleiser K, Ford J. Experiential and emotion-focused therapy. In: Courtois C, Ford JD, eds. *Complex Traumatic Stress Disorders: An Evidence-Based Clinician's Guide*. Chap 14. New York: Guilford Press; 2009:286–311.

Fosha D. Emotion and recognition at work: energy, vitality, pleasure, truth, desire and the emergent phenomenology of transformational experience. In: Fosha D, Siegel DJ, Solomon MR, eds. *The Healing Power of Emotion: Affective Neuroscience, Development, Clinical Practice*. New York: Norton; 2009.

Freud S. The resistances to psycho-analysis. In: Strachey J, ed. *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XIX (1923–1925): The Ego and the Id and Other works*. London: Hogarth; 1925.

Gerber AJ, Peterson BS. Measuring transference phenomena with fMRI [abstract]. *J Am Psychoanal Assoc*. 2006;54(4):1319–1325.

Hilsenroth MJ. A programmatic study of short-term psychodynamic psychotherapy: Assessment, process, outcome, and training. *Psychother Res*. 2007;7(1):31–45.

Hilsenroth MJ, Defife JA, Blake MM, Cromer TD. The effects of borderline pathology on short-term psychodynamic psychotherapy for depression. *Psychother Res*. 2007;17(2):172–184.

*Kohut H. *The Analysis of the Self*. New York: International University Press; 1971.

Leichsenring F, Rabung S, Leibing E. The efficacy of short-term psychodynamic psychotherapy in specific psychiatric disorders: a meta-analysis. *Arch Gen Psychiatry*. 2004;61(12):1208–1216.

Leichsenring F, Schauenburg H. Empirically supported methods of short-term psychodynamic therapy in depression—towards an evidence based unified protocol. *J Affect Disord*. 2014;169:128–143.

Malan D. *The Frontier of Brief Psychotherapy: An Example of the Convergence of Research and Clinical Practice*. New York: Plenum Medical Book Co; 1976.

Malan D. *Towards the Validation of Dynamic Psychotherapy*. New York: Plenum Press; 1976.

Milrod B, Leon AC, Busch F, Rudden M, Schwalberg M. A randomized controlled clinical trial of psychoanalytic psychotherapy for panic disorder. *Am J Psychiatry*. 2007;164(2):265–272.

*Peterson BS. Clinical neuroscience and imaging studies of core psychoanalytic constructs. *Clin Neurosci Res*. 2005;4:349–365.

Reneses B, Galian M, Serrano R, et al. A new time limited psychotherapy for BPD: preliminary results of a randomized and controlled trial. *Actas Esp Psiquiatria*. 2013;41(3):139–148.

Roffman JL, Marci CD, Glick DM, Dougherty DD, Rauch SL. Neuroimaging and the functional neuroanatomy of psychotherapy. *Psychol Med*. 2005;35(10):1385–1398.

Sifneos P. *Short-Term Anxiety Provoking Psychotherapy: A Treatment Manual*. New York: Basic Books; 1992.

Smith ML, Glass GV, Miller TI. *The Benefits of Psychotherapy*. Baltimore, MD: John Hopkins University Press; 1980.

Trujillo M. Treating pathological narcissism with short-term dynamic psychotherapy. In: Ogrodniczuk JS, ed. *Understanding and Treating Pathological Narcissism*. Washington, DC, US: American Psychological Association; 2013.

Trujillo M. Short-term dynamic psychotherapy of narcissistic personality disorders. In: Kaslow FW, ed. *Comprehensive Handbook of Psychotherapy*. Vol. I. New York: Wiley; 2002.

Winston A, Pollack J, McCullough L, Flegenheimer W, Kestenbaum R, Trujillo M. Brief psychotherapy of personality disorders. *J Nerv Ment Dis*. 1991;179(4):188–193.

▲ 33.11 Other Methods of Psychotherapy

ADAM M. BRENNER, M.D., AND E. WILL CLARK, M.D.

It is well established that a majority of the effectiveness of psychotherapeutic treatments can be accounted for by the common elements that all psychotherapies share. The features shared by all include a healer–patient relationship, in which the different roles carry different expectations and a differential balance of power; a nonjudgmental, supportive acceptance of the patient; and an alliance of working together on shared goals. These common elements are rooted in the universal need for attachment and connection. Human brains are designed to seek out attachment when they feel isolated or threatened, and to reciprocally be moved by the pain and loneliness of those who reach out to us. Resting on this foundation, psychotherapies of many different stripes are able to provide an experience—sometimes as quickly as the initial session—of feeling comforted, finding hope, and experiencing an amelioration of the sense of being alone or marginalized by one's presenting difficulties.

But if all therapies share these comforting elements, what accounts for the existence of so many different psychotherapies? The development of a specific psychotherapeutic method begins with the creation of a psychological model. Each individual model is built upon a unique conceptualization of how psychopathology evolves and what must happen to achieve an effective and satisfying life, either through the resolution of the psychopathology or in spite of the ongoing psychopathology. Thus, psychotherapy is always a learning process, taking advantage of the brain's neuroplasticity and the related human capacity for durable change at the most fundamental levels of being. Yet, as in formal education settings, learning within psychotherapy is a unique process among individuals and requires a diverse range of approaches. One way to understand the wide array of psychotherapeutic methods, therefore, is as an attempt to

accommodate the various learning styles.

In addition, what the therapist hopes the patient will learn in the course of treatment is remarkably diverse, and often rooted in very different fundamental ideas about the nature of psychopathology and about what defines an effective and satisfying life. As an example, the cognitive therapist believes patients need to learn about implicit maladaptive thoughts, which eventually, results in changes to deeply held belief systems that contribute to recurrent emotional distress and dysfunctional interpersonal engagement. The psychoanalytic therapist, on the other hand, believes the most effective learning approach involves exploration of unconscious wishes and conflicts, which allows for the release of long-standing, self-imposed obstacles that hold people back from living life in the fullest way possible. In both instances, the aim is to relieve suffering and promote the growth and development that hopefully, will prevent suffering in the future. How this aim is achieved and the specific goals that occur within the treatment itself are quite different.

Through the years, new psychotherapeutic approaches have been developed, usually borne out of a desire to fill an identified need in the treatment of mental illness. This section reviews several psychotherapies, some with a particular political slant or dealing with a special population and some older therapies that continue to have discrete followings. Each of them, if effective, has been able to generate and benefit from the comforting elements common to all therapies, but each of them has had very different ideas about what kind of learning will be of benefit to the patient and what the therapist should do to help "teach" the patient.

PARADOXICAL AND STRATEGIC THERAPIES

History

According to the principle of paradoxical intention, fighting the symptoms of performance anxiety increases their severity, but invoking them deliberately may yield the opposite. Milton Erickson elaborated those ideas into paradoxical therapy, and the method was further refined and promulgated widely by Jay Haley.

Erickson was a hypnotherapist whose prowess in the use of paradoxical techniques was reputed to be extraordinary. Haley, while working on the double-bind theory of schizophrenia, thought that perhaps there could be therapeutic double binds, as well as pathological ones. He attended one of Erickson's hypnosis workshops in 1953 and, with John Weakland, began a regular series of visits that endured for 20 years. Haley and his colleagues learned from Erickson and published what they had learned. Others made similar visits or were attracted by the ideas or the stories that they heard, and they added to the reputation and the literature of the growing movement. Haley later developed strategic therapy, essentially a method

that applied Ericksonian paradoxical principles, along with other directive techniques, to family therapy.

Interest in Erickson's paradoxical therapy grew over the years and reached its height in the late 1970s and early 1980s. By this time, the method claimed success in treating individuals, couples, and families for a range of problems, including phobias, compulsions, insomnia, schizophrenia, Tourette's syndrome, urinary retention, transvestic fetishism, anorexia nervosa, agoraphobia, childhood fecal incontinence, alcohol use disorders, depressive disorders, fire setting, asthma, constipation, truancy, and a variety of sexual performance and relationship difficulties.

Erickson died in 1980, and a foundation was formed to continue the work that he started. He was one of the first to show that brief therapy can be an effective and primary approach. He advocated that therapists should not be just passive listeners, but active interventionists and guides; focus on symptoms rather than on dynamics and on action rather than on understanding; prescribe, as part of the treatment, the very resistances and symptoms the patient brought to it; and use indirection and metaphor to achieve the goals of therapy.

Theoretical Issues

The overarching idea in paradoxical therapy is to embrace and amplify the notion of paradoxical intention and thereby to reduce performance anxiety or destabilize a fixed internal or social system. The paradox and the destabilization may be effective in themselves or may cause secondary reverberations, such as increased self-awareness and the evocation of new behavior or emotions that lead to helpful change.

Paradoxical therapy is relatively atheoretical about cause and is not much concerned with it. Behavioral prescription is emphasized, with the idea that altering internal and social feedback loops can alter the system. Symptoms that block change can be made vehicles for it if they are embraced and are prescribed or altered in their timing or the function that they serve. Instructing an impotent patient under no circumstances to try intercourse, for example, may dissipate anxiety and may allow him to do so; telling an insomniac to get up and polish the floors for the rest of the night may similarly invite sleep. Paradoxical therapy also recommends defining the problem clearly, establishing clear goals for the treatment, offering a plan, and enlisting the collaboration of the patient or the family in the suggested strategy.

Techniques

Some principles of paradoxical therapy can be directly translated into techniques: Using patients' tendencies toward compliance or defiance—for

example, by overtly expecting a relapse and being surprised about improvement; reframing (i.e., relabeling behavior and emotions so that their experience is changed from negative to positive or vice versa); using metaphor, indirect directives, and cryptic paradox; and distancing patients from their symptoms or challenging patients by humor or sarcasm. These techniques and the extent of the therapist activity encouraged in paradoxical therapy are illustrated in the following cases of Erickson and his followers.

Making the Activity Deliberate. An alcoholic wife secretly hides her bottles, and the husband searches for them with increasing anger. With the relationship at a breaking point, Erickson prescribed that she deliberately hide the bottles and that he search for them. If he found one, he could keep it. If it eluded him, she would be allowed to drink in peace. He also suggested that they buy a trailer and go fishing, an activity that they both disliked. The reported results were that the drinking abated, and the couple later happily used the trailer to go camping. The deliberate game apparently subverted the purpose of the prior secretive activity, and forcing the couple to share a disliked recreation brought them into close contact (with alcohol unavailable) and pressured them to communicate, to enjoy each other, and to make pleasant vacation plans.

Prescribing the Symptom. Instead of telling her husband what she felt, a woman would unconsciously start a fight with him when she wanted affection. She was instructed to start a fight on purpose and intensely the next time that she felt empty or unloved. Doing so, she noted the connection and realized that, under attack, her partner could not possibly respond with support or love. She also felt ashamed, and the experience helped her acknowledge and express her wishes more directly.

Aversive Conditioning. A woman consulted Erickson about her schizophrenic son, who did nothing but sit around the house and moan. Erickson recommended taking him for a drive; she was to drop him off, drive 3 miles farther, then sit and read while she waited for him. After doing that a few times, the son volunteered to walk without being dropped off, and later asked if he could go bowling while she read, because he liked bowling more than walking.

A second patient told Erickson that he could drive only on certain streets and that he would become nauseated, vomit, and faint if he tried to drive outside the city limits. Erickson required him to drive to the edge of town in his best clothes, to park and to lie down in a sandy ditch at the roadside, and to lie there for at least 15 minutes. He was then to drive one or two car lengths farther, to repeat the procedure, and do that until he could drive from one telephone pole to the next. He was also told to stop

immediately and to get into the ditch at the first sign of any symptoms. The patient reported: "The more I did it the madder I got. So I just quit and began to enjoy driving." Thirteen years later, Erickson reported, the former patient still had no trouble driving. Anger and defiance allegedly brought about a cure.

Destabilization. A boy who was 10 years of age engaged in compulsive, public masturbation; the compulsion had endured for years, despite many treatment efforts. Taking the behavior seriously, the therapist asked the boy to keep a list of when he masturbated and to note when he found it most pleasurable. With the idea that he should do it when he found it most pleasurable (on weekends), he was assigned to do it eight times on Sunday but not on Monday. When the boy failed and masturbated on Monday, the disappointed therapist prescribed that he do it 12 times the next Sunday. When the boy was not up to the task, the therapist suggested that, as punishment, the boy masturbate once daily in public and eight times on Sunday. Masturbation was also prescribed as a punishment for thumb sucking and as an alternative to homework. Gradually, the boy stopped and even enjoyed his homework.

Examples reported as successful include telling a woman with headaches, which were symptomatic of family relationship problems, that the headaches were probably incurable and that therapy should aim at helping her live with them; asking the parents of a child with severe and long-standing encopresis whether they could tolerate being normal and being better parents than their mothers (the father then threatened the child with laxatives, and the problem disappeared); and combining reframing and negative emphasis or challenge by telling a family with a school-phobic child: "Maybe it's better she stay at home. It keeps the family from being normal, and being normal may not be tolerable." (The therapist reported that the family did not get insulted and leave but changed spontaneously to prove him wrong.)

Clinical and Ethical Issues

Paradoxical therapy is clearly not a psychotherapy itself but a set of techniques potentially usable in therapy. Making explicit the assumptions that underlie the techniques reveals their range of usefulness and their pitfalls. In addition to the idea that prescribing or restraining the symptoms can reduce performance anxiety, the assumptions are that symptoms always reflect a problem in a social context or relationships, that they are somehow and ultimately volitional, and that they therefore respond to being emphasized or manipulated.

Such assumptions are probably acceptable in some cases, in which emphasizing the symptom amounts to flooding, an accepted and well-demonstrated technique of behavioral therapy, and in relationship

problems, in which prescribing the symptom or altering the context of its appearance may help the participants recognize their inappropriate emotions or behavior patterns.

Using paradoxical therapy in biologically based illnesses is far more of a problem. Expecting a patient with Tourette's syndrome to respond to such techniques alone suggests a volitional element greater than what is present. The basic symptoms of schizophrenia are unlikely to respond to paradoxical therapy, and, although behaviorally conditioned symptoms may respond, the case histories show the difficulty of distinguishing what is conditioned from what is biologically based. The literature about paradoxical therapy is primarily anecdotal, and the rare research study is less enthusiastic than authors who claim widespread success.

Most importantly, many of the case histories in the literature include manipulations of the patient or the families that may cover or even invite aggressive acting out by the therapist, and some case histories suggest a simplistic conception of illness in which punishment is curative. When the operating field permits behavior that is less than open and straightforward, therapists need even more vigilance and self-awareness than when procedural safeguards protect the patient.

CLIENT-CENTERED THERAPY

History

Client-centered therapy was created in 1940 by Carl Rogers, a psychologist. Believing that diagnosis, planning, and interpretation manipulate the patient to conform to the therapist's ideas, Rogers tried only to listen, understand, and reflect what he perceived the patient to feel. When Rogers did so, he found that the patient often responded with a deep feeling of personal involvement and the production of new and important material. That passive or nondirective therapy was later named *client-centered therapy* and was published as such in 1951. Rogers' ideas found wide acceptance, and a movement around the principles of the therapy gathered momentum and grew over the next three decades. Rogers himself continued to develop the method and its applications at the University of Wisconsin, the University of Chicago, and Ohio State University and in La Jolla, California, where he died in 1987.

Client-centered therapy is based on the ideas that each person has the ability to solve psychological problems and that feeling understood and highly valued aids the forces within to do so. Ideas such as an instinct toward self-realization and limitless inner powers had enormous appeal to the humanistic idealism that was developing at approximately the time that the method was published. A focus on the value of the individual and Rogers' later precept that long and particular training was not necessary for becoming a good therapist were similar rallying points for the

antiestablishment, antiauthority emotions of the 1960s and 1970s.

Rogers played a major role in wresting the practice of psychotherapy from psychiatry alone and distributing it into the hands of psychology and other disciplines. He is also credited with being the first to research his results, the first to record sessions word by word and in movies and on videotape, the first to publish transcripts of his failures to encourage the study of what goes wrong, and the first to put forward the thesis that one does not need academic degrees to become a therapist. All but the last idea have become well-accepted elements in the field.

Client-centered therapy has been reported to be useful for children and adults, problems of neuroses, situational problems, speech difficulties, psychosomatic problems (e.g., allergies), and, to some extent, psychosis. It has been practiced with individuals and groups, and its principles have been applied in industry, education, child rearing, and intensive encounter groups. Just before his death in 1987, Rogers was training policy makers and leading communication groups in South Africa and in workshops in Hungary and the former Soviet Union.

Theoretical Issues

The overarching presupposition of client-centered psychotherapy is that a basic, forward-moving force exists within—a tendency to grow, actualize, and integrate. A number of other principles flow from that idea: All people are believed to have the capacity to understand what factors cause them unhappiness and to overcome those factors; to do so, they must bring what is going on inside into awareness; and, to effectuate the circumstances that make awareness possible, the therapist must establish a relationship that is warm and accepting.

Rogers believed that human beings immediately attach meaning to whatever is perceived, so that perception becomes reality. A group of perceptions is organized as the self-concept—people's perceptions of their own characteristics, their relationships with others, and the value they place on those perceptions. Experiences contradicting the preferred perceptions of that organization cause anxiety and must be shut out. In therapy, the structure of the self is relaxed, so that the patient can drop the barrier and admit, "Yes, I do have those parts of myself that are unacceptable." When this happens, the previously denied parts are accepted and then integrated, and the result is a better-functioning person, one who is less tied up in denial and conflict, more effective, and more tolerant of others than in the past.

The aim of the therapy is to create an interpersonal environment in which the self-structure can relax and in which material can come into consciousness. The therapist sets the three conditions that foster relaxation and the personal growth that follows. The three conditions are *congruence* (or *genuineness*), the ability to be present as a person, to be real in the

relationship with the client; *unconditional positive regard,* which conveys to the client, "I accept you as you are, without making judgments or evaluations"; and real *empathic understanding,* a sense of standing in the client's shoes, seeing the world from that vantage point, and then communicating what is understood from that view to the client. When these conditions are met, the client feels released and understood and is able to accept and integrate, view the world differently, and continue the process by bringing out new material.

Client-centered therapy was first practiced in a course of 5 or 6 hours, then 15 or 20 hours, and, still later, 100 or more hours. Whether the shift was caused by clients' seeking greater levels of change, the enlistment in treatment of more difficult cases, an evolution in technique, or a fall-off in results after the initial enthusiasm is unclear.

Techniques

The techniques of client-centered psychotherapy serve the goal of having the therapist convey a sense of the worth and significance of the client and the idea that the therapist is trying to understand. Far from being laissez faire, the therapist needs to help actively and empathetically clarify what the patient feels. To be entirely nondirective, the therapist restates and reflects a close approximation of the client's feelings. When this is done correctly, clients feel as if the therapist were almost a part of them, working on their problems as they want to work on them; they feel accepted warmly, without any judgmental overtones. The acceptance and empathy are themselves therapeutic.

Rogers cited an example of an explosive 13-year-old boy who was changed and calmed by sessions in which he and the therapist only played tic-tac-toe or in which the therapist allowed silence while the patient sat at the window with his back to him, quietly counting cars. The therapist conveyed respect and belief in the child and the idea that the hour belonged to the child and that the adult would not be shocked by anything he did. The only limitation was that anger was allowed no physical expression, only symbolic and verbal expression.

The therapy aims not to interpret the client's unconscious motivations or conflicts but to reflect what the client feels. As the counselor and the client progressively clarify their feelings within the warm and accepting relationship, denied feelings come to light and can be dealt with and integrated.

Rogers described a male patient with considerable anger. As the therapist accepted the client's feelings and clarified at whom they were aimed, the client came to see which parts were rational and which were not. He was able to say, "If you are like my father, or if I only see you that way, I hate you as I hated him." When the unrealistic aspects were expressed openly, they dissolved, leaving the real relationship in place.

The theory also claims that resistance is overcome rather than interpreted. That is accomplished by the consistent acceptance and valuing of the client, the thoroughly undemanding nature of the relationship, and the empathic effort to clarify the client's feelings.

A woman with denied anger at her husband embarrassed him with dizziness and faintness at gatherings that were important to him. The therapist consistently reflected her feelings to her with such questions as "You seem to feel . . . ," "You're confused by . . . ," and "It seems to you that . . . " That clarification helped her to move from affirmations of a good relationship with her husband to recollections of old resentments and the recognition of present ill will. Keeping the focus on her also aborted any tendencies for her to project or to accuse the therapist of anything but interest.

Transference is allegedly avoided in a similar fashion. Rogers gave examples of an angry client:

Client: Am I supposed to tell you everything?
Therapist: It's up to you; you can talk about anything you want to.
Client: I had a big fight with you in my mind as I came here. I thought the whole thing was a fraud—I'm being very frank—you can talk like we do down here almost anywhere if you get someone to listen.
Therapist: It seemed it was kind of a cheat, something you could get anywhere.
Client: It's like advertisements you see where, for a few dollars, people will sit and listen to you.
Therapist: Just listen to your troubles.
Client: Of course, they don't have your extensive background, but they do the same thing. Then, after I talked it out with myself, I realized that the reason I was mad at you is because I feel nervous and jittery coming here (pause), because I'm facing something I don't like to, talking things out myself.
Therapist: So the feelings about someone listening to you being a fraud you recognize as being connected with your fear and irritation at having to face things in yourself?
Client: That's right. That's the real logic.
In another example with chaotic angry feelings, the following dialogue took place:
Client: I hate you! I wish you never were born!
Therapist: You just hate me bitterly.
Client: I think I'll throw you in the lake. I'll cut you up! You think people like you, but they don't. I wish you were dead!
Therapist: You detest me, and you'd like to get rid of me.
Client: You think my father did bad things with me, but he didn't! You think he wasn't good, but he was! You think I want intercourse, but I don't.
Therapist: You feel I absolutely misrepresent all your thoughts.

Through this type of work, the client later confessed a hallucination and said: "I had to get rid of this somewhere. I knew you'd understand. It's true I hate myself, but I couldn't say it, so I thought of all the ugly things I could say to you instead."

Rogers wrote that those interviews showed how feelings of transference were dealt with in the same way as other feelings and how a consistent reflection of the client's feelings essentially dissolved a transference issue by helping the client recognize her irrational emotions, after which she

could deal with them and what they hid.

Indications

The general indications for therapy are when clients feel that their organized self-structures are no longer effective in meeting their needs in the reality situation or when they perceive discrepancies in themselves or see that their behavior is out of control. The description generally covers problems in the neurotic range, but a client-centered approach has been used with children who are 2 years of age and adults who are 65 years of age, with people who have mild adjustment problems and disorders diagnosed as psychotic, and with persons of all social classes and levels of intelligence. The theory affirms that a trial of the therapy is unlikely to do harm.

Early in the development of the method, Rogers believed that diagnosis was inadequate and prejudicial. Placing the locus of evaluation in the hands of the therapist rather than the client had undesirable consequences, undermining the client's power of self-determination and raising the specter of misuse for social control. Nosological limitations may have supported this early position, and the social circumstances of the 1960s may have lent it credence, but Rogers' followers believed in it as late as 1987.

Ethical Issues, Research, and Evaluation

Rogers' ideas sparked and rode the crest of the humanistic movement in psychology. His emphasis on the value of clients and on the need to accept and respect them in therapeutic relationships was timely and had a refreshing influence on child rearing, education, and group relations, as well as on psychotherapy. His attempts to train nonprofessionals were less well advised, and the attempts foundered; the method became so simplified that blank verbal repetitions too easily replaced a true empathic congruence and effort to understand.

Practitioners of client-centered psychotherapy did not believe that diagnosis was particularly useful and did not follow the medical model of illness. That a warm and genuine relationship is sufficient to dissolve resistances that are dynamically motivated and tenaciously held is another tenet open to question. Reflecting what the client feels without interpreting where it comes from or the purposes that it may serve is uninformative and could drive the client to action as a further defense. Rogers admitted as much when he noted that the method is frequently unsuccessful with clients who feel incapable of managing themselves and who insist that the therapist take over; a different approach interprets the client's insistence and resentment, rather than just reflecting them. Similarly, reflection in the face of other intense transferences seems to be less effective than a focus on

the fantasies that lend force to the transference.

In its contributions to psychotherapy research, client-centered therapy has been outstanding. Rogers and his followers were the first to record and review transcripts of sessions, the first to document the process of treatment, and the first to do follow-up studies of results. They developed such innovations as the q-sort and scales to objectify the Rorschach test. They showed that, in successful therapy, negative attitudes about the self are replaced by positive attitudes, that acceptance of others occurs in parallel, that psychological tension decreases, and that plans take the place of old complaints. Although initial claims about success in treating schizophrenia patients had to be dropped, and although Rogers and his colleagues did not do controlled comparisons of their method with others, they were able to document that client-centered therapy is effective. By clearly delineating that a successful outcome is associated with genuine acceptance and an empathic congruence that facilitates the recognition and expression of emotion, they recognized two of the features that empower psychotherapy of whatever stripe. Perhaps most importantly, their emphasis on progressive clarification of the client's feelings sounds an insistent cautionary note about the need for clear understanding, rather than facile interpretation based on the assumptions of one's theory.

TRANSACTIONAL PSYCHOTHERAPY

History

Transactional analysis or psychotherapy was the creation of Eric Berne, a psychiatrist with psychoanalytic training who parted with the Freudian movement to focus on social interaction as the unit of analysis. *Transactional Analysis in Psychotherapy* was published in 1961, and *Games People Play* was published in 1964; *Games People Play* became a record bestseller, with sales of more than 2.5 million copies. Berne's thesis was that, in social interactions, people seek intimacy, comfort, or recognition, which they get through a series of mutual exchanges or *strokes*. If the ego states of the participants are adult and appropriate, the transaction is harmonious and realistic. Ulterior motives, sponsored by other than an adult ego state, cause discomfort in the transaction, and transactions habitually colored by such motives may become ritualized and predictable.

Written with wit and humor, Berne's books inspired considerable controversy. Some accused them of facile psychologizing better left to cocktail parties, rather than true contributions to psychotherapeutic theory and practice. The fact that the movement has endured, inspiring a journal, an accrediting body, and an occasional article in the psychiatric literature, suggests that a number of its ideas have merit.

Theoretical Issues

Transactional analysis rests first on the principle that human beings have a hunger for stimulation and contact in the infantile style. When such strivings are blocked, efforts are made to regain the intimate contact. Compromises and sublimations are made in that pursuit, so that a hint of the original intimacy may suffice, and the stimulus hunger may be satisfied symbolically. Thus, the original physical strokes of caress and comfort are replaceable by metaphorical stroking by which the person is recognized, flattered, or supported and even by acknowledgments that may appear to bring discomfort. A stroke is a fundamental unit of interaction. An exchange of strokes is a *transaction,* the unit of social intercourse; a stimulus from one person produces a corresponding response in another person. Because social stroking satisfies the original infantile need, it follows that any social intercourse is better than none.

In addition, social interactions fill time, provide structure, and help avoid boredom. Interactional programs, all of which contain stroking opportunities, may be *material* (dealing with reality issues); *social* (the norms and manners of a social order); or *individual* (incidents sneaked into the other programs to get extra, personalized strokes). The sequences of a person's program are developed in line with that person's *life script,* a set of overall goals and patterns, and these sequences are the individual *games*. The games range from harmless and pleasantly satisfying to dangerous and destructive. What develops in a game is assumed to be the purpose of the game, and what eventuates as its end result are the strokes that constitute the aim and the satisfaction.

To understand and to diagram the interactions in a game, the theory posits three ego states, the parent, the child, and the adult. In the parental ego state, the person reacts as the parent or as the parent would have directed. In the child state, one may be adaptively compliant or whiny and withdrawing or, as the natural child, rebellious, creative, or impulsive. The adult state appraises reality and acts in tune with it; the adult also moderates between the parent and the child. All three states are present and potentially active at all times.

When the adult in one person addresses the adult in another and is responded to in adult terms, the stimulus and the response are complementary, and the transaction is smooth. The same is true if the child addresses the parent, and the parent responds appropriately—for example, when a child (to parent) demands "Water!" and the parent (to child) responds "Here, baby." Such transactions are normal, two-stroke harmonious social interactions.

However, if one partner in a couple asks, "Where are my cuff links?" and the other responds, "You're always losing things!" the lines of communication are crossed; the query may be adult to adult, but the

response is parent to child, and the situation gets sticky. A comparable child-to-parent cross in response to an adult-to-adult stimulus may occur when one notes to a person driving, "You missed the turnoff," and the driver responds, "You're always criticizing me!"

Such ulterior transactions are the basis of games, which are ongoing series of complementary ulterior transactions progressing to a predictable defined outcome that satisfies some aspect of the person's life-script strivings. On the surface, the interactions may appear plausible, but their recurrent nature, concealed snares, and predictable payoffs define them as parts of a neurotic game.

The test of a game is that a respondent's refusal to play causes increasingly intense efforts to force its continuation; continued refusal may elicit despair or bewilderment. The cure for a game is recognition of its presence, its rules of play, and its motivation. A cure is signified when the initiator can note with a laugh, "There I go again," and can desist volitionally.

Techniques

Because games generally involve at least two parties, they reveal themselves and are best treated in couples or groups. Transactional analysis has therefore had its most widespread application in couples and group therapies, although it has also been seen as useful in individual treatment. The techniques aim to illustrate the game, dissect the transactions, and frustrate players in their pursuit of strokes. In doing so, the players may learn a considerable amount about their life scripts and how they evolved, although that is not necessarily the goal of treatment.

Games are named for convenience of recognition, with the name usually connoting something about the ends sought. Thus, in "if it weren't for you," a woman (usually) picks a domineering man to marry, who then forbids her to do things she allegedly wants. She blames him for her unhappiness, they quarrel, their sex life deteriorates, he gives her gifts to make up, and the cycle goes on. In reality, the woman fears the activities for which she says she longs, has a sexual inhibition (which the husband may share), and is pleased by the opportunity to prove that men are tyrannical. In the game of "schlemiel," a man (usually) messes up or breaks things to be castigated and then forgiven.

Other life games include "debtor," "alcoholic," and "see what you made me do." The aim of each is to obtain one or several ulterior satisfactions without accepting conscious responsibility. In "alcoholic," the players may include the alcoholic, a persecutor (usually, a spouse), a rescuer (often, a health professional), a patsy (who gives money or sympathy), and, sometimes, a connection (who may tempt the alcoholic to drink). Each player may also be in the game for idiosyncratic reasons, games, or strokes. For the alcoholic person, the payoff is the hangover and its chance to

indulge in psychological torment and suffering; for the rescuer, it is an opportunity to play "I'm only trying to help you"; for the persecutor, it is an opportunity to play "Look what you've done to me." The antithesis, the goal of treatment, is to have each one recognize the role played and frustrate the primary player (and the other persons addressed as well) by giving up the role assigned.

Groups may be randomly selected or homogeneous. In treatment sessions, the theory of transactional analysis is taught assiduously; with blackboard illustrations, the aims of treatment are outlined in operational terms; and contracts are made about individual large goals and are redrawn for small, particular steps in the course of therapy. The therapist's role is active and confrontational but egalitarian.

Research and Ethical Issues

Interest in transactional analysis peaked long before the advent of criterion-based diagnosis and controlled clinical trials, and little research on it can stand up to today's critical scrutiny. It was claimed to be effective for persons with neuroses, psychoses, personality disorders, sexual disorders, and mental retardation; for individuals and groups; and as a means of self-help for young children and teenagers. Like most therapies, however, its usefulness rests on clinical lore and a certain common-sense validity. What seems sensible includes an emphasis that is typical of later cognitive-based treatments—namely, identifying elements of behavior that reflect neurotic problems, so that their maladaptation can be acknowledged, their driving emotions and underlying assumptions can be recognized, and the pattern of behavior can be altered. Psychoanalytic therapies share the same aims but include the additional goal of genetic uncovering. Transactional analysis' emphases on confrontation, contracts, homework, active diagramming of conflict, and teaching about behavior are also more like cognitive therapies.

The weakness in the system and the potential it has for abuse in work with patients lie in its theoretical assumptions. *Neurotic behavior,* by definition, is not realistically aimed and is often maladaptive. To assume that all bad consequences are deliberately sought is incorrect and unfair. To the extent that such a view is held, important motivations may be missed, and therapy may be tainted by a tendency to level accusations. If this position is avoided, and complaints are not trivialized by being seen routinely as goals, the dissection of social transactions and the focus on personal responsibility offered by transactional therapy conform to what is generally considered to be of therapeutic value.

EXISTENTIAL PSYCHOTHERAPIES

History

Existential psychotherapies are a diverse group of treatments that were inspired by the writings of the existential philosophers and novelists. The therapies emphasize exploring and validating the individual's subjective experience with a minimum of theoretical expectations; fostering greater authenticity in affective experience and decision making; and confronting the unavoidable and painful limitations of human existence. Defining the essential underlying qualities of these treatments so as to create a unifying framework for practice and research has been difficult. There is no single founder of existential therapy to invoke as a defining authority; the philosophies in which existential therapies are rooted differ greatly and are open to multiple interpretations; and the existential stance tends to be suspicious of authority and of overarching systems of meaning, making it inherently inhospitable to attempts to codify the disparate approaches. Nonetheless, there are several significant elements these theoreticians and clinicians share, and which together yield a reasonably coherent body of theory and practice.

Existential philosophy arose in the late 19th and first half of the 20th centuries as a challenge to traditional religious and philosophic optimism, which held that truth could be attained through reason, and that individual values could be grounded in the ever-expanding knowledge of the universe. Existentialism asserted that reason has limits, that the universe was a place of arbitrary facts, and that the irrational elements of the human psyche always limit the potential for self-knowledge and social order. Existentialists were suspicious of abstract truths, and placed their trust in the reality of an individual's experience. Within this perspective existentialists showed great diversity and included religious thinkers such as Martin Buber, Soren Kierkegaard, and Paul Tillich as well as secular and atheistic philosophers such as Martin Heidegger, Friedrich Nietzsche, and Jean Paul Sartre. Some, like Sartre and Albert Camus, expressed their ideas in fictional writing that explored the individual's confrontation with the inherent meaninglessness of life and the responsibility to nonetheless make authentic choices.

Kierkegaard and Nietzsche were aware of the profound anxiety and confusion resulting from feeling uprooted from a grounding morality and faith, while at the same time being cognizant of mankind's potential for acts of destruction. Kierkegaard advocated a religious "leap of faith," while Nietzsche pronounced that "God is dead," which if true would leave mankind forced to create new systems of value. As different as they were, both thinkers shared a conviction that a meaningful life requires a profound effort and commitment on the part of the individual.

The clinicians in Europe who created existential psychotherapy felt that Freud did not adequately address the human need to find meaning in an innately meaningless world. They felt that Freud's technique was profoundly helpful, but that his theory of biological drives was reductive

and overly focused on the physical world. The school of Existential Analysis, represented prominently by Ludwig Binswanger and Medard Boss, grew out of the application of Heidegger's philosophy to problems of psychotherapy. Heidegger created the concept of "being in the world" to define what is unique to human existence, specifically, that man is always both a thing in the world and the creator of meaning and knowledge.

Logotherapy is an existential psychotherapy based on the work of psychologist Viktor Frankl, whose ideas were influenced by his experience as an inmate in Nazi concentration camps. He came to believe that while people cannot control what happens to them, they can choose the attitudes with which they face and respond to their circumstances. Frankl felt this to be a fundamentally hopeful stance, and termed it "tragic optimism."

Somewhat later Rollo May focused the attention of the psychiatric community on existential psychotherapy and generated substantial interest among American psychotherapists; his particular approach blended humanistic psychology, psychodynamic techniques, and an emphasis on the individual's alienation from his or her own passions and inner life. Irwin Yalom's case studies and his text *Existential Psychotherapy* were similarly influential in sustaining interest in this existential-humanistic tradition among a broad audience of psychotherapy practitioners.

Theoretical Issues

Existential psychotherapies draw on diverse theoretical roots, which include the elaboration of the phenomenological method, as well as explorations of alienation and anxiety. Phenomenologists attempt to step outside pre-existing conceptions and ideas about the world in order to experience each new object or person directly. The phenomenological method, therefore, inspired existential psychotherapists to seek a spontaneous empathic encounter with each new patient. Clinicians within this tradition emphasize an engagement with the patient in his or her unique wholeness. Therapists strive to avoid letting their preconceived ideas about patients and their pathology interfere with their understanding of the patient in the present moment. In their view, a focus on diagnosis, or on ego structure and defenses, is thought to divide the experience of the patient into pieces and miss the opportunity for an encounter with the whole person.

Existential therapists see anxiety and guilt as essential aspects of human existence, and not just responses to specific events or conditions. Anxiety and guilt can be realistic, neurotic, or existential. Realistic anxiety is the signal of actual danger, while neurotic anxiety follows from fantasized, projected, or displaced threats. Existential anxiety is seen as the experience of the dread of nonbeing—which includes death, but also a person's unrealized potential. Realistic guilt is in response to actual crimes or transgression, and neurotic guilt follows from fantasized, generally

unconscious, forbidden wishes. Existential guilt means guilt derived from suppressing one's real potential in favor of inauthentic conformity.

Techniques

Although the practice of therapists varies widely, there remain some common points of technical emphasis. First, existential therapists try to find the patient's unique world view and enter into this experience with the patient. They emphasize the importance of empathy with the totality of the patient—that is, actually "knowing" as opposed to "knowing about" the person. Second, existential therapists attempt to uncover and work through deep anxiety that stems from such universal threats such as aloneness, meaninglessness, and mortality. Third, existential therapists emphasize that all people are subject to inescapable reality-based limits and try to help their patient bear these facts. One such limit is the fact of being who one is and not someone else; another is the impossibility of fully knowing, or being known by, another person. Fourth, existential therapists consider the denial of one's mortality and the fear of death to be critical determinants of pathology and focus their attention on these issues. Finally, the therapists support patients recognizing their freedom to choose their actions within these limits, and their freedom to create their own meanings for the events of their lives. The existential therapist pays close attention to whether patients' choices represent their true desires and values of the patient, or whether they are better understood as distractions or narcotics.

Yalom described a patient who presented with compulsive sexual affairs. A dream of a magic bow and arrow was interpreted as representing the use of his phallus to transcend the ordinariness and banality of death and his affairs as providing him with the illusion of immortality. As the patient began to give up the compulsive behavior, it was possible to use therapy to also confront the question "What are people for?" and open up an exploration of existential isolation.

Existential analysis borrowed the technical structure of psychoanalysis (frequent sessions, use of the couch, free association, nondirective exploration), but remains squarely within the frame of the patient's current experience. Therapists eschew interpretations of unconscious conflict, symbolism, or transference and view the essential pathology as that of being closed off to parts of one's own experience. Thus, treatment aims to increase patients' openness to what is disavowed. To this end, some therapists use direct confrontation, paradoxical interventions, and active reframing. The techniques of logotherapy aim to promote the identification of meanings that transcend one's own self; this is done through "self-distancing," which seems very similar to developing and observing ego. Such specific maneuvers are meant to be used in the context of a "Socratic dialogue" that attempts to trigger the patient's own resources of wisdom.

Existential therapy, unlike psychoanalytic approaches, is not primarily

developmental in its thinking. The basic concerns of human existence—death, freedom, isolation, and meaning—underlie anxiety for patients of any age. And May, reversing the classical psychoanalytic tendency to find present meanings in autobiographical events, suggested that people's understanding of the past is actually determined by their experience of the present and future. As with many contributions of the existential clinicians, these ideas have since been integrated into the body of mainstream psychodynamic thought.

Indications

Existential psychotherapists claim success with this modality in the same range of conditions as dynamic therapies in general—depression and dysthymia, anxiety disorders, complicated bereavement, trauma, adjustment disorders, and the like. An extended application of existential therapy suggests that much of the depressed affect seen in bipolar patients may be better conceptualized as existential despair; thus, psychiatrists should address core existential issues such as the conflict of hope versus the demoralization. It has been reported also that logotherapy may be a useful adjunctive treatment for combat veterans whose difficulties include existential concerns such as a foreshortened future, survivor guilt, and a sense of nihilism. A variety of interventions grounded in existential concerns are reported in the psychosomatic psychiatry literature in recognition of the significance of such issues for patients coping with severe or chronic medical illness. Group existential psychotherapies are increasingly reported on as adjunctive treatment for cancer patients. Brief psychotherapeutic interventions for these and other medically hospitalized patients may help the patient move from a position of defeat and hopelessness to one of resilience, agency, and connection with others.

Ethical Issues, Research, and Evaluation

There has been little systematic study of the indications or outcomes of existential psychotherapies. It may be that the emphasis on the individual's authentic and autonomous creation of meanings is especially challenging to systematize, and this has created an impediment to empirical study.

The existential therapies aim to understand patients according to their unique life situation and to foster their authentic choices, without forcing them to fit into a preordained theory of mental structure or development. The inevitable influence of the therapist's subjectivity, however, has made it increasingly difficult to accept the existential assumption that the patient can be perceived uncontaminated by the therapist's theory and technique. Any attempt to restate the patient's narrative, distill central themes, or define affective currents is likely to be influenced by the therapist's theories of mind, conflict, development, and change. The existential

psychotherapists have, however, provided a valuable service to the field of psychotherapy by their impassioned attention to the individual patient's experience and choices, and their conviction that it is all too easy for therapists to collude in an avoidance of the existential challenges of mortality, meaning, and disconnection. As a result, the priorities, interests, and techniques of the existential therapies are today often integrated into the work of eclectic therapists of different persuasions.

GUIDED IMAGERY THERAPY

History

The use of guided imagery or guided affective imagery as a form of therapy was systematized by Hanscarl Leuner in 1954. Leuner considered himself to be an analyst, and he was influenced by Carl Gustav Jung's interest in active imagination and Johannes Heinrich Schultze's autogenic training. Having confronted himself with disturbing emotions and images during his self-analysis, Jung thought that spontaneous, inner, dream-like images were translations of unconscious emotions and that observing and interpreting them could bring unconscious contents to consciousness. He called the technique *active imagination,* and he made it a part of what clinicians following his analytical tradition would use. Schultze framed the principles of *autogenic training* as a form of self-hypnosis and relaxation training, a mind- and organ-strengthening system based on expansive powers alleged to reside in the mind.

Leuner found that resting comfortably with the eyes closed for some time led to the appearance of images and that an exploration of the images with a guide was a pleasant and beneficial experience. He induced the images actively to shorten the interval before their appearance, added analytical insight to the guided exploration, and so framed *guided affective imagery*.

The ideas of guided affective imagery caught on in the United States and other Western countries, and the treatment was reported to be useful in chronic neuroses, personality disorders, depression, phobias, OCDs, sexual disorders, hyperhidrosis, substance abuse, anorexia nervosa, ulcerative colitis, and hives. The niche it filled in organized psychiatry, however, was never more than a small one. The ancillary disciplines of medicine, particularly nursing, embraced it with more enthusiasm than did psychiatry, and positive claims have been made regarding its use in postsurgical pain control, in general adaptation to illness, and, particularly, in coping and dealing with incurable, progressive disease.

Theoretical Issues

Guided affective imagery attempts systematically and psychotherapeutically to use the human ability to create images. It holds

that so-called catathymic images, images appearing before one's closed eyes as if on a screen and ultimately coming together and taking on movement to tell a story, represent a person's interpersonal world in autosymbolic imaginal form. The images include color, plasticity, and action, and the more that one is absorbed in them, the richer, more colorful, and more subtle they become. The symbols and the stories project the person's conflicts. Psychoanalytically informed guidance as the person recounts the experience can materially enhance conflict resolution by manipulating the symbols or assisting the story toward a successful outcome. The crucial idea is that influencing the symbols and the story influences the conflict.

Techniques

The patient is placed in a comfortable chair and is taught to practice relaxation, with suggestions of warmth, heaviness, and regular breathing. The patient is then invited to imagine a harmless, pleasant image, such as a flower, and is asked to describe it in detail, to look in its center, or to touch or smell it. Involvement in the images heightens the relaxation and vice versa; the patient soon becomes able to induce a readiness for imagery quickly. The hypnoidal state is ended with a command, such as "Clench your fists, bend your arms, open your eyes, and you are back, awake and refreshed."

When such learning is well established, the therapist relaxes the patient once more and suggests an image to be used in the therapy—for example, a meadow. The patient is told to nod when it comes into view and then to describe the scene.

Specific images used in the therapy are classified as basic, intermediate, or advanced, according to the level of conflict that they are likely to involve and the skill and training required by the therapist. Basic images include a meadow, a brook, a mountain, a house, and a wood; intermediate images include a car ride and a rosebush (representing women's and men's sexual feelings toward the opposite sex); and advanced images are such things as a swamp, a volcano, and a lion.

Descriptions of a meadow or a brook (the patient may see a brook in a meadow spontaneously or may be invited to imagine one) reflect current psychological issues. The meadow may be rich and green or scorched to stubble; the brook may be cool and inviting or a raging torrent. The therapist follows and guides with questions, such as "Can you describe that?" "What do you think of it?" or "What do you see as you follow the brook?"

The therapist's understanding is informed by knowledge and theory and determines whether the therapist interprets to patients or leads them to helpful activity. The therapist may suggest, for example, that a darkening scene reflects the patient's inner mood or that too perfect a picture shows

an effort to mask distress. If patients encounter a fence in the meadow or steep rocks as they follow the brook, the therapist may suggest that the patients climb over or find a way around the obstacle.

A young male patient saw a mountain from which rose another mountain and an observation tower. The young man had problems with authority and would stutter in the presence of his father or his boss. The therapist interpreted the scene to himself to mean that the young man felt watched and that too much was expected of him. The therapist guided him to ascend the tower, which the patient did successfully but amid images of fearful wandering and threats of thunder; the symptoms resolved.

The suggested motifs represent different aspects of the patient's psychology: The house and the edge of the wood represent one's own person; the brook represents vitality and energy; the mountain and its panorama represent the patient's view of the past, present, and future. The therapist looks to see whether supplies and a warm fire are in the house, what comes out of the wood, and whether the brook is cool and refreshing or an icy harbor of dangerous creatures. The management principle is to resolve the conflict symbolically.

One young student got over a block about taking an oral examination by being led in fantasy to feed a dangerous animal (the professor) coming out of a wood; those fearful of their own impulses might do well to approach such animals, pat them, and make friends with them. The guide conducts the exploration, using the symbols as a protective disguise, with empathy and warmth, positive feedback, praise, and, at times, direct suggestion.

Clinical and Ethical Issues

Guided affective imagery combines hypnotherapy with certain psychoanalytic ideas. Its major assumption is that manipulation of the symbols of conflict can resolve the conflict, a tenet that is open to question. The supportive relationship, the sense of involvement and effectuating something through the process, and the direct suggestions must contribute considerably to whatever effectiveness the technique has.

Interest in guided image therapy as a formal, psychoanalytically based psychotherapy has waned. Truncated and modified, however, guided imagery has gained a following in pain control, holistic medicine, and the treatment of incurable illness. In pain control, it has been reported to be useful as a relaxation and distraction technique, with additional suggestions of finding a second and then several seconds of decreased pain, and so extending the period of pain relief. The appeal in holistic medicine is to the limitless powers of the mind, a hopeful approach that has found favor in the nursing of patients with cancer and other terminal illnesses.

FEMINIST PSYCHOTHERAPY

History

Feminist psychotherapy is a product of the women's movement. It is an eclectic mix of therapies rather than a particular form of treatment. Its uniting aims, shared with the feminist movement, are to empower women and to raise their self-esteem.

Theoretical Issues

The overarching principles guiding feminist psychotherapy are that the personal is political and that millennia of sexist attitudes in the socialization of women have harmed them. The aims of treatment are to recognize the harmful effects of the sexist society, to explore inherent contradictions in the roles prescribed for women, support women in the exploration of their inner resources and their capacities for being nurturing, and demystify and rectify power relationships.

Its subsidiary ideas are that oppression leads to feelings of impotence and true impotence and to low self-esteem and conflicts. Traditional therapy is seen as being a mechanism of social control in the service of the status quo and as being based on an imbalance, even in the therapeutic relationship. Feminist psychotherapy rejects such a model and rejects the idea of conflict developing and residing within the person without regard to the socioeconomic system. It tries, therefore, to equalize power by modeling, self-disclosure, and disclaiming expertise. It focuses on how the stresses of a patriarchal society cause symptoms—for example, how depression, a lack of assertiveness, and job difficulties may grow from assigned sex roles and sexual harassment, with inner conflict being only secondary to exterior conditions; how a lesbian lifestyle may be a choice to love and to commit to a woman and may have nothing to do with an unresolved dislike of men; and how, if sexual intercourse is male aggression, a woman's best defense may be in sexual symptoms, such as dyspareunia and vaginismus.

Feminist psychotherapy believes that women may inadvertently contribute to the transmission of oppressive attitudes by teaching their daughters to forgo their own needs and to nurture and to please others, as mothers do themselves in child rearing. Conflicts between autonomy and dependence result and are not resolved by therapies that help women endure their suffering, mask or suppress their anger, or rely on antidepressants and sedatives.

Techniques

Techniques vary with the type of therapy being adapted for feminist use. Psychodynamic therapy and expressive and supportive therapies are used, and guided imagery has been invoked. Most of the literature, however, places a heavy emphasis on the cognitive approach. Steps include

redefining power and reducing guilt; defining one's own feelings and respecting them, thus defining one's own personal reality; emphasizing responsibility for one's choices and actions; working through the issues raised in therapy; and terminating therapy when the patient accepts her own needs as legitimate, understands the power differentials that hinder her, can assert herself appropriately, and needs the therapist less than in the past.

Cognitive beliefs common to women (which are also common to men) include the idea that a woman must be loved by everyone, that others' needs count more than her own, that she needs the support of someone strong, and that she has no control over her emotions. Recognizing those beliefs, disputing them, and practicing behavior based on alternative views are parts of the treatment; assertiveness training is also commonly recommended.

Research and Evaluation

Many of the ideas that feminist therapy has generated have been integrated into the mainstream. Feminists called attention to the facts that women have suffered from unequal power relationships and that male dominance in a family is taken for granted. They also noted that the description *absent father* is a less extreme judgment of the person than is the description *maternal deprivation;* and that the term *dominant mother* carries a similar pejorative quality. Other ideas that seemed radical when first introduced but are now mainstream include—for example, that boys do not so much catch up in school performance as that girls fall behind, because their type of understanding (feelingful and intuitive) is less valued than male understanding in a structure that is focused on adapting to male-dominated ideas. The observation that societal oppression contributes to the ills women experience, however, need not exclude the possibility that women may have personal, internal conflicts that are based on their individual histories.

TRANSPERSONAL PSYCHOTHERAPY

History

Transpersonal psychology grew out of the humanistic movement of the 1960s. Defining itself as "the personal, plus," it means to include all facets of personal psychology and psychiatry *plus* deeper or higher, transcendent experiences. It thus embraces a range of religious and spiritual experiences and parapsychological phenomena, such as conversion; crises of faith; ecstasy, peak, and near-death experiences; spiritual mystical experiences; and past-life experiences. Although the term *transpersonal* was first used by William James and later was expanded on by Jung, it was writers such as Abraham Maslow, Robert Assagioli, Ken Wilber, and Stanislav Grof who

generated its theoretical underpinnings. Each of these authors developed theories of needs and motivations or of stages of development that encompassed and were roughly modeled on traditional developmental theory but added the extra dimensions required by the transpersonal emphasis. The *Journal of Transpersonal Psychology* started publication in 1969 and continues to exist. The first Council Grove Conference, sponsored by the Menninger Clinic that same year, focused on topics such as meditation, biofeedback, yoga, and psychedelic agents. Since then, several transpersonally oriented institutions were founded, mostly in California, and, for the past several years, approximately 15 dissertations dealing with transpersonal experiences are written yearly. The *Diagnostic and Statistical Manual of Mental Disorders,* Fifth edition (DSM-5) includes religious and spiritual problems under Other Conditions That May Be a Focus of Clinical Attention.

Theoretical Issues

In the free-for-all of theory building, several authors have contributed to transpersonal psychology and psychiatry, but no single contribution dominates. Thus, the theory is an overlapping pastiche rather than a coherent whole: Maslow studied a series of self-actualized persons and found them similar to traditional descriptions of enlightened persons. He concluded that human beings have a biologically based instinctive nature that is fulfilled in spiritual self-actualization, and he and Grof coined the term and founded the field of *transpersonal psychology*. Assagioli suggested that each person has many subpersonalities and that a spiritual psychosynthesis is needed to unleash the transforming and creative energies of a *super conscious,* which also exists. Wilber proposes prepersonal, personal, and transpersonal developmental stages; disorders of the prepersonal developmental level result in infantile psychoses or narcissistic or borderline personality disorders, disturbances at the personal level give rise to neuroses, and spiritual emergencies, kundalini crises, and dark nights of the soul rest on disturbances at the transpersonal level. Grof, during 17 years as scholar in residence at the Esalen Institute in California, experimented widely with psychedelics in psychotherapy and proposed that, in psychedelic and other altered states of consciousness, patients re-experience the trauma of their birth. He suggested the existence of four basic perinatal matrices (BPMs) representing experiences from the entry into the intrauterine state (BPM-I), entry to the birth experience (BPM-II), passage toward a light with intense anxiety and near-death experiences (BPM-III), and culmination in total surrender, a sense of overwhelming and unending love, light, expansion, and forgiveness (BPM-IV). After experiencing or re-experiencing BPM-IV, in adulthood, further transpersonal expansions may occur in two domains: expansion of ego boundaries in ordinary real terms and extensions beyond the usual

framework. The first includes a sense of identification with the earth and all of life in the universe, and the second includes spiritual communication, out-of-body experiences, and telepathic, precognitive, and other paranormal experiences. These experiences were often seen as psychotic by traditional psychologists and therefore constitute spiritual emergencies that require special help. Grof urged mental health professionals to learn to distinguish these nonordinary, transpersonal states from mental illness.

Techniques

Technique in transpersonal psychotherapy is an amalgam of the generally accepted and the esoteric, depending on the world view of the practitioner and the issues being addressed. The range is from the psychodynamic and behavioral, through meditative approaches or the recommendation of spiritual texts, to rapid-breathing techniques, guided imagery, and the use of psychedelic drugs to promote altered states of consciousness. The teachings are variously applied to a similar wide range of issues. *Defensive spirituality,* in which religious or spiritual beliefs are used to deny feelings or conflicts, or *offensive spirituality,* in which there is an obnoxious, self-reassuring, but intrusive stance of "having the one true path," may be amenable to dynamic techniques. Similarly, loss of faith or a conversion experience may be dealt with by a sensitive therapist to determine what factors are at work in the patient's life and whether the conversion, for example, is a legitimate search for truth and the comfort of faith or a magical denial doomed to fail. Expressive explorations of hallucinations or delusions or of a psychotic religious experience after the psychosis is controlled are recommended by some as a means of finding personal insights.

Transpersonal psychotherapists also try to distinguish genuine religious experiences from psychoses, the latter being often more terrifying, accompanied by deterioration in other areas of life, and containing delusions of being addressed by specific religious figures. The boundaries, however, are hard to define in *spiritual emergencies,* when the individual feels dazzled by truths too powerful to grasp, or in *kundalini awakenings,* states in which energies expand upward from the base of the spine and explosively alter consciousness. Supportive therapy is recommended in these cases, along with education about the experience. Meditation and the use of mantras, such as "May I be happy" and "May I be peaceful," are also reported to be useful if the patient is having anxiety over a crisis of faith. Near-death experiences may also lead to a transformed view of the world, and support is recommended along with help in "those parts of the ego that may have died." The traditional psychotherapist and psychiatrist may have difficulty in areas of transpersonal psychiatry that aim to help individuals recognize and realize a presumed "ultimate state that transcends the usual state of functioning, a state of mystical illumination or cosmic unity." Some

of these include experiences such as channeling, encounters with unidentified flying objects (UFOs), possession states, and past-life experiences. In these areas, the transpersonal literature is ambiguous, on the one hand, cautiously attesting that such experiences are meaningful to the person and embody content relevant to the individual's personality, and, on the other hand, accepting them as normative spiritual experiences, an ultimate state of mystic illumination or cosmic unity that is reflective of a reality beyond the norm.

Research and Evaluation

Research on the paranormal, although always of interest because of its supernatural allure, has failed to confirm the legitimacy of past-life experiences, precognition, psychokinesis, and similar phenomena. Also, because transpersonal psychotherapy is not a single discipline, there have been no systematic studies of its efficacy. The field emphasizes the personally meaningful and possibly transformative effects of things such as conversions or near-death experiences; the importance of religion and mysticism in the lives of many; the difficulty of teasing apart the boundaries between psychiatric illness, personal mystique, and cultural differences in expected experience; and the need for sensitivity in working with patients whose life and experiences are different from one's own.

DIGNITY THERAPY

History

Beginning in the 1960s in the United Kingdom and gaining traction in the United States in the 1980s, the modern hospice movement changed the way in which end-of-life care is delivered, placing an emphasis on maintenance of functioning and physical comfort over aggressive illness-focused treatment. The subsequent establishment of palliative care as a board-certified internal medicine sub-specialty in 2006 brought further credibility to the approach championed by the modern hospice movement. The traditional palliative care model uses a multidisciplinary approach to address the far-ranging issues associated with terminal illness, including pain management, assistance with the provision of home medical care, and education and support for caregivers. While the scientific evidence overwhelmingly demonstrates positive gains in quality of life for both patients and loved-ones when palliative care methods are utilized, observational studies continue to demonstrate high rates of depression and anxiety in patients with terminal illness. Complicating matters, research also suggests patients with terminal illness, especially those who become sick later in life, do not respond well to traditional antidepressant and anxiolytic treatment approaches.

Beginning in the late 1990s, a group at the University of Manitoba led

by Dr. Harvey Max Chochinov set out to address the both prevention of psychological distress and improving response to psychiatric treatment in end-of-life care. Dr. Chochinov and his research group designed and tested a brief, manualized psychotherapeutic approach coined "Dignity Therapy" that was shown to reduce psychological distress in both patients with terminal illness and their caregivers, leading to improved outcomes in quality of life and lower rates of depression and anxiety. Subsequent to the publication of Dr. Chochinov's research, other groups tested the approach and found similar rates of effectiveness across an even wider range of diagnoses and conditions. This growing body of evidence has led some to call for similar approaches to be incorporated not only into routine palliative care practices, but also into the care of those with debilitating but nonterminal chronic medical illness.

Theoretical Issues

It is not entirely clear why patients with terminal illness have a high prevalence of depression and anxiety, and why they may not respond as well to traditional therapies. A common formulation is that the decline in physical, cognitive, and social functioning associated with terminal illness leads to a loss of "dignity" on the part of the patient, which in turn, leads to feelings of helplessness, hopelessness, isolation, and despair. Therefore, in order to prevent the development of depression and anxiety and to treat the symptoms of these conditions if they emerge, one must find a way to maintain a person's dignity as he or she moves toward death.

One of the clear results from dignity research in terminal illness is the mere exploration of dignity with most dying patients leads to the desired therapeutic effect. For some patients, though, a more systematic process is required. When thinking about the development of a systematized management approach, a complex picture emerges. First, dignity is a highly subjective and variable concept, and so common elements must be identified that cut across populations within the context of end-of-life care. Then, careful analysis and categorization of these common elements is necessary so that a cohesive model for dignity can be formed. Finally, a therapeutic process that operationalizes this model must be tested to gauge the validity of the model and the effect of the therapeutic process on those parts of a person's identity that contribute to dignity. Finally, the model of dignity must be sufficiently dynamic to account for the physical, mental, and social changes that occur with terminal illness, and the therapeutic process must be both accessible and engaging for patients whose ability to participate in traditional psychotherapeutic approaches often is compromised.

Drawing on techniques of narrative, client-centered, and existential psychotherapies and set within a framework of dignity grounded in Eriksonian developmental ideals, Dignity Therapy seeks to assist patients

as they strive to resolve questions and internal conflict surrounding the terminal and often debilitating nature of their illness, their evolving role in pre-existing social constructs, and the impact these two factors have on self-identity.

Dignity Therapy aims to promote a lasting sense of generativity while simultaneously engaging in the development of ego integrity. The underlying idea is that very often, terminal illness develops during the final years of one's middle adulthood. It also, then, compresses late adulthood into a much shorter timeframe. One of the keys, therefore, to alleviate psychological distress is to help patients complete the developmental tasks of these life stages. In order to accomplish this feat, the patients are charged with telling their story and describing themselves in their own words, while being prompted by the interviewer with probing questions focused on the exploration of their life's meaning and being refocused continually on those things that appear most central and engender the greatest sense of pride. The end result is a retrospective process that helps patients see their life through a lens of contentment and satisfaction, along with the production of a written portrait, with which patients are allowed to do as they please.

Techniques

Dignity Therapy is a brief, manualized psychotherapeutic treatment intended to be delivered on an individual basis by a qualified mental health professional. The treatment typically requires three to four sessions, spaced over a week's time. The duration of each patient visit varies with the specific goals of the session and the volume of material discussed. Typically, though, sessions last between 30 and 60 minutes. Dignity Therapy has been validated in outpatient, residential, and inpatient settings.

The basic structure of a course of Dignity Therapy is as follows:

Session 1

The first session consists of four main components.

A. Administration of rating scales—These can be self-report or clinician-administered measurements. The tools used should include, at a minimum, screening questions for depression, anxiety, hopelessness, and suicidality, as well as ratings for sense of well-being and quality of life.
B. Introduction to the process and goals of therapy—Special attention should be given to the collaborative nature of the therapy and the production of the "generativity document," which is the edited written transcript provided to the patient at the conclusion of the therapy.
C. Informed consent for treatment and audio-recording—This will have

critical importance when considering the ethical implications of the treatment.

D. Written copy of the questions included in the Dignity Therapy protocol provided for review and consideration—The same nine questions are used in every course of treatment.

Session 2

The second session is the audio-recording session and begins with the therapist asking the patient the first question on the Dignity Therapy question protocol:

"Tell me a little about your life history; particularly the parts that you either remember most or think are the most important?"

The therapist uses the remaining questions included in the Dignity Therapy protocol in a flexible manner, following the patient's lead and paying particular attention to the patient's affect to determine when to ask relevant follow-up questions. The goal is not to obtain a response to each question included in the protocol, but to prompt the patient to verbalize those aspects of his or her life that are viewed as essential to positive self-identity. The therapist utilizes a combination of reflective listening, displays of empathy, dynamic clarification, and existential probing. Additionally, the use of guided-imagery and relaxation techniques can be utilized if patients struggle to put together their thoughts or experience acute psychological distress.

Following Session 2, the audio-recording is transcribed and edited into a coherent narrative that conveys the tone and message desired by the patient. To this effect, an essential component of this editing process is finding a single, concluding statement that captures the patient's desired lasting legacy.

Session 3

During the third session, the narrative produced from the edited transcript of audio-recorded session is read aloud to the patient. This allows the patient to provide feedback on the historical accuracy of the narrative and to determine if there are any anecdotes or information that should be removed or included. In addition, reflecting upon their own self-portrait evokes a strong emotional response, and this helps to solidify ego integrity.

Session 4

Any feedback provided by the patient is used to create a final draft of the narrative, which is termed the "generativity document." In the final session, this document is provided to the patient, and the patient's plans for the document are considered. Additionally, the patient is invited to reflect upon the therapeutic process. Any concerns expressed by the patient

are clarified and addressed. Finally, the initial rating scales are readministered to gauge the therapeutic effect of the treatment and to determine if additional treatment steps are necessary.

Clinical Issues

The original research validating Dignity Therapy was performed within a fairly narrow patient population, namely, adults from Canada and Australia in hospice-care settings with terminal illness (primarily cancer) and a life expectancy of less than six months. Further studies have included more variable demographics, different illness types, and lesser illness severity, including multiple studies done with elderly patients in long-term care facilities without a specific terminal illness.

Notably absent from most of these studies are specific inclusion or exclusion criteria for pre-existing psychiatric symptoms or psychological distress, although these were identified in a large proportion of patients as part of the research protocols. Since all of the studies demonstrated effectiveness across multiple domains (dignity, quality of life, depression, anxiety, hopelessness, suffering, etc.) for the whole of the included populations, the research findings suggest Dignity Therapy is indicated not just for those persons identified as depressed or anxious, but for all persons as they approach the end-of-life. It should be noted, however, that those persons identified during study enrollment as having significant psychiatric symptoms or psychological distress experienced the most significant positive effect from a course of Dignity Therapy.

While Dignity Therapy may have widespread application, there are two particular patient groups for whom the effectiveness of the treatment is unknown. Dignity Therapy has not been studied in pediatric or adolescent populations, or in patients with significant cognitive impairment. The current approach requires a level of patient engagement and awareness that excludes younger patients and those with cognitive deficits. It is possible that Dignity Therapy can be adapted to these populations, but additional study will be required.

Ethical Issues

Helping patients resolve inner psychological conflict that is influenced by their terminal condition is a central feature of Dignity Therapy. Although acceptance of death is not a primary goal of the treatment, this process often leads to a greater acceptance of human mortality. It is therefore possible to imagine situations in which Dignity Therapy could be misused with the purpose of influencing patients as they consider various treatment options to either prolong or foreshorten life. Another risk of Dignity Therapy is the potential impact on matters of estate and other legal issues of the generativity document and the audio-recording from which it is

transcribed. This is why it is essential to obtain the patient's informed consent for both treatment and audio-recording and to review the generativity document with the patient before it is revealed to others.

Research and Evaluation

The research generally demonstrates beneficial effects on psychiatric symptoms and psychological distress in those who are suffering. Additionally, Dignity Therapy leads to improvements in quality of life and overall sense of well-being in patients who are not struggling with the psychological impact of their illness. Finally, there is evidence Dignity Therapy has a positive impact on caregivers.

Future research considerations likely include expansion of the Dignity Therapy approach to nonterminal illness, the study of Dignity Therapy in adolescent and young adult populations, and the adaptation of Dignity Therapy for persons who are unable to participate in the process as it currently is constructed (e.g., children and those with cognitive impairments).

PSYCHOTHERAPY FOR THE DEAF

History

Psychotherapy for the deaf is not a particular type of psychotherapy but a group of therapies that are adapted to the needs of a special population—people who develop a profound hearing impairment before speech is established. The impairment is defined operationally as one that renders audition useless as a means of intercourse with the hearing world, and it generally involves a hearing loss of 80 dB or more across the speech frequencies, even with the aid of amplification. Its most obvious consequence is to interfere with the ability to learn language by simple auditory imitation and at the normal time and pace in development.

Psychiatric interest in the deaf began with the work of Franz Kallmann, John Rainer, and Kenneth Altshuler in the mid-1950s. Then the country's most prominent geneticist in psychiatry, Kallmann was approached by a branch of what is now the U.S. Department of Health and Human Services for assistance in obtaining case closure among psychiatrically ill deaf clients. Kallmann recruited Rainer and Altshuler and began a series of normative studies and studies of deaf persons with mental illness. The first effort established an outpatient psychiatric clinic, and subsequent projects developed an inpatient unit for statewide services, a vocational rehabilitation liaison and halfway house program, and consultation programs at a school for the deaf. Centers focused on psychiatry and the deaf were subsequently established in Chicago, San Francisco, Washington, and several other cities and countries around the world. At present, a small caucus of psychiatrists interested in working with the profoundly deaf

meets regularly at annual meetings of the American Psychiatric Association.

Theoretical Issues

To adapt a given psychotherapy to the needs of persons with profound early deafness demands that the therapist understand what the handicap means to one's life experience and that he or she take into account the limitations that this experience places on the therapeutic exchange.

A first guiding principle is that the simple absence of audition has no relevance to innate intelligence potentials; intelligence quotients (IQs) among the deaf follow the same bell-shaped curve as in the hearing. Among the deaf, performance scores are generally higher than verbal scores because of the lag in language development. Deaf children of deaf parents develop vocabulary at the same rate as their hearing peers, albeit in sign language, because their parents are able to communicate with them freely. Occasionally, an enlightened hearing parent learns sign language to provide the only free, two-way accessibility that the handicap allows.

Too often, however, parents are caught up in an educational controversy that has gone on since Alexander Graham Bell and the 1880s, about whether signing discourages the learning of speech. Evidence suggests that an educational approach that embraces all avenues of communication—combining oralism, speech and lipreading instruction, amplification, and sign language—gives the best results. Moreover, the combination supports the exercise of age-appropriate interests and language skills in accord with their natural developmental timetable. In the absence of a widespread combined educational approach, studies in the 1970s showed that the average deaf adolescent left high school with approximately fourth-grade and fifth-grade English and mathematical skills.

As a result of these findings and the urging of mental health professionals, since the 1970s, there has been a major shift in educational practices from a nearly exclusive emphasis on oralism toward some variety of the combined mode of instruction. Pockets of resistance exist and the recent availability of cochlear implants at a relatively early age may recast expectations and heat up the issue again.

A factor contributing to the endurance of the controversy is that parents are understandably hurt and saddened when the handicap is recognized, and they wish that it would go away. Motivated partly by denial and partly by the commendable desire that their child's life be like any other child's life, parents resist sign language as the visible symbol of the handicap and the harbinger of a separate and sequestered existence. What is lost in the face of these pressures is the recognition that language, generally mediated through audition, brings a tremendous amount of information and experience. Without hearing or the effective alternative of sign language, uncountable hours of speech and lipreading instruction can only provide

language late and often incompletely.

Sound also brings nuance and music, and it is the sole distance receptor that scans 360 degrees of the environment at all times, perhaps even in sleep. What may be associated with missing those elements of experience is uncertain.

A final concomitant of deafness, at least in the 90 percent of those born to hearing parents, is that the parents of deaf children feel confused, uncertain, and, often, guilty. Their own ambivalence, the conflicting advice that they may receive, and the stranger to communication in their midst often makes them overprotective, inconsistent in discipline, and erratic in their hopes and approach.

As a result of all of those factors but probably influenced primarily by the language lag and the lack of general experience, deaf persons have been described as more naïve, impulsive, concrete, and immature than their hearing counterparts. Such descriptions apply only as a composite, and the deaf should not be stereotyped. In a large enough number, more deaf persons may show more of those traits than their hearing peers, but the presence and the extent of the traits in a particular person are not predictable.

Techniques

Communication is the sine qua non of psychotherapy. The first requirement with the deaf is to understand what life experiences the handicap entails; the second requirement is to interact freely with patients so as to apprehend their personal and particular experiences and to exchange understanding about the details. By adulthood, virtually every person with early profound deafness is comfortable and fluent in sign language (barring an accompanying organic problem that interferes with language acquisition, regardless of mode). Having the therapist equally versed in manual language would be ideal but is rarely possible, despite the fact that a growing number of psychiatrists and others have learned to sign. The problem is lessened to the extent that the patient's language, speech, and lipreading skills are developed. Often, these skills are insufficient to allow the conduct of meaningful psychotherapy, even with exchanges in writing. In such instances, an interpreter is necessary. Deaf persons often rely on interpreters, and the therapist, rather than the patient, may be the one who is uncomfortable with the third party's presence. One must be alert, however, to the interpreter's tendency to interpret when the therapist wants an exact translation; for example, when the patient's thoughts are to be evaluated, the interpreter's efforts to make sense of them may be counterproductive. Supportive psychotherapy can be conducted with the aid of an interpreter, but substantial insightful work probably demands a subtlety of understanding that is only available with manual skill on the part of the therapist or adept speech and lipreading skills on the part of the

patient.

Clinical Issues

Less than profound deafness and deafness after speech is well established are problems of a different nature and require different adaptations. To the extent that language has been developed before the loss of hearing, and to the extent that the loss is remediable by amplification, the patient's life experiences are less affected, and therapy requires less in the way of modification. However, the loss itself, its cause, and the patient's changed life circumstances because of it are factors that affect the person's problems and the therapy.

A couple of additional points must be kept in mind in counseling the hearing parents of deaf children. Eager for their child to have a normal life, the parents may be impatient with modest results in attempts to remediate the handicap and its effects. Complaints of bad treatment at the hands of many institutions or agencies must be evaluated with that possibility in mind. Also, denial on the part of a family may be brought to an abrupt end with the child's impending adolescence. When parallel play and participation in simple games give way, and the horizons of social interaction loom large, the effects of the handicap become unavoidable, and parents may become depressed.

Research

The educational, developmental, and psychiatric literature on early profound deafness is considerable, but no controlled studies of psychotherapy have been reported. The psychiatric literature, mostly from the 1970s and before, suggests that major mental illness is no more prevalent among the deaf than in the hearing, that impulsive or aggressive behavior may be more common in the presentation of illness among the deaf, and that guilt and psychomotor retardation may be less frequent in cases of depressive disorders in the deaf. Anecdotal reports also indicate that various forms of psychotherapy are as useful in the deaf as in the hearing when modifications in view of the handicap are made.

COMPUTERS AND PSYCHOTHERAPY, PSYCHOTHERAPY BY NONPROFESSIONALS, AND PSYCHOTHERAPY UNDER MANAGED CARE

History and Techniques

Psychotherapy by means of a computer, psychotherapy by nonprofessionals, and brief psychotherapy under managed care are not particular schools of treatment. What they share are the aims of cutting costs, thereby making therapy more widely available. They lower cost by abbreviating treatments or by substituting less expensive providers for more expensive ones.

Computer programs have been written in the last several years to provide educational materials to patients or to interact with them in a more truly therapeutic way, for example, aiming at enhanced self-awareness or self-control. Psychotherapy by nonprofessionals uses common sense and supportive approaches that derive from general experience. Brief therapy under managed care rests on the thesis that dynamic understanding can be used episodically and that widely spaced sessions providing insight and support can help most patients overcome their acute problems in living.

Theoretical Issues

Each approach raises challenging issues of theory. Are the comfort and the flexibility of a real person requisite for effective psychotherapy, or can a computer be individually responsive enough, and can patients develop transferences and a sense of safety relating to it? Is professional training necessary to obtain satisfactory results, or can well-adjusted and mature adults do just as well? Are regular meetings, weekly or more often, cogent therapeutic agents, or can monthly meetings be equally efficacious?

Clinical and Ethical Issues

As with all therapies, these novel approaches have been reported to be successful and have gained a following. They pose considerable challenge to traditional therapies, calling into question many of the basic and revered assumptions on which the longer-established methods rest. Because the treatments are new, and research on any of them is in its infancy, it is likely that early claims will moderate over time, as research establishes their limits. Meanwhile, the fact that they are experimental and unproven should be kept in mind and clearly communicated to patients. There is little danger in the computer-driven and the nonprofessional efforts, because their innovative nature makes informed consent and scrutiny by institutional review boards mandatory before treatment under such conditions is initiated. The current emphasis on cost containment, however, does not afford the same protection in the case of the minimalist approaches being instituted by enthusiastic managers of care. In such cases, patients may need better information than is now offered, if they are to make an informed decision.

FUTURE DIRECTIONS

There already exists convincing evidence that psychotherapy is associated with measurable changes in the brain. Indeed, as the methods for capturing the neurobiological correlates of psychotherapy outcome become more refined and nuanced, it will allow clinicians to more easily discern those treatment approaches that have an effect beyond the “common factors” found in all forms of psychotherapeutic engagement. This will provide a

means of demonstrating the value, or lack thereof, in specific psychotherapeutic modalities. Additionally, as the current understanding of the role played by certain parts of the brain in psychopathology expands, gaining insight into the ways in which psychotherapy affects the brain might allow for the creation of new psychotherapies targeted at creating change in those parts of the brain most tightly linked with specific disease processes or illness states.

Perhaps even more exciting than the ability to measure or gauge treatment effect within the brain are the research efforts to define "biomarkers" for psychiatric illness. Proposed biomarkers range from blood tests to functional neuroimaging and advanced EEG techniques. While the development of biomarkers may lead to redefinitions of psychiatric conditions, these lines of research already show promise in defining subgroups of patients who may be responsive to particular therapeutic interventions. This suggests that in the near future, specific psychiatric treatments, including particular psychotherapeutic approaches, can be prescribed to individual patients based on phenotypic profiles of biomarkers that indicate a greater likelihood positive response.

References

Altshuler KZ. Studies of the deaf: relevance to psychiatric theory. *Am J Psychiatry*. 1971;127:1521–1526.

Altshuler KZ. Will the psychotherapies yield differential results: A look at assumptions in therapy trials. *Am J Psychother*. 1989;43:310–320.

Berne E. *Transactional Analysis in Psychotherapy*. New York: Grove; 1961.

Berne E. *Games People Play: The Psychology of Human Relationships*. New York: Castle; 1964.

Boorstein S. *Clinical Studies in Transpersonal Psychotherapy*. Albany, NY: State University of New York Press; 1997.

Bozarth JD, Zimring FM, Tausch R. Client centered psychotherapy: the evolution of a revolution. In: Cain DJ, ed. *Humanistic Psychotherapies: Handbook of Research and Practice*. Washington, DC: American Psychological Association; 2002.

Brown LS. Still subversive after all these years: the relevance of feminist therapy in the age of evidence-based practice. *Psychol Women Q*. 2006;30:15–24.

Brown LS. *Feminist Therapy*. Washington, DC: American Psychological Association Press; 2010.

Chochinov HM, Hack T, Hassard T, Kristjanson LJ, McClement S, Harlos M. Dignity therapy: a novel psychotherapeutic intervention for patients near the end of life. *J Clin Oncol*. 2005;20:5520–5525.

Cooper M. *Existential Therapies*. London: Sage Publications; 2003.

Dolezal-Wood S, Belar CD, Snibbe J. A comparison of computer-assisted psychotherapy and cognitive-behavioral therapy in groups. *J Clin Psycol Med Settings*. 1998;5:103–115.

Dutton-Douglas MA, Walker LE. *Feminist Psychotherapies: Integration of Therapeutic and Feminist Systems*. Norwood, NJ: Ablex; 1988.

Etkin A, Pittenger C, Polan J, Kandel ER. Toward a neurobiology of psychotherapy: basic science

and clinical applications. *J Neuropsychiatry Clin Neurosci*. 2005;17:145–158.

Gill IJ, Fox JR. A qualitative meta-synthesis on the experience of psychotherapy for deaf and hard-of-hearing people. *Ment Health Relig Cult*. 2012;15:637–651.

Greyson B. The near-death experience as a transpersonal crisis. In: Battista JR, Chinen AB, Scotton BW, eds. *Textbook of Transpersonal Psychiatry and Psychology*. New York: Basic Books; 1996.

Griffith JL, Gaby L. Brief psychotherapy at the bedside: countering demoralization from medical illness. *Psychosomatics*. 2005;46:109–116.

Haley J. *The Uncommon Therapy*. 2nd ed. New York; Norton; 1986.

Jones L, River EM. Current uses of imagery in cognitive and behavioral therapies. In: Vande Creek L, Knapp S, Jackson TL, eds. *Innovations in Clinical Practice: A Source Book*. Sarasota, FL: Professional Resource Press; 1997.

Lytle LR, Lewis JW. In: Glickman NS, Harvey MA, eds. *Culturally Affirmative Psychotherapy with Deaf Persons*. Mahwah, NJ: Lawrence Erlbaum Associates; 1996.

Markman KD, Proulx T, Lindberg MJ, eds. Psychotherapy and the restoration of meaning: existential philosophy in clinical practice. Washington, DC: American Psychological Association; 2013.

May R, Angel EE, Ellenberger HF. *Existence: A New Dimension in Psychiatry and Psychology*. New York: Simon and Schuster; 1958.

Meadow-Orlans KP. *Deafness and Child Development*. Berkeley, CA: University of California Press; 1980.

Perez VB, Swerdlow NR, Braff DL, Naatanen R, Ligh GH. Biomarkers in schizophrenia: using biomarkers to inform diagnosis, guide treatments, and track response to interventions in psychotic illnesses. *Biomark Med*. 2014;8:9–14.

Rainer JD, Altshuler KZ, Kallmann FJ, eds. *Family and Mental Health Problems in a Deaf Population*. New York: Springer; 1963.

Rogers CR. *Client-Centered Psychotherapy*. Boston, MA: Houghton Mifflin; 1951.

Schultz JH, Luthe W. *Autogenic Training: A Psychophysiologistic Approach in Psychotherapy*. New York: Grune and Stratton; 1959.

Shmukler D. Reflections on transactional analysis in the context of contemporary relational approaches. *Trans Anal J*. 2001;31:94–102.

Singer JL. *Imagery in Psychotherapy*. Washington, DC: American Psychological Association; 2005.

Southwick SM, Gilmartin R, Mcdonough P, Morrissey P. Logotherapy as an adjunctive treatment for chronic combat-related PTSD: A meaning-based intervention. *Am J Psychother*. 2006;60:161–174.

Summerton O. The development of gene analysis. *Trans Anal J*. 2000;30:207–218.

Szapocznik J, Williams RA. Brief strategic therapy: Twenty-five years of interplay among theory, research and practice in adolescent behavior and drug abuse. *Clin Child Fam Psychol Rev*. 2000;3:117–134.

Teasell RW, Shapiro AP. Strategic-behavioral intervention in the treatment of chronic nonorganic motor disorders. *Am J Physiol Med Rehab*. 1994;73:44–50.

Utay J, Miller M. Guided imagery as an effective therapeutic technique: A brief review of its

history and efficacy research. *J Instruct Psychol*. 2006;33:40–44.

Weeks GR, L'Abate L. *Paradoxical Psychotherapy: Theory and Practice with Individuals, Couples, and Families*. New York: Brunner/Mazel; 1984.

Yalom ID. *Existential Psychotherapy*. New York: Basic Books; 1980.

▲ 33.12 Narrative Psychiatry

BRADLEY LEWIS, M.D., PH.D.

INTRODUCTION

More than anything else psychiatrists do, they listen to stories. These stories so saturate the clinical encounter that it would be impossible to imagine a clinical encounter without them. In the very first meeting between psychiatrist and patient, the psychiatrist begins with an open-ended invitation to a story: "What brings you here?" or "What seems to be the problem?" Patients respond to these questions by telling psychiatrists about their life, their troubles, when the troubles began, what seems to have caused them, how they create difficulty, and what kinds of problem solving they have tried. Such stories may be rudimentary, they may be only partially worked out, and they may even be baffled and confused. The patient may even be perplexed enough to answer, "I don't know why I came" or "I'm not really sure what's wrong, my family sent me." Nonetheless, the patient's response to the psychiatrist's initial questions always involves a story.

As the interview proceeds, going into greater details about the person's life and troubles, the psychiatrist and patient never leave the story frame. After the initial broad questions, focused questions aimed at particular symptoms also invite stories. Even finely grained, highly medical sounding questions like "Do you have any medical problems?" or "Does depression run in your family?" involve stories. Moreover, the clinical evaluation is only the beginning of the story-centered encounter between psychiatrist and patient. Follow-up interviews and meetings go deep into a clinical treatment process, including those that only involve medication checks once every so often, are also saturated with stories. "How have you been?" "What has happened since we last spoke?" "Are you having any side effects with this medication?" These questions all elicit stories from the patient.

One might ask, however, what is the importance of these stories in clinical life? Until recently, little attention has been paid to the question of stories within clinical settings. A patient's narrative of his or her troubles were seen as merely a vehicle to get across the facts. The "art" of clinical work was to abstract away the facts embedded in patient's stories, but the stories themselves were secondary. In the last 30 years, however, the view of narrative as being merely functional has been thrown into question, and, increasingly, clinicians see stories as indispensable for understanding human phenomena. For these clinicians, stories are just as important as facts because it is the story that allows people to select, organize, and shape the facts that matter.

Narrative psychiatry emerges out of this increased interest in clinical stories. The two main tributaries that lead to narrative psychiatry come from the two different sides of psychiatry: "narrative medicine" and "narrative psychology and psychotherapy." Narrative psychiatrists are psychiatrists who combine the wisdom of these two domains.

Following the lead of narrative medicine, narrative psychiatrists recognize that psychiatric patients, like medical patients, come to clinics with intense stories to tell. Contemporary narrative medicine has developed from 30 years of work in bioethics and medical humanities devoted to humanizing the clinical encounter through a better understanding of patient stories. The term "narrative medicine" comes from Rita Charon, an internist and literary scholar, who used it to describe an approach to medicine that uses narrative approaches to augment scientific understandings of illness. Narrative medicine brings together insights from human-centered medical models, such as George Engel's biopsychosocial model and Eric Cassel's person-centered model, with research and insights from phenomenology, the humanities, and interpretive social sciences.

Narrative medicine uses these resources to better understand the illness experience, "to recognize, absorb, interpret, and be moved by the stories of illness." As Charon argued, when clinician's possess narrative competency, they can enter the clinical setting with a nuanced capacity for "attentive listening ..., adopting alien perspectives, following the narrative thread of the story of another, being curious about other peoples motives and experiences, and tolerating the uncertainty of stories." She further argued that doctors "*need* rigorous and disciplined training" in narrative reading and writing not just for their own sake (helping them to deal with the strains and traumas of clinical work), but also "*for the sake of their practice*." Without such narrative competency, clinicians lack the ability to fully understand their client's experience of illness. For Charon and others in narrative medicine, narrative study is not a mere adornment to a doctor's medical training; it is a crucial and basic science that must be mastered for medical practice.

A major task of narrative medicine, and therefore narrative psychiatry, is to be a good listener and to connect empathically with the patient's story. A narrative psychiatrist, like a narrative physician, seeks to understand the patient first and foremost. This understanding brings patient and clinician together into a shared experience of the patient's world. This narrative understanding is much more than a causal explanation of problem A or problem B that the patient might have. It does not simply abstract from the person's situation a categorical label that groups problems under a well-known abstract grid. Instead, narrative understanding tunes into the uniqueness of the individual and the unrepeatability of the person's experience and difficulties. Narrative understanding, in short, is a deep

appreciation of the person as a whole—what it feels like for this person, in this particular context, going through these particular problems.

By tuning in to the person's story, first and foremost, as a goal in and of itself, narrative psychiatrists, like their narrative medicine colleagues, create an environment in which the person can be known and his or her difficulties can be appreciated. This step can be one of the most healing acts a clinician can perform. By allowing the person to be known, the clinician creates a bond that reduces the patient's isolation because the patient immediately becomes less alone with the problem. Whatever else may happen in the clinical encounter, this sense of togetherness will reduce the person's suffering. The patient now has a comrade, a foxhole partner, someone who is in it with him or her. This partnership may not change the extent of the person's troubles, but it changes immensely the experience of the situation. Although clinicians are not the only ones who can perform this role, a good clinician can sometimes be one of the few people in the person's world who can truly understand. Clinicians have extensive experience with similar experiences of human suffering, as well as the advantage of working in clinical settings in which many of the blocks to discussing shameful or embarrassing problems often are removed.

In addition to following the lead of narrative medicine colleagues, narrative psychiatrists also follow the lead of contemporary colleagues in narrative psychology and psychotherapy. The history of narrative psychology and psychotherapy goes back to Sigmund Freud's early work at the inception of psychoanalysis. At that time, Freud lamented about how his case histories sounded more like narrative fictions than hard science. As Freud explained, "I have not always been a psychotherapist. Like other neuropathologists, I was trained to employ local diagnosis and electro-prognosis, and it still strikes me a strange that the case histories I write should read like short stories and that, as one might say, they lack the serious stamp of science."

Despite this awareness of the role of story in his clinical work, however, Freud never embraced a narrative theory of psychoanalysis, and throughout his career he remained attached to the natural science model he inherited from neurology.

Contemporary narrative psychology and psychotherapy's motivation for returning to the role of narrative comes partly from the broader turn to narrative in humanities and social science and partly from the history of psychotherapy since Freud. The last century of psychotherapy has been a century of strife, with one faction after another splitting off from psychoanalysis. Leading alternatives to psychoanalysis included behavioral, humanistic, family, cognitive, feminist, and interpersonal, just to name some. All of these splits are characterized by further splits within splits, which has fragmented the field of psychotherapy to the point that there are now more than 400 active approaches to psychotherapy. Narrative

approaches emerge at this particular moment as part of an important trend away from further fragmentation and toward psychotherapy integration. Narrative approaches are invaluable for psychotherapy integration because they provide a metatheoretical orientation from which to understand and practice psychotherapy. As Lynne Angus and John McLeod pointed out in *The Handbook of Narrative and Psychotherapy,* narrative approaches allow psychotherapists to understand the common factors within psychotherapy and to appreciate the diversity of psychotherapeutic models without attempting to close these differences down. Thus, in many ways, the contemporary emergence of narrative approaches brings the history of psychotherapy back full circle. More than 100 years since Freud introduced "talking therapy," psychotherapy has returned to its early appreciation of the similarity between case histories and narrative stories.

Similar to narrative psychotherapists, narrative psychiatrists recognize that the stories people tell about themselves not only describe their lives, they also shape their lives. The stories people use to frame and organize their lives are not simply incidental to their identity; they are the primary tools people have for configuring and reconfiguring their identity. Narrative psychiatrists recognize that this shaping of identity through narrative is a major tool for recovery. When people come for help in a therapeutic situation, they not only want to be understood, they also want help making changes in their lives. Narrative psychiatrists collaboratively encourage and foster these changes by helping coauthor new stories of purpose and thriving. As SuEllen Hamkins puts it in *The Art of Narrative Psychiatry,* a narrative approach helps patients discover "untold but inspiring stories of resiliency and skill in resisting mental health challenges while dismantling aspects of presenting narratives that fuel problems. It examines what the doctor's kit of psychiatry has to offer in light of the values and preferences of the person seeking consultation, authorizing the patient as the arbiter of what are helpful and what is not." Working in this way, narrative psychiatrists understand that collaboratively re-authoring stories people tell about themselves is a powerful way to make changes toward recovery goals.

WHAT IS NARRATIVE?

Both narrative medicine and narrative approaches to psychotherapy are deeply informed by narrative theory. But, how have writers and literary scholars tried to describe "narrative?" What are the important elements of narrative? How does narrative work in human life? Although these questions may sound straightforward, they are not so easy to answer. The literature on narrative theory is vast and it has many tributaries. Historian and narrative theorist Hadyen White has argued that to "raise the question of the nature of narrative is to invite reflection on the very nature of culture and, possibly, even on the very nature of humanity itself." For clinical work,

therefore, it is important select which aspects of narrative theory are most helpful.

Narrative physician and novelist Abraham Verghese provides a good place to start. Speaking to colleagues at the American Society of Internal Medicine, Verghese explains that "As physicians, most of us become involved in the stories of our patient's lives ... we become players in these stories. Our actions change the narrative trajectory ... and our patient's stories come to depend heavily on repetition of what we say." The inescapable implications for medicine are threefold:

> (1) *Story* helps us link and make sense of events in our lives; (2) we as physicians *create* stories as often as we record them ...; and (3) we are characters in [these] various stories, walking on and off the stage in tales that take place in our hospitals and clinics.

Verghese points to examples in which physicians reach narrative competence through years of attentive practice, but he finds a more direct route through combining medical practice with the "tools that the writer possesses," or what Verghese called the "storytelling craft." During Verghese's immersion into the craft of writing, he found that the pillars of narrative invariably involved the author's selection and organization of *metaphor, plot,* and *character*. Making selections from these narrative elements are "fundamental to good writing in the same way that internal medicine skills rest on understanding the mechanisms behind dyspnea, edema, polyuria, and other cardinal manifestations of disease."

Following Verghese's lead, narrative psychiatrists must at a minimum understand the narrative elements of metaphor, plot, character, and we can add to this list, list point of view. The following sections use narrative theory, particularly the richly synthetic efforts of philosopher Paul Ricoeur, to develop these narrative elements. Then, recent memoirs of mental illness are explored to see how people use narrative elements to tell the story of mental health concerns.

Metaphor

The stories people tell are rich with metaphor. Consider this example from clinical life: "My boss looks at me lately like I am a piece of garbage. In his eyes, I am used up, something to discard and throw away. My whole interactions with him leave me feeling like shit. I feel utterly worthless, and my life is meaningless." This snippet of a clinical story has at least four metaphors. The person compares herself to garbage, shit, monetary value, and a literary text. Commonsense would indicate that these kinds of metaphors have little importance. Metaphors, most people believe, are embellishments and ornamentations to the content of a story, but they do little else. From this perspective, the person behind the snippet feels bad and believes her boss does not value her. End of meaning. The literal paraphrase is all you need.

For influential literary theorists, however, such as I. A. Richards, a number of philosophers, particularly Max Black, Mary Hesse, Paul Ricoeur, and Mark Johnson, and increasingly for cognitive scientists like George Lakoff, metaphor has a much larger function than simple ornamentation. In their book *Metaphors We Live By,* Lakoff and Johnson put it most succinctly: "metaphor is not just a matter of language, that is, of mere words ... On the contrary, human thought processes are largely metaphorical." Lakoff and Johnson argue that metaphor is a pervasive aspect of human life, not just in the words that are used, but also in the very concepts: "Our ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature." By shaping concepts, metaphor structures the way in which the world is perceived, what is experienced, how people relate to other people, and the choices one makes.

Metaphor performs this function by allowing one to understand and experience one thing in terms of something else. Hesse uses the example, borrowed from Black, of "man is a wolf." Hesse explains that the metaphor works because it transfers ideas and associations from one term to the other. The metaphor selects, accentuates, and backgrounds aspects of two systems of ideas so that they come to be seen as similar: "Men are seen to be more like wolves after the wolf metaphor is used, and wolves seem to be more human." Hesse goes on, "Nature becomes more like a machine in the mechanical philosophy, and actual, concrete machines themselves are seen as stripped down to their essential qualities of mass in motion." In short, metaphor allows the two terms, or more precisely the two systems of terms, to interact and adapt to one another.

Understanding metaphor in this way connects to broader work in continental linguistic philosophy, and that work, as a whole, shifts standard ideas about truth and objectivity. It allows one to sidestep the usual binary traps between relativism (anything goes) and realism (there is only one correct or True way to describe the world). When the role of language and metaphor as a mediator between concepts and the world is understood, it no longer makes sense to think in these highly modernist either/or terms. Rather than using a rigid binary splitting between true and false, people can think instead in a postmodern language of semiotic realism and pluridimensional consequences. Or, as Paul Ricoeur puts it, people can let their notions of universal truth (or single objective truth) give way to ideas about "metaphorical truths."

The multiple metaphors that shape one's concepts and perception allow one to see and experience the world in a particular way and from particular perspectives. Ricoeur calls these particular ways "metaphorical truths" to emphasize that the metaphorical process does allow one to understand something real about the world. That "reality" is not absolute, and it is always in tension with other ways of understanding something equally real

about the world. However, these other ways of understanding are not available to unless people inhabit new or alternative metaphors for perception and description. For Ricoeur this means that the "seeing as" of metaphor generally also means a "being as" of lived experience.

Plot

Plot works like metaphor in that it also orders one's experience and provides form to narratives. Plot, or the process of emplotment, adds to metaphor two key dimensions: (1) it brings together what would otherwise be separate and heterogeneous elements, and (2) it organizes one's understanding and experience or time, or what could be called temporal perception.

The critical function of plot for narrative is that plot creates a narrative synthesis between multiple individual events and brings them together into a single story. It allows an intelligible connection to be made between them. Remarkably, plot can create a synthesis between events and elements that are surprisingly incongruous or heterogeneous—events that do not seem to fit together. In any one story, otherwise incommensurable elements like human desires, human rationality, human purpose, human agency, unconscious motives, unintended circumstances, unplanned accidents, other peoples' desires and rationality, their unconscious motives, material force, natural laws, divine intervention, discoveries, nightmares, historical chance, social prejudice, economic downturns, failures of communication, contemporary styles and fashions, the effects of collaboration and conflict, grassroots organizing, spiritual conversion, the weather, technological capacity, and cell phones (just to name some) can all be brought together through the process of emplotment. Without a plot, there is no way to make sense of such diversity of phenomena.

Plot also configures these multiple elements into a temporal order. This temporal order is of two sorts. First, each plot is composed of a discrete series of incidents, of theoretically infinite *nows*. Second, each plot takes these infinite nows, proceeding one after another in succession, and organizes them into a humanly manageable experience. As Ricoeur points out, without plot, clock time is only a series of nows. The aporia, or gap, between clock time and the experience of "human time" comes from a well-known paradox: "the future is not yet, the past is no longer, and the present does not remain." This paradox means that clock time cannot be directly experienced. Although humans have a sense of time, when they are asked to describe it, they are baffled.

Ricoeur finds a phenomenological solution to the paradox of human time in St. Augustine's "thesis of the threefold present." Augustine argued that present, past, and future can all exist simultaneously in the human mind through a coming together of a plural, or threefold, present: the present moment accompanied by *memory* and *expectation*. Memory in this

formula recalls the past, and expectation, in an analogous way, looks towards the future. The present, in other words, contains one's memory of the past and one's expectation of future events. For Ricoeur, the threefold present of time highlights a parallel between the experience of time and the structure of plot. Plot, like human time, is organized in a temporal order of beginning, middle, and end. Through plots, people configure their confused, unformed, and otherwise mute temporal experience.

Ricoeur sees this relation between time and plot as a two-way phenomenon: time becomes human time when it is organized by plot, and conversely, plot is meaningful because it portrays the features of temporal experience. Ricoeur admits that his conclusion is "undeniably circular," but he argues that this circularity is not in itself a problem: The circle of plot and time, he explains, is "not a vicious, but a healthy circle." Emplotment, from this perspective, can be seen as a kind of prosthetic device. It does not solve the riddle of time so much as it allows one to keep going without a solution. The riddle of human time, in other words, leaves humans with an inconclusive rumination to which the process of emplotment alone can respond: "Emplotment ... replies to the speculative aporia [of time] with a poetic making of something capable, certainly, of clarifying the aporia ..., but not of resolving it theoretically."

Character

In narrative theory, the concept of character connects directly to contemporary controversy surrounding the related and more basic concept of identity. The controversy around identity may be understood as a tension between essentialist and nonessentialist approaches. Essentialist notions of identity indicate that humans have a fixed personality, perhaps biologically stamped, that is authentically belongs to one's self and that is at the core of one's being. This "true self" or "core self" may be distorted or covered over, but it is nonetheless there for the discovery if people apply themselves patiently and persistently to the task. Nonessentialist critiques, however, have deconstructed this ideal of identity and its notion of an integral, originary, and unified self. This critique has been comprehensively advanced by influential philosophers like Roland Barthes, Jacques Derrida, and Michel Foucault and by psychoanalytic writers working in the tradition of Jacques Lacan politically inflected through cultural studies and feminism. As a group, these writers have deconstructed the essentialist ideal of identity and its similar notions of an integral, originary, and unified self.

One of the most productive ways to navigate the tension between essentialist and nonessentialist understandings of identity is to draw a comparison between identity (in life) and character (in fiction). Rather than adopting a linear logic that understands identity as a more fundamental concept to character, this approach uses a circular logic to

argue that people understand themselves in the same way they understand characters. Paul Ricoeur calls this approach to identity "narrative identity," and the basic idea has recently been developed into a school of psychology known as "narrative psychology." As Ricoeur puts it, "fiction, in particular narrative fiction, is an irreducible dimension of self-understanding ... fiction is only completed in life and life can be understood only through the stories that we tell about it."

Self-understanding, on this account, is an interpretive event, and narrative is the privileged form for this interpretation: "a life story [is] a fictional history or, if one prefers, a historical fiction, interweaving the historiographic style of biographies with the novelistic style of imaginary autobiographies." Literary theorist Peter Brooks sums up this idea beautifully in his book *Reading for the Plot: Design and Intention in Narrative:*

> Our lives are ceaselessly intertwined with narrative, with the stories that we tell and hear told, those we dream or imagine or would like to tell, all of which are reworked in that story of our own interrupted monologue. We live immersed in narrative, recounting and reassessing the meaning of our past actions, anticipating the outcome of our future projects, situating ourselves at the intersection of several stories not yet completed.

Narrative approaches to identity allow one to navigate the tension between essentialist and nonessentialist identities because narrative identity allows for a kind of continuity over time, a relative stability of self, without implying a substantial or essentialist core to this stability. People's interpretations of themselves use the cultural stories with which they are surrounded to tell a story of self that escapes the two poles of random change and absolute identity. In this way, a narrative identity is also a cultural identification. People's identifications may seem original, but they are narrated with the resources of history, language, and culture.

As cultural theorist Stuart Hall puts, identity is a "process of becoming rather than being: Not 'who we are' or 'where we came from,' so much as what we might become, how we have been represented and how that bears on how we might represent ourselves." Culturally narrated identity stories, Hall argues, involve "not the so-called return to roots, but a coming-to-terms with our 'routes.'" They involve a deep appreciation of what routes have been taken, the effects of those routes, and what new routes might be chosen ahead. This coming to terms with routes, although it involves a use of inessential cultural narratives, in no way undermines the personal or political efficacy and value of identity. Nor does it undermine a genuine sense of belonging that can come from a well-narrated self.

First-Person Narratives

Bringing this theory to psychiatry, the importance of understanding the elements of narrative can be highlighted by looking at first-person memoirs

of mental illness. As Hillary Clark, a scholar of mental illness memoirs, puts it: "One cannot feel well one day, numb and oppressed the next, suicidal after a few months have elapsed ... without seeking a narrative explanation —the cycle of bipolar disorder, for instance, or the Christian narrative of sin, repentance, and redemption—in order to make sense of it all, to trace a single self through all these changes." Clark finds that studying personal narratives are valuable because they "give voice to the ill, the traumatized, and the disabled," and they restore insights and wisdom they have achieved —wisdom that can be lost by medical research and clinical case studies that risk turning people into statistics and diagnostic labels.

Memoirs are also helpful because they show how models of mental illness work like the narrative element of metaphor in the process of creating illness stories. Clinical models may be understood as containing root metaphors developed over time through a research and practice community. In ordinary writing or conversation, people generally use metaphors to provide a quick glimpse: "This hotel is a palace." But, in a clinical model, research communities (e.g., biomedical, psychoanalytic, cognitive behavioral, family, social/political, art therapy, spiritual or pastoral counseling) develop metaphors into systematic models that frame their work and that explain the world through a metaphorical re-description. The metaphor, "Joe is a pig," re-describes Joe in a new way and allows for something new to be perceived about him. So too the disease model of mental illness—"mental illness is a disease"—re-describes mental difference and allows something new about it to be perceived. Philosophical psychiatrists Bill Fulford and Tony Colombo use these insights to develop the following definition of psychiatric models: "[Psychiatric] models ... are the conceptual frameworks, or sets of ideas, by which, in any given area, people structure and make sense of the world around them."

Memoirs reveal that the root metaphors which organize models of mental illness do not stay in the clinics; they seep into culture and become cultural resources that join with the other narrative elements for organizing self-stories. In times of trouble, people look through the metaphorical structures of mental models to perceive, select, and plot aspects of their lives that they believe to be important. These culturally located self-stories and the priorities within those stories combine with other cultural resources to scaffold narrative identity and provide a compass for living. They tell where people have been, where they are now, and they provide a trajectory into the future. Studying memoirs of mental illness, a wealth of first-person stories can be found that help one to see how people plot the models and metaphors of mental difference into their narrative identity and their stories of recovery.

Lesley Dormen, in her memoir "Planet No," emphasizes themes of loss and depression through a metaphor of unresolved grief common in

psychoanalytic models. Dormen explains that she initially spiraled into depression by two losses in her life: “An architect I’d loved for under a year left me for a woman he met in a supermarket parking lot. And my best friend, Tessa, a woman as central to my understanding of myself as my brown eyes and small feet, disappeared into marriage and withdrew from our friendship completely.” Dormen used psychotherapy to explore these losses and came to realize that they were compounded by earlier losses and traumas from her childhood. She lost her father after her parents divorced when she was 6 years old, she was sexually abused by her stepfather, and she had a painful emotional relationship with her mother. Dormen felt so isolated and rejected by these experiences that when her therapist asked her, “if [you] were a girl in a story, how would [you] see yourself?” Dormen answered: “Like an orphan.” Building on this insight, she eventually came to understand that her many ungrieved losses amplified one another, and these unresolved feelings were repeated in the present life. As an adult, she re-lived her past experiences of solitary isolation and failure as an unmarried woman in New York, feeling the pressure to be coupled like “normal and successful” people.

Psychoanalytic therapy and self-reflection helped Dormen to understand that her stepfather’s sexual abuse was the central “curve in [her] depression story. His touch had caused the spine of the story to veer off into a new direction, had reconfigured the terms of my loves and longings.” Most pernicious for her ongoing development, the abuse left her confused about her desires, and this confusion left her “frozen in a girlhood made curiously old.” The dense and painful emotional conflicts that grew up around the sexual abuse knotted together shame, illicit sexual pleasure, heartbreaking loss, pity, betrayal, guilt, duplicity, and an aching distance between her mother and herself. These emotional knots were so tight that she could not begin to open them or untie the feelings, but at the same time she could not forget them. She found herself constantly and repeatedly returning to them “with the vague uneasiness that was a familiar part of [her] internal landscape.”

Slowly and often painfully, therapy helped her move past these knots in her character that seemed to rope her into depression and despair. She was able to separate herself from this history: “Something happened—something that was separate from me. There was a me that existed before the something and a me that existed after. Before and after were separate, not the same.” No longer tied together in knots, “it was possible to stand outside those events, to observe them and draw a conclusion about them. It was possible not to just be them.” And, not only that, it became possible to move past the grief, to feel joy, and to make new connections: “I fell in love with New York, with the giddy green of the first spring after the worst of the depression lifted, with new friends, travel, and wonderful books. I looked around my apartment and saw that there was pleasure in things, in their

colors and shapes, and that one did not actually need a marriage certificate to buy good pots." Dormen's memoir suggests that she came to understand her troubles through the metaphors of psychoanalysis, the plot of psychotherapeutic change, and the narrative identity of someone who used psychotherapy for recovery and transformation.

In contrast to Dormen's psychoanalytically inflected story, William Styron's memoir, *Darkness Visible,* is paradigmatic of the disease model as narrative template. Styron describes himself as "laid low by the disease," a "major illness" of "horrible intensity," which came on him like a "brainstorm ... a veritable howling tempest in the brain, which is indeed what clinical depression resembles like nothing else." This connection between depression and bad weather perfectly captures the disease model of depression as lying outside the frame of human goals, desires, losses, and disappointments. Like the weather, depression comes from the material world of inhuman forces, from physical and chemical interactions. For a strong disease model advocate like Styron, it makes no more sense to give human meaning to depression than to tell an atheist that thunder is caused by God's anger toward us. Styron explains how he sees it: "I shall never learn what 'caused' my depression, as no one will ever learn about their own ... so complex are the intermingling factors of abnormal chemistry, behavior and genetics."

Styron's memoir was one of the first and most influential to adopt such a narrative frame. He relies on biological metaphors for depression that perfectly match the disease model in psychiatry as well as an avalanche of pharmaceutical marketing devoted to promoting antidepressant medications. Depression, Styron writes, "results from an aberration of biochemical process. [It] is chemically induced amid the neurotransmitters of the brain ... a depletion of the chemicals norepinephrine and serotonin." This "upheaval in the brain tissues" causes the mind "to feel aggrieved, stricken," and it creates "muddied thought processes" like the "distress of an organ in convulsion."

Styron is far from alone in his approach, and the disease metaphor runs through many other memoirs. For example, in a collection of Depression memoirs, *Unholy Ghost: Writers on Depression,* it shows up in several examples. Virginia Hefferman laments the "substandard physiology," which creates her depression. Chase Twitchell moves past the phrase, "chemical imbalance," to chime in with a more updated disease model explanation: "What happens in depression, for reasons that are still unknown is that the limbic-diencephalic system malfunctions." And similarly Russell Banks fears that his wife's "malfunctioning limbic diencephalic systems" has created a depression so severe it may result in her suicide. All of these memoirs show the metaphors of biology shaping the plots of recovery and the narrative identity of the people involved.

In contrast to biopsychiatry or psychoanalysis, David Budbill's memoir,

"The Uses of Depression: The Way Around is Through," provides a good example of alternative models and metaphors. He approaches painful periods of depression through the less common but still vital models of spirituality and creativity. Because of what he calls his "rebellious, contrary" personality, none of the mainstream approaches is suitable for him, and he decides to push past the mainstream to develop his "own way of dealing with and using depression" that he calls his "give in" method: "I discovered that the only way around my periods of depression is directly through them; in other words, the sooner I can resign myself to the Angel of Depression, the sooner she will be done with me and leave me alone."

Drawing from the narrative frames of Zen Buddhism and creative writing, Budbill discovers that if he lets go of his critique of depression and allows himself to be with the intensity of the feelings, the depressive episodes are not all negative. Indeed, depression can be experienced as generative (or positive) as well. Budbill finds that one of its most important positive dimensions is that it slows him down. Depression forces him to enter a "negative space" where he can neither work nor produce; instead, he tunes into a very different experience—"empty, open, quiet, passive, receptive, dark." These slothful, withdrawn periods function as a kind of dormancy where he can store up energy for times when the Angel of Depression lets him go. During such times, he can better hear his muse, and this "receptacle" period later emerges in his poetry. For Budbill, the metaphorical frames of spirituality and creative writing helped him cultivate a narrative identity of spiritual seeker and artist whose experiences show sensitivity rather than pathology. The metaphor of sensitivity can be seen to scaffold a recovery plot that involves tuning into his depression experience as a muse and a generator of wisdom rather than a clinical plot that works to remove or transform the experience.

Joshua Wolf Shenk gives a good sense in his memoir how a self-reflective awareness of the many models and metaphors of mental illness can be a tool of coping. Responding to his depressive experiences by researching the history of depression and the role of metaphor in shaping what is known about depression, Shenk concludes that "what we call 'depression,' like the mythical black bile, is a chimera ... It is cobbled together of so many different parts, causes, experiences, and affects as to render the word ineffectual and perhaps even noxious to a full, true narrative." Even so, the pain and suffering of depression is anything but a chimera, and Shenk understands the inevitable need to "abbreviate and simplify" in trying to understand it. He finds "no way around words like 'depression' and 'melancholy'... but it is one thing to use shorthand while straining against the limits of language. It is quite another to mistake such brevities for the face of suffering." For Shenk, simplified phrases like "biochemical malfunctions" and "biological brain diseases" favored by public health campaigns, well-intentioned clinicians, and pharmaceutical

companies can mystify as much as they inform. “When we funnel a sea of human experience into the linguistic equivalent of a laboratory beaker, when we discuss suffering in simple terms of broken or fixed, mad and sane, depressed and ‘treated successfully,’ we choke the long streams of breath needed to tell of a life in whole.”

This awareness allowed Shenk to consciously seek an array of different approaches to his depression. He gave medications a try (although found none that were helpful) and sought solace and guidance in several hundred afternoons of psychotherapy. He worked through the pain of his parents’ divorce when he was seven—“the slow leakage of affection and kindness from my parent’s marriage, the grim entrance of resentment, confusion, and anger,” and he overcame the automatic tendency from his family life that “forbade expressing these emotions.” All of this work went into the larger task of creating a life story that put his depression in perspective: “to find my story by living it, following moments of emotional clarity through life’s maze. I look for help in therapy, [but also] in relationships, and faith in its broadest sense—faith in the gardener, the faith of the lover, the faith of the writer. The faith that I can experience what is real about the world, that I can hurt plainly, love ravenously, feel purely, and be strong enough to go on.”

The story of the Icarus Project, a consumer and mad studies activist group, provides another instructive example of how people can bring conscious awareness of the diversity of narrative options into the process of recovery. The Icarus Project was started in 2002 by Sascha DuBrul and Ashley McNamara—two arts and humanities educated college students who were given diagnoses of “Bipolar Disorder” in their 20s and who found themselves struggling as much with their psychiatric treatment as with their emotional intensities. DuBrul and McNamara, similar to many others in consumer activist groups, felt that the core of the problem with the psychiatric perspectives they encountered was an overemphasis on scientific models of “disease” and “treatment.” DuBrul and McNamara did not respond against these models with an antipsychiatry perspective or dismiss psychiatric disease models outright, but they felt this way of telling the story of their difficulties missed something crucial. They agreed that their psychic differences could cause them tremendous difficulties similar to the way that medical illness can do, but they also found it a mistake to over emphasize a model of pathology and story of broken brains.

Instead of accepting or rejecting the prevailing biological narrative, DuBrul and McNamara set out to create a grass roots alternative community for healing and recovery that was open to the many stories of psychic differences. As McNamara explains, “We began The Icarus Project as a way of creating space for people to share their trajectories through this under-charted world of blackness and brilliance and the million shades of gray that the medical establishment has no idea how to describe.”

Throughout this process, Icarus Project organizing has been led by an insightful narrative organized around the story of the boy Icarus:

> Defining ourselves outside convention we see our condition as a *dangerous gift* to be cultivated and taken care of rather than [simply] a disease or disorder needing to be 'cured.' With this double edged blessing we have the ability to fly to places of great vision and creativity, but like the boy Icarus, we also have the potential to fly dangerously close to the sun—into realms of delusion and psychosis—and crash in a blaze of fire and confusion. At our heights we may find ourselves capable of creating music, art, words, and inventions which touch people's souls and change the course of history.

DuBrul and McNamara go on to acknowledge the risks in using the Icarus story and the possible perils of narrating their emotional difficulties as dangerous gifts. However, in spite of these risks, they also recognize how metaphors of mental and biological pathology and metaphors of creative, political, and spiritual sensitivity can be as close as a turn of the kaleidoscope. They recognize how stories centered on pathology can turn into stories of creativity and that this narrative flexibility can create "tools of inspiration and hope in this repressed and damaged society." As they put it, "We understand that we are members of a group that has been misunderstood and persecuted throughout history, but has also been responsible for some of its most brilliant creations. And we are proud."

The Icarus Project is similar to Shenk in that it consciously creates integrated narratives of psychic troubles that are open to pathological metaphors as well as creative, spiritual, and political metaphors. The difference is that Shenk is working on his story alone, and the Icarus project uses a mutual aid community. The goals for both, as the Icarus project puts it, are to recognize the dangers and pitfalls of mental difficulties, to celebrate and tap into the "potential that lies between brilliance and madness," to put in the necessary effort to stay balanced and grounded, and to plot recovery stories that lead to full and meaningful lives (which might include stories of psychiatric medications, psychotherapy, artistic practice, activism, or spirituality). By foregrounding the narrative process of recovery, the Icarus Project aims to separate mental health issues "from the reductionist framework offered by the current mental health establishment" and in doing so, learn "from each others' mistakes and victories, stories and art, and create a new culture and language that resonates with our actual experiences of 'disorder.'"

NARRATIVE PSYCHIATRY

The deep implication of narrative theory and memoirs of mental illness for psychiatric practice is that there are many ways to tell the story of mental difference and suffering—not just one right way and many other wrong ways. To be helpful for people with the tremendous diversity of values and life-style preferences seen in memoirs, psychiatrists cannot be limited to one or two storytelling options. Psychiatrist must be comfortable using a

range of metaphors and models to help people plot their recovery stories and narrative identities. To see how this works in practice, it is helpful develop a character from literary fiction as a "case example" for a narrative psychiatry approach.

Narrative psychiatrists find literary fiction an invaluable resource for "case examples" because literary examples, as a rule, provide much richer contextual data than clinical case studies. All too often clinical case studies tell a story that has been tightly organized along a particular narrative frame. Literary examples do this as well, but the framing is generally much less heavy handed and much more open to alternative frames and perspectives. That said, literary examples usually require additional imaginative work to make them suitable as clinical case histories.

To see how this works, consider Marina, a character in Claire Messud's 2007 novel *Emperor's Children*. Marina is a single woman, in her early 30s, lives with her parents, and is unemployed. She struggles throughout the novel with symptoms of depression and anxiety. To get to know Marina's story, start with a scene where she and her college friend Danielle are meeting for dinner. The dinner is at Marina's house, but Danielle arrives first, which leaves her spending several minutes alone with Marina's parents. Being alone with Marina's parents makes Danielle feel awkward and uncomfortable. Here's the passage:

> Danielle ... felt like a teenager, as she used to feel in the kitchen of her parents' house in Columbus, before the divorce of course, and she was suddenly, powerfully, aware of the profound oddity of Marina's present life. A life arrested at, or at least returned to, childhood. Danielle couldn't imagine eating nightly with her parents, not only because they now lived in different states and didn't speak to each other, but because she was entering the fourth decade of life and hadn't been through the wearying rigmarole of family life for anything more than an almost supportable day since she was seventeen and had gone off to college.

In this passage, Danielle uses the metaphor of childhood to organize her understanding of Marina's life. Through the childhood metaphor Danielle perceives Marina's adult life as "arrested, or at least returned to, childhood." If Danielle were to tell the story of Marina's life using this childhood metaphor, she might very well select events from Marina's past that led up to this moment. Perhaps she would recall a difficulty Marina had in committing to relationships, or a series of immature reactions to authority figures, or a perhaps several examples of her being dreamy and irresponsible. These diverse elements would come together in a plot that leads to the present moment of Marina as trapped in childhood. The story would not only make sense of the past and the present and provide a narrative identity for Marina, it would also provide clear direction for the future. In this narrative frame, Marina must "grow up." She must leave her parents, home, get a job, get married, and have children (or at least make steps in these directions).

This example highlights the importance of "point of view." Narratives

tell a story from a particular perspective. Even if an author uses the omniscient point of view, that is still not a story from nowhere; it remains a story told by an author who has a particular take on the events that he or she selects and frames. In the selection above from *Emperor's Children,* readers learn about Marina's life from Danielle's point of view. This passage does not tell the reader whether Marina would see herself as stuck in childhood; it only tells the reader that Danielle sees her this way.

As it turns out, a little further on in the story, readers learn that Marina does feel similarly about herself. Indeed, the perspective of herself as childish and immature has left her deeply unhappy and anxious about her life and even about her worth as a person. Not only this, but as the novel proceeds, readers see Marina's deep discontent with her "childish" self-perception leading her into a poorly thought out romance and decision to marry an older man, a magazine editor, after they had been going out for a few weeks. In the process of this marriage, Marina gains in quick succession not only a husband but also job (as she joins the editorial staff of her new husband's magazine). This rapid marriage and entrance into the job market effectively disrupts the childhood metaphor (Marina has now moved into "adulthood"), but, unfortunately, Marina's judgments turn out to have been potentially premature, as the magazine quickly fails and the relationship falters. By the end of the novel, it is not known whether Marina's marriage will work, but Messud leaves her readers feeling uncertain and doubtful.

To use this story as a literary "case example," the additional narrative work involves imagining Marina coming to see a psychiatrist at or near a moment of crisis. A common plot structure in narrative fiction involves what theorists call the Freitag triangle. In this plot structure, the story and its characters move from increasing tension and conflict to climactic crisis and then finally to resolution. If it is a novel or a complicated short story, the plot is often composed of multiple Freitag triangles that may or may not come together into an overarching Freitag triangle. In using literary material as case examples, it is helpful to imagine that a character in the story comes to see an imaginary clinician just at or near this point of climactic crisis. For example, imagine that Marina came to see a narrative psychiatrist for help when she was at the height of her crisis. This would be when she was in the middle of her most intense anxiety and feelings of inadequacy, and before these feelings were resolved (at least temporarily) by marrying the magazine editor and getting a job.

How would the narrative psychiatrist work with Marina in this situation? Like with narrative medicine and other person-centered approaches, the narrative psychiatrist's first and in many ways foremost task in helping Marina will be in getting to know her and developing a rich appreciation of her story. This empathic connection helps Marina become much less alone with her problems. She now has a nonjudgmental partner

with whom to share them and to relate them. Psychiatrist Robert Coles captured this deeply person-centered dimension of narrative psychiatry in his book *The Mind's Fate: A Psychiatrists Looks at His Profession*. Coles sees the person-centered core of narrative psychiatry as the "*ultimate mission* of those who are doctors of the mind." To be a psychiatrist is to be an "attentive listener first" and to be able to make sense of the "inside cry for a knowing assistance that is at the heart of a person's decision to visit a physician." This kind of attentive listening remains central to psychiatry regardless of evolving science and research into mind and brain because:

> Human beings are creatures of awareness, of language, of long and articulated memory—and also, creatures whose special skills and triumphs live side by side with a special capacity for tragedy. Losses, vulnerabilities, weak spots and soft spots, one's ultimate destiny of disappearance from this planet are known to humans, but they are not to all other forms of life on this earth. Through words, whether spoken to others or silently put to one's self, humans constantly struggle to make sense of things, even as they try (if they can) to prepare themselves for the worst: hunker or huddle down, reach out, whistle in the dark, while trembling mightily within themselves. Not rarely, at such moments, people have learned in this century to seek the counsel of psychiatrists, men and women presumed to understand one's many possible predicaments, one's wild or loony sides, one's lapses into incoherence or rage, into lusts and times of darkness that threaten one's reliability, one's stability, one's very own life (or the lives' of others).

For Coles, psychiatrists must understand that to be sought out at these moments, to be asked to "attend such a cry" is an honor, a privilege, and the "ultimate bedrock" of psychiatry's professional competence.

This empathic core, however, will only be part of the interaction between Marina and the narrative psychiatrist. In addition, Marina will likely wish that the psychiatrist recommend interventions and treatment options to help her move beyond her current crisis and conflict. Narrative psychiatrists are closely guided in this process of care planning by the bioethical principles of *respect for autonomy* and *informed consent*. Bioethicists Tom Beachamp and James Childress explain that respect for autonomy acknowledges the moral right of every individual to choose and follow his or her own plan of life and actions. Informed consent is the practical application of the respect for autonomy. Following Beauchamp and Childress, narrative psychiatrists agree with Johnson et al.'s description of informed consent in their *Clinical Ethics:*

> Informed consent constitutes a central feature of an encounter between physicians and patients, characterized by mutual participation, good communication, mutual respect, and shared decision making. Informed consent requires a dialogue between physician and patient leading to agreement about the course of medical care. After initial consent to treatment has occurred, and ongoing dialogue between patient and physician, concerning the patient's continuing medical needs, reinforces the original consent. A properly negotiated informed consent benefits both the physician and the patient—a therapeutic alliance is forged in which the physician's work is facilitated because the patient has realistic expectations about results of the treatment, is prepared for possible complications, and is more likely to be a willing collaborator in the treatment.

Following this approach to informed consent, the narrative psychiatrist

initiates a collaborative process of selecting strategies for change by giving a background on the array of treatment options available and explaining the benefits and risks of these options. Narrative psychiatry is not unique in this way, however narrative psychiatry does share bioethicists concern that despite a vast literature in bioethics that physicians in general and psychiatrists in particular are often deficient in observing the practice and spirit of informed consent. Narrative psychiatry, by contrast, puts this form of ethical alliance building and collaborative planning at the center of care. Narrative psychiatrist SuEllen Hamkins describes this stage of treatment as exploring with the patient the "doctor's tool kit of psychiatry." As she puts it, "when we have a collaborative therapeutic stance, we can look side by side without our patients at the wealth of treatment options that might be helpful and weigh the pros and cons together."

In the case of Marina, because her problems are relatively mild at this point, the narrative psychiatrist would generally suggest psychotherapy. However, which psychotherapy should Marina choose? The marketplace consists of a tremendous variety of options. The collection *Current Psychotherapies* features 14 different approaches and *Essential Psychotherapies* focuses on 12 different options. Both of these texts cover psychoanalysis, cognitive therapies, and humanistic therapies (among others), but at the same time they both leave out a myriad of additional options, such as feminist therapies, expressive arts therapies, and multicultural psychotherapies. In addition, both texts do not mention alternative approaches to psychic troubles, such as acupuncture, yoga, peer support, and mad pride activism. Although these last options are not psychotherapy by definition, they do represent viable approaches to psychic difficulties that Marina might find helpful. Taking this tremendous variability of psychotherapies and other treatment options into account leaves the narrative psychiatrist with two problems. How to make sense of this diversity and how to present these options to Marina in a way consistent with autonomy and informed consent?

Fortunately, one of the most helpful aspects of narrative theory for psychiatry is that it provides an overarching, or metatheoretical, rationale for understanding how these many psychotherapies work. From a narrative perspective, all therapies involve a process of storytelling and story retelling. No matter which style of psychotherapy one uses, the process of therapy involves an initial presentation of problems that the client is unable to resolve. The client and therapist work together to bring additional perspectives to these problems, allowing the client to understand them in a new way. These additional perspectives vary greatly depending on which style and therapeutic point of view is used. It matters whether the therapy is psychodynamic, cognitive, humanistic, feminist, spiritual, or expressive. But from the vantage point of narrative theory, what these different approaches all have in common is that they rework, or "re-author," the

patient's initial story into a new story. This new story allows new degrees of flexibility for understanding the past and provides new strategies for moving into the future.

Having a metatheoretical narrative understanding of the variety of psychotherapy options helps make sense of perhaps the most important empirical psychiatric finding to date—an array of alternative approaches to psychotherapy are all *equally effective*. This finding is known lovingly in the therapeutic research community as the "Dodo effect." The appellation comes from a line in *Alice in Wonderland:* "At last, Dodo said, everyone has won and all must have prizes." Although this finding is controversial and has been only partially integrated into clinical practice, empirical studies confirm with remarkable consistency that the positive effects of therapy (of which there are many) are not due to the specific interventions of the therapist. The benefits of therapy come instead from "common factors" of the therapeutic setting. These studies suggest that the process of setting up a strong therapeutic alliance with a quality therapist is much more important than the content of the specific models and theories from which the therapist works.

The Dodo effect suggests that a variety of theoretical models can be used to understand, cope with, and ameliorate painful emotional states. The Dodo effect does not mean that any old interpretation will do. Therapy does not work if the therapist says things completely out of context, such as "your sadness is because the cow jumped over the moon." On the contrary, for psychotherapy to be effective and for the alliance to be strong, the therapist and the client must have a sense of belief and confidence in the interpretive frame being used. In the words of Saul Rosenzweig, who first coined the term Dodo effect, "Whether the therapist talks in terms of psychoanalysis or Christian Science is ... relatively unimportant as compared with the formal consistency with which the doctrine employed is adhered to, for by virtue of this consistency the patient receives a schema for achieving some sort and degree of personality organization."

When the narrative psychiatrist applies these insights to Marina's situation, it suggests that an array of psychotherapy approaches might be helpful for Marina. Any number of different interpretations of Marina's situation could be developed and worked through in a therapeutic relationship with a well-meaning quality practitioner. If that were to happen, there is a very good chance that an array of different insights and narrative structures could effectively help Marina understand her anxiety and feelings of inadequacy. If the insights and narrative frames were not imposed on Marina but were ones in which she participated in creating and to whose meaning she felt strongly connected, the therapy would "work" through encouraging modifications in Marina's assumptive worlds. These modifications would transform the meanings of her painful feelings to more favorable ones and free her up for alternative forms of coping.

A narrative psychiatrist working with Marina would have all of this in the back of his or her mind as he or she began the process of making psychotherapeutic recommendations. Following the basic ethical principles of autonomy and informed consent even in the fine-grained choices between types of psychotherapy options, the psychiatrist would give Marina background on the array of psychotherapies so that Marina could make informed choices about which therapy to choose. Most narrative psychiatrists would explain to Marina that he or she practices a narrative approach that collaboratively selects options and resources from a variety of treatment approaches as they seem appropriate for the therapeutic process. This means that narrative psychiatry tends to use an integrative form of psychotherapy that works through collaborate eclecticism. The strategic and collaborative frame for selecting from a range of different therapeutic approaches is guided by narrative goal of creating preferred identities of meaning and thriving. Eclectic openness is consistent with narrative psychiatry because it appreciates that, not only are a variety of therapeutic models valuable for people, but also a variety of therapeutic insights and strategies can be woven together to form a hybrid narrative of a person's past, present, and future that can be helpful. These insights may come from psychoanalysis or cognitive–behavioral, feminist, or creative arts perspectives to name a few. Eclectic therapists are not constrained by a loyalty to one point of view over another and therefore can borrow and mix from a variety of schools.

If this form of therapy did not sound helpful to Marina, or if she had preferences for a more specific form of treatment beyond what the narrative psychiatrist felt comfortable, then the psychiatrist would work with her to find a suitable referral. If there was a match between what Marina would like to try and the narrative psychiatrist's collaborative eclectic approach, then they can continue working together. It is important to point out that although there are overlaps, a narrative approach to the many different psychotherapies is different from a specific narrative psychotherapy approach. A "branded" narrative psychotherapy approach would not have particular pride of place in the therapeutic process. Narrative psychotherapy as a particular form of psychotherapy, in other words, would be one approach among others, which Marina and the narrative psychiatrist might eclectically choose from in developing her preferred recovery stories.

However, for the sake of this overview of narrative psychiatry, it is useful to consider what narrative psychotherapy approaches might look like for Marina. Narrative psychotherapy is a form of psychotherapy that is in some ways similar to psychodynamic, or insight-oriented psychotherapy, and in other ways similar to humanistic and existential psychotherapy. The key idea is that clients come to therapy with, what narrative psychotherapists Michael White and David Epston call, "problem saturated

stories." The process of therapy deconstructs this presenting story by carefully and lovingly pulling it apart. In this example, Marina and the narrative psychiatrist would look together for gaps and breaks in her story, for inconsistencies, for left-out information, for overly emphasized events, and for metaphors and linguistic distinctions that ground the assumptions of the story. Finding these kinds of gaps and inessential authorial choices would help to establish that the Marina's presenting story is not an absolute necessity and that other stories could be told. Marina and the therapist would then work together to find a new story that would help Marina re-author her past and present and expand her future options in desirable ways.

This process is similar to the process of other psychotherapy approaches, except that it would be less likely than most to rely on therapeutic jargon and pretold story frames—such as are more common in psychoanalytic therapy, cognitive therapy, interpersonal therapy, or feminist therapy. Rather than using a therapeutic jargon, narrative psychotherapy would try to stay with the language of everyday meaning and folk psychology. In addition, narrative psychotherapy, more than most other therapy options, tends to be very conscious of cultural power dynamics and the way in which people's problems often have to do with internalized hierarchies that they have adopted from mainstream culture.

In Marina's case, the metaphor of "childhood" provides a good example of the pain cultural power dynamics can play. Marina lives in an elite and still deeply patriarchal culture, where "autonomy" and "independence" are prioritized over "connection" and "interdependence." These cultural norms are loaded into the meaning of the "childhood" metaphor. If Marina lived in a different culture or if she thought through these values preferences in the context of therapy, she might feel very differently about her current life and what it means to live at home. In a narrative therapy, Marina and the therapist would look for aspects of Marina's life—what narrative therapists call "unique outcomes"—that go against the grain of the childhood metaphor. They would look at the many ways in which Marina's life is very different from the life of a child, and they would deconstruct the wisdom of a cultural insistence on norms of autonomy and independence. Are these values preferences absolute, do they necessarily lead to a contented life, and are there examples in which people have followed these goals and ended up discontented and demoralized? Indeed, as it turns out in the novel, Marina's friend Danielle—the very one who could not imagine being in Marina's situation—is far from happy in following a path of relentless independence and "maturity." By working through this kind of narrative re-authoring, Marina could use narrative psychotherapy to reach a new level of peace with her current life and a new degree of flexibility as she makes plans for the future.

In Marina's "case," as it has been imagined it thus far, the main

treatment strategy has been psychotherapy. But it is important to point out that the process of using narrative as a guiding frame for treatment planning also includes more intensive psychiatric treatments such as medications, hospitalizations, and involuntary commitments. To enter into these additional aspects of narrative psychiatry, make a further imaginative adjustment to Marina's case history. Imagine that Marina did not enter therapy in the early stages of her crisis, but that she continued to suffer alone with her problems. Imagine that without a resolution for her crisis, her sense of inadequacy worsened to the point at which Marina became not only anxious, but also deeply sad and tearful, with intense feelings of worthlessness and a profound sense of hopelessness that she could ever move forward with her life. Imagine that she loses interest in the things that she had previously enjoyed and starts spending most of her time in bed watching television. These feelings go on for weeks, and it gets to the point at which Marina is unable to sleep, her appetite becomes erratic, and she even begins to fantasize about suicide.

If Marina were to see the narrative psychiatrist at this moment of worsened crisis, she would now have a presenting problem that fits the *Diagnostic and Statistical Manual of Mental Disorders,* fifth edition (DSM-5), diagnosis of Major Depression, severe without psychosis. With this diagnosis, Marina would be considered by many to have a "biological depression," which is thought to be a kind of "disease" not unlike diabetes or high blood pressure. How would the narrative psychiatrist respond in this situation? Would it differ significantly than in the previous situation? The answer is yes and no. As before, the first task for the narrative psychiatrist would be to get to know Marina's story in all of its unique complexity. The main difference in this new situation would be the added complication of Marina's suicidal ideas and the need to make an assessment of her danger to herself. If Marina met standard of care criteria for hospitalization, then the narrative psychiatrist would collaboratively work toward that goal. If the standard of care for hospitalization was met, but the patient refused and collaborative treatment planning broke down, the narrative psychiatrist would work toward involuntary commitment. This would unfortunately interrupt the collaborative process of meaning making as the narrative psychiatrist would have to act unilaterally. But the larger narrative process of the treatment would not stop. During the hospitalization and afterwards Marina and psychiatrist would continue to process the event. The eventual narrative meaning of the involuntary hospitalization would be open to ongoing storytelling. It may turn out that Marina comes to agree with that "danger to self" plotting of the story, or it may turn out that the psychiatrists comes to re-story the involuntary hospitalization as an overreaction and even a mistake.

If hospital care was not an issue, the outpatient process of assessment and treatment planning would be similar to before except for the question

of medication. After getting to know Marina, or even while this process was proceeding, the narrative psychiatrist would explain that in view of the severity of her depressive symptoms many, if not most, psychiatrists would recommend medication treatment along with psychotherapy. The narrative psychiatrist would explain the risks and benefits of the medication options along with a basic review of the outcome studies for the medicine's effectiveness.

If Marina opted for medication and it was indicated in the standard of care, the narrative psychiatrist would be comfortable with this choice. Narrative psychiatry, in other words, does not change its basic storytelling approach when it comes to biological explanations, disease models, or psychopharmacologic interventions for mental states. Narrative psychiatry does not resemble antipsychiatry approaches that understand disease models of mental illness as a myth. However, narrative psychiatry also does not essentialize biological or disease models of explanation and treatment as somehow more real and substantial than other models. For the narrative psychiatrist, biological models are not an exception to a general understanding of the ways in which metaphor and plot work to shape narrative identity because narrative psychiatry conceptualizes biological models not as myths or mirrors of truth, but as metaphors. Narrative psychiatry stands on the shoulders of work in philosophy of science and science studies to draw out the similarities between scientific models and metaphors more broadly. It recognizes that scientific models also create a kind of "seeing as." First-person narratives mental illness show that when people inhabit scientific frames they come to experience "seeing as" in a deeply naturalized way, to the point at which "seeing as" becomes "being as" and people experience themselves as having some form of "broken brain." However, this process is not different from other uses of metaphor to structure experience—it is just another example of the role of metaphor and linguistic structures in the formation of narrative identifications.

Biological models do pose a problem for informed consent, however, because of the degree the biological point of view is overhyped in the media and in the medical literature. Sociologist Nicolas Rose uses the term "neurochemical self" and anthropologist Emily Martin uses the phrase "pharmaceutical person" to describe the dramatic rise in neurochemically mediated self-identity. To become a neurochemical self one does not have to become a fundamentally different kind of human. One only has to internalize the metaphorical structures of neuroscience and the practices of psychopharmacology and use these metaphors and practices to organize one's sense of self. Consider the example of Gary, a character from Jonathan Franzen's book *The Corrections:*

> Although in general Gary applauded the modern trend toward individual self-management of retirement funds and long-distance calling plans and private-schooling options, he was less than thrilled to be given responsibility for his own personal brain chemistry, especially when certain

people in his life, notably his father, refused to take any such responsibility. But Gary was nothing if not conscientious. As he entered the darkroom, he estimated that his levels of Neurofactor 3 (i.e., serotonin: A very, very important factor) were posting seven-day or even thirty-day highs, that his Factor 2 and Factor 7 levels were likewise outperforming expectation, and that his Factor 1 had rebounded from an early-morning slump related to the glass of Armagnac he'd drunk at bedtime. He had a spring in his step, and agreeable awareness of his above-average height and his late summer suntan. His resentment of his wife, Caroline, was moderate and well contained. Declines led key advances in key indices of paranoia (e.g., his persistent suspicion that Caroline and his two older sons were mocking him), and his seasonally adjusted assessment of life's futility and brevity was consistent with the overall robustness of his mental economy. He was not the least clinically depressed.

Whatever else one might want to say about Gary, it is clear that the metaphors and plots of neuroscience have become a core part of his self-understanding. For another example, this time from nonfiction, consider this quote from neurophilosopher Patricia Churchland. The context is that she is describing her feelings to Paul Churchland (her neurophilosopher husband) after coming home from a particularly frustrating faculty meeting:

Paul, don't speak to me, my serotonin levels have hit bottom, my brain is awash in glucocorticoids, my blood vessels are full of adrenaline, and if it weren't for my endogenous opiates I'd have driven the car into a tree on the way home. My dopamine levels need lifting. Pour me a Chardonnay, and I'll be down in a minute.

Neurochemical selves, as these examples make clear, are selves like other selves. The only difference is that they have different metaphors, plots, and narrative identifications than nonneurochemical selves. Narrative psychiatrists have an ambivalent relationship with this rapidly emerging phenomenon. On one hand, narrative psychiatrists are open to this form of selfhood and they respect it as much as they would other forms of self. William Styron's memoir and the many similar ones show that people are clearly able to integrate the biological model and neurochemical self-identity into their lives in a helpful way. On the other hand, narrative psychiatrists are also aware of the deeply problematic power dynamics of this emerging form of self and that there are many people who prefer to avoid biological stories as much as possible.

Both Nicolas Rose and Emily Martin point to contemporary pharmaceutical practices as a major determinant of this new form of self-making. Since 1980, the pharmaceuticals have transformed themselves from research-driven companies to companies that operate more like Procter & Gamble, the maker of Tide. The pharmaceutical business plan of expanding markets through chronic, life style, conditions (such as arthritis, asthma, diabetes, mental illness) means that the industry has moved away from a focus on product per se and toward the creation of climate of opinion and desire surrounding the product. Jan Leshley, the former CEO of Squibb pharmaceuticals, explains: "We realized we are an information company more than we are a pill company." For Leschley, the modern

transnational drug company does not sell its products, the pills, as much as the information that surrounds their pills. Today's pharmaceutical industry has become an information industry, an industry of cultural shaping, more than a pill industry, and this cultural shaping very consciously includes the internal culture of psychiatric research, education, and practice. These powerful marketing strategies and tremendous promotion budgets that go with them means that pharmaceutically inspired neurochemical metaphors for self-making can often drown out other options.

All of this becomes grist for the mill, as psychoanalysts like to say, in the clinical setting. Narrative psychiatrists navigate their complicated relationship to pharmaceutical self-making by sharing with their clients their ambivalence during their discussion of these options. They let their clients know that there is considerable hype around these kinds of interventions, and they make a point to discuss the reality of side effect risks along with the ambiguity of most outcome studies for these medications. However, they also make the point that, for any one person, there is no way to know—except through a treatment trial—how much he or she might benefit or risk from a medication. Should the patient wish to undertake such a trial in the context of informed decision making, the narrative psychiatrist is happy to offer whatever assistance is possible. For narrative psychiatrists, these kinds of psychopharmacologic frames and strategies become part of an array of possible tools than can be woven together into an eclectic narrative of treatment, and the baggage of "big pharma" promotion dynamics does not necessarily rule out their usefulness for a particular person in a particular situation.

FUTURE DIRECTIONS

In summary, the time is right for a narrative turn in psychiatry. Recent work in narrative medicine, narrative psychology and psychotherapy, and narrative theory has opened the door for the development of narrative psychiatry. Narrative is particularly valuable for psychiatry at this moment in history because it helps self-correct from an overreliance on biological models without falling into a morass of anti-psychiatry and without throwing away the bioscience baby with the bathwater. Plus, the narrative ferment in the intellectual life surrounding psychiatry—the arts, humanities, social science, medicine, and psychotherapy—means there has never been better resources available for cultivating narrative psychiatry. Finally, psychiatry needs narrative to tell its own story because key identity questions for psychiatry as a field and for individual psychiatrists are narrative questions. Everyone involved in psychiatry must ask themselves: where have they been, where are they now, and where will they go in the future? All of these questions are about narrative identity. For the field, these questions are about how to prioritize future research, education, and practice models. For individual psychiatrists, these questions are about

determining who he or she wants to be and how he or she wants to practice. These kind narrative identity questions for the field and for individual psychiatrists are storytelling questions; they cannot be answered by science alone without the help of narrative meaning making.

None of this means the end to neuroscience, the biological model, psychiatric medications, or emergent neuroplastic training practices such as mindfulness-based stress reduction. Just the opposite. The biological model is now, and will remain, a critical tool for understanding and intervening in psychiatric troubles. This will only become truer as biology and neuroscience moves into a brave new world that maps out physical markers for mental states and mental processes. Taking a narrative turn in psychiatry does not mean abandoning the biological model, or any other model. A narrative turn says all models and metaphors are valuable if they are used self-reflexively in dialogue with patients and in the service of recovery. Ultimately, it is the patient's choice which models and metaphors fit best with their goals and desires. The task of psychiatrists is to help navigate the process and the options. This kind of narrative self-reflexivity helps psychiatrists remember that their allegiance is to their patients—not their models and metaphors. Recovery is the goal, not loyalty to particular narrative templates.

Narrative theory also does not take away from empirical research. Indeed, empirical studies of neural mapping and neuroplasticity are rapidly articulating a neurological basis for narrative theory itself. The work of neuroscientists Michael Gazzaniga, Joseph LeDoux, Antonio Damasio, V.S. Ramachandran, Jeffery Gray, and Gerald Edelman support the emerging consensus in neuroscience that self-understanding takes a narrative form. Philosopher Leslie Paul Theile summarizes the biological research that understands humans as storytelling animals this way: "Neural mapping is best understood as a narrative accounting of lived experience. It constitutes a silent, synaptic story. This narrative accounting is an 'obsession' for humans in the sense that, phylogenetically and ontogenetically, it constitutes our being. From the evolutionary perspective of the species and the developmental perspective of the individual organism, brain mapping makes us who we are. Synaptic storylines inventory the individual's life, serve as its engine of development, and foster a sense of self." In other words, at the neurological level, everyone can understood as fabricated characters. The brain constructs a model of the external world though synaptic maps and it constructs a model of the self as actor in the world. This emergent convergence of neuroscience and narrative theory can be invaluable in helping psychiatry move beyond deeply problematic mind/brain and science/humanities binaries. After a century of psychiatric swings between mind-centered and brain-centered versions of itself, narrative provides an opportunity for psychiatry to be true to both its biological dimensions and its human meaning-making dimensions.

Empirical research in the form of qualitative studies of recovery is also adding to an understanding of narrative processes in psychiatry. In one such study, Damien Ridge and Sue Ziebland interviewed 38 men and women who had recovered from depression to discover the strategies they used to revitalize their lives following depression. Ridge and Ziebland found that the people they interviewed did not use any single approach. Instead, they selected a variety of "narrative tools," such as talking therapy, medication, yoga, and complimentary therapies to regain mastery of their situation. Ridge explains: a "key finding…was that recovery tools considered effective by patients go hand in hand with telling a good story about recovering from depression." The very process of assembling a range of tools for recovery helped create the stories that were effective. The people interviewed found that therapists were most helpful if they could function not as experts but as "recovery allies." The most helpful health professionals were those who were empathic, good listeners, provided hope and encouragement, and who had a positive belief in the person going through the depression. A recovery ally of this type does not function in a "doctor knows best" capacity, but is someone who helps the person select the narratives and the recovery tools that feel most true to them.

Thus, when psychiatrists take a narrative turn, they do not throw out their other skills, knowledge, or research methods. The shift to narrative is, as much as anything else, an attitude shift and an opening out to additional sources of information. It starts by bringing to the foreground that the clinical encounter is a human encounter, and it follows by opening out to colleagues in the humanities, interpretive social sciences, and the arts to help to better understand this human encounter. In addition, narrative psychiatry joins with other contemporary efforts in psychiatry—such as in bioethics and the recovery movement—to make clinical encounters much more client focused and collaborative. Narrative psychiatry, at its core, recognizes that there are many ways to tell the story of one's life. The choice among these different options is a key way in which people create their identity. These choices should not be reduced to expert choices or scientific choices because they are always also personal and ethical choices. In the end, they are choices about what kind of life one wants to live.

For psychiatric practitioners, taking a narrative turn means joining with colleagues in narrative medicine and narrative psychotherapy to develop their own narrative competency. Narrative psychiatrists must understand the basics of narrative theory, and they must learn to read widely in a range of different contexts. They must come to appreciate the many stories of biopsychiatry, psychoanalysis, cognitive therapy, interpersonal therapy, family therapy, humanistic approaches, cross-cultural approaches, feminist approaches, mad pride and disability activist approaches, postmodern approaches, and spiritual approaches, just to name a few. Furthermore, they must come to understand the value of biography, memoir, and fiction

for developing a repertoire of narrative frames and options. In the end, narrative competency in psychiatry means a tremendous familiarity with the many possible stories of psychic pain and psychic difference. The more stories clinicians know, the more likely they are to help their clients to find a narrative frame that works for them.

For clients and potential service users, a narrative understanding means that there is a range of possible therapists and healing solutions that might be helpful. An approach that is right for one person may not be right for another. There must be a fit between the person and the approach, and people should feel empowered to take seriously their intuitions and feelings. If the person getting help does not feel this fit, he or she is likely right. There may well be another approach that would work better with the person's proclivities. Like everything else, however, judgment is critical. Therapeutic experiences of all kinds can be frustrating, slow, and uncertain. How, for example, does one know when an approach misses his or her needs and when it is something that will take time, patience, and perseverance to be helpful? From a narrative perspective, there can be no gold standard or simple answers. Only judgment, wisdom, and trial and error can decide.

References

Angell M. *The Truth about Drug Companies: How They Deceive Us and What to Do about It*. New York: Random House; 2005.

Angus L, McLeod J. Toward an integrative framework for understanding the role of narrative in psychotherapy process. In: Angus L, McLeod J, eds. *Handbook of Narrative and Psychotherapy: Practice, Theory and Research*. Thousand Oaks, CA: Sage; 2004.

Brooks P. *Reading for the Plot: Design and Intention in Narrative*. Cambridge, MA: Harvard University Press; 1984.

Budbill D. The uses of depression: the way around is through. In: Berlin R, ed. *Poets on Prozac: Mental Illness, Treatment and The Creative Process*. Baltimore, MD: Johns Hopkins University Press; 2008.

Casey N. *Unholy Ghost: Writers on Depression*. New York: Harper Perennial; 2002.

Cassel E. The nature of suffering and the goals of medicine. *N Engl J Med*. 1982;306(11):639–645.

Charon R, DasGupta S, Herman N, et al. *The Principles and Practice of Narrative Medicine*. Oxford: Oxford University Press; 2016.

Clark H. *Narratives and Depression: Telling the Dark*. Albany, NY: State University of New York; 2008.

Coles R. *The Mind's Fate: A Psychiatrist Looks at his Profession*. Boston, MA: Little, Brown, and Co; 1995.

Dormen L. Planet No. In: Casey N, ed. *Unholy Ghost: Writers on Depression*. New York: Harper Perennial; 2002.

DuBrul S. The Icarus Project: a counter narrative for psychic diversity. *Journal of Medical*

Humanities. 2014;35:257–271.

Engel G. The need for a new medical model: a challenge for biomedicine. *Science*. 1977;196(4286):129–136.

Franzen J. *The Corrections*. New York: Farrar, Straus & Giroux; 2001.

Fulford B, Colombo T. Six models of mental disorder: a study combining linguistic-analytic and empirical methods. *Philos Psychiatr Psychol*. 2004;11(2):129–144.

Hall S. Introduction: who needs identity. In: Hall S, du Gay P, eds. *Questions of Cultural Identity*. London: Sage; 1996.

Hamkins S. *The Art of Narrative Psychiatry*. Oxford: Oxford University Press; 2013.

Hesse M. The explanatory function of metaphor. In: McErlean J, ed. *Philosophies of Science: From Foundations to Contemporary Issues*. Belmont, CA: Wadsworth; 2000:349.

Hoshmand L. Narrative psychology. In: Kazdin E, ed. *Encyclopedia of Psychology*. Washington, DC: American Psychological Association; 2000:382.

Josen A, Siegler M, Winslade W. *Clinical Ethics: A Practical Approach to Ethical Decisions in Clinical Medicine:* 8th ed. New York: McGraw Hill; 2015.

Lakoff G, Johnson M. *Metaphors We Live By*. Chicago: University of Chicago Press; 1980.

LeFrancois BA, Menzies R, Reaume G. *Mad Matters: A Critical Reader in Canadian Mad Studies*. Toronto: Canadian Scholars Press; 2013.

Martin E. The pharmaceutical person. *BioSocieties*. 2006;1:273.

Messud C. *The Emperor's Children*. New York: Vintage Books; 2006.

Ricoeur P. *The Rule of Metaphor: Multi-Disciplinary Studies in the Creation of Meaning in Language*. Toronto: University of Toronto Press; 1977.

Ricoeur P. *Time and Narrative*. Vol 1. Chicago, IL: University of Chicago Press; 1984.

Ridge D, Zeibland S. 'The old me could never have done that': how people give meaning to recovery following depression. *Qual Health Res*. 2006;16(8):1038–1053.

Rose N, Abi-Rahced JM. *Neuro: The New Brain Science and the Management of the Mind*. Princeton: Princeton University Press; 2013.

Rosenzweig S. Some implicit common factors in diverse methods of psychotherapy: At last the Dodo said, 'Everybody has won and all must have prizes.' *Am J Orthopsychiatry*. 1936;6:412.

Sismondo S, Greene J. *The Pharmaceutical Studies Reader*. West Sussex: Wiley Blackwell; 2015.

Theile L. *The Heart of Judgment: Practical Wisdom, Neuroscience, and Narrative:* Cambridge: Cambridge University Press; 2010.

Verghese A. The physician as storyteller. *Ann Intern Med*. 2001;135(11):1012–1017.

Wampold BE, Imel Z. *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work*. 2nd ed. New York: Routledge; 2015.

White M, Epston D. *Narrative Means to Therapeutic Ends*. New York: WW Norton; 1990.

▲ 33.13 Positive Psychology

NANSOOK PARK, PH.D.

POSITIVE PSYCHOLOGY

Positive psychology is an umbrella term describing the scientific study of what makes life most worth living. The good life, that is the happy, healthy, productive, and meaningful life, entails more than the mere absence of disease, disorder, and problems. Traditional approaches in psychology largely focus on identifying problems and reducing them. However, it is limited in addressing how to build and sustain a good life, which requires additional skills and conditions. The field of positive psychology is the science and practice that has the potential to fill this need.

Research findings from positive psychology are intended to provide a more complete and balanced scientific understanding of the human experience. The new field of positive psychology calls for as much focus on strength as on weakness, as much interest in building the best things in life as in repairing the worst, and as much concern with making the lives of normal people fulfilling as with healing pathology.

Positive psychology does not replace business-as-usual psychology, which often focuses on people's problems and how to remedy them. Rather, positive psychology intends to complement and extend a problem-focused psychology. The attention of positive psychologists is increasingly turning to deliberate interventions that promote the well-being of individuals and groups, and the same qualifications apply.

For decades, scientific psychology has neglected the study of what can go right with people and often has had little more to say about the good life than do pop psychologists, inspirational speakers, and armchair gurus. More subtly, the underlying assumptions of psychology have shifted to embrace a disease model of human nature. Human beings are seen as flawed and fragile, victims of cruel environments or casualties of bad genetics, and if not in denial then at best in recovery. This worldview has even crept into the common culture, and some people have become self-identified victims, trying to survive but not to flourish. Positive psychology proposes that it is time to correct this imbalance.

Positive psychology studies what goes right in life, from birth to death. It is concerned with optimal experience, people being their best and doing their best. Everyone's life has peaks and valleys, and positive psychology does not deny the low points. Its signature premise is more nuanced: what is good about life is as genuine as what is bad and therefore deserves equal attention from psychologists. Positive psychology assumes that life entails more than avoiding or undoing problems and that explanations of the good life must do more than reverse accounts of distress and dysfunction.

HISTORY OF POSITIVE PSYCHOLOGY

The most basic assumption that positive psychology urges is that human goodness and excellence are as authentic as disease and disorder. Positive psychologists argue that these topics are neither secondary nor derivative. There are many good examples of psychological research, past and present,

that illustrate the fundamental premise of positive psychology, although only recently has there been explicit emphasis on their common concerns.

Accordingly, positive psychology has a very long past but only a short history. The long past of positive psychology stretches at least to the Athenian philosophers in the West and to Confucius and Lao-tzu in the East. In the writings of these great thinkers can be found the same questions posed by contemporary positive psychologists. What is the good life? Is virtue its own reward? What does it mean to be happy? Is it possible to pursue happiness directly, or is fulfillment a by-product of other pursuits? What roles are played by other people and society as a whole?

Early Ancestors

Within psychology, the premises of positive psychology were laid out long ago. More than 100 years ago, when psychology began as a formal discipline in Germany, it had a narrow focus on sensation and perception. Psychology soon spread. When it came to the United States, it was blended with American pragmatism by William James, John Dewey, and others to result in *functionalism,* which was concerned with the consequences—functions—of psychological states, traits, and processes. Inevitably, functionalism gave rise to a host of applied psychologies, out-of-the-laboratory endeavors with practical goals.

Among the early goals of applied psychology was to understand accomplishment and satisfaction in domains such as education, religion, work, and the family. The very first applied psychologists had the explicit intent to identify what was successful in these important areas of life and to apply this knowledge to make the lives of normal people more fulfilling.

The first psychological clinic in the United States was founded in 1896 by Lightner Witmer at the University of Pennsylvania. Witmer is credited with coining the term clinical psychology, a field which he defined as "the study of individuals, by observation or experimentation, with the intention of promoting change." Many of the clients at Witmer's clinic were school children with learning difficulties, and the goal of the clinical work was to make them better students. Sounding like a contemporary positive psychologist, Witmer argued that "The progress of psychology, as of every other science, will be determined by the amount of its contribution to the advancement of the human race."

Had this original mission of psychology been sustained, there would be no positive psychology today, although all of its current concerns would be well-represented. However, events occurred that changed psychology. With the rise of behaviorism and the positivism that provided its philosophical basis, a strict distinction between fact and value emerged, with fact being seen as the province of science and value as the province of philosophy. Many psychologists stopped studying things that mattered morally.

In the early 20th century, psychologists interested in individual

differences, what is now called personality, divided their subject matter into character—which could be changed—and temperament—which could not. The moral flavor of character was seen as a critical feature. Gordon Allport, the major personality trait theorist in the United States psychology at that time, was uncomfortable with the notion of character and actually banished the term from academic discourse concerning personality. Allport argued that character was the subject matter of philosophy and not psychology. The traits he urged psychologists to study were presumably objective entities stripped of moral significance and linked to "adjustment" but not endowed with inherent value.

Another turning point was a separation between basic psychology and applied psychology. Whereas the early functionalists made no such distinction, the next generation of psychologists in the United States began to draw a line between laboratory experiments investigating "basic" processes of perception, thought, and learning and field studies investigating outcomes of interest because of their substantive content. The former studies were regarded as more important, and the elite schools stopped recruiting applied psychologists to their departments. Applied work of course was continued by psychologists in schools of business, education, kinesiology, and elsewhere, but what they did was not part of the acknowledged mainstream.

As important as these events were, the biggest change in psychology was the result of an external event. The realities of World War II changed psychology and spurred the growth of clinical psychology, now the numerically dominant specialization within psychology. With an unprecedented proportion of the adult male population serving in the armed forces and experiencing the stresses of modern war, the need for acute clinical services outpaced what could be provided by available psychiatrists. Psychologists, heretofore involved in clinical work mainly in devising, administering, and interpreting psychological tests, were called upon to provide treatment. What resulted was a generation of basic psychologists who through necessity became practitioners.

The immediate aftermath of World War II saw no reduction in the need for clinical services. Accordingly, the new generation of scientist-practitioner psychologists consolidated, formalized, and eventually institutionalized an identity and a profession, encouraged and supported by the United States federal government. The National Institute of Mental Health (NIMH) was created to support research into disorders and their treatment, and stipends for graduate training programs in clinical psychology were part of the NIMH portfolio. The Veterans Administration (VA) provided clinical psychology internships for students in these programs.

The growth of clinical psychology as a profession has been phenomenal. The downside of the growth of clinical psychology is that its view of people

has come to dominate much of psychology and even the common culture. What students learn is that people are victims—of unfortunate genetics, cruel environments, and their own cognitive shortcomings. This characterization applies not just to clinical psychology but to the entire discipline. For example, the field of social psychology began in the 1940s with Kurt Lewin's stirring call for *action research,* using psychology to solve social problems, but it has since become a field that documents repeatedly what people cannot know and cannot do.

This is the context in which Martin Seligman, a mainstream academic psychologist, called for psychology to reclaim its original mission and study the rest of the human condition. The field was named in 1998 as one of the initiatives of Martin Seligman in his role as president of the American Psychological Association. A trigger for positive psychology was Seligman's realization that psychology since World War II had focused much of its efforts on human problems and how to remedy them.

The yield of this focus on pathology has been considerable, but there has been a cost. Much of scientific psychology has neglected the study of what can go right with people and often has little to say about the psychological good life. More subtly, the underlying assumptions of psychology have shifted to embrace a disease model of human nature. People are seen as flawed and fragile, casualties of cruel environments or bad genetics. This world view has even crept into the common culture, where many people have become self-identified victims.

Positive psychology is not intended to replace business-as-usual psychology but rather to complement and extend it, by drawing on past work into the psychological good life as well as doing new work from the positive psychology perspective.

In one sense, positive psychology is nothing new, and William James, John Dewey, Lightner Witmer, and other psychologists from a century ago can be considered positive psychologists. Yet, the label "positive psychology" has considerable value in the here and now because it draws together what had been scattered lines of inquiry within psychology and allied social science—e.g., investigations of life and work satisfaction; utopian visions of education like those of Alexander Sutherland Neill; primary prevention programs based on notions of wellness as pioneered by George Albee and Emory Cowen; work by Albert Bandura and others on human agency and efficacy; work by Mihalyi Csikszentmihalyi on creativity, flow, and optimal functioning; George Vaillant's longitudinal studies of the benefits of mature defense mechanisms; investigations of resilience; studies of giftedness; conceptions of intelligence as multiple; and assessments of the quality of life among medical and psychiatric patients that went beyond their symptoms and diseases.

Humanistic Psychology

Chief among the immediate ancestors of today's positive psychology is *humanistic psychology* as popularized by Carl Rogers and Abraham Maslow (who actually used the phrase "positive psychology" to describe his own approach). Humanistic psychology is one of psychology's venerable perspectives, popular in the 1960s and 1970s and still with adherents today. *Humanism* is the doctrine that the needs and values of human beings take precedence over material things, and further, that people cannot be understood simply as part of the material world. Humanists argue that scientific psychologists miss what is most important about people by focusing on the supposed causes of behavior, as if people were simply billiard balls, doing poorly or well depending on what other billiard balls happen to have ricocheted into them.

Both Rogers and Maslow emphasized that people strive to make the most of their potential in a process called *self-actualization*. Self-actualization can be thwarted by various conditions, but if these conditions are changed, then the potential within each individual will necessarily unfold.

This is a very different way of thinking about human nature than that embodied in psychoanalysis or behaviorism, dominant perspectives within psychology during the 20th century. Humanistic psychology stresses the goals for which people strive, their conscious awareness of this striving, the importance of their own choices, and their rationality. The attention of psychology is thereby directed away from mechanical causes and toward fundamental questions about existence and meaning.

In a 2000 discussion of positive psychology, Martin Seligman and Mihalyi Csikszentmihalyi tersely distanced the new field from humanistic psychology. They made two arguments. First, contemporary positive psychology regards both the good and the bad about life as genuine, whereas humanists often—but not always—assume that people are inherently good. Second, positive psychology is strongly committed to the scientific method, whereas humanists often—but again, not always—are skeptical of science and its ability to shed light on what really matters.

As points of relative and occasional contrast, these arguments are reasonable, and they highlight what is novel about positive psychology. Yet, as positive psychology has evolved and as positive psychologists have more carefully examined allied perspectives, positive psychology and humanistic psychology should be seen as close relatives.

Many humanists would agree that each person has the capacity for good and bad, just as positive psychology assumes. That "the good life is simply a matter of choice" seems to go too far, given the well-documented barriers to thriving posed by external circumstances like pestilence, poverty, and prejudice, but many positive psychologists nowadays acknowledge that notions of choice and will are indispensable ones. Finally, many humanistic psychologists, from decades ago to the present are as committed to science

as any positive psychologist—although one can still question the success of humanistic psychology in creating a cumulative database.

Positive Mental Health

One more important historical precursor of positive psychology needs to be mentioned. In 1958, the United States Joint Commission on Mental Illness and Health was authorized by a unanimous resolution of Congress. The commission published a monograph series, and the first book in the series was *Current Concepts of Positive Mental Health* by social psychologist Marie Jahoda. Foreshadowing the arguments of positive psychology decades later, Jahoda argued that mental health was not simply the absence of mental illness, and further that psychological well-being needed to be understood in its own right.

Jahoda surveyed what previous thinkers—mainly clinicians—had to say about positive mental health and integrated their views. She proposed half a dozen underlying processes that produce the states and traits usually identified as psychologically healthy:

- acceptance of oneself
- accurate perception of reality
- autonomy (freedom from social pressures)
- environmental mastery
- growth, development, becoming
- integration of personality

Note that she did not include happiness or satisfaction as part of her definition, believing that these were more properties of the situation than of the individual.

At the same time that Jahoda published her book, social psychologist William Scott surveyed the existing research literature on mental health, focusing on research definitions (measures) of well-being and empirically established correlates of these measures. Then as now, the majority of studies ostensibly looking at health were really studies of disorder. Scott could draw firm conclusions only about the factors that characterized the lack of pathology. Good social relationships were the most common correlate. Scott cited a personal communication from Jahoda in which she described her own attempts to develop measures of several of her criteria for mental health (i.e., accurate perception of reality and environmental mastery) as less than satisfactory, which may explain why her ideas did not stimulate further work for decades.

More recently, Carol Ryff and her colleagues repeated Jahoda's analysis by surveying and integrating what different theorists, again mostly clinicians, said about the psychological components of being and doing well—striving and thriving, as it were. They identified six points of convergence

across discussions of psychological well-being, which agree substantially with those specified decades earlier by Jahoda:

- autonomy
- environmental mastery
- personal growth
- positive relations with others
- purpose in life
- self-acceptance

Notable about Ryff's work is that she and her colleagues have solved the measurement problem and have created reliable and valid self-report surveys of these components of psychological well-being. Using these measures, they have explored the links between psychological well-being and physical health, finding positive associations and implicating the left prefrontal cortex—believed to aid in the organization of goal-directed activity—in these associations.

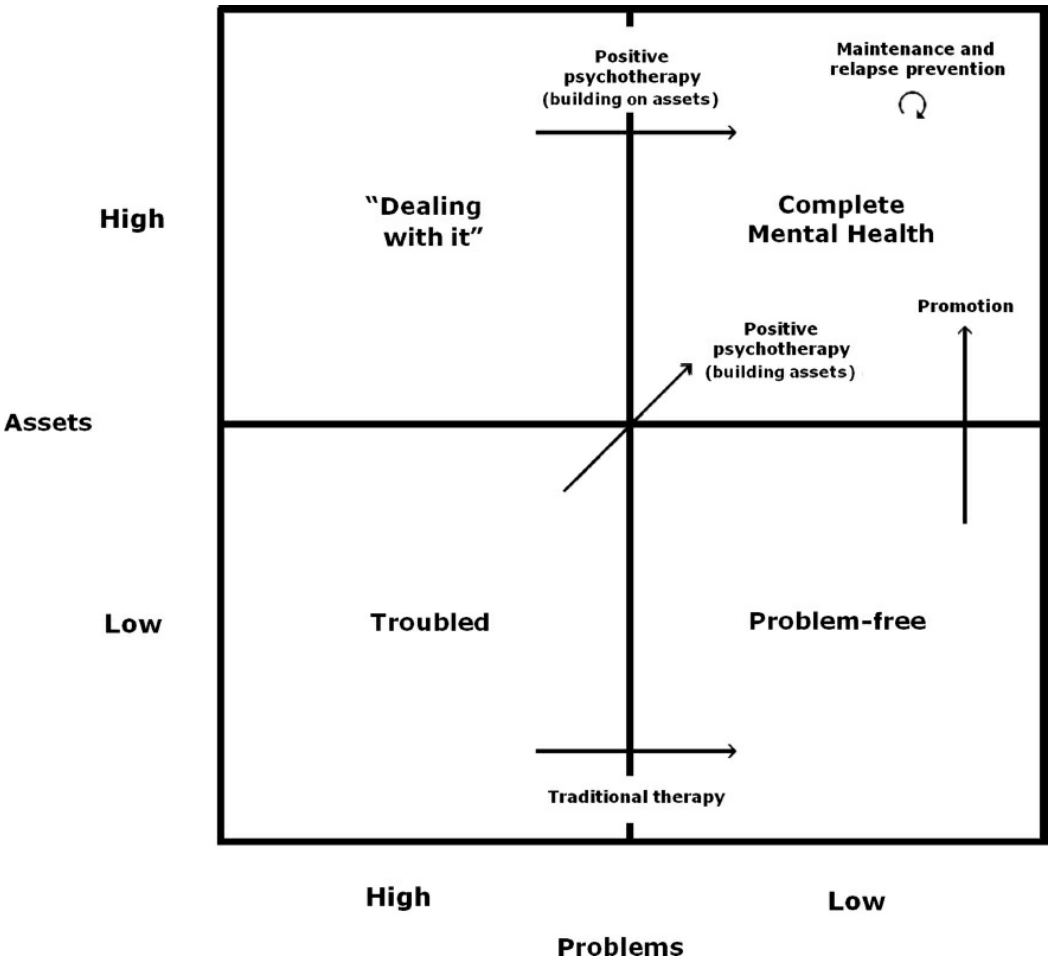

FIGURE 33.13–1. Mental health and mental illness. These are dimensions. Quadrants are shown for the purpose of illustration only.

Building on Jahoda's earlier work, contemporary positive psychologists

suggest that doing well (mental health) and doing poorly (mental illness) are not simple opposites. They argue that these are distinguishable dimensions in terms of which people can be described (Fig. 33.13–1). Complete mental health is defined by high levels of psychological well-being along with the absence of signs and symptoms of psychological disorder. Said in DSM terms, complete mental illness is described by one or more Axis I diagnoses and a low Axis V (global assessment of functioning) score.

Figure 33.13–1 represents how we think about intervention from a positive psychology perspective. It is based on one of the signature premises of positive psychology that well-being is not simply the absence of pathology. Both well-being and pathology are multidimensional, and their various components can and do coexist. The framework is shown in Figure 33.13–1. This framework depicts strengths and problems as in principle independent of one another.

Individuals can be placed in Figure 33.13–1 depending on their particular profiles of strengths and problems. In Quadrant 1 are strengths accompanied by problems. We like to describe individuals in Quadrant 1 as "dealing with it," meaning that they have and use strengths despite experiencing problems and challenges.

The fully-functioning person is found in Quadrant 2: high strengths and low problems. Here we find health as defined by in the constitution of the World Health Organization: "A state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity."

Individuals in Quadrant 3 are the ones most frequently studied by clinical psychologists, poorly-functioning individuals with psychiatric diagnoses or problems such as alienation, demoralization, and loneliness.

Quadrant 4 is both unremarkable and understudied, occupied by individuals with neither problems nor strengths, passing time quietly, complacent but not joyful. What may characterize Quadrant 4 is a history of insufficient challenge. Because life has been easy, strengths have not developed or have not needed to be displayed.

Traditional psychotherapy or consultation is usually concerned with solving problems, moving individuals from Quadrant 3 to Quadrant 4. Figure 33.13–1 suggests additional intervention goals. The ultimate goal of practitioners should be to move people and groups into Quadrant 2 and to keep them there. Depending on where individuals and groups start, different interventions are probably needed because change will entail different processes. In some cases, strengths need to be built from scratch, whereas in other cases, individuals and groups need to capitalize on existing strengths.

Even among individuals and groups already in Quadrant 2, their well-being can be further promoted. Promotion needs to be distinguished from prevention, the goal of which is to prevent problems by reducing risk

factors. Prevention is of course important, but attention to risk factors alone is not sufficient because people who develop problems may lack obvious risk factors. The positive psychology perspective is that building or enhancing strengths may be an additional and broader strategy for intervention.

When first introduced by Seligman, positive psychology was described as focusing on complete mental health and its further enhancement, but the scheme suggested by positive psychologists sensibly expands positive psychology and provides a useful way to describe the range of therapeutic interventions. For example, what can be called *positive psychotherapy* attempts to move people with problems closer to the state of complete mental health. The goal of *promotion* is to increase the well-being of those without problems. *Maintenance* keeps healthy people healthy. Given that existing treatments for psychological problems rarely if ever "cure" people in the sense of precluding recurrence of problems, it is surprising how little attention has been paid to the matter of maintenance. Along these lines, positive psychology may provide new insights into the challenging task of *relapse prevention*: helping people whose problems have been resolved to stay healthy. *Traditional therapy* simply tries to mitigate someone's psychological problems.

Incomplete mental illness and incomplete mental health are common states, and positive psychology has potential value in helping people who can be so described, as well as learning lessons from them. Incomplete mental illness might be a suitable label for people like Abraham Lincoln or Winston Churchill, who functioned well despite recurrent depression; somehow they had learned to "deal with" their problems. Incomplete mental health might be a suitable label for those who lead lives of quiet desperation or those who earn NOS (not otherwise specified) diagnoses should they enter the mental health system.

CONCERNS OF POSITIVE PSYCHOLOGY AND CRITICISMS

The framework of positive psychology provides a comprehensive scheme for describing and understanding the good life and how it can be attained and sustained, regardless of one's starting point. The concerns of the field can be divided into four related topics:

- positive subjective experiences (happiness, gratification, fulfillment, and flow)
- positive individual traits (strengths of character like humor, love, gratitude, forgiveness, modesty, spirituality, optimism, kindness, creativity, courage, and wisdom; talents; interests; values; and meaning and purpose)
- positive interpersonal relationships (friendship, marriage, and colleagueship)

- positive institutions (families, schools, businesses, communities, and societies)

A theory is implied here: positive institutions enable the development of positive relationships, which enable the display of positive traits, which in turn enable positive subjective experiences.

The word "enable" avoids strict causal language. It is possible for people to be happy or content even in the absence of good character, and good character can operate against the interpersonal and institutional grain. However, people are at their best when institutions, relationships, traits, and experiences are in alignment. Doing well in life represents a coming together of all four domains.

As a new field, positive psychology raises questions. Despite the pervasiveness of a victim mentality, everyday people seem to know that the elimination or reduction of problems is not all that is involved in improving the human condition. In contrast, the scientific community is more skeptical. Contributing to skepticism are widespread assumptions within the social sciences about human nature as brittle and broken, notions more explicit among social scientists than the general public.

Positive psychology is also criticized in some quarters for a relentless emphasis on being positive, that is, happy and cheerful. This criticism of the field reflects a misunderstanding, because positive psychologists merely propose that what is positive about life is worth studying, in addition to what is negative. Happiness is one of many topics of interest to positive psychology, but so too are virtues like gratitude and forgiveness, generosity, wisdom, optimism, and resilience, meaning and purpose, engagement, and good relationships.

Positive psychologists do not deny the problems that people experience, and positive psychologists do not ignore the negative—stress and challenge—in their attempts to understand what it means to live well. As already emphasized, positive psychology intends, by expanding the topics of legitimate study, to complement business-as-usual psychology, not to replace it.

Indeed, what is most troubling in life can set the stage for what is most fulfilling. Consider that complex emotional experiences often blend the positive and negative, that optimism is most apparent when people confront setbacks and failures, that crisis reveals strengths of character, and that ongoing challenge is a prerequisite to experience flow in the moment and to achieve something important in a lifetime. Along these lines, many positive psychologists believe that identifying and using what one does well can be an effective way to address and resolve psychological problems.

In fairness to critics, positive psychology has become so popular that at least some people appropriate the label and present "positive psychology"

as a simplified version of the power of positive thinking. All one can do here is assert that Pollyanna is not the poster child for positive psychology, that the fortunate in life—those who are happy—deserve no special congratulations, and that the unfortunate—those who are unhappy—deserve no special blame.

Positive psychology does not endorse wishful thinking, self-deception, or quick fixes. Rather, it uses the scientific method to understand and enhance well-being. The lesson of history is that patience is needed. Great thinkers for many centuries have pondered well-being, and progress in positive psychology will be slow and difficult.

EMPIRICAL FINDINGS

Positive psychology already has a canon of established findings worth considering. Here are some of the research findings that provide some insight into ways to build a happy, healthy, resilient, and thriving life.

In contrast to stereotypes that happy people are naïve, stupid, or in denial, empirical studies show that happiness and positive emotions are beneficial. The real worry should be about people who are unhappy and not satisfied with their life.

Most people are happy. When psychologists study self-reported happiness and life satisfaction, usually under the rubric of subjective well-being, they administer numerical rating scales. The consistent and perhaps surprising result is that most people in most circumstances most of the time score above the scale midpoint, whether they are multimillionaires in the United States or pavement dwellers in Calcutta. This conclusion holds across demographic characteristics like age, gender, ethnicity, and education, each of which has a surprisingly small association with avowed happiness.

Happiness is causal not epiphenomenal. People who are successful in life's venues are of course happy, but the less obvious and more interesting finding from experimental and longitudinal research is that happiness actually foreshadows success in academic, vocational, and interpersonal realms.

For example, researchers analyzed pictures from the 1958 and 1960 yearbooks of a women's college. All but three of the young women were smiling, but the smiles varied. Some displayed what is termed a Duchenne smile: a genuine, full-faced expression of happiness indexed by the degree to which the muscles surrounding one's eyes are contracted. Others showed the forced grimace known as a flight-attendant smile. The researchers chose these particular pictures for analysis because the women in them were participants in a long-term study of important life events. Specifically, the researchers knew—decades after their yearbook photos—whether the women were married and if they were satisfied with their marriage. As it turns out, ratings of the Duchenne-ness of their yearbook smiles predicted

both of these outcomes. Young women who expressed genuine positive emotions (happiness) in yearbook photos, and presumably in other areas of their lives, as middle-aged women had better marriages. The skeptic might wonder if these results reflect the operation of some confound like physical attractiveness. As it turns out, rated "prettiness" was independent of the Duchenne-ness rating and moreover did not predict who had an intact or satisfying marriage.

Happiness can be increased. The possibility of deliberately pursuing happiness is a relatively recent arrival on the historical scene, immortalized by Thomas Jefferson in the United States Declaration of Independence. Prior to this time, happiness was relegated to the afterlife or seen as a matter of luck. Even today, some psychologists are skeptical that happiness can be increased.

Several factors contribute to this skepticism. First is the well-documented fact that people adapt to pleasure. When they repeatedly encounter the same pleasure-producing stimulus, they experience increasingly less pleasure in response. Adaptation to pleasure is so common that theorists have proposed that people live on a *hedonic treadmill,* meaning that they continually adapt to improving circumstances to the point that they always return to a point of relative neutrality. Another reason for pessimism about the possibility of increasing happiness is its demonstrable heritability, which has led theorists to posit a genetically-determined *happiness set-point,* above which one cannot rise.

Research challenges these arguments. Current thinking about the hedonic treadmill and the happiness set-point holds that these are at best loose metaphors and hardly biological realities. Remember the finding that most people most of the time are happy, which means that the metaphorical treadmill is not set at neutral. Furthermore, longitudinal studies show that the metaphorical set-point can be altered by life events like marriage and unemployment.

Other studies show that happiness can be increased when people simply act differently. For example, researchers asked introverted and extraverted college students to keep a careful diary for 3 weeks of their activities and mood. Both introverts and extraverts were happiest when doing extraverted things, implying that anyone who behaves in an outgoing way will be happy as a consequence, whether or not this behavior comes naturally. Happiness results not from what one is predisposed to do but instead from what one actually does.

Along these lines, researchers have shown that people instructed to do the sorts of things that characterize happy people indeed become happier as a result:

- Being grateful (e.g., counting blessings on a daily basis)
- Creating flow (e.g., seeking out activities at work or at leisure that

challenge one's abilities)

- Giving service to others (e.g., performing volunteer work)
- Performing acts of kindness (e.g., doing small good deeds for others on a daily basis)
- Savoring good things that happen (e.g., sharing experiences and memories with others, taking souvenirs, congratulating the self)
- Identifying and using character strengths in novel ways (e.g., seeking out different ways to display one's humor, teamwork, or love of learning)
- Articulating one's hopes and wishes (e.g., writing about an envisioned and desired "possible self")

People are terrible at emotional forecasting. People of course know that a promotion at work will make them happy and that a divorce will make them sad, but research consistently shows that people cannot accurately estimate how long these emotional reactions will last. They consistently overestimate their duration and further seem not to learn from experience that they do so. Clinicians know that part of treating depression is addressing the convictions of clients that they will never feel any differently, when in fact most depressive episodes are self-limiting. Conversely, the positive psychology practitioner needs to convey the information that even the most exhilarating experiences and activities lose their magic, sometimes quickly.

Happiness leads to physical well-being. People who are happy recover more quickly from infectious disease and live longer than their peers. For example, in an experimental study, volunteers were exposed to the virus that causes the common cold. All participants became ill, but those who were happier had less severe symptoms and recovered more quickly. For another example, from the well-known longitudinal study of Catholic nuns, women who described positive emotions in essays they wrote in their early twenties were more likely to be alive decades later. Indeed, the happier nuns (those in the upper 25 percent of the essay writers) lived on average 10 years longer than their less happy counterparts (those in the bottom 25 percent). Positive psychology interventions that bolster happiness may be of considerable value in the applied fields of behavioral medicine and health psychology.

Positive emotions and negative emotions are distinct. Theorists who catalogue basic emotions usually include one or two positive ones—e.g., joy and sometimes surprise or curiosity—but their focus is on the negative ones. Part of the reason for neglecting positive emotions is the problem-focused nature of psychology in general; "emotional" problems are after all a synonym for psychological disorders. However, there are important differences between positive emotions and negative emotions. They have separate bases in brain chemistry, and they can coexist. They also function differently. Negative emotions are associated with specific behavioral

predispositions. Fear makes people want to run, and anger makes them want to attack. In contrast, positive emotions are not linked to such specific action tendencies. Positive emotions like joy may activate people, but in a much more vague and diffuse way.

Positive emotions broaden and build people's psychological and behavioral repertoires. Negative emotions—fear, anxiety, anger—alert people to danger. When one experiences a negative emotion, response options narrow, and the person acts with haste to avoid, escape, or undo whatever danger is signaled. In contrast, positive emotions signal safety, and the inherent response to them is not to narrow options but to broaden and build upon them. Laboratory studies have shown that inducing positive moods among research participants (by showing them brief video clips) leads to enhanced performance on tasks requiring creativity and flexibility. Perhaps it is advantageous to experience positive emotions because they lead people to engage in activities that add to their behavioral and cognitive repertoires, building so-called psychological capital.

Positive emotions undo the physiological effects of negative emotions. For example, in one study, college students were given just 1 minute to prepare a brief speech on "why you are a good friend" that would—so they were told—be videotaped and evaluated by their peers. They did so while having their heart rate, peripheral vasoconstriction, and blood pressure (all indices of sympathetic nervous system activation) measured. Following this preparation time, research participants saw one of four different films; two produced positive emotions (joy or contentment), one produced feelings of sadness, and the last one was emotionally neutral. None of these films in and of itself had an effect on sympathetic arousal. However, following the speech preparation exercise, research participants who then saw either of the positive emotions films showed more rapid cardiovascular recovery (return to baseline) than those who saw the sad film or the neutral film.

Engagement matters. Flow is the term used to describe the psychological state that accompanies highly engaging activities. During flow, time passes quickly for the engaged individual. Attention is focused on the activity itself. The sense of the self as a social actor is lost. The aftermath of the flow experience is invigorating. Flow is not to be confused with sensual pleasure. Indeed, flow in the moment is nonemotional and arguably nonconscious. People describe flow as highly and intrinsically enjoyable, but this is an after-the-fact summary judgment, and joy is not immediately present during the activity itself. Flow can be described as the experience of working at full capacity, which is why the concept figures prominently in positive psychology.

Unlike many positive psychology concepts, flow has been extensively investigated, using an experience sampling method which asks research participants to carry a pager that goes off at random intervals. They are then asked to describe what they are doing and how they are feeling at the

time. From studies of this sort, psychologists have learned that flow is most likely to occur when there is an optimal balance between skill and challenge. One need not be an expert in a domain to experience flow. All that matters is that the presented challenge meets one's skills, and vice versa.

Flow can be experienced in all sorts of activities, at work or play, but usually among those perceived as voluntary. For example, homework for most school children strikes a balance between challenge and skill, but because homework is regarded as coerced, it rarely produces the flow state. Indeed, flow is rarely experienced by youth during any school activity. The other common activities of adolescents in the United States—watching television or "hanging out" with friends—also fail to produce flow, because they fail to meet one or another of its enabling conditions. People differ greatly in the frequency with which they experience flow, for reasons not well understood, and young people who experience more flow during adolescence show long-term desirable consequences, such as achievement in creative domains. They may even have better physical health.

"Good days" have common features. Wishing someone a good day is a common conversational gambit, and positive psychologists have been able to describe what a good day entails. For most people, it is one in which they do things that make them feel autonomous, competent, and connected to others. Encouraging people to experience pleasant events has long been a staple of behavioral therapy for depression, and this research makes the intervention less a matter of trial-and-error.

A rosy view of matters is associated with physical, psychological, and social well-being. For example, optimism assessed early in life predicts health decades later and even whether someone is still alive. For another example, viewing one's spouse in a more positive way than do others—presumably an idealization—predicts a good and lasting marriage. Data showing that positive illusions like these are beneficial stand in sharp contrast to theoretical arguments mounted by typical clinical psychologists and psychiatrists that realism and accuracy are the hallmarks of health.

Positive expectations drive analgesic placebo effects through physiologic pathways. Specifically, dopamine—implicated in the experience of positive emotions—triggers the release of endorphins. Optimism and hope are not just in one's head but also in one's nervous system.

Most people are resilient. Stressful life events can take an enduring toll on well-being, and their consequences have been extensively studied. Usually not highlighted in research on post-traumatic stress disorder is that the majority of people exposed to life-threatening events do not experience lasting negative effects.

One of the effective strategies that contributes to resilience is the positive reframing of the unfortunate situation or events. One example is

the phenomenon of *downward social comparison* described by psychologist Shelley Taylor. Her research program began as an investigation of depression that followed in the wake of a diagnosis of breast cancer, a terrible and serious life event. However, Taylor had trouble identifying a sufficient number of cancer patients who were severely depressed. Instead, most were able to deal with the diagnosis by a process by thinking of someone who was faring more poorly than they were:

> I have breast cancer, but it could be worse ... after all, I could be younger ... the surgery required is a lumpectomy rather than a mastectomy, a single mastectomy rather than a double mastectomy ... I can tolerate the side effects of chemotherapy better than most ... I have a supportive family ... and so on.

A previous generation of psychologists might have concluded that these women were in denial, but they were open-eyed, lucid, and sober. This study foreshadowed the premises of positive psychology by concluding that resilience and the cognitive work that sustains it were the rule rather than the exception.

In some cases, people may even show psychological growth following adversity, evidencing improved relationships with others, openness to new possibilities, greater appreciation of life, enhanced personal strength, and spiritual development. Interventions that specifically target these components of posttraumatic growth may help people recover from the effects of traumatic events.

Virtue is more than its own reward. Positive psychologists have begun to study the sorts of character strengths enumerated by philosophers and theologians over the centuries. Several findings are important. First is that notions of good character are not idiosyncratic or culturally bound. There is a strong consensus about the qualities of a "good" person. Second is that character strengths have demonstrable consequences. For example, the strengths of love, hope, curiosity, gratitude, and zest are robustly linked to greater life satisfaction and less depression. Perseverance foreshadows successful at school and at work. Hope and optimism predict good health and a long life. So too does generosity as indexed by altruistic acts.

Meaning and purpose matter. According to positive psychologists, a meaningful life is one in which people feel connected to something larger than themselves. Research consistently links the presence of meaning to well-being. Individuals with a sense of purpose report greater life satisfaction, more positive affect, higher levels of optimism, and better self-esteem. They are less likely to have psychological problems. However, the search for meaning is related to lower life satisfaction. It seems that although having meaning and purpose is beneficial, the process of finding life meaning may entail challenge and confusion. Research further shows that a life framed by meaning is more satisfying than a life centered on pleasure.

For example, researchers asked undergraduate students to keep an

online diary for several weeks, noting the frequency they did pleasurable things or more meaningful things like helping others or simply listening to their problems, as well as how satisfied they were each day. The more people participated in meaningful activities, the happier they were. Pleasure, in contrast, was unrelated to satisfaction. In order to investigate the direction of these effects, the researchers looked at the sequencing of activities and satisfaction, finding that participants became happier after they did something good, suggesting that happiness does, in fact, result from meaningful activities and not vice versa.

Other people matter. Good relationships with other people are the single most important contributor to a satisfied life and may even be a necessary condition for happiness. Having a "best friend" at work is a strong predictor of work satisfaction and even productivity. A good relationship is one in which the amount of positive communication considerably outweighs the amount of negative communication.

Positive psychologists have taken a close look at the features of positive communication, describing four ways that a person can respond to someone else when he or she shares good news, including good events like a raise at work:

- Active-constructive responding—a positive and enthusiastic response: "That's great; I am so proud of you; tell me more about it."
- Active-destructive responding—a response that is critical or points out the potential downside: "Are they going to expect more of you now?"
- Passive-constructive-responding—a muted response: "That's nice, dear."
- Passive-destructive—a response that conveys disinterest: "It rained all day here; what's for dinner?"

Couples who use active-constructive responding have good relationships. The other responses, if they dominate, are associated with relationship troubles. Although this research has only been done in the context of romantic relations, it may well generalize to other relationships.

The important correlates of happiness are social in nature. In contrast to the modest demographic correlates of happiness and well-being, consider the following robust correlates:

- Number of friends
- Being married
- Being extraverted
- Being grateful
- Being religious/spiritual
- Pursuing leisure activities
- Employment (not income)

In a study that compared happy people to very happy people, there was

one striking difference: good relationships with other people. Of the very happy people in the sample, all had close relationships with others. Psychology research documents very few necessary or sufficient conditions, but these data suggest that good social relationships may be a necessary condition for extreme happiness.

Kindness and generosity to others matter. Studies consistently show that doing volunteer work, helping, and giving are associated with high life satisfaction and good health. The implication of these findings is that encouraging people who have experienced difficult events to turn their attention to others might be beneficial. Doing so changes how the givers think about themselves, in positive ways that make them happier and healthier. It is also conceivable that doing service for others, as well as the feeling of contributing to others and society, gives people life meaning, that is the sense that their life matters. In particular, when people experience adversity and uncontrollable events, they feel helpless and depressed.

Perhaps encouraging people to engage in prosocial or service activities would empower them and increase their sense of self-worthiness. The activities do not need to be organized volunteer activities. Even simple acts of kindness in the community, followed by reflection on them, as through journaling, could have therapeutic effectiveness. For example, in a study with Japanese college students, merely asking them to count the kind acts they performed during the day led to increased happiness and more kind behaviors.

Money makes an ever-diminishing contribution to happiness. People who are so poor that they cannot meet their basic needs of course are unhappy, but above extreme poverty, increased income has a surprisingly small relationship with life satisfaction. In fact, money was related to experiencing less daily sadness, but no more daily happiness. The relationship between money and happiness is more complex. What people do with money matters more for happiness. That is, spending money on experiences increased happiness, while spending money on buying things did not. Money has also shown to have a downside. People with larger amounts of money savored less positive experiences compared to their lower income counterparts.

Work matters. Despite the small contribution of income to well-being, whether or not someone is working is much more strongly related to happiness. People who are employed and engaged in what they do are happy, regardless of the status or compensation associated with their job. Happiness and engagement lead people to regard their work as a calling and to be more productive at whatever they do, take fewer sick days, and even postpone their retirement.

Eudaimonia trumps hedonism. According to Aristotle's notion of *eudaimonia*—being true to one's inner self (demon)—true happiness entails identifying one's virtues, cultivating them, and living in accordance

with them. Contrast this notion with the equally venerable idea of *hedonism*—pursuing pleasure and avoiding pain—that is the foundation for *utilitarianism,* which in turn provides the underpinning of psychoanalysis and all but the most radical of the behaviorisms. Research shows that eudaimonia consistently trumps pleasure as a predictor of life satisfaction. Those who pursue eudaimonic goals and activities are more satisfied than those who pursue pleasure. This is not to say that hedonism is irrelevant to life satisfaction, just that all things being equal, hedonism contributes less to long-term happiness than does eudaimonia.

The "heart" matters more than the "head." As already mentioned, positive psychologists have begun to study character strengths—individual differences like curiosity, creativity, kindness, and teamwork. Measures have been devised and administered to people across the lifespan. Research consistently shows that strengths of the "heart" that connect people together—like love and gratitude—are much more strongly associated with well-being than are strengths of the "head" that are individual in nature like creativity critical thinking, and aesthetic appreciation. Formal education of courses stresses the latter strengths, but if one goal of education is to encourage the good life, the research results suggest that the former strengths deserve attention as well.

Happiness, strengths of character, and good social relationships are buffers against the damaging effects of stressful life events. No one will go through life without problems and setbacks, but to the degree that people have more life satisfaction, greater character strengths, and better social support, they will experience fewer psychological or physical problems in the wake of difficulties and recover more quickly from them. These findings provide targets for primary prevention and may represent a unique contribution of the positive psychology perspective. Although all can agree that prevention of psychological problems is more efficient and humane than their treatment once they occur, understanding of prevention has lagged far behind what has been learned about the etiology of psychopathology. Perhaps researchers have been looking at the wrong things—risk factors rather than protective factors.

Positive institutions have common features. Over the last decade, research on positive institutions has been growing steadily. There is agreement that institutions that allow people to flourish—whether families, schools, workplaces, or even entire societies—share a core of common characteristics:

- Purpose—a shared vision of the moral goals of the institution, one reinforced by remembrances and celebrations
- Safety—protection against threat, danger, and exploitation
- Fairness—equitable rules governing reward and punishment and the means for consistently enforcing them

- Humanity—mutual care and concern
- Dignity—the treatment of all people in the institution as individuals regardless of their position

POSITIVE PSYCHOLOGY AND CLINICAL WORK

The goal of interventions based on positive psychology was not only relieving problems but also helping people with or without problems to lead a fulfilling life. The therapeutic goals were not to move people from −5 to 0 (problem-free)—the presumed goal of conventional psychology—but rather break through the zero point.

Interventions based on positive psychology can help people to do more than help people doing well in life do even better. These interventions may also help people with psychological problems lead a more thriving life, moving them from −5 to +2 or beyond. The core of positive psychology interventions is the strength-based approach. Whatever the presenting complaints, people also bring into therapy assets and strengths that can be used to resolve their problems. A crucial task of any treatment is therefore to identify a client's strengths and resources and encourage their use, not just to solve problems but more generally to encourage flourishing. Such a balanced approach should build rapport, bolster client confidence. It would certainly make sense to clients.

When asked how they can tell that treatment has been effective, individuals with DSM diagnoses of depressive disorder describe their own view of "remission" in positive psychology language, spontaneously mentioning that would be more optimistic and confident, would have more energy, and would function well. They did *not* mention merely the reduction of symptoms or a change in diagnostic status. That is, depressed people want to be more than just relieved of negative symptoms from their therapy. They also want to attain certain positive aspects of life.

Positive psychology interventions may also help psychologically healthy people stay that way—keeping them in the right-hand quadrants of Figure 33.13–1—by expanding the targets of prevention programs beyond risk factors to include those that maintain and create resilience. For those who have recovered from psychological difficulties, positive psychology interventions may prevent relapse, keeping them in the right-hand quadrants of Figure 33.13–1. Finally, these interventions may encourage and sustain physical health.

Perhaps the most general contribution of positive psychology to clinical work is to provide a vision of psychological health. Even if therapeutic concern is with traditional diagnostic categories, positive psychology provides substantive goals for any therapy that go beyond symptom reduction or removal.

Positive Psychology's Vision of Psychological Health

In its 1948 constitution, the World Health Organization (WHO) defined *health* as "a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity." In more recent years, this statement has been expanded to include the ability to lead a life that is socially and economically productive. This definition is an important declaration that health entails more than the absence of illness, but it is circular, inasmuch as "well-being" is a synonym for "health." Work by positive psychologists makes this definition more concrete and thus useful as a guide for research and intervention.

If one can extrapolate from the sorts of topics that have been studied, positive psychology assumes that people are doing well when they experience more positive feelings than negative feelings, are satisfied with their lives as they have been lived, have deep positive relationships, have identified what they do well and use these talents and strengths on an ongoing basis, are highly engaged in what they do, are contributing members of a social community, and have a sense of meaning and purpose in their lives. Physical health and safety of course provide an important context for psychological well-being. It is difficult to imagine a cultural group in which these components of the good life are not valued. Respect for human diversity need not entail extreme cultural relativism.

Note that this fuller characterization of health reflects the WHO definition and is drawn from research in all of the domains of concern to contemporary positive psychology. The relevant research cautions that health so defined is not unitary. No one can have it all, at least at the same time, given tradeoffs among the psychological states and traits that reflect doing well. Psychological health therefore needs to be described with a profile of features and not a single summary score (cf. DSM Axis V). DSM describes hundreds of psychological problems, and there may be just as many different manifestations of the good life.

The positive psychology vision of optimal functioning can also be used to make sense of the psychological problems that people do have. Problems (from a positive psychology perspective) represent shortcomings with respect to one or more of the components of the psychological good life. For example, people may experience more negative affect than positive affect; they may be dissatisfied with their lives; they may be socially estranged; they may be bored, alienated, or nihilistic; and so on.

These ways of thinking about problems of course differ from the DSM vision of psychopathology in terms of a set of symptom-defined disorders that are either present or absent. The positive psychology vision regards so-called "symptoms" as the problem and further expects them to exist in degrees. These visions need not compete. Traditional DSM diagnostic categories are not independent of the components of well-being implied by a positive psychology perspective, and a positive psychology perspective explains why given DSM symptoms and diagnoses constitute problems in

living by identifying the ways in which optimal personal and social functioning are compromised by the disorder and its symptoms.

For example, it can be difficult to say why depression *per se* is considered a disorder, especially in light of discussions about its occasional utility and its basis in the evolutionary history of the human species. Yet, if depression is examined further in terms of its association with such difficulties as disinterest, indifference, helplessness, isolation and loneliness, pessimism, and alienation, then one can readily understand when and why depression is a psychological problem. At least in principle, a person who is "depressed" only in the sense of being sad and subdued does not have a problem. This person might simply be solemn and cautious, and Prozac should not be prescribed to counteract what may really be strengths of character.

Positive psychology can be criticized for being value-infused. The offsetting virtue is that the values inherent in positive psychology are clearly stated. The real question is whether these values are idiosyncratic or culture-bound. At a certain level of abstraction, the vision of psychological well-being that emerges from positive psychology seems ubiquitous if not universal.

As a perspective on topics that deserve scientific study, positive psychology has no single theory. Like much of contemporary psychology, it instead relies on midrange theories that draw on a variety of larger perspectives, from evolutionary to behavioral to cognitive to sociocultural models, in order to make sense of specific phenomena. Different topics are explained with different theories. The eventual integration of psychology may be a worthy goal, but it has not yet been achieved.

Positive Psychology Assessment

Much of psychological assessment has been tilted toward identifying weaknesses, deficiencies, and problems. The positive psychology perspective is that problem-focused assessment should be expanded to include strengths, assets and competence. For example, low life satisfaction can occur in the absence of psychopathology, and it is nonetheless related to psychological and social problems. Conversely, high life satisfaction is linked to good functioning even in the presence of symptoms. Absence of problem is not equal to health. Problems and strengths can coexist.

Positive psychologists have already developed an impressive set of measurement instruments that allow someone doing assessment to break through the zero point of deficiency measures. For example, the healthiest that one can score on a typical measure of depression is 0, but this lumps together people who are blasé with those who are filled with zest and joy. The distinction seems well worth making, and the self-report surveys and interviews developed by positive psychologists allow it.

Space does not permit a detailed presentation of positive psychology

measures, but Table 33.13–1 provides some examples. For example, the Values in Action Inventory of Strengths (VIA-IS) is a comprehensive survey that measures an individual's character strengths, 24 positive traits like curiosity, social intelligence, hope, kindness, zest, and teamwork. Positive psychology interventions are often described as strengths-based, which means that the results of this survey can provide useful information for the therapist and the client in identifying the client's strengths and devising interventions.

Many of the popular positive self-report surveys are available online at no cost at www.authentichappiness.org. Upon completion of a survey, individual feedback is provided, which means that clients can be directed to this website and asked to keep a record of their scores.

Most of the existing positive psychology measures were developed for research purposes, and they are most valid when aggregated to yield conclusions about groups of people. They can also be used ipsatively, to describe the psychological characteristics of an individual and how they stay the same or change over time, but the cautious use of these descriptions is as a point of discussion and departure in treatment. They are not strong diagnostic tests, and none should be treated as if they are. They should not be used alone for important decisions. Prudence is necessary for its interpretation and use, especially for the clinical setting. Such prudence is appropriate for all psychological assessment, but it is worth emphasizing in the special case of positive psychology measures.

Table 33.13–1.
Positive Psychology Measures

Positive Affect
Positive and Negative Affect Schedule (PANAS)
Authentic Happiness Inventory
Orientations to Happiness Scale (measures endorsement of pleasure, engagement, meaning)
Subjective Vitality Scale
Savoring Beliefs Inventory
Life Satisfaction
Satisfaction with Life Scale
Marital satisfaction
Work satisfaction
Positive Traits
Values in Action Inventory of Strengths (VIA) (adult and youth version)
Ryff and Singer's Psychological Well-Being Scales
Search Institute's Developmental Assets (for youth)
The Life Orientation Test-Revised (measures optimism)
Attributional style questionnaire (measures optimistic explanatory style)
Values
Values Inventories of Rokeach, Schwartz, Scott, and others
The Gratitude Questionnaire
Transgression-Related Interpersonal Motivations Inventory (measures forgiveness)
Life Meaning and Purpose

The Meaning in Life Questionnaire

Interests

Strong–Campbell Vocational Interest Blank

Abilities

Multiple intelligences

Social Support and Attachment

The Multidimensional Scale of Perceived Social Support

Adult Attachment Style Questionnaire

Experiences in Close Relationships-Revised

Positive Psychology Techniques

Positive psychologists have shown that a variety of brief interventions can, in the short term, increase well-being as measured by happiness, satisfaction, and fulfillment. In some cases, there is also evidence that they can alleviate depression. These interventions have an obvious role not only in remedying distress following experience with adversity but also in building resources that blunt the negative effects of adversity in the first place.

When people experience adversity, they could be overwhelmed to appreciate good things in life. In a recent experimental study by Lauren Conklin and her colleagues, the depressed people showed no difficulties in learning negative information but had trouble in learning positive information. Depressed people may have trouble recognizing and remembering the good things in life. This suggests that therapists should consider helping clients to identify, appreciate, and celebrate positive aspects of their everyday life. Here are some promising positive psychology intervention techniques of potential value in prevention, promotion, and/or treatment interventions.

Savoring. Savoring is to one's awareness of pleasure and one's deliberate attempt to make it last. One way to cultivate positive emotion is to maximize when good things happen in one's life. Fred Bryant and Joseph Veroff examined the effects of savoring good events and found that people who do so are more satisfied. They also identified simple strategies for savoring, such as sharing good events with others, either in the moment or after-the-fact; building memories of the good events (e.g., photographs, diaries, souvenirs); congratulating the self when good things happen; sharpening perceptions during the experience of good events; and becoming fully absorbed in pleasure and not thinking about other matters. Professionals could also help their clients to use more of these strategies to maximize the even simple and routine life experiences such as eating or walking. People who habitually savor are happier and more optimistic, and less depressed than those who do not savor.

Counting Blessings. Across the lifespan, people who are grateful have

higher life satisfaction. Gratitude can be added to someone's repertoire by deliberate intervention, and a number of studies by different research groups have shown the effectiveness of so doing for both children and adults by increasing life satisfaction and optimism and reducing depression. Participants are usually asked to write down at the end of the day things that went well during the day and for which they are grateful. Details of this intervention may vary across the number of things listed and the frequency of listing them. This exercise can be modified depending on client age and situation. To avoid making this exercise burdensome to some people, especially children, it could be a few times a week and count one or two blessings instead of three blessings. Although more evidence is needed for the effect of different frequencies and amount of blessings on the outcome across different age groups, a study reported the negative effect when this exercise becomes excessive and as a result a burden.

In the authors' own work, they have asked participants to write down why each good thing happened, encouraging them to be more mindful.

> Every night for one week, set aside 10 minutes before you go to bed. Use that time to write down three things that went really well on that day and why they went well. You may use a journal or your computer to write about the events, but it is important that you have a physical record of what you wrote. It is not enough to do this exercise in your head. The three things you list can be relatively small in importance or relatively large in importance. Next to each positive event in your list, answer the question, "Why did this good thing happen?"

Participants, specifically those who continue to do this exercise, reported more happiness and less depression. The effect was maintained even in a 6-month follow-up. The mechanisms for these effects have not been closely studied, but there are likely several of them. This exercise could help people to pay more attention to the positive aspects of ongoing life and appreciate life more and be more optimistic and experience more positive emotion.

Strengths-Based Approach: Identifying and Using Strengths of Character in Novel Ways. Our multidimensional character strengths measures can be scored within the person (e.g., rank ordered)—to identify an individual's "signature strengths" relative to his or her other strengths. Everybody has strengths regardless of where they may stand relative to others. This strengths-based approach is particularly useful for working with people with a history of disability, low achievement or low self-esteem. They often have a hard time finding anything they are good at. However, if the 24 strengths *within* a person are compared, those strengths that are stronger than others can be identified. Subsequently, professionals can help them to use these strengths in their lives.

After clients identify top character strengths using online questionnaires (VIA-IS), they are encouraged to use them in novel ways in their daily life:

Every day for the next seven days use one of your top five strengths in a way that you have not before. You might use your strength in a new setting or with a new person. It's your choice.

Table 33.13–2.
Positive Psychology Techniques and Interventions

Exercises to increase positive feelings
Performing acts of kindness
Savoring
Writing a gratitude letter
Exercises to decrease negative feelings
Turning one's head to see the positive
Positive reframing
Exercises to increase life satisfaction
Counting one's blessings
Exercises to develop talents and strengths
Using talents and/or signature strengths of character in novel ways
Exercises to increase engagement
Finding a challenging hobby
Exercises to increase social connectedness
Being a good teammate
Active-constructive responding
Exercises to increase meaning and purpose
Performing secret good deeds
Writing one's own legacy
Working for a valued institution

Our research found that this exercise reduces depression and increases happiness and that the effect is evident even at a 6-month follow-up, so long as people continue to do the exercise.

Once individuals build their confidence by using their signature strengths, they can be taught how to use these strengths to work on weaknesses or less-developed strengths. If discussions and interventions start with the strengths of clients—things at which they are good—this can build rapport and increase motivation.

Research support is still accumulating. Enough outcome studies have been conducted to conclude that strengths-based approaches to change are more than just promising. Not known in most cases is how these expanded therapies fare in direct comparison to traditional treatments for problems and what are the mechanisms that are involved. Perhaps, attention to both strengths and weaknesses is critical and that no useful purpose is served by regarding these as mutually exclusive therapeutic goals.

See Table 33.13–2 for other examples.

Outcome research shows that a variety of psychotherapies are effective in alleviating problems and usually equally so, despite the different forms they take. One interpretation of the equal effectiveness of different therapies is that nonspecific factors common to all treatments are responsible. Perhaps the sorts of strategies being studied by positive

psychologists reflect these common factors and give names to them. Strategies like instilling hope and building strengths may be the critical in the effectiveness of any therapy.

Some qualifications are in order if these techniques are used in the context of treatment. First, they are not therapies *per se* but simply exercises to be deployed as part of an overall treatment program. The therapist must ascertain a client's readiness to change in the particular ways requested in the exercise as well as the client's capacity to make the change. Like any psychotherapeutic procedure, these techniques cannot be imposed on the unwilling or the unable.

Second, none of these techniques is akin to a crash diet or an antibiotic. To the degree that they have lasting effects, it is because clients integrate them into their regular behavioral routines. Counting blessings for a week will make a person happier for that week, but only if the person becomes habitually grateful will there be a more enduring effect. Research finds—not surprisingly—that the people who showed lasting benefits were those who continued to use the exercise.

Third, these exercises are typically presented as one-size-fits-all, but there is no reason to think that they are equally useful for all clients. Nothing is known about the match of an exercise with a client's particular presenting problems or goals or with a client's age, gender, social class, or ethnicity.

Fourth, little is known about the parameters of these interventions. For example, how many blessings should one count, and how frequently should this be done? With college students, counting blessings three times a week may be more effective in increasing happiness than counting them more frequently. Is this a general phenomenon or one specific to young adults attending college?

Fifth, all interventions run the risk of unintended harm, and while positive psychologists would like to believe that their techniques avoid iatrogenic effects, this assertion cannot be made with thorough confidence. For example, although optimism is related to mental and physical health, it would be simplistic and potentially hazardous to tell clients that positive expectations will solve all their difficulties. Along these lines, if a positive psychology intervention overemphasizes a client's choice and responsibility, considerable damage could be done in cases of abuse and victimization, where self-blame needs to be undone and certainly not encouraged. Interventions based on positive psychology should not preclude the use of existing therapeutic strategies when these are indicated.

Positive Psychotherapies

Positive psychotherapies are beginning to appear: therapeutic interventions based on the theories and findings of positive psychology. What distinguishes these emerging positive psychotherapies from

conventional treatments is that their stated goal is not just symptom reduction or relief but also enhanced happiness, life satisfaction, fulfillment, productivity, and the like—one or more components of positive psychology's vision of the good life. These new therapies target people with psychological problems as well as those without (see Fig. 33.13–1). In the latter case, positive psychotherapies make contact with life coaching.

However, there is a danger in the label because it may give the misleading impression that an altogether different approach to treatment has been created. Almost all treatments labeled as positive psychotherapy are derived from well-established cognitive–behavioral approaches. The authors prefer to describe these interventions as *informed* by positive psychology and to stress the continuity between them and existing approaches.

In identifying positive psychotherapies, definition is crucial but difficult. As noted, positive psychotherapies have the goal of helping people thrive. They attempt more than the solution, eradication, or mitigation of problems. The possible field of positive psychotherapies is so broad that it needs narrowing, and somewhat arbitrarily, the focus here is on approaches characterized by an explicit *therapeutic alliance* between the positive psychologist and the client. This feature goes by many names and has been variously defined, but its recurring themes include collaboration between therapist and client, an affective bond between them, and agreement on the goals and tasks of the intervention. Asking people to write about their goals or to perform acts of kindness, despite beneficial consequences, may or may not be instances of positive psychotherapy; what matters is the relational context of the request.

This more narrow definition of positive psychotherapy is met by several approaches, most of which are recent. These are described in the next section. In most cases, there is only limited evidence for their effectiveness. This is not to say that the evidence argues against their use but simply that not enough research has been done. The good news is that most of these positive psychotherapies have been devised by research-minded psychologists, who are in the process of putting their ideas to empirical test.

The further good news is that there are decades of available lessons from the psychotherapy outcome research literature to guide the evaluation efforts of positive therapists. The need for well-characterized samples, plausible comparison groups, random assignment, treatment manuals, objective outcome measures, effect size estimates, long-term follow-ups, and analyses of process is recognized. All are readily applicable to the investigation of positive psychotherapy effectiveness.

In the meantime, investigation of the effectiveness of conventional therapies will continue, and it would be desirable for researchers to add positive psychology measures to their assessment batteries. These

measures include a number that are brief yet well-validated. Although the inclusion of these measures would not transform conventional therapies into positive ones, they would provide important information for practitioners of all stripes. Do therapies that effectively treat anxiety or depression return clients to a typical level of life satisfaction or still leave them short? Might they in some cases lead to actual flourishing? Are the individual differences of concern to positive psychology moderators or mediators of traditional therapy effectiveness? One simply does not know, but these are easily answered questions given the availability of appropriate assessment techniques.

Positive Psychotherapy Examples. Here are descriptions of some of interventions that can be considered positive psychotherapies and—if the information is available—what is known about their effectiveness. As noted, positive psychology is a perspective, and no single theory underlies this perspective. When positive psychologists propose full-blown approaches to psychotherapy consistent with the premises of positive psychology, they turn to existing therapies for the overall model of treatment.

1. Decades ago, Michael Fordyce developed a *Personal Happiness Program* and showed it to be effective in boosting the long-term satisfaction of college students. Fordyce surveyed the research literature on happiness and identified predictors presumably under the short-term control of ordinary people—e.g., keeping busy, socializing with others, doing meaningful work, and making the pursuit of happiness a priority. This information was conveyed to individuals along with suggestions to carry out 14 different behavioral and cognitive exercises over the course of a semester. In seven studies that included no-treatment comparison groups, Fordyce found that self-reported happiness increased and feelings of depression decreased. Follow-up was as long as 18 months. Fordyce believed that mere exposure to what he called "happiness principles" drove these effects, but one can suspect that the exercises were crucial, as well their encouragement of positive emotions.
2. Already described were investigations of specific *positive psychology techniques* like counting one's blessings and using signature strengths in novel ways. These studies were part of a deliberate program to develop and evaluate a positive psychotherapy. Much like Fordyce, Seligman and his colleagues surveyed the scientific and popular literature for interventions purporting to increase happiness. More than 100 of these were identified, of which 40 were able to be manualized. Ongoing work is systematically testing the effects of these manualized interventions in randomized trials against a plausible placebo control (asking individuals to write about their early memories). The Internet is used to recruit participants, to deliver the intervention, and to collect data. So-called cybertherapy is obviously new and invites skepticism, but the only way to

evaluate what promises to be a cost-effective strategy for helping people is to do the relevant outcome studies. Two of the interventions so far evaluated—counting one's blessings and using signature strengths in novel ways—increase happiness and decrease depression through 6 months of follow-up, with effect sizes as large as those of conventional psychotherapy. A third intervention—writing and delivering a letter of gratitude to another person—had a very large effect on happiness and depression, but only for 1 month. Two other interventions—writing an essay about "you at your best" and identifying one's signature strengths (without the instruction to use them in novel ways)—did not affect well-being. As noted, research is ongoing, and it will be just as important to learn what does not work as to determine what does.

Table 33.13–3.
Idealized Sequence of Positive Psychotherapy Sessions

1. Orientation to therapy
2. Identifying strengths of character and approaches to happiness (pleasure, engagement, and meaning)
3. Using signature strengths of character in novel ways
4. Positive emotions in the past (reminiscence)
5. Undoing negative emotions in the past (forgiveness)
6. Gratitude
7. Taking stock
8. Meaning
9. Optimism and hope
10. Love and attachment
11. Family
12. Savoring
13. Giving to others
14. Integration: the full life

3. The *positive psychotherapy* sketched by Martin Seligman, Tayyab Rashid, and Acacia Park packages together different positive psychology techniques and uses them sequentially in one-to-one therapy with depressed clients (see Table 33.13–3). Out-of-session homework assignments are given, and their results are the focus of discussion in treatment. Pilot data suggest that this form of treatment reduces depression and increases life satisfaction. However, more research is needed to determine the critical techniques for therapeutic outcomes, how each technique contributes to its result, and whether the order in which different techniques are given matters.
4. *Mindfulness-based cognitive therapy* (MBCT) combines techniques of cognitive therapy with the Buddhist practice of mindfulness meditation, which encourages people to be more aware of what they are thinking, to live in the moment, and to refrain from evaluating judgments about their own thoughts. Whereas typical cognitive therapy tries to change the

content of thought, MBCT emphasizes nonjudgmental awareness. MBCT grew out of stress-reduction strategies developed by Jon Kabat-Zinn to help people with physical problems like chronic pain and hypertension. Zindel Segal, John Teasdale, and Mark Williams adapted these strategies to help individuals with recurring depression. MBCT entails weekly group sessions and homework assignments. Outcome studies show the effectiveness of MBCT in decreasing depression and anxiety. MBCT may also be more effective than conventional pharmacologic or psychotherapy treatment in reducing relapse rate. MBCT also improves the quality of life among medical patients, which makes it an instance of positive psychotherapy.

5. The *Penn Resiliency Program* (PRP) was developed by Jane Gillham and Karen Reivich to prevent depression by teaching children cognitive–behavioral skills before the development of problems. Its most-studied version is a 12-session program designed for group administration to middle-school students. PRP contains two main components, one cognitive and the other based on social problem–solving techniques. In the cognitive component, core cognitive techniques are translated, through the use of cartoons and skits, into a language that adolescents can understand and apply to their own lives. In the social problem–solving component, students learn skills that help them better interact with this world: assertiveness, negotiation, relaxation, avoidance of procrastination, social skills, decision-making, and problem-solving. PRP has been successfully evaluated in school and managed care settings in both the United States and China. Results indicated that prevention participants report fewer depressive symptoms and are less likely to report symptoms in the moderate to severe range through 2 years of follow-up. Researchers are in the process of investigating the effects of PRP on life satisfaction, which means that it can be construed as a positive psychotherapy.
6. *Action and commitment therapy* (ACT), developed by Steven C. Hayes and colleagues, has not been explicitly discussed by its originators as a positive psychotherapy, although its goals are unambiguously positive: to reduce psychological suffering of clients and to help them live a life of value. ACT is a short-term structured therapy that has been used with individuals and groups with a variety of psychological and medical problems. It provides exercises and assignments that encourage an individual to accept psychological pain as normal and indeed important but also to avoid increasing it unnecessarily. ACT takes issue with the premise of typical cognitive therapies that the direct targeting of problematic ways of thinking is helpful. Indeed, the proponents of ACT argue that too much attention to thinking is counterproductive and may even exacerbate problems. ACT distinguishes between language and experience and teaches clients mindfulness techniques that identify

experience without the entanglement of language. Cognitive defusion is the term used for learning to regard thoughts as mere thoughts. ACT also helps clients to identify their core values, to commit to them, and to take action to live in accordance with them. ACT has a Buddhist flavor acknowledged by its originator, but this therapy was derived from behavior therapy and an explicit psychological theory of thought and language. Of all the positive psychotherapies discussed here, ACT has been among the most frequently tested, and it is as effective in reducing psychological problems as conventional cognitive–behavioral therapies. Most outcome research to date has used pathology measures, although some studies have also used the well-being measures of interest to positive psychology.

7. C. Rick Snyder's *hope therapy* uses cognitive–behavioral techniques to increase hope through a process of setting and monitoring explicit goals. It consists of formal and informal strategies for building and enhancing hope, among children and adults. Versions of hope therapy exist for individuals and groups, and several studies demonstrate that by enhancing hope, hope therapy increases well-being and decreases symptoms of depression and anxiety.
8. *Well-being therapy* was developed by Giovanni Fava and is based on Ryff's model of psychological well-being. As described earlier, Ryff's model distinguishes several components of doing well and overlaps considerably with the positive psychology vision of psychological health. Well-being therapy is a short-term individual treatment that takes a thoroughly cognitive approach. In early sessions, clients are asked to identify and to situate episodes of doing well. Then they are asked to keep a diary or journal about such episodes and to identify the thoughts and beliefs that interrupt them. Finally, they are taught to challenge and to test these beliefs. Special attention is paid to the components of well-being that are frequently impaired, which makes well-being therapy more remedial than other positive psychotherapies. Well-being therapy is effective in reducing anxiety and depression, and it enhances well-being.
9. *Positive behavioral support* is a form of behavior therapy that focuses on the environment and specifically on problem contexts, defined as counterproductive and unfair situations that produce or exacerbate difficulties for people. Originally devised by Jacki Anderson, Edward Carr, Glen Dunlap, Robert Horner, Robert Koegel, Wayne Sailor, and Ann Turnbull to help people with disabilities, positive behavioral support has more general applicability. It attempts to impart to clients skills, coping strategies, and the motivation to change. It also enlists the support of family members and other key figures in the person's life. Positive behavioral support becomes an instance of positive psychotherapy when it expands its attention from what produces

problems to what promotes happiness and life satisfaction. Like positive psychology per se, positive behavioral support is based on the premise that happiness is not simply the absence of conditions that produce problems.

ɔ. *Quality-of-life therapy* was developed by Michael Frisch, who presented it as an instance of cognitive therapy. Its goal is to increase one's quality of life (happiness) by using the sorts of techniques developed by Aaron Beck and colleagues to decrease depression. Like other positive psychotherapies, quality-of-life therapy assumes that there are multiple routes to the good life, and it describes specific techniques for achieving these, from changing one's objective situation to revising one's standards for defining personal success. Quality-of-life therapy can be used with individual clients or groups. Different versions of quality-of-life therapy have been created for clinical versus nonclinical clients, and evidence is starting to accumulate that quality-of-life therapy is successful in its intended goal of increasing life satisfaction.

Taking Stock. The unique and explicit goal of emerging positive psychotherapies is to enhance well-being and to promote the good life among those with obvious psychological problems as well as those without. They are also similar to more established therapies. Positive psychotherapies are short-term structured interventions for individuals or small groups. Most can be placed in the cognitive–behavioral realm, although their techniques could be integrated easily into other treatment models. Most positive psychotherapies entail out-of-sessions exercises and homework assignments, the results of which are discussed in sessions. A number of positive psychotherapies ask clients to keep journals, and many of these therapies rely on ongoing assessment.

Like other cognitive–behavioral interventions, positive psychotherapies take issue with assumptions of the medical model that people in treatment are ill and that their problems are best described as discrete (present-or-absent) entities *à la* DSM. According to positive psychology, people's weaknesses and strengths exist in degrees.

As emphasized, research support is still accumulating. Enough outcome studies have been conducted to conclude that positive psychotherapies are more than just promising, with effect sizes in the small to moderate range typical of psychological interventions. Not known in most cases is how positive psychotherapies fare in direct comparison to conventional treatments for anxiety or depression. As already mentioned, the boundary conditions of effective positive psychotherapy are unknown.

In the past, it was thought that positive psychology interventions were inherently light-handed, meaning that they are met with little resistance and that they are self-sustaining. Why would someone *not* want to be happier and more fulfilled, especially given the additional—presumably

rewarding—benefits of so being?

Current thinking is that positive interventions can challenge both the client and the therapist. People do not always do the right or good (i.e., happy) thing, even when they know perfectly well what is to be done. Challenging one's skills is more exhilarating than watching television. Spending time with close friends is more fulfilling than surfing the Internet. Doing volunteer work pays more dividends than reading the tabloids. Living in accord with explicit values is more satisfying than constant compromising. So why do people live as they do?

This question is among the most crucial for positive psychologists to answer, especially as they develop and deploy positive psychotherapies. The not-so-good life has considerable inertia. In the case of happiness, theorists have made sense of this inertia by positing a hedonic set-point. This set-point may reflect genetic constraints, cognitive styles like defensive pessimism, or habits sustained by given environments. Although recent research shows that the hedonic set-point can be changed by major life events, nothing suggests that change for the better is typical or easy.

CONCLUSIONS

Positive psychology goes beyond a focus on problems and their absence to reflect health and well-being. Positive psychology focuses on identifying and capitalizing strengths and capacities. Evidence is accumulating that positive constructs such as positive emotions, good relationships, meaning and purpose, character strengths, and services for others play important roles in health and well-being, not only as broad protective factors, preventing or mitigating pathology and problems, but also as enabling conditions that promote thriving and flourishing life.

The good life can be taught and learned, but it requires hard work. Psychologists, at least in the United States, have long believed that the human condition can be improved by the intelligent application of what they have learned. Positive psychologists are no exception, and many have turned their attention to interventions that make people more happy, hopeful, virtuous, productive, and socially involved. In some cases, these applications are running in front of data that would support them, but in other cases, outcome research has been done. Even the most compelling research is not based on follow-up that extends beyond a few years, and the research participants are usually motivated and willing volunteers. How well these interventions will generalize—across diverse people and over time—is therefore a research topic of high priority.

It is difficult to predict the future of positive psychotherapies except to say that their success will hinge on the relevant research. Although many therapy techniques informed by positive psychology seem promising, so far only a few of them have been tested rigorously. More work needs to be done.

One point of view is that a thoroughly successful positive psychology will result in the fading away of the perspective, leaving a balanced psychology, one that recognizes the positive and the negative and of course their interplay. As emphasized, positive psychology does not deny the negative, and it may well be that what is most troubling in life can set the stage for what is most fulfilling. Complex emotional experiences often blend the positive and negative; optimism is most apparent when people confront setbacks and failures; crisis reveals strengths of character; ongoing challenge is a prerequisite to experience flow in the moment and to achieve something important in a lifetime; relationship success is foreshadowed not by the absence of problems but by how people resolve those that arise.

All people want to live happy and fulfilling life. It is a fundamental human desire and right. Positive psychology provides a valuable perspective on the good life as well as strategies for assessment, treatment, prevention, maintenance, and promotion. Positive psychology expands the view of psychological health beyond the absence of symptoms and disorders and provides hope that a healthy, fulfilled, and productive life is possible for all.

REFERENCES

Bryant FB, Veroff J. *Savoring: A New Model of Positive Experience*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2007.

Conklin LR, Strunk DR, Fazio RH. Attitude formation in depression: evidence for deficits in forming positive attitudes. *J Behav Ther Exp Psychiat*. 2009;40:120–126.

Gable SL, Reis HT, Impett EA, Asher ER. What do you do when things go right? The intrapersonal and interpersonal benefits of sharing positive events. *J Pers Soc Psychol*. 2004;87(2):228–245.

Jahoda M. *Current Concepts of Positive Mental Health*. New York: Basic Books; 1958.

Jeste DV, Palmer BW, Rettew DC, Boardman S. Positive psychiatry: its time has come. *J clin psychiatry*. 2015;76:675–683.

*Joseph S (ed.). *Positive Psychology in Practice: Promoting Human Flourishing in Work, Health, Education, and Everyday Life*. 2nd ed. New York: John Wiley & Sons; 2015.

Park N, Peterson C. Methodological issues in positive psychology and the assessment of character strengths. In: Ong AD, van Dulmen M, eds. *Handbook of Methods in Positive Psychology*. New York: Oxford University Press; 2006.

Park N, Peterson C. Positive psychology and character strengths: its application for strength-based school counseling. *J Prof Sch Couns*. 2008;12:85–92.

*Peterson C. *A Primer in Positive Psychology*. New York: Oxford University Press; 2006.

*Peterson C, Seligman MEP. *Character Strengths and Virtues: A Classification and Handbook*. Washington, DC: American Psychological Association/New York: Oxford University Press; 2004.

Quoidbach J, Mikolajczak M, Gross JJ. Positive interventions: an emotion regulation perspective.

Psychol bull. 2015;141:655–693.

Ryff CD, Heller AS, Schaefer SM, Van Reekum C, Davidson RJ. Purposeful engagement healthy aging, and the brain. *Curr Behav Neurosci Rep*. 2016;3(4):318–327.

Sansone RA, Sansone LA. Gratitude and well being: the benefits of appreciation. Psychiatry *(1550-5952)*. 2010;7(11):18–22.

Seligman ME, Csikszentmihalyi M. Positive psychology: an introduction. *Am Psychol*. 2002;55:5–14.

*Seligman ME, Steen TA, Park N, Peterson C. Positive psychology progress: empirical validation of interventions. *Am Psychol*. 2005;60:410–421.

*Snyder CR, Lopez SJ. *Handbook of Positive Psychology*. 2nd ed. New York: Oxford University Press; 2011.

Werner-Seidler A, Tan L, Dalgleish T. The vicissitudes of positive autobiographical recollection as an emotion regulation strategy in depression. *Clin Psychol Sci*. In Press 2017.

▲ 33.14 Psychodrama

EDWARD J. SCHREIBER, T.E.P., D.D.

PSYCHODRAMA, SOCIODRAMA, SOCIOMETRY, AND SOCIATRY

The biological picture of an individual places the psyche within the body (as an epiphenomenon). In the sociometric [and psychodramatic] picture of the individual (person), the psyche appears as outside of the body, the body is surrounded by the psyche, and the psyche is surrounded by, and interwoven, into the social and cultural atom.

—J. L. Moreno, *Sociometry and the Cultural Order*

Psychiatrist Jacob Levy Moreno (1889–1974) hypothesized that human beings learn primarily through interaction with relationship systems, which he called social atoms. Moreno believed that an encounter of the self with creative processes, which he defined as spontaneity-creativity, could deepen and encourage growth for individuals and groups. Psychodrama and related action methods of sociodrama and sociometry facilitate creativity, and the discovery of strengths, based upon J. L. Moreno's understanding of human growth, from infancy to the end of life, with spontaneity-creativity as a central factor.

Psychodrama is used as therapy, in various forms of education, and in training therapists, doctors, lawyers, teachers, social workers, political leaders, public speakers, businesses and organizations, and others. Psychodrama can often be used in conjunction with other forms of therapy.

In a psychodrama, one person, called the protagonist, works with the director and a group, or just individually with a director. In a group, participants may be asked to assume a role in the protagonist's exploration of an aspect of life. Those asked are not obligated to enact any role; it is a matter of choice. The director's role is complex and includes the facilitation of the protagonist, the role players, the group, the multi-phased psychodrama process and techniques, and the integration and sharing phase with protagonist and group. Psychodrama follows a sequence of phases, which includes the *warm up, action, and sharing*. In a training program there is often a *processing* phase for students.

Sociodrama is related to psychodrama, and is a process for a group, organization, or community to explore relevant concerns and issues. In sociodrama, the whole group is on the stage in a role-playing theme-, idea- and setting-based encounter in which participants have taken a role to play in the sociodrama. Sociodrama is used in educational settings, in professional development programs for a wide range of professions, and in clinical settings.

Sociometry, related to psychodrama and sociodrama, although a method of its own, is used to understand and evaluate, and to support, enhance, change, and develop relationship systems, sub-groups in organizations, communities, and entire organizations. Sociometry is a

tested process to address and improve communication, cohesion, functioning, and growth. Sociometry is rich with techniques to measure and examine the often unseen—but felt—networks very much alive in impact within a group and its members.

Sociatry (healing society) utilizes psychodrama, sociodrama, and sociometry to study the forces J. L. Moreno described as impacting all of humankind. J. L. Moreno understood there to be *an actual, underlying, dynamic, central structure impacting all formal and informal groups.* This underlying structure may be seen in groups looking through a sociometric lens, known within sociatry as the social microscope. The social microscope is an educational tool, used to show the forces J. L. Moreno named as the sociodynamic effect and the social and organic unity of mankind. Sociatry deepens relationship to the creative process and mankind's common interrelatedness.

HISTORY

J. L. Moreno was born in Bucharest, Romania and received his MD in Vienna in 1917. In Vienna he began the House of Encounter and formulated ideas about the nature of groups, healing, and society. In 1925, he immigrated to the United States, and in 1931, he inaugurated the Impromptu Theater in New York City and sociometric research at Sing Sing Prison, the Hudson School for Girls, and public schools in New York. In 1936, Moreno opened the Beacon Hill Sanatorium and the psychodrama theater in Beacon, New York. In 1954, he organized the First International Congress of Group Psychotherapy in Toronto, Canada. He is credited with many books, journals, and articles, along with the organizations he began, and his worldwide teaching and training.

J. L. Moreno was a doctor, scientist, psychiatrist, and mystic. To note his mysticism causes alarm for some, but is important in grasping the fullness of his ideas. It is from J. L. Moreno's integration of mysticism and science that this work has come. As a young man J. L. Moreno was fascinated with the cosmos, fixated with the realm of the unknown. He dedicated his life and work to understanding and mapping the connections people make and have in life. J. L. Moreno created dozens of techniques, a complex method to experience and to integrate this creative force—a kind of intelligence as he found it. He taught and trained thousands of practitioners in this method; how the force of spontaneity and creativity can be found; how to touch it, to the benefit of oneself and others. He was also a social scientist and put enormous effort into the investigation and cure of social problems. His goal was ultimately to realize an organic unity within humankind, which he notes in *Who Shall Survive?*

A truly therapeutic procedure cannot have less an objective than the whole of humankind. But no adequate therapy can be prescribed as long as humankind is not a unity in some fashion and as long as its organization is

unknown. It helped in the beginning to think, although there was no definite proof, that humankind was a social and organic unity.

Zerka Toeman Moreno began her study and work with J. L. Moreno in 1941 after having emigrated from England to the United States. Her path to finding J. L. Moreno and the Beacon, NY Hospital is a remarkable story written in her memoirs, *To Dream Again*. Zerka joined J. L. Moreno as his assistant, then as his wife and primary collaborator. J. L. Moreno's work might be just a footnote in history, were it not for the dedication, insight, and skills of Zerka Moreno. During the long period of their developing the method, Zerka emerged as co-developer, went on to teach thousands of students throughout the world, and helped heal many wounded souls.

Zerka Moreno co-authored and edited most of J. L. Moreno's books and articles and served as his primary teacher of the method. After his death, Zerka Moreno continued to teach and to develop psychodrama, and authored articles, chapters, and books, with *To Dream Again—A Memoir* as her last, with this chapter's author as editor.

J. L. Moreno and Zerka Moreno were world-minded. The issues of access to resources and the attention to those left out were of great importance to their work. J. L. Moreno created a therapy of experience, an experience of qualities of being human, to bring understanding and regeneration to individuals and groups.

The method created by J. L. and Zerka Moreno helps many problems. Work is being done throughout the world to bring the method and its practices to assist with the widespread issues that humans face as whole human family. The International Association of Group Psychotherapy and Process (IAGP) Psychodrama Division is a source of information regarding social applications.

METHODS, TECHNIQUES, CLINICAL ISSUES

The method includes psychodrama, sociodrama, sociometry, group psychotherapy, spontaneity testing and training, role testing and training, the therapeutic stage, the empty chair, role playing, role models, role relationships, and the exploration of subgroups and systems in organizations, groups, and communities. The method involves sociometry, surplus reality, tele, the cultural conserve, the social atom, the protagonist, the double, auxiliary egos, the canon of creativity, the director, role theory, role reversal, and a developmental theory based on spontaneity-creativity. There are aspects of the psychodrama method that are not as widely known, such as the cultural atom, the sociodynamic effect, sociogenetic law, sociostasis, socio-gravitation, the autonomous healing center, and sociatry.

Psychodrama

Psychodrama takes place on a stage. A stage can be like the three-level stage J. L. Moreno designed (relocated from the original hospital in Beacon, NY to Boughton Place in Highland, NY) or any space designated as a stage separated from the group. The stage is the location in which the psychodrama and sociodrama takes place. There is flexibility in the stage; the director and protagonist may adjust and change the stage when needed.

The protagonist works with the director in a group, or just individually with a director. In a group, participants may be asked to assume a role in the protagonist's exploration of an aspect of his or her life. The director works with the group and group process during psychodrama, although the focus of the group is on the psychodrama and the protagonist. It requires a combination of techniques and experiential methods to be complete. In psychodrama, a protagonist may examine aspects of life from the past, the present, or the future. The protagonist can explore a relationship, a moment in time, an interaction, a setting, or a scene from life.

According to Zerka Moreno, psychodrama is not so much a method of catharsis of abreaction, as of catharsis of integration—integration of what happened, what was needed but did not happen, what is needed now, or in the future, and how this is to be healed in some way. Integration may take place during the psychodrama or days later, and may include a new defining and developing of skills, relationships, and roles.

An initial task of the director is to discern the general frame and need the protagonist brings, and to determine a place (scene, setting, situation) to begin the psychodrama. There is a warm-up phase in which the protagonist emerges, and also a warm up between the director and the protagonist. The director may suggest an initial scene, often elicited during an interview with the protagonist at the beginning of the psychodrama. A psychodrama may commence with scene setting in which the protagonist creates in action and imagination a setting, a scene, or simply wants the empty stage. There is no requirement to begin with the past. Zerka Moreno suggests that what matters is not so much where the psychodrama begins, *but that it begin*. Psychodrama develops a protagonist's capacities and strengths in their work to create, and to recreate, in an authentic and adequate way.

The double technique in psychodrama is when a member of the group (who is guided by the director) offers statements and actions of support or stabilization, clarification, and amplification to the protagonist during the psychodrama. The double is role-played by someone from the group chosen by the protagonist or the director. The double can add depth, expand expression, extend, or clarify statements or feelings, as well as help the protagonist to gain perspective and integrate, hold, and contain the experiential-based learning.

The mirror technique is when the protagonist is outside of a scene, as though watching or seeing a scene from their life presented by others. A

group member may take the place of the protagonist in the scene, and members may be asked to assume other roles. The protagonist gains an observation, perspective, and emotional distance from the events.

Role reversal is a primary technique in psychodrama. It is a directed experience and, as with other techniques, can be used as therapy, in education, and for training programs. In role reversal, the protagonist presents what needs to be presented to another person, imagined, in a setting or scene on the stage in the psychodrama. The protagonist offers to "that other" what needs to be said. With direction, and when appropriate, the protagonist "takes the role" of that other. The protagonist moves to be in the role of that other, and to eventually respond from the role of the other. Role reversal can continue for a series of directed exchanges, but the protagonist finishes in his or her own space, place, and role as his or her self, with a period of sharing after the psychodrama has been finished. Role reversal may take place with an imagined person, an object, an idea, a place in time, etc. A protagonist may role-reverse with lessons in life, or, as J. L. Moreno described, with creation.

Role reversal expands creativity, though there are contradictions. If there is difficulty with role reversal, this may indicate that other techniques such as the double, the soliloquy, or the mirror may be more useful and appropriate. One needs to have strong and secure boundaries to return to. There is a contraindication and caution for any role reversal with trauma, perpetrators, or violence of any kind. Training is required to facilitate role reversal or any aspect of psychodrama and related methods.

Sociodrama

In psychodrama, one person is the protagonist; in sociodrama, the whole group participates in the action learning. In sociodrama, the group, organization, or community explores concerns, problems, situations, or issues relevant. Sociodrama is used in education, for professional development, and in clinical settings. The group is role-playing within a theme, exploring issues in an imagined setting. Refer to the reference list at the end of the chapter for how themes, issues, concerns, and settings are designed and chosen.

Sociodrama has been used to address global warming; with exchanges between people of differing cultures and religions; with survivors of the Holocaust on both sides; in working with issues and differences between the generations; and between police and gangs, etc. For the last several years a group of practitioners have on one same day, to offer Sociodrama Conference, which takes place every few years, sponsored in association with the International Association of Group Psychotherapy and Group Process (IAGP). Sociodrama develops a group in their capacity to double and reverse roles with others, to see the world from the point of view of another, to find new meaning and learning.

Sociodrama is used for training doctors to better understand the experience of being a patient, and parents to better know the experience of being a parent. Sociodrama increases communication and cohesion and *tele* (two-way empathy). Sociodrama includes techniques and directorial and group processes for how a sociodrama is directed, and how the concerns, themes, roles, and settings are chosen.

Psychodrama and sociodrama are experientially woven applications of the social atom, encounter, the canon of creativity, role theory, concretization, sociometry, catharsis of integration, group psychotherapy, and other techniques, which will be explained further in this chapter. It is human relationships in life that can be concretized and explored through psychodrama, sociodrama, and sociometry.

Sociometry and Tele

Moreno hypothesized that human society has an atomic structure, which corresponds to the atomic structure of matter, and that its existence can be brought to an empirical test by social microscopy, related to sociometry. Sociometry is a method for the examination, exploration, and alteration of the connections and dynamics taking place within a group, organization, or community. The connections are often felt but often not fully recognized. Sociometry makes the connections visible, approachable to growth and change, to increase communication, creativity, cohesion, and resolution.

Sociometry is a method with many instruments for the measurement of connection, the web of tele "meetings" in operation. Telic connections can impact the functioning of a group or organization and individuals within it, and the dynamics of cohesion, inclusion, exclusion, subgroups, etc. Telic constellations form between two or more people often unrealized in their capacity to harness and unleash a kind of inter-feeling that impacts a group's embrace of its members, its creativity, its functioning, cohesion, and learning. Tele is a phenomenon below transference not always seen, with multiple types of often interlocking relationships and groupings in an organization and community.

Telic constellations are made visible with sociometry. Sociometric structures and configurations can regenerate to increase communication, cohesion, creativity, and functioning. Sociometry is used in clinical settings, in training programs, with organizations, and in education.

Zerka Moreno explains tele:

> Empathy and transference are part of a more elementary and more inclusive process of tele. The process of reciprocation does not enter into the meaning of empathy, and transference is considered the factor responsible for dissociation and disintegration of social groups. Tele is the factor responsible for the increased mutuality of choices surpassing chance possibility and responsible for the increased rate of interaction among members of a group.

Telic structures are central to one's personal sociometry, noticed in the

feeling between oneself and others. For personal and professional development, sociometry evaluates and measures this. With a group, sociometry may knit positive telic connection and communication, and to address negative, ambivalent, non-mutual connections. Sociometry can unclog blocked communication by revealing to the group its telic structures. When the structures are more fully understood, they can be changed and developed in a productive way. Telic structures often operate below the formal organization as a matrix of cohesion or not, at a moment in time. Sociometry shows a group the living social atom, the dynamic telic processes commonly unseen but often felt, and the intelligence of life called spontaneity-creativity in action.

THEORIES THAT INFORM THE METHOD

The Social Atom

The social atom is used to depict the relationships of one's life and the nature of each relationship. The social atom can be an instrument used in a written format, or else a psychodrama enactment may depict a social atom by bringing it alive with others taking roles. The social atom can include one's relationship with individuals, groups, animals, communities, etc. The social atom may depict the past, the present, or the needed or hoped for relationships. It is a visual representation of positive, negative, neutral, ambivalent, mutual, and non-mutual relationships.

The social atom depicts those with whom one feels emotionally impacted, connected, or related. The social atom is a relationship map; it brings awareness to relationships. The social atom looks like a bull's-eye; the nucleus represents oneself. In the first circle beyond oneself, one notes those with whom there is such a significant impact. One draws a line connecting the self to that other, and there are specific line notations to show the nature of the relationship. Beyond that, in the next circle, one depicts those with impact best depicted at that level, and so on. *The social atom is a depiction of one's own truth and perception*. The spatial distance from the center to others reflects the perception of the nature and impact of those persons. A menu of symbols and lines are given that denote male and female, positive and negative, neutral, ambivalent, mutual, and one-way relationships.

There is flexibility with this tool, which provides information and insight and points to areas of further exploration. J. L. Moreno believed human are as closely tied to their social atoms as they are to their bodies.

Encounter

Encounter represents a theoretical and practical underpinning to psychodrama. Encounter is an action-instrument, a basic idea, for an existential dialogue with another, imagined or real. With the encounter one

finds oneself in the *psychodramatic experience of me and not me*. This is an existential fiber in life for what connects us, an experience often had with psychodrama. J. L. Moreno gave a name to the mutual feeling of the essentialness of another, calling that sense of another *tele,* noted earlier in this chapter. The interweaving of tele is explored in psychodrama, sociodrama, and sociometry.

J. L. Moreno wrote about the existence of a primordial nature in life, found within humans and within groups. In *To Dream Again—A Memoir,* Zerka Moreno named this primordial nature the *autonomous healing center* within. It is generated in psychodrama to craft a new, novel, and adequate response to life.

This is the case when the protagonist explores a role reversal, in taking another's role in language, behavior, and feeling. One finds a sense of what that other might feel, and what might be learned from this. It is in the encounter of oneself with others and life that spontaneity-creativity are discovered and strengthened. Facilitating role reversal and encounter requires training; for information regarding training and certification for psychodrama, refer to the American Board of Examiners in Psychodrama, Sociometry and Group Psychotherapy. There are similar boards throughout the world.

Canon of Creativity: Formula for Creativity

The Canon of Creativity is a visual diagram to represent a theoretical underpinning to the psychodrama method. The diagram identifies stages of creativity for a person and group. The Canon of Creativity is a theory actualized and brought to life by psychodrama, and includes the warm up where group members prepare for spontaneity-creativity in action. There is then a movement to create something new, which is met at first with resistance. Once resistance is addressed, then renewal and creativity can emerge. There is the movement to form—and presence—in the world and life. The renewal of society was J. L. Moreno's full vision, bring creativity into culture for its renewal.

Human Development

Zerka Moreno describes the perspective for a series of developmental stages through the course of life in *The Quintessential Zerka*. These are the periods from one's birth to one's ending, offering guidance for tasks to be considered for each developmental stage. The developmental stages and tasks are templates for a realization of creative potential and have life tasks for growth.

Psychodrama helps to experience the tasks, even those not originally completed. To become more able to role-reverse with others, including groups of people, is one of the developmental tasks. There is a task of

completing unfinished business with one's parents, and preparing to yield parents back to the cosmos, as people eventually ready themselves. Developmental stages include the taking of new roles, finishing others, and crafting one's continued path.

It is useful to consider that, in each of the developmental phases, different psychodrama tools are highlighted as instruments to help with that phase. It is also wise, in conducting psychodrama, to have as a mental construct the notion that what the protagonists are showing could be related to the tool needed for a developmental stage to be completed.

Role Theory

Role theory suggests that the self develops and emerges through the process and actualization of roles through one's life. Roles are internal and external aspects and functions of self. Role training, learning new roles in life, works according to a three-stage process: (1) deciding to take a role or not, addressing ambivalence and resistance; (2) assuming a role and finding role models and examples of how to function with increasing skill in a role; and (3) emerging more fully as oneself and with increasing skill by creating new and authentic expressions of oneself.

The method considers human psychological and emotional growth comes through how people start, how people learn, play, live, change, and finish the roles and role relationships throughout their lives. To attend to this consciously can be supported by the methods stated here. Psychodrama can be used to explore and prepare for the roles one wants, or wants to change, in some fashion.

Surplus Reality

Surplus reality is a wide-ranging tool for a protagonist and group to explore beyond a present reality. Surplus reality uses a variety of techniques and is a foundation for psychodrama. The protagonist creates a psychodrama that includes and can go beyond known reality, using imagination, perception, creativity, and psychodramatic activity. The director helps the protagonist to explore the real or imagined, and to forge an adequate response.

Techniques include role reversal, double, mirror, soliloquy, aside, concretization, presentation of oneself in the role of another, and changing time, space, and location. Surplus reality can involve empty chairs representing people, events, or aspects of oneself, and stages of life, etc. Surplus reality can offer to the protagonist an experience as it was needed but did not happen, aspects of the past, present, or future, and roles that may be required.

Surplus reality touches upon imagination and creativity. It opens to new worlds, exploring an event, redoing an event, or exploring an event that never happened. Auxiliary egos, as they are called, are group members

asked to take a part in the protagonist's exploration. Regardless of the interaction in many ways the protagonist explores strengths, capacities, and resources along with vulnerabilities and creativity. There is a great range of choice that surplus reality offers.

Universalia of Time, Space, Reality, and the Cosmos

The psychodrama method uses time, space, reality, surplus reality, and the existential realm of the cosmos as J. L. Moreno described it, as psychodramatic tools. The past, present, and future are reflected in the here and now, making time a flexible instrument. A protagonist may explore a scene that happened in childhood, then change scenes to one imagined in the future. A sequencing of scenes and times may be explored, for example, as markers of one's life.

Psychodrama explores life as it happened, or as it was needed or wished for. The director uses psychodrama to facilitate the protagonist's exploration and repair. The connection of human beings to life relates people to an inherent dignity. Many report psychodrama to be an experiential encounter that reveals this universal dimension.

Warm Up, Action, Sharing—the Structure of Psychodrama

Psychodrama is a three-step process of warm up, action, and sharing, with processing for those in training. The warm up constitutes approximately one-third of the session, the action phase one-third, and sharing phase one-third. There are many approaches to conducting a warm up, which can be quite organic to help the group meet one another, for participants to feel present, to make connections to others, and for the director to introduce themes. The warm up is a natural tendency in human life: people have a warm up, for example, as a preparation for going to sleep, or a warm up for going to work, or for addressing difficult issues with another. This phase is natural to how humans live.

The warm-up process introduces creativity in some form to the group and individuals. A warm up may be planned or spontaneous, may include movement, conversation, reflection, art and music, etc. The warm up is preparation for the action phase. The director weaves and orients the positive telic connections, enhancing the social atom of the group.

The action phase is the psychodrama or sociodrama itself. In psychodrama a protagonist may be chosen by the group, at times self-selected, or selected by the director. The protagonist enters the stage and there is often an initial dialogue between the protagonist and the director about goals, boundaries, time frames, and elements of the psychodrama.

The director works in collaboration with the protagonist. The group remains a group, even as some group members may be asked to enact a role in the psychodrama. The action phase often starts with a scene or some

expression of present concerns. There may be a number of scenes or just one scene. There is often a peak of action and integration of the experiential learning.

Sharing is the final third of psychodrama. With the psychodrama completed, group members share with the protagonist and group about how the psychodrama was related to one's life. Those who assumed a role in the psychodrama generally share first, initially from the experience of the role, then as it relates to their life. Sharing is not advice giving or analysis; the director may share as well.

Do–Undo–Redo

Do–Undo–Redo is one option for a methodologic sequence for a psychodrama, as a strategy for change, and a structure for the action phase. The protagonist may begin by showing a condition or concern, perhaps a stepping in and out of the concern, concretized in some way on the stage. The protagonist may encounter an aspect or aspects they wish to address or change. The protagonist can undo things and, following that, redo them differently or from a different vantage point. There can be a scene followed by the redone scene in which something new has been created. There can be a reconstruction of the situation or elements in it. This is a flexible structure adjusted to meet the needs of the protagonist.

One quite beautiful application of this has been the "moments from one's life" in which the protagonist (often many protagonists) may "re-live" an experience in life, often a joyful and enlightening one.

ETHICS IN TRAINING, EDUCATION, AND PRACTICE

A challenge in conducting psychodrama is keeping the experiential structures intact while having flexibility in the application. This places a demand on the director, educator, trainer, and practitioner to evaluate and apply the psychodrama method with a solid foundation of theory, training, education, practice, and certification. The strategic thinking for directing requires training, supervision, and professional development. The American Board of Examiners in Psychodrama, Sociometry and Group Psychotherapy sets standards and examines candidates for two levels of national certification: Certified Practitioner (CP) and Trainer-Educator-Practitioner (TEP). Similar certification boards exist in South America, Australia, Europe, and Asia. The American Board has adopted the ethical standards of the American Psychological Association. Psychodrama should not be practiced without board-approved training, in accordance with the established certification guidelines.

RESEARCH AND THE SOCIAL AND ORGANIC UNITY OF HUMANITY

For the application of research in the use of psychodrama, two sources are

recommended. The American Society of Group Psychotherapy and Psychodrama (ASGPP) is the membership organization of psychodrama in the United States. There are similar organizations throughout the world, for which research information would also be available. The American Board of Examiners in Psychodrama, Sociometry and Group Psychotherapy (ABE) is the certification organization for psychodrama in the United States. Similar boards exist throughout the world.

Social and Organic Unity of Humanity

J. L. Moreno wrote extensively about society and social forces impacting humankind. He makes reference to it in his book, *Who Shall Survive?*

"The discovery that human society has an actual, dynamic, central structure underlying and determining all its peripheral and formal groupings may one day be considered as the cornerstone of all social science. This central structure—once it has been identified—is either found or discernible in every form of human society, from the most primitive to the most civilized: it is in the genesis of every type of society. In addition, it exerts a determining influence upon every sphere in which the factor of human interrelations is an active agent in economics, biology, social pathology, politics, government, and similar spheres of social action."

To see society's actual, dynamic, central structure underlying and determining all peripheral and formal groupings offers a perspective on what impacts present situations. The situation calls for a willingness to reverse roles with others and other groups of people, and to find what is common among us, to find solutions that consider all of mankind and the impacted natural world. Sociatry endeavors to understand this underlying phenomenon to human society and to create education for personal and professional development about the underlying structures impacting humankind and the ecological fabric. It is a creative path for which J. L. Moreno offers:

"A truly therapeutic procedure cannot have less an objective than the whole of humankind. But no adequate therapy can be prescribed as long as humankind is not a unity in some fashion and as long as its organization is unknown. It helped in the beginning to think, although there was no definite proof, that humankind was a social and organic unity."

Based on this, it may be an idea developed of necessity. As the whole of mankind is a unity in some fashion, tendencies emerge between different parts, drawing some apart, drawing some together. This may be advantageous for some at times and disadvantageous to others. If it is true that there is a unity among us, to humanity, biologically or otherwise, if a unity of humankind does exist, how can one find and realize it? From this, further research could be generated.

References

Blatner A. *Acting-In: Practical Applications of Psychodramatic Methods*. New York: Springer; 1996.

Blatner A. *Foundations of Psychodrama: History, Theory and Practice*. New York: Springer; 2000.

Fox J, ed. *The Essential Moreno*. New York: Springer; 1987.

Homes P, Karp M. *Psychodrama: Inspiration and Technique*. London: Routledge; 1991.

Hovatin T, Schreiber E, eds. *The Quintessential Zerka: Writings by Zerka Toeman Moreno on Psychodrama, Sociometry and Group Psychotherapy*. New York: Routledge; 2006.

Marineau R. *Jacob Levy Moreno 1889–1974: Father of Psychodrama, Sociometry and Group Psychotherapy*. London: Routledge; 1989.

Minkin R. *Sociodrama For Our Time – Sociodrama Manual*. Philadelphia, PA: East West Sociodrama Institute; 2011.

Moreno JL. *Das Stegreiftheater*. Berlin: Kiepenheuer Verlag; 1924 [Trans. *The Theater of Spontaneity*. Beacon, NY: Beacon House; 1947].

Moreno JL. *Das Testament des Vaters*. Berlin: Kiepenheuer Verlag; 1920, 1922 [Trans. *The Words of the Father*. Beacon, NY: Beacon House; 1924].

Moreno JL. *Who Shall Survive?* Beacon, NY: Beacon House; 1934.

Moreno JL. *Sociometry and the Science of Man*. Beacon, NY: Beacon House; 1956.

Moreno JL, Moreno ZT. *Psychodrama*. Vol. 2. Beacon, NY: Beacon House; 1959.

Moreno JL, Moreno ZT. *Psychodrama*. Vol. 3. Beacon, NY: Beacon House; 1969.

Moreno ZT. *Surplus Reality and the Art of Healing*. London: Routledge; 2000.

Moreno ZT. *To Dream Again – A Memoir*. New Paltz, NY: Mental Health Resources; 2012.

▲ 33.15 Evaluation of Psychotherapy

MEGAN L. PETRIK, PH.D., AND STEPHEN M. SAUNDERS, PH.D.

PSYCHOTHERAPY RESEARCH

The principal goals of psychotherapy research are to understand what the effective components of psychotherapy are and how they work, to determine how patient and therapist factors influence outcome, to improve the effectiveness of psychotherapeutic interventions, to guide the development of new therapeutic techniques and evaluate their effectiveness, and to inform public policies that increase the availability of quality mental health care.

Past controversies regarding the ability to examine scientifically various psychotherapeutic techniques have largely been settled. Valid and reliable methods for measuring therapeutic events and their effects have been developed. These have included intensive analyses of patient and therapist variables, in-depth assessment of therapeutic processes, and implementation of outcome measures that assess general distress, symptoms related to specific disorders, and functional impairment in

emotional, cognitive, and behavioral domains. Controlled clinical trials comparing replicable, distinct psychotherapeutic interventions are normative, as are sophisticated analytic methodologies, including growth curve analysis, time-series panel analysis, and structural equation modeling. Most significantly, various methods of meta-analysis—techniques that combine results across different studies to evaluate the effectiveness of particular treatments for specific patients and problems—have been applied to psychotherapy research.

This section begins by briefly describing the characteristics of psychotherapy consumers. Definitions of efficacy and effectiveness are provided, and research evidence supporting both with regard to widely used psychotherapy models and techniques are presented, followed by a review of the benefits of combining psychotherapy and pharmacotherapy. Other treatment variables reviewed include the therapeutic alliance and therapist-specific interventions. Patient and therapist variables that have been studied are discussed. Emerging themes and unanswered questions regarding psychotherapy research are considered.

CONSUMERS OF PSYCHOTHERAPY

The effectiveness of psychotherapy must be understood within the context of who receives such help and how they obtain it. Several large-scale national surveys found that two-thirds of those who make at least one mental health visit in any given year are female. Ninety percent of psychotherapy patients are European American, approximately 50 percent have at least some college education, and more than one-half are between 21 and 40 years of age.

Approximately 25 percent of the US population meet criteria for a psychiatric disorder in any given year, but most do not seek any type of formal care. The likelihood of obtaining treatment increases with emotional distress and functional impairment. Accordingly, those with comorbid mental illnesses or with severely disabling disorders, such as schizophrenia, are the most likely to seek professional help. In contrast, a lack of insurance and concerns about how to pay for services decrease the likelihood of seeking treatment. However, the primary barriers to mental health care are psychological, perhaps primarily the stigma related to having a mental illness. Negative attitudes toward admitting that one has a psychiatric problem and toward seeking care from a specialty mental health provider have actually increased over the last five decades. Research suggests that this is related to the increasing conceptualization of mental health problems as medical, which may translate into the impression that the problem is permanent. Consistent with this, most who seek care increasingly do so in the general medical sector rather than from a mental health specialist.

Mary Beth, 36-year-old mother of three, had been struggling with an alcohol use disorder for the last 3 years. Her husband and her mother both discouraged her from drinking, but they also implied that it would be a mistake to seek treatment. They insisted that if she were strong enough, she could recover without help. Mary Beth was worried that if she admitted that she had an alcohol problem and needed treatment, her husband might sue for divorce and try to take her children away from her.

EFFICACY AND EFFECTIVENESS

Psychotherapy researchers distinguish between efficacy and effectiveness research. Each has a distinct purpose, methodology, and interpretative context.

Efficacy

Efficacy studies evaluate the sufficiency of a specific treatment to reduce distress, symptoms, and impairment with a group of patients having a particular psychiatric disorder. To minimize the influence of confounding factors, efficacy studies are conducted using randomized clinical trial methodology in a controlled setting. Patients are screened to control for excessive patient variability and are randomly assigned to interventions that are being compared. The treatment under investigation might be compared to no treatment, a waiting list, placebo, minimal intervention (e.g., psychoeducation), or an alternative treatment. Therapists are trained to conduct the competing treatments to maximize the integrity of putative therapeutic ingredients. To achieve this, most efficacy studies use treatment manuals, which direct (to some extent) the therapist with regard to the intervention offered. Most studies also perform integrity checks, such as reviews of taped sessions, to see whether the therapy under study was actually implemented.

Treatment manuals have been essential in establishing that certain psychotherapeutic techniques are efficacious. Prominent examples of treatment manuals include David Barlow and Jerome Cerny's *Psychological Treatment of Panic;* Aaron Beck, John Rush, Brian Shaw, and Gary Emery's *Cognitive Therapy of Depression;* and Lisa Najavit's *Seeking Safety: A Treatment Manual for PTSD and Substance Abuse*. Other manuals have been published in journals and as chapters in books, such as in David Barlow's *Clinical Handbook of Psychological Disorders*.

Effectiveness

Effectiveness studies are concerned with whether psychotherapy delivered in actual clinical settings is effective in reducing the symptoms, distress, and dysfunction associated with mental illness. The experimental controls used in efficacy studies are absent in effectiveness studies, just as they are absent in community settings. (Indeed, most effectiveness studies are conducted in the community.) For instance, in real-world clinics, patients

with comorbid disorders are commonplace, whereas they tend to be screened out of efficacy studies. Likewise, at the typical clinic, length of treatment is determined by patient need rather than treatment prescription. Thus, the goal of effectiveness research is to describe patient responses to the typical, community-based delivery of psychotherapy.

Effectiveness studies typically assess the characteristics of settings and providers, the generalizability of the patient samples, the features of treatments that are actually delivered, barriers to treatment experienced by patients, and the average costs associated with treatment. In one general theme of research, effectiveness studies have demonstrated that manualized treatments found to be efficacious in randomized clinical trials can be effective when applied in community settings. Research further suggests that the magnitude of treatment effects in effectiveness studies is similar to that found in efficacy studies.

Effectiveness studies have demonstrated a dose–response relation between the number of sessions attended and the probability of improvement. One study found that approximately 50 percent of patients are measurably improved by their eighth session of treatment, and that by session 26, three in four patients have improved. Other effectiveness studies have found that chronically distressed patients, including those with personality disorders, improve more slowly than acutely distressed patients.

The results of effectiveness studies are sometimes seen as more relevant to actual practice than are the results of efficacy studies, but both are important to the implementation of evidence-based practice. Although stringent experimental control reduces the likelihood of confounding variables in efficacy studies, the results can have limited external validity (generalizability) because of those very controls. Effectiveness research is more likely to be conducted in actual clinical settings, but the causal implications of such research must be interpreted with caution because of the likelihood that confounding variables influenced the results.

Systematic Reviews and Meta-Analyses

Common methods to summarize and integrate efficacy and effectiveness research are systematic reviews and statistical meta-analyses. Systematic reviews apply specific strategies to assemble, critically evaluate, and integrate relevant studies that address a specific clinical question. Systematic reviews are exemplified in treatment guidelines published by the American Psychiatric Association and other organizations. Meta-analysis is a particular type of systematic review that uses statistical methods to combine and summarize the results of original studies. The individual studies are quantified and coded, so that they can be compared with one another. Typically, an effect size is a measure of treatment effect that is calculated for each study by computing the size of the difference

between the treated and untreated (or control) groups.

Hundreds of meta-analyses of psychotherapy have been conducted, and most have reached the same general conclusion: Psychotherapy is an effective intervention for psychiatric illness across diverse populations and settings. Most meta-analyses have focused on specific disorders (e.g., the efficacy of psychological treatment for bulimia nervosa) or on specific therapeutic approaches for various disorders (e.g., comparing CBT to other therapies). In general, meta-analytic studies have shown that psychological treatments are vastly more effective than no treatment (e.g., wait-list controls or minimal interventions), are about as effective as biomedical treatments for most disorders, and are about equally effective when compared to each other.

EFFICACY AND EFFECTIVENESS OF COMMON EVIDENCE-BASED TREATMENTS

Effective psychotherapy requires a therapeutic alliance between the therapist and patient, which entails the sense on the part of the patient that the therapist understands, is capable of helping, wants to help, and genuinely cares for the well-being of the patient. This emotional aspect of the alliance enables agreement between the patient and therapist on the necessary therapeutic tasks to achieve the goals of symptom reduction, alleviation of distress, and/or improving life functioning. This section will review the most commonly used and studied psychotherapeutic interventions (i.e., cognitive therapy, behavioral therapy, interpersonal therapy, and psychodynamic therapy), as well as evidence supporting these interventions. There also will be a brief review of the evidence supporting acceptance and commitment therapy (ACT), which is a mindfulness-based cognitive-behavioral approach that has seen a large surge in support for its use over the last 15 years.

Cognitive Therapy

Cognitive therapy is designed to help patients recognize, challenge, and modify irrational or dysfunctional beliefs, ideas, perceptions, and self-statements that are assumed to cause maladaptive emotional reactions to negative or stressful events. The general term for modifying such beliefs is "cognitive restructuring," which entails identifying habitual thinking, understanding that thinking strongly influences emotions, evaluating the rationality of thinking, and modifying the thinking if it is found to be illogical (and pathognomonic). Cognitive therapy has been investigated intensively because it has very specifically detailed putative causes and cures, and as such, these components are easy to describe in treatment manuals and examine in outcome studies.

Cognitive therapy has been found to be effective in a variety of psychological disorders. A review of meta-analyses encompassing 325

studies of cognitive therapy concluded that its efficacy with mood and anxiety disorders in particular is well documented. The reported benefits of cognitive therapy for depression are comparable to those of pharmacotherapy, with some indication that a combination of psychotherapy and medication may be the most effective. However, cognitive therapy may be better tolerated by patients and has better long-term outcomes. Specific cognitive techniques found to be helpful in treating depressive symptoms include monitoring mood, identifying cognitive distortions, and creating rational responses. These interventions are typically conducted via daily or weekly logs that allow the patient to keep track of his or her fluctuating emotional states, to list the related maladaptive thoughts, and to dispute the identified irrational beliefs. Maintenance sessions of cognitive therapy following positive response to the treatment of depression have been shown to reduce relapse rates.

In the treatment of anxiety disorders, cognitive therapy has been shown to be effective with GAD, panic disorder, and both social and specific phobias. Patients with GAD benefit from identifying and challenging habitual and irrational beliefs about events and their ability to cope with worry. Similar results for panic disorder have been reported, in which clinically significant reduction in panic symptoms and panic-related cognitions has been demonstrated. Phobic disorders are also amenable, to a somewhat lesser extent, to modifications of anxiety-provoking cognitions. The effectiveness of cognitive therapy with various other disorders, including substance abuse, schizophrenia, personality disorders, and the eating disorders, has been documented, if not as extensively as with the mood and anxiety disorders.

Janice, a 29-year-old married mother of two school-age children, struggled with moderately severe depression on and off for several years before seeking treatment from a cognitive therapist. With the assistance of the therapist, she learned to become aware of negative self-statements and to identify irrational beliefs related to the requirements of being a "good mother." She came to recognize that these thoughts caused her to feel incompetent and generally upset about how her children were doing. She learned to evaluate and appreciate evidence that her children were, in reality, well adjusted and well liked by both teachers and peers. Via cognitive restructuring, she learned to challenge and modify her thinking so that she could judge herself more accurately as a good mother. Other long-held, negative beliefs that contributed to depressed mood were also identified and challenged. Four months after terminating treatment, she returned for some "booster sessions" that helped her to get even better at recognizing and challenging excessively negative self-evaluations.

Behavior Therapy

Behavior therapy uses operant conditioning, classic conditioning, and modeling to help patients to change maladaptive behavior patterns. Techniques include the use of extinction, positive and negative reinforcement, stimulus control, shaping, and vicarious conditioning.

One of the most elemental behavioral techniques is relaxation training,

which has been evaluated and found effective for various problems in hundreds of studies. Through repetition of relaxation-inducing strategies (such as deep breathing or alternating tensing and relaxing muscles), patients are trained to produce voluntarily the relaxation response, which can be used to alleviate anxious mood, depression, involuntary muscle tension, pain, somatic complaints, and various other stress reactions.

Exposure-based behavioral interventions are often combined with relaxation training, and they have been shown to be highly effective in the treatment of a variety of anxiety disorders. The typical behavioral intervention involves confronting a feared stimulus while preventing the usual anxiety-based reaction (e.g., fleeing, engaging in a compulsion). Research has shown that patients will habituate to the feared stimulus and that the fear-based response will extinguish. For example, exposure and response prevention for OCD is an extremely effective approach that involves presenting the anxiety-provoking stimuli and preventing the patient from performing the compulsions normally engaged to reduce anxiety. Exposure-based treatments have likewise been shown effective for the other anxiety disorders, including GAD, PTSD, panic disorder, and both social and specific phobias. Behavior therapy has also been shown to be effective in the treatment of pain disorders, sleep problems, behavioral disturbances (such as those associated with schizophrenia), and substance use disorders.

Behavior therapy is commonly used to treat childhood psychopathology. Behavioral interventions might be implemented directly with the child. For example, early intensive behavioral interventions, which heavily rely on principles of positive reinforcement, have been shown to improve social skills, intelligence test scores, and school performance for children with autism spectrum disorders. Alternatively, behavioral principles have been successfully taught to parents for the alleviation of distress, symptoms, and impairment related to attention-deficit/hyperactivity disorder (ADHD), oppositional defiant disorder, school phobia, separation anxiety disorder, and enuresis. Behavioral interventions have also been shown effective in reducing self-injury in children with intellectual disabilities.

Tony, a 10-year-old boy with ADHD, was being maintained on a tolerable dose of stimulant medication. On the advice of their pediatric psychiatrist, the family sought adjunctive behavior therapy to deal with Tony's poor school performance. The therapist and his parents developed a behavioral modification program to change his homework-related behavior. They developed a system for rewarding Tony if he engaged in adaptive behaviors, such as writing down homework in his assignment notebook, bringing home any necessary homework-related school materials, and completing and turning in homework assignments. His parents worked with the teacher to set up similar contingencies to reinforce Tony's compliance, positive peer-related behavior, and general attentiveness in the classroom. After several months, Tony's schoolwork improved dramatically. His parents and teacher noted that he appeared to enjoy school more, to be more self-confident, and to be better liked by classmates.

Cognitive-Behavior Therapy

CBT combines the effective elements of cognitive therapy and behavior therapy. CBT is intended to help a patient develop strategies to change maladaptive thoughts and behaviors. For example, CBT for anxiety might include cognitive restructuring to change irrational fears, relaxation training to reduce physiological arousal, and exposure to extinguish fear reactions.

CBT has been the most extensively studied psychological intervention, and there is ample evidence supporting CBT's efficacy and effectiveness with numerous psychiatric disorders. A recent review of 269 meta-analytic studies determined the average effect size for CBT for different disorders. Regarding depressive disorders, CBT has demonstrated clear benefits as compared to a waiting list or no treatment conditions, but meta-analyses have shown mixed results in suggesting CBT is more effective compared to other active treatments (i.e., psychodynamic or interpersonal therapies). CBT and psychopharmacology have been shown to both produce medium–large effect sizes in treating depression, although a combination of both treatments was more beneficial than either alone. Regarding subgroups of patients with depression, individual studies suggest that CBT is also effective with older adults, adolescents, patients with comorbid diagnoses, and low-income minority women.

The effect sizes for CBT for most anxiety disorders and somatoform were consistently medium to large. CBT has also been shown to have a strong effect in treating insomnia, anger, and bulimia. Meta-analytic reviews suggest that CBT has clear benefit for special subgroups of patients, including treating internalizing disorders in children and emotional distress in older adults. CBT has been shown to have small–medium effects for substance use disorders, with more benefit for cannabis and nicotine dependence, as well as for chronic pain. Although there are small–medium effects of CBT for treating bipolar disorder and schizophrenia, there is less support for the use of CBT without the combination of psychopharmacology for these disorders.

Recently Developed Therapies

Early cognitive-behavioral treatments focused on depression and anxiety. As just reviewed, numerous applications to other disorders have been developed. In addition, the methods themselves have undergone revision to incorporate greater focus on contextual, existential, and developmental issues. The latter approaches have sometimes been referred to as "third wave" cognitive-behavioral treatments.

ACT proposes that psychopathology is due to cognitive fusion (unquestioning adherence to the truth of one's thoughts) and efforts to avoid painful experiences, which lead to maladaptive behavioral restriction.

ACT attempts to inculcate a more skeptical perspective toward one's cognitive habits, a more accepting attitude toward the pains and foibles of life, and a commitment toward behaviors that are consistent with one's deeply held values. Controlled trials of ACT have found it efficacious for depression, OCD, substance use disorders, and PTSD.

DBT proposes that skill deficits in emotion regulation are core causal factors of various disorders. DBT was developed for patients with BPD, but it has also been applied to the substance use disorders and the eating disorders. Controlled trials have demonstrated its efficacy in reducing nonsuicidal self-injury and other aspects of BPD.

Several recently developed therapies utilize mindfulness techniques (e.g., mindfulness-based cognitive therapy [MBCT]). Increasing mindfulness is intended to reduce the pathognomonic habit of ignoring or suppressing one's emotions and thoughts. The practice essentially entails experiencing thoughts and feelings as mental events that may or may not be related to external reality, but which can be experienced without judgment or fear. Controlled trials of mindfulness practices have found it effective for alleviating symptoms of the depression, anxiety, and substance use disorders.

Dan had experienced disabling panic attacks for 20 of his 45 years but only sought treatment when they started to interfere with his career aspirations. Treatment using David Barlow and Jerome Cerny's manualized *Psychological Treatment of Panic* included information regarding the physiology of panic attacks, which helped to reassure Dan that he would not die of an attack. Based on the use of cognitive techniques, other fears and misperceptions regarding his body's reactions to stress were identified and modified (e.g., that increased heart rate was not a sign of impending heart attack but was a normal biological response to stress). Using behavioral techniques, Dan was taught via relaxation training to breathe normally rather than hyperventilate during a panic attack. Over the course of treatment, Dan learned to confront previously feared internal and environmental cues without experiencing a panic attack. After 3 months, Dan felt that he was managing his panic disorder and discontinued treatment.

Interpersonal Therapy

IPT is a structured, time-limited intervention that is based on the notion that interpersonal stress is a causal and maintaining factor in psychological distress. IPT was originally developed to treat depression and it targets the following four domains: interpersonal deficits, role disputes, role transitions, and complicated bereavement. Therapy involves three phases; the therapist initially links the diagnosis of depression to a specific interpersonal problem, implements strategies for resolving the problem in the middle phase, and finally supports the patient's self-efficacy to continually use the strategies at the end of the treatment.

One meta-analysis found that the effect size for IPT was large compared to placebo groups in the treatment of depression, but nonsignificant differential effect sizes when IPT was compared to other psychological treatments. However, other studies have found IPT to have slightly better

dropout and remission rates but may take longer to show a therapeutic effect when compared to CBT. When compared to pharmacotherapy, the combination of IPT and pharmacotherapy prevented relapse better than medication alone.

IPT has been successfully adapted to treat depression in particular subgroups, such as adolescents and women with postpartum depression. In addition, IPT has been applied to patients with eating disorders, particularly bulimia nervosa and binge eating disorder, and has shown modest outcomes. In relation to eating disorders, interpersonal therapy may be most appropriate for patients with bulimia nervosa who are also experiencing interpersonal relationship difficulties.

Maureen sought therapy for depression and anxiety at 45 years of age, 2 years after her youngest of four children departed for college. With her therapist, she was able to recognize that her successful rearing of her children had the unfortunate consequence of their departure from home. Maureen was able to identify and to pursue new activities outside of the home, to appreciate and transfer the skills that she had mastered as a successful mother to outside of the home, and to reestablish musical interest and pursuits that had waned while she raised her children. After several months of therapy, Maureen was less depressed and anxious, as she pursued a teaching certificate in music, with the encouragement of her husband, at her community college.

Psychodynamic Psychotherapy

Psychodynamic therapies derive from Sigmund Freud's psychoanalysis. They are intended to promote self-understanding and insight into recurring problems with others. Psychodynamic therapists typically promote the development of a therapeutic alliance, within which the interpersonal problems recurring in the patient's life develop. In therapy, these interpersonal problems, which are the result of defense mechanisms and intrapsychic conflicts, are identified, interpreted, and managed, with the intent that such understanding and new behaviors transfer to other relationships.

Psychodynamic psychotherapies can range in length and intensity from long-term psychoanalysis to short-term psychodynamic psychotherapy. Short-term psychodynamic therapies have been manualized (e.g., Lester Luborsky's *Principles of Psychoanalytic Psychotherapy: A Manual for Supportive-Expressive Treatment*) and evaluated in controlled trials. Research supports both the efficacy and effectiveness of the treatment, suggesting it is as helpful as cognitive and behavioral approaches. Some studies suggest that patients in psychodynamic therapy better retain long-term gains from treatment. Researchers have documented the effectiveness of psychodynamic therapy in the treatment of various disorders, including the depression, substance use, generalized anxiety, panic, and somatic disorders, as well as the personality disorders.

Dario, a 34-year-old married father of three boys, sought therapy at the insistence of his wife for recurrent problems with anger and depression, both of which seemed to stem from feelings of inadequacy. His marital relationship had worsened considerably since his wife had started a new job. Toward the beginning of psychodynamic therapy, Dario expressed frustration at the situation but insisted that he did not resent his wife's successful career. Dario's underlying feelings toward his wife became evident in the course of treatment, as the feelings were transferred toward his therapist. His therapist helped him to identify and interpret his feelings of anger, fear, and sadness and to recognize that they were a manifestation of feelings of intense inadequacy and fears leftover from an abusive childhood. Dario was able to confront honestly and openly these feelings with his therapist and, ultimately, with his wife. His anger and depression lessened considerably, and his marriage improved.

Psychotherapy and Pharmacotherapy

Perhaps the most consistent finding in comparative research into the treatment of mental illness is that combining psychological and biological treatments provides the maximum likelihood of benefit. Various meta-analyses have found psychotherapy and pharmacotherapy to be equivalent in efficacy at both posttreatment and follow-up, whereas combined psychotherapy and pharmacotherapy has been routinely found to be superior to either alone.

TREATMENT VARIABLES RELATED TO OUTCOME

Therapeutic Alliance

The therapeutic alliance entails the sense of mutual warmth, caring, understanding, empathy, and agreement. Humanistic therapists, such as Carl Rogers, proposed that the alliance is curative in and of itself, theorizing that the patient's experience of a nonjudgmental, warm, genuinely caring attitude on the part of the therapist mediates positive change. In contrast, psychoanalytic and cognitive-behavioral theorists argue that the relationship is only a necessary precondition for the implementation of the active ingredients of therapy, such as interpretation or cognitive restructuring, which are responsible for positive change.

Research suggests that both perspectives are correct. Several meta-analyses have concluded that there is a consistent relation between the quality of the alliance and positive outcome. A strong therapeutic alliance, whether measured from the perspective of nonparticipant observers, the therapist, or the patient, is related to therapy effectiveness. These findings cut across various interventions, suggesting that the alliance in and of itself is curative, rather than that it enhances interventions. Research shows that a positive alliance facilitates patient adherence to treatment, suggesting that it is a necessary condition of successful interventions. Research likewise shows that patients who dropped out of treatment liked the clinician less, felt less respected, experienced a weaker therapeutic alliance, viewed the psychotherapist as more passive, and regarded psychotherapy as being less potent than other interventions.

The alliance is similarly important in nonindividual interventions. Studies examining group therapy have indicated that the benefit of this treatment modality is more likely when there is a strong level of group cohesion throughout the therapeutic process. A family's or couple's alliance with the therapist is an important predictor of positive response. Research into the causes, uses, and effects of the alliance is continuing.

Therapist Interventions

Psychotherapy interventions and procedures have been extensively studied. Some research suggests a significant positive relation between compliance with a treatment manual and outcome, but other research suggests that the therapist's relationship skills and perceived supportiveness may be diminished by excessive compliance with a manual. Moreover, numerous reviews and meta-analyses have concluded that not all therapists obtain equivalent results even if implementing manual-based treatments. One study found that about 30 percent of therapists perform consistently better than their colleagues regardless of approach. Another found that a small portion of therapists accounts for most of the occasional detrimental effects observed in treatment. Therapist characteristics that predict effectiveness include skill (but not necessarily experience), attention to the therapeutic alliance, the ability to repair the so-called "ruptures" in the alliance, and capacity to question and adjust their approach. Research suggests that therapy is most likely to be successful if therapists attend to the therapeutic alliance and make efforts to address potential problem with it. Some research suggests that alliance problems are highly likely, as patients struggle with the effort to implement recommended changes.

The most commonly studied therapist variables are those proposed by Carl Rogers. Research has consistently found therapist in-session empathy, genuineness, and unconditional warmth to be predictive of treatment success. On the other hand, there is minimal evidence to suggest that therapist gender, age, race or ethnicity, socioeconomic background, or indeed therapeutic stance (i.e., preferred treatment modality) is predictive of treatment success.

PATIENT VARIABLES RELATED TO OUTCOME

Patient Demographic Characteristics

A meta-analysis found that there is no association between patient age or gender and treatment effectiveness. Research suggests that older adults benefit from various forms of psychotherapy as much as younger adults, which is important because older adults have been found to prefer therapy over medication.

In terms of ethnicity, research indicates that members of ethnic and racial minority groups obtain fewer benefits from conventional

psychotherapies (in contrast to treatments customized to their sociocultural background) than their European American counterparts. Consistent with this finding, dropout rates are relatively higher for minority groups, especially if the provider and patient come from dissimilar racial, ethnic and cultural backgrounds. Research also suggests that financial difficulties on the part of the patient are related to less likely treatment seeking, greater likelihood of treatment dropout, and worse outcomes, particularly in anxiety and substance abuse treatment. These findings have been shown to be related to lack of trust, a preference for self-reliance, greater skepticism about the benefits of treatment, more concerns about potential negative effects, and enhanced perceptions of stigma, all of which make both treatment initiation and treatment adherence less likely. Research further suggests that these concerns are not altogether invalid. For example, the quality of services available to ethnic and racial minority groups tends to be significantly more poor, including underdiagnosis and inappropriate, ineffective (e.g., nonspecialty), or culturally incompetent treatment.

Patient Cognitive Characteristics

Numerous cognitive variables have been found to be predictive of productively engaging in psychotherapy. The best-supported include motivation to change and willingness to engage in the therapeutic process. Numerous studies have documented the utility, in terms of maximizing the likelihood of successful treatment, of enhancing a patient's motivation to change, or to participate in therapeutic efforts. Patients bring into therapy an attachment style, which research consistently shows is related to the capacity to develop an effective therapeutic alliance. Some evidence suggests that positive outcomes are associated with ego strength (i.e., the capacity to question assumptions and perceptions) and intelligence. A great deal of evidence implicates client expectations of being helped in therapy as a requirement of effective treatment.

Patient Clinical Characteristics

Investigation of the relation between patient disturbance and outcome has yielded consistent findings. Patients with less disturbance at the beginning of treatment are relatively healthier at the end of treatment than their more distressed peers. This is regularly found in studies of the treatment of mood disorders, anxiety disorders, and substance use disorders. (Paradoxically, patient distress is the single best predictor of the likelihood of seeking professional treatment.) Specific to mood disorders, patients with high levels of impairment appear to require relatively long-term, intensive treatment paradigms. Research has also delineated that the diagnosis of a personality disorder either as the primary focus of treatment or as a

comorbid condition of other disorders is an indicator of relatively poorer treatment outcome.

The quality of the alliance appears to be influenced by a variety of intra- and interpersonal patient characteristics. Patients who have a better history of or capacity for social relationships tend to develop a better alliance than defensive or negativistic patients. Problem severity does not appear to affect the quality of the alliance. There is converging evidence that measures of the alliance early in treatment are more strongly related to outcome than later measures. An exciting area of research concerns tracking the alliance across therapy, including monitoring therapist interventions, when there are problems in the alliance; successful treatments are often characterized by rupture and repair of the therapeutic relationship via active intervention on the part of the therapist.

Patient–Therapist Matching

The age, gender and racial and ethnic identity of the patient and therapist do not seem to predict outcome, but patient and therapist similarity in these characteristics does. There is a modest relationship between the age similarity of therapist and patient and treatment outcome. Therapist–patient ethnic similarities (as well as the ethnic sensitivity of the therapist) have been shown to be related to outcome, although the advantages of matching based on ethnicity has been an issue of debate. Some point to findings that ethnic minority patients prefer a therapist that is ethnically similar. For instance, a study in Los Angeles County examined the effect of matching patient and therapist ethnicity in the treatment of Asian Americans, Mexican Americans, African Americans, and European Americans. In all ethnic groups, matched patients stayed longer in therapy, but matching was positively related to outcome only for the Mexican American group. Still, other studies have found that if the therapist makes an effort to addresses openly the dissimilarity of therapist–patient ethnicity, the therapeutic relationship is strengthened.

Perhaps more important than ethnic match is whether the therapist and patient have compatible beliefs about the causes and treatment of psychological problems. That is, the focus on analytical and scientific explanations to mental health conditions has, for the most part, ignored religious and indigenous cultural explanations to mental health problems. Thus, therapists should not only respect patient beliefs about the etiology of their condition, but should also take these attitudes into account when developing and implementing intervention strategies.

Patient–Treatment Matching

There have been efforts to optimize the effects of psychotherapy by matching patient variables to treatment interventions. The aptitude-by-

treatment interaction (ATI) approach to psychotherapy research attempts to identify treatment strategies and procedures that produce optimal outcomes for individual patients. ATI research describes patient and therapist dimensionally with quantitative differences (e.g., resistive, impulsive, extroverted, directive, supportive). The questions addressed in ATI research include the following: (1) What procedures are important to effective treatment in different contexts? (2) Which patients are likely to respond to certain therapist or therapy characteristics? (3) How can patients be effectively assigned to therapies and their responses monitored within those therapies? ATI research is fairly recent and, by definition, requires standardized therapies.

One study examined patient variables of coping style (i.e., externalization and defensiveness) as patients were administered different types of therapy for major depression. Results suggested that externalizing patients and patients low in defensiveness did better in cognitive therapy, which is fairly directive, whereas internalizing patients and patients high in defensiveness did better in supportive, nondirective (or self-directed) therapy. Other research has found that highly anxious, repressed patients are more responsive to emotion clarification and supportive approaches, whereas externally cued patients respond better to more directive therapies, such as cognitive-behavioral psychotherapy. Similarly, whereas supportive therapy has been shown to be more effective with patients having higher levels of distress, cognitive therapy has been shown to be more effective with patients presenting with lower levels of distress. Research has provided evidence that depressed caregivers of older adults are differently responsive to psychodynamic and cognitive-behavioral psychotherapy depending on the length of time in the caregiving role. There is some evidence that panic disorder responds best and most quickly to a combination of behavioral and cognitive interventions.

EMERGING THEMES AND UNANSWERED QUESTIONS

The question of whether psychotherapy works has been definitely answered. There is a plethora of evidence from efficacy and effectiveness studies indicating that therapy is effective in alleviating emotional distress and behavioral dysfunction. Questions being addressed by researchers include the relative importance of specific (e.g., interventions) versus nonspecific (e.g., the alliance) curative factors, differential effects of treatment techniques (i.e., are some interventions more powerful for some clients or conditions?), and the transfer of research methods and technology to actual practice.

Two models of psychotherapy—the medical model and the contextual model—have competed in the last two decades. The medical model suggests that psychotherapies have specific ingredients that address some problem within the patient and account for outcomes. In contrast, the contextual

model holds that therapies are effective owing to universal relational and psychological factors, which vary according to the skills and qualifications of therapists. Meta-analysis supports the contextual framework: There are small, if any, differences among treatments; there is little evidence for the specificity or necessity of any particular intervention; the therapeutic relationship is a necessary aspect of treatment, regardless of the type of psychotherapy; and therapists' qualifications and contributions are a critical factor in the outcome of psychotherapy.

There is emerging consensus on the importance of patient-focused research, which uses research to inform case-based treatment decisions about the effectiveness of a specific treatment with a specific patient. Several systems of case-based decision support have been developed and applied to patient monitoring and therapist feedback. These approaches include the use of a global measure of patient functioning, repeated measurement during the course of therapy, and norms based on large patient and nonpatient samples to assess reliable change and clinically significant change. One approach uses patient profiling techniques to predict the course of response to treatment, given the clinical characteristics of the patient. Monitored progress can be compared to the expected rate of change qualified by the patient's clinical characteristics.

The main questions remaining for psychotherapy researchers entail maximizing the likelihood of positive effects for as many persons in distress as possible. This involves determining the suitable intervention, applied by the properly trained practitioner, to the matching pathological condition in the appropriately identified and amenable patient. As the United States becomes an increasingly diverse society, culturally appropriate treatment paradigms become an essential aspect of psychotherapy research. Accounting for cultural factors involves developing new interventions and/or adapting established protocols to deal with language, ethnic, and generational differences between the patient and therapist. Understanding treatment failure is also a critically important issue, given that the failure rate for certain disorders, such as substance abuse disorders, remains high. Finally, although there has been some progress in promoting psychological interventions as legitimate and beneficial, there remains intense and widespread stigma associated with mental illness and its treatment.

One of the challenges facing the field at present is getting psychotherapists to implement evidence-based treatments that have been developed in the laboratory. Moreover, policy changes, such as the Patient Protection and Affordable Care Act of 2010, have strongly promoted pragmatic, effectiveness-based research designs. As such, there has been a large increase in the last several years in the number of studies that examine how efficacious clinical interventions can be delivered in usual care settings or target patient-centered outcomes. Translational research evaluates the effects of dissemination strategies, which are attempts to

promote the use of research-supported psychotherapy techniques.

References

*American Psychological Association. Recognition of psychotherapy effectiveness. *Psychotherapy (Chic)*. 2013;50:102–109.

*Barlow DH, ed. *Clinical Handbook of Psychological Disorders: A Step-by-Step Treatment Manual*. New York: Guilford Press; 2014.

Boumparis N, Karyotaki E, Kleiboer A, et al. The effect of psychotherapeutic interventions on positive and negative affect in depression: a systematic review and meta-analysis. *J Affect Disorders*. 2016;202:153–162.

Butler AC, Beck JS. Cognitive therapy outcomes: a review of meta-analyses. *J Norwegian Psychol Assoc*. 2000;37:1–9.

Cabral RR, Smith TB. Racial/ethnic matching of clients and therapists in mental health services: a meta-analytic review of preferences, perceptions, and outcomes. *J Couns Psychol*. 2011;58:537–554.

Carr A. The effectiveness of family therapy and systemic interventions for child-focused problems. *J Fam Ther*. 2009;31:3–45.

Castonguay LG, Beutler LE. Common and unique principles of therapeutic change: What do we know and what do we need to know? In: Castonguay LG, Beutler LE, eds. *Principles of Therapeutic Change That Work*. New York: Oxford University Press; 2006;353–369.

Chorpita BF, Daleiden EL, Ebesutani C, et al. Evidence-based treatments for children and adolescents: an updated review of indicators of efficacy and effectiveness. *Clin Psychol Sci Prac*. 2011;18:154–172.

Cuijpers P, Geraedts AS, van Oppen P, Andersson G, Markowitz JC, van Straten A. Interpersonal psychotherapy for depression: a meta-analysis. *Am J Psychiatry*. 2011;168:581–592.

Gelo O, Pritz A, Rieken B, eds. *Psychotherapy Research: Foundations, Process, and Outcome*. Berlin: Springer; 2015.

Gu J, Strauss C, Bond R, Cavanagh K. How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? *Clin Psych Review*. 2015;37:1–12.

Hanrahan F, Field AP, Jones FW, Davey GC. A meta-analysis of cognitive therapy for worry in generalized anxiety disorder. *Clin Psychol Rev*. 2013;33:120–132.

Hofmann SG, Asmundson GJ, Beck AT. The science of cognitive therapy. *Behav Ther*. 2013;44:199–212.

Hofmann SG, Asnaani A, Vonk IJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. *Cognit Ther Res*. 2012;36:427–440.

Kazak AE, Hoagwood K, Weisz JR, et al. A meta-systems approach to evidence-based practice for children and adolescents. *Am Psychol*. 2010;65:85–97.

Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62:593–602.

*Lambert MJ. The efficacy and effectiveness of psychotherapy. In: Lambert MJ, ed. *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*. 6th ed. New Jersey, NJ: Wiley;

2013.

Lemmens L, Müller V, Arntz A, Huibers M. Mechanisms of change in psychotherapy for depression. *Clin Psych Review*. 2016;50:95–107.

Mackenzie CS, Erickson J, Deane FP, Wright M. Changes in attitudes toward seeking mental health services: a 40-year cross-temporal meta-analysis. *Clin Psychol Rev*. 2014;34:99–106.

Mojtabai R, Olfson M, Sampson NA, et al. Barriers to mental health treatment: results from the National Comorbidity Survey Replication. *Psychol Med*. 2011;41:1751–1761.

Nathan PE, Gorman JM, eds. *A Guide to Treatments That Work*. 4th ed. New York: Oxford University Press; 2015.

*Norcross JC, ed. *Psychotherapy Relationships That Work: Evidence-Based Responsiveness*. 2nd ed. New York: Oxford University Press; 2011.

Olfson M, Marcus SC. National trends in outpatient psychotherapy. *Am J Psychiatry*. 2010;167:1456–1463.

Powers MB, Zum Vörde Sive Vörding MB, Emmelkamp PM. Acceptance and commitment therapy: a meta-analytic review. *Psychother Psychosom*. 2009;78:73–80.

Reichow B. Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. *J Autism Dev Disord*. 2012;42:512–520.

Shedler J. The efficacy of psychodynamic psychotherapy. *Am Psychol*. 2010;65:98–109.

Vocks S, Tuschen-Caffier B, Pietrowsky R, Rustenbach SJ, Kersting A, Herpertz S. Meta-analysis of the effectiveness of psychological and pharmacological treatments for binge eating disorder. *Int J Eat Disord*. 2010;43:205–217.

*Wampold BE, Imel ZE. *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work*. 2nd ed. New York: Routledge; 2015.

Wang PS, Demler O, Olfson M, Pincus HA, Wells KB, Kessler RC. Changing profiles of service sectors used for mental health care in the United States. *Am J Psychiatry*. 2006;163:1187–1198.

▲ 33.16 Individual Psychodynamic Psychotherapy and Cognitive Behavioral Therapy for Psychosis

MICHAEL D. GARRETT, M.D.

INTRODUCTION

This chapter reviews the individual psychodynamic psychotherapy (PDTp) and cognitive behavioral therapy (CBTp) of psychosis. It is organized historically, first describing clinically based psychoanalytic theories developed in the last century, then covering the newer, more research-based CBTp approaches that have evolved in the past 25 years, and finally suggesting the practical integration of both bodies of knowledge.

There is much overlap in these approaches, and there are some significant differences. Both place central importance on a close, collaborative therapeutic relationship, active listening, and a technique

guided by psychological theory. PDTp, framed in a theory that assumes an open-ended receptive technique, places greater emphasis on psychological conflict, dissociated affect, fantasy, and the patient's past, while CBTp, employing a more structured technique, places more emphasis on cognition and behavior in the patient's present life. Each treatment pivots on a different foundation aiming to arrive at a similar place. PDTp seeks to bring insight and reduce suffering in the present by showing patients how psychotic symptoms are disguised expressions of emotionally charged experiences from the past. CBTp seeks to reduce suffering by marshalling the patient's capacity to examine evidence to change delusional beliefs. PDTp therapists focus on the emotional truth of psychotic symptoms; CBTp therapists focus on their literal falsity. These approaches meet and can be integrated where the heart meets the mind. Both may be of value to the same patient; at different times in treatment, patients with different needs may benefit from either approach separately or in combination.

This chapter begins with two clinical vignettes, one of a man with psychosis treated with PDTp and one of a woman treated with CBTp. Both had a similar core psychological schema: abused as children, they were left with a psychological legacy of fear, need, rage, and expectations of mistreatment and abandonment that made it very difficult to establish stable, satisfying adult relationships.

THE MAN WHO CHOKED PEOPLE. First the patient treated with PDTp. Bertram Karon was invited to provide psychotherapy for a man with chronic psychosis, who had been hospitalized for 10 years, a man with a history of unprovoked violence against other people in the hospital. The patient was grossly incoherent, and when he became coherent, he stuttered very badly. Review of his chart revealed that he came from a poor family, his father was an alcoholic, he had developed the stutter as an adolescent, and he had reported to sick call in the Army with a venereal infection of his mouth just before his first assault on a stranger. He would creep up behind other patients and choke them, but he had never killed anyone, as he would drop his victims when he had rendered them unconscious. After establishing rapport with the patient, in the first session Karon discovered that when he was a little boy, his mother, for minor offenses like not eating, would put a cloth around his neck and choke him. After the patient had revealed this piece of history to the therapist, he stopped choking other patients.

The patient began a subsequent therapy hour by yelling, "Why did you do it to me, Dad?" Exploring the psychotic transference, Karon learned that the patient's alcoholic father had come home drunk one day and anally raped him. The patient's terrible stutter was revealed to have an extraordinary cause. Latin words were discernable in the middle of his stutter. When asked if he had been an altar boy, he said, "You swallow a snake, and then you stutter. You must not let anyone know." He was extremely ashamed and guilty at having been compelled to performed fellatio on a priest. The therapist offered direct reassurance that what he had done was all right, and said, "Anyone as hungry (for love) as you were would have done the same thing." At that point the stuttering stopped. When he started to stutter in later sessions, it was only necessary to repeat this interpretation, and the stuttering immediately ceased. As a boy, the patient had turned from his mother who choked him to a father who raped him to a priest who sexually abused him. This succession of horrors drove him out of his mind. Yet his hospital record contained not one word of this history, and 10 years of confinement and treatment as usual had done nothing to allay his primary symptoms. This case illustrates how seemingly bizarre symptoms can be understood to have meaning, and how psychotherapy can ameliorate them. There are many such reports of successful psychotherapy scattered through the literature.

Now the second patient, treated with CBTp by the author.

THE WOMAN WHOSE CAT SOUGHT TO MURDER HER. The patient was a 25-year-old woman divorced mother of three who had made several serious suicide attempts. She reported feeling depressed, hopeless, and terrified when she came to believe that her cat was planning to murder her in her sleep. In the weeks prior to the emergence of this delusion, she had been feeling unusually needy and had begun telephoning her boyfriend at work many times a day, until he began to set limits on the phone calls. She interpreted this as meaning he no longer loved her. She blamed her cat for the loss of his love. At night when she and her boyfriend went to bed, the cat would snuggle up between them, which the patient experienced as the cat "coming between" her and her boyfriend. The boyfriend showed the cat affection, which indicated to the patient that the cat's plan to steal her boyfriend's love was working. She believed the cat planned to murder her so as to have her partner all to herself. The psychodynamic meaning of the delusion was readily apparent. Having been abused as a child and hungry for love, she felt a compelling need to be in sole possession of her boyfriend's affection.

The therapist inquired about the details of how the patient imagined the cat would kill her. The patient responded as though this should be perfectly self-evident. "You see it all the time on TV! Lions and tigers bite their prey in the neck. They bite the jugular!" The patient expected the cat to bite her jugular when she was asleep. While respecting the patient's conviction that the attack was inevitable, the therapist inquired in an offhand manner if the patient knew where her jugular was located. She said she knew it was somewhere in the neck, but did not know exactly where. The therapist commented, again in an offhand manner, that the cat would have to have a much more precise knowledge of human anatomy than she did. While speculating where in the neck the jugular was the patient acknowledged that her being overweight and quite thick in the neck would make it especially difficult for her cat to locate her jugular. The therapist opined that the cat would certainly have one and only one chance to attack her precisely at the jugular, because as soon as she felt the cat's teeth on her neck she would surely awaken immediately and throw the cat off, compress the bite on her neck, and call an ambulance. The patient returned the next session less anxious and less depressed. When the therapist explored her change in mood, she said, "I still believe my cat plans to kill me, but I no longer believe she has the means to do it." This positive shift in mood and decrease in depression, while not resolving the delusion entirely, was a significant gain obtained in a single session. The psychotherapy then continued to explore her relationship with her boyfriend.

Chronic psychosis is the cause of enormous suffering and individual disability. In contrast to significant advances in the treatment of heart disease, cancer, and stroke that have dramatically improved quality of life and life expectancy, most chronically psychotic patients remain functionally disabled in their entire lives and live 15 to 20 years less than the general population. While medication is undoubtedly helpful to some patients at some times in some ways, compared with other branches of medicine, therapeutic outcomes in the treatment of chronic psychosis leave room for improvement. Psychotherapy can help achieve more positive results.

Psychotherapy (broadly defined) is a disciplined professional interaction between a trained psychotherapist and a patient that aims to improve the patient's quality of life by exploring the patient's thoughts, feelings, and behaviors, with the goal of facilitating new possibilities for

emotional growth and more adaptive thinking and functioning. This chapter will assume that the reader has at least some familiarity with such essential concepts in PDT as the existence of unconscious mental processes, unconscious fantasy, transference, and the therapist's interpretation of psychological meaning, and CBT concepts including the existence of cognitive biases in thinking and the A-B-C cognitive sequence in which an activating event is followed by a belief about the meaning of that event, which in turn leads to an emotional and/or behavioral consequence.

Although psychotherapy is widely practiced in the treatment of nonpsychotic patients, psychotherapy has for decades taken a backseat to pharmacological treatments of psychosis. The reason psychotherapy for psychosis is so little practiced is not that psychotherapy has no value. Research to be cited in a later section of this chapter demonstrates its efficacy. This chapter's exclusive attention to psychotherapy should not be taken to suggest that biological factors are less important than psychological factors. Why has psychotherapy lagged so far behind medication as a frontline treatment? In the author's view, the following issues have contributed to marginalizing psychotherapy for psychosis.

1. The biggest obstacle to providing psychotherapy to patients with psychosis is the widespread belief that psychotherapy is of no value. The stunning initial successes in the work of pioneers like Harry Stack Sullivan, Frieda Fromm-Reichmann, and Silvano Arieti and the solid evidence base for CBTp should have put this mistaken idea to rest, but few people have been trained in these approaches. It is not surprising that approaches that are helpful to other patients are insufficient with patients suffering from psychosis and have left clinicians with feelings of inadequacy, failure, and pessimism about therapy for psychosis.
2. The same symptom picture might be diagnosed as "brief reactive psychosis" or "schizophrenia" depending solely on length of illness. This convention implies that schizophrenia is by definition a bear of a disease that never gets better. Such fatalism embedded in diagnostic practice allows clinicians to be complacent about poor outcomes rather than to press hard for better treatments.
3. Understanding the psychology of psychosis requires the clinician to empathize with levels of anguish not often encountered in everyday life. Clinicians may take for granted that psychotherapy can help patients with nonpsychotic depression and anxiety because they know from personal experience that talking things through with the right person can help, but when they see little connection between ordinary mental life and psychotic symptoms, the utility of talk therapy for psychosis may not be intuitively obvious.
4. The introduction of the neuroleptic drug chlorpromazine (Thorazine) in

1954 led to the dramatic amelioration of acute psychotic symptoms in many patients. The observation that effective neuroleptics diminish dopamine transmission led to the dopamine theory of schizophrenia, which added credibility to the idea that schizophrenia is essentially a brain disease, and as such is more likely to respond to medication than to psychotherapy.

5. The obvious efficacy of neuroleptics in relieving some psychotic symptoms fueled a backlash against the unfortunate hypothesis that the "schizophrenogenic mother" was the cause of psychotic illness, but this needed course correction went too far in the other direction, lending support to the idea that psychological factors are of little importance.
6. Freud did not believe that psychotic patients formed a transference that could be analyzed in a conventional psychoanalytic treatment, and so had no faith in the efficacy of psychoanalysis, at least as he practiced it, for psychosis. With some notable exceptions, the psychoanalytic community mirrored Freud's attitude and largely abandoned the care of the severely mentally ill.
7. PDTp failed to establish itself not only because it was crowded from the scene by biological treatments, but because a classical psychodynamic approach is rarely effective without a modification of technique. Psychodynamic therapists have paid too much attention too early to the unconscious psychological meanings of psychotic symptoms and too little attention to the patient's conscious experience of psychotic symptoms. Pioneers of PDTp have often been mythologized as exotically talented clinicians who could do what the average well-trained clinician could not. In fact, doing effective PDTp and CBTp is well within the range of most clinicians in public clinics when they are provided with specialized training and supervision.

The time has come to give renewed consideration to the place of individual psychotherapy in the treatment of psychotic individuals. Not only is individual psychotherapy often effective, but there is also mounting evidence of the limitations of psychopharmacology. The clinical antipsychotic trials of intervention effectiveness (CATIE) study found that newer medications are not decisively more effective than older drugs and cause serious side effects of their own, including metabolic syndrome and diabetes. Two recent studies suggest that structural changes in the brain correlate not only with total period of untreated psychosis but also with total exposure to neuroleptics; that is, both psychosis and the drugs used to treat psychosis may damage the brain. A recent review of the traumagenic neurodevelopmental hypothesis suggests that the brain changes typically attributed to a primary brain disease may reflect changes in the hypothalamic–pituitary–adrenal axis and dopamine system in response to psychological trauma. These studies suggest that psychiatrists should aim

for the lowest dose of medication possible for the shortest period of time. Psychotherapy may help to achieve this goal.

THE WOMAN WHO HAS NEVER BEEN REHOSPITALIZED. In the author's experience, effective psychotherapy can reduce the necessity for dosage increase, polypharmacy, and inpatient admission. Consider an example that illustrates the possibilities for combining medication and psychotherapy. A 30-year-old woman with a chronic paranoid psychosis with a 6-year history of treatment as usual, with 3 to 4 hospitalizations a year and intermittent homelessness, was referred to the author for psychotherapy. She had failed multiple neuroleptics including clozapine (Clozaril) and was being prescribed a high dose of aripiprazole (Abilify) at the time of initiating psychotherapy. In the course of a 5-month, once-a-week CBTp treatment, the patient understood for the first time that her belief that she was being pursued by thugs, her conviction that she was being monitored through a computer chip in her body, and her experience of "voices" telling her she was a child abuser were all manifestations of an illness (the convention "voices" will refer to verbal auditory hallucinations). Aripiprazole was titrated downward to a lower dose, but could not be stopped without a flare-up of psychotic symptoms. She attends session every other week and has never been readmitted. During her 13 years in treatment she suffered two significant recurrences of paranoid ideas and ideas of reference, both of which were worked through psychotherapeutically without recourse to increased medication or hospitalization.

History

In the interest of brevity, one can divide the history of PDTp after its beginnings in the early 1900s into two schools, a European Kleinian School heavily influenced by the ideas of Melanie Klein that was active in the first half of the twentieth century, represented by Hannah Segal, Wilfred Bion, and Herbert Rosenfeld, and an American Interpersonal School based on the ideas of Harry Stack Sullivan that has been influential from the mid-twentieth century on, represented by Harry Stack Sullivan, Frieda Fromm-Reichmann, Silvano Arieti, and Harold Searles. See the review by Michael Stone (11) for a more comprehensive listing of contributors. The history of CBTp that begins in the early 1990s will be covered in a later section. "Cognitive Behavioral Therapy for Psychosis."

Psychodynamic thinking about schizophrenia began at the Burghölzli Mental Hospital in Zurich, Switzerland, when Eugen Bleuler, MD, became Director in 1898. He emphasized the importance of psychological "complexes" in the formation of psychotic symptoms. Carl Jung worked there during Bleuler's tenure. Jung recalled that most psychiatrists at the time paid attention to diagnosis but ignored the meaning of psychotic symptoms.

Jung described a psychotic woman who had spent many years in the hospital who made stereotyped, repetitive, seemingly meaningless gestures. He read her chart and found an entry years prior that described these gestures as a "cobbler's motion," i.e., like the movement a cobbler would make drawing a thread through leather in a shoe he was holding between his knees. The patient died shortly after Jung met her, and her elder brother came for the funeral. Jung asked him, "Why did you sister lose her sanity?" The brother told Jung that she had been in love with a shoemaker who had not wanted to marry her, and when he rejected her, she had "gone off the deep end." Jung concluded that the shoemaker's movements represented the heartbroken patient's identification with her beloved cobbler, a way of keeping her sweetheart

with her psychologically despite his having rejected her in reality. In short, he broke her heart, she lost her mind, and she remained "married" to him in gesture until death did her part.

Paul Federn was one of the few analysts in Freud's circle with significant experience treating psychotic patients. He extended Victor Tausk's concept of the "ego boundary," a fundamental idea in the psychoanalytic theory of psychosis. Federn believed that the fundamental problem in schizophrenia was that the ego boundary that ordinarily separates thoughts and feelings from perceptions is faulty, allowing the content of the unconscious to spill into consciousness where it is perceived as reality.

Theoretical Issues

In both psychodynamic and CBT literatures on borderline personality disorders (e.g., work by Kernberg, Fonagy, Meares, Linehan, and Young), one finds research-based efforts emphasizing specific deviations from ordinary psychoanalytic or cognitive behavioral treatment required for that population. Although there is a small, clinically based psychoanalytic literature on "supportive therapy" of schizophrenia, comparable refinements of theory and technique have yet to be fully realized and widely practiced in the treatment of individuals with psychosis.

As clinical technique should derive from research-supported theory, this chapter now turns to the psychodynamic theory of psychosis. Despite disagreements about the relative importance of biological and psychological factors in the etiology of psychosis, there is substantial agreement among psychodynamic theoreticians about its fundamental psychology: *Psychotic states represent the intrusion into adult mental life of primitive psychological processes that are ordinary in childhood; these early processes normally recede from consciousness by the time a person reaches later childhood (the "latency" period) but they re-emerge in psychosis.*

These intrusions of early mental life are of two sorts. First, individuals in psychotic states utilize *primitive cognitive processes* similar to those operative in the first years of life. Instead of thinking a thought and expressing it in verbal symbols, or feeling a feeling that is subjectively located within the psychological boundary of the self, psychotically disturbed individuals experience a body sensation or experience a distorted perception of the outside world that expresses feelings and thoughts that a nonpsychotic person would express in words. For example, instead of thinking the thought, "I am a bad person" a psychotic individual might *perceive* a "voice" saying, "You are a bad person." Federn called this failure to distinguish between subjective experience and perceptions of the outside world a *breakdown of the ego boundary*. This confusion of thought, feeling, and perception plays a significant role in the formation of ideas of

reference, hallucinations, and delusional beliefs.

Second, *primitive object-related phantasies* (more on this language later) that ordinarily operate in the first years of life re-enter consciousness in psychosis and dominate the person's experience of relationships with other people and the world, shaping the content of delusions and hallucinations and the narratives that patients construct to explain their lives. For example, while normal children might play at being king or queen, treating their parents as vassal subjects while resisting efforts to spoil the fun and expose the game as play-acting, people with schizophrenia might make a literal claim to being a king or queen, insisting that family members and clinical staff bow to their will, while resisting efforts to characterize their belief as illusory.

What remains unconscious in psychosis and in nonpsychotic conditions, respectively, is different. The psychotic person, while closer to primitive childhood mental life, is not consciously aware of using archaic mental processes to keep painful thoughts and feelings about real life outside conscious awareness.

Although there is widespread psychoanalytic agreement about the intrusion of primitive psychological processes in psychosis, historically there has been disagreement about how to deal with these intrusions psychotherapeutically (see Table 33.16–1). Because Federn viewed the psychotic process as a defect in the brain's ability to invest internal mental events with sufficient "ego feeling" to mark them as belonging to the self, he advised clinicians not to encourage the emergence of unconscious material and to harbor no ambition to interpret unconscious processes. Instead, they should attempt to compensate for the patient's defective ego boundary by supportive reality testing. His approach is echoed in current-day psychosocial programs such as day hospitals, supportive housing, and employment arrangements that center treatment in a clinical team rather than a single therapist. Federn's concern that exploratory psychodynamic treatment might undermine the reality-testing of psychotic individuals is mirrored in the empirically unproven belief of some in the 1980s that psychotherapy for patients with psychosis is ineffective or even harmful.

Table 33.16–1.
Supportive, Kleinian, and Sullivanian Approaches to Psychotherapy of Psychosis

	Focus of Theory	Focus of Interpretations	Handling of the Transference
Federn and the current clinical zeitgeist	A defective ego filter that fails to bar primitive unconscious content from	Do not encourage the emergence of unconscious content with the aim of interpreting it. Develop a supportive relationship that eventually allows the therapist to	Try to avoid a psychotic transference by avoiding free association and including other people as therapy aides to diffuse the intensity of the relationship with the

	consciousness	encourage reality testing.	therapist.
European Kleinian School Klein, Segal, Bion, Rosenfeld	Primitive unconscious object-related phantasies	Interpret the primitive unconscious phantasy that operates through mechanisms of projective identification.	Welcome the psychotic transference with the intent to interpret it as an essential opportunity for therapeutic change.
American Sullivanian School Sullivan, Fromm-Reichmann, Arieti, Searles	Failure to achieve independence and stable social competency within one's core family group	Show the patient how psychotic symptoms are meaningful reflections of painful early interpersonal failures.	Diminish the blank screen of classical analytic technique that encourages a psychotic transference by making a special effort to establish a real trusting relationship with the patient.

Followers of Melanie Klein, in contrast, advocated direct interpretation of primitive thinking with no modification of classical psychoanalytic technique. They regarded the "psychotic transferences" that then emerged in treatment as important opportunities to make interpretations revealing the unconscious ideas dominating patients' experiences.

Also in contradiction to Federn's ideas, analysts in the American Interpersonal School of Harry Stack Sullivan sought above all to establish trusting relationships with patients that would facilitate their ability to open a dialogue. They did not encourage the development of a psychotic transference, however, nor regard it as the essential focus of therapeutic change. Rather, they would interpret transference distortions when they interfered with the therapeutic alliance. Sullivanians made interpretations at the interface between patients' subjective experiences of painful situations in childhood, latency, and adolescence, and the meaningful expression of these experiences in the form of symptoms such as delusions and hallucinations.

The Psychology of Psychosis

These pioneers, despite their differences, all tried to reduce the apparent distance between psychotic processes and ordinary human psychology. If one cannot understand psychotic states psychologically, how can one hope to treat them psychologically? How can one relate the seemingly bizarre symptoms of psychosis to ordinary human psychology, and in so doing forward a psychological theory of psychosis and a rationale for psychological treatment? Addressing these questions begins in an understanding of the psychology of young children that can be inferred from direct child observation and clinical experience.

The experience of newborns and infants under 4 months old, as best as can be determined, consists of isolated fragments of somatic sensations

associated with perceptual images that have not yet been integrated within the boundary of an organized self. The newborn does not know yet that the hand in its immediate visual field of vision is *its* hand. Children do not recognize themselves in a mirror until roughly 18 months. Very young children do not know that sensations are part of their own personal subjective experience. As development proceeds, consciousness moves beyond somatic perceptions to increasingly complex "mentalized" forms of experience, allowing for symbolic representations, verbal thought, and conscious and unconscious phantasy.

Klein came to believe that early in life, experiences are organized and represented mentally in "phantasy" as *internal object relationships* between psychological "objects," where the self and an "other" (the *object* of the child's interest) are connected by an affectively charged link: the self is doing something to the object or the object is doing something to the self. This chapter adopts her use of "phantasy" to refer to primitive unconscious phantasies and "fantasy" to refer to conscious daydreaming. Given that human beings are fundamentally interdependent social animals, it is not surprising that human phantasy is organized in an object-related form. In object relations theory, the primitive self defends against painful mental experience by splitting into two primal psychological objects: a pristine self that attempts to preserve its freedom from pain by identifying itself with a natural state of being phantasied to be pain-free and a persecutory object that is causing the pain by intruding upon what would otherwise be the satisfactions of the primitive self. In this construction, human suffering does not arise from internal conflict but is rather caused by persecutory attacks of elements experienced as outside the self.

This basic tendency to split off the pristine self from a persecutory object is apparent in childhood stories and their adult derivatives. (Superman is all good; he does not cheat on his taxes. Lex Luther is all bad; he does not spend weekends volunteering at soup kitchens for the homeless.) This psychology leaves many residuals in adult mental life. For example, during a headache someone may say, "My head is killing me." In this case, two psychological objects are being imagined, a "me" that in its natural state would be headache-free were it not for the murderous activity of the head, and a "head," a persecutory object intruding pain into the otherwise pain-free "me." In psychotic illness, this split between the self and a persecutory object becomes the basic psychological structure of a paranoid delusion. The fundamental tendency to split the mind into contrasting mental representations of self and persecutory object is the source of the ubiquitous human tendency to blame others for one's failings and sorrows; for example, the train is responsible for my lateness (not my failure to leave early enough); a particular ethnic group is responsible for society's social ills, and so on.

Child observation suggests that phantasy begins preverbally. Susan

Isaacs described a 20-month-old child with delayed speech development who saw her mother wearing a shoe with a loose sole that was flapping about. The child screamed in terror, and for a week would shrink away if she saw her mother wearing any shoes at all. At 2 years 11 months, when she was able to express herself verbally, she suddenly asked in a frightened voice, "Where are Mummy's broken shoes?" Her mother reassured her that she had sent the frightening shoes away, to which the child replied "They might have eaten me right up." The child at 20 months had seen the shoe as a threatening mouth but could not express this phantasy verbally. Her conception of the shoe as a dangerous mouth can have no other origin than her psychology. Her oral aggressive impulses were projected outside the self and experienced as though residing in the shoe. In this phantasy, the internal danger posed by the child's aggression is transformed into a danger threatening from outside the self; that is, an intrapsychic conflict has become an interpersonal problem between two internal objects, with the persecutory object experienced as located in the outside world. A similar process happens in psychosis, when a man in a state of great internal turmoil conceives his difficulty as an external threat from the Mafia or a government spy agency.

In the first 3 years of life, the child's conception of the world is dominated by projected elements of its own mental life; for example, the shoe is a mouth. This developmental period can be characterized as a two-person psychology in which the child feels inherently related to particular persons. Young children live in a psychic world in which what they feel and think is experienced as real. The girl who was frightened of the shoe was not telling her mother "I am having a phantasy that your shoe is a scary mouth, but I know it is just a shoe." Rather, she *saw* the shoe as a threatening mouth. The child's state of feeling *is* the child's reality during this early phase, and its perception of the world is conditioned by its emotions. This is also the case in psychosis.

Peter Fonagy and Mary Target distinguish two modes of cognitive functioning in young children, the stage of "psychic equivalence" in which whatever is psychically experienced is felt as real, and the "pretend mode," in which ideas are representational but their correspondence with reality is not examined, allowing children to play at pretend games. Consistent with the preoccupations of young children, the primitive object-related phantasies that predominate in the first 3 years of life often have oral and anal themes, such as shoes that bite. The child's experience of eating becomes the body-based metaphor for the psychological processes of *introjection* and *identification,* via which aspects of the environment are taken into the self; defecating becomes the body metaphor for *projection,* in which aspects of the self are extruded and deposited in mental representations of psychological objects outside the self.

In effect, in the service of defense, the mind appropriates the body

processes of taking in and pushing out content and transforms these processes into psychological defenses. At this maturational stage, the mind is like an open international border where heavy trucks rumble back and forth day and night, exporting projected psychic content and importing it back again after it has been modified by the external environment. Residuals of these early, body-grounded mental processes remain in adult metaphors such as "That's food for thought!" where the comprehension of an idea is conceived as the ingestion of a satisfying food, a sort of mental "good breast" experience. When someone says, "I had to swallow my pride," they mean "I didn't abandon my pride. Rather, I temporarily hid it in an interior body space from which it can safely re-emerge when the coast is clear."

If young children are to achieve a firm sense of being real in a real world, they must grow out of their experience of the world as projected fragments of the self. How does this happen? Donald Winnicott, a practicing pediatrician and psychoanalyst, described how a secure attachment to an empathic mothering figure becomes the developmental base for a child's experience of reality. Winnicott's "good-enough mother" responds empathically to the child's authentic strivings while not imposing an alien agenda on the child's gradually emerging unique identity. The child's earliest experience of the external world is its encounter with the mother's body, especially the breast. The empathic mother, attuned to the baby, learns to read the infant's communications so as to feed the baby when it is hungry, and change the diaper when it is wet, and not the other way around.

As described by Winnicott, fancifully, the baby's experience of hunger and feeding might be, "I was feeling uncomfortable just now. I was starting to imagine something that would eliminate the discomfort, when suddenly, as if by magic, the breast appeared. Then I knew I was trying to imagine something like a breast all along. I am of this world. I am in synch with this world. The world sees my existence." The empathic mother delivers the real world to the child at an opportune moment of need, and thereby builds a bridge between the child's psychic reality and the external real world.

Children who find sufficiently resonant correspondences between their mental states and the outside world turn toward reality and begin to feel real in a real world. A child whose authentic stirrings are not met empathically may turn inward and begin living as a "false self" that merely reacts to the impingements of others rather than expressing its authentic nature. If the world returns no resonant ping to a child's sonar probe, the child adapts to living in a dead world that appears to not know the child's true self is alive. In some family systems, the child is recruited to fulfill the psychological needs of others in the family, in which case the family environment impinges upon and distorts the development of the child's authentic individuality. This impingement beggars the self, leaving the

child little base from which to build a sustaining sense of interpersonal competence and independence, a state of vulnerability that can predispose to psychosis.

In secure, empathic environments, the primitive unconscious phantasies of early childhood fade over time, but when the social environment impinges upon a child's authentic self when these phantasies are active, they do not fade. Cumulative disappointments give rise to painful conscious fantasies and unconscious phantasies that are reinforced. These will influence how later life experience is perceived and given meaning, and may leave people with such experiences feeling they have no space for tolerable living. What others have described as the prepsychotic child's difficulty with separation and individuation from a symbiotic relationship with the mother can be viewed as a massive impingement of caregivers on the authentic independence of the child.

Recall the little girl who was frightened that her mother's shoe would eat her. Oral experiences are important to young children, who often phantasize about biting and eating, and being bitten and eaten. Frightening oral themes of this sort appear in fairy tales and also in psychotic delusions. In "Hansel and Gretel," a famine descends upon the land and a terrible mother plans to leave her two children to die in order to conserve the limited food supplies for herself. The children narrowly escape being eaten by a witch. Little Red Riding Hood and her grandmother are both eaten by a wolf who has assumed the grandmother's clothing.

Compare the internal world of a psychotic man from a chaotic family with a drug-addicted mother and absent father, who suffered a delusion composed of primitive oral phantasies. He reported that he had developed a cancer when at birth his mother touched his penis (the mother has put a "persecutory bad object" into him in the form of the cancer), and that this cancer now coated all his internal organs (the "bad object" is about to engulf him). Because he feared that medication would disturb a delicate balance between his organs and the cancer, giving the cancer the upper hand, he refused all pharmacological treatment. In phantasy, taking medication would destabilize his resistance to the persecutory object. This man was frozen in a state of terror in which he feared his true self would disappear inside a bad object (the cancer) and be annihilated. In effect, his delusion is a horrifying version of "Hansel and Gretel" and "Little Red Riding Hood" with no happy ending, a personal fairy tale with a readership of one.

Similarly, a 30-year-old man with two prior psychotic episodes reported in a first psychiatric consultation an intense anxiety he associated with his belief that a terrible drought was about to occur, which had led him to fill his apartment with many gallons of stockpiled water. At the end of the consultation, the patient said, "You have asked me a lot of questions. Can I ask you some questions?" The therapist said he would be happy to answer

any questions the patient had relating to his treatment. The man then asked how old the therapist was, was he married, did he have children, and where did he live? Before offering any answer the therapist stated, "I am wondering why it is important for you to know the answer to these particular questions." The patient responded in a matter-of-fact tone as though the reason was perfectly obvious. "If I am to model myself after you, I must know everything about you! I must know details."

A primitive phantasy of being deprived of essential oral supplies, akin to the famine of fairy tales, emerged in his delusion of an impending drought. He imagined his relationship with the therapist as equivalent to swallowing him whole; that is, indiscriminately absorbing attributes of the therapist's identity into himself. Before swallowing the therapist and taking on his identity (just as the wolf swallowed Little Red's grandmother and assumed her identity), the patient sought to determine in detail what the therapist was made of, so as not to ingest a toxic persecutor inadvertently. Indeed, the devil is in the details.

In psychosis there is regression to Fonagy's "psychic equivalence" mode, in which what is felt to be real is experienced as being real. Segal called this cognitive process "symbolic equivalence"; other writers have referred to "concrete thinking" or the experience of projected parts of the personality as "things." Segal describes a women who, when nonpsychotic, wrote a story depicting her mother and the analyst as being *like* a Lancashire witch (a fantasy), but when psychotic, believed the analyst *was* a Lancashire witch (a concrete metaphor that constitutes a delusion). Franz Kafka's famous story *The Metamorphosis,* in which Gregor Samsa comes to experience himself as a monstrous insect, can be viewed as a psychotic sequence beginning with Samsa's feeling of being a nobody, an anonymous functionary in an industrial society, thus being *like* an insect, then experienced as the concrete metaphor that he *is* an insect. In the "psychic equivalence" mode, if two people or things share a common attribute, they are experienced not as being similar, but as being the same.

This primitive mode of psychic functioning can be observed in young children trying to figure out what categories different objects belong to. One imaginative 3-year-old who was delighted to receive a paper origami bird requested "More ohwogamee!" but burst into tears when offered an origami flower. "That not ohwogamee!" Are all instances of "ohwogamee" avian? Does a folded paper bird belong to the category "paper" or "bird"? In psychosis, instead of figurative metaphorical thoughts of the "X is like Y" variety that form a basis for consensual symbols, individuals aggregate disparate objects and people that share some quality of emotional significance to themselves in a single identity, like a child sorting experience according to categories different from those of most adults.

Such "X = Y" statements, when reported to clinicians, are seen as delusions. For example, a man diagnosed with schizophrenia reported that

a nurse's aide on the inpatient unit, who had a last name with four letters in common with his mother's name, was in fact his mother in disguise. This primitive type of cognition forms a basis for a variety of delusional beliefs and ideas of reference in psychosis.

In his description of what others have called "concrete thinking," Arieti describes how in psychosis, thinking regresses to "paleologic" or "predicate logic" (according to the Von Domarus principle), where parts or attributes of a person or thing become equated with the whole. This regression in psychosis has a psychodynamic purpose. When people in psychotic states try to think about their lives and feelings, clear thinking that maintains discernable symbolic emotionally charged connections between thoughts and events can be unbearably painful. Concrete thinking damps down the symbolic emotional resonance of thought, thereby blurring its clarity and emotional significance, making it harder for the patient to think clearly to a painful endpoint.

Concrete thinking also turns attention away from intolerable internal experience, outward to the perception of concrete objects and external entities. But the suffering that the patient wishes to avoid nevertheless persists in disguised form, hidden in plain sight in the suffering associated with psychotic symptoms. A psychotic person continues to suffer mightily, but achieves a modicum of relief by displacing the suffering away from the core interpersonal relationships that gave rise to the pain and projecting it outward into mental representations of the real world.

For example, a psychotic woman who was conflicted about her wish for a life insurance payout following her husband's death developed a delusion that her father-in-law intended to kill her to steal the money. Hospital staff were unable to confirm the existence of any insurance policy, and there was no evidence the father-in-law wished her harm. She reported seeing a red car drive by her home and concluded that the driver was her father-in-law (who had a red car), circling in the neighborhood with plans to kill her. Her paleological equation of any driver of any red car with her father-in-law degraded her ability to think clearly in her bereavement. Instead of feeling her pain over her husband's failure to provide a comfortable inheritance for her, she dealt with her guilt over her wish to profit from her husband's death by becoming preoccupied with an external danger. Greed and murderous feeling have gone missing from her mind and reappeared in the imagined mind of her father-in-law.

In the 4th and 5th years of life, the "psychic equivalence" and "pretend" modes evolve into a more fully developed "theory of mind," in which the child can distinguish between dream images, thoughts, and real things, and appreciate that others may have thoughts and feelings that differ from those of the child. The stage is set for the three-person psychology of Freud's oedipal period, where the child registers the parents' relations to each other and also begins relating to a wider network of people. Figure

33.16–1 depicts the normal course of development of children: mental life before age 3 is dominated by primitive phantasies and metaphors for body processes, especially taking in external objects and extruding internal ones. As the child grows up, the fantastic, primitive two-person psychology of early childhood is gradually replaced by more realistic mental representations of others.

Examples of the disappearance of fantastic conceptions of the world in favor of more realistic representations include the disappearance, normally around age 8, of beliefs in Santa Claus, the Easter Bunny, and the Tooth Fairy. Just as the Wizard of Oz is revealed as an ordinary man operating a hand crank, not a phenomenal being wielding a magic wand, such entities become understood as parental inventions. This change is shown in Figure 33.16–1 as the gradually widening swath of more realistic mental representations of others, progressing through latency and adolescence into young adulthood.

The Impact of Less than Optimal Family Relations in Solidifying Primitive Unconscious Phantasy

Tolstoy's observation that "Happy families are all alike; every unhappy family is unhappy in its own way" might be rephrased here as "Happy children are all alike, but unhappy children who are vulnerable to developing a psychotic illness are each unhappy in their own way." Because the failures in development that make people vulnerable to psychosis can take many forms, the roads to psychosis are as varied as there are different patterns of troubled development. Just as no single gene has been identified to account for all instances of psychosis, no single psychological dynamic has been implicated in all cases of psychosis. There are many fault lines along which the mind can shatter. In some cases there is a massive failure in the early caretaking environment that ushers in the psychosis, or an evident trauma such as physical or sexual abuse, rape, or persistent bullying at school, but in many cases no such massive trauma is apparent.

Arieti observed the so-called "schizophrenogenic mother" (a regrettable term coined by Fromm-Reichmann) in less than 20 percent of families, and found that most parents of children who became psychotic were loving and concerned. Karon makes the important point that psychological processes in parents that may predispose to psychosis in children often operate outside the parents' conscious awareness. The caregivers fail to give the child what is needed but do not harm the child intentionally; in such cases both child and parent may be victims of the parent's unconscious psychology. In most cases, the erosion of good-enough development occurs gradually, one disappointment at a time, one failed interpersonal exchange after another, until the young person's psyche can no longer hold on. After having failed to achieve a positive standing in the peer group despite repeated efforts to do so, the about-to-be psychotic person one day

surrenders and says simply, without much fanfare, "I'm done. I give up."

In her account of her psychosis and psychotherapy with Fromm-Reichmann in her book *I Never Promised You a Rose Garden,* Joanne Greenberg describes her mother and father as loving parents desperate to help her, but not knowing how. In her case, her psychosis issued from the accumulation of many factors—her younger sister's birth, traumatizing urological surgery at age 5, the impinging presence of her charismatic, demanding maternal grandfather (who to a great extent defined her parents identities and accordingly hers), a move away from a familiar neighborhood where she felt safe, and the antisemitic bullying she encountered in summer camp.

YOUNG ADULTHOOD

AGE 6 YEARS
- **Three person (Oedipal) psychology**
  - Primary anxiety-body injury; guilt, shame
  - Primary defense: repression
  - Fonagy-mentalization
  - Klein-depressive position

AGE 3 YEARS
- **Two person psychology**
  - Primary anxiety-loss of mother; shame
  - Primary defense-projective identification
  - Bowlby-attachment
  - Fonagy-pretend make believe
  - Klein-depressive position
  - Mahler-rapproachment and individuation
  - Erikson-autonomy

AGE 18 MONTHS
- **Psychology of the primal self**
  - Primary anxiety-annihilation of self
  - Primary defense: splitting, projection, denial
  - Fonagy-psychic equivalence
  - Klein-paranoid-schizoid position
  - Mahler-symbiosis
  - Erikson-basic trust

Stages of Development

PSYCHOSIS: Primitive Object-Related Phantasies Re-emerge in Young Adulthood
The Psychotic Man Who Feared a Drought
"If I am going to model myself after you I must know everything about you."

More realistic multi-dimensional images of other people as separate individuals. The ability to "mentalize" the feelings and points of view of other people.
Hansel and Gretel/Little Red Riding Hood

Diminishing influence of primitive internal object - relations.
The ability to pretend and play with reality.
"Mother's shoes might have eaten me."

Primitive internal object-relations consisting of split-off fragments of the self are projected into mental representations of others.

Internal Object-Relations

FIGURE 33.16–1. Stages of Development and Internal Object-Relations.

Greenberg describes her retreat from the real world, where she found it impossible to live, into a psychotic world she called Yr, inhabited by a host of magical characters with their own language, personas as beautiful, compelling, and terrifying as in any fairy tale. She hears the "voice" of a hallucinated being named Anterrabae, who says to her "Bird-one, you are not of them," which confirms her conclusion that she is impossibly misfit to the real world. As she descends into psychosis, her sense of self disintegrates, and her mind drifts away from verbal thought that would require her to think clearly about her pain, toward a succession of images and sensations characteristic of the preverbal mental processes of early childhood.

A black wind came up. The walls dissolved and the world became a combination of shadows. Seeking for the shadow of firm ground on which to stand, she was only deceived again when it

warped away like a heat mirage; she looked toward a landfall and the wind blew it away. All direction became a lie. The laws of physics and solid matter were repealed and the experience of a lifetime of tactile sensation, motion, form, gravity, and light were invalidated. She did not know whether she was standing or sitting down, which way was upright, and from where the light, which was a stab as it touched her, was coming. She lost track of the parts of her body; where her arms were and how to move them. As sight went spinning erratically away and back, she tried to clutch at thoughts only to find that she had lost all memory of the English language and that even Yri was only gibberish. Memory went entirely, and then mind, and then there was only the faster and faster succession of sensations, unidentifiable without words or thoughts by which to hold them. These suggested something secret and horrible, but she could not catch what it was because there was at last no longer a responding self. The terror, now, could have no boundary.

How can one conceive of the ways adverse experiences trigger psychosis? To be psychologically significant, an event must, by definition, have personal meaning for an individual, whether that meaning is good or bad. Because the unconscious phantasies laid down in childhood become the foundation of the adult's system of meanings, inferences about significant life events are inevitably linked to, and elaborated by, unconscious phantasy. If someone passes a test, the world appears to welcome them; if they fail a test, the world turns away, and so on, in the ongoing, never-ending cascade of meaning that is the human psyche. Unconscious phantasy always operates in the background of conscious experience, inevitably shaping one's thoughts, feelings, and actions over the course of an ordinary day. Most such phantasy is not experienced as an ego-dystonic intrusion, but rather as a source of value, meaning, and goal directedness.

When people have a sense that they are *over-reacting* to an event, they know they are experiencing the intrusion of an unconscious phantasy that is emotionally coloring their view of the event. They know that what has just happened is somehow not as bad as they have felt it to be. With cognitive effort, and in a brief time, people can usually temper their over-reaction. The unconscious phantasy that lent its overcharged emotions to the experience fades, and people go on with their day. But when a traumatic event or succession of events kindles powerful emotions that resonate with primitive unconscious phantasies that have been reinforced by early failures of development, people may become convinced that a scene expressive of their primitive unconscious phantasies is actually happening now, in reality. Any feeling person can empathize with the reactions of helplessness, rage, fear, and betrayal of someone raped by a parent or sexually abused by a priest—or both, as was true for Karon's patient.

But in psychosis the enormity of what has happened goes even deeper. A child who desperately wants a parent's love and needs the parent's protection with the intensity that only a vulnerable child can feel, who is instead used like a condom and thrown away, may not be able to bear such events within an integrated experience of self. As the trauma resonates with the deep object-related phantasies of early childhood, the mind begins to

vibrate with pain, then shake, then fall apart into pieces. At the deepest levels of the psyche, where the most joyous phantasies of loving others should ordinarily reside and give hope at times of adversity, a sort of wholesale genocide of good psychological objects occurs. The psychotic person's mind then becomes a fragmented collage of fantastic figures that tend to be all good or all bad.

This experience of one's worst unconscious anxieties coming true is accompanied by a pervasive terror that has been termed *annihilation anxiety*. On the brink of psychosis, the self sees itself beginning to fall apart and disintegrate. As individuals descend into the "psychic equivalence" mode, where what they feel is experienced as real, primitive unconscious phantasies surface in consciousness as delusions, hallucination, or other psychotic symptoms, whereby the psychotic person's thoughts and feelings are expressed in a psychotic "fairy tale," in idiosyncratic concrete metaphors, and in altered perceptions of other people and the outside world.

The author once supervised the psychotherapy of a young woman, who routinely sat down to what on the surface was expected to be a warm family dinner where one might be expected to say such things as "How was school today?" or "Can you pass the potatoes?" knowing that within three hours her father would likely be raping her. When she was unable to bind this experience within an integrated view of herself and the world, her mind fell apart, first into two pieces in addition to the self, and later into four. She began to hear the "voice" of a "good angel" on her right shoulder, and a "bad angel" on her left. The bad angel urged her to kill herself, the good angel to stay alive. As she descended further into psychosis, she came to hear three distinct "voices," that of "Darryl," who combined traits of her violent father and a fantasied older brother who might have defended her; "Baby," who cried and was weak, who represented the needy child who had been betrayed; and "Tamika," who was brash, irreverent, street-wise, so tough she could not be hurt. She described how the Tamika "voice," like in a professional wrestling match where one wrestler tags off to another, would command her to take certain actions when she felt at a loss as to what to say, feel, or do. Her psychotherapist established a trusting rapport, which over time allowed her to put Humpty Dumpty back together again. She came to understand that the states of mind composed of her cast of hallucinated characters were aspects of her own feelings. Darryl's violence and Tamika's indifference evolved into a confident self-assertiveness that allowed her to stand up for herself without need for recourse to auxiliary internal objects.

Here are examples of how different family constellations may predispose to psychosis by impinging upon the good-enough development of children, leaving them unhappy in their own way. In good-enough development, parents support independence when the growing child shows this developmental striving but allow dependence when needs for reassurance and security are paramount, as is particularly apparent in the practicing and rapprochement phases of development described by Mahler. If this empathically alternating relationship between parent and child is kept in mind as a central arc of development that leads from the nuclear family out eventually into the adult world, the disturbances that predispose to psychosis can be viewed, broadly defined, as a disturbance in the child's

separation and individuation from the parent; that is, a failure in the child's becoming an authentic, self-inhabiting person.

In the severe cases of childhood onset psychosis, florid acute psychosis in adults, and severe chronic psychosis, individuals may literally not know where other people end and they begin. More often, however, the disturbance is less dramatic but nevertheless profoundly important. Some parents need their child or imagine their child to be someone other than who the child authentically is. They impose their own mental representation of an internal object on the identity of the child. Rather than discovering who the child is as the child's identity unfolds, they decide who the child is and require the child to inhabit that identity as a condition of relating to the parent.

The father of a son might have wanted a daughter to avoid the competitive feelings he might experience with a son. Or a mother with an inadequate relationship with her husband may want an enmeshed relationship with a child who will remain dependent upon her, allaying the mother's loneliness and separation anxiety. Or a father may have wanted an athlete, but the son prefers to draw, or a mother who has lost a child in infancy may subtly communicate to another child in the family the parent's wish that it had been the other child who had died, or the next child born may be viewed as a replacement for the dead child, and so on. The possibilities for mismatches between the child's authentic identity and the imposition of a parent's need are many. These sorts of stories routinely emerge in the individual psychotherapy of psychosis.

Theodore Lidz described two basic patterns in which the psychology of one or both parents impinges upon and dominates the development of the child, patterns he termed the "skewed" and the "schismatic" family, respectively. The former serves the needs of one of the parents, most commonly the mother. The other parent and the children accept the dominance of one parent's needs, knowing that servicing them is necessary to the parent's psychological survival and therefore the survival of the family. In the latter, both parents attempt to impose their own differing and distorted view of the world on the family, and no consistent consensus is achieved. When two very different parental agendas are imposed upon the identity of the child, the child has difficulty forming a stable identity.

Arieti described a psychotic patient who grew up in a schismatic family environment. His father was a war hero killed in World War II. The father's loss so crippled the patient's mother that she was unable to invest emotionally in her son. The patient translated this seeming lack of interest into a feeling of his being a bad person. His maternal grandmother, however, encouraged him to believe he was an exceptional person who would grow up to be a hero like his father. In the face of these two contradictory imposed identities, the patient grew up with no core identity and no realistic self-esteem, a condition that left him vulnerable to psychosis.

Another patient Jason grew up in an atmosphere of constant bickering and fighting between his parents, who frequently urged him to take sides. The parents, who divorced when he was 16, forced him to decide which parent he wished to live with, making no effort to arrange an amicable joint custody. He chose his father, who promised Jason a motorcycle when he was 18.

The father had little interest in his son except to demonstrate the son's preference of him over his wife. Jason soon became disenchanted, and one evening returned to his mother's house, hat in hand. At first she barred him from her home, haranguing him about how he had betrayed her by choosing his father, saying he must live with his choice. She eventually relented and allowed him in, but she said, amidst angry declarations of what a monster the father was, that nothing would please her more than to read in the newspaper that Jason's father had been murdered. Jason got the message: The price of his return was to be his father's life. He was psychotic by daybreak. Both parents viewed Jason as belonging either to themselves or to the other parent, and if the other parent, their mental representation of Jason was infiltrated by their hatred of their ex-spouse.

Karon describes a mother–child dyad in which an insidious mismatch arose around feeding. A psychotic man who went on hunger strikes and believed he was being poisoned reported, "The Athenian girls are laughing at me. They say their breasts are poisoned." There was at first pass no apparent trauma in the patient's history. His mother, who had breast fed him for a year and weaned him gradually, asked her son "Did I give you enough milk?" He replied cryptically, "The cow gave her calf milk and then kicked it. She shouldn't do that. It's something that happened hundreds of times in the history of the world." Karon inferred that the cow represented his mother, the calf represented the patient, and the history of the world represented his life history. Thus, the patient was complaining that every time his mother fed him she got angry, or "kicked" him. This pattern was apparent in their relationship as adults. If the mother cooked a meal and he ate it, they quarreled afterwards. If she cooked and he was on a hunger strike, or if someone else cooked, there was no quarrel. When the patient was a child his need to be fed produced anxiety in the mother, which made her angry. The same sequence occurred in their adult relationship, in which she would find pretexts for a squabble after feeding her son.

The essential aim of PDTp can be stated rather simply in relationship to the theory outlined above. The therapist tries to help patients tell their stories, to say what has happened to them in their lives and how this has made them feel, and what thoughts and feelings are currently being expressed in the psychotic symptoms. In this sense, each patient's psychotic symptoms are a Rosetta Stone by which one person's emotional life can be translated into simple human terms.

TECHNIQUE

As noted, most psychoanalytic theorists would agree that psychotic mental states represent the intrusion of primitive psychological processes into adult waking life, and that those processes were originally most clearly explicated by Melanie Klein and her followers. Although Kleinian analysts made brilliant contributions to the psychoanalytic understanding of psychosis, their way of working using classical psychoanalytic technique has no direct application in the public mental health sector. But if one regards psychoanalysis as a body of ideas applicable in a variety of settings, not as a specific clinical technique, Kleinian ideas are a compass in psychotic chaos, invaluable tools for understanding the psychotic process and guiding care wherever it occurs with whatever technique is applied. This chapter begins with a short review of Kleinian technical ideas, as they illuminate some general principles of therapy for psychosis, followed by the recommendations of therapists of the American Sullivanian School.

The European Kleinian School

Segal describes the treatment of a chronically psychotic man with multiple delusions and hallucinations who believed that he was a prisoner, not a mental patient, and that the world had been changed and was being destroyed. His two core delusions, that he was in prison and that the world was changing, were connected to primitive internal object-related phantasies. He complained that he was being given plenty to eat in the hospital, but not enough to do or think about, which led him to believe his head, arms, and legs were shrunken while his stomach had become huge. Instead of saying he was bored and that he felt like he was being force fed, he expressed these sentiments in a somatic delusion (a concrete metaphor). Because in the hospital he thought about little else than eating, he believed the world was obsessed with eating, to the point that people had neglected to plant new crops, which would lead to famine and the destruction of the world. After Segal consistently interpreted his primitive object-related phantasies and analyzed the psychotic transference, within a year all his delusions disappeared.

Wilfred Bion extended the work of Freud, Klein and Segal in novel ways with his concept of *attacks on linking*. In psychosis, individuals turn away from painful realities by internally undermining (attacking) the connecting links between memories and verbal thoughts that give experience meaning. Instead of thinking a verbalizable thought, psychotically disturbed individuals project the part of their cognitive apparatus that would ordinary maintain a conscious awareness of reality outside the self into what Bion called a "bizarre object." The familiar computer chip in the brain, a common delusion, is a "bizarre object." The person's cognitive capacity to observe reality has gone missing from the person's mind and has been projected into a mental representation of the chip. The chip operates within the person's body but is separate from the person's ability to think. The capacity most often projected into the chip is that of monitoring thoughts and transmitting them to an external agency such as the FBI, in which case the chip has taken over one's capacity to know one's own thoughts and evaluate them.

Herbert Rosenfeld believed in the importance of analyzing the psychotic transference. In his experience, primitive phantasies of intruding upon others or being intruded upon to the extent of annihilating the self are common in the psychotic transference. He describes a patient who feared that one day she would begin to speak in a strange voice that belonged not to her but to Rosenfeld. In keeping with Klein's emphasis on primitive body-related phantasies in which children imagine taking something out of or putting something into the mother's body, or having the same done to themselves, Rosenfeld interpreted the negative psychotic transference to this woman, telling her she feared that he was going to force himself into her body and her mind and take control of her thoughts so as to cause her to lose her identity entirely.

The American Interpersonal School

Harry Stack Sullivan's ideas about the etiology and treatment of psychosis are important not only because they are persuasive, but because of their simplicity. Sullivan believed that disturbances in the early mother–child relationship predispose to psychosis. According to Sullivan, all human beings have basic needs for *security* and *satisfaction*. His concept of security roughly parallels current concepts of secure attachment, positive self-esteem, and associated expectation of esteem from others. His concept of satisfaction roughly corresponds to Freud's conception of drive satisfactions, including the gratification of sexual needs. In order to achieve security and satisfaction, children must cultivate sufficient interpersonal competence to ensure that these basic needs are met in childhood, latency, adolescence, and adulthood.

According to Sullivan, when an infant's need for maternal care makes the mother anxious, this anxiety is communicated to the child, undermining the child's capacity to feel secure and to develop the interpersonal competence necessary for independent living. This failure of relatedness arises in interactions with the small group of people that compose the child's core care-giving environment. He considered the developmental milestone of having a "chum" during latency to be an important marker of mental health, as it signals the growing person's ability to establish a stable and gratifying relationship outside the nuclear family.

Sullivan died at 57 before ever writing a comprehensive manual for technique. His recommendations are scattered throughout his writings. Much of his initial work with psychotic patients was done on a special six-bed inpatient unit that he designed at Sheppard Enoch Pratt Hospital, where, without medication, he reportedly achieved an average 85 percent recovery rate for acute psychosis. Consistent with his theory, he placed significant emphasis on the quality of interpersonal contacts in the ward milieu. He took care to alert ward staff to the exquisite sensitivity of psychotic patients, encouraging them to spend considerable time fostering a relationship in which the patient could feel secure. Believing that his patients felt interpersonally incompetent in their family settings, he placed them in a humane environment in which all patients could feel themselves as valued members of the community. He resisted objectifying diagnoses while teaching that, however, strange their symptoms, people with psychosis should be considered "more simply human than otherwise." (In a similar vein, CBTp recommends *normalizing* to emphasize the continuity between psychosis and ordinary mental life.)

A central aim in Sullivan's approach was to reconstruct the chronology of the onset of the psychosis, so as to impress upon patients that whatever psychotic reaction had befallen them was the result of their profound

disappointments in life experiences with a small number of people. Sullivan cautioned that it is extremely important to use simple, clear language because people in psychotic states have difficulty communicating in consensually familiar verbal symbols. Talking in too complex or fancy a way would be, in his words, "asinine behavior." With hebephrenic patients who had retreated into an autistic self-involvement, or paranoid patients with fixed delusions, he sought to revive a state of confusion present in an earlier phase of the psychosis. This goal is similar to the CBTp aim of encouraging a state of uncertainty in patients about their delusional beliefs. The therapist must spend sufficient time with the patient to let the treatment unfold. When the patient is not communicating clearly, the therapist should welcome the attempt to communicate but not pretend to understand what is still confusing.

Sullivan offers occasional examples of his work with patients. A psychotic man who had been severely disappointed in love said of the woman who had rejected him, she is "of the past" and "out of my mind" and someone "of no importance." Sullivan took advantage of this moment of denial to say something along the lines of, "Nonsense! You are happy with her." Having startled the patient into alertness, Sullivan continued, "and there is no reason on earth why the pleasure you had in her company should be thrown away just because the relationship didn't last forever."

If all goes well enough on the inpatient service and some degree of human contact can be re-established, allowing some degree of social recovery, the patient can be treated as an outpatient. Sullivan believed the training, experience, and personality traits of the therapist were crucial to a positive outcome. (Training and experience do matter, a fact often lost in RCTs of PDTp.)

Frieda Fromm-Reichmann worked closely with Sullivan. She distinguished *psychoanalytically oriented psychotherapy* from classical psychoanalytic technique and believed that even the most psychotic patients were capable of some degree of interpersonal relatedness. Her success as a clinician stemmed from her warmth, humility, and exquisite psychological intuition. She stressed that when the therapist inevitably disappoints the patient in some way, it will likely be experienced as a repetition of early traumas, and will provoke anger and intense hostility that can be understood and worked through. She emphasized how lonely psychotic people are, noting that in spite of their retreat from the real world, people in psychotic states nevertheless retain some dim notion of the unreality of their substitute delusional world. A longing for human contact and understanding remains, yet patients are afraid to admit this to themselves or to the therapist for fear of further disappointments.

Fromm-Reichmann emphasized that to establish trust the therapist must treat patients with kindness and consideration, without infantilizing them. She would explain to them that symptoms were ways of remodeling life experiences in consequence of, or in accordance with, their thwarted past or present relationships. She recommended telling hallucinating and

delusional patients that the therapist does not see, hear, and believe the same things the patient does, but that the two parties can investigate together the reasons for the differences in their experiences. She wanted patients to become aware of and to grieve the losses they had sustained in early life without distorting them symbolically as psychotic symptoms. This approach is reflected in the title of Greenberg's book *I Never Promised You a Rose Garden*. Interestingly, at one point in Greenberg's psychotherapy Fromm-Reichmann used CBTp-like technique to array evidence against her patient's belief she had tried to kill her sister.

Silvano Arieti won the National Book Award in Science in 1975 for his seminal work *Interpretation of Schizophrenia*. This 756 page book is arguably the most comprehensive and coherently organized text currently available on PDTp, with Karon and VandenBos's *Psychotherapy of Schizophrenia* a close second. Sullivan, Arieti, Fromm-Reichmann, and Karon share much in their clinical style, an approach which might be called an *interpersonal interpretive technique*. What follows is a description of this technique following Arieti's outline of three phases of treatment, written as an amalgam of his approach and others.

Interpersonal Interpretive Technique

Arieti conceptualized PDTp as proceeding in three stages. First, the *engagement* phase, in which the therapist attempts to connect with a person who has largely withdrawn from social contact and consensual reality. Next, a phase that focuses on showing patients how they express themselves in a variety of *psychotic mechanisms,* such as hallucinations, delusions, and ideas of reference. Specific techniques are required to bring the thoughts and feelings experienced as perceptions back within the boundary of the self, where they can be experienced as mental events and examined in more traditional psychodynamic fashion. CBTp techniques that can be a stand-alone practice can be extremely useful into this second phase, about which there is more in the CBTp portion of this chapter.

Third, when patients have learned to identify the psychotic mechanisms that disguise their thoughts and feelings, they can begin to reclaim their inner life as the therapist interprets the meaning of symptoms in relationship to disowned emotions, memories, distorted thinking, and current life events. The therapist attempts to help the patient see his psychotic symptoms as meaningful reactions to problems in living, without accepting the psychotic distortions or promising the patient compensation or an ideal world. These three phases of treatment will now be described in more detail.

Engaging the Patient and Establishing a Therapeutic Alliance

Establishing a trusting emotional relationship with a patient suffering

psychotically is the first and most important step in the psychotherapeutic process. The patient is caught in a devilish version of Schopenhauer's porcupine dilemma, in which too much closeness with other human beings brings the prick of the quill, and too much distance the chill of a cold that withers the spirit. Having found that attempts to relate to others have in the past occasioned intense anxiety and repeated interpersonal failures, the psychotic person withdraws from consensual reality and ordinary social contact into a solitary delusional world. Patient who can be engaged in psychotherapy must find a way to set the value of a relationship with the therapist against the unaltered continuation of their psychotic existence.

All clinicians, whether in psychiatry, internal medicine, or surgery, will see in their patients a balance of health-promoting factors and pathology. The clinician will not know all the variables involved, or their ultimate strength in the clinical equation defining outcome. The same is true in the initial engagement phase of therapy with patients suffering psychosis. The therapist cannot always know beforehand how the need for human contact will play out against the patient's aversion to the real world. Patients who lack basic trust will approach all relationships, including a relationship with the therapist, with trepidation. When a person's development has been significantly impinged upon by others, the patient may experience being influenced by the therapist in any way as an impingement.

For example, a psychotic woman revealed after months of psychotherapy that she did not think about what the therapist said during the sessions; instead, she tried to memorize the therapist's words so that she could digest them later, alone, when she felt she could exert more control and avoid the therapist's ideas taking over her mind. Patients who have had a symbiotic relationship with their primary caregiver may adopt an opposite course and rush to fuse with the therapist's identity in an attempt to reserve no mind of their own with which to feel their own pain.

The minimum conditions for psychotherapy require the patient to be able attend sessions, whether as an inpatient or outpatient, and to curtail violent behavior directed toward themselves, the therapist and others. The therapist cannot allow socially intolerable behaviors such as stealing from the therapist, inappropriate touching, or significant disruptions of the therapist's life. A minority of patients, even when properly approached, are too damaged or too terrified to engage in treatment, at least within the parameters of the current mental health system. The therapist needs to strike a flexible balance between authoritarian controls over the patient's behavior and sitting passively by while the patient contemplates some ruinous action.

Confidentiality should be discussed early in treatment. With respect to parents, significant others, and public authorities, the therapist should explain that the treatment will remain confidential except when the therapist believes that the patient presents a clear danger to self or others,

in which case the therapist may need to contact family, friends, or the authorities. If a patient cannot allow the therapist to call in help if need be, the patient may be best treated in a more structured clinical setting. The therapist can share a draft of insurance forms, psychosocial summaries, and disability applications prior to releasing confidential information.

Patients with psychosis do not necessarily expect to have rewarding conversations with mental health professionals. They may find that the delusions and hallucinations that are the most compelling concerns in their lives may be brushed aside as "imaginary" by others. After establishing rapport with his therapist, a psychotic man acknowledged that the most important relationship in his life was his ongoing dialogue with a female "voice" he believed to be the voice of God. He said, "I never tell the doctors much at all about God. I know what they are going to think! They will just think I am crazy. They don't want to talk. They are just Haldog (*sic*) terrorists. So I take the medication and then after a while they let me go, and I stop taking the medication after I leave. I trust you. Maybe I shouldn't, but I do. There is something about your way of speaking, something about your manner, that makes me trust you."

Establishing rapport is not a mechanical matter of technique, but fundamentally *it is a matter of the heart*. The therapist's relationship with the patient will reflect what the therapist genuinely thinks and feels about the psychotic person. Just as children readily discern when meeting an adult for the first time whether it feels safe to proceed, psychotic individuals will know by the therapist's words, gestures, facial expressions, and tone of voice, whether the therapist's greeting is a well-intentioned but generic professional posture, or whether the therapist is motivated by an affectionate, abiding respect for the patient's individual personhood. Patients must know that they are truly seen by the therapist as discrete individuals. Unlike nonpsychotic patients, who can explore their fantasies about who the therapist is, the psychotic patient cannot wait to know, for fear of the worst. Adrift in a sea of madness, the psychotic patient must find something real in the therapist to hold on to that is sufficiently reassuring to allow the patient to risk connecting.

Greenberg illustrates the sensitivity of psychotic patients to the therapist's true feelings in her description of the reaction of inpatients to a conscientious objector who was unhappily assigned to the hospital as a therapy aide in lieu of military service. The patients sensed that he was frightened of them and that he defensively held them in contempt. "His rigid fundamentalist beliefs made him see insanity is a just desert for its victims, as God's vengeance, or as the Devil's work, and sometimes as all three at once. As the days passed, his fear waned in time and his righteous wrath was at hand ... Against his loathing, the sick fought in their sick way. The literate rewrote the Bible or ridiculed its passages to make him horrified. Constantia made flagrant sexual advances to him. And Debra made a few pointed observations about the similarity between psychotics and religious fanatics."

The seeming failure of people around them to see the world as they do leads psychotic patients to believe that at best the therapist will not understand, and at worst the therapist is directly involved in the patient's delusions. The therapist's attempt to communicate effectively is paramount. Initially the therapist should not ask the patient any

complicated questions, because a person who is desperately fighting for psychological survival will be unable to attend to the therapist's queries. A patient's wish to talk to the therapist is a good sign, no matter the subject. When the patient is unable to talk directly about the psychosis, the therapist should encourage the patient to talk about something else.

When the therapist does not understand what the patient has said, the therapist should not pretend to understand, but rather say the therapist is trying to understand but is not yet certain what the patient is trying to say. When the patient cannot or will not speak, the therapist must be able to tolerate periods of silence, but the therapist should not allow protracted silence to develop. Based on whatever the therapist knows about the patient, the therapist can keep talking, musing, meandering, speculating, allowing the patient to feel the therapist's effort to communicate. The therapist should be emotionally engaged, firm, honest, and familiar in his manner as opposed to emotionally distant.

The therapist need not withhold direct reassurance. Arieti describes his approach with a patient with postpartum psychosis. "In many cases, I have told the patient that she is in great distress because she rejects the baby, she feels guilty about it and believes that the baby is going to suffer. I have reassured the patient by telling her that she must accept the fact that she rejects the baby and that the baby will not suffer. Adequate provisions are being made for the baby. We shall eventually discover why she feels this way and her feelings will change. In the meantime, she cannot feel guilty because she cannot help rejecting the baby. I have found that even very disturbed women understand these explanations, partially or totally, and are relieved by them." Karon routinely reassured terrified patients saying, "I will not let anyone kill you."

The therapist should ally with the nonpsychotic part of the psychotic patient's personality that remains capable of standing outside the psychosis and reflecting on it, thin ledge along a deep chasm though this foothold may be. The therapist should not collude with the patient's delusional beliefs, nor should the therapist side-step the issue of the patient's being ill, though the matter of what to call the patient's condition must be handled with finesse. The patient's condition may have been called "schizophrenia," "mental illness," "bipolar disorder," or "a chemical imbalance." It is important to understand the labels that have been applied, the meaning of these descriptors to the patient, and what words the patient might use to describe their condition. How the label is worded shapes patients' self-image and the way they believe the therapist will see them. Patients will sometimes talk comfortably, for example, about an "emotional breakdown" while being allergic to the word "schizophrenia."

A woman with a delusion about a man at her church was willing to consider that she might be mistaken about his intentions toward her, but she was absolutely certain that the label

"schizophrenia" did not apply to her. She had looked up "schizophrenia" on the Internet and had taken away the idea that "schizophrenics imagine things that are not real." This could not possibly apply to her, she thought, because she had no doubt that the man at church had in fact said certain things to her, as indeed he had. Although she granted that she may have misinterpreted what he said, she had certainly not *imagined* him saying what he said. She had understood the diagnostic label "schizophrenia" to apply to people who imagined things that never happened rather than people who misinterpreted things that did happen. The fact that her doctors had given her a diagnostic label that she believed clearly did not apply to her led her to conclude that they must have some malicious purpose. The therapeutic alliance foundered until this issue was identified and worked through.

Although it would not be accurate to characterize psychosis as a freely chosen, alternative way of living, people who have recovered from psychosis and written autobiographical accounts of their illness describe needing to give up on an impossibly painful real world and to surrender, subjectively voluntarily, to psychosis. Greenberg describes her escape from the real world into her delusional Yr, where in her mind she lived apart from the pain of the real world for years. Fromm-Reichmann stood patiently at the boundary between Yr and the real world, encouraging Greenberg to step across the boundary back into reality when she could, accepting Greenberg's need to step back into Yr when she needed to.

People with psychotic illnesses may feel they have less to lose than nonpsychotic individuals because they have already lost so much. Having already given up on life, such patients may feel freer to tell the truth about what they think has been going on in their core relationships, including with the therapist. When a patient with a psychotic condition comments accurately on something the therapist has said or done, the therapist must acknowledge the truth of the perception. As Arieti put it, "... on the one hand [the psychotic person] is very sensitive and is capable of seeing through a situation to perceiving the truth even more so than the normal person; on the other hand, what he does with what he sees is so distorted that it will increase, rather than decrease, his difficulties."

If attempts to engage the patient are successful, the clinician may recognize a "breaking through" from the profound isolation of psychosis. Arieti quotes a patient who describes such a breakthrough moment. "There was a reading room on the ward, with a piano and magazines. I went there and I looked at the cover of a magazine. The nurse stepped in. I had never seen her before. She started a conversation with me as if I were a normal person. I told her with tremendous emotion, 'You're the first person who has broken through to me.' She was not on guard. I was not on guard. She was an ordinary girl. She made me feel communication with people was worthwhile. Before that nothing was worthwhile. People were hateful."

Identifying Psychotic Mechanisms

Arieti distinguished between *psychotic mechanisms,* meaning the form the symptom takes, and the content and psychodynamic meaning of the psychotic symptom. For example, auditory hallucinations may be similar in form from patient to patient, but different in content and personal meaning. Arieti helped patients recognize when these psychotic

mechanisms were operating in their minds. CBTp-like techniques were thus essential to Arieti's method. Through a process that CBT therapists would now call "guided discovery" of a cognitive bias Arieti helped patients understand that they heard "voices" when they *expected* to hear them: "A patient goes home after work and expects the neighbors to talk about him. As soon as he expects to hear them, he hears them. In other words, he puts himself in what I have called '*the listening attitude*.'" The patient comes to see that there is an internal mental contingency to what would otherwise be experienced as an external event. A similar realization happens when patients in CBTp understand that their "voices" follow stress or lack of sleep.

Even after recognizing that their expectation of abuse leads them to imagine an abusive attitude in strangers, or to trigger critical "voices," patients may remain puzzled as to why specific charges have seemingly been brought against them. They have no conscious awareness of ever having had thoughts about themselves of the sort given voice in the attack. Klein's ideas about primitive splitting and projection are extremely useful in understanding and working with this state of mind. In psychosis, repudiated thoughts and feelings literally go missing from the patient's mind and reappear as infiltrations of the patient's perception of reality. The patient who reports being unaware of ever having any thoughts similar to the persecutory charges is telling the truth. Arieti offers the example of a woman who heard "voices" telling her she was a prostitute, who was puzzled because she had no conscious awareness of thinking such thoughts about herself, until in PDTp she became aware of a variety of sources of poor self-esteem that had condensed in her mind into the "prostitute" charge.

Arieti approached ideas of reference and delusions with a technique similar to his approach to auditory hallucinations, by helping the patient identify what he called a "referential attitude." A man enters a public park feeling a vague, abstract sense of foreboding. Once in the park he notices children running in a direction opposite from him, and he jumps to the conclusion that he is being perceived as a child molester and the children are running from him. What Arieti called a "referential attitude," CBT therapists would call a jumping-to-conclusions bias and a self-referential bias.

Arieti taught patients to recognize the vague antecedent referential attitude that preceded and became concretized in ideas of reference. PDTp attempts to reverse this sequence, to move from the perception of concrete exterior events to the mental interior of more specific thoughts and feelings, and finally to the developmental origins of these feelings.

For example, Arieti describes a patient who believed she was getting Morse code messages she could feel when she palpated pulsations in her wrist that gave her directions about what to do and not to do to remain safe. Arieti interpreted that that she was actually giving herself advice,

but she needed to experience the directives as originating in some august authority outside herself. He related this symptom to her parents never having provided the confident guidance she needed to grow into an independent person.

In some cases, the patient's delusions reflect a sensitive awareness of an actual social reality. Arieti calls this a "punctiform insight." He gives the example of a compliant woman who was in fact being overworked and taken advantage of by her employer, who received a present of yellow roses from her boss for her birthday. She associated the color yellow with jealousy and in a turn of "paleologic," she decided that yellow roses had been selected to show her that her office-mates knew she was jealous of the boss's wife. Arieti interpreted that she was in a state of anxiety at work relating to her self-doubts and poor self-esteem, and that she had used a special type of thinking to express her feelings of unworthiness, in which she attributed her thoughts about herself to others. The boss did in fact appear to be hostile and demanding, and it did in fact appear she was being exploited, but her punctiform insight into this social reality had been expressed in a widening circle of paleological ideas of reference in which the entire social surround was seen as being hostile.

Interpreting the Relationship between Psychotic Symptoms and the Patient's Current and the Past Life

Assuming that the therapist has been able to engage the patient and has made progress in the second phase of treatment in helping the patient to identify psychotic mechanisms and not regard their content reflexively as reality, the therapy can turn gradually toward a more traditional psychodynamic approach. Impingements upon the good-enough development of an authentic independent self have left the psychotic person with a socially incompetent personality incapable of meeting the person's adult needs for security and satisfaction. Painful, repressed elements of the person's mental life have so distorted the psychotic person's mental representations of other people and reality that it has become impossible for the person to form gratifying relationships with others. Psychotherapy offers patients the opportunity to rethink these distortions so as to do better with other people and have a second chance at finding their lost authentic self.

The therapist tries to create a psychological holding environment in which, session by session, the patient's authentic thoughts and feelings can emerge in treatment. Week after week, the patient speaks to the therapist of thoughts, feelings, and actions, and the therapist helps the patient by interpreting the unconscious meaning of the patient's productions, including especially the patient's psychotic symptoms The patient may say things that are fragmented and initially incomprehensible. The therapist must gather up these mental shards over time to make coherent sense of what the patient is saying. The therapist's genuine concern and ability to

bear empathic witness to the patient's emotional life provides a corrective emotional experience that offsets a childhood environment that failed to bear witness to the child's true self.

The third phase of PDTp, while similar in broad outline to work with nonpsychotic patients, poses unique challenges. The nonpsychotic person in psychotherapy has the challenge of trying to find links between an affectively charged experience in the present and an affectively charged experience in the past; such patients come to see that the way they are feeling in the present is an expression of a childhood state of mind. Once nonpsychotic persons can see that the possibilities of the present have been constrained by the past, they can work themselves free of the past and return to the full possibilities that life offers in the present. The psychotic person's task in PDTp involves an additional intermediate step that the therapist must help the patient accomplish. In psychosis, affectively charged experiences from the past are experienced as altered perceptions of the outside world that are elaborated in delusional accounts of reality. The psychotic person in psychotherapy must first translate a current perception, or a belief regarded as a real perception of the objective world, into an affectively charged thought and/or feeling about people and the world, and then link that affectively charged experience with an affectively charged experience from the past.

THE MAN WITH PEOPLE INSIDE HIM. Consider, for example, a typical sort of session that might be occurring in any number of public clinics across the country. The author supervised a social worker who had been seeing a 22-year-old man who had become psychotic at the age of 17 for supportive sessions once a month. He had been tried on adequate doses of several neuroleptics with no seeming impact on his core delusion that people he did not know were having sex inside his body, an activity he found very distressing. The patient reported vague somatic sensations that he took as indications of the movement of other people's bodies inside his body. The therapist was encouraged to see the patient more often and to take a careful developmental history, which he did.

Lo and behold, the therapist discovered that the patient had been seduced into an ongoing sexual relationship by his older sister when he was 8 years old, a fact not known previously, which appeared nowhere in the patient's chart. In supervision, the author suggested that the therapist might consider the patient's perceptions of movements inside his body as a kind of sexual daydream in which the patient was experiencing his thoughts and feelings as perceptions. Following up on the history he had obtained, the therapist remarked to the patient in an offhand manner that most young men his age have fantasies about sex from time to time, and asked if it was possible that his sense of movement inside his body, which he thought was sexual activity, might be a form of sexual daydreaming?

The patient took readily to this suggestion. As if to say "Now that you mention it," he shared with the therapist for the first time some doubts he had about his delusion. He was at a loss to understand how the people having sex entered and left his body without leaving scars. He acknowledged that when he imagined what the woman inside him might be like, he imagined her as having some of his sister's attributes. After exploring his memories, thoughts, and feelings about his relationship with his sister in four subsequent sessions, the body sensations faded, his delusion receded, and his interest turned to how he might get a job.

In psychoanalytic terms, his sister's early sexual overstimulation was an impingement upon what should have been a slower, more natural development of his sexuality. At age 8, he was not emotionally prepared to reject his sister's advances, nor was he able to absorb them emotionally. Unable to integrate these highly charged experiences within his latency age

identity, he found them persisting in the form of unabsorbed introjects that he experienced as moving around inside his body. Consistent with primitive object-related phantasy that imagines the movement of internal objects in and out of the body, the patient believed that two people were moving in and out of his physical self. This is an example of the meat-and-potatoes sort of psychotherapy work that can occur in public clinics where earnest, well trained clinicians are supervised by a clinician with more experience doing PDTp or CBTp.

The therapist will want to accumulate a chronology of significant life events and interpersonal relationships with the aim of relating this history to the patient's psychotic symptoms. It will likely not be possible to take an orderly history in a single sitting. In modern treatment-as-usual, the life histories of psychotic patients are sometimes reduced to a clinical shorthand that may be adequate for pharmacological treatment, but falls short of what is needed for PDTp. For example, one might read in a chart, "This is the fifth psychiatric hospitalization for this 26-year-old male with a diagnosis of chronic paranoid schizophrenia who presents with auditory hallucinations telling him to kill himself that recurred 10 days after he stopped taking his neuroleptic." This sentence contains enough information to sketch out a bare-bones pharmacological plan: restart his medication because it worked before and discharge him when the "voices" abate. A good clinical service will address the questions "Why five admissions?" and "Why did he stop his medication?" although important issues like these are difficult to work through in a 10-day admission, and this focus is often entirely lost in the transition to outpatient services. Rarely will the origins of the patient's self-hatred, so intense as to be expressed in a "voice" telling him to kill himself, be a central concern. Rarely will his self-hatred be the focus of a treatment plan that includes PDTp or CBTp as the treatment of choice to explore this issue.

While helping the patient translate altered perceptions and concrete metaphors into more ordinary human thoughts and feelings, the therapist must also help the patient deal with the strong affects that in the first place drove the patient's thoughts out of the patient's mind and into the patient's perceptions of the outside world. Repressed affects that commonly emerge in PDTp include intense terror, rage, shame, guilt, longing, loneliness, despair, and self-hatred. Andrew Lotterman advocates the therapist's occasionally disclosing the therapist's feelings to the patient as a way of modeling how a person can contain and verbally express distressing affects. For example, if the patient were to arrive quite late for a session, the therapist might say simply, "I was worried when you weren't at our appointment on time." The therapist's modest, measured disclosure of the therapist's feelings helps the patient to keep in mind the person the therapist really is.

In psychosis, affect may cease to operate as a signal triggering defense, but may course through the patient's mind like a runaway train, becoming traumatic in its own right. The patient can be taught that no emotion, no

matter how strong or seemingly out of place, is entirely irrational. If angry, it is likely the patient has felt hurt, though the hurt may not be conscious, or may relate to something in the past, or be symbolic in nature. All felt emotion is real in the sense it is really experienced, and all emotion is commensurate and in proportion to the mental events that aroused the emotion. The therapist can aim to say, "Given your experience of those events, it is not surprising you felt as angry [frightened, needed, guilty, etc.] as you did." In keeping with the playground adage, "Sticks and stones can break my bones, but words can never hurt me," the patient can be encouraged to understand that emotion alone does not harm.

Psychotic individuals inevitably suffer from poor self-esteem, and more often from self-hatred. People who have not felt their existence empathically recognized, who enter adolescence and young adulthood lacking the social competence to meet age expectable goals in social interaction will judge themselves to be a failure and a lost cause. Many delusions directly reveal patients' poor self-esteem, as when they think they have a vile odor, or skin eruptions, or are harangued by "voices" telling them that they are worthless and should kill themselves to spare the world their presence.

For example, a young man who regarded it as his duty to lead his family out of poverty and set a positive example for his siblings believed he was a total failure after this younger brother was killed in a drug-related slaying. The patient came to believe that dogs in the neighborhood were looking through his clothing, and with their eyes were mocking his puny body. This man's self-hatred had gone missing from his own mind and reappeared in the mocking eyes of the dog. The therapist's devotion of considerable time and effort on the patient's behalf can go a long way to shore up self-esteem. Unfortunately, it is difficult to spend much time with patients in current care delivery models.

As the patient comes to doubt that his problems are caused by a delusional persecutor, the patient may mount an antiparent campaign with a similar accusatory theme. Descriptions of parents that first emerge may be malignantly distorted images of the parents, infiltrated by primitive projections quite at variance with the therapist's impression of the parents on meeting them. Because the patient has come to give some credence to what the therapist says, the therapist can point out distortions in current events related to the parents. For example, a psychotic man, who had come to understand that his perceptions of what his students said in the classroom were distorted, accused his parents of tormenting him by the offhand mention at family dinners of the names of his prior homosexual lovers. In the course of therapy, he acknowledged that his parents could never have been aware of the existence of the lovers in question, and so could not have known their names. He was able to see that events similar to the "voice hearing" intrusions in the classroom were happening around the family dinner table.

When over time the therapist has come to a more measured view of the patient's development, the therapist should acknowledge the very real limitations of the parents and the hurts they inflicted. In some cases, the patient may be told that despite the parent's actions being hurtful, there is no clear evidence the parents intended to hurt the patient in the way they ultimately did. In keeping with Klein's conception of the depressive position as including complex, integrated images of other people that

tolerate ambivalence, the therapist endeavors to help patients away from the paranoid-schizoid position to the more mature depressive position, where they can see their parents as having both good and bad traits.

Patients can also be shown that their behavior may reinforce negative interactions with parents that foreclose the possibility of making whatever connection might be possible. As is true in the psychotherapy of nonpsychotic individuals, patients with psychosis must learn that they will not get what they want by putting their parents on trial with the expectation they will now make good on what was missing in childhood. Patients must face the necessity of going out into the world and finding their own adult satisfactions. When Greenberg ("Deborah" in her book) complained to her therapist about how hard it was to live in the real world, Fromm-Reichmann responded, "I never promised you a rose garden," from which the book took its title.

Over the many months of her treatment, Greenberg worked to understand the meaning of her psychotic symptoms in terms of her life experience. Well into treatment, she revealed to her therapist that she believed herself to be a Japanese soldier. She knew about the "Japanese menace" in World War II and had heard the phrase on the radio, "If you aren't with us you are against us." Socially estranged from her peer group, she heard her "voices" repeating "You are not of them," which she took to mean that people thought of her as they thought about the hated Japanese soldiers. Instead of feeling *like* a Japanese soldier (a figurative metaphor), in her delusion she believed she *was* a Japanese soldier (a concrete metaphor). As she imagined herself into the identity of the hated Japanese soldier, her urological surgery at age 5 became a soldier's war wound, and the soldier's wish to escape was her wish to escape from a painfully overwhelming real world. At a point late in her treatment, when Greenberg had recovered her will to live, she made an additional discovery regarding the Japanese soldier. Greenberg writes:

> "It's a description of ... why ... why, it's grandfather!" Deborah [Greenberg] cried, having unearthed the familiar tyrannical Latvian to whom she had given such an unrecognizable mask. It was a description of him and it fitted him better than height or weight or number of teeth. Dr. Fried [Fromm-Reichmann] then said, "Coming to see this ... does it hurt so much?" Deborah replied, "A good hurt." Then Dr. Fried continued, "The symptoms and the sickness and the secrets have many reasons for being. The parts and facets sustain one another, locking in and strengthening one another. If it were not so, we could give you a nice shot of this or that drug or a quick hypnosis and say, Craziness, begone!' and it would be an easy job. But these symptoms are built of many needs and serve many purposes, and that is why getting them away makes so much suffering."

In addition to disclosing emotions, Lotterman advocates structured modifications of free association that he calls "naming" and "enlargement" that can help the patient get in touch with repressed emotions and put them into words, as well as role playing, in which the therapist might say "Imagine you are a poet trying to put your feelings into worlds." For

example, when a patient offers a vague description of a feeling state such as "I felt bad," the therapist might encourage the patient to linger over the statement to name more specifically the emotions involved. The therapist may note a particular phrase that appears to have a significant emotional valence, as when a patient said her feelings had been "interred. "When the therapist drew her attention to this word, she at first dismissed it as just an expression, but when encouraged to stay on point and enlarge her associations to the word, she began to talk about how cut off from people she felt and how hopeless she felt about anyone ever discovering her; that is, about her ever being disinterred.

At the beginning of treatment, the relationship with the therapist as a supportive real presence is paramount. As awareness of psychotic mechanisms increases and examination of the patient's family relationships proceeds, treatment may pay more attention to transference distortions. For example, a patient who believed she had an offensive odor recalled that the therapist had opened the window to his office when she arrived—to air the stink out of the room, she assumed. She believed that the therapist and most strangers on the street thought she had a bad smell. The therapist was able to point out that he had actually *closed* the window to shut out noise from the air well, but she had distorted her memory of his actions in keeping with her negative expectations. The patient was able to work through her conviction that the therapist thought she had a bad smell (but withheld telling her because he did not want to hurt her feelings).

At the beginning of treatment, the therapist offers a sincere commitment to provide whatever support is needed for the patient to survive. Later in treatment, the therapist may move toward a more demanding stance that conveys the expectation the patient will be able to live in the real world. As is true in the back-and-forth movement between dependence and independence in childhood, the patient may resent and resist such encouragement. The therapist must go slowly, showing the patient that the expectation of independence is an expression of confidence.

Aside from the fear of psychosis that may inhibit a therapist's close work with psychotic people, and aside from feelings in the countertransference that can provide clues to the patient's preverbal states, several other forms of countertransference reactions may arise during the course of treatment. Therapists may feel so deeply gratified by the intimately meaningful reward of the work that they do not recognize that a patient has made significant progress toward termination. Or they may become so habituated to a patient's peculiarities that they fail to recognize the pervasiveness of a patient's psychosis. Or they may experience a diminished interest in a patient as the person becomes less psychotic. There is little idle chit-chat in work with people who suffer from psychosis. Therapists may find the courage and raw vitality in psychotic patients an

appealing environment in which to work, especially when compared with the deeply defended "normalcy" one often encounters working with personality disorders.

As in any psychotherapy, the patient changes not only from insight but from internalizing aspects of the therapist. The patient may identify with the therapist's accepting attitude such that the patient's excessively harsh superego is softened. The patient may take the relationship with the therapist as a model of how a human relationship might be. When therapy succeeds in reducing a patient's distortions of reality, expressed in psychotic symptoms, and when the relationship with the therapist provides a sufficient basis for self-esteem, the patient can get on with the sometimes difficult business of adult living. After all, the therapist never promised the patient a rose garden—just a better chance at living.

Cognitive Behavioral Therapy for Psychosis

Cognitive behavior therapy, as first described by Aaron Beck in 1979, was not applied to the treatment of psychosis in earnest until the early 1990s. The first comprehensive textbook of CBTp appeared in 1995, and many have appeared since. Interventions shifted from the primarily behavioral focus in token economies to cognition in psychosis, and a vibrant several decades of clinical research and refinement of technique ensued and continues. The following account of CBTp is written by a psychoanalyst who practices CBTp. The same material described by a therapist dedicated to cognitive behavioral therapy would read somewhat differently.

A basic tenant of cognitive theory is that a person's beliefs about events mediate between those events and the emotional and behavioral consequences that follow them. This paradigm is summarized in the A-B-C sequence familiar to all cognitive therapists, in which "A" is an activating event, "B" is a belief about the meaning of that event to the individual, and "C" is the emotional and behavioral consequence of that belief. Assuming that emotional distress and maladaptive behavior are a consequence of delusional beliefs, CBTp researchers set out to determine how delusional beliefs and beliefs about "voices" arise, hoping to find ways to help patients change these beliefs and reduce suffering.

Their research revealed that a number of cognitive biases found in ordinary people were exaggerated in people with psychosis, including attributional bias, self-referential and attentional bias, a tendency to jump to conclusions, dichotomous thinking, belief inflexibility, and theory of mind deficits. A cognitive model of psychosis emerged, in which prior expectations were seen to shape the cognitive processing of an external life event or an internal event like hearing a "voice." Informational events that have the qualities of being expected, external, clear, common, neutral in affective significance, voluntary, public, and irrelevant to the person, may be ignored. But when, as is often the case with psychotic experiences, an

event is unexpected, internal, ambiguous, unusual, affectively loaded, involuntary, private, and personally relevant, the event proceeds to further cognitive processing according to the individual's information-processing style under conditions of high emotional arousal. Delusional people tend to have a cognitive style that focuses on current stimuli as opposed to past regularities of experience, and they tend to make rapid, overconfident judgments that are lacking in nuanced cognitive complexity.

Descriptions of the cognitive biases operating in psychosis allowed clinicians to tailor CBTp interventions to research findings. Through a process of guided discovery, patients can be shown that their belief is not the only possible explanation for events. Patients were taught to identify the cognitive biases operating in their own thoughts and to question the conclusions that followed from these biases. Many of the cognitive biases revealed by empirical research are known in the psychoanalytic clinical literature by different names. For example, Klein's idea of a primal split, in which the self is identified with good elements of the psyche and bad elements are projected into a persecutory object, is a description of what CBTp therapists would call an "attributional bias." What Freud described in 1926 in *Inhibitions, Symptoms, and Anxiety* as a constriction of the ego in the service of defense, CBTp would call a "safety behavior." There are many such crosswalks between psychoanalytic and cognitive theory. It may be reassuring to experienced clinicians to find that the psychoanalytic and cognitive behavioral traditions have found their way to many of the same ideas, and that CBTp has demonstrated the value of those ideas in empirical research.

CBTp continues to advance at a rapid pace. In a so-called "third wave" of clinical interventions, new types of therapy continue to emerge. Over time, early attempts to mitigate suffering by changing beliefs have made room for more meta-cognitive approaches; that is, helping patients change the way they think about their own thinking. These innovative techniques include, among others, mindfulness-based approaches that draw upon the wisdom of Eastern traditions to foster a more accepting attitude toward suffering, with an emphasis on kindness and compassion. In contrast to traditional CBTp, which might encourage patients who are hearing distressing "voices" to turn their attention away from the "voice" and to change their beliefs about the "voice," a mindfulness approach would encourage them to adopt a more tolerant, more meditative, less emotionally reactive response to the presence of the "voice."

Acceptance and commitment therapy does not have the primary aim of changing beliefs, but rather hopes to encourage patients to identify and keep clearly in mind their personal values, and to act on these values rather than acting on the prompt of a psychotic symptom. Compassionate mind training aims to help patients increase self-esteem by resisting self-critical elements in their personality more effectively. For example, the "two chair

technique' allows the patient to imagine sitting in one chair while conducting a dialogue with an inner "bully." This technique echoes the voice dialogue approach developed in the late 1980s, and Gestalt approaches promulgated earlier, and is in keeping with a fundamental goal of most psychodynamic psychotherapy: diminishing the impact of an overly punitive superego.

CBTp offers a variety of effective interventions that can be offered in their own right. It can be effectively integrated into the second phase of PDTp to help patients recognize the operation of distorting psychotic mechanisms. This sets the stage for more effective PDTp work, if the patient needs, wants, and can use this treatment approach. Like psychodynamic approaches, CBTp therapies are too richly diverse to summarize. Below find a description of how a more traditional CBTp might be conducted in the following sequence: (1) engaging the patient, (2) eliciting the patient's story, (3) introducing CBTp with an emphasis on coping mechanisms, (4) reviewing the three treatment models underlying CBTp, (5) applying these models to the patient's experience, (6) challenging beliefs and reformulating the patient's personal life narrative, and (7) maintaining contact and employing "booster sessions" as needed.

Engaging the Patient in CBTp

The therapist should pay attention to how the patient came to be referred. Was it the patient's idea, or someone else's? What is the patient's understanding of the purpose of the meeting? Does the patient have any specific anxieties about the meeting? Patients may wonder why the therapist is offering to see them more frequently than they expected. One patient joked, "Do you have a new supervisor?" The therapist might say, "If it is okay with you, we can meet a few times and talk about what has been happening in your life and see if we can figure out a way for things to be less distressing." Does the patient have an immediate, over-riding concern that must be addressed before the patient can focus on the therapy? A patient who is worried about being evicted from home, or who does not have enough to eat, will have difficulty paying attention until these needs are addressed.

The therapist will want to know about prior interactions with the mental health system, particularly those that have been notably problematic for the patient. Allowing the patient a virtual "panic button" that the patient can activate if the patient gets overly anxious in the session may be helpful. The therapist can ask occasionally, "How are we doing so far?" The therapist must be sensitive to the fact that the patients may believe they are putting their life at risk by even talking with a therapist. If the patient is reluctant to talk about what the therapist sees as psychotic symptoms, the therapist can find something else the patient is willing to discuss. In summary, as was emphasized by the American Sullivanians,

contemporary CBTp therapists establish a relaxed conversational manner that levels the power relationship between patient and therapist.

Eliciting the Patient's Story

It is vital to ask what the patient thinks is happening and why. What evidence does the patient have for a problematic idea? The patient will often point either to one seminal event that fostered the delusional belief or to a series of events given meaning by ideas of reference, in which case the patient may have pondered, "If this happened once or twice, it might be a coincidence, but it has happened over and over again, so there must be something going on." It is extremely important to elicit any aspects of the patient's experience about which the patient feels confused, as these doubts can often be used to forward later work.

For example, a man who had no doubt that he was an angel reported he had had an experience when he was being taken to the hospital that made no sense to him. The ambulance had passed children playing in school yard. He saw a good aura around the heads of some of the children, and a bad aura around others, but he knew there must have been something wrong with his perception because he believed all children were innocent and good. To encourage a therapeutic alliance based on evaluating data, the therapist might ask such a patient to rate his degree of preoccupation and conviction regarding a delusion on a 1 to 100 percent scale. If the patient reports being 99 percent certain of the delusion, the therapist might quip, "So we can work with the 1 percent."

Because the wish to reduce distress drives treatment, the therapist must determine what the patient finds most distressing about the current situation. While a diagnostically oriented clinician might aim at eliminating symptoms like "voices," the patient may be initially unconcerned about the "voices" and yet quite distressed that the patient's family regards them as imaginary. The therapist will want to establish rapport by empathizing with the patient's beliefs and feelings from the patient's point of view. The therapist might say, "Having heard what you went through, I can see why you came to the conclusions you did. When you first noticed strangers in your neighborhood looking at you, you concluded that you were under surveillance. I can see how that made sense to you at the time." The therapist may go on to ask what other people think about the patient's belief. What does the patient's family say? Does anyone share the patient's point of view?

The therapist can record the patient's life history on a *time line* sketched out along a horizontal line on paper and add to the time line as history accumulates. Putting the patient's history down on paper will frame the therapeutic alliance around a shared data set. In many cases, the time line will be the first occasion the patient has ever constructed any organized, linear sequence of personal history. Life events that psychotic individuals

have trouble connecting in mind can sometimes be connected on paper by noting their physical proximity on the page. For example, the therapist might say, "We marked your parent's getting divorced right here when you were away at school, and I see the 'voices' started up a short time after that."

Temporal sequences, and therefore causal links, can often be better established on paper than in the flow of conversation. The use of paper also gets the patient used to working with diaries and documents in homework assignments. At a later time in treatment, it will be useful for the therapist to draw two columns representing alternative beliefs so that evidence for and against each belief can be arrayed in each column. It is important to inquire about a history of physical or sexual abuse or bullying, but it is wise to establish some rapport with the patient before venturing into this territory. Even when the patient has been abused, the patient may not tell the therapist early in treatment. Substance abuse is usually more easily acknowledged.

If the patient hears "voices," the therapist should inquire about the identity, purpose, power relationship to the patient, consequences of noncompliance with the "voices," and whether the patient regards them as omnipotent or omniscient. As noted earlier, the therapist should inquire what diagnosis the patient has been given, by whom, and what it means to the patient. The therapist will want to know what the patient's daily life is like, and what relationships are important to the patient. The therapist should consult the chart and collaterals for additional information, and, of course, screen for danger to self or others.

Introducing the Principles Underlying CBTp with an Initial Emphasis on Coping Mechanisms

After the therapist has constructed a comprehensive time line and identified the primary sources of the patient's distress, goals for treatment with the patient can be addressed. The fundamental aim is to reduce distress rather than necessarily to eliminate symptoms. Goals might include an increased understanding of what triggers psychotic symptoms, improved coping with residual symptoms, reduction in stress related to delusions and hallucinations, and maintaining gains and preventing relapse. The therapist should inquire as to what coping mechanisms the patient currently uses.

If a patient has abandoned hope of coping, this may be a grave sign. The therapist may ask patients if they would be interested in learning how other patients cope with distressing experiences. Most will say yes. This conveys the implicit message that what is happening to the patient has happened to other people: the patient is not alone. Most coping techniques involve the redirection of attention away from distressing symptoms or toward more rewarding interests. Patients may distract themselves by listening to music

or practice focusing away from symptoms and on to pleasant imagery. The therapist can go through a list of possible coping techniques with the patient and choose several to try as a homework assignment. Coping cards that the patient can carry are useful. By highlighting the patient's coping capacity, the therapist underscores that symptoms are not entirely outside the person's control.

The therapist can explain the three basic models underlying the CBTp approach: (1) the stress/vulnerability model that emphasizes the importance of stressful life events in the genesis of distressing experiences, (2) the continuum model that finds analogies to psychotic symptoms in ordinary mental life, and (3) the A-B-C model noted earlier. The therapist should be fluent in each of these models and be able to explain them clearly to the patient. The stress/vulnerability model is a good place to start. The therapist should avoid appearing to be an expert lecturing the disempowered student. One might say, "Everyone experiences stress at some point in life. Stress can result in worries and lack of sleep. I have certainly felt the effects of stress myself."

Discussing the Normalizing and A-B-C Models in More Detail

Normalizing psychotic symptoms equalizes the power relationship between the patient and therapist, which encourages a good-faith exploration of the patient's symptoms without serious risk that the patient will feel diagnostically stigmatized. The patient will hear the therapist saying, "We all have a bit of what you have, though you have more than most." Normalizing encourages patients to place their symptoms in a psychological realm rather than to regard them as problematic realities. The therapist should be prepared with a variety of personal examples of psychosis-related mental events that can be shared. One might normalize hearing "voices" by describing hallucinations of everyday life such as hearing one's name on the street when no one is speaking, or by describing how "voice" hearing is a variant of the self-talk and internal dialogues everyone engages in from time to time, as when people say to themselves subvocally, "You idiot!" or when, on a diet, they say to themselves "Who are you kidding! You don't need dessert!"

Although affect and cognition interact in a dialectic manner, it is useful to teach patients the A-B-C model that emphasizes beliefs. This can be taught first in abstract outline, then from the therapist's personal examples, then from mundane examples from the patient's experience, and eventually in relationship to the patient's core delusional beliefs. The A-B-C model can be used to analyze a relatively neutral event in the patient's life. This gives patient and therapist practice with the concept. The patient is invited to examine a distressing "A," ideally an event about which the patient already expresses some doubt.

As patient and therapist begin to use the A-B-C model to examine the

patient's core psychotic experiences, the therapist can return to normalizing to emphasize the range of beliefs and perceptions that occur in ordinary mental life. The author sometimes tells patients the story of how in dim light he "hallucinated" the image of a black dog curled up on a bed that turned out to be a black sweatshirt thrown on the bed in such a way that the arms of the sweat shirt looked like the folded legs of a curled up dog. This example illustrates how people sometimes "see" what they expect to see, or do not want to see.

Applying These Models to the Patient's Experience and Begin Investigating Core Psychotic Beliefs

The therapist can with advantage adopt a "Columbo-style," offhand way of talking. "Columbo" was a detective on television, played by Peter Falk, who spoke in a casual manner that belied what he knew. The therapist fosters a therapeutic alliance not based on the therapist's claim to know what is in the patient's mind, but on the joint investigation of data consciously available to both patient and therapist. The treatment team consists of Watson and Crick sitting side by side, rather than Freud sitting behind Dora (one of his first patients).

The A-B-C model can be used to examine core psychotic experiences. The therapist can put two columns down on a piece of paper, to examine alternative theories. First, the therapist and patient clarify the "A": What precisely did the patient experience? What precisely did the patient hear? The patient may have actually experienced only a muffled sound, but later concluded it was a threat. An "A" may be a mental event, like a "voice," or and external event, like the glance of a stranger. Patients often merge the "A" and the "B" such that the event and the patient's belief about an event are a simultaneous, equated experience. CBTp aims to separate the "A" from the "B" so that the patient's beliefs can be examined. The therapist might say, "If your explanation for 'A' is correct, your reaction is understandable. If there is any other way of understanding what is happening to you that would be less distressing, it would be important for us to discover it. Let's look at the evidence, and see what we can figure out."

A rich toolbox of CBTp techniques has been developed to examine beliefs; for example, the practice of rating the patient's degree of conviction on a scale from 1 to 100 percent. If patients can be made aware of their convictions changing, the therapist can point out that while facts do not change, beliefs do, thereby categorizing the delusion as a belief rather than a fact. Other techniques include rating the likelihood of a belief, rating the value of evidence, agreeing to disagree about final conclusions pending further assessment, peripheral questioning, informational handouts that increase real-world knowledge, googling topics during session, inference chaining ("If this turned out to be true, what would it mean to you? Which explanation would you prefer to be true?"), homework assignments and

diaries, reality-testing experiments, and self-esteem building exercises.

Challenging Delusional Beliefs and Reformulate the Patient's Personal Life Narrative

On the basis of their joint investigation, the therapist can offer a tentative formulation based on the stress/vulnerability, cognitive, or biopsychosocial model that leads to a more adaptive alternative belief. A formulation using the stress/vulnerability model might be: "Your mother's death was very stressful for you. You weren't sleeping well, and were using drugs to ease the pain. In that stressed-out state, you had some disturbing experiences, like thinking your neighbor was trying to poison you. We can now see that those beliefs were the result of the upset state of mind you were in at the time." Or, a formulation based on identifying cognitive biases: "You have a cognitive bias that leads you sometimes to jump to the conclusion that people know you and want to hurt you. You tend to personalize events, believing people are referring to you when they are not." Or, a fuller biopsychosocial formulation: "You felt that if you had set a better example for your brother, he might not have died so young. You felt guilty, and when you couldn't get your life together after the Navy, you felt ashamed. You criticized yourself. Over the years those feelings of guilt and shame built up, and you came to feel that shame everywhere you went, including in the eyes of the dogs in your neighborhood." The therapist should determine whether the patient engages in safety behaviors that constrict life experience, precluding positive experiences that would contradict delusional beliefs.

As described in the earlier part of the section on technique, the CBTp work outlined above can lead into a final phase of therapy that seeks to find psychological explanations for the formation of the patient's psychotic symptoms. CBTp is the preferred method for showing patients the *literal falsity* of their delusional beliefs. PDTp is the preferred method for exploring the *figurative truth* of their delusional beliefs.

Maintaining Contact and Employ "Booster Sessions" as Needed

Therapists should develop with each patient a description of possible stressors and early signs of relapse to look out for. They need to anticipate what the patient should do if these signs appear. Working collaboratively, they should encourage the patient to practice the ability to identify signs of relapse and to implement the recommended course of action. Continuing to see the patient on a regular basis is critical for supporting positive changes.

ETHICAL ISSUES

The ethical constraints, such as maintaining professional boundaries, that govern the practice of psychotherapy with nonpsychotic patients are well

known, and apply equally to the treatment of psychotic patients, and so will not be addressed here. Because most individuals with chronic psychosis are treated in the public psychiatry sector, or in hospitals or clinics funded by government third party insurance, their well-being is directly affected by how resources are allocated and what kinds of care are reimbursed, a circumstance that poses ethical concerns.

In the study of ethics, there is a time-honored tradition of using a hypothetical situation to think through an issue. Consider the following hypothetical: Assume that chronic psychosis is a complex biopsychosocial disorder that has no single cause. Assume that no mental health system has developed a way of treating psychosis that avoids the prescription of neuroleptic medications in all patient cases, despite evidence that these drugs may cause structural damage to the brain. Also assume that individual psychotherapy is an effective treatment of psychosis that in some cases may resolve psychotic symptoms unresponsive to medication, improve quality of life, and reduce or even on occasion eliminate the need for medication. If these assumptions are made in the author's thought experiment, and assuming that physicians are ethically bound to provide patients access to quality care, what ethical implications follow from the above assumptions?

The author suggests that psychiatry has an ethical obligation to consider the following:

1. At present there are insufficient numbers of therapists trained in psychotherapy for psychosis to meet patient need, and there are not enough supervisors who can train additional therapists. Federal and state agencies should fund pilot training programs with a research component to develop supervisors and therapists with the needed skills. Psychiatric residency programs should find ways to teach psychotherapy for psychosis.
2. In understaffed public clinics, where too many patients are assigned to too few clinicians, administrators should pilot psychotherapy initiatives coupled with research that allow clinicians to see at least a few patients frequently enough to provide PDTp and/or CBTp.
3. To foster a therapeutic alliance that creates a bridge between inpatient acute care and outpatient PDTp and CBTp, suitable patients nearing discharge on inpatient units should begin working with an outpatient therapist while the patient is still in the hospital.
4. Granting agencies should encourage research that explores whether PDTp and CBTp can help reduce the total neuroleptic exposure patients receive.

Despite the enormous amount of information accumulated about the functioning of the brain, most psychiatrists regard neuroscience as still in its infancy. Psychiatry should also regard the psychotherapy of psychosis as

still in its infancy, but as a childhood with much promise. Improvements in outcome seen in the last two decades with CBTp suggest there is much progress to be made in the psychotherapy of psychosis when a critical mass of researchers and clinicians put their minds to the challenge.

Since Bleuler's time, there has always been a small chorus of voices in the clinical community who were convinced from personal experience that psychotherapy can be an effective mode of treatment of psychosis. If it is true (as personal experience and research evidence suggest) that psychotherapy can be an effective element of treatment, psychiatry has an ethical obligation to understand why patients with psychotic disorders have little or no access to psychotherapy, and to devise the psychiatric educational, political, economic, and clinical vehicles that would allow psychotherapy to be delivered. As Karon put it a decade ago, "The real tragedy of schizophrenia is not the severity of the symptoms and the suffering that results for patients and for their families, but that we know psychotherapies that work and we are not using them."

RESEARCH AND EVALUATION

There is solid evidence for the efficacy of CBTp, supported by multiple RCTs and meta-analyses. An early meta-analysis of 34 CBTp clinical trials found an effect size of .4, in the range of effect sizes for some neuroleptics (typically in the .4 to .5 range, with the exception of clozapine at .8). A more recent meta-analysis of CBTp RCTs showed a somewhat smaller average pooled effect size of .33. CBTp is included in the British National Institute for Health and Care Excellence (NICE) guidelines. A recent review of six different psychological interventions for psychosis found that CBTp was significantly more efficacious than other interventions in reducing positive symptoms, while social skills training showed an advantage in improving negative symptoms. Thus, it is clear that CBTp works. There are no head-to-head RCTs comparing psychopharmacology alone against CBTp alone.

The many published case histories, autobiographical accounts, and the testimony of many therapists and patients bear witness to the efficacy of PDTp, and the positive experiences of the Soteria Project and the Open Dialogues program lend support to treating psychosis without neuroleptics, but if one considers RCTs to be the gold standard of evidence, the RCT base for PDTp is small, methodologically flawed, and outdated. The Michigan State Psychotherapy Project, in which hospitalized patients with psychosis were randomly assigned to either individual psychotherapy without medication, therapy with medication, or medication without therapy, reported that psychotherapy produced significantly greater improvement than medication alone and was particularly effective in changing thought disorder. The investigators also concluded that more experienced therapists produced more improvement. A more recent study compared expressive insight-oriented therapy (EIO) with reality adaptive supportive

therapy (RAS) and found that both were effective. RAS had a clearly greater impact on recidivism and social functioning, while EIO had a modestly more positive effect on ego functioning and cognition. The conclusion of a 2001 Cochran Database Review of randomized individual psychodynamic psychotherapy and psychoanalysis for schizophrenia, however, did not support the efficacy of those approaches. The review concludes by saying there is an urgent need for research in this area. It is apparent that while RCT support for CBTp is strong, PDTp has been insufficiently studied.

FUTURE DIRECTIONS

CBTp is structured and brief compared with PDTp, which is less structured and more open-ended. This makes PDTp more difficult to study than CBTp. Developing new ideas and conducting research requires a critical mass of clinicians trained in the techniques to be studied. At present there exists more than a critical mass of CBTp trained clinicians working and doing research in many centers. The same cannot be said for PDTp, where more teaching and practice opportunities need to be developed.

At this point, what is known about the comparative efficacy of psychopharmacology and psychotherapy? It is known that the majority of patients treated in a care system with a predominantly pharmacological paradigm remain chronically disabled. It is known that all neuroleptics have potentially significant side effects and that psychotherapy for psychosis has never been proven to have any. It is known that roughly a quarter of patients initially treated with medication after years of pharmacological treatment can do well off medication, but how to identify these people is unknown, or how to wean them off medication even if they could be identified.

It is not known how to teach psychotherapy and integrate it into current care systems. Whether psychotherapy combined with medication can reduce dosage creep or reduce total neuroleptic exposure for patients on maintenance medication is unknown. It is not known if the current psychotherapeutic techniques are the best that can ever be devised. Whether psychotherapy is more effective for some symptoms and medication is more effective for others remains unclear. Many important questions remain to be answered, but until there is a more vibrant clinical community doing psychotherapy with psychotic patients that can collaborate doing research with psychopharmacologists, these questions will likely go unanswered.

THE WOMAN WHO MURDERED HER PARENTS. The author will now describe the psychotherapy of a psychotic woman who, prior to her beginning an integrated CBTp/PDPp treatment, had been an inpatient in a state hospital for 15 years, after she stabbed both her parents to death while in an acute psychotic state. At the time she began psychotherapy, she had been taking 600 mg of clozapine for the prior 8 years, with no change in her core delusions, and accordingly no hope of ever leaving the hospital. After a successful course of

psychotherapy, she was discharged to a community residence.

Loretta was an only child who returned home after college and worked a number of low-demand jobs. She joined an Evangelical church and became preoccupied with church activities. Soon afterward, she suffered her first psychotic episode and psychiatric admission. After discharge, she was followed in a community clinic on psychotropic medication. Five years later, she again became preoccupied with the church. One evening after an altercation with her father, she murdered her parents. She recalled that prior to the homicides, "one of the Evangelicals had looked at me with a piercing stare" that she believed had placed her in a hypnotic trance that led her to murder her parents. Staff believed that she might seek revenge if she were discharged. In 15 years in the hospital, she had never been involved in a violent incident, though she feared assault by other patients and believed that nursing staff were intentionally provoking her to violence.

The therapist followed a CBTp approach for the first 20 sessions. At one point during this period, the therapist focused in detail on the subjective events that had occurred during the psychotic episode when Loretta killed her parents, framing these as "A"s to be investigated in A-B-C format. Loretta acknowledged that people might stare for a variety of reasons. The therapist drew two columns of paper: one for the "Evangelical" explanation and one for the "mental health" explanation. He then asked which Loretta hoped was true. He had been imagining that the patient might welcome the mental illness explanation, but much to his surprise, she said the opposite. "If the Evangelicals hypnotized me, I was just stupid. But if I killed my parents because of mental illness, I am a piece of garbage and I should rot in hell!"

Moving into the PDTp phase of treatment with more focus on life history and affect, the therapist conceived a role-playing exercise to explore the patient's dissociated guilt. The therapist asked if she were familiar with the case of the woman who claimed to be David Letterman's wife, who kept breaking into his house, and she said she was. The therapist then pretended to be an attorney who presented three different scenarios to Judge Loretta, asking whether she would recommend treatment or punishment. In the first scenario the woman resists police, but no one is hurt. She recommended treatment. In the second scenario, a policeman receives a minor physical injury; she again recommended treatment. In the third scenario, the woman strikes a policeman in the head with a lamp and kills him. Judge Loretta paused. She said she needed more information. Was the attack premeditated? If there was premeditation, Judge Loretta would recommend punishment. If not, she would rule for treatment. The therapist then asked Loretta if she had premeditated the murder of her parents. As if taken aback, she replied "No" with confidence and calm. In 15 years of inpatient treatment-as-usual, no one had ever given her an opportunity to talk about her terrible guilt over murdering her parents.

The next week as the therapist turned the key and stepped through the door into Loretta's ward, she rushed to greet him, a broad smile on her face, a brand new shawl around her shoulders. She announced she had recovered the ability to pray, and that she felt once again able to communicate with God! The therapist feared that the previous session had undone her, and that she was now hallucinating. But this was not the case. She explained that after she murdered her parents, she had a feeling that God had turned his back and would no longer listen to her prayers. In the previous week, she had not heard the "voice" of God, but rather she had recovered a feeling that God was once again listening to her! This was a deeply moving moment. Her psychotherapy had begun to mitigate the agonizing guilt that she had carried for 15 years in the Evangelical delusion. Her obsession with the Evangelicals faded, staff grew increasingly confident in her recovery, and she was discharged to a community residence within a year. The therapist met Loretta at Dunkin'Donuts a year after her discharge. She was doing well. Both ordered coffee and a jelly-filled donut, and both agreed the donuts were delicious.

Loretta's story and the stories of many patients like her who have benefited from PDTp and CBPp constitute evidence, even if those stories have not been studied via RCTs, of the importance and efficacy of psychotherapy in the treatment of individuals with psychotic suffering.

References

Andreasen NC, Liu D, Ziebell S, Vora A, Ho BC. Relapse duration, treatment intensity, and brain tissue loss in schizophrenia: a prospective longitudinal MRI study. *Am J Psychiatry*. 2013;170(6):609–615.

*Arieti S. *Interpretation of Schizophrenia,* 2nd ed. New York: Basic Books, Inc; 1974.

*Beck AT, Rector N, Stolar N, Grant PM. *Schizophrenia: Cognitive Theory, Research and Therapy*. New York: Guilford Press; 2009.

*Bunker N, Summers A, (eds). *The International Society for Psychological and Social Approaches to Psychosis Book Series:* London: Routledge.

Chadwick PD, Birchwood M, Trower P. *Cognitive Therapy for Delusions, Voices and Paranoia*. Chichester: Wiley; 1996.

Federn P. *Ego Psychology and the Psychoses*. Weiss E, ed. Oxford: Basic Books, Inc.; 1952.

Fonagy P, Target M. Playing with reality: I. Theory of mind and the normal development of psychic reality. *Int J Psychoanal*. 1996;77 (Pt 2):217–233.

Fromm-Reichmann F. Transference Problems in Schizophrenics. *Psychoanal Quart*. 1939;8:412–426.

Garety PA, Kuipers E, Fowler D, Freeman D, Bebbington PE. A cognitive model of the positive symptoms of psychosis. *Psychol Med*. 2001;31(2):189–195.

Garrett M, Turkington D. CBT for psychosis in a psychoanalytic frame. *Psychosis*. 2011;3(1):2–13.

*Greenberg J. *I Never Promised You a Rose Garden*. New York: St Martins Press; 1964.

Gunderson JG, Frank AF, Katz HM, Vannicelli ML, Frosch JP, Knapp PH. Effects of psychotherapy in schizophrenia: II. Comparative outcome of two forms of treatment. *Schizophr Bull*. 1984;10(4):564–598.

Ho BC, Andreasen NC, Ziebell S, Pierson R, Magnotta V. Long-term antipsychotic treatment and brain volumes: a longitudinal study of first-episode schizophrenia. *Arch Gen Psychiatry*. 2011;68(2):128–137.

Isaacs S. The Nature and Function of Phantasy. *Int J Psychoanal*. 1948;29:73–97.

Karon BP. The tragedy of schizophrenia without psychotherapy. *J Am Acad Psychoanal Dyn Psychiatry*. 2003;31(1):89–118.

*Karon BP, VandenBos GR. *Psychotherapy of Schizophrenia. The Treatment of Choice*. Lanham, MA: Rowman & Littlefield; 2004.

Lidz T. *The Origin and Treatment of Schizophrenic Disorders:* New York: Basic Books; 1973.

Lotterman A. *Psychotherapy for People Diagnosed with Schizophrenia: Specific Techniques*. East Essex and New York: Routledge; 2015.

Malmberg L, Fenton M. Individual psychodynamic psychotherapy and psychoanalysis for schizophrenia and severe mental illness. *Cochrane Database Syst Rev*. 2011;(3):CD001360.

Marcus ER. *Psychosis and Near Psychosis*. Revised 2nd ed. Madison: International Universities Press; 2003.

Morrison AP, Renton JC, French P, Bentall RP. *Think You're Crazy? Think Again: A Resource Book for Cognitive Therapy for Psychosis*. London: Routledge; 2008.

Read J, Fosse R, Moskowitz A, Perry BD. The traumagenic neurodevelopmental model of psychosis revisited. *Neuropsychiatry*. 2014;4(1):656–679.

Segal H. Notes on symbol formation. *Int J Psychoanal*. 1957;38(6):391–397.

Stone MH. The history of the psychoanalytic treatment of schizophrenia. *J Am Acad Psychoanal*. 1999;27(4):583–601.

Sullivan HS. *Schizophrenia as a Human Process*. New York: Norton; 1962.

Summers A, Rosenbaum B. Psychodynamic psychotherapy for psychosis. Empirical evidence. In: Read J, Dillon J, eds. *Models of Madness:Psychological, Social and Biological Approaches to Psychosis*. London: Routledge; 2013.

Tai S, Turkington D. The evolution of cognitive behavior therapy for schizophrenia: Current Practice and Recent Developments. *Schizopr Bull*. 2009;35(5):865–873.

Turner DT, van der Gaag M, Karyotaki E, Cuijpers P. Psychological interventions for psychosis: a meta-analysis of comparative outcome studies. *Am J Psychiatry*. 2014;171(5):523–538.

Winnicott DW. Transitional objects and transitional phenomena; a study of the first not-me possession. *Int J Psychoanal*. 1953;34(2):89–97.

Wykes T, Steel C, Everitt B, Tarrier N. Cognitive behavior therapy for schizophrenia: effect sizes, clinical models, and methodological rigor. *Schizophr Bull*. 2008;34(3):523–537.

▲ 33.17 Hallucinogen-Assisted Psychotherapy

STEPHEN ROSS, M.D., AND MICHAEL P. BOGENSCHUTZ, M.D.

INTRODUCTION AND DEFINITION

In this section, the historical and recent literature on the clinical use of the classic or serotoninergic hallucinogens (SHs) (i.e., those hallucinogens that are thought to exert their psychedelic or consciousness-altering effects primarily through agonist or partial agonist activity at serotonin 2A [5HT2A] receptors) as adjuncts to psychotherapy will be reviewed. Although the SHs have been used by humans for millennia, known to science for over a century, and subjected to extensive basic research for decades, the rigorous study of their clinical use as adjuncts to psychotherapy is in its infancy. Promising early-phase research conducted from the 1950s through the early 1970s was discontinued before any conclusions could be reached concerning the efficacy of any of the classic hallucinogens for any clinical condition. However, the research that was conducted in that era strongly suggests that classic hallucinogens have clinically relevant effects, particularly in the case of lysergic acid diethylamide (LSD) treatment of alcoholism and cancer-related psychological and existential distress.

In the past decade, clinical trials have resumed investigating the effects of classic hallucinogens in the treatment of various mental illnesses (i.e., cancer-related psychological and existential distress, treatment-resistant depression, OCD) and addictive disorders (i.e., alcohol, nicotine). The studies that have been completed to date are small open label or randomized controlled trials (RCTs), and although the results have been very promising, they are not sufficient to definitively establish treatment efficacy. The encouraging results that have been seen historically with classic hallucinogen treatment across alcohol, nicotine, and opioid dependence suggest the possibility that other chemical and behavioral addictions may also respond to treatment with these agents. While it cannot be assumed that addictions are interchangeable or that classic hallucinogen all have comparable efficacy in the treatment of addiction, it seems reasonable to pursue additional indications. One such possibility is currently being explored in an ongoing RCT that is the first to evaluate the effects of psilocybin-assisted treatment of cocaine addiction (NCT 02037126). Other promising psychiatric indications that are worth exploring using hallucinogen-assisted psychotherapy could include: PTSD, Generalized Anxiety Disorder, eating disorders, anti-social personality disorder/criminal recidivism, and narcissistic personality disorder. If the data ultimately supports treatment efficacy, hallucinogen-assisted

psychotherapy would truly constitute a novel psychopharmacological–psychosocial treatment paradigm to treat these disorders, although the risk of adverse psychological events would have to be controlled through a careful screening process.

The clinical safety of SHs, well documented by extensive research in the 1950s through early 1970s, has been confirmed in the rigorous clinical work which began again in the 1990s. Although physical and psychological risks exist, they can be minimized through the safety procedures followed in recent clinical studies, including careful screening of participants, thorough training of therapists, close attention to dose, set, and setting, preparation, support, monitoring, and follow-up of participants, and predetermined procedures for dealing with medical and psychiatric emergencies. Since the resumption of clinical research with classic hallucinogens in the 1990s to date, there have been no reported hallucinogen-related serious adverse events or cases of persistent harm to participants. Although it would be unrealistic to believe that any medication is harmless, this fact is encouraging. Future research must continue to make every effort to maximize safety. Should a classic hallucinogen be approved for clinical use in the future as an adjunct to psychotherapy, it will also be critical that rigorous safety procedures be built into the approved conditions of clinical use.

Although research has elucidated many of the acute neurobiological and psychological effects of classic hallucinogens on humans, animals, and in vitro systems, the mechanisms of clinically relevant persisting effects remain poorly understood. Although neuroimaging studies have begun to characterize the acute effects of psilocybin on brain function, how these changes neuro-biologically mediate therapeutic benefit is unclear at present. Changes in 5HT2A receptors, serotonin transporters, neurotrophic factors, and other relevant brain chemicals have not been measured in the context of clinical studies of classic hallucinogen treatments. At this point, the causal mechanisms with the strongest support relate to neuroplasticity, in terms of intensely memorable and highly meaningful experiences, particularly those of a mystical character. Much more work is necessary better to understand the mechanisms by which such experiences, and their neurobiological correlates, can lead to persisting psychological, behavioral and neurobiological change.

Although the small trials that have been conducted have provided some information concerning the psychological processes that may underlie behavior change, future trials should continue to characterize both the acute medication effects and the longer-term psychological changes that are associated with desired clinical outcomes.

Although interest in clinical use of SHs has increased dramatically in the past decade, significant barriers remain to a full exploration of their clinical potential. Due to the high costs, almost all new drugs today are

developed through investment by private, for-profit corporations. Such investment is unlikely in the case of classic hallucinogens such as psilocybin and LSD, because the compounds have been known for over 60 years and are not patentable. For this reason, funding through federal agencies such as the National Institutes of Health is particularly important. No trial has yet received federal funding, although preclinical work and neuroimaging studies have been supported. To date, funding of clinical research with classic hallucinogens has come entirely from private, nonprofit organizations such as the Heffter Research Institute. Currently the Usona Research Institute and the Heffter Research Institute are planning development of psilocybin for the indication of treatment of anxiety and depression associated with a cancer diagnosis and are in discussions with FDA about a potential phase 3 trial.

HISTORY

Early Academic Research with Hallucinogens

Although classic hallucinogens have been used by humans for at least 5 millennia, western scientific study of hallucinogens dates to the late 1800s, when mescaline was isolated and its effects described by Arthur Heffter. In 1943, Albert Hofmann discovered the psychoactive effects of LSD when he accidentally ingested traces of an ergot derivative he had synthesized, and confirmed the effects through self-experimentation. He quickly recognized the therapeutic potential of LSD, within psychiatry as did other investigators. Following Gordon Wasson's report of hallucinogenic mushroom use by the Mazatec tribe in Mexico, Hofmann isolated psilocybin from samples of the mushrooms and synthesized it in the laboratory in 1958.

Following the discovery or LSD, research on LSD and other classic hallucinogens expanded rapidly. Initially, the main focus of research was on the use of LSD as a model of psychosis. However, clinicians and clinical scientists in Europe, Canada, and the United States soon began to explore the use of these unregulated chemicals, in combination with various forms of psychotherapy, in the treatment of alcohol and drug addiction, existential crisis related to death in the terminally ill, personality disorders (in particular antisocial personality disorder), and various neurotic conditions being addressed in the context of psychodynamic psychotherapy. From the early 1950s to the mid-1960s, over 1,000 scientific articles were published reporting on the treatment of over 40,000 patients with classic hallucinogens.

In reaction to the cultural upheaval of the 1960s and concern about the dangers of wide-spread illicit use of psychedelics during that era, clinical research on hallucinogens came to halt in the early 1970s. The Controlled Substances Act of 1970 placed LSD, psilocybin, and related compounds in

Schedule I, and human research was not allowed to continue. Basic research continued, and many additional compounds were and continue to be discovered, most of which have never been subjected to the most basic pharmacological study, let alone human laboratory studies or clinical trials. Illicit use for recreational, therapeutic, and/or spiritual purposes continued.

Reemergence of Research Using Hallucinogens

After a hiatus of about two decades, human research on psychedelics resumed in the early1990s with Rick Strassman's studies of the subjective and physiological effects of intravenous dimethyl tryptamine (DMT) on normal volunteers. Since the beginning of the 21st century, there has been a sharp increase in research in this area especially pertaining to the use of psilocybin-assisted psychotherapy for psychiatric conditions such as psychological distress associated with cancer and addiction (tobacco and alcohol). The safety of classic hallucinogens (particularly psilocybin) in clinical research settings has been well documented. A considerable body of work has characterized the acute effects of classic hallucinogens on physiology, cognition, emotion, and brain function. All of these factors have motivated the resumption of research into the possible clinical applications of these drugs.

THEORETICAL ISSUES

Soon after Hoffman's accidental discovery of LSD in 1943, mental health practitioners and researchers began to explore the therapeutic potential of LSD, and later other hallucinogens, as adjuncts to psychotherapy for a range of psychiatric disorders. Although classic hallucinogens were sometimes used in a strictly biological or pharmacological model of treatment, hallucinogen-assisted psychotherapy became a newly developed model within psychiatry in Europe, Canada, and the United States. There was excitement that it represented a novel pharmacological–psychotherapeutic technique that could expedite the psychotherapeutic process by shortening the length of traditional treatment and that it could be used to help patients that did not respond well to traditional psychotherapies at the time (i.e., addiction, existential anxiety associated with terminal cancer). The two models used in the 1950s through early 1970s were *psycholytic* and *psychedelic* therapy. Although both combined administration of hallucinogens with psychotherapy to achieve therapeutic change, they emphasized different processes in bringing about these changes.

Psycholytic Therapy

In the *psycholytic* method, low-to-moderate doses of hallucinogens were

administered on multiple occasions to facilitate therapy that was based on traditional psychoanalytic principles, that is, helping the patient to become aware of unconscious desires, emotions, attachments, self-representations, and intra-psychic conflicts.

Psychedelic Therapy

The *psychedelic* method was a new and radical treatment model for psychiatry that used higher doses of LSD administered on no more than a few occasions, with the goal of occasioning a "peak-psychedelic" or mystical/spiritual experiences that were felt to be at the core of clinical improvement in conditions such as death anxiety in the context of cancer, and alcoholism. These experiences are classically characterized by the phenomenology of unity (sense of oneness), transcendence of the ordinary experience of space and time, sense of sacredness, sense of deep truth or ultimate meaning (noetic quality), deeply felt positive mood, and ineffability. It was held that such experiences often facilitated lasting change in habitual patterns of thought, behavior, experience of emotion, and even personality. Although these two models are conceptually distinct, some clinicians and investigators used both, or created hybrid models.

TECHNIQUES AND GOALS OF THERAPY

The first attempts at using hallucinogens (LSD) to treat psychiatric distress were based on a purely pharmacological model of intervention that did not account for set, setting, or psychotherapeutic preparation or integration. It was applied to patients with major depression, traumatic spectrum disorders, pain and existential distress associated with terminal cancer and patients with refractory personality or neurotic spectrum disorders. This technique was ultimately abandoned by most researchers as it became apparent to clinicians and researchers that the therapeutic outcome of the use of the hallucinogens was critically dependent on a variety of extrapharmacological factors. After trial and error, a new technique of hallucinogen-assisted psychotherapy became a newly developed model within psychiatry in Europe, Canada, and the United States. This new treatment model established the importance of ***set*** (psychological frame of mind, intention), excluding participants with major mental illness (especially psychotic illness or family history of psychotic illness), ***setting*** (environment/room on dosing days), and ***drug dose*** (and associated intensity of experience). It combined the administration of hallucinogens with psychotherapy before, during, and after the drug experience, believing that the subjective experience during the drug's acute effects, and the successful integration of these experiences, was crucial to the therapeutic benefit of the treatment.

Psycholytic Therapy

Psycholytic, and especially psychedelic therapy, had elements that made them unique forms of psychotherapy. Psycholytic therapy represented an extension and modification of psychoanalytically oriented psychotherapy, involved administration of low-to-medium doses of a hallucinogen (usually LSD), included a variable number of sessions depending on the therapeutic indication (range between 15 and 100 sessions), was conducted over months to sometimes more than 1 year, and had an emphasis on the phenomena emanating from the drug sessions. The drug dosing sessions would occur in a darkened, quiet room prepared to look like someone was in their home and would usually involve more than one therapist with sessions lasting a variable amount of time depending on the hallucinogen but typically for 8 hours or more since LSD was the most predominant drug used. General principles of psychoanalytically oriented therapy would be applied during sessions and during interim nondrug psychotherapy sessions with an intensification of the usual elements of psychoanalytic technique such as exploration of unconscious material (from the drug dosing session), interpretation and analysis of defenses and transference. The technique tended to be different from the more orthodox applications of Freudian psychoanalysis and would be more along the lines of supportive psychodynamic treatment with more personal involvement, physical contact, and genuine human support on the part of the treatment providers.

Psychedelic Therapy

Psychedelic therapy differed from psycholytic therapy in several important aspects and represented a completely new pharmacological–psychotherapeutic model when it was first introduced into psychiatry in the 1950s. It was developed based on the observation of sudden and dramatic clinical improvements associated with high-dose hallucinogen administration and quantum shifts in perception and cognition in patients who typically experienced “transcendental” “peak” experiences marked by highly salient spiritual, religious, or mystical phenomenon. One of the main goals of psychedelic therapy is to create an optimal treatment environment during the preparatory psychotherapy and during dosing sessions to foster such “peak” experiences and for the patient to experience a sense of unity of consciousness, egolysis, and an experience of transcendence associated with a sense of death and rebirth. During preparatory psychotherapy, the therapist reviews the patient’s life narrative; explores the narrative of the psychopathological disease state being treated (e.g., addiction, terminal cancer distress); focuses on personality factors and resistances to achieving a peak state; focuses on philosophical/spiritual/religious beliefs as well as existential themes and hierarchical life values; reviews the range of phenomenological effects of the drug experience ranging from esthetic changes to psychodynamic to archetypal to mystical; reviews techniques

used to deal with difficult psychological experiences, and emphasizes the importance of "letting go" and yielding to the experience knowing that skilled therapists will be with the patient for the duration of the experience and that the patient will return to their normative consciousness after the effects of the drug wear off. Psychedelic therapy emphasizes creating an optimal esthetic environment for the dosing sessions which are usually conducted in comfortable living room–like settings decorated with flowers, fruit, esthetically pleasing and evocative art, and items (i.e., photographs, art, spiritual/religious icons) that have a highly personally meaningful and symbolic value for the patient. Unlike conventional individual psychotherapy conducted by one therapist, there is commonly a male–female psychotherapy dyad that treats the patient during the preparatory, dosing, and integrative phases of the treatment. During the initial part of the dosing sessions (which last for hours depending on the hallucinogen used—i.e., 6 to 8 hours for psilocybin and 8 to 12 hours for LSD), the dyad team typically will ask the patient to state their specific intention for the dosing experience and day. After receiving the drug, the patient is instructed to lie down on a coach (with a comfortable blanket and pillow), is given eyeshades and instructed to focus internally on their experience. Preselected music accompanies the experience and is selected in relation to varying stages of the experience (i.e., before, during, and post the peak of the experience). The default position is for the patient to lie supine with eyeshades on and focusing internally, rather than with eyes open attending to external stimuli or interacting or talking with the therapists. Toward the end of the dosing session, the treatment providers will typically ask the patient to recount their experiences which can be difficult to describe, ineffable, transient, and quick to fade from memory. They will also typically instruct the patient to reflect on the experience of the day and write down as much of the experience as they can as a way to further facilitate encoding and remembering of the experience so as to optimally guide postsession integrative psychotherapy. At the end of the dosing session, the treatment providers assure that the patient has returned to their normal state of consciousness and is safe from a psychological and medical perspective before being discharged home, usually accompanied by a family member or a friend. Compared to conventional psychotherapy, psychedelic therapy is more concerned about transcending psychopathology than necessarily an interest in its analysis, although postsession integrative psychotherapy occurs to incorporate the experiences during the dosing sessions into further psychotherapy with the goal of consolidating psychodynamic changes that have occurred, and fostering a change in psychopathological symptoms.

CLINICAL ISSUES

Indications

Hallucinogen-assisted psychotherapy has historically been applied to a variety of psychiatric conditions including: addiction, existential distress associated with cancer, mood and anxiety disorders, and personality disorders. The first three conditions mentioned have the most historical and current empirical data assessing their use in these disorders (see Research and Evaluation section).

Limitations

There are several key limitations to providing hallucinogen-assisted psychotherapy. For one, all of the SHs are currently schedule I drugs that are not available for clinical use and are only able to be used in clinical research treatment studies. Given the onerous nature of obtaining approval for schedule I research with hallucinogens and the lack of federal funding in this area, currently these types of treatments are very limited to a few major academic medical centers in the United States and Europe and confined to a small group of psychopathological conditions. If the SHs were rescheduled and available as adjuncts to psychotherapy, there would be significant limitations to their broad use among mental health practitioners. For one, the time commitment is vastly different than the typical 45- to 50-minute individual psychotherapy session and the therapist would need to spend the entire day with the patient (considering drugs like psilocybin and LSD). Also, it is unlikely the drug would be available by prescription to acquire at a pharmacy to be taken at home by the patient. It is more likely that the treatment would be restricted to specially designated and regulated psychedelic treatment clinics. Financial issues such as the cost of treatment (relative to the intensity of therapist resources necessary) and whether insurance companies would reimburse such treatments may further limit the broader use of this type of treatment. Also, given that this is a highly specialized form of psychotherapy, there might be a requirement for therapists to undergo specific training in a psychedelic psychotherapy training program and to obtain special certification to use the treatment. However, on the other hand, if limited pharmacological-type interventions (i.e., one to three drug sessions) can cause immediate and sustained clinical improvements, they would be ideal for insurance companies to reimburse because of the potential cost savings of such an intervention model especially compared to the normative daily and extended use of most pharmacological agents (i.e., antidepressants) for most psychiatric disorders.

COMPLICATIONS

Physiological Toxicity

In general, classic hallucinogens possess remarkably low physiological toxicity and are not associated with end-organ damage, carcinogenicity,

teratogenicity, lasting neuropsychological deficits or overdose fatalities. The classic hallucinogens produce sympathomimetic effects and can moderately increase pulse as well as diastolic and systolic blood pressure, although this has not been associated with cardiac, neurological, or other organ damage. Common physiological side effects of the classic hallucinogens include: mydriasis, blurry vision, dizziness, tremors, weakness, paresthesias, and increased deep tendon reflexes.

Psychological Toxicity

The acute psychological and behavioral effects of the classic hallucinogens are greatly influenced by set (personality and expectations of the individual), setting (environmental conditions and context of use) and dose, with the factors combining to influence the valence (positive or negative) of the experience. Affective changes can range from euphoric or ecstatic spiritual states to anxiety, terror, and panic. Perception is intensified and amplified with alterations in time, space, and boundaries between self and others. Synesthesia is common with mixing of various sensory stimuli (e.g., hearing colors). Sensory illusions (e.g., walls breathing) are common, and frank hallucinations can occur, though less frequently. Thought processes are loosened, with effects ranging from increased creativity to thought disorder. Cognition is altered and can range from increased and sudden insight ("noetic" effect) to confusion and disorientation. The sum total of the experience can range from positive mystical type experiences associated with enduring positive changes in affect/cognition/behavior to very unpleasant experiences dominated by fear and dysphoria. Severe adverse psychological experiences ("bad trips") tend to occur in poorly prepared individuals who use the substance in an uncontrolled setting and who have psychological risk factors (e.g., severe mental illness, recent trauma). These experiences typically include anxiety, panic, dysphoria, depersonalization, paranoid ideation, fear that the experience will never end, and fear of losing one's mind. Despite such adverse reactions, users usually retain insight into the fact that their symptoms are related to drug ingestion and they usually respond to verbal reassurance. Classic hallucinogens can acutely engender frank psychosis marked by hallucinations, thought disorder, and delusions, although this is rare in individuals without underlying psychotic spectrum illness. Such adverse psychological experiences can potentially lead to dangerous behavior toward self or others especially when used by vulnerable individuals without proper preparation and supervision. In a review of LSD research conducted during the 1950s in thousands of research participants who were carefully screened and prepared for dosing sessions, the reported suicide attempt and completed suicide rates were very low at 1.2 and 0.4, respectively, per 1,000 patients.

In the modern era, over 2,000 doses of psilocybin (ranging from low-to-

high doses) have been administered safely in the United States at major academic medical centers (Johns Hopkins, NYU, UCLA, University of New Mexico, University of Arizona) with no reports of any medical or psychiatric serious adverse events attributed to psilocybin (including acute or prolonged psychosis or HPPD). In addition, approximately 200 doses of intravenous DMT were administered at the University of New Mexico to normal participants without any SAEs and LSD has now been administered to approximately 2 dozen research participants in Europe in the last several years (Switzerland, England) without any SAEs reported. The key commonality among all these studies is the careful attention to set, setting, dose, session preparation, and integration.

Contraindictions

Given the known link between SHs and their ability to initiate or exacerbate psychosis in patients with underlying psychotic spectrum illness, they are contraindicated in anyone with a significant history of psychosis whether it be from a primary psychotic spectrum illness (i.e., schizophrenia, schizoaffective disorder), affective psychotic illnesses (i.e., bipolar I disorder with psychotic features, major depression with psychotic features), severe personality disorders associated with micro or frank psychosis, patients with borderline to psychotic level ego function, severe anxiety disorders associated with psychotic phenomenon (OCD with psychotic features), or chronic psychosis due to drugs (i.e., alcohol-induced psychotic disorder, Korsakoff psychosis) or due to general medical conditions. Furthermore, they would be contraindicated in patients at high risk for psychosis (i.e., strong family history of psychosis), children and adolescents. Age limitations should exist (i.e., ê25 years old), given the potential risk for precipitation of psychosis (i.e., schizophrenia) in vulnerable individuals (even those without any family history). The SHs are also contraindicated in individuals who have chronic HPPD.

Case Examples

Because the process of treatment with classic hallucinogens is so different from other pharmacotherapies, it may be useful to provide descriptions of the treatment process in specific cases. Below two brief case reports from the authors' clinical trials are presented: one of a cancer patient struggling with depression and anxiety, and the other a patient treated for alcohol addiction. Details of these cases have been altered to obscure the identity of these individuals. The cases include elements of both psychedelic and psycholytic-assisted psychotherapies.

PSILOCYBIN-ASSISTED TREATMENT OF CANCER-RELATED PSYCHOLOGICAL AND EXISTENTIAL DISTRESS

E is a 22-year-old woman, a first year law student, with recent diagnosis of leukemia. She had

no prior medical or psychiatry history. She was treated successfully for her leukemia with chemotherapy and was told by her oncologist she had a very good prognosis and was likely cured of her cancer. For the first 6 months following diagnosis and treatment, she did not experience any adverse psychological effects of the cancer diagnosis. However, at the 6-month point of remission, E started to become frightened that the cancer could recur. Even though she understood the good prognosis, she experienced increasing anxiety about the cancer returning and what that would mean for her. These thoughts continued to get worse to the point where E started experiencing panic attack–like events related to an acute fear of death. She had always thought of death as a distant event, something that she would not have to contend with for decades. However, the cancer provoked a kind of crisis and she began to feel that there was now a short distance between herself and death. She began to experience a dysphoric and anxious mood accompanied by anhedonia, poor concentration, preoccupation with death, and a desire for a hastened death because she felt that the end was imminent. From a Catholic faith background, she became angry at God for allowing her to get cancer and started for the first time in her life to have doubts about the existence of God. She stopped attending church and felt she had lost access to the nurturing parts of her religion and spirituality. She told her oncologist about these symptoms, and her oncologist referred her to the NYU Psilocybin Cancer Project. E was a casual user of alcohol and had no prior use of other drugs including hallucinogens and marijuana. She was intrigued by the use of psilocybin to treat anxiety but admitted that she was skeptical that it would actually do anything to decrease her anxiety or fear of death related to her cancer diagnosis. She felt it was worth a try given her level of distress and because she trusted her oncologist's judgment.

E underwent her first dosing session and happened to be randomized to the group that received psilocybin (0.3 mg/kg) first. On the day of the session, she reported having some trouble sleeping the night before and being somewhat apprehensive about the dosing session. She wanted to make sure that she would return to normative reality when the session ended. Her treatment dyad team did some deep breathing with her before the session and reassured her of all the ways the team would address any difficult or anxious psychological experiences. E received her psilocybin capsule at 9 AM. The treatment team checked in every hour to assess E's mental state. For the first 3 hours, E reported nothing suggestive of a psychoactive experience (unusual for those that receive psilocybin whose effects usually are experienced in the first hour postdosing). She told the dyad team that she was sure she must have gotten the placebo. At approximately 12:30, E sat up and took off her eye shades. She reported that the music sounded "trippy." When the dyad team inquired more, she said that she felt anxious and wondered if she would feel normal again. After verbal reassurance, E lay back down and placed the eye shades on. At approximately 2 pm, she sat up and reported the following visions: she journeyed back to parts of her childhood and came upon multiple scenes of her and her family (two parents, and younger brother) joined together in laughter and love that made her feel a sense of unity and support; she reported a scene where she saw herself standing in front of a hospital and saw a cauldron of black, hot smoke circling inside of her body (which she interpreted as her cancer), followed by her boyfriend and family forming a loving circle around her which caused the black smoke inside of her to disappear outside of her body and into the sky (she interpreted this as a cathartic experience of the fear of cancer leaving her); she reported seeing her cancer-riddled body die in front of her eyes while the treatment providers and a transcendental entity together solved a complicated mathematical equation (she interpreted this as God and the treatment team working together to figure out the biological nature of her cancer-related anxiety). At the end of the experience, E declared that she felt "reborn" and freed from the anxiety and fear of death caused by her cancer diagnosis. Her scores on the main outcome measures of anxiety and depression (i.e., Hospital Anxiety and Depression Scale, Beck Depression) dropped dramatically from very high clinically significant values at baseline to almost zero, 7 hours postdosing with psilocybin. Her follow-up distress scores on all of these scales remained close to zero for the next 6 months until the final follow-up point. It is now 3.5 years after her treatment, and she continues to report no anxiety or distress associated with her cancer diagnosis. She is flourishing as a law student and reports a sense of calmness and psychological well-being that she continues to attribute to the psilocybin dosing experience. She has taken up a meditation practice and still can remember salient aspects of the experience that have stuck in her memory.

PSILOCYBIN-ASSISTED TREATMENT OF ALCOHOLISM

C is a 59-year-old divorced mother of two who had struggled with alcohol since age 15. Her drinking had led to problems including recurrent physical violence, multiple arrests, poor work history, and intermittent homelessness. She had suffered severe abuse in the context of relationships with partners who also drank, including being beaten unconscious and suffering from intracranial bleeding on at least one occasion. She had made several past attempts to stop drinking, with little success. When she volunteered for the psilocybin trial, she had been sober for 11 days, and had been drinking 7 out of the past 84 days, an average of 16.9 standard drinks per drinking day. During the preparatory phase of treatment in the study, she stated a goal of total abstinence, and rated the importance of abstinence as high and her readiness for abstinence as high, but her confidence in achieving it was low. She said that she wanted to understand why she drank, and hoped that this would help her stay sober. She listed God's will, forgiveness, humility (to be) loved, and self-control as her most important values, and saw clearly that her drinking was in conflict with these values.

During her first psilocybin session, she reported that she experienced powerful feelings of sorrow and remorse regarding the course of her life, and particularly concerning her perceived failures as a parent as a result of her drinking. This experience was quite painful, and she believed that she was sobbing uncontrollably during much of this time, although she was actually lying quietly on the couch at the time. After the session, she felt a sense of relief, and said that she had been able to let go of these feelings and experience a sense of forgiveness. She was hopeful that the experience would help her stay sober, and had no desire to drink after the session. C remained sober between the first and second psilocybin session. During the second session, she reported that she experienced a visual image of a small child lying "broken" on the floor. She realized that this child was her, and experienced herself as a 3-year-old child, devastated by abandonment by her father, an issue that she had not discussed in the preparatory sessions. After this, she began to perceive a white light, which she called "God's healing light," and felt a profound sense of love. She felt that she had been healed by this experience, and that she now felt "whole" and worthy of love. In discussing these experiences afterward, C said that she thought her drinking had been an attempt to escape the painful feelings of being unworthy of love, as well as the painful feelings of shame and loss related to her life as an alcoholic. She had avoided these feelings, believing that she would "fall apart" if she faced them. Following the sessions, she now felt that she was strong enough to face these feelings, and that she was a whole person, worthy of love. At her most recent follow-up, 5 months after the first psilocybin session, she remained abstinent and continued to feel that her life had been transformed, in spite of the unexpected death of a close family member during the interim.

Ethical Issues

The ethical concerns involved in the hallucinogen-assisted therapies are the same as those for other forms of psychotherapy and pharmacotherapy, but there are particular issues that require special attention due to the nature of the process and the possible effects. Patients must be informed of the risks and benefits of any therapy so they can make an informed decision about whether to proceed. Although the medical risks of hallucinogen administration are minor and easily described, the nature of the experience and the attendant psychological risks are much harder to describe, particularly if the patient has no prior experience with hallucinogens. It is important that the therapists describe the full range of possible experiences. This should be done in a way that prepares the patient as well as possible for both pleasant and challenging experiences, without minimizing possible difficult aspects or unnecessarily heightening anxiety about them. Undergoing any form of psychotherapy can result in changes

that are difficult to predict, and may not always be desired. These could include recognition of painful or traumatic experiences, relationship difficulties, or strong emotions that were unconscious or dormant. In hallucinogen-assisted therapy, mystical experiences can result in changes in beliefs about the nature of reality. All of these changes can have unpredictable consequences, for example, desire for changes in relationships, career, or other important life domains. None of this is unique to hallucinogen-assisted therapy, but such changes, if they occur, can happen very rapidly due to the nature of the process. Finally, as in all therapy relationships, there is a power differential, and patients are vulnerable to exploitation. In hallucinogen-assisted therapy, patients are often incapacitated during the drug administration sessions, and are in a state of heightened suggestibility during and possibly after the sessions. Therapists must resist any temptation to misuse power during the therapy, including being overly directive during the drug administration sessions.

RESEARCH AND EVALUATION

Use of Psychedelics in the Treatment of Addiction

Lysergic Acid Diethylamide. The use of LSD in the treatment of alcoholism was studied extensively in the 1950s through early 1970s. The earliest investigations in Saskatchewan, led by the pioneering psychedelic researcher Humphrey Osmond, were based on a model which held that the LSD experience mimicked the experience of delirium tremens, and that this negative experience of "hitting bottom" could induce abstinence in some cases. However, based on their experience and the recommendations of A. M. Hubbard, the treatment was modified to facilitate the positive experience of self-surrender, consistent with the psychedelic model. Treatment of alcoholism with LSD became an accepted clinical treatment in Saskatchewan and several thousand patients would end up being treated by Dr. Osmond and colleagues using this model. Uncontrolled trials with severe, chronic alcoholic patients using a single high-dose LSD session had variable but generally encouraging results.

Controlled trials of LSD for alcoholism were largely underpowered and had mixed results, leading reviewers to conclude that few conclusions could be drawn from these studies. However, a recent meta-analysis revealed consistent and clinically meaningful effects of LSD over control treatment in the six randomized trials of LSD for alcohol dependence that reported drinking outcomes. Across the 6 studies, 325 participants received active treatment with LSD, and 211 received control treatment. In all of the studies, LSD was administered in a single high-dose session, in doses ranging from about 210 to 800 mcg. At the first posttreatment follow-up, LSD-treated patients were more likely to show significant improvement in drinking outcomes (odds ratio = 1.96, 95 percent CI 1.36 to 2.84, p =

0.0003). Treatment effects remained significant at 6 months. The effect of LSD treatment was homogeneous across the six studies, in spite of great variability in the psychotherapeutic components of treatment and the control treatments employed. These promising findings strongly support renewed clinical investigation of LSD for the treatment of alcoholism.

Studies were also conducted using LSD as a component of treatment for opioid addiction. Seventy "postnarcotic drug addicts" at the US Public Health Service Hospital in Lexington, Kentucky received a single 2- to 3-hour therapeutic session in which they were randomly assigned to receive one of five treatments: (1) insight-oriented psychotherapy, (2) hypnotherapy (psychotherapy conducted under hypnosis), (3) LSD with no psychotherapeutic intervention, (4) LSD with psychotherapy, or (5) LSD with hypnotherapy. At 2-month follow-up, all groups showed significant improvement on a questionnaire designed to measure various dimensions of psychopathology, with greater improvement in the group that received LSD with hypnotherapy. Drug use behavior after discharge from the hospital was not investigated. The dose of LSD used in this study (0.2 mcg/kg, or 140 mcg for a 70-kg person) was lower than the doses used in the controlled alcohol trials summarized above. Savage and McCabe randomly assigned 78 incarcerated male heroin addicts to outpatient treatment as usual versus psychedelic therapy followed by treatment as usual. The psychedelic therapy included a single high-dose (300 to 500 mcg) LSD session in the context of 4 to 6 weeks of residential treatment that included extensive psychotherapy before and after the session. Participants who received psychedelic therapy had better outcomes during the 12-month follow-up period (25 percent vs. 5 percent continuous abstinence). The design of this study did not separate the effects of the LSD from other aspects of the residential treatment.

Psilocybin. The effects of psilocybin were described over 50 years ago and at least 2,000 individuals have received psilocybin in clinical and laboratory studies (since the early 1990s) using the *psychedelic* and *psycholytic* models of treatment. However, other than very brief reports of psilocybin used in combination with LSD the first trial of psilocybin used to treat alcoholism was completed very recently. In this proof-of-concept study, ten volunteers with DSM-IV alcohol dependence received orally administered psilocybin in one or two supervised sessions scheduled 4 weeks apart. Psilocybin was administered in the context of a 12-week manualized therapy program, in doses of 0.3 mg/kg and 0.4 mg/kg. Drinking decreased significantly following psilocybin administration, and gains remained significant during 36 weeks of follow-up. The intensity of self-reported effects during the first psilocybin session at week 4 strongly predicted improvement in drinking during weeks 5 to 8 (r = 0.76 to r = 0.89). A double-blind, multi-center, efficacy RCT using a similar model of

treatment is currently under way being conducted at the New York Langone Medical Center (NCT02061293) and the New Mexico School of Medicine. Although there were not published trials of classic hallucinogens administered in the treatment of nicotine addiction, a recent open-label pilot study, utilizing a model of high-dose psilocybin assisted CBT, has reported remarkably positive outcomes using psilocybin in the treatment of this addiction. Fifteen nicotine-dependent smokers received psilocybin (20 to 30 mg/70 kg in 2 to 3 sessions) during a 15-week course of manualized therapy. At 6-month follow-up, 12/15 (80 percent) were abstinent based on self-report and verified by biological measures (urinary cotinine and breath carbon monoxide). This success rate far exceeds those seen in clinical trials of any currently available pharmacotherapies for tobacco addiction. Measures of mystical experience during psilocybin sessions and of spiritual significance and personal meaning attributed to the sessions were significantly correlated with smoking outcomes. Based on these very promising results, a comparative efficacy trial (psilocybin vs. nicotine replacement) is now being conducted at the Johns Hopkins School of Medicine (NCT01943994). In addition, a pilot RCT is underway investigating the effects of psilocybin-assisted psychotherapy in the treatment of cocaine addiction (NCT02037126) at the University of Alabama School of Medicine. Given the persistent lack of effective medications to treat cocaine addiction, this study will be the first to use a classic hallucinogen in the treatment of cocaine use disorder.

Use of Psychedelics to Ameliorate Cancer-Related Psychological and Existential Distress

Psychological and Existential Distress in Patients with Cancer. With limited evidence from randomized controlled trials (RCTs) in cancer populations, there are no well-established evidence-based treatment algorithms to guide the optimal pharmacological or psychosocial treatment of cancer patients with depressive and anxiety spectrum disorders; rather, treatment guidelines have had to be derived from the available research conducted in cancer populations as well as from the more extensive research in psychiatric and other medical illness populations. Also, given the relationship between spiritual well-being and improved psychiatric outcomes in patients with advanced cancer (most importantly decreased depression, decreased hopelessness, decreased desire for hastened death, and possibly decreased suicide), it would make sense to develop treatment (psychosocial, pharmacological, or combined pharmacological–psychosocial) paradigms to specifically target such spiritual/existential distress when it manifests. Although a handful of manualized existentially oriented psychotherapies have been developed to address the existential/spiritual issues faced by patients with advanced/terminal cancer, with some empirical support from clinical trials, there are no

current pharmacologic-psychosocial interventions to treat this typology of distress in cancer patients.

History of Psychedelic Therapy with Terminally Ill Patients. The first scientific exploration of psychedelics as stand-alone pharmacological interventions to help the dying in a medical setting occurred in the early 1960s with Eric Kast, M.D., an internist and psychiatrist at the University of Chicago Medical School. Kast was a specialist in pain medicine and became interested in studying LSD as a novel analgesic agent in patients with terminal cancer and pain syndromes by exploring LSDs ability to alter pain attention and perception. In his first paper published in 1964 on the topic, Kast conducted a blinded comparative efficacy trial of LSD (100 mcg orally) compared to Demerol 100 mg and dilaudid 2 mg in a sample of 50 gravely ill patients, mostly with terminal cancer but also included those with severe burns and infectious illnesses. Kast reported that the LSD group had statistically significant reductions in pain compared to the two opioid groups, from 3 hours postdosing to up at least 19 hours postdosing. In addition, he noted that the patients in the LSD group appeared to display a type of detachment from the fear of dying that he thought was useful given their grave medical conditions. Based on this observation, he extended his LSD research by administering LSD 100 mcg to over 200 patients with terminal cancer and pain syndromes in an open-label design and paid increasing attention to psychological phenomenon (in addition to pain perception) such as sleep, affective changes, and attitudes toward death and dying. In further published research, Kast confirmed his earlier findings by observing: decreased pain perception acutely and lasting up to 2 weeks postdosing, improved sleep and mood, improved communication between treatment provider and patients, enhanced morale and outlook on life, reports of mystical-type experiences ("happy, oceanic feelings"), enhanced philosophical and spiritual states, and decreased fear of cancer diagnoses and death. It is important to note that Dr. Kast viewed LSD treatment in his patients as a type of pure pharmacological intervention and did not account for or introduce the well-known components of set, setting, dose titration, and psychotherapeutic preparation and integration relative to the dosing sessions. The other significant historical research utilizing psychedelics to treat psychological distress associated with advanced or terminal cancer occurred from the early 1960s to the mid-1970s at the Spring Grove State Hospital in Maryland, a research affiliate of the Johns Hopkins School of Medicine and the University of Maryland School of Medicine at the time. Starting in 1963, a group of psychiatrists, psychologists, nurses, and social workers began examining LSD-assisted psychotherapy to alcoholics. In 1965, when one of the nurses on the research team became ill with metastatic breast cancer, one of the research psychologists (Sidney Wolf) suggested that a course of psychedelic therapy

might help alleviate her anxiety associated with cancer. After she underwent LSD-assisted treatment and reported relief, the researchers at Spring Grove decided to embark on research utilizing LSD-assisted psychotherapy (a pharmacological–psychosocial paradigm) to help patients with terminal cancer and psychological distress. This research project began in 1967 and was headed by the psychiatrists Stanislav Grof and Walter Pahnke. The treatment model utilized was significantly different from the pharmacological only model developed by Dr. Kast. It was based on Dr. Grof's LSD research in administering moderate-to-high doses of LSD to normal volunteers (developed while doing research at the Psychiatric Research Institute in Prague, Czechoslovakia), paying careful attention to set, setting, as well as session preparatory psychotherapy and postsession integrative psychotherapy. The treatment model drew from psychoanalytic theory as well as transpersonal psychology, and Grof commented that "many individuals who had the experience of death and rebirth sometimes accompanied by feelings of cosmic unity independently reported that their attitudes towards dying and their concepts of death underwent dramatic changes." Fear of their own physiological demise diminished, they became open to the possibility of consciousness existing after clinical death, and tended to view the process of dying as an adventure in consciousness rather than the ultimate biological disaster. A total of 60 patients participated in the open-label experimental trial with 44 of the patients receiving LSD (200 to 500 mcg orally) and 19 patients receiving dipropyltryptamine (DPT) (60 to 105 mg IM). Systematic, longitudinal analyses performed on a subsample of 31 participants found statistically significant prepost reductions in the following domains: depression, anxiety, pain, isolation, and fear of death; in addition, a global index of improvement broke down as follows (29 percent dramatically improved, 42 percent moderately improved, 23 percent unimproved, 6 percent worse).

Phase II LSD-Assisted Psychotherapy for Anxiety Associated with Life-Threatening Disease. The reemergence of LSD-assisted psychotherapy clinical trials to treat psychosocial distress associated with life-threatening illnesses has resumed with a recently completed trial in Switzerland. This double-blind, randomized, placebo-controlled phase II trial with a crossover design (2 months) included a sample of 12 participants with life-threatening medical disorders (including cancer and other medical conditions) who received either the experimental condition (LSD 200 mcg orally) or the active control (LSD 20 mcg orally) in conjunction with psychotherapy. Safe administration of LSD was established in this small cohort with no reported serious adverse events. Compared to the active control, the experimental group had significant short-term (2-month follow-up) reductions in anxiety as measured by the State-Trait Anxiety Inventory (STAI). A recent 12-month follow-up study of participants from

the above trial reported sustained reductions in anxiety (as measured by the STAI), increases in quality of life, and no reports of any adverse psychological or medical sequelae.

PHASE II AND PHASE III PSILOCYBIN-ASSISTED PSYCHOTHERAPY FOR CANCER-RELATED PSYCHOLOGICAL DISTRESS RANDOMIZED CONTROLLED TRIALS. RCTs utilizing psilocybin-assisted psychotherapy to treat psycho-social-spiritual distress (depression, anxiety, existential suffering) associated with advanced or terminal cancer have resumed in the United States within the last two decades as part of a research collaboration between UCLA, NYU, and Johns Hopkins. At all three sites, the treatment model was similar and was based on the Spring Grove psychedelic-assisted psychotherapy model, developed by Dr. Grof and colleagues. Important similarities among the three sites were as follows: (1) Double-blind design methodology; (2) Randomization among groups; (3) Use of validated outcome measures for cancer-related psychological and existential distress; (4) The use of exclusion criteria such as major unstable medical illnesses (i.e., cardiac/renal/hepatic failure) and major mental illness (i.e., schizophrenia, bipolar spectrum disorder, or a family history of these disorders); (5) The participants were carefully prepared for the experimental sessions by trained monitors after a thorough life review and review of their cancer diagnosis narrative with an emphasis on spiritual/existential themes of distress; (6) Participants underwent the dosing sessions in a comfortable living room–like setting designed for maximal comfort and safety; (7) To increase the likelihood of the induction of mystical states of consciousness, participants were instructed to lie down on a couch with eyeshades to reduce external visual distractions with a preselected music program played and with instructions to focus on their inner experiences during the session; (8) Integration of the experience during the dosing sessions as part of postintegrative psychotherapy. The three studies differed by: (1) Dose of psilocybin—UCLA (0.2 mg/kg); NYU (0.3 mg/kg)/JHU (0.43 mg/kg); (2) the active control: niacin (UCLA, NYU), low-dose psilocybin (JHU); (3) Inclusion of nonterminally ill cancer patients (NYU, JHU).

The first double-blind placebo-controlled phase II trial of single low-dose (0.2 mg/kg) psilocybin-assisted psychotherapy for advanced-stage cancer was conducted at Harbor-UCLA (N = 12) and suggested an acute and short-term sustained antidepressant and anxiolytic effect associated with psilocybin versus placebo (Grob, 2011). Two other RCTs of single-dose psilocybin-assisted psychotherapy for patients with cancer related psychosocial and existential distress (NYU, N = 29; Johns Hopkins, N = 51) have recently been published. The findings are strikingly similar in a number of ways in terms of the efficacy of psilocybin: (1) Rapid/acute acting anti-depressant and anxiolytic responses with substantial effect sizes and clinically meaningful results (i.e., approximately 80 percent with

clinically significant anti-depressant or anxiolytic responses); (2) Enduring clinically significant anti-depressant and anxiolytic effects for as long as 5–7-weeks, but possibly up to 6–8 months with 60–80 percent meeting criteria for clinically meaningful anti-depressant or anxiolytic responses at 6–6.5 month final follow-up; (3) Decreases in existential distress (i.e., cancer-related demoralization, hopelessness, meaninglessness of life); (4) Improved anxiety or attitudes towards death; (5) Improved quality of life; (6) Mystical experiences induced by psilocybin; (7) Psilocybin-induced mystical experience partially mediated the therapeutic effect of psilocybin on anxiety and depression; (8) Substantial personal meaningfulness of the experiences; (9) Improved spiritual well-being. Overall, the systematic replication of the main outcomes across two sites contributes to the confidence of the robustness of the finding that a single dose of psilocybin has rapid, substantial, and enduring anti-depressant and anxioloytic benefits in patients with life-threatening cancer diagnoses.

Phase III Trial. The Heffter Research Institute (the sponsor of the above-mentioned phase II RCTs) and the Usona Research Institute are currently in discussions with FDA regarding the optimal design of a potential phase III RCT of psilocybin-assisted psychotherapy for psychological distress associated with a cancer diagnosis. If such a trial would demonstrate a statistically and clinically meaningful therapeutic effect of psilocybin-assisted treatment (vs. placebo) on psychological/existential distress in cancer, it would potentially lead to the historic rescheduling of psilocybin from schedule I to another schedule (likely II or III) to be made available as a prescription medication to be used (likely only to be dispensed and administered in specialized designated and regulated psychedelic treatment clinics) as an adjunct to psychotherapy to treat psychological and existential/spiritual distress associated with cancer, whether it occurs early in the illness or as part of the death and dying process. Training of therapists and fidelity to treatment models would have to be considered in and it would be important to establish specific psychedelic therapy training programs. Oversight, quality control, and strict measures to reduce adverse psychological events would have to be key components of maintaining a psychedelic treatment clinic.

Treatment Resistant Depression (TRD). The acute and sustained anti-depressant effects of psilocybin in the above 2 psilocybin cancer trials are consistent with a recently published open-label study (N = 12) of oral psilocybin treatment in patients with treatment-resistant depression (TRD), where psilocybin (25 mg) was associated with 1-week and 3-months post-psilocybin anti-depressant effects (Carhart-Harris et al., 2016). As a follow-up to this open-label trial, several academic medical centers (NYU,

Hopkins, UCSF, Imperial College, U Zurich) are planning RCTs using oral psilocybin treatment for TRD.

REFERENCES

Abuzzahab FS Sr, Anderson BJ. A review of LSD treatment in alcoholism. *Int Pharmacopsychiatry*. 1971;6(4):223–235.

*Bogenschutz MP, Forcehimes AA, Pommy JA, Wilcox CE, Barbosa PC, Strassman RJ. Psilocybin-assisted treatment for alcohol dependence: A proof-of-concept study. *J Psychopharmacol*. 2015;29:289–299.

Carhart-Harris RL, Bolstridge M, Rucker J, et al. Psilocybin with psychological support for treatment-resistant depression: an open-label feasibility study. *Lancet Psychiatry*. 2016;3(7):619–627.

Carhart-Harris RL, Kaelen M, Whalley MG, Bolstridge M, Feilding A, Nutt DJ. LSD enhances suggestibility in healthy volunteers. *Psychopharmacology*. 2015;232(4):785–794.

Carhart-Harris RL, Leech R, Hellyer PJ, et al. The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Front Hum Neurosci*. 2014;8:20.

Chwelos N, Blewett DB, Smith CM, Hoffer A. Use of d-lysergic acid diethylamide in the treatment of alcoholism. *Q J Stud Alcohol*. 1959;20:577–590.

Cohen S. Lysergic acid diethylamide: side effects and complications. *J Nerv Ment Dis*. 1960;130:30–40.

Gasser P, Holstein D, Michel Y, et al. Safety and efficacy of lysergic acid diethylamide-assisted psychotherapy for anxiety associated with life-threatening diseases. *J Nerv Ment Dis*. 2014;202(7):513–520.

Gasser P, Kirchner K, Passi T. LSD-assisted psychotherapy for anxiety associated with a life-threatening disease: a qualitative study of acute and sustained subjective effects. *J Psychopharmacol*. 2015;29(1):57–68.

Griffiths RR, Johnson MW, Carducc MA, et al. Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: a randomized double-blind trial. *J Psychopharmacol*. 2016;30(12):1181–1197.

Griffiths RR, Johnson MW, Richards WA, Richards BD, McCann U, Jesse R. Psilocybin occasioned mystical-type experiences: immediate and persisting dose-related effects. *Psychopharmacology (Berl)*. 2011;218(4):649–665.

*Griffiths RR, Richards WA, Johnson MW, McCann UD, Jesse R. Mystical-type experiences occasioned by psilocybin mediate the attribution of personal meaning and spiritual significance 14 months later. *J Psychopharmacol*. 2008;22(6):621–632.

Griffiths RR, Richards WA, McCann U, Jesse R. Psilocybin can occasion mystical-type experiences having substantial and sustained personal meaning and spiritual significance. *Psychopharmacology (Berl)*. 2006;187(3):268–283; discussion 284–292.

Grinspoon L, Balakar JB. *Psychedelic Drugs Reconsidered*. New York: The Lindesmith Center; 1997.

*Grob CS, Danforth AL, Chopra GS, et al. Pilot study of psilocybin treatment for anxiety in patients with advanced-stage cancer. *Arch Gen Psychiatry*. 2011;68(1):71–78.

Grof S, Halifax J. *The Human Encounter with Death*. New York: EP Dutton; 1977.

Grof S, Soskin RA, Richards WA, Kurland AA. DPT as an adjunct in psychotherapy of alcoholics. *Int Pharmacopsychiatry*. 1973;8(1):104–115.

Halpern JH, Sherwood AR, Passie T, Blackwell KC, Ruttenber AJ. Evidence of health and safety in American members of a religion who use a hallucinogenic sacrament. *Med Sci Monit*. 2008;14(8):SR15–SR22.

Hollister LE, Shelton J, Krieger G. A controlled comparison of lysergic acid diethylamide (LSD) and dextroamphetmine in alcoholics. *Am J Psychiatry*. 1969;125(10):1352–1357.

*Johnson MW, Garcia-Romeu A, Cosimano MP, Griffiths RR. Pilot study of the 5-HT2AR agonist psilocybin in the treatment of tobacco addiction. *J Psychopharmacol*. 2014;28(11):983–992

Johnson M, Richards W, Griffiths R. Human hallucinogen research: guidelines for safety. *J Psychopharmacol*. 2008;22(6):603–620.

Kast E. LSD and the dying patient. *Chic Med Sch Q*. 1966;26(2):80–87.

Kast EC, Collins VJ. Study of lysergic acid diethylamide as an analgesic agent. *Anesth Analg*. 1964;43:285–291.

*Krebs TS, Johansen PO. Lysergic acid diethylamide (LSD) for alcoholism: meta-analysis of randomized controlled trials. *J Psychopharmacol*. 2012;26(7):994–1002.

LeMay K, Wilson KG. Treatment of existential distress in life threatening illness: a review of manualized interventions. *Clin Psychol Rev*. 2008;28(3):472–493.

Li M, Fitzgerald P, Rodin G. Evidence-based treatment of depression in patients with cancer. *J Clin Oncol*. 2012;30(11):1187–1196.

Ludwig A, Levine J, Stark L, Lazar R. A clinical study of LSD treatment in alcoholism. *Am J Psychiatry*. 1969;126(1):59–69.

Pahnke WN, Kurland AA, Goodman LE, Richards WA. LSD-assisted psychotherapy with terminal cancer patients. *Curr Psychiatr Ther*. 1969;9:144–152.

Passie T, Seifert J. Schneider U, Emrich HM. The pharmacology of psilocybin. *Addict Biol*. 2002;7(4):357–364.

Ross S, Bossis A, Guss J, et al. Rapid and sustained symptom reduction following psilocybin treatment for anxiety and depression in patients with life-threatening cancer: a randomized controlled trial. *J Psychopharmacol*. 2016;30(12):1165–1180.

Strassman RJ, Qualls CR. Dose-response study of N,N-dimethyltryptamine in humans. I. Neuroendocrine, autonomic, and cardiovascular effects. *Arch Gen Psychiatry*. 1994;51(2):85–97.

Studerus E, Kometer M, Hasler F, et al. Acute, subacute and long-term subjective effects of psilocybin in healthy humans: a pooled analysis of experimental studies. *J Psychopharmacol*. 2011;25:1434–1452.

Traeger L, Greer JA, Fernandez-Robles C, Temel JS, Pirl WF. Evidence-based treatment of anxiety in patients with cancer. *J Clin Oncol*. 2012;30(11):1197–1205.

▲ 33.18 Mentalization-Based Treatment

ANTHONY W. BATEMAN, F.R.C.PSYCH., AND PETER FONAGY, PH.D.

INTRODUCTION AND DEFINITION

Mentalization-based treatment (MBT) is an evidence-based treatment for BPD. Randomized controlled trials (RCTs) have demonstrated its effectiveness in reducing the core symptoms of BPD and it is currently being studied for other conditions, including antisocial personality disorder (ASPD) and avoidant personality disorder, substance abuse, depression, and eating disorders.

This chapter will outline the theoretical basis of MBT and the core treatment model in relation to BPD.

What Is Mentalizing?

Mentalizing is the social cognitive ability to understand actions by other people and oneself in terms of mental states, including thoughts, feelings, wishes, and desires; it is a very human capability that underpins everyday interactions. In nontechnical language, it is attentiveness to thinking and feeling in oneself and others. It is beyond question that mental states influence behavior. Beliefs, wishes, feelings, and thoughts, whether within or outside people's awareness, always influence what people do. Mentalizing involves a whole spectrum of capacities: critically, this includes the ability to experience one's *own* behavior as coherently organized by mental states, and to differentiate oneself psychologically from others. These capacities are reduced in individuals with a personality disorder, who lose cognitive and emotional coherence, particularly at moments of interpersonal (relational) stress, which challenge one's mentalizing capacities. In the authors' view, many symptoms characteristic of BPD emerge in association with a distortion or reduction in mentalizing. Based on this understanding of BPD, MBT is a psychotherapy that focuses specifically on the mentalizing vulnerabilities of the patient in the context of an understanding of attachment process, which is the developmental context in which mentalizing is originally acquired.

HISTORY

The word "mentalizing" has been in existence for two centuries and in the *Oxford English Dictionary* for the past century. French psychoanalysts introduced the concept into the professional psychotherapy literature in the second half of the 20th century, and mentalizing came into the English professional literature on the cusp of the final decade, conceived of initially as a deficit in autism and as a transiently impaired process associated with profound insecurity in attachment relationships in the developmental psychopathology of BPD. This proposal was the origin for the development of MBT. The approach is rooted in attachment theory and psychoanalytic ideas, but in the interest of parsimony sheds many of the core assumptions of both theoretical approaches while negating neither body of work. Its

historical origin is born from the authors' wish to extend what they saw as the clear benefits of using a psychotherapeutic approach to professionals (nurses, rehabilitation or addiction counsellors, activity therapists) who normally did not have opportunities to undertake extensive training in particular psychotherapeutic modalities. The authors wished to extract core elements of a therapeutic approach that could be linked to experimentally observed distortions of mental function in personality disorders.

THEORETICAL ISSUES

The Role of Mentalizing in Therapy

MBT is based on the assumption that failure of mentalizing, while common to everyone, becomes a dominant feature in individuals with personality disorders, leading to both serious interpersonal problems and profound psychological distress. A number of therapies that have been shown to be successful in addressing the difficulties experienced by individuals with BPD appear to strengthen the patient's capacity to mentalize; for example, conversational therapy, cognitive analytic therapy, and certain aspects of dialectical behavior therapy. MBT is unique in attempting to understand the problematic and distressing aspects of severe personality disorder in terms of a failure of mentalizing (see later) and to focus a structured therapeutic approach on addressing problems of mentalizing as they occur in a therapy session. The overarching principle of MBT is to enhance mentalizing in the context of the therapeutic relationship (in both individual and group sessions) by systematically addressing instances of nonmentalizing and using these moments of discourse as opportunities to work with the patient to achieve a fuller psychological understanding of behavior. The significance of this process rests in the hope of generalization from the therapeutic situation to the wider social context. It is in this wider context that, from the perspective of the MBT clinician, the key difficulties blocking the possibility of change lie for the patient with BPD.

The Multidimensional Nature of Mentalizing

Mentalizing is not an entirely stable, consistent, or unidimensional capacity. Neuroscience has identified four distinct components to mentalizing, which the authors have organized into dimensions that they see as helpful for therapists to identify in the clinical practice of MBT:

1. Automatic versus controlled mentalizing
2. Mentalizing the self versus others
3. Mentalizing with regard to internal versus external features
4. Cognitive versus affective mentalizing.

To mentalize effectively requires the individual not only to be able to maintain a balance across these dimensions of social cognition but also to

apply them appropriately according to context. Consistent favoring of one or other side (or *pole*) of these dimensions leads to distorted understanding of mental states, associated with profound social and emotional difficulties.

In an adult with personality disorder, consistent distortions of social cognition consequent on imbalanced mentalizing on at least one of these four dimensions would be evident. Commonly, one or more of the dimensions underperforms at one end, and consequently the opposite pole comes to dominate social cognition. For example, excessively emotional thinking emerges in the absence of cognitive mentalizing, or the influence of others dominates if subjective experience of self-states is reduced. From this perspective, different types of psychopathology can be distinguished on the basis of different combinations of impairments along the four dimensions. In other words, personality disorders (and, to some extent, other psychiatric disorders) can be understood according to different characteristic *mentalizing profiles*.

Automatic versus Controlled Mentalizing. The most fundamental dimension to mentalizing is the spectrum between *automatic* (or *implicit*) and *controlled* (or *explicit*) mentalizing.

Controlled mentalizing reflects a serial and relatively slow process, which is typically verbal and demands reflection, attention, awareness, intention, and effort. For example, a person might misunderstand someone, so they stop them and ask them to explain more clearly what is underlying their statements, or they focus their own mind to work out what their opinion may be. People tell stories to others about their mental states; they think back into their past and report how they felt long ago; they enjoy autobiographical coherence in their personal lives and tell others about it.

The opposite pole of this dimension, *automatic mentalizing,* involves much faster processing, tends to be reflexive, and requires little or no attention, awareness, intention, or effort. In day-to-day life and ordinary social interaction, most mentalizing tends to be automatic because most straightforward exchanges do not require more attention. People in conversation naturally take turns, adapt their tone and posture to others' emotional states, and reflexively take into account their knowledge.

Self versus Others. This mentalizing dimension involves the capacity to mentalize one's own state—the *self* (including one's own physical experiences)—or the state of *others*. The two are closely connected, and an imbalance signals vulnerability in mentalizing of both others and/or the self. Individuals with mentalizing difficulties are likely to preferentially focus on one end of the spectrum, although they may be impaired at both. Awareness of the mental states of others is in part mediated by neural structures that organize one's own actions (the *mirror neuron system*), moderated by explicit reflective processes that reinforce the self–other

distinction. If explicit mentalizing is weakened, the influence of the current mental state of the other will increase.

Internal versus External Mentalizing. Mentalizing can involve making inferences on the basis of the *external* indicators of a person's mental states (e.g., facial expressions, tone of voice, body posture) or figuring out someone's *internal* experience from what one knows about them and the situation they are in. This dimension does not simply refer to a process of focusing on the externally visible manifestations versus the internal mental state of others, it also applies to the self—it includes thinking about oneself and one's own mind state versus considering one's current (external) situation and physical (interoceptive) state. Someone who has poor access to and great uncertainty about their subjective experience, for example, as is often seen in individuals with BPD, may come to a conclusion about what they are feeling from the reactions of others as well as from observing their own behavior: for example, their legs are restless, therefore they must be feeling anxious.

Cognitive versus Affective Mentalizing. *Cognitive* mentalizing involves the ability to name, recognize, and reason about mental states (in both oneself and others), whereas *affective* mentalizing involves the ability to experience and understand associated *feelings* (again, in both oneself and others). Both are required for any genuine experience of empathy or true sense of self-coherence. Some individuals give undue weight to either cognitive or affective mentalizing. People with obsessional characteristics may be masters at explicating in detail the internal states of themselves and others, but this may be devoid of emotional content and meaning. Conversely, people with BPD are flooded with emotion and so prone to automatic process, reactivity, emotional contagion, and poor self–other differentiation. Intense emotion can disrupt the process of cognitive appraisal that normally helps to regulate it.

Dimensional Mentalizing Profile and BPD

Particularly when arousal increases, as is typical in the context of intense attachment relationships, individuals with BPD easily find themselves switching to automatic mentalizing. Stress and arousal, especially in an attachment context, bring automatic mentalizing to the fore and disengage the neural systems that are associated with controlled mentalizing. Under these conditions, interactions become nonquestioning precisely when they need to be more controlled and contextualized. Thinking becomes impulsive: the individual makes quick assumptions about others' thoughts and feelings, which are not reflected upon or tested. Logic is intuitive, unreasoned, and nonverbal; it is marked by an unwarranted certainty, which betrays its unreflective origin. As a consequence, patients show

severe impairments in interpersonal and intimate relationships; for example, they may be overly distrustful (paranoid) or, indeed, overly trustful (naive).

Patients with BPD may show excessive concern about their own internal state, that is, they *hypermentalize* in relation to the self. At the same time, these views of the self develop without reference to social reality, namely an awareness of how others perceive one. Failure to balance self-perception with sincere curiosity about how one is perceived by others can lead to exaggerations of the self-image, in both positive and negative directions. A balanced, adaptive form of self-mentalizing conditioned by the social context is lost.

Patients with BPD pay more attention to external indicators of mental states, and their initial ideas arising from automatic mentalizing are left unchecked by controlled/reflective mentalizing. For example, if the clinician frowns, perhaps pensively, the patient may interpret this as looking angry or disgusted with them; seeing the clinician look at the clock during a session can stimulate an internal state of overwhelming unease and an experience that the clinician wants to be rid of them when, in reality, the internal state of the clinician is concern about the time left to work on the issue being discussed. A focus on external features, in the absence of reflective mentalizing, makes an individual highly vulnerable in social contexts, as it generates interpersonal hypersensitivity.

Table 33.18–1.
Modes of Nonmentalizing: Psychic Equivalence

Clinical appearance	Certainty/suspension of doubt Absolute Reality is defined by self-experience Finality—"It just is" Internal is seen as equivalent to external
Clinician's experience	Puzzled Wish to refute Statement appears logical but obviously overgeneralized Not sure what to say Angry or fed up and hopeless
Intervention	Empathic validation with subjective experience Curious—"How did you reach that conclusion?" Presentation of clinician's puzzlement (marked) Linked topic (diversion) to trigger mentalizing then return to psychic equivalent area
Iatrogenic	Argue with patient Excessive focus on content Cognitive challenge

The Reemergence of Nonmentalizing Modes and BPD

While the dimensions of mentalizing can reflect anomalies in terms of

mechanisms, that is not what the clinician sees. What the patient and the mentalizing clinician experience is a product of a malfunctioning mentalizing system, driven by imbalances in the dimensions of mentalizing. The outcomes of these malfunctions can be grouped into three typical modes of subjectivity for the purpose of illuminating clinical experience. The modes are termed *psychic equivalence mode, teleological mode,* and *pretend mode*. These modes are summarized in Tables 33.18–1 through 33.18–3, which also outline MBT interventions to address each nonmentalizing mode.

The nonmentalizing modes are important for the clinician to recognize and understand, as they tend to emerge in the consulting room and reflect core aspects of the patient's experience. It is important to address them because they cause considerable interpersonal difficulties and result in destructive behaviors.

Table 33.18–2.
Modes of Nonmentalizing: Pretend Mode

Clinical appearance	Inconsequential talk/groundless inferences about mental states Lack of affect. Absence of pleasure Circularity without conclusion—"spinning in sand" (hypermentalizing) No change Dissociation—self-harm to avoid meaninglessness Body and mind decoupled
Clinician's experience	Boredom Detachment Patient agrees with clinician's concepts and ideas Identification with clinician's model Feels progress is made in therapy
Intervention	Probe extent Counterintuitive Challenge
Iatrogenic	Nonrecognition Joining in with acceptance as real Insight-orientated/skill acquisition intervention

Table 33.18–3.
Modes of Nonmentalizing: Teleological Mode

Clinical appearance	Expectation of things being "done" Outcomes in physical world determine understanding of inner state—"I took an overdose; I must have been suicidal" Motives of others are based on what actually happens Only actions can change mental process "What you do and not what you say"
Clinician's experience	Uncertainty and anxiety Wish to do something—medication review, letter, telephone call, extend session
Intervention	Empathic validation of need Do (or do not do) according to exploration of need

	Affect focus of dilemma of doing
Iatrogenic	Excessive "doing"
	"Prove" clinician cares in the belief it will induce positive change
	Elasticity (extending what clinician does, e.g., providing extra sessions, only to rebound with extra constraints) rather than flexibility

Psychic Equivalence Mode. In the *psychic equivalence mode*, thoughts and feelings become "too real" to a point where it is extremely difficult for the patient to entertain possible alternative perspectives. When mentalizing gives way to psychic equivalence, what is thought is experienced as being real and true, leading to what clinicians describe as "concreteness of thought" in their patients. Patients with BPD who are in this mode describe an overriding sense of certainty about their beliefs, for example, "the therapist does not like me" or "I am a wicked person." Such a state of mind can be extremely frightening, adding a powerful sense of drama and risk to life experiences. The sometimes extreme reactions of patients are justified by the seriousness and realness with which they can experience their own and others' thoughts and feelings. The vividness and bizarreness of subjective experience can appear as quasipsychotic symptoms and are also manifested in the physically compelling memories associated with trauma. More important to appreciate is that negative affect in this mode will be overwhelming and cause very deep distress to the sufferer. Psychic equivalence permits no alternative perspectives to be taken. The patient can feel locked into an extraordinarily profound sense of pain without the slightest possibility that matters could be any different from how they currently are. This state of hopelessness makes suicidality comprehensible.

Teleological Mode. In the *teleological mode,* states of mind are recognized and believed only if their outcomes are physically observable. Hence, the individual can recognize the existence and potential importance of states of mind, but this recognition is limited to very concrete situations. For example, affection is perceived to be true only if it is accompanied by physical contact such as a touch or caress. The teleological mode shows itself in patients who are imbalanced toward the external pole of the internal–external mentalizing dimension—they are heavily biased toward understanding how people (and they themselves) behave and what their intentions may be in terms of what they physically do. Impulsivity involves function in teleological mode and a heavy emphasis on the automatic pole of the automatic–controlled dimension. There is insufficient reflection concerning the impact of one's actions on others, or on oneself. In teleological mode, the individual cannot accept anything other than a physical action as a true expression of the other person's intentions. Friends of a patient might constantly assure the patient of their love and support, and yet none of that feels real: it does not address the "hole" that the patient falls into at certain, especially lonely, times when they feel

terrible emptiness. Feeling that interpersonal affection can only be real if it is accompanied by physical behavior explains some risky sexual behavior, but also the need to create physical distractions that help with the feeling that all verbal expressions of interpersonal affection are without real meaning (see case example, later). The teleological mode makes the need for action and generation of "real" change overwhelming. The so-called "manipulativeness" of patients with personality disorder is little more than the experience of a pressing need to feel genuine reactions from others in terms of actions rather than words.

Pretend Mode. In the *pretend mode,* thoughts and feelings become severed from reality. Taken to an extreme, this may lead to feelings of derealization and dissociation. Patients in pretend mode can discuss experiences without contextualizing these in any kind of physical or material reality, as if they were creating a pretend world. The patient may *hypermentalize* or *pseudomentalize,* a state in which they may say much about states of mind but with little true meaning or connection to reality. Attempting psychotherapy with patients who are in this mode can lead to lengthy but inconsequential discussions of internal experience that have no link to genuine experience and will achieve no change. The limited capacity to experience a sense of an internal world gives rise to a deep and extremely distressing sense of emptiness. Dramatic action—for example, sometimes violent action against the self—may be experienced as the best way of addressing such feelings.

In summary, imbalances within the dimensions of mentalizing predictably generate the nonmentalizing modes. Psychic equivalence is inevitable if emotion (affect) dominates cognition. Teleological mode follows from an exclusive focus on external features to the neglect of the internal. Pretend mode thinking and hypermentalizing are unavoidable if reflective, explicit, controlled mentalizing is not well established.

Attachment

It is a central tenet of the mentalization-based approach that a sense of self and the capacity to mentalize both develop in the context of attachment relationships. The child observes, mirrors, and then internalizes their attachment figures' ability to represent and reflect mental states. So the reflections need to be *contingent*—that is, related to the child's internal experience—accurate, and *marked*—that is, indicating, for example, using "motherese" (a special tone of voice), that what is being expressed is a representation of the mind of the child and not that of the caregiver. Hence, the self and others—and the capacity to reflect on the self and others—are inevitably closely intertwined. Disorders that are characterized by severe impairments in feelings of self-identity, a central tenet of the pathology of BPD—, are also characterized by severe deficits in the ability to reflect

about others' mental states. However, this should not be taken to mean that an individual whose capacity to mentalize themselves is impaired will always show similar impairments in their ability to mentalize others. For example, individuals with ASPD can often be surprisingly skilled in "reading the mind" of others, but typically lack any real understanding of their own inner world.

In BPD there is commonly a history of early (in particular emotional) neglect, and a disrupted early social environment in general, and this may contribute to undermining the ability of some individuals to develop full mentalizing capacities. Subsequent adversity or trauma may further disrupt mentalizing, in part as an adaptive maneuver on the part of the individual to limit exposure to a brutalizing psychosocial environment, and in part because the high level of arousal generated by attachment hyperactivation and disorganized attachment strategies serve to disrupt less well-practiced and less robustly established higher cognitive capacities. In addition, genetic influences may be expressed through the mediation of mentalizing.

In sum, the mentalizing model is not strictly an etiological model, although it clearly prioritizes a social psychiatry perspective; it points to a final common developmental pathway that a range of biological, family, and broader social contextual influences may take to generate the range of difficulties that are normally considered under the term "personality disorder."

TECHNIQUES

MBT is a group and individual treatment. It is anticipated that, at times, the patient will experience strong affect while focusing on identified problems in treatment sessions and their mentalizing will be limited or failing, and/or the patient's understanding of the way mental states link to behavior is inadequate.

The clinician addresses this by a structured process (the sessional intervention trajectory) of:

1. Empathy and validation about problem areas
2. Clarification, exploration, and, where necessary, challenge
3. Following a structured process to gently expand mentalizing and encourage the patient to identify the mental states previously outside their awareness.

The process is primarily in the here and now of the session but increasingly, as the patient's mentalizing improves, it comes to concern core attachment relationships, including how these are activated with the clinician and key figures in the patient's life and how they influence mentalizing itself. Gradually, improvements in mentalizing serve to enable the patient to address their distorted representations of personal and social relationships.

Therapeutic Alliance and Engagement in the Model

The assessment and introductory process in MBT facilitates the alliance between patient and clinician (see Table 33.18–4) and introduces the patient to the treatment frame. An MBT-Introductory group of 10 to 12 sessions assists in the development of the formulation and facilitates the alliance. This psychoeducational intervention covers all areas of mentalizing, attachment processes, personality disorder, emotion management, and treatment itself. This preparatory work aims to ensure patients know what they are facing in trying to address their problems and are fully aware of the method and focus of treatment.

Table 33.18–4.
Alliance Building in MBT

Identification of patient's mentalizing vulnerabilities in an understandable form
Formulation of problems—agreed between patient and clinician
Identification of patient's risk profile and crisis management strategies
Agreement of short-term and long-term goals

MBT is collaborative. Nothing can occur without joint discussion, taking into account the mental experiences and ideas of both patient and clinician. The process of mentalizing requires an authentic desire to understand the mental processes of oneself and others. This applies as much to the clinician as to the patient. So the MBT clinician focuses on the patient's mind and attempts to understand their experience. Similarly, the patient is asked to aim to do the same in relation to the clinician—for example, the patient's perspective "Why does my clinician want me to focus on this at the moment?" may be paired with the clinician's, "Why does my patient *not* want to focus on this at the moment?" The therapeutic process has to become a shared endeavor aimed at extending the influence of explicit, reflective, cognitive, internally focused mentalizing. Initial goals, on the road to improved mentalizing, are jointly developed and focused on. The goals cannot solely be those of the patient, although the patient's aims take priority unless they are antithetical to the whole process of treatment. The sharing of responsibility for the therapeutic process is at the core of the effectiveness of the treatment approach in the pursuit of improved mental state understanding.

Assessment involves delineation of the patient's mentalizing vulnerabilities and mentalizing profile, identification of nonmentalizing cycles (see case example, later) and a shared formulation, which includes specific detail of attachment patterns and areas of vulnerability to emotional dysregulation. This has to be understood by the patient and is for both patient and clinician. The formulation identifies common relational fears, for example, abandonment, which stimulate the patient's

attachment system and result in the use of maladaptive attachment strategies in interpersonal interactions. In brief, the pattern of the patient's relationships informs an understanding of the relationship in treatment, and the relationship in treatment is used to reappraise the relationships in life outside treatment. Finally, it is important that the patient and clinician consider establishing a goal of improving social function. This will include work, social activity, voluntary work, education and other constructive life-affirming activity.

Clinical Principles in MBT

Clinicians follow a number of principles when treating patients with MBT (see Table 33.18–5). MBT recommends an authentic "not-knowing" stance that forms the bedrock for exploration of the patient's perspective. The not-knowing stance refers to respecting the opacity of mental states. Minds can never be "known" and it is important that the clinician recognizes that mental processes generate experiences imbued with uncertainty. The clinician's task is to take an inquisitive stance—a wish to inquire and a willingness to be surprised by the patient's response—with the aim of facilitating the patient's increased awareness of their internal states through a social process. This may be particularly important around the point at which experiences of ideas and feelings begin to collapse into nonmentalizing modes, leading to destructive behaviors or intolerable feeling states. It is not for the clinician to compensate for the patient's mentalizing failure with their own high-level mentalizing, "explaining" to the patient what they may have experienced. Nonmentalizing in the patient cannot be met by mentalizing in the clinician; it can be met only by "switching on" mentalizing in the patient via the range of techniques that MBT uses in the clinical situation.

Table 33.18–5.
Clinical Principles in MBT

The clinician must:

- ► Remain alert to imbalances of the mentalizing dimensions and emergence of nonmentalizing modes
- ► Monitor the patient's arousal levels to maintain optimal mentalizing
- ► Seek moments of mentalizing vulnerability related to events in the patient's life or in the session itself
- ► Maintain their own mentalizing
- ► *Not* meet nonmentalizing in the patient with high-level mentalizing
- ► Match interventions to the patient's mentalizing capacity

Primarily, the clinician is alert to nonmentalizing in terms of the different manifest modes of nonmentalizing, but also in terms of the indications that the patient's functioning is fixed at one pole of any of the dimensions of mentalizing. In general, as described earlier, mentalizing is

optimal when the dimensions—for example, emotion and cognition, or representation of self and other—are in balance and nonmentalizing modes are inactive. The key for the clinician is to be constantly aware of imbalance and lack of flexibility in terms of the dimensions and if any dimension is operating in a nonmentalizing mode. Nonmentalizing in a dimension or mode is an indication that intervention is necessary. An exclusive focus on feelings and ideas in relation to the self should suggest to the clinician that exploring the mental states of the other is called for. Similarly, exclusive concern with emotions suggests the need to bring cognitions into the foreground through judicious inquiry.

Second, the clinician monitors arousal levels carefully, ensuring that anxiety is neither too low nor too high, as both interfere with mentalizing. Similarly, if attachment feelings (e.g., in relation to the therapist) become too strong, a shift to less charged relationships may help restore mentalizing.

Third, the focus of a session is maintained through the clinician always seeking moments of mentalizing vulnerability either in relation to events in the patient's life or in the session itself. Mentalizing failure is best indicated by the clinician's experience of being challenged to understand the patient and consequently feeling a degree of confusion about how to respond. At these times, rewinding to moments when shared understanding characterized the discourse is the best solution.

Fourth, the clinician makes sure that their own mentalizing is maintained. It is not possible to deliver effective treatment if the clinician's mentalizing is compromised. So the MBT clinician always monitors their own capacities and may even have to explicitly own this experience, for example, by saying that their mind has become muddled and they cannot think. This type of self-disclosure of the mind state of the clinician should not be confused with sharing personal information. Sharing the effect that a patient's actions and state of mind is having on the clinician acts in the service of asking the patient to consider another mind as well as their own and has the implicit aim of enhancing mentalizing.

Sessions are focused. They do not consist of free associative dialogue that seeks to illuminate unconscious process. The target area is working memory or preconsciously held experience. It is expected that a focus for a given session will have been achieved after 10 to 15 minutes of the session, and this focus will then become the pivotal point around which the clinician and patient orient themselves, returning to it whenever nonmentalizing comes to dominate the interaction.

Finally, interventions are carefully matched to the mentalizing capacities of the patient. It is no good offering complex interventions that require considerable thought and appraisal to an individual functioning in psychic equivalence mode. At best, this serves to take over their mentalizing for them, rather than facilitating its rekindling. As stated

earlier, nonmentalizing in the patient cannot be recovered by mentalizing in the clinician, but only by reactivating mentalizing in the patient. This is achieved through a series of steps, which underpin the trajectory of every MBT session and may recur several times within each session.

Trajectory of MBT Session and Interventions

Empathic Validation. The initial step in a session is listening to the patient's narrative. Listening to the patient's story allows the clinician to begin by empathic validation. Empathic validation and establishing a shared affective platform held between patient and clinician increases the patient's experience that they are not alone and indicates that another mind can be useful to clarify mental states and increase a sense of agency. Increasing focus on affect and interpersonal interaction during a session and over time provides the context in which to explore ever more complex states of mind within an attachment context that would normally trigger loss of mentalizing.

Empathic validation requires the clinician to find something in the story that they can empathize with. This is not the same as behaving in a sympathetic manner or saying things that repeat the patient's story. Empathic validation seeks to engender in the patient a sense that the clinician really understands the patient, their internal state, and the issue they are talking about. This is the clinical equivalent of attachment-based contingent responsiveness. Validation is an affectively based intervention with important cognitive components; the key component is the creation of a sense of alignment with the patient's internal emotional state by demonstrating an appreciation of the experience and the consequent secondary emotions triggered by a powerful emotional reaction (e.g., the patient's fury with a partner engenders deep distress and anxiety). A lack of appreciation of the patient's emotional experience and its impact on the patient's current state (noncontingent responsiveness on the part of the clinician) is likely to trigger nonmentalizing (e.g., pseudomentalizing) or generate avoidant or other insecure and disorganized attachment strategies. Once contingent responding has increased collaboration and reduced the patient's arousal, maintaining emotions at a manageable level, the clinician can consider sensitive but less contingent responses to try to stimulate mentalizing about the story the patient brings.

Clarification and Exploration. The second step is clarification and exploration (see case example, later). The "story" the patient brings is clarified. This is not clarification of facts about the narrative or events, although this must also take place. The clinician establishes the facts as quickly as possible. For example, if the patient speaks about an act of self-harm or a suicide attempt, a drunken brawl or an emotional outburst, the clinician quickly elicits when it occurred, who was there, what were the

circumstances, and so on. This will indicate the level of risk and provide other important information. More than this, though, the MBT clinician wishes to contextualize events with mentalizing. Clarification establishes the reflection the patient has on the events—what was their "premorbid" state of mind, what were their hopes, what was their experience when they were waiting for their boyfriend to return home, what thoughts intruded into their mind, what feelings did they identify, and can they reflect on it differently now? This process of clarification, in the service of engaging mentalizing, links inextricably with affect identification and exploration, which is the third step.

Affects and Affect Identification. Affects and interpersonal relationships reciprocally interact and are core to the personality problems characteristic of BPD. Unmanageable emotions impinge on relationships, and relationships stimulate powerful feelings. Patients may not be able to identify their feelings accurately but experience them primarily as inchoate bodily experiences. Working with the patient to identify a range of feelings is part of the clarification and exploration component of MBT (see case example, later). Sometimes, emotions in specific contexts have to be normalized. Too often, patients feel that their experience is wrong; in effect, they invalidate their own internal perceptions and feel ashamed. It may be that their feeling is appropriate but excessive, or at other times inexplicably absent. All this becomes clear if the clinician systematically focuses on affects. Initially, the affects associated with the events are established, then any reflection on those feelings are clarified, followed by eliciting current concerns about the events.

Clarification of current affect in the session is included as the third step if the patient and clinician retain their capacity to jointly mentalize around the focus. This is more than asking the patient how they feel at the moment or how they "feel about" an experience they reported, although this may be an initial component. The process requires the identification of *current* affect *related to talking about events to the clinician in the session*.

This expansion from identifying affects in relation to events to affect experienced while talking to the clinician about the events is named the *affect focus* of the session. The patient may initially bring intense rage into the session in relation to an experience of rejection, which is gradually clarified as having been caused by characteristic inappropriate behaviors on the part of the patient. The affect focus turns out to be a sense of humiliation the patient experiences in once again failing to manage their own actions in a more effective manner. The aim throughout this process is to build a robust platform of mentalized experience jointly recognized and shared by the patient and clinician, which incorporates the complexity of mental states and moves the patient beyond a nonmentalizing, narrow oversimplification of their uncontextualized feelings.

The affect focus, that is, identification of the interpersonal interaction in the session and the associated affect, if accurate, heightens the focus on the clinician–patient interaction in the moment of the session. This often indicates that a patient's attachment strategies and relational patterns, or possibly those of the clinician, are being activated. So it allows a move toward *mentalizing the relationship,* the final step of the intervention trajectory.

Presenting Alternative Perspectives. Before extending to incorporate the relationship with the clinician into the patient's narrative, the clinician endeavors to enhance the patient's mentalizing by broadening the patient's perspectives on an event. This follows naturally from clarification and the affect focus, and entails elaborating the mental states of the participants in an event. However, it gradually moves the focus from emotion to cognition. The clinician explores alternative ways of looking an event, at first playfully but sometimes in a mildly challenging manner. If the affect focus helped reestablish more balanced mentalizing, then considering additional possibilities of what may have happened, at the level of thoughts and feelings, helps further reengage the patient's mentalizing processes.

Alternative perspectives may entail reconsiderations of entire scenarios from the point of view of the patient's hypotheses about the thoughts and feelings of others, or a reevaluation of the putative sequence of the patient's reactions—reframing their version of events in the light of alternative thoughts and feelings which they may have also been experiencing. The aim of alternative perspectives is the gentle expansion of the patient's mentalizing, to move from the certainty of affect toward the doubt of cognition. In general, the move is also from situational (external determinants) to exploring putative internal states. If the patient's focus was exclusively on their own state of mind the alternative perspective may bring in the other, while if the focus was on all protagonists other than the patient, the alternative perspective can focus on the self. Overall, the clinician engenders a recognition that a pause for reflection and explicit mentalizing may be of value in addition to intuition and the certainty automatic mentalizing can bring.

Mentalizing the Relationship. The groundwork for mentalizing the relationship will have been done through the development of the formulation, which identifies the predominant attachment strategies of the patient.

Mentalizing the relationship in a session is conceptualized as a training ground for managing difficult feelings in interpersonal situations in daily life through maintaining mentalizing while within an emotional interaction. The authors have identified a number of steps for the clinician to consider, which follow closely the three steps of technique described

above. First, the clinician has to empathically validate the patient's perception of the clinician. If the patient says that they experience the clinician in a particular way, then the clinician needs to find part of that experience that they can validate. The clinician actively avoids invalidating the patient's experience. Second, the clinician needs to work out their contribution to the patient's experience of the clinician. The clinician does this explicitly by thinking aloud about it and asking the patient to explain how they have come to that conclusion. This questioning must be authentic and genuinely curious, and must not come from a perspective that implies that the patient's experience is inaccurate or a reemergence of the past distorting the present. Such an invalidating attitude will lead to therapeutic rupture because invalidation (a noncontingent response) will lead to increased arousal and a consequent reduction in mentalizing. Mentalizing the relationship can meaningfully take place only in the context of mentalizing. Once the clinician has accepted their role in the relational process, the next step of more detailed exploration can occur. In this step, the aim is to generate a more complex understanding of the relationship, to see it from a different angle, and to see what its relevance is for the patient's life. It is not to engender insight in the sense of understanding the operation of the past in the present, although this may arise as part of the patient's broader understanding of their emotional experience. Critically, the past is used only sparingly as an explanation of the present, in case considering the past leads to a nonmentalizing, reductionist "short-cut" that obscures rather than elaborates the patient's current experience. In MBT, the clinician is cautioned about offering historical interpretation in the context of nonmentalizing process. For the exploration of the therapeutic relationship to contribute to enhanced mentalizing, the clinician works toward increased complexity and the establishment of multiple mental models of relationships.

Mentalizing the *counter-relationship,* or the feeling in the clinician, is the counterweight to mentalizing the relationship. The feelings and mind state of the clinician are given considerable weight in MBT—not as representing the patient's projected feeling, but as a meaningful aspect of an interactive relationship, to be used to demonstrate how minds affect minds. MBT clinicians monitor their experience of the patient. Not knowing what to say to the patient may be the best indicator of the patient's inadequate mentalizing. The dominance of the pretend mode in the patient may be indicated by a sense of boredom, while teleological mode may be indicated by a sense of confusion and anxiety. The interaction becomes the subject of useful concern, removing obstacles to mentalizing and enhancing accurate mentalizing of the relationship. For example, if the clinician is frightened of a patient with ASPD, this is an important feeling in the clinician that interferes with treatment and dictates the form of the therapeutic relationship. The clinician finds a way of expressing their

experience of the patient in a way that makes it palatable and recognizable as something worth exploring. The authors recommend that this is done through a number of steps. First, the clinician works out exactly what their feeling is and what it relates to in the patient–clinician interaction. Second, the clinician considers the patient's likely response to an explicit statement of this current feeling state, and states this before talking about the current feeling. Third, in the dialogue, the clinician identifies the experience as their own, marks it, and, finally, monitors the patient's reaction to the statement:

> "It is possible that what I am going to say may make you feel I am telling you off or being critical but I assure you that is not the case [anticipating the response of the patient].
>
> The problem is that when you sit forward like that and raise your voice I start to feel anxious and under threat [identifying the behavioral evidence and focus of external mentalizing, presenting his own affect and the effect it has on him].
>
> I realize that this may be me [marking the feeling] but it makes it difficult for me to concentrate on what you are talking about [additional effect on him, interfering with the relationship]."

From here the patient's reaction can be taken into account and the session can continue. But if the threatening attitude and angry presentation is something that permeates all the patient's relationships, then further exploration is essential.

In conclusion, the key to MBT is to:

1. Develop a focused narrative around the problems of the patient, especially interpersonal issues
2. Infuse the narrative with mentalizing process and prevent a collapse into nonmentalizing experience
3. Work with the patient to reinstate mentalizing when it is lost, to prevent destructive behavior and personal distress. Mentalizing oneself with others is the basis of satisfactory social and personal relationships, which so often is the goal of people with psychiatric problems and yet seems to them so unreachable.

CLINICAL ISSUES

Indications

MBT is effective in treatment for severe BPD. Patients in the early studies of the intervention (see Research and Evaluation, later) had made serious suicide attempts, been admitted to psychiatric hospital for risk, and/or had self-harmed. Both men and women were included in trials and patients showed high levels of comorbidity. Analysis of the data suggested that patients who showed comorbidity for a number of personality disorders, including ASPD, did preferentially better in MBT than comparison treatment. At a clinical level, patients with marked interpersonal problems who have a personality disorder rooted in mentalizing vulnerability and

attachment problems may benefit from MBT.

Limitations

Patients treated with MBT show a reduction in life-threatening behaviors and distressing psychiatric symptoms at the end of treatment, require less mental health care, and demonstrate improved social and interpersonal functioning. Nevertheless, long-term follow-up shows that patients continue to underfunction in their personal lives. Follow-up over up to 8 years shows that individuals remain with lower levels of social and relational satisfaction than expected, although the benefit of the therapy remains possible to detect.

It should be noted that the evidence for mentalizing mediating therapeutic change in MBT is currently limited, and more evidence is available in relation to mentalizing mediating change in other treatment modalities.

Complications

Psychotherapy can be harmful and MBT is probably no exception. However, MBT pays particular attention to ways in which patients may be harmed by treatment. On the basis that patients with BPD are uniquely sensitive to attachment process, and the stimulation of attachment reduces mentalizing, the MBT clinician focuses on levels of arousal in treatment sessions, constantly trying to balance arousal and mentalizing, ensuring that the therapeutic relationship is not a source of excessive attachment stress. Overstimulation of patients with avoidant attachment patterns is likely to trigger retreat and drop-out. Drop-out rates in MBT in clinical services in the United Kingdom are around 15 percent of people offered treatment, although in Scandinavia drop-out is as low as 2 percent. This suggests that the treatment overall is acceptable to people with BPD.

The aim of treatment is to increase the robustness of the patient's mentalizing capacity, and yet clinicians in many psychotherapeutic modalities may often tell patients "how they feel" or "what they are really saying." This undermines the mentalizing of the patient and so is avoided in MBT.

Contraindications

MBT is a generic treatment constructed to optimize access both by creating a low-demand treatment protocol and by facilitating access to training by a range of professionals. Individuals who have problems with mentalizing rooted in nonattachment contexts—for example, those with ASPD or psychosis—may not benefit from MBT; if they do, the mechanism of change should be assumed to be different from that in BPD. Individuals with relatively simple problems, such as phobias or uncomplicated depression,

may do better with more direct approaches such as cognitive-behavioral therapy. Even within a population of patients with BPD, patients with more complex presentations (multiple personality disorder diagnoses) are more likely to require MBT than those with a single BPD diagnosis, who may do as well in structured clinical management.

A 24-year-old patient, Sharon, was referred following a number of suicide attempts. She self-harmed, engaged in frequent polydrug misuse, and described emotionally volatile and occasionally violent relationships with men. She had a 2-year-old daughter who had recently been removed by child protection services. The clinician was able to take a detailed history, which elicited a number of vulnerability factors. Sharon was taken into care at the age of 5 and had a number of foster placements. She experienced recurrent sexual abuse from a carer around the age of 8. Her behavior was described in social reports as "over-sexualized" around the time of puberty and she attended an adolescent unit from the age of 13 to 15 years. She left school at 16 years having not attended for the previous year. The formulation was developed, which included these vulnerability factors but also identified the interpersonal vulnerabilities that led to mentalizing failures. The clinician explored two relational events with her boyfriend in which she had been violent, and two contexts in which she had tried to kill herself. Unsurprisingly, these events were interlinked

FORMULATION AND NONMENTALIZING INTERACTIONAL PATTERNS

In the formulation in MBT, mentalizing vulnerability points are identified and placed in the context of attachment strategies. In this patient there was evidence that she had an insecure attachment with marked ambivalence. She would seek something from her boyfriend, but when he did not respond she became desperate and increasingly demanding and clinging, eventually attacking him. She stated that if she could not "have him" she would "prefer to die," or alternatively she would "trap him" forever by getting pregnant. This nonmentalizing interactional cycle of need, demand, rejection, coercion, and the associated nonmentalizing modes were jointly explored, written down and shared with the patient. It is anticipated that the pattern will occur in relation to the clinician in a less intense way—the patient might seek some reassurance, for example, and the clinician may fail to respond contingently, triggering feelings of rejection. Identification and reduction of the nonmentalizing interactional process was an initial goal of treatment because the interactions occurred frequently and were an important area of vulnerability leading to suicide attempts.

CLARIFICATION AND ELABORATION AND AFFECT IDENTIFICATION

Sharon reported that she had been in a fight with her boyfriend. She had telephoned her boyfriend to find out when he would be back from work because she was looking forward to seeing him and to tell him that she loved him. He said that he was leaving soon and would be back within an hour. A few minutes later, a friend telephoned and, in their conversation, said that she had seen Sharon's boyfriend in a bar with a blonde woman only 30 minutes ago. When the boyfriend arrived home, Sharon asked why he had said that he was at work when she knew he was in a bar with a blonde woman. He said that he had called in for a drink on the way home. They had an argument during which Sharon attacked him and then, after smashing some crockery, cut herself.

Clarification of mental states about this event suggested that Sharon initially felt that she was looking forward to seeing her boyfriend, which was why she phoned him. His response was "contingent" with her feeling for him—he said he would be home soon. When her friend reported having seen him in a bar, this feeling was replaced with doubt about his love for her and a sense of rejection. Rapidly, she started to have thoughts that he was having an affair with the blonde woman, and in the context of feeling a loss of being loved this was experienced in psychic equivalence—"I am unlovable"—with some elements of hypermentalizing about his motives—"he was having an affair and I am certain of it." This was the nonmentalizing cycle and associated modes identified in her formulation.

When Sharon's boyfriend arrived home, there was no doubt in her mind about what had happened. He was unable to persuade her it was not true. In nonmentalizing states, actions are

the only meaningful way of communicating and she was therefore left with being coercive to "make" him love her. But to do this she tried to lock him in the house so he could not go out and she demanded sex. He resented this, refused, and the violence occurred. At this point her demand was driven by teleological belief that only his actions could prove his love for her.

The task in MBT is now to identify how Sharon manages her initial excitement about seeing her boyfriend, the sudden collapse of these feelings, and her collapse into psychic equivalent thoughts about his infidelity. To do this the clinician engages in detailed delineation of her mental states while asking her to re-present them to herself to instill a sense of uncertainty: can she manage emotional turmoil between excitement and disappointment; can she question her certainty over her boyfriend's activities and motives; can she engage with him to establish a more robust way of managing the distrust in their relationship other than teleological demand?

AFFECT SESSIONAL FOCUS AND MENTALIZING THE RELATIONSHIP

The clinician asked Sharon how she felt about her boyfriend now and she said that she still felt that he did not care for her. She was miserable. But, in addition, she felt that she had created problems, which was "typical of me." This is identification of current affect in relation to the focus. But, as she was talking to the clinician, she said that she felt that the clinician was judging her and would give up on treating her as it had happened again. She was a failure. This is the identification of current affect in the session and was quickly explored. At this point Sharon became coercive to some degree, suggesting that the clinician could not stop seeing her and to prevent it she would take an overdose. It becomes possible at this point to mentalize the relationship. Sharon thinks that the clinician will not meet her need and so automatically engages in a nonmentalizing interactional process. This needs to be discussed in the context of the clinician stating that it is not in his mind to stop seeing her in treatment (statement of clinician counter-relationship), so it is important to understand what is fuelling her belief and to question it. It is not for the clinician to interpret the repetition in the session of the interaction with the boyfriend but to define the emotional need that Sharon reacts to that makes her become coercive.

Goals of Treatment

The primary aims of MBT are to rekindle mentalizing when it is lost, to maintain it when it is present, and to increase the resilience of the individual's capacity to keep it going when it would otherwise be lost. But mentalizing is about something. So, the development and maintenance of mentalizing is initially focused on the core symptoms of BPD (or another condition), such as suicide attempts, self-harm, and other self-destructive behaviors and emotional instability. This is followed by emphasis on interpersonal problems because the key area of vulnerability to losing mentalizing in BPD is the interpersonal domain, especially when attachment processes are activated. So the clinician–patient relationship is a significant area of scrutiny. Patient and clinician increase attentiveness to mental states and interpersonal contexts in which they become disordered.

ETHICAL ISSUES

A key feature of MBT is the collaborative stance, which engages patients in a mutually agreed protocol, each phase of which focuses on achieving shared understanding between the patient and the clinician. In other words, it may be argued that MBT clinicians are less likely to encounter ethical issues than proponents of other therapeutic approaches in which collaboration is perhaps seen as a necessary condition but not the aim of

the treatment.

However, MBT clinicians should be well aware that the nature of the clinical problems they are dealing with invariably and unavoidably creates ethical issues daily. Working with a person whose capacity to represent themselves in an agentive way is lacking, whose attachment system is disorganized, and who is extremely vulnerable to creating a dependent relationship, will present a consistent problem for the clinician to avoid exerting undue influence. MBT clinicians—like therapists working in other orientations—can prolong relationships beyond their useful timespan, citing the patient's need for their input as justification. In particular, when patients pay for their treatment either themselves or through a third party, financial exploitation is a real and present risk. Beyond this, undue influence may manifest through the uncritical acceptance of the clinician's frame of reference: a nonmentalizing individual has no alternative and gratefully grabs hold of a powerful, coherent model when it is presented to them. The teleological predisposition of patients sometimes leads clinicians to find themselves caught up in action-oriented attempts to address the patient's distress, offering more than they should, yet delivering less than they could.

RESEARCH AND EVALUATION

MBT for BPD

There have been several recent reviews of psychosocial interventions for BPD. These recognize the evidence base for MBT for BPD as generally inferior only to that for dialectical behavior therapy, although not from a health economic perspective. Encouragingly, a large qualitative study of patients' treatment goals established that the goals of MBT were closely allied with what patients hoped to gain from their therapies.

A small number of RCTs and a number of naturalistic studies have tested the effectiveness of the MBT approach for BPD patients.

In an RCT of MBT for BPD in a partial hospital setting, an 18-month program achieved significant and enduring changes in mood states and interpersonal functioning. Outcome measures included frequency of suicide attempts and acts of self-harm, number and duration of inpatient admissions, service utilization, and self-report measures of depression, anxiety, general symptom distress, interpersonal function, and social adjustment. The benefits, relative to treatment as usual (TAU), were large, with a number needed to treat of approximately two; in addition, the benefits were observed to increase during the follow-up period of 18 months. Analysis of participants' healthcare use suggested that day hospital treatment for BPD was no more expensive than general psychiatric care and showed considerable cost savings after treatment. A follow-up study of BPD patients 5 years after all treatment was completed and 8 years after

initial entry into treatment, comparing patients treated with MBT and those receiving TAU, found that those who received MBT remained better than the TAU group. Superior levels of improvement were shown on levels of suicidality (23 percent in the MBT group vs. 74 percent in the TAU group), diagnostic status (13 percent vs. 87 percent), service use (2 years vs. 3.5 years), and other measurements such as use of medication, global functioning, and vocational status.

Two well-controlled single-blind trials of outpatient MBT have been conducted, one with adults with BPD and the second with adolescents presenting to clinical services with self-harm, the vast majority of whom met BPD criteria. In both trials, MBT was found to be superior to TAU in reducing self-harm, including suicidality, and depression. Importantly, in the adult trial, the control group received a manualized, highly efficacious treatment, structured clinical management; MBT was superior to this intervention, particularly in the long term. A posthoc analysis of moderators found that the number of personality disorder diagnoses in addition to BPD was the key indicator of severity that predicted the need for the MBT approach, as structured clinical management appeared to have little benefit on most outcome measures among these patients. Furthermore, in the trial with an adolescent sample, improvements generated by MBT appear to have been mediated by improved levels of mentalizing, reduced attachment avoidance, and amelioration of participants' emergent BPD features; participants treated with MBT showed a recovery rate of 44 percent, compared with 17 percent of those who received TAU. Ongoing follow-ups of both these trials indicate that improvements in the MBT groups have been at least maintained, and in most cases improvements continued after treatment termination, and differences relative to the comparison group remain significant.

Three recent studies provide further support for the efficacy of MBT in BPD. An RCT from Denmark investigated the efficacy of MBT versus a less intensive, manualized supportive group therapy in patients diagnosed with BPD. Patients were randomly allocated to MBT (n = 58) or the manualized supportive therapy (n = 27). Each intervention was delivered in combination with psychoeducation and medication. Both the combined MBT treatment and the less intensive supportive therapy brought about significant improvements on a range of psychological and interpersonal measures (e.g., general functioning, depression, and social functioning) and decreased the number of diagnostic criteria met for BPD; effect sizes were large (d = 0.5 to 2.1). The combined MBT was superior to the less intensive supportive group therapy on clinician-rated Global Assessment of Functioning. An 18-month naturalistic follow-up found that treatment effects at termination were sustained at 18 months. Half of the patients in the MBT group met criteria for functional remission at follow-up, compared with less than one-fifth in the supportive therapy group, but

three-quarters of both groups achieved diagnostic remission, and almost half of the patients had attained symptomatic remission. A limitation of this study is that the same clinicians delivered both interventions (and thus there was a high risk of spillover effects between the two treatments); incomplete data was a further significant limitation. In a second study from Denmark, a cohort of patients treated with partial hospitalization followed by group MBT showed significant improvements after treatment (average length 2 years) on a range of measures including Global Assessment of Functioning, hospitalizations, and vocational status, with further improvement at 2-year follow-up.

A quality improvement study examined the outcomes for BPD patients treated in an MBT program in a Norwegian specialist treatment unit compared with a former psychodynamic treatment program. This longitudinal comparison had a sample of 345 BPD patients, including 282 patients treated on the psychodynamic program and 64 who received MBT, who had comparable baseline severity and impairments of functioning on all measures. Outcome measures included Symptom Checklist-90 symptom distress, interpersonal problems, and global functioning assessed routinely throughout treatment, and suicidal/self-harming acts, hospital admissions, medication, and occupational status assessed at baseline and discharge. The change in program from traditional psychodynamic therapy to MBT led to a reduction in unplanned discharges (MBT had a low drop-out rate of 2 percent). Measured benefits from the change of program included greater improvements in symptom distress and interpersonal, global, and occupational functioning. Although the change was associated with the introduction of MBT, specific causal attributions are hard to establish in such a design.

A naturalistic study in the Netherlands investigated the effectiveness of an 18-month manualized program of MBT in 45 patients diagnosed with severe BPD. There was a high prevalence of comorbidity of DSM-IV Axis I and Axis II disorders. Results showed significant positive change in symptom distress, social and interpersonal functioning, and personality pathology and functioning; effect sizes were moderate to large (d = 0.7 to 1.7). The study also showed that the use of additional treatments and psychiatric inpatient admissions decreased significantly during and after treatment. The lack of a control group in this study limits the ability to draw conclusions about the efficacy of MBT. Another study applied propensity score matching to determine the best matches for 29 MBT patients from within a larger (n = 175) group who received other specialized psychotherapeutic treatments. These other specialized treatments yielded improvement across domains, which was generally only moderate; in contrast, pre–post-effect sizes were consistently large for MBT, with Cohen's d for reduction in psychiatric symptoms of −1.06 and −1.42 at 18 and 36 months, respectively, and ds ranging from 0.81 to 2.08 for

improvement in domains of personality functioning. Given the nonrandomized study design and the variation in treatment dose received by participants, the between-condition difference in effects should be interpreted cautiously. A multisite randomized trial by the same group comparing intensive outpatient and partial hospitalization-based MBT for patients with BPD is currently underway.

A recent naturalistic pilot trial studied the feasibility and effectiveness of an inpatient adaptation of MBT in 11 female adolescents (14 to 18 years) with borderline symptoms. One year after the start of treatment, significant decreases in symptoms and improvements in personality functioning and quality of life were observed; effect sizes were between d = 0.58 and 1.46, representing medium to large effects. Further, 91 percent (n = 10) of the adolescents showed reliable change on the Brief Symptom Inventory and 18 percent (n = 2) moved to the functional range on this measure. A report of the application of MBT principles to a therapeutic community also yielded positive results. Patients who completed 18 months of treatment showed significant self- and clinician-rated symptomatic improvement and significant change on clinician-administered measures of social and occupational functioning.

MBT for ASPD

Research into treatment for ASPD up to 2009 is summarized in the United Kingdom's National Institute for Health and Clinical Excellence clinical guideline for ASPD, which confirmed that interventions for ASPD are poorly researched and that evidence on its treatment is scarce. The authors of the guideline concluded that the evidence for treatments for ASPD was extremely limited, and did not support the development of any guidance on treatment recommendations. Two Cochrane reviews concluded that there was no consistent evidence for any intervention for ASPD, and recommended that research to test interventions for the disorder is urgently needed.

A feasibility study of MBT for ASPD reports findings from a small sample (n = 9) receiving group and individual MBT. Preliminary results on the Overt Aggression Scale suggested that the participants rated the severity of their aggression toward others and themselves as decreasing over the first 6.5 months of treatment; in contrast, their rating of irritability did not change. Psychiatric symptom severity on the Brief Symptom Inventory showed a reduction in the distress participants experienced in relation to their symptoms at a 6-month follow-up, with participants reporting greatest decreases in distress resulting from symptoms of depression, anxiety, and hostility.

Finally, the authors note that a significant subsample of the participants in the outpatient treatment trial of MBT for BPD described earlier also met criteria for ASPD. A separate analysis of these individuals with comorbid

ASPD revealed that they benefited significantly from MBT.

REFERENCES

Allen JG. *Mentalizing in the Development and Treatment of Attachment Trauma*. London, UK: Karnac Books; 2013.

*Allen JG, Fonagy P, Bateman AW. *Mentalizing in Clinical Practice*. Washington, DC: American Psychiatric Publishing; 2008.

Asen E, Fonagy P. Mentalization-based therapeutic interventions for families. *J Fam Ther*. 2012;34:347–370.

Bateman AW. Treating borderline personality disorder in clinical practice. *Am J Psychiatry*. 2012;169:560–563.

Bateman A, Bolton R, Fonagy P. Antisocial personality disorder: a mentalizing framework. *FOCUS*. 2013;11:178–186.

Bateman A, Fonagy P. 8-year follow-up of patients treated for borderline personality disorder: mentalization-based treatment versus treatment as usual. *Am J Psychiatry*. 2008;165:631–638.

*Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry*. 2009;166:1355–1364.

Bateman AW, Fonagy P, eds. *Handbook of Mentalizing in Mental Health Practice*. Washington, DC: American Psychiatric Publishing; 2012.

*Bateman A, Fonagy P. *Mentalization-Based Treatment for Personality Disorders: a Practical Guide*. Oxford, UK: Oxford University Press; 2016.

Bateman AW, Gunderson J, Mulder R. Treatment of personality disorder. *Lancet*. 2015;385:735–743.

Bateman A, Krawitz R. *Borderline Personality Disorder: an Evidence-Based Guide for Generalist Mental Health Professionals*. Oxford, UK: Oxford University Press; 2013

Bateman A, O'Connell J, Lorenzini N, Gardner T, Fonagy P. A randomised controlled trial of mentalization-based treatment versus structured clinical management for patients with comorbid borderline personality disorder and antisocial personality disorder. *BMC Psychiatry*. 2016;16:304.

Bevington D, Fuggle P, Fonagy P. Applying attachment theory to effective practice with hard-to-reach youth: the AMBIT approach. *Attach Hum Dev*. 2015;17:157–174.

*Fonagy P, Luyten P. A multilevel perspective on the development of borderline personality disorder. In: Cicchetti D, ed. *Developmental Psychopathology,* 3rd ed. New York: John Wiley & Sons; 2016.

*Fonagy P, Luyten P, Allison E. Epistemic petrification and the restoration of epistemic trust: a new conceptualization of borderline personality disorder and its psychosocial treatment. *J Pers Disord*. 2015;29:575–609.

Gergely G. Ostensive communication and cultural learning: The natural pedagogy hypothesis. In: Metcalfe J, Terrace HS, eds. *Agency and Joint Attention*. Oxford, UK: Oxford University Press; 2013.

Ha C, Sharp C, Ensink K, Fonagy P, Cirino P. The measurement of reflective function in

adolescents with and without borderline traits. *J Adolesc*. 2013;36:1215–1223.

National Institute for Health and Clinical Excellence. *Antisocial personality disorder: treatment, management and prevention*. London: The British Psychological Society and the Royal College of Psychiatrists; 2010.

Robinson P, Hellier J, Barrett B, et al. The NOURISHED randomised controlled trial comparing mentalisation-based treatment for eating disorders (MBT-ED) with specialist supportive clinical management (SSCM-ED) for patients with eating disorders and symptoms of borderline personality disorder. *Trials*. 2016;17:549.

Rossouw TI, Fonagy P. Mentalization-based treatment for self-harm in adolescents: A randomized controlled trial. *J Am Acad Child Adolesc Psychiatry*. 2012;51:1304–1313.

Sharp C, Ha C, Carbone C, et al. Hypermentalizing in adolescent inpatients: Treatment effects and association with borderline traits. *J Pers Disord*. 2013;27:3–18.

Sharp C, Fonagy P. Practitioner Review: Borderline personality disorder in adolescence – recent conceptualization, intervention, and implications for clinical practice. *J Child Psychol Psychiatry*. 2015;56:1266–1288.

Sharp C, Venta A. Mentalizing problems in children and adolescents. In: Midgley N, Vrouva I, eds. *Minding the child: Mentalization-based interventions with children, young people and their families*. London, UK: Routledge; 2012.

34 Biological Therapies

▲ 34.1 General Principles of Psychopharmacology

OLUSOLA AJILORE, M.D., PH.D., AND MARK M. RASENICK, PH.D.

Psychopharmacology is the use of medications as a treatment of mental disorders. Previous editions of this textbook have chronicled the evolution of psychopharmacology as a medical discipline. The first edition of *Kaplan and Sadock's Comprehensive Textbook of Psychiatry,* published in 1967, appeared a decade after the appearance of modern antidepressants, antipsychotics, mood stabilizers, and anxiolytics. Of the nearly 1,700 pages in that first edition, a mere 25 pages were devoted to drugs as treatments of mental disorders. Chlorpromazine (Thorazine), approved by the U.S. Food and Drug Administration (FDA) 15 years earlier, was alone among the drugs acknowledged to be a significant therapeutic breakthrough. Other psychotropic agents, including the benzodiazepines (introduced in 1959), were characterized as playing a minor therapeutic role, serving as adjuncts to psychotherapy or as a last resort for patients failing to respond to therapy. There was no intimation that benzodiazepines would become the most widely prescribed drugs of any type within a few years. Lithium (Eskalith), which had been described as an effective treatment of mania in 1949, was mentioned in a brief paragraph. Indeed, in reading the first edition of *Kaplan and Sadock's Comprehensive Textbook of Psychiatry* and other textbooks of that time, it would have been difficult to predict how psychotropic drugs would soon transform the treatment of mental disorders. In a reversal of positions, psychoanalytic theory, which once served to define the practice of psychiatry, has been supplanted by psychopharmacology as the most widely used form of treatment of psychiatric disorders. However, there is concern by many in the field that psychopharmacology has stalled in the last few years. Since the last edition of *Kaplan and Sadock's Comprehensive Textbook of Psychiatry* published in 2009, there have been only 13 new FDA-approved agents for psychiatric indications that have come to the market, compared to over 20 for neurology. In addition, none of the new agents have involved mechanisms

with novel therapeutic targets. In 2014 alone, there were 20 new drugs for cancer and a single, non-novel agent for psychiatry. This crisis is due, in part, to the high cost of failure for new drug trials, particularly as it is difficult to separate out from placebo for many psychiatric drugs. This crisis is also due to declining government budgets for basic research. As researchers fail to find new targets, the ability of pharmaceutical companies to develop truly novel psychiatric drugs is in jeopardy. It is important to understand the present state of psychopharmacology in order to advance the current knowledge in the field. In this context, this chapter begins by describing basic principles of both pharmacology and psychopharmacology.

In order to engender complete understanding of the basic principles of psychopharmacology, it is first necessary to understand the principles of pharmacology—particularly those relating to receptors and cellular signaling. This chapter is organized into two sections: the first section will introduce those basic principles of pharmacology and the second section will build upon them to discuss general issues with psychotropic drugs.

RECEPTOR PHARMACOLOGY

The majority of drugs used for palliation of psychiatric symptoms either mimic, block, or modify the actions of neurotransmitters, and most do so by binding to receptors for those transmitters. While receptors can be located either on the plasma membrane or intracellularly, the neurotransmitter receptors relevant for the drugs to de discussed in this chapter are transmembrane molecules, with outer surfaces forming binding pockets for neurotransmitter.

The receptors discussed in this chapter are either ion channels (ionotropic) or metabotropic (G protein coupled) and many neurotransmitters (e.g., acetylcholine, glutamate, gamma aminobutyric acid [GABA], and serotonin) act as agonists for receptors in each class. This chapter will discuss the basic biology of many of these transmitters and receptors. It begins with a description of characterization and quantitation of different receptors.

Receptor Analysis

There are two basic mechanisms for studying receptors. One is to measure the binding of a ligand (often radioactively labeled) and the other is to measure the physiological consequences of agonist binding (and the effects of antagonist).

Binding Assays. In order to study receptors with radioligand binding, hormones, neurotransmitters, or their analogs have been labeled with a variety of radioactive compounds. Membranes prepared from cultured cells or brain tissue (or intact cells) can be assayed for their ability to bind these

radioactive drugs. The radioactive compounds can be either agonist or antagonist, and through the use of these compounds, a quantitation of both the number of receptors and the affinity of those receptors for the drug can be determined.

In order to quantify binding, ligands are added to membrane preparations at known concentrations and competed with nonradioactive ligands in order to determine the extent to which binding is specific. Membranes are trapped on filters, the filters are washed and radioactivity is determined by scintillation spectrometry. Typical results are illustrated in Figure 34.1–1. Specific binding can be transformed by Scatchard analysis (Fig. 34.1–2) to reveal the existence of multiple affinity states for agonist and Hill analysis to determine the whether binding is cooperative.

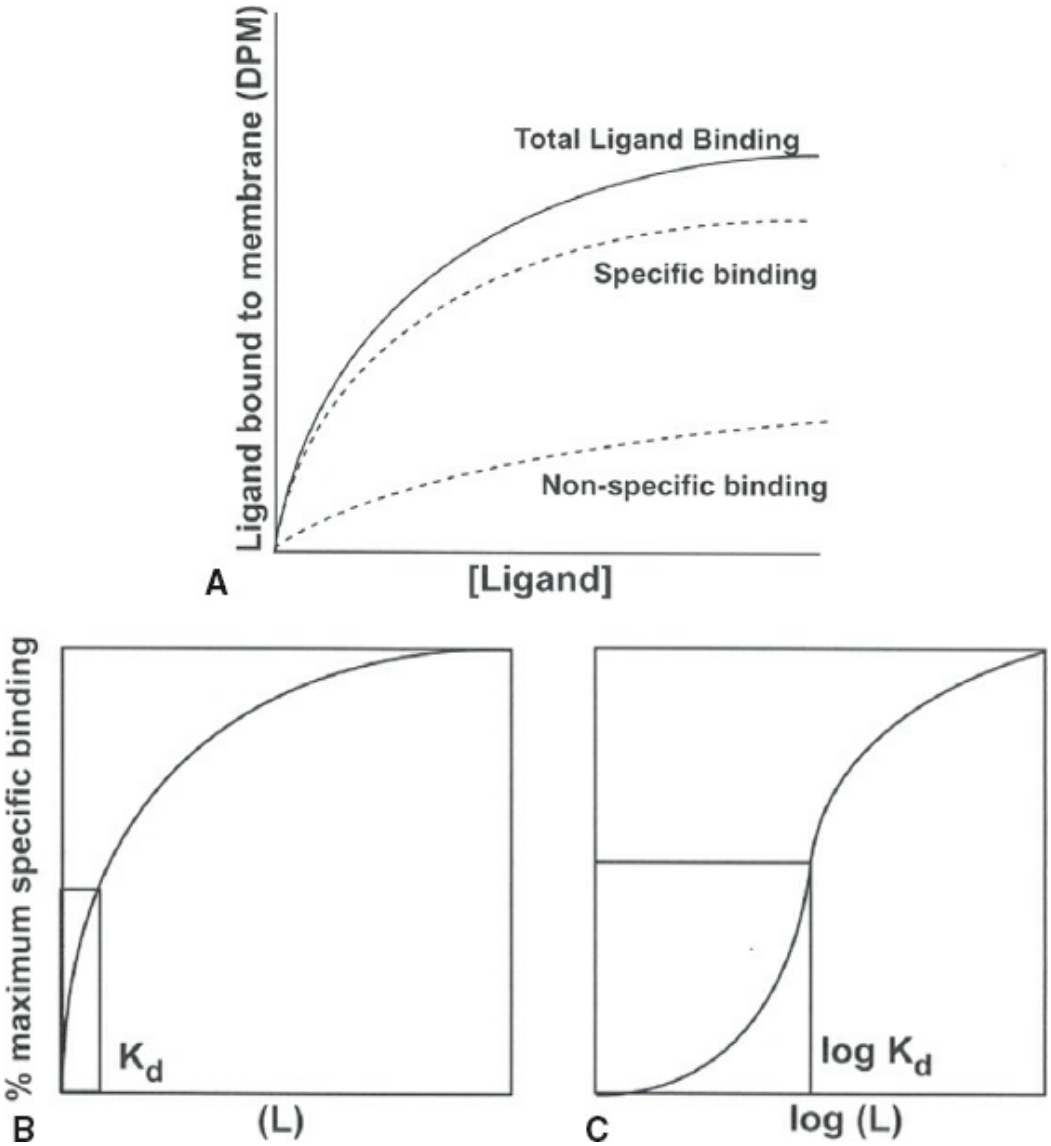

FIGURE 34.1–1. Kinetics of ligand (agonist or antagonist) binding. **A:** Measurement of receptors using radioligand binding. Membranes (or cells) are incubated with radioactive agonist or antagonist and collected by sedimentation or filtration. Radioactivity in particulate fractions is calculated and plotted as shown above. Nonspecific binding refers to that portion of the total binding that persists in the presence of an excess of unlabeled

compound. Specific binding is saturable, within the range of the experiment shown, but nonspecific binding does not saturate. **B:** Two plots of specific binding versus concentration of ligand. **C:** The approximate Kd is represented by the ligand concentration at 50 percent saturation of specific binding. DPM, disintegrations per minute. (From Felig P, Frohman LA. *Endocrinology and Metabolism*. McGraw-Hill Education; 2001.)

Some receptors are oligomeric. Many ionotropic receptors have multiple subunits and two or more of these bind agonist, often displaying cooperative binding. G protein–coupled receptors (GPCRs) have been shown to dimerize as well, and this can take the form of homodimers or heterodimers, in the case of multiple receptor species. Although dimerization is obligatory only for GABAB receptors, many other forms of receptors have been demonstrated to dimerize.

Biological (Efficacy) Assays. In addition to direct determination of agonist or antagonist binding to receptor, the biological consequences of this binding can be assessed. One of these methods involves a biological assay. A classical biological assay utilizes an isolated organ strip, such as the guinea pig ileum, and measures changes in both the frequency and strength of contraction of the intestinal smooth muscle in response to the presence of an agonist. Such a system allows the investigator to coordinate the dose of hormone with the strength of the response, but quantification is often imprecise.

More recent biological assays rely on enzyme "effectors" activated subsequent to receptor activation. In the case of the guinea pig ileum, this might involve measuring the cyclic adenosine monophosphate (cAMP) produced by the enzyme adenylyl cyclase subsequent to the activation of the β-adrenergic receptor. Higher throughput assays designed to screen drugs that may act as agonists or antagonists for a given receptor might involve fluorescent sensors detecting an increase in Ca^{2+} subsequent to activation of a receptor that activates phospholipase c (e.g., 5HT2A), a process that ultimately increases cellular calcium. Such assays are referred to by the acronym FLPR (fluorescence imaging plate reader) and are used routinely in the process of drug discovery and development. These high-throughput assays employ fluorescent calcium (or other small molecule) sensors to determine GPCR activation and determine the efficacy of an agonist or potency of an antagonist at a given receptor.

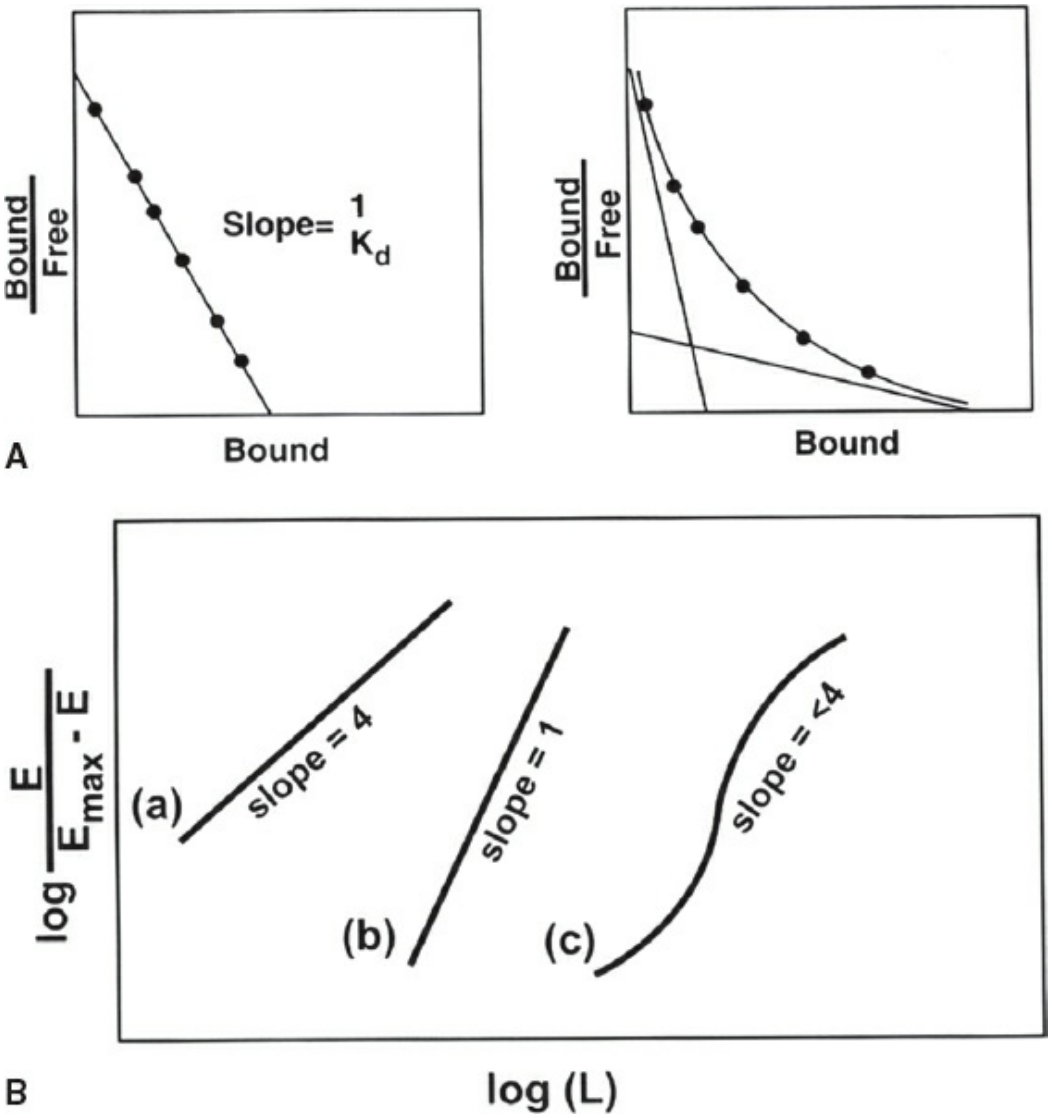

FIGURE 34.1–2. Computational analysis of receptor binding. **A:** Scatchard plots. (Left) a receptor with a single affinity state for the binding of ligand; (right) two-site fit. **B:** Hill plots for the determination of cooperativity in ligand–receptor binding. Three Hill plots are shown here. (**a**) Four binding sites (possibly a tetramer); (**b**) is representative of a single site; (**c**) cooperative binding of fewer than four but more than one site. (From Felig P, Frohman LA. *Endocrinology and Metabolism*. New York: McGraw-Hill Education; 2001.)

Efficacy could also be plotted in a manner similar to that used in Figure 34.1–1. The x-axis would continue to represent the amount of drug required (potency) and the degree of activation (efficacy) would be represented on the y-axis.

G Protein Modulation of Agonist Binding. GPCRs show a significant decrease in agonist affinity following the introduction of guanosine 5′-triphosphate (GTP) into the medium. This is demonstrated in Figure 34.1–3. Isoproterenol, a β-adrenergic agonist, displaces iodocyanopindolol (a radioactive antagonist) with less potency in the presence of GTP. It is suggested that the complex between receptor and G protein represents the

high-affinity state for agonist binding. Agonist allows the active receptor to induce the G protein to bind GTP (this will be explained more completely below), which then causes the G protein to dissociate from the receptor. This "free receptor" has a lower affinity for agonist than does the receptor–G protein complex. Note that some antagonists also show the ability to undergo this "GTP shift." As will be noted below, some antagonists also modify the conformation of certain other GPCRs.

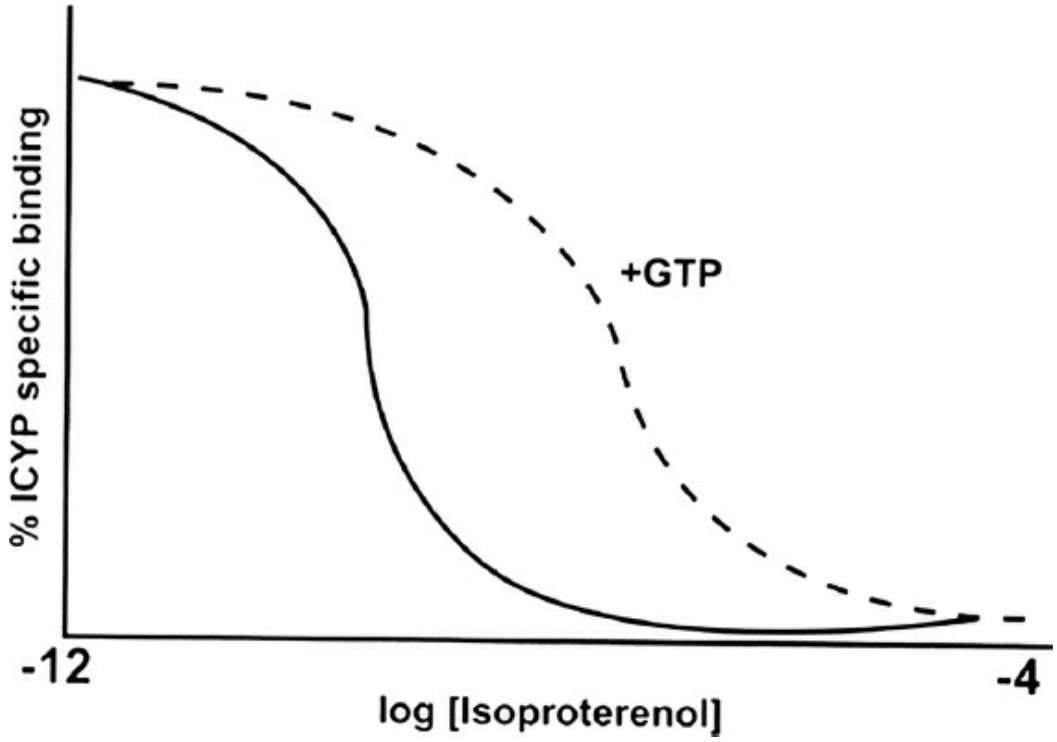

FIGURE 34.1–3. Guanosine 5'-triphosphate (GTP) decreases agonist affinity. In the example shown, the β-adrenergic agonist, isoproterenol, is used to displace the labeled β-adrenergic antagonist, 125I-iodocyanopindolol (ICYP). In the presence of GTP significantly more isoproterenol is required to displace the iodocyanopindolol. This "right shift" of the binding curve indicates that, in the presence of GTP, affinity for agonist by the receptor is decreased. (From Felig P, Frohman LA. *Endocrinology and Metabolism*. New York: McGraw-Hill Education; 2001.)

Structure and Function of Receptors

Ionotropic receptors were first to be characterized structurally, as the nicotinic acetylcholine receptors (nAChRs) are present in large concentration in the electroplax organs of electric fish and eel, and proved rich sources of receptor that was modeled by x-ray diffraction. Many years later, outer segments of retinal rod photoreceptors proved a similarly rich source of rhodopsin, the first GPCR to be crystalized by Palczewski and colleagues. While solving the structures of GPCRs more relevant to psychiatry was daunting, Kobilka, Stevens and their colleagues devised a mechanism for modifying these receptors and a number of pharmacologically-relevant GPCRs have now been determined,

structurally.

In the case of ionotropic receptors, the binding of agonist elicits a cascade that initially creates a channel for ion conductance. Second messenger systems (elicited subsequent to GPCR activation) modify channel function considerably. In the case of GPCR, activation subsequent to agonist binding results in a change in receptor morphology that releases GDP from the bound G protein α subunit. The subsequent binding of GTP activates the G protein, allowing α subunits to engage intracellular effector enzymes such as adenylyl cyclase or phospholipase and β/γ subunits to engage voltage-dependent ion channels.

Ionotropic Receptors

VOLTAGE GATED. These receptors, which are intrinsic ion channels, are either voltage gated or ligand gated. The first class, consisting of Na^+, K^+, and Ca^{2+} channels, possesses sensors (arrays of charged amino acids within transmembrane regions). Many of these receptors have multiple (4 or 5) subunits that combine to form a functional receptor. Several mutations have been identified for voltage-sensitive channels and these result in a variety of maladies (channelopathies) ranging, in the case of calcium channels, from familial migraine to dystonia (Fig. 34.1–4).

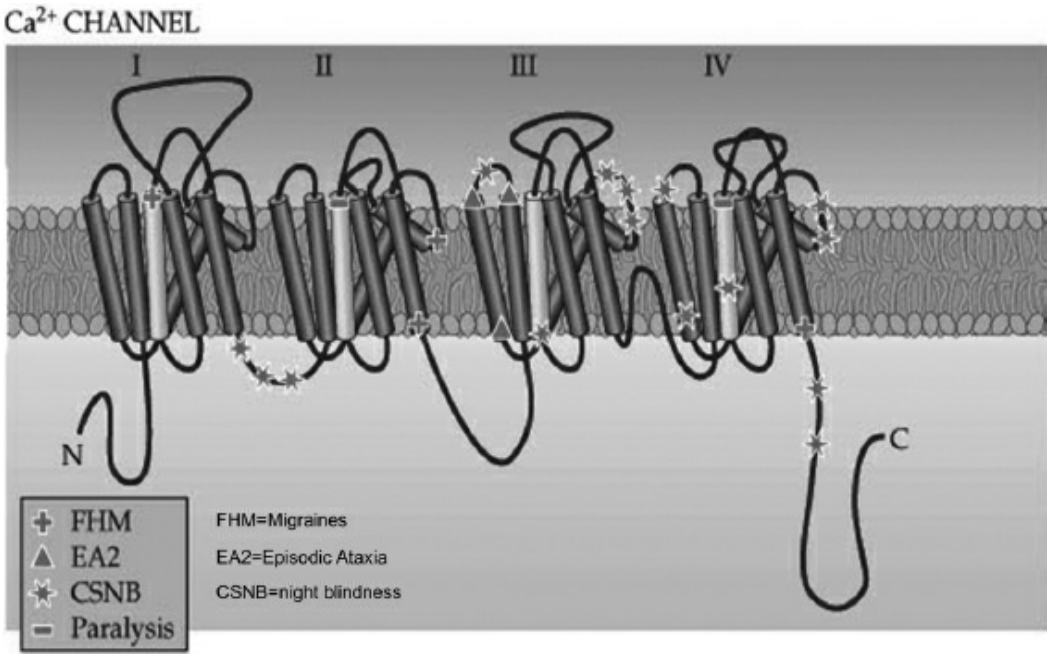

FIGURE 34.1–4. Mutations in Ca^{2+} channels and disease phenotypes (channelopathies). This model of a voltage-dependent Ca^{2+} channel shows clusters of mutations and the relevance of mutations in a particular region for a particular malady. FHM, Familial Hemiplegic Migrane; EA2, Episodic ataxia type 2; CSNB, X-linked Congenital Stationary Night Blindness. (Adapted from Purves D, Augustine G, Fitzpatrick D, Hall W, LaMantia A, White L. *Neuroscience*. 5th ed. Sunderland, MA: Sinauer; 2012.)

LIGAND GATED. These receptors, which also form ion channels that open in response to agonist, often share the quatrameric or pentameric structure of the voltage-gated receptors, but have specialized regions that bind one or two molecules of agonist. nAChRs, GABAA receptors, glutamate receptors (either kainate or *N*-methyl-D-aspartate [NMDA]), and 5HT3 receptors fall into this class (Fig. 34.1–5). While some of these receptors (e.g., nAChR) are regulated by binding of agonist, others have more complex regulation, GABAA have binding sites for benzodiazepines that increase agonist potency as well as binding sites within the channel region that decrease chloride conductance, hence acting as convulsants (Strychnine) (Fig. 34.1–6). Similarly, NMDA receptors, which normally act as Na^+ channels, conduct Ca^+ subsequent to membrane depolarization (Fig. 34.1–7) and can even participate in rapid remodeling of the synapse. Curiously, glycine, which is an agonist for a Cl-channel receptor similar to GABAA (but found primarily in spinal cord), also acts as to promote activity of glutamate at the excitatory NMDA receptor.

Receptors for Growth Factors. Receptors for neurotrophins, such as nerve growth factor (NGF) and brain-derived neurotrophic factor (BDNF), have intrinsic tyrosine kinase activity and dimerize upon agonist binding (Fig. 34.1–8). These receptors include trkA, trkB, and p75NTR. Regardless of the specific receptor type, most of the members of this receptor family consist of two subunits. Sometimes receptors form a homodimer and, at other times, a heterodimer. Formation of a dimer, which is a result of receptor activation, leads to autophosphorylation or transphosphorylation of the receptor within its cytoplasmic domain. Consequently, other substrates for that receptor phosphate are found and phosphorylate. Signal propagation often is a result of an activation of a series of kinase cascades resulting in cell growth or differentiation.

Metabotropic (G Protein–Coupled) Receptors. GPCRs are a large family of receptors for light, ordorants, small molecule neurotransmitters as well as a variety of peptide agonists. It is estimated that somewhere between 50 and 70 percent of current drugs target GPCRs, elevating their importance for psychopharmacology. GPCRs regulate a multitude of physiological functions through activation of heterotrimeric G protein α and βγ subunits that directly bind to target proteins to initiate cellular signal transduction cascades, a process that is critically important for neurons and neurotransmitter action. Nearly 800 GPCRs have been identified in the human genome and include both chemosensory receptors and receptors for hormones and neurotransmitters. Approximately 100 of these GPCRs have not been linked to an endogenous ligand. These are known as orphan receptors. All GPCRs share a similar structure, which includes seven transmembrane domains, an extracellular amino terminal, an intracellular

carboxyl terminal, three extracellular loops, and three cytosolic loops (Fig. 34.1–9). While each of the seven transmembrane domains comprises somewhere between 20 and 27 amino acids, the end-terminal segments as well as either the extracellular or cytoplasmic loops can vary greatly among the different receptor types.

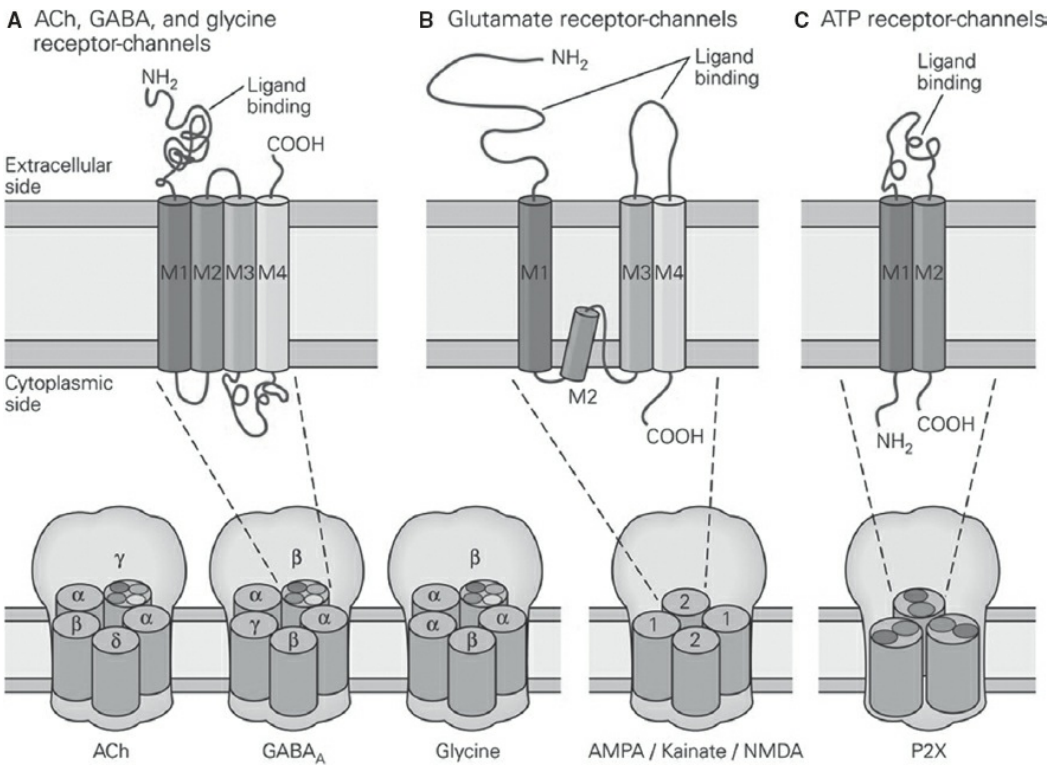

FIGURE 34.1–5. Families of ligand-gated ion channels. **A:** Nicotinic AChR, GABA, and glycine receptor channels are all pentamers and are comprised of many different types of subunits. There actions are explained in the text. **B:** Glutamate receptors are tetramers, usually comprised of two each of similar subunits. **C:** Purinergic (ATP) receptors are trimmers, each subunits with 2 membrane-spanning domains. (Adapted from Kandel ER, Schwartz JH, Jessell TM, Siegelbaum SA, Hudspeth AJ, eds. *Principles of Neural Science*. 5th ed. New York: McGraw Hill; 2013.)

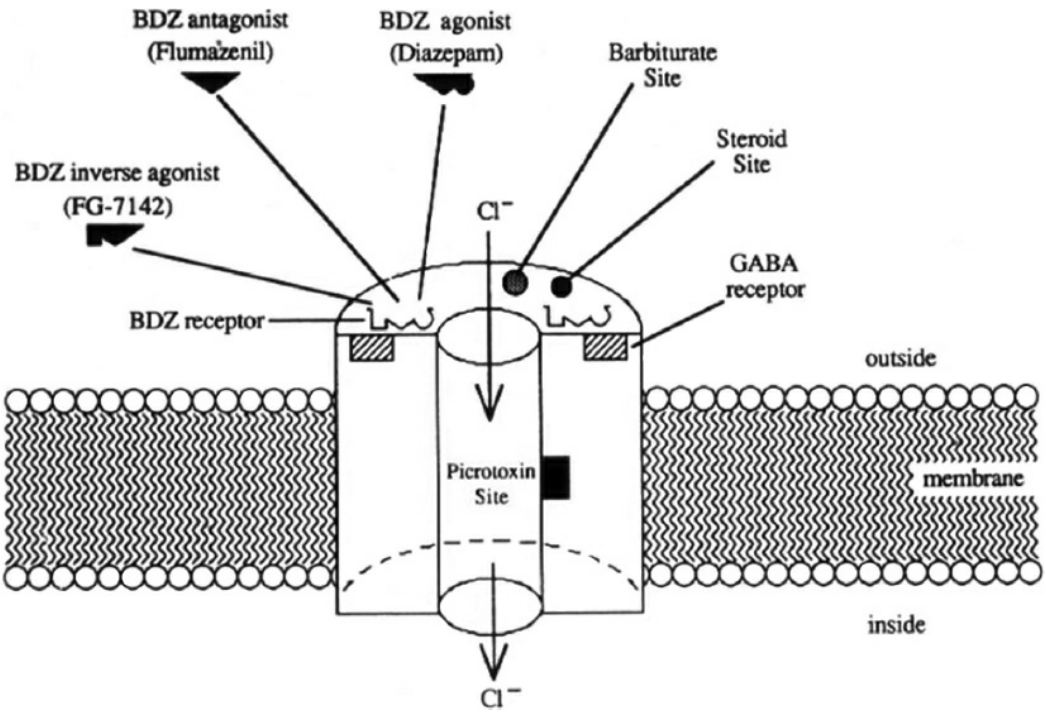

FIGURE 34.1–6. Structure and function of the GABAA receptor. A model for the receptor shows binding sites for GABA as well as compounds that affect GABA signaling. Activation of the receptor increases chloride conductance. Actions of these compounds affect either agonist binding or channel conductance. BDZ, benzodiazepine. (Adapted from Hudspeth AJ, Jessell TM, Kandel ER, Schwartz JH, Siegelbaum SA, eds. *Principles of Neural Science*. New York: McGraw Hill; 2013.)

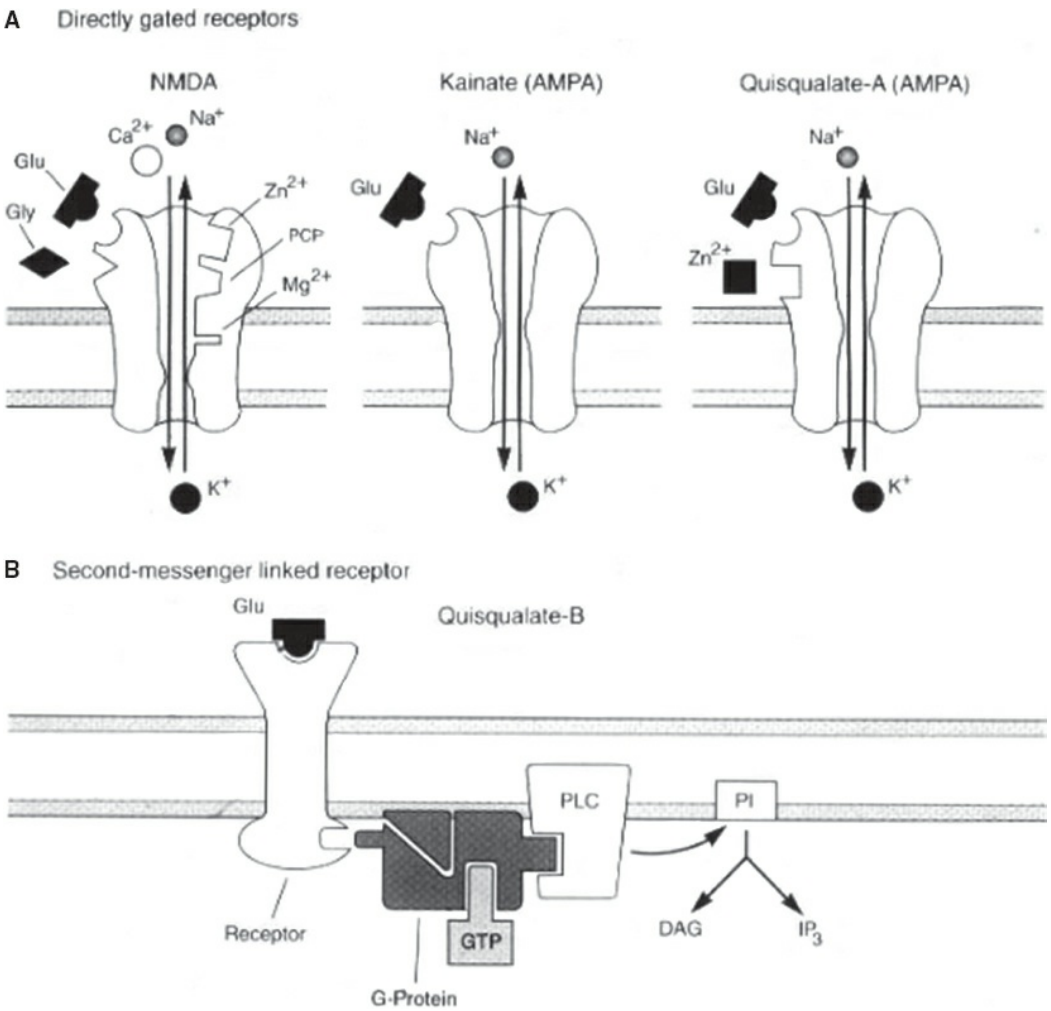

FIGURE 34.1–7. Glutamate receptors. Three classes of glutamate receptor regulate excitatory synaptic actions in neurons in the spinal cord and brain. Two types of ionotropic glutamate receptor gate ion channels directly. Two subtypes of non-*N*-methyl-D-aspartate (NMDA) receptor bind the glutamate agonists kainate or alpha-amino 3-hydroxy-5--methylisoxazole-4-proprionate (AMPA) and regulate a channel permeable to Na^+ and K^+. The NMDA receptor (left) regulates a channel permeable to Ca^{2+} and Na^+ and has binding sites for glycine, Zn^{2+}, phencyclidine (PCP or "angel dust"), MK801, and Mg^{2+}, which modulate the functioning of this channel in the direction indicated.

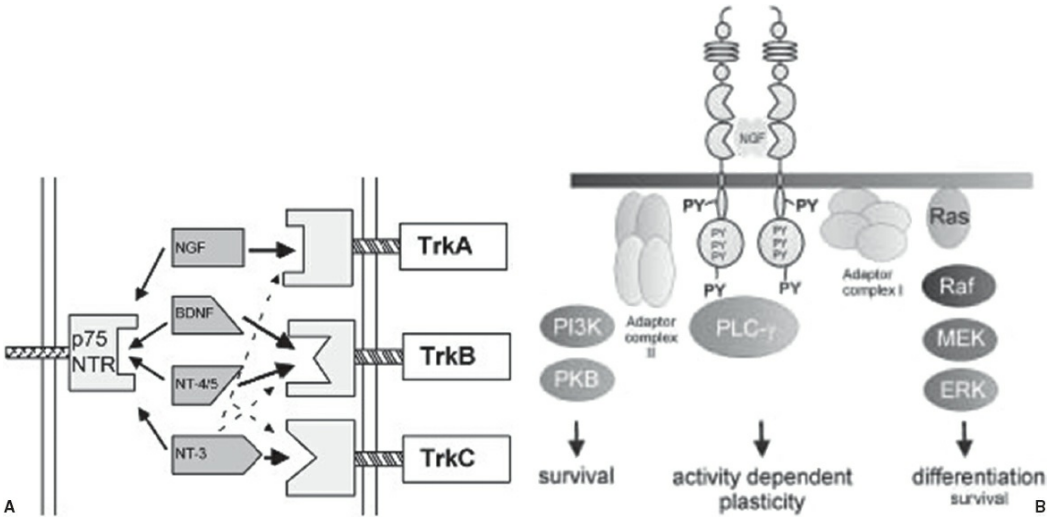

FIGURE 34.1–8. Neurotrophin receptors. **A:** All neurotrophins bind p75NTR, a receptor lacking tyrosine kinase activity. The receptors with tyrosine kinase activity are selective for the molecules indicated in the figure. **B:** Upon activation, the three Trk receptors dimerize, autophosphorylate, and phosphorylate intracellular targets. A variety of survival or differentiation pathways can be activated.

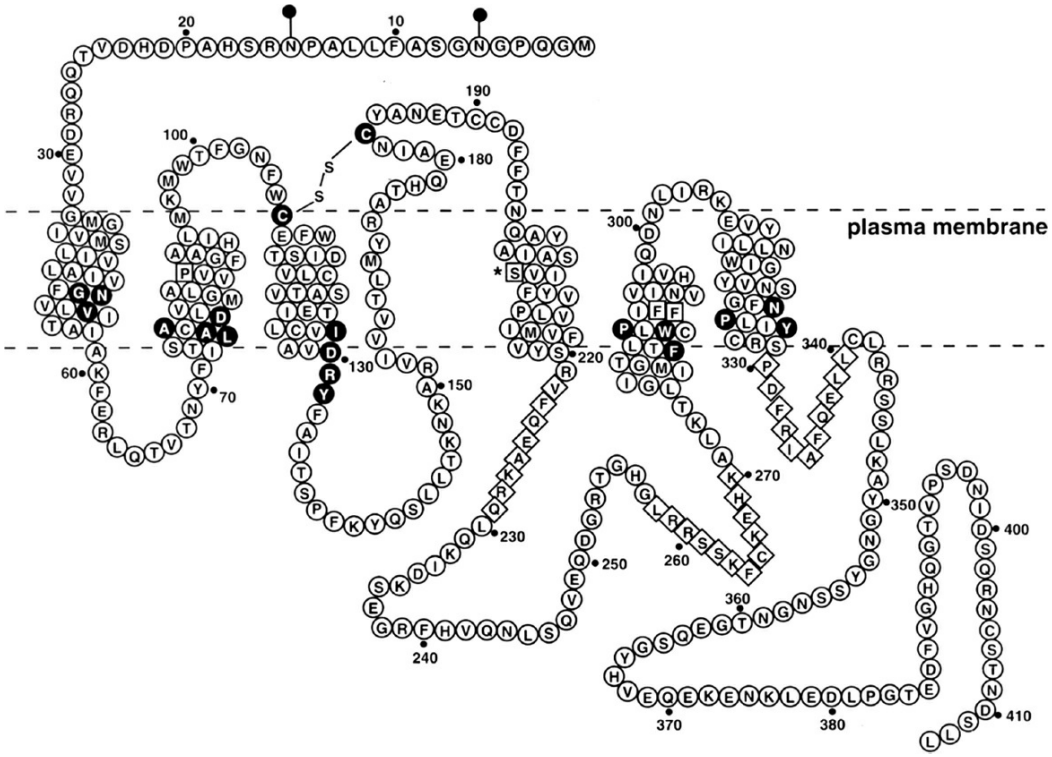

FIGURE 34.1–9. Model of a "classic" G protein–coupled receptor. The β1-adrenergic receptor is shown. Highly conserved residues (among all G protein–coupled receptors) are shown in black. Boxed residues are conserved in catecholamine receptors and the serine marked with * is conserved in all G protein–coupled serotonin receptors. Diamond residues are involved in G protein coupling. (Modified from Probst WC, Snyder LA, Schuster DI, Brosius J, Sealfon SC. Sequence alignment of the G-protein coupled receptor superfamily. *DNA Cell Biol.* 1992;11(1):1–20.)

Changes in Receptor Conformation upon Agonist Activation. Mutational analysis suggested several regions of GPCRs were involved in coupling to and activating G proteins. Curiously, the elucidation of the crystal structure of the β_2 receptor, suggested that there were distinct pattern of among GPCRs and their cognate G proteins.

Receptors with constitutive activity, which are active in the absence of agonist, can be generated by mutation. The importance of the C-terminal portion of the third intracellular loop was demonstrated in studies on the adrenergic receptor that is linked to norepinephrine-induced activation of phosphatidylinositol metabolism.

Another region relevant for G protein coupling is located at the junction between the third transmembrane segment and the second intracellular loop of many GPCRs. This region is delineated by a motif known as E/DRY (Glu/Asp-Arg-Tyr), which is highly conserved among GPCRs and

mutations result in constitutive activation of many different GPCRs.

Some naturally occurring mutations result in constitutive receptor activity. Shenker and colleagues (1993) identified a mutation (Asp578Gly/Tyr) in the sixth transmembrane domain of the luteinizing hormone receptor, which confers constitutive activation. Males carrying this mutation display precocious puberty. Some types of GPCRs display constitutive activity in their normal state. While D1a dopamine receptors require the presence of agonist to activate G protein, D1b receptors show high constitutive activity, almost completely activating G protein in the absence of agonist.

INVERSE AGONISM. The demonstration of naturally occurring constitutive activity for GPCRs raises the possibility that inhibition of such constitutive activity should have a functional effect. Thus, for a number of receptors, compounds exist that simply inhibit the binding of agonist (neutral antagonists) and other compounds exist that inhibit the activity of a constitutively active receptor (inverse agonists). This was demonstrated for the δ opiate receptor in studies which revealed that the "inverse agonist" ICI 174864 lowered the basal activation of Gi by these receptors. Considering the model of receptor–G protein complexes presented above, it would be predicted that inverse agonists would show preferential binding to receptors that were uncoupled to G proteins. Inverse agonism has been demonstrated at a number of receptor types.

BIASED AGONISM. Endogenous agonists for most GPCRs (with the possible exception of Wnt) activate the receptor so that it engages both a G protein and β-arrestin. While the G protein pathway often generates cellular second messengers, the β-arrestin pathway evokes receptor internalization and a series of β-arrestin-mediated intracellular events. Over the past few years, compounds have been developed that select (are biased for) either the G protein pathway or the β-arrestin pathway. While the usefulness for compounds that modify cardiac contractility without modifying hypertrophy are readily apparent, the therapeutic potential for biased agonism in psychiatry remains unestablished (Fig. 34.1–10).

RECEPTOR DESENSITIZATION. Receptors are also capable of becoming desensitized. This process involves the attenuation of receptor response in the continued presence of agonist. The process is quite different for ion channels and ionotropic receptors versus GPCRs. In the case of voltage-gated ion channels, there are processes that either desensitize or actually inactivate channels, and some of these events result from second messengers created subsequent to activation of GPCRs. Similarly, ionotropic receptors are often modified, either by altering affinity for agonist or by altering channel characteristics. Again, this is due to second messengers, often through phosphorylation events.

GPCRs, however, have specific accessory proteins that mediate their desensitization. During the process, G protein receptor kinases (GRKs) an enzyme family, phosphorylate receptors in the activated (agonist-bound) conformation and these receptors become a target for β-arrestins, which herd them into clathrin-coated pits from whence they internalize. Once internalized, they either recycle back to the surface or are degraded. Figure 34.1–10 illustrates the process.

Heterotrimeric G Proteins. Table 34.1–1 reveals the current list of G protein α subunits. Currently, there have been identified 17α, 5β, and 13γ genes in humans. Alternative splicing increases the number of available proteins. Each of these G proteins interacts with specific effector molecules. These can be enzymes like adenylyl cyclase, which makes cAMP; phospholipase C, which generates lipid messengers and increases intracellular calcium; or voltage-gated ion channels (particularly K^+ or Ca^{2+}), which are modified by Gβγ subsequent to activation of Gα.

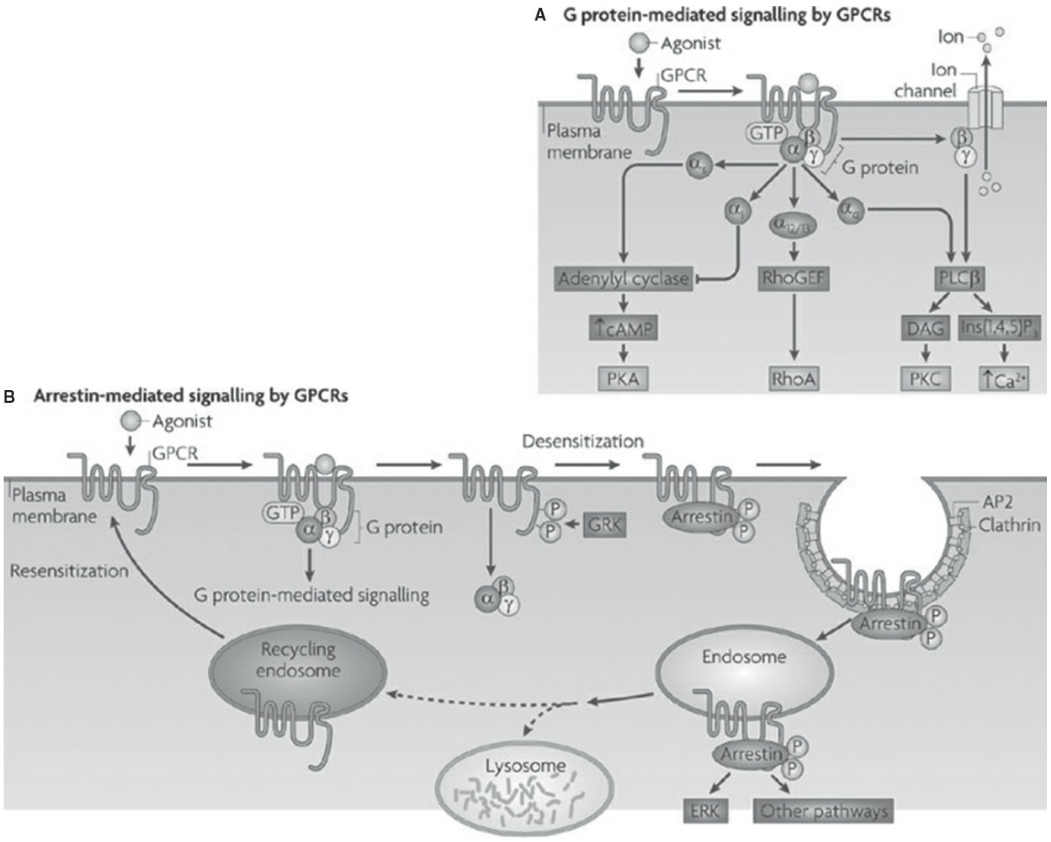

FIGURE 34.1–10. G protein–coupled receptor (GPCR) internalization and biased agonism. In the long-established model of GPCR signaling, an agonist binds to extracellular and/or transmembrane regions of the receptor, leading to its activation of G protein that results from exchange of GDP for GTP on Gα subunit, creating active Gα and Gβγ subunits. Termination of receptor activation can occur through GPCR phosphorylation by GPCR kinases (GRKs) and concomitant GPCR association with arrestins, which interact with clathrin to internalize GPCRs in endosomes. Recruitment of arrestins to activated GPCRs can also lead to the initiation of distinct arrestin-mediated signaling pathways, including activation of the mitogen-activated protein kinase extracellular signal-regulated kinase

(ERK) pathway. This can be done absent involvement of G protein, and agonists engaging either G protein or arrestin pathways are referred to as biased. Subsequent to internalization, GPCRs can be trafficked to lysosomes, where they are ultimately degraded, or to recycling endosomes for reinsertion to the plasma membrane. (From Ritter SL, Hall RA. Fine-tuning of GPCR activity by receptor-interacting proteins. *Nat Rev Mol Cell Biol.* 2009;10(12):819–830.)

Table 34.1–1.
Mammalian Heterotrimeric G Protein ` Subunits

Family	Subtype	Effectors
Gs(S)α	Gsα	Adenylyl cyclase, GTPase of tubulin
	Golfα	Adenylyl cyclase (olfactory)
Gi/oα	Go1α	Adenylyl cyclase inhibition (Gi,o,za) Go2
	Gi1-i3α	K^+ channels; Ca^{2+} channels inhibition (Gβγ)
	Ggust α	K^+ channels; Ca^{2+} channels inhibition (Gβγ) (transducin) cGMP-PDE in retina
	Gtα1/2	
Gq/11α	Gqα	Phospholipase C
	G11α	p63-RhoGEF
	G14α	
	G16α	
G12/13a	G12α	p115RhoGEF and related proteins
	G13α	

THE ACTIVATION CYCLE. Prior to binding agonist, GPCRs exist in a complex with heterotrimeric G proteins (αβγ subunits). The latter two exist as a "fused complex" and the former releases GDP and binds GTP upon receptor activation. This shifts the Gα to an active conformation and alters the association between Gα and Gβγ in such a manner that both engage effector molecules. Hydrolysis of GTP, by a GTPase intrinsic to the α subunit, terminates the activation cycle (Fig. 34.1–11).

RGS Proteins. For many Gα proteins, the GTPase is relatively slow, and regulators of G protein signaling (RGS) proteins facilitate the process. RGS proteins are a family of GTPase-activating proteins with conserved RGS domains (also called the RGS box) of about 120 amino-acid residues in length. RGS proteins interact directly with a subunit of heterotrimeric G proteins to catalyze GTP hydrolysis by enhancing the intrinsic GTPase activities of Gα, thus converting G proteins from a GTP-bound active state to a GDP-bound inactive state and terminating G protein–mediated signaling. While most G protein α subunits have been demonstrated to have their GTPase activated by a member of this large family of proteins, Gsα may be excluded. There are two sorting nexins that associate with Gsα (SNX13 and SNS14) that bind specifically with Gsα and are also known as RGSPX1 and RGSPX2, respectively. It is not clear that these proteins activate GTPase, and RGSPX2 clearly lacks this effect. Curiously, RGSPX2, a brain specific protein, helps mediate the trafficking of Gsα from the cell

membrane and also specifically modulates the cell-surface expression of the Gs-associated 5HT6 receptor.

Bacterial Toxins and G Proteins. Two bacterial toxins are known to modify G proteins by covalently ADP-ribosylating select amino acids within those G proteins. Cholera toxin inactivates the intrinsic GTPase of Gsα that results in sustained (permanent) activation of Gs and subsequent hyperactivation of adenylyl cyclase. Pertussis toxin primarily acts on Gi and Go proteins and uncouples these proteins from the receptors that activate them. Since a prominent effect of Gi/o is inhibition of adenylyl cyclase, pertussis toxin will attenuate this inhibition with the net effect of increasing cAMP.

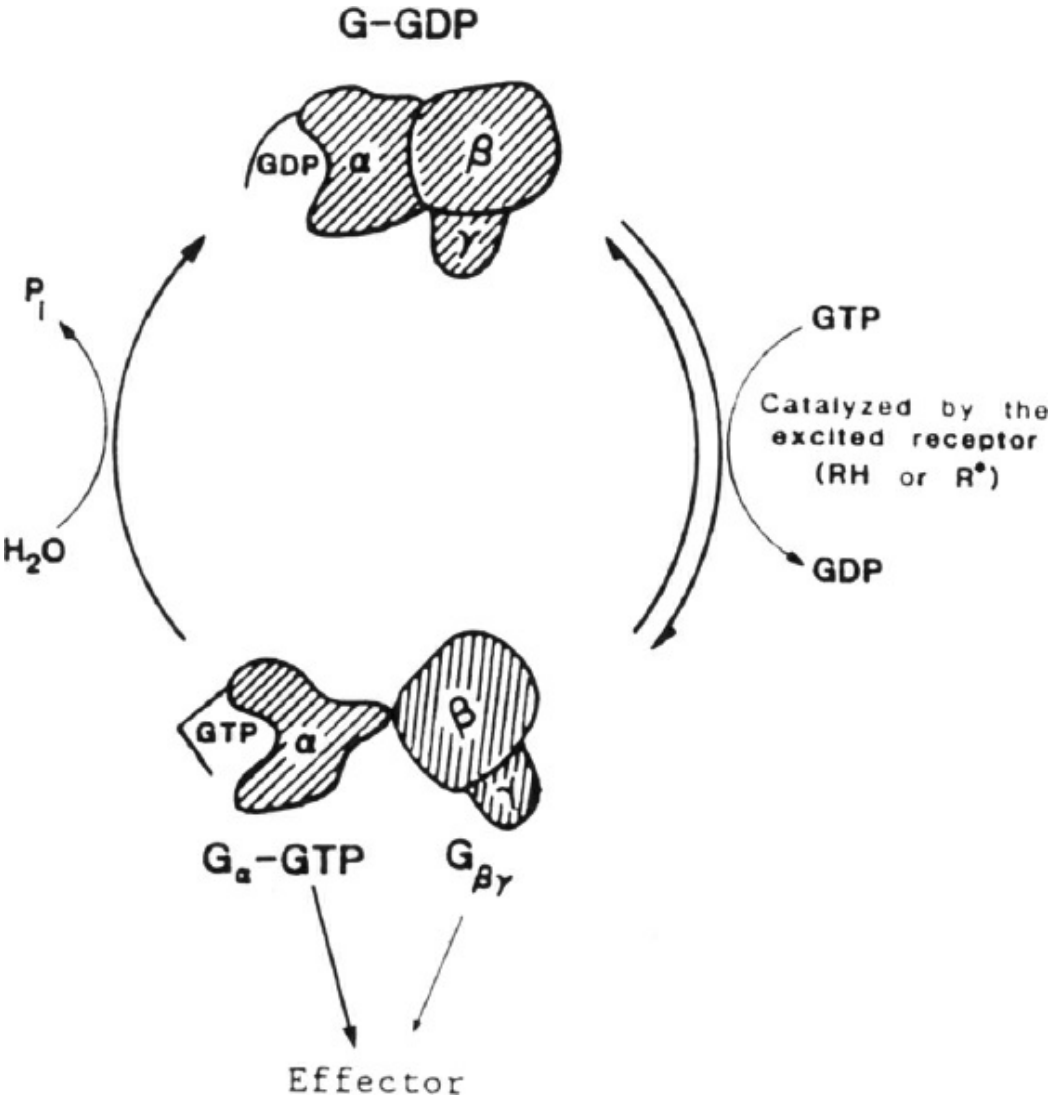

FIGURE 34.1–11. Activation cycle for G protein heterotrimers. Activation of the GPCR by agonist (described in Fig. 34.1–10) evokes release of GDP by G protein α subunit, allowing the binding of GTP. The GTP-bound conformation of Gα activates Effector molecules such as adenylyl cyclase of Phospholipase C. The activation of Gα is terminated by the hydrolysis of GTP for GDP. While the GTPase is intrinsic, there are molecules, such as RGS proteins, that facilitate the process.

Microdomains in the Plasma Membrane and GPCR Signaling

The plasma membrane, while comprised of lipids and proteins, enjoys a more complex organization than would be predicted from proteins floating in a lipid bilayer. Membranes are highly compartmentalized and are rife with domains of highly ordered and less ordered membrane lipids. The more highly ordered regions tend to be cholesterol-rich and these domains are normally associated with cytoskeletal proteins as well. The term used to describe these areas is "lipid rafts" and they are thought to range in diameter between 20 and 100 nM. Some of these structures contain the protein caveolin and form invaginated structures known as caveolae (little caves). At this point, these structures are conceptual, since they are below the level of resolution of the light microscope. Regardless, they can be manipulated experimentally, and it now appears that neurotransmitter action is, depending on the ligand and second-messenger pathway, facilitated or inhibited in caveolae and lipid rafts. Figure 34.1–12 shows the postulated organization of signaling molecules into lipid rafts and caveolae. These membrane microdomains have been suggested to be important sites of action for antidepressants and may be instrumental in evoking the disrupted signaling seen in depression.

G PROTEIN EFFECTORS

A vast array of cellular processes is mediated by G proteins. These molecules can be either activated or inhibited by G proteins and either α or βγ subunits have been demonstrated to function as activators. A list of some of these G protein–mediated events is given in Table 34.1–1 and a scheme is shown in Figure 34.1–13.

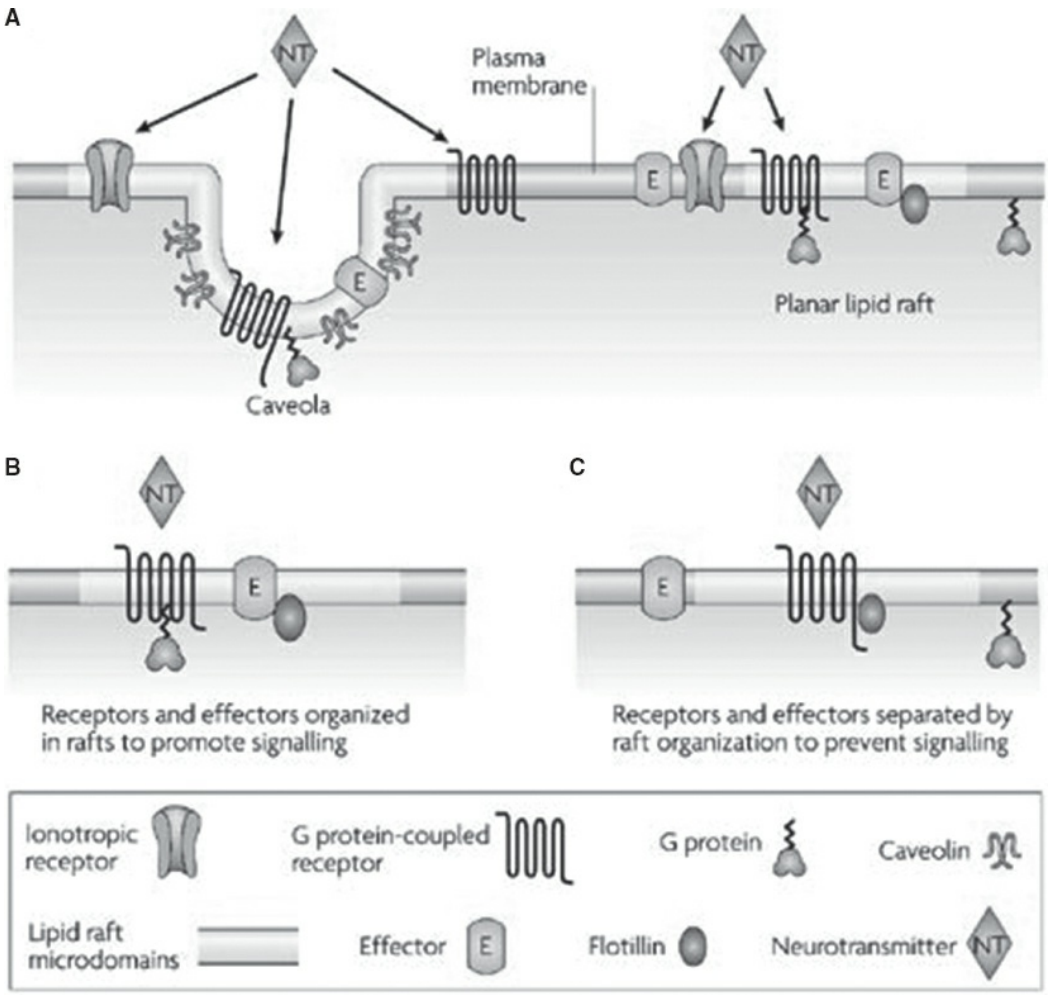

FIGURE 34.1–12. Lipid raft microdomains and membrane organization of neurotransmitter signaling molecules. **A:** There are two common raft domains in mammalian cells: planar lipid rafts and caveolae. Both possess a similar lipid composition. Planar rafts are essentially continuous with the plane of the plasma membrane and lack distinguishing morphological features. By contrast, caveolae are small, flask-shaped membrane invaginations of the plasma membrane that contain caveolins. Caveolins and flotillin can recruit signaling molecules into lipid rafts. Many neurotransmitter receptors (both ionotropic and G protein-coupled), G proteins, and signaling effectors such as second-messenger-generating enzymes are found in lipid rafts. Neurotransmitters might activate receptors that are located both within and outside lipid rafts. **B:** The lipid raft signaling hypothesis proposes that these microdomains organize, spatially, signaling molecules at the membrane, perhaps in complexes, to promote interactions that are necessary for signal transduction. **C:** Alternately, lipid rafts might inhibit interactions by separating signaling molecules, disrupting complexes and damping neurotransmitter response. (From Allen JA, Halverson-Tamboli RA, Rasenick MM. Lipid raft microdomains and neurotransmitter signalling. *Nat Rev Neurosci.* 2007;8:128–140.)

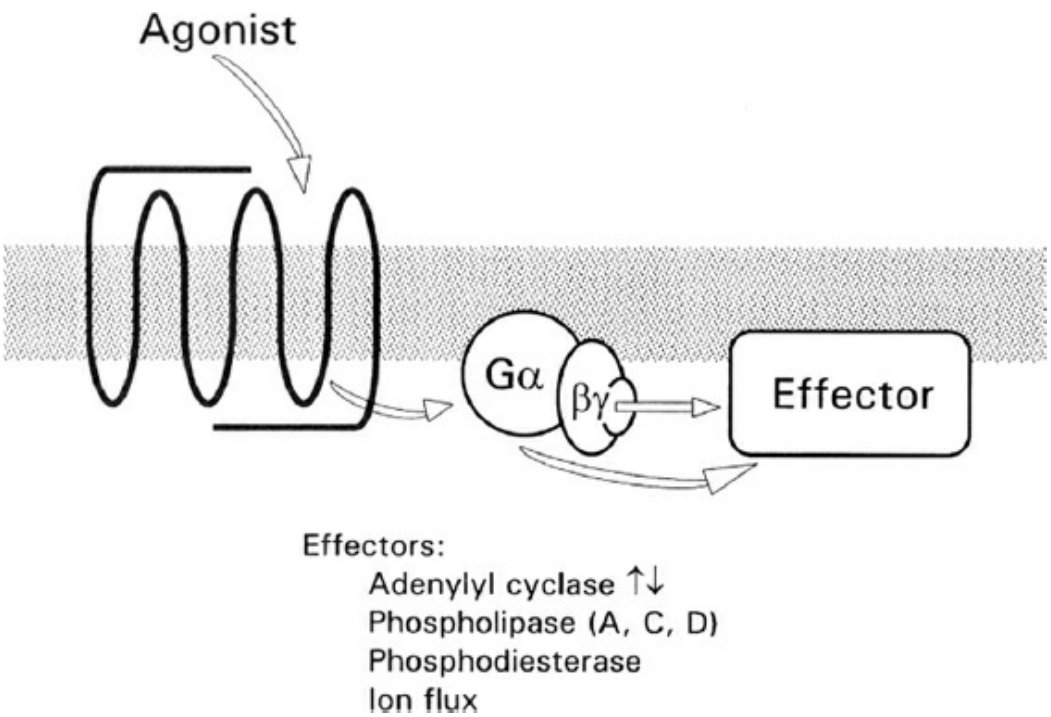

FIGURE 34.1–13. Coupling of G proteins to receptors and effectors. The activation of a G protein–coupled receptor (GPCR) by agonist allows -activation of the G protein (the α subunits exchange guanosine diphosphate [GDP] for guanosine 5'-triphosphate [GTP]) and subsequent activation of an effector molecule by α or βγ subunits. Several types of effector are listed. (From Felig P, Frohman LA. *Endocrinology and Metabolism*. New York: McGraw-Hill Education; 2001.)

Adenylyl Cyclase

This is the effector system that was being studied by Sutherland and Rall in their pioneering work on the activation of glycogen phosphorylase—work that led to the identification of cAMP and the concept of intracellular messenger molecules. Adenylyl cyclase, which generates cAMP from adenosine triphosphate (ATP), is a family of nine membrane-bound enzymes (plus a soluble version), each with a different mode of regulation. A common feature of all the adenylyl cyclase enzymes is activation by Gsα and (to varying degrees) by the diterpene, forskolin (Table 34.1–2). While Giα (three genes encode different forms of protein) was called "i" because it was inhibitory to adenylyl cyclase, only a few forms of adenylyl cyclase are inhibited directly by Giα. Calcium can act to stimulate some forms of the enzyme (in conjunction with calmodulin) and other forms are inhibited by calcium. While all forms of adenylyl cyclase are found in brain, these calcium-regulated types may be particularly relevant for psychopharmacology, as neurotransmitters that increase calcium (glutamate or acetylcholine) can also modify cAMP indirectly.

Table 34.1–2.
Adenylyl Cyclase Isoforms and Their Regulation

Regulator	Effect	Cyclase Isoform
Gsα	Stimulatory	All
Giα	Inhibitory	I, V, VI
GOα	Inhibitory	I
Gzα	Inhibitory	I, V
Gβγ	Inhibitory	I
Gβγ	Stimulatory[a]	II, IV, VII
Forskolin	Stimulatory[b]	All
Ca^{2+}/calmodulin	Stimulatory	I, III, VIII
Ca^{2+}	Inhibitory	V, VI, IX[c]
PKC	Stimulatory	I, II, III, V, VII
PKC	Inhibitory	IV
Calmodulin Kinase	Inhibitory	III

[a]Enzyme must first be stimulated by Gsα.

[b]Very weak stimulation for type IX.

[c]Requires calcineurin.

Phospholipases

Phospholipases responsive to G proteins fall into three families: phospholipase A, phospholipase C, and phospholipase D (Fig. 34.1–14). The products of phospholipase A are metabolites of arachidonic acid and are known as eicosanoids. The eicosanoids fall into the subgroups of prostaglandins and leukotrienes, and these molecules function as extracellular signals. The next group of phospholipids includes those that result from the breakdown of diacylglycerol and the inositol phosphates and these are the result of the actions of the phospholipase C enzymes. Gq and G11 activate phospholipase C, which has the net effect of increasing intracellular cellular calcium, both through the release of Ca^{2+} from intracellular stores and increasing capacitive calcium entry. 5HT2A receptors activate phospholipase C through a Gq intermediary. Finally, phospholipase D, which uses phosphatidyl choline as a substrate, can generate diacylglycerol and phosphatidic acid as well as lysophosphatidic acid. This latter molecule also has extracellular signaling properties.

G Protein–Activated Ion Channels

Direct activation of potassium channels by G proteins is important for the agonist-induced, rapid inhibition of membrane excitability in nervous tissue, and is essential for modifying both excitation and neurotransmitter release. As noted above, activation of GPCRs induces activation of G proteins, and both Gα and Gβγ can activate effector molecules, in this case, channel proteins. These channels (G protein–gated inward rectifier Kþ channels, GIRK) have been shown to bind directly and become activated by Gβγ. They are strictly voltage-independent, and channel gating requires association with intracellular proteins. Just as prominently (and also

resulting in an inhibition of a voltage-activated process), Gβγ subunits inhibit voltage-gated Ca^{2+} channels, attenuating neurotransmitter release. In these processes, the G protein, itself, is the second messenger (Fig. 34.1–15). Thus, the activation/inhibition of ion channels is somewhat more rapid than the generation of a second messenger and subsequent modification of channel function by phosphorylation.

INTRACELLULAR SIGNALING PATHWAYS

cAMP is synthesized by adenylyl cyclase and degraded by cAMP phosphodiesterase (PDE), which causes cAMP to be hydrolyzed to 5′AMP. Initial hormonal signals are more likely to regulate cAMP by the activation or inhibition of adenylyl cyclase than they are by altering PDE, but this is primarily because the receptors are coupled to G proteins that then directly stimulate or inhibit adenylyl cyclase. However, activation of a Gq-coupled receptor an NMDA receptor or a nAChR will increase cellular Ca^{2+}, which will activate calcium-sensitive variants of PDE, decreasing cAMP.

Regulation of cAMP and Ca^{2+} signaling begins with the activation of specific protein kinases sensitive to those messengers. For cAMP, the primary kinase is protein kinase A (PKA) (Fig. 34.1–16), while Ca^{2+} activates either calcium/calmodulin-activated kinases (CaM kinase) or protein kinase C, which is activated by Ca^{2+} and diacylclycerol. Each of these have a series of targets that engage a series of intracellular pathways, including modification of ion channels, modification of cytoskeletal elements (and cellular fine structure) and modification of neurotransmitter synthesis, transport, release or catabolism as well as modification of gene expression via cyclic AMP response element (CREB) phosphorylation (Fig. 34.1–17).

As a rule, intracellular messages require a specific termination switch, and phosphoprotein phosphatases served this purpose for the phosphorylations described above. Some of these enzymes are specific for PKA (protein phosphatase IIb) or CaM kinase (calcineurin).

Regulation of adenylyl cyclase has been discussed above. The regulation of PDEs is quite complex, and much of it is due to the exquisite and unique subcellular localization of a myriad of PDE species. Further, the localization of cAMP-activated protein kinases (PKA) and a variety of PDEs is regulated by a family of proteins known as A kinase anchoring proteins (AKAPs). These proteins, often referred to as cellular scaffolds, provide a spatial dimension to second messenger signaling as well as a temporal discretion provided by close physical relationship of the relevant proteins.

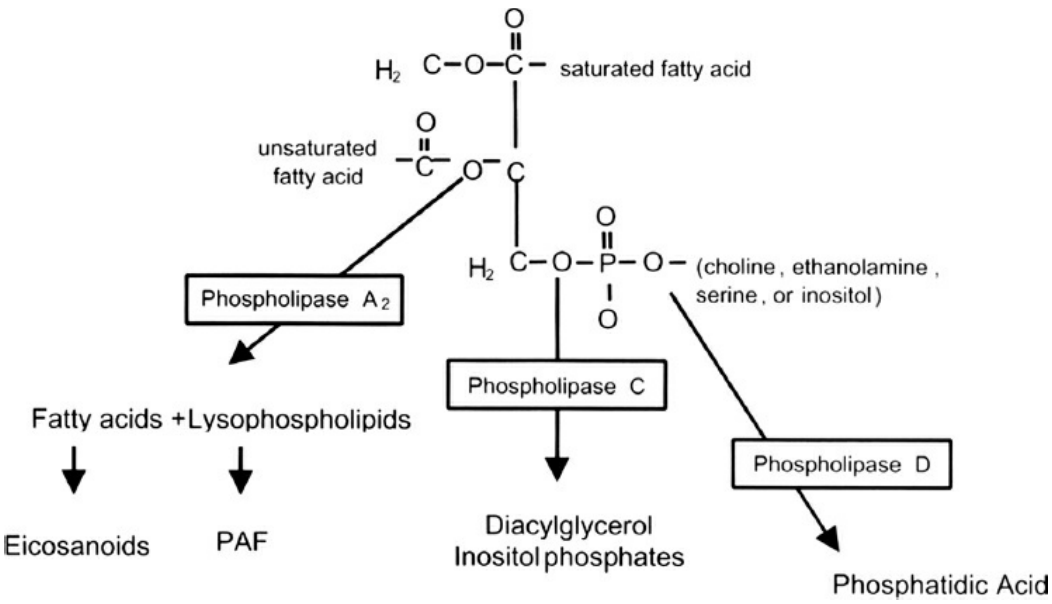

FIGURE 34.1–14. Substrates for, and products of, G protein–coupled phospholipases. PAF, platelet activating factor. (From Felig P, Frohman LA. *Endocrinology and Metabolism*. New York: McGraw-Hill Education; 2001.)

A Direct opening of the GIRK channel by a G protein

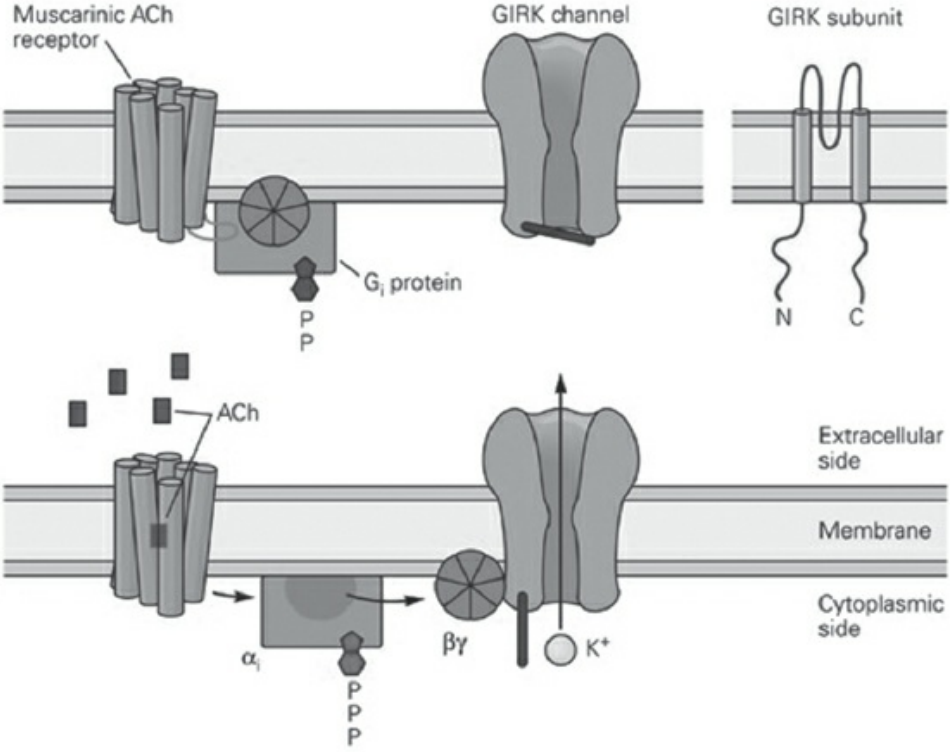

B Opening of GIRK channels by ACh hyperpolarizes cardiac muscle cells

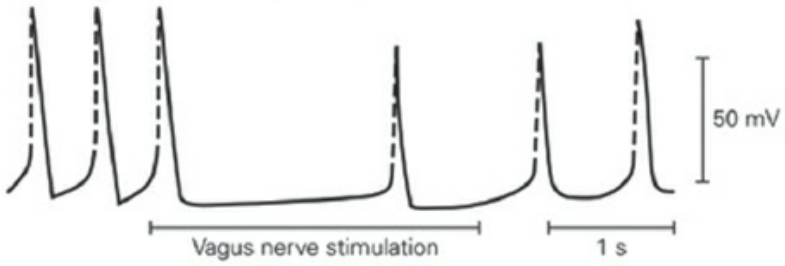

C Opening of GIRK channels by ACh does not require second messengers

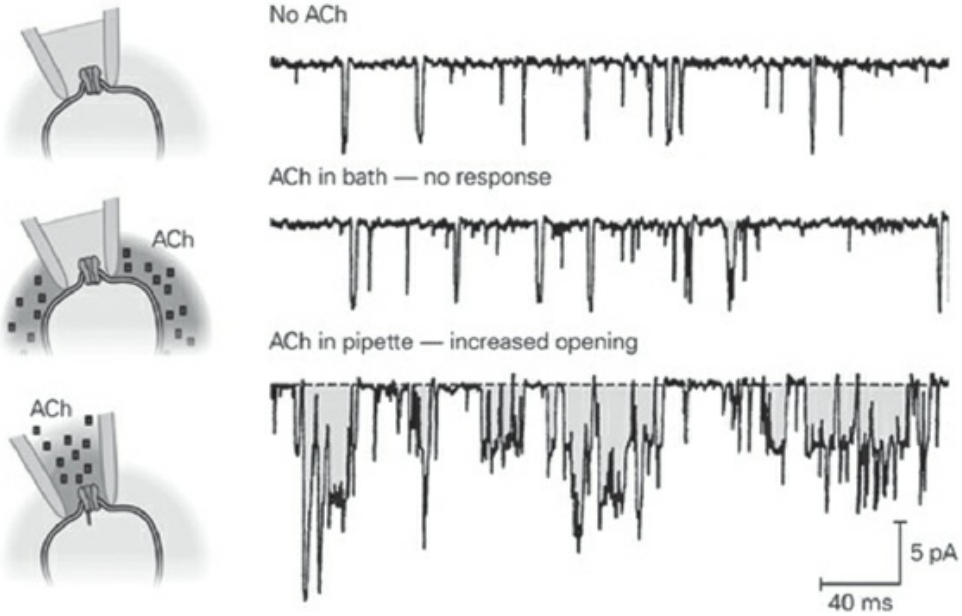

FIGURE 34.1–15. Direct opening of an ion channel by a G protein. In this instance, activation of an mAChR by acetylcholine allows Gβγ to bind to and open a K^+ channel known as a G protein–coupled inwardly rectifying K^+ channel (GIRK). Similarly, Gβγ can bind to and inhibit voltage gated Ca^{2+} channels.

Epac, Another cAMP Target

The default cAMP pathway is PKA, but it is noteworthy that other cellular proteins are regulated by cAMP independent of phosphorylation. One of these cAMP targets is a family of proteins known as Epac (exchange proteins activated by cAMP). These elicit guanine nucleotide exchange and subsequent activation of a number of small G proteins involved in cell growth and synaptogenesis. Furthermore, there are ion channels (particularly important in olfaction) that are directly opened subsequent to the binding of cAMP.

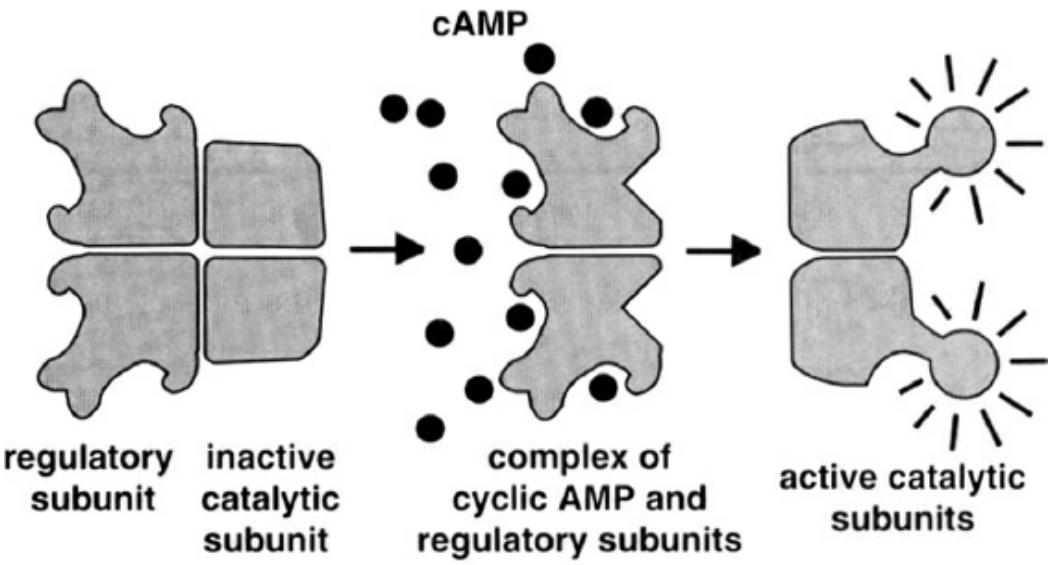

FIGURE 34.1–16. Activation of cyclic adenosine monophosphate (cAMP)-dependent protein kinase (PKA). Binding of cAMP to the regulatory subunits induces a conformational change, causing these subunits to dissociate from the complex, thereby activating the catalytic subunits. Each regulatory subunit has two cAMP-binding sites, and the release of the catalytic subunits requires the binding of more than two cAMP molecules to the tetramer. This greatly sharpens the response of the kinase to changes in cAMP concentration. There are at least two types of A kinase in most mammalian cells: type I is mainly in the cytosol, whereas type II is bound via its regulatory subunit to the plasma membrane, nuclear membrane, and microtubules. In both cases, however, once the catalytic subunits are freed and active, they can migrate into the nucleus (where they phosphorylate gene-regulatory proteins such as CREB; see Fig. 34.1–17). (From Felig P, Frohman LA. *Endocrinology and Metabolism*. New York: McGraw-Hill Education; 2001.)

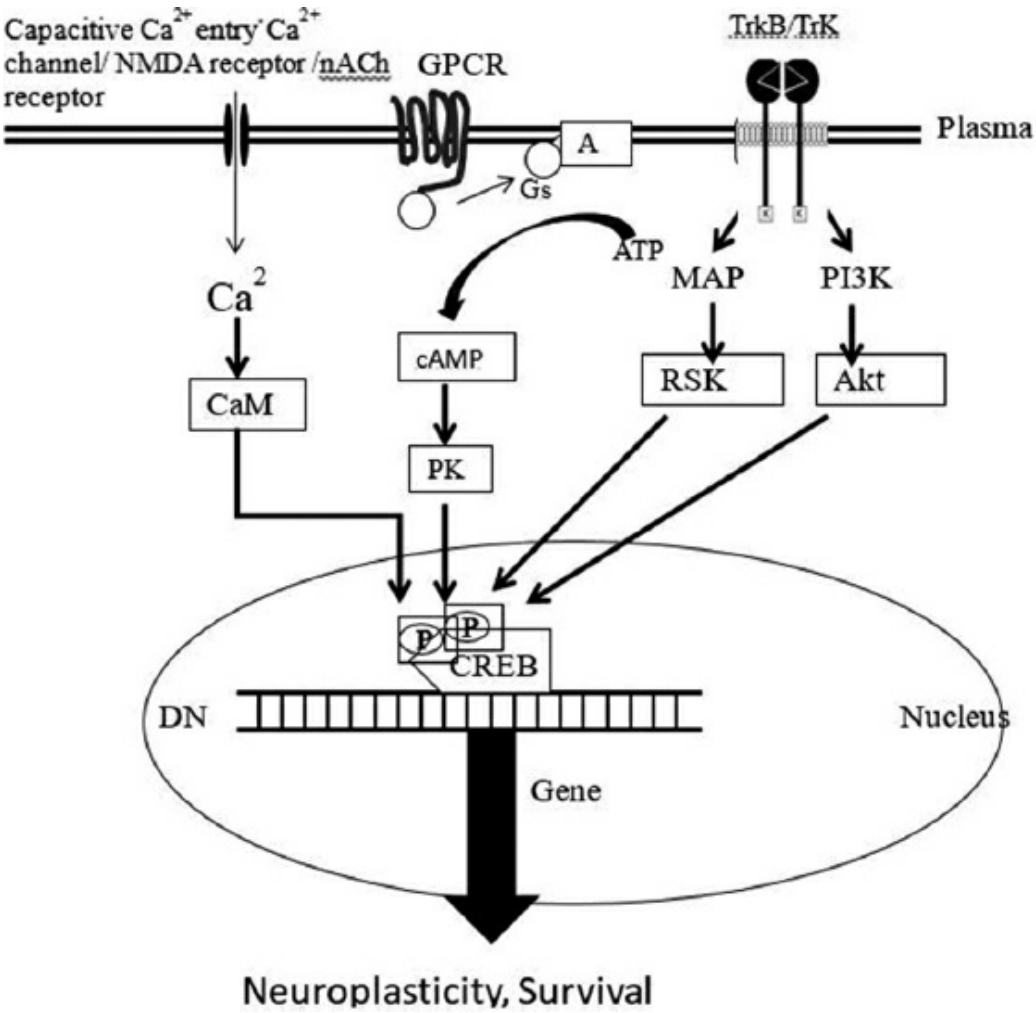

FIGURE 34.1–17. cAMP and Calcium signaling orchestrate gene expression. Mechanisms for cAMP generation, calcium entry, and regulation of nuclear and intracellular calcium-responsive proteins are demonstrated. This scheme also suggests an interface for the action of growth factor receptors, GPCRs and Ca^{2+}.

Table 34.1–3. Psychotropic Drugs with Postmarketing Black Box Warnings

Drug	Psychiatric Application	Basis for Warning
Clozapine (Clozaril)	Refractory schizophrenia	Myocarditis
Droperidol (Inapsine)	Sedation of agitated patients	Torsades de pointes or QT and QTc prolongation, or both, and other cardiovascular events
Thioridazine (Mellaril)	Psychosis	Torsades de pointes or QT and QTc prolongation, or both, and other cardiovascular events
Divalproex (Depakote)	Acute mania	Hemorrhagic pancreatitis and hepatic failure
Nefazodone (Serzone)	Depression	Acute hepatic failure
Pemoline (Cylert)	Attention-deficit/hyperactivity disorder	Acute liver failure
Antidepressants (in patients <24 years old)	Depression	Increased risk of suicidality

Antipsychotics (for dementia-related psychoses)	Psychosis	Increased risk of mortality

INTEGRATION OF NEUROTRANSMITTER SIGNALS

Most diagrams of synaptic signaling, including those in this chapter, depict a single neurotransmitter, engaging a single receptor, which then evokes a programmed series of responses through a set pathway. While as a didactic principle this is useful, neurotransmitter signaling, both at extracellular pathways and within cells, enjoys a far greater complexity. First, hundreds (even thousands) of axons converge upon a single dendrite (or even on a single dendritic spine). Second, a single transmitter, acting through a myriad of receptor types, could evoke multiple intracellular pathways. This is illustrated for serotonin in Table 34.1–3. Finally, each of these pathways are interdependent, as illustrated for Ca^{2+} and cAMP signaling pathways in Figure 34.1–18.

GENERAL ISSUES IN CLINICAL PSYCHOPHARMACOLOGY

Reflecting the expanded role of psychopharmacology, as well as the variety and sheer number of psychotropic drugs available, the sections on psychotropic drugs that follow represent a virtual textbook within a textbook. These sections are also complemented by discussions in other chapters that focus on the treatment of specific disorders and special patient populations. This section is intended as a review of important principles underlying the selection and use of psychotropic medications. It provides a review of common issues that arise in the clinical setting, as well as considerations for more effective use of available agents. Every effort has been made to provide a comprehensive discussion of the properties of the drugs, as well as guidance on their most appropriate use. Nevertheless, it needs to be emphasized that psychopharmacology is a rapidly evolving field and that many patients experience idiosyncratic or paradoxical responses to treatment.

History

For centuries, people have used plant extracts and naturally occurring substances to influence mood and anxiety, to modulate sleep and arousal, to produce euphoria, and to alter consciousness. Notable examples include alcohol, coffee, tea, opium preparations, bromides, and hallucinogens. Use of cocaine-containing soothing syrups and elixirs was common. In the nineteenth and early twentieth centuries, several types of *sedatives* were developed. Some of these compounds, chloral hydrate (Aquachloral Supprettes), and the barbiturates, are still in use. From the 1930s to the 1950s, the most effective biological treatment of a mental disorder was electroconvulsive therapy (ECT), but the image of the so-called shock

therapy was unpalatable to many patients. In 1931, two psychiatrists in India, Ganpath Sen and Kartik Chandra Bose, described the ability of *Rauwolfia serpentina* to reduce symptoms of psychosis, but this report was overlooked by Western medicine for many decades. When rediscovered, reserpine (Serpalan) became one of the first agents to offer benefit to patients with schizophrenia. Understanding of reserpine's effects on monoamine neurotransmitters spurred further research into the involvement of norepinephrine, serotonin, and dopamine in the treatment of psychiatric disorders, and the synthesis of drugs that targeted these agents. Similarly, published findings on the antimanic properties of lithium salts appeared in 1949, but were neglected by most psychiatrists. It was more than two decades before lithium became accepted as the standard treatment of bipolar disorder.

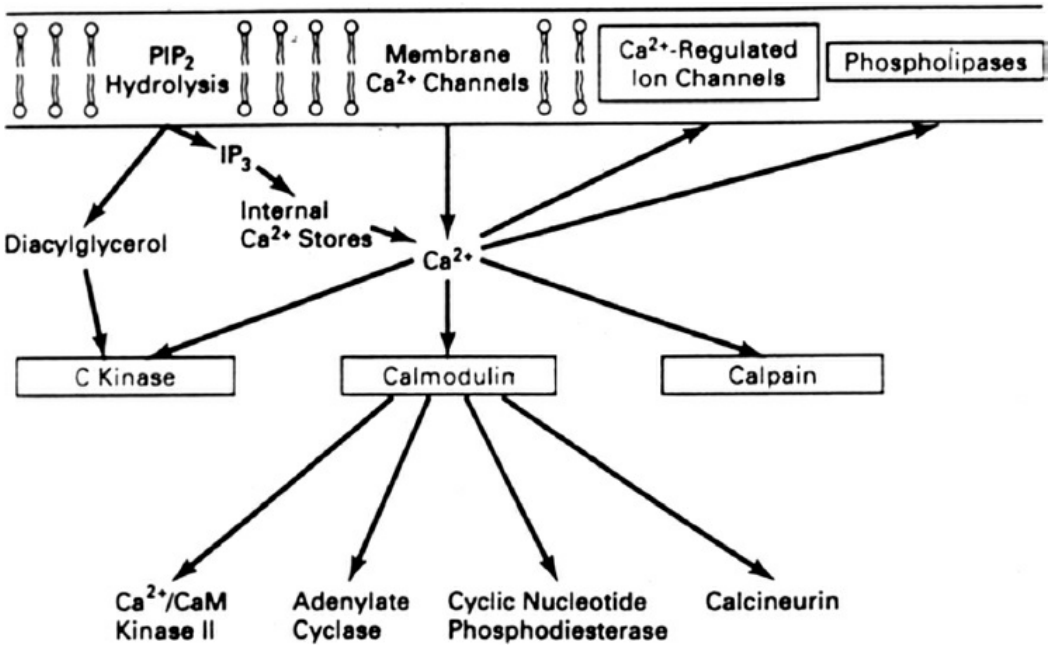

FIGURE 34.1–18. Interface among second messenger systems: many roles for calcium. Several conduits for Ca^{2+} signaling are shown and the interactions among them and cAMP signaling pathways are described in the text.

The field of psychopharmacology evolved rapidly during the second half of the 20th century. It was the nearly simultaneous discoveries of several classes of medications that began the transformation of psychiatric treatment from a psychoanalytic to a biological orientation. The first breakthrough occurred between 1950 and 1952 with the observation by Jean Delay and Pierre Deniker that chlorpromazine—a presurgical antihistamine—brought about dramatic control of psychotic symptoms associated with mania and schizophrenia. The effect of the drug was described as "miraculous" and "astonishing." Chlorpromazine was released for clinical use in 1954, and within 8 months, was administered to more than 2 million patients. Patients with long histories of hospitalization

owing to disorganized thinking and bizarre behavior demonstrated sufficient improvement to be discharged from their overcrowded institutions. For a time, there were expectations that chlorpromazine and its analogs, then referred to as *tranquilizers,* could eradicate or permanently cure schizophrenia. Although this turned out to be a false hope, newer versions of these drugs improve the quality of life of many patients, enabling them to be productive and to live in the community.

In 1952, iproniazid (Marsilid), a drug being tested as a treatment of tuberculosis, was observed to cause a dramatic elevation in mood. Over the next few years, psychiatrists came to recognize this antidepressant action of iproniazid and similar isonicotinic hydrazide analogs. Cases of unexpected hypertensive episodes associated with monoamine oxidase inhibitors (MAOIs) were reported soon after the introduction of these drugs. Some patients died of strokes. An explanation for this so-called cheese reaction did not emerge until 1963 in an article by Barry Blackwell.

Shortly after the discovery of MAOI antidepressant activity, imipramine (Tofranil), a derivative of chlorpromazine, was observed to alleviate depressive symptoms. Imipramine, the first tricyclic antidepressant (TCA), and iproniazid, the first MAOI, became available for clinical use in 1957. Within 1 year of its release, more than 400,000 patients received imipramine for depression.

Despite the progress that has been made in more than a half century, no antidepressants have surpassed the TCAs and MAOIs in efficacy. Nevertheless, these drugs have never lived up to their full potential because of troublesome side effects and concerns about safety. The dietary restrictions reduced the acceptability of MAOI treatment. The anticholinergic side effects of TCAs and their toxicity in overdose also had a restraining effect on their use.

The availability of meprobamate (Equanil) in 1955 marked the first time that a treatment of anxiety had a relatively wide margin of safety. Compared to the barbiturates, meprobamate caused less dose escalation and a less severe withdrawal syndrome. By the end of that decade, an even better tolerated and safer alternative emerged. In 1957, the benzodiazepine chlordiazepoxide (Librium) was found to produce sedative, muscle relaxant, and anticonvulsant properties but without significant risk of serious toxicity in overdose or during withdrawal. By 1960, chlordiazepoxide was marketed, soon followed by diazepam (Valium). It is worth noting that chlordiazepoxide was introduced into clinical practice just 2.5 years after its pharmacological studies were started.

In 1958, Paul Janssen synthesized another important chemical family of compounds used to treat psychoses—the butyrophenones. The prototype of this class—haloperidol (Haldol)—remains one of the most popular drugs used for the management of psychosis. The first atypical antipsychotic was synthesized in the late 1950s, but it was another of Paul Janssen's

compounds, risperidone (Risperdal), that became the first widely prescribed *atypical* neuroleptic or serotonin-dopamine antagonist.

Attitudes toward psychotropics were transformed in the early 1980s after reports were published that alprazolam (Xanax) could suppress panic attacks. Until that time, only TCAs and MAOIs had been shown to control panic symptoms, and benzodiazepines were not thought to be effective in treating this severe form of anxiety.

Based on the newly gained understanding that TCAs block not only norepinephrine reuptake but also serotonin, Arvid Carlsson and his colleagues set out to develop a selective inhibitor of serotonin reuptake. Zimelidine, the first of these selective serotonin reuptake inhibitors (SSRIs), was developed in the early 1970s, and was found to be a potent antidepressant. It was marketed in Europe but was quickly withdrawn from clinical use because of a rare neurological side effect. It was not until the introduction of fluoxetine (Prozac) late in 1987 that an SSRI became available in the United States. Early clinical experience with fluoxetine was followed by the surprising realization by patients and physicians that the drug sometimes produced dramatic responses without the same type of tolerability problems associated with the TCAs. Quickly, first-line treatment of depression quickly shifted from the TCAs to fluoxetine and other SSRIs. Even around the start of the 21st century, the majority of new prescriptions for antidepressant treatments were for SSRIs.

The first psychotropic drugs were thus discovered through observations of behavioral effects in animals or humans. As the mechanisms underlying the actions of these drugs have been determined, it has facilitated rational drug development of new compounds, such as the SSRIs, with more selective action. It also has led to the identification of disorders that are preferentially responsive to specific agents. Patients now have many more treatment options; the drugs used are better understood. One of the more remarkable transformations in the field of psychopharmacology has been the degree to which psychotropic drugs are now routinely prescribed by nonpsychiatrists.

Two large studies have helped to provide a perspective on how difficult it is to differentiate the growing number of psychotropic agents in terms of their overall effectiveness. The findings of these studies, the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) and the Sequential Treatment Alternatives to Relieve Depression (STAR*D) trial, point out that patient responses to medications tend to be unpredictable, that no interventions are effective in all patients with a specific disorder, and that tolerability and safety often become the determinant of treatment success or failure.

The STAR*D trial was a large, federally funded study that investigated treatment effectiveness in a “real-world” sample of depressed outpatients treated in psychiatric and primary care practices. There were 3,671

participants in the trial. It found that remission was more common among patients in their first or second treatment trial (37 and 31 percent, respectively) than among those requiring three or four trials (14 and 13 percent, respectively). Patients needing more treatment steps tended to have more chronic depression and more coexisting psychiatric and general medical disorders. Overall, 67 percent of patients who completed up to four treatment steps reached remission. Results for STAR*D revealed higher relapse rates for those who required more initial treatment steps. Patients were allowed to choose among groups of subsequent treatments, which included cognitive therapy alone or combined with medication, as well as several different antidepressants used alone or in combination. The study failed to show any advantage, in terms of switching or combining drugs, for one strategy over another. The study found higher relapse rates among patients who failed to reach full remission and among those who went through a higher number of treatment steps to reach remission.

The CATIE study, funded by the National Institutes of Health's National Institute of Mental Health (NIMH), analyzed the effectiveness of an older (first-generation) antipsychotic medication perphenazine versus the newer (second-generation) medications during initial treatment. It found that the older medication might be more cost-effective for some patients with schizophrenia. Perphenazine was less expensive and no less effective than the newer drugs used in the trial. In phase 1 of the trial, patients were randomly assigned to receive olanzapine, perphenazine, quetiapine, risperidone, or ziprasidone for up to 18 months. Clozapine was included in phase 2 of the study. The study found that olanzapine had the longest time to discontinuation in phase 1 but caused the most weight gain and metabolic syndrome. Perphenazine did not significantly differ in overall effectiveness from quetiapine, risperidone, and ziprasidone. There were no differences in neurocognitive or psychosocial functioning in response to medications.

Classification

Traditionally, classification and categorization of psychotropic medications have been based on the primary indication for these medications. With each subsequent edition of this textbook, the classification and categorization of psychotropic drugs have changed. This is not merely a consequence of the development of new compounds. It is also the result of advances in the understanding of how these compounds produce their clinical effects. It was not long ago that psychotropic drugs were broadly defined as either stimulants or sedatives. The term *tranquilizer* was applied to antianxiety/sedative drugs (minor tranquilizers) and antipsychotics (major tranquilizers). At first, the only antidepressants were TCAs or MAOIs. However, when newer antidepressant drugs emerged in the 1970s and 1980s, they were labeled as second- or third-generation

antidepressants.

Defining drugs by their major indications is problematic. For example, the standard categories of drugs are (1) antipsychotic drugs or neuroleptics used to treat psychosis, (2) antidepressant drugs used to treat depression, (3) mood stabilizers used to treat bipolar disorder, (4) antianxiety or anxiolytic drugs used to treat anxious states, (5) hypnotic agents to promote sleep, and (6) psychostimulants to treat attention-deficit disorder. As a rule, clinicians think of drugs primarily in terms of their major clinical application. However, few psychotropic drugs are diagnosis specific, and most of the newer drugs assume multiple therapeutic roles over time. Benzodiazepine "anxiolytics," for example, are also very effective as hypnotics. Antidepressant drugs have a wide spectrum of action. Tricyclics like amitriptyline, which have significant histamine H1 antagonism, have been used in urticarial and other dermatologic disorders. Agents such as the SSRIs and venlafaxine (Effexor), all originally used to treat depression, have gained FDA approval as treatments of disorders as diverse as panic disorder, generalized anxiety disorder, social anxiety disorder, obsessive-compulsive disorder (OCD), posttraumatic stress disorder (PTSD), and premenstrual dysphoric disorder and chronic pain. Similarly, all atypical antipsychotic agents have become commonly used as mood stabilizers and as augmentation agents to treat depression and OCD. Such drugs as clonidine (Catapres), propranolol (Inderal), verapamil (Isoptin), and pregabalin (Lyrica), all not officially recognized as psychotropic agents, are widely used to manage a variety of psychiatric disorders. As a result, there is a growing movement to reclassify psychotropic drugs according to mechanism of action rather than by indication. The major neuropsychopharmacology societies around the world have come together to create a pharmacologically driven nomenclature for psychiatric drugs. While this system has yet to widely accepted and adopted, it suggests a more rational way of conceptualizing these medications.

REGULATORY ISSUES

In the United States, the FDA determines the safety and efficacy of medication. The FDA has the authority to approve a drug for clinical use if it is found to be "safe for use" and "effective in use." The FDA is also charged with ensuring that product labeling is truthful and contains all information pertinent to the safe and effective use of that drug. This includes monitoring the postmarketing experience with a drug, proposing changes in the product information, and overseeing the way in which drugs are manufactured. Antidepressants, antipsychotics, hypnotics, mood stabilizers, and psychostimulants have been among the drugs with additional warnings due to postmarketing monitoring. In addition to the FDA, another level of government regulation is directed by the Drug Enforcement Agency (DEA), which classifies drugs according to their abuse

potential. Clinicians are advised to exercise increased caution when prescribing controlled substances (control levels II through IV).

Prescribing Information

FDA-approved product information for marketed drugs appears as a package insert. It is also referred to as the *product labeling*. The *Physicians' Desk Reference* (PDR) is an annual collection of product information for prescription drugs. The product information for most drugs—not just psychotropic agents—contains an extensive list of potential side effects, drug interactions, the need for special monitoring, and restrictions for use. This information reflects the results of critical evaluations submitted to the FDA about the safety and efficacy of the drug for the approved indications. Postmarketing experience is added to the product information over time. The limitations of the prescribing information are evident in the many adverse events that are underreported during clinical trials and the frequent use of doses of medications in clinical practice that exceed the recommended dosing limits. Product information is also widely available through online resources such as drugs.com and Epocrates.

Black Box Warning

The prescribing information for a drug sometimes contains a *black box warning* to alert physicians about potentially important safety information. The term describes a black box at the beginning of the package insert that encloses the warning. It is considered the strongest warning, being more serious than just bolded text. A black box warning may be included in the product information when a drug is approved by the FDA or it can be added later. The presence of a black box warning deters some clinicians from using a drug as a first-choice agent. In some cases, manufacturers have decided against bringing a drug to market because of the competitive disadvantage that a warning label confers. A controversial example is the black box warning for the increased risk of suicidality in young patients taking antidepressants issued in 2004. The biostatistician Robert Gibbons and his colleagues noted that prescribing rates decreased and suicide rates increased since the publication of the warning. Studies such as this led to revised wording of the warning by the FDA so that clinicians would not be discouraged from using antidepressants when clinically indicated. In 2005, the FDA issued a black box warning about increased risk of mortality with the use of atypical antipsychotics for dementia-related psychosis. This was expanded to include conventional antipsychotics in 2008.

The longer that a drug has been in clinical use, the better is the understanding of its side effect profile. Even though new agents undergo extensive scrutiny for safety and tolerability, clinical trials do not typically

capture the full range and severity of all side effects. Most clinical trials are of too short duration and lack sufficient power to detect the full extent of side effect risk. Approximately 10 percent of all drugs approved in the United States in recent years have acquired a black box warning after marketing or have been withdrawn. In terms of timing, one half of black box additions occurred within the first 7 years after a drug was marketed. Psychotropic drugs with added black box warnings (Table 34.1–3) include nefazodone (Serzone), pemoline (Cylert), droperidol (Inapsine), lamotrigine, clozapine (Clozaril), divalproex (Depakote), and thioridazine (Mellaril). In the last-named case, the warning was added after the drug had been in use for many decades. In the context of safer alternatives, the benefits of thioridazine no longer outweighed the risk of cardiac conduction abnormalities. A warning about the risk of stroke in elderly patients with dementia has been added to the prescribing information for risperidone.

Approximately one half of drug withdrawals due to safety issues occur within the first 2 years after a drug is marketed. Between 1993 and 2002, there were seven drug recalls that followed reports of death or severe health complications. None of these involved psychotropic drugs. This is a remarkable testimonial to an often-overlooked fact: Psychotropic drugs are among the safest drugs available. Since nomifensine (Merital) was withdrawn in 1986, within 1 year of its marketing, no other psychotropic agent has been withdrawn because of serious side effects.

PHARMACOLOGICAL ACTIONS

Heterogeneity of response characterizes all psychotropic drugs. Among other considerations, genetic and environmental factors influence individual response to and tolerability of psychotropic agents. Thus, a drug that may not prove effective in large numbers of patients with a disorder can dramatically improve symptoms in some. In these cases, the identification of characteristics that might predict potential candidates for that drug becomes important but often remains elusive. Attention is being focused on gaining a better understanding of the thousands of genes that regulate the expression of drug-metabolizing enzymes and of the complex functions of the brain.

Even drugs within the same class are distinguished from one another by often subtle differences in molecular structure, types of interactions with neurotransmitter systems, differences in pharmacokinetics, the presence or absence of active metabolites, and protein binding. These differences, combined with the biochemistry of the patient, account for the profile of efficacy, tolerability, and safety and the risk–benefit ratio for the individual. These multiple variables, some poorly understood, make it difficult to predict a drug's effect with certainty. Nevertheless, knowledge of the nature of each property increases the likelihood of successful treatment. The clinical effects of drugs are best understood in terms of

pharmacokinetics, which describes what the body does to a drug, and pharmacodynamics, which describes what the drug does to the body.

Pharmacokinetics and pharmacodynamics need to be seen in the context of the underlying variability among patients with respect to how drug effects are expressed clinically. Patients differ in their therapeutic response to a drug and the experience of side effects. It is increasingly clear that these differences have a strong genetic basis. Pharmacogenetics research is attempting to identify the role of genetics in drug response.

DRUG FACTORS

Pharmacodynamics

The time course and intensity of a drug's effects are referred to as its *pharmacodynamics*. Major pharmacodynamic considerations include receptor mechanisms, the dose–response curve, the therapeutic index, and the development of tolerance, dependence, and withdrawal phenomena. Drug mechanism of action is subsumed under pharmacodynamics. The clinical response to a drug, including adverse reactions, results from an interaction between that drug and a patient's susceptibility to those actions. Pharmacogenetic studies are beginning to identify genetic polymorphisms linked to individual differences in treatment response and sensitivity to side effects.

Mechanisms

The mechanisms through which most psychotropic drugs produce their therapeutic effects remain poorly understood. Standard explanations focus on ways that drugs alter synaptic concentrations of dopamine, serotonin, norepinephrine, histamine, GABA, or acetylcholine. These changes are said to result from receptor antagonists or agonists, interference with neurotransmitter reuptake, enhancement of neurotransmitter release, or inhibition of enzymes. Specific drugs are associated with permutations or combinations of these actions. Thus, a drug can be an agonist for a receptor, thus stimulating the specific biological activity of the receptor, or an antagonist, thus inhibiting the biological activity. Some drugs are partial agonists, in that they are not capable of fully activating a specific receptor. Partial agonists may also act as a competitive antagonist in the presence of full agonist. Some psychotropic drugs also produce clinical effects through mechanisms other than receptor interactions. For example, lithium may act by directly inhibiting the enzyme inositol-1-phosphatase. Some effects are closely linked to a specific synaptic effect. For example, most medications that treat psychosis share the ability to block the dopamine type 2 (D_2) receptor. Similarly, benzodiazepine agonists bind a receptor complex that contains benzodiazepine and GABA receptors.

It is important to note that action mechanisms for most psychotropic

drugs are poorly understood and that even the current classification of drugs (e.g., SSRI and 5HT2A antagonist) is problematic. There are two reasons for this. Studies of binding properties of many drugs have revealed complex patterns of receptor agonism and antagonism. For example, olanzapine has 5HT2 antagonist properties, but it also is an antagonist for 5HT 1B, 1D, 2A, 2C; α1 adrenergic; H1 histamine and D2 dopamine receptors while acting as a partial agonist for dopamine D1 and muscarinic AChR m1 receptors. This polyspecificity has been extensively by Roth and colleagues. The second issue is the assignation of mechanism based upon a single target. For example, while SSRIs do, indeed, inhibit 5HT transport, they do so within hours after being administered, while the clinical effects of these compounds require weeks to develop. Furthermore, at least some biochemical properties have been attributed to multiple classes of antidepressant drugs. SSRIs can elicit these in cells lacking 5HT transporters.

Thus, accounts of so-called mechanisms of action should nevertheless be kept in perspective. Explanations of how psychotropic drugs actually work that focus on synaptic elements represent an oversimplification of a complex series of events. If merely raising or lowering levels of neurotransmitter activity is associated with the clinical effects of a drug, then all drugs that cause these changes should produce equivalent benefits. This is not the case. Multiple obscure actions, several steps removed from events at neuronal receptor sites, may well be responsible for the therapeutic effects of psychotropic drugs. These *downstream* elements are postulated to represent the actual reasons that these drugs produce clinical improvement.

Pharmacokinetics

Pharmacokinetic drug interactions are the effects of drugs on the plasma concentrations of each other, and pharmacodynamic drug interactions are the effects of drugs on the biological activities of each other. Pharmacokinetic concepts are used to describe and to predict the time course of drug concentrations in different parts of the body, such as plasma, adipose tissue, and the central nervous system (CNS). From a clinical perspective, pharmacokinetic methods help explain or predict the onset and duration of drug activity and interactions between drugs that alter their absorption, distribution, metabolism or excretion.

Pharmacogenetic research focuses on finding variant alleles that alter drug pharmacokinetics and pharmacodynamics. Researchers are attempting to identify genetic differences in how enzymes metabolize psychotropics, as well as CNS proteins directly involved in drug action. It is likely that the identification of patient genotypes will facilitate the prediction of clinical response to different types of drugs.

Drugs differ in their propensity to interact with other drugs. Awareness

of potential drug–drug interactions has grown dramatically over the last two decades. Most clinicians need to consult charts or online resources to determine when potential interactions may occur and, if so, how clinically relevant they may be. Whenever possible, it is preferable to use a medication that produces minimal risk of drug interactions. In addition, it is recommended that prescribers know the interaction profiles of the drugs that they most commonly prescribe.

Examples of pharmacokinetic interactions include one drug increasing or decreasing the concentrations of a coadministered compound. These types of interactions can also lead to altered concentrations of metabolites. In some cases, there also may be interference with the conversion of a drug to its active metabolite. There is enormous variability among patients with respect to pharmacokinetic parameters, such as drug absorption and metabolism. Another type of interaction is represented by interactions involving the kidney. Commonly used medications, such as angiotensin-converting enzyme (ACE) inhibitors, nonsteroidal anti-inflammatory drugs (NSAIDs), and thiazides decrease the renal clearance of lithium, increasing the likelihood of its severe elevation. Drug interactions may be pharmacokinetic or pharmacodynamic.

Pharmacogenetics is being used to study why patients differ in the way that they metabolize drugs. In patients who are ultrarapid or extensive metabolizers, the concentrations of a drug may be lower than expected. There are now commercially available assays that determine the potential of individual subjects for metabolizing various psychiatric drugs, in an attempt to create a profile of agents to be used and those to be avoided. Unfortunately, the actual therapeutic efficacy of a compound may not be dictated simply by the ability of the patient to process it, since an individual's profile of receptors, transporters and intracellular signaling cascades (described above) is likely essential to the action of a given agent.

PATIENT-RELATED FACTORS

Response to medication, as well sensitivity to side effects, is influenced by factors related to the patient. This is why there is no one-size-fits-all approach to pharmacological treatment. This is reflected in the finding that, although drugs used to treat a specific disorder demonstrate comparable efficacy in head-to-head studies, these drugs are not equally effective in all patients. Patient-related variables include diagnosis, genetic factors, lifestyle, overall medical status, concurrent disorders, and past history of drug response. A patient's attitude toward medication in general, aversion to certain types of side effects, and preference for a specific agent also need to be considered.

Diagnosis

Accurate diagnosis is an integral component of the treatment process. Failure to correctly diagnose a disorder diminishes the likelihood of optimal drug selection. Misdiagnosis not only can result in a missed opportunity, but at times it also can produce a worsening of symptoms. Inadvertently diagnosing and treating a patient in the depressed phase of bipolar disorder as having unipolar depression can result in the induction of mania or rapid cycling. Treatment failure or exacerbation of symptoms should prompt a reassessment of the working diagnosis.

Past Treatment Response

A specific drug should be selected according to the patient's history of drug response (compliance, therapeutic response, and adverse effects), the patient's family history of drug response, the profile of adverse effects for that drug with regard to the particular patient, and the prescribing clinician's usual practice. If a drug has previously been effective in treating a patient or a family member, the same drug should be used again. While this makes intuitive clinical sense, the evidence base for this has not been well-established, particularly for family history of treatment response. One study by Fava and colleagues demonstrated the response to a second trial of fluoxetine after placebo was similar to an initial trial with fluoxetine, suggesting that past performance predicts future results. Current studies like the PReDICT trial, examining predictors of depression remission, will be key in determining relevant patient-related factors that predict treatment response. A history of severe adverse effects from a specific drug is a strong indicator that the patient would not be compliant with that particular drug.

It is helpful if patients can recall the details of past psychotropic drug treatment: The drugs prescribed, in what doses, for how long, and in what combinations. Often while patients may have tried a medication, they may not have had an adequate trial, either by dose or duration. Unfortunately, because of their mental disorders, many patients are poor historians. If possible, patients' medical records should be obtained to confirm their reports. Family members are a good source of collateral information.

Concurrent Medical or Psychiatric Disorders

Initial assessment should elicit information about coexisting medical disorders. In some cases, there may be a medical disorder that is responsible for the symptoms. Patients with thyroid disease who are not adequately treated may appear depressed. Sleep apnea produces depression and cognitive impairment. Rare conditions, such as Kleine–Levin syndrome, can mimic bipolar disorder. A drug should be selected that minimally exacerbates any preexisting medical problems that a particular patient may have.

Recreational drug use, excessive consumption of alcohol, and frequent ingestion of caffeine-containing beverages can complicate and even undermine psychotropic drug treatment. These compounds possess significant psychoactive properties and, in some cases, may represent the source of the patient's symptoms. It is reasonable to ask patients to abstain from use of these substances at least until the benefits of psychotropic drug treatment have been unequivocally established. Gradual reintroduction of moderate amounts of alcohol, tea, and coffee can then take place. Patients can then observe whether there are any untoward effects on their clinical status.

Drug Preference

Some prescribers have obvious favorites among the available psychotropic agents. This bias or *allegiance effect* regarding a treatment can stem from a combination of factors. A dramatically successful patient response during early experience with a new drug can produce a *conversion experience,* in which the prescribing physician forms a highly positive opinion of that drug. Patients being treated by that physician are more likely to receive that medication than alternatives. Previous studies have cited physician experience and prior treatment history as among the most highly ranked factors in treatment selection.

INFORMED CONSENT AND PATIENT EDUCATION

Establishing trust and providing motivation to comply with the medication regimen are essential components of successful treatment. Patients should be informed about treatment options and the probable side effects and unique benefits of each treatment. Patient preference should be respected, unless there is a compelling advantage involving efficacy, tolerability, or safety with an alternative agent. If a particular medication is being recommended, the reasons for this recommendation should be explained. Patients are more likely to continue taking their medication if they fully understand the reasons why it is being prescribed.

Physicians no longer have a monopoly over the knowledge regarding psychotropic drugs. Many patients now routinely search the Internet for information about available treatments. What physicians do have is an understanding of the clinical circumstances that would lead to the choice of one drug among the many that are available.

Medication treatment benefits from a strong therapeutic alliance between a clinician and a patient. Given the unpredictability of medication response, the frequent occurrence of side effects, and underlying ambivalence about or fear of taking medication, a positive, trusting relationship serves to improve patient compliance. Repeated failed trials may be needed before a response is seen. A patient's confidence in the

physician's knowledge and judgment enables medication trials and more complex regimens, such as the use of multiple medications.

Discussions about drug selection should be documented in notes. In particular, patients who are informed of potential adverse effects report a higher incidence of side effects but do not have higher rates of premature discontinuation.

How the patient and family are engaged in the treatment plan can determine the success of treatment. With the patient's consent, relatives and other clinicians should be instructed on the reasons for the drug treatment, as well as the expected benefits and potential risks. In addition, the clinician may also find it useful to explain the theoretical basis for pharmacotherapy to the patient and other significant parties.

DOSING, DURATION, AND MONITORING

Dosing

The clinically effective dose for treatment depends on the characteristics of the drug and patient factors, such as inherited sensitivity to and ability to metabolize a drug, concurrent medical disorders, use of concurrent medication treatments, and history of exposure to previous medications.

Doses used in clinical practice often differ from those used in pivotal studies, which, in turn, form the basis for dosing recommendations that eventually appear in the product information. The reason for this divergence is that patients in pivotal trials are highly selected and are not representative of the people who are ultimately treated. Once a drug achieves widespread use in real-world patients, it might be discovered that the doses needed for optimal efficacy need to be higher than suggested by clinical trial findings. Similarly, the doses needed for tolerability might be lower than thought. Even when a disparity on dosing between actual practice and clinical trial is established, the package insert is rarely, if ever, changed. This would involve the conduct of costly and time-consuming studies and submission of an application for change. At times, it may be clinically beneficial to exceed the upper dose limit for a drug—so-called high-dose therapy. However, it is best not to exceed the recommended upper dose threshold without a clear understanding of the potential benefits and risks involved. This caution is particularly important when a drug is associated with a severe concentration-dependent side effect.

Plasma concentrations of many psychotropics can vary up to 10-fold. Thus, to some extent, the optimal dose for an individual is ultimately determined by trial and error, guided by the empirical evidence of the usual dose range for that drug. Some drugs demonstrate a clear relationship between increases in dose and clinical response. This dose–response curve plots the drug concentration against the effects of the drug.

The *potency* of a drug refers to the relative dose required to achieve

certain effects, not to its efficacy. Haloperidol, for example, is more potent than chlorpromazine because approximately 5 mg of haloperidol is required to achieve the same therapeutic effect as 100 mg of chlorpromazine. However, these drugs are equal in their clinical efficacy—that is, the maximum clinical response achievable by administration of a drug.

Drugs must be used in effective doses for sufficient periods. Although drug tolerability and safety are always a consideration, subtherapeutic doses and incomplete therapeutic trials should be avoided. The use of inadequate doses merely exposes the patient to the risk of side effects without providing the probability of therapeutic benefit. In view of the wide margin of safety associated with most currently prescribed medications, there is more risk in underdosing than in overshooting the recommended dose range.

Time of dosing is usually based on the plasma half-life of a drug and its side effect profile. Sedating drugs are given all at night or with disproportionate daily doses at night. The opposite is true with activating drugs. The frequency of dosing is less clear-cut. Most dosing regimens of psychotropic drugs, such as once-a-day versus divided doses, are based on measurements of plasma concentrations rather than receptor occupancy in the brain. Evidence suggests that there is a significant dissociation between brain and plasma kinetics. Reliance on plasma kinetics as the basis for dosing regimens leads to misunderstanding of necessary schedules.

Duration of Treatment

A common question that a patient typically asks is "How long do I need to take the medication?" The answer, of course, depends on multiple variables, including the nature of the disorder, the duration of symptoms, the family history, and the extent to which the patient tolerates and benefits from the medication. Patients can be given a reasonable explanation of the probabilities but should be told that it is first best to see whether the medication works for him or her and whether the side effects are acceptable. Any more definitive discussion of duration of treatment can be held once the degree of success is clear. Even patients with a philosophical aversion to the use of psychotropic drugs may elect to stay on medication indefinitely if the magnitude of improvement is great. Most psychiatric disorders have high rates of chronicity and relapse. Because of this, long-term treatment is often needed to prevent recurrence. Nevertheless, the fact remains that psychiatric treatments are not said to cure the disorders they treat, but rather to help control the symptoms.

Treatment is conceptually broken down into three phases: The initial therapeutic trial, the continuation, and the maintenance phase. The initial period of treatment should last at least several weeks because of the delay in therapeutic effects that characterizes most classes of psychotropic drugs.

The required duration of a *therapeutic trial* of a drug should be discussed at the outset of treatment, so that the patient does not have unrealistic expectations of an immediate improvement in symptoms. Patients are more likely to experience side effects early in the course of pharmacotherapy than any relief from their disorder. In some cases, medication may even exacerbate some symptoms. Patients should be counseled that a poor initial reaction to medication is not an indicator of the ultimate outcome of treatment. For instance, many patients with panic disorder develop jitteriness or an increase in panic attacks after starting on tricyclic or SSRI treatment. Benzodiazepine agonists are an exception to the rule that there is a delay in clinical onset. In most cases, their hypnotic and antianxiety effects are evident immediately.

Ongoing use of medication does not, however, provide absolute protection against relapse. Continuation therapy provides clinically and statistically significant protective effects against relapse. The optimal duration of continuation or maintenance therapy is variable and depends on the clinical history of the patient. Early-onset chronic major depression, for example, has a more severe course and greater comorbidity than late-onset chronic major depression. In addition to early onset, a history of multiple past episodes, severity of the current episodes, and length of current episode would make longer, even indefinite, treatment appropriate.

TREATMENT OUTCOMES

The goal of psychotropic treatment is to eliminate all manifestations of disorder, thus enabling the patient to regain the ability to function and to enjoy life as fully as before he or she became ill. This degree of improvement below the syndromal threshold is defined as *remission*.

Response and Remission

Remission is the preferred outcome of treatment not only because of the immediate effect on functioning and state of mind, but also because emerging evidence suggests that remitted patients are less likely to experience relapse and recurrence of their disorder.

Patients who improve but do not experience a full resolution are considered to be responders. They may exhibit significant improvement but continue to experience symptoms. In depression studies, *response* is usually defined as a 50 percent or greater decrease from baseline on a standard rating scale, such as the Hamilton Depression (HAMD) Scale or the Montgomery-Asberg Depression Rating Scale (MADRS). *Remission* is defined as an absolute score of 7 or less on the HAMD Scale or 10 or less on the MADRS. Expectations about the likely degree of improvement should be based on what is known about the responsiveness of specific disorders to medication therapy. OCD and schizophrenia, for example, are more likely

to be associated with residual manifestations of illness than major depression or panic disorder. The probability of full remission from OCD with SSRI treatment alone over a 3-year period is 38 percent, and the probability of partial remission is approximately 65 percent.

Treatment Failure

The initial treatment plan should anticipate the possibility that the medication may be ineffective. A next-step strategy should be in place at the initiation of treatment. Repeated drug failures should prompt reassessment of the patient. First, was the original diagnosis correct? In answering this question, the clinician should include the possibility of an undiagnosed medical condition or recreational drug use as the cause of the psychiatric symptoms. Second, are the observed symptoms related to the original disorder or are they actually adverse effects of the drug treatment? Some antipsychotic drugs, for example, can produce akinesia, which resembles psychotic withdrawal, or akathisia and neuroleptic malignant syndrome, which resemble increased psychotic agitation. Long-term use of SSRIs can produce emotional blunting, which can mimic depression.

Intolerance of side effects may be the most common reason for treatment failure. Third, was the drug administered at an appropriate dose for a sufficient length of time? Because patients can vary greatly in their absorption and metabolism of drugs, the clinician may need to measure plasma levels of a drug to ensure a sufficient dose of the drug. Fourth, did a pharmacokinetic or pharmacodynamic interaction with another drug that the patient was taking reduce the efficacy of the newly prescribed drug? Fifth, did the patient take the drug as directed? Drug noncompliance is a common clinical problem that arises as a result of complicated drug regimens (more than one drug in more than one daily dose), adverse effects (especially if unnoticed by the clinician), and poor patient education about the drug treatment plan. Patients may discontinue medication when they recover, thinking that they are cured and no longer benefiting from the medication.

Treatment Resistance

Some patients fail to respond to repeated trials of medication. No single factor can explain the ineffectiveness of the various interventions in these cases. Strategies in these cases include the use of drug combinations, high-dose therapy, and the use of unconventional drugs. There is limited evidence on the comparative success rates associated with any given strategy. This is one area that illustrates the critical need to develop biomarkers to determine trait, state, and therapeutic response. Often patients are given extended (2 months) drug trials before a decision is made to revise therapy. Thus, it may take 8 months to declare a patient

"treatment-resistant." As indicated above, it is not known whether a given patient is inherently treatment resistant or whether treatment resistance evolves as a result of repeated failures.

Tolerance

The development of tolerance is marked by a need, over time, to use increased doses of a drug to maintain a clinical effect. This decreased responsiveness to a drug occurs after repeated doses. Tolerance also describes decreased sensitivity to adverse effects of the drug, such as nausea. This phenomenon is used as the basis for starting some drugs at subtherapeutic doses, with the plan to adjust the schedule once the patient can tolerate higher doses. Clinical tolerance appears to represent changes in the CNS, such as altered receptor configuration or density. Drugs with similar pharmacological actions often exhibit cross-tolerance.

Sensitization

Clinically manifested as the reverse of tolerance, sensitization is said to occur when there is increased sensitivity to a drug effect over time. In these cases, the same dose typically produces more-pronounced effects as treatment progresses. Conversely, if the therapeutic index is low, as it is for lithium, careful monitoring of serum levels is essential. For most psychotropic drugs, the therapeutic index is high.

Withdrawal

The development of physiological adaptation to a drug, with a subsequent risk of withdrawal symptoms, has been reported for many classes of psychotropic drugs. Technically, withdrawal should be considered a side effect. The probability and severity of these reactions are remote with most drugs and more common with others. As a general rule, the more abruptly a drug is stopped and the shorter its elimination half-life, the more likely it is that clinically significant withdrawal symptoms will occur. With some short-acting drugs, withdrawal reactions can result from missed doses and during daily intervals between doses. Gradual tapering of medications after prolonged use is recommended when possible. Although this reduces the risk of withdrawal reactions, it does not ensure that they will not occur. So-called sedative–hypnotics and opiates are the agents most often associated with mentally and physically distressing discontinuation reactions. In some cases, such as the barbiturates, withdrawal can be fatal.

There are marked differences among agents, even within a given class, with respect to the probability and severity of discontinuation effects. For example, among the benzodiazepines, alprazolam commonly produces more-immediate and intense withdrawal symptoms than other compounds. Among the SSRIs, there is a well-described withdrawal

syndrome that appears to be more frequent and severe with paroxetine (Paxil). It can, however, occur with any SSRI. Even fluoxetine may be associated with discontinuation symptoms, but the symptoms may be delayed and attenuated due to the long elimination half-life of its active metabolite. These manifestations are subtle and are delayed for weeks after the last dose. Venlafaxine also produces a severe SSRI-like withdrawal syndrome.

In addition to half-life, many variables can influence the likelihood and degree of discontinuation symptoms. Changes in the rate of drug metabolism, for example, can play a role. Paroxetine is primarily metabolized by the cytochrome P450 (CYP) 2D6 isozyme. However, paroxetine is also a potent inhibitor of CYP 2D6. This results in *autoinhibition,* a dose-dependent inhibition of its own metabolism, with a subsequent increase in plasma concentrations of paroxetine. If the dose of paroxetine is decreased or the drug is stopped, the decline in its plasma concentrations can be steep, causing withdrawal to occur. Withdrawal can occur in rare cases in which the dose of a drug is not decreased but a second agent that had been inhibiting its metabolism is stopped. For example, paroxetine is metabolized via the CYP 2D6 enzyme system. Bupropion inhibits that enzyme. If a patient taking both agents for several weeks discontinues the bupropion, it could result in a rapid increase in the rate of paroxetine metabolism and a consequent drop in plasma concentrations.

The development of sustained-release versions of drugs such as alprazolam, paroxetine, and venlafaxine has not reduced the severity of their withdrawal reactions. The prolonged half-life of those agents results from delayed absorption rather than prolongation of the elimination phase. The frequency of drug dosing is reduced but not the rate of fall-off in plasma concentrations.

Poor bioavailability with a generic agent may account for an unexpected loss of clinical effect in the emergence of withdrawal symptoms. The occurrence of these events soon after refilling a prescription should prompt examination of the new medication. It should be confirmed whether the dispensed medication and dose are both correct. It is difficult to ascertain whether generic medications are truly equivalent, so the possibility exists that differences in potency may underlie adverse changes in clinical status.

Withdrawal symptoms invariably occur hours or days after dose reduction or discontinuation. Symptoms resolve within a few weeks, so the persistence of symptoms argues against withdrawal. Although depletion studies have been shown to provoke a rapid return of symptoms, in clinical practice, psychotic and mood symptoms do not usually reappear abruptly after long-term treatment.

COMBINATION OF DRUGS

According to the American Psychiatric Association Practice Guidelines for the Treatment of Psychiatric Disorders, “the use of multiple agents should be avoided if possible” in the treatment of psychiatric disorders. Although *monotherapy* represents the ideal, *polypharmacy*—the simultaneous use of psychotropic medications—has been commonplace since chlorpromazine was combined with reserpine in the early 1950s. The practice of combining drugs is commonplace, and the merits of various *augmentation* or *combination* strategies are routinely discussed in the literature and at scientific meetings. The mean number of simultaneously prescribed medications has increased in recent decades. In response to some of these trends, The Joint Commission has recently instituted quality measures to reduce the inappropriate use of antipsychotic polypharmacy.

Sometimes distinctions are made between augmentation and combination therapy. When two psychotropics with the same approved indications are used concurrently, this is termed *combination therapy*. Adding a drug with another indication is termed *augmentation*. Augmentation often entails use of a drug that is not primarily considered a psychotropic. For example, in treating depression, it is common to add thyroid hormone to an approved antidepressant. Some strategies are more popular or are preferred but are not necessarily more effective than any other. Augmentation was built into the treatment algorithm for the STAR*D trial, where there was option to add either lithium or thyroid hormone. In a STAR*D study that examined these augmentation options, the authors found that lithium and thyroid for equally effective in improving response rates in patients that failed monotherapy, but that thyroid hormone was better tolerated.

The widespread reliance on multiple-drug therapy, in the face of official discouragement and the absence of objective evidence supporting this practice, is easy to explain: Many patients do not become well with monotherapy with existing drugs. In a typical scenario, a patient has little or no response to a medication, so the physician adds a second agent to induce a better response. This need not be a second drug, as some studies suggest improved efficacy for SSRIs with the addition of omega 3 polyunsaturated fatty acids (fish oil).

In some cases, the use of multiple medications is the rule. Almost all patients with bipolar disorder are taking more than one psychotropic agent. Combination treatment with drugs that treat depression and a dopamine-receptor antagonist or serotonin-dopamine antagonist has long been held as preferable in patients with psychotic depression. Similarly, SSRIs typically produce partial improvement in patients with OCD, so that the addition of a serotonin-dopamine antagonist may be helpful.

Medications may be also be combined to counteract side effects, to treat specific symptoms, and as a temporary measure to transition from one drug to another. It is common practice to add a new medication without the

discontinuation of a prior drug, particularly when the first drug has provided partial benefit. This can be done as part of a plan to transition from an agent that is not producing a satisfactory response or as an attempt to maintain the patient on combined therapy.

One limitation is that there is increased noncompliance and adverse effects and that the clinician may not be able to determine whether it was the second drug alone or the combination of drugs that resulted in a therapeutic success or a particular adverse effect. Combining drugs can create a broad-spectrum effect and also changes the ratio of metabolites.

The merits of going to a single drug with a different pharmacological profile include lower risk of drug–drug interactions, simplicity, and lower cost. It is less burdensome to be taking one medication than two or three and is less likely to meet resistance by the patient. Many patients are ambivalent about taking even one medication, let alone two. This practice is less justifiable in newly treated patients than in those who have failed repeated trials.

SPECIAL POPULATIONS

Although every patient brings a unique combination of demographic and clinical variables to the clinical setting, certain patient populations require special consideration. When treating the young, the elderly and those with medical disorders, awareness of risks associated with medication assumes increased importance. Data derived from clinical trials are of limited value in guiding many decisions because populations in these studies consisted of healthy young adults. Studies of children and adolescents have become more common, so that understanding of treatment effects in this population has grown.

Children

Understanding of the safety and efficacy of most psychotropic drugs when used to treat children is based more on clinical experience than on evidence from large clinical trial data. Other than attention-deficit/hyperactivity disorder (ADHD) and OCD, commonly used psychotropic drugs have no labeling for pediatric use, so results from adult studies are extrapolated to children. This is not necessarily appropriate because of developmental differences in pharmacokinetics and pharmacodynamics. Dosing is another special consideration in drug use with children. Although the small volume of distribution suggests the use of lower doses than those used in adults, a child's higher rate of metabolism suggests that a higher ratio of milligrams of drug to kilograms of body weight should be used. In practice, it is best to begin with a small dose and to increase it until clinical effects are observed. However, the clinician should not hesitate to use adult doses in children if these doses are effective and the adverse effects are acceptable.

The paucity of research data is a legacy of many years in which manufacturers avoided conducting trials in children because of liability concerns, small market share, and, hence, limited profit potential represented by this population. To correct this problem, the FDA Modernization Act (FDAMA) of 1997 provided for special encouragement and incentives to study drugs for pediatric use. Since then, there have been expanded indications for FDA-approved medications to treat psychiatric conditions in children, such as risperidone for the treatment of schizophrenia in adolescents and manic/mixed episodes of bipolar I disorder in children and adolescents ages 10 to 17.

Some drugs used to treat psychiatric disorders carry a higher risk of serious side effects. Lamotrigine, for example, is associated with a substantially higher incidence of serious rash among children than among adults.

Elderly Patients

Elderly persons may be more susceptible to adverse effects (particularly cardiac effects) and may metabolize and excrete drugs more slowly, thus requiring lower doses of medication. In practice, clinicians should begin treating geriatric patients with a small dose—usually approximately one half of the usual starting dose. The dose should be raised in small increments, more slowly than for middle-aged adults, until a clinical benefit is achieved or unacceptable adverse effects appear. Although most geriatric patients require a small dose of medication, there are others who require a full therapeutic dose. Thus, sound clinical judgment and close monitoring are necessary.

A major concern about psychotropic drugs in the elderly is the well-established causal relationship between the use of these agents and falls. It has been shown that elderly users of psychotropic medications have a 47 percent increased risk of falls. This risk of falls and hip fractures may be minimized by selecting compounds that produce less unwanted sedation, dizziness, parkinsonism, and postural hypotension.

Historically, antipsychotics have been used to treat nonpsychotic, agitated patients. This practice has been called into question primarily as a result of the CATIE-AD trial that showed limited benefit of atypical antipsychotics in the treatment of agitation in dementia. Although they may reduce paranoia, there was no significant benefit in improving function or quality of life and there is evidence of worsening cognitive outcomes with these medications. There is an FDA warning regarding the use of atypical antipsychotics in the treatment of behavioral disorders in elderly patients with dementia. Studies have shown a higher death rate associated with their use. Accordingly, antipsychotic drugs should be reserved for psychosis.

Response to treatment is less predictable among older adults than

younger adults. Age-related changes in renal clearance and hepatic metabolism make it more important to be conservative with the starting doses of medication, as well as with the rate of dose titration. The frequent presence of medical disorders and the use of medications to treat them make it necessary to recognize how psychotropic drugs interact with the nonpsychiatric disorders and medications. Within any class of agents, those with fewer drug interactions or side effects with potentially serious consequences, such as hypotension, cardiac conduction abnormalities, anticholinergic activity, and respiratory depression, are not suitable choices. Drugs that cause cognitive impairment, such as benzodiazepines and anticholinergics, may mimic or exacerbate symptoms of dementia. Similarly, dopamine-receptor antagonists can worsen or induce Parkinson disease, another age-related disorder. Some side effects, such as SSRI-associated syndrome of inappropriate secretion of antidiuretic hormone (SIADH) and oxcarbazepine-associated hyponatremia, occur more commonly in older patients. A common ethical dilemma with the medically ill elderly or those with dementia is the question of their capacity to give informed consent before treatment with psychotropic drugs or ECT.

Medically Ill Patients

When selecting and monitoring psychiatric drugs, the patient's medical disorders and their treatment need to be considered. Most important, the medical disorder or disorders and their treatments should be ruled out as a cause of the observed psychiatric symptoms. For example, patients with neurological or endocrine disorders or human immunodeficiency virus (HIV)–infected patients may experience disturbances of mood and cognition. Common medications, such as corticosteroids and L-dopa, are associated with induction of mania.

Although obvious, when various options are available, a psychotropic should be chosen that does not exacerbate or counteract the treatment of the existing medical disorder. For example, a psychotic patient with diabetes mellitus is better treated with an agent without the risk of weight gain or glucose dysregulation. Depending on the diagnosis, drugs that might treat the primary psychiatric disorder and also cause weight loss, drugs such as bupropion and topiramate (Topamax) can be prescribed for these patients. Patients with obstructive pulmonary disease should not be given sedating drugs, which raise the arousal threshold and suppress respiration. Patients with medical disorders are also taking other medications, which can result in pharmacodynamic and pharmacokinetic interactions. Combined treatment with an inducer of multiple CYP enzymes and a drug that is a substrate for those enzymes could result in subtherapeutic levels, leading to inadequate symptom control. Use of the tuberculosis treatment rifampicin (Rifadin) with carbamazepine is an example. Use of drugs that inhibit CYP 2D6—agents such as paroxetine and

fluoxetine—can prevent the conversion of hydrocodone (Robidone) and other opiates into an active analgesic form.

NSAIDs are also a rare cause of perceptual disturbances and psychotic symptoms. Other issues include a potentially increased sensitivity to adverse effects, including increased or decreased metabolism and excretion of the drug, and interactions with other medications. Drug interactions are an obvious concern when drugs with a narrow therapeutic range are being used. Any change in the rate of metabolism or interference with the formation and elimination of metabolites can profoundly influence the activity of that drug. Similarly, interactions that interfere with drug metabolism can produce an increase in side effects and toxicity.

As with children and geriatric patients, the most reasonable clinical practice is to begin with a small dose, increase it slowly, and watch for clinical benefit and adverse effects. Determining the plasma drug concentrations may be helpful for such patients, but therapeutic blood concentrations for most psychotropic drugs are neither necessary nor routinely available. Special caution is needed with potential drug–disease interactions. In patients with diabetes mellitus, drugs such as valproic acid or olanzapine, which are associated with increasing body weight, and insulin resistance, should be avoided. Patients with seizure disorder should not receive bupropion, or clomipramine (Anafranil), which lower the seizure threshold.

Substance Abuse

Many patients who seek or need treatment of a psychiatric disorder engage in chronic use of illicit substances or drink excessive amounts of alcohol. Marijuana is the most commonly used illicit drug in the United States.

Discontinuation of chronic drug or alcohol use can result not only in craving but also in clinically significant psychiatric and physiological withdrawal symptoms. For many patients, successful treatment of their underlying psychiatric disorder may not be possible in the presence of ongoing marijuana, cocaine, and alcohol use. If several trials of medications fail, hospitalization for detoxification may be necessary. There is little research and no consensus about how to use psychotropic agents in patients who are regular users of cocaine, marijuana, or other recreational drugs. If these agents are used as a method of coping—so-called avoidance coping—it appears to not be effective.

FUTURE DIRECTIONS

After decades of progress in bringing new treatments to the clinical setting, there has been a marked decline in the development or discovery of truly innovative drugs. Although researchers continue to search for medications with new modes of action and that are both more effective and better

tolerated than existing treatments, there needs to be a better understanding of the underlying neurochemical abnormalities that cause psychiatric disorder. Researchers may be close to significant breakthroughs by which research findings can be translated into better treatments. It is also possible that it will be years before new options are available to clinicians. There is a need to be able to predict more reliably which individuals will respond to specific agents because it is clear that there is no "one-size-fits-all" approach to psychopharmacology. There also needs to be more systematic postmarketing surveillance of drugs for potentially serious side effects. In the meantime, it is imperative that clinicians remain current with the clinical literature and remain thoughtful about the drug choices that they make.

References

Allen JA, Halverson-Tamboli RA, Rasenick MM. Lipid raft microdomains and neurotransmitter signalling. *Nat Rev Neurosci*. 2007;8(2):128–140.

Berwian IM, Walter H, Seifritz E, Huys QJ. Predicting relapse after antidepressant withdrawal—a systematic review. *Psychol Med*. 2016:1–12.

Bogetto F, Bellino S, Revello RB, Patria L. Discontinuation syndrome in dysthymic patients treated with selective serotonin reuptake inhibitors: a clinical investigation. *CNS Drugs*. 2002;16(4):273–283.

Bull SA, Hunkeler EM, Lee JY, et al. Discontinuing or switching selective serotonin-reuptake inhibitors. *Ann Pharmacother*. 2002;36(4):578–584.

Burchett SA, Volk ML, Bannon MJ, Granneman JG. Regulators of G protein signaling: rapid changes in mRNA abundance in response to amphetamine. *J Neurochem*. 1998;70(5):2216–2219.

Catapano F, Perris F, Masella M, et al. Obsessive-compulsive disorder: a 3-year prospective follow-up study of patients treated with serotonin reuptakeinhibitors OCD follow-up study. *J Psychiatr Res*.2006;40(6):502–510.

Cocchi M, Tonello L, Rasenick MM. Human depression: a new approach in quantitative psychiatry. *Ann Gen Psychiatry*. 2010;9:25.

Costa T, Herz A. Antagonists with negative intrinsic activity at delta opioid receptors coupled to GTP-binding proteins. *P Natl Acad Sci U S A*. 1989;86(19):7321–7325.

Czysz AH, Schappi JM, Rasenick MM. Lateral diffusion of gαs in the plasma membrane is decreased after chronic but not acute antidepressant treatment: role of lipid raft and non-raft membrane microdomains. *Neuropsychopharmacology*. 2015;40(3):766–773.

Donati RJ, Dwivedi Y, Roberts RC, Conley RR, Pandey GN, Rasenick MM. Postmortem brain tissue of depressed suicides reveals increased Gs alpha localization in lipid raft domains where it is less likely to activate adenylyl cyclase. *J Neurosci*. 2008;28(12):3042–3050.

Eisensamer B, Uhr M, Meyr S, et al. Antidepressants and antipsychotic drugs colocalize with 5-HT3 receptors in raft-like domains. *J Neurosci*. 2005;25(44):10198–10206.

Fuhrer C, Gautam M, Sugiyama JE, Hall ZW. Roles of rapsyn and agrin in interaction of postsynaptic proteins with acetylcholine receptors. *J Neurosci*. 1999;19(15):6405–6416.

Gardner DM, MacKinnon N, Langille DB, Andreou P. A comparison of factors used by physicians and patients in the selection of antidepressant agents. *Psychiatr Serv*. 2007;58(1):34–40.

Gibbons RD, Brown CH, Hur K, et al. Early evidence on the effects of regulators' suicidality warnings on SSRI prescriptions and suicide in children and adolescents. *Am J Psychiatry*. 2007;164(9):1356–1363.

Goodwin FK, Ghaemi SN. The impact of the discovery of lithium on psychiatric thought and practice in the USA and Europe. *Aust N Z J Psychiatry*. 1999;33(Suppl):S54–S64.

Gurevich VV, Gurevich EV. How and why do GPCRs dimerize? *Trends Pharmacol Sci*. 2008;29(5):234–240.

Ha CM, Park D, Kim Y, et al. SNX14 is a bifunctional negative regulator for neuronal 5-HT6 receptor signaling. *J Cell Sci*. 2015;128(9):1848–1861.

Hurst JH, Hooks SB. Regulator of G-protein signaling (RGS) proteins in cancer biology. *Biochem Pharmacol*. 2009;78(10):1289–1297.

Johnell K, Jonasdottir Bergman G, Fastbom J, Danielsson B, Borg N, Salmi P. Psychotropic drugs and the risk of fall injuries, hospitalisations and mortality among older adults. *Int J Geriatr Psychiatry*. 2016.

Kang J, Bowman FD, Mayberg H, Liu H. A depression network of functionally connected regions discovered via multi-attribute canonical correlation graphs. *Neuroimage*. 2016;14:431–441.

Khan A, Khan SR, Leventhal RM, Krishnan KR, Gorman JM. An application of the revised CONSORT standards to FDA summary reports of recently approved antidepressants and antipsychotics. *Biol Psychiatry*. 2002;52(1):62–67.

Kobilka BK. Structural insights into adrenergic receptor function and pharmacology. *Trends in Pharmacol Sci*. 2011;32(4):213–218.

Kroeze WK, Roth BL. Screening the receptorome. *J Psychopharmacol*. 2006;20(4 Suppl):41–46.

Lam RW, Wan DD, Cohen NL, Kennedy SH. Combining antidepressants for treatment-resistant depression: a review. *J Clin Psychiatry*. 2002;63(8):685–693.

Langeberg LK, Scott JD. Signalling scaffolds and local organization of cellular behaviour. *Nat Rev Mol Cell Biol*. 2015;16(4):232–244.

Lieberman JA, Stroup TS, McEvoy JP, et al; Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Investigators. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med*. 2005;353(12):1209–1223.

Meltzer HY. The rise of psychopharmacology and the story of CINP. *J Clin Psychiat*. 2000;61:65–66.

Murray AJ, Tucker SJ, Shewan DA. cAMP-dependent axon guidance is distinctly regulated by Epac and protein kinase A. *J Neurosci*. 2009;29(49):15434–15444.

Nierenberg AA, Fava M, Trivedi MH, et al. A comparison of lithium and T(3) augmentation following two failed medication treatments for depression: a STAR*D report. *Am J Psychiatry*. 2006;163:1519–1530; quiz 1665.

Rasmussen SG, DeVree BT, Zou Y, et al. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature*. 2011;477(7366):549–555.

Reiter E, Ahn S, Shukla AK, Lefkowitz RJ. Molecular mechanism of β-arrestin-biased agonism at seven-transmembrane receptors. *Annu Rev Pharmacol Toxicol*. 2012;52:179–197.

Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed

outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry*. 2006;163(11):1905–1917.

Samama P, Cotecchia S, Costa T, Lefkowitz RJ. A mutation-induced activated state of the beta 2-adrenergic receptor. Extending the ternary complex model. *J Biol Chem*. 1993;268(7):4625–4636.

Sarris J, Murphy J, Mischoulon D, et al. Adjunctive nutraceuticals for depression: a systematic review and meta-analyses. *Am J Psychiatry*. 2016;173(6):575–587.

Shenker A, Laue L, Kosugi S, Merendino JJ Jr, Minegishi T, Cutler GB Jr. A constitutively activating mutation of the luteinizing hormone receptor in familial male precocious puberty. 1993;365(6447):652–654.

Sultzer DL, Davis SM, Tariot PN, et al. Clinical symptom responses to atypical antipsychotic medications in Alzheimer's disease: phase 1 outcomes from the CATIE-AD effectiveness trial. *Am J Psychiatry*. 2008;165(7):844–854.

Swartz MS, Stroup TS, McEvoy JP, et al. What CATIE found: results from the schizophrenia trial. *Psychiat Serv*. 2008;59(5):500–506.

The Joint Commission, Specifications Manual for Joint Commission National Quality Measures (v2016A1). *Measure Set: Hospital Based Inpatient Psychiatric Services (HBIPS), Set Measure ID: HBIPS-5*.

Unwin N. Refined structure of the nicotinic acetylcholine receptor at 4 angstrom resolution. *J Mol Biol*. 2005;346(4):967–989.

Vigen CL, Mack WJ, Keefe RS, et al. Cognitive effects of atypical antipsychotic medications in patients with Alzheimer's disease: outcomes from CATIE-AD. *Am J Psychiatry*. 2011;168(8):831–839.

Viktorin A, Lichtenstein P, Thase ME, et al. The risk of switch to mania in patients with bipolar disorder during treatment with an antidepressant alone and in combination with a mood stabilizer. *Am J Psychiatry*. 2014;171(10):1067–1073.

Zohar J, Nutt DJ, Kupfer DJ, et al. A proposal for an updated neuropsychopharmacological nomenclature. *Eur Neuropsychopharmacol*. 2014;24(7):1005–1014.

▲ 34.2 Drug Development and Approval Process in the United States

AMIR KALALI, M.D., DAVID AKERS HOSFORD, M.D., PH.D., AND SARAH J. RICHERSON, PH.D.

SCIENTIFIC CHALLENGES IN CNS DRUG DEVELOPMENT

Anyone following the news will be aware of challenges faced by the pharmaceutical industry. The financial challenges are driven by many factors including the "patent cliff," where the expiration of the patents on many recent blockbuster treatments has led to severely reduced revenues for many companies resulting in cutbacks in research and development. Increased scrutiny by payors and governments, as to the value and novelty of new treatments, as well as price controls in many countries, has added to these challenges.

The latest figures estimate that the cost to develop a new drug is an average of $2.6 billion. Drugs that treat CNS diseases take more than a year longer to develop and are less than half as likely to obtain marketing approval as other drugs, according to a newly completed study by the Tufts Center for the Study of Drug Development.

In addition to the financial challenges, companies must leverage basic science to develop new and innovative treatments that are more effective or safer than current treatments available as generics. This has proved particularly challenging in psychiatric disorders for several reasons including an incomplete understanding of the basic science, and lack of a diagnostic system based on etiology, leading to a heterogeneous population in clinical studies.

Add to this the methodological challenges in conducting clinical trials in psychiatry, including subjective outcome measures, lack of medication adherence and increasing placebo response, it is not a surprise that this area is seen as high risk and has seen several companies abandon the area. These factors will be further elaborated below.

The good news is that due to the clear unmet patient need, many new companies are entering this area of drug development.

The Basic Science Challenge

In spite of research conducted in both academic and industry settings, including under the auspices of the Decade of the Brain at the end of the 20th century, and subsequent large initiatives involving academia, government, and corporate pharmaceutical entities, the pathophysiological underpinnings of most conditions affecting the CNS are virtually unknown. This lack of understanding of etiology complicates attempts to construct in vitro or in vivo nonclinical models with which to identify or to test therapeutic molecules before taking them into the clinical phases of drug development. In addition, when potential compounds are tested in accepted nonclinical models, the benefits conferred in these models do not predict similar benefits in man. An example would be Alzheimer disease. For more than 40 years, a deficit of brain acetylcholinergic function has been hypothesized to play a role in the condition, and therefore it was rational to test compounds that diminished the metabolism of acetylcholine: acetylcholinesterase inhibitors. Resulting from a global drug development process that tested many compounds with this pharmacological mechanism of action, four acetylcholinesterase inhibitors have been approved by regulatory authorities to treat dementia of the Alzheimer type. This might be viewed as a triumph of disease understanding, except that these agents offer symptomatic treatment of lesser magnitude and shorter duration than was predicted on the basis of the known acetylcholinergic deficit.

The Biomarker Conundrum

A more cautionary tale is the meager success of very expensive and long-duration clinical trials that have tested monoclonal antibodies designed to clear beta-amyloid from the brains of patients with the more advanced stages of Alzheimer disease. Beta-amyloid has been considered by some to be a surrogate biomarker for the Alzheimer disease process because the progressive brain deposition of beta-amyloid plaques has been a well-known pathophysiological hallmark of the disease since its description by Alois Alzheimer more than 100 years ago. Indeed, this lack of translation from bench to bedside is the rule in the field of Alzheimer disease, where more than 99 percent of clinical trials have failed. Clearly more predictive nonclinical models and a better choice of surrogate biomarkers are needed.

THE CLINICAL SCIENCE CHALLENGE: FACTORS IMPACTING SIGNAL DETECTION IN CLINICAL TRIALS

There is no one reason some CNS clinical trials fail. There are multiple factors that occur to one extent or another in any clinical trial. The following section reviews some of the most important of these.

Study Design

Clearly the design of any clinical trial is a key factor in its success or failure. In general, there are some basic trial principles, such as blinding and randomization that are followed in most drug development clinical trials. However, sometimes small design changes can have unintended consequences, which emphasize the importance of thoughtful clinical trial design.

Effect Size

In general, the effect sizes of compounds developed so far for psychiatric disorders are not large, which makes signal detection more challenging. One may argue that if treatments had larger effect sizes, other methodological issues would not matter.

Dose

Animal models for psychiatric disorders are not good at predicting dosing in humans, and there is a history in CNS drug development of the optimal clinical dose of a compound not being established during the development stage. Several compounds have had negative programs only to be later shown to be effective at different doses.

Diagnosis and Appropriate Patient Selection

Translation from bench to bedside is only one aspect of the issues

confronting CNS drug development. Even if clinical research investigators knew a priori that they were about to test the right therapeutic molecules, enrolling the right patients can be just as much of a challenge to successful trial outcomes. The *Diagnostic and Statistical Manual of Mental Disorders* (DSM) provides a standard language by which clinicians, researchers and public health officials in the US communicate about mental health disorders. The DSM Fifth Edition (DSM-5), published in 2013, marked the first significant revision since the DSM-IV in 1994.

Although useful for administrative and clinical reasons, as DSM is not based on biological etiology of psychiatric disorders, the subjects selected for a trial are clearly a heterogeneous group, and it is not surprising that this dilutes drug effects shown. There have been recent attempts to categorize patients with new models of diagnosis such as the NIMH Research Domain Criteria Project (RDoC).

Because there is little understanding of the pathophysiology of CNS disorders, it is not known whether the phenotypically similar features of clinically diagnosed patients may in fact result from a multiplicity of underlying genotypes. Compounds with discrete mechanisms of action may offer benefit to patients with only one or a few, but likely not all of these underlying genetic etiologies. Given this etiological heterogeneity of the population with a clinical disease phenotype, how does the clinical trial investigator choose the right patient who might benefit from a promising new compound? Greater knowledge of the heterogeneous pathophysiological causes and the contributing roles of genetic variants or environmental factors on CNS conditions are necessary.

Another question to ask is whether one treatment can be realistically expected to treat all aspects of a disorder such as depression, schizophrenia, or dementia? However, currently in psychiatry, the labels are granted by regulatory agencies map to the DSM.

Other fields such as oncology have faced these issues earlier and it has been new knowledge about the basic science and etiology that has led to breakthrough success.

Expectation Bias

This effect describes a situation where either the staff conducting the trial or subjects in the trial, or both groups, expect that the test drug will confer a benefit or that it is unlikely to do so. In the case of a positive bias, this expectation creates a cognitive bias that increases the effect that is being measured. Suppose a large clinical trial of a new antidepressant molecule (Drug X) is being conducted. Going into the trial, it is already known that Drug X has a mechanism of action that is similar to other drugs which produce true antidepressant effects. It is also known that in a prior, smaller clinical trial, Drug X produced a measureable effect when compared to placebo. These facts are known by the staff conducting the new and larger

trial, and they are also explained to each subject enrolling in the trial, as a part of informed consent. These facts may create an expectation of drug benefit in the staff who use rating scales to measure the symptoms produced by depression; and so the staff may thereby have an unconscious bias toward rating more beneficial effects than are actually present. What might be the outcome of this larger trial? Drug X produces the same benefit as it did in the earlier trial, but placebo produces a greater effect in this trial, shrinking or completely eliminating any measureable difference in the effects produced by Drug X and placebo.

Subjective Outcome Measures

Unlike objective measure used in other therapeutic areas such as blood pressure measurements, death, or the presence of a microorganism, psychiatry trials currently rely on outcome measures based on questions that subjects answer on a rating scale. These outcome measures are prone to a greater degree of uncertainty (and often a greater variance), as well as issues such as recall bias, and inter-rater reliability in a multicenter clinical trial.

Regression to the Mean

As already discussed, variance is a way to capture the random change in repeated measurements of the same variable. In a group if people in whom repeated measurements are performed, some people will have values that are far from the mean on one measurement, and values that are quite close to the mean on another measurement—all due to the randomness with which repeated measurements may vary. The presence of an extreme value (one far from the mean) on one measurement, and a value much closer to the mean at another measurement, is what is termed regression to the mean. Extreme values tend to move closer (regress) to the mean on different measurements.

As many trials enroll patients during an exacerbation of their illness, the phenomenon would tend to mean that future measurements would be less extreme even if the subject is on placebo. This phenomenon will therefore act to diminish any drug placebo difference that might be present.

Increased Interaction with Health Care Professionals

During the conduct of a clinical trial, subjects are seen much more frequently than in clinical practice and the visits can last several hours to conduct all the assessments required by the protocol. This increased interaction is well known to increase placebo response.

Medication Nonadherence in Clinical Trials

Another important reason for failed clinical trials is nonadherence to

clinical trial medication. Reasons for this are complex and multifactorial in nature. Whether intentional or unintentional, nonadherence to medication has major consequences in clinical trials leading to a lack of good quality data and poor signal detection, which can often times lead to an incorrect assumption that a drug is ineffective when it just has not been taken correctly or taken at all.

New digital and mobile technologies are emerging where adherence can be monitored remotely, which allow for active intervention by Investigators with nonadhering patients. While these technologies have great potential, it is important to remember that these are only the instruments which will aid investigators in oversight of treatment adherence and enhanced patient engagement.

Statistical Considerations

The proper use and interpretation of statistical methods is important in interpreting any clinical trial result.

FEW USEFUL DEFINITIONS

Positive, Negative, and Failed Trials

How does the scientist or clinician detect whether a therapeutic molecule (we will call it Drug X) offers a benefit (i.e., greater efficacy) in a clinical trial? One method is to perform a test to determine if the effect of Drug X on a variable of interest (like blood pressure) has a greater benefit when compared to the effect of either an inert agent that cannot be visually distinguished from the molecule (i.e., a placebo), or a greater benefit when compared to some other compound (Drug Y) that is already known to produce a beneficial effect. But how does one then detect whether there is a beneficial difference between Drug X and placebo, or between Drug X and Drug Y?

Before this question is answered, one needs to look a little more closely at what happens when a variable is measured over and over again. Again, take blood pressure as an example. Successive measurements of blood pressure in the same person, even when performed just moments apart, will yield different values with each measurement—tiny differences, perhaps, but real ones. A portion of this difference may be explained by physiological changes in the blood pressure from moment to moment, but another portion of the difference arises simply from small errors in the measurement itself. If this is the case, then how does one decide if Drug X produces the same benefit to drop high blood pressure as Drug Y, if measurements of either drug's effect on blood pressure are subject to error? If one does detect a difference in their measured effects on blood pressure, is that difference real, or could it be due completely to the variation in repeated measurements? To circumvent this inherent difficulty in

comparing two measurements, statisticians have defined methods to measure the extent to which multiple measurements will result in a spread of their values around the true average (mean) of that measurement. Not surprisingly, one of these methods to detect the random differences between repeated measurements is termed variance. Signal detection methods take the variance of two sets of measurements (say, from two groups of patients, each group administered a different treatment) into account to determine if there is a true difference conferred by the two treatments. There are many statistical methods to compare groups and to test whether the differences appear to be real, and these methods are chosen depending on the type of clinical trial, and on the outcome being studied (e.g., a change in a numerical value versus a discrete endpoint, such as death). Typically, the statistician hypothesizes that there is no difference in the signal produced by two treatments, and then the statistical tests are performed to see if the signal appears to be present after all, suggesting that the initial hypothesis is false (i.e., that there appears to be a real difference between the treatments).

Given these definitions, signal detection methods may, in principle, give rise to three sets of clinical trial outcomes. If the hypothesis is that two treatments are the same, then a positive trial outcome is one in which the effect of Drug X is statistically different from and better than that of another treatment (better than Drug Y, or better than placebo). A negative trial outcome is one in which this statement cannot be made; namely, Drug X produces an effect that is not statistically different than that of another treatment, thereby suggesting no additional benefit of Drug X.

But how does one know that a negative outcome is truly due to a nonefficacious drug, and not simply to some arbitrary factor such as a poor clinical trial design or a poorly chosen patient population—one that does not allow detection of a positive signal even when a positive signal should be present? To answer this question, some clinical trials will add a third treatment that produces a well-known clinical effect. For example, a test drug, Drug X, will be compared to placebo; and in the same trial, a known beneficial drug, Drug Z, is also compared to placebo. If Drug Z produces a signal that is significantly different from placebo, as it is should, and then if the effect produced by Drug X is not different than placebo, then this is a well-designed clinical trial that yields a very interpretable result: namely, Drug X does not produce an effect. However, what if the trial is so poorly designed or conducted that Drug Z itself fails to separate from placebo? This is an example of a failed clinical trial—one that is poorly designed or conducted, and which is therefore not even interpretable because even a drug with a known effect cannot be detected.

Repeated Measures

Sometimes companies conducting trials wish to explore the effect of

different doses of a compound on multiple symptom domains. Although this sounds like an efficient way to conduct an experiment, it turns out that if there is a small chance that a measured effect of the drug on a single variable is due to chance (and therefore that this is a true effect); it is unfortunately also true that many measurements of many doses on many effects may appear to produce a measureable effect of the drug, and some of these effects are false. Why? Random chance has been multiplied across all of these tests due to multiple measurements, a factor termed repeated measurements. Various statistical techniques have been constructed to overcome the effect of repeated measurements.

A more recent development in statistical design of clinical trials has been alternative clinical trial designs such as adaptive designs (Bowalekar, 2011) and Bayesian methods, but there has been few examples of their adoption in the CNS clinical trials so far.

FUTURE DIRECTIONS FOR DRUG DEVELOPMENT

The cost of developing a new treatment has been increasing to the point that further increases are unsustainable. There is a realization in the pharmaceutical industry that the current model has to change. This will drive a much more synergistic model that includes a whole network of collaborators who by fostering collective scientific creativity will ensure future long-term success. Competitive and technological changes in the pharmaceutical industry—from powerful new drug chemistries to innovative R&D partnerships—have the potential to improve productivity that will in turn increase output with greater efficiency, less cost, and more value.

Collaborative Networks

Some clinical trial goals require larger and longer trials, namely those designed for disease modification or for long-term follow-up to examine durability of effect. For example, in examining clinical trial outcome measures that target negative symptoms in schizophrenia, Laughren and Levin (2011) proposed that for FDA regulatory purposes, a 6-month or longer treatment period was desirable in order to examine durability of effect of the therapeutic molecule proposed for that indication.

Taking this trend toward longer trial durations much further, trials designed to examine the effects of treatments for prodromal Alzheimer disease may last 12 to 24 months, and involve a large sample size in order to detect small changes in neuropsychological measures and/or in functional or clinical outcomes. To accomplish these large trials, involving many thousands of patients and years of treatment duration, collaborative networks have been established between one or more biopharmaceutical companies, academia, and government agencies. These large collaborations

will not only facilitate the conduct of the trials but also decrease the per-entity cost and spread the risk of negative trial outcomes.

Globalization

There has been a marked trend toward drug development programs becoming more and more global. This is occurring for several reasons. Regulators and patient advocates are rightly demanding data that is obtained in ethnically relevant research subjects that address the possibility of differential efficacy and safety, and optimal dosing in these populations.

This trend to globalization has its challenges, including more complexity, longer regulatory reviews in some regions, different standards of medical care in some countries, and the transcultural issues in ensuring consistent reliable administering of subjective outcome measures.

Personalized Technology

CNS trials often use outcome measures that require patient entries into diaries. These can include subjective judgments on sleep quality, symptoms of depression, quality of life, and functional measures. In the past, these entries were made using paper diaries, but increasingly trials have turned to electronic devices that patients use for their entries. Device-based entries offer a number of advantages over the older paper-based methods. For example, entries can be both time stamped and immediately uploaded to a trial database. This speeds the acquisition of data, and it also allows the trial analysis staff to determine that entries were made at the times and dates specified by the protocol, rather than just before a patient's next clinic visit. Finally, electronic devices are increasingly familiar to people through their use of personal smart phones and other handheld devices; this tends to enhance patient interest in and familiarity with the use of these devices.

The introduction of Apple's ResearchKit (http://www.apple.com/researchkit/), holds the promise of taking clinical research into the mainstream with the availability of clinical research to everyone via their mobile devices, with enabling modules such as eConsent. Digital Medicine is profoundly influencing health care delivery systems and disease management.

Personalized Medicine

Currently most treatments available to patients were developed based on their effect on a patient population. It has been a long held aspiration to develop treatments that are personalized to patients based on factors such as genetics. This goal can be termed personalized medicine.

An example of an efficacy-targeted success in personalized medicine is the use of HER2 screening in patients with metastatic breast cancer.

Patients who over-express (are positive for) HER2 have a greater likelihood of responding to trastuzumab (Herceptin).

A similar example of a safety-targeted success is the use of CYP450 testing in order to predict drugs that should be dosed at different strengths in order to avoid excess or poor metabolism of drugs handled by those metabolic paths; and also to avoid drug: drug interactions. How does this affect the use of psychiatric drugs? Fluoxetine inhibits activity of CYP2D6, which metabolizes more than 100 commonly used medications. This means that patients taking fluoxetine may be at risk of receiving a functional overdose of other compounds that are metabolized by 2D6, unless the potential for a drug interaction is avoided by measuring 2D6 function in patients taking fluoxetine, or by tailoring the doses of coadministered drugs that are metabolized by 2D6.

The field of personalized medicine extends much further than these two examples about its applicability to efficacy and safety. Two basic approaches will be considered: tailoring drugs for the individual and identification of genetic susceptibility to disease.

Numerous large clinical trials targeting CNS and other indications routinely collect genetic information on patients in those trials, allowing the retrospective identification of a gene variant profile that correlates with efficacy or safety signals observed during the trial. If identified, then prospective use of these gene variant profiles is feasible, by screening patients for these gene variants in order to predict the efficacy or safety of the drug in question before it is administered to them.

Whole genome sequencing (WGS) is now available through many avenues for the individual patient. WGS approaches conducted in large populations can help identify variants that confer susceptibility to a variety of conditions, and then individual WGS results can be correlated with these population databases to seek clues about that individual's likelihood of harboring or developing those conditions. Caution must be exerted in applying this information to individual genomic data because of the ethical and interpretative issues that may be encountered. For example, if an asymptomatic individual has the variant profile consistent with schizophrenia, how likely is it that this person will ever be symptomatic? And how will this information influence: that person's lifestyle; the extent to which this information is shared with family members, friends, work colleagues, or insurance providers; the advisability of parenting; or if already a parent, the extent to which this information is communicated to the offspring? These issues and many others must be carefully considered by each individual in consultation with his/her health care providers.

Biomarkers

In the broadest sense, biomarkers are any measureable feature of a patient that can serve to predict disease likelihood and severity, disease course,

disease response to treatment, or the current general health status of the patient. Optimally, a surrogate biomarker for disease severity with disease severity can be objectively measured; and so can serve as a clinical trial outcome measure in place of a more subjective "soft" outcome measure such as a symptom rating scale. Biomarkers for psychiatric conditions have been sought for decades. Laughren (2010) discussed the last 50 years of psychiatric drug development and concluded that biomarker approaches were the most likely methods to help advance the field. This has been echoed by many in the research community (e.g., Filiou and Turck, 2011; Li et al., 2012; Nutt and Attridge, 2014). The challenge has been to identify and validate biomarkers meeting these criteria that can be used as trial outcome measures and that would also be accepted by regulatory authorities for drug approvals. To date, this has had no more than limited success in most indications, and no success yet in psychiatric indications. Hopefully the wide variety of imaging, genetic, and other approaches that are underway will soon change this landscape.

Patient Centricity

Traditional drug development processes have not been particularly patient centric. As an example, research protocols had little, if any input from patients, resulting at times in complex burdensome protocols that either excluded most patients or negatively impacted recruitment and retention of research subjects, and the success of the programs.

More recently, there has been a realization that this needs to change. Also in some disease areas, especially rare and orphan diseases, patient advocacy groups are themselves funding clinical trials and having a stronger voice.

More "Real-World" Data

One of the criticisms of drug development programs has been that the data generated is not particularly helpful to clinicians in helping them make treatment decisions and choices. This is driven by a misunderstanding of the purpose of phase 2 and phase 3 clinical trials. These programs are conducted to meet the requirements of regulatory authorities and traditionally it was the postapproval phase 4 studies that were conducted to address questions that would help prescribers understand the optimal use of the new treatment.

As payors across the world have increased their demand for data that justifies pricing above generic treatments, it has become imperative to generate this data earlier in development of new treatments. This will lead to a modification of the issues that are addressed in drug development in preapproval programs.

MORE RELEVANT OUTCOME MEASURES

Traditionally in psychiatric drug development, outcome measures used, measured symptom severity and the impact of new treatment on these symptoms. However, it has become clear that reducing symptoms is not enough. Both payors and patients are looking for the impact on functioning and other novel domains. This will lead to new ways of assessing the impact of a new treatment on patients.

Disease Interception

Throughout the history of medicine, the main focus has been on treating diseases once they have become manifest. It is clear that in many conditions this may be too late, the clearest example currently being Alzheimer disease. There are efforts underway across neuroscience to identify patients that may be at risk and developing interventions that would prevent the disease before it develops. This approach has implications both for regulatory bodies and payors if physicians are to intervene that early in the natural history of a disease.

Regulation of Drug Approvals

When clinical trials yield results that demonstrate the potential benefits of efficacious and safe compounds, these data must be submitted to regulatory authorities for their scrutiny before approval for marketing is granted. It is important to understand the landscape in which regulatory guidelines have developed in order to put the challenges of drug development into perspective. This section reviews the history and the current US regulatory environment for drug development.

Background

The FDA, according to its mission statement, "is responsible for protecting the public health by assuring the safety, efficacy, and security of human and veterinary drugs, biological products, medical devices, our nation's food supply, cosmetics, and products that emit radiation. The FDA is also responsible for advancing the public health by helping to speed innovations that make medicines and foods more effective, safer, and more affordable; and helping the public get the accurate, science-based information they need to use medicines and foods to improve their health." The FDA plays a role in overseeing and guiding development of the medications and devices that are used in the practice of psychiatry and ensuring that they are effective, offer a favorable risk to benefit ratio, are accompanied by adequate directions for use, can be manufactured with reliable quality, and are not promoted in a misleading fashion.

HISTORY OF DRUG REGULATION

The history of American drug regulation features a series of legislative responses to public concerns. A century ago there were no restrictions on what could be sold as medicine and what claims could be made about the products. *Patent* medicines were widely sold for a variety of ailments. Labels did not list ingredients. Some medicines contained ethanol, opium, morphine, heroin, or cocaine; others contained harmless (and ineffective) ingredients. Popular press accounts of the harms of these *tonics* sold to unsuspecting customers, along with concerns about food adulteration and the dangers of chemical additives, led to the passage of the Pure Food and Drug Act of 1906, championed by Harvey Wiley, head of the Bureau of Chemistry in the Department of Agriculture.

Pure Food and Drug Act of 1906

Wiley law, the Pure Food and Drug Act of 1906, was only six pages long. It prohibited the manufacture and interstate shipment of "adulterated" and "misbranded" foods and drugs. This gave the government the authority to go to court against illegal products but did not require any premarketing approval. Initially, the law also did not outlaw false therapeutic claims. A 1912 amendment declared such claims illegal but required that the government prove fraudulent intent. Thus, the only requirement was that the purveyor of a medication believes that the drug was effective. Regarding labeling, the law required only that a contents statement, if used, must be truthful. Not until a 1913 amendment was a statement of contents required.

Elixir Sulfanilamide Tragedy of 1937

The shortcomings of the 1906 law limited the government's ability to ensure the safety of drugs. Although the need for reform of drug regulation was apparent by the early 1930s, legislation proposing expansion of the regulatory authority of the FDA was held up by debate in Congress and opposition from the food and drug industries. Finally, in 1937, a public health scandal emerged that generated the political will to pass the bill. A chemist at Massengill, attempting to develop a palatable formulation for a sulfa antibiotic, created a product that used diethylene glycol as a solvent. No premarket testing, in humans or animals, was required under the laws then prevailing, and none was undertaken. As a result of exposure to this solvent (akin to antifreeze), more than 100 people died, and many of them were children. As it happens, the Pure Food and Drug Act gave the government the authority to take action against the manufacturer, only because the product was labeled an *elixir*. By definition, an *elixir* uses ethanol as a solvent; had the manufacturer labeled the product a *solution,* it would not have been in violation of any law.

Food, Drug, and Cosmetic Act of 1938

Although there had been other crises associated with the marketing of untested, unsafe drugs, the elixir sulfanilamide tragedy spurred the passage of the Food, Drug, and Cosmetic Act of 1938. This law contained a number of important provisions concerning drug regulation. The 1938 law required that new drugs undergo safety testing, the results of which would be submitted to the FDA in a new drug application (NDA). Premarketing approval of the NDA was required before a drug could be sold. The law also required that drugs have adequate labeling for safe use. Furthermore, it clearly prohibited false therapeutic claims for drugs, without the requirement that the government show fraudulent intent. In addition, the FDA was given authority to conduct factory inspections.

Thalidomide Tragedy of 1962

By the early 1960s, Senate hearings on strengthening the regulation of drugs were once again under way. In the midst of these hearings, chaired by Senator Estes Kefauver, came the news of the devastating effects of thalidomide (Thalomid), which was being marketed in Europe. Although the NDA had been submitted to the FDA for review, the medical officer responsible for the review, Frances Kelsey, had requested additional safety information, and the drug had not yet been approved. Kelsey's role in protecting the American public from the dangers of thalidomide was widely reported and generated support for passage of what became known as the Kefauver-Harris Amendments in 1962. These amendments required drug manufacturers to prove to the FDA that their products were safe *and* effective before marketing. They established that the evidence of efficacy must be derived from "adequate and well-controlled trials." Furthermore, the amendments also addressed the use of drugs in clinical trials. Under the Kefauver-Harris Amendments, the current investigational new drug (IND) process was established. Under the law, the FDA had to be provided with full details of the clinical investigations, including drug distribution, and the clinical studies had to be based on previous animal investigations to ensure safety.

After the addition of the efficacy requirement, the FDA commissioned a review by the Institute of Medicine of all drugs approved between 1938 and 1962 to determine whether they were effective as well as safe. This process, known as the drug efficacy study implementation (DESI), resulted in the removal of some drugs from the market, whereas others were determined to be effective. Nonprescription drugs were also evaluated under a similar process, which resulted in publication of monographs establishing which ingredients may be "generally recognized as safe and effective" for various conditions.

IND Regulations

After the 1962 amendments, the FDA issued and implemented regulations concerning the investigational use of drugs; these regulations were updated and revised in 1987. The revisions were intended to "ensure FDA's ability to monitor carefully the safety of patients participating in clinical investigations, while also facilitating the development of new beneficial drug therapies." The regulations serve to "focus FDA's attention during the early phase of clinical research on protecting the safety of human test subjects and to give sponsors greater freedom to design, revise, and implement clinical research studies" and to "facilitate consultation between FDA and drug sponsors, once the preliminary human studies have been completed and a drug appears to have marketing potential, to help ensure that the design of major clinical trials is acceptable and will support marketing approval if the test results are favorable." And beginning in the mid-1970s, the FDA began giving advice on the kinds of studies and data that would be needed to support drug approval. Clinical guidances have been published, setting forth the data needed for drugs in various therapeutic categories.

PDUFA, FDAMA, and Pediatric Provision

Thirty years after the 1962 amendments, frustration over lengthy review times led to the passage of the Prescription Drug User Fee Act (PDUFA) of 1992, which charged fees for NDA review to facilitate the expansion of the FDA's review resources and, in return, ensured action on an application in a specified period of time. PDUFA was renewed with the passage of the FDA Modernization Act (FDAMA) of 1997, which further shortened review times and made additional changes in the review process. The authority to collect user fees was again reauthorized as part of the Public Health Security and Bioterrorism Preparedness Act of 2002 and the Food and Drug Administration Amendments Act of 2007 (FDAAA).

Among the provisions of FDAMA were the pediatric exclusivity provisions. These offered incentives, in the form of an additional period of protection from generic competition, to sponsors for fulfilling an FDA request for pediatric data on their products. In 2002, the Best Pharmaceuticals for Children Act (BPCA) was passed, providing further clarification of the pediatric exclusivity program and establishing a system for conducting studies through the National Institutes of Health for high-priority drugs that were already off-patent or for which the commercial sponsor declined to respond to the agency's request. In the first 10 years after the passage of FDAMA, over 150 submissions of data in response to written requests were evaluated to determine whether pediatric exclusivity should be granted. Among these were studies of a number of psychotropic agents, including antidepressants, antipsychotics, and anxiolytics. In many

cases, the results of the studies led to labeling changes stating that safety and efficacy in children had not been established; yet pediatric exclusivity is awarded even in these cases. In addition, under the provisions of the BPCA, upward of 50 drugs have been included on a priority list, including lithium (Eskalith), bupropion (Wellbutrin), and clonidine (Catapres).

Although FDAMA and BPCA established incentives for pharmaceutical companies to conduct pediatric studies, such studies remained optional. Subsequently, in 2003, the Pediatric Research Equity Act (PREA) authorized the agency to require sponsors to conduct pediatric studies as part of their development program. The requirements may be waived if the drug is not likely to be used in a substantial number of pediatric patients or does not represent a meaningful therapeutic benefit for pediatric patients. In addition, the agency may also waive the requirement if "there is evidence strongly suggesting that the drug or biological product would be ineffective or unsafe in all pediatric age groups," or on the grounds that the studies are impossible or highly impracticable. The agency may require that the pediatric studies are submitted as part of the initial application or may defer all or some of the pediatric studies until after the application is approved for use in adults.

The FDAAA of 2007 reauthorized PDUFA, PREA, and BPCA as well as creating new programs to enhance postmarketing safety surveillance and the FDA's ability to communicate safety information.

STRUCTURE OF THE FDA

There are five headquarters centers in the Washington, D.C., area: The Center for Drug Evaluation and Research (CDER), the Center for Biologics Evaluation and Research (CBER), the Center for Devices and Radiological Health (CDRH), the Center for Veterinary Medicine (CVM), and the Center for Food Safety and Applied Nutrition (CFSAN). Staff in these offices work primarily on product review and regulatory policy. In addition, a network of field offices throughout the United States focuses on inspection and surveillance, laboratory work, and public and industry education.

Most relevant to psychiatric treatment is the work of CDER, where drugs and therapeutic biological treatments are reviewed, and CDRH, which has jurisdiction over medical devices such as those used in ECT. This section will address the review of drugs and biological therapeutic products as undertaken in CDER.

NEW DRUG DEVELOPMENT

The development of a new pharmacological treatment of a psychiatric (or any other) disorder usually begins with exploration of a compound's actions in in vitro and animal investigations. These investigations may be carried out before involving the FDA. Indeed, initial safety testing in animals is required before beginning human studies that are subject to

FDA review. Most compounds are abandoned in these early stages of development. However, those that appear promising and are reasonably tolerated in animal safety studies may go forward in development.

Definition of a Drug

The Food, Drug, and Cosmetic Act defines drugs as:

1. Articles recognized in the official U.S. Pharmacopeia, the official Homeopathic Pharmacopeia of the United States, or the official National Formulary
2. Articles *intended for* [emphasis added] use in the diagnosis, cure, mitigation, treatment, or prevention of disease
3. Articles (other than food) intended to affect the structure or function of the body
4. Articles intended for use as a component of any article specified previously

Therefore, it is the *intended use* of the substance that defines it as a drug. As soon as a substance is intended for human use toward one of these objectives (i.e., when human studies begin), the substance is, by definition, a drug and is subject to regulation under the drug provisions of the Food, Drug, and Cosmetic Act.

Note that therapeutic biological products are a subset of drugs. Although the term "drug" is often used to designate a typical small-molecule product and the terms "biologic" or "therapeutic biologic" to designate the large, biologically derived substances used therapeutically such as monoclonal antibodies, both types of products are considered drugs outside the FDA. In this section, the term *drug* will be used to encompass all products meeting the statutory definition.

Phases of Drug Development

Drug development to support a commercial marketing application proceeds in somewhat overlapping phases. Nonclinical (also called *preclinical*) studies begin before human studies but continue throughout the clinical development period. Each phase of human testing is supported by animal tests that are appropriate to the extent of planned human exposure in the ensuing studies. Human studies proceed from basic characterization of the drug product to establishment of efficacy and elucidation of the adverse event profile in the intended population of use. Although the categorizations phase 1, phase 2, and phase 3 are used commonly, the development of a drug does not proceed linearly through these phases. For example, studies of drug–drug interaction and studies of pharmacokinetics in renal or hepatic impairment are both considered phase 1 studies, but often occur late in drug development. Finally,

unanswered questions or new indications may be explored after the drug has been approved for marketing and are commonly called phase 4 studies.

Nonclinical Studies

Nonclinical studies that are sufficient to establish a tolerable dose and to identify the target organs of toxicity for a new drug must be conducted before first use of a new chemical entity in humans. A standard battery of animal studies and in vitro studies are required to support a marketing application, with the duration and timing of these studies related to the proposed duration of use of the drug in the clinical program. Thus, acute toxicology studies are needed to support single-dose human administration, whereas longer studies must be completed before using the drug in human studies involving chronic exposure.

Phase 1. Phase 1 studies represent the initial introduction of the new drug into humans. These studies, usually conducted in healthy volunteers, typically in closely monitored (often inpatient) settings, serve to characterize the absorption, distribution, metabolism, and excretion of the compound, to identify overt toxicities associated with drug administration, and to establish a tolerable dose for use in further studies. In some circumstances, phase 1 studies may involve patients and may be designed to provide early indications of the drug's efficacy; examples include drugs intended as chemotherapeutic or anti-HIV agents. Phase 1 also includes studies in which investigational drugs are used as research tools to explore biological phenomena or disease processes. The objective of phase 1 in commercial drug development is to obtain sufficient information about the drug's pharmacokinetics and pharmacological effects to permit the design of well-controlled, scientifically valid, phase 2 studies. The total number of subjects and patients included in phase 1 studies varies with the drug but is generally in the range of 20 to 80.

Phase 2. Phase 2 includes the initial controlled clinical efficacy studies. These studies typically include carefully selected patients with the disease or condition under study and are usually well controlled, closely monitored, and optimized for the collection of efficacy data. In phase 2, exploratory work is undertaken to help determine the optimal doses of the drug. In conditions for which no standard outcome measures have been defined, studies may explore different approaches to capturing the effect of the drug. Safety data are collected in this phase to determine the common short-term side effects and risks associated with the drug.

Phase 3. After preliminary evidence suggesting effectiveness of the drug has been established in phase 2 trials, additional information about effectiveness and safety is needed to evaluate the overall risk to benefit

relationship of the drug and to provide an adequate basis for product labeling. Phase 3 studies provide this information. The inclusion and exclusion criteria in phase 3 studies are usually broader than in phase 2, in order to select a population that is more representative of the patients who will be exposed to the drug during marketing. Phase 3 studies may include from several hundred to several thousand subjects, including individuals with a broad range of typical comorbid conditions and concomitant medications seen in the target population.

Phase 4. Once the drug has been approved, subsequent postmarketing activities may be conducted in phase 4. Studies to elucidate specific issues about a drug, which are considered important for optimal use but not essential to know before drug approval, may be performed. Studies to support approval for additional indications (e.g., use in pediatrics, if not previously established) may also be conducted. Programs of active surveillance for adverse events may be undertaken, as well as the required passive collection of adverse event reports from patients, health care providers, and other sources, which manufacturers must report to the FDA periodically.

IND Review Process. Before beginning human studies, the commercial drug sponsor must submit a request for an IND exemption from interstate shipping laws, or an IND, for review by the FDA. Once the IND is submitted, the sponsor must wait 30 calendar days before initiating any clinical trials. During this time, the FDA has an opportunity to review the IND for safety to ensure that research subjects are not subjected to unreasonable risk.

REVIEW EMPHASIS AND PROCEDURES. According to the Code of Federal Regulations:

FDA's primary objectives in reviewing an IND are, in all phases of the investigation, to assure the safety and rights of subjects, and, in phases 2 and 3, to help assure that the quality of the scientific evaluation of drugs is adequate to permit an evaluation of the drug's effectiveness and safety. Therefore, although FDA's review of phase 1 submissions will focus on assessing the safety of phase 1 investigations, FDA's review of phases 2 and 3 submissions will also include an assessment of the scientific quality of the clinical investigations and the likelihood that the investigations will yield data capable of meeting statutory standards for marketing approval.

Thus, the reviewers pay close attention in all reviews in determining whether the risks of the study are appreciated, anticipated, communicated, and managed by the sponsor. For later phase studies, they also consider whether the study design is adequate to meet the scientific and regulatory objectives of the study. During the review of an initial IND submission (as

well as the review of subsequent submissions of additional protocols and amendments), the interdisciplinary review team approaches these issues from various angles including: clinical study results; chemistry/product quality; pharmacology/toxicology information; clinical pharmacology data; and biostatistics use and results.

Clinical Hold. If no concerns are identified by the interdisciplinary review team, the IND goes into effect, and the studies proposed in the initial submission are permitted to proceed. Once the IND is in effect, additional protocols, study results, and safety information are submitted. For each new clinical study proposed, the review process is similar, but the sponsor is not required to wait 30 days before initiating subsequent studies.

However, at the time of initial IND review or at any time during drug development, the FDA can place human studies under an IND on *clinical hold*. A clinical hold, according to the Code of Federal Regulations, is "an order issued by FDA to delay a proposed clinical investigation or to suspend an ongoing investigation. The clinical hold order may apply to one or more of the investigations covered by an IND."

Under the Code of Federal Regulations, the FDA may impose a clinical hold on an IND for any of the following reasons:

- Human subjects are or would be exposed to an unreasonable and significant risk of illness or injury.
- The clinical investigators named in the IND are not qualified by reason of their scientific training and experience to conduct the investigation described in the IND.
- The investigator brochure is misleading, erroneous, or materially incomplete.
- The IND does not contain sufficient information required under §312.23 to assess the risks to subjects of the proposed studies.
- The plan or protocol for a phase 2 or 3 study is clearly deficient in its design to meet its stated objectives. When a clinical hold is issued, a sponsor must address the issues that are the basis of the hold before the order is removed, and the study may proceed.

Marketing Application Submission and Review. The culmination of the development program is the submission of the NDA for a small-molecule drug product or a biologics licensing application (BLA) for a therapeutic biological product. FDA staff have 10 months (6 months for a drug designated as a priority application owing to its potential to offer a breakthrough in treatment) to assess thoroughly the raw data and to compare their own conclusions about the data to the applicant's conclusions presented in written reports. The objective is to determine

whether the sponsor (now called the *applicant*) has met the regulatory burden for marketing: Demonstrating efficacy through "substantial evidence" from "adequate and well-controlled investigations, including clinical investigations, by experts qualified by scientific training and experience to evaluate the effectiveness of the drug involved" and demonstrating safety through "all tests reasonably applicable." Therefore, in the course of review, FDA staff assess whether the clinical studies were indeed appropriately designed and conducted, whether the data support the conclusion that the drug is effective, whether the proper human and animal safety data ("all tests reasonably applicable") were collected and illustrate a favorable risk to benefit ratio, and whether the manufacturing methods and controls are sufficient to ensure consistency of identity, quality, purity, and strength of the product.

SAFETY REVIEW. The purpose of the safety review is twofold: First, to determine the adequacy of the extent of exposure and the specific tests carried out in the safety evaluation (such as laboratory tests, electrocardiograms [ECGs], and adverse event collection) and to assess whether it meets the standard of "all tests reasonably applicable." For example, approval of a drug in a class known to be associated with a prolongation of the QT interval might be withheld if safety review determined that the effects of the drug on cardiac conduction were not thoroughly evaluated. The second purpose is to characterize the safety profile, so that the risks of the drug may be weighed against the benefits identified in the efficacy review. Characterization of the safety profile includes identifying and estimating the frequency of common adverse events that are, or may be, causally related to the use of the drug; identifying factors that predict the occurrence of adverse events, such as age, gender, concomitant medication, or illness; and examining the data for events that suggest problems that are so serious that the use of the drug should be limited or approval withheld altogether. The clinical reviewer evaluates all available safety information and integrates it into an overall picture of the drug's effects.

EFFICACY REVIEW. The review of efficacy, conducted jointly or in parallel by the medical officer and biostatistician, uses raw data from the pivotal clinical trials in the application to confirm the applicant's conclusions about the study results. Any deviations from the protocol-specified study plan are evaluated for possible impact on the outcome of the studies. Reviewers analyze the applicant's assumptions about outcome for protocol violators, subjects who withdraw early from studies, or subjects with missing data. They test the impact of alternative assumptions on study outcomes. The objective of the efficacy review is to determine whether the studies submitted were "adequate and well controlled" and whether they

provide "substantial evidence" of efficacy. For the most important trials, the review also includes clinical trial site inspections by the Division of Scientific Investigation (DSI). By examining the trial site, documentation, source materials, and equipment, the DSI establishes whether there are important violations that call into question the results of the study. Such findings are unusual but do occur as a result of deliberate fraud, error, or incompetence.

Abuse Liability Assessment. For most psychotropics, the NDA also includes a section addressing the abuse potential of the product. Assessment of abuse potential is required for many drugs with CNS activity. This testing may include a battery of in vitro receptor-binding studies, various animal studies of the rewarding properties of the drug, and, if preliminary studies suggest the potential for abuse, human laboratory studies. Data from these studies, as well as other information concerning the chemical structure of the compound, and any other relevant data from literature or epidemiological reports are reviewed as part of the NDA and are considered in writing the relevant sections of the product package insert. Importantly, these data are also taken into consideration in preparing a recommendation concerning control of the substance under the Controlled Substances Act. This recommendation is prepared by a team within the Office of the Director of CDER at the FDA and is forwarded through the secretary of Health and Human Services to the Drug Enforcement Administration.

PHARMACOLOGY AND TOXICOLOGY. By the time that the NDA is submitted, the nonclinical data include a variety of animal studies. There should be both acute and multidose toxicity studies, the durations of which are determined by the intended use of the product. Also included are reproductive toxicology studies, mutagenicity studies, carcinogenicity studies for chronically used products, and any special toxicity studies (such as dermal sensitization studies for a transdermal product and blood compatibility studies for an intravenous [IV] product). There also should be studies of the absorption, distribution, metabolism, and elimination of the drug in animals, because these data are needed to compare the dose response in animals to the dose recommended for humans. In the course of review, the pharmacologist examines the results of these studies and also critically evaluates the study design and conduct to determine if the data may be accepted as valid.

CLINICAL PHARMACOLOGY. The clinical pharmacology reviewer examines data relevant to understanding all aspects of the absorption, distribution, metabolism, and excretion of the drug. This includes information on any important influences on these parameters, such as food effects, drug interactions, or effects of age, gender, or hepatic or renal impairment. This

information is included in the clinical pharmacology section of the label and often forms the basis for important precautions to ensure safe use of the product. Explorations of exposure–response relationships for efficacy and safety end points are also an important part of the clinical pharmacology review. Some pharmacogenomic analyses (analyses of genetically based differences in pharmacology), such as explorations of the role of metabolic polymorphisms, may also be undertaken.

CHEMISTRY/PRODUCT QUALITY. In reviewing an NDA, the chemistry reviewer examines the manufacturing information to determine whether there are appropriate specifications and test methods in place to ensure identity, quality, purity, and strength, as well as consistency from batch to batch. For a BLA, the quality reviewer, similarly, is responsible for ensuring the quality, safety, efficacy, availability, and security of therapeutic protein and monoclonal antibody products. The reviewer also evaluates information about stability, including the adequacy of the stability data to support product use, expiration dating, and any concerns about packaging or shipping that might affect stability. In addition, the reviewer takes into account the findings of inspections of the manufacturing facilities, which are undertaken to ensure adherence to the principles of good manufacturing practices. Reviewers of biological products participate in the actual preapproval and biennial facility inspections. The FDA may withhold approval if inspections show significant procedural problems at the facilities that are responsible for manufacturing or packaging the product.

REGULATORY DECISION. All of the discipline reviews are taken into consideration in making an assessment of whether the drug's benefits outweigh its risks for the intended use in the intended population. If a drug offers significant benefits for a narrowly defined population but carries significant risks, the FDA may decide to approve the drug subject to special conditions to ensure that physicians understand and comply with using the drug only for those in whom the anticipated benefits can be expected to be greater than the possible risks.

If there are unanswered questions or unresolved issues in an application or concerns about safety or lack of efficacy, the FDA may withhold approval. The applicant is issued a letter explaining the deficiencies precluding approval and instructions on how to address these deficiencies.

PRODUCT LABELING. For the physician and the patient, the most obvious product of the NDA review process (apart from the regulatory decision that makes the product available on the market) is the package insert or label. The multidisciplinary review team carefully examines language drafted by the NDA applicant and ensures that the statements throughout the label are supported by the submitted data. The format of the label is set forth in

the Code of Federal Regulations (21 CFR §201).

Postapproval (Phase 4). Despite careful premarket review, postmarketing surveillance of drug adverse effects is also essential. It is evident that all possible side effects of a drug cannot be anticipated based on preapproval studies. The total exposure to an investigational drug during its development depends, to some extent, on the novelty of the compound and on the presence of any specific safety concerns. However, for practical reasons, it is rarely possible to have sufficient premarket exposure to characterize the occurrence of rare adverse events. According to the current recommendations of the International Conference on Harmonization, for new chemical entities that are to be used chronically (or repeatedly but intermittently) in the treatment of conditions that are not immediately life threatening, the number of patients exposed should be sufficient to detect serious adverse events occurring at a rate of approximately 1 per 1,000 in short-term use. It is recommended that the safety database include at least 1,500 subjects. Of these, 300 to 600 should be treated for at least 6 months, and 100 should be treated for 1 year. Applying the so-called rule of three, if a particular serious adverse drug effect is not observed in the 100 subjects exposed for 1 year, this can provide reasonable assurance that the true cumulative 1-year incidence is no greater than 3 percent. However, this rule of three represents only 95 percent confidence limits. It must be noted that even a database of 3,000 subject exposures stands a 5 percent chance of missing events that occur at a rate of 1 per 1,000, and such a rate may translate into a significant number of occurrences when the drug is introduced into a population of hundreds of thousands.

Therefore, the FDA maintains a system of postmarketing surveillance and risk assessment programs to identify adverse events that did not appear during the drug approval process. The agency monitors adverse reactions reported by drug companies, consumers, and physicians and seeks out signals suggestive of drug-related events. The agency uses this information to update drug labeling, to identify unanticipated drug–drug or drug–disease interactions, to spot common medication errors and to take action to prevent them, and, on rare occasions, to re-evaluate the approval or marketing decision.

POSTMARKETING SAFETY. Due to practical limitations on how many patients can be studied in clinical trials for a new drug, it is inevitable that some safety concerns will come to light until after a new drug is used in a broader population after approval. Unlike Frances Kelsey, the modern FDA reviewer often does not have foreign postmarketing data available to supplement the clinical trial data submitted in the drug application since drugs are now frequently approved first in the United States. In the post-

PDUFA era, critics of the FDA raised concerns that the emphasis on bringing new drugs to market more rapidly had led to a failure to identify important safety issues. Several high-profile drug safety issues led the agency to examine its procedures and to launch several initiatives aimed at improving the safe use of medicines, including a dedicated Safety Policy and Communication staff that ensures that physicians, health provider groups, and patients are alerted to important new safety information in a timely manner. The Safety Policy and Communications staff includes a Drug Safety Oversight Board (known as the DSB) that was established to provide independent oversight and advice to the center's director on the management of important and emerging drug safety issues. The DSB provides perspective from scientists not involved in the initial review and approval of a particular marketing application and provides a venue for adjudicating organizational disputes concerning the management of drug safety issues. In 2007, the Safety Policy and Communications Staff launched a new tool for communicating safety issues, the FDA Drug Safety Newsletter, which is distributed electronically and viewable at the FDA's Web site.

FUTURE DIRECTIONS FOR FDA

In 2004, the FDA launched a new initiative known as the *Critical Path Initiative,* which is intended to "stimulate and facilitate a national effort to modernize the scientific process through which a potential human drug, biological product, or medical device is transformed from a discovery or 'proof of concept' into a medical product." By identifying key areas in need of attention and partnering with other agencies and organizations to initiate research, the FDA is working to improve the tools used to evaluate the safety and efficacy of human and veterinary products as well as the safety and nutrition of food and food ingredients. A list of areas representing high-priority opportunities for improvement was disseminated and work has been initiated in a variety of projects.

One project specific to psychiatry is an effort under the aegis of the *Biomarkers Consortium,* a public/private biomedical research partnership supported by the Foundation for the National Institutes of Health (FNIH). Among their first projects is an effort to identify genomic biomarkers for treatment response in major depressive disorder. Many other critical path projects are broadly relevant to psychotropic drugs, including work on examining the genetic basis of adverse drug events; work intended to broaden understanding of the biochemical and genetic basis of drug-induced liver injury; a database of digital ECGs collected in drug-approval studies to be used to improve methods of measuring drug effects on cardiac conduction; and efforts to identify nephrotoxic biomarkers that predict the effects of drugs on kidney function.

As drug development becomes more and more of a global endeavor, the

need for regulatory authorities in different countries and regions to interact and discuss regulatory issues and challenges becomes more important.

CONCLUSION

Drug development in every field has become increasingly complex, costly and challenging, and in no other therapeutic area is this more evident than in psychiatric conditions. The challenges are enormous in translating from the Gordian knot of poorly understood CNS pathophysiological processes and nonpredictive models into clinical trials that will have a good probability of success. Nevertheless, there are new technological and scientific approaches that provide a basis for increased hope as clinical trials for psychiatric conditions move into the 21st century. The goal is clear. Psychiatrists must discover, develop, and obtain regulatory approval for new medicines that will improve treatment of many unmet or poorly served medical needs in the psychiatric field. Patients are waiting.

REFERENCES

Apple ResearchKit. Available at http://www.apple.com/researchkit/.

Brayanov JB, Smith MA. Bayesian and "anti-Bayesian" biases in sensory integration for action and perception in the size-weight illusion. *J Neurophysiol*. 2010;103(3):1518–1531.

Chapman SC, Horne R. Medication nonadherence and psychiatry. *Curr Opin Psychiatry*. 2013;26(5):446–452.

Code of Federal Register Volume 71, No. 15, Tuesday, January 24, 2006, Rules and Regulations, Requirements on Content and Format of Labeling for Human Prescription Drugs Title 21: Food and Biological Products Drugs, Chapter 1—Food and Drug Administration, Department of Health and Human Services, Subchapter D—Drugs for Human Use, Part 314, Applications for FDA approval to market a new drug.

*Code of Federal Regulations, Title 21: Food and Drugs, Chapter 1—Food and Drug Administration, Department of Health and Human Services, Subchapter C—Drugs: General, Part 201, Labeling and Subchapter D—Drugs for Human Use: Part 312, Investigational new drug application.

Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimers Res Ther*. 2014;6(4):37.

Cutler AJ, Tran-Johnson T, Kalali A, Aström M, Brecher M, Meulien D. A failed 6-week, randomized, double-blind, placebo-controlled study of once-daily extended release quetiapine fumarate in patients with acute schizophrenia: lessons learned. *Psychopharmacol Bull*. 2010;43(4):37–69.

Daley A. Chief Technology Investment Strategist – Could Big Pharma's Patent Collapse Sink your Portfolio. 2012. Available at http://www.caseyresearch.com/articles/could-big-pharmas-patent-collapse-sink-your-portfolio. Accessed October 5, 2016.

De D. *The Transformation Challenges Facing Pharmaceutical Executives. Life Sciences Consulting*. Tata Consultancy Services Ltd; 2014.

Doug H. *Is the Patient Cliff a Lethal Blow to Big Pharma*. 2013. Available at

https://www.caseyresearch.com/articles/is-the-patent-cliff-a-lethal-blow-to-big-pharma. Accessed October 5, 2016.

Elenko E, Underwood L, Zohar D. Defining digital medicine. *Nat Biotechnol*. 2015;33(5):456–461.

Federal Food, Drug, and Cosmetic Act: Public Law No. 75–717, 52 Stat. 1040 (1938), 21 USC §§301 et. seq., Chapter II—Definitions, Sec. 201 [321] Definitions.

Federal Register Vol 52, No. 53. (1987): Rules and Regulations, Department of Health and Human Services, Food and Drug Administration, 21 CFR Parts 312, 314, 511, and 514 [Docket No. 82N-0394], New Drug, Antibiotic, and Biologic Drug Product Regulations, Final Rule.

Filiou MD, Turck CW. General overview: biomarkers in neuroscience research. *Int Rev Neurobiol*. 2011;101:1–17.

Foley KF, Quigley DI. Pharmacogenomic potential of psychiatric medications and CYP2D6. *MLO Med Lab Obs*. 2010;42(1):32–34.

George M, Selvarajan S, S SK, Dkhar SA, Chandrasekaran A. Globalization of clinical trials—where are we heading? *Curr Clin Pharmacol*. 2013;8(2):115–123.

Ghaemi SN. DSM-5 and the miracle that never happens. *Acta Psychiatr Scand*. 2014;129(6):410–412.

*Guidance: Drug Safety Information—FDA's Communication to the Public: US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, 2007.

Hellewell JS. Patients' subjective experiences of antipsychotics: clinical relevance. *CNS Drugs*. 2002;16(7):457–471.

International conference on harmonisation; guidance on statistics principles for clinical trials; availability–FDA. Notice. *Fed Regist*. 1998;63(179):49583–49598. Available at http://www.ich.org/products/guidelines/efficacy/efficacy-single/article/statistical-principles-for-clinical-trials.html. Accessed October 5, 2016.

Kemp AS, Schooler NR, Kalali AH, et al. What is causing the reduced drug-placebo difference in recent schizophrenia clinical trials and what can be done about it? *Schizophr Bull*. 2010;36(3):504–509.

Laughren TP. What's next after 50 years of psychiatric drug development: an FDA perspective. *J Clin Psychiatry*. 2010;71(9):1196–1204.

Laughren T, Levin R. Food and drug administration commentary on methodological issues in negative symptom trials. *Schizophr Bull*. 2011;37(2):255–256.

Library of Congress: Project on Decade of the Brain. 2000. http://www.loc.gov/loc/brain/. Accessed October 5, 2016.

McCoy TH, Castro VM, Rosenfield HR, Cagan A, Kohane IS, Perlis RH. A clinical perspective on the relevance of research domain criteria in electronic health records. *Am J Psychiatry*. 2015;172(4):316–320.

Michopoulos V, Norrholm SD, Jovanovic T. Diagnostic biomarkers for posttraumatic stress disorder: promising horizons from translational neuroscience research. *Biol Psychiatry*. 2015;78(5):344–353.

Nutt DJ, Attridge J. CNS drug development in Europe—past progress and future challenges. *Neurobiol Dis*. 2014;61:6–20.

Paul SM, Mytelka DS, Dunwiddie CT, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Durg Discov*. 2010;9(3):203–214.

Riffenburgh RH. *Statistics in Medicine*. 3rd ed. Amsterdam: Elsevier Press; 2012.

Rutherford BR, Wall MM, Glass A, Stewart JW. The role of patient expectancy in placebo and nocebo effects in antidepressant trials. *J Clin Psychiatry*. 2014;75(10):1040–1046.

Schneider LS, Mangialasche F, Andreasen N, et al. Clinical trials and late-stage drug development for Alzheimer's disease: an appraisal from 1984 to 2014. *J Intern Med*. 2014;275(3):251–283.

Slamon D, Pegram M. Rationale for trastuzumab (Herceptin) in adjuvant breast cancer trials. *Semin Oncol*. 2001;28(1 Suppl 3):13–19. Review.

Sullivan P, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet*. 2012;13(8):537–551.

Tufts Center for the Study of Drug Development, Volume 16, Number 6. 2014.

Wilson BJ, Nicholls SG. The Human Genome Project, and recent advances in personalized genomics. *Risk Manag Health Policy*. 2015;8:9–20.

Yarbrough C. Adaptive clinical design addresses the uncertainties. *Clinical Leader*. 2013;3.

Zygmunt A, Olfson M, Boyer CA, Mechanic D. Interventions to improve medication adherence in schizophrenia. *Am J Psychiatry*. 2002;159:1653–1664.

▲ 34.3 Medication-Induced Movement Disorders

PHILIP G. JANICAK, M.D., AND KUMAIL HUSSAIN, M.D.

All antipsychotics can induce movement disorders. Thus, the *Diagnostic and Statistical Manual of Mental Disorders,* 5th edition (DSM-5) continues to characterize these agents as "neuroleptics" in recognition of this property. Although controversial, these agents are categorized as first-generation (or typical) and second-generation (or atypical) antipsychotics with one potential differentiation being their propensity to produce movement disorders. In addition to antipsychotics, a variety of other psychotropic and nonpsychotropic agents can also produce these symptoms (Table 34.3–1). For example, lithium (Eskalith), anticonvulsants, or antidepressants can all induce tremors; dopamine-receptor agonists can induce dyskinesia; some antidepressants can induce akathisia or produce abnormal movements upon their discontinuation; and prochlorperazine (Compazine) or metoclopramide (Reglan) can induce pseudoparkinsonism. Based on the specific agent and type of movement problem, various strategies are employed to counter these adverse effects.

HISTORY

An important observation is that spontaneous movement disorders (e.g., blunted affect, psychomotor slowing, buccal–lingual–masticatory movements, catatonia, alogia) can occur in patients with schizophrenia even in the absence of exposure to antipsychotic medications. For example, one review found a 4 to 7 percent prevalence of dyskinetic movements in patients with schizophrenia not on an antipsychotic. This risk increases with age and illness chronicity. Thus, an association between psychotic illness and movement abnormalities was described years before the introduction of antipsychotics.

The first antipsychotic, chlorpromazine (Thorazine), entered clinical practice in 1952. As a specific and effective medication for psychosis, it revolutionized treatment for millions of patients who until then had received little beyond custodial care. Additional agents were synthesized in the phenothiazine class and eventually several other first-generation antipsychotics (FGAs) became available. All appeared equally effective if given in sufficient bioequivalent doses.

Problems with movement disorders went relatively unnoticed with the early FGA, low-potency agents. Symptoms such as sedation, anticholinergic effects, and hypotension, however, encouraged the development of more potent and specific dopamine 2 (D2) receptor antagonists, culminating with the introduction of haloperidol (Haldol) in 1968. Since this agent had

little effect on other receptors, previous troublesome adverse effects could usually be avoided or substantially reduced. Before the introduction of the second-generation antipsychotics (SGAs), a common strategy was to use higher doses of a high-potency FGA with the belief that optimal efficacy could only be obtained at the cost of some extrapyramidal symptoms (EPS). This approach resulted in an increased incidence of acute and late-onset movement syndromes (e.g., pseudoparkinsonism and tardive dyskinesia). Not surprisingly, adherence was compromised because patients often felt the benefit of treatment was outweighed by such adverse effects. Adjunctive agents (e.g., benztropine [Cogentin]) were often used to control these problems but possessed their own set of uncomfortable (or even dangerous) anticholinergic effects. Coming full circle, older, lower-potency FGAs were then used as an alternative approach to minimize acute movement syndromes. Some psychiatrists also maintained that weaker D2 antagonists could produce superior outcomes for certain patients while minimizing these risks. This position was supported with the release of the lower-potency D2 antagonist, clozapine (Clozaril) in 1990.

DIAGNOSTIC CRITERIA

The DSM-5 lists 10 categories under "Medication-induced Movement Disorders and Other Adverse Effects of Medication." Clinical syndromes are divided into those associated with antipsychotic (neuroleptic) use and those associated with other psychotropic and nonpsychotropic agents. Neuroleptic-related syndromes are divided according to the usual time of onset in relation to starting medication or increasing the dose. In general, SGAs cause less frequent and less severe movement disorders than their first-generation counterparts. In sensitive individuals, however, these symptoms also occur with SGAs. Further, the reduction in incidence and severity of abnormal movements may delay recognition when they do occur. The specific categories include neuroleptic or other medication-induced parkinsonism; neuroleptic malignant syndrome (NMS); medication-induced acute dystonia and akathisia; tardive dyskinesia; tardive dystonia and akathisia; medication-induced postural tremor; other medication-induced movement disorder; and antidepressant discontinuation syndrome.

The term "neuroleptic" (from the Greek word meaning to "clasp the neuron") includes existing antipsychotics, certain dopamine-blocking drugs used to treat gastrointestinal symptoms (e.g., metoclopramide) and amoxapine (Asendin), an antidepressant whose active metabolite has dopamine antagonist activity. The remaining categories include movement disorders such as tremor or parkinsonism induced by antipsychotics or other medications. Table 34.3–2 lists a brief description of the 10 categories provided in the DSM-5.

CENTRAL NERVOUS SYSTEM CONTROL OF MOVEMENT

Cortical motor areas, the cerebellum, basal ganglia, and related subcortical nuclei control normal movement. Other CNS areas which govern motivation are also important for the spontaneity and nature of movement. The classic antipsychotic-induced movement disorders are mediated primarily by their impact on D2 receptors located in the extrapyramidal tract. Agents with high D2 receptor affinity are associated with a greater risk of developing acute and late-onset movement disorders.

Table 34.3–1.
Selected Medications Associated with Movement Disorders: Impact on Relevant Neuroreceptors

Type (Subtype)	Generic (Brand) Name	D_2 Blockade	5-HT_2 Blockade	mACh Blockade
Antipsychotics				
Phenothiazine (Aliphatics)	Chlorpromazine (Thorazine)	Low	High	High
Phenothiazine (Piperidines)	Thioridazine (Mellaril)	Low	Med	High
	Mesoridazine (Serentil)	Low	Med	High
Phenothiazine (Piperazines)	Trifluoperazine (Stelazine)	Med	Med	Med
	Fluphenazine (Prolixin)	High	Low	Low
	Perphenazine (Trilafon)	High	Med	Low
Thioxanthenes	Thiothixene (Navane)	High	Med	Low
	Chlorprothixene (Taractan)	Med	High	Med
Dibenzoxazepines	Loxapine (Loxitane)	Med	High	Low
Butyrophenones	Haloperidol (Haldol)	High	Low	Low
	Droperidol (Inapsine)	High	Med	—
Diphenylbutylpiperidines	Pimozide (Orap)	High	Med	Low
Dihydroindolones	Molindone (Moban)	Med	Low	Low
Dibenzodiazepines	Clozapine (Clozaril)	Low	High	High
Benzisoxazoles	Risperidone (Risperdal)	High	High	Low
	Paliperidone (Invega)	High	High	Low
	Iloperidone (Fanapt)	High	High	Low
Thienobenzodiazepines	Olanzapine (Zyprexa)	Low	High	High
Dibenzothiazepines	Quetiapine (Seroquel)	Low/med	Low/med	Low
Benzisothiazolyls	Ziprasidone (Geodon)	Med	High	Low
Dibenzo-oxepino pyrroles	Asenapine (Saphris)	Med	High	Low
Quinolones	Aripiprazole (Abilify)	High (as partial agonist)	High	Low
	Brexpiprazole (Rexulti)	High (as partial agonist)	High	Low
	Cariprazine (Vraylar)	High (as partial	High	Low

		agonist)		
Benzisothiazol	Lurasidone (Latuda)	Med	High	Low
Nonantipsychotic psychotropics				
Metals	Lithium (Eskalith)	N/A	N/A	N/A
Anticonvulsants	Valproic acid (Depakene)	Low	Low	Low
Antidepressants	SSRIs	Low (except amoxapine)	(Varies)	(Varies)
Nonpsychotropics				
Antiemetics	Prochlorperazine (Compazine) (Gastrointestinal motility agents)	High	Med	Low
GIT motility	Metoclopramide (Reglan)	High	High	—

D_2, dopamine type 2; 5-HT_2, 5-hydroxytryptamine type 2; mACh, muscarinic acetylcholine; N/A, not applicable.

Adapted from Janicak PG, Marder S, Pavuluri M. *Principles and Practice of Psychopharmacotherapy*. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2011.

Table 34.3–2. DSM-5 Categories for Medication-Induced Movement Disorders

Neuroleptic or other medication-induced parkinsonism	Tremor, muscular rigidity, bradykinesia, or akinesia developing within a few weeks of starting or raising the dose of a neuroleptic medication or after reducing a medication used to treat extrapyramidal symptoms.
Neuroleptic malignant syndrome	Severe muscle rigidity, elevated temperature, altered consciousness, autonomic dysregulation, and other related findings (e.g., diaphoresis, dysphagia, incontinence, mutism, elevated or labile blood pressure, elevated creatine kinase levels) developing in association with the use of certain medications.
Medication-induced acute dystonia	Abnormal positioning or spasm of the muscles of the head, neck, limbs, or trunk developing within a few days of starting or raising the dose of certain medications or after reducing a medication used to treat extrapyramidal symptoms.
Medication-induced acute akathisia	Subjective complaints of restlessness accompanied by observed movements (e.g., movement of the legs, rocking from foot to foot, pacing, or inability to sit or stand still) developing within a few weeks of starting or raising the dose of certain medications or after reducing a medication used to treat extrapyramidal symptoms.
Tardive dyskinesia	Involuntary choreiform, athetoid, or rhythmic movements of the tongue, jaw, or extremities lasting at least a few weeks and developing in association with certain medications for at least a few months. May be for a shorter period of time in elderly individuals.
Tardive dystonia	Dystonia which occurs late in the course of treatment and persists for months to years.
Tardive akathisia	Akathisia which occurs late in the course of treatment and persists for months to years.
Medication-induced postural	Fine tremor which develops during attempts to maintain posture in association with the use of certain medications.

tremor	
Other medication-induced movement disorder	This category is for medication-induced movement disorders not explained by any of the disorders listed above. Examples include (1) parkinsonian, acute akathisic, acute dystonic, or dyskinetic movements associated with a nonneuroleptic medication; (2) a presentation that resembles NMS but is associated with a nonneuroleptic; and (3) tardive dystonia.
Antidepressant discontinuation syndrome	Movement symptoms which occur upon discontinuation or dose reduction of an antidepressant in use for at least one month.

Data from American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

These adverse effects are categorized into *hypokinetic disorders* (e.g., akinesia, bradykinesia) which involve a paucity of movement; or *hyperkinetic disorders* (e.g., dyskinesia, akathisia) which involve excessive movement. Such disorders are primarily mediated by the basal ganglia which includes the corpus striatum (caudate nuclei, putamen), globus pallidus (internal, external segments), substantia nigra (pars compacta, pars reticulata), and the subthalamic nucleus.

These structures modulate motor activity by a set of complex mechanisms, ultimately balancing inhibitory and stimulatory impulses via dopamine, as well as other neurotransmitters. The latter include serotonin (5-HT), acetylcholine (Ach), GABA and glutamate (Glu). In this context, the optimal range of D2 blockade in drug-naive patients may be very narrow. For example, studies indicate that D2 receptor occupancy in the corpus striatum with FGAs exceeding 78 percent is associated with increased risk of EPS, while levels less than 65 percent are associated with suboptimal efficacy. An important finding is that these levels can be achieved in first-episode patients with haloperidol doses as low as 1 to 5 mg per day.

NEUROLEPTIC OR OTHER MEDICATION-INDUCED PARKINSONISM

Clinical Presentation

The clinical presentations of antipsychotic-induced parkinsonism and idiopathic primary Parkinson disease are indistinguishable. Further, patients with subclinical Parkinson disease often experience a worsening of the disorder when exposed to an antipsychotic. The classic triad of symptoms includes *muscular rigidity* which can be continuous ("lead pipe") or cogwheeling (a discontinuous, ratchet-like muscle motion); *akinesia or bradykinesia* which are characterized by psychomotor retardation, a minimization of spontaneous motor activity, slowed speech, decreased arm movements during walking, shuffling gait, and masked facies. A *resting tremor* usually presents with rhythmic oscillations (3 to 6 Hz) affecting the head, tongue, jaw, legs, and extremities in a unilateral or symmetrical fashion. Although all three symptoms may be present, only

one is required for the diagnosis. Symptoms can develop within days to weeks of starting or increasing the dose of medication or after reduction of agents used to treat EPS. Not included in the criteria are other parkinsonian symptoms (e.g., poor balance, micrographia, shuffling gait) which are often seen in more severe cases and mimic the progression of Parkinson disease in a dose-related fashion. Women are more susceptible than men and the incidence is higher for those older than 60 years. Symptoms are not better explained by a mental or generalized medical disorder or an adverse effect of another medication.

Pathophysiology

Drug-induced parkinsonism results from a decrease in nigrostriatal dopamine activity. This can be induced by various actions, including depletion of dopamine in presynaptic stores (e.g., reserpine); dopamine-blocking agents (e.g., antipsychotics); and the atypical calcium blocking agent (e.g., cinnarizine [Stugeron]).

Treatment Options and Implications

This syndrome usually subsides with dose reduction or medication cessation. Since this is not always an option, alternative strategies include switching from a high-potency D2 blocking agent (e.g., fluphenazine [Prolixin]) to a lower-potency FGA (e.g., perphenazine [Trilafon]) or to an SGA (e.g., quetiapine [Seroquel]; aripiprazole [Abilify], a partial D2 agonist). Adjunctive anticholinergic agents such as benztropine (Cogentin) in divided doses of 2 to 8 mg per day; trihexyphenidyl (Artane) in divided doses of 6 to 10 mg per day; or diphenhydramine hydrochloride (Benadryl) in divided doses of 25 to 100 mg per day are usually effective in ameliorating this syndrome. Dopamine agonists such amantadine (Symmetrel) in divided doses of 100 to 300 mg per day may also be helpful if anticholinergics are not tolerated or in elderly patients. Thus, adequate management is usually achieved by modification of the activity of dopamine, acetylcholine, or both. In addition, preliminary data indicate that the addition of oral dehydroepiandrosterone (DHEA) to existing antipsychotic regimens may also improve pseudoparkinsonism symptoms. After control of these adverse effects is achieved, the gradual withdrawal of antiparkinsonian medication may be achieved without recurrence of EPS in some patients.

NEUROLEPTIC MALIGNANT SYNDROME

Clinical Presentation

The core symptoms of NMS include a generalized “lead pipe” *muscle rigidity* and an *elevated temperature* which may range from modest to severe (e.g., 106°F). *Altered consciousness* and *autonomic system*

dysregulation are also common characteristics. These complications typically occur within days after exposure to an antipsychotic. NMS, however, can vary in terms of symptom presentation (e.g., core symptoms are not always present with certain SGAs); progression (e.g., rapid to delayed in onset); and outcome (e.g., full recovery to death, particularly in unrecognized cases). Thus, the disorder can exhibit rapid or gradual onset of the following symptoms:

- Autonomic instability:
  - Diaphoresis
  - Tachycardia
  - Elevated or labile blood pressure
  - Urinary incontinence
  - Pallor
  - Sialorrhea
- Other symptoms:
  - Dysphagia
  - Dysarthria
  - Trismus
  - Myoclonus
  - Akinesia
  - Tremor
  - Changes in the level of consciousness, ranging from confusion to stupor to coma
- Laboratory findings:
  - Leukocytosis
  - Metabolic acidosis
  - Evidence of muscle injury (e.g., creatine kinase elevation)
  - Liver enzyme elevation

NMS-like symptoms should not be due to a general medical (e.g., thyrotoxicosis, malignant hyperthermia [MH]) or neurological (e.g., parkinsonian hyperthermia) disorder; the effects of another type of medication (e.g., serotonin syndrome); or be primary to the underlying psychiatric illness (e.g., malignant catatonia) (Table 34.3–3).

Table 34.3–3.
Neuroleptic Malignant Syndrome: Differential Diagnosis

Primary CNS disorders	CNS vasculitis Infarctions Infections Parkinson disease Status epilepticus Trauma

	Tumors
Systemic disorders	Acute porphyria
	Autoimmune disorders
	Dehydration
	Heat stroke
	Hyperthyroidism
	Infections
	Pheochromocytoma
	Tetanus
Psychiatric disorders	Idiopathic lethal catatonia
Medication-related disorders	Anticholinergic syndrome
	Drug intoxication
	Levodopa syndrome
	Malignant hyperthermia
	Serotonin syndrome

CNS, central nervous system.

Data from Zarrouf FA, Bhanot V. Neuroleptic malignant syndrome: don't let your guard down yet. *Curr Psychiatry.* 2007;6(8):89–94.

Pathophysiology

A hypodopaminergic state which interferes with temperature regulation is the postulated mechanism for the development of this syndrome. NMS is associated with dopamine antagonists, the abrupt discontinuation of antiparkinsonian medications and, rarely, the abrupt discontinuation of a dopaminergic agent. This suggests that fluctuation in dopamine binding may be important in its etiology. Because dopamine receptors are widely distributed in the CNS and spinal cord, it is likely that the various symptoms seen in NMS are accounted for by dopamine dysregulation in different areas. Thus, blockade of dopamine receptors in the corpus striatum causes muscle contraction and rigidity which then initiates heat generation; whereas blockade in the hypothalamus may lead to impaired heat regulation. The autonomic dysfunction commonly seen in advanced cases may be caused by dopamine blockade of spinal neurons. The fluctuations in mental status (e.g., delirium, coma) may be secondary to hyperthermia and/or effects on the mesocortical dopamine tract.

Abnormal calcium availability in muscle cells may also contribute, given the potential for muscle damage reflected by high creatine kinase elevations. There are no direct data which link NMS and MH, although cautious use of antipsychotics is urged in patients with a history of MH. NMS may also be due to a dysregulation of the sympathetic nervous system.

The occasional familial clustering of NMS leads to an analysis of genetic factors which may predispose to this condition. For example, the A1 allele of the dopamine D2 receptor gene was observed to be overrepresented in patients who develop NMS compared with those who do not. This allele reduces the density and function of the dopamine receptor in the striatum

and caudate nuclei, leading to reduced dopaminergic metabolism in those CNS areas with high concentrations of the neurotransmitter. By contrast, polymorphisms of the serotonin receptor gene did not appear to be associated with NMS but may increase the risk of developing acute EPS.

Treatment Options and Implications

NMS can be fatal or produce residual sequelae (e.g., cognitive or other neurological deficits). Therefore, management involves the aggressive use of supportive measures, as well as more specific interventions. Early identification is important and patients at higher risk should be carefully monitored. Factors which may increase the risk of developing NMS include:

- Concurrent medical and neuropsychiatric issues:
  - Dehydration
  - Psychomotor agitation
  - Exhaustion
  - Comorbid neurological disorders such as encephalitis and traumatic brain injury
  - Low serum iron levels
- Psychiatric diagnosis and symptom history:
  - Mood disorders
  - Pre-existing catatonia
  - History of NMS
- Medication-related:
  - Parenteral administration
  - High potency FGAs (e.g., haloperidol)
  - High doses of an antipsychotic
  - Dose increases or intermittent nonadherence
  - Adjunctive lithium treatment
- Demographic factors:
  - Male gender
  - Younger age

One easily treatable risk factor is dehydration. Thus, as a preventative measure, severely ill or catatonic patients on an antipsychotic should receive oral fluids or intravenous (IV) hydration. In addition, patients should avoid excessively high temperatures. Very mild cases of low-grade temperature increases or mild creatine kinase elevations may be treated in a psychiatric unit, while a rapid onset or worsening of symptoms requires transfer to a medical intensive care unit. Re-exposure to an antipsychotic can be considered after full resolution of the episode. While switching to an SGA (e.g., clozapine) is often recommended, NMS also occurs with these

agents. A past history of this syndrome is another reason patients are started or switched to SGAs. The relative risk of developing NMS between the two generations of antipsychotics, however, is difficult to compare.

Although NMS is one of the few life-threatening psychotropic-induced complications, its infrequent occurrence and severe nature make controlled studies impossible. The best available data support the following strategies:

1. *Prevention:* Antipsychotics should be used cautiously in patients with dehydration, catatonic symptoms, basal ganglionic disease, or a history of NMS.
2. *Early identification:* Patients with early signs should have their antipsychotic immediately discontinued, supportive measures initiated, and careful reintroduction of these agents considered only if absolutely necessary.
3. *Treatment:* The most important step is early recognition and prompt withdrawal of the offending agent. All cases require supportive interventions which may include IV hydration, temperature control and, in more severe cases, ventilator support or dialysis. In advanced cases, admission to an intensive care unit is required with a fatal outcome occurring in up to 20 percent of cases. A variety of other treatments may also help, including the use of dopamine agonists (e.g., bromocriptine, amantadine); muscle relaxants (e.g., benzodiazepines, dantrolene); and, in selected patients, ECT (Table 34.3–4). In patients who require an antipsychotic for their ongoing psychosis after an episode of NMS, several steps may reduce the chance of recurrence. These include waiting at least 2 weeks after symptoms resolve before reintroduction is attempted; resolving prior risk factors (e.g., dehydration, avoiding concomitant lithium); and considering alternative therapies when appropriate (e.g., mood stabilizers, ECT). If rechallenge with an antipsychotic is necessary, choosing an alternate agent, starting with low doses and titrating slowly can also reduce the risk of recurrence.

Table 34.3–4.
Neuroleptic Malignant Syndrome: Treatment

Intervention	Dosing	Effectiveness
Amantadine	200–400 mg/day by mouth in divided doses	Beneficial as monotherapy or in combination; decreases death rate
Bromocriptine	2.5 mg twice a day or three times a day by mouth; may increase to a total of 45 mg/day	Mortality reduced when used as a single agent or in combination; may worsen psychosis, reduce blood pressure; induce emesis
Levodopa/carbidopa	Levodopa 50–100 mg/day IV as continuous infusion	Case reports of dramatic improvement; may worsen psychosis, reduce blood pressure, or induce emesis
Dantrolene	1 mg/kg/day parenterally for 8 days;	Benefits may occur in minutes or hours

	continue by mouth for 7 additional days	as a single agent or in combination
Benzodiazepines	1–2 mg intramuscular as a test dose; if effective, switch to oral route; consider if catatonic symptoms are present	Has been reported effective when other agents fail
Electroconvulsive therapy	Reports of good outcome with both unilateral and bilateral treatment; response may occur in as few as three treatments	Effective when medications have failed; also may treat underlying psychiatric disorder; may worsen cognitive function
Supportive measures	IV hydration and oxygenation	Often effective as an initial approach early in the episode
	Cooling blankets	Can be used initially to lower temperature
	Ice packs	A second line measure if a more aggressive approach is needed to lower temperature
	Ice water enema	Another second line measure if a more aggressive approach is needed to lower temperature
	Antipyretics	May be used to reduce fever although its benefit is uncertain

Adapted from Berman BD. Neuroleptic malignant syndrome: a review for neurohospitalists. *Neurohospitalist.* 2011;1(1):41–47, Figure 31.1.

MEDICATION-INDUCED ACUTE DYSTONIA

Clinical Presentation

Acute dystonia occurs in up to 10 percent of patients, especially young men receiving a high potency FGA. Up to seven different symptoms may occur, although only one is required for the diagnosis. These symptoms include abnormal positioning of the head and neck (e.g., torticollis; retrocollis); spasms of the jaw muscles (e.g., trismus); impaired swallowing, speaking or breathing; thickened or slurred speech; tongue protrusion or dysfunction; deviated eyes in any direction (e.g., oculogyric crisis); and abnormal positioning of limbs or trunk (e.g., opisthotonos). These symptoms are usually quite frightening to patients who may develop the syndrome within minutes of receiving parenteral, high-potency FGAs; more commonly, however, symptoms usually develop within days of starting or increasing the oral antipsychotic dose or reducing a medication used to treat EPS. Rarely, dystonia can be life-threatening if laryngeal muscles are affected.

Another complication with acute dystonic reactions is the frequent occurrence of long-term treatment nonadherence. After such an experience, patients commonly report being “allergic” to the offending medication. Careful history taking often reveals the true nature of the event, but patients are understandably reluctant to resume treatment with the implicated drug. Such a history may preclude future use of specific

agents unless considerable time and energy are spent educating the patient regarding the probable cause of the reaction and how it can be prevented with anticholinergic agents. If a patient has a history of acute dystonia, the prophylactic use of such medications should be considered, even with the use of less potent antipsychotics.

The risk of acute dystonia is increased in young males. Recent use of cocaine, which depletes dopamine, is also associated with an increased risk. Some patients who are covertly nonadherent to oral tablets or capsules may develop acute dystonia when they are switched to a liquid concentrate or parenteral formulation. The risk of dystonia diminishes quickly after a steady dose is reached and ongoing use of anticholinergic medication is usually not required. Patients at risk for developing dystonia, however, are also at risk for developing other forms of EPS which may require continued anticholinergic management. A family history of movement disorders, as well as a prior history of dystonia, is predictive of a higher rate of acute dystonia and other forms of EPS.

Pathophysiology

Although the mechanism of acute dystonia is not understood, the sudden onset of symptoms and its limitation to the initiation of treatment or dose increase suggest that acute saturation of D2 receptors plays a role. Preclinical studies suggest sigma-1 and sigma-2 receptors, found in the motor regions of the brain and involved in the control of posture and movement, may be involved in all forms of dystonia, including medication-induced.

Treatment Options and Implications

Patients should always be informed of the risk of muscle spasms associated with these drugs and encouraged to report such events immediately. With severe reactions or laryngeal dystonia which can compromise breathing, immediate recognition and treatment are essential. Parenteral (intramuscular [IM] or IV) benztropine (1 to 2 mg) or diphenhydramine (50 mg) typically produce rapid relief. Prophylactic use of oral anticholinergic medication, as well as the ready availability of the parenteral form, is strongly encouraged when use of a high-potency FGA is considered.

Anticholinergic medication has the greatest effect on antipsychotic-induced acute dystonia, less of an effect on tardive dystonia, and little effect on idiopathic dystonia. Because tardive dystonia can be a particularly disabling disorder, there is some support for the use of clozapine which may even improve this condition. Botulinum toxin (Botox) injections also offer symptomatic improvement for tardive dystonia. The prevalence of this adverse effect diminishes when acute agitation is treated with a

combination of intramuscular haloperidol (e.g., 2 to 5 mg) plus lorazepam (Ativan) (e.g., 2 mg). It is likely that both the lower haloperidol dose and lorazepam's sedative and peripheral muscle-relaxant effects contribute to attenuation of this painful and frightening event. Ziprasidone (Geodon), olanzapine (Zyprexa), and aripiprazole all have approved acute parenteral (1M) formulations which appear to carry a lower risk for acute dystonia and may be alternatives to haloperidol plus lorazepam.

MEDICATION-INDUCED ACUTE AKATHISIA

Clinical Presentation

Akathisia (from the Greek word meaning "never to sit down") usually develops after starting or increasing the dose of medication or after decreasing the use of medications used to treat EPS. The incidence is estimated to be between 20 and 45 percent for patients on an antipsychotic. This syndrome is also reported with various antidepressants, particularly the SSRIs, possibly mediated by serotonin agonism of the dopamine system. It is also associated with severe distress, suicidality, a higher risk for developing tardive dyskinesia, and a poor treatment outcome.

Acute akathisia typically occurs within weeks of starting a medication, lasts less than 6 months and is characterized by intense dysphoria and an inner sense of restlessness. In addition, at least one of the following observable movements should be present: fidgeting or swinging of the legs, rocking from foot to foot while standing, pacing, lifting the feet as if marching in place, crossing and uncrossing the legs when sitting, or the inability to sit or stand in one place for several minutes. *Chronic akathisia* involves persistent symptoms (i.e., 6 months or longer) after the introduction of or last change in medication dose. It is characterized by stereotypic movements often in the presence of limb or orofacial dyskinesias. *Tardive akathisia* is often associated with tardive dyskinesia and manifests with a delayed onset which may persist even after medication cessation. *Withdrawal akathisia* can occur within 6 weeks of switching or stopping an antipsychotic or upon stopping an anticholinergic agent. The most common metric used to assess for this problem is the Barnes Akathisia Scale (BAS) which quantifies both subjective and objective components.

Severe cases are easily recognized, whereas mild cases are more difficult to separate from symptoms of the primary psychiatric disorder, especially those in which anxiety and motor restlessness are prominent. Thus, patients who present with pacing or running may be misdiagnosed with manic excitement or psychotic agitation and receive additional doses of an antipsychotic, risking exacerbation of the movement disorder. By contrast, lowering the dose may unmask tardive akathisia. This adverse effect is also

more common with high-potency FGAs and may lead to inadequate adherence with future treatment recommendations. The symptoms of akathisia should not be better explained as part of a mental, neurological, or general medical disorder.

Pathophysiology

The pathophysiology of akathisia is uncertain. There is evidence, however, that decreasing dopamine tone in the ventral tegmental area and substantia nigra, which in turn alters interactions between the mesolimbic, mesocortial, and spinal dopamine/norepinephrine systems, is contributory. GABAergic, cholinergic, adrenergic, serotonergic, and neuropeptide mechanisms are also implicated. Risk factors can be divided into:

- Drug-related:
  - Higher doses
  - Rapid dose escalation
  - Higher potency agents
- Individual-related:
  - Older age
  - Female
  - Mood disorder
  - Preponderance of negative symptoms
- Symptom-related:
  - Iron deficiency
  - Cognitive dysfunction

Treatment Options and Implications

A slow dose titration schedule, dose reduction, or switching to a less potent FGA or an SGA are the preferred initial options. This should be done under close supervision to avoid exacerbation of psychosis or precipitating withdrawal symptoms, including akathisia. Clozapine may be an option for patients with chronic, persistent akathisia. In cases in which there is uncertainty, benzodiazepines are preferable to increasing the antipsychotic dose.

β-*Blockers* (e.g., propranolol [Inderal]) are the most commonly prescribed medication for this problem. Propranolol's mechanism of action may relate to its alteration of both noradrenergic and serotonergic inputs into the DA system. *5-HT2a/c antagonists* such as mirtazapine (15 mg per day), trazodone (100 mg per day), mianserin (15 mg per day), or cyproheptadine (8 to 16 mg per day) may produce benefits comparable to propranolol with better possible tolerability. *Anticholinergic* agents (e.g., trihexyphenidyl [2 to 10 mg per day]; benztropine [1 to 8 mg per day]) and

benzodiazepines (e.g., lorazepam at 1 to 2 mg per day) are second and third level choices, respectively.

In resistant cases, other *GABAergic agents* (e.g., gabapentin; 3,000 mg per day); *amantadine* (a weak glutamate receptor antagonist and DA-enhancing agent; 100 mg per day) or *clonidine* (a selective α_2-adrenergic agonist; up to 0.15 mg per day) may help. In two randomized controlled trials (RCTs), the addition of *vitamin B6* (1,200 mg per day) for 5 days significantly improved subjective and global subscales of the BAS.

TARDIVE DYSKINESIA

Clinical Presentation

Symptoms of tardive dyskinesia should develop after at least 2 months of exposure to a dopamine blocking agent (or 1 month if the patient is 60 years or older); while the patient is on medication or within 4 weeks of withdrawal from an oral agent (8 weeks from long-acting injectable medication); and be present for at least 4 weeks. The patient usually demonstrates involuntary movements of the tongue (e.g., twisting, protrusion), jaw (e.g., chewing), lips (e.g., smacking, puckering), trunk, or extremities. Patterns of movements may include rapid, jerky, and nonrepetitive movements (i.e., choreiform); slow, continuous, and sinuous movements (i.e., athetoid); or rhythmic movements.

Onset of the disorder may be gradual, occurring over a period of months to years. Symptoms are often seen in older patients who have poorly fitting dentures or other reasons for dyskinetic movements. In older patients (55 years or older), the onset of tardive dyskinesia is 25 percent after 1 year, 34 percent after 2 years, and 53 percent after 3 years. This rate of onset is three to five times higher than in younger patients. Patients with multi-infarct dementia and a psychiatric diagnosis are at greater risk than patients with Alzheimer disease or other organic mental disorders. Tardive dyskinesia may also occur in never-medicated patients with schizophrenia. Risk factors include increasing age, pre-existing parkinsonism, previous brain damage, antipsychotic treatment duration, and exposure to FGAs. These symptoms should not be due to a general medical or neurological disorder, or present as an adverse effect of another class of medication.

Pathophysiology

Sustained D2 receptor blockade resulting in receptor hypersensitivity is the most common hypothesis explaining the development of tardive dyskinesia. Clozapine's decreased propensity to induce tardive dyskinesia and its multiple neuroreceptor effects (e.g., binding to 5-HT1C, 5-HT2, and 5-HT3 receptors) suggest a role for serotonin as well. Serotonin-containing neurons interact with dopaminergic neurons in the substantia nigra and ventral tegmental areas. Genetic studies indicate a possible relationship

with polymorphisms in the DA2 receptor, DA3 receptor, dopamine transporter (DAT1), and the serotonin 2A receptor genes. Oxidative stress and cell death secondary to increased glutamatergic neurotransmission caused by blockade of presynaptic dopamine receptors is also postulated.

Treatment Options and Implications

Since the advent of SGAs, fewer episodes and less severe symptoms of tardive dyskinesia are now seen. For this reason, patients on FGAs are generally switched to a newer agent if possible. Further, there is evidence that agents such as clozapine may improve symptoms of tardive dyskinesia without causing progression of the disorder. For example, some studies of clozapine report reduced symptom severity in tardive syndromes by up to 50 percent. Patients on an FGA who develop tardive dyskinesia and are switched to an SGA have demonstrated improvement in their symptoms, but which specific agents are preferable is still unclear. Although a small percentage of cases progress to florid and potentially disabling or life-threatening intensity, others remain stable and some show spontaneous improvement.

Other strategies to address the problem include a *gradual discontinuation* in patients whose need for antipsychotic medication is not clearly established (e.g., a psychotically depressed patient who is maintained on an antipsychotic after apparent full recovery). Because a certain number of these patients will relapse, however, medication discontinuance requires close monitoring. Patients taken off high-potency FGAs usually require a cross-taper because tardive dyskinesia symptom exacerbation may occur after an abrupt change to a low-potency agent. *Discontinuance, dose reduction, or switch* to another agent may improve, stabilize, or worsen symptoms. Use of adjunctive agents is of limited benefit. These include *vitamin B6* and *procholinergic agents* (e.g., donepezil [Aricept]). *Ondansetron* (Zofran), a selective 5-HT3 receptor antagonist, may reduce symptoms of tardive dyskinesia and also exert a beneficial effect on psychotic symptoms. Preliminary studies indicate *cyproheptadine* (Periactin), a 5-HT and histamine antagonist, and *levetiracetam* (Keppra) may improve tardive dyskinesia. The use of *deep brain stimulation* (DBS) for severe and refractory tardive dyskinesia offers hope to those who are severely disabled. Systematic reviews and studies of galantamine (Razadyne), vitamin E, calcium channel blockers, benzodiazepines, lithium, estrogen, and essential fatty acids do not support their use for tardive dyskinesia.

MEDICATION-INDUCED POSTURAL TREMOR

Clinical Presentation

Motor tremors can develop after starting a medication such as lithium,

valproic acid (Depakene), α-adrenergic blockers, stimulants, dopaminergic agonists, caffeine, theophylline (Bronkodyl), an antipsychotic, or an antidepressant. Typically, it is a fine 8- to 12-Hz postural tremor affecting the limbs, head, mouth, or tongue. Other medical causes should be excluded and the tremor should not be part of an antipsychotic-induced pseudoparkinsonism which usually occurs at rest, is slower and, unlike true parkinsonian tremor, tends to worsen with movement.

Pathophysiology

The hypothesized mechanism of medication-induced postural tremor varies based on the class of drug implicated. For example, stimulants may cause tremors due to the resulting hyperadrenergic state.

Treatment Options and Implications

This diagnosis is commonly associated with the use of mood stabilizers such as lithium. For example, tremor is often present, may be dose related and can herald the onset of early lithium toxicity. Thus, checking the lithium plasma level and looking for other signs of toxicity is advisable. If the tremor is an isolated finding not associated with toxicity, it may improve with time. Although the dose of lithium may also be reduced, this may compromise its therapeutic effect. Tremor is often worsened by emotional factors such as anxiety, stress or fatigue; medical issues such as alcohol withdrawal, hypoglycemia, thyrotoxicosis, pheochromocytoma, or hypothermia; or by the addition of other medications. In cases where a pre-existing tremor worsens with the addition of a medication, the diagnosis of other medication-induced movement disorder should be considered. *Reduction of dose or switching* to another agent in a different class may help. In some cases, a *benzodiazepine or β-blocker* can ameliorate the tremor.

TARDIVE DYSTONIA/TARDIVE AKATHISIA

Clinical Presentation

Tardive dystonia and tardive akathisia are movement symptoms which arise later in treatment with an antipsychotic and may persist for months or years despite the reduction or discontinuation of the patient's medication. Young male patients have a higher risk. This syndrome is differentiated from acute dystonia and acute akathisia based on the time to onset. Both tardive dystonia and akathisia are associated with a higher incidence of tardive dyskinesia.

Pathophysiology

There is little known about the pathophysiology of tardive dystonia and

tardive akathisia. Tardive dystonia appears to involve impairment of the basal ganglia and tardive akathisia may involve dysfunction of the pontine region of the brain.

Treatment Options and Implications

Both of these movement disorders are preventable with the use of SGAs (e.g., clozapine; quetiapine). If the patient has symptoms while on or after discontinuing an antipsychotic, an increase in its dose or immediate reintroduction of the agent may be necessary to increase dopamine blockade. Such interventions can achieve short-term relief.

One potential treatment for refractory tardive dystonia is DBS. In one report, 60 patients with symptoms of various types of dystonia underwent DBS with leads placed in the globus pallidus pars interna. Using the Burke–Fahn–Marsden Dystonia Rating Scale (BFM), patients were assessed preoperatively, postoperatively, and for up to 5 years of follow-up. Improvement in the BFM severity and disability scores were 43 percent and 27 percent, respectively by 6 months and were sustained thereafter. Thus, DBS may be a potential option for patients with severe, treatment-refractory, medication-induced tardive dystonia.

OTHER MEDICATION-INDUCED MOVEMENT DISORDERS

Clinical Presentation

These disorders do not fit into more specific categories. For example, acute dystonia may occasionally follow the use of a benzodiazepine (e.g., midazolam) or antidepressants (e.g., amoxapine). Other forms of EPS which may occur as an idiosyncratic reaction to a nonantipsychotic medication are also coded under this category. NMS-like syndromes and tardive dystonia not associated with antipsychotic exposure may be coded under this category as well.

Pathophysiology

Genetic variability may account for idiosyncratic reactions seen in some patients. The ongoing research and rapid growth in knowledge about these variations in the CNS and their relationship to treatment response and adverse effects will help to clarify these questions.

Treatment Options and Implications

These disorders are infrequent and discontinuation of the offending agent is always the preferred option. Careful history taking and collateral information may reveal exposure to more typical causes of movement disorders (e.g., patient may have taken another's medication or is exposed to an antiemetic phenothiazine).

ANTIDEPRESSANT DISCONTINUATION SYNDROME

Clinical Presentation

While movement symptoms are not common withdrawal phenomena with antidepressants, abrupt discontinuation or rapid dose reduction may produce them. Usually, the patient is typically on the drug for a month or more and movement symptoms arise within days after discontinuation or tapering. This may occur with various classes of antidepressants, including SSRIs (e.g., citalopram, fluoxetine); serotonin and noradrenergic reuptake inhibitors (e.g., duloxetine, venlafaxine); TCAs (e.g., desipramine, imipramine); MAOIs (e.g., tranylcypromine, isocarboxazid); as well as other agents (e.g., mirtazapine). Common symptoms include *somatic* (e.g., dizziness, headache, nausea); *sensory* (e.g., numbness, ringing in the ears, "electric shocks," flashes of light); *mood* (e.g., anxiety, irritability); and *abnormal movements* (e.g., EPS). Some movement symptoms occur with all antidepressants, while others are specific to the class of agent and include:

- SSRIs:
  - Tremor
  - Akathisia
  - Parkinsonism
- TCAs:
  - Tremor
  - Akathisia
  - Parkinsonism
- MAOIs:
  - Catatonia
  - Myoclonic Jerks
- Serotonin and norepinephrine reuptake inhibitors (SNRIs):
  - Akathisia

Most of these symptoms occur upon discontinuation of an SSRI or TCA. The latter class has a higher tendency to cause parkinsonian symptoms which are managed with anticholinergics (e.g., benztropine) if a patient does not wish to restart their antidepressant. MAOIs, especially tranylcypromine, may cause more severe movement symptoms such as myoclonic jerks and catatonia which require acute management with benzodiazepines (e.g., lorazepam). Akathisia may occur with discontinuation from SSRIs, SNRIs, and TCAs and is managed with β-blockers (e.g., propranolol), anticholinergics, and in some cases, benzodiazepines.

Pathophysiology

While there are several hypotheses as to the cause of discontinuation symptoms with antidepressants, there is no clear consensus on their pathophysiology. There is agreement, however, that an increased propensity to develop these symptoms is associated with shorter half-life agents. For example, paroxetine is associated with the highest risk for the discontinuation syndrome.

Treatment Options and Implications

Reinstatement of the antidepressant typically resolves these symptoms. When no further treatment with an antidepressant is required, however, management of these symptoms depends on their severity. Mild symptoms often resolve spontaneously and require no treatment beyond counseling the patient. More moderate-to-severe symptoms may require reinstating the agent to quickly alleviate symptoms. The original antidepressant can be reintroduced or another antidepressant with a similar mode of action may be substituted. Once stabilized, a more gradual tapering is then initiated. It is important to take various factors into account such as the dose and type of antidepressant agent to determine the appropriate tapering duration.

While the symptoms of the discontinuation syndrome are manageable, they can be prevented by educating the patient on the importance of gradually tapering the dose of medication. It is also essential to educate patients on the importance of adhering to their medication and the potential for adverse events upon their cessation.

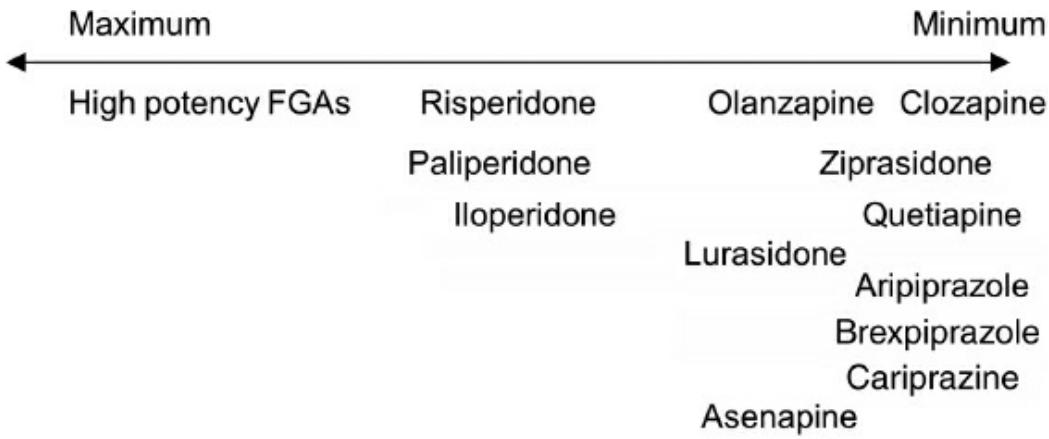

FIGURE 34.3–1. Continuum of acute extrapyramidal symptoms (EPS). (Adapted from Janicak PG, Marder S, Pavuluri M. *Principles and Practice of Psychopharmacotherapy*. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2011.)

SUMMARY

The impact of medication-induced movement disorders should not be underestimated. These conditions are often debilitating, even life-threatening; diminish the effectiveness of treatment by precluding an adequate trial; contribute to nonadherence; simulate or accentuate

negative, cognitive or mood symptoms; and increase the risk for irreversible, later-onset abnormal movements. Also of concern is that some of these adverse effects (e.g., pseudoparkinsonism, tardive dyskinesia, akathisia) are associated with an increase in suicidality (e.g., ideation, behavior, attempt). As a result, the choice of antipsychotic is often based in large part on its potential to produce such adverse effects. Since SGAs usually produce fewer acute and chronic movement symptoms than FGAs, patients are more likely to initially receive one of these agents. Further, patients with concurrent movement disorders (e.g., Parkinson disease), significant drug-induced movement adverse effects, or who are at increased risk for developing tardive syndromes are often switched to an SGA. The substantial variation in the experience of such adverse effects, however, always requires an individualized risk–benefit assessment before choosing a specific agent. Figure 34.3–1 depicts the relative risk of acute EPS with these agents in comparison to high-potency FGAs.

Diagnosis, phase of the illness, age, gender, and race affect the likelihood of developing movement disorders. For example, individuals with first- or early-onset psychosis are at greatest risk for developing movement symptoms than more chronically ill patients; older white females with a mood disorder are at higher risk to develop tardive dyskinesia; younger black males are at greatest risk for developing acute dystonia; and more severe acute EPS in the context of catatonic symptoms may predispose to developing NMS.

The management of medication-induced movement disorders includes the appropriate prescription of antipsychotic agents (e.g., accurate diagnosis); optimal dosing; medication change; or adjunctive treatments (e.g., anticholinergic agents, benzodiazepines, and β-blockers which are most effective for antipsychotic-induced pseudoparkinsonism). Other agents may benefit NMS, including bromocriptine (Parlodel) and dantrolene (Dantrium). Management may also include the selection of an SGA with less likelihood of producing or even possibly improving these syndromes.

FUTURE DIRECTIONS

Medication-induced movement disorders require a careful consideration of the risk–benefit ratio when initiating antipsychotics or other potential offending agents, limiting their use, and encouraging the development of safer alternatives. Developments in the current understanding of brain function through advancements in neuroimaging and genetics will lead to safer and better tolerated agents. For example, pharmacogenetics and pharmacogenomics are increasingly helping to explain how the various alleles of receptor-related genes impact response and/or influence the development of acute and long-term medication-related adverse effects. Since medications which are life-saving for some patients but may be

disabling or even life-threatening for others, such advances will permit a better risk assessment and a better selection process from the increasing number of available agents. For now, predicting and minimizing drug-related movement disorders are primarily based on multiple clinical factors (e.g., diagnosis; phase of the illness; race; age). Although the recognition and management of medication-induced adverse effects are important, it is always better to avoid or minimize them with the judicious choice of treatment(s). The goal is to improve adherence, thus maximizing efficacy, functionality, and quality of life.

REFERENCES

*American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

*Aquino CC, Lang AE. Tardive dyskinesia syndromes: current concepts. *Parkinsonism Relat Disord*. 2014;20 Suppl 1:S113–S117.

*Atsariyasing W, Goldman M. Management of medication-related adverse effects. In: Janicak PG, Marder SR, Tandon R, Goldman M, eds. *Schizophrenia: Recent Advances in Diagnosis and Treatment*. New York: Springer; 2014:225–262.

Bai YM, Yu SC, Chen JY, Lin CY, Chou P, Lin CC. Risperidone for pre-existing severe tardive dyskinesia: a 48-week prospective follow-up study. *Int Clin Psychopharmacol*. 2005;20(2):79–85.

Berman BD. Neuroleptic malignant syndrome: a review for neurohospitalists. *Neurohospitalist*. 2011;1(1):41–47.

Bratti IM, Kane JM, Marder SR. Chronic restlessness with antipsychotics. *Am J Psychiatry*. 2007;164(11):1648–1654.

Buckley PF. Treating movement disorders and akathisia as side effects of antipsychotic pharmacotherapy. *J Clin Psychiatry*. 2008;69(5):e14.

Chang EF, Schrock LE, Starr PA, Ostrem JL. Long-term benefit sustained after bilateral pallidal deep brain stimulation in patients with refractory tardive dystonia. *Stereotact Funct Neurosurg*. 2010;88:304–310.

Damier P, Thobois S, Witjas T, et al. Bilateral deep brain stimulation of the globus pallidus to treat tardive dyskinesia. *Arch Gen Psychiatry*. 2007;64(2):170–176.

Davis JM, Janicak, PG, Sakkas P, Gilmore C, Wang Z. Electroconvulsive therapy in the treatment of the neuroleptic malignant syndrome. *Convuls Ther*. 1991;7(2):111–120.

Dayalu P, Chou KL. Antipsychotic-induced extrapyramidal symptoms and their management. *Expert Opin Pharmacother*. 2008;9:1451–1462.

FitzGerald JJ, Rosendal F, de Pennington N, et al. Long-term outcome of deep brain stimulation in generalised dystonia: a series of 60 cases. *J Neurol Neurosurg Psychiatry*. 2014;85(12):1371–1376.

Gunes A, Dahl ML, Spina E, Scordo MG. Further evidence for the association between 5-HT2C receptor gene polymorphisms and extrapyramidal side effects in male schizophrenic patients. *Eur J Clin Pharmacol*. 2008;64:477–482.

Gunes A, Scordo MG, Jaanson P, Dahl ML. Serotonin and dopamine receptor gene

polymorphisms and the risk of extrapyramidal side effects in perphenazine-treated schizophrenic patients. *Psychopharmacology(Berl)*. 2007;190:479–484.

Güzey C, Scordo MG, Spina E, Landsem VM, Spigset O. Antipsychotic-induced extrapyramidal symptoms in patients with schizophrenia: associations with dopamine and serotonin receptor and transporter polymorphisms. *Eur J Clin Pharmacol*. 2007;63:233–241.

Haddad PM, Anderson IM. Recognizing and managing antidepressant discontinuation symptoms. *Advances in Psychiatric Treatment*. 2007;13:447–457.

Hosenbocus S, Chahal R. SSRIs and SNRIs: a review of the discontinuation syndrome in children and adolescents. *J Can Acad Child Adolesc Psychiatry*. 2011;20:60–67.

Janicak PG, Marder S, Pavuluri M. *Principles and Practice of Psychopharmacotherapy*. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2011:65–180.

*Kapur S, Zipursky R, Jones C, Remington G, Houle S. Relationship between dopamine D(2) occupancy, clinical response, and side effects: a double-blind PET study of first-episode schizophrenia. *Am J Psychiatry*. 2000;157:514–520.

*Koning JP, Tenback DE, van Os J, Aleman A, Kahn RS, van Harten PN. Dyskinesia and parkinsonism in antipsychotic-naive patients with schizophrenia, first-degree relatives and healthy controls: a meta-analysis. *Schizophr Bull*. 2010;36(4):723–731.

Kornhuber J, Schultz A, Wiltfang J, et al. Persistence of haloperidol in human brain tissue. *Am J Psychiatry*. 1999;156:885–890.

Krystkowiak P, du Montcel ST, Vercueil L, et al. Reliability of the Burke–Fahn–Marsden scale in a multicenter trial for dystonia. *Mov Disord*. 2007;22(5):685–689.

Lencer R, Eismann G, Kasten M, et al. Family history of primary movement disorders as a predictor for neuroleptic-induced extrapyramidal symptoms. *Br J Psychiatry*. 2004;185:465–471.

Matsumoto RR, Pouw B. Correlation between neuroleptic binding to sigma(1) and sigma(2) receptors and acute dystonic reactions. *Eur J Pharmacol*. 2000;401:155–160.

*Meltzer HY, Alphs L, Green AI, et al. Clozapine treatment for suicidality in schizophrenia: International Suicide Prevention Trial (InterSePT). *Arch Gen Psychiatry*. 2003;60:82–91.

*Miller DD, Caroff SN, Davis SM, et al. Extrapyramidal side-effects of antipsychotics in a randomised trial. *Br J Psychiatry*. 2008;193:279–288.

Modi S, Dharaiya D, Schultz L, Varelas P. Neuroleptic malignant syndrome: complications, outcomes, and mortality. *Neurocrit Care*. 2016;24(1):97–103.

Pappa S, Dazzan P. Spontaneous movement disorders in antipsychotic-naive patients with first-episode psychoses: a systematic review. *Psychol Med*. 2009;39(7):1065–1076.

Psychoactive Drug Screening Program (PDSP) Database. Available at http://kidb.cwru.edu/pdsp.php. Accessed on July 2014.

Rajamani B, Kumar Y, Rahman SM. Neuroleptic malignant syndrome. *J Family Med Prim Care*. 2016;5(1):178–180.

*Rana AQ, Chaudry ZM, Blanchet PJ. New and emerging treatments for symptomatic tardive dyskinesia. *Drug Des Devel Ther*. 2013;7:1329–1340.

*Renoir T. Selective serotonin reuptake inhibitor antidepressant treatment discontinuation syndrome: a review of the clinical evidence and possible mechanisms involved. *Front Pharmacol*. 2013;4(45):1–10.

Sethuram K, Gedzior J. Akathisia: case presentation and review of newer treatment agents. *Psychiatr Ann*. 2014;44(8):391–396.

Strous RD, Stryjer R, Maayan R, et al. Analysis of clinical symptomatology, extrapyramidal symptoms and neurocognitive dysfunction following dehydroepiandrosterone (DHEA) administration in olanzapine treated schizophrenia patients: a randomized, double-blind placebo controlled trial. *Psychoneuroendocrinology*. 2007;32:96–105.

Suzuki T. Which rating scales are regarded as 'the standard' in clinical trials for schizophrenia? A critical review. *Psychopharmacol Bull*. 2011;44(1):18–31.

Suzuki A, Kondo T, Otani K, et al. Association of the TaqI A polymorphism of the dopamine D(2) receptor gene with predisposition to neuroleptic malignant syndrome. *Am J Psychiatry*. 2001;158:1714–1716.

Tse L, Barr AM, Scarapicchia V, Vila-Rodriguez F. Neuroleptic malignant syndrome: a review from a clinically oriented perspective. *Curr Neuropharmacol*. 2015;13(3):395–406.

Wilson E, Lader MA. review of the management of antidepressant discontinuation symptoms. *Ther Adv Psychopharmacol*. 2015;5(6):357–368.

▲ 34.4 α2-Adrenergic Receptor Agonists: Clonidine, Clonidine Extended Release, Guanfacine, and Guanfacine Extended Release

ERIC HOLLANDER, M.D., AND JENNIFER N. PETRAS, M.D.

HISTORY

Clonidine (Catapres) was synthesized in the early 1960s and found to produce vasoconstriction mediated by α-receptors. The drug was initially developed as a topical nasal decongestant. During clinical testing it was found to cause hypotension, sedation, and bradycardia, and it was subsequently used as a centrally active antihypertensive. Due to its effects on norepinephrine and other neurotransmitter systems, and functioning of the prefrontal cortex (PFC), clonidine has been studied in several psychiatric disorders including tic disorders, ADHD, PTSD, and opiate, nicotine, and alcohol withdrawal. Guanfacine (Tenex) is also an α-adrenergic receptor agonist. It has a longer half-life and may have fewer side effects due to its differential affinity for certain receptor subtypes.

CHEMISTRY

Clonidine is an imidazoline derivative and exists as a mesomeric compound. Its chemical name is 2-(2,6-dichlorophenylamino)-2-imidazoline hydrochloride; the structural formula is shown in Figure 34.4–1. Clonidine hydrochloride is an odorless, bitter, white, crystalline substance soluble in water and alcohol. Its molecular weight is 266.56.

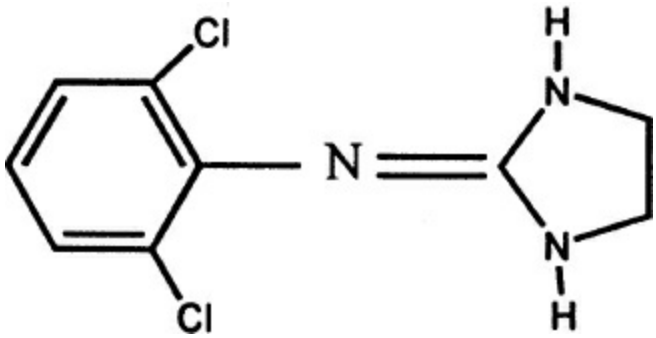

FIGURE 34.4–1. Molecular structure of clonidine.

The chemical name of guanfacine is *N*-amidino-2-(2,6-dichlorophenyl) acetamide monohydrochloride, and its molecular weight is 282.56. Guanfacine hydrochloride is a white to off-white powder sparingly soluble in water and alcohol and slightly soluble in acetone. Guanfacine's structural formula is shown in Figure 34.4–2.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

Absorption. Clonidine immediate release (CLON-IR) is well absorbed after oral administration and its bioavailability is nearly 100 percent. The plasma level peaks between 1 and 3 hours after an oral dose. Clonidine extended release (Kapvay) (CLON-XR) is well absorbed after oral administration with a bioavailability of about 90 percent, but is designed to be absorbed more slowly with peak plasma levels achieved at 6 to 7 hours. Due to differing pharmacokinetic characteristics, utilization of the same dose of CLON-IR and CLON-XR will result in different drug exposure. Comparisons across studies suggest peak plasma concentrations are 50 percent lower for CLON-XR compared with CLON-IR. Absorption of CLON-XR is not impacted by food intake. Guanfacine immediate release (GUAN-IR) is also well absorbed orally with a bioavailability of about 80 percent and peak plasma levels are seen in 1 to 4 hours. Guanfacine extended release (Intuniv) (GXR) has a relative bioavailability of 57 percent compared with the GUAN-IR formulation, therefore the XR and IR formulations are not interchangeable. GXR reaches peak plasma levels in 5 to 6 hours after oral administration. Absorption of GXR is significantly affected by food intake, with a 75 percent increase in peak plasma levels when administered with a high-fat meal. Administration with a high-fat meal should be avoided.

Distribution. Clonidine and guanfacine formulations cross the blood–brain barrier. Clonidine is present in breast milk. Clonidine is between 20 and 40 percent protein bound and has a volume of distribution of 2.1 L/kg. Guanfacine is approximately 70 percent protein bound and has a volume of

distribution of 6 L/kg, which suggests a high distribution of drug to the tissues. Clonidine is known to cross the placenta and is present in breast milk. It is not known if guanfacine crosses the placenta or is excreted in breast milk, although guanfacine has been shown to cross the placenta in rats.

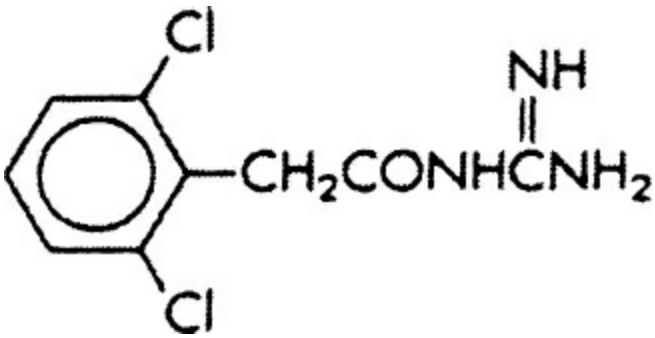

FIGURE 34.4–2. Molecular structure of guanfacine.

Metabolism and Elimination. The elimination half-life (time required for the concentration of the drug to reach half its original value) of both the immediate-release and extended-release clonidine range from 6 to 24 hours, with a mean of 12 hours. After oral administration of clonidine, approximately 40 to 60 percent of the dose is present, unchanged in the urine. When renal function is severely impaired, the half-life can increase to approximately 41 hours. About 50 percent of the absorbed dose is metabolized in the liver. The elimination half-life of both guanfacine immediate release and guanfacine extended release is approximately 17 hours, and steady-state blood levels are reached within 4 days. Guanfacine is metabolized by hepatic cytochrome p450 (CYP) 3A4 microsomal enzymes. Exposure to guanfacine may potentially be increased by concurrent use of CYP3A4 inhibitors (clarithromycin, grapefruit juice, itraconazole) or decreased by concurrent use of CYP3A4 inducers (rifampin, St. John wort, phenytoin). Approximately 50 percent of the dose is eliminated in the urine as unchanged drug. In renal insufficiency clearance is reduced, but drug levels are only slightly increased. Guanfacine is poorly dialyzed. Both clonidine and guanfacine are metabolized more rapidly in children and, therefore, the half-life is decreased. When the clonidine transdermal patch is removed, plasma concentrations remain stable for approximately 8 hours and then decline gradually over a few days.

Pharmacodynamics

Mechanism of Action. The α_2-receptor is one of five adrenergic receptors (α_1, α_2, β_1, β_2, β_3). The α_1-receptors are located in the periphery, while the α_2-receptors are located centrally and peripherally. In the CNS,

they are located both pre- and postsynaptically. The α_2-receptors have three distinct subtypes (A, B, and C), each with its own area of anatomical predominance. The α_2A-adrenoreceptor is the most predominant and broadly distributed subtype being concentrated in the PFC, locus ceruleus (LC), amygdala, septum, hippocampus, and hypothalamus. The α_2B-adrenoreceptor is mainly restricted to the thalamus. The α_2C-adrenoreceptor is present in high densities in the hippocampus, the olfactory system, the striatum, and the cerebral cortex. Each of these subtypes is a Gi/Go protein–coupled receptor, which, on a cellular level, when activated by endogenous agonists (epinephrine) or exogenous agonists (clonidine), suppresses adenylyl cyclase activity which then leads to decreased production of cAMP. This in turn decreases feedforward cAMP–calcium signaling and prevents the opening of potassium channels. Continued exposure to α_2-agonists leads to downregulation of all three receptor subtypes.

Clonidine behaves as a partial agonist at all three α_2-receptor subtypes. Guanfacine is a selective α_2A-adrenoreceptor agonist that has a 15- to 20-fold increased selectivity for the α_2A-receptor over α_2B- and α_2C-receptor subtypes. The majority of the pharmacologic and physiologic functions are attributed to the α_2A-adrenoreceptor subtype.

Presynaptic α_2A-adrenoreceptor activation leads to an autoinhibition of norepinephrine release with activation of presynaptic α_2-receptors in the LC and cortex, and of α_2B-receptors in the thalamus likely accounting for the sedative effects of the α_2-agonists. Guanfacine is less sedating than clonidine, which is likely due to a weaker action at α_2B- and α_2C-receptors. The majority of α_2A-receptors are postsynaptic, including receptors on the dendritic spines of pyramidal cells in the PFC. Activation of these receptors and the resulting decreases in cAMP enhances prefrontal neuronal firing and strengthens prefrontal network connectivity, allowing for greater control of behavior including cognitive enhancing effects. The activation of this postsynaptic receptor also has a stimulatory effect on growth hormone secretion. Less is known about the effects of agonists on the α_2B- and α_2C-adrenoreceptors.

Clonidine, an imidazoline derivative, also binds to imidazoline-binding sites (IBS) located throughout the body and in the CNS. There are three subtypes of IBS: I1, I2, and I3. Clonidine has a greater affinity for the I1 IBS, which is involved in the control of blood pressure, and there is evidence to support that the hypotensive effect of clonidine is mediated via the IBS I1. Hypotensive side effects are less common with guanfacine, likely due to its weaker affinity at the IBS I1.

Blood Concentrations and Relation to Action. Concentrations of both

clonidine and guanfacine vary between individuals based on many factors, most notably age and renal function. A positive correlation between plasma concentrations of clonidine and its pharmacological effects has been demonstrated. In clinical practice, however, plasma levels are rarely obtained.

DESIGN AND INTERPRETATION OF CLINICAL DRUG STUDIES

The α_2-adrenergic receptor agonists have been evaluated in several psychiatric disorders with proposed underlying dysregulation of the adrenergic system or sympathetic nervous system. More recently, animal and neuroimaging research has led to further understanding of the neurobiology of PFC and the impact of the PFC and its connections on attention, emotion, and behavior. This has provided therapeutic targets to treat psychiatric disorders with a proposed underlying dysregulation of PFC. The design of clinical drug studies has been largely open studies, retrospective reviews, case reports, and small controlled studies. These studies have shown the α_2-adrenergic receptor agonists to have varying degrees of efficacy in the treatment of a variety of psychiatric disorders, including tic disorders, ADHD, PTSD, and opiate, nicotine, and alcohol withdrawal. Over the past 5 years, with the development of the extended-release formulations of clonidine and guanfacine, there have been multiple larger randomized, placebo-controlled studies. These studies have shown the extended-release α_2-agonists to be safe and efficacious for use in children and adolescents for the treatment of ADHD. In addition, the possible neuroprotective effects, especially in times of uncontrollable stress, of the α_2-adrenergic receptor agonists continues to be evaluated in humans because evidence of neuroprotective effects in animals has continued to expand.

EFFECT ON SPECIFIC ORGANS AND SYSTEMS

The α_2-agonists are active antihypertensive medications that directly affect the cardiovascular system by stimulating central α-adrenoreceptors, leading to reduced sympathetic nerve impulses from the vasomotor center to the heart and blood vessels. This results in a decrease in peripheral vascular resistance, renal vascular resistance, and a reduction in heart rate and blood pressure. Given these direct effects, measurement of heart rate and blood pressure prior to initiation, following dose increases and periodically while on therapy is recommended. Renal blood flow and glomerular filtration rate remain essentially unchanged. Normal postural reflexes remain intact; therefore orthostatic symptoms are mild and infrequent. Normal hemodynamic responses to exercise are maintained. Both clonidine and guanfacine acutely stimulate growth hormone release in both children and adults, but do not produce a chronic elevation of growth

hormone with long-term use.

THERAPEUTIC INDICATIONS

Clonidine extended release and guanfacine extended release are FDA approved in the treatment of ADHD as monotherapy and as adjunctive therapy to stimulant medication in children and adolescents. Immediate-release clonidine and guanfacine are FDA approved only for the treatment of hypertension, as monotherapy, or in combination with other antihypertensive medications. These medications are also used in a number of off-label uses for the treatment of numerous psychiatric conditions.

Tourette Syndrome and Other Tic Disorders

A tic is a sudden, rapid, recurrent, nonrhythmic motor movement or vocalization. Tourette syndrome (TS) is a neuropsychiatric syndrome characterized by both multiple motor and one or more vocal tics that may wax and wane in frequency but have persisted for more than 1 year since first tic onset, with first tic onset before the age of 18 years.

Clonidine and guanfacine are considered by many to be a first-line treatment for tics and TS, with multiple positive double-blind randomized placebo-controlled studies in the literature. Although haloperidol (Haldol) and pimozide (Orap) have been shown to be more effective than the α_2-agonists, concerns about their side-effect profiles, specifically NMS, tardive dyskinesia, and QTc prolongation (especially pimozide), have limited their use. Among the newer antipsychotics, the use of risperidone (Risperdal) has the strongest evidence for efficacy, yet its use is limited due to issues of weight gain, sedation, and risks of extrapyramidal symptoms and hyperprolactinemia. Preliminary data have shown clonidine and risperidone (Risperdal) to be equally efficacious. Preliminary data have also shown the clonidine transdermal patch to be an effective option for the treatment of TS and other tic disorders. When using clonidine for tics and TS it is important to note that there can be a delayed response, sometimes taking 2 to 6 weeks to have clinical effect.

Greater than 50 percent of children with TS have a comorbid diagnosis of ADHD. In this population, clonidine is considered first-line treatment. Several studies in this population report a significant improvement in impulsive-hyperactive symptoms as well as tic symptoms with both clonidine and guanfacine. In addition, there is evidence that the combination of clonidine and methylphenidate (Ritalin) is particularly efficacious and may not increase the severity of the tic disorder.

A standard titration of clonidine for the treatment of tics is 0.05 mg before bed, increasing by 0.05 mg every 3 to 7 days until the dose is 0.05 mg in the morning and midday, and 0.1 mg at bedtime. Total daily dose of clonidine for the treatment of tics is 0.1 to 0.3 mg. A standard titration of

guanfacine for the treatment of tics is 0.5 mg before bed, increasing by 0.5 mg every 3 to 7 days until the dose is 1 mg twice per day. In studies of tics or TS and comorbid ADHD, the dose range was 3 to 5 μg/kg per day. Measurement of heart rate and blood pressure prior to initiation, following dose increases and periodically while on therapy is recommended.

There are no published studies on the use of the extended-release formulations of clonidine and guanfacine in TS or tic disorders.

Attention-Deficit/Hyperactivity Disorder

Dysregulation of the central noradrenergic network has long been hypothesized to be an important aspect of the pathophysiology of ADHD. Ongoing structural and functional imaging studies along with neuropsychological studies have demonstrated delayed or inadequate maturation of the PFC, particularly in the right inferior regions needed for inhibitory control and executive cognitive functioning. Moderate levels of norepinephrine are critical for proper PFC function, while high and low levels may impair function. Action at presynaptic α_2-receptors in the LC leads to decreased cell firing and release of norepinephrine modulating and optimizing catecholamine influences to remediate symptoms of ADHD. Ascending norepinephrine afferent projections from the LC to the PFC enhance functional connectivity between the PFC and other brain regions. Guanfacine mimics norepinephrine at the postsynaptic α_2A-receptors.

The α_2-agonist clonidine has been used in the treatment of ADHD for more than 20 years. Yet, the effect size for stimulant medications in ADHD is larger than that for the α_2-agonists, and clonidine and guanfacine are therefore considered second-line agents for the treatment of ADHD. The α_2-agonists are most often used for augmentation of stimulants or when stimulant medications have been ineffective or poorly tolerated. There is also a particular interest in the use of α_2-agonists as monotherapy or in combination with stimulant medications for patients with comorbid ADHD and tic disorders. In these patients, data have shown moderate efficacy for some symptoms of ADHD with clonidine and guanfacine monotherapy, and with clonidine in combination with methylphenidate. For children with ADHD and comorbid oppositional defiant disorder or conduct disorder, there is increasing evidence that the addition of clonidine to ongoing psychostimulant treatment reduces symptoms of oppositional defiant disorder and conduct disorder, particularly symptoms of aggression.

The recent development and approval of the extended-release formulations of clonidine (CLON-XR) in 2010 and guanfacine (GXR) in 2009 has led to multiple positive randomized placebo-controlled studies as both monotherapy and adjunctive therapy for children and adolescents with ADHD. These trials support the efficacy and safety of extended-release α_2-receptor agonists for the symptomatic treatment of ADHD both as

monotherapy and adjunctive therapy with stimulant medication.

Highly comorbid with ADHD are oppositional symptoms such as anger, hospitality, irritability, arguing, and refusing to comply with adult directives. In studies of GXR, improvements in the oppositional subscales showed correlation with improvements in ADHD symptoms. Preliminary data of GXR plus a stimulant have shown significant reduction in oppositional symptoms, compared with placebo plus a stimulant in children and adolescents with ADHD and a suboptimal response to stimulant alone.

A standard CLON-IR dose for the treatment of ADHD is approximately 3 to 5 μg/kg per day, with a daily dose range of 0.1 to 0.3 mg per day. It is recommended to start with low doses and increase by 0.05 mg every 3 days. A standard dose of GUAN-IR for the treatment of ADHD ranges from 0.5 to 4 mg per day. The dose can be increased by 0.5 mg every 3 to 4 days. CLON-IR is administered on a schedule of three to four times a day, while GUAN-IR with its longer duration of action is generally administered on a twice a day schedule.

For CLON-XR as monotherapy for ADHD or as adjunctive therapy to a psychostimulant, dosing should be initiated with 0.1 mg at bedtime and adjusted in increments of 0.1 mg per day at weekly intervals until desired dose response is achieved. Dosing should consist of twice a day with either an equal split or the higher dosage at bedtime to a maximal total dose of 0.4 mg per day. For GXR as monotherapy for ADHD or as adjunctive therapy to a psychostimulant, dosing should be initiated once daily, either morning or evening with a dose of 1 mg per day and adjusted in increments of 1 mg per day at weekly intervals until desired dose response is achieved. Doses above 4 mg per day have not been systematically studied in controlled studies.

Concerns regarding the safety of the combination of stimulants and α_2-agonists remain after the sudden death of four children taking methylphenidate and clonidine was reported in 1995, although the specific contribution of clonidine in these deaths remains controversial. There have been no further reports of sudden death, and studies with children on this combination have been conducted with baseline, interval, and final visit ECGs and pulse and blood pressure monitoring. Before the initiation of an α_2-agonist, a cardiologist should evaluate patients with known cardiac problems or abnormal ECG. Measurement of heart rate and blood pressure prior to initiation, following dose increases, and periodically while on therapy is also recommended.

Opiate Withdrawal

Clonidine has been used for the treatment of opiate withdrawal since the 1980s and continues to be used in most detoxification protocols. Opiate

withdrawal is manifested by symptoms of irritability, anxiety, sweating, sniffling, rhinorrhea, nausea, diarrhea, lacrimation, abdominal pain, headache, muscular pain, and insomnia. During withdrawal, neural activity in the LC is greatly increased, and this surge is thought to be responsible for many of the symptoms seen during withdrawal. Clonidine's ability, through its preferential action on presynaptic α_2-neurons, to inhibit noradrenergic activity has been consistently shown to be effective in decreasing these symptoms by counteracting catecholamine release and decreasing the sympathetic hyperactivity in opiate withdrawal.

Clonidine has been used as a single agent in the treatment of opiate withdrawal and in combination with methadone (Dolophine), or buprenorphine (Buprenex), a partial opiate agonist. Clonidine is also used in rapid opiate detoxification under anesthesia (RODA), which is a procedure using opiate antagonists, such as naloxone/naltrexone (Narcan/Revia), under general anesthesia to rapidly reverse opioid dependence. Preliminary evidence finds that clonidine may be particularly effective in managing the gastrointestinal (GI) side effects associated with RODA. Some authors have suggested that withdrawal-induced diarrhea may be mediated through α_2-adrenoreceptors located on intestinal smooth muscles.

Although studies have shown clonidine to be effective in the management of opiate withdrawal, in comparison to methadone and buprenorphine, the treatment retention rate of patients treated with clonidine only is markedly lower. This is believed to be because patients treated with clonidine, as compared to methadone or buprenorphine, experience the most intense symptoms of opiate withdrawal earlier in the detoxification process. In addition, these patients experience and report greater medication side effects. These side effects are sedation, lethargy, fatigue, dry mouth, and hypotension. Withholding doses or decreasing the clonidine dose according to blood pressure changes usually manages the hypotensive side effects. Although the treatment retention rates and the percentage of patients who complete detoxification is higher when methadone, buprenorphine, or RODA is used, some recent work has suggested that at 6- and 12-month follow-up, the relapse rates with these other medications are equal to, or higher than, the clonidine only group. In adolescents, one study found buprenorphine with behavioral intervention to be significantly more efficacious than clonidine and behavioral treatment in the treatment of opiate withdrawal. In particular populations, such as prison inmates, where policies restrict opiate-assisted detoxification, surveys indicate clonidine is utilized in approximately 50 percent of jails. There is preliminary evidence that the use of standard clonidine protocols reduces inmate subjective symptoms of opiate withdrawal.

Several standardized protocols for the use of clonidine in opiate

withdrawal are commonly used. Generally, clonidine therapy is initiated at a dose of 0.1 to 0.2 mg orally up to six times per day and is titrated based on severity of symptoms and side effects. Maximum doses of clonidine are usually approximately 1 mg per day. The maximal doses are used for 2 to 3 days during the most acute withdrawal period. When the patient is stabilized, the clonidine is tapered over a 7- to 10-day period. Some protocols involve a slower taper or transition to a transdermal patch in order to address chronic symptoms of withdrawal such as sleep disturbance. Clonidine is usually withheld if the systolic blood pressure is less than 90 or if the patient shows marked bradycardia. Other regimens include basing the daily clonidine dose schedule on an as-needed basis and then decreasing the amount given over the next week.

Different formulations of clonidine are used in different protocols; RODA protocols use intravenous clonidine, although most protocols use tablets or a combination of tablet with a transition to the transdermal patch. Protocols using a combination of an opiate agonist, partial opiate agonist, or opiate antagonist along with clonidine have higher rates of completion of withdrawal.

At this time there are insufficient data available to make a conclusion regarding the efficacy of guanfacine in opiate withdrawal, but preliminary findings suggest that guanfacine may be as effective and associated with fewer side effects when compared with clonidine. In the comparison of clonidine and lofexidine (BritLofex), another α_2-receptor agonist, these drugs appear to have similar effectiveness in terms of reducing overall intensity of withdrawal, with those treated with lofexidine experiencing fewer side effects.

Nicotine Withdrawal

Clonidine is considered a second-line treatment for nicotine withdrawal. Although it has been shown to be effective in promoting smoking cessation, there is a high incidence of adverse effects, such as hypotension, dry mouth, and sedation. In light of these adverse effects, clonidine is typically reserved for smokers in whom sedation would be a desirable effect, such as in a smoker who experiences extreme agitation or anxiety during nicotine withdrawal. Clonidine also has a role in treating patients withdrawing from multiple drugs of abuse.

The recommended dose of clonidine for nicotine withdrawal from available studies is 0.1 mg twice a day with a titration to 0.4 mg daily as tolerated. Clonidine should be started 48 to 72 hours prior to cessation of nicotine use to allow time to reach steady state before the onset of nicotine withdrawal. Nicotine withdrawal symptoms may last 3 to 4 weeks after cessation of smoking, and then the clonidine dose is tapered gradually over several days in order to avoid clonidine withdrawal symptoms, such as

rebound hypertension.

Autism Spectrum Disorders

Results of surveys of medication prescribing patterns in children with autism spectrum disorders reveal that immediate release clonidine and guanfacine are being prescribed for between 10 and 15 percent of children with pervasive developmental disorders, and these rates have been increasing over time. In clinical practice, α_2-agonists are considered a second-line treatment for managing symptoms of hyperactivity, inattention, impulsivity, aggression, explosive outbursts, self-injury, and sleep dysfunction in autism spectrum disorders.

Overall empirical support for the use of the α-agonists in autism spectrum disorders is limited. There are three published trials of guanfacine in children with autism spectrum disorders. Preliminary evidence supports the efficacy and safety of guanfacine in the treatment of hyperactivity, inattention, and aggression in particular subsets of the autism spectrum disorders population. Clonidine has been studied as a sedative agent for autism spectrum disorder patients undergoing electroencephalogram (EEG), as a treatment for hyperactivity and inattention, and for sleep disturbance in these patients. Transdermal clonidine has also been investigated in small studies for the treatment of impulsivity and hyperactivity in children with autism spectrum disorders. At this time larger, controlled studies in this population are required. Of note, the individuals with autism spectrum disorders are generally more sensitive to the side effects of irritability and sedation when treated with α_2-agonists. There are case reports of use of extended-release guanfacine in children with autism spectrum disorders for symptoms of inattention and hyperactivity.

Posttraumatic Stress Disorder and Other Anxiety Disorders

Increased and altered noradrenergic neurotransmission and reactivity in the CNS has been associated with a variety of hyperarousal symptoms and sleep disturbances in PTSD. Clonidine's action at the postsynaptic α_2-adrenergic receptors in the CNS reduces sympathetic activity, possibly reducing the arousal symptoms of PTSD. Guanfacine, with the same mechanism as clonidine, has also been used for treatment of the symptoms of PTSD, including nightmares. Clinically there are reports of effectiveness in child refugees and veterans, although there are no randomized, placebo-controlled trials of clonidine in PTSD, and there are reports that it loses its effectiveness over time. In an 8-week, double-blind, randomized-controlled trial, immediate-release guanfacine was found to be ineffective in treating PTSD in veterans. In a small 8-week open label study of extended-release guanfacine for traumatic stress, generalized anxiety, and functional

impairment in traumatized children and adolescents, significant differences compared to placebo were seen on all symptom severity measures suggesting need for further investigation of GXR for traumatic symptomatology in children and adolescents.

Clonidine has also been studied in other anxiety disorders. In patients with panic disorder, clonidine appears to have an acute anxiolytic effect that is transient. Some debate exists as to whether clonidine has anxiolytic effects by proxy of its sedative side effects. The α_2-agonists do not appear to be effective in OCD.

Cognition and Neuroprotective Effects

The relationship between the noradrenergic system and executive and memory functions has been extensively investigated in animal models with positive results. Investigations of the cognitive enhancing potential of α_2-agonists via its effects on the PFC in humans have been inconclusive, and more recent studies found no significant clinical effects on adult healthy controls. There is evidence for the use of clonidine and guanfacine to enhance cognition in patients with schizophrenia and Korsakoff syndrome; however, its use remains limited by issues of tolerability.

Research on the neuroprotective effects of α_2-adrenergic receptor agonists began approximately 20 years ago after studies in rats indicated a protective effect of clonidine after cerebral ischemia. These neuroprotective effects have also been seen with the α_2-agonist brimonidine (Alphagan P) in glaucoma. At this time, despite extensive animal research, minimal investigation has been done to evaluate the neuroprotective effects of these drugs in humans.

Other Disorders

Clonidine has been used in the treatment of alcohol withdrawal to target the symptoms thought to be secondary to adrenergic hyperactivity and CNS excitation. There is no evidence suggesting that these medications are effective in reducing the incidence of seizures or delirium associated with alcohol withdrawal, thus reducing their use to an adjunctive treatment to benzodiazepines and typical antipsychotics. Animal studies have also looked at clonidine in decreasing voluntary alcohol drinking. There are also reports of using low-dose clonidine (0.025 mg three times per day) in hallucinogen persisting perception disorder, where symptoms may resolve within 2 months of initiating treatment.

There has also been a positive small double-blind randomized placebo-controlled trial investigating the use of clonidine for hyperarousal symptoms in patients with borderline personality disorder with and without PTSD.

PRECAUTIONS AND ADVERSE REACTIONS

Adverse effects of clonidine include dry mouth (40 percent), drowsiness (33 percent), dizziness (16 percent), constipation (10 percent), and sedation (10 percent). There are also numerous, albeit less frequent, side effects that have been reported in which a causal relationship has not been established, including nervousness, agitation, sleep disturbance, impotence, loss of libido, weight gain, headache, nausea, and vomiting. There are isolated reports of transient abnormalities in liver function tests, urinary dysfunction, thrombocytopenia, transient elevation of blood glucose, muscle pain, and leg cramps. Abnormalities on ECG such as sinus node arrest, junctional bradycardia, high-degree AV block, and arrhythmias have also been reported. Marked bradycardia is observed in some patients. Localized contact dermatitis is an adverse reaction to the clonidine transdermal therapeutic system (TTS) for approximately 15 to 20 percent of patients, particularly when using the optional adhesive cover. This can be addressed by alternating the site of the patch. The clonidine TTS contains aluminum and therefore should be removed before undergoing magnetic resonance imaging (MRI). Adverse reactions with use of extended-release clonidine include somnolence (30 percent), fatigue (15 percent), constipation (5 percent), and nightmares (3 to 9 percent), although there is some variability depending on dose. For guanfacine, the adverse effects are similar to those for clonidine, although hypotension, sedation, and dry mouth are somewhat less frequent. Adverse reactions with the use of extended-release guanfacine include somnolence (35 percent), fatigue (14 percent), and hypotension (7 percent). For these medications, there is evidence that the severity of dose-dependent adverse effects lessen over time.

Clonidine is considered a class C drug in pregnancy, which means that the safety is uncertain and that well-controlled studies in humans do not exist. Clonidine is excreted in breast milk, and caution should be exercised when administering clonidine in pregnancy and during lactation. Studies have not found evidence to suggest carcinogenesis and mutagenesis as a result of clonidine. Guanfacine is considered a class B drug in pregnancy, meaning that studies in humans have shown no evidence of risk to the fetus, although there have been adverse findings in animal studies. The excretion of guanfacine in breast milk is unknown, but in light of the fact that most medications are eliminated in breast milk, caution is suggested.

Clonidine is known to have a withdrawal syndrome. With sudden cessation of clonidine there may be an associated increase in plasma and urinary catecholamines, and associated symptoms such as nervousness, restlessness, insomnia, headache, increased heart rate, sweating, increased salivation, nausea, and flushing. There can also be episodes of rebound hypertension upon cessation of the medication. This syndrome is more pronounced after abrupt cessation of long-term therapy, in patients on a

higher total dose, and in patients also receiving β-blockers. A taper of the clonidine is required when the duration of use has been greater than 1 week, with a reduction in dose of 0.05 mg every other day. When the duration of use has been longer than 4 weeks, the dose reduction should be every third day. If there are symptoms of withdrawal present, they can be alleviated with clonidine. There are fewer data regarding a withdrawal syndrome with guanfacine, and although it remains a possibility, its longer duration of action makes withdrawal less likely.

DRUG INTERACTIONS

Clonidine and guanfacine cause sedation, especially early in therapy, and when administered with other centrally active depressants, such as barbiturates, alcohol, and benzodiazepines, the potential for additive sedative effects should be considered and dose reduction may be required. Due to a potential for additive effects, caution is required in patients receiving agents known to affect AV node and sinus node conduction such as β-blockers, calcium channel blockers, and digitalis. This combination increases the risk of AV block and bradycardia. When utilized along with antihypertensive drugs, use caution as there is the potential for additive effects. Guanfacine is metabolized by CYP3A4 microsomal enzymes and plasma concentrations can be affected significantly by CYP3A4 inhibitors or inducers. Guanfacine does not significantly affect exposures of methylphenidate and lisdexamfetamine when coadministered.

LABORATORY INTERFERENCE

Neither clonidine nor guanfacine is known to interfere with any laboratory tests.

DOSAGE AND ADMINISTRATION

Clonidine is supplied in 0.1-, 0.2-, and 0.3-mg tablets. It is also available in TTS that delivers 0.1, 0.2 or 0.3 mg daily (TTS-1, TTS-2, TTS-3). The dosing for the TTS is equivalent to the total oral daily dose. Multiple patches can be used at one time in order to administer the required dose. When administered for the treatment of hypertension, clonidine is initiated at 0.1 mg twice daily with increases of 0.1 mg at weekly intervals until desired response is achieved. Therapeutic doses for the treatment of hypertension range from 0.2 to 0.6 mg per day in divided doses. The maximum effective daily dose is 2.4 mg. It is suggested to administer a larger portion of the oral daily dose at bedtime to minimize adverse effects of dry mouth and sedation. The clonidine TTS is applied once every 7 days to a clean, dry, hairless area of intact skin on the upper outer arm, chest, or back. In children, a new patch is applied every 5 days due to their more rapid hepatic metabolism. Each new application should be on a different

skin site from the prior patch. If the clonidine TTS becomes loose from the skin, the adhesive cover can be placed over the system to ensure adhesion.

Clonidine extended release is available in 0.1- and 0.2-mg tablets. The tablets must be swallowed whole and never cut, crushed, or chewed.

Guanfacine is supplied in 1- and 2-mg tablets. When administered for the treatment of hypertension, guanfacine is initiated at 1 mg at bedtime to reduce sedation. The daily dose may be raised to 2 to 3 mg after 3 to 4 weeks of therapy. For the treatment of hypertension, the most clinical effect is seen at the 1- to 3-mg dose range, although the maximum recommended dose is 6 mg daily.

Guanfacine extended release is available in 1-, 2-, 3-, and 4-mg tablets. The tablets must be swallowed whole and never cut, crushed, or chewed.

REFERENCES

Arnsten AF. The use of α_2A adrenergic agonists for the treatment of attention-deficit/hyperactivity disorder. *Expert Rev Neurother*. 2010;10:1595–1605.

Arnsten AF, Jin LE. Guanfacine for the treatment of cognitive disorders: a century of discoveries at Yale. *Yale J Biol Med*. 2012;85(1):45–58.

Blankenship K, Erickson CA, Stigler KA, Posey DJ, McDougle CJ. Guanfacine extended release in two patients with pervasive developmental disorders. *J Child Adolesc Psychopharmacol*. 2011;21(3):287–290.

Bukstein OG, Head J. Guanfacine ER for the treatment of adolescent attention-deficit/hyperactivity disorder. *Expert Opin Pharmacother*. 2012;13(15):2207–2213.

*Connor DF, Arnsten AF, Pearson GS, Greco GF. Guanfacine extended release for the treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Expert Opin Pharmacother*. 2014;15(11):1601–1610.

Connor DF, Findling RL, Kollins SH, et al. Effects of guanfacine extended release on oppositional symptoms in children aged 6–12 years with attention-deficit hyperactivity disorder and oppositional symptoms: a randomized, double-blind, placebo-controlled trial. *CNS Drugs*. 2010;24(9):755–768.

Connor DF, Grasso DJ, Slivinsky MD, Pearson GS, Banga A. An open-label study of guanfacine extended release for traumatic stress related symptoms in children and adolescents. *J Child Adolesc Psychopharmacol*. 2013;23(4):244–251.

Croxtall JD. Clonidine extended-release: in attention-deficit hyperactivity disorder. *Pediatr Drugs*. 2011;13(5):329–336.

Daviss WB, Patel NC, Robb AS, et al. Clonidine for attention-deficit/hyperactivity disorder: II. ECG changes and adverse events analysis. *J Am Acad Child Adolesc Psychiatry*. 2008;47(2):189–198.

Elbe D, Reddy D. Focus on guanfacine extended-release: a review of its use in child and adolescent psychiatry. *J Can Acad Child Adolesc Psychiatry*. 2014;23(1):48–60.

Findling RL, McBurnett K, White C, Youcha S. Guanfacine extended release adjunctive to a psychostimulant in the treatment of comorbid oppositional symptoms in children and adolescents with attention-deficit/hyperactivity disorder. *J Child Adolesc Psychopharmacol*.

2014;24(5):245–252.

Fiscella K, Moore A, Engerman J, Meldrum S. Management of opiate detoxification in jails. *J Addict Dis*. 2005;24(1):61–71.

Fresquez-Chavez KR, Fogger S. Reduction of opiate withdrawal symptoms with use of clonidine in a county jail. *J Correct Health Care*. 2015;21(1):27–34.

*Gamo NJ, Arnsten AF. Molecular modulation of prefrontal cortex: rational development of treatments for psychiatric disorders. *Behav Neurosci*. 2011;125(3):282–296.

Gourlay SG, Stead LF, Benowitz NL. Clonidine for smoking cessation. *Cochrane Database Syst Rev*. 2004;(3):CD000058.

Gowing L, Farrell M, Ali R, White JM. Alpha$_2$-adrenergic agonists for the management of opioid withdrawal. *Cochrane Database Syst Rev*. 2016;3:CD002024.

*Hirota T, Schwartz S, Correll CU. Alpha-2 agonists for attention-deficit/hyperactivity disorder in youth: a systematic review and meta-analysis of monotherapy and add-on trials to stimulant therapy. *J Am Acad Child Adolesc Psychiatry*. 2014;53(2):153–173.

Jain R, Segal S, Kollins SH, Khayrallah M. Clonidine extended-release tablets for pediatric patients with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2011;50(2):171–179.

Kollins SH, Jain R, Brams M, et al. Clonidine extended-release tablets as add-on therapy to psychostimulants in children and adolescents with ADHD. *Pediatrics*. 2011;127:e1406–e1413.

Müller U, Clark L, Lam ML, et al. Lack of effects of guanfacine on executive and memory functions in healthy male volunteers. *Psychopharmacology (Berl)*. 2005;182:205–213.

Newcorn JH, Stein MA, Childress AC, et al. Randomized, double-blind trial of guanfacine extended release in children with attention-deficit/hyperactivity disorder: morning or evening administration. *J Am Acad Child Adolesc Psychiatry*. 2013;52(9):921–930.

Palumbo DR, Sallee FR, Pelham WE Jr, Bukstein OG, Daviss WB, McDermott MP. Clonidine for attention-deficit/hyperactivity disorder: I. Efficacy and tolerability outcomes. *J Am Acad Child Adolesc Psychiatry*. 2008;47(2):180–188.

*Rizzo R, Gulisano M, Calì PV, Curatolo P. Tourette Syndrome and comorbid ADHD: current pharmacological treatment options. *Eur J Paediatr Neurol*. 2013;17:421–428.

*Sallee F, Connor DF, Newcorn JH. A review of the rationale and clinical utilization of α2-adrenoceptor agonists for the treatment of attention-deficit/hyperactivity and related disorders. *J Child Adolesc Psychopharmacol*. 2013;23(5):308–319.

Sallee FR, McGough J, Wigal T, et al. Guanfacine extended release in children and adolescents with attention-deficit/hyperactivity disorder: a placebo-controlled trial. *J Am Acad Child Adolesc Psychiatry*. 2009;48(2):155–165.

*Stigler KA. Psychopharmacologic management of serious behavioral disturbance in ASD. *Child Adolesc Psychiatr Clin N Am*. 2014;23:73–82.

Strawn JR, Geracioti TD Jr. Noradrenergic dysfunction and the psychopharmacology of posttraumatic stress disorder. *Depress Anxiety*. 2008;25(3):260–271.

*Waldon K, Hill J, Termine C, Balottin U, Cavanna AE. Trials of pharmacological interventions for Tourette syndrome: a systematic review. *Behav Neurol*. 2013;26:265–273.

Weber B, Steinfath M, Scholz J, Bein B. Neuroprotective effects of alpha 2-adrenergic receptor agonists. *Drug News Perspect*. 2007;20(3):149–154.

Ziegenhorn AA, Roepke S, Schommer NC, Merkl A, Danker-Hopfe H. Clonidine improves hyperarousal in borderline personality disorder with or without comorbid posttraumatic stress disorder: a randomized, double-blind, placebo-controlled trial. *J Clin Psychopharmacol*. 2009;29:170–173.

▲ 34.5 β-Adrenergic Receptor Antagonists

ROGER S. MCINTYRE, M.D.

INTRODUCTION AND HISTORY

The first nontoxic β-adrenergic receptor antagonist, propranolol, was isolated by the Scottish pharmacologist James W. Black, earning him the Nobel Prize in Medicine in 1988. β-Adrenergic receptors are ubiquitously distributed throughout the body including the heart, lung, autonomic, and CNS. Although β-adrenergic receptor antagonists are not currently FDA approved for use in any psychiatric disorder, these agents have been used to treat disparate psychiatric conditions such as performance anxiety, PTSD panic attacks, major depressive disorder, aggressive and violent behavior, behavioral symptoms associated with dementia, alcohol and cocaine withdrawal, lithium-induced postural tremor, antipsychotic-induced akathisia, and migraine.

The five β-adrenergic receptor antagonists most frequently studied for psychiatric application are propranolol (Inderal), metoprolol (Lopressor, Toprol XL), atenolol (Tenormin), nadolol (Corgard), and pindolol (Visken).

CHEMISTRY

Propranolol is chemically described as 1-(isopropylamino)-3-(naphthalene-1-yloxy)propan-2-ol. Its molecular formula is $C_{16}H_{21}NO_2$. Propranolol is a stable, white, crystalline solid that is readily soluble in water and ethanol. Its molecular mass is 295.34. Its molecular structure is shown in Figure 34.5–1.

Metoprolol tartrate is chemically described as (1-[4-(2-methoxyethyl)phenoxy]-3-propan-2-ylamino-propan-2-ol). It is a white, practically odorless, crystalline powder with a molecular mass of [267.36] 684.82. Its molecular formula is $C_{15}H_{25}NO_3$. It is very soluble in water; freely soluble in methylene chloride, chloroform, and alcohol; slightly soluble in acetone; and insoluble in ether. The molecular structure is shown in Figure 34.5–1.

Atenolol may be chemically described as benzeneacetamide, 2-[4-[2-hydroxy-3-(1-methylethylamine)propoxy]phenyl]ethanamide. The molecular formula is $C_{14}H_{22}N_2O_3$. Atenolol has a molecular weight of 266.34. It is a relatively polar hydrophilic compound with a water solubility

of 26.5 mg/mL at 37°C and a log partition coefficient (octanol/water) of 0.23. It is freely soluble in 1N hydrochloric acid (300 mg/mL at 25°C) and less soluble in chloroform (3 mg/mL at 25°C). The molecular structure is given in Figure 34.5–1.

Nadolol is chemically described as 2R,3S 5-[2-hydroxy-3-3(tert-butylamino)-propoxy]tetralin-2,3-diol. Its molecular formula is $C_{17}H_{27}NO_4$. Nadolol has a molecular weight of 309.40. It is a white, crystalline powder that is freely soluble in alcohol, propylene glycol, and ethanol; slightly soluble in chloroform and water; very slightly soluble in sodium hydroxide; and insoluble in acetone, benzene, ether, and hexane. The molecular structure is shown in Figure 34.5–1.

Pindolol is chemically described as 1-(1H-indol-4-ylox)-3-(1--methylethylamino)propan-2-ol. Its molecular formula is $C_{14}H_{20}N_2O_2$. Pindolol has a molecular mass of 248.32. Pindolol is a white to off-white crystalline powder with a faint odor. It is practically insoluble in water, slightly soluble in methanol, and very slightly soluble in chloroform. The molecular structure is shown in Figure 34.5–1.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

The pharmacokinetics of atenolol, metoprolol, nadolol, propranolol, and pindolol are listed in Table 34.5–1.

Absorption. Absorption of propranolol (>95 percent) and pindolol (~95 percent) is extensive. Atenolol (~50 percent) and metoprolol (~50 percent) are intermediate, while absorption for nadolol is variable, averaging approximately 30 percent.

Distribution. β-Adrenergic receptor antagonists are categorized on the basis of their lipophilicity. The ability of β-adrenergic receptor antagonists to traverse the blood–brain barrier increases as a function of their lipophilicity. Propranolol and metoprolol are highly lipophilic and as a result, have both potent central and peripheral effects. Pindolol is of intermediate lipophilicity; nadolol and atenolol are the least lipophilic and have minimal central effects.

Propranolol is largely bound to circulating proteins (>90 percent). Propranolol exhibits extensive first-pass metabolism, with approximately 25 percent of an orally administered dose reaching the systemic circulation. Metoprolol is extensively metabolized by cytochrome P450 enzymes of the liver. With repeated oral administration, increases in a dose-dependent fashion. Metoprolol exhibits approximately 50 percent first-pass metabolism and relatively low protein binding (10 to 15 percent).

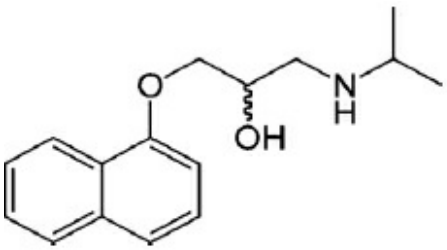

Propranolol

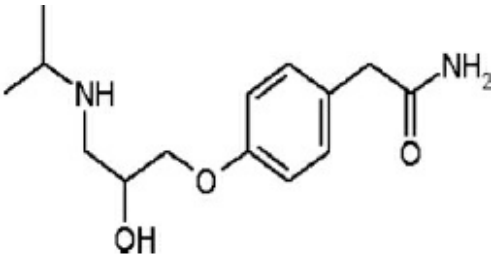

Atenolol

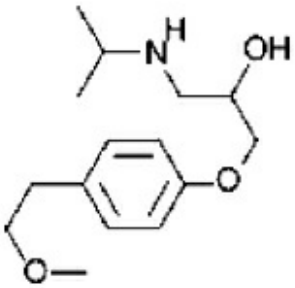

Metoprolol

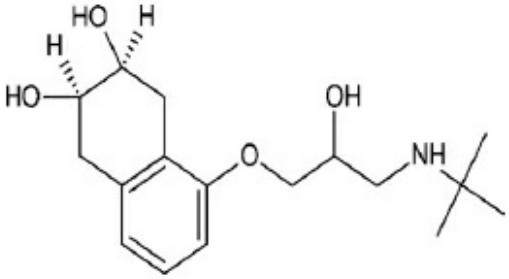

Nadolol

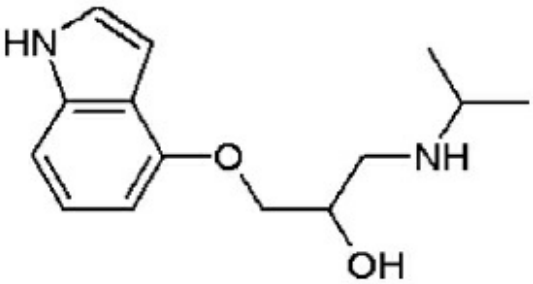

Pindolol

FIGURE 34.5–1. Molecular structures of β-adrenergic receptor antagonists.

Table 34.5–1. Pharmacokinetics

Drug	Pregnancy Category	Trade Name	Protein Binding (%)	Lipophilic	ISA	Metabolism	Receptor Selectivity	Half Life (hou
Atenolol (Tenormin)	D	Tenormin	6–16	No		Renal	$\beta_1 > \beta_2$	6–9
Metoprolol (Lopressor)	C	Lopressor	5–10	Yes		Hepatic	$\beta_1 > \beta_2$	3–4
Nadolol	C	Corgard	30	No		Renal	$\beta_1 = \beta_2$	14–

(Corgard)								
Propranolol (Inderal)	C	Inderal	>90	Yes		Hepatic	$\beta_1 = \beta_2$	3–6
Pindolol (Visken)	B	Visken	40	Yes	Minimal	Hepatic	$\beta_1 > \beta_2$	3–4

ISA, intrinsic sympathomimetic activity.

Atenolol plasma levels are proportional to dose over the range of 50 to 400 mg with relatively low protein binding (5 to 16 percent). Pindolol exhibits substantial interindividual variation in plasma concentrations. Pindolol has relatively low protein binding (~40 percent). Pindolol undergoes extensive metabolism in animals and man. In man, 35 to 40 percent is excreted unchanged in the urine and 60 to 65 percent is metabolized primarily to hydroxy-metabolites, which are excreted as glucuronides and ethereal sulfates. Approximately 30 percent of nadolol in serum is reversibly bound to plasma proteins. Nadolol is not metabolized by the liver and is excreted unchanged principally by the kidneys.

Elimination. The β-adrenergic receptor antagonists exhibit considerable variation in elimination half-life (see Table 34.5–1).

PHARMACODYNAMICS

Mechanism of Action

β-Adrenergic receptor antagonists block the action of endogenous catecholamines (primarily norepinephrine, as well as epinephrine) on β-adrenergic receptors. There are three types of β-adrenergic receptors, each of which exhibits tissue selectivity. β_1-Adrenergic receptors are localized in the myocardium and kidney; β_2-adrenergic receptors are located in the gastrointestinal tract, lung, liver, uterus, and vascular smooth muscles; and β_3-adrenergic receptors are located in fat cells. Norepinephrine is produced in the CNS by neurons in the locus ceruleus, a phylogenetically old monoamine system located in the dorsal pons with diffuse projections to subcortical and cortical regions.

There are also three types of α_1- and α_2-adrenoreceptors. Both α- and β-adrenergic receptors are G protein–coupled receptors. α_1-Adrenergic receptors are located in smooth muscle, notably in skin, where they mediate vasoconstriction. α_2-Adrenergic receptors are located pre- and postsynaptically in nerve terminals, where they mediate synaptic transmission. α_1-Adrenergic receptors are located in both the sympathetic and CNS. Clonidine is a selective agonist at the inhibitory α_2-adrenergic autoreceptor, while yohimbine is a selective antagonist. Clonidine, by decreasing central noradrenergic neurotransmission, exerts a centrally mediated hypotensive effect.

β-Adrenergic receptor antagonists are categorized as selective or nonselective on the basis of their affinity for β_1 and β_2 receptors, respectively. For example, propranolol and nadolol are nonselective, whereas atenolol, pindolol, and metoprolol exhibit differential affinity for the β_1-adrenergic receptor. Pindolol also exhibits intrinsic sympathomimetic activity, which implies modest agonist activity at the β-adrenergic receptor. None of the β-adrenergic receptor antagonists used in psychiatry is considered mixed antagonist, which implies antagonism of both β_1- and α_1-adrenergic receptors.

The β-adrenergic system participates in the formation and consolidation of emotional memories in humans. Pre- and postlearning stimulation of the β-adrenergic system enhances memory for the events that trigger stressful experiences, whereas β-adrenergic receptor antagonists administered prior to learning may inhibit declarative stressful memories. Propranolol affects context conditioning but has less impact on the acquisition or retention of cued fear conditioning.

Blood Levels and Relation to Action

Both propranolol and metoprolol exhibit variations in plasma concentrations due to their hepatic clearance. The bioavailability of propranolol increases with food ingestion and long-term administration. The lower lipid solubility of nadolol, pindolol, and atenolol reduces their interindividual variation.

DESIGN AND INTERPRETATION OF CLINICAL DRUG STUDIES

β-Adrenergic receptor antagonists are used in psychiatry to treat a variety of conditions. They are most frequently used as treatment for performance anxiety, lithium-induced postural tremor, and neuroleptic-induced akathisia. β-Adrenergic receptor antagonists are also used occasionally as adjunctive agents for the treatment of major depressive episodes, anxiety disorders, notably performance-based anxiety and panic attacks, aggression, behavioral disorders in the elderly, posttraumatic stress, and alcohol and cocaine withdrawal.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

The ubiquity of β-adrenergic receptors throughout the body implies a panoply of physiological effects on organs and systems (Tables 34.5–2 and 34.5–3). The systems most pertinent are cardiovascular, respiratory, and metabolic. Psychiatric populations commonly treated with β-adrenergic receptor antagonists (e.g., individuals with mood, anxiety, and psychotic disorders) are differentially affected by disorders in each of these systems. For individuals without pre-existing cardiovascular disease (e.g., hypertension), β-adrenergic receptor antagonists rarely cause hypotension

at doses prescribed. Individuals with pre-existing hypertension, however, exhibit an increased propensity to the hypotensive effects of β-adrenergic receptor antagonists.

Table 34.5–2.
Side Effects and Toxicity of β-Adrenergic Receptor Antagonists

Cardiovascular
Hypotension
Bradycardia
Congestive heart failure (in patients with compromised myocardial function)
Gastrointestinal
Anorexia
Nausea
Vomiting
Diarrhea
Abdominal pain
Metabolic
Weight gain
Worsened hypoglycemia in diabetic patients on insulin or oral agents
Respiratory
Asthma (less risk with β_1-selective drugs)
Shortness of breath
Wheezing
Sleep disturbances
Vivid dreams and nightmares with or without insomnia
Central nervous system
Dizziness
Drowsiness
Lassitude
Fatigue
Dysphoria
Bronchospasm
Lethargy
Depression
Psychosis (rare)
Dermatological
Rashes
Urticaria
Sweating
Sexual function
Impotence
Other (rare)
Raynaud phenomenon
Peyronie disease
Withdrawal syndrome
Rebound worsening of pre-existing angina pectoris when β-adrenergic receptor antagonists are discontinued

β-Adrenergic receptor antagonists minimally affect pulmonary function in healthy persons, but they may result in severe life-threatening bronchoconstriction in patients with chronic obstructive pulmonary

disease and asthma. As β-adrenergic receptor antagonists may alter the metabolism of carbohydrates and lipids and mask the premonitory signs of hypoglycemia, they should be used with caution in individuals with diabetes mellitus.

Table 34.5–3.
a-Adrenergic Receptor Antagonist Preparations

	Tablets (mg)	Extended Release (mg)	Solution	Parenteral
Atenolol	25, 50, 100			
Metoprolol	50, 100	200		1 mg/mL
Nadolol	20, 40, 80, 120, 160			
Propranolol	10, 20, 40, 80, 120	60, 80, 120, 160		1 mg/mL
Pindolol	5, 10, 15			

From Arana GW, Santos AB: β-adrenergic receptor antagonists. In: Kaplan HI, Sadock BJ, eds. *Comprehensive Textbook of Psychiatry*. 6th ed. Baltimore, MD: Williams & Wilkins; 1995

TERATOGENIC EFFECTS

None of the available β-adrenergic receptor antagonists has been adequately tested with controlled trials in pregnant women. Perinatal complications (e.g., placental insufficiency and intrauterine growth retardation) are reported with propranolol. Some infants born to mothers administered propranolol during pregnancy exhibited hypoglycemia and/or bradycardia.

Propranolol is the most commonly employed β-adrenergic receptor antagonist during pregnancy and is not known to exert a teratogenic effect in humans. β_1-Selective adrenergic antagonists such as atenolol or metoprolol may be preferred as they interfere less with β_2 receptor–mediated uterine relaxation. There are reports associating β-adrenergic receptor antagonists to fetal bradycardia, hypotonia, hypoglycemia, and intrauterine growth retardation.

Both metoprolol and pindolol cross the placental barrier; they are excreted in breast milk in small amounts. Transplacental passage of atenolol also occurs in pregnant women; fetal drug serum levels equal to the mother have been reported. Patients administered atenolol in the third trimester of pregnancy are at risk for low–birth-weight newborns and placental insufficiency. Nadolol crosses the placenta and is distributed into breast milk.

THERAPEUTIC INDICATIONS

Indications for β-adrenergic receptor antagonists are listed in Table 34.5–4.

Lithium Tremor

Lithium may result in an action tremor at therapeutic nontoxic levels. A stable lithium tremor in the absence of toxicity may, however, be distressing and impairing to the individual and require intervention. Therapeutic guiding principles are to ensure adequate blood levels necessary to achieve efficacy and to ensure that no other concomitant agents are contributory.

Consolidation of the dosing regimen to once a day at night ensures maximum lithium concentration during nonwaking hours (and may also reduce lithium-induced polyuria). If tremor persists, propranolol is most often used at a dose of 20 to 160 mg per day given in b.i.d. or t.i.d. dosing. β-Adrenergic receptor antagonists that bypass hepatic biotransformation (e.g., nadolol) may be preferred in individuals with compromised liver function. Less lipophilic β-adrenergic receptor antagonists may also be effective in lithium-induced tremor.

Table 34.5–4.
Indications for a-Adrenergic Receptor Antagonists in Psychiatric Practice

Effective
Social phobia (performance type)
Obsessive-compulsive disorder
Lithium-induced postural tremor
Neuroleptic-induced acute akathisia
Accelerate antidepressant action
Probably effective
Control of impulsive aggression or violence
Adjunct to benzodiazepines in treatment of alcohol withdrawal
Adjunct in treatment of cocaine and benzodiazepine withdrawal
Possibly effective
Generalized anxiety disorder
Schizophrenia
Treatment-resistant depression
Panic disorder
Posttraumatic stress disorder

Major Depressive Disorder

Pharmacological blockade of 5-HT_{1A} autoreceptors is thought to mimic the effect of autoreceptor desensitization and hasten the antidepressant effect of SSRIs. Results from open and randomized controlled clinical trials indicate that pindolol may accelerate the onset of antidepressant action in individuals with major depressive disorder. Extant studies suggest that pindolol's efficacy as an adjunct to SSRIs may be limited to the first 2 weeks of treatment. Pindolol's efficacy in treatment-resistant major depressive disorder, however, is less consistent. One possible explanation for inconsistent efficacy with pindolol may relate to drug–drug interactions as a consequence of a patient's pharmacokinetic profile (i.e., slow vs. fast

metabolizer).

The possibility that pindolol dosing in resistant populations was suboptimal (i.e., at 2.5 mg t.i.d.) has not been excluded. It is presumed that pindolol's mechanism of action relates to its affinity for somatic 5-HT_{1A} autoreceptors. When combined with SSRIs, pindolol also increases rapid eye movement (REM) suppression, notably an increase in REM onset latency and reduction in slow-wave sleep.

For patients receiving ECT, β-blockers are often an administered agent as part of an anesthesia regimen. It has been reported that the use of β-blocker therapy during ECT may be effective in reducing cardiovascular responses, but may also in some cases, may shorten seizure duration.

Substance Use Disorders

Available data suggest that blockade of β-adrenergic receptors potentiates cocaine-induced elevation of dopamine transmission in the nucleus accumbens, which is associated with an increase in cocaine-induced motor activity and a decrease in cocaine self-administration. Propranolol may attenuate the responsitivity of noradrenergic neurons, resulting in decreased anxiety from abstinence from cocaine, alcohol, and benzodiazepines. Propranolol may also selectively block memories associated with arousal in humans, which is salient to cue-triggered relapses in substance use disorders. Propranolol is superior to placebo in reducing cocaine withdrawal symptoms, but is not reliably efficacious in promoting abstinence from cocaine.

β-Adrenergic receptor antagonists have been used adjunctively for alcohol withdrawal and relapse prevention. Preclinical data indicate that the combination of propranolol with prazosin, an α_1-adrenergic receptor antagonist, significantly reduced alcohol drinking during both alcohol withdrawal and following imposed abstinence when compared to either drug alone. The foregoing observation suggests that contemporaneous blockade of central α_1- and β-adrenergic receptors may be a more robust strategy to reduce alcohol intake. β-Adrenergic receptor antagonists do not replace benzodiazepines for alcohol withdrawal. Available evidence indicates that atenolol 50 to 100 mg per day may be effective. β-Adrenergic receptor antagonists are not administered during alcohol withdrawal in individuals with a heart rate of less than 50 beats per minute.

Violence

Results from case reports and randomized, placebo-controlled trials indicate that β-adrenergic receptor antagonists are modestly effective and well tolerated for individuals with impulsive violence and disruptive behaviors associated with impulsive control disorders, brain injury, neurodegenerative disorder, and general medical conditions affecting the

CNS. The noradrenaline system has been implicated in the pathophysiology of behavioral disorders in Alzheimer disease. For example, individuals with Alzheimer disease evince an increased concentration of cerebrospinal fluid and norepinephrine and its metabolites.

Stimulation of CNS norepinephrine release by yohimbine, the α_2-adrenergic antagonist, engenders disruptive behaviors in Alzheimer disease. Propranolol 30 to 120 mg per day in divided doses is more effective than placebo in improving disruptive behaviors as measured by the Neuropsychiatric Inventory (NPI). In individuals with psychological impairment and severe aggression, propranolol 40 to 600 mg per day in divided doses has frequently been used as an adjuvant to antipsychotics, anticonvulsants, and lithium.

Performance Anxiety

The use of propranolol for performance anxiety as part of social phobia is the most familiar application of β-adrenergic receptor antagonists in anxiety disorders. Performance anxiety includes symptoms of dry mouth, coarse voice, increased heart rate, difficulty breathing, tremor, weakness, and dizziness. The somatic features often predispose and portend anticipatory anxiety for subsequent events. Performance anxiety can be severely impairing to affected individuals. When appropriately dosed, the β-adrenergic receptor antagonists have minimal central side effects and may facilitate performance. Benzodiazepines may result in disinhibition and impair cognitive performance.

Single dosing of propranolol 10 to 40 mg per day given approximately 20 to 30 minutes before performance is most frequently employed. It is generally recommended to consider a test dose prior to the performance event. The mechanism of action of β-adrenergic receptor antagonists relates to their effect on the peripheral sympathetic nervous system. In keeping with that view, less lipophilic agents may also be effective.

Anxiety

It has been postulated that disturbances in β-adrenergic receptor activity are associated with several anxiety disorders, notably panic disorder. Moreover, the anxiety behaviors are subserved by neural structures and circuits that involve the locus ceruleus. Human volunteer studies using the International Affective Pictures paradigm indicate that propranolol decreases basolateral amygdala responses to facial expressions independent of their emotional valence. Several randomized, controlled trials have established the efficacy of β-adrenergic receptor antagonists in the treatment of disparate anxiety disorders. Most studies have included individuals with generalized anxiety disorder, panic disorder, and OCD. It is generally understood that β-adrenergic receptor antagonists are

differentially efficacious for the somatic-autonomic symptoms (vs. psychological symptoms). Pindolol 2.5 mg per day t.i.d. has been used successfully as an adjunct to SSRIs in the treatment of OCD. For generalized anxiety disorder and panic attacks, propranolol doses of 40 to 320 mg per day have been employed.

Posttraumatic Stress Disorder

Disruptions in the catecholaminergic system in stress models are a highly replicated preclinical observation. Moreover, it has been reported in preclinical models that β-adrenergic receptor antagonism indirectly targets protein synthesis implicated in the reconsolidation of fear memories. It is established that disrupting reconsolidation targets original fears, and as a consequence, prevents the return of fear. Notwithstanding, a recent Cochrane collaboration review had concluded that β-adrenergic receptor antagonists have insufficient efficacy for the prevention of PTSD in individuals who are identified as at-risk.

Neuroleptic-Induced Akathisia

β-Adrenergic receptor antagonists may be more efficacious than anticholinergics and benzodiazepines for medication-induced movement disorders such as akathisia. Adequately powered comparator controlled studies, however, have not been completed. β-Adrenergic receptor antagonists are not considered efficacious for other neuroleptic-induced movement disorders such as acute dystonia and parkinsonism. In some resistant cases, β-adrenergic receptor antagonists may be combined with benzodiazepines and anticholinergics. Propranolol 10 to 30 mg per day has been the most frequently reported dosage; nadolol may also be employed.

PRECAUTIONS AND ADVERSE REACTIONS

Symptoms of propranolol overdose include bradycardia, cardiac failure, hypotension, and bronchospasm. The introduction of selective β-adrenergic receptor antagonists with less affinity at β_2 receptors has decreased, but not eliminated, the hazard of bronchospasm associated with nonselective agents. For patients with asthma or chronic obstructive pulmonary disease, selective β_1-adrenergic receptor antagonists should be used. Depression of the myocardium leading to cardiac failure has also been described. Abrupt discontinuation of propranolol therapy has been associated with severe exacerbation of angina and the occurrence of myocardial infarction. Skin rashes and conjunctival xerosis are rarely associated with β-adrenergic receptor antagonists. Aggravation of congestive heart failure has also been described. β-Adrenergic receptor antagonists may mask clinical signs of developing or continuing hyperthyroidism, as well as block premonitory signs and symptoms of

acute hypoglycemia.

Other adverse events include CNS depression, fatigue, nightmares, insomnia, vivid dreaming, and hallucinations. β-Adrenergic receptor antagonists do not adversely affect cognitive function and may enhance performance in measures of memory, learning, and psychomotor speed. The probability that a β-adrenergic receptor antagonist exerts central side effects is a function of its lipophilicity and ability to readily cross the blood–brain barrier. The amount of β-adrenergic receptor antagonist absorbed into brain tissue, rather than cerebrospinal fluid concentration, may be more pertinent to the liability of central side effects.

The incidents of CNS side effects is lowest with the most hydrophilic agents (e.g., atenolol) and highest with pindolol and propranolol, while metoprolol is intermediate. Other side effects attributable to β-adrenergic receptor antagonists are sexual dysfunction and gastrointestinal disturbances; hypersensitivity reactions are rare. Incident depressive symptoms are hypothesized to be due to the inhibitory actions of β-adrenergic receptor antagonists on β-adrenergic and serotonin receptors in the CNS. Studies evaluating this association have produced divergent results, however, which may be related to methodological factors such as small sample sizes, nonrandomized data, and lack of controlled groups.

β-Adrenergic receptors are located on osteoblasts; conflicting evidence suggests that β-adrenergic receptor antagonists may increase bone mineral density and reduce fracture risk. Individuals with sinus bradycardia, congestive heart failure, greater than first-degree atrioventricular block, severe chronic obstructive pulmonary disease, and bronchial asthma should not be prescribed β-adrenergic receptor antagonists.

The predominant symptoms reported following β-adrenergic receptor antagonist overdose are lethargy, respiratory depression, wheezing, sinus pause, and bradycardia. Additionally, a common effect associated with overdose of any β-adrenergic blocking agent is congestive heart failure. Treatment should be symptomatic and supportive and directed to the removal of any unabsorbed drug by induced emesis or administration of activated charcoal. Further consideration should be given to dehydration, electrolyte imbalance, and hypotension by established procedures.

DRUG INTERACTIONS

Drug interactions between β-adrenergic receptor antagonists and other agents may be either pharmacokinetic or pharmacodynamic. Lipid-soluble β-adrenergic receptor antagonists (e.g., propranolol, metoprolol) are extensively metabolized and have a high first-pass hepatic clearance. Conversely, atenolol, nadolol, and pindolol are eliminated unchanged by the kidney. Coadministered drugs that induce hepatic oxidative enzymes may increase the clearance of propranolol and metoprolol. Individuals coadministered metoprolol with a cytochrome P450 2D6-inhibiting SSRIs,

or those who are slow metabolizers of debrisoquine hydroxylase, may exhibit elevated plasma concentrations of metoprolol.

Coadministered agents that alter hepatic blood flow are also predicted to increase plasma concentrations of orally administered metoprolol and propranolol. Agents that inhibit hepatic enzymes and/or reduce hepatic blood flow will exert an opposite effect.

β-Adrenergic receptor antagonists may alter pharmacokinetics of other agents by reducing hepatic blood flow and inhibiting hepatic oxidative metabolism. Propranolol may increase the blood concentration of chlorpromazine by inhibiting its metabolism. Propranolol may also inhibit the metabolism of coumarin anticoagulants.

Pharmacodynamic interactions with β-adrenergic receptor antagonists include cardiac depression effects and potentiation of postural hypotension with coadministered antiarrhythmics and vasodilators.

LABORATORY INTERFERENCE

Pindolol may elevate liver transaminases, alkaline phosphatase, lactate dehydrogenase, and serum uric acid; pindolol rarely reduces bilirubin. Cases of thrombocytopenia and leucopenia have been described with β-adrenergic receptor antagonists.

DOSAGE AND ADMINISTRATION

β-Adrenergic receptor antagonists that act peripherally may be safer with regard to CNS adverse events when compared to agents with both peripheral and central action. For disorders in which a central effect is desired (e.g., violence), a more lipophilic agent (e.g., propranolol) should be considered. For peripheral side effects associated with psychotropic medication use, less lipophilic agents should be considered.

Propranolol is usually initiated orally at 10 mg t.i.d. or 20 mg b.i.d. increasing by approximately 20 to 30 mg per day until the desired effect is achieved. The patient's pulse and blood pressure readings should be taken regularly and the agent should be withheld if pulse is less than 50 beats per minute, the systolic blood pressure is less than 90 mm Hg, or the patient manifests CNS symptomatology. Propranolol should be tapered by approximately 40 to 60 mg daily until 60 mg per day is reached, after which time the drug should be tapered 10 to 20 mg per day every 3 to 4 days. β-Adrenergic receptor antagonists should not be discontinued abruptly as severe exacerbation of angina and the occurrence of myocardial infarction have been reported in patients with pre-existing cardiovascular disease following abrupt discontinuation of propranolol therapy.

Pindolol is initiated at 2.5 mg t.i.d. and increase is tolerated up to a maximum of 60 mg per day. Nadolol is initiated at 40 mg per day, increasing to 80 mg per day as tolerated. Atenolol is initiated at 50 mg per day and can be increased to 100 mg per day. Rarely, an atenolol dose of up

to 200 mg per day is used. The effective dosing range of metoprolol is 100 to 450 mg. The initiation dose is 100 mg; dosing beyond 450 mg has not been studied.

REFERENCES

Amos T, Stein DJ, Ipser JC. Pharmacological interventions for preventing post-traumatic stress disorder (PTSD). *Cochrane Database Syst Rev*. 2014;7:CD006239.

Antonelli-Incalzi R, Pedone C. Respiratory effects of beta-adrenergic receptor blockers. *Curr Med Chem*. 2007;14(10):1121–1128.

*Arana GW, Santos AB. B-Adrenergic receptor antagonists. In: Kaplan HI, Sadock BJ, eds. *Comprehensive Textbook of Psychiatry*, 6th ed. Baltimore, MD: Williams & Wilkins; 1995.

*Baker JG. The selectivity of beta-adrenoceptor antagonists at the human beta1, beta2 and beta3 adrenoceptors. *Br J Pharmacol*. 2005;144(3):317–322.

Ballesteros J, Callado LF. Effectiveness of pindolol plus serotonin uptake inhibitors in depression: A meta-analysis of early and late outcomes from randomised controlled trials. *J Affect Disord*. 2004;79(1–3):137–147.

Bell C, Wilson S, Rich A, Bailey J, Nutt D. Effects on sleep architecture of pindolol, paroxetine and their combination in healthy volunteers. *Psychopharmacology (Berl)*. 2003;166(2):102–110.

Boere E, Birkenhager TK, Groenland TH, van den Broek WW. Beta-blocking agents during electroconvulsive therapy: a review. *Br J Anaesth*. 2014;113(1):43–51.

**Compendium of Pharmaceuticals and Specialties*. Ottawa: Canadian Pharmacist Association; 2007.

de Quervain DJ, Aerni A, Roozendaal B. Preventive effect of beta-adrenoceptor blockade on glucocorticoid-induced memory retrieval deficits. *Am J Psychiatry*. 2007;164(6):967–969.

Hurlemann R, Walter H, Rehme AK, Kukolja J, Santoro SC, Schmidt C, et al. Human amygdala reactivity is diminished by the beta-noradrenergic antagonist propranolol. *Psychol Med*. 2010;40(11):1839–1848.

*Kaplan HI, Sadock BJ, Grebb JA. B-Adrenergic receptor antagonists. In: Kaplan HI, Sadock BJ, Grebb JA, eds. *Synopsis of Psychiatry*, 7th ed. Baltimore, MD: Williams & Wilkins; 1994:891.

Kindt M, Soeter M, Sevenster D. Disrupting reconsolidation of fear memory in humans by a noradrenergic beta-blocker. *J Vis Exp*. 2014;(94):doi: 10.3791/52151.

Martiny K, Lunde M, Bech P, Plenge P. A short-term double-blind randomized controlled pilot trial with active or placebo pindolol in patients treated with venlafaxine for major depression. *Nord J Psychiatry*. 2012;66(3):147–154.

Peskind ER, Tsuang DW, Bonner LT, et al. Propranolol for disruptive behaviors in nursing home residents with probable or possible Alzheimer disease: A placebo-controlled study. *Alzheimer Dis Assoc Disord*. 2005;19(1):23–28.

Rasmussen DD, Beckwith LE, Kincaid CL, Froehlich JC. Combining the alpha1-adrenergic receptor antagonist, prazosin, with the beta-adrenergic receptor antagonist, propranolol, reduces alcohol drinking more effectively than either drug alone. *Alcohol Clin Exp Res*. 2014;38(6):1532–1539.

Tsai SY, Lee HC, Chen CC. Hyperinsulinaemia associated with beta-adrenoceptor antagonist in

medicated bipolar patients during manic episode. *Prog Neuropsychopharmacol Biol Psychiatry*. 2007;31(5):1038–1043.

▲ 34.6 Anticholinergics and Amantadine

SAMOON AHMAD, M.D.

HISTORY

Anticholinergics block the binding of acetylcholine to its receptor present in nerve cells and are used to treat numerous diseases like asthma, motion sickness, symptoms of Parkinson disease and incontinence, and many others. In nature, anticholinergic substances exist as alkaloids and are most commonly found in perennial herbaceous plants such as *Atropa belladonna* (deadly nightshade), *Mandragora officinarum* (mandrake), *Hyoscyamus niger* (henbane), and Datura species. The active agents in belladonna with anticholinergic properties include atropine, hyoscine (scopolamine), and hyoscyamine. Atropine is the most common muscarinic antagonist and is used in cosmetics, for medicinal purposes, homeopathy, and as a recreational drug. In addition to being hallucinogenic, *Atropa belladonna* is one of the most toxic plants found in the Eastern hemisphere.

Anticholinergics fall into two categories: the antimuscarinic and antinicotinic. Antimuscarinic agents affect both the central and peripheral nervous system with varying selectivity among different agents. Antimuscarinic agents that block central cholinergic activity are used as antiparkinsonian drugs while those with nonselective antagonism have various clinical uses including bronchospasm, antispasmodic, urinary incontinence, and in anesthesia. The prototype of these agents is atropine.

ATROPINE

Atropine is a tropane alkaloid extracted from *Atropa belladonna* and other plants of the family *Solanaceae*. The solubility varies from being highly soluble in ethanol to being insoluble in ethyl ether. It is a secondary metabolite of these plants and serves as a drug with a wide variety of effects. Being potentially deadly, it derives its name from *Atropos*, one of the three fates who, according to Greek mythology, chose how a person was to die. Atropine is used in ophthalmic examination for pupil dilation and has been historically used by Italian women (*Bella Donna* is Italian for beautiful lady) to dilate their eyes for aesthetic reasons. Atropine is also used to treat poisoning by cholinesterase inhibitors, reverses the effects of nerve agents, and has cardiac and GI effects. Small doses can cause a marked increase in heart rate. It plays an important role in reversing reflex vagal cardiac slowing or asystole caused by various agents. Atropine is also administered in small doses before general anesthesia to lessen oral and air-passage secretions. Its ability to reduce these secretions is also utilized

in several preparations commonly used for symptomatic relief of colds. It also acts as a specific antidote to rapid type of mushroom poisoning. Blockade of muscarinic receptors in the CNS by atropine can lead to restlessness and excitement, and large doses can cause hallucinations. At the same time, these drugs are found to be therapeutically useful in Parkinson disease.

Atropine also serves as an antidote for poisoning by organophosphate insecticides and nerve gases. Troops who are likely to be attacked with chemical weapons often carry autoinjectors with atropine, which can be quickly injected into the thigh (Fig. 34.6–1).

Other commonly used anticholinergics include benztropine mesylate (Cogentin), diphenhydramine (Benadryl), trihexyphenidyl (Artane), and procyclidine (Kemadrin).

ANTICHOLINERGIC AGENTS

Chemistry

Artane (trihexyphenidyl hydrochloride) is the substituted piperidine salt, 3-(1-piperidyl)-1-phenyl-cyclohexyl-1-propanol hydrochloride, which exerts a direct inhibitory effect on the parasympathetic nervous system. It is used in the treatment of parkinsonism and extrapyramidal disorders. It also exhibits a direct spasmolytic action and weak mydriatic, antisialogogue and cardiovagal blocking effects (Fig. 34.6–2).

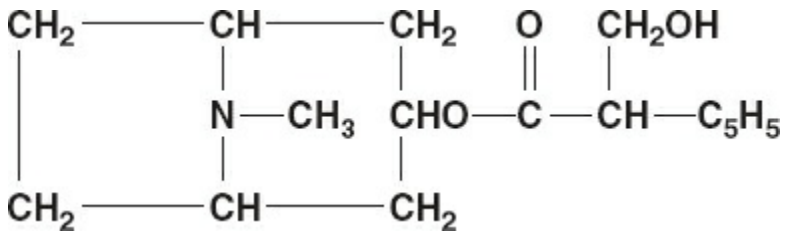

FIGURE 34.6–1. Structure of atropine.

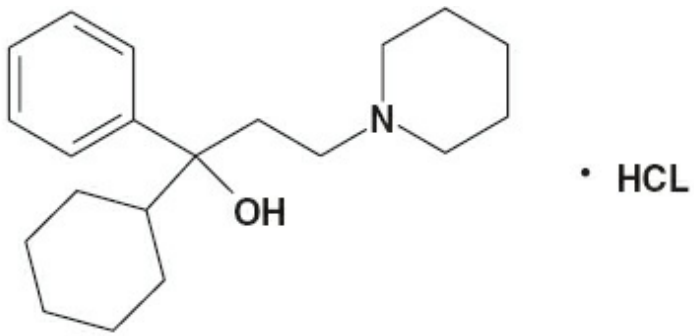

MOLECULAR STRUCTURE OF TRIHEXYPHEIDYL HCL

FIGURE 34.6–2. Molecular structure of trihexyphenidyl HCL.

Benztropine mesylate is designated chemically as 8-azabicyclo[3.2.1]octane, 3-(diphenylmethoxy)-*endo*, methanesulfonate. Its empirical formula is $C_{21}H_{25}NO \cdot CH_4O_3S$. It is a crystalline, white powder, very soluble in water, and has a molecular weight of 403.54. It possesses both anticholinergic and antihistaminic effects, although only the former has been established as therapeutically significant in the management of parkinsonism (Fig. 34.6–3).

Diphenhydramine hydrochloride is an antihistamine that occurs as a white, crystalline powder and is freely soluble in water and alcohol. It has a molecular weight of 291.82 (Fig. 34.6–4). It is used in the treatment of runny nose, sneezing, and itching of the nose or throat.

Procyclidine hydrochloride occurs as a white, tasteless, crystalline powder that is soluble in water. It is a synthetic antispasmodic compound useful for the symptomatic treatment of parkinsonism (paralysis agitans) and extrapyramidal dysfunction. Some find fewer side effects with this agent. The dose is begun at 2.5 mg per day and increased gradually to a maximum of approximately 30 mg per day. It is known chemically as α-cyclohexyl-α-phenyl-1-pyrrolidinepropanol hydrochloride. Its structural formula is shown in Figure 34.6–5.

Biperiden is an analog of trihexyphenidyl. It is approved for the treatment of parkinsonism and has greater pharmacological anticholinergic activity than trihexyphenidyl.

Pharmacological Actions

There is limited information available regarding pharmacokinetics of these drugs especially in the elderly and may limit the safe and effective use in this population. According to Goodman and Gilman's Pharmacological Basis of Therapeutics 8th edition, "There is surprisingly little information on the pharmacokinetic properties of these anticholinergic drugs." There are numerous factors contributing to this limited data including newer and more effective treatments, loss of patent for older drugs, and very low dosages used for treatment.

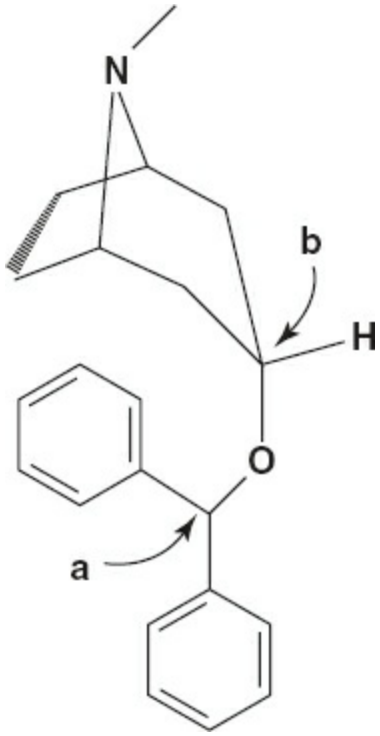

FIGURE 34.6–3. Molecular structure of benztropine mesylate.

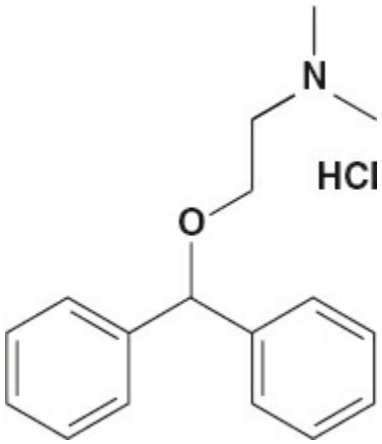

FIGURE 34.6–4. Molecular structure of diphenhydramine HCL.

Absorption and Distribution. Anticholinergic agents are widely distributed after absorption and achieve significant levels in the CNS within 30 minutes to 1 hour, and this high concentration in brain tissue may be related to their high lipophilicity. In humans, the bioavailability of these drugs is not reported and plasma protein binding is also poorly understood. All of these drugs are extensively metabolized, and all unchanged drugs as well as metabolites are excreted in the bile and urine.

Trihexyphenidyl is rapidly absorbed from the GI tract. After oral administration, the onset of action occurs within 1 hour, peak effects last 2 to 3 hours, and the duration of action is 6 to 12 hours. It is excreted in the urine.

Benztropine is well absorbed by the GI tract. However, it exhibits a

more rapid onset of action when administered either by the intravenous (IV) or intramuscular (IM) route, although there is no difference in onset of action between these two parenteral routes. All of these agents readily cross the blood–brain barrier (BBB).

Pharmacodynamics

Anticholinergic drugs act by blocking acetylcholine receptors. Acetylcholine acts in the nervous system at two major classes of cholinergic receptors: nicotinic and muscarinic. Distribution is quite different with the muscarinic receptors being located in the CNS, smooth muscle, cardiac muscle, and parasympathetically innervated glands, such as the salivary glands. Nicotinic receptors are located on striated muscle at the neuromuscular junction (motor end plate), the autonomic ganglia, and in the CNS. The muscarinic receptors are G-protein–coupled receptors. There are numerous muscarinic receptor subtypes (M1 to M5), and all are present in the brain. These receptor subtypes have selectivity for various brain regions with M1 receptor being the most abundant subtype in the CNS, and are predominantly located in the neocortex, neostriatum, and hippocampus. The M2 and M3 receptors are also widely distributed throughout the brain, but M3 receptor has a relatively lower density. The M4 receptors are mostly present in the neostriatum, while M5 receptors are localized in the hippocampus and projection neurons of substantia nigra, pars compacta, and ventral tegmental nuclei. Anticholinergic medications, such as benztropine, trihexyphenidyl, procyclidine, and biperiden, act as competitive antagonists at muscarinic receptors. The beneficial effect of agents in Parkinson disease is mediated by CNS muscarinic receptor blockade.

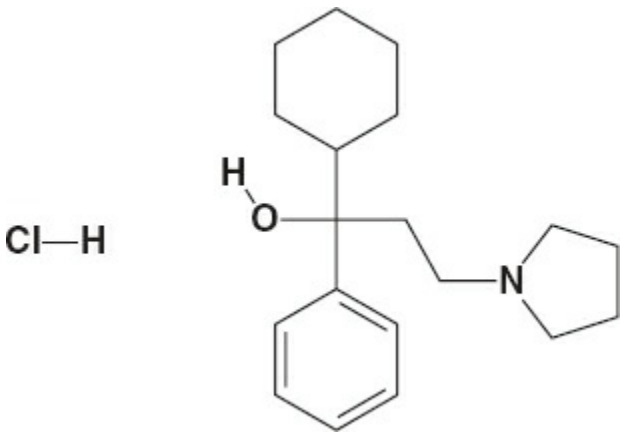

FIGURE 34.6–5. Molecular structure of procyclidine HCL.

The anticholinergic activity is independent of the serum drug levels of anticholinergic medications because drug levels vary up to 100-fold, and hence, are not routinely measured. There is evidence that there is a

threshold serum level for anticholinergics, and once this threshold is achieved, it is unlikely for extrapyramidal side effects to occur. This threshold level is a result of muscarinic receptor blockade. Consequently, the decrease in extrapyramidal side effects is not related to the dose of anticholinergic medication nor the dose or serum level of antipsychotic medication, but only to the serum level of anticholinergic activity.

Clinical Studies

The antimuscarinic drugs cause wide variety of CNS side effects, and this depends on their ability to penetrate the BBB as well as the activity on the receptors. Only small lipid-soluble molecules are able to pass from capillary blood to brain tissue. The cognitive impairment associated with these drugs in particular memory loss, may result from antagonism of M1 and, to some extent, M2 or M4 receptors in the CNS. Clinical studies suggest that older patients are more vulnerable to these cognitive side effects especially those with early stages of dementia, pre-existing cognitive impairment, and age associated memory decline.

Many recent studies have looked at the impact of these agents on cognitive functioning. A double-blind, crossover trial showed that agents that decrease cholinergic transmission impair storage of new information into long-term memory, but have little effect on retrieval from memory or on tasks involving only immediate memory. Another study evaluated the effects of discontinuation of anticholinergic medication on cognitive functioning in a group of elderly chronic schizophrenia patients. Significant improvement was demonstrated in the cognitive total score, as well as in the ideational praxia and orientation subscales.

These agents have also been implicated in worsening psychoses. However, many recent studies challenge this hypothesis; other recent studies show great variability in the patient population studied. A few recent small double-blind, placebo-controlled crossover studies have looked at anticholinergic medications and their effects on both positive and negative symptoms of schizophrenia. These agents appear to have no effect on the positive symptoms of schizophrenia in a stabilized patient, possibly because the presumed dopamine hyperactivity has been corrected. In schizophrenic patients who are acutely ill, studies consistently find that anticholinergic medications worsen positive symptoms.

Limited data exist regarding the utility of serum anticholinergic levels. One study compared verbal memory and reaction time of ten schizophrenic patients at two different serum anticholinergic levels. Verbal recall was worse at higher drug levels, while reaction time tended to be improved by anticholinergic treatment.

The effect of anticholinergics on tardive dyskinesia is also a controversial subject. Though the exact mechanism is unknown research studies have postulated that tardive dyskinesia may be related to an

abnormal influence of dopamine on striatal dopaminergic receptors. In one double-blind, placebo-controlled study, 10 chronic schizophrenic patients with pronounced symptoms of tardive dyskinesia were withdrawn from anticholinergic medication. The severity of tardive dyskinesia decreased significantly in nine patients within 2 weeks; this improvement, most pronounced in the oral region, persisted during a 6-week placebo period. Furthermore, acetylcholine and dopamine exert opposite effects on the striatum and a delicate balance is necessary to maintain normal functioning. The neuroleptic-induced dopaminergic blockage is considered to be a chemical denervation and analogous to pathological anatomical denervation of idiopathic Parkinson disease. Studies have shown the incidence of tardive dyskinesia to be higher in patients receiving both neuroleptic and anticholinergic medications. This hypothesis suggests that anticholinergic medications may predispose patients to the development of tardive dyskinesia. It is recommended that the use of anticholinergic medications should be kept to a minimum in patients receiving long-term neuroleptic therapy, and discontinuation of anticholinergic medication is a possible therapeutic approach in patients with tardive dyskinesia.

Effects on Specific Organs and Systems

Central Nervous System. In the usual recommended dosages, these drugs have a mild stimulant effect and a slower, longer-lasting sedative effect. The initial central vagal stimulant effect causes bradycardia, but as the antimuscarinic effects take place, it is supplanted by tachycardia. These effects are a result of the effect these agents have on the vagal nucleus in the medullary center and the sinoatrial node. Scopolamine is used for motion sickness and produces drowsiness and can cause amnesia in sensitive individuals. Toxic doses can cause agitation, hallucinations, excitement, and coma.

Parkinson disease is a result of excess of cholinergic activity and a deficiency of dopaminergic activity. The centrally acting antimuscarinic agents reduce the tremor, and when used in combination with a dopaminergic drug (levodopa [Dopar]) can have more effective response.

Eye. These agents cause cycloplegia (paralysis of the ciliary muscle), leading to loss of accommodation and the inability of the eye to focus for near vision. These medications also block the response of the sphincter muscle of the iris, which causes dilation of the pupil.

Anticholinergic medications have little effect on intraocular pressure, except in patients with narrow-angle glaucoma, in whom the pressure may rise to dangerous levels.

Cardiovascular System. In the heart, the atria are innervated by the parasympathetic (vagal) nerve fibers. The anticholinergic agents cause

tachycardia when moderate to high therapeutic doses are used. In toxic doses, these agents can cause intraventricular conduction block.

Respiratory System. The secretory glands and the smooth muscle of the airways are innervated by the muscarinic receptors. These agents inhibit secretions of the nose, mouth, pharynx and bronchi, usually clinically manifesting as dry mouth.

Gastrointestinal Tract. The GI tract is innervated by both cholinergic and noncholinergic fibers and is also modulated by local hormones. The net effect of these agents is decreased motility via inhibition of parasympathetic control, which may result in constipation or paralytic ileus. These agents also cause dry mouth and delay the gastric emptying time.

Genitourinary Tract. Anticholinergics cause relaxation of the smooth muscle in the ureter and the bladder wall and are used in the treatment of spasm caused by inflammatory conditions. At the same time, this can lead to urinary hesitancy, particularly in elderly men with prostatic hypertrophy.

Therapeutic Indications

Central Nervous System

EXTRAPYRAMIDAL SIDE EFFECTS. Extrapyramidal side effects (acute dystonias, parkinsonism, and akathisia) are caused by neuroleptic blockade of dopaminergic nigrostriatal tracts, leading to an increase in cholinergic activity. Other mechanisms involved include decreased serum calcium levels, which affect the cholinergic and dopaminergic systems. Acute dystonic reactions usually appear early and may occur within 1 to 4 days of neuroleptic therapy, while parkinsonism and akathisia usually manifest within the first 3 months of treatment. Anticholinergics reduce cholinergic activity and restore the cholinergic and dopaminergic equilibrium, leading to improvement in extrapyramidal side effects.

Acute Dystonia. Acute dystonia is characterized by a spastic contraction of muscle groups and often affects the neck, eyes, and trunk and develops early in the treatment or after rapid dose escalation of antipsychotic medication. The various forms of dystonia include torticollis, retrocollis, trismus or lockjaw, grimacing, dysarthria, blepharospasm, oculogyric crisis, and opisthotonos. These involuntary muscle contractions that occur suddenly are frightening and painful to the patient and may, in rare cases, be life-threatening (e.g., laryngospasm). In most cases, this side effect occurs within the first couple of days of treatment. It can be effectively and rapidly treated with medication. Treatment of a dystonic reaction is

typically with benztropine (1 to 2 mg) or diphenhydramine (25 to 50 mg), either IM or IV. Onset of action of these medications, IM or IV, is rapid, with resolution of the symptoms often within 10 to 15 minutes. As dystonic reactions may recur, prophylaxis with oral medications such as benztropine (1 to 2 mg twice daily) or trihexyphenidyl (2 mg three times a day) is recommended. Tardive dystonia may develop after months or years of treatment with antipsychotic medications and usually does not respond rapidly to anticholinergic medication.

Akathisia. Akathisia is characterized by an internal sense of motor restlessness often described as an inability to resist the urge to move. The most common form of akathisia involves pacing and an inability to sit still and most often occurs during the first few months of conventional antipsychotic use. At times, patients may have difficulty explaining the sensation of akathisia and can be often misdiagnosed as agitated, leading to an increase in antipsychotic medication and hence worsening the akathisia. Anticholinergics have limited utility in the treatment of akathisia; they should be considered second-line therapy and used when first-line agents, such as adrenergic receptor antagonists (propranolol), either fail or are contraindicated.

Parkinsonism. Parkinson disease and parkinsonism secondary to the use of antipsychotics can be indistinguishable. The symptoms include generalized slowing of movement (akinesia), rigidity, resting tremor, hypersalivation, cog-wheeling, and masked facies. These symptoms are caused by a blockade of dopamine receptors of nigrostriatal neurons, which terminate in the caudate nucleus. These symptoms can be mistaken for the negative symptoms of schizophrenia. Parkinsonism is more common with the high-potency antipsychotics and is seen much less frequently with atypical antipsychotics. These symptoms may manifest any time during antipsychotic treatment, but usually begin within the first 3 months of treatment and can be managed by lowering the antipsychotic dose to the lowest effective dose.

Anticholinergics are found to be useful in treating the symptoms of neuroleptic-induced parkinsonism. The additions of benztropine (0.5 to 6.0 mg per day), trihexyphenidyl (2 mg, two to four times daily), procyclidine (5 to 20 mg per day), and biperiden (2 to 24 mg per day) have been found to be effective therapies.

Long-Term Use of Anticholinergics

Patients with schizophrenia often require long-term treatment with conventional antipsychotics. As noted above, these medications have serious adverse effects such as extrapyramidal side effects, including acute dystonic reactions, akathisia, parkinsonism, and tardive dyskinesia. Much controversy exists in the prophylactic use of anticholinergic medication to

reduce or minimize these adverse events. As mentioned before, earlier reports stated that long-term use of anticholinergics might predispose to tardive dyskinesia. Anticholinergics also have serious side effects: long-term use may affect memory, and cognitive impairment can be seen as early as 60 days if these medicines are taken continuously; overdose of these agents creates a dangerous condition with tachycardia, hyperthermia, mydriasis, hallucinations, seizure, and coma; and anticholinergics may be abused. These findings led to recommendations by the World Health Organization, which has released a consensus statement that "the prophylactic use of anticholinergics in patients on neuroleptic treatment is not recommended, and may be justified only early in treatment (after which it should be discontinued and its need should be reevaluated)." Based on the above, since many patients will not develop parkinsonism, they do not need these medications.

Later studies suggest that there is evidence to support the use of initial prophylaxis, especially in young men receiving high-potency antipsychotics and who remain at high risk for developing acute dystonia. They also reviewed the literature for continuing prophylactic treatment and concluded that continuing prophylaxis is helpful for many patients because a large number of patients will develop parkinsonism upon withdrawal of anticholinergic medication. The prophylactic use of these agents is not routinely indicated and should be used on a case-to-case basis for patients with high risk of developing acute dystonic reactions.

Precautions and Adverse Reactions

Anticholinergic medications have CNS as well as peripheral adverse reactions due to their widespread antagonism of muscarinic receptors in the periphery. These agents produce clinically meaningful effects on the cardiovascular, GI, autonomic, and urogenital systems. The major adverse reactions of these medications are summarized in Table 34.6–1.

Anticholinergic agents block muscarinic acetylcholine receptors throughout the body and produce system-wide side effects. At low doses, it causes dry mouth, but at higher dosages it can cause constipation, which can progress to intestinal pseudo-obstruction and bowel dilation. Effects on the bladder can result in urinary retention, especially in male patients. Nausea, drowsiness, abdominal cramps, and mild tachycardia can also occur. Dry skin and impaired sweating may lead to heat stroke. Anticholinergics are contraindicated in patients with closed-angle glaucoma. These drugs can cause confusion and memory loss, and in the more advanced and demented patients, hallucinations and disorientation can occur. Lowering or discontinuing the drug can treat the side effects.

Table 34.6–1.

Adverse Effects of Anticholinergics
Cardiovascular system
► Tachycardia
► Digestive system
► Constipation
► Dry mouth
► Nausea
Nervous system
► Confusion
► Disorientation
► Memory impairment
Visual system
► Blurred vision
► Dilated pupils
► Worsening of narrow-angle glaucoma
Urogenital system
► Urinary retention

Toxic doses or an overdose of anticholinergic medication may lead to an atropine-like poisoning, which may manifest as confusion, nervousness, psychosis, ataxia, dry mouth, blurred vision, tachycardia, elevated blood pressure, GI upset, anhidrosis, hyperthermia, and respiratory collapse. In general, supportive measures are used after an overdose. However, 1 to 2 mg of physostigmine salicylate (Antilirium), an acetylcholinesterase inhibitor, may be used to reverse the anticholinergic symptoms of overdose. If required, a second dose of physostigmine may be administered 2 hours after the first dose.

Anticholinergic Abuse. These agents have an abuse potential secondary to their euphoriant and hallucinogenic effects, especially when combined with other street drugs. The prevalence of anticholinergic abuse varies from 0 to 34 percent. The most common reason for abuse is the drug-induced euphoria. Trihexyphenidyl is the most commonly abused anticholinergic for the reason that it tends to be the most stimulating of all anticholinergics.

Drug Interactions

Any medication with anticholinergic or antihistaminic properties may be pharmacodynamically additive to the effects of anticholinergic medications and, thus, would predispose to side effects and toxicity. Pharmacokinetic interactions with these agents occur as well. For example, the coprescription of benztropine and digoxin (Digitek) has been reported to increase digoxin levels (Table 34.6–2).

Table 34.6–2.

Anticholinergic Preparations and Dosages

Generic Name	Brand Name	Preparations	Daily Oral Dose
Benztropine	Cogentin	0.5, 1.0, 2.0 mg tablets 1 mg/mL injection	0.5–2.0 mg b.i.d. or t.i.d.
Trihexyphenidyl	Artane	2, 5 mg tablets 2 mg/5 mL elixir	2 mg b.i.d. to q.i.d.
Biperiden	Akineton	2 mg tablets 5 mg/mL injection	2 mg t.i.d. to q.i.d.

Any medication with anticholinergic properties may also potentiate the anticholinergic effects of amantadine. Due to the renal clearance of amantadine, medications that decrease renal clearance may result in higher plasma concentrations, and consequently, a greater risk for side effects. These include triamterene hydrochlorothiazide (Maxzide), trimethoprim sulfamethoxazole (Bactrim), and quinidine.

Laboratory Interference

The anticholinergic medications and amantadine are not known to alter the results of any standard laboratory tests and values.

Dosage and Administration

Anticholinergic doses can vary widely among individuals and must be empirically determined. In general, prophylactic doses are the same as the doses usually needed to treat parkinsonism. For acute dystonias, an IM or IV medication such as benztropine (1 to 2 mg IM) or the antihistamine diphenhydramine (25 to 50 mg IM or IV) is usually given. The starting dose for trihexyphenidyl is 0.5 to 1 mg per day and the dose should be increased no faster than every 7 to 14 days. Dose increase of 1 mg per day is reasonable and a dose of 2 mg three to four times per day is the usual goal.

The dose of benztropine is begun at 0.5 mg per day and can be raised to 4 to 8 mg per day and requires dosing two to three times per day.

The dose of procyclidine is begun at 2.5 mg per day and increased gradually to a maximum of approximately 30 mg per day.

The starting dose of diphenhydramine is 25 mg per day, and it can be titrated to 150 to 200 mg per day in divided doses.

SCOPOLAMINE

Introduction and Chemistry

Scopolamine is a tropane alkaloid drug and a competitive acetylcholine antagonist with significant affinity for the M3 receptor and has minimal effect at nicotinic receptors. It has been in use in its plant-based form since prehistoric times but has only been used in its purified forms (including

hydrobromide, hydrochloride, hydroiodide, and sulfate) since 1880. It is more potent then atropine for its antimuscarinic effect on the eye and salivary glands and less potent on the heart, respiratory and intestinal smooth muscle (Fig. 34.6–6).

Pharmacological Actions

Scopolamine is well absorbed from the GI tract primarily from the upper small intestine. Its onset of action varies depending on the formulation but ranges from 15 to 30 minutes for an injectable form to 4 hours for the transdermal form. The therapeutic effects last up to a maximum of 72 hours with the transdermal patch. Scopolamine crosses the BBB as well as the placenta and is secreted in to the milk and is classified as a pregnancy category C drug. It is reversibly bound to plasma proteins and principally eliminated in the urine. The transdermal formulation has fewer side effects and longer duration of action compared to the oral form. It is also used preoperatively to inhibit gastric and respiratory tract secretions.

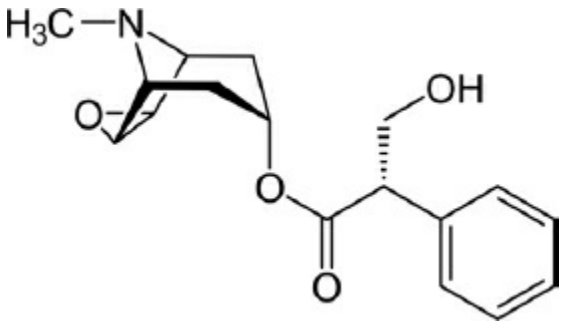

FIGURE 34.6–6. Molecular structure of scopalamine.

Dosage and Administration

Scopolamine can be administered orally, as a transdermal patch or injected IM, direct IV or sub-Q. The oral adult dosage ranges from 0.4 to 0.8 mg initially and up to three times daily as needed. The injectable dose ranges from 0.32 to 0.65 mg and can be repeated up to three times daily.

Therapeutic Indications

Scopolamine is mostly used for motion sickness and postoperative nausea and vomiting. Recent studies suggest its role in depressive disorders as well. Scopolamine is available in various forms including transdermal, injectable, and oral formulations.

Precautions and Adverse Reactions

CNS effects: Drowsiness, amnesia, disorientation, dizziness, and confusion.

Ocular effects: Possible increased intraocular pressure requiring monitoring and may require adjustment of therapy.

Idiosyncratic reactions: Acute toxic psychoses, somnolence, dilated pupils, delirium and dryness of the mouth, and increased pulse rate.

Other possible adverse effects: Decreased GI motility, possible urinary retention, and decreased bronchial secretions with formation of bronchial plugs.

Design and Interpretation of Clinical Drug Studies

During the last few years, the role of anticholinergic medications has expanded to include treatment of depressive disorders. Researchers initially proposed the idea of adrenergic cholinergic imbalance as early as 1972, and though initially ignored, most of the recent body of research is based on this hypothesis. It is well known that increasing cholinergic activity using physostigmine (an anticholinesterase inhibitor) may lead to exacerbation of acute depressive symptoms in patients with major depressive disorder (MDD). In addition, this cholinergic increase may also induce depressive symptoms in currently manic patients with bipolar disorder (BD). Numerous studies since then have shown depressive symptoms associated with increased cholinergic activity while the inverse is true for improvement in major depression. It is also well established that REM sleep changes are associated with MDD, and REM sleep deprivation as induced by anticholinergic medications leads to reduction in these symptoms.

Among the anticholinergic medications, scopolamine in particular has been shown to improve symptoms of moderate to severe depression. The first evidence of this finding was suggested in 1991 research conducted on a small group of patients with moderate depression using intramuscular scopolamine and showed a statistical difference. However, it was not until 2006 that investigators established the efficacy of scopolamine in depressive disorders while studying the cognitive effects of intravenous scopolamine and incidentally found antidepressant effects of this drug. To validate these findings the investigators designed two crossover randomized controlled trials, and showed a significant robust and rapid antidepressant response of intravenous scopolamine compared to placebo. A later study by the same researchers in 2010 validated these findings specifically in patients with unipolar depression.

Another study conducted at the National Institute of Mental Health enrolled patients with MDD who received multiple sessions of placebo or scopolamine hydrobromide (4 μg/kg) and were randomized to a placebo/scopolamine or scopolamine/placebo sequence of infusions 3 to 5 days apart. Patients were evaluated using the Montgomery-Asberg Depression Rating Scale and the Hamilton Anxiety Rating Scale to assess response to scopolamine. Patients showed reductions in depression and anxiety from baseline during scopolamine infusions compared to placebo infusions. The antidepressant and antianxiety effects occurred in less than

3 days compared to conventional antidepressants which may take up to 4 weeks. The group receiving scopolamine infusions first followed by placebo infusions continued to show response suggesting alternative mechanisms besides effect on neuroreceptors. This persistence of improvement in mood beyond the initial infusion period also negates the hypothesis of anticholinergic euphoria and suggests a sustained antidepressant response. This study showed a rapid antidepressant response to scopolamine in depressed patients that is well tolerated and safe relative to other rapid antidepressant treatments including ketamine and high-dose TCA therapy. Scopolamine has high affinity for M3 receptor and is 30 times more potent than amitriptyline.

Recognizing the fact that a significant number of patients with MDDs do not respond to antidepressant medications, the need for augmentation strategies from the beginning are being increasingly entertained. Considering the role of scopolamine as discussed above, it is no surprise that it may have a substantial role in such treatments. The major limitation of using scopolamine has been the intramuscular and intravenous route of administration, which has a limited utility in every day practice settings. To overcome this barrier and assess the tolerability and efficacy of oral scopolamine, investigators designed a randomized, placebo-controlled, double-blind, parallel-group study to evaluate the role of oral scopolamine as an adjuvant therapy to citalopram in moderate to severe MDD.

Forty patients were randomized into two groups, the active group receiving citalopram and scopolamine hydrobromide (containing 0.5-mg scopolamine) with the second group receiving citalopram and placebo. The total dose of scopolamine used was 1 mg/day (0.5 mg b.i.d.) while citalopram was adjusted from an initial dose of 20 mg in the first week to 40 mg/day for the next 5 weeks. The patients were assessed using 17-item Hamilton Depression Rating Scale (HDRS) for 6 weeks with assessments at day 4, 7, 14, 28, and 42. By week 6, all patients in the scopolamine group and most patients in the placebo group had achieved a 50 percent reduction in the HDRS while the remission rate by week 6 were significantly higher (65 percent) in the scopolamine group compared to the placebo group (20 percent). Another study at the National Institute of Mental Health showed that intravenous scopolamine had a significant effect in unipolar and bipolar depression.

The results of this study validated the role of scopolamine as an adjunct to antidepressant therapy when initiated from the beginning of the treatment with a high response and remission rate and few adverse effects. These studies suggest that scopolamine as monotherapy as well as when used as an adjunct is a safe and effective strategy in MDDs.

In recent years, the need for better treatment options has been directed to biomarkers to investigate rapid treatment response especially with antidepressants. Scopolamine has been shown to induce rapid

antidepressant response through its effect on the cholinergic system; a study in 2012 showed that neural responses in the visual cortex especially to emotional stimuli may be a potential biomarker for patients with favorable response to scopolamine. The study conducted at National Institute of Mental Health suggests differences in treatment response based on brain-based variables and requires a more personalized approach to treatment for depressive disorders.

Various theories abound regarding the antidepressant efficacy and precise mechanism of action of scopolamine in MDDs, and to date, none have been established. Some of these theories suggest alterations in type 2 and 3 muscarinic receptors as well as polymorphisms in type 2 muscarinic receptor gene, and determine that scopolamine has a high affinity for type 3 muscarinic receptor. Other possible mechanisms include a hypercholinergic system, REM sleep deprivation as well as hyperactivity of the glutamatergic system which has been implicated in the development of MDD. Scopolamine decreases the glutamatergic transmission, thereby alleviating depressive symptoms.

Scopolamine seems to be an appropriate treatment option for difficult to treat patients with major and bipolar depression. Its relatively simple dosing and minor side effect profile with no significant impact on cognition based on the low doses used in the studies suggest its place as an adjunct to antidepressant treatment. Moreover, it should be noted that combination therapy from the beginning of treatment results in greater improvement than monotherapy. Some early studies reported less robust antidepressant effects but those results may have been due to the uncontrolled nature of the studies as well as lower dose range utilized. Despite the positive implications of the studies, numerous limitations remain regarding its impact on cognition (confusion and delirium), dosing, and safety, and more research is needed to study its efficacy in the long term and rapid antidepressant response as an adjunct when added later in the treatment.

AMANTADINE

Amantadine (Symmetrel; 1-aminoadamantane hydrochloride) is an antiviral drug that was approved in 1976 for the prophylaxis and treatment of influenza A. It was by chance found to have antiparkinsonian properties and is now used to treat the symptoms of Parkinson disease, drug-induced extrapyramidal symptoms such as parkinsonism, akinesia, and the rabbit syndrome (focal perioral tremor of the choreoathetoid type).

Chemistry

Amantadine is a tricyclic symmetric amine and is a stable, white, or nearly white, crystalline powder that is freely soluble in water and soluble in alcohol and in chloroform. Its molecular weight is 187.71 with a molecular

formula $C_{10}H_{18}NCl$ (Fig. 34.6–7).

Pharmacological Actions

Absorption and Distribution. Amantadine is well absorbed orally. The peak plasma concentrations of 0.2 to 0.4 μg/mL are reached in 1 to 4 hours after an oral dose and are directly related to dose for doses up to 200 mg per day. It reaches steady state in 4 to 7 days. Doses above 200 mg per day may result in a greater than proportional increase in maximum plasma concentrations. It is approximately 67 percent bound to plasma protein.

Metabolism and Elimination. Amantadine is primarily excreted unchanged in the urine by glomerular filtration and tubular secretion. Increasing plasma concentrations can lead to toxicity, but no absolute values of amantadine concentrations associated with adverse effects have been fully established. Mean half-life ranges from 10 to 14 hours. One study involving 19 healthy volunteers showed the plasma half-life averaged 16 ± 6 hours (range 9 to 31 hours).

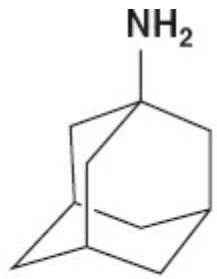

CHEMICAL FORMULA

$C_{10}H_{17}N$

FIGURE 34.6–7. Amantadine structure.

SPECIAL CONSIDERATION IN THE ELDERLY AND RENALLY IMPAIRED. In healthy elderly individuals age 60 and older, the oral plasma clearance of amantadine is reduced and the plasma half-life and plasma concentrations are increased. It is unclear whether these changes are due to decline in renal function or other age-related factors. In adult patients with renal insufficiency, the clearance of amantadine is significantly reduced, and the elimination half-life increases two- to threefold or greater when creatinine clearance is less than 40 mL/1.73 m^2 per minute. Acidic urine increases the excretion rate of amantadine, and urine-acidifying drugs may increase the elimination of the drug from the body. Amantadine is removed in negligible amounts by hemodialysis.

Pharmacodynamics

Mechanism of Action. The exact mechanism of action of amantadine in the treatment of Parkinson disease and drug-induced extrapyramidal reactions is not known. It has preferential selectivity for central catecholamine neurons and acts by potentiating the release of endogenous catecholamines and dopamine from presynaptic vesicles, blocking reuptake of dopamine and having an agonist effect on postsynaptic dopamine receptors. Amantadine is a weak, noncompetitive NMDA receptor antagonist, and some studies suggest that it may have direct and indirect effects on dopamine neurons. It has no anticholinergic activity, and though it is only 1/209,000 as potent as atropine, it does cause side effects like anticholinergics such as constipation, dry mouth, and urinary retention. When used alone or in combination for mild parkinsonism, many patients derive only transient benefit, if any at all.

Design and Interpretation of Clinical Trials

Amantadine has been used as an antiviral drug for more than 30 years; as early as 1976, positive effects of amantadine on neuroleptic extrapyramidal side effects have been described.

It is also used in Parkinson disease and parkinsonism, but its efficacy in this arena remains questionable. Many studies have compared the effectiveness of amantadine and anticholinergic agents, but to date, there are no conclusive data to establish its efficacy as a first-line agent. A double-blind study compared the effects of amantadine and biperiden in treating extrapyramidal side effects in 42 schizophrenic subjects treated with haloperidol (Haldol). Of note, the treatment effect of amantadine and biperiden was similar. This study supports the fact that amantadine is a useful alternative to anticholinergics. In another double-blind, placebo-controlled study, amantadine was compared to trihexyphenidyl in treating drug-induced parkinsonism. The study showed that amantadine not only produced far fewer adverse effects but also was equally effective as trihexyphenidyl and is a useful alternative to atropine-like drugs. Some case reports show that amantadine may alleviate symptoms of tardive dyskinesia. This is a controversial issue, and there are no controlled studies to replicate these results.

A new form of extended-release formulation of amantadine is currently under investigation and phase 2/3 clinical trials are underway to evaluate its efficacy in Parkinson disease. Preliminary data suggest that two of the tested doses resulted in a statistically significant reduction in dyskinesia. This study suggests that the extended-release formulation of amantadine is effective in decreasing motor fluctuations and has a side-effect profile similar to regular available form of amantadine. The new extended-release formulation is taken at bedtime and is slowly released into the bloodstream overnight. As a result, the peak levels of the drug are present when it is time to wake up.

Recent studies have also looked at weight loss potential, reversal of sexual dysfunction, and improvement in cognitive functioning by amantadine. One study assessed the weight loss potential of amantadine in patients who were treated with olanzapine (Zyprexa). Significantly fewer subjects taking amantadine gained weight, compared to the placebo group. Amantadine stabilized weight in patients treated with olanzapine but did not improve fasting insulin or lipid levels.

Amantadine has a predilection for the nigrostriatal dopaminergic system compared to the mesolimbic system and enhances the CNS dopaminergic activity. This can lead to delusion, paranoia, alterations and elevation of mood, confusion, and disorientation. A few case reports suggest psychotic decompensation in patients treated with amantadine. The vast majority who are treated for extrapyramidal symptoms show a lack of psychotic relapse. Clinicians should be alerted to the fact that patients treated with amantadine are vulnerable to psychotic decompensation.

A double-blind study showed that amantadine did not cause memory impairment in patients treated with the medication, and, therefore, it is a good alternative for the elderly who are more prone to these side effects.

Effects on Specific Organs and Systems

Amantadine relieves parkinsonian symptoms by increasing dopaminergic function and increasing the release of catecholamines. These effects can be manifold and manifest as adverse reactions. These include restlessness, depression, irritability, insomnia, agitation, excitement, and occasionally psychotic decompensation. Amantadine is excreted unchanged by the kidneys, and patients with renal dysfunction should be monitored closely.

Therapeutic Indications

Drug-Induced Extrapyramidal Reactions. Amantadine is indicated in the treatment of drug-induced extrapyramidal reactions, but there are no data that support the use of this drug as a first-line agent particularly for acute dystonic reactions. Some studies suggest that it is as effective as anticholinergics for extrapyramidal signs and symptoms but is not effective in treating tardive dyskinesia and akathisia. Although side effects have been noted with amantadine, the incidence of anticholinergic-type side effects is lower than that observed with the anticholinergic antiparkinsonian drugs. This is particularly useful in the elderly who are more sensitive to the anticholinergic effects, both in the CNS, such as anticholinergic delirium, and in the peripheral nervous system, such as urinary retention.

Parkinson Disease and Parkinsonism. The usefulness of amantadine in

the treatment of parkinsonism is controversial; its efficacy in comparison with the anticholinergic antiparkinsonian drugs has not yet been established. In a controlled trial of 17 patients with parkinsonism, amantadine hydrochloride in a dose of 100 mg b.i.d. showed a minor beneficial effect on physical signs but failed to demonstrate any significant effect on functional disabilities. This improvement in physical signs was only sustained for 8 weeks but numerous patients noticed deterioration in functional performance after 2 to 5 weeks of treatment.

Amantadine is less effective than levodopa but is indicated in the treatment of idiopathic Parkinson disease (paralysis agitans), postencephalitic parkinsonism, and symptomatic parkinsonism, which may follow injury to the nervous system by carbon monoxide intoxication. It is indicated in those elderly patients believed to have developed parkinsonism in association with cerebral arteriosclerosis.

Amantadine has been reported to be beneficial in the treatment of selective serotonin reuptake inhibitor–associated side effects, such as lethargy, fatigue, anorgasmia, and ejaculatory inhibition.

Brain Injury. Traumatic brain injury (TBI) is one of the leading causes of death and disability in persons between 15 and 30 years of age. Currently there are no interventions or treatment to change the prognosis of this condition, and it leads to chronic disability in more than 90,000 survivors of brain injury. The trauma to the brain usually impacts the frontal region, disrupting neurotransmitters, including dopamine, and causes frontal lobe dysfunction. This syndrome includes problems with aggression, disinhibition, attention, impulsivity, planning, perseveration, amotivation, apathy, memory deficits, mood dysregulation, and problem solving. Other causes of brain injury with similar sequelae include stroke, encephalitis, and aneurysm. The role of dopamine and lowered levels of its metabolites is well established in brain injury, specifically in the frontal lobe regions. Multiple studies have shown that dopamine is involved in the regulation of cognition, mood, and behavior. In patients with brain injury, there is deficiency and dysregulation of dopamine in the frontal lobes.

Amantadine was first used in 1988 in two patients for the treatment of TBI. Subsequently, a study assessed 30 patients aged 5 to 59 with TBI with behavior problems including agitation, mood lability, and inattention. Treatment with amantadine led to improvement in 63 percent of the patients. These positive results led to the hypothesis that since amantadine impacted the dopaminergic system (presynaptic and postsynaptic), it would be an appropriate treatment for patients with TBI with compromised dopaminergic systems. Numerous studies since then have validated these results, and amantadine is now considered safe and beneficial for arousal, agitation, and problems with executive functioning in patients with head injury. During the past 20 years, amantadine has been used frequently for

the treatment of head injuries.

Amantadine is one of the most frequently used medications for patients undergoing neurorehabilitation. Various authors have reported the use of amantadine alone or in combination therapy with positive response in patients with brain injury, specifically in the frontal lobe syndrome. Recent studies suggest that amantadine and other dopaminergic agents may substantially improve the level of executive functioning in these patients. The clinical effects can be seen within a few days to a week. In one study of 184 patients with TBI who were minimally conscious 1 to 4 months after the brain injury, amantadine use after 4 weeks showed a significantly faster recovery compared to the placebo group. Of note, despite the high risk of seizures in this population, amantadine did not increase the risk of adverse effects including any medical or neurological events.

Another study showed improvements in attention span, speed of information processing, learning, and alertness. Other studies have shown improvements in apathy, amotivation, and perseveration. Greatest improvements have been seen in reduction of agitation, physical aggression, distractibility, and mood swings. More recently, amantadine has shown promising results in improving clinical status in dementia patients. According to one report, it is also useful in treating fatigue and depression of multiple sclerosis. Despite the fact that these case reports show promising results, there are limitations to these small and uncontrolled studies, which underscores the need for double-blind, large, and controlled studies.

The exact mechanism of action remains unclear, but amantadine appears to act as an NMDA antagonist and indirect dopamine agonist. Amantadine may increase dopaminergic activity by increasing presynaptic release and blocking reuptake postsynaptically targeting the nigrostriatal, mesolimbic, and frontostriatal circuits. The final result is increased arousal, drive, and attention.

These results suggest that patients with severe TBI can be treated with amantadine safely at doses between 200 and 400 mg.

Treatment in Children

Head injury is a common problem in children: nearly 300 out of 100,000 children suffer TBI by the age of 18. Most children are able to survive these injuries yet are left with deficits including inattention, difficulties with executive functioning, and academic performance. The use of agents to treat these behavior disturbances secondary to head injuries in pediatric population is less well known, but a small number of studies suggest the use of amantadine. Despite the limited body of literature on the use of amantadine in pediatric population, it is a useful and beneficial treatment in children with head injury and symptoms of arousal and alertness.

Pathological Gambling and Drug Addiction

Pathological gambling (PG) is a persistent gambling behavior with devastating consequences and is identified as one of the substance abuse and addictive disorders. Despite significant research, there have been no FDA-approved treatments for this condition. Considering the fact that PG and addictive disorders may have similar clinical and biological underpinnings, it may be appropriate to use drugs for the treatment of PG that are used for drug addiction. Numerous studies suggest the role of dopamine in addictive, reinforcing, and rewarding behaviors. The pleasure and reward seeking behavior present in drug addiction is linked to hypodopaminergic state and may also be responsible for gambling behavior triggering the release of dopamine (DA) and consequently sense of pleasure. In addition, imbalance in glutamate (Glu) homeostasis may also be responsible for reward seeking behaviors.

One study suggested that amantadine may be effective in reducing PG and showed symptoms resolving after 4 days of treatment. Another case study of a male patient with PG, comorbid depression, and alcohol dependence showed that amantadine 100 to 150 mg daily had a 43 to 64 percent reduction in the Gambling Symptom Assessment Scale (G-SAS).

The data suggest that glutamatergic system may play an important role in PG, and amantadine could be a potential treatment option for this condition.

Amantadine exerts its effects on dopamine neurotransmission by increasing DA release, DA reuptake inhibition, and acts as an antagonist at NMDA glutamatergic receptors. These effects may explain its role as an antidepressant. Numerous studies have suggested that its dopaminergic properties may be beneficial in cocaine-dependent patients with withdrawal symptoms, but more studies are required to validate these results.

Precautions and Adverse Reactions

The most frequently reported CNS adverse reactions (5 to 10 percent) are insomnia, dizziness, and impaired concentration. Less frequently (1 to 5 percent) reported CNS side effects include restlessness, irritability, agitation, excitement, depression, anxiety, and confusion. More severe CNS side effects are seizures and psychotic symptoms, and overdose may produce acute toxic psychosis.

Patient treated with anti-Parkinson medications including amantadine may have increased sexual urges and to spend money uncontrollably, urges to gamble, and other intense urges. Medications that increase central dopaminergic tone like amantadine may impair a patient's ability to control such urges. Prescribers should routinely inquire about the development of these urges (gambling, sexual, uncontrolled spending or other urges) and

lower the dose of amantadine or completely stop the medication if a patient develops such urges.

Other less frequently (1 to 5 percent) reported adverse reactions are anorexia, dry mouth, peripheral edema, orthostatic hypotension, congestive heart failure, headache, somnolence, nervousness, dream abnormality, dry nose, diarrhea, and fatigue. Livedo reticularis of the legs (purple discoloration of the skin caused by dilatation of the vessels) may occur in patients and usually clears within a month after discontinuation of the drug.

Rare (less than 0.1 percent) occurring adverse reactions are instances of convulsion, leukopenia, neutropenia, eczematoid dermatitis, oculogyric episodes, suicidal attempt, suicide, and suicidal ideation.

Amantadine is relatively contraindicated in patients with known history of seizures or renal disease. It should also be used with caution in persons with congestive heart failure or edema. There is evidence to suggest that it is teratogenic, and hence, should not be used by pregnant women or women who are breastfeeding.

Amantadine should not be discontinued abruptly in patients with Parkinson disease, since a few patients have experienced a parkinsonian crisis (i.e., a sudden marked clinical deterioration upon abrupt cessation of this agent) (Table 34.6–3).

Deaths

Amantadine has been associated with deaths secondary to overdose in amounts as low as 2 g. Anticholinergics effects may be responsible for these deaths secondary to cardiac, respiratory, renal, or CNS toxicity.

Table 34.6–3.
Adverse Effects of Amantadine

Cardiovascular system
► Arrhythmia
► Tachycardia
Nervous system
► Delirium
► Paranoia
► Paresthesias

Drug Interaction

Coadministration of amantadine with phenelzine (Nardil) or other MAOIs may result in a significant increase in blood pressure. There is an increased likelihood of side effects such as confusion, hallucinations, nightmares, dry mouth, and blurred vision if administered with anticholinergics. The coadministration with CNS stimulants can result in insomnia, irritability,

nervousness, and possibly seizures.

Laboratory Interferences

Amantadine is not known to alter the results of any standard laboratory tests and values.

Dosage and Administration

Adult. The usual starting dose of amantadine hydrochloride is 100 mg twice a day when used alone and can be cautiously increased to 200 mg twice a day in patients whose responses are not optimal. Amantadine has an onset of action usually within 48 hours.

The initial dose of amantadine hydrochloride is 100 mg daily for patients with serious associated medical illnesses or who are receiving high doses of other anti-Parkinson drugs. After one to several weeks at 100 mg once daily, the dose may be increased to 100 mg twice daily, if necessary.

Some patients experience a fall-off of effectiveness after a few months and may benefit by increasing the dose to 300 mg daily or temporarily discontinuing amantadine for several weeks, followed by reinitiation of the drug.

Dosage for Concomitant Therapy

When patients do not respond to monotherapy with anticholinergic anti-Parkinson drugs, switching to or concomitant therapy with amantadine may produce additional benefits. In some cases when amantadine and levodopa are initiated concurrently, the patient can exhibit rapid therapeutic benefits. Amantadine should be held constant at 100 mg daily or twice daily while the daily dose of levodopa is gradually increased to optimal benefit.

Dosage for Drug-Induced Extrapyramidal Reactions

The usual starting dose of amantadine hydrochloride is 100 mg twice a day. Occasionally, patients whose responses are not optimal with amantadine hydrochloride at 200 mg daily may benefit from an increase up to 300 mg daily in divided doses (Table 34.6–4).

Table 34.6–4.
Dosage for Impaired Renal Function

Creatinine Clearance (mL/min/1.73 m^2)	Amantadine Hydrochloride Dosage
30–50	200 mg 1st day and 100 mg each day thereafter
15–29	200 mg 1st day followed by 100 mg on alternate days
<15	200 mg every 7 days

The recommended dosage for patients on hemodialysis is 200 mg every 7 days.

REFERENCES

Buhrich N, Weller A, Kevans P. Misuse of anticholinergic drugs by people with serious mental illness. *Psychiatry Serv*. 2000;51:928–929.

Caligiuri MR, Jeste DV, Lacro JP. Antipsychotic-induced movement disorders in the elderly: Epidemiology and treatment recommendations. *Drugs Aging*. 2000;17:363–384.

Comings DE, Wu S, Rostamkhani M, McGue M, Iacono WG, MacMurray JP. Association of the muscarinic cholinergic 2 receptor (CHRM2) gene with major depression in women. *Am J Med Genet*. 2002;114(5):527–529.

Dose M, Tempel HD. Abuse potential of anticholinergics. *Pharmacopsychiatry*. 2000;33:43–46.

Drevets WC, Furey ML. Replication of scopolamine's antidepressant efficacy in major depressive disorder: a randomized, placebo-controlled clinical trial. *Biol Psychiatry*. 2010;67(5):432–438.

*Drimer T, Shahal B, Barak Y. Effects of discontinuation of long-term anticholinergic treatment in elderly schizophrenia patients. *Int Clin Psychopharmacol*. 2004;19(1):27–29.

Furey ML, Drevets WC. Antidepressant efficacy of the antimuscarinic drug scopolamine: a randomized, placebo-controlled clinical trial. *Arch Gen Psychiatry*. 2006;63(10):1121–1129.

Furey ML, Drevets WC, Hoffman EM, Frankel E, Speer AM, Zarate CA Jr. Potential of pretreatment neural activity in the visual cortex during emotional processing to predict treatment response to scopolamine in major depressive disorder: *JAMA Psychiatry*. 2013;70(3):280–290.

Furey ML, Khanna A, Hoffman EM, Drevets WC. Scopolamine produces larger antidepressant and anti-anxiety effects in women than in men. *Neuropsychopharmacology*. 2010;35(12):2479–2488.

Gelenberg AJ, Van Putten T, Lavori PW, et al. Anticholinergic effects on memory: Benztropine versus amantadine. *J Clin Psychopharmacol*. 1989;9(3):180–185.

Giacino JT, Whyte J, Bagiella E, et al. Placebo-controlled trial of amantadine for severe traumatic brain injury. *N Engl J Med*. 2012;366(9):819–826.

Graham KA, Gu H, Lieberman JA, Harp JB, Perkins DO. Double-blind, placebo-controlled investigation of amantadine for weight loss in subjects who gained weight with olanzapine. *Am J Psychiatry*. 2005;162(9):1744–1746.

Han C, Pae CU. Oral scopolamine augmentation for major depression. *Expert Rev Neurother*. 2013;13(1):19–21.

Hodgins DC, Stea JN, Grant JE. Gambling disorders. *Lancet*. 2011;378(9806):1874–1884.

Hughes S, Colantonio A, Santaguida PL, Paton T. Amantadine to enhance readiness for rehabilitation following severe traumatic brain injury. *Brain Inj*. 2005;19(14):1197–1206.

Janowsky DS. Serendipity strikes again: scopolamine as an antidepressant agent in bipolar depressed patients. *Curr Psychiatry Rep*. 2011;13(6):443–445.

Janowsky DS, Overstreet DH, Nurnberger JI Jr. Is cholinergic sensitivity a genetic marker for the affective disorders? *Am J Med Genet*. 1994;54(4):335–344.

Ji CX, Zhang JJ. Effect of scopolamine on depression in mice [in Chinese]. *Yao Xue Xue Bao*. 2011;46(4):400–405.

Kampman KM, Volpicelli JR, Alterman AI, Cornish J, O'Brien CP. Amantadine in the treatment of cocaine-dependent patients with severe withdrawal symptoms *Am J Psychiatry*.

2000;157(12):2052–2054.

Krystal JH. N-methyl-D-aspartate glutamate receptor antagonists and the promise of rapid-acting antidepressants. *Arch Gen Psychiatry*. 2010;67(11):1110–1111.

Krystal JH, Sanacora G, Blumberg H, et al. Glutamate and GABA systems as targets for novel antidepressant and mood-stabilizing treatments: *Mol Psychiatry*. 2002;7(suppl 1):S71–S80.

Mamo DC, Sweet RA, Keshavan MS. Managing antipsychotic-induced parkinsonism. *Drug Saf*. 1999;20(3):269–275.

Paul IA, Skolnick P. Glutamate and depression: clinical and preclinical studies. *Ann N Y Acad Sci*. 2003;1003(1):250–272.

Pettorruso M, Martinotti G, Di Nicola M, et al. Amantadine in the treatment of pathological gambling: a case report. *Front Psychiatry*. 2012;3:102.

Potenza MN. The neurobiology of pathological gambling and drug addiction: an overview and new findings. *Philos Trans R Soc Lond B Biol Sci*. 2008;363(1507):3181–3189.

Schatzberg AF. Drugs to treat extrapyramidal side effects. In: Schatzberg AF, Nemeroff CB, eds. *The American Psychiatric Publishing Textbook of Psychopharmacology*. 3rd ed. Washington, DC: American Psychiatric Publishing; 2004:284.

Staskin DR, Zoltan E. Anticholinergics and central nervous system effects: are we confused? *Rev Urol*. 2007;9(4):191–196.

Stromberg U, Svensson TH, Waldeck B. On the mode of action of amantadine. *J Pharm Pharmacol*. 1970;22(12):959–962.

Tandon R, Dequardo JR. Treatment of schizophrenia with anticholinergic medications. *Am J Psychiatry*. 1995;152(5):814–815.

Wang JC, Hinrichs AL, Stock H, et al. Evidence of common and specific genetic effects: association of the muscarinic acetylcholine receptor M2 (CHRM2) gene with alcohol dependence and major depressive syndrome: *Hum Mol Genet*. 2004;13(17):1903–1911.

Williams SE. Amantadine treatment following traumatic brain injury in children. *Brain Inj*. 2007;21(9):885–889.

▲ 34.7 Anticonvulsants: Gabapentin, Topiramate, Tiagabine, Levetiracetam, Zonisamide, and Pregabalin

PO W. WANG, M.D., AND TERENCE ARTHUR KETTER, M.D.

Newer anticonvulsants, marketed over the past two decades, have generally been developed to have enhanced tolerability, simpler kinetics, and fewer drug–drug interactions. Chemical structures (Fig. 34.7–1) and pharmacodynamic effects of these drugs are diverse. As some purported mechanisms of action include enhancing neuronal inhibition by increasing GABAergic function or decreasing neuronal excitation by decreasing glutamatergic function, these agents have potential psychotropic benefits, and these amino acid neurotransmitters have been implicated in psychiatric disorders. Thus, newer anticonvulsants appear to have not only distinctive efficacy profiles for the treatment of epilepsy but also diverse

psychotropic and analgesic effects. A few of these agents can provide psychotropic benefits, and have received FDA indications for acute mania (valproate and carbamazepine) and bipolar maintenance treatment (lamotrigine).

The FDA announced in 2008 a class warning regarding increased risk of suicidality (suicidal behavior or ideation) in patients with epilepsy as well as psychiatric disorders based on data from diverse anticonvulsants, six of which are described in detail in this chapter (gabapentin [Neurontin], topiramate [Topamax], tiagabine [Gabitril], levetiracetam [Keppra], zonisamide [Zonegran], and pregabalin [Lyrica]) and five of which are not described in this chapter (carbamazepine [Tegretol], felbamate [Felbatol], lamotrigine [Lamictal], oxcarbazepine [Trileptal], and valproate [Depakote]). This warning has been applied to seven subsequent anticonvulsants, which are not described in this chapter (rufinamide [Banzel], lacosamide [Vimpat], vigabatrin [Sabril], esogabine [Potiga], clobazam [Onfi], perampanel [Fycompa], and eslicarbazepine acetate [Aptiom]). In the FDA's analysis, anticonvulsants compared to placebo yielded approximately twice the risk of suicidality (0.43 percent vs. 0.22 percent). The relative risk for suicidality was higher in patients with epilepsy compared to psychiatric disorders.

The FDA has also stipulated additional behavioral toxicity warnings for seven diverse anticonvulsants (gabapentin, topiramate, tiagabine, levetiracetam, zonisamide, esogabine, and perampanel), including for perampanel a boxed warning for treatment-emergent irritability and homicidal ideation. However, additional behavioral toxicity warnings are lacking for 11 other diverse anticonvulsants (carbamazepine, valproate, lamotrigine, felbamate, oxcarbazepine, pregabalin, rufinamide, lacosamide, vigabatrin, clobazam, and eslicarbazepine acetate). Thus, to date, anticonvulsants with FDA approvals for bipolar disorder (carbamazepine, valproate, and lamotrigine) all lack behavioral toxicity warnings.

GABAPENTIN

Introduction and History

Gabapentin (Neurontin), which has a biochemical structure similar to the inhibitory amino acid neurotransmitter GABA, was developed by Warner–Lambert in efforts to create an agent for the treatment of spasticity. Gabapentin proved to have greater efficacy as an anticonvulsant than antispasticity agent. Gabapentin was found effective in several animal models of epilepsy, developed as an antiepileptic drug, and marketed for the adjunctive treatment of partial seizures in the United States in 1994 by Parke–Davis.

Due to generally good tolerability, fewer clinically significant drug–drug interactions than older anticonvulsants, and initial, postmarketing, open,

uncontrolled clinical experience suggesting efficacy in a variety of neurological and psychiatric conditions, gabapentin became increasingly used in conditions other than seizure disorders. In fact, nonepilepsy uses substantially exceeded those in epilepsy. Double-blind, placebo-controlled studies later confirmed gabapentin efficacy for some, but not all, of the neurological and psychiatric conditions suggested in open studies. Specifically, gabapentin appeared effective in neuropathic pain and was marketed for the management of postherpetic neuralgia in adults. Gabapentin also appeared effective in anxiety disorders (social phobia and panic disorder), but not as a main intervention in mania or treatment-resistant mood disorders. Thus, gabapentin has emerged as a safe and effective agent in epilepsy and postherpetic neuralgia, and has more widespread use adjunctively to relieve symptoms and pain, and improving, anxiety, and insomnia in a dimensional fashion across diverse neurological and psychiatric conditions. Moreover, gabapentin enacarbil (a prodrug of gabapentin) was effective in moderate to severe restless legs syndrome. Finally, recent controlled data suggested that gabapentin could be effective for relapse prevention in people with alcohol dependence.

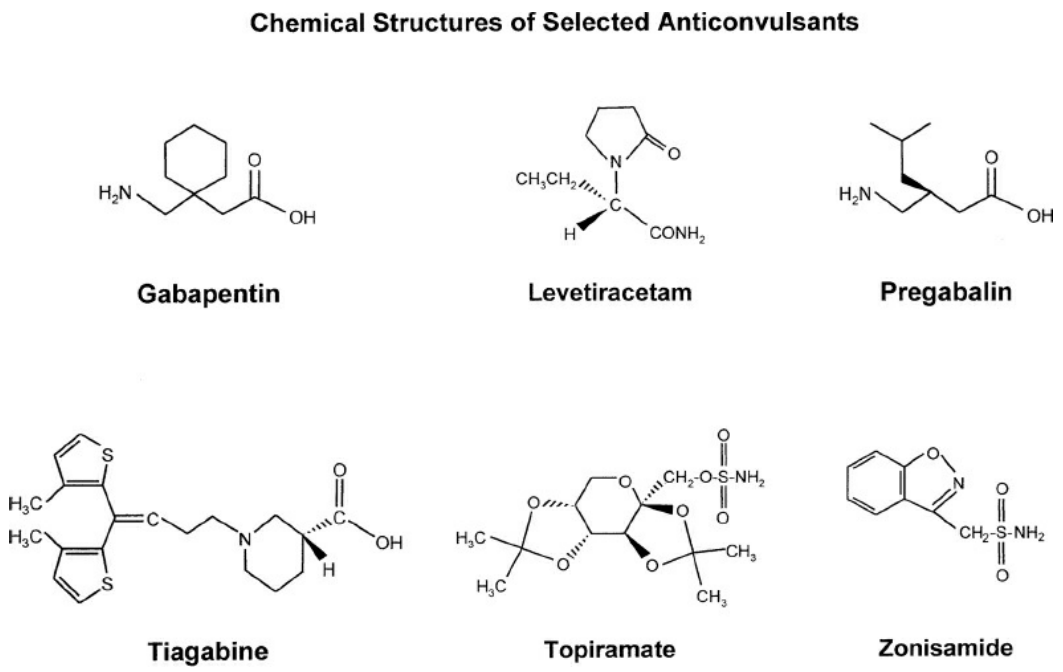

FIGURE 34.7–1. Chemical structures of selected anticonvulsants.

Chemistry

Gabapentin, 1-(aminomethyl)-cyclohexane acetic acid, is a novel amino acid derived from adding a cyclohexyl group to GABA (Fig. 34.7–1), with some structural similarity to other anticonvulsants that resemble GABA, such as valproate. Immediate-release gabapentin preparations currently marketed in the United States include: 100-, 300-, and 400-mg capsules; 600- and 800-mg tablets; and a 250-mg/5-mL oral solution. Generic

immediate-release gabapentin is available in comparable formulations. Extended-release gabapentin and gabapentin enacarbil (a prodrug of gabapentin) are available in 300- and 600-mg tablets.

Pharmacological Actions

Pharmacokinetics. Gabapentin is available in immediate-release (Neurontin™) and sustained-release (Gralise™) formulations. In addition, gabapentin enacarbil is available in an extended-release formulation (Horizant™). Immediate-release gabapentin has sublinear kinetics due to saturable gastrointestinal absorption, with oral bioavailability of about 60 percent with 300 mg three times a day that decreases to 33 percent with 1,200 mg two times a day. In practice, maximum individual doses of immediate-release gabapentin should be no more than 1,200 mg. Sustained-release gabapentin can be titrated up to 1,800 mg once daily with the evening meal. Gabapentin enacarbil was designed to have high capacity absorption compared to gabapentin, even in doses greater than 1,200 mg, but the recommended dosing is only 600 mg once or twice daily. Sustained-release gabapentin and gabapentin enacarbil are not interchangeable with other gabapentin products because of differing pharmacokinetic profiles that affect frequency of administration. Food only modestly affects the rate and extent of immediate-release gabapentin absorption, but more significantly affects absorption of sustained-release gabapentin and gabapentin enacarbil, which need to be taken with food. Gabapentin circulates in the blood largely unbound, with less than 3 percent bound to plasma proteins, and has a moderate volume of distribution of about 1 L/kg. In epilepsy patients, cerebrospinal fluid gabapentin concentrations are about 20 percent of plasma levels. Immediate-release gabapentin has a half-life of about 6 hours and thus is given in three divided doses in epilepsy patients. Sustained-release gabapentin and gabapentin enacarbil have an elimination half-life of approximately 6 hours, but due to differential absorption and indications, are dosed once daily and once to twice daily, respectively. Gabapentin is not appreciably metabolized in humans, is eliminated unchanged by renal excretion, and can be removed by hemodialysis. Gabapentin clearance correlates with that of creatinine, so dosage adjustments are required in patients with impaired renal function. Clearance may be increased in children under age 5 years and decreased in the elderly, requiring dosage adjustments.

Pharmacodynamics. Like most other anticonvulsants, gabapentin has multiple biochemical actions, and the relationships between these and direct clinical effects remain to be established. Although gabapentin is a structural analog of GABA, it appears to lack direct GABAergic actions. Gabapentin does appear to have indirect GABAergic effects, as both acute

and chronic gabapentin increase cerebral GABA in humans. Gabapentin may also have antiglutamatergic actions, mediated by inhibition of glutamate synthesis by branched chain amino acid aminotransferase. It also modulates (but does not directly block) sodium channels and increases human whole blood serotonin concentrations. Gabapentin has high-affinity regional (including neocortical and hippocampal) binding to $\alpha_2\delta$ subunits of voltage-activated calcium channels and thus may modulate calcium channels to reduce monoamine release. Several of these biochemical mechanisms may underlie gabapentin's psychotropic effects.

Gabapentin has antiseizure activity in maximal electroshock, pentylenetetrazole, amygdala kindled, and other animal models of seizures, antinociceptive effects in several animal models of neuropathic pain, antispasticity activity in several animal models, and anxiolytic-like effects in animal models of anxiety disorders such as the rat conflict and elevated plus-maze, and the mouse light/dark box paradigms. Gabapentin was not active in the behavioral despair model of depression, or the Sidman avoidance paradigm, which screens for antipsychotic activity.

Clinical Drug Studies

Seizure Disorders. Adjunctive immediate-release gabapentin was more effective than adjunctive placebo in three multicenter, double-blind, placebo-controlled trials involving adults and adolescents (age 12 years or more) with refractory partial seizures, and in one trial in children (age 3 to 12 years) with refractory partial seizures. In a small trial, immediate-release gabapentin monotherapy did not differ from placebo in children (age 4 to 12 years) with newly diagnosed absence epilepsy.

Postherpetic Neuralgia. Immediate-release gabapentin appeared more effective than placebo in two multicenter, double-blind, placebo-controlled trials in adults with postherpetic neuralgia, relieving pain and sleep interference, and improving mood and quality of life. Two more recent multicenter, double-blind, placebo-controlled trials with gabapentin extended-release formulation in adults with postherpetic neuralgia demonstrated efficacy. Moreover, gabapentin enacarbil, appeared more effective than placebo in three multicenter, double-blind, placebo-controlled trials in adults with postherpetic neuralgia.

Other Pain Syndromes. Gabapentin appears to have analgesic effects in other pain syndromes, particularly in neuropathic pain. Thus, immediate-release gabapentin was more effective than placebo in relieving pain and in some instances yielded benefits on other measures such as sleep, mood, and quality of life in diabetic neuropathy, trigeminal neuralgia, neuropathic cancer pain (when added to opioids), Guillain–Barré syndrome, fibromyalgia, painful human immunodeficiency virus (HIV)-associated

sensory neuropathy, and diverse neuropathic pain. Immediate-release gabapentin was also more effective than placebo in relieving postamputation pain (in one of two studies), postradical mastectomy pain, postlumbar discectomy pain (when added to fentanyl [Actiq] but not when added to morphine [Roxanol]), postabdominal hysterectomy pain (when added to rofecoxib [Vioxx]), and catheter-related bladder discomfort and postoperative pain in percutaneous nephrolithotomy patients. Immediate-release gabapentin was more effective than placebo in chronic masticatory myalgia, complex regional pain syndrome type one (formerly known as reflex sympathetic dystrophy), and in suppressing nausea and vomiting and postoperative rescue analgesic requirements in laparoscopic cholecystectomy patients. Immediate-release gabapentin compared to placebo reduced migraine frequency and intensity in migraine patients, relieved symptoms and improved function and quality of life in chronic daily headache patients, and may attenuate hypobaric hypoxia-induced headache (a syndrome commonly experienced by newcomers to high-altitude areas). However, immediate-release gabapentin was no better than placebo for chemotherapy-induced peripheral neuropathy.

Restless Legs Syndrome. Immediate-release gabapentin appeared more effective than placebo in one small cross-over design study of adults, reducing periodic leg movements, pain, sleep quality and sleep architecture, and overall clinical impression, and in another small cross-over design study of adult hemodialysis patients. More recently, gabapentin enacarbil was more effective than placebo in two multicenter, double-blind, placebo-controlled trials in adults with restless legs syndrome, relieving restless legs movements and improving sleep quality.

Other Neurological Disorders. Limited data suggest gabapentin may have efficacy in relieving symptoms of other neurological disorders. Immediate-release gabapentin compared to placebo relieved essential tremor in two of three studies, primary orthostatic tremor in one study, and impairment of spasticity in multiple sclerosis in one study. Also, compared to placebo, immediate-release gabapentin attenuated rigidity, bradykinesia, and tremor in Parkinson disease. Immediate-release gabapentin compared to no treatment improved lower but not upper extremity function in spinal muscular atrophy patients. However, compared to placebo, immediate-release gabapentin failed to slow the decline of arm strength in amyotrophic lateral sclerosis patients in two studies and failed to increase muscle strength in spinal muscular atrophy patients in another.

Other Syndromes. Immediate-release gabapentin was more effective than placebo in treating pruritus related to uremia (but not related to cholestasis) and in controlling hot flashes in healthy postmenopausal

women and in women with breast cancer.

Alcohol and Substance Use Disorders. Four randomized, placebo-controlled studies in alcohol use disorders have found immediate-release gabapentin efficacious as monotherapy or in combination with either flumazenil or naltrexone, specifically associated with more abstinence days and longer time to relapse. Efficacy in opiate use disorders was found in two randomized, placebo-controlled add-on (to methadone or buprenorphine) studies with less opiate use and less withdrawal symptoms, but not in one study of monotherapy gabapentin. Efficacy in stimulant (methamphetamine or cocaine) use disorders is not well established for gabapentin where only one randomized, placebo-controlled study adding immediate-release gabapentin to flumazenil had less drug use and cravings, but three controlled studies (monotherapy and add-on) found no difference from placebo. One randomized, placebo-controlled study of immediate-release gabapentin monotherapy for marijuana use disorder found less use and improved executive functioning on neuropsychological testing. Thus, gabapentin has the strongest evidence of efficacy for alcohol use disorders, with less robust evidence in other substances of abuse.

Anxiety Disorders. Controlled studies have confirmed immediate-release gabapentin utility in some anxiety disorders, with double-blind, placebo-controlled trials demonstrating efficacy in social phobia (social anxiety disorder), and in a post hoc analysis, in moderate to severe panic disorder. Efficacy of immediate-release gabapentin in other anxiety disorders as suggested by open reports remains to be confirmed in controlled studies.

Mood Disorders. Two controlled studies found immediate-release gabapentin was ineffective, despite early, open, uncontrolled reports suggested that adjunctive immediate-release gabapentin was well tolerated and may offer benefit in patients with refractory mood disorders. In a double-blind, placebo-controlled study of bipolar patients suffering from mania or hypomania while taking lithium (Eskalith), valproate, or lithium and valproate, adjunctive placebo tended to lower mania scores better than adjunctive immediate-release gabapentin. In treatment-resistant (primarily rapid-cycling bipolar) mood disorders, immediate-release gabapentin monotherapy was no more effective than placebo, while lamotrigine monotherapy was more effective than both immediate-release gabapentin and placebo. However, a small (25 patients), randomized clinical trial suggested that adjunctive immediate-release gabapentin compared to adjunctive placebo might provide some prophylactic benefit in patients with bipolar disorder. Although antidepressant efficacy of adjunctive immediate-release gabapentin in bipolar disorder patients has been

suggested in several open trials, controlled studies have not supported the primary use of gabapentin for mood disorders.

Therapeutic Indications

Seizure Disorders. Immediate-release gabapentin is indicated as adjunctive treatment of partial seizures with and without secondary generalization in epilepsy patients over age 12 years, and as adjunctive treatment of partial seizures in pediatric patients age 3 to 12 years.

Postherpetic Neuralgia. Immediate- and extended-release gabapentin and gabapentin enacarbil are indicated for the management of postherpetic neuralgia in adults.

Restless Legs Syndrome. Gabapentin enacarbil is indicated for the treatment of moderate to severe restless legs syndrome in adults.

Precautions and Adverse Reactions

In general, gabapentin's adverse effects are mild to moderate, can be attenuated by dosage rescheduling or decreases, and do not typically require discontinuation. In adult epilepsy patients, the most common adverse effects with adjunctive immediate-release gabapentin were somnolence, dizziness, ataxia, fatigue, and nystagmus. In adult postherpetic neuralgia patients, the most common adverse effects with immediate-release gabapentin were dizziness, somnolence, and peripheral edema. In pediatric (age 3 to 12 years) epilepsy patients, the most common adverse effects with adjunctive gabapentin are viral infection, fever, nausea or vomiting, somnolence, and hostility. Interestingly, psychiatric adverse effects in pediatric epilepsy patients have been sufficiently troublesome that the immediate-release gabapentin United States Prescribing Information (USPI) includes warnings of the risks of (1) emotional lability or behavioral problems in 6 percent (vs. 1.3 percent with placebo); (2) hostility or aggressiveness in 5.2 percent (vs. 1.3 percent with placebo); (3) thought disorder, concentration, or academic performance problems in 1.7 percent (vs. 0 percent with placebo); and (4) hyperkinesia, restlessness, or hyperactivity in 4.7 percent (vs. 2.9 percent with placebo). As with other anticonvulsants, gabapentin has a warning of increased risk of suicidality (suicidal behavior or ideation) in patients with epilepsy as well as psychiatric disorders. Gabapentin can also cause driving impairment. Acute oral overdoses of immediate-release gabapentin (up to 49 g) have been associated with diplopia, slurred speech, drowsiness, lethargy, and diarrhea, but all patients recovered.

Gabapentin was classified as pregnancy category C, as fetal loss and teratogenicity have been demonstrated in animals, but to date not in

humans. Gabapentin was classified as "Probably Safe" for lactation, though it is excreted in breast milk, with nursing infants receiving about 1 mg/kg per day, so that gabapentin should be administered to women who are breastfeeding only if the potential benefits clearly outweigh the (unknown) risks.

Gabapentin is contraindicated in patients with known hypersensitivity to gabapentin or its ingredients.

Drug Interactions

Gabapentin bioavailability may decrease as much as 20 percent when administered at the same time as antacids (such as Maalox [Novartis]), so that gabapentin should be taken at least 2 hours following antacids. Hydrocodone and morphine may increase blood gabapentin concentrations. However, as gabapentin is renally excreted, it generally lacks clinically significant hepatic drug–drug interactions. In addition, gabapentin may not have major drug–drug interactions with renally excreted drugs. For example, chronic gabapentin did not appear to alter the single dose pharmacokinetics of lithium in subjects with normal renal function.

Dosage and Administration

The immediate-release gabapentin USPI recommends that adjunctive therapy in adolescents and adults with epilepsy should start at 300 mg three times a day and be increased up to 1,800 mg per day in three divided doses. However, in adults with epilepsy, immediate-release gabapentin has been generally well tolerated at doses as high as 3,600 and 2,400 mg per day in short- and long-term studies, respectively. Adjunctive immediate-release gabapentin in pediatric seizures is recommended to start with 10 to 15 mg/kg per day in three divided doses and to increase over three days to 25 to 40 mg/kg per day in three divided doses. A long-term epilepsy study suggested that immediate-release gabapentin doses up to 50 mg/kg per day may be well tolerated. Plasma gabapentin concentrations in epilepsy patients commonly ranged between approximately 12 and 20 μg/mL (70 and 120 μM/L), but did not appear to be systematically related to antiseizure effects and thus are not routinely monitored.

Immediate-release gabapentin for adults with postherpetic neuralgia or neuropathic pain is recommended to start with 300 mg on day 1, increase to 600 mg on day 2, 900 mg on day 3, and subsequently increase as needed to relieve pain up to 1,800 mg per day in three divided doses. Extended-release gabapentin for postherpetic neuralgia is started at 300 mg daily, increased to 600 mg on day 2, then further increased 300 mg every 4 days, with maximum dose of 1,800 mg. Gabapentin enacarbil for postherpetic neuralgia is started at 600 mg daily with food for 3 days, then increased to

600 mg twice daily, and for restless legs syndrome is started and maintained at 600 mg per evening with food.

In patients with renal impairment, gabapentin doses may need to be decreased 30 to 80 percent, depending on the degree of renal insufficiency.

Off-label gabapentin is typically used adjunctively with similar doses. In some (nonepilepsy) applications, clinicians may administer more of the dosage at bedtime to enhance sleep and decrease daytime somnolence. However, individual doses of immediate-release gabapentin above 1,200 mg are divided due to saturable absorption. In psychiatric outpatients, immediate-release gabapentin started at 300 mg at bedtime and increased by 300 mg every 4 days as necessary, initially targeting a dose of 1,200 mg at bedtime, is typically well tolerated. Occasionally, elderly patients or patients on complex combination therapies may need doses as low as 100 to 200 mg due to sedation. Final total daily doses tend to be between 1,200 and 2,400 mg per day, but may range as high as 4,800 mg per day.

TOPIRAMATE

Introduction and History

Topiramate (Topamax) was developed by R. W. Johnson Pharmaceutical Research Institute. After effectiveness was seen in several animal models of epilepsy, topiramate was studied as an antiepileptic drug, and then marketed for the adjunctive treatment of epilepsy in the United States in 1997 by Ortho–McNeil Pharmaceutical, Inc.

Double-blind, placebo-controlled studies revealed that topiramate was useful for some, but not all, of the neurological and psychiatric conditions reported in initial postmarketing, open, clinical studies. Topiramate appeared effective in migraine prevention and was approved for this indication. Also, topiramate appeared effective in obesity, bulimia, binge eating disorder with obesity, and alcohol dependence. In contrast, topiramate was not effective as a main intervention in mania in adults with bipolar disorder.

Chemistry

Topiramate, 2,3:4,5-bis-*O*-(1-methylethylidene)β-*D*-fructo-pyranose sulfamate, has a novel sulfamate-substituted monosaccharide structure (Fig. 34.7–1). Two oral topiramate preparations are currently marketed in the United States: 25-, 50-, 100-, and 200-mg tablets; and 15- and 25-mg sprinkle capsules, which can be swallowed whole or opened and sprinkled on food. Generic topiramate is available.

Pharmacological Actions

Pharmacokinetics. Topiramate has oral bioavailability of about 80 percent and linear kinetics that are not significantly altered by food.

Topiramate is only 15 percent bound to plasma proteins and has a moderate volume of distribution of about 0.8 L/kg. Topiramate is about 70 percent excreted unchanged in the urine, so in the absence of enzyme inducers, clearance does not extensively involve hepatic mechanisms. However, in the presence of hepatic enzyme inducers (such as carbamazepine or phenytoin), hepatic metabolism increases and renal clearance decreases about 50 percent. Topiramate has a half-life of approximately 24 hours with monotherapy; half-life decreases about 50 percent with enzyme inducers, so that topiramate is given in two divided doses. Topiramate clearance is generally about 25 mL per minute. Enzyme inducers can almost double clearance rates, and moderate to severe renal impairment decreases clearance rates by about 50 percent. Topiramate clearance may be also decreased in patients with hepatic impairment, requiring dosage adjustment. Thus, dosage needs to be increased with concurrent inducers, decreased with renal insufficiency, and possibly decreased in patients with hepatic impairment.

Pharmacodynamics. Like most other anticonvulsants, topiramate has multiple biochemical actions, and the relationships between these and clinical effects remain to be established. Topiramate is a state-dependent sodium channel blocker. Topiramate also has GABAergic effects; as it positively modulates $GABA_A$ receptors and both acute and chronic administration appear to increase cerebral GABA in humans. Topiramate has antiglutamatergic effects, as it blocks kainite and α-amino-3-hydroxy-5-methylisoxazole-4-proprionic acid (AMPA) receptors. Topiramate also weakly inhibits carbonic anhydrase CA-II and CA-IV isoenzymes.

Topiramate has antiseizure activity in maximal electroshock, amygdala kindled, and other animal models of seizures, cerebroprotective actions in several animal models of ischemia, activity in animal models of obesity, and some evidence of antinociceptive activity in animal models of neuropathic pain.

Clinical Drug Studies

Seizure Disorders. Topiramate was superior to placebo as initial monotherapy in adults and children 10 years of age and older with partial-onset or primary generalized seizures in a multicenter, double-blind, placebo-controlled trial. Adjunctive topiramate appeared more effective than adjunctive placebo in several multicenter, double-blind, placebo-controlled epilepsy trials including six multiple-dose and four single-dose trials in patients with partial-onset seizures, one trial in children age 2 to 16 years with partial-onset seizures, one trial in patients age 2 years or older with primary generalized tonic–clonic seizures, and one trial in patients age 2 years or older with Lennox-Gastaut syndrome.

Migraine Prevention. Topiramate was more effective than placebo in two multicenter, double-blind, placebo-controlled trials, decreasing frequency of migraine headaches. Topiramate also appeared to improve daily activities and patient functioning.

Pain Syndromes. Adjunctive topiramate was more effective than placebo in painful diabetic neuropathy in one of four studies and in chronic low back pain in one study.

Other Neurological Disorders. Adjunctive topiramate was more effective than placebo in attenuating essential tremor in two of three studies, but was ineffective in patients with amyotrophic lateral sclerosis.

Weight Loss, Obesity, Bulimia, and Binge Eating Disorder with Obesity. Topiramate has been associated with weight loss in controlled studies of patients with epilepsy, psychiatric illness, and obesity and has shown efficacy in the treatment of adiposity in women undergoing olanzapine (Zyprexa) therapy, bulimia, and binge eating disorder associated with obesity in controlled trials. In 2012, the combination of topiramate with phentermine was FDA approved as an adjunct to a reduced-calorie diet and increased physical activity for chronic weight management in adults with obesity.

Alcohol Use Disorder. Topiramate was superior to placebo in patients with alcohol dependence in two studies.

Smoking Cessation. In a controlled trial, topiramate promoted smoking abstinence among alcohol-dependent smokers.

Substance Use Disorders. In a controlled trial, topiramate was not better than placebo in promoting methamphetamine abstinence in patients with methamphetamine dependence. However, a separate controlled study found that topiramate improved some (retainment in treatment) but not all measures (alcohol or cocaine use) in patients with comorbid cocaine and alcohol dependence.

Bipolar Disorders. Although early open reports were encouraging, four double-blind, placebo-controlled trials failed to demonstrate topiramate monotherapy efficacy in acute mania in adults. A similar trial in children and adolescents was inconclusive, perhaps due to premature termination. In a controlled trial, adding topiramate to lithium or valproate was no better than adding placebo in patients with acute manic or mixed episodes. In one small single-blind study, adjunctive topiramate and adjunctive sustained-release bupropion (Wellbutrin) had similar efficacy in bipolar depression, although topiramate appeared to yield more weight loss and

adverse effects.

Psychotic Disorders. Case reports and small case series in patients with treatment-resistant schizophrenia suggested that open topiramate added to antipsychotics may help attenuate clozapine (Clozaril) adverse effects such as weight gain and seizures. In a small controlled study, adjunctive topiramate compared to adjunctive placebo in patients with schizophrenia yielded greater improvement in general psychopathological symptoms, but not positive or negative symptoms. However, adjunctive topiramate failed to attenuate schizophrenia symptoms in one small case series and may have even yielded deterioration in another. In a small controlled study, adjunctive topiramate was no better than adjunctive placebo in patients with schizoaffective disorder, bipolar type.

Posttraumatic Stress Disorder. Open, primarily adjunctive topiramate appeared effective in chronic civilian PTSD. A small controlled study suggested topiramate monotherapy was superior on some outcome measures in patients with civilian posttraumatic stress disorder, but another controlled study in male veterans with PTSD did not show benefit with topiramate.

Therapeutic Indications

Seizure Disorders. Topiramate is indicated as adjunctive treatment for adults and pediatric patients ages 2 to 16 years with partial-onset seizures or primary generalized tonic–clonic seizures, and in patients age 2 years or older with Lennox–Gastaut syndrome. Topiramate is also indicated as initial monotherapy in patients 2 years of age and older with partial-onset or primary generalized tonic–clonic seizures.

Migraine Prevention. Topiramate is indicated for adolescents (age ≥12 years) and adults for the prevention of migraine headaches.

Precautions and Adverse Reactions

The most common adverse effects with topiramate 400 mg per day monotherapy in epilepsy patients were paresthesia, weight decrease, somnolence, anorexia, dizziness, and memory problems in adults (21 percent adverse effect discontinuation rate); and weight decrease, upper respiratory tract infection, paresthesia, anorexia, diarrhea, and mood problems in children (12 percent adverse effect discontinuation rate). The most common adverse effects with adjunctive topiramate at 200 to 400 mg per day in adult epilepsy patients were somnolence, dizziness, ataxia, speech problems, psychomotor slowing, visual problems, memory problems, paresthesia, and diplopia (11 percent adverse effect discontinuation rate). The most common adverse effects with adjunctive

topiramate at 5 to 9 mg/kg per day in pediatric epilepsy patients were fatigue, somnolence, anorexia, nervousness, concentration or attention problems, memory problems, aggression, and weight loss (0 percent adverse effect discontinuation rate). The most common adverse effects with topiramate monotherapy in migraine prevention trials were paresthesia, fatigue, nausea, anorexia, upper respiratory tract infection, taste perversion, and dizziness (25 percent adverse effect discontinuation rate vs. 10 percent with placebo). In many cases, topiramate adverse effects are mild to moderate, and can be attenuated by dosage rescheduling or decrease.

Topiramate can cause appetite suppression and weight loss in epilepsy, psychiatric, and obesity patients. Although this is clearly a desirable effect in overweight or obese patients, it may be problematic in underweight patients. However, in epilepsy studies, weight loss with topiramate was more evident in patients with higher baseline weights. In these studies, topiramate-induced weight loss was more evident in women compared to men with similar baseline weights.

The topiramate USPI includes a warning of the risk of hyperchloremic, nonanion gap metabolic acidosis (low serum bicarbonate without chronic respiratory alkalosis), which can occur in a dose-related fashion with topiramate and yield hyperventilation, fatigue, anorexia, increased risk of nephrolithiasis, nephrocalcinosis, and osteomalacia; decreased growth rate; and rarely cardiac arrhythmias and stupor. Measurement of baseline and periodic serum bicarbonate is recommended, and if metabolic acidosis is detected, decreasing or discontinuing topiramate or treating with alkali should be considered. Topiramate has been rarely associated with a syndrome of acute myopia with secondary angle closure glaucoma, characterized by acute onset of decreased visual acuity or ocular pain. Symptoms can occur in pediatric as well as adult patients, typically present within 1 month of starting topiramate, and reverse with discontinuation. Other warnings in the topiramate USPI include the risks of oligohydrosis and hyperthermia, unexplained death in epilepsy, visual field defects, hyperammonemia and encephalopathy with or without concomitant valproic acid, hypothermia with and without hyperammonemia during concurrent treatment valproic acid, and seizures after abrupt discontinuation in patients with or without epilepsy. As with other anticonvulsants, topiramate has a warning of increased risk of suicidality (suicidal behavior or ideation) in patients with epilepsy as well as psychiatric disorders.

An important warning in the topiramate USPI concerns the risks of cognitive or neuropsychiatric adverse events. These are common and include cognitive problems (confusion, psychomotor slowing, and concentration or attention, memory, speech, or language problems), psychiatric problems (depression, mood problems), sedation, and fatigue.

Topiramate has also been associated with the onset of psychosis in case reports.

About 1.5 percent of adults exposed to topiramate develop nephrolithiasis. Maintaining adequate fluid intake may help decrease the risk of nephrolithiasis. Topiramate is a weak carbonic anhydrase inhibitor and hence decreases urinary citrate excretion and increases urinary pH, perhaps contributing to the increased risk of nephrolithiasis with this medicine. Carbonic anhydrase inhibition may also contribute to the increased risk of paresthesias with topiramate.

In patients with or without a history of epilepsy, topiramate should be gradually withdrawn to minimize the potential for seizures.

Acute oral overdoses of topiramate have been associated with convulsions, drowsiness, speech disturbance, blurred vision, diplopia, impaired mentation, lethargy, impaired coordination, stupor, hypotension, abdominal pain, agitation, dizziness, depression, and severe metabolic acidosis. In most instances, clinical consequences were not severe, but deaths have been reported in polydrug overdoses involving topiramate.

Topiramate was classified as pregnancy category C, as fetal loss and teratogenicity have been demonstrated in animals, and use during pregnancy may cause cleft lip and/or palate in humans. It is not known the degree to which topiramate is excreted in breast milk, so it was classified as "Safety Unknown" for lactation, and should be administered to women who are breastfeeding with caution only if the potential benefits clearly outweigh the (unknown) risks.

Topiramate is contraindicated in patients with known hypersensitivity to topiramate or its ingredients.

Drug Interactions

The large renal component of topiramate disposition yields it resistant to effects of enzyme inhibitors, yet the limited monotherapy hepatic metabolism component makes it susceptible to enzyme inducers. For example, in epilepsy patients, carbamazepine yields clinically significant decreases in topiramate serum concentrations, but topiramate does not yield clinically significant changes in carbamazepine serum concentrations. In contrast, valproate and lamotrigine do not appear to have clinically significant pharmacokinetic interactions with topiramate in patients with epilepsy. However, concomitant administration of topiramate and valproate has been associated with hyperammonemia with and without encephalopathy, and hypothermia (with and without hyperammonemia), most likely due to pharmacodynamic drug interactions. In view of its substantial renal clearance, topiramate could affect lithium levels. However, the clinical significance of such putative effects remains to be definitively established. Specifically, although topiramate at doses over 200 mg/day may affect serum lithium concentrations, the clinical importance of

this phenomenon remains to be adequately characterized. Nevertheless, the topiramate USPI recommends monitoring lithium concentrations when coadministered with high-dose topiramate. The effects of lithium on topiramate pharmacokinetics remain to be established. Although levetiracetam does not alter topiramate concentrations, the combination may yield adverse effects through a pharmacodynamic interaction. In addition, combination of topiramate with other carbonic anhydrase inhibitors, such as acetazolamide (Diamox) and zonisamide, is not recommended due to concerns about increasing the risks of metabolic acidosis, nephrolithiasis or heat-related problems (oligohydrosis and hyperthermia). Also, should topiramate cause metabolic acidosis, use of metformin is contraindicated. Topiramate is a mild enzyme inducer and can decrease blood concentrations of hormonal contraceptives, potentially compromising their efficacy. This effect may be less problematic at doses of 200 mg per day or less. Because of the potential of topiramate to cause CNS depression, as well as other cognitive and/or neuropsychiatric adverse reactions, combination with alcohol and other CNS depressants should only be done with caution.

Dosage and Administration

The topiramate USPI recommends that topiramate monotherapy in adults and children 10 years of age or older with epilepsy start with 25 mg twice daily for the first week, increase to 50 mg twice daily for week 2, 75 mg twice daily for week 3, 100 mg twice daily for week 4, 150 mg twice daily for week 5, and 200 mg twice daily for week 6 onward. In adult (age 17 years or older) epilepsy patients, topiramate adjunctive therapy is recommended to start with 25 to 50 mg per day and be increased by 25 to 50 mg on a weekly basis, up to 400 mg per day in two divided doses. In pediatric (age 2 to 16 years) epilepsy patients, topiramate adjunctive therapy is recommended to start with 1 to 3 mg/kg per day at bedtime and increase by 1 to 3 mg/kg on a weekly basis, up to 5 to 9 mg/kg per day. Plasma topiramate concentrations in epilepsy patients commonly range between 5 and 20 μg/mL (15 to 60 μM/L), but to do not appear to be systematically related to antiseizure effects and thus are not routinely monitored. For migraine prevention, topiramate is recommended to start with 25 mg in the evening for the first week, increase to 25 mg twice daily for week 2, 25 mg in the morning plus 50 mg in the evening for week 3, and 50 mg twice daily for week 4 onward. Doses may need to be decreased by 50 percent in patients with moderate to severe renal impairment (creatinine clearance less than 70 mL/1.73 m^2 per minute), and doubled in patients receiving concurrent enzyme inducers such as carbamazepine and phenytoin. Topiramate doses may need to be decreased in patients with hepatic impairment.

Off-label topiramate is typically used adjunctively, starting with 25 mg at bedtime and increasing weekly by 25 mg, as necessary and tolerated.

Dosage increases of 25 mg per day every 2 to 4 weeks during initiation may sometimes improve tolerability. Unfortunately, tolerability may be a problem for some patients even with doses as low as 25 mg per day. Final doses in efforts to promote weight loss are often between 75 and 200 mg per day at bedtime. Lower doses appear to less consistently suppress appetite, and higher doses, at least early in therapy, tend to not be tolerated due to sedation or cognitive difficulties. However, continuing to gradually increase topiramate 25 mg per day on a monthly basis may allow some patients to eventually achieve doses of 400 mg per day. In some (nonepilepsy) applications, clinicians may administer most or all of the dosage toward bedtime to enhance sleep and decrease daytime somnolence.

TIAGABINE

Introduction and History

Tiagabine (Gabitril), a lipophilic derivative of nipecotic acid, was developed by Novo Nordisk. The addition of a lipophilic group to nipecotic acid (a hydrophilic GABA reuptake inhibitor that does not cross the blood–brain barrier) allowed tiagabine to more easily enter the central nervous system. Tiagabine was effective in several animal models of epilepsy. Abbott Laboratories developed it as an antiepileptic drug, initially marketed for the adjunctive treatment of partial seizures in the United States in 1997, and later marketed by Cephalon, Inc.

Limited, postmarketing, open, uncontrolled clinical experience suggested efficacy in some psychiatric conditions. Importantly, safety concerns were raised paradoxically for this antiepileptic drug concerning the risk of seizures in patients with bipolar disorder, particularly if the drug was dosed aggressively. Limited double-blind, placebo-controlled studies revealed that tiagabine may be useful for generalized anxiety disorder and primary insomnia. Safety concerns, along with a lack of controlled efficacy data, have limited the use of tiagabine in disorders other than epilepsy.

Chemistry

Tiagabine, (-)-(R)-1-[4,4-Bis(3-methyl-2-thienyl)-3-butenyl] nipecotic acid, is a lipophilic derivative of the hydrophilic cyclic amino acid nipecotic acid (comprising nipecotic acid linked to a lipophilic group by an aliphatic chain), which inhibits GABA reuptake in neurons and glial cells (Fig. 34.7–1). Tiagabine is available in 2-, 4-, 12-, and 16-mg tablets for oral administration.

Pharmacological Actions

Pharmacokinetics. Tiagabine is well absorbed with a bioavailability of about 90 percent, extensively (96 percent) bound to plasma proteins, and

has linear kinetics at doses between 2 and 24 mg per day. Giving tiagabine with meals decreases rate of absorption and delays and lowers peak plasma concentrations, but does not significantly affect extent of absorption. In monotherapy, tiagabine has a half-life of about 8 hours and clearance of about 110 mL per minute. In the presence of enzyme inducers (as in the pivotal epilepsy studies), the half-life decreased to about 4 hours and clearance doubled to about 220 mL per minute. Tiagabine is a CYP3A substrate and is extensively transformed into inactive 5-oxo-tiagabine and glucuronide metabolites, with only 2 percent being excreted unchanged in the urine. The remainder is excreted as metabolites in the feces (65 percent) and the urine (25 percent).

Pharmacodynamics. Tiagabine has a putative specific single major biochemical action (as opposed to multiple diverse actions with most other anticonvulsants) that may account for its clinical effects. Tiagabine acts on the GABA transporter, blocking reuptake of the inhibitory amino acid neurotransmitter GABA into neurons and glia, and thus enhancing the inhibitory action of GABA. Actions at both $GABA_A$ and $GABA_B$ receptors putatively yield anticonvulsant and antinociceptive effects, respectively. Tiagabine has antiseizure activity in the pentylenetetrazole mouse, amygdala kindled rat, and other animal models of seizures, but only has weak activity in the maximal electroshock mouse model. Tiagabine has antinociceptive activity in animal models of neuropathic pain and anxiolytic-like effects in the rat open field and elevated plus-maze models of anxiety.

Clinical Drug Studies

Seizure Disorders. Adjunctive tiagabine appeared more effective than adjunctive placebo in three multicenter, double-blind, placebo-controlled trials involving patients with refractory partial seizures who were taking at least one enzyme-inducing anticonvulsant, as well as in two small placebo-controlled crossover studies.

Pain Syndromes. Limited data suggest possible utility of tiagabine in bruxism and associated temporomandibular joint pain, painful tonic spasms in multiple sclerosis, chronic pain, and chronic painful neuropathy. Large controlled studies are needed to confirm these encouraging preliminary findings.

Mood Disorders. There are no controlled studies of tiagabine in bipolar disorders. Moreover, there is an apparent lack of compelling and consistent open data suggesting efficacy of tiagabine in bipolar disorders. Some experience with open low-dose tiagabine in bipolar disorders may have been encouraging, but other case series suggested significant

problems with both efficacy and tolerability. For example, rapid loading of open primarily adjunctive tiagabine, starting at 20 mg per day (five times the recommended starting dose) in eight acute mania inpatients not only was ineffective, but also yielded unacceptable adverse effects, with one patient having a seizure. This suggested not only inefficacy in mania, but also that loading or rapid dosage escalation should be avoided and caution exercised in using tiagabine, particularly in the initial titration phase. However, even more gradual initiation may be problematic. Thus, open primarily adjunctive tiagabine initiated at 4 mg per day and increased weekly as tolerated by 4 mg per day (mean dose 8.7 mg per day, mean duration 38 days) in outpatients with treatment-refractory bipolar disorder had limited efficacy and raised tolerability concerns, as 2 of 13 patients (15 percent) had seizures attributed as likely due to the medication. Thus, controlled studies are lacking and open studies have raised both efficacy and tolerability concerns regarding tiagabine in bipolar disorders.

Cocaine Use Disorder. Although preliminary studies were somewhat encouraging, in a larger controlled trial, tiagabine at 20 mg per day was no better than placebo as treatment for cocaine dependence.

Other Psychiatric Disorders. Limited controlled data suggested that low-dose (16 mg or less per day) tiagabine was generally well tolerated and yielded some benefit in primary insomnia. In three controlled trials tiagabine was no better than placebo for generalized anxiety disorder.

Therapeutic Indications

Seizure Disorders. Tiagabine is indicated as adjunctive treatment of partial seizures adults and adolescents 12 years of age and older with epilepsy.

Precautions and Adverse Reactions

In general, in patients with epilepsy who are also taking enzyme-inducing anticonvulsants, tiagabine adverse effects are mild to moderate, can be attenuated by dosage rescheduling or decreases, and do not typically require discontinuation. In epilepsy patients, the most common adverse effects with adjunctive tiagabine were dizziness or lightheadedness, asthenia or lack of energy, somnolence, nausea, nervousness or irritability, tremor, abdominal pain, and thinking abnormally or difficulty with concentration or attention.

Importantly, the tiagabine USPI dosing recommendations were based on epilepsy patients taking enzyme-inducing anticonvulsants, which lower tiagabine plasma levels by approximately 50 percent. Administering tiagabine without enzyme inducers will approximately double tiagabine

plasma levels compared to those attained in the published epilepsy trials, which provided guidance for dosing. Thus, tiagabine given without enzyme inducers needs to have doses approximately 50 percent of those in the USPI. If this adjustment is not made, patients are at increased risk for adverse effects, some of which are serious.

The tiagabine USPI includes a boldface warning regarding the risk of seizures in patients without epilepsy. As noted above, rapid loading of open primarily adjunctive tiagabine in patients with acute mania, starting at 20 mg per day, yielded unacceptable adverse effects, with one patient having a seizure, and even more gradual initiation in outpatients with bipolar disorder at 4 mg per day and increased weekly as tolerated by 4 mg per day (mean final dose 8.7 mg per day) raised tolerability concerns, with 2 of 13 patients having seizures.

Other warnings in the tiagabine USPI include the risks of withdrawal seizures, cognitive or neuropsychiatric problems (impaired concentration, speech or language problems, somnolence, and fatigue), status epilepticus, and sudden unexpected death in epilepsy. As with other anticonvulsants, tiagabine has a warning of increased risk of suicidality (suicidal behavior or ideation) in patients with epilepsy as well as psychiatric disorders.

Acute oral overdoses of tiagabine have been associated with seizures, status epilepticus, coma, ataxia, confusion, somnolence, drowsiness, impaired speech, agitation, lethargy, myoclonus, stupor, tremors, disorientation, vomiting, hostility, temporary paralysis, and respiratory depression. Deaths have been reported in polydrug overdoses involving tiagabine.

Tiagabine was classified as pregnancy category C, as fetal loss and teratogenicity have been demonstrated in animals, but to date not in humans. It is not known the degree to which tiagabine is excreted in breast milk, so it was classified as “Safety Unknown” for lactation, and should be administered to women who are breastfeeding with caution only if the potential benefits clearly outweigh the (unknown) risks.

Tiagabine is contraindicated in patients with known hypersensitivity to tiagabine or its ingredients.

Drug Interactions

Tiagabine neither induces nor inhibits hepatic enzymes, but it is a target for pharmacokinetic drug interactions. Thus, enzyme inducers such as carbamazepine can yield clinically significant decreases in serum tiagabine concentrations, possibly by induction of CYP3A. Valproate does not yield clinically significant changes in serum total tiagabine concentrations, but appears to displace tiagabine from protein-binding sites, yielding an increase in free tiagabine, and thus potentially higher tiagabine toxicity. Tiagabine does not yield clinically significant changes in the pharmacokinetics of carbamazepine, phenytoin, and valproate. The

presence or absence of pharmacokinetic interactions between tiagabine and lamotrigine and tiagabine and lithium remains to be established. Tiagabine does not yield clinically significant changes in the kinetics of hormonal contraceptives.

Dosage and Administration

Most importantly, tiagabine should not be rapidly loaded or rapidly initiated due to the risk of serious adverse effects, namely potential seizures. The tiagabine USPI recommends that in adults and adolescents 12 years of age or older with epilepsy who were also taking enzyme inducers, tiagabine be initiated at 4 mg per day and increased weekly by 4 mg per day during the first month, and then increased weekly by 4 to 8 mg per day for weeks 5 and 6, yielding 24 to 32 mg per day administered in two to four divided doses by week 6. In adults (but not adolescents), tiagabine doses may be further increased weekly by 4 to 8 mg per day to as high as 56 mg per day. Plasma tiagabine concentrations in epilepsy patients commonly range between approximately 20 and 100 ng/mL (50 and 250 nM/L), but do not appear to be systematically related to antiseizure effects, and thus are not routinely monitored.

In pivotal epilepsy trials, tiagabine was given with meals, which decreases the rate (later, lower peak plasma concentration) but not the extent of absorption. Thus, in order to limit the risk of adverse effects, the USPI recommends that tiagabine be given with meals.

If tiagabine is administered in the absence of enzyme-inducing anticonvulsants, the dosing needs to be even more conservative (50 percent lower doses) in order to limit the risk of adverse effects. Thus, it is prudent that tiagabine be initiated at no more than 2 mg per day and increased weekly by 2 mg per day during the first month, and then increased weekly by 2 to 4 mg per day for weeks 5 and 6, yielding 12 to 16 mg per day administered in two to four divided doses by week 6. In adults (but not adolescents), tiagabine doses may be further increased weekly by 2 to 4 mg per day to as high as 28 mg per day.

Patients with hepatic impairment compared to those without may require up to 60 percent lower doses of tiagabine.

LEVETIRACETAM

Introduction and History

Levetiracetam (Keppra), developed in efforts to create second-generation nootropic (memory enhancing) medications, proved to be a potent anticonvulsant in the audiogenic-susceptible mouse model of epilepsy. In 1999, UCB, Inc. began marketing levetiracetam for the adjunctive treatment of refractory partial seizures in the United States.

Postmarketing, open, uncontrolled clinical experience suggested

potential benefit in patients with depressive symptoms with seizure disorders, bipolar disorder, and premenstrual dysphoric disorder. On the other hand, there also were reports of levetiracetam-related depression and aggressive behavior, and placebo-controlled studies did not find efficacy of levetiracetam for behavioral or overall clinical improvement in children with autism. Controlled studies have been lacking to assess the effects of levetiracetam in patients with bipolar disorder and other psychiatric disorders.

Chemistry

Levetiracetam, (-)-(*S*)-α-ethyl-2-oxo-pyrrolidine acetamide, is the (*S*)-enantiomer of the ethyl analog of the nootropic agent piracetam (Fig. 34.7–1) and has a structure unrelated to other anticonvulsants. Levetiracetam preparations currently marketed in the United States include: 250-, 500-, 750-, and 1,000-mg immediate-release tablets, 500- and 750-mg extended-release tablets; a 100-mg/mL oral solution; and a 100-mg/mL injectable intravenous solution. Generic immediate-release and extended-release levetiracetam are available.

Pharmacological Actions

Pharmacokinetics. Levetiracetam has a favorable pharmacokinetic profile, with absorption unaffected by food, bioavailability greater than 95 percent, binding to plasma proteins less than 10 percent, moderate volume of distribution of about .6 L/kg, linear elimination pharmacokinetics at clinically relevant doses (between 500 and 5,000 mg per day), half-life of approximately 8 hours, and clearance of 40 mL per minute. Levetiracetam is 66 percent excreted unchanged in the urine and is 24 percent hydrolyzed (in blood rather than in liver) to an inactive metabolite. Dosage may need to be decreased in patients with impaired renal function. However, levetiracetam has a low potential for pharmacokinetic drug interactions, because only a minor metabolic component involves hydrolysis (rather than CYP [cytochrome P450] oxidation or glucuronidation), and levetiracetam does not yield clinically significant changes in hepatic oxidation or glucuronidation.

Pharmacodynamics. Like most other anticonvulsants, levetiracetam has multiple biochemical actions, and the relationships between these and clinical effects remain to be established. Levetiracetam has GABAergic and antikindling effects, reduces voltage-dependent potassium currents, and binds to the synaptic vesicle protein SV2A. Levetiracetam lacks antiseizure activity in the pentylenetetrazole and maximal electroshock acute seizure models, but is potently active in chronic seizure models such as kindled seizures and genetic animal models. Levetiracetam also has antinociceptive

activity in animal models of neuropathic pain, and anxiolytic-like effects in animal models of anxiety disorders such as the rat conflict and elevated plus-maze.

Clinical Drug Studies

Seizure Disorders. Adjunctive levetiracetam appeared more effective than adjunctive placebo in three multicenter, randomized, double-blind, placebo-controlled trials involving adults with refractory partial seizures with or without secondary generalization; in one multicenter, randomized, double-blind, placebo-controlled trial in children (age 4 to 16 years) with refractory partial seizures; in one multicenter, randomized, double-blind, placebo-controlled trial in patients 12 years of age and older with juvenile myoclonic epilepsy; and in one multicenter, randomized, double-blind, placebo-controlled trial in patients 6 years of age and older with idiopathic primary generalized tonic–clonic seizures.

Other Neurological Disorders. In controlled pilot studies, levetiracetam appeared superior to placebo for migraine prevention, but was no better than placebo for essential tremor and Tourette syndrome.

Psychiatric Disorders. Levetiracetam was effective in an animal model of mania, and a series of uncontrolled case reports and case series suggested that levetiracetam may have efficacy in patients with bipolar disorder. In a small open-label study, patients with premenstrual dysphoric disorder had some improvement with levetiracetam. However, in controlled pilot studies, levetiracetam was no better than placebo in patients with social anxiety disorder and autism. Furthermore, there are reports of levetiracetam-related depression and aggressive behavior. Adequately powered controlled studies are lacking to assess the effects of levetiracetam in psychiatric disorders.

Therapeutic Indications

Seizure Disorders. Levetiracetam is indicated as adjunctive treatment of (1) partial-onset seizures in adults and children 4 years of age and older with epilepsy; (2) myoclonic seizures in adults and adolescents 12 years of age and older with juvenile myoclonic epilepsy; and (3) primary generalized tonic–clonic seizures in adults and children 6 years of age and older with idiopathic generalized epilepsy.

Precautions and Adverse Reactions

Levetiracetam is generally well tolerated but can cause somnolence and fatigue, coordination problems, and mood, anxiety, or behavioral adverse effects. In adult epilepsy patients, somnolence was seen in 14.8 percent (vs.

8.4 percent with placebo), asthenia occurred in 14.7 percent (vs. 9.1 percent with placebo), and coordination problems were evident in 3.4 percent (vs. 1.6 percent with placebo). In adult epilepsy patients, psychosis occurred in 0.7 percent (vs. 0.2 percent with placebo), and other behavioral problems were seen in 13.3 percent (vs. 6.2 percent with placebo). Behavioral problems included aggression, agitation, anger, anxiety, apathy, depersonalization, depression, emotional lability, hostility, and irritability. Behavioral adverse effects occurred early in treatment (often in the first month of therapy), and led to discontinuation in 1.7 percent (vs. 0.2 percent with placebo). Suicide attempts occurred in 0.5 percent (vs. 0 percent with placebo).

In pediatric epilepsy patients, somnolence was seen in 22.8 percent (vs. 11.3 percent with placebo), and asthenia occurred in 8.9 percent (vs. 3.1 percent with placebo). In pediatric epilepsy patients, behavioral problems were seen in 37.6 percent (vs. 18.6 percent with placebo). Behavioral problems in pediatric epilepsy included agitation, anxiety, apathy, depersonalization, depression, emotional lability, hostility, hyperkinesia, nervousness, neurosis, and personality disorder, leading to discontinuation in 10.9 percent (vs. 6.2 percent with placebo).

Levetiracetam-related behavioral problems appeared more common in patients with epilepsy than in patients with cognitive or anxiety disorders.

As with other anticonvulsants, levetiracetam has a warning of increased risk of suicidality (suicidal behavior or ideation) in patients with epilepsy as well as psychiatric disorders. Acute oral overdoses of levetiracetam have been associated with somnolence, agitation, aggression, respiratory depression, and coma.

Levetiracetam was classified as pregnancy category C, as fetal loss and teratogenicity have been demonstrated in animals, but to date not in humans. Levetiracetam was classified as “Probably Safe” for lactation, though it is excreted in breast milk, so that levetiracetam should be administered to women who are breastfeeding only if the potential benefits clearly outweigh the (unknown) risks.

Levetiracetam is contraindicated in patients with known hypersensitivity to levetiracetam or its ingredients.

Drug Interactions

Levetiracetam’s high degree of renal clearance and hydrolytic metabolism yield a low potential for pharmacokinetic drug interactions. Hence, levetiracetam does not appear to have clinically significant pharmacokinetic interactions with valproate or lamotrigine in epilepsy patients. Levetiracetam does not significantly alter serum concentrations of carbamazepine or carbamazepine epoxide in epilepsy patients; however, levetiracetam plus carbamazepine may still yield adverse effects through pharmacodynamic interactions. Carbamazepine (Tegretol) may decrease

serum levetiracetam concentrations, but with variable reports of potential clinical significance. Oxcarbazepine (Trileptal) may also modestly decrease serum concentrations of levetiracetam. Levetiracetam does not appear to have clinically significant pharmacokinetic interactions with gabapentin or topiramate in epilepsy patients, although the levetiracetam plus topiramate combination may yield adverse effects through a pharmacodynamic interaction. Levetiracetam does not yield clinically significant alterations in serum concentrations of phenytoin (Dilantin), phenobarbital (Bellatal), or primidone (Mysoline) in epilepsy patients. Despite both being renally cleared, levetiracetam and lithium pharmacokinetic interactions remain to be established. Levetiracetam does not yield clinically significant changes in the metabolism of hormonal contraceptives.

Dosage and Administration

In adult epilepsy patients, adjunctive levetiracetam is commonly started at 500 to 1,000 mg per day and increased as necessary and tolerated every 2 weeks by 1,000 mg per day, up to a maximum 3,000 mg per day in two divided doses. In pediatric epilepsy patients, adjunctive levetiracetam is commonly started at 20 mg/kg per day and increased as necessary and tolerated every 2 weeks by 20 mg/kg per day, up to a maximum 60 mg/kg per day in two divided doses. Plasma levetiracetam concentrations in epilepsy patients commonly range between approximately 6 and 20 μg/mL (35 and 120 μM/L), but to do not appear to be systematically related to antiseizure effects and thus are not routinely monitored. In view of the substantive renal clearance, levetiracetam dosage needs to be adjusted in patients with impaired renal function, based on the degree of renal insufficiency.

ZONISAMIDE

Introduction and History

Zonisamide (Zonegran) was synthesized by Dainippon Pharmaceutical Company while screening arylsulfonamides for anticonvulsant effects, putatively related to carbonic anhydrase inhibition. Although zonisamide only weakly inhibited carbonic anhydrase, it proved effective in several animal models of epilepsy. Zonisamide was subsequently developed as an antiepileptic drug, and marketed for over a decade in Japan and South Korea prior to being marketed for the adjunctive treatment of partial seizures in the United States in 2000 by Elan Pharmaceuticals, and more recently by Eisai, Inc.

Limited, open, uncontrolled clinical experience and controlled trials suggest potential efficacy in some other neurological and psychiatric conditions. Large controlled trials are needed to assess these encouraging

preliminary findings.

Chemistry

Zonisamide, (3-[sulfamoylmethyl]-1,2-benzisoxazole), is a 1,2-benzisoxazole-3-methanesulfonamide, which is structurally different from other marked anticonvulsants (Fig. 34.7–1). Zonisamide is available in 25- and 100-mg capsules for oral administration. Generic zonisamide is available.

Pharmacological Actions

Pharmacokinetics. Zonisamide has generally favorable pharmacokinetic properties, with good (close to 100 percent) bioavailability, only 40 to 60 percent binding to plasma proteins (but extensive binding to erythrocytes), a clearance of 20 mL per minute, and a moderate volume of distribution of 1.5 L/kg. Zonisamide has linear pharmacokinetics between 200 and 400 mg per day, but has disproportionate serum concentrations at 800 mg, presumably related to saturable binding to carbonic anhydrase in erythrocytes. As monotherapy, zonisamide has a half-life of 60 hours, which decreases to 30 hours if administered with enzyme inducers. Zonisamide is reduced to 2-sulfamoylacetylphenol (SMAP), primarily by CYP3A4, but is also 30 percent excreted unchanged in the urine. Marked renal impairment (creatinine clearance less than 20 mL per minute) increases plasma zonisamide concentrations by 35 percent.

Pharmacodynamics. Like most other anticonvulsants, zonisamide has multiple biochemical actions, and the relationships between these and clinical effects remain to be established. Zonisamide blocks sodium channels, reduces T-type calcium currents, facilitates dopaminergic and serotonergic neurotransmission, and is a weak carbonic anhydrase inhibitor. Although zonisamide binds to the GABA/benzodiazepine receptor complex, it does not alter chloride currents and does not appear to have major GABAergic activity. Zonisamide has antiseizure activity in the maximal electroshock, pentylenetetrazole (against maximal but not minimal seizures), fully kindled seizure, and several other animal models of seizures and antinociceptive activity in animal models of neuropathic pain.

Clinical Drug Studies

Seizure Disorders. Adjunctive zonisamide appeared more effective than adjunctive placebo in three multicenter, double-blind, placebo-controlled trials in patients with refractory partial seizures, with and without secondary generalization.

Diabetic Neuropathy. Zonisamide was no better than placebo for diabetic neuropathy in a small pilot study.

Other Neurological Disorders. A multicenter, randomized, double-blind, placebo-controlled study suggested adjunctive zonisamide had efficacy in treating motor symptoms in patients with Parkinson disease. However, zonisamide was no better than placebo for essential tremor in a small pilot study.

Eating Disorders. In controlled trials, zonisamide was superior to placebo for obesity and for binge eating disorder associated with obesity.

Bipolar Disorder. Open, uncontrolled reports suggest that zonisamide may have some utility in bipolar disorder, including for treatment of weight gain. Controlled trials are needed to assess these preliminary findings.

Therapeutic Indications

Seizure Disorders. Zonisamide is indicated as adjunctive treatment of partial seizures in adults with epilepsy.

Precautions and Adverse Reactions

In epilepsy registration trials, adverse effects with adjunctive zonisamide were generally mild to moderate, commonly attenuated by dosage rescheduling or decreases, and did not typically require discontinuation (adverse effect discontinuation rate 12 vs. 6 percent with placebo). The most common adverse effects with adjunctive zonisamide in epilepsy trials were somnolence, anorexia, dizziness, headache, nausea, and agitation or irritability. The appetite suppression effect of zonisamide has been used to therapeutic advantage in patients with obesity and binge eating disorder.

In patient populations other than epilepsy, zonisamide tolerability may be more challenging, with reports of higher adverse effect discontinuation rates in patients with binge eating disorder with obesity (27 vs. 13 percent with placebo), diabetic neuropathy (38 vs. 0 percent with placebo), essential tremor (30 vs. 0 percent with placebo), and bipolar disorder (32 percent due to worsening mood symptoms with open treatment).

Zonisamide has been associated with nephrolithiasis in approximately 4 percent of patients (but only symptomatic in approximately 1 percent of patients), presumably due to carbonic anhydrase inhibition. Zonisamide increases serum creatinine and blood urea nitrogen by approximately 8 percent, but does not appear to alter glomerular filtration rate (GFR) or yield progressive renal damage. Zonisamide is not recommended in patients with renal failure (GFR less than 50 mL per minute) due to toxicity risks. Sudden unexplained death in epilepsy, status epilepticus, and rare

creatine phosphokinase elevation and pancreatitis have been reported.

The zonisamide USPI includes a warning regarding the risks of severe reactions (Stevens–Johnson syndrome, toxic epidermal necrolysis, fulminant hepatic necrosis, agranulocytosis, aplastic anemia, and other blood dyscrasias) seen with sulfonamides, so if signs of such reactions occur, zonisamide needs to be discontinued immediately. Rash was seen in 2.2 percent of patients taking zonisamide in pivotal trials and resulted in discontinuation in 1.4 percent of patients. Other warnings include the risks of metabolic acidosis and nephrolithiasis, oligohydrosis and hyperthermia in pediatric patients, withdrawal seizures, teratogenicity, cognitive or neuropsychiatric problems (depression, psychosis, psychomotor slowing, impaired concentration or speech, language, word finding), somnolence, fatigue, and rare serious hematological events (two cases of aplastic anemia and one case of agranulocytosis in the first 11 years of marketing in Japan, and no case of aplastic anemia but two cases of agranulocytosis in the United States, European, and Japanese development programs). As with other anticonvulsants, zonisamide has a warning of increased risk of suicidality (suicidal behavior or ideation) in patients with epilepsy as well as psychiatric disorders.

Experience with zonisamide overdose is limited. Acute oral overdoses of zonisamide have been associated with coma, bradycardia, hypotension, and respiratory depression.

Zonisamide was classified as pregnancy category C, as fetal loss and teratogenicity have been demonstrated in animals, but to date not in humans. It is not known the degree to which zonisamide is excreted in breast milk, so it was classified as “Safety Unknown” for lactation, and should be administered to women who are breastfeeding with caution only if the potential benefits clearly outweigh the (unknown) risks.

Zonisamide is contraindicated in patients with known hypersensitivity to zonisamide or its ingredients. In addition, zonisamide has a sulfonamide functional group, so is contraindicated in patients with sulfa drug allergies.

Drug Interactions

Zonisamide does not inhibit CYP450 isoenzymes, does not induce its own metabolism, and is a target rather than an instigator of drug interactions. Carbamazepine decreases serum zonisamide concentrations, while zonisamide does not appear to yield clinically significant changes in carbamazepine pharmacokinetics. Zonisamide and valproate and zonisamide and lamotrigine do have clinically significant pharmacokinetic interactions in patients with epilepsy. The presence or absence of pharmacokinetic interactions between zonisamide and lithium remains to be established. Phenytoin and phenobarbital decrease serum zonisamide concentrations. Combination of zonisamide with other carbonic anhydrase inhibitors such as acetazolamide and topiramate is not recommended due

to concerns about increasing the risk of nephrolithiasis or heat-related problems (oligohydrosis and hyperthermia). Zonisamide does not alter hormonal contraceptive pharmacokinetics.

Dosage and Administration

In patients with epilepsy, adjunctive zonisamide is started at 100 mg per day, and increased every 2 weeks (no faster due to a 60-hour half-life) as necessary and tolerated by 100 mg per day, with final doses commonly ranging between 300 and 600 mg per day in two divided doses. Adverse effects were more common at doses of 300 mg per day or more. Plasma zonisamide concentrations in epilepsy patients commonly range between approximately 10 and 38 μg/mL (45 and 180 μM/L) but do not appear to be systematically related to antiseizure effects and thus are not routinely monitored. In controlled studies, mean final zonisamide doses were 160 mg per day for essential tremor, 427 mg per day for obesity, 436 mg per day for binge eating disorder with obesity, and 540 mg per day for painful diabetic neuropathy. As zonisamide undergoes both hepatic metabolism and renal excretion unchanged, patients with hepatic or renal disease may require more conservative dosing and titration and more frequent monitoring.

PREGABALIN

Introduction and History

Pregabalin (Lyrica) was developed in a fashion similar to that of gabapentin, as a compound with structure similar to the inhibitory amino acid neurotransmitter GABA that could easily cross the blood–brain barrier. Although these agents lacked direct GABAergic effects, they were effective in animal models of pain and epilepsy (with pregabalin being more potent) and developed as analgesic and antiepileptic drugs. Pregabalin was marketed for the treatment of neuropathic pain associated with diabetic peripheral neuropathy in the United States in 2005 by the Parke–Davis division of Pfizer, Inc. Pregabalin subsequently received approval for the treatment of postherpetic neuralgia, adjunctive treatment of adults with partial-onset seizures, for the treatment of fibromyalgia, and for the treatment of neuropathic pain associated with spinal cord injury. Pregabalin also appeared effective in controlled trials for generalized anxiety disorder and social anxiety disorder, but as of late 2014, had not been approved for these indications.

Chemistry

Pregabalin, *S*-(+)-3-isobutylGABA, is a novel amino acid derived from GABA (Fig. 34.7–1), with a structure and mechanisms that overlap those of gabapentin and some structural similarity to other anticonvulsants that resemble GABA, such as valproate. Pregabalin is currently marketed in the

United States as an oral preparation in 25-, 50-, 75-, 100-, 150-, 200-, 225-, and 300-mg capsules, and as a 20-mg/mL oral solution.

Pharmacological Actions

Pharmacokinetics. Pregabalin appears to have generally favorable pharmacokinetic properties, similar to gabapentin. Pregabalin is not bound to plasma proteins, has a moderate volume of distribution of 0.5 L/kg, a half-life of 6 hours (requiring twice daily dosing), is greater than 95 percent excreted unchanged in urine, and has clearance of 80 mL per minute that varies with creatinine clearance. Pregabalin clearance is highly dependent on renal function, and dosage reduction may be necessary in patients with renal impairment. Pregabalin lacks hepatic metabolism, does not induce or inhibit hepatic metabolism, and lacks metabolic drug interactions. Unlike gabapentin, pregabalin has higher (greater than 90 percent) bioavailability and absorption that is proportional to dose, without the gastrointestinal absorption saturation of gabapentin. Food has no clinically relevant effect on pregabalin absorption.

Pharmacodynamics. Like most other anticonvulsants, pregabalin has multiple biochemical actions, and the relationships between these and clinical effects remain to be established. Pregabalin, like gabapentin, is a structural analog of GABA, but these agents appear to lack direct GABAergic actions. However, both of these agents have selective inhibitory effects on voltage-gated calcium channels containing the $\alpha_2\delta$ subunit, reducing depolarization-induced calcium influx at nerve terminals, resulting in decreased release of excitatory neurotransmitters such as glutamate and noradrenaline. In rat forebrain, these drugs acutely decreased cellular glutamate but did not alter GABA or glutamine concentrations.

Pregabalin has antiseizure activity that is more potent than gabapentin in the maximal electroshock, pentylenetetrazole, hippocampal kindled, and other animal models of seizures. Antinociceptive activity is more potent than gabapentin in animal models of neuropathic pain. Pregabalin has anxiolytic-like effects in the rat conflict and elevated plus-maze animal models of anxiety disorders.

Clinical Drug Studies

Seizure Disorders. Adjunctive pregabalin was more effective than adjunctive placebo in three multicenter, double-blind, placebo-controlled trials in adults with partial-onset seizures.

Postherpetic Neuralgia. Pregabalin was more effective than placebo in three multicenter, double-blind, placebo-controlled trials in adults with

postherpetic neuralgia. Interestingly, mood, sleep, and quality of life measures improved, in addition to improved pain problems.

Diabetic Peripheral Neuropathy. Pregabalin was more effective than placebo in three multicenter, double-blind, placebo-controlled trials in adults with diabetic peripheral neuropathy. Similar to postherpetic neuralgia, pain, mood, sleep, and quality of life measures all improved.

Neuropathic Pain Associated with Spinal Cord Injury. Pregabalin was more effective than placebo in two multicenter, double-blind, placebo-controlled trials in adults with neuropathic pain associated with spinal cord injury. Similar to postherpetic neuralgia, pain, mood, sleep, and quality of life measures all improved.

Fibromyalgia. Pregabalin was more effective than placebo for fibromyalgia in one 14-week multicenter, double-blind, placebo-controlled trial (with superior relief of pain, disturbed sleep, and fatigue), and in one 6-month randomized withdrawal study.

Other Pain Syndromes. Pregabalin was effective for central neuropathic pain associated with spinal cord injury, improving pain, sleep, anxiety, and overall patient status. Pregabalin appeared effective for pain associated with chronic pancreatitis, postoperative mastectomy, and neuropathic pain from severe burns. Pregabalin was not found to be more effective than placebo for pain associated with prostatitis and HIV.

Other Neurological Disorders. In a pilot, double-blind, controlled, randomized trial, pregabalin was superior to placebo in the treatment of essential tremor.

Anxiety Disorders. In eight randomized, double-blind, placebo-controlled trials, pregabalin was effective for adult and older adult patients with generalized anxiety disorder and social anxiety disorder in short- and longer-term treatment.

Therapeutic Indications

Diabetic Peripheral Neuropathy. Pregabalin is indicated for the treatment of diabetic peripheral neuropathy.

Postherpetic Neuralgia. Pregabalin is indicated for the treatment of postherpetic neuralgia.

Neuropathic Pain Associated with Spinal Cord Injury. Pregabalin is indicated for the treatment of neuropathic pain associated with spinal cord

injury.

Seizure Disorders. Pregabalin is indicated for the adjunctive treatment of adults with partial-onset seizures.

Fibromyalgia. Pregabalin is indicated for the treatment of fibromyalgia.

Precautions and Adverse Reactions

In general, pregabalin adverse effects are mild to moderate, can be attenuated by dosage rescheduling or decreases, and do not typically require discontinuation. In diabetic neuropathy patients, the most common adverse effects with pregabalin were dizziness, somnolence, peripheral edema, dry mouth, and asthenia. In postherpetic neuralgia patients, the most common adverse effects with pregabalin were dizziness, somnolence, peripheral edema, infection, and headache. In partial-onset seizure patients, the most common adverse effects with adjunctive pregabalin were dizziness, somnolence, ataxia, weight gain, blurred vision, diplopia, and accidental injury. In fibromyalgia patients, the most common adverse effects with pregabalin were dizziness, somnolence, headache, weight gain, blurred vision, and dry mouth. The pregabalin USPI includes warnings regarding the risks for several of the above adverse effects (dizziness, somnolence, peripheral edema, weight gain, and blurred vision), as well as angioedema, hypersensitivity, withdrawal symptoms, tumorigenic potential (hemangiosarcoma in mice), elevated creatine kinase, thrombocytopenia, and PR interval prolongation. As with other anticonvulsants, pregabalin has a warning of increased risk of suicidality (suicidal behavior or ideation) in patients with epilepsy as well as psychiatric disorders.

Pregabalin is a schedule V controlled substance, with abuse and dependence potential concerns, based on the potential to yield euphoria, and the potential for withdrawal reactions.

There are limited data regarding the clinical consequences of acute oral overdoses of pregabalin.

Pregabalin was classified as pregnancy category C, as fetal loss and teratogenicity have been demonstrated in animals, but to date not in humans. It is not known the degree to which pregabalin is excreted in breast milk, so it was classified as “Safety Unknown” for lactation, and should be administered to women who are breastfeeding with caution only if the potential benefits clearly outweigh the (unknown) risks.

Pregabalin is contraindicated in patients with known hypersensitivity to pregabalin or its ingredients.

Drug Interactions

Like gabapentin, in view of the absence of hepatic metabolism, pregabalin lacks metabolic drug interactions. Thus, in patients with epilepsy, pregabalin lacks clinically significant pharmacokinetic drug interactions with carbamazepine, valproate, and lamotrigine. The presence or absence of pregabalin pharmacokinetic interactions with lithium remains to be established. In patients with epilepsy, pregabalin lacks clinically significant pharmacokinetic drug interactions with gabapentin and topiramate. Although tiagabine does not affect pregabalin pharmacokinetics, the presence or absence of a pregabalin effect on tiagabine pharmacokinetics remains to be established. The presence or absence of pregabalin pharmacokinetic interactions with oxcarbazepine, levetiracetam, and zonisamide remains to be established. Pregabalin does not yield clinically significant changes in the kinetics of hormonal contraceptives.

Dosage and Administration

The pregabalin USPI recommends that therapy in diabetic neuropathy, postherpetic neuralgia, fibromyalgia, partial-onset seizures, and neuropathic pain associated with spinal cord injury should start at 150 mg per day in two to three (two for fibromyalgia and neuropathic pain associated with spinal cord injury) divided doses, and be increased within 1 week (based on efficacy and tolerability) to 300 mg per day in two or three (two for fibromyalgia and neuropathic pain associated with spinal cord injury) divided doses. Patients with insufficient efficacy may be further increased to 450 mg per day in divided doses for fibromyalgia, and 600 mg per day in divided doses in partial-onset seizures, postherpetic neuralgia, and neuropathic pain associated with spinal cord injury. Plasma pregabalin concentrations in epilepsy patients do not appear to be systematically related to antiseizure effects and thus are not routinely monitored. In controlled trials, pregabalin doses have ranged from 150 to 600 mg per day in generalized anxiety disorder and social anxiety disorder, though pregabalin lacks a specific indication for anxiety disorders. In patients with renal impairment, pregabalin doses may need to be decreased, depending on the degree of renal insufficiency.

NEWEST ANTICONVULSANTS

The seven newest anticonvulsants (rufinamide, lacosamide, vigabatrin, esogabine, clobazam, perampanel, and eslicarbazepine acetate) also have diverse efficacy and safety/tolerability profiles, but detailed review of these agents is beyond the scope of this chapter. Thus, a few anticonvulsants may yield efficacy for some psychiatric and pain disorders. Clinicians may also find adjunctive uses for anticonvulsants, such as for managing weight gain. Knowledge of these differences, combined with familiarity with the pharmacodynamics, dosing, pharmacokinetics, drug interactions, adverse

effects, and their management for these agents can help clinicians provide safer and more effective therapeutics when newer anticonvulsants are used in the management of not only epilepsy, but also psychiatric and pain disorders.

Table 34.7–1.
Evidence-Based Uses of Selected Anticonvulsants

Gabapentin	Levetira-cetam	Pregabalin	Tiagabine	Topiramate	Zonisamide
Seizure disorders	Seizure disorders	Seizure disorders	Seizure disorders	Seizure disorders	Seizure disorders
Postherpetic neuralgia		Postherpetic neuralgia		Migraine prevention	
Restless legs syndrome		Diabetic peripheral neuropathy			
		Neuropathy associated with spinal cord injury			
		Fibromyalgia			
Neuropathic pain	*Migraine prevention*	*Generalized anxiety disorder*[a]		*Obesity*	*Obesity*
Social anxiety disorder[a]		*Social anxiety disorder*[a]		*Binge eating disorder with obesity*[a]	*Binge eating disorder with obesity*[a]
Panic disorder[a]				*Alcohol use disorders*[a]	
Alcohol use disorders					
Other substance use disorders					

Adapted from U.S. Food and Drug Administration (FDA) indicated uses in plain font. *Off-label uses in italic font.*

[a]Potential psychiatric use.

References

Ahmed M, Hays R, Steven Poceta J, Jaros MJ, Kim R, Shang G. Effect of gabapentin enacarbil on individual items of the international restless legs study group rating scale and post-sleep questionnaire in adults with moderate-to-severe primary restless legs syndrome: pooled analysis of 3 randomized trials. *Clin Ther*. 2016;38(7):1726–1737.

*Bray GA, Hollander P, Klein S, Kushner R, Levy B. A 6-month randomized, placebo-controlled, dose-ranging trial of topiramate for weight loss in obesity. *Obes Res*. 2003;11(6):722.

Coyle H, Clough P, Cooper P, Mohanraj R. Clinical experience with perampanel: focus on psychiatric adverse effects. *Epilepsy Behav*. 2014;41:193–196.

Frampton JE. Pregabalin: a review of its use in adults with generalized anxiety disorder. *CNS Drugs*. 2014;28(9):835–854.

*Gadde KM, Kopping MF, Wagner HR 2nd, Yonish GM, Allison DB, Bray GA. Zonisamide for weight reduction in obese adults: a 1-year randomized controlled trial. *Arch Intern Med*. 2012;172(20):1557–1564.

Johnson BA, Rosenthal N, Capece JA, Wiegand F, Mao L. Topiramate for treating alcohol dependence: a randomized controlled trial. *JAMA*. 2007;298(14):1641.

Knapp CM, Ciraulo DA, Sarid-Segal O, et al. Zonisamide, topiramate, and levetiracetam: efficacy and neuropsychological effects in alcohol use disorders. *J Clin Psychopharmacol*. 2015;35(1):34–42.

Kushner SF, Khan A, Lane R, Olson WH. Topiramate monotherapy in the management of acute mania: results of four double-blind placebo-controlled trials. *Bipolar Disord*. 2006;8(1):15.

Lee JJ, Song HS, Hwang YH, Lee HW, Suh CK, Park SP. Psychiatric symptoms and quality of life in patients with drug-refractory epilepsy receiving adjunctive levetiracetam therapy. *J Clin Neurol*. 2011;7(3):128–136.

Mason BJ, Quello S, Goodell V, Shadan F, Kyle M, Begovic A. Gabapentin treatment for alcohol dependence: a randomized clinical trial. *JAMA Intern Med*. 2014;174(1):70–77.

*McElroy SL, Hudson JI, Capece JA, Beyers K, Fisher AC. Topiramate for the treatment of binge eating disorder associated with obesity: a placebo-controlled study. *Biol Psychiatry*. 2007;61(9):1039.

McElroy SL, Kotwal R, Guerdjikova AI, Welge JA, Nelson EB. Zonisamide in the treatment of binge eating disorder with obesity: a randomized controlled trial. *J Clin Psychiatry*. 2006;67(12):1897.

Molokwu OA, Ezeala-Adikaibe BA, Onwuekwe IO. Levetiracetam-induced rage and suicidality: two case reports and review of literature. *Epilepsy Behav Case Rep*. 2015;4:79–81.

Pollack MH, Tiller J, Xie F, Trivedi MH. Tiagabine in adult patients with generalized anxiety disorder: results from 3 randomized, double-blind, placebo-controlled, parallel-group studies. *J Clin Psychopharmacol*. 2008;28:308.

*Rickels K, Pollack MH, Feltner DE, Lydiard RB, Zimbroff DL. Pregabalin for treatment of generalized anxiety disorder: a 4-week, multicenter, double-blind, placebo-controlled trial of pregabalin and alprazolam. *Arch Gen Psychiatry*. 2005;62(9):1022.

Sadeghian H, Motiei-Langroudi R. Comparison of levetiracetam and sodium valproate in migraine prophylaxis: a randomized placebo-controlled study. *Ann Indian Acad Neurol*. 2015;18(1):45–48.

Shin JH, Gadde KM. Clinical utility of phentermine/topiramate (Qsymia) combination for the treatment of obesity. *Diabetes Metab Syndr Obes*. 2013;6:131–139.

Silberstein S, Lipton R, Dodick D, et al. Topiramate treatment of chronic migraine: a randomized, placebo-controlled trial of quality of life and other efficacy measures. *Headache*. 2009;49(8):1153–1162.

Sun Y, van Valkenhoef G, Morel T. A mixed treatment comparison of gabapentin enacarbil, pramipexole, ropinirole and rotigotine in moderate-to-severe restless legs syndrome. *Curr Med Res Opin*. 2014;30(11):2267–2278.

Vieta E, Manuel Goikolea J, Martinez-Aran A, Comes M, Verger K. A double-blind, randomized, placebo-controlled, prophylaxis study of adjunctive gabapentin for bipolar disorder. *J Clin Psychiatry*. 2006;67(3):473.

Wang PW, Yang YS, Chandler RA. Adjunctive zonisamide for weight loss in euthymic bipolar disorder patients: a pilot study. *J Psychiatr Res*. 2008;42:451.

▲ 34.8 Antihistamines

ROGER S. MCINTYRE, M.D.

HISTORY

Histamine is a phylogenetically old amine widely distributed throughout the CNS across many species. In the early 1900s, histamine's role in immunomodulation and smooth muscle relaxation was originally described. Charles Best and colleagues were the first to isolate histamine (H-substance) in 1927. The antihistamines were first discovered by the Swiss-born Italian pharmacologist Daniel Bovet in 1937. The first nontoxic histamine-blocking drug, pyrilamine (Poly-Histine-D), was discovered in 1944. In the 1950s, diphenhydramine (Benadryl) and tripelennamine (Pyribenzamine) were introduced as highly effective histamine antagonists. The sedative and anticholinergic activity of antihistamines provided the basis for their frequent prescription for a panoply of psychiatric symptoms. The more recently introduced nonsedating antihistamines are less frequently prescribed in psychiatric practice.

CHEMISTRY

With approximately 50 H_1 antihistamines currently available worldwide, antihistamines are among the most widely utilized class of agents in the world (Table 34.8–1; Fig. 34.8–1). H_1 antihistamines have been reclassified as inverse agonists rather than competitive antagonists. Antihistamines are classified into six chemical classes: Phenothiazines (e.g., promethazine [Phenergan]), piperidines (e.g., cyproheptadine [Periactin]), ethanolamines (e.g., diphenhydramine), piperazines (e.g., hydroxyzine HCI [Atarax]), ethylenediamines (e.g., pyrilamine), and alkylamines (e.g., chlorpheniramine [Chlor-Trimeton). Although the antihistamines are structurally unrelated, they share a common ethylamine moiety. The typology, first- and second-generation antihistamines, refers to sedating and nonsedating agents, respectively. More recently, nonsedating (i.e., they are substrates of P-glycoprotein) antihistamines with less arrhythmogenic potential (e.g., fexofenadine [Allegra], loratadine [Claritin], desloratadine [Clarinex], and cetirizine [Zyrtec]) have been referred to as third-generation antihistamines.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

Some general properties of important histamine antagonists are shown in Table 34.8–1.

Absorption

The first-generation H_1 antagonists are readily absorbed via the GI tract. Most agents achieve maximum plasma levels within 1 to 3 hours (T_{max}).

Distribution

Antihistamine binding of plasma protein varies from approximately 50 to 98 percent. The H_1 antihistamine's volume of distribution varies more than 100-fold from 0.33 L/kg (i.e., levocetirizine [Xyzal]) to greater than 100 L/kg (i.e., loratadine). Sedating antihistamines are widely distributed in the body and CNS. The antihistamine CNS adverse event profile is a consequence of its relative ability to traverse the BBB. Factors that influence BBB permeability are lipophilicity, molecular weight, and recognition by the P-glycoprotein efflux pump expressed on luminal surfaces of nonfenestrated endothelial cells in the vascular wall of the CNS. Radioligand studies utilizing carbon-11 doxepin [Sinequan] indicate that first-generation H_1 antihistamines occupy approximately 50 to 90 percent of the H_1 receptors (H_1R) in the frontal and temporal cortices, hippocampus, and pons. Second-generation H_1 antihistamines occupy approximately 0 to 30 percent of CNS H_1R_5 receptors largely due to their minimal lipophilicity and affinity for P-glycoproteins.

Table 34.8–1.
Pharmacokinetics and Pharmacodynamics of Oral H_2 Antihistamines in Healthy Young Adults[a]

H_2 Antihistamine (Metabolic)	T_{max} after a Single Dose*b* (hours)	Terminal Elimination Half-Life	Elimination Unchanged in Urine/Feces (%)	Clinically Relevant Drug–Drug Interaction	Onset, Duration of Action*c* (hours)	Usual Adult Dose	Co Th Re Ac
First-generation drug[d]							
Chlorpheniramine	2.8 ± 0.8	27.9 ± 8.7	—	Possible	3, 24	4 mg three or four times daily 12 mg (sustained-release formulation) twice daily	
Diphenhydramine	1.7 ± 1.0	9.2 ± 2.5	—	Possible	2, 12	25–50 mg three or four times daily or at bedtime	He
Doxepin	2	13	—	Possible	—	25–50 mg three times	He

						daily or at bedtime	
Hydroxyzine	2.1 ± 0.4	20.0 ± 4.1	—	Possible	2, 24	25–50 mg three times daily or at bedtime	He
Second-generation drug							
Acrivastine	1.4 ± 0.4	1.4–3.1	59/0	Unlikely	1, 8	8 mg three times daily	—
Cetirizine	1.0 ± 0.5	6.5–10	60/10	Unlikely	1, 24	5–10 mg daily	Re
Desloratadine	1–3	27	0	Unlikely	2, 24	5 mg daily	Re
Ebastine*e* (carebastine)	(2.6–5.7)	(10.3–19.3)	(75–95/0)	—	2, 24	10–20 mg daily	Re
Fexofenadine*f*	2.6	14.4	12/80	Unlikely	2, 24	60 mg twice daily or 120 or 180 mg daily	Re
Levocetirizine*e*	0.8 ± 0.5	7 ± 1.5	86/13	Unlikely	1, 24	5 mg daily	Re
Loratadine (descarboethoxy-loratadine)	1.2 ± 0.3 (1.5 ± 0.7)	7.8 ± 4.2 (24 ± 9.8)	Trace	Unlikely	2, 24	10 mg daily	He
Mizolastine*e*	1.5	12.9	0.5/0	—	1, 24	10 mg daily	—

[a]Unless otherwise indicated, results are expressed as means ± SD or as a range of values. Parentheses denote active metabolites.

[b]T_{max} denotes time from oral intake to peak plasma concentration.

[c]Onset and duration of action are based on wheal-and-flare studies.

[d]Studies were performed decades after some of these drugs were introduced. Empirical dosage regimens persist.

[e]Second-generation H_2 antihistamines were not available in the United States at the time of publication.

[f]Fexofenadine should not be administered within 15 minutes after the ingestion of antacids that contain aluminum and magnesium as it may decrease absorption.

METABOLISM

Most antihistamines are subject to phase I metabolism via hepatic cytochrome P450 systems. In vitro studies evaluating human liver microsomes transfected with CYP2D6 complementary deoxyribonucleic acid (cDNA) indicate that several antihistamines (e.g., diphenhydramine, chlorpheniramine, clemastine [Dayhist-1], perphenazine [Trilafon], hydroxyzine, tripelennamine) inhibit the metabolism and are hepatically biotransformed by CYP2D6 and CYP3A4. Second- and third-generation

antihistamines are neither potent inhibitors nor inducers of the cytochrome P450 system. The H_1-antihistamine terminal elimination half-life ($t_{1/2}\beta$) ranges from approximately 2 to 27 hours. Pediatric populations rapidly metabolize antihistamines; individuals who are slow metabolizers may exhibit drug accumulation.

ELIMINATION

Most antihistamines are excreted largely unchanged in the urine and feces.

PHARMACODYNAMICS

Histamine is a pleiotropic, low–molecular-weight amine synthesized endogenously in diverse cell types from L-histidine by the enzyme histidine decarboxylase. Histamine is produced in the CNS in neurons of the tuberomammillary nucleus of the posterior hypothalamus with projections to both cortical and subcortical regions. Although no precise topographical organization of the histaminergic projection field has been elucidated, the cerebral cortex, amygdala, substantia nigra, and striatum receive moderate to dense histaminergic innervation. Innervations to the hippocampus, thalamus, retina, and spinal cord are variable. Histaminergic networks interact and modulate classical monoamine networks. Afferent projection to tuberomammillary neurons includes infralimbic cortex, lateral septum, preoptic nucleus, and brainstem (from adrenergic cell group C1 to C3, noradrenergic group A1 to A3, and serotonergic group B5 to B9).

In addition to histamine, the tuberomammillary neurons also contain GABA–synthesizing enzyme, glutamic acid decarboxylase (GAD), galanin, thyrotropin-releasing hormone, proenkephalin-derived peptides, substance P, and adenosine deaminase. The CNS histamine pool is also comprised of mast cells, glial cells, and vascular endothelium, which are activated by inflammatory effectors.

H_1 and H_2 receptors exert mostly excitatory actions on neurons, while H_3-receptor activation results in autoinhibition of tuberomammillary neurons and reduced neurotransmitter release. H_1-histamine receptors participate in synaptic plasticity and long-term potentiation (LTP) via their effect on intracellular Ca^{2+} and protein kinase C (PKC). H_2-receptor activation potently stimulates the cAMP pathway; a critical substrate for late-phase NMDA receptor-dependent LTP. H_2-receptor activation also participates in the induction of mossy-fiber LTP. High histamine concentration in the CA1 region of the hippocampus facilitates the induction of LTP in the presence of H_1 and H_2 antagonists, suggesting direct action at the NMDA receptors.

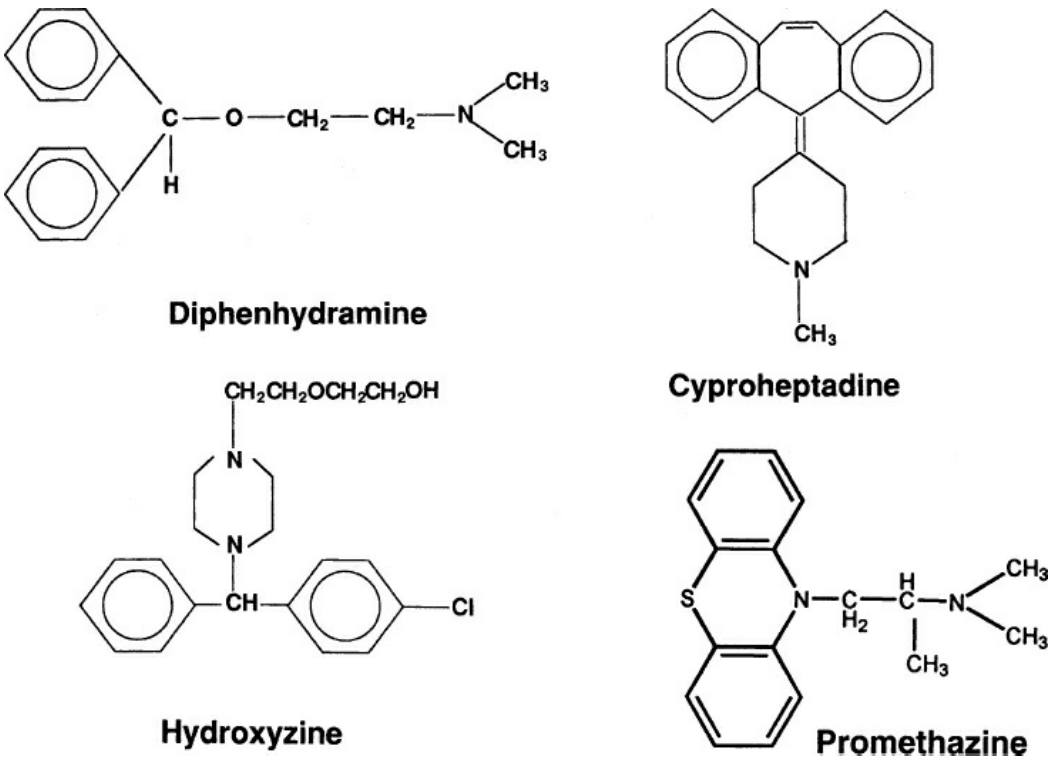

FIGURE 34.8–1. Molecular structures of antihistamines used in psychiatric practice.

Histamine plays a critical role in CNS arousal. Lesioning of the posterior hypothalamus results in hypersomnia and decreased psychomotor activity. Histamine also regulates cortical activity by stimulating cholinergic neurons of the basal forebrain and serotonergic neurons of the dorsal raphe nucleus. The wakefulness-promoting effects of orexins may be mediated by increased histamine release as inactivation mutations of the H_1R decrease the orexin-mediated arousal effect. Histamine is also involved in vasopressor effects, temperature regulation, and analgesia.

Histamine regulates energy homeostasis via effects on leptin networks. Central administration of leptin increases histamine turnover in the hypothalamus. Leptin-induced suppression of food intake and upregulation of uncoupling protein 1 expression in brown adipose tissue is attenuated in H_1R knockout mice. H_1R as well as POMC-MC4-R regulate energy homeostasis in the brain. Histamine's effect on drinking behavior and osmotic homeostasis is mediated by its stimulatory effect on vasopressin-producing neurons in the supraoptic nucleus. Histamine also regulates neuroendocrine systems (e.g., prolactin, growth hormone, adrenocorticotropin, and thyrotropin-releasing hormone).

Unlike other aminergic systems, histamine and the level of its metabolites increase in the CSF as a function of age. In Alzheimer disease, neurofibrillary tangles occur prominently in the tuberomammillary nucleus, with a corresponding decrease in the number of tuberomammillary neurons. Histamine levels and histidine decarboxylase

activity are diminished in brain regions affected by Alzheimer disease and Down syndrome. In Parkinson patients, histamine levels are selectively increased in the putamen, substantia nigra, and external globus pallidus. Levels of tele-methylhistamine (t-MeHA), the primary metabolite of histamine, are unchanged in substantia nigra. In contradistinction, t-MeHA is elevated in the CSF of individuals with schizophrenia, particularly those with prominent negative symptoms. These observations indicate that increased histamine release and turnover are pathophysiological features in schizophrenia, providing the basis for evaluating the efficacy of H_2-receptor antagonists. Interestingly, results from relatively small placebo-controlled clinical trials in schizophrenia indicate that adjunctive famotidine was superior to placebo in reducing overall objectively measured negative symptom severity. In addition, global psychopathology in famotidine-treated subjects may exhibit significantly greater reduction when compared to placebo treatment.

During the past two decades, the human H_1-, H_2-, H_3-, and H_4-histamine receptors have been cloned and characterized, each differing in structure, function, localization, and pharmacology. Immunocytochemical studies have isolated H_3 receptors throughout the CNS. Conformationally, the histamine receptor is a heptahelical protein that exists in both active and inactive forms in dynamic equilibrium. Several isoforms for each of the histamine receptors have been identified with reverse transcription polymerase chain reaction. The four major histamine receptors transduce extracellular signals via G proteins to intracellular second-messenger systems. Histamine receptors are constitutively active and capable of independent signaling from agonists both in vitro and in vivo.

The human H_1R have approximately 45 percent homology with muscarinic receptors. Inactivation mutations of the H_1-receptor gene result in abnormal learning, memory, immunological abnormalities, and locomotor activity. The H_1R also mediate the increased permeability of small vessels seen in allergic responses. Extravascular smooth muscle (e.g., bronchiole smooth muscle) contraction is also mediated by H_1R. H_3 receptors are presynaptic autoreceptors on histamine-containing neurons; they mediate the synthesis, release, and inhibition of histamine. H_3 receptors are also involved in modulating the release of several other monoamines and neuropeptides.

DESIGN AND INTERPRETATION OF CLINICAL DRUG STUDIES

A substantial body of evidence has established the efficacy of H_1 antihistamines for allergy-related conditions as well as motion sickness. There are no contemporary studies that have established the efficacy of H_1 histamine antagonist for any *Diagnostic and Statistical Manual of Mental*

Disorders (DSM-5)–defined mental disorders. H_1 antagonists are frequently prescribed to achieve ataraxia and to reduce psychomotor agitation and sleep latency. The anticholinergic effects of H_1 antihistamines have provided the basis for prescribing these agents in Parkinson disease and for antipsychotic-induced movement disorders. Equivocal evidence supports the use of H_2-antihistamine receptor antagonists in the treatment of schizophrenia, Alzheimer disease, medication-associated weight loss, anxiety disorders, and autism.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

The effects of H_1 antihistamines are a direct consequence of their pharmacological activity at their canonical receptors. First-generation H_1 antihistamines result in sedation, somnolence, and decreased reaction time. In some cases, paradoxical excitation characterized by restlessness, agitation, and nervousness has also been reported. The stimulating effects of H_1 antihistamines have been particularly apparent in tasks that require a high level of attention (e.g., divided attention tasks, vigilance tasks, and driving tasks). In some cases, the stimulating effects of antihistamines are dose-dependent. The precise biochemical mechanism subserving the excitatory effects has not been elucidated; it is hypothesized that neurotransmitters such as dopamine and GABA are salient or that the effect is mediated by the H_3-histamine receptor.

Antihistamines also have prominent autonomic nervous system effects. They exert a drying effect on mucous membranes in the respiratory system, and their effects on smooth muscles result in bronchodilation and the inhibition of vasoconstriction. H_1 antihistamines do not significantly affect blood pressure; parenteral administration of antihistamines results in transient lowering of blood pressure.

THERAPEUTIC INDICATIONS

Histamine's critical role in immunomodulation provides the basis for its clinical application in allergic rhinoconjunctivitis, asthma, urticaria, and other allergic or immunological disorders. H_1 antihistamines are widely prescribed in the treatment of insomnia, analgesia, akathisia, serotonin syndrome, anxiety, motion sickness, vertigo, antiemetic, and for bruxism. For example, preliminary evidence suggests that hydroxyzine may be superior to placebo at reducing overall severity of bruxism in children. They are often included in many over-the-counter preparations. Benzodiazepines and other classes of psychotropic agents that reduce anxiety and improve sleep efficiency have largely replaced antihistamines in clinical practice.

The rationale for considering antihistamines is supported by evidence

that the histaminergic systems are critical to wakefulness promotion. Doxepin, a TCA, is FDA approved for the treatment of insomnia. Doxepin's principle pharmacodynamic mode of action in insomnia relates to its potent antihistaminergic effects. Doxepin (and diphenhydramine) have been reported to induce non-REM sleep by blockade of H_1R. A highly replicated body of evidence indicates that low-dose doxepin's beneficial effect on sleep maintenance and duration has a medium effect size in comparison to placebo.

Diphenhydramine

Oral diphenhydramine is efficacious in the treatment of antipsychotic-induced movement disorders. Diphenhydramine is indicated by the FDA for the treatment of antipsychotic-induced parkinsonism, allergic and anaphylactic reactions, as well as motion sickness. When prescribed for parkinsonism, diphenhydramine is often combined with other anticholinergic agents. The parenteral availability of diphenhydramine provides for it a viable treatment for Parkinson patients with difficulty swallowing and in the treatment of acute dystonic reactions. Diphenhydramine has historically been prescribed as a nonspecific sedative for children and the elderly. Reports of paradoxical excitation as well as cognitive impairment (e.g., anticholinergic effect) reduce its therapeutic index.

Hydroxyzine

Hydroxyzine is indicated for the short-term treatment of anxiety and tension due to psychological factors. Although hydroxyzine's efficacy in the treatment of generalized anxiety is supported with comparator- and placebo-controlled trials, there is a paucity of rigorous, controlled studies in DSM-IV-TR–defined generalized anxiety disorder. Hydroxyzine is also indicated for use as an antipruritic and as a sedative as part of a premedication regimen in conjunction with general anesthesia.

Promethazine

Promethazine is indicated before or after surgery and as an antiemetic in postoperative patients. Like diphenhydramine, it is also used for allergic reactions and to treat motion sickness. Promethazine is indicated for sedation in both pediatric and adult populations. It is also used as an adjuvant for analgesia.

Cyproheptadine

Cyproheptadine is indicated for allergic reactions. Cyproheptadine is used for disparate purposes in psychiatric practice (e.g., to reverse antidepressant-associated sexual dysfunction or stimulant-induced weight

loss or insomnia), and for the symptomatic treatment of schizophrenia, autistic disorders, and serotonin syndrome.

H_2-Receptor Antagonists

H_2-receptor antagonist use in psychiatric practice has primarily been for weight loss in overweight individuals and persons experiencing antipsychotic-associated weight gain. Initial studies indicated that the H_2-receptor antagonist nizatidine (Axid) was efficacious as a secondary prevention strategy for olanzapine (Zyprexa)-associated weight gain. Subsequent placebo-controlled studies with this agent as a primary or secondary prevention strategy for olanzapine and quetiapine (Seroquel)-associated weight gain have produced more equivocal results. The mechanism mediating nizatidine's modest weight loss effect is hypothesized to be via central appetite–regulating centers. Pharmacological managements related to weight gain from psychotropic agents are likely best treated with other better proven strategies (e.g., metformin).

Ecological studies suggested that H_2-receptor antagonists could possibly decrease incident Alzheimer disease. Subsequent randomized, double-blind, placebo-controlled trials with nizatidine, however, failed to demonstrate a therapeutic effect in the treatment of cognitive deterioration as part of Alzheimer disease. Notwithstanding, several lines of preclinical evidence suggest that antihistaminergic therapy may facilitate reduction in amyloid-beta neuropathology and prevention of autophagic failure.

PRECAUTIONS AND ADVERSE EVENTS

The most common side effect of first-generation H_1-antagonist receptors is CNS depression, manifesting as sedation, somnolence, cognitive impairment, and reduction in psychomotor activity. Simple and complex tests of psychomotor performance indicate that second-generation antihistamines are less impairing than first-generation antihistamines. Although second- and third-generation antihistamines are less sedating, when coadministered with other agents that alter P-glycoprotein function, their propensity for CNS adverse events is increased.

The coadministration of alcohol, benzodiazepines, or other CNS depressants may result in additive sedation, requiring dose adjustment. Other reported CNS side effects include cognitive impairment, blurred vision, nervousness, and tremors. Other non-CNS side effects associated with antihistamines are dry mouth, increased appetite, nausea, vomiting, diarrhea, or constipation. H_1-receptor antagonists also have anticholinergic side effects such as dry mouth and blurred vision and should be used with caution for individuals with bladder neck obstruction or enlarged prostate.

Use of antihistamines in the elderly could result in confusion and

disorientation. Pediatric populations and individuals overdosing with H_1-receptor antagonists are at risk for excitatory effects. Excitatory symptoms include ataxia, agitation, incoordination, and seizures. Other features seen in overdose resemble atropine poisoning and include flushed face, dilated pupils, tachycardia, and fever. Treatment of overdose is generally through supportive medical therapy.

Of particular concern with H_1 antihistamines is the possibility of cardiotoxicity, particularly in overdose or when coprescribed with other agents. The arrhythmogenic potential of several second-generation H_1 antihistamines is mediated by nonhistaminergic receptors. Reports in the late 1980s of cardiac arrhythmia associated with intentional overdose and therapeutic use of terfenadine (Seldane) drew attention to the hazards of drug–drug interactions via P450 (CYP) oxidative enzymes. The first-generation H_1 antihistamines exhibit both muscarinic and α-adrenergic blocking effects as well as dose-dependent prolongation of QT intervals.

Originally introduced second-generation H_1 antihistamines astemizole (Hismanal) and terfenadine are known to block the rapid component of delayed rectifier potassium current (I_{Kr}) with 50 percent inhibitory in nanomolar range: This results in prolongation of the monophasic cardiac action potential and QT interval, engendering early depolarization and dispersion increasing risk for torsades de pointes. Terfenadine is similar in potency to quinidine at inhibiting the delayed rectifier potassium channel in cardiac tissue. Subsequently introduced second-generation antihistamines exhibit less potency in blocking the I_{Kr}, bestowing these agents with less cardiac toxicity potential.

Unintentional or intentional overdose of antihistamine can be fatal. The abuse liability of H_1 antihistamines is generally low, although some first-generation agents may be used for recreational purposes (e.g., dimenhydrinate [Dramamine]). Some first-generation H_1 antihistamines have been associated with sudden infant death syndrome. Other adverse events include fixed-drug eruption, photosensitivity, fever, hepatotransaminase elevation, and agranulocytosis. Some first- and second-generation H_1 antihistamines are FDA category B, while others are category C, indicating they should generally be avoided if possible.

DRUG INTERACTIONS

Pharmacodynamic drug interactions are not uncommon and usually have additive CNS depressant and anticholinergic effects. Many over-the-counter agents contain antihistamines. Coprescribing antihistamines with MAOIs may enhance anticholinergic effects.

LABORATORY INTERFERENCES

Promethazine has been reported to interfere with pregnancy tests and increase glucose levels. Hydroxyzine may falsely elevate urinary 17-hydroxycorticosteroid levels. All available H_1-receptor antagonists may interfere with allergic skin tests, resulting in a false-negative test.

DOSAGE ADMINISTRATION

Table 34.8–2 outlines the dosing of commonly prescribed first-generation H_1-receptor antagonists. Superficial administration of intramuscular formulations may result in local irritation, hence deep injection is recommended. Most antihistamines are prescribed in divided doses throughout the day; when used as a hypnotic, a single evening dose is recommended.

Table 34.8–2.
Dosage and Administration of Common Histamine Antagonists

Medication	Route	Preparation	Common Dosage
Diphenhydramine (Benadryl)	Oral	Capsules and tablets: 25 mg, 50 mg	Adults: 25–50 mg three to four times per day
		Liquid: 12.5 mg/5.0 mL	Children: 5 mg/kg three to four times per day, not to exceed 300 mg/day
	Deep IM or IV	Solution: 10 or 50 mg/mL	Same as oral
Hydroxyzine Hydrochloride (Atarax)	Oral	Tablets: 10, 25, 50, and 100 mg	Adults: 50–100 mg three to four times daily
		Syrup: 10 mg/5 mL	Children younger than 6 yrs of age: 2 mg/kg per day in divided doses
			Children older than 6 yrs of age: 12.5–25.0 mg, three to four times daily
	IM	Solution: 25 or 50 mg/mL	Same as oral
Pamoate (Vistaril)	Oral	Suspension: 25 mg/mL	Same as dosages for hydrochloride
		Capsules: 25, 50, and 100 mg	
Promethazine (Phenergan)	Oral	Tablets: 15.2, 25.0, and 50.0 mg	Adults: 50–100 mg three to four times daily for sedation
		Syrup: 3.25 mg/5 mL	Children: 12.5–25.0 mg at night for sedation
	Rectal	Suppositories: 12.5, 25.0, and 50.0 mg	
	IM	Solution: 25 and 50 mg/mL	
Cyproheptadine (Periactin)	Oral	Tablets: 4 mg	Adults: 4–20 mg/day
		Syrup: 2 mg/5 mL	Children 2–7 years of age: 2 mg two to three times daily (maximum of 12 mg/day)

	Children 7–14 years of age: 4 mg two to three times daily (maximum of 16 mg per day)

IM, intramuscular; IV, intravenous.

REFERENCES

Armstrong SC, Cozza KL. Antihistamines. *Psychosomatics*. 2003;44(5):430–434.

Brown RE, Stevens DR, Haas HL. The physiology of brain histamine. *Prog Neurobiol*. 2001;63(6):637–672.

Camelo-Nunes IC. New antihistamines: a critical view. *J Pediatr (Rio J)*. 2006;82(5 Suppl):S173–S180.

Fernández-Novoa L, Cacabelos R. Histamine function in brain disorders. *Behav Brain Res*. 2001;124(2):213–233.

Ghanizadeh A, Zare S. A preliminary randomised double-blind placebo-controlled clinical trial of hydroxyzine for treating sleep bruxism in children. *J Oral Rehabil*. 2013;40(6):413–417.

Haas H, Panula P. The role of histamine and the tuberomamillary nucleus in the nervous system. *Nat Rev Neurosci*. 2003;4(2):121–130.

*Kaplan HI, Sadock BJ, Grebb JA. Antihistamines. In: Kaplan HI, Sadock BJ, Grebb JA, eds. *Synopsis of Psychiatry*. 7th ed. Baltimore, MD: Williams & Wilkins; 1994:900.

Krystal AD, Richelson E, Roth T. Review of the histamine system and the clinical effects of H1 antagonists: basis for a new model for understanding the effects of insomnia medications. *Sleep Med Rev*. 2013;17(4):263–272.

Linneke JL, Jørgensen GK, Csillag C. Metformin for weight loss and control in patients with mood disorder. *J Clin Psychiatry*. 2014;75(10):e1140–e1141.

Linnet K, Ejsing TB. A review on the impact of P-glycoprotein on the penetration of drugs into the brain. Focus on psychotropic drugs. *Eur Neuropsychopharmacol*. 2008;18(3):157–169.

Meskanen K, Ekelund H, Laitinen J, et al. A randomized clinical trial of histamine 2 receptor antagonism in treatment-resistant schizophrenia. *J Clin Psychopharmacol*. 2013;33(4):472–478.

Montoro J, Sastre J, Bartra J, et al. Effect of H1 antihistamines upon the central nervous system. *J Investig Allergol Clin Immunol*. 2006;16 Suppl 1:24–28.

*Simons FE. Comparative pharmacology of H1 antihistamines: clinical relevance. *Am J Med*. 2002;113 Suppl 9A:38S–46S.

*Simons FE. Advances in H1-antihistamines. *N Engl J Med*. 2004;351(21):2203–2217.

*Thomas WU, Manuel ET. Antihistamines. In: Kaplan HI, Sadock BJ, eds. *Comprehensive Textbook of Psychiatry*. 6th ed. Baltimore, MD: Williams & Wilkins; 1995:1923.

Wang YQ, Takata Y, Li R, et al. Doxepin and diphenhydramine increased non-rapid eye movement sleep through blockade of histamine H1 receptors. *Pharmacol Biochem Behav*. 2015;129:56–64.

*Welch MJ, Meltzer EO, Simons FE. H1-antihistamines and the central nervous system. *Clin Allergy Immunol*. 2002;17:337–388.

Yeung WF, Chung KF, Yung KP, Ng TH. Doxepin for insomnia: a systematic review of randomized placebo-controlled trials. *Sleep Med Rev*. 2015;19:75–83.

▲ 34.9 Barbiturates and Similarly Acting Substances

STEVEN L. DUBOVSKY, M.D.

HISTORY

Stories of substances that produce sleep or unconsciousness date back to antiquity. The first commercially available sedative–hypnotics were the bromides, which were introduced in 1853 and were frequently taken in overdose. In 1864, Adolph von Baeyer discovered barbituric acid. The name barbituric acid comes from a combination of uric acid and the name Barbara. It is believed that the name Barbara was picked either because it was the name of a close friend of von Baeyer or because he celebrated discovering the new compound by visiting a tavern on St. Barbara's Day. Barbituric acid had no sedative or hypnotic properties, but diethylbarbituric acid or barbital was introduced in 1903 as the first sedative barbiturate. Many other barbituric acid derivatives were subsequently developed, beginning with phenobarbital (Luminal) in 1912, followed by a number of related compounds such as chloral hydrate (Aquachloral, Noctec) and paraldehyde (Paral).

The barbiturates were widely used as sedative–hypnotic agents in the first half of the 20th century. However, by the 1950s and 1960s, these medications became the most common agents in successful suicides, as was publicized in the deaths of celebrities such as Marilyn Monroe and Judy Garland. After additional problems such as high addiction potential and a narrow therapeutic index became apparent, "safer" alternatives such as methyprylon (Noludar), glutethimide (Doriden), ethchlorvynol (Placidyl), and methaqualone (Quaalude) were introduced. These drugs turned out to have the same problems, however, and the entire group of agents was eventually supplanted by the benzodiazepines, the first one of which (chlordiazepoxide [Librium]) was synthesized in 1955 and marketed in 1960. Glutethimide, methaqualone, and other derivatives were withdrawn from the market because of significant toxicity and risks of misuse. Barbiturates are still occasionally used, mostly in anesthesia, including for ECT, for detoxification from alcohol or other sedative–hypnotics, and as anticonvulsants.

CHEMISTRY

Modern barbiturates are derivatives of barbituric acid (2,4,6-trioxo-hexahydropyrimidine), the structure of which is illustrated in Figure 34.9–1. Barbituric acid itself has no CNS depressant properties, but the presence of an alkyl or aryl group at position 5 confers sedative–hypnotic and anticonvulsant activity and various side chains alter lipophilicity, potency, and elimination. For example, a sulfur group at position 2 produces

thiobarbiturates, which have different lipid solubility and therefore a different duration of action, latency of action, metabolism, and potency.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

Absorption. The barbiturates are administered orally when used as sedative–hypnotics and chronic anticonvulsants, and parenterally when used for anesthesia and for the treatment of status epilepticus. Oral doses of barbiturates are rapidly and completely absorbed from the small intestine. Onset of action ranges from 10 to 60 minutes, depending on the preparation and stomach contents. Barbiturates that are more lipid soluble, more protein bound, and have a high pK_a have a more rapid onset and shorter duration of action. When administered intravenously (IV), barbiturates have an immediate onset of action.

Distribution. Barbiturates are widely distributed and readily cross the placental barrier. Barbiturates vary in their lipid solubility and their binding to plasma proteins. More lipophilic compounds such as thiopental (Pentothal) and methohexital (Brevital) are highly protein bound (80 percent), whereas less lipophilic compounds (such as phenobarbital) are only 20 to 40 percent protein bound. More lipid-soluble barbiturates are more rapidly distributed to the brain, resulting in a rapid onset of action. Rapid efflux from the brain along a concentration gradient then leads to a short duration of action. Some weak acids, such as warfarin (Coumadin) or aspirin may displace barbiturates from plasma proteins, thus augmenting their CNS effects.

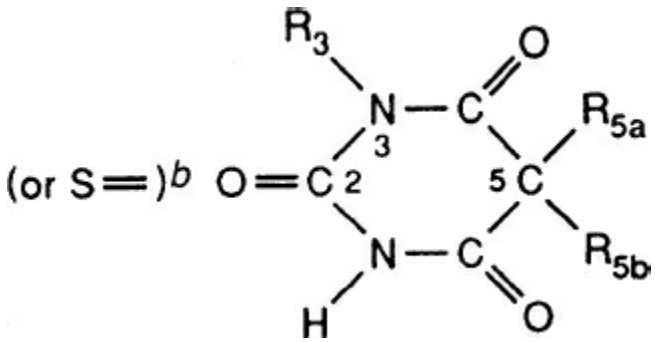

FIGURE 34.9–1. Molecular structure of barbiturates that are currently available in the United States.

Metabolism and Elimination. Most barbiturates are metabolized almost completely in the liver by cytochrome P450 microsomal enzymes. However, a significant amount of phenobarbital is excreted unchanged by the kidney. Urinary excretion of phenobarbital, which has a low pK_a, can be increased five- to tenfold by alkalinization of the urine. This approach will

not work for short-acting barbiturates because they have higher pK_as, more protein binding, and more metabolism to inactive products in the liver.

The most important biotransformation that is responsible for the termination of barbiturate activity is the oxidation of radicals at C5. N-glycosylation is also an important metabolic pathway and, in some instances, yields active metabolites. Metabolism and elimination are more rapid in young people than in the elderly, infants, and pregnant women. Metabolism is slowed, and the half-life is increased, in chronic liver disease. Elimination half-lives for some barbiturates are summarized in Table 34.9–1.

Many barbiturates induce their own metabolism as well as the metabolism of other drugs. The elimination half-life may therefore decrease significantly with repeated administration. Phenobarbital induces CYP2 isoenzymes, resulting in substantially increased metabolism and decreased effect of β-blockers, corticosteroids, doxycycline (Vibramycin), estrogen, phenothiazines, tricyclic antidepressants and some newer antidepressants, quinidine, and other anticonvulsants.

Pharmacodynamics

Mechanism of Action. Through a separate binding site allosterically linked to other binding sites on the ion channel, barbiturates increase the duration of opening of the chloride ion channel in the γ-aminobutyric acid type A receptor ($GABA_A$) complex, whereas benzodiazepines increase the frequency of channel openings. These receptors are pentameric structures surrounding a chloride ion channel on postsynaptic membranes (see Section 31.10). There are three families of receptor subunits: α, β, and γ. Most sedative–hypnotics bind to $GABA_A$ receptors containing the α_1 subunit, but etomidate (Amidate) produces sedation at the β_2 subunit and anesthesia at the β_3 subunit. Barbiturates, etomidate, and propofol (Diprivan) inhibit NMDA and amino-3-hydroxy-5-methyl-4-isoxazol-propionic acid (AMPA) receptors and reduce glutamate release and increase presynaptic glutamate reuptake. Barbiturates and related drugs inhibit excitatory nicotinic cholinergic receptors by acting at binding sites in the desensitized state of the receptor. At higher concentrations used for anesthetic purposes, barbiturates appear to reduce repetitive firing of neurons by inhibiting voltage-dependent Na^+ channels and, possibly, K^+ channels. Barbiturates and related drugs may also affect binding of other drugs for their binding sites. For example, propofol potentiates the effect of pentobarbital (Nembutal) on the $GABA_A$ receptor.

Blood Concentrations. Although barbiturates have greater effects at higher plasma concentrations, tolerance to these drugs is common; therefore, the same plasma level may produce drastically different actions

at different times in the same individual or from one person to another. Tolerance to barbiturates is associated with downregulation and desensitization of receptors, structural changes in receptors, and uncoupling of the drug from the receptor in addition to declining serum levels due to autoinduction of metabolism. Tolerance to anesthetics such as thiopental may occur within minutes. Tolerance may not have a linear relationship to dose. The Mellanby effect is acute tolerance that results in more clinical effect when serum levels are increasing than when they are decreasing. Cross-tolerance of all drugs acting on the $GABA_A$ receptor is common. For example, tolerance to benzodiazepines is associated with decreased affinity of barbiturate-binding sites.

DESIGN AND INTERPRETATION OF CLINICAL DRUG STUDIES

Because barbiturates were mostly used for insomnia and anxiety before modern protocols for diagnosis and treatment response, there are few reliable controlled studies of these indications. Although extensive clinical experience indicates that barbiturates are effective as sedative–hypnotics, the rapid development of tolerance, frequent rebound insomnia and REM sleep rebound, multiple interactions, and narrow therapeutic index make these drugs poor initial choices for insomnia. Barbiturates and related compounds are only approved for short-term treatment of insomnia and are at best second- or third-line treatment for severe acute insomnia after the benzodiazepines and nonbenzodiazepine $GABA_A$ agonists and possibly sedating antidepressants or antipsychotic drugs.

Table 34.9–1.
Dosage, Half-Life, Onset, and Duration of Action of Selected Barbiturates

Barbiturate	Dosage Range	Half-Life (hours)	Onset (minutes)	Duration (hours)
Amobarbital (Amytal)	50–300 mg PO (hypnotic); 65–500 IV (anesthesia; interview)	8–42	60	10–12
Aprobarbital (Alurate)[a]	40–120 mg	14–34	45–60	6–8
Butabarbital (Butisol)	15–120	34–42	45–60	6–8
Mephobarbital (Mebaral)	100–600	5–6	60+	10–12
Methohexital (Brevital)	0.2–1% IV infusion; 1 mg/kg for ECT	3–6	1	<0.25
Pentobarbital (Nembutal)	25–400 IV or IM 100 mg IV each minute up to 500 mg	15–50	15–48	3–4
Phenobarbital (Luminal)	90–300	80–120	80–120	10–12
Secobarbital (Seconal)	100–200	15–40	10–15	3–4

Thiopental (Pentothal)	3 mg/kg per hr	6–46	5–15	1–4
Related compounds				
Chloral hydrate (Aquachloral)	500–2,000	4–9.5	15–30	6–8
Ethchlorvynol (Placidyl)[a]	—	10–25	15–30	6–8
Etomidate (Amidate)	0.2–0.6 mg/kg IV	2.9–5.3	1–5	5–30
Glutethimide (Doriden)[a]	—	5–22	30–60	3–6
Meprobamate (Miltown)[a]	—	6–17	30–60	4–8
Methaqualone (Quaalude)[a]	—	19	15–30	2–6
Paraldehyde (Paral)[a]	—	7	30–60	4–8
Propofol (Diprivan)	5–50 mcg/kg per min IV infusion	4–23	1–3	3–8 min

PO, by mouth; IV, intravenous; ECT, electroconvulsive therapy; IM, intramuscular.

[a]No longer available in United States.

Although the sedative effect of barbiturates may reduce anxiety, there is little empirical evidence of a reliable antianxiety effect from controlled studies in any specific anxiety disorder. Benzodiazepines and antidepressants, especially the SSRIs, are generally recommended as first-line agents in the treatment of anxiety disorders.

There is variable evidence supporting the use of barbiturates in the treatment of a number of neurological conditions.

Phenobarbital is considered a safe, economical, and effective treatment for generalized tonic–clonic seizures and simple and complex partial seizures. Several studies have shown that barbiturate therapy reduces intracranial pressure (ICP), and it is therefore sometimes used in the management of acute head injury. However, a systematic review found no evidence that reduction of ICP by barbiturates is associated with reduction of morbidity or mortality. In addition, hypotension caused by the barbiturates can reduce cerebral perfusion pressure. Barbiturate-containing analgesic products are often prescribed for headaches (including vascular, tension, and mixed migraines) as well as premenstrual and postpartum pain. A Canadian review of the literature found that there was no evidence of a clinically important enhancement of analgesia by adding a barbiturate to an established analgesic, and the risk of dependence and withdrawal outweighs any benefit from sedation caused by the barbiturate. In addition, use of barbituates 5 days or more per month is associated with progression of migraines due to rebound headaches.

Barbiturates are still routinely used to induce anesthesia. Methohexital is the most frequently utilized agent for anesthesia for ECT. Propofol, etomidate, and thiopental have also been found to be effective for this indication. Propofol has increasingly been used for outpatient anesthesia

(e.g., for colonoscopy), but this application requires careful monitoring and benzodiazepines such as midazolam may be safer. Fentanyl appears to be an equivalent anesthetic agent to propofol, but it is associated with better postoperative pain control.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Barbiturates depress the activity of all excitable tissues. The CNS is more sensitive to these effects than are peripheral tissues. Barbiturates produce varying degrees of CNS depression, from mild sedation to general anesthesia. Although all barbiturates have anticonvulsant properties at anesthetic doses, only phenobarbital and mephobarbital (Mebaral) have anticonvulsant properties at subhypnotic doses. Barbiturates depress the sensory cortex, decrease motor activity, alter cerebellar function, and produce drowsiness, sedation, and sleep. At subanesthetic doses, barbiturates may increase the reaction to painful stimuli by suppressing inhibitory pain centers.

Barbiturates decrease sleep latency and the number of awakenings, and they increase total sleep time, but they also decrease REM sleep as well as stages III and IV sleep. Tolerance develops to the hypnotic effect, often within a few days. Within 2 weeks of continued use, the effects of barbiturates on prolonging sleep are virtually eliminated and doses must be increased to obtain the same effect. Abrupt discontinuation of barbiturates results in insomnia, increased dreaming, and increased nightmares (REM rebound).

In the peripheral nervous system, barbiturates selectively inhibit transmission in autonomic ganglia and reduce nicotinic excitation. This results in reduction in blood pressure, at times with clinically significant hypotension.

Respiratory drive and mechanisms for sleep rhythms are depressed by barbiturates. At therapeutic doses, the neurogenic drive for wakefulness is only reduced to that seen during normal sleep. However, at three times the hypnotic dose, the wakefulness drive is virtually eliminated. These doses also appear to reduce the hypoxic drive for breathing, exacerbating sleep apnea. Protective reflexes are not depressed until a level of sedation is reached that is associated with severe respiratory depression; this results in one of the most serious complications of barbiturate anesthesia, namely laryngospasm. These mechanisms may have contributed to the death of the entertainer Michael Jackson after escalating unmonitored self-administration of propofol in an attempt to suppress tolerance, rebound and withdrawal.

Barbiturate effects on the cardiovascular system are generally mild. At sedative–hypnotic doses, barbiturates produce a slight decrease in blood pressure and heart rate but usually no more than that seen during normal sleep. This reduction in blood pressure and heart rate can be exaggerated in

patients with congestive heart failure or hypovolemic shock. IV barbiturate administration may increase the incidence of ventricular arrhythmias.

The primary effect of barbiturates on the liver is induction of microsomal enzymes. Induction of cytochrome P450 isoenzymes inhibits metabolism of a number of medications and substances, including many anticonvulsants, antidepressants, antipsychotic drugs, and steroid hormones. In addition, autoinduction of P450 enzymes results in acceleration of metabolism of the barbiturates themselves.

THERAPEUTIC INDICATIONS

Many barbiturates, including amobarbital (Amytal), butabarbital (Butisol), aprobarbital (Alurate), pentobarbital, and secobarbital (Seconal), are indicated for oral use as sedatives or hypnotics. However, their use in insomnia should be limited to 2 weeks or less because they lose much of their effectiveness after this period of time. Secobarbital and pentobarbital are also approved as preanesthetics. Methohexital has been approved for use in ECT anesthesia.

Seizure Disorders

Barbiturates are standard anticonvulsants in the prophylaxis of seizures and the treatment of status epilepticus. These medications inhibit epileptic foci and reduce the spread of abnormal activity. Phenobarbital is the most commonly used barbiturate and is often administered along with other anticonvulsant classes.

Phenobarbital is effective in treating simple and complex partial seizures, as well as generalized tonic–clonic seizures. For status epilepticus, benzodiazepines and phenytoin (Dilantin) are first-line treatments. However, if those treatments fail and seizure activity continues, IV phenobarbital (20 mg/kg) at a rate no faster than 100 mg/min may be administered. IV benzodiazepines may also be effective. If status epilepticus continues, the patient may be anesthetized with pentobarbital. Because it raises the seizure threshold, ECT can be used for refractory status epilepticus.

Electroconvulsive Therapy

In general, the brief duration of ECT requires shorter-acting agents. Methohexital is usually the first-line agent for anesthesia during ECT. The rapid onset and offset of action of etomidate make it an appropriate choice. Thiopental is longer acting and its anticonvulsant properties can interfere with induction of therapeutic seizures. Propofol may also shorten seizure duration. The proconvulsant action of ketamine (Ketalar) makes it useful when seizures are difficult to induce, and ketamine may have acute antidepressant effects but it can cause psychosis and it may have

cardiovascular side effects. A randomized comparison of etomidate, propofol, and thiopental found all to be relatively free of important cardiovascular effects in otherwise healthy patients undergoing ECT. Etomidate raises the seizure threshold less than propofol, and its use is associated with a requirement for a lower stimulus dose to produce adequate seizures.

Cerebral Lateralization

The intracarotid amobarbital procedure (IAP or Wada test) is used to determine lateralization of memory and language in order to predict the risks of amnesia and aphasia following epilepsy surgery. Usually, amobarbital (100 to 200 mg) is administered directly into the internal carotid artery first contralateral, and then on the next day ipsilateral, to the epileptic focus. When one hemisphere is anesthetized, all functions that localize to that hemisphere are temporarily interrupted, and emergence of aphasia or memory impairment on tests such as naming, repeating, following commands, and remembering objects indicates that such functions are localized to that hemisphere. By comparing the effects of the two injections, it can be determined whether surgery on the side of the focus is likely to interrupt these functions. A recent study suggested that intracarotid methohexital was not as reliable for this purpose.

Narcoanalysis

Since the 1930s, interviews with amobarbital have been performed to distinguish between psychogenic and neurological symptoms and to elicit a better history. The amobarbital interview has been used as a diagnostic tool in cases of mutism, stupor, and suspected malingering or conversion disorder. The procedure involves administering IV amobarbital slowly in doses up to a maximum of 500 mg. Initial responses include dizziness, blurred vision, and fatigue, followed by drowsiness and mild euphoria. It is during this phase that the interview may be most fruitful. If the dose is increased further, the patient is likely to fall asleep. Controlled studies of the amobarbital interview are difficult to conduct, but clinical experience suggests that it may help an uncommunicative patient provide a history or clarify confusing psychogenic symptoms. The amobarbital interview should only be performed if personnel and equipment for resuscitation are present and increased ICP is not suspected.

Diagnosing Withdrawal from Sedative–Hypnotics

Since CNS depressants are cross-reactive through actions at different sites on the $GABA_A$ receptor, the barbiturate tolerance test can be used to determine the extent of tolerance to barbiturates, related drugs, benzodiazepines, and alcohol alone or in combination. This procedure

involves a test dose of 200 mg of pentobarbital or 30 to 100 mg of phenobarbital for patients who are thought to be experiencing CNS depressant withdrawal (and therefore will be tolerant to the barbiturate. If the patient is grossly intoxicated an hour later, tolerance is not significant and a withdrawal syndrome is unlikely. For less marked intoxication after the test dose (e.g., dysarthria and ataxia without depressed consciousness), further doses are administered every 1 to 2 hours for pentobarbital, or every 1 to 4 hours for phenobarbital, until signs and symptoms of withdrawal are suppressed, distinct intoxication emerges, or until a maximum dose (e.g., 480 mg of phenobarbital or 500 mg of pentobarbital) is reached. The substituted drug is then administered in divided dose and the dose is decreased by about 10 percent per day as determined each day by emergence of signs of toxicity (e.g., sedation, nystagmus) or withdrawal (e.g., hypertension, hyperreflexia) charted before and after each dose of barbiturate. Because it is available in an injectable form and it has a longer duration of action, phenobarbital is preferred as the withdrawal agent (with 30 mg of phenobarbital being roughly equivalent to 100 mg of pentobarbital), whereas either drug can be used to diagnose tolerance.

Delirium Tremens

Oral or IV phenobarbital, sometimes in doses as high as 1,500 to 2,000 mg in the first day, have been used as alternatives to benzodiazepines in the treatment of alcohol withdrawal delirium. Propofol in doses up to 4 mg/kg/hr has been used for severe intractable withdrawal delirium, especially if the patient is intubated. In such severe cases, ECT might also be considered.

Neuroprotection

Because they reduce neuronal electrical activity and energy demands and antagonize NMDA and AMPA receptors, barbiturates alone or with hypothermia have been used for years to protect the brain during surgery and sometimes following CNS injury, especially intraoperative stroke. If damage is not severe brain oxygen levels are increased by pentobarbital. However, any protective action is short term because these medications also have systemic immunosuppressive properties related to inhibition of T-cell calcineurin/calmodulin function, and their use for this purpose has been supplanted by newer anesthetic agents such as isoflurane (Forane). Barbiturates are used in many settings to reduce ICP following TBI, but these medications can also cause hemodynamic instability that reduces cerebral perfusion. Data on improved outcomes after barbiturate treatment following TBI have not been robust. Animal studies suggest that propofol and related medications may reduce beta amyloid-induced mitochondrial dysfunction and improve cognition. However, demented patients may be

more susceptible to adverse cognitive effects, although they appear not to have increased vulnerability to respiratory depression after anesthesia. By combining nucleic acid bases, heterocyclic amines, and barbiturate derivatives, a group of nucleobase-based barbiturates have been developed that appear to have antioxidant properties that protect against DNA damage in laboratory studies.

PRECAUTIONS AND ADVERSE REACTIONS

Barbiturate overdoses can be dangerous: 1 g (10 pills) causes serious toxicity and 2 g or more may be fatal. Acute overdose is manifested by dysarthria, ataxia, incoordination, CNS depression, and respiratory depression with respiratory acidosis, tachycardia, hypotension, hypothermia, and oliguria. Overdoses of specific drugs are associated with additional features. For example, chloral hydrate and ethchlorvynol produce unique odors; meprobamate (Miltown) and chloral hydrate cause cardiac arrhythmias; methaqualone, propofol, and etomidate induce myoclonus; chloral hydrate and methaqualone can cause GI bleeding; and glutethimide produces anticholinergic signs. Intravenous overdoses additionally produce toxicity associated with the material in which the drug is diluted. Propylene glycol in intravenous formulations of etomidate (and lorazepam [Ativan]) causes hypotension, hyperosmolar states, and metabolic acidosis.

Because death from barbiturate overdose is caused by cardiorespiratory collapse, treatment involves supporting airway, oxygenation, and cardiovascular function. Volume expansion is used to treat severe hypotension, with vasopressors only used when this approach does not work or causes volume overload. β-Adrenergic antagonists may be helpful for tachyarrhythmias. Activated charcoal with a protected airway can be helpful, although controlled research does not necessarily confirm the benefit of this practice. Hemodialysis is used for patients with chloral hydrate overdoses with life-threatening arrhythmias and in cases of prolonged intubation after ingesting large quantities of phenobarbital. The benzodiazepine receptor antagonist flumazenil (Romazicon), which can reverse benzodiazepine overdose, is not useful for overdoses of barbiturates, which act on a different receptor on the GABA receptor chloride ion channel complex. Severe barbiturate overdose can be successfully treated with hemodialysis.

The elderly, children, and those with brain damage may develop paradoxical excitation in response to barbiturates and other CNS depressants. Cases have been reported in which phenobarbital caused or exacerbated hyperactivity when used chronically in younger children, and the hyperactivity did not remit when the medication was withdrawn.

As noted earlier, barbiturates are often prescribed in combination with analgesics, but they may exacerbate pain in some patients. There is little

evidence of an analgesic effect of this class of medication. On the other hand, headache preparations containing butalbital are clearly implicated in medication overuse headaches. In this syndrome, which usually involves migraine headaches, continued use of the medication results in rebound headaches, which are then treated with more self-administered medication. Discontinuing the medication can be a challenge, since rebound headaches may persist for some time after detoxification and relapse of medication use is common.

Hypersensitivity reactions can occur with the use of barbiturates. These reactions are most likely to occur in patients with asthma, urticaria, or angioedema. Allergic reactions range from local swelling to exfoliative dermatitis. Bullous lesions have occasionally been caused by phenobarbital, more frequently but not exclusively in comatose patients.

Barbiturates are contraindicated in patients with porphyria variegata or acute intermittent porphyria because they enhance porphyrin synthesis and aggravate the illness.

Barbiturate Dependence

Barbiturates rapidly induce tolerance and dependence. Dependence is most common with shorter-acting agents such as pentobarbital and secobarbital. Pre-existing substance use increases the risk of barbiturate dependence. The risk of dependence on barbiturates increases with higher doses taken for longer periods of time (e.g., 400 mg per day of secobarbital equivalent taken for 90 days or more). Patients who have become dependent on barbiturates take on average 1,500 mg per day.

Rapid discontinuation of barbiturates induces withdrawal syndromes that include anxiety, tremor, myoclonus, dizziness, nausea, vomiting, insomnia, and hypotension, progressing in more severe cases to delirium, seizures, coma, and death. Withdrawal begins within a few hours of discontinuation and lasts days; signs and symptoms appear and abate more rapidly after discontinuing shorter-acting agents. As was described earlier, withdrawal of barbiturates with or without other CNS depressants acting at the $GABA_A$ receptor complex can be diagnosed with the barbiturate tolerance test.

Teratogenicity

Barbiturates inhibit progesterone synthesis by inhibiting cholesterol transport, and they could be risky to the pregnancy itself. Because barbiturates readily cross the placental barrier, caution should be used in prescribing them to pregnant women. There is some evidence that maternal consumption of barbiturates is associated with fetal abnormalities, especially cardiac anomalies. Moreover, newborns of mothers who use barbiturates in the third trimester may have withdrawal

symptoms after birth. Barbiturates are excreted in breast milk and may produce sedation and withdrawal syndromes in nursing infants. Concerns have been expressed about potential adverse effects of barbiturate and related anesthetics during pregnancy, but exposure is brief, and any risk is balanced by adverse effects of the condition for which anesthesia is administered.

Psychomotor Impairment

As is true of benzodiazepines, barbiturates can impair driving in addition to causing sedation and an intoxicated feeling. The degree of impairment is similar to that with alcohol.

DRUG INTERACTIONS

Complications from the concomitant use of barbiturates and other medications most often result from additive CNS depressant effects or from the induction of hepatic microsomal enzymes. Additive CNS depression occurs with alcohol, antihistamines, benzodiazepines, and other sedative–hypnotics. Metabolism of vitamins D and K is accelerated by barbiturates, which may reduce calcium absorption. Induction of hepatic metabolism may decrease the effectiveness of oral anticoagulants such as warfarin, corticosteroids, and estrogen. The latter effect may result in failure of oral contraceptives. Phenobarbital interferes with oral absorption of griseofulvin (Grifulvin V), thereby decreasing blood levels and possibly decreasing therapeutic response. Phenobarbital induces metabolizing enzymes for doxycycline, phenytoin, TCAs, and neuroleptics. Valproate (Depakote, Depakene) may decrease barbiturate metabolism, thus reducing the dose of barbiturate needed, while MAOIs may prolong the effects of barbiturates by inhibiting their metabolism. In view of frequent induction of CYP450 enzymes, it is important to look for altered efficacy of medications to which barbiturates are added.

LABORATORY INTERFERENCES

Routine toxicology screening does not identify barbiturates. Barbiturates have been reported to produce a false-positive phentolamine (Regitine) test result. Phenobarbital accelerates metabolism of dexamethasone (Decadron), resulting in a false-positive dexamethasone suppression test. Induction of metabolism of metyrapone by barbiturates may also produce false-negative responses in the metyrapone (Metopirone) test resulting from decreased levels of the test medication.

DOSAGE AND ADMINISTRATION

Dosages and routes of administration are given in Table 34.9–1. Table 34.9–2 lists some doses of preparations that contain barbiturates in

combination with other medications.

Table 34.9–2.
Barbiturate-Containing Medications

Brand-Name Product	Barbiturate	Other Contents
Fiorinal with codeine	Butalbital (Butisol), 50 mg	Aspirin (Bayer), 325 mg; caffeine, 40 mg; codeine, 30 mg
Fioricet with codeine	Butalbital, 50 mg	Acetaminophen (Tylenol), 325 mg; caffeine, 40 mg; codeine, 30 mg
Esgic	Butalbital, 50 mg	Caffeine, 40 mg; acetaminophen, 325 mg
Donnatal	Phenobarbital (Luminal), 16.2 mg	Atropine, 0.02 mg; hyoscyamine (Anaspaz), .1 mg; scopolamine (Transderm-Scop), 6.5 µg

NONBARBITURATE, NONBENZODIAZEPINE SEDATIVE–HYPNOTIC AGENTS

Chloral Hydrate

Chloral hydrate was introduced in 1832. It was widely used in the late 19th century as a sedative and hypnotic. Chloral hydrate was also used to treat many psychiatric illnesses, such as anxiety, psychosis, and mania, until more effective treatments were introduced. Soon after its introduction, there were reports of overdose deaths and by 1910, it was considered to be more dangerous than heroin or opium. The routine use of chloral hydrate as a sedative–hypnotic has mostly been replaced by benzodiazepines and related medications, but it is still used regularly to sedate children for medical procedures. In addition, some clinicians prefer to prescribe it because it is relatively inexpensive, and it is a schedule IV medication. Probably as a matter of tradition, chloral hydrate is also often used as a hypnotic agent during industry sponsored clinical trials of novel psychotropic drugs.

Chemistry. Chloral hydrate is formed by attaching a water molecule to the carbonyl group of chloral,2,2,2-trichloroacetaldehyde. The hypnotic effect of chloral hydrate is attributable to its first active metabolite, trichloroethanol.

Pharmacological Actions

PHARMACOKINETICS. Chloral hydrate is rapidly and completely absorbed from the GI tract after oral administration, but it causes GI irritation. It is rapidly converted by hepatic dehydrogenase to trichloroethanol, which has an elimination half-life of 4 to 12 hours and is metabolized by alcohol and aldehyde dehydrogenases to the inactive trichloroacetic acid; it is also conjugated with glucuronide to urochloralic acid, which is excreted by the kidney. About 10 percent of the total chloral hydrate dose is excreted

unchanged. Plasma concentrations reach their peak in 30 to 60 minutes. Chloral hydrate and trichloroethanol are highly lipid soluble, readily penetrate the BBB and the placenta, and are excreted in breast milk. In newborns, the half-life of trichloroethanol is around 28 hours, and it drops to 10 hours in toddlers. Repetitive dosing in infants causes drug accumulation, and trichloroacetic acid has been detected 6 days after administration to infants is discontinued.

PHARMACODYNAMICS. Most of the pharmacological action of chloral hydrate is attributable to trichloroethanol. Trichloroethanol acts through the $GABA_A$ receptor complex in a manner similar to barbiturates, increasing chloride ion influx and reducing neuronal excitability.

Design and Interpretation of Clinical Drug Studies. Chloral hydrate was used at one time to treat alcohol withdrawal, but this use was never supported by controlled trials. Comparisons of chloral hydrate and benzodiazepines suggested equivalent efficacy for short-term treatment of insomnia, but the benefits of chloral hydrate diminished after 2 weeks due to tolerance.

Chloral hydrate is often used as a sedative during pediatric procedures such as echocardiography, MRI, and other radiological procedures. A randomized study found that chloral hydrate was as effective as midazolam (Versed) in producing adequate sedation during pediatric echocardiography.

Effects on Specific Organs and Systems

CENTRAL NERVOUS SYSTEM. The CNS depressant action of chloral hydrate produces drowsiness, sedation, lightheadedness, malaise, ataxia, and nightmares. Hangover effects seen with other hypnotics are rare with chloral hydrate. Chloral hydrate has few analgesic properties and is occasionally associated with agitation in the presence of pain. CNS effects of overdose of chloral hydrate are similar to overdose of other CNS depressants.

CARDIOVASCULAR SYSTEM. Chloral hydrate and trichloroethanol reduce myocardial contractility, shorten the refractory period, and increase sensitivity of the myocardium to endogenous catecholamines. As a result, most deaths from chloral hydrate poisoning are caused by cardiac dysrhythmias that are refractory to standard antiarrhythmics. β-Adrenergic blocking agents may be helpful in this setting.

RESPIRATORY SYSTEM. In typical clinical doses, there is little effect of chloral hydrate on respiration. However, overdose may cause respiratory depression.

GASTROINTESTINAL SYSTEM. Chloral hydrate may cause GI irritation, resulting in nausea and vomiting. This is most common when the medication is ingested on an empty stomach. Overdoses can cause hemorrhagic gastritis and occasionally gastric or intestinal necrosis with perforation, as well as esophagitis with stricture.

Therapeutic Indications

INSOMNIA. Chloral hydrate is indicated for the short-term treatment of insomnia. Most patients respond to a dose of 500 to 1,500 mg, with sleep taking place within 1 hour. Chloral hydrate may be effective in patients who are medically ill, the young, and elderly patients who become agitated or confused with a benzodiazepine.

SEDATION FOR PEDIATRIC PROCEDURES. Chloral hydrate is used in children undergoing a variety of diagnostic and minor surgical procedures. It can be administered orally or rectally, and the onset of its effects is rapid. Generally, 25 to 100 mg/kg is used in children, up to a maximum of 1 g per dose. In children younger than 2 years of age, chloral hydrate appears to be less effective. The dose of chloral hydrate may be repeated, but the child must be monitored until fully awake.

PREOPERATIVE USE. Chloral hydrate can be used preoperatively to reduce anxiety associated with surgery and to induce sleep without depressing respiration or the cough reflex if the dose is not excessive.

Precautions and Adverse Reactions

CENTRAL NERVOUS SYSTEM. Chloral hydrate can cause sedation, dizziness, lightheadedness, confusion, and, occasionally, paradoxical excitement. The CNS depressant effects of chloral hydrate may be additive to those of other CNS depressants, especially barbiturates and alcohol.

ABUSE AND DEPENDENCE. Similar to many other CNS depressants, chloral hydrate may be habit forming, particularly in those who have a history of alcohol or other drug dependence. Tolerance and physiological dependence may develop if chloral hydrate is used for long periods of time on a regular basis. Sudden withdrawal of chloral hydrate can precipitate a syndrome that is similar to barbiturate withdrawal and can be diagnosed and treated using the barbiturate tolerance test described above.

OVERDOSE. Signs and symptoms of overdose are similar to those seen in barbiturate overdose. Symptoms of overdose begin with ataxia and lethargy and may progress to stupor or coma. Initially, pupils may be constricted but then dilate as coma deepens. Toxic doses for adults are approximately 10 g. Treatment of overdose includes gastric lavage and supportive

measures. Hemodialysis may be effective in promoting the clearance of trichloroethanol.

ALLERGIC REACTIONS. Allergic skin reactions have been reported with chloral hydrate, including hives, eczematoid dermatitis, and urticaria. Occasionally, its use has been associated with leukopenia and eosinophilia.

GASTROINTESTINAL SYSTEM. Chloral hydrate can irritate the GI system, particularly when taken on an empty stomach. Symptoms can include nausea, vomiting, diarrhea, and an unpleasant taste.

PORPHYRIA. There have been reports of chloral hydrate precipitating attacks of acute intermittent porphyria, and, therefore, it should not be used in patients with porphyria.

CARCINOGENICITY. There have been several laboratory studies suggesting carcinogenicity of chloral hydrate. However, there is currently no evidence that chloral hydrate is carcinogenic in humans.

Drug Interactions. Chloral hydrate competes with ethanol for alcohol dehydrogenase, resulting in an increase of the half-life of ethanol by concurrent chloral hydrate administration. At the same time, ethanol inhibits the conversion of chloral hydrate to trichloroethanol as well as the conjugation of trichloroethanol. The combination of ethanol and chloral hydrate is known as a Mickey Finn (knockout drops) and produces rapid unconsciousness. Chloral hydrate has an unpredictable effect on warfarin use and may cause excessive hypoprothrombinemia. Patients who also take furosemide (Lasix) may experience vasomotor instability due to the displacement of furosemide from protein-binding sites resulting in increased free drug. Symptoms may include vasodilation, flushing, tachycardia, sweating, and hypotension.

Laboratory Interference. Because it is radiopaque, chloral hydrate may be detected on abdominal x-ray. Chloral hydrate may interfere with copper sulfate tests for glycosuria. It may also interfere with tests for urine catecholamines or urinary 17-hydroxycorticosteroid determinations.

Dosage and Administration. Chloral hydrate is available in an oral form, as well as a suppository. Chloral hydrate capsules are formulated in 250- or 500-mg strengths. Chloral hydrate syrup is available in a 500 mg per 5 mL concentration. The oral syrup should be mixed with 4 to 8 oz of water or juice to reduce GI upset. Suppositories are formulated in 325- and 650-mg strengths.

The usual adult dose for insomnia is 500 to 1,000 mg taken approximately 15 to 30 minutes before bedtime. Single, daily, adult doses

should not generally exceed 2 g. Pediatric doses vary from 25 to 100 mg/kg per dose, with a maximum dose of 1 g.

OTHER SEDATIVE–HYPNOTICS

Over the years, a number of medications with varying structures have been used as sedative–hypnotics. Glutethimide was introduced as a substitute for barbiturates. Its absorption was erratic and it had more than 14 metabolites, some contributing to toxicity. In addition, recirculation in the enterohepatic circulation contributed to fluctuating, prolonged toxicity. Glutethimide was unique in causing anticholinergic side effects, leading to thick bronchial secretions that further compromised respiration after overdose. Psychosis, seizures, peripheral neuropathy, and ataxia were associated with prolonged use. Glutethimide therefore was withdrawn in the United States.

Meprobamate was introduced in 1950 as a muscle relaxant. Carisoprodol (Soma), which was introduced 5 years later, is metabolized to meprobamate. Both of these propanediol carbamates have actions on the $GABA_A$ receptor and on NMDA receptors similar to the barbiturates. Meprobamate is more likely than other drugs in the group to cause euphoria. Meprobamate pills can agglutinate in the stomach to form concretions, leading to recurrent toxicity after overdose. Paraldehyde was used parenterally for agitation in alcohol withdrawal, but this approach caused sterile abscesses and the drug was withdrawn. Ethchlorvynol caused severe withdrawal syndromes and was particularly lethal after overdose.

Anesthetic Agents

Propofol. Propofol is a $GABA_A$ agonist. It also induces presynaptic release of GABA and dopamine (the latter possibly through an action on $GABA_B$ heteroreceptors on dopaminergic neurons) and is a partial agonist at dopamine D_2 and NMDA receptors. Because it is very lipid soluble, it crosses the BBB readily and induces anesthesia in less than 1 minute. Rapid redistribution out of the CNS results in offset of action within 3 to 8 minutes after the infusion is discontinued. It is well tolerated when used for conscious sedation, but it has a potential for acute adverse effects including respiratory depression, apnea, and bradyarrhythmias, and prolonged infusion can cause acidosis and mitochondrial myopathies. The carrier used for the infusion is a soybean emulsion that can be a culture medium for various organisms. The carrier also can impair macrophage function and cause hematological and lipid abnormalities and anaphylactic reactions. Recent concerns about unmonitored outpatient use for indications other than anesthesia were mentioned earlier.

Etomidate. Etomidate is a carboxylated imidazole that acts at the β_2

and β_3 subunits of the $GABA_A$ receptor. It has a rapid onset (1 minute) and short duration (less than 5 minutes) of action. The propylene glycol vehicle has been linked to hyperosmolar metabolic acidosis, as has been the case for parenteral lorazepam (see Chapter 34.10). It has both proconvulsant and anticonvulsant properties, and it inhibits cortisol release, with possible medical consequences after long-term use.

References

Alali AS, Naimark DM, Wilson JR, et al. Economic evaluation of decompressive craniectomy versus barbiturate coma for refractory intracranial hypertension following traumatic brain injury. *Critical Care Medicine*. 2014;42:2235–2243.

Anonymous. Chloral and chloral hydrate. *IARC Monogr Eval Carcinog Risks Hum*. 1995;63:245–269.

*Bigal ME, Lipton RB. Excessive acute migraine medication use and migraine progression. *Neurology*. 2008;71:1821–1828.

Browne ML, Van Zutphen AR, Botto LD, Louik C, Richardson S, Druschel CM. Maternal butalbital use and selected defects in the national birth defects prevention study. *Headache*. 2014;54:54–66.

Caksen H, Odabas D. Perioral skin and mucosal lesions due to chloral hydrate. *Pediatr Dermatol*. 2001;18:454–455.

Da Silva AN, Lake AE 3rd. Clinical aspects of medication overuse headaches. *Headache*. 2014;54:211–217.

Dhorajiya BD, Dholakiya BZ, Ibrahim AS, Badria FA. Nucleobase-based barbiturates: their protective effect against DNA damage induced by bleomyciun-iron, antioxidant, and lymphocyte transformation assay. *BioMed Res Int*. 2014;2014:898670.

Ferrari A, Tiraferri I, Palazzoli F, et al. Hair analysis to monitor abuse of analgesic combinations containing butalbital and propyphenazone. *J Pharm Biomed Anal*. 2015;115:576–579.

*Flomenbaum NE, Goldfrank LR, Hoffman RS, Howland MA, Lewin NA. *Goldfrank's Toxicologic Emergencies*. 8th ed. New York: McGraw-Hill; 2006.

Guillen Sans R, Guzman Chozas M. Historical aspects and applications of barbituric acid derivatives. A review. *Pharmazie*. 1988;43:827–829.

Hutto B, Fairchild A, Bright R. γ-Hydroxybutyrate withdrawal and chloral hydrate. *Am J Psychiatry*. 2000;157:1706.

McCarron MM, Schulze BW, Walberg CB, Thompson GA, Ansari A. Short acting barbiturate overdosage: correlation of intoxication score with serum barbiturate concentration. *JAMA*. 1982;248:55–61.

*Pershad J, Palmisano P, Nichols M. Chloral hydrate: The good and the bad. *Pediatr Emerg Care*. 1999;15:432–435.

Roberts I, Sydenham E. Barbiturates for acute traumatic brain injury. *Cochrane Database Syst Rev*. 2012;12:CD000033.

*Rosa MA, Rosa MO, Marcolin MA, Fregni F. Cardiovascular effects of anesthesia in ECT: a randomized, double-blind comparison of etomidate, propofol and thiopental. *J ECT*. 2007;23:6–8.

Salmon AG, Kizer KW, Zeise L, Jackson RJ, Smith MT. Potential carcinogenicity of chloral hydrate—a review. *J Toxicol Clin Toxicol*. 1995;33:115–121.

Shao H, Zhang Y, Dong Y, Yu B, Xia W, Xie Z. Chronic treatment with anesthetic propofol improves cognitive function and attenuates caspase activation in both aged and Alzheimer's disease transgenic mice. *J Alzheimers Dis*. 2014;41:499–513.

*Smith DE, Wesson DR. Phenobarbital technique for treatment of barbiturate dependence. *Arch Gen Psychiatry*. 1971;24:56–60.

Stephen MC, Mathew J, Varghese AM, Kurien M, Mathew GA. A randomized controlled trial comparing intranasal midazolam and chloral hydrate for procedural sedation in children. *Otolaryngol Head Neck Surg*. 2015;153(6):1042–1050.

Weaver MF. Prescription sedative misuse and abuse. *Yale J Biol Med*. 2015;88:247–256.

*Wheeler DS, Jensen RA, Poss WB. A randomized, blinded comparison of chloral hydrate and midazolam sedation in children undergoing echocardiography. *Clin Pediatr*. 2001;40:381–387.

Yaekel JW, Logan BK. Butalbital and driving impairment. *J Forensic Sci*. 2013;58:941–945.

Zeller A, Arras M, Jurd R, Rudolph U. Identification of a molecular target mediating the general anesthetic actions of pentobarbital. *Mol Pharmacol*. 2007;71(3);852–859.

Zenk KE. High dose barbiturate therapy in children. Part 2. Pharmacology and pharmacokinetics. *Infusion*. 1983;7:124.

▲ 34.10 Benzodiazepine Receptor Agonists and Antagonists

STEVEN L. DUBOVSKY, M.D.

HISTORY

Alcohol has been used as a sedative since biblical times. Bromides were introduced as hypnotics in the mid-19th century. Barbital, the first barbiturate, was introduced in 1903, and phenobarbital (Barbita) was synthesized in 1912. Along with the later introduction of related drugs, such as meprobamate (Equanil), the barbiturates were the mainstays of the pharmacological treatment of anxiety and insomnia until the benzodiazepines were developed in the late 1950s.

The first benzodiazepine to be introduced was chlordiazepoxide (Librium), which was synthesized in 1959 at Roche Laboratories. In 1963, diazepam (Valium) was developed in the same laboratory. Over the next three decades, superior safety and tolerability helped the benzodiazepines replace the older antianxiety and hypnotic medications. Barbiturates and related medications are now rarely, if ever, used for anxiety and only occasionally used for insomnia, whereas 10 to 15 percent of the US population takes benzodiazepines in a given year. Recently, antidepressants, especially the SSRIs, have been used in place of benzodiazepines for some chronically anxious patients, and sedating

antidepressants, such as trazodone (Desyrel) and mirtazapine (Remeron), have replaced benzodiazepines, to some extent, as hypnotics.

Studies of the actions of the benzodiazepines and barbiturates have led to the elucidation of the structure and function of their primary target, the GABA–benzodiazepine receptor complex. This research has led to the development of other medications that act on this complex as agonists, antagonists, and inverse agonists. Understanding benzodiazepine receptors and their interaction with linked receptors makes it straightforward to predict the actions of a diverse group of medications.

GABA–Benzodiazepine Receptor Complex

GABA, the major inhibitory neurotransmitter in the CNS, interacts with inotropic GABA type A ($GABA_A$) heteroreceptors located on neurons in arousal and motor systems to decrease excitability of these neurons. The pentameric GABA receptor complex is a member of the nicotinicoid superfamily and includes six subunit families (α, β, γ, ρ, δ, θ) and 16 specific subunits that are encoded by separate genes and are allosterically linked to form a chloride ion channel. Interaction of the $GABA_A$ receptor with GABA increases chloride channel openings as described below, allowing more negative charges into the neuron, hyperpolarizing it (Fig. 34.10–1A,B). A $GABA_B$ receptor, which is linked by a G protein to calcium and potassium channels, is a presynaptic autoreceptor that modulates release of GABA, glutamate, and monoamines. A structurally and functionally distinct $GABA_C$ (also called $GABA_P$) receptor has recently been identified in the retina, spinal cord, superior colliculus, pituitary, and gut. These receptors appear to be involved in the regulation of vision, memory, and the sleep–wake cycle. The balance between inhibitory GABAergic, inhibitory glycinergic, and excitatory glutaminergic systems is fine tuned to set the responsiveness of the CNS to a host of neurotransmitters and neuromodulators.

The benzodiazepine receptor is contiguous to the $GABA_A$ receptor and is located at the interface between the α and γ subunits. Occupancy of this receptor does not affect the chloride ion channel directly but changes the conformation of the GABA receptor to increase its affinity for GABA. Benzodiazepine receptor agonists therefore have the effect of increasing the action of each available molecule of GABA in enhancing chloride ion influx and hyperpolarizing neurons on which the receptor complex is located (Fig. 34.10–1C). A picrotoxin site associated with the benzodiazepine–GABA receptor complex reduces GABA receptor affinity, reducing chloride ion influx and tending to depolarize the same neuron. A number of other substances, including inhaled and intravenous (IV) anesthetics, barbiturates, neurosteroids, and alcohol, alter $GABA_A$ receptor affinity at various binding sites, not all of which have been identified.

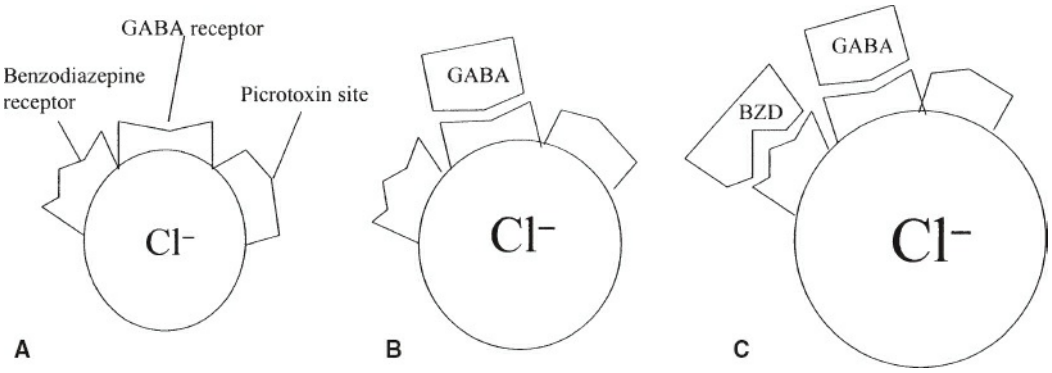

FIGURE 34.10–1. Benzodiazepine (BZD) receptor actions. The BZD receptor is allosterically linked to a γ-aminobutyric acid (GABA) type A receptor; both receptors outline a chloride ion (Cl^-) channel **(A)**. When they are stimulated by their agonist, GABA receptors increase Cl^- channel openings and hyperpolarize neurons **(B)**. Interaction of the BZD receptor with an agonist increases the affinity of GABA receptors, resulting in more Cl^- channel opening for each molecule of GABA **(C)**.

Three major benzodiazepine receptor subtypes have been described, two of them called central benzodiazepine receptors and one called a peripheral benzodiazepine receptor (PBR). Benzodiazepine-1 (Ω-1) receptors contain the α1 subunit and are located throughout the CNS. These receptors are thought to mediate the sedative, hypnotic, and antianxiety effects of benzodiazepine receptor agonists. Benzodiazepine-2 (Ω-2) receptors, which are a heterogeneous group of receptors containing α2, α3, or α5 subunits, are located in the cortex, hippocampus, striatum, and spinal cord and on pyramidal neurons, and mediate anxiolysis, muscle relaxation, CNS depression sedation, and psychomotor impairment and contribute to the anticonvulsant effect.

Benzodiazepine-3 (Ω-3) receptors (PBRs) are 18-kDa proteins with a structure and function distinct from that of the central (i.e., benzodiazepine-1 and benzodiazepine-2) receptors. PBRs are found on glial and other brain cells as well as throughout the body. These receptors, which bind benzodiazepines and the endogenous inverse agonist, are primarily located in the outer mitochondrial membrane and may contribute to tolerance and withdrawal. They are also involved in a number of complicated processes, including synthesis of neurosteroids implicated in some psychiatric disorders, and regulation of apoptosis. The mitochondrial PBR is linked to a mitochondrial permeability transition pore (MPTP) consisting of a nonselective ion channel, which alters the transmembrane potential, leading to release of proapoptotic proteins from the mitochondrion and other actions. The PBR plays an important role in protecting neurons from oxidative damage, which is implicated in Parkinson disease, Alzheimer disease, and negative schizophrenia. At micromolar concentrations at the receptor, PBR ligands may promote apoptosis, whereas nanomolar concentrations may, under some circumstances, interfere with it. An agonist effect at this receptor could

protect against oxidative damage, whereas an inverse agonist effect could play a role in cancer chemotherapy, which enhances apoptosis. Carbamazepine (Tegretol) binds to PBRs in lymphocytes, which may contribute to altered chemotaxis by these cells in carbamazepine-treated patients.

Benzodiazepine Receptor Ligands

Compounds that bind to the benzodiazepine receptor can have five different kinds of action. Nonselective full agonists (e.g., benzodiazepines) cause all benzodiazepine receptors to increase affinity of GABA receptors, which increases the frequency of chloride channel openings and hyperpolarizes neurons. Selective full agonists (e.g., zolpidem [Ambien]) have the same effect, but this action is selective for the benzodiazepine-1 receptor. Clobazam (Frisium) is the only benzodiazepine with more affinity for benzodiazepine-2 than for benzodiazepine-1 receptors. Partial agonists bind to all benzodiazepine receptors, but they are less effective than full agonists. If low levels of an endogenous agonist for a receptor are present, partial agonists occupy the receptor and stimulate it, although less so than the full agonist. If levels of the endogenous agonist are high, a partial agonist competes for receptors but is less effective than the full agonist, essentially reducing activity at the receptor. In the case of the benzodiazepine receptor, partial agonists have a less pronounced therapeutic effect at the receptor than full agonists, but they also have fewer adverse effects.

Inverse agonists occupy receptors in the same manner as full agonists, but they have the opposite effect. An inverse agonist of the benzodiazepine receptor binds to the same site as the benzodiazepines and reduces GABA receptor affinity, decreases the frequency of chloride channel openings, and depolarizes neurons. Inverse agonists produce arousal and anxiety and sometimes cause seizures and psychosis. Antagonists occupy a receptor without affecting its activity, thereby blocking the action of agonists and inverse agonists. Flumazenil (Romazicon) is a benzodiazepine receptor antagonist used to treat benzodiazepine overdose.

A number of medications and substances that depress or activate the CNS act on the benzodiazepine–GABA receptor complex, and many of these agents are cross-reactive. For example, picrotoxin binds to the chloride channel pore and reduces chloride channel openings to induce seizures. Alcohol and barbiturates have direct effects on GABA receptors and chloride channels and possibly on the picrotoxin site, and the reduction of cortical and limbic neuronal activity results in sedation and psychomotor impairment. Valproate (Depakene, Depakote) enhances GABA release, and gabapentin (Neurontin) has unclear actions on GABAergic transmission. Propofol (Diprivan), an IV anesthetic discussed in Chapter 34.9, acts at a site other than the benzodiazepine receptor to

increase GABA affinity synergistically with benzodiazepines at higher doses. Unidentified constituents of valerian, an herbal preparation used for centuries as a hypnotic and anxiolytic, seem to be GABA receptor agonists. Zopiclone (Imovane), which is widely used in Europe, and eszopiclone (Lunesta), which is available in the United States, are nonbenzodiazepines that act as partial agonists on a site close to the benzodiazepine receptor, with effects that are virtually the same as but less marked than those of direct benzodiazepine receptor agonists. The anticonvulsant tiagabine (Gabitril) inhibits GABA reuptake, increasing GABA action at all GABA receptors, and the anticonvulsant vigabatrin (Sabril) reduces GABA degradation. Baclofen (Lioresal) is a $GABA_B$ agonist that reduces spasticity via an action on presynaptic receptors on excitatory spinal cord neurons. This section reviews medications that act primarily on the benzodiazepine receptor on the $GABA_A$ receptor complex.

BENZODIAZEPINES

Chemistry

Benzodiazepines contain two benzene rings and a seven-member diazepine ring (Fig. 34.10–2). Because all benzodiazepines in clinical use contain a 5-aryl substituent ring and a 1,4-diazepine ring, they are 5-aryl-1,4-benzodiazepines. A fused triazolo ring at positions R1 and R2 on the diazepine ring in triazolobenzodiazepines such as alprazolam (Xanax), triazolam (Halcion), and estazolam (ProSom), increases potency; the same is true of a fused imidiazo ring at these positions in midazolam (Versed). Replacement of ring C with a keto substituent in position 5 and a methyl substituent in position 4 confers benzodiazepine antagonist features to flumazenil. Clobazam, the only 1-5-benzodiazepine, is not yet approved in the United States.

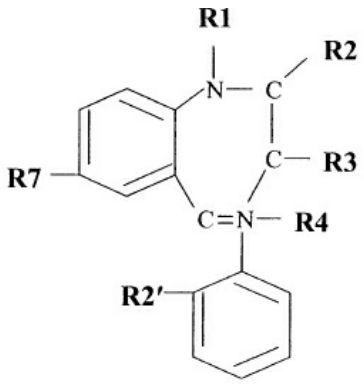

Benzodiazepine	R1	R2	R3	R7	R2′
Alprazolam	Fused triazolo ring		-H	-Cl	-H
Chlordiazepoxide	-	$-NHCH_3$	-H	-Cl	-H
Clonazepam	-H	=O	-H	NO_2	-Cl
Diazepam	$-CH_3$	=O	-H	-Cl	-H

FIGURE 34.10–2. Basic benzodiazepine structure.

Pharmacological Actions

Absorption. All benzodiazepines except clorazepate (Tranxene) are completely absorbed after oral administration and reach peak serum levels within 30 minutes to 2 hours. Metabolism of clorazepate in the stomach converts it to desmethyldiazepam, which is then completely absorbed. Intramuscular (IM) absorption of benzodiazepines other than lorazepam (Ativan) is slower than oral absorption. The onset of action is nearly immediate with IV administration of high-potency benzodiazepines, such as midazolam.

Distribution. All benzodiazepines are lipid-soluble, but the degree of lipid solubility varies considerably between drugs. Benzodiazepines and their active metabolites bind to plasma proteins, the extent of binding being proportional to their lipid solubility. The amount of protein binding varies from 70 to 99 percent.

Distribution, onset, and termination of action after a single dose are determined by benzodiazepine lipid solubility. Preparations with high lipid solubility, such as diazepam and alprazolam, are absorbed rapidly from the GI tract and distribute rapidly to the brain by passive diffusion along a concentration gradient, resulting in a rapid onset of action. However, as the concentration of the medication increases in the brain and falls in the bloodstream, the concentration gradient reverses itself, and these medications leave the brain rapidly, resulting in fast cessation of drug

effect. Drugs with longer elimination half-lives, such as diazepam, may remain in the bloodstream for a substantially longer period of time than their actual pharmacological action at benzodiazepine receptors with single or intermittent dosing because the concentration in the brain falls rapidly below the level necessary for a noticeable effect. In contrast, lorazepam, which has a shorter elimination half-life than diazepam but is less lipid soluble, has a slower onset of action after a single dose, because the drug is absorbed and enters the brain more slowly. However, the duration of action after a single dose is longer because it takes longer for lorazepam to leave the brain and for brain levels to fall below the concentration that produces an effect. In chronic dosing, some of these differences are not as apparent because brain levels are in equilibrium with more consistent steady-state blood levels, but additional doses still produce a more rapid but briefer action with diazepam than with lorazepam.

Benzodiazepines are distributed widely in adipose tissue. As a result, some of these medications may persist in the body after discontinuation longer than would be predicted from their elimination half-lives. In addition, the dynamic half-life (i.e., duration of action on the receptor) may be longer than the elimination half-life.

Metabolism. Some benzodiazepines (e.g., oxazepam [Serax]) are conjugated directly by glucuronidation to a water-soluble form and are excreted. Most benzodiazepines are oxidized first by cytochrome P450 (CYP) 3A4 and CYP 2C19, often to active metabolites (Fig. 34.10–3). These metabolites may then be hydroxylated to another active metabolite. For example, diazepam is oxidized to desmethyldiazepam, which, in turn, is hydroxylated to produce oxazepam. These products undergo glucuronidation to inactive metabolites. A number of benzodiazepines (e.g., diazepam, chlordiazepoxide) have the same active metabolite (desmethyldiazepam), which has an elimination half-life of more than 120 hours. Flurazepam (Dalmane), a lipid-soluble benzodiazepine used as a hypnotic that has a short elimination half-life, has an active metabolite (desalkylflurazepam) with a half-life greater than 100 hours. The duration of action of a benzodiazepine therefore may not correspond to the half-life of the parent drug.

Because administration of medications more frequently than the elimination half-life leads to drug accumulation, medications such as diazepam and flurazepam accumulate with daily dosing, eventually resulting in increased daytime sedation. Clorazepate, halazepam (Paxipam), and prazepam (Centrax; no longer available in the United States) are not pharmacologically active but are prodrugs for desmethyldiazepam (nordiazepam). In contrast, temazepam (Restoril), oxazepam, and lorazepam are conjugated directly and do not have active metabolites. Unlike barbiturates, benzodiazepines do not induce their own

metabolism.

The combination of lipid solubility, potency, and elimination half-life (Fig. 34.10–4) predicts onset and duration of action in clinical settings as well as optimal frequency of drug administration. Benzodiazepines that are high in lipid solubility and potency and have a short half-life have a rapid onset but a brief duration of action and produce more intense and abrupt discontinuation syndromes. For example, midazolam, a high-potency and highly lipid-soluble nonselective full agonist with a short elimination half-life, has an almost immediate onset of action when used IV for conscious sedation or anesthesia. Sedation diminishes rapidly and is followed by withdrawal symptoms (usually amnesia) in approximately 70 percent of patients after a single dose. This profile makes midazolam useful in anesthetic practice but impractical for chronic treatment. Alprazolam, which is high in lipid solubility and potency, reduces anxiety rapidly but can produce interdose withdrawal when it is administered in accordance with its intermediate half-life (three to four times per day), because it leaves the brain more rapidly than it is eliminated from the body. In contrast, chlordiazepoxide is low in lipid solubility and potency and has a long elimination half-life. This medication has a slow onset of action after a single dose and is administered less frequently in chronic treatment. Withdrawal from long half-life, low-potency, low lipid-soluble benzodiazepines such as chlordiazepoxide results in discontinuation syndromes that begin later after discontinuation and are less intense but more prolonged than discontinuation syndromes produced by withdrawal of alprazolam.

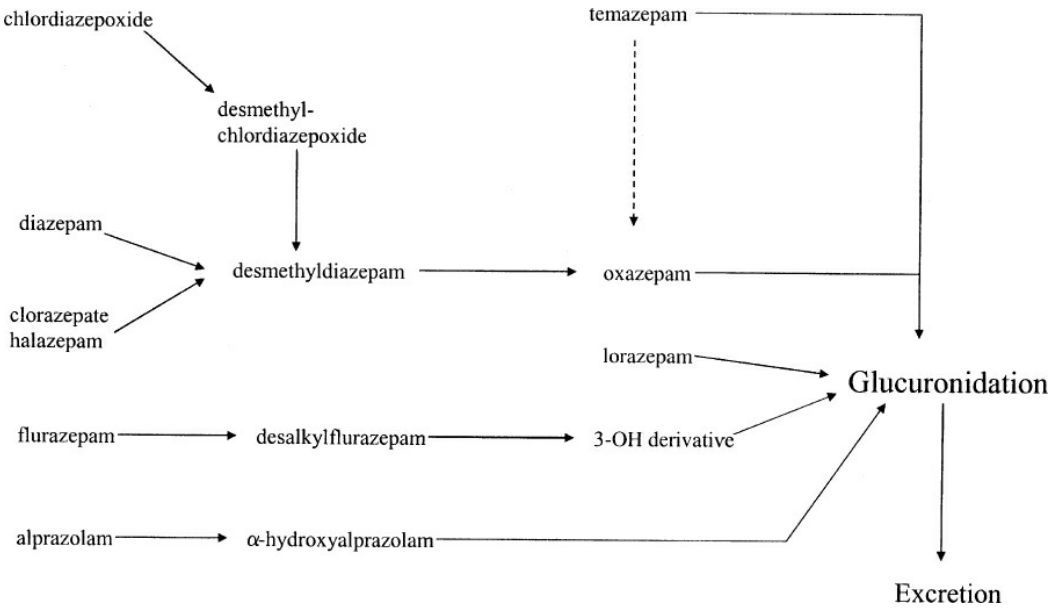

FIGURE 34.10–3. Simplified benzodiazepine metabolism.

Pharmacodynamics

Mechanism of Action. The affinity of clinically active benzodiazepines for the benzodiazepine receptor raises the question of whether there is a naturally occurring agonist for this receptor. An endogenous agonist (endozepine) has been identified, but mammals cannot synthesize it and they probably obtain it from plants. Conversely, a gene has been identified for a mammalian endogenous inverse agonist called diazepam-binding inhibitor (DBI). DBI may serve to upregulate the gain on the CNS by lowering the threshold for depolarization in arousal systems. DBI levels are increased in refractory epilepsy, suggesting that this could be a mechanism of uncontrolled gross neuronal activation. Along with cortisol, increased DBI activity during spontaneous labor (as opposed to caesarian section with anesthesia), indicating that this may be a mechanism of arousal during any kind of stress. One cause of chronic anxiety may be excessive activity of DBI, which is antagonized when benzodiazepines compete for benzodiazepine receptor occupancy and have a hyperpolarizing, rather than a depolarizing effect through opposite effects on the GABA receptor and, hence, on chloride ion influx. However, even if anxiety is caused by excessive excitability of arousal systems that is primary or that is secondary to some other influence, such as a hyperactive suffocation response, which has been demonstrated in panic disorder, benzodiazepines can reduce anxiety by enhancing hyperpolarization of systems that underlie anxiety and arousal.

The same action that results in the therapeutic action of benzodiazepines on benzodiazepine-1 receptors causes adverse effects when it is too pronounced and when benzodiazepines act on other receptors. For example, actions of benzodiazepines on benzodiazepine receptors in the cortex can cause cognitive and psychomotor impairment. Actions on pyramidal benzodiazepine receptors have an anticonvulsant effect and can also cause muscle relaxation and incoordination. Benzodiazepine hypnotics act on receptors in the reticular activating system, but if suppression of this system is prolonged, daytime sedation results. Suppression of the reticular activating system can induce delirium in demented patients and aggravate delirium if it is already present. However, benzodiazepines do not produce the same degree of CNS depression as do the barbiturates.

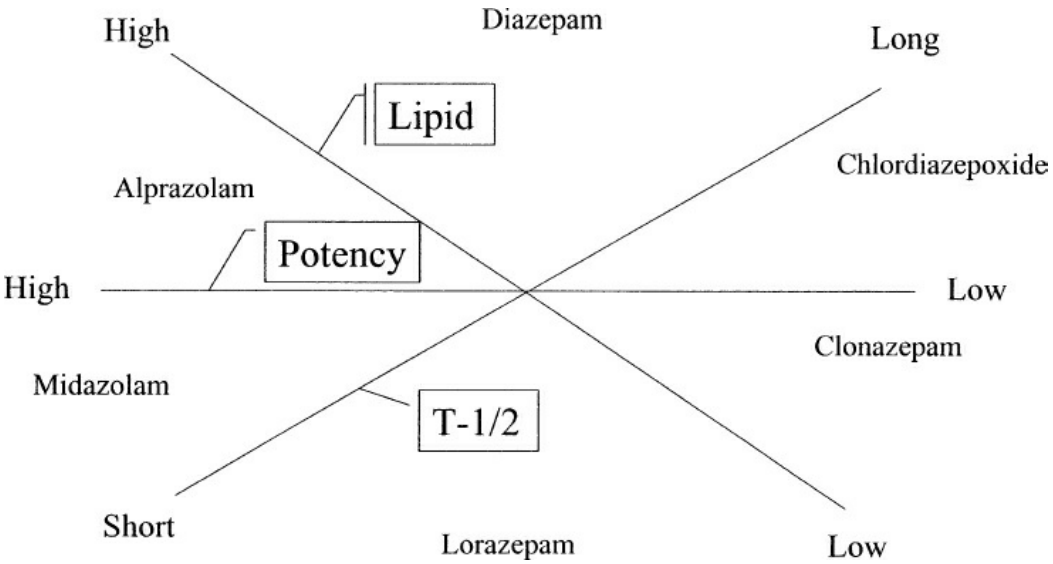

FIGURE 34.10–4. Potency, lipid solubility, and half-life (T-1/2) of some benzodiazepines.

Blood Concentration and Relation to Action. Concentrations of benzodiazepines in the CSF are equal to plasma concentrations of the free drug. Because of the complex metabolism of long-acting benzodiazepines, serum levels vary considerably between individuals. Levels of parent compounds have not been correlated with therapeutic response, and, because many of these medications have active metabolites, it is unlikely that such a correlation exists. Although many shorter-acting benzodiazepines do not have active metabolites, a correlation between blood levels and clinical response has not been demonstrated for these preparations either. There may be a correlation between levels of alprazolam and both adverse effects and efficacy for panic attacks, with response more likely at plasma concentrations around 20 to 40 ng/mL, which correspond to doses of 1.5 to 6.0 mg per day. However, such levels are rarely measured in clinical practice, and levels would be expected to fluctuate substantially between doses, making the optimal time to measure the trough level difficult to identify. On the other hand, undetectable levels in an agitated patient who has been taking benzodiazepines raise the suspicion of a withdrawal syndrome. The primary uses of serum levels therefore are for toxicology screening and identification of abstinence syndromes.

Design and Interpretation of Clinical Drug Studies

Studies of acute situational anxiety are difficult to control because exogenous stresses vary, as do past histories and medical and psychiatric comorbidities of acutely anxious patients. The fluctuating nature and high placebo response rate of primary anxiety disorders, such as GAD, social phobia, and panic disorder, require placebo controls and longer trials to

establish drug effectiveness. A number of well-validated diagnostic and assessment instruments exist that facilitate blinded outcome ratings in such trials. Studies of the usefulness of benzodiazepines as hypnotics are more difficult and expensive because ratings of amount and quality of sleep are not reliably correlated with objective sleep measures. Clinical experience suggests that factors listed in Table 34.10–1 predict a good response in the treatment of acute anxiety.

Placebo-controlled trials have clearly established the efficacy of benzodiazepines in the treatment of generalized anxiety. Multicenter, double-blind studies have also demonstrated that alprazolam is effective for panic disorder. Smaller studies indicate that, in equivalent doses, other benzodiazepines can also reduce the frequency of panic attacks, although such doses are much higher than ordinary doses for generalized anxiety. For example, 60 mg of diazepam would be equivalent to the reported antipanic dose of alprazolam of close to 6 mg per day in a pivotal multicenter trial. Although patients report significantly prolonged sleep with benzodiazepine hypnotics and these medications shorten sleep onset in the sleep laboratory compared to placebo, total sleep time is increased by an average of only 30 minutes with these drugs.

Table 34.10–1.
Predictors of a Good Response to a Benzodiazepine

Acute anxiety
Clear precipitant
Previous good response to a benzodiazepine
Patient awareness that symptoms are psychological
Patient desire for a benzodiazepine
Absence of history of substance misuse

One factor that can complicate interpretation of studies of therapeutic and adverse effects of benzodiazepines is that anxious patients and normal subjects differ in their sensitivity to both effects. For example, without any difference in pharmacokinetics, alprazolam produced more of an impact on mood and EEG but fewer adverse effects in patients with panic disorder than in normal controls.

Effects on Specific Organs and Systems

The primary action of benzodiazepines is on the CNS, where they have an anticonvulsant and anxiolytic effect and produce muscle relaxation and CNS depression. Decreased depolarization of neurons in the brainstem can cause respiratory depression, primarily at high doses. Diazepam increases coronary blood flow and has negative inotropic effects. Other benzodiazepines used as preanesthetics reduce peripheral vascular resistance, thereby decreasing blood pressure and increasing heart rate.

Benzodiazepines cross the placenta and enter breast milk.

Benzodiazepines such as diazepam also bind to PBRs. The therapeutic potential of these interactions has not been extensively explored, but it is possible that stabilizing the mitochondrial ion channel could reduce neurotoxicity in some degenerative disorders.

Therapeutic Indications

Benzodiazepines have largely replaced the older barbiturates and related medications as anxiolytics and hypnotics. However, concerns about addiction have led to unwarranted regulation of these medications and declining use for chronic anxiety disorders in some settings. The year after a requirement for special prescribing procedures for benzodiazepines was introduced in the New York State in 1989, benzodiazepine prescriptions decreased by 60 percent while prescriptions for the more dangerous meprobamate, chloral hydrate (Aquachloral), and butabarbital (Butisol) increased by 125 percent, 136 percent, and 30 percent, respectively. Not surprisingly, overdoses of the latter drugs also increased.

Although they have clear advantages of effectiveness for most forms of anxiety, generally good tolerability, and ease of prescription, benzodiazepines (and other medications) have the potential to reinforce a passive approach to symptom relief in which one relies on something outside oneself for symptom relief and does not necessarily acquire an internal sense of mastery of stressful situations. For chronic anxiety, behavioral therapies such as cognitive behavior therapy and desensitization should therefore be considered along with or as alternatives to benzodiazepines. Benzodiazepines can facilitate exposure to phobic and other anxiety-provoking situations, resulting in more enduring benefit.

Generalized Anxiety Disorder. Benzodiazepines are clearly effective for GAD. Despite the chronic and relapsing nature of this condition, experts frequently recommend that benzodiazepines be taken for brief periods of time. This approach frequently results in relapse of anxiety, which is mistaken for drug dependence. Antidepressants, especially the SSRIs, have increasingly been used as replacements or supplements for benzodiazepines in the treatment of GAD. However, a significant number of patients require ongoing therapy with benzodiazepines.

Panic Disorder. As was mentioned previously, alprazolam has been found to be superior to placebo in reducing the frequency of panic attacks, usually in doses higher than those used to treat generalized anxiety. Other benzodiazepines appear to be equally effective in equivalent doses. Antidepressants other than bupropion (Wellbutrin) are equally effective for this condition and are probably better tolerated in chronic use. Some patients do best on combinations of antidepressants and benzodiazepines.

Other Anxiety Disorders. Benzodiazepines are used as adjuncts to reduce arousal in PTSD and are sometimes used intermittently to reduce social anxiety, although chronic treatment with SSRI and MAOI antidepressants is more effective. β-Blockers are more useful acutely for performance anxiety than benzodiazepines because they are less likely to impair psychomotor function and they directly reduce autonomic arousal that interferes with performance. Benzodiazepines are appropriate initial treatments for acute situational anxiety, especially in settings in which physiological arousal could be dangerous, such as in a patient with a damaged myocardium on the coronary care unit who will be vulnerable to the arrhythmogenic effects of catecholamines released in response to anxiety.

Major Depressive Disorder. Alprazolam has been found to have some antidepressant properties, although monotherapy with this benzodiazepine is not as effective as antidepressants for severe depression. Alprazolam might be considered for a depressed patient who could not tolerate any antidepressant and who did not respond to established psychotherapies and who refused or was not a good candidate for ECT or another instrumental therapy. Antidepressants frequently increase anxiety acutely in patients with primary anxiety disorders and with mixed anxiety and depression because of an initial increase in noradrenergic activity, even with medications that are purely serotonergic in the laboratory. To block early excessive activation by antidepressants, benzodiazepines are often prescribed first. Once anxiety and other symptoms of arousal are suppressed, the antidepressant is better tolerated. Eventually, noradrenergic alerting systems are down-regulated by the antidepressant, at which point the benzodiazepine can be withdrawn. In some cases of mixed anxiety and depression, both medications are continued.

Conventional wisdom holds that benzodiazepines other than alprazolam aggravate or cause depression. A Cochrane Review meta-analysis of nine randomized, controlled studies involving 679 unipolar depressed patients found that patients treated with combinations of an antidepressant and a benzodiazepine were significantly more likely than those treated with an antidepressant alone to respond during the first 4 weeks of treatment. In these studies, adding a benzodiazepine to an antidepressant conveyed as much additional benefit as chlorpromazine (Thorazine) over placebo for acute schizophrenia and almost as much as an SSRI over placebo for major depression or lithium (Eskalith) over placebo for mania. Much of the benefit was a function of fewer early dropouts among patients receiving a benzodiazepine, probably because of more rapid reduction of anxiety and more tolerability of the antidepressant. After 6 weeks of treatment, there was no longer any difference between those taking and those not taking benzodiazepines along with antidepressants.

Adverse benzodiazepine effects did not outweigh the therapeutic benefit. One comparison of diazepam, the MAOI moclobemide (Manerix), and placebo found the benzodiazepine to be superior to the antidepressant, as well as the placebo, for anxious depression. Conversely, a prospective study of elderly (mean age 85 years at baseline) patients suggested that those taking benzodiazepines were more likely to develop new depressive symptoms and had slightly higher depression rating scale scores. No controlled data are available about the development of a new episode of major depression in nondepressed patients taking benzodiazepines.

Insomnia. Benzodiazepines increase quality and length of sleep, decrease sleep latency, and slightly reduce REM sleep. However, tolerance develops to the hypnotic effect after 1 to 2 months. Efficacy can be retained by intermittent use. Gaboxadol, an experimental nonselective extrasynaptic $GABA_A$ agonist, reduces sleep latency and increases slow wave sleep. Sedating antidepressants, especially trazodone, and the atypical antipsychotic drug quetiapine (Seroquel), are currently used more frequently than benzodiazepines to treat chronic insomnia. Benzodiazepines are useful for insomnia caused by acute stress, such as hospitalization. Hospitalized patients should be asked if they would prefer a sleeping pill rather than routinely prescribing these medications. In many cases, chronic insomnia is probably better treated with a sedating antidepressant.

Flurazepam itself has a short elimination half-life, but its major active metabolite, desalkylflurazepam, has a half-life longer than 1 day. Persistence of the drug effect makes this medication most useful for patients with daytime anxiety, and high lipid solubility can also result in a rapid onset of action. Accumulation with repeated dosing often leads to daytime sedation. The relatively slow onset of action of temazepam, which is related to its lower lipid solubility, combined with an intermediate half-life make it most appropriate for patients with middle or terminal but not initial insomnia. Triazolam, on the other hand, has a rapid onset of action and a short elimination half-life, making it useful for patients who have difficulty falling asleep who want to avoid daytime impairment. However, the same features lead to a high frequency of withdrawal symptoms, especially amnesia, the next morning. In addition, concerns have been raised about the efficacy of lower doses (0.125 mg) and the safety of higher doses (0.5 mg).

Gamma-hydroxybuturate (GHB, Xyrem), which is approved for the treatment of narcolepsy and improves slow wave sleep, is also an agonist at the $GABA_A$ receptor, where it binds to specific GHB receptors. GHB has the capacity both to reduce drug craving and to induce dependence, abuse, and absence seizures as a result of complex actions on tegmental dopaminergic systems. Such adverse effects, as well as reports of drug misuse,

substantially limit use of this drug. The application of other nonbenzodiazepines acting on the benzodiazepine receptor in the treatment of insomnia is discussed later in this chapter.

Benzodiazepine and other sleeping pills are often used to treat or prevent jet lag and work shift-related sleep phase changes. However, the sleep–wake cycle is more effectively shifted by exposure to bright light at the time of waking up than by the time of falling asleep. If a hypnotic agent is used, it should be combined with getting out in the sunlight or using artificial bright light the next morning. Melatonin, or ramelteon (Rozerem), a melatonin M1 and M2 receptor agonist, is more likely to shift the onset of sleepiness than is a primarily sedating medication.

The decision to promote a particular benzodiazepine receptor agonist as an anxiolytic or a hypnotic is based on marketing strategies and the kinds of studies that the manufacturer chose to perform to obtain FDA approval, not the pharmacology of the drug. There is no reason, for example, why preparations with similar onset and duration of action (e.g., temazepam, lorazepam, oxazepam) could not be used for the same indication.

Bipolar Disorder. Clonazepam (Klonopin) is frequently used for anxiety and insomnia in patients with unstable bipolar mood disorders. Although some authorities hold that a mildly serotonergic action of this medication could contribute to an antidepressant effect, clinical experience suggests that it may induce depression or at least oversedation with chronic use. Lorazepam, with or without antipsychotic drugs, is still used frequently as an adjunctive treatment for mania. In view of the risk that antidepressants can destabilize mood in bipolar disorder, benzodiazepines other than alprazolam (which has antidepressant properties) are safer anxiolytics for patients with bipolar illness.

Schizophrenia. Perhaps as a result of GABAergic inhibition of dopamine release, benzodiazepines have been found somewhat useful as adjuncts to antipsychotic drugs, primarily in reducing agitation and anxiety, but not as monotherapy for positive symptoms of schizophrenia. A Cochrane Review of 31 studies involving over 2,000 patients concluded that there are insufficient data to recommend benzodiazepines alone or as adjuncts in the treatment of schizophrenia.

Catatonia. Lorazepam, sometimes in low doses (less than 5 mg per day) and sometimes in very high doses (12 mg per day or more), is regularly used to treat acute catatonia, which is more frequently associated with bipolar disorder than with schizophrenia. Other benzodiazepines have also been said to be helpful. However, there are no valid controlled trials of benzodiazepines in catatonia. Chronic catatonia does not respond as well to benzodiazepines. The definitive treatment for catatonia is ECT.

Seizure Disorders. Because they reduce neuronal hyperactivity in cortical and pyramidal cells, all benzodiazepines have anticonvulsant properties. Its longer duration of action makes clonazepam the most frequently used benzodiazepine for adjunctive chronic seizure control. IV diazepam, lorazepam, and midazolam are used for status epilepticus. However, the longer a seizure persists, the more likely it is to become resistant to the anticonvulsant effect of the benzodiazepine, as has also been reported with barbiturates. Animal studies suggest that cannabinoid receptor antagonists may not exhibit rapid development of tolerance in the treatment of status epilepticus. Clobazam, which is approved in 100 countries for the treatment of Lennox-Gastaux syndrome, a form of intractable childhood epilepsy, is undergoing phase III testing in the United States. A group of 2,3-benzodiazepines that are also noncompetitive antagonists of the excitatory amino acid α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor are being investigated as antiepileptic and neuroprotective agents. As noted earlier, anticonvulsants such as valproate and tiagabine alter GABA dynamics rather than acting on the benzodiazepine receptor. The same appears to be true of pregabalin and gabapentin.

Acute Agitation. A rapid onset of action with minimal residual impairment has made IV midazolam useful for control of severe acute agitation in the emergency department. Life-threatening agitation caused by delirium or other neurological conditions can be controlled with a few doses of IV lorazepam, which is often combined in the same syringe with haloperidol (Haldol), and ECT is also rapidly effective for this purpose. Nonspecific agitation is treated in many centers with a combination of intramuscular lorazepam and haloperidol, but the advantage of this combination over either drug alone is limited to remaining asleep longer, and midazolam acts more rapidly with less residual sedation. A Cochrane Review of 11 studies with a total of 648 patients found no difference between haloperidol plus lorazepam and lorazepam alone in improvement, need for additional medication, or incidence of extrapyramidal side effects. Lorazepam is used as an adjunctive treatment to reduce excitement, improve sleep in manic patients, and treat agitation in schizophrenic patients. Although benzodiazepines can reduce the dose of antipsychotic medications in agitated manic and schizophrenic patients, there is no evidence that they have an adjunctive antipsychotic effect. Increased use of atypical antipsychotic drugs, which are better tolerated than the older neuroleptics, has reduced the need for adjunctive benzodiazepines.

The parenteral form of lorazepam contains a small amount of propylene glycol that is not problematic for isolated doses but can cause metabolic acidosis if given in high enough doses for a long enough period of time, especially in patients with renal or hepatic impairment. Doses greater than

0.1 mg/kg/hr over 1 to 2 days, which may be used as a continuous infusion for sever alcohol withdrawal, can result in propylene glycol levels of 25 mg/dL or more.

Alcohol Withdrawal. Half of people with alcohol use disorders develop withdrawal syndromes when they decrease or discontinue drinking; delirium tremens and/or seizures complicate 3 to 5 percent of cases of alcohol withdrawal. Benzodiazepines are standard treatments for alcohol withdrawal because they reduce arousal and act at a different site on the same receptor complex. No single benzodiazepine has been found to be superior to any other. Benzodiazepines with low lipid solubility and long elimination half-lives, such as chlordiazepoxide, are most frequently used because the blood level slowly declines after a decreasing dose administered over several days. Clorazepate, which has an intermediate half-life, is also used because it is a prodrug with a long half-life active metabolite. Lack of hepatic metabolism of lorazepam, which can be increased by alcohol, make this an appealing choice, but more frequent dosing is necessary because of its shorter elimination half-life.

In the treatment of delirium tremens, very high benzodiazepine doses may be needed. For example, as much as 2,000 mg of diazepam (Valium) on the first day, or 2,800 mg of midazolam over 50 days, have been used in ICU settings. One protocol recommends up to 8 mg of lorazepam (Ativan) orally, IM or IV every 15 minutes, followed by 10 to 30 mg/hour after a final 8-mg bolus. Such dosing usually requires careful monitoring in a hospital setting.

Addiction. Baclofen, a $GABA_B$ agonist, reduced the reinforcing effects of substances and food in animal studies. This or another agonist could prove useful for smoking cessation.

Neuroleptic Side Effects. Benzodiazepines are more effective than antiparkinsonian drugs in treating akathisia, but they are not as useful for parkinsonism. There is some evidence from a single study of 24 patients that benzodiazepines may reduce manifestations of tardive dyskinesia, but pooled data from other studies indicate no benefit.

Anesthesia and Conscious Sedation. Midazolam is frequently used for brief anesthesia and for conscious sedation. Other highly lipid-soluble benzodiazepines, such as diazepam, are used adjunctively for conscious sedation.

Acute Agitation and Violence. Medications that improve agitated and aggressive behavior do so primarily as a result of sedation. Even psychotic agitation is ameliorated by sedative effects of antipsychotic drugs, since

reduction of psychosis itself is not immediate. IM lorazepam (Ativan) in doses of 0.5 to 2 mg and sometimes as high as 4 mg, often in combination with 5 mg of haloperidol, is commonly used to treat acute agitation in the emergency room and in psychiatric inpatient settings. In most studies, 2 mg of lorazepam has been found to be as effective as 5 mg of haloperidol, primarily as a function of sedation. Clonazepam has also been found to be equivalent to haloperidol, although it is not available in an IM form. IM midazolam in a dose of 5 mg is very appealing as monotherapy for agitation because it has a rapid onset of action. Oversedation is not as significant a concern because of the rapid offset of action and the short half-life of midazolam; the same features may require more frequent dosing. In general, benzodiazepines are preferable to antipsychoptic drugs for most forms of agitation, because they do not cause neurological side effects or QT prolongation. The latter is an important consideration in emergency settings, where an electrocardiogram that could reveal pre-existing QT prolongation cannot realistically be performed. They are obviously useful for agitation caused by CNS depressant withdrawal, but they may aggravate other forms of delirium.

Antispasmodic, Chronic Pain and Tinnitus. Baclofen is commonly used to treat spasticity. It is also effective for pain, but tolerance develops rapidly. A small single-blind study reported significant reduction of severe disabling tinnitus of cochlear origin by benzodiazepines and other GABAergic drugs compared to placebo.

Infectious Disease. Benzodiazepines may inhibit the respiratory syncytial virus, an action that seems attributable to the S-enantiomer. The in vitro action of benzodiazepines for leishmaniasis is equivalent to the standard treatment for this illness, sodium stibogluconate.

Precautions and Adverse Effects

Most adverse effects of the benzodiazepines involve the CNS.

Sedation. Although some benzodiazepines (e.g., alprazolam) are less sedating than others (e.g., chlordiazepoxide), all benzodiazepines are CNS depressants and can cause intolerable sedation. Tolerance often develops to this side effect if the dose is raised slowly. Respiratory depression can occur as a result of depression of the central respiratory drive in high doses or in patients with liver disease. This mechanism, along with the muscle relaxant effect, makes nighttime benzodiazepine use risky for patients with sleep apnea and chronic obstructive pulmonary disease.

Cognitive and Psychomotor Impairment. Retrograde amnesia, caused by withdrawal from short-acting, high-potency benzodiazepines such as

triazolam and midazolam, can occur after a single dose. Withdrawal probably also accounts for amnesia following a dose of the high-potency lipid-soluble "date rape drug" flunitrazepam (Rohypnol), especially when it is combined with alcohol. All benzodiazepines can cause anterograde amnesia and can interfere with new learning, although they do not usually interfere with memories that are already acquired.

All patients taking benzodiazepines should be warned about the risk of impaired psychomotor performance. Impairment of driving caused by a single 10-mg dose of diazepam is equivalent to impairment caused by a blood alcohol level sufficient to qualify for a citation for driving under the influence. Patients whose driving is impaired often are not aware of the problem. Tolerance to psychomotor impairment does not occur as readily as does tolerance to sedation.

At one time it was thought that an endogenous benzodiazepine receptor ligand caused hepatic encephalopathy. However, a careful study of controls and cirrhotic patients with and without encephalopathy found that increased concentrations of benzodiazepine receptor ligands in encephalopathy patients were entirely explained by administration of benzodiazepines for sedation or endoscopy.

In some cases, persistent delirium associated with alcohol withdrawal or mixed CNS depressant use is caused by benzodiazepines, especially those with longer durations of action. This condition can be diagnosed by improvement with the benzodiazepine receptor antagonist flumazenil if delirium has persisted for 5 days. Flumazenil can precipitate benzodiazepine withdrawal if benzodiazepine use has continued.

Aggravation of Dementia. Cognitive side effects of benzodiazepines can exacerbate the cognitive deficits of dementia. Interference with balance and psychomotor performance can also lead to falls, which can be devastating in elderly, frail, patients. These medications therefore should be used with great caution if at all in elderly patients, and in 2012 the American Geriatrics Society included benzodiazepines in its list of drugs that are inappropriate for older patients because of cognitive side effects. If a benzodiazepine is indicated, preparations with shorter half-lives and simpler metabolism such as lorazepam or oxazepam in lower doses are preferable. Nevertheless, a recent study found that benzodiazepine use increased from 2.6 percent of the population aged 18 to 26 years to 8.7 percent of those aged 65 to 80 years, with 24 percent of all benzodiazepine users taking long-acting preparations, especially when nonpsychiatric physicians prescribed them.

Concern about the risk of Alzheimer disease in patients taking benzodiazepines was raised by a recent case-control study of 1,796 patients with Alzheimer disease who had been followed prospectively for at least 6 years prior to diagnosis, and 7,184 matched controls (age >66 years). After

controlling for relevant risk factors and anxiety, insomnia, and depression, Alzheimer disease was 1.32 times more likely in patients taking benzodiazepines for 3 to 6 months and 1.84 times more likely in patients taking benzodiazepines for more than 6 months; less than 3 months use of a benzodiazepine was significantly associated with Alzheimer disease. The association was stronger for long half-life (1.70 times increase) than for short half-life benzodiazepines (1.43 times increase). Although not all research has supported this finding, another study found a 50 percent increased risk of Alzheimer disease within 15 years after starting a benzodiazepine. Such results could indicate that benzodiazepine use reduces cognitive reserve sufficiently for Alzheimer disease to be expressed earlier, that using a benzodiazepine for extended periods of time is a marker of an as yet unidentified risk factor for Alzheimer disease, or that benzodiazepines themselves increase the risk of Alzheimer disease through an unclear mechanism.

Disinhibition. Demented patients may become more confused when they take benzodiazepines because further depression of the reticular activating system induces delirium. These patients become agitated as a direct result of the reduced organizing function of brainstem structures and because they become more confused and frightened. Some hypervigilant individuals become more suspicious or agitated when their ability to keep track of their environment is impaired by a benzodiazepine.

Abuse and Dependence. A 1990 population survey found that 10 to 15 percent of the population used benzodiazepines during the previous year, and 2 percent used benzodiazepines chronically. Benzodiazepine dependence and abuse are more prevalent in individuals who abuse other substances. By contrast, misuse of benzodiazepines is considered rare in medical practice, occurring primarily in people with a history of misusing these or other substances. However, a Canadian study of more than 400,000 hospitalizations of elderly patients who had not previously taken benzodiazepines found that 1.5 percent had new chronic use of benzodiazepines that were started in the hospital. The likelihood of new chronic benzodiazepine use was highest in women, patients admitted to the intensive care unit or nonsurgical wards, those with longer hospital stays, patients with more comorbidity, a prior diagnosis of alcoholism, and patients taking more medications. Benzodiazepine use by narcotic addicts increases complexity and refractoriness of the substance use disorder.

A comprehensive interview study of 33 primary care physicians found them to be reluctant to address discontinuing long-term benzodiazepine use with elderly patients because they felt unskilled in doing so and because the patients usually resisted the change. Similarly, a study by the same group of elderly chronic benzodiazepine users in primary care practice

found them to be unwilling to consider discontinuing or tapering the medication because they felt that it reduced stress and in some cases even prolonged life. Conversely, a Spanish study found that with biweekly follow-up and a standard withdrawal schedule, patients taking benzodiazepines chronically were almost five times as likely to discontinue or substantially reduce benzodiazepine use as they were with usual practice (number needed to treat = 3). Most studies demonstrate that gradual taper of benzodiazepines is more likely to result in successful discontinuation than rapid withdrawal. There is not enough research to determine the efficacy of the common practice of switching to a long-acting benzodiazepine before withdrawing short-acting benzodiazepines, but this practice may result in fewer dropouts. It is important to distinguish between a desire to prevent relapse and ongoing excessive benzodiazepine misuse before embarking on discontinuing benzodiazepines. For patients who need ongoing benzodiazepine use but should not be exposed to the continued risk of these medications (especially older patients), adding an antidepressant first may facilitate discontinuation of the benzodiazepine. It should at least be possible to cross over to a benzodiazepine with a shorter half-life in older patients.

In studies of addict ratings of the degree to which a particular drug would be sought and emergency department admissions for drug abuse or dependence, diazepam, alprazolam, and lorazepam have been judged to be most prone to misuse. This finding probably reflects the high potency and lipid solubility of the first two medications, resulting in a rapid onset of action perceived as a high, and interdose withdrawal leading to escalation of the dose. Lorazepam does not have a rapid onset of action, but the combination of high potency and a short elimination half-life can lead to dosage escalation to avoid withdrawal.

Dysphoria. Patients who feel intensely anxious may find the rapid onset of high-potency lipid-soluble benzodiazepines to be reassuring. Those who are afraid of losing control may become dysphoric when they feel rapidly calmed or sedated by the same medication and may feel more in control with the slow onset of action of less lipid-soluble preparations.

Pregnancy and Lactation. Benzodiazepines readily cross the placenta. Concerns have been raised that exposure to benzodiazepines in utero causes cleft lip and palate, but a large case-controlled study of diazepam and chlordiazepoxide found no teratogenicity and a case control study of more than 38,000 pregnancies found no association of teratogenicity with any of five benzodiazepines. A critical review of the literature on the teratogenic risks of benzodiazepines during the first trimester of pregnancy from 2001 to 2011 found no evidence of congenital malformations and noted that studies were limited by lack of detailed data about

benzodiazepine use during pregnancy, possible recall bias, and lack of control for confounding factors. The authors recommended diazepam and chlordiazepoxide as first choices during pregnancy. All benzodiazepines are excreted in breast milk, in which concentrations are close to serum concentrations for more lipid-soluble preparations. Although the total dose of benzodiazepine at any time to the nursing infant is low, sedation, hypotonia, and decreased feeding may occur. Some infants chronically exposed to benzodiazepines may develop discontinuation syndromes when the mother stops taking the benzodiazepine or stops breastfeeding.

Discontinuation Syndromes. Three types of syndromes have been clearly described when benzodiazepines are discontinued. All are more pronounced when the medication is withdrawn more rapidly and with shorter half-life, high-potency, and lipid-soluble preparations.

RELAPSE. Return of symptoms of the condition that the medication was treating is called *relapse*. Relapse of anxiety disorders begins over weeks to months and lasts indefinitely with fluctuating severity unless it is treated again because most anxiety disorders are chronic or recurrent.

REBOUND. *Rebound* refers to an intensification of the original symptoms. Rebound symptoms, which are worse than was the case before treatment began, start within days to weeks of drug discontinuation and last from days to weeks. The level of symptomatology then returns to the baseline (i.e., rebound is replaced by relapse). Rebound anxiety may occur between doses of highly lipid-soluble, high-potency, short half-life benzodiazepines, such as alprazolam. Taking such medications every 2 to 3 hours is not necessarily a sign of misuse and may simply reflect the rapid onset and termination of drug effect.

WITHDRAWAL. *Withdrawal* syndromes involve new autonomic signs and symptoms (Table 34.10–2) that are not components of the original disorder. Withdrawal begins within hours to days of discontinuation of short-acting benzodiazepines but may not occur for 1 to 2 weeks after abrupt discontinuation of longer-acting preparations. Withdrawal typically lasts for a few days for shorter-acting medications to 1 or 2 weeks for longer-acting drugs, but attenuated withdrawal symptoms have been observed 1 year after discontinuation of long half-life benzodiazepines. Slow downward taper of the benzodiazepine dose reduces the risk of withdrawal symptoms but does not prevent them. Some patients experience interdose withdrawal in addition to rebound with high-potency short-acting benzodiazepines. Withdrawal may occur after 4 to 8 months of treatment with typical benzodiazepine doses or after half that duration of treatment with higher doses.

Table 34.10–2.
Signs and Symptoms of Benzodiazepine Withdrawal

Anxiety	Tremor
Irritability	Depersonalization
Insomnia	Hyperesthesia
Hyperacusis	Myoclonus
Nausea	Delirium
Difficulty concentrating	Seizures

Withdrawal can probably best be understood as a natural consequence of benzodiazepine receptor adaptation to chronic benzodiazepine use. Benzodiazepine receptors downregulate when presented with a large number of benzodiazepine molecules. At the same time, the inverse agonist DBI is upregulated to compensate for depression of CNS excitability. When the benzodiazepine is abruptly withdrawn, there are more DBI molecules than there are benzodiazepine molecules competing for fewer benzodiazepine receptors. This results in a shift of the benzodiazepine receptor toward inverse agonism, causing decreased, rather than increased, GABA receptor affinity and reduced, rather than enhanced, chloride influx. CNS excitability increases abruptly, resulting in manifestation of arousal ranging from jitteriness and tremor to myoclonus and seizures. Slower benzodiazepine withdrawal allows the benzodiazepine receptor to upregulate, while DBI is gradually downregulated, but the unpredictability of this accommodation may lead to withdrawal symptoms with a slow taper or even with a reduction in benzodiazepine dose. An additional contributing factor to rebound insomnia after discontinuation of long-term benzodiazepine hypnotic use may be dysregulation of endogenous melatonin rhythms.

Withdrawal from benzodiazepines, as well as many other CNS depressants, is treated by reinstituting the previous dose and withdrawing the medication more slowly or by substituting a longer-acting benzodiazepine or another CNS depressant with a known pattern of withdrawal, preferably one that acts at the same receptor complex. A barbiturate tolerance test with phenobarbital or pentobarbital (Nembutal) can be used to more precisely determine the dose of CNS depressant necessary to suppress withdrawal, as described in Chapter 34.9. Standardized withdrawal protocols using one of these barbiturates can then be applied. In a recent study, addition of 2 mg of sustained-release melatonin was superior to placebo in promoting withdrawal of benzodiazepine hypnotics in 19 of 24 patients who were able to discontinue the benzodiazepine and were able to maintain good sleep quality without return to benzodiazepine use. However, patients were then maintained on melatonin.

As chronic doses of alprazolam are reduced below approximately 1 mg,

the benzodiazepine receptor undergoes paradoxical downregulation, which can cause rebound and withdrawal symptoms, even with an extremely slow taper. Some clinicians find that substitution of clonazepam facilitates withdrawal of the final 1.0 to 1.5 mg, but others have found that clonazepam does not reliably suppress alprazolam withdrawal.

A confusing clinical picture may emerge when alcohol-dependent patients who have also been taking long-acting benzodiazepines are admitted to the hospital. The onset of withdrawal seizures or alcohol withdrawal delirium may be postponed up to several weeks because it is prevented by the benzodiazepine, which has a longer duration of action at the GABA–benzodiazepine receptor complex. When the serum level of the benzodiazepine eventually declines, typical alcohol withdrawal emerges, sometimes complicated by benzodiazepine withdrawal. The late onset of symptoms may result in withdrawal delirium being mistaken for postoperative psychosis or a complication of a medical illness.

Drug Interactions

Antacids may reduce GI absorption of benzodiazepines. Otherwise, benzodiazepine interactions may be pharmacokinetic or pharmacodynamic. Inhibitors of CYP 3A4, such as erythromycin (E-Mycin), ketoconazole (Nizoral), cimetidine (Tagamet), estrogen, disulfiram (Antabuse), verapamil (Isoptin), nefazodone (Serzone), and grapefruit juice, increase benzodiazepine levels, whereas carbamazepine and St. John wort, which induce this enzyme, lower benzodiazepine levels. These are not issues for benzodiazepines with minimal hepatic metabolism, such as oxazepam and lorazepam.

Pharmacodynamic interactions mainly involve additive CNS depression with other sedating agents, such as antihistamines and sedating antidepressants. Combining benzodiazepines and alcohol results in additive psychomotor impairment. Adding benzodiazepines to clozapine can produce respiratory and cardiovascular depression. There are some reports of respiratory arrest in patients taking buprenorphine and parenteral benzodiazepines, but this interaction does not seem to have been as important with lower doses of oral benzodiazepines.

Laboratory Interferences

Benzodiazepines do not affect most laboratory tests.

Dosage and Administration

Usual benzodiazepine doses for the treatment of anxiety are summarized in Table 34.10–3. Even medications with long elimination half-lives are often given in divided doses to avoid oversedating peak levels. In chronic treatment, especially at higher doses, it is sometimes necessary to

administer alprazolam frequently (e.g., every 2 to 3 hours) to avoid interdose withdrawal and rebound anxiety. An alprazolam sustained-release preparation (Xanax XR) has acceptable steady-state kinetics with three daily doses. Slow absorption of temazepam administered at bedtime makes this approach most useful for patients who have no trouble falling asleep but wake up early. To reduce sleep latency, this medication should be taken 1 to 2 hours before bedtime.

The common approach of withdrawing benzodiazepines after 2 months of treatment usually results in relapse in chronic and relapsing anxiety disorders, such as GAD and panic disorder. However, it is often possible to suppress anxiety initially with a benzodiazepine and then add an antidepressant, which is better tolerated when arousal has been controlled acutely with the benzodiazepine. It may then be possible to withdraw the benzodiazepine. Some patients who have been taking stable doses of benzodiazepines for years without evidence of impairment or misuse and who cannot tolerate other treatments should simply continue the benzodiazepine rather than undergo multiple unsuccessful attempts at drug discontinuation.

PARTIAL AGONISTS AND SELECTIVE BENZODIAZEPINE RECEPTOR FULL AGONISTS

Chemistry

Zolpidem (Ambien), an imidazopyridine, and zaleplon (Sonata), a pyrazolopyrimidine, are not benzodiazepines in that they lack a diazepine ring (Fig. 34.10–5). Although it is technically accurate to describe these medications as nonbenzodiazepines, they act on benzodiazepine receptors, and their clinical effect can be blocked by the benzodiazepine receptor antagonist flumazenil. However, they are selective for the benzodiazepine-1 receptor and have no significant effect on other benzodiazepine receptor subtypes at usual doses (Table 34.10–3); at higher doses, they are less selective. Zolpidem has an elimination half-life of approximately 2 hours, with a bioavailability of 70 percent; ingestion with food slows absorption. Zolpidem is oxidized to inactive metabolites. Zaleplon is rapidly absorbed and has a half-life of just 1 hour, with a 30 percent bioavailability because of extensive presystemic metabolism. Zaleplon is almost entirely metabolized to inactive metabolites. Alpidem (Ananxyl) is a selective benzodiazepine-1 agonist that is not available in the United States and is used as an anxiolytic.

Table 34.10–3.
Preparations and Doses of Medications Acting on the Benzodiazepine Receptor

Medication	Brand Name	Dose Equivalent	Usual Adult Dose (mg)	How Supplied

Diazepam	Valium	5	2.5–40.0	2-, 5-, and 10-mg tablets 15-mg slow release
Clonazepam	Klonopin	0.25	0.5–4.0	0.5-, 1.0-, and 2.0-mg tablets
Alprazolam	Xanax	0.5	0.5–6.0	0.25-, 0.5-, 1.0-, and 2.0-mg tablet 1.5-mg sustained-release tablet
Lorazepam	Ativan	1	0.5–6.0	0.5-, 1.0-, and 2.0-mg tablets 4 mg/mL parenteral
Oxazepam	Serax	15	15–120	7.5-, 10.0-, 15.0-, and 30.0-mg capsules 15-mg tablet
Chlordiazepoxide	Librium	25	10–100	5-, 10-, and 25-mg capsules and tablets
Clorazepate	Tranxene	7.5	15–60	3.75-, 7.50-, and 15.00-mg tablets 11.25- and 22.50-mg slow-release tablets
Halazepam	Paxipam	20	60–160	20- and 40-mg tablets
Midazolam	Versed	0.25	1–50	5 mg/mL parenteral 1-, 2-, 5-, and 10-mL vials
Flurazepam	Dalmane	15	15–30	15- and 30-mg capsules
Temazepam	Restoril	15	7.5–30.0	7.5-, 15.0-, and 30.0-mg capsules
Triazolam	Halcion	0.125	0.125–0.250	0.125- and 0.250-mg tablets
Estazolam	ProSom	1	1–2	1- and 2-mg tablets
Quazepam	Doral	5	7.5–15.0	7.5- and 15.0-mg tablets
Zolpidem	Ambien	10	5–10	5- and 10-mg tablets
	Ambien CR		6.25–12.5	6.25- and 12.5-mg tablets
	Zolpimist		5–10	5-mg/0.1-mL spray
	Intermezzo		1.75–3.5	1.75- and 3.5-mg sublingual tablet
Zaleplon	Sonata	10	5–20	5- and 10-mg capsules
Eszopiclone	Lunesta	1	1–3	1-, 2- and 3-mg tablets
Flumazenil	Romazicon	0.05	0.2–0.5 per min	0.1 mg/mL 5- and 10-mL vials
Flunitrazepam[a]	Rohypnol	0.5	—	—

[a]Not available in the United States.

The hypnotic drug quazepam (Doral) is selective for the benzodiazepine-1 receptor, but it is metabolized to desalkylflurazepam, a nonselective benzodiazepine metabolite of flurazepam that accounts for 40 percent of total drug activity. With chronic dosing, this medication therefore is similar to other long-acting benzodiazepine hypnotics.

Zopiclone is a benzodiazepine receptor partial agonist that is in widespread use in Europe but is not available in the United States. Its S-enantiomer, eszopiclone (Lunesta), is marketed in the United States. Eszopiclone is extensively metabolized by CYP 3A4 and 2E1; less than 10 percent of an orally administered dose is excreted in the urine, mainly as racemic zopiclone. It does not appear to stimulate or inhibit any CYP450 enzymes.

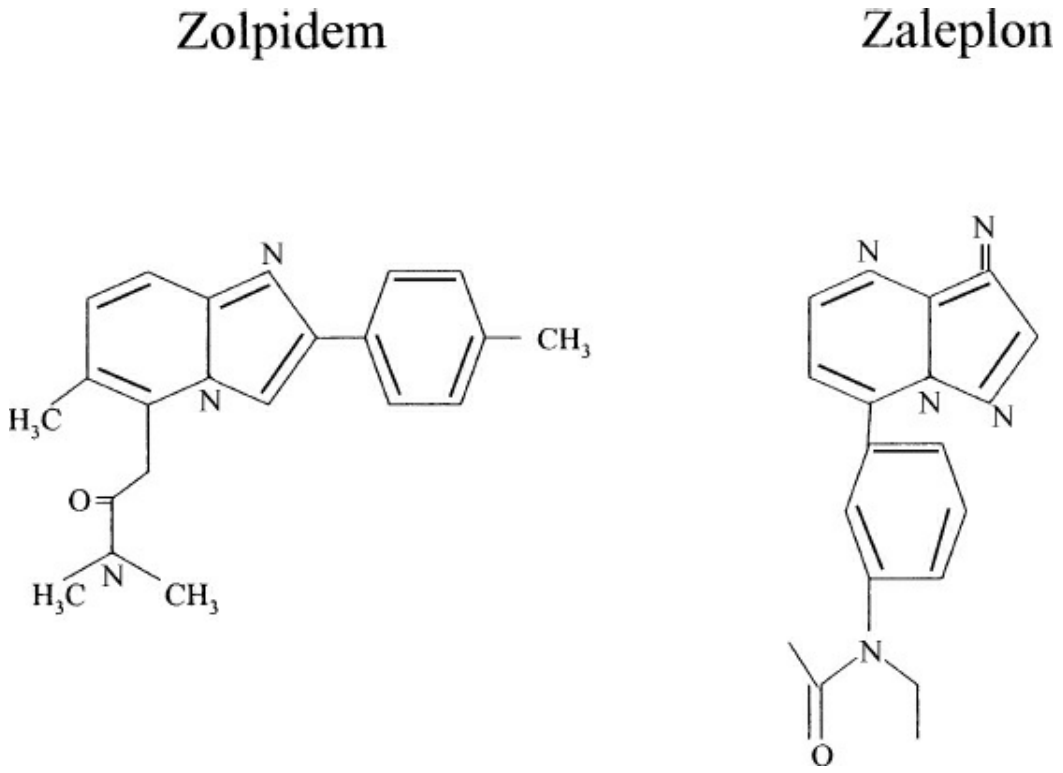

FIGURE 34.10–5. Selective benzodiazepine-1 receptor agonists.

Pharmacological Actions

Zolpidem is rapidly absorbed after oral administration and reaches peak plasma levels in 2 to 3 hours. Zolpidem has an elimination half-life of 2 to 3 hours and has no active metabolites. As a result, in most cases, a bedtime dose is completely eliminated during the night, and there is no daytime hangover. The pharmacology of zolpidem therefore predicts greater usefulness for difficulty falling asleep than for interrupted sleep and early morning awakening. A controlled release preparation (Ambien CR) has a more prolonged duration of action and may help patients stay asleep. Zaleplon, which reaches peak plasma levels 1 to 2 hours after dosing and has an elimination half-life of approximately 1.5 hours, is primarily useful for difficulty falling asleep and usually does not cause a hangover.

Selectivity for the benzodiazepine-1 receptor results in hypnotic and some anxiolytic action without muscle relaxant or anticonvulsant effects for zolpidem and zaleplon. Lack of effects on benzodiazepine-3 receptors appears to result in a lower incidence of withdrawal and rebound symptoms, although these may occur especially at higher doses and with more chronic treatment. Eszopiclone is rapidly absorbed from the GI tract. Its elimination half-life is 5 to 7 hours. About 50 to 60 percent is weakly bound to plasma protein.

Therapeutic Indications

Comparisons of zolpidem and eszopiclone to cognitive behavior therapy

(which includes sleep hygiene) find the latter to be more effective, with more prolonged benefit, than medications. Medications may be more appropriate initial choices for acute insomnia associated with a time limited stress, but behavioral approaches should be considered for chronic insomnia.

The "Z drugs," zolpidem, zaleplon, and eszopiclone, are primarily used to treat insomnia. Zolpidem improves sleep latency and total sleep time to an extent similar to benzodiazepines, without reduction of REM or delta sleep or other changes in sleep architecture. Zolpidem accounts for approximately one-third of all hypnotic prescriptions, having eclipsed benzodiazepines because of its more favorable side-effect profile. An extended release (Ambien CR) and a formulation in which the drug is released longer after ingestion (Intermezzo) have been released for patients with difficulty remaining asleep. A sublingual tablet (Edluar) and oral spray (Zolpimist) are also available. As is true of other hypnotics, it is usually recommended that zolpidem not be administered for more than 2 to 8 weeks. However, one study showed no tolerance during 6 months of treatment.

Zaleplon in a 10-mg dose reduces sleep latency to an extent similar to 0.25 mg of triazolam. Usual doses increase total sleep time by approximately 42 minutes. At higher doses (40 to 60 mg), slow wave sleep increases and REM percentage decreases. Eszopiclone improves sleep efficiency, sleep latency, number of awakenings, total sleep time, and quality and depth of sleep in adults and elderly patients, in whom it also decreases the number of daytime naps. Because the manufacturer chose to conduct 6-month studies, eszopiclone is approved for long-term treatment of insomnia. Sales of this medication have increased substantially as a result of this indication and extensive advertising.

The "Z drugs" are frequently used to treat jet lag and to facilitate shift changes. However, all sleeping pills do is induce sleep; they do not reset the sleep–wake cycle. The time of waking up generally determines the time of falling asleep, and the most effective cue for a sleep–wake cycle that is out of phase with the environment is exposure to bright light at the time of waking up. Therefore, the best approach to preventing jet lag and work shift sleep changes is to take melatonin at the desired time of going to sleep at the new location or shift, and to go outside or have 30 minutes of exposure to artificial bright light at an intensity of 10,000 lux at or near the desired time of waking up.

Precautions and Adverse Reactions

A survey of British primary care physicians found that despite any evidence supporting the belief, the majority believed that the "Z drugs" were safer and more effective than the benzodiazepines. Although tolerance, withdrawal, and rebound insomnia have been found to be rare with

zolpidem, this finding is based on short-term studies. Clinical experience suggests that tolerance can develop, and rebound insomnia can occur after chronic use. Because of its short elimination half-life, zolpidem usually does not produce much daytime impairment. However, memory impairment similar to that seen with triazolam occurs approximately 1.5 hours after dosing, and complex automatisms including driving and eating, during sleep have been reported with amnesia for the behavior, reflecting the peak plasma level of the drug. Adverse effects are more common at doses exceeding 20 mg. Zolpidem is excreted in breast milk and may cause sedation in nursing infants. Concerns about adverse effects and withdrawal have led to reduction in the recommended dose of zolpidem to 5 mg for the immediate release form and 6.25 mg for the extended release form.

In doses up to 10 mg, zaleplon has been found to cause dizziness, headache, and somnolence. Psychomotor impairment has also been reported. However, these side effects are usually not problems during the day because they wear off within 2 to 3 hours as a result of the short half-life of the drug. At doses of 60 mg, psychomotor impairment persists for 25 hours. Abuse potential of zaleplon and zolpidem is low.

In addition to expected side effects such as somnolence and dizziness, eszopiclone can also cause a bitter taste. An Irish study found that almost one-fourth of patients in a methadone maintenance program were abusing zopiclone, often along with benzodiazepines. Abuse of eszopiclone has been reported at doses two to four times the recommended maximum. Although tolerance and rebound insomnia have been rare in clinical trials, rebound insomnia is more common than other discontinuation syndromes with eszopiclone.

Drug Interactions and Laboratory Interferences

Zolpidem, zaleplon, and eszopiclone can have additive sedative effects with other CNS depressants. L-type calcium channel blockers and ritonavir (Norvir) may increase zolpidem effects. No effects on laboratory tests have been described.

Dosage and Administration

The initial dose of zolpidem, sublingual Edluar, and oral spray Zolpimist is 5 mg. The dose can be increased to 10 mg at bedtime. The controlled-release formulation of zolpidem is started at a dose of 6.25 mg, and the dose can be increased to 12.5 mg for males if necessary. The initial dose of Intermezzo is 1.75 mg for women and 3.5 mg for men. Zaleplon is started at a dose of 5 mg, and the dose can be increased to 10 to 20 mg. The short duration of action of zaleplon makes it applicable for patients who have to get back to sleep after waking up in the middle of the night. Eszopiclone is started at a dose of 1 mg, and the dose is increased as needed to 2 to 3 mg at

bedtime.

BENZODIAZEPINE RECEPTOR ANTAGONISTS

Antagonists occupy receptors but have no effect on receptor signaling. They are most frequently used to block the action of a pharmacological agonist, usually a benzodiazepine. The only benzodiazepine antagonist in common use is flumazenil, which has a benzodiazepine structure.

Pharmacological Actions

Flumazenil is only used IV because extensive presystemic metabolism results in low bioavailability after oral dosing, and effective oral doses are too high to tolerate. With IV use, flumazenil has an elimination half-life of approximately 15 minutes. Flumazenil is metabolized to a free acid and a glucuronide, which are excreted in the urine. The drug is completely eliminated within 72 hours. Age and gender do not affect pharmacokinetics, but liver disease slows metabolism and prolongs half-life and action of flumazenil.

Flumazenil reverses CNS depression caused by benzodiazepines. Although alcohol and barbiturates act on the same complex, their primary action is not at the benzodiazepine receptor, and sedation and other signs of CNS depression caused by these substances are not antagonized by flumazenil. The effects of opioids, which act at an entirely different site, are not affected by flumazenil. Emergency treatment of a mixed overdose therefore may require opioid as well as benzodiazepine receptor antagonists.

Therapeutic Indications

Flumazenil is primarily indicated for reversal of benzodiazepine-induced CNS depression caused by intentional overdose or excessive administration during anesthesia. It has also been used experimentally to treat hepatic coma.

Precautions and Adverse Reactions

Common adverse effects include nausea, vomiting, dizziness, agitation, cutaneous vasodilation, headache, agitation, visual problems, memory impairment, and emotional lability. Benzodiazepine withdrawal may be precipitated abruptly, especially in patients who are tolerant to benzodiazepines. When administered to patients taking benzodiazepines, especially if they already have epilepsy, seizures may occur. Seizures can also be precipitated by flumazenil in patients who have had overdoses of TCAs.

Drug Interactions and Laboratory Interferences

The primary interaction of flumazenil is with benzodiazepines. In overdoses that include benzodiazepines and other medications, reversal of sedation from the benzodiazepine may uncover symptoms of intoxication of the other substances. If substances that induce seizures have been ingested along with the benzodiazepine, flumazenil may reverse the anticonvulsant, as well as the depressant, effect of the benzodiazepine, resulting in the emergence of seizures.

No interferences of flumazenil with laboratory tests have been reported.

Dosage and Administration

The starting dose of flumazenil is 0.2 mg (2 mL) given IV over 30 seconds. A 0.3-mg (3-mL) dose is added if the patient does not regain consciousness within 1 minute. This dose can be repeated up to a maximum total dose of 5 mg over approximately 5 minutes. If unconsciousness is not reversed with this dose, it is probably not caused by benzodiazepines, and further treatment with flumazenil is not likely to be of any additional benefit.

INVESTIGATIONAL DRUGS

Abecarnil is a beta-carboline benzodiazepine receptor partial agonist that has anxiolytic and anticonvulsant actions with minimal sedation and psychomotor impairment. This medication appears useful in treating GAD and uncomplicated alcohol withdrawal. Withdrawal symptoms have occurred after 3 to 6 months of treatment with higher doses, but abuse liability appears to be low. Abecarnil has not been released in the United States. Bretazenil, an imidazobenzodiazepine partial benzodiazepine receptor agonist, was found to be effective as monotherapy or adjunctive treatment in reducing symptomatology of severe refractory chronic schizophrenia. The antipsychotic effect was not a function of sedative or antianxiety actions. There was no evidence of withdrawal when the medication was withdrawn.

A benzodiazepine S-enantiomer, (S)-1-(2-fluorophenyl)-3-(2-oxo-5-phenyl-2,3-dihydro-1H-benzo[e][1,4]-diazepin-3-yl)-urea, or 17b (RSV604), was found to inhibit respiratory syncytial virus, a common cause of lower respiratory tract infections, and is now in clinical trials.

Naturally occurring flavones bind to the benzodiazepine receptor, and multiple synthetic flavonoids have been designed to have agonist and inverse agonist effects. The flavonoid 6-bromo-3′-nitroflavone, which has a structure derived from plants used in folk medicine as tranquilizers, is a high-affinity partial benzodiazepine receptor agonist that is ten times as potent as diazepam as an anxiolytic in animal studies. Sedation produced by this preparation occurs at doses much higher than doses that reduce anxiety, suggesting that it could be a well-tolerated antianxiety medication at lower doses. Perhaps because of differential affinities for

benzodiazepine-1 and benzodiazepine-2 receptors, it seems to have minimal muscle relaxant and amnestic effects.

Studies of point mutations of $GABA_A$ receptor subunits are beginning to identify targets for more specific drug action. For example, L838,417 acts on α2, α3, and α5 subunits of the benzodiazepine-$GABA_A$ receptor complex, possibly conveying an antianxiety effect without sedation.

Experimental ligands of the PBR receptor may prove useful as adjuncts in the treatment of illnesses in which enhancement or suppression of apoptosis may be useful. Micromolar concentrations of RO54864 and PK11195 have been found to enhance apoptosis induced by other agonists of the PBR site in a variety of tumor cell lines. Conversely, nanomolar concentrations of the same ligands were found to interfere with apoptosis induced by tamoxifen (Nolvadex) in breast cancer cells with and without estrogen receptors. This discrepancy may reflect differing effects (e.g., agonist vs. inverse agonist) at different concentrations, differential effects in different cancers, or variations in method. However, research is consistent in suggesting that medications that act as antagonists of the PBR may enhance apoptosis and may play an eventual adjunctive role in the treatment and possibly the prevention of cancer. An agonist effect at the PBR could help protect the brain from oxidative damage and could possibly reduce progression of some neurodegenerative diseases.

Indiplon, which is selective for the benzodiazepine-1 receptor, may prove useful for insomnia. Ocinaplon is a full agonist at benzodiazepine-1 receptors and a partial agonist at benzodiazepine-2 and -3 receptors that is in development as a treatment for GAD. It has fewer of the usual sedative side effects, but it may cause hepatotoxicity. ELB139 is a benzodiazepine-3 selective agonist with potential applications in anxiety disorders. Pagoclone is a cyclopyrroline that is a full agonist at benzodiazepine-3 receptors and a partial agonist at benzodiazepine-1 and -2 receptors that is well tolerated and is effective for panic attacks. SGS742 is an experimental $GABA_B$ agonist under study for improvement of cognition in Alzheimer disease and for absence epilepsy. Conversely, based on observations that baclofen causes cognitive impairment, $GABA_B$ antagonists are under development as potential cognitive enhancers. SGS742 is one such compound that is in phase II clinical trials for mild cognitive impairment, and CGP36742 has been found to improve cognitive function in animal studies. Results of phase II studies have not yet been reported for either compound. In human and animal studies, $GABA_B$ agonists have reduced substance craving, suggesting an interesting direction for new drug development.

References

Bellantuono C, Tofani S, Di Sciascio G, Santone G. Benzodiazepine exposure in pregnancy and risk of major malformations: a critical overview. *Gen Hosp Psychiatry*. 2013;35:3.

Billioti de Gage S, Moride Y, Ducruet T, et al. Benzodiazepine use and risk of Alzheimer disease. *Br Med J*. 2014;349:g5205

Brands B, Blake J, Marsh DC, Sproule B, Jeypalan R, Li S. The impact of benzodiazepine use on methadone maintenance treatment outcomes. *J Addictive Disease*. 2008;27:37.

*Casellas P, Galiegue S, Basile AS. Peripheral benzodiazepine receptors and mitochondrial function. *Neurochem Int*. 2002;40:475.

Cook JM, Marshall R, Masci C, Coyne JC. Physicians' perspectives on prescribing benzodiazepines for older adults: A qualitative study. *J Gen Intern Med*. 2007;22:303.

Decaudin D, Castedo M, Nemati F, Beurdeley-Thomas A, de Pinieux G. Peripheral benzodiazepine receptor ligands reverse apoptosis resistance of cancer cells in vitro and in vivo. *Cancer Res*. 2002;62:1388.

Dubois S, Bedard M, Weaver B. The impact of benzodiazepine on safe driving. *Traffic Injury Prev*. 2008;9:404.

*Furukawa TA, Streiner DL, Young LT. Antidepressant and benzodiazepine for major depression (Cochrane Review). *Cochrane Database Syst Rev*. 2002;1:CD110126.

Gallarda T. Paradoxical aggressive reactions to benzodiazepines use: A review. *Encephale*. 2008;34:330.

Gibson RC, Walcott G. Benzodiazepines for catatonia in people with schizophrenia and other serious mental illness. *Cochrane Database Syst Rev*. Issue 4. Art. No.: CD006570. DOI: 10.1002/14651858.CD006570.pub2.

*Gillies D, Beck A, McCloud A, Rathbone J. Benzodiazepines alone or in combination with antipsychotic drugs for acute psychosis. *Cochrane Database Syst Rev*. 2007;3:1.

Katsura M. Functional involvement of cerebral diazepam binding inhibitor (DBI) in the establishment of drug dependence. *Nippon Yakurigaku Zasshi*. 2001;117:159.

Korpi ER, Matilla MJ, Wisden W, Luddens H. GABA(A)-receptor subtypes: Clinical efficacy and selectivity of benzodiazepine site ligands. *Ann Med*. 1997;29:275.

Kriikku P, Hurme H, Wilhelm L, et al. Sedative-hypnotics are widely abused by drivers apprehended for driving under the influence of drugs. *Ther Drug Monit*. 2015;37:339–346.

Lima AR, Soares-Weiser K, Bacaltchuk J, Barnes TR. Benzodiazepines for neuroleptic-induced acute akathisia (Cochrane Review). *Cochrane Database Syst Rev*. 2002;1:CD001950.

Lukens TW, Wolf SJ, Edlow JA, et al. Clinical policy: critical issues in the diagnosis and management of the adult psychiatric patient in the emergency department. *Ann Emerg Med*. 2006;47:79.

Mehta AK, Ticku MK. An update on $GABA_A$ receptors. *Brain Res Rev*. 1999;29:196.

*Moller H-J. Effectiveness and safety of benzodiazepines. *J Clin Psychopharmacol*. 1999;19[Suppl 2]:2S.

Najib J. Eszopiclone, a nonbenzodiazepine sedative-hypnotic agent for the treatment of transient and chronic insomnia. *Clin Ther*. 2006;28:490.

Olfson M, King M, Schoenbaum M. Benzodiazepine use in the United States. *JAMA Psychiatry*. 2014;72(2):1763. doi:10.1001/jamapsychiatry.

Penninga EL, Graudal N, Ladekarl MB, Jurgens G. Adverse effects associated with flumazenil treatment for the management of suspected benzodiazepine intoxication-A systematic review with meta-analysis. *Basic Clin Pharmacol Toxicol*. 2016;118:37–44.

Rudolph U, Knoflach F. Beyond classical benzodiazepines: novel therapeutic potential of GABAA receptor subtypes. *Nat Rev Drug Discov*. 2011;10:685–697.

Rund DA, Ewing JD, Mitzel K, Botolato N. The use of intramuscular benzodiazepines and antipsychotic agents in the treatment of acute agitation or violence in the emergency department. *J Emergency Med*. 2006;31:317.

Saarelainen L, Taipale H, Koponen M, et al. The incidence of benzodiazepine and related drug use in persons with and without Alzheimer's disease. *J Alzheimers Dis*. 2016;49(3):809–818.

Sarter M, Bruno JP, Berntson GG. Psychotogenic properties of benzodiazepine receptor inverse agonists. *Psychopharmacology*. 2001;156:1.

Schweizer E, Rickels K. Benzodiazepine dependence and withdrawal: A review of the syndrome and its clinical management. *Acta Psychiatr Scand*. 1998;393[Suppl]:95.

Shuckit MA: Recognition and management of withdrawal delirium (delirium tremens). *N Engl J Med*. 2014;371:2109.

Strohmmeir R, Roller M, Sanger N, Knecht R, Kuhl H. Modulation of tamoxifen-induced apoptosis by peripheral benzodiazepine receptor ligands in breast cancer. *Biochem Pharmacol*. 2002;64:99.

*Tanaka E. Clinically significant pharmacokinetic drug interactions with benzodiazepines. *J Clin Pharmacol Ther*. 1999;24:347.

Van Vilet P, van der Mast RC, van den Broek M, Westendorp, RG, de Craen AJ. Use of benzodiazepines, depressive symptoms and cognitive function in old age. *Int J Geriatric Psychiatry*. 2009;24:500.

Volz A, Khorsand V, Gillies D, Leucht S. Benzodiazepines for schizophrenia. *Cochrane Database Syst Rev*. 2007;1:CD006391.

Wagner J, Wagner ML, Hening WA. Beyond benzodiazepines: Alternative pharmacologic agents for the treatment of insomnia. *Ann Pharmacother*. 1998;32:680.

Wright BT, Gluszek CF, Heldt SA. The effects of repeated zolpidem treatment on tolerance, withdrawal-like symptoms, and GABAA receptor mRNAs profile expression in mice: comparison with diazepam. *Psychopharmacology*. 2014;231(15):2967–2979.

▲ 34.11 Bupropion

CHARLES DEBATTISTA, M.D., AND ALAN F. SCHATZBERG, M.D.

Originally introduced in the 1980s bupropion (Wellbutrin) is one of the more commonly prescribed antidepressants in the United States. As a result of a unique pharmacology, bupropion is often used as an alternative to SSRIs and serotonin norepinephrine reuptake inhibitors (SNRIs). In addition to its FDA approval in the treatment of depression, bupropion is also approved for smoking cessation and is the only antidepressant to receive FDA approval for the preventive treatment of seasonal affective disorder (SAD). Additionally, bupropion is used for numerous off-label indications including the treatment of ADHD, sexual dysfunction, obesity, and fatigue related to nonpsychiatric medical condition. Currently, all formulations of bupropion are generically available and the major metabolite of bupropion, hydroxybupropion, is being investigated as a potential antidepressant.

PHARMACOLOGY

Pharmacodynamics

The pharmacodynamics of bupropion remains in debate. Although frequently described as a noradrenergic agonist/dopamine reuptake inhibitor, the human in vivo effects of bupropion on dopamine appear rather modest. Animal studies have indicated that bupropion is associated with an increase in nucleus accumbens dopamine concentrations and it appears to inhibit the reuptake of dopamine more potently than do tricyclics. However, human studies have not consistently shown a dopamine effect. For example, some PET studies have shown that bupropion increases occupancy of the dopamine transporter as a dopamine reuptake inhibitor should but others have shown no effect. Likewise, concentrations of homovanillic acid (HVA), a major metabolite of dopamine, do not appear to rise in clinical responders, although they may rise in nonresponders.

The effects of bupropion on noradrenergic activity appear more consistent than its dopamine effects. Bupropion reduces the 24-hour excretion of norepinephrine and its metabolites. Bupropion and its active metabolites, hydroxybupropion, erythrobupropion, and threohydroxybupropion, inhibit the reuptake of norepinephrine in rat cortex, although much less potently than does imipramine. Downregulation of beta-adrenergic receptors and decreased firing of norepinephrine cells in the locus ceruleus have been reported in animal studies. In addition, chronic administration of bupropion appears to modestly inhibit the

tyrosine hydroxylase activity.

Bupropion became the focus of investigation by medicinal chemists in the 1960s when it appeared to be effective in standard screening models of antidepressants but lacked many of the properties of the TCAs and MAOIs that make both classes problematic to use. These include muscarinic side effects, tyramine pressor effects, hypotension, and so forth. Bupropion is not a significant inhibitor of serotonin reuptake, an agonist of 5-HT_{1A} receptors, or an antagonist of 5-HT_2 receptors. The lack of significant serotonergic effects is thought to be the major factor in bupropion's lack of sexual side effects and its general lack of efficacy in anxiety disorders.

Pharmacokinetics

Bupropion is a racemic mixture of enantiomers and is rapidly absorbed. This absorption is not substantially affected by concurrent food of anticholinergic administration. The peak blood level occurs within 2 hours of administration, and bupropion is tightly bound to protein (80 to 90 percent). Protein binding for hydroxybupropion is similar to bupropion, while the other metabolites are about half as tightly bound as bupropion.

Elimination of bupropion is biphasic, consistent with what has been referred to as a "two compartment model," that is, after distribution serum levels are in equilibrium between the intravascular and extravascular compartments. The terminal mean half-life is about 21 hours, while the distribution phase half-life is about 4 hours. Hydroxybupropion may have a half-life as long as 45 hours.

Bupropion undergoes extensive first-pass metabolism with only 1 percent of the drug excreted unchanged in the urine. Most elimination is fecal. Bupropion is a substrate for the CYP450 2B6 isoenzyme, which is not commonly inhibited or induced by other drugs. Oncology agents such as thiotepa (Thioplex) for meningeal neoplasms are among the rare potent inhibitors of 2B6, while barbiturates, such as phenobarbital (Bellatal), may induce the enzyme.

CLINICAL USES

Depressive Disorders

Bupropion has long been known to be an effective antidepressant. It appears to be comparable in efficacy to both TCAs and SSRIs. Most studies with bupropion have involved outpatients with mild to moderate depression. However, inpatient studies with more severely depressed patients have also indicated benefit. Small open-label studies in children and adolescents suggest benefit of bupropion in pediatric depression, but no controlled studies are available. Geriatric depression studies indicate that bupropion in lower doses may be effective and well tolerated. Case reports suggest occasional worsening of psychotic features in patients with

delusional depression.

Unlike most antidepressants, the efficacy of bupropion in treating MDD with concomitant anxiety or anxiety disorders is not well established. Some studies in depressed patients have reported comparable anxiolytic effects of bupropion to SSRIs. However, most studies of bupropion in the treatment of panic disorder have failed to show benefit. Likewise, no controlled trials have been completed that demonstrate clear benefit of bupropion monotherapy for GAD, PTSD, or other anxiety disorders. On the other hand, anxiety may be a consequence of bupropion treatment on occasion.

Among the most common applications of bupropion is in the augmentation of response to other antidepressants. Thus, it is frequently combined with SSRIs, SNRIs, or mirtazapine in patients who have shown inadequate responses to monotherapy. It makes intuitive sense that the combination of a noradrenergic and perhaps dopaminergic molecule such as bupropion might.

Among the largest trials to evaluate the efficacy of bupropion augmentation to date is the Sequenced Treatment Alternatives in Resistant Depression (STAR*D) trial. In that study, patients who had failed citalopram (Celexa) treatment could opt to be randomized to either buspirone (BuSpar) or bupropion augmentation. Both treatments were equally effective with about 30 percent of patients achieving remission after failing citalopram monotherapy. Bupropion was somewhat more effective than buspirone in overall reduction of depressive symptoms. However, the lack of a placebo arm makes it difficult to determine whether either augmenting strategy was more effective than would be longer treatment with citalopram alone.

Given the low remission rates to monotherapy of major depression with antidepressants, attempts have also been made to initiate combined treatment with bupropion and an SSRI as a way of enhancing the speed of onset and remission rates. A randomized controlled study of 245 MDD patients compared the combination treatment with escitalopram and bupropion to monotherapy with either escitalopram or bupropion. No advantage was found for the combination treatment with bupropion to either monotherapy in speed of response or overall remission rates.

Bupropion is commonly used adjunctively with antidepressants to treat SSRI-induced sexual side effects. Again, the rationale for using a dopaminergic agent, even a modest one, to impact sexual side effects appears sound. Stimulants and other dopamine-enhancing agents enhance libido. Although case reports and open studies supported the use of bupropion, several placebo-controlled trials have not demonstrated substantial benefit in depressed patients. Two studies at 150 mg per day of bupropion sustained release (SR) added to an SSRI did not report significant benefit over placebo in treating sexual side effects; one study at 300 mg per day did demonstrate mild efficacy. The benefit of adding

bupropion to a serotonergic antidepressant may be more evident in women than men. A randomized controlled study of 218 women with antidepressant-induced sexual side effects found that 300 mg per day of bupropion added to an existing antidepressant was significantly more effective than placebo in increasing libido, and to a lesser extent, arousal. Switching from a serotonergic antidepressant to bupropion also appears to be an effective strategy to reduce sexual side effects in both men and women.

Bupropion was approved for the preventive treatment of SAD in 2006. In a series of three trials, 1,042 patients who met criteria for SAD were treated in the autumn while they were still well. Patients were randomized to receive either 150 to 300 mg per day of bupropion extended release (XL) or placebo and then followed through the spring. On average, patients had 13 previous seasonal depressive episodes. Patients on placebo were significantly more likely to experience fall/winter depressive episodes than were bupropion-treated patients. Placebo-treated patients were also more likely to have a recurrence of depression sooner than were those on bupropion. To date, no other antidepressants have received an indication in the treatment of SAD.

Bupropion has been favored as an adjunctive treatment for bipolar depression secondary to the perception that it might be less likely to induce switches into mania or more rapid cycling. This perception, however, is based on very few data. Few large prospective studies have evaluated the utility and safety of bupropion in bipolar depression. The largest study to date to evaluate bupropion in bipolar disorder was a part of the NIMH-funded Systematic Treatment Enhancement Program for Bipolar Disorder (STEP BD) study. In that study, patients with bipolar depression on a mood stabilizer were randomized to the addition of bupropion or paroxetine (Paxil) versus the addition of placebo. Patients were treated for 26 weeks and evaluated to determine if they had at least 8 weeks of recovery. Neither bupropion nor paroxetine was significantly better than placebo on achieving a durable recovery from the bipolar depressive episodes. In fact, on some measures placebo-treated patients fared somewhat better. The rate of switches into mania also did not differ between bupropion, paroxetine, and placebo in patients on therapeutic doses of standard mood stabilizers. A study supported by the Stanley Foundation likewise showed relatively low switch rates of bupropion in an acute bipolar trial. The response rate in this study was about 50 percent with about a 10 percent switch rate. Thus, the utility of bupropion in bipolar depression is still uncertain, but switch rates of bupropion when combined with standard mood stabilizers appear low.

Smoking Cessation

Beyond the treatment of depression, the only other FDA indication for

bupropion (SR) is for the promotion of smoking cessation. Early conceptualization about the mechanism for its efficacy in nicotine withdrawal and cravings included bupropion's effects on dopamine reuptake and an influence on reward mechanisms in addictive states. Indeed, there is evidence that dopamine polymorphisms have some predictive value in determining the efficacy of bupropion in smoking cessation. However, more current thinking is that bupropion may have significant, although indirect, effects on nicotine receptors. For example, bupropion may mimic nicotinic effects on dopamine and norepinephrine, which may attenuate withdrawal symptoms. In addition, there is some evidence of direct antagonism of nicotine receptors, which might prevent relapse by diminishing the reinforcing properties of nicotine. In addition, bupropion appears to act as a noncompetitive α3β4 nicotinic antagonist. Despite additional understanding, the mechanism in bupropion for smoking cessation remains unclear, but bupropion appears unlike the nicotine-specific therapies such as the nicotine substitutes (nicotine patch/gum) or the partial nicotine receptor agonist varenicline (Chantix). There is evidence that the major metabolite of bupoprion, hydroxybupropion may mediate the smoking cessation effects of drug. Since there are polymorphisms for the CYP2B6 isoenzyme that is responsible for the metabolism of bupropion to hydroxybupropion, slow metabolizers of CYP2B6 may require higher doses to achieve a serum level of 0.7 mcg/mL as these doses may predict better response to bupropion therapy for smoking cessation.

Overall, bupropion appears to be about twice as effective as placebo in reducing acute nicotine withdrawal symptoms, nicotine cravings, and maintaining abstinence. For example, bupropion has been associated with less mood instability and weight gain after acute withdrawal of tobacco. In one pivotal trial, only 27 percent of patients experienced nicotine cravings 1 month after smoking cessation on bupropion 300 mg per day versus 56 percent of those on placebo. About 30 percent of patients on bupropion are able to quit after 1 to 3 months versus about 15 percent of placebo-treated patients. After 1 year of treatment, patients who remain on bupropion are about 1.5 times as likely to remain abstinent as placebo-treated patients. Bupropion also appears effective in specific populations with nicotine addictions including those with chronic obstructive pulmonary disease (COPD), cardiovascular illness, concurrent alcoholism, schizophrenia, and those with comorbid depression.

Direct comparisons with nicotine-specific therapies are few. Bupropion appears to be at least as effective as nicotine replacement with a nicotine patch, but the combination of bupropion and the nicotine patch may not offer benefit over either therapy alone. For example, the addition of bupropion to a nicotine patch did not offer benefit over the addition of placebo in a 26-week study of smoking cessation in adolescents. At least

two head-to-head comparisons with varenicline tartrate 1 mg twice daily indicated that varenicline may be somewhat more effective in the acute and long-term management of smoking cessation. The combination of varenicline and bupropion may be more effective than either treatment alone.

The usual starting dose of bupropion SR for smoking cession is 150 mg per day, with a target dose of 300 mg per day within 7 days. An acute course of bupropion tends to last 8 to 12 weeks with tobacco stopped within the first 2 to 3 weeks of treatment. Responders may continue bupropion treatment chronically to prevent relapse.

Obesity and Eating Disorders. The FDA approved the combination of bupropion and naltrexone for the treatment of obesity in 2014. Four large registration trials involving over 4,500 patients were completed in overweight BMI >27 to obese BMI >30 patients (mean BMI = 36). The general design of these studies was 3 weeks of dose titration followed by approximately 1 year of treatment. Subjects were randomized to the combination of naltrexone (15 to 60 mg) and bupropion (300 to 400 mg) along with dietary modifications (e.g., reducing daily calories by 500 kcal/day), behavioral counseling, and an increase in exercise or physical activity versus the control group included with these interventions with a placebo instead of the active drug. In all four trials, subjects randomized to the active group were more likely to lose clinically meaningful amount of weight defined as >5 percent or >10 percent of total body weight. In addition, subjects on the combination of bupropion (mean dose 360 mg per day) and naltrexone (mean dose 32 mg per day) lost about 5 percent more of total body weight versus subjects in the control group.

A large percentage of patients in both the active and control groups of the naltrexone/bupropion obesity studies dropped out before completing the 1 year end point. In fact, approximately 45 percent of both groups dropped out early with subjects in the active group twice as likely to drop out as subjects in the control group (24 percent vs. 12 percent). The most common side effects leading to discontinuation of the combination treatment were nausea, headache, and vomiting. Other common side effects of the combination of naltrexone/bupropion constipation,

Bupropion alone also has mild anorexiant effects that have been evaluated in nondepressed obese patients. Approximately 25 percent of patients in acute depression trials lost 5 lb or more while taking bupropion. Several controlled studies have evaluated 300 to 400 mg per day of bupropion in the treatment of obesity. These studies have indicated that bupropion-treated patients tend to lose 1.1 to 4.5 kg at 6 to 12 months, and this is significantly greater than placebo-treated patients. Bupropion has also been investigated in the treatment of binge eating disorder. The largest randomized, placebo-controlled trial to date (n = 61) did not find the

bupropion to reduce binging or to improve weight loss in women with the disorder.

Other Clinical Uses

In addition to the FDA-approved indications in the treatment of major depression and for smoking cessation, bupropion has been clinically tested for a number of other conditions. For example, as with other noradrenergic antidepressants such as nortriptyline (Aventyl), bupropion may be useful in the treatment of adult ADHD. In pediatric populations, the only large controlled trial did not show significant benefit of bupropion on the primary outcome measures. However, in this trial, teachers rated the bupropion-treated children as less aggressive and hyperactive relative to placebo-treated children. Clinicians and parents did not observe a significant difference. There a few direct comparisons of bupropion with stimulants; however, these studies revealed that bupropion is either about as effective or somewhat less effective. A small randomized trial (n = 44) comparison of bupropion and methylphenidate found that both drugs were equally effective and well tolerated. Studies of bupropion in adults with ADHD have been more consistently positive. For example, in a study of 162 adults with ADHD 53 percent of patients responded to bupropion XL versus 31 percent of placebo-treated patients. Benefits over placebo were noted as early as week 2, and bupropion, with an average dose of 450 mg per day, was well tolerated.

Finally, bupropion has been suggested to be of benefit in a number of small studies and case series for a broad spectrum of conditions. These have included sexual dysfunction in nondepressed patients, fatigue in cancer patients, the treatment of amphetamine addiction, lowering tumor necrosis factor-α, and improving concentration in patients without ADHD.

SIDE EFFECTS

The pharmacodynamics effects of bupropion result in a side-effect profile different from most antidepressants. Many of the adverse events reported with bupropion are associated with its noradrenergic, and to a lesser extent, dopaminergic effects. For example, CNS activation is associated with bupropion treatment. Insomnia is one of the common side effects of bupropion, affecting at least 11 percent of patients in clinical trials. Anxiety and agitation may occur as treatment emergent effects. At least 6 percent of patients in clinical trials also experienced a mild to moderate tremor. Treatment emergent psychotic symptoms have rarely been reported in depressed patients.

The risk of seizure with the immediate release (IR) formulation of bupropion became evident in studies of eating disorder patients initially and lead to withdrawal of the drug in 1986. Patients in the eating disorder

trials were treated with up to 600 mg per day of bupropion. Analysis of data sets revealed an enhanced risk of seizure in patients with a history of eating disorders, head trauma, or previous seizure disorder. Also, doses above 450 mg per day or 200 mg at one time (for the IR formulation) enhanced the risk of seizure. For extended-release formulations at doses of 450 mg per day or less, the seizure risk is estimated at 0.1 to 0.2 percent, which is comparable to the rates reported with SSRIs and may be less than the rates reported with some TCAs and maprotiline (Novo-Maprotiline).

Other common side effects reported in trials included nausea, dry mouth, excessive sweating, tinnitus, and rash. The most common reasons patients discontinued bupropion in acute clinical trials in depression were nausea and rash.

Bupropion has a favorable cardiovascular profile with few cardiac effects. Bupropion is not known to have clinically significant effects on blood pressure, heart rate, or cardiac conduction. Thus, bupropion is commonly used in patients with known cardiac pathology. However, there have been rare reports of significant hypertension associated with bupropion in patients with and without a pre-existing history of high blood pressure.

Side effects common to many antidepressants (including sexual side effects and weight gain) appear uncommon with bupropion. As noted earlier, bupropion has been associated with weight loss, and there are many reports of improved sexual functioning, including increased libido and arousal, with the drug.

Other side effects such as hepatoxicity, hematopoietic changes, and severe headaches occur rarely in bupropion-treated patients. Thus, the side-effect profile makes bupropion an important alternative to many patients who cannot tolerate other antidepressants.

DRUG INTERACTIONS

The limited cytochrome P450 interactions of bupropion contribute to its few drug interactions. Since bupropion and hydroxybupropion are metabolized via the cytochrome P2B6, there is a theoretical risk of an interaction between oncology agents such as IV thiotepa and the muscle relaxant orphenadrine (Norflex). Hydroxybupropion is a weak to moderate inhibitor of cytochrome P450 2D6 and has been associated with increasing serum levels of some 2D6-dependent drugs such as venlafaxine (Effexor) and desipramine (Norpramin). Rare seizures that have occurred in combination of fluoxetine (Prozac) and bupropion have sometimes been attributed to inhibition of the metabolism of hydroxybupropion by 2D6 inhibitors, but seizures can occur independent of these potential interactions, so it is unclear whether this mechanism was at play.

Since bupropion may lower the seizure threshold, particularly when the IR form is used in higher doses, combination with other drugs that lower

the seizure threshold, such as some tricyclics, amoxapine (Asendin), clozapine (Clozaril), and theophylline (Elixophyllin), have sometimes been associated with seizures.

As with other antidepressants, the concurrent administration with MAOIs is contraindicated. However, there have been few reports in the literature of serious interactions with MAOIs. Oral selegiline (Carbex) at doses of 10 mg per day or less have sometimes been combined with bupropion without incident.

OVERDOSE

Bupropion has rarely been associated with fatalities in monotherapy overdoses. Overdoses of bupropion as low as 10 g have sometimes been associated with deaths. Causes of death in these overdoses have included complications of status epilepticus and arrhythmias. Although therapeutic doses of bupropion are rarely associated with cardiovascular effects, overdoses have sometimes been associated with tachycardia, intraventricular conduction delays, and QRS complex widening. Treatment for overdoses includes lavage, seizure treatment or prophylaxis, cardiac monitoring, and vital sign support. Most patients who overdose on bupropion recover.

DOSING AND FORMULATIONS

Bupropion hydrochloride is available in IR, SR, and XL forms. As of 2006, all of these formulations had become available as generics in the United States. The usual starting dose in all formulations is 150 mg per day taken in the morning, but the IR form was often started at 100 mg twice a day. The average therapeutic dose of bupropion for the treatment of depression is 300 to 450 mg per day. Doses above 200 mg at once are associated with greater seizure risk in the IR form. Thus, at maximum doses, compliance to bupropion IR prescriptions was somewhat difficult when dosed at 150 mg three times a day. In contrast, the typical therapeutic dose of the SR form was 150 or 200 mg twice a day. The XL form could be taken all in one morning dose, which was said to have more consistent blood levels and perhaps less sleep disruption. A cellulose coating in XL formulation allowed for slow absorption.

In smoking cessation, bupropion SR is typically started at 150 mg per day and then increased to the target dose of 300 mg per day.

REFERENCES

Caixàs A, Albert L, Capel I, Rigla M. Naltrexone sustained-release/bupropion sustained-release for the management of obesity: review of the data to date. *Drug Des Devel Ther*. 2014;8:1419–1427.

Cinciripini PM, Robinson JD, Karam-Hage M, et al. Effects of varenicline and bupropion

sustained-release use plus intensive smoking cessation counseling on prolonged abstinence from smoking and on depression, negative affect, and other symptoms of nicotine withdrawal. *JAMA Psychiatry*. 2013;70(5):522–533.

Clayton AH, Croft HA, Horrigan JP, et al. Bupropion extended release compared with escitalopram: Effects on sexual functioning and antidepressant efficacy in 2 randomized, double-blind, placebo-controlled studies. *J Clin Psychiatry*. 2006;67(5):736–746.

Clayton AH, Montejo AL. Major depressive disorder, antidepressants, and sexual dysfunction. *J Clin Psychiatry*. 2006;67[Suppl 6]:33–37.

Davidson JR, Connor KM. Bupropion sustained release: A therapeutic overview. *J Clin Psychiatry*. 1998;59[Suppl 4]:25–31.

DeBattista C. Augmentation and combination strategies for depression. *J Psychopharmacol*. 2006;20[3 Suppl]:11–18.

DeBattista C, Solvason B, Poirier J, Kendrick E, Loraas E. A placebo-controlled, randomized, double-blind study of adjunctive bupropion sustained release in the treatment of SSRI-induced sexual dysfunction. *J Clin Psychiatry*. 2005;66(7):844–848.

DeBattista C, Solvason HB, Poirier J, Kendrick E, Schatzberg AF. A prospective trial of bupropion SR augmentation of partial and non-responders to serotonergic antidepressants. *J Clin Psychopharmacol*. 2003;23(1):27–30.

Ebbert JO, Hatsukami DK, Croghan IT, et al. Combination varenicline and bupropion SR for tobacco-dependence treatment in cigarette smokers: a randomized trial. *JAMA*. 20148;311(2):155–163.

Foley KF, DeSanty KP, Kast RE. Bupropion: Pharmacology and therapeutic applications. *Expert Rev Neurother*. 2006;6(9):1249–1265.

Friedman ES, Davis LL, Zisook S, et al. Baseline depression severity as a predictor of single and combination antidepressant treatment outcome: results from the CO-MED trial. *Eur Neuropsychopharmacol*. 2012;22(3):183–199.

Jefferson JW, Pradko JF, Muir KT. Bupropion for major depressive disorder: Pharmacokinetic and formulation considerations. *Clin Ther*. 2005;27(11):1685–1695.

Killen JD, Robinson TN, Ammerman S, Hayward C, Rogers J. Randomized clinical trial of the efficacy of bupropion combined with nicotine patch in the treatment of adolescent smokers. *J Consult Clin Psychol*. 2004;72(4):729–735.

Kuperman S, Perry PJ, Gaffney GR, et al. Bupropion SR vs. methylphenidate vs. placebo for attention deficit hyperactivity disorder in adults. *Ann Clin Psychiatry*. 2001;13(3):129–134.

Lee MS, Yang JW, Ko YH, et al. Effects of methylphenidate and bupropion on DHEA-S and cortisol plasma levels in attention-deficit hyperactivity disorder. *Child Psychiatry Hum Dev*. 2008;39(2):201–209.

Levin FR, Evans SM, Brooks DJ, Kalbag AS, Garawi F, Nunes EV. Treatment of methadone-maintained patients with adult ADHD: Double-blind comparison of methylphenidate, bupropion and placebo. *Drug Alcohol Depend*. 2006;81(2):137–148.

Modell JG, Rosenthal NE, Harriett AE, et al. Seasonal affective disorder and its prevention by anticipatory treatment with bupropion XL. *Biol Psychiatry*. 2005;58(8):658–667.

Morris DW, Budhwar N, Husain M, et al. Depression treatment in patients with general medical conditions: results from the CO-MED trial. *Ann Fam Med*. 2012;10(1):23–33.

Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR*D report. *Am J Psychiatry*. 2006;163(11):1905–1917.

Rush AJ, Trivedi MH, Wisniewski SR, et al. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. *N Engl J Med*. 2006;354(12):1231–1242.

Sachs GS, Nierenberg AA, Calabrese JR, et al. Effectiveness of adjunctive antidepressant treatment for bipolar depression. *N Engl J Med*. 2007;356(17):1711–1722.

Solhkhah R, Wilens TE, Daly J, Prince JB, Van Patten SL, Biederman J. Bupropion SR for the treatment of substance-abusing outpatient adolescents with attention-deficit/hyperactivity disorder and mood disorders. *J Child Adolesc Psychopharmacol*. 2005;15(5):777–786.

Taylor MJ. Strategies for managing antidepressant-induced sexual dysfunction: a review. *Curr Psychiatry Rep*. 2006;8(6):431–436.

Thase ME, Friedman ES, Biggs MM, et al. Cognitive therapy versus medication in augmentation and switch strategies as second-step treatments: A STAR*D report. *Am J Psychiatry*. 2007;164(5):739–752.

Trivedi MH, Fava M, Wisniewski SR, et al. Medication augmentation after the failure of SSRIs for depression. *N Engl J Med*. 2006;354(12):1243–1252.

Upadhyaya HP, Brady KT, Wang W. Bupropion SR in adolescents with comorbid ADHD and nicotine dependence: A pilot study. *J Am Acad Child Adolesc Psychiatry*. 2004;43(2):199–205.

White MA, Grilo CM. Bupropion for overweight women with binge-eating disorder: a randomized, double-blind, placebo-controlled trial. *J Clin Psychiatry*. 2013;74(4):400–406.

Wilens TE, Haight BR, Horrigan JP, et al. Bupropion XL in adults with attention-deficit/hyperactivity disorder: a randomized, placebo-controlled study. *Biol Psychiatry*. 2005;57(7):793–801.

Wilens TE, Hammerness PG, Biederman J, et al. Blood pressure changes associated with medication treatment of adults with attention-deficit/hyperactivity disorder. *J Clin Psychiatry*. 2005;66(2):253–259.

Wilens TE, Spencer TJ, Biederman J, et al. A controlled clinical trial of bupropion for attention deficit hyperactivity disorder in adults. *Am J Psychiatry*. 2001;158(2):282–288.

Zisook S, Rush AJ, Haight BR, Clines DC, Rockett CB. Use of bupropion in combination with serotonin reuptake inhibitors. *Biol Psychiatry*. 2006;59(3):203–210.

▲ 34.12 Buspirone

MOLLY M. MCCARTHY, D.O., AND REGINA M. CARNEY, M.D.

INTRODUCTION AND HISTORY

The azapirones are a unique pharmacological class of medications that includes buspirone, an azaspirodecanedione derivative, and its second-generation compound gepirone. Two other related compounds are ipsapirone and tandospirone. Since its synthesis in 1968 in the laboratories

of Mead Johnson, this class of medications has been studied for its potential antidepressant and anxiolytic effects. Although it was initially developed as an antipsychotic drug thought to act primarily on dopamine receptors, buspirone was later found to exert its clinical effects by working on the serotonergic system, specifically on the serotonin 1A ($5\text{-}HT_{1A}$) receptor subtype. Studies conducted in the 1970s revealed that a single dose of buspirone improved aggressive behavior in rhesus monkeys. Furthermore, buspirone showed marked anxiolytic effects in behavioral models of anxiety in rodents in the early 1980s. Eventually, a phase II study in patients with DSM-II anxiety disorder showed improvement in symptoms with treatment of buspirone as compared with placebo, leading to FDA approval for generalized anxiety disorder (GAD) in 1986. By that time, the newest DSM classification system, DSM-III, had replaced anxiety neurosis with the diagnostic category of GAD. FDA approval for the treatment of GAD with buspirone reflected this change. This drug was the only nonbenzodiazepine medication approved by the FDA for treatment of the disorder. Other potential clinical uses have since been explored, such as treatment for other anxiety disorders (e.g., social anxiety disorder, OCD, PTSD, panic disorder), depressive disorders, eating disorders, and attention-deficit/hyperactivity disorder (ADHD). To date, buspirone has been indicated for the treatment of GAD and anxiety associated with depressive symptoms.

PHARMACOLOGICAL ACTIONS

Buspirone hydrochloride (BuSpar) belongs to a group of compounds known as the azaspirodecanediones and has been shown to act on two types of receptors: Serotonin (5-HT) and dopamine (D). It has high affinity for the $5\text{-}HT_{1A}$ serotonin receptor, acting as an agonist or partial agonist, and moderate affinity for the D_2-dopamine receptor, acting as both an agonist and antagonist. The approved indication for this psychotropic drug is for the treatment of GAD. It was initially believed to be a better alternative to the benzodiazepine drug group since buspirone does not exert strong sedative effects and does not possess the same potential for dependence. However, it is more currently believed that although the azapirones, including buspirone, are better tolerated than benzodiazepines, combined evidence reveals that they are not more effective.

PHARMACOKINETICS

Absorption

Buspirone undergoes extensive first-pass metabolism when ingested orally, with an absolute bioavailability of 4 percent. A 20-mg dose of buspirone led to peak plasma concentrations of 1 to 6 μg/L (C_{max}) from 0.7 to 1.5 hours

following administration (T_{max}). Studies in humans have shown that buspirone exhibits linear pharmacokinetics over the 10- to 40-mg dose range. Outside of this dosing range, buspirone no longer holds this relationship. Taking buspirone with food affects bioavailability, increasing the area under the plasma concentration-time curve (AUC) by 84 percent and the C_{max} by 116 percent. Patients with hepatic impairment, defined by the Child–Pugh classification system, had a 13-fold increase in AUC.

Distribution

Following first-pass metabolism, the remaining drug enters the systemic circulation. The volume of distribution is 5.3 L/kg following intravenous (IV) administration. Buspirone and its major active metabolites 6-OH-buspirone and 1-pyrimidinylpiperazine (1-PP), all readily cross the BBB. A recent in vivo study demonstrated that its active metabolite 6-OH-buspirone localized in the hippocampus and the dorsal raphe nucleus. Buspirone and its metabolites have also been reported in the lungs and kidneys. In vitro studies have indicated that 86 to 95 percent of the drug is bound to plasma protein. In vivo studies in rats have indicated that both buspirone and its metabolites are found in maternal milk and cross the placenta, although it is not known to what extent both these events occur in humans.

METABOLISM AND ELIMINATION

Buspirone is metabolized through oxidation by cytochrome P450 3A4 (CYP3A4). CYP3A4, the most abundant form of cytochrome P450 in the human adult liver, is polymorphic and thus shows wide interindividual variation in its activity. Main metabolic pathways include hydroxylations on the 6′-position on the spiro ring system (6-OH buspirone) and the 5′-position on the pyrimidine ring of the drug complex (5-OH buspirone), as well as a *N*-dealkylation of a butyl side group, forming 1-PP. Approximately 26 percent of buspirone was found to be converted to 1-PP, and this metabolite had a 49-fold greater AUC than that of buspirone but conversely had 25 percent of the potency of the parent compound.

Another major active metabolite is 6-OH-hydroxybuspirone. This metabolite was found to have 12 times greater mean AUC in the brain compared to the parent compound. Similar to the parent compound, 6-OH-hydroxybuspirone had four times greater affinity for presynaptic 5-HT_{1A} receptors in the dorsal raphe nucleus compared to postsynaptic 5-HT_{1A} located in the hippocampus. As expected, given that CYP3A4 is polymorphic, the mean half-life ($t_{1/2}$) of buspirone varied greatly between healthy individuals, with a minimum of 2 to 3 hours and a maximum of 11 hours after a single dose of 10 to 40 mg. The mean half-life of 1-PP was 6 hours following a 20-mg dose. In subjects with renal impairment, defined

as improper creatinine clearance, the half-life of buspirone increased, on average, by 2 hours.

PHARMACODYNAMICS

Mechanism of Action

Buspirone has a unique mechanism of action from other anxiolytics, such as the widely used benzodiazepines, as it does not bind to the GABA receptor. Although the exact mechanism of action is not fully understood, it is thought to involve a combination of systems, mainly, the serotonergic and dopaminergic systems.

Serotonin System

Buspirone functions as a full agonist at presynaptic 5-HT_{1A} autoreceptors, as a partial agonist at postsynaptic 5-HT_{1A} receptors, and as a weak antagonist at 5-HT_{2C} receptors. The activation of somatodendritic 5-HT_{1A} autoreceptors in the midbrain raphe nuclei has been demonstrated to decrease the synthesis of serotonin (5-HT) in rats. This is believed to be the mechanism of action that may explain buspirone's therapeutic effect. Moreover, activation of postsynaptic 5-HT_{1A} receptors located in the hippocampus resulted in decreased neuron excitability, which could explain the antidepressant effects of buspirone. A combination of buspirone and the SSRI citalopram (Celexa) was observed to cause a downregulation in 5-HT_{2A} and 5-HT_{2C} and may be the source of the effectiveness of buspirone in augmentation therapy for nonresponders to SSRI treatment.

Dopaminergic System

The dopaminergic system is another site of action of buspirone. Buspirone has been categorized as both an agonist and antagonist to the D_2 subtype of the dopamine receptor. Buspirone acts as an antagonist to presynaptic dopamine antagonist autoreceptors, but some evidence has also pointed to the involvement of postsynaptic DA receptors. The activation of these receptors increased dopamine antagonist synthesis and metabolism in the striatum and nucleus accumbens. Other studies have suggested that buspirone can function as a full dopamine agonist and could also antagonize D_1-dopamine receptors.

GABAergic System

As previously mentioned, buspirone has no affinity for the benzodiazepine-binding site on the GABA receptor, nor does it affect the binding of this drug group or the binding of the GABA receptor to its natural ligand.

Despite this fact, there has been some suggestion that buspirone can indirectly affect the GABAergic system, as buspirone was found to increase the level of binding of flunitrazepam in an in vivo model. Further studies are needed to explore the implications of this effect.

Norepinephrine System

Administration of buspirone in humans has been found to interact with both the α_1-adrenergic and β-adrenergic systems, with plasma levels of noradrenaline increased significantly along with a decrease in systolic blood pressure and heart rate. One of the major metabolites of buspirone, 1-PP, has antagonistic properties to α_2-adrenergic receptors.

Cholinergic System

Binding evidence has shown that buspirone has no binding affinity for muscarinic cholinergic receptors, but some evidence indicates the possibility of an indirect relationship with this system. Administration of buspirone caused an efflux of acetylcholine in the rat hippocampus, which led to a modest increase in motor activity and was observed to have a trophic effect on septal cholinergic-containing neurons.

Blood Concentration and Relation to Action

There is no evidence tightly linking plasma concentrations to the therapeutic effects of buspirone.

CLINICAL TRIALS

Generalized Anxiety Disorder

Several double-blind, randomized, placebo-controlled trials have established the anxiolytic efficacy of buspirone. Most monotherapy studies have focused on the treatment of GAD, with the majority of trials finding buspirone to be significantly associated with greater reduction in anxiety symptoms compared to placebo. Standard outcome measures of severity of anxiety symptoms (Hamilton Anxiety [HAM-A] scale) and global improvement (Clinical Global Impressions [CGI] scale) were generally employed. A recent clinical trial examined the use of buspirone or placebo in conjunction with anxiety management therapy (AMT) or nondirective therapy (NDT). All four treatment groups exhibited a significant reduction in anxiety symptoms, but there was no significant difference between any of these groups. Although each of these trials required a diagnosis of GAD, it is worthwhile to note that the definition of GAD has shifted over time, with the required duration of anxiety symptoms lengthened from 1 month (third edition of *Diagnostic and Statistical Manual of Mental Disorders* [DSM-III]) to 6 months (DSM-III-R, DSM-IV). Early buspirone clinical

trials used DSM-III criteria for GAD, while more recent clinical trials utilized DSM-III-R or DSM-IV criteria for GAD. It is possible that those initial studies may have had many subjects with an anxiety disorder that would not have been sufficiently chronic to be currently classified as GAD. Early studies found buspirone to be statistically inferior or equivalent to benzodiazepines in the acute reduction of anxiety symptoms in GAD. One clinical trial found buspirone to have significantly less anxiolytic efficacy than either lorazepam (Ativan) or alprazolam (Xanax). However, another study did not find any statistical difference between buspirone and alprazolam with respect to reduction of anxiety symptoms and the same conclusion was reached in a trial between buspirone and diazepam (Valium). It has also been demonstrated that recent benzodiazepine treatment negatively predicts treatment response to buspirone. This study found no significant difference in response to buspirone or diazepam in patients with no prior exposure to benzodiazepines, but that diazepam was significantly more effective than buspirone at reducing anxiety symptoms in patients with recent exposure to benzodiazepines. These buspirone versus benzodiazepines studies all made use of the DSM-III criteria for GAD. There have been no recent studies comparing buspirone to benzodiazepines. A recent study found no statistical difference between buspirone and venlafaxine XR (Effexor XR) in reducing anxiety symptoms in GAD. Serotonergic antidepressants are often used as first-line treatments for GAD, and the paucity of studies comparing buspirone to SSRI antidepressants is a large gap in the available literature. Further trials are required to compare the use of standard treatment for GAD (i.e., SSRIs) with buspirone to adequately understand the role of buspirone for this indication. An interesting recent study found no statistical difference between buspirone and kava-kava with respect to reducing anxiety symptoms in GAD. Kava-kava is an herbal medication with constituents that putatively act both as a monoamine oxidase B (MAO-B) inhibitor and as a norepinephrine reuptake inhibitor. Another trial found buspirone to be equivalent to hydroxyzine (Vistaril) in reducing anxiety symptoms in GAD. Hydroxyzine is a first-generation antihistamine of the piperazine class that acts via antagonism of the H_1 receptor. Further data comparing buspirone to other alternative pharmacological and nonpharmacological interventions are warranted.

Other Anxiety Disorders

Several clinical trials have also examined the role of buspirone in the treatment of other anxiety disorders such as panic disorder OCD, and PTSD. Double-blind, randomized, placebo-controlled trials found that buspirone was not significantly superior to placebo in treating patients with panic disorder. In addition, one of these trials also found alprazolam to be

superior to buspirone and placebo with respect to its antipanic and anxiolytic effects. With respect to PTSD, smaller, open-label trials found buspirone to be effective in alleviating symptoms such as anxiety, insomnia, flashbacks, and depressive state. In one of these studies, seven of eight patients exhibited a significant reduction in symptoms as measured by the Structured Inventory for PTSD (SI-PTSD) and the Beck Depression Inventory (BDI) after usage of buspirone as a monotherapy. In another trial, 12 combat-related PTSD patients exhibited a 40 percent reduction in reexperiencing symptoms after usage of buspirone as an adjunct therapy.

The promising results from these smaller, open-label trials highlight the need for larger, placebo-controlled studies. Several double-blind, randomized, placebo-controlled trials have evaluated buspirone augmentation of antidepressant treatment for patients with OCD. These trials found no significant improvement in Yale-Brown Obsessive Compulsive Scale (Y-BOCS) scores associated with the addition of buspirone. Two trials added buspirone to SSRI therapy (fluoxetine [Prozac] and fluvoxamine [Luvox]), and another used buspirone as an adjuvant therapy to the TCA clomipramine (Anafranil). The use of buspirone as adjunct treatment for OCD is therefore not supported by the current evidence.

Major Depressive Disorder

Several large, double-blind, randomized, placebo-controlled trials have examined the efficacy of buspirone in treating patients with MDD. These clinical trials have generally shown buspirone to be significantly superior to placebo in reducing depressive symptoms. One study compared buspirone to placebo as adjunct therapy to citalopram in patients who were treatment refractory to this SSRI. This trial found buspirone to be significantly more effective at reducing depressive symptoms in these patients than placebo. Most of these trials used the HAM-D scale as the standard outcome measure of severity of depressive symptoms, although the Montgomery-Asberg Depression Rating Scale (MADRAS) was also employed in this capacity. More recent studies have evaluated buspirone as augmentation to SSRI treatment for MDD. The NIMH-sponsored Sequenced Treatment Alternatives in Resistant Depression (STAR*D) trial included an arm of the study in which patients who did not remit during a treatment course with citalopram were treated with the addition of either bupropion sustained release (SR) (Wellbutrin XL) or buspirone. Although there was no significant difference between response to buspirone and bupropion SR, both adjuvant therapies exhibited a clinically meaningful reduction in HAM-D scores in these treatment-refractory MDD patients, with remission rates of 30.1 and 29.7 percent, respectively. Table 34.12–1 provides a summary of clinical trials of buspirone.

Other Uses

Alcohol and Cannabis Dependence. Buspirone usage has also been examined in alcoholism. One clinical trial found that while buspirone reduced anxiety symptoms in male veterans, this change was not significantly different from the reductions associated with placebo. In addition, that study also found buspirone to be statistically equivalent to placebo with respect to standard measures of alcohol consumption, such as time to first drink and number of standard drinks per drinking day. Similarly, a meta-analysis of five double-blind, RCTs concluded that buspirone was not statistically significantly superior to placebo in reducing alcohol consumption. That report did, however, highlight the positive effects of buspirone on treatment retention and anxiety. The authors suggested that while buspirone did not affect alcohol consumption, it may mediate associated psychopathological symptoms. Given its favorable side-effect profile and lack of interaction with alcohol, buspirone may have a role as an adjunct therapy in the treatment of comorbid anxiety and alcoholism.

Treatment of cannabis (marijuana) dependence with buspirone was compared to placebo in a single study of 50 individuals titrated to a target dose of 60 mg daily and treated for 12 weeks. There were no significant differences between the buspirone-treated individuals and those receiving placebo on self-reported and urine drug detection measures of cannabis use. Secondary outcome measures (time to first negative urine drug screen result, summary measures of self-reported use, and measures of marijuana craving/withdrawal symptoms and anxiety) suggested a trend toward faster time to first negative urine drug screen and a reduction in overall positive urine drug screens for those individuals in the buspirone-treated group.

Smoking Cessation. The efficacy of buspirone as a means of promoting smoking cessation has been evaluated in open-label trials and a small number of RCTs. The evidence of efficacy remains inconclusive. A comparison between buspirone and nicotine transdermal patches found both treatments had similar cessation rates, while another study looking exclusively at buspirone found that the greatest improvement was observed in patients who had high anxiety. In a more recent study, buspirone coadministered with sertraline (Zoloft) was found to lead to improved abstinence rates (43.5 percent for the drug group compared to 17.3 percent for the control) at 26 weeks following smoking cessation (n = 98).

Table 34.12–1.
Buspirone Clinical Trials

	Number of Subjects	Duration	Dose	Significant Difference vs. Comparator

Generalized Anxiety Disorder (GAD)

Buspirone vs. Placebo				
Ross and Matas, 1987	21	4 weeks	5–40 mg	NS
Enkelmann, 1991	62	6 weeks	18.7 mg (mean)	+
Gammans et al., 1992	Meta-analysis			+
Chiaie et al., 1995	44	8 weeks	15 mg	+
Sramek et al., 1996	162	6 weeks	5–45 mg	+
Lader and Scotto, 1998	163	4 weeks	20 mg	+
DeMartinis et al., 2000	487	4 weeks	30 mg (up to)	+
Bond et al., 2002	44	8 weeks	30 mg (up to)	+
Majercsik et al., 2003	52	6 weeks	30 mg	+
Buspirone vs. Benzodiazepine				
Cohn et al., 1986 (lorazepam)	40	4 weeks	15–50 mg	—
Cohn et al., 1986 (alprazolam)	40	4 weeks	15–50 mg	—
Ross and Matas, 1987 (diazepam)	19	4 weeks	5–40 mg	NS
Enkelmann, 1991 (alprazolam)	63	6 weeks	18.7 mg	NS
DeMartinis et al., 2000 (diazepam)	483	4 weeks	30 mg (up to)	—
Buspirone vs. Hydroxyzine				
Lader and Scotto, 1998	162	4 weeks	20 mg	NS
Buspirone vs. Venlafaxine XR				
Davidson et al., 1999	267	8 weeks	30 mg	NS
Buspirone vs. Kava-Kava				
Boerner et al., 2003	85	8 weeks	10 mg	NS
Panic Disorder				
Buspirone vs. Placebo				
Pohl et al., 1989 (+ imipramine)	60	8 weeks	29.5 mg (mean)	NS
Sheehan et al., 1990	52	8 weeks	57.2 mg	NS
Sheehan et al., 1993 (+ alprazolam)	92	8 weeks	61 mg (mean)	—
Posttraumatic Stress Disorder (PTSD)				
Buspirone (open-label trial)				
Wells et al., 1991	3	NK	35–60 mg	+
Duffy and Malloy, 1994	8	NK	5–30 mg	+
Fichtner and Crayton, 1994	12	NK	46 mg (mean)	+
Obsessive-Compulsive Disorder (OCD)				
Buspirone vs. Placebo				
Piggott et al., 1992 (+ clomipramine)	14	10 weeks	57 mg (mean)	—
McDougle et al., 1993 (+ fluvoxamine)	33	6 weeks	60 mg (up to)	—

Grady et al., 1993 (+ fluoxetine)	13	4 weeks	60 mg	—
Major Depressive Disorder (MDD)				
Buspirone vs. Placebo				
Fabre, 1990	140	8 weeks	54 mg (mean)	+
Rickels et al., 1990	155	8 weeks	90 mg (up to)	+
Robinson et al., 1990	382	NK	90 mg (up to)	+
Appelberg et al., 2001 (+ citalopram or fluoxetine)	102	6 weeks	10–30 mg	+
Buspirone vs. Bupropion SR				
Trivedi et al., 2006 (+ citalopram)	565	6 weeks	41 mg (mean)	NS
Alcohol Dependence				
Buspirone vs. Placebo				
Malcolm et al., 1992	67	24 weeks	45–60 mg	NS
Malec et al., 1996	Meta-analysis	—	—	NS
Cannabis Dependence				
Buspirone vs. Placebo				
McRae-Clark et al., 2009	50	12 weeks	60 mg (up to)	NS

NS, no significant difference; (+), significantly greater improvement in primary efficacy measure for buspirone arm; (−), significantly greater improvement in primary efficacy measure for comparator arm; NK, not known (unpublished).

Agitation. The possible role of buspirone in the management of behavioral and psychological symptoms of dementia (BPSD), in the context of Alzheimer disease (AD), has been examined. Buspirone was compared to trazodone (Desyrel) and placebo in AD patients with behavioral symptoms (n = 10) in a 12-week study. Only subjects taking trazodone, not buspirone, showed improvement compared to placebo. In another study looking at an AD population, buspirone (n = 12) was compared to haloperidol (Haldol) (n = 14) for the treatment of agitation. Subjects taking buspirone demonstrated a near-significant drop in tension when compared to subjects taking haloperidol (11 percent reduction in tension scores for buspirone versus 1.6 percent for haloperidol).

Psychosis. Buspirone has not demonstrated any clear efficacy in the treatment of psychosis. In the context of schizophrenia, efficacy has not been observed in the treatment of either positive or negative symptoms. An RCT using a population with a diagnosis of schizophrenia found that patients taking buspirone (n = 30) had a significantly greater improvement

in cognitive scores compared to placebo (n = 29) following 6 months of treatment, although no significant improvement in psychotic symptoms was observed. A number of open-label studies in the schizophrenic population have shown contradictory evidence, where a 6-week trial found that subjects taking buspirone (n = 20), in addition to a stable dose of an antipsychotic, demonstrated significant improvement in outcome scores, while showing no worsening of tardive dyskinesia and akathisia. Conversely, a smaller study (n = 7) found no benefit and worsening of akathisia.

Somatic Dysfunction. Fibromyalgia patients (n = 22) were found to have an augmented prolactin response compared to healthy controls (n = 14) when challenged with a single dose of buspirone. The researchers concluded that dopaminergic neurotransmission was altered in fibromyalgia via the D_2 receptor and that the onset of this disease may be associated with anxiety. Buspirone has also been investigated in its role for altering colonic sensorimotor function in healthy subjects. In a randomized control study (n = 60), buspirone was compared to venlafaxine XR and placebo in its ability to alter colonic function. It was found that individuals taking buspirone had no significant changes in colonic compliance, tone, and sensation when compared to those on placebo.

Sexual Dysfunction. It has been previously reported that in patients with GAD suffering from decreased sexual function, an improvement in sexual dysfunction was seen after 4 weeks of buspirone. Sexual dysfunction is a common side effect experienced by antidepressant users. Buspirone has been previously investigated in order to determine if this side effect could be reversed. Subjects who met the criteria for a major depressive episode (MDE) and were on a stable dose of an SSRI for at least 4 weeks were given either buspirone or placebo to determine if an improvement in sexual function would occur. Only women in weeks 1 and 3 had a significant improvement in sexual function. An 8-week RCT in females with an MDE being treated with fluoxetine compared buspirone to the antiviral medication amantadine (Symmetrel) and found that neither drug significantly improved sexual dysfunction compared to placebo.

Premenstrual Syndrome. Buspirone has demonstrated some benefit in the treatment of premenstrual syndrome (PMS) in two RCTs. The first RCT compared buspirone (n = 19) to nefazodone (n = 22) and placebo (n = 22) in subjects suffering from premenstrual dysphoria, and found that subjects taking buspirone had a greater global improvement and improvement in irritability compared to placebo. The second RCT demonstrated that subjects on buspirone (n = 17) had a significant improvement in PMS-related symptoms, including aches and pains, cramping, impaired social

interaction, and irritability.

Sleep Apnea. Buspirone as a treatment option for sleep apnea has been studied but no benefit has been observed. One randomized control trial (n = 5) found that patients on buspirone scored less on the Apnea Hypopnea Index (AHI); however, this difference did not reach significance. One animal model study demonstrated efficacy of buspirone in the treatment of apnea, where morphine-induced apnea was reversed with the IV administration of buspirone.

Attention-Deficit/Hyperactivity Disorder. An open-label study of children and adolescents (n = 50) with a diagnosis of both ADHD and oppositional defiant disorder (ODD) were given buspirone for a minimum of 6 months. Eighty-three percent of the study population improved moderately or better with symptoms related to ODD, reducing irritability, aggression, and temper outbursts. An RCT (n = 57) administering buspirone via a transdermal patch did not observe a significant decrease in ADHD-related symptoms in the buspirone group compared to placebo.

Withdrawal from Other Medications/Substances of Abuse. An RCT looking at the withdrawal of opiates found that buspirone had similar efficacy to methadone in alleviating withdrawal symptoms in patients discontinuing heroin. Another RCT looking at the discontinuation of long-term benzodiazepine use in patients with panic disorder, comparing imipramine (Tofranil) to buspirone, found that buspirone did not significantly decrease withdrawal symptoms compared to placebo. A recent open-label study looking at withdrawal from marijuana found that patients on buspirone (n = 11) had a significant decrease in both cravings for drug and irritability.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Nervous System

The most frequent adverse events affecting the nervous system reported by buspirone users included dizziness, headache, drowsiness, lightheadedness, insomnia, fatigue, nervousness, decreased concentration, abnormal thinking, excitement, depression, confusion, vivid dreams, and hostility. Less frequent events included noise intolerance, euphoria, fear, apathy, seizures, and suicidal thoughts.

Gastrointestinal System

Users of buspirone report instances of nausea, GI distress, diarrhea, and vomiting. Less frequent events include anorexia and rectal bleeding.

Cardiovascular System

Tachycardia or palpitations was the most frequently listed cardiovascular effect with the use of buspirone. Less common adverse events included syncope, hypotension, and hypertension. Rare cardiovascular effects include congestive heart failure, cerebrovascular accident, myocardial infarction, cardiomyopathy, bradycardia, and ECG change.

Dermatological Effects

The most frequent dermatological side effect reported with buspirone use was skin rash, while facial edema, easy bruising, blisters, hair loss, and dry skin were reported infrequently. Rare events included acne and thinning of nails.

Endocrine System

Patients taking buspirone reported, infrequently, the following adverse events: change in libido, change in weight, and menstrual irregularity.

THERAPEUTIC INDICATIONS

The FDA has approved buspirone for the management of anxiety disorders or for short-term treatment of anxiety-related symptoms. It is recommended for use in individuals who present symptoms that correspond with GAD based on the DSM criteria, for at least 1 month. There are no data available pertaining to the long-term (greater than 4 weeks) use of buspirone.

PRECAUTIONS AND ADVERSE REACTIONS

Patients with a hypersensitivity to buspirone should not take this medication. Pharmacological studies have indicated that liver and kidney impairment affect drug metabolism and drug clearance, respectively. As such, patients with either condition should receive less than the recommended dose. Buspirone is not recommended for patients with severe liver or kidney impairment. In vivo animal studies have indicated a lack of teratogenic, carcinogenic, and fetal impairment effects of buspirone administration, but similarly well-controlled studies in humans have not been performed, resulting in a current classification by the U.S. Food and Drug Administration of Pregnancy Category B. In rats, buspirone and its metabolites are found in maternal milk and cross the placenta, although it is not known to what extent both these events occur in humans.

The most commonly listed adverse reactions to buspirone that were significantly higher when compared to placebo included dizziness, nausea, headache, lightheadedness, excitement, and sweating or clamminess. A study in a child and adolescent population found that the most frequently reported side effects were lightheadedness, headache, and dyspepsia, while

studies in geriatric populations found no difference in adverse events reported in comparison to a younger population. Rare side effects include congestive heart failure, bradycardia, slurred speech, claustrophobia, tingling of limbs, burning tongue, and photophobia.

DRUG INTERACTIONS

Buspirone is a substrate for CYP3A4. Concurrent use of buspirone and MAOIs) has been associated with a significant increase in blood pressure. Verapamil (Verelan), an inhibitor of CYP3A4, was found to increase the AUC and C_{max} of buspirone by 4.1- and 5.0-fold, respectively. Erythromycin (ERY-C), an inhibitor of CYP3A4, increased both AUC and C_{max} by 19- and 16-fold, respectively. Rifampiin (Rifadin), an inducer of CYP3A4, decreased the C_{max} and AUC by 84 to 90 percent and the half-life from 2.8 to 1.3 hours. With the coadministration of haloperidol and buspirone in subjects with Schizophrenia, buspirone increased the plasma concentration of both haloperidol and its metabolite, although this study did not examine the concentration of buspirone in the presence of haloperidol. Grapefruit juice, an inhibitor of CYP3A4, increased the AUC and C_{max} of buspirone by 9.2- and 4.3-fold, respectively. Other known inducers or inhibitors of CYP3A4 that may affect buspirone plasma concentration include ketoconazole (Nizoral), dexamethasone (Decadron), phenytoin (Dilantin), phenobarbital (Bellatal), and carbamazepine (Tegretol). Concurrent use of certain drugs has been found to change the percentage of protein-bound buspirone. Acetylsalicylic acid (ASA) increased plasma levels of free buspirone by 23 percent, while flurazepam (Dalmane) decreased levels by 20 percent. Other drugs that alter the free plasma level of this drug include desipramine (Norpramin), diazepam, ibuprofen (Advil), propranolol (Inderal), thioridazine (Mellaril), and tolbutamide (Orinase).

DOSAGE AND ADMINISTRATION

Buspirone is available as an oral tablet and is provided in the following strengths: 5, 10, 15, and 30 mg. The recommended initial dose listed by the FDA is 15 mg daily (7.5 mg twice a day). Patients may be titrated to an optimal dose within 2 to 3 days by increasing the dose by 5 mg daily. The maximum daily dose should not surpass 60 mg.

REFERENCES

*Appelberg BG, Syvälahti EK, Koskinen TE, et al. Patients with severe depression may benefit from buspirone augmentation of selective serotonin reuptake inhibitors: results from a placebo-controlled, randomized, double-blind, placebo wash-in study. *J Clin Psychiatry*. 2001;62:448–452.

Apter JT, Allen LA. Buspirone: future directions. *J Clin Psychopharmacol*. 1999;19(1):86–93.

Benyamina A, Lecacheux M, Blecha L, Reynaud M, Lukasiewcz M. Pharmacotherapy and psychotherapy in cannabis withdrawal and dependence. *Expert Rev Neurother*. 2008;8:479–491.

Blier P, Bergeron R, de Montigny C. Selective activation of postsynaptic 5-HT1A receptors induces rapid antidepressant response. *Neuropsychopharmacology*. 1997;16(5):333–338.

Brody D, Adler LA, Kim T, Angrist B, Rotrosen J. Effects of buspirone in seven schizophrenic subjects. *J Clin Psychopharmacol*. 1990;10(1):68–69.

Cantillon M, Brunswick R, Molina D, Bahro M. Buspirone vs. haloperidol: a double-blind trial for agitation in a nursing home population with Alzheimer's disease. *Am J Geriatr Psychiatry*. 1996;4(3):263–267.

Carrão JL, Moreira LB, Fuchs FD. The efficacy of the combination of sertraline with buspirone for smoking cessation: a randomized clinical trial in nondepressed smokers. *Eur Arch Psychiatry Clin Neurosci*. 2007;257(7):383–388.

*Chessick CA, Allen MH, Thase M, et al. Azapirones for generalized anxiety disorder. *Cochrane Database Syst Rev*. 2006;3:CD006115.

Chial HJ, Camilleri M, Ferber I, et al. Effects of venlafaxine, buspirone, and placebo on colonic sensorimotor functions in healthy humans. *Clin Gastroenterol Hepatol*. 2003;1(3):211–218.

Cinciripini PM, Lapitsky L, Seay S, Wallfisch A, Meyer WJ 3rd, van Vunakis H. A placebo-controlled evaluation of the effects of buspirone on smoking cessation: differences between high- and low-anxiety smokers. *J Clin Psychopharmacol*. 1995;15(3):182–191.

*DeMartinis N, Rynn M, Rickels K, Mandos L. Prior benzodiazepine use and buspirone response in the treatment of generalized anxiety disorder. *J Clin Psychiatry*. 2000;61:91–94.

Eison AS. Azapirones: history of development. *J Clin Psychopharmacol*. 1990;10(3 Suppl):2S–5S.

*Fava M, Rush AJ, Alpert JE,, et al. Difference in treatment outcome in outpatients with anxious versus nonanxious depression: a Star*D report. *Am J Psychiatry*. 2008;165:342–351.

Goff DC, Midha KK, Brotman AW, McCormick S, Waites M, Amico ET. An open trial of buspirone added to neuroleptics in schizophrenic patients. *J Clin Psychopharmacol*. 1991;11(3):193–197.

Gross MD. Buspirone in ADHD with ODD. *J Am Acad Child Adolesc Psychiatry*. 1995;34(10):1260.

Hilleman DE, Mohiuddin SM, Delcore MG. Comparison of fixed-dose transdermal nicotine, tapered-dose transdermal nicotine, and buspirone in smoking cessation. *J Clin Pharmacol*. 1994;34(3):222–224.

Imai H, Tajika A, Chen P, et al. Azapirones versus placebo for panic disorder in adults. *Cochrane Database Syst Rev*. 2014;(9):CD010828.

Jaron´czyk M, Chilmonczyk Z, Mazurek AP, et al. The molecular interactions of buspirone analogues with the serotonin transporter. *Bioorg Med Chem*. 2008;16:9283–9294.

Kim JA, Druse MJ. Protective effects of maternal buspirone treatment on serotonin reuptake sites in ethanol-exposed offspring. *Brain Res Dev Brain Res*. 1996;92(2):190–198.

Landén M, Eriksson E, Agren H, Fahlén T. Effect of buspirone on sexual dysfunction in depressed patients treated with selective serotonin reuptake inhibitors. *J Clin Psychopharmacol*. 1999;19(3):268–271.

Landén M, Eriksson O, Sundblad C, Andersch B, Naessen T, Eriksson E. Compounds with affinity for serotonergic receptors in the treatment of premenstrual dysphoria: a comparison of buspirone, nefazodone and placebo. *Psychopharmacology (Berl)*. 2001;155(3):292–298.

Lawlor BA, Radcliffe J, Molchan SE, Martinez RA, Hill JL, Sunderland T. A pilot placebo-controlled study of trazodone and buspirone in Alzheimer's disease. *Int J Geriatr Psychiatry*. 1994;9:55–59.

Malec TS, Malec EA, Dongier M. Efficacy of buspirone in alcohol dependence: a review. *Alcohol Clin Exp Res*. 1996;20:853–858.

McRae-Clark AL, Carter RE, Killeen TK, et al. A placebo-controlled trial of buspirone for the treatment of marijuana dependence. *Drug Alcohol Depend*. 2009;105(1–2):132–138.

Myers RA, Plym MJ, Signor LJ, Lodge NJ. 1-(2-pyrimidinyl)-piperazine, a buspirone metabolite, modulates bladder function in the anesthetized rat. *Neurourol Urodyn*. 2004;23(7):709–715.

Navines R, Martin-Santos R, Gomez-Gil E, Martinez De Osaba MJ, Gasto C. Interaction between serotonin 5-HT1A receptors and beta-endorphins modulates antidepressant response. *Prog Neuropsychopharmacol Biol Psychiatry*. 2008;32:1804–1809.

Peroutka SJ. Selective interaction of novel anxiolytics with 5-hydroxytryptamine1A receptors. *Biol Psychiatry*. 1985;20(9):971–979.

Riblet LA, Taylor DP, Eison MS, Stanton HC. Pharmacology and neurochemistry of buspirone. *J Clin Psychiatry*. 1982;43(12 Pt 2):11–18.

Rickels K, Freeman E, Sondheimer S. Buspirone in treatment of premenstrual syndrome. *Lancet*. 1989;1(8641):777.

*Robinson DS, Rickels K, Yocca FD. Buspirone and gepirone. In: Schatzberg AF, Nemeroff CB, eds. *The American Psychiatric Publishing Textbook of Psychopharmcology*. 4th ed. Arlington, VA: American Psychiatric Publishing; 2009:487–501.

Sahibzada N, Ferreira M, Wasserman AM, Taveira-DaSilva AM, Gillis RA. Reversal of morphine-induced apnea in the anesthetized rat by drugs that activate 5-hydroxytryptamine(1A) receptors. *J Pharmacol Exp Ther*. 2000;292(2):704–713.

Salazar DE, Frackiewicz EJ, Dockens R, et al. Pharmacokinetics and tolerability of buspirone during oral administration to children and adolescents with anxiety disorder and normal healthy adults. *J Clin Pharmacol*. 2001;41(12):1351–1358.

Syvälahti E, Penttilä J, Majasuo H, Pälvimäki EP, Laakso A, Hietala J. Combined treatment with citalopram and buspirone: effects on serotonin 5-HT2A and 5-HT2C receptors in the rat brain. *Pharmacopsychiatry*. 2006;39(1):1–8.

Wong H, Dockens RC, Pajor L, et al. 6-Hydroxybuspirone is a major active metabolite of buspirone: assessment of pharmacokinetics and 5-hydroxytryptamine1A receptor occupancy in rats. *Drug Metab Dispos*. 2007;35(8):1387–1392.

▲ 34.13 Calcium Channel Inhibitors

STEVEN L. DUBOVSKY, M.D.

HISTORY

The complex psychobiology of mood disorders is unlikely to be explained by changes in a single neurotransmitter or receptor. Intracellular messengers, however, regulate multiple aspects of neuronal function, and abnormal regulation of one messenger can have multiple downstream effects on affective, cognitive, and behavioral systems directly and through actions on gene expression. The intracellular calcium ion is particularly interesting in this regard because it regulates activity of multiple neurotransmitters, such as serotonin and dopamine, and has a biphasic action (Fig. 34.13–1). Moderate increases in free intracellular calcium ion concentration ($[Ca^{2+}]_i$) stimulate cellular activity, whereas greater elevations inhibit the same functions and even higher levels within a physiologic range induce apoptosis. A hyperactive intracellular Ca^{2+} signal that contributes to the activation of mania at the same time that it suppresses behavioral and other affect regulation systems depending on the sensitivity of those systems to increased intracellular signaling could explain why so many bipolar patients have rapid fluctuations between mania and depression or experience both states at the same time. The same explanation could apply to observations that hyperparathyroidism can induce both depression and mania. The contribution of excessive intracellular calcium signaling to programmed cell death could be a component of the neuronal loss that has been observed with recurrent affective episodes, and hyperactive calcium signaling in other cellular systems could play a role in the susceptibility of patients with bipolar mood disorders to hypertension and cardiovascular disease. Correction of hyperactive intracellular calcium signaling could explain why treatments such as lithium (Eskalith) can improve both mania and bipolar depression.

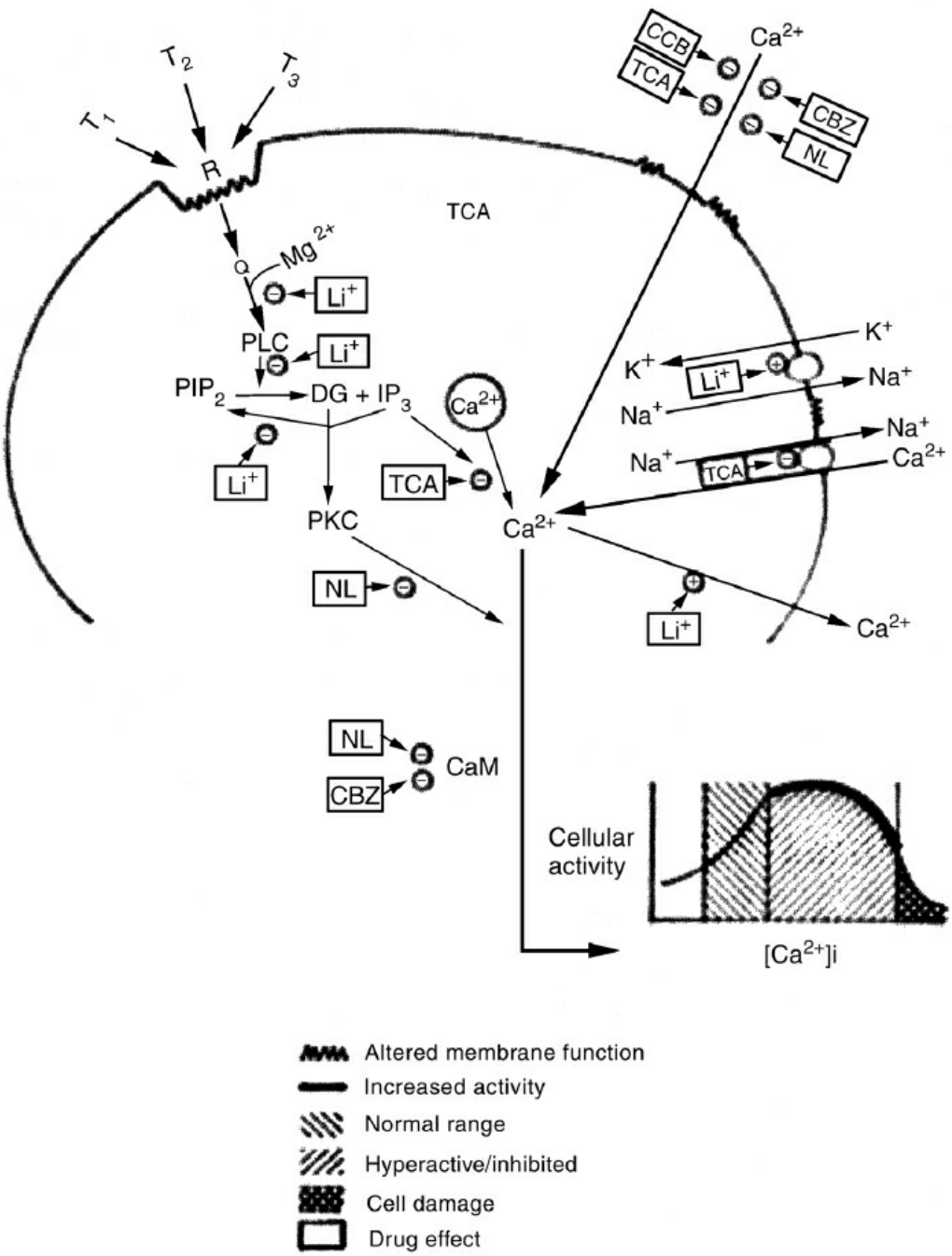

FIGURE 34.13–1. Intracellular calcium ion signaling. Intracellular calcium ion concentration ($[Ca^{2+}]_i$) is increased by entry into the extracellular space and by release from intracellular stores. Calcium enters the cytosol through receptor operated (e.g., *N*-methyl-D-aspartate) and potential-dependent ion-specific channels, as well as in exchange for intracellular sodium. Stimulation of a variety of G-protein–associated receptors (e.g., serotonin type 2 and dopamine type 2) activates turnover of the membrane phosphoinositol system, with production of inositol triphosphate (IP_3) and diacylglycerol (DG). IP_3 releases calcium ions from intracellular stores, and DG draws protein kinase C (PKC) to the cell membrane, where it is activated by Ca^{2+}. Moderate increases in $[Ca^{2+}]_i$ stimulate cellular processes by activating PKC and a variety of other enzymes that are activated by the calcium–calmodulin (CaM) complex; greater increases suppress the same processes. CBZ, carbamazepine; CCB, calcium channel blocker; Li, lithium; NL, neuroleptic; PIP_2, phosphatidylinositol biphosphate; PLC, phospholipase C; TCA, tricyclic antidepressant.

Free intracellular Ca^{2+} concentration is normally regulated tightly at approximately 100 nM, or 0.0001 of the Ca^{2+} concentration in the extracellular fluid. Compared with controls, elevations of resting ($[Ca^{2+}]_B$) free intracellular calcium ion concentration have been found in affectively ill patients with bipolar disorder in blood platelets, lymphocytes, B-lymphoblast cell lines (BLCLs), olfactory neurons, and neuroblastoma cells, as well as elevations of intracellular calcium ion concentration stimulated by agonists ($[Ca^{2+}]_S$) including thrombin, platelet-activating factor (PAF), serotonin, dopamine, thapsigargin, and various mitogens. Most studies that demonstrate increased resting and/or agonist stimulated $[Ca^{2+}]_i$ in bipolar disorder have reported similar elevations in mania and bipolar depression. When bipolar I and bipolar II disorders have been compared, the results have been similar in some if not all studies. These changes tend to normalize with euthymia after treatment with various medications or ECT (Fig. 34.13–2). There are a few reports of greater serotonin-stimulated $[Ca^{2+}]_S$ in unipolar depressed patients than in controls, with blunting of the excessive intracellular signal by serotonin reuptake inhibitors, possibly because these samples included patients with unrecognized or subsyndromal bipolar disorder. This possibility is supported by the finding in 53 patients who had a diagnosis of unipolar depression with melancholia and who had elevated platelet $[Ca^{2+}]_S$ in response to serotonin that levels were significantly higher in 7 patients whose diagnosis was changed over the next 8 to 10 years to bipolar disorder. Hyperactive intracellular Ca^{2+} signaling appears to be specific to bipolar disorder and is not found in schizophrenia. This finding is not a function of psychosis, which is common in bipolar disorder, as $[Ca^{2+}]_B$ and $[Ca^{2+}]_S$ are found in nonpsychotic as well as psychotic bipolar disorder.

The same finding with different agonists in different cells from patients with bipolar disorder suggests that hyperactive intracellular Ca^{2+} signaling is not a function of a specific receptor but a primary change in intracellular calcium ion dynamics. The impression of a diffuse change in cellular function rather than an effect of a circulating factor such as cortisol is strengthened by the finding of lack of increase in $[Ca^{2+}]_i$ in control platelets after incubation with a plasma ultrafiltrate from bipolar disorder patients with elevated platelet $[Ca^{2+}]_i$.

The immediate source of increased $[Ca2+]_i$ in bipolar disorder has been thought to be increased release from intracellular calcium stores promoted by inositol triphosphate (IP_3), a product of membrane phosphoinositide (PI) turnover. Proposed mediators of increased release of stored intracellular calcium include increased PI turnover; hyperactivity of the calcium-binding protein neuronal calcium sensor-1 (NCS-1), which is activated by IP_3; a hyperactive G-protein, mitochondrial dysfunction;

increased production of protein kinase C (PKC) from membrane PI turnover; secondarily altering Ca2+ efflux from intracellular stores; and decreased expression of the antiapoptotic and neurotrophic B-cell CLL/lymphoma (Bcl-2) gene, the product of which reduces intracellular Ca2+ release by IP_3.

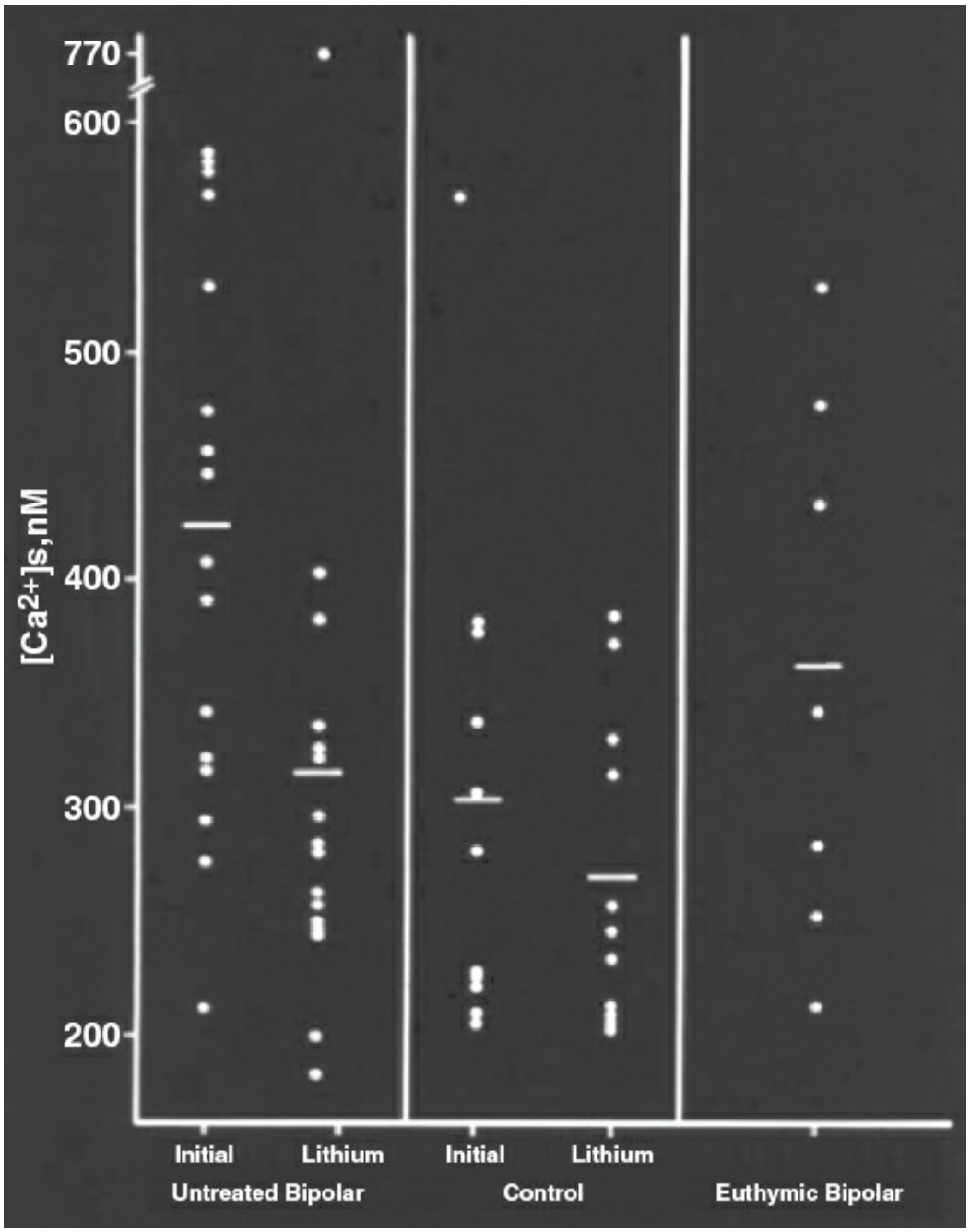

FIGURE 34.13–2. Elevated platelet intracellular calcium ion concentration in bipolar disorder and normalization by lithium (Eskalith). Resting ($[Ca^{2+}]_B$) and, in this figure, stimulated ($[Ca^{2+}]_S$), as well as interindividual variability, are increased in mania and bipolar depression. Hyperactive intracellular calcium signaling is normalized by in vitro incubation with therapeutic concentrations of lithium, and bipolar patients who are euthymic after treatment with lithium or anticonvulsants have mean $[Ca^{2+}]_S$ levels that are statistically the same as in controls.

Heterozygous Myshkin (Myk/+) mice, which have a missense mutation of the neuron-specific sodium/potassium ATPase, exhibit mania-like behavior such as hyperambulation, altered sleep–wake cycle, and increased reward preference. The manic-like behavior, which is reduced by lithium and valproate, is associated with increased release of Ca^{2+} from intracellular stores in cortical neurons via activation of IP_3 receptors, as well as increased glutamate-induced Ca^{2+} currents. In humans, polymorphisms of the CACNA1C gene, which encodes the alpha subunit of the calcium channel Cav1.2, have been associated with bipolar disorder in a number of studies. Such results suggest that altered intracellular calcium homeostasis could be involved in the actual pathophysiology of bipolar disorder and implies that this might be a target of effective thymoleptic treatments.

Lithium has a number of effects on intracellular Ca^{2+} signaling. For example, it inhibits a crucial step in membrane PI turnover (inositol-1-monophosphatase), which could reduce production of IP_3, thereby decreasing calcium mobilization from intracellular stores; this possibility is suggested by the findings that lithium-treated bipolar patients, but not controls, have reduced platelet phosphatidylinositol biphosphate (PIP_2), that lithium increases accumulation of IP_1, and that lithium reduces concentrations of myoinositol in the frontal cortex of bipolar patients. Since lithium is an uncompetitive inhibitor of this enzyme, its inhibitory activity increases with higher rates of PI turnover. Another recently demonstrated action by which lithium could dampen hyperactive intracellular calcium signaling is antagonism of the intracellular calcium-binding protein NCS-1, which enhances IP_3 calcium channel activity, promoting efflux of Ca^{2+} from intracellular stores to the cytosol. Lithium may also interfere with calcium channel activity or it may enhance calcium efflux, perhaps through an effect on sodium countertransport. Additional relevant lithium actions include antagonism of the hyperactive G protein by competition for magnesium ion binding required for G protein dissociation, induction of Bcl-2, and inhibition of glycogen synthase kinase-3β, an effect that enhances membrane stability. Inhibition of PKC, a ubiquitous protein activated by intracellular Ca2+ that is increased in bipolar disorder, may be another mechanism relevant to calcium signaling. Carbamazepine (Tegretol) inhibits Ca^{2+} currents in a variety of models, and the time course of suppression of calcium-dependent potentials is comparable to that produced by the calcium channel inhibitor verapamil (Calan, Isoptin). In vitro incubation with lithium or carbamazepine lowers platelet and lymphocyte $[Ca^{2+}]_i$ significantly in ill bipolar patients but not in controls or euthymic bipolar patients (Fig. 34.13–2). Lamotrigine (Lamictal), another medication that has been found useful in bipolar disorder, also has calcium channel blocking properties, as do valproate

(Depakote), gabapentin (Neurontin), and pregabalin (Lyrica), although these effects are not nearly as prominent as with lithium and carbamazepine. Attenuation of hyperactive intracellular calcium signaling by antimanic drugs may have a number of different mechanisms (Fig. 34.13–1). Findings of elevated intracellular Ca^{2+} activity in mania and bipolar depression that is attenuated by established mood stabilizing medications warrant investigation of medications that act primarily on intracellular Ca^{2+}.

CHEMISTRY

Verapamil, the first calcium channel antagonist to be introduced, was synthesized in 1962. A phenylalkylamine derivative of papaverine (Pavarine), verapamil was found to have negative inotropic effects that, in 1967, were postulated to be due to reduction of excitation–contraction coupling caused by inhibition of calcium influx into cardiac cells. Three additional classes of calcium channel inhibitors were subsequently identified, with significantly different structures (Fig. 34.13–3), leading to differential activity in various tissues depending on their binding site on the calcium channel. In addition to the phenylalkylamine verapamil, currently available calcium channel blocker (CCB) classes include benzothiazepine (diltiazem), and dihydropyridine (nifedipine, nimodipine, amlodipine, felodipine, isradipine, nicardipine, and nisoldipine). Because of their heterogeneity of structure and action, these medications are not interchangeable. However, they all have the capacity to reduce excessive excitability of diverse cellular systems.

Verapamil and diltiazem (Cardizem) are crystalline powders that are soluble in water and chloroform. Nifedipine (Procardia) is insoluble in water, but it is soluble in ethanol. Nimodipine (Nimotop), a 1,4-dihydropyridine (DHP) that is insoluble in water but is highly lipophilic, has two closely related congeners: isradipine (DynaCirc) and amlodipine (Norvasc). Amlodipine is a white crystalline powder that is slightly soluble in water and ethanol.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

Absorption. The CCBs are nearly completely absorbed after oral administration, but bioavailability is decreased by extensive first-pass hepatic and probably intestinal metabolism. First-pass metabolism also results in considerable interindividual variability in blood level at a given dose. As noted in Table 34.13–1, oral bioavailability of the CCBs is generally less than that of other drugs used to treat mania, but effective concentrations for their primary indication are an order of magnitude lower. Optimal serum concentrations for mood disorders or any other

conditions have not been investigated.

Verapamil is metabolized by D-dealkylation and *O*-demethylation to at least 12 metabolites. One of these, norverapamil, has approximately 20 percent as much cardiovascular activity as the parent drug in animal models and a half-life of 10 hours. Diltiazem, which is metabolized by *O*-deacetylation or *N*-demethylation, has active metabolites that do not appear to be clinically important. Nifedipine, nicardipine (Cardene), and most of the other dihydropyridines have inactive metabolites. The majority of people are rapid metabolizers of nifedipine and possibly other DHPs; approximately 17 percent are slow metabolizers. Isradipine is metabolized to inactive monoacids and a cyclic lactone.

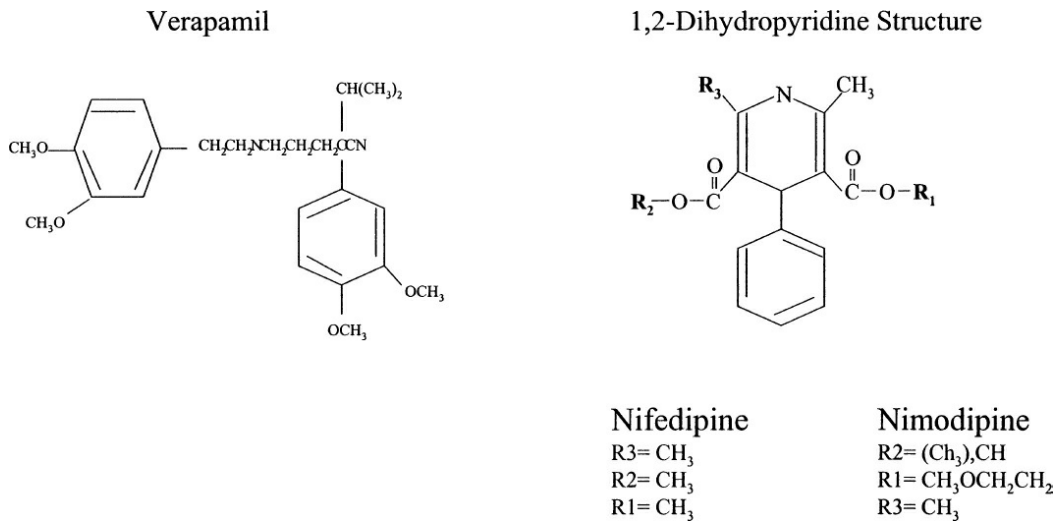

FIGURE 34.13–3. Structure of some calcium channel inhibitors.

Table 34.13–1.
Comparative Chemistry of Some Calcium Channel Blockers and Some Mood Stabilizers

Drug	Oral Availability (%)	Plasma Protein Binding (%)	Volume of Distribution (L/kg)	Elimination Half-Life (hours)	Effective Concentration for Primary Indication
Verapamil (Calan, Isoptin)	22 ± 8	90	5.0 ± 2.1	4.0 ± 1.5	40–120 ng/mL
Diltiazem (Cardizem)	44 ± 10	78	3.0 ± 1.2	6.0 ± 1.2	Unknown
Nifedipine (Procardia)	50 ± 13	96	0.8 ± 0.22	3.0 ± 1.3	27–67 ng/mL
Lithium (Eskalith)	100	0	0.8 ± 0.34	22.0 ± 8.0	0.5–1.5 mEq/L
Carbamazepine (Tegretol)	70	74	1.4 ± 0.4	15.0 ± 5.0	4–12 µg/mL
Divalproex (Depakote)	100	93	0.2 ± 0.07	14.0 ± 2.2	50–125 µg/mL

Distribution. Peak plasma levels of the majority of CCB agents are

achieved within 30 minutes to 3 hours after oral dosing, but amlodipine does not reach peak plasma concentrations for 6 hours. All CCBs are extensively bound to plasma proteins (70 to 98 percent). Most CCBs are excreted in breast milk. Verapamil and norverapamil can be recovered from human cerebrospinal fluid (CSF) after oral administration, and the concentration of phenylalkylamines in the brain is sufficient to be protective after cerebral ischemia in rats. However, the more lipophilic nimodipine crosses the blood–brain barrier more readily.

As noted in Table 34.13–2, L-type channels are distributed in multiple organ systems, whereas many of the other calcium channels are found primarily in neurons. Since all of the CCBs that have been used for psychiatric indications bind to L-type channels, their actions are not limited to the CNS, and central actions can be accompanied by cardiovascular and other effects that may be troublesome.

Elimination. Most CCBs have elimination half-lives in the range of 1.3 to 6.0 hours, except for amlodipine, which has a half-life of 35 to 50 hours. Although amlodipine can be given once daily, the short half-lives of most CCBs require multiple daily doses. However, because repeated administration saturates hepatic metabolizing enzymes, bioavailability and half-lives increase with chronic administration, and the dosing interval can be increased. For example, verapamil is administered four times per day initially, but after 1 to 2 months the dosing interval can often be reduced to twice a day. Intestinal and hepatic metabolism of many CCBs by CYP3A4 enables increased bioavailability by 3A4 inhibitors such as grapefruit juice.

Pharmacodynamics

Mechanism of Action. Calcium ions enter the cytosol through receptor-operated channels gated by receptors for hormones and transmitters, such as some catecholamines and the excitatory amino acids, and through potential-dependent channels (PDCs or voltage-gated channels), which are gated by membrane potential. In the brain, potential-dependent calcium channels are localized in regions rich in synapses, perhaps because high $[Ca^{2+}]_i$ must be produced rapidly to regulate neurotransmitter release. Additional pathways for calcium entry include leakage through an ungated channel and exchange of extracellular calcium ions for intracellular sodium ions (Na^+) (Fig. 34.13–1). Under physiological conditions, PDCs can be regulated by receptor-mediated events, such as the production of IP_3, and receptor-operated channels can be gated by voltage dependent events. In addition, extracellular Ca^{2+} entering the cell may release calcium ions from intracellular stores, and trigger amounts of calcium ions released from intracellular stores may facilitate calcium channel opening. Even subtle alterations of the function of one kind of calcium channel can therefore

have significant effects on the overall balance of calcium-dependent cellular activity.

CCB agents interact with the PDC. This type of channel consists of five allosterically linked subunits: α_1, α_2, β, γ, and ε; the α_1 channel has a hydrophobic region that spans the cell membrane and outlines and gates the actual calcium pore (Fig. 34.13–4). The α_1 subunit has at least four and possibly seven or more separate binding sites for different CCB classes. Thus far, four separate genes coding for the L-channel α_1 subunit have been identified, producing different versions of CCB-binding sites. The other subunits, and possibly other endogenous circulating factors, can alter the conformation of the α_1 subunit, changing the affinity of the L-channel for a given CCB. Endogenous regulators unrelated to neurotransmitters appear to modulate PDC gating and CCB binding and may be altered in disease states.

Six types of potential-dependent calcium channels have been characterized (Table 34.13–2). The L (for long-lasting) channel requires significant depolarization for Ca^{2+} entry and inactivates slowly. L-type calcium channels do not participate directly in neurotransmitter release, but they play an important role in determining neuronal excitability. Currently available CCBs bind to L-channels, which are found in the cardiovascular, endocrine, and nervous systems, and in muscle. Because of the widespread distribution of L-channels, these medications have multiple effects. The T (transient) channel is activated by small depolarizations and inactivates rapidly. T-type calcium channels may be involved in the action of ethosuximide (Zarontin), valproic acid (Depakene), and divalproex (Depakote) in the treatment of absence seizures. The N (neither L nor T) channel, which is primarily found on CNS neurons, is unresponsive to CCBs. Levetiracetam (Keppra), an anticonvulsant used for refractory epilepsy, antagonizes N-type channels. This action does not appear to be associated with reduction in elevated $[Ca^{2+}]_i$ in platelets of bipolar disorder patients or clinically with a mood stabilizing effect. Ziconotide (Prialt), a peptide toxin used for chronic pain, also antagonizes N-channels. A rapidly inactivating P (Purkinje cell) channel identified in cerebellar Purkinje cells is insensitive to the DHP CCBs. Q and R channels, like N and P channels, participate in neurotransmitter release.

Table 34.13–2.
Calcium Channel Subtypes

Channel Type[a]	L	T	N	P	Q	R
Locations	CV, neurons, endocrine, skeletal muscle	Neurons, smooth muscle, SA node	Neurons	Neurons	Neurons	Neurons

Functions	Excitation-contraction coupling, NT release, vascular tone	SA node activity, repetitive neuron depolarization, hormone release	NT release	NT release	NT release	NT release

[a]Most currently available calcium channel blockers bind to the L-channel but do not have binding sites on other channels.

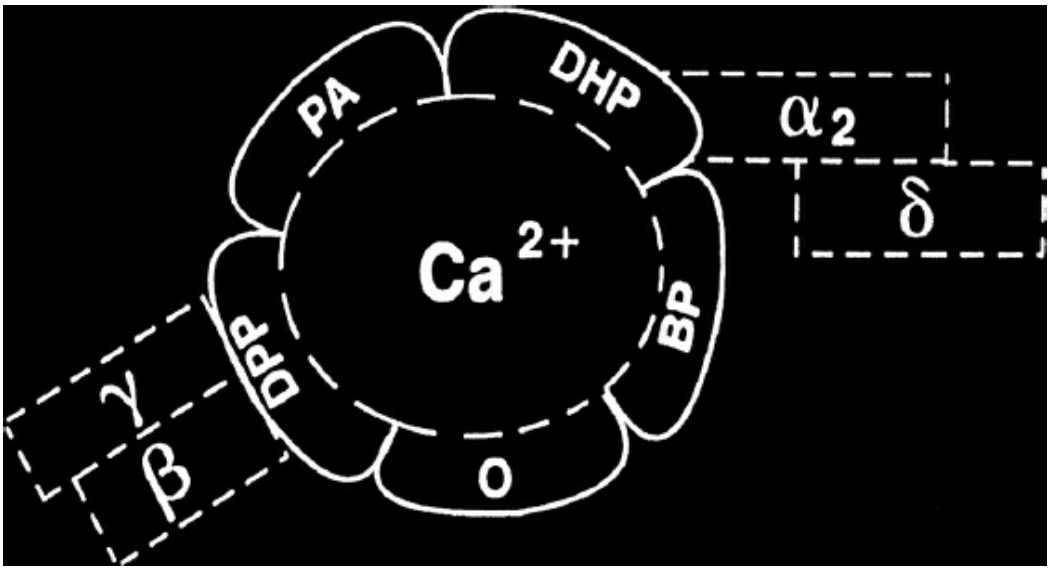

FIGURE 34.13–4. Structure of the calcium channel blocker–binding site. BP, benzodiazepine-binding site; DHP, 1,4-dihydropyridine-binding site; DPP, diphenylpiperazine-binding site; O, other calcium channel blocking agent binding sites; PA, phenyl alkylamine–binding site.

In addition to binding to L-channels, nimodipine, nicardipine, methoxyverapamil, flunarizine (Adalat), and cinnarizine (Sturgeon) may also antagonize T-channels, and a phenylalkylamine-binding site exists on the inner mitochondrial membrane. Variations in CCB structure produce variations in potency at different L-channels in different tissues, and depending on the distribution of binding sites for a given chemical class of CCBs, that class may be more or less active in a particular tissue or region. For example, dihydropyridines are more selective for vascular tissue, whereas verapamil and diltiazem have more prominent antiarrhythmic properties. In animal studies, nifedipine, diltiazem and verapamil blocked the effect of subconvulsive doses of pentylenetetrazole (PTZ) to the same extent as valproate, and diltiazem enhanced the anticonvulsant effect of valproate. In other studies, nifedipine, verapamil, and diltiazem increased the threshold for PTZ-induced clonic seizures, and amlodipine blocked PTZ-induced seizures in mice. However, verapamil did not improve epilepsy control in phenobarbital-resistant canine epilepsy. Possibly as a result of T-channel antagonism, nimodipine has more of a potential for anticonvulsant actions.

Several additional features of the calcium channel contribute to activity

profiles of CCBs of different structure. Along with allosteric regulation and the influence of circulating modulators, membrane potential determines the conformation and, therefore, the affinity of CCB-binding sites. Dihydropyridine binding is enhanced by membrane depolarization, the CCB stabilizing the channel in an inactivated state. Most DHPs are uncharged at physiologic pH and easily gain access to the receptor through the cell membrane if it is sufficiently depolarized. Charged DHPs, such as amlodipine, have slower onsets of action because of interactions with negatively charged phosphate head groups in the cell membrane. Verapamil and diltiazem, which are also charged at physiological pH, gain access to the receptor through open channels and therefore are more active when channel openings are more frequent. The activity dependence of charged and uncharged CCBs makes these drugs much more potent in hyperactive tissue than in normal tissue. As a result, actions measured in normal cell preparations may not reflect activity in pathological states, a point emphasized by the finding that lithium attenuates elevated $[Ca^{2+}]_i$ in platelets of bipolar disorder patients but has no significant effect on $[Ca^{2+}]_i$ in normal platelets.

Several objections can be raised to the hypothesis that attenuation of excessive rises in $[Ca^{2+}]_i$ is a primary mechanism of action of CCBs in bipolar illness. First, as noted previously, L-channels do not appear to be involved in the action potential-dependent release of neurotransmitters, and calcium channel subtypes that are more specific to the CNS are not affected by L-type CCBs. However, neurotransmitters activate L-channels, making them an important component of neuronal signaling, and it is not certain that neurotransmitter dysfunction is the primary pathophysiology of mood disorders anyway. Similarly, low concentrations of CCBs do not affect neurotransmitter release or neuronal Ca^{2+} currents, because depolarization in these cells is normally too brief to permit sufficient binding of the CCBs, but higher concentrations of CCBs may alter neuronal activity in other fundamental ways, and CCBs have greater effects on neurons with prolonged or marked depolarization. Reduction in tonic calcium influx may reduce the intracellular pool from which more immediate release results in rapid neuronal activation.

It could be argued that it is the anticonvulsant property of some CCBs and not their effect on Ca^{2+} currents that accounts for their antimanic activity, and there is some overlap between intracellular calcium dynamics in epilepsy and bipolar disorder. For example, mutations of the CACNA1A calcium channel gene have also been associated with idiopathic epilepsy, and prolonged elevations of $[Ca^{2+}]_i$ can contribute to the development and maintenance of seizure disorders through changes in GABA receptor recycling, gene transcription, protein synthesis, and neuronal sprouting. Nimodipine, a CCB with anticonvulsant properties, seems particularly

useful in the treatment of mood disorders, although this may be more a function of its high lipid solubility and penetration into the CNS. The finding noted earlier that levetiracetam, which is clearly an anticonvulsant, does not normalize increased $[Ca^{2+}]_i$ in bipolar disorder and has little efficacy in this condition, could indicate that a combination of anticonvulsant and calcium antagonist properties may be associated with a mood stabilizing effect. Some CCBs also antagonize nicotinic cholinergic receptors, and some reduce dopamine D2 receptor signaling, which could also contribute to their clinical effects. Inhibition of PKC may be an additional independent action of CCBs relevant to an antimanic effect, as suggested by studies demonstrating an antimanic effect of tamoxifen, which has PKC inhibitory properties.

DESIGN AND INTERPRETATION OF CLINICAL DRUG STUDIES

The first type of calcium antagonist used in psychiatry was calcitonin (Cibacalcin), which lowers $[Ca^{2+}]_i$ by driving Ca^{2+} into intracellular storage sites. John Scott Carman and Richard A. Wyatt found that three manic patients had temporary reduction in agitation after injections of subcutaneous salmon calcitonin (Calcimar) but not after administration of placebo. In a subsequent trial, a group of nine patients with "chronic psychopathological suffering" accompanied by depression or anxiety but no formal diagnoses felt tranquillized and demonstrated a reduction in Brief Psychiatric Rating Scale scores after receiving an unspecified dose of salmon calcitonin for 20 days.

The first report of the use of a CCB in the treatment of mania was a double-blind, placebo-controlled trial of verapamil in a single manic patient. The rationale for choosing this particular CCB was that it was the only CCB that the manufacturer would make available at a time at which none of these medications had yet been approved for use in the United States. Based on initial positive experience in case reports and open trials, double-blind studies of verapamil in bipolar disorder have been conducted. Because of lack of interest of CCB manufacturers in psychiatric applications of the CCBs, there have not been any large, multicenter trials of any of these medications in mood disorders.

In initial studies, verapamil was superior to placebo and no treatment in mania, and two studies reported equivalent antimanic efficacy with lithium. A 2-week open trial of diltiazem in seven manic patients, a 3-year follow-up of a bipolar patient for whom flunarizine was successfully substituted for lithium, a 1-week trial of nimodipine, and a prolonged double-blind follow-up of three bipolar patients receiving high doses of nimodipine suggested possible antimanic applications of these medications. In 10 manic patients who did not respond to 3 weeks of treatment with lithium, switching to verapamil for 3 weeks did not produce

better results, but another 3 weeks of treatment with the combination of lithium and verapamil led to a robust response.

Verapamil has been noted to prevent antidepressant-induced hypomania. Extended controlled follow-up after acute treatment demonstrated the usefulness of nimodipine as maintenance therapy in bipolar illness, and verapamil also has been found effective in the clinical setting in preventing affective recurrence. Careful longitudinal observations have suggested possible usefulness of nimodipine for ultradian cycling and recurrent brief depression. Blinded monotherapy with nimodipine was more effective than placebo for 10 of 30 patients with refractory, highly recurrent bipolar and unipolar mood disorders; augmentation of nimodipine with carbamazepine substantially improved the status of 4 of 14 partial nimodipine responders. The highly refractory nature of these patients' illnesses and the careful evaluation of response using Post's life-charting method increase confidence in the results, even though the sample was small.

In one of two negative trials of verapamil in mania, single-blind lithium produced significantly more improvement in mania rating scales than verapamil in a 28-day randomized trial of 40 manic patients. However, doses and frequency of administration of verapamil were lower than are usually effective, and both groups were still quite symptomatic at the end of the trial. The second negative study was a random assignment trial of 32 acutely manic patients, most of them psychotic, to placebo or verapamil given three times per day. Both groups had similarly slight and statistically insignificant reductions in mania and depression rating scale scores. However, patients only remained in the treatment for a mean of 9.3 (verapamil) to 13.3 (placebo) days, a duration of treatment that would not be considered adequate for any antimanic drug. Most patients in clinical practice with psychotic mania would be treated with antipsychotic drugs along with or before an antimanic drug. In addition, although the authors stated that verapamil doses could be adjusted to as much as 480 mg per day, actual doses received were not noted, and verapamil may not have been administered frequently enough to achieve a steady state acutely. Finally, no data were presented to determine whether patients had been refractory to lithium or other treatments; patients who have responded to verapamil monotherapy in most studies have also been responsive to lithium.

Most antimanic and mood-stabilizing medications are effective for some patients and not others, and negative studies have been published of all currently used mood stabilizers. Possibly because CCBs have not been popularized by manufacturers to the same extent as anticonvulsants and some atypical antipsychotic drugs, the two negative antimanic trials of verapamil led to conclusions by some authorities that the entire class of CCBs is ineffective for bipolar disorder, a generalization that would not be

applied to other medication classes and that is not justified for verapamil, given the limitations noted previously. It is also worth noting that studies supporting the thymoleptic properties of popular medications such as the atypical antipsychotic drugs have utilized sample enrichment, while CCB studies have enrolled more heterogeneous samples of patients who have not been demonstrated a priori to respond to the experimental treatment before a randomized trial is conducted.

As is true of lithium, the usefulness of CCBs in the treatment of depression is not as clear as their potential applications in mania. Nifedipine reduces immobility in the behavioral despair test, which is believed to be an animal model of antidepressant activity. Flunarizine was thought to reduce depressive recurrences in one patient, and one patient with recurrent brief depression responded to blind trials of nimodipine and verapamil. Verapamil may have acute antidepressant properties in some bipolar and psychotically depressed patients. However, a Czech study found that overall verapamil was no more effective than placebo and was inferior to amitriptyline (Elavil) in the treatment of depression, although a few patients did seem to have a genuine antidepressant response to verapamil in this study and in a previous report by the same authors. As the investigators did not differentiate between bipolar and unipolar depression, it is possible that, as would be expected with lithium, those with bipolar depression were more likely to respond to verapamil than were those with unipolar depression. Verapamil appeared to augment the antidepressant action of imipramine (Tofranil) in a patient with unipolar depression who had previously responded to lithium augmentation; when imipramine was discontinued 6 months later, the patient remained euthymic on verapamil alone. Compared to placebo, treatment with nicardipine during ECT resulted in a better antidepressant response to ECT. Conversely, investigators thought that nifedipine caused depression itself or resistance to antidepressants in five unipolar depressed patients. In a blinded protocol involving 10 patients with bipolar depression generally of more than 6 months duration, acute treatment with isradipine resulted in significant improvement in 6 patients and remission in two.

For several reasons, CCB studies have been considerably smaller than studies of other putative antimanic and mood-stabilizing agents. Manufacturers have not been particularly interested in possible psychiatric implications of these medications. In addition, clinical testing of hypotheses of actions on second messengers and other intracellular processes has not been of as much interest as studies of neurotransmitter, receptor, and anticonvulsant hypotheses and development of “me too” products. These conceptual and financial limitations on clinical studies of CCBs have constrained the data available to reliably assess the place of CCBs in the treatment of bipolar disorder.

As is true of other medications used to treat bipolar mood disorders,

insufficient numbers of patients with different bipolar subtypes have been studied to understand which patients are most likely to benefit. The scatter in actual levels of $[Ca^{2+}]_i$ in bipolar patients (Fig. 34.13–2) undoubtedly reflects the biological and clinical heterogeneity of the disorder, and patients with hyperactive calcium signaling may be more responsive to CCBs than those without this finding. This possibility is suggested by findings in a few small studies that bipolar patients with elevated $[Ca^{2+}]_i$ had a better response to lithium than those with normal levels. The possibility remains that CCBs may be more effective in bipolar disorder accompanied by elevated $[Ca^{2+}]_i$.

Another issue to consider is that the heterogeneity of CCB-binding sites provides for different spectra of action of various CCBs. Just as patients with one bipolar variant may do better with CCBs than those with another variant, it is possible that patients who do not respond to one CCB class may have a good response to another. Combinations of CCBs with each other and with other antimanic drugs could also produce additive effects on Ca^{2+}-dependent and other systems to make them useful in complex treatment-resistant states. However, such possibilities have not been investigated formally.

Bearing these limitations in mind, the available data suggest that verapamil is most likely to be useful for mania and hypomania in patients who are responsive to lithium but cannot tolerate it, whereas patients who do not respond to lithium seem less likely to respond to verapamil. The only available data on combinations of verapamil and other medications involve a few case reports and small studies suggesting that addition of verapamil to lithium was helpful to some patients who did not respond to lithium alone. Nimodipine seems to have potential for treating refractory and complex bipolar disorder, especially in combination with other thymoleptic medications, and there are fewer drug–drug interactions with nimodipine than verapamil. In addition, nimodipine may be useful in bipolar depression. However, larger controlled studies are necessary to assess its precise role. The few small studies of isradipine suggest that this medication, which is closely related to nimodipine, may have similar applications. Verapamil does not appear to be predictably useful for depression, and some CCBs may inhibit an antidepressant response in unipolar depression. However, a single study suggested that verapamil could be a useful adjunct in panic disorder.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Animal studies and clinical experience with verapamil when it is used to treat premature labor and fetal arrhythmias suggest that it is relatively safe in pregnancy, at least in later trimesters. Reports have been published of three patients for whom verapamil was effective for mania during

pregnancy without any evidence of teratogenicity. A study of 28 pregnant women with acute episodes of mania or bipolar depression and nine euthymic pregnant women who switched to verapamil maintenance from another mood stabilizer found no complications of pregnancy, as well as a 100 percent response rate in mania, a 77 percent response rate in mixed states, and a 39 percent response rate in depression; six of eight patients who continued verapamil maintenance therapy remained well. A few case reports of similar outcomes with nimodipine during pregnancy have emerged, and nimodipine and isradipine have been used to treat or prevent pre-eclampsia without adverse effects on the mother or newborn. Obviously, such experience is insufficient to ensure its safety in this setting, but thus far evidence of teratogenicity has not emerged. There is not enough information about effects of other CCBs during pregnancy.

In a prospective study, elderly patients taking short-acting CCBs, especially dihydropyridines like nifedipine, had a greater risk of myocardial infarction (MI) than patients taking other antihypertensives. However, as severity of hypertension was not controlled, and CCBs are usually used as second-line treatments of more severe hypertension, the possibility could not be excluded that patients treated with CCBs were more likely to have this complication because their hypertension was more severe. Although the investigators speculated that coronary vessel steal caused by diffuse vasodilatation accounted for the increased risk of MI, the large Danish Verapamil Infarction Trial (DAVIT) found that CCBs reduced the risk of recurrent MI when used prophylactically after an index MI.

A review of Medicare files of elderly patients treated for hypertension found a somewhat increased risk of a variety of cancers in patients taking short-acting CCBs compared with those taking other antihypertensives. The investigators argued that CCBs could interfere with apoptosis (programmed cell death), which is dependent on increased Ca^{2+} influx that activates proteolytic enzymes. However, cell growth is also calcium dependent and antihypertensive use was assessed years before the presence of cancer was recorded. Although carcinogenesis is a prolonged process, there was no way of knowing when cancer developed, whether antihypertensives were taken during carcinogenesis, or whether these medications were still being taken when cancer was initiated or diagnosed. Furthermore, the CCBs have been found to have potential applications in the treatment of cancer discussed below.

THERAPEUTIC INDICATIONS

Because of their capacity to reduce cellular hyperactivity without suppressing normally active cells, CCBs have been used to treat angina, hypertension, migraine headaches, Raynaud phenomenon, esophageal spasm, premature labor, and epilepsy; verapamil and diltiazem, but not the DHPs, have class IV antiarrhythmic activity and are used to treat

supraventricular arrhythmias. Several CCBs have been found to retard the development of atherosclerosis in animals. In view of the role of excessive elevations of $[Ca^{2+}]_i$ in apoptosis, nimodipine, which is used to block cerebral arteriospasm after subarachnoid hemorrhage, also has the potential to reduce anoxic brain damage. Preliminary studies suggested that this feature could slow progression of Alzheimer disease, but a large multicenter trial found no difference between nimodipine and placebo in this disorder.

In the few reported cases in which verapamil or diltiazem was administered to brain-damaged manic patients, the CCBs were well tolerated. CCBs might be considered in the presence of physical conditions in which other medications would be more dangerous (e.g., pregnancy) or in which a CCB might be helpful, such as hypertension, migraine headaches, tardive dyskinesia, Tourette syndrome, or supraventricular tachyarrhythmias. Since the CCBs usually do not cause weight gain or metabolic changes, they may be better tolerated by patients for whom such effects would be particularly troublesome. Lack of a need for blood tests is appealing to some patients.

Nifedipine chewed and placed under the tongue has been used to treat hypertensive crises associated with MAOIs. However, absorption of nifedipine is poor from the buccal mucosa, and this approach has been found not to be pharmacologically effective. A study of 18 normal subjects found that isradipine reduced the subjective response to methamphetamine and craving for the drug, suggesting that it could have applications in the treatment of stimulant abuse.

One mechanism of resistance to cancer chemotherapy is induction of the *MDR1* gene, which codes for P-glycoprotein, a drug extrusion protein that removes chemotherapy drugs from cancer cells. Verapamil, which inhibits P-glycoprotein, has been used to reduce chemotherapy resistance. Since cell proliferation is a calcium-dependent process, CCBs can inhibit tumor cell growth, and the ability of T-channels to generate rapid calcium currents at resting membrane potentials may account for findings that experimental T-channel antagonists reduce cancer cell proliferation in vitro. An epidemiologic study found an inverse relationship between the use of verapamil and the risk of prostate cancer.

PRECAUTIONS AND ADVERSE REACTIONS

The most common CCB side effects involve consequences of excessive vasodilatation, such as dizziness, headache, skin flushing, tachycardia, nausea, lightheadedness, weakness, and digital dysesthesia; aggravation of myocardial ischemia may rarely occur because of coronary steal, as discussed previously. Precapillary dilatation with reflex postcapillary constriction can increase capillary hydrostatic pressure and can cause peripheral edema. Verapamil and diltiazem are more likely than other

preparations to produce sinus bradycardia and atrioventricular block and may aggravate heart failure. Coughing, shortness of breath, wheezing, rashes, somnolence, constipation, seizures, and, rarely, depressed level of consciousness or psychosis may be caused by the CCBs. Possibly because of inhibition of calcium signaling in postsynaptic dopamine receptors, akathisia, parkinsonism, and delirium have occasionally been associated with use of calcium channel antagonists. Toxic serum concentrations of CCBs have not been defined. Possibly because of compensatory changes in calcium channels or CCB receptors, tachyphylaxis and withdrawal do not occur. CCBs do not alter exercise tolerance. Treatment of severe toxicity is with calcium, glucagon, volume expansion, and vasopressors.

Although inhibition of hyperactive intracellular calcium signaling may reduce a downstream effect of excitotoxicity in adult neurons, postnatal brain development involves increased intracellular calcium ion activity and animal studies suggest that CCBs that inhibit calcium signaling can injure neonatal neurons. The use of CCBs in very young children therefore requires caution.

DRUG INTERACTIONS

Adding CCBs to β-adrenergic blocking agents can result in additive cardiac slowing and atrioventricular block, whereas additive effects with α-adrenergic blocking agents may produce hypotension. Verapamil and nitrendipine increase plasma concentrations of digoxin (Lanoxin) and produce bradycardia, hypotension, or atrioventricular block when combined with quinidine (Cardioquin). Because verapamil and nifedipine inhibit cytochrome P450 (CYP) 3A4, they raise serum concentrations of CYP 3A4 substrates, such as carbamazepine and the triazolobenzodiazepines; this interaction can lead to carbamazepine neurotoxicity. Neurotoxicity also occurs occasionally when verapamil is combined with lithium. Additive suppression of the spontaneous rate of depolarization of the sinus node and of conduction through the atrioventricular node can increase the risk of bradycardia when lithium is combined with verapamil.

LABORATORY INTERFERENCES

As mentioned previously, verapamil raises serum levels of carbamazepine, digoxin, and other CYP 3A4 substrates. CCBs do not primarily affect other laboratory tests.

DOSAGE AND ADMINISTRATION

Verapamil is available in 40-, 80-, and 120-mg tablets. A 120-mg sustained release preparation used to treat hypertension was effective for mania in one report of three cases, but, because peak serum levels are not as high, it

generally is not as reliably effective for mood disorders as the immediate release forms. The usual daily doses of verapamil in the treatment of mania have been between 240 and 480 mg per day, although some patients may need higher doses because of low bioavailability. The short elimination half-life of verapamil requires three to four times per day dosing. Nimodipine is available only as a 30-mg capsule and is administered in doses of 90 mg per day for dementia and stroke. However, doses for bipolar illness have ranged from 360 mg to as much as 720 mg per day. The latter dose is substantially greater than the maximum dose approved by the FDA, although they are usually well tolerated. Nimodipine must be administered four to six times per day to sustain therapeutic levels. Isradipine, which is available as a tablet or capsule, is available in 2.5-, 5.0-, and 10.0-mg doses and is administered in doses of 2.5 to 20.0 mg per day in a frequently divided dose. Amlodipine, the only CCB that is effective when it is administered once daily, is available in 2.5-, 5.0-, and 10.0-mg tablets and is given in doses of 5 to 15 mg per day. However, there is little experience with this medication in mood disorders.

INVESTIGATIONAL DRUGS AND PREDICTION OF RESPONSE

Research into psychiatric applications of CCBs has been limited by relative lack of interest of manufacturers of these agents. Nimodipine and related dihydropyridine CCBs will probably turn out to have greater potential for psychiatric applications, because they cross the blood–brain barrier more readily than verapamil and they act on T- as well as L-type channels. Future development of CCBs for psychiatric indications will focus on antagonists of channels that are primarily expressed in the CNS such as N-channels, since such agents would not be expected to be limited by cardiovascular effects as could occur with high doses of L-channel antagonists. Such agents may be found to have applications in a variety of disorders characterized by neuronal hyperactivity. Investigations of medications with novel calcium antagonist properties may also prove fruitful. For example, as noted earlier, tamoxifen (Nolvadex), which is primarily used as an estrogen receptor antagonist in prevention of breast cancer, also inhibits PKC and has been found in preliminary trials to have antimanic properties.

Imaging of receptors for specific CCBs may prove useful in predicting which CCB might be effective for which patients. A study of 22 patients with bipolar or unipolar depression found that patients who responded to nimodipine had lower CSF somatostatin levels than patients who did not respond, and that nimodipine raised CSF somatostatin levels. However, this does not appear to be a practical approach to predicting treatment response. A positron emission tomography (PET) study finding that nimodipine responders had greater cerebral (including left insular) hypometabolism than nonresponders, whereas carbamazepine responders

had the opposite finding, may indicate a direction for further research into mechanisms of response but probably would not be used clinically. A 2-week trial in which manic patients received 360 mg per day of verapamil found a 61 percent reduction in manic symptoms and a correlation between the degree of improvement of mania and an increase in plasma calcium levels. A correlation has been suggested between the degree to which in vitro incubation with lithium lowered platelet $[Ca^{2+}]_i$ and predicted an antimanic action of lithium. Similarly, elevated platelet or lymphocyte $[Ca^{2+}]_i$ that is reduced by incubation with a CCB may turn out to be the best predictor of response to that class of medication.

References

Bachmann RF, Schloesser RJ, Gould TD, Manji HK. Mood stabilizers target cellular plasticity and resilience cascades. *Mol Neurobiol*. 2005;32:173–202.

Casamassima F, Hay AC, Benedetti A, Lattanzi L, Cassano GB, Perlis RH. L-type calcium channels and psychiatric disorders: a brief review. *Am J Med Genet B Neuropsychiatr Genet*. 2010;153B:1373–1390.

Catterall WA, Swanson TM. Structural basis for pharmacology of voltage-gated sodium and calcium channels. *Mol Pharmacol*. 2015;88:141–150.

Delisi SM, Konopka LM, O'Connor FL, Crayton JW. Platelet cytosolic calcium responses to serotonin in depressed patients and controls: relationship to symptomatology and medication. *Biol Psychiatry*. 1998;43:327–334.

DeLorenzo RJ, Sun DA, Deshpande LS. Cellular mechanisms underlying acquired epilepsy: the calcium hypothesis of the induction and maintenance of epilepsy. *Pharmacol Ther*. 2005;105:229.

Dubovsky SL, Buzan RD, Thomas M, Kassner C, Cullum CM. Nicardipine improves the antidepressant action of ECT but does not improve cognition. *J ECT*. 2001;17:3–10.

Dubovsky SL, Daurignac E, Leonard KE. Increased platelet intracellular calcium ion concentration is specific to bipolar disorder. *J Affective Disord*. 2014;164:38–42.

Dubovsky SL, Daurignac E, Leonard KE, Serotte JC. Levetiracetam, calcium antagonism, and bipolar disorder. *J Clin Psychopharmacol*. 2015;35:422–427.

Dubovsky SL, Lee C, Christiano J. Lithium decreases platelet intracellular calcium ion concentrations in bipolar patients. *Lithium*. 1991;2:167.

Dubovsky SL, Lee C, Christiano J, Murphy J. Elevated intracellular calcium ion concentration in bipolar depression. *Biol Psychiatry*. 1991;29:441–450.

*Dubovsky SL, Murphy J, Christiano J, Lee C. The calcium second messenger system in bipolar disorders: data supporting new research directions. *J Neuropsychiatry Clin Neurosci*. 1992;4:3–14.

*Dubovsky SL, Murphy J, Thomas M, Rademacher J. Abnormal intracellular calcium ion concentration in platelets and lymphocytes of bipolar patients. *Am J Psychiatry*. 1992;149:118–120.

Dubovsky SL, Thomas M, Hijazi A, Murphy J. Intracellular calcium signaling in peripheral cells of patients with bipolar affective disorder. *Eur Arch Psychiatry Clin Neurosci*. 1994;243:229–

234.

Emamghoreishi M, Schlichter L, Li PP, Parikh S, Sen J. High intracellular calcium concentrations in transformed lymphoblasts from subjects with bipolar I disorder. *Am J Psychiatry*. 1997;154:976–982.

Garza-Trevino ES, Overall JE, Hollister LE. Verapamil versus lithium in acute mania. *Am J Psychiatry*. 1992;149:121–122.

Goodnick PJ. Nimodipine treatment of rapid cycling bipolar disorder. *J Clin Psychiatry*. 1995;56:330.

Ikeda A, Kato T. Biological predictors of lithium response in bipolar disorder. *Psychiatry Clin Neurosci*. 2003;57:243–250.

Jaffe EH. Ca^{2+} dependency of serotonin and dopamine release from CNS slices of chronically isolated rats. *Psychopharmacology*. 1998;139:255–260.

Janicak PG, Sharma RP, Pandey GN, Davis JM. Verapamil for the treatment of acute mania: a double-blind, placebo-controlled trial. *Am J Psychiatry*. 1998;155:972–973.

Kato T, Ishiwata M, Mori K, Washizuka S, Tajima O. Mechanisms of altered Ca^{2+} signaling in transformed lymphoblastoid cells from patients with bipolar disorder. *Int J Neuropsychopharmacol*. 2003;6:379–389.

Ketter TA, Kimbrell TA, George MS, Willis MW, Benson BE. Baseline cerebral hypermetabolism associated with carbamazepine response, and hypometabolism with nimodipine response in mood disorders. *Biol Psychiatry*. 1999;46:1364–1374.

Mallinger AG, Thase ME, Haskett R, et al. Verapamil augmentation of lithium treatment improves outcome in mania unresponse to lithium alone: preliminary findings and a discussion of therapeutic mechanisms. *Bipolar Disord*. 2008;10:856–866.

*Manji HK, Chen G, Shimon H, Hsiao JK, Potter WZ. Guanine nucleotide-binding proteins in bipolar affective disorder. Effects of long-term lithium treatment. *Arch Gen Psychiatry*. 1995;52:135–144.

Nahorski SR. Pharmacology of intracellular signaling pathways. *Br J Pharmacol*. 2006;147:S38–S45.

Ostacher MJ, Iosifescu DV, Hay A, et al. Pilot investigation of isradipine in the treatment of bipolar depression motivated by genome-wide association. *Bipolar Disord*. 2014;16:199–203.

*Pazzaglia P, Post RM, Ketter TA, et al. Nimodipine monotherapy and carbamazepine augmentation in patients with refractory recurrent affective illness. *J Clin Psychopharmacol*. 1998;18:404–413.

Soares JC, Mallinger AG, Dippold CS, Frank E, Kupfer DJ. Platelet membrane phospholipids in euthymic bipolar disorder patients: are they affected by lithium treatment? *Biol Psychiatry*. 1999;45:453–457.

Suzuki K, Kusumi I, Sasaki A, Koyama T. Serotonin-induced platelet intracellular calcium mobilization in various psychiatric disorders: is it specific to bipolar disorder? *J Affect Disord*. 2001;64:291–296.

Triggle DJ. Calcium channel antagonists: clinical uses—past, present and future. *Biochem Pharmacol*. 2007;74:1–9.

Uemura T, Green M, Corson TW, Perova T, Li PP, Warsh JJ. Bcl-2 SNP rs956572 associated with disrupted intracellular calcium homeostasis in bipolar I disorder. *Bipolar Disorders*.

2011;13:41–51.

Wang H, Friedman E. Increased association of brain protein kinase C with the receptor for activated C kinase-1 (RACK1) in bipolar affective disorder. *Biol Psychiatry*. 2001;50:364–370.

Wisner KL, Peindl KS, Perel JM, Hanusa BH, Piontek CM. Verapamil treatment for women with bipolar disorder. *Biol Psychiatry*. 2002;51:745–752.

Yingling DR, Utter G, Vengalil S, Mason B. Calcium channel blocker, nimodipine, for the treatment of bipolar disorder during pregnancy. *Am J Obstet Gynecol*. 2002;187:1711–1712.

▲ 34.14 Carbamazepine

ROBERT M. POST, M.D., AND MARK A. FRYE, M.D.

Carbamazepine (Tegretol) is an iminostilbene drug possessing some structural similarity to the tricyclic antidepressant imipramine (Tofranil), but having a very different clinical spectrum of efficacy. It was approved for use in the United States for the treatment of trigeminal neuralgia in 1968 and for temporal lobe epilepsy (complex partial seizures) in 1974. Interestingly, carbamazepine was first synthesized as a potential antidepressant, but, because of its atypical profile in a number of animal models, it was initially developed for use in pain and seizure disorders. It is now recognized in most guidelines as a first- or second-line mood stabilizer useful in the treatment and prevention of both phases of bipolar affective disorder. A long-acting sustained release formulation (Equetro) was approved by the FDA for the treatment of acute mania in 2002.

The structure and chemistry of carbamazepine are closely related to the more recently introduced anticonvulsant oxcarbazepine (Trileptal). As illustrated in Figure 34.14–1, whereas carbamazepine has a double bond in its middle ring, oxcarbazepine has a keto-oxygen attached. These differences result in very different routes and profiles of drug metabolism, yet, apparently, a generally similar clinical spectrum of efficacy in seizure, pain, and mood disorders. Oxcarbazepine is not FDA approved for acute mania. Although double-blind comparisons showed equal response to oxcarbazepine and haloperidol (Haldol) or lithium (Eskalith), these results were contaminated by the concomitant use of high potency benzodiazepines. One study suggested better efficacy of oxcarbazepine in less severe mania compared with more severe mania, while carbamazepine is effective in severe forms of mania.

CHEMISTRY

Carbamazepine is oxidized by cytochrome P450 oxidase of the 3A4 subtype to an active 10,11-epoxide metabolite that also possesses both anticonvulsant and antinociceptive properties. This epoxide is stable but can be further hydrolyzed to an inactive dihydroxy compound via an enzyme called epoxide hydroxylase. In contrast, oxcarbazepine is metabolized directly to its monohydroxy metabolite, which is also metabolically and physiologically active. Dissimilar from carbamazepine, neither oxcarbazepine nor its active monohydroxy metabolite is readily oxidized by P450 liver enzymes, resulting in a lesser number of pharmacokinetic interactions with other drugs than with carbamazepine. These and other differences are highlighted in Table 34.14–1.

Carbamazepine is highly insoluble in water and difficult to dissolve in most preparations. It is commonly given in preclinical studies to animals in a solution of ethanol and propylene glycol, but no clinically available intravenous formulation exists.

Two long-acting preparations are now available. One of these preparations, extended release carbamazepine, needs to be taken with food, as its absorption depends on normal gastrointestinal transit time, and it leaves a residual capsule ghost. Another long-acting preparation (Carbatrol or Equetro) uses a microtrol delivery system with 25 percent uncoated intermediate release beads, which are absorbed almost as soon as the pill is swallowed; 40 percent extended release beads, which dissolve gradually over a period of 8 to 12 hours because of a polymer coating that allows carbamazepine to diffuse slowly through tiny micropores; and 35 percent enteric release beads, which release the drug very slowly as a result of special pH-sensitive coating.

Such a formulation makes the long-acting forms suitable for single bedtime administration in affectively ill patients (although this is not recommended in epilepsy).

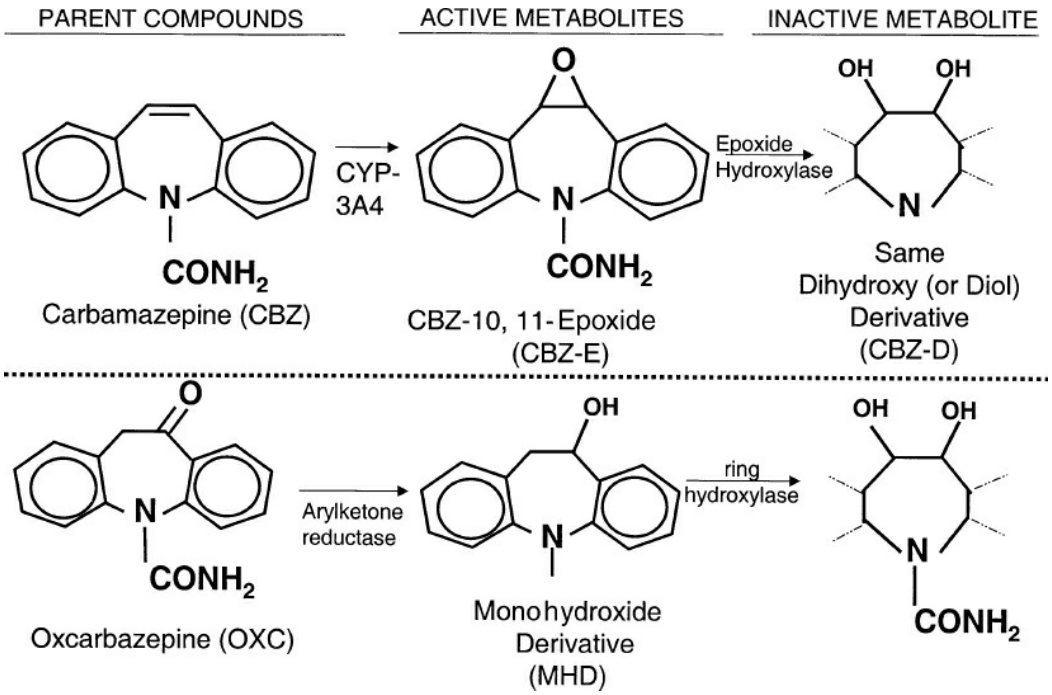

FIGURE 34.14–1. Comparative metabolism of carbamazepine and oxcarbazepine. Both parent and first metabolite are active anticonvulsants—the top two are potent enzyme inducers; the bottom two are not.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

Carbamazepine is absorbed slowly and somewhat erratically through the gastrointestinal tract. Bioavailability is 80 percent, and the drug should not

be exposed to high humidity because this can decrease bioavailability. Peak plasma levels are reached 2 to 8 hours after a single dose. Carbamazepine is 70 to 80 percent bound to plasma protein. The average half-life after a single dose is 26 hours, with a range of 18 to 54 hours. However, with long-term administration in humans, the half-life decreases to approximately 12 hours (range, 10 to 25 hours). This change occurs with the induction of hepatic P450 3A4 enzymes, which increase the rate of metabolism of carbamazepine itself (autoinduction) and a variety of other agents, as referred to in Table 34.14–2.

Table 34.14–1.
Comparative Effects of Carbamazepine (CBZ) and Oxcarbazepine (OXC)

	CBZ	OXC	Comment on OXC
Hepatic enzyme 3A4 induction	+++	+	No autoinduction
Decrease WBC	+++	±	No WBC suppression
Rash	6.5%	2.5%	Lower rate of rash[a]
Hyponatremia	+	++	More hyponatremia
LFT↑	+	±	Less liver enzyme elevation
Teratogenic	++	±	Lesser incidence of spina
Agranulocytosis	1/10,000	±	Lower incidence of hematologic reaction
Aplastic anemia			

[a]83/106 with benign rash on CBZ did not rerash on OXC. 3/3 with exfoliative rash did, however.

The degree of carbamazepine autoinduction is somewhat dose dependent but is usually complete after 3 to 5 weeks of treatment. Thus, doses that are not initially tolerated may become so after several weeks, and blood levels usually fall despite the use of the same daily dose, necessitating later dose increases to maintain similar blood levels and clinical effect. Long-term use of carbamazepine is associated with a markedly increased ratio of the amount of the 10,11-epoxide to carbamazepine. The half-life of the 10,11-epoxide is approximately 6 hours.

Table 34.14–2.
Drugs Whose Serum Concentrations Are Decreased by Carbamazepine Via Prominent Induction of Cytochrome P450 3A4

Drug Group	Drugs
Analgesics	Alfentanil; buprenorphine; fentanyl (?); levobupivacaine; methadone; tramadol
Anticoagulants	Dicoumarol (?); phenprocoumon; warfarin[a]
Antiepileptics	Carbamazepine[a]; ethosuximide; felbamate; lamotrigine[a]; oxcarbazepine; phenytoin; primidone; tiagabine; topiramate; valproic acid[b]; zonisamide
Antidepressants	Bupropion; citalopram; mirtazapine; tricyclic antidepressants
Antimicrobials	Caspofungin; doxycycline
Antipsychotics	Aripiprazole; clozapine; fluphenazine (?); haloperidola; lurasidone[a]; olanzapine; quetiapine; risperidone; thiothixene (tiotixene); ziprasidone

Antivirals	Delavirdine; protease inhibitors
Anxiolytics/	Alprazolam (?); buspirone
Sedatives	Clonazepam; midazolam; zopiclone (?)
Immunosuppressants	Cyclosporine (?); sirolimus (rapamycin); tacrolimus
Muscle relaxants	Doxacurium chloride; pancuronium bromide; rapacuronium bromide; rocuronium bromide; vecuronium bromide
Estrogen; steroids	Estrogen in hormonal contraceptives[a]; dexamethasone[b]; mifepristone; prednisolone
Stimulants	Methylphenidate[b]; modafinil
Others	Bepridil; oxiracetam (?); paclitaxel; quinidine; remacemide (?); repaglinide; theophylline (?); thyroid hormones

[a]Clinical most important.

[b]Important.

(?) Signifies doubtful evidence/clinical significance.

Carbamazepine is metabolized in the liver and then excreted by the kidneys, as only 1 percent is eliminated by biliary excretion. Oxcarbazepine is a white to faintly orange crystalline powder and, like carbamazepine, is practically insoluble in ethanol, ether, or water. It is slightly soluble in chloroform, acetone, and methanol.

Because there is no or minimal autoinduction with oxcarbazepine, its half-life remains relatively stable over time. The half-life of oxcarbazepine is approximately 2 hours, and, because of the rapid conversion to the monohydroxide, which has a half-life of approximately 9 hours, the monohydroxide is responsible for most of the anticonvulsant activity. Dosing is usually twice or three times daily. There is an FDA approved long-acting preparation of oxcarbazepine (Oxtellar) for complex partial seizures, but efficacy in bipolar disorder has not been demonstrated. Oxcarbazepine is 96 percent absorbed and 60 percent bound to plasma proteins. Its monohydroxy derivative is 40 percent bound to plasma proteins. The monohydroxy derivative, eslicarbazepine, was FDA approved for seizures in 2013, although the results of earlier trials in mania have not been reported.

PHARMACODYNAMICS

Mechanisms of Antinociceptive, Antimanic, and Antidepressant Action

The onset of anticonvulsant and antinociceptive effects of carbamazepine is rapid and can occur after a single dose in appropriate animal model studies, or after 24 to 48 hours in most clinical studies of paroxysmal pain syndromes, such as trigeminal neuralgia (Fig. 34.14–2). Its antimanic actions can also be relatively rapid in onset, beginning within the first several days of treatment, but several weeks are typically required for maximum treatment response in patients. In contrast, the antidepressant effects of carbamazepine are often delayed, with 3 to 4 weeks required

before maximum effects are observed. However, because carbamazepine can be relatively sedating, an improvement in sleep may occur much earlier, yet as with other antidepressant modalities, the degree of improvement in sleep is not correlated with the degree of subsequent antidepressant response.

In assessing mechanism of action related to clinical effects, the actions of carbamazepine at clinically relevant blood levels and in the appropriate time frame should be considered, as schematized in Figure 34.14–2. Immediate effects would be more likely related to the anticonvulsant and antinociceptive profile of the drug, whereas effects requiring more chronic administration would be those more likely linked to the parallel time course of antidepressant effects of carbamazepine. The anticonvulsant effects of carbamazepine have been most closely linked to blockade of type 2 or batrachotoxin-sensitive sodium channels. When these channels are inactivated, they tend to decrease release of excitatory amino acids such as glutamate and inhibit sustained rapid neuronal firing in a variety of in vitro models. This action of sodium channel blockade is shared by oxcarbazepine, as well as a variety of other anticonvulsant compounds, including phenytoin (Dilantin), lamotrigine (Lamictal), and topiramate (Topamax), and, to a lesser extent, valproate (Depakene; Depakote). Although these are acute effects of the drug and are associated with decreased neurotransmitter release, Robert Belmaker and colleagues have postulated that the potency of sodium channel blockade is related to the antimanic efficacy of the anticonvulsant compounds as a group. This potential relationship, which is discordant for the time frame of clinical action and violated by topiramate (which is not antimanic), remains to be further investigated.

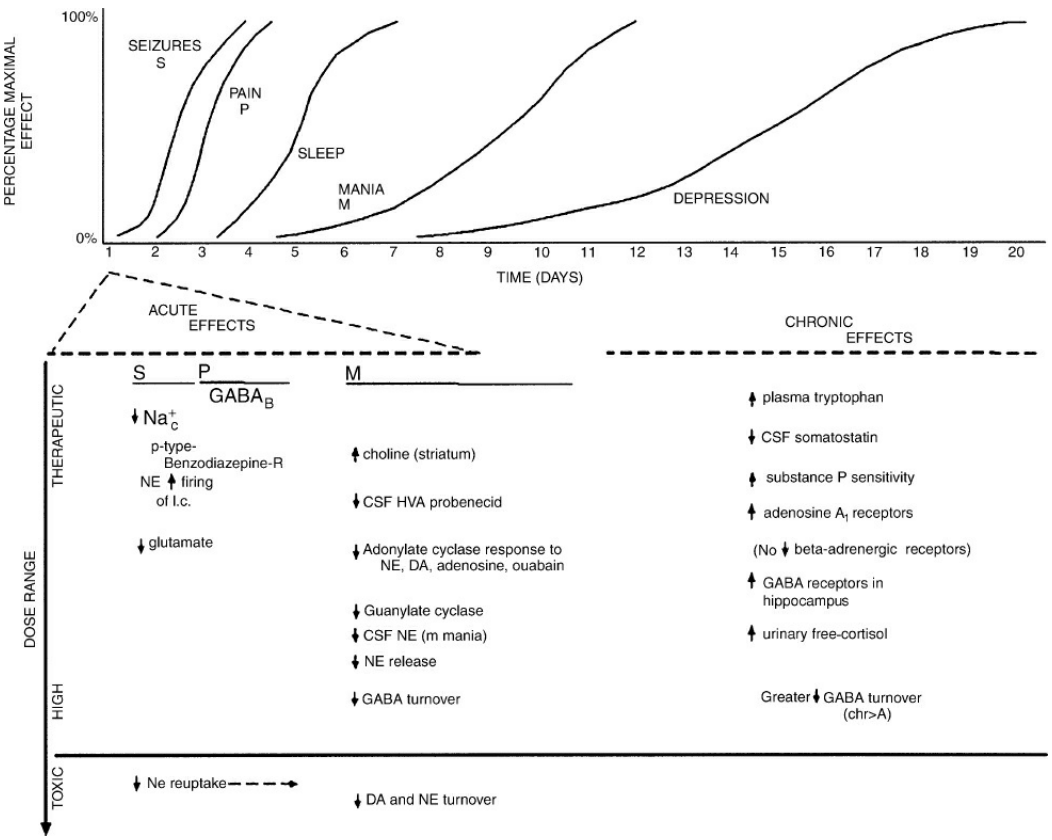

FIGURE 34.14–2. Schematic and time course of clinical and biochemical effects of carbamazepine. Seizure(s) and paroxysmal pain(p) syndromes respond rapidly to carbamazepine, whereas maximal effects in mania(m) and depression are delayed 2 to 3 weeks or more, respectively. The acute effects observed in animal and humans are probably related to anticonvulsant and antinociceptive efficacy, whereas chronic effects are more likely to be related to antidepressant mechanisms. Candidate mechanisms for each clinical effect are listed according to their occurrence at clinically relevant concentrations. Chr>A, chronic greater than acute; CSF, cerebrospinal fluid; DA, dopamine; GABA, γ-aminobutyric acid; HVA, homovanillic acid; l.c., locus ceruleus; NE, norepinephrine.

The acute anticonvulsant effects of carbamazepine have been closely linked to those that can be readily studied in a variety of animal models, including amygdala-kindled seizures. Carbamazepine appears to exert its actions on amygdala-kindled seizures through the so-called peripheral-type or mitochondrial benzodiazepine receptor. It is at this receptor that the peripheral-type ligand Ro 5-4864 is proconvulsant and an agonist, whereas PK1195 is anticonvulsant and an antagonist. Carbamazepine appears to share PK1195-like antagonist activities at the peripheral-type benzodiazepine receptor.

In contrast to the anticonvulsant effects of diazepam (Valium), which are reversed by a pure central benzodiazepine antagonist such as Ro 15-1788 or flumazenil (Romazicon), carbamazepine actions are reversed by Ro 5-4864 in a PK1195-blockable fashion, strongly implicating this peripheral-type benzodiazepine receptor in carbamazepine's anticonvulsant actions on amygdala-kindled seizures (as illustrated in Fig. 34.14–4A). Whether this peripheral-type effect is related to any of carbamazepine's psychotropic

properties remains to be further investigated.

The peripheral-type benzodiazepine receptors have been linked to cholesterol transport and calcium influx through mitochondrial membranes. Cholesterol is converted to pregnenolone in mitochondria, and this compound subsequently becomes transformed into neuroactive steroids allopregnanolone and allotetrahydro-DOC (which enhance γ-aminobutyric acid type A [$GABA_A$] receptors) and pregnanolone sulfate (which inhibits TBPS [t-(35)S]butyl-bicyclophosphorothionate] binding and voltage gated Ca^{++} channels, but enhances NMDA receptors) and dehydroepiandrosterone sulfate (which inhibit the barbiturate binding site). Thus, it remains possible that some of carbamazepine's actions initiated at peripheral-type benzodiazepine receptors may be achieved indirectly at other receptors through these effects on neurosteroid metabolism.

In addition to its blockade of glutamate release, carbamazepine is also potent at adenosine A_1 receptors. Its agonist properties at these receptors could account for not only some of its anticonvulsant properties but also its sedative properties.

It is also of interest that a variety of mood stabilizers beyond carbamazepine, such as lithium, valproate, and lamotrigine, are able to weakly block calcium influx through glutamate receptors of the NMDA subtype. Thus, the possibility remains that carbamazepine could exert relatively antiglutamatergic effects both by decreased release of glutamate and by relative decreases in glutamate's postsynaptic efficacy by inhibiting calcium influx. Glutamate NMDA receptor antagonist actions in general have been increasingly linked to potential antidepressant (ketamine) and possible mood-stabilizing effects (memantine), but further investigation of this possible linkage to carbamazepine's actions is required.

In common with both lithium and valproate, carbamazepine blocks inositol transport at the myoinositol transporter. Other common and differential effects of carbamazepine compared with lithium and valproate are illustrated in Figure 34.14–3.

The antinociceptive effects of carbamazepine have been most closely linked to effects at $GABA_B$ or baclofen-like receptors. Carbamazepine and baclofen share structural similarities, and the antinociceptive actions of both drugs are specifically inhibited by $GABA_B$ antagonists. It is also of interest that chronic but not acute administration of lithium, carbamazepine, and valproate are all associated with upregulation of $GABA_B$ receptors in the hippocampus, yielding this action as a potential convergent mechanism for mood stabilization. In contrast, unimodal antidepressants appear to upregulate $GABA_B$ receptors in frontal cortex.

Carbamazepine also exerts a variety of interesting effects on aminergic and indolaminergic systems. Carbamazepine is a very weak blocker of

norepinephrine reuptake, which is probably not clinically relevant, but it does decrease induced release of norepinephrine at clinically relevant concentrations. These preclinical findings converge with carbamazepine's ability to decrease CSF norepinephrine in manic patients.

Actions at noradrenergic α_2 receptors are of special interest in that the α_2 agonist in clonidine (Catapres) reverses some of carbamazepine's effects in some seizure models, whereas, paradoxically, the α_2 antagonist yohimbine reverses carbamazepine's anticonvulsant effects in other models. In epilepsy-prone animals, carbamazepine's ability to increase serotonin in the hippocampus appears to be important, as depletion of serotonin by several different mechanisms is associated with a loss of carbamazepine's anticonvulsant effects in this model.

Consistent with other mood-stabilizing drugs (lithium and valproate), carbamazepine is also able to decrease dopamine turnover, an action that could be related to its antimanic and mood-stabilizing properties. In contrast to the antipsychotics that directly block dopamine receptors, carbamazepine is not a direct antagonist at dopamine receptors and exerts its actions on the dopamine system through some other series of mechanisms. Dopamine levels have been reported to be increased in the prefrontal cortex after chronic carbamazepine administration. Carbamazepine does decrease levels of the dopamine metabolite homovanillic acid (HVA) in the CSF of affectively ill patients following probenecid administration, consistent with its ability to decrease dopamine turnover in animals.

Taken together, these data suggest that different mechanisms of carbamazepine's actions could be related differentially to its anticonvulsant efficacy in different seizure models. To the extent that this is clearly delineated in epilepsy research, as, for example, in relationship to α_2 receptors, it raises the possibility that different components of carbamazepine's actions could be related to different aspects of its efficacy and side effect profile in the affective disorders.

In addition to those mentioned above, carbamazepine's actions on a variety of neuropeptides appear to require chronic administration and, thus, could be pertinent to antidepressant actions. Chronic, but not acute, drug administration increases substance P levels and sensitivity in the striatum, findings of particular interest, as they are shared by lithium. Substance P antagonists have been studied in depression, but have a mixed record for effectiveness.

Carbamazepine decreases levels of somatostatin in CSF, in contrast to nimodipine (Nimotop; Nymalize), which increases CSF somatostatin. These differential effects of carbamazepine and nimodipine exist not only in terms of their blockade of calcium influx via NMDA type glutamate receptors versus L-type calcium channels, respectively, but also in terms of

their physiological effects on brain metabolism, as assessed by PET. Ketter and associates have found a differential relationship of metabolism in paralimbic areas, particularly in the left insula, to the degree of antidepressant response to carbamazepine, with higher baseline levels of activity being associated with better antidepressant response. In contrast, low levels of metabolism in the left insula are associated with better antidepressant response to the dihydropyridine calcium channel blocker nimodipine. In both instances, response to the drug was associated with a relative normalization of the baseline abnormalities in the activity of insula and related paralimbic areas of brain.

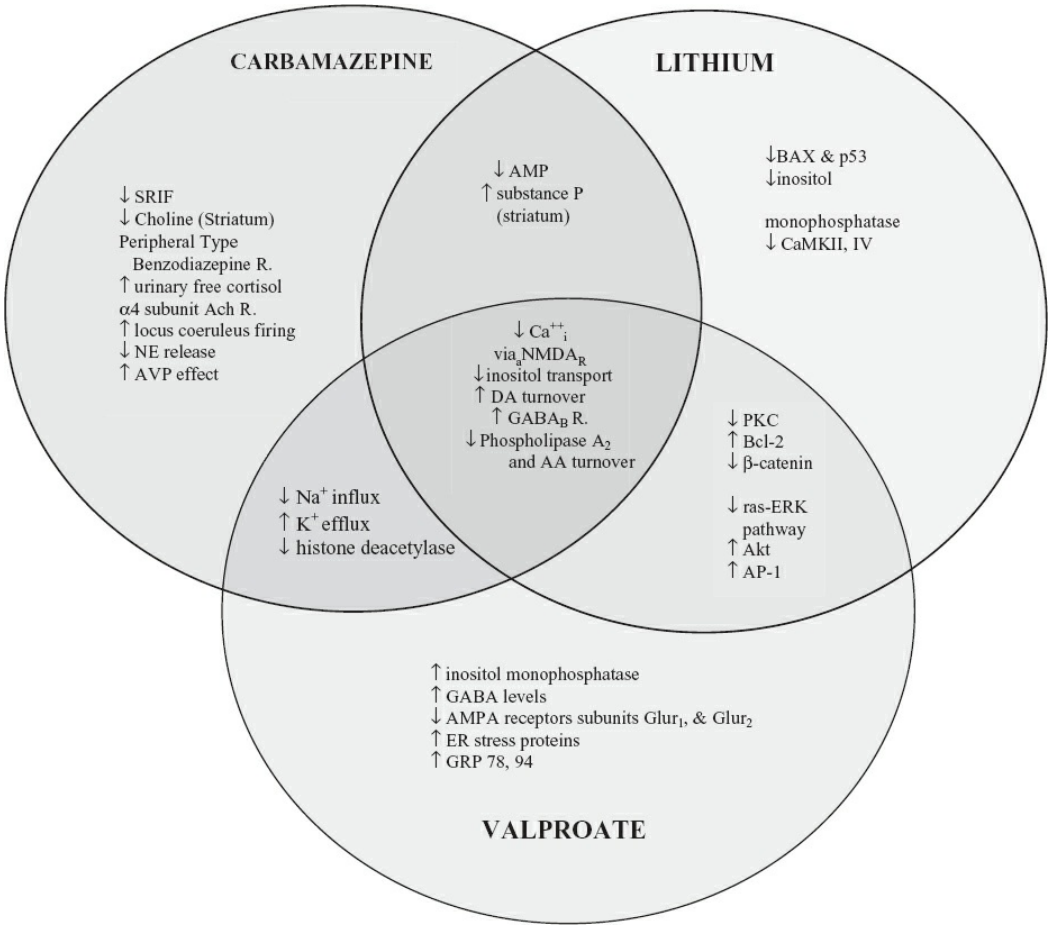

FIGURE 34.14–3. Overlapping, differential, and unique effects of mood stabilizers. SRIF, somatostatin; R, receptors; NE, norepinephrine; AVP, arginine vasopressin; NMDA, *N*-methyl-D-aspartate; cAMP, cyclic adenosine monophosphate; PKC, protein kinase C; GABA, γ-aminobutyric acid; ER, endoplasmic reticulum.

Carbamazepine's antidepressant effects may be mediated by those that are atypical of many other antidepressant modalities. For example, most antidepressants decrease β-adrenergic receptors in the prefrontal cortex, whereas carbamazepine upregulates them. Most antidepressants upregulate mineralocorticoid and glucocorticoid receptors and, thus, decrease cortisol, whereas carbamazepine increases plasma and urinary

free cortisol. This unusual profile of carbamazepine may explain why it is effective in some depressed patients with affective syndromes that are resistant to more traditional antidepressant modalities. A definitive link of carbamazepine to antidepressant response in depressed and PTSD patients with low basal cortisol or increased dexamethasone suppression has not been adequately studied, however.

As noted earlier, carbamazepine has both convergent and contrasting effects with other mood stabilizers, such as lithium and valproate (see Section 34.34, Table 34.34–1 in the section on valproate). Whereas lithium and carbamazepine block cAMP and G proteins, their effects on phosphoinositide turnover diverge. Lithium inhibits several inositol phosphatases, whereas valproate does not, and carbamazepine may enhance their activity. This contrasts with the common effects of the three drugs in blocking the inositol transporter and in decreasing arachidonic acid turnover. Carbamazepine, as well as the mood stabilizers lithium, valproate, and most antidepressant modalities increase BDNF. This action may be of considerable importance because the Val66 met allele of pro-BDNF and early life adversity can reduce hippocampal volume and yield low levels of BDNF in hippocampus and cortex in adulthood. Stressors, affective episodes (both mania and depression), and substances of abuse all can lower BDNF in hippocampus or serum of bipolar patients. As such, effective prophylaxis will indirectly prevent episode-related BDNF decrease, and mood stabilizer treatment may also directly increase BDNF and help correct persisting BDNF decreases.

Whereas lithium is a blocker of vasopressin effects via its inhibition of the second messenger cAMP, carbamazepine appears to exert indirect facilitatory effects at the vasopressin receptor. Therefore, because carbamazepine's effects are at the level of the receptor and lithium's blockade is below the receptor (in the transduction system), carbamazepine is unable to reverse the effects of lithium on the development of diabetes insipidus.

Carbamazepine, like phenytoin and lamotrigine, is not effective in blocking the initial developmental phase of electrical kindling of the amygdala, but is a highly effective anticonvulsant in the midphase of full-blown kindled seizures. Conversely, chronic treatment with carbamazepine does block the initial or developmental phase of pharmacological kindling with the local anesthetics lidocaine (Xylocaine) or cocaine. Chronic carbamazepine should thus make the development of cocaine-related seizures in humans less likely as well. This effect could contribute to carbamazepine's utility in helping with cocaine abstinence, an effect that has only been demonstrated compared with placebo in those with concomitant affective disorders.

Mechanisms of Tolerance Development

It is of interest that carbamazepine is acutely effective against completed amygdala-kindled seizures, but tolerance rapidly develops to this effect with repeated daily administration (Fig. 34.14–4A). This form of tolerance is called contingent tolerance as it develops only when the drug is given before the amygdala stimulation, but not when the drug is given immediately after the stimulation. In this way tolerance is contingent on having the drug present during the seizure-inducing stimulation and is clearly based on pharmacodynamics and not pharmacokinetic mechanisms. This tolerance that develops to the anticonvulsant effects of carbamazepine may provide an indirect but useful model for considering mechanisms potentially associated with clinical tolerance development in seizure and mood disorders and ways of slowing its development or reversing it once tolerance has occurred.

A gradual loss of efficacy after many years of successful treatment has been reported in a subgroup of patients via an apparent pharmacodynamic tolerance mechanism in patients with: paroxysmal pain syndromes such as trigeminal neuralgia, the prophylaxis of bipolar illness (Fig. 34.14–5), and to a lesser extent in clinical epilepsy syndromes. Some of the ways of slowing tolerance development in the preclinical seizure model could be tested clinically for their relevance to human therapeutics, either in epilepsy patients specifically or in other syndromes such as bipolar disorder as well.

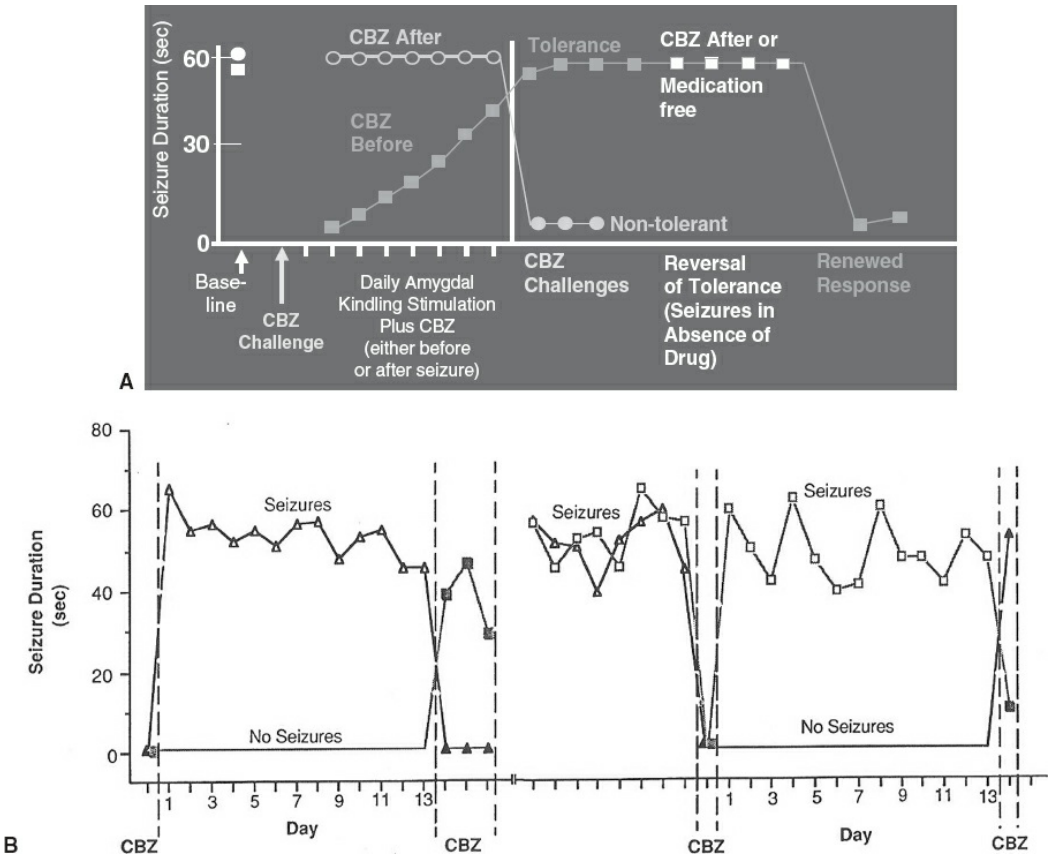

FIGURE 34.14–4. **A:** Schematic illustration of contingent tolerance to carbamazepine (CBZ): development and reversal. In fully kindled animals (*open circles, squares*), carbamazepine treatment inhibits kindled seizures (*filled squares*). Repeated drug administration before (*filled squares*), but not after (*filled circles*), stimulation results in tolerance development. Tolerance induced in this matter (*filled squares*) can be reversed by a period of kindled seizures with drug administration after each seizure (*open squares, dotted lines*). **B:** Seizures enable carbamazepine's anticonvulsant effect. *Left side:* In fully kindled animals, CBZ was administered (15 mg/kg, i.p.) and produced complete seizure suppression (CBZ-first day). Animals were then divided into two groups that received either 13 daily amygdala-kindled seizures (at 800 μA [*open triangles*]) or no treatment. Following this, both groups of animals received 3 days of kindling stimulation preceded by CBZ administration on each day (15 mg/kg). The rats that did not experience seizures for the 13-day interval showed only a partial response to CBZ (*closed squares*), whereas the rats that underwent seizures again showed complete seizure suppression by CBZ. Both groups were then given repeated seizures (*open squares and triangles*) to reinstate the original CBZ effect. *Right side:* A crossover design was used to replicate the initial finding. Animals that previously were given no treatment now received 13 days of seizures (*open squares*) and vice versa. Again, and even more robustly, the animals that did not experience seizures for 13 days showed no anticonvulsant response to CBZ administration (day 14), whereas the animals experiencing seizures showed almost complete seizure suppression.

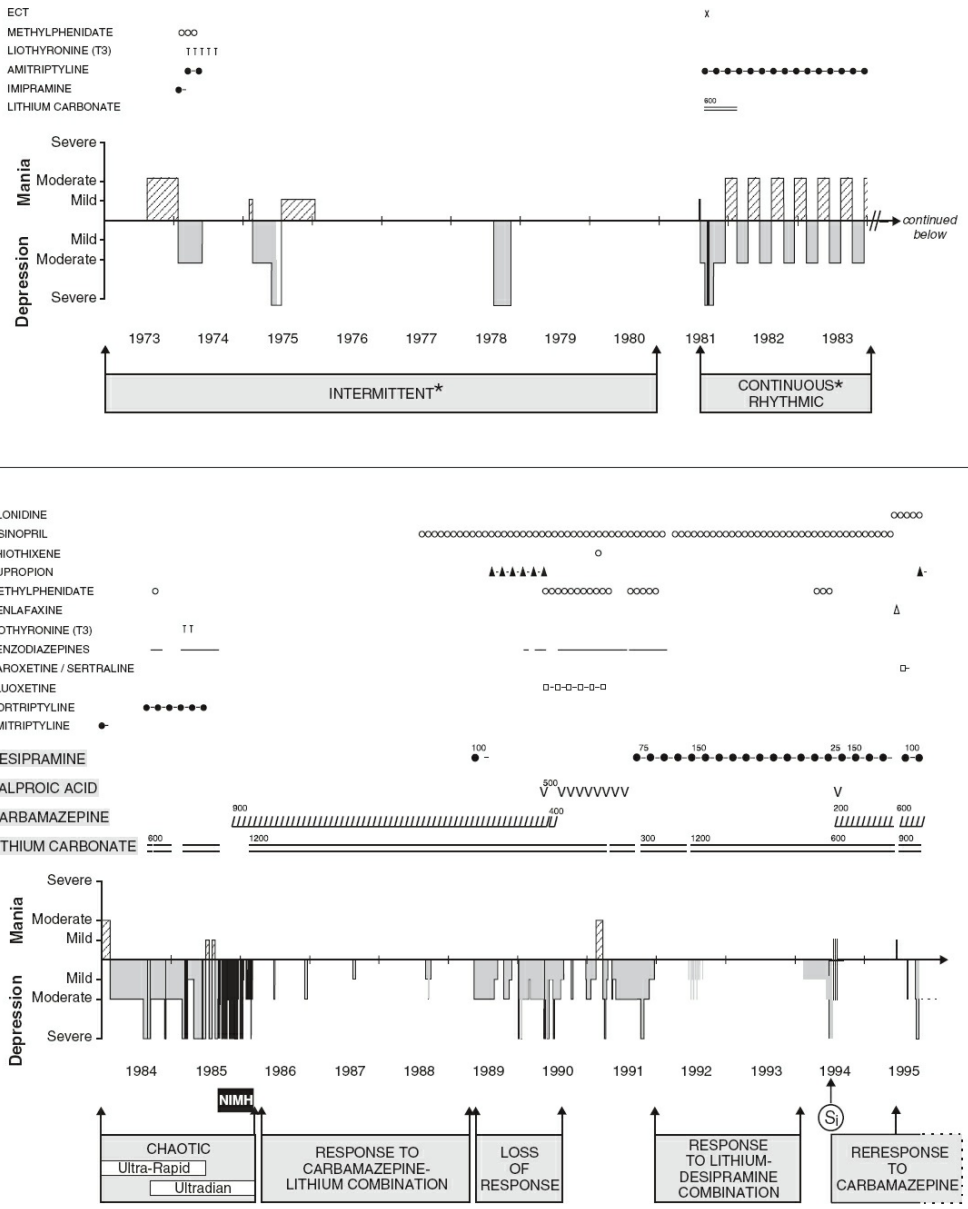

FIGURE 34.14–5. Life chart showing the phases in evolution and treatment response in a female patient with bipolar II disorder. This life chart illustrates the phenomena of illness progression, probable antidepressant-induced cycle acceleration (m 1981), lack of response to lithium in a rapid cycler (1984–5), response to the combination therapy of lithium and carbamazepine, development of loss of responsiveness to this treatment regimen via tolerance (1989), and cross-tolerance to valproate (1990). Key: Lithium (= = = =); carbamazepine (//////); valproic acid (VVVV); tricyclics (-.-.-.); benzodiazepines (–.–.–.); thyroid (TTTT); other (oooo). CBZ, carbamazepine; Dep, depression; ECT, electroconvulsive therapy; VPA, valproic acid.

As noted above, the tolerance that develops to carbamazepine's anticonvulsant effects on amygdala-kindled seizures is not related to pharmacokinetics or a drop in carbamazepine blood levels. This can be demonstrated by the fact that tolerance develops only when the drug is given to the animal just before each amygdala stimulation, but not when the same dose of drug is administered immediately after each kindled seizure has occurred. Animals administered daily carbamazepine just after

each seizure has occurred (and drug levels again become negligible prior to the next stimulation the next day) never become tolerant. In this fashion, tolerance development is contingent on the drug's being in the brain concurrently with the amygdala stimulation. This type of contingent tolerance can be reversed either by a period off the drug (and the occurrence of seizures in the medication-free state) or, remarkably, by a period of continuing to give the drug daily, but now after (rather than before) each of a series of several once-daily seizures has occurred.

The neurobiology of this anticonvulsant tolerance development has been preliminarily explored in the kindled seizure model and appears to involve differential regulation of a variety of seizure-induced adaptations at the level of gene expression. Seizures ordinarily induce a variety of neurotransmitter and receptor effects that have been linked either to the primary pathology of kindled seizure progression or to more short-lived secondary compensatory or endogenous anticonvulsant effects (Table 34.14–3, left column). For example, thyrotropin-releasing hormone (TRH) peptide and its messenger ribonucleic acid (mRNA) are increased for approximately 4 days after an amygdala-kindled seizure. However, TRH does not increase after a seizure that occurs when animals are tolerant to carbamazepine.

TRH is likely an anticonvulsant adaptation based on the finding that its bilateral injection of TRH into the hippocampus decreases amygdala-kindled seizures. Further supporting this perspective are the observations that after a period of 4 or more days after the last amygdala-kindled seizure, endogenous TRH levels and other putative endogenous anticonvulsant adaptations dissipate and carbamazepine loses its anticonvulsant effectiveness (even in nontolerant animals) (Fig. 34.14–4B; day 14). These data suggest that TRH is one example of many endogenous substances that are induced postictally in an attempt to prevent further seizures from occurring and at the same time are also required for carbamazepine's efficacy. Thus, the failure of seizures to induce TRH and other anticonvulsant peptides in carbamazepine tolerant animals could be relevant to this loss of efficacy. TRH also has antidepressant effects when administered intravenously or intrathecally to depressed patients. Therefore, it remains a possibility to be further studied that loss of depressive-episode induced increases in TRH could be related to clinical tolerance development in the mood disorders.

Similarly, in carbamazepine-tolerant animals, the breakthrough seizures selectively fail to induce other endogenous adaptations, such as increases in $GABA_A$ receptors, but increases in benzodiazepine receptors continue to occur (Table 34.14–3, right columns). The inability of GABA receptors to show their normal increase after a seizure is also observed at the level of the $GABA_A$ receptor subunit mRNAs. The mRNA for the α_4 subunit is one of several subunits that ordinarily increases after seizures,

but it fails to do so in the carbamazepine-tolerant state. Thus, in addition to an anticonvulsant neuropeptide such as TRH, transient seizure-induced increases in other inhibitory neurotransmitter receptors, such as those of $GABA_A$, and its α_4 subunit are also selectively suppressed during carbamazepine tolerance, which could contribute to its loss of anticonvulsant efficacy. That is, some of the needed episode-induced adaptations fail to occur.

These are just two examples of several neural systems that are selectively impaired after seizures in carbamazepine-tolerant animals but are consistently induced by seizures in nontolerant animals that have been exposed to the same number of prior seizures and the same amount of carbamazepine (but given after the seizure has occurred) (Fig. 34.14–4A). Also, supporting this view that seizure-induced adaptations are important to carbamazepine's efficacy is the observation that carbamazepine again becomes effective in tolerant animals, after a series of seizures are given in the absence of medication (which then re-induce the full range of adaptive effects on gene expression).

Clinical case studies suggest that this renewal of efficacy may also be seen after a period of time off medications in the instance of clinical tolerance development in seizure or mood disorders, but it would be preferable to find a new treatment agent that did not show cross-tolerance to carbamazepine so as to avoid the necessity of experiencing further episodes while off medication.

Table 34.14–3.
Selective Failure of Some Kindled Seizure-Induced Neurochemical Changes in CBZ-Contingent Tolerance

Seizure-Induced Adaptation (in Nontolerance Animals)	Seizure-Induced Adaptations Present in CBZ-Tolerant Animals	Seizure-Induced Adaptations Absent in CBZ-Tolerance Animals
↑ c-fos mRNA		No ↑ c-fos
↑ DZP receptors	↑ DZP-R	
↑ GABA receptors (3H) muscimol		No ↑ GABA-R
↑ Subunit		No ↑ subunit
↑ Subunit	↑ subunits	
↑ TBPS binding		No ↑ TBPS
↑ Glucocorticoid-R mRNA		No ↑ Glucocorticoid-R
↑ Mineralocorticoid-R mRNA	↑ Mineralocorticoid-R	
↑ BDNF mRNA		No ↑ BDNF
↑ NT_3 mRNA	↑ NT_3	
↑ TRH mRNA		No ↑ TRH
↑ CRH mRNA		No ↑ CRH
↑ CRH-BP mRNA		No ↑ CRH-BP
↑ NPY mRNA		No ↑ NPY

↑ Enkephalin mRNA		No ↑ Enkephalin
↑ Dynorphin mRNA	↑ Dysmorphia	

Treated with CBZ after each daily amygdala kindling stimulation; nontolerance animals were matched for amount of drug and number of seizures seen in tolerant animals.

CBZ, carbamazepine; DZP, diazepam; GABA γ-aminobutyric acid; TBPS (35S) t-butyl-bicyclophosphorothionate; BDNF, brain-derived neurotrophic factor; NT_3, neurotrophin-3; TRF, thyrotropin-releasing hormone; CRH, corticotropin-releasing hormone; CRH BP, corticotropin-releasing hormone-binding protein.

In the preclinical model of anticonvulsant tolerance to carbamazepine on amygdala-kindled seizures, cross-tolerance is observed to the anticonvulsant effects of PK11195 (a peripheral type of benzodiazepine antagonist), lamotrigine, and, surprisingly, valproate and levetiracetam (Keppra), but not to diazepam or phenytoin. How these cross-reactivities for seizure tolerance in the rodent relate to those in the clinical condition of recurrent affective disorders remains to be further studied. Finding ways to address clinical tolerance development is potentially important because a substantial group of treatment-refractory patients (perhaps some 20 to 40 percent) who initially experience an excellent prophylactic response to carbamazepine begin to show breakthrough episodes consistent with a tolerance process, which often begins after a period of 2 to 3 years of effective treatment (see years 1986 to 1988 in Fig. 34.14–5).

Interestingly, as in the seizure model of cross-tolerance between carbamazepine and valproate, the patient illustrated in Figure 34.14–5 failed to respond to valproate when she was becoming tolerant to carbamazepine. Further exploration of ways to circumvent tolerance development and ascertain which drugs do not possess cross-tolerance is of considerable interest. A number of approaches have been examined that slow tolerance development in the amygdala kindled seizure model. These include using sustained higher rather than marginal and then increasing doses of drug, combining carbamazepine with valproate, treating earlier rather than later in the course of kindling, and reducing the stimulation intensity (or lowering illness drive). Whether some of these are also pertinent to slowing tolerance development during clinical treatment of seizure and mood disorders remains to be studied systematically, but they are listed here for their potential heuristic value.

CLINICAL DRUG STUDIES

Acute Mania

Numerous double-blind studies of various designs have compared carbamazepine to placebo or other active agents that indicate antimanic efficacy. Two large placebo-controlled, parallel-group design trials have recently revealed that the long-acting formulation of carbamazepine

(Equetro) was significantly more effective than placebo in acute mania. Similar acute antimanic effects of carbamazepine that roughly parallel those of lithium and antipsychotics are illustrated in Figure 34.14–6.

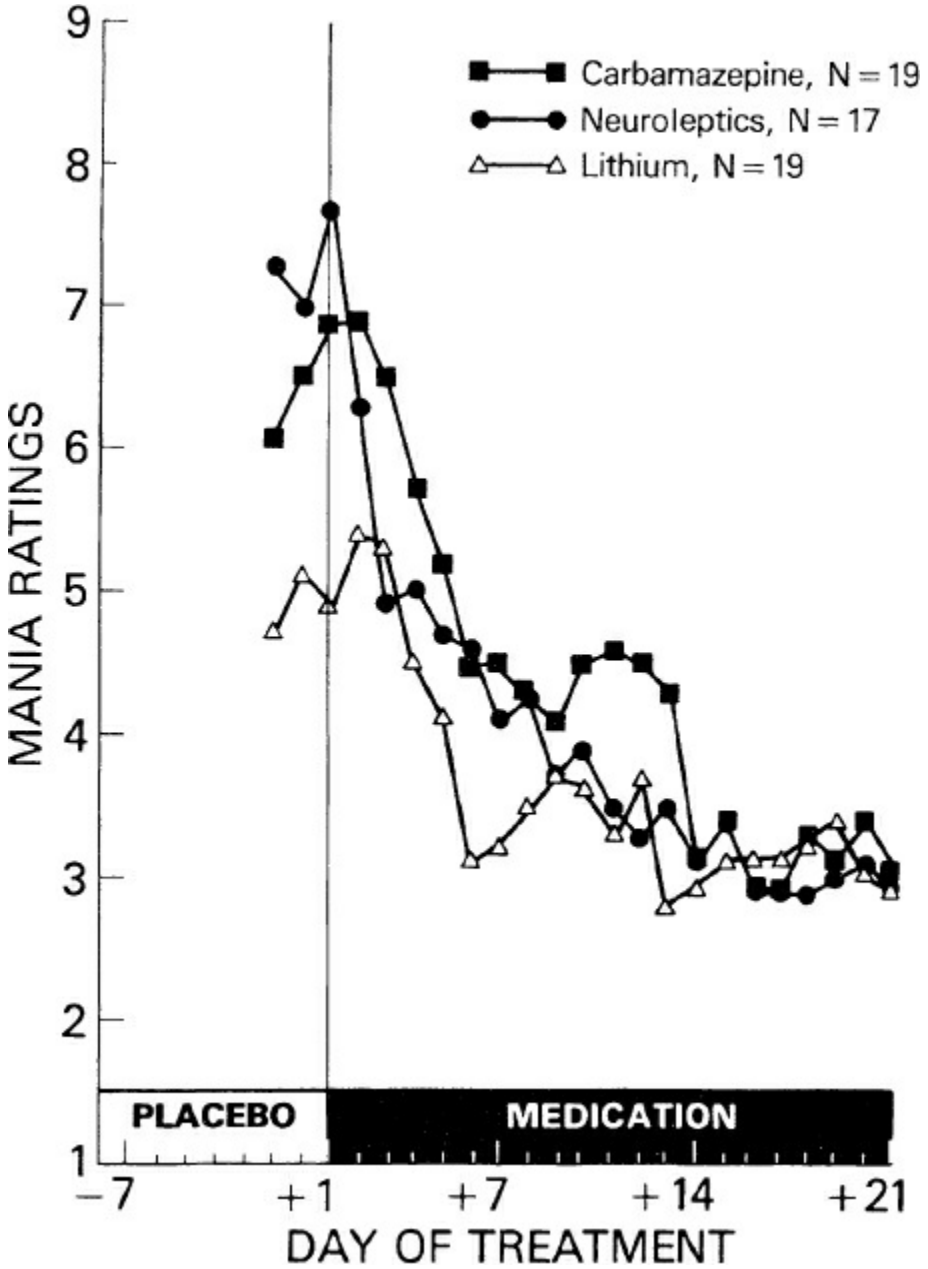

FIGURE 34.14–6. Time course of double-blind acute antimanic effects of carbamazepine compared to lithium and neuroleptics. The incidence and time course of antimanic effects of carbamazepine appear comparable with those achieved by neuroleptics, although several studies have suggested a slightly faster onset of efficacy of neuroleptics in the first 3 days of treatment, a difference that was gone by 7 days. Even patients requiring seclusion for acute, fulminant, psychotic manic episodes may rapidly improve after initiation of treatment to carbamazepine alone.

Moreover, acute antimanic responsivity has been documented in double-blind, off-on–off-on study designs as illustrated in Figure 34.14–7, in a patient who was a clear nonresponder to lithium, as revealed by a

breakthrough episode during carefully monitored lithium treatment upon hospital discharge. The patient was readmitted in a full-blown manic episode despite lithium treatment and showed a third excellent re-response to active carbamazepine and, subsequently, has been able to remain out of the hospital during long-term prophylaxis with carbamazepine as supplemented by other agents. Response to carbamazepine but not to valproate has also been seen in some patients (Figs. 34.14–3 and 34.14–9). Other patients with bipolar disorder have shown the opposite pattern of clear-cut responses to valproate, but not carbamazepine, emphasizing the principle that response to one anticonvulsant is not predictive of response to another.

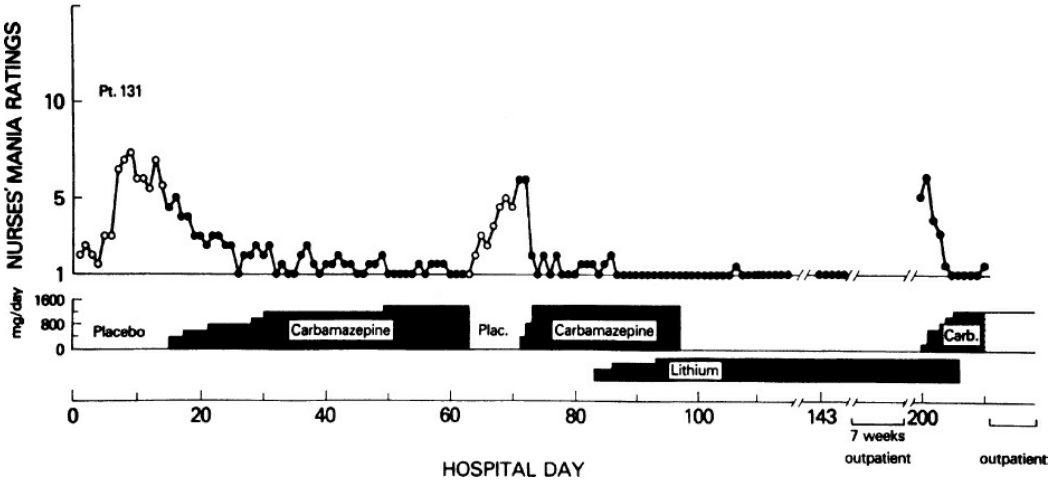

FIGURE 34.14–7. Unequivocal antimanic response to carbamazepine in a double-blind, off-on–off-on clinical trial, a subsequent manic episode on lithium, and a third "on" to carbamazepine.

Table 34.14–4.
Response[a] Rates in Bipolar Depression (Zhang et al., 2007)

LOCF Analysis

Placebo	CBZ[b]	CBZ + FEWP[bc]
34.8%	63.8%	84.8%
(*N* = 23)	(*N* = 47)	(*N* = 46)

Observed Cases

Placebo	CBZ	CBZ + FEWP[b]
40.0%	61.1%	82.5%
(*N* = 15)	(*N* = 36)	(*N* = 40)

[a]≥50% reduction in Ham-D scores after 12 weeks.

[b]$p < 0.05$ vs. PBO.

[c]Preparation of 11 herbal materials: 4 dry roots; 1 fresh root; 1 dry ripe fruit; 1 dry root bark; 1 peeled dry root; 1 dry root and stem; 1 dry root and leaf; and 1 dry sclerotium.

Acute Depression

Fewer studies have addressed the acute antidepressant effects of carbamazepine, but several are highly suggestive of efficacy. One large parallel-group study of carbamazepine compared with placebo has been completed recently by Zhang and associates (Table 34.14–4). These authors found 50 percent reductions in Hamilton Depression Rating Scale scores in 34.8 percent on placebo ($N = 23$), 63.8 percent on carbamazepine ($N = 47$), and 84.8 percent on carbamazepine plus a Chinese herbal preparation (Free and Easy Wanderer Plus [FEWP]) ($N = 46$; $p < 0.05$), further supporting the acute antidepressant effects of the drug. Zhang et al. have also demonstrated acute antidepressant effects of carbamazepine compared with placebo in unipolar depression.

Table 34.14–5.
Possible Correlates of Carbamazepine Response to Treatment in the Affective Disorders

Depression
Greater acuity and severity
Low homovanillic acid levels in CSF
Higher epoxide/carbamazepine ratio
Greater decreases in thyroxine and free thyroxine decrease during carbamazepine treatment
Frontal and paralimbic (especially left insula) hypermetabolism on positron emission tomography
History of alcoholism
Comorbid posttraumatic stress disorder
Prophylaxis
Lack of family history of bipolar illness in first-degree relatives
Lack of prior rapid cycling
Nonclassical presentations:
Bipolar II, non-euphoric mania
Schizoaffective (especially depressive with mood incongruent delusions)
Substance abuse comorbidities

Moreover, the authors have observed that a group of 10 patients at the National Institute of Mental Health (NIMH) who were initial good acute antidepressant responders to carbamazepine on a double-blind basis subsequently relapsed on placebo and then re-responded during a second blind treatment trial with carbamazepine (Fig. 34.14–8). This off-on-off–on documentation of responsivity indicates that carbamazepine has unequivocal antidepressant effects in a subgroup of highly treatment-refractory depressed patients. While good acute antidepressant effects occurred in about one-third of patients exposed to carbamazepine, the percentage of antidepressant response to this agent and the particular characteristics and predictors of response remain to be further delineated.

Initial data summarized in Table 34.14–5 suggest that those with more severe episodic and less chronic depressions were among the best responders to the drug, as were those with low baseline CSF HVA, the

major metabolite of dopamine. Carbamazepine responders also showed a pattern of baseline hypermetabolism on PET scans that is unusual and contrasts with the more typical hypofrontality of major depression. Some data also suggest that those with a history of depression in association with prior alcoholism are among those who respond well to the drug. This is of considerable interest because carbamazepine can block the effects of alcohol withdrawal in experimental animals, and the drug is widely used in the acute and long-term treatment of those with alcohol-related dysphoria in Scandinavian countries.

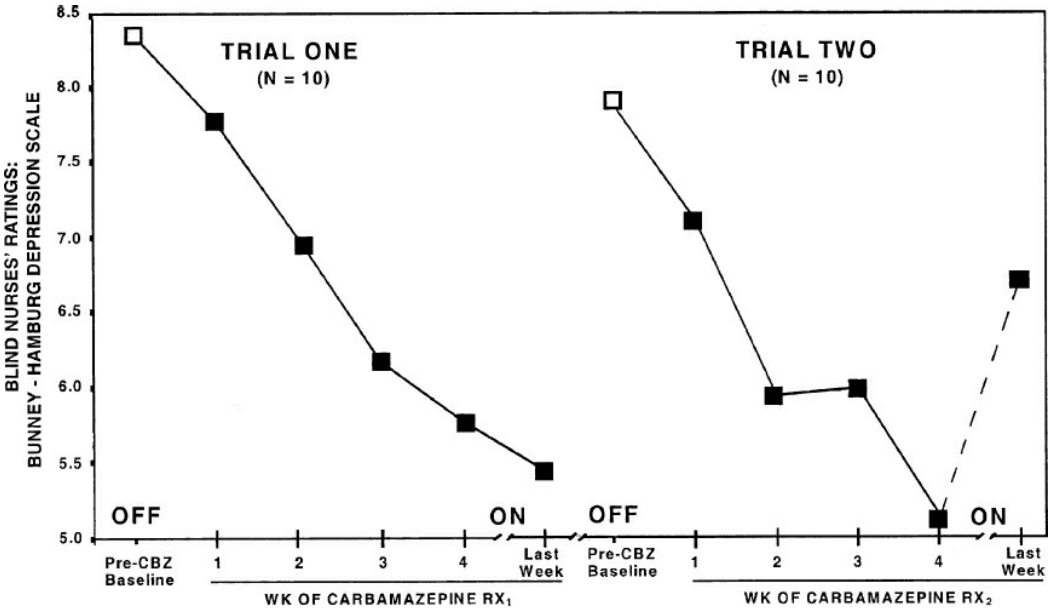

FIGURE 34.14–8. Confirmation of antidepressant response to carbamazepine (CBZ) in an off-on–off-on design in a subgroup of ten treatment-refractory patients. Seventeen of the 54 depressed patients studied in a blind fashion (31.5%) showed an initial good response (≥2 points on the Bunney–Hamberg Depression scale).

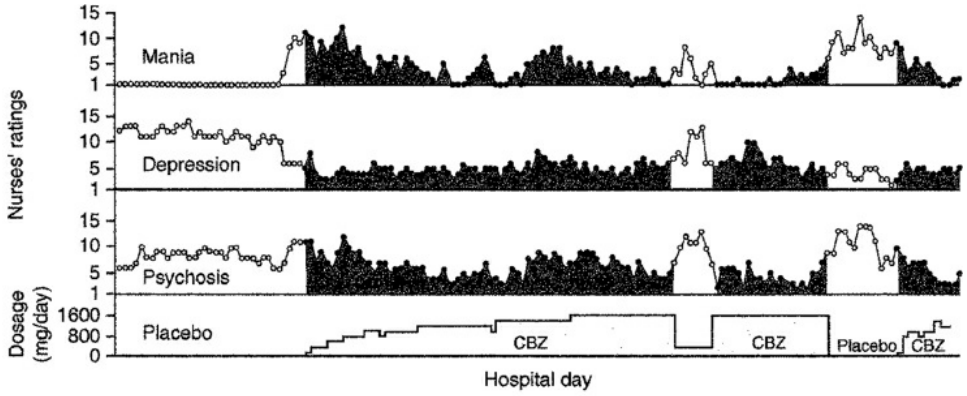
Nurses' ratings
Mania
Depression
Psychosis
Dosage (mg/day)
Placebo
CBZ
CBZ
Placebo
CBZ
Hospital day

A

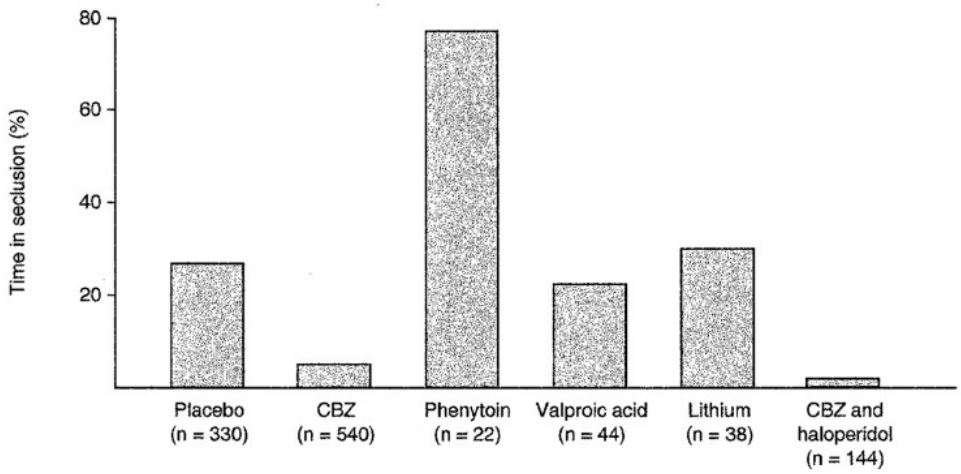
Time in seclusion (%)
Placebo (n = 330)
CBZ (n = 540)
Phenytoin (n = 22)
Valproic acid (n = 44)
Lithium (n = 38)
CBZ and haloperidol (n = 144)

B

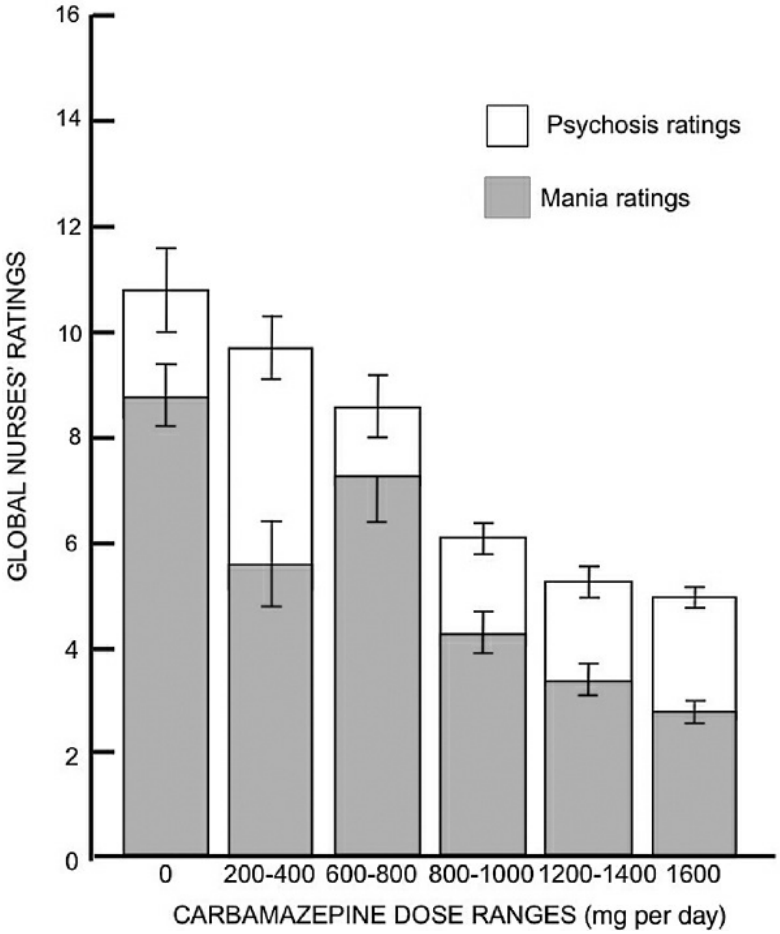
Psychosis ratings
Mania ratings
GLOBAL NURSES' RATINGS
0
200-400
600-800
800-1000
1200-1400
1600
CARBAMAZEPINE DOSE RANGES (mg per day)

C

FIGURE 34.14–9. Efficacy of carbamazepine (CBZ) in preventing manic and depressive episodes in a patient whose symptoms were previously refractory to lithium and antipsychotic treatment **(A).** CBZ administered on a blind basis acutely suppressed manic and psychotic symptoms. During periods of CBZ dose reduction or placebo substitution, exacerbation of manic and psychotic symptoms was consistently noted and reinstitution of active CBZ treatment led to improvement on three different occasions. Each rating circle is one hospital day **(B).** The prophylactic effect of CBZ during double-blind, clinical trials in a treatment-refractory, rapidly cycling patient who required seclusion 26 percent of the time (based on each 24-hour day) while taking placebo or 30 percent of the time while taking lithium, but only 4 percent of the time while taking CBZ (*n* = 540 days) at doses greater than 800 mg per day. The patient was not responsive to two other antiepileptic drugs, phenytoin (seclusion 77 percent) or valproic acid (seclusion 22 percent), but did exceptionally well and rarely required seclusion (1.6 percent) while treated with the combination of CBZ and haloperidol. *n,* number of days in treatment. **C:** Dose-dependent therapeutic effects of carbamazepine during a prolonged manic episode in the same patient.

Prophylaxis

More than a dozen partially controlled studies (involving either randomization or double-blind comparison with other agents such as lithium) reveal that carbamazepine is approximately as effective as lithium in the prevention of affective episodes, although several caveats remain. In these controlled studies of carbamazepine prophylaxis, 224 of 382 (59 percent) showed a moderate to marked response to carbamazepine. Similarly, lithium response in these same studies occurred in 241 of 395 patients (61 percent). However, in several studies, there was better antimanic efficacy of lithium compared with carbamazepine.

The German collaborative study indicated several different predictors of response to these two agents. Lithium was substantially more effective than carbamazepine in those with a classic bipolar I diagnosis, with euphoric mania, no thought disorder or comorbid substance abuse, as well as (in other studies) a positive family history of affective disorder. Conversely, there was a trend toward better prophylactic efficacy of carbamazepine in those with bipolar II illness, schizoaffective diagnosis, dysphoric mania, and substance abuse comorbidity (current or prior), as well as (in other studies) a negative family history of bipolar disorder in first-degree relatives.

Case studies performed in a double-blind fashion have also documented that some patients respond to carbamazepine and they fail to respond to divalproex sodium (Depakote) (Fig. 34.14–9B) or vice versa. These observations are consistent with the notion that response to the anticonvulsants does not occur as a class effect, and that further elucidation of the clinical and biological predictors of response to carbamazepine would be of great clinical importance.

It is noteworthy that within individual patients there appears to be dose-related antimanic and antipsychotic responsiveness (Fig. 34.14–9C). However, there is little evidence of a relationship between blood levels across individuals and the degree of clinical antidepressant or antimanic

response (Fig. 34.14–10).

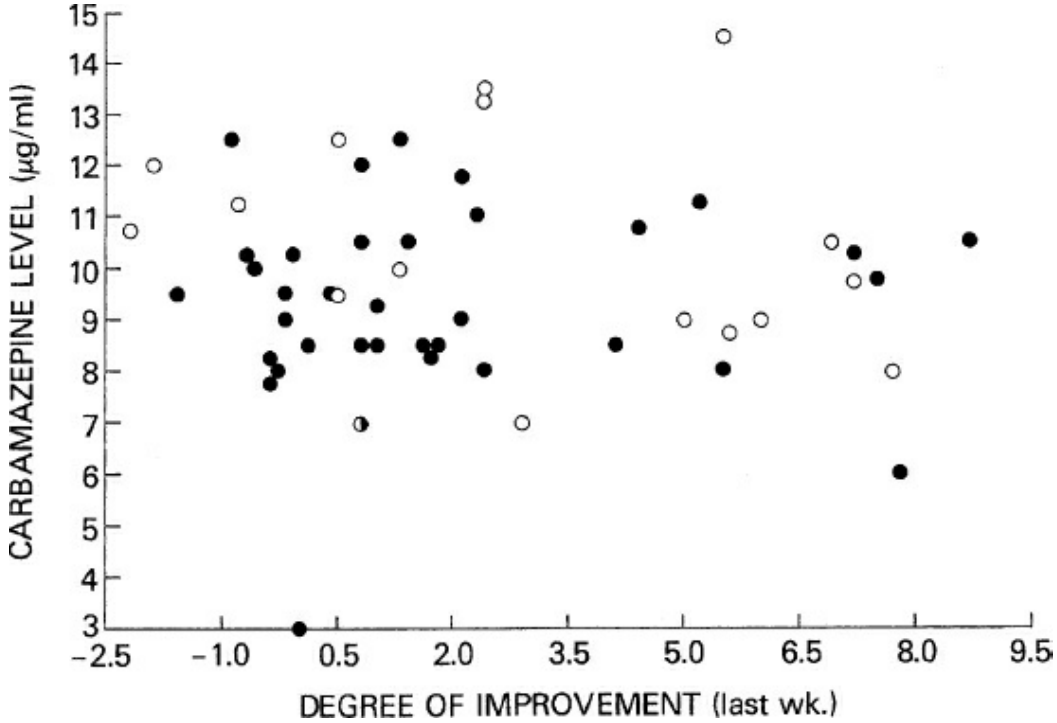

FIGURE 34.14–10. Lack of relationship between plasma levels and degree of clinical improvement across patients in mania (*open circles*) or depression (*filled circles*).

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Central Nervous System

Typical of many anticonvulsant drugs, carbamazepine can exert a variety of untoward effects on CNS function. Most commonly these include dizziness, ataxia, and diplopia (characteristically occurring approximately 2 hours after a given dose of the regular immediate-release formulation). These common side effects can largely be avoided with slow upward titration of the dose to just below a patient's side-effects threshold, which can vary greatly, as illustrated in Figure 34.14–10. Given this wide range of individual proclivity to side effects (Fig. 34.14–11) and the lack of a clear anticonvulsant dose–response relationship, the importance of individualizing dose titration is emphasized as preferable to attempting to achieve a set blood level within the conventional 4 to 12 μg/mL "therapeutic" window.

In overdose, carbamazepine produces drowsiness, stupor, coma, sinus tachycardia, atrial ventricular conduction blocks, hypotension or hypertension, seizures, nystagmus, and hyporeflexia or hyperreflexia. Hypothermia, oral dyskinesias, and respiratory depression can also occur, yet few fatalities have been reported even after ingestion of as high as 10 to 20 g of carbamazepine.

In overdose, supportive techniques include gastric lavage and

instillation of activated charcoal (10 to 100 g and then 12.5 g per hour) until the patient is symptom free. Consideration of charcoal hemoperfusion before the patient becomes hypotensive has also been recommended. Although flumazenil (Ro 15-1788) has been recommended by some, the preclinical pharmacology noted above would not support such a suggestion. Instead, the $GABA_A$ receptor partial inverse agonist Ro 15-4513, which has been shown to reverse alcohol intoxication and is also effective in reducing the sedation and lethality of carbamazepine overdose in rodents, should be considered.

Carbamazepine's effects on memory and psychomotor performance have generally been minor, and improvement in a few indices has even been reported during treatment. However, higher doses or those that exceed an individual's toxic threshold can be associated with impairment. A similar range of side effects can occur with oxcarbazepine, although it is generally better tolerated than carbamazepine with the exception of an increased incidence of hyponatremia (Fig. 34.14–1).

Blood

As summarized in Table 34.14–6, the hematopoietic system is a major target of carbamazepine's benign and serious adverse actions. A statistically significant but usually clinically insignificant suppression of the white blood cell (WBC) count occurs in a substantial proportion of patients and is thought to be based on the drug's ability to inhibit colony-stimulating factor. Lithium enhances colony-stimulating factor and will reverse the benign WBC suppression of carbamazepine.

In contrast, the potentially life-threatening occurrence of agranulocytosis, pancytopenia, or aplastic anemia is a rare phenomenon and thought to occur in approximately 1 in 25,000 to 100,000 patients. Because of the rarity of these idiosyncratic events, many authorities recognize the inadequacy of routine blood count monitoring and suggest advising patients that, in the face of clinical signs of potential hematological dysfunction (such as fever, sore throat, rash, petechiae, bruising, or overt bleeding), they should immediately contact their physician. It may be wise to monitor the complete blood count (CBC) during initial phases of carbamazepine administration so as to assess the degree of benign white count suppression or any evidence of more problematic changes, such as drops in platelet or red cell counts, which deserve closer monitoring and attention.

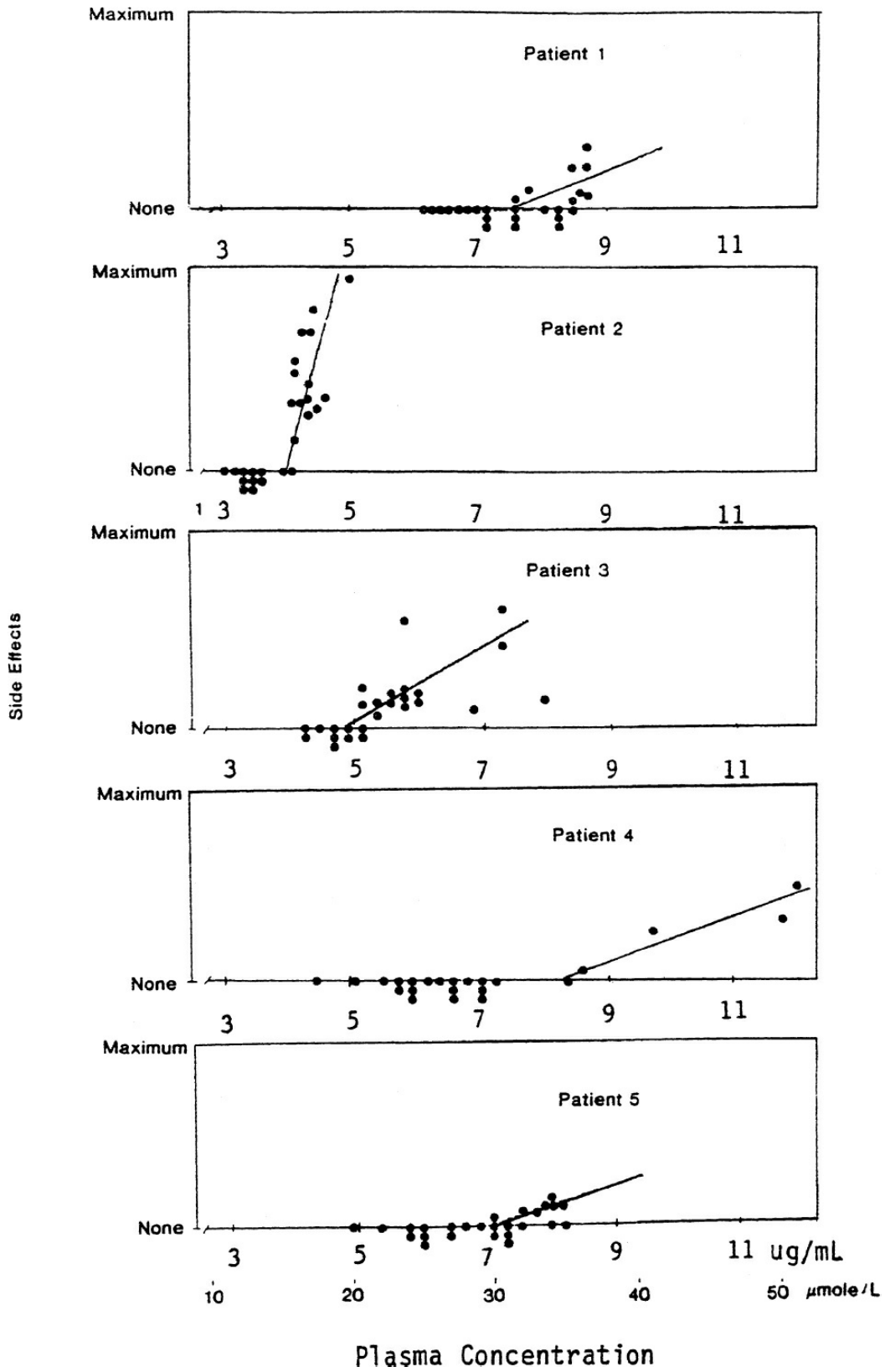

FIGURE 34.14–11. The lack of relationship between carbamazepine concentrations and adverse effects emergence (i.e., highly individualized blood concentration to adverse-effects relationships exist within different patients). (Adapted from Tomson T. Interdosage fluctuations in plasma concentration determine intermittent adverse effects. *Arch Neurol*. 1984;41(8):830–834. Copyright 1984, the American Medical Association.)

Table 34.14–6.
Some Side Effects of Carbamazepine and Their Possible Management

Side Effect (Frequency, Type)	Possible Approach (Strength of Evidence or Practice)	Comment
Dizziness, ataxia, diplopia (C, D)	Reduce peak dose effects by ↓ dose (+++); switch to all HS sustained release formulation (+)	Will occur 2–3 hours after IR dose if patient is near their side effects threshold (+++)
Fatigue, sedation (C, D)	Switch more of IR dose to HS (+++) or all of ER dose to HS (+); ↓ Dose or wait for adaptation or autoinduction (++)	Start with low HS dose; given wide individual variability in side effects threshold, titrate slowly to clinical effect or to side effects limitation (+++)
Benign rash (VC, I)	Discontinue drug (+); antihistamines (+); steroids (+)	Possible switch to oxcarbazepine, 2/3 no rash (++); or can restart CBZ slowly under coverage of prednisone (++)
Severe rash (R, I, S)	Discontinue drug (+++)	Obtain medical/dermatological consultation (+++); rechallenge not recommended (even with steroids)
Benign WBC suppression (VC, D)	Lower dose (+); add lithium (+++)	Lithium increases CSF (+++), whereas CBZ suppresses CSF (+++); 2,500 to 3,000 WBC lower limit of tolerability (±)
vs. Agranulocytosis	Discontinue drug (+++)	Medical support and consultation required (+++)
Aplastic anemia	Discontinue drug (+++)	Medical support and consultation required (+++)
Weight gain (UC, I)	Consider topiramate or zonisamide augmentation (++) or switch to weight-neutral drug such as LTG, OXC (+)	Rarely a problem
Hyponatremia (O, D)	↓ Dose (++); add lithium (++); add demeclocycline (++)	Women, older age, and use of higher doses are vulnerability factors (+++)
Thyroid hormone suppression (VC, D)	None needed as ↓ T_4 may correlate with better response (++)	No ↑ TSH; no change BMR with CBZ (+); thyroid supplements rarely needed (+++); ↓ T_4 and ↓ T_3 may be addictive with lithium (++)
Tremor (R, D)	↓ Dose (+)	Check ammonia levels when combined with VPA if tremor is coarse or flapping) (±)
Teratogenic (O, D) Spina bifida ≈1% Cranial facial abnormality	Pretreat with ↑ folate (+)	Avoid combination Rx if possible, especially with VPA (+); minor developmental delays do not persist
Hepatitis (R, I)	Discontinue if enzymes ↑ to three times normal (++) Discontinue if patient is symptomatic (+++)	
Memory disturbance	↓ Dose (+); switch to all HS with long-acting preparation (±)	CBZ potentiation of vasopressin receptors may prevent memory loss (+/−)
High cholesterol	Treat with statins (+)	Switch to non-enzyme inducer (±)

C, common; CSF, colony stimulating factor; D, dose related; hs, at night; I, idiosyncratic; O, occasional; R,

rare; S, sensitivity may cross to other antiepileptic drugs; T_3, triiodothyronine (liothyronine); T_4, thyroxine; UC, uncommon; VC, very common; WBC, white blood cell; ↑, indicates increase; ↓, indicates decrease; (+) likely works; (++) many case reports; (+++) well supported, controlled data; (±) questionable or hypothetical.

In contrast to carbamazepine, the routine benign suppression of the WBC has not been reported with oxcarbazepine, and the incidence of more serious blood dyscrasias does not appear elevated over baseline. No hematological monitoring is required with this agent.

Hormonal Effects

Carbamazepine increases the 24-hour excretion of urinary free cortisol and notably induces a robust escape from dexamethasone suppression. These effects are thought to be not entirely related to heightened metabolism of steroid compounds (such as lowering of dexamethasone levels) and may suggest actions of carbamazepine that are contrary to those of more traditional antidepressants that downregulate the hypothalamic–pituitary–adrenal axis and increase glucocorticoid hormone receptors upon chronic administration. Whether this side effect of increasing cortisol could be clinically useful in the treatment of patients with PTSD who have abnormally low basal secretion of urinary free cortisol and show enhanced sensitivity to dexamethasone suppression compared with normal or depressed patients remains to be systematically explored.

Carbamazepine decreases circulating levels of thyroxine (T_4), free T_4, and triiodothyronine (T_3) without substantially altering thyroid-stimulating hormone. Studies indicate that it does not affect basal metabolic rate and that thyroid replacement is almost never required. In fact, patients with the largest decrements in T_4 or free T_4 showed the greatest degree of antidepressant response to carbamazepine in one series. For example, the degree of improvement in depression with carbamazepine was inversely related to the degree of decrease in free T_4 from baseline ($r = -0.56$; $N = 36$; $p = .0004$).

Renal Effects (Hyponatremia)

Carbamazepine has direct or indirect agonist-like effects at the vasopressin receptor, accounting for its occasional use in the treatment of diabetes insipidus. It will not reverse the lithium-induced variety of this syndrome, however. This carbamazepine-increased resorption of water predisposes patients to the development of hyponatremia and, in rare instances, water intoxication. Elderly patients, women, and those treated with high doses are more prone to this side effect. Several small case series suggest that carbamazepine-related hyponatremia can be prevented or reversed by the administration of lithium or demeclocycline (Declomycin).

Hyponatremia is the one area in which oxcarbazepine has a higher incidence of side effects than carbamazepine. Approximately 3 to 5 percent of patients treated with oxcarbazepine develop hyponatremia, and whether lithium or demeclocycline would be effective treatments is not known.

The elimination half-life of oxcarbazepine is prolonged in those with renal impairment.

Skin

Carbamazepine can induce a variety of dermatological syndromes—from a common benign macular papular pruritic rash typically emerging in the first 9 to 23 days after drug initiation to more serious conditions. Drug cessation usually results in regression of the rash, but, in rare instances, it can progress to exfoliative dermatitis, erythema multiforme, Stevens–Johnson syndrome (SJS), or toxic epidermal necrolysis (TEN). Given the inability to distinguish which rashes might evolve toward these serious end points, most clinicians recommend stopping the drug in the face of even the benign presentation of a carbamazepine-related rash, which can be expected to occur in some 5 to 15 percent of individuals. However, a substantial group of patients re-exposed to carbamazepine will not have a re-emergence of the rash. The rash rate may be lower with the new extended release preparations.

In those individuals who appear to respond well to carbamazepine but not other agents, one can attempt to suppress the benign dermatological reaction with prednisone (Cordrol) (40 mg per day). However, several case reports have indicated that, in the face of evidence of more serious systemic allergy (such as fever, lymphadenopathy, and pneumonitis), such steroid co-treatment may not be successful.

Oxcarbazepine appears to have a reduced frequency of inducing both the benign and serious forms of rash. Moreover, the majority of patients who have a rash on carbamazepine, it will not recur on oxcarbazepine, but approximately 25 to 30 percent will show such a cross-sensitivity.

Considerable progress has been made in studying the pharmacogenomics of serious rash (SJS/TEN) on carbamazepine. The human lymphocyte antigen (HLA-B*1502 has been associated with this syndrome in patients of Asian descent (including Han Chinese, Thai, Malaysian, Filipino, and Taiwan) where the frequency of HLA-B*1502 is about 10 to 15 percent in the general population. If one has this antigen the odds ratio (OR) is 79.8. A direct interaction of HLA-B*1502 with cytotoxic T lymphocytes has been reported.

In contrast, in those of northern European decent, the risk of SJS/TEN is lower and of the order of 1 to 6 per 10,000 and is associated with HLA-A*3101 (which occurs in 2 to 5 percent of Europeans) with an OR of 25.9. 9 percent of Japanese have the HLA-A*3101 and it is associated with SJS/TEN OR = 10.8. In addition, preliminary evidence links HLA-B*1511 in

severe rash in Japanese. There is some controversy about the clinical utility of doing these genetic screens for potential carbamazepine-treated patients, but consensus is emerging that it would be worthwhile in those of Asian descent. In Thailand, the positive predictive value (PPV) was 1.92, while the negative predictive value (NPV) was 99.96.

There is another hypersensitivity syndrome (also called DRESS for drug reaction with eosinophilia and systemic symptoms) which occurs with a maculopapular rash (77 percent), increased liver enzymes (71 percent), eosinophilia (64.3 percent), fever (61 percent), lymphadenopathy (55 percent), and increased WBC (43 percent). Treatment with *N*-acetylcysteine (NAC) and/or intravenous immunoglobulin (IVIG) has anecdotally been reported effective as has treatment without the use of steroids. In those with the CBZ-induced hypersensitivity syndrome, CBZ-specific CD4(+) and CD8(+) T cells have been shown to persist some years after the resolution of the syndrome.

Liver

Like other anticonvulsants, carbamazepine has been associated with enzyme elevations. However, the outright appearance of hepatitis is extremely rare. As with oxcarbazepine, routine monitoring does not appear indicated.

Carbamazepine increases total serum cholesterol by an average of 20 points. It increases both high- and low-density lipoproteins, such that deleterious cardiovascular effects have not been unambiguously demonstrated. However, in those at high risk, preventive treatment with statins may be indicated. Homocysteine increases and folate decreases have been reported, so treatment with folate may also be helpful. Vitamin D levels may also decrease on carbamazepine with or without associated increases in osteoporosis.

Weight

In epilepsy patients, long-term treatment with carbamazepine is associated with small degrees of weight gain, which are significantly less than those observed on valproate. Oxcarbazepine is reported to be weight neutral. Carbamazepine decreases plasma levels of ghrelin, which also decreases with obesity.

Heart

The cardiac effects of carbamazepine are generally not prominent, although the drug does decrease atrial-ventricular conduction and is relatively contraindicated in patients with heart block. Carbamazepine has been used to treat digitalis-induced arrhythmias, including ventricular tachycardia. Bradycardia and syncopal episodes have been reported in rare instances,

with elderly women apparently at increased risk. Effects on blood pressure are typically minimal, and orthostatic hypotension is not a common accompaniment (as it is with several of the structurally related tricyclic antidepressants).

THERAPEUTIC INDICATIONS

Bipolar Disorder

Carbamazepine has been approved by the FDA for use in seizure disorders and trigeminal neuralgia. Like many treatments of bipolar disorder, carbamazepine is widely used in the treatment of acute mania and in long-term prophylaxis, but is only FDA approved as the long-acting preparation Equetro in acute mania. Carbamazepine is approved in many European countries, Japan, and Canada, however. It is considered a first- or second-line treatment in many guidelines and can be used as a substitute for lithium or as an adjunct to it for the partial responder (Table 34.14–4). Use in combination with lithium appears especially helpful in rapid cyclers.

Oxcarbazepine is thought to have effects similar to carbamazepine, but these have not yet been documented in large placebo studies and the drug is not FDA approved for bipolar illness. One study reported that that overall oxcarbazepine was no more effective than placebo, although it did exceed placebo in younger children with childhood-onset mania and not in adolescent-onset acute mania.

Carbamazepine can also be used in conjunction with unimodal antidepressants and antipsychotic agents, although several instances of drug–drug interactions are of some importance—for example, blood levels of nefazodone (Serzone), haloperidol, aripiprazole (Abilify), lurasidone (Latuda), and quetiapine (Seroquel) and lamotrigine are substantially reduced. Carbamazepine is not recommended for use in conjunction with the atypical antipsychotic clozapine (Clozaril) for fear of an additive risk of agranulocytosis, although the mechanisms for WBC suppression appear to differ. Use of carbamazepine in conjunction with valproate has not been as systematically explored in the refractory mood disorders as it has in seizure disorders. The combination usually requires a decrease in the dose of carbamazepine (because valproate displaces carbamazepine from its protein-binding sites and also increases the levels of the 10,11-epoxide by virtue of its inhibition of the epoxide hydroxylase). Case reports show efficacy of the combination of carbamazepine and valproate when either drug alone was not efficacious.

As antimanic effects are often observed within the first several days of treatment with carbamazepine, a several weeks' trial in acute mania is usually sufficient to ascertain whether the patient is going to be a responder. Longer periods of evaluation may be required to assess its potential antidepressant efficacy, however. Exploration of its potential

efficacy in long-term prophylaxis requires the observation of a consistent reduction of episode severity or the prevention of episodes for a duration several times longer than the previous usual "well interval." Thus, the faster the cycling at baseline, the more rapid can be the assessment of prophylactic effectiveness.

Although 4 to 12 μg/mL has been identified as a therapeutic anticonvulsant level, serum levels for antimanic efficacy have not been established, and there is no relationship of blood levels to the degree of acute antimanic or antidepressant response. One study suggested that 1,200 mg per day was associated with an antimanic response, although trough serum levels of 8 to 12 μg/mL are generally associated with good anticonvulsant and antimanic effects (Fig. 34.14–10).

Schizophrenia and Schizoaffective Disorder

A series of double-blind studies with few exceptions have indicated that the addition of carbamazepine to antipsychotics is more effective than placebo, particularly in those patients with more florid positive symptoms and problems with aggression and dyscontrol. It does not appear useful as an adjunct in chronic, negative, symptom-prominent schizophrenia. Carbamazepine is helpful in schizoaffective disorder, particularly in depressed phases. One study suggested that, in contrast to carbamazepine, oxcarbazepine was less effective in those with more fulminant psychotic manic symptoms.

Impulse Control Disorders

A very substantial case series literature indicates that carbamazepine is highly effective in a wide range of aggressive syndromes and impulse-dyscontrol disorders that can be manifest in those with diagnoses of borderline and other personality disorders, affective disorders, and schizophrenia. Its efficacy in these dyscontrol syndromes appears independent of the presence or absence of an associated EEG abnormality.

Other Disorders

Carbamazepine is as effective in acute alcohol withdrawal syndromes as diazepam, and it does not have the abuse liability of the benzodiazepines. Greater anxiolytic and global distress decreases have also been reported in withdrawing alcoholics treated with carbamazepine compared with benzodiazepines. In addition, placebo-controlled prophylactic studies of alcohol dependence have suggested greater relapse prevention of heavy drinking with carbamazepine. Although carbamazepine inhibits the development of cocaine-kindled seizures and their associated lethality in rodents, controlled clinical studies in cocaine abusers did not indicate that it is more successful in maintaining abstinence or drug-free urine than

placebo, except in a subgroup with concomitant mood disorder.

Open studies suggest the potential use of carbamazepine in the treatment of PTSD, particularly the paroxysmal recurrent components and the associated sleep disorder. Isolated case reports, but no controlled clinical trials, indicate its adjunctive use in other anxiety disorders and in OCD.

PRECAUTIONS AND ADVERSE REACTIONS

The *Physicians' Desk Reference* indicates that aplastic anemia and agranulocytosis occur in conjunction with the use of carbamazepine at a rate five to eight times greater than that in the general population. That incidence is approximately six patients per 1,000,000 of the population per year for agranulocytosis and two patients per 1,000,000 population per year for aplastic anemia. If, in the course of treatment, a patient exhibits low or decreased WBC or platelet counts, that individual should be monitored closely in the face of stable benign WBC suppression; the usual minimum safe total WBC count is approximately 3,000 and absolute neutrophil count approximately 1,000 to 1,500 neutrophils/mm^3. If there is a concomitant suppression of the red cell or platelet series or other indications of bone marrow depression, the drug should be stopped immediately.

Carbamazepine will increase the metabolism of estrogen in birth control formulations, resulting in occasional unwanted pregnancies. Therefore, a higher dosage form of the estrogen component or other contraceptive devices should be used.

Carbamazepine is a category D drug and should be avoided during pregnancy. There is a several percentage risk of spina bifida with the use of carbamazepine, and this appears to increase with anticonvulsant combination treatment. In one series, carbamazepine was reported to decrease head size and body length, although the developmental lag was completely compensated for after a period of several months. Although carbamazepine and its epoxide metabolite are transferred to breast milk during lactation, their concentrations are approximately 50 percent of those in maternal plasma. There are no adequate studies of the clinical effects of oxcarbazepine in pregnant women; given its similar structure to carbamazepine and evidence of problems in animal studies, so one should also consider it teratogenic in humans until more systematic data to the contrary are acquired.

DRUG INTERACTIONS

In contrast to oxcarbazepine, carbamazepine is a potent inducer of cytochrome P450 oxidase enzymes, particularly of the 3A4 category. This leads to its own autoinduction after 2 to 3 weeks, as well as a reduction in the levels of a variety of other compounds listed in Table 34.14–2.

If a patient is switched from carbamazepine to oxcarbazepine, deinduction of the 3A4 enzyme should be expected, and substances that would have their levels reduced by carbamazepine may increase. This is particularly important for lamotrigine, as carbamazepine approximately halves lamotrigine levels and substantial increases in lamotrigine levels would occur with carbamazepine discontinuation. This could also occur with dicumarol and warfarin (Coumadin), increasing proclivity for excess anticoagulant activity and increased bleeding.

Of potential importance for the generation of carbamazepine's toxicity are a series of agents that inhibit either cytochrome oxidase 3A4 or the epoxide hydroxylase, as schematized in Figure 34.14–12. Many commonly used agents markedly increase blood levels of carbamazepine (Table 34.14–7). These include the antibiotic erythromycin and all of its related congeners, as well as the calcium channel blockers verapamil (Verelan) and diltiazem (Cardizem), but not nifedipine (Adalat), nimodipine, or amlodipine (Norvasc). The antituberculosis drug isoniazid (Nydrazid) will also markedly increase carbamazepine levels, but more minor effects are observed with dextropropoxyphene (Darvon), fluoxetine (Prozac), and nicotinamide (Nicomide), and minimal effects occur with cimetidine (Tagamet).

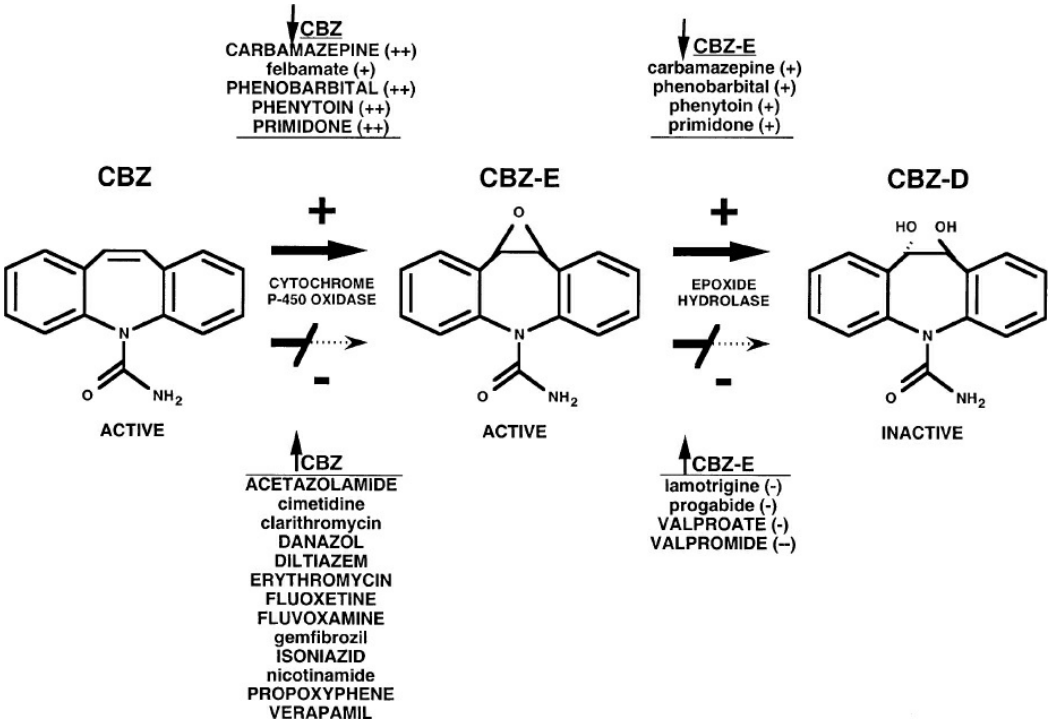

FIGURE 34.14–12. Carbamazepine (CBZ) metabolism. *Bottom left:* Drugs that inhibit cytochrome P450 increase CBZ. ++, drugs that decrease the level of the active metabolite carbamazepine through induction of its oxidative metabolic pathway cytochrome P450; +, drugs that decrease the level of the active metabolite carbamazepine-10, 11-epoxide (CBZ-E) through induction of the epoxide hydroxylase and production of the inactive dihydroxy metabolite (CBZ-D); –, drugs that increase the level of carbamazepine-10,11-epoxide through blockade of the epoxide catabolism. Drug names in capital letters indicate common

or more important interactions.

There are fewer drug–drug interactions of importance for oxcarbazepine. However, strong inducers of cytochrome P450 enzymes (such as carbamazepine) decrease plasma levels of the monohydroxy derivative metabolite by 29 to 40 percent. As noted in Table 34.14–1, there is no autoinduction, and few major drug–drug interactions have been reported. However, oxcarbazepine reduces estrogen levels in oral contraceptives, and its use, as with carbamazepine, will necessitate higher dosage forms or alternative contraceptive devices.

Use of carbamazepine fell off in the 1990s after a decade of considerable use in the 1980s. One reason was the advent of valproate and the notion that it was easier to use. However, now that it is clear that some patients are relatively selective responders to carbamazepine, it becomes increasingly important to demystify some of the important details about its use. A primer employing a dozen key principles is included in Table 34.14–8.

Table 34.14–7.
Drugs That Increase Concentrations of Carbamazepine

Drug Class	Drug
Antidepressants	Fluoxetine, flavoxamine, nefazodone[a,b]
Antiepileptics	Valproate (increases CBZ-10, 11-epoxide[c])
Antimicrobials	Isoniazid[d], quinupristin/dalfopristin
Ca^+ channel blockers	Diltiazem[c], verapamil[c,d], (but not nifedipine or nimodipine)
Lipid active	Gemfibrozil[a], nicotinamide[a]
Macrolide antibiotics	Clarithromycin[c,d], erythromycin[c,d], flurimycin[c,d], josamycin[c,d], midecamycin acetate, ponsinomycin[c,d], troleandromycin, triacetyleandronyion[c,d]
Anti-HIV	Delavirdine[c,d], ritonavir[c,d]
Others	Acetazolamide[a], cimetidine[a], danazol, delavirdine[c,d], dextropropoxyphene[a], gestodene[b], mifepristone, omeprazole[a], ticlopidine[a]

[a]Weak effect.

[b]Mild to moderate.

[c]Potential for toxicity.

[d]Important or severe.

Table 34.14–8.
CBZ in Bipolar Illness: A Brief Users Guide

Start with low (200 mg) H.S. doses in depression, euthymia; higher doses (600–800 mg/day in nighttime divided doses) in manic inpatients.

All H.S. dosing is reasonable with CBZ ER (Equetro) preparation.

Titrate slowly to an individual's response or side effects threshold (as there is no correlation with blood levels and degree of response).

Hepatic enzyme CYP 450 (3-A-4) induction and autoinduction occurs in 2–3 weeks; slightly higher doses may be needed/tolerated at that time.

Employ adequate birth control methods, including higher dosage forms of estrogen (as CBZ will lower estrogen levels).

Major increases in CBZ blood levels and potential toxicity are caused by erythromycin, verapamil, isradipine.

Warn regarding benign rash 5–10%; rare progression to severe rash, so drug should be discontinued. HLA-A*3101 predicts severe rash in northern Europeans and Japanese. HLA-B-1502 predicts in those of Asian descent.

Benign WBC decreases occurs regularly (and are usually inconsequential). Lithium will prevent or reverse these.

Rare agranulocytosis and aplastic anemia (several per million new exposures); warn regarding appearance of fever, sore throat, petechiae, bleeding gums, and to check with physician to obtain immediate CBC.

Avoid CBZ in pregnancy: Spina bifida 0.5%; other adverse outcomes about 8%.

Some people will response well to CBZ and not other mood stabilizers (lithium) or anticonvulsants (VPA).

Combination treatment often required to maintain remission and prevent loss of effect via tolerance.

H.S., nighttime or hours of sleep.

References

Alvarez G, Marsh W, Camacho IA, Gracia SL. Effectiveness and tolerability of carbamazepine vs. oxcarbazepine as mood stabilizers. *Clin Res Reg Affairs*. 2003;20:365.

Benedetti A, Lattanzi L, Pini S, Musetti L, Dell'Osso L, Cassano GB. Oxcarbazepine as add-on treatment in patients with bipolar manic, mixed, or depressive episode. *J Affect Disord*. 2004;79:273–277.

*Bertilsson L, Tomson T. Clinical pharmacokinetics and pharmacological effects of carbamazepine and carbamazepine-10,11-epoxide: an update. *Clin Pharmacokinet*. 1986;11:177–198.

*Denicoff KD, Smith-Jackson EE, Disney ER, Ali SO, Leverich GS, Post RM. Comparative prophylactic efficacy of lithium, carbamazepine, and the combination in bipolar disorder. *J Clin Psychiatry*. 1997;58(11):470–478.

Emrich HM. Studies with (Trileptal) oxcarbazepine in acute mania. *Int Clin Psychopharmacol*. 1990;5[Suppl 1]:83.

Fattore C, Cipolla G, Gatti G, et al. Induction of ethinylestradiol and levonorgestrel metabolism by oxcarbazepine in healthy women. *Epilepsia*. 1999;40(6):783–787.

Ghaemi NS, Berv DA, Klugman J, Rosenquist KJ, Hsu DJ. Oxcarbazepine treatment of bipolar disorder. *J Clin Psychiatry*. 2003;64:943–945.

Glauser TA, Nigro M, Sachdeo R, et al. Adjunctive therapy with oxcarbazepine in children with partial seizures. The Oxcarbazepine Pediatric Study Group. *Neurology*. 2000;54(12):2237–2244.

*Greil W, Kleindienst N, Erazo N, Muller-Oerlinghausen B. Differential response to lithium and carbamazepine in the prophylaxis of bipolar disorder. *J Clin Psychopharmacol*. 1998;18(6):455–460.

Hartong EG, Moleman P, Hoogduin CA, Broekman TG, Nolen WA, LitCar Group. Prophylactic efficacy of lithium versus carbamazepine in treatment-naive bipolar patients. *J Clin*

Psychiatry. 2003;64:144–151.

Hebeisen S, Pires N, Loureiro AI, et al. Eslicarbazepine and the enhancement of slow inactivation of voltage-gated sodium channels: a comparison with carbamazepine, oxcarbazepine and lacosamide. *Neuropharmacology*. 2015;89:122–135.

Isojarvi JI, Turkka J, Pakarinen AJ, Kotila M, Rattya J, Myllyla VV. Thyroid function in men taking carbamazepine, oxcarbazepine, or valproate for epilepsy. *Epilepsia*. 2001;42(7):930–934.

Joca SR, Skalisz LL, Beijamini V, Vital MA, Andreatini R. The antidepressive-like effect of oxcarbazepine: Possible role of dopaminergic neurotransmission. *Eur Neuropsychopharmacol*. 2000;10(4):223–238.

Ketter TA, Kimbrell TA, George MS, Willis MW, Benson BE. Baseline cerebral hypermetabolism associated with carbamazepine response, and hypometabolism with nimodipine response in mood disorders. *Biol Psychiatry*. 1999;46(15):1364–1374.

Pazzaglia PJ, Post RM, Ketter TA, et al. Nimodipine monotherapy and carbamazepine augmentation in patients with refractory recurrent affective illness. *J Clin Psychopharmacol*. 1998;18(5):404–413.

Post RM, Altshuler L, Leverich GS, et al. Clinical correlates of sustained response to individual drugs used in naturalistic treatment of patients with bipolar disorder. *Compr Psychiatry*. 2016;66:146–156.

Post RM, Berrettini W, Uhde TW, Kellner C. Selective response to the anticonvulsant carbamazepine in manic-depressive illness: a case study. *J Clin Psychopharmacol*. 1984;4(4):178–185.

Post RM, Leverich GS. *Treatment of Bipolar Illness: A Casebook for Clinicians and Patients,* Chapters 14–20. NY: W.W.Norton & Company, Inc; 2008:133–190.

Post RM, Uhde TW, Roy-Byrne PP, Joffe RT. Antidepressant effects of carbamazepine. *Am J Psychiatry*. 1986;143:29–34.

Raja M, Azzoni A. Oxcarbazepine vs. valproate in the treatment of mood and schizoaffective disorders. *Int J Neuropsychopharmacol*. 2003;6:409–414.

Rattya J, Turkka J, Pakarinen AJ, et al. Reproductive effects of valproate, carbamazepine, and oxcarbazepine in men with epilepsy. *Neurology*. 2001;56(1):31–36.

Rosa FW. Spina bifida in infants of women treated with carbamazepine during pregnancy. *N Engl J Med*. 1991;324:674–677.

Shimoyama R, Ohkubo T, Sugawara K. Monitoring of carbamazepine and carbamazepine 10,11-epoxide in breast milk and plasma by high-performance liquid chromatography. *Ann Clin Biochem*. 2000;37(Pt 2):210–215.

Wagner KD, Kowatch RA, Emslie GJ, et al. A double-blind, randomized, placebo-controlled trial of oxcarbazepine in the treatment of bipolar disorder in children and adolescents. *Am J Psychiatry*. 2006;163(7):1179–1186.

Wang PW, Ketter TA, Post RM. Carbamazepine and oxcarbazepine. In: Schatzberg AF, Nemeroff CB, eds: *Textbook of Psychopharmacology*. Washington, DC: American Psychiatric Press; 2004:581.

*Weisler RH, Kalai AK, Ketter TA, SPD417 Study Group. A multicenter, randomized, placebo-controlled trial of extended-release carbamazepine capsules as monotherapy for bipolar

disorder patients with manic or mixed episodes. *J Clin Psychiatry*. 2004;65(4):478–484.

Weiss SR, Clark M, Rosen JB, Smith MA, Post RM. Contingent tolerance to the anticonvulsant effects of carbamazepine: relationship to loss of endogenous adaptive mechanisms. *Brain Res Brain Res Rev*. 1995;20:305–325.

Zhang ZJ, Kang WH, Tan QR, et al. Adjunctive herbal medicine with carbamazepine for bipolar disorders: a double-blind, randomized, placebo-controlled study. *J Psychiatr Res*. 2007;41(3–4):360–369.

Zhang ZJ, Postma T, Obeng K, Russell S, Weiss SR, Post RM. The benzodiazepine partial inverse agonist Ro-15-4513 alters anticonvulsant and lethal effects of carbamazepine in amygdala-kindled rats. *Neurosci Lett*. 2002;329:253–256.

Zhang ZJ, Xing GQ, Russell S, Obeng K, Post RM. Unidirectional cross-tolerance from levetiracetam to carbamazepine in amygdala-kindled seizures. *Epilepsia*. 2003;44:1487–1493.

▲ 34.15 Cholinesterase Inhibitors and Memantine

ELIZABETH A. CROCCO, M.D., AND SAMIR A. SABBAG, M.D.

In 2010, the total number of people with dementia worldwide was 35.6 million. With the aging of the population, particularly in the United States, these numbers are expected to nearly double in the next 20 years. The incidence of Alzheimer disease (AD) increases with age. Epidemiological studies in the United States report its prevalence to be about 3 percent at age 65, which increases to greater than 30 percent by age 85. As the population ages, the prevalence of the disease will increase and significantly impact society and its health care systems.

Currently, cholinesterase inhibitors and memantine (Namenda) are the only medications approved by the U.S. Food and Drug Administration (FDA) for the treatment of AD and other dementias (Table 34.15–1). Since their introduction in the early 1990s, cholinesterase inhibitors were indicated only for mild to moderately ill patients with AD. Recently, their indication has expanded to include patients with moderate to severe AD and patients with dementia associated with Parkinson disease (PD). These agents are also being studied as potential treatment for other types of non-AD dementias, but they presently lack an FDA indication. Memantine, an *N*-methyl-D-aspartate (NMDA) receptor antagonist, has been approved for moderate to severe AD. It is routinely used with a cholinesterase inhibitor in clinical practice, as studies have demonstrated how this combination provides a potentially better response compared to monotherapy with a cholinesterase inhibitor.

Both the histopathology and clinical features of AD was first described over 100 years ago by Alois Alzheimer. Even today, its underlying pathophysiology and genetic basis continue to be investigated. Amyloid-β peptide and phosphorylated tau proteins make up the neuritic plaques and

neurofibrillary tangles seen in AD. These proteins play a central role in the two leading theories behind the pathogenesis of the disease. Additionally, loss of cholinergic neurons and impairment in cholinergic function have also been demonstrated in AD. Early subcortical changes occur in the nucleus basalis of Meynert, which provides cholinergic innervation to the hippocampus and cortex. Decline of acetylcholine in this area has been linked to the cognitive impairment seen in AD, and has been the basis for cholinergic therapies.

Currently, the most effective approach to increasing acetylcholine activity is achieved with the use of cholinesterase inhibitors. Cholinesterase inhibitors deter the activity of the enzymes that break down acetylcholine at the synapse. Two enzymes are known to hydrolyze acetylcholine: acetylcholinesterase (AChE) and butyrylcholinesterase (BuChE). AChE has two binding sites—anionic and esteratic—and is present in several areas of the brain, as well as in erythrocytes, platelets, lymphocytes, and skeletal muscle. BuChE has a similar structure to that of AChE but lacks the peripheral anionic binding site. It is synthesized in the liver, secreted into the plasma, and found in the gastrointestinal tract and heart, comprising approximately 10 percent of the total amount of cholinesterase in the central nervous system (CNS) present in glial cells. In vitro studies indicate that cholinesterase inhibitors can selectively inhibit AChE and BuChE to varying extents. Currently, three cholinesterase inhibitors are available in the United States: donepezil (Aricept), rivastigmine (Exelon), and galantamine (Razadyne). The first cholinesterase to be approved for treatment of dementia, tacrine (Cognex), was recently removed from the US market due to its high side-effect profile.

Table 34.15–1.
Medications Used in Neurocognitive Disorders

Drug (Brand)	Formulation	Indications	Starting Dose	Titration	Dosage Range
Tacrine (Cognex)	10-, 20-, 30-, 40-mg capsules	Mild to moderate AD dementia	10 mg 4 times a day	Increase by 10 mg 4 times a day (40 mg total daily) every 4 weeks	40–160 mg daily with 4 times a day dosing
Donepezil (Aricept)	5-, 10-, 23-mg tablets 5-, 10-mg disintegrated tablets	Mild to moderate and severe AD dementia	5 mg daily	Increase to 10 mg after 4 weeks. Increase to 23 mg daily after 3 months	10–23 mg daily
Rivastigmine (Exelon, Exelon	1.5-, 3-, 4.5-, 6-mg oral solution	Mild to moderate AD dementia Major	4 mg twice a day 8 mg daily	By 1.5 mg twice a day every 4	3–6 mg daily for oral preparation.

Patch)	4.6-, 9.5-, 13.3-mg transdermal patch	neurocognitive disorder due to Parkinson disease	extended release	weeks. Increase to 9.5 mg in 4 weeks and then to 13.3 mg after 4 weeks for transdermal patch	4.6 – 13.3 mg daily for transdermal patch
Galantamine (Razadyne, Razadine ER)	4-, 8-, 12-mg tablets 8-, 16-, 24-mg extended-release capsules 4-mg/mL oral solution	Mild to moderate AD dementia	1.5 mg twice a day 4.6-mg transdermal patch	Increase by 4 mg twice a day every 4 weeks. Increase by 8 mg every 4 weeks for extended release	4–12 mg in divided doses for oral preparation. 8–24 mg daily for extended release
Memantine (Namenda, Namenda XR)	5-, 10-mg tablets 7-, 14-, 21-, 28-mg extended-release capsules 10-mg/5-mL oral solution	Moderate to severe AD dementia	5 mg daily 7-mg extended-release formula	By 5 mg daily in weekly intervals. Increase by 7 mg daily extended-release formula	10–20 mg daily in divided dose for oral preparation. 7–28 mg daily for extended release
Memantine XR and donepezil combination (Namzaric)	14-mg memantine ER/10-mg donepezil capsules 28-mg memantine ER/10-mg donepezil capsules	Moderate to severe Alzheimer disease when patient already stabilized on memantine ER and donepezil	14-mg memantine ER/10-mg donepezil daily 28-mg memantine ER/10-mg donepezil daily in severe renal impairment		

Tacrine inhibits BuChE 18 times more potently than AChE, while donepezil inhibits AChE with higher selectivity. Rivastigmine inhibits both AChE and BuChE, and galantamine selectively inhibits AChE, while also producing allosteric modulation of nicotinic receptors. Although the actions of AChE have been thoroughly studied, the relevance of BuChE and nicotinic receptor modulation has become increasingly apparent to the overall function of cholinergic neurons. BuChE has been shown to be important in the development of amyloid plaques. When a polymorphic mutation of BuChE, known as the K variant, is combined with the apolipoprotein E4 (APOE4), this may result in an increased risk factor for developing AD. Nicotinic modulation results in increased ion channel

opening, enhancing cholinergic transmission independent of enzymatic inhibition. Nicotinic receptors are found in presynaptic and postsynaptic neurons. Presynaptic nicotinic receptors regulate the release of acetylcholine, glutamate, and γ-aminobutyric acid (GABA), while postsynaptic nicotinic receptors mediate cholinergic transmission in the hippocampus and cerebral cortex. Both receptor types play an important role in memory and learning.

Cholinesterase inhibitors can be classified into three categories: reversible, pseudoirreversible, and irreversible. These categories are based on the duration of enzyme inhibition. Reversible inhibitors bind to the enzyme for a very short time (seconds to minutes). Pseudoirreversible inhibitors bind for hours, but new enzymes are not developed to assist in the body's biochemical functions. Irreversible inhibitors bind to the enzyme and prevent further activity, prompting the body to synthesize new enzymes.

Prior to the appearance of medications for cognitive impairment, the estimated total direct and indirect costs for the treatment and care of Alzheimer patients in the United States was nearly $100 billion in 1990. Their introduction was associated with a reduction in costs of medical and social services, as well as a slight increase in functioning and a delay in nursing home placement. These medications helped bring to a lesser degree acute hospital visits, outpatient clinic encounters and unpaid caregiver time.

In addition to AD, there are other neurodegenerative disorders that can commonly lead to a major neurocognitive disorder. Major neurocognitive disorders due to PD, Lewy Body Disease (LBD), vascular dementia (VD), as well as frontotemporal dementia (FTD) have distinct pathophysiological mechanisms, although some overlap may occur. Dementia can occur in up to 40 percent of idiopathic PD patients who demonstrate decline in cholinergic function. The cholinesterase inhibitor, rivastigmine, is indicated by the FDA for the treatment of cognitive impairment in PD, and there is some evidence that other cholinesterase inhibitors like donepezil may also improve major neurocognitive disorder due to PD. Diffuse LBD, although a distinct illness clinically, has similar pathophysiological mechanisms to PD and may possibly benefit from cholinesterase inhibitors. VD may produce cerebrovascular damage causing cholinergic dysfunction. Additionally, the added presence of AD pathology rather than pure VD may explain why some studies demonstrate higher rates than others in improvement in VD with the use of cholinesterase inhibitors. Currently, there are no adequate randomized controlled trials (RCTs) showing efficacy of cholinesterase inhibitors in the treatment of FTD and similar dementias with tau neuropathology. Cholinesterase inhibitors are commonly used off-label in clinical practice in select non-AD dementia conditions, but tolerability and adverse reactions must be carefully taken into account.

Patients with mild neurocognitive disorders, including mild cognitive impairment (MCI), are a group of patients who display a modest decline in cognition. They do not meet criteria for dementia, as this decline does not preclude them from independent performance of everyday activities. These individuals pose a challenging dilemma, as cholinesterase inhibitors are not approved for the treatment of their illness, even though the annual conversion rate into a major neurocognitive disorder is approximately 10 to 15 percent. Trials with these medications have not been found to decrease the rate of progression to dementia, and at this time, there is weak support for using them in these conditions.

Very few studies have looked into the efficacy of cholinesterase inhibitors in treating behavioral and psychiatric symptoms, such as agitation, visual hallucinations or apathy in patients with AD. The evidence is limited due to methodological problems with these studies. Currently, despite not having FDA approval for treatment of behavioral or psychiatric problems, their use could be an alternative in the absence of other safe and effective options.

Cholinesterase inhibitors and memantine are moderately effective in treating the symptoms of AD including cognitive, functional, and behavioral impairment. Unfortunately, these agents are not disease-modifying and risks versus benefit must be carefully weighed by the clinician before starting treatment. Clinicians, caregivers, and patients should be made aware of and educated on realistic goals of therapy. The overall results of the pharmacological intervention should be to maintain for a longer period of time the patient's quality of life and functional ability. By detecting, diagnosing, and treating AD early during its development, there is an opportunity for clinicians and caregivers to optimize drug therapy and other psychosocial interventions that are important in the disease process. Cholinesterase inhibitors are indicated for all stages of the disease and early therapeutic interventions can maximize pharmacological efficacy with these agents.

TACRINE

The first cholinesterase inhibitor approved by the FDA for the treatment of AD, tacrine was introduced to the market in 1993. Tacrine is a reversible cholinesterase inhibitor that works by making intrasynaptic acetylcholine readily available in the brain of patients with AD. Despite being the first agent developed for AD, the use of tacrine was inconvenient due to its multiple daily dosing of three to four times a day and the likelihood of hepatoxicity, which required constant monitoring of liver function during its use. In 2013, tacrine was removed from the US market due to concerns with safety and is no longer available.

Chemistry

The molecular structure of tacrine is shown in Figure 34.15–1. It belongs to the family of the aminoacridines and consists of a 1,2,3,4-tetrahydro-9-acridinamine monohydrochloride monohydrate. This component is hydrolyzed to a variety of metabolites, including velnacrine maleate. This metabolite was under investigation as a possible treatment for AD, but studies were discontinued due to side effects, including elevation of transaminases. Other hydroxylated metabolites—2-OH and 4-OH—in tacrine possess weak cholinesterase inhibition activity.

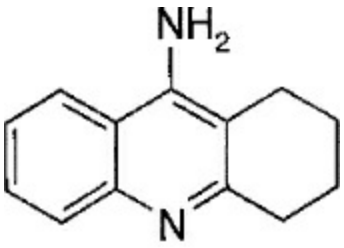

Tacrine

FIGURE 34.15–1. Molecular structure of tacrine.

Pharmacological Actions

Pharmacokinetics. Tacrine possesses a low bioavailability (Mean [M] = 17 percent) despite being well absorbed from the gastrointestinal tract. Its time to peak plasma concentration ranges from 1 to 2 hours. Absorption rate and level of tacrine is decreased up to 40 percent by food intake, however food does not affect absorption if the medication is taken at least 1 hour before a meal. The elimination half-life of tacrine is 1.3 to 7 hours and reaches steady-state concentrations in 24 to 36 hours. Protein binding is at about 75 percent with a steady-state volume of distribution of 300 L. Tacrine is primarily metabolized by hepatic cytochrome P450 (CYP) 1A2 and to a minor extent, by CYP 2D6 isozymes. Its disposition is not affected by renal impairment and dose adjustments are not needed. There is a nonlinear relationship between tacrine dose and steady-state concentration. In women, plasma concentrations of the drug are almost twice as in men when comparable doses are given. This difference could be due to a lower CYP 1A2 activity in women. Smoking is a CYP 1A2 inducer that significantly lowers plasma drug concentrations in smokers compared to nonsmokers by 33 percent. Dose adjustments are not necessary when considering age, as this factor is not significant in tacrine pharmacokinetics.

Pharmacodynamics. Tacrine is a noncompetitive, centrally active reversible cholinesterase inhibitor that inhibits BuChE activity 18 times more than AChE activity in the cerebral cortex.

Design and Interpretation of Clinical Drug Studies

Two pivotal studies positioned tacrine as an effective treatment of mild to moderate AD, one study with a 12-week duration and the other one spanning 30. Both clinical trials found statistically significant differences between tacrine and placebo based on two primary outcome measures: the cognitive subscale of the Alzheimer's Disease Assessment Scale (ADAS-cog) and the Clinician's Interview-Based Impression of Change (CIBIC).

The 12-week study included comparison groups taking placebo, 20, 40, and 80 mg of tacrine per day in an outpatient setting at 23 different centers. 468 patients participated in this study. Statistically significant differences were detected on both primary outcome measures in the group taking 80 mg/day when compared to placebo. Of note, reversible asymptomatic elevations of liver transaminases more than three times the normal value were seen in 25 percent of participants.

The 30-week double-blind, placebo-controlled trial, which involved 663 patients, evaluated tacrine effect at doses of 80, 120, and 160 mg/day. All groups started at 40 mg and the daily dose was increased at 6-week intervals in order to reach the desired goal. Those on the 120-mg group received this dose for the total final 18 weeks and the 160-mg group for the last 12 weeks. The groups using tacrine 120 and 160 mg/day demonstrated statistically significant improvement in the cognitive subscale of the ADAS-cog compared to the placebo group at 30 weeks. A similar finding occurred with the CIBIC scores with these doses. Forty percent of the tacrine 160-mg/day patients showed improvement in ADAS-cog scores by at least four points at week 30. Significant differences between the 80-mg/day group and placebo were not found.

Effects on Specific Organs and Systems

A dose-dependent effect appears to take place for therapeutic response, and the typical time frame necessary for response occurs after at least 12 weeks. Patients experienced elevated liver transaminases, sometimes up to 20 times the upper limit of normal, which represented a common cause for tacrine discontinuation. Several cases of liver toxicity have been reported, with elevated serum bilirubin, hepatitis, and liver failure as hallmarks. Most cases were reversible, but some deaths occurred. Constant liver function monitoring was necessary during treatment. Tacrine could potentially cause vagotonic cardiovascular effects and needed to be used with caution in patients with bradycardia and/or heart block. Due to increased cholinergic activity, tacrine can increase gastric acid secretion, placing patients at increased risk for developing ulcers. Patients who had a history of peptic ulcer disease or were taking nonsteroidal anti-inflammatory drugs (NSAIDs) needed monitoring of gastrointestinal symptom exacerbation when tacrine was being prescribed.

Therapeutic Indications

Tacrine was indicated for mild to moderate AD until it was removed from the market.

Precautions and Adverse Reactions

Clinical trials have reported frequent adverse side effects from tacrine, most notably increased liver transaminases, as mentioned in the effects on specific organs and systems section. These elevations were often asymptomatic but required monitoring, with some patients experiencing liver toxicity. Due to its hepatotoxic profile, tacrine was not indicated for patients who suffered from hepatic impairment. Gastrointestinal problems, including nausea, vomiting, and loose stools, were the second most frequent side effect observed.

Hepatic Monitoring and Rechallenge

Serum transaminase levels, specifically alanine transaminase (ALT), needed to be monitored every other week for 16 weeks, then monthly for 2 months, and then once every 3 months for patients without significant elevations. Dose adjustments were not needed until a value doubling the upper limit of ALT was reached. Weekly monitoring was recommended when two to three times the upper ALT limit occurred. When a level three to five times the upper limit of ALT was reached, tacrine dosage needed to be decreased by 40 mg/day with weekly ALT monitoring. If ALT levels normalized, dosage increase was resumed with every other week monitoring. If ALT level was greater than five times the upper limit, tacrine was discontinued and patients were monitored for symptoms of hepatitis. If jaundice or total bilirubin levels were greater than 3 mg/dL or there was presence of rash, eosinophilia, hepatitis, or fever, patients needed to permanently stop tacrine and not be rechallenged.

Drug Interactions

Being metabolized primarily via CYP IA2, drug–drug interactions could occur when given with agents that underwent the same metabolism. In conjunction with theophylline, tacrine increased its elimination half-life, doubling its concentration in plasma. Cimetidine coadministration results in tacrine increases of C_{max} and AUC by approximately 54 percent and 64 percent, respectively. Fluvoxamine is a potent CYP 1A2 inhibitor that decreased tacrine clearance by eightfold. Close monitoring was needed when prescribing fluvoxamine and tacrine together. Patients on theophylline demonstrated a 50 percent reduction in clearance when tacrine was taken at the same time. Tacrine could potentiate the effects of succinylcholine by inhibition of BuChE, the enzyme responsible for

succinylcholine breakdown. Women being prescribed estrogen replacement therapy (ERT) had enhanced clinical response to tacrine, which could have been due to increased tacrine plasma concentrations produced by ERT. Tacrine could also lengthen the action of succinylcholine-type muscle relaxants during anesthesia. All patients taking cholinesterase inhibitors need to be aware of concomitant use of anticholinergic medications, as they may interfere with their activity due to their mechanism of action.

Laboratory References

Elevation of the serum transaminase ALT is associated with the use of tacrine and occurred in about 28 percent of patients taking the drug. It does not interfere with complete blood counts, electrolytes, urinalysis, or other laboratory tests.

Dosage and Administration

Tacrine was available in 10-, 20-, 30-, and 40-mg capsules. Initial dosing was 10 mg four times a day, which was increased until a maximum of 40 mg four times a day was achieved, based on tolerance of adverse side effects. If liver transaminase elevation occurred, the dose was maintained or reduced. If tacrine was discontinued for longer than 4 weeks, the patient needed to be restarted at the initial dose of 10 mg four times a day with careful monitoring of liver enzymes. If gastrointestinal problems occurred, tacrine could be given with food, but this could result in a decrease in drug bioavailability.

DONEPEZIL

Originally developed in 1986, Donepezil was approved by the FDA in 1997 for its use in mild to moderate AD. It was later approved in 2006 for the treatment of severe AD as well. Unlike its predecessor tacrine, this medication immediately appealed to clinicians thanks to its once a day administration, lack of hepatotoxicity and not having to measure serum liver transaminase levels constantly.

Chemistry

Figure 34.15–2 shows the molecular structure of donepezil hydrochloride. It is a piperidine compound that consists of (±) –2,3 dihydro-5,6-dimethoxy-2-[[1-(phenylmethyl)-4-piperidinyl]methyl]-1H-inden-1-one hydrochloride.

Pharmacological Actions

Pharmacokinetics. Donepezil is not affected by food intake in its regular or oral disintegrating tablet presentations. The relationship between donepezil dose and its area under the plasma concentration time

curve (AUC) is linear at a dose range of 1 to 10 mg. It is bound to plasma proteins by approximately 96 percent (75 percent to albumin and 21 percent to α_1-acid glycoprotein). The donepezil 10-mg formulation reaches peak plasma concentration at approximately 3 hours, compared to 8 hours in the 23-mg tablet. Its volume of distribution at steady state is between 12 and 16 L/kg, reaching steady state after 2 weeks of use.

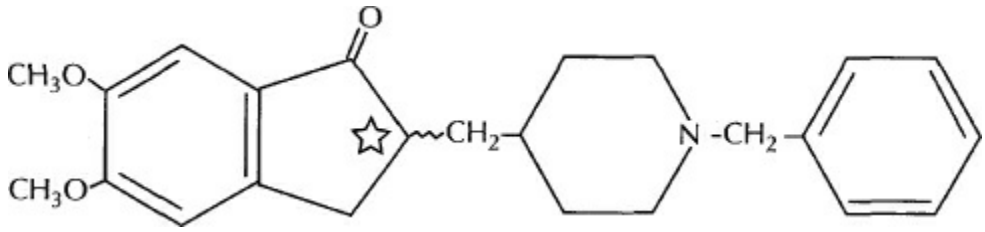

FIGURE 34.15–2. Molecular structure of donepezil.

Donepezil undergoes extensive hepatic metabolism by the CYP 2D6 and 3A4. 6-O-desmethyl-donepezil is its major active metabolite and possesses equal potency. It is also converted into an inactive *cis-N*-oxide metabolite and two glucuronide conjugates. The elimination half-life of donepezil is approximately 70 hours, allowing for once-daily administration. Age, gender, hepatic disease, and renal impairment have not been shown to affect donepezil disposition to any significant degree. Due to this, dose adjustments based on renal or liver function are not necessary. Donepezil is classified as a category C drug. There are currently no adequate trials of donepezil in pregnant women however, animal trials revealed that it may cause fetal harm. It therefore should not be used in pregnant women. It is unknown if it is excreted in breast milk.

Pharmacodynamics. Donepezil enhances cholinergic neurotransmission by increasing the concentration of acetylcholine through reversible, noncompetitive inhibition of AChE, the main enzyme involved in the hydrolysis of acetylcholine. Unlike tacrine, it is not highly selective for inhibition of butyrylcholinesterase (BuChE). It is a potent selective inhibitor that acts on the CNS, and it does not act on the peripheral nervous system.

Design and Interpretation of Clinical Drug Studies

A substantial number of randomized, double-blind, placebo-controlled studies have provided evidence of the effectiveness of donepezil as a treatment for AD at all of its stages. The results of these studies confirmed that donepezil was effective and well tolerated in treating the symptoms of mild to moderate AD, leading to its initial FDA approval. Additionally, the effectiveness of donepezil for the treatment of severe AD has been demonstrated by studies using both 10-mg daily doses and more recently,

23 mg.

To establish donepezil's effect on mild to moderate AD, two large multicenter clinical trials were conducted, both using similar study methods. One of them was a US study that took place during a 15-week period (12 weeks of active treatment followed by a 3-week blind placebo washout period), the other one a multinational study done during a 30-week period (24 weeks of treatment followed by a 6-week washout period). Patients were randomized into three groups: placebo, donepezil 5 mg, and donepezil 10 mg. In both studies, at the end of each treatment period (weeks 12 and 24), participants treated with placebo had a decline in their ADAS-cog score, while donepezil-treated patients showed statistically significant dose-dependent improvement. A scale measuring clinician's assessment of functionality with the help of information provided by a caregiver, called the Clinician's Interview-Based Impression of Change with caregiver input (CIBIC-plus), revealed scores at weeks 12 and 24 in each study that showed statistically significant improvement in the donepezil-treated groups. The 12-week study presented no difference between the 5-mg and 10-mg groups. On the other hand, the international study group demonstrated a statistically significant difference between the two doses, with the 10-mg donepezil group showing greater benefits in all outcome measures when compared to the 5-mg group. After the placebo washout periods, ADAS-cog scores declined in both donepezil-treated groups of the two studies, reaching scores that could not be differentiated from the placebo groups.

A 6-month randomized, double-blind, parallel-group, placebo-controlled study of 248 nursing home dwelling patients with severe AD (Mini-Mental State Examination [MMSE] score of 1 to 10), evaluated the effect of donepezil on cognition and daily functioning. Patients were assigned to donepezil or placebo. Donepezil dose was 5 mg/day for 30 days and then increased to 10 mg/day for the remainder of the study. The primary study end points were changed from baseline to 6 months on the Severe Impairment Battery (SIB) and the Alzheimer's Disease Cooperative Study–Activities of Daily Living–Severe Impairment Version (ADCS-ADL-SIV). The SIB uses a validated 40-item scale developed for the evaluation of the severity of cognitive dysfunction in more advanced AD patients. ADCS-ADL-SIV is a caregiver-based ADL scale composed of 19 items developed for use in clinical studies of dementia. When compared to baseline at 6 months, patients treated with donepezil displayed more improvement in SIB scores and less decline in the ADCS-ADL scores than placebo in severe AD subjects.

The use of donepezil 23 mg in moderate to severe AD was approved by the FDA in 2010, based on the conclusions of a multinational, randomized, double-blind study that included patients who were already being treated with a 10-mg dose for at least 3 months. 1,467 patients were randomly

assigned to receive either a 23 mg once a day dose of donepezil or a standard dose of 10 mg for 24 weeks. At week 24, changes in cognition and global functioning were assessed with the SIB and the CIBIC-plus, respectively. This study revealed a statistically significant benefit in cognition with the 23-mg dosing, but did not find significant change in global functioning. Side effects were also monitored, revealing increased incidence of gastrointestinal symptoms, which included nausea, vomiting, and diarrhea, in the higher 23-mg dose group.

One randomized placebo-controlled study was conducted over a 1-year period, evaluating the preservation of function in 431 patients with probable AD who received either placebo or donepezil. Participants randomized into the active treatment arm were started on 5 mg with a posterior dose increase to 10 mg on day 28. Function was assessed by evaluating basic activities of daily living (ADLs) and instrumental ADLs through the AD Functional Assessment and Change Scale. The results indicated that after 1 year, donepezil had a 38 percent reduction in the risk of functional decline when compared to placebo. Moreover, this study showed how donepezil delayed functional decline by 5 months in patients with AD. At 48 weeks, the probability of survival with no evidence of functional loss was 51 percent for those taking donepezil, compared to only 35 percent for those on placebo. These results demonstrated that continued use of donepezil benefits patients over a period of 1 year, and that untreated patients tend to decline in their functional ability.

Effects on Specific Organs and Systems

In the elderly, hepatic function is frequently decreased, leading to reduced tolerability of certain medications. Donepezil does not induce elevation of liver transaminases or causes hepatotoxicity and is, in general, well tolerated. Even though kidney function may also decrease with aging, donepezil may be safely administered in patients with moderate to severe renal impairment. This medication may produce vagotonic effects in the sinoatrial and atrioventricular nodes by acting on the muscarinic receptors, causing slowing of the heart rate. Like all cholinesterase inhibitors, it should be used with caution in patients with cardiac arrhythmias and significant cardiac dysfunction. Due to its increased cholinergic activity, donepezil may also augment gastric acid secretion in a dose-dependent fashion, causing gastrointestinal complications that potentially may increase the risk of gastric disease and gastrointestinal bleeding.

Therapeutic Indications

Donepezil is indicated for the treatment of mild, moderate, and severe AD. Additionally, select randomized, double-blind, placebo-controlled studies have demonstrated some benefit in cognition and functioning when

treating vascular dementia. Donepezil has not been approved by the FDA for this use, as there is concern for its safety in this population due to cardiovascular and cerebrovascular risks. Similarly, small studies have shown improvement of cognition in patients with LBD, but the use of donepezil in this population remains off-label.

Precautions and Adverse Reactions

The most common adverse reactions leading to discontinuation and poor tolerability of dose optimization in clinical settings include diarrhea, nausea, vomiting, insomnia, muscle cramps, fatigue, and anorexia. These side effects are usually mild and transient and have not resulted in a significant dropout rate during clinical trials with donepezil. Less commonly, donepezil can potentially cause more severe reactions such as seizures, cerebrovascular disease, or exacerbation of chronic obstructive pulmonary disease (COPD).

The incidence of gastrointestinal adverse side effects did not differ between patients with mild to moderate dementia when compared to severely ill patients. Patients who suffer from peptic ulcer disease, or are taking NSAIDs, need to be carefully monitored for gastrointestinal bleeding. Caution should also be exerted when using the 23-mg dose in patients who weigh less than 55 kg, as there have been increased reports of nausea, vomiting, and weight loss in this group. Due to donepezil's effect on cardiac conduction, donepezil should be used with caution in patients with cardiac arrhythmias, especially sick sinus syndrome. Patients may be at risk of developing bradycardia. It is recommended to titrate donepezil slowly to minimize the incidence of side effects. Additionally, donepezil is contraindicated in patients with known hypersensitivity to the drug or to piperidine derivatives.

Drug Interactions

Although metabolized mainly through the CYP 2D6 and 3A4, and highly protein bound at 96 percent, donepezil has minimal influence on drugs metabolized through these pathways. Pharmacokinetic studies have been conducted assessing the induction effects of donepezil on cimetidine, ketoconazole, risperidone (which is also metabolized through 2D6 and 3A4), warfarin and digoxin, and no effects have been observed on the pharmacokinetics of these drugs.

In vitro studies of inhibitors of CYP 3A4, such as ketoconazole, and 2D6, such as quinidine, have shown inhibition of donepezil metabolism. These studies have not translated into clinically significant drug–drug interaction. Phenytoin, carbamazepine, dexamethasone, rifampin, and phenobarbital, which are inducers of 2D6 and 3A4 pathways, can potentially increase donepezil's elimination rate. Donepezil is not affected

by the concomitant use of cimetidine or digoxin. As with all cholinesterase inhibitors, caution is advised when patients are taking anticholinergic medications, as donepezil may interfere with their activity due to its mechanism of action. Donepezil may also lengthen the action of succinylcholine-type muscle relaxants during anesthesia.

Laboratory Interferences

Donepezil has not been shown to interfere with any laboratory test including biochemical, urinalysis, or spectrometry.

Dosage and Administration

Donepezil is available in 5-, 10-, and 23-mg tablets and in an orally disintegrating tablet form of 5 and 10 mg. The recommended starting dose is 5 mg once a day at bedtime. After 4 to 6 weeks, the dose can be increased to 10 mg depending on tolerability. For mild to moderate AD, 5 and 10 mg are the recommended effective doses. In patients presenting with moderate to severe dementia, clinical trials have found doses of 10 and 23 mg effective. In the moderate to severe group, for those who are tolerating it well, the dose may be increased to 23 mg after a minimum of 3 months on a 10-mg dose. The medication should be administered at bedtime so that peak plasma concentrations occur during the night and potential adverse effects are minimized. As food intake does not affect its absorption, donepezil may be taken with or without food.

RIVASTIGMINE

Rivastigmine has been approved for the treatment of mild to moderate AD since 2000 by the FDA. It is also the only cholinesterase inhibitor approved, since 2006, for the treatment of mild to moderate dementia due to major neurocognitive disorder related to PD. Additionally, in 2013, rivastigmine was approved for severe AD. Rivastigmine is a second-generation cholinesterase inhibitor and has a different pharmacological profile in comparison to the earlier cholinesterase inhibitors, tacrine and donepezil.

Chemistry

Rivastigmine is composed of the chemical compound (S)-3-[1-(dimethylamino)ethyl]phenylethylmethylcarbamate. Its molecular structure is shown in Figure 34.15–3.

Pharmacological Actions

Pharmacokinetics. Rivastigmine is rapidly absorbed in its oral formulation, with a mean time to maximum plasma concentration (T_{max}) of

1 hour (range 0.8 to 1.7 hours). Additionally, it has a mean absolute bioavailability of 60 percent in comparison to intravenous drug administration. When ingested with food, its C_{max} is decreased by 30 percent and increases the total amount of drug in the body (AUC) by 30 percent. Because its bioavailability is unaffected, it is recommended that oral rivastigmine always be given with food to potentially prevent the drug's most common side effects, which are nausea, vomiting, and other forms of gastrointestinal distress. Rivastigmine is only 40 percent bound to plasma proteins with a volume of distribution between 1.8 and 2.7 L/kg. The relationship between drug dose and AUC is linear with a mean elimination half-life of 1.6 hours. The drug easily crosses the blood–brain barrier. Cerebrospinal fluid (CSF) rivastigmine concentrations are rapidly reached with a T_{max} of 1.4 to 3.8 hours, and it is quickly eliminated from the CSF with a half-life of 0.31 to 2.95 hours. Rivastigmine is converted after activity to an inactive metabolite which undergoes *N*-demethylation and sulfate conjugation prior to elimination. Rivastigmine is mainly eliminated through renal excretion, and it is not metabolized by hepatic oxidative CYP isozymes. Age, hepatic impairment, and renal insufficiency do not significantly affect rivastigmine disposition and dosage adjustments are not necessary in those individual patients. Additionally, the drug is not affected by inducers or inhibitors of CYP450 metabolism. There are currently no adequate trials of rivastigmine in pregnant women; however, animal trials have revealed no evidence of teratogenicity. It is unknown if it is excreted in breast milk.

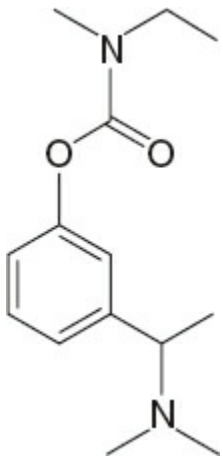

FIGURE 34.15–3. Molecular structure of rivastigmine.

A rivastigmine transdermal patch (5 to 20 cm²) was developed and approved by the FDA in 2007. The 10-cm² patch pharmacokinetic profile is comparable to the oral 6 mg twice-daily dose, while the 20-cm² patch is similar to 9 to 10 mg twice-daily oral dose. The patch showed a 30 percent lower C_{max}, a longer T_{max} (8 hours vs. 1 hour), and a 1.8-fold greater

increase in AUC compared to the oral dose. Markedly less fluctuation in plasma drug concentrations between peak and trough levels were found with the patch. The 10-cm^2 patch showed adequate efficacy with fewer side effects than the 20-cm^2 patch, and therefore was initially recommended as the primary target dose in mild to moderate AD and major neurocognitive disorder due to PD. A higher, 15-cm^2 transdermal patch formulation is currently available and is the FDA recommended dose for patients with severe AD. Drug exposure is optimal when the rivastigmine patch is applied to the upper arm, chest, or back. These locations are the recommended application sites.

Pharmacodynamics. Rivastigmine is a noncompetitive, pseudoirreversible cholinesterase inhibitor due to its prolonged enzyme inhibition activity, where a single dose was shown to produce inhibition for as long as 10 to 12 hours. Rivastigmine displays equal potency in inhibiting both AChE and BuChE reported from both in vitro and in vivo data. This period of activity inhibition indicates that twice a day dosing is optimal. A linear relationship was reported between rivastigmine dose and AChE and BuChE inhibition measured in CSF. Plasma BChE inhibition rose slowly with the patch compared to two distinct peaks with the oral twice-daily dosing.

Design and Interpretation of Clinical Drug Studies

There have been a number of randomized, placebo-controlled, double-blind studies that have demonstrated the efficacy of rivastigmine in the treatment of AD. The majority of trials have been conducted on community dwelling subjects with mild to moderate AD, and were of 26-week duration or shorter. Doses ranged from 6 to 12 mg daily for oral doses and 5 to 20 cm^2 for the transdermal patch. The more efficacious oral dose was found to be 6 to 12 mg daily, higher doses demonstrating increased side effects and worse tolerability. For the transdermal patch, doses ranged from 5 to 20 cm^2, with 10 cm^2 having better tolerability than higher doses with significant clinical outcome. Moderate improvement in cognitive and functional scales occurred, as well as impression of improvement as assessed by the subjects' clinician. These improvements demonstrated equivalent efficacy with donepezil and other cholinesterase inhibitors.

Specifically for the oral drug form, two placebo-controlled multicenter clinical studies with rivastigmine (N = 1,424) were conducted prior to FDA approval. Subjects were randomized into placebo group, rivastigmine 1 to 4 mg/day, and rivastigmine 6 to 12 mg/day. Drug doses were adjusted weekly for the first 7 weeks as subjects reached the minimum assigned dose range. From weeks 8 to 26, subjects could be maintained on a flexible dose but in their assigned range to maximize therapeutic benefit and minimize side effects. Both studies showed a dose-dependent improvement effect on

ADAS-cog scores from baseline to the study end point. The placebo group showed a decline of two to four points in the ADAS-cog, while the rivastigmine 6 to 12-mg/day group had about a one-point increase in ADAS-cog score. The rivastigmine 1 to 4-mg/day group had approximately a one- to two-point decline in ADAS-cog scores. Both rivastigmine groups showed significant benefit in the CIBIC-plus scores compared to placebo.

Additionally, the subjects in one of these large trials continued into an open-label extension study where both, the placebo and rivastigmine groups, were treated with rivastigmine 6 to 12 mg/day for an additional 26 weeks. In this study, both the placebo and rivastigmine groups showed improvement at the study end point, compared to the projected rate of decline in untreated subjects. However, their scores were lower than the rivastigmine 6 to 12 mg/day group with about a three- to four-point decrease in ADAS-cog scores at week 52. Subjects treated with rivastigmine 6 to 12 mg/day had almost maintained their ADAS-cog scores from baseline with an approximate two-point decrease after 52 weeks. These results imply that with a 6-month delay in treatment or inadequate treatment with low doses, subjects may not be able to demonstrate that same level of improvement as patients with therapeutic treatment from the start. For maximum benefit, when a diagnosis of AD is established, treatment with cholinesterase inhibitors such as rivastigmine, when indicated, should be started as soon as possible. Dosing at suboptimal levels for extended periods can also hinder drug efficacy.

In a 24 week, randomized double-blind trial, patients with severe AD were randomized, receiving a 4.6 mg/24 hr patch versus a 13.3 mg/24 hr rivastigmine patch. There was an 8-week titration phase and a 16-week maintenance phase. As in the donepezil severe AD clinical trial, the SIB as well as the ADCS-ADL-SIV were used as clinical end points. Decline in both the SIB and the ASCS-ADL-SIV were both significantly less at 16 and 24 weeks for the 13.3 mg/24 hr dose.

In a 24-week, double-blind, placebo-controlled trial, rivastigmine with doses ranging from 3 to 12 mg was evaluated in 541 patients with PD dementia. Diagnosis was based on a 2-year or greater history of idiopathic PD before the onset of dementia, and patients did not meet criteria for other etiological causes of dementia. Based on ADAS-cog scores comparing baseline to week 24, rivastigmine-treated PD dementia patients had a mean improvement of 2.1 points, whereas the placebo group declined by a mean of 0.7 points. The Alzheimer's Disease Cooperative Study Clinical Global Impressions of Change (ADCS-CGIC) scores also showed benefit for rivastigmine, as 19 percent had improvement compared to 13 percent for placebo. Clinically meaningful worsening was reported to be only 13 percent for rivastigmine compared to a higher occurrence of 23 percent with placebo. Of note, the most frequent adverse side effects for rivastigmine were gastrointestinal, with nausea in 29 percent and vomiting

in 16 percent, as well as tremors in 10 percent of patients. Worsening tremor may signify increased clinical relevance in PD dementia patients and warrants closer monitoring with rivastigmine treatment.

Effects on Specific Organs and Systems

Rivastigmine does not produce hepatotoxicity or liver transaminase elevations, and therefore no monitoring of liver function tests are required with its use. Like other cholinesterase inhibitors, rivastigmine can produce vagotonic cardiovascular effects and should be used cautiously in patients with cardiac arrhythmias and cardiac conduction delays. Gastrointestinal problems can occur with all cholinesterase inhibitors, but is particularly seen with rivastigmine use in a more dose-dependent manner. Rivastigmine can also increase gastrointestinal acid production. Patients on NSAIDs, with peptic ulcer disease or other gastrointestinal illnesses should be carefully monitored, as they may be at greater risk of gastrointestinal bleeding.

Therapeutic Indications

Rivastigmine is indicated for the treatment of mild to moderate AD. It is also the only cholinesterase inhibitor with an FDA indication for mild to moderate dementia associated with PD. Similarly, rivastigmine has demonstrated benefits in LBD.

Precautions and Adverse Reactions

The most common adverse side effects with rivastigmine are typical of all the cholinesterase inhibitors. These include diarrhea, nausea, vomiting, insomnia, muscle cramps, fatigue, and anorexia. These side effects are usually mild and transient but in clinical trials, rivastigmine-treated groups had a higher drop-out rate than placebo, usually due to GI side effects. Additionally, in a study comparing rivastigmine with donepezil in AD patients, oral rivastigmine-treated participants had a significantly higher drop-out rate due to side effects. Gastrointestinal side effects with rivastigmine are not different when comparing mild to moderate individuals with those with moderate to severe AD. The transdermal patch formulation demonstrates much less gastrointestinal problems, most likely because of reduced plasma drug concentration, due to longer T_{max} and lower C_{max} when compared to the oral preparation. Less commonly, rivastigmine can potentially cause adverse reactions such as seizures, cerebrovascular disease or worsen COPD. It can also worsen urinary obstruction. Reanalysis of the electrocardiograph data from the clinical trial data revealed no significant differences between rivastigmine and placebo in heart rate, PR, QRS, and QTc interval. However, patients with cardiac conduction delays and significant heart disease should use

rivastigmine with caution, given the potential for bradycardia with cholinesterase inhibitors. The transdermal patch formulation reported skin erythema barely visible or mild that self-resolved within 24 hours. The upper back site, arms, and chest had a lower incidence of erythema compared to the abdomen and thigh sites.

Drug Interactions

Because rivastigmine is not metabolized by hepatic oxidative CYP isozymes, it does not interact with substrates of CYP450. Subsequently, there have been no drug interactions reported with digoxin, warfarin, diazepam, or fluoxetine. A retrospective analysis of the clinical trials did not find an increase in adverse side effects in patients who received various medications, including β-blockers, antihypertensives, and other drugs that may potentially produce pharmacodynamic interactions such as bradycardia. As with all cholinesterase inhibitors, caution is advised when patients are taking anticholinergic medications. Rivastigmine may also increase the action of succinylcholine-type muscle relaxants during anesthesia.

Laboratory Interferences

Rivastigmine has not been shown to interfere with any laboratory test, urinalysis, or spectrometry.

Dosage and Administration

Oral rivastigmine is available in 1.5-, 3.0-, 4.5-, and 6.0-mg capsules. Based on clinical trials, suggested dosage range is 3 to 12 mg/day. Patients should be initiated at 1.5 mg twice a day and gradually increased every 2 to 4 weeks by 1.5 mg daily until the maximum daily dose of 12 mg/day is reached or titrated to the patient's tolerability. It is highly recommended to take food with rivastigmine to minimize side effects. The transdermal patch preparation (5 to 20 cm^2) is able to be applied once a day due to its pharmacokinetic and pharmacodynamic profile for cholinesterase inhibition, reported to produce a slow and consistent release of the medication. The average exposure of transdermal patch dose of 10 cm^2 was comparable to an oral dose of 9.5 mg/24 hr. Patients should be initiated at a dose of 5 cm^2/4.6 mg daily and increased to 10 cm^2/9.5 mg in 4 weeks. Once this dose is maintained for 4 weeks and the patient tolerates it well, the maximum effective dose of 13.3 mg/day may be given. The slow release of rivastigmine from the patch should improve its gastrointestinal tolerability.

GALANTAMINE

Galantamine was approved by the FDA for the treatment of AD in 2001 and

is the fourth cholinesterase inhibitor made available in the United States. Galantamine is extracted from daffodils of the snowdrop plant *Galanthus nivalis* and is a phenanthrene alkaloid similar to codeine.

Chemistry

Galantamine is a (4aS,6R,8aS)–4a,5,9,10,11,12-hexahydro-3-methoxy-11-methyl-6H-benzofuro [3a,3,2–ef][2]benzazepin-6-ol hydrobromide compound. Its molecular structure is shown in Figure 34.15–4.

Pharmacological Actions

Pharmacokinetics. Galantamine immediate release (IR) has an 85 to 100 percent bioavailability in comparison of tablets to oral solution, and is readily absorbed with a T_{max} from 0.5 to 2.0 hours. C_{max} is decreased by 25 percent with a slight delay of T_{max} by 1.5 hours with food intake. Coadministration with food does not affect galantamine IR AUC. The volume of distribution at steady state was 2.6 L/kg with a very low protein binding of 18 percent. The elimination half-life of galantamine IR is about 7 hours. The bioequivalence of galantamine extended release (ER) 24 mg was reported to be within the 90 percent confidence intervals for AUC_{24} and C_{min} of galantamine IR 12-mg tablets twice daily. The mean time to peak plasma concentration was 4.4 hours with an elimination half-life of 8.3 hours after fasting and 8.0 hours after food intake. Galantamine ER formulation does not differ in its pharmacokinetic profile whether giving it with food or while fasting.

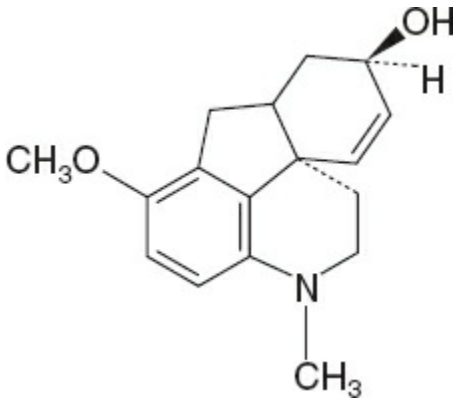

FIGURE 34.15–4. Molecular structure of galantamine.

Galantamine is mainly metabolized by hepatic CYP 2D6 and CYP 3A4 isozymes. Sanguinine (*O*-desmethylgalantamine) is formed by CYP 2D6 and accounts for 20 percent of the parent drug. Sanguinine is its major active metabolite and was shown to be three times more potent that galantamine in producing AChE inhibition. Other metabolites include galanthaminone and *N*-desmethylgalantamine, but they are much less

potent than sanguinine. Although a 25 percent decrease in drug clearance of galantamine IR was found between poor metabolizers and extensive metabolizers of CYP 2D6, and a 50 percent higher drug exposure in galantamine ER, these findings are probably not significant due to the large interpatient variability in galantamine disposition. The doses can be titrated to individual tolerability.

Age and mild hepatic impairment (defined as a Child–Pugh score of five to six) was shown not to affect galantamine pharmacokinetics. Patients with moderate hepatic impairment (Child–Pugh score of seven to nine) had galantamine AUC and elimination half-life increased by 30 percent. Patients with mild renal impairment had galantamine concentrations similar to Alzheimer patients without renal impairment. No dosage adjustments are needed for galantamine in elderly patients or in patients with mild hepatic impairment. A maximum dose of 16 mg/day is recommended for patients with moderate hepatic impairment. Galantamine is contraindicated in patients with severe renal (creatinine clearance less than 9 mL/min) and severe hepatic impairment (Child–Pugh score greater than nine). Galantamine is classified as a category C drug with no trials in pregnant women and animal trials revealing potential teratogenicity. Additionally, it is unknown if it is excreted in breast milk.

Pharmacodynamics. Galantamine is a competitive, reversible inhibitor that is significantly more selective for AChE over BuChE based on in vitro data. Galantamine also produces allosteric modulation of nicotinic receptors that are distinct from acetylcholine and nicotinic agonists or antagonists. Allosteric modulators upregulate or potentiate ion channel activity of nicotinic receptors in response to acetylcholine. This mechanism of action may play a unique clinical role as AD is associated with a deficit in nicotinic transmission. By enhancing the release of acetylcholine, this action can additionally modulate the release of other neurotransmitters such as GABA, serotonin, and glutamate.

Design and Interpretation of Clinical Drug Studies

A number of large, multicenter clinical trials with galantamine have been conducted. 978 subjects were evaluated in the main clinical trial that established the minimum effective galantamine dose. After a 4-week single-blind placebo treatment, patients were randomized into four treatment groups: placebo, galantamine 8, 16, and 24 mg/day. Patients were initiated on 8 mg/day and the dose was increased by 8 mg/day after 4 weeks until the target dose was achieved. The total study duration for the randomized groups was 5 months. All three groups showed improvement in ADAS-cog scores from baseline to the study end point when compared to placebo. During the intent-to-treat analysis, the 16- and 24-mg/day groups were found to be superior to placebo, but not the 8-mg group. The CIBIC-plus

evaluations showed that 64 to 68 percent of the patients in the 16- and 24-mg/day groups remained stable or improved after 5 months, versus 47 percent of the placebo group. The 8-mg/day group had only a 53 percent reported incidence of stability or improvement.

Several multicenter double-blind, placebo-controlled clinical trials have evaluated galantamine at higher doses of 18, 24, and 32 mg/day. Each study employed a different dosing strategy to reach these three dose groups, with a starting dose of 8 mg/day. Doses were either increased weekly or every 2 to 3 days until the assigned target dose was reached. The duration of the studies ranged from 12 weeks to 6 months. Each trial revealed that the 18- and 24-mg/day dose groups were superior to placebo. The 32-mg/day group also showed improvement compared to placebo, but did not reveal increased benefit when compared to the 24-mg group. The incidence of gastrointestinal adverse effects was highest in the 32-mg/day group, thereby limiting its usefulness. The results of these studies suggested that the optimal dose of galantamine is 24 mg daily.

One of these clinical trials reported a 6-month extension where patients assigned to the placebo group were placed on galantamine 24 mg/day. After the subsequent 6 months, the placebo/galantamine group had ADAS-cog scores that resembled the galantamine 32-mg group, which had its dose decreased to the 24-mg/day dose group, and above the projected rate of decline in untreated patients. The original 24-mg/day dose group at 6 months had ADAS-cog scores above baseline and at the 12-month time slightly decreased to baseline scores. The two other groups (placebo/galantamine and galantamine 32 mg or 24 mg) had ADAS-cog scores below the original 24-mg dose group at the 12-month time period. These results suggest that 24-mg dose is the optimal galantamine dose and that a delay of 6 months by the placebo group could worsen outcome in AD.

A 24-week double-blind, placebo-controlled study compared galantamine IR to galantamine ER in 971 patients with mild to moderate AD. Galantamine IR and ER were started at 8 mg/day and to 16, and then to 24-mg by 8-mg increments every 4 weeks as tolerated. The IR and ER were significantly better than placebo in regard to cognition, measured by the ADAS-cog, but not in global change (CIBIC-Plus), where they did not differ from placebo. Although these findings with CIBIC-Plus differ from the original IR studies, the secondary ADCS-ADL measure favored ER over placebo.

Effects on Specific Organs and Systems

Galantamine does not elevate hepatic transaminases and is not known to cause hepatotoxicity. As the other cholinesterase inhibitors, this medication may cause vagotonic cardiovascular effects and should be used cautiously in patients with cardiac arrhythmias. Gastrointestinal problems may occur in a dose-dependent manner, and patients with peptic ulcer

disease and other gastrointestinal problems should be carefully monitored.

Therapeutic Indication

Galantamine is indicated for the treatment of mild to moderate AD.

Precautions and Adverse Reactions

The most commonly reported side effects of galantamine are dizziness, headache, nausea, vomiting, diarrhea, anorexia, and weight loss. These side effects are mild and transient and did not produce a significant dropout rate during clinical studies. As previously stated, patients with renal and hepatic impairment may have substantially increased levels of galantamine. A maximum daily dose of 16 mg/day is the recommended dose for patients with moderate hepatic impairment. Galantamine is contraindicated in patients with severe renal and severe hepatic impairment. Less commonly, galantamine can potentially cause more severe reactions such as seizures, cerebrovascular disease, or exacerbation of COPD. The adverse side-effect profile of galantamine ER is similar to that of the IR formulation.

Drug Interactions

Galantamine is metabolized by hepatic CYP 2D6 and CYP 3A4 isozymes and could have significant drug–drug interactions. Galantamine AUC was reported to be increased by 40 percent when coadministered with paroxetine, 30 percent when given with ketoconazole, and 12 percent when erythromycin was coadministered. Patients should be monitored closely when known CYP 2D6 and CYP 3A4 inhibitors are given along with galantamine. As with all cholinesterase inhibitors, caution is advised when patients are taking anticholinergic medications, as galantamine may interfere with their activity due to its mechanism of action. It may also lengthen the action of succinylcholine-type muscle relaxants during anesthesia.

Laboratory Interferences

Galantamine has not been shown to interfere with any laboratory test, including urinalysis, spectrometry and biochemical.

Dosage and Administration

Galantamine IR is available in 4-, 8-, and 16-mg tablets. The suggested dose range is 16 to 32 mg/day given twice a day in a divided dose. It should be given with food. The 32-mg dose is less well tolerated than lower doses, and clinical trial data do not indicate significant benefit. If patients tolerate the 24-mg dose, an increase to the 32-mg dose can be warranted. The initial dose is 8 mg/day, and after a minimum of 4 weeks, the dose can be

increased. Further dose increases should be based on tolerability, with a minimum time period of 4 weeks on the previous dose. An oral solution is also available in a 4 mg/mL presentation. It should be taken as the IR form. Galantamine ER is available in 8-, 16-, and 24-mg tablets. The ER formulation should be taken with food with an initial dose of 8 mg daily in the morning and increasing the dose to 16 mg daily after a minimum of 4 weeks.

MEMANTINE

Memantine is the only noncholinesterase inhibitor indicated for the treatment of AD. It was first synthesized in the early 1960s as a possible hypoglycemic agent, but was found to be devoid of such activity. It was not until 2003, after multiple studies revealed its utility in dementia, that the FDA approved memantine for the treatment of moderate to severe AD. This agent is not indicated for mild to moderate AD as are the cholinesterase inhibitors.

Chemistry

Memantine is a 1-amino-3,5-dimethyladamantane hydrochloride, a derivative of amantadine (Symmetrel) (1-amino-adamatane). Its molecular structure is shown in Figure 34.15–5.

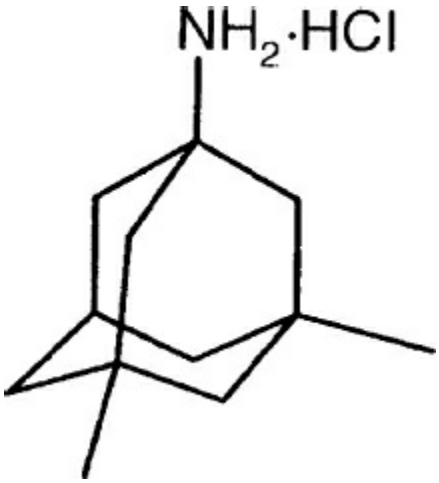

FIGURE 34.15–5. Molecular structure of memantine.

Pharmacological Actions

Pharmacokinetics. Memantine is absorbed completely from the gastrointestinal tract and achieves peak T_{max} within 3 to 7 hours. Protein binding is low from 42 to 45 percent and readily crosses the blood–brain barrier with a CSF serum concentration of 0.52. The elimination half-life of memantine is between 60 and 80 hours and is excreted unchanged in the urine. Steady-state conditions are reached in about 21 days. In a study

comparing 28 mg once daily of memantine XR to 10 mg twice daily of memantine, C_{max} and AUC0–24 values were 48 and 33 percent higher for the XR dosage regimen, respectively. Memantine is largely excreted unchanged in the urine and the remaining is converted to three minimally active polar metabolites. Memantine is a weak base with a pKa of 10.27, and urinary pH can affect renal clearance. Reduced renal clearance occurs at alkaline urinary pH (pH 8) and results in increased plasma concentrations of drug. Memantine has linear pharmacokinetics and with doses of 10 to 20 mg/day, has produced minimal inhibitions of CYP enzymes, having no significant interactions with CYP inhibitors. Pharmacokinetic studies in healthy young adults and AD patients indicate that memantine can be safely coadministered with cholinesterase inhibitors. With repeated administration, total serum concentrations reach 0.5 μM, which corresponds to the Ki value of memantine at the MK-801 binding site in the human frontal cortex, indicating that NMDA receptor antagonism occurs at therapeutic concentrations. There are currently no adequate trials of memantine in pregnant women, and animal trials have revealed minimal evidence of teratogenicity. It is unknown if it is excreted in breast milk.

Pharmacodynamics. The pharmacological effects of memantine are believed to result from the selective uncompetitive inhibitor blockade of the NMDA receptor class of glutamate receptors. Memantine does block the binding of glutamate to the receptor, but it blocks the ion current (Ca^{+2}, Na^{+}) through the receptor channel activated by glutamate. Memantine blocks the NMDA receptor with a strong voltage dependence and rapid blocking and unblocking kinetics, dissociating from the receptor in response to membrane depolarization. Blockade of the NMDA receptor channel during chronic, low-level excitation is thought to be the therapeutic benefit in neurodegenerative diseases such as AD.

Design and Interpretation of Clinical Drug Studies

Memantine indication for moderate to severe AD was based on two multicenter clinical trials. The first was a double-blind, placebo-controlled study comparing memantine 20 mg/day versus placebo in moderately to severely ill Alzheimer patients (N = 252). They had MMSE scores between 3 and 14 or a stage 5 or 6 on the Global Deterioration Scale (GDS). The study duration was 28 weeks, with an optional 24-week open-label extension phase. Patients receiving memantine had a significantly better outcome than placebo patients measured by the CIBIC-plus, ADCS-ADLsev and SIB at 28 weeks. The second study was a double-blind, placebo-controlled trial, but it had a single-blind, placebo lead-in period of 1 to 2 weeks to assess compliance. Afterward, patients were randomized to drug or placebo for the next 24 weeks. Memantine dose was initiated at 5

mg/day, with weekly increases to 20 mg/day at the beginning of week 4. The primary outcome measures were the SIB, CIBIC-plus, and the ADCS-ADLsev at week 24. Patients were allowed to remain on stable donepezil doses of 5 to 10 mg/day throughout the trial if medication was prescribed for at least 6 months prior to study entrance. The change in SIB, CIBIC-Plus, and ADCS-ADLsev outcome scores were all statistically significant in favor of memantine versus placebo. The SIB scores had a slight increase above baseline at week 12 and then remain above baseline by a mean of 1 point at 24 weeks. ADCS-ADLsev scores showed at gradual decline from baseline to 24 weeks for the placebo group, while the drug group displayed a slower rate of decline.

A 24-week open-label extension trial with moderate to severely ill patients reported continued benefit and a much slower decline in ADCS-ADL scores in those patients who were initially treated with memantine. Patients assigned to the placebo group were then placed on memantine. The newly treated patients showed less worsening of SIB, CIBIC-plus, and ADCS-ADL scores compared to the projected decline in untreated patients. At the end of 52 weeks, both groups had similar SIB, CIBIC-Plus, and ADCS-ADL scores and significantly higher scores than the projected scores for untreated patients. These findings indicate that memantine-treated patients continued to benefit from medication and had significantly delayed impairment compared to projections in untreated patients.

Although memantine is indicated only for moderate to severely ill patients, trials have been conducted on mild to moderate AD patients. These trials did not result in FDA indication for the mild stage of AD. A 24-week double-blind, placebo-controlled trial (N = 403) compared 20-mg memantine per day to placebo in mild to moderately ill AD patients, defined by MMSE scores of 10 to 22 points. The outcome measures included the ADAS-cog, CIBIC-plus, and ADCAS-ADL rating scales, comparing baseline to week 24. The study results indicated that only modest treatment effect in favor of memantine had occurred. The mean ADAS-cog score for the placebo group increased by 1.1 points, while the memantine group score decreased by 0.8 point (P = 0.003). The mean end point CIBIC-plus scores for placebo and memantine groups were 4.52 and 4.20, respectively (P = 0.004). A significant difference was not found in ADCS-ADL scores between the two groups.

Effects on Specific Organs and Systems

Memantine does not cause an increase in hepatic transaminases or hepatotoxicity. Memantine XR terminal elimination half-life increased by 16 percent in patients with moderately impaired liver function compared to controls. It is recommended that patients with severe hepatic impairment take memantine ER with extreme caution. Memantine is mainly excreted through the kidneys. Even though there is scarce data about renally

impaired patients taking memantine, drug doses should be reduced in patients with mild to moderate renal impairment. For patients with severe renal impairment, memantine is not recommended.

Therapeutic Indications

Memantine is indicated for the treatment of moderate to severe AD.

Precautions and Adverse Reactions

The most common side effects observed with memantine are constipation, dizziness, headache, and confusion. These effects are usually mild and transient and did not result in a significant dropout rate during clinical trials. With the use of memantine ER, the most common side effects are headache, diarrhea, and dizziness.

Drug Interactions

The use of memantine along with other NMDA antagonists such as amantadine and dextromethorphan has not been evaluated and this combination should be used with caution. In vitro studies with CYP substrates 1A2, 3A4, 2A6, 2C9, 2D6, 2E1, and 3A4 showed minimal inhibition by memantine. Therefore, no pharmacokinetic interaction by these enzymes should occur. Using memantine with donepezil did not affect the pharmacokinetics of either compound. The adverse side-effect profile of the combination of memantine and donepezil was similar to that of patients treated with donepezil alone. Memantine is eliminated by renal tubular secretion and, if administered with other drugs that are excreted through the same mechanism, the plasma concentration of both agents could be altered. Examples of these drugs include hydrochlorothiazide, triamterene, cimetidine, ranitidine, quinidine, and nicotine. Concomitant use of memantine, hydrochlorothiazide, and triamterene did not affect the disposition of memantine or triamterene, but the bioavailability of hydrochlorothiazide was decreased by 20 percent. Memantine clearance was reduced by 80 percent with alkaline urine conditions at pH 8. If the urine pH is altered, drug accumulation may occur. This condition could be caused by diet, drugs like carbonic anhydrase inhibitors and sodium bicarbonate, and medical conditions such as renal tubular acidosis. Memantine should be cautiously used in this situation.

Laboratory Interferences

Memantine has not been shown to interfere with any laboratory test, including biochemical, urinalysis, and hematology. No clinically significant electrocardiographic changes were noted during clinical trials.

Dosage and Administration

Memantine is available in 5- and 10-mg tablets, with a recommended starting dose of 5 mg daily. The recommended target dose is 20 mg/day. Drug is administered twice daily in separate doses with 5-mg increment increases weekly depending on tolerability. Memantine ER is available in 7-, 14-, 21-, and 28-mg capsules. Starting dose is 7 mg daily, with a target dose of 28 mg daily. The ER formulation should be increased no greater than 7 mg weekly. Patients with renal impairment should have a target dose of 14 mg daily. A drug is now available as a combination of memantine XR and donepezil (Namzaric). It is available in 14-mg/10-mg and 28-mg/10-mg formulations. It is only indicated for patients with severe AD already stabilized on memantine ER 14 or 28 mg and donepezil 10 mg.

FUTURE DIRECTIONS

In more recent years, drugs have been actively developed with different pharmacodynamic mechanisms to modify and delay the onset of AD. Multiple treatment trials, with the goal of slowing or even reversing the course of cognitive and functional decline among AD subjects are ongoing. They include secretase inhibitors, amyloid-β inhibitors, tau and amyloid-β-aggregation inhibitors as well as vaccines against amyloid-β and tau. These potentially disease-modifying therapies focus mainly on agents that will decrease amyloid-β content and tau hyperphosphorylation. To date, trials have not demonstrated any significant improvement in the decline of patients already suffering the full clinical picture of AD. Current research has shifted to better recognition of AD in its earliest clinical phases. The aim is to enlist those at risk populations with the mildest symptoms or even those still clinically asymptomatic into active clinical trials to potentially reverse or even prevent AD as well as other neurodegenerative disorders.

REFERENCES

Alzheimer's Disease International World Alzheimer report 2015: the global impact of dementia. 2015. Available at https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf. Accessed December 12, 2016.

*American Psychiatric Association, Rabins PV, Blacker D, Rovner BW, et al. Practice guideline for the treatment of patients with Alzheimer's disease and other dementias. Second edition. *Am J Psychiatry*. 2007;164(12):5–56.

Auchus AP, Brashear HR, Salloway S, et al. Galantamine treatment of vascular dementia: a randomized trial. *Neurology*. 2007;69(5):448–458.

Bhasin M, Rowan E, Edwards K, McKeith I. Cholinesterase inhibitors in dementia with Lewy bodies—a comparative analysis. *Int J Geriatr Psychiatry*. 2007;22(9):890–895.

*Birks J. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database Syst Rev*. 2006;1:CD005593.

Black SE, Doody R, Li H, et al. Donepezil preserves cognition and global function in patients with severe Alzheimer disease. *Neurology*. 2007;69(5):459–469.

Boinpally R, Chen L, Zukin SR, McClure N, Hofbauer RK, Periclou A. A novel once-daily fixed-dose combination of memantine extended release and donepezil for the treatment of moderate to severe Alzheimer's disease: two phase I studies in healthy volunteers. *Clin Drug Investig*. 2015;35(7):427–435.

Brodaty H, Corey-Bloom J, Potocnik FC, Truyen L, Gold M, Damaraju CR. Galantamine prolonged-release formulation in the treatment of mild to moderate Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2005;20:120–132.

Chen S, Zhang XJ, Li L, Le WD. Current experimental therapy for Alzheimer's disease. *Curr Neuropharmacol*. 2007;5(2):127–134.

Corey-Bloom J, Anand R, Veach J; ENA 713 B352 Study Group. A randomized trial evaluating the efficacy and safety of ENA 713 (rivastigmine tartrate), a new acetylcholinesterase inhibitor, in patients with mild to moderately severe Alzheimer's disease. *Int J Ger Psychopharmacol*. 1998;1:55–65.

Cummings J, Lefevre G, Small G, Appel-Dingemanse S. Pharmacokinetic rationale for rivastigmine patch. *Neurology*. 2007;69(4 Suppl 1):S10–S13.

Emre M, Aarsland D, Albanese A, et al. Rivastigmine for dementia associated with Parkinson's disease. *N Engl J Med*. 2004;351(24):2509–2518.

Farlow M, Gracon SI, Hershey LA, Lewis KW, Sadowsky CH, Dolan-Ureno J. A controlled trial of tacrine in Alzheimer's disease. The Tacrine Study Group. *JAMA*. 1992;268(18):2523–2529.

Farlow MR, Miller ML, Pejovic V. Treatment options in Alzheimer's disease: maximizing benefit, managing expectations. *Dement Geriatr Cogn Disord*. 2008;25(5):408–422.

Grossberg GT, Manes F, Allegri RF, et al. The safety, tolerability, and efficacy of once-daily memantine (28 mg): a multinational, randomized, double-blind, placebo-controlled trial in patients with moderate-to-severe Alzheimer's disease taking cholinesterase inhibitors. *CNS Drugs*. 2013;27(6):469–478.

Kavirajan H, Schneider LS. Efficacy and adverse effects of cholinesterase inhibitors and memantine in vascular dementia: a meta-analysis of randomized controlled trials. *Lancet Neurol*. 2007;6(9):782–792.

Lefèvre G, Sedek G, Jhee SS, et al. Pharmacokinetics and pharmacodynamics of the novel daily rivastigmine transdermal patch compared with twice-daily capsules in Alzheimer's disease patients. *Clin Pharmacol Ther*. 2008;83(1):106–114.

Matsunaga S, Kishi T, Iwata N. Combination therapy with cholinesterase inhibitors and memantine for Alzheimer's disease: a systematic review and meta-analysis. *Int J Neuropsychopharmacol*. 2015;18(5).

Miranda LF, Gomes KB, Tito PA, et al. Clinical response to donepezil in mild and moderate dementia: relationship to drug plasma concentration and CYP2D6 and APOE genetic polymorphisms. *J Alzheimers Dis*. 2016;55(2):539–549.

*O'Brien JT, Burns A; BAP Dementia Consensus Group. Clinical practice with anti-dementia drugs: a revised (second) consensus statement from the British Association for Psychopharmacology. *J Psychopharmacol*. 2011;25(8):997–1019.

Panek D, Wie˛ckowska A, Wichur T, et al. Design, synthesis and biological evaluation of new phthalimide and saccharin derivatives with alicyclic amines targeting cholinesterases, beta-secretase and amyloid beta aggregation. *Eur J Med Chem*. 2017;125:676–695.

Qassem A, Snow V, Cross JT Jr, et al. Current pharmacologic treatment of dementia: a clinical practice guideline from the American College of Physicians and the American Academy of Family Physicians. *Ann Intern Med*. 2008;148(5):370–378.

Raskind MA, Peskind ER, Wessel T, Yuan W. Galantamine in AD. A 6-month randomized placebo controlled trial with a 6-month extension. The Galantamine USA-1 Study Group. *Neurology*. 2000;54(12):2261–2268.

Reisberg B, Doody R, Stöffler A, Schmitt F, Ferris S, Möbius HJ; Memantine Study Group. Memantine in moderate-to-severe Alzheimer's disease. *N Engl J Med*. 2003;348(14):1333–1341.

Rogers SL, Farlow MR, Doody RS, Mohs R, Friedhoff LT. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. *Neurology*. 1998;50(1):136–145.

*Rolinski M, Fox C, Maidment I, McShane R. Cholinesterase inhibitors for dementia with Lewy bodies, Parkinson's disease dementia and cognitive impairment in Parkinson's disease. *Cochrane Database Syst Rev*. 2012;3:CD006504.

*Russ TC, Morling JR. Cholinesterase inhibitors for mild cognitive impairment. *Cochrane Database Syst Rev*. 2012;9:CD009132.

Tariot PN, Solomon PR, Morris JC, Kershaw P, Lilienfeld S, Ding C. A 5-month, randomized placebo-controlled trial of galantamine in AD. *Neurology*. 2000;54(12):2269–2276.

Winblad B, Kilander L, Eriksson S, et al. Donepezil in patients with severe Alzheimer's disease: double-blind, parallel group, placebo-controlled study. *Lancet*. 2006;367(9516):1057–1065.

Winblad B, Grossberg G, Frölich L, et al. IDEAL: a 6-month, double-blind, placebo-controlled study of the first skin patch for Alzheimer disease. *Neurology*. 2007;69(4 Suppl 1):S14–S22.

▲ 34.16 Disulfiram and Acamprosate

ALLISON K. UNGAR, M.D., KATHERINE S. PIER, M.D., AND ILIYAN IVANOV, M.D.

DISULFIRAM

History

The agent disulfiram was serendipitously discovered in 1937 for possible medical use by Dr. E. E. Williams, a physician to a rubber company. Dr. Williams observed that workers in the vulcanized rubber industry who were exposed to tetraethylthiuram disulfide became physically ill after alcohol consumption. The tetraethyl analog of the vulcanization chemical had also been used topically since 1942 to treat scabies. In 1945 while investigating this chemical to treat parasitic stomach infections, Erik Jacobsen and Jens Hald tested the compound on themselves only to become ill after drinking alcohol. They observed and reported symptoms such as nausea, vomiting, facial flushing, headache, and dizziness. Two years later, they collaborated with Dr. Oluf Martensen-Larsen and designed a study for disulfiram as a treatment for alcohol dependence. In 1951, it was the first medication approved by the FDA for that indication.

Over the years it has been recognized that the very same properties that constitute disulfiram's main therapeutic effect (i.e., the ability to produce unpleasant symptoms after alcohol intake also known as disulfiram–alcohol reaction) may also have potentially dangerous consequences. These are frequently related but not limited to dehydration and electrolyte imbalance secondary to vomiting, a symptom of the disulfiram–alcohol reaction. In the most severe cases serious clinical conditions such as respiratory depression, myocardial infarction, arrhythmias, heart failure, seizure, loss of consciousness, and death have been documented. These potential complications as well as the more recent development of alternative alcohol use disorder medications have been the limiting factors for a wider use of disulfiram.

Chemistry

Disulfiram, when it is in its pure state, is a white to off-white, odorless, almost tasteless powder, which is soluble in water and alcohol. Its chemical name is bis(diethylthiocarbamoyl) disulfide; its structural formula is shown in Figure 34.16–1.

Pharmacological Actions

Pharmacokinetics

ABSORPTION. After ingestion, disulfiram is rapidly converted to its diethyldithiocarbamate (DDC) copper complex. Subsequent absorption and distribution via the GI mucosa into the blood involves both the parent drug and the DDC copper complex. Absorption of disulfiram from the GI tract is rapid but incomplete (e.g., approximately 80 to 95 percent is absorbed through the GI tract). The concentration of disulfiram's active metabolite diethylthiocarbaminic acid methyl ester (Me-DTC) reaches its plasma peak about 4 hours after absorption. It is hypothesized that disulfiram tablets may be more bioavailable when taken with food. It can be crushed and mixed with water, coffee, tea, milk, soft drinks, or fruit juice.

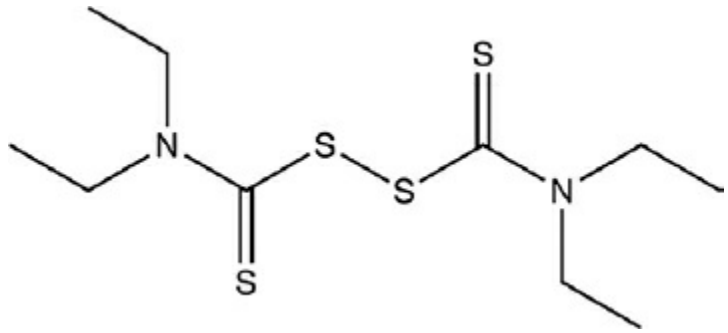

FIGURE 34.16–1. Chemical structure of disulfiram.

DISTRIBUTION. Because of its high lipid solubility, disulfiram is widely distributed and accumulates in various fat deposits. Its safety during pregnancy has not been established, and it is labeled as a class C drug. In experimental animals, administration of disulfiram with ethanol potentiates the embryotoxic effects of ethanol. It is not clear whether the same potentiation occurs in humans. There are several case reports of normal pregnancy outcome in humans exposed to disulfiram without exposure to ethanol. It is not known whether it is excreted in breast milk.

METABOLISM AND ELIMINATION. Due to partial absorption, approximately 5 to 20 percent of disulfiram is excreted in the feces. In the blood, disulfiram is rapidly metabolized to DDC, which is partly excreted as carbon disulfide in the expired air and is partly metabolized in the liver to diethyldithiomethylcarbamate (Me-DDC). Further metabolic changes result in the formation of the active metabolite Me-DTC. This metabolic cascade is not appreciably affected by a mild to moderate decrease in hepatic function. Disufiram's metabolites are chiefly excreted with the urine and a part is expired as carbon disulfide. Overall elimination of disulfiram is slow, and up to 20 percent of a single dose may remain in the body for a period of 7 days or more.

Pharmacodynamics

MECHANISM OF ACTION. Disulfiram is a prodrug. It has been shown to inhibit the hepatic enzyme aldehyde dehydrogenase (ALDH) in vitro. However, in vivo the drug is too short-lived to inhibit the enzyme directly. The net effect of irreversible inhibition of ALDH is achieved by disulfiram's active metabolite Me-DTC.

The pharmacological action of disulfiram and its metabolites is based on the inhibition of enzymes involved in ethanol catabolism. Disulfiram is a sulfhydryl (-SH, thiol) group reagent and inhibits enzymes concerned with oxidation of active (-SH group) sites on enzymes' protein molecules. Normally, ethanol is metabolized to carbon dioxide and water, but in the presence of disulfiram the metabolic chain of reactions stops after the production of acetaldehyde, leading to the accumulation of this substance. The acetaldehyde concentration in the blood may increase to 5 to 10 times than occurring without disulfiram. Acetaldehyde is an important factor for the development of disulfiram–alcohol reaction, which occurs within the course of 10 to 30 minutes after the ingestion of alcohol. The clinical picture of disulfiram–alcohol reaction includes a number of unpleasant symptoms: intense flushing of the face, a feeling of difficulty in breathing, palpitations, a throbbing headache, nausea and vomiting. Ingestion of large amounts of alcohol may cause a significant fall of the blood pressure, resulting in fainting and the risk of cardiovascular shock.

Re-establishment of the ALDH enzymatic activity is dependent on new synthesis which occurs gradually during the course of up to 2 weeks. Although the plasma half-life of disulfiram's active metabolite Me-DTC is about 10 hours, its inhibiting effect on ALDH may last considerably longer, persisting for 7 to 14 days after discontinuation. Prolonged use of disulfiram does not result in the development of tolerance; conversely, it has been shown to produce a heightened sensitivity to alcohol.

In addition to its effect on ALDH, disulfiram and some of its metabolites (e.g., DDC) also inhibit the enzyme dopamine-beta-hydroxylase (DBH), thus reducing the synthesis of noradrenaline. This mechanism is suggested to play a role in the disulfiram–alcohol reaction. Disulfiram also diffuses readily into cells and raises intracellular -SH levels; therefore, it can affect intracellular oxidation reactions. Notably, disulfiram has been shown to inhibit the enzymes xanthine oxidase and succinoxidase. Disulfiram also possesses antithyroid properties as it reacts with free iodine to form a stable molecular complex.

BLOOD CONCENTRATIONS AND RELATION TO ACTION. Plasma concentrations of disulfiram may vary between individuals due to a number of factors, most notably age and hepatic function. In general, the severity of disulfiram–alcohol reaction has been shown to be proportional to the amount of the ingested disulfiram and alcohol. However, disulfiram plasma levels are rarely utilized in clinical practice. Mild effects may occur at blood alcohol

concentrations of 5 to 10 mg per 100 mL. When alcohol levels reach 50 mg per 100 mL, symptoms will have fully developed. Unconsciousness may occur when alcohol levels reach 125 to 150 mg per 100 mL.

Design and Interpretation of Clinical Drug Studies

To date, the Veterans Administration Cooperative Study remains the largest disulfiram treatment study, conducted in a cohort of 605 males diagnosed with alcoholism. Participants were randomly assigned to three treatment arms: 202 participants received disulfiram 250 mg daily, 204 participants were administered disulfiram 1 mg QD, and 200 participants were prescribed riboflavin. After 1 year of treatment, there was no difference in the percentage of patients from the three groups in total abstinence, time to first drink, employment, or social stability. There was a significant relationship between adherence to any of the drug regimens and complete abstinence. Among those who did drink, those in the disulfiram 250 mg group reported significantly fewer drinking days. Also, in the subpopulation of patients who attended all the scheduled follow-up research appointments, treatment with disulfiram 250 mg daily significantly decreased drinking days when compared low-dose disulfiram and riboflavin. In this large trial, compliance with the study medication was a notable barrier, with only 19 percent of patients taking disulfiram consistently.

The first systematic meta-analysis of its disulfiram's effectiveness was published in 2011. Eleven randomized controlled trials were included with a total of 1,527 patients. Studies were published from 1979 to 2010 in India, Denmark, USA, Austria, Finland, and Italy. This review concluded that disulfiram had some short-term clinical benefits including short-term abstinence, number of drinking days, and days until relapse. However, possible confounds were related to the fact that five of the six efficacious studies also had supervised disulfiram administration and that all studies were short-term treatment protocols. Others have pointed to possible selection bias since the authors included one nonefficacy study while five eligible trials were absent.

Moreover, the aforementioned meta-analysis utilized both open and blind studies. This issue was addressed by a second meta-analysis that compared blind versus open-label studies of disulfiram. The analysis, which included 2,414 participants from 22 studies showed no difference between disulfiram and control groups. The authors explained this finding on the premise that the clinical effect of disulfiram is directly proportional to the expectation of a disulfiram–alcohol reaction, meaning a blinded placebo group has the same psychological threat as the treatment arm. Moreover, they point out that even in a blinded trial, the participants can unmask themselves easily with minimal alcohol. In the same meta-analysis, the results from only open-label studies indicate disulfiram is more

effective than naltrexone, acamprosate, and control group treatment conditions. Finally, the results suggested that disulfiram is significantly superior to placebo condition only with supervised administration.

There are only few long-term studies of supervised disulfiram administration and there have not been randomized controlled trials. The Outpatient Long-Term Intensive Therapy for Alcoholics (OLITA) study presents a biopsychosocial approach, including therapeutic elements like frequent participant contact, therapist rotation, support of social reintegration, aggressive aftercare, supervised use of disulfiram, and regular urine analysis for alcohol and other drugs of abuse. The group reported a main outcome of 52 percent cumulative 9-year abstinence rate.

By expert consensus, optimum disulfiram response is achieved when treatment is conducted in a specialized facility for substance use disorders. Aspects of care that can increase disulfiram efficacy by way of adherence include: use of incentives, contracting with the patient and a significant other, providing regular reminders, patient behavioral training, and social support networks. The need for adherence and supervised administration of this medication provides some framework to understand that disulfiram has documented superior efficacy over placebo in the treatment of alcohol abuse in adolescents or in combination with spouse-observed administration. This also helps conceptualize why court-ordered disulfiram therapy promotes efficacy by increasing adherence to the regimen. Additional research indicates that, in general, older med involved in a structured social network like AA, are more likely to remain compliant to treatment regimens, and thus, derive clinical benefit.

As aforementioned, spousal involvement can be advantageous in the utility of disulfiram. Several studies have documented the benefit of behavioral marital therapy as adjunct to disulfiram treatment. A pivotal component in this form of combined therapy is the disulfiram contract—a time-limited, renewable agreement by the couple to practice daily a behavioral interaction of observed disulfiram administration. It is believed that on individual level the contract aims to increase patient adherence to the medicine. On the couple's level, the goal is to restructure behaviorally the relationship dynamics of early sobriety, often characterized by negative emotions like distrust, shame, and occasionally violence. Several trials have evaluated the efficacy of this approach as the couples received the combined treatment for periods of 10 to 24 weeks and were followed as long as 104 weeks after that. Overall the combined treatment has been shown to produce greater disulfiram adherence, better drinking outcomes, higher marital satisfaction, and less domestic violence.

There are few comparisons between disulfiram and the other FDA-approved medications for the treatment of alcoholism, naltrexone, and acamprosate. Some of these trials are also potentially confounded by the nonspecific effects of frequent supervision with disulfiram but not the other

agents. One randomized, open-label trial aimed to compare the effects in alcohol-dependent patients of three pharmacotherapies, disulfiram, naltrexone, and acamprosate. It included 243 voluntary treatment-seeking alcohol-dependent adult outpatients. It had two phases: first, a 12-week continuously supervised medication phase, followed by up to a year of unsupervised medication. Participants were followed for up to 2.5 years. All treatment arms included a brief manual-based cognitive-behavioral therapy. The investigators found that all three medication interventions demonstrated a reduction in drinking, from baseline to the end of the study. Disulfiram was significantly superior to the other two agents (i.e., naltrexone and acamprosate) during the phase of supervised treatment; it demonstrated greater reduction in heavy drinking days, reduction in weekly consumption, abstinence days, and an increased time to the first drink. During the unsupervised period of the study, there were no significant differences between the groups in the heavy drinking days or time to first drink, although disulfiram displayed a significantly greater efficacy in the number of abstinence days.

Another study analyzed retrospectively collected data (i.e., from 2002 to 2007) in 353 outpatient alcohol-dependent individuals and compared the efficacy of acamprosate and disulfiram. Two groups were identified, individuals receiving three-time-a-week appointments for supervised disulfiram and once-weekly appointments for acamprosate. Both groups were individuals who had completed an inpatient detox prior to the medication intervention. They assessed abstinence by several means: breathylizer, self-report, urine and serum assays, and physician assessment. The investigators identified supervised disulfiram treatment to be more effective than the treatment with acamprosate, particularly in patients with a longer duration of alcohol dependence.

Efficacy of disulfiram has been also studied in individuals with addiction and co-occurring Axis I disorders (duel-diagnosis), including major depression, PTSD bipolar disorder, schizophrenia, and schizoaffective disorder. One study enrolled 254 patients with alcoholism and a co-occurring Axis I disorder and randomized participants into four groups: (1) naltrexone alone, (2) placebo alone, (3) disulfiram and naltrexone, and (4) disulfiram and placebo. There was open randomization to disulfiram or no disulfiram, and double-blind randomization to naltrexone or placebo. All patients were treated for 12 weeks in an outpatient setting. Participants treated with disulfiram had significantly more consecutive weeks of abstinence and less craving than those treated with placebo, but there were no significant group differences in other measures of alcohol consumption (percent of days abstinent, percent of heavy drinking days, achieving total abstinence). There was no advantage of the combination of both medications. Of note, individuals with psychotic spectrum disorders did not have a change in psychotic symptoms with the

intervention groups.

Disulfiram may be more effective in promoting short-term abstinence and treatment retention after detoxification, rather than long-term relapse prevention. As such, some have suggested that this medication be solely used to assist in the initial period of abstinence to facilitate the client's involvement in psychosocial interventions. Further, some experts dismiss disulfiram as a viable treatment option, considering the severe adverse effects from the disulfiram–alcohol reaction and also the potentially serious side effects from the drug itself. Others have questioned the effectiveness of disulfiram because the time between alcohol ingestion and the disulfiram–alcohol reaction can be up to 30 minutes and the reaction intensity can be variable. Patients with severely impaired judgment, significant cognitive impairment, or those who are highly impulsive may be inappropriate candidates for this treatment.

Effects on Specific Organs and Systems

Deaths from disulfiram treatment have become rare due to treatment with lower dosages and exclusion of patients with significant cardiac disease. Disulfiram has a black box warning to never administer to a patient who is in a state of alcohol intoxication or without the patient's full knowledge.

Adverse effects of disulfiram have been mostly documented on the gastrointestinal and nervous systems. Of particular clinical importance is the hepatotoxic effect of disulfiram. The symptoms of hepatotoxicity may be mild and reversible (e.g., jaundice, cholestasis, and altered liver function tests) or more serious such as acute hepatitis, hepatic necrosis, and hepatic failure that may require liver transplantation. These effects may occur after many months of disulfiram treatment and have been documented in both individuals with and without prior history of abnormal liver function. Estimates of disulfiram-induced hepatitis are between 1 in 25,000 and 1 in 30,000 patients treated per year.

Disulfiram effects on the nervous system may involve optic neuritis, resulting in eye pain or tenderness and changes in vision, as well as peripheral neuropathy (i.e., parasthesias—numbness, tingling, pain, or weakness in hands or feet). The latter is mainly reported in patients who have been maintained on relatively high dose of disulfiram (e.g., 500 mg daily). Peripheral neuropathy tends to improve with discontinuation of disulfiram treatment. More rare side effects include the antithyroid effect of disulfiram and skin eruptions that are usually controlled with antihistamine treatment.

Therapeutic Indications

Alcohol Use Disorder. In clinical practice disulfiram is recommended as an aid in the management of selected patients with alcohol use disorder.

It is emphasized that the selected patients should be highly motivated to maintain a state of sobriety. The agent has little effect on the long-term drinking pattern of the affected individuals. It has been found to be most appropriate and efficacious in the short-term after detoxification, to assist in the initial period of abstinence to facilitate the client's involvement in psychosocial interventions including psychotherapy. Optimum response to disulfiram seems to be achieved in facilities specializing in the treatment of substance use disorders.

The disulfiram–ethanol reaction is an aversive physiological state that functions through negative reinforcement and behavioral counterconditioning to extinguish an unwanted behavior (alcohol consumption). When an individual takes this medication, they acknowledge that possible alcohol consumption will entail such reaction and this knowledge is the motivation for abstinence. It is important that patients should be educated about the medication and informed consent should be obtained. Baseline blood tests include blood alcohol or breathalyzer (to confirm abstinence), liver function tests, a complete blood count, serum chemistries including blood urea nitrogen (BUN) and creatinine, and a pregnancy test. Follow-up liver function tests are recommended 10 to 14 days after therapy initiation, then monthly for a period of 6 months, to monitor for possible hepatic dysfunction. After the first 6 months, liver function testing should be obtained every 3 months. Monthly pregnancy tests are indicated for women of childbearing age. An ECG should be obtained at baseline if clinically indicated by a history of heart disease.

Disulfiram should be prescribed for individuals who are abstinent from alcohol for at least 12 hours. On occasions, patients who appear to be on adequate maintenance dose of disulfiram may report that they can consume alcohol without any negative reactions. In such cases the physician should presume that patient's compliance with the treatment has been compromised. Before the patient is observed to reliably administer the medicine (preferably crushed and mixed with water) it may not be concluded that disulfiram is ineffective.

Disulfiram Test Reaction. During the earlier use of disulfiram it was thought advisable to have patients experience at least one supervised disulfiram–alcohol reaction, a so-called "alcohol challenge." This practice is currently not recommended; moreover, it should never be considered for patients over the age of 50 years. At present it is believed that a clear, convincing, and detailed description of the disulfiram–alcohol reaction is considered sufficient.

Management of Disulfiram–Alcohol Reaction. The disulfiram–alcohol reaction is diagnosed when a patient with a history of disulfiram use and

possible exposure to ethanol presents with symptoms of the characteristic hypotensive flushing reaction. Blood disulfiram levels are not of value in diagnosis or treatment. Blood acetaldehyde levels may be elevated during the disulfiram–ethanol reaction, but this information is of little value in acute management. In a severe reaction, all measures should be taken to stabilize the patient's respiratory and cardiovascular systems, this may include supplemental oxygen, intubation, IV fluids, and/or pressor agents. Other interventions include intravenous use of vitamin C, ephedrine sulfate, and antihistamines. Since hypokalemia has been reported, potassium levels should be monitored, particularly in patients treated with digitalis. Other useful laboratory studies include electrolytes, glucose, BUN, creatinine, liver aminotransferases, and ethanol level. Some protocols include treating vomiting with metoclopramide. There is no specific antidote. Decontamination procedures are not likely to be of benefit once the reaction begins. Hemodialysis is not indicated for disulfiram overdose.

Management of Disulfiram Overdose (without Alcohol). Disulfiram overdose (without alcohol ingestion) is rare and its clinical presentation is one of neurological adverse effects from the medication: confusion, lethargy, headache, ataxia, seizures, and coma. Brain lesions in the basal ganglia have also been reported. A patient may also present with a garlic-like breath odor, vomiting, and hypotension. The diagnosis of overdose is made based on history and presenting symptoms. In some circumstances, activated charcoal may be clinically indicated. The treatment is focused on maintaining a patent airway, stabilizing the cardiovascular system, and treating seizures.

Other Disorders. Disulfiram is hypothesized to be potentially effective as a treatment for cocaine use disorder, as it inhibits dopamine-beta-hydroxylase resulting in an excess of dopamine and decreased synthesis of norepinephrine. Cocaine, on the other hand, increases dopamine through inhibition of the dopamine transporter, blocking reuptake. There is evidence that disulfiram could be efficacious for cocaine dependence alone and also for comorbid cocaine and alcohol dependence. Usually interventions were paired with some form of psychotherapy or 12-step program. A recent Cochrane review included seven controlled studies with 492 participants receiving disulfiram, a placebo, no pharmacological treatment, or naltrexone in addition to psychosocial treatment. The review showed "low evidence" in support of disulfiram for the treatment of cocaine dependence. Current investigations are examining safety profiles for the combination of disulfiram with alcohol and cocaine.

Some concern has been raised about the combination of the cardiovascular response to cocaine and the cardiovascular effects of a disulfiram–alcohol reaction. Safety trial data from few studies with small

sample sizes suggest that the severity of physiological effects is related to the doses of any of the three agents (e.g., disulfiram, alcohol, cocaine); however, cocaine does not appear to exacerbate the disulfiram–alcohol reaction. Disulfiram does inhibit the metabolism of cocaine, increasing cocaine plasma levels.

In nonalcohol-dependent individuals, several studies have demonstrated a benefit in reduction of cocaine use with comorbid opiate-dependent individuals receiving opiate agonist maintenance treatment (methadone and buprenorphine). However, one double-blind, randomized, placebo-controlled clinical trial in a population of methadone stabilized, cocaine-dependent individuals, reported that there may be a disulfiram dose-related benefit. The study found cocaine-positive urines increased in the 62.5- and 125-mg disulfiram groups. Disulfiram at 250 mg was similar to placebo. The authors suggest using higher doses of disulfiram (greater than 250 mg) in individuals with comorbid alcohol, cocaine, and opioid use disorders, which requires further investigation.

Precautions and Adverse Reactions

Precautions. All patients who are prescribed disulfiram should be fully informed of the disulfiram–ethanol reaction and cautioned against the possible consequences of taking alcohol either surreptitiously or unwittingly. Patients receiving disulfiram should be strongly advised not to use any alcohol-containing preparations, for example, certain cough syrups, sauces, vinegar, tonics, foods prepared with wine, and even should avoid the use of aftershave lotions and moisturizers containing alcohol. In addition, patients should be advised to obtain and carry an Identification Card describing the symptoms of disulfiram–alcohol reaction and containing contact information in cases of emergency. Disulfiram should not be coadministered with paraldehyde or metronidazole.

Candidates for disulfiram should be evaluated for hypersensitivity to other thiuram derivatives used in pesticides and rubber vulcanization. Concomitant conditions such as severe myocardial disease or ischemic heart disease, pregnancy, advanced hepatic and renal disease, psychotic states, and serious organic brain damage are contraindications for the use of disulfiram.

Adverse Reactions. Most common side effects associated with disulfiram include drowsiness, headache, fatigue, allergic dermatitis, stomach upset, dizziness, acneform eruption, halitosis, and impotence. Metallic or garlic-like aftertaste has been reported during the first 2 weeks of therapy. These complaints disappear spontaneously or by reducing dosage.

Sensorimotor peripheral neuropathy, frequently associated with combined treatment with metronidazole or isoniazid, may present with

numbness, tingling, pain, or weakness in hands or feet. In addition, symptoms of optic neuritis may occur, indicated by eye pain and changes in vision. Most of these symptoms improve when disulfiram is discontinued.

More serious but rare side effects include disulfiram-induced hepatitis with an incidence between 1 in 25,000 and 1 in 30,000 patient years of treatment. Varying degrees of hepatotoxicity have been reported including altered liver function tests, jaundice, and hepatic necrosis. Psychiatric side effects like mood liability and mental status changes (e.g., psychosis, mania, delirium) have also been reported with the administration of disulfiram.

Drug Interactions

Concomitant ingestion of antacids containing divalent cations as well as large doses of ferrous salts may reduce absorption of disulfiram. Other clinically salient drug–drug interactions of disulfiram may occur through several separate physiological mechanisms (i.e., inhibition of hepatic enzymes). For instance, disulfiram may affect the metabolic rate of other drugs, in particular phenytoin and theophylline. The concomitant use of disulfiram with either of these two agents should be used with caution since it may result in an elevated phenytoin or theophylline plasma level and possibly lead to phenytoin or theophylline toxicity. It is recommended that baseline phenytoin and theophylline levels be obtained before initiation of disulfiram with subsequent regular monitoring of phenytoin and theophylline levels.

Disulfiram may prolong prothrombin time when coadministered oral anticoagulants (e.g., warfarin), as disulfiram inhibits warfarin's metabolism and regular assessments of anticoagulation profiles are recommended.

Another clinical interaction with less clearly understood mechanism occurs between isoniazid and disulfiram. Their coadministration may cause ataxia and even mental status changes, both of which usually resolve after discontinuation of disulfiram. Another antituberculosis agent, rifampin, also interacts with disulfiram, resulting in inhibition of rifampin's oxidation and renal excretion. Mental status changes (e.g., psychosis and confusion) have also been demonstrated with concurrent use of metronidazole.

Disulfiram can differentially potentiate and prolong benzodiazepines' suppressing effect on the CNS—for instance this effect is more pronounced with coadministration of chlordiazepoxide and diazepam, but not oxazepam or lorazepam.

Animal research has documented that in the GI tract disulfiram may interact with nitrates and form nitrosamine in the GI tract, which has been shown to cause stomach tumors. The toxicity of certain centrally acting drugs (e.g., morphine, pethidine, amphetamine, and barbiturates) has been potentiated by disulfiram in rats; however, the importance of this

observation for humans is unclear.

Laboratory Interference

Disulfiram is not known to interfere with any laboratory tests.

Dosage and Administration

Dosing for the first 1 to 2 weeks is 250 mg in the morning or the evening (based on whether the patient has sedation). The average maintenance dose is 250 mg daily, with a range of 125 to 500 mg. The dose of disulfiram should never exceed 500 mg per day.

Disulfiram should be taken for 6 weeks to 6 months as required. A review of the effectiveness of therapy should be undertaken before continuation on a longer-term basis. A few patients would require higher doses of disulfiram and longer duration of treatment under close supervision.

Administration. Disulfiram should be taken preferably on waking, although in patients who experience a sedative effect, it may be taken at bedtime. Alternatively, to minimize the sedative effect, the initial dosage may be reduced.

ACAMPROSATE

History

Acamprosate was developed in response to a growing literature demonstrating that chronic alcohol consumption and withdrawal fundamentally change neurochemistry and electrophysiology. Consequently, a major therapeutic challenge in the treatment of alcohol use disorders is the maintenance of abstinence once it has been achieved and regaining sobriety after relapse. Over 30 years ago several European pharmaceutical companies started synthesizing amino acids neuromediators, with the hope of ameliorating the pathological hyperglutamatergic state that ensues from protracted alcohol use, thought to perpetuate alcohol use and contribute to relapse after abstinence. In collaboration with several universities in France, pharmaceutical companies investigated a number of taurine and homotaurine compounds in the 1980s. An animal model of alcohol relapse conducted at the University of Rouen in France supported the utility of calcium acetyl-homotaurinate in reducing alcohol consumption in dependent rats. This initial preclinical trial was again validated in another animal model, which revealed significant reduction in voluntary alcohol consumption.

These preclinical results catalyzed the first multicenter double-blind, placebo-controlled study in 569 detoxified alcohol-dependent male patients in Europe, again showing decreased alcohol consumption in patients

receiving calcium acetyl-homotaurinate compared to placebo. The compound was approved for the treatment of alcohol dependence in 1989 in France under the trade name Aotal. Extensive clinical studies involving over 4,000 subjects throughout Europe demonstrated the safety and efficacy of acamprosate across clinical settings. By 1996, the medication became available in all European Union countries for the treatment of alcohol use disorder, promulgating additional investigations into its utility in the pharmacological treatment of alcoholism.

In July 2004, the FDA approved acamprosate for the maintenance of abstinence from alcohol in detoxified patients with alcohol use disorder. It became available for use in January 2005 under the trade name Campral. Since its introduction in 1989, acamprosate has been used to treat over 1.5 million patients worldwide. It is now the most widely prescribed agent for the treatment of alcohol use disorders in the United States.

Chemistry

Acamprosate is the calcium salt of N-acetyl-homotaurine (Ca-AOTA), a synthetic analog of the endogenous amino acid taurine with additional carboxyl and methyl groups in its chain. Acamprosate calcium is a white, nearly odorless powder, which is readily soluble in water and practically insoluble in absolute ethanol. Its chemical name is calcium 3-acetylaminopropane-1-sulfonate and its structure is shown on Figure 34.16–2.

Pharmacological Actions

Pharmacokinetics

ABSORPTION. Acamprosate is a dimer of acetyl-homotaurine linked by a calcium salt, formulated to assist in its absorption across the GI tract. The amount of acamprosate actually absorbed is quite low, however, with a bioavailability of ~11 percent in the fasting state. Food consumption reduces acamprosate absorption but not sufficiently to justify dose adjustments. Acamprosate is so slowly absorbed that it acts as if it were a modified-release medication. Peak plasma concentrations are usually reached within 3 to 8 hours after dosing and steady-state levels are achieved by the fifth to seventh day of dosing. Its terminal half-life ranges from 20 to 33 hours subsequent to the administration of a standard 666 mg enteric-coated formulation.

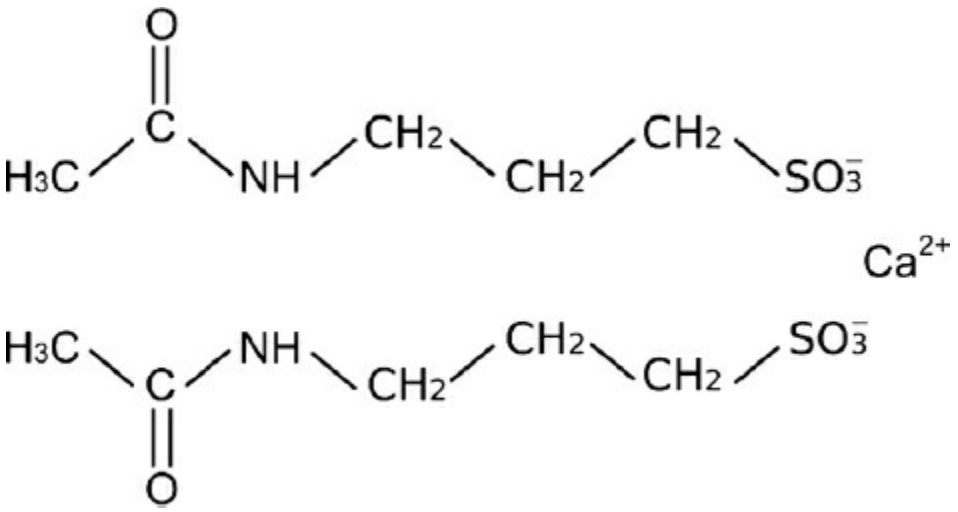

FIGURE 34.16–2. Chemical structure of acamprosate.

DISTRIBUTION. The amount of acamprosate which is protein bound is negligible, such that the active drug that permeates the intestinal mucosa is entirely dissociated in plasma. While its volume of distribution following intravenous acamprosate administration is high, estimated to be 72 to 109 L (approximately 1 L/kg), its oral bioavailability is relatively low. Only 6.5 to 11 percent of what is ingested crosses the GI tract, which explains why it must be given in high doses. In contrast to orally administered acamprosate, its concentrations are quite high when administered intravenously, but its half-life is approximately 3 hours as compared to 20 to 33 hours when taken orally. This is an example of "flip-flop" pharmacokinetics in which the concentration and half-life are limited by the rate and extent of absorption rather than volume of distribution or excretion.

Acamprosate has been assigned a category C labeling in pregnancy. Teratogenicity has been demonstrated in rat and rabbit fetuses though there have been no large-scale studies in humans to date. It is also unknown whether the medication is excreted through breast milk.

METABOLISM AND ELIMINATION. Importantly, acamprosate bypasses the liver and is eliminated unchanged in the urine and the feces. The majority of the acamprosate does not permeate the intestinal mucosa, and is therefore removed through the feces. Once the medication is available via paracellular transport, renal clearance is expedient, occurring at approximately 6 mL/min. The portion of the drug that is absorbed (ranging from 6.5 to 11 percent) does not undergo metabolism and is excreted in the urine as acamprosate. There is a linear relationship between renal clearance and plasma half-life of acamprosate and creatinine clearance values when taken orally. No significant pharmacokinetic differences between male and female subjects have been found for acamprosate.

Pharmacodynamics

MECHANISM OF ACTION. Acamprosate bears structural similarities to several amino acids, including GABA, glycine, aspartate, and glutamate; its mechanism of action is therefore thought to be mediated, at least in part, through numerous neurotransmitter systems.

Initial studies focused on acamprosate's in vitro affinity for $GABA_A$ and $GABA_B$ receptors, which appears to decrease electrical excitability without changing the membrane potential of the neuron. It is postulated that the therapeutic effect of acamprosate is partially due to its potentiation of the GABA network, the main neuronal inhibitory system. Unlike other GABA agonists, however, acamprosate doesn't exert anxiolytic, hypnotic, or muscle relaxant properties.

Other research has shifted focus to the structural similarities between acamprosate and the excitatory neurotransmitter, glutamate. It has been proposed that acamprosate may decrease the excitatory effects of glutamate, especially during the period after abrupt discontinuation of alcohol. Since alcohol inhibits the activity of the NMDA receptors, chronic alcohol consumption leads to the overexpression of these receptors. Chronic alcohol intake also affects the release of glutamate—in order to compensate for the alcohol-induced inhibition at the NMDA receptors, the amount of glutamate released into the synaptic cleft increases over time. Thus, abrupt discontinuation of alcohol results in unopposed neuronal excitation due to supranormal levels of glutamate acting on upregulated NMDA receptors, which may contribute to the pathophysiology of the alcohol withdrawal syndrome, including delirium tremens and seizures.

Other research supports the hypothesis that acamprosate interacts with NMDA receptors in a more complex way; for example, studies show that the drug is an NMDA partial coagonist. This means that at lower concentrations, acamprosate activates NMDA receptors when receptor activity is low. Conversely, at higher concentrations and when NMDA receptor activity is high, acamprosate is inhibitory, thereby helping to restore equilibrium in this disrupted neuronal system. Furthermore, it has been shown that acamprosate induces the release of taurine, an inhibitory neuromodulator, which also decreases neuroexcitability. In early alcohol withdrawal, the drug also decreases dopamine hyperexcitability in the nucleus accumbens, as well as the expression of brain c-fos, a gene expressed during early alcohol withdrawal.

While it was previously thought that N-acetyl-homotaurine was the biologically active compound mediating acamprosate's effects, recent literature has demonstrated that calcium may be the moiety responsible for acamprosate's success in the treatment of alcohol use disorder. Studies have demonstrated that treatment with calcium-acamprosate decreases voluntary drinking, while sodium-acamprosate was not superior to placebo

in attenuating alcohol consumption after prolonged alcohol exposure. Further research showed no difference in the bioavailability of sodium and calcium-acamprosate, such that this finding cannot be explained by differential absorption. Other studies reveal that higher brain and plasma calcium levels correlate with acamprosate's efficacy in treating alcohol use disorder. Interestingly, studies on the use of calcium supplementation were conducted though never completed in the 1950s; initial findings of "intensive calcium therapy" resulted in fewer days hospitalized and better success during the transition of psychotherapy to sustain abstinence.

It is clinically important to emphasize that despite acamprosate's modulation of glutamate dysregulation in rodents, there is no clinical improvement in patients who begin acamprosate during acute detoxification. Acamprosate does not eliminate or diminish withdrawal symptoms and is not indicated for the treatment of delirium tremens. On the contrary, available research suggests that patients who receive acamprosate during the detoxification period drink statistically more alcohol in 10-week follow-up than those administered placebo.

Acamprosate is most beneficial in maintaining abstinence and helping patients regain abstinence after relapse. The medication is thought to alleviate the negative affect associated with early stages of alcohol withdrawal, reverse alcohol-related changes in sleep architecture and also to decrease alcohol cravings once abstinence is achieved. It additionally decreases cravings in response to conditioned environmental cues. Unlike disulfiram, which relies on decreasing alcohol use by inducing distinctly unpleasant effects or naltrexone, which decreases the pleasurable effects of alcohol, acamprosate's primary mechanism appears to lie in its neuromodulatory effect—its ability to normalize disturbed brain chemistry and electrophysiology during acute and protracted withdrawal. Some studies suggest that it may reduce the rewarding effects of alcohol.

The medication has not been associated with the development of tolerance or dependence following prolonged exposure and evidence suggests that it should be continued in the event that relapse occurs. Acamprosate does not cause alcohol aversion or produce a disulfiram-type reaction in the event of relapse.

Preclinical research provides evidence that acamprosate has neuroprotective properties in neurons exposed to alcohol. The hyperglutamatergic imbalance that takes place during alcohol withdrawal is associated with neuronal excitotoxicity, which acamprosate has been shown to mitigate. Results from cell culture and animal studies have demonstrated decreased cell death in alcohol-exposed neurons in the cerebral cortex and hippocampus. Research supports that excessive glutamate activity continues after acute withdrawal, inducing ischemia, as measured by lactate dehydrogenase levels. Furthermore, acamprosate has been shown to inhibit *trans*-ACPD, a metabotropic glutamate type receptor

(mGluR) agonist, responsible for some of the neurotoxicity that glutamate exerts. The hyperglutamatergic syndrome that occurs during alcohol withdrawal is also thought to be mediated by calcium entry into neurons, which acamprosate alleviates. The relevance of these finding for treatment in humans awaits further research.

BLOOD CONCENTRATIONS AND RELATION TO ACTION. Plasma concentrations of acamprosate may significantly vary between individuals due to multiple factors, most notably individual rates of absorption. Pharmacokinetic studies indicate that the concomitant use of acamprosate and alcohol doesn't result in any noticeable changes in the pharmacokinetics of either substance. At least one double-blind, placebo-controlled randomized controlled trial in a sample of 56 patients seeking treatment for alcohol use disorder found a negative correlation between alcohol craving following a priming drink (i.e., first drink after 21 days of continuous acamprosate treatment) and acamprosate plasma levels. This suggests that higher acamprosate blood concentrations may increase the chances of regaining abstinence after mild relapse.

Design and Interpretation of Clinical Drug Studies

Since its inception, there have been more than 20 double-blind, placebo-controlled protocols investigating the utility and safety of acamprosate in the pharmacological treatment of alcohol use disorders. The length of any individual trial ranged from 12 to 52 weeks with some protocols offering additional follow-ups. With few exceptions, patients were detoxified prior to treatment, dosing was based on weight and ranged between four and six 333-mg tablets. Treatment took place in outpatient programs with concurrent psychosocial interventions. The subjects across trials were predominantly male and mean age fell in the fourth decade of life.

Cumulative sample size from a meta-analysis of eight positive randomized controlled trials was 2,430, sufficiently powered to support some clinically relevant conclusions. Six of these protocols followed patients for over 52 weeks, providing a relatively long duration to study adverse effects over time. Acamprosate treatment separated from placebo most consistently when acamprosate was initiated after the acute withdrawal phase and in patients receiving concurrent psychosocial interventions. Acamprosate outperformed placebo in terms of continuous abstinence duration (time to first drink) as well as cumulative abstinence proportion (number of days abstinent out of the total days patients were in treatment) measured at 6 months. These results support the notion that acamprosate may help patients achieve complete abstinence and regain abstinence if relapse occurs. Another meta-analysis assessed acamprosate in 27 studies across 7,519 subjects. Sample sizes ranged from 21 to 1,383, treatment durations were 12 to 52 weeks and mean age was in the 40s. The

participants met *Diagnostic and Statistical Manual of Mental Disorders* (DSM) criteria for alcohol dependence and the majority of studies included concurrent psychosocial interventions. This meta-analysis found that the number needed to treat to prevent return to any drinking for acamprosate was 12, suggesting modest efficacy of acamprosate treatment.

In protocols that included patients who had not undergone detoxification or were acutely withdrawing from alcohol, acamprosate failed to demonstrate superiority over placebo. Further, in another study where patients initiated treatment 25 days after recovery from physical withdrawal symptoms, acamprosate was not more effective than placebo. One meta-analysis was designed to study differential outcomes as a consequence of duration of detoxification and abstinence prior to initiating treatment. Nineteen trials that used acamprosate alone and in combination with various other interventions (naltrexone with and without adjunctive psychotherapies) were analyzed. In the 5,559 subjects included in this part of the meta-analysis, detoxification prior to treatment was associated with larger medication effects. These data indicate that the outcome of acamprosate treatment is time sensitive and may depend on the timing of intervention in relation to acute detoxification.

Several studies have attempted to evaluate and compare the usefulness of different psychotherapies (e.g., dynamically oriented individual psychotherapy, cognitive-behavioral individual coping skills therapy, group therapy, or various brief interventions) in the treatment of abstinence from alcohol when used in combination with acamprosate. To date, it remains unclear whether any particular psychotherapy is superior to another. The literature does substantiate that acamprosate is more effective when administered concurrently with psychotherapy. Brief motivational enhancement therapy (MET) is readily learned and can be implemented in psychiatric and primary care settings alike. This intervention focuses on improving a person's motivation to achieve and sustain abstinence. Given that acamprosate has been shown to be more effective when abstinence is the goal of treatment, MET holds promise as a psychotherapeutic adjunct to acamprosate.

Additional studies have investigated the efficacy of acamprosate in comparison, or in conjunction with, naltrexone as compared to acamprosate alone, acamprosate and psychosocial interventions, and placebo. These studies have not shown any efficacy of acamprosate over placebo with or without psychotherapy. Results from a 16-week randomized controlled trial, including 1,383 recently alcohol-abstinent volunteers, demonstrated no evidence of efficacy of acamprosate alone or together with combined behavioral intervention (CBI) when compared to other treatments (e.g., naltrexone, CBI and placebo alone or in various combinations). The largest limitation of this study was that subjects were only given 4 days for detoxification in an outpatient program. Studies that

reported no additional benefit from the combined acamprosate and psychotherapy treatment had additional limitations including the following: (1) lack of randomization in one open-label naturalistic study of 753 alcohol-dependent patients and (2) presenting the participants with the option to be randomized to the no psychotherapy arm, which may have created a selection bias for participants with a particular interest in receiving psychotherapy as documented in one randomized, multicenter study of 248 patients.

Another protocol studied 426 subjects meeting DSM-IV criteria for alcohol dependence when the diagnosis was established using the structural clinical interview for DSM-IV and *International Statistical Classification of Diseases*, Tenth Revision (ICD-10). All subjects were recruited from inpatient wards for medical detoxification. Following a mean of 18 days' inpatient detoxification, eligible patients were treated for 12 weeks in the outpatient setting with acamprosate, naltrexone, acamprosate and naltrexone, or placebo in conjunction with brief psychotherapy. Exclusion criteria included dependence or abuse of another substance other than cannabis or nicotine, a comorbid psychiatric disorder for which subjects were receiving medication and unstable medical conditions. The study showed no effect on abstinence in those who received treatment as compared to those who received placebo. Limitations were related to higher acuity since participants were recruited from an inpatient setting, in addition to relatively long detoxification period followed by relatively short outpatient treatment duration. Furthermore, these data likely are not generalizable as patients with alcohol dependence often have comorbid psychiatric and medical conditions. Out of all the patients screened, for example, only 43 percent met inclusion criteria and only 31 percent of all eligible patients chose to participate, which is suggestive that selection bias may have influenced the results.

One study investigated the efficacy of acamprosate in respect to gender using gender-specific meta-analysis of individual patient data obtained from double-blind, randomized controlled trials with inpatients carrying diagnoses of alcohol dependence based on DSM-III or IV criteria. Records were obtained from 1,317 women and 4,793 men from 22 studies across 18 countries. Exclusion criteria were pregnancy, lactation, or refusal to use birth control in women of childbearing age and potential, significant comorbid psychiatric disorders for which patients were prescribed medication, medical disorders such and renal failure, and concurrent use of another pharmacological intervention for alcohol use. Fourteen trials had durations of at least 6 months and five lasted longer than 1 year. Patients needed to have at least one quantitative measure of drinking behavior to assess treatment efficacy. The four end points variables included percentage of abstinent days, percentage of no heavy drinking days, rates of complete abstinence and rates of no heavy drinking over the study's

duration. There was an incremental gain of 10.4 percent in percentage abstinent days, a gain of 11 percent in percentage of no heavy drinking days, an odds ratio of 1.9 for complete abstinence rates and an odds ratio of 1.9 for rate of no heavy drinking over the studies' durations. Importantly, outcomes in acamprosate efficacy, safety or tolerability did not differ between men and women.

These data are limited in that the randomized controlled trial with the highest power failed to demonstrate the medication's efficacy as compared to placebo or naltrexone and that many studies supporting acamprosate's efficacy are short. Taken cumulatively, however, acamprosate appears to be safe and effective adjunct to psychosocial alcohol rehabilitation programs. It is most useful for patients receiving concomitant psychotherapy and whose stated goal upon initiating treatment is abstinence.

Effects on Specific Organs and Systems

Most serious adverse effect of acamprosate may be related to the urogenital system since there have been several documented cases of patients developing acute renal failure while receiving acamprosate treatment. However, no causal relationship has been established between acamprosate and any specific physiological mechanism that may result in acute kidney impairment. Less severe but more frequent are acamprosate's effects on the cardiovascular (palpitation and syncope), respiratory (increased cough, rhinitis, dyspnea), digestive (diarrhea, nausea, vomiting, and abdominal pain), musculoskeletal (myalgias and arthralgias), nervous (somnolence, amnesia, tremor) and dermatological (pruritus) systems. Mild diarrhea or loose stool was the only statistically significant adverse event documented in a large meta-analysis. It is further estimated that up to 10 percent of patients may experience mild diarrhea; importantly, those who suffered from diarrhea were no more likely than their counterparts to discontinue the medication.

Therapeutic Indications

Acamprosate is an FDA-approved agent for maintenance of abstinence from alcohol in patients with alcohol dependence who are abstinent at treatment initiation. The efficacy of acamprosate in promoting abstinence has not been demonstrated in subjects who have not undergone detoxification and not achieved alcohol abstinence prior to beginning acamprosate treatment. The efficacy of acamprosate in promoting abstinence from alcohol in patients with polysubstance abuse has not been adequately assessed. Treatment with acamprosate should be a part of a comprehensive management program that includes psychosocial support.

Other Disorders. Acamprosate is only FDA approved for the treatment

of alcohol dependence. The literature on other uses for acamprosate over the last few years is expanding. The medication appears to hold promise in the treatment of fragile-X-syndrome (FXS)-associated autistic spectrum disorders (ASD) and tinnitus.

Excessive glutamatergic and deficient GABAergic neurotransmission is found in patients with both FXS and with idiopathic ASD. Extensive literature supports a correlation between excessive mGluR5 and deficient $GABA_A$ neurotransmission in FXS-associated ASD. Preliminary studies showed that three adults with FXS and ASD, who were treated with a mean dose of 1,221 mg/day of acamprosate for a mean duration of 21.3 achieved clinical and statistical improvement in communication. A subsequent 10-week prospective open-label trial of acamprosate in 12 patients (mean age 11.9 years old) with comorbid FXS and ASD showed improvement in social impairment and inattention/hyperactivity in 75 percent of the patients.

A derivative of amyloid-β precursor protein (APP), secreted APPα (sAPPα), has neuroproliferative effects and has been shown to be elevated in the plasma of persons with ASD compared to control subjects. A pilot study of 15 subjects treated with acamprosate for 10 weeks showed a significant decrease in sAPP-total and sAPPα in subjects with FXS-associated or idiopathic ASD ($p < 0.05$), proposing a mechanism by which acamprosate might treat these disorders on the molecular level. The largest limitation of this study was its very low power and that 13 of the patients enrolled were receiving other pharmacological therapies. There were no adverse effects reported in patients receiving acamprosate. The medication may benefit patients with FXS and ASD, though no high-powered randomized controlled trials have been conducted to date.

A prospective, randomized, placebo-controlled, double-blind, cross-over study investigated the use of acamprosate in treating 40 subjects with sensorineural tinnitus. The patients were randomized to receive three 333-mg tablets of acamprosate daily or placebo for 45 days, followed by a washout period of 7 days and then crossed over to acamprosate or placebo for another 45 days. 92.5 percent of subjects receiving acamprosate demonstrated objective and subjective improvement in tinnitus as opposed to 12.5 percent subjective improvement seen with placebo. Objective improvement of tinnitus was measured by tinnitus matching for loudness by decibel. On this objective measure of tinnitus, 80 percent of the treatment group showed improvement as compared to 0.0 percent of placebo group. Numerous case reports also reflect subjective improvement in tinnitus with acamprosate. Larger studies are needed for FDA approval, but evidence suggests some efficacy of acamprosate in the treatment of tinnitus.

There is insufficient data to determine acamprosate's utility in the treatment of other substance use disorders. This may be because the medication is relatively new and small pilot studies have failed to

demonstrate superiority to placebo in the treatment of stimulant (cocaine) use disorder while preliminary studies in rodents were promising. This topic awaits further research.

Precautions and Adverse Reactions

Precautions. The dose of acamprosate will need to be adjusted for patients with moderate renal impairment (i.e., creatinine clearance of 30 to 50 mL/min). Patients with severe renal impairment (creatinine clearance of ≤30 mL/min) should not be prescribed acamprosate. Because acamprosate does not rely on hepatic metabolism, it can be used in patients with mild to moderate hepatic impairment (Child-Pugh groups A and B). It has not been used in Child-Pugh C liver disease and is not approved for use in this population.

In placebo-controlled trials, adverse events such as suicidal ideation, suicide attempts, and even completed suicides were reported more frequently in acamprosate-treated patients than in patients who received placebo (1.4 percent vs. 0.5 percent in studies of 6 months or less; 2.4 percent vs. 0.8 percent in year-long studies). Completed suicides in the pooled acamprosate group were 0.13 percent (3 out of 2,272 patients) and 0.10 percent (2 out of 1,962) in the placebo group. Depression was reported at similar rates in both treatment groups. The interrelationship between alcohol dependence, depression, and suicidality is well recognized and complex, and since many of these adverse events occurred in the periods of addiction relapse, it was hard to determine any direct link between the clinical course of recovery from alcoholism and the emergence of suicidality. Patients with alcohol use disorders, regardless of whether or not they are treated with acamprosate, should be monitored for the development of symptoms of depression and suicidality since suicide attempts and completions occur more frequently in patients with alcohol use disorders.

Although the safety of acamprosate during pregnancy has not been determined, the agent has been shown to have a teratogenic effect on the fetus in rodents. Acamprosate is a category C drug during pregnancy. The benefit of acamprosate use during pregnancy should be carefully weighed against possible adverse effects. It is also not recommended for use in nursing mothers.

Adverse Reactions. More frequent and less serious side effects may include headaches, insomnia, stomach upset with nausea, vomiting, and abdominal pain, as well as sexual dysfunctions (e.g., decreased libido, frigidity, and impotence). In clinical trials the most commonly reported side effects were diarrhea and pruritus, though these adverse effects were usually transient and in recent meta-analyses, patients with these adverse effects are no more likely than those receiving placebo to discontinue the

medication. Those who developed anxiety and depression while taking acamprosate were most likely to discontinue the medication, but it is difficult to parse out whether these mood symptoms are directly attributable to acamprosate or to underlying untreated mood disorders or mood disorders that developed as a result of prolonged alcohol use. More serious side effects include allergic reaction to the drug, arrhythmias, and fluctuations of blood pressure.

Drug Interactions

Because acamprosate does not rely on hepatic metabolism, it has relatively few drug–drug interactions. Coadministration of benzodiazepines, disulfiram, antidepressants or alcohol do not change acamprosate's pharmacokinetic properties, which is fortunate and these medications are commonly prescribed in patients with alcohol use disorders. However, coadministration of acamprosate with naltrexone may result in a 33 percent increase in the C_{max} of acamprosate, which requires no adjustment of the dosage. Similarly, the pharmacokinetics of naltrexone and its major metabolite 6-beta-naltrexol are unaffected when given concomitantly with acamprosate. Patients taking acamprosate with antidepressants more commonly reported both weight gain and weight loss in comparison to patients taking either medication alone.

The absorption of several agents (tetracyclines, quinolones, levothyroxine, atenolol, iron, alendronate, sodium fluoride, and zinc) may be decreased by the calcium component of acamprosate during concurrent administration. The absorption of acamprosate can also be decreased by food. Acamprosate has shown no inhibiting or inducing potential both in vitro and in vivo on the enzymes of the cytochrome P450 system.

Laboratory Interference

Acamprosate is not known to interfere with any laboratory tests.

Dosage and Administration

Initiation of therapy should begin after alcohol detoxification has taken place and the patient has abstained from alcohol for 5 to 7 days. Just as it should not be initiated during acute detoxification, a review of the literature substantiates that it may not be superior to placebo if the medication is initiated more than 25 days postacute alcohol detoxification. Acamprosate treatment should not be interrupted if relapses occur and is recommended to proceed for 12 months.

Acamprosate is dosed by weight: patients >60 kg should receive 1,998 mg/day in divided doses with meals, while patients <60 kg should receive 1,332 mg/day in three divided doses with meals. The drug has not been sufficiently studies in pediatric and geriatric populations.

A.B. is a 54-year-old male with a life-long history of alcoholism. Through his involvement with Alcoholic Anonymous (AA) he managed to maintain abstinence for close to 2 decades before relapsing few months back. The current relapse was accompanied by increased feelings of loneliness and sadness along with self-isolation and binging episodes. Remarkably his work performance has been sustained, but he felt that "things will collapse soon" and self-enrolled in an outpatient 12-step based program. A.B. articulately participated in group sessions twice a week, but shared that "I have heard all of that many times before" referring to his prior experience with AA and felt that he is not "getting much out of it." Although he was able to maintain longer nondrinking intervals between relatively short-lived binges he was very concerned that "next big relapse is about to happen at any time with no warning" and continued to feel unable to enjoy some of his work-related successes. More specifically, he was recently offered a very high level position in a not-for-profit organization that was involved in a very high-end political fund-raising. On several occasions he stated that he cannot allow this opportunity to fail because of his inability to stay abstinent.

Within a month after A.B. joined the outpatient program he requested an appointment with the psychiatrist to discuss medication treatment in respect to his worsening depression and ongoing tendency to relapse on short-lived binges. He had used an SSRI medication for several years in the past and has stopped it about a year ago "because I did not need it anymore." He has considered restarting the medicine and was in agreement with such recommendation. He has also researched medications for alcoholism such as disulfiram (Antabuse), naltrexone (Revia) and acamprosate (Campral). He seemed concerned about Antabuse's "serious side effects" and decided on starting acamprosate (Campral). He found the treatment easy to tolerate and probably effective since he was abstinent for about 4 weeks until he had yet another relapse that he acknowledged was somewhat premeditated. Needless to say, A.B. felt very discouraged by his self-sabotaging behavior and possibly compelled to some extend by desperation decided to try disulfiram (Antabuse). To assure success he made a significant effort to be consistent with it on daily basis and shared that "as long as I take it I know I cannot drink." On a later date he inquired on the compatibility of disulfiram (Antabuse) and acamprosate (Campral) since the latter "seemed to have reduced the sudden and difficult to control urges to drink," which he found so distressing. About 2 months after he started Antabuse A.B. added Campral to his daily regimen. Ultimately the combination of disulfiram (Antabuse) and acamprosate (Campral) turned to successfully help A.B. in his motivation to consistently maintain sobriety.

REFERENCES

*Center for Substance Abuse Treatment. *Incorporating Alcohol Pharmacotherapies Into Medical Practice: A review of the Literature. Treatment Improvement Protocol (TIP) Series, No. 49s*. Rockville, MD: Substance Abuse and Mental Health Services Administration (US); 2009.

Diehl A, Ulmer L, Mutschler J, et al. Why is disulfiram superior to acamprosate in the routine clinical setting? A retrospective long-term study in 353 alcohol-dependent patients. *Alcohol*. 2010;45:271.

Erickson CA, Ray B, Maloney B, et al. Impact of acamprosate on plasma amyloid-β precursor protein in youth: a pilot analysis in fragile X syndrome-associated and idiopathic autism spectrum disorder suggests a pharmacodynamics protein marker. *J Psychiatr Res*. 2014;49:220.

Fuller RK, Branchey L, Brightwell DR, et al. Disulfiram treatment of alcoholism. A Veterans Administration cooperative study. *JAMA*. 1986;256(11):1449.

Geller RJ. Chapter 63. Disulfiram. In: Olson KR, ed. *Poisoning & Drug Overdose,* 6e; 2012.

Hammarberg A, Jayaram-Lindstöm N, Beck O, Franck J, Reid MS. The effects of acamprosate on alcohol-cue reactivity and alcohol priming in dependent patients: a randomized controlled

trial. *Psychopharmacology*. 2009;205(1):53–62.

*Jonas DE, Amick HR, Feltner C, et al. Pharmacotherapy for adults with alcohol use disorders in outpatient settings: a systematic review and meta-analysis. *JAMA*. 2014;311(18):1889–1900.

Jørgensen CH, Pedersen B, Tønnesen H. The efficacy of disulfiram for the treatment of alcohol use disorder. *Alcohol Clin Exp Res*. 2011;35:1749–1758.

*Karila L, Reynaud M, Aubin HJ, et al. Pharmacological treatments for cocaine dependence: is there something new? *Curr Pharm Des*. 2011;17(14):1359–1368.

Kiefer F, Mann K. Acamprosate: how, where and for whom does it work? Mechanism of action, treatment targets and individualized therapy. *Curr Pharm Des*. 2010;16(9):2098–2102.

Krampe H, Stawicki S, Hoehe MR, Ehrenreich H. Outpatient Long-term Intensive Therapy for Alcoholics (OLITA): a successful biopsychosocial approach to the treatment of alcoholism. *Dialogues Clin Neurosci*. 2007;9:399–412.

Kufahl PR, Watterson LR, Olive MF. The development of acamprosate as a treatment against alcohol relapse. *Expert Opin Drug Discov*. 2014;9(11):1355–1369.

Laaksonen E, Koski-Jännes A, Salaspuro M, Ahtinen H, Alho H. A randomized, multicentre, open-label, comparative trial of disulfiram, naltrexone and acamprosate in the treatment of alcohol dependence. *Alcohol Alcohol*. 2008;43:53–61.

Maisel NC, Blodgett JC, Wilbourne PL, Humphreys K, Finney JW. Meta-analysis of naltrexone and acamprosate for treating alcohol use disorders: when are these medications most helpful. *Addiction*. 2013;108(2):275–293.

Malcolm R, Olive MF, Lechner W. The safety of disulfiram for the treatment of alcohol and cocaine dependence in randomized clinical trials: guidance for clinical practice. *Expert Opin Drug Saf*. 2008;7(4):459–472.

Mann K, Kiefer F, Smoka M, Gann H, Wellek S, Heinz A. Searching for responders to acamprosate and naltrexone in alcoholism treatment: rationale and design of the PREDICT study. *Alcohol Clin Exp Res*. 2009;33(4):674–683.

Mann K, Lemenager T, Hoffman S, et al. Results of a double-blind, placebo-controlled pharmacotherapy trial in alcoholism conducted in Germany and comparison with the US COMBINE study. *Addict Biol*. 2013;18(6):937–946.

Mason BJ, Goodman AM, Chabac S, Lehert P. Effect of oral acamprosate on abstinence in patients with alcohol dependence in a double-blind placebo-controlled trial: the role of patient motivation. *J Psychiatr Res*. 2006;40(5):383–393.

Mason BJ, Heyser CJ. The neurobiology, clinical efficacy and safety of acamprosate in the treatment of alcohol dependence. *Expert Opin Drg Saf*. 2010;9(11):177–188.

Mason BJ, Lehert P. Acamprosate for alcohol dependence: a sex-specific meta-analysis based on individual patient data. *Alcohol Clin Exp Res*. 2012;36(3):497–508.

O'Farrell TJ, Fals-Stewart W. Family-involved alcoholism treatment. An update. *Recent Dev Alcohol*. 2001;15:329–356.

Oka M, Hirouchi M, Tamura M, Sagahara S, Oyama T. Acamprosate {monocalcium bis(3-acetaminopropane-1-sulfonate)} reduces ethanol-drinking behavior in rats and glutamate-induced toxicity in ethanol-exposed primary rat cortical neuronal cultures. *Eur J Pharmacol*. 2013;718:323.

Pani PP, Trogu E, Vacca R, Amato L, Vecchi S, Davoli M. Disulfiram for the treatment of cocaine

dependence. *Cochrane Database Syst Rev*. 2010;CD007024.

Petrakis IL, Nich C, Ralevski E. Psychotic spectrum disorders and alcohol abuse: a review of pharmacotherapeutic strategies and a report on the effectiveness of naltrexone and disulfiram. *Schizophr Bull*. 2006;32(4):644–654.

Reprotox. Agent Detail - DISULFIRAM. (2014, June). Available at http://www.reprotox.org/Members/AgentDetail.aspx?a=1228. Accessed on November, 2014.

Roache JD, Kahn R, Newton TF, et al. A double-blind, placebo-controlled assessment of the safety of potential interactions between intravenous cocaine, ethanol, and oral disulfiram. *Drug Alcohol Depend*. 2011;119(1–2):37.

*Rösner S, Hackl-Herrwerth A, Leucht S, Lehert P, Vecchi S, Soyka M. Acamprosate for alcohol dependence. *Cochrane Database Syst Rev*. 2010;8(9).

Sharma DK, Kaur S, Singh J, Kaur I. Role of acamprosate in sensorineural tinnitus. *Indian J Pharmacol*. 2012;44(1):93.

*Skinner MD, Lahmek P, Pham H, Aubin HJ. Disulfiram efficacy in the treatment of alcohol dependence: a meta-analysis. *PLoS One*. 2014;9(2):e87366.

Spanagel R, Vengeliene V, Jandeleit B, et al. Acamprosate produces its anti-relapse effects via calcium. *Neuropsychopharmacology*. 2014;39(4):783.

▲ 34.17 First-Generation Antipsychotics

MARTIN T. STRASSNIG, M.D., AND PHILIP D. HARVEY, PH.D.

HISTORY

There is now emerging agreement that antipsychotic drugs should be differentiated based on their individual side-effect profiles in addition to their efficacy. Although 5-HT_{2A} and 5-HT_C inverse agonism (or antagonism) in addition to dopamine D2 antagonism has been a mainstay of dividing newer from older antipsychotics, several of the (older) first-generation agents (e.g., clozapine [Clozaril], thioridazine [Mellaril], melperone [Buronil], and cyamemazine) have considerable 5-HT_{2A} antagonism as well, making the distinction less useful. Moreover, it has become increasingly clear that SGAs differ among each other in many properties and are not a homogeneous class either.

The FGAs represent the first group of effective agents for schizophrenia and other psychotic illnesses. They include all of the antipsychotics in the following groups: phenothiazines, butyrophenones, thioxanthenes, dibenzoxazepines, dihydroindoles, and diphenylbutylpiperidines. Although the FGAs have gradually been replaced by the SGAs in the United States, the FGAs remain the most commonly prescribed antipsychotics in many parts of the world, and as cost considerations have become more prominent, FGAs remain frequently prescribed in the United States as well.

These agents are highly effective for nearly all disorders that result in

psychotic symptoms. However, these drugs have important limitations: they can cause acute and chronic neurological symptoms that, especially at higher doses, significantly limit their effectiveness; they are ineffective or only partially effective for a substantial proportion of psychotic patients; the minimally effective dosage has frequently not been determined and therefore patients may be over- or undermedicated; and they have limited efficacy for some of the most important signs and symptoms of the most common psychotic illness, schizophrenia. In higher dosages, the motor symptoms may induce "secondary" negative symptoms, such as akinesia, emotional blunting, apathy, and cognitive impairment.

Despite their limitations, FGAs have played an important role in the management of psychosis. The modern era of psychopharmacology began with the discovery of chlorpromazine (Thorazine), the first effective antipsychotic, in the early 1950s. The annual increase in state and county mental hospital beds in the United States that had occurred since the beginning of the 20th century reversed shortly after the arrival of antipsychotics. This occurred because most psychotic patients with schizophrenia and other psychotic illnesses became clinically stable enough to be discharged from hospitals after they received drug treatment. Moreover, patients with stabilized illnesses became candidates for psychosocial treatments and rehabilitation. There is also evidence suggesting that treatment with antipsychotics can improve the long-term course of schizophrenia.

The FGAs have wider applicability than for the treatment of schizophrenia only. They are effective for treating psychosis regardless of its cause. As a result, antipsychotics are standard treatments for many patients with psychoses associated with physical origins, bipolar illness, major depressions with psychotic features, and schizoaffective disorder.

The first phenothiazines were the first effective antipsychotic drugs. Drugs from this class were originally used as antihelminthics in veterinary medicine and as urinary antiseptics in humans. In 1950, Paul Charpentier, at the Rhóne-Poulenc Laboratories in Paris, synthesized chlorpromazine, a drug that appeared notable as a mild antihistaminic that was also sedating. Henri Laborit, a surgeon with an interest in drugs that would decrease preoperative anxiety and prevent postsurgical shock, convinced the Rhóne-Poulenc Laboratories of the importance of identifying compounds that would relax patients and make surgical shock less likely. Chlorpromazine appeared to fulfill these criteria, but it also appeared qualitatively different from other sedating drugs. Patients appeared indifferent toward the environment and more tranquil. These properties led Laborit to convince psychiatrists, including Jean Delay and Pierre Deniker, to administer chlorpromazine to patients with mania and schizophrenia. In 1952, these authors reported their observation that chlorpromazine was effective in treating mania and schizophrenia. Within a year, the atmosphere in Paris's

psychiatric hospitals had improved substantially. The discovery of chlorpromazine's effectiveness for psychosis represents the beginning of modern psychopharmacology in that it indicated that drugs could have specific effects on the symptoms of mental disorders that went beyond sedation. Moreover, the availability of drugs with direct effects on mental disorders became a powerful tool for laboratory studies with animals, leading to a greater understanding how drugs affect the brain and information about the nature of the illnesses that were affected.

The use of chlorpromazine was reported by Lehman and Hanrahan from Montreal in 1954 and it rapidly captured the interest of US psychiatrists. In 1955, Henry Brill, the assistant state mental health commissioner in New York, initiated the general use of chlorpromazine and reserpine in state hospitals in that state. Although reserpine was also an effective antipsychotic, its side effects led clinicians to prefer chlorpromazine and, as a result, it is no longer used for treating psychosis because of its side effects, such as depression, suicidality, dysphoria, and orthostatic hypotension.

In 1953, reports surfaced indicating that chlorpromazine could cause neurological adverse effects, predominantly parkinsonian symptoms. Although some individuals warned that these symptoms were evidence of neural toxicity, others suggested that the dosage should be increased until patients experienced a mild form of parkinsonism, termed the "neuroleptic threshold" dose. Later that year, clinicians noted that chlorpromazine could also result in forms of persistent dyskinesia, a syndrome that would later be called tardive dyskinesia (TD). Thus, information about the side effects of FGA medications has been available since their introduction.

The introduction of chlorpromazine was followed by the introduction of other phenothiazines including perphenazine (Trilafon) and fluphenazine (Prolixin). In 1958, the first effective butyrophenone, haloperidol (Haldol), was introduced by Paul Janssen from Janssen Pharmaceutica in Beerse, Belgium. The first thioxanthene antipsychotics were introduced the same year by Peterson and his coworkers. Although many other FGAs were introduced in subsequent years, most of these agents were similar in effectiveness and differed only in their propensity to cause side effects. The first long-acting antipsychotic, fluphenazine enanthate, was introduced in the early 1960s. Although this and subsequent long-acting agents, such as haloperidol decanoate, had similar effectiveness and side effects to other antipsychotics, they provided a method for administering an antipsychotic without requiring patients to take daily oral medication, improving treatment adherence and reducing relapse and rehospitalization rates.

CHEMISTRY

The FGAs have been referred to by a number of titles including neuroleptics, major tranquillizers, dopamine receptor antagonists, and

antipsychotics (typical or FGAs). The designation neuroleptic was suggested by Deniker who noted the tendency of these agents to have neurological side effects and thus appear to "take the neuron," based on the observation that neurological side effects are inevitably associated with antipsychotic activity. However, the effectiveness of drugs such as clozapine suggests that neurological side effects are not necessary to treat psychosis. Thus, referring to a side effect of these agents rather than the primary clinical effect may be misleading. The term *major tranquillizer* referred to the strange quietness or blandness (or *ataraxia,* Latin) that had been associated with those agents. However, the suggestion that these agents function by their tranquillizing effects is also misleading.

In this section, the dopamine receptor antagonists are subclassified according to their chemical structure. Another system for subclassifying this group is to use the antipsychotic potency of these agents (Table 34.17–1), dividing them into low-, mid-, and high-potency agents. This latter method is probably more useful to clinicians since this classification provides information about the amount of drug that is required for a clinical effect and the likelihood of important side effects. The antipsychotic drugs come from a number of chemical classes, each with somewhat different pharmacokinetic properties. The differences can have important effects on the clinical use of the drugs. The phenothiazines and the thioxanthenes share structural similarities and have a number of shared metabolic pathways. The same is true for the butyrophenones and the diphenylbutylpiperidines. This overview will focus on chlorpromazine and haloperidol as representative compounds. Since the pharmacokinetics of the long-acting agents differs substantially from the short-acting agents, they are discussed separately.

Phenothiazines

All of the phenothiazines have the same three-ring structure with different side chains joined at the nitrogen atom of the middle ring (Fig. 34.17–1). The activity of the group can be affected by substitutions at position 2 or 10. The phenothiazines are usually categorized into three classes based on substitutions at position 10. These substitutions have important effects on the pharmacological characteristics of the phenothiazines. Thus, substituting an *aliphatic* side chain results in a drug with relatively low antipsychotic potency such as chlorpromazine. These agents also tend to be sedating and to cause anticholinergic and hypotensive effects at their effective dosages. Chlorpromazine has a chlorine atom at position 2. If the chlorine atom is eliminated, the resulting drug is promazine (Sparine), a weak antipsychotic. Substituting a *piperidine* ring at position 10 leads to an effective group of antipsychotics characterized by thioridazine and mesoridazine (Serentil). These drugs have similar potency and side effects when compared with the aliphatic phenothiazines. Fluphenazine and

trifluoperazine (Stelazine) are representative antipsychotics with a *piperazine* group substituted at the 10 position. Piperazine phenothiazines have fewer anticholinergic and autonomic effects, but greater affinity for D_2 sites, and as a result produce more EPS. A number of piperazine phenothiazines have been esterified at a free hydroxyl group with ethanoic or decanoic acid to produce injectable long-acting antipsychotics.

Table 34.17–1.
Dopamine Receptor Antagonists: Potency and Adverse Effects

			Side Effects		
Drug Name	**Chemical Classification**	**Therapeutically Equivalent Oral Dosage (mg)**	**Sedation**	**Autonomic[a]**	**Extrapyramidal Reactions[b]**
Pimozide[c] Orap	Diphenylbutylpiperidine	1.5	+	+	+++
Fluphenazine Permitil Prolixin	Phenothiazine: piperazine compound	2	+	+	+++
Haloperidol Haldol	Butyrophenone	2	+	+	+++
Thiothixene Navane	Thioxanthene	4	+	+	+++
Trifluoperazine Stelazine	Phenothiazine: piperazine compound	5	++	+	+++
Perphenazine Trilafon	Phenothiazine: piperazine compound	8	++	+	++/+++
Molindone Moban	Dihydroindolone	10	++	+	+
Loxapine Loxitane	Dibenzoxazepine	10	++	+/++	++/+++
Prochlorperazine[c] Compazine	Phenothiazine: piperazine compound	15	++	+	+++
Acetophenazine Tindal	Phenothiazine: piperazine compound	20	++	+	++/+++
Trifluprom azine Vesprin	Phenothiazine: aliphatic compound	25	+++	++/+++	++
Mesoridazine	Phenothiazine:	50	+++	++	+

	piperidine compound				
Serentil					
Chlorpromazine	Phenothiazine: aliphatic compound	100	+++	+++	++
Thorazine					
Chlorprothixene	Thioxanthene	100	+++	+++	+/++
Taractan					
Thioridazine	Phenothiazine: piperidine compound	100	+++	+++	+
Mellaril					

+, minimally present; ++, moderately present; +++, present in most patients.

[a]Anti–α-adrenergic and anticholinergic effects.

[b]Excluding tardive dyskinesia, which appears to be produced to the same degree and frequency by all agents with equieffective antipsychotic dosages.

[c]Pimozide is used principally in the treatment of Tourette syndrome; prochlorperazine is used rarely, if ever, as an antipsychotic agent.

Adapted from American Medical Association. *AMA Drug Evaluations: Annual 1992*. Chicago: American Medical Association; 1992.

PHENOTHIAZINES

Aliphatic

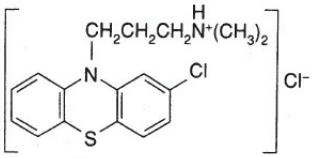

Chlorpromazine

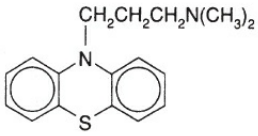

Promazine

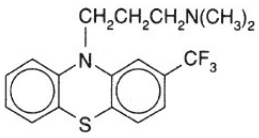

Trifluprom azine

Piperazine

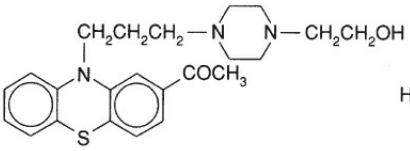

Acetophenazine

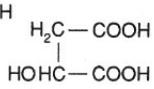

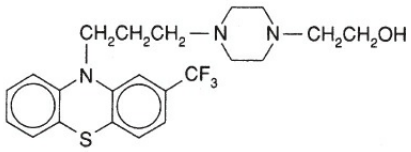

Fluphenazine

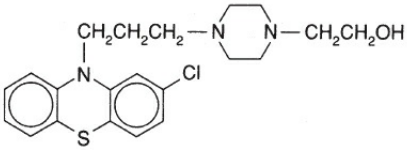

Perphenazine

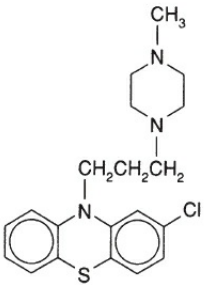

Prochlorperazine

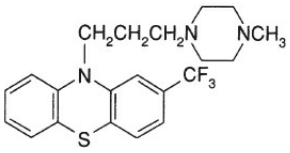

Trifluoperazine

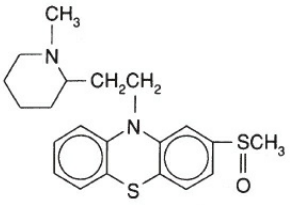

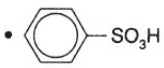

Mesoridazine

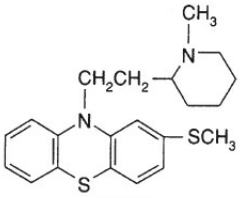

Thioridazine

THIOXANTHENES

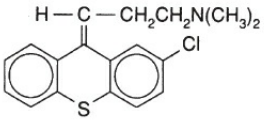

Chlorprothixene

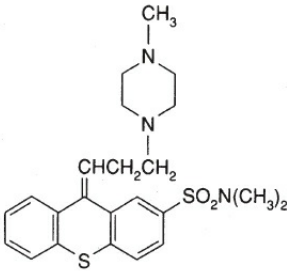

Thiothixene

A

Dibenzoxazepine

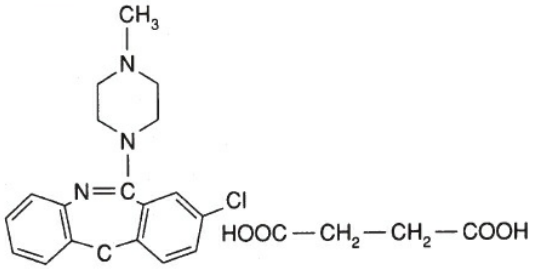

Loxapine

Dihydroindole

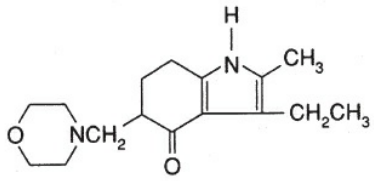

Molindone

Butyrophenones

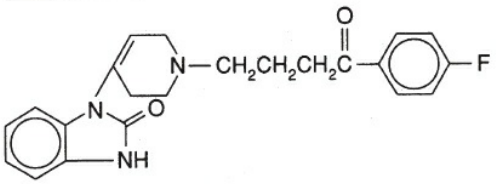

Droperidol

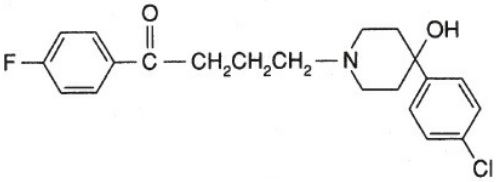

Haloperidol

Diphenylbutylpiperidine

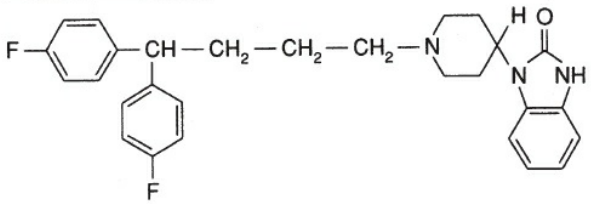

Pimozide

Benzamide

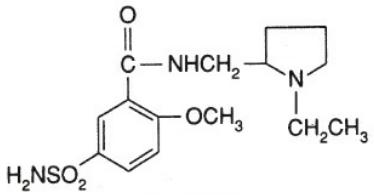

Sulpiride (not available in U.S.)

Rauwolfia Alkaloid

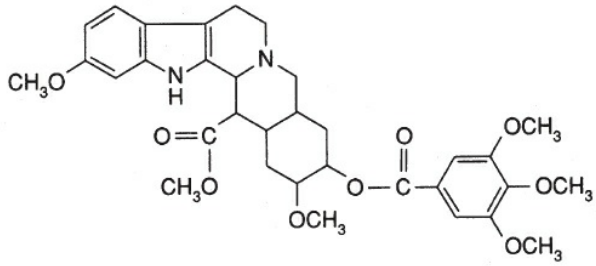

Reserpine

B

FIGURE 34.17–1. Molecular structure of dopamine receptor antagonists and reserpine.

Thioxanthenes

This group shares a similar three-ring structure with the phenothiazines, but has a carbon atom substituted for nitrogen at position 10. Chlorprothixene (Taractan) is an effective low-potency *aliphatic* thioxanthene agent with an antipsychotic potency and a side-effect profile similar to its phenothiazine counterpart, chlorpromazine. Thiothixene (Navane), clopenthixol (Sordinol), and flupenthixol (Depixol) are thioxanthenes with a *piperazine* substitution. Again, their clinical characteristics resemble those of their phenothiazine counterparts. As a result of their double bond at position 10, these agents have geometric isomers with the *cis* isomers demonstrating greater antipsychotic activity.

Butyrophenones

This class of antipsychotic drugs is characterized by a substituted phenyl ring, which is attached to a carbonyl group attached by a 3-carbon chain to a tertiary amino group. Most of the clinically useful butyrophenones have a piperidine ring attached to the tertiary amino group. Haloperidol is the representative drug for this class. It and other butyrophenones tend to be potent D_2 antagonists and have minimal anticholinergic and autonomic effects, as well as reduced sedation compared to phenothiazines. Haloperidol, a substituted piperidine, is the most commonly used (and oldest) drug from this class. Other important butyrophenones include droperidol (Dropletan), a short-acting, sedating compound that is commonly used to control agitation; and spiperone (Spiroperidol), an agent with very high affinity for D_2, 5-HT_{2a}, and 5-HT_{2b} receptors that is commonly utilized for labeling receptors in animal studies and in PET scanning.

Dibenzoxazepines

This class of antipsychotic drugs has a three-ring structure with a seven-member center ring. Loxapine (Loxapac), a dibenzoxazepine, is the only drug from this group that is available in the United States. Clozapine, a dibenzodiazepine, differs from loxapine in having a nitrogen instead of an oxygen atom in the middle ring, as well as differences in side chains.

Dihydroindoles

These derivatives are structurally related to serotonin, melatonin, and indole hallucinogens, such as dimethyltryptamine. The only dihydroindole available in the United States was molindone (Moban), the only antipsychotic agent without known weight gain liability. Molindone has

since been discontinued from the market in 2010.

Diphenylbutylpiperidines

This group of agents is similar in structure to the butyrophenones. Pimozide (Orap), the only diphenylbutylpiperidine available in the United States (indication for Tourette syndrome), is derived from the butyrophenone benperidol. It has a long half-life: 20 to 26 hours. Doses of 20 mg or higher are associated with seizure risk and QTc prolongation. This agent has been removed from the market for psychotic conditions because of those reasons. Penfluridol (Semap) and fluspirilene (experimental) are antipsychotics from this group that are not available in the United States.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

Plasma concentrations of short-acting drugs rise rather rapidly during what is referred to as the absorption phase and then decline during a distribution, metabolic, and elimination phase. Each of these three phases is influenced by different factors related to the chemical structure of the agent (Table 34.17–2).

Absorption

In general, the FGAs are well absorbed when they are administered orally or parenterally. As with most drugs, oral administration leads to less predictable absorption than parenteral administration. Liquid concentrates are absorbed slightly more rapidly than pills. Plasma concentrations of the drugs usually reach peak levels 1 to 4 hours after ingestion and 30 to 60 minutes after intramuscular (IM) administration. In general, IM preparations reach their peak concentrations sooner than oral drugs and as a result have an earlier onset of action. For example, IM administration of most antipsychotics results in peak plasma levels in about 30 minutes with clinical effects emerging within 15 to 30 minutes. Oral administration of most FGAs results in a peak plasma level in 1 to 4 hours. A number of factors can interfere with the GI absorption of these drugs including antacids, coffee, smoking, and fatty foods.

Table 34.17–2.
Pharmacokinetics of Selected Antipsychotics

Drug	Bioavailability (%)	Protein-Binding (%)	Vd (L/kg)	$t_{1/2}$ (hours)	Active Metabolites	Cp 12 hours Postdose (ng/mL)
Chlorpromazine	10–33	90–95	7–20	8–35	7-Hydroxy, others	100–300

Thioridazine	25–33	99	—	9–30	Mesoridazine, Sulfordiazine	200–800
Perphenazine	~25	90–95	10–35	8–21	None known	0.5–5.0
Fluphenazine	20–50	90–95	—	14–24	7-Hydroxy	0.2–3.0
Haloperidol	40–70	92	10–35	12–36	Reduced haloperidol	2.0–18.0
Thiothixene	~50	90–95	—	34	None known	1.0–5.0

Vd, volume distribution; Cp, concentration in plasma after 12 hours.

Adapted from Ereshefsky L. Pharmacokinetics and drug interactions: Update for new antipsychotics. *J Clin Psychiatry*. 1996;57 Suppl 11:12–25.

Drugs usually reach steady state in about three to five half-lives. Thus, steady state levels for chlorpromazine, haloperidol, and most other FGAs are reached in about 3 to 5 days since their half-lives are about 24 hours. The bioavailability (i.e., the amount of drug reaching the systemic circulation) is substantially greater—as much as 10-fold—when FGAs are administered parenterally. This difference may result from the incomplete absorption of the drug in the GI tract and from extensive metabolism of oral drugs during the first pass mainly through the liver.

Distribution

The initial decline in the plasma concentration of an antipsychotic (or any drug) results primarily from drug distribution into various bodily compartments. Since antipsychotics are highly lipophilic, they tend to accumulate in tissues such as fat, lungs, and brain. Brain levels of antipsychotics tend to be substantially higher than plasma levels. Since plasma and brain levels are in equilibrium, measuring plasma levels is sufficient for estimating brain levels. However, behavioral effects may last substantially longer than the blood levels indicate. For instance, sexual side effects may last for several weeks after the last dose of chronic treatment.

Most FGAs are highly protein bound. For example, more than 90 percent of drugs such as fluphenazine and haloperidol are bound to plasma protein. The remaining unbound portion is the drug that is available for passing through the blood–brain barrier. In theory, conditions that alter the amount of plasma protein such as malnutrition will also alter the amount of bioavailable antipsychotic drug. This is potentially important for drugs that are protein bound for 98 percent or more.

Metabolism and Elimination

The metabolism of antipsychotic drugs is largely hepatic and occurs through conjugation with glucuronic acid, hydroxylation, oxidation, demethylation, and sulfoxide formation. The metabolism of the phenothiazines and thioxanthenes is particularly complex. For example, chlorpromazine has more than 100 different potential metabolites, with

some metabolites having significant amounts of pharmacological activity. The case of thioridazine, a substantial amount of the drug's activity, is from metabolites that may be more active than thioridazine itself. One metabolite, mesoridazine, is also marketed as an antipsychotic. However, haloperidol has only one major metabolite, reduced haloperidol, which has substantially less antidopaminergic activity than the parent compound. However, reduced haloperidol is converted back to the parent compound and in this manner may contribute to antipsychotic activity.

Most FGAs are metabolized by the cytochrome P450 2D6 and P450 3A subfamilies in the liver. Since these same isoenzymes also metabolize a number of drugs that are commonly combined with antipsychotics, a number of important drug–drug interactions are possible. These are discussed in the section on drug–drug interactions.

Patients demonstrate substantial differences in the extent to which they metabolize these compounds. This explains, in part, the marked differences among patients in the oral neuroleptic dosages required for treatment. This, along with differences in absorption, may also explain the 20-fold variation in plasma concentrations that can be found among patients treated with the same dosage of an antipsychotic.

The systemic clearance of FGAs is high as the result of a high hepatic extraction ratio. As a result, only negligible amounts of the unchanged drug are excreted by the kidneys. Phenothiazines, thioxanthenes, and their metabolites are excreted in the urine and the feces.

Pharmacokinetics of Long-Acting Antipsychotics

The pharmacokinetics of long-acting injectable antipsychotics differs markedly from those of short-acting oral and injectable drugs. Long-acting fluphenazine and haloperidol are administered as esters dissolved in sesame oil. The oil is injected into a muscle and the drug gradually diffuses from the oily vehicle into the surrounding tissues. The rate-limiting step appears to be the rate of diffusion, since once the drug enters the tissue it is rapidly hydrolyzed and the parent compound is released. Long-acting compounds, however, are absorbed continually during the interval between injections. Moreover, for patients who have received multiple injections, the drug will be absorbed from multiple injection sites simultaneously. As a result, it takes long-acting compounds much longer to reach steady state, and they are eliminated much more slowly than their oral counterparts. For example, the decanoate forms of haloperidol and fluphenazine each require about 3 months to reach steady state, and substantial plasma concentrations can be detected up to several months after therapy has been discontinued. This has consequences for prolonged side effects, and potentially for TD risk, even after discontinuation of therapy.

ELDERLY PATIENTS. Elderly patients require lower than usual dosages of

FGAs for a number of reasons. These include decreased renal clearance, decreased cardiac output, decreased liver size, and decreased CYP450 enzyme. In addition, elderly patients may be more sensitive to the parkinsonian side effects of the drugs. Recently, FGAs (and SGAs) have been associated with increased mortality risk in elderly patients with dementia.

PHARMACODYNAMICS

The pharmacodynamics of an antipsychotic describes its activity at important sites of action, particularly at its brain receptors. Recent findings in neuroscience have provided important information about the effects of FGAs at a number of different receptor sites, how these effects influence groups of neurons, and how activity at receptors is translated into intracellular events.

MECHANISM OF ACTION OF DOPAMINE RECEPTOR ANTAGONISTS

The Dopamine Hypothesis

The dopamine hypothesis states that antipsychotics reduce psychotic symptoms by decreasing dopamine activity. It was originally proposed by Arvid Carlsson from Sweden and other basic scientists based on the observation that haloperidol and chlorpromazine raised the levels of dopamine metabolites in dopamine-rich areas of mouse brain, interpreted as dopamine antagonism, and compensation of self-regulating systems by increasing dopamine production. Other investigators noted that FGAs inhibit the activation and stereotypical behaviors of rodents induced by amphetamine. This effect has been found to be a reasonably reliable predictor of antipsychotic activity in patients. Moreover, the dose–response relationships in animals provide information about the likely clinical dosage range for patients. These behaviors are thought to be mediated by dopamine and have been used as screening devices to identify compounds that are likely to have clinical efficacy as antipsychotics. Other evidence supporting a role of dopamine in psychosis comes from the observation that all drugs—at least to this date—capable of reducing psychosis are also dopamine antagonists, including partial dopamine agonists. Moreover, agents that are associated with increased dopamine activity such as amphetamine, methylphenidate, or cocaine tend to increase dopamine neurotransmission and also tend to worsen symptoms in patients with schizophrenia. At persistent high dosages, these drugs may cause psychosis in individuals without pre-existing psychiatric disorders.

A role for dopamine in the activity of antipsychotics, originally proposed by Carlsson, the Nobel laureate, was given strong support by the work of Solomon Snyder from Johns Hopkins University and Philip Seeman from the University of Toronto in the mid-1970s. Using similar receptor binding

methods, both found that the antipsychotic potency of a number of drugs was highly correlated with the drug's affinity for D_2 receptors. More recent studies, using highly selective D_2 receptor ligands and PET indicate that dopamine receptor antagonists are effective when approximately 80 percent of D_2 receptors in the striatum are occupied. Higher occupancy rates may be associated with more parkinsonian side effects without greater effectiveness. In contrast, clozapine is effective when only 40 to 60 percent of D_2 receptors are occupied (Table 34.17–3). Although it remains unclear why clozapine is effective at lower levels of D_2 occupancy, it has been suggested that activity at other receptors, particularly 5-HT_{2A} receptors, may contribute to clozapine's clinical activity, and to its major active metabolite, *N*-des-methyl clozapine (NDMC), which has partial $D_{2/3}$ receptor agonist activity in addition to being a strong 5-$HT_{2A/2C}$ inverse agonist (or antagonist). The 5-HT_{2A} antagonism by itself has been suggested to have antipsychotic effects, albeit not as strong as $D_{2/3}$ antagonism (Fig. 34.17–2). NMDA receptor agonism at the glycine site may increase its efficacy as well.

Table 34.17–3.
Dopamine Receptor Occupancy of Currently Used Typical Antipsychotics in Relation to Clozapine

Drug	Daily Dosage (mg)	Serum Concentration (nmol/L)	Occupancy (%)	
			D_2	D_1
Phenothiazines				
Chlorpromazine	100 × 2	100	78	—
Thioridazine	150 × 2	620	74	—
Thioridazine	200 × 2	900	81	30
Perphenazine enanthate (7 days)	100	5	76	—
Perphenazine	8 × 2	4	—	0
Butyrophenones				
Haloperidol	6 × 2	19	84	—
Haloperidol	3 × 2	13	89	—
Haloperidol	3 × 2	9	84	—
Haloperidol	2 × 2	6	75	—
Haloperidol	2 × 2	11	84	—
Haloperidol	3 × 2	[a]	86	3
Haloperidol decanoate (28 days)	50	9	85	—
Haloperidol decanoate (28 days)	70	[a]	74	5
Diphenylbutyls				
Pimozide	4 × 2	19[a]	79	—

Dibenzodiazepines				
Clozapine	300 × 2	600	63	—
Clozapine	150 × 2	960	40	38
Clozapine	250 × 2	1,500	38	42
Clozapine	250 × 2	2,500	46	36
Clozapine	250 × 2	1,600	55	52

[a]Sample lost during transportation.

Data from Farde L, Nordstroem AL, Wiesel FA, et al. Positron emission tomographic analysis of central D_1 and D_2 dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine: relation to extrapyramidal side effects. *Arch Gen Psychiatry.* 1992;49:538–544; Wiesel FA, Farde L, Nordstrom AL, Sedvall G. Central D1- and D2-receptor occupancy during antipsychotic drug treatment. *Prog Nueropsychopharmacol Biol Psychiatry.* 1990;14(5):759–767.

Several studies using PET scanning have found that there is no difference in D_2 receptor occupancy in the striatum between individuals who respond to an antipsychotic and those who do not. This finding suggests either that processes beyond the dopamine receptor affect response or that dopamine blockade in the striatum is not the mediator of behavior change.

The Effects of FGAs on the Dopamine System

Dopamine neurons with cell bodies in the ventral tegmental area (the A_{10} cell group in rat brain) project via the meso-limbo-cortical dopamine system to the nucleus accumbens, amygdala, and the neocortex. Dopamine cells in the substantia nigra (the A_9 cell group) project to the caudate putamen via the nigrostriatal system. It has been proposed that activity in the meso-limbo-cortical system explains the antipsychotic activity of drugs, whereas neurological side effects such as EPS and TD result from activity in the nigrostriatal system. Thus, drugs with preferential activity at meso-limbo-cortical sites would result in clinical effectiveness with fewer neurological side effects. However, the reason that the SGAs have fewer EPS has been attributed to having additional effects on the $5\text{-}HT_{2A}$ receptors, which increases dopamine turnover.

Advances in molecular biology have led to the cloning of a number of dopamine receptor subtypes. The discovery of these subtypes has raised the question of their possible role in the treatment of psychosis. Dopamine receptors can be classified into two groups: D_1-like (D_1 and D_5) and D_2-like receptors (D_2, D_3, D_4). Moreover, each of these receptor subtypes may have variants that have the potential for explaining heterogeneity in drug responses. As mentioned previously, the antipsychotic potency of dopamine receptor antagonists is highly correlated with their affinity for D_2 receptors. This relationship has not been established with other dopamine

receptor subtypes. Nevertheless, certain characteristics of receptor subtypes suggest that they may contribute to clinical responses. For example, a role for D_4 receptors has been suggested by clozapine's high affinity for this subtype. Unfortunately, multiple attempts to bring D_4 antagonists to the market have failed for lack of antipsychotic efficacy. The localization of D_3 receptors in limbic areas suggests that these subtypes may be clinically important for antipsychotic effects. In contrast, D_1 receptor blockade may be undesired, as D_1 stimulation in the cortex is associated with better cognitive functioning and D_1 antagonists have been reported to cause suicidality.

Studies of the dopamine system in animals and man have suggested important interactions between the subcortical and prefrontal dopamine systems that may affect responses to FGAs. Lesions in the prefrontal cortex of rats result in increases in dopamine activity in the striatum, suggesting that these two systems have important functional linkages. Similarly, increases in dopamine activity in the prefrontal cortex are associated with decreases in striatal activity. Also, a number of investigators propose that negative symptoms of schizophrenia are associated with decreased dopamine antagonist prefrontal activity, whereas psychotic symptoms are generated by subcortical increases in dopamine antagonist activity (temporal lobe). In this model, an effective antipsychotic would increase dopamine activity in prefrontal areas and decrease activity in subcortical areas.

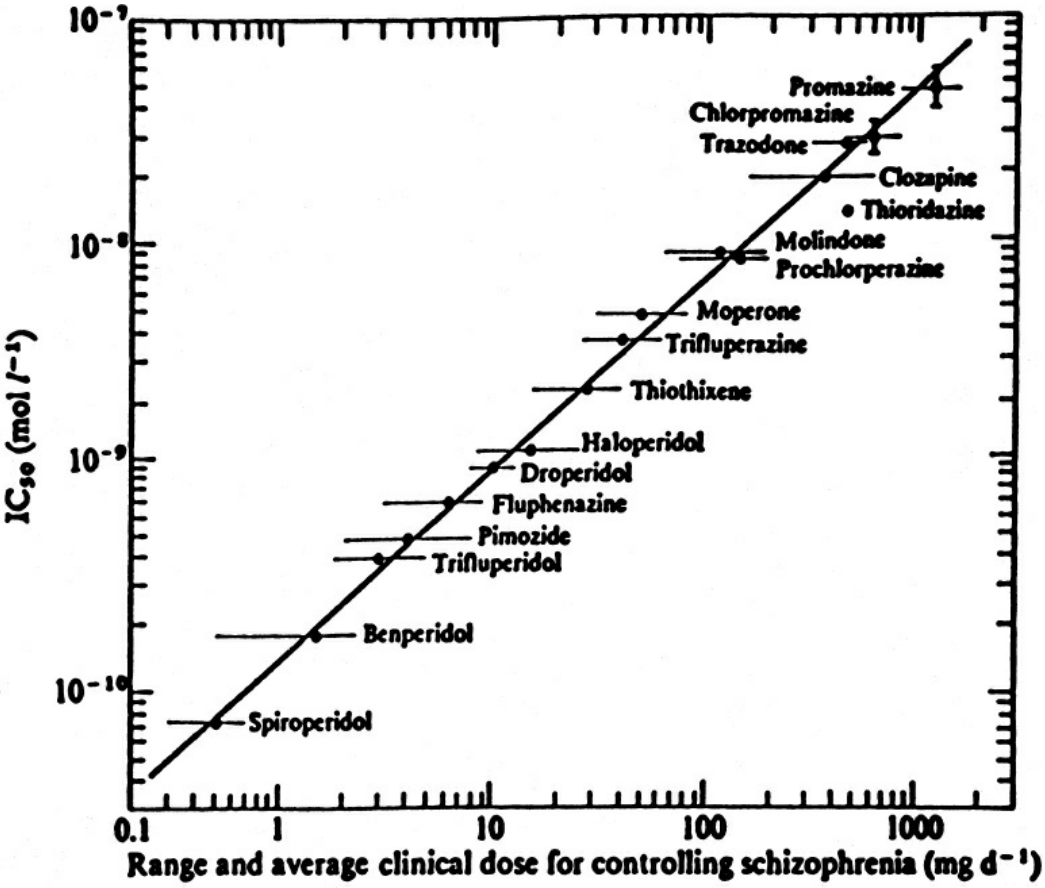

FIGURE 34.17–2. Blockage of dopamine release. (From Seeman P, Lee T, Chau-Wong M, Wong K. Antipsychotic drug doses and neuroleptic/dopamine receptors. *Nature.* 1976;261(5562):717–719.)

Other clinical evidence for a role of dopamine in the antipsychotic effects of dopamine receptor antagonists comes from studies of homovanillic acid (HVA), a metabolite of dopamine. In these investigations, changes in plasma HVA on an antipsychotic are used as reflecting changes in dopamine activity in the brain. These studies indicate that plasma HVA rises soon after patients begin treatment with an antipsychotic and then falls below pretreatment levels in patients who improve. Moreover, a number of studies indicate that patients with high levels of plasma HVA prior to drug treatment are more likely to improve than those with lower levels. These findings suggest that improvement on an antipsychotic agent is associated with an overall reduction in dopamine activity. Since subcortical areas have the greatest contribution to dopamine activity in the brain, this finding may not reflect cortical dopamine activity.

Studies that monitor the intracellular firing rates of dopamine neurons have also contributed to an improved understanding of the effects of antipsychotics. Chiodo, Grace, and Bunney have reported that the administration of haloperidol to rats resulted in a short-term increase in the firing of midbrain dopamine neurons followed by a prolonged decrease in firing that has been referred to as “depolarization inactivation.” It has

been proposed that the antipsychotic effects of these drugs are associated with the decrease in firing rate. Taken together with the studies of plasma HVA, these findings suggest that for an antipsychotic response to occur, the blockade of D_2 receptors is an initial response that must be followed by a decrease in dopamine activity.

DESIGN AND INTERPRETATION OF CLINICAL DRUG STUDIES

Design

The FGAs were the first specific compounds for treating a major mental illness. Their effectiveness, conclusively proven, challenged the views of some clinicians who believed that psychosis was a psychological reaction related to early experiences and that only psychological treatment could have antipsychotic effects. This initial skepticism was valuable in that it helped inspire the development of controlled clinical drug trials aimed at demonstrating the effectiveness of antipsychotics. Early studies, such as those developed by the Department of Veterans Affairs (VA) and the National Institute of Mental Health (NIMH), were influential in establishing objective research methods. One principle, now widely acknowledged, is for an antipsychotic (or any other new drug) to be considered effective; it should yield a better outcome than a placebo. That principle controls for nonspecific improvement that occurs because of a number of factors, such as (1) spontaneous fluctuations in symptoms; (2) the rate of spontaneous improvement for the illness; (3) the expectation of improvement, which is related to starting a new drug; (4) the subject's protection from environmental and psychosocial stresses that may occur with hospitalization; and (5) other environmental factors such as the hospital milieu and psychological treatments that can lead to improvement.

Two other principles that reduce sources of bias are random assignment to a treatment group and double-blind evaluation methods. Random assignment ensures that differences in outcome between the groups (usually between those receiving a drug and those receiving a placebo or between two different active groups) are not attributable to patients with different expected outcomes or drug responses being assigned to one group. An example would be a study in which patients who are more severely ill are assigned to a particular group or if patients who are predicted to improve spontaneously are assigned preferentially to the drug group rather than the placebo group. Double-blind methods were developed to ensure that bias by patients or by those who evaluate the outcome does not influence the study. For example, a person who performs a clinical rating may be prejudiced toward rating patients who are receiving a drug as more improved than patients who are receiving a placebo. Double-blind methods also protect against bias in the manner in which treatment is administered. For example, a treating clinician may give more

attention to a patient who is receiving a certain drug.

A number of modifications in trial methods have been developed since the first studies of FGAs, although the principles of random assignment and blindness remain important. For example, treating acutely ill patients with schizophrenia with FGAs has become a standard of care and has led some to question the ethical use of placebos when effective treatments are available. Alternative methods for demonstrating the efficacy of a compound have included showing that it is as effective as an antipsychotic agent with proven efficacy or showing that a dosage thought to be subtherapeutic is less effective than a higher dosage. Since blindness is often difficult to preserve when a drug has known side effects, active placebos with side effects but without CNS effects have been used. However, inexperienced evaluators tend to overinterpret side effects as being secondary to active treatment rather than to placebo, which also is associated with similar side effects, even so less frequent. Recent research have indicated that high interrater reliability and centralized ratings may decrease variance even further, while increased number of sites and more rating scales and raters may increase variance and therefore decrease the likelihood of separation from placebo.

INTERPRETATION

Acute Psychosis

Numerous double-blind clinical trials have demonstrated that FGAs are more effective than placebo for reducing symptoms of psychosis. The initial studies were carried out in the 1960s by the NIMH, reviewed by Donald Klein and John Davis who concluded that compared more than one antipsychotic and found that with the exception of mepazine (Pacatal) and promazine, all of these agents were effective. These and subsequent studies also demonstrated that antipsychotics are effective for every subgroup of patients with schizophrenia. In 2009, the Schizophrenia Patient Outcomes Research Team (PORT) updated prior reviews and concluded that there is clear evidence that FGAs are effective for reducing positive symptoms during acute psychotic episodes. PORT recommendations also suggest that choosing antipsychotics should be based on individual preference, prior treatment response and adherence history, medical history and risk factors, the individual medication's side-effect profile, and long-term treatment planning.

Controlled clinical trials have demonstrated that antipsychotic medications are effective in reducing psychotic symptoms, such as hallucinations and paranoid ideation to a greater extent than anxiety and depression. Improvements individual patient experiences on these agents vary considerably. For many patients, FGAs can lead to a complete remission of psychotic symptoms. The early studies from the NIMH

indicate that 60 percent of drug-treated patients improved to the extent that they achieved a complete remission (albeit temporarily) or experienced only mild symptoms; the remaining 40 percent will continue to experience psychotic symptoms, and approximately 8 percent of them are not improved with drug treatment.

Positive psychotic symptoms, which include hallucinations and delusions, are more responsive to dopamine receptor antagonists than are negative symptoms, such as blunted affect, emotional withdrawal, and lack of social or vocational interest or disorganized symptoms, such as impaired attention and impaired communication. Although these latter two groups of symptoms are less responsive than psychotic symptoms, double-blind studies indicate that these symptoms improve to some degree. EPS can cause secondary negative symptoms, compounding symptom burden. Patients with parkinsonism may demonstrate fewer expressive gestures and diminished interest in their environment. Anticholinergic antiparkinsonian medications, aimed at mitigating parkinsonian and other EPS symptoms, sometimes recommended prophylactically when used with a (high potency) FGA, may impair memory in patients who require these drugs, while EPS may worsen cognitive test scores that require intact motor function.

The symptomatology of all forms and types of schizophrenia improves with FGAs. Although it was proposed that low-potency, sedating FGAs were more effective for agitated patients and high-potency, nonsedating drugs were more effective for more withdrawn patients, this has never been substantiated by controlled trials. As a group, women tend to respond better and to require lower dosages than men. Unfortunately, there are no biological or clinical indicators that can be used to predict treatment response other than the patient's prior responses.

Maintenance Therapy

A number of studies have compared FGAs to placebos for preventing relapse in patients who have been stabilized following a psychotic episode. These studies provide strong support for continuing FGAs when patients have recovered from a psychotic episode. A review of 24 double-blind studies of maintenance antipsychotic therapy found that in all studies many more patients relapsed when they were not receiving an active drug. A meta-analysis that included most of these studies reported that approximately 72 percent of patients will relapse in a year on placebo compared to only 23 percent on a conventional antipsychotic. A more recent review cited 66 studies in which antipsychotics were withdrawn. The mean relapse rate was 53 percent for patients withdrawn from medication and 16 percent for patients maintained on an antipsychotic. The more robust the remission prior to discontinuation (the lower the psychosis ratings) is, the longer the time until relapse.

A number of investigations have focused on the usefulness of continuing medication for patients who have recovered from a first episode of schizophrenia. Again, maintenance medication showed a substantial advantage, although relapse rates were lower for remitted first-episode patients when compared to multiepisode patients. Other studies indicate that patients who have been well and stable on an antipsychotic for years have a high likelihood of relapse when their medications are discontinued.

Earlier studies by Donald Johnson suggest other advantages for maintenance antipsychotics in schizophrenia. He found that patients who relapsed when they were receiving FGAs had episodes that were less severe than those in patients who had discontinued their drugs. Drug-maintained patients were less likely to have episodes with self-destructive behaviors, violence, and antisocial acts when compared with patients who were not receiving medication. Patients who were not receiving medication were also more likely to require involuntary hospitalization. Patients who discontinued their medications ended up, on the average, receiving more total medication because the dosage needed to treat acute relapses is much higher than the dosages usually prescribed for preventing relapse. Many of these findings have since been replicated. The common denominator is that maintenance therapy with antipsychotics benefits patients with schizophrenia. Antipsychotic maintenance treatment with both FGAs and SGAs substantially reduces relapse risk and reduces hospitalization rates. Depot antipsychotics appear to be more effective for relapse prevention than oral medications. This would suggest there is a more prominent place for depot antipsychotics in relapse prevention in schizophrenia, including haloperidol and fluphenazine decanoate. One recent study pointed out that there was no difference in relapse rates between the long-acting FGA haloperidol and SGA paliperidone over 24 months, although side effects rates differed. Haloperidol decanoate caused more and more frequent EPS, while paliperidone palmitate caused more hyperprolactinemia (McEvoy et al., 2014).

Medication compliance is a serious problem for patients with schizophrenia, with up to 50 percent of patients failing to comply with recommended (oral) treatments. Importantly, only a fraction of new patients adhere to their initial antipsychotic medication after their first hospitalization, or even collect their initial prescription; the initiation of both FGA and SGA depot antipsychotics during hospitalization significantly reduces rehospitalization rates as compared to oral antipsychotics.

A number of statistics also indicate the limitations of pharmacotherapy for chronic schizophrenia. Even among treatment responders, as many as 40 percent are permanently disabled, and most of the other patients fail to recover. In addition, more than 80 percent of patients are not fully employed and require public support. In large cities, individuals with

schizophrenia comprise 14 to 50 percent of the homeless. The rate of completed suicide is about 10 percent, with suicidal ideation and suicide attempts occurring in 70 percent or more of cases and the mortality rate is increased due to greater accident proneness.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Effects of FGAs on Motor Systems

Animals, administered high doses of a dopamine receptor antagonist, develop a syndrome consisting of immobility, increased muscle tone, and abnormal postures called catalepsy. In addition, these agents decrease spontaneous motor activity in rodents. These motor effects in animals and humans are caused by dopamine receptor blockade in the striatum and the substantia nigra.

In humans, all FGAs produce EPS, in line with their antipsychotic potency, including parkinsonism (i.e., stiffness, akinesia, bradykinesia, tremor, and shuffling gait), dystonia (i.e., abrupt onset, sometimes bizarre muscular spasms affecting mainly the musculature of the head and neck), and akathisia (with objective and subjective restlessness). These effects differ from antipsychotic effects in that they have a more rapid onset, and are potentially amenable to adjunct treatment. In addition, prolonged treatment with dopamine receptor antagonists can lead to the development of late appearing abnormal movements of TD. Both short- and long-term motor effects in patients are discussed in a later section.

Neuroendocrine Effects

Antipsychotic drugs influence the secretion of hormones in the pituitary and elsewhere mainly due to their blockade of dopamine receptors. Prolactin secretion is tonically inhibited by dopamine in the pituitary gland. As a result, dopamine receptor antagonists can increase serum prolactin concentration in the clinical dosage range, in line with their antipsychotic potency. This elevation sometimes produces gynecomastia, galactorrhea, and in men erectile dysfunction. FGAs also suppress levels of luteinizing hormone (LH) and follicle-stimulating hormone (FSH). These changes can lead to amenorrhea and inhibition of orgasms in women. There is also evidence indicating that FGAs can inhibit the release of growth hormone (GH).

Table 34.17–4.
Indications for Dopamine Receptor Antagonists

Acute psychotic episodes in schizophrenia and schizoaffective disorder
Maintenance treatment in schizophrenia and schizoaffective disorders
Mania

Depression with psychotic symptoms
Delusional disorder
Borderline personality disorder
Substance-induced psychotic disorder
Delirium and dementia
Mental disorders due to a medical condition
Childhood schizophrenia
Pervasive developmental disorder
Tourette syndrome
Huntington disease

THERAPEUTIC INDICATIONS

Indications for dopamine receptor antagonists are shown in Table 34.17–4.

Acute Psychotic Episodes in Schizophrenia and Schizoaffective Disorder

Schizophrenia is usually an illness that is characterized by periods of relapse and remission. Acute psychotic episodes are usually identified as periods during which patients demonstrate the emergence or worsening of positive or disorganized symptoms of schizophrenia such as hallucinations, delusions, and disorganized thoughts.

Attention has focused on whether SGAs should be considered the preferred antipsychotics for acute episodes of schizophrenia. This issue was considered by a group of experts who met in 2001 to consider important issues in the pharmacotherapy of schizophrenia. The consensus opinion at that time was that SGAs should be favored over FGAs for first-episode patients or when a history of drug responses was unavailable. The group supported the use of FGAs over SGAs for patients who have responded well to FGAs without experiencing EPS; for patients who have responded better to FGAs than SGAs; and for patients who respond better to depot forms of FGAs.

Between 2005 and 2007, the results of large multicenter clinical trials provided important information about the relative effectiveness and side effects of SGAs and FGAs. These results challenged the view of some that the newer agents were more effective and had fewer adverse effects than their predecessors. The most important studies were an NIMH trial known as CATIE (Clinical Antipsychotic Trial of Intervention Effectiveness), an open trial from the United Kingdom National Health Service known as CUtLASS (Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study), and EUFEST (European First Episode Schizophrenia Trial), another open trial completed in Europe. These trials are important because they were each relatively large for a nonpharmaceutical drug study and they were neither sponsored nor designed by pharmaceutical company sponsors.

The first phase of CATIE compared four SGAs, olanzapine (Zyprexa; 7.5

to 30 mg per day), quetiapine (Seroquel; 200 to 800 mg per day), risperidone (Risperdal; 1.5 to 6.0 mg per day), or ziprasidone (Geodon; 40 to 160 mg per day), to one FGA, perphenazine (8 to 32 mg per day), a midpotency phenothiazine. The subjects were 1,493 patients with schizophrenia who on average were in their 40s and who had been ill for more than 14 years. The primary outcome was the time that patients remained on their assigned drug. The results, reported by Jeffrey Lieberman and coworkers in 2005, demonstrated similar efficacy to the SGAs in terms of time to discontinuation. This finding was surprising to some, but consistent with a number of meta-analyses indicating that differences between first- and second-generation drugs were small if they existed at all. However, patients with TD were excluded from the perphenazine group, which may have given this agent an advantage. Patients treated with perphenazine were more likely to discontinue their medications for EPS. Patients receiving some SGAs, particularly clozapine and olanzapine, experienced weight gain and lipid elevations. Despite modest prolactin elevation, risperidone was the best tolerated antipsychotic. The trial was too brief to compare the risk of TD among the different agents.

CUtLASS randomized 227 patients to either an SGA or an FGA. Prior to randomization, the patient's psychiatrist selected the SGA or FGA that the patient would be assigned to. The results were published in a report by Peter Jones and coworkers. There was no significant difference in scores on the Quality of Life Scale, the primary outcome instrument. Moreover, there were trends suggesting advantages for the FGAs. However, 49 percent of patients in the FGA group received sulpiride (Meresa), an agent that is not available in the United States and differs substantially from other FGAs. Therefore, the results are more meaningful in settings such as the United Kingdom where sulpiride is available. The size of the sample was too small to compare the effectiveness of individual antipsychotics.

EUFEST compared several SGAs (amisulpride, olanzapine, quetiapine, ziprasidone) with a low dose (1 to 4 mg) of the high-potency FGA, haloperidol, in the treatment of first episode schizophrenia. The pragmatic trial found high discontinuation rates (highest with haloperidol), and symptom reductions and rehospitalization rates were virtually the same in all groups.

The results of these trials do not mean that older and newer antipsychotics are equivalent in their side-effect profiles. The differences in EPS liability may have been underestimated in CATIE. First, eliminating patients with evidence of TD may have excluded those who were most vulnerable to discomforting EPS side effects. Moreover, while there were no differences in EPS ratings for parkinsonism and akathisia between perphenazine and the SGAs, it is interesting to note that the higher rate of discontinuations for EPS among the perphenazine-treated patients

suggests that ratings may have not been adequately sensitive. However, most of the studies documenting large differences between FGAs and SGAs may have used what would now be considered excessive doses of haloperidol, a drug with an incidence of EPS, which may have exaggerated the difference between the two types of agents.

The current body evidence does not support the earlier belief that the SGAs are more effective than FGAs. However, the FGAs are more likely to cause EPS and, as noted below, are probably more likely to cause TD. Perhaps FGAs should be prescribed in lower dosages than traditionally has been done. If patients are treated with FGAs, they should be regularly monitored for EPS, particularly drug-induced parkinsonism and akathisia, and they should be assessed for the evidence of TD.

Maintenance Treatment in Schizophrenia and Schizoaffective Disorders

FGAs are effective for decreasing the risk of psychotic relapse in patients who have recovered from a psychotic episode. National and international guidelines indicate that after the acute phase of treatment for a first episode, treatment should continue for at least 1 year, but there is no consensus regarding the total duration of treatment if the patient remains symptomatic. In case of multiple episodes, maintenance treatment should be continued for at least 3 to 5 years, and often indefinitely.

Mania

Dopamine receptor antagonists are effective for reducing excitement and psychotic symptoms in mania. FGAs tend to have a more rapid onset of action than antimanic drugs including lithium (Eskalith), valproic acid (Depakene) (divalproex [Depakote]), and carbamazepine (Tegretol). This observation has led to the widespread practice of combining antipsychotic and antimanic drugs during the first days of treatment of severe excited states, before the antimanic compound has its onset of action. When the antimanic compound has become effective, the dosage of antipsychotic can usually be reduced or the drug can be discontinued. Although there is only limited evidence supporting the use of FGAs for the long-term management of bipolar mood disorder, it is common clinical practice to manage some patients with severe illness with maintenance antipsychotics combined with antimanic agents. Some studies indicate that patients with mood disorders are more vulnerable to developing TD than schizophrenia patients, suggesting the importance of using FGAs as briefly as possible and only when clearly indicated. The advent of the SGAs has preempted these arguments, because some of these agents are approved for acute as well as maintenance treatment of bipolar disorder.

Depression with Psychotic Symptoms

Patients with major depressive disorder may develop psychotic symptoms such as hallucinations or delusions. These patients tend to demonstrate a more robust improvement when an antipsychotic is added to an antidepressant agent. Although most studies have focused on perphenazine, fluphenazine, and haloperidol, there is no evidence that any single dopamine receptor antagonist is more or less effective than any other for this indication, barring side-effect considerations. The added benefits of FGAs are most apparent for those patients who are tormented by severe delusions. When the psychotic symptoms have remitted, the antipsychotic should be withdrawn.

Delusional Disorder

Delusional disorder is characterized mostly by nonbizarre delusions, that is, delusions that are within the realm of possibility for the patient. Moreover, patients rarely hallucinate and their behavior is rarely bizarre. Patients with delusional disorder will often demonstrate substantial improvement when they are treated with a dopamine receptor antagonist.

Borderline Personality Disorder

Borderline personality disorder is characterized by a pervasive pattern of instability that affects mood, interpersonal relationships, and self-image. Some patients with severe forms of this disorder may experience transient psychotic states particularly when they are under stress. These states may be characterized by suspiciousness, ideas of reference, impulsiveness, and aggressiveness.

Low dosages of thiothixene and haloperidol have been found to be effective in some of these patients, although there is no evidence indicating that other FGAs are less effective. Unfortunately, most trials have been relatively brief, and it is unclear how medications affect the overall course of this disorder and how medications interact with psychotherapeutic management that these patients usually require.

Substance-Induced Psychotic Disorder

A number of drugs including amphetamines, cocaine, alcohol, and phencyclidine can cause psychotic symptoms that occur while the patient is intoxicated or during drug withdrawal. These symptoms are usually time limited and remit within hours or days depending on the concentration of drug ingested and the rate elimination of the agent FGAs may reduce these substance induced psychotic symptoms, reducing dangerous behaviors or unusual suffering. Benzodiazepines may be helpful in calming the patient until psychotic symptoms remit. Some authors recommend prescribing benzodiazepines rather than dopamine receptor antagonists for phencyclidine intoxications. In some circumstances it may be helpful to

follow the patient's untreated course in order to differentiate drug-induced syndromes from schizophrenia or bipolar disorder.

FGAs are also useful for patients with hallucinations or delusions that occur with alcoholism, particularly for chronic alcoholic hallucinosis. Benzodiazepines are more appropriate than antipsychotics for uncomplicated alcohol withdrawal, and FGAs can be prescribed as adjunct if withdrawal delirium develops. FGAs should be prescribed carefully since these drugs may increase the likelihood of seizures that occur during alcohol withdrawal.

Impulse Control Disorders

There is a long-standing, but controversial, practice of prescribing dopamine receptor antagonists for individuals with extremely poor impulse control and a propensity for violent behavior. This practice is not supported by controlled clinical trials and may be appropriate only when other measures such as the use of SSRIs, carbamazepine, lithium, or β-blockers has failed. In such instances, nonweight gain–inducing SGAs may be more appropriate than FGAs.

Delirium and Dementia

Low of doses of high-potency FGAs, especially if administered intravenously, can be useful in the treatment of the psychotic symptoms and agitation associated with delirium that may be due to organic conditions. Interactions with the anticholinergic properties of low-potency drugs may exacerbate or prolong toxic deliriums caused by anticholinergic agents.

FGAs can be useful for reducing agitation and psychotic symptoms that are related to dementia, Wernicke syndrome, and Korsakoff syndrome. A number of carefully done double-blind studies have found that FGAs are superior to placebo for treating agitated elderly patients. A review by Janicak found that 66 percent of geriatric patients with "organic" psychosis improved on an antipsychotic compared to 35 percent who improved on a placebo. Low dosages of high-potency FGAs are usually preferable for elderly patients with dementia. Low-potency medications can result in orthostatic hypotension and falling episodes. In addition, the anticholinergic effects of low-potency FGAs can worsen cognitive and memory impairments. The treatment of elderly patients is complicated since many take multiple medications and are vulnerable to drug–drug interactions. Nevertheless, the adverse effects of these drugs should be weighed against the dangers of agitation in elderly patients. Clinicians should be aware that akathisia can present itself as agitation in these patients.

Childhood Schizophrenia

Schizophrenia infrequently has its onset during childhood (before age 17). Termed early onset schizophrenia (EOS) and associated with higher rates of premorbid abnormalities, children with schizophrenia may have symptoms that are similar to those in adults with schizophrenia, with similar thought disorder and perceptual abnormalities but less systematized delusions. Clinical experience indicates that children improve with antipsychotics, and there are a number of controlled trials available to demonstrate this effect. It is particularly important to treat children with the lowest effective dosage of an antipsychotic because of increased vulnerability to side effects. Children may are more prone to developing extrapyramidal symptoms associated with FGAs, but are also more likely to experience weight gain and metabolic abnormalities with SGAs. A brief, 8-week federally funded study comparing risperidone, olanzapine and molindone did not show differences in efficacy to molindone, an FGA, but weight gain was greater with the SGAs. Antipsychotic drugs may impair learning.

Pervasive Developmental Disorder

Patients with pervasive developmental disorders may demonstrate periods of hyperactivity, screaming, and agitation with combativeness. Although there are few controlled studies for this disorder, patients tend to improve with high-potency agents. Since these agents may impair learning, there is concern that they are overprescribed in some settings.

Tourette Syndrome

Tourette syndrome is a neurobehavioral disorder that is associated with motor and vocal tics. Motor tics may involve a single muscle group or several, thus appearing as purposeful movements. Similarly, vocal tics may involve an inarticulate sound or actual words. Often the words are socially inappropriate or even obscene. Milder tics may not require pharmacotherapy. However, when the tics are disabling, an antipsychotic medication can be effective in reducing the severity of both motor and vocal tics. Although haloperidol and pimozide are the most commonly used agents, there is little evidence indicating that other agents are less effective.

Huntington Disease

Huntington disease is a genetically transmitted disease that follows an autosomal dominant type of heritability with 100 percent penetrance. The median age of onset is 35 to 40 years, but childhood onset (high number of CAG repeats) and much later onsets also occur. Once the diagnosis is made, death will follow in 15 to 20 years. Patients who eventually develop Huntington disease may display psychiatric symptoms and cognitive

impairment years before the diagnosis of Huntington disease is made. The most common psychiatric symptoms are cognitive impairment, depression, irritability, apathy, obsessive–compulsive symptoms, and suicidality. These may progress to a psychotic delusional state or sometimes are followed by manic or hypomanic episodes. Some individuals may demonstrate symptoms that closely resemble schizophrenia with delusions and hallucinations (less than 10 percent).

Antipsychotics are believed to be useful during the early stages of the illness to reduce chorea. However, double-blind studies are lacking. They also may be effective in managing irritability, paranoia, hallucinations, violent tendencies, and bizarre behaviors. High-potency agents are usually preferred, but all agents appear to be effective. Clinicians should be aware of acute EPS, particularly in the rigid forms of the disorder. There is no dosing guidance. At present it is believed that the long-term benefit does not merit their use, particularly as TD, other EPS, and akathisia may resemble Huntington disease's later stage symptomatology. Recently, the FDA-approved tetrabenazine (Xenazine, an amine depletor, like reserpine) for the treatment of chorea in Huntington disease. Tetrabenazine should not be used to prevent the occurrence of choreoathetotic symptoms as there is evidence that tetrabenazine has no neuroprotective benefits. However, dysphoria, depression, suicidality, and dysphagia have been noted. No long-term studies are available.

Parkinson Psychosis and Lewy Body Disease

FGAs are contraindicated in these conditions, because they have the potential to worsen parkinsonian syndromes and may lead to precipitous dystonic reactions.

Psychiatric Disorders due to a Medical Condition

Dopamine receptor antagonists are effective for a broad range of medical conditions that result in psychosis. This category includes endocrinopathies and psychoses associated with temporal lobe epilepsy. FGAs are also helpful for patients with psychotic symptoms that can result from steroids.

Other Indications

Antipsychotics are also effective for treating ballismus, an illness characterized by abnormal activity of the axial and proximal musculature. Because of their antiemetic effects, dopamine receptor antagonists are sometimes prescribed for severe emesis associated with chemotherapy. Patients with intractable hiccups are sometimes aided by FGAs. Metoclopramide (Reglan), a dopamine receptor antagonist, is commonly prescribed for patients with gastroesophageal reflux and delayed gastric emptying. For obvious reasons, psychosis associated with Parkinson

disease, cannot be treated with these agents.

PRECAUTIONS AND ADVERSE REACTIONS

Neurological Side Effects

The most serious side effects of dopamine receptor antagonists are neurological and are largely confined to the extrapyramidal motor system (Table 34.17–5). These EPS are, in turn, separated into acute effects that occur in the first days or weeks of treatment and chronic effects that occur after patients have received medications for months or years.

Acute Extrapyramidal Syndromes

All of the dopamine receptor antagonists are associated with EPS. Since EPS, particularly parkinsonism, is believed to be caused by decreased dopamine activity in the basal ganglia, it suggests that this toxic effect is inescapable given that blockade of D_2 receptors is associated with the antipsychotic effect of these agents. Also, this relationship explains why the potency of FGAs for causing EPS is related to their affinity for D_2 receptors in the basal ganglia. Acute EPS, which include dystonia, akathisia, and drug-induced parkinsonism, have their onset relatively soon after the initiation of antipsychotic drug treatment. They remit soon after the drugs are discontinued. In contrast, tardive EPS syndromes occur after months of treatment and commonly persist after medications are discontinued.

Table 34.17–5.
Neurological Side Effects of Dopamine Receptor Antagonists

Acute extrapyramidal syndromes
Akathisia
Acute dystonia
Drug-induced parkinsonism
Neuroleptic malignant syndrome
Chronic extrapyramidal syndromes
Tardive dyskinesia
Perioral tremor

Akathisia. Akathisia comes from the Greek word *akathizein,* which means inability to sit still. It is the most common acute manifestation of EPS and often the most distressing. One study found that 41 percent of patients treated with a dopamine antagonist had mild akathisia and an additional 21 percent had moderate to severe akathisia. The most common clinical manifestations consist of shifting the weight from foot to foot, walking on the spot, inability to keep the legs still, feelings of inner restlessness, and the shifting of body positions in a chair. Patients with akathisia may describe a compelling urge to walk or to initiate movement.

In mild cases, patients may experience a subjective feeling of restlessness, but not show increased motor activity. Akathisia may appear in the second or third day of antipsychotic treatment, but more frequently has its onset after 5 days.

Differentiating akathisia from behaviors that are related to the psychotic illness may be difficult. Feelings of restlessness or agitation may result in irritability or anxiety or even aggressive behavior. Psychotic patients with akathisia may appear more hostile or belligerent. This could lead clinicians to incorrectly conclude that the patient requires an increase in medication, which, in turn, could worsen this side effect. Akathisia is also thought to be a correlate of poor antipsychotic drug response. The use of these agents by inexperienced practitioners in nursing homes may lead to behavior deterioration.

Acute Dystonia. Acute dystonias consist of intermittent and sustained spasms of the muscles of the head and the neck, leading to involuntary movements. They can be experienced as the most frightening side effect of FGAs, particularly when patients experience the sudden onset of what may appear to be a form of paralysis.

Ten percent of drug-induced dystonias occur during the first hours of drug treatment, and 90 percent occur within the first 3 days. Common types of dystonia include opisthotonos, a rigid contraction of the back muscles with arching; retrocollis and torticollis of the neck; oculogyric crisis, a spasm in which an eye or both eyes are turned upward; macroglossia and tongue protrusion, which can lead to choking; and laryngeal dystonias. Rarely, dystonias of laryngeal or pharyngeal muscles can lead to sudden death.

Younger patients, particularly young males, are more likely to develop dystonias. They are more common when patients are treated with large dosages of high-potency FGAs. Dystonias almost always respond rapidly to antiparkinsonian medications and can usually be prevented by either pretreatment with antiparkinsonian medications or by limiting the antipsychotic dosage prescribed.

Drug-Induced Parkinsonism. Patients on dopamine receptor antagonists may experience all of the common motor symptoms of idiopathic parkinsonism including rigidity, bradykinesia, shuffling gait, and tremor. These side effects commonly occur during the first 5 to 30 days of treatment and may persist until the dosage is lowered, treatment is initiated, or FGAs are discontinued. Examination will usually reveal a positive glabella tap. The first evidence of drug-induced parkinsonism may be a diminished arm swing or decreased facial expressiveness. Risk factors for drug-induced parkinsonism include increasing age, dosage, a history of parkinsonism, and underlying basal ganglia damage (e.g., vascular insult).

In its milder forms, drug-induced parkinsonism may appear as a decrease in spontaneous gestures, a masked facial expression, apathy, unspontaneous speech, and difficulty in initiating usual activities, so-called akinesia. These symptoms may be difficult to distinguish from the negative or deficit symptoms of schizophrenia. Patients with akinesia may also appear to be depressed. The tendency of patients with akinesia to not cross their legs can be helpful in diagnosing this side effect.

Management of Acute Extrapyramidal Syndromes

Anticholinergic agents are the most effective treatment of EPS. For the acute dystonias, IM or intravenous (IV) diphenhydramine (Benadryl) (50 mg, 25 mg in children) or IM or IV benztropine (Cogentin) (0.5 to 2 mg IM or IV) or other anticholinergics alleviate this uncomfortable and often frightening side effect within minutes after the drug enters the bloodstream. Oral anticholinergic agents are usually prescribed for the management of parkinsonism and often for akathisia. Reducing the dosage of the antipsychotic is often effective, but it is common for patients to experience EPS at the minimum dosage that is needed to manage psychosis. Amantadine (Symmetrel), an antiviral agent that affects dopamine release, is also effective in reducing most EPS. This antiviral agent should not be used for this indication as a first-line agent. Direct dopamine agonists (e.g., bromocriptine [Parlodel]) and indirect dopamine agonists (e.g., levodopa [Dopar]) are not preferred, as they may aggravate psychotic symptoms.

Anticholinergics are frequently, but not always, effective for treating akathisia. β-Blockers, particularly propranolol (Inderal; 10 to 80 mg a day), have been found to be effective for treating akathisia and are preferred by many clinicians. The effectiveness of β-blockers may result from akathisia being a disturbance in the normal balance between dopamine and norepinephrine. Studies have been shown that 5-HT_{2A} antagonists can reduce haloperidol-induced akathisia.

Anticholinergic antiparkinsonian medications should be prescribed with caution in combination with low-potency FGAs and some antidepressants since the combined anticholinergic activity may lead to a number of side effects including hyperthermia and delirium. A sustained tachycardia may also occur. Further caution is warranted in the use of these agents in individuals with glaucoma, ileus, urinary retention, or with prostatic hypertrophy. Because many over-the-counter cold remedies, sedatives, and sleep aids contain anticholinergic agents and antihistamines, patients should be warned about the interaction between these agents and anticholinergic antiparkinsonian drugs.

Some controversy has arisen about the appropriateness of prescribing antiparkinsonian medications prophylactically when FGAs are started.

Reports indicate that 30 to 50 percent of patients receiving long-term treatment do not need anticholinergics. In 1990, the World Health Organization (WHO) published a consensus statement on the use of anticholinergics in patients receiving long-term antipsychotic treatment. The statement argued against the prophylactic use of these agents because of some of their side effects (e.g., the possible worsening of TD, the risk of anticholinergic toxicity, the impairment in memory, and the possible pharmacokinetic interactions with the primary treatment agent). The consensus group concluded that anticholinergics are overprescribed, that they should be given only as needed, and side effects should be treated first by lowering the antipsychotic dosage or by using other agents. However, reports indicate that anticholinergics may be helpful in managing forms of EPS that may not be obvious to the clinician and side effects. That is, milder forms of akathisia may appear as anxiety or irritability and milder forms of akinesia may appear as negative symptoms such as decreased motivation and a lack of affect. In these individuals, long-term anticholinergics may be useful for optimal pharmacotherapy.

There are other circumstances in which prophylactic antiparkinsonian medications may be helpful. Because of the discomfort and anxiety associated with acute dystonias, prophylaxis is prudent when there is a substantial risk of this side effect, for example, when young patients are treated with high-potency FGAs. Some patients and staff may interpret a dystonia as evidence that they are allergic to a particular drug, which may lead to drug refusal and an "allergy" label in the patient chart. Prudent practice may include using prophylactic antiparkinsonian drugs in high-risk patients, with gradual lowering of the dosage of this drug with regular evaluations for obvious and subtle manifestations of EPS over a 2- to 3-month period.

The availability of serotonin–dopamine antagonists provides an important alternative for patients with EPS. Symptoms of acute EPS who cannot be adequately controlled with dosage reduction or an adjunctive medication will probably improve if they are changed to an SGA.

Neuroleptic Malignant Syndrome. NMS is an uncommon but potentially fatal complication of FGAs. Its main clinical features include (1) hyperthermia; (2) severe muscular rigidity; (3) autonomic instability including hyperthermia, tachycardia, increased blood pressure, tachypnea, and diaphoresis; and (4) changing levels of consciousness. It usually presents as muscular rigidity and progresses to elevated temperature, fluctuating consciousness, and unstable vital signs. These symptoms are often associated with elevations in creatine phosphokinase (CPK) and aldolase. Less common are elevations in liver transaminases, leukocytosis, myoglobinemia, and myoglobinuria. Acute renal failure may also occur. Mortality in well-developed cases ranges from 20 to 30 percent and may be

higher when depot forms are used. More expedient early recognition of NMS has led to a substantially decreased mortality rate. Readers may find the Neuroleptic Malignant Syndrome Information Service (NMSIS) (nmsis.org) website helpful for specific questions or for their physician hotline to aid in the diagnosis of NMS.

NMS is more common when high-potency FGAs are prescribed in high doses and when dosage is escalated rapidly. The syndrome is twice as common in men as in women and is more likely to be present in younger patients. Clinicians should be concerned about any patient who demonstrates severe muscular rigidity and a rising body temperature, because early diagnosis and treatment can be life-saving.

Management of NMS. When NMS is diagnosed or suspected, antipsychotics should be discontinued and supportive and symptomatic treatment should begin. This may include treating EPS with antiparkinsonian medications, correcting fluid and electrolyte imbalances, treating fevers, and managing cardiovascular complications such as hyper- or hypotension. Dantrolene (Dantrium) may be effective for treating severe NMS. Treatment begins with IV administration of 0.8 to 2.5 mg/kg every 6 hours, with a maximum of 10 mg/kg daily. When the symptoms subside and the patient can swallow, dantrolene is administered orally in dosages of 100 to 200 mg daily. Bromocriptine can then be added at doses of 20 to 30 mg daily in four divided doses. Amantadine may be helpful if the other agents are not sufficient. The course of treatment is commonly 5 to 10 days unless a long-acting injectable agent has been used.

Although NMS is considered by some to be an extreme form of EPS, its cause and biological risk factors are unknown. Most patients will require antipsychotic medications after they recover from NMS. Low-potency drugs may be preferable for patients who have previously had NMS. However, even clozapine has been demonstrated to be a potential cause of the syndrome. At this time there are no data available to indicate whether clozapine or other serotonin–dopamine antagonists are associated with a lower risk of NMS than SGAs.

CHRONIC EXTRAPYRAMIDAL SYNDROMES

Tardive Dyskinesia

TD is a movement disorder that can follow chronic treatment with antipsychotic medications. Patients with TD may have any or all of a number of abnormal movements, including mouth and tongue movements (lip smacking, sucking, and puckering) as well as facial grimacing. Other movements may include irregular movements of the limbs, particularly choreoathetoid-like movements of the fingers and toes; and slow, writhing movements of the trunk. Although seriously disabling TD is uncommon, a

small proportion may affect walking, breathing, eating, and talking.

The differential diagnosis of TD includes a number of disorders of the basal ganglia (including Huntington disease, Wilson disease, and Sydenham chorea), striatal hypercalcifications, hyperthyroidism, hypoparathyroidism, tardive Tourette disorder, and dyskinesias related to other drugs, such as L-dopa and amphetamines. Abnormal movements were also described in patients with schizophrenia before the introduction of antipsychotic medications.

Although antipsychotic medications will frequently suppress or reduce TD, dose increases are only appropriate when the movements are severe and incapacitating. If patients have their dosages of antipsychotic medication either decreased or discontinued, abnormal movements may worsen temporarily, or appear for the first time and then diminish, termed withdrawal emergent dyskinesias. Antiparkinsonian medications will often worsen TD. However, when patients suffer from combined TD and acute EPS, prescribing antiparkinsonian medication is often unavoidable.

Although the biological mechanism underlying TD is still controversial, it is likely that the disorder is related to an increased sensitivity of dopamine receptors in the basal ganglia. This is supported by the observation that drugs that block dopamine, such as antipsychotic medications, suppress TD, whereas dopamine agonists worsen the disorder. These observations are consistent with findings from preclinical laboratories indicating that blocking dopamine receptors can result in their upward regulation.

Approximately 10 to 20 percent of patients treated with FGAs for more than a year develop TD. In chronically institutionalized patients the prevalence of TD is between 15 and 20 percent, but increases with each year of treatment. Recent prospective studies indicate that approximately 4 to 5 percent of patients who are undergoing antipsychotic drug treatment will develop TD each year, with this risk markedly compounded in elderly individuals whose first treatment with antipsychotics is after age 65. Certain populations are at a greater risk than others for developing TD. Elderly women are particularly vulnerable. Moreover, TD in the elderly is less likely to remit when FGAs are discontinued. Patients with affective disorders may also be at a greater risk of TD when they are treated with dopamine receptor antagonists. Other possible risk factors for TD are the dosage of antipsychotic medications and the duration of treatment. There is also evidence indicating that the presence of acute EPS increases the risk of developing TD later.

TD does not appear to be a progressive disorder for most patients. The disorder seems to develop rather rapidly and to then stabilize, and often improves with time. The consensus of a number of long-term studies is that the severity of TD in most patients improves even if their FGAs are continued. Moreover, this improvement can be clinically meaningful in

some patients. Individuals who are most likely to remit are those with a recent onset and those with an onset prior to age 40. When FGAs are discontinued, a substantial proportion of patients with TD will demonstrate a remission. Reinstatement of antipsychotics does not mean an automatic reoccurrence. This is particularly likely for those with a rather recent onset of TD. Sometimes, this remission will take several months to occur.

Perioral Tremor

Perioral tremor is an uncommon side effect of dopamine receptor antagonists that appears after months or years of drug treatment. The peculiar tremors of the mouth have led to this disorder being called rabbit syndrome.

Management of TD. Patients receiving FGAs for more than 6 months should have regular and systematic evaluations for TD every 6 months. Although instruments such as the Abnormal Involuntary Movement Scale (AIMS) were originally developed as research tools, they also provide a structured system for evaluating and recording abnormal movements.

Once TD is diagnosed, the clinician should evaluate the need for continued antipsychotic treatment. This decision requires consideration of the dangers of both the psychotic illness and TD. It is a decision that, whenever possible, should be made in collaboration with the patient and his or her family. The results of this discussion should be documented in the patient's medical record. For most patients with chronic schizophrenia, the best decision will probably involve continuing the antipsychotic. If this is the case, there is some evidence that treating patients with the lowest effective dosage may minimize the risk of TD. Moreover, prescribing the lowest effective dosage of antipsychotic medication may also decrease the patient's risk of developing the disorder. Patients with severe TD may also be candidates for a trial of clozapine or other SGAs. Most clinical reports indicate that clozapine can suppress TD.

The American Psychiatric Association task force on TD recommended a number of steps for the prevention and management of the disorder, including (1) establishing objective evidence that antipsychotic medications are effective; (2) using the minimum effective dosage for long-term treatment; (3) exercising caution with children, the elderly, and patients with mood disorders; (4) examining patients on a regular basis for evidence of dyskinesia and noting the results of the examination in the medical record; (5) if TD is diagnosed, considering alternatives to antipsychotics, obtaining informed consent, and considering dosage reduction; and (6) if the dyskinesia worsens, considering the discontinuation of the antipsychotic, switching to a different drug, or considering a trial of clozapine.

In addition, there is substantial data indicating that clozapine and other

SGAs are less likely to cause the disorder than other antipsychotic medications, and, therefore, it is promising for the long-term management of patients who have developed TD. In addition, studies using animal models indicate that clozapine does not result in the increased sensitivity of striatal D_2 receptors, which is associated with typical neuroleptics and has been theorized to result in TD.

Table 34.17–6.
Behavioral Side Effects of Dopamine Receptor Antagonists

Dysphoric adverse effects
Secondary negative symptoms
Cognitive impairments

Current evidence suggests that SGAs are associated with a lower risk of TD than the FGAs. This has been proposed since the decreased tendency of these agents to cause acute EPS would predict a lower risk of TD. An accumulation of controlled and uncontrolled studies led a consensus group of experts in the pharmacotherapy of schizophrenia to conclude that SGAs were associated with a lower risk of TD. Because most of these studies were not designed to focus on TD, this remains a controversial issue. A Cochrane review did not find sufficient evidence to conclude that there was an advantage for SGAs, while other meta-analyses concluded that the available evidence supported an advantage. Still, despite high penetration of SGAs into clinical practice, the incidence and prevalence of TD appears relatively unchanged since the 1980s. Since it would be difficult to perform an adequate study of sufficient duration to prove that there is a difference, this issue remains unresolved.

BEHAVIORAL SIDE EFFECTS

FGAs may cause a number of subjective and behavioral effects that are not accompanied by obvious motor effects (Table 34.17–6). Subtle parkinsonian side effects may lead to decreases in emotional expressiveness, apathy, and depression. These symptoms can easily be confused with the negative symptoms of schizophrenia. In addition, subjective akathisia may cause feelings of restlessness that can cause anxiety or even agitation. High dosages of FGAs may also lead to cognitive impairment and attentional deficits. These behavioral symptoms can often be managed by decreasing the antipsychotic dosage.

CARDIOVASCULAR EFFECTS

Cardiac Side Effects

Animal studies have shown that antipsychotic agents may decrease cardiac contractility, disrupt enzyme activity in cardiac cells, decrease tissue levels

and increase circulating levels of catecholamines, and prolong atrial and ventricular conduction time and refractory periods. Low-potency agents, particularly chlorpromazine and thioridazine, can induce electrocardiogram (ECG) abnormalities, such as prolongation of the QT and PR intervals, blunting of the T waves, and depression of the ST segment. QT prolongation is a particular concern since this can lead to a ventricular arrhythmia (known as torsades de pointes), which can progress to ventricular fibrillation and sudden death. There are indications that these arrhythmias have occurred in patients treated with thioridazine and chlorpromazine. A recent study monitored the QT interval in patients treated with a number of SGAs as well as haloperidol and thioridazine. The results led the FDA to conclude that thioridazine and mesoridazine (which is metabolized to thioridazine) were associated with substantial QTc prolongation and a risk of torsades de pointes, with subsequent high risk of sudden death. As a result, these agents are recommended for patients who are refractory to other agents. Clinicians should consider obtaining an ECG before and during treatment with thioridazine and mesoridazine. These agents should not be prescribed for patients with prolonged QTc intervals or those who are receiving other agents that have a tendency to prolong the QTc interval or that decrease the metabolism of such agents. Attempts are under way to identify patients at risk of QTc prolongation through pharmacogenomics. Recently the FDA and Johnson & Johnson have informed health care providers of a revision to the prescribing information for haloperidol, haloperidol decanoate, and haloperidol lactate to include a new cardiovascular subsection highlighting the risk of sudden death, QT prolongation, and torsades de pointes, particularly when given intravenously or at dosages higher than those recommended. Pimozide is also associated with QTc problems (and seizures) when given in doses of 20 mg or higher. Thioridazine (Mellaril) has the greatest risk of QTc prolongation among all currently available antipsychotics.

Elderly patients and those with dementia are at a greater risk of sudden cardiac death, presumably because of slower metabolism (higher blood levels) and concomitant polypharmacy for multiple aging related illnesses and symptoms. This polypharmacy may lead to drug–drug interactions.

OTHER DOPAMINE RECEPTOR ANTAGONISTS EFFECTS

Orthostatic (Postural) Hypotension

Low-potency dopamine receptor antagonists (e.g., chlorpromazine and thioridazine) may cause orthostatic hypotension particularly when these agents are first prescribed. This side effect can be dangerous when it results in fainting or falls. Both elderly patients and individuals with impaired balance should be monitored carefully for orthostatic changes. Since tolerance may develop gradually, patients should be initially treated with a

relatively low dose of these agents (e.g., 25 mg orally twice daily) and the dose should be titrated upward over several days until a target dose is reached. When appropriate, patients should be warned of this side effect and told to sit or stand up slowly, to sit with their feet outside the bed when arising, to wait for as long as a minute, and to sit or lie down if they feel faint, particularly during the first 2 weeks of treatment. Some patients, particularly the elderly, may have persisting orthostatic hypotension and, as a result, and may benefit from a change to a high-potency antipsychotic. Younger patients usually adapt to hypotension and may be asymptomatic, even though their blood pressure is low.

Orthostatic hypotension is more likely when low-potency agents are administered intramuscularly. Thus, clinicians should consider measuring the patient's blood pressure (both lying and standing) before and 1 to 2 hours after the first dose and before subsequent doses during the first days of treatment.

If hypotension does occur, even with higher-potency agents, it can usually be managed by having the patient lie down and instructing the individual to arise very slowly. Some patients may benefit from stockings, volume expansion, or pressor agents. The use of epinephrine paradoxically worsens antipsychotic-induced hypotension through β-stimulation and is, therefore, contraindicated. Pure α-adrenergic agents, such as metaraminol (Aramine), and norepinephrine, are the drugs of choice for this condition.

Sudden Unexplained Death

In spite of reports that dopamine receptor antagonists have been associated with sudden unexplained death, no evidence indicates that the incidence of such deaths changed in psychiatric patients after the introduction of antipsychotics. Similarly, no difference in sudden unexplained death is seen in patients on high-potency versus low-potency drugs. An American Psychiatric Association task force report concluded:

[Al]though a relationship between the use of antipsychotic drugs and sudden death has not been firmly established, it has also not been disproved. From a neurocardiologic perspective, these drugs have the potential for both increasing and decreasing the risk of sudden death. Ultimate outcome is probably determined by a multitude of interacting factors, and the role played by the drug in a given individual is difficult, if not impossible to determine.

Reports of sudden death in patients treated with antipsychotics could be explained by cardiac arrhythmias such as torsades de pointes secondary to QTc prolongation or other causes. For example, seizures, asphyxiation, heat stroke, malignant hyperthermia, and NMS could result in relatively sudden death. Individuals who may be more vulnerable to sudden death include the elderly, physically debilitated individuals, and agitated patients (particularly those who are physically restrained). Increasing evidence of

QTc prolongation caused by antipsychotic agents is surfacing, presumably responsible for torsades de pointes.

Gastrointestinal Tract

The most common GI side effects of dopamine receptor antagonists are related to their peripheral anticholinergic effects. These include dry mouth, constipation, occasional diarrhea, and urinary retention. Nausea and vomiting are less common. Anticholinergic effects are most prevalent with low-potency FGAs such as chlorpromazine or thioridazine and are uncommon with high-potency drugs. These effects can be exacerbated when other anticholinergic drugs such as tricyclic antidepressants are combined with a low-potency antipsychotic.

Patients who experience a severe dry mouth should be advised to rinse their mouth frequently and to chew sugarless gum or candy. Sugar-containing candies may lead to fungal infections in the mouth and increase the risk of dental caries. Constipation can be managed with stool softeners or laxatives. Clinicians should manage severe constipation carefully since it can progress to paralytic ileus, a potentially fatal condition.

Liver

Dopamine receptor antagonists may cause transient abnormalities in liver function tests. These are seldom of serious medical concern. Obstructive or cholestatic jaundice is a serious condition that was commonly associated with chlorpromazine treatment in the past. It is currently very uncommon.

Blood

Dopamine receptor antagonists may be associated with temporary and transient depression of leukopoieses. Mild leukopenia may occur for prolonged periods, although it is more common for the leukocyte to eventually return to normal. Agranulocytosis is rare with dopamine receptor antagonists, occurring in less than 1 in 10,000 cases. Thrombocytopenia and pancytopenia may also occur.

Kidney and Urinary Function

Chlorpromazine has weak diuretic effects. The anticholinergic effects of low-potency drugs may lead to urinary retention.

Skin and Eye Effects

Cutaneous reactions seen during the first weeks of antipsychotic drug treatment include urticarial, maculopapular, petechial, and edematous reactions. These reactions usually improve when drugs are discontinued, and sometimes clear even when drugs are continued. Patients receiving

low-potency FGAs, particularly chlorpromazine, may develop photosensitivity reactions consisting of severe sunburn or rash. Patients should be instructed to avoid direct sunlight or to use sunscreens.

Low-potency FGAs, especially chlorpromazine, are associated with an uncommon discoloration of the skin. Skin areas that are exposed to sunlight, particularly the face and neck develop blue–gray metallic discoloration. This skin reaction is usually associated with long-term treatment with high-drug dosages. Changing to a high-potency drug will usually lead to gradual improvement.

Patients receiving long-term treatment with chlorpromazine may develop granular deposits in the anterior lens and posterior cornea. These deposits, visualized on slit lamp examination, seldom affect the patient's vision. Changing the patient to another drug will usually result in a gradual improvement in the condition.

High doses of thioridazine—above 1,000 mg daily—can result in altered retinal pigmentation (retinitis pigmentosa). This condition can lead to serious visual impairment or blindness. Moreover, this condition may not remit when thioridazine is discontinued. As a result, thioridazine should not be prescribed at doses over 800 mg daily.

Endocrine Effects

Dopamine receptor antagonists elevate circulating levels of prolactin. In women, elevated prolactin can lead to breast enlargement, galactorrhea, and irregular menses, including anovulatory cycles and infertility, menses with abnormal luteal phases, or frank amenorrhea and hypoestrogenemia. The elevation of prolactin associated with FGAs may also suppress testosterone in men. Prolactin elevation occurs as a direct result of D_2 blockade in the tuberoinfundibular dopamine pathway and the pituitary. Since prolactin is under tonic inhibitory control by dopamine, blockade of dopamine receptors will elevate prolactin. Some patients, but not all, develop tolerance to the prolactin-elevating effects of these drugs after several weeks. Dosage reduction may relieve symptoms that are related to prolactin elevation.

Effects on Sexual Function

FGAs can have important effects on sexual functioning in both men and women. In men these agents can result in decreased libido, and ejaculatory and erectile disturbances. Erectile dysfunction has been reported to occur in 23 to 54 percent of men receiving an antipsychotic. This can be an important cause of medication noncompliance in men, and clinicians should inquire about it. Retrograde ejaculation has been reported with some FGAs, particularly thioridazine. In addition, these agents can affect ejaculatory latency and the quantity of ejaculate. Since sexual functioning is

affected by dopamine, serotonergic, and adrenergic systems, it is unclear how sexual dysfunction resulting from FGAs is mediated. Dosage reduction may be helpful in some individuals.

Women who receive dopamine receptor antagonists may experience decreased libido, anorgasmia, and decreased lubrication. Although reports indicate that a number of adjunctive medications have been reported, there is little evidence that any are effective. As with men, reducing the dosage of the antipsychotic may be helpful. Presumably agents with high 5-HT_{2A} blockade should be able to counter these effects.

Pregnancy

Although dopamine receptor antagonists cross the placenta, there is very little evidence to suggest that there is an association between prenatal exposure to an antipsychotic and an increased incidence of congenital malformations. A meta-analysis by Altshuler and colleagues found that the baseline incidence of congenital anomalies is 2 percent and the incidence with antipsychotics is 2.4 percent. This suggests that there is a very small additional risk to the fetus associated with these agents. This meta-analysis found no evidence that high-potency drugs increase the risk to the fetus.

Prenatal exposure to dopamine receptor antagonists in rats may lead to a decreased number of brain dopamine receptors. Other animal studies indicate that prenatal haloperidol exposure may raise cholesterol levels in offspring. It is unclear if these animal findings indicate risks in humans.

In developing a treatment of pregnant patients with schizophrenia, the relatively low risk of congenital abnormalities associated with antipsychotics should be weighed against the dangers of psychotic episodes. Clinicians may attempt to discontinue these medications during the first trimester if this is feasible and should consider carefully whether the risks of prescribing these drugs during the remainder of a patient's pregnancy are justified by the likely benefits. If patients have a history of relapse when FGAs are discontinued, the drugs are relatively safe. If FGAs are prescribed, high-potency compounds are probably safer.

Lactation

Dopamine receptor antagonists are secreted in low concentrations with lactation, suggesting that patients should be advised against breastfeeding during antipsychotic treatment.

Seizure Threshold

Dopamine receptor antagonists tend to lower the seizure threshold. As a result, they are associated with an increased risk of seizures. This risk is usually clinically important in individuals with a prior history of seizures. Low-potency drugs have a greater tendency to lower the seizure threshold.

Therefore, high-potency FGAs may be preferable for seizure-prone individuals.

Temperature Regulation

Some patients become poikilothermic, that is, they become hypothermic in the winter and overheat in the summer, with chlorpromazine and other low-potency agents. This indicates that special precautions should be taken for these patients during exposure to extremes in temperature.

Overdose

Overdoses with FGAs are usually characterized by exaggerations of their usual side effects. Thus, high-potency drugs are associated with severe extrapyramidal reactions, including dystonias and severe rigidity, as well as hypotension and sedation. Lower-potency drugs are more likely to be associated with CNS depression, sedation, anticholinergic effects, and hypotension. The clinical picture may also include agitation, restlessness, convulsions, fever, autonomic reactions such as dry mouth and ileus, ECG changes, and cardiac arrhythmias. In severe overdoses with FGAs the pupils are mydriatic, deep tendon reflexes are decreased and areflexia may occur, tachycardia and hypotension are present, and the EEG shows diffuse slowing and low voltage. This picture may progress to delirium and coma, with respiratory depression and hypotension.

The management of FGA overdose is symptomatic and supportive since overdoses with these agents are seldom lethal. However, the outcome may be unfavorable when these agents are combined with others, particularly CNS depressants such as alcohol, barbiturates, and benzodiazepines. If the overdose is recent, gastric lavage is helpful. If the patient can swallow, activated charcoal is indicated. Emetics should not be administered since their effects may be attenuated by the antipsychotic and because they are associated with a risk of aspiration pneumonia. Aspiration may be a serious risk if the patient develops a dystonic reaction of the head or neck. Forced diuresis and hemodialysis are not effective because FGAs are highly protein bound and lipophilic.

When hypotension occurs, standard measures for managing circulatory shock should be implemented. The hypotension may occur when, during α-receptor blockade, β-receptors are unopposed. Therefore, the clinician should consider norepinephrine or dopamine. Epinephrine is contraindicated.

Because of their quinidine-like effects, overdoses with thioridazine and mesoridazine may lead to heart block and ventricular fibrillation. Since these agents may have long half-lives, patients recovering from substantial overdoses may require cardiac monitoring for several days.

Tolerance, Dependence, and Withdrawal Symptoms

Some tolerance develops to the sedative, the anticholinergic, the hypotensive, and some other effects of FGAs. Tolerance does not develop to the antipsychotic effects. Some patients experience a decrease in EPS over time, suggesting tolerance to the motor side effects of these agents. Dopamine receptor antagonists are not addictive. As a result, most patients can be withdrawn from these agents without difficulty. Rapid withdrawal of low-potency FGAs can lead to symptoms related to their anticholinergic effects. These symptoms are commonly manifestations of "cholinergic rebound" including nausea, vomiting, and diarrhea. Other reported withdrawal symptoms include headache, sweating, restlessness, and insomnia. These symptoms tend to occur within 2 to 3 days of drug withdrawal and may last as long as 2 weeks. Most patients who are withdrawn from high-potency FGAs do not show withdrawal symptoms. These observations suggest that patients who are withdrawn from low-potency agents should be withdrawn gradually over 1 to 2 weeks or longer.

DRUG INTERACTIONS

Clinicians should be aware of several types of interactions (Table 34.17–7). These include drugs that interact with the metabolism of FGAs; age and gender, which may affect metabolic pathways; medical conditions, which may affect drug protein binding, absorption, and fecal or renal clearance; individuals who are dehydrated or malnourished or who have impaired liver and kidney function and are at risk of altered sensitivity and metabolism.

Table 34.17–7.
Factors Influencing the Pharmacokinetics of Antipsychotics

Age	Elderly patients may demonstrate reduced clearance rates
Medical condition	Decreased hepatic blood flow can reduce clearance Hepatic disease can decrease clearance
Enzyme inducers	Carbamazepine, phenytoin, ethambutol, barbiturates
Clearance inhibitors	Include selective serotonin reuptake inhibitors (SSRIs), tricyclic antidepressants, cimetidine, β-blockers, isoniazid, methylphenidate, erythromycin, triazolobenzodiazepines, ciprofloxacin, ketoconazole
Changes in binding protein	Hypoalbuminemia can occur with malnutrition or hepatic failure

Adapted from Ereshefsky L. Pharmacokinetics and drug interactions: update for new antipsychotics. *J Clin Psychiatry*. 1996;57 Suppl 11:12–25.

Metabolism

FGAs may have clinically important interactions with drugs that are

metabolized by the same cytochrome P450 enzymes. For the conventional antipsychotics, particularly haloperidol, perphenazine, thioridazine, and others, this involves CYP 2D6. Since some SSRIs—particularly fluoxetine and paroxetine—are potent inhibitors of this enzyme, plasma levels of the dopamine antagonist are likely to be increased when an SSRI is added. This may result in an increase in a worsening of EPS. Other drugs that are substrates for 2D6, including heterocyclic antidepressants, β-blockers, and cimetidine (Tagamet), may also increase antipsychotic plasma concentrations.

Other drugs, particularly phenytoin (Dilantin), carbamazepine, and barbiturates, can reduce antipsychotic plasma concentrations by increasing the metabolism of the dopamine antagonist. Reports indicate that carbamazepine can reduce the plasma concentration of haloperidol by 50 percent. Cigarette smoking can also reduce the plasma levels of FGAs. Phenobarbital (Bellatal) has been shown to increase urinary excretion and to decrease the plasma levels of chlorpromazine.

There are other common interactions that should be of concern to clinicians. Antacids, activated charcoal, cimetidine, cholestyramine (Questran), kaolin, and pectin can decrease the absorption of antipsychotic from the gut. Increased absorption may occur with drugs that decrease GI motility such as digoxin (Digitek) and steroids.

Coffee and tea may precipitate FGAs. Although this has been demonstrated in vitro, its clinical significance is unclear. Trifluoperazine and chlorpromazine have been reported to precipitate with lithium citrate. The clinical relevance is unclear.

Antipsychotic drugs antagonize the effects of dopamine agonists or levodopa when these drugs are used to treat parkinsonism. Chlorpromazine, haloperidol, and thiothixene can block the antihypertensive effects of guanethidine (Ismelin). FGAs may also enhance the effects of CNS depressants such as analgesics, anxiolytics, and hypnotics.

FGAs may enhance the effects of oral hypoglycemics. Phenothiazines enhance the effects of oral anticoagulants, although haloperidol may decrease these effects. Cigarette smoking may lower dopamine antagonist blood levels by enzyme induction.

Side-Effect Interactions

Narcotics, captopril (Capoten), trazodone (Desyrel), heterocyclic antidepressants, hydralazine (Apresoline), and minoxidil (Rogaine) may increase hypotension. The combination of a low-potency dopamine antagonist and another sedating drug may lead to excessive sedation. Combining low-potency FGAs with other anticholinergic compounds can lead to peripheral anticholinergic symptoms such as dry mouth and urinary retention and central effects such as cognitive deficits and delirium

(atropine psychosis). Patients should be warned about the use of over-the-counter medications that contain anticholinergics.

Amoxapine (Asendin), fluoxetine, chloroquine (Aralen), and lithium may increase extrapyramidal symptoms. Hydroxyzine (Vistaril) potentiates thioridazine and mesoridazine toxicity. Diuretics and corticosteroids enhance cardiotoxic side effects through hypokalemia, and quinidine enhances the quinidine-like effects of these agents.

LABORATORY INTERFERENCES

Chlorpromazine and perphenazine may give false-positive or false-negative results on pregnancy tests. These drugs may also give inaccurate reports of bilirubin (with reagent test strips) and urobilinogen (with Ehrlich's reagent test). Phenothiazines may produce false-positive results in tests for phenylketonuria. Phenothiazines may also interfere with 17-ketosteroids (with the Haltorff-Koch modification of the Zimmermann reaction) and 17-OH corticosteroids (with the modified Glenn-Nelson technique).

DOSAGE AND ADMINISTRATION

Dopamine receptor antagonists are available in a wide range of formulations and dose sizes (Table 34.17–8).

TREATMENT OF ACUTE PSYCHOSIS

Initiation of Treatment

Treatment with an FGA is usually initiated soon after a diagnosis is made. Whenever possible, patients should receive a psychiatric evaluation including a history and physical examination and a laboratory evaluation that includes complete blood cell count; blood electrolytes; glucose, liver, renal, and thyroid function tests; and other appropriate evaluations (including ECG, pregnancy tests, and screens for alcohol and drugs). The results of this evaluation can be useful in confirming the diagnosis and establish a pretreatment baseline. Under most conditions, it is not necessary to delay treatment while awaiting the results from these tests. Moreover, because there are few contraindications to receiving FGAs, these evaluations can be postponed when patients refuse an evaluation or when treatment is urgent.

Choice of Agent

There is no evidence for differential effectiveness of FGAs, with the exception of clozapine, which is the most effective antipsychotic available. Tolerability may differ between FGAs, influenced by their potency. Low-potency FGAs may not have a higher EPS risk than their high-potency counterparts, but may not be as effective at low doses, and can be quite sedating at high doses. Since thioridazine and mesoridazine have been

associated with prolongation of the QT interval of the ECG, these agents should be reserved for patients who fail to respond to all other antipsychotics. The choice among the other FGAs is based on prior responses, both objective improvement and the patient's subjective response to the drug, available routes of administration and dosage forms, side-effect profiles, and cost.

Table 34.17–8.
Dopamine Receptor Antagonist Preparations

Drug Name	Tablets (mg)	Capsules (mg)	Solution	Parenteral	Rectal Suppositories (mg)
Chlorpromazine	10, 25, 50, 100, 200	30, 75, 150, 200, 300	10 mg/5 mL, 30 mg/mL, 100 mg/mL	25 mg/mL	25, 100
Chlorprothixene	10, 25, 50, 100	—	100 mg/5 mL (suspension)	12.5 mg/mL	—
Droperidol	—	—	—	2.5 mg/mL	—
Fluphenazine	1, 2.5, 5, 10	—	2.5 mg/5 mL, 5 mg/mL	2.5 mg/mL (IM only)	—
Fluphenazine decanoate	—	—	—	25 mg/mL	—
Fluphenazine enanthate	—	—	—	25 mg/mL	—
Haloperidol	0.5, 1, 2, 5, 10, 20	—	2 mg/mL	5 mg/mL (IM only)	—
Haloperidol decanoate	—	—	—	50 mg/mL, 100 mg/mL (IM only)	—
Loxapine	—	5, 10, 25, 50	25 mg/5 mL	50 mg/mL	—
Molindone	5, 10, 25, 50, 100	—	20 mg/mL	—	—
Perphenazine	2, 4, 8, 16	—	16 mg/5 mL	5 mg/mL	—
Pimozide	2	—	—	—	—
Prochlorperazine	5, 10, 25	10, 15, 30	5 mg/5 mL	5 mg/mL	2.5, 5.0, 25.0
Thioridazine	10, 15, 25, 50, 100, 150, 200	—	25 mg/5 mL, 100 mg/5 mL, 30 mg/mL, 100 mg/mL	—	—
Thiothixene	—	1, 2, 5, 10, 20	5 mg/mL	5 mg/mL (IM only), 2 mg/mL (IM only)	—
Trifluoperazine	1, 2, 5, 10	—	10 mg/mL	2 mg/mL	—

IM, intramuscular.

The introduction of the SGAs (risperidone, olanzapine, quetiapine, ziprasidone, aripiprazole, iloperidone, lurasidone, zotepine, iloperidone, amisulpride at this time) has complicated the decision-making process. These agents, in themselves quite heterogeneous, have different side-effect profiles when compared with (especially high potency) FGAs. The most important advantage is that most SGAs cause substantially less EPS than high-potency FGAs. Patients who experience EPS at dosages of FGAs needed to treat their psychosis may or may not experience EPS with an effective dosage of an SGA. This was evaluated in a phase of the CATIE trial in which patients who discontinued perphenazine were subsequently randomized to risperidone, olanzapine, or quetiapine.

There is also evidence from a number of well-designed clinical trials indicating that patients who have only partial responses to an FGA due to lack of efficacy or side effects may improve if changed to an SGA. The evidence is compelling for clozapine, but there are also some reports supportive of other SGAs. However, most of the studies reporting this advantage of newer drugs were carried out in the late 1980s and 1990s when patients had only been exposed to FGAs. More recently, there are clinical reports of patients who had only received SGAs who improved when they were changed to an FGA. Moreover, cost considerations have become more important, especially in the United States. Many of the SGAs are patent protected and are considerably more expensive than SGAs which are usually available as generics. Risks and benefits need to be considered carefully when choosing antipsychotics, especially in light of the different TD risks with FGAs versus SGAs.

An additional advantage of some SGAs is that they are less likely to elevate prolactin. As a result, these agents may be helpful for women who experience galactorrhea and irregular menses with FGAs.

High-potency FGAs, such as haloperidol and fluphenazine, are more commonly prescribed than low-potency compounds. High-potency drugs are more strongly antidopaminergic and thus more likely to cause EPS than the low-potency medications such as chlorpromazine or thioridazine. However, EPS in some individuals may be easier to manage than the sedation and orthostatic hypotension associated with low-potency medications. High-potency medications can more readily be administered intramuscularly, since they seldom cause problems with blood pressure. In addition, because of sedation, orthostatic hypotension, and lethargy, the dosage of low-potency FGAs should be gradually titrated upward, whereas patients on high-potency medications can usually be treated with a therapeutic dosage within a day or two. Midpotency drugs such as loxapine and molindone are options for patients who have difficulty tolerating the side effects of high- or low-potency drugs.

Route of Administration

FGAs can be administered orally, or as short- or long-acting injectables (Table 34.17–9). Once patients are stabilized on a dosage, most patients can be managed with a single daily dose. Short-acting injectables are administered intramuscularly for calming acutely anxious patients or for administering a drug for patients who refuse oral medication. The clinical effects of an IM dose may appear as soon as 15 minutes after the injection. Since the FGAs are most effective when patients have reached a steady state for several days or even weeks, patients should be changed to oral administration as soon as possible.

FGAs can also be administered as long-acting injectables. In the United States, only two FGAs, haloperidol and fluphenazine, are available in long-acting formulations. Long-acting antipsychotics are usually prescribed in stabilized patients during maintenance treatment, after tolerance to the oral formulation has been established. It takes several months to reach steady state and oral supplementation may be necessary during that time. For these reasons, long-acting depot drugs better suited for long-term relapse prevention. They are cost-effective as well, but their TD risk may be higher.

Dosage

Patients demonstrate marked differences in the dosage of a dopamine receptor antagonist that is optimal. For example, some patients will find it difficult to tolerate the side effects from 1 mg of haloperidol, whereas others will tolerate 50 mg without suffering any side effects. Finding the best dosage for a dopamine receptor antagonist is both important and difficult. The importance stems from the tendency of FGAs to cause side effects at their effective dosages. Oftentimes the clinician must weigh the therapeutic advantages of a particular dosage against discomforting or disabling side effects. As many of the FGAs were developed prior to the required rigorous pharmacological methodology of today, a minimally effective or lowest dose has usually not been established. As a result, the most reasonable practice may be to treat a patient with a dosage that is likely to be effective and to wait an adequate amount of time, 4 to 6 weeks at most, to determine if the patient is responding.

Table 34.17–9.
Dopamine Receptor Antagonists

	Name			Adult Dose Range (mg/day)	
Class	**Generic or Chemical**	**Trade**	**Manufacturer**	**Acute**	**Maintenance**
Phenothiazine					
Aliphatic	Chlorpromazine	Thorazine	SmithKline Beecham	100–1,600	50–400 oral

				oral 25–400 IM	
Piperazine	Prochlorperazine	Compazine	SmithKline Beecham	15–200 oral 40–80 IM	15–60 oral
	Perphenazine	Trilafon	Schering	12–64 oral 15–30 IM	8–24 oral
	Trifluoperazine	Stelazine	SmithKline Beecham	4–40 oral 4–10 IM	5–20 oral
	Fluphenazine	Prolixin	Apothecon	2.5–40.0 oral 5–20 IM	1.0–15.0 oral 12.5–50.0 IM (decanoate or enanthate, weekly or biweekly)
Piperidine	Thioridazine	Mellaril	Sandoz	200–800 oral	100–300 oral
	Mesoridazine	Serentil	Boehringer Ingelheim	100–400 oral 25–200 IM	30–150 oral
Nonphenothiazine					
Butyrophenones	Haloperidol	Haldol	McNeil	5–20 oral 5–20 IM	1–5 oral 25–200 IM (decanoate, monthly)
Thioxanthenes	Chlorprothixene	Taractan	Roche	75–600 oral 75–200 IM	50–400
	Thiothixene	Navane	Roerig	6–100 oral 8–30 IM	6–30
Dibenzoxazepines	Loxapine	Loxitane	Lederle	20–250 20–75 IM	20–100
Diphenylbutylpiperidines	Pimozide	Orap	Gate	0.5–20	0.5–5.0
Benzamides	Sulpiride	Dogmatil	Sesif (France)	600–1,800	

IM, intramuscular.

Dosage comparison studies provide guidance to clinicians who are making decisions regarding the dosages of FGAs for most patients with schizophrenia or schizoaffective disorder. Doses below 300 mg of chlorpromazine (or 5 mg of fluphenazine or 2 mg of haloperidol) are likely to be too low for many psychotic patients. At the same time doses above 1,000 mg of chlorpromazine (or 10 mg of haloperidol or 20 mg of fluphenazine) are seldom necessary and may lead to substantial side effects and even fewer antipsychotic effects. In other words, most patients will respond to doses between 5 and 20 mg of fluphenazine or 2 and 10 mg of haloperidol or equivalent amounts of other FGAs. Studies that focused on haloperidol suggest that doses below 8 mg are adequate for most patients and that these doses are associated with less EPS than 20 mg doses. First-episode patients should be treated with lower doses, and children and elderly patients should be treated with much lower doses.

The issue of dosing was recently highlighted by the FDA when it added a "Black Box" warning to the use of SGAs in elderly patients with dementia (increased mortality risk). Similar increased mortality risk has been associated with FGAs. Whether this is secondary to higher blood levels and the use of polypharmacy in the elderly, exposing them to an increased risk of QTc prolongation, which may lead to torsades de pointes, arrhythmias, and subsequent death, has not been determined but seems likely. As with any condition that requires multiple medications, practitioners need to be aware of potential drug–drug interactions.

Some patients tolerate very high dosages of FGAs, particularly high-potency drugs, leading some clinicians to raise the prescribed doses in the hope that "more is better." This resulted in a substantial increase in the average dose of antipsychotic prescribed in the United States during the 1970s and 1980s. However, dosage comparison studies failed to support the routine use of higher doses. For example, when groups of patients were assigned to high doses such as over 2,000 mg of chlorpromazine or 40 mg of haloperidol, the commensurate improvements were no greater than in those assigned to more moderate doses. Moreover, these higher doses were associated with more EPS. Current evidence does not support the use of excessive doses of antipsychotic medications.

As-Needed Medication

Prescribing medication as needed is recommended only for the initial phases of treatment. Although FGAs can result in nearly immediate calming for agitated individuals, this response appears to differ from the true antipsychotic effect of these medications, which has a delayed onset. Treating excitement and agitation with the addition of benzodiazepines such as lorazepam (Ativan) has the advantage of calming patients without causing EPS.

Blood Levels and Clinical Activity

Drug response is an interaction between dosage and time on drug and individual pharmacokinetic and dynamic variables. Finding the best dosage of an antipsychotic for an individual patient often requires weeks or months of adjusted treatment. The clinician usually selects a drug dosage that is effective for most patients with the same illness or a dosage that has been effective for the patient at an earlier time. Once a decision is made about the dosage and the plan implemented, the clinician is forced to wait days or weeks until it is clear whether the dosage is effective, toxic, or ineffective. If measuring a patient's plasma or blood level enabled clinicians to select a drug dosage for an individual scientifically, this could decrease the amount of time it takes to arrive at an appropriate dosage.

Early studies of plasma levels indicated that patients who received the same dosage of a drug had large variations in their plasma levels. These findings suggested that studying individual differences in plasma levels would provide information that could be useful to clinicians. However, these early studies were limited to drugs such as chlorpromazine that have complex metabolisms. This creates a problem since some of the antipsychotic activity of these agents could result from metabolites rather than the parent drug. This and other problems including flawed research methods may explain why early studies failed to find a reliable relationship.

A number of more recent studies have focused on other dopamine receptor antagonists and have shown more promising results. Haloperidol has received the most attention for the study of relationships between plasma concentration and clinical response. An advantage of haloperidol for these studies is that its metabolism includes only a single important metabolite, reduced haloperidol, which does not have significant antipsychotic activity. As a result, measuring the plasma level of haloperidol may provide an accurate representation of the effective amount of the drug. At least five research groups, but not all, have reported a therapeutic window relationship between haloperidol levels and clinical response. Table 34.17–10 gives reported plasma concentration ranges for four commonly prescribed dopamine receptor antagonists. Although all of these ranges are supported by one or more high-quality study, the strongest evidence is for haloperidol.

Table 34.17–10.
Reported Therapeutic Plasma Concentration Ranges for Antipsychotic Drugs

Therapeutic Plasma Concentration (ng/mL)
30–100[a]
0.2–2.0[b]
2–15[a]

	0.8–2.4[b]

[a]Consider dosage reduction in patients with plasma concentration above upper limit.

[b]Range of concentration in which good response without debilitating adverse effects has been found; no evidence that reduction of plasma concentration above the upper limit improves the antipsychotic effect.

Adapted from Van Putten T, Marder SR, Wirshing WC, Aravagiri M, Chabert N. Neuroleptic plasma levels. *Schizophr Bull.* 1991;17:197–216.

Studies do not support the routine use of plasma level measurement with the exception of clozapine. Still, a level may be helpful in the patient who is receiving what appears to be an adequate drug dosage but who is not responding. A low level may suggest that the patient is not complying or that poor absorption or extensive drug metabolism is interfering with an adequate level. In other cases, the clinician may find it difficult to interpret anxiety or agitation and may consider whether psychotic agitation could come from an inadequate dosage or whether akathisia may result from an excessive dosage. A high or low plasma level could provide valuable information in these circumstances.

Plasma levels should only be obtained after steady-state levels have been reached, preferably during the late phase of the plasma level curve. This usually means after 5 days on a stable oral dose and 10 to 12 hours after the last dose.

Adjunctive Medications

Antipsychotic medications are frequently supplemented by other psychotropic medications. These are aimed at managing comorbid conditions that frequently occur with schizophrenia. There is substantial evidence indicating that depression is underdiagnosed and undertreated, while antidepressants may be effective for treating depressive symptoms in patients with schizophrenia and schizoaffective disorder. Mood stabilizing drugs including lithium, carbamazepine, and valproate (Depakote) are often used for patients with unstable moods. However, carbamazepine may induce the metabolism of antipsychotic agents, lowering plasma concentrations of the parent drug. Benzodiazepines are often effective for treating anxiety as well as agitation and excitement. Newer sleeping agents may assist patients to sleep better with lower risks than older medications.

MAINTENANCE TREATMENT

Duration of Treatment

During the maintenance phase of treatment, the treatment goals include preventing psychotic relapse and supporting treatments aimed at improving the patient's level of functioning and quality of life. Although

FGAs are effective in delaying or preventing relapse, they are also associated with short- and long-term side effects. For most patients with schizophrenia or schizoaffective illness, the analysis of costs and benefits will favor continuing the patient on an antipsychotic, including depot antipsychotics. Patients who have recovered from a single psychotic episode have a substantial risk of relapse if they are withdrawn from medications. As a result, first-episode patients should receive at least 1 year of maintenance therapy if asymptomatic. This is often a substantial point of contention with parents and their families, but the evidence is very clear that relapse will occur in 95 percent of patients eventually. Patients who have had multiple episodes should continue to receive antipsychotic maintenance for at least 2 to 5 years, potentially indefinitely. Present data indicate that there is a 10 percent relapse rate per year, often precipitated by nonadherence with medication, suggesting that continuous, moderate dosage maintenance FGAs afford the best chance of avoiding relapse in psychotic patients. Patients with a history of violent, aggressive behavior or serious suicide attempts should receive antipsychotics indefinitely. If a decision is made to discontinue medications, the dosage of medication should gradually be reduced and the patient should be carefully observed. However, as with many chronic medical diseases, patients do discontinue or take medications irregularly on their own.

Long-Acting Injectable Antipsychotics

There is substantial evidence indicating that long-acting injectable antipsychotics have advantages for patients, including first episode patients during the maintenance stage of treatment. Depot antipsychotics should be considered for each patient, starting early in the course of illness, because they have been shown to reduce relapse more than oral antipsychotics, potentially improving long-term prognosis as well as preventing both psychiatric readmissions and arrests. One advantage is that these agents are a partial solution to the problem of poor medication adherence in schizophrenia. Clinicians can administer a drug without depending on patients to take their pills. In addition, because depot drugs bypass variations in drug absorption and first-pass hepatic and gut metabolism, drug levels may be more consistent. A number of studies have compared the relative effectiveness of long-acting and oral drugs for preventing relapse in stabilized patients. Studies that closely resemble treatment in the community have found that patients are better protected by depot medications. A number of carefully controlled studies and meta-analyses have compared the risk of psychotic relapse in patients who were randomly assigned to receive either oral or depot medication, with depot antipsychotics reducing relapse rates to a greater extent than oral antipsychotics.

Treatment Strategies during Maintenance

Concerns about the side effects of antipsychotics and the risk of TD have led to a number of proposed strategies for safely managing patients on the lowest possible amount of a dopamine antagonist. A number of studies have found that patients can be safely managed with substantially lower dosages of a long-acting depot antipsychotic than are commonly prescribed. These studies found that most patients did well when their doses were reduced by 80 percent from doses that are usually prescribed for acute treatment. Patients treated with lower dosages experienced a small increase in the risk of relapse that was balanced in some studies by reductions in side effects, improved compliance, and suggestions of improved community adjustment.

Other studies have focused on a strategy called *targeted* or *intermittent therapy,* in which an antipsychotic is gradually reduced and eventually discontinued. Patients are monitored carefully for early or prodromal signs of relapse. The antipsychotic is reintroduced when these signs appear. A number of studies, including an NIMH collaborative study, have found that targeted treatment is associated with high relapse rates, indicating that this strategy is not advisable.

References

Baldessarini RJ, Cohen BM, Teicher MH. Significance of neuroleptic dose and plasma level in the pharmacological treatment of psychoses. *Arch Gen Psychiatry*. 1988;45:79–91.

Buchanan RW, Kreyenbuhl J, Kelly DL, et al. Schizophrenia PORT psychopharmacology treatment recommendations and summary statements. *Schizophr Bull*. 2010;36:71–93.

Chiodo LA, Bunney BS. Population response of midbrain dopaminergic neurons to neuroleptics: further studies on time course and nondopaminergic neuronal influences. *J Neurosci*. 1987;7:629–633.

Creese I, Burt DR, Snyder SH. Dopamine receptor binding predicts clinical and pharmacologic potencies of antischizophrenic drugs. *Science*. 1976;192:481–483.

De Leon J, Simpson GM. Assessment of neuroleptic-induced extrapyramidal symptoms. In: Kane JM, Lieberman JA, eds. *Adverse Effects of Psychotropic Drugs*. New York: Guilford; 1992:218–234.

*Farde L, Nordstrom AL, Wiesel FA, Pauli S, Halldin C, Sedvall G. Positron emission tomographic analysis of central D and D_1 and D_2 dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine. Relation to extrapyramidal side effects. *Arch Gen Psychiatry*. 1992;49:538–544.

Geyer MA, Krebs-Thomson K, Braff DL, Swerdlow NR. Pharmacological studies of prepulse inhibition models of sensorimotor gating deficits in schizophrenia: a decade in review. *Psychopharmacology (Berl)*. 2001;156(2–3):117–154.

Gilbert PL, Harris MJ, McAdams LA, Jeste DV. Neuroleptic withdrawal in schizophrenic patients. A review of the literature. *Arch Gen Psychiatry*. 1995;52:173–188.

Hartling L, Abou-Setta AM, Dursun S, Mousavi SS, Pasichnyk D, Newton AS. Antipsychotics in

adults with schizophrenia: comparative effectiveness of first-generation versus second-generation medications: a systematic review and meta-analysis. *Ann Intern Med*. 2012;157(7):498–511.

Herz MI, Marder SR. *Schizophrenia: Comprehensive Treatment and Management*. Philadelphia, PA: Lippincott Williams & Wilkins; 2002.

Janicak PG, Davis JM, Preskorn SH, Ayd FJ. *Principles and Practice of Psychopharmacology*. 3rd ed. Baltimore, MD: Lippincott Williams & Wilkins; 2001.

Johnson DA, Pasterski G, Ludlow JM, Street K, Taylor RD. The discontinuance of maintenance neuroleptic therapy in chronic schizophrenic patients: drug and social consequences. *Acta Psychiatr Scand*. 1983;67:339–352.

Jones PB, Barnes TR, Davies L, et al. Randomized controlled trial of the effect on quality of life of second- vs first-generation antipsychotic drugs in schizophrenia: Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1). *Arch Gen Psychiatry*. 2006;63:1079–1087.

Joyce JN, Meador-Woodruff JH. Linking the family of D_2 receptors to neuronal circuits in human brain: insights into schizophrenia. *Neuropsychopharmacology*. 1997;16:375–384.

Kahn RS, Davis KL. New developments in dopamine and schizophrenia. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: The Fourth Generation of Progress*. New York, NY: Raven; 1995:1193–1203.

Kahn RS, Fleischhacker WW, Boter H, et al. Effectiveness of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: an open randomized clinical trial. *Lancet*. 2008;371:1085–1097.

Kane JM, Kishimoto T, Correll CU. Assessing the comparative effectiveness of long-acting injectable vs. oral antipsychotic medications in the prevention of relapse provides a case study in comparative effectiveness research in psychiatry. *J Clin Epidemiol*. 2013;66(8 Suppl):S37–S41.

Lerner PP, Miodownik C, Lerner V. Tardive dyskinesia (syndrome): current concept and modern approaches to its management. *Psychiatry Clin Neurosci*. 2015;69(6):321–334.

Leucht S, Cipriani A, Spineli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*. 2013;382(9896):951–962.

Leucht S, Corves C, Arbter D, Engel R, Li C, Davis J. Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. *Lancet*. 2009;373:31–41.

Leucht C, Heres S, Kane JM, Kissling W, Davis JM, Leucht S. Oral versus depot antipsychotic drugs for schizophrenia—a critical systematic review and meta-analysis of randomized long-term trials. *Schizophr Res*. 2011;127:83–92.

*Leucht S, Pitschel-Walz G, Abraham D, Kissling W. Efficacy and extrapyramidal side-effects of the new antipsychotics olanzapine, quetiapine, risperidone, and sertindole compared to conventional antipsychotics and placebo. A meta-analysis of randomized controlled trials. *Schizophr Res*. 1999;35(1):51–68.

Leucht S, Tardy M, Komossa K, et al. Antipsychotic drug versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. *Lancet*. 2012;379:2063–2071.

Leucht S, Wahlbeck K, Hamann J, Kissling W. New generation antipsychotics versus low-potency

conventional antipsychotics: a systematic review and meta analysis. *Lancet*. 2003;361:1581–1589.

Lieberman JA, Stroup TS, McEvoy JP, et al. Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) investigators. *N Engl J Med*. 2005;353:1209–1223.

Marder SR, Hubbard JW, Van Putten T, Midha KK. The pharmacokinetics of long-acting injectable neuroleptic drugs: clinical implications. *Psychopharmacology (Berl)*. 1989;98:433–439.

Mattai AK, Hill JL, Lenroot RK. Treatment of early-onset schizophrenia. *Curr Opin Psychiatry*. 2010;23:304–310.

McEvoy JP, Byerly M, Hamer RM, et al. Effectiveness of paliperidone palmitate vs haloperidol decanoate for maintenance treatment of schizophrenia: a randomized clinical trial. *JAMA*. 2014;311(19):1978–1987.

Moore TA, Buchanan RW, Buckley PF, et al. The Texas Medication Algorithm Project antipsychotic algorithm for schizophrenia: 2006 update. *J Clin Psychiatry*. 2007;68(11):1751–1762.

National Institute for Health and Clinical Excellence (NICE). Schizophrenia: core interventions in the treatment and management of schizophrenia in adults in primary and secondary care (Update), final version (NICE Clinical Guideline 82). London, NICE, March 2009.

Sachdev P, Kruk J. Clinical characteristics and predisposing factor in acute drug-induced akathisia. *Arch Gen Psychiatry*. 1994;51:963–974.

Schooler NR, Keith SJ, Severe JB, et al. Relapse and rehospitalization during maintenance treatment of schizophrenia. The effects of dose reduction and family treatment. *Arch Gen Psychiatry*. 1997;54:453–463.

Seeman P, Lee T, Chau-Wong M, Wong K. Antipsychotic drug doses and neuroleptic/dopamine receptors. *Nature*. 1976;261:717–719.

Sikich L, Frazer JA, McClellan J, et al. Double-blind comparison of first- and second generation antipsychotics in early-onset schizophrenia and schizoaffective disorder: findings from the treatment of early-onset schizophrenia spectrum disorders (TEOSS) study. *Am J Psychiatry*. 2008;165(11):1420–1431.

Smith RC, Segman RH, Golcer-Dubner T, Pavlov V, Lerer B. Allelic variation in ApoC3, ApoA5 and LPL genes and first and second generation antipsychotic effects on serum lipids in patients with schizophrenia. *Pharmacogenomics J*. 2008;8(3):228–236.

Tiihonen J Haukka J, Taylor M, Haddad P, Patel MX, Korhonen P. A nationwide cohort study of oral and depot antipsychotics after first hospitalization for schizophrenia. *Am J Psychiatry*. 2011;168:603–609.

Woods SW, Morgenstern H, Saksa JR, et al. Incidence of tardive dyskinesia with atypical versus conventional antipsychotic medications: a prospective cohort study. *J Clin Psychiatry*. 2010;71(4):463–474.

▲ 34.18 Lamotrigine

PO W. WANG, M.D., AND TERENCE ARTHUR KETTER, M.D.

HISTORY

Lamotrigine (Lamictal) was developed as a result of screening folate antagonists as anticonvulsants. The observations that folates could produce seizures in animals and some anticonvulsants such as phenytoin (Dilantin) and phenobarbital (Solfoton) could impair folate metabolism led to the hypothesis that folate antagonists could be anticonvulsants. Further research has failed to establish anti-folate as treatments for epilepsy in humans, nor has the exact role of folate in epileptogenesis been established. Lamotrigine also appears to lack clinically significant antifolate effects in humans. Nevertheless, lamotrigine did prove effective in several animal models of epilepsy. Burroughs Wellcome (later Glaxo Wellcome, and currently GlaxoSmithKline) developed lamotrigine as an antiepileptic drug, and it was marketed for the adjunctive treatment of partial seizures in the United States in 1995.

Initial, postmarketing, open, clinical experience suggested efficacy, coupled with generally good tolerability (aside from the risk of rash), in a variety of neurological and psychiatric conditions. Later double-blind, placebo-controlled studies revealed that lamotrigine was useful for some of the neurological and psychiatric conditions suggested in open studies. Lamotrigine demonstrated efficacy and was approved for maintenance treatment of bipolar I disorder in 2003. Lamotrigine also had some efficacy in acute bipolar depression, but the magnitude of the effect was only modest. Limited data also suggested lamotrigine had potential utility in rapid cycling bipolar disorder. However, lamotrigine did *not* receive approval for the treatment of acute bipolar depression nor rapid cycling bipolar disorder. Lamotrigine was *not* effective as a main intervention in acute mania. In summary, lamotrigine emerged as an agent that appears to "stabilize mood from below," such that it may maximally impact the depressive component of bipolar disorders and have some mood stabilizing effects as well.

CHEMISTRY

Lamotrigine, 3,5-diamino-6-(2,3-dichlorophenyl)-*as*-triazine, is a structurally novel phenyltriazine anticonvulsant (Fig. 34.18–1). Three oral lamotrigine preparations are currently available in the United States: 25-, 100-, 150-, and 200-mg immediate- or extended-release tablets and 50-, 250-, and 300-mg extended-release tablets; and 2-, 5-, 25-, 50-, 100-, and 200-mg chewable dispersible tablets. By mid 2016, generic formulations of

all of the above (with the exception of the 2- and 5-mg chewable dispersible tablets) were available.

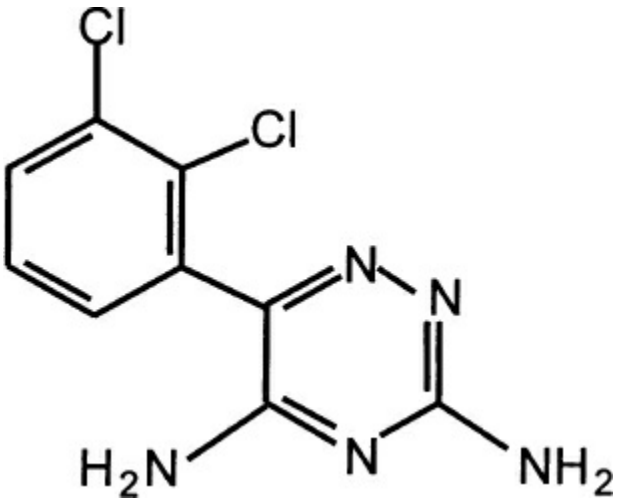

FIGURE 34.18–1. Molecular structure of lamotrigine.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

Lamotrigine has oral bioavailability of about 98 percent and linear kinetics that are not significantly altered by food. Lamotrigine is 55 percent bound to plasma proteins and has a moderate volume of distribution of about 1 L/kg. A large percentage (approximately 85 percent) of lamotrigine is conjugated to yield inactive glucuronide metabolites, with only about 10 percent excreted unchanged in the urine. In monotherapy its half-life is around 28 hours and clearance is about 40 mL per minute. Combined with hepatic enzyme inducers such as carbamazepine (Tegretol, Equetro) or phenytoin (Dilantin), lamotrigine's half-life falls about 50 percent to approximately 14 hours and clearance doubles to approximately 80 mL per minute. Estrogen-containing hormonal contraceptives also induce lamotrigine metabolism. Combined with the enzyme inhibitor valproate (Depakote), lamotrigine's half-life doubles to approximately 56 hours, and clearance falls approximately 50 percent to around 20 mL per minute. Thus, in patients with epilepsy immediate-release lamotrigine is given in two divided doses, unless valproate is present, which may allow single daily dosing. In bipolar disorder patients where stable blood levels are less critical, single daily dosing of immediate-release lamotrigine is used, unless an enzyme inducer (such as carbamazepine [Tegretol]) is present, in which case two divided doses are recommended. An extended-release formulation of lamotrigine has been approved for once daily administration in patients with epilepsy, but not bipolar disorder. Lamotrigine clearance is decreased with hepatic insufficiency, which may necessitate decreased dosage. Clearance is decreased to a lesser extent with renal insufficiency, but

dosage decreases may still be needed.

Pharmacodynamics

Like other anticonvulsants, lamotrigine has multiple biochemical actions, and the relationships between these and clinical effects remain to be established. Lamotrigine blocks voltage-sensitive sodium channels, modulating release of the excitatory amino acids glutamate and aspartate, and may have more limited effects on calcium channels. Lamotrigine also weakly blocks 5-HT_3 serotonergic receptors, has weak effects at σ-opioid receptors, and modestly inhibits serotonin uptake in human platelets and rat brain synaptosomes and noradrenaline and dopamine uptake in rat brain synaptosomes. In contrast, lamotrigine does not bind with high affinity to GABAergic, NMDA, glutamatergic, adrenergic, 5-HT_2 serotonergic, dopaminergic, cholinergic, histaminergic, adenosinergic, or κ-opiate–binding sites, and does not affect dihydropyridine calcium channels. Although lamotrigine inhibits dihydrofolate reductase in vitro, it does not appear to have clinically significant antifolate effects in humans.

Lamotrigine has antiseizure activity in maximal electroshock, pentylenetetrazol, amygdala-kindled, and other animal models of seizures; cerebroprotective actions in several animal models of ischemia; analgesic effects in animal models of neuropathic pain; and activity in the conflict animal model of anxiety.

Lamotrigine blood concentrations are not systematically related to antiseizure effects or efficacy in bipolar disorders. However, as other drugs may alter lamotrigine levels, monitoring of lamotrigine levels may have certain clinical uses.

CLINICAL DRUG STUDIES

Seizure Disorders

Immediate-release lamotrigine monotherapy appeared effective in a multicenter, double-blind controlled study of converting adults with partial seizures taking carbamazepine, phenytoin, phenobarbital, or primidone monotherapy to lamotrigine monotherapy. In a single study with less rigorous methodology, extended-release lamotrigine monotherapy also appeared effective for conversion to monotherapy in adolescents and adults with partial seizures. Adjunctive immediate-release lamotrigine appeared more effective than adjunctive placebo in three multicenter, double-blind, placebo-controlled trials in adults with refractory partial seizures, in one multicenter, double-blind, placebo-controlled trial in pediatric (age 2 to 16 years) partial seizure disorder patients, in one multicenter, double-blind, placebo-controlled study in adult and pediatric (age 3 to 25 years) patients with Lennox–Gastaut syndrome, and in one multicenter, double-blind, placebo-controlled trial in pediatric and adult patients with primary

generalized tonic–clonic seizures. Extended-release adjunctive lamotrigine appeared effective for primary generalized tonic–clonic seizures and partial-onset seizures in adolescents (age ≥13 years) and adults in multicenter, double-blind, randomized, placebo-controlled studies.

Pain Syndromes

Lamotrigine may have analgesic effects, as it was effective in placebo-controlled studies of posttransurethral prostatectomy pain, central poststroke pain, and HIV painful distal sensory polyneuropathy. In the latter study there was a higher rash frequency than seen in epilepsy studies. However, in placebo-controlled studies lamotrigine had inconsistent efficacy in diabetic neuropathy and was ineffective in neuropathic pain and spinal cord injury pain.

Other Neurological Disorders

Although open reports were encouraging, a double-blind, placebo-controlled study failed to show benefit in migraine. Lamotrigine was also ineffective in placebo-controlled studies in amyotrophic lateral sclerosis, Huntington disease, and Parkinson disease.

Bipolar Disorders

In two 18-month multicenter, randomized, double-blind, placebo-controlled trials (one after depressive episodes, one after manic or hypomanic episodes), both lamotrigine and lithium, compared with placebo, delayed time to intervention for any mood episode in bipolar I patients. Lamotrigine mood prophylaxis effect was driven by more robust delay of depression, so lamotrigine was predominantly effective against depressive relapse. Lithium mood prophylaxis was driven by more robust delay of mania, so lithium was predominantly effective against mania. Although an early double-blind, placebo-controlled study suggested lamotrigine monotherapy had some efficacy in the acute treatment of bipolar depression (based on a secondary outcome measure) four subsequent trials were negative. However, a meta-analysis including all five controlled trials suggested a statistically significant but only clinically modest efficacy in acute bipolar depression. One double-blind, placebo-controlled, industry-sponsored study suggested that the olanzapine (Zyprexa) plus fluoxetine (Prozac) combination compared to lamotrigine monotherapy had modestly superior efficacy for acute bipolar depression, but also more adverse effects. A small controlled trial suggested that adjunctive lamotrigine compared to adjunctive risperidone (Risperdal) might have utility in bipolar disorder patients who remained depressed despite taking a mood stabilizer plus an antidepressant. In treatment-resistant (primarily rapid cycling bipolar) mood disorder inpatients, about

50 percent responded to lamotrigine monotherapy, compared to only about 25 percent with gabapentin (Neurontin) monotherapy or placebo. In less treatment-resistant community sample of patients with bipolar disorder and rapid cycling course of illness, lamotrigine monotherapy appeared more effective than placebo bipolar II disorder, but not bipolar I disorder. Adjunctive (added to lithium) lamotrigine compared to adjunctive placebo was superior in acute bipolar depression, and also appeared to provide continued benefit in longer-term treatment. Adding lamotrigine to quetiapine yielded acute and longer-term benefit in patients with quetiapine-resistant bipolar depression. Lamotrigine monotherapy has not been shown efficacious in controlled studies of acute mania. In summary, these controlled studies support the notion that lamotrigine may "stabilize mood from below" in the sense that it may maximally impact the depressive component of bipolar disorders.

Major Depressive Disorders

There are fewer data suggesting lamotrigine efficacy in (unipolar) major depressive disorders. One of two double-blind, placebo-controlled monotherapy trials suggested possible efficacy. In a placebo-controlled study, adjunctive lamotrigine appeared to accelerate the onset but not degree of paroxetine (Paxil) response in depressed (mainly unipolar) patients. In another study, adjunctive lamotrigine compared to placebo yielded similar improvements in depression ratings, but better improvement in global (Clinical Global Impression) ratings, in (mainly unipolar) patients with depression nonresponsive to fluoxetine monotherapy.

Other Psychiatric Disorders

There are only limited data regarding the utility of lamotrigine in other psychiatric disorders. In controlled studies, chronic lamotrigine was no better than placebo in children with autism or adults with depersonalization disorder, and was only numerically but not statistically superior to placebo in adults with PTSD. Single-dose lamotrigine compared to placebo did not significantly attenuate the effects of acute cocaine or naloxone (Narcan)-induced opiate withdrawal. In hospitalized treatment-resistant schizophrenia patients on clozapine (Clozaril), adding lamotrigine compared to placebo reduced general psychopathological and positive (but not negative) symptoms. A subsequent similar trial of adding lamotrigine to conventional and atypical antipsychotics failed to replicate this in the primary (last observation carried forward) analysis, but was positive in a posthoc less rigorous observed case analysis. Interestingly, in a small controlled trial, lamotrigine appeared superior to placebo in the treatment of anger in women with borderline personality disorder. Lamotrigine was

also superior to adding placebo to SSRI treatment for Obsessive-Compulsive Disorder.

Obesity

A small controlled study in patients with obesity suggested that lamotrigine compared to placebo might yield modestly greater decreases in BMI, but not weight.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Lamotrigine has been associated with common benign rashes, uncommon serious rashes, and rare hypersensitivity reactions. Lamotrigine binds to melanin-containing tissues and thus has the potential to accumulate in the eyes. However, to date lamotrigine has not been associated with ocular problems. Lamotrigine inhibits dihydrofolate reductase in vitro but does not appear to have clinically significant antifolate effects in humans.

THERAPEUTIC INDICATIONS

Seizure Disorders

Immediate-release lamotrigine is indicated as adjunctive treatment for adults and children over age 2 years with partial seizures and for pediatric (over age 2 years) and adult patients with Lennox–Gastaut syndrome, and for conversion to monotherapy in adults with partial seizures who are receiving valproate or a single enzyme inducing antiepileptic drug such as carbamazepine, phenytoin, phenobarbital, or primidone in patients ≥16 years old. Extended-release lamotrigine is indicated in adolescents and adults as adjunctive therapy for primary generalized tonic–clonic seizures and partial-onset seizures, as well as for conversion to monotherapy in patients with partial seizures who are taking a single anticonvulsant.

Pain Syndromes

Emerging controlled data suggest lamotrigine may be effective in posttransurethral prostatectomy pain, diabetic neuropathy, central poststroke pain, and HIV painful distal sensory polyneuropathy. However, as of late 2016 lamotrigine lacked FDA approval for such uses.

Bipolar Disorders

Lamotrigine is indicated for the maintenance treatment of bipolar I disorder to delay the time to occurrence of mood episodes in patients treated for acute mood episodes with standard therapy. Limited controlled data suggested lamotrigine might have at least some efficacy in the acute treatment of bipolar depression and in rapid cycling bipolar disorder. In view of such data, lamotrigine was considered to be an important treatment

option for bipolar maintenance treatment and acute bipolar depression in the 2013 revision of the Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders' update of the CANMAT guidelines for the management of patients with bipolar disorder. However, as of late 2016, lamotrigine continued to lack FDA approval for the treatment of acute bipolar depression.

PRECAUTIONS AND ADVERSE REACTIONS

The prescribing information for lamotrigine includes a boxed warning due to the risk of serious rashes (including Stevens–Johnson syndrome) requiring hospitalization and discontinuation of treatment. The incidence of serious rashes was estimated in 1997 to be 1 percent in prospectively monitored pediatric (age less than 16 years) and 0.3 percent in premarketing clinical trials of adult patients. Rare cases of toxic epidermal necrolysis and rash-related death have been reported, but numbers are too few to allow precise estimates of rates. Other than age, possible factors increasing the risk of rash include factors that increase lamotrigine doses too quickly: concurrent administration of valproate (which doubles lamotrigine blood concentrations), exceeding the recommended initial lamotrigine dose, or exceeding the recommended lamotrigine dose escalation. Nearly all life-threatening rashes have occurred between 2 and 8 weeks after lamotrigine initiation, but isolated cases have been reported after prolonged treatment. If patients do not take lamotrigine for 5 days or more (approximately five times the half-life of lamotrigine), the drug needs to be reintroduced according to the gradual initial escalation regimen, as serious rashes can occur with rapidly restarting the medication.

Benign rashes can occur in as many as 10 percent of patients taking lamotrigine, and the prescribing information recommends that lamotrigine be ordinarily discontinued at the first sign of rash, unless the rash is clearly not drug related. Determination of the latter can contribute substantively to important clinical decisions. Thus, any rash in a patient on lamotrigine requires immediate evaluation. Some clinicians report that "dermatology precautions" and initial titration even slower than that described in the prescribing information may decrease the rate of benign rash. Dermatology precautions include not starting lamotrigine within 2 weeks of having a rash, viral syndrome, or vaccination, and during the first 3 months on lamotrigine to avoid ingesting other new medicines or new foods or utilizing new cosmetics, conditioners, deodorants, detergents, or fabric softeners, or stimulating their immune systems by developing sunburn or poison ivy/oak exposure. However, the utility of dermatology precautions was not confirmed in a controlled study. The utility of initial titration even slower than that described in the prescribing information remains to be established in a controlled study.

Any patient developing a rash accompanied by eye, mouth, or bladder

discomfort, fever, or lymphadenopathy should go immediately the emergency room for dermatological assessment; as such symptoms could herald the onset of a serious rash. Rashes with more benign presentations require medical evaluation as soon as possible (and before taking the next dose of lamotrigine) for assessment of the risk of continuing lamotrigine treatment. Rare hypersensitivity reactions with multiorgan failure/dysfunction have been reported, so that patients who develop fever or lymphadenopathy require emergent evaluation. Rare blood dyscrasias have been reported. In view of the potential risk of withdrawal seizures, lamotrigine should be tapered over a period of at least 2 weeks unless safety concerns require a more rapid lamotrigine discontinuation.

With the exception of the risk of rash, lamotrigine is generally well tolerated. Although transient difficulties with headache may be encountered during the initiation phase, other difficulties such as sedation, neurocognitive adverse effects, sexual problems, or weight gain are not typically seen.

The most common lamotrigine adverse effects during dose escalation (added to standard therapy) phase in adult bipolar disorder patients were headache, rash, dizziness, diarrhea, dream abnormality, and pruritus. In the lamotrigine monotherapy phase in maintenance treatment studies in adult bipolar disorder patients, the most common adverse effects were nausea, insomnia, somnolence, back pain, fatigue, rash, and rhinitis, and the adverse effects most commonly leading to discontinuation were rash and manic/hypomanic/mixed mood symptoms.

The most common lamotrigine monotherapy adverse effects in adult epilepsy patients were vomiting, incoordination, dyspepsia, nausea, dizziness, rhinitis, anxiety, insomnia, infection, pain, weight loss, chest pain, and dysmenorrhea. The most common adjunctive lamotrigine adverse effects in adult epilepsy patients were dizziness, ataxia, somnolence, headache, diplopia, blurred vision, nausea, vomiting, and rash. The most common adjunctive lamotrigine adverse effects in pediatric epilepsy patients were infection, vomiting, rash, fever, somnolence, accidental injury, dizziness, diarrhea, abdominal pain, nausea, ataxia, tremor, asthenia, bronchitis, flu syndrome, and diplopia.

In late 2008, the FDA announced a class warning regarding increased risk of suicidality (suicidal behavior or ideation) in patients with epilepsy as well as psychiatric disorders for lamotrigine as well as 10 other anticonvulsants (carbamazepine, felbamate [Felbatol], gabapentin [Neurontin], levetiracetam [Keppra], oxcarbazepine [Trileptal], pregabalin [Lyrica], tiagabine [Gabitril], topiramate [Topamax], valproate, and zonisamide [Zonegran]). In the FDA's analysis, anticonvulsants compared to placebo yielded approximately twice the risk of suicidality (0.43 percent vs. 0.22 percent). The relative risk for suicidality was higher in the patients with epilepsy compared to psychiatric disorders. Acute oral overdoses of

lamotrigine up to 15 g have caused central nervous system and cardiac adverse effects, and some have been fatal.

Lamotrigine was classified as pregnancy category C. Although to date teratogenicity has not been definitively demonstrated in animals or humans, the ability of lamotrigine to inhibit dihydrofolate reductase in vitro has raised concerns regarding the risk of malformations. Limited data suggest the possibility of a modest association between lamotrigine and cleft lip or palate, but this remains to be definitively established. Specifically, the North American Antiepileptic Drug Registry found a higher risk of oral cleft, which was not found in other pregnancy registries. Lamotrigine is excreted in breast milk, considered as "unsafe" for lactation, and thus is contraindicated with breastfeeding.

Lamotrigine is contraindicated in patients with known hypersensitivity to lamotrigine or its ingredients. The lamotrigine prescribing information concludes with the recommended wording for a leaflet to provide to patients.

DRUG INTERACTIONS

Lamotrigine is mainly metabolized by glucuronidation with minimal involvement of oxidative metabolism, and is primarily a target rather than an instigator of drug–drug interactions. Thus, enzyme inducers such as carbamazepine, phenytoin, phenobarbital, primidone (Mysoline), rifampin (Rifadin), and estrogen-containing (but not progesterone) hormonal contraceptives can decrease blood lamotrigine concentrations by 40 to 50 percent. The carbamazepine derivatives, oxcarbazepine (Trileptal) and eslicarbazepine acetate (Aptiom), appear to less robustly decrease lamotrigine levels (so that dosage adjustment of lamotrigine administered with these agents is not recommended). The enzyme inhibitor valproate approximately doubles blood lamotrigine concentrations. Also, sertraline (Zoloft) may increase lamotrigine levels.

Lamotrigine appears to have at most modest effects on the pharmacokinetics of other drugs. For example, lamotrigine can yield modest (25 percent) decreases in valproate levels. Although lamotrigine can exacerbate the neurotoxic effects of carbamazepine, this may be due to a pharmacodynamic interaction, as only equivocal evidence suggests that lamotrigine could increase levels of the epoxide metabolite of carbamazepine. Lamotrigine does not alter lithium pharmacokinetics, which is important, given lamotrigine's utility in bipolar disorders. Lamotrigine does not appear to have clinically significant drug interactions with clozapine, olanzapine, risperidone, or aripiprazole, although a therapeutic drug monitoring database study suggested that lamotrigine might decrease quetiapine concentrations. Lamotrigine does not appear to have clinically significant antifolate effects in humans, although folate may interfere with the benefit of adding lamotrigine to quetiapine.

LABORATORY INTERFERENCES

Lamotrigine has not been associated with interference with laboratory tests.

DOSAGE AND ADMINISTRATION

It is crucial to gradually introduce lamotrigine in order to attenuate the risk of rash.

Because valproate increases and enzyme inducers (such as carbamazepine, phenytoin, phenobarbital, primidone, and rifampin) decrease lamotrigine blood a concentration, the lamotrigine prescribing information includes detailed instructions for gradual lamotrigine introduction (Table 34.18–1).

In maintenance treatment of adults with bipolar I disorder, when given without valproate or enzyme-inducing anticonvulsants, lamotrigine is started at 25 mg per day for weeks 1 and 2, increased to 50 mg per day for weeks 3 and 4, increased to 100 mg per day for week 5, and then increased to the 200 mg per day target dose. Some clinicians use even more gradual initial titration in efforts to further decrease the risk of rash, starting with 25 mg per day for the first 2 weeks, increasing to 50 mg per day for the next 2 weeks, and thereafter increasing by 25 mg per day on a weekly basis.

When added to valproate, lamotrigine is started at 25 mg every *other* day (or 12.5 mg per day) for weeks 1 and 2, increased to 25 mg per day for weeks 3 and 4, increased to 50 mg per day for week 5, and then increased to the 100 mg per day target dose. Some clinicians use even more gradual initial titration in efforts to further decrease the risk of rash, starting with 12.5 mg per day for the first 2 weeks, increasing to 25 mg per day for the next 2 weeks, and thereafter increasing by 12.5 mg per day on a weekly basis.

When added to enzyme-inducing anticonvulsants (such as carbamazepine), lamotrigine is started at 50 mg per day for weeks 1 and 2, increased to 100 mg per day in divided doses for weeks 3 and 4, increased to 200 mg per day in divided doses for week 5, increased to 300 mg per day in divided doses for week 6, and then increased to 400 mg per day in divided doses to the target dose. Some clinicians use even more gradual initial titration in efforts to further decrease the risk of rash, starting with 50 mg per day for the first 2 weeks, increasing to 100 mg per day for the next 2 weeks, and thereafter increasing by 50 mg per day on a weekly basis.

Based on controlled trials, the prescribing information describes target doses of 100 mg per day (with valproate), 400 mg per day (with enzyme inducers), and 200 mg per day (with neither), and does not recommend higher doses. However, some clinicians choose to gradually further increase doses as necessary and tolerated to about twice these figures. Although a therapeutic plasma concentration range has not been

established for lamotrigine for bipolar disorder or for epilepsy, if patients attain high doses (e.g., 400 mg per day without inducers), it may be helpful to assess a 12-hour trough lamotrigine level to determine if there is a pharmacokinetic cause to apparent nonresponse. To date, lamotrigine use in bipolar disorders has been limited in patients under age 16 years, due to the lack of an FDA indication and the potential for increased risk of rash.

Table 34.18–1.
Gradual Introduction of Lamotrigine in Adults with Bipolar Disorders

	Lamotrigine with Valproate (mg/day)	Lamotrigine with Enzyme Inducers (mg/day)	Lamotrigine with neither Valproate nor Enzyme Inducers (mg/day)
Weeks 1 and 2 dose	12.5	50	25
Weeks 3 and 4 dose	25	100	50
Week 5 dose	50	200	100
Week 6 dose	100	300	200
Week 7 and thereafter (FDA target dose)	100	400	200
Typical final dose range	100–200	400–800	200–400

Adapted from FDA, U.S. Food and Drug Administration.

In adults and adolescents over age 12 years with epilepsy taking valproate, adjunctive lamotrigine should start at 25 mg every *other* day for weeks 1 and 2, be increased to 25 mg per day for weeks 3 and 4, and increased every 1 to 2 weeks thereafter by 25 to 50 mg per day to 100 to 200 mg per day in one or two divided doses. In adults and adolescents over age 12 years with epilepsy on enzyme-inducing anticonvulsants (in the absence of valproate) adjunctive lamotrigine should start at 50 mg per day for weeks 1 and 2, be increased to 100 mg per day for weeks 3 and 4, and increased every 1 to 2 weeks thereafter by 100 mg per day to 300 to 500 mg per day in two divided doses. The prescribing information describes weight-proportional gradual adjunctive lamotrigine introduction in pediatric (age 2 to 12 years) epilepsy patients, with maintenance doses of 1 to 5 mg/kg per day (maximum 200 mg per day in one or two divided doses) in patients taking valproate, and 5 to 15 mg/kg per day (maximum 400 mg per day in two divided doses) in patients on enzyme-inducing anticonvulsants in the absence of valproate. The recommended lamotrigine monotherapy maintenance dose in epilepsy is 500 mg per day in divided doses.

Although estrogen-containing hormonal contraceptives can decrease blood lamotrigine concentrations by as much as 50 percent, no adjustment

to the lamotrigine dose escalation guidelines is recommended solely based on the use of estrogen-containing hormonal contraceptives. However, the maintenance dose of lamotrigine may need to be increased as much as twofold.

When estrogen-containing hormonal contraceptives are added to lamotrigine, lamotrigine doses may be increased by 50 to 100 mg per day on a weekly basis. If adverse effects related to increased lamotrigine concentrations such as dizziness, ataxia, and diplopia consistently occur during the week of inactive hormonal preparation (i.e., the "pill-free" week), it may be necessary to decrease the overall lamotrigine maintenance dosage, but dose adjustments confined to the "pill-free" week are not recommended.

When estrogen-containing hormonal contraceptives are discontinued, lamotrigine doses may be decreased by 25 percent the first week and by another 25 percent the second week.

In patients with hepatic impairment, lamotrigine doses may need to be decreased by 25 percent in patients with moderate to severe hepatic impairment without ascites, and 50 percent in patients with severe hepatic impairment with ascites. Lamotrigine dosage may need to be decreased in patients with impaired renal function.

If patients do not take lamotrigine for 5 days or more (i.e., for five times the half-life of lamotrigine or more), the drug needs to be reintroduced according to the gradual initial escalation regimen as serious rashes can occur with rapidly restarting the medication.

In conditions other than seizure disorders, lamotrigine is commonly given in one or two divided doses. Single daily dosing is convenient, enhances adherence, and in contrast to epilepsy, breakthrough symptoms due to the resulting lower trough blood concentrations are less of a concern. In some (nonepilepsy) applications, clinicians may administer most or all of the dosage in the morning or at bedtime if individual patients find the drug activating or sedating, respectively.

References

Agarwal N, Joshi M. Effectiveness of amitriptyline and lamotrigine in traumatic spinal cord injury-induced neuropathic pain: a randomized longitudinal comparative study. *Spinal Cord*. 2016.

Barbee JG, Thompson TR, Jamhour NJ, et al. A double-blind placebo-controlled trial of lamotrigine as an antidepressant augmentation agent in treatment-refractory unipolar depression. *J Clin Psychiatry*. 2011;72(10):1405–1412.

Biederman J, Joshi G, Mick E, et al. A prospective open-label trial of lamotrigine monotherapy in children and adolescents with bipolar disorder. *CNS Neurosci Ther*. 2010;16(2):91–102.

*Bowden CL, Calabrese JR, Sachs G, et al; Lamictal 606 Study Group. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently manic or hypomanic patients with bipolar I disorder. *Arch Gen Psychiatry*. 2003;60(4):392–400.

Brown E, Dunner DL, McElroy SL, et al. Olanzapine/fluoxetine combination vs. lamotrigine in the 6-month treatment of bipolar I depression. *Int J Neuropsychopharmacol.* 2009;12(6):773–782.

*Calabrese JR, Bowden CL, Sachs G, et al; Lamictal 605 Study Group. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently depressed patients with bipolar I disorder. *J Clin Psychiatry*. 2003;64(9):1013–1024.

Calabrese JR, Huffman RF, White RL, et al. Lamotrigine in the acute treatment of bipolar depression: results of five double-blind, placebo-controlled clinical trials. *Bipolar Disord.* 2008;10:323–333.

*Calabrese JR, Suppes T, Bowden CL, et al. A double-blind, placebo-controlled, prophylaxis study of lamotrigine in rapid-cycling bipolar disorder. *J Clin Psychiatry*. 2000;61(11):841–850.

Findling RL, Chang K, Robb A, et al., Adjunctive maintenance lamotrigine for pediatric bipolar I disorder: a placebo-controlled, randomized withdrawal study. *J Am Acad Child Adolesc Psychiatry*. 2015;54(12):1020–1031.

*Geddes JR, Calabrese JR, Goodwin GM. Lamotrigine for treatment of bipolar depression: independent meta-analysis and meta-regression of individual patient data from five randomised trials. *Br J Psychiatry*. 2009;194:4–9.

Geddes JR, Gardiner A, Rendell J, et al. Comparative evaluation of quetiapine plus lamotrigine combination versus quetiapine monotherapy (and folic acid versus placebo) in bipolar depression (CEQUEL): a 2 × 2 factorial randomised trial. *Lancet Psychiatry*. 2016;3(1):31–39.

Goldberg JF, Bowden CL, Calabrese JR. Six-month prospective life charting of mood symptoms with lamotrigine monotherapy versus placebo in rapid cycling bipolar disorder. *Biol Psychiatry*. 2008;63:125–130.

Ittasakul P, Johnson KR, Srivastava S, et al. Effectiveness of quetiapine plus lamotrigine maintenance therapy in challenging bipolar disorder patients. *J Affect Disord.* 2012;137(1–3):139–145.

Kessing LV, Hellmund G, Andersen PK. An observational nationwide register based cohort study on lamotrigine versus lithium in bipolar disorder. *J Psychopharmacol.* 2012;26(5):644–652.

Ketter TA, Brooks JO, Hoblyn JC, et al. Effectiveness of lamotrigine in bipolar disorder in a clinical setting. *J Psychiatr Res*. 2008;43:13–23.

*Ketter TA, Calabrese JR. Stabilization of mood from below versus above baseline in bipolar disorder: a new nomenclature. *J Clin Psychiatry*. 2002;63(2):146–151.

Ketter TA, Manji HK, Post RM. Potential mechanisms of action of lamotrigine in the treatment of bipolar disorders. *J Clin Psychopharmacol.* 2003;23(5):484–495.

Khalkhali M, Aram S, Zarrabi H, Kafie M, Heidarzadeh A. Lamotrigine augmentation versus placebo in serotonin reuptake inhibitors-resistant obsessive-compulsive disorder: a randomized controlled trial. *Iran J Psychiatry*. 2016;11(2):104–114.

Normann C, Hummel B, Scharer LO, Horn M, Grunze H, Walden J. Lamotrigine as adjunct to paroxetine in acute depression: A placebo-controlled, double-blind study. *J Clin Psychiatry*. 2002;63(4):337–344.

Sajatovic M, Gildengers A, Al Jurdi RK, et al. Multisite, open-label, prospective trial of lamotrigine for geriatric bipolar depression: a preliminary report. *Bipolar Disord.*

2011;13(3):294–302.

Solmi M, Veronese N, Zaninotto L, et al. Lamotrigine compared to placebo and other agents with antidepressant activity in patients with unipolar and bipolar depression: a comprehensive meta-analysis of efficacy and safety outcomes in short-term trials. *CNS Spectr*. 2016;21(5):403–418.

Suppes T, Marangell LB, Bernstein IH, et al. A single blind comparison of lithium and lamotrigine for the treatment of bipolar II depression. *J Affect Disord*. 2008;111:334–343.

Tiihonen J, Hallikainen T, Ryynanen OP, et al. Lamotrigine in treatment-resistant schizophrenia: A randomized placebo-controlled crossover trial. *Biol Psychiatry*. 2003;54(11):1241–1248.

Tritt K, Nickel C, Lahmann C, et al. Lamotrigine treatment of aggression in female borderline-patients: a randomized, double-blind, placebo-controlled study. *J Psychopharmacol*. 2005;19(3):287–291.

van der Loos ML, Mulder PG, Hartong EG, et al; LamLit Study Group. Efficacy and safety of lamotrigine as add-on treatment to lithium in bipolar depression: a multicenter, double-blind, placebo-controlled trial. *J Clin Psychiatry*. 2009;70(2):223–231.

van der Loos ML, Mulder P, Hartong EG, et al; LamLit Study Group. Long-term outcome of bipolar depressed patients receiving lamotrigine as add-on to lithium with the possibility of the addition of paroxetine in nonresponders: a randomized, placebo-controlled trial with a novel design. *Bipolar Disord*. 2011;13(1):111–117.

Vestergaard K, Andersen G, Gottrup H, Kristensen BT, Jensen TS. Lamotrigine for central poststroke pain: a randomized controlled trial. *Neurology*. 2001;56(2):184–190.

Yatham LN, Kennedy SH, Parikh SV, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2013. *Bipolar Disord*. 2013;15(1):1–44.

Zoccali R, Muscatello MR, Bruno A, et al. The effect of lamotrigine augmentation of clozapine in a sample of treatment-resistant schizophrenic patients: a double-blind, placebo-controlled study. *Schizophr Res*. 2007;93(1–3):109–116.

▲ 34.19 Lithium

CHARLES L. BOWDEN, M.D.

HISTORY

John Cade, a psychiatrist working in a psychiatric hospital in Australia, was the first person to treat a patient with lithium and observe its unequivocal beneficial effects in previously chronically manic patients. The clinical observations and associated publications by Cade were followed shortly thereafter by systematic clinical trials in Europe led by Schou, Baastrup, and Gershon in Australia which in turn led shortly thereafter to its approval by regulatory agencies for the treatment of mania. In the United States, the FDA approved its labeling for the treatment of mania in 1970. The only

other approval indication came in 1974 for maintenance therapy in bipolar patients with a history of mania. The story of its development therefore forms a major part of the foundation for the psychopharmacology revolution experienced in the last half of the 20th century.

CHEMISTRY

Lithium follows hydrogen and helium on the periodic table, making it the third simplest element (atomic number 3, atomic weight 6.94) and the first solid element. Whether lithium is essential to normal human growth and development has not been established. There is no evidence of a lithium deficiency state in humans. Experimentally produced lithium deficiency in goats and rats led to reduced fertility, retarded development, and reduced longevity.

PHARMACOLOGICAL ACTIONS

Lithium is rapidly and completely absorbed, with serum concentrations peaking in 1 to 1.5 hours with standard preparations and in 4 to 4.5 hours with slow and controlled-release forms. Unlike most psychiatric drugs, lithium has no clinically important protein-binding properties and no metabolites. It is excreted by the kidneys, although small amounts are also lost in sweat and feces. A substantial amount of filtered lithium is reabsorbed (primarily in the proximal tubules), so that renal lithium clearance is about one-fifth of creatinine clearance. The elimination half-life of lithium is about 18 to 24 hours, although it is considerably longer in the elderly because of the age-related decrease in GFR. Consequently, elderly patients generally need lower than usual doses to reach a given serum concentration and require a longer period to reach steady state. For example, thyroid and renal concentrations exceed serum level, whereas red blood cell, spinal fluid, and brain concentrations do not. Lithium enters and leaves the CNS slowly, which may explain why acute overdoses with relatively high blood levels are sometimes well tolerated and why clinical manifestations of chronic intoxications often persist long after blood levels have decreased substantially.

Pharmacodynamics

Li, acting via glycogen synthase kinase-3 (GSK3β) inhibition, induces the transcription and expression of neurotrophic and neuroprotective proteins linked to the activation of cell survival signaling cascades. Li has also been shown to have other important cell signaling effects related to abnormalities previously linked to bipolar disorder (BD) pathogenesis, several of which have been studied in lymphoblastoid cell lines: Li appears to decrease intracellular inositol levels, subsequently preventing the formation of inositol phosphate 3 (IP3) and diacylglycerol, thus blocking

transmembrane signaling. Particularly, through inositol depletion, lithium reduces IP3-mediated calcium release from endoplasmic reticulum. Li may decrease the phosphorylation of cAMP-response-binding-protein (CREB), which interferes with expression of the genes regulated by this transcription factor, and modulates protein kinase A and protein kinase C pathways. Additional evidence linking Li effects to signaling pathways comes from genome wide association studies reporting BD-associated SNPs in the diacylglycerol kinase gene, which encodes a key protein in the lithium-sensitive phosphatidyl inositol pathway.

Lithium has neurotropic and neuroprotective activity that may extend its spectrum of use beyond bipolar disorder. For example, it has been shown to increase levels of Bcl-2 in various areas of rat brain, to induce expression of BDNF, to stimulate neurogenesis in adult rat hippocampus, and to increase total brain gray matter volume (by 3 percent) in eight of ten bipolar patients after 4 weeks of treatment.

Design and Interpretation of Clinical Drug Studies

The history of lithium encompasses most of the history of clinical drug trials in psychiatry. The research that established its antimanic activity to the satisfaction of the FDA was meager compared to today's large, double-blind, placebo-controlled parallel design, multicenter trials. Although the response rate consistently favored lithium over placebo, early studies were small, short term, and of crossover design. Subsequently, however, large double-blind, placebo-controlled, parallel designs using lithium as an active comparator to divalproex (Depakote and Depakote ER), aripiprazole (Abilify), and quetiapine (Seroquel) more firmly established its effectiveness as an antimanic agent.

In the 1960s, open studies by experienced clinicians comparing episode frequency before and during lithium treatment did much to establish the long-term effectiveness of the drug. Citing lack of scientific rigor in the design of these studies, critics cast doubt on the conclusion and stimulated conduct of double-blind, placebo-controlled studies (both prospective and discontinuation designs) that convincingly established the long-term prophylactic efficacy of lithium in bipolar disorder. Recent maintenance studies of lithium and other drugs have used time to recurrence as the main outcome measure. Less examined, and often not at all, is whether treatment reduces episode frequency, duration, and severity.

In the first randomized, placebo-controlled, 12-month maintenance study of lithium and divalproex designed not to enrich either active drug the two drugs were non-significantly superior to placebo on time to recurrence of any mood episode. Median time to 50 percent survival without a mood episode was 40 weeks for divalproex, 24 weeks for lithium, and 28 weeks for placebo. Another study showed that the combination of lithium and valproate was superior to valproate and trended to be superior

to lithium alone.

THERAPEUTIC INDICATIONS

Bipolar Disorders

Mania. Lithium is an established, first line, antimanic drug with effectiveness greater than placebo. Currently it is but one of many FDA-approved drugs for treating manic episodes, but it continues to be listed among the first choices. Its onset of action is relatively slow, with clinical improvement usually occurring over the first 1 to 3 weeks of treatment. Although a small open study suggested that oral lithium loading might produce improvement in a few days, this approach has yet to be substantiated. Instead, the concomitant use of a benzodiazepine or an antipsychotic drug is often necessary during the initial phase of treatment. Among antipsychotics, second generation drugs have largely replaced conventional antipsychotic drugs in terms of preference and recognition of somewhat more benign adverse effects. Since disturbed sleep may ignite or fuel a manic episode, early treatment of insomnia may be useful.

As with other treatments, not all patients with acute mania respond equally well to lithium. Response rates of around 80 percent in initial research trials contrast with findings of less success when patients with mixed or dysphoric mania, rapid cycling, comorbid substance abuse, or organic brain disorders are included. FDA-approved antimanic alternatives include divalproex, carbamazepine (Tegretol, Equetro), aripiprazole, olanzapine (Zyprexa), quetiapine, risperidone (Risperdal), and ziprasidone (Geodon). Combining lithium with any one of these is a common practice, supported by good evidence for divalproex, quetiapine, and olanzapine. In contrast, carbamazepine combined with olanzapine reduces olanzapine drug levels to therapeutically inadequate, therefore should be avoided. Whether another medication would be preferable to lithium as the initial treatment of mania, or at least for some subtypes of mania, has not been established in well-designed studies. Divalproex has become a preferred initial treatment of mixed mania.

Starting with the 1994, three arm study comparing valproate, lithium and placebo, most studies meet current methodological standards as comprehensively reviewed and published by the World Federation of Societies of Biological Psychiatry, whose expert guideline task force is international in membership. Lithium and valproate were significantly more effective than placebo in the 1994 study. Lithium was superior to placebo in studies for which quetiapine, topiramate, and aripiprazole were the investigational drugs. A meta-analysis of six randomized controlled trials including lithium reported a medium effect size for improvement with lithium of 0.4, which corresponded to a number needed to treat to see one patient significantly better than with placebo of 6 patients. In

secondary analyses of the studies that included valproate and quetiapine, lithium, valproate, and quetiapine were similarly superior to placebo, and to one another, in the subgroup of patients with psychotic symptoms.

Depression. Lithium is not approved by the FDA for the acute treatment of bipolar depressive episodes (in contrast to quetiapine, lurasidone and olanzapine/fluoxetine combination [Symbyax]). A large 2008 study did not show a difference versus placebo for low-dose lithium. A generally recommended approach for patients with bipolar depression who are already taking lithium is to increase dosage to the maximum tolerated level up to 1 mM/L either alone or, for more severe depressions, in combination with an antidepressant. This recommendation is mainly based on clinical experience. This strategy is to some extent contradicted by a study of Severus et al. indicating that high lithium serum levels are associated with an increased rate of relapse into bipolar depression. Treating bipolar depression with a conventional antidepressant alone does not protect against mania and may even precipitate mania or rapid cycling. Consequently, the presence of a mood stabilizer is a necessity in all but occasional low-risk bipolar II patients.

Antidepressants with some evidence for lower rates of worsening of manic symptoms include quetiapine, bupropion (Wellbutrin), lurasidone (Latuda), and MAOIs. Based on the results of one study (van der Loos et al., 2009), lamotrigine might be initiated if lithium optimization is unsuccessful.

If breakthrough depression occurs during lithium maintenance, lithium-induced hypothyroidism, poor compliance, and substance abuse must be considered as possible causes or contributors (Fig. 34.19–1). In the figure, derived from the two lithium, lamotrigine, placebo maintenance studies, note that patients taking lithium who required thyroid supplementation during the 18-month studies had essentially twice as high thyroid-stimulating hormone (TSH) levels as did lithium treated patients who did not require thyroid supplementation, whereas TSH was not elevated in subjects taking lamotrigine or placebo. If there is a margin of safety and tolerability, increasing the serum lithium level may be sufficient to bring the episode under control. Even if results of thyroid function testing are within normal limits, supplementing with thyroid hormone may be of value. There is some evidence that bipolar depression responds more slowly to antidepressant treatment in patients who have a lower free thyroxine index (FTI) and higher TSH levels, even within the range of normal. If optimizing the lithium blood level is ineffective, an antidepressant, a second mood stabilizer, or an atypical antipsychotic could be added. At present, many clinicians will follow the more conventional route and add an antidepressant, but others argue that two mood stabilizers might be preferable. One small short-term study found no

difference in response rates to bupropion or topiramate (Topamax) and no manic switches with either. In another small short-term study of patients on lithium or divalproex who had breakthrough depression, no difference in efficacy was found when either paroxetine (Paxil) or the alternate mood stabilizer (lithium or divalproex) was added. However, a placebo-controlled study of bipolar depression found no benefit from an antidepressant added to a mood stabilizer in maintaining 8 consecutive weeks of euthymia. All in all, the role that antidepressants should or should not play in the treatment of bipolar disorder remains a matter of debate. Since lamotrigine is beneficial for the prevention of recurrent bipolar depression with long-term treatment, combining it with lithium has appeal.

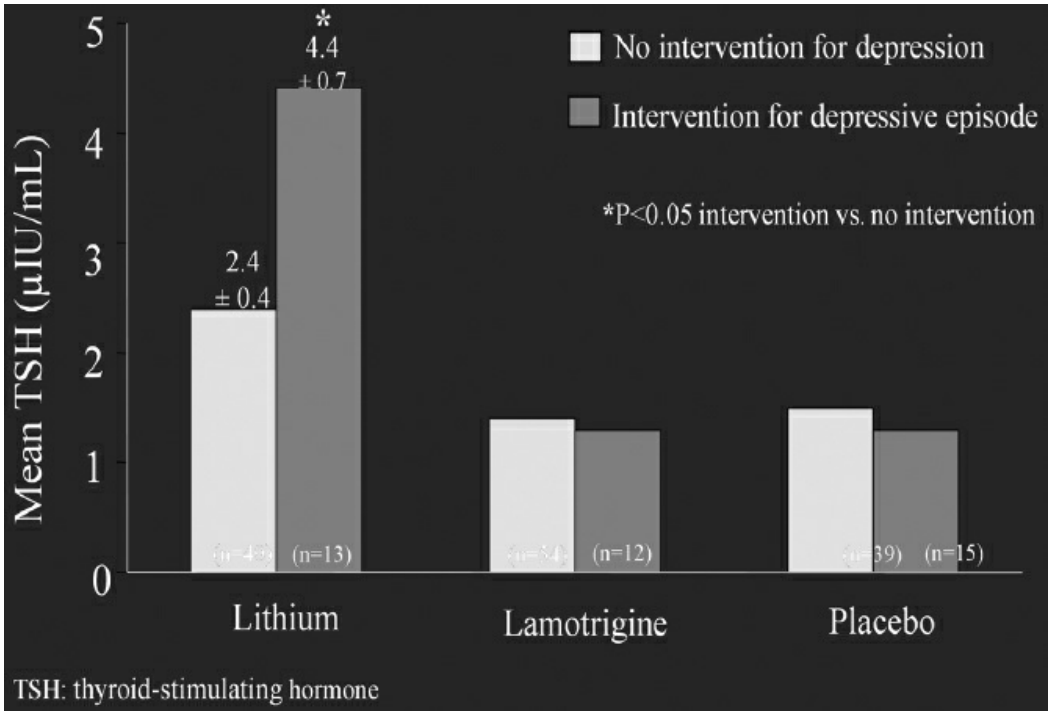

FIGURE 34.19–1. Elevated TSH levels from lithium predispose to development of depression. (From Frye M, Yatham L, Ketter TA, Goldberg J, Suppes T, Calabrese JR, et al. Depressive relapse during lithium treatment associated with increased serum thyroid-stimulating hormone: results from two placebo-controlled bipolar I maintenance studies. *Acta Psychiat Scan*. 2009;120(1):10–13.)

ECT is an effective treatment of bipolar depression, although it tends to be used only after medications have failed. Concern has been expressed that the presence of lithium during ECT increases the risk and severity of cognitive dysfunction, but whether this is true remains controversial.

Maintenance Therapy

Effective maintenance treatment, that is, reducing subsyndromal depression, mania and mixed states, along with preventing or minimizing the impact of new episodes, is the principal pathway to functional well-

being and productive lives for persons with bipolar disorders and all kinds of complications and disablement. Bipolar disorder ranks worldwide among the top ten of the most disabling disorders in adults.

Most long-term studies have used Kaplan–Meier (KM) survival analyses based on time to intervention. However, some studies use as study end point "any reason of failure" (inefficacy as indicated by new mood episodes or need for additional treatments or hospitalization, adverse events, withdrawal of consent, lost to follow up) as primary outcome. Some use discontinuation for emerging new mood episodes defined either by symptomatic *Diagnostic and Statistical Manual of Mental Disorders* (DSM) criteria or/and by clinical rating scale thresholds. An intrinsic problem with KM survival analytic techniques is that they measure the occurrence of a predefined event, for example, intervention, discontinuation, only at two time points: baseline (absence of the event) and end point (occurrence of event). This might be suitable if between these time points there is only one state possible, for example, "absolutely healthy." This is not the case in bipolar disorder, where subsyndromal fluctuations of mood, impairing functionality and quality of life, are more the rule than exception. In addition, other clinically valuable information, for example, tolerability and impact of medication on physical health will not be fully captured. Another issue in survival analysis is that the risk of censoring should be independent of the risk of the event in question, which most often is not the case. One reason why survival analysis has gained popularity in pivotal trials is that it is more sensitive for measuring differences than the more traditional counting of failures.

Based on trials that used both CGI-BP and YMRS and MADRS, it appears that a cut-off score of <5 on the MADRS and <4 on the YMRS approximates a CGI-BP of 1 for a meaningful definition of remitted status. Clinical meaningful remission is rarely achieved in published controlled trials; for example, in the lamotrigine/lithium long-term studies, remission as entry criterion for the double-blind phase was defined as having a CGI-S score of ≤3 for 4 consecutive weeks.

Most recent maintenance studies have used enriched discontinuation designs wherein the patient's acute symptoms had to respond to the given medication during open label treatment to the point of syndromal remission before randomization, resulting in sample "enrichment" for responders to the open label treatments allowed.

Long-term treatment with lithium is an effective way to reduce the frequency, severity and duration of manic, and, to a lesser degree, depressive episodes in patients with bipolar disorder. More recent placebo-controlled studies compared lithium to divalproex, lamotrigine, or olanzapine and utilized survival analysis to evaluate outcomes. In the two lamotrigine studies, lithium was superior to placebo in time to intervention for a mood episode and in time to a manic episode, but it did not differ

significantly in time to a depressive episode. The olanzapine study, which lacked a placebo control, found no significant overall difference, while prevention of manic relapse favored olanzapine. Treatment response to lithium monotherapy is usually less than complete, and supplemental pharmacotherapies are often necessary either intermittently or continuously. Nonetheless, there is good evidence that long-term lithium therapy reduces mortality rates. In general, bipolar disorder carries with it an increased risk of premature death, particularly but not entirely from suicide. Most studies comparing mortality rates in lithium-treated patients to the general population, to patients following discontinuation of lithium, and to patients treated with alternative prophylactic medications support the mortality-lowering and suicide-reducing effect of lithium. A meta-analysis of 31 studies of major affective disorders found a fivefold reduction in suicides and attempted suicides during lithium treatment compared to those not taking the drug. In a retrospective study, patients at high risk of suicidal behavior were grouped into excellent, moderate, and poor lithium responders. No further suicide attempts occurred in 93.3 percent, 82.7 percent, and 48.8 percent, respectively, suggesting that the antisuicidal effect of lithium extends beyond its mood stabilizing property (the beneficial role of psychosocial support could also be a contributing factor).

Work by Grof and others to predict a favorable response to lithium maintenance found that the diagnosis of a mood disorder, an episodic course with euthymic intervals, and the absence of rapid cycling were the main predictors. Additional factors that have been associated with inadequate response to lithium in maintenance use include many prior episodes, poor interepisode functioning, comorbid substance abuse, and comorbid personality disorder. It is likely that such limitations apply to other mood stabilizing agents, not just lithium. A double-blind study without enrichment favoring either active agent found no significant difference in rates of all cause termination in the lithium–valproate–placebo trial. However, secondary outcome indices, for example, termination for depression or for intolerable adverse effects favored valproate assigned patients (Fig. 34.19–2). The decision to apply long-term treatment with lithium should not be a difficult one in most circumstances. One would expect some unequivocal benefit in an initial acute illness state, most likely mania or hypomania. Unless the patient is tolerating the lithium well, it makes no sense to plan on its long-term use. However, there is sufficient evidence that in aggregate samples higher range dosage of lithium provides mostly worse tolerability, not greater healthiness. Licht et al. conducted an unusual study of 5 years duration commencing with randomization to either lithium or lamotrigine, but subsequently allowing any psychiatrist determined adjunctive medication use. As shown in Figure 34.19–3, a large majority of patients assigned to either drug required additional medication, generally within the first 2 years of follow-up

treatment.

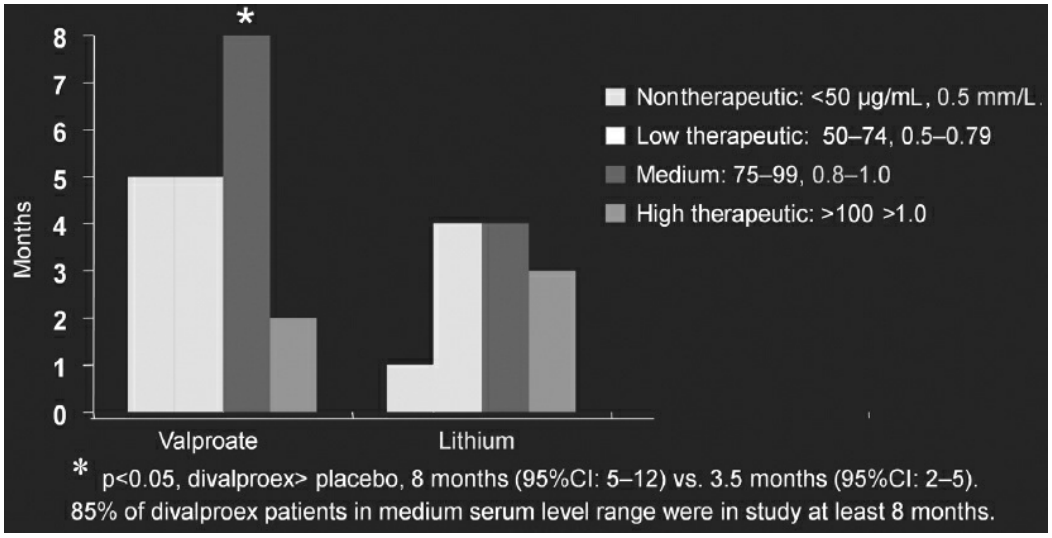

FIGURE 34.19–2. Months to a mood episode or early discontinuation in relationship to serum levels of valproate or lithium compared with placebo. (From Keck PE, Bowden CL, Meinhold JM, et al. Relationship between serum valproate and lithium levels and efficacy and tolerability in bipolar maintenance therapy. *Int J Psych Clin Pract*. 2005;9(4):271–277.)

The process of considering use of adjunctive or cotherapy for overall mood stabilization or for some bipolar illness aspect that is not within the spectrum of benefit of lithium is generally straightforward. If the symptom is outside the spectrum of benefit of lithium (or for that matter any other agent with which maintenance treatment is commenced), add what is. Even when exactly when to begin long-term lithium treatment has not been fully resolved, although most recommendations suggest earlier rather than later —possibly after only one episode of mania and certainly after two. Soft evidence suggests that episodes beget episodes (perhaps a form of kindling or behavioral sensitization), with each episode increasing the likelihood of the next and possibly shortening the interval to the next; hence, the recent trend toward earlier initiation of maintenance therapy. Obviously, many individual factors must be taken into consideration, including the abruptness, severity, and frequency of episodes; the presence of risk factors; and attitudes toward medication.

A patient had two manic episodes in 2 years, one resolving spontaneously and one with ECT, followed by 18 euthymic years without any treatment before the next episode occurred. Had she been treated with lithium after the second manic episode, the subsequent 18 years would have been considered a testimonial to the therapeutic effectiveness of lithium.

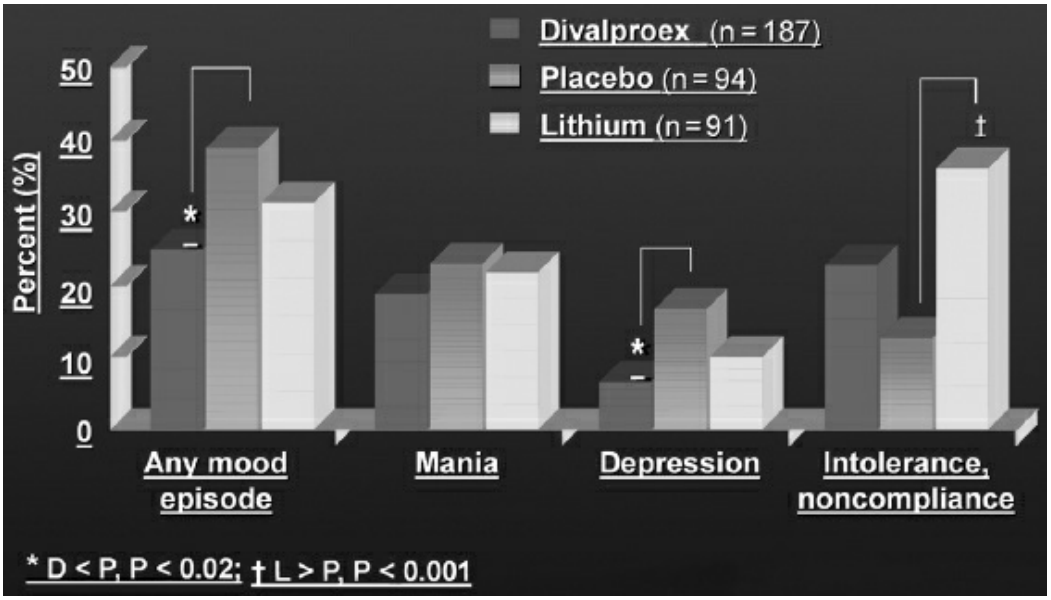

FIGURE 34.19–3. Early termination in 1 year randomized maintenance study. (From Bowden CL, Calabrese JR, McElroy SL, et al. A randomized, placebo-controlled 12-month trial of divalproex and lithium in treatment of outpatients with bipolar I disorder. Divalproex Maintenance Study Group. *Arch Gen Psychiatry*. 2000;57(5):481–489.)

The maximum benefits of lithium maintenance may not be immediate; with continued treatment, relapses sometimes become less severe and less frequent. Some patients appear to develop a tolerance to lithium after several years of successful use; this tolerance can sometimes be overcome by the addition of divalproex, carbamazepine, lamotrigine, or an atypical antipsychotic. Similarly, patients who are only partial responders to lithium maintenance may benefit from the addition of one of these drugs. Just as breakthrough depression may respond to a temporary increase in serum lithium concentration, so too may breakthrough mania. At such times, however, supplementing lithium with an antipsychotic drug or one targeted for sleep, anxiety control, or another discrete symptom will be desirable. A series of studies from 1983 through the present have consistently reported that serum lithium concentrations of 0.5 to 0.8 mmol/L range are at least as efficacious in maintenance as higher concentrations, and are much better tolerated and adhered to than higher ones. In a 6-month study (Fig. 34.19–4) among those randomized to lithium, only serum levels below 0.5 were associated with shorter time in study. No systematic, randomized study provides support for higher serum concentrations. Given the consistent evidence that lithium is frequently poorly tolerated, with consequent low adherence, attention to this strategy, particularly early in maintenance phase treatment should be routine.

The year 2000 valproate–lithium–placebo three arm study provides substantial clinically useful evidence consequent to its design, which per protocol allowed addition of a selective serotonin reuptake inhibitor (SSRI) if clinically significant depression occurred. The study shows the

proportion of patients who developed depression to have been highest with placebo and significantly less if receiving divalproex. The results or the study regarding profiles of maintenance benefit indicate that significantly fewer patients receiving valproate plus SSRI had early termination for any mood episode and for depression than occurred with placebo. Lithium resulted in nonsignificantly lower rates than placebo. This has been the only maintenance study to date that allowed treatment of breakthrough depression with an antidepressant (either sertraline or paroxetine). Valproate or lithium plus an SSRI provided longer time in study without discontinuation for depression than did placebo plus an SSRI. Fewer patients discontinued prematurely among valproate plus SSRI-treated patients than among placebo-treated patients. These results indirectly suggest that the combination of valproate and an SSRI in acute bipolar depression is more effective than SSRI monotherapy.

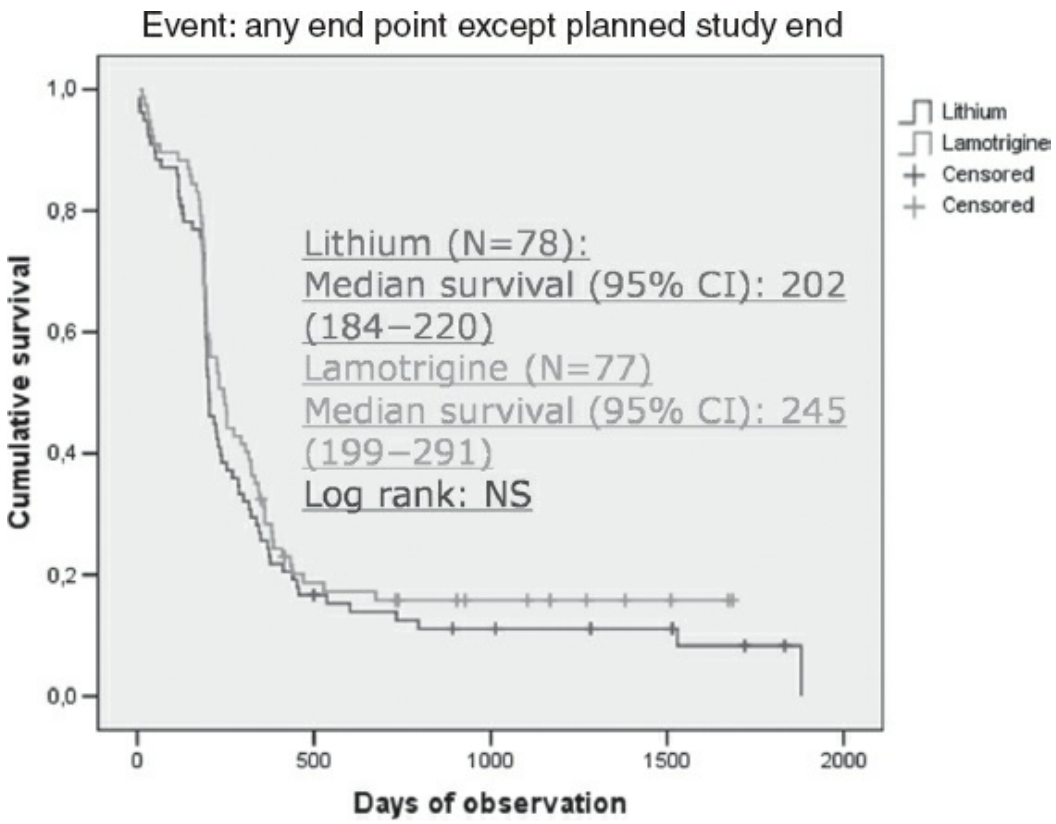

FIGURE 34.19–4. Danish study of lithium versus lamotrigine monotherapy maintenance. (From Licht RW, Nielsen JN, Gram LF, Vestergaard P, Bendz H. Lamotrigine versus lithium as maintenance treatment in bipolar I disorder: an open, randomized effectiveness study mimicking clinical practice. The 6th trial of the Danish University Antidepressant Group (DUAG-6). *Bipolar Disord*. 2010;12(5):483–493.)

Both lithium and lamotrigine were associated with longer time in remitted status and lower time in subsyndromal mania. Lithium, but not lamotrigine, was also associated with significantly less time in syndromal and subsyndromal mania. Neither lithium nor lamotrigine reduced time

depressed compared with placebo. Neither lithium nor lamotrigine differed from placebo on the proportion of days with subsyndromal or syndromal depression or mixed states, which constituted almost one quarter of time in study for both drugs. This analysis by Multiple Outcome Assessment of Treatment methodology discussed earlier in this chapter also can be combined by latent class analysis with indices of tolerability to provide an aggregate measure of benefit and tolerability.

Lithium was associated with both significant benefit in time manic and worse tolerability compared to placebo. In summary, lamotrigine was associated with limited therapeutic benefit but not harm; lithium with both benefit and harm; and placebo with neither benefit nor harm.

The multi-state outcome analysis of treatments (MOAT) analytic approach that the authors developed to overcome limitations of survival analyses yielded the results displayed in Table 34.19–1. Both lithium and lamotrigine provided a significantly longer term in remitted status than did placebo. Lithium also provided significantly less time in a manic state. When syndromal and subsyndromal manic times were combined, the lithium superiority versus placebo continued. Lithium randomized patients did not differ for any facet of depressive time or time in a mixed state. Lamotrigine trended to less time manic but not after taking into account multiple comparisons.

Long-term treatment does not cure bipolar disorder. Following discontinuation of successful lithium therapy, the risk of recurrence increases substantially. For example, one review of discontinuation studies of patients with bipolar disorder found the risk of recurrence to be 23 times higher subsequent to discontinuing. In addition, rapid discontinuation of lithium appears to be associated with a rebound phenomenon characterized by increased risk of early recurrence, especially mania. One study found in the first year after stopping lithium that rapid discontinuation (1 to 14 days) was associated with twice as many manic and depressive episodes as was gradual discontinuation (15 to 30 days).

Table 34.19–1.
MOAT Estimates of Mean State Durations ± Bootstrap Standard Error

State	Mean Duration ± Standard Error	Cochran's F-value (P-value)	PBO Lithium	LTG
Remitted	99.0 ± 8.2	131.3 ± 9.7	137.0 ± 7.4	6.38 (0.002)[a]
SS mania	11.8 ± 2.5	5.4 ± 1.1	8.5 ± 1.3	3.70 (0.026)[b]
SS depression	44.1 ± 4.9	53.4 ± 6.6	55.7 ± 4.9	1.52 (0.220)
SS mixed	7.2 ± 1.9	6.0 ± 2.0	6.6 ± 1.5	0.09 (0.913)
SYN mania	6.5 ± 1.3	3.8 ± 1.2	5.0 ± 1.0	1.13 (0.325)
SYN depression	16.4 ± 2.3	17.3 ± 2.9	16.0 ± 1.9	0.08 (0.924)
SYN mixed	5.9 ± 1.5	4.2 ± 1.5	4.4 ± 1.2	0.40 (0.667)

Totals				
SYN/SS mania	18.3 ± 2.8	9.2 ± 1.8	3.5 ± 1.8	3.91 (0.021)[c]
SYN/SS depression	60.5 ± 5.4	70.7 ± 8.1	71.6 ± 5.5	1.18 (0.308)
SYN/SS mixed	13.1 ± 2.8	10.2 ± 2.8	11.0 ± 2.0	0.29 (0.746)
Total days	190.8 ± 8.9	221.4 ± 9.5	233.2 ± 8.9	6.00 (0.003)[d]

Note: Cochran's F-value has a numerator df = 2 and the denominator df range from 351 to 377. *P*-values are unadjusted for multiple testing.

Pairwise comparisons: [a] Lithium-PBO $z = 2.53$, $p = 0.011$; LTG-PBO $z = 3.43$, $p = 0.0006$.

[b]Lithium-PBO $z = -2.38$, $p = 0.018$.

[c]Lithium-PBO $z = -2.71$, $p = 0.007$.

[d]Lithium-PBO $z = 2.35$, $p = 0.019$; LTG-PBO $z = 3.37$, $p = 0.0008$.

LTG, lamotrigine; MOAT, multi-state outcome analysis of treatments; PBO, placebo; SS, subsyndromal; SYN, syndromal.

Major Depressive Disorder

Although episodes of major depressive disorder may respond to lithium alone, traditional antidepressant medications are considerably more effective and are the preferred treatment. The major value of lithium in major depressive disorder patients with acute depression is as an augmenting agent when antidepressants alone have been ineffective. Approximately half of patients respond when lithium is added to a wide variety of antidepressant drugs. Favorable responses have been described in both psychotic and nonpsychotic depression. Benefit has been reported with most all antidepressants, but evidence is most convincing with tricyclics. Although alternative explanations may apply in some situations (spontaneous remission), few negative studies for this use have been published. Experts disagree, however, as to where lithium augmentation should fall in the hierarchy of treatment resistance strategies.

Schou stated: "Doubt about the efficacy of prophylactic lithium treatment in unipolar illness seems almost exclusively a U.S. phenomenon." Indeed, most studies have found lithium to be more effective than placebo and as effective as antidepressant drugs for the long-term management of major depressive disorder. This may reflect the conceptualization that highly recurrent major depressive disorder and bipolar disorder are both components of manic-depressive illness. Lithium has not, however, demonstrated clear advantages over conventional antidepressant drugs, and since long-term treatment is most likely to involve the drug that ended the acute episode (an antidepressant), lithium maintenance for major depressive disorder is likely to be reserved for antidepressant treatment failures. However, a recent meta-analysis found support for antisuicidal effects of lithium in patients with recurrent major depressive disorder.

Schizoaffective Disorder and Schizophrenia

In general, the less affective and the more schizophrenic an illness is, the less likely it is to respond to lithium. The same cannot be said of an episode because the acute manifestations of mania and schizophrenia may be indistinguishable. Although not approved by the FDA for schizoaffective disorder or the revisions in DSM-5, lithium is generally accepted to be of value for treating this condition, especially in combination with antipsychotic drugs, and especially if the affective component is prominent.

Antipsychotic drugs remain the treatment of choice for schizophrenia, although patient response is often limited. The addition of lithium to an antipsychotic agent remains a treatment strategy for antipsychotic-resistant schizophrenia, although the availability of atypical antipsychotics such as clozapine (Clozaril), aripiprazole, olanzapine, quetiapine, paliperidone (Invega), risperidone, lurasidone, iloperidone, and ziprasidone makes this less necessary. In addition, while several small early studies suggested benefit from lithium augmentation of antipsychotic medication, more recent data have been less convincing. Should there be a coexisting mood disorder, lithium might play a more prominent role. A 2007 Cochrane Collaboration review concluded: "there is no evidence that lithium on its own is effective for people with schizophrenia or schizoaffective disorder. There is some evidence for the effectiveness of lithium as an adjunctive treatment to antipsychotic drugs, but this result was inconclusive."

Aggression

Lithium appears to have antiaggressive effects independent of its mood-stabilizing action. Several controlled studies have found reduced aggression in subjects recruited from prison populations. Behavioral dyscontrol and self-mutilation in intellectual disabled patients have also been successfully treated with lithium. However, lithium has not been used extensively to treat aggression associated with head trauma or epilepsy, and the results have been mixed. No work has been done to evaluate the effect of lithium in those patients defined by DSM-IV-TR as having intermittent explosive disorder. Aggressive outbursts are also treated with medications such as anticonvulsants, β-adrenergic receptor antagonists, benzodiazepines, buspirone (BuSpar), antipsychotics, and antidepressants. The relative ranking of lithium among those other drugs must be determined on an individual basis.

Alcohol-Use Disorders

Lithium's potential for treating alcohol dependence and abuse is based in part on the close association between mood disorders and alcohol-use disorders and in part on animal research that found reduced alcohol intake

in rodents receiving lithium. Efforts to establish lithium as a useful treatment of alcoholism have been largely unsuccessful. A large (N = 457), multicenter, double-blind, placebo-controlled 1-year study found that lithium did not alter the course of alcoholism in depressed or nondepressed alcoholics. Although not all research reports have been that discouraging, lithium is unlikely to play a major role in the treatment of alcoholism. Some patients may be exceptions, but clinicians have found no way to identify them prospectively.

Other Indications

The philosophy that "lithium is good for what ails you" led to its use in a wide variety of psychiatric (Table 34.19–2) and nonpsychiatric medical conditions (Table 34.19–3). In general, reports in these areas involve small numbers of patients in studies often lacking in scientific rigor. In such situations, both positive and negative outcomes may be inconclusive and even misleading. For example, two double-blind, placebo-controlled trials found lithium ineffective in OCD. Only 16 patients were involved and the treatment duration of 2 weeks was insufficient to evaluate the drug adequately. Two more recent well-designed, placebo-controlled studies of adequate duration found lithium to be ineffective in augmenting conventional drugs for OCD in treatment-resistant patients.

Table 34.19–2.
Psychiatric Uses of Lithium

Historical
Gouty mania
Well established (FDA-approved)
Manic episode
Bipolar maintenance therapy
Reasonably well established
Bipolar disorder
Depressive episode
Bipolar II disorder
Rapid-cycling bipolar I disorder
Cyclothymic disorder
Major depressive disorder
Acute depression (as an augmenting agent)
Maintenance therapy
Schizoaffective disorder
Evidence of benefit in particular groups
Schizophrenia
Aggression (episodic), explosive behavior, and self-mutilation
Conduct disorder in children and adolescents
Mental retardation
Cognitive disorders
Prisoners
Anecdotal, controversial, unresolved, or doubtful

Alcohol and other substance-related disorders
  Cocaine abuse
  Substance-induced mood disorder with manic features
Anxiety disorders
  Obsessive–compulsive disorder
  Phobias
  Posttraumatic stress disorder
Attention-deficit/hyperactivity disorder
Eating disorders
  Anorexia nervosa
  Bulimia nervosa
Impulse-control disorders
Kleine–Levin syndrome
Mental disorders due to a general medical condition (e.g., mood disorder due to a general medical condition with manic features)
Periodic catatonia
Periodic hypersomnia
Personality disorders (e.g., antisocial, borderline, emotionally unstable, and schizotypal)
Premenstrual dysphoric disorder
Sexual disorders
  Transvestic fetishism
  Exhibitionism
  Pathological hypersexuality

Efforts to establish lithium as a useful treatment of personality disorders have been largely unsuccessful. When favorable outcomes do occur, it is likely that a comorbid mood disorder has responded followed by indirect improvement in "personality." Following FDA approval of lithium for mania lithium was used with some success to treat inpatients with emotionally unstable character disorder. The findings were never replicated or extended to outpatients, and the disorder failed to appear in subsequent editions of DSM. Today, the substantial mood component of borderline personality disorder suggests a potential role for lithium in its treatment, but controlled trials have not been conducted.

Table 34.19–3.
Nonpsychiatric Uses of Lithium[a]

Historical
Gout and other uric acid diatheses
Lithium bromide as anticonvulsant
Neurological
Epilepsy
Headache (chronic cluster, hypnic, migraine, particularly cyclic)
Ménière disease (not supported by controlled studies)
Movement disorders
Huntington disease
Levodopa-induced hyperkinesias
On-off phenomenon in Parkinson disease (controlled study found decreased akinesia, but development of dyskinesia in a few cases)
Spasmodic torticollis

Tardive dyskinesia (not supported by controlled studies, and pseudo-parkinsonism has been reported)
Tourette disorder
Pain (facial pain syndrome, painful shoulder syndrome, fibromyalgia)
Periodic paralysis (hypokalemic and hypermagnesic but not hyperkalemic)
Hematological
Aplastic anemia
Cancer—chemotherapy-induced and radiotherapy-induced neutropenia (one study found increased risk of sudden death in patients with pre-existing cardiovascular disorder)
Drug-induced neutropenia (e.g., from carbamazepine, antipsychotics, immunosuppressives, and zidovudine)
Felty syndrome
Leukemia
Endocrine
Thyroid cancer, as adjunct to radioactive iodine
Thyrotoxicosis
Syndrome of inappropriate antidiuretic hormone secretion
Cardiovascular
Antiarrhythmic agent (animal data only)
Dermatological
Genital herpes (controlled studies support topical and oral use)
Eczematoid dermatitis
Seborrheic dermatitis (controlled studies support)
Gastrointestinal
Cyclic vomiting
Gastric ulcers
Pancreatic cholera
Ulcerative colitis
Respiratory
Asthma (controlled study did not support)
Cystic fibrosis
Other
Bovine spastic paresis

[a]All the uses listed here are experimental and do not have FDA-approved labeling. There are conflicting reports about many of these uses, some have negative findings in controlled studies, and a few involve reports of possible adverse effects.

There has been considerable interest in lithium's potential for treating a wide variety of nonpsychiatric medical conditions (Table 34.19–3). The granulocyte-stimulating effect of the drug has been used to increase white blood cell count in conditions such as Felty syndrome, idiopathic neutropenia, chemotherapy-induced granulocytopenia, clozapine-induced neutropenia, and immunosuppressant drug-induced leukopenias. Lithium's antithyroid effects have been utilized investigationally to treat hyperthyroidism. Neurologists have found lithium useful as a treatment of chronic and episodic cluster headaches. The observation that lithium inhibits replication of DNA viruses such as human herpes virus 1 and 2 led to successful double-blind, placebo-controlled trials of both topical and oral preparations to treat herpes genitalis. Seborrheic dermatitis has also been shown in controlled studies to respond to a lithium-containing ointment that is marketed outside the United States.

PRECAUTIONS

Children and Adolescents

At least one-third of patients with bipolar I disorder have their first episode before the age of 20. This percentage may be even higher if age of onset is dated from first appearance of symptoms rather than from first well-defined manic or depressive episode or first hospitalization. Although adolescent onset is common, prepubertal onset is far less so, but its prevalence may be underestimated because of atypical presentations and reluctance to make the diagnosis in children.

In general, diagnostic criteria for bipolar disorder in the young are the same as in adults; however, atypical presentations (especially in younger children) and comorbidity, especially with attention-deficit/hyperactivity disorder (ADHD) and conduct disorder, may make recognition more difficult. Furthermore, extrapolating adult diagnostic criteria to youth may be problematic. Lack of a well-established longitudinal history further complicates the diagnosis and may contribute to miscategorizing first-episode psychotic mania as schizophrenia.

Another diagnostic issue is whether a first episode of depression will follow a bipolar or unipolar course. The following factors may be predictive of a bipolar outcome and influence early treatment with lithium or another mood stabilizer: (1) rapid onset with psychosis and psychomotor retardation, (2) family history of bipolar disorder, and (3) antidepressant-induced hypomania. One study followed 72 children diagnosed with prepubertal major depressive disorder into early adulthood and found that a third had bipolar I disorder and an additional 15 percent had bipolar II disorder or hypomania.

Despite a substantial literature on the use of lithium to treat children and adolescents with conditions such as bipolar disorder, aggression, and behavioral problems associated with intellectual disability and developmental disorders, there remains a paucity of well-controlled studies. Nevertheless, lithium is FDA-approved for treating mania in adolescents (ages 12 years and older). The range of serum lithium concentrations in adolescents is similar to that in adults (although its elimination half-life may be shorter) and the likelihood of responding appears the same. The adverse effect profile of lithium is also the same across age groups, although in view of adolescent concern about body image, adverse events such as acne and weight gain may be poorly tolerated. Also, nontoxic cognitive dulling induced by therapeutic amounts of lithium may adversely impact academic performance.

Whether to begin long-term lithium therapy is an even more difficult decision when treating young patients than when treating adults. The markedly disruptive effects of episodes in youth and the highly recurrent nature of bipolar disorder must be balanced against the potentially

deleterious adverse effects of long-term lithium.

Elderly Patients

In the absence of organicity and substance abuse, most evidence suggests that advanced age alone does not compromise responsiveness to lithium. However, the use of lithium in the elderly is complicated by factors that include associated medical illnesses and medications, special diets, age-related reduction in GFR, and increased sensitivity to adverse effects.

Whether the elderly as a group respond to lower serum concentrations of lithium than do their younger counterparts is not known. Increased sensitivity to adverse effects and toxicity confine many elderly patients to relatively low doses and blood concentrations, which may prevent them from benefiting from the drug. However, this is not always the case, as demonstrated in the case study presented here.

After 10 hospitalizations for mania and depression beginning at age 20, a 53-year-old woman was stabilized on 1,500 mg a day of lithium carbonate and remained on this regimen for the next 15 years with serum concentrations ranging between 0.6 and 0.75 mEq/L. It was only at age 68 that an increase in serum concentration led to a dose reduction to 1,200 mg daily, and at age 70 to a further reduction to 900 mg per day.

In general, elderly patients should be started on lower-than-usual dosages, with dosage changes occurring less frequently than in younger patients. The elimination half-life of lithium increases with age, and the time required to reach steady state is much longer in the elderly. If lithium is stopped, serum levels fall more slowly and the resolution of adverse effects and toxicity may be prolonged.

A 60-year-old man was treated with 900 mg a day of lithium carbonate. The dosage was continued unchanged for 10 years, despite laboratory evidence of gradually increasing serum lithium and creatinine concentrations. Even as clinical symptoms of toxicity were being reported, a thiazide diuretic was added to treat hypertension. Three weeks later, the patient was hospitalized with a serum lithium concentration of 4.2 mEq/L and marked neurological impairment that never fully resolved.

With appropriate monitoring and compliant patients, lithium use in the elderly can be both safe and effective. Unfortunately, there continues to be a lack of randomized controlled studies of any medication in elderly bipolar patients. Whether the putative neuroprotective effects of lithium would have practical utility in this population remains an open question.

Pregnant Patients

Lithium is in FDA pregnancy category D, which includes drugs for which there is evidence of human fetal risk, but whose potential benefits may outweigh the risk in some pregnant women. Almost three decades ago, the

International Register of Lithium Babies reported an 11 percent incidence of major malformations in babies exposed to lithium during the first trimester, with a disproportionate representation of cardiovascular malformations, especially Ebstein anomaly of the tricuspid valve. These data exaggerate risk because of the greater likelihood of reporting abnormal outcomes, failing to control for other medications and medical conditions, and failing to control for the effect of the illness itself on outcome. Recent studies have not confirmed such a high incidence of malformations, although there still appears to be an increased risk of cardiovascular anomalies in humans (the incidence of Ebstein anomaly is between 1 and 2 per 1,000 which is 10 to 20 times greater than in the general population). Although the teratogenic risk of lithium is still greater than that found in the general population, it is lower than that posed by carbamazepine and valproate. Fetal echocardiography is advised to screen for cardiovascular malformations in women exposed to lithium during the first trimester of pregnancy.

Physiological changes accompanying pregnancy alter maternal lithium metabolism. The GFR increases 30 to 50 percent over baseline, and plasma volume increases by 50 percent. The filtered sodium load increases markedly (5,000 to 10,000 mEq daily), as does the renal tubular reabsorption of sodium, resulting in cumulative sodium retention of close to 1,000 mEq. Complications of pregnancy, such as hypertension and edema and their treatments, may further complicate lithium therapy. Polyhydramnios in a pregnant woman taking lithium was ascribed to lithium-induced fetal polyuria. Following delivery, the altered renal physiology of pregnancy returns rapidly to normal. These factors and their variability among individual patients dictate that lithium use during pregnancy be closely supervised. Fetal and maternal blood concentrations are similar, so women should receive minimum effective dosages. To reduce the risk of toxicity in the newborn, clinicians should markedly reduce or possibly temporarily discontinue the drug shortly before delivery. In view of evidence that abrupt or rapid discontinuation of lithium is associated with a high risk of relapse (which may be increased further postpartum), dosage reduction rather than discontinuation may be more appropriate as delivery approaches. Full prophylactic concentrations should be re-established following delivery because this is a time of great vulnerability to recurrence. Because lithium appears in breast milk and in the blood of breast-fed babies, the American Academy of Pediatrics Committee on Drugs recommends that lithium be used with caution in nursing mothers. One study of 10 breastfeeding infants found serum levels to average 0.16 mEq/L (range 0.05 to 0.23). The merits of breastfeeding are considerable, and actual reports of infant lithium toxicity are limited to one or two cases; consequently, in some situations the benefits outweigh the risks.

Medically Ill Patients

The patient with bipolar disorder and coexisting medical illnesses presents a particularly difficult therapeutic challenge. Untreated or inadequately treated mania or depression can adversely affect a medical illness (e.g., mania and congestive heart failure) and decrease treatment adherence, and lithium may be more difficult or impossible to use in the presence of a medical illness (e.g., severe renal disease). Because medically ill patients are receiving other medications, their risk of adverse drug interactions is increased. Consequently, treatment of such patients requires close collaboration between psychiatric and nonpsychiatric physicians.

Such was not the case when an elderly woman became lethargic, confused, and anorectic after her internist prescribed a thiazide diuretic and her serum lithium level increased from 0.5 to 1.1 mEq/L.

Rapid Cycling Patients

Rapid cycling has been defined as four or more mood disorder episodes per year. Up to 20 percent of bipolar disorder patients experience rapid cycling, the majority of whom are women. Risk factors associated with rapid cycling include a depression-hypomania course, antidepressant drug therapy, thyroid abnormalities, and neurological disease. Rapid cycling appears not be a persisting condition in the majority of patients who experience it. Ultra-rapid cyclers have 24- to 48-hour cycles, and the concept of ultra-ultra rapid cycling has been introduced to describe patients with frequent mood shifts within a 24-hour period. Given increasing evidence that mood lability (also named affective lability) is increasingly recognized as a major symptom in bipolar disorder, some of what was previously labeled as rapid cycling subtype may be more fundamentally be viewed as a prominent bipolar symptom. Affective lability is closely associated with mixed states.

Overall, rapid-cycling appears to be more difficult to treat with most drugs and regimens, including lithium. Prior treatment with antidepressant drugs may predict a poorer response to lithium. Also, because antidepressants may induce rapid cycling, simply discontinuing the antidepressant agent may be sufficient to slow the cycling pattern (although at times worsening depression is the result). Correcting thyroid abnormalities is an important aspect of treating rapid cycling. Preliminary research suggests that hypermetabolic doses of thyroid hormone may be useful in dealing with treatment-resistant rapid cycling. If lithium alone is ineffective, treatment considerations include some anticonvulsants (particularly lamotrigine and valproate) either alone or combined with lithium. Atypical antipsychotics are additional considerations, as is ECT in more extreme cases.

ADVERSE REACTIONS

Adverse effects are to be expected during lithium therapy. Fewer than 20 percent of patients have no adverse effects, but only about 30 percent have more than minor complaints. The incidence of severe adverse effects may be underestimated, however, because patients who do not tolerate therapy may not be available for comment. Recognizing and minimizing adverse effects can do much to enhance compliance with lithium treatment (Table 34.19–4).

Neurological Reactions

Patients often complain of dysphoria, lack of spontaneity, slow reaction times, intellectual inefficiency, and impairment of memory. These complaints are a common cause of poor adherence. In assessing such complaints, psychiatrists should also assess for other contributors to cognitive impairment including depression, lithium-induced hypothyroidism or hypercalcemia, other illnesses, and other drugs. Single daily dosing at bedtime or a dosage reduction may be helpful, although switching to another mood stabilizer may be necessary.

After 19 years of successful lithium maintenance, a 37-year-old woman was switched to carbamazepine because of polyuria and proteinuria. Only in retrospect did she and her physician become aware of the cognitive dulling and impaired concentration that had been present during her many years on lithium.

Table 34.19–4.
Adverse Effects of Lithium

Neurological
Generally benign, nontoxic: Dysphoria, lack of spontaneity, slowed reaction time, memory difficulties
Tremor: Postural, occasional extrapyramidal
Toxic: Course tremor, dysarthria, ataxia, neuromuscular irritability, seizures, coma, death
Miscellaneous: Peripheral neuropathy, benign intracranial hypertension, myasthenia gravis-like syndrome, reduced creativity, lowered seizure threshold
Endocrine
Thyroid: Goiter, hypothyroidism, exophthalmos, hyperthyroidism (rare)
Parathyroid: Hyperparathyroidism, adenoma
Cardiovascular
Benign T-wave changes, sinus node dysfunction
Renal
Concentrating defect, morphological changes, polyuria (nephrogenic diabetes insipidus), reduced glomerular filtration rate, nephrotic syndrome, renal tubular acidosis
Dermatological
Acne, hair loss, psoriasis, rash
Gastrointestinal
Appetite loss, nausea, vomiting, diarrhea
Miscellaneous
Altered carbohydrate metabolism, weight gain, fluid retention

Tremor. The most common tremor associated with lithium use is a

postural tremor in the arms and hands. Because the tremor worsens during activities requiring fine motor control, it can be socially embarrassing and occupationally troublesome. Tremor sometimes improves spontaneously, but if it does not, benefit may be obtained from dose reduction, use of a slow-release lithium preparation, elimination of dietary caffeine, discontinuation of other medications, and treatment of associated anxiety. Medications useful in treating lithium tremor include β-adrenergic receptor antagonists such as propranolol (Inderal), as well as primidone (Mysoline), and possibly gabapentin (Neurontin). With long-term lithium therapy, a tremor with parkinsonian characteristics may occur occasionally. Worsening of tremor at any time during the course of lithium therapy may be an indication of impending lithium intoxication, and severe tremor should be considered due to lithium toxicity until proven otherwise.

Other Nontoxic Effects. Lithium use has been associated rarely with peripheral neuropathy, downbeat nystagmus, benign intracranial hypertension (pseudotumor cerebri), a myasthenia gravis–like syndrome, and lowering of the seizure threshold.

Lithium Intoxication

Lithium intoxication is primarily a neurotoxicity that can lead to death or permanent neurological damage (often cerebellar as characterized by dysarthric speech, tremor, and wide-based gait). Cardiovascular, gastrointestinal, and renal manifestations may also be present. Factors contributory to toxicity include, of course, excessive intake (accidental or deliberate), reduced excretion, kidney disease, low-sodium diet, drug interaction, reduced volume of distribution (dehydration), and individual sensitivity (e.g., elderly and the neurologically impaired). Absorption can sometimes be delayed substantially if slow-release preparations are involved.

A woman with lithium-induced nephrogenic diabetes insipidus was hospitalized for the treatment of a manic episode. Antipsychotic drug-induced sedation impaired her ability to replace fluids, and she quickly became lithium toxic and hypernatremic. Permanent cerebellar damage was the result.

Severe toxicity has been described in patients whose serum concentrations were within the so-called therapeutic range of .5 to 1.5 mEq/L. Some patients have tolerated high concentrations after deliberate overdoses without noticeable clinical consequences. In general, however, the higher the concentration and the longer the exposure, the more serious the intoxication. Early clinical manifestations include gastrointestinal upset, dysarthria, ataxia, and coarse tremor. Ominous findings include impaired consciousness, neuromuscular irritability

(fasciculations, myoclonus), seizures, and coma. Some cases of lithium intoxication have resembled Creutzfeldt–Jakob disease, nonconvulsive status epilepticus, or neuroleptic malignant syndrome.

Treatment is directed at the removal of lithium from the body. In mild cases, nothing more may be necessary than discontinuation of intake. After an overdose, gastric lavage is indicated and may have to be repeated. Although activated charcoal is not helpful following an acute lithium overdose, preliminary reports using polystyrene sulfonate (Kayexalate), a cation exchange resin, or whole bowel irrigation with polyethylene glycol solution (GoLYTELY) have been promising. In the presence of normal renal function, mild-to-moderate toxicity often responds to correcting dehydration and maintaining proper fluid and electrolyte balance. Whether forced diuresis provides additional benefit is open to debate. Hemodialysis is the treatment of choice for severe intoxication (peritoneal dialysis is considerably less efficient). Lithium is removed effectively by dialysis because it is water soluble, unbound to plasma protein, without metabolites, and of simple chemical structure (an element). Redistribution of lithium from tissues to blood after dialysis usually results in a rebound increase in its blood level; this may necessitate further dialysis. Hemodialfiltration has also been used, sometimes in conjunction with hemodialysis to minimize the likelihood of rebound. Because of lithium's tissue-toxic effects and gradual removal from the CNS, clinical improvement may lag several days behind the lowering of serum concentration.

Thyroid Impairment

Transient mild abnormalities in thyroid function testing are common early in the course of lithium treatment but are usually of little or no clinical consequence. Some patients, however, develop goiter or clinical hypothyroidism sometime during the course of treatment. Women, persons with pre-existing thyroid dysfunction, and those from iodine deficient areas are more than usually susceptible. Among lithium's many effects on thyroid function is its impedance of release of hormone from the gland. The goiter that has been described in about 5 percent of patients taking lithium (prevalence figures vary widely) is rarely of cosmetic or obstructive importance; however, an ultrasound study found thyroid gland enlargement in 44 percent of patients on lithium for 1 to 5 years in contrast to 16 percent of a control group.

Clinical hypothyroidism prevalence is directly associated with serum lithium concentration. Once diagnosed, it can be treated with supplemental levothyroxine (Synthroid) at a dosage that returns the TSH concentration to normal. Subclinical hypothyroidism (elevated TSH, normal free thyroxine) is considerably more common than clinical hypothyroidism. As noted earlier, elevated TSH predisposes to clinical depression. Substantial

elevations of TSH are likely to progress to clinical hypothyroidism and thus should be treated with exogenous thyroid hormone. Of course, the principal action warranted is reduction in lithium dosage. Subclinical hypothyroidism may contribute to reduced response to the treatment of bipolar depression.

Exophthalmos and hyperthyroidism have been described during lithium therapy. Whether these occur more often than in the general population is not established. Baseline assessment (including antithyroid antibody titers) and periodic monitoring of thyroid function (usually measurement of serum TSH concentration with or without triiodothyronine [T_3] and thyroxine [T_4] concentrations) is an integral part of lithium therapy. Whether it is done routinely (every 6 or 12 months) or only when clinical suspicion is aroused is a matter of preference.

Cardiovascular Reactions

Benign, reversible T-wave changes on the ECG are common during lithium therapy and are of no clinical consequence. Varying degrees of heart block and bradyarrhythmias as well as QTc prolongation and T-wave inversion have been reported during lithium intoxication and occasionally during its therapeutic use. Electrophysiological studies show that lithium can impair sinus node function, and sinus dysrhythmias, sometimes associated with syncopal episodes, have been reported in a few patients. Lithium does not have clinically important effects on blood pressure.

Although lithium is not contraindicated in the presence of cardiovascular disease, it should be used with caution. A pacemaker may be necessary if lithium is used in patients with clinically significant sinus node disease. Low-salt diets, certain diuretics, ACE inhibitors, angiotensin II receptor type-1 antagonists, fluid-electrolyte imbalances, and impaired renal function all predispose to lithium toxicity.

Renal Reactions

Lithium has complex effects on renal function. In the majority of patients, lithium impairs renal concentrating ability, which in itself is of no clinical importance. Nephrogenic diabetes insipidus occurs in around a third of patients taking lithium. The concentrating defect is not always reversible after lithium discontinuation, suggesting that both functional and structural changes have occurred in the distal tubules and collecting ducts.

Polyuria is the most clinically troublesome renal effect of lithium. Figures vary considerably, but a 24-hour urine volume of more than 3 L (1 to 2 L is normal) has been reported in as many as 35 percent of patients taking lithium. Lithium increases urine volume primarily by inhibiting the effect of antidiuretic hormone on the kidneys, although additional renal (including disruption of the aquaporin-2 shuttle) and CNS mechanisms

have been identified. When severe, polyuria can be socially and occupationally compromising; a cause of insomnia, weight gain (through consumption of high-calorie beverages), poor nutrition, and noncompliance; and potentially dangerous if dehydration occurs. Treatment considerations include (1) adequate fluid replacement, (2) use of the lowest effective dosage, and (3) counteractive medications such as thiazides or potassium sparing diuretics, or indomethacin (Indocin). Amiloride (Midamor) or an amiloride-hydrochlorothiazide combination (Moduretic) is preferred to minimize the risk of hypokalemia. Whether single-daily dosing is beneficial remains controversial, as do treatments such as inositol, potassium supplementation, and desmopressin (DDAVP). Polyuria does not always resolve following discontinuation of lithium.

Lithium can alter kidney morphology. The most common finding is a nonspecific interstitial fibrosis, although a lithium-distinctive microcystic lesion has also been described. In the absence of lithium intoxication and renal insufficiency, the morphological changes attributed to lithium tend to be mild, especially when compared with appropriately selected controls.

An issue not resolved is the extent to which nontoxic use of lithium is associated with eventual renal insufficiency. After correcting for age, one study of 142 patients on lithium for a minimum of 15 years found reduced GFR in 21 percent and increased serum creatinine in 12 percent. Another study of 114 patients on long-term lithium reported increased serum creatinine levels in 21 percent—the so-called creatinine creep. Both higher dosage and duration of treatment increase risk. Progression of renal insufficiency has been described despite discontinuation of lithium.

By age 62, a man successfully stabilized on lithium for 27 years had experienced a gradual increase in serum creatinine to 3.6 mg/dL (range of normal 0.6 to 1.3). A thorough evaluation by a nephrologist produced no other explanation, and lithium was felt to be the most likely cause. The drug was stopped, but over the next several years, kidney function worsened further and hemodialysis was started.

There have been occasional well-documented reports of lithium-related nephrotic syndrome and of incomplete distal renal tubular acidosis.

Regular monitoring of renal function is an essential part of long-term lithium therapy. For most patients, periodic measurement of serum creatinine, a urinalysis, and a clinical estimate of urine volume is sufficient. If clinically warranted, a 24-hour urine volume, protein, and creatinine clearance will provide more comprehensive information. Because there are many causes of impaired kidney function, consultation with a nephrologist is indicated if abnormalities are found.

Other Reactions

Weight gain is a common adverse effect of lithium and may be due to the

drug's complex effects on carbohydrate metabolism. Other possible causes include lithium-induced hypothyroidism, fluid retention, and increased caloric intake from thirst-quenching beverages.

Gastrointestinal adverse effects—including nausea, appetite loss, and loose stools—are common but usually mild. Severe gastrointestinal symptoms may portend impending lithium intoxication.

The most common hematological effect of lithium is a benign, reversible, and usually mild granulocytosis induced by stimulation of granulocyte production. An increase in platelet count is noted occasionally, but there are no clinically important effects on red blood cell count.

Dermatological adverse effects include the first occurrence or worsening of acne, psoriasis, and follicular keratosis, scattered reports of rashes of various types, and hair loss that is only occasionally related to lithium-induced hypothyroidism.

Through several possible mechanisms (antagonism of the calcium sensing receptor, direct stimulation of parathyroid hormone production, and decreased renal calcium excretion), lithium can cause hypercalcemia and hyperparathyroidism. Hypercalcemia is usually mild, but there have been reports of surgery being required to treat parathyroid hyperplasia or adenoma. Some authorities advise baseline and periodic monitoring of serum calcium levels.

Sexual dysfunction, including decreased libido and erectile difficulties, has been attributed to lithium, but well-documented cases are quite uncommon.

There is an inverse relationship between dietary sodium intake and serum lithium concentration. Low-sodium diets result in increased serum lithium concentrations and may cause lithium intoxication unless the dosage of lithium is reduced. Diets that severely restrict fluid intake place patients at risk of dehydration and subsequent lithium intoxication.

DRUG INTERACTIONS

Antipsychotics

Lithium and antipsychotic drugs have been used together safely and effectively for years; however, concern has been expressed that the combination may cause an encephalopathy in certain persons. Although the interaction has been reported most frequently with haloperidol (Haldol), a number of other antipsychotic drugs have also been implicated. A clear cause-and-effect relationship has not been established; nonetheless, caution is advised, and the simultaneous use of high doses of the two drugs should be avoided. The risk of extrapyramidal reactions may be increased as might the risk of neuroleptic malignant syndrome when lithium and antipsychotic drugs are combined.

Anticonvulsants

In treatment-resistant cases, lithium is often combined with another mood stabilizer such as divalproex, carbamazepine, lamotrigine, or oxcarbazepine. Neurotoxic interactions have been described but are usually avoidable with careful clinical management. When two drugs are used together, shared adverse effects may be more pronounced; for example, tremor and weight gain from lithium and divalproex. In addition, lithium offers some protection against carbamazepine-induced granulocytopenia, although whether it does the same for the rare agranulocytosis and aplastic anemia is unknown.

Antidepressants

In general, lithium and antidepressant combinations are safe and effective, although there have been occasional reports of a serotonin-like syndrome when lithium was combined with potent SSRIs. Serum lithium concentrations are not altered in any substantial way by antidepressant drugs.

Anti-Inflammatory Drugs

Most nonsteroidal anti-inflammatory drugs reduce renal lithium clearance and increase the serum lithium concentration in a way that is clinically important and potentially dangerous. Among the drugs that cause this interaction are indomethacin (Indocin), phenylbutazone (Butazolidin), diclofenac (Voltaren), ketoprofen (Orudis), meloxicam (Mobic), oxyphenbutazone (Tandearil), ibuprofen (Motrin, Advil, Nuprin), piroxicam (Feldene), and naproxen (Naprosyn). Aspirin and possibly sulindac (Clinoril) appear to be exceptions. The newer cyclooxygenase-2 (COX-2) inhibitors, such as celecoxib (Celebrex) and rofecoxib (Vioxx, no longer available), are no exception. Risk factors include high doses of the anti-inflammatory drug, increased age, and renal impairment. Because many of these drugs are now available without prescription, it is especially important that patients be aware of the possibility of an interaction.

Diuretics

Thiazide diuretics decrease renal lithium clearance and increase the serum lithium concentration. Lithium toxicity has been known to occur. If a thiazide is prescribed, lithium dosage reduction is often necessary. If a thiazide is discontinued, lithium dosage may need to be increased to avoid subtherapeutic levels. Although less well established, potassium-sparing diuretics may also cause lithium retention. However, loop diuretics such as furosemide (Lasix) do not reduce and may actually increase renal lithium clearance. Osmotic and xanthine diuretics (caffeine, theophylline [Theo-dur], and aminophylline) also increase lithium clearance. Because diuretics

are often given to patients who are medically ill with unstable fluid-electrolyte status, interactions with lithium may not be predictable, and close monitoring is advised.

Angiotensin-Converting Enzyme Inhibitors and Angiotensin II Receptor Antagonists

Both of these drug classes are used extensively to treat cardiovascular disease, and both have been reported to cause lithium retention and toxicity in some patients. Package inserts note this in the precautions section.

Other Drug Interactions

Adverse interactions with varying degrees of substantiation have been reported between lithium and a number of other drugs (Table 34.19–5). These interactions may involve increases or decreases in serum lithium concentrations or neurotoxic symptoms in the presence of unchanged serum levels. Questions about potential interactions should be addressed by consulting a knowledgeable pharmacist, a handbook on drug interactions, a regional drug information center, or the Lithium Information Center at the Madison Institute of Medicine.

Table 34.19–5.
Drug Interactions with Lithium

Drug Class	Reaction
Antipsychotics	Case reports of encephalopathy, worsening of extrapyramidal adverse effects, and neuroleptic malignant syndrome. Inconsistent reports of altered red blood cell and plasma concentrations of lithium, antipsychotic drug, or both.
Antidepressants	Occasional reports of a serotonin-like syndrome with potent serotonin reuptake inhibitors.
Anticonvulsants	No significant pharmacokinetic interactions. Occasional reports of neurotoxicity; combinations helpful for treatment resistance.
Nonsteroidal anti-inflammatory drugs	May reduce renal lithium clearance and increase serum concentration. Toxicity reported (exceptions are aspirin and possibly sulindac [Clinoril]).
Diuretics	
Thiazides	Reduced renal lithium clearance and increased serum concentration. Toxicity reported.
Potassium-sparing	Limited data may increase lithium concentration.
Loop	Lithium clearance unchanged (some case reports of increased lithium concentration).
Osmotic (mannitol, urea)	Increase renal lithium clearance and decrease lithium concentration.
Xanthine (aminophylline,	Increase renal lithium clearance and decrease lithium concentration.

caffeine, theophylline)	
Carbonic anhydrase inhibitors	Increase renal lithium clearance (acetazolamide).
Angiotensin-converting enzyme (ACE) inhibitors	Reports of reduced lithium clearance, increased concentrations and toxicity.
Angiotensin II receptor type-1 (AT-1) antagonists	Reports of increased lithium concentrations and toxicity.
Calcium channel inhibitors	Case reports of neurotoxicity; no consistent pharmacokinetic interactions.
Miscellaneous	
Succinylcholine, pancuronium	Reports of prolonged neuromuscular blockade.
Metronidazole	Increased lithium concentration.
Methyldopa	Few reports of neurotoxicity.
Sodium bicarbonate	Increased renal lithium clearance.
Iodides	Additive antithyroid effects.
Propranolol	Used for lithium tremor. Possible slight increase in lithium concentration.

LABORATORY INTERFERENCES

Lithium does not appear to interfere in any clinically important way with other laboratory tests. A distinction must be made, however, between interfering with laboratory tests and causing alterations in laboratory tests. Lithium can cause alterations in a number of areas, including white blood cell count (increase), thyroid function (decrease), and serum calcium (increase). Those changes are induced by lithium rather than by an artifactual interference with the tests themselves. Serum lithium concentrations will be elevated spuriously if blood has been collected in a lithium–heparin anticoagulant tube.

Laboratory Monitoring

Lithium has been measured in virtually every body fluid, but only blood (serum or plasma) is used in clinical practice. Despite considerable research, red blood cell, saliva, and tear lithium measurements and ratios such as red blood cell to plasma have failed to gain widespread acceptance. Methods used to measure serum lithium include flame photometry, atomic absorption spectrophotometry, and ion-selective electrode analysis. Any of these methods in competent hands yields results that are acceptable for both therapeutic and toxic clinical situations. More recently, colorimetric assays have been developed. One uses finger stick blood and returns a plasma lithium level within minutes. The analyzers used in clinical

laboratories are usually not capable of measuring the low levels of naturally occurring lithium in the blood, but more sensitive measures such as flameless atomic absorption spectroscopy and mass spectrometry are available for that purpose. Promising research has been conducted using Li-7 magnetic resonance spectroscopy to measure lithium concentrations in human brain.

The periodic measurement of serum lithium concentration is an essential aspect of patient care, but it should always be combined with sound clinical judgment. A laboratory report listing the therapeutic range as 0.5 to 1.5 mEq/L may lull a clinician into disregarding early signs of lithium intoxication in patients whose levels are less than 1.5 mEq/L. Clinical toxicity, especially in the elderly, has been well documented within this so-called therapeutic range.

The US package inserts for lithium products list effective serum concentrations for mania between 1.0 and 1.5 mEq/L (usually achieved with 1,800 mg of lithium carbonate daily) and for long-term maintenance between 0.6 and 1.2 mEq/L (usually achieved with 900 to 1,200 mg of lithium carbonate daily). The dose–blood level relationship may vary considerably from patient to patient. The likelihood of achieving a response at levels above 1.5 mEq/L is usually outweighed greatly by the increased risk of toxicity, although rarely a patient may both require and tolerate a higher-than-usual blood concentration.

What constitutes the lower end of the therapeutic range remains a matter of debate. A prospective 3-year study found patients maintained between 0.4 and 0.6 mEq/L (mean 0.54) were 2.6 times more likely to relapse than those maintained between 0.8 and 1.0 mEq/L (mean 0.83). However, the higher blood concentrations produced more adverse effects and were less well tolerated. A reanalysis of these data concluded that lower maintenance levels per se were not responsible for the higher relapse rate, but rather that the risk of relapse was increased only in patients who were abruptly randomized from higher prestudy lithium levels to the lower level group. An international group of experts concluded recently that a good prophylactic effect can usually be obtained within the range of 0.4 to 0.8 mEq/L; unfortunately, trial and error is the only way to establish the minimum effective level for a given patient.

Essential to interpreting serum lithium levels is knowing the sampling interval (the time between the last dose and the drawing of blood), the dose form, and the dosage schedule. Blood concentrations obtained an hour or so after a dose are considerably higher than those obtained after a long sampling interval. The 12-hour interval that has been adopted as standard has been defined as follows: (1) the blood should be drawn in the morning, 12 hours (± 30 minutes) after the last dose; (2) a multiple-dose regimen should be used; and (3) a steady-state condition should exist (skipped or extra doses within 4 or 5 days should be avoided). There may be variations

in the 12-hour serum level when switching from immediate release to a slow or controlled release preparation. Although the effectiveness of lithium does not depend on the daily dosing schedule, 12-hour serum concentrations will be higher with single as opposed to multiple daily dosing. The difference may range between 10 and 30 percent and cannot be easily predicted. Consequently, treatment is usually initiated with at least twice-daily dosing, although switching later to single-daily dosing may have the advantage of improved compliance and fewer adverse effects.

Even when these guidelines are followed, the serum level may fluctuate from measurement to measurement, reflecting factors such as dietary sodium intake, mood state, activity level, body position, and improper sample collection (use of a lithium-heparin anticoagulant will increase the concentration by more than 1 mEq/L).

Dosage Prediction

In an effort to make lithium as therapeutically efficient as possible, investigators have developed several approaches to predicting the daily dosage required to produce the desired serum concentration. Dosage prediction techniques include giving a single dose of lithium carbonate and using the serum concentration after a given interval (usually 24 hours) to predict the maintenance dose; measuring renal lithium clearance; and the use of mathematical models. Thus far, prediction techniques have not been widely accepted, and most clinicians continue to adjust the dosage based on the patient's clinical condition and previous blood concentrations. Nonetheless, "loading" a drug to produce a more rapid onset of action has become more appealing because of economic pressures to minimize the length of hospital stay. Although two small studies found loading doses of 20 mg/kg per day and 30 mg/kg per day of slow-release lithium carbonate to be well tolerated, it remains to be determined whether lithium loading can safely reduce time to remission or duration of hospitalization.

DOSAGE AND ADMINISTRATION

The extent of medical evaluation necessary before starting lithium is determined by individual patient characteristics such as age, associated illnesses, and the presence of other drugs; consequently, a thorough medical history is essential. The history determines the need for and the extent of physical and laboratory evaluations. All patients, however, should have renal and thyroid function tested before beginning treatment. A pregnancy test would be appropriate for women of childbearing potential, as would an ECG in older patients.

In the United States, lithium preparations include immediate release generic lithium carbonate, slow release (Lithobid and generic lithium carbonate), controlled release (generic lithium carbonate), and lithium

citrate syrup. There are no parenteral forms. Both 300 mg of the carbonate and 5 mL of the citrate contain about 8 mEq (mmol) of lithium (Table 34.19–6). Elsewhere in the world lithium acetate, glutamate, gluconate, orotate, and sulfate preparations have been available or are currently available. Clinicians should be aware the *lithium* and *lithium carbonate* are not interchangeable (300 mg of lithium is equivalent to 1,597 mg of lithium carbonate).

Table 34.19–6.
Lithium Preparations Available in the United States

Lithium carbonate capsules	150, 300, 600 mg
Lithium carbonate tablets	300 mg
Lithium carbonate controlled-release tablets	450 mg
Lithium carbonate slow-release tablets	300 mg
Lithium citrate syrup	8 mEq/5 mL

Starting Dosage

Lithium therapy is initiated in divided doses. Once a patient is stabilized, single daily dosing is sometimes more convenient. In the presence of normal kidney function, a total daily dose of 1,200 to 1,800 mg of lithium carbonate generally produces an antimanic serum concentration of 0.8 to 1.2 mEq/L. Maintenance levels of 0.6 to 1 mEq/L can usually be attained with 900 to 1,200 mg daily. In general, a conservatively low dose is started, perhaps 300 mg two or three times daily, a serum concentration is obtained after steady state is reached (in 4 or 5 days), and the dose is adjusted accordingly. Patients vary widely in the dose required to reach a desired blood concentration, although the elderly usually require lesser amounts. Consequently, these figures must be considered guidelines rather than absolutes.

Maintenance Dosage

During maintenance therapy, patients must be evaluated clinically, lithium levels determined periodically, and appropriate laboratory tests performed at regular intervals. Monthly visits are common early in treatment if the clinical course is uncomplicated. Patients who have been stable for extended periods may be seen at intervals of 3, 4, or even 6 months. Patient education, including written materials and support programs, such as those by the Depression and Bipolar Support Alliance (DBSA) (formerly the National Depressive and Manic Depressive Association), plays an important role in ensuring safe, effective, and compliant long-term treatment.

Discontinuation

Lithium should be discontinued if it is ineffective or not tolerated. Patients may stop the drug for other reasons, such as a perceived or real loss in creativity, feeling cured, or a dislike for feeling controlled by a medicine. After a period of stability with maintenance therapy, a trial off lithium may be considered, although the risk of recurrence is substantial (especially if there have been several prior episodes), and there have been reports of failure to respond to lithium when treatment is reinstituted. Discontinuation should be gradual over many weeks because more abrupt discontinuation appears to be associated with a higher likelihood of early recurrence, especially of mania. Teaching patients and significant others to recognize early signs of recurrence is an important part of the discontinuation process.

REFERENCES

Baldessarini RJ, Tondo L, David P, Pompili M, Goodwin FK. Decreased risk of suicides and attempts during long-term lithium treatment: a meta-analytic review. *Bipolar Disord*. 2006;8:625.

*Bauer M, Grof P, Müller-Oerlinghausen B. *Lithium in Neuropsychiatry: The Comprehensive Guide*. Oxon, UK: Informa UK; 2006.

Bearden CE, Thompson PM, Dalwani M, Hayashi KM, Lee AD. Greater cortical gray matter density in lithium-treated patients with bipolar disorder. *Biol Psychiatry*. 2007;62:7.

Bendz H, Sjüdin I, Aurell M. Renal function on and off lithium in patients treated with lithium for 15 years or more. A controlled, prospective lithium-withdrawal study. *Nephrol Dial Transplant*. 1996;11:457.

Bowden CL, Calabrese JR, McElroy SL, Gyulai L, Wassef A. A randomized, placebo-controlled 12-month trial of divalproex and lithium in treatment of outpatients with bipolar I disorder. *Arch Gen Psychiatry*. 2000;57:481.

Bowden CL, Mintz J, Tohen M. Multi-state outcome analysis of treatments (MOAT): application of a new approach to evaluate outcomes in longitudinal studies of bipolar disorder. *Mol Psychiatry*. 2016;21(2):237–242.

Caldieraro MA, Walsh S, Rabideau, D, et al. Clinical correlates of acute bipolar depressive episode with psychosis. *J Affective Disorders*. (In press)

Cohen LS, Friedman JM, Jefferson JW, Johnson EM, Weiner ML. A reevaluation of risk of in utero exposure to lithium. *JAMA*. 1994;271:146.

Collins J, McFarland B. Divalproex, lithium and suicide among Medicaid patients with bipolar disorder. *J Affective Disord*. 2008;107:23.

Einat H, Manji HK. Cellular plasticity cascades: genes-to-behavior pathways in animal models of bipolar disorder. *Biol Psychiatry*. 2006;59:1160.

Geddes JR, Burgess S, Hawton K, Jamison K, Goodwin GM. Long-term lithium therapy for bipolar disorder: systematic review and meta-analysis of randomized controlled trials. *Am J Psychiatry*. 2004;161:217.

Gitlin M. Lithium and the kidney. An updated review. *Drug Safety*. 1999;20:231.

Gonzalez Arnold J, Salcedo S, Ketter TA, et al. An exploratory study of responses to low-dose

lithium in African Americans and Hispanics. *J Affect Disord*. 2015;178:224–228.

*Goodwin FK, Jamison KR. *Manic-Depressive Illness*. 2nd ed. New York: Oxford University Press; 2007.

Goodwin GM, Bowden CL, Calabrese JR, Grunze H, Kasper S. A pooled analysis of 2 placebo-controlled 18-month trials of lamotrigine and lithium maintenance in bipolar I disorder. *J Clin Psychiatry*. 2004;65:432.

Harwood AJ. Lithium and bipolar mood disorder: the inositol-depletion hypothesis revisited. *Mol Psychiatry*. 2005;10:117.

Johnson FN. *The History of Lithium Therapy*. London: Macmillan; 1984.

Livingston C, Rampes H. Lithium: a review of its metabolic adverse effects. *J Psychopharmacol*. 2006;20:347.

Müller-Oerlinghausen B. Drug interactions with lithium—a guide for clinicians. *CNS Drugs*. 1999;11:41.

Ostacher MJ, Nierenberg AA, Rabideau D, et al. A clinical measure of suicidal ideation, suicidal behavior, and associated symptoms in bipolar disorder: psychometric properties of the Concise Health Risk Tracking Self-Report (CHRT-SR). *J Psychiat Research*. 2015;71:126–133.

Presne C, Fakhouri F, Noël LH, Stengel B, Even C. Lithium-induced nephropathy: rate of progression and prognostic factors. *Kidney Int*. 2003;64:1328.

*Raedler TJ, Wiedemann K. Lithium-induced nephropathies. *Psychopharmacol Bull*. 2007;40:134.

Rowe MK, Wiest C, Chuang D-M. GSK-3 is a viable potential target for therapeutic intervention in bipolar disorder. *Neurosci Biobehav Rev*. 2007;31:920.

Shaltiel G, Chen G, Manji HK. Neurotrophic signaling cascades in the pathophysiology and treatment of bipolar disorder. *Curr Opin Pharmacol*. 2007;7:22.

Suppes T, Dennehy EB, Hirschfeld RMA, Altshuler LL, Bowde CL. The Texas implementation of medication algorithms: update to the algorithms for treatment of bipolar I disorder. *J Clin Psychiatry*. 2005;66:870.

Sylvia LG, Shelton RC, Kemp DE, et al. Medical burden in bipolar disorder: findings from the Clinical and Health Outcomes Initiative in Comparative Effectiveness for Bipolar Disorder study. *Bipolar Disord*. 2015;17:212–223.

Viguera AC, Newport DJ, Ritchie J, Stowe Z, Whitfield T. Lithium in breast milk and nursing infants: clinical implications. *Am J Psychiatry*. 2007;164:342.

Waring WS. Management of lithium toxicity. *Toxicol Rev*. 2006;25:221.

Yonkers KA, Little BB, March D. Lithium during pregnancy. Drug effects and their therapeutic implications. *CNS Drugs*. 1998;9:261.

▲ 34.20 Melatonin Receptor Agonists

SAMANTHA B. SALTZ, M.D., KAREN M. PADRON, M.D., PH.D., BRETT LLOYD, JR., M.D., PH.D., AND JEFFREY NEWPORT, M.D.

INTRODUCTION AND HISTORY

Insomnia is a debilitating condition affecting 34 percent of adults in the United States with chronic insomnia affecting 10 to 15 percent of the population. Insomnia disorder is characterized in *DSM-5* as difficulty initiating or maintaining sleep, or early morning awakening with an inability to return to sleep, occurring three or more nights per week for at least 3 months and resulting in significant psychosocial impairment. Historically, many classes of pharmacologic agents have been used to treat insomnia including classic benzodiazepines, other benzodiazepine receptor agonists (e.g., zolpidem), antihistamines, naturopathic agents, and even antidepressants and antipsychotics. Potential side effects of such drugs include residual daytime sedation, gait instability, memory impairment, weight gain and metabolic complications, and abuse and withdrawal potential (Lieberman, 2007). For this reason, safe, well-tolerated medications with low side-effect profiles for insomnia are imperative.

Although the pineal gland, as an anatomical structure, has been recognized since antiquity, its function remained a mystery until the 20th century. In 1917, McCord and Allen reported that the dermal pigmentation of tadpoles, when placed in water seeded with crushed bovine pineal gland, was altered. This provided initial evidence of the existence of the neurohormone now known as melatonin. Forty years would pass before *N*-acetyl-5-methoxytryptamine would be isolated and named melatonin. Nearly 40 additional years would elapse before melatonin receptors in the human were characterized.

Melatonin is released from the pineal gland in accordance with a circadian rhythm producing high nocturnal levels (400 pM) peaking at around 2 AM and low daytime levels (40 pM). It soon emerged that, in most vertebrate species, nocturnal melatonin secretion varies inversely with the length of day, thereby regulating, in some species, seasonal variations in reproductive behavior. It was not until nearly 1990 that the importance of melatonin in synchronizing not only seasonal but circadian rhythmicity became a focus of inquiry. It is now known that the suprachiasmatic nucleus (SCN) of the hypothalamus synchronizes melatonin synthesis to the light/dark cycle, transmitting neural signals from retinal neurons to the pineal gland.

With this realization, increased attention has been focused upon melatonergic dysfunction as a pathophysiologic mechanism underlying disruptions in the circadian sleep–wake cycle associated with insomnia, depression, and other disorders, and upon melatonergic ligands as pharmacologic agents to not only promote sleep but to normalize circadian rhythmicity.

Finally, within the last decade, medicinal use of melatonin, categorized in the United States as a nutritional supplement, has rapidly expanded and has been accompanied by the arrival of several melatonin receptor agonists with others in development. Agents currently approved for use in the

United States or Europe include melatonin, agomelatine, ramelteon, and tasimelteon.

CHEMISTRY

Melatonin is derived from serotonin within the pineal gland and other sites. As a result, the molecular structure of melatonin (Fig. 34.20–1) closely resembles that of serotonin, both molecules comprised of an indole aromatic ring (i.e., a 6-member benzene ring fused with a 5-member nitrogen-containing pyrrole ring) with a side chain at position 3. To form melatonin, the terminal amine group on serotonin's side chain is acetylated producing an acylaminoethyl side chain, and serotonin's hydroxyl group is methylated producing a methoxy group at position 5. These changes render melatonin a more lipophilic molecule that, unlike serotonin, readily crosses the blood–brain barrier.

All currently available melatonin receptor agonists are structurally similar to melatonin. Most similar is agomelatine, the only structural difference being replacement of melatonin's indole (fused benzene and pyrrole rings) aromatic ring with a naphthalene (two fused benzene rings) aromatic ring (Fig. 34.20–2). Ramelteon and tasimelteon preserve melatonin's acylaminoethyl side chain but replace the indole with a benzofuran aromatic ring in which a 6-member benzene ring is fused with a 5-member oxygen-containing furan ring (Figs. 34.20–3 and 34.20–4).

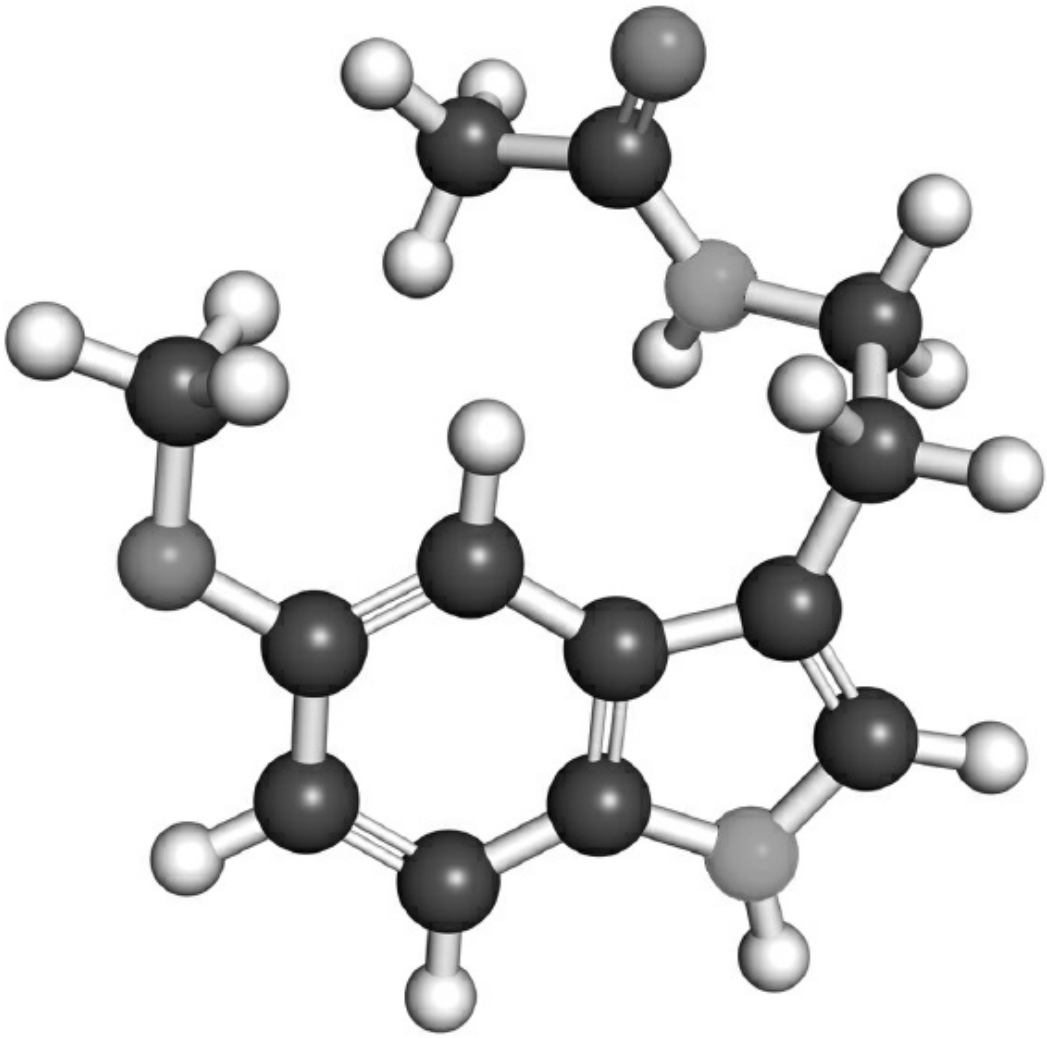

FIGURE 34.20–1. Molecular structure of melatonin. Light gray, nitrogen; gray, carbon; medium gray, oxygen; white, hydrogen.

PHARMACOLOGIC ACTIONS

Pharmacokinetics

Pharmacokinetic profiles for melatonin and the three commercially available melatonin receptor agonists are remarkably similar (Table 34-20–1). Each is readily and rapidly absorbed when administered orally, reaching peak plasma concentrations within an hour of administration. Concomitant food intake lowers absorption of tasimelteon (44 percent), ramelteon (22 percent), and melatonin (13 percent), but not agomelatine.

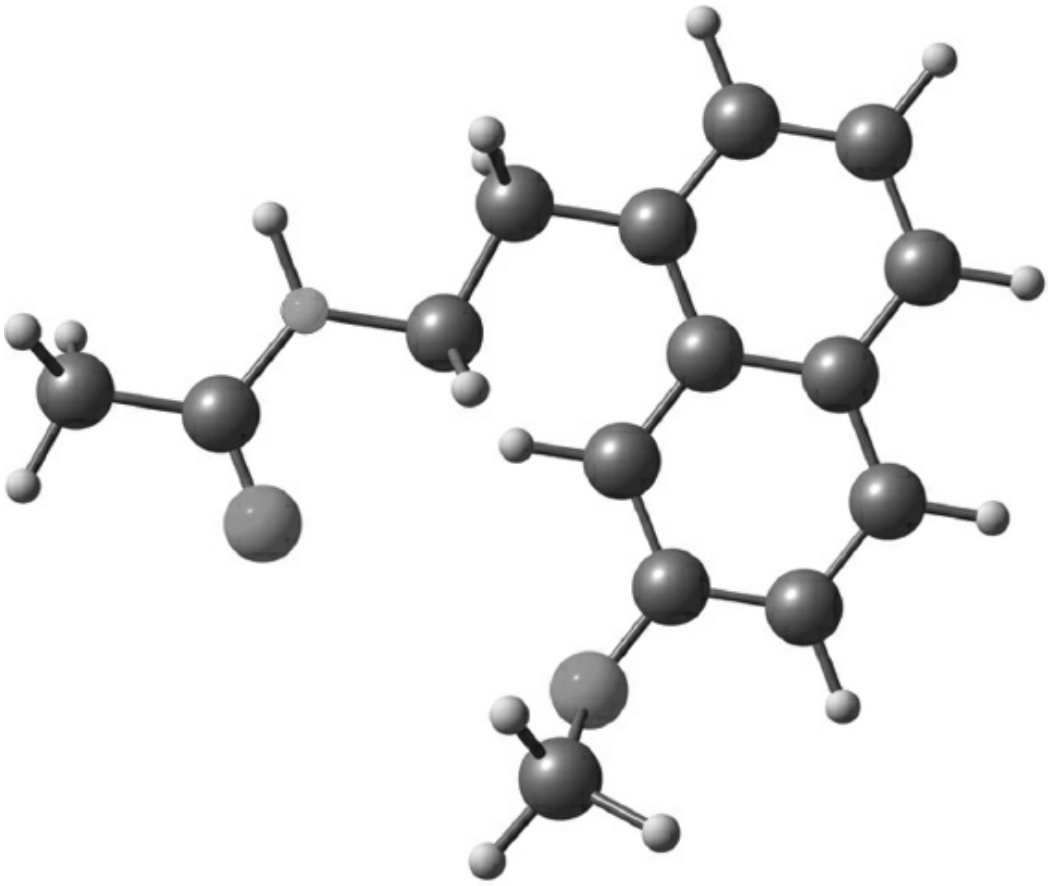

FIGURE 34.20–2. Molecular structure of agomelatine. Light gray, nitrogen; gray, carbon; medium gray, oxygen; white, hydrogen.

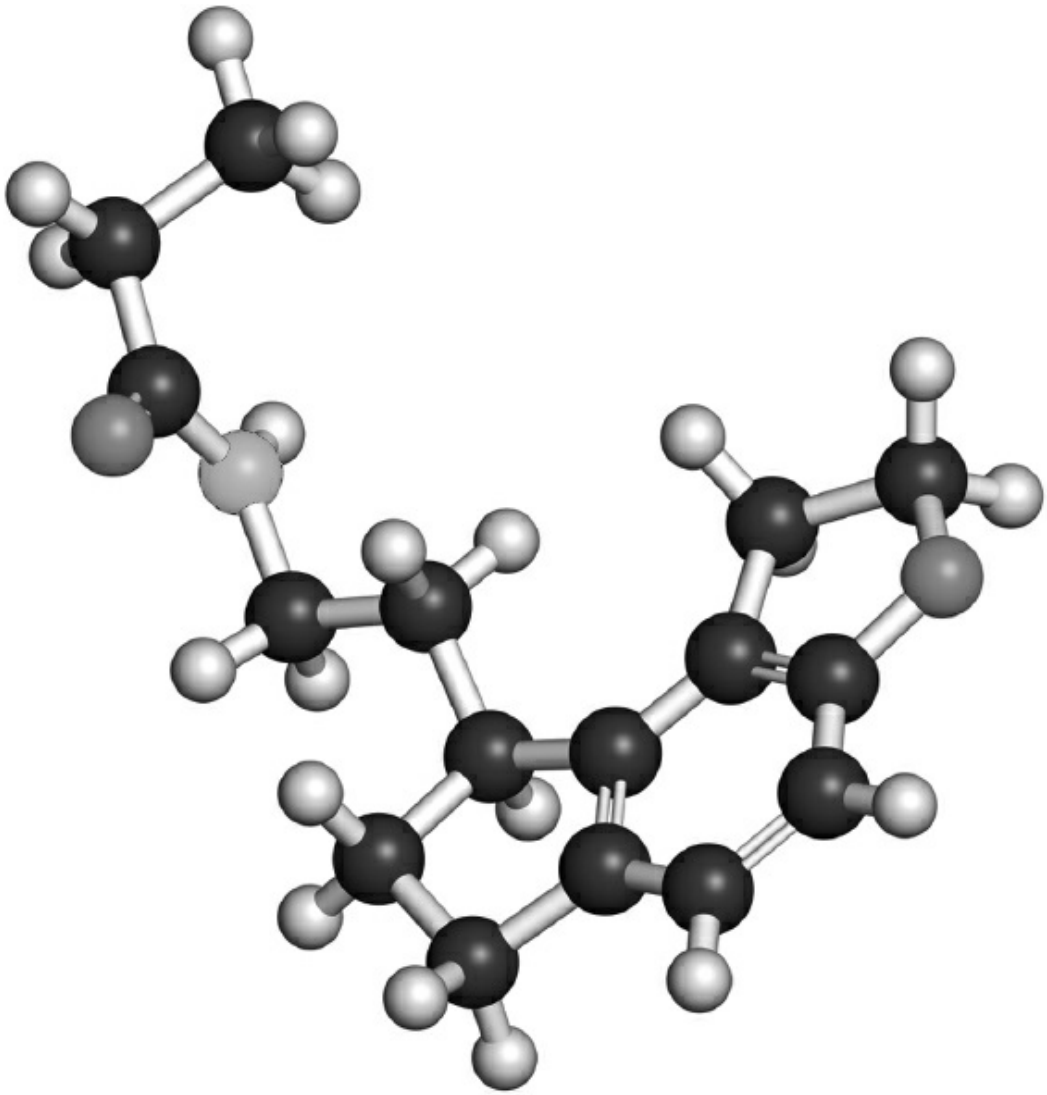

FIGURE 34.20–3. Molecular structure of ramelteon. Light gray, nitrogen; gray, carbon; medium gray, oxygen; white, hydrogen.

Despite a uniformly high absorption rate (≥80 percent), absolute oral bioavailability is only 38 percent for tasimelteon, 15 percent for melatonin, and <5 percent for the others due to extensive hepatic first pass metabolism. All four are rapidly and extensively metabolized with elimination half-lives less than 2 hours and no more than 2 percent of the parent compound excreted unchanged in urine.

All four agents are primarily metabolized via hepatic oxidation, most extensively by the 1A2 cytochrome P450 isoenzyme. Ramelteon and tasimelteon are also substrates for the 3A4 isoenzyme, whereas melatonin, agomelatine, and ramelteon are substrates for the 2C9/C19 isoenzymes.

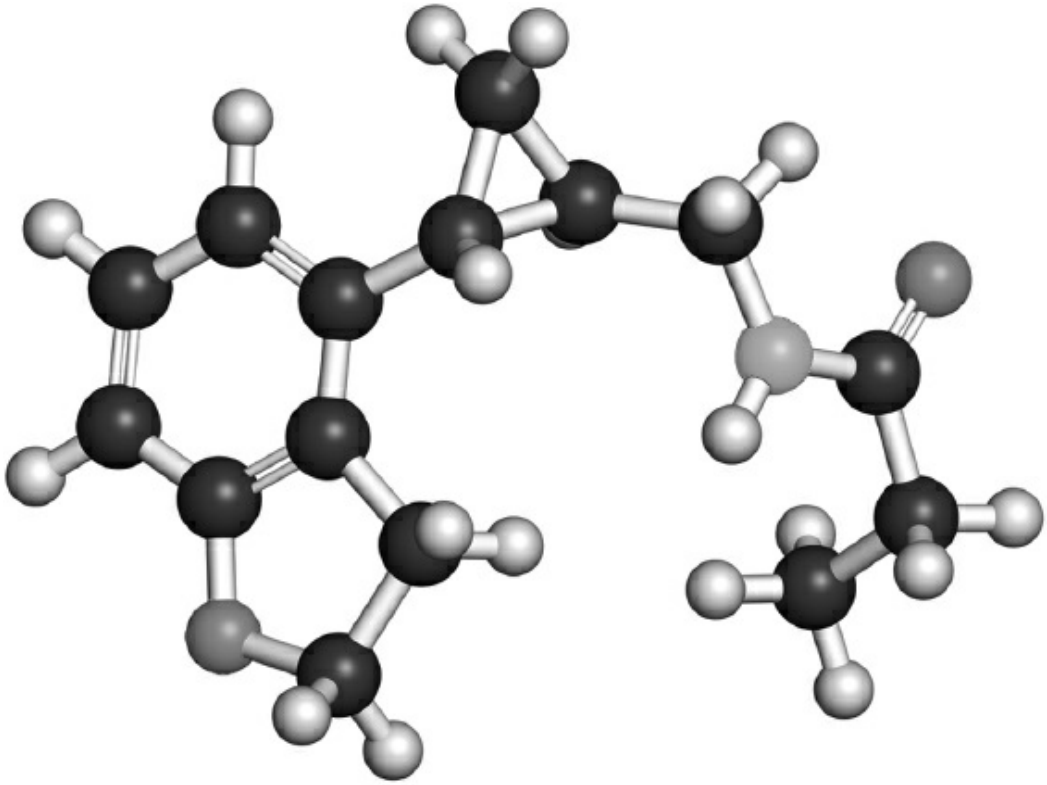

FIGURE 34.20–4. Molecular structure of tasimelteon. Light gray, nitrogen; gray, carbon; medium gray, oxygen; white, hydrogen.

Table 34.20–1.
Pharmacokinetic Profiles of Melatonin Receptor Agonists

Generic Name	Oral Absorption (%)	Time to Peak Concent-ration (hours)	Absolute Oral Bioavail-ability (%)	Protein Binding (%)	Elimination Half-Life (hours)	Renal Excretion of Parent Compound (%)	Metabolism (Cytochrome P4 Substrate)	
							1A2	2C9/19
Agomelatine	≥80	1–2	<5	95	1.0–2.0	<1	+++	+
Melatonin	100	0.75	15	60	0.8	2	+++	+
Ramelteon	84	0.75	1.8	82	1.3	<1	+++	+
Tasimelteon	≥90	0.5–3	38	90	1.3	<1	+++	

Renal impairment does not change the rate of drug clearance for any of the four melatonin receptor agonists. Conversely, mild and moderate hepatic impairment produces 70- and 140-fold increases in agomelatine exposure, respectively, and 4- and 10-fold increases in ramelteon exposure. Hepatic impairment is not associated with any changes in tasimelteon exposure or clearance. The effect of hepatic impairment upon the clearance of exogenous melatonin has not been formally evaluated; however, it has been reported that hepatic impairment reduces clearance of endogenous melatonin.

The pharmacokinetic profiles for melatonin, ramelteon, and tasimelteon are significantly affected by age, with elderly patients

exhibiting higher peak concentrations, diminished clearance rates, and longer elimination half-lives. Age-associated pharmacokinetic changes have not been observed with agomelatine.

There is a distinct gender difference in pharmacokinetic profile, with higher peak concentrations and/or longer elimination half-lives reported among women for all four agents. This is likely a consequence of estradiol's known inhibition of the 1A2 cytochrome P450 isoenzyme, the primary metabolic enzyme for all four. If so, reported gender differences are likely most pronounced among premenopausal women.

Finally, all of the agents are highly lipophilic compounds that readily traverse the blood–brain barrier. Of note, melatonin is less tightly protein bound (60 percent) than the three pharmaceutical agents.

Pharmacodynamics

Melatonin's physiologic activity is primarily mediated by two membrane-bound G protein–coupled receptors, melatonin receptor 1 (MT_1) and melatonin receptor 2 (MT_2), which are located in numerous brain regions and peripheral tissues. Both MT_1 and MT_2 receptors are found in the SCN of the hypothalamus, a region adjacent to the optic chiasm that is well recognized for its role in regulating circadian rhythms. Additional MT_1 receptors are located in the *pars tuberalis* of the pituitary. MT_2 receptors are also found in the hippocampus, cerebellum, and retina. Although their respective roles remain somewhat obscure, MT_1 receptors are thought to be involved in promoting sleep and inhibiting firing of the SCN, and MT_2 receptors are believed to play a major role in melatonin's resynchronizing activity. A purported melatonin receptor 3 (MT_3) has been demonstrated to be the enzyme quinone reductase 2 and may thus underlie the antioxidant properties of melatonin.

Other physiologic roles ascribed to melatonin include regulation of metabolic, cardiovascular, and immune systems with reported antioxidant, antinociceptive, and anticancer effects. These are likely mediated, at least in part, by activation of MT_1 and MT_2 receptors present in various peripheral tissues.

Melatonin and the other commercially available agents (i.e., agomelatine, ramelteon, tasimelteon) are all considered nonselective MT_1/MT_2 receptor agonists, albeit with distinct affinity profiles (Table 34.20–2). None has pharmacologically meaningful affinity for benzodiazepine, histaminic, or cholinergic receptors. The affinities of agomelatine, ramelteon, and tasimelteon for the so-called MT3 receptor are obscure. Among the current agents, ramelteon has the greatest MT_1 selectivity with six times melatonin's affinity for the MT_1 receptor. Tasimelteon, conversely, is the only current agent with greater affinity for

the MT_2 than the MT_1 receptor, theoretically underlying its particular utility in entrainment to remedy circadian rhythm disturbances.

Agomelatine, approved in Europe for the treatment of depression, is not only an MT_1 and MT_2 agonist but an antagonist at the serotonin-2C ($5HT_{2C}$) receptor. It remains a matter of debate whether the antidepressant efficacy of agomelatine is a consequence solely of its action on $5HT_{2C}$ receptors or of a synergy between its effects on MT_1, MT_2, and $5HT_{2C}$ receptors.

CLINICAL DRUG STUDIES

Numerous randomized clinical trials have demonstrated the beneficial effects of melatonin and other melatonin receptor agonists for sleep disturbances. For example, recent placebo-controlled trials of controlled-release melatonin (2 mg) in the treatment of primary insomnia have demonstrated significant improvements in self-reported quality of sleep in a study of 170 elderly subjects and reduction in sleep latency (–19 vs. –2 minutes) in 722 adults.

Ramelteon is approved in the United States for the treatment of insomnia characterized by difficulty with sleep onset. A meta-analysis encompassing 13 trials and 5,812 patients found ramelteon to improve sleep quality and reduce sleep latency by 4.30 minutes in a non–dose-dependent fashion with somnolence as the only significant side effect. A similar meta-analysis including eight trials, also found improvements in subjective and polysomnographic sleep latency, total sleep time, and latency to REM with ramelteon. Moreover, improvements in sleep latency produced by ramelteon may be enduring, with one study demonstrating reduced sleep latency compared to baseline at week 1, and months 1, 3, 5, and 6. Tasimelteon is approved in the United States for the treatment of non–24-hour sleep–wake disorder in blind adults. In a controlled trial of 84 non-24 patients, reported in ClinicalTrials.gov (NCT01163032), tasimelteon significantly outperformed placebo in synchronizing the master body clock to a 24-hour day (20 percent vs. 2.6 percent), increasing nighttime sleep time (50- vs. 22-minute change), and decreasing daytime nap time (–49- vs. –22-minute change). Moreover, a non-24 withdrawal study (NCT01430754) reported greater maintenance of 24-hour synchronization among those maintained on tasimelteon than those randomized to placebo (90 percent vs. 20 percent). Similar benefits have been observed for transient insomnia following a sleep-time shift. A pooled report (Rajaratnam et al., 2009) of phase II and phase III trials enrolling 39 and 411 patients, respectively, confirmed that tasimelteon improved sleep efficiency, reducing sleep latency and wakefulness after sleep onset. A clinical trial (NCT00548340) for primary insomnia (n = 321) demonstrated significantly reduced sleep latency times for both 20-mg and 50-mg

tasimelteon doses (−45 and −46 minutes) compared to placebo (−28 minutes).

Table 34.20–2.
Receptor Affinities of Melatonin Receptor Agonists

	Receptor Affinity (K_i)		
Generic Name	**MT_1 (pM)**	**MT_2 (pM)**	**MT_1/MT_2 K_i Ratio**
Agomelatine	62	268	0.231
Melatonin	81	383	0.211
Ramelteon	14	112	0.125
Tasimelteon	350	170	2.059
TIK-301	81	42	1.929

MT_1, Melatonin 1 Receptor; MT_2, Melatonin 2 Receptor; $5HT_{2C}$, Serotonin 2C Receptor.

Melatonin receptor agonists have consistently outperformed placebo in clinical trials of insomnia and other circadian sleep disturbances. Unfortunately, tolerance to these agents has not been systematically assessed, nor have they been scrutinized in head-to-head comparisons against other classes of hypnotic agents. Nevertheless, a recent meta-analysis reported that polysomnographically determined reductions in sleep latency and time awake after sleep onset for extended-release melatonin (2 mg) and ramelteon (16 mg) were less than that of benzodiazepines but comparable to changes produced by other benzodiazepine receptor agonists. The same meta-analysis reported that changes produced by ramelteon at an 8-mg dose were less than both benzodiazepines and other benzodiazepine receptor agonists. The utility of melatonin receptor agonists in the management of mood disorders has also been a focus of inquiry, albeit with mixed results. A controlled trial (NCT00705003) evaluating melatonin (3 mg) augmentation of buspirone for depression demonstrated no evidence of therapeutic efficacy. Similarly, a controlled trial of ramelteon (8 mg) augmentation of escitalopram for depression (NCT00642694) was terminated due to lack of efficacy. Finally, in an 8-week depression trial (NCT01428661) including 507 subjects, 20-mg tasimelteon failed to separate from placebo in reducing depressive symptom severity. Controlled trials of ramelteon (8 mg) augmentation for bipolar disorder have produced mixed results with one study reporting a significantly longer time to relapse than placebo (188 vs. 84 days, respectively) among euthymic patients with bipolar disorder but another reporting no difference from placebo in reducing manic symptoms or sleep disturbance in manic outpatients.

The strongest evidence for efficacy in the treatment of mood disorders is found with agomelatine, which, in Europe, is indicated for the treatment

of depression. It is suggested that the antidepressant efficacy of agomelatine stems both from its effects on mood and on circadian resynchronization. A recent meta-analysis incorporating 20 placebo-controlled agomelatine trials and 7,460 subjects demonstrated antidepressant efficacy, with an odds ratio for therapeutic response 1.25 [95 percent CI: 1.11–1.40]. A recent Cochrane review comparing agomelatine to SSRI antidepressants and venlafaxine found no advantage or disadvantage in achieving response or remission in unipolar depression. Interestingly, in a meta-analysis of 14 antidepressants, agomelatine was the only agent less likely to induce insomnia than placebo, yet it also had the third lowest rate of daytime somnolence.

The foundation for the role of chronobiotic therapies for mood disorders has largely emerged from the study of SAD. Most patients with SAD have a phase delay, or later secretion of dim light melatonin, due to the absence of early morning light during the winter months in the northern hemisphere. Both the dose and timing of chronobiotic administration are arguably keys to successful treatment. Proper treatment results in phase advancement with improvements in mood symptoms, while incorrect dosing or timing may actually worsen insomnia and mood disturbances.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Cardiovascular System

Endogenous melatonin is likely involved in the circadian regulation of the cardiovascular with decreases in cardiac output, blood pressure, and heart rate when melatonin levels are peaking during the evening hours. Rodent studies indicate similar effects of exogenous administration of melatonin.

Clinical trial data have not indicated any adverse effects upon cardiovascular function among the four commercially available melatonin receptor agonists. In fact, it has been posited that melatonin receptor agonists should be studied to identify possible beneficial cardiovascular effects.

Central Nervous System

As a group, the melatonin receptor agonists compare favorably to benzodiazepines and other hypnotics with respect to adverse effects on the central nervous system. Moreover, a risk of abuse liability for this class of medications has not been found in small cross-over trials.

Endocrine System

Preclinical studies indicate that supraphysiologic administration of exogenous melatonin may disrupt secretory profiles for numerous hormones including corticosteroids, gonadal steroids, luteinizing hormone,

insulin, prolactin, and thyroid hormone.

Two clinical trials have evaluated the endocrine effects of ramelteon administration. Neither study demonstrated an impact of ramelteon on adrenal or thyroid axis function. However, one of the studies indicated significant effects on reproductive endocrine function, increasing serum prolactin levels by 34 percent, particularly among females, following 6 months of ramelteon therapy.

Endocrine function has not been specifically examined in clinical trials of the other commercially available melatonin receptor agonists.

Hepatic and Renal Systems

There is no evidence that clearance of any of the four available melatonin receptor agonists is altered by pre-existing renal impairment. Nor is there evidence that these agents impair renal function. In fact, some preclinical studies suggest that melatonin may protect the kidneys from renal ischemia and reperfusion injury, exposure to carcinogens, hemorrhagic shock, or other injuries.

Agomelatine, unfortunately, has been associated with hepatic dysfunction. During clinical trials, 1.4 percent of patients receiving agomelatine 25 mg and 2.5 percent of those receiving agomelatine 50 mg experienced elevations in serum transaminase concentrations that were more than three times the upper limit of normal. Rare occurrences of hepatitis, increased γ-glutamyl transferase concentrations to three times the upper limit of normal, jaundice, and hepatic failure have been reported in association with agomelatine therapy. Hepatic injury has not been reported in association with any of the other three approved melatonin receptor agonists.

Table 34.20–3.
Formulations of Commercially Available Melatonin Receptor Agonists

Generic Name	Brand Name	Primary Indication	Usual Adult Dose (mg)	How Supplied
Agomelatine	Valdoxan	Major depressive disorder	25–50	25-mg tablets
Melatonin		Insomnia	0.3–5	Varied
Melatonin (prolonged release)	Circadin	Insomnia (patient >55 years old)	2	2-mg tablets
Ramelteon	Rozerem	Insomnia (sleep onset)	8	8-mg tablets
Tasimelteon	Hetlioz	Non–24-hour sleep–wake disorder	20	20-mg capsules

Respiratory System

In contrast to benzodiazepines and other hypnotics, melatonin receptor agonists appear to be devoid of respiratory depressant effects. For example,

in a study of patients with mild to moderate sleep apnea or chronic obstructive pulmonary disease, ramelteon did not adversely impact apnea–hypopnea indexes or arterial oxygen saturation levels.

THERAPEUTIC INDICATIONS

As shown, melatonin agonists provide a novel approach to the treatment of insomnia and other circadian sleep disturbances (Table 34.20–3). In addition to its effect on the circadian rhythm, melatonin has also been reported to possess anti-inflammatory, antitumor, vascular, retinal, antioxidant, neuroprotective, and pain modulatory effects. It has antiepileptic, antinociceptive, neurotrophic, antioxidant, and anxiolytic properties, perhaps providing other avenues for which to explore its therapeutic potential. For example, melatonin may be beneficial in promoting healing for tissue injury secondary to radiation exposure. Following radiation injury, TNF-α levels are elevated, resulting in increased levels of reactive oxygen species, which are scavenged by melatonin.

Melatonin receptor agonists also hold potential utility in the treatment of depression and other mental disorders associated with disturbances of circadian rhythmicity. Indeed, agomelatine is indicated for the treatment of major depression in Europe. It is unclear, however, the extent to which, if any, agomelatine's MT_1/MT_2 agonism contribute to its antidepressant efficacy.

Emerging evidence suggests that melatonin agonists may play an important role in the treatment and reduction of delirium. In a recent study, ramelteon was associated with a lower incidence of delirium when administered to older adults admitted to an acute care setting. Based on reviews of controlled trials and case series, melatonin may also be useful in the treatment of delirium in hospitalized older adults.

PRECAUTIONS AND ADVERSE REACTIONS

In many respects, the side-effect profile of the melatonin receptor agonists compare favorably to other sleep agents. Unlike benzodiazepine hypnotics, melatonin agonists have not been associated with prominent cognitive impairment, fall risk, or discontinuation-associated withdrawal syndromes, or rebound insomnia. The melatonin agonists are also devoid of prominent anticholinergic (e.g., dry mouth, constipation, blurred vision) or antihistaminic (e.g., carbohydrate craving, weight gain) side effects associated with other hypnotic agents.

Pregnancy. Endogenous maternal melatonin plays an important role during pregnancy. Although the fetal pineal gland does not yet produce melatonin (infant melatonin production does not begin until 2 to 3 months after delivery), there are more melatonin binding sites in both the fetal pituitary and fetal hypothalamus than in the adult. It is speculated that

maternal melatonin, which has been shown to cross the placenta, establishes periodicity and entrains the circadian pacemaker. For example, in evening deliveries, significantly elevated melatonin levels are detected via sampling of the umbilical artery or vein at birth. One hypothesis is that maternal melatonin might communicate important information about maternal sleep–wake cycles and entrain the fetal pineal gland, so that fetal circadian rhythms may be found as early as 3 months. Preclinical evidence suggests other roles for endogenous melatonin during fetal life including influencing the development of reproductive organs, regulating gonadal hormone activity, and protecting the fetal brain from oxidative damage during episodes of fetal hypoxia.

Despite the growing evidence that endogenous melatonin benefits the fetus, exogenous melatonin administration, which likely produces markedly higher fetal circulating concentrations, should not carry a presumption of safety. For example, a series of rodent studies, administering exogenous melatonin during pregnancy at doses as low as 2 to 10 times recommended human doses have demonstrated decreased luteinizing hormone levels, decreased ovarian and uterine size, and delayed vaginal opening in the offspring. Following delivery, offspring exposed to exogenous melatonin exhibit disruption in estrous cycles and blunted responses to testosterone, suggesting impairment in maturation of the hypothalamic–pituitary gonadal and prolactin axes.

There are no controlled clinical data addressing the reproductive safety of exogenous administration of melatonin or other melatonin receptor agonists.

In summary, there is no clear evidence of teratogenicity, though there is preclinical evidence that increased levels of melatonin may interfere with regulation of sex hormone levels and cycles, as well as delay sexual maturation. Pending further data, exogenous administration of melatonin and other melatonin receptor agonists should be avoided during pregnancy.

Lactation

Existing studies have examined levels of endogenous melatonin in breast milk but not at exogenous levels. Cyclic levels of melatonin in breast milk replicate maternal patterns including elevations in the evening and decreases with light with endogenous melatonin breast milk levels ranging from 35 to 80 percent of maternal serum levels. Circadian fluctuations in breast milk melatonin concentrations may dictate nursing schedules to the infant.

Agomelatine is detectable in breast milk with peak concentrations occurring 1 to 2 hours after the drug was administered and milk:plasma ratio equaling 0.19. Of note, 4 hours after the drug was administered, agomelatine was not detectable in breast milk. Clinical breast milk

excretion data have not been published for other melatonin receptor agonists.

In a recent edition of Hale's standard lactation reference, both melatonin and ramelteon are assigned a moderately safe (L3) rating.

DRUG INTERACTIONS

All currently available melatonin receptor agonists, including melatonin itself, are predominantly metabolized by the hepatic cytochrome P450 1A2 isoenzyme and to a lesser extent, the 2C9/2C19 and/or 3A4 isoenzymes. Thus, there are potential interactions with drugs that inhibit or induce the activity of these P450 isoenzymes (Table 34.20–4).

Of greatest concern are drug interactions mediated by the 1A2 isoenzyme. For example, coadministration with fluvoxamine, a potent 1A2 inhibitor, produces a 7-fold increase in total tasimelteon exposure, a 17-fold increase in melatonin exposure, a 60-fold increase in agomelatine exposure, and a whopping 160-fold increase in ramelteon exposure. By comparison, coadministration with ketoconazole, a potent 3A4 inhibitor, produces a 0.84-fold increase in ramelteon exposure and a 0.50-fold increase in tasimelteon exposure. Fluconazole, a potent 2C9 inhibitor, produces a 1.5-fold increase in ramelteon exposure and was not observed to alter agomelatine exposure.

Table 34.20–4.
Potential Drug Interactions for Melatonin Receptor Agonists

	Cytochrome P450 Isoenzyme		
Pharmacokinetic Activity	**1A2**	**2C9/2C19**	**3A4**
Substrates[a]	Agomelatine Melatonin Ramelteon Tasimelteon	Agomelatine Melatonin Ramelteon	Ramelteon Tasimelteon
Inhibitors[b]	Estrogen[c] Fluvoxamine Quinolones[d]	Amiodarone Azole antifungals[e] Fluoxetine Fluvoxamine Omeprazole, Esomeprazole Tricyclic antidepressants[f]	Antiretrovirals[g] Azole antifungals[e] Conivaptan Diltiazem Fluvoxamine Macrolides[h] Nefazodone Verapamil
Inducers[b]	Carbamazepine Montelukast Phenobarbital Phenytoin Tobacco (Smoking)	Carbamazepine Phenytoin Rifampin	Carbamazepine Efavirenz Etravirine Modafinil Nafcillin Oxcarbazepine Phenobarbital Phenytoin

	Rifampin St. John wort Topiramate

[a]Substrate status for melatonin receptor agonists.

[b]Representative lists of strong and moderate potency inhibitors and inducers.

[c]Estrogen-containing agents include oral contraceptives and hormone replacement therapies.

[d]Quinolone antibiotics include ciprofloxacin and ofloxacin.

[e]Azole antifungals include fluconazole, itraconazole, ketoconazole, miconazole, and voriconazole.

[f]Tricyclic antidepressants include amitriptyline, clomipramine, and imipramine.

[g]Antiretroviral agents include atazanavir, boceprevir, darunavir, fosamprenavir, indinavir, lopinavir, nelfinavir, ritonavir, saquinavir, and telaprevir.

[h]Macrolide antibiotics include clarithromycin, erythromycin, and telithromycin.

The magnitude of the drug interaction observed between fluvoxamine and the four melatonin receptor agonists raises concern for interactions with other agents that alter 1A2 activity. For example, coadministration with estrogen-containing compounds (e.g., oral contraceptives, hormone replacement therapies) may produce similar increases in levels of exposure to these agents. Unfortunately, no data are available to clarify the magnitude of this potential interaction. Tobacco, a known inducer of the 1A2 isoenzyme, may produce the opposite effect, that is, lowering exposure to the melatonergic agents. Indeed, smoking has been associated with a 3.7-fold lowering of cumulative exposure to agomelatine.

None of the four melatonergic agents, at therapeutic concentrations, have been observed to alter P450 enzyme activity; therefore, they seem unlikely to alter the clearance of other medications. However, in vitro research indicates that melatonin, at supratherapeutic concentrations, induces activity of the 3A4 isoenzyme. The clinical relevance of melatonin's impact of 3A4 activity, if any, is unknown. An additive impact upon psychomotor performance when taken with alcohol has been reported for the melatonin receptor agonists.

LABORATORY INTERFERENCE

None of the four approved agents (i.e., agomelatine, melatonin, ramelteon, tasimelteon) have been shown to have significant laboratory interferences. They also have not been shown to interfere with drug urine toxicology screens or to produce false-positive results.

DOSAGE AND ADMINISTRATION

Extended-release melatonin is administered using a 2-mg bedtime dose. Exogenous melatonin should not be administered to patients with moderate or severe hepatic impairment.

Ramelteon is usually administered at an 8-mg dose 30 minutes prior to bedtime, but there have been reports of efficacy between 4 mg and 64 mg

(Arendt, 2008). No dosing adjustment is needed for renal impairment. Ramelteon should not be administered to patients with moderate or severe hepatic impairment.

Tasimelteon is administered at a 20-mg dose at bedtime. There is no recommended dosing adjustment for renal or hepatic impairment.

Agomelatine is administered for depression using a 25-mg bedtime dose. No dose adjustment is needed for renal impairment. Agomelatine should not be administered to patients with even mild hepatic impairment.

INVESTIGATIONAL DRUGS

Two novel melatonin receptor agonists are currently in clinical trials. TIK-301, formerly known as LY-156735 and PD-6735, was initially developed by Eli Lilly and is currently licensed to Tikvah Therapeutics where phase III trials are on-going. The pharmacodynamic profile of TIK-301 is similar to agomelatine, a nonselective MT_1/MT_2 agonist (Table 34.20–3) as well as a $5HT_{2B}$ and $5HT_{2C}$ antagonist. It has potential utility for insomnia, circadian rhythm disorders, and depression.

Piromelatine, also known as Neu-P11, is a nonselective MT_1/MT_2 agonist and $5HT_{1A}/5HT_{1D}$ agonist under development by Neurim Pharmaceuticals for sleep and cognitive function in association with insomnia and Alzheimer disease. Piromelatine outperformed placebo in a 4-week phase II trial (NCT01489969) for insomnia. Preclinical research indicates that piromelatine may improve cognitive function in animal models of Alzheimer disease and prevent insulin resistance and metabolic derangements associated with sleep deprivation.

A number of other melatonin receptor ligands are in various stages of development and/or preclinical testing. Melatonin agonists in development exhibit one of four pharmacodynamic profiles: (1) nonselective MT_1/MT_2 agonists; (2) nonselective MT_1/MT_2 agonists with agonist or antagonist activity at one or more serotonin receptor subtypes; (3) MT_1 selective agonists; and (4) MT_2 selective agonists. By convention, a selective must display at least 50 times greater affinity for one receptor subtype.

Two MT_2 selective partial agonists, UCM765 and UCM924, have demonstrated antinociceptive efficacy in animal models of neuropathic pain. In other preclinical studies, UCM765 showed anxiolytic effects and promoted non-REM sleep. Production of MT_1 selective ligands has proven more challenging. Thus, progress toward development of MT_1 selective agonists is lagging.

REFERENCES

Al-Aama T, Brymer C, Gutmanis I, Woolmore-Goodwin SM, Esbaugh J, Dasqupta M. Melatonin decreases delirium in elderly patients: a randomized, placebo-controlled trial. *Int J Geriatr*

Psychiatry. 2011;26:687–694.

Alberti S, Chiesa A, Andrisano C, Serretti A. Insomnia and somnolence associated with second-generation antidepressants during the treatment of major depression: a meta-analysis. *J Clin Psychopharmacol*. 2015;35(3):296–303.

Asarnow LD, Soehner AM, Harvey AG. Circadian rhythms and psychiatric illness. *Curr Opin Psychiatry*. 2013;26(6):566–571.

Carocci A, Catalano A, Sinicropi MS. Melatonergic drugs in development. *Clin Pharmacol*. 2014;6:127–137.

Claustrat B, Brun J, Chazot G. The basic physiology and pathophysiology of melatonin. *Sleep Med Rev*. 2005;9:11–24.

ClinicalTrials.gov [Internet]. Bethesda, MD: National Library of Medicine (US); 2000–15.

Gahr M, Freudenmann RW, Connemann BJ, Hiemke C, Schonfeldt-Lecuona C. Agomelatine and hepatotoxicity: implications of cumulated data derived from spontaneous reports of adverse drug reactions. *Pharmacopsychiatry*. 2013;46(6):214–220.

Guaiana G, Gupta S, Chiodo D, Davies SJ, Haederle K, Koesters M. Agomelatine versus other antidepressive agents for major depression. *Cochrane Database Syst Rev*. 2013;12:1–132.

Hale TW. *Medications and Mother's Milk: A Manual of Lactational Pharmacology 2012*. Amarillo, TX: Hale Publishing; 2012.

Hatta K, Kishi Y, Wada K, et al. Preventive effects of ramelteon on delirium: a randomized placebo-controlled trial. *JAMA Psychiatry*. 2014;71(4):397–403.

Kuriyama A, Honda M, Hayashino Y. Ramelteon for the treatment of insomnia in adults: a systematic review and meta-analysis. *Sleep Med*. 2014;15(4):385–392.

Laudon M, Frydman-Marom A. Therapeutic effects of melatonin receptor agonists on sleep and comorbid disorders. *Int J Mol Sci*. 2014;15(9):15924–15950.

Lemoine P, Nir T, Laudon M, Zisapel N. Prolonged-release melatonin improves sleep quality and morning alertness in insomnia patients aged 55 years and older and has no withdrawal effects. *J Sleep Res*. 2007;16(4):372–380.

Lerner AB, Case JD, Takahashi Y, Lee TH, Mori W. Isolation of melatonin, the pineal gland factor that lightens melanocytes. *J Am Chem Soc*. 1958;80(10):2587.

Lewy AJ, Emens JS, Songer JB, et al. Winter depression: integrating mood, circadian rhythms, and the sleep/wake and light/dark cycles into a bio-psycho-social-environmental model. *Sleep Med Clin*. 2009;4(2):285–299.

Lieberman JA. Update on the safety considerations in the management of insomnia with hypnotics: incorporating modified-release formulations into primary care. *Prim Care Companion J Clin Psychiatry*. 2007;9(1):25–31.

Liu J, Wang LN. Ramelteon in the treatment of chronic insomnia: systematic review and meta-analysis. *Int J Clin Pract*. 2012;66(9):867–873.

Mayer G, Wang-Weigand S, Roth-Schechter B, Lehmann R, Staner C, Partinen M. Efficacy and safety of 6-month nightly ramelteon administration in adults with chronic primary insomnia. *Sleep*. 2009;32(3):351–360.

McCord CP, Allen FP. Evidences associating pineal gland function with alterations in pigmentation. *J Exp Zool*. 1917;23:207–224.

McElroy SL, Winstanley EL, Martens B, et al. A randomized, placebo-controlled study of

adjunctive ramelteon in ambulatory bipolar I disorder with manic symptoms and sleep disturbance. *Intl Clin Psychopharmacol*. 2011;26(1):48–53.

Norman TR. The effect of agomelatine on 5HT2C receptors in humans: a clinically relevant mechanism? *Psychopharmacology*. 2012;221:177–178.

Norris ER, Burke K, Correll JR, et al. A double-blind, randomized, placebo-controlled trial of adjunctive ramelteon for the treatment of insomnia and mood stability in patients with euthymic bipolar disorder. *J Affect Disord*. 2013;144:141–147.

Paulis L, Simko F, Laudon M. Cardiovascular effects of melatonin receptor agonists. *Expert Opin Investig Drugs*. 2012;21(11):1661–1678.

Rajaratnam SM, Polymeropoulos MH, Fisher DM, et al. Melatonin agonist tasimelteon (VEC-162) for transient insomnia after sleep-time shift: two randomised controlled multicentre trials. *Lancet*. 2009;373(9662):482–491.

Reppert SM, Weaver DR, Ebisawa T. Cloning and characterization of a mammalian melatonin receptor that mediates reproductive and circadian responses. *Neuron*. 1994;13(5):1177–1185.

Schmidt FM, Lichtblau N, Uribe MM, Kirchherr H, Himmerich H. Agomelatine in breast milk. *Int J Neuropsychopharmacol*. 2013;16:497–499.

Taylor D, Sparshatt A, Varma S, Olofinjana O. Antidepressant efficacy of agomelatine: meta-analysis of published and unpublished studies. *BMJ*. 2014;348:g1888.

Vella F, Ferry G, Delagrange P, Boutin JA. NRH: quinone reductase 2: an enzyme of surprises and mysteries. *Biochem Pharmacol*. 2005;71:1–12.

Voiculescu SE, Zygouropoulos N, Zahiu CD, Zagrean AM. Role of melatonin in embryo fetal development. *J Med Life*. 2014;7(4):488–492.

Wade AG, Ford I, Crawford G, et al. Nightly treatment of primary insomnia with prolonged release melatonin for 6 months: a randomized placebo controlled trial on age and endogenous melatonin as predictors of efficacy and safety. *BMC Med*. 2010;8(51):1–18.

Winkler A, Auer C, Doering BK, Rief W. Drug treatment of primary insomnia: a meta-analysis of polysomnographic randomized controlled trials. *CNS Drugs*. 2014;28(9):799–816.

Zawilska JB, Skene DJ, Arendt J. Physiology and pharmacology of melatonin in relation to biological rhythms. *Pharmacol Rep*. 2009;61(3):383–410.

Zlotos DP. Recent progress in the development of agonists and antagonists for melatonin receptors. *Curr Med Chem*. 2012;19(21):3532–3549.

▲ 34.21 Mirtazapine

RADU V. SAVEANU, M.D., MOUSA SOBHY AZER BOTROS, M.D., AND MICHAEL E. THASE, M.D.

Mirtazapine (Remeron) is a tetracyclic piperazinoazepine compound (Fig. 34.21–1) that was approved by the FDA in 1995 for the treatment of depression. Unlike most newer antidepressants, mirtazapine has virtually no effect on monoamine uptake. Rather, therapeutic effects are thought to be mediated by inhibition of α2-adrenergic receptors and blockade of postsynaptic serotonin type 2 (5-HT2) and type 3 (5-HT3) receptors. The former action diminishes the inhibitory effects of both α2-autoreceptors (located on presynaptic norepinephrine [NE] neurons) and α2-heteroreceptors (located on cell bodies of dorsal raphe 5-HT neurons). Inhibition of inhibitory cellular mechanisms thus increases the firing rates of noradrenergic and 5-HT neurons. This effect differs markedly from that of most other antidepressants, which tend to suppress cell firing rates until autoreceptors are downregulated.

Blockade of 5-HT2 and 5-HT3 receptors may enhance therapeutic effects directly (e.g., relief of insomnia or enhancement of appetite) or indirectly by dampening side effects attributable to disinhibited 5-HT release (e.g., agitation or nausea). Antagonism of 5-HT2 receptors also may enhance the availability of neurotransmitters to interact with serotonin type 1 (5-HT1) receptors. Mirtazapine is the only newer antidepressant that is a potent antagonist of histamine-1 receptors, which adds to sedative and appetite-enhancing properties. The sleep effects of mirtazapine in major depression have been studied. Mirtazapine therapy has been shown to lessen nocturnal awakenings and prolong total sleep time. Mirtazapine administration increased total slow wave sleep (SWS), REM latency, the duration of the first REM episode, and stage 3 sleep. It decreased the number of REM episodes. Simultaneously, it reduced wake time after sleep onset.

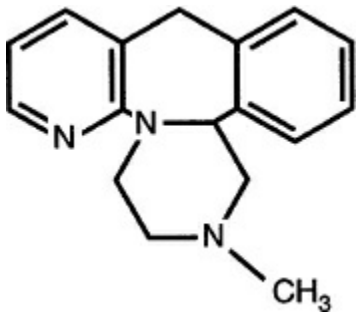

FIGURE 34.21–1. Chemical structure of mirtazapine (Remeron).

First synthesized in the early 1980s, mirtazapine is related to one older

antidepressant, mianserin (Tolvon). Mianserin is available throughout most of the world (although not in the United States or Canada). Mianserin is a more potent antagonist of α1-receptors than the newer compound, which is thought to explain why mianserin has less effect on 5-HT neurotransmission than mirtazapine.

PHARMACOLOGY

Mirtazapine has a half-life of approximately 30 hours and reaches steady state after 6 days of therapy. Linear pharmacokinetics is observed across the therapeutic dose range (i.e., 15 to 60 mg/day). Mirtazapine is essentially completely absorbed after oral dosing, and food has little effect on the rate or extent of absorption. Approximately 85 percent of circulating drug is bound to plasma protein.

Mirtazapine is metabolized in the liver via cytochrome P450 (CYP) 2D6, CYP 1A2, and CYP 3A4 pathways. The primary metabolite, *N*-desmethylmirtazapine, is substantially less active than the parent compound and has trivial pharmacological effects. Cimetidine, which inhibits CYP 2D6, CYP 1A2, and CYP 3A4, raises plasma levels of mirtazapine by more than 50 percent. Even greater increases in mirtazapine concentrations result from concomitant administration of the potent CYP 1A2 inhibitor fluvoxamine (Luvox).

The original formulation of mirtazapine is available in 15-, 30-, and 45-mg tablets. An alternate vehicle, an orally disintegrating tablet (ODT), was introduced in 2001 to facilitate compliance and to aid therapy for patients who have difficulty swallowing. The ODT is available in 30- and 45-mg strengths.

Antidepressant Efficacy

The minimum effective dose of mirtazapine in placebo-controlled clinical trials of major depressive disorder (MDD) is 15 mg/day. The maximum approved dose ranges from 45 mg/day (United States) to 60 mg/day (some European countries). There is some evidence of an ascending dose–response relationship for antidepressant efficacy, although this has not been definitively established.

The relative efficacy of mirtazapine has been established against older and newer antidepressants. Mirtazapine was found to be significantly more effective than trazodone (Desyrel) and as effective as the TCAs: amitriptyline (Elavil), doxepin (Sinequan), clomipramine (Anafranil), and nortriptyline (Aventyl). Therapy with the TCA, imipramine (Tofranil), was significantly more effective than mirtazapine treatment in one inpatient study. However, in this trial, dosages of both drugs were titrated in relation to plasma levels, which is a valid strategy for the TCA but of uncertain value for mirtazapine.

There are a number of published studies comparing mirtazapine and SSRIs, including citalopram (Celexa), paroxetine (Paxil), sertraline (Zoloft), fluvoxamine, and fluoxetine (Prozac). Several meta-analyses have been completed and have confirmed that, on average, the two types of antidepressants are comparably effective (less than 5 percent difference in response or remission rates) after 6 to 8 weeks of therapy, although patients treated with mirtazapine received earlier therapeutic benefit. It is not clear if this temporal difference is primarily due to a relatively nonspecific advantage in improving sleep and appetite or if it is attributable to the difference in mechanism of action.

Two studies have contrasted mirtazapine with the serotonin norepinephrine reuptake inhibitor (SNRI) venlafaxine (Effexor). In the first study, mirtazapine (mean dose: 50 mg/day) and the immediate release formulation of venlafaxine (mean dose: 255 mg/day) were compared in 157 depressed inpatients with melancholic features. Trends favored the mirtazapine-treated group on most outcome measures, but only the difference on the Hamilton Rating Scale for Depression (HAMD) sleep disturbance factor reached statistical significance at study endpoint. Significantly fewer patients did, however, withdraw from the mirtazapine group during the first 6 weeks of therapy because of side effects. The second study, which was conducted in 242 depressed outpatients, contrasted the ODT formulation of mirtazapine (45 mg/day) with the extended release formulation of venlafaxine (225 mg/day). Although the efficacy of the study drugs again did not differ at the end of the trial, results consistently favored mirtazapine therapy during the first 2 weeks of double-blind therapy, with significant differences in overall symptom improvement, response rates, and changes in sleep.

Combining SSRIs/SNRIs and Mirtazapine

Almost from the time of introduction, it was speculated that mirtazapine could be combined with SSRIs to obtain synergistic or additive effects. Subsequently, a number of studies have examined the clinical utility—both effectiveness and tolerability—of combining mirtazapine with other antidepressants, including SNRIs and bupropion as well as SSRIs.

The first RCT to evaluate mirtazapine in combination with other antidepressants (predominantly SSRIs) studied patients with a documented history of nonresponse to antidepressant therapy in the current depressive episode. In this pilot study, 26 outpatients were randomly assigned to 4 weeks of additional double-blind therapy with either active mirtazapine (15 to 30 mg/day) or matching placebo, in addition to ongoing therapy with the first antidepressant. The investigators found that the patient group receiving adjunctive therapy with active mirtazapine had significantly greater improvement in depressive symptoms than the group receiving adjunctive placebo. Although the difference in

remission rates was not statistically significant in this small study, the magnitude of the difference (45 percent vs. 13 percent)—if replicable in a larger study—would be clinically meaningful.

Another study tested the hypothesis that starting patients on the combination of mirtazapine and an SSRI would deliver a more rapid and/or larger response than monotherapy. This outpatient RCT randomly assigned 61 participants with MDD to 6 weeks of double-blind therapy with one of three options: (1) mirtazapine (30 to 60 mg/day) plus placebo; (2) paroxetine (20 to 40 mg/day) plus placebo; or (3) mirtazapine plus paroxetine. The group receiving combined therapy of mirtazapine obtained a significantly larger reduction in depressive symptoms than the groups that received mirtazapine and paroxetine monotherapy groups. The remission rate for the combined therapy group at week 6 was 43 percent, as compared to 19 and 26 percent for the mirtazapine and paroxetine monotherapy groups, respectively. Although again not statistically significant, such a difference in remission rates would be considered clinically significant if confirmed in a larger study.

A second study of the value of combining antidepressants from the outset of therapy randomly assigned 105 MDD outpatients to receive 6 weeks of therapy with one of four strategies. One group received fluoxetine (20 mg/day) plus placebo; this group served as the active comparator. A second group received fluoxetine (20 mg/day) in combination with mirtazapine (30 mg daily); this group directly tested the value of combining mirtazapine with the SSRI. Two other groups received mirtazapine (30 mg/day) in combination with either venlafaxine (225 mg/day) or bupropion (150 mg/day); these groups permitting a preliminary comparison of the utility of combining mirtazapine with other first-line antidepressants. This study again tested the hypothesis that patients receiving combination antidepressant therapy would show a more rapid or more complete remission than a monotherapy from the outset of therapy. At the end of 6 weeks of double-blind therapy, the fluoxetine plus placebo group had a remission rate of 25 percent. Across all outcome measures, trends favored the groups that received mirtazapine in combination with fluoxetine (52 percent remission rate), venlafaxine XR (58 percent remission rate), or bupropion (46 percent remission rate).

The combination of mirtazapine and venlafaxine was evaluated in two other studies. The first study was part of the massive STAR*D project. In this open label trial, 109 MDD outpatients who had not obtained adequate benefit from three rounds of antidepressant therapy were randomized to receive 12 weeks of therapy with mirtazapine (mean dose: 36 mg/day) and venlafaxine XR (mean dose: 210 mg/day) or the monoamine oxidase inhibitor tranylcypromine (mean dosage: 37 mg/day). Outcomes were assessed by self-report and by blinded independent evaluators. At study endpoint, neither group did particularly well, though the remission rate in

the mirtazapine plus venlafaxine XR group was twice that of the group that received the MAOI (14 vs. 7 percent). Patients receiving combination therapy were more likely to complete the trial and did show a statistically significant difference in improvement in depressive symptoms. A relatively few patients completed 12 weeks treatment with tranylcypromine at a therapeutic dose (i.e., at least 40 mg/day); one implication is that physicians in contemporary practice are better able to implement a trial of combined therapy than execute a successful switch to an MAOI.

A subgroup of the STAR*D investigators completed a second, larger study of the value of combining antidepressants at the beginning of treatment. This trial, known by the acronym COMED, randomized 220 MDD outpatients to 12 weeks of treatment with escitalopram plus placebo (i.e., the active comparison group), mirtazapine (mean dose: 18 mg/day) plus venlafaxine XR (mean dose: 178 mg/day), or escitalopram plus bupropion (i.e., an alternate widely used combination that had performed reasonably well in STAR*D). This trial found no evidence to support the routine use of antidepressant combinations from the outset of therapy, which reported significantly greater side effects.

Prophylactic Therapy

Two studies have evaluated longer-term treatment with mirtazapine. The first report examined the outcomes of patients who received extended therapy after responding to double-blind, acute-phase therapy with mirtazapine, amitriptyline, or placebo. Results indicated that patients who remained on mirtazapine had a significantly lower risk of relapse than patients who continued to take double-blind placebo. The amitriptyline-treated group had an intermediate outcome.

The second study was a formal relapse prevention trial in which 156 patients who had remitted during 8 to 12 weeks of acute-phase therapy were randomized to remain on active mirtazapine (N = 76) or to switch to placebo (N = 80). Across the subsequent 40 weeks of double-blind therapy, 44 percent of the placebo group relapsed, compared to 20 percent of the group receiving active continuation-phase therapy. The effects of continuation-phase mirtazapine therapy were of the same magnitude observed in studies of other effective antidepressants.

OTHER INDICATIONS FOR MIRTAZAPINE

Mirtazapine is not formally indicated for treatment of any other disorder. Nevertheless, mirtazapine may be useful for the treatment of insomnia, panic disorder, generalized anxiety disorder, PTSD, and OCD, as well as autism and other pervasive developmental disorders. A number of small studies and case series have evaluated the utility of mirtazapine in treating anxiety and depression in cancer patients. Blockade of 5-HT3 and

histamine receptors may have particularly useful effects in the symptomatic management of anxiety and depression in cancer patients. Improvements in pain, nausea, appetite, insomnia, depression, and anxiety have been reported.

A study of 12 adult outpatient subjects with comorbid MDD/alcohol dependence was recently conducted. Subjects were given 15 mg of mirtazapine for the first 2 weeks and 30 mg for the last 6 weeks of the medication trial. Mirtazapine improved depression on the Beck Depression Inventory (BDI), and decreased the number of alcohol drinks per week. These preliminary findings suggest that mirtazapine may be efficacious in reducing both depressive symptoms and excessive alcohol use in comorbid MDD and alcohol dependence. Two years later, 10 of the 12 patients who entered the acute-phase study participated in the follow-up study. The significant improvements initially observed in depressive symptoms (BDI), drinking, and sleep disturbance persisted at the follow-up evaluation. A study of depressed cocaine-dependent subjects who received either mirtazapine (11 subjects) or placebo (13 subjects) for 12 weeks suggested that mirtazapine is superior to placebo in improving sleep in patients with comorbid depression and cocaine dependence, but is not more effective than placebo in reducing cocaine use.

In another study, sixteen agitated patients with Alzheimer disease (AD) were enrolled in a 12-week open-label pilot study. Patients were given 15 to 30 mg/day of mirtazapine. The study found that agitated patients with AD could be treated with mirtazapine without significant adverse events or cognitive deterioration. Larger-scale studies are now warranted.

MIRTAZAPINE USE DURING PREGNANCY AND LACTATION

In a study of eight women who were taking a median dose of 38 mg of mirtazapine per day while breastfeeding, concentrations of mirtazapine or desmethylmirtazapine were measured in milk and plasma. All infants were clinically examined and four infants had blood drawn for analysis. No adverse effects were seen. Mirtazapine was detected in only one of four infants tested. The study suggested that mirtazapine use by lactating women is safe for the breast-fed infant. Nevertheless, each decision to breastfeed should always be made on the basis of an individual risk–benefit analysis.

It is not known whether mirtazapine may cause harm to the unborn baby but a study assessed the risks associated with exposure to this antidepressant. Pregnant depressed women taking mirtazapine were compared to pregnant women taking other antidepressants and pregnant women exposed to nonteratogens. About 100 pregnant women in each group were followed. Mirtazapine did not appear to increase the rate of major malformations. However, the rate of prematurity was higher in the mirtazapine group compared to pregnant women exposed to

nonteratogens.

TOLERABILITY, SAFETY, AND TOXICITY

Attrition from mirtazapine therapy because of adverse events generally was less than that observed in comparison groups treated with tricyclic antidepressants, but somewhat greater than that observed in comparative studies versus SSRIs. Mirtazapine and venlafaxine have comparable attrition rates in controlled studies. The specific side-effect profile of mirtazapine is quite different, however, from that associated with SSRIs and SNRIs.

The most common side effects reported during mirtazapine therapy are sedation, increased appetite, weight gain, and dry mouth. The drug's antihistaminic effects primarily cause these side effects. Approximately 50 percent of patients report sedation during the first week of mirtazapine therapy. This side effect can be clinically significant, as demonstrated in studies of reaction time or automotive driving performance. Sedation usually is not minimized by slow upward titration of the dose, and lower and higher starting doses of mirtazapine have similar rates of sedation. If early sedation can be tolerated, however, most patients adapt to this side effect without difficulty. When compared to SSRIs or SNRIs, mirtazapine therapy is associated with significantly lower rates of gastrointestinal disturbance and a lower incidence of treatment-emergent sexual dysfunction.

Mirtazapine provoked periodic leg movements (PLMS) in 67 percent of young healthy males. The effect was most pronounced in the first few days but gradually decreased in the first week. The indirect increase in central dopamine via 5-HT1A receptor activation may explain why mirtazapine-induced PLMS decreased over time. Further, diphenhydramine, another first-generation H1 antagonist, provoked PLMS during wake in subjects with restless legs syndrome, pointing to a possible antihistaminergic mechanism in the induction of PLMS.

Mirtazapine may cause withdrawal syndrome upon sudden discontinuation. Symptoms include depression, anxiety, restlessness, irritability, decreased appetite, insomnia, diarrhea, nausea, and vomiting. Therefore, reduction in dose should be gradual and slow.

Mirtazapine is not approved for use in children under the age of 18 years.

Perhaps 15 to 25 percent of patients gain an unacceptable amount of weight during mirtazapine therapy. This annoying adverse effect may occur more commonly among younger patients, and, interestingly, weight gain was not a problem in one study of depression in late life. The likelihood of problematic weight gain usually can be predicted within the first 4 to 6 weeks of therapy. A small study of seven women with MDD treated with mirtazapine for 6 weeks concluded that mirtazapine was associated with a

significant increase in body weight, body mass fat and leptin concentration. Interestingly, glucose metabolism remained unaffected in this study, in contrast to other psychotropic medications inducing weight gain.

The effects of mirtazapine and venlafaxine on leptin plasma levels were recently contrasted. After 28 days of treatment, mirtazapine treatment was strongly associated with an increase in leptin concentrations. There was also a significant effect on the leptin/body mass index (BMI) ratio elevated in mirtazapine-treated patients and slightly declining in venlafaxine-treated patients. Blockade of histamine synthesis may contribute to leptin resistance, as well as via possible histamine H1-receptor–mediated dysregulation of hypothalamic nuclei integrating central and peripheral signals relevant to energy balance.

The only significant effect of mirtazapine therapy on clinical chemistry profiles has been increased blood lipid levels. As this effect occurred with approximately twice the frequency seen in patients randomly assigned to treatment with amitriptyline, weight gain alone does not appear to be the cause.

The most controversial severe adverse effect possibly associated with mirtazapine therapy is agranulocytosis. Although concern has existed about this as a potential problem in the early (premarketing) studies, a definite association has not been established over the subsequent two decades.

Mirtazapine is remarkably safe in overdose. The principal problems following ingestion of as much as 5 g of mirtazapine are sedation, aspiration pneumonia, delirium, and tachycardia. No cases of seizure or abnormal heart rhythms have been reported. In the only reported fatal overdose, the patient also had ingested toxic levels of several other drugs.

FUTURE DIRECTIONS

Mirtazapine is distinctly different from the remainder of the newer antidepressants. It is at least as effective as the SSRIs, and most likely conveys earlier symptomatic benefit. It also differs dramatically from the SSRIs and venlafaxine with respect to tolerability. Virtually free from causing side effects commonly associated with 5-HT uptake inhibitors (e.g., gastrointestinal disturbances, insomnia, and sexual dysfunction), mirtazapine therapy is complicated by sedation (in as many as half of all patients) and weight gain (in as many as a fourth of all patients). Mirtazapine is increasingly used for treatment of older depressed people, for whom these effects are more likely to be therapeutic than problematic.

One study demonstrated that the interaction between HTR1A+272G>A and 5-HTT VNTR genotypes may be involved in the response to mirtazapine treatment in MDD patients. The response was greater for the combination than for individual genotypes. This observation should be replicated and further expanded in larger-scale studies.

REFERENCES

Afshar M, Knapp CM, Sarid-Segal O, et al. The efficacy of mirtazapine in the treatment of cocaine dependence with comorbid depression. *Am J Drug Alcohol Abuse*. 2012;38(2):181–186.

Anttila SA, Leinonen EV. A review of the pharmacological and clinical profile of mirtazapine. *CNS Drug Rev*. 2001;7(3):249–264.

Benkert O, Muller M, Szegedi A. An overview of the clinical efficacy of mirtazapine. *Hum Psychopharmacol*. 2002;17(Suppl 1):S23–S26.

Benkert O, Szegedi A, Kohnen R. Mirtazapine compared with paroxetine in major depression. *J Clin Psychiatry*. 2000;61(9):656–663.

Benkert O, Szegedi A, Philipp M, et al. Mirtazapine orally disintegrating tablets versus venlafaxine extended release: a double-blind, randomized multicenter trial comparing the onset of antidepressant response in patients with major depressive disorder. *J Clin Psychopharmacol*. 2006;26(1):75–78.

Blier P, Gobbi G, Turcotte JE, et al. Mirtazapine and paroxetine in major depression: a comparison of monotherapy versus their combination from treatment initiation. *Eur Neuropsychopharmacol*. 2009;19:457–465.

Blier P, Ward HE, Tremblay P, Laberge L, Hébert C, Bergeron R. Combination of antidepressant medications from treatment initiation for major depressive disorder: a double-blind randomized study. *Am J Psychiatry*. 2010;167(3):281–288.

Bruijn JA, Moleman P, Mulder PG, et al. A double-blind, fixed blood-level study comparing mirtazapine with imipramine in depressed in-patients. *Psychopharmacology (Berl)*. 1996;127(3):231–237.

Carpenter LL, Yasmin S, Price LH. A double-blind, placebo-controlled study of antidepressant augmentation with mirtazapine. *Biol Psychiatry*. 2002;51:183–188.

Chang HS, Lee HY, Cha JH, et al. Interaction of 5-HTT and HTR1A gene polymorphisms in treatment responses to mirtazapine in patients with major depressive disorder. *J Clin Psychopharmacol*. 2014;34(4):446–454.

Clayton AH, Montejo AL. Major depressive disorder, antidepressants, and sexual dysfunction. *J Clin Psychiatry*. 2006;67(Suppl 6):33–37.

Cornelius JR, Douaihy AB, Clark DB, et al. Mirtazapine in comorbid major depression and alcohol use disorder: a long-term follow-up study. *J Addict Behav Ther Rehabil*. 2013;3(1):1–7.

de Boer T. The pharmacologic profile of mirtazapine. *J Clin Psychiatry*. 1996;57(Suppl 4):19–25.

Djulus J, Koren G, Einarson TR, et al. Exposure to mirtazapine during pregnancy: a prospective, comparative study of birth outcomes. *J Clin Psychiatry*. 2006;67(8):1280–1284.

Fava M, Rush AJ, Wisniewski SR, et al. A comparison of mirtazapine and nortriptyline following two consecutive failed medication treatments for depressed outpatients: a STAR*D report. *Am J Psychiatry*. 2006;163(7):1161–1172.

Fulda S, Kloiber S, Dose T, et al. Mirtazapine provokes periodic leg movements during sleep in young healthy men. *Sleep*. 2013;36(5):661–669.

Guelfi JD, Ansseau M, Timmerman L, Kørsgaard S; Mirtazapine-Venlafaxine Study Group. Mirtazapine versus venlafaxine in hospitalized severely depressed patients with melancholic features. *J Clin Psychopharmacol*. 2001;21(4):425–431.

Kim SW, Shin IS, Kim JM, Park KH, Youn T, Yoon JS. Factors potentiating the risk of mirtazapine-associated restless legs syndrome. *Hum Psychopharmacol*. 2008;23(7):615–620.

Kristensen JH, Ilett KF, Rampono J, Kohan R, Hackett LP. Transfer of the antidepressant mirtazapine into breast milk. *Br J Clin Pharmacol*. 2007;63(3):322–327.

Laimer M, Kramer-Reinstadler K, Rauchenzauner M, et al. Effect of mirtazapine treatment on body composition and metabolism. *J Clin Psychiatry*. 2006;67(3):421–424.

McGrath PJ, Stewart JW, Fava M, et al. Tranylcypromine versus venlafaxine plus mirtazapine following three failed antidepressant medication trials for depression: a STAR*D report. *Am J Psychiatry*. 2006;163(9):1531–1541.

Montgomery SA, Reimitz PE, Zivkov M. Mirtazapine versus amitriptyline in the long-term treatment of depression: a double-blind placebo-controlled study. *Int Clin Psychopharmacol*. 1998;13(2):63–73.

Nutt D. Mirtazapine: pharmacology in relation to adverse effects. *Acta Psychiatr Scand Suppl*. 1997;391:31–37.

Papakostas GI, Homberger CH, Fava M. A meta-analysis of clinical trials comparing mirtazapine with selective serotonin reuptake inhibitors for the treatment of major depressive disorder. *J Psychopharmacol*. 2008;22(8):843–848.

Rush AJ, Trivedi MH, Stewart JW, et al. Combining medications to enhance depression outcomes (CO-MED): acute and long-term outcomes of a single-blind randomized study. *Am J Psychiatry*. 2011;168:689–701.

Schatzberg AF, Kremer C, Rodrigues HE, Murphy GM Jr; Mirtazapine vs. Paroxetine Study Group. Double-blind, randomized comparison of mirtazapine and paroxetine in elderly depressed patients. *Am J Geriatr Psychiatry*. 2002;10(5):541–550.

Schilling C, Gilles M, Blum WF, et al. Leptin plasma concentrations increase during antidepressant treatment with amitriptyline and mirtazapine, but not paroxetine and venlafaxine: leptin resistance mediated by antihistaminergic activity? *J Clin Psychopharmacol*. 2013;33(1):99–103.

Schittecatte M, Dumont F, Machowski R, et al. Mirtazapine, but not fluvoxamine, normalizes the blunted REM sleep response to clonidine in depressed patients: implications for subsensitivity of alpha(2)-adrenergic receptors in depression. *Psychiatry Res*. 2002;109(1):1–8.

Shen J, Chung SA, Kayumov L, et al. Polysomnographic and symptomatological analyses of major depressive disorder patients treated with mirtazapine. *Can J Psychiatry*. 2006;51(1):27–34.

Stahl S, Zivkov M, Reimitz PE, Panagides J, Hoff W. Meta-analysis of randomized, double-blind, placebo-controlled, efficacy and safety studies of mirtazapine versus amitriptyline in major depression. *Acta Psychiatr Scand Suppl*. 1997;391:22–30.

Thase ME, Nierenberg AA, Keller MB, Panagides J; Relapse Prevention Study Group. Efficacy of mirtazapine for prevention of depressive relapse: a placebo-controlled double-blind trial of recently remitted high-risk patients. *J Clin Psychiatry*. 2001;62(10):782–788.

Thase ME, Nierenberg AA, Vrijland P, van Oers HJ, Schutte AJ, Simmons JH. Remission with mirtazapine and selective serotonin reuptake inhibitors: a meta-analysis of individual patient data from 15 controlled trials of acute phase treatment of major depression. *Int Clin Psychopharmacol*. 2010;25:189–198.

Wheatley DP, van Moffaert M, Timmerman L, Kremer CM. Mirtazapine: efficacy and tolerability in comparison with fluoxetine in patients with moderate to severe major depressive disorder. Mirtazapine-Fluoxetine Study Group. *J Clin Psychiatry*. 1998;59(6):306–312.

▲ 34.22 Monoamine Oxidase Inhibitors

ETHAN F. KARSEN, M.D., AND PAUL E. HOLTZHEIMER, M.D., M.S.

INTRODUCTION AND HISTORY

Monoamine oxidase inhibitors (MAOIs) were among the first antidepressants found to be clinically effective in humans. Although MAOI use has declined over time due to the emergence of newer antidepressants with equivalent efficacy and more tolerable side effects, research interest in MAOIs has been rekindled by the development of selective MAOIs, new delivery systems, and the discovery of their neuroprotective role in neurodegenerative diseases.

Development of MAOIs

In 1957, at a regional meeting of the American Psychiatric Association (APA) in Syracuse, New York, two independent research groups presented findings showing that the antitubercular medication iproniazid had antidepressant effects in patients with tuberculosis and comorbid depression. These findings represented a culmination of years of laboratory and clinical research. Emil Fischer first isolated hydrazines, the class to which early MAOIs belong, in the 1870s. In the 1920s and 1930s, several research groups helped identify the enzyme monoamine oxidase (MAO), and showed that it is capable of metabolizing adrenaline through oxidation. In 1952, Albert Zeller demonstrated that iproniazid acts as an MAO inhibitor. The same year, physicians at Staten Island's Seaview Hospital observed that iproniazid had a more powerful stimulating effect than isoniazid in the patients they treated for tuberculosis. The first study of MAOI use in patients without tuberculosis was also presented at the APA meeting in 1957. The study showed substantial improvement of symptoms in 70 percent of a group of psychotically depressed patients taking iproniazid. By 1958, over 400,000 patients were being treated with iproniazid for depression, then marketed as Marsilid. The clinical use of iproniazid, however, was short-lived as reports of jaundice and hepatotoxicity accumulated, and in 1961 it was withdrawn from the US market.

While iproniazid was entering and leaving the market, pharmaceutical companies were developing several new MAOIs. Hydrazine derivatives similar to iproniazid included: isocarboxazid (Marplan), phenelzine (Nardil), nyalamide, mebanazine, and pheniprazine. Like iproniazid, many

of these agents showed hepatotoxicity and limited clinical use. In 1959, researchers discovered that tranylcypromine (Parnate), a cyclopropylamine structurally unrelated to the hydrazines, had MAOI action. Soon after its introduction, however, reports appeared of hypertensive crises and severe headaches associated, with subarachnoid intracranial hemorrhage. Barry Blackwell, a psychiatric resident at Maudsley Hospital in London, traced these effects back to ingestion of certain cheeses in patients taking MAOIs, thereafter known as the "cheese effect." Further investigation showed that food rich in amines, especially tyramine, could precipitate hypertensive crises in patients treated with MAOIs. Due to the concern for hemorrhage, tranylcypromine was briefly removed from the US market in 1964, only to be reintroduced a few months later after postmarketing analysis showed an incidence rate of less than 0.001 percent. Thereafter, all MAOIs included labeling to include precautions on adverse reactions and suggested dietary restrictions.

The first-generation MAOIs are irreversible inhibitors of MAO; they covalently bond to and inactivate the enzyme. After the discover of two isoforms of MAO, MAO-A and MAO-B (discussed in more detail later in this chapter), in the 1970s, researchers sought to synthesize MAOIs that would be reversible and selective inhibitors with fewer adverse effects. The use of platelets helped identify molecules selective for MAO-B, with evidence suggesting correlation between platelet MAO-B inhibition and antidepressant response to phenelzine. Pargyline was the first FDA-approved selective MAO-B inhibitor, although it did not show significant antidepressant effects in clinical trials and is no longer available. In 1972, researchers described L-deprenyl (Eldepryl), an irreversible but selective MAO-B inhibitor. However, in clinical trials L-deprenyl demonstrated antidepressant effect only at doses at which it loses its selectivity for MAO-B. This finding suggested that inhibition of MAO-A, alone or in combination with MAO-B inhibition, is necessary for antidepressant effect. Interestingly, experimental findings from the 1970s found L-deprenyl to be an effective adjunct to L-dopa in treatment of patients with Parkinson disease (PD). The FDA licensed the drug in 1996 and renamed it selegiline. Rasagiline (Azilect), with theoretically less adverse effects than selegiline, followed as another irreversible MAO-B inhibitor for use in PD.

Following the suggested association between MAO-A inhibition and antidepressive effects, researchers in 1979 identified clorgyline, an irreversible but selective MAO-A inhibitor. It proved effective for depression, but caused hypertensive crises as it irreversibly bound MAO in the gut. Moclobemide, initially developed for hyperlipidemia, became the first reversible and selective MAO-A inhibitor; it is currently available for use in depression in over 50 countries worldwide. Reversible inhibitors of monoamine oxidase type A (RIMA) are less likely to cause hypertensive crises because they are readily displaced by tyramine in the gut. There are

no RIMAs currently available in the United States (see Table 34.22–1 for MAOI drugs available worldwide). The five MAOIs currently available in the United States are isocarboxazid, phenelzine, tranylcypromine, selegiline, and rasagiline (Fig. 34.22–1).

Table 34.22–1.
Classification of Monoamine Oxidase Inhibitors

Drug	Reversibility	Selectivity	Indication
Iproniazid[a] (Marsilid)	Irreversible	MAO-A MAO-B	Depression
Isocarboxazid (Marplan)	Irreversible	MAO-A MAO-B	Depression
Phenelzine (Nardil)	Irreversible	MAO-A MAO-B	Depression
Tranylcypromine (Parnate)	Irreversible	MAO-A MAO-B	Depression
Isoniazid (Nydrazid)	Irreversible	MAO-A MAO-B	Antituberculosis
Nialamide[a] (Niamid)	Irreversible	MAO-A MAO-B	Depression
Procarbazine (Matulane)	Weak irreversible	MAO-A MAO-B	Antineoplastic
Clorgyline[b]	Irreversible	MAO-A only	Depression
Selegiline (Deprenyl)	Irreversible	MAO-B only	Depression, Parkinson disease
Rasagiline (Azilect)	Irreversible	MAO-B only	Parkinson disease
Pargyline[a] (Eutonyl)	Irreversible	MAO-B only	Antihypertensive
Linezolid	Reversible	MAO-A MAO-B	Antibiotic
Lazabemide[c] (Pakio)	Reversible	MAO-B only	Parkinson disease
Moclobemide[c] (Aurorix/Manerix)	Reversible	MAO-A only	Depression
Brofaromine[c] (Consonar)	Reversible	MAO-A only	Depression
Befloxatone[c] (Synthelabo)	Reversible	MAO-A only	Depression
Pirindole[c] (Pirazidol)	Reversible	MAO-A only	Depression
Toloxatone[c] (Humoryl)	Reversible	MAO-A only	Depression

[a]Discontinued.

[b]Used in research only.

[c]Available outside the United States only.

CHEMISTRY

MAO is a flavoprotein enzyme system located on the outer mitochondrial membrane of both neurons and nonneuronal cells. In neurons, MAO metabolizes three major neurotransmitters: serotonin, dopamine, and norepinephrine. MAO also serves several other important roles, including the metabolism of trace amines in the brain, the peripheral inactivation of biogenic and dietary amines, phase I metabolism, and as an imidazoline

receptor.

MAO contains an intracellular active site that functions to metabolize amines through oxidation (Fig. 34.22–2). MAO transfers two hydrogen atoms to a flavin cofactor (FAD). The reduced flavin is reoxidized by oxygen producing hydrogen peroxide plus the imine, which spontaneously hydrolyzes to the aldehyde (or a substituted amine, in the case of secondary amines like adrenaline) plus ammonia. The aldehyde product is further metabolized for urine excretion.

MAO comes in two forms, MAO-A and MAO-B, encoded by separate genes on chromosome X, originally distinguished by their differing sensitivities to the inhibitors clorgyline and l-deprenyl. However, researchers did not publish structures of MOA-A and MAO-B (which share 70 percent DNA similarity) until 2005 and 2002, respectively.

Researchers have identified four polymorphisms in the gene coding for MAO-A as possible risk factors for psychiatric disorders. The most studied of these polymorphisms is MAO-A-uVNTR, which has been associated with depression in male patients, aggression and impulsivity levels, autism spectrum disorders, alcoholism and suicide attempts. Studies have also found polymorphisms in intron 13 and exon 14 of MAO-B to be associated with increased risk for PD.

In the central nervous system (CNS), both MAO-A and MAO-B serve to regulate the content of intracellular amine stores, terminate actions of amine neurotransmitters, and protect neurons from exogenous amines. In addition to MAO metabolism, neurotransmitter action may be terminated by catecholamine-o-methyl-transferase (COMT) or vesicular repackaging. Under normal conditions, neurotransmitters are preferentially stored in vesicles, where they are not exposed to MAO. As a result, MAO maintains low neurotransmitter concentration within the cytoplasm of neurons.

MAO-A preferentially deaminates serotonin and norepinephrine and predominates in catecholaminergic neurons in the substantia nigra, locus cerulerus, and periventricular regions of the hypothalamus. MAO-B preferentially deaminates phentylamine and benzylamine and predominates in serotonergic neurons in the dorsal raphe nucleus. Both forms metabolize dopamine, tyramine, and tryptamine. That MAO subtypes populate neurons with neurotransmitters of low affinity, suggests a role to prevent build-up of nonnative neurotransmitters.

In addition to the major monoamines, MAO also serves to metabolize other amines: phenylethanolamine, meta- and para-tyramine and octamine, which are found in the brain at lower levels and collectively known as trace amines. These trace amines can markedly affect uptake and release of catecholamines and serotonin at nerve endings. There has been increased research into the signaling roles of trace amines with the discovery of specific G-protein-coupled receptors for trace amines in the mammalian brain, which may help regulate activity of monoamine

transporters.

In the periphery, MAO serves to rapidly inactivate biogenic and dietary amines. The most well-known example is MAO-A's role in the intestinal tract. MAO-A inactivates ingested tyramine from the diet, preventing its release into the general circulation. When absorbed, tyramine causes significant release of norepinephrine for peripheral adrenergic neurons leading to vasoconstriction, increased heart rate, and elevated blood pressure. MAO also deactivates peripheral serotonin, epinephrine, and norepinephrine as it does in the CNS.

In addition to its role in amine regulation, MAO is involved in phase I metabolism of xenobiotics. Like the cytochrome P450 system, MAO can decrease bioactivity. For example, MAO-B in the liver oxidizes sertraline, citalopram, sumatriptan, and phenylephrine.

Finally, recent research has identified MAO as an imidazoline (I2) receptor (IR). The clinical relevance of this finding remains unclear, however there is some emerging evidence that it may play a role in depression. While IRs are found throughout the brain and on platelets, compounds that have I2-selectivity appear to allosterically inhibit both MAO isozymes, which may explain some anxiolytic and antidepressant effects seen in mice. Additionally, rat models show that chronic treatment with irreversible MAOIs reduces the density of I2 sites. Finally, studies found increased density of 45-kd I2 receptors in platelets of depressed patients and prefrontal cortices of suicide victims when compared to controls.

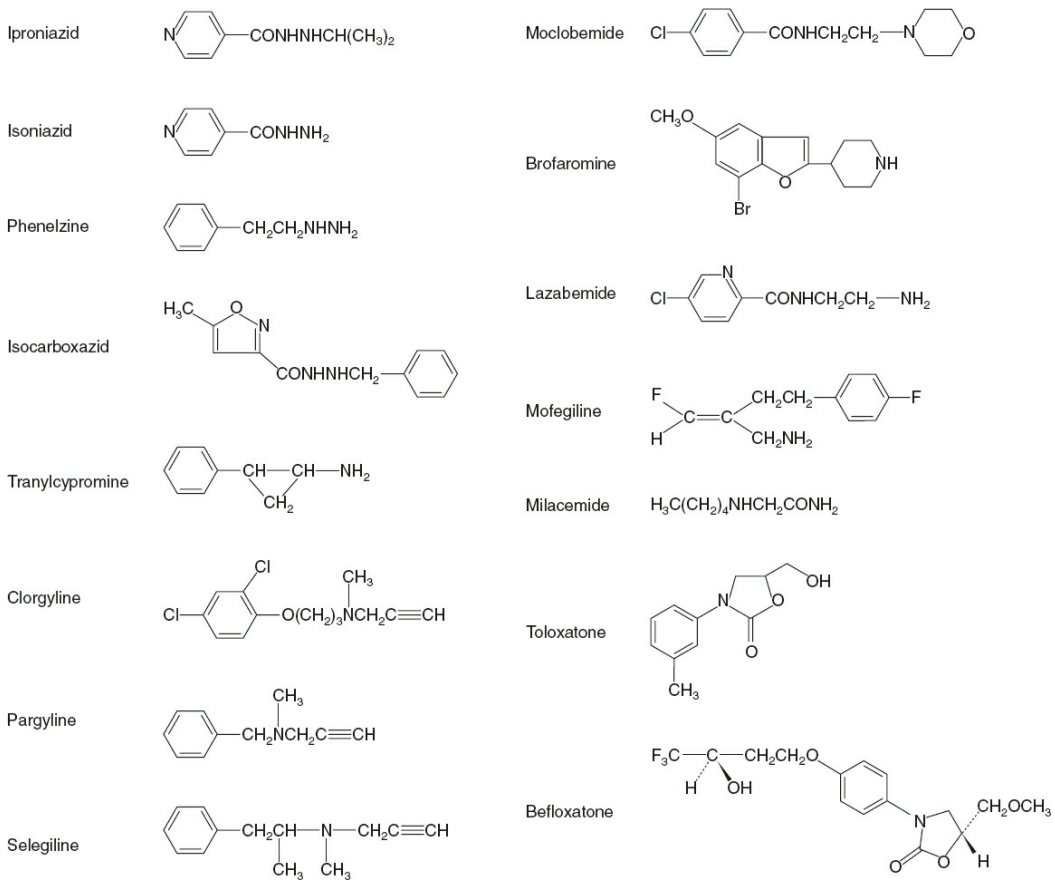

FIGURE 34.22–1. Molecular structures of monoamine oxidase inhibitors.

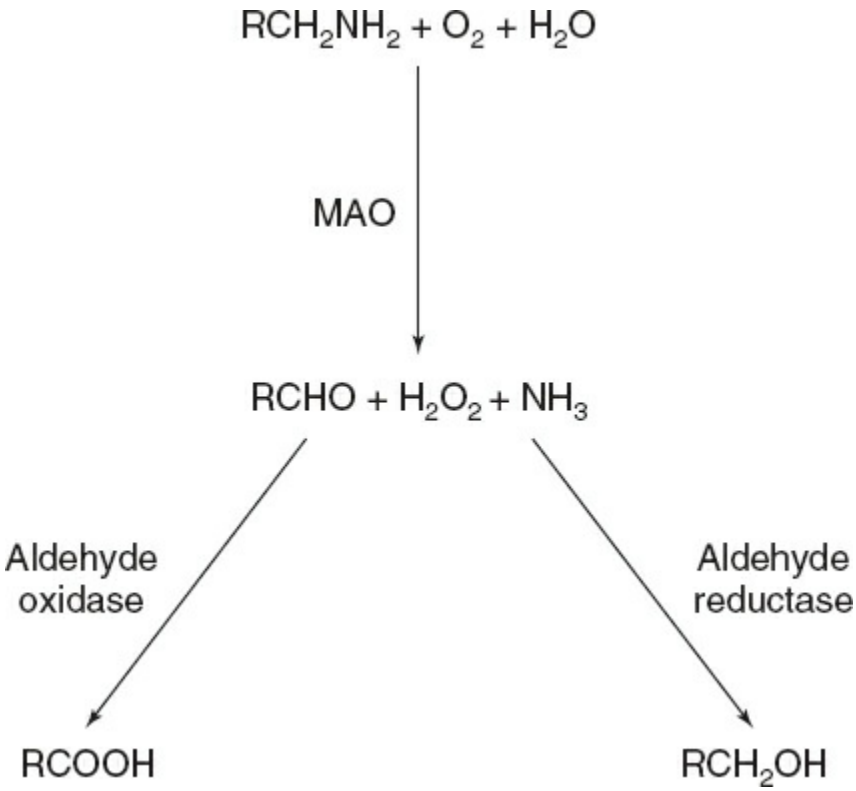

FIGURE 34.22–2. The reaction catalyzed by monoamine oxidase (MAO), acting on a primary amine.

MECHANISM OF ACTION

The proposed mechanism of MAOIs efficacy in depression has evolved since their discovery and expanded with use in other neuropsychiatric disorders.

Several early observations led to the "catecholamine hypothesis" or "monoamine theory" of depression, including the effectiveness of MAOIs in depression, the nearly concurrent discovery of TCAs and their mechanism of norepinephrine reuptake inhibition, and reserpine's depressogenic effect via monoamine depletion. The theory postulates that depression in humans is caused by decreased levels of norepinephrine in the synaptic cleft, and that MAOIs alleviate depression by direct inhibition of MAO and an increase in catecholamines available within the synapse.

Monoamines, especially dopamine, serotonin and norepinephrine have long been known to play an important role in the behavioral regulation of mammals. For instance, in humans, low levels of amines in the brain are associated with hyperactivity and depression, and elevated amines are associated with risk-taking behavior. However, increased monoamine levels can be measured within hours after the first dose of an MAOI, while antidepressant effects take 2 to 4 weeks. Further studies suggest that secondary adaptations to long-term medication therapy may be responsible for the antidepressant effect including: reduced density of alpha-2, beta-adrenergic and 5-HT1 and 5-HT2 receptors; correction of hypothalamic–pituitary–adrenal (HPA) axis overactivation; increases in BDNF mediated signaling; and induction of hippocampal neurogenesis. There is also evidence that tranylcypromine at sufficiently high concentrations may act as a monoamine reuptake inhibitor and phenelzine may inhibit GABA transaminase, leading to increased levels of GABA within the brain.

Despite evidence questioning a simplistic monoamine deficiency model, recent imaging studies have renewed interest in the monoamine theory of depression. Researchers using positron emission tomography (PET) have shown increased MAO-A binding (suggesting decreased presence of monoamines) in the brains of medication-free patients during a major depressive episode compared with healthy controls. A follow-up study found that successful SSRI treatment did not result in normalization of MAO-A levels and that patients who experienced recurrent depressive episodes had increased MAO-A binding in the prefrontal cortex and anterior cingulate cortex compared to those who maintained remission. The authors go on to suggest an advanced monoamine model for depression that relies on chronic monoamine loss and associated changes as the primary mechanism.

As clinicians found use for selegiline in PD, researchers sought a mechanism of action different from that for depression. MAO activity is thought to have an important role in pathogenesis of PD. Postmortem studies have shown increased MAO-B activity in the substantia nigra in

patients with PD. It is postulated that heightened enzyme activity increases reactive oxygen species formed by metabolism of dopamine, leading to oxidative damage to neurons. Increased MAO-B activity also promotes aging of the brain by increasing hydrogen peroxide levels and activating other neurotoxins.

MAO-B inhibitors such as selegiline and rasagiline are thought to be effective for treating PD by enhancement of striatal dopamine activity, through antioxidation, and antiapoptotic activity. Chronic use of both selegiline and rasagiline has been shown to increase dopamine levels in rat striatum. The proposed mechanism is prevention of metabolism of B-phenylethylamine, a trace amine that stimulates release of dopamine and inhibits neuronal uptake. Selegiline and rasagiline may also protect neurons against damage from reactive oxygen species generated by hydrogen peroxide created by MAO-B metabolism of dopamine. Rasagiline may also possess added neuroprotective properties related to antiapoptotic pathways through activation of Bcl-2 and Bcl-X1 and preventing collapse of the mitochondrial membrane.

Finally, selegiline is metabolized to amphetamine compounds, which may explain its effectiveness in the treatment of attention deficit hyperactivity disorder (ADHD).

CLASSIFICATION

MAOIs may be classified according to chemical structure, reversibility, or selectivity (Table 34.22–1). Isocarboxazid and phenelzine are hydrazine compounds, containing a nitrogen–nitrogen pair on their side chain. All other MAOIs are classified as nonhydrazines. Irreversible MAOIs covalently bind to the enzyme, permanently inhibiting MAO function until new MAO can be produced, usually taking 2 to 3 weeks. Reversible MAOIs can be displaced by tyramine from the active site of MAO-B, allowing it to be metabolized. Tyramine is not present in high concentrations in the brain, and therefore does not inhibit MAOI effects on neurotransmitters. Nonselective MAOIs target both MAO-A and MAO-B, while selective MAOIs target a single isozyme.

Pharmacological Actions of Specific Agents

Tranylcypromine (Parnate). Tranylcypromine is Food and Drug Administration (FDA)-approved for use in major depressive episode without melancholia. Off-label uses include: other depressive disorders, panic disorder, social anxiety disorders, and post-traumatic stress disorder (PTSD).

PHARMACODYNAMICS

Mechanism of Action. Tranylcypromine is a irreversible and nonselective

nonhydrazine MAOI. It covalently binds to both MAO-A and MAO-B. Of note, as a nonhydrazine, its antidepressant effects tend to have more rapid onset.

PHARMACOKINETICS

Absorption. Tranylcypromine is readily absorbed from the GI tract.

Distribution. There is no clear data on distribution of tranylcypromine.

Metabolism and Elimination. Tranylcypromine has peak plasma concentrations 1 to 3 hours after ingestion. Its elimination half-life is 2.5 hours. Metabolites are excreted in the urine.

DOSING. Dosing generally starts at 10 mg twice a day (see Table 34.22–2). If there is no improvement within 2 to 3 weeks, then the dose can be increased by 10 mg every 2 to 3 weeks with a total maximum daily dose of 60 mg.

Tranylcypromine is contraindicated in patients with abnormal liver function tests, hepatic impairment, or a history of liver disease. Tranylcypromine is also contraindicated in patients with severe renal impairment. In patients with mild to moderate renal impairment, tranylcypromine should be used with caution due to the potential for accumulation of active metabolites.

Table 34.22–2.
Typical Dosage Forms and Recommended Dosages for Currently Available Monoamine Oxidase Inhibitors

Drug	Usual Dose (mg/day)	Maximum Dose (mg/day)	Dosage (Oral) Formulation
Tranylcypromine (Parnate)	20–60	60	10-mg tablets
Phenelzine (Nardil)	30–60	90	15-mg tablets
Isocarboxazid (Marplan)	20–40	60	10-mg tablets
Selegiline (Eldepryl)	10	30	5-mg tablets
Transdermal selegiline (Emsam)	6	12	6, 9 or 12 mg patch
Rasagiline	0.5–1.0	1.0	0.5- or 1.0-mg tablets

MONITORING PARAMETERS. Complete metabolic panel at baseline, 1 and 3 months.

PREGNANCY AND LACTATION. Tranylcypromine is classified as FDA pregnancy risk category C. According to the manufacturer, tranylcypromine is excreted in human milk and the potential benefit of the drug should be weighed against the potential hazards to the mother and infant during breast-feeding.

COMMON ADVERSE EFFECTS. Agitation, anxiety, constipation, diarrhea, dizziness, headache, impotence, insomnia, loss of appetite, mania, nausea, orthostatic hypotension, somnolence, weight gain, and xerostomia.

Phenelzine. Phenelzine is FDA-approved for use in depressed patients clinically characterized as "atypical," "nonendogenous," or "neurotic." Off-label uses include: other depressive disorders, panic disorder, social anxiety disorder, and obsessive-compulsive disorder (OCD).

PHARMACODYNAMICS

Mechanism of Action. Phenelzine is an irreversible and nonselective hydrazine MAOI. It covalently bonds both MAO-A and MAO-B.

PHARMACOKINETICS

Absorption. Phenelzine is well-absorbed following oral administration.

Distribution. There is insufficient data on phenelzine's distribution.

Metabolism and Elimination. Peak levels are reached in approximately 2 to 4 hours. The mean half-life is approximately 12 hours. Phenelzine is eliminated in the urine as metabolites.

DOSAGE. Initial dosage of phenelzine is generally 15 mg two to three times per day. It may be increased rapidly, as tolerated, to a maximum dose of 30 mg three times a day. Many patients do not show response until treatment at 60 mg has been continued for at least 4 weeks. Historically, it has been recommended that after the maximum effect is achieved, the dosage should be slowly reduced to the lowest dose needed to maintain clinical effectiveness.

Due to the potential for drug accumulation and associated effects, phenelzine is not recommended in patients with hepatic disease or severe renal impairment.

PREGNANCY AND LACTATION. Phenelzine is classified as FDA pregnancy risk category C. According to the manufacturer, the safe use of phenelzine during breast-feeding has not been established, and the potential benefit of the drug should be weighed against the possible hazards to the mother and nursing infant.

MONITORING PARAMETERS. Complete metabolic panel at baseline, 1 and 3 months.

COMMON ADVERSE EFFECTS. Constipation, disorder of ejaculation and/or orgasm, dizziness, edema, fatigue, headache, hyperreflexia, impotence, elevated values on liver function tests, orthostatic hypotension, sleep disorders, somnolence, tremor, weight gain, and xerostomia.

Isocarboxazid. Isocarboxazid is FDA-approved for use in major depressive disorder. Off-label use includes: panic disorder and social anxiety disorder.

PHARMACODYNAMICS

Mechanism of Action. Isocarboxazid is an irreversible and nonselective hydrazine MAOI. It covalently bonds both MAO-A and MAO-B.

PHARMACOKINETICS

Absorption. Isocarboxazid is readily absorbed from the GI tract.

Distribution. There is insufficient data on distribution of isocarboxazid.

Metabolism and Elimination. Isocarboxazid reaches peak plasma concentrations in 3 to 5 hours. The half-life is approximately 2.5 hours. It is metabolized by the liver and excreted as metabolites in the urine.

DOSAGE. Initial dose of isocarboxazid is 10 mg twice a day. The dose may be increased 10 mg daily up to maximum dosage of 60 mg total daily in divided doses. If therapeutic response is not seen in 3 to 6 weeks, it is unlikely to develop. After maximum benefit is achieved, the dose should be reduced to the lowest effective dose.

Isocarboxazid is not recommended in hepatic impairment or renal impairment.

PREGNANCY AND LACTATION. Isocarboxazid is classified as FDA pregnancy risk category C. According to the manufacturer, the amount of excretion of isocarboxazid and its metabolites in human milk have not been determined; therefore, isocarboxazid should only be used during breast-feeding if clearly needed.

MONITORING PARAMETERS. Complete metabolic panel at baseline, 1 and 3 months.

COMMON ADVERSE EFFECTS. Anxiety, blurred vision, constipation, dizziness, headache, insomnia, mania, somnolence, weight gain, and xerostomia.

Selegiline. Selegiline is FDA-approved for use in PD in combination with levodopa. Off label use includes: Alzheimer's disease, depression, and ADHD.

PHARMACODYNAMICS

Mechanism of Action. Selegiline is an irreversible and selective MAO inhibitor. At doses lower than 10 mg it covalently binds MAO-B. At doses above 10 mg it covalently binds both MAO-A and MAO-B.

PHARMACOKINETICS

Absorption. Selegiline is readily absorbed in the GI tract.

Distribution. Up to 85 percent of plasma selegiline is reversibly bound to proteins.

Metabolism and Elimination. Peak serum concentrations are found in 0.5 to 2 hours. At steady state, elimination half-life is approximately 10 hours. The drug is rapidly and completely metabolized to three active derivatives with the following half-lives: N-desmethyldeprenyl, 2 hours; 1-amphetamine, 17.7 hours; and 1-methamphetamine, 20.5 hours.

DOSAGE. Effective antidepressant doses of oral selegiline are not usually achieved unless doses in the range of 20 to 30 mg/day are used, at which point selegiline loses its selectivity and requires dietary monitoring

In PD, selegiline dosing is 5 mg after breakfast and 5 mg after lunch.

For ADHD, trials have used 2.5 mg twice a day for children under 5 years old, and 5 mg twice a day for children over 5 years old.

Specific guidelines for dosage adjustments of selegiline in hepatic impairment and renal impairment are not available.

PREGNANCY AND LACTATION. Selegiline is classified as FDA pregnancy risk category C. According to the manufacturer, it is unknown if selegiline is excreted into breast milk, and a decision should be made whether to discontinue the drug or discontinue breast-feeding, taking into account the importance of the drug to the mother.

MONITORING PARAMETERS. None.

COMMON ADVERSE EFFECTS. Nausea, dizziness, and insomnia.

Transdermal Selegiline. Transdermal selegiline is FDA- approved for use in major depressive disorder.

PHARMACODYNAMICS

Mechanism of Action. Transdermal selegiline is an irreversible and selective MAOI. At doses less than 6 mg it covalently binds MAO-B only. At doses above 9 mg it covalently binds MAO-A and MAO-B.

PHARMACOKINETICS

Absorption. Approximately 23 to 30 percent of the medication is absorbed through the skin over 24 hours.

Distribution. Once absorbed it is 90 percent protein bound, rapidly distributed throughout the body, and crosses the blood–brain barrier. Transdermal administration avoids first-pass metabolism and results in

much higher bioavailability (74 percent) compared to the oral form (4 percent).

Metabolism and Elimination. Selegiline is metabolized via N-dealkylation or N-depropargylation by several CYP450 enzymes (2A6, 2B6, 2C9, 3A4, and 3A5). Under steady-state conditions, the median elimination half-life increases to 10 hours. The half-life is approximately 18 to 25 hours. The final metabolites are methamphetamine and amphetamine, which are excreted in urine.

DOSAGE. Initial dose is 6 mg/24 hr. Although maximum dose is 12 mg, there is little evidence for doses above 6 mg. If clinically indicated, dose increases should be made in 3 mg increments every 2 to 3 weeks. In elderly patients, the recommended dose is 6 mg per 24 hours. At dose above 9 mg, the manufacturer suggests following a diet modified for reduced tyramine intake.

Dose adjustments for transdermal selegiline are not needed in patients with mild to moderate hepatic impairment or any degree of renal impairment.

PREGNANCY AND LACTATION. Transdermal selegiline is classified as FDA pregnancy risk category C. According to the manufacturer, it is not known if transdermal selegiline is excreted into human milk, and caution is advisable if administering selegiline to a nursing woman.

MONITORING PARAMETERS. None.

COMMON ADVERSE EFFECTS. Application site reaction, decreased systolic blood pressure, diarrhea, headache, indigestion, insomnia, orthostatic hypotension, weight loss, and xerostomia.

Rasagiline. Rasagiline is FDA-approved for use in PD as monotherapy or an adjunct to levodopa.

PHARMACODYNAMICS

Mechanism of Action. Rasagiline is an irreversible and selective MAOI. It covalently binds MAO-B.

PHARMACOKINETICS

Absorption. Rasagiline is rapidly absorbed and reaches peak levels in 1 hour.

Distribution. Plasma protein binding ranges from 88 to 94 percent.

Metabolism and Elimination. Rasagiline is metabolized by the cytochrome P450 system in the liver, predominately CYP1A2, and is mainly excreted via

urine.

DOSING. For monotherapy in PD, rasagiline may be started at 1 mg daily. In patients taking levodopa, the recommended initial dose is 0.5 mg daily, which if tolerated, may be increased to 1 mg daily if needed.

In patients taking a CYP1A2 inhibitor such as ciprofloxacin, dosages should not exceed 0.5 mg daily.

In patients with mild hepatic impairment, dosage should not exceed 0.5 mg daily and it should not be used in patients with moderate or severe hepatic impairment.

No dose adjustment is required for patients with mild or moderate renal impairment and it has not been studied in patients with severe renal impairment.

At recommended doses, rasagiline retains MAO-B selectivity and although the manufacturer does not recommend dietary restrictions, it suggests foods with high tyramine content may cause hypertensive cries.

PREGNANCY AND LACTATION. Rasagiline is pregnancy category C and there are no adequate studies of use of the medication in pregnant women. It is not known whether rasagiline is excreted in human milk.

MONITORING PARAMETERS. Liver function test at baseline, 1 and 3 months.

COMMON ADVERSE EFFECTS. Weight loss, orthostatic or postural hypotension, dry mouth, abnormal involuntary movements, and hallucinations.

Drug Level Monitoring

MAOI plasma levels may be measured but there is little evidence that levels correspond to treatment response and may be difficult to interpret given the short half life of MAOIs.

There is evidence that platelet MAO inhibition is a predictor of response, however its use is not recommended. Several studies have shown that patients taking phenelzine who have more than 80 percent platelet MAO inhibition after 2 weeks of treatment show higher response rates. However, researchers point out the moderate effect size means that this accounts for only 9 to 16 percent of treatment outcomes. MAO-B platelet inhibition is most likely an epiphenomenon as it has been shown in studies of selegiline that 100 percent inhibition can occur in the absence of antidepressant effects.

Laboratory Interference

Laboratory interference is not a major problem with use of MAOIs. They may lower blood glucose values. They have also been associated with a

minimal false elevation in thyroid function test results.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Central Nervous System

Commonly reported CNS effects include: dizziness, headache, drowsiness, sleep disturbance, fatigue, weakness, twitching, myoclonus, and hyperreflexia. Uncommon but serious CNS effects include: ataxia, shock-like coma, toxic delirium, acute anxiety reaction, psychosis, transient respiratory and cardiovascular depression following electroconvulsive therapy, confusion, memory impairment, hypomania, and mania. Rasagiline may cause increased impulsive or compulsive behaviors such as gambling, sexual activity, and binge eating.

Cardiovascular

Common side effects include: orthostatic and postural hypotension, syncope, palpitations, and tachycardia.

Skin

Adverse dermatological effects include rash, pruritus, and hyperhidrosis. Rasagiline may increase risk for melanoma and regular screening is recommended.

Gastrointestinal

Commonly reported GI effects include: constipation, diarrhea, nausea, xerostomia, and elevated hepatic enzymes. Severe but less frequent events include glottis edema, progressing necrotizing hepatocellular damage, and reversible jaundice, the last two more likely with hydrazine derivatives.

Genitourinary

Common urinary symptoms include urinary retention and hesitancy. Commonly reported sexual side effects include impotence, ejaculation dysfunction, and anorgasmia. Clinical experience shows less sexual side effects with tranylcypromine. There have been case reports of priapism in patients taking phenelzine. Impaired water excretion compatible with syndrome of inappropriate secretion of antidiuretic hormone (SIADH) has been associated with antidepressant use.

Optic

Glaucoma and blurred vision have been infrequently reported.

Hematologic

There are reports of rare occurrences of anemia, leukopenia, agranulocytosis, and thrombocytopenia.

Table 34.22–3.
Relative Restrictions in Food and Beverages with Irreversible MAOI Use

Restriction	Foods
Absolute (high risk)	Aged cheeses; aged and cured meats; banana peel; broad bean pods; improperly stored or spoiled meats, poultry or fish; Marmite; sauerkraut; soy sauce and soybean condiments; tap beer
Moderate (moderate risk)	Red or white wine; bottle or canned beer
Unnecessary (low to no risk)	Avocados; bananas; beef/chicken bouillon; chocolate; fresh and mild cheeses; fresh meat, poultry and fish; gravy; monosodium glutamate; peanuts; properly stored pickled or smoked fish; soy milk; yeast extracts other than Marmite

Neurologic

Tremor is a commonly reported side effect of phenelzine. Peripheral neuropathy, sometimes associated with a pyridoxine deficiency, has been reported in patients receiving phenelzine. There are case reports that suggest an association between MAOI use and pseudoparkinsonism with symptoms of bradykinesia, tremor, cogwheel rigidity, and dysarthria.

WARNINGS AND PRECAUTIONS

Food-Drug Interactions

Nonselective and irreversible MAOIs, when combined with food rich in tyramine, can cause increased release of catecholamines, resulting in hypertensive crisis characterized by an occipital headache, palpitations, nausea, vomiting, diaphoresis, and tachycardia. Severe reactions can lead to cardiac arrhythmias, stroke, and intracranial hemorrhage. Prior to recommendations on dietary restrictions, rates of hypertensive reactions were estimated to occur in 2.4 to 25 percent of patients treated with MAOIs.

After observation of hypertensive reactions in the 1960s, clinical use of MAOIs dropped dramatically. At the height of concern over drug–food interactions, more than 70 different types of food and beverages were on the list of foods to be avoided. However, over the past 20 years, critical review of reports of adverse effects and analysis of tyramine levels in food and beverages has refined the list of dietary restrictions (Table 34.22–3).

Drug–Drug Interactions

Hypertensive Crisis. In addition to tyramine-induced hypertensive crises, there are similar dangers in drug–drug interactions. Any medication

that increases adrenergic stimulation by a mechanism other than MAO inhibition can lead to hypertensive crisis (Table 34.22–4). The list includes decongestants that contain alpha-1 agonists, stimulants that increase release of norepinephrine, antidepressants that block norepinephrine reuptake, and appetite suppressants and analgesics with norepinephrine-reuptake inhibition.

Serotonin Syndrome. Concomitant administration of serotonergic agents with MAOIs, especially those with MAO-A inhibition, can result in dangerous levels of serotonin leading to serotonin syndrome. Increased activation at postsynaptic serotonin receptors causes a disruption in thermoregulation leading to a syndrome of fever, altered mental status, myoclonus, diaphoresis, and diarrhea. In extreme cases symptoms may include hypothermia, seizures, coma, cardiovascular collapse, permanent brain damage, and death. MAOIs should not be prescribed in combination with SSRIs, SNRIs, TCAs, certain narcotics, and other medications with serotonergic properties (Table 34.22–5).

Table 34.22–4.
Medications Contraindicated with MAOI Use due to Risk of Hypertensive Crisis

Decongestants: phenylephrine, oxymetazoline, ephedrine, pseudoephedrine, phenylpropanolamine
Stimulants: methylphenidate, amphetamines
Antidepressants: SSRIs, SNRIs, TCAs, bupropion, buspirone
Appetite suppressants: sibutramine, phentermine
Analgesics: tramadol
Centrally acting antihypertensives: guanethidine, methyldopa
Antiparkinson agents: levodopa

Table 34.22–5.
Medications Contraindicated with MAOI Use due to Risk for Serotonin Syndrome

Antidepressants: SSRIs, SNRIs, TCAs
Other TCAs: cyclobenzaprine, carbamazepine
Appetite suppressant: sibutramine
Cough suppressant: dextromethorphan
Opioids: meperidine, tramadol, methadone, propoxyphene
Antihistamines: chlorpheniramine, brompheniramine

In addition, care must be taken when changing antidepressants to or from MAOIs. A washout period of 2 weeks is required for all SSRIs prior to starting an MAOI, except for fluoxetine, which due to its long half-life, requires a 5-week washout period. A 2- to 4-week waiting period is also

required when switching from an MAOI to another antidepressant or when discontinuing an MAOI prior to ending the restricted diet. The selegiline patch loses its selectivity at higher doses, so the manufacturer recommends dietary restrictions with doses of 9 mg or higher and for 2 weeks following any discontinuation.

Weight Gain

In a retrospective review, weight gain occurred in 35.7 percent of patients taking phenelzine and 10.5 percent taking tranylcypromine.

Orthostatic Hypotension

All nonselective MAOIs may reduce blood pressure, likely due to central alpha-adrenergic agonist effects of increased norepinephrine levels. Rates of orthostasis may be higher in phenelzine. Hypotension usually develops gradually over first 2 to 3 weeks of treatment and correlates with peak plasma levels. The side effect may be reduced with divided dosing and resolves hours to days after discontinuation.

Insomnia

MAOIs, especially tranylcypromine, are associated with insomnia. MAOIs are known to reduce the time spent in REM sleep. Transdermal selegiline also shows this side effect in a dose-dependent relationship. Insomnia may be alleviated by avoiding evening dosing.

Application Site Reaction

The most common side effect to transdermal selegiline is an application site reaction, which occurs in a dose-dependent relationship, and occurs in 24 percent of patients.

Hepatotoxicity

Postmarketing monitoring showed increased risk for hepatotoxicity for hydrazine (isocarboxazid and phenelzine) MAOIs compared to the nonhydrazine (tranylcypromine). However absolute risk appears to be low (less than 1 in 30,000).

Psychosis

The combination of MAOIs and dextromethorphan has been reported to cause brief episodes of psychosis and bizarre behavior.

Overdose

Signs and symptoms of overdose include: dilated pupils, fever, diaphoresis, drowsiness, agitation, headache, hallucinations, trismus, opisthotonos,

seizures, coma, fever, tachycardia, hypertension, hypotension, respiratory depression/failure, shock, and death. Signs and symptoms may be absent or minimal during the initial 12-hour period following ingestion and may develop slowly, reaching a maximum in 24 to 48 hours. In some cases, symptoms may persist for 8 to 14 days. Disseminated intravascular coagulation, rhabdomyolysis, and acute tubular necrosis can also occur. Treatment of overdose is mainly supportive. Inducing vomiting or gastric lavage with charcoal may be helpful in early poisoning.

Withdrawal

Common withdrawal symptoms include nausea, vomiting, and malaise. Abrupt withdrawal for patients on long-term treatment with MAOIs can result in a withdrawal syndrome characterized by increased arousal, mood disturbance, psychosis, and convulsions within 24 to 72 hours after drug discontinuation.

Contraindications

Contraindications to MAOI use include cardiac disease, cerebrovascular disease, hypertension, headache, hepatic disease, pheochromocytoma, renal disease, and general anesthesia with certain opioid analgesics.

THERAPEUTIC INDICATIONS AND CLINICAL DRUG STUDIES

A significant number of studies on MAOIs occurred prior to publication of DSM-II, making interpretation difficult. Reviewed here are studies with clear inclusion criteria, and when possible, a control group. Limitations are noted where they arise. Meta-analyses are generally not used as they usually include RIMA studies. Evidence for efficacy of MAOIs that are not available in the United States, most notably moclobemide and brofaromine, is not discussed.

MAOIs were first used psychiatrically to treat patients with depression. However, given concerns over efficacy, side effects, drug–drug interactions, and dietary restrictions, MAOIs as a class generally fell out of favor. SSRIs have largely replaced MAOIs for psychiatric use. MAOIs are most commonly prescribed for atypical major depressive disorder and treatment-resistant depression. Less common uses include typical major depressive disorder, psychotic depression, dysthymia, bipolar depression, seasonal affective disorder, borderline personality disorder (BPD), eating disorders, social anxiety disorder, panic disorder, obsessive-compulsive disorder, and ADHD. Selective MAO-B inhibitors are used regularly in treatment of PD and Alzheimer disease (AD).

Atypical Depression

Early observations on groups of depressed patients who responded to

MAOIs but not to TCAs sparked interest in identifying the characteristics of this subset of depressed patients. Studies reported that these "atypical" patients had depressed mood with anxiety and phobia or reactive mood with reversed vegetative symptoms. Subsequent controlled trials by a group at Columbia codified patients with atypical features of depression, defining criteria as reactive mood plus at least two of four associated features: hyperphagia, hypersomnolence, leaden paralysis, and sensitivity to rejection. They also identified a second group characterized by depression plus panic attacks for which MAOIs also appeared to show advantage of TCAs. However replication studies by the same group failed to show a difference between MAOIs and TCAs in mixed depressive/anxiety states.

Acute Treatment. The Columbia group summarized their findings in a 1993 paper. From data of 269 patients, they compared response rates on Clinical Global Improvement scales for patients taking phenelzine (72 percent), imipramine (44 percent), and placebo (26 percent). They concluded that rather than enhanced response to MAOIs patients with atypical depression show substandard response to TCAs. Through stratification, they also demonstrated that presence of anxiety symptoms did not factor into response rates, despite previous suggestions of superior MAOI response in depression with anxiety.

Since that review, two controlled studies have been published, one comparing MAOI to cognitive-behavior therapy (CBT) and another to an SSRI. A 1999 study randomized patients with atypical depression to phenelzine, CBT, or placebo for 10 weeks. Response rates for phenelzine (58 percent) and CBT (58 percent) were equivalent to each other and superior to placebo (28 percent). Analysis found similar results for reduction in Hamilton Rating Scale for Depression (HAM-D) scores. Patients taking phenelzine had higher rates of marked side effects (92 percent) compared to placebo (53 percent).

The only trial that has compared an MAOI available in the United States to an SSRI is a randomized, double-blind, non–placebo controlled trial comparing fluoxetine and phenelzine. Forty patients with atypical depression randomized to fluoxetine or phenelzine showed no difference in response (80 percent vs. 85 percent) or remission (80 percent vs. 70 percent). The only patient who dropped out due to an adverse effect (confusion) was taking phenelzine. Dizziness, asthenia, arthralgias, somnolence, and blood pressure changes were more common in phenelzine group. Tremor and weight gain were more common in the fluoxetine group.

Maintenance Treatment. In addition to acute atypical depression treatment, there is one long-term study of relapse prevention in atypical depression. In this study, 60 patients who had improved with phenelzine or imipramine over 6 months were randomly assigned to same medication or

placebo for an additional 6 months. Recurrence rates were 23 percent for those receiving phenelzine, 41 percent for imipramine, 47 percent for those who switched from imipramine to placebo, and 87 percent for those who switched from phenelzine to placebo.

Treatment-Resistant Depression

Recently, the use of MAOIs in typical depression has been limited to treatment-resistant depression (TRD). Most controlled trials have focused on tranylcypromine after early case studies found benefit from tranylcypromine for patients with TCA-refractory depression. Response rates in these trials vary greatly, which may reflect the difficulty in defining this group, and it is a challenge to make firm conclusions about specific advantages to MAOIs in TRD.

In a 1988 study, 21 patients with major depressive disorder who had not responded to oxaprotiline of fluvoxamine were randomized to tranylcypromine or nomifensine. Compared to the nomifensine group, the tranylcypromine group showed statistically superiority in reduction on HAM-D scores but not response rate (45 percent vs. 10 percent) at 4 weeks.

A 1993 study compared brofaromine and tranylcypromine in 39 patients with major depressive disorder who had previously not responded to 4 weeks of maprotiline or nortriptyline. Of note, 12 patients had psychotic features and 23 melancholic features. There were no difference in response rates for brofaromine (45.5, 59.1 percent) and tranylcypromine (52.9, 58.9 percent) as measured by Hamilton Depression Subscale and Clinical Global Improvement, respectively. Melancholic features did not predict response. Brofaromine was better tolerated.

A second study in 1993 also compared brofaromine and tranylcypromine. In this case, 93 inpatients who had failed two classes of antidepressants were treated for 6 weeks. There were no differences between the two groups on reduction in HAM-D scores or response rates (nearly 74 percent for both groups). There were no differences in tolerability.

A 38-year-old divorced mother of two teenage daughters presented at the Mood Disorders Clinic with a history of failed antidepressant treatment trials. Two years earlier, she had become increasingly irritable, exhausted, and tearful after the breakup of her marriage. She could not tolerate fluoxetine (Prozac) or sertraline and discontinued each of these SSRIs after only a few days because of GI symptoms. Although she could tolerate desipramine (Norpramin) up to 300 mg daily, she felt no benefit after 4 months and was reluctant to continue this TCA medication because of adverse effects (weight gain and sweating, in particular). Because of her increasing fatigue, increasing sleep needs, and generally atypical profile, she was offered a trial of phenelzine. After 4 weeks, at a dosage of 60 mg daily, she reported significant improvement in her energy level, mood, and sleep pattern. Two years later, she continues to take 45 mg of phenelzine daily with some adverse effects (occasional insomnia and anorgasmia).

A 2004 study enrolled 77 patients who had failed either of two phases: (1) a double-blind randomized trial of imipramine or fluvoxamine, followed by augmentation with lithium or (2) an open trial with a TCA. They were randomized to either tranylcypromine or phenelzine for 5 weeks. Results showed no difference in response rate between tranylcypromine (44 percent) and phenelzine (47 percent) or remission rates (18 and 11 percent, respectively). About 21 percent of patients reported "severe" side effects including dizziness, agitation, and insomnia, though there were no differences between the two groups.

Level 4 of the STAR*D trial compared tranylcypromine versus extended-release venlafaxine plus mirtazapine in 109 patients who had not responded to three prior medication trials. The results showed no difference in response (13.8 percent vs. 15.7 percent) or remission (6.9 percent vs. 13.7 percent) between the two groups. Patients in the tranylcypromine group were more likely to exit the study (41.4 percent vs. 21.6 percent) and almost half had less than 6 weeks of treatment, suggesting that MAOI treatment may less tolerable than other antidepressant treatment. However, patients did not differ in ratings on side effect burden scales. Critics of the study point out that average tranylcypromine levels were only 36.9 mg, less than in similar studies, and may explain the lower response rates.

Finally, a crossover study compared selegiline to placebo in 16 treatment-resistant (failed at least two clinical trials of antidepressant) older patients with depression. Patients received selegiline 60 mg or placebo in two 3-week phases. During the active treatment phase, patients showed significant improvement in HAM-D scores. Overall 8 of 16 patients responded to treatment. Responders were more likely to be older, less depressed, less anxious, and more cognitively intact. Researchers also performed cognitive tests and with no changes during active treatment. There were no differences in side effect rates in each group.

Major Depressive Disorder

Early clinical observation, suggested that MAOIs tend to be more effective than TCAs in treatment of atypical depression and less effective than TCAs in melancholic or endogenous depressions. A particularly large and influential early study from the United Kingdom showed phenelzine 45 mg to be marginally inferior to placebo in a group of depressed inpatients of unspecified type. Subsequent trials with higher doses showed more equivalent efficacy for MAOIs to TCAs in treatment of endogenous or melancholic depression. Discussed here are trials investigating groups of patients with unspecified type of major depressive disorder or endogenous/melancholic subtype. Of note, there are no trials comparing MAOIs available in the United States with second-generation antidepressants.

Acute Treatment. Early studies of MAOI use in depression were characterized by limited size, short duration, and low dosages. A 1960 study crossed over 25 moderately or severely endogenously depressed patients to receive phenelzine 45 mg or placebo for 2 weeks each. Improvement rates were significantly higher for the second phenelzine group compare to placebo, but not the first. In a 1961 study, 21 consecutively admitted patients diagnosed with depression (13 with endogenous type) were given phenelzine or placebo for 3 weeks and then switched. On a seven-point improvement scale, a higher proportion of patients showed "complete recovery" or "marked improvement" in the phenelzine (50 percent) group versus placebo (15 percent). Fifty-three percent of endogenous depressed patients improved or remitted.

Later trials appear to show equivalence for acute depression, in endogenous or mixed groups, when comparing MAOIs to TCAs, without clear evidence of difference in side-effect burden. In 1981, researchers randomized 49 inpatients diagnosed with depression (33 nonendogenous) to phenelzine or imipramine for 3 weeks after a 1-week placebo-response period. Both groups showed an equal and statistically significant response as measured by the HAM-D. There was no subgroup analysis.

A 1986 study of 75 depressed patients over age 55 compared phenelzine and nortriptyline to placebo. Response rates found phenelzine and nortriptyline superior to placebo (59.1 percent vs. 60.0 percent vs. 10.7 percent) as measured by HAM-D. A majority of patients in each group had endogenous subtype but a subgroup analysis was not performed. Of note, only two patients from each active arm dropped out of the study due to side effects.

A 1987 paper investigated use of phenelzine or imipramine for 6 weeks after a 1-week placebo-response period in 83 patients with a diagnosis of major depression with melancholia or dysthymic disorder. In subgroup analysis for melancholia, reduction in scores on the HAM-D did not differ between phenelzine (58.2 percent) and imipramine (71.4 percent) treatment. Side effects were not reported.

Finally, two short-term studies established the efficacy of transdermal selegiline for major depressive disorder. A 6-week trial of 6 mg versus placebo showed significant results (9 vs. 6.5 point change on the HAM-D) and an 8-week trial with doses from 6 to 12 mg also showed significant changes (10.9 vs. 8.6 point changes on HAM-D). Oral selegiline has been found to be effective in treating depression only at doses higher than 20 mg daily, a dose where selective enzyme inhibition is lost, and tyramine-restricted diet required.

Maintenance Treatment. Few studies have investigated long-term use of MAOIs for the prevention of recurrent depressive episodes.

A 1991 study looked at maintenance on phenelzine versus placebo. After

acute unblinded treatment with phenelzine, responders were entered into a 16-week continuation phase followed by double-blind randomization to phenelzine 60 mg, phenelzine 45 mg or placebo for a 2-year maintenance phase. Forty-seven patients entered the maintenance phase. Both phenelzine doses showed significant protection against relapse within 3 months but not against recurrence (symptom onset beyond 3 months of continuation of treatment). Generalization from this study is limited as 36 percent of patients who initially responded to treatment discontinued during continuation therapy prior to enrollment in maintenance phase.

Fifty-two adults over 55 years old with depression who had responded to acute treatment with phenelzine or nortriptyline in the 1986 study were randomized to maintenance on the study drug or tapered to placebo. Results showed significantly less recurrence for patients taking phenelzine (13.3 percent) versus nortriptyline (53.8 percent) and placebo (65.2 percent) over 1 year. The phenelzine group had the only two dropouts due to side effects.

A 52-week study investigated relapse rates for 322 patients who had responded to acute treatment with transdermal selegiline 6 mg daily, who were randomized to continued treatment or placebo. Significantly fewer patients on transdermal selegiline (16.8 percent) relapsed compared to placebo (30.7 percent).

Psychotic Depression

It is unclear that patients diagnosed with psychotic depression in early trials would meet contemporary criteria, and these early trials found no efficacy for MAOIs in treatment of psychotic depression. There are no controlled studies. The only recent study is, an open-label trial examining use of phenelzine in depressed inpatients. The researchers classified patients into three groups: definitely psychotic, probably psychotic, and nonpsychotic. Response rates were 22, 43, and 68 percent, respectively. This inpatient study showed increased effectiveness for phenelzine in nonpsychotic versus psychotic patients. MAOIs are not usually considered for this indication.

Dysthymia

There is limited evidence for the use of phenelzine in dysthymic disorder. Three prospective studies have investigated the use of MAOIs for dysthymia, but only one used an MAOI available in the United States. A 1993 study comparing phenelzine to imipramine in dysthymic disorders over 6 weeks found superior response in the phenelzine group when measured by response on the HAM-D.

A retrospective analysis by the Columbia group studying atypical depression identified 37 patients meeting DSM-III criteria for dysthymic

disorder. The researchers found response rates of 22, 50, and 17 percent for phenelzine, imipramine, and placebo, respectively. However this study was limited by a small number of subjects, short duration of only 6 weeks for a chronic illness, and exclusion of patients with melancholic features. The same group found in a pooled group analysis that phenelzine was more effective than imipramine and placebo (83 percent vs. 45 percent vs. 21 percent) in double depression.

Bipolar Depression

There are three positive prospective studies for MAOI use in bipolar depression, one negative prospective study and one positive retrospective analysis.

A 1982 study on "anergic" depression, defined as depressed mood with motor retardation, fatigue, volitional inhibition, and reversed neurovegetative features, included 59 patients, 29 of which met criteria for bipolar disorder. The other 30 were classified as "pseudounipolar" or "borderline syndromes." Researchers randomized patients to tranylcypromine or placebo. They did not report subgroup analysis. Twenty of twenty-four patients on an active drug responded and only 4 of 31 on placebo. The same group, in 1991, conducted a double-blind study comparing tranylcypromine to imipramine for patients with anergic bipolar depression. Fifty-six patients were treated with either tranylcypromine or imipramine without a mood stabilizer over 6-week acute treatment phase and 10-week maintenance phase. Tranylcypromine outperformed imipramine in acute response (81 percent vs. 48 percent) and in sustained response (71 percent vs. 20 percent). Of note, 21 percent of tranylcypromine and 25 percent of imipramine patients developed mania or hypomania. Significantly more (7 vs. 2) patients taking imipramine dropped out than those taking tranylcypromine. They also conducted a crossover trial involving 16 patients who did not respond to original treatment. Nine of twelve patients who failed imipramine responded to tranylcypromine and had categorical improvement at similar rates to the original tranylcypromine group (75 and 81 percent). Two developed hypomania. Only one of four patients who crossed over to imipramine responded; two withdrew due to side effects and the other withdrew due to a mixed state.

The other prospective trial for bipolar disorder failed. Researchers compared tranylcypromine to lamotrigine in patients with bipolar depression who had previously failed a 6-week trial with an SSRI, TCA, venlafaxine, or bupropion added to an adequate dose of mood stabilizer. This study was severely limited by the number enrolled (20) and the number completed (11). However response, rates, though not statistically significant partly due to the small power of the study, were consistent with the response rates of trials discussed above: 62.5 percent for

tranylcypromine and 36 percent for lamotrigine. Two patients in the lamotrigine group developed mania.

Finally, a retrospective analysis of data from the Maintenance Therapies in Bipolar Disorder (MTBD) study compared the addition of MAOIs or paroxetine in 52 patients with bipolar depression who had already failed to respond on a mood stabilizer alone (usually lithium). Rates of durable recovery (defined as a period of 8 weeks of euthymia starting within the first 16 weeks of treatment) were significantly higher in the MAOI group (70 percent) versus the paroxetine group (33 percent). The authors did not report adverse events.

Seasonal Affective Disorder

In the only study of an MAOI for treatment of seasonal affective disorder, an open trial of tranylcypromine showed remission in 12 of 14 patients treated for "winter depression" as defined by the National Institute of Mental Health (NIMH).

Borderline Personality Disorder

Given overlapping symptoms with atypical depression, there has been some interest in the study of MAOIs for BPD. However only two prospective trials and a reanalysis have been completed. Overall there is weak evidence for use of MAOIs in BPD. The two prospective and one retrospective trial showed improvement on nonspecific or subscales only, without improvement on core borderline personality symptoms.

A 1988 placebo-controlled crossover study investigated four medications for the treatment of BPD: alprazolam, carbamazepine, trifluoperazine, and tranylcypromine. The study assigned 16 patients to complete 6-week trials of each medication. Completion rates were low, an average of 55 percent per arm. Tranylcypromine showed statistically significant improvement compared to placebo when measured by CGI, depression and anxiety subscales, but not on anger, rejection resistivity, impulsivity, suicidality, or behavioral dyscontrol. Tranylcypromine showed a nonsignificant trend toward increased number of physical side effects. The study is limited by its crossover design, short duration, and number of completers.

The second prospective trial, a 1993 study, compared phenelzine and haloperidol to placebo in 108 consecutively admitted patients with BPD over 5 weeks. Phenelzine was superior to placebo on a hostility rating scale but no difference was found on the Borderline Syndrome index, depression, atypical depression, anxiety, or anger scales. A follow-up paper described patients continuing on treatment for an additional 16 weeks. Results found phenelzine to be superior to haloperidol and placebo only on depression and irritability subscales.

Finally, the Columbia group reanalyzed data from their study on treatment of atypical depression with phenelzine, imipramine, or placebo. Diagnosis of comorbid BPD was made with a checklist. They found response rates on CGI for this group of phenelzine (89 percent), imipramine (38 percent), and placebo (20 percent). However, the measures did not include scales for borderline personality symptoms specifically.

Eating Disorders

Two controlled trials and a reanalysis have found efficacy of phenelzine and isocarboxazid in bulimia nervosa. Despite concerns about adherence to diet, the studies reported no hypertensive crises. There are no controlled trials of MAOI use in anorexia nervosa or binge eating disorder.

A 1986 crossover study compared isocarboxazid to placebo over 12 weeks in patients with bulimia nervosa. Isocarboxazid significantly decreased the number of binge episodes (7.0 to 2.8) compared to placebo (7.5 to 4.5). There were no adverse food or drug interactions but 40 percent experienced insomnia or hypotension.

A 1987 study enrolled 82 patients with bulimia to receive phenelzine or placebo over 8 weeks following a 2-week placebo washout phase. Twenty dropped out before randomization and another 12 did not complete at least 2 weeks of treatment. Although number of binges per week at study end did not differ between the two groups, the phenelzine group showed statistically significant reduction in change in binge frequency (64 percent) versus placebo (5 percent). A covariate analysis showed that this response was not related to antidepressant effects. Nine of thirty-one patients taking phenelzine did not complete the 8-week trial due to side effects. There were no reports of hypertensive crisis.

Finally, the Columbia group reanalyzed their data on treatment with phenelzine, imipramine, or placebo for atypical depression. They identified 24 patients who met criteria for comorbid bulimia. The study is limited by outcome measures that relied chart review for notes indicating at least a 50 percent reduction in bingeing and purging and the overeating question on the Hopkins Symptoms Checklist (SCL-90). Phenelzine was found to be statistically superior to both imipramine and placebo on both endpoints.

Posttraumatic Stress Disorder

There are two controlled trials, one positive and one negative, of MAOI use for PTSD. A 1987 study enrolled 13 patients with PTSD in a double-blind crossover trial comparing phenelzine and placebo. Response to phenelzine and placebo did not differ as measured by either PTSD Scale or Impact of Events Scale. Only 10 patients completed treatment and each phase lasted only 5 weeks. A second trial compared phenelzine and imipramine to placebo. At 8-week follow-up, phenelzine showed superiority to both

imipramine and placebo on the Impact of Events Scale, with 45, 25, and 5 percent reductions, respectively. Of note, PTSD was combat-related for all patients in this trial.

Social Anxiety Disorder

Until the introduction of SSRIs, MAOIs were widely considered to be the most effective medication for the treatment of social anxiety disorder. Four of five controlled trials have found positive results for use of phenelzine in social anxiety disorder or social phobia.

In a 1992 study, 74 patients with social phobia were randomized to phenelzine, atenolol, or placebo for 8 weeks. Response rates for phenelzine (64 percent) were statistically higher than atenolol (30 percent) and placebo (20 percent).

A study compared phenelzine and moclobemide to placebo over three 8-week phases. At week 8, both phenelzine and moclobemide showed statistically superior results compared to placebo on the Social Phobia Scale. During the next 8 weeks, mean scores continued to improve in both active treatments with no difference between the two with remittance rates of phenelzine (91 percent), moclobemide (82 percent), and placebo (43 percent). Finally in the last phase during which patients were randomized to continue active treatment or switch to placebo, relapse rates were significantly higher for those switched from either group to placebo. Finally, the phenelzine group showed highest side effect rates and had two severe adverse events, urinary retention and hypertensive crisis, that led to discontinuation of treatment.

The negative result came in a 1991 study that randomized 65 patients with social phobia to weekly cognitive-behavioral group therapy (CBGT) or alprazolam, phenelzine, or placebo with instruction for self-exposure to phobic stimuli for 12 weeks. On self-rating scales of social avoidance, no difference was found between groups. On physician-rated Work and Social Disability Scale, phenelzine and alprazolam were found to be superior to placebo (CBGT was not included as it could not be blinded).

Two more studies compared phenelzine to CBGT. A 1998 study randomized 133 patients with social phobia to CBGT, phenelzine, placebo, or education-supportive group therapy (ES) over 12 weeks. Independent ratings on the Social Phobic Disorders Severity and Change Form found response rates to be statistically significant for phenelzine (65 percent) and CBGT (58 percent) compared to placebo (33 percent) and ES (27 percent). On self-report measures, phenelzine was found to be superior on four of six anxiety rating scales with no difference between the other three groups. No serious adverse effects occurred. Researchers then continued responders in the phenelzine and CBGT group in 6 months of maintenance treatment and 6 months of treatment-free follow-up. Of 28 patients in the maintenance phase, relapse rates did not differ significantly between groups: phenelzine

(21 percent) and CBGT (14 percent). In the follow-up phase there was a nonsignificant trend toward less relapse in CBGT (0 percent) versus phenelzine (30 percent).

The same researchers followed up with a 2011 study that randomized 166 patients to one of four groups: phenelzine, cognitive-behavioral group therapy (CBGT), phenelzine plus CBGT, or placebo. Remission rates on the Liebowitz Social Anxiety Scale (LSAS) were measured at 12 weeks, and found that combined therapy (59.4 percent) was superior to placebo (11.1 percent), which also could not be distinguished from monotherapy phenelzine (20.0 percent) or CBGT (20.6 percent). Insomnia, light-headedness, dry mouth, weight gain, constipation, anorgasmia, and nervousness rates were all higher in patients taking phenelzine.

A 2003 meta-analysis of pharmacological treatment for social anxiety disorder combined results from the two phenelzine trials. The moderate effect size (0.66) was not statistically different from SSRIs, gabapentin, or clonazepam. Despite this equivocal result, the authors concluded that given SSRIs tolerability and safety they should remain first-line treatment for social anxiety disorder.

Obsessive-Compulsive Disorder

Early case reports described patients with OCD who responded to MAOI treatment, however TCAs emerged as the first-line treatment. Surprisingly, there are only two controlled studies, one comparing an MAOI and a TCA, the other an MAOI and an SSRI, with overall inconclusive results.

A 1992 study compared clomipramine to phenelzine without placebo control in 30 patients with OCD. No significant difference was found; each medication decreased OCD symptoms 40 to 50 percent over the 12-week trial period. A larger study of 64 patients compared phenelzine and fluoxetine to placebo. Fluoxetine was found to be superior to both on the Yale-Brown Obsessive Compulsive Scale but not the OCD Scale, NIMH OCD scale, or CGI. No difference was found between phenelzine and placebo on any scale.

Despite these inconclusive results, recently published guidelines on pharmacological management of treatment-resistant OCD list MAOIs as a third-line agent.

Panic Disorder and Agoraphobia

Early studies on MAOI treatment for panic disorder are difficult to interpret as they often clumped phobias together or examined mixed anxiety and depressive states. However the result of these early studies was that clinicians regarded MAOIs as more effective antipanic agents than TCAs.

The first prospective study on “endogenous anxiety” (characterized by

spontaneous panic attacks and agoraphobia) in 1980, compared imipramine and phenelzine to placebo. All three groups received supportive group therapy. At 12 weeks, results showed both active groups to be superior to placebo. The phenelzine group showed trends toward superiority on all measures and statistically significant improvement on Work and Social Disability Scale and Symptom Severity and Phobic Avoidance Scale.

A double-blind study compared four treatment conditions: phenelzine/exposure, phenelzine/no exposure, placebo/exposure and placebo/no exposure in a group of 40 patients with agoraphobia and social phobia. Exposure, with or without phenelzine, was found to be superior in reducing anxiety symptoms compared to no exposure groups.

The Columbia group compared response to phenelzine, imipramine, and placebo in 104 patients with atypical depression with and without spontaneous panic attacks. In patients with panic, response at 6 weeks was significantly higher for phenelzine (77 percent) compared to imipramine (2 percent) and placebo (15 percent) as measured by Clinical Global Improvement. It is unclear from the data whether response is related to decreased panic or depression, although stratified subgroup analysis suggests that response to phenelzine was independent of pretreatment anxiety levels.

Since the introduction in 1981 of panic disorder to DSM-III, there have been no published prospective trials on MAOIs. However, there are two studies that may provide some information.

A 1987 open-label study treated 40 patients with panic disorder or agoraphobia with phenelzine. Thirty-four of thirty-five patients who completed the study showed statistically significant response on the Hamilton-Rating for Anxiety. This study is limited by its open-label design and its high response rate suggests an unaccounted for confounder. Patients also showed significant adverse effects with postural hypotension (64.3 percent), insomnia (44.3 percent), dry mouth (42.9 percent), sexual dysfunction (27.3 percent men, 16.7 percent women), and weight gain (17.1 percent).

A 2010 study compared tranylcypromine 30 mg and 60 mg without placebo control in a group of patients with panic disorder and comorbid social anxiety disorder. Both doses reduced panic attacks about 70 percent from baseline, without statistical significance between the two. However, the higher dose reduced social anxiety symptoms by double the lower dose. Again, adverse effect rates were significant: orthostatic dizziness (58.8 percent), insomnia (47.1 percent), decreased libido (41.2 percent), and somnolence (35.3 percent).

Overall, the evidence is mixed, without clear antipanic effects from phenelzine. Studies are difficult to interpret, given inclusion of comorbidities and mixed treatment modalities. Still, guidelines suggest

MAOIs as third-line or fourth-line agents for panic disorder.

Attention-Deficit/Hyperactivity Disorder

As selegiline is metabolized to amphetamine and methamphetamine, there has been interest in its use for ADHD. Three randomized control trials show selegiline to be superior to placebo in treatment of ADHD and three drug–drug trials found MAOIs to be noninferior to stimulants.

A small crossover trial included 11 children with ADHD (nine with combined type). Each received selegiline 5 mg twice a day or placebo for 4 weeks and then crossed over. While on active treatment, children showed statistically significant improvement on attention and memory but not impulsivity tasks, and improvement on parent–teacher rating scales for inattention and hyperactivity but not impulsivity. There were no differences in side effects.

Another crossover design study used selegiline 5 mg twice a day and placebo in two 8-week periods for children with ADHD and Tourette syndrome. Active treatment showed benefit on ADHD scales only for the MAOI group in the first phase. There was a trend toward improvement in tics that was not statistically significant.

The efficacy of selegiline could not be reproduced for adults with ADHD. Twenty-four adults were randomized to selegiline 60 mg, selegiline 20 mg, or placebo. There was no difference found on ADHD measures after 6 weeks of treatment. Side effects were more frequent in the high-dose group.

Finally, two studies found no difference between selegiline and methylphenidate in treatment of ADHD in children. Importantly, selegiline was better tolerated than methylphenidate, the latter of which had increased rates of headache, decreased appetite, and insomnia. An earlier study also found tranylcypromine, which has similar amphetamine-like properties to selegiline, to be equally efficacious to dextroamphetamine in a crossover study of 14 children with ADHD.

Given positive results versus placebo and noninferiority to stimulants, selegiline appears to be an effective medication for ADHD in children. Importantly, it also is tolerated well. Use in adults is not yet supported. Though one study found tranylcypromine to be noninferior, selegiline has more evidence and is most likely a safer choice for children who do not respond to stimulants.

Neurodegenerative Disorders

Evidence for the use of MAOIs in the treatment of Parkinson and Alzheimer's disease will be covered in detail in another chapter. Please see Chen et al and Youdim et al in the reference section for two recent reviews.

FUTURE DIRECTIONS AND INVESTIGATIONAL DRUGS

Despite evidence for MAOI effectiveness in a number of psychiatric illnesses and consensus guidelines that continue to recommend their use as third- or fourth-line agents, MAOIs have fallen out of favor as SSRI use has increased. A population-based study in Ontario showed that from 1997 to 2007, only 348 older adults were started on an MAOI, accounting for only 1/500 older adults started on an antidepressant. Review of active clinical trials shows only one recently completed study on a RIMA for treatment-resistant depression (Efficacy, Safety, and Tolerability of CX157 in Treatment-Resistant Depression). There is no evidence that RIMAs will be approved for use in the United States.

Although MAOIs have fallen out of favor with prescribing clinicians, research continues in the area. Between 2012 and 2014, researchers filed 36 patents for compounds, with inhibitory effects on MAO. Recent clinical trials have focused on new indications for MAOIs including hair regrowth, prostate cancer, muscular dystrophies, negative symptoms in schizophrenia, and cocaine addictions. Though not covered in this chapter, selective MAO-B inhibitors continue to be investigated for use in Parkinson's disease, Alzheimer's disease, and other neurodegenerative disorders.

REFERENCES

Akhondzadeh S, Tavakolian R, Davari-Ashtiani R, Arabgol F, Amini H. Selegiline in the treatment of attention deficit hyperactivity disorder in children: a double blind and randomized trial. *Prog Neuropsychopharmacol Biol Psychiatry*. 2003;27(5):841.

Amsterdam JD, Bodkin JA. Selegiline transdermal system in the prevention of relapse of major depressive disorder: a 52-week, double-blind, placebo-substitution, parallel-group clinical trial. *J Clin Psychopharmacol*. 2006;26(6):579.

*Atiwannapat P, Arden PC, Stewart JW. Monoamine oxidase inhibitors for various psychiatric disorders and conditions. *Psych Ann*. 2014;44(12):567.

Birkenhäger TK, van den Broek WW, Mulder PG, Bruijn JA, Moleman P. Efficacy and tolerability of tranylcypromine versus phenelzine: a double-blind study in antidepressant-refractory depressed inpatients. *J Clin Psychiatry*. 2004;65(11):1505.

Blanco C, Heimberg RG, Schneier FR, et al. A placebo-controlled trial of phenelzine, cognitive behavioral group therapy, and their combination for social anxiety disorder. *Arch Gen Psychiatry*. 2010;67(3):286.

Blanco C, Schneier FR, Schmidt A, et al. Pharmacological treatment of social anxiety disorder: a meta-analysis. *Depress Anxiety*. 2003;18(1):29.

Buigues J, Vallejo J. Therapeutic response to phenelzine in patients with panic disorder and agoraphobia with panic attacks. *J Clin Psychiatry*. 1987;48(2):55.

Carradori S, Petzer JP. Novel monoamine oxidase inhibitors: a patent review (2012–2014). *Expert Opin Ther Pat*. 2015;25(1):91.

Chen JJ, Swope DM, Dashtipour K. Comprehensive review of rasagiline, a second-generation

monoamine oxidase inhibitor, for the treatment of Parkinson's disease. *Clin Ther*. 2007;29(9):1825.

Citrome L, Goldberg JF, Portland KB. Placing transdermal selegiline for major depressive disorder into clinical context: Number needed to treat, number needed to harm, and likelihood to be helped or harmed. *J Affect Disord*. 2013;151(2):409.

Cornelius JR, Soloff PH, Perel JM, Ulrich RF. Continuation pharmacotherapy of borderline personality disorder with haloperidol and phenelzine. *Am J Psychiatry*. 1993;150(12):1843.

Cowdry RW, Gardner DL. Pharmacotherapy of borderline personality disorder. Alprazolam, carbamazepine, trifluoperazine, and tranylcypromine. *Arch Gen Psychiatry*. 1988;45(2):111.

Davidson JR, McLeod MN, Turnbull CD, Miller RD. A comparison of phenelzine and imipramine in depressed inpatients. *J Clin Psychiatry*. 1981;42(10):395.

Dilsaver SC, Jaeckle RS. Winter depression responds to an open trial of tranylcypromine. *J Clin Psychiatry*. 1990;51(8):326.

Ernst M, Liebenauer LL, Jons PH, Tebeka D, Cohen RM, Zametkin AJ. Selegiline in adults with attention deficit hyperactivity disorder: clinical efficacy and safety. *Psychopharmacol Bull*. 1996;32(3):327.

Feigin A, Kurlan R, McDermott MP, et al. A controlled trial of deprenyl in children with Tourette's syndrome and attention deficit hyperactivity disorder. *Neurology*. 1996;46(4):965.

Gardner DM, Shulman KI, Walker SE, Tailor SA. The making of a user friendly MAOI diet. *J Clin Psychiatry*. 1996;57(3):99.

Gelernter CS, Uhde TW, Cimbolic P, et al. Cognitive-behavioral and pharmacological treatments of social phobia. A controlled study. *Arch Gen Psychiatry*. 1991;48(10):938.

Georgotas A, McCue RE, Cooper TB. A placebo-controlled comparison of nortriptyline and phenelzine in maintenance therapy of elderly depressed patients. *Arch Gen Psychiatry*. 1989;46(9):783.

Georgotas A, McCue RE, Friedman E, Cooper TB. Response of depressive symptoms to nortriptyline, phenelzine and placebo. *Br J Psychiatry*. 1987;151:102.

Gillman PK. Advances pertaining to the pharmacology and interactions of irreversible nonselective monoamine oxidase inhibitors. *J Clin Psychopharmacol*. 2011;31(1):66.

Halaris A, Piletz JE. Relevance of imidazoline receptors and agmatine to psychiatry: a decade of progress. *Ann N Y Acad Sci*. 2003;1009:1.

Heimberg RG, Liebowitz MR, Hope DA, et al. Cognitive behavioral group therapy vs phenelzine therapy for social phobia: 12-week outcome. *Arch Gen Psychiatry*. 1998;55(12):1133.

Henkel V, Mergl R, Allgaier AK, Kohnen R, Möller HJ, Hegerl U. Treatment of depression with atypical features: a meta-analytic approach. *Psychiatry Res*. 2006;141(1):89.

Himmelhoch JM, Fuchs CZ, Symons BJ. A double-blind study of tranylcypromine treatment of major anergic depression. *J Nerv Ment Dis*. 1982;170(10):628.

Himmelhoch JM, Thase ME, Mallinger AG, Houck P. Tranylcypromine versus imipramine in anergic bipolar depression. *Am J Psychiatry*. 1991;148(7):910.

Janicak PG, Pandey GN, Davis JM, Boshes R, Bresnahan D, Sharma R. Response of psychotic and nonpsychotic depression to phenelzine. *Am J Psychiatry*. 1988;145(1):93.

Jarrett RB, Schaffer M, McIntire D, Witt-Browder A, Kraft D, Risser RC. Treatment of atypical depression with cognitive therapy or phenelzine: a double-blind, placebo-controlled trial. *Arch*

Gen Psychiatry. 1999;56(5):431.

Jenike MA, Baer L, Minichiello WE, Rauch SL, Buttolph ML. Placebo-controlled trial of fluoxetine and phenelzine for obsessive-compulsive disorder. *Am J Psychiatry*. 1997;154(9):1261.

Kennedy S, Piran N, Garfinkel PE. Isocarboxazid in the treatment of bulimia. *Am J Psychiatry*. 1986;143(11):1495.

Kosten TR, Frank JB, Dan E, McDougle CJ, Giller EL Jr. Pharmacotherapy for posttraumatic stress disorder using phenelzine or imipramine. *J Nerv Ment Dis*. 1991;179(6):366.

Krishnan KR. Revisiting monoamine oxidase inhibitors. *J Clin Psychiatry*. 2007;68(Suppl 8):35.

Krishnan V, Nestler EJ. The molecular neurobiology of depression. *Nature*. 2008;455(7215):894.

Liebowitz MR, Heimberg RG, Schneier FR, et al. Cognitive-behavioral group therapy versus phenelzine in social phobia: long-term outcome. *Depress Anxiety*. 1999;10(3):89.

Liebowitz MR, Schneier F, Campeas R, et al. Phenelzine vs atenolol in social phobia. A placebo-controlled comparison. *Arch Gen Psychiatry*. 1992;49(4):290.

*López-Muñoz F, Alamo C, Juckel G, Assion HJ. Half a century of antidepressant drugs: on the clinical introduction of monoamine oxidase inhibitors, tricyclics, and tetracyclics. Part I: monoamine oxidase inhibitors. *J Clin Psychopharmacol*. 2007;27(6):555.

Mallinger AG, Frank E, Thase ME, et al. Revisiting the effectiveness of standard antidepressants in bipolar disorder: are monoamine oxidase inhibitors superior? *Psychopharmacol Bull*. 2009;42(2):64.

Mann JJ, Aarons SF, Wilner PJ, et al. A controlled study of the antidepressant efficacy and side effects of (I)-Deprenyl. *Arch Gen Psychiatry*. 1989;46(1):45.

McGrath PJ, Stewart JW, Fava M, et al. Tranylcypromine versus venlafaxine plus mirtazapine following three failed antidepressant medication trials for depression: a STAR*D report. *Am J Psychiatry*. 2006;163(9):1531.

Meyer JH, Ginovart N, Boovariwala A, et al. Elevated monoamine oxidase a levels in the brain: an explanation for the monoamine imbalance of major depression. *Arch Gen Psychiatry*. 2006;63(11):1209.

Meyer JH, Wilson AA, Sagrati S, et al. Brain monoamine oxidase A binding in major depressive disorder: relationship to selective serotonin reuptake inhibitor treatment, recovery, and recurrence. *Arch Gen Psychiatry*. 2009;66(12):1304.

Middlefell R, Frost I, Egan GP, Eeaton H. A report on the effects of phenelzine (nardil), a monoamine oxidase inhibitor, in depressed patients. *J Ment Sci*. 1960;106:1533–1538.

Mirowska-Guzel D, Balkowiec-Iskra E. The role of monoamine oxidase in humans and its metabolism. *Psych Ann*. 2014;44(11):495.

Mohammadi MR, Ghanizadeh A, Alaghband-Rad J, Tehranidoost M, Mesgarpour B, Soori H. Selegiline in comparison with methylphenidate in attention deficit hyperactivity disorder children and adolescents in a double-blind, randomized clinical trial. *J Child Adolesc Psychopharmacol*. 2004;14(3):418.

Nakash-Eisikovits O, Dierberger A, Westen D. A multidimensional meta-analysis of pharmacotherapy for bulimia nervosa: summarizing the range of outcomes in controlled clinical trials. *Harv Rev Psychiatry*. 2002;10(4):193.

Nolen WA, Haffmans PM, Bouvy PF, Duivenvoorden HJ. Monoamine oxidase inhibitors in

resistant major depression. A double-blind comparison of brofaromine and tranylcypromine in patients resistant to tricyclic antidepressants. *J Affect Disord*. 1993;28(3):189.

Nolen WA, Kupka RW, Hellemann G, et al. Tranylcypromine vs. lamotrigine in the treatment of refractory bipolar depression: a failed but clinically useful study. *Acta Psychiatr Scand*. 2007;115(5):360.

Nolen WA, van de Putte JJ, Dijken WA, et al. Treatment strategy in depression. II. MAO inhibitors in depression resistant to cyclic antidepressants: two controlled crossover studies with tranylcypromine versus L-5-hydroxytryptophan and nomifensine. *Acta Psychiatr Scand*. 1988;78(6):676.

Pande AC, Birkett M, Fechner-Bates S, Haskett RF, Greden JF. Fluoxetine versus phenelzine in atypical depression. *Biol Psychiatry*. 1996;40(10):1017.

Parsons B, Quitkin FM, McGrath PJ, et al. Phenelzine, imipramine, and placebo in borderline patients meeting criteria for atypical depression. *Psychopharmacol Bull*. 1989;25(4):524.

Pohl R, Berchou R, Rainey JM Jr. Tricyclic antidepressants and monoamine oxidase inhibitors in the treatment of agoraphobia. *J Clin Psychopharmacol*. 1982;2(6):399.

Quitkin FM, McGrath PJ, Stewart JW, et al. Atypical depression, panic attacks, and response to imipramine and phenelzine. A replication. *Arch Gen Psychiatry*. 1990;47(10):935.

*Quitkin FM, Stewart JW, McGrath PJ, et al. Columbia atypical depression. A subgroup of depressives with better response to MAOI than to tricyclic antidepressants or placebo. *Br J Psychiatry Suppl*. 1993;(21):30.

*Ramsay RR. Monoamine oxidases: the biochemistry of the proteins as targets in medicinal chemistry and drug discovery. *Curr Top Med Chem*. 2012;12(20):2189.

Rapaport MH. Dietary restrictions and drug interactions with monoamine oxidase inhibitors: the state of the art. *J Clin Psychiatry*. 2007;68(Suppl 8):42.

Rees L, Davies B. A controlled trial of phenelzine ("Nardil") in the treatment of severe depressive illness. *J Ment Sci*. 1961;107:560.

Robinson DS, Lerfald SC, Bennett B, et al. Continuation and maintenance treatment of major depression with the monoamine oxidase inhibitor phenelzine: a double-blind placebo-controlled discontinuation study. *Psychopharmacol Bull*. 1991;27(1):31.

Rothschild R, Quitkin HM, Quitkin FM, et al. A double-blind placebo-controlled comparison of phenelzine and imipramine in the treatment of bulimia in atypical depressives. *Int J Eat Disord*. 1994;15(1):1.

Rubinstein S, Malone MA, Roberts W, Logan WJ. Placebo-controlled study examining effects of selegiline in children with attention-deficit/hyperactivity disorder. *J Child Adolesc Psychopharmacol*. 2006;16(4):404.

Shestatzky M, Greenberg D, Lerer B. A controlled trial of phenelzine in posttraumatic stress disorder. *Psychiatry Res*. 1988;24(2):149.

Shulman KI, Fischer HD, Herrmann N, Huo CY, Anderson GM, Rochon PA. Current prescription patterns and safety profile of irreversible monoamine oxidase inhibitors: a population-based cohort study of older adults. *J Clin Psychiatry*. 2009;70(12):1681.

Shulman KI, Herrmann N, Walker SE. Current place of monoamine oxidase inhibitors in the treatment of depression. *CNS Drugs*. 2013;27(10):789.

Soloff PH, Cornelius J, George A, Nathan S, Perel JM, Ulrich RF. Efficacy of phenelzine and

haloperidol in borderline personality disorder. *Arch Gen Psychiatry*. 1993;50(5):377.

Solyom C, Solyom L, LaPierre Y, Pecknold J, Morton L. Phenelzine and exposure in the treatment of phobias. *Biol Psychiatry*. 1981;16(3):239.

Stewart JW, McGrath PJ, Quitkin FM, et al. Chronic depression: response to placebo, imipramine, and phenelzine. *J Clin Psychopharmacol*. 1993;13(6):391.

Sunderland T, Cohen RM, Molchan S, et al. High-dose selegiline in treatment-resistant older depressive patients. *Arch Gen Psychiatry*. 1994;51(8):607.

Thase ME, Mallinger AG, McKnight D, Himmelhoch JM. Treatment of imipramine-resistant recurrent depression, IV: A double-blind crossover study of tranylcypromine for anergic bipolar depression. *Am J Psychiatry*. 1992;149(2):195.

*Thase ME, Trivedi MH, Rush AJ. MAOIs in the contemporary treatment of depression. *Neuropsychopharmacology*. 2015;12(3):185.

Vallejo J, Olivares J, Marcos T, Bulbena A, Menchón JM. Clomipramine versus phenelzine in obsessive-compulsive disorder. A controlled clinical trial. *Br J Psychiatry*. 1992;161:665.

Versiani M, Nardi AE, Mundim FD, Alves AB, Liebowitz MR, Amrein R. Pharmacotherapy of social phobia. A controlled study with moclobemide and phenelzine. *Br J Psychiatry*. 1992;161:353–360.

Volz HP, Faltus F, Magyar I, Möller HJ. Brofaromine in treatment-resistant depressed patients–a comparative trial versus tranylcypromine. *J Affect Disord*. 1994;30(3):209.

Walsh BT, Gladis M, Roose SP, Stewart JW, Stetner F, Glassman AH. Phenelzine vs placebo in 50 patients with bulimia. *Arch Gen Psychiatry*. 1988;45(5):471.

Youdim MB. Multi target neuroprotective and neurorestorative anti-Parkinson and anti-Alzheimer drugs ladostigil and m30 derived from rasagiline. *Exp Neurobiol*. 2013;22(1):1.

Youdim MB, Bakhle YS. Monoamine oxidase: isoforms and inhibitors in Parkinson's disease and depressive illness. *Br J Pharmacol*. 2006;147(Suppl 1):S287.

Youdim MB, Edmondson D, Tipton KF. The therapeutic potential of monoamine oxidase inhibitors. *Nat Rev Neurosci*. 2006;7(4):295.

Zametkin A, Rapoport JL, Murphy DL, Linnoila M, Ismond D. Treatment of hyperactive children with monoamine oxidase inhibitors. I. Clinical efficacy. *Arch Gen Psychiatry*. 1985;42(10):962.

▲ 34.23 Nefazodone

AMIR A. KHAN, M.D., AND SUSAN G. KORNSTEIN, M.D.

Nefazodone (Serzone) was developed to improve the side-effect profile of the analogous, second-generation antidepressant trazodone (Desyrel) by using receptor-binding techniques. It has unique effects on the serotonergic system that are distinct from other antidepressants. It became available in the United States in early 1995. Following the issuance of a black box warning regarding rare but potentially life-threatening hepatotoxicity and the availability of generic nefazodone, the original manufacturer of Serzone (Bristol–Myers Squibb) discontinued the production of Serzone in June 2004, citing declining market share as a primary reason. Generic

nefazodone remains available in the US market though its clinical utilization continues to decline. Sales of nefazodone in Canada were discontinued in 2003.

Nefazodone is a phenylpiperazine compound, with some structural similarity to trazodone. It is structurally unrelated to other antidepressants. Nefazodone's mechanism of action is attributed to its antagonism of the 5-hydroxytryptamine, type 2A (5-HT_{2A}) receptor and inhibition of the serotonin transporter. It has limited affinity for the α_1-adrenergic receptor, hence the decreased incidence of postural hypotension compared with trazodone.

CHEMISTRY

Nefazodone is a synthetically derived phenylpiperazine antidepressant. The chemical name for nefazodone hydrochloride is 2-[3-[4-(3-cholorophenyl)-1-piperazinyl]propyl]-5-ethyl-2,4-dihydro-4-(2-phenoxyethyl)-3H-1,2,4-trizol-3-one monohydrochloride. The molecular formula is $C_{25}H_{32}CIN_5O_2$•HCl with a molecular weight of 506.5 (Fig. 34.23–1).

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

Absorption. Nefazodone is rapidly and completely absorbed after oral intake. The absolute bioavailability, however, is low at about 20 percent and variable because of its extensive metabolism. Presence of food in the GI tract delays the absorption and further reduces the bioavailability. The clinical significance of this finding is unclear and probably not important. Peak plasma concentration occurs in about 1 hour.

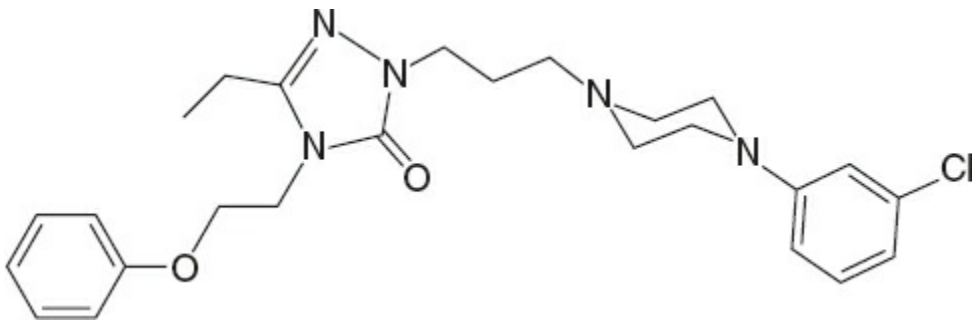

FIGURE 34.23–1. Molecular structure of nefazodone (Serzone).

Distribution. Nefazodone is widely distributed in the body tissues, including the CNS. The volume of distribution for nefazodone is 0.22 to 0.87 L/kg. It readily crosses the blood–brain barrier. Nefazodone is extensively (>99 percent) but loosely bound to plasma proteins. It does not tend to displace other protein-bound drugs when studied in vitro, but this

property in vivo is not well studied. It does decrease the protein binding of haloperidol by 5 percent, which is unlikely to be of any clinical importance.

Metabolism and Elimination. Nefazodone is extensively metabolized by n-dealkylation and aliphatic and aromatic hydroxylation. Less than 1 percent of the drug is excreted unchanged in urine. Nefazodone's three major pharmacologically active metabolites are hydroxynefazodone, meta-chlorphenylpiperazine (mCPP), and triazole-dione. Both nefazodone and hydroxynefazodone have a relatively short half-life (1.5 to 4 hours) and exhibit nonlinear kinetics. Compared to a single dose, multiple dosing as well as dosage increase produce a disproportional increase in the maximum concentration (C_{max}) of the drug. This property signifies greater-than-predicted accumulation of the drug and its hydroxy metabolite with multiple dosing. Nefazodone and its metabolites attain steady-state concentration within 4 to 5 days of initiation of BID dosing or changes in previous regimens.

The pharmacological profile of hydroxynefazodone is similar to that of nefazodone. mCPP is also somewhat similar to nefazodone, with the notable difference of possessing agonist activity at some serotonergic receptor subtypes. It has a half-life of 4 to 8 hours. The actions of triazole-dione are not yet well characterized, however it has a longer half-life of 18 hours.

Nefazodone should not be used in patients with hepatic disease. Pharmacokinetic studies have shown reduced drug clearance in patients with hepatic cirrhosis; a dose reduction should be considered in such patients, as the steady-state concentrations of nefazodone and hydroxynefazodone are 25 percent higher than in normal controls. Exposure to mCPP is about two- to threefold higher in patients with hepatic cirrhosis. Patients with renal dysfunction do not show a significantly different pharmacokinetic profile. Elderly patients show a 10 to 20 percent increase in steady-state plasma levels; therefore, the starting dose for older patients, particularly women, should be half the usual dose. Nefazodone does not utilize the CYP 2D6 metabolic pathway; hence, dosage adjustment for poor metabolizers is not required.

Pharmacodynamics

Nefazodone has a multimodal, complex mechanism of action. It inhibits neuronal uptake of serotonin and norepinephrine; however, its antidepressant effects are presumably exerted only through the serotonergic system. Presynaptically, it inhibits 5-HT reuptake, which increases the 5-HT level within the synapse and results in increased availability of serotonin. Postsynaptically, it occupies 5-HT_{2A} receptors as an antagonist, hence downregulating the receptor. The latter property is

believed to be responsible for the antidepressant action of nefazodone. Its ability to block 5-HT_{2A} receptors also provides a potential mechanism for the low likelihood of treatment-related activation and sexual dysfunction.

Hydroxynefazodone has an almost identical pharmacological profile to the parent compound and probably contributes to the antidepressant action of nefazodone. mCPP is a minor metabolite, but it appears to have important pharmacological properties. It binds potentially to all 5-HT_1, 5-HT_2, and 5-HT_3 receptors. It acts as an agonist at the 5-HT_{1A}, 5-HT_{2C}, and 5-HT_3 receptors and as a mixed agonist and antagonist at the 5-HT_{2A} site. It is believed to possess anxiogenic properties. Animal studies as well as the direct administration of mCPP to patients with anxiety disorders and normal controls generally support this finding.

Nefazodone has weak α_1-adrenergic antagonism that may cause postural hypotension, though this effect is much less pronounced compared to trazodone. Nefazodone does not have significant affinity for α_2-adrenergic, β-adrenergic, cholinergic, dopaminergic, 5-HT_{1A}, or benzodiazepine receptors. This receptor affinity profile offers distinct clinical advantages with regard to minimizing potential side effects.

THERAPEUTIC INDICATIONS

Nefazodone was approved for the treatment of major depression in the United States in 1994. This approval was granted based upon the results of five well-controlled premarketing clinical trials. Since its approval, many small and large scale studies including both inpatient and outpatient populations have documented its efficacy in the treatment of depression.

Major Depression

Nefazodone was evaluated in four phase III, double-blind, active (imipramine) and placebo-controlled trials to study its efficacy and safety as an antidepressant in outpatients with major depression. Primary efficacy measures were the 17-item Hamilton Rating Scale for Depression (HAM-D) and the Clinical Global Impression (CGI) scale. Three of the four studies showed nefazodone to be superior to placebo and comparable to imipramine in efficacy. The fourth trial failed to show a significant difference between both active drugs and placebo. The incidence of side effects and premature treatment discontinuation was higher with imipramine than nefazodone. Two dose-finding studies as well as other phase II/III studies determined the optimal treatment dosage range to be between 300 and 500 mg/day.

Nefazodone has been directly compared with the SSRIs in both short-term (6 to 8 weeks) and maintenance studies (up to 46 weeks). A meta-analysis of six studies reported that nefazodone (n = 523) and SSRIs (fluoxetine, paroxetine, and sertraline; n = 513) were comparable in efficacy

for the treatment of moderate to severe major depression. Differences were noted in the side-effect profile favoring nefazodone, particularly with regard to sleep disturbances, weight gain, and sexual dysfunction. An additional meta-analysis involving nine studies comparing nefazodone/trazadone with SSRIs for the treatment of major depression (n = 998) confirmed these findings.

There is considerable evidence for the efficacy of nefazodone in the treatment of various subtypes of depression. A meta-analysis of four clinical trials comparing nefazodone (n = 247), imipramine (n = 166), and placebo (n = 251) showed that nefazodone is effective for the treatment of severe depression as well as melancholic and recurrent depression. A placebo-controlled study evaluated the efficacy of nefazodone in 120 hospitalized patients with severe depression. Significant improvement was noted in patients treated with nefazodone compared to placebo (CGI-I response rates of 50 percent vs. 29 percent) in this 6-week study.

Nefazodone has been studied in direct comparison with psychotherapy and combination treatment in patients with chronic depression. In this study, 681 outpatients with chronic, nonpsychotic major depression were assigned to 12 weeks of treatment with nefazodone, The Cognitive Behavioral Analysis System of Psychotherapy, or their combination. In patients who completed the study (n = 519), the response rates were 55 percent in the nefazodone group, 52 percent in the psychotherapy group, and 85 percent in the combined treatment group. The response rates for the monotherapy groups are comparable to other studies of chronic depression. The level of response for combination treatment is the highest ever reported in this patient population. A subset of these patients who did not respond to the initial trial of either psychotherapy or nefazodone treatment showed improved response when treatments were switched (from nefazodone to psychotherapy and vice versa).

Other Mood Disorders

Nefazodone has not been systematically evaluated for the treatment of bipolar depression. One small open-label trial studied nefazodone as an add-on treatment for bipolar depression in 13 patients and showed positive results. The reason to use nefazodone, despite its well-known hepatotoxicity, was to study the underlying mechanism of postsynaptic 5-HT(2A) antagonism as a potential mechanism to develop future therapies for bipolar depression. There are several case reports of nefazodone-induced mania. Caution should be exercised in using nefazodone for the treatment of bipolar depression.

Two small open-label studies reported the efficacy of nefazodone in the treatment of premenstrual dysphoric disorder (PMDD). One small randomized, controlled trial compared its efficacy in PMDD versus buspirone and placebo. This study did not find significant efficacy for

nefazodone in the treatment of PMDD, except that it was modestly effective in controlling PMDD-associated irritability; however, the small sample size renders these findings inconclusive.

Symptoms of anxiety during depressive episodes are common. There is evidence from controlled trials that nefazodone helps the comorbid symptoms of anxiety in depressed patients. A meta-analysis of six randomized, controlled trials of nefazodone in patients with major depression showed that nefazodone produced earlier and more sustained improvement in the symptoms of agitation and somatic anxiety compared to imipramine and placebo. This finding was replicated by another retrospective analysis of over 300 patients with major depression. In this analysis, imipramine and nefazodone were both effective in alleviating depressive symptoms; however, nefazodone showed superior response for comorbid anxiety symptoms. Additional evidence from the nefazodone chronic depression study also supports this beneficial effect of nefazodone on comorbid anxiety, particularly early (first or second week) in the course of treatment.

Anxiety Disorders

Although a number of studies have looked at anxiety symptoms during an episode of major depression, there has been little study of nefazodone for the treatment of primary anxiety disorders. Most of the evidence available is from small, open-label trials.

A meta-analysis of six open-label trials consisting of 105 patients with PTSD (both civilians and combat veterans) showed that nefazodone was effective for a broad spectrum of PTSD symptoms at a dosage of up to 600 mg/day. This finding has been replicated in four additional, small, open-label trials. Randomized controlled trials are needed to determine if nefazodone is superior to placebo and comparable to other treatments for PTSD. An interesting finding in one of the open-label trials was that although patients with chronic PTSD reported improvement in the symptoms of nightmares and other sleep disturbance as well as depression and other PTSD symptoms, the objective, polysomnographic sleep measures did not change, which is in contrast to the sleep studies of depressed patients on nefazodone.

There are two open-label studies of nefazodone use in panic disorder and an additional study in panic disorder comorbid with depression. The total number of patients enrolled in these studies was 39, with study durations of 8 to 12 weeks. Two of the three studies found nefazodone to be beneficial for the treatment of panic disorder, with or without agoraphobia and with or without depression. However, one study reported that the response rate of nefazodone was lower compared to some other anti-panic medications such as paroxetine or low-dose venlafaxine.

A clinical case series of five patients suggested the potential efficacy of

nefazodone for the treatment of social phobia. An open-label study consisting of 23 patients reported a response rate of about 70 percent (n = 16) for social phobia. However, a recent double-blind placebo-controlled study for the treatment of social phobia (n = 102) reported no significant difference in response rate between nefazodone and placebo-treated patients. Another open-label trial of nefazodone was conducted in 21 patients with generalized anxiety disorder (GAD). Of 15 subjects who completed the trial, 12 were considered significantly improved. An open-label study of nefazodone in patients with depression and comorbid OCD did not report significant antiobsessional effects. Another retrospective study evaluated the efficacy of nefazodone for the treatment of nonparaphilic compulsive sexual behavior. Of the 14 subjects studied, 6 reported good control of sexual obsessions and compulsions and 5 reported remission of those symptoms.

In summary, nefazodone can potentially be useful for a variety of anxiety disorders; however, the existing evidence is not rigorous. If a patient fails the standard treatment, then nefazodone can be used as a second- or third-line treatment, with the understanding that the data regarding its efficacy are very limited.

Substance Abuse and Dual Diagnosis

There is one double-blind, placebo-controlled study of nefazodone in patients with comorbid major depression and alcohol dependence. This study enrolled 64 subjects who were actively drinking alcohol and had a history of at least one episode of depression when abstinent. Patients were randomized to receive 12 weeks of treatment with nefazodone or placebo. All patients continued once-weekly psychoeducational group therapy. The results indicated that nefazodone is efficacious for the treatment of depression in this population, but it did not show any advantage over placebo with regard to alcohol outcome. There was also an increased incidence of nonserious, drug-related side effects (sedation, visual trails, dizziness, and dry mouth). Another, larger study of 183 patients with alcohol dependence compared nefazodone with naltrexone and placebo. This study did not find any beneficial effects of nefazodone for the treatment of alcohol dependence. There is one case series reporting successful treatment with nefazodone for depression in four patients who were on methadone maintenance treatment and whose depression had not responded to other medications.

Two double-blind, placebo-controlled studies investigated the use of nefazodone in cocaine-dependent patients. The first study included 69 subjects with cocaine dependence and depressive symptoms, and the primary objective was to determine the efficacy of nefazodone in reducing cocaine use. The second study included 210 subjects. Both trials indicated no specific efficacy of nefazodone in treating cocaine dependence.

A recent double-blind study investigated the efficacy of nefazodone and bupropion-sustained release for treating cannabis dependence. Sample size was relatively small. The study did not show any specific benefit of nefazodone or bupropion for the treatment of cannabis dependence in the studied population.

From these reports, it can be concluded that nefazodone does not possess specific efficacy for the treatment of alcohol or substance abuse. However, it can be used for the treatment of comorbid depression in this population. Extreme caution should be exercised in using nefazodone for patients who have alcohol use disorders because of the rare but life-threatening adverse effect of liver failure.

Pain

One open-label study evaluated the efficacy of nefazodone for the treatment of chronic daily headaches in 52 patients. Most of the patients enrolled (n = 48) had a history of migraine, based upon the criteria established by International Headache Society. The remainder of the patients had chronic tension-type headaches. The median dose used was 300 mg/day. At the end of the 12th week, 71 percent of the patients showed at least a 50 percent improvement in their headache index. Patients who were depressed or anxious showed significantly greater improvement in their headache as well as in the symptoms of depression and anxiety. There is another small, open-label study of 10 patients with diabetes mellitus type II with painful diabetic neuropathy but who did not meet criteria for major depression. These patients received up to 450 mg/day of nefazodone for 8 weeks. At the end of the treatment period, the patients showed improvement in pain, parasthesias and numbness as well as slightly improved diabetic control. Both of these studies indicate a potential analgesic effect of nefazodone. However, whether this effect is independent of antidepressant action remains unclear.

Other Uses

There are case reports and small, open-label studies of nefazodone use for the treatment of mood disorders in children and adolescents. The results show that nefazodone is effective for the symptoms of depression in this population and is well tolerated. However, these studies also indicate that the pharmacokinetic properties of nefazodone in children and adolescents are highly variable.

There are also small open-label trials of nefazodone use as an adjunct treatment for schizophrenia, in patients who were on a stable dose of antipsychotic medication. There is also one study of nefazodone use in somatization disorder. These studies not only suffer from the limitations common to open-label trials, but they also do not have adequate power to

establish the efficacy of nefazodone independent of the comorbid depression present.

A recently published analysis of nefazodone and cognitive behavioral analysis system of psychotherapy trial showed an interesting finding supporting patient preference as a potent moderator of treatment response for patients with chronic forms of MDD (i.e., patients who preferred psychotherapy responded better to psychotherapy and patients who preferred medication treatment responded better to that intervention).

From the above discussion, it is reasonable to conclude that nefazodone is comparable to other treatments for major depression. However, there is clearly a need for randomized, controlled trials before nefazodone can be recommended for routine use for conditions other than depression.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

General

Premarketing clinical trials with nefazodone in almost 3,500 patients showed that it has a relatively low incidence of adverse events. In these trials, about 16 percent of participating patients discontinued treatment due to adverse experiences. The commonly observed adverse events that differed significantly from placebo were somnolence (25 percent vs. 14 percent), dry mouth (25 percent vs. 13 percent), nausea (22 percent vs. 12 percent), dizziness (17 percent vs. 5 percent), constipation (14 percent vs. 8 percent), asthenia (11 percent vs. 5 percent), light headedness (10 percent vs. 3 percent), blurred vision (9 percent vs. 3 percent), confusion (7 percent vs. 2 percent), and abnormal vision, consisting of scotoma and visual trails (7 percent vs. 1 percent). The most common reasons for treatment discontinuation were nausea, dizziness, asthenia, and agitation. The incidence of adverse events increases with the dosage of nefazodone. An analysis of three placebo-controlled trials in which two different dosage ranges were used (low ≤300 mg/day; high 300 to 600 mg/day) showed that the incidence of somnolence, nausea, dizziness, constipation, abnormal vision, confusion, and tinnitus was comparable to placebo in the low dosage range. However, the incidence of these adverse events increased significantly in patients treated at the higher dosage range (300 to 600 mg/day).

In 2004, FDA mandated that a boxed warning be placed on the labels of all antidepressants indicating an increased risk of suicidal thoughts and behavior in youth taking these medications and in 2006, added young adults age 24 years or less to this list. There are no data indicating a specific increase in this risk by nefazodone; however, caution should be exercised, particularly in children.

Central Nervous System

The most common adverse events pertaining to the CNS are somnolence, dizziness, confusion, insomnia, and light headedness. Less frequently reported events include memory impairment, paresthesias, flushing and feeling warm, abnormal dreams, ataxia, and tremors. There is at least one report of a nefazodone-induced acute dystonic reaction, which responded well to standard anticholinergic treatment.

There are reports of rare incidences of seizures after nefazodone administration, particularly in patients who have a seizure disorder, however a causal relationship has not been established. Controlled data for combining antidepressant drug therapy and ECT are sparse. In a case report from Australia, one patient with treatment-resistant dysthymia underwent ECT while taking 600 mg/day of nefazodone and 15 mg of midazolam without significant problems. There was no adverse outcome, and the patient showed a good response.

Sleep disturbance is one of the most frequent and distressing symptoms of depression. Many of the antidepressants tend to cause or worsen sleep disturbances. Their effects include decreased REM sleep, increased REM latency, increased stage I sleep, and decreased sleep efficiency. Nefazodone, on the other hand, appears to possess unique and potentially beneficial effects on sleep architecture. It shows minimal effects on REM sleep, decreases stage I sleep, and improves sleep efficiency. These effects have been observed in both normal as well as depressed subjects. In comparison with fluoxetine, nefazodone produced significantly greater improvement in sleep on both subjective and objective (sleep EEG) measures. The beneficial effects of nefazodone on sleep tend to appear early in treatment, during the first 2 weeks. The differences between nefazodone and SSRIs on sleep parameters have been shown to remain significant even after patients have responded to the antidepressant treatment. However, a recent study found no difference in subjective sleep improvement between nefazodone and paroxetine after 8 weeks of treatment.

Preclinical studies on human volunteers were carried out to determine the effects of nefazodone on daytime sleepiness, highway driving, and cognitive performance. Results of these early studies showed minimal impact of nefazodone on these measures. Clinical experience in patient populations shows that somnolence is frequently observed, particularly early in the treatment, however it does not appear to result in an actual increase in daytime sleepiness. The incidence of adverse events related to motor performance and cognitive tasks is also minimal but has a dose–response relationship and tends to increase for doses over 300 mg/day.

Cardiovascular System

Nefazodone has limited effects on the cardiovascular system. It can cause a slightly increased incidence of postural hypotension and hypotension. These effects are much less pronounced compared to the TCAs or

trazodone. In a pooled analysis of almost 800 patients, the incidence of postural hypotension with nefazodone was 4 percent, compared to 1 percent with placebo. In the same study, the incidence of hypotension was 2 percent.

There were no clinically significant ECG changes noted in placebo-controlled trials involving 1,153 patients taking nefazodone. Asymptomatic sinus bradycardia was noted in 1.3 percent of nefazodone-treated patients, compared to 0.4 percent of placebo-treated patients. The specific cardiac effects of nefazodone in patients with existing cardiovascular disease, particularly of recent onset, are not fully known. There is a potential for QTc prolongation if nefazodone is combined with terfenadine, astemizole, cisapride, or pimozide.

Genitourinary System

Antidepressant medications are often associated with treatment-related sexual dysfunction, including effects on desire, arousal, and orgasmic function. These side effects are particularly common with the SSRIs. In clinical studies, nefazodone has been associated with a low incidence of sexual side effects in both men and women. In one study comparing nefazodone and sertraline in depressed outpatients, 47 percent women on sertraline reported difficulty achieving orgasm compared to 17 percent of women on nefazodone. In the same study, significantly more men taking sertraline (67 percent) reported ejaculatory difficulty than did men taking nefazodone (19 percent). The incidence of sexual side effects reported with nefazodone was also low (3.5 percent) in a recent study of patients with chronic depression.

Priapism, a recognized side effect of trazodone, had not been reported with nefazodone until recently. There is one case report of nefazodone-precipitated clitoral priapism. The paucity of other case reports indicates that priapism is a possible but very rare adverse event associated with nefazodone use. It is believed that the relatively low affinity of nefazodone for α_1-adrenergic receptors is the reason for the low likelihood of treatment-related priapism or postural hypotension.

In premarketing clinical trials, the incidence of urinary frequency, retention, and infection was slightly increased with nefazodone compared to placebo. However, the overall incidence of these adverse events was low.

Gastrointestinal System

Rare cases of life-threatening hepatic failure have been reported in patients treated with nefazodone. The reported rate in the United States is about one case of liver failure resulting in death or transplant per 250,000 to 300,000 patient-years of nefazodone treatment. This rate is considered an underestimate by some because of potential under-reporting. This finding

prompted a mandatory black box warning by the FDA and a revision of the product labeling and patient information packet in 2002. This warning is based upon the postmarketing experience of over seven million patients in the United States. The reported rate represents about three to four times the estimated background rate of liver failure. Most liver failures were reported between 1 and 8 months after drug initiation. The common clinical symptoms of hepatotoxicity included jaundice and increases in ALT (10x), AST (10x), totalbilirubin (mostly conjugated), and prothrombin time. Dark urine and clay-colored stool were also reported. Available histological evaluations of livers demonstrated centrilobular necrosis and bile-duct proliferation with cholestasis. Inhibition of bile acid transport by nefazodone has been proposed as an indicator of potential hepatotoxicity in animal models. The radiochemical in vitro methods are available to detect and quantify covalent binding (CB) to liver proteins for a drug or its metabolites and are being studied to predict the risk of liver injury. The potential clinical application of these techniques remains to be seen. The risk factors for drug-induced hepatic injury include age and polypharmacy. Given the seriousness of hepatotoxicity, it is clinically important to obtain a complete pretreatment history of any underlying liver disease. Periodic serum testing of liver enzymes has not been shown to prevent serious liver injury, however, early detection of drug-induced hepatic injury and immediate withdrawal of the suspected drug is likely to improve the potential for recovery. Nefazodone should be discontinued in patients showing signs of hepatic toxicity. These patients should not be rechallenged with nefazodone.

More common GI problems associated with nefazodone use include dry mouth (25 percent), nausea (22 percent), and constipation (14 percent).

Other Systems

Nefazodone is generally considered to be weight-neutral. Evidence from both short-term and long-term studies comparing nefazodone with either placebo or other antidepressants supports this conclusion. A pooled analysis of six clinical trials comparing nefazodone (n = 523) and SSRIs (fluoxetine, sertraline, or paroxetine n = 523) as well as three trials comparing nefazodone (n = 224) with imipramine (n = 224) evaluated the impact of these antidepressants on weight. The criterion for clinically significant weight change was a 7 percent increase or decrease in weight from the baseline measurement. The results showed that in short-term trials, SSRI-treated patients showed more weight loss than nefazodone-treated patients; however, during long-term treatment, more patients gained weight on SSRIs compared to nefazodone. Imipramine was also found to be associated with more weight gain than nefazodone.

Abnormal vision, usually consisting of scotoma and visual trails (palinopsia), has been associated with nefazodone use. This adverse event

was noted in up to 10 percent of the patients taking nefazodone, compared to 2 to 3 percent of patients on placebo. It is generally considered benign and is probably dose-dependent. It can usually be managed by lowering the medication dose or changing the time of the day the dose is given.

In a recent, retrospective study, joint disorders were considerably more frequently reported adverse effects of 5-HT(2A)-blocking antidepressants, including nefazodone, than of other comparable drugs, suggesting a possible association between the use of 5-HT(2A)-blocking antidepressants and joint disorders.

PRECAUTIONS AND ADVERSE EFFECTS

Overdose

The limited information available regarding the effects of nefazodone overdose during premarketing trials consisted of seven reports of overdose ranging from 1,000 to 11,200 mg, alone or in combination with other drugs. Since its approval, one study has analyzed data on nefazodone from the American Association of Poison Control Centers. It contained over 1,300 human exposures to nefazodone overdose, after excluding patients with concomitant agent overdose. Commonly reported symptoms of overdose included drowsiness (17.3 percent), nausea (9.7 percent), and dizziness (9.5 percent). The most common serious clinical effect was hypotension (1.6 percent). None of these patients died. Most of them were treated with GI decontamination. The study concluded that nefazodone appears to be of low toxicity during poisonings.

There are rare reports of fatalities, predominantly in patients taking an overdose of nefazodone in combination with alcohol and/or other substances. There is no specific antidote available to treat a nefazodone overdose. The mainstay of treatment consists of supportive measures, including adequate airway management and monitoring cardiac rhythm and vital signs.

Used in medically compromised patients during premarketing trials, the pharmacokinetics of nefazodone and its metabolites were studied in a small number of normal healthy volunteers, both younger and older, as well as in patients with biopsy-proven hepatic cirrhosis and patients with impaired renal function. The results of these studies indicate that elderly or cirrhotic patients may require lower doses of nefazodone to achieve and maintain therapeutic effectiveness. Patients with mild to moderate renal impairment (based upon creatinine clearance) appear to tolerate the usual dose without difficulty.

There are no systematic studies of nefazodone treatment of depression in patients with compromised hepatic or renal function. Given the potential for life-threatening hepatotoxicity, nefazodone should be avoided in patients with existing liver disease. It should be used with caution in

patients with renal impairment and in patients on hemodialysis. There is one case report of the use of nefazodone to treat depression in a 59-year-old patient on hemodialysis who had not responded to an adequate trial of fluoxetine. Treatment was initiated with nefazodone at a dose of 50 mg BID, which was increased to 100 mg after 1 week and 150 mg after 4 weeks. This patient responded well to nefazodone treatment with minimal adverse events. Pre- and postdialysis nefazodone serum concentrations in this patient suggested that drug levels may be affected by this procedure. Additional studies are needed to confirm the efficacy and safety of nefazodone for the treatment of depression in such patients.

One open-label study evaluated the efficacy of nefazodone in 15 HIV-seropositive patients with major depression. Eleven patients completed the 12-week study. The results showed that HIV-seropositive outpatients respond well to nefazodone and comparably to other outpatient populations. Potential drug interactions with protease inhibitors should be considered before initiating nefazodone therapy in this patient population.

There are anecdotal reports of nefazodone use for depression in patients with cancer on chemotherapy and in patients with epilepsy and with cardiac problems, however systematic studies are not available. Aside from taking routine precautions, increased potential for drug–drug interactions as well as the need for lower dosage and slow titration should be important considerations in instituting nefazodone therapy in medically ill patients.

Use during Pregnancy and Lactation

Nefazodone is classified as a pregnancy category C medication. It has shown to increase early pup mortality in rabbits at a dose approximately five times the human dose. A decrease in pup weight was also seen at this dose as well as at lower doses. There are no systematic human studies evaluating nefazodone during pregnancy. Nefazodone should be used in pregnancy only if the potential benefit justifies the potential risk to the fetus.

Very little information is available regarding the excretion of nefazodone in human breast milk. One small study of two lactating mothers reported that nefazodone is excreted in human milk; no metabolites were found. The levels of nefazodone in breast milk were not related to the maternal medication dose. Drug levels in the infants were not assayed; however, there was no indication of developmental difficulty in those infants. A case report of a premature, breast-fed infant described drowsiness, lethargy, poor feeding, and difficulty maintaining body temperature attributed to maternal use of nefazodone. Until more information becomes available, a careful risk–benefit analysis must be performed before recommending the use of nefazodone in lactating mothers.

Discontinuation Syndrome

Animal studies using monkeys trained for self-administration of cocaine as a reinforcer did not show this behavior when nefazodone was substituted for cocaine, suggesting that nefazodone has no abuse liability. Its potential for tolerance, physical dependence, or withdrawal has not been studied systematically in humans.

There are case reports suggesting possible withdrawal symptoms resulting from abrupt discontinuation of nefazodone. Typical symptoms reported include dizziness, nausea, vomiting, fatigue, insomnia, and paresthesias. Some unusual symptoms such as brief moments of disorientation and shock-like or electrical sensations have also been reported. Patients generally recover within a few days without any serious problems and the symptoms can usually be prevented by tapering the medication. The incidence of discontinuation symptoms with nefazodone appears to be much less common than with the SSRIs and venlafaxine. In a relapse prevention study of 259 patients who had been maintained on nefazodone treatment for 16 weeks and were randomly assigned to continuation treatment with nefazodone or placebo, there were no significant differences in new-onset or worsened adverse experiences between the nefazodone and placebo groups at 14 and 28 days after randomization.

DRUG INTERACTIONS

There are various mechanisms involved in the drug–drug interactions of nefazodone. Most of these interactions involve Cytochrome P450 (CYP) isoenzymes. Another potential mechanism involves drugs that are highly bound to plasma proteins, such as warfarin. Different metabolites of nefazodone also play a role in drug interactions; for example, the plasma concentration of mCPP is increased in individuals who have decreased activity of CYP2D6.

Psychotropic Agents

Nefazodone and its metabolites, hydroxynefazodone, mCPP, and trizoledione, inhibit CYP3A4, both in vitro and in vivo. This inhibition can result in increased bioavailability of medications metabolized by this pathway. Therefore, caution should be exercised when administering nefazodone with other compounds metabolized by CYP3A4. This interaction becomes even more important for drugs that are exclusively or predominantly metabolized by the CYP3A4 pathway. Two such benzodiazepines, alprazolam and triazolam, show a substantial and clinically important increase in blood levels when administered with nefazodone. When triazolam was coadministered with nefazodone, triazolam half-life increased by threefold, AUC increased by fourfold, and

peak concentration increased by 1.7-fold. Likewise, there was a 1.6-fold increase in peak concentration and a twofold increase in the AUC of alprazolam. Nefazodone plasma concentration was unaffected by either medication. Given these interactions, the manufacturer recommends that triazolam be avoided during nefazodone treatment in most patients. If alprazolam is used with nefazodone, a 50 percent reduction in the initial dose of alprazolam is recommended. This effect has not been observed with lorazepam.

One of the metabolites of nefazodone, mCPP, is present in relatively small quantities and is metabolized by the CYP2D6 pathway. When a patient is switched from an antidepressant that is an inhibitor of CYP2D6 (e.g., fluoxetine) to nefazodone, the metabolism of mCPP may be inhibited for some time. This can result in an increased level of mCPP, which is an anxiogenic compound, and may lead to increased anxiety and agitation. This problem can be minimized by using an adequate washout period before switching antidepressants. The interaction between nefazodone and MAOIs has not been systematically studied in humans or animals. The existing evidence with SSRIs or TCAs when combined with MAOIs suggests a potential for a serious, sometimes fatal, reaction. This serotonin syndrome consists of hyperthermia, rigidity, myoclonus, autonomic instability with rapid fluctuation of vital signs, and mental status changes that may include extreme agitation progressing to delirium and coma. This reaction has also been reported in patients who have recently discontinued an SSRI or TCA and have been started on MAOI. Some cases present with symptoms resembling neuroleptic malignant syndrome, including severe hyperthermia and seizures. Given the existing evidence with other antidepressants, it is recommended that nefazodone not be used in combination with an MAOI or within 14 days of discontinuing treatment with an MAOI. In addition, a minimum of 1 week should be allowed after discontinuing nefazodone before starting an MAOI.

Coadministration of lithium and nefazodone resulted in only a slight decrease in the concentration of two metabolites of nefazodone, mCPP and triazoledione, which was not clinically significant. The combination of nefazodone and carbamazepine, although well tolerated in small number of healthy volunteers, resulted in a very significant decrease in the steady-state C_{max} and AUC of nefazodone (86 and 93 percent, respectively), along with a similar but modest reduction in its metabolites. At the same time, the steady state C_{max} and AUC of carbamazepine were each increased by 23 percent, while the steady-state C_{max} and AUC of its metabolite, 10,11-epoxycarbamazepine, decreased by 21 and 20 percent, respectively. Because the decrease in nefazodone and its metabolites is likely to result in a loss of antidepressant efficacy, this combination should be avoided. There are no systematic studies to evaluate the impact of coadministration of

valproic acid and nefazodone. One case report suggested a potentially dangerous interaction of these two medications in a patient with bipolar disorder. This patient suffered from symptoms that were suggestive of serotonin syndrome. Whether this occurrence was incidental or represents a real interaction remains unclear. The impact of combining phenytoin and nefazodone has not extensively been studied. It appears that a single-dose administration of phenytoin with steady state nefazodone does not alter the pharmacokinetic properties of phenytoin. However, due to the nonlinear pharmacokinetics of phenytoin, this combination should be avoided if possible or the dose should be adjusted carefully, according to usual clinical practices.

Nefazodone tends to decrease the protein binding of haloperidol by 5 percent, however this interaction is unlikely be of any clinical significance. Human experience with clozapine and nefazodone in a small number of patients suggests that clozapine levels are increased by 4 percent and norclozapine levels by 16 percent. This relatively modest increase in levels is considered to be clinically insignificant. There are other small studies of concomitant antipsychotic and nefazodone use that have shown no significant difficulties. However, an important and potentially dangerous interaction may occur between nefazodone and pimozide resulting in a prolonged QT interval, as described in detail in the following section.

Other Medications

Terfenadine, astemizole, cisapride, and pimozide are metabolized via the CYP3A4 pathway. Drugs such as ketoconazole or erythromycin inhibit this enzyme and block the metabolism of terfenadine, astemizole, cisapride, and pimozide, resulting in increased plasma concentrations of these medications. Increased plasma concentrations of these four medications are associated with QT prolongation and rare cases of cardiovascular adverse events, including death, due principally to ventricular tachycardia of the torsades de pointes type. Nefazodone is a potent inhibitor of CYP3A4. Consequently, nefazodone must not be used in combination with terfenadine, astemizole, cisapride or pimozide. At the present time, terfenadine, astemizole and cisapride have been removed from the US market by the FDA.

Nefazodone was coadministered with digoxin to healthy volunteers who were genotyped as extensive CYP2D6 metabolizers. This combination resulted in an increased C_{max}, C_{min}, and AUC of digoxin by 29, 27, and 15 percent, respectively. There was no effect on the pharmacokinetics of nefazodone or its active metabolites. Since digoxin has a narrow therapeutic index, caution should be exercised in combining it with nefazodone. Plasma level monitoring of digoxin is also recommended. Coadministration of nefazodone and propranolol in healthy volunteers

resulted in a decrease in the level of propranolol, but did not impact the level of nefazodone or its active metabolites with the exception of mCPP, which was increased. Hence, initial dosage adjustment is not needed for either of the medications, however; subsequent dosage should be adjusted based upon the clinical response.

Simvastatin and atorvastatin are substrates of CYP3A4. When nefazodone was combined with either of these medications, it resulted in a 20-fold increase in the concentration of simvastatin and simvastatin acid and a fourfold increase in atorvastatin and atorvastatin lactone. This effect appears to result from the inhibition of CYP3A4 by nefazodone. Nefazodone did not significantly affect the level of pravastatin, another HMG-CoA reductase inhibitor. Metabolic interactions are unlikely between nefazodone and HMG-CoA reductase inhibitors such as pravastatin, fluvastatin, or rosuvastatin that undergo little or no metabolism by the CYP3A4 isoenzyme; therefore, dosage adjustment for these agents should not be necessary. However, dose adjustment is needed for simvastatin and atorvastatin because the increased concentrations of these medications might result in a greater risk of rhabdomyolysis and other adverse events associated with HMG-CoA reductase inhibitors.

There have been reports of increased blood concentrations of cyclosporin and tacrolimus into toxic ranges in patients who received concomitant nefazodone. These are immunosuppressive agents that are substrates of CYP3A4. If nefazodone is combined with either of these medications, the level should be monitored and the dose should be adjusted accordingly.

When nefazodone was administered with warfarin, there was no change in prothrombin or bleeding time despite a modest decrease in C_{max} of warfarin. This finding suggests a clinically insignificant interaction.

DOSAGE AND ADMINISTRATION

Nefazodone is indicated for the treatment of major depression. Based upon the results of clinical trials, the recommended starting dose of nefazodone is 200 mg/day in divided doses (BID). The effective dosage range is usually between 300 mg to 600 mg/day in divided doses. The starting dose can be increased by 100 to 200 mg/day at intervals of no less than 1 week. In order to minimize adverse events, it is preferable to start at 100 mg/day and then increase by 100 mg/day every week. The usual therapeutic dosage range is 400 to 600 mg/day. Doses higher than 600 mg/day are not recommended. As with all antidepressants, several weeks of treatment at a reasonable therapeutic dose are required for antidepressant response. As with other antidepressants, long-term study (up to 1 year) with nefazodone as a maintenance treatment suggests that the therapeutic dose should be continued as the maintenance dose. Once-daily dosing has been attempted in patients who were previously maintained on BID dosing as well as in

patients starting nefazodone. The results did not show significant differences with once- versus twice-daily dosing of nefazodone. Since these were open-label trials, this finding needs to be replicated in double-blind studies.

The recommended starting dose for elderly or debilitated patients is 50 mg BID. The increase in dosage should be slower and tailored to the individual patient's symptoms and side effects. The final therapeutic range in elderly patients may be similar to younger patients.

Nefazodone is supplied in 50, 100, 150, 200, and 250 mg strengths.

REFERENCES

Benson BE, Mathiason M, Dahl B, et al. Toxicities and outcomes associated with nefazodone poisoning: an analysis of 1,338 exposures. *Am J Emerg Med*. 2000;18(5):587–592.

Carpenter KM, McDowell D, Brooks DJ, et al. A preliminary trial: double-blind comparison of nefazodone, bupropion-SR, and placebo in the treatment of cannabis dependence. *Am J Addict*. 2009;18(1):53–64.

Ciraulo DA, Knapp C, Rotrosen J, et al. Nefazodone treatment of cocaine dependence with comorbid depressive symptoms. *Addiction*. 2005;100(Suppl 1):23–31.

Choi JM, Oh SJ, Lee JY, et al. Prediction of drug-induced liver injury in HepG2 cells cultured with human liver microsomes. *Chem res Toxicol*. 2015;28(5):872–885.

DeVane CL, Grothe DR, Smith SL. Pharmacology of antidepressants: focus on nefazodone. *J Clin Psychiatry*. 2002;63(1):10–17.

Dunner DL, Laird LK, Zajecka J, Bailey L, Sussman N, Seabolt JL. Six-year perspectives on the safety and tolerability of nefazodone. *J Clin Psychiatry*. 2002;63(1):32–41.

Feiger AD, Bielski RJ, Bremner J, et al. Double-blind, placebo-substitution study of nefazodone in the prevention of relapse during continuation treatment of outpatients with major depression. *Int Clin Psychopharmacol*. 1999;14(1):19–28.

Feighner J, Targum SD, Bennett ME, et al. A double-blind, placebo-controlled trial of nefazodone in the treatment of patients hospitalized for major depression. *J Clin Psychiatry*. 1998;59(5):246–253.

Fontaine R, Ontiveros A, Elie R, et al. A double-blind comparison of nefazodone, imipramine, and placebo in major depression. *J Clin Psychiatry*. 1994;55(6):234–241.

Gillin JC, Smith-Vaniz A, Schnierow B, et al. An open-label, 12-week clinical and sleep EEG study of nefazodone in chronic combat-related posttraumatic stress disorder. *J Clin Psychiatry*. 2001;62(10):789–796.

Hicks JA, Argyropoulos SV, Rich AS, et al. Randomised controlled study of sleep after nefazodone or paroxetine treatment in out-patients with depression. *Br J Psychiatry*. 2002;180:528–535.

Keller MB, McCullough JP, Klein DN, et al. A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *N Engl J Med*. 2000;342(20):1462–1470.

Kocsis JH, Leon AC, Markowitz JC, et al. Patient preference as a moderator of outcome for chronic forms of major depressive disorder treated with nefazodone, cognitive behavioral

analysis system of psychotherapy, or their combination. *J Clin Psychiatry*. 2009;70(3):354–361.

Kodesh A, Katz S, Lerner AG, Finkel B, Sigal M. Intermittent, luteal phase nefazodone treatment of premenstrual dysphoric disorder. *J Psychopharmacol*. 2001;15(1):58–60.

Kostrubsky SE, Strom SC, Kalgutkar AS, et al. Inhibition of hepatobiliary transport as a predictive method for clinical hepatotoxicity of nefazodone. *Toxicol Sci*. 2006;90(2):451–459.

Kranzler HR, Modesto-Lowe V, Van Kirk J. Naltrexone vs. nefazodone for treatment of alcohol dependence. A placebo-controlled trial. *Neuropsychopharmacology*. 2000;22(5):493–503.

Landen M, Eriksson O, Sundblad C, Andersch B, Naessen T, Eriksson E. Compounds with affinity for serotonergic receptors in the treatment of premenstrual dysphoria: a comparison of buspirone, nefazodone and placebo. *Psychopharmacology (Berl)*. 2001;155(3):292–298.

Marcus RN, Mendels J. Nefazodone in the treatment of severe, melancholic, and recurrent depression. *J Clin Psychiatry*. 1996;57(Suppl 2):19–23.

Papakostas GI, Fava M. A meta-analysis of clinical trials comparing the serotonin (5HT)-2 receptor antagonists trazodone and nefazodone with selective serotonin reuptake inhibitors for the treatment of major depressive disorder. *Eur Psychiatry*. 2007;22(7):444–447.

Passos SR, Camacho LA, Lopes CS, dos Santos MA. Nefazodone in out-patient treatment of inhaled cocaine dependence: a randomized double-blind placebo-controlled trial. *Addiction*. 2005;100(4):489–494.

Ren Z, Chen S, Zhang J, Doshi U, Li AP, Guo L. Endoplasmic reticulum stress induction and ERK1/2 activation contribute to nefazodone-induced toxicity in hepatic cells. *Toxicol Sci*. 2016;154(2):368–380.

Roy-Byrne PP, Pages KP, Russo JE, et al. Nefazodone treatment of major depression in alcohol-dependent patients: a double-blind, placebo-controlled trial. *J Clin Psychopharmacol*. 2000;20(2):129–136.

Rush AJ, Armitage R, Gillin JC, et al. Comparative effects of nefazodone and fluoxetine on sleep in outpatients with major depressive disorder. *Biol Psychiatry*. 1998;44(1):3–14.

Schatzberg AF, Prather MR, Keller MB, Rush AJ, Laird LK, Wright CW. Clinical use of nefazodone in major depression: a 6-year perspective. *J Clin Psychiatry*. 2002;63(1):18–31.

Schatzberg AF, Rush AJ, Arnow BA, et al. Chronic depression: medication (nefazodone) or psychotherapy (CBASP) is effective when the other is not. *Arch Gen Psychiatry*. 2005;62(5):513–520.

Silva AM, Barbosa IA, Seabra C, et al. Involvement of mitochondrial dysfunction in nefazodone-induced hepatotoxicity. *Food Chem Toxicol*. 2016;94:148–158.

Sussman N, Ginsberg DL, Bikoff J. Effects of nefazodone on body weight: a pooled analysis of selective serotonin reuptake inhibitor- and imipramine-controlled trials. *J Clin Psychiatry*. 2001;62(4):256–260.

Usui T, Mise M, Hashizume T, Yabuki M, Komuro S. Evaluation of the potential for drug-induced liver injury based on in vitro covalent binding to human liver proteins. *Drug Metab Dispos*. 2009;37(12):2383–2392.

Van Ameringen M, Mancini C, Oakman J, et al. Nefazodone in the treatment of generalized social phobia: a randomized, placebo-controlled trial. *J Clin Psychiatry*. 2007;68(2):288–295.

Voican CS, Corruble E, Naveau S, Perlemuter G. Antidepressant-induced liver injury: a review for

clinicians. *Am J Psychiatry*. 2014;171(4):404–415.

▲ 34.24 Opioid Maintenance Treatments: Methadone and Buprenorphine

ANDREW J. SAXON, M.D., AIMEE L. MCRAE-CLARK, PHARM.D., AND KATHLEEN T. BRADY, M.D., PH.D.

INTRODUCTION

It has long been recognized that after ongoing, regular use of opioids a use disorder including maladaptive use, functional impairment, withdrawal symptoms, and tolerance may develop. Commonly misused opioids include heroin, morphine, hydromorphone (Dilaudid), codeine, meperidine (Demerol), oxycodone, and hydrocodone (Robidone). The prevalence of opioid use and use disorder particularly in regard to prescription opioids has risen in recent years. The 2013 National Survey on Drug Use and Health reported that approximately 4.5 million Americans used prescription pain relievers for nonmedical reasons in the past month. Heroin use has also risen, with estimates of more than 4.8 million Americans having ever used heroin and 681,000 having used it in the last year. Sustained opioid-use disorder is associated with several negative outcomes, including early mortality, increased rates of hepatitis, HIV infection, sexually transmitted diseases, and other health problems, as well as criminal justice system involvement.

Physical Dependence and Withdrawal

If an individual is physiologically dependent on opioids, opioid discontinuation or diminution in the amount consumed often leads to the development of an opioid withdrawal syndrome. Physical signs and symptoms that frequently occur include sweating, piloerection, yawning, rhinorrhea, sneezing, lacrimation, pupillary dilation, nausea, vomiting, abdominal cramps, diarrhea, tremor, muscle aches or cramps, agitation, dysphoric mood, fever, and sleep disturbances. Although opioid withdrawal is not considered life threatening, it is highly physically and subjectively uncomfortable, such that an individual continues to use opioids to avoid withdrawal symptoms.

The goal of most medically supervised withdrawal (detoxification) protocols is to limit patient discomfort by decreasing or ameliorating withdrawal symptoms. Medications commonly used in the treatment of acute opioid withdrawal include the α_2-adrenergic agonist clonidine (Catapres), cyclobenzaprine (Flexeril) or benzodiazepines for muscle cramps, the antispasmodic dicyclomine (Bentyl), antidiarrheals such as loperamide, antiemetics such as prochlorperazine (Compazine) or ondansetron (Zofran), and sedatives such as zolpidem (Ambien) or

benzodiazepines. Another option is to convert the patient to an equivalent dose of another opioid, typically either methadone or buprenorphine and then to gradually reduce the dose to minimize withdrawal. If methadone is used, it must be done under the auspices of a federally licensed program, and if buprenorphine is used, it must be prescribed by a qualified, trained physician with a special Drug Enforcement Agency (DEA) number. If the goal is to expedite the withdrawal process, especially in the case of a long-acting opioid, administration of an opioid antagonist such as naloxone (Narcan) or naltrexone (Revia), hastens the onset of the withdrawal syndrome. Caution must be taken when using this approach because of the abrupt onset of symptoms that can appear in precipitated withdrawal. Again, adjunctive medications, such as clonidine, cyclobenzaprine, and dicyclomine, may also be used for the symptomatic treatment of some opioid withdrawal symptoms. The use of *ultrarapid opioid detoxification* which involves the administration of general anesthesia or heavy sedation in individuals undergoing withdrawal, although effective in reducing physiological dependence, has been demonstrated to lead to a greater number of serious adverse events than the other approaches without superior outcomes, and so it is not recommended. Once a patient has been completely withdrawn from opioids, naltrexone, an orally and parenterally active μ-opioid antagonist, may also be used to help prevent relapse. Prior to starting naltrexone, a naloxone challenge test should be performed to assure that no residual physiological dependence remains, or naltrexone could cause a prolonged, precipitated withdrawal episode. Oral naltrexone may be dosed at 50 mg/day or 100 to 150 mg two to three times a week. Positive results have been found with using naltrexone in highly motivated patients, but many opioid-dependent patients tend to drop out of naltrexone therapy. A long-acting injectable form of naltrexone (Vivitrol) that provides blockade of the μ-opioid receptor for about 30 days is also available for treatment of opioid-use disorder and may have a better efficacy profile than the oral form, but more research needs to be done to determine its efficacy.

Opioid Maintenance Therapies

Because outcomes for medically supervised opioid withdrawal are generally suboptimal, an alternative and generally preferred approach to the treatment of opioid-use disorder is opioid maintenance treatment. The goal of maintenance treatment is to provide a physician ordered or prescribed opioid (either methadone or buprenorphine) to deter and ideally eliminate the use of illicit opioids. Early researchers in this area speculated that an effective pharmacological treatment for opioid-use disorder must eliminate the euphoric appeal of heroin and the abstinence symptoms that draw addicts back to drug use; it must be sufficiently free from toxic dysphoric effects that patients will continue with treatment; it must be orally

effective, long-acting, medically safe, and compatible with normal performance at work and at school and with responsible behavior in society.

Methadone (Dolophine) which came into widespread use in the 1970s was the first medication to meet these requirements. Levomethadyl (ORLAAM) was available from 1993 until 2003 as an opioid maintenance therapy; however, production of levomethadyl was discontinued by the manufacturer due in part to reports of prolonged QT intervals and risk of torsades de pointes. In 2002, buprenorphine and buprenorphine/naloxone (Suboxone, Zubsolv, Bunavail) became the first opioid maintenance medications approved for office-based treatment of opioid-use disorder.

Among the more than 2 million current, long-term opioid users in the United States, as of 2013 only 330,308 were receiving methadone and 12,872 buprenorphine through federally licensed opioid treatment programs with an additional 35,276 receiving buprenorphine from other types of treatment programs. An undetermined number received buprenorphine through physicians' offices. Barriers to patients entering licensed opioid treatment programs include state and federal regulations for program admittance, treatment costs, and inadequate expansion of treatment centers since the 1980s. The social stigma associated with methadone treatment also may deter opioid-dependent individuals from entering treatment. Office-based treatment with buprenorphine, which appears to have nearly comparable effectiveness to clinic-based treatment, is making opioid maintenance therapy more accessible. As of 2014, 25,732 physicians had obtained waivers to prescribe buprenorphine. No clear data exist on the total number of patients treated since FDA approval, but it certainly exceeds 1,000,000. The following discussion reviews the history and current status of methadone and buprenorphine. A discussion on the use of opioid agonists in the treatment of chronic pain is also included.

METHADONE

History

In response to high rates of heroin-related deaths in the 1960s, researchers began investigating the feasibility of opioid maintenance therapy. Vincent Dole, Marie Nyswander, and Mary Jeanne Kreek evaluated several short-acting opioids as possible replacement medications but found that patients were overly sedated by the narcotic effects and still craved heroin. Moreover, patients required frequent dosing—every 4 to 6 hours—to prevent the emergence of withdrawal symptoms. Searching for a longer-acting alternative, the researchers administered methadone, which had been previously shown to be useful in heroin withdrawal. They discovered that, with methadone, it was possible to maintain a heroin-dependent individual "between the limits of narcosis and abstinence [where] there is a

functional zone."

Rapid expansion of methadone maintenance programs occurred in the 1970s during the presidency of Richard Nixon. At this time, politicians were concerned about the high rate of heroin-related crimes and were encouraged by the decrease in illegal activities reported by individuals receiving methadone. However, initial rapid expansion was soon curtailed by government regulation in an attempt to maintain some control over the distribution of methadone. Today, the DEA and the Center for Substance Abuse Treatment, which is part of the federal Substance Abuse and Mental Health Services Administration, both have roles in regulating licensed opioid treatment programs. There remains substantial community resistance to the placement and existence of these programs, which limits the availability of services in some areas. Critics often suggest that methadone maintenance is, in essence, trading one chemical dependency for another and is therefore not appropriate treatment. Finally, there is concern that methadone diversion may occur, further worsening opioid-use disorder epidemics or contributing to opioid overdoses. In reality evidence accrued thus far suggests that diverted methadone comes mainly from methadone prescribed to treat pain, not from methadone dispensed in licensed opioid treatment programs.

Chemistry

Methadone is a synthetic opioid of the phenylheptylamine class structurally unrelated to morphine (Fig. 34.24–1). Methadone consists of a racemic mixture; the *l*-isomer is responsible for the drug's analgesic and other opioid agonist effects. The *d*-isomer exhibits substantially less analgesic activity and does not have respiratory depression activity or addiction liability, but it does have antitussive effects.

Pharmacological Actions

Pharmacokinetics. Methadone is rapidly absorbed after oral administration. It is highly plasma protein bound, with an approximate volume of distribution of 4 L/kg. Methadone is demethylated in the liver by several cytochrome P450 (CYP) enzymes, primarily 3A4, and its inactive metabolite, 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine (EDDP), is excreted in the urine. Of interest, women may metabolize methadone more rapidly than men. After oral administration, methadone has a biexponential half-life with an initial phase range of 12 to 24 hours and a secondary phase of 55 hours. Methadone obviously crosses the blood–brain barrier as well as the placenta and is found in maternal milk. The benefits of breastfeeding far outweigh concerns about exposing infants to the methadone in breast milk, and mothers on methadone maintenance are encouraged to breastfeed.

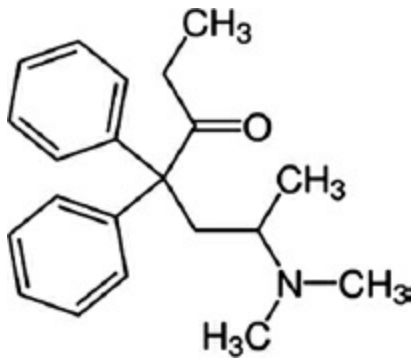

FIGURE 34.24–1. Molecular formula for methadone.

Pharmacodynamics. The major pharmacodynamic action of methadone is agonism at the μ-opioid receptor, but methadone, unlike most other opioids, also acts as an antagonist at the NMDA receptor and blocks serotonin and norepinephrine transporters. In addition, methadone inhibits the cardiac potassium channel, hERG, which can cause a prolonged QT interval on the ECG. There is often not a clear linear relationship between methadone blood levels and its clinical effects. Therefore, it is rarely necessary to obtain blood levels during clinical care of patients on methadone maintenance. Suggested minimally therapeutic trough serum levels of methadone range from 100 to 400 ng/mL. Peak serum levels of methadone vary considerably in stable patients. A peak-to-trough serum ratio of greater than 2:1 is often consider indicative of inadequate pharmacological coverage throughout the 24-hour dosing interval. Patients who display such a pattern may be rapid metabolizers of methadone who would ideally do best on a split-dosing regimen rather than on a higher single daily dose.

Design and Interpretation of Clinical Studies on Methadone

Although methadone maintenance treatment (MMT) was introduced in the 1970s, few randomized, controlled trials have evaluated its efficacy. Although these studies had small sample sizes and rarely followed patients for longer than 1 year, they reported positive results. The majority of data supporting the use of methadone maintenance comes instead from the many observational studies that have been conducted to evaluate this treatment modality. These studies support the findings of clinical trials in demonstrating that methadone treatment reduces heroin or other opioid use and criminal activity. Furthermore, MMT has been shown to decrease mortality among opioid-dependent patients. Mortality estimates indicate that heroin-dependent individuals in MMT are one-fourth as likely to die as dependent individuals not in treatment. MMT also appears to decrease drug use practices, such as needle sharing, that are associated with the transmission of HIV. Methadone treatment appears to be cost-effective, in that the cost per year of life saved is less than that incurred for other

common medical therapies such as coronary bypass surgery or renal dialysis. A review that included seven randomized controlled trials and two quasi-experimental studies of MMT from 1995 to 2012 confirmed all the benefits described above, particularly when the dose of methadone exceeded 60 mg/day and also confirmed improved maternal and fetal outcomes for pregnant women with opioid-use disorder when provided MMT.

Adverse Effects

General. Because methadone shares the adverse effects and toxic potential of other opioid agonists, the usual precautions of opioid agonist therapy should be observed. Common side effects of opioids include sedation, constipation, sweating, nausea, dizziness, and hypotension. During methadone maintenance, most adverse effects disappear over the course of several weeks. However, constipation and excessive sweating often persist even with long-term methadone administration. Recent case reports also indicate that, as noted above, methadone, especially in high doses, although occasionally in moderate doses, can impair cardiac conduction, prolong the QT interval, and in rare instances lead to torsades de pointes. Like other opioids methadone does suppress the hypothalamic–pituitary–gonadal axis and thus can cause hypogonadism with low testosterone levels in men and anovulation and lack of menstruation in women. A potential downstream effect of hypogonadism could be osteoporosis. Methadone does not cause chronic adverse effects on the heart, lungs, liver, or kidneys.

Overdose. Severe intoxication or overdose with opioids constitutes a medical emergency. Miosis (known as *pinpoint pupils*) is a recognized effect of most opioids and can be used to identify opioid intoxication. Of most concern in opioid overdose is potentially fatal respiratory depression from direct suppression of respiratory centers in the midbrain and medulla. For severe cases, the administration of the pure opioid antagonist naloxone in combination with general supportive measures is indicated. Owing to the long half-life of methadone, repeat administration of naloxone may be necessary. An overdose of methadone may also involve other sedative agents that may affect treatment decisions. In 2006, the FDA added a black box warning to the package insert for methadone emphasizing the risks of respiratory depression and QT prolongation. Naloxone is not effective orally but can be prescribed to patients on methadone maintenance either as an auto-injection device (Evzio), by a nasal spray device (Narcan), or in generic form for injection or by nasal spray for administration outside of the health care setting in the event of overdose.

Drug Interactions

Pharmacokinetic interactions are mainly related to methadone being a substrate for CYP450 3A4. Inducers of this enzyme such as the anticonvulsants, phenytoin and carbamazepine, the antituberculous medication rifampin, and antiretroviral medications, nevirapine, efavirenz increase the metabolism of methadone so that the patient could go into opioid withdrawal suddenly and become destabilized. Inhibitors of the enzymes such as the antidepressant fluvoxamine, the antibiotic, erythromycin, and antifungal medications, like ketoconazole have the potential to reduce methadone metabolism and lead to side effects or toxicity, though clinically meaningful effects are rarely observed with inhibitors. Pharmacodynamic interactions obviously can occur with opioid antagonists, such as naloxone or naltrexone, which would cause a patient on methadone maintenance to go into severe, precipitated withdrawal. Other opioids, sedative-hypnotics, and alcohol also have a pharmacodynamic interaction with methadone leading to additive excess sedation and greater risk of suppression of respiratory drive and overdose.

Laboratory Interference

None known.

Dosing

Federal eligibility requirements for MMT stipulate that individuals entering treatment have a minimum 1-year history of opioid-use disorder. The minimum age requirement is 18 years of age; however, individuals between 16 and 18 years of age may enter methadone treatment under certain conditions. There is no upper age limit on program entry. Pregnant women with opioid-use disorder may be accepted for methadone treatment, which is considered the standard of care for these women. Individuals with major medical conditions are also eligible for treatment, as are polydrug users who have opioid-use disorder as one component of their addictive disease.

Methadone dose determination should always be performed by an experienced physician based on his or her clinical assessment of the patient. Data indicate that adequate methadone doses are correlated with reduced illicit drug use and improved patient retention in treatment. A comparison by Eric C. Strain and colleagues of moderate daily doses of methadone (40 to 50 mg) and high doses (80 to 100 mg) showed significantly lower illicit opioid use in the high-dose group. Similar findings have been reported in regard to treatment retention, with one study showing that patients receiving 80 mg/day of methadone were twice as likely to remain in treatment as those receiving 60 to 79 mg, who were, in turn, twice as likely to remain in treatment as patients receiving less than 60 mg. The lowest effective dose of methadone is now widely accepted to be approximately 60 mg, with an optimum average dose for most patients of

approximately 80 mg (±20 mg), although some patients require a dose higher than this range.

In general, methadone can be initiated in an individual with well-established current physiological dependence at a dose of 20 to 30 mg. If there is any doubt as to the degree of an individual's tolerance to opioids, an initial dose of 10 mg is indicated. Under federal regulations the maximum allowable first dose is 30 mg, and the maximum total dose for the first day of treatment is 40 mg unless the physician documents the fact that 40 mg is insufficient to suppress withdrawal symptoms. Dose escalation should be done in small increments (usually 5 to 10 mg at a time) under close physician supervision until an optimum maintenance dose is achieved. Given methadone's long half-life, it is advisable to observe the patient for at least 3 to 4 days before an additional dose increase is provided. The optimum dose is achieved when (1) the patient evidences no withdrawal signs or symptoms throughout the 24-hour dosing period, (2) the patient reports an absence of craving for opioids, (3) adequate tolerance is obtained such that the patient experiences little or no reinforcement from use of other opioids, and (4) self-report and urine testing indicate the absence of illicit opioid use. Ideally these beneficial effects need to be balanced against side effects, so that small dose decrements may be required if unpleasant or intolerable side effects occur.

The duration of maintenance therapy may range from a few months to an indefinite period. In general, methadone maintenance should be continued as long as the patient is continuing to benefit from and wishes to remain in treatment. It should be noted that methadone administration should ideally be combined with comprehensive counseling and rehabilitative services.

BUPRENORPHINE

History

Reckitt and Coleman Ltd. developed buprenorphine in the 1970s as an analgesic because of the widespread interest in developing mixed agonist and antagonist opioid agents for pain relief. It was hoped that these agents would offer analgesia with greater safety and less abuse liability than the pure opioid agonists. Buprenorphine is available in many countries as an analgesic. It was first marketed in the United Kingdom in 1978 for injection and in 1981 as a sublingual tablet.

The potential for the use of buprenorphine in the treatment of opioid addiction was first recognized in the late 1970s when Donald Jasinski and colleagues demonstrated that buprenorphine substituted for morphine and suppressed opioid withdrawal while having few withdrawal symptoms. Human laboratory studies have demonstrated that buprenorphine suppresses heroin self-administration. A number of published clinical

observations have also supported the usefulness and general safety of buprenorphine in clinical settings. However, in clinical practice, buprenorphine has been widely used as an injectable analgesic, and intravenous (IV) buprenorphine has been reported as a significant drug of abuse in Europe, New Zealand, Australia, and the Republic of Georgia. In 2002, the FDA approved the use of buprenorphine for the treatment of opioid dependence. It is commercially available as a generic sublingual tablet containing buprenorphine only or a combination of buprenorphine and naloxone which can come as a sublingual tablet, a sublingual film strip, or a buccal patch. Only the sublingual formulation or the buccal patch, and not the parenteral or transdermal formulations, are approved for the treatment of opioid-use disorder. In 2016, the FDA-approved a subdermal formulation of buprenorphine requiring surgical implantation that has a 6 month duration of action. This formulation is approved for patients who have been stable on a dosage of sublingual buprenorphine of 8 mg daily or less for a minimum of 90 days.

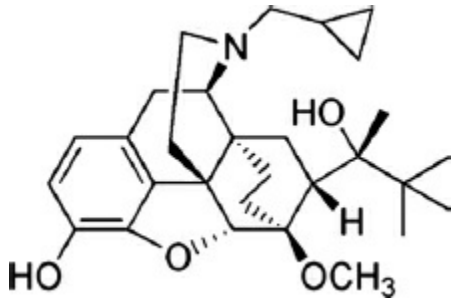

FIGURE 34.24–2. Molecular formula for buprenorphine.

Chemistry

Buprenorphine is a semisynthetic derivative of the opioid alkaloid, thebaine (Fig. 34.24–2).

Pharmacological Actions

Pharmacokinetics. Because of extensive metabolism by the intestine and liver, buprenorphine is poorly absorbed by the oral route and is therefore best administered using IV or sublingual, buccal, or transdermal dosing. When administered by the IV route, buprenorphine is readily distributed into tissues, resulting in a large volume of distribution. The elimination is biphasic, with a rapid distribution phase ($T_{1/2}$ = 2 minutes) and a slower elimination phase ($T_{1/2}$ = 3 to 5 hours). The rate of absorption is slower by the sublingual route, with a time to reach maximum drug concentration (T_{max}) of approximately 60 minutes. It takes 3 to 5 minutes to dissolve, and sublingual bioavailability is approximately 55 percent. There is considerable between-subject variability in bioavailability, but the

data suggest greater bioavailability of buprenorphine/naloxone versus buprenorphine alone. Buprenorphine has a variable (range 3 to 44 hours) but typically long half-life. Following three-times-per-day dosing, steady-state levels of buprenorphine are reached at approximately 4 days. Buprenorphine is 96 percent bound to plasma proteins. Buprenorphine undergoes both glucuronidation and *N*-dealkylation, with *N*-dealkylation catalyzed by the liver enzyme CYP450 3A4 producing an active metabolite, norbuprenorphine. The main route of elimination for both parent drug and metabolites is fecal with lesser amounts excreted by the kidneys. Buprenorphine of course crosses the blood–brain barrier as well as the placenta and is found in maternal milk. The benefits of breastfeeding far outweigh concerns about exposing infants to the buprenorphine in breast milk, especially since buprenorphine has such low bioavailability by the oral route, so mothers on buprenorphine maintenance are encouraged to breastfeed.

Pharmacodynamics. Pharmacological studies have shown that buprenorphine combines a strong affinity for the μ-opioid receptor with low intrinsic activity. As such, it is considered a partial μ agonist. Buprenorphine is also an antagonist at the κ receptor. Because buprenorphine is a partial agonist, it has a considerable margin of safety with little chance of lethal overdose.

While blood levels of buprenorphine and norbuprenorphine are generally proportional to dose, to date no specific correspondence between blood levels and therapeutic effect has been demonstrated.

Design and Interpretation of Clinical Studies on Buprenorphine

A series of well-controlled clinical trials has clearly established the safety and clinical efficacy of buprenorphine. These trials have demonstrated the efficacy of buprenorphine, as compared to placebo, in decreasing illicit opioid use in opioid-dependent individuals. Subsequent trials have established the safety and efficacy of buprenorphine in maintenance treatment of opioid dependence in doses ranging from 2 to 32 mg/day.

Controlled clinical trials have demonstrated that buprenorphine and moderate-dose methadone (50 to 60 mg/day) have similar effects on retention and illicit drug use in maintenance programs. Comparisons of buprenorphine with higher doses of methadone suggest that high-dose methadone may have a slight advantage in terms of treatment retention and suppression of illicit opioid use. One recent study suggests that higher buprenorphine doses are associated with better treatment retention. Studies by Walter Ling and colleagues have demonstrated successful implementation and use of buprenorphine in private physicians' offices. Further studies have indicated that buprenorphine works as well or better for patients who primarily misuse prescription opioids as opposed to

heroin; that for patients without severe psychiatric illness outcomes for buprenorphine are as good with quality medical management alone as compared to addiction-focused counseling or psychotherapy added to medical management; and that outcomes with maintenance buprenorphine treatment are superior to outcomes when buprenorphine is tapered.

The Drug Addiction Treatment Act, which was signed into law in October 2000, marked the beginning of a process designed to allow qualified physicians to prescribe approved narcotic drugs for the treatment of opioid-use disorder in office-based settings. There are a number of approved training programs in place to teach physicians about the use of buprenorphine. This is the first time in modern medicine that physicians in the United States practicing in a variety of clinical settings, including office-based practice, are able adequately to treat opioid-use disorder with pharmacotherapy. This initiative has increased access to treatment for some individuals with opioid-use disorder, but in the face of an epidemic of prescription opioid and heroin use, a treatment gap still remains.

Adverse Effects

Because buprenorphine is a partial agonist rather than a full agonist at the μ receptor, and it is a weak antagonist at the κ receptor, this agent produces a milder withdrawal syndrome and has a wider margin of safety than the full μ-agonist compounds generally used in treatment. Buprenorphine has a ceiling effect beyond which dose increases prolong the duration of action of the drug without further increasing the agonist effects. Because of this, buprenorphine has a high clinical safety profile, with limited respiratory depression, therefore decreasing the likelihood of lethal overdose. Buprenorphine does have the capacity to cause typical side effects associated with opioids, including sedation, nausea and vomiting, constipation, dizziness, headache, and sweating. Buprenorphine may have the potential to suppress the hypothalamic–pituitary–gonadal axis, but this possible adverse effect requires further study. Buprenorphine does not cause chronic adverse effects on the heart, lungs, or kidneys. Some small early studies and case reports raised a concern that it could be hepatotoxic, but more recent, larger studies show no evidence of a hepatotoxic effect.

Drug Interactions

Theoretical pharmacokinetic interactions have been proposed mainly related to buprenorphine also being a substrate for CYP450 3A4. However, to the extent that these theoretical interactions have been studied in the human laboratory, no clinically meaningful effects have been observed, and no isolated clinical case reports have surfaced. It has been proposed that because buprenorphine has high affinity for and slow dissociation from the μ-opioid receptor, changes in plasma levels have little effect on efficacy.

Pharmacodynamic interactions obviously can occur with opioid antagonists, such as naloxone or naltrexone, although because of buprenorphine's high affinity and slow dissociation, there is competition for the receptor with likelihood of precipitated withdrawal being low. Other opioids, sedative-hypnotics, and alcohol could also have a pharmacodynamic interaction with buprenorphine leading to additive excess sedation and greater risk of suppression of respiratory drive and potential overdose. This risk is less than with full-agonist opioids, and overdoses in adults on buprenorphine alone are nearly nonexistent.

Dosing

To reduce the likelihood of abusing buprenorphine via the IV route, buprenorphine has been combined with the opioid antagonist naloxone for sublingual administration. Because naloxone is poorly absorbed by the sublingual route, when the combination drug is taken sublingually, there is no effect of the naloxone on the efficacy of buprenorphine. If an individual with opioid-use disorder injects the combination medication, the naloxone may precipitate a withdrawal reaction and/or block the intended effects of the buprenorphine being sought, therefore reducing the likelihood of illicit injection use of the sublingual preparation.

Inducting and stabilizing a patient on buprenorphine is analogous to inducting and stabilizing a patient on methadone except that, as a partial agonist, buprenorphine has the potential to cause precipitated withdrawal in patients who have recently taken full-agonist opioids. Thus, a patient must abstain from use of short-acting opioids for 12 to 24 hours prior to starting buprenorphine and from longer-acting opioids such as methadone for 24 to 48 hours or longer. The physician must assess the patient clinically and determine that the patient is in mild to moderate opioid withdrawal with objectively observable withdrawal signs prior to initiating buprenorphine. In most instances a relatively low dose of buprenorphine (2 to 4 mg) can then be administered with additional doses given in 1 to 2 hours if withdrawal signs persist. The goal for the first 24 hours is to suppress withdrawal signs and symptoms, and the total 24-hour dose to do so can range from 2 to 16 mg on the first day. In subsequent days the dose can be adjusted upward or downward to resolve withdrawal fully and, as with methadone, to achieve an absence of craving, adequate tolerance to prevent reinforcement from the use of other opioids, and ultimately abstinence from other opioids while minimizing side effects. Dose-ranging studies have demonstrated that doses of 6 to 16 mg/day are associated with improved treatment outcomes, as compared to lower doses of buprenorphine (1 to 4 mg). Sometimes patients seem to need doses higher than 16 mg/day, although there is no evidence for any benefit of doses beyond 32 mg/day. For the treatment of opioid-use disorder, a dose of approximately 4 mg of sublingual buprenorphine is the equivalent of a

daily dose of 40 mg of oral methadone. It has also been demonstrated that daily, alternate-day, or three-times-per-week administration have equivalent effects in suppressing symptoms of opioid withdrawal in physiologically dependent individuals. The buprenorphine/naloxone combination formulations are recommended for most clinical purposes, including induction and maintenance. The buprenorphine mono should be used only for pregnant patients or for patients who have a documented anaphylactic reaction to naloxone.

Newer forms of buprenorphine delivery including a long-acting depot intramuscular injection that provides therapeutic plasma levels for several weeks are being investigated. Such a delivery system, as well as the subdermal implants mentioned above, could obviate the need for taking medications daily while virtually eliminating the risk of medication nonadherence or diversion.

OPIOID AGONISTS IN THE TREATMENT OF PAIN IN INDIVIDUALS WITH ADDICTIVE DISORDERS

Opioid agonists are efficacious agents for the treatment of acute pain and possibly for some individuals with chronic pain syndromes, although their long-term use in chronic pain has never been demonstrated in randomized clinical trials, and the use of opioid agonists in the treatment of pain in individuals with substance use disorders is controversial. Untreated pain can be a powerful trigger for relapse. On the other hand, exposure to opioid analgesics for some individuals with addictive disorders could trigger reminders of past use and place them at risk for relapse. In addition, some of the physiological aspects of addictive disorders may make pain more difficult to treat in the addicted individual than in the nonaddicted individual. Finally, it can be difficult to distinguish distress related to pain from distress related to drug craving in individuals with opioid-use disorder. In spite of these potential problems, individuals with substance use disorders develop both acute and chronic pain disorders that warrant treatment, and total prohibitions against the use of opioids is neither practical nor good medical practice. However, pain management in individuals with addictions must attend to the treatment of both the pain and the addictive disorder to be successful. Ongoing support for maintaining recovery is essential for patients taking opioids for pain, including attendance at 12-step meetings and/or regular follow-up with a treatment professional who understands the issues involved in the treatment of pain in individuals with addictions.

Acute Pain Management

Undertreatment of pain associated with trauma and surgery has been demonstrated to increase morbidity and to lengthen hospital stays. In

individuals with addictive disorders, acute pain conditions increase distress and anxiety and may become a risk factor for relapse. Physicians treating patients in acute pain who have addictive disorders should reassure the patient that his or her addiction is not an obstacle to the appropriate and effective treatment of the pain condition.

For patients with opioid-use disorder, the physician should not consider tapering or discontinuing opioid medications until the acute pain situation is resolved. The use of nonmedication pain treatment approaches, such as ice or a transcutaneous electrical nerve stimulation (TENS) unit, should be optimized. When medications are necessary, it is best to use those least likely to alter mood or to produce physical dependence, if they are effective. Nonsteroidal anti-inflammatory drugs, tricyclic antidepressants, anticonvulsants, and topical agents all have demonstrated efficacy in certain pain conditions. However, when opioid agents are necessary to manage pain conditions, they should be used effectively. In the management of moderate to severe acute pain, opioids are generally the mainstay of treatment in addicted or nonaddicted individuals. Scheduled administration of opioids is preferred to as-needed administration in individuals with addictive disorders, providing several advantages. If the patient needs to ask for medications, this may be misconstrued as drug-seeking behavior. In addition, when drug administration is time dependent rather than symptom dependent, reinforcement of pain symptoms is minimized.

Individuals on methadone maintenance or physically dependent on other opioids should have their baseline opioid requirements met in addition to the medications provided for pain treatment. At times these individuals require more opioid pain medication than the average individual because of their opioid tolerance and associated hyperalgesia. When possible, patients on methadone should receive a different opioid medication for the control of acute pain or have additional methadone prescribed on a three- or four-times-daily basis. Although an increase in methadone in divided doses throughout the day is generally effective in acute pain management, this strategy makes it easy to confuse issues of pain treatment with issues of addiction treatment. When the acute pain condition resolves, the taper of methadone back to the original maintenance dose can at times prove difficult for the patient.

The management of patients on buprenorphine who develop an acute pain condition is somewhat more complex, with less data and clinical experience to guide it. Although buprenorphine is an effective analgesic, it may not have the analgesic potency of a full-agonist opioid. There are essentially three options. One is to increase the dose of buprenorphine and give it in divided doses throughout the day. Another option is to try adding a full-agonist opioid to the regimen. It is unclear how efficacious a full agonist might be in the face of receptor occupancy by the partial-agonist

buprenorphine. A third option is to discontinue buprenorphine and switch temporarily to a full-agonist opioid during the time acute pain treatment is needed. When the acute pain subsides, the patient can be reinducted on buprenorphine, with care taken not to administer buprenorphine too soon after ingestion of the full agonist to avoid precipitating withdrawal.

Management of Chronic Pain

The treatment of chronic pain in individuals with addictive diseases also does not differ significantly from the treatment of these pain conditions in nonaddicted individuals. The goals of treatment are the reduction of pain, improvement in associated features, and restoration of function. Generally, these goals are approached initially through the use of nonpharmacological strategies, such as a TENS unit, peripheral nerve blocks, or trigger point injections. The mainstay of medication treatment for chronic, nonmalignant pain includes the nonsteroidal anti-inflammatory drugs, tricyclic antidepressants, and anticonvulsants. Anticonvulsants (pregabalin) and antidepressant agents (duloxetine) have received FDA approval for specific chronic nonmalignant pain syndromes. There is now a Centers for Disease Control and Prevention Guideline for Prescribing Opioids for Chronic pain which emphasizes the importance of using a standardized approach so that decisions about the use of opioids and dose increases can be clearly justified. Clinicians should be aware that long-term opioid use may be associated with the development of abnormal sensitivity to pain or hyperalgesia, which can manifest clinically as needing increasing doses of opioids to maintain the same level of analgesia. However, dose escalation should be approached with caution, and ongoing assessment to determine whether the increased dosing is meeting the goals of treatment (decreased pain, improved functionality) is essential. Failure to meet goals requires reevaluation and change in treatment plan. Medication agreements may be a useful tool in the medication management of pain in individuals with addictions. A medication agreement is a written, signed statement that reinforces the patient's responsibility as a prerequisite for receiving controlled substance prescriptions and clarifies the boundaries of treatment. These can be useful in avoiding misunderstanding and clarifying decision points concerning opioid use in the treatment process.

Even when medications, including opioids, are deemed necessary in the treatment of chronic pain syndromes, the treatment plan should be multidimensional. Specifically, it should include a careful history of the development of the pain problem, factors that contribute to the patient's pain and distress, and a plan for dealing with each of these contributing factors. The treatment plan may include the physical treatments of pain mentioned previously (i.e., TENS unit, nerve blocks), as well as behavioral interventions, such as relaxation training, biofeedback, and changes in behavioral responses to pain. The use of long-term opioids as one

component in the treatment of chronic pain is generally a reasonable approach when other treatment regimens have failed. When opioids are used, it is important to document their benefit by observing that functional status is improving. The majority of chronic pain patients achieve relief of pain through multimodal approaches that do not include the use of opioids. Often, patients report that they feel better with the elimination of opioid use. For most patients, a trial period with the use of coordinated alternative treatments should be tried before long-term opioid therapy is used.

FUTURE DIRECTIONS

The efficacy of opioid agonist medications in the treatment of opioid dependence has been clearly established. In spite of the data demonstrating that maintenance programs are effective in decreasing opioid use, improving employment and prosocial behaviors, and enhancing public health and safety, only a small percentage of opioid-dependent individuals in the United States are in treatment. It is hoped that legislation making it possible to prescribe opioid agonist therapies in a variety of settings and the development of agonist agents with a greater margin of safety and less abuse potential will increase the number of opioid-dependent individuals receiving treatment.

REFERENCES

Ball JC, Ross A. *The Effectiveness of Methadone Maintenance Treatment: Patients, Programs, Services, and Outcome*. New York: Springer-Verlag; 1991.

Ballantyne JC, Mao J. Opioid therapy for chronic pain. *N Engl J Med*. 2003;349:1943.

Barnett PG, Zaric GS, Brandeau ML. The cost-effectiveness of buprenorphine maintenance therapy for opiate addiction in the United States. *Addiction*. 2001;96:1267–1278.

*Center for Substance Abuse Treatment. *Clinical Guidelines for the Use of Buprenorphine in the Treatment of Opioid Addiction*. Treatment Improvement Protocol (TIP) Series 40, HHS Publication No. (SMA) 04–3939. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2004.

*Center for Substance Abuse Treatment. *Medication-Assisted Treatment for Opioid Addiction in Opioid Treatment Programs*. Treatment Improvement Protocol (TIP) Series 43, HHS Publication No. (SMA) 05–4048. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2005.

Collins ED, Kleber HD, Whittington RA, Heitler NE. Anesthesia-assisted vs. buprenorphine- or clonidine-assisted heroin detoxification and naltrexone induction: a randomized trial. *JAMA*. 2005;294(8):903–913.

Dole VP, Nyswander M. A medical treatment for diacetylmorphine (heroin) addiction. A clinical trial with methadone hydrochloride. *JAMA*. 1965;193:646–650.

Doran CM, Shanahan M, Mattick RP, Ali R, White J, Bell J. Buprenorphine versus methadone maintenance: a cost-effectiveness analysis. *Drug Alcohol Depend*. 2003;71:295–302.

*Effective medical treatment of opiate addiction. National Consensus Development Panel on

Effective Medical Treatment of Opiate Addiction. *JAMA*. 1998;280:1936–1943.

Ehret GB, Voide C, Gex-Fabry M, et al. Drug-induced long QT syndrome in injection drug users receiving methadone: high frequency in hospitalized patients and risk factors. *Arch Intern Med*. 2006;166(12):1280–1287.

Elkader A, Sproule B. Buprenorphine: clinical pharmacokinetics in the treatment of opioid dependence. *Clin Pharmacokinet*. 2005;44(7):661–680.

Fiellin DA, Moore BA, Sullivan LE, et al. Long-term treatment with buprenorphine/naloxone in primary care: results at 2–5 years. *Am J Addict*. 2008;17:116–120.

Fiellin DA, Schottenfeld RS, Cutter CJ, Moore BA, Barry DT, O'Connor PG. Primary care-based buprenorphine taper vs maintenance therapy for prescription opioid dependence: a randomized clinical trial. *JAMA Intern Med*. 2014;174:1947–1954.

Fudala PJ, Bridge TP, Herbert S, et al. Office-based treatment of opiate addiction with a sublingual-tablet formulation of buprenorphine and naloxone. *N Engl J Med*. 2003;349:949–958.

Fullerton CA, Kim M, Thomas CP, et al. Medication-assisted treatment with methadone: assessing the evidence. *Psychiatr Serv*. 2014;65:146–157.

Gibson A, Degenhardt L, Mattick RP, Ali R, White J, O'Brien S. Exposure to opioid maintenance treatment reduces long-term mortality. *Addiction*. 2008;103:462–468.

Heit HA, Gourlay DL. Buprenorphine: new tricks with an old molecule for pain management. *Clin J Pain*. 2008;24:93–97.

Hser YI, Hoffman V, Grella CE, Anglin MD. A 33-year follow-up of narcotics addicts. *Arch Gen Psychiatry*. 2001;58:503–508.

Hser YI, Saxon AJ, Huang D, et al. Treatment retention among patients randomized to buprenorphine/naloxone compared to methadone in a multi-site trial. *Addiction*. 2014;109:79–87.

Jasinski DR, Pevnick JS, Griffith JD. Human pharmacology and abuse potential of the analgesic buprenorphine: a potential agent for treating narcotic addiction. *Arch Gen Psychiatry*. 1978;35(4):501–516.

Johnson RE, Jaffe JH, Fudala PJ. A controlled trial of buprenorphine treatment for opioid dependence. *JAMA*. 1992;267:2750–2755.

Likar R, Kayser H, Sittl R. Long-term management of chronic pain with transdermal buprenorphine: a multicenter, open-label, follow-up study in patients from three short-term clinical trials. *Clin Ther*. 2006;28(6):943–952.

*Ling W, Charuvastra C, Collins JF, et al. Buprenorphine maintenance treatment of opiate dependence: a multicenter, randomized clinical trial. *Addiction*. 1998;93:475–486.

Ling W, Hillhouse M, Ang A, Jenkins J, Fahey J. Comparison of behavioral treatment conditions in buprenorphine maintenance. *Addiction*. 2013;108:1788–1798.

Mattick RP, Kimber J, Breen C, Davoli M. Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. *Cochrane Database Syst Rev*. 2014;(2):CD002207.

Dowell D, Haegerich, TM, Chou R. CDC guideline for prescribing opioids for chronic pain–United States, 2016. *MMWR Recomm Rep*. 2016;65(1).

Saxon AJ, Ling W, Hillhouse M, et al. Buprenorphine/naloxone and methadone effects on

laboratory indices of liver health: a randomized trial. *Drug Alcohol Depend*. 2013;128:71–76.

Sigmon SC, Moody DE, Nuwayser ES, Bigelow GE. An injection depot formulation of buprenorphine: extended bio-delivery and effects. *Addiction*. 2006;101(3):420–432.

Strain EC, Bigelow GE, Liebson IA, Stitzer ML. Moderate- vs high-dose methadone in the treatment of opioid dependence: a randomized trial. *JAMA*. 1999;281:1000–1005.

Strain EC, Moody DE, Stoller KB, Walsh SL, Bigelow GE. Relative bioavailability of different buprenorphine formulations under chronic dosing conditions. *Drug Alcohol Depend*. 2004;74:37–43.

Substance Abuse and Mental Health Services Administration. *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings,* NSDUH Series H-48, HHS Publication No. (SMA) 14–4863. Rockville, MD: Department of Health and Human Services; 2014.

Thomas CP, Fullerton CA, Kim M, et al. Medication-assisted treatment with buprenorphine: assessing the evidence. *Psychiatr Serv*. 2014;65:158–170.

Verebely K, Volavka J, Mulé S, Resnick R. Methadone in man: pharmacokinetic and excretion studies in acute and chronic treatment. *Clin Pharmacol Ther*. 1975;18:180–190.

Weiss RD, Potter JS, Fiellin DA, et al. Adjunctive counseling during brief and extended buprenorphine-naloxone treatment for prescription opioid dependence: a 2-phase randomized controlled trial. *Arch Gen Psychiatry*. 2011;68:1238–1246.

▲ 34.25 Opioid Receptor Antagonists: Naltrexone and Nalmefene

SRINIVAS B. MUVVALA, M.D., SUCHITRA KRISHNAN-SARIN, PH.D., AND STEPHANIE S. O'MALLEY, PH.D.

INTRODUCTION AND HISTORY

Naltrexone, an opioid antagonist, was synthesized in 1965 by H. Blumberg and H. B. Dayton and proved to have a number of advantages over earlier opioid antagonists, including cyclazocine, nalorphine, and naloxone. Specifically, naltrexone had a relatively long half-life, was orally effective, was not associated with dysphoria, and could be administered once a day. Whereas earlier narcotic antagonists—principally naloxone—were used to reduce narcotic overdose, these characteristics of naltrexone made it suitable for use in preventing relapse to opiate use in detoxified opiate addicts. In 1984, naltrexone (under the original trade name Trexan) was approved for providing blockade of the effects of exogenously administered opiates in detoxified opiate addicts. Over the ensuing years, naltrexone has been tried for the treatment of a wide range of psychiatric disorders, including, among others, eating disorders, autism, self-injurious behavior, cocaine dependence, gambling, and alcoholism. Naltrexone (ReVia; Depade) was approved for the treatment of alcohol dependence in 1994. A number of generic formulations are also available. An extended-release, once-a-month injectable suspension (Vivitrol) was also approved in 2006 to treat alcohol dependence and in 2010 to treat opioid dependence. In

2014, FDA approved Contrave, which is a trade name for the combination of naltrexone hydrochloride and bupropion hydrochloride in an extended--release formulation, for treatment of obesity. Nalmefene (Revex) is indicated for the complete or partial reversal of opioid drug effects and in the management of known or suspected opioid overdose. An oral formulation of Nalmefene is available in some countries but not in the United States. Naloxone, an opiate antagonist, has been FDA approved for reversal of opioid overdose. It is an effective medication in reversing respiratory depression and preventing fatalities secondary to opioid overdose.

CHEMISTRY

Naltrexone, or 17-(cyclopropylmethyl)-4,5a-epoxy-3,14-dihydroxymorphinan-6-one hydrochloride, is a competitive opioid antagonist that was developed by substituting an *N*-methyl group on the opioid agonist oxymorphone with a methylcyclopropyl moiety. Naltrexone is available as a white, crystalline hydrochloride salt that is soluble in water, with a melting point of 275°C and a pKa of 8.13 at 37°C. Its molecular formula is $C_{20}H_{23}NO_4 \cdot HCl$. Nalmefene is a 6-methylene analog of naltrexone with a molecular formula $C_{21}H_{25}NO_3$, and it is also available as a hydrochloride salt with a pKa of 7.6 (Fig. 34.25–1).

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

Naltrexone is rapidly and completely absorbed following oral administration of a 100-mg dose, with peak plasma concentrations of 19 to 44 mg/L being achieved within 1 hour. Oral bioavailability of naltrexone is 60 percent, with 20 percent of the drug being bound to plasma proteins. A single dose of naltrexone has a volume of distribution of 16.1 L/kg, which decreases to 14.2 L/kg following repeated administration.

Naltrexone is primarily recovered in plasma, urine, and feces as conjugated naltrexone and conjugated and unconjugated 6β-naltrexol along with minor quantities of 2-hydroxy-3-*O*-methyl-6β-naltrexol and 2-hydroxy-3-*O*-methylnaltrexone. Approximately 60 percent of the drug is recovered in the urine in 48 hours following oral administration and in 72 hours following intravenous administration. Naltrexone has a plasma half-life of 10.5 hours regardless of whether it is administered via the oral or intravenous route. Conjugated 6β-naltrexol has been shown to have a plasma half-life of 14 to 19 hours, which has been proposed to contribute to the longer duration of action of naltrexone. Studies indicate that after a 100-mg dose, naltrexone can be detected in plasma for 30 hours, whereas 6β-naltrexol can be detected for up to 125 hours. The extended-release injectable formulation (Vivitrol), containing 380 mg of naltrexone, was

created using Medisorb drug-delivery technology for once-a-month intramuscular (IM) injections. Following administration of this formulation, an initial peak is seen within 2 hours, followed by a second peak 2 to 3 days later, with a slow decline in concentrations after 14 days. Total naltrexone exposure is three- to fourfold higher than that observed with daily oral dosing over a 28-day period. The elimination half-life of both naltrexone and β-naltrexol in this form is 5 to 10 days.

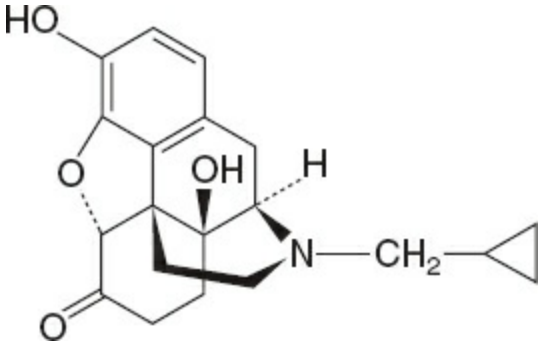

FIGURE 34.25–1. Molecular structure of naltrexone (ReVia).

Nalmefene, following administration of a 1-mg intravenous dose, has a volume of distribution of 8.6 L/kg and a terminal half-life of 10.8 hours. Nalmefene is completely bioavailable following IM and subcutaneous (SC) routes, with time-to-peak plasma concentrations of 2.3 (±1.1) and 1.5 (±1.2) hours, respectively. Because of variability in the speed of absorption of IM and SC doses, repeated dosing should be done with great care. Nalmefene is primarily metabolized though glucuronide conjugation in the liver and is excreted in the urine. Both the glucuronide and the dealkylated metabolite (which is formed in trace amounts) have minimal pharmacological activity. Nalmefene when administered orally has a pharmacokinetic profile that is similar to that of naltrexone. A metabolite, nornalmefene, has been identified that also exists in glucuronidated and alkylated forms. Hepatic or renal impairment decreases clearance of both naltrexone and nalmefene.

Pharmacodynamics

Naltrexone is a nonspecific opioid antagonist that binds to all three opiate receptor sites (μ, κ, δ) as a function of dose administered. Naltrexone has minimal opioid agonistic effects, and its relative opiate antagonistic potency is 12 times that of nalorphine and 2.5 times that of naloxone or cyclazocine. The time course of action of oral naltrexone has been shown to be about three times longer than that of naloxone in rodents. No significant species differences in the binding of [^{3}H]naltrexone have been documented. An oral 100-mg dose of naltrexone was found almost completely to antagonize the subjective and objective responses of 25 mg of IV heroin.

Similarly, following a 1-mg IV dose, nalmefene blocks greater than 80 percent of brain opioid receptors within 5 minutes after administration. Whereas nalmefene has been shown to have greater antagonistic potency at the δ-opioid receptors, it has also been shown following oral administration of a dose of 20 mg to produce complete and prolonged blockade of the μ-opioid receptors.

Acute administration of 50 mg of naltrexone has been shown to precipitate an abstinence syndrome in morphine-dependent individuals, an effect that decreases with chronic use. This decrease is believed to be due to increases in the number and functional activity of brain opioid receptors with long-term naltrexone administration.

DESIGN AND INTERPRETATION OF CLINICAL DRUG STUDIES

Naltrexone and nalmefene have demonstrated efficacy in blocking exogenously administered opiates and were used in protocols for rapid opioid detoxification. Both agents have also been extensively examined as aids in preventing relapse in patients with alcohol dependence. Both agents, particularly naltrexone have shown benefit in reducing heavy drinking in numerous studies. Extended-release formulations of naltrexone have been evaluated to prevent relapse in both alcohol and opiate use disorders. Naltrexone has also been evaluated as a treatment for self-injurious behavior and other autistic behaviors in patients with mental retardation and autism. Support for use in patients with mental retardation and autism is based only on small studies and case reports that have not been confirmed by larger, double-blind clinical trials. Case reports and small open-label trials have also examined naltrexone's utility in treating compulsive sexual behavior and kleptomania. Open-label and randomized, controlled trials using relatively small samples have examined the efficacy of naltrexone in reducing binge eating. Randomized placebo-controlled trials have evaluated the extended-release combination of naltrexone and bupropion for the treatment of obesity. Use of nalmefene and naltrexone for pathological gambling disorder is supported by some randomized clinical trials. Naltrexone has been evaluated for cocaine use disorders without much success. Although randomized clinical trials suggest the possible utility for naltrexone as adjunctive therapy for smoking cessation, the effects are inconsistent and may vary by gender.

THERAPEUTIC INDICATIONS

Alcohol Use Disorders

The efficacy of naltrexone was originally demonstrated in the 1990s in two double-blind, placebo-controlled trials of naltrexone in alcohol-dependent patients. In each of these studies, treatment with an oral dose of 50 mg daily was initiated following a minimum of 7 days of abstinence or

completion of medical detoxification and was provided with concurrent psychosocial interventions. In a combined analysis of these studies, naltrexone-treated patients were more likely to remain abstinent (54 percent vs. 31 percent), to experience fewer days of drinking (3.1 percent vs. 7.5 percent), and to report lower rates of relapse to heavy drinking (25 percent vs. 52 percent). Furthermore, participants who lapsed (i.e., were nonabstinent) were less likely to experience heavy drinking if they were taking naltrexone rather than placebo. The COMBINE Study, a large multisite trial combining pharmacotherapy and behavioral interventions, conducted in 11 academic sites across the United States showed that 100-mg dose of naltrexone, when given along with medical management was more effective than placebo in increasing percent days of abstinence and reducing heavy drinking days. In a recent meta-analysis of pharmacotherapies for alcohol dependence, 53 trials studied the efficacy of naltrexone (n = 9,140), and treatment duration ranged from 12 to 52 weeks. Naltrexone was shown to have benefit in various consumption outcomes. Naltrexone is particularly shown to have efficacy in preventing return to heavy drinking and reducing percent drinking days and percent heavy drinking days.

An extended-release formulation of naltrexone has also shown similar efficacy in reducing the rate of heavy drinking. A large multisite double-blind placebo-controlled study evaluated the efficacy and tolerability of 2 doses of intramuscular long-acting naltrexone injections (380 mg and 190 mg) versus placebo in alcohol-dependent patients. The long-acting preparation was well tolerated over a 6-month period and the 380-mg injectable naltrexone preparation showed 25 percent reduction in number of heavy drinking days. In the subset of patients who were abstinent 4 or more days prior to treatment initiation, the 380-mg preparation also prolonged abstinence and reduced the number of heavy drinking days and drinking days. Some research also suggests that naltrexone may have utility when used to control progression to heavier drinking behavior without a goal of abstinence.

Despite a substantial literature supporting the use of naltrexone for reducing drinking behavior, the efficacy of this agent is modest. This evidence has motivated researchers to examine predictors of response to naltrexone. Recent evidence suggests that the efficacy of naltrexone may be more pronounced in certain subgroups of patients, such as male drinkers, those with a positive family history of alcoholism, those who have antisocial traits, or smokers. Individuals with high levels of craving also seem to benefit from naltrexone. The predictive utility of the Asn40Asp polymorphism of the μ-opioid receptor (OPRM1) gene in treatment response is being debated with some studies showing a stronger effect of naltrexone in Asp40 carriers while others do not. As a result, the data are still insufficient to recommend using these moderators in predicting

response in clinical practice. Because of the good safety profile of naltrexone and support for its effectiveness in a range of patient populations, naltrexone should be offered as one of the treatment options to all patients with alcohol dependence who do not have contraindicating conditions.

Nalmefene is now approved by the European Medicines Agency for reduction of alcohol consumption in patients with high-risk drinking who do have alcohol withdrawal symptoms and who do not require alcohol detoxification. The studies supporting the approval tested nalmefene taken as need (i.e., in high-risk situations, when drinking is imminent, or as soon as possible after drinking) compared to placebo. Although there have been fewer studies of nalmefene, a recent meta-analysis concluded that nalmefene led to a reduction in heavy drinking days per month and the number of drinks per drinking day.

Several small trials have also examined the combined use of opioid antagonists with other agents. Preliminary small placebo-controlled studies suggest that combination of the 5-HT_3 antagonist ondansetron and naltrexone, when compared with placebo, may produce greater suppression of drinking outcomes in early-onset alcoholics. Combined use of naltrexone with disulfiram has shown modest efficacy in increasing abstinence rates in patients with concurrent axis I psychiatric disorders and alcohol dependence. A large, multisite trial indicated no significant advantage of using naltrexone in combination with acamprosate for the treatment of alcohol dependence. Combination of low-dose naltrexone and prazosin has shown some promise in patients with comorbid PTSD. Addition of gabapentin to naltrexone has shown to improve drinking outcomes than naltrexone alone in the first 6 weeks of abstinence possibly secondary to beneficial effects with symptoms such as insomnia and mood instability.

Narcotic Addiction

Although naltrexone was originally developed for the treatment of opiate addiction, the value of naltrexone for this indication has been limited primarily by two issues: (1) the need for opiate detoxification prior to induction and (2) poor adherence compared to agonist therapy. The newer extended-release formulations address to a large extent the issue of adherence at least for the 30 days between injections. The advantages and disadvantages of naltrexone for the treatment of narcotic addiction are discussed below.

As an opioid antagonist, naltrexone selectively competes with exogenous opioids for central nervous system (CNS) and non-CNS opioid receptors and blocks their activity. This blockade lasts for up to 72 hours after naltrexone administration and prevents patients from achieving sought-after reinforcement from opioid use and from resuming physical

dependence if prolonged opioid use is resumed. This blockade can have positive psychological consequences, including reduced craving related to reduced immediate availability of opioid effects and deconditioning of cue-related craving as patients are exposed to people, places, and things formerly associated with opioid use. The opioid antagonist properties of naltrexone do not appear to be subject to the development of tolerance because potency has been demonstrated after maintenance for 1 year or more. Naltrexone does not have opioid agonist properties, and patients do not obtain a psychoactive reward ("high") from using naltrexone. It is also not associated with significant withdrawal symptoms if use is discontinued. However, patients may be more sensitive to lower doses of opioids once naltrexone is discontinued due to the loss of tolerance to exogenous opiates.

In carefully selected patients who are motivated to take naltrexone, maintenance on naltrexone, an opioid antagonist, can be used as an alternative to maintenance treatment with opioid agonists, such as methadone or the partial agonist buprenorphine. As a nonscheduled medication without psychoactive or addictive properties, naltrexone can be administered as a maintenance treatment in a range of medical settings without any special licensing requirements. Because diversion is not an issue, take-home medications can be provided. As a drug without agonist effects that blocks the effects of opioids, naltrexone can be used only after completion of a 7- to 10-day period of complete opioid abstinence or after a rapid detoxification procedure to avoid precipitating opiate withdrawal. Hence, patients must start treatment with a goal of achieving total opioid abstinence and a willingness to undergo detoxification. Unlike patients maintained on opioid agonists, patients on naltrexone achieve no rewarding effects of taking medications and experience no withdrawal effects if doses are missed. Thus, there are no pharmacological barriers to dropping out of treatment. In addition, some patients continue to complain of at least mild side effects or opioid withdrawal effects throughout the first several weeks of treatment.

However, oral naltrexone has been associated with high rates of discontinuation. A Cochrane review of randomized controlled clinical trials of naltrexone maintenance treatments for opiate addiction suggested that oral naltrexone was not more effective than placebo or no pharmacological agent in preventing relapse mainly due to low adherence and retention rates. Greater success has been achieved with patients motivated by powerful contingencies such as opioid-dependent health care professionals for whom opioid abstinence is a condition for regaining medical licensure or incarcerated addicts for whom treatment is a requirement for probation or parole. Other strategies to enhance compliance include the use of voucher-based contingency management and the dispensing of naltrexone that is directly observed by either a medical professional or a significant

other. The importance of retention in therapy for opiate addiction is highlighted by a review of coronial data in Australia in which opiate agonist treatment was associated with lower opiate overdose deaths compared to oral naltrexone, likely due to better retention in treatment with agonist therapy.

Extended-release formulations, both injectable and surgically implantable, address some of the shortcomings of the oral naltrexone related to adherence and retention. The extended-release injectable formulation of naltrexone (XR-NTX) was approved by FDA in 2010 to treat and prevent relapse after patients with opioid dependence have undergone detoxification treatment. A double-blind placebo-controlled multicenter trial in Russia showed that XR-NTX in conjunction with psychosocial treatment is efficacious in preventing relapse on opiates in detoxified patients. The median proportion of weeks of abstinence was 90 percent in 380-mg extended-release naltrexone group versus 34 percent in the placebo group. In addition, treatment retention was much higher and opioid craving was much lower in the active medication group. The 1-year open-label extension phase of this study showed that 62.3 percent clients remained compliant out of which 50.9 percent patients remained abstinent from opiate use and the safety of use of XR-NTX is further established during this phase. A recent study demonstrated the usefulness of XR-NTX as a therapeutic option in the treatment of opioid dependence in individuals involved with the criminal justice system. XR-NTX patients were shown to have lower health care utilization costs and longer periods of medication refills than patients taking oral naltrexone or acamprosate. Opioid-dependent patients maintained on XR-NTX were shown to have both lower inpatient substance abuse–related utilization as well as total cost of treatment as compared to methadone. Newer sustained-release formulations have been developed such as naltrexone implants but further research is needed to establish their safety and efficacy in clinical use.

XR-NTX appears to be well tolerated and concerns about hepatotoxicity have not been realized. No differences have been found between XR-NTX and placebo on the proportion of patients showing elevations in liver enzymes and XR-NTX has been found to be safe in patients with mild to moderate chronic HCV and/or HIV infections. Opioid detoxification is key prior to naltrexone induction to avoid precipitating opiate withdrawal. This can be done by a gradual opioid or buprenorphine taper or by more rapid discontinuation with the help of adjunctive nonopioid medications for symptomatic relief. The latter is more suitable for patients with low physiological dependence. For patients who are at risk of moderate to severe withdrawal symptoms, detoxification is best carried out in an inpatient detoxification setting or a partial hospital program with inpatient backup.

Opiate Detoxification Protocols in Combination with Naltrexone

As a method for rapid detoxification, outpatient protocols have been developed combining naltrexone with clonidine, a selective α2-adrenergic antagonist initially used as an antihypertensive agent, and other adjunctive medication. This combined treatment allows patients to be completely opioid free from the first day of treatment, to have completed detoxification within 48 to 72 hours, and to be able to commence naltrexone maintenance at that time. In addition, as a method of initiating maintenance naltrexone treatment, this mode of detoxification removes the need for an unprotected 7- to 10-day opioid-free period prior to starting naltrexone, a time when vulnerability to relapse is high. Because blood pressure may be substantially lowered by clonidine administration, patients must be monitored for up to 8 hours on the initial detoxification day, limiting this procedure to settings in which treatment rooms and adequate nursing staff are available. As this procedure may result in significant withdrawal symptoms, this method is best used in patients with low level of physiological dependence to opiates.

A variant method of "ultrarapid" inpatient detoxification involves administering naltrexone to opioid-dependent patients who are placed under light sedation or general anesthesia (including mechanical ventilation) followed by detoxification using full doses of naltrexone. This method was associated with more incidences of life-threatening adverse events and therefore this form of "ultrarapid" detoxification cannot be recommended at this time.

Opioid Overdose Prevention

Deaths from opioid overdoses have reached epidemic levels in the United States. Naloxone hydrochloride (Narcan) a potent opioid antagonist has been widely used in various health care settings and by emergency medical personnel to reverse opioid overdose and the associated respiratory depression. Naloxone can be administered by multiple routes in a hospital setting, but it is available to be administered intramuscularly or intranasally in the community. Some community-based overdose prevention programs have been distributing naloxone to patients at risk of overdose since 1996. Subsequently, several agencies have been training clinicians to screen patients for overdose risk, and to educate them and their kin in recognizing and responding to overdoses. With increase in these initiatives focusing on naloxone distribution, there have been reports of thousands of overdose rescues in the United States. Many states have passed legislations permitting physicians to prescribe or dispense Narcan to prevent or reverse opioid overdose and for bystanders to administer Narcan in the event of an opioid overdose.

Cocaine Use Disorders

Based on preclinical studies showing that opioid antagonists can attenuate many effects of cocaine, including acquisition and maintenance of self-administration, a randomized, placebo-controlled clinical trial evaluated its efficacy in a 50-mg/day dose alone and in combination with relapse prevention therapy. Naltrexone was associated with modestly improved cocaine use outcomes only when combined with relapse prevention therapy. However, trials in patients with cocaine and alcohol dependence found no benefit of naltrexone in reducing cocaine use. A prior open-label study and another double-blind placebo-controlled study found disulfiram to be superior to naltrexone for patients with combined cocaine and alcohol abuse. Hence, the evidence does not justify a recommendation for naltrexone use in cocaine dependence.

Nicotine Use Disorders

Several lines of evidence suggest that the opioid system may be involved in nicotine dependence. Nicotine administration is known to stimulate the release of endogenous opioids and to attenuate the perception of pain via opioid mechanisms. Clinical laboratory studies examining the effects of opiate antagonists, including naloxone and naltrexone, on smoking behavior have been inconclusive, with some studies showing reductions in smoking and others finding no effect. Prior clinical studies suggested that naltrexone may reduce urges or the desire to smoke and may also reduce weight gain following smoking cessation. The clinical utility of naltrexone as an aid for smoking cessation is supported by two randomized clinical trials, both of which tested the efficacy of naltrexone when used in combination with the nicotine patch for smoking cessation. One study found preliminary efficacy for 50-mg naltrexone primarily in women. The other trial, which combined naltrexone (25 to 100 mg/day) with a 21-mg nicotine patch, found some advantage for the 100-mg dose of naltrexone as an augmentation strategy for smoking cessation. However, a recent meta-analysis of opioid antagonists for smoking cessation concluded that naltrexone alone or as an adjunct to nicotine replacement has no beneficial effect in short- or long-term smoking abstinence. The two clinical trials mentioned above found significant reductions in weight gain following smoking cessation, suggesting that naltrexone may have clinical utility as an adjunctive treatment for smokers who are concerned about weight gain following smoking cessation.

Self-Injurious Behavior

Naltrexone has been proposed as a treatment for self-injurious behavior and for autistic behaviors in patients with mental retardation and autism. Clinically significant symptom reduction has been reported from numerous

open case studies and trials with small sample sizes. Larger studies using more rigorous experimental designs have failed to replicate these reports. Hence, naltrexone's utility for these symptoms is not established.

Eating Disorders/Obesity

Numerous studies in rodents have suggested an association between altered endogenous opioid activity and obesity and other eating disorders. Exogenous and endogenous opioid administration increases food intake, whereas opiate antagonists block the increase in food intake. The clinical data on the use of opiate antagonists in eating disorders is controversial, and the literature is replete with studies that demonstrate differential effects of these agents based on the type of eating disorder. For example, a laboratory study demonstrated that naloxone reduces the consumption of sweet, high-fat foods in both obese and lean binge eaters compared with obese and lean nonbingers suggesting a role of the opioid system in binge eating but not in obesity. In bulimia nervosa, a handful of small sample-controlled clinical trials have been conducted, with some, but not all, reporting reductions in binge symptomatology. With regard to anorexia nervosa, naloxone infusion produced a 10-fold increase in weight gain per week, an effect that was ended when the naloxone infusion was stopped in patients who were also concurrently treated with the antidepressant amitriptyline Therefore, the role for naltrexone and other opiate antagonists in the treatment of eating disorders remains unclear, and larger, more definitive, controlled clinical trials are needed.

Naltrexone, when used alone has failed to produce clinically significant weight loss in controlled studies. However, in translational studies the combination of naltrexone and bupropion (NB) produced a synergistic reduction of food consumption in rodents. This was further proven by studies in overweight and obese individuals in which weight loss produced by the combination of bupropion and naltrexone was shown to be greater than the sum of individual monotherapies. A large, multicenter, randomized, double-blind, placebo-controlled trial showed that NB is a useful therapeutic option for treatment of obesity. The mean change in body weight was –6.1 percent in the sustained-release naltrexone 32 mg plus sustained-release bupropion 360-mg/day group (NB32) and –5 percent in the sustained-release naltrexone 16 mg plus sustained-release bupropion 360-mg/day group (NB16) as compared to –1.3 percent in placebo. Similarly, 48 percent of participants assigned to NB32 group and 39 percent of participants assigned to NB16 group had greater than 5 percent weight loss, as compared to only 16 percent in the placebo group. Another large 56-week, randomized, double-blind, placebo-controlled study in obese or overweight participants with dyslipidemia and/or hypertension compared NB32 to placebo. 55.6 percent of the participants treated with NB32 had greater than 5 percent weight loss as compared to

17.5 percent in the placebo group. This study also showed that NB32 produced greater improvement in weight-related quality of life, control of eating, and various cardiometabolic risk markers without any increase in depression or suicidality. Another double-blind placebo-controlled study in overweight/obese type 2 diabetics showed that NB was associated with improvements in glycemic control and select cardiovascular risk factors along with promoting weight loss. FDA recently, in 2014, approved the combination of naltrexone hydrochloride and bupropion hydrochloride in an extended-release formulation for the treatment of obesity.

Naltrexone showed promise in reducing long-term smoking cessation-related weight gain in women. In recent pilot studies naltrexone has shown promise in arresting antipsychotic-associated weight gain in patients with a diagnosis of schizophrenia and schizoaffective disorder. It is hypothesized that naltrexone could help reduce food cravings in this population by interacting with the dopamine–opioid system.

Pathological Gambling

Opioid antagonists have the ability to decrease dopamine release in the nucleus accumbens through the disinhibition of GABAergic input to dopamine neurons in the ventral tegmental area. As such, it has been hypothesized that these agents might be effective in the treatment of impulse control disorders in general and pathological gambling in particular. Two open-label case series and a moderately sized, placebo-controlled, randomized, double-blind study supported the efficacy of naltrexone in the short-term treatment of pathological gambling. In these studies, naltrexone was used at higher doses than generally used in the treatment of alcohol or opiate dependence. In the placebo-controlled, double-blind study, for example, the dose of 250 mg/day was targeted, and completers had an end-of-study dose of 188 mg/day. Consistent with the use of these higher doses, adverse effects were not infrequently reported, the most significant of which included elevated liver function tests (LFTs). An examination of the individuals experiencing naltrexone-induced increases in LFTs suggested that concurrent use of nonsteroidal anti-inflammatory drugs (ibuprofen, aspirin) was associated with LFT increase. A double-blind placebo-controlled trial demonstrated that 50, 100, and 150 mg doses of naltrexone were equally efficacious in reducing gambling urges and behavior in pathological gambling.

There has been more research examining the use of nalmefene for the treatment of pathological gambling. Two multicenter investigations suggested that the use of 25 mg of nalmefene daily may effectively reduce the severity of gambling in those with pathological gambling disorder; there was 59 percent improvement in those who received nalmefene versus 34 percent in those who received placebo. Higher doses of nalmefene (50 and 100 mg/day) were associated with intolerable side effects. Another

double-blind placebo-controlled showed that at doses of 20 and 40 mg, nalmefene was not superior to placebo in the treatment of pathological gambling.

Currently there is no FDA-approved medication for treatment of gambling disorder. A recent meta-analysis investigating the efficacy of pharmacological treatments found that opiate antagonists have small but significant benefit than placebo in treatment of pathological gambling. However, the study did not support the use of the opiate antagonists at this time citing that some of the studies have not adhered to intent-to-treat principles. Further studies are needed to examine the safety and efficacy of nalmefene and naltrexone in the treatment of pathological gambling, especially in individuals with co-occurring substance use disorders and also in combination with nonpharmacological treatment modalities.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Although this section is written specifically for naltrexone, the effects of nalmefene are similar except where noted.

Respiratory System

In opiate-dependent individuals, the administration of naltrexone has been shown to produce a variety of nasal symptoms, such as nasal congestion, sneezing, rhinorrhea, sore throat, cough, and shortness of breath. In clinical trials with the injectable formulation Vivitrol, one diagnosed case and one suspected case of eosinophilic pneumonia were observed.

Heart

Naltrexone has increased blood pressure in opiate-dependent individuals, and it can also produce palpitations and tachycardia. Naltrexone has also been shown to decrease fibrinogen levels in opiate addicts.

Smooth Muscles

Naltrexone has produced muscle tension, tremors, and twitching in opioid-dependent subjects. Incidences of back pain/stiffness, muscle cramps, and joint stiffness have also been reported with Vivitrol.

Liver

While naltrexone has not been shown to be hepatotoxic when administered in the recommended dose of 50 mg/day, hepatotoxic potential has been documented for a sixfold higher dose of 300 mg/day in obese individuals and in patients with Alzheimer disease. In an open-label safety study in which the typical dose was 50 mg, naltrexone was not associated with increased risk of hepatotoxicity in more than 500 alcohol-dependent patients treated for up to 12 weeks. However, in another trial, a daily dose

of 100 mg was associated with a small but statistically significant increase in the number of patients who experienced elevated liver function tests during a 16-week period. Thus, naltrexone is contraindicated in patients with acute hepatitis, liver failure, or significant liver disease. However, mild to moderate transaminase elevations associated with alcohol or drug use are not a contraindication. Similarly XR-NTX is reported to have lower likelihood of hepatotoxicity than previously suspected even in individuals with HIV and HCV co-infection. Moreover, naltrexone has also shown some utility in the treatment of pruritis observed in cholestatic patients.

Endocrine System

Changes in baseline levels of some hypothalamic, pituitary, and gonadal hormones have been reported with naltrexone. Naltrexone has been shown to produce significant increases in plasma luteinizing hormone (LH) and testosterone in opiate-dependent individuals and to reverse heroin-induced decreases in such individuals. In healthy adult men, naltrexone has been shown to increase plasma prolactin, LH, testosterone, adrenocorticotropin (ACTH), and cortisol levels.

Gastrointestinal Symptoms

Abdominal pain and cramps have been documented following naltrexone administration in both opiate-dependent and nondependent individuals. Some individuals also report feeling nauseated and/or vomiting and suffering loss of appetite, as well as diarrhea or constipation.

Carcinogenesis and Mutagenesis

Carcinogenicity studies with high doses in rats have documented a small increase in mesotheliomas and vascular tumors. No evidence of mutagenicity has been documented with naltrexone or nalmefene.

Pregnancy

Naltrexone, when administered at doses 140 times the therapeutic dose in humans, has been shown to decrease pregnancy rates in rats and to have an embryocidal effect in rats and rabbits. Doses of 100 mg per kg in rats and 60 mg per kg in rabbits have been shown to have similar effects. In preclinical studies, naltrexone has been shown to pass through the placenta to the fetus. There are no adequate or controlled studies of naltrexone in oral or injectable form in pregnant or nursing women. Similarly, preclinical reproduction studies with nalmefene indicate no evidence of impaired fertility or fetal harm. However, because there are no adequate trials in pregnant women, this drug should be used during pregnancy or in nursing women only if it is clearly needed.

Genitourinary

Less than 10 percent of individuals have reported increased frequency of or discomfort during urination, increased frequency of erections, or alteration in sexual interest.

Psychiatric/Neurological

Non–opiate-dependent individuals have reported feeling fatigued, tired, and listless following the administration of naltrexone. Some individuals have reported an increase in irritability, restlessness, mental confusion, and slowing following naltrexone administration. A recent meta-analyses of medications for alcohol use disorders found that both naltrexone and nalmefene have higher risk of withdrawal due to adverse events like dizziness, nausea and vomiting but the latter has a higher risk difference for incidence of dizziness as compared to placebo and also has higher rates of incidence of headache. Substance abusers have an increased risk of suicide, and this risk is not reduced by naltrexone.

Skin/Local Injection Site

Vivitrol (XR-NTX) injections can result in pain, tenderness, induration, swelling, erythema, bruising and pruritus at the injection site. Proper injection technique is essential as inadvertent subcutaneous administration may increase likelihood of severe injection site reactions.

General Problems

Some individuals have also reported blurred eyes and increased light sensitivity, as well as tinnitus.

THERAPEUTIC INDICATIONS

FDA-approved indications for naltrexone include the use of naltrexone as a maintenance treatment for detoxified patients with opioid dependence and to prevent relapse in patients with alcohol dependence. Naltrexone has established utility when combined with clonidine as part of outpatient detoxification protocols. Extended-release injectable formulation of naltrexone has been approved for maintenance treatment of both opiate and alcohol dependence. Fixed-dose combination of naltrexone and bupropion in an extended-release formulation has been approved for the treatment of obesity. Naloxone hydrochloride injection has been approved by FDA for opioid overdose reversal or for suspected acute opioid overdose.

PRECAUTIONS AND ADVERSE REACTIONS

One of the most important risks with this class of agents is the risk of precipitated withdrawal in those who have received high doses of opioids

for therapeutic purposes or in those who use opioids recreationally. As a result, patients should be abstinent from short-acting narcotics (e.g., heroin) for at least 5 to 7 days and from longer-acting opiates (e.g., methadone) for 10 days or more based on self-report and urine toxicology. If there is a question of occult opioid dependence in an individual who has a negative urine toxicology screen for opiates and is not showing signs of withdrawal, a naloxone challenge test is recommended in which the patient is observed for signs and symptoms of opiate withdrawal for 20 minutes following a subcutaneous injection of 0.8-mg naloxone. Otherwise, naltrexone is generally well tolerated, and the majority of patients who have no reaction to a test dose of naloxone report no side effects. The profile of reported side effects has considerable similarity to the symptoms of opioid withdrawal and includes gastrointestinal distress (nausea, vomiting, diarrhea, and abdominal pain), anxiety and restlessness, dysphoria, mild hypertension, headache, and insomnia. These side effects are most prominent in the first several days of use and improve rapidly for patients remaining in treatment. However, a substantial minority of those with a diagnosis of opiate dependence (5 to 10 percent) may refuse a second dose, and an additional 10 to 20 percent may discontinue medication in the first month of treatment, often citing continued discomfort reminiscent of opioid withdrawal throughout the first 1 to 3 weeks of treatment. Side effects can be minimized by ensuring an adequate interval between cessation of opioid use and initiation of naltrexone and by a graduated initial dose schedule starting with 12.5 or 25 mg daily for the first several days. A prior study which used naltrexone depot treatment showed that 32 percent of opioid-dependent patients discontinued treatment after the first injection. Six-month retention rates of oral naltrexone in opiate-dependent patients are generally less than 20 percent following induction. However, injectable naltrexone has been shown to have higher 6-month retention rates as compared to placebo in a recent large Russian study. Studies examining the medication refill persistence showed that injectable naltrexone is consistently associated with better adherence rates in alcohol-dependent patients as compared to the oral agents.

Patients should be advised that taking large quantities of exogenous opioids in an attempt to overcome the blockade provided by opioid antagonists can be extremely dangerous. Life-threatening opiate intoxication may result, including respiratory arrest or circulatory collapse. In addition, patients should be advised that once these agents are discontinued, they may be more sensitive to the effects of opioids due to loss of tolerance.

Among alcohol-dependent patients, nausea is the most common adverse reaction to naltrexone, can be mild to severe, and typically occurs within 90 minutes of administration. In a large multisite safety study, side

effects that occurred in more than 2 percent of patients included nausea (10 percent), headache (7 percent), dizziness (4 percent), nervousness (4 percent), fatigue (4 percent), insomnia (3 percent), vomiting (3 percent), anxiety (2 percent), and somnolence (2 percent). In a large clinical trial for alcohol dependence with a 100-mg daily dose of naltrexone a similar side-effect profile was observed. The effect of naltrexone on acute alcohol withdrawal has not been investigated. The rate of early termination due to adverse events has been somewhat higher in studies in which naltrexone was administered to actively drinking alcoholics. Women and younger patients appear to be at somewhat increased risk of side effects.

The potential for hepatotoxicity has been identified when naltrexone is administered at high doses. However, convincing reports of elevated liver function test results have been limited to patients treated for eating disorders with doses of 250 to 300 mg daily, which are five to six times higher than the recommended maintenance dose for opioid dependence. In a clinical trial of alcohol dependence with a 100-mg/day dose, a 2 percent rate of increases in liver enzymes (aspartate or alanine aminotransferases) to greater than five times the upper limit of normal was observed; this increase resolved on discontinuation of naltrexone. Therefore, as a precaution, patients should receive a full battery of liver function tests prior to receiving naltrexone, and naltrexone is contraindicated in patients with liver failure or acute hepatitis. The moderate transaminitis often seen in alcoholism and narcotic addition is not a contraindication; however, baseline bilirubin and serum transaminases and periodic monitoring of transaminases is indicated. If persistent elevations in liver enzymes occur, naltrexone should be discontinued. The possibility of medication noncompliance and continued alcohol or drug use should be ruled out when mild transaminase elevations (less than threefold) occur.

Because this class of agents is excreted through the kidneys and metabolized by the liver, caution should be used in patients with renal impairment or severe liver disease. Among patients with compensated and decompensated liver cirrhosis, a 5- to 10-fold increase in naltrexone levels has been found relative to individuals with normal liver function. The increased systemic availability of naltrexone, which is believed to be due to lower rates of conversion of naltrexone to 6β-naltrexol, appear to be related to the severity of liver disease.

Naltrexone has not been studied in pregnant or nursing women, and an embryocidal effect has been demonstrated in rats and rabbits when given dose approximately 140 times the human recommended dose. As a result, naltrexone should be used during pregnancy only when the potential benefit clearly outweighs the potential risks to the fetus. Similarly, there are no studies of nalmefene in pregnant women, and hence this agent should be used in pregnant women only when clearly needed.

DRUG INTERACTIONS

Because naltrexone and nalmefene are opioid antagonists, they block the pain-relieving effects of opioid analgesics. Therefore, in patients maintained on these agents, nonopioid pain medications or approaches should be tried first when possible. For elective surgery, cessation of oral naltrexone treatment 72 hours prior to surgery is recommended. Naltrexone does not reverse the effects of nerve root blocks or spinal anesthesia and does not antagonize agents used for general anesthesia. Intravenous nalmefene can be used to reverse respiratory depression produced by these agents without blocking the analgesic effects of subsequently administered opioids. In patients treated with opioid antagonists, if opioid medications are necessary for the treatment of severe pain, the blockade can be overridden with a rapidly acting analgesic, but treatment requires close monitoring in a setting equipped and staffed for cardiopulmonary resuscitation because the resulting respiratory depression may be deep and prolonged. Patients should be given a card identifying that they are on naltrexone to alert medical personnel in case of an emergency.

Recent trials of alcohol dependence have combined naltrexone and disulfiram without any significant increase in adverse effects or in hepatic dysfunction, and some clinicians use the two drugs concomitantly for brief periods of time. However, given the potential for hepatotoxicity with both these agents, frequent monitoring of hepatic function is recommended. Although formal drug interaction studies have not been conducted with antidepressant medications, the incidence of adverse events reported by patients who started naltrexone while on ongoing antidepressant therapy was similar to that of patients receiving naltrexone only in a large safety study. Simultaneous initiation of antidepressant therapy and naltrexone, however, has been found to result in higher rates of nausea than use of naltrexone alone. Increased lethargy and somnolence has been reported with the combined use of naltrexone and thioridazine. Combined use of high-dose naltrexone (e.g., 250 mg daily) with nonsteroidal anti-inflammatory agents (ibuprofen, aspirin) may be associated with elevations in liver function tests.

Intravenous nalmefene has been administered after benzodiazepines, inhalational anesthetics, muscle relaxants, and muscle relaxant antagonists administered in conjunction with general anesthetics without any adverse reactions. Care should be taken when using flumazenil and nalmefene together because both of these agents have been shown to induce seizures in preclinical studies.

LABORATORY INTERFERENCES

Thin-layer, gas-liquid, and high-pressure liquid chromatographic methods

used for the detection of opiates in the urine are not interfered with naltrexone. The potential for a false-positive urine for opiates using less specific urine screens such as enzyme multiplied immunoassay technique (EMIT) may exist, given that naltrexone and nalmefene are derivatives of oxymorphone.

DOSAGE AND ADMINISTRATION

Naltrexone is prepared in scored, 50-mg, capsule-shaped tablets. The recommended dose for adults is 50 mg/day for alcoholism and narcotic addiction. Lower initial doses can be considered over the first few days for patients at increased risk of adverse effects, with the dose increased as tolerated. In a large multisite study of alcoholism treatment, effect sizes for the 100-mg dose (compared to placebo) were not larger than those observed in earlier studies with a 50-mg daily dose. Morning administration is recommended for improved medication compliance.

Although daily oral administration is clearly effective for the treatment of alcoholism, there is some evidence that targeting naltrexone use to the days when drinking is anticipated may also be useful for some patients. The injectable formulation Vivitrol is available in single-use kits that contain 380 mg of Vivitrol microspheres and a diluent. The microspheres need to be suspended in the diluent provided prior to administration, and this agent is administered as an intramuscular gluteal injection once every 4 weeks or once a month.

Naltrexone should not be administered until there is no reasonable possibility of opiate use within the past 7 to 10 days or unless the patient has completed a detoxification and abstinence has been verified with urine testing for opiates. If there is a question about continued opiate use, opiate-free status should be confirmed by a negative response to an IM naloxone test dose of 0.8 mg. If substantial opioid withdrawal symptoms do emerge, initiation of naltrexone should be delayed at least 1 additional day. Nalmefene is available as a sterile solution for intravenous, intramuscular, and subcutaneous administration in two concentrations, containing 100 μg/mL or 1.0 mg/mL of nalmefene. The dose of this agent is titrated to achieve the desired reversal of the opioid effects. The oral form of nalmefene is not available in the United States. Naloxone hydrochloride injection is available to be administered intravenously, intramuscularly, or subcutaneously. An initial dose of 0.4 mg to 2 mg can be administered in adults, and may be repeated at 2- to 3-minute intervals to achieve reversal of opioid effects. If reversal is not achieved after administering 10 mg of naloxone, the diagnosis of opioid overdose should be questioned. Although it is not yet FDA approved, naloxone hydrochloride is also available in intranasal preparations for opioid overdose reversal.

INVESTIGATIONAL DRUGS

Naltrexone and nalmefene are nonspecific opioid antagonists that have been shown to dose dependently bind to all three types of opioid receptors. The highest antagonistic potency has been found to be at the μ receptor, which is 5 to 10 times greater than that at the δ receptor and 13 to 33 times greater than that at the κ receptor. Opioid antagonists that have specific affinity of specific opioid receptor types have been developed, which have enormous use as selective tools to identify specific receptor types involved in specific pharmacological endpoints. Most of the new specific opioid antagonists are nonpeptides that can easily penetrate the blood–brain barrier and therefore can be administered via peripheral routes. All of these are in the different phases of drug testing.

References

Anton RF, Myrick H, Wright TM, et al. Gabapentin combined with naltrexone for the treatment of alcohol dependence. *Am J Psychiatry*. 2011;168(7):709–717.

*Anton RF, O'Malley SS, Ciraulo DA, et al. Combined pharmacotherapies and behavioral interventions for alcohol dependence: the COMBINE study: a randomized controlled trial. *JAMA*. 2006;295(17):2003–2017.

Apovian CM, Aronne L, Rubino D, et al. A randomized, phase 3 trial of naltrexone SR/bupropion SR on weight and obesity-related risk factors (COR-II). *Obesity (Silver Spring)*. 2013;21(5):935–943.

Bartley CA, Bloch MH. Meta-analysis: pharmacological treatment of pathological gambling. *Expert Rev Neurother*. 2013;13(8):887–894.

Clark AK, Wilder CM, Winstanley EL. A systematic review of community opioid overdose prevention and naloxone distribution programs. *J Addict Med*. 2014;8(3):153–163.

David SP, Chu IM, Lancaster T, Stead LF, Evins AE, Prochaska JJ. Systematic review and meta-analysis of opioid antagonists for smoking cessation. *BMJ Open*. 2014;4(3):e004393.

*Garbutt JC, Greenblatt AM, West SL, et al. Clinical and biological moderators of response to naltrexone in alcohol dependence: a systematic review of the evidence. *Addiction*. 2014;109(8):1274–1284.

Garbutt JC, Kranzler HR, O'Malley SS, et al. Efficacy and tolerability of long-acting injectable naltrexone for alcohol dependence: a randomized controlled trial. *JAMA*. 2005;293(13):1617–1625.

Gibson AE, Degenhardt LJ. Mortality related to pharmacotherapies for opioid dependence: a comparative analysis of coronial records. *Drug Alcohol Rev*. 2007;26(4):405–410.

Gonzalez JP, Brogden RN. Naltrexone. A review of its pharmacodynamic and pharmacokinetic properties and therapeutic efficacy in the management of opioid dependence. *Drugs*. 1998;35(3):192–213.

Gormez A, Rana F, Varghese S. Pharmacological interventions for self-injurious behaviour in adults with intellectual disabilities: abridged republication of a Cochrane systematic review. *J Psychopharmacol*. 2014;28(7):624–632.

Grant JE, Kim SW, Hartman BK. A double-blind, placebo-controlled study of the opiate antagonist naltrexone in the treatment of pathological gambling urges. *J Clin Psychiatry*.

2008;69(5):783–789.

Grant JE, Odlaug BL, Potenza MN, Hollander E, Kim SW. Nalmefene in the treatment of pathological gambling: multicentre, double-blind, placebo-controlled study. *Br J Psychiatry*. 2010;197(4):330–331.

*Greenway FL, Fujioka K, Plodkowski RA, et al. Effect of naltrexone plus bupropion on weight loss in overweight and obese adults (COR-I): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet*. 2010;376(9741):595–605.

Hartung DM, McCarty D, Fu R, Wiest K, Chalk M, Gastfriend DR. Extended-release naltrexone for alcohol and opioid dependence: a meta-analysis of healthcare utilization studies. *J Subst Abuse Treat*. 2014;47(2):113–121.

Hollander P, Gupta AK, Plodkowski R, et al. Effects of naltrexone sustained-release/bupropion sustained-release combination therapy on body weight and glycemic parameters in overweight and obese patients with type 2 diabetes. *Diabetes Care*. 2013;36(12):4022–4029.

*Jonas DE, Amick HR, Feltner C, et al. Pharmacotherapy for adults with alcohol use disorders in outpatient settings: a systematic review and meta-analysis. *JAMA*. 2014;311(18):1889–1900.

Jones JD, Comer SD, Kranzler HR. The pharmacogenetics of alcohol use disorder. *Alcoholism, clinical and experimental research*. 2015;39(3):391–402.

Krupitsky E, Nunes EV, Ling W, Gastfriend DR, Memisoglu A, Silverman BL. Injectable extended-release naltrexone (XR-NTX) for opioid dependence: long-term safety and effectiveness. *Addiction*. 2013;108(9):1628–1637.

*Krupitsky E, Nunes EV, Ling W, Illeperuma A, Gastfriend DR, Silverman BL. Injectable extended-release naltrexone for opioid dependence: a double-blind, placebo-controlled, multicentre randomised trial. *Lancet*. 2011;377(9776):1506–1513.

Krystal JH, Cramer JA, Krol WF, Kirk GF Rosenheck RA. Naltrexone in the treatment of alcohol dependence. *N Engl J Med*. 2001;345(24):1734–1739.

Lee JD, Friedmann PD, Boney TY, et al. Extended-release naltrexone to prevent relapse among opioid dependent, criminal justice system involved adults: rationale and design of a randomized controlled effectiveness trial. *Contemporary Clinical Trials*. 2015;41:110–117.

Lim JK, Bratberg JP, Davis CS, Green TC, Walley AY. Prescribe to prevent: overdose prevention and naloxone rescue kits for prescribers and pharmacists. *Journal of Addiction Medicine*. 2016;10(5):300–308.

Lobmaier PP, Kunøe N, Gossop M,Waal H. Naltrexone depot formulations for opioid and alcohol dependence: a systematic review. *CNS Neurosci Ther*. 2011;17(6):629–636.

Minozzi S, Amato L, Vecchi S, Davoli M, Kirchmayer U, Verster A. Oral naltrexone maintenance treatment for opioid dependence. *Cochrane Database Syst Rev*. 2011;(4): CD001333.

Mitchell MC, Memisoglu A, Silverman BL. Hepatic safety of injectable extended-release naltrexone in patients with chronic hepatitis C and HIV infection. *J Stud Alcohol Drugs*. 2012;73(6):991–997.

O'Malley SS, Corbin WR, Leeman RF, et al. Reduction of alcohol drinking in young adults by naltrexone: a double-blind, placebo-controlled, randomized clinical trial of efficacy and safety. *J Clin Psychiatry*. 2015;76(2):e207–e213.

O'Malley SS, Garbutt JC, Gastfriend DR, Dong Q, Kranzler HR. Efficacy of extended-release naltrexone in alcohol-dependent patients who are abstinent before treatment. *J Clin*

Psychopharmacol. 2007;27(5):507–512.

Pettinati HM, Kampman KM, Lynch KG, et al. A double blind, placebo-controlled trial that combines disulfiram and naltrexone for treating co-occurring cocaine and alcohol dependence. *Addict Behav*. 2008;33(5):651–667.

Sigmon SC, Bisaga A, Nunes EV, O'Connor PG, Kosten T, Woody G. Opioid detoxification and naltrexone induction strategies: recommendations for clinical practice. *Am J Drug Alcohol Abuse*. 2012;38(3):187–199.

Woody GE, Krupitsky E, Zvartau E. Antagonist models for relapse prevention and reducing HIV risk. *J Neuroimmune Pharmacol*. 2016;11(3):401–407.

▲ 34.26 Selective Serotonin-Norepinephrine Reuptake Inhibitors

MICHAEL E. THASE, M.D., AND PIERRE BLIER, M.D., PH.D.

The term *serotonin-norepinephrine reuptake inhibitor* (SNRI) describes a group of medications that have therapeutic effects that are presumably mediated by concomitant blockade of neuronal serotonin (5-HT) and norepinephrine (NE) uptake transporters. The SNRIs are also sometimes referred to as *dual reuptake inhibitors*, a broader functional class of antidepressant medications that includes tricyclic antidepressants (TCAs) such as clomipramine and, arguably, imipramine and amitriptyline. What distinguishes the SNRIs from TCAs is selectivity, which in this context refers to a relative lack of affinity for other receptors, especially muscarinic, histaminergic, and the families of α- and β-adrenergic receptors. This distinction is an important one because the SNRIs have a more favorable tolerability profile than the older TCA dual reuptake inhibitors.

Clinical interest in the SNRIs has largely stemmed from the hypothesis that direct modulation of these two monoamine systems conveys greater antidepressant effects than exerted by medications that selectively enhance only 5-HT or NE neurotransmission. Greater therapeutic benefits could result from an acceleration of postsynaptic adaptation to increased neuronal signaling, simultaneous activation of two pathways for intracellular signal transduction, additive effects on the activity of relevant neuronal factors such as brain-derived neurotrophic factor or, quite simply, broader coverage of depressive symptoms. Clinical evidence supporting this hypothesis first emerged in a pair of studies conducted by the Danish University Antidepressant Group, which found an advantage for the dual reuptake inhibitor clomipramine in comparison to the SSRIs, citalopram and paroxetine. A third report, which compared the results of a group of patients prospectively treated with the combination of the TCAs, desipramine and fluoxetine with a historical comparison group treated with desipramine alone, provided additional support. A meta-analysis of 25 inpatient studies comparing the efficacy of TCAs and SSRIs yielded the

strongest evidence. Specifically, although the TCAs were found to have a modest overall advantage, superiority versus SSRIs was almost entirely explained by the studies that used the TCAs that are considered to be dual reuptake inhibitors—clomipramine, amitriptyline, and imipramine.

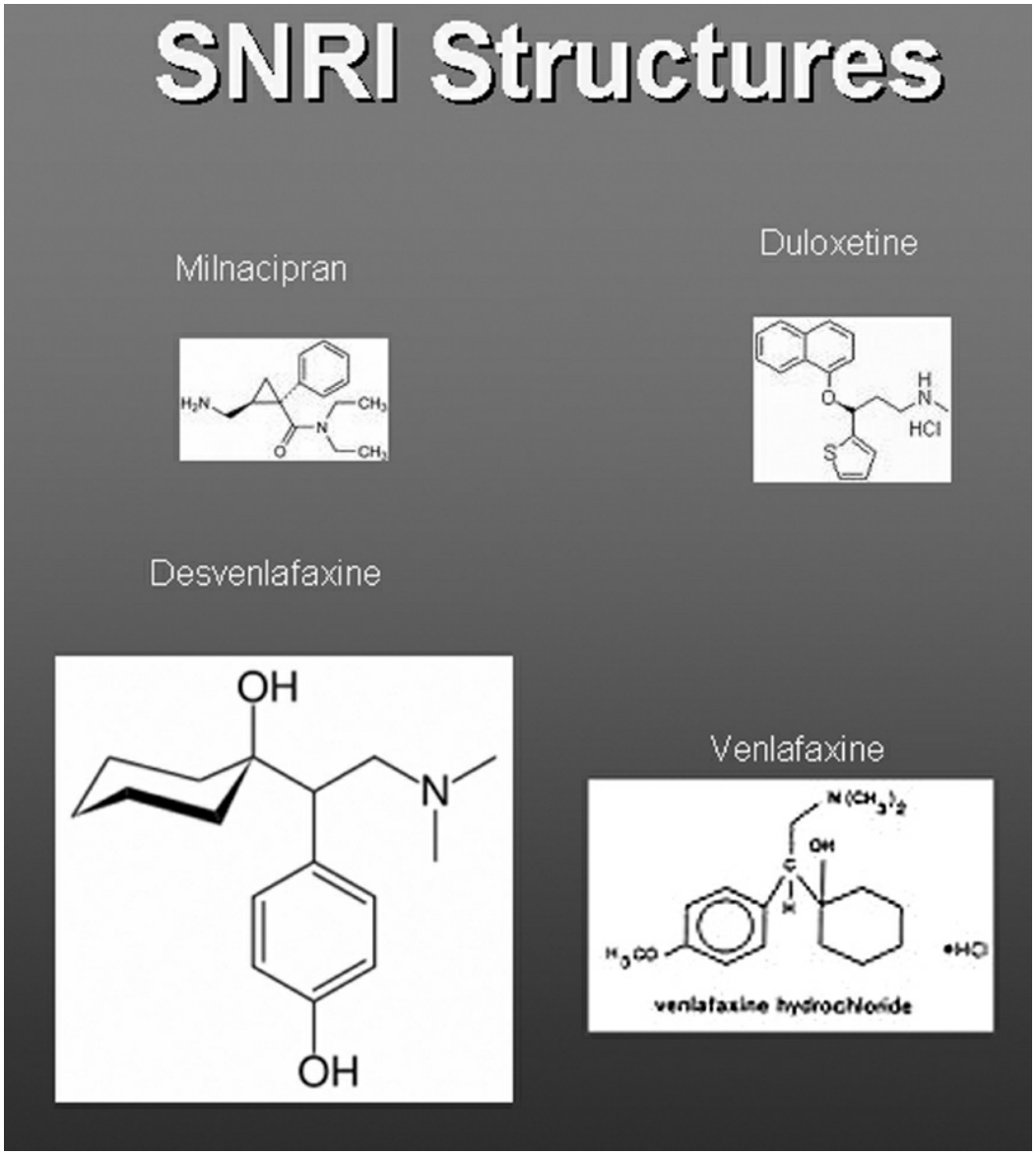

FIGURE 34.26–1. Chemical structures of the serotonin-norepinephrine reuptake inhibitors (SNRIs) venlafaxine, duloxetine, milnacipran, and desvenlafaxine.

There are currently five SNRIs approved for use in the United States: Venlafaxine (Effexor and Effexor XR), desvenlafaxine succinate (Pristiq), duloxetine (Cymbalta), milnacipran (Savella), and levomilnacipran (Fetzima). The racemic agent, milnacipran is indicated for major depressive disorder in a number of countries, including France and Japan, but in the United States only for fibromyalgia.

VENLAFAXINE AND DESVENLAFAXINE

Venlafaxine (Fig. 34.26–1) and desvenlafaxine are bicyclic phenylethylamine compounds that are structurally unrelated to other available antidepressants and anxiolytics. First synthesized in the late 1970s, venlafaxine was found to block uptake of 5-HT and NE in vitro, with a greater affinity for the 5-HT transporter. Unlike the TCAs, venlafaxine was found to have virtually no affinity for muscarinic, histaminergic, or α- or β-adrenergic postsynaptic receptors. Venlafaxine thus was predicted to exert antidepressant activity comparable to the prototypic dual reuptake inhibitors (i.e., amitriptyline and clomipramine), yet affording a safety and tolerability profile that was more similar to the SSRIs.

After efficacy was established versus placebo in the requisite phase II and phase III RCTs, the original immediate-release (IR) formulation of venlafaxine was approved by the FDA in 1994 for treatment of major depressive disorder. A microencapsulated extended-release (XR) formulation was introduced in 1997, which permitted once-daily dosing. Generic formulations of venlafaxine are now available. An extended-release formulation of the succinate salt of *O*-desmethylvenlafaxine (ODV), the principal active metabolite of venlafaxine (DVS; Pristiq), was approved by the FDA for treatment of major depressive disorder in early 2008.

Pharmacology

Venlafaxine and ODV inhibit neuronal uptake of 5-HT and, with significantly lower in vitro affinity, NE. In addition to a virtual lack of affinity for muscarinic, histamine 1 (H_1), and adrenergic receptors, neither the parent drug nor the metabolite affect sodium fast channel activity. Venlafaxine and ODV do not inhibit the activity of monoamine oxidases.

It is widely believed that the relative affinities of venlafaxine for inhibition of 5-HT and NE uptake underpin an ascending dose–response relationship. Specifically, venlafaxine is essentially an SSRI at the lowest therapeutic dose (i.e., 75 mg/day) and that noradrenergic effects increase progressively as the dose is advanced. Although supportive experimental and clinical data can be marshaled, this hypothesis cannot be tested definitively until a specific ^{11}C-ligand for NE transporters is developed so that in vivo studies correlating clinical response with NE transporter occupancy can be undertaken. Venlafaxine at higher doses has also been proposed to be a triple reuptake inhibitor on the basis of in vitro work documenting weak inhibition of dopamine reuptake. However, this effect is probably not clinically significant at usual therapeutic doses.

Orally administered venlafaxine and DVS are well absorbed from the GI tract and undergo extensive first-pass metabolism in the liver. During venlafaxine therapy, ODV accumulates in higher concentrations (typically about twofold higher) than the parent drug. It thus can be said that ODV accounts for a majority of the therapeutic activity of venlafaxine.

Coadministration with food decreases the rate, but not the extent, of absorption of the IR formulation and does not appreciably alter absorption of venlafaxine XR or ODV. Peak plasma concentrations are achieved within 2 to 3 hours after ingestion of the IR formulation of venlafaxine. During therapy with the XR formulation, peak plasma concentrations are achieved within 5.5 hours for venlafaxine and 9 hours for ODV. When compared to the initial formulation, therapy with venlafaxine XR results in lower peak plasma concentrations and a flatter plasma drug concentration curve when compared to venlafaxine IR and yet virtually identical bioavailability. During therapy with DVS, peak plasma concentrations of ODV are achieved within about 7 to 9 hours. Formulation of ODV as a succinate salt significantly increases the bioavailability of DVS as compared to venlafaxine ER.

Venlafaxine and ODV have linear kinetics over a dose range of 75 to 450 mg/day. Venlafaxine and ODV are primarily eliminated by the kidneys. The elimination half-lives of venlafaxine and its metabolite are short (i.e., 4 and 10 hours, respectively). Neither venlafaxine nor ODV is highly bound to plasma albumin (i.e., 25 to 30 percent protein binding at therapeutic concentrations).

Clearance of venlafaxine and ODV is reduced among patients with cirrhosis and severe renal disease, necessitating use of lower doses for such patients. Without data from formal studies of treatment of patients with severe hepatic or renal disease, a "half-dose–half-speed" approach is clinically reasonable. Adjustments in dosing on the basis of age alone typically are not necessary. However, a low starting dose and slow titration are recommended for treatment of frail or medically compromised older patients.

Therapeutic Indications

Venlafaxine is approved by the FDA for treatment of four therapeutic disorders: major depressive disorder, generalized anxiety disorder, social anxiety disorder, panic disorder, and attention deficit disorder. Major depressive disorder is currently the only FDA-approved indication for DVS.

Depression. The recommended starting dose of both formulations of venlafaxine, 75 mg/day, is the minimum effective dose for treatment of depression. It should be administered with meals, on a two- or three-times-a-day basis for the IR. The XR formulation is available in 37.5-, 75-, and 150-mg capsules. The minimum therapeutic dose of DVS is 50 mg/day and it is available in 50- and 100-mg tablets. Venlafaxine XR and DVS may be taken in the morning or evening, with or without food, as clinically indicated.

In ambulatory clinical trials of venlafaxine using flexible doses, a modal daily dose of 150 mg/day is typically observed, regardless of the

formulation used. When therapy at modest doses is ineffective, further increases up to 375 mg/day can be considered, as tolerated. However, the maximum recommended daily dose of the XR formulation is only 225 mg because of a lack of data on once-daily ingestion of higher dosages. Consistent with this general approach to therapy, higher mean doses are observed in studies of treatment-resistant depression. Despite some concerns that noradrenergic activity may exacerbate anxiety, efficacy has been established in studies of depressed people with significant concomitant symptoms of anxiety.

In placebo-controlled RCTs of DVS, doses ranging from 50 mg/day to 400 mg/day have shown efficacy, with essentially no evidence of greater efficacy at higher doses, although higher regimens were introduced rather abruptly.

Venlafaxine is one of the most extensively studied newer antidepressants. In published studies using active comparators, venlafaxine therapy has been shown to be at least as effective as treatment with the TCAs—imipramine, clomipramine, dothiepin, and amitriptyline (seven studies total). Research also suggests comparable overall efficacy with two other multiaction antidepressants, mirtazapine (Remeron) and bupropion (Wellbutrin), although the former drug was found to have a better effect on sleep and the latter drug was found to have fewer sexual side effects. As described later in the chapter, a pair of RCTs has contrasted venlafaxine XR and duloxetine.

For both conceptual and commercial reasons, comparisons of venlafaxine and various SSRIs have garnered considerable attention. Looking at published fluoxetine (Prozac)-controlled studies, the differences in remission rates have ranged from nil to as much as 30 percent favoring venlafaxine; in aggregate, the difference in effect sizes was statistically significant. Although there are fewer published studies using other SSRIs, venlafaxine therapy similarly has ranged from comparable to more effective than citalopram, fluvoxamine, paroxetine, and sertraline. Only two studies have directly compared escitalopram and venlafaxine: the SNRI showed no efficacy advantage in these trials and tolerability indices favored the SSRI.

As few comparative clinical trials have sufficient statistical power to distinguish between two effective antidepressants, meta-analyses are increasingly used to examine relative efficacy. One such meta-analysis used the entire set of double-blind RCTs comparing venlafaxine IR and XR versus SSRIs conducted by the manufacturer of venlafaxine through 2005, focusing primarily on remission rates after 8 weeks of therapy. In this analysis of more than 7,000 depressed patients, those treated with venlafaxine were about one-third more likely to achieve remission than patients treated with SSRIs (odds ratio = 1.31), with an absolute advantage in remission rates of about 6 percent. Inspection of the results of all studies comparing venlafaxine and SSRIs, including those conducted by other

sponsors, largely confirm this finding, although again the average difference is small (odds ratio = 1.26) and the numeric difference favored the SSRI in approximately 20 percent of the studies (Fig. 34.26–2).

Despite a noteworthy amount of comparative data, venlafaxine has not been studied extensively in bipolar and psychotic subtypes of major depressive disorder. There are only three randomized studies of bipolar depression, including one comparison versus the SSRI paroxetine (bipolar I only), one versus the SSRI sertraline and bupropion (bipolar I and bipolar II), and one versus lithium (bipolar II only). In the first two studies, which required that patients take concomitant mood stabilizers, there were no significant differences in efficacy, although venlafaxine therapy was associated with a higher rate of treatment-emergent affective switches (TEAS) than the comparators. In the third study which compared monotherapies, venlafaxine was significantly more effective than lithium and was not associated with an increased risk of TEAS.

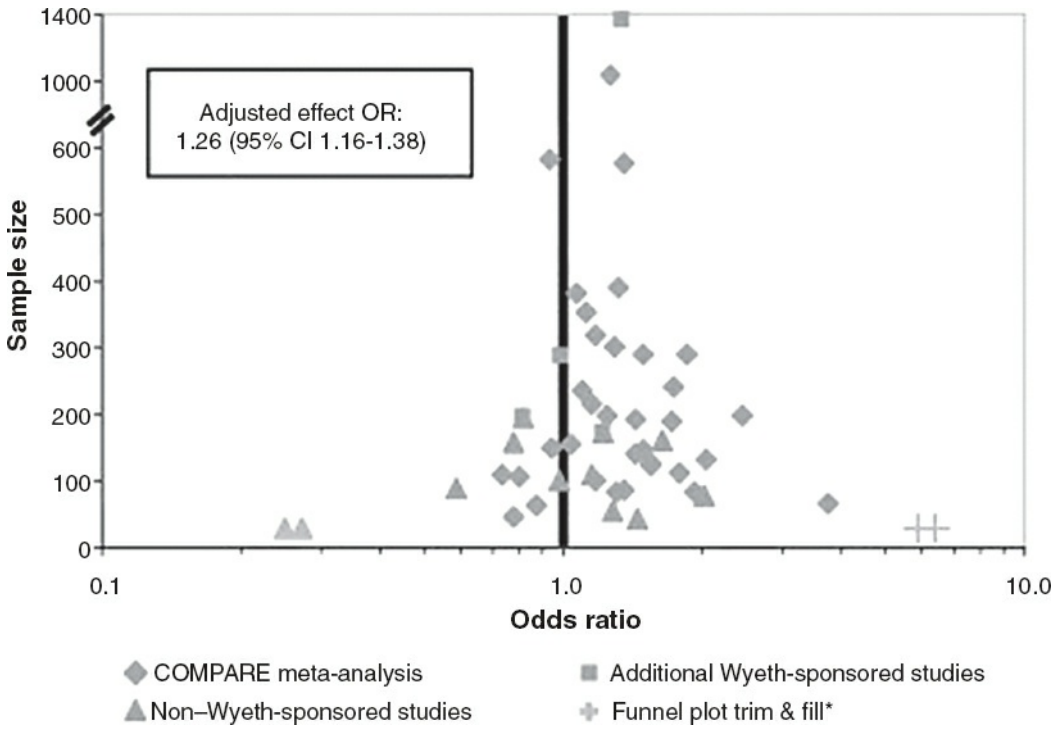

FIGURE 34.26–2. Funnel plot of randomized controlled trials comparing venlafaxine and selective serotonin reuptake inhibitors. CI, confidence interval. (From Nemeroff CB, Entsuah R, Benattia I, Demitrack M, Sloan DM, Thase ME. Comprehensive analysis of remission [COMPARE] with venlafaxine versus SSRIs. *Biol Psychiatry*. 2008;63:424–434.)

Two long-term studies using more methodologically rigorous placebo-controlled discontinuation designs have confirmed efficacy for prevention of relapse (i.e., 6 months after acute-phase treatment) and recurrence (12 and 24 months thereafter).

Generalized Anxiety Disorder. Venlafaxine XR received FDA approval

for treatment of generalized anxiety disorder in 1999. Efficacy and tolerability data for this indication are documented in four published randomized, controlled trials. In the one published study that used an active comparator, venlafaxine was found to be significantly more effective than buspirone (BuSpar) which, in turn, was not significantly more effective than placebo.

To date, higher-dose venlafaxine therapy for generalized anxiety disorder has not been studied systematically, nor have there been any studies using another antidepressant as the active comparator. The regulatory requirements of the first wave of double-blind studies also necessitated exclusion of patients with significant comorbid depression. Nevertheless, efficacy has been established in studies of anxious patients with a primary diagnosis of major depression.

Social Anxiety Disorder. The FDA approved venlafaxine for treatment of social anxiety disorder in early 2003. To date, the results of three double-blind, placebo-controlled studies have been presented at scientific meetings. These trials demonstrated efficacy and tolerability of venlafaxine at doses of 75 to 225 mg/day; both therapeutic effects and tolerability appear to be comparable to those of paroxetine.

Other Psychiatric Disorders. With respect to panic disorder, efficacy has been established in two RCTs, one including a 40 mg/day regimen of the SSRI paroxetine. In the latter study, 225 mg/day of venlafaxine outperformed 75 mg/day and the paroxetine treatment. Despite the obvious parallels with clomipramine, the antiobsessional effects of venlafaxine have not been as extensively studied. Two large RCTs have shown the efficacy of venlafaxine XR in PTSD, but it is an off-label use.

There are studies pertaining to treatment of various other mood and anxiety disorders, ADHD, and other psychiatric disorders. Premenstrual dysphoric disorder may be thought of as a gender-specific form of brief recurrent depression, and it is known to be responsive to SSRIs. A randomized, double-blind, placebo-controlled study evaluated flexible doses of venlafaxine therapy (range from 50 to 200 mg/day) across four menstrual cycles among 157 women with premenstrual dysphoric disorder. There was significantly greater improvement in the venlafaxine group compared to the placebo group across measures of emotion, function, physical symptoms, and pain.

Tolerability

Venlafaxine has a safety and tolerability profile that approaches that of the more widely prescribed SSRI class. For example, in the pooled data set of Michael Thase and colleagues, 9 percent of venlafaxine-treated patients withdrew because of adverse events, as compared to 7 percent of SSRI-

treated patients.

Nausea is the most frequently reported treatment-emergent adverse effect associated with venlafaxine therapy. The mechanism mediating this side effect is thought to be stimulation of central 5-HT type 3 (5-HT_3) receptors. With an incidence of as great as 35 percent in short-term placebo-controlled trials, nausea is approximately one-third more likely than seen with SSRIs. This is presumed to be because low–protein-binding results in relatively high concentrations of free (biologically active) venlafaxine in the CNS shortly after dosing. Nausea reliably diminishes in prevalence and severity during the first few weeks of therapy, presumably because of desensitization of 5-HT_3 receptors. Use of the XR formulation somewhat reduces the incidence and intensity of nausea, as does taking the medication with meals. Initiating therapy at lower dosages (37.5 mg/day) also may help attenuate nausea. When extremely problematic, treatment-induced nausea can be controlled by prescribing a selective 5-HT_3 antagonist or mirtazapine.

As with other 5-HT reuptake inhibitors, venlafaxine therapy is associated with sexual side effects, predominantly decreased libido, and a delay to orgasm or ejaculation. The incidence of these side effects is underreported in clinical trials and may exceed 30 to 40 percent when there is direct, detailed assessment of sexual function. Although the exact mechanisms of the sexual side effects are not known, it is thought that 5-HT–mediated inhibition of NE and dopamine effects is implicated.

Other common (reported by >10 percent of subjects in published studies) side effects include headache, insomnia, somnolence, dry mouth, dizziness, constipation, asthenia, sweating, and nervousness. Although several side effects are suggestive of anticholinergic effects, the drug has no affinity for muscarinic or nicotinic receptors. Thus, noradrenergic agonism is likely to be the culprit.

Hypertension

Higher-dose venlafaxine therapy is associated with an increased risk of sustained elevations of blood pressure. Experience with the IR formulation in studies of depressed patients indicated that sustained hypertension was dose related; increasing from 3 to 7 percent at doses of 100 to 300 mg/day and to 13 percent at doses greater than 300 mg/day. In this data set, venlafaxine therapy did not adversely affect blood pressure control of patients taking antihypertensives and actually lowered mean values of patients with elevated blood pressure readings before therapy. In controlled studies of the XR formulation, venlafaxine therapy resulted in only approximately a 1 percent greater risk of high blood pressure when compared to placebo. Arbitrarily capping the upper dose of venlafaxine used in these studies thus greatly attenuated concerns about elevated blood

pressure. When higher doses of the XR formulation are used, however, monitoring of blood pressure is recommended.

Discontinuation Syndrome

Venlafaxine is one of the antidepressants most commonly associated with a discontinuation syndrome. This syndrome is characterized by the appearance of a constellation of adverse effects during a rapid taper or abrupt cessation, including dizziness, dry mouth, insomnia, nausea, nervousness, sweating, anorexia, diarrhea, somnolence, and sensory disturbances. As the XR formulation does *not* affect the elimination half-life of the compound, it does not decrease the potential for discontinuation symptoms. Patients need to be informed that venlafaxine therapy should not be stopped abruptly and that even missing a few days of therapy may provoke discontinuation symptoms. It is recommended that, whenever possible, a slow taper schedule should be used when longer-term treatment must be stopped (e.g., reducing the daily dose by no more than 75 mg each week). On occasion, substituting a few doses of fluoxetine may help to bridge this transition, especially at the last step of discontinuation.

Overdose

There were no overdose fatalities in premarketing trials, although ECG changes (e.g., prolongation of QT interval, bundle branch block, QRS interval prolongation), tachycardia, bradycardia, hypotension, hypertension, coma, 5-HT syndrome, and seizures were reported. Fatal overdoses have been documented subsequently, typically involving venlafaxine ingestion in combination with other drugs or alcohol, or both. In an analysis of fatal antidepressant overdoses in the United Kingdom from 1993 to 1999, Nicholas A. Buckley and Peter R. McManus found that venlafaxine had a significantly greater fatal toxicity index (a measure of overdose deaths/million prescriptions) than the SSRIs (13.2 vs. 0.7 to 3.0). These data could raise some concerns about the first-line use of venlafaxine for treatment of suicidal patients. However, it should be noted that the authors did not take into account that psychiatrists wrote most of the venlafaxine prescriptions in the United Kingdom during the time in which these data were collected, whereas the SSRIs were widely prescribed by primary care providers. Thus, the higher fatality index may be describing more the patient profile treated with venlafaxine (i.e., inpatients or those with more chronic, severe, or treatment-resistant mood disorders), rather than the toxicity of the drug.

Drug–Drug Interactions

Venlafaxine is metabolized in the liver primarily by the cytochrome P450 (CYP) 2D6 isoenzyme. Because the parent drug and principal metabolite

are essentially equipotent, medications that inhibit this isoenzyme usually do not adversely affect therapy. In vitro and in vivo studies have shown venlafaxine to cause little or no inhibition of CYP 1A2, CYP 2C9, CYP 2C19, and CYP 3A4.

Venlafaxine is contraindicated in patients taking MAOIs because of the risk of a pharmacodynamic interaction (i.e., serotonin syndrome). An MAOI should not be started for at least 7 days after stopping venlafaxine. There are now considerable data from trials of augmentation of inadequate response to antidepressants by atypical antipsychotics with no apparent complications. Regarding the combination of venlafaxine with lithium (Eskalith) and anticonvulsants, clinical judgment should be exercised when combining medications given the paucity of data.

DULOXETINE

Duloxetine hydrochloride is a propanamine compound that was synthesized in the 1980s by the same drug discovery team that identified the SSRI fluoxetine (Fig. 34.26–1). Like fluoxetine, duloxetine is a potent 5-HT uptake inhibitor and lacks affinity for muscarinic, histaminic, and α- and β-adrenergic receptors. Unlike fluoxetine, duloxetine is also a relatively potent inhibitor of NE reuptake and in in vitro studies, duloxetine was found to be a substantially more potent NE reuptake inhibitor than venlafaxine. If these studies predict effects in the brain, duloxetine should be expected to produce dual reuptake inhibition at lower absolute doses than observed with venlafaxine. Duloxetine was approved by the FDA for treatment of major depressive disorder in 2004; other indications include generalized anxiety disorder, fibromyalgia, neuropathic pain associated with diabetic peripheral neuropathy, and chronic musculoskeletal pain.

Pharmacology

Orally administered duloxetine is well absorbed, and peak plasma levels are achieved approximately 3 hours after ingestion. Duloxetine has no active metabolites and exhibits linear pharmacokinetics within a dose range of 20 to 120 mg. Steady-state concentrations are achieved within 3 days of oral dosing, and the drug has an elimination half-life of approximately 12 hours (Fig. 34.26–3).

In the plasma, approximately 90 percent of the drug is protein bound. After hepatic oxidation, metabolites of duloxetine are primarily eliminated by the kidneys. Duloxetine is metabolized by CYP 1A2 and is a moderately potent inhibitor of the CYP 2D6 isoenzyme.

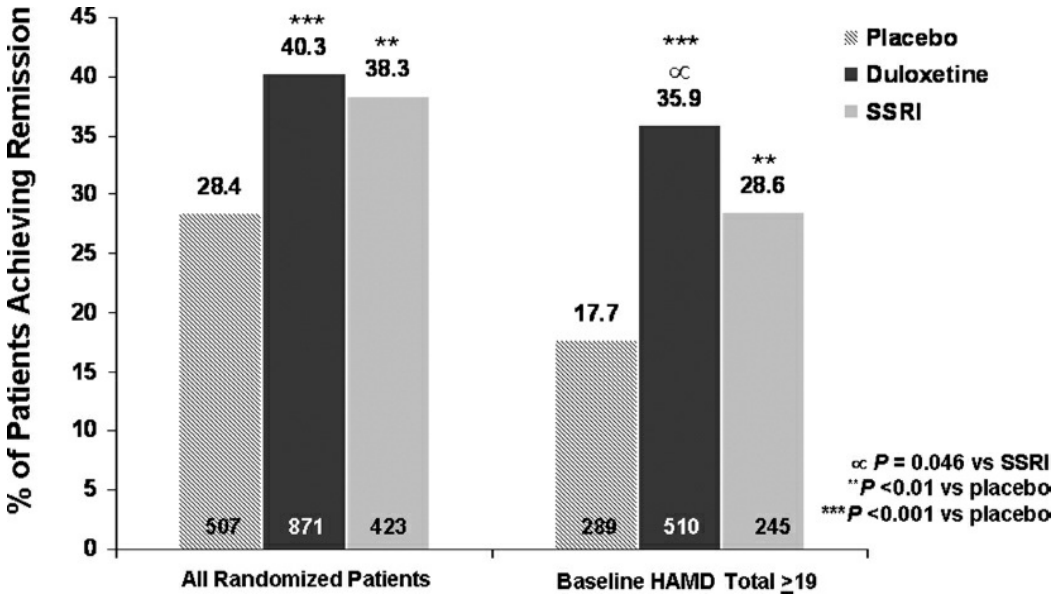

FIGURE 34.26–3. Meta-analytic comparison of duloxetine (40 to 120 mg/day) and two selective serotonin reuptake inhibitors (SSRIs; fluoxetine and paroxetine; 20 mg/day) in six placebo-controlled studies of major depressive disorder. Remission defined as Hamilton Depression (HAMD) Rating Scale Score <7 at end point. Both antidepressants were more effective than placebo overall, with a significant advantage for duloxetine over the SSRIs among the subset of patients scoring 19 or higher on the 17-item HAMD. (From Thase ME, Pritchett YL, Ossanna MJ, Swindle RW, Xu J, Detke MJ. Efficacy of duloxetine and selective serotonin reuptake inhibitors: comparisons as assessed by remission rates in patients with major depressive disorder. *J Clin Psychopharmacol*. 2007;27:672–676.)

Antidepressant Efficacy

Eight phase II and phase III placebo-controlled, randomized, controlled trials were completed to establish efficacy and safety; these studies formed the basis for the FDA's approval of duloxetine therapy of major depressive disorder. No evidence of a dose–response relationship was observed between 60 mg/day and 120 mg/day; somewhat weaker antidepressant effects were observed at 40 mg/day. These clinical results are fully consistent with the observations that 40 mg of duloxetine reached the 80 percent occupancy level of the 5-HT transporter in the brain, achieved with the minimal effective doses of SSRIs and venlafaxine, but that 60 mg once daily was necessary to maintain it. Studies using in vivo methods to assay NE transporter occupancy rates are needed to determine the extent to which NE uptake inhibition contributes to therapeutic efficacy at various doses.

Six of the initial registration studies of duloxetine (40 to 120 mg/day) included an SSRI as an active comparator. A meta-analysis confirmed that both duloxetine and the comparators were effective and, among the subset of patients with moderate-to-severe depressive symptoms, there was a significant advantage favoring the SNRI in remission rates (Fig. 34.26–3). However, none of these studies utilized the 60 mg/day dose of duloxetine and all of these studies used low (20 mg/day) doses of the comparators

(fluoxetine or paroxetine). Three studies subsequently compared duloxetine (60 mg/day) versus escitalopram (10 to 20 mg/day). No advantage in efficacy was observed for the SNRI and tolerability indices favored escitalopram. A pooled analysis of a pair of 6-week double-blind studies contrasting duloxetine (60 mg/day) and venlafaxine XR (150 mg/day) likewise found significant differences in efficacy. Duloxetine therapy was associated with a significantly higher incidence of nausea; venlafaxine therapy was associated with a higher incidence of treatment-emergent high blood pressure and, following withdrawal of study medication, discontinuation emergent symptoms.

Tolerability

The discontinuation rate due to adverse events in the studies summarized by Charles Nemeroff and colleagues was 15 percent in duloxetine-treated patients and 5 percent in placebo-treated patients. Side-effect attrition was slightly higher in the duloxetine groups in the studies that used fluoxetine (10 vs. 6 percent) and paroxetine (14 vs. 10 percent) comparison groups, although these differences were not statistically significant in individual studies. The most commonly reported side effects during duloxetine therapy were nausea (22 percent), dry mouth (16 percent), fatigue (11 percent), dizziness (11 percent), and somnolence (8 percent). Incidence of sexual dysfunction appears to be comparable to that of the SSRIs. Of note, the incidence of nausea was similar for duloxetine (dosed twice daily) to fluoxetine and paroxetine (dosed every day) in head-to-head double-blind trials. Despite the drug's potential usefulness for treatment of urinary incontinence, urinary hesitancy and urinary retention were not common occurrences in studies of depressed patients.

Like other 5-HT uptake inhibitors, duloxetine therapy has been associated with sexual dysfunction and, after abrupt cessation, a characteristic discontinuation syndrome. It is not yet possible to determine the relative risks of these phenomena during duloxetine therapy on a continuum with the other newer antidepressants.

Duloxetine therapy resulted in a small increase in resting pulse rate (approximately two beats per minute) and an average increase in blood pressure of approximately 2 mm Hg in the studies reviewed by Nemeroff and colleagues. Of note, the risk of sustained high blood pressure during duloxetine therapy was approximately 1 percent greater than observed in placebo-treated patients. Curiously, the risk of increased blood pressure does not appear to be dose dependent.

Although the CYP enzyme inhibition profile of the drug looks generally favorable, duloxetine levels are about 30 percent lower in smokers, likely as a result of CYP 1A2 induction, and increased fourfold by potent CYP 1A2 inhibitors, like fluvoxamine and some fluoroquinolone antibiotics. A lethal interaction with MAOIs is expected with duloxetine and the guidelines

suggested previously for venlafaxine therapy must be followed.

MILNACIPRAN AND LEVOMILNACIPRAN

The SNRI, milnacipran is available in Japan and several European countries for the treatment of major depressive disorder, but in the United States only for the treatment of fibromyalgia (Fig. 34.26–1). Milnacipran is distinguished as a dual reuptake inhibitor that on the basis of in vitro and in vivo studies may be viewed as the converse of that of duloxetine and venlafaxine, an NSRI rather than an SNRI. Specifically, milnacipran is several times more potent for inhibition of NE uptake than for 5-HT reuptake inhibition. Milnacipran thus can be thought of as a relatively selective NE uptake inhibitor at the minimum therapeutic dosage (50 mg/day) that exerts progressively greater inhibitory effects on 5-HT reuptake at higher doses (i.e., 200 mg/day, its maximal recommended dose for fibromyalgia). Like the other SNRIs, milnacipran has virtually no affinity for muscarinic, histaminic, or α- and β-noradrenergic receptors.

Pharmacology

Milnacipran has a half-life of approximately 8 hours and shows linear pharmacokinetics between doses of 50 and 250 mg/day. Metabolized in the liver, milnacipran has no active metabolites. Milnacipran is primarily excreted by the kidneys. Most orally administered milnacipran is bioavailable, and only 13 percent of milnacipran is bound to plasma protein. Milnacipran is not a potent inhibitor of any of the CYP isoenzymes.

Antidepressant Efficacy

The countries that have approved milnacipran for general use as an antidepressant have not required completion of a large number of placebo-controlled studies, and hence, efficacy is not as well established as is the case for venlafaxine or duloxetine. A meta-analysis of 16 RCTs (including all published and unpublished studies) encompassing 2,277 patients concluded that milnacipran did not have significantly different efficacy, acceptability, and tolerability when compared to other antidepressant agents. However, when compared to TCAs, fewer patients dropped out due to side effects on milnacipran.

Tolerability

In comparative clinical trials, milnacipran therapy was associated with significantly fewer anticholinergic and antihistaminic side effects than the tertiary amine TCAs, and fewer GI side effects than the SSRIs. Noteworthy side effects (i.e., incidence greater than an SSRI) included dizziness, sweating, and urinary hesitancy. Neither prospective data nor any expert opinion on the incidence of sexual dysfunction, especially when contrasted

to the SSRIs, could be located. Likewise, potential for lethality in overdose has not been systematically examined. Like 5-HT reuptake inhibitors, milnacipran should not be prescribed in proximity to an MAOI.

Characteristics of Levomilnacipran

Levomilnacipran is the active enantiomer of the racemic mixture of milnacipran. It is available in an extended release formulation and has a higher maximal plasma concentration (peaks in 6 to 8 hours), a greater area under the curve, and a longer half-life (12 hours) than its much less active dextrorotatory enantiomer. It is only 22 percent bound to plasma proteins and excreted by the kidneys 58 percent unchanged. It is recommended to initiate it at 20 mg/day for 2 days and gradually increase the dose to 40, 80, and 120 mg/day according to tolerability and response. It is effective in the 40 to 120 mg/day based on a data set of five placebo-controlled studies on 2,598 patients aged 18 to 80. Interestingly, it was effective in patients aged 60 and older, a population that is generally resistant to other medications. There are no as yet active comparator studies. In both acute and long-term studies, there was no evidence for weight gain. On average, there was approximately a 7 beats/min increase in hear rate, a thumbprint of NE reuptake inhibition. Only prospective studies using specific sexual function scales, and not merely treatment-emergent adverse events, will determine whether it is less problematic than other more potent 5-HT reuptake inhibiting medications. As for other SNRIs, it is recommended to utilize a gradual taper to discontinue the medication.

The SNRIs venlafaxine, desvenlafaxine, duloxetine, milnacipran, and levomilnacipran are effective antidepressants. There is evidence of an ascending dose–response relationship for venlafaxine and, possibly, milnacipran. A substantial body of research documents that venlafaxine therapy (at higher doses) has stronger antidepressant effects than the SSRIs; some data suggest a similar advantage for duloxetine and, for patients with severe depression, milnacipran. The efficacy of venlafaxine also has clearly been established in controlled studies of generalized anxiety, social anxiety, and panic disorders. Treatment with SNRIs is generally well tolerated. Tolerability and safety profiles are superior to the TCAs and approach that of the SSRIs. A higher initial incidence of nausea and, at higher doses, elevated blood pressure may be distinguishing characteristics of venlafaxine therapy. Apart from duloxetine, the four other medications have a low potential for CYP-mediated drug–drug interactions.

References

Allgulander C, Hackett D, Salinas E. Venlafaxine extended release (ER) in the treatment of generalised anxiety disorder: twenty-four-week placebo-controlled dose-ranging study. *Br J*

Psychiatry. 2001;179:15–22.

*Amsterdam JD, Wang CH, Shwarz M, Shults J. Venlafaxine versus lithium monotherapy of rapid and non-rapid cycling patients with bipolar II major depressive episode: a randomized, parallel group, open-label trial. *J Affect Disord*. 2009;112(1–3):219–230.

Anderson IM. SSRIs versus tricyclic antidepressants in depressed inpatients: a meta-analysis of efficacy and tolerability. *Depress Anxiety*. 1998;7:11–17.

Asnis GM, Henderson MA. Levomilnacipran for the treatment of major depressive disorder: a review. *Neuropsychiatr Dis Treat*. 2015;11:125–135.

Auclair AL, Martel JC, Assié MB, et al. Levomilnacipran (F2695), a norepinephrine-preferring SNRI: profile in vitro and in models of depression and anxiety. *Neuropharmacology*. 2013;70:338–347.

Buckley NA, McManus PR. Fatal toxicity of serotoninergic and other antidepressant drugs: analysis of United Kingdom mortality data. *BMJ*. 2002;325:1332–1333.

Bymaster FP, Dreshfield-Ahmad LJ, Threlkeld PG, et al. Comparative affinity of duloxetine and venlafaxine for serotonin and norepinephrine transporters in vitro and in vivo, human serotonin receptor subtypes, and other neuronal receptors. *Neuropsychopharmacology*. 2001;25:871–880.

Deecher DC, Beyer CE, Johnston G, et al. Desvenlafaxine succinate: A new serotonin and norepinephrine reuptake inhibitor. *J Pharmacol Exp Ther*. 2006;318:657–665.

Frampton JE, Plosker GL. Duloxetine: a review of its use in the treatment of major depressive disorder. *CNS Drugs*. 2007;21:581–609.

Guelfi JD, Ansseau M, Timmerman L, Kørsgaard S; Mirtazapine-Venlafaxine Study Group. Mirtazapine versus venlafaxine in hospitalized severely depressed patients with melancholic features. *J Clin Psychopharmacol*. 2001;21:425–431.

Katzman MA, Bleau P, Blier P, et al. Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. *BMC Psychiatry*. 2014;14 Suppl 1:S1.

Keller MB, Trivedi MH, Thase ME, et al. The Prevention of Recurrent Episodes of Depression with Venlafaxine for Two Years (PREVENT) Study: outcomes from the 2-year and combined maintenance phases. *J Clin Psychiatry*. 2007;68:1246–1256.

Kennedy SH, Eisfeld BS, Dickens SE, Bacchiochi JR, Bagby RM. Antidepressant-induced sexual dysfunction during treatment with moclobemide, paroxetine, sertraline, and venlafaxine. *J Clin Psychiatry*. 2000;61:276–281.

Lam RW, Andersen HF, Wade AG. Escitalopram and duloxetine in the treatment of major depressive disorder: a pooled analysis of two trials. *Int Clin Psychopharmacol*. 2008;23(4):181–187.

*Liebowitz MR, Manley AL, Padmanabhan SK, Ganguly R, Tummala R, Tourian KA. Efficacy, safety, and tolerability of desvenlafaxine 50 mg/day and 100 mg/day in outpatients with major depressive disorder. *Curr Med Res Opin*. 2008;24:1877–1890.

Montgomery SA, Baldwin DS, Blier P, et al. Which antidepressants have demonstrated superior efficacy? A review of the evidence. *Int Clin Psychopharmacol*. 2007;22:323–329.

*Nemeroff CB, Entsuah R, Benattia I, Demitrack M, Sloan DM, Thase ME. Comprehensive analysis of remission (COMPARE) with venlafaxine versus SSRIs. *Biol Psychiatry*.

2008;63:424–434.

Owens MJ, Krulewicz S, Simon JS, et al. Estimates of serotonin and norepinephrine transporter inhibition in depressed patients treated with paroxetine or venlafaxine. *Neuropsychopharmacology*. 2008;33:3201–3212.

Pae CU, Lim HK, Ajwani N, Lee C, Patkar AA. Extended-release formulation of venlafaxine in the treatment of post-traumatic stress disorder. *Expert Rev Neurother*. 2007;7:603–615.

Papakostas GI, Fava M. A meta-analysis of clinical trials comparing milnacipran, a serotonin–norepinephrine reuptake inhibitor, with a selective serotonin reuptake inhibitor for the treatment of major depressive disorder. *Eur Neuropsychopharmacol*. 2007;17:32–36.

Papakostas GI, Thase ME, Fava M, Nelson JC, Shelton RC. Are antidepressant drugs that combine serotonergic and noradrenergic mechanisms of action more effective than the selective serotonin reuptake inhibitors in treating major depressive disorder? A meta-analysis of studies of newer agents. *Biol Psychiatry*. 2007;62:1217–1227.

Perahia DG, Pritchett YL, Kajdasz DK, et al. A randomized, double-blind comparison of duloxetine and venlafaxine in the treatment of patients with major depressive disorder. *J Psychiatr Res*. 2008;42:22–34.

Pollack M, Mangano R, Entsuah R, Tzanis E, Simon NM, Zhang Y. A randomized controlled trial of venlafaxine ER and paroxetine in the treatment of outpatients with panic disorder. *Psychopharmacology (Berl)*. 2007;194:233–242.

*Rynn M, Russell J, Erickson J, et al. Efficacy and safety of duloxetine in the treatment of generalized anxiety disorder: a flexible-dose, progressive-titration, placebo-controlled trial. *Depress Anxiety*. 2008;25:182–189.

Smith D, Dempster C, Glanville J, Freemantle N, Anderson I. Efficacy and tolerability of venlafaxine compared with selective serotonin reuptake inhibitors and other antidepressants: a meta-analysis. *Br J Psychiatry*. 2002;180:396–404.

Smith T, Nicholson RA. Review of duloxetine in the management of diabetic peripheral neuropathic pain. *Vasc Health Risk Manag*. 2007;3:833–844.

Takano A, Suzuki K, Kosaka J, et al. A dose-finding study of duloxetine based on serotonin transporter occupancy. *Psychopharmacology (Berl)*. 2006;185:395–399.

Thase ME. Effects of venlafaxine on blood pressure: a meta-analysis of original data from 3,744 depressed patients. *J Clin Psychiatry*. 1998;59:502–508.

Thase ME, Entsuah AR, Rudolph RL. Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors. *Br J Psychiatry*. 2001;178:234–241.

Thase ME, Friedman ES, Howland RH. Venlafaxine and treatment-resistant depression. *Depress Anxiety*. 2000;12 Suppl 1:55–62.

Thase ME, Pritchett YL, Ossanna MJ, Swindle RW, Xu J, Detke MJ. Efficacy of duloxetine and selective serotonin reuptake inhibitors: comparisons as assessed by remission rates in patients with major depressive disorder. *J Clin Psychopharmacol*. 2007;27:672–676.

*Thase ME, Sloan DME. Venlafaxine. In: Schatzberg AF, Nemeroff CB, eds. *Textbook of Psychopharmacology*. 4th ed. Washington, DC: American Psychiatric Publishing, Inc.; 2009:465.

Thase ME, Tran PV, Wiltse C, Pangallo BA, Mallinckrodt C, Detke MJ. Cardiovascular profile of duloxetine, a dual reuptake inhibitor of serotonin and norepinephrine. *J Clin*

Psychopharmacol. 2005;25:132–140.

van den Broek WW, Mulder PG, van Os E, Birkenhäger TK, Pluijms E, Bruijn JA. Efficacy of venlafaxine compared with tricyclic antidepressants in depressive disorder: a meta-analysis. *J Psychopharmacol*. 2009;23:708–713.

Vieta E, Martinez-Arán A, Goikolea JM, et al. A randomized trial comparing paroxetine and venlafaxine in the treatment of bipolar depressed patients taking mood stabilizers. *J Clin Psychiatry*. 2002;63:508–512.

▲ 34.27 Selective Serotonin Reuptake Inhibitors

NORMAN SUSSMAN, M.D.

INTRODUCTION AND OVERVIEW

As a class, the selective serotonin reuptake inhibitors (SSRIs) have been the most widely used psychopharmacological agents for depression worldwide for nearly three decades. They are still extensively prescribed, in part because drugs of this class have been found to be effective in treating a range of disorders and symptoms other than depression. The most important of these are the anxiety disorders. From a historic perspective, the most significant impact of fluoxetine and subsequent SSRIs may have been the attention they generated in the popular press, which helped to reduce the longstanding stigma and fear of taking antidepressants. Patients became more amenable to come forward for treatment.

Clinicians also became less reluctant to prescribe antidepressants, because the side-effect profile of the SSRIs were devoid of side effects associated with the earlier antidepressant drugs—the tricyclics (TCAs) and MAOIs. These included dry mouth, constipation, sedation, orthostatic hypotension, and tachycardia. The ease of SSRI dosing led to clinicians other than psychiatrists prescribing SSRIs. Another very important consideration in the adoption of SSRIs as first-line treatments was their wide therapeutic index. Deaths from antidepressant overdoses plummeted after of SSRIs supplanted TCAs as the mainstays of treatment.

It needs to be kept in mind that the SSRIs have never been shown to be more effective than any other antidepressants. All of the SSRIs, starting with fluoxetine and followed by sertraline (Zoloft), paroxetine (Paxil), fluvoxamine (Luvox), citalopram (Celexa), and escitalopram (Lexapro), have been shown to be equally effective as antidepressants. Neither has the atypical antidepressant vortioxetine been found to produce higher response and remission rates than the SSRIs. There are, however, meaningful differences in pharmacodynamics and pharmacokinetics that might affect clinical responses among individual patients. In fact, some patients respond better to a particular SSRI than another.

SSRIs have proven more problematic in terms of side effects than the

original clinical trials suggested. In clinical practice, quality-of-life-associated adverse effects such as sexual dysfunction, weight gain, and emotional blunting often mitigate the therapeutic benefits of the SSRIs. These side effects were not captured in the short-term pivotal trials of these agents. It should also be noted that some patients report a diminished response or total loss of response to SSRIs with recurrence of depressive symptoms while remaining on a full dose of medication. The exact mechanism of this so-called "poop-out" is unknown, but the phenomenon is commonly seen in clinical practice.

Included in this section is vortioxetine (Brintellix), even though it is not an SSRI. Vortioxetine is classified as an atypical antidepressant, more specifically, a serotonin modulator and stimulator.

THERAPEUTIC INDICATIONS

The extensive use of SSRIs also derives from the fact that they are approved by the FDA for the following indications: Major depression, OCD, PTSD, PMDD, panic disorder, and social phobia (social anxiety disorder). Not all SSRIs have FDA approval for some of these indications, but it is commonly assumed that they in fact have a common spectrum of action.

Major Depression

Thousands of patients have participated in dose-ranging studies and active comparator trials of SSRIs. They may well be the most-studied drugs in the history of psychiatry in terms of the number of patients who have taken them in controlled clinical trials. The majority of studies included outpatients between 18 and 65 years of age with major depressive disorder of at least moderate severity (Hamilton depression [HAMD] Scale >20). These studies show that there is a wide range of doses effective in the treatment of moderate or severe depression.

Meta-analyses of studies that compared the SSRIs have found them to show equal effectiveness as antidepressants. In these analyses, about one-third of patients in short-term (6- to 8-week) placebo-controlled trials achieved full remission as determined by rating scales, about one-third had a partial response and one-third had no response.

One meta-analysis of SSRI showed the drugs had little benefit over placebo in mild and moderate depression but clinically significant benefits over placebo in patients with more severe depression. On the other hand, there have been numerous studies demonstrating the effectiveness of SSRIs in the treatment of dysthymia and other forms of chronic mild depression.

An FDA systematic review of all SSRI maintenance trials submitted to the agency between 1985 and 2012 found that maintenance treatment reduced the risk of relapse by 52 percent compared to patients treated with placebo.

Generalized Anxiety Disorder

Studies of SSRIs in the treatment of GAD have shown the effectiveness of these agents to be greater than the effectiveness of placebo in treating this condition. SSRIs have also been demonstrated to be effective in preventing relapse after the successful treatment of GAD. There may be continuing improvement over a 6-month period in those receiving SSRIs, suggesting the usefulness of ongoing treatment.

Obsessive-Compulsive Disorder

Placebo-controlled studies have demonstrated SSRIs to be efficacious for the treatment of OCD in adults and children. In some but not most cases, doses higher than the maximum recommended for depression are needed. Full response to treatment may take longer than is the case in depressive disorders.

Intravenous administration of citalopram has been studied as a treatment for OCD. Among patients with moderate to severe DSM-IV-diagnosed OCD who had failed at least two adequate treatment trials of SSRIs, excluding citalopram, about 60 percent of patients had a significant decrease in OCD symptoms after 3 weeks of intravenous citalopram.

Panic Disorder

SSRIs decrease the frequency of panic attacks by about 80 percent and decrease associated anticipatory anxiety. Sertraline results in an improvement of quality of life on most parameters. Previous unsuccessful treatment with benzodiazepines does not diminish their efficacy. The response rate may be lower for patients with more severe panic. Patients experience distinct withdrawal, expressed primarily as insomnia and dizziness, and a significant part of the relapse rate among the discontinued patients could possibly be accounted for by the withdrawal syndrome. When stopping SSRIs in patients with panic disorder, gradual discontinuation is recommended.

Social Phobia

SSRIs have been successfully used for the treatment of social phobia. Improvement may be achieved slowly, becoming evident during the second month of treatment and continuing to increase until the third month. The combination of SSRIs and cognitive-behavior therapy is more rapid and more complete than the use of an SSRI alone.

Premenstrual Dysphoric Disorder

SSRIs are effective in alleviating the symptoms of PMDD. Improvement may be seen during the first week of treatment. Mood, irritability, and

anxiety improve. Physical symptoms, such as swelling, bloating, and breast tenderness, are less responsive to treatment.

Posttraumatic Stress Disorder

Placebo-controlled studies have confirmed the efficacy of SSRIs for severe chronic PTSD. Discontinuation of the successful treatment after 6 months might result in the return of the PTSD symptoms in about half of patients.

Other Indications

SSRIs can be used for the treatment of premature ejaculation. In this case, paroxetine may be the most useful of these drugs. However, they require continuous daily treatment to delay ejaculation significantly.

PHARMACODYNAMICS

All SSRIs bind to and inhibit the action of serotonin transporters. This raises serotonin levels in the CNS and throughout the body. The SSRIs act not only on the 5-HT transporter in presynaptic neurons, but also at other sites, some well characterized, others not as clearly delineated. It is not clear whether the degree of SSRI selectivity offers any therapeutic benefit. Studies have failed to demonstrate differential efficacy of any SSRI in the treatment of psychiatric disorders. One must be cautious, however, in extrapolating group effects to those of an individual patient, given that studies have shown that patients who fail to respond to one may respond to another of these drugs.

Medications other than the SSRIs that demonstrate appreciable inhibition of the 5-HT transporter are clomipramine (Anafranil), duloxetine (Cymbalta), milnacipran (Savella available in the United States only for the treatment of fibromyalgia), levomilnacipran (Fetzima), and venlafaxine (Effexor). These agents also act through norepinephrine reuptake inhibition. SSRIs are essentially devoid of any clinically significant degree of receptor antagonistic properties at the muscarinic, cholinergic, H_1, M1-adrenergic, M2-adrenergic, β-adrenergic, 5-HT_1, and 5-HT_2 receptors. Sertraline is approximately 4-fold less potent than paroxetine and 60-fold less potent than amitriptyline at inhibiting [^{3}H]quinuclidinyl benzilate binding, which is a measure of muscarinic cholinergic receptor antagonism. However, paroxetine has the greatest affinity for muscarinic cholinergic receptors of the available SSRIs and therefore produces, in a small percentage of patients, dry mouth, constipation, blurred vision, or urinary hesitancy. However, the affinity of paroxetine for muscarinic receptors as measured by inhibition of [^{3}H]quinuclidinyl benzilate binding is 15-fold less potent than that of amitriptyline, and its anticholinergic properties may underlie its lack of side effects such as diarrhea and loose stools observed with other SSRIs

and its occasional sedating properties.

PET studies of the effect of antidepressants have been used for in vivo confirmation of in vitro and ex vivo derived hypotheses of mechanisms of action as well as for clinical dose finding. For example, one study used PET imaging to measure occupancies of citalopram, fluoxetine, sertraline, paroxetine, or extended-release venlafaxine at minimum therapeutic doses. It investigated the relationship between dose and occupancy. Striatal 5-HTT binding potential was measured before and after 4 weeks of medication administration. Healthy subjects received subtherapeutic doses; subjects with mood and anxiety disorders received therapeutic doses. Percent reduction in 5-HTT binding potential for each medication and dose was calculated. The findings suggested that 80 percent 5-HTT blockade is necessary for an antidepressant effect. A later study looked at a sample of patients in remission from depression, and found 5-HTT occupancy was significantly lower (61 to 74 percent, 95 percent confidence intervals) in the remitted patients.

THERAPEUTIC INDICATIONS

Although not all of the SSRIs have obtained the needed approval for treatment of each of these disorders, the spectrum of action of the SSRIs appears to be a class effect. The current approved indications for the SSRIs are shown in Table 34.27–1.

Table 34.27–1.
Approved Indications of the Selective Serotonin Reuptake Inhibitors in the United States for Adult and Pediatric Populations

	Citalopram (Celexa)	Escitalopram (Lexapro)	Fluoxetine (Prozac)	Fluvoxamine (Luvox)[a]	Paroxetine Paxil, (Pexeva)[b]	Sertraline (Zoloft)
Major depressive disorder	Adult	Adult	Adult and pediatric	—	Adult	Adult
Generalized anxiety disorder	—	Adult	—	—	Adult	—
Obsessive-compulsive disorder	—	—	Adult and pediatric	Adult and pediatric	Adult	Adult and pediatric
Panic disorder	—	—	Adult	—	Adult	Adult
Posttraumatic stress disorder	—	—	—	—	Adult	Adult
Social anxiety disorder	—	—	—	—	Adult	Adult
Bulimia nervosa	—	—	Adult	—	—	—
Premenstrual dysphoric disorder	—	—	Adult	—	Adult	Adult

[a]Available as immediate and controlled release.

[b]Available as immediate- and controlled-release paroxetine HCl as well as immediate-release HCl and mesylate salt.

The SSRIs, as a class, clearly demonstrate superiority in terms of safety when compared to the older antidepressant medications. SSRIs differ from TCAs is that they are devoid of cardiac toxicity. In addition, in contrast to the TCAs, MAOIs, and trazodone (Desyrel), the SSRIs are not antagonists at α-adrenergic receptors and therefore do not produce orthostatic hypotension. In general, with the exception of paroxetine, the SSRIs are relatively devoid of anticholinergic activity and do not demonstrate typical anticholinergic side effects, such as dry mouth, pupillary dilatation, constipation, urinary retention, or confusion. Even though paroxetine has more anticholinergic activity than the other SSRIs, it is only a fraction of that seen with the TCAs.

Each of the SSRIs is structurally unique (Fig. 34.27–1), although escitalopram is the *S*-enantiomer of the racemic mixture that is citalopram, and these differences are seen in the selectivity of reuptake inhibition, drug–drug interactions, and pharmacokinetic profiles. **Studies of the reuptake inhibitory potencies of the SSRIs in vitro have been performed. Although different studies show different values, in general, the K_i values tend to be similar and of the same relative ordering across studies. When interpreting in vitro data and attempting to extrapolate findings regarding relative degrees of reuptake inhibition to in vivo predictions, pharmacokinetic data such as steady-state concentrations, protein binding, molecular weights, and the presence or absence of active metabolites must all be accounted for.**

The combination of fluoxetine and olanzapine (Symbyax) is approved for the treatment of bipolar depression. All of the SSRIs, with the exception of fluvoxamine, were initially approved for the treatment of major depressive disorder. Fluvoxamine is unique in that it is only indicated for the treatment of OCD in the United States. This is a consequence of commercial considerations, not of any significant difference in its clinical profile. It shares the same therapeutic spectrum of activity with the other SSRIs and is widely used as a treatment of major depressive disorder in other countries worldwide.

EFFICACY

As already noted, innumerable clinical trials have failed to produce conclusive evidence that any SSRI is more effective or produces a greater degree or more rapid clinical improvement than any other agent in the class. Moreover, there are no proven differences in effectiveness between

SSRIs and other classes of antidepressants. Ultimately, treatment decisions should be based on considerations of relative patient acceptability, toxicity, and cost. There is also no compelling evidence that any one SSRI possess greater efficacy as an antidepressant or antianxiety treatment than another.

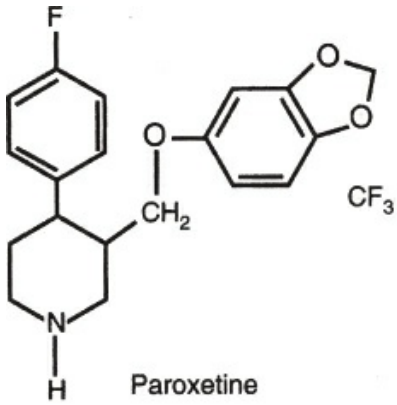

Paroxetine

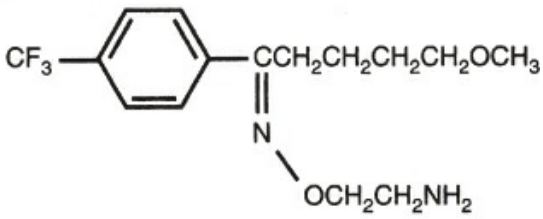

Fluvoxamine

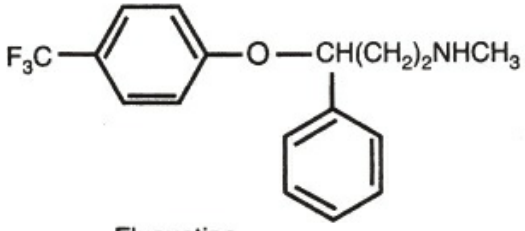

Fluoxetine

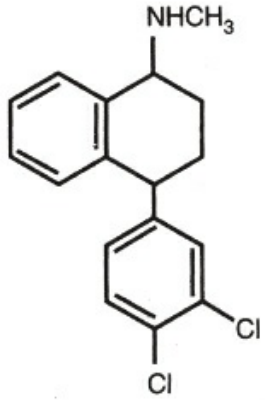

Sertraline

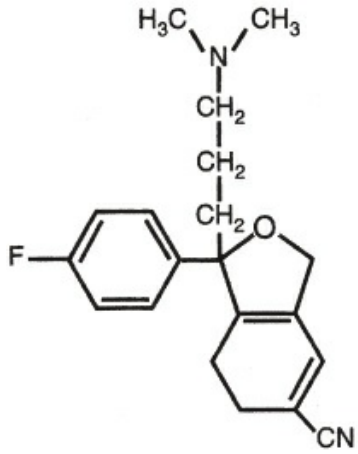

Citalopram

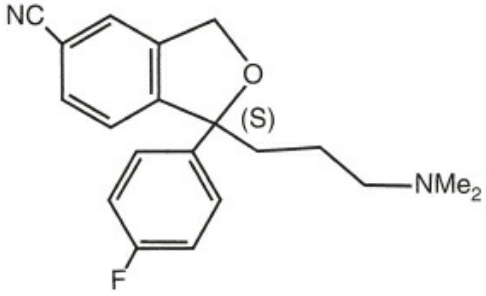

Escitalopram

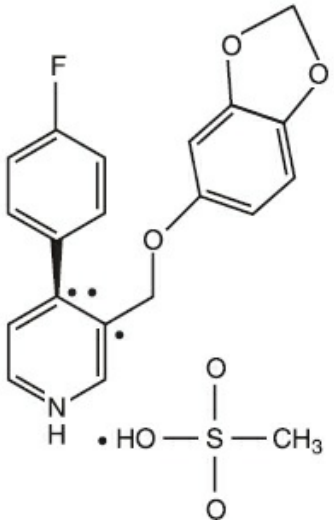

FIGURE 34.27–1. Structures of selective serotonin reuptake inhibitors.

In clinical trials for depression, the number of responders and remitters to the SSRIs are similar across clinical trials. About one-third of patients achieve response, one-third of patients achieve remission, and one-third show no response during an initial treatment period of 6 to 12 weeks. Some patients may demonstrate a better response to one of these agents than to another. The reason for this differential responsiveness is not known, and it is not possible to predict treatment response when initiating treatment.

TOLERABILITY

Both the therapeutic and adverse event profiles of SSRIs are the result of their blockade of serotonin transporter function. They raise serotonin levels in the CNS and throughout the body, with this action being particularly evident in the GI tract. As a group, the SSRIs possess the adverse effects listed in Table 34.27–2.

When reviewing product information for the SSRIs, both the frequency and severity of side effects need to be considered.

Gastrointestinal

Nausea occurs in at least 25 to 40 percent of patients and can range from mild discomfort to, in some patients, vomiting. It is the most common early side effect of SSRI, and most often diminishes during the first weeks of treatment. Nausea at the initiation of treatment is generally dose related. Accordingly, typical starting doses that are used to reduce this side effect are half of the usual starting dose. After 2 to 4 days at the lower doses, or when GI side effects resolve, the dose size is usually doubled. Flatulence may be another SSRI side effect. Unlike nausea, it tends to persist and prove socially embarrassing.

Table 34.27–2.
Selected Treatment-Emergent Adverse Events

Autonomic nervous system disorders
Dry mouth
Increased sweating
Central and peripheral nervous system disorders
Dizziness
Emotional blunting
Mild forgetfulness
Gastrointestinal disorders
Nausea
Diarrhea
Constipation
Indigestion
Abdominal pain

Psychiatric disorders
- Insomnia
- Somnolence
- Decreased appetite
- Decreased libido

Respiratory system disorders
- Rhinitis
- Sinusitis

Urogenital
- Ejaculation disorder
- Impotence
- Anorgasmia

General
- Influenza-like symptoms
- Fatigue
- Yawning
- Rash

Sexual Dysfunction

Sexual dysfunction, such as delayed ejaculation in men, delayed orgasm or anorgasmia in women, and diminished libido is a common side effect of 5-HT reuptake inhibition. Antidepressant-associated sexual dysfunction is the most frequent side-effect–associated cause of premature medication treatment discontinuation. Package insert data, almost always derived from phase III clinical trials, have underestimated the true incidence of sexual dysfunction, which is seen in clinical practice in nearly half of all treated patients. Reliable estimates of the incidence and severity of untoward experiences involving sexual desire, performance, and satisfaction are difficult to obtain, however, in part because patients and physicians may be reluctant to discuss them. Accordingly, estimates of the incidence of untoward sexual experience and performance cited in product labeling are likely to underestimate their actual incidence.

In some but not most instances, sexual dysfunction appears to be a dose-related effect, and occasionally decreasing the dose might be helpful. There have been a number of strategies suggested to deal with this important adverse effect, including decreased dose, drug holidays for shorter half-life SSRIs, or addition of buspirone, amantadine, yohimbine, cyproheptadine, stimulants, or bupropion. Anecdotally, selective inhibitors of cGMP-specific phosphodiesterase type 5 (PDE5), such as sildenafil (Viagra), have been reported by some patients to be helpful. One large study that used sildenafil treatment of sexual dysfunction in women taking SRRIs found that its use was associated with a reduction in adverse sexual effects and an improvement in treatment adherence rate. These findings need to be replicated. In fact, none of these approaches appears to be completely successful in reversing sexual side effects.

Sweating

Excessive sweating is not uncommon while being treated with SSRIs. The sweating is unrelated to ambient temperature. Nocturnal sweating may drench bed sheets and require a change of nightclothes. Terazosin, 1 or 2 mg per day, is often dramatically effective in counteracting sweating.

Discontinuation Syndrome

Drugs that have 5-HT reuptake inhibition have the potential to produce a discontinuation syndrome if stopped abruptly. Longer durations of treatment appear, at least in clinical observation, to be associated with a higher probability of discontinuation syndrome than do shorter treatment periods. Similarly, drugs with shorter half-lives are more likely to produce a discontinuation syndrome than those with longer half-lives, although it may also be the factor of discontinuation syndromes being less readily recognized in the latter group. In general, good clinical practice dictates a gradual tapering down of an antidepressant after a successful treatment.

The incidence of this syndrome is inversely proportional to the half-life of the compound. Typical symptoms include asthenia, anxiety, agitation, GI distress, myalgias, sometimes a flu-like syndrome but with the absence of a fever, disequilibrium, and, in extreme cases, an experience described as electrical shocks coursing through the arms and legs. This can often be prevented by advising patients to avoid abrupt discontinuation of the antidepressant (i.e., away for a weekend and neglecting to take the medicine bottle with them or skipping 2 or 3 days of treatment) and, when discontinuing the medication, to do so over the course of weeks rather than days. Drug tapers that become problematic as the last dose is stopped only to see an emergence of the discontinuation syndrome can often be accomplished by prescribing fluoxetine 20 mg per day for 2 weeks and then stopping that medication. Presumably, the longer half-lives of fluoxetine and norfluoxetine counter the discontinuation effects to a significant degree. It is not clear whether there is a discontinuation syndrome that emerges on the discontinuation of fluoxetine. If present, it is most likely delayed compared to that with the other SSRIs.

Although all SSRIs are associated with the possibility of a withdrawal syndrome, the onset is more immediate and more intense with some of the drugs. Decreasing doses every 2 to 3 days by small increments decreases the chances of encountering a discontinuation syndrome. Perhaps one of the more common scenarios in which this is encountered is with the patient who goes away for the weekend and neglects to take his or her prescription. Patient counseling is an important preventative measure in these cases. Occasionally, although rarely, there are patients who find it difficult to complete a taper off shorter–half-life drugs, such as paroxetine, sertraline, citalopram, escitalopram, or fluvoxamine. Anecdotal evidence suggests that a dosing of fluoxetine, 20 mg per day for 2 weeks, which is then stopped, may be helpful for those patients.

Bleeding

Both case reports and epidemiological studies confirm that SSRIs alone may increase bleeding risk in some patients. There have been numerous case reports and case series describing abnormal bruising, bleeding, and heavy menses in patients taking SSRIs. Serotonin is released from platelets following vascular injury. It promotes vasoconstriction and a change in the shape of the platelets that leads to aggregation. Platelets cannot themselves synthesize serotonin, so SSRIs deplete serotonin in the platelets, which may decrease coagulation, leading to excessive bleeding diathesis in some patients. Other physiological mechanisms may also be involved. There is an additive increase in bleeding risk when SSRIs are used concurrently with other drugs that can cause bleeding. SSRIs, for example, increase the incidence of serious bleeding in patients taking warfarin. Risk is also increased with concurrent use of nonsteroidal anti-inflammatory drugs, aspirin, or clopidogrel. A large general practice database search of an analysis data involving nearly 400,000 patients found that the number needed to harm for upper GI bleeding with SSRI treatment in a low-risk population was 3,177; and in a high-risk population it was 881. Thus, this risk associated with SSRI use is modest.

Hyponatremia

There is an increased risk of hyponatremia in patients taking SSRIs. Risk factors include older age, female gender, volume depletion, and use of other medications or medical conditions causing hyponatremia. SSRI-induced hyponatremia typically begins during the first weeks of treatment. In mild cases medication should be stopped (if possible) and water restriction started. The condition should normalize in about 2 weeks. Sodium replacement, if necessary, should not be too aggressive treatment (3 percent saline, with or without furosemide) unless the symptoms are acute or sodium is less than 110 mEq/L. Replacement should be no faster than 0.5 mEq/L per hour correction (to avoid the risk of central pontine myelinolysis). Once sodium reaches 120 mEq/L, there should be water restriction only. This side effect is reversed on discontinuation of the drug. This may be secondary to the syndrome of inappropriate antidiuretic hormone. Groups at risk include elderly patients, patients concomitantly treated with diuretics, and patients who were volume depleted.

Cardiovascular Effects

All SSRIs can lengthen the QT interval in otherwise healthy people and cause drug-induced long QT syndrome, possibly leading to torsade de pointes, especially when taken in overdose. The risk of QTc prolongation increases when an antidepressant and an antipsychotic are used in combination, an increasingly common practice. In 2011, the FDA has

issued a safety warning about high-dose citalopram treatment. It made the following recommendation regarding citalopram use.

- 20 mg/day is the maximum recommended dose for patients with hepatic impairment, who are older than 60 years of age, who are CYP2C19 poor metabolizers, or who are taking concomitant cimetidine.
- No longer prescribe at doses greater than 40 mg/day.
- Do not use in patients with congenital long QT syndrome.
- Correct hypokalemia and hypomagnesemia before administering citalopram.
- Monitor electrolytes monitored as clinically indicated.
- Consider more frequent EKGs in patients with CHF, bradyarrhythmias, or patients on concomitant medications that prolong the QT interval.

This was based on a QT study assessing the effects of 20- and 60-mg doses of citalopram on the QT interval in adults, compared with placebo, found a maximum mean prolongation in the individually corrected QT intervals were 8.5 milliseconds (ms) for 20-mg citalopram and 18.5 ms for 60 mg. For 40 mg, prolongation of the corrected QT interval was estimated to be 12.6 ms.

Comments

A 2013 study done at the Ann Arbor VAMC and the University of Michigan failed to confirm an increased risk for arrhythmias or death associated with daily doses of more than 40 mg. These findings do raise questions about the continued legitimacy of the FDA warning.

The study looked at the records of more than 600,000 VHA patients who received citalopram prescriptions between 2004 and 2009, as well as patient outcomes for more than 300,000 treated with, sertraline, which does not have an FDA warning.

According to study results, citalopram daily doses greater than 40 mg were actually associated with lower risks of ventricular arrhythmia (adjusted hazard ratio = 0.68, 95 percent CI = 0.61 to 0.76), all cause mortality (adjusted hazard ratio = 0.94, 95 percent CI = 0.90 to 0.99) and noncardiac mortality (adjusted hazard ratio = 0.90, 95 percent CI = 0.86 to 0.96) compared with daily doses of 1 to 20 mg. Overall, no increased risks of cardiac mortality were found, according to the researchers.

Physicians should consider whether the benefits of androgen deprivation therapy outweigh the potential risks in SSRI-treated patients with prostate cancer as reductions in androgen levels can cause QTc interval prolongation.

Dextromethorphan/Quinidine (Nuedexta) is available as a treatment for pseudobulbar affect, which is defined by involuntary, sudden, and frequent episodes of laughing and/or crying that are generally out of

proportion or inappropriate to the situation. Quinidine that can prolong the QT interval is a potent inhibitor of CYP2D6. It should not be used with other medications that prolong the QT interval and are metabolized by CYP2D6. This drug should be used with caution with any medications that can prolong the QT interval and inhibit CYP3A4, particularly in patients with cardiac disease.

Antepartum use of SSRIs is sometimes associated with QTc interval prolongation in exposed neonates. In a review of 52 newborns exposed to SSRIs in the immediate antepartum period and 52 matched control subjects, the mean QTc was significantly longer in the group of newborns exposed to antidepressants as compared with control subjects. Five (10 percent) newborns exposed to SSRIs had a markedly prolonged QTc interval (>460 ms) compared with none of the unexposed newborns. The longest QTc interval observed among exposed newborns was 543 ms. All of the drug-associated repolarization abnormalities normalized in subsequent electrocardiographic tracings.

Weight Gain

During early treatment, many patients lose weight. Over time, many patients regain this weight, and it is not uncommon for there to be a significant increase in weight over the following months. In some instances it becomes very difficult for patients to reverse this weight gain, even with rigorous dietary measures and exercise.

Diabetes

Population-based studies have shown that SSRIs may increase the risk of diabetes. An analysis was done based on three large well-known US datasets: the Nurses' Health Studies I and II and the Health Professionals Follow-up Study, a total of 139,000 women and 30,000 men. During over 1.6 million person-years of follow-up the analysis documented a total of 6,641 new cases of type 2 diabetes and found that self-reported antidepressant use was associated with an approximately 10 to 30 percent excess risk of developing diabetes. This association was robust to adjustment for age, ethnicity, marital and living status, smoking habits, alcohol intake, physical activity, family history of diabetes, diet, BMI and major comorbidities (e.g., hypertension and dyslipidemia), and their treatments. There needs to be additional evidence that this increased risk is due to associated factors or SSRI treatment, but these findings suggest that it is prudent to monitor patients periodically for changes in blood glucose levels.

Sleep Disturbances and Daytime Somnolence

Up to one-fourth of persons taking SSRIs note trouble sleeping, excessive

somnolence, or overwhelming fatigue. Bruxism is another side effect. Nocturnal motor activity, including nocturnal myoclonus, may be increased during SSRI treatment. Many persons taking SSRIs report recalling extremely vivid dreams or nightmares. They describe sleep as "busy." Some persons benefit from taking their SSRI dose before going to bed, but other prefer to take it in the morning. SSRI-induced insomnia can be treated with benzodiazepines, trazodone (Desyrel) (clinicians must explain the risk of priapism), or other sedating medicines. Significant SSRI-induced somnolence often requires switching to use of another SSRI or bupropion.

Emotional Blunting

Emotional blunting or indifference are largely overlooked side effects associated with chronic SSRI use. Patients report an inability to cry in response to emotional situations, a feeling of apathy or indifference, or a restriction in the intensity of emotional experiences. This side effect often leads to treatment discontinuation even when the drugs provide relief from depression or anxiety. The frequency of this syndrome is unknown since few empirical studies are available. It has a delayed onset, appears to be related to dosing, and may resolve with a dose reduction or discontinuation of the SSRI.

Yawning

Close clinical observation of patients taking SSRIs reveals excessive yawning as a side effect. This side effect is not well known by practitioners, so it is often overlooked. A 6-week, placebo-controlled, 446 subject trial of citalpram and placebo found 2 percent of participants in the citalopram arm developed yawning, as compared to <1 percent of patients taking placebo. Another study found that 7 percent of patients with OCD, 11 percent of patients with bulimia, and 1 percent of patients with panic disorder receiving fluoxetine experienced yawning as a treatment-emergent side effect of treatment. None of the placebo-treated patients reported this side effect. It should be noted that SSRI-related yawning is not a reflection of fatigue or poor nocturnal sleep but is postulated to be the result of SSRI effects on the hypothalamus. In rare instances, yawning is associated with spontaneous orgasms.

Extrapyramidal Symptoms

The SSRIs may rarely cause akathisia, dystonia, tremor, cogwheel rigidity, torticollis, opisthotonos, gait disorders, and bradykinesia. Rare cases of tardive dyskinesia have been reported. Some people with well-controlled Parkinson disease may experience acute worsening of their motor symptoms when they take SSRIs.

Laboratory Interferences

SSRIs do not interfere with any laboratory measures. Sertraline was found, in premarketing trials, to possess a weak uricosuric effect, with a mean decrease of uric acid by 7 percent.

CHEMISTRY AND PHARMACOKINETICS OF INDIVIDUAL SELECTIVE SEROTONIN REUPTAKE INHIBITORS

The SSRIs as a class exhibit a broad range of chemical structures and serum half-lives ranging from hours to days. There is no evidence that the structure of the drug molecules differentiate the clinical benefits and risks of the drugs in any meaningful way. The wide range of serum half-lives does manifest clinically. This is true not only in terms of time to steady state and the time required for the drug to exit the body after it is stopped, but also in differing potentials for discontinuation syndrome on stopping the SSRI or, more commonly, after missing occasional doses. The SSRIs demonstrate moderate to high protein binding. Table 34.27–3 summarizes the pharmacokinetic properties of the SSRIs. The contribution of protein binding to potential drug–drug interactions is somewhat unclear. The only fraction of drug that is available to exert biological activity is the free drug in the plasma. Protein-bound drug does not equilibrate into the CNS, but neither is it available to be metabolized. The compartment of free drug can enter the CNS but is also subject to being metabolized. SSRIs, such as paroxetine, sertraline, and fluoxetine, are tightly protein bound, and only approximately 1/20 of the drug in the system exists as free drug, the balance being nonspecifically bound to serum proteins. It remains unknown whether clinically significant displacement of protein binding by the SSRIs occurs during concomitant therapy with other medications, such as warfarin (Coumadin). It can be argued that clinicians should actively look for such drug–drug interactions following the addition of one drug with high protein-binding characteristics to another, especially when functional assays (such as with warfarin) are available. Another area that requires exploration is the determination of the level at which protein binding becomes saturated and accumulation of free drug becomes a nonlinear process. This may have important implications for agents that demonstrate different mechanisms of action across low-to-high dosing ranges. The chemical structures of the individual SSRIs are illustrated in Figure 34.27–1, and their pharmacokinetic profiles are shown in Table 34.27–3.

Table 34.27–3.
Pharmacokinetic Profiles of the Selective Serotonin Reuptake Inhibitors

	Time to Peak Plasma	Half-	Half-Life	Time to Steady State	Plasma Protein

Drug	Concentration (hrs)	Life	Metabolite	(days)	Binding (percent)
Citalopram (Celexa)	4	35 hrs	3 hrs	7	80
Escitalopram (Lexapro)	5	27–32 hrs	—	7	56
Fluoxetine (Prozac)	6–8	4–6 days	4–16 days	28–35	95
Fluvoxamine (Luvox)	3–8	15 hrs	—	5–7	80
Fluvoxamine CR (Luvox CR)					
Paroxetine (Paxil)	5–6	21 hrs	—	5–10	95
Paroxetine CR (Paxil CR)	6–10	15–20 hrs	—	<14	95
Paroxetine mesylate (Pexeva)		21 hrs			
Sertraline (Zoloft)	4.5–8.5	26 hrs	62–104 hrs	7	95

CR, controlled release.

FLUOXETINE

Chemistry

Fluoxetine's empirical formula is $C_{17}H_{18}F_3NO \cdot HCl$, and it has a molecular weight of 345.79. Its chemical designation is (±)-*N*-methyl-3-phenyl-3-[(α–α–α-trifluoro-*p*-toylyl)-oxy]-propylamine hydrochloride.

Pharmacological Actions

Absorption. Fluoxetine is well absorbed orally, with 72 to 90 percent systemic availability, reflecting the combination of almost complete absorption and first-pass metabolism by liver hepatocytes. Maximum plasma concentration is achieved within 6 to 8 hours of ingestion after a single 40-mg dose. Peak plasma concentrations are not affected by food, but absorption may be delayed 1 to 2 hours. The delayed-release, enteric-coated preparation of once-weekly fluoxetine also demonstrates peak plasma levels 6 to 8 hours after absorption, although absorption does not begin until entry into the portion of the GI tract where the pH exceeds 5.5. There is more variability of peak and trough levels with the once-weekly preparation compared to once-daily dosing. In addition, the C_{max} after administration of the 90-mg, once-weekly preparation 1 day after 20-mg, once-daily dosing was 1.7 times higher than the C_{max} observed with once-daily dosing. When the transition to once-weekly dosing was delayed by 1 week after the last once-daily dose was given, the observed C_{max} was approximately equal. Fluoxetine is available as pulvules (10, 20, or 40 mg),

tablets (10 mg), an oral concentrate (20 mg/5 mL), and in an enteric-coated, delayed-release, once-weekly form (90 mg).

Distribution. In vitro, fluoxetine is approximately 95 percent bound to serum proteins (albumin and α_1–acid-glycoproteins) at concentrations of 200 to 1,000 ng/mL. The volume of distribution for fluoxetine and norfluoxetine, the biologically active metabolite, ranges from 20 to 45 L/kg. The rate of plasma clearance is 20 and 9 L/hr for fluoxetine and norfluoxetine, respectively. Fluoxetine is widely distributed, including excretion into breast milk.

Metabolism and Elimination. Fluoxetine is a racemic mixture of *R*-fluoxetine and *S*-fluoxetine in equal proportions. Both are approximately equipotent in 5-HT reuptake inhibition activity, although the *S*-fluoxetine enantiomer is more slowly eliminated and is therefore the predominant form in plasma at steady state. Fluoxetine undergoes demethylation in the liver to norfluoxetine and to a number of other unidentified minor metabolites of fluoxetine, although norfluoxetine is the only metabolite identified that possesses biological activity. In animals, the *S*-enantiomer of norfluoxetine is essentially equivalent in activity to that of *R*- or *S*-fluoxetine, but *R*-norfluoxetine has significantly less activity. The kidneys excrete inactive metabolites produced by hepatic metabolism.

The elimination half-life of fluoxetine is 1 to 3 days after acute administration and 4 to 6 days after chronic administration. Norfluoxetine has an elimination half-life of 4 to 16 days, independent of duration of administration. Steady-state levels of fluoxetine are higher than would be predicted based on single-dose administration. This is likely due to saturation of the cytochrome P450 (CYP) 2D6 isozyme and subsequent metabolism by alternate nonsaturable pathways. Steady state is achieved owing to the presence of the nonsaturable pathways. Norfluoxetine, in contrast, demonstrates a linear dose-proportion relationship. The long half-lives of fluoxetine and norfluoxetine indicate the possibility of drug–drug pharmacokinetic interactions for a sustained period of time after the discontinuation of the drug.

Individuals who are deficient in CYP 2D6 activity—so-called poor metabolizers—constitute 6 to 8 percent of the population and metabolize *S*-fluoxetine at a slower-than-expected rate. In contrast, metabolism of *R*-fluoxetine is normal. Owing to the slower metabolism of *S*-fluoxetine, it is present at higher-than-expected concentrations, but *S*-norfluoxetine is found at lower concentrations. The sum total of *S*-fluoxetine and *R*-fluoxetine and *S*-norfluoxetine and *R*-norfluoxetine is approximately equal between poor metabolizers and those with a full complement of CYP 2D6 activity.

Impairment of hepatic function is associated with impaired metabolism

of fluoxetine. In a series of subjects with cirrhosis of the liver, the mean half-life of fluoxetine was increased to 7.6 days, compared to 2 to 3 days for healthy subjects. The half-life of norfluoxetine was similarly increased to a mean of 12 days, compared to 7 to 9 days for healthy subjects. This would indicate the necessity of dose reduction in individuals with significant liver disease.

Renal disease does not appear to impair the metabolism of fluoxetine or norfluoxetine in single-dose studies. One small study (N = 12) of patients on dialysis receiving fluoxetine at 20 mg per day for 2 months found steady-state fluoxetine and norfluoxetine levels close to those observed in patients with intact renal function.

Geriatric patients (60 years of age and older) who were depressed demonstrated fluoxetine levels of 209.3 ± 85.7 ng/mL after dosing of 20 mg per day for 6 weeks. Children with OCD or major depressive disorder (6 to less than 13 years of age; N = 10), compared to adolescents (13 to less than 18 years of age; N = 11), receiving fluoxetine at doses of 20 mg per day showed steady-state (60 days) fluoxetine levels of 171 and 86 ng/mL, respectively. Norfluoxetine levels were also higher in children than in adolescents, 195 and 113 ng/mL, respectively. These differences appear to be primarily related to relative body mass differences, and the levels normalized when this was taken into account.

SERTRALINE

Chemistry

The empirical formula of sertraline is $C_{17}H_{17}NCl_2 \cdot HCl$; it has a molecular weight of 342.7. The structure of sertraline is distinct from that of other 5-HT reuptake inhibitors, and it is a naphthylamine derivative. The chemical designation is (1*S*-*cis*)-4-(3,4-dichlorophenyl)-1,2,3,4-tetrahydro-*N*-methyl-1-naphthalenamine hydrochloride.

Pharmacological Actions

Absorption. Sertraline is well absorbed orally, with equivalent absorption observed with tablets and the oral concentrate. Maximum systemic concentration is achieved within 4.5 to 8.4 hours of ingestion after 14 days of dosing in the range of 50 to 200 mg daily. When administered with food, the mean peak plasma concentration is increased by 25 percent with the tablet preparation, and the peak plasma concentration is achieved in 5.5 hours after ingestion, compared to 8 hours when given without food. Whether this food-related increase in sertraline plasma levels is clinically significant in terms of adverse events is unclear, although many clinicians advocate the initial administration of sertraline, like other that of SSRIs, after a meal. Administration of the oral concentrate preparation of sertraline with food is associated with a delay in the time to peak plasma

concentration of from 5.9 to 7.0 hours.

Distribution. In vitro, sertraline is approximately 98 percent bound to serum proteins at concentrations of 20 to 500 ng/mL. The volume of distribution for sertraline is approximately 20 L/kg, and the rate of plasma clearance is 96 L/hr. The extent of distribution of sertraline into breast milk has been studied in detail. The data suggest that sertraline is present in breast milk and in infant serum, albeit in low concentrations.

Metabolism and Elimination. Sertraline undergoes demethylation in the liver to *N*-desmethylsertraline. The half-life of sertraline is 26 hours, and the half-life of *N*-desmethylsertraline is 62 to 104 hours; however, the metabolite is 5 to 10 times less potent in inhibiting 5-HT reuptake than the parent compound. Desmethylsertraline displays a five- to ninefold increase in area under the curve (AUC) and peak and minimum plasma concentrations after 14 days of sertraline administration. In healthy volunteers, sertraline demonstrates linear pharmacokinetics over the range of 50 to 200 mg per day. Based on the use of radiolabeled sertraline in two healthy male subjects, unchanged sertraline was found to constitute less than 5 percent of the plasma radioactivity. After 9 days, 40 to 45 percent of the radioactivity was recovered in urine, but no unchanged sertraline was detected. Another 40 to 45 percent was recovered in feces, of which 12 to 14 percent was unchanged sertraline.

The pharmacokinetics of sertraline in children (6 to 12 years of age) and adolescents (13 to 17 years of age) is similar to that seen in adults when adjustments are made based on body mass. Pediatric patients with major depressive disorder or OCD who were administered sertraline (42 days of sertraline therapy of which at least the final 11 days were at a dose of 200 mg per day) had mean C_{max} (maximal concentrations) and half-lives of 165 ng/mL and 26.2 hours and 123 ng/mL and 27.8 hours for children and adolescents, respectively. In comparison, a separate study involving healthy adults who received sertraline (200 mg per day for 30 days) had a mean C_{max} of 142 ng/mL and a half-life of 27.2 hours. Children and adolescents have a smaller (by approximately 22 percent) AUC with sertraline than adults, suggesting the potential for more rapid metabolism in the younger age group. Even so, initial dosing strategies should typically be adjusted downward for younger patients, who typically have a lower body mass than that of adults.

Liver disease is associated with impaired metabolism of sertraline. In a series of subjects with mild, stable cirrhosis of the liver, the mean half-life of sertraline was increased to 52 hours, compared to 22 hours for normal subjects. This suggests the necessity of dose reduction or increased intervals between dosings in individuals with hepatic impairment.

The presence of renal impairment (mild, moderate, or severe) does not

appear to affect the multiple-dose pharmacokinetics of sertraline. This is perhaps due to renal excretion providing a relatively minor pathway for the elimination of unchanged drug.

Aging is associated with a decrease in sertraline clearance. In 16 elderly patients treated with sertraline (100 mg per day for 14 days), the clearance of sertraline was 40 percent lower than that observed in younger subjects (25 to 32 years of age). The establishment of steady-state drug levels is delayed in this group to a time frame of 14 to 21 days of medication administration. Desmethylsertraline clearance was decreased in older men but not women.

PAROXETINE AND PAROXETINE CR

Chemistry

Paroxetine is the hydrochloride salt of a phenylpiperidine compound and is structurally distinct from other antidepressants. Paroxetine is designated as (–)-trans-4*R*-(4′-fluorophenyl)-3*S*-([3′,4′-methylenedioxyphenoxy]methyl) piperidine hydrochloride hemihydrate. The empirical formula is $C_{19}H_{20}FNO_3 \cdot HCl \cdot \frac{1}{2}H_2O$, and the molecular weight of the free base is 374.8. Paroxetine CR is the same salt as the immediate-release (IR) formulation but is in a degradable polymeric matrix with an enteric coating that delays drug release until the tablet has left the stomach.

Pharmacological Actions

Absorption. Paroxetine is well absorbed orally, and food does not alter absorption. Maximum systemic concentration is achieved approximately 5.2 hours and 6 to 10 hours after ingestion for the IR and CR formulations, respectively. The CR formulation demonstrates a decrease in bioavailability of approximately 25 percent compared to the IR preparation, and thus equivalent doses of paroxetine IR and CR would be 10.0- and 12.5-mg, 20.0- and 25.0-mg, and 30.0- and 37.5-mg tablets, respectively.

Distribution. Paroxetine, like sertraline and fluoxetine, is tightly bound to serum proteins in vitro, with binding of 95 and 93 percent at plasma levels of 100 and 400 ng/mL, respectively. Paroxetine has not been reported to alter the in vitro protein binding of warfarin or phenytoin (Dilantin). The volume of distribution is approximately 15 L/kg, and the rate of plasma clearance is close to 79 L/hr. Most individuals achieve steady-state plasma concentrations within 5 to 10 days of dosing of the IR formulation and within 2 weeks with paroxetine CR.

Metabolism and Elimination. The principal path of metabolism for paroxetine is oxidation and methylation. A number of major metabolites

have been identified, although none appears to be any greater than 1/50 as potent as the parent compound in the in vitro reuptake inhibition of 5-HT. A single test dose of paroxetine (30 mg) was recovered as 64 percent in the urine (2 percent as parent compound and 62 percent as metabolites over 10 days). In the feces, 36 percent was excreted, with less than 1 percent as the parent compound.

Paroxetine saturates the CYP 2D6 isozyme of the CYP system and therefore exhibits nonlinear pharmacokinetics. Steady-state plasma maximal and minimal concentrations were elevated 6 and 14 times, respectively, over single-dose values after reaching steady state. This translates into a 24-hour AUC value close to eightfold higher at steady state than would be predicted from a single dose.

In elderly patients, dosing of paroxetine at 20, 30, or 40 mg per day produced plasma concentrations approximately 70 to 80 percent greater than those observed in younger patients. Patients with a creatinine clearance of less than 30 mL/min demonstrated a fourfold elevation in plasma levels compared to healthy controls. Patients with moderate renal impairment (creatinine clearance of 30 to 60 mL/min), as well as patients with moderate hepatic impairment, had a twofold increase in plasma concentrations. All of the previously mentioned groups should receive reduced starting doses of paroxetine, for example, 10 mg per day of paroxetine or 12.5 mg per day of paroxetine CR.

FLUVOXAMINE AND FLUVOXAMINE CR

Chemistry

Fluvoxamine is a member of the 2-aminoethyloxime ethers of the aralkyl ketone family. The empirical formula of fluvoxamine is $C_{15}H_{21}O_2N_2F_3 \cdot C_4H_4O_4$, and it has a molecular weight of 434.4. The chemical designation is 5-methoxy-4′-(trifluoromethyl)valerophenone-(E)-*O*-(2-aminoethyl)oxime maleate (1:1).

Pharmacological Actions

Absorption. The bioavailability of fluvoxamine maleate is 53 percent by oral administration, and its absorption is essentially unaffected by food. At steady state, maximal plasma concentrations are achieved 3 to 8 hours after dosing.

Distribution. In vitro assays demonstrate protein binding, predominantly albumin, of approximately 80 percent for fluvoxamine at concentrations of 20 to 2,000 ng/mL. The volume of distribution for fluvoxamine is approximately 25 L/kg. The extent of excretion of fluvoxamine into breast milk is unclear.

Metabolism and Elimination. The primary route of metabolism for fluvoxamine is oxidative demethylation and deamination in the liver. The main human metabolite, as determined by radiotracer studies, is fluvoxamine acid, which, along with its *N*-acetylated analog, accounts for approximately 60 percent of urinary products. In this same assay, fluvoxethanol comprised approximately 10 percent. Fluvoxamine acid is 10- to 100-fold less potent in an in vitro assay of 5-HT reuptake inhibition than the parent compound.

Fluvoxamine demonstrates nonlinear pharmacokinetics. In 30 normal volunteers receiving fluvoxamine at 100, 200, or 300 mg per day over 10 days, the maximum plasma concentration was 88, 283, and 546 ng/mL, respectively. This represents increases of 160 and 207 percent for the 200- and 300-mg doses, respectively, over what would have been predicted from the lowest dose. The mean plasma half-life at steady state is 15.6 hours in young healthy volunteers receiving 100 mg per day and 13.6 hours in those receiving 50 mg per day. After repeated dosing, the half-life increases to 17 to 22 hours. Steady state is typically achieved in approximately 7 days.

Elderly patients show a prolongation in the half-life of fluvoxamine. Compared to younger patients, in the elderly, the half-life at 50 mg per day goes from 13.6 to 17.4 hours, and it goes from 15.6 to 25.9 hours for dosing of 100 mg per day.

Similar to adults, children (6 to 11 years of age) and adolescents (12 to 17 years of age) display nonlinear, multiple-dose pharmacokinetics with fluvoxamine. Steady-state plasma concentrations of fluvoxamine were two- to threefold higher in children than in adolescents after administration of 100 mg twice a day. The AUC, C_{max}, and minimum drug concentration (C_{min}) were similarly elevated in children compared to adolescents. Female children, in a small sample (N = 3), displayed higher plasma levels than male children. Adolescents are much more similar to adults with regard to the pharmacokinetic profile of fluvoxamine and thus may require higher doses than those necessary for the treatment of children.

Hepatic dysfunction is associated with an approximately 30 percent decrease in fluvoxamine clearance compared to healthy controls. In contrast, renal impairment (creatinine clearance of 5 to 45 mL/min) was not associated with decreased clearance.

CITALOPRAM AND ESCITALOPRAM

Citalopram was a late addition to the US market when it received approval for the treatment of major depressive disorder in 1998. Escitalopram is the *S*-enantiomer of racemic citalopram and is the most recent agent approved for the treatment of major depressive disorder in the United States, having been introduced in 2002.

Chemistry

Citalopram is the hydrobromide salt of a racemic mixture of a bicyclic phthalane derivative designated chemically as (±)–1-(3-dimethylaminopropyl)–1–(4-fluorophenyl)–1,3-dihydroisobenzofuran-5-carbonitrile, HBr. The molecular formula is $C_{20}H_{22}BrFN_2O$, and it has a molecular weight of 405.35.

Escitalopram is the oxalate salt of the *S*-enantiomer of racemic citalopram.

Pharmacological Actions

Absorption. Citalopram is well absorbed orally, and absorption is unaffected by food. The bioavailability of oral citalopram (40 mg, single dose) is approximately 80 percent of that observed after intravenous (IV) administration. Maximum systemic concentration is achieved approximately 4 hours after ingestion. The tablet and oral preparations of citalopram are bioequivalent.

Peak serum levels of escitalopram are seen approximately 5 hours after ingestion. The tablet and oral preparations of escitalopram are bioequivalent.

Distribution. Citalopram has less protein binding than some of the other SSRIs, with citalopram and its two primary metabolites demethylcitalopram (DCT) and didemethylcitalopram (DDCT) displaying in vitro protein binding of 80 percent. The volume of distribution of citalopram is 12 L/kg, and the clearance rate is 19.8 L/hr, of which approximately 20 percent is renal clearance. The half-life of citalopram is approximately 35 hours, and most patients should achieve steady-state concentrations in approximately 1 week.

Escitalopram has protein binding, measured in vitro, of approximately 56 percent. The clearance rate of escitalopram is 36 L/hr, with approximately 7 percent due to renal clearance. The volume of distribution is not well characterized. The mean terminal half-life is approximately 27 to 32 hours, so steady-state concentrations should be achieved after approximately 7 days of dosing.

Metabolism and Elimination. The principal pathway of metabolism for citalopram is by *N*-demethylation to DCT, and the primary CYP isozymes involved are CYP 3A4 and CYP 2C19. Citalopram is the compound found in highest concentrations in humans, with DCT and DDCT present at concentrations of approximately 50 and 10 percent, respectively, of that of citalopram. The metabolites DCT and DDCT possess only approximately one-eighth the potency of inhibition of the 5-HT transporter that citalopram does and so likely contribute little if any to the therapeutic activity. Two other compounds are products of citalopram metabolism in addition to DCT and DDCT; these are citalopram-*N*-oxide and a

deaminated propionic acid derivative. Citalopram displays linear and dose-proportional pharmacokinetics for single and multiple doses in the dose range of 10 to 60 mg per day.

Citalopram pharmacokinetics has been studied in different population subgroups. Normal controls 60 years of age and older in a single-dose study exhibited increases in AUC and half-life by 30 and 50 percent, respectively. The same population after receiving multiple doses had increases of 23 and 30 percent, respectively. Thus, 20 mg is the recommended daily dose in the geriatric population.

Gender differences have been examined in citalopram metabolism as well. In pharmacokinetic studies, citalopram AUC in women was 150 to 200 percent that seen in men. In contrast, however, five other pharmacokinetic studies, as well as the pivotal clinical trials, failed to show differences in steady-state citalopram levels between men and women. There were no differences seen in levels of DCT or DDCT between genders.

Mild to moderate renal impairment is associated with a decreased clearance rate of 17 percent compared to healthy controls. The effect of severe renal impairment on citalopram clearance is not well characterized. Patients with reduced hepatic function have demonstrated a 37 percent decrease in clearance of orally administered citalopram and a doubling of half-life. The recommended dose for patients with hepatic impairment is 20 mg per day.

Escitalopram is metabolized to *S*-demethylcitalopram (*S*-DCT) and *S*-didemethylcitalopram (*S*-DDCT). Similar to the metabolism of citalopram, escitalopram is the predominant compound found in human plasma, with *S*-DCT being present at approximately 30 percent of the concentration of the parent compound. The level of *S*-DDCT is typically below limits of detection. Escitalopram inhibition of the 5-HT transporter in vitro is at least 7 and 27 times more potent than inhibition displayed by *S*-DCT and *S*-DDCT, respectively. The primary CYP isozymes thought to be involved in escitalopram metabolism are CYP 3A4 and CYP 2C19. Escitalopram displays linear and dose-proportional pharmacokinetics for single and multiple doses in the dose range of 10 to 30 mg per day.

Examination of escitalopram clearance in subjects 65 years of age and older in single-dose and multiple-dose studies showed increases of approximately 50 percent in AUC and half-life. Thus, 10 mg is the recommended daily dose in the geriatric population. Similar to citalopram, there have been no gender differences noted in the metabolism of escitalopram. The effect of reduced renal or hepatic function on escitalopram clearance is poorly characterized, but dose adjustments are extrapolated from data available for the racemic compound citalopram.

DRUG INTERACTIONS

SSRIs can produce pharmacodynamic or pharmacokinetic interactions

with other medications. Categories of potential pharmacokinetic interactions between the SSRIs and other agents are enzyme inhibition through the CYP isozyme system and protein binding. Fortunately, the majority of drug–drug interactions are not fatal, but they may insidiously compromise the patient's quality of life or overall tolerability of a therapeutic regimen. Thus, it is not simply an issue of whether drug–drug interactions are seen in a particular area of practice; it is often a question of whether they are recognized, given that they do occur on a regular basis. Ultimately, it is the subsequent pharmacodynamic consequence that dictates the importance of the pharmacokinetic interaction.

Interactions involving SSRIs are particularly relevant clinically because major depressive disorder often requires more than one step of treatment to elicit a remission of symptoms. Frequently, a second medication is added to augment the first. Augmentation of an initial SSRI with an almost endless list of other agents, psychotropic and otherwise, is standard practice. The STAR*D study looked at the "next-step" benefits of switching medications or adding on to an SSRI. The various steps and agents used in this study are shown in Figure 34.27–2. The results found that among patients who do not have a remission during up to 14 weeks of an initial SSRI or who could not tolerate such therapy, switching to another SSRI or augmentation of SSRIs with any agent resulted in similar symptom remission. All strategies used to augment therapy with the SSRI citalopram had similar remission rates. Factors to be considered when selecting augmentation treatments include efficacy, tolerability, burden of side effects, interactions among drugs, dosing convenience, adherence, and cost. These results do not, however, answer the question of whether to use augmentation agents (or other treatment combinations) as first-line treatment in an attempt to achieve greater remission rates sooner in more patients than with SSRIs alone.

Pharmacodynamic Interactions

Symptoms such as autonomic instability, abdominal pain, diarrhea, confusion or delirium, increased motor activity, myoclonus, hyperpyrexia, cardiovascular shock, or death—the serotonin syndrome—are greatest when SSRIs are combined with MAOIs. Accordingly, with the exception of fluoxetine, SSRIs should not be administered within 14 days of discontinuing treatment with an MAOI. At least 14 days should pass before an SSRI (again, with the exception of fluoxetine) is started after treatment with an MAOI is stopped. Because of its long elimination half-life, the prudent washout period for starting fluoxetine after an MAOI would be at least 3 weeks. Five weeks is the minimum recommended interval between stopping fluoxetine and starting an MAOI owing to the long half-life of norfluoxetine. Tryptophan in combination with SSRIs also has the potential to produce the 5-HT syndrome.

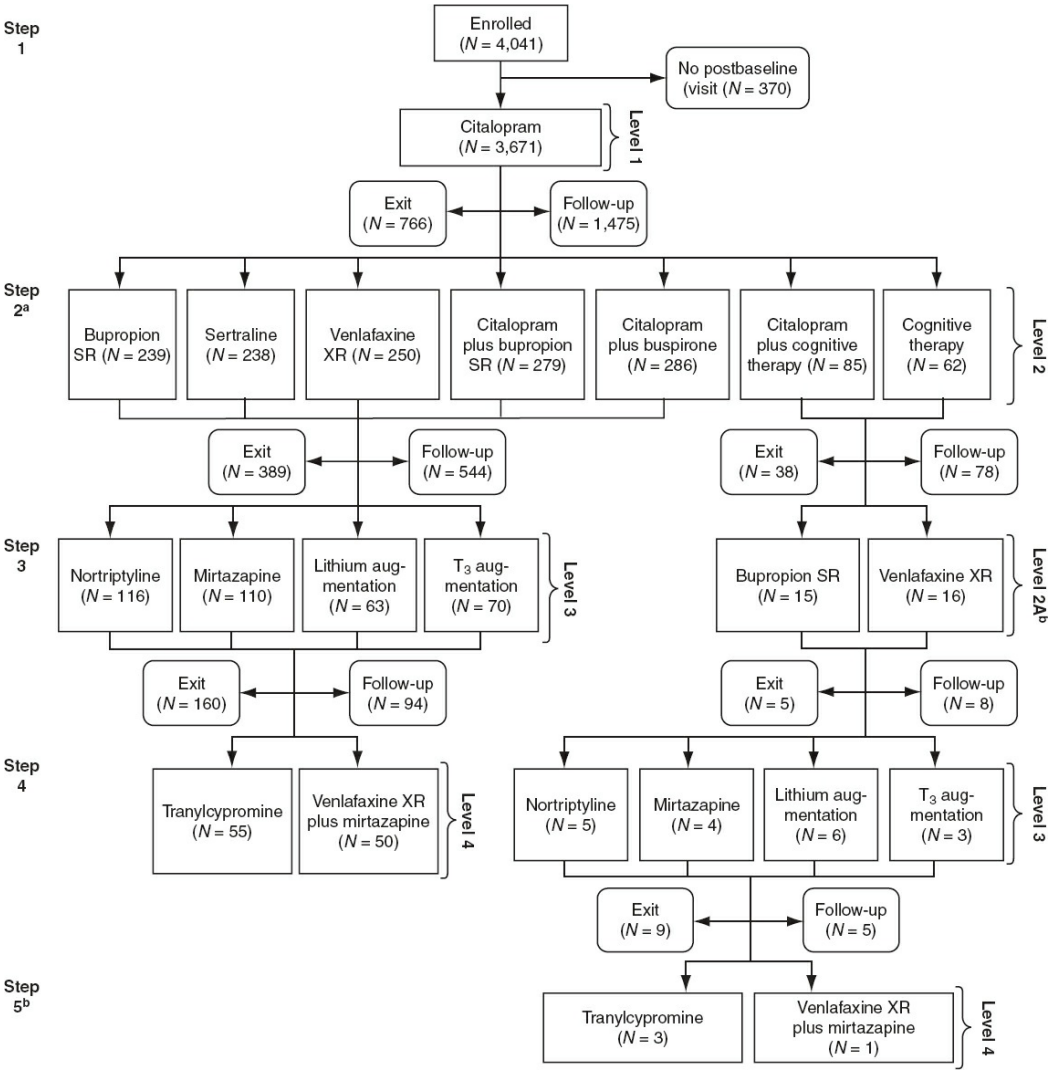

FIGURE 34.27–2. Steps in the STAR*D study.

Because SSRIs may reduce dopamine turnover, use in combination with dopamine-receptor antagonists can raise prolactin levels, even causing breast enlargement and galactorrhea and produce extrapyramidal side effects.

A frequent drug combination in clinical practice is the use of a triptan to treat migraine along with an SSRI. Concurrent use of a triptan and an SSRI is normally uneventful, but serotonin toxicity occurs occasionally. There may be additive effects on the serotonin system, as well as inhibition of triptan metabolism by SSRIs. Responding to multiple reports of serious reactions, in 2006 the FDA issued a Public Health Advisory warning of the possibility of serotonin syndrome when triptans and SSRIs are taken together.

According to the FDA advisory, physicians prescribing a triptan and SSRI should weigh the potential risk of serotonin syndrome against the expected benefit of using a triptan with an SSRI and follow patients closely if a triptan and an SSRI are used together, particularly during treatment

initiation, with dose increases, or with the addition of another serotonergic medication. Patients who take a triptan and an SSRI together should be advised to seek medical attention immediately if they experience the symptoms of serotonin syndrome.

Concurrent administration of an SSRI with an MAOI, L-tryptophan, or lithium can result in risk of *serotonin syndrome*. Its symptoms in order of appearance as the condition worsens are (1) diarrhea; (2) restlessness; (3) extreme agitation, hyperreflexia, and autonomic instability with possible rapid fluctuations in vital signs; (4) myoclonus, seizures, hyperthermia, uncontrollable shivering, and rigidity; (5) delirium, coma, status epilepticus, cardiovascular collapse, and death.

Treatment of the serotonin syndrome consists of removing the offending agents and promptly instituting comprehensive support care with nitroglycerine, cyproheptadine, methysergide (Sansert), cooling blankets, chlorpromazine (Thorazine), dantrolene (Dantrium), benzodiazepines, anticonvulsants, mechanical ventilation, and paralyzing agents.

Pharmacokinetic Interactions

The ability of antidepressants to exert a pharmacokinetic effect by enzyme inhibition is dependent on the extent of inhibition that is produced. Typically, classification schemes of antidepressants denote individual drugs as high, medium, or low or minimal potential for isozyme inhibition, as shown in Table 34.27–4.

Predicting the potential clinical relevance of isozyme inhibition requires knowledge of not only the potential inhibitory properties of drugs, but also the clinically relevant substrates for the isozymes. Examples of clinically relevant interactions include fluvoxamine and theophylline (Theo-Dur) through CYP 1A2 inhibition; clozapine (Clozaril) and fluvoxamine through CYP 1A2 inhibition; fluoxetine, sertraline, or paroxetine through CYP 2D6 inhibition; and fluvoxamine with alprazolam (Xanax) or clonazepam (Klonopin) through CYP 3A4 inhibition. Examples of commonly prescribed medications and the isozymes that metabolize them are listed in Table 34.27–4. The combination of a substrate and a potential inhibitor does not contraindicate their simultaneous use.

Fluoxetine may displace other tightly protein-bound drugs, such as warfarin or digitoxin (Digitaline). Good clinical care calls for careful monitoring of potential elevated drug concentrations of agents when tightly protein-bound drugs are coadministered, especially if one agent has a low therapeutic index.

In the CYP isozyme system, fluoxetine is considered to be a relatively potent inhibitor of the CYP 2C and CYP 2D6 isozymes but with no appreciable activity in the CYP 3A4 isozyme system, although norfluoxetine does demonstrate some inhibition of the CYP 3A4 isozyme in vitro. Table

34.27–4 contains a partial listing of substrates at a variety of CYP isozymes.

Fluoxetine, similar to paroxetine, sertraline, and citalopram, inhibits the CYP 2D6 isozyme. Fluoxetine, however, has a significantly longer half-life than the other SSRIs. A study to determine the onset and offset of CYP 2D6 inhibition by fluoxetine (20 mg per day), paroxetine (20 mg per day), or sertraline (50 mg per day for 3 days, then 100 mg per day) found the inhibition half-life for fluoxetine to be longer (7.0 ± 1.5 days) than paroxetine (2.9 ± 1.9 days) or sertraline (3 ± 3 days). There was also a significantly longer period of time required to reverse the CYP 2D6 inhibition after discontinuation of fluoxetine (extrapolated time to baseline of 63.2 ± 5.6 days) compared to paroxetine (20.3 ± 6.4 days) and sertraline (25 ± 11 days). There exists then a potential for drug–drug interactions for an extended period of time after discontinuation.

Fluoxetine may produce hypoglycemia in patients with diabetes who are treated with insulin or oral hypoglycemics. Dose reduction of the hypoglycemic or insulin may be required.

Sertraline, like fluoxetine and paroxetine, is tightly bound to serum proteins, and the manufacturer warns about a potential interaction between sertraline and other tightly protein-bound drugs, such as warfarin and digitoxin. Good clinical practice dictates checking serum levels or activity when an antidepressant with high protein binding is added to a tightly protein-bound drug with a low therapeutic index. Most assays do not differentiate between bound and free drug, so functional assays, such as prothrombin time with warfarin, are a better indicator.

Potential medication interactions with paroxetine include protein-binding interactions or enzymatic inhibition mediated by the CYP isozyme system. Paroxetine, like many of the other SSRIs, is tightly protein bound and may potentially displace other highly protein-bound drugs, such as warfarin or digitoxin. There may be a pharmacokinetic interaction between paroxetine and warfarin that prolongs bleeding, although the prothrombin time is unchanged. The addition of paroxetine, like that of other SSRIs, to a drug regimen in which tightly protein-bound medications are concurrently prescribed warrants careful observation during dose adjustment.

Paroxetine is a relatively potent inhibitor of the CYP 2D6 isozyme system but appears to be without clinically significant effects in the CYP 1A2, CYP 2C, or CYP 3A4 isozyme systems. Coadministration of paroxetine with other medications that are substrates of the CYP 2D6 isozyme should be approached cautiously with an awareness of potential increases in serum concentrations of drugs that primarily use the CYP 2D6 pathway for metabolism. Examples of CYP 2D6 substrates include the type 1C antiarrhythmics—encainide (Enkaid), flecainide (Tambocor), propafenone (Rythmol), and mexiletine (Mexitil)—and β-blockers—labetalol (Normodyne), metoprolol, propranolol, and timolol (Blocadren). Atenolol (Tenormin), a cardioselective β-blocker, is renally excreted and exhibits no

pharmacokinetic interactions. Codeine is inactive as an analgesic and requires metabolism to morphine via the CYP 2D6 isozyme to be effective in reducing pain. Inhibition of the CYP 2D6 pathway can result in loss of analgesic properties. Paroxetine is a substrate and an inhibitor of the CYP 2D6 isozyme system, but there exists an alternative pathway for its metabolism that is not saturated.

Table 34.27–4.
Cytochrome P450 (CYP) Inhibitory Potential of Commonly Prescribed Antidepressants

Relative Rank	CYP 1A2	CYP 2C	CYP 2D6	CYP 3A
Higher	Fluvoxamine (Luvox) Grapefruit	Fluoxetine Fluvoxamine	Bupropion Fluoxetine Paroxetine	Fluvoxamine Grapefruit Nefazodone Tricyclics
Moderate	Tertiary amine tricyclics Fluoxetine (Prozac)	Sertraline	Secondary amine tricyclics Citalopram (Celexa) Escitalopram (Lexapro) Sertraline	Fluoxetine Sertraline
Low or minimal	Bupropion (Wellbutrin) Mirtazapine (Remeron) Nefazodone (Serzone) Paroxetine (Paxil) Sertraline (Zoloft) Venlafaxine	Paroxetine Venlafaxine (Effexor)	Fluvoxamine Mirtazapine Nefazodone Venlafaxine	Citalopram Escitalopram Mirtazapine Paroxetine Venlafaxine

Citalopram enjoys a lower degree of protein binding (80 percent) than many other SSRIs, and escitalopram has an even lower degree of protein binding at approximately 56 percent. Thus, protein-binding interactions would not appear to be a problem with citalopram or escitalopram for the majority of patients, although the clinical trials for these two compounds, as with other antidepressants, are done typically in a relatively healthy population, and the pharmacokinetics of SSRIs when combined with multiple other agents is an underresearched area. The combination of citalopram with a highly protein-bound drug, such as warfarin, resulted in a prolongation of prothrombin time by only 5 percent.

Citalopram and escitalopram are modest inhibitors of the CYP 2D6 isozyme, based on in vivo testing. Coadministration of citalopram (40 mg per day) for 22 days resulted in a doubling of the plasma levels of metoprolol (Lopressor), a β-adrenergic blocker that is a CYP 2D6 substrate. In addition, coadministration of citalopram (40 mg per day) for 10 days along with a single dose of imipramine (Tofranil) (100 mg) did not alter the

plasma concentrations of imipramine but did increase plasma levels of the primary metabolite, desipramine (Norpramin), by 50 percent. Similarly, escitalopram (20 mg per day) for 21 days resulted in increases of C_{max} (50 percent) and AUC (82 percent) of metoprolol given as a single dose of 100 mg. Likewise, escitalopram (20 mg per day) for 21 days resulted in increases of C_{max} (40 percent) and AUC (100 percent) of desipramine given as a single dose of 50 mg.

In vitro enzyme assays suggest that citalopram is not an inhibitor for the CYP 3A4, CYP 2C9, or CYP 2E1 isozymes but is a weak inhibitor for CYP 1A2 and CYP 2C19, in addition to the data presented previously for the CYP 2D6 isozyme. Escitalopram does not appear to have in vitro inhibition of the CYP 3A4, CYP 1A2, CYP 2C9, CYP 2C19, and CYP 2E1 isozymes, although in vivo data are needed to confirm this.

Neither citalopram nor escitalopram, like other SSRIs, should be administered within 14 days of discontinuing treatment with an MAOI. At least 14 days should pass before either is started after treatment with an MAOI is stopped.

Fluvoxamine has moderate protein binding and may be less likely to produce displacement of tightly protein-bound drugs than some of the others in this class. In the CYP system, fluvoxamine is an inhibitor of the CYP 1A2, CYP 2C9, and CYP 3A4 isozyme systems. Theophylline, aminophylline (Phyllocontin), warfarin, propranolol (Inderal), and caffeine metabolism is inhibited through the CYP 1A2 isozyme by fluvoxamine. Warfarin metabolism is also inhibited through the CYP 2C9 isozyme system. Benzodiazepines, including alprazolam, clonazepam, triazolam (Halcion), and midazolam (Versed), are metabolized by hepatic oxidation via the CYP 3A4 isozyme. Approximately twofold elevations in serum concentrations of alprazolam have been observed after coadministration of fluvoxamine. Increased serum levels of carbamazepine (Tegretol), clozapine, methadone (Dolophine), propranolol, amitriptyline, clomipramine, and imipramine have been reported after the addition of fluvoxamine treatment. Thioridazine (Mellaril) should not be combined with fluvoxamine owing to the potential drug–drug interaction and the potential for cardiac QTc interval prolongation secondary to the inhibition of thioridazine metabolism.

The manufacturer warns against the coadministration of fluvoxamine with terfenadine (Seldane), astemizole (Hismanal), cisapride (Propulsid), and pimozide (Orap). All four compounds are inactive prodrugs on ingestion and require CYP 3A4 activity to produce the active metabolite. Inhibition of metabolism produces an accumulation of precursor that may produce cardiac conduction delay and a potentially fatal torsades de pointes type of arrhythmia.

In premarketing trials, smokers demonstrated a 25 percent increase in fluvoxamine metabolism compared to nonsmokers, presumably through

induction of the CYP 1A2 isozyme.

DOSING AND ADMINISTRATION

The therapeutic dose ranges for the SSRIs have been well established. In general, therapeutic effects are not dose related. There are no data that suggest that the presence of therapeutic levels of any SSRI correlates with plasma levels and subsequent clinical response. Nevertheless, patients may vary considerably in the amount of medication they need to derive clinical benefit.

Fluoxetine

Fluoxetine is available in 10- or 20-mg pulvules, as 10-mg tablets, as 90-mg enteric-coated capsules for once-weekly administration, and as an oral concentrate of 20 mg/5 mL. The suggested starting dose of fluoxetine in patients with major depression is 20 mg per day. Effective doses for the treatment of major depression have been reported, ranging from 5 to 80 mg per day in dose-finding trials. Owing to the long half-life, achieving a new steady state after dose adjustment takes longer than that seen with most other antidepressants. Decreasing the dose to treat emerging side effects, particularly if it has been several weeks since initiation of treatment or the last dose adjustment, is a reasonable approach to managing many side effects.

Preclinical trials of fluoxetine in bulimia nervosa showed efficacy greater than placebo at 60 mg per day. Treatment of OCD is also usually at higher doses than those used in major depression, with doses of 20 to 80 mg per day most commonly used. However, SSRIs take longer to reduce OCD symptoms than symptoms of depression, so time for the drug at normal doses should be allowed. The long half-life of fluoxetine and norfluoxetine does make interpretations of what may be the minimal therapeutic dose somewhat complicated. Fortunately, because fluoxetine is generally well tolerated, this becomes less of an issue than it might be with older antidepressants, such as TCAs.

Treatment of panic disorder is often initiated at less than 5 to 10 mg per day with a gradual upward dose titration. Treatment of social anxiety disorder or premenstrual dysphoric disorder is often started at 10 to 20 mg per day. Dose is adjusted upward as needed.

Starting doses of 10 mg per day are often used in treating children, adolescents, or elderly patients. Adjustment upward is based on clinical response and tolerance of side effects.

Paroxetine

Paroxetine HCl is available as 10-, 20-, 30-, and 40-mg-size tablets. The 20-mg-size tablets are scored. Paroxetine CR is available in 12.5-, 25-, and

37.5-mg tablets. Paroxetine is also available in an oral suspension at a concentration of 10 mg/5 mL. Paroxetine mesylate is available as 10-, 20-, 30-, and 40-mg tablets.

The recommended starting dose for paroxetine is 20 mg per day for all approved indications, except in panic disorder, for which an initial dose of 10 mg per day is suggested. Starting paroxetine at 10 mg per day is also suggested for the elderly, medically ill, or those with significant renal or hepatic impairment. Paroxetine CR starting dose recommendations are similar, with 25 mg per day in major depressive disorder and 12.5 mg per day in patients with panic disorder, the elderly, the medically ill, or those with significant renal or hepatic impairment. Many clinicians find starting the more anxious or somatically preoccupied patient at a dose of 10 mg of paroxetine or 12.5 mg of paroxetine CR for a few days to be of benefit.

Taking paroxetine with meals, at least initially, reduces the incidence of nausea, and the majority of patients describe nausea, if present, to be a transient side effect. Patients with major depression who do not respond to 20 mg per day of paroxetine or 25 mg per day of paroxetine CR can have their doses increased by 10-mg (paroxetine) or 12.5-mg (paroxetine CR) intervals to as much as 50 mg per day (paroxetine) or 62.5 mg per day (paroxetine CR). In studies that established paroxetine to have efficacy in maintenance treatment, patients received a mean dose of approximately 30 mg per day.

The target dose of paroxetine in the treatment of OCD is 40 mg per day, with most responders receiving 40 to 60 mg per day. Presumably, doses of paroxetine CR need to be approximately 25 percent higher. The patient's response to treatment of OCD is often slower than what is seen in major depressive disorder.

Patients with panic disorder usually tolerate the side effects of paroxetine, like antidepressants in general, if started at a lower dose than is typically used in treating depression. Most panic disorder patients do better if started at an initial dose of 10 mg per day, with gradual titration to a target dose of 40 mg per day. Dose increases can be made at the rate of 10 mg per week. The maximal dose studied in premarketing clinical trials of paroxetine was 60 mg per day. The dosing strategy with paroxetine CR is similar, starting at 12.5 mg per week, and the maximal dose studied in clinical trials was 75 mg per day.

Dosing for other disorders is usually done with an initial dose of 10 to 20 mg per day (12.5 to 25.0 mg per day for paroxetine CR) and a gradual upward titration.

Among the SSRIs, paroxetine is associated with the most severe withdrawal syndrome and carries the strongest FDA warning about use in pregnancy. It carries a category D rating from the FDA.

Sertraline

Sertraline is available in scored tablets of 25-, 50-, and 100-mg strengths. The oral concentrate is formulated at a concentration of 20 mg/mL with 12 percent alcohol content and must be diluted before use. Owing to the alcohol content in the oral concentrate, it is contraindicated in conjunction with disulfiram (Antabuse). Initial dosing for major depressive disorder and OCD is typically at 50 mg per day, with escalation to 100 mg per day after 4 to 7 days of treatment. Initial treatment of panic disorder, social anxiety disorder, and PTSD is more typically 25 mg per day. Children in OCD studies were started at 25 mg per day, whereas adolescents in OCD studies were started at 50 mg per day. Treatment of premenstrual dysphoric disorder is usually started at 50 mg per day given daily or during the 2 weeks of the luteal phase. Individual patients may demonstrate improved tolerability starting at lower doses. Not uncommonly, patients require an upward titration of the dose.

The dose range in most patients with anxiety or affective disorders is typically 100 to 200 mg a day, although some patients respond to 50 mg per day. Patients with OCD often require higher doses for full therapeutic effect.

Citalopram and Escitalopram

Citalopram is available in 10-, 20-, and 40-mg-size tablets. The 20- and 40-mg-size tablets are scored. Citalopram is also available in an oral solution at a concentration of 10 mg/5 mL. Escitalopram is available in 10- and 20-mg-size tablets, both of which are scored. An oral solution of 5 mg/5 mL is available as well.

The manufacturer recommends a starting dose for citalopram of 20 mg per day, with the expectation of generally increasing the dose to 40 mg per day, which, in clinical practice, is typically done after at least 1 week on the initial dosing. Highly anxious patients or those with increased sensitivities to side effects of medications may benefit from starting with the 10-mg tablet. The recommended starting dose of escitalopram is 10 mg per day. In clinical trials, the 20-mg-per-day dose did not offer any added benefit, but individual patients are often observed to benefit from the higher dose. If the dose is to be increased to 20 mg, this should generally occur no sooner than after 1 week on the initial dose. Because the 10-mg tablet is scored, 5 mg as a starting dose, if required, is available.

Many patients are observed to do well if citalopram or escitalopram is taken after meals for the first few days. After that period, they can take the medication with or without food.

Fluvoxamine

Fluvoxamine is available in 25-, 50-, and 100-mg scored tablets and 100- and 150-mg controlled release capsules. Adults with OCD are often started

at a dose of 50 mg administered at bedtime. The dose may be increased by 50 mg every 4 to 7 days. The effective dosing range in premarketing trials with OCD was 100 to 300 mg per day, although, clearly, some patients do respond to higher doses. The manufacturer suggests dividing the daily dose if it exceeds 100 mg per day, although many patients tolerate doses larger than 100 mg as a single dose. Initial upward titration is often easier with a split-dose regimen.

Treatment of OCD in adolescents and children should start at 25 mg per day with 25-mg-per-day increases every 4 to 7 days. The effective dosing range in premarketing studies was 50 to 200 mg per day. Adolescents often require a dose close to that used in adults, whereas children, perhaps owing to higher plasma fluvoxamine levels, may respond to the lower portion of the dosing range. Due to its short half-life, total daily doses of fluvoxamine greater than 50 mg should be administered as a split dose. Using the controlled release formulation obviates this need for divided doses.

Dosing for disorders other than those for which fluvoxamine is approved in the United States is usually initiated at 25 to 50 mg per day. Patients with a major depressive disorder typically show responses over the range of 100 to 300 mg per day, although response is often seen in the range of 100 to 200 mg per day. Due to possible early activation of anxiety symptoms, treatment of panic disorder is best started at a dose of 25 mg per day for most patients, owing to increased sensitivity of side effects in this population. The typical effective dose is usually 100 to 200 mg per day.

SELECTIVE SEROTONIN REUPTAKE INHIBITORS AND SUICIDALITY

In the early 1990s, a debate arose as to whether fluoxetine treatment was associated with an increase in suicidal ideation and suicide attempts in some patients. A comparison of the emergence of suicidal ideation in patients receiving fluoxetine, TCAs, or placebo suggested that fluoxetine-treated patients had less emergent suicidal ideation than those treated with TCAs. Both antidepressant groups showed less emergent suicidal ideation than the group treated with placebo. At times, fluoxetine produces symptoms, often described as restlessness, that are similar to akathisia in some patients. In the spring of 2004, the FDA asked manufacturers of the SSRIs, along with those of the other newer-generation antidepressants, to include a warning in their labeling recommending close observation of both adult and pediatric patients for signs or symptoms of worsening depression, including suicidal ideation, early on in therapy with these medications or on titration of dosing. This followed earlier recommendations concerning the potential of newer-generation antidepressants to be associated with a higher rate of agitation, suicidal thoughts, or both in clinical trials when these rates were compared to treatment with placebo in studies of children and adolescents with major depressive disorder. Many clinicians believe that the recommendations to

observe patients for worsening of symptoms during the initiation of therapy represent a recommendation for sound clinical care. The issue regarding treatment of children and adolescents remains somewhat unclear. One distinct possibility is that the 2 percent rate of agitation that was generally observed in studies that have been made public may represent pediatric patients who would ultimately go on to develop bipolar affective disorder.

A problem with the FDA advisory is that it was based on short-term trials, and did not include suicidal patients. A study—the Adolescent Depression and Psychotherapy Trial (ADAPT)—actively included suicidal patients, and found that the propensity for depressive relapse and self-reported suicidality in adolescents were reduced with SSRI administration. Furthermore, this study found that the reduction in clinical symptoms with SSRIs appeared to be a long-term effect. The authors of that study concluded that SSRIs are not detrimental to adolescent brain development and that therapeutic benefits outweigh the potential developmental risks.

Since the first press reports appeared, and especially since the new warning about suicidality in children and adolescents, there has been a drop in both SSRI prescriptions for that group and a decrease in the diagnosis of depression. Unfortunately, studies in the United States, Canada, and The Netherlands have shown a dramatic increase starting in 2003 to 2004 in suicide attempts and completed suicides. Analysis involving older adults shows that antidepressants reduce the risk of suicidal thoughts for 25- to 64-year-olds and may reduce the risk of both suicidal thoughts and behavior for people older than 65 years. The obvious conclusion is that although SSRIs might increase suicidality in a small percentage of patients, the use of these drugs is associated with significant decreases in suicide attempts. Health advisories and warnings issued by the FDA may in fact have produced an unintended withholding of appropriate care and caused unexpected harm.

PREGNANCY AND BREAST-FEEDING

SSRIs are the most frequently used class of antidepressant medications among women of childbearing age. It is therefore not surprising that they are the most commonly used psychotropic drugs during pregnancy. With the exception of paroxetine, the SSRIs are considered relatively safe to take during pregnancy when deemed necessary for treatment of the mother. There are adverse effects, in terms of both serotonin toxicity and withdrawal.

Women who are pregnant or planning a pregnancy and are taking SSRIs should consult with their physician about whether to continue taking their medication. As with all medications that treat serious disorders, such as those treated by SSRIs, the potential benefits and risks of SSRI treatment need to be weighed against the potential effect of not treating,

given that pregnancy does not protect against the serious disorders these drugs treat.

A Centers for Disease Control and Prevention (CDC) epidemiological study on the causes of birth defects looked at the use of fluoxetine, sertraline, paroxetine, and citalopram during pregnancy. It found that these drugs did not significantly increase the risk for most birth defects. The study did find an association between SSRI use and a slightly increased risk for three specific birth defects: A defect of the brain, one type of abnormal skull development, and a GI abnormality. These increases in risk were minimal and have not been found before or since.

The CDC study also looked at the risk for congenital heart defects associated with paroxetine use. In 2005, at the FDA's recommendation, paroxetine's pregnancy category was changed from C to D and new data added the Warnings section of paroxetine's prescribing information. The FDA's conclusions and changes in paroxetine's prescribing information were based on preliminary analyses of two unpublished epidemiology studies: In a study using Swedish national registry data, women who received paroxetine in early pregnancy had an approximately twofold increased risk for having an infant with a cardiac defect compared to the entire national registry population (the risk of a cardiac defect was about 2 percent in paroxetine-exposed infants vs. 1 percent among all registry infants). In a separate study using a US insurance claims database, infants of women who received paroxetine in the first trimester had a 1.5-fold increased risk for cardiac malformations and a 1.8-fold increased risk for congenital malformations overall compared to infants of women who received other antidepressants in the first trimester. The risk of a cardiac defect was about 1.5 percent in paroxetine-exposed infants versus 1 percent among infants exposed to other antidepressants.

Most of the cardiac defects observed in these studies were atrial or ventricular septal defects, conditions in which the wall between the right and left sides of the heart is not completely developed. In general, septal defects are among the most common types of congenital malformations. They range from those that are symptomatic and may require surgery to those that are asymptomatic and may resolve on their own. It is of note that the data in these studies were limited to first trimester exposures only, and there are not data to address whether this or any other risk extends to later periods of pregnancy. Based on this labeling, it was recommended that paroxetine should not be initiated in women who are in their first trimester of pregnancy or in women who plan to become pregnant in the near future.

Transient QTc prolongation has been noted in newborns whose mother was being treated with an SSRI during pregnancy.

Another possible pregnancy-related side effect in offspring of SSRI-treated mothers is persistent pulmonary hypertension of the newborn (PPHN), a condition associated with substantial infant mortality and

morbidity. A study has shown that infants with PPHN who were exposed to an SSRI after the completion of the 20th week of gestation, as compared with infants who were not exposed to an SSRI or non-SSRI before the 20th week of gestation or the use of non-SSRI antidepressant drugs at any time during pregnancy, had an increased risk of PPHN. Some neonates exposed to SSRIs late in the third trimester developed complications requiring prolonged hospitalization, respiratory support, and tube feeding. These are listed as precautions in the product information for all SSRIs. Physicians are advised in the labeling to consider tapering SSRIs in the third trimester, but it is not known whether stopping treatment in the last trimester avoids this problem. In addition, the clinical need for treatment of the mother—the risk of nontreatment—needs to be weighed. For example, pregnant women with untreated depression have a greater risk for obstetrical complications, premature delivery, and pre-eclampsia. Newborns of these women are at greater risk to be low birth weight, exhibit lower activity and interactive behaviors, have increased irritability, and demonstrate delayed language and behavioral difficulties.

Excretion of SSRIs in Breast Milk

SSRIs appear in breast milk to varying extents, and data from the literature differ in their findings with regard to the relative dose to the infant. Various reports about drug levels in milk are unreliable because of variations in milk collection procedures and because antidepressants tend to be lipophilic, so that breast milk concentrations for some drugs may vary more than twofold between the watery fore milk and the fattier hind milk, whereas others may show little variation. Accordingly, clinical monitoring of maternal plasma or breast milk is not routinely indicated.

In some studies, concentrations are lowest for fluvoxamine and sertraline. Low levels of both sertraline and its metabolite are found in the sera of infants. Concentration of sertraline and desmethylsertraline in breast milk is highly variable and, on average, is of the same order of magnitude as their concentration in the blood plasma of the mother. As a result, more than half of breast-fed babies receive less than 2 mg/day of sertraline and desmethylsertraline combined, and in most cases these substances are undetectable in their blood. No changes in serotonin uptake by the platelets of breast-fed infants occur, as measured by blood serotonin levels before and after beginning sertraline treatment. Sertraline is more likely to be detected in an infant if the maternal daily dose is 100 mg or higher. Data on paroxetine are mixed, with some studies showing somewhat higher breast milk concentrations and others reporting paroxetine to have the lowest transfer into breast milk. According to a pooled analysis of available data, breast-feeding infants exposed to paroxetine or sertraline are least likely to demonstrate detectable or elevated plasma drug levels, whereas infants exposed to fluoxetine appear

to be at higher risk of developing elevated levels. Most data show that the relative infant dose of citalopram is intermediate between those of paroxetine and fluoxetine. High doses of citalopram should also be used with caution. Escitalopram (Lexapro) and its demethyl metabolite into milk, the absolute and relative infant doses via milk and to assess any unwanted effects in the breast-fed infant.

Because its absolute infant dose is lower than that for an equivalent antidepressant dose of citalopram, it may be preferred over citalopram in treating depression in lactating women.

If treatment with an SSRI is started during the postpartum period, fluoxetine should not be the first alternative. Fluoxetine transfers to a greater extent and its active metabolite, norfluoxetine, has a long half-life of 1 to 2 weeks and may accumulate in a breast-fed infant. When the use of an SSRI is clearly indicated in a breast-feeding woman, the data generally indicate that the positive effects of breast-feeding outweigh the risks for pharmacological effects in the infant. Adverse effects are rarely reported in breast-fed infants. Some short-term adverse effects that have been reported in association with fluoxetine include colic, seizures, irritability, withdrawal symptoms, and cyanosis.

Adverse effects, although rare, may occur, so parents and caretakers should report irritability, poor feeding, or poor sleep or any other worrisome signs or symptoms in the infant. Premature infants or those with impaired hepatic or other metabolite disturbances need to be monitored more closely if the mother is taking SSRIs and breast-feeding.

USE IN CHILDREN

Use of SSRI antidepressants in children and adolescents has been controversial. Few studies have shown clear-cut benefits from the use of these drugs. However, some children and adolescents do exhibit dramatic response to these drugs in terms of depression and anxiety. Fluoxetine has most consistently demonstrated effectiveness in reducing symptoms of depressive disorder in both children and adolescents. This may be a function of the quality of the clinical trials involved. Sertraline has been shown to be effective in treating social anxiety disorder in this population, especially when combined with cognitive-behavioral therapy. Given the potential negative effect of untreated depression and anxiety in a young population and the uncertainty about many aspects of how children and adolescents might react to medication, any use of SSRIs should be undertaken only within the context of comprehensive management of the patient.

VORTIOXETINE

Vortioxetine (Brintellix) is an atypical antidepressant (a serotonin modulator and stimulator) approved by the FDA in 2013 for the treatment

of MDD in adults. Some experts think the drug should be classified as an SSRI because its major pharmacological effect is to inhibit the serotonin transporter. The clinical development program for vortioxetine yielded six positive 6- to 8-week studies, including one study in elderly people, and another positive maintenance study in adults. Vortioxetine was investigated as a treatment for GAD, but was found to be no better than placebo.

Pharmacodynamics

Vortioxetine is classified as a "serotonin modulator and stimulator." Some regard it as a multi-modal 5-HT_3, 5-HT_7, and 5-HT_{1D} receptor antagonist, 5-HT_{1B} receptor partial agonist, 5-HT_{1A} receptor agonist, and inhibitor of the 5-HT transporter. More specifically, it has been shown to possess the following pharmacological actions:

- Serotonin transporter blocker—K_i (binding affinity) = 1.6 nM,IC_{50} = 5.4 nM
- Norepinephrine transporter blocker—K_i = 113 nM
- 5-HT_{1A} receptor high-efficacy partial agonist/near-full agonist—K_i = 15 nM, IA = 80 percent
- 5-HT_{1B} receptor partial agonist—K_i = 33 nM
- 5-HT_{1D} receptor antagonist—K_i = 54 nM
- 5-HT_{3A} receptor antagonist—K_i = 3.7 nM
- 5-HT_7 receptor antagonist—K_i = 19 nM
- β_1-Adrenergic receptor ligand—K_i = 46 nM

Vortioxetine also increases extracellular levels of acetylcholine and histamine in the rat medial prefrontal cortex.

Therapeutic Indications

Being relatively new in clinical practice, most of the data about the efficacy of vortioxetine in treating MDD derives from multiple short-term clinical trials and one long-term relapse prevention trial. Of nine short-term (6- to 8-week) placebo-controlled trials in nonelderly patients, five were considered positive. Two trials were positive for vortioxetine 20 mg/day, but not for 10 or 15 mg/day. One trial failed because the active control did not separate from placebo. Another trial failed because the active control (duloxetine) did separate from placebo. Two other trials failed to demonstrate statistical superiority for vortioxetine over placebo. In the short-term placebo-controlled studies, a statistically significant antidepressant effect was observed at week 2, with additional improvement in subsequent weeks. The full antidepressant effect of vortioxetine was

typically not seen until week 4 or later.

Based on analysis of evaluable trials, the NNT for response with vortioxetine versus placebo is 7 and for remission versus placebo is 11. The most robust NNT for response versus placebo was for the 15-mg/day dose, with an NNT of 6; the most robust NNT for remission versus placebo was for the 5-mg/day dose. In studies with the active controls duloxetine 60 mg/day and venlafaxine 225 mg/day, the control drugs exhibited more robust effect sizes for response and remission compared with vortioxetine at any dose.

An MDD relapse-prevention (maintenance) trial was conducted involving subjects aged 18 to 75 years. Inclusion criteria were a current major depressive episode of at least 4 weeks' duration and at least one prior episode and a MADRS total score of at least 26. Subjects received a 12-week, open-label trial of vortioxetine 5 or 10 mg/day. Subjects who achieved considered full remission were randomized to either continue receiving vortioxetine or to receive placebo in a double-blind fashion. The dose of vortioxetine received at the last assessment in the open-label period was 5 mg/day for 35 percent of the subjects and 10 mg/day for the remaining 65 percent. Of the patients randomized, 75 percent in the vortioxetine group and 66 percent in the placebo group took vortioxetine 10 mg/day at the end of the open-label period. Time to relapse was the primary efficacy outcome and demonstrated superiority for vortioxetine versus placebo.

Pharmacokinetics

The pharmacological activity of vortioxetine is entirely derived from the parent molecule. It has no active metabolites. In normal volunteers, single daily doses of vortioxetine exhibits linear pharmacokinetics for the dose range of 2.5 to 60 mg. It has a mean terminal half-life of about 66 hours. Vortioxetine reaches peak plasma concentration within 7 to 11 hours postadministration. Steady-state plasma concentrations are usually achieved within 2 weeks. In studies, no differences were observed between the fed and fasting states. Accordingly, it may be taken with or without food. The absolute bioavailability of vortioxetine is 75 percent, and it has an extended absorption phase, a medium clearance and a large volume of distribution. C_{max} is reached in 7 to 11 hours. Vortioxetine is extensively metabolized, mainly through the CYP450 isozymes 2D6, 3A4/5, 2C19, 2C9, 2A6, 2C8, and 2B6. After oxidation it undergoes glucuronic acid conjugation. It is converted to a pharmacologically inactive, carboxylic acid metabolite through CYP2D6 oxidation. Poor metabolizers of CYP2D6 have nearly twice the plasma concentration vortioxetine than do extensive metabolizers. No significant impact on plasma protein binding or clearance has been demonstrated by hepatic or renal disease. Insignificant amounts

of unchanged vortioxetine are excreted in the urine. Concurrent use of vortioxetine has no effect on plasma levels of ethinyl estradiol/levonorgestrel or 5′-hydroxyomeprazole or bupropion. However, steady-state AUC and C_{max} of vortioxetine are increased when coadministered with bupropion fluconazole and ketoconazole. These are decreased, when vortioxetine is used with rifampicin.

Precautions and Adverse Reactions

The most common side effects reported with vortioxetine are nausea, diarrhea, xerostomia, constipation, vomiting, flatulence, dizziness, and sexual dysfunction. Based on clinical trial data, the incidence of sexual dysfunction is reportedly higher in patients taking vortioxetine than in patients taking placebos, but lower than in patients taking venlafaxine. Vortioxetine used alone in high dose or in combination with other medications, such as other antidepressants, can produce serotonin syndrome.

There are no published data on the efficacy and safety of vortioxetine use in children and adolescents. It has been studied in elderly patients, with no evidence of unusual adverse events. Vortioxetine is classified as FDA pregnancy category C.

Dosing

Patients typically initiate drug therapy at 10 mg daily, which may be subsequently increased to 20 mg daily. If 10 mg dosing is poorly tolerated, the dose can be lowered to 5 mg per day, though available evidence demonstrating clinical efficacy at this lower dosage is not entirely conclusive and is questionably reliable.

Reported efficacy data across clinical trials failed to indicate any linear correlation to patient dosing. Clinical trial data do not indicate a clinically significant difference response and remission rates at doses above 20 mg/day. In fact, during clinical trials, clinical improvement noted in those patients receiving a higher dose of 20 mg/day, compared to 10 mg/day, was insignificant. Of course, there are always outliers who may benefit from higher doses.

Available Formulations

In the United States vortioxetine is available in 5-, 10-, 15-, and 20-mg tablets.

References

Baldessarini RJ, Pompili M, Tondo L. Suicidal risk in antidepressant drug trials. *Arch Gen Psychiatry*. 2006;63:246–248.

Barbui C, Esposito E, Cipriani A. Selective serotonin reuptake inhibitors and risk of suicide: a

systematic review of observational studies. *CMAJ*. 2009;180:291–297.

Cipriani A, Barbui C, Brambilla P, Furukawa TA, Hotopf M. Are all antidepressants really the same? The case of fluoxetine: a systematic review. *J Clin Psychiatry*. 2006;67:850–864.

Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet*. 2009;373:746–758.

Cousins L, Goodyer IM. Antidepressants and the adolescent brain. *J Psychopharmacol*. 2015;29(5):545–555.

Diem SJ, Blackwell TL, Stone KL, et al. Use of antidepressants and rates of hip bone loss in older women: the study of osteoporotic fractures. *Arch Intern Med*. 2007;167:1240–1245.

Emslie GJ, Heiligenstein JH, Wagner KD, et al. Fluoxetine for acute treatment of depression in children and adolescents: a placebo-controlled, randomized clinical trial. *J Am Acad Child Adolesc Psychiatry*. 2002;41:1205–1215.

Fenger-Grøn J, Thomsen M, Andersen KS, Nielsen RG. Paediatric outcomes following intrauterine exposure to serotonin reuptake inhibitors – a systematic review. *Dan Med Bul*. 2011;58:A4303.

Glassman AH, O'Connor CM, Califf RM, et al; Osteoporotic Fractures in Men Study Group. Association of low bone mineral density with selective serotonin reuptake inhibitor use by older men. *Arch Intern Med*. 2007;167(12):1246–1251.

Glassman AH, O'Connor CM, Califf RM, et al; Sertraline Antidepressant Heart Attack Randomized Trial (SADHEART) Group. Sertraline treatment of major depression in patients with acute MI or unstable angina. *JAMA*. 2002;288:701–709.

*Hu X-Z, Rush AJ, Charney D, et al. Association between a functional serotonin transporter promoter polymorphism and citalopram treatment in adult outpatients with major depression. *Arch Gen Psychiatry*. 2007;64:783–792.

Kivimaki M, Batty DG. Antidepressant drug use and future diabetes risk. *Diabetologia*. 2012;55:10–12.

*Kocsis JH. Review: SSRIs and TCAs equally effective at treating chronic depression and dysthemia; SSRIs are associated with fewer adverse events than TCAs. *Evid Based Mental Health*. 2013;16:82.

Kornstein SG, Schatzberg AF, Thase ME, et al. Gender differences in treatment response to sertraline versus imipramine in chronic depression. *Am J Psychiatry*. 2000;157:1445–1452.

Kuo CK, Liao YT, Chen VC. Risk of upper gastrointestinal bleeding when taking SSRIs with NSAIDs or aspirin. *Am J Psychiatry*. 2014;171:582–582.

Leon AC, Keller MB, Warshaw MG, et al. Prospective study of fluoxetine treatment and suicidal behavior in affectively ill subjects. *Am J Psychiatry*. 1999;156:195–201.

Liebowitz MR, Stein MB, Tancer M, Carpenter D, Oakes R. A randomized, double-blind, fixed-dose comparison of paroxetine and placebo in the treatment of generalized social anxiety disorder. *J Clin Psychiatry*. 2002;63:66–74.

Liebowitz MR, Turner SM, Piacentini J, et al. Fluoxetine in children and adolescents with OCD: a placebo-controlled trial. *J Am Acad Child Adolesc Psychiatry*. 2002;41:1431–1438.

Looper KL. Potential medical and surgical complications of serotonergic antidepressant medications. *Psychosomatics*. 2007;48:1–9.

Loprinzi CL, Sloan JA, Perez EA, et al. Phase III evaluation of fluoxetine for treatment of hot

flashes. *J Clin Oncol*. 2002;20:1578–1583.

Lundberg J, Tiger M, Landén M, Halldin C, Farde L. Serotonin transporter occupancy with TCAs and SSRIs: a PET study in patients with major depressive disorder. *Int J Neuropsychopharmacol*. 2012;15:1167–1172.

Meyer JH, Wilson AA, Sagrati S, et al. Serotonin transporter occupancy of five selective serotonin reuptake inhibitors at different doses: an [11C]DASB positron emission tomography study. *Am J Psychiatry*. 2004;161:826–835.

Miner C, Brown E, McCray S, Gonzales J, Wohlreich M. Weekly luteal-phase dosing with enteric-coated fluoxetine 90 mg in premenstrual dysphoric disorder: a randomized, double-blind, placebo-controlled clinical trial. *Clin Ther*. 2002;24:417–433.

Pal S, Padala PR. A case of excessive yawning with citalopram. *Prim Care Companion J Clin Psychiatry*. 2009;11(3):125–126.

Pan AQ, Sun Q, Okereke OI, et al. Use of antidepressant medication and risk of type 2 diabetes: results from three cohorts of US adults. *Diabetologia*. 2012;55:63–72.

Phillips KA, Albertini RS, Rasmussen SA. A randomized placebo-controlled trial of fluoxetine in body dysmorphic disorder. *Arch Gen Psychiatry*. 2002;59:381–388.

Sansone RA, Sansone LA. SSRI-induced indifference. *Psychiatry (Edgmont)*. 2010;7:14–18.

Teicher MH, Glod C, Cole JO. Emergence of intense suicidal preoccupation during fluoxetine treatment. *Am J Psychiatry*. 1990;147:207–210.

Thase ME, Haight BR, Richard N, et al. Remission rates following antidepressant therapy with bupropion or selective reuptake inhibitors: a meta-analysis of original data from 7 randomized controlled trials. *J Clin Psychiatry*. 2005;66:974–981.

Trivedi MH, Rush AJ, Wisniewski SR, et al; STAR*D Study Team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. *Am J Psychiatry*. 2006;163:28–40.

Van Ameringen MA, Lane RM, Walker JR, et al. Sertraline treatment of generalized social phobia: a 20-week, double-blind, placebo-controlled study. *Am J Psychiatry*. 2001;158:275–281.

Zivin K, Pfeiffer PN, Bohnert AS, et al. Evaluation of the FDA warning against prescribing citalopram at doses exceeding 40 mg. *Am J Psychiatry*. 2013;170:642–650.

▲ 34.28 Second-Generation Antipsychotics

STEPHEN R. MARDER, M.D., AND MICHAEL C. DAVIS, M.D., PH.D.*

INTRODUCTION

The second-generation antipsychotics (SGAs) are a group of antipsychotic agents that were introduced into clinical psychiatry during the early 1990s. In the United States and most parts of the world, these agents are the most widely prescribed agents for schizophrenia and other illnesses associated with psychotic symptoms. In prior editions of this text, these agents were referred to as serotonin-dopamine antagonists (SDAs) based on the belief that they could be differentiated from the dopamine receptor antagonists based on their high affinity for serotonin 2A (5-HT_{2A}) as well as dopamine (D_2) receptors. However, as noted below, this explanation for differentiating the effects of these agents remains controversial. As a result, this term has gradually been replaced by the term second-generation antipsychotics or SGAs, which will be used in this section.

The first agent in this group was clozapine (Clozaril), an antipsychotic that was introduced in the United States in 1990, although it had been available in some countries in Europe since the early 1970s. Clozapine was considered an "atypical" antipsychotic because it was an effective agent at doses that caused negligible EPS. Moreover, it had broad effectiveness for a range of mood and anxiety symptoms that were common in schizophrenia. The enthusiasm for its effectiveness was balanced by a number of side effects (described later in this section). Attempts to duplicate clozapine's effectiveness in agents with milder side effects led to the development of a group of agents that included risperidone (Risperdal), olanzapine (Zyprexa), quetiapine (Seroquel), ziprasidone (Geodon), and aripiprazole (Abilify). Since the last edition, iloperidone (Fanapt), asenapine (Saphris), and lurasidone (Latuda) have been introduced in the United States. Cariprazine (Vraylar) and Brexpiprazole (Rexulti) have been introduced now. Unfortunately, none of these drugs duplicated clozapine's effectiveness in patients who are partially responsive to FGAs. However, these drugs are associated with antipsychotic effects at doses less likely to cause EPS. The term *atypical* is probably outdated since these agents are currently the most prescribed group of antipsychotics.

Since the previous edition of this text, new information about the effectiveness and tolerability of SGAs has become available as a result of large trials comparing these drugs and FGAs. New formulations of existing SGAs have been introduced and new indications for these agents have been approved by the FDA. In addition, the prescribing of SGAs for off-label conditions such as anxiety disorders, OCD, insomnia, and agitation in

elderly patients has increased despite very limited data indicating that these agents are both safe and effective for these conditions.

COMPARISONS OF SGAS

The results from meta-analyses have provided important information about the relative effectiveness of the SGAs and how they compare to FGAs. One large meta-analysis included more than 43,000 subjects and compared 13 SGAs and 2 FGAs (haloperidol and chlorpromazine). The study did not find a pattern indicating that there was a difference in efficacy between first- and second-generation agents. Rather, there were small differences in efficacy among the antipsychotics with clozapine, amisulpride (which is not available in the United States), olanzapine, and risperidone having an efficacy advantage. There were substantial differences among the agents in side effects, however, with clozapine, sertindole (which is not available in the United States), olanzapine, and quetiapine showing the least EPS and haloperidol the most. Haloperidol, ziprasidone, and lurasidone resulted in the least weight gain and clozapine and olanzapine caused the most. Paliperidone and risperidone led to the largest increases in prolactin and aripiprazole the least. These effects are consistent with results from a maintenance treatment trial that compared long-acting formulations of paliperidone and haloperidol. The effectiveness of the two agents for preventing relapse was similar. However, paliperidone resulted in greater prolactin elevation and more weight gain whereas haloperidol led to more akathisia. Taken together these results do not indicate a clear advantage for SGAs, except for likely a lower propensity for causing tardive dyskinesia. Instead, clinicians will need to consider a number of factors in selecting an antipsychotic that is suited for an individual patient.

SGAS AND MORTALITY

One study evaluated the risk of sudden cardiac death in a large population of patients in Tennessee. Although the study was unable to explain the cause of these deaths, the authors believe that they were most likely due to ventricular arrhythmias that can occur when drugs block potassium channels and prolong cardiac repolarization. The study found that both FGAs and SGAs were associated with an increased risk and that the newer drugs had a slightly higher risk. The incidence rate ratios for SGAs were 3.67 for clozapine, 2.91 for risperidone, 2.04 for olanzapine, and 1.88 for quetiapine. The study was unable to evaluate the risks of ziprasidone, aripiprazole, and paliperidone. The risk of cardiac death increased with both age and drug dose. Overall, the findings point to a relatively low risk of cardiac death, although the risk is greater than the risk of death from agranulocytosis with clozapine. Clinicians should consider obtaining an ECG before and after treating patients with an FGA or SGA. A prolonged

QT interval would signal that the patient is at risk.

A population-based cohort study from Finland evaluated mortality on different antipsychotics during an 11-year period from 1996 to 2006. Clozapine, an antipsychotic that had been considered a drug associated with high risk, was actually associated with the lowest risk of death and quetiapine was associated with the greatest risk. There was no indicating that the introduction of SGAs during the 1990s resulted in increased mortality. Long-term treatment with an antipsychotic in schizophrenia was associated with lower mortality when compared to not receiving an antipsychotic. These studies support the relative safety of long-term treatment with an antipsychotic.

BLACK BOX WARNINGS

Increased Mortality in Elderly Patients with Dementia

In 2005, the FDA mandated a black box warning for all SGAs for increased mortality in elderly patients with dementia-related psychosis. This was based on meta-analysis of 17 placebo-controlled trials, which demonstrated a risk of death of 1.6 to 1.7 times in the drug-treated patients versus the placebo-treated patients. The causes of death were variable but mostly cardiovascular or infectious in nature. It is noted that it is not clear whether the increased mortality can be directly attributed to the SGA as opposed to other patient characteristics.

Suicidal Thoughts and Behaviors

SGAs that carry therapeutic indications for the treatment of depression (e.g., quetiapine, aripiprazole, lurasidone, olanzapine, and brexpiprazole) carry the black box warning that is associated with all antidepressant medications. This warning is based on pooled analyses of short-term placebo-controlled trials of antidepressants (24 trials, 9 antidepressants, over 4,400 patients). These analyses found that antidepressant treatment was associated with an increased risk of suicidal thinking in behavior in children, adolescents, and young adults up to age 24. There was no increased risk of suicidality associated with antidepressant use in adults older than 24 (295 short-term trials, 11 antidepressants, over 77,000 patients), and there was a reduction in suicidality in adults 65 and older. It is recommended that all patients started on antidepressants be monitored for the emergence of agitation, irritability, unusual behavior, as well as suicidal thoughts and behaviors.

CLOZAPINE

History

Clozapine was the first of the SGAs. It was discovered in 1958 in Bern,

Switzerland, and was first studied in animal experiments in 1960. Soon after clinical trials with clozapine began, the drug appeared to be an effective antipsychotic agent that did not cause extrapyramidal side effects. The first studies showed that patients became calmer in the first 5 days, and psychotic excitement decreased from day 5 onward. Day 20 effects on psychosis were substantial, and negative symptoms improved at 40 days. However, early studies with clozapine in Europe were met with skepticism because this agent was inconsistent with the dogma that effective antipsychotic agents were associated with extrapyramidal effects. This dogma was challenged by a number of European authorities, including Hans Hippius, and, as a result, clozapine was eventually approved in several countries in Europe in 1972.

The enthusiasm about clozapine was dampened when it was discovered that the drug was associated with significant hematological toxicity. In 1975, clozapine was introduced in Finland, where 16 of 1,600 treated patients developed granulocytopenia (1,600 cells or less per mm^3). Eight of the 13 patients in whom granulocytopenia progressed to agranulocytosis died of infectious diseases. After 50 patients around the world had died, clozapine was withdrawn from most European markets, and research with this drug came to a virtual halt. Some patients who were receiving clozapine before its withdrawal from the market continued on the drug because of its clear advantages over prior treatment. Regular monitoring of white blood cell (WBC) counts in these individuals showed that clozapine-induced agranulocytosis was reversible when drug use was stopped. In addition, controlled trials in Europe showed that clozapine was more effective than conventional agents for treating schizophrenic patients who were either severely ill or resistant to other antipsychotic agents.

Discussions with the FDA led to a US multicenter trial to determine whether clozapine was effective for treatment-refractory patients. The effectiveness of clozapine was compared with that of chlorpromazine (Thorazine) in 268 schizophrenia patients with thoroughly documented failure to respond to conventional antipsychotic agents. The study, reported in 1988 by John Kane, Herbert Meltzer, and others, found that clozapine was more effective than chlorpromazine in terms of a number of important outcomes, including psychosis, mood, and anxiety. The results from this study and others led to the approval of clozapine in the United States in 1990 for schizophrenia patients who are resistant to treatment with other antipsychotic drugs or who are unable to tolerate conventional drugs because of extrapyramidal side effects or severe tardive dyskinesia.

Treatment resistance in schizophrenia is a poorly defined concept that has evolved since clozapine's introduction. At that time, treatment resistance referred to a failure to respond to a conventional dopamine receptor antagonist. Because of clozapine's side-effect profile, it was hoped that other SGAs would replace it as the agent of choice for treatment-

resistant patients and, to some extent, this appears to have happened, as clozapine use continues to be restricted to a relatively small proportion of patients with schizophrenia and other psychotic illnesses. As noted below, there is substantial evidence, including the results from phase 2E of Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) as well as meta-analyses, indicating that clozapine is probably the most effective antipsychotic.

Chemistry

Clozapine is a five-membered heterocyclic compound, 8-chloro-11-(4-methyl-1-piperazinyl)-5*H*-dibenzo-[*b,e*][1,4-]diazepine. Its chemical formula is $C_{18}H_{19}N_4Cl$ with a molecular weight of 326.8. Substitutions for the A-ring have led to other clinically effective clozapine-like agents; however, loxapine (Loxitane), an antipsychotic with a clozapine-related molecular structure, has a clinical and side-effect profile similar to other traditional antipsychotic drugs. Other clozapine-derived antipsychotic agents are olanzapine and quetiapine, but being a derivative does not translate to similar efficacy. Chemical structures for clozapine and the other SGAs discussed in this chapter are depicted in Figure 34.28–1.

Pharmacological Actions

Pharmacokinetics. The pharmacokinetic profiles of clozapine and other SGAs in special populations are summarized in Table 34.28–1.

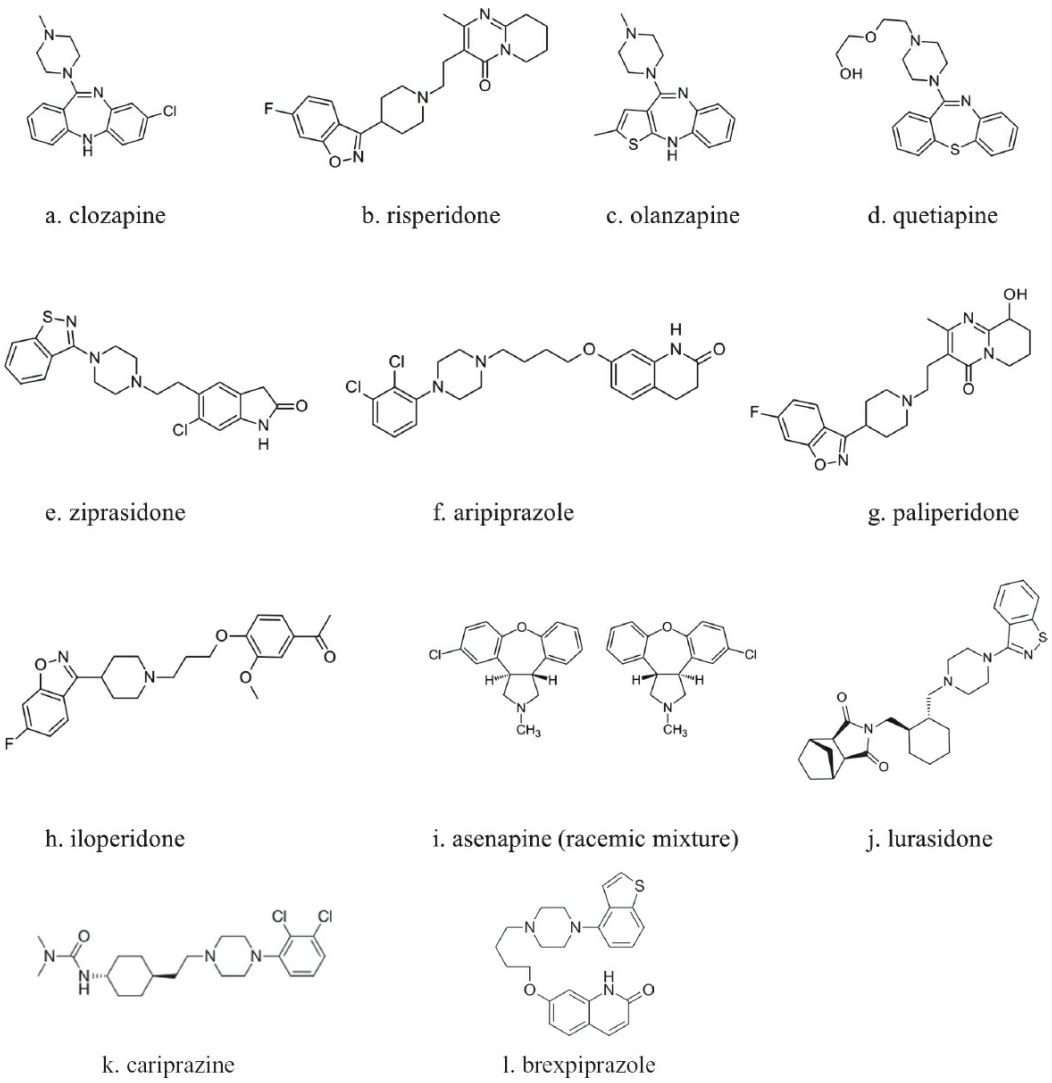

FIGURE 34.28–1. Molecular structures of SGAs discussed in this chapter. Panels a-l by Fuse809, via Wikimedia Commons. Licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license. Panel j in public domain.

ABSORPTION. Clozapine is only available as an oral preparation. As with other antipsychotic agents, peak plasma levels are reached approximately 2 hours after oral administration (range, 1.1 ± 0.5 to 3.0 ± 1.5 hours). The elimination half-life is approximately 12 hours (between 10 and 16 hours; range, 6 to 33 hours). As a result, twice-daily dosing results in steady-state plasma concentrations in less than 1 week. Based on animal studies, coadministration of highly protein-bound drugs may increase free clozapine concentrations, although the total concentration (free plus bound) may be unchanged.

DISTRIBUTION. Clozapine has a large volume of distribution (mean of 2.0 to 5.1 L/kg [range, 1.0 to 10.2 L/kg]), although this is lower than other antipsychotic drugs.

METABOLISM AND ELIMINATION. Clozapine undergoes extensive first-pass

metabolism in liver and gut. Absolute bioavailability (percentage of clozapine reaching the systemic circulation unchanged) after oral administration varies from 27 to 47 percent. The primary metabolic routes are demethylation, *N*-oxidation, oxidation of the aromatic ring, and conjugation. Reportedly 80 percent of administered clozapine appears in the urine or feces as metabolites. Less than 5 percent of the parent compound is found unchanged in the urine.

Table 34.28–1.
Changes in Pharmacokinetics: Special Populations

Drug	Elderly	Renal Impairment	Hepatic Impairment	Genetic Variants of Metabolic Enzymes	Notes
Clozapine (Clozaril)	↓↓ Dosage	No δ dosage	↓ Dosage	Consider ↓ Dosage if CYP2D6 poor metabolizer	
Risperidone (Risperdal)	↓↓ Dosage	Dosage	↓ Dosage	Consider ↓ Dosage if CYP2D6 poor metabolizer	
Olanzapine (Zyprexa)	↓ Dosage (35%)	No δ dosage	No δ dosage		Clearance 30% lower in women and 40% higher in smokers
Quetiapine (Seroquel)	↓↓ Dosage	No δ dosage	Slight ↓ dosage		
Ziprasidone (Geodon)	No δ dosage	No δ dosage	↓ Dosage		
Aripiprazole (Abilify)	No δ dosage	No δ dosage	No δ dosage	↓ Dosage if CYP2D6 poor metabolizer	
Paliperidone (Invega)	No δ dosage	↓ Dosage	No δ dosage		
Iloperidone (Fanapt)	?	No δ dosage	No δ dosage	↓ Dosage if CYP2D6 poor metabolizer	
Asenapine (Saphris)	↓↓ Dosage	No δ dosage	Do not use with severe hepatic impairment		
Lurasidone (Latuda)	No δ dosage	↓ Dosage	↓ Dosage		
Cariprazine	No δ	CrCl <30 mL/min:	Severe impairment		

(Vraylar)	dosage	Use not recommended	(Child-Pugh class C): Use not recommended.	
Brexpiprazole (Rexulti)	No δ dosage	CrCl <60 mL/min: Maximum dose: Major depressive disorder: 2 mg Schizophrenia: 3 mg	Moderate to severe impairment (Child-Pugh class B or C): Maximum dose: Major depressive disorder: 2 mg Schizophrenia: 3 mg	å dose if poor CYP2D6 metabolizer

CYP2D6, cytochrome P450 system 2D6 polymorphism. Poor metabolizers indicated by this genetic polymorphism include ~8% of Caucasians and smaller proportions of Asians and African Americans. PD, pharmacodynamic; PK, pharmacokinetic; δ, change in; ↓,↓↓, decrease;?, uncertain.

Adapted from Ereshefsky L. Pharmacokinetics and drug interactions: Update for new antipsychotics. *J Clin Psychiatry*. 1996;57(Suppl):12 and information provided by manufacturers.

The *N*-desmethylclozapine metabolite represents 40 to 80 percent of the concentration of clozapine. It is a muscarinic type 1 (M_1) receptor agonist, whereas clozapine itself has muscarinic antagonist effects. According to Herbert Meltzer, this metabolite may be a major contributor to clozapine's purported superior effects on cognition and negative symptoms compared to the other agents.

Pharmacodynamics

MECHANISM OF ACTION. The explanation for clozapine's unique clinical profile remains controversial. Preclinical studies in rats have shown that clozapine blocks the conditioned avoidance response, indicating that the drug has potential antipsychotic activity. However, it did not produce catalepsy in animals, suggesting that it is unlikely to produce EPS in humans. This made it different from conventional antipsychotic medications, which all cause extrapyramidal side effects as a direct result of their D_2 receptor blockade in the forebrain.

The relatively low affinity for D_2 receptors may explain clozapine's lack of extrapyramidal side effects. In patients treated with most FGAs and SGAs, PET scan studies have shown that, with doses comparable to 10 mg of haloperidol (Haldol), more than 80 percent of striatal D_2 receptors are occupied. Higher doses do not increase receptor occupancy further. In clinically effective doses, clozapine only occupies 40 to 50 percent of D_2 receptors. This weak effect at the D_2 receptor site may be related to clozapine's unusual biological profile that combines a low D_2 affinity with activity at a number of other receptor sites, including serotonergic (e.g., 5-HT_2 and 5-HT_{1C}), adrenergic, and cholinergic sites.

Several hypotheses address the mechanism of action of clozapine.

Because its D_2 receptor activity is lower than that of traditional antipsychotic agents and its 5-HT_2 receptor blockade is stronger than that of haloperidol, Herbert Meltzer has proposed that the differential antipsychotic effect of clozapine is related to its low D_2 to 5-HT_2 ratio. Clozapine's affinity for the 5-HT_2 receptor is among the highest for antipsychotic agents. An alternative hypothesis proposes that clozapine is effective because it has a selective preference for the mesolimbic dopamine system. D_2 receptor blockade in the striatum caused by conventional agents may interfere with their antipsychotic effects and induce negative symptoms. Others have focused on clozapine's stronger 5-HT_3 blockade and its stronger α_1- and α_2-adrenergic receptor blockade. Other differences include its stronger affinities for histamine type 1 (H_1) receptors and muscarinic receptors but lower affinity for sigma receptors than haloperidol.

As mentioned above, the clozapine metabolite *N*-desmethylclozapine has M_1 agonist activity. Studies suggest that patients who have higher ratios of this metabolite to the parent compound tend to have better clinical outcomes. This suggests that M_1 agonism may contribute to clozapine's unique effectiveness.

Philip Seeman and Shitij Kapur have proposed that the lack of EPS with clozapine and other SGAs is related to their fast dissociation from the D_2 receptor. Conventional dopamine receptor antagonists tend to bind more tightly to dopamine receptors than dopamine itself. Clozapine and nearly all of the other SGAs bind more loosely and tend to come off the receptor more readily in the presence of dopamine. As a result, these newer agents may block these receptors more transiently, coming off the receptor to permit more normal dopamine transmission. Clozapine also has a 10-fold higher affinity for the D_4 receptor than do other antipsychotics. The D_4 receptor is present in the cortex and less so in the striatal areas. However, other agents with high affinity for D_4 receptors have failed to demonstrate that they are effective antipsychotics.

BLOOD CONCENTRATIONS AND RELATION TO ACTION. As with most antipsychotic drugs, patients display considerable variability in their clozapine plasma concentrations among patients receiving the same daily dose. This probably results from differences in both absorption and metabolism. Variability is also observed within patients, which is presumably related to variable absorption rates rather than changes in hepatic metabolism and clearance. Some reports show that women have slightly higher and smokers slightly lower plasma levels (20 to 30 percent). The plasma levels of older adults may be twofold higher than those of young adults.

Several investigators have studied the relationship between plasma

concentrations of clozapine and clinical response. Their findings suggest that monitoring clozapine plasma concentrations may be useful under certain conditions. A wide range of clozapine concentrations appears to be associated with a clinically meaningful response. Plasma concentrations of clozapine average approximately 10 to 80 ng/mL/mg of drug given per kilogram of weight. Thus, a typical daily dosage of 300 to 400 mg (approximately 5 mg/kg) is associated with plasma concentrations ranging from 200 to 400 ng/mL. Studies indicate that patients are more likely to respond when their clozapine plasma concentrations are >350 ng/mL. If patients with a plasma concentration of 250 ng/mL have not responded by 6 weeks, the clinician should consider adjusting the dosage to raise the concentration to approximately 350 ng/mL. The correlations between plasma concentration and clinical improvement or side effects are weak, as is observed with other antipsychotic agents.

Design and Interpretation of Clinical Drug Studies

Clozapine for Acute Schizophrenia. Open-label clinical trials conducted in the early 1970s suggested that clozapine was an effective antipsychotic drug that lacked extrapyramidal side effects. These findings were then reinforced with double-blind trials comparing clozapine with other agents. In 79 percent of the controlled trials that compared clozapine with another antipsychotic drug, clozapine was superior. In reviewing these studies, Ross Baldessarini cautions that the differences between clozapine and the comparator drug were relatively small. Nevertheless, certain patterns emerge. One comparison with chlorpromazine found a small advantage favoring clozapine when a heterogeneous group of patients was studied, but differences favoring clozapine were clearer when the most severely ill subgroup was studied. Another study compared haloperidol with clozapine in more severely ill and treatment-refractory patients and found a clear advantage of clozapine over haloperidol. One study comparing clozapine with risperidone found no differences in clinical response at the end of 6 weeks. Response to risperidone came sooner, but some extrapyramidal effects were observed with risperidone. Few treatment-refractory patients were included. In summary, these studies suggest that clozapine may be particularly effective for a subgroup of patients with schizophrenia who are both severely ill and refractory to treatment with conventional drugs. Whether clozapine is more effective in acutely psychotic patients who are likely to respond to conventional antipsychotics is unclear, although a study suggests that clozapine also has advantages in these patients. In a US multicenter study of patients who had tardive dyskinesia or were sensitive to extrapyramidal side effects, clozapine was superior to chlorpromazine.

Efficacy in Treatment-Refractory Patients. Early clinical experience with clozapine suggested that it was more effective than conventional drugs for

patients who were unresponsive to conventional antipsychotic agents. Moreover, the studies cited in the previous section suggested that the differences between clozapine and other antipsychotics were greater in patients who were severely ill and refractory to treatment with conventional dopamine receptor antagonists. Before marketing clozapine for these patients, the FDA and the manufacturer of clozapine developed a study comparing clozapine with chlorpromazine plus benztropine (Cogentin). The 268 patients who participated in this trial had previously responded poorly in at least three prior antipsychotic trials and a prospective single-blind trial of haloperidol. This study compared the effectiveness of chlorpromazine and clozapine in 248 patients who had thorough documentation of poor responsiveness to conventional antipsychotic agents. During this 6-week trial, the clozapine-treated patients did better on a range of measures of both positive and negative symptoms. Thirty percent of clozapine-treated patients met a priori criteria for significant improvement, in contrast to only 4 percent of chlorpromazine-treated patients. Other studies show that this 6-week trial may have significantly underestimated the proportion of patients who respond to clozapine, and that up to 60 percent of patients may display substantial improvement after a 6-month trial. Clozapine in this trial was superior to chlorpromazine on 14 of 18 Brief Psychiatric Rating Scale (BPRS) items, as well as on the Clinical Global Impression (CGI) and nursing ratings such as social competence, social interest, personal neatness, irritability, manifest psychosis, and motor retardation. When patients were classified as having improved (by prospective criteria) to a clinically significant extent, 30 percent of clozapine patients improved within 6 weeks versus only 4 percent of those receiving chlorpromazine plus benztropine ($p < 0.001$). Based on this trial, the FDA approved clozapine for limited use in the United States.

As noted in another section of this chapter, the NIMH CATIE study compared clozapine (open label) to double-blind risperidone, olanzapine, or quetiapine in patients who had their phase 1 medication changed for lack of efficacy. Patients assigned to clozapine had the best outcome with 56 percent of clozapine patients being discontinued compared to 71 percent on olanzapine, 86 percent on risperidone, and 93 percent on quetiapine. These results were also supported by a study from the UK by Shon Lewis that compared clozapine to an SGA selected prior to the randomization and found an advantage for clozapine.

The effectiveness of clozapine in treatment-refractory schizophrenia has also been compared to the use of SGA polypharmacy, which is a nonrecommended strategy that is often used in this population. In this study, the utilization of emergency department and hospital services, as well as cost outcomes, was compared between Medicaid beneficiaries receiving clozapine monotherapy and those receiving two or more SGAs. It

was found that Medicaid beneficiaries receiving clozapine monotherapy between 2006 and 2009 (N = 479) had a lower risk for emergency department use as well as lower Medicaid payments (both all-cause and mental disorder related) than patients receiving SGA polypharmacy.

Altogether, these studies support the view that patients with schizophrenia who continue to demonstrate psychotic symptoms that are bothersome should not be considered refractory to treatment unless they have had an adequate trial with clozapine. An adequate trial is probably at least 3 months of clozapine at an adequate dose.

Long-Term Effectiveness. Long-term studies have supported the concept of maintaining patients on clozapine. A 29-week study reported in 2001 found that 60 percent of moderately refractory outpatients with schizophrenia treated with clozapine met improvement criteria compared with 12 percent of haloperidol-treated patients. In a study of patients with schizophrenia in Connecticut state hospitals, those who were discharged on clozapine treatment were more likely to remain out of the hospital when compared with individuals discharged on treatment with other agents. A Department of Veterans Affairs Cooperative Study comparing 1 year of treatment with clozapine to that of haloperidol also found that clozapine was associated with a decreased need for hospitalization. Moreover, patients treated with clozapine demonstrated a greater likelihood of participating in psychosocial treatments and rehabilitation programs. As mentioned in an earlier section, there is evidence that patients who are unresponsive to treatment in adequate trials of conventional dopamine receptor antagonists may respond when treated with SGAs other than clozapine. As a result, patients who respond poorly to conventional dopamine receptor antagonists may benefit from a trial of at least one of these agents before receiving a clozapine trial. Clinicians should also consider that prolonged sequential trials with a number of antipsychotics may take months or even years, thus delaying treatment with an agent that may be effective.

Severe Tardive Dyskinesia. Patients with severe tardive dyskinesia should also be considered for a clozapine trial. Most clinical reports indicate that clozapine can suppress abnormal movements in tardive dyskinesia. This is not surprising, as conventional antipsychotic agents are effective for the short-term suppression of tardive dyskinesia. However, clozapine is probably less likely to cause the disorder and, therefore, is promising for the long-term management of patients, particularly those who already have tardive dyskinesia. Although rare cases of tardive dyskinesia have been found in patients who receive clozapine, the risk of developing this disorder is substantially reduced. Studies using animal models indicate that clozapine does not result in the increased sensitivity of

striatal D_2 receptors that is associated with typical antipsychotic agents and has been theorized to result in tardive dyskinesia. It remains to be established whether a positive effect on tardive dyskinesia results from a direct suppressive effect of clozapine on certain transmitter systems or whether clozapine spares the striatal D_2 receptors, allowing spontaneous remission of tardive dyskinesia. The former interpretation is supported by the reports of a return of tardive dyskinetic movements after clozapine withdrawal.

Clozapine in Hostile and Aggressive Patients. Clozapine has also been studied in patients with schizophrenia and other psychotic disorders who display hostile and aggressive behaviors. A study of 157 inpatients by Citrome found that clozapine treatment led to greater reductions in ratings of hostility than other antipsychotics. Similarly, Chengappa found that clozapine led to a reduction in the need for seclusion and restraint. These findings suggest that clozapine may be of particular benefit to refractory patients who demonstrate hostile and aggressive behaviors.

Patients at Risk for Suicide. Clozapine may also be a preferred agent for patients with schizophrenia who are at a higher risk of suicide. Large epidemiological studies found what appeared to be a reduction in death from suicide in patients receiving clozapine. This and other observations led Herbert Meltzer and colleagues to design the International Suicide Prevention trial (or InterSePT). The study compared clozapine and olanzapine in 956 patients with schizophrenia or schizoaffective disorder who were considered at risk for suicide. Clozapine was more effective in delaying the time to recurrent suicidal behavior. In addition, over a 2-year period, 24 percent of clozapine and 32 percent of olanzapine patients were hospitalized for suicidal risk.

Schizophrenia with Substance Abuse. Although the effectiveness of clozapine in patients with comorbid substance abuse has not been demonstrated by randomized controlled trials, there is some supporting evidence from naturalistic studies. A retrospective study by Green and coworkers found that patients treated with clozapine were more likely to abstain from alcohol and cannabis use than patients treated with risperidone. A prospective study by the same investigators found that patients treated with clozapine were better able to reduce their substance abuse.

Treatment-Resistant Mania. Clozapine can be an effective treatment of manic symptoms that are not responsive to other agents. Early clinical case reports indicated that clozapine was effective both as monotherapy and as an adjunct to mood-stabilizing agents in this population. An open-label,

randomized trial in acutely manic patients found that clozapine had a more rapid onset of effect than that of chlorpromazine. In another open-label trial involving patients with schizoaffective disorder or bipolar disorder who were randomized to either clozapine or treatment as usual, clozapine-treated patients demonstrated greater improvements. Although the available data support clozapine's effectiveness in mania, none of these trials have used a double-blind design.

Severe Psychotic Depression. The effectiveness of clozapine in the treatment of patients with major depression with psychotic features is supported by clinical case reports alone.

Neurological Illnesses. Clozapine has also found uses in patients with neurological disorders. It has been used in idiopathic Parkinson disease as a treatment of psychotic symptoms that are secondary to dopaminergic agents that are used to treat the disease. These patients may respond to relatively low dosages of clozapine (doses of 25 to 75 mg). Higher dosages (e.g., 100 to 250 mg a day) may exacerbate the parkinsonian symptoms. The effectiveness of clozapine in patients with idiopathic Parkinson disease is supported by at least two double-blind studies. A report from the American Academy of Neurology recommends clozapine as a preferred agent for psychosis in Parkinson disease. Patients with benign essential tremor or with Huntington disease have also responded positively to clozapine.

Polydipsia–Hyponatremia Syndrome. Case reports suggest that clozapine may have advantages over other antipsychotic agents in patients who tend to intoxicate themselves with water. This syndrome—called *polydipsia–hyponatremia syndrome*—can result in seizures.

Treatment Resistance in Other Disorders. Treatment-resistant patients with a pervasive developmental disorder, autism of childhood, or OCD have responded well to clozapine, either by itself or in addition to their regular treatment.

Effects on Specific Organs and Systems

General. Sedation is the most common adverse effect of clozapine treatment. It occurs early in treatment, and tolerance usually develops after the first few days or weeks. Giving the larger part of the daily dose at bedtime may reduce daytime sedation. However, many individuals are still unable to tolerate the sedative effects and should, therefore, be treated with another antipsychotic. Dizziness or vertigo may be related to orthostatic hypotension.

Disturbances in Temperature Regulation. Benign hyperthermia is occasionally observed during the first 3 weeks of clozapine treatment, with a peak incidence on the tenth day. The increase in temperature is usually not more than 1 to 2°F and resolves over a few days. Occasionally, however, temperature readings above 101°F have been seen, which require temporary interruption of clozapine treatment and frequent hematological monitoring. Differential diagnoses include drug fever, intercurrent infection, infection secondary to agranulocytosis, dehydration, heat stroke, lethal catatonia, and, possibly, NMS. Patients who develop a high temperature can be given the drug again but with a more gradual increase in dosage. If this strategy does not work, clozapine may need to be discontinued. As is also seen with chlorpromazine, mild hypothermia is frequently observed (87 percent) after clozapine treatment. This common effect may indicate the poikilotherm effect of antipsychotic agents, presumably mediated by the hypothalamus.

Cardiovascular. The cardiovascular side effects most frequently observed with clozapine are tachycardia and postural hypotension. The tachycardia is probably a direct effect of the vagolytic properties of the drug. It can be present in the supine position and is thus not the result of orthostatic changes. Increased pulse rates of 20 to 25 beats per minute are encountered when clozapine reaches 300 mg per day or more over a 7-day period (e.g., sinus tachycardia). Tachycardia related to clozapine may be treated with a β-blocker such as atenolol (Tenormin), although patients should then be monitored for hypotension. Propranolol (Inderal) should not be prescribed because it may increase the risk of agranulocytosis.

Reversible nonspecific ST-T wave changes, T-wave flattening, or inversions (repolarization effects) are seen infrequently with clozapine and are usually of no clinical significance. These changes are similar to those seen with other antipsychotic agents and are dose dependent. Clozapine can increase the QT interval, although there are very limited data to determine if these increases are associated with a risk of ventricular arrhythmias. Orthostatic hypotension is more likely to occur during initial titration of clozapine in association with rapid dose escalation but may occur after the first dose. In one report, initial dosages of clozapine as low as 12.5 mg were associated with orthostatic collapse and respiratory arrest. The risk of postural hypotension progressing to orthostatic collapse is substantially greater if the initial dosage of clozapine exceeds 75 mg per day. Tolerance to the hypotensive effects of clozapine often develops over time, and the effects can usually be managed if the initial dosage is low (e.g., 25 mg per day) and is increased gradually. When hypotension occurs, lowering the dosage to the previous level and slowing the rate of upward titration usually limits problems. Support stockings, increased sodium intake, or fludrocortisone (Florinef) may be helpful. Rarely (approximately

1 case per 3,000 or 4,000 patients), orthostatic collapse can be profound, accompanied by respiratory or cardiac arrest.

Paradoxically, hypertension has also been observed with clozapine (less than 4 percent). Clozapine induces significant changes in hemodynamic mechanisms; it should be used with caution in patients with pre-existing cardiovascular disease, such as a history of myocardial infarction or of arrhythmias. The authors recommend careful monitoring of vital signs during the first few weeks of treatment.

Clozapine has been associated with idiopathic myocarditis. The section on warning and precautions below will discuss this further.

Gastrointestinal. Constipation, most probably due to the antimuscarinic effects of clozapine, is the most common gastrointestinal adverse effect and, if left untreated, can progress to intestinal obstruction. Clinicians treating patients with clozapine should monitor the presence of constipation and intervene proactively at the earliest sign that this symptom has developed. Interventions may include modifications in the patient's diet, use of stool softeners, increased fluid intake, and exercise.

Nausea and vomiting may occur during the initial stages of clozapine treatment. Patients who experience the side effects can often, but not always, be managed by dosage reduction.

Urogenital. Urogenital effects include enuresis, frequency and urgency, hesitancy, urinary retention, and impotence. All of these are independent of dose or blood concentrations of clozapine. Urinary incontinence or enuresis can be managed by setting an alarm clock to ring at night or administering intranasal desmopressin (DDAVP) or oxybutynin (Ditropan) (5 to 15 mg per day). So far, a few cases of clozapine treatment-associated priapism have been noted. Because patients may not volunteer these complaints, clinicians should ask about them.

Therapeutic Indications

Treatment-Resistant Schizophrenia. Clozapine is indicated for patients who have failed trials with other antipsychotics. Given the serious side effects associated with clozapine treatment, including seizures and agranulocytosis, most guidelines recommend a trial of a minimum of two other antipsychotics, one of which should be an SGA. Studies have found that clozapine's advantage when compared with other antipsychotics extends from patients who continue to experience severe symptoms to those who have moderate symptoms.

Reduction in the Risk of Recurrent Suicidal Behavior in Schizophrenia or Schizoaffective Disorder. The InterSePT study as well as epidemiological data indicate that clozapine is indicated for those patients who are at risk of

suicidal behaviors. It is the only antipsychotic with this particular indication.

Precautions and Adverse Reactions

The only contraindications to the use of clozapine are history of hypersensitivity to clozapine or tablet components, a WBC count below 3,500 cells per mm^3, a previous bone marrow disorder, a history of agranulocytosis during clozapine treatment, or the concomitant use of another bone marrow suppressant drug such as carbamazepine (Tegretol). History of myocarditis is not a labeled contraindication, but the authors recommend against use of clozapine in these patients. Although the mechanism of action is unknown, there is a slight risk of respiratory depression or collapse if treatment is initiated while patients are taking benzodiazepines.

Black Box Warnings

AGRANULOCYTOSIS. The risk of clozapine-induced agranulocytosis is 0.73 percent during the first year of treatment and 0.07 percent in the second year. Review of the clozapine registry indicates that the incidence may be as low as 0.038 percent. The risk of this adverse effect is greatest during the first 3 months of treatment, but a significant risk remains during the first year of treatment. For neutropenia, the risk is 2.32 percent in the first year and 0.69 percent in the second year.

Increased familiarity of clinicians with agranulocytosis (absolute neutrophil count [ANC] less than 500 cells per mm^3 or WBC less than 1,000 cells per mm^3) decreased the incidence of agranulocytosis by more than 50 percent from before 1995 to after 1997.

Agranulocytosis due to clozapine is a potentially fatal condition that requires immediate medical attention. The risk of agranulocytosis increases with age and is higher in women. Therefore, before treating patients with clozapine, clinicians must register patients with a clozapine monitoring system. The United Kingdom and Australia require, in addition to the 3,500 WBC per mm^3, that granulocyte count (ANC) is 2,000 cells per mm^3 or higher when clozapine is initiated. Patients who are treated with clozapine in the United States must receive the following monitoring protocol:

- WBC is no longer required to be monitored—only ANC.
- Initial ANC must be $\geq 1{,}500/mm^3$, and in patients with benign ethnic neutropenia it must be $\geq 1{,}000/mm^3$.
- Monitoring is weekly for the first 6 months, every 2 weeks for the second 6 months, and monthly after 12 months. The threshold minimum ANC that is not classified as neutropenia is now $1{,}500/mm^3$ rather than

1,000/m3.

The requirement for ANC monitoring is continued for as long as patients receive the drug and for at least 1 month after it is discontinued. Current guidelines specify that any fever or sign of infection (e.g., pharyngitis) is an immediate indication for a WBC count, particularly in the first 18 weeks of treatment. The revised requirements indicate that clozapine treatment should be interrupted for ANC <1,000/mm3. Reports of clozapine re-exposure after hematologic recovery have shown that all such patients again experienced agranulocytosis, but sooner and at lower doses than during initial treatment. As a result, patients who develop agranulocytosis should not receive clozapine again. Although a WBC count below 3,500 cells per mm3 is not a proven risk factor, clozapine should be administered cautiously in patients with such low counts because of the difficulty in early identification of agranulocytosis. Rechallenge is not dictated by WBC any longer, only ANC. Rechallenge is allowed following recovery from moderate neutropenia (500 to 999/mm3) but not after recovery from severe neutropenia (<500/mm3) unless prescriber determines benefits outweigh risks. Clozapine therapy should be discontinued whenever the total WBC count falls below 3,000 cells per mm3 or the granulocyte count falls below 1,500 cells per mm3.

Other hematological changes have been reported with clozapine, including leukocytosis (0.6 percent), eosinophilia (1 percent), and leukopenia, neutropenia, decreased WBC count (3 percent), and, rarely, thrombocytopenia. If the total eosinophil count rises above 4,000, clozapine should be discontinued until the count falls below 3,000.

ORTHOSTATIC HYPOTENSION. On rare occasions clozapine can cause severe episodes with orthostatic hypotension and possible syncope. This has been associated with cardiac or respiratory arrest. These serious events are more likely to occur early in treatment when the dose is being increased. Also, these events have been associated with benzodiazepine use and suggest that combining clozapine and benzodiazepines early in treatment should be done cautiously. Dosage should be increased slowly, with divided doses, and clozapine should be used with caution in patients with cardiovascular or cerebrovascular disease or conditions predisposing them to hypotension.

SEIZURES. Five to 10 percent of patients treated with clozapine will develop seizures. The risk is clearly related to dosage, with an increased risk at dosages greater than 600 mg. Rapid escalation of dosage during the first weeks of clozapine treatment may increase the likelihood of seizures. A pre-existing seizure disorder or head trauma places the patient at even greater risk.

A number of strategies may be effective for managing the risk of seizures. Monitoring clozapine plasma concentrations in patients receiving

higher dosages may ensure that concentrations do not become excessive. Patients with a history of head injury should have baseline electroencephalographs (EEG) performed before the initiation of clozapine. Clinicians may also consider an EEG before increasing clozapine dosages above 600 mg. In addition, when patients receive higher doses or have abnormal EEGs, adding an anticonvulsant such as sodium valproate (Depakene) may be helpful.

The occurrence of seizures in patients treated with clozapine should not be viewed as a clear reason to stop clozapine. Lowering the dosage and adding an anticonvulsant is usually sufficient to prevent further seizures. Sodium valproate, gabapentin (Neurontin), and topiramate (Topamax) have been effective in treating clozapine-induced seizures. Carbamazepine should not be prescribed because it can also cause neutropenia or granulocytopenia. Once patients have begun anticonvulsant therapy, the dosage of clozapine can be retitrated to an optimal level.

MYOCARDITIS AND CARDIOMYOPATHY. Clozapine is associated with an increased risk of myocarditis. Although the incidence of myocarditis is very low (approximately 2.8/100,000 patient-years), it can result in death. Clinicians should consider the possibility of myocarditis when patients receiving clozapine present with unexplained chest pain, tachycardia, palpitations, dyspnea, fever, flu-like symptoms, or hypotension. Unfortunately, some of these symptoms are common in individuals who are in the early stages of clozapine treatment, and 80 percent of clozapine-induced cases of myocarditis occurred during the first 6 weeks of treatment. Because mortality from myocarditis approaches 40 percent, early detection and treatment are important. If this syndrome is suspected, clozapine should be discontinued and evaluations including an ECG, creatine kinase level, and erythrocyte sedimentation rate should be performed.

INCREASED MORTALITY IN ELDERLY PATIENTS WITH DEMENTIA. Clozapine and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis compared with placebo.

Other Adverse Events

SIALORRHEA. Sialorrhea is an uncomfortable adverse effect that develops early in treatment. Hypersalivation is most profound during sleep, and patients may complain that their pillow is wet in the morning. Placing a towel over the pillow may be of some help. Clonidine (Catapres) (0.1- or 0.2-mg patch once a week) or amitriptyline (Elavil, Endep) at bedtime has been used to treat sialorrhea. Studies suggest that clozapine does not increase saliva flow but decreases swallowing, which becomes more pronounced during sleep.

PERIPHERAL ANTICHOLINERGIC EFFECTS. Dry mouth, blurred vision, constipation, and urinary retention are commonly observed with clozapine. Constipation can be severe, particularly when patients fail to exercise or drink sufficient fluids. Other side effects of the autonomic nervous system, such as increased sweating, have also been reported.

WEIGHT GAIN. Clozapine can be associated with substantial weight gain; a meta-analysis of short-term trials with clozapine found an average weight gain of 4.45 kg over 10 weeks. This exceeded the weight gain observed with all of the other medications evaluated in the trials. Other studies indicate that patients can continue to gain weight over a period of 5 years. The amount of weight that patients gain can be substantial and can result in some patients becoming obese. Younger patients, including adolescents, may be more vulnerable to gaining weight on clozapine.

These observations demonstrate the importance of preventing substantial weight gain for patients receiving clozapine. Physicians should weigh patients frequently during the first months of clozapine treatment and should encourage patients to monitor their own weight. If weight gain is observed, it is important for clinicians to intervene with counseling about nutrition, weight, and exercise. If these interventions are unsuccessful, changing medication to an agent less likely to cause weight gain should be considered.

DIABETES MELLITUS. Clozapine, along with many other SGAs, can increase the risk of type 2 diabetes mellitus. In a naturalistic study by David Henderson that monitored 81 patients over a 5-year period, 36.6 percent developed diabetes. Case reports suggest that new cases of diabetes usually occur during the first 6 months of clozapine treatment. Michael Sernyak and his coworkers examined the records of more than 30,000 veterans and found that the prevalence of diabetes was higher in patients receiving clozapine, olanzapine, or quetiapine. It is also concerning that a number of case reports indicate that patients receiving clozapine developed diabetic ketoacidosis. Although these forms of evidence remain inconclusive, they strongly suggest that clinicians should be aware of the possible association of clozapine with an increased risk of diabetes. This may lead clinicians to monitor patients for diabetes, with fasting blood glucose determinations at regular intervals. It may also be helpful to inquire if patients are experiencing clinical signs of diabetes, such as polydipsia, polyuria, or weight loss.

LIVER FUNCTION DISTURBANCES. Liver function disturbances are generally mild and transient.

EXTRAPYRAMIDAL SIDE EFFECTS. Clozapine has a much lower incidence of extrapyramidal side effects than other antipsychotics. Akathisia (6

percent), tremors (6 percent), and rigidity (3 percent) are considerably less frequent with clozapine than with FGAs. The authors are not aware of reports of acute dystonia with clozapine. Case reports suggest that tardive dyskinesia may improve with clozapine.

NEUROLEPTIC MALIGNANT SYNDROME. Consistent with clozapine's low affinity for D_2 receptors, there have been few reports of NMS developing in patients who were treated solely with clozapine, although case reports of this syndrome developing in patients when concomitantly treated with lithium (Eskalith) exist. Although two patients had experienced NMS with conventional antipsychotics before, one case has been reported of a patient with a history of NMS that was successfully treated with clozapine. The authors recommend, therefore, that lithium should not be used as adjunctive treatment with clozapine in patients who have previously experienced NMS.

Because clozapine treatment is accompanied by side effects that resemble several signs and symptoms of NMS (e.g., hyperthermia, cardiovascular effects, delirium, increased sweating, elevations in creatine phosphokinase, and leukocytosis), clinicians should be alert for the development of a potential NMS during clozapine treatment. Some case reports have described NMS without rigidity in patients treated with clozapine, although the lack of rigidity could preclude the diagnosis of this syndrome.

PREGNANCY AND LACTATION. Clozapine is rated as pregnancy category B. Studies have not revealed evidence that clozapine causes harm to the fetus. However, there are no adequate and well-controlled studies in pregnant women. Studies in animals suggest that clozapine is excreted in breast milk. As a result, women taking clozapine should not breastfeed.

Discontinuing Clozapine

If a decision is made to change a patient to a different antipsychotic, it should not be made abruptly unless patients have developed agranulocytosis. In some individuals, the sudden discontinuation of clozapine has led to a severe exacerbation of their underlying illness. In addition, withdrawal of a potent antimuscarinic agent such as clozapine can lead to cholinergic rebound, with symptoms such as nausea and vomiting. For these reasons, the authors recommend that patients are gradually cross-tapered from clozapine to another antipsychotic agent over 2 to 4 weeks.

Drug Interactions

Because the cytochrome P450 (CYP) mitochondrial system is involved in the metabolism of clozapine, ethnic differences and interactions with drugs

that inhibit or stimulate this system will affect clozapine plasma concentrations. For instance, caffeine *N*-3-demethylation, which is a measure of CYP isoenzyme 1A2 (CYP1A2) activity, accounts for 70 percent of the variance in clozapine clearance. This suggests that administration of fluvoxamine (Luvox), a potent inhibitor of CYP1A2, will elevate clozapine concentrations, whereas fluoxetine (Prozac) administration may only lead to changes in higher dosage ranges. Cimetidine (Tagamet), SSRIs, tricyclic drugs, and valproic acid decrease clearance, presumably through inhibition of CYP1A2 and CYP2D6, whereas phenytoin (Dilantin) and carbamazepine induce CYP2C and CYP3A4 hepatic isoenzymes and may decrease concentrations of clozapine.

Furthermore, when risperidone, which weakly inhibits CYP2D6, is added, clozapine concentrations may increase. Erythromycin (E-Mycin) and ketoconazole (Nizoral) may increase clozapine concentrations.

Laboratory Interference

Clozapine does not appear to interfere with other laboratory tests.

Dosage and Administration

Initiating Treatment. Clozapine treatment should begin with 12.5 to 25 mg per day and increased to 25 or 50 mg per day on the second day. The daily dose, which should be administered at least twice daily initially, can usually be increased by 25 mg every 1 or 2 days until a target dose of 300 to 450 mg per day is reached. The rate of increase is often limited by sedation and orthostatic hypotension. Because the titration phase for clozapine may require several weeks in some individuals, these patients should be continued on their current antipsychotic agent, which should be gradually decreased as the clozapine dosage increases. Patients who fail to respond to respond in the 300 to 450 mg per day range and who do not have serious side effects can be increased gradually to doses as high as 900 mg per day. However, clinicians and patients should be aware of the risk of seizures at higher levels and when doses are escalated more rapidly.

Duration of Treatment. Once the appropriate dosage has been reached and the patient improves, the medication should be continued. There is evidence that patients who receive clozapine will continue to improve over a 3- to 6-month period. If there are concerns that patients are experiencing excessive side effects or an inadequate response, measurement of plasma concentration may be helpful. Levels above 350 ng/mL are an indication that patients are receiving adequate dosages of clozapine.

RISPERIDONE

History

During the early 1980s scientists at Janssen Pharmaceutica were interested in developing an antipsychotic that would have a clinical profile similar to clozapine but without a tendency to cause agranulocytosis. The company developed a number of drugs with 5-HT_2 antagonism with and without D_2 activity. One compound, setoperone, had weak dopamine activity and was a relatively weak antipsychotic. Ritanserin was developed as a 5-HT_2 antagonist (inverse agonist) that could be added to a D_2 antagonist to decrease the amount of extrapyramidal side effects. Although ritanserin had anti-negative symptom effects, improved sleep, and some weak antipsychotic effects, it also had significant effects on QTc interval. The synthesis of risperidone in 1984 led the company to focus on the development of this agent since it combined strong D_2 and 5-HT_2 antagonism. Subsequent safety and efficacy studies with risperidone led to its approval in the United States in 1994.

Risperidone was the first SGA medication to be approved following clozapine. Whereas clozapine was reserved for patients with illnesses that were poorly responsive to available antipsychotic drugs, risperidone was a first-line antipsychotic that could be administered to nearly every patient with a psychotic illness. Following its introduction it was hoped that risperidone would share clozapine's efficacy for a broad range of psychopathology in schizophrenia and that it would be more effective than the FGAs. Although this promise has not been met, risperidone has been widely accepted and is among the most prescribed antipsychotics worldwide.

Chemistry

Risperidone is a benzisoxazole derivative (Fig. 34.28–1). The molecular formula is $C_{23}H_{27}FN_4O_2$ and its molecular weight is 410.49.

Pharmacological Action

Pharmacokinetics

ABSORPTION AND DISTRIBUTION. Risperidone has a bioavailability of 70 percent, and studies show that all oral forms of risperidone are bioequivalent. Risperidone is metabolized in the liver to 9-hydroxyrisperidone (or paliperidone [Invega]), which has generally the same pharmacological profile as the parent compound. After ingestion, peak plasma levels of the parent compound occur within 1 hour and within 3 hours for 9-OH-risperidone. Steady state is expected to be reached by 5 days (with a range of 1 to 5 days, depending on the CYP2D6 activity in the individual patient). Food does not affect the rate or extent of absorption in the gut.

The volume of distribution is 1 to 2 L/kg. Risperidone is 90 percent

plasma protein bound, and its metabolite is 77 percent protein bound. However, there is only a slight increase in the free fraction of risperidone and paliperidone with high therapeutic doses of warfarin (Coumadin) and other highly protein-bound medications.

METABOLISM AND ELIMINATION. Risperidone undergoes extensive metabolism, primarily by CYP2D6 (see section on drug interactions). The major metabolite is 9-OH-risperidone, which has similar pharmacologic activity as risperidone. The half-life of risperidone is 3 hours in rapid metabolizers and 20 hours in poor metabolizers. The half-life of 9-OH-risperidone is 21 hours in rapid metabolizers and 30 hours in poor metabolizers. The combined half-life of risperidone and its metabolites has a mean of 20 hours. Seventy percent of risperidone is excreted in the urine.

Pharmacodynamics

MECHANISM OF ACTION. Risperidone has high affinity for dopamine D_2 receptors (Table 34.28–2) and for serotonin 5-HT_{2A} receptors, and its potent antagonism at these receptors is thought to mediate its therapeutic effects. Risperidone also demonstrates high affinity for adrenergic α_1- and α_2-receptors and histaminergic H_1 receptors. It has moderate affinity for serotonin 5-HT_{1C}, 5-HT_{1D}, and 5-HT_{2A} receptors, and weak affinity for dopamine D_1 receptors. Risperidone has no affinity for cholinergic muscarinic receptors or adrenergic β_1 and β_2 receptors. Although risperidone has high affinity for D_2 receptors, it does not have the rates of EPS that the high-potency FGA medications have. This is most likely due to the dopamine-promoting effects of 5-HT_{2A} antagonism. Risperidone blocks 65 percent of D_2 receptors (the lowest threshold percentage for antipsychotic efficacy) at an average dose of 2 mg per day. At an average of 6 mg per day, 80 percent of D_2 receptors are blocked, and EPS may occur. At the 2-mg dose, 5-HT_{2A} effects may not be optimal.

Table 34.28–2.
Side Effects Associated with Receptor Antagonism

Item	D_2	D_1	5-HT_{2A}	H_1	`1	`2	ACh
Extrapyramidal symptoms (parkinsonism, dystonia, akathisia)	+	0	−	0	0	0	−
Tardive dyskinesia	+	0	0	0	0	0	0
Sedation, somnolence	+	0	0	+	+	+	+
Neuroleptic malignant syndrome	+	0	0	0	0	0	0
Orthostatic hypotension	0	0	0	+	+	+	+
Tachycardia	0	0	0	+	+	+	+

Anticholinergic side effects (dry mouth, blurred vision, constipation, difficulty urinating)	0	0	0	+	0	0	+
Prolactin increase	+	0	0	0	0	0	0
Decreased ejaculatory volume	+	0	0	0	0	0	0
Weight gain	0	0	+	+	0	0	0
Sexual dysfunction	+	0	0	0	0	0	+

BLOOD CONCENTRATIONS AND CLINICAL ACTIVITY. The clinical activity of risperidone results from the combination of risperidone and 9-OH-risperidone. No clear evidence exists supporting the usefulness of measuring these levels in managing patients on oral or long-acting risperidone. As the ratio of 5-HT_{2A} over D_2 binding is constant across doses for the risperidone molecule, the individual antipsychotic response and side-effect profile may vary among patients based on their own 5-HT_{2A} and D_2 receptor activities (e.g., for some patients 6 mg may provide too much D_2 but not enough 5-HT_{2A} antagonism, while others may respond to only 2 mg or up to 10 mg per day).

Design and Interpretation of Clinical Drug Studies

Design. Since risperidone has become available, there have been numerous studies evaluating its efficacy, safety, and tolerability. In this section, the focus mostly will be on randomized double-blind, placebo- or active-controlled trials.

SCHIZOPHRENIA AND PSYCHOSIS: ACUTE TREATMENT. Multiple studies with brief treatment trials have confirmed risperidone's efficacy in treating the acute phase of psychosis associated with schizophrenia, schizoaffective disorder, schizophreniform disorder, and acute psychotic disorder. Risperidone has consistently been shown to be superior to placebo, superior or equivalent to FGAs such as haloperidol and perphenazine, and equivalent to other SGAs with the exception of clozapine. Early multicenter trials in the United States and Canada compared risperidone 2, 6, 12, or 16 mg, haloperidol 20 mg, or placebo in a combined total of 513 patients with schizophrenia. The results from the US and Canadian studies were very similar. In the Canadian study, risperidone 2 mg per day was not shown to be superior to placebo on measures of efficacy. Risperidone was superior to placebo at all higher doses and superior to haloperidol on a variety of measures, most noticeably the Positive and Negative Syndrome Scale (PANSS). Risperidone 6 mg was as efficacious as the higher doses. In the US sample, there was no significant difference in the rates of EPS between patients on risperidone 6 mg and placebo. There was a linear trend toward more EPS as risperidone doses increased. A factor analysis using the combined results of the Canadian and US placebo-controlled studies examined the effects of

risperidone on the five dimensions of schizophrenia (negative symptoms, positive symptoms, disorganized thought, uncontrolled hostility and excitement, and anxiety and depression). Risperidone (at all dosages studied) produced significantly greater improvements than haloperidol on all dimensions of schizophrenia ($p \leq 0.05$). The difference was most pronounced on three dimensions—hostility and excitement, anxiety and depression, and negative symptoms. Another large multicenter, multinational study of 1,362 patients with schizophrenia compared risperidone 1, 4, 8, and 12 mg per day to haloperidol 10 mg for 8 weeks. At doses of 4 and 8 mg, the response rates were 63 percent and 66 percent, respectively. There was no significant difference between response to haloperidol (response rate 59 percent) and response to risperidone. Risperidone was associated with lower EPS side effects at all doses compared with haloperidol. An 8-week study comparing risperidone (dose range 2 to 6 mg per day) to olanzapine (dose range 5 to 20 mg per day) in 377 patients with schizophrenia found similar improvements in PANSS scores between the two drugs. Risperidone showed a slight advantage in positive symptom and anxiety or depressive symptom reduction. Rates of EPS were similar between the two drugs, and rates of weight gain of over 7 percent of baseline body weight were 12 and 27 percent for risperidone and olanzapine, respectively. A Cochrane review of risperidone versus olanzapine treatment concluded that the two medications have equivalent short-term efficacy, and that more patients gained weight or experienced metabolic side effects on olanzapine than on risperidone. Both medications were associated with EPS side effects; 25 percent of risperidone users required anticholinergic medication versus 18 percent of people on olanzapine. Two studies comparing risperidone to quetiapine found risperidone improved PANSS scores significantly more than did quetiapine at 8 and 12 weeks.

Taken together these studies indicate that risperidone at doses of 4 to 8 mg has similar effectiveness to other SGAs. Using group data, the findings suggest that there is little advantage to increasing the dose above 4 to 6 mg. This observation, and the tendency of risperidone to be associated with dose-dependent EPS, has led many clinicians to target risperidone in the 3 to 5 mg range.

SCHIZOPHRENIA AND PSYCHOSIS: MAINTENANCE TREATMENT. In a randomized, nonblinded study of 184 patients, Bouchard and colleagues compared the long-term (12 months) efficacy of risperidone and conventional neuroleptics for treating patients with schizophrenia. After 12 months' treatment, risperidone (mean dose of 5.5 mg per day) demonstrated significantly greater clinical improvement than did conventional neuroleptics, both in terms of average change in score from baseline on the PANSS ($p = 0.006$) and the proportion of good responders (defined as

those with at least a 20 percent decrease in PANSS total scores, $p = 0.03$). Twice as many risperidone patients responded (30 percent) as patients treated with conventional neuroleptics (15 percent, $p = 0.03$). In a double-blind clinical trial of 555 first-episode schizophrenia patients randomized to haloperidol or risperidone for long-term treatment, patients in both groups showed similar levels of clinical improvement. Time to relapse was significantly longer with risperidone treatment (median = 466 days) versus haloperidol treatment (median = 205 days), and fewer patients relapsed on risperidone than on haloperidol. Marder and colleagues compared the effectiveness and tolerability of risperidone (mean dose 5.7 mg per day) and low-dose haloperidol (mean dose 4.5 mg per day) in 63 patients over 2 years. There were no significant differences between the treatment conditions for the risk of psychotic exacerbation or study discontinuation. Both groups showed improvements on the BPRS over time, although there was a significant advantage to risperidone treatment of the anxiety/depression cluster. Furthermore, risperidone-treated subjects had lower rates of EPS than subjects on low-dose haloperidol.

SCHIZOPHRENIA AND PSYCHOSIS: ADOLESCENTS. Two trials evaluating the efficacy of risperidone in adolescents aged 13 to 17 with schizophrenia compared risperidone at doses ranging from 0.15 to 6 mg per day with placebo. Risperidone treatment significantly improved PANSS total scores. Doses below 0.6 mg per day were less effective than higher doses. Doses above 3 mg per day were not more efficacious and resulted in a greater side-effect burden.

BIPOLAR MANIA: ADULTS. In two 3-week studies comparing risperidone treatment with placebo in patients with acute mania, risperidone treatment was associated with significantly greater improvements on the Young Mania Rating Scale (YMRS) versus placebo. Time to onset of improvement was as brief as 3 days in the first study. A recent meta-analysis by Scherk et al. was conducted on the effects of SGAs in the treatment of mania. It concluded that risperidone was significantly more efficacious than placebo in the treatment of mania, as efficacious as haloperidol, and more effective when used in combination with a mood stabilizer than was a mood stabilizer alone. A small study of 65 patients compared risperidone with haloperidol and lithium and found all medications were similarly efficacious.

BIPOLAR MANIA: CHILDREN AND ADOLESCENTS. An 8-week open-label trial of 30 children and adolescents (aged 6 to 17) found that risperidone monotherapy (mean dose 1.25 mg per day) resulted in clinically significant improvements in the YMRS scores, although at end point there were mild residual symptoms present. A double-blind trial of children and adolescents aged 10 to 17 years compared risperidone 0.5 to 2.5 mg per

day, risperidone 3 to 6 mg per day, and placebo in the treatment of manic or mixed episodes of bipolar I disorder. Risperidone in both dosage groups significantly improved YMRS scores compared to placebo. Doses above 2.5 mg per day were not more efficacious than lower doses.

IRRITABILITY ASSOCIATED WITH AUTISM: CHILDREN AND ADOLESCENTS. A study of 101 children and adolescents aged 5 to 17 years compared risperidone treatment (dose range 0.5 to 3.5 mg per day) with placebo on measures of irritability and global improvement. Sixty-nine percent of risperidone-treated patients versus 12 percent of placebo-treated patients were classified as responders. Hyperactivity and stereotypy scores also improved significantly in the risperidone group. Patients treated with risperidone gained more weight (mean = 2.7 kg) than did patients on placebo (mean = 0.8 kg) in the 8-week trial. EPS rates were similar, although parents reported a higher incidence of tremor in the risperidone group. Of the risperidone responders followed in open-label treatment for the next 6 months, 68 percent maintained those gains. A continuation of the same study followed 63 of the original patient sample for 4 months on open-label risperidone treatment, followed by an 8-week placebo-controlled discontinuation protocol. During the 4-month risperidone continuation phase, 83 percent of the patients maintained the gains in irritability diminution they had achieved during the original 8-week trial. Overall the mean irritability ratings increased from 9.5 to 10.8, an increase the authors concluded was clinically insignificant. Thirty-eight subjects were enrolled in the placebo-controlled discontinuation phase. The relapse rate was 62.5 percent in the placebo group, versus 12.5 percent in the risperidone group. A randomized double-blind, placebo-controlled trial of 40 children with autism (aged 2 to 9 years) found that risperidone 1 mg per day produced improvements of 20 percent or greater on the Childhood Autism Rating Scale and the Children's Global Assessment Scale in 63 and 89 percent, respectively, versus 0 and 10 percent of children treated with placebo. Risperidone was also associated with improvements in social responsiveness, nonverbal communication, hyperactivity, and aggression.

OTHER DIAGNOSES. Risperidone has been used off-label for multiple different symptoms and diagnoses. Approximately a dozen small studies, both open-label and double-blind, initially evaluated the use of risperidone in the treatment of combat- and non-combat-related PTSD. Early studies reported improvements in symptom scores with risperidone treatment, both as adjunctive treatment and as monotherapy. Subsequently, a larger VA Cooperative Study was conducted of a 6-month randomized controlled trial of risperidone (up to 4 mg daily) versus placebo, as adjunctive treatment to ongoing pharmacologic and psychosocial treatments of chronic PTSD. In this study (N = 247), risperidone did not show significant

difference from placebo on the primary outcome measure (Clinician-Administered PTSD Scale) or scales assessing depression, anxiety, and quality of life, but it caused more adverse events.

Risperidone augmentation has demonstrated efficacy in symptom reduction in treatment-refractory patients with major depressive disorder in a few studies. Case reports, open-label trials, and small randomized double-blind trials have shown that risperidone has efficacy as an adjunctive treatment in treatment-refractory OCD. There are a few case reports and a small study demonstrating improvements in various symptoms domains, including aggression and mood, in patients with borderline personality disorder treated with risperidone. Children with severe disruptive behavioral disorders and conduct disorders have been shown to benefit from treatment with risperidone, with important positive effects of risperidone on destructive behaviors such as aggression and self-injury. Additionally, risperidone is efficacious and well-tolerated in children with Tourette syndrome. There is a growing body of data that shows risperidone is useful in children with disruptive behaviors and below average intelligence. In this population, effective doses range from 0.01 to 0.06 mg/kg per day. In smaller studies, risperidone has been used successfully in patients with psychotic disorder secondary to brain trauma (i.e., hemorrhage) and drug-induced psychotic disorder. It has also demonstrated clinical efficacy in those with acquired immunodeficiency syndrome (AIDS)–related dementia and psychosis or with agitation in elderly patients with schizophrenia. It is not effective or well tolerated in the treatment of psychosis in Parkinson disease. Open studies indicate that elderly patients may require lower dosages (0.25 to 3.0 mg per day). Other case reports indicate that behavioral disturbances in Huntington disease and levodopa-induced hallucinations may respond favorably as well.

RISPERIDONE LONG-ACTING INJECTABLE MICROSPHERES. A double-blind, placebo-controlled study of 640 patients compared oral risperidone therapy with risperidone long-acting microspheres for 12 weeks. There was no significant difference in PANSS total score or symptom factor scores. There was no significant difference in adverse events between the two groups. A randomized, controlled, open-label study of 377 patients compared long-acting risperidone microspheres (25 or 50 mg) with olanzapine (5 to 20 mg per day) for 53 weeks. The initial assessment at 13 weeks demonstrated noninferiority of risperidone microspheres (mean dose 40.7 mg per injection) to olanzapine (mean dose 14.6 mg per day). At 12 months, the risperidone group had significantly greater improvement on the disorganized thought factor of the PANSS, while olanzapine showed greater improvement on the anxiety/depression factor. There was a trend toward greater clinical improvement in the risperidone microspheres group at 12 months. There was a greater rate of EPS in the risperidone group; the

olanzapine group had higher rates of weight gain. Another trial of 188 clinically stable patients evaluated clinical improvement on risperidone microspheres after switching from a first-generation depot medication. The authors reported that treatment with risperidone microspheres resulted in clinical improvement, defined by over 20 percent improvement in PANSS total scores, achieved in 51.5 percent of patients switched from first-generation depot medication to risperidone microspheres. There was also an improvement in EPS.

The available data with risperidone microspheres fail to address some of the important issues related to long-term treatment. First, there are no studies available that address whether this new formulation has any advantages in efficacy or tolerability when compared with other available long-acting compounds such as fluphenazine (Prolixin Decanoate) or haloperidol decanoate. In addition, at this time, the longest comparison with oral risperidone was only 12 weeks.

CATIE TRIAL. In CATIE phase I, risperidone, like the other SGAs other than olanzapine, did not significantly lengthen time to treatment discontinuation compared to perphenazine or the other second-generation medications. The mean dose of risperidone was 3.9 mg per day. Risperidone was associated with the lowest rate of discontinuation secondary to intolerable side effects (10 percent). Risperidone and olanzapine were associated with fewer concomitant prescriptions for antianxiety medications than the other antipsychotics. In CATIE phase II, risperidone was associated with gains in psychosocial functioning consistent with the gains made by the other medications in the CATIE trial. In patients who had discontinued perphenazine, risperidone was associated with a significantly shorter time to discontinuation than olanzapine or quetiapine. In patients who had discontinued an SGA, risperidone and olanzapine had significantly longer times to discontinuation than did quetiapine or ziprasidone. Of the patients who discontinued the original second-generation medication due to intolerability, there was a trend toward longer time to discontinuation with risperidone compared to the other second-generation medications.

Effects on Specific Organs and Systems

Cardiovascular System. Tachycardia and palpitations are the most commonly reported cardiovascular side effects, followed by hypotension and postural hypotension. This is most likely related to α-receptor blockade. Palpitations are noted to be dose related. There have been rare reports of atrioventricular and bundle branch block.

Respiratory System. Dyspnea and cough are reported with some frequency. Reports of rhinitis and upper respiratory tract infection are

most common in the pediatric population.

Gastrointestinal System. Abdominal pain and salivary hypersecretion are the most commonly reported gastrointestinal side effects reported in adults taking risperidone. There are also reports of dyspepsia, dry mouth, nausea, and diarrhea. In children, salivary hypersecretion, abdominal pain, nausea, dyspepsia, vomiting, and diarrhea are commonly reported side effects.

Hematological and Lymphatic System. Epistaxis and anemia are rare side effects in the adult population. No hematological effects have been reported in the pediatric population.

Central Nervous System. The majority of common CNS effects are EPS related. These include parkinsonism, dystonia, tremor, akathisia, and dyskinesia. These are also common side effects in the pediatric population, as are automatisms (or compulsive symptoms). Dizziness is also frequently reported. The presence of akathisia probably indicates that not enough 5-HT_{2A} blockade is present.

Genitourinary System. Ejaculatory disorders, priapism, and sexual dysfunction have been reported in men. Nonpuerperal lactation and menstrual irregularities are reported in women and adolescent girls. Sexual side effects may be related to risperidone effect on prolactin levels. There are some reports of urinary tract infections and urinary incontinence with risperidone therapy.

Dermatological System. Rash has been reported with risperidone use. There have been case reports of recurrent giant urticaria and angioedema associated with risperidone use.

Therapeutic Indications

Schizophrenia. Risperidone is effective for both first and later episodes of psychosis in schizophrenia and schizoaffective disorder, as well as other psychoses with prominent positive or negative symptoms. A fast-dissolving tablet (Risperdal M-tab) is available and is especially useful in the treatment of acute psychosis. Risperidone has also been found to be effective for long-term maintenance treatment of schizophrenia.

Bipolar Mania. Risperidone is indicated in adults and children aged 10 to 17 years as monotherapy or adjunctive therapy with lithium or valproate in the treatment of acute manic or mixed episodes associated with bipolar I disorder.

Irritability Associated with Autistic Disorder. Risperidone is indicated for irritability associated with autism in children and adolescents aged 5 to 17 years.

Other Diagnoses. A number of controlled studies have also been conducted in other indications, including schizoaffective disorder and schizotypal personality disorder, bipolar maintenance therapy, PTSD, treatment-resistant depression, borderline personality disorder, and aggression. One patient with Lesch–Nyhan syndrome showed fewer episodes of self-mutilation during treatment with risperidone.

Precautions and Adverse Reactions

Black Box Warning. Risperidone and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis compared with placebo.

Cerebrovascular Adverse Events. In placebo-controlled trials, elderly patients with dementia-related psychosis treated with risperidone had significantly higher incidences of cerebrovascular accidents and other related adverse events than did patients treated with placebo.

Neuroleptic Malignant Syndrome. Risperidone has been associated with NMS, possibly at rates higher than most other SGA medications. NMS is a serious and potentially life-threatening condition. Patients with suspected cases of NMS should be immediately discontinued from the antipsychotic agent and aggressively treated with supportive measures, symptomatic treatment, and medical monitoring. If the patient requires antipsychotic treatment after recovery from NMS, close monitoring and gradual reintroduction is essential.

Tardive Dyskinesia. All antipsychotic medications (with the possible exceptions of clozapine and quetiapine) carry a risk of tardive dyskinesia. Tardive dyskinesia risk is thought to be related to a combination of factors, the most important of which is unopposed D_2 blockade in the basal ganglia. Risperidone has a high affinity for D_2 receptors, but its action at the 5-HT_{2A} receptors may partially mitigate some of the striatal D_2 blockade. In one trial of risperidone treatment, 5 of 453 patients developed symptoms of tardive dyskinesia, and 14 patients with pre-existing tardive dyskinesia improved. In another study of risperidone microspheres, the rate of tardive dyskinesia emergence in 530 patients after 50 weeks of treatment was 1.19 percent. A review article that considered the second-generation medications together reported a rate of tardive dyskinesia with risperidone, amisulpride (Solian), olanzapine, and quetiapine ranging from 0.5 to 1.5

percent. The rates of tardive dyskinesia emergence are slightly higher in elderly populations, especially among women. Among 133 elderly adults administered risperidone the incidence of persistent tardive dyskinesia was 2.6 percent, and at doses of 0.75 to 1.5 mg per day the incidence dropped to 1 percent.

Hypercholesterolemia, Hyperglycemia, and Diabetes Mellitus. Studies indicate that risperidone has a lower risk of hyperglycemia and diabetes mellitus than some of the other atypical antipsychotic medications, in particular olanzapine. Multiple studies have concluded that risperidone has a risk of developing diabetes comparable to the FGAs. A population-based case-control study evaluating 19,637 patients in the United Kingdom concluded that risperidone use was associated with a nonsignificant increased risk of developing diabetes versus no antipsychotic treatment. Another study analyzing health plan claims for 4,308 patients taking antipsychotic medications concluded that the frequency of new onset diabetes type 2 was not significantly different in patients treated with risperidone versus untreated patients (N = 3,625). Results from the CATIE study indicate that risperidone produced changes in glucose level comparable to perphenazine and quetiapine, with an increase lower than that with olanzapine treatment and higher than that with ziprasidone treatment. Some SGA medications have also been associated with elevations in total cholesterol, low-density lipoprotein (LDL), and triglycerides. Studies examining the rates of lipid elevations with risperidone use have generally concluded that risperidone has a rate of hyperlipidemia similar to that of high-potency FGAs or untreated controls. The CATIE data demonstrate little or no change in lipid parameters with risperidone treatment. However, despite the statistically lower rates of hyperglycemia and hyperlipidemia with risperidone treatment versus some other second-generation agents, regular monitoring of fasting cholesterol and glucose levels is appropriate in all patients on SGA medications.

Changes in Weight. In the CATIE study, patients on risperidone gained an average of 0.8 lb during treatment, and 14 percent gained more than 7 percent of their body weight. There was no significant difference between risperidone, quetiapine, perphenazine, and ziprasidone on measures of body weight change. The percentage of patients gaining over 7 percent of their body weight during the 6- to 8-week trials was 18 percent versus 9 percent for placebo. Other studies indicate that weight gain on risperidone is highest in people with a low baseline body mass index (BMI) and lowest in people with a high baseline BMI.

Movement Side Effects and Disorders. Double-blind trials indicate that risperidone treatment is associated with a dose-dependent increase in

extrapyramidal side effects. Although most adult patients may not experience these effects at the recommended daily dosage of 6 mg, dystonic reactions have occurred with dosages between 4 and 16 mg per day, and akathisia has been described. Many patients require only 2 to 4 mg per day of risperidone.

Orthostatic Hypotension and Syncope. Risperidone has antagonistic activity at α-receptors, which contributes to its risk of orthostatic hypotension, dizziness, and syncope. In phase II and III trials among adult patients with schizophrenia, 0.2 percent of patients reported syncope on risperidone. Limiting the initial dose to 2 mg per day either in a single or divided dose may minimize these side effects in healthy adults. In patients with hepatic or renal insufficiency, or elderly patients, the initial dose should be reduced to 0.5 mg twice a day to reduce the potential for hypotension and syncope. Risperidone should be used with caution in people with known cardiovascular disease or individuals with predispositions to hypotension.

Hyperprolactinemia. Risperidone increases prolactin levels. In the CATIE study, risperidone increased prolactin a mean of 15.4 ng/dL, and it was the only study medication to significantly increase prolactin levels. Hyperprolactinemia suppresses gonadotropin-releasing hormone (GnRH), which reduces gonadotropin release and can impair gonadal steroidogenesis in men and women. This elevation is sometimes but not always associated with inhibited reproductive function and symptoms such as galactorrhea, amenorrhea, gynecomastia, erectile dysfunction, and anorgasmia. Long-term hyperprolactinemia may lead to decreased bone density in both sexes when associated with hypogonadism (although no direct evidence exists that would suggest that risperidone decreases bone density, regardless of prolactin level). Approximately one-third of breast cancers are prolactin dependent. The manufacturer suggests that patients with a history or family history of breast cancer should be monitored for emergent malignancy, especially if they develop hyperprolactinemia while on risperidone.

Pregnancy. Risperidone is considered pregnancy category C. A retrospective and prospective case series of 713 pregnancies with risperidone exposure found that in the prospective data, the percentage of congenital malformations (3.6 percent) and spontaneous abortions (16.9 percent) was consistent with background rates in the general population. In the retrospectively identified cases, the percentages of congenital malformation (7.6 percent) and spontaneous abortion (16.7 percent) were higher, but retrospective reporting is subject to bias. Thirty-seven of 197 pregnancies retrospectively had perinatal syndromes, 21 of which were

behavioral or motor disorders. These syndromes involved symptoms of tremor, jitteriness, irritability, somnolence, and feeding problems, consistent with possible risperidone withdrawal or risperidone-induced movement side effects. All but one of these children recovered completely. Use of risperidone in pregnant women should be considered carefully after weighing the risks of fetal exposure with the risks of untreated mental illness in the mother. If risperidone is used during pregnancy, the medication should be slowly titrated down prior to delivery.

Lactation. Risperidone is excreted in breast milk. Studies of the distribution of risperidone and its metabolite, 9-OH-risperidone, reveal that the milk/plasma concentration ratio for both agents was less than 0.5. Relative infant doses of risperidone ranged from 2.3 to 4.7 percent of the maternal dose. In infants who were breastfed, risperidone and 9-OH-risperidone were not detected in the plasma of the infants. However, the possibility of infant exposure and possible EPS/NMS risk cannot be ruled out. The decision to breastfeed while on risperidone should be made only after careful risk–benefit analysis.

Fertility and Teratogenicity. Risperidone increases prolactin, which can decrease fertility in men and women. Meta-analysis of 15 antipsychotic medications across 212 clinical trials found risperidone and its metabolite paliperidone caused the greatest elevation in prolactin level of the drugs studied.

In rat populations, risperidone doses of 0.16 to 5 mg/kg impaired mating in female rats, but not fertility. Beagles given risperidone showed decreases in testosterone levels and sperm motility, which continued at an attenuated level at least 2 months after drug discontinuation. In animal studies, risperidone was associated with a higher number of rat pup deaths in both the perinatal and neonatal periods. Risperidone doses administered to the rat mothers were between 0.1 and 3 times the maximum recommended human dose. There was no increase in the incidence of malformations in the rat pups compared to pups born to untreated mothers.

Carcinogenesis. Risperidone has been associated with increases in pituitary gland adenomas, endocrine pancreas adenomas, and mammary gland adrenocarcinomas in mice and rats. The no-effect dose for the tumors is less than or equal to the maximum human recommended dose. The increase in the above tumors is considered to be associated with prolonged D_2 antagonism and hyperprolactinemia. The relevance to human populations of these findings in rodents is unknown.

Mutagenesis. There is no evidence of mutagenesis with risperidone.

Special Populations

PEDIATRIC USE. Risperidone has been approved for use in children aged 5 to 16 years old for irritability associated with autism, aged 10 to 17 years for bipolar disorder, manic, or mixed episodes, and aged 13 to 17 years for schizophrenia. Overall, over 1,850 pediatric patients have been treated with risperidone in industry-sponsored clinical trials. The rate of tardive dyskinesia in this population was 0.1 percent, and it resolved with discontinuation of risperidone. In a review of studies lasting more than 11 months, Correll and Kane reported an annual tardive dyskinesia incidence in children and adolescents of 0.3 percent. Weight gain among children and adolescents must be monitored with regard to expected weight gain for the appropriate age group. Risperidone is associated with larger than expected levels of weight gain in these populations. In a study of 103 adolescents with diagnoses of schizophrenia, the average weight gain was 9.0 kg over 8 months. In several studies of children with autistic disorder or other disorders, there was an average weight gain of 7.5 kg over 1 year, compared to the expected gains of 3 to 3.5 kg per year, depending on age. Risperidone increases prolactin in children and adolescents, as it does in adults. Rates of prolactin elevation vary from 49 to 87 percent of pediatric patients taking risperidone compared to 2 to 7 percent of patients on placebo. Galactorrhea was reported in 0.8 percent of pediatric patients on risperidone; and gynecomastia was reported in 2.3 percent of risperidone users in the pediatric population. The long-term effects of risperidone treatment on growth and sexual maturation have not been assessed. Transient somnolence lasting a median of 16 days after treatment initiation was reported frequently in the child and adolescent population.

Risperidone use in the elderly should be initiated cautiously. Elderly patients are more prone to orthostatic hypotension, which can put them at greater risk of falls. Furthermore, elderly patients tend to have reduced pharmacokinetic clearance, higher rates of hepatic or renal insufficiency, and greater likelihood of polypharmacy. Elderly patients started on risperidone should be given a lower initial dose of 0.5 once or twice daily, and titration should be slower than in the adult population. Potential drug interactions should be carefully monitored, and any medications significantly metabolized by CYP2D6 should be coadministered with caution. In particular, two placebo-controlled trials found that concomitant administration of risperidone and furosemide (Lasix) increased mortality in elderly patients with dementia-related psychosis, compared to patients treated with risperidone or furosemide alone. Furosemide should not be combined with risperidone in the elderly.

Drug Interactions

Risperidone is extensively metabolized by the cytochrome P450 2D6

enzyme, where it is converted to its major metabolite 9-OH-risperidone. There are identified genetic polymorphisms for the CYP2D6 enzyme, and some percentage of people (6 to 8 percent in Caucasians; higher in African Americans) have little to no 2D6 activity. These people are poor metabolizers of drugs hydroxylated through the 2D6 pathway. Similarly, some individuals are rapid metabolizers at this enzyme site. Risperidone metabolism can also be affected by drugs that are inhibitors (paroxetine [Paxil], fluoxetine, quinidine [Cardioquin]) or inducers (carbamazepine, phenytoin, rifampin [Rifadin], phenobarbital [Bellatal]) of the 2D6 enzyme. Medications that inhibit the 2D6 enzyme (Table 34.28–3) increase plasma levels of risperidone and decrease the rate of conversion to 9-OH-risperidone. Dosing of 20 mg each of fluoxetine and paroxetine (strong inhibitors of 2D6) raised plasma levels of risperidone by 2.5 times and 3.9 times, respectively. Because risperidone and its major metabolite have similar if not identical pharmacologic effects, the result of 2D6 inhibition may not be clinically meaningful. However, risperidone dosage should be reassessed whenever a medication with 2D6 inhibition is initiated, discontinued, or changed. Medications that induce 2D6 metabolism can substantially decrease the steady-state levels of risperidone and 9-OH-risperidone. Carbamazepine (a strong 2D6 inducer) lowers the steady-state concentrations of risperidone and its metabolite by up to 50 percent. Patients coadministered risperidone and any 2D6 inducer, like carbamazepine, should be closely monitored when the inducing medication is initiated, titrated, or discontinued. In particular, if a 2D6-inducing medication is discontinued, risperidone may need to be decreased to avoid plasma concentration increases to toxic levels.

Other potential drug interactions include up to 20 percent higher levels of valproic acid with risperidone coadministration and the additive effects of medications with the potential to cause orthostatic hypotension.

Dosage and Administration

Schizophrenia. For adults starting oral risperidone (available as tablets, fast-dissolve tablets, or a solution), the initial dose should be no higher than 2 mg daily in a single or divided dose. Higher starting dosages are associated with tachycardia and orthostatic hypotension. Dose increases of 1 to 2 mg per day are acceptable but should not be spaced less than 24 hours apart. Most patients are evaluated on 4 to 6 mg for the first 4 to 6 weeks. Many patients may be optimally treated with an average 4 mg daily dose, but doses ranging from 2 to 16 mg have been found effective. Several studies have concluded that doses above 6 mg per day in divided doses were not more effective than lower doses and were associated with higher rates of EPS and other dose-dependent side effects. However one study assessing once-daily doses concluded that 8 mg per day was more efficacious than 4 mg per day. Given the risk of EPS, tardive dyskinesia,

and prolactin elevation, efforts should be made to maintain patients on the lowest effective dose. Lorazepam (Ativan) may be added during dose titration if the patient shows agitation in the early phases of treatment. Most of the improvement occurs in the first 8 weeks. Risperidone can be given once daily, which is as effective as twice-daily dosing.

Table 34.28–3.
Overview of Potential Cytochrome P450 Interactions with Second-Generation Antipsychotics

SGA	Major P450 Pathway[a]	Drugs that ↑ SGA Levels[b]	Drugs that ↓ SGA Levels[b]	SGA Effect on Other Med Levels[b]	Comments
Clozapine (Clozaril)	1A2, 2D6, 3A4	Fluvoxamine (paroxetine, fluoxetine, bupropion, other 2D6 → modest ↑), cimetidine, caffeine, ciprofloxacin, erythromycin	Nicotine, carbamazepine, modafinil, St. John wort, omeprazole, phenytoin, rifampin	Mild 2D6 inhibitor. Caution advised with other meds metabolized by 2D6. Levels of both may be ↑.	May decrease tobacco use, which can ↑ clozapine leve caution when patients change tobacc use; due to high toxicity and multiple P450 pathway check all coadministere medicines for interactions and monitor clozapine leve with medicine changes.
Risperidone (Risperdal)	2D6	Fluoxetine, paroxetine, bupropion, quinidine	Carbamazepine, phenytoin, phenobarbital, rifampin	Valproate ↑ by 20%	Risperidone and its metabolite are both active
Olanzapine (Zyprexa)	1A2	Fluvoxamine, ciprofloxacin, fluoxetine (small)	Nicotine, carbamazepine, modafinil, omeprazole, rifampin	N/A	
Quetiapine (Seroquel)	3A4	Divalproex (small), cimetidine (small), nefazodone, ketoconazole, clarithromycin, erythromycin, ritonavir	Carbamazepine, thioridazine, phenytoin, modafinil, St. John wort	Slight decrease in lorazepam clearance; Slight decrease in valproic acid exposure	
Ziprasidone	3A4	Nefazodone,	Carbamazepine,	N/A	Concomitant Rx

(Geodon)		ketoconazole, clarithromycin, erythromycin, protease inhibitors	phenytoin, modafinil, St. John wort		of medicines that ↑ QTc is contraindicate or should be used with caution.
Aripiprazole (Abilify)	2D6, 3A4	Fluoxetine, paroxetine, bupropion, quinidine, nefazodone, ketoconazole, clarithromycin, erythromycin, protease inhibitors	Carbamazepine, rifampin, phenytoin, modafinil, St. John wort	N/A	
Paliperidone (Invega)	N/A	Paroxetine (small), divalproex	N/A	N/A	As the metabolit of risperidone, paliperidone undergoes onl limited P450 metabolism.
Iloperidone (Fanapt)	2D6, 3A4	Fluoxetine, paroxetine, ketoconazole, itraconazole	Carbamazepine, rifampin, phenytoin, modafinil, St. John wort	N/A	Recommend not using with other drugs which prolong QTc interval.
Asenapine (Saphris)	1A2	Fluvoxamine, imipramine (small)	Paroxetine (small), carbamazepine (small)	Weak 2D6 inhibitor → double paroxetine exposure	
Lurasidone (Latuda)	3A4	Ketoconazole, clarithromycin, ritonavir, voriconazole, mibefradil, diltiazem, atazanavir, erythromycin, fluconazole, verapamil, grapefruit juice	Rifampin, avasimibe, St. John wort, Phenytoin, carbamazepine	Small increase in midazolam exposure (CYP3A4 substrate)	
Cariprazine (Vraylar)	3A4, 2D6	Ketoconazole, erythromycin	Carbamazepine, phenytoin, modafinil, St. John wort		
Brexpiprazole (Rexulti)	2D6, 3A4	Fluoxetine, paroxetine, bupropion, ketoconazole, quinidine,	Carbamazepine, phenytoin, rifampin, modafinil, St. John wort		

	erythromycin

Rx, prescription; N/A, not available.

[a]Medications primarily metabolized by only one cytochrome system, especially CYP2D6, are subject to more numerous and more significant potential drug–drug interactions than medications with multiple pathways.

[b]This is not a comprehensive list of potential interactions. Clinicians should check for possible drug interactions whenever patients' medication regimens are changed, including by other providers.

Adolescents with schizophrenia should be started on no more than 0.5 mg once daily. Dose increases should be made in increments of 0.5 to 1 mg per day, no less than 24 hours apart. The recommended daily dosage for adolescents with schizophrenia is 3 mg per day. Although a dose range of 1 to 6 mg per day has demonstrated efficacy, doses higher than 3 mg per day did not provide additional benefit and were associated with higher rates of side effects. If side effects are experienced, tolerability may be improved by dividing the dose into twice-daily dosing.

Elderly and Parkinson disease patients should start with 0.5 mg once or twice daily. Doses should be divided into twice-daily dosing to reduce the risk of orthostatic hypotension. Dose increases should be spaced 2 to 3 days apart, and dose increases above 1.5 mg twice a day should be made no sooner than 1 week after the last dosage change. This population may need maintenance doses of 1 mg or less to prevent other side effects.

Bipolar Mania. Adults with acute bipolar mania should be started at a dose of 2 to 3 mg per day with once-daily dosing. Dose increments should not be higher than 1 mg per day or spaced less than 24 hours apart. In this population, doses higher than 6 mg per day have not been fully evaluated.

Children and adolescents with bipolar mania should be started at doses of 0.5 mg once daily and titrated up by 0.5 to 1 mg per day no less than 24 hours apart to a recommended dose of 2.5 mg per day. Doses as high as 6 mg per day have been studied in this population, but efficacy was no greater at doses above 2.5 mg per day. The dose can be divided to reduce side effects, should they emerge. Given the unknown long-term impact of risperidone on the neurological, endocrinological, and sexual maturation of children, clinicians should strive to find the lowest effective dose and regularly re-evaluate risperidone's efficacy and appropriateness for each patient.

Irritability Associated with Autism. Children and adolescents with autistic disorder should be initiated on 0.25 mg per day for patients under 20 kg and 0.5 mg per day for patients over 20 kg. No further dose increase should be initiated for a minimum of 4 days, after which time the dose can be titrated up to 0.5 mg per day for patients under 20 kg, and 1 mg per day for patients over 20 kg. This dose should be maintained for 2 weeks. If

there is no or partial response at that dose, increases of 0.25 mg per day for patients below 20 kg and 0.5 mg per day for patients over 20 kg should occur at intervals no briefer than 2 weeks. In clinical trials of children and adolescents with autistic disorder, 90 percent of the medication responders responded to doses between 0.5 and 2.5 mg per day. In patients over 45 kg, the optimal dose may be slightly higher, approximately 3 mg per day. As in the pediatric bipolar population, given the unknown long-term impact of risperidone on the neurological, endocrinological, and sexual maturation of children, clinicians should strive to find the lowest effective dose and regularly re-evaluate risperidone's efficacy and appropriateness for each patient.

Overdose. Reports of overdose up to 360 mg of risperidone are associated with signs and symptoms such as sedation, tachycardia, hypotension, EPS, prolonged QT interval, and convulsions. A study evaluating the symptoms associated with overdose has shown that risperidone is associated with fewer side effects than clozapine and olanzapine (particularly anticholinergic side effects like agitation, confusion, and dysarthria).

Long-Acting Risperidone Microspheres. In 2003, Janssen released a novel long-acting intramuscular (IM) preparation of risperidone (Consta), the first SGA available in long-acting injectable form. Uniquely, among IM formulations, the long-acting preparation of risperidone is an aqueous suspension, with risperidone contained in a matrix of glycolic acid–lactate polymer "microspheres" (25 to 150 μm). This preparation minimizes pain and inflammation at the injection site. Gradual hydrolysis of the copolymer at the site of injection ensures slow, steady, and smooth release of risperidone over several weeks. The administration of risperidone in this long-acting, bimonthly preparation removes the need to remember once-daily dosing.

Multiple-dose pharmacokinetic studies have indicated that patients can be smoothly transitioned from oral to long-acting risperidone while maintaining stable plasma drug levels. Once administered, the main drug release commences 3 weeks after injection, is sustained for 4 to 6 weeks, and subsides by 7 weeks. To maintain steady plasma levels, oral risperidone is coadministered with the injectable risperidone microspheres for the first 3 weeks, after which the oral preparation can be gradually discontinued. Risperidone microspheres achieve steady-state concentrations after about four injections, and therapeutic concentrations are maintained for 4 to 6 weeks after the last injection. Dosing risperidone microspheres every 2 weeks ensures sustained therapeutic concentrations without significant fluctuations in plasma level. Doses range from 12.5, 25, 37.5, and 50 mg per 2 weeks. Studies evaluating doses higher than 50 mg have found no

additional clinical benefit and a higher rate of side effects and adverse events. A recent 48-week randomized, single-blind pharmacokinetic study evaluating switching dose from oral to long-acting risperidone had the following recommendations: For patients on 3 mg or less of oral risperidone, administer a 25-mg dose of the microsphere preparation; for 3 to 5 mg of oral medication, use 37.5 mg of the preparation; above 6 mg of oral risperidone, begin injections at the 50-mg dose. Patients with hepatic or renal impairment should be initiated at a 12.5-mg dose. Risperidone microspheres should be injected into alternating gluteal muscles. Detailed instructions on proper injection procedure are included with the medication. Patients obtaining the medication in advance of the injection should be advised to keep it refrigerated until their appointment. Risperidone microspheres have a lower incidence of side effects such as EPS, insomnia, orthostatic hypotension, sexual dysfunction (less than 2 percent), and weight gain (0.5 kg in the 25-mg dose group and 1.4 kg in the 50-mg group in 12-week trials). If a good symptomatic response is obtained with oral risperidone or other oral SGA agents but side effects are intolerable, switching to risperidone microspheres may be preferable. Because the smooth, continuous pharmacokinetic profile of this formulation is associated with a positive tolerability profile, it is likely to promote enhanced patient compliance and autonomy as well as improved outcomes. Hospitalization rates at 1 year were found to be lower after treatment with long-acting IM risperidone than would normally be anticipated after treatment with oral atypical antipsychotic agents.

OLANZAPINE

Olanzapine is an effective agent for the treatment of psychosis in schizophrenia and bipolar disorder. It has a mildly sedating, calming effect that has led to its acceptance in many acute psychiatric settings. The results from recent large trials, including CATIE, suggest an important dilemma for clinicians and patients. On the one hand, olanzapine may have marginal efficacy advantages when compared with some other antipsychotics. On the other hand, olanzapine is also associated with serious side effects including weight gain, lipid elevation, and an increased risk of diabetes.

History

Eli Lilly and Company formulated olanzapine in Great Britain in 1982. Since olanzapine is similar in structure to clozapine, its early development was encouraged by the hope that it would have similar efficacy with milder side effects. Phase III clinical trials were completed in July 1995, and olanzapine became available on the US and European markets in the fall of 1996.

Chemistry

Olanzapine is a thienobenzodiazepine (Fig. 34.28–1). It can be considered a derivative of clozapine with substitution of a thieno ring for clozapine's carbonyl ring.

Pharmacological Action

Pharmacokinetics

ABSORPTION. Olanzapine is well absorbed when administered orally. Peak plasma concentrations are reached in about 6 hours. Food has relatively little influence on its absorption. Olanzapine's half-life is influenced by smoking, gender, and age with a mean of about 31 hours (range of 21 to 54 hours). This allows for once-daily dosing. Patients reach steady plasma concentrations in about 7 days. IM administration of olanzapine leads to peak plasma concentrations within 15 to 45 minutes.

DISTRIBUTION. Olanzapine is widely distributed throughout the body. The drug is 93 percent protein bound to plasma proteins.

METABOLISM AND ELIMINATION. The primary metabolic pathways for olanzapine are direct glucuronidation and P450-medicated oxidation. None of olanzapine's metabolites are active. Although in vitro studies indicate that CYP1A2 and 2D6 and the flavin-containing mono-oxygenase system are involved in its oxidation, CYP1A2 appears to be the major pathway.

Pharmacodynamics

MECHANISM OF ACTION. Olanzapine is a high affinity antagonist at 5-$HT_{2A/2C}$, 5-HT_6, D_{1-4}, H_1, and adrenergic α_1-receptors and a moderate affinity antagonist at M_{1-5} and 5-HT_3 receptors. It resembles most of the SGAs by having a combination of high 5-HT_2 and D_2 activity. PET studies indicate that typical doses of olanzapine such as 10 to 20 mg result in 68 to 84 percent D_2 occupancy. Its 5-HT_2 activity is approximately eight times as strong as its dopamine receptor blockade. Compared with risperidone, quetiapine, and ziprasidone, olanzapine has greater muscarinic and histaminergic receptor antagonism. This histamine activity may explain its tendency to cause weight gain.

BLOOD CONCENTRATIONS AND CLINICAL ACTIVITY. A study by Perry and coworkers measured plasma concentrations in a large early olanzapine trial. Their findings indicate that patients with plasma concentrations above 23 ng/mL were more likely to have a favorable clinical response.

Design and Interpretation of Clinical Drug Studies

Acute Treatment of Schizophrenia. Double-blind studies have

compared once-daily olanzapine at dosages ranging from 2.5 to 17.5 mg per day, with placebo and haloperidol (10 to 20 mg per day). In one study, olanzapine was found to be superior to placebo and comparable to haloperidol in treatment of overall psychopathology (mean CGI-S score), and positive psychotic symptoms (positive BPRS and PANSS scores) at intermediate (7.5 to 12.5 mg per day) and high (12.5 to 17.5 mg per day) dosages. Olanzapine was also found to be superior to haloperidol in negative psychotic symptom response.

In another 6-week, fixed-dose, double-blind study involving 1,996 patients in North America and Europe, olanzapine showed significantly greater clinical improvement over haloperidol (total BPRS score, $p < 0.02$). Significant improvement was also seen in both positive symptom and negative symptom scores ($p = 0.03$) as well as depression rating ($p = 0.001$).

Two other dose-ranging studies have also shown olanzapine (1 to 15 mg per day) to have significant antipsychotic activity against both positive and negative symptoms of schizophrenia. Acute EPS were less frequent with olanzapine at all doses compared with haloperidol. Olanzapine has also been shown to be superior to haloperidol in treating schizophrenia patients experiencing first-episode psychosis. One study has reported that olanzapine has a significantly better risk–benefit profile than haloperidol in this group of patients.

Double-blind studies have also evaluated the effect of a rapid-acting IM formulation of olanzapine in the treatment of acute agitation in schizophrenia. IM olanzapine was found to be effective against positive symptoms in acutely agitated patients, superior to placebo and similar to IM haloperidol.

Olanzapine in Children and Adolescents. There is relatively few data evaluating the effectiveness of olanzapine for children or adolescents with schizophrenia or bipolar illness. A study in adolescents reported by Kryzhanovskaya and coworkers found that olanzapine was more effective than placebo, but associated with somnolence, liver enzyme elevations, prolactin elevation, and weight gain. An 8-week study that compared olanzapine, risperidone, and haloperidol in 50 patients aged 8 to 19 years with psychosis found that the three conditions were equally effective. However, weight gain and EPS on olanzapine and risperidone appeared to be greater than had been reported in adults. A trial by Linmarie Sikich and coworkers compared risperidone and olanzapine to molindone (Moban) in patients between 8 and 19 years of age. The NIMH Data and Safety Monitoring Board closed randomization to the olanzapine arm due to weight gain. These concerns are mirrored from a study of bipolar adolescents (13 to 17 years) who were randomized in a 3-week comparison of olanzapine or placebo. Although olanzapine was effective in reducing

symptoms, it was also associated with increases in weight, liver enzymes, fasting cholesterol, prolactin, and fasting glucose.

Taken together, all of these studies support very serious safety concerns related to treatment with olanzapine in children and adolescents. In the opinion of the authors, olanzapine should be considered for this population only when other antipsychotics have failed.

Olanzapine in First-Episode Schizophrenia. A study of 263 patients with first-episode schizophrenia reported by Jeffrey Lieberman found that olanzapine had advantages over haloperidol in total PANSS scores and depression. An interesting finding from this study indicated that first-episode patients who received haloperidol experienced losses in gray matter, which did not occur in patients who received olanzapine. An analysis of data over 2 years found that fewer patients discontinued olanzapine. However, olanzapine was associated with greater weight gain and elevations in liver enzymes. A 1-year double-blind study reported by Joseph McEvoy randomized 400 recent-onset patients to olanzapine, quetiapine, or risperidone. Discontinuation rates were similar on the three drugs. However, there were advantages for olanzapine on positive symptoms. As in other studies, weight gain was highest with olanzapine.

Comparisons with Clozapine. One published study has compared olanzapine with clozapine. In the double-blind study, 180 patients with chronic and resistant schizophrenia randomly received olanzapine (15 to 25 mg per day) or clozapine (200 to 600 mg per day) for 18 weeks. Both drugs were equally effective in reducing overall psychopathology. A significantly greater number of patients discontinued treatment with clozapine compared with olanzapine ($p = 0.022$). However, other studies have not supported this claim of superior efficacy.

Efficacy in Prodromal Schizophrenia. Thomas McGlashan and a group of North American investigators evaluated the effectiveness of olanzapine versus placebo for delaying the onset of psychosis in individuals who had prodromal symptoms. During a year of double-blind treatment 16.1 percent of olanzapine and 37.9 percent of placebo patients converted to psychosis. There was also an advantage of olanzapine for decreasing the severity of prodromal positive symptoms. However, olanzapine patients experienced a mean weight gain of 8.79 kg compared to 0.30 kg on placebo. This finding mirrors other studies in which the advantages of olanzapine are balanced by its tendency to cause substantial weight gain, particularly in younger patients.

Efficacy in Treatment-Refractory Patients. Olanzapine is potentially of interest for treatment-refractory patients because it has a biochemical

profile similar to that of clozapine and is more effective at treating negative symptoms than haloperidol. One study reported that olanzapine was as effective as clozapine in treatment-resistant patients. Alan Breier and his colleagues demonstrated that olanzapine provided a significantly greater mean improvement than haloperidol for key symptom domains (PANSS negative symptoms, comorbid depressive symptoms) and side effects (EPS, akathisia) in treatment-resistant patients.

Long-Term Effectiveness. A year-long maintenance study found fewer relapses with olanzapine than with placebo or haloperidol. At higher dosages of olanzapine, fewer patients dropped out of the study than with haloperidol.

Efficacy for Mania and Mixed States in Bipolar Disorder. Two studies found that olanzapine monotherapy at doses between 5 and 20 mg was superior to placebo at reducing manic symptoms as measured by the YMRS. A third study failed to find this advantage for olanzapine. These studies support the use of olanzapine in bipolar patients with mania or in mixed states. A study by Perlis and coworkers compared olanzapine and risperidone and found that the two SGAs had similar effectiveness in reducing manic symptoms. Two trials evaluated the effectiveness of adding olanzapine or placebo to either lithium or valproate when patients had symptoms that were inadequately controlled on the first drug. In both trials, the addition of olanzapine was more effective than the addition of placebo in reducing manic symptoms.

Agitation in Schizophrenia and Bipolar Disorder. Studies of IM olanzapine focused on the short-term reduction in agitation in patients with schizophrenia or bipolar disorder. They were double-blind and consisted of patients being assigned to receive up to three injections in the first 24 hours. Two studies compared IM olanzapine to placebo in patients with schizophrenia who demonstrated symptoms of excitement. Olanzapine at doses of 2.5, 5, 7.5, and 10 mg was more effective than placebo at decreasing levels of excitement. The higher olanzapine doses resulted in greater reductions in excitement. The advantage of olanzapine was apparent for all doses at 2 hours and for 5 to 10 mg at 30 minutes. These improvements were similar to those with IM haloperidol or IM lorazepam in separate studies. Another study found that 10 mg of IM olanzapine was more effective than placebo and oral olanzapine for reducing excitement in bipolar patients with mania or mixed states.

Effects on Specific Organs and Systems

CARDIOVASCULAR. Olanzapine is associated with orthostatic hypotension,

tachycardia, and possible syncope. This probably results from its α_1-adrenergic antagonistic properties. IM olanzapine has been associated with a syndrome that includes hypotension with bradycardia.

GASTROINTESTINAL. The anticholinergic effects of olanzapine can lead to constipation in some patients.

CENTRAL NERVOUS SYSTEM. Olanzapine is somewhat sedating and may affect motor skills.

Therapeutic Indications

Schizophrenia. Olanzapine is an effective agent for treating acute psychotic symptoms in schizophrenia. Since it is mildly sedating, psychiatrists in acute settings often favor olanzapine for patients who will benefit from a calming agent.

Olanzapine is also effective for maintaining an antipsychotic response in stabilized patients. In large studies such as CATIE, olanzapine patients tended to have lower discontinuation rates when compared with other first- and second-generation drugs. However, weight gain, lipid elevations, and insulin resistance can be serious problems in some patients. As a result, patients who continue in long-term treatment should be carefully monitored for these metabolic effects.

Bipolar Disorder. Olanzapine is effective for the acute treatment of mania and mixed states, both as monotherapy and as adjunctive therapy to lithium or valproate. It is notable that olanzapine is effective for both psychotic and nonpsychotic patients.

Olanzapine is also effective as monotherapy for maintaining response in patients who have been treated for mania or mixed states. As with schizophrenia, patients receiving long-term treatment with olanzapine should be carefully monitored for its metabolic adverse effects.

Agitation Associated with Schizophrenia and Bipolar I Mania. IM olanzapine is indicated for agitation that can be associated with schizophrenia or mania. The olanzapine product information defines psychomotor agitation as "excessive motor activity associated with a feeling of inner tension." It should be noted that oral olanzapine in its different formulations as well as other antipsychotics are also effective for agitation. The important advantage of the IM route is a more rapid onset of action.

Other Possible Indications. In double-blind studies, olanzapine has been effective in patients with schizoaffective disorder, Tourette syndrome, and for psychosis in dementia. However, there are serious concerns about increased mortality in dementia patients treated with olanzapine. In open-

label studies, olanzapine has been found to be effective in autism and in producing weight gain in patients with anorexia nervosa. Olanzapine is also used in combination with fluoxetine (Symbyax), which is indicated for depressive episodes associated with bipolar I disorder as well as treatment-resistant depression.

Precautions and Adverse Reactions

Black Box Warning

INCREASED MORTALITY IN ELDERLY PATIENTS WITH DEMENTIA. Olanzapine and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis compared with placebo.

Weight Gain. A large number of studies indicate that olanzapine is associated with greater weight gain than other SGAs with the possible exception of clozapine. In the NIMH CATIE study, olanzapine patients gained approximately 2 lb per week. Studies during long-term therapy with olanzapine found that 56 percent of patients gained greater than 7 percent of baseline weight. Weight gain occurred across all BMI levels indicating that low, normal, overweight, and obese patients were vulnerable to weight gain. As mentioned earlier, children and adolescents are more vulnerable to weight increase with olanzapine treatment.

Hyperlipidemia. Patients treated with olanzapine may experience serious increases in total cholesterol, LDL cholesterol, and triglycerides. Triglycerides are most likely to be elevated, sometimes to levels above 500 mg/dL. In the NIMH CATIE study, the mean increase in triglycerides in patients taking olanzapine was 40.5 mg/dL. For comparison, quetiapine was associated with an increase of 21.2, risperidone with a decrease of 2.4, perphenazine with an increase of 9.2, and ziprasidone with a decrease of 16.5. Since some patients were receiving olanzapine at the start of the trial and may have already experience increases, this may be an underestimate of triglyceride elevation on olanzapine. Patients with no prior history of hyperlipidemia were particularly vulnerable to increases during olanzapine treatment. These concerns about hyperlipidemia suggest that patients who are treated with olanzapine should have lipid levels monitored before and after olanzapine treatment is initiated and at regular intervals.

Hyperglycemia, Insulin Resistance, Diabetes Mellitus. Several studies suggest that hyperglycemia and diabetes mellitus are potential side effects that follow treatment with olanzapine. A number of studies, including CATIE, found greater elevations in blood glucose on olanzapine than on other antipsychotics. These effects were apparent in individuals with normal glucose regulation as indicated by normal fasting blood glucose as

well as in individuals with diabetes mellitus or other evidence of glucose dysregulation. Other studies have found that olanzapine is associated with an increase in insulin resistance in both obese and nonobese patients. This suggests that the effects of olanzapine are not just secondary to weight gain.

There is also substantial evidence indicating that olanzapine treatment is associated with an increased risk of the development of type 2 diabetes mellitus. Using the United Kingdom General Practice Research Database, Koro and coworkers found that olanzapine was associated with a significantly increased risk of diabetes when compared with patients who were not on antipsychotics (odds ratio 5.8) or taking FGAs. This finding has been supported by other large epidemiological studies. The increased risk could be related to weight gain, which is highly associated with the development of diabetes. However, there are descriptions of normal weight patients on olanzapine who develop diabetes.

There is also evidence that olanzapine can result in elevations in blood glucose that are life-threatening. In a retrospective epidemiological survey of spontaneously reported adverse events related to olanzapine therapy, 237 patients with olanzapine-associated diabetes or hyperglycemia were identified. Seventy-three percent of all cases of hyperglycemia appeared within 6 months of initiating olanzapine therapy. Eighty patients had metabolic acidosis or ketosis, 41 had glucose levels of 1,000 mg/dL or greater, and 15 patients died. When olanzapine was discontinued or the dosage was decreased, 78 percent of patients had improved glycemic control. Hyperglycemia recurred in eight of ten cases with rechallenge.

These concerns regarding glycemic control with olanzapine indicate that other antipsychotics may be preferred for patients who already have diabetes as well as those with risk factors for diabetes such as obesity or a family history of diabetes. In addition, the American Psychiatric Association and Mount Sinai Guidelines recommend that clinicians evaluate fasting blood glucose before and after initiating treatment with olanzapine.

Neuroleptic Malignant Syndrome and Tardive Dyskinesia. The product information for olanzapine has the same warnings regarding tardive dyskinesia and NMS as other FGAs and SGAs. However, most studies of tardive dyskinesia indicate a reduced risk with olanzapine compared to FGAs. In addition, there are few reports of NMS with olanzapine.

Elevated Hepatic Enzymes. Placebo-controlled trials of olanzapine found a small incidence of elevated liver enzymes. For example, 2 percent of olanzapine-treated patients had clinically significant alanine aminotransferase (ALT) elevations, which were defined as three times the upper level of normal. Patients with these elevations did not develop jaundice.

Pregnancy. Olanzapine is considered pregnancy category C. In a study that compared exposure to all SGAs to no exposure during pregnancy, the authors reported no increased risk of malformation or neonatal complications, except for lower birth weight in the exposed group. Although patients on SGAs had a higher risk of excessive weight gain during pregnancy, olanzapine was not associated with an increased risk among the SGAs. A 2000 article by the manufacturer reported on 37 prospective and 11 retrospective pregnancies with exposure to olanzapine. The rate of spontaneous abortion and malformation was at population norms. Of the prospective pregnancies, 84 percent had normal delivery and postnatal course. The remaining 16 percent had a number of problems, including prematurity, postmaturity, and low and high birth weight. Data on the metabolic status of the mothers were not reported; however, at least two of the mothers were noted to have gestational diabetes. Another study measured placental passage of medication from mother to fetus and documented neonatal outcomes in 54 women followed through pregnancy. Newport and colleagues concluded that olanzapine has the highest rate of placental passage, compared with haloperidol, risperidone, and quetiapine. Furthermore, neonates exposed to olanzapine showed trends toward lower birth weights and more neonatal intensive care unit admissions than neonates exposed to other antipsychotic medication. Case reports have indicated increased risk of gestational diabetes among pregnant women treated with olanzapine. Concerns of excessive weight gain during pregnancy and its concomitant risks to mother and fetus must be considered. The benefits of olanzapine treatment should be weighed against the potential risks to the fetus.

Nursing Mothers. Olanzapine is excreted in breast milk, but the concentration of olanzapine in breast milk appears to be low. A study was conducted by Gardiner and colleague that measured transfer of olanzapine to breast milk and then infant plasma in seven mother–infant pairs. The measured concentration of olanzapine in breast milk was approximately 1 percent of maternal dose. Olanzapine was undetectable in the plasma of exposed infants, and all infants were healthy without observable side effects. As with any medication use during breastfeeding, mothers should be advised of the risks and benefits of breastfeeding while on medication and the potential consequences of olanzapine exposure to their infant.

Drug Interactions

Agents that induce CYP1A2 such as omeprazole (Prilosec) and rifampin may decrease olanzapine plasma concentrations. Smoking also induces CYP1A2 and may have similar effects. Carbamazepine can increase the clearance of olanzapine and may also decrease plasma concentrations. Fluvoxamine, an inhibitor of CYP1A2, can cause increases in olanzapine

levels.

Laboratory Interference

Olanzapine does not interfere with common laboratory tests.

Dosage and Administration

Olanzapine is available as 2.5-, 5-, 7.5-, 10-, 15-, and 20-mg tablets; as 5-, 10-, 15-, and 20-mg orally disintegrating tablets (Zydis); as vials with 10 mg of olanzapine for IM injection; and as vials with 150-, 210-, 300-, or 450-mg olanzapine pamoate extended release for IM injection (Relprevv).

Treatment of Schizophrenia. Olanzapine is effective at dosages between 7.5 and 20.0 mg per day administered once daily without regard to meals. A single starting dose of 5 to 10 mg is well tolerated in most adult patients. The recommended initial dose is then adjusted as necessary within the range of 5 to 20 mg per day. Although clinical studies suggest that there is not an advantage for prescribing doses above 10 mg daily, the impression of many clinicians is that doses of 15 to 25 mg or more are optimal for acute schizophrenia. This is reflected in the CATIE study where the mean modal daily dose of olanzapine was 20.1 mg daily. In other words, the mean dose was above the maximum dose of 20 mg that is recommended in the product information. In adolescents, the dose range of 2.5 to 20 mg/day demonstrated efficacy, with the mean dose being 11.1 mg/day.

Maintenance in Schizophrenia. Long-term studies indicate that doses of 10 to 20 mg daily of olanzapine are sufficient to prevent psychotic relapse in stabilized patients.

Treatment of Bipolar Disorder. Patients with mania or mixed states can usually be started on 10 to 15 mg daily of oral olanzapine. Doses can be increased to 20 or 25 mg daily. Similar doses can be prescribed when olanzapine is added to lithium or valproate. Patients with bipolar disorder who have been stabilized can usually be maintained on 5 to 10 mg of olanzapine daily.

Special Populations. Patients who are debilitated or vulnerable to hypotension should be started on 5 mg or less. Doses of 5 to 10 mg orally may effective for schizophrenia or bipolar disorder in these individuals.

IM Olanzapine in Agitation Associated with Schizophrenia or Bipolar Disorder. IM doses of 2.5 to 10 mg have been effective in reducing agitation. In most cases, 10 mg IM is an appropriate initial dose. Lower doses may be appropriate for some patients. Doses of 10 mg may be repeated after 2 hours if the patient remains agitated. As noted earlier, the

advantage of IM administration is a more rapid onset of calming. Therefore, cooperative patients may be changed to oral administration with rapid dissolve or regular tablets.

Olanzapine Pamoate Long-Acting Injectable. In 2009, the FDA-approved olanzapine pamoate as a long-acting IM injectable agent indicated for the treatment of schizophrenia. Its efficacy was demonstrated in an 8-week randomized controlled trial (*N* = 404), in which patients with schizophrenia were discontinued from their prior oral antipsychotics and underwent a 2 to 7 day washout, followed by initiation of olanzapine long-acting injectable 210 mg every 2 weeks, 405 mg every 4 weeks, 300 mg every 2 weeks, or placebo. The primary outcome measure was the change in PANSS total score from baseline to end point. Olanzapine long-acting injectable showed statistically significant improvement over placebo at all three dosages. A subsequent longer-term trial enrolled 1065 adults with schizophrenia who were stable for 4 to 8 weeks on oral olanzapine. They were randomized to either continue their current oral olanzapine or switch to olanzapine in long-acting injectable form. The primary outcome measure was the time to exacerbation of schizophrenia, which was defined by prespecified increases in BPRS positive symptom score or hospitalization. In this study, olanzapine long-acting injectable at dosages of 150 mg every 2 weeks, 405 mg every 4 weeks, and 300 mg every 2 weeks were superior to the lower dose of 45 mg every 4 weeks.

Of particular importance, this medication has a black box warning for post-injection/sedation syndrome, which can present with signs and symptoms similar to overdose of olanzapine (e.g., may include sedation, delirium, EPS, dysarthria, ataxia, aggression, dizziness, weakness, hypertension, and convulsions). In its most extreme presentation, it can lead to coma and require intubation. This syndrome generally occurs within the first 3 hours after injection, but can also occur after this period. The frequency of this syndrome is reported to be <0.1 percent of injections and in roughly 2 percent of patients who received injections for up to 46 months.

The FDA mandates that this medication be administered at registered facilities with access to emergency services if needed. After injection, patients must be under observation for at least 3 hours before leaving the site. If this medication is used, the prescriber, health care facility, patient, and pharmacy must be enrolled in a patient care program to facilitate medication distribution.

Patients being prescribed olanzapine long-acting injection should already be taking olanzapine orally to ensure medication tolerability. This medication is recommended to be administered every 2 to 4 weeks by deep IM gluteal injection. The dosage schedules specify initial dose during the first 8 weeks, followed by maintenance dosing thereafter. To achieve levels

roughly equivalent to a 10 mg/day oral dose, patients should be initially administered 210 mg every 2 weeks or 405 mg every 4 weeks, followed by maintenance dosing of 150 mg every 2 weeks or 300 mg every 4 weeks. To achieve levels roughly equivalent to a 15 mg/day oral dose, patients should be initially administered 300 mg every 2 weeks, following by maintenance dosing of 210 mg every 2 weeks or 405 mg every 4 weeks. To achieve levels roughly equivalent to a 20 mg/day oral dose, patients should be initially administered 300 mg every 2 weeks, following by maintenance dosing of 300 mg every 2 weeks. Patients who are debilitated or more sensitive to olanzapine should be started at 150 mg every 4 weeks.

QUETIAPINE

Since its release in 1998, quetiapine has become a first-line SGA medication. In clinical trials, efficacy in schizophrenia has been demonstrated at a dosage range of 150 to 800 mg per day. The efficacy of quetiapine is superior to placebo and generally equivalent to other antipsychotic medications. Quetiapine can be distinguished by a lower incidence of extrapyramidal side effects than any of the other SGAs, with the possible exception of clozapine. Other side effects include drowsiness, orthostatic hypotension, increased heart rate, dry mouth, constipation, weight gain, hyperlipidemia, and hyperglycemia. No blood dyscrasias or effects on prolactin concentrations have been observed in long-term trials. In addition to its indication for treatment of schizophrenia, quetiapine is FDA approved for the treatment of depressive episodes in bipolar I and II and manic episodes in bipolar I. An extended release formulation is also available, which allows once-daily dosing rather than the recommended twice-daily dosing for the original formulation.

History

Quetiapine was developed by Zeneca Laboratories. Its new drug application was filed with the FDA in 1996 and approved in 1998.

Chemistry

Quetiapine is a dibenzothiazepine with more potent 5-HT_2 than D_2 receptor-blocking properties. Its molecular formula is $C_{42}H_{50}N_6O_4S_2 \cdot C_4H_4O_4$. Quetiapine is related to the clozapine, olanzapine (Fig. 34.28–1), and fluperlapine molecules.

Pharmacological Action

Pharmacokinetics

ABSORPTION. The time to maximum concentration (T_{max}) after oral intake is 1.5 hours, with an estimated half-life of 6 hours. No difference was

found between twice- and three-times-daily dosing, in spite of a steady-state half-life of 6.9 hours.

DISTRIBUTION. Steady-state levels are reached in 48 hours. Quetiapine has a wide volume of distribution, approximately 10 L/kg.

METABOLISM AND ELIMINATION. Quetiapine is extensively metabolized by the liver. It has many metabolites, although only a small percentage of these are active. The primary metabolic pathway is through CYP3A4. Ninety-five percent of [^{3}H] quetiapine is recovered as metabolites in urine and feces, but only 1 percent is recovered as quetiapine itself.

Pharmacodynamics

MECHANISM OF ACTION. Quetiapine is a multitransmitter antipsychotic agent: It has a high affinity for 5-HT_2, histamine H_1, α_1, and α_2 receptors; a moderate affinity for 5-HT_{1A} receptors; a moderate affinity for D_2 and a low affinity for D_1 receptors. Quetiapine has a very low affinity for cholinergic M_1 and D_4 receptors (despite its low affinity for cholinergic receptors, quetiapine is associated with several typically anticholinergic side effects, such as dry mouth and constipation).

Quetiapine's low incidence of EPS may be explained by its relatively high 5-HT_{2A} binding compared to its activity at D_2 receptors, according to Meltzer. An early PET study demonstrated that 2 hours after a single clinical dose, quetiapine occupies 44 percent of D_2 receptors and 72 percent of 5-HT_2 receptors. Other PET studies demonstrate rapid dissociation of quetiapine from the D_2 receptor. In one study, D_2 receptor occupancy 2 hours post dose ranged from 58 to 64 percent, and fell to 20 to 0 percent 12 to 24 hours post dose for single doses of 400 and 450 mg, respectively. This supports the hypothesis of Kapur and colleagues who propose that quetiapine's "fast dissociation" (i.e., a high D_2 occupancy which decreases to low levels by the end of the dosing interval) may be sufficient to induce antipsychotic response while minimizing EPS.

DESIGN AND INTERPRETATION OF CLINICAL DRUG STUDIES

Schizophrenia: Acute Treatment. The earliest clinical trials demonstrated that quetiapine is significantly superior to placebo at reducing positive, negative, and depressive symptoms of schizophrenia. A post-hoc analysis of early placebo-controlled trials found that quetiapine separated from placebo on multiple clinical measures by the end of the first week of treatment. Later trials compared quetiapine to other FGAs and SGAs. Some of the early comparison trials suffered from quetiapine underdosing, with some study doses as low as 150 to 300 mg per day.

These studies generally found quetiapine to have comparable efficacy to haloperidol in reducing positive symptoms, but slightly less efficacy in reducing negative symptoms. Later studies began assessing quetiapine treatment at higher doses, up to 800 mg per day. Such studied mostly found comparable efficacy between quetiapine and haloperidol on positive and negative symptoms, and virtually no EPS or persistent prolactin elevation with quetiapine. In two separate 6- to 8-week, double-blind comparative studies, quetiapine (300 to 600 mg per day) had comparable efficacy to haloperidol in reducing psychotic symptoms. Both quetiapine and haloperidol produced clear reductions in mean PANSS scores, with more quetiapine-treated patients showing a clinical response (greater than 20 percent reduction in total PANSS score, $p = 0.043$). In addition, patients treated with quetiapine had a greater reduction in EPS ($p = 0.005$) and fewer treatment-emergent extrapyramidal-related adverse events ($p < 0.001$) compared with haloperidol-treated patients. In another study comparing multiple fixed doses of quetiapine to haloperidol and placebo, a dose range of 150 to 750 mg per day of quetiapine was comparable to haloperidol 12 mg per day in reducing positive symptoms, and doses of quetiapine above 300 mg per day were comparable in reducing negative symptoms. A post-hoc analysis of this data concluded that patients treated with quetiapine had significantly greater reductions of hostility and agitation compared to both placebo- and haloperidol-treated patients that were independent of the improvements in psychotic symptoms. An 8-week double-blind trial of 673 patients comparing quetiapine (mean dose 525 mg per day) and risperidone treatment (mean dose 5.2 mg per day) found that overall treatment efficacy did not differ between the two drugs, although there were some data to support greater positive symptom reduction with risperidone. Quetiapine had significantly lower rates of EPS and prolactin elevation but higher rates of somnolence. Another 6-week, double-blind trial in recently hospitalized patients with schizophrenia found that treatment with risperidone, but not quetiapine, was associated with a statistically significant separation from placebo by week 2 of treatment. Patients on quetiapine required higher rates of antipsychotic polypharmacy to control symptom exacerbation.

Schizophrenia: Maintenance Treatment. Quetiapine does not have FDA approval for the maintenance treatment of schizophrenia. However, quetiapine is frequently used off-label for maintenance treatment. Studies of quetiapine treatment in schizophrenia have followed patients for up to 18 months and found continued symptomatic improvement of positive, negative, and affective symptoms on par with improvements seen in acute treatment with quetiapine and the maintenance effects of other FGAs and SGAs. A small open-label study of 25 patients compared treatment with oral quetiapine (mean dose 493 mg per day) to treatment with haloperidol

decanoate (mean dose 170 mg per injection). Quetiapine treatment was associated with comparable prevention of symptom exacerbation over 48 weeks, less rigidity and akathisia, and greater improvement in negative symptoms over haloperidol decanoate treatment. There have been few direct comparisons between quetiapine and the other SGA medications. A large multinational naturalistic prospective study of 5,883 participants compared the efficacy of haloperidol, olanzapine, risperidone, and quetiapine at 6 months of treatment. Quetiapine was found to have comparable efficacy to haloperidol but was less efficacious than olanzapine and risperidone. However, quetiapine was likely underdosed, with a mean study dose of 200 mg per day at baseline and 300 mg per day by 6 months.

An open-label study of 728 patients with schizophrenia compared quetiapine to risperidone treatment for 4 months. Quetiapine- and risperidone-treated patients showed similar improvements on positive and negative symptoms scales, and quetiapine-treated patients had greater improvements on depressive symptom scales. Quetiapine was not associated with any extrapyramidal side effects but had higher rates of dry mouth, somnolence, and dizziness. A 6-month study comparing treatment with quetiapine or olanzapine in 346 patients found that both medications demonstrated similar efficacy in the treatment of negative symptoms, but that olanzapine therapy was associated with significantly greater positive symptom reduction, improvements in measures of functional outcome, and reduced rates of medication discontinuation. Finally, a 1-year study comparing olanzapine, risperidone, and quetiapine in the treatment of 400 patients with early onset psychosis found no differences in the all-cause discontinuation rate among the drugs. However, positive symptoms reductions were greater in the olanzapine and risperidone groups at some (but not all) time points.

These studies suggest that when appropriately dosed, quetiapine has equivalent or near equivalent efficacy compared with FGAs or other SGAs. Quetiapine appears to demonstrate a tendency toward slightly weaker symptom control but a better movement disorder profile than most other antipsychotic medications.

CATIE Trial. Quetiapine did not separate from fluphenazine, risperidone, or ziprasidone on the primary measure of time to discontinuation in the first phase of CATIE. Treatment with quetiapine (mean dose 543.4 mg per day) was associated with higher rates of side effects like dry mouth and constipation. In phase II, patients who had discontinued an SGA due to lack of efficacy had a significantly longer time to discontinuation on olanzapine and risperidone than on quetiapine or ziprasidone. However, in patients who had discontinued the first-generation medication due to lack of efficacy, time to discontinuation was longer in the quetiapine and olanzapine groups than the risperidone group.

This latter finding suggests that quetiapine, because of its low incidence of EPS, may have advantages for patients who have difficulty tolerating FGAs.

Bipolar Depression. Two 8-week, placebo-controlled trials of identical design evaluated treatment response to quetiapine at 300 mg per day and 600 mg per day fixed doses in 542 and 509 outpatients with bipolar I or bipolar II depression. The results of both studies were similar. Both fixed doses demonstrated significant improvements in depressive symptoms, which separated from placebo by the first week of treatment. There were no statistically significant differences between the dosage groups. There was no difference between quetiapine- and placebo-treated patients in terms of manic switch rate. A subanalysis of one of the studies examined patients with rapid-cycling course and concluded that quetiapine 300 or 600 mg per day was significantly more efficacious than placebo at reducing symptoms of depression, anxiety, and related symptoms in this patient population.

Bipolar Mania. Two identical studies comparing adjunctive quetiapine treatment versus adjunctive placebo treatment in 191 and 402 bipolar manic patients on lithium or valproate found that adjunctive quetiapine treatment (mean dose 584 mg per day and 492 mg per day, respectively) was associated with significantly greater reductions in manic symptoms and significantly higher rates of remission than adjunctive treatment with placebo. In the second study, a statistically greater degree of improvement in manic symptoms was evident in the adjunctive quetiapine group by the end of the first week of treatment. A study of 302 patients randomized to quetiapine, lithium, or placebo monotherapy treatment concluded that quetiapine was significantly more effective at reducing manic symptoms than placebo. Although the study was not powered to detect differences between lithium and quetiapine treatment, post-hoc analyses found no significant differences in symptom reduction between patients treated with quetiapine versus lithium. Another placebo-controlled study using haloperidol as the active control found that quetiapine was slightly less effective at reducing symptoms and achieving remission than haloperidol. The comparative analysis was post-hoc and some differences between the medications were not statistically significant. A placebo-controlled trial of 403 patients with mania concluded that quetiapine was associated with higher rates of symptom remission than was placebo at both 3 and 12 weeks.

Other Diagnoses. Quetiapine is used off-label for a number of other diagnoses and symptoms, despite limited controlled data. Three small open-label studies have evaluated the effects of treatment with quetiapine for chronic PTSD. In all three studies, quetiapine significantly reduced the

severity of PTSD symptoms in patients with and without psychotic features.

There are several small trials of quetiapine treatment of OCD. A small single-blind, placebo-controlled study found that adjunctive treatment with quetiapine and an SSRI was more efficacious at reducing symptoms of OCD than placebo and SSRI.

There are a few case reports and small open-label trials of quetiapine in the treatment of borderline personality disorder. The dose ranges in these studies varied from about 200 to 800 mg per day. Overall, the studies reported that treatment with quetiapine improved multiple symptom clusters, most notably mood, impulsivity, and aggression. There have been no randomized, double-blind, controlled trials in this population, to date.

Several studies have assessed quetiapine's effect on reducing aggression associated with psychiatric illness. A small study of adolescents with both bipolar disorder and disruptive behavior disorders found that quetiapine and divalproex had equivalent efficacy in reducing impulsivity and reactive aggression.

Anecdotally, quetiapine is considered the antipsychotic drug of choice to treat psychosis in Parkinson disease. Although quetiapine does not appear to worsen parkinsonism, in two placebo-controlled trials it failed to show efficacy in treating the psychotic symptoms associated with dopaminergic treatment of Parkinson disease. One study comparing quetiapine to clozapine (the only antipsychotic medication with proven efficacy and tolerability in Parkinson psychosis) found comparable levels of efficacy between the two drugs. Because of the high 5-HT_{2A} blockade and relatively low D_2 binding, quetiapine has been used off-label as first-line treatment psychosis in Parkinson disease. Quetiapine is more easily administered than clozapine in patients with Parkinson disease, but treatment efficacy has not been sufficiently established. The sedating effects and potential EPS are treatment limiting in this population.

Quetiapine Extended Release Formulation. A study comparing quetiapine XR (once daily dosing) to quetiapine immediate release (twice daily dosing) and placebo found comparable improvements with both forms of quetiapine preparation compared to placebo. There were no adverse events specific to the XR formulation.

Effects on Specific Organs and Systems

Cardiovascular System. Postural hypotension, likely due to quetiapine's high affinity for adrenergic α_1 receptors, is the most common cardiovascular effect. Tachycardia also occurs at rates above placebo; whether the tachycardia is secondary to postural hypotension or a separate effect is not clear. Quetiapine is not associated with clinically important increases in the QTc interval.

Respiratory and Ear, Nose, and Throat System. Dyspnea, cough, pharyngitis, rhinitis, and nasal congestion have been reported with rates slightly above placebo, which is observed with all SGAs.

Gastrointestinal System. Dry mouth, constipation, and dyspepsia are commonly reported side effects of quetiapine, with dry mouth being an especially frequent side effect. Abdominal pain is less common, but it (along with dyspepsia) is a dose-related effect.

Hematological and Lymphatic System. Leukopenia and neutropenia in conjunction with quetiapine therapy have been observed. Very rare cases of agranulocytosis, including fatal cases, have also been described.

Central Nervous System. Sedation, somnolence, and dizziness are frequently reported effects, with lethargy slightly less so. Headache and to a lesser degree asthenia are also commonly reported.

Metabolic and Nutritional System. Peripheral edema has been described with quetiapine and other antipsychotics. Hyperlipidemia, hypertriglyceridemia, and hyperglycemia are common effects of quetiapine. These have been attributed to the H_1 and 5-HT_{2C} binding as in other multireceptor antipsychotics.

Therapeutic Indications

Schizophrenia. Quetiapine is indicated for the acute treatment of schizophrenia, although in clinical practice it is used both acutely and for maintenance treatment.

Bipolar Disorder. Quetiapine is indicated for the acute treatment of manic episodes in bipolar I disorder, either as monotherapy or as adjunctive treatment with lithium or valproate. It is also indicated for monotherapy for the acute treatment of depressive episodes associated with bipolar I or II disorder. Quetiapine is indicated for the maintenance treatment of bipolar I disorder, as an adjunct to lithium or divalproex.

Off-Label Uses. Quetiapine is used in clinical practice to treat a number of symptoms. Typically in these cases clinicians are taking advantage of quetiapine's sedative and somnolent properties to treat insomnia, anxiety, and agitation. Usually the doses used for such symptomatic indications are between 12.5 and 100 mg per day, and insomnia will often respond to a dose range of 25 to 50 mg at bedtime. Sleep induction may be mediated by H_1 blockade, while improvement in sleep maintenance insomnia (i.e., decreased slow wave sleep) may be caused by 5-HT_{2A} antagonism. This mechanism in insomnia is similar to

the effects of low-dose doxepin (Sinequan). Although quetiapine benefits from having limited abuse potential, the possible metabolic risks of quetiapine use must be weighed against the benefits of a nonaddictive medication. Having lower doses for these indications means that the metabolic risks are also lower, but still require regular assessment.

Quetiapine is used to treat nonpsychotic disorders such as major depressive disorder, OCD, and borderline personality disorder.

Anecdotally, quetiapine is considered the antipsychotic drug of choice to treat psychosis in Parkinson disease, although efficacy failed to be established in two placebo-controlled trials.

Precautions and Adverse Reactions

Black Box Warnings

INCREASED MORTALITY IN ELDERLY PATIENTS WITH DEMENTIA-RELATED PSYCHOSIS. Quetiapine and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis compared with placebo.

SUICIDAL THOUGHTS AND BEHAVIORS. After quetiapine received an indication for the treatment of depression in bipolar disorder, it received the same black box warning as other antidepressant medications. In 2004, the FDA issued a warning about suicidal ideation and antidepressant therapy in the pediatric and young adult population. Although quetiapine was not one of the agents included in the 24 trials that led to the FDA black box warning, the impact of quetiapine on suicidal ideation in patients with depression or anxiety has not been assessed. Any patient, especially patients younger than 24 years, treated with medication for depression or severe anxiety should be monitored for the emergence or worsening of suicidal ideation. Emergent or worsening symptoms of anxiety, panic attacks, agitation, irritability, hostility, aggressiveness, impulsivity, akathisia, insomnia, hypomania, and mania may precede the onset of suicidal ideation and should be evaluated carefully at each visit. Patients started on treatment of depression or severe anxiety should be monitored frequently for the first 2 to 3 months of treatment or until remission occurs.

Neuroleptic Malignant Syndrome. Despite quetiapine's very low rate of EPS, rare cases of NMS have been reported with quetiapine use. Physicians should remain aware that even medications with low to nonexistent EPS risk can still occasionally cause NMS. Furthermore, several cases of high creatine kinase levels in the absence of clear NMS have been reported with quetiapine.

Tardive Dyskinesia. Quetiapine has a very low rate of tardive

dyskinesia. Although there are some reported cases of presumed quetiapine-induced tardive dyskinesia, there are many more reports of quetiapine being an effective treatment of tardive dyskinesia, much like clozapine. However, the data supporting quetiapine as a treatment of tardive dyskinesia are equivocal.

Movement Side Effects and Disorders. Since its release in 1998, most studies suggest that quetiapine, along with clozapine, has the lowest rate of extrapyramidal side effects of all the SGAs. This is thought to be the result of quetiapine's low and transient D_2 occupancy and its relatively high 5-$HT_{2A/2C}$ binding.

Hypercholesterolemia and Hypertriglyceridemia. In the industry-sponsored efficacy and safety trials of quetiapine, 16 percent of patients with schizophrenia and 9 percent of patients with bipolar disorder developed hypercholesterolemia, compared with 7 and 6 percent of placebo-treated patients. The rates of hypertriglyceridemia were higher; 23 percent of people with schizophrenia and 14 percent of people with bipolar disorder, versus placebo rates of 16 and 9 percent. In phase I of the CATIE trial, quetiapine-treated patients experienced exposure-adjusted mean increases in cholesterol and triglycerides of 6.6 and 21.2 mg/dL, respectively. These increases were second only to olanzapine, and clinically significantly higher than those seen with risperidone or perphenazine. Severe cases of hypertriglyceridemia, with levels above 400 mg/dL, are not uncommon with quetiapine therapy. Patients on quetiapine should be monitored regularly (at baseline and every 6 months) for changes in lipid parameters.

Hyperglycemia and Diabetes Mellitus. Quetiapine is associated with increases in plasma glucose and an increased risk of diabetes, but the degree of glucose increase is less dramatic than that of lipid and triglyceride increases with quetiapine. In two long-term treatment studies, 10.7 percent of patients on quetiapine developed fasting glucose measurements greater than 126 mg/dL, versus 4.6 percent of placebo-treated patients. During phase I of the CATIE trial, patients treated with quetiapine had exposure-adjusted mean increases in blood glucose of 7.5 mg/dL, an increase that was not significantly larger than that seen with risperidone or perphenazine. Patients treated with quetiapine should be monitored regularly for increases in fasting glucose above 100 mg/dL.

Changes in Weight. Quetiapine is associated with significant increases in weight, likely related in part to its high affinity for histaminergic H_1 receptors. In four 3- to 6-week trials of patients with schizophrenia, 23 percent of quetiapine-treated patients and 6 percent of placebo-treated

patients gained more than 7 percent of their body weight. Similar rates were seen in patients with bipolar disorder treated with either quetiapine or placebo. However, the increases are not as extreme as those seen with olanzapine or clozapine. In phase I of the CATIE study, 16 percent of patients treated with quetiapine gained over 7 percent of their body weight. The mean weight gain for quetiapine-treated patients was 1.1 lb from baseline to last observation.

Orthostatic Hypotension and Syncope. Quetiapine is a potent antagonist at the α_1 adrenergic receptor, contributing to its risk of orthostatic hypotension and syncope. Other related effects include dizziness and tachycardia. These symptoms are most pronounced during the initial dose titration and are minimized by starting doses of 25 mg twice a day or less. In the elderly, even lower initial doses may be indicated. Patients with known cardiovascular disease, cerebrovascular disease, or conditions that increase the risk of hypotension should be initiated on quetiapine with caution and frequent monitoring.

Cataracts. Studies on dogs at quetiapine doses four times the maximum human dose found associations between long-term (6- or 12-month) quetiapine use and cataracts. Studies in humans have failed to find similar associations between cataracts and quetiapine use. The product insert for quetiapine recommends slit lamp examinations at treatment initiation and every 6 months thereafter; it is unclear if these recommendations are followed regularly in clinical care or whether the health care costs of such frequent assessment are justified by the hypothetical human risk of cataracts.

Thyroid Hormone Changes. Quetiapine affects total and free thyroxine (T_4) levels in a dose-dependent manner. At the high therapeutic dose range of quetiapine, T_4 levels may be decreased by as much as 20 percent. These changes are generally not clinically significant, nor do they typically result in changes to TSH or thyroid-binding globulin (TBG). Less than 1 percent of patients had changes in TSH while on quetiapine. T_4 levels generally return to baseline with discontinuation of quetiapine. Patients on lithium therapy who are started on adjunctive treatment with quetiapine are likely at greater risk of thyroid hormone fluctuation than those not started on quetiapine adjunctively.

Elevated Transaminases. Quetiapine has been associated with transient and asymptomatic elevations in serum transaminases, especially ALT. In 3- to 6-week trials of patients with schizophrenia, 6 percent of people on quetiapine had elevations in transaminases of over three times the normal limit, compared with 1 percent of people on placebo. The transaminase

elevations typically occurred by the third week of treatment, and levels returned to baseline with continued quetiapine administration.

Somnolence and Sedation. Quetiapine can cause significant sedation and somnolence, especially during initial dose titration. In industry-sponsored studies, the rates of somnolence in the schizophrenia, bipolar mania, and bipolar depressed groups ranged from 18 to 34 percent of patients treated with quetiapine, versus 7 to 9 percent of patients treated with placebo. The percentage of patients on quetiapine reporting somnolence in the CATIE trial was 31 percent, but this was not significantly different from the other antipsychotic treatment groups. The sedation and somnolence are likely secondary to quetiapine's high affinity for histaminergic H_1 receptors.

Pregnancy. Quetiapine is rated pregnancy category C. Three cases studies of pregnant women treated with quetiapine throughout all or part of their pregnancies reported no noticeable effects of quetiapine treatment on the neonates. A study that compared the pregnancy outcomes of 151 women on all SGA medications found no increased risk of malformation or spontaneous abortion, but a small but statistically significant risk of low birth weight in neonates prenatally exposed to SGAs. In a study measuring the degree of placental passage of SGAs, the authors concluded that quetiapine had the lowest placental passage, as compared to haloperidol, olanzapine, and risperidone. The majority of the evidence available suggests that quetiapine is safe to use during pregnancy. However, the data are limited to case studies and small trials, and infant outcomes have not been assessed beyond the 3- to 6-month mark. Physicians should carefully discuss and document the risks and benefits to antipsychotic therapy during pregnancy with their patients, and pregnant patients treated with antipsychotic medication should be frequently monitored by both a psychiatrist and obstetrician throughout the pregnancy.

Lactation. A single case study of quetiapine excretion in human breast milk concluded that by 2 hours post dose, the level of quetiapine in breast milk was almost at the pre dose level. The calculated infant dose based on the levels collected pre dose and again at 1, 2, 4, and 6 hours post dose ranged from 0.09 to 0.43 percent of the weight-adjusted maternal dose. The infant breastfed by this mother showed no signs of adverse effects from the mother's continued quetiapine treatment. Another case study on a breastfed infant found the weight-adjusted infant dose was .1 percent of the maternal dose, and no adverse effects on the infant were noted with continued breastfeeding. The available data on the excretion of quetiapine in human breast milk and the consequent effects on infant development are sparse and limited to case studies. Patients should be advised about the

potential risks of breastfeeding while on psychotropic medications.

Dosage and Administration

Schizophrenia. Quetiapine dosages should be titrated to avoid excessive somnolence, postural hypotension, and syncope. The manufacturer's recommended dosing schedule is as follows: 25 mg twice daily on day 1, 75 to 100 mg in divided doses by day 2 or 3, for a total dose of 300 to 400 mg in divided doses by day 4. Dose increments above day 4 should occur no more rapidly than every 2 days (however, other studies have shown that a more aggressive titration is well tolerated and associated with a good response). The target dose range is 150 to 800 mg per day. All efficacy studies occurred with twice or three times a day dose schedules, but in the community quetiapine is frequently administered once daily at bedtime. Dosages above 800 mg have not been systematically evaluated, but doses up to 1,200 mg per day have been used clinically.

Mania. In acute mania, the manufacturer recommends an initial dose of 100 mg in a divided dose, with a rapid titration of up to 100 mg increases per day in divided doses, to a total dose of 400 mg by day 4. The dose can be escalated by up to 200 mg increases on day 5 and 6, for a total dose of 800 mg by day 6. Children and adolescents (10 to 17 years) should start with 25 mg twice daily and should reach a total dose of 400 mg by day 5. Patients should be monitored for orthostatic hypotension and syncope. The effective dose range for mania is 400 to 800 mg per day (400 to 600 mg per day in children and adolescents). Inpatients may benefit from the daytime sedative effects of twice-daily dosing, and as an outpatient, community practice is once a day at bedtime dosing.

Depression. The recommended titration schedule for bipolar depression is as follows: Day 1, 50 mg at bedtime; day 2, 100 mg at bedtime; day 3, 200 mg at bedtime; day 4, 300 mg at bedtime. Patients being titrated beyond 300 mg can receive 400 mg by day 5 and 600 mg by day 8. In efficacy trials no additional antidepressant benefit was noted above doses of 300 mg. The efficacy trials for bipolar depression dosed quetiapine once daily at night.

The manufacturer's recommended titration schedules should be modified for people who demonstrate significant orthostatic hypotension, somnolence, or have episodes of syncope.

Quetiapine Extended Release. The extended release formulation of quetiapine is intended for once-daily dosing at bedtime. The recommended initial dose is 300 mg. The dose can be titrated up to 800 mg as rapidly as by day 3. The tablets should not be crushed or cut. It is recommended that quetiapine extended release be taken without food or with no more than

300 calories at time of administration, as high calorie or fat meals cause increased plasma concentrations.

Dosing in Special Populations

ELDERLY. Elderly patients are generally more sensitive to quetiapine's effect on blood pressure. In this population orthostatic hypotension or syncope can increase the risk of falls. Furthermore, metabolism of all drugs is slowed in geriatric populations. The titration schedule in the elderly should be slowed and the target doses lower. Initial doses should be 12.5 to 25 mg per day, and increments of no more than 25 mg should occur every 1 to 3 days, depending on tolerability. Quetiapine is not indicated for patients with psychosis due to dementia.

HEPATIC IMPAIRMENT. Patients with hepatic impairment should be initiated at doses no higher than 25 mg per day. The dose can be increased daily by 25 to 50 mg to an effective dose.

Drug Interactions

Quetiapine is metabolized by the cytochrome P450 system. CYP 3A4 is the major metabolic pathway, and CYP 2D6 is a minor pathway. Quetiapine is neither an inhibitor nor an inducer of any CYP enzyme system. Cimetidine (a nonspecific enzyme inhibitor) did not inhibit quetiapine metabolism. Phenytoin, a strong inducer of CYP3A4, caused a fivefold increase in quetiapine clearance. Similarly, ketoconazole, a strong CYP3A4 inhibitor, increased quetiapine levels by 335 percent. Valproic acid increased quetiapine plasma levels by a clinically insignificant amount. Lithium, lorazepam (or other benzodiazepines), risperidone, haloperidol, fluoxetine, and imipramine (Tofranil) had no effects on the pharmacokinetics of quetiapine and vice versa (i.e., did not require dose adjustments). Concomitant administration of thioridazine (Mellaril), however, increased oral clearance of quetiapine by 65 percent.

Overdose

Quetiapine is relatively safe in overdose. Acute nonfatal overdoses up to 30 g have been described. One fatal overdose involved ingestion of over 13 g of quetiapine. The effects of overdose are exaggerations of the medication's side-effect profile: Sedation and somnolence, hypotension, and tachycardia.

ZIPRASIDONE

History

Ziprasidone was the fifth SGA approved in the United States. It was

synthesized in 1987 by Harry Howard of Pfizer Pharmaceuticals. Tom Seeger screened the compound and identified the activity at D_2 and 5-HT_2 receptors. Early on it was noted that ziprasidone differed from the other agents in this class by not being associated with weight gain and having relatively little sedative side effect. Its approval in the United States was delayed due to concerns regarding its tendency to prolong the QT interval on the ECG. The FDA required Pfizer to document the extent of the problem in a large trial that evaluated the effects of several antipsychotics on the QT interval. The results of this study led to ziprasidone's approval in the United States in 2001.

Concerns about weight gain, diabetes, and lipid elevations led to an increase in the use of ziprasidone. For example, in the NIMH CATIE trial, patients randomized to ziprasidone showed negligible changes in weight, glucose, or lipids. Moreover, patients who gained weight on other antipsychotics were likely to show substantial weight loss when they were switched to ziprasidone.

Chemistry

Ziprasidone is a benzisothiazolyl piperazine (Fig. 34.28–1) with a structure unlike any of the other SGAs.

Pharmacological Action

Pharmacokinetics of Oral Ziprasidone

ABSORPTION. Ziprasidone is well absorbed when orally administered. Bioavailability doubles when ziprasidone is administered with food. Therefore, patients should be instructed to take ziprasidone with a meal. After administration of multiple doses of ziprasidone under fed conditions, peak serum concentrations occur at 6 to 8 hours.

DISTRIBUTION. Ziprasidone has a mean volume of distribution of 1.5 L/kg. It is more than 99 percent protein bound.

METABOLISM AND ELIMINATION. Ziprasidone is extensively metabolized, with less than 1 percent excreted unchanged in the urine and less than 4 percent in feces. In vitro studies indicate that the formation of the major metabolites is mediated by CYP3A4, although there is a small contribution from CYP1A2. Ziprasidone's metabolites are for the most part inactive, representing less than 1 percent of the binding affinity of the parent compound for D_2 and 5-HT_{2A} receptors. Unmetabolized ziprasidone represents 44 percent of the circulating drug-related material.

Steady state is attained after 1 to 3 days of dosing. Overall exposure at steady state is relative to the dose over the range of 20 to 80 mg twice a day. The mean terminal plasma half-life for ziprasidone at steady state

ranges from 5 to 10 hours in volunteers and in schizophrenia patients. The 0.5- to 1-hour lag in reaching peak plasma concentrations has been attributed to delayed gastric emptying. Age, gender, or mild to moderate renal or hepatic impairment have no significant effect on the pharmacokinetics of ziprasidone, making dosage adjustments unnecessary.

Pharmacokinetics of IM Ziprasidone

ABSORPTION AND DISTRIBUTION. Intramuscularly administered ziprasidone, as with all parenterally administered drugs, is 100 percent bioavailable. Serum concentrations of ziprasidone reach their peak approximately 1 hour after IM administration, with a half-life from 2 to 5 hours.

METABOLISM AND ELIMINATION. After IM administration, metabolism and elimination are anticipated to occur via the same pathways as with oral administration.

Pharmacodynamics

MECHANISM OF ACTION. In vitro studies show that while ziprasidone is strong D_2 antagonist, it is an even more potent antagonist at the 5-HT_{2A} receptor, with a very high 5-HT_{2A} to D_2 ratio of 11. PET studies in healthy volunteers confirm that 5-HT_2 occupancy substantially exceeds D_2 occupancy. Ziprasidone also has potent affinity for the D_3 receptor and moderate affinity for the D_4 receptor, similar to haloperidol; the affinity for the D_1 receptor is 100-fold lower than for other dopamine receptors. Ziprasidone is an agonist at the 5-HT_{1A} receptor and a potent antagonist at 5-HT_{2C} (five times greater than clozapine and 19 times greater than risperidone) and 5-HT_{1D} receptors. It also moderately affects norepinephrine and serotonin (and dopamine) uptake sites. Ziprasidone has only moderate affinity for the α_1 receptor and low affinity for the H_1 receptor. Ziprasidone is the only agent providing uptake inhibition of norepinephrine and serotonin. Stimulation of 5-HT_{1A} receptors is shared with clozapine.

BLOOD CONCENTRATIONS AND RELATION TO ACTION. With twice-daily administration of oral ziprasidone, estimates based on systemic exposure and PET scan data predict that 5-HT_2 and D_2 receptor occupancy ranges from 80 to 90 percent and 45 to 75 percent, respectively. Ziprasidone at serum concentrations of 30 to 40 ng/mL is likely to be associated with approximately 75 percent 5-HT_2 and more than 65 percent D_2 receptor occupancy, as determined from PET scan data.

Early Clinical Trials in Schizophrenia. Results from 4- to 6-week, fixed-dosage, double-blind clinical trials indicate that oral ziprasidone is effective in the treatment of positive, negative, and affective symptoms in patients with schizophrenia and schizoaffective disorder. In two separate 4- to 6-week clinical trials in patients with acute psychotic exacerbation of schizophrenia, oral ziprasidone (80 to 160 mg per day) was significantly superior to placebo, and, in a 4-week trial, oral ziprasidone (160 mg per day) was comparable to haloperidol (15 mg per day) in reducing psychotic symptoms. Statistically significant, dose-dependent improvements in negative symptoms, as measured by the PANSS, have been observed with oral ziprasidone (80 and 160 mg per day) compared with placebo in a 6-week trial. Oral ziprasidone (160 mg per day) also significantly improved depressive symptoms in patients with clinically significant symptoms at baseline (Montgomery–Asberg Depression Rating Scale [MADRS] 14 or more; baseline mean, 23.5) compared with placebo at 6 weeks. The statistically significant improvement has also been observed in affective symptoms at 4 weeks in patients with a BPRS anxiety-depression cluster baseline score greater than or equal to 18. Further evidence for efficacy in negative symptoms comes from a 4-week, placebo-controlled trial in which a significant improvement in the BPRS anergia factor was observed with ziprasidone at 120 mg per day. A dose-dependent effect was not observed, which suggests that the negative symptom reduction effect may be related to its lack of extrapyramidal side effects.

A careful look at these short-term trials provides valuable information about the dosing of ziprasidone. In most of the fixed dose studies, less than 40 mg twice daily was not superior to placebo; results for 40 mg twice daily were mixed. This suggests that for most patients, dosing of ziprasidone should be targeted at above 40 mg twice daily.

CATIE Trial and Ziprasidone. In phase I of CATIE, discontinuation rates on ziprasidone were similar to those for patients treated with the other SGAs and perphenazine. Ziprasidone had the best metabolic profile among the phase I agents. Ziprasidone-treated patients showed slight weight loss, as well as small reductions in cholesterol and triglycerides and trivial increases in glucose levels. Among patients who gained more than 7 percent of their baseline body weight during phase I, 42 percent of those assigned to ziprasidone during phase II lost 7 percent of their new weight.

Long-Term Effectiveness in Schizophrenia. Oral ziprasidone appears to be effective in the long-term treatment of stable patients with chronic schizophrenia according to double-blind, randomized studies. Results from a 28-week, fixed-dose comparative study indicate that ziprasidone is comparable to haloperidol in reducing overall psychopathology during long-term treatment of schizophrenia. Ziprasidone, however, demonstrated

superior tolerability and control of negative symptoms.

Results from a prospective, 1-year double-blind, fixed-dose ziprasidone study in patients with chronic or subchronic schizophrenia showed that oral ziprasidone (40 to 160 mg per day) is associated with significantly lower relapse rates compared with placebo. This group of stable patients experienced a significant reduction in negative symptoms from baseline compared with placebo, as measured with the negative symptom subscale of the PANSS. This improvement in negative symptoms was maintained throughout 12 months of treatment, with a tendency for continuing improvement in those patients completing the study.

Bipolar Mania. A study reported by Paul Keck and coworkers randomized 210 patients with mania or mixed states to a 3-week trial of ziprasidone or placebo. Dosing was flexible. From day 8 to 14 the mean ziprasidone dose was 139 mg; between days 15 to 21, it was 130.1. Ziprasidone-treated patients demonstrated significant improvements compared to placebo-treated patients on a Mania Rating Scale and the CGI. Improvements on ziprasidone emerged by day 2. Advantages on the same instruments were found in a second 3-week controlled study.

Agitation in Schizophrenia. Two double-blind studies evaluated the efficacy of IM ziprasidone for reducing agitation in schizophrenia during a 24-hour period. Patients could receive up to four injections 4 hours apart. Both studies used the Behavioral Activation Scales (BARS) to evaluate outcome. One study found that 20 mg was more effective than the 2-mg dose and the other found that 10 mg was superior to 2 mg. Two other studies were open-label and compared 3 days of IM ziprasidone or haloperidol followed by 4 days on the same drug. A 6-week study administered IM ziprasidone or haloperidol for 3 days followed by oral treatment. Each of these studies found that IM ziprasidone had similar or better efficacy against agitation than haloperidol and caused fewer extrapyramidal side effects.

Effects on Specific Organs and Systems

General. Ziprasidone causes drowsiness in a relatively small proportion of patients.

Dermatological System. In premarketing trials 5 percent of patients treated with ziprasidone developed a rash or urticaria.

Cardiovascular System. Ziprasidone has substantial α_1-adrenergic antagonist effects, which may explain a tendency to cause orthostatic hypotension, tachycardia, and dizziness in some individuals. Ziprasidone can cause prolongation of the QTc interval (see Precautions and Adverse

Reactions).

Gastrointestinal System. Occasionally, patients experience anorexia and nausea.

Endocrine. Ziprasidone can increase prolactin in some patients. However, the effects are less prominent than FGAs or risperidone.

Therapeutic Indications

Schizophrenia. Ziprasidone is effective for the treatment of schizophrenia. The evidence supporting its efficacy is from 4- to 6-week trials indicating greater efficacy than placebo. A large 52-week double-blind trial also found that ziprasidone was more effective than placebo for preventing relapse in stabilized patients with schizophrenia.

Bipolar I Disorder. Ziprasidone is also indicated as monotherapy for the acute treatment of manic or mixed episodes in bipolar I disorder. It is also indicated for the maintenance treatment of bipolar I disorder, as an adjunct to lithium or valproate. The efficacy of ziprasidone as monotherapy for maintenance therapy has not been systematically evaluated.

Acute Treatment of Agitation in Schizophrenia. IM ziprasidone is effective for managing agitation in patients with schizophrenia. The advantage of IM medication is a more rapid onset of effects. The product information for ziprasidone has the following description of the type of individual where ziprasidone may be useful: "Schizophrenic patients experiencing agitation often manifest behaviors that interfere with their diagnosis and care, for example, threatening behaviors, escalating or urgently distressing behavior, or self-exhausting behavior, leading clinicians to the use of IM antipsychotic medications to achieve immediate control of the agitation."

Precautions and Adverse Reactions

Black Box Warning. Ziprasidone and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis compared with placebo.

QT Prolongation. Ziprasidone is associated with a dose-related prolongation of the QT interval on the ECG. This is a concern since a prolonged QT interval can lead to the development of a characteristic ventricular tachycardia known as torsade de pointes. This arrhythmia can result in hypotension, or more seriously, ventricular fibrillation and death. There is a concern that if ziprasidone was administered in combination with other agents associated with QTc prolongation, the effects could be

additive. As a result, the product information for ziprasidone warns against combining it with other drugs that may increase the QT interval including dofetilide (Tikosyn), sotalol (Betapace), quinidine, other class Ia and III antiarrhythmics, mesoridazine (Serentil), thioridazine, chlorpromazine, droperidol (Inapsine), pimozide (Orap), sparfloxacin (Zagam), gatifloxacin (Tequin), moxifloxacin (Avelox), halofantrine (Halfan), mefloquine (Lariam), pentamidine (Nebupent), arsenic trioxide (Trisenox), levomethadyl acetate (Orlaam), dolasetron mesylate (Anzemet), probucol (Lorelco), or tacrolimus (Prograf). Ziprasidone should also not be prescribed for patients with existing QT prolongation. In addition, it should not be prescribed in patients who have conditions such as hypokalemia and hypocalcemia, which can prolong the QT interval.

An interesting study that was required prior to the approval of ziprasidone by FDA gives a perspective on the amount of QT prolongation associated with ziprasidone and other antipsychotics. Patients receiving six antipsychotics had their QT interval measured at their maximal plasma concentration. This was done in the presence and absence of a metabolic inhibitor of the antipsychotic in question (in the presence of an inhibitor, the plasma concentration of the antipsychotic drug increases). Ziprasidone administration resulted in a mean increase in the QTc (that is the QT interval corrected for heart rate) of 9 to 14 ms greater than four of the comparator drugs (risperidone, olanzapine, quetiapine, and haloperidol) but 14 ms less than thioridazine. This finding was sufficiently reassuring to lead to the approval of ziprasidone.

Use in Pregnancy and during Lactation. Ziprasidone is considered category C for pregnancy. Although there are no well-controlled studies on the risk of ziprasidone to the fetus, animal studies suggest that there is a potential for teratogenic effects from ziprasidone. Unlike the other older SGAs, little is published on the effects of ziprasidone during pregnancy and lactation. Therefore, ziprasidone should only be taken during pregnancy if the clinical benefits justify the risk to the fetus.

Other Side Effects. Double-blind comparisons of ziprasidone and placebo have not found a statistically significant difference in EPS or akathisia. However, case reports suggest that some patients experience an increase in restlessness that may be mild akathisia.

Drug Interactions

Ziprasidone should not be combined with drugs that prolong the QT interval. Otherwise, ziprasidone has low potential for clinically significant drug interactions. In vitro studies indicate that ziprasidone is not a substrate for CYP2D6, CYP2C9, CYP2C19, or CYP1A2 and has low potential to inhibit CYP1A2, CYP2C9, CYP2C19, CYP2D6, and CYP3A4. The lack of

clinically relevant CYP2D6 inhibition has been confirmed in healthy volunteers, in whom ziprasidone did not inhibit metabolism of dextromethorphan (Benylin) (a CYP2D6 model substrate) to its major metabolite, dextrorphan. Although ziprasidone is a substrate for CYP3A4, cimetidine (a CYP3A4 inhibitor) does not inhibit its metabolism, suggesting that significant metabolism occurs by an alternative pathway. Carbamazepine induces CYP3A4 and lowers ziprasidone plasma concentrations. In healthy subjects, clinically significant interactions have not been reported with lithium, cimetidine, antacids (aluminum and magnesium hydroxides), ketoconazole, or oral contraceptives. Ziprasidone did not affect steady-state lithium concentrations.

Laboratory Interference

Ziprasidone does not affect common laboratory tests.

Dosage and Administration

Schizophrenia. For most adult patients, ziprasidone should be initiated at 20 mg twice a day with food. Dosage may be increased every 2 days. As mentioned earlier, trials in schizophrenia indicate that patients receiving a total daily dose of 100 mg or greater are more likely to respond. Daily dosage can be adjusted on the basis of clinical status up to a maximum of 80 mg twice a day.

Duration of Treatment. The benefits of ziprasidone in schizophrenia have been found to be maintained for periods of up to 1 year when given at dosages ranging from 20 to 80 mg twice a day, although in these studies no additional benefits were seen at dosages higher than 20 mg twice a day. Ziprasidone has been administered safely for up to 2 years.

Bipolar I Disorder. For the acute treatment of manic or mixed episodes, ziprasidone should be given at an initial dose of 40 mg twice daily with food. It may be increased to 60 mg or 80 mg twice daily on day 2, and titrated according to efficacy and tolerance. In flexible-dose trials, the mean daily dose used was around 120 mg. For maintenance treatment as an adjunct to lithium or valproate, patients should be continued at the same dose as initial stabilization, within the range of 40 to 80 mg twice daily with food.

Acute Treatment of Agitation in Schizophrenia (IM Formulation). The recommended IM dosage is 10 to 20 mg administered every 2 hours for the 10-mg dose and every 4 hours for the 20-mg dose, up to a maximum dosage of 40 mg per day. IM administration has not been studied for longer than 3 consecutive days, and it is recommended that oral therapy be initiated if longer-term treatment is necessary.

ARIPIPRAZOLE

Aripiprazole differs from other SGAs due to its partial agonist activity at D_2 and 5-HT_{1A} receptors. It is similar to other SGAs in having a relatively high affinity for both 5-HT_{2A} and D_2 receptors. However, in contrast to other agents that are full antagonists at D_2 receptors, aripiprazole has substantial intrinsic activity. The development of this drug was based on the hope that its partial agonist activity would result in decreased dopamine activity in limbic brain areas and increased activity in frontal areas. This pattern of dopamine activation, it was hypothesized, would lead to an effective antipsychotic that could improve negative symptoms and improved cognition. The clinical importance of this different mechanism of action is as yet unclear. The available data indicate that aripiprazole is an effective agent with minimal sedation and metabolic effects. It has a relatively long half-life that allows once-daily dosing.

Aripiprazole has become a common alternative when patients and clinicians are concerned about metabolic effects. Although the introduction of a short-acting IM formulation has increased the usefulness of aripiprazole in acute settings, its lack of ability to sedate and its tendency to cause psychomotor restlessness in some individuals have led to some reluctance to prescribe aripiprazole for the most seriously ill psychotic patients.

History

The Nobel laureate Arvid Carlsson, most responsible for identifying the importance of dopamine in antipsychotic therapeutics, suggested that a dopamine partial agonist that acted at dopamine autoreceptors may be effective in treating psychosis. This led to a search by Tetsuro Kikuchi and Yasuo Oshiro at the Otsuka Pharmaceuticals laboratory in Tokushima, Japan, for partial dopamine agonists. Aripiprazole was synthesized in 1988 and entered phase I and early phase II studies in Japan in 1990, with later phase II trials in the United States beginning in 1993. In 1999, a collaborative arrangement was developed between Otsuka Pharmaceuticals and Bristol-Meyers Squibb for the worldwide development of aripiprazole.

Chemistry

Aripiprazole is a highly lipid-soluble quinolinone derivative. Its chemical structure does not resemble any of the available antipsychotics (Fig. 34.28–1).

Pharmacological Action

Pharmacokinetics

ABSORPTION. Aripiprazole is well absorbed after oral administration. Peak plasma concentrations are reached within 3 to 5 hours after administration. The bioavailability of aripiprazole after oral administration is 87 percent. Aripiprazole's absorption is not affected by administration with food.

IM administration of aripiprazole (short acting form) leads to peak plasma concentrations at 1 to 3 hours. The bioavailability is 100 percent.

DISTRIBUTION. Aripiprazole is more than 99 percent bound to serum proteins. Studies of its steady-state volume of distribution indicate that aripiprazole has extensive extravascular distribution.

METABOLISM AND ELIMINATION. The activity of aripiprazole is primarily due to the parent compound, although some activity is related to dehydroaripiprazole metabolite. The mean elimination half-life of aripiprazole is approximately 75 hours, and the mean half-life for dehydroaripiprazole is 96 hours. These relatively long half-lives support once-daily dosing for aripiprazole and indicate that patients should attain a steady-state plasma concentration within 14 days. Aripiprazole is extensively metabolized in the liver by dehydrogenation, hydroxylation, and *N*-alkylation. Metabolism is primarily by CYP 3A4 and CYP 2D6 enzymes.

Pharmacodynamics

MECHANISM OF ACTION. A partial agonist is an agent that binds to a receptor and induces a signal that is somewhat reduced compared with the endogenous ligand; this differs from an antagonist that occupies a receptor and completely blocks the signal from the endogenous ligand. The putative advantage of a partial agonist is that it may stabilize a neurotransmitter system by decreasing the amount of signaling when there is excessive ligand (by occupying receptors and transmitting a reduced signal compared to the endogenous ligand) and by increasing the signal when there is reduced activity (by acting as an agonist). For aripiprazole, D_2 partial agonism could translate into an agent that acts as a functional antagonist under hyperdopaminergic conditions and as an agonist under hypodopaminergic conditions. An advantage of this mechanism is that it could address concerns that schizophrenia may be an illness that is associated with excessive dopamine activity at sites in limbic areas and decreased dopamine activity in frontal and prefrontal areas. In addition, a partial D_2 agonist may not result in the severe hypodopaminergic state in striatal areas that results in extrapyramidal side effects.

The evidence that aripiprazole acts as a partial agonist at D_2 receptors is substantial. In vivo binding studies have found that aripiprazole binds with

high affinity to D_2 receptors. Studies in human volunteers have found that, at its clinically effective doses, aripiprazole occupies a similar proportion of D_2 receptors to that of other antipsychotic agents. Aripiprazole blocks stereotypy induced by apomorphine in rodents, as do other antipsychotics. Agonist effects of aripiprazole were identified in a study in which its administration resulted in a release of prolactin from rat pituitary slices. This property was supported in a model of dopamine hypoactivity in which aripiprazole decreased the reserpine-induced DOPA accumulation that results from stimulation of presynaptic D_2 receptors.

Aripiprazole also appears to be a partial agonist at 5-HT_{1A} receptors. The importance of this activity in schizophrenia is unclear. Because buspirone (BuSpar), also a 5-HT_{1A} partial agonist, is an effective anxiolytic, it is conceivable that aripiprazole may also decrease anxiety. Aripiprazole also binds with high affinity to D_3, 5-HT_{2A}, 5-HT_{2C}, and H_1 receptors.

Design and Interpretation of Clinical Drug Studies

Efficacy for Acute Schizophrenia. Double-blind trials have compared aripiprazole with placebo, risperidone, and haloperidol. The earliest published study evaluated 15 and 30 mg of aripiprazole, 10 mg of haloperidol, and placebo in 414 patients with acute schizophrenia and schizoaffective disorder. Over the 4 weeks of the trial, all three active groups were significantly superior to placebo and similar to one another on PANSS total and PANSS positive symptoms. For PANSS negative symptoms, 15 mg of aripiprazole and 10 mg of haloperidol resulted in significantly greater improvements compared with placebo. A subsequent study compared 20 and 30 mg of aripiprazole to placebo and 6 mg of risperidone in 406 patients with schizophrenia or schizoaffective disorder. All of the active conditions were superior to placebo for both positive and negative symptoms on the PANSS. Another study that evaluated 10, 15, and 20 mg of aripiprazole and placebo in 420 patients also found that all three active doses were more effective than placebo. Moreover, there were no significant differences among the three doses. The findings from these acute studies indicate that doses of 10, 15, 20, and 30 mg of aripiprazole are more effective than placebo. There is also no indication that 30 mg is more effective than the lower doses. These results are consistent with a study indicating that 10 mg of aripiprazole results in 85 percent D_2 occupancy.

A study reported by John Kane and collaborators evaluated aripiprazole in patients who were treatment resistant. Treatment resistance was confirmed in an open-label trial of risperidone or olanzapine. After 6 weeks of double-blind treatment with aripiprazole or perphenazine, there were no significant differences in the number of patients who met improvement criteria. However, in both groups, a substantial proportion of patients, 27 percent of aripiprazole and 25 percent of perphenazine patients, met these

criteria. This suggests that either strategy may be useful for patients who fail to respond well to risperidone or olanzapine.

Efficacy for Acute Mania. Aripiprazole at 30 mg daily was compared to placebo in two studies of bipolar disorder patients in an acute manic or mixed episode. Aripiprazole was superior to placebo in both improvement on the YMRS and in the proportion of patients who demonstrated substantial improvement during the trial.

Long-Term Effectiveness in Schizophrenia. Two trials have evaluated the long-term effectiveness of aripiprazole in schizophrenia. In one study, stabilized patients (N = 310) were randomly assigned to either 15 mg of aripiprazole or placebo and observed for relapse over a 6-month period. Sixty-four percent of aripiprazole-treated patients survived the 6-month period without relapsing ($p < 0.01$) compared with 43 percent of placebo-treated patients. This study indicates that aripiprazole is an effective agent for preventing relapse in stabilized patients with schizophrenia.

The second study evaluated 1,294 patients who had experienced a psychotic relapse. Patients were randomly assigned to 30 mg of aripiprazole or 10 mg of haloperidol and followed for 1 year to evaluate their recovery from acute psychosis and the subsequent maintenance of the response. Aripiprazole-treated patients were more likely to remain on their medication for the full 52 weeks. Moreover, aripiprazole treatment was associated with greater improvements over the course of the study in both positive and negative symptoms, as well as depression as measured by the MADRS.

Long-Term Effectiveness in Bipolar Disorder. Patients with bipolar disorder who were stabilized on aripiprazole following a recent episode of acute or mixed mania were randomly assigned to continue on aripiprazole or change to a placebo. Aripiprazole was superior to placebo in delaying manic and depressed episodes.

Agitation in Schizophrenia or Bipolar Disorder. Clinicians often prescribe sedating antipsychotics for patients who appear agitated. This common practice as well as the tendency of aripiprazole to cause restlessness in some individuals has led some to avoid prescribing this agent in agitated patients. A recent meta-analysis combining the data from several trials indicated that aripiprazole was effective in patients with relatively high and low levels of agitation.

A short-acting IM formulation of aripiprazole was evaluated in two studies of agitated patients with schizophrenia or schizoaffective disorder. In one study, patients (N = 357) were randomly assigned 1.0, 5.25, 9.75, or 15 mg of aripiprazole, 7.5 of haloperidol, or placebo. Although all of the

treatment groups except 1 mg of aripiprazole were more effective than placebo, the 9.75 mg had the best response. Patients in this group demonstrated improvement by 45 minutes following their injection. A second study (*N* = 448) compared 9.75 mg of aripiprazole, 6.5 mg of haloperidol, and placebo. Both aripiprazole and haloperidol were superior to placebo.

Adjunctive Treatment in Major Depression. Two double-blind, placebo-controlled trials evaluated adjunctive aripiprazole (2 to 15 mg per day) in patients with major depressive disorder who had inadequate responses to escitalopram (Lexapro), fluoxetine, paroxetine controlled-release, sertraline (Zoloft), or venlafaxine extended release (Effexor XR). Improvement on the MADRS was greater in the group that received aripiprazole. However, rates of akathisia and restlessness were higher in the aripiprazole group. The average doses in the two trials were 10.7 and 11.4 mg daily.

Other Studies of Aripiprazole. Aripiprazole is the only antipsychotic that may reduce levels of prolactin. A study from South Korea found that elevated levels of prolactin in patients receiving haloperidol could be reduced by adding 30 mg daily of aripiprazole. This suggests an interesting possible use of aripiprazole for patients who show their best antipsychotic responses to drugs that elevated prolactin and who have prolactin-related side effects.

Effects on Specific Organs and Systems

General. Aripiprazole is among the least sedating of available antipsychotics. Aripiprazole is also associated with only minimal weight gain; in the comparison with haloperidol and placebo and risperidone and placebo, there were no significant differences in weight change among the treatment groups. It also appears to have minimal effects on lipids and glucose. This suggests that aripiprazole may have an important advantage for patients in whom there is concern about weight gain. Since aripiprazole was not included in the CATIE study, there is insufficient information to compare its weight gain and lipid liabilities with other antipsychotics.

Neurological. Aripiprazole appears to be associated with only minimal amounts of extrapyramidal side effects. In the comparison of aripiprazole (15 or 30 mg), 10 mg of haloperidol, and placebo, only haloperidol resulted in a significantly greater increase in EPS as measured by the Simpson-Angus Scale compared with placebo. Patients treated with haloperidol were more likely to receive treatment of extrapyramidal side effects than were patients in the other groups. Thirty percent of patients receiving haloperidol, 12 percent on placebo, 8 percent on 15 mg of aripiprazole, and

15 percent on 30 mg of aripiprazole received benztropine. However, there are a number of reports suggesting that aripiprazole-treated patients may experience restlessness. It is unclear if this experience is related to stimulating effects of aripiprazole or to akathisia.

Cardiac. Aripiprazole is not associated with prolongation of the QT interval as assessed by ECG.

Therapeutic Indications

Schizophrenia. Aripiprazole is effective for the treatment of acute psychosis in schizophrenia. It is also effective for preventing relapse in stabilized patients with schizophrenia. This is supported by a 26-week trial that found higher relapse rates on placebo compared to aripiprazole. The 52-week comparison with haloperidol also provides evidence that patients tend to remain stable once they have responded to aripiprazole. Its effectiveness appears to be similar to that of haloperidol and risperidone. Aripiprazole is also indicated for adolescents aged 13 to 17 with schizophrenia.

Bipolar I Disorder. Aripiprazole is effective for treating acute manic and mixed episodes, as monotherapy or as adjunctive therapy with lithium or valproate. It is also indicated for children and adolescents (age 10 to 17 years).

Adjunctive Treatment of Depression. Aripiprazole is indicated as an adjunctive treatment of patients who have inadequate responses to antidepressants.

Irritability Associated with Autistic Disorder. Aripiprazole is indicated in the treatment of irritability associated with autistic disorder. Trials demonstrating efficacy were conducted in patients 6 to 17 years of age.

Tourette Disorder. Aripiprazole is indicated in the treatment of Tourette disorder. Studies evaluating efficacy were conducted in pediatric patients aged 6 to 18 years of age.

Agitation Associated with Schizophrenia or Bipolar Mania. Short-acting IM aripiprazole has been approved for acute agitation. A dose of 10 mg of aripiprazole was effective after 45 minutes versus 7.5 mg IM haloperidol at 100 minutes.

Precautions and Adverse Reactions

Black Box Warning

INCREASED MORTALITY IN ELDERLY PATIENTS WITH DEMENTIA-RELATED PSYCHOSIS. Aripiprazole and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis compared with placebo.

SUICIDAL THOUGHTS AND BEHAVIORS. As with other SGAs with indications for depression (and all other antidepressants), aripiprazole has a black box warning indicating antidepressants may be associated with an increased risk of suicidal thinking and behavior in children, adolescents, and young adults.

Pregnancy and Lactation. Aripiprazole is designated pregnancy category C. There have been limited reports documenting neonatal and maternal outcomes to aripiprazole exposure during pregnancy. In one report, mother and neonate had no pregnancy or labor-related complications, and the infant seemed to be developing normally at a 6-month follow-up. However, the mother was unable to lactate, raising concerns about potential hypoprolactinemia. Aripiprazole has been associated with teratogenicity in animal studies at three to ten times the maximum recommended dose for humans. Given the paucity of published data on the effect of aripiprazole on fetal and neonatal development, this medication should be prescribed during pregnancy only after a careful risk–benefit analysis. Currently there is no published information on the effects of aripiprazole use in nursing mothers.

Drug Interactions

Aripiprazole is metabolized by CYP3A4 and CYP2D6. Because carbamazepine induces CYP3A4, coadministration with this agent reduces aripiprazole concentrations. Agents that inhibit CYP3A4, such as ketoconazole, may reduce the metabolism of aripiprazole, leading to an increase in concentrations. Similarly, agents that inhibit CYP2D6, such as quinidine, may raise aripiprazole concentrations (Table 34.28–4).

Laboratory Interference

There is no evidence that aripiprazole interferes with laboratory tests.

Dosage and Administration

Aripiprazole can be administered orally as a conventional tablet, an orally disintegrating tablet, or as an oral solution. It is also available as a short-acting IM formulation in single dose vials containing 9.75 mg of aripiprazole and in a long-acting IM formulation (Maintena) in both single dose vials and prefilled syringes containing 300 mg or 400 mg of lyophilized aripiprazole.

Schizophrenia. Aripiprazole is administered at dosages between 10 and 30 mg daily. Two studies compared 10 mg of aripiprazole with placebo; in one study, aripiprazole was more effective than placebo, whereas, in the other, it was not. This suggests that 15 mg administered once a day is a reasonable dosage for initiating treatment. There is no evidence that higher dosages are associated with a greater likelihood of improvement. As a result, patients should probably be treated with 15 mg per day for a sufficient amount of time (e.g., 2 to 4 weeks) before increasing the dosage. Oral aripiprazole can be administered once daily without regard to meals.

Studies in adolescents found that aripiprazole was well tolerated when patients were started at 2 mg and raised to 5 mg in 2 days and to the target dose of 10 mg in 2 additional days.

Bipolar Disorder. Aripiprazole is effective at doses of 15 to 30 mg for patients with mania and mixed states. The recommended starting dose in adults is 15 mg daily if given as monotherapy or 10 to 15 mg daily if given as adjunctive therapy with lithium or valproate. In pediatric patients aged 10 to 17 years, it should be started at 2 mg/day, titrated to 5 mg/day after 2 days, and titrated to the target dose of 10 mg/day after 2 additional days. Further dose increases can be made by increments of 5 mg/day.

Adjunctive Treatment of Major Depressive Disorder. The recommended starting dose of aripiprazole in adults with inadequate responses to antidepressants is 2 to 5 mg daily. The effective dose appears to be between 10 and 15 mg.

Irritability Associated with Autistic Disorder. In patients aged 6 to 17 years, aripiprazole should be started at 2 mg/day, then increased to 5 mg/day, and subsequently increasing to 10 or 15 mg/day based on efficacy and tolerability.

Tourette Disorder. For pediatric patients (age 6 to 18 years) weighing less than 50 kg, aripiprazole can be started at 2 mg/day, increased to 5 mg/day after 2 days, and subsequently increased to 10 mg/day based on efficacy and tolerability. For patients weighing more than 50 kg, aripiprazole should be titrated to a target dose of 10 mg/day on day 8, and the dose can be increased up to 20 mg/day if needed for control of tics.

Agitation Associated with Schizophrenia or Bipolar Mania (IM Formulation). The recommended IM dose for treating agitation in schizophrenia or bipolar mania is 9.75 mg. There is no indication that higher doses are more effective. As with other IM antipsychotics, if repeated injections are required, patients should be treated with an oral formulation.

There is no indication that adjustments in dosage are required for age, gender, race, or smoking.

Long-Acting Injectable. In 2013, a long-acting form of aripiprazole was approved for IM injection, with indication for the treatment of schizophrenia. As with all long-acting injectables, it is recommended to first assess medication tolerability with the oral formulation. The recommended starting and maintenance dose is 400 mg IM monthly. If this dose is not tolerated, it can be reduced to 300 mg IM monthly. The median T_{max} of this formulation is 5 to 7 days and the terminal half-life of a 400 mg IM dose is 46.5 days. Thus, it is recommended that when starting this medication, oral antipsychotic should be continued for 14 days before discontinuation, given the gradual increase in plasma levels after initial injection.

Table 34.28–4.
Comparison of Usual Dosing[a] for the Currently Available Second-Generation Antipsychotics in Schizophrenia

Antipsychotic	Typical Starting Dosage	Maintenance Therapy Dose Range	Titration/Notes	Maximum Recommended Dosage
Clozapine (tablets) (Clozaril)	12.5 mg once or twice a day	150–450 mg per day in divided doses or as a single bedtime dose	The dosage should be increased to 25–50 mg on the second day. Further increases may be made in daily increments of 25–50 mg to a target dose of 300–450 mg per day. Subsequent dosage increases should be made no more than once or twice weekly in increments of no more than 100 mg.	900 mg per day
Risperidone (tablets, orally disintegrating tablets, and oral solution) (Risperdal)	1 mg twice a day or 2 mg at night	2–8 mg per day, in divided doses or as a single bedtime dose	Increase by increments of 1–2 mg per day. In some patients (elderly, hepatic or renal impairment), a slower titration may be appropriate.	16 mg per day
Risperidone IM long acting (Consta)	25 mg every 2 wks	25–50 mg every 2 wks	Continue oral antipsychotic for 3 wks. Titrate by 12.5 mg every 4 wks.	50 mg every 2 wks
Olanzapine (tablets and orally disintegrating tablets) (Zyprexa)	5–10 mg per day	10–20 mg per day	Dosage increments of 5 mg, at change intervals of not less than 1 wk.	20 mg per day

Olanzapine IM long acting (Relprevv)	210 mg/2 wks or 405 mg/4 wks (for target10 mg/day); 300 mg/2 wks (for target 15–20 mg/day)	150 mg/2 wks or 300 mg/4 wks (for target 10 mg/day); 210 mg/2 wks or 405 mg/4 wks (for target 15 mg/day); 300 mg/2 wks (for target 20 mg/day) Requires a 3 hour observation period after injections.	Adjust dosage at intervals of 2 or 4 wks, depending on dosing schedule and maintenance therapy target dose.	300 mg IM every 2 wks
Quetiapine (tablets) (Seroquel)	25 mg twice a day	400–800 mg per day	Increase by 25–50 mg in divided doses for first several days, to target of 300–400 mg per day by day 4. Adjustments can be made by 25–50 mg twice daily, every 2 days	800 mg per day
Quetiapine extended release (Seroquel XR)	300 mg nightly	400–800 mg per day	Can increase by up to 300 mg per day at intervals as short as 1 day	800 mg per day
Ziprasidone (capsules) (Geodon)	20 mg twice a day with food	20–80 mg twice a day with food	Adjust at intervals not less than 2 days	80 mg twice a day
Ziprasidone (intramuscular)	For acute agitation: 10–20 mg, as required, up to a maximum of 40 mg per day	N/A	For acute agitation: Dosages of 10 mg may be administered every 2 hrs and dosages of 20 mg may be administered every 4 hrs up to a maximum of 40 mg per day.	For acute agitation: 40 mg per day, for not more than 3 consecutive days
Aripiprazole (tablets) (Abilify)	10–15 mg once per day	10–30 mg per day	Dosage increases should not be made at intervals less than every 2 wks	30 mg per day
Aripiprazole IM long acting (Maintena)	400 mg IM monthly	300–400 mg IM monthly	Continue oral antipsychotic for 14 days before discontinuing.	400 mg IM monthly
Paliperidone (tablets) (Invega)	6 mg once per day	3–12 mg per day	Efficacy beyond 6 mg per day has not been shown. Dose increases of 3 mg should occur at intervals no less than 5 days.	12 mg per day
Paliperidone palmitate (Invega Sustenna)	234 mg IM on day 1, then 156 mg IM 1	39–234 mg IM monthly (117 mg recommended)	Make adjustments no more frequently than monthly	234 mg IM monthly

	wk later			
Iloperidone (Fanapt)	1 mg twice daily	6–12 mg twice daily	Can make dose increases no more frequently than by 2 mg twice daily every day	12 mg twice daily
Asenapine (Saphris)	5 mg twice daily	5–10 mg twice daily	Can increase to 10 mg twice daily after 1 wk	10 mg twice daily
Lurasidone (Latuda)	40 mg daily with food	40–160 mg daily with food	Not specified; tablets are available in 20, 40, 60, 80, and 120 mg strengths	160 mg daily
Cariprazine (Vraylar)	1.5 mg once per day	1.5–6 mg once daily	Start with 1.5 mg for 1 day with increases in 1.5 mg increments; due to the long half-life of cariprazine and its active metabolites, changes in dose will not be fully reflected in plasma for several weeks	6 mg
Brexpiprazole (Rexulti)	1 mg once per day	2–4 mg daily for schizophrenia; 2–3 mg daily for adjunctive treatment of depression	For schizophrenia start 1 mg qd for 4 days, then 2 mg for 3 days, then 4 mg qd. For adjunctive treatment of depression, start 0.5–1 mg qd, then increase dose to 2 mg qd after 1 wk.	4 mg for schizophrenia; 3 mg for adjunctive treatment of depression

Information taken from U.S. Prescribing Information for individual agents.

[a]Dosage adjustments may be required in special populations.

The clinical efficacy of aripiprazole long-acting injectable was established in two clinical trials. The first trial (N = 339) was a 12-week randomized controlled trial in adults with schizophrenia who were having an acute relapse. The primary outcome measure used was the change in PANSS total score from baseline to week 10. Participants were randomized to receive aripiprazole long-acting injectable 400 mg or placebo on day 0, 28, and 56, with the dose of aripiprazole being allowed to be adjusted to 300 mg if clinically indicated. Aripiprazole showed statistically significant improvement on the primary outcome measure as compared to placebo. A second trial assessed longer-term efficacy. In this study, adults with schizophrenia were converted and stabilized on oral aripiprazole (10-30 mg daily). Next, participants were stabilized on aripiprazole long-acting injection (400 mg or 300 mg monthly) for at least 12 weeks, discontinuing oral aripiprazole after the first 2 weeks. In the final double-blind placebo-controlled randomized withdrawal phase, patients (N = 403) were randomized to either continue aripiprazole long-acting injection or switch to placebo injection. The primary outcome measure was the time to relapse, defined as worsening on prespecified CGI scales, prespecified PANSS items, hospitalization, or violent behavior. Participants randomized to continue aripiprazole long-acting injectable had a longer time to relapse than those randomized to placebo.

PALIPERIDONE

Paliperidone is the major active metabolite of risperidone and shares risperidone's receptor profile. Paliperidone extended release tablets use a pressure-based osmotic release oral system (OROS) delivery system to provide continual drug delivery for over 24 hours (see below). Because paliperidone extended release is gradually released into the plasma without peaks in concentration, it permits treatment initiation without titration. As the primary metabolite of risperidone, paliperidone does not undergo further significant hepatic metabolism. Therefore, it suffers from fewer potential drug–drug interactions.

OROS Delivery

The OROS delivery system is capsule shaped and consists of an osmotically active trilayer core surrounded by a subcoat and semipermeable membrane. The trilayer core has two drug layers and a push layer containing osmotically active components. In the gastrointestinal tract, the color overcoat of the capsule dissolves quickly and water enters the semipermeable membrane. The semipermeable membrane controls the rate at which water enters the tablet core and so determines the rate of drug delivery. The hydrophilic polymers in the core hydrate and swell and create a gel that contains paliperidone, which is pushed out of two laser-drilled orifices in the tablet shell. The osmotic delivery system reduces or eliminates plasma concentration fluctuations, allowing for once-daily dosing without fluctuating plasma levels.

Chemistry

Paliperidone belongs to the class of benzisoxazole derivatives (Fig. 34.28-1). The chemical name of paliperidone is (±)-3-[2-[4-(6-fluoro-1,2-benzisoxazol-3-yl)-1-piperidinyl]ethyl]-6,7,8,9-tetrahydro-9-hydroxy-2-methyl-4H-pyrido[1,2-a]pyrimidin-4-one. Its molecular formula is $C_{23}H_{27}FN_4O_3$ and its molecular weight is 426.49. Paliperidone contains a racemic mixture of (+) and (–) paliperidone.

Pharmacological Action

Pharmacokinetics

ABSORPTION. Peak plasma concentration of paliperidone extended release is reached approximately 24 hours after a single dose. The half-life of paliperidone is about 23 hours, which allows for once a day dosing. Absorption is increased by approximately 60 percent when paliperidone extended release is taken with food.

DISTRIBUTION. Steady-state concentrations are reached after 4 to 5 days of daily dosing. Paliperidone has a volume of distribution of 487 L and is 74 percent plasma protein bound.

METABOLISM AND ELIMINATION. In vitro studies demonstrate that paliperidone is a substrate at CYP2D6 and CYP3A4, but in vivo studies do not support a clinically significant role for the P450 system in the metabolism of paliperidone. For instance, there was no difference in the exposure or clearance rates of subjects who are rapid or poor metabolizers of CYP2D6 substrates. Fifty-nine percent of paliperidone is excreted unchanged in the urine, and clearance is decreased in patients with moderate-to-severe renal impairment. The tablet shell is excreted whole in the stool and patients should be advised not to worry if they see what looks like an undissolved tablet in their stool.

Pharmacodynamics

MECHANISM OF ACTION. As mentioned above, paliperidone is the major active metabolite of risperidone. Both drugs, therefore, share the same receptor affinities and putative mechanism of action. Paliperidone is an antagonist at dopamine D_2 and serotonin 5-HT_{2A} receptors. In a small sample pharmacokinetics study, PET imagery demonstrated a median D_2 receptor occupancy of 64 percent 22 hours after a single dose of 6 mg paliperidone extended release. The 5-HT_{2A} receptor occupancy was 65 percent at 4.5 hours after a dose. Paliperidone also has antagonist activity at α_1, α_2, and H_1 receptors.

Design and Interpretation of Clinical Drug Studies

Acute Treatment Trials. Three 6-week randomized double-blind, placebo-controlled efficacy and safety studies were conducted using fixed doses of paliperidone extended release. Kane and colleagues administered placebo, 6, 9, or 12 mg of paliperidone extended release to 628 patients with schizophrenia. Davidson and colleagues treated 618 subjects with schizophrenia with placebo or 3, 9, or 15 mg of paliperidone extended release. Marder and colleagues evaluated 444 patients administered placebo, 6, or 12 mg of paliperidone extended release. In each of the three studies a 10-mg olanzapine group was also included for trial validity assessment only. In all three studies, the PANSS total score decreased significantly in all paliperidone arms compared with placebo. In the study by Kane et al., there was a trend toward greater improvement with the 12-mg paliperidone group. Davidson's group reported that as early as day 4 the three paliperidone arms had separated significantly from placebo on PANSS total score. The study by Marder and colleagues found significant improvement by day 4 for the 6-mg arm and by day 15 for the 12-mg arm.

In all three studies, the three paliperidone arms showed significant or trend improvement from baseline compared with placebo on positive, negative, disorganized, hostility or agitation, and anxiety or depression factor scores. The paliperidone groups also improved compared with placebo on measures of personal and social functioning, although whether this change was more reflective of improving psychosis is unclear. Paliperidone was associated with some metabolic and movement-related adverse events, especially in the 9-, 12-, and 15-mg groups.

Maintenance Treatment Trials. Kramer and colleagues treated 205 patients in a study with an 8-week open-label run-in phase, followed by a 6-week open-label stabilization period before entering a variable duration double-blind treatment phase where patients were randomized to either placebo or paliperidone 3 to 15 mg daily. The primary efficacy variable was time to first recurrence of schizophrenia symptoms (e.g., hospitalization, changes in symptom scores, aggressive, self-injurious, or suicidal behavior). Time to recurrence was significantly longer in the paliperidone group; at the final analysis 52 percent of patients on placebo had a recurrence versus 22 percent on paliperidone. Twenty-five percent of patients on placebo had a recurrence event by day 23, whereas 25 percent of paliperidone-treated patients had a recurrence event by day 83.

Effect on Sleep Architecture. In a randomized double-blind, placebo-controlled trial, Luthringer and his colleagues evaluated the effect of paliperidone extended release on sleep architecture in 36 patients with schizophrenia and insomnia. Using a significance level of 0.1 percent, they found that 9 mg of paliperidone extended release significantly increased overall sleep duration, and stage 2 and rapid eye movement (REM) sleep duration over treatment with placebo. Paliperidone treatment was associated with an 11 percent increase in sleep efficiency index. In this study, paliperidone was not associated with daytime somnolence.

Effects on Specific Organs and Systems

General. None of the randomized control studies reported statistically significant changes in glucose, insulin, or lipid levels in the paliperidone arms versus placebo. However, there were isolated cases of hyperglycemia reported. Across the studies, approximately 7 to 12 percent more patients in the higher dose paliperidone arms versus the placebo arms experienced a weight gain over 7 percent of their body weight during the course of the medication trial. This suggests that paliperidone has a weight gain profile similar to that of risperidone.

Cardiovascular System. Tachycardia and palpitations are relatively frequent side effects of paliperidone, as they are with risperidone.

Tachycardia does not appear to be dose related. Less frequently reported is a risk of bradycardia. Paliperidone has a modest effect on the QT interval. Vascular ischemia and venous thrombosis occurred rarely.

Respiratory System. Dyspnea not associated with further adverse outcome occurred with some frequency during the phase II and III clinical trials of paliperidone. Pulmonary emboli occurred in less than 0.001 percent of patients on paliperidone.

Gastrointestinal System. Abdominal pain and salivary hypersecretion were noted to occur with paliperidone administration during early trials. Salivary hypersecretion appeared to be dose related, with most cases occurring at the 12-mg dose.

Hematological and Lymphatic System. Thrombocytopenia was a rare occurrence, occurring in less than 1 of 1,000 patients. Edema was noted in less than 1 percent of patients.

Central Nervous System. Headache and insomnia were frequent side effects of paliperidone treatment in the randomized clinical trials. Confusion, abnormal gait, and impaired coordination were noted infrequently in patients taking therapeutic doses of paliperidone.

Immunologic System. Anaphylactic reactions occurred in less than 1 of 1,000 people who took paliperidone.

Endocrine. Paliperidone increases prolactin levels. In the randomized clinical trials, prolactin levels were highest in those on the greatest doses of paliperidone extended release. In the study by Kane and colleagues, mean plasma prolactin levels rose 27 ng/mL in men and 86 ng/mL in women. In Davidson's group, male patients on 15 mg of paliperidone experienced mean plasma prolactin level increases of 35.3 ng/mL, and female patients had mean increases of 106.9 ng/mL. In all studies, 4 percent or less of patients on paliperidone reported prolactin-related side effects, such as galactorrhea, gynecomastia, amenorrhea, anorgasmia, and abnormal sexual function. However, spontaneous reports of side effects may underestimate the number of patients experiencing these effects.

Therapeutic Indications

Schizophrenia. Paliperidone extended release is indicated for the acute and maintenance treatment of schizophrenia. Efficacy was also established in adolescent subjects aged 12 to 17.

Schizoaffective Disorder. Paliperidone extended release is also

indicated for the treatment of schizoaffective disorder, both as monotherapy or as adjunct therapy to mood stabilizers and/or antidepressants.

Precautions and Adverse Reactions

Black Box Warning. Paliperidone and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis compared with placebo.

QT Prolongation. Paliperidone immediate release has been shown to have a modest effect on the QT interval; with a mean increase of 12.3 ms at the 8-mg dose (which has a mean steady-state peak plasma concentration over twice the exposure of a 12-mg dose of paliperidone extended release). In the QT study and the three fixed-dose studies, only one patient had a change exceeding 60 ms at one time point. No patient on paliperidone had a QTc exceeding 500 ms at any time in the studies. Paliperidone use should be avoided or initiated with care in patients on class IA or class III antiarrhythmic medications, on other medications that increase the QT interval, or with other cardiac arrhythmias including congenital QT prolongation.

Neuroleptic Malignant Syndrome. One patient in the five clinical trials investigating the safety and efficacy of paliperidone extended release had symptoms of NMS after 19 days on paliperidone extended release treatment.

Tardive Dyskinesia. All antipsychotic medications (with the possible exception of clozapine and quetiapine) carry a risk of tardive dyskinesia. Risk of tardive dyskinesia is difficult to assess in the course of 6-week trials. However, acute dyskinesias occurred in 7.7 percent and 8.7 percent of patients in the 9- and 12-mg paliperidone extended-release groups versus 3.4 percent in the placebo group.

Hypercholesterolemia, Hyperglycemia, and Diabetes Mellitus. In the early clinical trials, paliperidone extended release was not associated with statistically significant changes in cholesterol, glucose, or insulin levels compared to placebo. Given that paliperidone is the metabolite of risperidone, it is reasonable to predict the two medications will have similar metabolic profiles. However, many of the SGA medications have been associated with increases in fasting serum glucose levels and rates of diabetes mellitus type 2. Some SGA medications have also been associated with elevations in total cholesterol, LDL, and triglycerides. Regular monitoring of fasting cholesterol and glucose levels is appropriate in all patients on SGA medications.

Gastrointestinal. The paliperidone extended release tablet does not deform or change shape in the gastrointestinal tract. Therefore, there is a potential risk of obstruction in patients with pre-existing and severe gastrointestinal tract narrowing. Patients with such conditions should not be administered paliperidone extended release.

Cerebrovascular. Trials of other SGA medications, such as risperidone, aripiprazole, and olanzapine, have demonstrated higher incidences of strokes and transient ischemic attacks in elderly patients with dementia compared to patients on placebo. No data exist on paliperidone in this population, but SGAs are not indicated in elderly patients with dementia.

Movement Side Effects and Disorders. In all studies, paliperidone extended release was noted to be associated with higher rates of movement side effects in dosages above 9 mg. Reported symptoms include hypertonia, dystonia, tremor, hyperkinesia, and oculogyric crisis. Serious movement side effects were rare, although one study reported a possible case of NMS after 19 days of paliperidone extended release treatment. Generally, across the studies, patients on paliperidone doses above 6 mg required more doses of anticholinergic medication than did patients on placebo.

Orthostatic Hypertension and Syncope. Like risperidone, paliperidone has antagonistic activity at the α receptors, contributing to its risk of orthostatic hypotension and syncope. Paliperidone extended release should be used with caution in people with known cardiovascular disease or individuals with predispositions to hypotension.

Hyperprolactinemia. Paliperidone increases prolactin levels through D_2 antagonism (see above section on metabolic effects). Hyperprolactinemia suppresses GnRH, which reduces gonadotropin release and can impair gonadal steroidogenesis in men and women. The result can be inhibited reproductive function and symptoms such as galactorrhea, amenorrhea, gynecomastia, impotence, and anorgasmia. Long-term hyperprolactinemia may lead to decreased bone density in both sexes when associated with hypogonadism. Approximately one-third of breast cancers are prolactin dependent, and patients with a personal or family history of breast cancer should be monitored closely, especially if they develop hyperprolactinemia while on paliperidone extended release.

Pregnancy. Paliperidone is rated pregnancy category C. In rats and rabbits given paliperidone up to eight times the maximum recommended human dose (MRHD), there were no increases in fetal abnormalities. However, reproduction studies with risperidone have demonstrated increases in pup deaths at oral doses less than the maximum recommended

doses in humans. Use of FGAs during the third trimester has been associated with self-limited EPS in the neonate. Whether paliperidone will produce similar symptoms in the neonate is still unknown. Given the limited information available on the effects of paliperidone on pregnant women and neonates, it should be used only after careful consideration of the risks and benefits.

Lactation. Paliperidone is excreted in breast milk. Studies of the distribution of risperidone and its metabolite, paliperidone, reveal that the milk/plasma concentration ratio for both agents was less than .5. Relative infant doses of risperidone ranged from 2.3 to 4.7 percent of the maternal dose. In infants who were breastfed, risperidone and paliperidone were not detected in the plasma of the infants. However, the possibility of infant exposure and possible EPS or NMS risk cannot be ruled out. The decision to breastfeed while on paliperidone should be made only after careful risk–benefit analysis.

Fertility. In female rats on doses up to 2.5 mg/kg per day the percentage of pregnancies was not affected, but pre- and postimplantation loss was increased and the number of live embryos was decreased. In male rats given doses of paliperidone up to 2.5 mg/kg per day fertility was not affected, but sperm analysis was not conducted. Beagles given risperidone showed decreases in testosterone levels and sperm motility, which continued at an attenuated level at least 2 months after drug discontinuation.

Carcinogenesis. Carcinogenesis studies of paliperidone have not yet been performed. Risperidone was associated with increases in pituitary gland adenomas, endocrine pancreas adenomas, and mammary gland adrenocarcinomas in mice and rats. The no-effect dose for the tumors was less than or equal to the maximum human recommended dose. The increase in the above tumors is considered to be associated with prolonged D_2 antagonism and hyperprolactinemia. The relevance to human populations of these findings in rodents is unknown.

Mutagenesis. There is no evidence of mutagenesis with paliperidone.

Geriatric Use. Paliperidone use in geriatric populations was assessed in the 6-week placebo-controlled studies that included 125 elderly subjects over age 65 with schizophrenia. No overall differences in safety and effectiveness were observed between elderly and younger subjects when renal function was not impaired. SGA medications are associated with greater risk of death compared with placebo in elderly patients with dementia-related psychosis.

Overdose. There are limited data on overdose with paliperidone. The highest estimated ingestion was 405 mg, which resulted in signs of gait unsteadiness and EPS. As paliperidone is the active metabolite of risperidone, overdosage effects are likely to be similar.

Drug Interactions

Although in vitro studies demonstrate that paliperidone is a substrate of both CYP3A4 and CYP2D6, in vivo studies do not demonstrate clinically significant metabolism by 2D6 or 3A4 isozymes. Therefore, paliperidone is unlikely to be affected by drugs metabolized by the P450 system. Similarly, paliperidone is not an inhibitor or inducer of any of the P450 isozymes and is not expected to inhibit or induce drug clearance of any P450 substrates.

As with risperidone, paliperidone's effects on α_1 and α_2 receptors can lead to orthostatic hypotension. Other drugs with this side effect should be used with caution given potential additive effects.

Dosage and Administration

Paliperidone extended release comes in 3-, 6-, and 9-mg tablets. The recommended dosage range for paliperidone extended release is 3 to 12 mg, although some clinical trials included a 15-mg paliperidone group. The recommended starting dose is 6 mg and no titration is needed. The existing clinical trials found no significant difference in antipsychotic response between the 6-mg dose and higher dosages; however, there was a trend toward better response at higher dosages. Beyond the 6-mg starting dose, upward titration by 3 mg every 5 days is recommended. In adolescents aged 12 to 17, paliperidone should be started at 3 mg daily. If dose increases are necessary, they should occur by no more than 3 mg/day every 5 days.

Dosing in Special Populations. No dosage adjustments are needed for those with hepatic impairment. People with mild renal impairment should not be prescribed more than 6 mg of paliperidone a day. People with moderate-to-severe renal impairment should not take more than 3 mg per day. No adjustments are necessary for elderly populations beyond the above recommendations regarding renal status. Paliperidone has not yet been evaluated in pediatric populations.

Long-Acting Injectable. Paliperidone also comes in a long-acting injectable formulation (Sustenna), with kits containing needles prefilled syringes containing 39-, 78-, 117-, 156-, or 234-mg paliperidone palmitate.

In 2009, the FDA approved this medication as a long-acting injectable formulation of paliperidone. Its indications are for the treatment of schizophrenia and for the treatment of schizoaffective disorder, both as monotherapy and as an adjunct to mood stabilizers or antidepressants.

This medication can be administered by gluteal or deltoid IM injection.

As with all long-acting injectable medications, it is recommended that tolerability is established by taking oral paliperidone (or risperidone due to its metabolism into paliperidone) prior to switching to the injectable formulation. The oral antipsychotic can then be discontinued immediately after the first injection. The dosage recommendations consist of an initiation phase (injections on day 1 and day 8), followed by a monthly maintenance phase. The recommended dosage is 234 mg IM on day 1, followed by 156 mg IM on day 8, followed by 39 to 234 mg (for schizophrenia) or 78 to 234 mg (for schizoaffective disorder) 4 weeks later, then monthly. The recommended initial maintenance dose for the treatment of schizophrenia is 117 mg IM monthly and this can be adjusted within specified ranges according to tolerability and effectiveness. In patients with mild renal impairment (creatinine clearance 50 to 79 mL/min), the day 1 injection should be 156 mg, followed by 117 mg on day 8, and then using 78 mg monthly maintenance doses. Paliperidone palmitate should not be used in patients with creatinine clearance <50 mL/min.

Efficacy was established for the acute treatment of schizophrenia in one 9-week and three 13-week trials of relapsed adult inpatients. The primary outcome measures were the change in PANSS total score from baseline to end point. Dosages tested in the trials included 39, 78, 156, 234 mg, and all dosages were found to be statistically superior to placebo in one or more trials. Efficacy for maintenance monotherapy treatment of schizophrenia was established in a double-blind placebo-controlled randomized withdrawal study. In this study, adults with schizophrenia who were stabilized on paliperidone palmitate (N = 410) were randomized to continue paliperidone long-acting injection or switch to placebo injections. They were then monitored until they experienced a schizophrenia exacerbation, which was defined as either psychiatric hospitalization or prespecified increase in PANSS score. The primary outcome measure was time to relapse. In this study, paliperidone palmitate was statistically superior to placebo in increasing time to relapse.

Efficacy for the maintenance treatment of schizoaffective disorder (either as monotherapy or as adjunct to mood stabilizer or antidepressant) was established in one long-term randomized withdrawal study. In this study, adults with DSM-IV schizoaffective disorder (bipolar or depressive types) were first stabilized on open-label flexible-dose paliperidone palmitate in a 13-week lead-in. Participants could be taking concomitant mood stabilizers or antidepressants. Subjects (N = 334) who met stabilization criteria for at least 12 weeks (by prespecified PANSS, YMRS, and HAM-D scores) were randomized (double-blind) to continue paliperidone palmitate or switch to placebo injections, then monitored for 15 months. The primary outcome measure was time to relapse, as defined

by either psychiatric hospitalization, intervention used to avert hospitalization, self-injury, suicidal or homicidal ideation, violent behavior, prespecified increase on certain PANSS items, or increase in CGI-S-SCA overall score. Paliperidone palmitate was statistically superior to placebo in increasing time to relapse.

ILOPERIDONE

Iloperidone is an antipsychotic with high affinity for 5-HT_{2A} and D_2 receptors. Important clinical properties include a low liability to cause EPS, but a risk of orthostatic hypotension requiring a titration period of several days.

History

Initial research and development of iloperidone occurred in the early to middle 1990s. FDA approval for the treatment of schizophrenia was finally granted in 2009, after an initial new drug application submitted in 2007 was denied due to need for additional clinical trial support.

Chemistry

Iloperidone belongs to the class of piperidinyl-benzisoxazole derivatives (Fig. 34.28–1). The chemical name of iloperidone is 4′-[3-[4-(6-Fluoro-1,2-benzisoxazol-3-yl)piperidino]propoxy]-3′-methoxyacetophenone. Its molecular formula is $C_{24}H_{27}FN_2O_4$ and its molecular weight is 426.48.

Pharmacological Action

Pharmacokinetics. According to the product label, the mean elimination half-lives for iloperidone and its two major metabolites (P88 and P95) are 18, 26, and 23 hours, respectively, in CYP2D6 extensive metabolizers. In poor metabolizers, the half-lives are 33, 37, and 31 hours. Steady-state concentrations in plasma are reached after 3 to 4 days of dosing.

ABSORPTION. Iloperidone is well-absorbed with oral administration, with peak plasma levels occurring 2 to 4 hours after ingestion. Administration with a high-fat meal delayed T_{max} for iloperidone by 1 hour and its major metabolites, P88 and P95, by 2 and 6 hours, respectively. As a high-fat meal does not significantly affect the maximum concentration or area under the curve of iloperidone or its major metabolites, product labeling indicates it can be administered without regard to meals.

DISTRIBUTION. Iloperidone has an apparent volume of distribution of 1,340 to 2,800 L and it, as well as its major metabolites, is approximately 95 percent serum protein-bound.

METABOLISM AND ELIMINATION. Iloperidone is metabolized by three pathways: CYP2D6-mediated hydroxylation, CYP3A4-mediated O-demethylation, and carbonyl reduction. There are two major metabolites of iloperidone, P95 and P88, with P95 exposure being greater in CYP2D6 extensive metabolizers and P88 being greater in CYP2D6 poor metabolizers. Elimination of metabolites is primarily through the urine, with feces accounting for a lesser proportion.

Pharmacodynamics

MECHANISM OF ACTION. Iloperidone functions as an antagonist with high affinity at 5-HT_{2A}, D_2, D_3, α_1, and α_{2c} adrenergic receptors; intermediate affinity at D_4, 5-HT_6, and 5-HT_7 receptors; and low affinity for 5-HT_{1A}, D_1, and H_1 receptors. Iloperidone possesses negligible affinity for muscarinic receptors. Its metabolite P88 has receptor affinities "generally equal or less than that of the parent compound" according to product labeling. Its metabolite P95 possesses high affinity for 5-HT_{2A}, α_{1A}, α_{1B}, α_{1D}, and α_{2C} receptors only.

Design and Interpretation of Clinical Drug Studies

Acute Treatment Trials. The efficacy of iloperidone was supported by two large trials—one 4 weeks and one 6 weeks in duration. In the 4-week trial, 604 participants with DSM-IV schizophrenia were randomized to receive iloperidone 24 mg daily (titrated over course of 1 week, with twice-daily dosing), ziprasidone (160 mg daily), or placebo. The primary end point was the change in PANSS total score from baseline to end of treatment (day 28). In this study, both iloperidone and ziprasidone were statistically superior to placebo in improving PANSS scores. Iloperidone was noted to have a different adverse event profile than ziprasidone, with iloperidone having less sedation, somnolence, EPS, akathisia, agitation, and restlessness and causing more weight gain, tachycardia, orthostatic hypotension, dizziness, and nasal congestion. In the 6-week trial, 706 participants with DSM-IV schizophrenia were randomized to receive either one of 2 flexible dose ranges of iloperidone (12 to 16 mg/day or 20 to 24 mg/day), risperidone (6 to 8 mg/day), or placebo. The primary end point was the change in BPRS total score from baseline to end of treatment (day 42). Both dose ranges were superior to placebo in improving BPRS scores. While the efficacy of risperidone appeared greater than iloperidone within the first 2 weeks of treatment, the efficacy between treatments was similar in patients who remained on treatment of at least 2 weeks. This differential efficacy in the first 2 weeks was attributed to the slower dose titration schedule necessary for iloperidone.

Two additional large short-term studies were performed to assess the

efficacy of iloperidone for the treatment of schizophrenia or schizoaffective disorder. In the first study (6 weeks), iloperidone 12 mg/day (but not 4 mg or 8 mg/day) statistically improved the PANSS and BPRS total scores as compared to placebo. In the second 6-week study, both iloperidone 4 to 8 and 10 to 16 mg/day outperformed placebo for improving the BPRS total score. For the PANSS, the 4 to 8 mg/day group outperformed placebo only on the positive symptom subscale, while the 10 to 16 mg/day group outperformed placebo on the PANSS total as well as positive and negative symptom subscales.

Maintenance Treatment Trials. While iloperidone is currently only indicated "for the treatment of adults with schizophrenia," based on two large short-term studies, limited long-term maintenance studies have been performed. Kane et al. reported the data from three trials. Subjects from three short-term studies (N = 1,644), who had experienced at least a 20 percent reduction in PANSS total score in the initial 6-week study, were randomized to receive iloperidone 4 to 16 mg/day or haloperidol 5 to 20 mg/day for 46 weeks. The primary outcome measure was the time to relapse, defined as a ≥25 percent increase in PANSS total score, discontinuation due to lack of efficacy, aggravated psychosis with hospitalization, or 2-point increase on the CGI-C after week 6. Iloperidone was found to be statistically equivalent to haloperidol for relapse prevention.

A second long-term extension study was published by Cutler et al., in which participants with a CGI-C score of at least 4 (indicating improvement) in the acute efficacy trial could continue in a 25-week open-label extension study. In this study, participants were all treated with iloperidone 12 mg twice daily, with the dose permitted to be decreased to 12 mg nightly by investigator discretion. There was no significant worsening in PANSS total score at 25 weeks (using last observation carried forward analyses), though only 72 out of 173 patients completed the entire 25-week study.

Therapeutic Indications

Schizophrenia. Iloperidone is indicated for the treatment of schizophrenia, with efficacy established based on two short-term studies.

Precautions and Adverse Reactions

Black Box Warning. Iloperidone and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis compared with placebo.

QT Prolongation. Iloperidone, at the dose of 12 mg twice daily, was

found to increase the QTc interval by 9 ms in an open-label study of 160 individuals. This QTc prolongation was increased to 19 ms in the presence of CYP2D6 inhibitor (paroxetine) or CYP3A4 inhibitor (ketoconazole). It is recommended that iloperidone be avoided in combination with other medications known to increase the QTc interval, in individuals with congenital long QT syndrome, history of cardiac arrhythmias, or histories of significant cardiovascular illness. It is also recommended that individuals at risk of electrolyte disturbances have potassium and magnesium levels monitored periodically, as low levels can increase the risk of QTc prolongation and arrhythmia. Iloperidone should be discontinued if QTc is found to be greater than 500 ms.

Neuroleptic Malignant Syndrome. NMS has been associated with the use of antipsychotic medications, including iloperidone. If NMS is suspected, iloperidone and other antipsychotic medications should be discontinued immediately and the patient should undergo intensive supportive care and medical monitoring.

Extrapyramidal Symptoms. Iloperidone is associated with a low risk of EPS, including akathisia, bradykinesia, dyskinesia, dystonia, parkinsonism, and tremor. When considering all EPS events in four short-term trials, 15.1 percent of subjects taking iloperidone 20 to 24 mg/day (N = 391), 13.5 percent of subjects taking 10 to 16 mg/day (N = 483), and 11.6 percent of subjects taking placebo (N = 587) reported any type of EPS. Parkinsonism was reported in 0.3 percent of subjects taking iloperidone 20 to 24 mg/day as compared to 0 percent in subjects taking placebo.

Tardive Dyskinesia. Tardive dyskinesia can develop in all patients treated with antipsychotic drugs. Its prevalence is highest among the elderly. The propensity of iloperidone to cause tardive dyskinesia is not known, but expected to be low due to the minimal extrapyramidal side effects of this medication. In general, SGAs have a lower tardive dyskinesia risk than FGAs.

Metabolic Changes. Iloperidone has been associated with hyperglycemia, dyslipidemia, and weight gain, and patients should be monitored accordingly. In a pooled analysis of longer term trials, iloperidone 10 to 16 mg/day was associated with a 1.8 mg/dL glucose increase at 3 to 6 months (N = 773) and a 5.4 mg/dL glucose increase at >12 months (N = 425). Iloperidone 10 to 16 mg/day was associated with a –3.9 mg/dL change in total cholesterol and –8.9 mg/dL change in triglycerides at 3 to 6 months (N = 783) and –7.7 mg/dL change in total cholesterol and –17.7 mg/dL change in triglycerides at >12 months (N = 428). Of note, the observed decrease in lipid parameters could be a

consequence of the participants switching to iloperidone from another antipsychotic with higher propensity for dyslipidemia. Iloperidone has also been associated with weight gain in a dose-dependent manner, with 18 percent of subjects taking iloperidone 20 to 24 mg/day and 12 percent of those taking 10 to 16 mg/day gaining ≥7 percent of bodyweight at 4 to 6 weeks.

Seizures. In 4 to 6 weeks placebo controlled trials, iloperidone was associated with seizure in 1 of 1,344 patients treated with iloperidone (0.1 percent) and 2 of 587 patients treated with placebo (0.3 percent).

Orthostatic Hypotension and Syncope. Iloperidone has been associated with orthostatic hypotension with dizziness, tachycardia, and syncope, and this is thought to be related to its property as an α1-antagonist. In short-term studies where the dose was increased slowly, as per label recommendations, orthostatic hypotension occurred in 5 percent of patients taking iloperidone 20 to 24 mg/day, 3 percent taking iloperidone 10 to 16 mg/day, and 1 percent of patients taking placebo. It is noted that more rapid increases in iloperidone dose would be expected to increase the risk of orthostatic hypotension and syncope.

Hyperprolactinemia. In a 4-week placebo-controlled trial of iloperidone, prolactin increased 2.6 ng/mL in the group taking iloperidone 24 mg/day, as compared to a decrease of 6.3 ng/mL in the placebo group. In longer-term trials (N = 3,210), 0.1 percent of males taking iloperidone developed gynecomastia as compared to 0 percent in the placebo group, and 0.2 percent of females taking iloperidone developed galactorrhea as compared to 0.5 percent in the placebo group.

Priapism. Three cases of priapism were observed in the drug development process, according to manufacturer's prescribing information. This is thought to be related to the α-antagonist properties of iloperidone.

Potential for Cognitive and Motor Impairment. Iloperidone can cause somnolence and sedation, which could impair cognitive and motor skills. In short-term trials, iloperidone at dosages 10 mg/day or greater was associated with somnolence in 11.9 percent of patients as compared to 5.3 percent in the placebo group.

Pregnancy. Iloperidone is currently designated "Category C" for use in pregnant women, meaning that animal studies have shown an adverse effect on the fetus, there are no adequate and well-controlled studies in humans, and potential benefits may warrant use despite potential risks. In pregnant rat studies, a dose of iloperidone 26 times the MRHD (in

mg/kg/day) caused increased intrauterine deaths, decreased fetal weight and length, decreased fetal skeletal ossification, and increased incidence of minor fetal skeletal anomalies. In pregnant rabbits, a dose of iloperidone 20 times the MRHD increased early intrauterine deaths and decreased fetal viability at term. In pregnant rats, doses of 1.6 to 4.8 times the MRHD caused no adverse effects to the pups except for an increase in stillbirth rates at the dose of 1.6 times MRHD. The iloperidone metabolite P95 as given to pregnant rats at 20, 80, and 200 mg/kg/day during the period of organogenesis, and no teratogenic effects were seen; however, delayed skeletal ossification occurred.

Nursing Mothers. While iloperidone has been found to be excreted in milk of lactating rats, it is not known if this occurs in humans. It is currently recommended that women taking iloperidone should not breast-feed.

Dose-Related Adverse Reactions in Clinical Trials. Based on the pooled data from four short-term trials, adverse reactions that occurred in >2 percent of iloperidone patients and which at least doubled in the 20 to 24 mg/day group as compared to the 10 to 16 mg/day group included: abdominal discomfort, dizziness, hypotension, musculoskeletal stiffness, tachycardia, and weight gain.

Common and Drug-Related Adverse Reactions in Clinical Trials. Based on the pooled data from four short-term trials, adverse reactions that occurred in ≥5 percent of iloperidone patients at one dose and at least twice the placebo rate included: dizziness, dry mouth, fatigue, nasal congestion, somnolence, tachycardia, orthostatic hypotension, and weight gain.

Drug Interactions

Iloperidone dose should be reduced by ½ when administered along with strong CYP2D6 inhibitors (e.g., fluoxetine and paroxetine), CYP3A4 inhibitors (e.g., ketoconazole and clarithromycin), or poor CYP2D6 metabolizers. If CYP2D6 or CYP3A4 inhibitors are discontinued, the iloperidone dose can be increased accordingly.

Dosage and Administration

Iloperidone is available in tablets of 1, 2, 4, 6, 8, 10, and 12 mg. Iloperidone should be up-titrated slowly to minimize orthostatic hypotension side effects, given its α-adrenergic antagonist properties. It is recommended to start at 1 mg twice daily, increasing the dose (by no more than 4 mg total/day) to a target dose of 6 to 12 mg twice daily (12 to 24 mg total/day). Efficacy in pivotal trials as demonstrated in the range of 6 to 12 mg twice daily. Iloperidone does not need to be administered with meals. As with all

antipsychotics, it is recommended that patients with initial positive response to the drug be continued beyond the acute response. If have discontinued iloperidone for more than 3 days, it is recommended that the initiation titration schedule be followed.

Dosing in Special Populations

PEDIATRIC USE. Safety and efficacy data supporting the use of iloperidone in child and adolescent patients has not been established.

GERIATRIC USE. Safety and efficacy data supporting the use of iloperidone in geriatric populations has not been established. Only 25 of 3,210 participants of iloperidone premarketing trials were 65 years old or greater.

RENAL IMPAIRMENT. Iloperidone undergoes almost total hepatic metabolism, with <1 percent excreted unchanged. Creatinine clearance <30 mL/min had minimal effect on C_{max} of iloperidone or its P88 and P95 metabolites. However, the area under the curve (AUC) of iloperidone increased by 24 percent, P88 decreased by 6 percent, and P95 increased by 52 percent, so prescribers should monitor patients with renal impairment with greater vigilance.

HEPATIC IMPAIRMENT. No dosing adjustments to iloperidone are needed for patients with mild hepatic impairment. It is recommended that prescribers exercise caution when administering iloperidone to patients to moderate hepatic impairment, and not prescribe iloperidone to patients with severe hepatic impairment. Specifically, in patients with mild hepatic impairment, there were no significant differences in pharmacokinetics of iloperidone or its major metabolites as compared to healthy controls. With moderate hepatic impairment, a twofold higher exposure to the P88 metabolite was observed as compared to healthy controls. Adequate data is not available for patients with severe hepatic impairment.

SMOKERS. Iloperidone is not a substrate for CYP1A2, and it is noted that smoking should not have an effect on iloperidone pharmacokinetics.

ASENAPINE

Asenapine is an antipsychotic with high affinity for 5-HT_{2A} and D_2 receptors, as well as several other serotonergic, dopaminergic, and adrenergic receptors. Important clinical properties include sublingual administration (with temporary oral hypoesthesia as a possible side effect) and minimal anticholinergic side effects. It is associated with mild to moderate sedation (roughly equivalent to olanzapine and haloperidol) and mild changes in body weight.

History

Asenapine was initially developed by Organon. It was approved by the FDA in 2009, with indications for the acute treatment of both schizophrenia and bipolar I disorder (manic or mixed episodes). In 2010, the approval indications were expanded to include maintenance treatment of schizophrenia as well as use as adjunctive therapy (to lithium or valproate) for the acute treatment of manic or mixed episodes in bipolar I disorder.

Chemistry

Asenapine belongs to the class of dibenzo-oxepino pyrroles (Fig. 34.28–1). It was derived chemically by modifying the structure of the tetracyclic antidepressant, mianserin. The chemical name of asenapine is (3aRS,12bRS)-5-Chloro-2-methyl-2,3,3a,12b-tetrahydro-1Hdibenzo[2,3:6,7]oxepino[4,5-c]pyrrole (2Z)-2-butenedioate (1:1). Its molecular formula is $C_{17}H_{16}ClNO{\cdot}C_4H_4O_4$ and its molecular weight is 401.84.

Pharmacological Action

Pharmacokinetics

ABSORPTION. Asenapine is absorbed sublingually and is 35 percent bioavailable by this route. The bioavailability if swallowed is less than 2 percent. Drinking water 2 to 5 minutes after asenapine administration decreases absorption, so it is recommended that eating and drinking are avoided for 10 minutes after administration. Consumption of food immediately prior to sublingual administration decreases asenapine exposure by 20 percent and consumption of food 4 hours after administration decreases exposure by 10 percent; this is due to increased hepatic blood flow.

DISTRIBUTION. Asenapine is rapidly distributed, has a volume of distribution approximately 20 to 25 L/kg, and is highly protein-bound in plasma. Asenapine reaches maximal plasma concentration approximately 1 hour after administration and with twice-daily dosing, reaches steady-state within 3 days.

METABOLISM AND ELIMINATION. The primary metabolic pathways for asenapine are glucuronidation by UGT1A4 and oxidative metabolism by cytochrome P450 isozymes (primarily CYP1A2; CYP3A4 and CYP2D6 to a lesser extent). Asenapine is a weak inhibitor of CYP2D6. Though tobacco smoking induces CYP1A2, the clearance of asenapine in smokers has not been found to be altered. After single-dose administration, 50 percent of labeled asenapine metabolites were recovered in urine and 40 percent in stool. The mean terminal half-life of asenapine is approximately 24 hours.

Pharmacodynamics

MECHANISM OF ACTION. Asenapine has high affinity for: serotonin 5-HT_{1A}, 5-HT_{1B}, 5-HT_{2A}, 5-HT_{2B}, 5-HT_{2C}, 5-HT_{5A}, 5-HT_6, and 5-HT_7 receptors; dopamine D_{2A}, D_{2B}, D_3, D_4, and D_1 receptors; adrenergic α_{1A}, α_{2A}, α_{2B}, and α_{2C} receptors, and histamine H1 receptors, acting as an antagonist at all of these receptors. It has moderate affinity for histamine H2 receptors and minimal affinity for muscarinic receptors. Its activity is mostly due to parent asenapine rather than metabolites.

Design and Interpretation of Clinical Drug Studies

Schizophrenia. The efficacy of asenapine for acute treatment of schizophrenia was supported by two of three 6-week randomized controlled trials of adults with DSM-IV schizophrenia who were having an illness exacerbation. The primary outcome measure on all three studies was the change in PANSS total score. In the first study (N = 174), asenapine 5 mg twice daily showed statistically significant improvement on the primary outcome measure, as well as the PANSS positive, negative, and general psychopathology subscales, as compared to placebo. The active comparator risperidone did not separate from placebo. In the second study (N = 448), asenapine 5 mg but not 10 mg, was statistically superior to placebo on the primary outcome measure. In the third study (N = 386), which was considered negative, neither 5- or 10-mg asenapine were superior to placebo, as compared to olanzapine 15 mg which was superior to placebo.

The efficacy of asenapine for maintenance treatment of schizophrenia was demonstrated in a randomized withdrawal design trial. Seven hundred patients with schizophrenia entered open-label asenapine treatment for 26 weeks. Three hundred-eighty-six patients who met stability criteria were randomized to either continue asenapine or withdraw to placebo. The primary outcome measure was time to relapse. Asenapine was statistically superior to placebo in increasing time to relapse as well as impending relapse (defined by prespecified changes in PANSS and CGI-S scores).

Bipolar Disorder. The efficacy for asenapine in treating acute mania was established in two trials in adults with DSM-IV bipolar I disorder who were having acute manic or mixed episodes. After a 7-day placebo washout period, patients were randomized to receive asenapine 5 or 10 mg twice daily, olanzapine 5 to 20 mg/day, or placebo. The primary outcome measure was the change in YMRS from baseline to day 21. Both asenapine and olanzapine were statistically superior to placebo on the primary outcome measure as well as the CGI-BP Severity of Illness score.

The efficacy of asenapine as an adjunctive therapy in treating acute mania was established in a 12-week placebo-controlled trial. Patients

experiencing manic or mixed episodes of bipolar I disorder with only partial response to lithium or valproic acid monotherapy were randomized to receive asenapine (5 to 10 mg twice daily, *N* = 154) or placebo (*N* = 162). Adjunctive asenapine was statistically superior to placebo at week 3 on the primary outcome measure (change in YMRS score).

Therapeutic Indications

Schizophrenia. The efficacy of asenapine for the treatment of schizophrenia was established from the results of two 6-week trials and one maintenance trial.

Bipolar Disorder

MONOTHERAPY. Asenapine is indicated for the acute treatment of manic or mixed episodes associated with bipolar I disorder. Efficacy was established in two 3-week monotherapy trials.

ADJUNCTIVE THERAPY. Asenapine is indicated as adjunctive therapy with either lithium or valproate for the acute treatment of manic or mixed episodes associated with bipolar I disorder. Efficacy was established in one 3-week trial.

Precautions and Adverse Reactions

Black Box Warning. Asenapine and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis compared with placebo.

QT Prolongation. Asenapine has been specifically assessed for its effect on QT interval, at doses of 5, 10, 15, and 20 mg twice daily versus placebo (*N* = 151). Asenapine was associated with an increase in QTc of 2 to 5 ms as compared to placebo and no patients showed greater than 60 ms increase or reached QTc ≥500 ms. In other asenapine premarketing clinical trials, there were no differences in QTc ≥500 ms between asenapine and placebo-treated patients, and there were no reports of Torsade de Pointes or other related events. It is recommended, however, that asenapine should not be combined with other medications known to increase the QTc interval.

Neuroleptic Malignant Syndrome. NMS has been associated with the use of antipsychotic medications, including asenapine. If NMS is suspected, asenapine and other antipsychotic medications should be discontinued immediately and the patient should undergo intensive supportive care and medical monitoring.

Tardive Dyskinesia. Tardive dyskinesia can develop in all patients

treated with antipsychotic drugs. Its prevalence is highest among the elderly. The propensity of asenapine to cause tardive dyskinesia is not known. In general, SGAs have a lower tardive dyskinesia risk than FGAs.

Metabolic Changes. The precise risk with asenapine of hyperglycemia-related adverse reactions, such as ketoacidosis or hyperosmolar coma, is not known. However, hyperglycemia-related adverse reactions have been associated with other SGAs.

In pooled 6-week trials, 5 to 10 mg asenapine twice daily (*N* = 377) was associated with a 3.2 mg/dL increase in glucose, as compared to –0.2 mg/dL in the placebo group (*N* = 232). In a 52-week trial of asenapine, with comparator control, the mean increase of fasting blood glucose was 2.4 mg/dL.

Dyslipidemia has also been associated with asenapine use. In pooled 6-week trials, 5- to 10-mg asenapine twice daily was associated with a 0.4 mg/dL increase in total cholesterol (*N* = 539), 1.3 mg/dL increase in LDL (*N* = 465), 0.5 mg/dL increase in high-density lipoprotein (HDL) (*N* = 480), and 3.8 mg/dL increase in triglycerides (*N* = 380). In the same group of pooled trials, the placebo group showed a 2.2 mg/dL decrease in total cholesterol (*N* = 351), 0.1 mg/dL increase in LDL (*N* = 285), 0.5 mg/dL increase in HDL (*N* = 290), and 7.6 mg/dL decrease in triglycerides (*N* = 233). In a 52-week trial of asenapine, with comparator control, asenapine was associated a 6 mg/dL decrease in total cholesterol and 9.8 mg/dL decrease in triglycerides. Of note, these observed decreases in lipid parameters over 52 weeks could be a consequence of the study participants switching to asenapine from another antipsychotic with higher propensity for dyslipidemia.

Asenapine may also affect body weight. In pooled 6-week studies, asenapine 5 to 10 mg twice daily (*N* = 532) was associated with a 1.1 kg increase in body weight as compared to a mean change of 0 kg in the placebo group (*N* = 348). In a comparator-controlled 52-week trial (*N* = 887), asenapine was associated with a 0.9 kg increase in body weight, with 14.7 percent of participants increasing ≥7 percent of their baseline body weight. The weight gain was most pronounced in participants with baseline BMI <23.

Seizures. In premarketing clinical trials, seizures were reported in 0 percent of patients taking asenapine 5 mg twice daily (*N* = 572), 0.3 percent of patients taking asenapine 10 mg twice daily (*N* = 379), as compared to 0 percent of patients taking placebo (*N* = 503 in short-term schizophrenia trial and *N* = 203 in short-term bipolar mania trial).

Orthostatic Hypotension and Syncope. Asenapine possesses α1-adrenergic antagonist properties and has been reported to cause orthostatic

hypotension and syncope in some patients, particularly early in treatment. In 6-week schizophrenia trials, 1/527 patients taking asenapine experienced syncope, as compared to 1/378 of patients taking placebo. In 3-week bipolar disorder trials, 1/379 patients taking asenapine experienced syncope, as compared to no patients taking placebo. When including all participants in premarketing clinical trials (including 52-week treatment), syncope was reported in 11/1,953 (0.6 percent) of patients taking asenapine.

Hyperprolactinemia. As with most antipsychotic medications, asenapine can increase prolactin levels. In premarketing clinical trials for schizophrenia and bipolar disorder, adverse events related to hyperprolactinemia were found in 0.4 percent of subjects in the asenapine group, as compared to 0 percent in the placebo group.

Leukopenia/Neutropenia. Leukopenia and neutropenia events have been infrequently reported to be temporally associated with asenapine use. It is recommended that patients with low WBC or history of drug-induced leukopenia or neutropenia have their blood counts monitored frequently during the first few months of treatment and that asenapine be discontinued if a WBC decrease is observed that cannot be explained by other factors. Patients with ANC <1,000/mm^3 should discontinue asenapine until it has recovered.

Potential for Cognitive and Motor Impairment. In short-term premarketing schizophrenia trials, somnolence occurred in 15 percent of patients taking asenapine 5 mg twice daily (N = 274) and 13 percent of patients taking 10 mg twice daily (N = 208), as compared to 7 percent of patients taking placebo (N = 389). Somnolence was noted to be most significant during the first week of treatment and led to discontinuation in 0.6 percent of all patients in short-term placebo-controlled trials (N = 1,953).

Body Temperature Regulation. In premarketing short-term trials of asenapine for schizophrenia and bipolar disorder, adverse events suggesting elevated body temperature were uncommon (≤1 percent) and comparable to placebo.

Dysphagia. Dysphagia was reported in 0.2 percent of patients with schizophrenia who were taking asenapine (N = 572), as compared to 0 percent of patients taking placebo (N = 378). There were no reported incidents of dysphagia in premarketing trials for bipolar mania.

Pregnancy. Asenapine is currently designated category C, indicating

that there are no adequate and well-controlled studies of asenapine in pregnant women. In animal studies, asenapine increased postimplantation loss and decreased rodent pup weight and survival though was not found to cause other structural abnormalities. It was found to be nonteratogenic in rat and rabbit studies. It is recommended to be only used during pregnancy if the potential benefit justifies the potential risk to the fetus.

Lactation. Asenapine is excreted in the milk of rats during lactation. It is unknown if this also occurs in humans. It is recommended that women receiving asenapine not breast feed.

Common and Drug-Related Adverse Reactions in Clinical Trials. Adverse reactions that occurred in ≥5 percent of short-term premarketing trial patients receiving asenapine (in either schizophrenia or bipolar disorder trials) and at ≥2 times the rate in the placebo groups included: EPS (both *akathisia* and *other than akathisia* types), oral hypoesthesia, somnolence, dizziness, and weight gain. In short-term premarketing schizophrenia trials (N = 572 receiving asenapine and N = 378 receiving placebo), the only adverse reaction that was dose-related was akathisia.

Oral hypoesthesia is an adverse reaction associated with asenapine but no other SGAs. This can occur immediately after sublingual administration and usually resolves within 1 hour. Oral hypoesthesia was associated with asenapine discontinuation in short-term premarketing trials for bipolar mania, occurring in 1.1 percent of the asenapine group patients (N = 379) and 0 percent of the placebo group (N = 203). In postmarketing monitoring of asenapine use, other oral reactions have been reported to have occurred with asenapine sublingual administration, including oral ulcers, blisters, peeling/sloughing, and inflammation.

Drug Interactions

Asenapine is primarily metabolized by UGT1A4 (direct glucuronidation) and by several cytochrome P450 isozymes (mostly CYP1A2). As such, administration of asenapine 5 mg with the CYP1A2 inhibitor fluvoxamine (25 mg twice daily) for 8 days increased the asenapine AUC by 29 percent; caution is indicated for coadministration of asenapine and fluvoxamine. Coadministration with the CYP2D6 inhibitor paroxetine was associated with a 9 percent decrease in the asenapine AUC. As asenapine is a weak inhibitor of CYP2D6 itself, it is recommended to be used with caution with drugs that are both substrates and inhibitors of CYP2D6 (such as paroxetine). Though valproic acid is an inhibitor of UGT1A4, coadministration of asenapine with valproic acid 500 mg twice daily did not significantly change asenapine levels.

Dosage and Administration

Asenapine is available as sublingual tablets of 5 or 10 mg. Patients should place the tablet under the tongue and allow it to completely dissolve, which occurs within seconds. Product labeling indicates asenapine tablets should not be split, crushed, chewed, or swallowed, and patients should not eat or drink for 10 minutes after administration.

For the treatment of acute schizophrenia, the recommended starting and target dose of asenapine is 5 mg twice daily. Based on clinical response, the dose can be increased to 10 mg twice daily if tolerated. For maintenance treatment of schizophrenia, the starting dose is recommended to be 5 mg twice daily, and in 1 week it can be increased up to 10 mg twice daily if tolerated.

For the acute treatment of manic or mixed episodes associated with bipolar I disorder, the recommended starting dose is 10 mg twice daily. If not tolerated, the dose can be decreased to 5 mg twice daily. For adjunctive therapy with either lithium or valproate, the starting dose of asenapine is recommended to be 5 mg twice daily, and based on clinical response, the dose can be increased to 10 mg twice daily if tolerated. Asenapine is not currently indicated for the maintenance treatment of bipolar disorder.

Asenapine administration is contraindicated in patients with a known hypersensitivity to the medication. Hypersensitivity reactions in clinical trials have included anaphylaxis, angioedema, hypotension, tachycardia, swollen tongue, dyspnea, wheezing, and rash, and some of these cases occurred after the first dose.

Dosing in Special Populations

GERIATRIC USE. Only 25 of 2,250 (1.1 percent) premarketing trial patients receiving asenapine were 65 years or older. In elderly patients, the asenapine AUC was on average 40 percent higher than younger than adults, so elderly patients taking asenapine should be monitored carefully.

HEPATIC IMPAIRMENT. Hepatic impairment has been found to increase the exposure to asenapine, according to the extent of hepatic impairment. Severe hepatic impairment (Child-Pugh C) was associated with a sevenfold increase in asenapine concentrations. Dosage adjustments are not required for mild (Child-Pugh A) or moderate (Child-Pugh B) hepatic impairment, but asenapine is not recommended to patients with severe hepatic impairment.

RENAL IMPAIRMENT. Renal impairment (mild, moderate, or severe, but not on dialysis) was not found to affect exposure to asenapine, so dosage adjustment is not required.

LURASIDONE

Lurasidone is an antipsychotic that is a potent antagonist at 5-HT_{2A}, D_2,

and 5-HT_7 receptors and a partial agonist at the 5-HT_{1A} receptor. It is among the SGAs (along with aripiprazole and ziprasidone) that have a relatively low liability for metabolic side effects. Patients should be cautioned to take lurasidone with food. Its most common side effects include sedation, extrapyramidal side effects (including akathisia), and nausea.

History

Lurasidone was developed by Dainippon Sumitomo Pharma and marketed by Sunovion in the United States. It was approved by the FDA for the treatment of schizophrenia in adults in 2010. Subsequently approval was granted in 2013 for the treatment of adults with bipolar depression (as both monotherapy and as an adjunct to lithium or valproate therapy).

Chemistry

Lurasidone belongs to the class of benzisothiazol derivatives, with the full chemical name of (3aR,4S,7R,7aS)-2-[((1R,2R)-2-{[4-(1,2-benzisothiazol-3-yl)-piperazin-1-yl]methyl}cyclohexyl)methyl]hexahydro-1H-4,7-methanisoindol-1,3-dione. Its molecular formula is $C_{28}H_{36}N_4O_2S$ and its molecular weight is 492.676 (free base). It is currently used in the form of its hydrochloride salt.

Pharmacological Action

Pharmacokinetics

ABSORPTION. Lurasidone has the relatively low oral bioavailability of 9 to 19 percent and its absorption is affected by food intake. Administration of lurasidone with a meal of at least 350 calories increases drug exposure approximately.

DISTRIBUTION. Peak serum concentrations are reached in 1 to 3 hours. The volume of distribution is 6,173 L and lurasidone is highly bound to serum proteins.

METABOLISM AND ELIMINATION. Lurasidone is metabolized mostly by CYP3A4 and is not a substrate for CYP1A1, CYP1A2, CYP2A6, CYP4A11, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, or CYP2E1. Its major biotransformation pathways are oxidative *N*-dealkylation, hydroxylation of its norbornane ring, and S-oxidation. It has two major active and two major inactive metabolites. Lurasidone is eliminated primarily in the feces (80 percent), with 9 percent of labeled metabolites recovered in urine. The mean elimination half-life was found to be 18 hours.

Pharmacodynamics

MECHANISM OF ACTION. Lurasidone is an antagonist with high affinity for the dopamine D_2 receptor as well as serotonin 5-HT_{2A} and 5-HT_7 receptors. It has been found to occupy 60 to 70 percent of D_2 receptors at a 40-mg dose. In addition, it is an antagonist with moderate affinity for α_{2A} and α_{2C} adrenergic receptors and is a partial agonist at the 5-HT_{1A} receptor. It has low affinity for the serotonin 5-HT_{2C} receptor and negligible affinity for histamine H_1 and muscarinic M_1 receptors.

Design and Interpretation of Clinical Drug Studies

Schizophrenia. Efficacy was established for the treatment of schizophrenia by five 6-week randomized controlled trials, with two including active comparators. The studies were conducted in adult patients with DSM-IV schizophrenia. Two of these studies (*N* = 145 and B = 180) used the change in Brief Psychiatric Rating Scale derived (BPRSd) as the primary outcome measure, and lurasidone significantly outperformed placebo at 40, 80, and 120 mg/day dosages. A third study (*N* = 489) used the change in PANSS total score as the primary outcome measure, and lurasidone 80 mg/day was found significantly superior to placebo, but the 40 and 80 mg/day dosages did not separate from placebo. A fourth study (*N* = 473) used olanzapine 15 mg/day as an active comparator. Lurasidone 40 and 120 mg/day, as well as olanzapine 15 mg/day, were significantly more effective than placebo on the primary outcome measure of change in PANSS total score. In the fifth study (*N* = 482), patients were randomized to receive lurasidone 80 mg/day, lurasidone 160 mg/day, quetiapine extended-release 600 mg/day, or placebo. The primary outcome measure was the change in PANSS total score. Both doses of lurasidone, as well as quetiapine extended-release, were associated with significant improvement in the primary outcome measure as compared to placebo. A sixth study (*N* = 353) involved random assignment to lurasidone 20, 40, and 80 mg/day or haloperidol 10 mg/day as an active comparator. This study was considered failed, because lurasidone at all dosages, as well as haloperidol, did not separate from placebo on the primary outcome measure (BPRSd).

Long-term effectiveness of lurasidone has been assessed in a 52-week controlled trial in which 639 stable outpatients with schizophrenia were randomized to receive lurasidone 40 to 120 mg/day or risperidone 2 to 6 mg/day. The two groups did not statistically differ in relapse rate (defined as a 30 percent increase in PANSS total score compared to baseline). Both groups also showed statistically equivalent improvement on PANSS total score over the 12 months.

Depressive Episodes Associated with Bipolar I Disorder

MONOTHERAPY. Efficacy was established in a 6-week randomized

controlled trial (*N* = 485) of adults with DSM-IV bipolar I disorder in a major depressive episode, with or without rapid cycling, and without psychotic features. Participants were randomized to receive flexible-dose lurasidone 20 to 60, 80 to 120 mg/day, or placebo. The primary outcome measure was the change in MADRS score from baseline to week 6. Both dose ranges of lurasidone were superior to placebo, and the effectiveness of the lower- versus higher-range dosing of lurasidone were not statistically different.

ADJUNCTIVE THERAPY. Efficacy was established in a 6-week randomized controlled trial (*N* = 340) of adults with DSM-IV bipolar I disorder in a major depressive episode, with or without rapid cycling, and without psychotic features. Patients who were still symptomatic despite lithium or valproic acid treatment were randomized to receive flexible-dose lurasidone 20 to 120 mg/day or placebo. The primary outcome measure was the change in MADRS from baseline to week 6. Lurasidone, as an adjunct to valproic acid or lithium, was statistically superior to placebo in the reduction of MADRS score.

Therapeutic Indications

Schizophrenia. Lurasidone is indicated for the treatment of patients with schizophrenia. Its efficacy was established in five short-term studies of adults with schizophrenia. Lurasidone is not indicated at the time of writing for longer-term use (greater than 6 weeks).

Depressive Episodes Associated with Bipolar I Disorder

MONOTHERAPY. Lurasidone is indicated as monotherapy for the treatment of patients with major depressive episodes associated with bipolar I disorder. Its efficacy for this indication was established in a 6-week monotherapy study of adults with bipolar depression.

ADJUNCTIVE THERAPY WITH LITHIUM OR VALPROATE. Lurasidone is indicated as adjunctive therapy (with either lithium or valproate) for the treatment of patients with major depressive episodes associated with bipolar I disorder. Its efficacy for this indication was established in a 6-week study of adults with bipolar depression who were being treated with lithium or valproate.

Lurasidone is not indicated at time of writing for longer-term (greater than 6 weeks) treatment of bipolar depression. Lurasidone is also not indicated at time of writing for the treatment of mania associated with bipolar I disorder.

Precautions and Adverse Reactions

Lurasidone is contraindicated in individuals with hypersensitivity to

lurasidone or any components in the tablets. Lurasidone is also contraindicated in patients using strong CYP3A4 inhibitors (e.g., ketoconazole, clarithromycin, ritonavir, voriconazole, and mibefradil) or CYP3A4 inducers (e.g., rifampin, avasimibe, St. John wort, phenytoin, and carbamazepine).

Black Box Warnings

INCREASED MORTALITY IN ELDERLY PATIENTS WITH DEMENTIA. Lurasidone and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis compared with placebo.

SUICIDAL THOUGHTS AND BEHAVIORS. Lurasidone also carries a black box warning for suicidal thoughts and behavior in children, adolescents, and young adults. This warning is associated with all antidepressant medications, including SGAs with indications for the treatment of depression.

Neuroleptic Malignant Syndrome. NMS has been associated with the use of antipsychotic medications, including lurasidone. If NMS is suspected, lurasidone and other antipsychotic medications should be discontinued immediately and the patient should undergo intensive supportive care and medical monitoring.

Tardive Dyskinesia. Tardive dyskinesia can develop in all patients treated with antipsychotic drugs. Its prevalence is highest among the elderly. The propensity of lurasidone to cause tardive dyskinesia is not known. In general, SGAs have a lower tardive dyskinesia risk than FGAs.

Metabolic Changes. SGAs have been associated with hyperglycemia, including severe cases with ketoacidosis, hyperosmolar coma, and death. Insufficient data is available to assess whether lurasidone is associated with these risks. When interpreting changes of metabolic parameters in antipsychotic clinical trials, it is important to consider that observed improvements in parameters could be a consequence of participants switching to lurasidone from other antipsychotics with higher propensities for adverse metabolic effects.

FASTING PLASMA GLUCOSE. In pooled short-term schizophrenia studies, changes of fasting serum glucose with placebo (N = 680) and lurasidone 20 (N = 71), 40 (N = 478), 80 (N = 508), 120 (N = 283), and 160 mg/day (N = 113) were 0, −0.6, +2.6, −0.4, +2.5, and +2.5 mg/dL, respectively. In uncontrolled longer-term schizophrenia studies, lurasidone was associated with a change in fasting glucose of +2.3 mg/dL at week 52 (N = 307). In the short-term bipolar depression monotherapy study, changes of fasting

serum glucose with placebo (*N* = 148) and lurasidone 20 to 60 (*N* = 140) and 80 to 120 mg/day (*N* = 143) were +1.8, -0.8, and +1.8 mg/dL, respectively. In the uncontrolled longer-term bipolar depression monotherapy study, lurasidone was associated with a +1.2 mg/dL change in fasting glucose at week 24 (*N* = 129). In the pooled short-term bipolar depression adjunctive therapy study, changes of fasting serum glucose with placebo (*N* = 302) and lurasidone 20 to 120 mg/day (*N* = 319) were −0.9 and 1.2 mg/dL, respectively. In the uncontrolled longer-term bipolar depression adjunctive therapy study, lurasidone was associated with a +1.7 mg/dL change in glucose at week 24 (*N* = 88).

FASTING LIPIDS. In pooled short-term schizophrenia studies, changes of fasting total cholesterol with placebo (*N* = 660) and lurasidone 20 (*N* = 71), 40 (*N* = 466), 80 (*N* = 499), 120 (*N* = 268), and 160 mg/day (*N* = 115) were −5.8, −12.3, −5.7, −6.2, −3.8, and −6.9 mg/dL, respectively. In uncontrolled longer-term schizophrenia studies, lurasidone was associated with a change in fasting total cholesterol of −6.9 mg/dL at week 52 (*N* = 307). In the short-term bipolar depression monotherapy study, changes of fasting total cholesterol with placebo (*N* = 147) and lurasidone 20 to 60 (*N* = 140) and 80 to 120 mg/day (*N* = 144) were −3.2, +1.2, and −4.6 mg/dL, respectively. In the uncontrolled longer-term bipolar depression monotherapy study, lurasidone was associated with a −0.5 mg/dL change in fasting glucose at week 24 (*N* = 130). In the short-term bipolar depression adjunctive therapy study, changes of fasting serum total cholesterol with placebo (*N* = 303) and lurasidone 20 to 120 mg/day (*N* = 321) were −2.9 and -3.1 mg/dL, respectively. In the uncontrolled longer-term bipolar depression adjunctive therapy study, lurasidone was associated with a -0.9 mg/dL change in fasting total cholesterol at week 24 (*N* = 88).

WEIGHT CHANGE. In pooled short-term schizophrenia studies, changes of body weight with placebo (*N* = 696) and lurasidone 20 (*N* = 71), 40 (*N* = 484), 80 (*N* = 526), 120 (*N* = 291), and 160 mg/day (*N* = 114) were −0.02, −0.15, +0.22, +0.54, +0.68, and +0.60 kg, respectively. In uncontrolled longer-term schizophrenia studies, lurasidone was associated with a change in body weight of −0.73 kg at week 52 (*N* = 377). In the short-term bipolar depression monotherapy study, changes of body weight with placebo (*N* = 151) and lurasidone 20 to 60 (*N* = 143) and 80 to 120 mg/day (*N* = 147) were −0.04, +0.56, and +0.02 kg, respectively. In the uncontrolled longer-term bipolar depression monotherapy study, lurasidone was associated with a −0.02 kg change in body weight week 24 (*N* = 130). In the short-term bipolar depression adjunctive therapy study, changes of body weight with placebo (*N* = 307) and lurasidone 20 to 120 mg/day (*N* = 327) were +0.16 and +0.11 kg, respectively. In the uncontrolled longer-term bipolar

depression adjunctive therapy study, lurasidone was associated with a +1.28 kg change in body weight at week 24 (*N* = 86).

QT Prolongation. Lurasidone was evaluated in a randomized controlled thorough QT study in 43 patients with schizophrenia or schizoaffective disorder. Participants were administered lurasidone 120 or 600 mg and ECGs were assessed 2 to 4 hours after dosing. Patients who received the 120-mg dose experienced a 7.5 ms increase in QTc, and those who received the 600 mg experienced a 4.6 ms increase. There was no observed dose–response relationship. In short-term schizophrenia and bipolar depression clinical trials, no study participants experienced QTc prolongation to greater than 500 ms.

Seizures. In short-term premarketing schizophrenia studies, seizures occurred in 0.1 percent (2/1,508) of patients treated with lurasidone as compared to 0.1 percent (1/708) patients treated with placebo. In short-term premarketing bipolar depression studies, no patients experienced seizures.

Orthostatic Hypotension and Syncope. Lurasidone may cause orthostatic hypotension and syncope. In pooled premarketing short-term schizophrenia studies, orthostatic hypotension was reported in 0.3 percent of lurasidone and 0.1 percent of placebo patients; syncope was reported in 0.1 percent of lurasidone and no placebo patients. In the premarketing short-term bipolar depression study, there were no reported adverse events of orthostatic hypotension and syncope. However, orthostatic hypotension, as assessed by vital signs, occurred in 0.6 percent of lurasidone patients as compared to 0 percent receiving placebo. In the premarketing short-term adjunctive therapy bipolar depression study, there were no reported adverse events of hypotension and syncope. Orthostatic hypotension, as assessed by vital signs, occurred in 1.1 percent of lurasidone patients and 0.9 percent of placebo patients.

Hyperprolactinemia. As with most antipsychotic medications, lurasidone can increase prolactin levels. In pooled short-term schizophrenia studies, changes in prolactin level with placebo (*N* = 672) and lurasidone 20 (*N* = 70), 40 (*N* = 476), 80 (*N* = 495), 120 (*N* = 284), and 160 mg/day (*N* = 115) were −1.9, −1.1, −1.4, −0.2, +3.3, and +3.3 ng/mL, respectively. In uncontrolled longer-term schizophrenia studies, lurasidone was associated with a change in prolactin level of −2.2 ng/mL at week 52 (*N* = 307). In the short-term bipolar depression monotherapy study, changes in prolactin level with placebo (*N* = 147) and lurasidone 20 to 60 (*N* = 140) and 80 to 120 mg/day (*N* = 144) were +0.3, +1.7, and +3.5 ng/mL, respectively. In the uncontrolled longer-term bipolar depression

monotherapy study, lurasidone was associated with a –1.15 ng/mL change in prolactin level week 24 (*N* = 130). In the short-term bipolar depression adjunctive therapy study, changes in prolactin level with placebo (*N* = 301) and lurasidone 20 to 120 mg/day (*N* = 321) were 0 and +2.8 ng/mL, respectively. In the uncontrolled longer-term bipolar depression adjunctive therapy study, lurasidone was associated with a –2.9 ng/mL change in prolactin level at week 24 (*N* = 88).

Potential for Cognitive and Motor Impairment. Lurasidone can cause somnolence, which can impair judgment, thinking, and motor skills. In short-term premarketing schizophrenia studies, somnolence was reported by 17 percent of patients receiving lurasidone (*N* = 1,508) and 7.1 percent of patients receiving placebo (*N* = 708). In the premarketing bipolar depression monotherapy study, 6.5 percent of the placebo group (*N* = 168), 7.3 percent of the lurasidone 20 to 60 mg/day group (*N* = 164), and 13.8 percent of the lurasidone 80 to 120 mg/day group (*N* = 167) reported somnolence. In the premarketing bipolar depression adjunctive therapy study, somnolence was reported by 11.4 percent of the lurasidone 20 to 120 mg/day group (*N* = 360) as compared to 5.1 percent of the placebo group (*N* = 334).

Suicide. In short-term premarketing schizophrenia studies, the incidence of suicidal ideation was 0.4 percent (6/1,508) in lurasidone-treated patients as compared to 0.8 percent (6/708) in placebo-treated patients. There were no suicide attempts or completed suicides in these studies. In the short-term bipolar depression monotherapy study, there were no reported events of suicidal ideation, attempts, or completed suicides in either group. In the short-term bipolar depression adjunctive therapy study, the incidence of suicidal ideation was 1.1 percent (4/360) in lurasidone-treated patients and 0.3 percent (1/334) in placebo-treated patients. There were no suicide attempts or completed suicides in this study.

Serum Creatinine. In short-term schizophrenia trials, serum creatinine levels increased 0.05 mg/dL in lurasidone-treated participants as compared to 0.02 mg/dL in placebo-treated participants. Changes in creatinine levels appeared to be dose-dependent, with 7 percent of patients receiving lurasidone 160 mg/day (*N* = 121) shifting from normal to abnormally high creatinine levels as compared to 2 percent of those receiving placebo (*N* = 708).

Extrapyramidal Symptoms. In premarketing schizophrenia studies, EPS (excluding akathisia) was reported in 13.5 percent of patients receiving lurasidone as compared to 5.8 percent of patients receiving placebo.

Akathisia was reported in 12.9 percent of patients receiving lurasidone as compared to 3 percent of patients receiving placebo. These effects were dose-dependent, with the highest percentage of events occurring in the group receiving lurasidone 120 mg/day. EPS were also reported in short-term bipolar depression studies with similar frequencies. Dystonic symptoms occurred in 4.2 percent of patients in premarketing schizophrenia trials versus 0.8 percent of those receiving placebo, and 7/1,508 participants (0.5 percent) discontinued the clinical trial due to dystonic events, which occurred during the first few days of treatment. Dystonia occurred less frequently with lurasidone in bipolar depression studies though still occurred more often than the placebo groups. No participants in bipolar depression studies discontinued participation due to dystonic symptoms.

Pregnancy. Lurasidone is classified as pregnancy category B, which indicates that animal reproductive studies did not identify a specific risk to the fetus, but there are no adequate and well-controlled studies in pregnant women. Lurasidone should only be used in pregnancy if the potential benefit justifies the potential risk to the fetus.

Nursing Mothers. Rat studies demonstrated excretion of lurasidone in milk during lactation. It is unknown whether this also occurs in humans. Mothers should consider discontinuation of nursing due to the potential for adverse reactions in infants.

Common and Drug-Related Adverse Reactions in Clinical Trials. In short-term premarketing schizophrenia studies, the most common adverse reactions associated with lurasidone (defined as incidence ≥5 percent and at least twice the rate of placebo) were somnolence, akathisia, EPS, and nausea. In the bipolar depression monotherapy study, the most common adverse reactions associated with lurasidone were akathisia, EPS, somnolence, nausea, vomiting, diarrhea, and anxiety. In the bipolar depression adjunctive therapy study, the most common adverse reactions associated with lurasidone were akathisia and somnolence. In premarketing studies for both schizophrenia and bipolar disorder, there were no adverse reactions associated with lurasidone discontinuation (defined as ≥2 percent occurrence and at least twice the rate of placebo).

Drug Interactions

It is recommended that lurasidone not be used in conjunction with strong CYP3A4 inhibitors (e.g., ketoconazole, clarithromycin, ritonavir, voriconazole, and mibefradil). If lurasidone is used with moderate CYP3A4 inhibitors (e.g., diltiazem, atazanavir, erythromycin, fluconazole, and verapamil), the lurasidone dose should be reduced by ½. Grapefruit and

grapefruit juice should be avoided in conjunction with lurasidone use.

Lurasidone is contraindicated for use with strong CYP3A4 inducers (e.g., rifampin, avasimibe, St. John wort, phenytoin, and carbamazepine). If lurasidone is used in conjunction with a moderate CYP3A4 inducer, it may be necessary to increase lurasidone dosage after at least 7 days of treatment with the CYP3A4 inducer.

Lithium and valproic acid dosages do not need to be adjusted if used concomitantly with lurasidone. Dosages of P-glycoprotein substrates (i.e., digoxin) and other CYP3A4 substrates (e.g., midazolam or oral contraceptives) also do not need to be adjusted if used with lurasidone.

Dosage and Administration

Lurasidone is available in 20-, 40-, 60-, 80-, and 120-mg tablets. It is recommended that lurasidone is taken with at least 350 calories of food to improve absorption.

For the treatment of schizophrenia, it is recommended to start lurasidone at 40 mg once daily. Lurasidone has been found effective for schizophrenia in the dose range of 40 to 160 mg daily, and the maximum recommended dose is 160 mg daily.

For the treatment of bipolar I depression, the recommended starting dose is 20 mg once daily, given as either monotherapy or as an adjunct to lithium or valproate. Lurasidone has been found effective for bipolar I depression in the dose range of 20 to 120 mg daily, and the maximum recommended dose for both monotherapy and adjunct therapy is 120 mg daily.

Dosing in Special Populations

GERIATRIC USE. There were insufficient numbers of patients aged 65 and older in lurasidone trials to determine whether they respond differently to the medication than younger adults. In a limited sample of elderly patients with psychosis (65 to 85 years of age), plasma concentrations of lurasidone after 20 mg/day administration were similar to those in younger adults.

HEPATIC IMPAIRMENT. In patients with moderate (Child-Pugh Score of 7 to 9) or severe hepatic impairment (Child-Pugh score of 10 to 15), dose reduction in lurasidone is recommended. The recommended starting dose is 20 mg daily and the maximum recommended dose in moderate or severe hepatic impairment is 80 and 40 mg/day, respectively.

RENAL IMPAIRMENT. In patients with moderate or severe renal impairment (creatinine clearance less than 50 mL/min), dose reduction in lurasidone is recommended. In this population, the recommended starting dose is 20 mg daily, and the maximum recommended dose is 80 mg daily.

OTHER NOVEL ANTIPSYCHOTIC AGENTS

Two new antipsychotics have been approved by the FDA in 2015. Cariprazine is an antipsychotic that is a partial agonist at D_2 and D_3 receptors with substantially greater affinity for D_3R. It is unclear how its clinical properties will differ from those of aripiprazole, the only available dopamine receptor partial agonist. Brexpiprazole is a partial agonist at D_2 and 5-HT_{1A} receptors as well as an antagonist at 5-HT_{2A} receptors. There is also an interest in developing new long-acting formulations of current antipsychotics that could be administered at longer intervals.

Although there is an interest in developing antipsychotic agents that are not dopamine receptor antagonists, there has been only limited progress in finding effective agents. Pomaglumetad, an agonist for metabotropic glutamate receptors 2 and 3, appeared to be effective as an antipsychotic in an initial trial, but failed in subsequent phase II trials. Sarcosine, an agent that inhibits the glycine transporter (GlyT-1), has been reported to be effective as an antipsychotic. The results from initial studies will need to be replicated. Another GlyT-1 inhibitor, bitopertin did not appear to be an effective antipsychotic. Pimavanserin, an inverse agonist at the 5-HT_{2A} receptor, is being studied as a possible adjunctive agent that could be added to a subtherapeutic dose of an antipsychotic. Phosphodiesterase inhibitors are being studied as agents that may be effective when added to antipsychotic. A number of these agents are in early trials.

CARIPRAZINE

Cariprazine is a partial agonist at dopamine D_3 and D_2 receptors and serotonin 5-HT_{1A} receptors. The binding affinity for D_3 is an order of magnitude higher than the affinity for D_2. It is plausible that the preference for D_3 might result in improvements in cognition, mood, and negative symptoms.

History

Cariprazine was developed by Gedeon Richter, a pharmaceutical company located in Budapest, Hungary. It was subsequently licensed to Allergan (which was previously known as Actavis) for the United States and Canada. Cariprazine was initially approved for the treatment of schizophrenia and bipolar disorder in September, 2015, under the trade name Vraylar.

Pharmacokinetics

ABSORPTION. Cariprazine is well absorbed following oral administration, reaching peak concentrations between 3 and 6 hours after dosing. Absorption is not substantially affected by food.

Pharmacokinetics. The activity of orally administered cariprazine is mediated by the parent drug and its two major active metabolites, desmethyl cariprazine (DCAR) and didesmethyl cariprazine (DDCAR). The metabolites appear to be equipotent to cariprazine at target therapeutic receptors. After multiple dose administrations, mean cariprazine and DCAR concentrations reached steady state between 1 and 2 weeks; mean DDCAR concentrations reach steady state in 4 to 8 weeks. The mean half-life of cariprazine is 2 to 5 days, whereas the half-life of DDCAR is considerably longer and results in substantial plasma accumulation. This very long half-life of the active metabolite indicates that patients who are started on cariprazine are likely to have a gradual increase in active drug over several weeks. Moreover, when the drug is discontinued, the active moiety is likely to remain in plasma for weeks.

METABOLISM AND ELIMINATION. Cariprazine is extensively metabolized by CYP3A4 and CYP2D6 to DCAR and DDCAR. Approximately 21 percent of the daily dose is found in urine. The excretion half-life is 2 to 4 days for the parent drug and 1 to 3 weeks for the active metabolite.

PHARMACODYNAMICS AND MECHANISM OF ACTION. In vitro studies indicated that cariprazine has greater affinity for D_3 compared with D_2 receptors. A recent PET study found that subjects treated with 1 mg/day of cariprazine had mean D_3 and D_2 receptor occupancies of 76 and 45 percent, respectively, indicating selectivity for D_3 over D_2 receptors at low doses.

Cariprazine is unique among available antipsychotics in having much greater affinity for D_3 as compared to D_2 receptors. In addition, it has greater affinity for D_3 receptors than dopamine itself. It is unclear whether this activity is related to its antipsychotic properties since studies in rodents indicate that D_3 antagonists are not effective in preclinical models of schizophrenia. It is possible that the combination of D_3 and D_2 activity of cariprazine results in antipsychotic activity, and potentially confers additional activity for other domains of psychopathology.

Design and Interpretation of Clinical Drug Studies

The efficacy of cariprazine in acute exacerbations of schizophrenia was demonstrated in three positive clinical trials. In a Phase II trial, patients treated with 1.5, 3, or 4.5 mg daily of cariprazine or 4 mg of risperidone showed significantly greater improvement in total PANSS compared with placebo. Another trial with a similar population found that 3 and 6 mg of cariprazine as well as 10 mg of aripiprazole resulted in greater improvement than placebo. A third trial in acute schizophrenia found that patients dosed from 3 to 6 mg and 6 to 9 mg demonstrated greater improvements than placebo-treated patients.

A separate trial in schizophrenia randomized patients with predominant negative symptoms to 26 weeks of double-blind treatment with cariprazine (mean dose, 4.2 mg/day) or risperidone (mean dose 3.8 mg/day). Patients assigned to cariprazine demonstrated significantly greater improvement in negative symptoms than patients assigned to risperidone, as measured by the PANSS negative symptoms factor score.

Cariprazine was also evaluated in 3-week trials for bipolar I disorder patients with manic or mixed states. In each of the trials, improvement was evaluated with the Young Mania Rating Scale (YMRS). In two flexible dose studies (cariprazine doses of 3 to 12 mg/day), cariprazine was more effective than placebo. A third study compared cariprazine at 3 to 6 and 6 to 12 mg/day with placebo. Both dose ranges of cariprazine were superior to placebo and there was no efficacy advantage for the higher-dose range.

A study of patients with major depressive disorder (MDD) and an inadequate response to an antidepressant randomized subjects to 8 weeks of treatment with adjunctive cariprazine 1 to 2 mg/day, 2 to 4.5 mg/day, or placebo. Patients in the 2 to 4.5 mg group demonstrated greater improvement on the Montgomery-Asberg Depression Rating Scale (MADRS).

Finally, a study in patients with bipolar I depression randomized subjects to treatment with cariprazine 0.75, 1.5, or 3.0 mg/day or placebo. The 1.5- and 3-mg/day doses were associated with greater improvement on the MADRS, but the 0.75-mg/day dose was not more effective than placebo.

Therapeutic Indications. Cariprazine is FDA approved in the United States for the treatment of adult patients with schizophrenia (1.5 to 6.0 mg/day) and acute manic or mixed episodes associated with bipolar I disorder (3.0 to 6.0 mg/day).

Precautions and Adverse Reactions

Cariprazine is contraindicated in individuals with hypersensitivity to the drug or components in the tablets. If cariprazine is prescribed with a strong CYP3A4 inhibitor, the dosage should be reduced.

Cariprazine and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis. Other precautions for all antipsychotic medications include: neuroleptic malignant syndrome; tardive dyskinesia, metabolic changes including hyperglycemia/diabetes mellitus, dyslipidemia, and weight gain. Due to cariprazine's long half-life (related to its active metabolites), patients should be monitored for adverse reactions and patient response for several weeks after starting this drug.

Metabolic Changes. In 3- and 6-week trials for acute psychosis and

mania, shifts in fasting glucose were similar for cariprazine and placebo. In long-term open-label trials, 4 percent of patients with normal hemoglobin A1c baseline values developed elevated levels. In acute studies, the shifts in fasting total cholesterol, LDL, HDL, and triglycerides were also similar in patients treated with cariprazine and placebo. There was a small amount of weight gain associated with cariprazine treatment in acute studies. For long-term, studies in schizophrenia, the mean changes in bodyweight from baseline at 12, 24, and 48 weeks were 1.2 kg, 1.7 kg, and 2.5 kg, respectively.

Extrapyramidal Symptoms and Akathisia. In clinical trials, extrapyramidal symptoms and akathisia were reported at rates ≥5 percent and at least twice the rate of placebo. Both of these effects were more common at higher doses of cariprazine.

Pregnancy. As with other antipsychotic drugs, neonates exposed to cariprazine during the third trimester of pregnancy are at risk for extrapyramidal and/or withdrawal symptoms following delivery. There are insufficient data on cariprazine to determine if it is associated with a risk for birth defects or miscarriage. There are animal data suggesting that cariprazine may cause fetal harm.

Nursing Mothers. There is insufficient information about the presence of cariprazine in human milk, the effects on the breastfed infant, or the effects on milk production.

Dosage and Administration. Cariprazine is available as 1.5-, 3-, 4.5-, and 6-mg capsules. It can be administered as a single daily dose.

For treating schizophrenia, it is recommended that patients be started on 1.5 mg for the first day. The dose can be increased in 1.5-mg increments. The recommended dose range is 1.5 to 6 mg/day. There is no indication that doses above 6 mg/day are associated with greater benefit.

For treating bipolar I mania and mixed states, it is also recommended that patients be started on 1.5 mg the first day, followed by an increase to 3 mg. The recommended dose range is 3 to 6 mg/day.

Dosing in Special Populations. The safety and effectiveness of cariprazine in pediatric populations have not been assessed. There are inadequate data regarding the safety of cariprazine in elderly patients. No dosage adjustment is recommended for patients with mild to moderate renal or hepatic disease.

BREXPIPRAZOLE

Brexpiprazole was developed as a 5-HT_{1A} and dopamine D_2 partial agonist that would have similar efficacy to its predecessor, aripiprazole and a

reduced tendency to cause akathisia, restlessness, and insomnia.

History

Brexpiprazole was discovered by Otsuka Pharmaceutical and was co-developed with Lundbeck. It received FDA approval in July, 2015 for the treatment of schizophrenia and as an adjunctive treatment for depression. In September, 2016, brexpiprazole was approved for the maintenance treatment of schizophrenia.

Pharmacological Action

Pharmacokinetics

ABSORPTION. After a single dose of brexpiprazole, peak plasma concentrations are reached within 4 hours. Steady-state concentrations are achieved within 10 to 12 days of multiple-dose administration. Levels are not affected by the ingestion of food.

METABOLISM AND ELIMINATION. Brexpiprazole is largely metabolized by CYP3A4 and CYP2D6 enzymes. Its main metabolite does not contribute to its clinical activity. After a single oral dose of brexpiprazole, approximately 25 percent of the dose is recovered in the urine and 46 percent in feces.

PHARMACODYNAMICS AND MECHANISM OF ACTION Brexpiprazole is a partial agonist at serotonin 5-HT_{1A} and dopamine D_2 receptors, and an antagonist at serotonin 5-HT_{2A} and adrenergic $\alpha_{1B/2C}$ receptors. The intrinsic activity of brexpiprazole at D_2 receptors is higher than that of pure antagonists but lower than that of aripiprazole. This limited intrinsic activity has the potential for reducing adverse reactions associated with other D_2 antagonists, such as extrapyramidal side effects and prolactin elevation. At the same time, brexpiprazole's lower intrinsic activity compared to aripiprazole may result in fewer adverse reactions related to D_2 receptor agonism such as akathisia, insomnia, restlessness, and nausea.

Design and Interpretation of Clinical Drug Studies

The efficacy of brexpiprazole in the treatment of acute exacerbations of schizophrenia was demonstrated in two pivotal 6-week, fixed-dose studies and a 6-week, flexible-dose trial with an active control, quetiapine XR.

In the fixed-dose studies, patients were randomized to brexpiprazole 2 or 4 mg once per day or placebo. One of the studies included a dose of 0.25 mg and the other 1 mg. The low doses were predicted to be ineffective and were not included in the analyses. Patients who were assigned to brexpiprazole were started on 1 mg for the first 4 days, increased to 2 mg on days 5 to 7, and then treated with 2 or 4 mg for the 5 remaining weeks.

In both of the studies, 4 mg of brexpiprazole resulted in a statistically significantly greater improvement in PANSS total score than placebo. In one of the studies, 2 mg of brexpiprazole also demonstrated a greater improvement than placebo.

In the flexible-dose study, patients were randomized to flexible doses of 2- to 4-mg brexpiprazole, 400- to 800-mg quetiapine XR, or placebo. The mean doses of brexpiprazole and quetiapine at the last study visit were 3.5 mg/day and 674.4 mg/day, respectively. Quetiapine resulted in greater improvement compared to placebo while brexpiprazole came very close to separating from placebo ($p = 0.056$).

The efficacy and tolerability of brexpiprazole as a maintenance treatment for schizophrenia were also demonstrated in a 52-week maintenance study. Patients who had been stabilized on brexpiprazole for at least 12 weeks were randomly assigned to receive brexpiprazole or placebo for 52 weeks. The primary efficacy endpoint was the time from randomization to the exacerbation of psychotic symptoms or an impending relapse. This study was terminated at an interim analysis when brexpiprazole demonstrated a significantly longer time to relapse.

The efficacy of brexpiprazole as an adjunctive treatment of MDD was evaluated in two 6-week, double-blind, placebo-controlled, fixed-dose trials. Subjects for both studies met criteria for MDD and had an inadequate response to prior antidepressant therapy and an inadequate response to 8 weeks of prospective antidepressant treatment (with escitalopram, fluoxetine, paroxetine controlled-release, sertraline, duloxetine delayed release, or venlafaxine extended-release). In one study, patients were randomly assigned to 2 mg once a day or placebo, and in the other to 1 mg, 3 mg, or placebo (added to their antidepressant). For both studies, patients received 0.5 mg once daily during week 1, 1 mg during week 2, and the assigned dose for the 4 remaining weeks. After 6 weeks of treatment, adjunctive brexpiprazole 2 or 3 mg/day provided significantly greater improvements from baseline in mean MADRS total scores.

Therapeutic Indications. Brexpiprazole is FDA approved for the treatment of schizophrenia (2 to 4 mg/day) and as an adjunctive treatment for MDD (2 mg/day).

Precautions and Adverse Reactions

Brexpiprazole is contraindicated in patients with a known hypersensitivity to brexpiprazole or any of its components.

Brexpiprazole and the other SGAs carry a black box warning for increased mortality in elderly patients with dementia-related psychosis. Other precautions for all antipsychotic medications include: neuroleptic malignant syndrome; tardive dyskinesia, metabolic changes including hyperglycemia/diabetes mellitus, dyslipidemia, and weight gain. It also

carries the same warning as other antidepressant medications for an increased suicidal risk in children, adolescents, and young adults with major depressive or other psychiatric disorders.

Metabolic Changes. In the 6-week trials for acute psychosis and depression, shifts in fasting glucose were similar for brexpiprazole and placebo. In the long-term open-label trials, 9 percent of subjects with depression and 10 percent of schizophrenia subjects experienced shifts to high fasting glucose from normal or borderline. The shifts to elevated levels of fasting total cholesterol, LDL, HDL, and triglycerides were also similar to placebo in the acute studies of schizophrenia and depression. In the 6-week trials in schizophrenia, 11 percent of patients who received 2 mg of brexpiprazole and 10 percent of those who received 4 mg had a weight gain equal to or greater than 7 percent compared to 4 percent on placebo. For the acute depression studies, 5 percent of 1 mg patients, 5 percent of 2 mg patients, and 2 percent of patients on 3 mg had a weight gain equal to or greater than 7 percent compared to 2 percent on placebo. This suggests a relatively modest risk for metabolic changes compared to other antipsychotics.

Extrapyramidal Symptoms and Akathisia. Both objective ratings and reported events suggest that brexpiprazole is associated with a small risk for extrapyramidal symptoms (EPS) and akathisia. For example, in depression trials the incidence of reported EPS, excluding akathisia, was 6 percent for brexpiprazole patients and 3 percent for placebo patients. The incidence of akathisia events for brexpiprazole patients was 9 percent versus 2 percent for placebo patients. In schizophrenia trials, the incidences of reported EPS and akathisia were similar for brexpiprazole and placebo.

Other Side Effects. Other notable adverse effects which occurred in ≥1 percent of subjects and at greater frequency than placebo included blurry vision, nausea, prolactin elevation, myalgia, somnolence, and insomnia.

Pregnancy. Neonates exposed to brexpiprazole during the third trimester of pregnancy are at risk for extrapyramidal and/or withdrawal symptoms following delivery. There are insufficient data on brexpiprazole to determine if it is associated with a risk for birth defects or miscarriage.

Nursing Mothers. There is insufficient information about the presence of brexpiprazole in human milk, the effects on the breastfed infant, or the effects on milk production.

Drug Interactions. The dosage of brexpiprazole should be reduced

when it is coadministered with strong CYP3A4 inhibitors or CYP2D6 inhibitors.

Dosage and Administration. Brexpiprazole is available as 0.25-, 0.5-, 1-, 2-, 3-, and 4-mg tablets.

The recommended dose range for schizophrenia is 2 to 4 mg administered as a single daily dose. Patients should be started on 1 mg daily for 4 days, increased to 2 mg for 3 days and then increased to 4 mg if necessary.

The recommended dose for the adjunctive treatment of depression is 2 mg daily. Patients should be started on 0.5 to 1 mg and then increased if necessary.

For both indications, the dose should be decreased 50 percent in CYP2D6 poor metabolizers.

Dosing in Special Populations. The safety and effectiveness of brexpiprazole in pediatric populations have not been assessed. There are inadequate data regarding the safety of brexpiprazole in elderly patients. Patients with moderate to severe hepatic impairment (Child-Pugh score ≥7) should be prescribed a maximum dose of 2 mg once daily for patients with MDD, and 3 mg once daily for patients with schizophrenia. For patients with moderate, severe, or end-stage renal impairment (creatinine clearance CrCl <60 mL/min), the maximum recommended dosage is 2 mg once daily for patients with MDD and 3 mg once daily for patients with schizophrenia.

FUTURE DIRECTIONS

To date, all of the available antipsychotics (except the recently approved 5-HT2A antagonist pimavanserin) target D2 receptors. It is possible that other drugs targeting this mechanism can only be marginal improvements. Drug development has also focused on agents that are effective for other domains of psychopathology in schizophrenia such as cognition or negative symptoms. These agents could be added to an antipsychotic if the drug is not effective for psychotic symptoms. Most of the focus has been on drugs that can enhance cognition. At the time of this writing, a number of drugs are in development with a number of molecular targets.

REFERENCES

*Buchanan RW, Kreyenbuhl J, Kelly DL, et al. The 2009 schizophrenia PORT psychopharmacological treatment recommendations and summary statements. *Schizophr Bull*. 2010;36:71–93.

Cohen LS, Viguera AC, McInerney KA, et al. Reproductive safety of second-generation antipsychotics: current data from the Massachusetts general hospital national pregnancy registry for atypical antipsychotics. *Am J Psychiatry*. 2016;173:263–270.

Correll CU, Skuban A, Hobart M, et al. Efficacy of brexpiprazole in patients with acute schizophrenia: review of three randomized, double-blind, placebo-controlled studies. *Schizophr Res*. 2016;174:82–92.

Correll CU, Skuban A, Ouyang J, et al. Efficacy and safety of brexpiprazole for the treatment of acute schizophrenia: a 6-week randomized, double-blind, placebo-controlled trial. *Am J Psychiatry*. 2015;172:870–880.

Davidson M, Emsley R, Kramer M, et al. Efficacy, safety and early response of paliperidone extended-release tablets (paliperidone ER): results of a 6-week, randomized, placebo-controlled study. *Schizophr Res*. 2007;93(1–3):117–130.

Durgam S, Earley W, Lipschitz A, et al. An 8-week randomized, double-blind, placebo-controlled evaluation of the safety and efficacy of cariprazine in patients with bipolar I depression. *Am J Psychiatry*. 2016;173:271–281.

Frieling H, Hillemacher T, Ziegenbein M, Neundorfer B, Bleich S. Treating dopamimetic psychosis in Parkinson's disease: Structured review and meta-analysis. *Eur Neuropsychopharmacol*. 2007;17(3):165–171.

Harrigan EP, Miceli JJ, Anziano R, Watsky E, Reeves KR. A randomized evaluation of the effects of six antipsychotic agents on QTc, in the absence and presence of metabolic inhibition. *J Clin Psychopharmacol*. 2004;24(1):62–69.

Jones PB, Barnes TR, Davies L, et al. Randomized controlled trial of the effect on Quality of Life of second- vs first-generation antipsychotic drugs in schizophrenia: Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1). *Arch Gen Psychiatry*. 2006;63(10):1079–1087.

Kahn RS, Fleischhacker WW, Boter H, et al; EUFEST study group. Effectiveness of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: an open randomised clinical trial. *Lancet*. 2008;371(9618):1085–1097.

Kane J, Canas F, Kramer M, et al. Treatment of schizophrenia with paliperidone extended-release tablets: a 6-week placebo-controlled trial. *Schizophr Res*. 2007;90(1–3):147–161.

Kane JM, Meltzer HY, Carson WH Jr, McQuade RD, Marcus RN; Aripiprazole Study Group. Aripiprazole for treatment-resistant schizophrenia: results of a multicenter, randomized, double-blind, comparison study versus perphenazine. *J Clin Psychiatry*. 2007;68(2):213–223.

Kane JM, Skuban A, Ouyang J, et al. A multicenter, randomized, double-blind, controlled phase 3 trial of fixed-dose brexpiprazole for the treatment of adults with acute schizophrenia. *Schizophr Res*. 2015;164:127–135.

Kane JM, Zukin S, Wang Y, et al. Efficacy and safety of cariprazine in acute exacerbation of schizophrenia: results from an international, phase iii clinical trial. *J Clin Psychopharmacol*. 2015;35:367–373.

Keefe RS, Bilder RM, Davis SM, et al; Neurocognitive Working Group. Neurocognitive effects of antipsychotic medications in patients with chronic schizophrenia in the CATIE Trial. *Arch Gen Psychiatry*. 2007;64(6):633–647.

Kramer M, Simpson G, Maciulis V, et al. Paliperidone extended-release tablets for prevention of symptom recurrence in patients with schizophrenia: a randomized, double-blind, placebo-controlled study. *J Clin Psychopharmacol*. 2007;27(1):6–14.

Kumra S, Kranzler H, Gerbino-Rosen G, et al. Clozapine and "high-dose" olanzapine in refractory

early-onset schizophrenia: a 12-week randomized and double-blind comparison. *Biol Psychiatry*. 2008;63(5):524–529.

Kumra S, Oberstar JV, Sikich L, et al. Efficacy and tolerability of second-generation antipsychotics in children and adolescents with schizophrenia. *Schizophr Bull*. 2008;34(1):60–71.

*Leucht S, Cipriani A, Spineli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*. 2013;382:951–962.

*Leucht S, Corves C, Arbter D, Engel RR, Li C, Davis JM. Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. *Lancet*. 2009;373(9657):31–41.

*Leucht S, Komossa K, Rummel-Kluge C, et al. A meta-analysis of head-to-head comparisons of second-generation antipsychotics in the treatment of schizophrenia. *Am J Psychiatry*. 2009;166(2):152–163.

Leucht S, Tardy M, Komossa K, et al. Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. *Lancet*. 2012;379:2063–2071.

Lewis SW, Barnes TR, Davies L, et al. Randomized controlled trial of effect of prescription of clozapine versus other second-generation antipsychotic drugs in resistant schizophrenia. *Schizophr Bull*. 2006;32(4):715–723.

Lieberman JA. Comparative effectiveness of antipsychotic drugs. A commentary on: Cost Utility of The Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1) and Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE). *Arch Gen Psychiatry*. 2006;63(10):1069–1072.

Mamo D, Graff A, Mizrahi R, Shammi CM, Romeyer F. Differential effects of aripiprazole on D(2), 5-HT(2), and 5-HT(1A) receptor occupancy in patients with schizophrenia: a triple tracer PET study. *Am J Psychiatry*. 2007;164(9):1411–1417.

Marder SR, Hakala MJ, Josiassen MK, et al. Brexpiprazole in patients with schizophrenia: overview of short- and long-term phase 3 controlled studies. *Acta Neuropsychiatr*. 2016:1–13.

McEvoy JP, Lieberman JA, Perkins DO, Hamer RM, Gu H. Efficacy and tolerability of olanzapine, quetiapine, and risperidone in the treatment of early psychosis: a randomized, double-blind 52-week comparison. *Am J Psychiatry*. 2007;164(7):1050–1060.

McEvoy JP, Byerly M, Hamer RM, et al. Effectiveness of paliperidone palmitate vs haloperidol decanoate for maintenance treatment of schizophrenia: a randomized clinical trial. *JAMA*. 2014;311:1978–1987.

McEvoy JP, Lieberman JA, Stroup TS, et al; CATIE Investigators. Effectiveness of clozapine versus olanzapine, quetiapine, and risperidone in patients with chronic schizophrenia who did not respond to prior atypical antipsychotic treatment. *Am J Psychiatry*. 2006;163(4):600–610.

Novick D, Haro JM, Suarez D, Vieta E, Naber D. Recovery in the outpatient setting: 36-month results from the Schizophrenia Outpatients Health Outcomes (SOHO) study. *Schizophr Res*. 2009;108(1–3):223–230.

Petrides G, Malur C, Braga RJ, et al. Electroconvulsive therapy augmentation in clozapine-resistant schizophrenia: a prospective, randomized study. *Am J Psychiatry*. 2015;172:52–58.

Rajkumar AP, Horsdal HT, Wimberley T, et al. Endogenous and antipsychotic-related risks for diabetes mellitus in young people with schizophrenia: a Danish population-based cohort study.

Am J Psychiatry. 2017 Jan 20:appiajp201616040442.

Ray WA, Chung CP, Murray KT, Hall K, Stein CM. Atypical antipsychotic drugs and the risk of sudden cardiac death. *N Engl J Med*. 2009;360(3):225–235.

Schooler NR, Marder SR, Chengappa KN, et al. Clozapine and risperidone in moderately refractory schizophrenia: a 6-month randomized double-blind comparison. *J Clin Psychiatry*. 2016;77:628–634.

Shim JC, Shin JG, Kelly DL, et al. Adjunctive treatment with a dopamine partial agonist, aripiprazole, for antipsychotic-induced hyperprolactinemia: a placebo-controlled trial. *Am J Psychiatry*. 2007;164(9):1404–1410.

Sikich L, Frazier JA, McClellan J, et al. Double-blind comparison of first- and second-generation antipsychotics in early-onset schizophrenia and schizo-affective disorder: findings from the treatment of early-onset schizophrenia spectrum disorders (TEOSS) study. *Am J Psychiatry*. 2008;165(11):1420–1431.

Stroup TS, Gerhard T, Crystal S, Huang C, Olfson M. Comparative effectiveness of clozapine and standard antipsychotic treatment in adults with schizophrenia. *Am J Psychiatry*. 2016;173:166–173.

Stroup TS, Lieberman JA, McEvoy JP, et al. Results of phase 3 of the CATIE schizophrenia trial. *Schizophr Res*. 2009;107(1):1–12.

Tandon R, Belmaker RH, Gattaz WF, et al. World Psychiatric Association Pharmacopsychiatry Section statement on comparative effectiveness of antipsychotics in the treatment of schizophrenia. *Schizophr Res*. 2008;100(1–3):20.1–6

*Tiihonen J, Lönnqvist J, Wahlbeck K, et al. 11-year follow-up of mortality in patients with schizophrenia: a population-based cohort study (FIN11 study). *Lancet*. 2009;374:620–627.

---

*The views expressed in this chapter do not necessarily represent the views of the FDA or the United States.

▲ 34.29 Central Nervous System Sympathomimetic Compounds

MANPREET KAUR SINGH, M.D., AND YURIY DOBRY, M.D.

INTRODUCTION

Sympathomimetic amines are compounds that activate the sympathetic central nervous system (CNS). Endogenous catecholamines, such as norepinephrine and dopamine, are direct sympathomimetic amines and exert their effects by binding to dopamine (DA) and norepinephrine (NE) receptors. Psychostimulants are indirect sympathomimetic amines with both central and peripheral effects. Psychostimulants enhance motivation, attention, concentration, wakefulness, and stimulate the cardiovascular system. They also suppress appetite and interfere with sleep. Psychostimulants achieve these effects by inhibiting the release or slowing the uptake of catecholamines, which leads to build up of DA and NE in the extracellular space.

Amphetamine (AMPH) was the first formulation of stimulants to be synthesized in 1927. Stimulants have been historically used for the treatment of "hyperkinetic" behavior in children, parkinsonism, epilepsy, and for some time, barbiturate poisoning. The second formulation of stimulants, methylphenidate (MPH), was synthesized in 1944, but was not identified as a stimulant until 1954. It was first used to treat barbiturate-induced coma and memory deficits in the elderly. Beginning in the 1960s, methylphenidate was used to treat children with ADHD, known at the time as "hyperactivity or minimal brain dysfunction (MBD)." Together, the AMPH- and MPH-based formulations constitute the psychostimulant medication class, which is currently approved by the United States Food and Drug Administration (US FDA) for the treatment of ADHD and narcolepsy for all age groups. Despite their clear therapeutic benefit, psychostimulants may be highly addictive, and are classified as, schedule II medications by the Drug Enforcement Agency (DEA). Although the immediate-release AMPH and MPH are prototypic psychostimulants, intermediate and long-acting formulations are more frequently prescribed due to the convenience of increased adherence and improved pharmacological coverage. See Table 34.29–1 for a complete list of available psychostimulants, strengths, and dosing recommendations for ADHD.

ETIOLOGY AND PATHOGENESIS OF ATTENTION DEFICIT/HYPERACTIVITY DISORDER

Research in the pathogenesis and etiology of ADHD may inform the pharmacological properties of psychostimulants. ADHD is a heritable,

chronic, neurobehavioral disorder that is commonly characterized by hyperactivity, impulsivity, and inattention. Clinical characteristics of ADHD correspond with aberrant catecholamine neurotransmission in the prefrontal cortex (PFC) and functionally associated cortical and subcortical areas. To sustain attention, regulate impulse control, facilitate working memory and cognitive flexibility, the PFC exerts a "top-down" control of the temporal cortex, parietal cortex, basal ganglia, amygdala, cerebellum, and corpus callosum. Dysfunction of catecholamine neurotransmission in the PFC compromises these cognitive functions, and can lead to clinical manifestations of ADHD.

AMPH and MPH change the dynamics of catecholamine transport system, which allows for an increase of NE and DA levels in the PFC area, and improves cognitive deficits associated with ADHD. Using animal models, and microdialysis techniques, Volkow and colleagues (2011) demonstrated that psychostimulants work by correcting the catecholamine dysfunction in the PFC, and a related "attention" network. They have shown that MPH can increase DA levels in the ventral striatum, the PFC, and temporal cortex; dysfunction which is implicated in ADHD. Further, MPH increases catecholamine levels in the fronto-striatal-thalamo-cerebellar circuit implicated in executive functions (EFs), and is positively associated with improvement in working memory, reaction time, and focus.

Table 34.29–1.
Clinical Indications of Psychostimulants

Medication	Dosing	Supplied	Average Therapeutic Dose	Maximum Daily Dose	US I Indi
Generic name: Mixed amphetamine and dextro-amphetamine salts **Brand name:** Adderall **Formulation:** Immediate release tablet	**Adults** *ADHD:* Initial: 5 mg once or twice daily Increase daily dose by 5 mg weekly intervals Narcolepsy: Initial: 10 mg daily Increase daily dose in 10 mg weekly intervals **Pediatric** Not FDA approved for children ≤3 years of age *ADHD:* **3–5 yo** Initial 2.5 mg once every morning Increase daily dose	5, 7.5, 10, 20, or 40 mg	**Narcolepsy:** 5–60 mg/day in 1–3 divided doses 2.5–20 mg b.i.d., daily	**Adults** *ADHD* 40 mg *Narcolepsy:* 60 mg **Pediatric** **3–17 yo** *ADHD* 40 mg *Narcolepsy* 60 mg **Not FDA approved for children Ä3 years of age**	Mor fo N

	in 2.5 mg weekly intervals **6–12 yo** Initial: 5 mg once or twice daily Increase daily dose in 5 mg weekly intervals **Adolescents 13–17 yo** Initial: 5 mg once or twice daily Increase daily dose in 5 mg weekly intervals *Narcolepsy:* **6–12 years:** Initial: 5 mg daily Increase daily dose in 5 mg weekly intervals **Children >12 years: Refer to adult dosing.** **Children: <3 years not recommended**				
Generic name: Dextroamphetamine sulfate **Brand name:** Dexedrine, ProCentra, Zenzedi **Formulation: immediate release tablet or extended release capsule**	**Immediate release** **Adults** **Narcolepsy:** Initial: 10 mg once daily Increase in increments of 10 mg at weekly intervals until optimal response is obtained **Pediatric** ADHD: **Children 3 to 5 years:** **Initial:** 2.5 once daily **Increase** daily 2.5 mg dose at weekly intervals *(Current AAP guidelines do not recommend use in children ≤5 years due to insufficient evidence)	5-, 10-, 15-mg capsules	**Adults, children 12 and older:** Narcolepsy: 5 to 60 mg in divided doses	**Immediate Release** **Adults** **ADHD:** 40 mg **Narcolepsy:** 60 mg **Pediatric** **ADHD:** 40 mg **Narcolepsy:** 60 mg **Extended Release** **Pediatric** **ADHD:**40 mg (60 mg in divided doses can be tried in children >50 kg)	Mor A N

	Children ≥6 years and adolescents: Initial: 5 mg once or twice daily; may increase in increments of 5 mg at weekly intervals **Narcolepsy** Children 6 to 12 years: Initial: 5 mg once daily Increase in increments of 5 mg at weekly intervals **Children >12 years and adolescents** Refer to adult dosing. **Extended-release capsules:** Initial: 5 mg once or twice daily Increase in increments of 5 mg weekly				
Generic name: Mixed amphetamine and dextro-amphetamine salts **Brand name:** Adderall XR **Formulation:** Extended-release capsule	**Adults-** *ADHD:* Initial: 20 mg once daily in the morning Not FDA approved for children ≤6 years of age **Pediatric** *ADHD:* Initial: 5–10 mg once daily in the morning Increase daily dose in 5–10 mg increments at weekly intervals	5-, 10-, 15-, 20-, 25- or 30-mg capsule	20–30 mg	30 mg **Children: <6 years not recommended**	Mor A
Generic name: Lisdexamfetamine **Brand name:** Vyvanse **Formulation:** Capsule with inactive pro-drug compound	**Adults** **ADHD:** Initial: 30 mg once daily in the morning Increase in increments of 10 or 20 mg daily at weekly intervals	20-, 30-, 40-, 50-, 70-mg capsule	**ADHD:** 20–70 mg daily **Binge Eating Disorder:** 50–70 mg daily	**Adults** *ADHD* and *BED* 70 mg **Pediatric** *ADHD* 60 mg **Not FDA approved for children Ä6**	Mor A B

	Binge eating disorder (BED)**:** **Pediatric** *ADHD:* Children 6–12 years: Initial: 20 to 30 mg once daily in the morning Increasing: weekly increases of 10 mg daily up to a maximum of 60 mg daily Adolescents: Refer to adult dosing			**years of age**	
Generic: Methylphenidate hydrochloride **Brand Name:** Ritalin **Formulation:** Immediate release tablet	**Adults** **ADHD:** Initial: 5 mg twice daily Increase by 5–10 mg daily at weekly intervals Narcolepsy: Initial: 5 mg twice daily Increase by 5–10 mg daily at weekly intervals Pediatric Not FDA approved for children ≤6 years of age Children ≥6 years and Adolescents: Initial: 5 mg twice daily Increase by 5–10 mg daily at weekly intervals **Off-label use:** Depression in medically ill older adults, adult patients with terminal illness, receiving palliative care: 2.5–5 mg once daily or twice daily Increase by 2.5–5 mg daily every 1–3 days in divided doses	5, 10, 20 mg	20–30 mg Some patients might require 40–60 mg	**Adults** **ADHD** and **Narcolepsy** 60 mg daily (in 2–3 divided doses) **Depression in medically ill older adults, adult patients with terminal illness, receiving palliative care:** 20–40 mg daily **Pediatric** ADHD: 60 mg daily (in 2–3 divided doses)	Mor A N

Generic: Methylphenidate hydrochloride **Brand Name:** Methylin **Formulation:** Immediate release oral solution	Adults Children (6 years and over) Initial: 5 mg twice daily Increase by 5 mg daily at weekly intervals	Methylin Oral Solution 5 mg/5 mL Methylin Oral Solution 10 mg/5 mL	20–30 mg daily	60 mg Not FDA approved for children ≤6 years of age	Mor A N
Generic Methylphenidate hydrochloride **Brand Name:** Ritalin SR **Formulation:** Immediate release tablet or extended release capsule	**Adults and Pediatric** ADHD and Narcolepsy: Equivalent to Ritalin immediate release	20 mg	20–60 mg	60 mg	Mor A N
Generic: Methylphenidate hydrochloride **Brand Name:** Concerta **Formulation:** Extended-release tablet	**Children 6–12 yo, adolescents 12–17 yo, Adults** ADHD: Initial: 18 mg daily (18–36 mg adults) Increased by 18 mg/day at weekly intervals	18, 27, 45, 54 mg	**Adults** 18–72 mg daily **Pediatric** 18–54 mg daily	54 mg (6–12 yo) 72 mg (adolescents and adults) **Not FDA approved for children Ä6 years of age or elderly**	Mor A
Generic: Methylphenidate hydrochloride **Brand Name:** Daytrana **Formulation:** Extended-release transdermal system	**6–12 yo, adolescents-** ADHD: **week 1:** 10 mg/9 hours (1.1 mg/hour) **week 2:** (1.6 mg/hour), 15 mg/9 hours **week 3:** (2.2 mg/hour), 20 mg/9 hours **week 4–7:**30 mg/9 hours (3.3 mg/hour)	Transdermal patch: 10 mg/9 hours (1.1 mg/hour) 15 mg/9 hours (1.6 mg/hour) 20 mg/9 hours (2.2 mg/hour) 30 mg/9 hours (3.3 mg/hour)	1.1–3.3 mg/hour	3.3 mg/hour **Not FDA approved for children Ä6 years of age or elderly** **No indication in adults**	Mor A
Generic: Methylphenidate hydrochloride **Brand name:** Metadate CD **Formulation:** Expended-release capsule	Children 6–12 yo, adolescents 12–17 yo, adults- ADHD: Initial: 20 mg daily; Increase in 10–20 mg increments at weekly intervals	10, 20, 30, 50, mg	60 mg	**Children >6, adolescents, adults** **ADHD: 60 mg**	Mor A N
Generic: Methylphenidate hydrochloride **Brand name:** Metadate ER	**Adults and Pediatric** ***ADHD and Narcolepsy:*** Equivalent to	20 mg	10–60 mg	60 mg	Mor A N

Formulation: Extended-release tablets	Ritalin immediate release				
Generic: Methylphenidate hydrochloride **Brand Name:** Quillivant XR **Formulation:** Extended-release powder for suspension	Not approved by FDA for children <6 years of age Children 6–12 years, adolescents, and adults: ADHD: Initial: 20 mg daily; Increase in 10–20 mg increments at weekly intervals	Bottles of 300-, 600-, 750-, 900-mg powder (5 mg/mL)	60 mg	60	Mor A N
Generic: Dexmethylphenidate **Brand Names:** Focalin and Focalin XR **Formulation:** Immediate release tablet and extended release capsule	Adults: ADHD: Immediate release: Initial: 2.5 mg twice daily Increase in increments of 2.5–5 mg at weekly intervals Extended release: Initial: 10 mg once daily Increase in increments of 10 mg/day at weekly intervals (maximum dose: 40 mg/day) Pediatric ADHD: Children ≥6 years: Initial: 2.5 mg twice daily Increased in increments of 2.5–5 mg at weekly intervals Initial: 5 mg once daily; dosage may be adjusted in increments of 5 mg/day at weekly intervals	Focalin: 2.5, 5, 10 mg Focalin XR: 5, 10, 15, 20, 30, 35, 40 mg	Focalin: 2.5 to 10 mg twice a day (doses should be taken at least 4 hours apart) Focalin XR 20–40 mg daily (adults) and 10–30 mg daily (pediatric)	**Focalin immediate release** 20 mg/day **Focalin extended release** 30 (pediatric) and 40 (adults)	Mor A
Generic: Atomoxetine HCl **Brand name:** Strattera **Formulation:** Immediate release capsules	Adults ADHD Initial: 40 mg/day Increase to ~80 mg/day after minimum of 3 days. May	10, 18, 25, 40, 60, 80, 100 mg	Children and adolescents ≥70 kg: 100 mg Children and adolescents ≤70 kg: 1.2	1.4 mg/kg or 100 mg	Mor A

	increase to 100 mg/day in 2–4 additional weeks to achieve optimal response. Pediatric ADHD: *Children ≥6 years and ≤70 kg* Initial 0.5 mg/kg/day Increase to ~1.2 mg/kg/day after a minimum of 3 days		mg/kg	
Generic Modafinil **Brand name:** Provigil **Formulation:** Immediate release tablet	**Narcolepsy, obstructive sleep apnea/hypopnea syndrome (OSAHS):** Initial: 200 mg daily in 2 divided doses Increase: in 100-mg increments as needed and tolerated **Shift work sleep disorder (SWSD):** Initial: 200 mg as a single dose taken ~1 hour prior to start of work shift	Supplied: 100-, 200-mg tabs	Narcolepsy 200–300 mg/day **SWSD: 200** mg/day	400 mg

Neurobiology of ADHD

Neuroimaging studies have now demonstrated aberrant brain structural and functional characteristics associated with key cognitive deficits in ADHD. Compared to healthy controls, individuals with ADHD across the lifespan have decreased volumes in frontal, parietal, and temporal cortices, as well as in the basal ganglia, limbic system, hypothalamus, cerebellum, and corpus callosum. Impairments in these frontal-striatal-cerebellar networks that are known to regulate attention, impulse control, and other key EFs may lead to deficits in these cognitive functions in ADHD. Importantly, more granular findings have been demonstrated in key functional subregions of the PFC, including the orbitofrontal cortex (OFC; control and processing of emotion and motivation), the inferior frontal cortex (IFC; regulates inhibition impulsive thoughts or feelings, and maintains cognitive flexibility essential for decision making), dorsolateral prefrontal cortex (DLPFC; key for sustained attention, planning and

working memory), dorsal anterior cingulate cortex (dACC; integrates information and controls executive and inhibitory functions; dACC facilitates sensory target detection, proper response selection, error detection, and action monitoring), and the anterior PFC (controls impulsivity of the motor behavior). Collectively, these regions function to filter irrelevant information or "noise" and amplify relevant signal needed to execute a specific task.

Biochemistry of Catecholamine Neurotransmission

ADHD-related dysfunction may be related to disruption of PFC fibers that project catecholamines to cortical and subcortical regions. This model of ADHD has also been supported by animal and human lesion and neuroimaging (diffusion tensor imaging) studies that show that disruption of key PFC projections result in attention, emotion, and locomotor dysregulation. Further, mutating genes that encode for the dopamine transporter protein (DAT) and enzymes that metabolize DA and NE, have conferred behavioral traits in animals that mimic ADHD-like traits. Indeed, both DA and NE play an important role in the pathophysiology of ADHD, and are candidate targets for treatment with psychostimulants. DA enhances working memory through its interaction with D1 receptors. NE functions to increase arousal, excitability, and memory storage and retrieval, and to modulate executive function, including cognitive organization and problem solving through its interactions of the postsynaptic α2A-adrenoceptors. Thus, deficits in this catecholamine system within the PFC have been associated with decreased control over irrelevant information, leading to inattention, impulsivity, and hyperactivity associated with ADHD. As above, psychostimulants are thought to inhibit or reverse the direction of the NE/DA membrane transporter proteins (NET/DAT), which increase catecholamines in the synaptic space, leading to enhanced function in the PFC to restore attention.

Genetics

Family, adoption, and twin studies suggest that the mean heritability estimate of ADHD is ~76 percent in children and adolescents, and ~30 percent in adults. Gene variants associated with the ADHD phenotype include mutations of genes that encode for the proteins involved in the catecholamine neurotransmission system. Unfortunately, the strength of the association between ADHD and likelihood of carrying mutations of most candidate genes that have been identified so far has not been adequately validated.

MECHANISMS OF ACTION OF PSYCHOSTIMULANTS

As above, psychostimulants exert clinical effects by raising extracellular catecholamine levels in the context of catecholamine deficit. The chemical compositions of psychostimulants AMPH and MPH are different. AMPH belongs to the phenethylamine chemical class. The asymmetric α-carbon atom gives rise to two AMPH enantiomers, the L-stereoisomer and the D-stereoisomer. MPH has four isomers but only L-MPH and D-MPH are part of the currently available MPH preparations, and only D threomethylphenidate exhibits clinically useful pharmacological activity.

Just as chemical compositions vary, the mechanisms of action of various psychostimulants also differ significantly. Whereas MPH blocks the DAT and NET, and increases DA and NE from the synaptic vesicle pool by inhibiting endocytosis, AMPH compounds use DAT and NET, along with several other transporters, to actively unload vesicular catecholamine stores. Using PET and radioligands specific to the DAT, NAT, and DA and NE receptors, D-methylphenidate has been shown to rapidly cross the blood–brain barrier (BBB), and bind to the DAT and NET. To reach efficacy threshold, MPH needs to occupy approximately 70 to 80 percent of NET and 60 to 70 percent of DAT, in order to increase synaptic DA in the PFC leading to increased attention. The rate of extracellular DA rise is proportional to the rate at which methylphenidate crosses the BBB.

In contrast to these MPH mechanisms, AMPH functions as a potent full agonist of trace amine-associated receptor 1 (TAAR1), which inhibits vesicular monoamine transporter function (VMAT), DAT (especially D1, the predominant receptor type in the PFC), and NET. AMPH also functions by reducing the DA and NE vesicular pH gradient, leading to reversal of catecholamine transport and inhibition of the monoamine oxidase (MAO) and catechol-o-methyltransferase (COMT) enzymes that degrade catecholamines. This confers more potency in AMPH compared to MPH, which translates to more rapid extracellular increases in DA and NE concentrations compared to MPH. Indeed, microdialysis studies in rat brains have demonstrated that AMPH induces three times greater behavioral effects and greater synaptic DA and NE boost compared to the catecholamine rise caused by much higher doses of MPH. Thus, AMPH may be a more effective indirect sympathomimetic agent than MPH, and may require lower effective doses to induce MPH-equivalent physiological effects.

CLINICAL PHARMACOLOGY OF PSYCHOSTIMULANTS

Psychostimulants, amphetamine and methylphenidate compounds, are the most commonly used psychotropic medications nationwide, and have US FDA–indicated treatment for ADHD and narcolepsy in children, adolescents, and adults. Please refer to Table 34.29–1 for a list of clinical indications of psychostimulants. This section will review the pharmacokinetics (see Table 34.29–2), pharmacodynamics, interactions,

efficacy, and tolerability (including adverse effects, contraindications, and precautions) of psychostimulants, and then conclude with an overview on these characteristics for nonstimulant medications, modafinil and atomoxetine (ATM).

Pharmacokinetics

Immediate-Release Formulations. Due to their high degree of lipophilicity, psychostimulants rapidly cross the gastrointestinal tract and the BBB. The absorption phase parallels the rate of release of the catecholamine neurotransmitters into the CNS synaptic clefts.

AMPH-related compounds typically reach peak blood levels in 2 to 3 hours, with a variable effect of several hours after oral administration. Amphetamine is metabolized by deamination, oxidation, and hydroxylation. Deamination produces an inactive metabolite, phenylacetone, which is further oxidized and then excreted in urine as hippuric acid and glucuronide conjugates. Amphetamine is also converted to norephedrine by oxidation and then this metabolite and the parent compound are p-hydroxylated. Several metabolites, including norephedrine, its hydroxymetabolite, and hydroxyamphetamine, are all pharmacologically active. The excretion of amphetamine depends on urinary pH, which provides a basis for using ammonium chloride to acidify the urine to manage an amphetamine overdose, which markedly increases excretion. Approximately 70 percent of the dose is excreted in a 24-hour urine collection with 30 percent excreted as unchanged drug. This increases to 74 percent under acidic conditions and decreases to 1 percent in alkaline urine.

The MPH tablet releases equal amounts of D-MPH and L-MPH, at a nearly identical, linear rate. The behavioral effects of MPH are exerted by its d-MPH enantiomer whereas L-MPH has little or no behavioral effect. MPH and related compounds reach peak blood concentration in 2 hours and have a short half-life of 1 to 2 hours. Concentrations of MPH in the brain exceed concentrations in the blood. Metabolism of MPH is highly enantio-specific, favoring L-MPH over D-MPH. Plasma half-life (t1/2) of D-MPH is markedly longer than that of L-MPH, and oral bioavailability is 22 percent for D-MPH and 2 to 8 percent for L-MPH. MPH undergoes hydrolysis an deesterification by a liver and stomach enzyme, CES-1. Another route of metabolism is hepatic deesterification and oxidation. Ritalinic acid is the major metabolite of MPH that is renally excreted.

Long-Acting Formulations. Methylphenidate and amphetamine long-acting products extend the apparent plasma half-life of methylphenidate and amphetamine by altering the absorption kinetics. Methylphenidate sustained release (Ritalin SR) was the first available long-acting formulation of methylphenidate. It delivers a single, prolonged pulse of

MPH via a wax-matrix–based technology. This formulation has a problematic delay in release rate, onset of behavioral effects, significant intrasubject variability and has been shown to have inferior clinical properties compared to MPH administered several times per day. Due to the availability of new multiple long-lasting release formulations of MPH, Ritalin-SR is rarely used. Controlled-release formulations of methylphenidate include Metadate CD, Focalin XR, and Ritalin LA. Metadate CD is the first extended-release formulation to use beaded technology. The MPH compound is stored in porous beads, with pore diameter controlling its release. The porous beads produce a bimodal MPH distribution, with an immediate burst of a D- and L-methylphenidate mixture, and an extended release of the remaining MPH approximately 5 hours later. The Focalin XR formulation also demonstrates similar bimodal kinetics. It provides 8-hour coverage without significant troughs. Focalin XR releases D-threo-methylphenidate, which is approximately twice as potent as L-threo-methylphenidate.

Mixed salts of a single-entity amphetamine product (Adderall XR) contain two types of beads in a 50:50 ratio. The immediate-release beads create an almost immediate burst of AMPH, and the delayed-release beads provide the second burst approximately 4 hours after oral ingestion. Extended-release methylphenidate, Ritalin LA, uses a spheroidal oral drug absorption system (SODAS) technology to evenly distribute MPH over approximately 8 hours. It has a bimodal distribution of a mixture of D- and L-threo-methylphenidate. Extended-release methylphenidate (Concerta) uses an oral osmotic release system (OROS), which relies on an osmotic gradient to provide extended coverage. The capsule is composed of a medication layer next to a layer of inactive matrix material. A third layer contains active osmotic components, which regulate the rate of release. In an aqueous environment such as gastrointestinal tract, the matrix layer dissolves within 1 hour, providing an initial dose of methylphenidate. Water expands the osmotic layer, and then the remaining amount of MPH is released, controlled by the diameter of the osmotic layer orifices.

Table 34.29–2.
Pharmacological Properties of Psychostimulants

Medication	T_{max} (h)	Half-life (hours)	Duration of Action (hours)	C_{max} (ng/mL)
Mixed (R,S)-amphetamine salts, short acting (Adderall) Racemic mixture of D:L-amphetamine isomers in a 3:1 ratio Each Adderall XR capsule	2	For D-Isomer ~11–13 L-Isomer ~15–17 (if urine is acidic-> both	6	110

contains a 1:1 ratio of immediate- to delayed-release mixed amphetamine salts beads.		isomers' half-lives are reduced to 7 hours)		
Mixed (R,S)-amphetamine salts, extended release (Adderall XR) Racemic mixture: D:L amphetamine isomers in a 3:1 ratio Each Adderall XR capsule contains a 1:1 ratio of immediate- to delayed-release mixed amphetamine salts beads.	Children 6–12 years: D-amphetamine: 9 hours L-amphetamine: 11 hours Adolescents 13–17 years: D-amphetamine: 11 hours L-amphetamine: 13–14 hours Adults D-amphetamine: 10 hours L-amphetamine: 13 hours An acidic diet: 8–11 hours An alkaline diet 16–31 hours	Mean 11.5	10–12	40.10
Dextroamphetamine sulfate extended release (Dexedrine)	First peak: 1.6 hours Second peak: 6.5 hours	9.4–13	10	36.6
Dextroamphetamine sulfate **immediate release**	1–2	9.4–13	10	36.6
Lisdexamfetamine dimesylate(Vyvanse)	3.8–4.7	1	10–12	58.2 +/– 28.1 ng/mL
Methylphenidate hydrochloride (Ritalin)	2	2–3	4	3.5 +/– 0.4 ng/mL
Methylphenidate sustained release (Ritalin SR)	3–8	Unknown	~8	4.84
Methylphenidate Extended-Release Oral Suspension (MEROS) Long-acting liquid preparation with MEROS delivery system. MEROS consists of drug–polymer complexes 80% of drug–polymer complexes are coated with an extended-release coating of varying thicknesses, which allows for release of methylphenidate throughout the day 20% of drug–polymer complexes are uncoated and	First peak ~2 Second peak 5	5–6	10–12	3.6 (± 5.8) ng/mL

act as immediate-release methylphenidate				
Methylphenidate hydrochloride oral solution (Methylene)	2	2–3	4	3.5 +/– 0.4 ng/mL
Extended-release methylphenidate hydrochloride (Metadate CD) Contains two types of coated beads, one type with immediate-release *methylphenidate* and second with extended-release *methylphenidate*. 30% of beads deliver d-threo- and l-threo-methylphenidate immediately The remaining 70% of beads are extended release, delivering the mixture of L-threo-methylphenidate content and D-threo-methylphenidate content several hours after the beads with methylphenidate have been released.	First peak ~2 Second peak 4–6	6.8	8–9	24.7–40.9 ng/mL
D-Threo-Methylphenidate (Focalin)	2	2–3	4	3.5 +/– 0.4 ng/mL
Dextra-methylphenidate extended release (Focalin XR) Formulation provides coverage of up to 12 hours. d- Focalin XR contains only the *d*-MPH which is twice as potent as Focalin immediate release. It delivers 50% of its d-threo-MPH initially, and 3 to 4 hours later. This creates an approximate 8–10 coverage without significant troughs. Thus, the closest concentration curve can be recreated by methylphenidate twice daily (50% of the dose).	First peak 1–2 Second peak 5–6	3.5	8–10	14–16
Methylphenidate transdermal system (Daytrana) Delivers methylphenidate into the systemic circulation directly through the skin,	10 hours after single application 8 hours after repeat patch	3–4	9	5.7 ± 9.39:12.5 CM^2– 542.9 ± 22.4:37.5

reducing first-pass metabolism by the liver.	applications			
Extended-release methylphenidate (Ritalin LA) Delivers MPH via spheroidal oral drug absorption system (SODAS) technology to produce a bimodal distribution of 1:1 racemic mixture of d-and l-threo-MPH enantiomers. 50% is released initially and another 50% approximately 3–4 hours later	First peak 1–2 Second peak 5–6	3.5	6–8	16.4
Extended-release methylphenidate (Concerta) It delivers methylphenidate via oral osmotic release system (OROS), which is composed of two drug layers containing the drug and excipients, and a push layer with active osmotic components. It delivers 1:1 racemic mixture of l:d-threo-methylphenidate. In the gastrointestinal tract, the drug overcoat dissolves within 1 hour, providing an initial dose or 22% of the medication. Water then permeates the tablet core, osmotically active polymer, releasing the remaining 78% of the mediation through the orifice	8–12	3.5	9–11	3.3–7.8
Extended-delivery methylphenidate (MEROS or Quillivant XR) is a powder that, after reconstitution with water, forms an extended-release oral suspension formulation of methylphenidate intended for once daily oral administration. It contains approximately 20% immediate-release and 80% extended-release methylphenidate	5.0	5.2		13.6 (± 5.8) ng/mL
Modafinil (Provigil)	2–3	D-modafinil: 4–5 hours	6–14	6.4 µg/mL

		L-modafinil: 13–16 hours L-modafinil constitutes 90% of C_{max}		
Atomaxetine (Strattera)	1–2	5.2 (extensive metabolizers) 21.6 (poor metabolizers)	24	584 ± 268

C_{max}, maximum plasma methylphenidate concentration
T_{max}, time to maximum plasma methylphenidate concentration

The methylphenidate transdermal system (Daytrana) delivers methylphenidate through a transdermal matrix. It does not undergo first-pass metabolism by CES-1 in the liver and stomach, and has a relatively constant release of both D- and L-methylphenidate throughout the entire patch over time, resulting in constant plasma methylphenidate levels. Daytrana may be useful for patients who have problems tolerating oral medications. In addition, extended-delivery methylphenidate (MEROS or Quillivant XR) is a powder formulation of MPH that forms an extended-release oral suspension for once per day oral administration. It contains approximately 20 percent immediate-release and 80 percent extended-release methylphenidate. This formulation might be advantageous for patients who cannot tolerate pill and prefer a liquid form. Finally, lisdexamfetamine dimesylate (LDX; Vyvanse) is the first known inactive pro-drug psychostimulant that is activated by a chemical modification in erythrocytes. LDX has high aqueous solubility, partially due to low lipophilicity. It is likely that in the small intestines LDX is transported by peptide transport proteins such as PEPT1 and PEPT2. Unlike other amphetamines, it consistently increases the extracellular concentration of serotonin in the PFC and striatum.

Pharmacodynamics

Traditional stimulants are all considered indirect-acting sympathomimetics because they do not act directly on receptors, as do natural catecholamines. Stimulant action occurs through simultaneous neurotransmitter release at the synapse and block of reuptake, thereby slowing metabolic breakdown of active neurotransmitters. Dextroamphetamine is the most potent by virtue of its ability to release both DA and NE and its capacity to release newly formed soluble pools and granule-stored pools of catecholamines in nerve endings. MPH acts on DA (but not NE) through release of DA from granular pools and blockage of DA reuptake. This may explain why

amphetamine is a more powerful stimulant and is more likely to activate latent hypertension than is methylphenidate. See Table 34.29–3 for important psychostimulant drug–drug interactions.

Table 34.29–3.
Important Psychostimulant Drug–Drug Interactions

Methylphenidate	**Inhibits metabolism of:**
	Anticoagulants (e.g., Coumadin)
	Tricyclic drugs (e.g., imipramine, clomipramine, desipramine)
	Anticonvulsants (e.g., phenobarbital, phenytoin, primidone)
	CYP2D6 substrates (e.g., flecainide, haloperidol, metoprolol, propafenone, risperidone, thioridazine SSRIs, SNRIs)
	Cyclosporine
	Glipizide
	Induces metabolism of:
	Methylphenidate
	Metabolism is inhibited by:
	Bupropion
	CYP2D6 inhibitors (e.g., chlorpromazine, delaviridine, fluoxetine, miconazole, paroxetine, pergolide, quinidine, quinine, ritonavir, and ropinirole)
	Metabolism is induced by: Carbamazepine
Amphetamines	**Acetazolamide**
	Acetazolamide decreases clearance of amphetamine
	Chlorpromazine
	Decreased amphetamine effectiveness
	Donepezil
	Increased norepinephrine availability lowers seizure threshold
	Guanethidine
	Decreases guanethidine effectiveness
	Opipramol, **Pargyline**, **Phenelzine**, **Procarbazine**, **Protriptyline**, **Tranylcypromine, Doxepin, Trimipramine, Toloxatone, Rasagiline**, **Selegiline**, **Sibutramine, Amitriptyline:** Increased norepinephrine availability can result in a hypertensive crisis (headache, hyperpyrexia, hypertension)
	The concurrent use of amphetamine and monoamine oxidase inhibitors (MAOIs) is contraindicated. At least 14 days should elapse following the discontinuation of MAOIs before amphetamine therapy is instituted
	Desvenlafaxine, Dolasetron, Lofepramine, Lorcaserin, Hydroxytryptophan, Imipramine, Iproniazid, Almotriptan, Clorgyline, Desipramine, Isocarboxazid, Meperidine, Moclobemide, Nialamide, Nortriptyline Tramadol,Trazodone, Vilazodone, Vortioxetine, Ziprasidone, Fentanyl, Fluoxetine, Furazolidone, Mirtazapine, Granisetron, Levomilnacipran, Palonosetron: increased risk of serotonin syndrome from additive serotonergic effects

Efficacy and Tolerability of Stimulants in the Treatment of ADHD and Other Conditions

Several randomized controlled trials (RCTs) have characterized the efficacy, tolerability, and adverse effects of stimulants for ADHD and for

other psychiatric disorders. In total, approximately 5,899 children, adolescents, and adults have been involved in RCTs examining the efficacy and safety of stimulants in treatment of ADHD. Outcomes of these studies pooled together have demonstrated that 65 to 75 percent of individuals in the treatment arm improved in core ADHD symptoms (inattentiveness, hyperactivity, impulsivity), compared to 5 to 30 percent of individuals who improved when assigned to the placebo arm, with large effect sizes (Cohen's d = 0.7 to 1.5). Psychostimulants are considered highly effective and relatively safe for the treatment of ADHD in all age groups. Based on such robust evidence, the American Academy of Pediatrics and the American Academy of Child and Adolescent Psychiatry (AACAP) have adopted guidelines that regard stimulant medications as a cornerstone treatment for ADHD.

Data in Children and Adolescents with ADHD. The Collaborative Multisite Multimodal Treatment Study of Children with ADHD (MTA) is the first major RCT that demonstrated long-term efficacy and safety of methylphenidate in the treatment of children diagnosed with ADHD, combined type. MTA investigators followed 579 children, 7 to 9.9 years of age. Participants were randomized to three carefully designed treatment groups and a control group that received treatment in the community. Treatment groups included one group of participants who were treated with immediate-release methylphenidate (MPH-IR) monotherapy, a second group treated with MPH-IR in combination with behavioral and community interventions, and a third group treated with behavioral and community interventions only. Participants were assessed using parent- and teacher-reported ADHD rating scales at baseline and after 14 months of treatment. Adverse effects were also assessed and reported. MPH-IR was administered 3 times/day, with a mean total daily dose was 32.1 ± 15.4 mg. Subjects treated with MPH-IR with or without behavioral or community interventions showed significant improvement in core ADHD symptoms such as inattention, impulsivity, and hyperactivity. In addition, children randomized to the two groups with methylphenidate treatment demonstrated significant improvement in family, social, and academic functioning, as well as in externalizing symptoms. The symptomatic and functional improvements persisted for the duration of the treatment.

A 93 percent retention rate was reported 24 months after the initiation of the MTA trial when ADHD symptoms and level of social, academic, and family functioning were reassessed. Most subjects continued to benefit from methylphenidate or from the combination of methylphenidate and behavioral interventions. The difference in ADHD symptoms and functioning among subjects treated with methylphenidate, subjects receiving behavioral intervention alone, or subjects treated in a community setting decreased significantly over time. Improvement seen in the two

groups treated with methylphenidate remained, but the effect size decreased by 50 percent. Twenty-four months after initiation of treatment, 21.8 percent of youth in the combination methylphenidate and behavioral treatment group stopped treatment versus 25.8 percent of subjects in the methylphenidate-only group. In the behavioral therapy-only group, at 24-month assessment, 68.2 percent of subjects added methylphenidate to their treatment regimen.

At the end of 14 months, 77 percent children in the methylphenidate monotherapy treatment group maintained their original dose of methylphenidate. In 23 percent of subjects who did not maintain the original dose, initial titration did not lead to an adequate response, and 12 percent did not maintain satisfactory long-term therapeutic gains. For the roughly 5 percent of youth who did not respond to initial titration of methylphenidate, switching to dextroamphetamine was effective. Methylphenidate was later found to be effective in a minority of children who did not respond to the initial methylphenidate titration, reflecting a false-negative titration outcome. Thus, some proportion of ADHD patients do not respond to initial stimulant treatment. However, a significant percentage of these patients will likely respond to another stimulant. Ninety percent of all MTA subjects assigned to methylphenidate treatment (with or without behavioral intervention) responded well to either the first stimulant trial with methylphenidate (79 percent) or a second stimulant trial with dextroamphetamine (11 percent).

The Preschool ADHD Treatment Study (PATS) is a double-blind, placebo-controlled, crossover trial that examined the effects of immediate-release methylphenidate in 165 children, ages 3 to 5.5 years, diagnosed with severe ADHD. Subjects were randomly assigned to four treatment arms, each at a different methylphenidate dose (1.25, 2.5, 5, or 7.5 mg, three times daily), or placebo. Significant improvements in ADHD symptoms were found with an average methylphenidate dose of 14.2 ± 8 mg/day. This average effective dose was considerably lower than the effective dose reported in the MTA study of older children (30.5 ± 14.2 mg/day). Twenty-one percent of the total subjects in the trial were reported to be “excellent responders.” During the 10-month maintenance phase of the PATS, 95 of the 140 children demonstrated continued improvement in global functioning, including improvement in social skills. However, parent and teacher symptom rating scales declined over the same time period for about one-third of these children. PATS had a relatively lower effect size range (Cohen’s d = 0.22 to 0.87) compared to the MTA study (Cohen’s d = 0.8 to 1.2).

Several adverse effects were reported in the PATS. Approximately 30 percent of PATS subjects developed moderate to severe side effects, and up to 11 percent of the subjects discontinued treatment due to medication intolerability. Side effects included insomnia, appetite loss, emotional

outbursts, repetitive thoughts/behaviors, and irritability. Thus, based on findings from the MTA and PATS studies, it appears that compared to school-aged children, psychostimulants have lower short-term efficacy and lower tolerability in preschoolers. One-third of preschool-aged subjects discontinued stimulants due to adverse events such as worsening behaviors, while others discontinued to switch to a long-acting stimulant or discontinued due to inadequate benefit.

Data in Adults with ADHD. A Cochrane review of 7 RCTs, with 1,091 subjects combined examined the efficacy and tolerability of three different amphetamine agents (dextroamphetamine, lisdexamfetamine, and mixed amphetamine salts [MAS]) in adults with ADHD. While dexamphetamine, lisdexamfetamine, and MAS all improved core ADHD symptoms, MAS was associated with lowest incidence of adverse effects and highest retention in treatment. In other pooled analyses of 19 trials in adults, effect sizes varied based on whether the psychostimulants were long-acting (pooled Cohen's d = 0.73) or immediate-release (pooled Cohen's d = 0.96) formulations. LDX had a significantly higher pooled effect size (Cohen's d = 1.5) based on several RCTs compared to other psychostimulants.

Other Clinical Indications of Psychostimulants

Beyond their use for the treatment of ADHD, psychostimulants have been used in several conditions including oppositional defiant disorder (ODD), conduct disorder (CD), symptoms of inattention and hyperactivity associated with autism, cognitive deficits associated with traumatic brain injury (TBI), apathy, depression, and cognitive deficits associated with seizure disorder and epilepsy. Stimulant use in these other disorders has been less extensively studied compared to ADHD, and the general recommendation to providers is to administer stimulants at lower doses than typically prescribed for treatment of ADHD. The following sections consider use of psychostimulants in each of these disorders.

Oppositional Defiant Disorder and Conduct Disorder Comorbid with ADHD

Up to 70 percent of ADHD patients also carry a diagnosis of ODD or CD. In a 4-week, multicenter, randomized, double-blind, parallel-group, forced dose-escalation study, 308 children and adolescents diagnosed with ODD with and without ADHD were treated daily with 30 or 40 mg daily of MAS XR (Adderall XR). Independent of the ADHD diagnosis, treatment with Adderall XR led to a reduction of ODD symptoms measured with the SNAP-IV parent rating subscale. Subjects who received both 30-mg/day (Cohen's d = 0.52) and 40-mg/day (Cohen's d = 0.56) MAS XR showed efficacy in the medium effect size range. Adverse effects were mild to moderate in severity and included anorexia/decreased appetite (25.3

percent), insomnia (19.5 percent), headache (18.5 percent), and abdominal pain (10.7 percent).

In a similar trial of 84 children between the ages of 6 and 15 years with cooccurring CD in two-thirds of children with ADHD, participants were randomized to methylphenidate hydrochloride (average methylphenidate hydrochloride dose was 41.3 mg/day) or placebo for 5 weeks. Behavior was evaluated by parent, teacher, and clinician reports and by direct classroom observations. Compared with the placebo group, methylphenidate showed significant benefit for ADHD symptoms, and CD symptoms including cheating and stealing, obscene language, attacks on others, vandalism of property, and deliberate cruelty.

Treatment of Neurodevelopmental Disorders with ADHD Symptoms. ADHD symptoms occur in 41 to 78 percent of children with autism spectrum disorders (ASDs). These symptoms often affect quality of life, interfering with learning or interventions that target primary ASD symptoms. Children who have developmental disabilities (including ASD, intellectual disabilities, Fragile X syndrome), have lower effect sizes with these medications and are more sensitive to adverse effects, including agitation and emotional lability, compared with typically developing children with ADHD. One large RCT of methylphenidate in children with ASD found that only 49 percent of children responded to stimulants compared to the 69 percent response rate in the MTA study. In addition, 18 percent of the children with ADHD and ASD discontinued participation due to adverse events (prominently irritability) compared to 1.4 percent in children with ADHD only.

NEUROPSYCHIATRIC DEFICITS SECONDARY TO TRAUMATIC BRAIN INJURY. The overall risk of TBI in the general population has been estimated at 2 to 5 percent. In the United States, approximately 230/100,000 of children under 15 years of age, and 294/100, 000 adolescents between 15 and 19 years of age experience a TBI each year. A significant proportion of patients who sustained a TBI will experience disturbance in academic, emotional, and social functioning. In a research setting, the most common psychiatric complication of TBI was a cluster of symptoms consistent with ADHD. A double-blind, placebo-controlled, crossover-study design used to investigate the efficacy of methylphenidate in children with ADHD symptoms secondary to a TBI, has demonstrated that methylphenidate was superior to placebo to benefit attention and concentration. This finding is in line with other RCTs that have demonstrated benefit of other stimulants, including pemoline and dextroamphetamine, for symptoms of ADHD secondary to TBI.

APATHY. Apathy is a common symptom of neurodegenerative diseases such as Alzheimer disease (AD), Parkinson disease (PD), as well as an

adverse effect of the sedative effects of pain medications, or toxic effects of chemotherapy. Unfortunately, few studies have investigated the benefit of psychostimulants in treating apathy secondary to these conditions. One RCT examining the efficacy of methylphenidate for the treatment of apathy associated with AD demonstrated greater reduction of apathy with methylphenidate 10 mg twice daily compared with placebo on Apathy Evaluation Scale change scores. Thus, efficacy studies of psychostimulants for the treatment of secondary neuropsychiatric complications are all limited by a small sample size, and necessitate larger trials to guide evidence-based practice.

DEPRESSION. Antidepressants are first-line treatment for moderate or severe depression, but depression resistant to antidepressants presents a therapeutic challenge. A Cochrane review article assessed evidence from RCTs on the effectiveness and safety of stimulants, such as methylphenidate and methylamphetamine, and modafinil and compared them to placebo in the treatment of major depressive disorder, either as monotherapy or as adjuncts to antidepressants. Several placebo-controlled parallel design studies have demonstrated benefit-adjunctive psychostimulants to reduce depressed mood in the short term. Further examination and pooling of longer-term outcomes for other symptoms of depression suggest that psychostimulants reduce fatigue associated with depression, but do not effectively target other neurovegetative symptoms. Other smaller short-term trials of psychostimulants used adjunctively versus placebo have not found significant differences in overall depression scores. In addition, modafinil is no more effective than placebo in targeting symptoms of depression including fatigue or hypersomnia. The overall quality of the extant data was suboptimal due to small sample sizes, selection bias due to inadequate randomization, attrition, and performance bias.

Treatment of Comorbid ADHD and Seizure Disorder. A retrospective study of youth with comorbid epilepsy and ADHD found that being seizure-free did not predict response to stimulant treatment. An MPH preparation had a 5.57-fold greater chance of treatment response compared to an AMPH preparation. Importantly, cognitive level may play a role in treatment response, but age, type of seizure, and success of previous stimulant trials were not associated with response to stimulants in this population. Stimulant discontinuation rates were comparable in seizure-free group (35 percent) and not-seizure-free groups (53 percent). In the not-seizure-free group, 3 out of the 19 patients had a worsening of seizures while on a stimulant.

Tolerability of Psychostimulants: Adverse Effects, Contraindications, and

Precautions

Psychostimulants are generally well tolerated. Different formulations of methylphenidate and amphetamine have similar side-effect profiles that present similarly in pediatric and adult populations, with some exceptions that will be noted below. Common side effects associated with stimulants include insomnia, decreased appetite, weight loss, headache, stomachache, and mild increases in heart rate and blood pressure (see Table 34.29–4 for a comprehensive list of adverse effects). Short-term or acute stimulant side effects are more likely to be related to higher doses, and careful dose titration mitigates these acute adverse effects.

Certain patient groups appear to be more sensitive to stimulants at doses lower than the general population, including preschool children with ADHD, children with developmental delays, and the elderly. It is recommended that stimulants be used with caution in the following populations: children under 5 years of age with IQ ≤5 years old, severe tics, or in children with histories of severe anxiety, mood, or psychotic symptoms. Further consideration of lower doses should be given in the elderly population. Nevertheless, stimulants may be first-line treatments for ADHD symptoms in these populations, despite a relatively small number of RCTs, response rates of 50 percent, and higher rates of side effects. Absolute and relative contraindications for stimulant use are listed in Table 34.29–5. Specific adverse reactions to stimulants will now be discussed in more detail.

Table 34.29–4.
Comprehensive List of Adverse Events

Adverse Effects	
Common:	***Uncommon:***
MPH/AMPH	Psychosis
Anorexia, weight loss (13–20%)	Euphoria/mania Sadness/isolation
Insomnia (11–30%)	Major depressive episodes Cognitive impairment Growth retardation
Headaches (12–15%)	Tic disorders (i.e., Tourette syndrome) Increased heart rate (clinically significant)
Abdominal pain (6–12%)	Impaired liver functioning (pemoline only) Increased blood pressure
Emotional lability (2–10%)	Dizziness, lethargy, fatigue Nausea, constipation Rash/hives
Slight increase in HR, BP	Hyperacusis

***Severe side effects occur in 4–10% of children treated with stimulants**	Formication Necrotizing angitis brain (IV amphetamine) Neuroleptic malignant syndrome?

Specific Adverse Events

Tics. Significant concerns regarding the development of tics and Tourette disorder due to use of psychostimulants in children and adolescents were initially raised in the 1980s. For many years, stimulants had been thought to induce or worsen tics in youth and presence of tics constituted a contraindication to psychostimulant use. In the last decade, several studies have suggested that stimulants do not worsen or reactivate tics, and might even improve pre-existing tics in many patients.

Table 34.29–5.
Contraindications to Psychostimulants

Absolute Contraindications	Relative Contraindications
1. Known hypersensitivity to stimulants 2. Structural cardiac defects 3. Narrow-angle glaucoma 4. Moderate to severe substance abuse disorder 5. Current use of MAOI or within 14 days after an MAOI has been discontinued	1. Presence of psychotic disorders like Schizophrenia, because of stimulants' potential to worsen or even precipitate psychotic symptoms 2. A severe tic or Tourette syndrome is a relative contraindication. This contraindication does not apply to mild tics in patients with severe ADHD symptoms 3. Unstable hypertension is a relative contraindication. Blood pleasure should be normalized and well controlled before administering stimulant medications

A Cochrane review analyzed eight RCTs for a potential relation between psychostimulant use and the development of tic disorders. One RCT in children with cooccurring ADHD and Tourette syndrome found a positive treatment effect with methylphenidate compared to placebo and with clonidine plus methylphenidate on measures of ADHD and tic severity. In the methylphenidate group only, 20 percent of participants reported worsening of tics as an adverse event compared to 22 percent receiving placebo. In a small placebo-controlled crossover study of dextroamphetamine that randomized 20 children to 3 dosage titrations of dextroamphetamine (target doses of 7.5, 15, or 22.5 mg twice daily) for 3 weeks, dextroamphetamine decreased hyperactivity but there was no significant interaction between drug and dose, indicating that additional improvements in hyperactivity were not observed for higher doses. Tic severity was significantly greater during the second (15 mg twice daily) and third weeks (22.5 mg twice daily) of dextroamphetamine treatment compared to weeks on placebo. Drug dose did not worsen tics, however. Larger studies are needed to definitively determine whether psychostimulants worsen tics or reduce pre-existing tics in youth with

comorbid ADHD and a tic disorder.

The increased sensitivity of dopaminergic brain systems has been widely implicated in the pathophysiology of tics. To date, no clear genetic relation between tics and ADHD on the emergence of tics during stimulant trials has been established. Moreover, there is commonly a delay between stimulant exposure and tic onset making it difficult to establish a causal association between stimulant exposure and tic outcome. In individuals with a personal or family history of tics or Tourette disorder and ADHD, stimulant medications may exacerbate motor and vocal tics. However, the data are not conclusive and in most cases, the benefit of treating ADHD outweighs the risk of stimulants triggering tics.

Cardiovascular Effects. A risk of sudden cardiac death from psychostimulant use emerged around the time they became a mainstay treatment for ADHD. The US FDA identified 11 sudden deaths in pediatric patients taking methylphenidate from 1992 to 2005. However, the rate of sudden cardiac death among patients with ADHD is equivalent to that of the general population and comparable to mortality due to motor vehicle accidents in children. In large community samples of children and adolescents, stimulants have not been shown to increase risk of cardiovascular complication, including risk of sudden death. Risk for stimulant-induced adverse cardiovascular events appears to be elevated in children with ADHD and a family history significant for early cardiac death, arrhythmia, cardiogenic syncope, or a personal history of structural cardiac abnormalities, exercise-induced syncope, chest pain, or palpitations. According to the AACAP practice guidelines, in such cases an assessment by a pediatric cardiologist prior to stimulant treatment is recommended. Finally, authors of the large prospective MTA study investigated stimulant-related changes in vital signs in youth treated for ADHD for up to 10 years. Treatment was associated with increases in heart rate at 14-month, 3-year, and 8-year follow-ups suggesting chronic induction of the sympathomimetic system by stimulant treatment. Stimulants were not associated with elevations in systolic or diastolic blood pressure.

Growth Velocity. Long-term stimulant use has been associated with interfering with typical growth trajectories in children and adolescents with ADHD. Specifically, MTA subjects reported significant decreases in height velocity with long-term stimulant treatment. Maximal reduction in annual growth rate occurred in the first year, decreased in the second year, and then subsided completely by the third year of treatment. Matched subjects with history of treatment before entry into the MTA study did not show differences in height velocity compared to treatment-naive youth over 3 years of follow-up. Treatment-naive youth had small growth acceleration at

2-year follow-up that was equivalent to the untreated group, and the treatment-naive group on stimulants did not compensate for the slowdown of growth observed initially. Compared to typically developing healthy comparison youth without a diagnosis of ADHD, ADHD youth not treated with stimulants were taller and heavier and grew at a faster rate. It is hypothesized that stimulant-related reductions in growth may be related to DA-induced decrease of prolactin secretion by the pituitary gland or suppression of growth hormone (GH) and insulin-like growth factors (IGFs).

The PATS study reported significant growth velocity suppression in preschoolers treated with stimulants for at least 1 year, with a 20 percent less than expected height gain (−1.38 cm/year) and 55 percent less than expected weight gain (−1.32 kg/year). Children under 6 years of age (mean age = 4.4 years) with ADHD were measured up to 29 times for at least 10 months in study follow-up. The height and weight of 140 stimulant-naive preschool-age children with a diagnosis of ADHD were much larger (an average of 2.04 cm taller and 1.78 kg heavier) than expected than referenced Centers for Disease Control (CDC) norms. After initiation of stimulant medication, the height and weight velocity slowed, and for the children who remained on medication, this resulted in annual gain that was 20 percent less for height and 55 percent less for weight than expected by CDC standards. These findings are consistent with the growth velocity findings reported in the MTA study. Moreover, greater stimulant-related reductions in height velocity have been observed with longer exposure to psychostimulants, in younger children compared to older adolescents, and for taller and heavier children. The magnitude of weight and height deficits does not appear to differ between MPH and AMPH.

In the long term, growth deficits normalize so that final height is not affected. In some children with ADHD, deficits in physical growth constitute a developmental delay much like core behavioral and cognitive signs of the disorder. Indeed, stimulant-associated growth deficits in children with ADHD may represent early temporary manifestations of ADHD itself. In addition, a failure to gain weight may impact height velocity although no studies have directly examined this possibility in youth with ADHD or have followed patients into adulthood to investigate a complete height trajectory. Stimulant effects on child growth should be part of a risk/benefit discussion with the child and family as part of the informed consent, and clinicians should monitor growth parameters in all patients taking stimulants. Change in weight or height crossing two percentile lines on a standardized growth curve suggest a deficit in growth that should prompt a clinical response. Possible responses might include stimulant drug holiday, nutritional supplementation, dose reduction, or change to a nonstimulant ADHD medication.

Abuse and Dependence. Risk of substance use, abuse or dependence in individuals diagnosed with ADHD complicates the diagnosis and is of great interest to clinicians and researchers. In a 10-year follow-up study, children and adolescents diagnosed with ADHD had an increased risk of a substance use disorder (SUD) compared to matched controls. Factors such as ADHD severity and duration, comorbid mood disorders, and CD are independently associated with a higher risk for SUD compared to controls. Presence or absence of treatment did not appear to change the risk of substance use, abuse, or dependence. In the MTA group, there were no protective or predisposing factors associated with SUD at 3-year follow-up. At the 6-year follow-up (mean age = 15 years), 35 percent of the ADHD struggled with an SUD versus 20 percent of controls. Thus, according to the MTA study, stimulants neither protect from, nor increase risk for, SUD. Finally, in a prospective study, a risk of SUD increased with age at initiation of methylphenidate, suggesting that delays in treatment of ADHD may be associated with the development of SUD.

Misuse and Diversion. When taken as indicated, stimulants are generally safe and highly effective for the treatment of ADHD symptoms. However, stimulant-induced elevations in DA in key brain regions appear to be integral to both efficacy in ADHD and potential for abuse. The nonmedical use of stimulants has received increasing attention, and has prompted consideration of nonstimulant options for the treatment of ADHD. Data from the National Survey on Drug Use and Health (NSDUH) database tracks the diversion and misuse of stimulants. In 2002 to 2003, the NSDUH found that approximately 6 percent of 16 to 25 year olds reported lifetime nonmedical misuse. Between 2003 and 2009, the lifetime prevalence of nonmedical misuse of methylphenidate was stable at lower rates ranging between 1.8 and 2.1 percent. Among 18 to 25-year-old youths, over the same time period, prevalence rates were around 5 percent.

Benson and colleagues (2015) published a meta-analysis on the misuse and diversion in the college student population. They found that a substantial number of college students are involved in nonmedical misuse and diversion of stimulants. Diversion is defined as giving away, selling, or otherwise distributing one's own medication. In 2014, 36 percent of stimulant prescription holders diverted their stimulant, and Adderall was the most frequently diverted stimulant. Fifty-six percent of prescription holders had been approached to divert their medication in the last 6 months. Factors associated with misuse of stimulants were male gender, being a member of a social college organization, having a diagnosis of CD and/or SUD, using immediate-release formulations, or using alcohol or illicit substances.

Thus, stimulant diversion and misuse among student populations is growing and is higher than that among general population samples.

However, misuse appears to be most prevalent and problematic among adults with ADHD and illicit drug users. Based on these data, the AACAP practice parameters recommend that clinicians obtain a careful personal and family history of substance abuse or dependence prior to prescribing a stimulant, and should educate populations most vulnerable to diversion, including high school or college students.

MODAFINIL (PROVIGIL)

Modafinil is a wakefulness-promoting agent with behavioral actions similar to psychostimulants, but is different pharmacologically not to be classified as a psychostimulant. Physiologically, modafinil can increase energy, concentration, and induce wake-promoting effects. Modafinil is approved for treatment of excessive sleepiness associated with narcolepsy, obstructive sleep apnea/hypopnea syndrome (OSAHS) or shift work sleep disorder (SWSD). In addition, multiple studies have investigated the efficacy and safety of modafinil in the treatment of ADHD, mood disorders, schizophrenia, and SUD.

MECHANISM OF ACTION

Modafinil is a racemic compound. The chemical name for modafinil is diphenylmethyl sulfinyl acetamide and its molecular formula is C15H15NO2S. Modafinil and its R-enantiomer, armodafinil, exert clinical effects by enhancing catecholamine neurotransmission. Modafinil has been found to activate lateral hypothalamus orexin (hypocretin) neurons that send axons to numerous brain regions, including major nuclei that have been implicated in sleep regulation. However, because its pharmacological profile is unique from stimulants, it is not classified as such. In addition to the wake-promoting effects, modafinil increases motor activity, euphoria, and alterations in perception, emotions, and cognition typical of CNS stimulants. Its optical enantiomer, armodafinil, has similar pharmacological actions in animals.

Modafinil, like other psychostimulants, increases the release of monoamines, but also elevates histamine, epinephrine, γ-aminobutyric acid, and glutamate levels. As a result, some researchers consider modafinil a “wakefulness-promoting agent” rather than a classic amphetamine-like psychostimulant. For example, mice lacking dopamine transporters (DATs) do not respond to the wake-promoting effects of modafinil, and this is also true for mice lacking D1 and D2 receptors. Further, microdialysis studies have reported that modafinil increases extracellular DA. PET with a D2/D3 radioligand (sensitive to changes in endogenous DA) and [11C] cocaine (DA transporter radioligand) was used to measure the effects of modafinil on extracellular DA and on DAT in 10 healthy males. A significant correlation between the percentage DA transporter occupancy and plasma modafinil

concentration in the caudate was demonstrated. Other studies have demonstrated that in mice engineered without DAT, modafinil failed to induce wake-promoting activity, suggesting that DAT was essential for its ability to exert its effects. Unlike those of amphetamine, however, the function of modafinil was not inhibited by a DA receptor antagonist. These studies suggest that despite some functional mimicry, modafinil acts through different cellular mechanisms than through stimulants.

Studies have also demonstrated that modafinil has moderate interaction with NE and serotonin receptors. However, unlike stimulants, modafinil does not release NE at a high rate. It has been hypothesized that this difference in effect on the catecholamine system might be responsible for lower efficacy in reducing of ADHD symptoms, compared to psychostimulants.

Pharmacological Properties of Modafinil

Modafinil is rapidly absorbed, and the absorption of the medication is not affected at steady state; a peak plasma concentration of 6.4 mg/mL was reached in 2.7 hours with modafinil 200 mg once daily. Approximately 90 percent of a modafinil dose is metabolized hepatically, primarily to two inactive metabolites. Less than 10 percent of an administered dose is excreted as the parent compound. The pharmacokinetics of modafinil may be altered in older males and in patients with chronic hepatic impairment. In vitro, modafinil has been shown to inhibit the cytochrome P450 (CYP) isoenzyme CYP2C19 (moderately potent, reversible inhibition), modestly induce CYP1A2, CYP3A4/5, and CYP2B6 and suppress CYP2C9.

US FDA–Approved Uses of Modafinil

Narcolepsy. Narcolepsy is a common, disabling sleep disorder characterized by excessive daytime sleepiness (EDS) and rapid eye movement (REM) sleep manifestations that intrude into wakefulness, including cataplexy (sudden loss of muscle tone triggered by strong emotions), sleep paralysis, hypnagogic/hypnopompic hallucinations (sensory events that occur at the transition from wakefulness to sleep), and sleep-onset REM periods. Narcolepsy affects approximately 0.02 percent of the population worldwide and causes disability in 24 percent of the affected subjects.

Currently, modafinil is considered a first-line therapy for EDS in narcolepsy. Four randomized, double-blind, placebo-controlled trials have assessed the usefulness of modafinil in treatment of EDS in narcolepsy and have demonstrated that treatment with modafinil 200 to 400 mg/day had significant positive impact on improved EDS. Modafinil does not appear to benefit cataplexy. Abrupt discontinuation may cause rebound sleepiness but does not appear to be associated with any amphetamine-like

withdrawal symptoms.

Obstructive Sleep Apnea. Efficacy and safety of modafinil treatment for obstructive sleep apnea (OSA) in patients who are not benefiting from CPAP has been demonstrated in a meta-analysis of 3 RCTs in this population. Modafinil is also effective in improving intentional or unintentional sleep in SWSD, which translated into fewer accidents and near accidents on the commute home but not during the night shifts.

Non-US FDA–Indicated Use of Modafinil

ADHD. Although sympathomimetic stimulants are the most common treatment for ADHD, a sizable minority of children and adolescents with ADHD do not respond adequately to psychostimulants or have intolerable side effects. Thus, additional treatment options are needed. Given the attention and wakefulness-enhancing properties of modafinil, it has been considered a candidate medication for ADHD. By 2007, multiple trials had examined the clinical utility of modafinil for ADHD. More than 900 children and adolescents were examined, and demonstrated the efficacy and tolerability of modafinil for the treatment of ADHD.

In 2005, 250 6- to 17-year-old youth, normalized for demographic factors, were involved in a 9-week, multicenter, randomized, double-blind, placebo-controlled, flexible-dose study of the film-coated tablet formulation of modafinil for treatment of ADHD symptoms. Modafinil was titrated to an optimal dose based on efficacy and tolerability (range: 170 to 425 mg once daily). Efficacy was assessed by clinicians (who completed the Attention-Deficit/Hyperactivity Disorder Rating Scale-IV (ADHD-RS-IV), interviews with teachers (who completed School Version of the ADHD-RS-IV) and parents (who completed Home Version of the ADHD-RS-IV) and the Clinical Global Impression of Improvement scale. Safety assessment was based on adverse events reported, laboratory tests, and vital signs. The effect size for improvement of core ADHD symptoms was moderate to large (Cohen's d = 0.69). Thus, once-daily modafinil was effective on the core symptoms of inattention, hyperactivity, and impulsivity, both at school and at home. Among the most common adverse events reported were insomnia and decreased appetite, and were mild to moderate in severity.

Cocaine Addiction. Several small RCTs have examined the use of modafinil to prevent cravings in patients with cocaine addiction. In a parallel group study, 62 treatment-seeking cocaine-dependent subjects without significant medical or psychiatric problems were randomized to receive modafinil 200 to 400 mg/day or placebo for 8 weeks. All subjects also received cognitive behavioral therapy twice a week. The efficacy measures were clean urine toxicology, subjective craving, and withdrawal reporting. Even though there were significantly more negative urine results

in patients taking modafinil, the subjective reporting of cocaine use, dollars spent, and cocaine craving remained unchanged. At present, there is no consistent evidence that modafinil is beneficial for treatment of cocaine craving.

Cognitive Enhancement. A number of neurodevelopmental and neuropsychiatric disorders are associated with cognitive impairment. Impaired cognition can lead to poor functional outcome and may limit the benefit gained by psychosocial interventions. Modafinil has recently emerged as a possible medication to improve cognition for certain patients. It has been shown to improve cognitive function including memory, planning and attention in animal studies, as well as reaction time, logical reasoning and short-term working memory, planning and mental flexibility in nonsleep-deprived and sleep-deprived humans. Modafinil may offer advantages over compounds such as amphetamine and L-dopa particularly if repeated administration is needed, as it is associated with low risk of abuse and has a relatively mild side-effect profile.

Adverse Effects. Modafinil is generally well tolerated in patients with excessive sleepiness associated with OSAHS or SWSD. Adverse events are generally of mild to moderate severity, and include headache, nausea, and nervousness.

ATOMOXETINE (STRATTERA)

ATM was the first US FDA medication approved for ADHD in children older than 6 years of age, adolescents, and adults. Similar to psychostimulants, ATM shows potent and very selective NE transporter binding. It selectively blocks NE reuptake at the noradrenergic neuron and is classified as a selective norepinephrine reuptake inhibitor (SNRI). However, it is not classified as a psychostimulant, and is not a controlled substance.

Mechanism of Action

ATM has been shown to increase extracellular availability of both NE and DA in the rat PFC. ATM improves measures of PFC cognitive function in rats, monkeys, and humans. In order to improve executive function in primates, ATM has to produce an optimal dose of catecholamines in the synaptic space; higher than optimal doses may impair executive function. The enhancing effects are blocked by coadministration of either an NE alpha-2 or DA D1 antagonist, suggesting that the NE and DA neurotransmission complement are essential for ATM function. Interestingly, optimal doses vary widely, which may explain variability of drug response in ATM compared to psychostimulants.

Pharmacological Properties and Drug Interactions of Atomoxetine

ATM has high aqueous solubility and lipophilicity that facilitates its rapid and complete GI absorption after oral administration. In the serum, 98 percent of the medication is bound primarily to albumin. Three oxidative metabolic pathways are involved in the hepatic metabolism of ATM: aromatic ring hydroxylation, benzylic hydroxylation and *N*-demethylation. Aromatic ring hydroxylation results in the formation of the primary oxidative metabolite of ATM, 4-hydroxyatomoxetine, which is subsequently excreted in urine. The formation of 4-hydroxyatomoxetine is primarily mediated by P450 (CYP) 2D6. This results in two distinct populations of individuals: those exhibiting active metabolic capabilities (CYP2D6 extensive metabolizers) and those exhibiting poor metabolic capabilities (CYP2D6 poor metabolizers) for ATM.

The oral bioavailability and clearance of ATM are influenced by the activity of CYP2D6. In extensive metabolizers, ATM has a plasma half-life of 5.2 hours. Poor metabolizers, ATM has a plasma half-life of 21.6 hours. As a result, the average steady-state plasma concentrations are approximately 10-fold higher in poor metabolizers compared with extensive metabolizers. ATM accumulates in plasma of poor metabolizers, but very little accumulation in extensive metabolizers. The volume of distribution in ATM suggests that it is distributed in total-body water, in both extensive and poor metabolizers. ATM is highly bound to plasma albumin (~99 percent bound in plasma). ATM does not inhibit or induce the clearance of other medications metabolized by the CYP system. In extensive metabolizers, potent and selective CYP2D6 inhibitors reduce ATM clearance; however, administration of CYP inhibitors to poor metabolizers has no effect on the steady-state plasma concentrations of ATM.

Albuterol used concomitantly with ATM may increase heart rate and blood pressure. ATM is contraindicated in patients currently receiving or having recently received (i.e., within 2 weeks) MAO inhibitor therapy. The starting dose of ATM should be reduced when potent CYP2D6 inhibitors including fluoxetine, paroxetine, or quinidine are administered concomitantly with ATM.

References

Atomoxetine/Methylphenidate Comparative Study Group; Newcorn JH, Kratochvil CJ, Allen AJ, et al. Atomoxetine and osmotically released methylphenidate for the treatment of attention deficit hyperactivity disorder: acute comparison and differential response. *Am J Psychiatry*. 2008;165(6):721–730.

Bates G. Medication in the treatment of the behavioural sequelae of traumatic brain injury. *Dev Med Child Neurol*. 2006;48(8):697–701.

Biederman J, Swanson JM, Wigal SB, et al; Modafinil ADHD Study Group. A comparison of

once-daily and divided doses of modafinil in children with attention-deficit/hyperactivity disorder: a randomized, double-blind, and placebo-controlled study. *J Clin Psychiatry*. 2006;67(5):727–735.

Candy M, Jones L, Williams R, Tookman A, King M. Psychostimulants for depression. *Cochrane Database Syst Rev*. 2008;(2):CD006722.

Castells X, Ramos-Quiroga JA, Bosch R, Nogueira M, Casas M. Amphetamines for attention deficit hyperactivity disorder (ADHD) in adults. *Cochrane Database Syst Rev*. 2011; (6):CD007813.

Clemow DB, Walker DJ. The potential for misuse and abuse of medications in ADHD: a review. *Postgrad Med*. 2014;126(5):64–81.

Garnock-Jones KP, Keating GM. Atomoxetine: a review of its use in attention-deficit hyperactivity disorder in children and adolescents. *Paediatr Drugs*. 2009;11(3):203–226.

Gonzalez-Heydrich J, Hsin O, Gumlak S, et al. Comparing stimulant effects in youth with ADHD symptoms and epilepsy. *Epilepsy Behav*. 2014;36:102–107.

Gould MS, Walsh BT, Munfakh JL, et al. Sudden death and use of stimulant medications in youths. *Am J Psychiatry*. 2009;166(9):992–1001.

Greenhill L, Kollins S, Abikoff H, et al. Efficacy and safety of immediate-release methylphenidate treatment for preschoolers with ADHD. *J Am Acad Child Adolesc Psychiatry*. 2006;45:1284–1293.

Heal DJ, Smith SL, Gosden J, Nutt DJ. Amphetamine, past and present–a pharmacological and clinical perspective. *J Psychopharmacol*. 2013;27(6):479–496.

Herrmann N, Rothenburg LS, Black SE, et al. Methylphenidate for the treatment of apathy in Alzheimer disease: prediction of response using dextroamphetamine challenge. *J Clin Psychopharmacol*. 2008;28:296–301.

Keating GM, Raffin MJ. Modafinil: a review of its use in excessive sleepiness associated with obstructive sleep apnoea/hypopnoea syndrome and shift work sleep disorder. *CNS Drugs*. 2005;19(9):785–803.

Kuczenski R, Segal DS. Effects of methylphenidate on extracellular dopamine, serotonin, and norepinephrine: comparison with amphetamine. *J Neurochem*.1997;68(5):2032–2037.

Li Z, Chang SH, Zhang LY, Gao L, Wang J. Molecular genetic studies of ADHD and its candidate genes: a review. *Psychiatry Res*. 2014;219(1):10–24.

Mahajan R, Bernal MP, Panzer R, et al. Clinical practice pathways for evaluation and medication choice for attention-deficit/hyperactivity disorder symptoms in autism spectrum disorders. *Pediatrics*. 2012;130 Suppl 2:S125–S138.

Mannuzza S, Klein RG, Truong NL, et al. Age of methylphenidate treatment initiation in children with ADHD and later substance abuse: prospective follow-up into adulthood. *Am J Psychiatry*. 2008;165(5):604–609.

McGough JJ, Biederman J, Greenhill LL, et al. Pharmacokinetics of SLI381 (ADDERALL XR), an extended-release formulation of Adderall. *J Am Acad Child Adolesc Psychiatry*. 2003;42(6):684–691.

Molina BS, Hinshaw SP, Eugene Arnold LE, et al. Adolescent substance use in the multimodal treatment study of attention-deficit/hyperactivity disorder (ADHD) (MTA) as a function of childhood ADHD, random assignment to childhood treatments, and subsequent medication. *J*

Am Acad Child Adolesc Psychiatry. 2013;52(3):250–263.

MTA Cooperative Group; A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. The MTA Cooperative Group. Multimodal Treatment Study of Children with ADHD. *Arch Gen Psychiatry*. 1999;56(12):1073–1086.

MTA Cooperative Group; National Institute of Mental Health Multimodal Treatment Study of ADHD follow-up: 24-month outcomes of treatment strategies for attention-deficit/hyperactivity disorder. *Pediatrics*. 2004;113(4):754–761.

Newcorn JH, Spencer TJ, Biederman J, Milton DR, Michelson D. Atomoxetine treatment in children and adolescents with attention-deficit/hyperactivity disorder and comorbid oppositional defiant disorder. *J Am Acad Child Adolesc Psychiatry*. 2005;44(3):240–248.

Pliszka S; AACAP Work Group on Quality Issues. Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2007;46(7):894–921.

Pringsheim T, Steeves T. Pharmacological treatment for attention deficit hyperactivity disorder (ADHD) in children with comorbid tic disorders. *Cochrane Database Syst Rev*. 2011; (4):CD007990.

Sagvolden T, Russell VA, Aase H, Johansen EB, Farshbaf M. Rodent models of attention-deficit/hyperactivity disorder. *Biol Psychiatry*. 2005;57(11):1239–1247.

Schneider MF, Krick CM, Retz W, et al. Impairment of fronto-striatal and parietal cerebral networks correlates with attention deficit hyperactivity disorder (ADHD) psychopathology in adults—a functional magnetic resonance imaging (fMRI) study. *Psychiatry Res*. 2010;183(1):75–84.

Spencer RC, Devilbiss DM, Berridge CW. The cognition-enhancing effects of psychostimulants involve direct action in the prefrontal cortex. *Biol Psychiatry*. 2015;77(11):940–950.

Swanson J, Greenhill L, Wigal T, et al. Stimulant-related reductions of growth rates in the PATS. *J Am Acad Child Adolesc Psychiatry*. 2006;45(11):1304–1313.

Swanson JM, Elliott GR, Greenhill LL, et al. Effects of stimulant medication on growth rates across 3 years in the MTA follow-up. *J Am Acad Child Adolesc Psychiatry*. 2007;46(8):1015–1027.

*Volkow ND, Wang GJ, Tomasi D, et al. Methylphenidate-elicited dopamine increases in ventral striatum are associated with long-term symptom improvement in adults with attention deficit hyperactivity disorder. *J Neurosci*. 2012;32(3):841–849.

Wigal SB, Gupta S, Greenhill L, et al. Pharmacokinetics of methylphenidate in preschoolers with attention-deficit/hyperactivity disorder. *J Child Adolesc Psychopharmacol*. 2007;17(2):153–164.

▲ 34.30 Dopamine Receptor Agonists

CARLOS A. ZARATE, JR., M.D., AND LAWRENCE T. PARK, M.D.

INTRODUCTION AND HISTORY

Since Carlsson's discovery of the physiological functions of dopamine (3-hydroxytyramine) in 1957, the dopaminergic system has been extensively studied and is among the best understood of all neurotransmitter systems. In humans, four major dopaminergic pathways have been identified: nigrostriatal (A9), mesolimbic (A10), mesocortical (A10), and tuberoinfundibular (A8). These pathways have been implicated in motor activity, affect, reward mediation, attention, cognitive functioning, appetite, and sleep. Five subtypes of dopamine (D) receptors have been identified (D1, D2, D3, D4, D5). They are all G protein–coupled receptors and are classified into D1 class (D1 and D5) which increases cAMP production via positive coupling to adenylyl cyclase and D2 class (D2, D3, D4) which is associated with inhibition of adenylyl cyclase and decreases cAMP production.

Dopamine receptor agonists (DA) bind to and activate dopamine receptors to reproduce the activity of the endogenous neurotransmitter. Classic DAs were initially developed during the 1970s to treat advanced Parkinson disease (PD). They are typically classified into nonergot and ergot derivatives. Nonergot DAs include apomorphine (Apokyn), pramipexole (Mirapex), and ropinirole (Requip). Though successfully used as monotherapy for new-onset PD, they are more commonly used as adjunctive treatment with levodopa (a dopamine precursor); as a result, smaller doses of levodopa are required and complications are delayed or reduced. Apomorphine has D1 and D2 agonist effects, and though it has demonstrated efficacy in PD, it is not used in the United States due to its short half-life. However, it is used in Europe for PD via subcutaneous and intravenous pump administration. Pramipexole and ropinirole are FDA approved for use in PD and restless legs syndrome (RLS), and have been studied as an adjunctive treatment for unipolar and bipolar depression. Amantadine (Symmetrel), an adamantane derivative, is FDA approved to treat PD, drug-induced extrapyramidal symptoms (EPS), and influenza A infection. It is also used off-label to treat neuroleptic malignant syndrome (NMS) and Cotard syndrome, a psychiatric condition that involves a delusional belief that one is dead. The ergot class includes bromocriptine (Parlodel), cabergoline (Dostinex, Caberlin), and pergolide (Permax). Bromocriptine is FDA approved for use in PD, galactorrhea (due to hyperprolactin conditions), and acromegaly, and is used off-label to treat RLS and NMS. Cabergoline is FDA approved for the treatment of

hyperprolactinemic disorders, either idiopathic or due to pituitary adenomas. In addition, a number of studies have suggested that cabergoline may be effective in treating PD or preventing ovarian hyperstimulation syndrome. Pergolide, an ergot D1/2 agonist was FDA approved for PD and studied as an adjunctive treatment of treatment-resistant depression (TRD). It is important to note that the entire class of ergot derivatives has been associated with cardiac fibrosis. As a result, the FDA issued a warning for the risk of cardiac valvulopathy for cabergoline in 2011 and pergolide was withdrawn from the US market in 2007 because of an increased risk of valvular heart disease.

Other classes of drugs may also exert agonist effects at the dopamine receptor. The wakefulness-promoting agents, modafinil (Provigil) and armodafinil (Nuvigil), have complex prodopaminergic activity, including dopamine transporter inhibition. They may also have dopamine (D2) receptor partial agonist effects. These drugs have been FDA approved to improve wakefulness in adult patients with excessive sleepiness associated with narcolepsy, obstructive sleep apnea (OSA), or shift work disorder (SWD). They have been studied as an adjunctive treatment for unipolar and bipolar depression, and attention-deficit/hyperactivity disorder (ADHD). These drugs have been used in nonpathological settings to optimize normal performance.

Finally, certain drugs have demonstrated partial agonist/antagonist properties at the dopamine receptor. For instance, aripiprazole (Abilify), a second-generation antipsychotic drug approved for the treatment of schizophrenia, acute manic/mixed episodes of bipolar I disorder, and irritability associated with autistic disorder in children and adolescents, is also approved as an adjunctive treatment for major depressive disorder (MDD). Cariprazine, a partial D2/3 agonist (with D3 selectivity), has a pending New Drug Application (NDA) with the FDA for schizophrenia and bipolar I disorder and is being investigated as an adjunctive treatment for MDD.

This chapter describes the key therapeutic DAs with significant psychiatric applications: pramipexole, ropinirole, amantadine, bromocriptine, modafinil, armodafinil, and aripiprazole.

CHEMISTRY

Pramipexole is a nonergot synthetic aminothiazole derivative. The chemical name is (S)-2-amino-4,5,6,7-tetrahydro-6-(propylamino)benzothiazole dihydrochloride monohydrate (IUPAC nomenclature). The molecular formula is $C_{10}H_{17}N_3S.2HCl.H_2O$ and the molecular weight is 302.27. It is a white to off-white crystalline powder, freely soluble in water, soluble in methanol, sparingly soluble to slightly soluble in ethanol, and practically insoluble in methylene chloride. The molecular structure is shown in Figure 34.30–1.

Ropinirole is a nonergot hydrochloride salt. The chemical name is 4-[2-(dipropylamino)ethyl]-1,3-dihydro-2*H*-indol-2-one. The molecular formula is C16H24N2O.HCl and the molecular weight is 296.84. It is a pale cream to yellow powder, soluble in water and methanol, and very slightly soluble in ethyl alcohol. The molecular structure is shown in Figure 34.30–2.

Amantadine is a nonergot adamantane derivative. The chemical name is adamantan-1-amine. The molecular formula is C10H17N and the molecular weight is 151.24868. It is a stable white or nearly white crystalline powder, freely soluble in water and soluble in alcohol and chloroform. The molecular structure is shown in Figure 34.30–3.

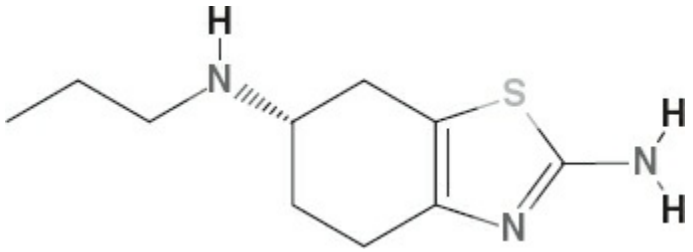

FIGURE 34.30–1. Molecular structure of pramipexole. (From National Center for Biotechnology Information. PubChem Compound Database; CID = 119570. http://pubchem.ncbi.nlm.nih.gov/compound/119570. Accessed May 13, 2015.)

Bromocriptine is a semisynthetic ergotamine alkaloid. The chemical name is (5′α)-2-bromo-12′-hydroxy-5′-(2-methylpropyl)-2′-(propan-2-yl)-3′,6′,18-trioxoergotaman.The molecular formula is C32H40BrN5O5 and the molecular weight is 654.5945. It is a white or slightly colored fine crystalline powder, soluble in methyl alcohol, and practically insoluble in water, chloroform, and benzene. The molecular structure is shown in Figure 34.30–4.

Modafinil is an oxazolidinone compound, which is a racemic compound consisting of a 1:1 mixture of R- and S-isomers (enantiomers). Armodafinil is the R-isomer of the racemic mixture. The chemical name of modafinil is 2-[(diphenylmethyl)sulfinyl]acetamide and armodafinil is (–)-2-[(*R*)-(diphenylmethyl)sulfinyl]acetamide. For both, the molecular formula is C15H15NO2S and the molecular weight is 273.35. Modafinil (and armodafinil) is a white to off-white crystalline powder that is sparingly to slightly soluble in methanol and acetone, and practically insoluble in water and cyclohexane. The molecular structure is shown in Figure 34.30–5.

Aripiprazole is a second-generation antipsychotic benzisoxazole derivative. The chemical name is 7-[4-[4-(2,3-dichlorophenyl) piperazin-1-yl]butoxy]-3,4-dihydro-1H-quinolin-2-one. The molecular formula is C23H27Cl2N3O2 and the molecular weight is 448.38538. Aripiprazole is a white to off-white crystalline powder. It is freely soluble in methylene chloride; sparingly soluble in toluene; insoluble in methanol and in water. The molecular structure is shown in Figure 34.30–6.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

After oral administration, pramipexole is rapidly absorbed with >90 percent bioavailability. Absorption is not significantly affected by food. It undergoes little first pass metabolism. The time to maximum concentration (T_{max}) is approximately 2 hours and the half-life ($T_{1/2}$) is approximately 8 hours (12 hours in elderly patients). It undergoes minimal metabolism with 90 percent of the drug being excreted unchanged in the urine.

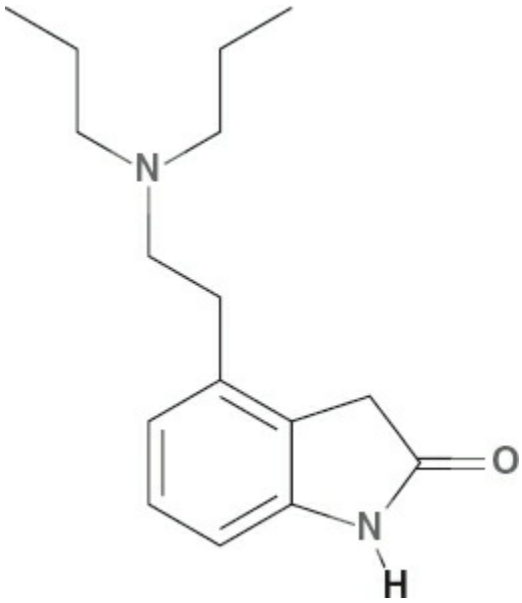

FIGURE 34.30–2. Molecular structure of ropinirole. (From National Center for Biotechnology Information. PubChem Compound Database; CID = 5095. http://pubchem.ncbi.nlm.nih.gov/compound/5095. Accessed May 13, 2015.)

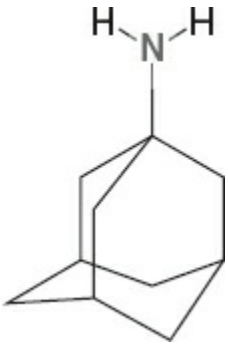

FIGURE 34.30–3. Molecular structure of amantadine. (From National Center for Biotechnology Information. PubChem Compound Database; CID = 2130. http://pubchem.ncbi.nlm.nih.gov/compound/2130. Accessed May 13, 2015.)

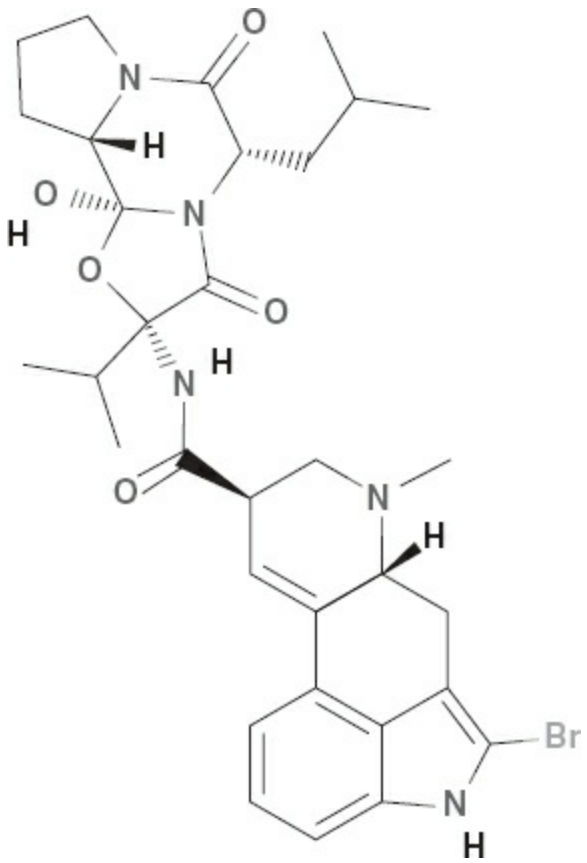

FIGURE 34.30–4. Molecular structure of bromocriptine. (From National Center for Biotechnology Information. PubChem Compound Database; CID = 31101, http://pubchem.ncbi.nlm.nih.gov/compound/31101. Accessed May 13, 2015.)

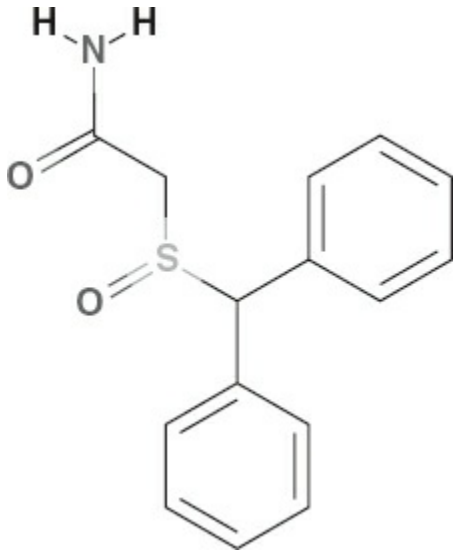

FIGURE 34.30–5. Molecular structure of modafinil (armodafinil). (From National Center for Biotechnology Information. PubChem Compound Database; CID = 4236, http://pubchem.ncbi.nlm.nih.gov/compound/4236. Accessed May 13, 2015.)

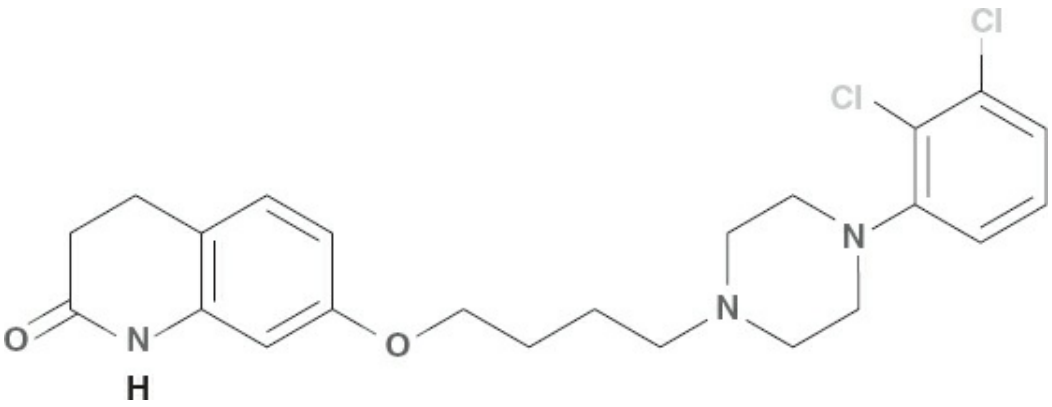

FIGURE 34.30–6. Molecular structure of aripiprazole. (From National Center for Biotechnology Information. PubChem Compound Database; CID = 60795, http://pubchem.ncbi.nlm.nih.gov/compound/60795. Accessed May 13, 2015.)

Ropinirole is rapidly absorbed after oral administration. Absorption is not significantly affected by food. Absolute bioavailability is 55 percent indicating significant first pass metabolism. The time to maximum concentration (T_{max}) is approximately 1 to 1.5 hours and half-life ($T_{1/2}$) is approximately 6 hours. Steady-state concentrations are expected within 2 days of dosing. It is primarily metabolized by CYP1A2 to inactive metabolites and excreted in the urine. Clearance is reduced by approximately 30 percent in patients above 65 years of age.

Amantadine is well absorbed after oral administration. The time to maximum concentration (T_{max}) is on average 3.3 hours and the half-life ($T_{1/2}$) is approximately 17 hours. It is primarily excreted unchanged in the urine. Clearance of amantadine is significantly reduced in patients with renal insufficiency.

The oral absorption rate of bromocriptine is approximately 28 percent. Absorption is not significantly affected by food though it is recommended that the drug be administered with food to reduce possible nausea/vomiting. Absolute bioavailability is 6 percent indicating a significant first pass effect. The time to maximum concentration (T_{max}) is approximately 1 to 2 hours and half-life ($T_{1/2}$) is 2.5 to 5 hours. It is primarily metabolized by CYP3A4 and excreted in the feces via biliary secretion; 6 percent is eliminated in the urine.

Modafinil is a racemic compound whose isomers have different kinetics. Armodafinil is the R-isomer of modafinil. Both isomers are rapidly absorbed after oral administration. The time to maximum concentration (T_{max}) is 2 to 4 hours for modafinil and 2 hours for armodafinil. Absorption is not significantly affected by food. The half-life of the R-isomer is approximately three times that of the S-isomer, so that exposure to the R-isomer is about three times that of the S-isomer. While the effective half-life ($T_{1/2}$) of both modafinil and armodafinil is approximately 11 to 15 hours after multiple doses, maximum plasma drug concentrations of armodafinil

remain higher for a longer period of time because the S-isomer is cleared more quickly than the R-isomer. Modafinil is predominantly metabolized by the liver via glucuronide conjugation (also CYP3A4/5 metabolism) and is excreted in the urine.

Aripiprazole is well absorbed after oral administration with a bioavailability of 87 percent. It can be administered with or without food. The time to maximum concentration (T_{max}) is 3 to 5 hours and the half-life ($T_{1/2}$) is 75 hours for extensive metabolizers and 146 hours for poor metabolizers. It is extensively metabolized by CYP2D6 and 3A4 isoenzymes and is primarily excreted in the urine.

Pharmacodynamics

Pramipexole and ropinirole are nonergot DAs with high specificity and full intrinsic activity at the D2 subtype receptors, with higher affinity binding at D3 (compared to D2 and D4) receptor subtypes. Both have been shown to stimulate dopamine receptors in the striatum. Pramipexole may also exert robust neurotropic effects. Ropinirole demonstrates moderate in vitro affinity for opioid receptors.

Amantadine is a nonergot adamantane derivative that may have direct and indirect effects on the dopamine neuron, including dopamine receptor agonist, serotonergic antagonist, and glutamatergic (*N*-methyl-D-aspartate [NMDA]) antagonist properties. The antiviral effects are thought to be mediated by interfering with the viral M2 protein. The mechanism of action for the treatment of PD and EPS is unknown, but may involve effects on the dopamine and/or NMDA system.

Bromocriptine is a semisynthetic ergotamine alkaloid with D2 and D3 receptor agonist activity. It is thought that postsynaptic D2 stimulation is primarily responsible for the antiparkinsonian effect while presynaptic D2 stimulation is associated with a neuroprotective effect. It also exhibits agonist activity at 5-HT1D, 5-HT1A, 5-HT2A, 5-HT1B, and 5-HT2C receptors and antagonist activity at α2A, α2C, α2B, and D1 receptors. It demonstrates partial agonist activity at 5-HT2B receptors and inactivated D4 and 5-HT7 receptors. In the tuberoinfundibular pathway, bromocriptine inhibits prolactin secretion by pituitary lactotrophs. Hallucinations caused by bromocriptine are thought to arise due to 5-HT2A agonism while pulmonary fibrosis may be associated with 5-HT1B and 5-HT2B activities.

Modafinil and its corresponding R-isomer, armodafinil, are oxazolidinone compounds that have complex neurotransmitter effects, including dopamine transporter blockade, dopamine reuptake inhibition, and possibly dopamine receptor agonist properties (including in the nucleus accumbens). Modafinil has been found to activate orexin (hypocretin)-containing neurons, which are located exclusively in the

lateral hypothalamus and send axons to numerous brain regions, including major nuclei implicated in sleep regulation. Animal studies have found that modafinil increases extracellular glutamate levels in the medial preoptic area and posterior hypothalamus, as well as decreasing γ-aminobutyric acid (GABA) levels in these areas. Evidence suggests that modafinil acts robustly on catecholamines, glutamate, GABA, serotonin, orexin, and histamine systems in the brain.

Aripiprazole is a second-generation antipsychotic benzisoxazole derivative that has high affinity for D2 and 5HT2 receptors. It demonstrates partial agonist properties at the D2 receptor with antagonist effects in the mesolimbic pathway and partial agonist activity in D2 mesocortical pathway. It is thought that antipsychotic effects are primarily mediated via D2 partial agonist effects. It also demonstrates partial agonist effects at the 5-HT1A receptor and antagonist effects at 5-HT2A. It also has moderate affinity for alpha-adrenergic and H1 histaminic receptors.

CLINICAL DRUG STUDIES

Major Depressive Disorder

Several lines of evidence suggest that the dopaminergic system may be an effective target for the treatment of depression. Depressed patients demonstrated decreased levels of dopamine metabolites (homovanillic acid —HVA) in cerebrospinal fluid and blood. Though results have been mixed, some SPECT studies have indicated higher in vivo availability of D2 receptor binding in depressed subjects compared to controls. In postmortem brains, lower levels of dopamine transporter binding were found in basal and central amygdaloid nuclei, whereas increased D2/3 receptor binding was found in basal, central, and lateral amygdaloid nuclei in major depression compared with controls. This finding supports the theory that a deficiency in mesolimbic dopamine may result in certain associated symptoms of depression (particularly anhedonia and loss of motivation). Finally, certain prodopaminergic agents such as bupropion (a dopamine reuptake inhibitor) have demonstrated efficacy and been approved for the treatment of MDD.

Pramipexole has been the most studied DA for the treatment of MDD. Several open-label studies have reported the efficacy of pramipexole in augmenting tricyclic antidepressants and SSRIs in TRD. One small open-label trial of 10 patients with treatment-resistant MDD (who had failed at least 2 prior antidepressant trials) demonstrated that pramipexole augmentation lead to significant improvement in depressive symptoms from baseline to posttreatment, with 6 subjects meeting criteria for treatment response (i.e., ≥50 percent reduction in Montgomery-Asberg Depression Rating Scale [MADRS] score). A randomized double-blind placebo-controlled trial of pramipexole augmentation in 60 treatment-

resistant MDD outpatients demonstrated modest but statistically significant benefit (p = 0.038). In another double-blind, comparative study, 39 subjects who had failed a standard antidepressant trial were randomized to receive pramipexole, escitalopram, or the combination. After 6 weeks of treatment, subjects receiving pramipexole monotherapy demonstrated significantly lower MADRS scores than the combination group (p = 0.01). In the combination group, however, only 15 percent of subjects tolerated regularly scheduled increases in pramipexole. This study demonstrated that concurrent initiation and titration of SSRI and pramipexole resulted in side effects often leading to cessation of treatment. Adjunctive use of pramipexole appeared to be better tolerated when it was added to ongoing SSRI treatment.

Ropinirole has also been studied as an augmentation strategy for standard antidepressants. An open-label trial of 10 subjects (7 with treatment-resistant MDD) demonstrated a statistically significant 4.5-point decrease in MADRS score from baseline to after 16 weeks of treatment. Another study demonstrated the effectiveness of ropinirole for antidepressant-induced sexual dysfunction.

Modafinil augmentation for MDD has also been extensively studied. A meta-analysis of six randomized controlled trials (four MDD trials) demonstrated that modafinil augmentation was effective for acute MDD, including for symptoms of fatigue, and was well tolerated.

Aripiprazole has an FDA-approved indication for the adjunctive treatment of MDD. Approval for this indication was based on two positive 6-week placebo-controlled trials (n = 381, n = 362) of treatment-resistant MDD (in the current episode) that demonstrated the drug was superior to placebo in reducing MADRS scores. In a meta-analysis of atypical antipsychotic treatment for MDD, aripiprazole demonstrated a statistically significant increased likelihood of achieving remission (OR, 2.01) compared to placebo, but was also associated with adverse events (akathisia, sedation, and weight gain) Table 34.30–1.

Bipolar Depression

There is growing support for the use of pramipexole in bipolar depression. Several case reports, case series, and open-label trials have reported the efficacy of pramipexole in augmenting tricyclic antidepressants and SSRIs in treatment-resistant depression; approximately half of the reported patients were diagnosed with bipolar disorder. One randomized controlled trial found that pramipexole was associated with superior response compared with placebo (60 percent vs. 9 percent) for bipolar II depression when added to therapeutic levels of lithium or valproate. Overall, pramipexole was found to have a large effect size (0.6 to 1.1) in the treatment of both MDD and bipolar depression, and was associated with a low switch rate (1 percent mania, 5 percent hypomania). In a functional

neuroimaging study of 15 bipolar II depressed subjects treated with 6 weeks of pramipexole augmentation of mood stabilizers versus placebo, pramipexole treatment was associated with reductions in normalized metabolism in bilateral orbitofrontal cortex, left ventrolateral prefrontal cortex, and right anteromedial prefrontal cortex. Voxel-wise analyses additionally showed decreased normalized metabolism in the left inferior parietal cortex and medial frontopolar cortical (BA 10P) area of the anteromedial prefrontal cortex.

Ropinirole has also been reported to be effective in the augmentation of standard antidepressants. Limited evidence supports its off-label efficacy for augmenting antidepressant response in patients with treatment-resistant bipolar depression or for alleviating antidepressant-induced sexual dysfunction. In one open-label trial, ropinirole was used in combination with lithium for acutely depressed bipolar patients and demonstrated a significant treatment effect. No switching into hypomania or mania was observed.

Table 34.30–1.
Management of Adverse Events of Sympathomimetic Drugs

Adverse Event	Management
For all adverse events	Unless severe, allow 7–10 days for tolerance to develop
	Evaluate dose–response relationships
	Evaluate time–action effects and then adjust dosing intervals or switch to sustained-release preparation
	Evaluate for concurrent conditions, including comorbidities and environment stressors
	Consider alternative treatments
Abuse/Dependence	Strict medication supervision
	Consider switching to lisdexamfetamine, atomoxetine, or other drug
	Abuse/dependence evaluation and treatment
	Discontinue stimulant treatment
Cardiovascular	Monitor blood pressure and pulse
Including: hypertension, tachycardia, arrhythmias, myocardial ischemia/infarction	Cardiac evaluation, including 12-lead ECG
	Lower dose
	Consider switch to other stimulant, atomoxetine
Anorexia or dyspepsia	Administer before, during, or after meals
Weight loss	Give drug after breakfast and after lunch
	Implement calorie-enhancement strategies
	Give brief drug holidays
Slowed growth	Apply weight loss remedies
	Give weekend and vacation (longer) drug holidays
	Consider another stimulant or nonstimulant drug
Dizziness	Monitor blood pressure and pulse
	Encourage adequate hydration
	If associated with only T_{max}, change to sustained release

	preparation
Insomnia or nightmares	Administer earlier in day
	Omit or reduce last dose
	If giving sustained preparation, switch to tablet drug
	Consider adjunctive antihistamine or clonidine
Dysphoric mood or emotional constriction	Reduce dose or switch to long-acting preparation
	Switch stimulants
	Consider comorbidity requiring alternative or adjunctive treatment
Rebound	Switch to sustained-release preparation
	Combine long- and short-acting preparations
Tics	Firmly establish correlation between tics and pharmacotherapy by examining dose–response relationship, including no-medication condition
	If tics are mild and abate after 7–10 days with medications, reconsider risks and benefits of continued stimulant treatment and renew informed consent
	Switch stimulants
	Consider nonstimulant treatment (e.g., clonidine or tricyclic antidepressant)
	If tic disorder and attention-deficit/hyperactivity disorder are severe, consider combining stimulant with high-potency neuroleptic drug
Psychosis	Discontinue stimulant treatment
	Assess for comorbid thought disorder
	Consider alternative treatments

T_{max}, time of maximal concentration.

Adapted from Creenhill LL, Halperini M, March JS. Psychostimulants. In: Tasman A, Kay J, Lieberman JA, eds. *Psychiatry*. Philadelphia, PA: WB Saunders; 1997.

Modafinil and armodafinil have been studied in the treatment of bipolar depression. A meta-analysis of six randomized controlled trials (two bipolar depression trials) demonstrated that modafinil augmentation was effective for acute bipolar depression, including for symptoms of fatigue, and was well tolerated. One placebo-controlled trial found that modafinil was effective as monotherapy for bipolar depression. A randomized, double-blind, placebo-controlled, multicenter trial found that armodafinil is a safe and effective adjunctive treatment for bipolar depression.

While aripiprazole has an FDA-approved indication for the adjunctive treatment of MDD, it has been widely studied for the treatment of bipolar depression as well. Despite fairly common use as an adjunctive strategy for acute bipolar depression and two initial positive uncontrolled open-label reports, two placebo-controlled trials were negative. Other analyses suggest that aripiprazole has not demonstrated efficacy in acute monotherapy or adjunctive therapy, but may be effective for maintenance treatment of bipolar depression.

Mood and Fatigue Related to PD (and Other Medical Conditions)

According to some estimates, depressive symptoms may occur in up to 45 percent of patients with PD. Depressive symptoms are an important factor affecting the quality of life of PD patients. Given the dysfunction of the dopaminergic nigrostriatal pathways in PD, it has been thought that dopaminergic mesolimbic reward pathways may be similarly affected, resulting in depressive symptoms such as anhedonia or lack of motivation. Several studies have supported the efficacy of DAs in treating depressive symptoms associated with PD. In a large open-label trial of 657 patients with depressive symptoms associated with PD, anhedonia and depression were reduced with adjunctive treatment of L-dopa and pramipexole. In addition, in a study comparing pramipexole and sertraline, pramipexole demonstrated superior benefit in the treatment of depression associated with PD, and was also reported to reduce anhedonia in PD patients. Another randomized double-blind placebo-controlled 12-week trial of 323 mild–moderate PD patients demonstrated that pramipexole was associated with greater decrease in depression scores (on Beck Depression Inventory) compared to placebo ($p = 0.01$). A meta-analysis of the effect of pramipexole on mood and motivation symptoms in PD concluded from seven randomized controlled trials that the drug had a beneficial effect on mood and motivational symptoms in PD patients who did not have MDD. However, a systematic review of the literature (by the same authors) concluded that given the quality of the data, further studies are required to determine if DAs are effective to treat depressive symptoms in PD.

DA agonists have also been used in an attempt to improve cognitive recovery and rehabilitation in traumatic brain injury patients. In a systematic review of methylphenidate, amantadine, and bromocriptine, however, no positive trend could be established for any of the drugs.

Armodafinil has been studied as a treatment for sleepiness and fatigue related to PD and HIV. In one study, armodafinil effectively treated 10 PD patients with excessive daytime sleepiness and was noted to increase tolerance for higher doses of dopamine agonists, which may cause sleepiness as a side effect. Also, a placebo-controlled, randomized, double-blind trial of 70 HIV-positive subjects found that armodafinil successfully reduced HIV-related fatigue. Although armodafinil did not improve depressive symptoms in the absence of fatigue improvement, it did improve depressive symptoms in those with an axis I depressive disorder who experienced an increase in energy.

Restless Leg Syndrome

Pramipexole and ropinirole have been FDA-approved for use in treating moderate-to-severe primary RLS. Pramipexole demonstrated effectiveness in four randomized controlled trials (two fixed dose studies, one flexible

dose study, and one treatment withdrawal with survival analysis) across approximately 1,000 subjects. Ropinirole demonstrated effectiveness in three randomized, double-blind placebo-controlled studies involving over 500 subjects.

Extrapyramidal Symptoms

Amantadine (Symmetrel) is FDA-approved for the treatment of drug-induced EPS. The mechanism of action of this effect is unclear. A double-blind study of amantadine compared to benztropine for drug-induced EPS; amantadine was found to be comparable in effect to benztropine, but with fewer side effects.

Neuroleptic Malignant Syndrome

There are no currently approved drug treatments for NMS. Treatment generally consists of immediate removal of the presumed offending agent as well as any other dopamine antagonists, and supportive care. Though data are limited, pharmacological considerations may include benzodiazepines (for milder cases), DAs, dantrolene, and electroconvulsive therapy. Bromocriptine (starting at 2.5 mg b.i.d.–t.i.d., maximum total daily dose 45 mg) and amantadine (200 to 400 mg/day in divided doses) are DAs that may enhance presynaptic release of dopamine to treat the presumed underlying cause of NMS. They may be particularly helpful if NMS was caused by withdrawal of medication to treat PD. However, bromocriptine may induce or exacerbate serotonin syndrome so this diagnosis should be clearly ruled out.

Excessive Daytime Sleepiness

Modafinil has been approved for use in excessive daytime sleepiness secondary to narcolepsy, shift work sleep disorder, and OSA. Two large placebo-controlled, randomized, double-blind parallel group studies found modafinil to be effective and well tolerated in the treatment of excessive daytime sleepiness secondary to narcolepsy, with effects being maintained for at least 40 weeks. Modafinil was also effective in 10 patients with excessive daytime sleepiness secondary to head injury and two patients with excessive somnolence due to sedating drugs at doses of 400 mg; no side effects or evidence of loss of effect or tolerance were noted over 5 to 13 months of continued use. Another open-label study of 72 patients with fatigue associated with multiple sclerosis found that 200 mg per day of modafinil was well tolerated and significantly improved fatigue.

Armodafinil has also been approved to treat excessive daytime sleepiness secondary to narcolepsy, shift work sleep disorder, and OSA/hypopnea syndrome. One large, placebo-controlled, multicenter trial found that armodafinil significantly improved ability to remain awake

compared to placebo in individuals with narcolepsy and shift work sleep disorder. In addition, two large, placebo-controlled trials showed that armodafinil significantly improved ability to remain awake compared to placebo in subjects with OSA/hypopnea syndrome.

Performance Enhancement

In schizophrenia, modafinil and armodafinil have been studied as an augmentation strategy targeting negative symptoms, fatigue/sleepiness, and cognition. A systematic review and meta-analysis concluded that both agents were well tolerated and reduced negative symptoms, though the effect size was small. Their effects on fatigue/sleepiness and cognition have been less clear. Various studies have investigated the ability of these agents to enhance certain aspects of cognition (i.e., alertness, attention, executive functioning, memory) in conditions such as substance abuse/dependence, neurological conditions (e.g., multiple sclerosis, dementia, brain tumors, stroke), and cancer.

While the traditional goal of medicine has been to ameliorate pathological conditions, modafinil and armodafinil have increasingly been used to enhance normal human functioning. From initial use to increase alertness and physical performance in military personnel (sometimes taking the place of psychostimulants), off-label use of these agents is expanding to enhance performance in a variety of settings; reports have noted their use by military/rescue personnel, athletes, academicians, university students, and video-game participants. The development and use of these agents as performance enhancers in nonpathological conditions has prompted a public discussion of ethical questions such as: who should control access to these agents, what risks accompany their use, how does the absence of disease affect the benefit risk assessment, what are the appropriate parameters of their use, is use fair in competitive settings, what are the limits to the enhancement effects, and how may enhancement effects influence human experience.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Central Nervous System

DAs display prominent central nervous system effects. From their initial development as treatment for PD, the prodopaminergic activity of the classic DAs underlies their effectiveness for RLS, mood disorders, EPS, and NMS. Central prodopaminergic activity also mediates many of the adverse events, including depression, irritability, confusion, hallucinations, dyskinesias, and impulsive/compulsive behavior. Interestingly, while administration of aripiprazole may be associated with akathisia, tardive dyskinesia, and NMS, full DAs may be used to treat NMS, and abrupt discontinuation of certain DAs may induce parkinsonian/extrapyramidal

and NMS-like symptoms. The increased wakefulness, energy, initiative, self-confidence, positive mood, and enhanced physical and mental performance associated with the wakefulness-promoting agents, modafinil and armodafinil, can often be dramatic and of rapid onset. As a result, the potential for abuse and the danger of dependence or addiction must be kept in mind. Finally, the partial dopamine agonist/antagonist properties of aripiprazole and cariprazine are powerful mechanisms for the treatment of mood and psychotic disorders. With aripiprazole, elderly patients may be at increased risk of cerebrovascular events.

Cardiovascular System

The entire class of ergot DA agonists has been associated with cardiac fibrosis. As a result, pergolide was withdrawn from the US market in 2007 because of an increased risk of valvular heart disease. Cabergoline remains on the market, but the FDA issued a warning in 2011 for the risk of cardiac valvulopathy. Bromocriptine has been associated with pericardial and pulmonary fibrosis particularly with long-term high-dose treatment, and increased risk of myocardial infarction in postpartum women. Congestive heart failure in conjunction with amantadine administration has also been reported. In general, most DAs can produce hypotension.

Modafinil and armodafinil are less commonly associated with cardiovascular adverse events.

Immunological Effects

Modafinil and armodafinil have been associated with certain hypersensitivity reactions, particularly in pediatric patients. 0.8 percent of pediatric subjects in premarket studies developed a rash resulting in drug discontinuation; one case proceeded to Stevens–Johnson syndrome while another developed multiorgan hypersensitivity reaction. There have also been one case of angioedema reported with modafinil, and one case of hypersensitivity, and one case of angioedema reported with armodafinil. Based on the possibility of hypersensitivity reactions, modafinil did not receive FDA approval for the treatment of ADHD, despite good effectiveness data.

Endocrine Effects

Aripiprazole may be associated with hyperglycemia and diabetes mellitus. It is recommended that blood glucose should be monitored during the time of administration.

Hematological Effects

On rare occasions, aripiprazole, as well as other second-generation antipsychotic drugs, may be associated with leukopenia, neutropenia, and

agranulocytosis. It is recommended that blood counts be closely monitored, especially during the first few months of administration.

THERAPEUTIC IMPLICATIONS

Pramipexole and ropinirole are FDA-approved for use in PD and RLS, and have been studied as an adjunctive treatment for unipolar and bipolar depression. Amantadine is FDA approved to treat PD, drug-induced EPS, and influenza A infection. It is also used off-label to treat NMS and Cotard syndrome, a psychiatric condition that involves a delusional belief that one is dead. Bromocriptine is FDA approved for use in PD, galactorrhea (due to hyperprolactin conditions) and acromegaly, and is used off-label to treat RLS and NMS. Cabergoline is FDA approved for the treatment of hyperprolactinemic disorders, either idiopathic or due to pituitary adenomas. In addition, a number of studies have suggested that cabergoline may be effective in treating PD or preventing ovarian hyperstimulation syndrome. Modafinil and armodafinil have been FDA approved to improve wakefulness in adult patients with excessive sleepiness associated with narcolepsy, OSA, or SWD. They have been studied as an adjunctive treatment for unipolar and bipolar depression, ADHD, and treatment of fatigue and cognitive dysfunction associated with medical conditions. They also have been used in nonpathological settings to optimize normal performance. Aripiprazole is FDA approved for the treatment of schizophrenia, acute manic/mixed episodes of bipolar I disorder, and irritability associated with autistic disorder in children and adolescents, and as an adjunctive treatment for MDD. Cariprazine, a partial D2/3 agonist (with D3 selectivity), was approved in 2015 by the FDA for the acute treatment of manic and mixed episodes associated with bipolar I disorder and the treatment of schizophrenia. It is also under investigation as an adjunctive treatment for MDD.

PRECAUTIONS AND ADVERSE REACTIONS

Each class of DAs has characteristic side effects that should be considered and discussed with patients before they are prescribed. Table 34.30–2 lists common and potentially serious adverse events associated with each drug.

Common side effects of the nonergot drugs, pramipexole and ropinirole, include somnolence, orthostatic hypotension, dyskinesias, and hallucinations. Use of these drugs has been associated with the possibility of sudden onset of sleep in the midst of daily activities, including driving. In addition, rare cases of impulsive/compulsive behavior (i.e., pathological gambling, hypersexuality, compulsive eating [including binge eating] and increased spending) have been reported. Some of these side effects may be dose related, while some may be related to individual susceptibility. One meta-analysis of nonergot dopamine agonists in PD found increased incidence of nausea and other symptoms compared to placebo. In addition,

a greater incidence of side effects such as dizziness, somnolence, constipation, vomiting, and insomnia were only found in early PD, while higher risks of dyskinesia and hallucination were only found in advanced PD. Pramipexole and ropinirole are classified as pregnancy risk C.

Common side effects of amantadine include nausea, dizziness, and insomnia. Psychiatric side effects such as depression, anxiety, irritability, hallucinations, and confusion have also been reported. Care should be taken in patients with seizure disorders as amantadine may increase seizure activity. Several cases of congestive heart failure in conjunction with amantadine administration have also been reported. Deaths from overdose using amantadine have also occurred with the lowest reported acute lethal dose of 1 g. When the drug is tapered or discontinued, care should be taken to decrease the dose gradually. Abrupt discontinuation may precipitate a parkinsonian crisis. In addition, sporadic cases of possible NMS have been reported. Amantadine is classified as pregnancy risk C.

Common side effects of bromocriptine include nausea, headache, dizziness, fatigue, light-headedness, vomiting, abdominal cramps, nasal congestion, constipation, diarrhea, drowsiness. Somnolence and symptomatic hypotension may occur especially in PD populations. Episodes of sudden sleep onset may occur and PD patients taking or prescribed bromocriptine should be cautioned not to drive or operate machinery. Bromocriptine use in pregnancy is contraindicated because the safety to mother and fetus has not been established. In addition, rare instances of myocardial infarction, seizures (including status epilepticus), and strokes have been reported in postpartum women receiving bromocriptine. Like other ergot derivatives, bromocriptine increases the risk of fibrotic complications. Pleural and pericardial effusions, as well as pleural and pulmonary fibrosis and constrictive pericarditis, and retroperitoneal fibrosis have been reported, especially with long-term, high-dose treatment. In PD patients, there is an increased risk of psychiatric complications such as confusion, mental disturbance, hallucinations, impulse control (with intense urges to gamble, increased sexual urges, intense urges to spend money uncontrollably and other intense urges). Hallucinations typically resolve with dose reduction. Caution should be taken in PD patients with a history of myocardial infarction who have residual, atrial, nodal, or ventricular arrhythmia. In acromegalic patients, there have been reports of cold-sensitive digital vasospasm, possible tumor expansion, and severe GI bleeding from peptic ulcers (with some fatalities). Bromocriptine is classified as pregnancy risk B.

Table 34.30–2.

Adverse Events of Dopamine Receptor Agonist Drugs

Drug	Common Adverse Events	Other Potential Serious Adverse Events
Pramipexole Pregnancy risk: C Ropinirole Pregnancy risk: C	Dizziness, somnolence, orthostatic hypotension, dyskinesias and hallucinations, constipation, vomiting, and insomnia	Episodes of sudden sleep onset, impulsive/compulsive behavior (i.e., pathological gambling, hypersexuality, compulsive eating (including binge eating) and increased spending)
Amantadine Pregnancy risk: C	Nausea, dizziness, and insomnia, depression, anxiety, irritability, hallucinations, and confusion	Congestive heart failure, increase in seizure activity Abrupt discontinuation may result in parkinsonian crisis, NMS
Bromocriptine (ergot) Pregnancy risk: B	Nausea, headache, dizziness, fatigue, light-headedness, vomiting, abdominal cramps, nasal congestion, constipation, diarrhea, drowsiness, confusion, mental disturbance, hallucinations, impulse control (intense urges to gamble, increased sexual urges, intense urges to spend money uncontrollably and other intense urges) Somnolence, hypotension (in PD)	Pleural/pericardial effusion, pleural/pulmonary fibrosis, constrictive pericarditis, retroperitoneal fibrosis Episodes of sudden sleep onset (PD) Cold sensitive digital vasospasm, tumor expansion, GI bleed from peptic ulcer (acromegaly)
Modafinil Pregnancy risk: C Armodafinil Pregnancy risk: C	Headache, nausea, nervousness, rhinitis, diarrhea, back pain, anxiety, insomnia, dizziness, and dyspepsia, mania, delusions, hallucinations, suicidal ideation, and aggression	Stevens–Johnson syndrome, multiorgan hypersensitivity reaction, angioedema Substance abuse
Aripiprazole Pregnancy risk: C	Nausea, vomiting, constipation, headache, dizziness, akathisia, anxiety, insomnia, and restlessness Somnolence, extrapyramidal symptoms, increased appetite, nasopharyngitis, and weight increase (pediatric patients)	Increased mortality (dementia-related psychosis in elderly patients) Increased risk of suicidal thinking/behavior (children, adolescents, and young adults) Cerebrovascular events (elderly) NMS, tardive dyskinesia, hyperglycemia, diabetes mellitus, orthostatic hypotension, seizures, cognitive and motor impairment, dysphagia, body temperature dysregulation, leukopenia, neutropenia, agranulocytosis

Modafinil and armodafinil are generally well-tolerated drugs. Common side effects include headache, nausea, nervousness, rhinitis, diarrhea, back pain, anxiety, insomnia, dizziness, and dyspepsia. Psychiatric

complications of mania, delusions, hallucinations, suicidal ideation, and aggression have been reported. There have also been some reports of abuse, however, the abuse potential of this class of drugs is significantly lower than with psychostimulants. In pediatric patients, a rash resulting in drug discontinuation occurred in 0.8 percent of patients (i.e., 13 of 1,585 subjects). Of these cases, one proceeded to Stevens–Johnson syndrome and there was also one report of multiorgan hypersensitivity reaction. The median time to rash was 13 days. No rashes or Stevens–Johnson syndrome have been observed in adults. In postmarket surveillance, modafinil has been associated with one case of angioedema and armodafinil has been associated with one case of hypersensitivity and one case of angioedema. Modafinil and armodafinil are classified as pregnancy risk C.

Aripiprazole has two black box warnings associated with its use. The first black box warning states that it is not approved for dementia-related psychosis in elderly patients due to an increased risk of mortality. The second black box warning states that use in children, adolescents, and young adults with depression may increase the risk of suicidal thinking and behavior. In adult patients, common side effects of aripiprazole include nausea, vomiting, constipation, headache, dizziness, akathisia, anxiety, insomnia, and restlessness. Pediatric patients may also experience somnolence, EPS, increased appetite, nasopharyngitis, and weight increase. Elderly patients may be at increased risk of cerebrovascular events. NMS has been reported with aripiprazole use; if this should occur, the drug should be immediately discontinued and the patient should be closely monitored. Tardive dyskinesia, hyperglycemia (and diabetes mellitus), orthostatic hypotension, seizures, cognitive and motor impairment, dysphagia, and body temperature dysregulation have also been reported. There is also a potential for the development of leukopenia, neutropenia, and agranulocytosis with the use of aripiprazole. Blood counts should be monitored, especially in the first few months of use. Aripiprazole is classified as pregnancy risk C.

DRUG INTERACTIONS

Pramipexole is not significantly metabolized by CYP enzymes and does not have significant effects on the CYP enzymes. It is renally cleared. Coadministration of cimetidine, an inhibitor of renal tubular secretion, may increase pramipexole C_{max} by 40 percent and half-life ($T_{1/2}$) 40 percent.

Ropinirole is an inhibitor of CYP1A2, therefore coadministration with CYP1A2 substrates may result in higher concentrations of the substrate drug. For instance, in one study, coadministration with L-dopa increased C_{max} by 20 percent.

Bromocriptine is a CYP3A4 substrate. Therefore, CYP3A4 inducers such

as macrolide antibiotics, azole antimycotics, and HIV protease inhibitors may increase serum levels. In one study, coadministration of bromocriptine and erythromycin increased bromocriptine levels by a factor of four. For all DA's coadministration with dopamine antagonists may result in decreased levels and decreased effect.

Amantadine is primarily excreted unchanged in the urine, therefore, clearance is significantly reduced in patients with renal insufficiency. Coadministration of amantadine with quinine or quinidine may reduce renal clearance of amantadine by up to 30 percent. When amantadine is administered with anticholinergics, anticholinergic-like side effects of amantadine may be potentiated. Coadministration of thioridazine with amantadine may worsen tremor in elderly patients with PD. Use of amantadine may potentially interfere with intranasal live attenuated influenza vaccine (though this has not been evaluated). However, there is no interaction between amantadine and trivalent inactivated influenza vaccine.

Modafinil and armodafinil are reversible moderate inhibitors of cytochrome P450 (CYP) 2C19, a drug-metabolizing enzyme, and may increase blood levels of diazepam (Valium), phenytoin, propranolol (Inderal), omeprazole, and clomipramine. In individuals deficient in the enzyme CYP2D6 (~7 to 10 percent of the Caucasian population and similar or lower for other populations), drugs such as tricyclic antidepressants, SSRIs, and bupropion may be increased. Chronic administration of modafinil may lead to a modest induction of CYP3A4 (armodafinil is associated with moderate enzyme induction), which may reduce levels of steroid contraceptives, triazolam, cyclosporine (Neoral), and theophylline (Bronkodyl), and dose adjustments of these medications may be required. Neither modafinil nor armodafinil have been studied in combination with MAOI medications; caution should thus be used when concomitantly administering MAOIs and these drugs. Because modafinil and armodafinil are partially cleared by CYP3A enzymes, coadministration of potent CYP3A4/5 inducers (e.g., carbamazepine, phenobarbital, and rifampin) or inhibitors (e.g., ketoconazole, erythromycin) may significantly alter drug plasma levels.

Table 34.30–3.
Dosing of Dopamine Receptor Agonist Drugs for Neuropsychiatric Indications

Drug	Indication	Dose
Pramipexole	PD	0.125–1.5 mg t.i.d. (0.375–4.5 mg total daily dose)
	RLS	0.125–0.5 mg (2–3 hours before bedtime)
	Depression (investigational)	0.375–3 mg (total daily dose)
Ropinirole	PD	0.25–1 mg t.i.d. (0.75–3 mg total daily dose)
	RLS	0.25–4 mg at bedtime

	Depression (investigational)	Up to 200 mg total daily dose
Amantadine	PD	100–400 mg total daily dose
	EPS	100–300 mg total daily dose
	NMS (off-label)	200–400 mg total daily dose
Bromocriptine	PD	1.25–100 mg total daily dose
	NMS (off-label)	2.5 mg b.i.d.–15 mg t.i.d. (5–45 mg total daily dose)
Modafinil	All indications	200–400 mg total daily dose
Armodafinil	All indications	150–250 mg total daily dose
Aripiprazole	Adult indications	10–30 mg total daily dose
	Pediatric/adolescent indications	2–30 mg total daily dose

Aripiprazole is metabolized by CYP3A4 and CYP2D6. Strong 3A4 inhibitors (e.g., ketoconazole) and 2D6 inhibitors (e.g., fluoxetine, quinidine) may reduce aripiprazole levels by up to 50 percent. Strong 3A4 inducers (e.g., carbamazepine) will decrease aripiprazole levels and higher doses of aripiprazole may be considered in these situations (e.g., use double the usual dose). Aripiprazole is unlikely to affect levels of other coadministered drugs.

LABORATORY INTERFERENCE

No significant laboratory interference has been reported.

DOSAGE AND ADMINISTRATION

Typical doses of medications for given indications are presented in Table 34.30–3.

For pramipexole, the recommended dose range for PD is 0.125 mg t.i.d. (total daily dose 0.375 mg) to 1.5 mg t.i.d. (total daily dose 4.5 mg). The recommended dose range for RLS is 0.125 mg daily (2 to 3 hours before bedtime) to 0.5 mg daily. For adjunctive treatment of depression, studies have used 0.375 to 3 mg (total daily dose).

For ropinirole, the recommended dose range for PD is 0.25 t.i.d. (total daily dose 0.75 mg) to 1 mg t.i.d. (total daily dose 3 mg). The recommended dose range for RLS is 0.25 mg to 4 mg at bedtime. For adjunctive treatment of depression, studies have investigated doses up to 200 mg total daily dose.

For amantadine, the recommended dose range for PD 100 to 400 mg total daily dose. For EPS, the recommended dose range is 100 to 300 mg total daily dose. For influenza A, the recommended adult dose range is 100 to 200 mg total daily dose. For pediatric patients (1 to 9 years of age), the total daily dose should be calculated on the basis of 2 to 4 mg/lb/day (4.4 to 8.8 mg/kg/day), not to exceed 150 mg total daily dose. For NMS, off-label use has typically administered 200 to 400 mg total daily dose.

For bromocriptine, the recommended dose range for hyperprolactinemia is 1.25 to 15 mg total daily dose. For acromegaly, the recommended dose range is 1.25 to 100 mg total daily dose. For PD, the

recommended dose range is 1.25 to 100 mg total daily dose. For NMS, off-label use has typically administered 5 mg (2.5 mg b.i.d./t.i.d.) to 45 mg total daily dose.

Dose ranges for modafinil are 200 to 400 mg total daily dose and armodafinil are 150 to 250 mg total daily dose.

For aripiprazole, recommended adult dose ranges are 10 to 30 mg total daily dose, and 2 to 30 mg total daily dose for pediatric/adolescent patients.

REFERENCES

Aiken CB. Pramipexole in psychiatry: a systematic review of the literature. *J Clin Psychiatry*. 2007;68(8):1230–1236.

Barone P, Poewe W, Albrecht S, et al. Pramipexole for the treatment of depressive symptoms in patients with Parkinson's disease: a randomized, double-blind, placebo-controlled trial. *Lancet Neurol*. 2010;9(6):573–580.

Barone P, Scarzella L, Marconi R, et al; Depression/Parkinson Italian Study Group. Pramipexole versus sertraline in the treatment of depression in Parkinson's disease: a national multicenter parallel-group randomized study. *J Neurol*. 2006;253(5):601–607.

*Beaulieu JM, Gainetdinov RR. The physiology, signaling and pharmacology of dopamine receptors. *Pharmacol Rev*. 2011;63(1):182–217.

Brooks DJ. Dopamine agonists: their role in the treatment of Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 2000;68(6):685–689.

Calabrese JR, Frye MA, Yang R, Ketter TA. Armodafinil Treatment Trial Study Network. Efficacy and safety of adjunctive armodafinil in adults with major depressive episodes associated with bipolar I disorder: a randomized, double-blind, placebo-controlled multicenter trial. *J Clin Psychiatry*. 2014;75(10):1054–1061.

Cassano P, Lattanzi L, Fava M, et al. Ropinirole in treatment-resistant depression: a 16-week pilot study. *Can J Psychiatry*. 2005;50(6):357–360.

Cole MG, Elie LM, McCusker J, Bellavance F, Mansour A. Feasibility and effectiveness of treatments for post-stroke depression in elderly inpatients: systematic review. *J Geriatr Psychiatry Neurol*. 2001;14(1):37–41.

Cusin C, Ioveieno N, Iosifescu DV, et al. A randomized, double-blind, placebo-controlled trial of pramipexole augmentation in treatment-resistant major depressive disorder. *J Clin Psychiatry*. 2013;74(7):e636–e641.

*Dell'Osso B, Ketter TA. Assessing efficacy/effectiveness and safety/tolerability profiles of adjunctive pramipexole in bipolar depression: acute versus long-term data. *Int Clin Psychopharmacol*. 2013;28(6):297–304.

DiMascio A, Bernardo CL, Greenblatt DJ, Marder JE. A controlled trial of amantadine in drug-induced extrapyramidal disorders. *Arch Gen Psychiatry*. 1976;33(5):599–602.

*Franco-Chaves FA, Mateus CF, Luckenbaugh DA, Martinez PE, Mallinger AG, Zarate CA Jr. Combining a dopamine agonist and selective serotonin reuptake inhibitor for the treatment of depression: a double-blind randomized pilot study. *J Affect Disord*. 2013;149:319–325.

Frenette AJ, Kanji S, Rees L, et al. Efficacy and safety of dopamine agonists in traumatic brain

injury: a systematic review of randomized controlled trials. *J Neurotrauma*. 2012;29:1–18.

Goldberg JF, Burdick KE, Endick CJ. Preliminary randomized, double-blind, placebo-controlled trial of pramipexole added to mood stabilizers for treatment-resistant bipolar depression. *Am J Psychiatry*. 2004;161(3):564–566.

*Goss AJ, Kaser M, Costafreda SG, Sahakian BJ, Fu CH. Modafinil augmentation therapy in unipolar and bipolar depression: a systematic review and meta-analysis of randomized controlled trials. *J Clin Psychiatry*. 2013;74(11):1101–1107.

Inoue T, Kitaichi Y, Masui T, et al. Pramipexole for stage 2 treatment-resistant major depression: an open study. *Prog Neuro-psychopharmacol Biol Psychiatry*. 2010;34(8):1446–1449.

*Klimek V, Schenck JE, Han H, Stockmeier CA, Ordway GA. Dopaminergic abnormalities in amygdaloid nuclei in major depression: a postmortem study. *Biol Psychiatry*. 2002;52(7):740–748.

Kulkarni SK, Dhir A. Current investigational drugs for major depression. *Expert Opin Investig Drugs*. 2009;18(6):767–788.

*Lam YW. Clinical pharmacology of dopamine agonists. *Pharmacotherapy*. 2000;20(1 Pt 2):17S–25S.

Leentjens AF. The role of dopamine agonists in the treatment of depression in patients with Parkinson's disease. *Drugs*. 2011;71(3):273–286.

Leentjens AF, Koester J, Fruh B, Shephard DT, Barone P, Houben JJ. The effect of pramipexole on mood and motivational symptoms in Parkinson's disease: a meta-analysis of placebo-controlled studies. *Clin Ther*. 2009;31(1):89–98.

Lemke MR, Brecht HM, Koester J, Kraus PH, Reichmann H. Anhedonia, depression, and motor functioning in Parkinson's disease during treatment with pramipexole. *J Neuropsychiatry Clin Neurosci*. 2005;17(2):214–220.

Mah L, Zarate Ca Jr, Nugent AC, Singh JB, Manji HK, Drevets WC. Neural mechanisms of antidepressant efficacy of the dopamine receptor agonist pramipexole in treatment of bipolar depression. *Int J Neuropsychopharmacol*. 2011;14(4):545–551.

National Center for Biotechnology Information. PubChem Compound Database. http://pubchem.ncbi.nlm.nih.gov/

Nelson JC, Mankoski R, Baker RA, et al. Effects of aripiprazole adjunctive to standard antidepressant treatment on the core symptoms of depression: a post-hoc, pooled analysis of two large, placebo-controlled studies. *J Affect Disord*. 2010;120(1–3):133–140.

Rakofsky JJ, Holtzheimer PE, Nemeroff CB. Emerging targets for antidepressant therapies. *Curr Opin Chem Biol*. 2009;13(3):291–302.

Rektorova I, Rektor I, Bares M, et al. Pramipexole and pergolide on the treatment of depression in Parkinson's disease: a national multicentre prospective randomized study. *Eur J Neurol*. 2003;10(4):399–406.

Selle V, Schalkwijk S, Vazquez GH, Baldessarini RJ. Treatments for acute bipolar depression: meta-analyses of placebo-controlled monotherapy trials of anticonvulsants, lithium and antipsychotics. *Pharmacopsychiatry*. 2014;47(2):43–52.

Sevransky JE, Bienvenu JO, Neufeld KJ, Needham DM. Treatment of four psychiatric emergencies in the intensive care unit. *Crit Care Med*. 2012;40(9):2662–2670.

Spielmans GI, Berman MI, Linardatos E, Rosenlicht NZ, Perry A, Tsai AC. Adjunctive atypical

antipsychotic treatment for major depressive disorder: a meta-analysis of depression, quality of life and safety outcomes. *PLoS Med*. 2013;10(3):e1001403.

*Strawn JR, Keck PE Jr, Caroff SN. Neuroleptic malignant syndrome. *Am J Psychiatry*. 2007;164(6):870–876.

Tang H, Hunter T, Hu Y, Zhai SD, Sheng X, Hart RJ. Cabergoline for preventing ovarian hyperstimulation syndrome. *Cochrane Database Syst Rev*. 2012;(2):CD008605.

U.S. National Library of Medicine. *Dailymed*. http://dailymed.nlm.nih.gov/dailymed/index.cfm. Accessed May 28, 2015.

Veselinovic T, Paulzen M, Gründer G. Cariprazine, a new orally active dopamine D2/3 receptor partial agonist for the treatment of schizophrenia, bipolar mania and depression. *Expert Rev Neurother*. 2013;13(11):1141–1159.

Vieta E, Valenti M. Pharmacological management of bipolar depression: acute treatment, maintenance and prophylaxis. *CNS Drugs*. 2013;27(7):515–529.

Worthington JJ 3rd, Simon NM, Korbly NB, Perlis RH, Pollack MH; Anxiety Disorders Research Program. Ropinirole for antidepressant-induced sexual dysfunction. *Int Clin Psychopharmacol*. 2002;17(6):307–310.

Zarate CA Jr, Payne JL, Singh J, et al. Pramipexole for bipolar II depression: a placebo-controlled proof of concept study. *Biol Psychiatry*. 2004;56(1):54–60.

Zarate CA Jr, Quiroz JA, Singh JB, et al. An open-label trial of the glutamate-modulating agent riluzole in combination with lithium for the treatment of bipolar depression. *Biol Psychiatry*. 2005;57(4):430–432.

▲ 34.31 Thyroid Hormones

RUSSELL T. JOFFE, M.D.

INTRODUCTION

Various hormonal systems have been implicated in the biology of the stress response and in the pathophysiology of mood disorders. The adrenal axis has been the focus of most attention, but the thyroid axis has also been established as an important endocrine system.

Thyroid hormones have long been of interest to those who study and treat mood disorders, both unipolar major depression and bipolar disorder. There have been two major areas of research study. The first is whether thyroid hormones are implicated in the etiology of mood disorders and whether there is a specific abnormality of the thyroid axis in primary mood disorders. The second, following on the first, is whether thyroid hormones may be useful in the treatment of unipolar or bipolar mood disorders.

For more than a century, depressive symptoms have been known to be one of the most prominent features of clinical thyroid disease, particularly hypothyroidism. Beginning in the late part of the 19th century, depression was seen as a prominent clinical component of clinical hypothyroidism,

although other psychiatric symptoms were often present. Clinical hyperthyroidism has been observed to have a much more heterogeneous presentation with a wider variety of psychiatric symptoms reported. Earlier studies and clinical observations also noted that when the particular thyroid deficiency or excess was corrected, depressive or other psychiatric symptoms frequently, but not invariably, resolved. As a result of the strong inverse relationship between thyroid function and depression observed in endocrine patients, it was hoped that thyroid hormones—particularly thyroid hormone deficiency—might be a biological factor in the etiology of unipolar and bipolar mood disorders.

Although thyroid hormone deficiency has been excluded as a major contributor to the biological basis of primary mood disorders, particularly depression, its potential role in the cascade of biological events that leads to depressive and bipolar disorders remains to be clarified. Thyroid hormones have extensive effects on various cellular mechanisms and may still be implicated in the biology of these disorders.

The early link established between abnormal mood symptomatology and thyroid deficiency in endocrine patients prompted the investigation of the use of thyroid hormones to treat depression. Earlier studies produced ambiguous findings, which may have been contributed to by methodological limitations in the studies done, especially lack of controlled designs and heterogeneous groups of patients. In the 1960s, levothyroxine (Synthroid), synthetic forms of thyroxine (T_4) and liothyronine (Cytomel), the synthetic levorotatory isomer of triiodothyronine (T_3), were developed and replaced desiccated thyroid as a primary replacement therapy for thyroid disease. This allowed a more refined approach to examining the efficacy of thyroid hormones in the treatment of mood disorders, and the therapeutic use of thyroid hormones in the treatment of depression became the subject of study. Although thyroid hormones have been used to treat a range of psychiatric illnesses, systematic research data have been collected primarily on unipolar mood disorders and to a lesser extent, on bipolar mood disorders.

CHEMISTRY

The source of thyroid hormones in the body is the thyroid gland. Several thyroid hormones are produced by the thyroid gland. The two main circulating hormones are thyroxine (T_4) and triiodothyronine (T_3). T_4 is derived mainly from the process of iodination of tyrosine in the thyroid gland, whereas the main source of T_3, approximately 80 percent, is by monodeiodination of T_4 in extrathyroidal tissue by deiodinase enzymes. Only a limited amount of T_3 is produced in the thyroid gland. Other thyroid hormones produced include reverse T_3, 3-diiodothyronines, 3-monoiodothyronines, and the acetate metabolites of T_4 and T_3—tetrac and

triac, respectively. These other thyroid hormone products are produced in very small amounts and have unknown physiological significance.

The secretion of thyroid hormones by the thyroid gland is regulated by the hypothalamic–pituitary–thyroid axis. This involves the classic negative feedback regulatory mechanism, which applies to most of the major endocrine systems.

The cells of the hypothalamus secrete TRH, which regulates the secretion of thyrotropin (thyroid-stimulating hormone [TSH]) from the cells of the anterior pituitary via the portal venous blood system, which connects the hypothalamus and the pituitary gland. The TSH released from the pituitary then acts on the thyroid gland to stimulate the production of thyroid hormones, including T_4 and T_3. A negative feedback system operates at both the levels of the pituitary and the hypothalamus to modulate circulating levels of T_4 and T_3 and thereby regulates the activity of the thyroid axis.

The physiological action of T_4 is largely dependent on its extrathyroidal conversion to T_3, which then acts at the T_3 receptor in various tissues to initiate the physiological actions of thyroid hormones. For the most part, the conversion of T_4 to T_3 occurs in the circulation so that the thyroid hormone available to peripheral tissues is largely T_3. In the brain to a greater extent and pituitary to a lesser extent, the main source of thyroid hormone for action at the receptor is T_4, which gets converted locally within the neuron or cell to T_3 before exerting its hormonal action.

PHARMACOLOGICAL ACTION

Pharmacokinetics

Levothyroxine and liothyronine are both relatively small molecules that are easily absorbed across the gastrointestinal tract. Both hormones are strongly protein bound and are widely distributed in tissues. There is an active transport mechanism for the uptake of thyroid hormones across the blood–brain barrier. There is also a transport mechanism at the choroid plexus to ensure the wide distribution of thyroid hormones within the CNS. Several deiodinase enzymes, particularly the type III deiodinase enzyme, convert T_3 to smaller molecules including reverse T_3, which are then excreted in both urine and feces. Levothyroxine has a long half-life of approximately 7 days, whereas liothyronine has a half-life of approximately 8 to 12 hours.

Pharmacodynamics

Thyroid hormones have widespread metabolic effects in numerous organs and tissues. In the mature adult, they play key roles in both protein synthesis and calorigenesis and they also have an important role in the

development of the maturing infant.

Until the mid-1970s, it was unclear whether thyroid hormones had a substantial role in mature brain function. Studies since then have indicated that thyroid hormones play an active role in the mature brain and that they are very highly regulated to ensure that brain thyroid hormone levels remain within very narrow limits.

CLINICAL HORMONE STUDIES

The vast majority of patients with primary major depression are euthyroid. A large number of studies evaluating baseline thyroid function tests have produced inconsistent findings. The most consistent findings to emerge are higher measures of T_4 or free T_4 in the depressed state as compared with recovery or to healthy controls with much more inconsistent findings for measures of circulating T_3 or free T_3. Up to one-third of depressed patients may have a blunted TSH response to TRH which often normalizes with treatment. However, due to low sensitivity and moderate specificity for major depression the TRH test is not diagnostic for depression. Although a specific abnormality of thyroid function within the normal range has not been identified, there is some limited data on baseline thyroid function and the relationship to antidepressant response. In some studies, higher mean T_4 levels and in other studies higher mean TSH levels have been associated with antidepressant response. A reduced TSH response to TRH has also been associated with treatment response. There are preliminary data that polymorphisms of the deiodinase enzymes, which are responsible for the extrathyroidal conversion of T_4 to T_3 (type I and type II deiodinase) and for the inactivation of T_3 by conversion to reverse T_3 (type III deiodinase), may be associated with antidepressant response in euthyroid-depressed patients.

Subclinical hypothyroidism (SCH), defined as elevated TSH levels with normal circulating T_4 and T_3 levels, occurs in a minority of patients, at most 10 to 15 percent, with primary mood disorders. SCH is more common with advanced age and in women. SCH is associated with a lower response rate to antidepressants and there is recent evidence that even the presence of thyroid autoimmunity as measured by positive anti-thyroperoxidase antibodies with normal TSH may be associated with lower rates of antidepressant response. There are no data to guide clinical decision making about whether thyroid replacement would enhance antidepressant response so the decision on thyroid replacement should be consistent with current guidelines for the treatment of SCH.

All peptides and hormones of the hypothalamic–pituitary–thyroid axis have been studied for the treatment of mood disorders. These include TRH and TSH, as well as the peripheral hormones T_4 and T_3. The best evidence for therapeutic efficacy is the use of liothyronine in unipolar major

depression, although there are less data on the role of levothyroxine in the treatment of depression. Preliminary data suggest that levothyroxine may have a role as a mood stabilizer in bipolar disorder.

In major depression, thyroid hormones have been used in several ways. First, in very early studies, they have been used as monotherapy for treatment of depression with uncertain benefit. Second, they have been used at the outset of an antidepressant trial, primarily with tricyclic antidepressants, to accelerate the response to antidepressant treatment. Third, they have been added to antidepressants in patients who have failed to have a satisfactory response to an antidepressant trial—that is, for augmentation therapy. Fourth, they have been used at the initiation of an antidepressant trial to determine whether they would accelerate and enhance antidepressant response in depressed subjects requiring antidepressant therapy. Finally, they have been used as mood stabilizers in bipolar disorders, especially rapid-cycling forms.

Thyrotropin-Releasing Hormone

TRH is a peptide released from the hypothalamus that regulates thyroid hormone secretion by the thyroid gland through its effects on pituitary TSH release. In addition, TRH is a peptide that circulates in the brain and has behavioral effects. These include reversal of drug-induced sedation or anesthesia and stimulation of locomotor activity in the rat. In addition, TRH also has effects on cardiorespiratory, gastrointestinal, and other neurological functions. Because of the widespread effects of TRH on basic physiological functions in the CNS and elsewhere, as well as its effects on the thyroid axis, it has been tested as an antidepressant. Most studies have involved monotherapy, but there have also been studies of the use of TRH together with electroconvulsive therapy (ECT). TRH has been administered to patients by both intravenous and oral routes for depression. TRH has been administered intravenously either as a single dose or as several doses over a period of 3 to 4 days, and transient antidepressant effects have been demonstrated. These studies are summarized in Table 34.31–1. However, at least half of these studies have reported no effect or very minimal therapeutic response to administration of either intravenous or oral TRH administered as monotherapy with duration of treatment ranging from a single dose up to 30 days. Although a positive effect on depressed mood cannot be definitively excluded, it is very difficult to determine whether TRH has a significant and specific therapeutic role in the treatment of depression. Moreover, the transient effects noted may have little to do with the thyroid axis and may be a nonspecific activating effect of the peptide.

Table 34.31–1.
Antidepressant Response to Thyrotropin-Releasing Hormone

Number of Studies	Total N	Route of Administration	Duration	Response
3	36	Intravenous	Single dose	1 of 3 responded to TRH > placebo
6	86	Intravenous	3–4 days	1 of 6 showed transient response
3	52	Oral	7–30 days	1 of 3 responded to TRH = amitriptyline

The effect of TRH on mood and cognitive side effects of electroconvulsive therapy (ECT) have also been studied. In a small, randomized, double-blind, placebo-controlled crossover study in which 500 μg of TRH was administered to eight depressed patients who also received ECT for their depression, TRH administered intravenously before the ECT yielded greater arousal and improved cognitive functioning when compared with placebo. TRH did not have any substantial effect on any seizure variables. This is a small study that requires replication before one can assess whether there is any clinical applicability of these findings.

Thyrotropin

TSH has also received limited attention as an antidepressant. The rationale for its use is based on its stimulation of the thyroid gland and, thus, would achieve an effect similar to the administration of peripheral thyroid hormones for the treatment of depression. One study to date has examined the antidepressant effect of TSH by administering 10 IU of TSH intravenously to 20 depressed women 1 day before beginning an antidepressant trial with the tricyclic imipramine (Tofranil). When compared to a placebo control group, the TSH-treated patients had a rapid antidepressant response. These data, while of interest, require replication with larger samples, current methodology for clinical trials, and the use of SSRIs. Such study designs would allow a more appropriate evaluation of the efficacy of TSH in accelerating the response to antidepressants.

Liothyronine

Liothyronine is the most broadly studied thyroid hormone for the treatment of depression. In early studies, it was used as monotherapy for the treatment of depressed patients. The data from these studies were largely inconclusive, as they involved small patient samples, inadequate clinical trial designs by current standards, and the use of heterogeneous patient groups who, by today's diagnostic criteria, would not necessarily be diagnosed with major depression. There have been no well-designed studies of liothyronine monotherapy to date, and, therefore, its use as a single treatment for depression has not gained any clinical use.

Liothyronine has been used in earlier studies to accelerate the response to tricyclic antidepressants. Liothyronine was administered at the outset of

a tricyclic antidepressant trial and a shorter lag in onset of therapeutic effect as compared with placebo was noted in some but not all studies. This acceleration effect was observed particularly in women as compared with men. All of these studies had major methodological flaws, including small numbers of patients, poorly defined patient groups, and relatively brief duration of treatment, as well as the fact that these earlier studies involved tricyclic antidepressants, as SSRIs were not available, usually at suboptimal doses by current standards. Nonetheless, a recent meta-analysis of these earlier studies showed a significant acceleration of antidepressant response during the initial 4 weeks of tricyclic treatment, for liothyronine over placebo.

The third approach to the use of liothyronine in depressive disorders has been to augment response to antidepressants in those who failed to respond to a tricyclic antidepressant trial. These studies are reviewed in Table 34.31–2. Open and controlled studies show that approximately 50 to 60 percent of patients who do not respond to an antidepressant trial with tricyclics will respond within 2 to 3 weeks after the addition of liothyronine. Liothyronine augmentation has also been found to be comparable to lithium augmentation of both tricyclics and SSRIs. In view of the limitations of the studies of liothyronine augmentation to date, a meta-analysis of these studies concluded that liothyronine may increase response rates and decrease severity of depression scores in patients who are refractory to tricyclic antidepressant treatment. Patients with liothyronine augmentation were approximately twice as likely to respond as were controls.

In the past few years there have been emerging data on the use of liothyronine to augment SSRIs in antidepressant nonresponders. The findings with the SSRIs are consistent with those for the tricyclics and are included in Table 34.31–3. Both open and controlled studies are generally positive and indicate that liothyronine may be an effective augmentation agent for SSRI nonresponders. In recent data from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial, liothyronine augmentation had comparable remission rates to other augmentation options and a more favorable adverse event dropout rate.

There are also preliminary data that if liothyronine is administered at the outset of a course of ECT, a reduced number of treatments are required to achieve therapeutic response, and there may be a greater sparing of cognitive difficulties.

In recent studies, there has been another use of liothyronine, that is, it has been studied at the outset of an SSRI trial to determine whether it would reduce delay in therapeutic onset and increase rates of treatment response in patients not necessarily refractory to antidepressant treatment. The studies to date have generally shown no benefit of liothyronine addition to the SSRI.

There is preliminary evidence suggesting that liothyronine may be effective in treatment-resistant patients. Larger studies are required to address several issues. First, further studies of liothyronine augmentation with SSRIs with large-scale, well-controlled designs are needed. Second, issues such as dosing and duration of treatment need to be addressed to provide solid data to guide the use of liothyronine in a broad-based safe clinical approach to treatment-resistant depression.

Table 34.31–2.
Liothyronine Augmentation of Antidepressants

Study	N	T_3 Dose (lg/day)	Design	Result
Earle, 1970	25	25	Open	14 of 25 responded
Ogura et al., 1974	44	20–30	Open	29 of 44 responded
Banki, 1975	52	20–40	Open	39 of 52 responded
Banki, 1977	33	20	Partially controlled	23 of 33 responded
Tsutsui et al., 1979	11	5–25	Open	10 of 11 responded
Goodwin et al., 1982	12	25–50	Double-blind	8 of 12 responded
Schwarcz et al., 1984	8	25–50	Open	4 of 8 responded
Gitlin et al., 1987	16	25	Double-blind, crossover	T_3 = placebo
Thase et al., 1989	20	25	Open	5 of 20 responded
Joffe and Singer, 1990	38	37.5	Double-blind	9 of 17 responded
Joffe et al., 1993	51	37.5	Double-blind, placebo-controlled	10 of 17 responded to T_3, T_3 > placebo, T_3 = (lithium?)
Birkenhager et al., 1997	14	37.5	Open	No response—all inpatients
Agid and Lerer, 2003	74	25–50	Open, added to SSRI	44 of 74 responded at 4 weeks
Abraham et al., 2006	12	50	Open, added to SSRI	Three remissions and five responses
Joffe et al., 2006	33	37.5	Double-blind, T_3 vs. lithium vs. placebo, added to SSRI	No difference between treatment groups
Nierenberg et al., 2006 STAR*D Trial	142	50	Open, vs. lithium, added to SSRI	Remission 25%-T_3 and 16%-lithium. More side effects with lithium
Posternak et al., 2007	50	50	Double-blind, placebo-controlled added to SSRI	Response 57%-T_3 vs. 33%-placebo

SSRI, selective serotonin reuptake inhibitor; STAR*D, sequenced treatment alternatives to relieve

depression; T_3, triiodothyronine.

Levothyroxine

This thyroid hormone has been used in two ways in the treatment of mood disorders. First, it has been used as an augmentation agent for the treatment of tricyclic-refractory depression and, second, as a mood stabilizer for rapid-cycling bipolar disorder.

Levothyroxine has received less attention as an augmentation agent, as most of the initial studies were carried out using liothyronine. However, there is a database on the use of levothyroxine in refractory depression. These studies are summarized in Table 34.31–3. These studies generally involve open designs using various diagnoses, including both unipolar and bipolar subjects in one study and chronic depression or dysthymic patients in another study. All involve small sample designs. In each of these studies, levothyroxine showed augmentation effect in antidepressant nonresponders.

Levothyroxine has also been used in high doses as a mood stabilizer. There are now five case studies or case series, four involving rapid-cycling bipolar patients and one involving nonrapid-cycling bipolar patients. The total patient sample in the five studies consists of 30 patients. These studies are reviewed in Table 34.31–4. Doses of levothyroxine ranged from 150 to 500 μg/day. Of the 30 patients, 27 were reported to respond to levothyroxine or augmentation of mood stabilizers. Clearly, these data are of considerable interest, especially because rapid-cycling bipolar patients represent a substantial therapeutic challenge. A large-scale systematic control trial would clarify the efficacy of levothyroxine in bipolar disorder and its role relative to other mood-stabilizing interventions.

Liothyronine has been preferably used in psychiatric disorders whereas thyroxine is the standard thyroid hormone used for replacement for hypothyroidism. Liothyronine has a lower therapeutic index and a much shorter half-life, making thyroid axis suppression—a goal of replacement therapy—unreliable. In the last 20 years, there has been greater appreciation of substantial dissatisfaction on a minority of patients who receive thyroxine replacement. Many of them complain of symptoms in the psychiatric domain including fatigue, lethargy, depression, and anxiety. Several meta-analyses have failed to demonstrate a significant advantage for combined thyroxine and liothyronine over thyroxine replacement, which remains the standard treatment. However, there is now increasing recognition that a selected group of patients may benefit from the combination and these are likely identified by various polymorphisms of the deiodinase enzymes which would shift the ratio of circulating T_4 and T_3.

Table 34.31–3.
Antidepressant Augmentation with Levothyroxine

Study	N	Patients	T_4 Dose (lg/day)	Design	Response
Joffe and Singer, 1990	38	Unipolar major depression	150	Double-blind vs. T_3	4 of 21 responded to T_4, $T_4 < T_3$
Bauer et al., 1998	17	Unipolar (or?) bipolar depression	300–500	Open	10 of 17 showed full or partial response
Spoov and Lahdelma, 1998	22	Major depression	200	Crossover vs. lithium	T_4 > lithium
Rudas et al., 1999	9	Chronic depression and dysthymia	150–300	Open	6 of 9 full or partial response
Lojko and Rybakowski, 2007	17	Unipolar and bipolar depression	100	Open, on SSRI	11 remission and 16 response

T_3, triiodothyronine; T_4, thyroxine; SSRI, selective serotonin reuptake inhibitor.

Table 34.31–4.
Levothyroxine as Adjunctive Mood Stabilizer in Bipolar Disorder

Study	N	T_4 Dose (lg/day)	Type	Design	Response
Stancer and Persad, 1982	10	240–500	Rapid cycling	Open	8 of 10
Leibow, 1983	1	400	Rapid cycling	Open	1 of 1
Bauer and Whybrow, 1990	11	150–400	Rapid cycling	Open	10 of 11
Horowitz and Leibowitz, 1993	2	200–350	Rapid cycling	Open	2 of 2
Baumgartner et al., 1998	6	250–500	Nonrapid cycling	Open	6 of 6

T_4, thyroxine.

Therapeutic Indications

None of the peptides or hormones of the hypothalamic–pituitary–thyroid axis have received formal approval by regulatory authorities in North America or worldwide for use in the treatment of depression. Preliminary evidence suggests that these hormones may have clinical use in the treatment of different types of mood disorders. Liothyronine may be useful as an augmentation agent in refractory depression and may also have some use in accelerating antidepressant response. The use of levothyroxine in refractory depression requires further study, and levothyroxine may be an adjunct to mood-stabilizer treatment, particularly in rapid-cycling patients. For thyroid hormones to have broad clinical application, empirical data are required to clarify their efficacy and to determine their role relative to other options for these various types of mood disorders. Until such data are obtained, liothyronine remains one of many possibilities to consider for use in patients who do not respond to their first or subsequent antidepressant

trial. Moreover, levothyroxine may be considered as an option for the refractory bipolar patient, especially when conventional mood stabilizers have proven to be ineffective or only partially effective.

Precautions and Adverse Reactions

Liothyronine administered in the usual doses for augmentation treatment should not lead to symptoms of thyrotoxicosis. A dosage of 25 μg/day is below the replacement dosage for this hormone; thus, symptoms of thyroid toxicity are extremely rare. A dose of 50 μg/day is closer to the daily T_3 production range; thus, symptoms of thyroid excess could occur. It should be noted that there are no well-demonstrated dose–response studies for liothyronine; therefore, doses in excess of 25 μg should be used with caution and are not generally recommended.

Liothyronine is generally well tolerated when administered to depressed patients. A transient headache is the most commonly reported adverse effect, and it usually subsides after 1 or 2 days, either spontaneously or with the administration of mild analgesics. Allergic reactions are extremely uncommon. Liothyronine should be used with great caution in patients, particularly the elderly and those with various medical disorders, including endocrine and cardiovascular disease, and should be avoided in pregnancy.

The same precautions as noted for liothyronine apply to levothyroxine, especially since this compound is commonly used in supraphysiological doses. A recent study has demonstrated that depressed patients are able to tolerate supraphysiological doses of levothyroxine much better than are healthy controls. Long-term thyroid hormone treatment may be associated with bone demineralization, resulting in osteoporosis in women. Although this association is still controversial, one should approach the long-term use of thyroid hormones with caution.

Dosage and Administration

Liothyronine is generally administered as a single dose in the morning to avoid insomnia. There is a good rationale for divided administration of liothyronine because of its short half-life, but a single daily dose is commonly used to improve compliance. Levothyroxine has been administered in dosages that produce a circulating levothyroxine concentration, which exceeds or is at the upper range of normal for patients with bipolar mood disorder. These dosage recommendations apply to healthy adults and should be modified for older patients or those with significant medical illnesses. There are no data to support the efficacy of thyroid hormone treatment of depression in children.

References

*Altshuler LL, Bauer M, Frye MA, et al. Does thyroid supplementation accelerate tricyclic

antidepressant response? A review and meta-analysis of the literature. *Am J Psychiatry*. 2001;158:1617–1622.

Appelhof BC, Brouwer JP, van Dyck R, et al. Triiodothyronine addition to paroxetine in the treatment of major depressive disorder. *J Clin Endocrinol Metab*. 2004;89:6271–6276.

*Aronson R, Offman HJ, Joffe RT, Naylor CD. Triiodothyronine augmentation in the treatment of refractory depression: a meta-analysis. *Arch Gen Psychiatry*. 1996;53:842–848.

Banki CM. Cerebrospinal fluid amine metabolites after combined amitriptyline-triiodothyronine treatment of depressed women. *Eur J Clin Pharmacol*. 1977;11:311–315.

Banki CM. Triiodothyronine in the treatment of depression. *Orv Hetil*. 1975;116:2543–2546.

Bauer M, Baur H, Berghöfer A, et al. Effects of supraphysiological thyroxine administration in healthy controls and patients with depressive disorders. *J Affect Disord*. 2002;68:285–294.

Bauer MS, Hellweg R, Gräf KJ, Baumgartner A. Treatment of refractory depression with high-dose thyroxine. *Neuropsychopharmacology*. 1998;18:444–455.

Bauer MS, Whybrow PC. Rapid cycling bipolar affective disorder. II: Treatment of refractory rapid cycling with high-dose levothyroxine: a preliminary study. *Arch Gen Psychiatry*. 1990;47:435–440.

Baumgartner A. Thyroxine and the treatment of affective disorders: an overview of the results of basic and clinical research. *Int J Neuropsychopharmacol*. 2000;3:149–165.

Baumgartner A, Bauer M, Hellweg R. Treatment of intractable non-rapid cycling bipolar affective disorder with high-dose thyroxine: an open clinical trial. *Neuropsychopharmacology*. 1994;10:183–189.

*Cooper-Kazaz R, Apter JT, Cohen R, et al. Combined treatment with sertraline and liothyronine in major depression: a randomized, double-blind, placebo-controlled trial. *Arch Gen Psychiatry*. 2007;64:679–688.

Cooper-Kazaz R, van der Deure WM, Medici M, et al. Preliminary evidence that a functional polymorphism in type I deiodinase is associated with enhanced potentiation of the antidepressant effect of sertraline by triiodothyronine. *J Affect Disord*. 2009;116:113–116.

Coppen A, Whybrow PC, Noguera R, Maggs R, Prange AJ Jr. The comparative antidepressant value of L-tryptophan and imipramine with and without attempted potentiation by liothyronine. *Arch Gen Psychiatry*. 1972;26:234–241.

Earle BV. Thyroid hormone and tricyclic antidepressants in resistant depressions. *Am J Psychiatry*. 1970;126:1667–1669.

Eller T, Metsküla K, Talja I, Maron E, Uibo R, Vasar V. Thyroid autoimmunity and treatment response to escitalopram in major depression. *Nord J Psychiatry*. 2010;64:253–257.

Flach FF, Celian CI, Rawson RW. Treatment of psychiatric disorders with triiodothyronine. *Am J Psychiatry*. 1958;114:841–842.

Garlow SJ, Dunlop BW, Ninan PT, Nemeroff CB. The combination of triiodothyronine (T3) and sertraline is not superior to sertraline monotherapy in the treatment of major depressive disorder. *J Psychiatr Res*. 2012;46:1406–1413.

Gitlin MJ, Weiner H, Fairbanks L, Hershman JM, Friedfeld N. Failure of T3 to potentiate tricyclic antidepressant response. *J Affect Disord*. 1987;13:267–272.

Goodwin FK, Prange AJ Jr, Post RM, Muscettola G, Lipton MA. Potentiation of antidepressant effects by L-triiodothyronine in tricyclic nonresponders. *Am J Psychiatry*. 1982;139:34–38.

Joffe RT, Singer W. A comparison of triiodothyronine and thyroxine in the potentiation of tricyclic antidepressants. *Psychiatry Res*. 1990;32:241–251.

*Joffe RT, Singer W, Levitt AJ, MacDonald C. A placebo-controlled comparison of lithium and triiodothyronine augmentation of tricyclic antidepressants in unipolar refractory depression. *Arch Gen Psychiatry*. 1993;50:387–393.

Joffe RT, Sokolov ST, Levitt AJ. Lithium and triiodothyronine augmentation of antidepressants. *Can J Psychiatry*. 2006;51:791–793.

Łojko D, Rybakowski JK. L-thyroxine augmentation of serotonergic antidepressants in female patients with refractory depression. *J Affect Disord*. 2007;103(1–3):253–256.

Nierenberg AA, Fava M, Trivedi MH, et al. A comparison of lithium and T(3) augmentation following two failed medication treatments for depression: a STAR*D report. *Am J Psychiatry*. 2006;163:1519–1530.

Ogura C, Okuma T, Uchida Y, Imai S, Yogi H. Combined thyroid (triiodothyronine)-tricyclic antidepressant treatment in depressive states. *Folia Psychiatr Neurol Jpn*. 1974;28:179–186.

Posternak M, Novak S, Stern R, et al. A pilot effectiveness study: placebo-controlled trial of adjunctive L-triiodothyronine (T3) used to accelerate and potentiate the antidepressant response. *Int J Neuropsychopharmacol*. 2008;11(1):15–25.

*Prange AJ Jr, Wilson IC, Rabon AM, Lipton MA. Enhancement of imipramine antidepressant activity by thyroid hormone. *Am J Psychiatry*. 1969;126:457–469.

Rudas AS, Schmitz M, Pichler P, Baumgartner A. Treatment of refractory chronic depression and dysthymia with high-dose thyroxine. *Biol Psychiatry*. 1999;45:229–233.

Schwarcz G, Halaris A, Baxter L, Escobar J, Thompson M, Young M. Normal thyroid function in desipramine nonresponders converted to responders by the addition of L-triiodothyronine. *Am J Psychiatry*. 1984;141:1614–1616.

Thase ME, Kupfer DJ, Jarrett DB. Treatment of imipramine-resistant recurrent depression: I. An open clinical trial of adjunctive L-triiodothyronine. *J Clin Psychiatry*. 1989;50:385–388.

Tsutsui S, Yamazaki Y, Namba T, Tsushima N. Combined therapy of T3, and antidepressants in depression. *J Int Med Res*. 1979;7:138–146.

van de Ven AC, Muntjewerff JW, Netea-Maier RT, et al. Association between thyroid function, thyroid autoimmunity, and state and trait factors of depression. *Acta Psychiatr Scand*. 2012;126:377–384.

Wheatley D. Potentiation of amitriptyline by thyroid hormone. *Arch Gen Psychiatry*. 1972;26:229–233.

Wiersinga WM. Paradigm shifts in thyroid hormone replacement therapies for hypothyroidism. *Nat Rev Endocrinol*. 2014;10:164–174.

Wiersinga WM, Duntas L, Fadeyev V, Nygaard B, Vanderpump MP. 2012 ETA Guidelines: The use of L-T4 + L-T3 in the treatment of hypothyroidism. *Eur Thyroid J*. 2012;1:55–71.

Wilson IC, Prange AJ Jr, McClane TK, Rabon AM, Lipton MA. Thyroid-hormone enhancement of imipramine in nonretarded depressions. *N Engl J Med*. 1970;282:1063–1067.

▲ 34.32 Trazodone

HISTORY

Developed in Italy in 1966, trazodone (Desyrel) is considered the first of the SGAs, possessing a novel triazolopyridine chemical structure that distinguished it from the TCAs that preceded it. This structural uniqueness was associated with pharmacological properties that also differentiated it from the TCAs in its performance on animal antidepressant screens, proposed mechanism of action, and adverse effect profile. Trazodone was first marketed in Europe in the 1970s and received FDA approval in 1981 for the treatment of major depression in the United States. Its improved safety and tolerability compared to TCAs was most likely responsible for trazodone's rapid rise in clinical use after its introduction. However, with the increased popularity of SSRI antidepressants, trazodone has fallen out of favor as a first-line treatment for depression in the United States. A once-a-day formulation of trazodone (Oleptro—trazodone/Contramid) was introduced by Labopharm Inc. in 2010 and approved by the FDA for the treatment of major depressive disorder.

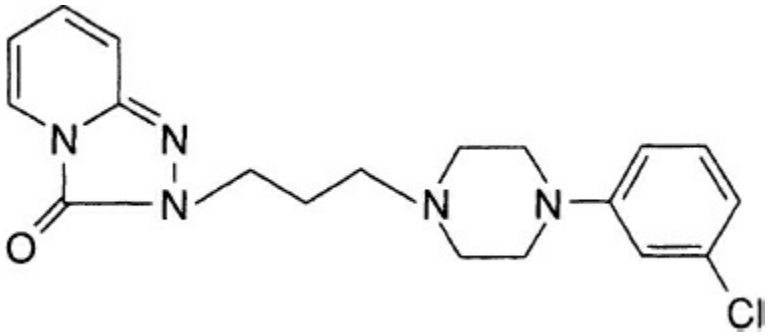

FIGURE 34.32–1. Molecular structure of trazodone (Desyrel).

CHEMISTRY

Trazodone is a triazolopyridine derivative, designated as 2-[3-[4-(3-chlorophenyl)-1-piperazinyl] propyl]-1,2,4-triazolo[4,3-a]pyridin-3(2H)-one hydrochloride. Its empirical formula is $C_{19}H_{22}ClN_5O.HCl$, and its molecular structure is depicted in Figure 34.32–1. Trazodone's molecular weight is 408.3. It is soluble in water, possessing a pKa of 6.7.

PHARMACOLOGY

Pharmacokinetics

Trazodone is absorbed efficiently by the human digestive tract after oral administration, reaching peak plasma concentration in about 1 hour. When consumed with food, the total amount of drug absorption is increased, maximum plasma concentration is decreased, and time to peak

concentration is doubled. Since some adverse effects such as dizziness are related to peak plasma levels, this strategy has been shown to increase tolerability. Trazodone exhibits a linear pharmacokinetic profile in the dosage range of 50 to 150 mg twice daily. When administered twice daily, initial steady-state concentrations are achieved in approximately 36 hours, with restabilization at new levels in 12 hours after dosage adjustment. There may be up to a fourfold difference in steady-state plasma levels between individuals maintained on identical dosages. Steady-state AUC of trazodone IR 100 mg three times daily and Oleptro 300 mg daily were similar, although steady-state C_{max} and C_{min} after 1 week were not found to be equivalent.

Trazodone is approximately 90 percent bound to plasma proteins and distributes without selective localization in any tissue. Trazodone and its major metabolite, meta-chlorophenylpiperazine (mCPP), both cross the blood–brain barrier. Trazodone is also excreted into breast milk, with one study reporting that the concentration is about 14 percent of the maternal plasma concentration.

Trazodone is primarily metabolized via hydroxylation and oxidation by the hepatic cytochrome P450 (CYP) system. Several studies implicate the CYP 3A4 microsomal enzyme in the production of its major, active metabolite, mCPP, while reports on the role of CYP 2D6 are conflicting. The elimination of trazodone is biphasic, with half-lives of 3 to 6 hours and 5 to 9 hours, respectively. It is eliminated primarily via the kidneys as metabolites, with only about 1 percent excreted as the parent compound. The half-life of mCPP is about 14 hours. The half-life of Oleptro is reported at 10 hours.

Pharmacodynamics

Trazodone's main pharmacological effects are believed to be mediated through its actions at postsynaptic serotonin (5-hydroxytryptamine [5-HT]) 5-HT_2 receptors. Although it exhibits weak in vitro effects at the presynaptic serotonin transporter, which is responsible for the main therapeutic effects of the SSRIs, in vivo studies suggest that this is of negligible significance in humans. Trazodone is a potent 5-HT_{2A} receptor antagonist, with a 100-fold selectivity for the 5-HT_2 relative to the 5-HT_1 binding site. Chronic administration causes a downregulation of 5-HT_2 as well as β-adrenergic receptors, similar to the actions of some other antidepressant compounds.

Trazodone's major metabolite, mCPP, is also active at serotonin receptors. It is an agonist at 5-HT_{1B} and 5-HT_{1C}, and possibly, 5-HT_{1A} and 5-HT_{1D} receptors. Research suggests that the parent compound, and not its metabolite, is responsible for the antidepressant effects. In fact, mCPP has been shown to be behaviorally activating and anxiogenic in experimental

animals and human subjects.

Trazodone is an antagonist at peripheral α-adrenergic receptors, which is most likely responsible for its adverse effects of orthostatic hypotension and dizziness, dry mouth, and priapism. It has no significant activity at noradrenergic or dopaminergic receptors. Antagonism at acetylcholine and histamine type-1 (H_1) receptors is negligible. One in vitro study suggests that inhibition of T-type calcium channels may contribute to trazodone's psychiatric and neurological effects.

There is no established consensus regarding the relationship between steady-state blood levels of trazodone and therapeutic response. One study in elderly subjects reported that a threshold concentration of 650 ng/mL predicts clinical efficacy, but studies in other populations produced conflicting results. An open-label, 3-week treatment study in a Japanese population of 26 adult patients with major depression reported a linear relationship between trazodone (but not mCPP) steady-state plasma concentration and both the final Montgomery Asberg Depression Rating Scale (MADRS) score and the percentage improvement at study completion. Also, the proportion of responders was significantly higher in the patients whose plasma concentration exceeded 714 ng/mL, consistent with the above study. No relationship between adverse effects and either trazodone or mCPP levels was seen. This finding suggests that, although not common in clinical practice, measuring trazodone levels to ensure that they exceed a threshold of 650 to 700 ng/mL may help improve outcome in patients with inadequate response to the usual dosage regimen.

DESIGN AND INTERPRETATION OF CLINICAL DRUG STUDIES

Major Depressive Disorder

The majority of clinical research on trazodone has focused on its use in the treatment of major depressive disorder. After a series of open trials suggesting the efficacy of trazodone in depressive syndromes, double-blind comparison trials of trazodone were carried out first in Europe and then in the United States. Some, but not all, of the studies included a placebo comparison arm, as some investigators considered it unethical to withhold active medication from seriously depressed patients. Comparator drugs included amitriptyline (Elavil), imipramine (Tofranil), doxepin (Sinequan), mianserin (Tolvon), dothiepin (Prothiaden), and diazepam (Valium). The studies were conducted in inpatient, outpatient, and general practice settings. Most, but not all, of these studies found trazodone to possess efficacy that is comparable to the TCAs. Several reviews and meta-analyses that attempted to overcome the limitations of the individual studies concluded that trazodone was indeed as efficacious as TCAs for the acute treatment of major depression. One such meta-analysis calculated the pooled odds ratio for response to trazodone compared to imipramine as

0.93 (95 percent confidence interval: 0.64 to 1.36), supporting the hypothesis of near-equivalent efficacy of the two medications. In addition, a 12-month, double-blind maintenance study found trazodone to be superior to imipramine on end-point efficacy measures.

Later studies have compared trazodone to other newer antidepressants. Two double-blind, randomized, 6-week comparisons between trazodone and fluoxetine (Prozac) in 43 and 126 patients, respectively, reported equivalent response rates for the two medications, although trazodone demonstrated better HAM-D sleep subscale scores in the first study. One double-blind, placebo-controlled multicenter trial compared trazodone (50 to 300 mg) and venlafaxine (Effexor) (25 to 150 mg) in 225 patients in a 6-week acute treatment phase followed by a 1-year continuation phase for responders. Subjects in the venlafaxine group were more likely to respond sufficiently to enter the continuation phase and more likely to remain in the continuation phase. Trazodone proved more effective on measures of sleep and venlafaxine proved more effective on measures of cognition and psychomotor slowing. Another 6-week, parallel, double-blind, randomized trial compared trazodone 50 to 300 mg to bupropion (Wellbutrin) 225 to 450 mg in 126 moderately to severely depressed outpatients, finding no significant differences in efficacy measures between the two treatments at end point but greater improvement with trazodone at 7 days based on sleep measures. However, a 6-week, double-blind comparison of trazodone (150 to 450 mg per day) and mirtazapine (Remeron) (24 to 72 mg per day) in 200 inpatients with a *Diagnostic and Statistical Manual of Mental Disorders*, third edition (DSM-III) diagnosis of moderate to severe major depressive episode found mirtazapine to be superior to trazodone on most efficacy measures. Since the introduction of controlled-release trazodone in 2000, two small, 6-week, double-blind studies reported equal efficacy in response and remission rates between this and both paroxetine (Paxil) and sertraline (Zoloft). A more recent once-a-day formulation of trazodone (Oleptro) was studied in a randomized double-blind, placebo-controlled multicenter trial in 2009 indicating statistically significant improvement in HAM-D 17 score versus placebo. Additionally, the antidepressant efficacy was shown to be independent of insomnia presence at baseline and independent of the overall reduction of insomnia. A 2007-published meta-analysis detected no significant differences between trazodone and its sister 5-HT_2 receptor antagonist, nefazodone (Serzone), compared with SSRIs for the treatment of major depression.

Trazodone has been systematically studied in elderly depressed patients, as its low potential for anticholinergic and cardiovascular side effects suggested greater tolerability compared with tricyclics in this population. Trials comparing trazodone with various tricyclic and heterocyclic antidepressants found comparable efficacy and a generally lower side-effect burden for trazodone. The most frequently reported side

effect in the elderly was dizziness, which may have been the result of orthostatic hypotension, expected from trazodone's pharmacological profile. This symptom can be significant in the elderly for its potential to produce falls and, therefore, may limit the use of trazodone in those most sensitive to it, particularly if exacerbated by physical illness or concomitant medications such as antihypertensives. In addition, studies have compared antidepressants for their effects on cognitive performance in the elderly. One review of such studies found evidence of impaired ability to concentrate or sustain attention for sedating antidepressants including trazodone, but little or no effects on memory with trazodone compared with tricyclics. In comparison, that review reported no cognitive adverse effects for MAOIs or SSRIs, suggesting that trazodone may not be the best choice for geriatric depression since the introduction of the SSRIs. Regarding the treatment of symptoms of depression in children, it should also be noted that a 2003 survey of 1,273 child and adolescent psychiatrists indicated that trazodone was the most frequently prescribed medication for insomnia for children with sleep disturbance associated with mood disorders despite very limited evidence for the efficacy and safety of trazodone for use in children.

During clinical studies of trazodone's antidepressant properties, several double-blind trials compared the efficacy of single bedtime dosing and divided daily dosing. One smaller study compared single bedtime dosing with thrice-daily dosing in a 4-week trial of 56 patients with major depression. That study reported equivalent therapeutic efficacy and tolerability for the two dosing regimens; a trend for greater improvement in onset and duration of sleep was noted for the bedtime dose group, as well as more side-effect complaints in the thrice-daily dosing group. In a larger trial of 146 depressed patients, twice-daily dosing was the comparison with single bedtime dosing over a period of 6 weeks. That study similarly reported equivalent antidepressant efficacy for both regimens as well as improvements in sleep for both overall. During the early weeks of the study, the single bedtime dosing group reported significantly superior sleep, while the twice-daily dosing group reported significantly greater daytime fatigue.

Overall, trazodone is an effective treatment for major depression. There is no clear indication that its effectiveness differs from that of the TCAs or SSRIs. Some studies suggest that its tolerability, and perhaps its overall effectiveness, is maximized by limiting doses to the more moderate ranges (300 mg per day or less). Adverse effects of sedation can be minimized via bedtime dosing, which also utilizes its beneficial effects for early relief of depression-related insomnia.

Anxiety Disorders

Although most clinical trials with trazodone have been undertaken for

major depression, many of those studies gave early indication that trazodone possesses antianxiety as well as antidepressant effects as evidenced by significant reductions in anxiety-related symptom scales. In addition, small, uncontrolled studies in patients with prominent anxiety symptoms or anxiety neurosis suggested its potential as an anxiolytic. However, few controlled trials of trazodone in samples meeting diagnostic criteria for anxiety disorders exist.

Generalized Anxiety Disorder. Two hundred and thirty patients with DSM-III GAD, in whom major depression and panic disorder were excluded, were treated with trazodone, imipramine, or diazepam in a randomized, double-blind, placebo-controlled 8-week trial. Although diazepam produced the greatest improvement in anxiety ratings during the first 2 weeks, both trazodone and imipramine attained comparable efficacy to diazepam in weeks 3 through 8, but only imipramine was significantly better than placebo at end point. Among completers, 73, 69, and 66 percent of imipramine-, trazodone-, and diazepam-treated patients reported moderate to marked improvement, respectively, compared to only 47 percent of patients treated with placebo. The antidepressants tended to provide greater improvement in the psychic symptoms of worry, tension, and apprehension compared to diazepam.

Panic Disorder. Data regarding the efficacy of trazodone for panic disorder is limited and conflicting. A small, uncontrolled, single-blind 8-week trial of trazodone in 11 patients with DSM-III diagnoses of panic disorder or agoraphobia with panic attacks reported significant improvements in ratings of panic and generalized anxiety, phobias, avoidance, and depression. However, in a double-blind study of 74 panic disorder patients, comparing trazodone to imipramine and alprazolam (Xanax), only the latter two effectively reduced panic and anxiety symptoms, with a high dropout rate for trazodone-treated patients primarily due to nonresponse.

Obsessive-Compulsive Disorder. Similarly, the data for trazodone's antiobsessional efficacy are mixed. Several case reports and small case series reported trazodone to be helpful for patients with obsessive-compulsive neurosis or disorder (OCD), including one report of nine patients who had failed to respond to clomipramine (Anafranil). However, a double-blind, placebo-controlled trial of trazodone in 21 patients with OCD reported no significant difference in OCD symptoms between those treated with trazodone (mean daily dose 235 mg) and those receiving placebo.

Posttraumatic Stress Disorder. Several small trials of trazodone use in

patients with PTSD have been published. One study retrospectively surveyed 74 chronic PTSD patients who had ever been prescribed trazodone for insomnia or nightmares. Of the 60 patients who had tolerated trazodone, 72 percent reported it to be helpful for nightmares and 92 percent reported improvement in sleep. A correlation of 0.57 between these two measures suggests that trazodone's beneficial effect on nightmares in PTSD may be partially independent of its sleep benefits. An open-label trial of trazodone in six patients with combat-related PTSD rated four of the six as much improved after 4 months of treatment. A 2011 literature review informing an algorithm for the treatment of PTSD proposed the use of low-dose trazodone following a trial of prazosin (Minipress) for nightmares and nighttime hyperarousal or for (persisting) initial insomnia (without nightmares) or after resolution of nightmares.

Adjustment Disorders

Two nearly identical studies of trazodone versus clorazepate (Tranxene) for the treatment of adjustment disorders in medically ill patients have been published. Both were randomized, double-blind, 28-day trials in patients meeting DSM-III-R criteria for adjustment disorder with anxious or depressed mood or mixed disturbance of emotion and conduct but without significant history of more serious psychiatric disorders. One population consisted of 18 patients receiving treatment for breast cancer, and the other 21 HIV-positive subjects. The outcomes were similar, each reporting significant symptom benefit from both medications, with nonsignificant trends suggesting greater efficacy from trazodone.

Bulimia Nervosa

Like other serotonergic antidepressants, trazodone has been tested for the treatment of bulimia nervosa, with some success. Several case studies have reported benefit on a range of illness measures. In an open-label, flexible-dose study of 13 subjects diagnosed with DSM-III bulimia, trazodone produced a significant decrease in the number of bingeing and purging (vomiting) episodes, carbohydrate cravings, and urges to binge eat. Bingeing and purging episodes were reduced to zero in four of ten subjects available for analysis and decreased by more than 50 percent in two others after about 8 weeks on the medication. In a 6-week, placebo-controlled study of 42 women meeting DSM-III-R criteria for bulimia nervosa, trazodone was superior to placebo in decreasing the frequency of bingeing and purging. In both of these trials, adverse effects were relatively mild and consistent with other studies of trazodone (sedation and dizziness) and in no cases resulted in withdrawal from the study, despite some isolated case reports of delirium in bulimic patients treated with trazodone. No trials exist to date that compare trazodone to other agents, such as the SSRIs.

Schizophrenia

Although it has no significant effects on the dopamine system, trazodone's actions at 5-HT_2 receptors are similar to those of atypical antipsychotics, suggesting a potential for efficacy in psychosis. However, this has not been borne out in clinical studies. One research group tested the efficacy of trazodone in the management of chronic schizophrenia in two small double-blind, controlled trials of 20 patients each. The first reported that trazodone produced significantly less deterioration than placebo in the Excitement and Hallucinatory Behavior item scores of the Brief Psychiatric Rating Scale (BPRS) over a 12-week period. The second compared trazodone to chlorpromazine (Thorazine), finding that chlorpromazine was significantly superior to trazodone in total BPRS score and several individual factor scores, although trazodone did improve depression scores on the BPRS. Subsequently, another group examined the potential for adjunctive trazodone use to reduce negative symptoms in chronic, residual schizophrenia using a double-blind, placebo-controlled design. Those investigators found a modest (10 to 15 percent) but beneficial effect for trazodone on two of three negative symptom scales, with only moderate clinical improvement seen in 3 of 26 trazodone-treated patients. Trazodone did not appear to worsen positive symptoms of schizophrenia. Of additional note, trazodone 100 mg has demonstrated some efficacy in reducing acute neuroleptic-induced akathisia in a small, placebo-controlled, double-blind crossover study in 2010.

Insomnia

From 1999 through 2010, trazodone was found to hold a stable (off-label) second position to the increased use of nonbenzodiazepine GABA-A omega-1 agonists as most commonly prescribed medication for insomnia (from the National Health and Nutrition Examination Survey of 32,328 adults) despite limited availability of double-blind, placebo-controlled trials demonstrating clear efficacy and safety of trazodone for insomnia treatment. This reflects physician prescriptions of trazodone for the treatment of primary insomnia, depression/anxiety/withdrawal-related insomnia, and antidepressant-induced insomnia, especially for patients at higher risk of sedative/hypnotic/anxiolytic use disorders or patients with comorbid depression. While trazodone is not associated with physiological dependence or withdrawal, a 2008 randomized, placebo-controlled trial of trazodone at 50 to 150 mg over the course of 12 weeks in patients with alcohol use disorder advised caution. The study found that trazodone was helpful for the treatment of insomnia during early recovery with rates of relapse similar to placebo. However, at 6 months postdetoxification the prior trazodone-treatment group demonstrated increased number of drinks per drinking day compared to placebo.

Trazodone is a 5-HT_{2A} antagonist that increases slow wave sleep, shortens sleep latency, improves sleep continuity and does not suppress REM sleep. In a 3-week, within-subjects randomized, double-blind trial, 16 patients with primary insomnia were given trazodone 50 mg or placebo 30 minutes before bedtime for 7 days and administered polysomnograms and MSLT on days 1 and 7. Trazodone produced small but significant impairments of short-term memory, verbal learning, equilibrium, and arm muscle endurance across time points. Compared to placebo, trazodone was associated with less middle insomnia, decreased stage 1 sleep, increased slow-wave sleep (by day 7), less daytime drowsiness (by day 7), and improved subjective experience of sleep by self-report. In another study of 14 patients with primary insomnia and Beck Depression Inventory (BDI) score >10 and 15 patients with primary insomnia and BDI score <10, trazodone CR 25 to 150 mg daily for 3 months improved sleep; however, for patients with elevated BDI scores, discontinuation of trazodone resulted in worsening sleep, while improvement in sleep quality was sustained for patients with lower BDI scores with trazodone. Compared to other off-label 5-HT 1/2A antagonists such as the atypical antipsychotic quetiapine, trazodone may have favorable efficacy. In a 2013 study (although nonrandomized, unblinded, and not placebo-controlled), of 30 inpatients receiving trazodone (mean dose 102 mg) and 31 inpatients receiving quetiapine (mean dose 182 mg), mean total sleep time hours (by subjective report and nursing sleep logs) were longer and nighttime awakenings fewer in the trazodone group compared to the quetiapine group. However, the trazodone group reported more frequent GI side effects. Trazodone has also been applied with some efficacy to treatment of insomnia in the postpartum period. In 54 postpartum patients with insomnia during the third trimester of pregnancy but without depression, EPDS scores were lower in patients treated with trazodone 50 mg or diphenhydramine 25 mg compared to placebo with no significant difference between the two active treatment groups.

Pain

Trazodone is among the various antidepressants that have been examined for their analgesic potential in a wide variety of chronic pain syndromes. Two case series of open-label trazodone trials for painful diabetic neuropathy, with sample sizes of 6 and 31 patients, respectively, reported fairly dramatic symptom relief in the majority of patients. Trazodone has also been reported to be beneficial in a double-blind, comparative study with amitriptyline for deafferentation pain and in a double-blind, placebo-controlled study of pediatric migraine prophylaxis, although two systematic literature reviews question these latter findings. Two uncontrolled open-label studies in 2010 and 2011 suggest that trazodone may help with

aspects of fibromyalgia. In one study, pain interference with daily activities, sleep, and depression were improved without direct mitigation of pain intensity. In the other study, pregabalin (Lyrica) added to trazodone reduced fibromyalgia overall severity, depression, pain interference with daily activities, and pain intensity. Trazodone did not prove to be effective in double-blind, placebo-controlled trials for fibromyalgia, chronic low-back pain, dysesthetic pain following traumatic myelopathy, or rheumatoid arthritis. Thus, little evidence exists supporting the use of trazodone in most chronic pain syndromes.

Behavioral Disturbances

Trazodone and other serotonergic medications have been used in place of sedative hypnotics and antipsychotics to manage aggression and other behavioral disturbances in certain patient populations, such as children with disruptive behavioral disorders and geriatric patients with cognitive disorders. It is postulated that trazodone targets disturbed serotonergic brain pathways that are believed to underlie such behaviors.

Childhood Disruptive Behavioral Disorders. One uncontrolled open trial studied the effectiveness of trazodone in the management of impulsive and aggressive behaviors in an inpatient group of 22 children diagnosed with DSM-III-R disruptive behavioral disorders (ADHD, conduct disorder, and oppositional-defiant disorder) comorbid with other psychiatric symptomatology. These children, aged 5 to 12, had been unresponsive to other inpatient and outpatient therapeutic interventions, including an average of 4.7 documented medication trials. Thrice-daily dosing of trazodone at an average dose of about 5 mg/kg per day for a duration of 3 to 4 weeks resulted in clinical improvement in 65 percent of the children, with most maintaining this benefit (per parental report) at 3 to 14 months' follow-up. The medication was well tolerated by the majority of the children, with transient orthostatic hypotension being the most frequently noted adverse effect (50 percent). A larger (also uncontrolled) study examining the connection between sleep and behavioral disturbances in 51 young children with opsoclonus-myoclonus syndrome reported that trazodone improved sleep in 95 percent and reduced rage attacks in 33 percent of those children treated.

Neurocognitive Disorders. Several open-label case series and double-blind, controlled trials have been published examining the effect of trazodone on agitation in elderly patients with neurocognitive disorders. These case series, typically small and with a mixed population of elderly patients with various neurocognitive disorders, overall reported beneficial effects of trazodone. A Cochrane Database review of the existing controlled studies found similarity in efficacy and tolerability of trazodone to placebo

and haloperidol as a treatment of behavioral disturbances of neurocognitive disorders, and subsequently, the review concluded that trazodone might improve some measures of agitation and psychosis for some patients with neurocognitive disorders.

Sexual Dysfunction

Trazodone has been associated with an increased risk of priapism and also has been the subject of several case reports of enhanced sexual desire and functioning in both sexes. These findings have led researchers to evaluate its potential benefit for treating disorders of sexual function, particularly male erectile dysfunction. In general, studies of treatments for sexual disorders in women have been disproportionately sparse; however, trazodone is being studied in combination with bupropion for the treatment of female hypoactive sexual desire disorder (see DSM-IV-TR definition, results pending as of November 2014). In 2009, one small open-label preliminary study of SSRI-induced sexual dysfunction in men and women with depressive/anxiety/somatic-symptom disorders found significant improvements in both men and women for all domains of sexual function and satisfaction measured. In this study, the improvement was additionally noted to be independent of improvement on measures of depressive/anxiety symptoms. In vitro, trazodone has been shown to have a high affinity for human penile α_1 receptors and moderate affinity for α_2-adrenergic receptors, lending experimental support to this endeavor.

Trazodone reportedly improved objectively assessed erectile measures in several small, uncontrolled case series of patients with impotence. However, double-blind, placebo-controlled trials of trazodone use for erectile dysfunction present different results. Three separate double-blind, placebo-controlled trials, the first in 69 patients taking 150 mg daily for 4 weeks, the second in 51 patients with 50 mg daily dosing for 3 months, and the third in 34 patients taking 200 mg daily for 4 weeks, all failed to find significant differences between active medication and placebo. The results of a review and meta-analysis of the literature on trazodone treatment for erectile dysfunction reported a nonsignificant trend of improvement with trazodone, particularly at higher doses, with a subgroup analysis favoring benefit if the dysfunction is psychogenic in origin.

Other Disorders

Trazodone has been studied as a potential treatment for a wide range of other psychiatric and medical syndromes, mostly via case reports or open-label trials. These include dysthymia, depression following stroke, tic disorders, chest pain due to esophageal dysmotility, sleep apnea, menopausal symptoms, and benzodiazepine, alcohol, and opiate use disorders or withdrawal. Although varying degrees of efficacy have been

reported, the overall lack of data, particularly from trials with more stringent study methodology, prevents one from drawing any firm conclusions regarding trazodone's potential use in these conditions.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

In both premarketing trials and postmarketing surveillance, the adverse effects seen with trazodone therapy have generally been benign, mostly limited to dry mouth, headache, nausea, drowsiness, fatigue, and dizziness, the latter in part secondary to orthostatic hypotension. Unlike TCAs, trazodone is virtually devoid of anticholinergic side effects, including major cardiac conduction abnormalities, thus explaining its benign profile following overdose. The adverse reaction for which trazodone is most widely recognized, although infrequent, is priapism.

Cardiovascular System

Trazodone appears to have a relatively benign cardiovascular profile. The most troublesome cardiovascular adverse effect is postural hypotension, which can be severe enough to produce syncope in susceptible patients, particularly the elderly or those with concurrent medical illness or taking other medications with cardiovascular effects. Adjustments in the dosage or dosing intervals of trazodone or concurrent medications, together with precautions regarding postural changes, can usually mitigate this problem if trazodone use is otherwise warranted.

No consistent effects of trazodone on cardiac contractility or the interventricular conduction system in humans or animals have been noted, except for possible nonspecific lengthening of the PR interval. However, several case reports of cardiac abnormalities associated with trazodone use exist in the literature, including ventricular arrhythmias, QT prolongation, and complete heart block. Most, but not all, of these effects were seen either associated with trazodone overdose or in individuals with pre-existing heart disease, suggesting caution in this latter category of patients.

Genitourinary System

The most concerning side effect attributed to trazodone is priapism, a painful unremitting erection that requires emergent medical attention. It can occur as a complication of various medical illnesses such as sickle cell disease, but a majority of cases have implicated medication use. The α-adrenergic receptor antagonists, such as trazodone, chlorpromazine, and prazosin, are hypothesized to induce priapism by inhibiting sympathetically controlled detumescence. The incidence of priapism during trazodone therapy is estimated at 1 in 1,000 to 1 in 10,000 cases. Although it is most likely to occur within the first 28 days of initiating treatment with trazodone, it has occurred as late as 18 months into treatment and at a wide

range of therapeutic dosages. A past history of prolonged erections appears to be a general risk factor for priapism. At least two cases of clitoral priapism associated with trazodone use have also been reported in the literature.

All patients on trazodone therapy need to be warned of the risk and medical significance of priapism, as patients may be reticent to acknowledge it and seek treatment. Once priapism develops, it is considered a urological emergency requiring immediate medical intervention. Patients should be advised to discontinue trazodone and seek emergency care in the event of an erection lasting more than 1 hour. The likelihood of morbid complications of impotence, cavernosal fibrosis, and gangrene is directly related to a delay in initiating treatment. Various medical interventions (ice packs, oral adrenergic medications) have produced inconsistent results, but are usually attempted before more definitive surgical procedures are instituted. One of two surgical interventions (intracavernosal irrigation with an α-adrenergic agonist or shunting between the corpus cavernosa and spongiosum) will usually produce detumescence, but impotence may still be the outcome in as many as half of these cases.

Central Nervous System

Trazodone frequently produces CNS side effects of drowsiness (20 to 40 percent) and dizziness or light-headedness (20 to 30 percent), particularly early in treatment. Drowsiness has been the side effect most frequently leading to discontinuation, although it can be a welcome benefit for those patients with severe insomnia. Headache, fatigue, nervousness, and confusion have lower incidence. Isolated cases of delirium induced by trazodone have been reported in patients with pre-existing cerebral or metabolic abnormalities.

Trazodone shows little, if any, risk for lowering the seizure threshold in patients. There have been isolated case reports of increased seizure activity associated with trazodone use, although these are far less prevalent than with TCAs. In patients with a known history of seizure disorder, it is prudent to initiate therapy with trazodone at lower doses and titrate the dose more gradually.

As is the case with other antidepressants, trazodone has been reported to precipitate manic episodes in patients with diagnoses of both major depressive disorder and bipolar disorder. Trazodone's potential for increasing the risk for hypomania or mania appears to be lower than that of TCAs. Finally, trazodone also carries the black box warning (applied to all antidepressant medications) for increased risk of suicidal ideation, especially in children, adolescents, and young adults.

THERAPEUTIC INDICATIONS

Trazodone has an FDA indication for major depressive disorder, and research supports its general efficacy and tolerability for this indication. Currently, the most common off-label use of trazodone is for the treatment of insomnia, particularly insomnia associated with the use of other antidepressants such as the SSRIs, for which it remains a reasonable option for some patients. For other conditions, as are described above, the findings are tentative and would require replication in larger samples before treatment recommendations can be made.

PRECAUTIONS AND ADVERSE REACTIONS

The most common adverse effect reported by patients using trazodone is drowsiness. Patients should be warned that this side effect may be exacerbated by alcohol or other concomitant medications, thus impairing their ability to safely engage in potentially hazardous tasks, such as driving. Orthostatic hypotension resulting in syncope can occur, particularly in the elderly or medically ill, and may cause significant injury through falls. Adjusting the dosage of trazodone or concurrent medications and ingesting trazodone with meals may diminish this effect, but continued caution should be exercised when prescribing trazodone to elderly patients. There are no reported withdrawal or discontinuation syndromes associated with trazodone. A 2014 review found some cases of dose-independent trazodone-induced liver injury (hepatocellular by immuno-allergic mediation), with one resulting in liver transplant. A 2011 meta-analysis indicated that trazodone was correlated with higher all-cause mortality compared to SSRIs and TCAs in treatment of elder depressed patients.

Priapism, although infrequent, is the side effect that is most concerning due to its potential for producing serious, irreversible injury. All patients should be strongly cautioned to discontinue trazodone and seek immediate medical attention in the event of prolonged or painful erections.

There is no evidence that trazodone possesses mutagenic or carcinogenic potential based on studies in rats receiving doses up to 300 mg/kg for 18 months. Trazodone is designated pregnancy category C, as it has produced fetal anomalies in animal studies at 15 to 50 times the maximum human dose, but it has not been adequately evaluated in human pregnancy. One multicenter study compared the pregnancy outcomes of 147 women who had exposure to either trazodone (n = 58) or the related phenylpiperazine antidepressant, nefazodone (n = 89), with the outcomes from women exposed to other antidepressant (n = 147) or nonteratogenic medications (n = 147). They found no statistically significant differences among the three groups for any of the end points of interest, including miscarriage rate, rate of major malformations, gestational age, and birth weight. As with all medications that have an unclear potential for fetal effects, use in pregnancy should only be considered if the potential benefit outweighs the risk. Trazodone is excreted into human breast milk, with one

study in six lactating women reporting a milk/plasma ratio of about 15 percent, indicating that exposure to nursing infants would be low.

Trazodone has a very high therapeutic index, with an LD_{50} of around 500 mg/kg in experimental animals. Data from hundreds of documented cases of overdose support its safety in all but a small fraction of cases, with the majority of those involving other concurrent medications. The main symptoms that occur when trazodone is taken in overdose are excessive drowsiness and vomiting, although isolated cases of seizures, hyponatremia, priapism, respiratory arrest, QT prolongation, cardiac arrest, and coma have been reported. One postmortem toxicological review of trazodone-related deaths concluded that trazodone played a primary role in the deaths by massive overdose of only three individuals who attained blood concentrations above 9 mg/L.

DRUG INTERACTIONS

Significant changes in prothrombin time (PT) and international normalized ratio (INR) in patients receiving warfarin (Coumadin) therapy and increased levels of digoxin (Lanoxin) and phenytoin (Dilantin) have been reported in isolated cases of patients receiving trazodone concurrently, most likely due to interactions via trazodone's high binding to plasma proteins. An in vitro study reported that the metabolism of trazodone to its principal metabolite (mCPP) via CYP3A-mediated clearance was significantly inhibited by ketoconazole (Nizoral), ritonavir (Norvir), and indinavir (Crixivan), suggesting caution in prescribing trazodone to individuals treated with CYP 3A4 inhibitors, including HIV-infected individuals who are commonly prescribed these medications. One patient taking fluoxetine and trazodone developed sedation and an unsteady gait, with the finding of a 30 percent increase in trazodone levels. However, this was an isolated report, as most patients tolerate this combination without difficulty.

As discussed earlier, trazodone may potentiate the sedative or hypotensive effects of other pharmacological agents, requiring caution when used in such combinations. In combination with MAOIs to treat insomnia, trazodone appears to be safe. No cases of hypertensive crisis have been reported, and one reported case of nocturnal myoclonus resolved after discontinuation of trazodone. There have been several isolated reports of serotonin syndrome associated with trazodone use. Most of these occurred when trazodone was prescribed in combination with a variety of serotonergic psychotropic medications, including buspirone (BuSpar), fluoxetine, sertraline, paroxetine, venlafaxine, nefazodone, and amitriptyline with lithium (Eskalith). One case occurred when trazodone alone was prescribed to a 73-year-old patient who was later suspected to suffer from Lewy body dementia.

LABORATORY INTERFERENCE

Trazodone generally exhibits no interference with urinalysis, spectrometry, or other laboratory tests. There is one isolated report of trazodone-related decreases in hematocrit, hemoglobin, red blood cell count, serum cholesterol, calcium, and albumin during a 6-week comparison trial with fluoxetine, with the development of clinically insignificant pseudoanemia in 36 percent of the trazodone-treated and 20 percent of the fluoxetine-treated groups, respectively. Trazodone has also been reported to result in false-positive urine drug screen for amphetamine.

DOSAGE AND ADMINISTRATION

The manufacturer's package insert recommends initiating trazodone at 150 mg per day in divided doses (after meals), increasing by 50 mg per day in intervals of 3 to 4 days up to a maximum dose of 400 mg per day in outpatients and 600 mg per day in inpatients. Elder patients should use a starting dose of 100 mg daily. Once-a-day trazodone (Oleptro) can be started at 150 mg and increased as needed by 75 mg every third day to a maximum dose of 375 mg daily for the treatment of major depressive disorder. Oleptro is recommended to be taken on an empty stomach at bedtime, and it is noted the tablets may be bisected if needed without adversely affecting the extended-release properties of the tablet. There are not yet approved dosage recommendations for children or adolescents.

Trials of trazodone for the treatment of major depression suggest that the majority of patients begin to respond within the first 2 weeks of therapy, although this finding was in the context of objective, controlled studies, many with higher initial dosages. However, patients with significant insomnia or agitation may show an early response to trazodone's sedating effects. For the treatment of primary or medication-related insomnia, doses in the 50- to 100-mg range are usually sufficient.

Trazodone is available in scored 50- and 100-mg tablets and in 150- and 300-mg tablets that can be divided into halves or thirds for convenient dosing adjustment and administration. Oleptro is available in 150- and 300-mg tablets.

REFERENCES

Ansari A. The efficacy of newer antidepressants in the treatment of chronic pain: A review of current literature. *Harv Rev Psychiatry*. 2000;7(5):257–277.

Bajor LA, Ticlea AN, Osser DN. The psychopharmacology algorithm project at the Harvard South Shore Program: An update on posttraumatic stress disorder. *Harv Rev Psychiatry*. 2011;19(5):240–258.

Belkin ZR, Krapf JM, Goldstein AT. Drugs in early clinical development for the treatment of female sexual dysfunction. *Expert Opin Investig Drugs*. 2014:1–9.

Bertisch SM, Herzig SJ, Winkelman JW, Buettner C. National use of prescription medications for

insomnia: NHANES 1999–2010. *Sleep*. 2014;37(2):343–349.

*Bossini L, Casolaro I, Koukouna D, Cecchini F, Fagiolini A. Off-label uses of trazodone: a review. *Expert Opin Pharmacother*. 2012;13(12):1707–1717.

Brahm NC, Yeager LL, Fox MD, Farmer KC, Palmer TA. Commonly prescribed medications and potential false-positive urine drug screens. *Am J Health Syst Pharm*. 2010;67(16):1344–1350.

Bryant SG, Hokanson JA, Brown CS. A drug utilization review of prescribing patterns for trazodone versus amitriptyline. *J Clin Psychiatry*. 1990;51 Suppl:27–29.

Coupland C, Dhiman P, Morriss R, Arthur A, Barton G, Hippisley-Cox J. Antidepressant use and risk of adverse outcomes in older people: population based cohort study. *BMJ*. 2011;343:d4551.

De Wit S, Cremers L, Hirsch D, Zulian C, Clumeck N, Kormoss N. Efficacy and safety of trazodone versus clorazepate in the treatment of HIV-positive subjects with adjustment disorders: a pilot study. *J Int Med Res*. 1999;27(5):223–232.

Doroudgar S, Chou TS, Yu J, Trinh K, Pal J, Perry PJ. Evaluation of trazodone and quetiapine for insomnia: An observational study in psychiatric inpatients. *Prim Care Companion CNS Disord*. 2013;15(6):PCC.13m01558.

*Fagiolini A, Comandini A, Dell'Osso MC, Kasper S. Rediscovering trazodone for the treatment of major depressive disorder. *CNS Drugs*. 2012;26:1033–1049.

Herr J, Caspi A. Oleptro™ (trazodone hydrochloride) extended-release tablets. *P T*. 2011;36(2):2–18.

Khazaiea H, Ghadamia MR, Knight DC, Emamiana F, Tahmasian M. Insomnia treatment in the third trimester of pregnancy reduces postpartum depression symptoms: A randomized clinical trial. *Psychiatry Res*. 2013;210(3):901–905.

Martinez MA, Ballesteros S, Sanchez de la Torre C, Almarza E. Investigation of a fatality due to trazodone poisoning: Case report and literature review. *J Anal Toxicol*. 2005;29(4):262–268.

*Mittur A. Trazodone: properties and utility in multiple disorders. *Expert Rev Clin Pharmacol*. 2011;4(2):181–196.

Owens JA, Rosen CL, Mindell JA, Kirchner HL. Use of pharmacotherapy for insomnia in child psychiatry practice: A national survey. *Sleep Med*. 2010;11(7):692–700.

Patten SB. The comparative efficacy of trazodone and imipramine in the treatment of depression. *CMAJ*. 1992;146(7):1177–1182.

Pranzatelli MR, Tate ED, Dukart WS, Flint MJ, Hoffman MT, Oksa AE. Sleep disturbance and rage attacks in opsoclonus-myoclonus syndrome: response to trazodone. *J Pediatr*. 2005;147(3):372–378.

Razavi D, Kormoss N, Collard A, Farvacques C, Delvaux N. Comparative study of the efficacy and safety of trazodone versus clorazepate in the treatment of adjustment disorders in cancer patients: a pilot study. *J Int Med Res*. 1999;27(6):264–272.

Roth AJ, McCall WV, Liguori A. Cognitive, psychomotor, and polysomnographic effects of trazodone in primary insomniacs. *J Sleep Res*. 2011;20(4):552–558.

Rotzinger S, Fang J, Baker GB Trazodone is metabolized to m-chlorophenylpiperazine by CYP3A4 from human sources. *Drug Metab Dispos*. 1998;26(6):572–575.

Saletu-Zyhlarz GM, Anderer P, Arnold O, Saletu B Confirmation of the neurophysiologically predicted therapeutic effects of trazodone on its target symptoms depression, anxiety and

insomnia by postmarketing clinical studies with a controlled-release formulation in depressed outpatients. *Neuropsychobiology*. 2003;48(4):194–208.

*Schatzberg AF. Trazodone A 5-year review of antidepressant efficacy. *Psychopathology*. 1987;20 Suppl 1:48–56.

Seitz DP, Adunuri N, Gill SS, Gruneir A, Herrmann N, Rochon P. Antidepressants for agitation and psychosis in dementia. *Cochrane Database Syst Rev*. 2011;(2):CD008191.

Thompson JW Jr, Ware MR, Blashfield RK: Psychotropic medication and priapism: A comprehensive review. *J Clin Psychiatry*. 1990;51(10):430–433.

Victor S, Ryan SW: Drugs for preventing migraine headaches in children. *Cochrane Database Syst Rev*. 2003;(4):CD002761.

Voican CS, Corruble E, Naveau S, Perlemuter G. Antidepressant-induced liver injury: a review for clinicians. *Am J Psychiatry*. 2014;171(4):404–415.

Wichniak A, Wierzbicka A, Jernajczyk W. Sleep and antidepressant treatment. *Curr Pharm Des*. 2012;18:5802–5817.

*Workman EA, Short DD. Atypical antidepressants versus imipramine in the treatment of major depression: A meta-analysis. *J Clin Psychiatry*. 1993;54(1):5–12.

Zubieta JK, Alessi NE. Acute and chronic administration of trazodone in the treatment of disruptive behavior disorders in children. *J Clin Psychopharmacol*. 1992;12(5):346–351.

▲ 34.33 Tricyclics and Tetracyclics

J. CRAIG NELSON, M.D.

HISTORY

In 1957, Roland Kuhn, a Swiss psychiatrist, observed that imipramine (Tofranil) had antidepressant effects in a small sample of 40 patients. This serendipitous observation, like those before for lithium (Eskalith) and chlorpromazine (Thorazine), had far reaching import. It demonstrated that medications could be used to treat depression. It led to the development of a new class of antidepressant compounds. This clinical observation, coupled with Julius Axelrod's report of imipramine's role in blocking reuptake of norepinephrine, laid the foundation for the norepinephrine hypothesis of depression.

The study of the treatment of depression with imipramine and related compounds would drive the development of new methodology. For example, objective measures were required to assess change during treatment. Max Hamilton developed the Hamilton Depression Rating Scale to monitor change during administration of these new compounds. Other scales followed. Better methods of defining depression were needed, and the Washington University group proposed criteria for primary affective disorder, criteria that became the basis for the category major depression in the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III). Observation of substantial placebo effects in depression during clinical trials led to the appreciation of the need for placebo controls.

The field also learned the basics of pharmacological management of depression with imipramine and other tricyclics. The necessity for administering an adequate dose for an adequate duration was described. The need to continue treatment to prevent relapse and recurrence was identified. The observation of the substantial variability of the tricyclic concentrations in plasma stimulated a whole area of study of the pharmacology of these compounds. This research introduced psychiatry to the concept of *genetic polymorphisms* and the possible consequences of drug interactions.

After the introduction of imipramine, several other antidepressant compounds were developed that shared a basic tricyclic structure and relatively similar secondary or side effects. Later, other heterocyclic compounds were also marketed that were somewhat similar in structure and that had relatively comparable secondary properties.

CHEMISTRY

Tricyclic and tetracyclic compounds are categorized primarily on the basis of their chemical structure (Fig. 34.33–1). The *tricyclics* have a three-ring structure, hence, their name. The tertiary amine tricyclics, such as amitriptyline (Elavil) and imipramine, have two methyl groups at the end of the side chain. These compounds can be demethylated to secondary amines, such as desipramine (Norpramin, Pertofrane) and nortriptyline (Aventyl, Pamelor). The tetracyclic compounds, maprotiline (Ludiomil) and amoxapine (Asendin), have a four-ring central structure. Five tertiary amines have been marketed in the United States—amitriptyline, clomipramine (Anafranil), doxepin (Sinequan), imipramine, and trimipramine (Surmontil). The three secondary amine compounds are desipramine, nortriptyline, and protriptyline (Vivactil).

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

Absorption. Absorption of the tricyclic and tetracyclic drugs occurs in the small intestine and is reasonably complete and rapid. Peak levels after ingestion occur within 2 to 8 hours. Exceptions are protriptyline, which is absorbed more slowly, with peak levels obtained between 6 and 12 hours, and maprotiline, which has peak levels occurring in 8 hours or more. Because the antidepressant action of these drugs occurs over several weeks, the timing of peak levels is relatively unimportant for efficacy but does have implications for side effects, especially during initial dosing.

Distribution. The tricyclic and tetracyclic compounds are basic lipophilic amines. As a result, they are concentrated in a variety of tissues throughout the body and have a high volume of distribution. In some organs, the heart for example, the drug is concentrated at levels exceeding those in plasma. Because of their lipid solubility, these compounds are extensively bound to plasma proteins. Binding of the tricyclic and tetracyclic drugs is generally 90 percent or greater. Exceptions are the hydroxymetabolites of the tricyclics, which have lower plasma protein binding.

Tricyclic compounds cross the placenta. Transfer of the tertiary compound amitriptyline is slightly higher than that for the secondary compound nortriptyline. Neonatal drug withdrawal can occur. The tricyclics are excreted in breast milk at concentrations similar to plasma. The actual quantity delivered, however, is small so that drug levels in the infant are usually undetectable or very low.

Metabolism and Elimination. After absorption, the tricyclics are taken up in the circulation, pass first through the liver, and metabolism of the drug begins. As a result of this *first-pass effect* the amount of the

compound that enters the systemic circulation is reduced.

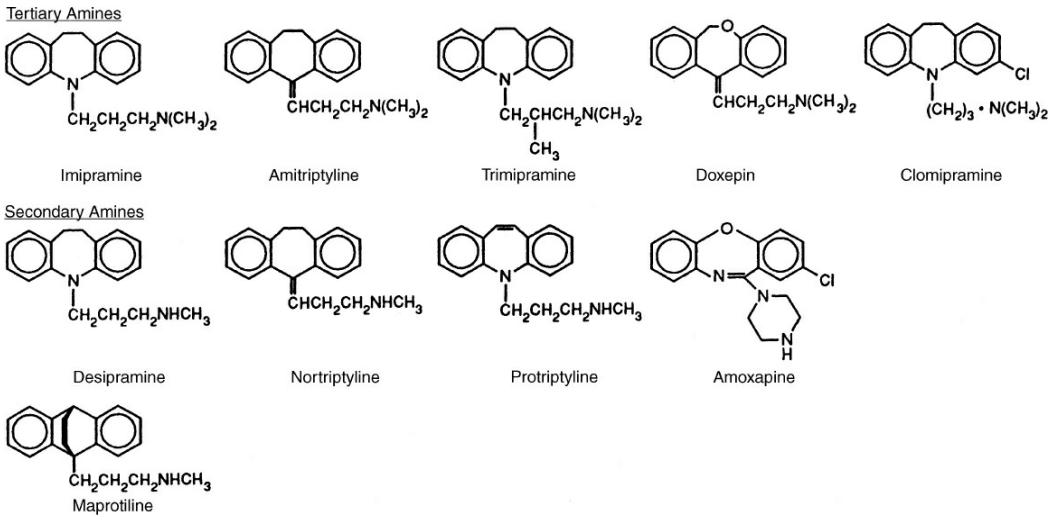

FIGURE 34.33–1. Molecular structure of tricyclics and tetracyclics.

Table 34.33–1. Receptor Affinity or Potency of Cyclic Antidepressants in Human Brain

	Receptor Affinity						Potency–U	
Drug	**α_1**	**α_2**	**Histamine Type 1**	**Muscarinic Type 1**	**Serotonin Type 1**	**Serotonin Type 2**	**Serotonin**	**Norep**
Tertiary tricyclic drugs								
Amitriptyline (Elavil)	3.7	0.11	91	5.6	0.53	3.4	23	
Clomipramine (Anafranil)	2.6	0.03	3.2	2.7	0.01	3.7	360	
Doxepin (Sinequan)	4.2	0.09	420	1.2	0.35	4.0	1.5	
Imipramine (Tofranil)	1.1	0.03	9.1	1.1	0.01	1.2	71	
Trimipramine (Surmontil)	4.2	0.15	370	1.7	0.01	3.1	0.67	0
Secondary tricyclic drugs								
Desipramine (Norpramin, Pertofrane)	0.77	0.01	0.91	0.50	0.01	0.36	5.7	
Nortriptyline (Aventyl, Pamelor)	1.70	0.04	10.0	0.67	0.32	2.3	5.6	
Protriptyline (Vivactil)	0.77	0.02	4.0	4.0	0.03	1.5	5.1	
Tetracyclic drugs								
Amoxapine (Asendin)	2	0.04	4.0	0.10	0.46	170	1.7	
Maprotiline (Ludiomil)	1.1	0.01	50.0	0.18	0.01	0.83	0.017	
Reference compounds								

Phentolamine (Regitine)	6.7	—	—	—	—	—	—
Yohimbine (Actibine)	—	62	—	—	—	—	—
Diphenhydramine (Benadryl)	—	—	7.1	—	—	—	—
Atropine	—	—	—	42	—	—	—
Buspirone (BuSpar)	—	—	—	—	26	—	—
Methysergide (Sansert)	—	—	—	—	—	26	—

Note: Affinity and potency = $10^{-7} \times 1/K_d$. K_d = equilibrium dissociation constant in molarity. Larger numbers indicate greater activity or potency.

Uptake potency data adapted from Tatsumi M, Groshan K, Blakely RD, Richelson E. Pharmacological profile of antidepressants and related compounds at human monoamine transporters. *Eur J Pharmacol.* 1997;340:249–258. Receptor affinity data adapted from Richelson E, Nelson A. Antagonism by antidepressants of neurotransmitter receptors of normal human brain in vitro. *J Pharmacol Exp Ther.* 1984;230:94–102.

Clearance of the tricyclic and tetracyclic compounds is principally the result of hepatic metabolism. Renal clearance accounts for only a small portion of drug elimination. Rates of hepatic metabolism vary widely from person to person. Elimination half-lives for most of the tricyclic and tetracyclic compounds are approximately 24 hours or longer, thus the drugs can be given once a day (Table 34.33–1). Amoxapine, which has a shorter half-life, is an exception.

There are two principal metabolic pathways: demethylation of the side chain and hydroxylation of the ring structure. Demethylation converts the tertiary amine to a secondary amine. Examples are the conversion of amitriptyline to nortriptyline and imipramine to desipramine. This process alters the activity of the compound. The tertiary amines are relatively more serotonergic, whereas the demethylated amines are relatively more noradrenergic. Hydroxylation of the compounds also occurs and results in hydroxymetabolites. The concentration 10-hydroxynortriptyline usually exceeds that of the parent. For desipramine, 2-hydroxydesipramine is present at levels of approximately 45 percent of the parent, but these ratios are quite variable. The ratio of hydroxymetabolite to parent increases in rapid metabolizers. Hydroxyimipramine and hydroxyamitriptyline are present at low concentrations and are clinically unimportant. The hydroxymetabolites are then conjugated and excreted. The conjugated metabolites are not active.

Hydroxynortriptyline and hydroxydesipramine are active compounds. Both have antidepressant activity. Hydroxydesipramine is comparable to the parent, in terms of norepinephrine reuptake blockade. There are two isomers of hydroxynortriptyline: E- and Z-10-hydroxynortriptyline. E-10-hydroxynortriptyline is present at levels four times higher than the Z-

isomer. E-10-hydroxynortriptyline blocks uptake of norepinephrine but is approximately 50 percent as potent as nortriptyline. The clinical significance of high levels of less potent hydroxynortriptyline is not entirely clear, but if high levels of hydroxynortriptyline interfere with the action of nortriptyline, this might explain the therapeutic window described for this drug. The hydroxymetabolites of nortriptyline and desipramine are less anticholinergic than their parent compounds but may have other effects. They appear to contribute disproportionately to the cardiac conduction abnormalities associated with the tricyclic antidepressants.

The principal metabolic pathway for amoxapine is hydroxylation. 7-Hydroxyamoxapine and 8-hydroxyamoxapine are produced. These compounds differ in their clinical activity. 7-Hydroxyamoxapine has high-potency neuroleptic properties but has a short half-life. 8-Hydroxyamoxapine is metabolized more slowly and appears to contribute to the drug's antidepressant action.

The tricyclics are metabolized by the cytochrome 450 (CYP) system. The CYP 2D6 pathway appears responsible for hydroxylation of desipramine and nortriptyline. Demethylation of the tertiary amine compounds appears to involve a number of CYP isoenzymes, including CYP 1A2, CYP 3A4, and CYP 2C19. These hepatic isoenzymes are under the control of specific genes, and the gene loci have been identified for several of these isoenzymes, including CYP 2D6. Approximately 5 to 10 percent of Caucasians are homozygous for the recessive autosomal CYP 2D6 trait, resulting in deficient hydroxylation of desipramine and nortriptyline. These individuals are known as poor metabolizers, whereas those with adequate CYP 2D6 isoenzyme activity are extensive metabolizers. Approximately 50 percent of those of Asian descent carry the 2D6*10 allele which results in reduced clearance by the 2D6 enzyme and are intermediate metabolizers. In addition, about 20 percent of Asians have a genetic polymorphism resulting in deficient CYP 2C19 metabolism. This pathway is involved in the metabolism of the tertiary tricyclic compounds.

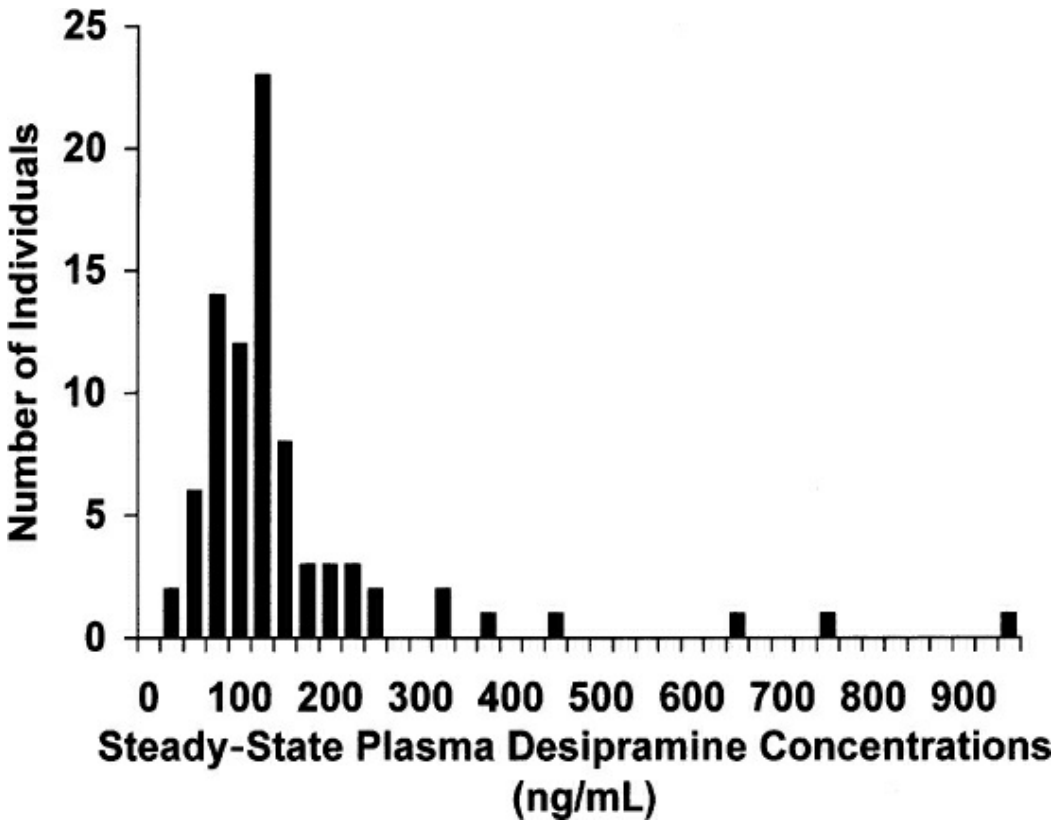

FIGURE 34.33–2. Distribution of steady-state desipramine (Norpramin, Pertofrane) concentrations in 83 inpatients receiving desipramine (2.5 mg/kg/day). (Adapted from Nelson JC. Use of desipramine in depressed patients. *J Clin Psychiatry.* 1984;45:10–16.)

The variability in plasma concentrations that results from these metabolic differences is substantial, often described as a 40-fold variation among individuals (Fig. 34.33–2). Even among the extensive metabolizers, there can be moderate interindividual variability in the rates of metabolism.

The concept of *steady state* is important for clinicians to understand if plasma concentration monitoring is used. Steady state is that point, on a fixed dose, at which plasma concentrations of the drug plateau. Steady state is achieved after five times the half-life.

If blood level monitoring is used, a trough sample (before the next dose) is drawn when the patient has reached steady state. At steady state, drug concentrations are relatively stable, as long as the dose is constant, the patient is compliant, and no drugs with interactive effects are added.

When the drug concentration is measured, the total of the free and bound tricyclic fractions are reported. Few laboratories are prepared to measure free levels, and yet free levels are directly related to the concentration of drug in the cerebrospinal fluid (CSF). The free concentration is dependent on dose and hepatic clearance and is not affected by absorption, the volume of distribution, or plasma protein binding. The latter is often misunderstood. If the binding protein concentration falls, the total (free and bound) concentration declines, but

the absolute free concentration is unaffected. If another drug affects binding, the absolute free concentration remains unaffected. In these instances, the *free fraction* may rise because the bound portion decreases, not because there is a change in the free concentration.

The day-to-day biological variability of drug concentrations is not frequently described. In outpatients, the day-to-day variability of nortriptyline levels (the coefficient of variation equals the standard deviation divided by the mean plasma concentration) has been reported to be 25 percent. In inpatients at steady state, the coefficient of variation of desipramine is approximately 10 to 15 percent. (If the average plasma concentration is 150 ng/mL, two-thirds of samples obtained will be between 135 and 165 ng/mL.) Research studies of drug concentrations have frequently used the average of two or three plasma samples drawn at steady state to estimate the level more precisely. For clinical purposes, usually one sample is drawn, and the clinician needs to remember that, even if the laboratory error is low, the biological variability is sufficiently great that blood levels are better viewed as estimates than precise measures.

Recently assays have become available that can genotype specific isoenzyme pathways such as 2D6. Genotyping might be useful for predicting how an individual would metabolize a group of drugs that are cleared by a specific pathway. Alternatively, a plasma drug concentration is more informative about the actual amount of a specific drug present and takes into account variations in absorption, distribution, and adherence in addition to differences in metabolic rate.

For most of the tricyclics, kinetics are relatively linear, although there are some exceptions. Desipramine, for example, appears to have nonlinear kinetics in the usual dose range. The patients most vulnerable to nonlinear changes are the rapid metabolizers—patients who, at usual doses, have low bloods levels. In cases of overdose, nonlinear changes are more likely to occur, and the clinician cannot assume that usual rates of drug elimination are maintained.

The pharmacokinetics of drug treatment in the elderly are complicated. A number of changes occur. For example, the ratio of fat to lean body mass increases, and cardiac output and hepatic blood flow decrease. There may be further changes associated with medical illness. Yet, the magnitude of these changes is usually relatively small in relation to the dramatic variability of hepatic metabolism. Metabolic changes with age appear to vary with the isoenzymes involved. The CYP 3A4 pathway, for example, slows with age. Most studies of the tertiary amines, such as imipramine, suggest that concentrations of these drugs are increased somewhat in the elderly. Concentrations of nortriptyline and desipramine, compounds metabolized primarily by CYP 2D6, are relatively unaffected by aging. Similarly, the relationship of drug concentrations to therapeutic effects appears to be relatively similar in younger and older adults for these

agents. Renal clearance of the hydroxymetabolites does decrease with age. Concentrations of hydroxynortriptyline may be substantially elevated in older patients.

In children, the clearance of tricyclic compounds is increased. Half-lives of imipramine are shorter, and ratios of desmethylimipramine to imipramine are higher, consistent with more rapid metabolism.

Pharmacodynamics

Mechanisms of Action. The tricyclic and tetracyclic antidepressants block the transporter sites for serotonin and norepinephrine and thus reduce uptake of these amines into the presynaptic neuron (Table 34.33–2). The tertiary amines have greater affinity for the serotonin transporter. The secondary amines primarily block the uptake of norepinephrine. Because the tertiary amines are demethylated to secondary amines, during the administration of amitriptyline, imipramine, or clomipramine, secondary amines are also present, thus serotonergic and noradrenergic effects occur. Because dopamine is inactivated by norepinephrine transporters in the frontal cortex, the tricyclics would be expected to increase dopamine concentrations in that region.

After the initial reuptake blockade, a sequence of events occurs. The tertiary tricyclic compounds, particularly clomipramine, inhibit the uptake of serotonin and concentrations of serotonin rise. In response, the presynaptic somatodendritic serotonin type 1A ($5\text{-}HT_{1A}$) autoreceptor reduces the firing rate of the presynaptic serotonin neuron, and concentrations of 5-hydroxyindoleacetic acid (5-HIAA), the major metabolite of serotonin, fall rapidly. Over a 2-week period, the presynaptic autoreceptor is desensitized, resulting in a normalization of the firing rate. At this point, serotonin transmission is enhanced.

Table 34.33–2.
Pharmacokinetic Parameters and Clinical Information for Tricyclic and Tetracyclic Drugs

Drug	Half-Life (hours)	Plasma Clearance (L/hour)	Antidepressant Dosage Range[a] (mg/day)	Presumed Therapeutic Plasma Concentrations (ng/mL)
Tertiary tricyclics				
Amitriptyline (Elavil)	5–45	20–70	150–300	—
Clomipramine (Anafranil)	15–60	20–120	150–300	>150[b]
Doxepin (Sinequan)	10–25	40–60	150–300	—
Imipramine (Tofranil)	5–30	30–100	150–300	>200[b]
Trimipramine (Surmontil)	15–40	40–105	—	—
Secondary tricyclics				

Desipramine (Norpramin, Pertofrane)	10–30	80–170	75–300	>115
Nortriptyline (Aventyl, Pamelor)	20–55	15–80	50–150	50–150
Protriptyline (Vivactil)	55–200	5–25	15–60	—
Tetracyclics				
Amoxapine (Asendin)	5–10	225–275	150–300	—
Maprotiline (Ludiomil)	25–50	15–35	100–225	—

[a]Dose range is a crude guideline; some fast or slow metabolizers need doses outside this range.

[b]Includes parent compound plus desmethylmetabolite.

The tricyclic agents as a group sensitize or upregulate *postsynaptic* 5-HT_{1A} receptors. Changes in the sensitivity of the presynaptic and postsynaptic receptors occur over a 2-week period. The onset of these changes is more consistent with the timing of antidepressant response than the initial uptake blockade.

Depression is associated with an increase in postsynaptic serotonin type 2 (5-HT_2) receptor density. A variety of antidepressants, including the tricyclics, have been shown to downregulate the 5-HT_2 receptors. The 5-HT_2 receptors principally mediate excitatory effects, whereas the 5-HT_{1A} receptors have an inhibitory effect in the brain, thus these two systems act in opposition. In preclinical experiments, administration of a 5-HT_2 antagonist enhances the effects of serotonin. Thus, the downregulation of 5-HT_2 receptors with tricyclic treatment appears to facilitate the effects of serotonin at postsynaptic 5-HT_{1A} receptors. In addition doxepin, clomipramine, amitriptyline, trimipramine, and nortriptyline act as antagonists at the 5-HT_2 receptor.

The sequence of events in the noradrenergic system is more complicated and less well understood. Following reuptake inhibition, concentrations of 3-methoxy-4-hydroxyphenylglycol (MHPG), a metabolite of norepinephrine that reflects the turnover of this neurotransmitter, fall rapidly, within 24 hours. Presumably, this is mediated by the presynaptic somatodendritic α_2-autoreceptor, which, sensing an increase in norepinephrine, lowers the firing rate of the presynaptic neuron. Different from the 5-HT system, the firing rate of the presynaptic noradrenergic neuron remains inhibited with chronic treatment. Norepinephrine concentrations, however, do increase at postsynaptic sites in the hippocampus and frontal cortex and norepinephrine transmission is increased. This may be mediated by desensitization of the *terminal* α_2-autoreceptor.

During chronic treatment the postsynaptic β-adrenergic receptor is downregulated or decreased in density. This downregulation was once proposed as a mechanism of tricyclic action. Instead, it appears more likely that it is a compensatory change. In fact, during continued administration of a noradrenergic reuptake inhibitor, the system remains upregulated.

This is reflected in the enhanced formation of the second messenger cyclic adenosine monophosphate (cAMP). Clinically, this is reflected as a persistent increase in heart rate.

Early theories of depression posited depletion of serotonin or norepinephrine, or both of these neurotransmitters. Administration of the tricyclic agents was thought to enhance the availability of both neurotransmitters. Subsequent theories suggested that the changes in presynaptic and postsynaptic receptors were a likely explanation for the delayed effects of these drugs in depression. Recent depletion studies suggest that the effects of antidepressant agents are mediated by serotonin and norepinephrine. Administration of a tryptophan-free diet rapidly depletes serotonin. Depressed patients, who have been successfully treated, experience transient return of symptoms after tryptophan depletion. Furthermore, relapse was most likely in those treated with serotonergic agents, whereas those treated with norepinephrine agents are relatively unaffected. Alternatively, administration of α-methyl-*para*-tyrosine, which interrupts the synthesis of catecholamines, causes relapse in patients successfully treated with noradrenergic agents, whereas those receiving serotonergic drugs are relatively unaffected. Tryptophan depletion in untreated depressed patients, however, has no effect on the patient's depression. These studies support the view that the actions of these two types of drugs are, in fact, mediated by serotonin and norepinephrine, but they do not necessarily imply that tricyclics are correcting *deficits* in these neurotransmitter systems.

Early theories of antidepressant action suggested that it might be possible to identify serotonergic and noradrenergic depressions based on the relative deficiencies of metabolites associated with each of these neurotransmitters and thus predict which type of antidepressant would be most effective. A number of studies investigated the predictive value of MHPG levels pretreatment, but a clear predictive link with a type of antidepressant was not established. These studies were, in part, hampered by the use of agents such as amitriptyline and imipramine, which are not selective.

The possible role of anticholinergic effects of the tricyclics in the mechanism of antidepressant action has not received much attention. These effects usually have been associated with unwanted adverse effects. Yet, as early as 1972, theories about the role of the cholinergic system in depression were advanced. Scopolamine, an anticholinergic agent, has rapid antidepressant effects given intravenously. Physostigmine and donepezil, which have cholinergic effects, have been noted to exacerbate depression or increase risk of relapse in patients with recurrent depression. Amitriptyline has anticholinergic effects at therapeutic doses and it is likely these actions would contribute to antidepressant activity; however, it would be difficult to disentangle the contribution of these effects relative to the

effects of monoamine reuptake inhibition.

Other possible mechanisms of action in depression have been explored including the effects of tricyclics on postsynaptic signal transduction, glucocorticoid receptors, neurotrophic factors, and gene expression. The antidepressants, including the tricyclics, appear to have neuroprotective effects. They raise levels of BDNF and promote neurogenesis in the hippocampus. A particularly intriguing finding is the observation of an association of decreased hippocampal volume with lifetime duration of *untreated* depression, implying that antidepressant treatment might protect against this outcome. Yet, at the present time, the role of these neuroprotective effects in mediating the actual antidepressant effects of tricyclics remains unclear.

The reuptake inhibition of the tricyclics and tetracyclic compounds also occurs in descending pathways of the spinal cord and contributes to analgesic effects. Animal studies suggest that serotonin and norepinephrine have independent synergistic effects on pain. In addition, the effects of tricyclics on sodium channels also contribute to their analgesic effects although these same effects appear to be responsible for the dose-dependent delay in cardiac conduction.

The tricyclic and tetracyclic compounds have *secondary* effects on a variety of other receptors (Table 34.33–2). In particular, these compounds block muscarinic receptors, producing anticholinergic effects. They block histamine-1 (H_1) receptors, and they block α_1- and α_2-adrenergic receptors. They affect fast sodium channels, and this explains their adverse cardiac effects. The potency of these various secondary effects varies considerably. Amitriptyline is the most anticholinergic. Doxepin is the most potent H_1 antagonist among the tricyclics, although it is less potent at this receptor than mirtazapine (Remeron).

Blood Concentrations and Relation to Response. Knowledge of the wide interindividual variability of tricyclic plasma concentrations led to the hope that drug level monitoring might ensure that therapeutic blood levels were achieved and might help to avoid toxic levels. In carefully selected endogenous or melancholic inpatients with major depression, treatment with adequate levels of imipramine or desipramine did result in response rates of approximately 85 percent. These studies suggested that adequate blood levels would help maximize response. Alternatively, the careful selection of patients in these studies may have resulted in samples that were not representative of typical clinical practice.

An American Psychiatric Association (APA) task force reviewing the evidence concluded that a relationship with blood levels and response had been demonstrated for imipramine, desipramine, and nortriptyline (Table 34.33–1). For imipramine, drug levels greater than 200 ng/mL were more effective than lower levels. For desipramine, levels greater than 115 ng/mL

were more effective. For desipramine and imipramine, blood levels in excess of 300 ng/mL were more likely to be associated with serious side effects. Effective plasma concentrations were also established for nortriptyline, but the relationship appeared to be curvilinear: Plasma levels between 50 ng/mL and 150 ng/mL were more effective than levels outside this window.

For amitriptyline, it has been difficult to establish a therapeutic relationship of levels and response. In part, this may be related to the fact that during amitriptyline administration, amitriptyline, nortriptyline, and hydroxynortriptyline are all present and active; and it is unclear if their effects are additive or if there is a more complicated relationship. During amitriptyline administration, responders usually have total amitriptyline and nortriptyline levels in the neighborhood of 150 to 250 ng/mL, but there is no good agreement between studies. For clomipramine, blood levels of 150 to 300 ng/mL (total of clomipramine and desmethylclomipramine) have been suggested for antidepressant effectiveness. Higher doses and levels are usually used in the treatment of OCD. The data relating blood levels and response are limited for the other tricyclic and tetracyclic compounds.

The clinical value of blood level monitoring has been the subject of much debate. Relationships between blood levels and response have been best demonstrated in melancholic inpatients. It has proven difficult to demonstrate a relationship of blood levels and response in depressed outpatients. In this group, placebo rates are high, the difference between drug and placebo is substantially reduced, and the effect of drug treatment is harder to detect. Moderately depressed outpatients may be a heterogeneous group and include patients who are not responsive to drug treatment. Finally, most studies are not designed to detect blood level–response relationships. Fixed dosing is required, and a significant portion of the sample must achieve plasma concentrations below the suspected threshold. If all patients achieve plasma drug concentrations that are above an effective threshold, no relationship with response will be found. It is logical to conclude that blood level relationships determined in severely depressed inpatients might be used as a guide for treatment of outpatients, but that has not been empirically confirmed. The case study provided here gives an example of blood level monitoring.

A 74-year-old man was admitted to the hospital with severe depression. Irritability was prominent. Desipramine was started but on 150 mg/day he had no response whatsoever. He was adherent to treatment. Some of his symptoms seemed atypical and diagnosis was questioned. A blood sample revealed a very low drug concentration. In fact the level suggested that if kinetics were linear, he would need a dose of 900 mg/day to achieve a therapeutic level. On 400 mg/day, the approved maximum dose, his level was still subtherapeutic. On 600 mg his level rose disproportionately, he achieved an adequate level, and responded. The case illustrates several issues. Blood levels can be useful for guiding dose adjustment in nonresponding patients. The extremely low plasma concentration on a conventional dose

suggested he might be an ultrafast metabolizer. Nonlinear kinetics are most likely to occur in fast metabolizers. Ironically, this 74-year-old man had the lowest concentration to dose ratio that these physicians had encountered in treating over 400 patients with desipramine.

The alternative question is whether blood level monitoring might help avoid toxicity. A variety of data support this view. The risk of delirium is substantially increased at amitriptyline plasma concentrations greater than 450 ng/mL and moderately increased at levels greater than 300 ng/mL. However, amitriptyline is the most anticholinergic tricyclic and is most likely to produce delirium. The risk of cardiac arrhythmia also is substantially increased at higher blood levels. First-degree atrioventricular block is increased with plasma concentrations of imipramine greater than 350 ng/mL and substantially increased at plasma concentrations greater than 450 ng/mL. The risk of seizures also increases in patients at higher doses and, presumably, higher blood levels, although a clear plasma level threshold has not been demonstrated for seizures. The risk of seizures for clomipramine is 1.66 percent with doses greater than 250 mg/day but is approximately 0.50 percent for doses less than or equal to 250 mg/day. In one large sample of patients with OCD at steady state on 250 mg/day of clomipramine, the average blood level was 155 ng/mL for clomipramine and 335 ng/mL for desmethylclomipramine, or a total of 490 ng/mL. For the older drugs, amitriptyline and imipramine, the risk of seizures also appears to be increased at doses greater than 250 mg/day. For patients who are especially vulnerable to delirium, delayed cardiac conduction, or seizures, adverse events can occur at levels below those cited above. After overdose, tricyclic blood levels greater than 1,000 ng/mL can be achieved, and the risks of delirium, stupor, cardiac abnormalities, and seizures are all substantially increased.

One of the problems in appreciating the relationship of toxicity to increased blood levels is that, although the risk may increase, actual rates are still low, and large samples are required to demonstrate any difference in risk. For certain types of toxicity, for example, delirium, there may be early warning signs, and the dose can be reduced unless the symptoms of delirium are mistaken for symptoms of the underlying psychiatric disorder. Alternatively, there may be no early signs of impending seizures or cardiac arrhythmia to alert the clinician to reduce the dose. Blood level monitoring might be most useful to reduce the risk of these particular side effects.

DESIGN AND INTERPRETATION OF CLINICAL DRUG STUDIES

Design

Early in the history of antidepressant studies relatively high rates of response to placebo were noted. As a consequence, open-label studies proved to be of limited value. Because of side effects, the tricyclics were started at low dose and then increased slowly. This tended to result in

inadequate dosing. Subsequently blood levels were employed to improve dosing adequacy; however, determination of effective blood levels itself proved challenging. Some early studies were quite brief (e.g., 3 to 4 weeks) and may have underestimated drug effects. Alternatively, many of the early studies were conducted in inpatients, in which drug–placebo differences were greater. As experience with the tricyclics grew, subtypes of depression were identified that were associated with greater or reduced response.

Interpretation

The clinical efficacy of the tricyclic and tetracyclic compounds in major depression is well established and has been extensively reviewed by John Davis and Alexander Glassman in the fifth edition of this text and in the Agency for Health Care Policy and Research (AHCPR) Clinical Practice Guideline. Imipramine is the best-studied compound. In 44 placebo-controlled studies, imipramine was more effective than placebo in 30 studies. Overall, in studies reporting response rates (N = 1,334), 65 percent of the patients *completing* treatment with imipramine were substantially improved, whereas 30 percent of those on placebo were similarly improved. In a meta-analysis of *intent-to-treat* response rates for placebo-controlled studies of tricyclics in outpatients reported in the AHCPR guideline, response rates were 51 percent for imipramine and 30 percent for placebo. In most comparison studies, the other tricyclic and tetracyclic antidepressants have been found to be comparable to imipramine in efficacy.

In addition to efficacy in acute depression, the tricyclic compounds are also effective for maintenance treatment. Several early studies demonstrated that maintenance treatment with a tricyclic would reduce the relapse rate associated with placebo by approximately 50 percent. These studies, however, often used low doses. Subsequently, an important study by the Pittsburgh group found that, if the full dose of imipramine used for acute treatment was continued, nearly 80 percent of the depressed patients maintained their improvement, whereas only 10 percent of those on placebo remained well at the end of a 3-year period. In this study, maintenance psychotherapy had an intermediate effect, with approximately 30 percent of the patients remaining well. This study again confirmed the value of maintenance treatment and underscored the importance of continuing the full dose of imipramine. The magnitude of the findings may, in part, relate to the sample selected; namely, patients with recurrent depression who had been symptom free between prior episodes. In clinical practice, if patients with chronic depression, prior residual symptoms, or comorbid psychiatric disorders are included, the number of patients who remain in remission may be lower.

Efficacy in Subtypes

The efficacy of the tricyclic compounds varies in different subtypes of depression. Kuhn thought it was essential to establish the efficacy of imipramine in *endogenous* inpatients. The tricyclic compounds have been frequently studied in severely depressed inpatients and were found to be effective. It has been suggested that the tricyclics may be more effective than the SSRIs in severe major depressive disorders with melancholic features. The issue remains unsettled, in part, because of the paucity of studies of the SSRIs in severe melancholic depression. In addition, it was unclear if the suggested advantage of some tricyclics in severe depression was unique to tricyclics as a class or applied to those tricyclics, particularly clomipramine, that had potent effects on both serotonin and norepinephrine.

The tricyclic and tetracyclic compounds have often been studied in anxious depression. Three of the compounds—doxepin, amoxapine, and maprotiline—have FDA approval for use in patients with depression and symptoms of anxiety. For many years, clinical lore suggested that amitriptyline was most effective for anxious depression. Direct comparison studies, however, found little evidence that these compounds differ in their efficacy for treatment of anxiety associated with depression. Alternatively three studies found that, among depressed patients, those with high levels of anxiety responded less well to amitriptyline, imipramine, or desipramine than those with lower levels of anxiety. Yet, these drugs are still more effective than placebo in anxious-depressed patients, and it is not established that other classes of antidepressants are more effective in anxious-depressed patients.

In patients with atypical features (defined by DSM-IV criteria), imipramine is more effective than placebo but significantly less effective than the MAOI phenelzine. Little information is available regarding comparisons of the tricyclics and the SSRIs in atypical depression.

Major depression with psychotic features is less responsive than nonpsychotic depression to tricyclic antidepressants used alone. Several open studies and one prospective study found that the tricyclics, when combined with an antipsychotic, were effective in psychotic depression. Yet other trials did not find an antidepressant/antipsychotic combination more effective than nortriptyline, imipramine, or trimipramine monotherapy. One double-blind study compared the efficacy of amoxapine with the combination of perphenazine (Trilafon) and amitriptyline, finding amoxapine to have comparable efficacy in psychotic depression. It was suggested that the antipsychotic action of 7-hydroxyamoxapine might explain the efficacy of this compound in psychotic depression.

In controlled studies of patients with bipolar depression, imipramine was found to be less effective than tranylcypromine (Parnate) or fluoxetine (Prozac). These studies and others suggest that the tricyclics appear to be more likely to induce mania than the serotonergic agents or bupropion

(Wellbutrin).

In the AHCPR review of drug trials in late-life depression, tricyclic antidepressants appear to be effective, but, overall, drug and placebo response rates appear to be lower than in nonelderly patients. Efficacy in the elderly may have, in part, reflected cautious dosing because of safety and tolerance concerns. In children and adolescents, the tricyclic antidepressants have not demonstrated superiority over placebo. Recently, an FDA review of clinical trials of second-generation antidepressants revealed an increased risk of suicidal thinking and suicide-related behaviors in children, adolescents, and young adults under the age of 25. Although this risk has not been established for the tricyclics, the warning has been extended to all antidepressants.

The tricyclic antidepressants are effective in dysthymia. Controlled trials suggest both imipramine and desipramine are effective for acute treatment of dysthymia and for maintenance treatment.

EFFECTS ON SPECIFIC ORGANS AND SYSTEMS

Somatic symptoms that occur during treatment are the final result of the interaction of direct effects of the medication on specific organs and indirect effects of the medication on depression and its associated somatic symptoms. Depression is accompanied by a variety of somatic symptoms. Headache, constipation, poor memory, drowsiness, and difficulty sitting still have been observed in more than 50 percent of untreated inpatients with major depression. Other symptoms, such as palpitations, dry mouth, tremors, light-headedness, and poor coordination, are reported by more than 40 percent of depressed patients before treatment. The determination of whether a somatic symptom is an adverse drug effect or a symptom of depression involves a judgment about whether the symptom is new or has worsened during drug treatment. Severely depressed patients may not remember accurately the severity of a symptom prior to treatment. During treatment, the strongest predictor of overall somatic symptom severity is the level of depression at the time of assessment.

Antidepressant drugs do have adverse effects. The in vitro potency or affinity of antidepressant compounds for various receptor sites (Table 34.33–1) is one method for comparing the likelihood that various agents will produce side effects. The other pharmacological factor that needs to be considered is the bioavailability of the drug.

Central Nervous System

In the CNS, the principal action of the tricyclic and tetracyclic agents is to alleviate depression. These agents do not have a direct mood-elevating effect. They do not produce euphoria. Rather, they reduce the symptoms of depression. Apparently related to their effects on brain amines, these

compounds can cause or aggravate psychotic symptoms and can induce mania.

The central anticholinergic and antihistaminic effects of the tricyclics and tetracyclics can produce confusion or delirium. The incidence of delirium is dose dependent and increases at blood levels greater than 300 ng/mL and may rise to 67 percent in patients with blood levels greater than 450 ng/mL and in patients receiving the tertiary amines, particularly amitriptyline. Patients with dementia or early Alzheimer disease are more vulnerable to the anticholinergic effects of the tricyclics, and delirium can occur in these patients at lower drug concentrations.

Seizures can occur with all of the tricyclic and tetracyclic agents. Seizures are dose- and blood-level related. For clomipramine, the risk for seizures is reported to be 0.5 percent in doses up to 250 mg/day. At doses greater than 250 mg/day, the seizure risk increases to 1.67 percent. For maprotiline, the overall risk of seizures is reported to be 0.4 percent, but, again, the risk of seizures increases at doses greater than the maximum recommended dose of 225 mg/day. Determination of seizure risk for some of the older compounds was hampered by less rigorous data collection at the time of marketing. In one meta-analysis of data for imipramine, patients receiving less than 200 mg had a seizure rate of 0.1 percent. In patients receiving more than 200 mg/day, the rate was 0.6 percent. In patients receiving 250 to 450 mg/day, reported rates varied from 1 to 4 percent, but these samples were often small, thus, these rates had wide confidence intervals. Seizure rates for the secondary amines have not been well described. The mechanism by which tricyclics produce seizures is not well understood, although it is proposed that antidepressant drugs may act at the GABA receptor chloride–ionophore complex, inhibiting chloride conductance and consequently inducing convulsions. The risk of convulsions is clearly increased in patients with predisposing factors, including a prior history of seizures, brain injury, and concurrent use of neuroleptics. The risk of seizures is substantially increased after overdose.

A fine rapid tremor can occur with the tricyclic agents, tends to be dose dependent, and may be a clinical indication of an elevated blood level. Because the 7-hydroxymetabolite of amoxapine has neuroleptic properties, administration of amoxapine carries the potential risk of neuroleptic malignant syndrome, which has been reported, and tardive dyskinesia.

A 73-year-old male was referred for evaluation of his antidepressant. He had had a good response to desipramine but it had taken some time to achieve it. He now had a fine rapid tremor that was quite bothersome. Both the patient and his physician were concerned that reducing the dose would compromise antidepressant efficacy. A blood sample revealed a desipramine concentration over 300 ng/mL. His dose was adjusted to achieve a level of about 150 ng/mL. His tremor resolved but the antidepressant effect was maintained.

Several of the tricyclic compounds and maprotiline have clinically

significant antihistaminic effects. Doxepin is the most potent H_1-receptor blocker of the tricyclics, although mirtazapine and olanzapine are more potent. Central H_1-receptor blockade can contribute to sedation and delirium. Antihistaminic effects also appear to be related to the increased appetite and associated weight gain.

Autonomic Nervous System

The tricyclics block muscarinic receptors and can cause a variety of anticholinergic side effects. These include dry mouth, constipation, blurred vision, urinary hesitancy, and an ocular crisis in patients with narrow-angle glaucoma. The tricyclic and tetracyclic compounds vary substantially in their muscarinic potency (Table 34.33–2). Amitriptyline is the most potent, followed by clomipramine. Desipramine is the least anticholinergic. Amoxapine and maprotiline also have minimal anticholinergic effects. Anticholinergic effects can contribute to tachycardia, but tachycardia commonly occurs in patients receiving desipramine, which has minimal anticholinergic effects. It seems likely that noradrenergic effects also contribute to increased pulse. Increased sweating is another secondary effect of the tricyclic compounds that can occur. The mechanism for this symptom is unclear but may be associated with noradrenergic effects.

Cardiovascular Effects

Orthostatic hypotension is one of the most common reasons for discontinuation of tricyclic antidepressant treatment. Orthostatic hypotension can occur with all of the tricyclics but is less pronounced with nortriptyline. It has been suggested that orthostatic hypotension is caused by α_1-adrenergic blockade; however, it is the postural reflex that is primarily affected. Resting supine blood pressure may be unaffected or may be raised. Orthostatic hypotension is most likely to occur or is most severe in patients who have pre-existing orthostatic hypotension.

Desipramine has been reported to raise supine blood pressure in younger patients, although it is not clear that this effect is limited to that age group. This effect may be similar to that reported for venlafaxine (Effexor).

Tachycardia occurs with all the tricyclics, not just the more anticholinergic agents. Supine and postural pulse changes can occur. Patients do not accommodate to the pulse rise, which can persist for months. Tachycardia is more prominent in younger patients, who appear to be more sensitive to sympathomimetic effects, and is one of the most common reasons for drug discontinuation in this age group.

The exploration of the effect of tricyclic antidepressants on cardiac conduction has received much attention. *Cardiac arrhythmia* is the principal cause of death after overdose. As a result, for many years, there

was great concern about the use of tricyclic antidepressants in patients with heart disease. The effect of these agents has now been well described. These agents have type I antiarrhythmic qualities or quinidine (Cardioquin)-like effects. Apparently, through inhibition of Na^+:K^+ adenosine triphosphatase (ATPase), the tricyclics stabilize electrically excitable membranes and delay conduction, particularly His-ventricular conduction.

At therapeutic blood levels, the tricyclics can have beneficial effects on ventricular excitability. Alternatively, in patients with pre-existing conduction delay, the tricyclic antidepressants can further delay conduction and cause heart block. High drug plasma levels further increase the risk of cardiac toxicity. For example, first-degree atrioventricular heart block is increased with imipramine plasma concentrations greater than 350 ng/mL and is increased more than 30-fold in patients with plasma concentrations greater than 450 ng/mL.

The tricyclic antidepressants do not appear to adversely affect cardiac contractility or cardiac output. Studies using radionuclide angiography indicate no adverse effect of imipramine or doxepin on cardiac output, even in patients with diminished left ventricular ejection fractions. However, it is noted that severe orthostatic hypotension was common in these patients.

Another potential hazard of tricyclic treatment was suggested by the Cardiac Arrhythmia Suppression Trial (CAST) studies, which found that type I antiarrhythmic drugs given after myocardial infarction actually increased the risk of sudden death. Alexander Glassman and associates suggested that the tricyclics may pose similar risks. Thus, the clinician considering long-term use of the tricyclics in patients with ischemic heart disease needs to weigh this possible risk against possible benefits.

Hepatic Effects

Drug-induced hepatitis can occur with these agents. Mild increases of liver enzymes are not uncommon and usually can be monitored safely over a period of days or weeks without apparent harmful consequences. However, acute hepatitis can occur quickly and is associated with high enzyme levels, for example, serum aspartate aminotransferase (AST) levels may be elevated 20-fold. The rise in enzymes can be hepatocellular or cholestatic in form and may precede clinical symptoms, especially in the hepatocellular form. This is a dangerous and potentially fatal condition. The antidepressant must be discontinued. The etiology is not well established but, in some cases, appears to be a hypersensitivity reaction. The risk of severe drug-induced hepatitis is not well established but appears to be between 1 in 1,000 and 1 in 100.

THERAPEUTIC INDICATIONS

Indications

The tricyclic and tetracyclic compounds have a variety of clinical uses. In fact, it is interesting to consider what these agents might have been called had their effectiveness in panic disorder, OCD, pain, or other syndrome been described before their antidepressant action was observed.

The FDA approved all of the tricyclic and tetracyclic compounds discussed in this section for use in the United States for the treatment of depression, with the exception of clomipramine. In Europe, clomipramine is also used for depression and, in fact, is regarded by many as the most potent antidepressant. Clomipramine acts on serotonin and norepinephrine, and the empirical data suggest that it is more effective than selective serotonin agents in severely depressed inpatients.

Although a variety of antidepressants appear to be effective in depressive disorders, this is not the case for OCD. In this instance, the serotonergic antidepressants, such as clomipramine, appear to be substantially more effective than noradrenergic tricyclics, such as desipramine. Among the tricyclic drugs, clomipramine is the only agent approved by the FDA for use in OCD.

Other Indications

The tricyclic agents have been used for a variety of other clinical indications without FDA approval. The tricyclic antidepressants are effective in panic disorder; in fact, imipramine was the first drug found to be effective in these patients.

Amitriptyline has been a popular drug for treatment of chronic pain syndromes and migraine headache prophylaxis. Over 60 placebo-controlled clinical trials of tricyclics or tetracyclics have been conducted in various chronic pain syndromes including headache, back pain, fibromyalgia, gastrointestinal discomfort, idiopathic pain, and tinnitus. In fact, the magnitude of their effects in chronic pain appears greater than that in depression. For example, in pain studies the average drug–placebo difference in response rates was 32 percent. The analgesic effects appeared, at least in part, to be independent of antidepressant effects. In humans, comparison studies suggested tricyclics with effects on both serotonin and norepinephrine were more effective than agents selective for one neurotransmitter. During treatment of pain, the usual dose of amitriptyline is lower (e.g., 75 mg/day) than that in depression, and effects occur more quickly, usually within the first 2 weeks. These differences suggest a different mode of action of amitriptyline in pain syndromes. The tricyclics appear to be of some benefit in chronic fatigue syndrome.

For many years, desipramine was commonly prescribed for ADHD in children. Reports of cases of sudden death, apparently of cardiac origin, in children younger than 12 years of age substantially reduced the use of this drug in younger patients. Imipramine has been used with FDA approval for treatment of nocturnal enuresis in children. Usual doses are 25 to 50 mg at

bedtime. It is doubtful, however, that the risk of cardiac problems would be substantially less with imipramine than with desipramine in young children.

Nortriptyline has been shown to be helpful for smoking cessation in controlled studies, but it is not approved for use in this syndrome nor has it been actively marketed for this condition.

Taking advantage of its anti-histaminic effects, topical doxepin (Zonalon) is FDA approved for use in pruritus associated with atopic dermatitis and lichen simplex chronicus. Doxepin (Silenor) is approved for use in insomnia at doses of 3 or 6 mg/day.

Tricyclic antidepressant drugs have been extensively studied in schizophrenic patients. However, in the absence of a major depressive syndrome, they appear to be of limited value, and their use for treatment of negative symptoms has been eclipsed by the introduction of the atypical neuroleptics.

PRECAUTIONS AND ADVERSE REACTIONS

Central Nervous System Effects

The central effects of the tricyclics can produce delirium and seizures. Both of these adverse effects are dose dependent and become more frequent at elevated blood levels. The incidence of delirium increases at blood levels greater than 300 ng/mL and becomes common in levels greater than 450 ng/mL, especially with amitriptyline. The clinician should be aware of this possibility because the signs of delirium may be mistaken for a worsening of the depression. This diagnostic dilemma can be especially problematic in patients with psychotic depression. Patients with concurrent dementia are particularly vulnerable to the development of delirium. Intramuscular (IM) or intravenous (IV) physostigmine (Eserine Salicylate) can be used to reverse or to reduce the symptoms of delirium. However, the short duration of action of this agent makes its continued use difficult.

Because delirium and seizures are dose- and blood-level dependent, the first goal is to reduce the risk by avoiding unnecessarily high serum concentrations. Some experts advocate routine blood level monitoring. This question continues to be debated. A strong case can be made for blood level monitoring in patients at increased risk, for example, the elderly, patients with dementia, or patients with a history of seizures. Blood levels can also be helpful in patients who fail conventional doses to determine that high doses are required to reach therapeutic levels.

The tricyclic antidepressants can cause a fine rapid tremor. This particular adverse event is often associated with elevated plasma concentrations and may improve with dose reduction.

Because the 7-hydroxymetabolite of amoxapine has neuroleptic effects, this agent can potentially cause tardive dyskinesia and neuroleptic

malignant syndrome. Although the occurrence of these adverse events is rare, the seriousness of the risk serves as a relative contraindication to this product, unless the clinical condition warrants the use of an agent with antipsychotic properties.

Autonomic Effects

The tricyclics can cause a variety of anticholinergic side effects, including dry mouth, constipation, visual changes, urinary hesitancy, delirium, and an ocular crisis in patients with narrow-angle glaucoma. Use of nortriptyline or desipramine, which is less anticholinergic, can help reduce the likelihood of these problems.

Usually, anticholinergic effects are annoying but not serious. In some circumstances, however, these effects can be severe. An ocular crisis in patients with narrow-angle glaucoma is an acute condition associated with severe pain. Urinary retention can be associated with stretch injuries to the bladder. Constipation can progress to severe obstipation. (Paralytic ileus has been described but is rare.) Medication must be discontinued, and appropriate supportive measures must be instituted. Elderly patients appear most vulnerable to this adverse consequence. The frequency of severe anticholinergic adverse reactions is increased by concomitant neuroleptic administration.

Some anticholinergic effects may benefit from other interventions. Patients with urinary hesitancy may benefit from the use of bethanechol (Urecholine) at a dose of 25 mg three or four times daily. Patients with constipation often benefit from regular use of stool softeners. Patients with narrow-angle glaucoma, who are receiving pilocarpine (Pilocar) eye drops regularly, can be treated with a tricyclic, as can those who have had an iridectomy. Tricyclic agents do not affect patients with chronic open-angle glaucoma.

Cardiovascular Effects

Orthostatic hypotension is the most common reason for discontinuation of tricyclic antidepressant treatment. It is most likely to be problematic in patients with pre-existing orthostatic hypotension. The elderly are more likely to have pre-existing hypotension and are also more vulnerable to the consequences of orthostatic hypotension—falls and hip fractures. Orthostatic hypotension can occur at blood levels that are relatively low. Although the subjective experience of light-headedness usually improves with time, the actual orthostatic blood pressure change does not accommodate within a reasonable period of time (e.g., 4 weeks). As a consequence, unless the plasma level is elevated and the dose can be reduced, patients who experience serious symptomatic orthostatic hypotension may not be treatable with a tricyclic antidepressant.

Fludrocortisone (Florinef) has been used to raise blood pressure, but it is not very effective. If patients are taking an antihypertensive, it may be possible and helpful to reduce these agents. Orthostatic hypotension is less common with nortriptyline.

Tachycardia occurs with all the tricyclics, not just the more anticholinergic agents. The supine pulse rate commonly increases 8 to 10 beats/min. The standing pulse can be markedly elevated. Patients do not accommodate to the pulse rise, which can persist for months. Although an 8 to 10 beat/min increase in heart rate may be subjectively tolerated by the patient, in the elderly, a sustained increase in cardiac work can be detrimental. Tachycardia is more prominent in younger patients, who appear to be more sensitive to the sympathomimetic effects of these agents, and is one of the most common reasons for drug discontinuation in this age group. Propranolol (Inderal) at low doses may help to reduce the tachycardia, unless it is associated with significant hypotension.

Cardiac Conduction Effects

The tricyclic antidepressants have type I antiarrhythmic effects. In patients without conduction delay, the tricyclics are well tolerated. They may be helpful in patients with premature ventricular beats. However, in patients with pre-existing conduction delay, the tricyclic antidepressants can further delay conduction and can cause heart block. Patients with a QTc interval of 450 msec or greater are at increased risk and are not candidates for tricyclic antidepressant treatment. The risk of cardiac toxicity is increased in patients with high plasma levels. For example, first-degree atrioventricular heart block is increased with imipramine plasma concentrations greater than 350 ng/mL and is increased more than 30-fold in patients with plasma concentrations greater than 450 ng/mL.

Hepatic Effects

Mild increases of liver enzymes usually can be monitored safely over a period of days or weeks without apparent harmful consequences. However, acute hepatitis develops quickly, over 3 to 4 days, and is associated with high enzyme levels, for example, a 20-fold increase in AST levels. This can be life-threatening, and the antidepressant must be discontinued. If it is clear which drug caused the hepatitis, that drug should not be introduced again, because the next reaction may be more severe. However, it is not uncommon for the patients to be receiving several medications, so that the offending agent is hard to identify. If a random blood test indicates mildly elevated liver enzymes, enzyme levels can be followed for a few days to distinguish the more severe form of hepatitis from mild enzyme elevations.

Other Adverse Effects

Sexual dysfunction is a potential problem with a variety of antidepressant drugs but appears to be more frequent with clomipramine. Unlike many of the side effects that are more likely to be reported early in treatment, this side effect may be more bothersome to the patient in the second or third month of treatment after the depression has improved. Numerous pharmacological antidotes have been described but without much success.

Tricyclic antidepressants can cause allergic skin rashes. They are sometimes associated with photosensitivity reactions. Various blood dyscrasias also have been reported, but fortunately, these are very rare.

Overdose

Because the risk of overdose is increased in depressed patients, the target population for which antidepressants are used and the lethality of antidepressant drugs in overdose is of great concern. A tricyclic overdose of 10 times the daily dose can be fatal. Death most commonly occurs as a result of cardiac toxicity. However, seizures and CNS depression can occur. Amitriptyline is one of the leading causes of death by overdose for any single drug in the United States. A five year study (2000 to 2004) of single drug ingestions found more deaths from amitriptyline than all other TCAs combined and it accounted for 80 percent of fatalities from all antidepressant overdoses. The other tricyclic and tetracyclic compounds are also dangerous in overdose even if their use has diminished. Desipramine has the highest fatality rate (deaths per ingestions) among the tricyclics. Amoxapine has been reported to produce high rates of seizure in overdose. Because the second-generation antidepressants are safer in overdose, subtle differences in lethality between the tricyclic and heterocyclic compounds have become relatively unimportant.

Teratogenicity and Pregnancy-Related Risks

A definitive link between the tricyclic compounds and tetracyclic compounds, and teratogenic effects has not been established, although isolated reports of morphogenesis have been reported. Tricyclics cross the placenta and neonatal drug withdrawal can occur. This syndrome includes tachypnea, cyanosis, irritability, and poor sucking reflex. If possible, tricyclic and tetracyclic medications should be discontinued 1 week before delivery. Recently, norepinephrine and serotonin transporters have been identified in the placenta and appear to play an important role in the clearance of these amines in the fetus. The understanding of the effects of reuptake inhibitors on these transporters during pregnancy is limited, but one study compared intelligence and language development in 80 children exposed to tricyclics during pregnancy with 84 children exposed to other nonteratogenic agents and found no deleterious effects of the tricyclics. The tricyclics are excreted in breast milk at concentrations similar to plasma.

The actual quantity delivered, however, is small so that drug levels in the infant are usually undetectable or very low. Because the risk of relapse is a serious concern in patients with recurrent depression, and these risks may be increased during pregnancy or the postpartum period, the risks and benefits of continuing or withdrawing treatment need to be discussed with the patient and weighed carefully.

DRUG INTERACTIONS

There are two principal types of drug interactions: pharmacodynamic and pharmacokinetic. Pharmacodynamic interactions are those in which the actions of two drugs interact. The effects may be deleterious. Because the tricyclics have acute effects on amine uptake, they are potentially dangerous when given with the MAOIs. The most dangerous sequence is to give a large dose of a tricyclic to a patient already on an MAOI. This can result in a sudden increase in catecholamines and a potentially fatal hypertensive reaction. If used cautiously, these two compounds can be used together to treat patients with refractory depression. Treatment is begun with lower doses, and the two compounds are started together, or the tricyclic is started first. Once begun, coadministration may reduce the risk of tyramine reactions; however, because the protective effect is variable and unpredictable, the usual MAOI diet is maintained.

Other pharmacodynamic interactions can occur. By blocking the transporters, the tricyclics block the uptake and thus interfere with the action of guanethidine (Ismelin). Quinidine poses two problems. First, the tricyclics have quinidine-like effects, thus, their effects on cardiac conduction are potentially additive with those of quinidine. In addition, quinidine is a potent enzyme inhibitor that can raise tricyclic levels, further adding to the problem.

Pharmacokinetic Interactions

A number of drugs can block the metabolic pathways of the tricyclics, resulting in higher and potentially toxic levels. Desipramine has been of particular interest, because its metabolism is fairly simple via the CYP 2D6 isoenzyme. Because there are no major alternative pathways, inhibition of CYP 2D6 can result in high desipramine plasma levels. Quinidine, mentioned previously, is a potent CYP 2D6 inhibitor. Several psychotropic drugs are 2D6 inhibitors including fluoxetine, paroxetine (Paxil), duloxetine (Cymbalta), and bupropion. For example, fluoxetine 20 mg/day, on average, raises desipramine levels three- to fourfold in extensive metabolizers. In slow metabolizers, enzyme inhibitors have less of an effect, because the patient is already deficient in this enzyme. At moderate doses sertraline (Zoloft), citalopram (Celexa), and escitalopram (Lexapro) have less potent effects on CYP 2D6 but at high doses may have appreciable

effects in some patients. Venlafaxine, nefazodone (Serzone), and mirtazapine appear to have minimal effects on CYP 2D6. It would be expected that CYP 2D6 inhibitors would also block nortriptyline metabolism, but the magnitude of this interaction has not been well studied.

The other group of drugs commonly used in psychiatry, which interact with the tricyclics, is the antipsychotic drugs. Chlorpromazine and perphenazine, at antipsychotic doses, inhibit CYP 2D6. The magnitude of the effect appears to be somewhat less than that of fluoxetine. For example, perphenazine, on average, raises desipramine levels twofold, but this effect varies with dose and varies with the neuroleptic used.

Because the tertiary tricyclics are metabolized by several pathways (CYP 1A2, CYP 3A4, and CYP 2C19), inhibition of one of these pathway has less of an effect on these compounds. Drug interactions with the tertiary amines do occur but appear to be less robust than that for desipramine.

Enzyme induction also occurs. Unlike enzyme inhibition, which occurs quickly, enzyme induction requires synthesis of new enzyme. As a result, the full effect of an enzyme inducer may take 2 to 3 weeks to develop and, similarly, if the inducer is discontinued, takes 2 to 3 weeks to dissipate. Barbiturates and carbamazepine (Tegretol) are two potent inducers commonly encountered in psychiatry. Although CYP 2D6 is a noninducible isoenzyme, phenobarbital (Barbita) reduces the availability of desipramine by approximately 40 percent. Apparently, CYP 3A4 is induced, becomes a more important pathway, and metabolism of desipramine is increased. It can be difficult to attain an effective blood level with a tricyclic in the presence of barbiturates or carbamazepine. Phenytoin (Dilantin) is also an enzyme inducer, but the magnitude of its effects on the tricyclics appears to be less dramatic. Nicotine induces the CYP 1A2 pathway and may affect the tertiary tricyclics, but the secondary tricyclics, such as desipramine, are less affected.

Alcohol has a complicated interaction with the tricyclics. Acute ingestion of alcohol can reduce first-pass metabolism, resulting in higher tricyclic levels. Because of the association of alcohol use with overdose, this can be an important factor resulting in higher tricyclic levels after overdose. Alternatively, chronic use of alcohol induces hepatic isoenzymes and may lower tricyclic levels.

Although most of the known drug–drug interactions have focused on interactions involving the CYP450 system, there is a growing body of literature examining the role of P-glycoproteins (P-gp) in drug interactions. P-gp play an active role at the blood–brain barrier in the efflux of some drugs from the CNS. Several of the tricyclics—amitriptyline, desipramine, doxepin, imipramine, nortriptyline, and trimipramine—are P-gp substrates. Drugs such as verapamil and quinidine inhibit P-gp, potentially leading to increased CNS concentrations of P-gp substrates. Some

antidepressants such as sertraline and paroxetine are also P-gp inhibitors. The clinical consequences of these potential interactions in humans are not well established. Polymorphisms of the ABCB1 gene that encodes P-gp have been associated with differences in response to antidepressants that are P-gp substrates suggesting that changes in P-gp activity are clinically relevant.

LABORATORY INTERFERENCES

The tricyclic compounds are present at low concentrations (ng/mL) and are not likely to interfere with other laboratory assays. It is possible that they may interfere with the determination of conventional neuroleptic blood concentrations because of their structural similarity and the low concentrations of some neuroleptics.

Two types of assays are performed for the tricyclic and tetracyclic antidepressants. After overdose, an immunoassay is commonly performed to determine if a tricyclic is present and to estimate the level. This assay is sensitive, simple to perform, and can be done quickly. The immune assays, however, are not specific, can cross-react with metabolites and other drugs, and thus are not precise quantitative measures. For therapeutic drug monitoring, a quantitative measure is needed and a slower turnaround time for the laboratory is usually acceptable. For this purpose, high-performance liquid chromatography (HPLC) is sufficiently precise and is usually employed. In a competent laboratory, the coefficient of variation is less than 10 percent. This assay is relatively specific; however, other drugs can interfere. Because there are many modifications of the HPLC technique, and because the drugs that interfere depend on these modifications (the extraction method, the column used), the interfering drugs vary by site. A clinician who frequently obtains tricyclic plasma levels is best advised to use one laboratory and to find out from that laboratory what agents interfere with his or her assay.

DOSAGE AND ADMINISTRATION

At one time, amitriptyline and imipramine were the two most commonly prescribed antidepressants in the United States, but because of their anticholinergic and antihistaminic side effects, their use declined and nortriptyline and desipramine became more popular. Nortriptyline has the least effect on orthostatic hypotension, and desipramine is the least anticholinergic. Because therapeutic plasma concentrations have been established for nortriptyline and desipramine, blood level monitoring can be used to ensure that a therapeutic level is achieved.

Most studies of nortriptyline find that an average dose needed to reach a blood level of 100 ng/mL is approximately 75 mg/day; however, there is considerable variability. For desipramine, the average dose required to reach a target level of 150 ng/mL is 3.0 mg/kg or 180 mg/day for a 60-kg

patient. Nortriptyline and desipramine can be adjusted to a full dose quickly, but the clinician needs to assess factors such as medical health, age, other medications present, and the severity of the depression, when deciding how quickly to adjust dose. The dose can be given once a day at bedtime and if tolerated, can be increased every 3 days. The target dose can then be determined in one of two ways. An average dose can be targeted (e.g., nortriptyline 75 mg/day) and blood levels employed only if the patient fails to respond. Alternatively a blood level can be drawn after 5 to 7 days on a moderate dose, 50 to 75 mg/day, and used to guide further dose adjustment for the trial. Because the kinetics of nortriptyline are linear, a proportional adjustment in dose can be made to reach a target level of 100 ng/mL. The advantage of the latter method is that one drug trial at an adequate blood level should establish if the drug is effective. Because nortriptyline has a therapeutic window, if the plasma level is near the upper limit of 150 ng/mL, a reduction in dose could be considered. Desipramine dose can also be adjusted relatively proportionately, with the exception that patients with low initial plasma levels (rapid metabolizers) are likely to have nonlinear changes. Plasma concentration–guided dose adjustment should reduce the time to achieve an effective dose (as opposed to a series of trials at gradually increasing doses) and should be of benefit in severe depression. Once the patient is at an adequate dose, treatment should then be continued for an adequate duration. At the time of the initial clinical trials of the tricyclics, trial duration was often short, 4 to 6 weeks. Recently, trials of up to 12 weeks have been advocated to achieve remission. Although it was suggested that tricyclics were more rapid acting than the SSRIs, this has not been confirmed. A 12-week comparison trial of nortriptyline and sertraline in late-life depression found no difference in the timing of response. A retrospective analysis of that data concluded that for both drug classes, a 12-week trial was useful to achieve remission, but that patients showing minimal improvement after 4 to 6 weeks were not likely to remit and other treatments should be considered. Doses and therapeutic plasma concentrations for other tricyclic and tetracyclic compounds are shown in Table 34.33–1.

If patients experience side effects, this approach may be altered. If patients on desipramine report insomnia, the dose can be given early in the day. Sedation and anticholinergic effects may diminish with time, and gradual dose adjustment may help manage these effects. Although subjective light-headedness may improve with time, blood pressure and pulse changes do not usually accommodate during the first 4 weeks, and gradual dosing does not help avoid these effects. In patients at high risk for side effects, the dose can be increased gradually to assess tolerance. If anticholinergic side effects occur, allowing more time between dose increases allows the patient to accommodate to these side effects.

Table 34.33–3.
Tricyclic and Tetracyclic Drug Preparations

Drug	Tablets	Capsules	Parenteral	Solution
Imipramine (Tofranil)	10, 25, and 50 mg	75, 100, 125, and 150 mg	12.5 mg/mL	—
Desipramine (Norpramin, Pertofrane)	10, 25, 50, 75, 100, and 150 mg	—	—	—
Trimipramine (Surmontil)	—	25, 50, and 100 mg	—	—
Amitriptyline (Elavil)	10, 25, 50, 75, 100, and 150 mg	—	10 mg/mL	—
Nortriptyline (Aventyl, Pamelor)	—	10, 25, 50, and 75 mg	—	10 mg/5 mL
Protriptyline (Vivactil)	5 and 10 mg	—	—	—
Amoxapine (Asendin)	25, 50, 100, and 150 mg	—	—	—
Doxepin (Sinequan)[a] (Silenor)	— 3, 6 mg	10, 25, 50, 75, 100, and 150 mg	—	10 mg/mL
Maprotiline (Ludiomil)	25, 50, and 75 mg	—	—	—
Clomipramine (Anafranil)	—	25, 50, and 75 mg	—	—

[a]Doxepin (Zonalon) is supplied as a 5% cream (30 g or 45 g).

Prospective Dosing Techniques

Conventional dosing requires administration of a given dose for a long enough period of time to determine if that dose is effective. Sometimes two or even three trials are needed to determine the effective dose. Once therapeutic blood levels were established for the tricyclics, methods were described for rapid dose adjustment to achieve an effective drug concentration quickly. Various methods were reported. They all involved obtaining one or more blood samples after an initial test dose from which a therapeutic dose could be estimated. Once the therapeutic dose of nortriptyline or desipramine was determined, the full dose could be given the next day. Prospective dosing methods were also reported for amitriptyline and imipramine; however, sedation and anticholinergic side effects limited rapid dose escalation of these drugs. Rapid prospective dosing methods are limited, however, by laboratory capability. The laboratory performing the assay must be prepared to determine drug concentrations accurately at low levels (below the therapeutic range) and must be able to report results quickly. Most laboratories are not prepared to do either. A more practical and clinically feasible method is to start the patient on a low, moderate, fixed dose, obtain a blood sample after 5 to 7 days on that dose, and then make further adjustments based on that level.

Exceptions

In elderly depressed patients, gradual dosing may be required to assess tolerance, but studies of desipramine and nortriptyline indicate that the final doses and blood levels required for response will be similar to those needed in younger patients.

In panic patients, lower starting doses are used, and dose is increased slowly to avoid exacerbation of panic attacks. In OCD, higher doses of clomipramine are used than those used in depression. In part, this is because OCD responds more slowly and less completely than depression and, thus, the dose is raised to ensure that the most effective treatment has been given.

If amitriptyline is used for pain, lower doses (e.g., 75 mg/day) are used. When effective, the action of the drug is observed within the first 2 weeks.

If doxepin is used for sleep, doses lower than those in depression are used, for example, 3 to 6 mg/day.

At this time, it is not clear when the tricyclic antidepressants would be used in children. It is recommended that desipramine should not be used in children younger than 12 years of age because of possible adverse cardiac effects, but the same logic applies to other tricyclics also. The tricyclic antidepressants have not been shown to be more effective than placebo in adolescents with depression, although this probably relates, in part, to diagnostic problems. The lethality of tricyclics in overdose should be kept in mind. Higher doses, on a milligram-per-kilogram basis, are likely to be required in younger patients than in adults to reach the same plasma level, but effective plasma level ranges have not been established in children.

Tricyclic and heterocyclic preparations are shown in Table 34.33–3.

References

*Agency for Health Care Policy and Research. *Clinical Practice Guideline: Depression in Primary Care: Treatment of Major Depression*. Vol 2. Rockville, MD: U.S. Government Printing Office; 1993.

Akamine Y, Yasui-Furukori N, Ieiri I, Uno T. Psychotropic drug-drug interactions involving P-glycoprotein. *CNS Drugs*. 2012;26:959–973.

Anderson IM. Selective serotonin reuptake inhibitors versus tricyclic antidepressants: a meta-analysis of efficacy and tolerability. *J Affect Disord*. 2000;58:19–36.

*APA Task Force. Tricyclic antidepressants–blood level measurements and clinical outcome: an APA Task Force report. Task force on the use of laboratory tests in psychiatry. *Am J Psychiatry*. 1985;142:155–162.

Chan CH, Janicak PG, Davis JM, Altman E, Andriukaitis S, Hedeker D. Response of psychotic and nonpsychotic depressed patients to tricyclic antidepressants. *J Clin Psychiatry*. 1987;48:197–200.

Charney DS, Delgado PL, Price LH, Heninger GR. The receptor sensitivity hypothesis of antidepressant action: a review of antidepressant effects on serotonin function. In: Brown SL, van Praag HM, eds. *The Role of Serotonin in Psychiatric Disorders*. New York:

Brunner/Mazel; 1991:29–56.

Danish University Antidepressant Group. Paroxetine: a selective serotonin reuptake inhibitor showing better tolerance, but weaker antidepressant effect than clomipramine in a controlled multicenter study. *J Affect Disord*. 1990;18:289–299.

*Davis JM, Glassman AH. Antidepressant drugs. In: Kaplan HI, Sadock BJ, eds. *Comprehensive Textbook of Psychiatry*. Baltimore, MD: Williams & Wilkins; 1989:1627–1655.

Drevets WC, Zarate CA Jr, Furey ML. Antidepressant effects of the muscarinic cholinergic receptor antagonist scopolamine: a review. *Biol Psychiatry*. 2013;73:1156–1163.

Duman RS, Heninger GR, Nestler EJ. A molecular and cellular theory of depression. *Arch Gen Psychiatry*. 1997;54:597–606.

Elkin I, Shea MT, Watkins JT, et al. National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. *Arch Gen Psychiatry*. 1989;46:971–982.

Evans WE, Relling MV. Pharmacogenomics: translating functional genomics into rational therapeutics. *Science*. 1999;286:487–491.

Farahani A, Correll CU. Are antipsychotics or antidepressants needed for psychotic depression? A systematic review and meta-analysis of trials comparing antidepressant or antipsychotic monotherapy with combination treatment. *J Clin Psychiatry*. 2012;73:486–496.

*Frank E, Kupfer DJ, Perel JM, et al. Three-year outcomes for maintenance therapies in recurrent depression. *Arch Gen Psychiatry*. 1990;47:1093–1099.

*Glassman AH, Bigger JT Jr. Cardiovascular effects of therapeutic doses of tricyclic antidepressants. A review. *Arch Gen Psychiatry*. 1981;38:815–820.

Glassman AH, Perel JM, Shostak M, Kantor SJ, Fleiss JL. Clinical implications of imipramine plasma levels for depressive illness. *Arch Gen Psychiatry*. 1977;34:197–204.

Kuhn R: The imipramine story. In: Ayd FJ, Blackwell B, eds. *Discoveries in Biological Psychiatry*. Philadelphia, PA: JB Lippincott Co; 1970:205–217.

Nelson JC, Jatlow PI, Quinlan DM. Subjective complaints during desipramine treatment. Relative importance of plasma drug concentrations and the severity of depression. *Arch Gen Psychiatry*. 1984;41:55–59.

Nelson JC, Mazure C, Jatlow PI. Antidepressant activity of 2-hydroxydesipramine. *Clin Pharmacol Ther*. 1988;44:283–288.

O'Malley PG, Jackson JL, Santoro J, Tomkins G, Balden E, Kroenke K. Antidepressant therapy for unexplained symptoms and symptom syndromes. *J Fam Pract*. 1999;48:980–990.

Potter WZ, Calil HM, Sutfin TA, et al. Active metabolites of imipramine and desipramine in man. *Clin Pharmacol Ther*. 1982;31:393–401.

Preskorn SH, Fast GA. Therapeutic drug monitoring for antidepressants: efficacy, safety, and cost effectiveness. *J Clin Psychiatry*. 1991;52 Suppl:23–33.

Roose SP, Laghrissi-Thode F, Kennedy JS, et al. Comparison of paroxetine and nortriptyline in depressed patients with ischemic heart disease. *JAMA*. 1998;279:287–291.

Rosenstein DL, Nelson JC, Jacobs JC. Seizures associated with antidepressants: a review. *J Clin Psychiatry*. 1993;54:289–299.

Ryan ND. The pharmacologic treatment of child and adolescent depression. *Psychiatr Clin North Am*. 1992;15:29–40.

Spiker DG, Weiss JC, Dealy RS, et al. The pharmacologic treatment of delusional depression. *Am J Psychiatry*. 1985;142:430–436.

Tremblay P, Blier P. Catecholaminergic strategies for the treatment of major depression. *Curr Drug Targets*. 2006;7:149–158.

White N, Litovitz T, Clancy C. Suicidal antidepressant overdoses: a comparative analysis by antidepressant type. *J Med Toxicol*. 2008;4:238–250.

Yan L, Wang Q, Fu Q, Ye Q, Xiao H, Wan Q. Amitriptyline inhibits currents and decreases the mRNA expression of voltage-gated sodium channels in cultured rat cortical neurons. *Brain Res*. 2010;1336:1–9.

Yoshimura M, Furue H. Mechanisms for the anti-nociceptive actions of the descending noradrenergic and serotonergic systems in the spinal cord. *J Pharmacol Sci*. 2006;101:107–117.

▲ 34.34 Valproate

ROBERT M. POST, M.D., AND MARK A. FRYE, M.D.

Valproic acid (Depakene), and its formulation derivatives divalproex sodium (Depakote), extended-release (ER) divalproex sodium (Depakote ER), and valproate sodium injection (Depacon), represent one of the first-generation mood-stabilizing anticonvulsants. Valproate is currently approved by the U.S. Food and Drug Administration (FDA) for (1) monotherapy or adjunctive therapy of complex partial seizures that occur in isolation or in conjunction with other types of seizures, (2) monotherapy and adjunctive therapy of simple and complex absence seizures, and (3) adjunctive therapy for patients with multiple seizures that include absence seizures. Divalproex has an FDA indication in these same epilepsy syndromes but additionally has indications for (4) prophylaxis of migraine and (5) treatment of acute manic episodes associated with bipolar disorder. Intravenous (IV) valproate sodium is currently FDA approved as an alternative in patients for whom oral administration is temporarily not feasible (i.e., unresponsive or otherwise on nothing-by-mouth [NPO] status). ER divalproex sodium is currently FDA approved for migraine prophylaxis and monotherapy and adjunctive therapy in complex partial seizures and acute mania.

HISTORY

The first clinical observations of valproate's anticonvulsant activity were in Europe in the 1970s. Shortly thereafter, the first reported evaluation of valpromide (the prodrug of valproate) in bipolar disorder was that of O. Lambert in France. Since that time, valproate has been extensively evaluated in bipolar disorder (particularly the manic phase as exemplified with the 1995 indication for acute mania) and, to a much lesser extent, in

other psychiatric conditions, such as schizophrenia, schizoaffective disorder, bipolar depression, panic disorder, posttraumatic stress disorder (PTSD), alcohol withdrawal, and target symptoms of impulsivity, agitation, and mood lability.

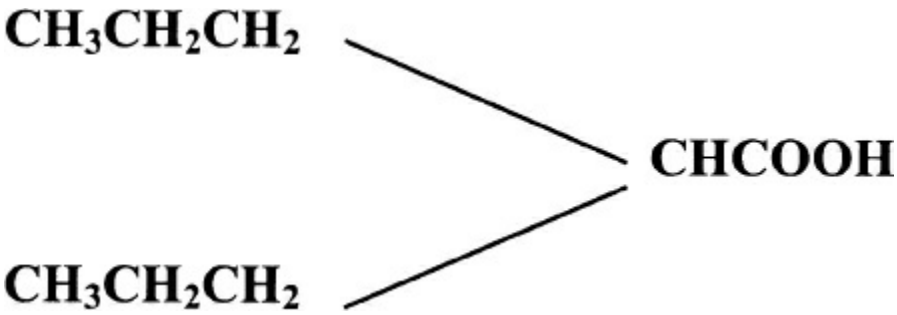

FIGURE 34.34–1. Molecular structure of valproate.

CHEMISTRY

Valproic acid (*N*-dipropylacetic acid) is a simple-chain branched-chain carboxylic acid (Fig. 34.34–1).

The commercial formulation derivatives available include valproic acid, which is available in syrup and capsules; valpromide (amide prodrug of valproic acid used in Europe); valproate sodium injection; divalproex sodium, which is an enteric-coated delayed-release complex of valproic acid and sodium valproate in a 1 to 1 ratio that is available in tablet and sprinkle formulation (can be opened and spread on food); and an ER formulation of divalproex sodium.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics

All formulations of valproate are rapidly absorbed, with peak plasma concentrations obtained 3 to 4 hours after oral administration. The plasma half-life of valproate is 10 to 16 hours; when dosed in combination with drugs that induce cytochrome P450 enzymes (i.e., carbamazepine [Tegretol]), the half-life is shorter.

Table 34.34–1.
Common and Differential Effects of Valproate (VPA) Compared to Carbamazepine (CBZ) or Lithium (Li)

	CBZ	VPA	Li
Common effects			
GABA turnover	↓	↓	↓
$GABA_B$ receptors in hippocampus (chronic)	↑	↑	↑
GABA levels	(↑)	↑	(↑)
↓ GSK-3B	+	+	+

Brain tryptophan and 5-HT	↑	↑	↑
Dopamine turnover	↓	↓	↓
NMDA-mediated Ca++ currents	↓	↓	↓
Arachidonic acid turnover	↓	↓	↓
Myoinositol transport (SMIT)	↓	↓	↓
Stroke neuroprotection	+	+	+
BDNF	↑	↑	↑
Lithium-shared effects			
Inhibition of ERK MAP kinase path	−	+	+
↓ PKC	−	+	+
↓ MARCKS levels	−	+	+
↑ Bcl-2	−	+	+
↑ Neurogenesis	−	+	+
↓ β-Catenin	−	+	+
↑ AP-1 Binding	−	+	+
Carbamazepine-shared effects			
Na_i+ block	+	(+)	−
K efflux	↑	↑	−
[^{3}H] picrotoxinin binding	↓	↓	−
Release of aspartate, glutamate	↓	↓	−
Somatostatin levels	↓	↓	−
Histone, deacetylase inhibition	+	+	−
Block completed kindled seizures	+	+	−
Unique effects of each drug			
Inositol Monophosphatases	−	↓	↑
Absence seizures	↑	↓	−
Blocks kindling development	−	+	−
Inhibit after discharge spread in hippocampus in cortex	−	+	−
	+	−	−
Potential anticancer effects (AML)	−	+	−
Forms ester with coenzyme A	−	+	−
↑ PPARs	−	+	−

GABA, γ-aminobutyric acid; GSK, glycogen synthase kinase; 5-HT, serotonin; NMDA, *N*-methyl-D-aspartate; SMIT, *myo*-inositol transport system; BDNF, brain-derived neurotrophic factor; ERK, extracellular-signal-regulated kinases; PKC, protein kinase C; MARCKS, myristoylated alanine-rich C kinase substrate; AP, activator protein; AML, acute myelogenous leukemia; PPARs, peroxisome proliferator-activated receptors.

The ER formulation of divalproex sodium was designed for once-daily dosing, thereby maintaining lower maximum drug concentration (C_{max}) and higher minimum drug concentration (C_{min}) between dosing intervals. This would appear to be useful clinically, as treatment adherence is linked to the simplicity of treatment regimen, and some side effects correlate with peak drug levels. The ER formulation has 80 percent bioavailability of direct release (DR) divalproex sodium. A small (N = 35), healthy control, pharmacokinetic comparison study of two ER (given every day) versus DR

(given twice a day) dosages reported that there was no significant difference in area under the curve (AUC) when comparing 1,000 mg of ER to 875 mg of DR; however, there was a greater AUC in the dosage of 1,500 mg of ER versus 1,250 mg of DR. The ER formulation obtained a lower C_{max} and higher C_{min} than the DR formulation with less peak-trough fluctuations. Further study is encouraged in epileptic or bipolar patient populations at higher dosages, but increasing the DR dose by 20 percent to an approximately equivalent ER dose with clinical monitoring would be a reasonable switch strategy.

Valproate is highly protein bound, primarily to albumin. It is generally thought that, at serum levels less than 45 to 50 μg/mL, these binding sites are unsaturated; at higher serum levels, these binding sites become saturated, with the unbound valproate level increasing. The unbound portion of valproate is considered to be pharmacologically active and can cross the blood–brain barrier. When used in combination with other psychotropic agents, valproate has a higher degree of plasma protein binding such that it can displace these drugs and increase their free levels (i.e., lamotrigine [Lamictal] and carbamazepine).

Valproate is metabolized primarily by hepatic glucuronidation and mitochondrial β-oxidation. Several of valproate's metabolites, 2-propyl-2-pentenoic acid and 2-propyl-3-oxopentanoic acid, have potent anticonvulsant activity.

Pharmacodynamics

Valproate's anticonvulsant and mood-stabilization mechanisms of action have not been specifically identified and may not necessarily be the same. Its anticonvulsant effects are rapid in onset, while antimanic and antidepressant effects are slower in onset and appear to require chronic administration. Several hypotheses include enhancement of brain γ-aminobutyric acid (GABA) levels via multiple actions of synthesis and degradation, and modulation of other neurotransmitters, voltage-sensitive Na^+ channels, extra hypothalamic neuropeptides, secondary messenger systems, and neuroprotection. Comparative effects with two other antimanic mood stabilizers, carbamazepine and lithium (Eskalith), are listed in Table 34.34–1.

The GABA deficit hypothesis, as originally proposed by Emrich in 1980, suggested that unipolar and bipolar illnesses were characterized by a relative deficiency of GABA (the major inhibitory neurotransmitter in cortex), and that this deficit could be ameliorated with valproate treatment. Preclinical studies have shown that valproate enhances GABAergic function in specific regions of brain by a variety of presynaptic or postsynaptic mechanisms, or both, such as enhancement of GABA synthesis through activation of glutamate decarboxylase, prevention of GABA catabolism

through inhibition of GABA transaminase, direct release of GABA in response to calcium influx, reversal of the GABA transporter, and postsynaptic GABA type A ($GABA_A$) receptor activation effects. However, whether GABA enhancement is directly related to the antimanic effects of valproate remains controversial, and indirect inferences from the effects of other potent anticonvulsants that enhance brain GABA further suggest that this is unlikely. For example, other potent GABAergic drugs, tiagabine (Gabitril Filmtabs), gabapentin (Neurontin), and topiramate (Topamax), are not effective antimanic treatments.

In addition, GABA deficits, as measured in cerebrospinal fluid (CSF), by magnetic resonance spectroscopy (MRS), or plasma analysis, have not been consistently reported in clinical studies of affectively ill patients. GABA assessments after drug treatment have also produced inconsistent results. For example, lithium treatment normalized a relative deficit of CSF GABA in 23 euthymic bipolar patients. By CSF analysis, valproate treatment increased GABA in 31 pediatric epileptic patients and six schizophrenic patients; by MRS analysis, valproate increased occipital GABA in six euthymic bipolar patients. However, no increase in GABA, as measured by CSF analysis in six schizophrenic patients or by MRS analysis in five refractory epileptic patients treated with valproate has also been reported. Plasma GABA levels have been reported to be increased after treatment with valproate in three studies (two studies of normal controls and one study of pediatric epilepsy) and decreased in one study of acute mania. In this latter largest study to date, higher baseline pretreatment GABA levels in those patients randomized to valproate (N = 19) was associated with better clinical response; this relationship was not observed in patients randomized to lithium (N = 13) or placebo (N = 31).

The discordance between the previously mentioned CSF, MRS, and plasma studies remains unresolved; different study populations, collection and storage times of samples, lengths of treatment, and methodologies of GABA assay need to be standardized to better elucidate the role of GABA in bipolar disorder and the role for GABAergic mechanisms in the therapeutic effects of valproate in the various neuropsychiatric syndromes for which it is indicated.

Valproate's effect on less well-studied neurotransmitters and neuropeptides has included glutamate (increased release in mouse brain), serotonin (increased extracellular level in rat hippocampus and striatum), dopamine (increased extracellular level in rat), somatostatin (decreased CSF level), and corticotropin-releasing factor (CRF) (decreased CRF in median eminence, raphe nuclei, and decreased CRF messenger ribonucleic acid [mRNA] in the central nucleus of amygdala and paraventricular nucleus of hypothalamus). Although these neurotransmitters and neuropeptides have all been implicated in the pathophysiology of mania, depression, psychosis, or a combination of these, it is not known whether

valproate's modulation of these neurotransmitters and neuropeptides contributes to its mood-stabilizing properties.

Intracellular signaling systems similar to those of lithium have also been observed with valproate. Like lithium, valproate stimulates glutamate release and, through activation of the *N*-methyl-D-aspartate (NMDA) receptor, has been shown to increase inositol-1,4,5-triphosphate. Additionally, lithium, valproate, and carbamazepine, but not amitriptyline (Elavil), have been shown to inhibit the high affinity *myo*-inositol transport system in astrocytes; this inhibition of inositol and subsequent reduced intraneuronal level of inositol (*inositol depletion*) may be related to the mood-stabilization mechanism of valproate. However, although lithium decreases inositol monophosphatase activity, valproate increases it. Both inhibit the *myo*-inositol transport system (SMIT) as does carbamazepine.

Similar actions have also been observed with lithium and valproate in downregulation of protein kinase C (PKC) activity, downregulation of myristylated alanine-rich C kinase substrate (MARCKS) activity (a major substrate of PKC), and increased level of Bcl-2 (a neuroprotective protein). Further study is required to link any of these common actions of two drugs with different chemical structures but similar mood-stabilizing properties. PKC inhibition is a likely contributor to antimanic efficacy. In one open and two double-blind, placebo-controlled studies, the potent PKC inhibitor tamoxifen (Nolvadex), which also has antiestrogen receptor effects, has proven antimanic efficacy.

Like lithium, valproate increases Bcl-2, brain-derived neurotrophic factor (BDNF), and neurogenesis, but the increases in cortical gray matter seen with lithium have not been consistently observed with valproate. However, patients on valproate had increased volumes of their cingulate cortices and increased *N*-acetyl-aspartate (NAA) to creatine (Cre) or choline (Cho) ratios in hippocampus compared with those not treated with valproate. Since BDNF decreases during episodes of depression and mania, valproate's ability to increase BDNF could be important to its clinical actions.

Valproate is also positive in animal models of stroke and surprisingly is effective not only when given before middle cerebral artery ligation, but (like lithium) when given shortly after as well. In contrast to lithium, valproate inhibits histone deacetylase which could account for many of its effects on gene transcription. It is also possible that this effect contributes to valproate's teratogenicity (as described below) or to its possible therapeutic actions in conjunction with hydralazine (an inhibitor of DNA methylation) in reversing treatment resistance to chemotherapy of malignancies.

CLINICAL DRUG STUDIES

Acute Mania

The early uncontrolled studies of valproate (divalproex or valpromide) in acute mania suggested efficacy. Subsequent controlled studies of valproate in acute mania confirmed this view; these studies are reviewed in Table 34.34–2 and have included one single-site, placebo-controlled study; one multisite, placebo-controlled, parallel design study with lithium as an active comparison; one placebo-controlled adjunctive study to typical neuroleptic therapy; and six controlled, active comparison studies to lithium (two studies), haloperidol (Haldol) (one study), carbamazepine (one study), and olanzapine (Zyprexa) (two studies).

The first placebo-controlled studies by Harrison G. Pope and colleagues (1991) and Charles L. Bowden et al. (1994) were pivotal in receiving an FDA indication for acute mania. Post hoc analyses of the Bowden study have suggested that, although there was no difference in the response rate for patients with euphoric mania, patients with depressive symptoms (i.e., dysphoric mania) or a cumulative lifetime episode burden of greater than four prior depressive episodes or 11 prior manic episodes responded poorly to lithium. Furthermore, four distinct manic subtypes (anxious-depressed, psychotic, classic, and irritable) based on pretreatment symptom cluster analyses confirmed that the classic manic subtype responded similarly to divalproex and lithium, but the irritable manic subtype responded significantly better to divalproex than lithium ($P < 0.004$) or placebo ($P < 0.04$). The three treatments did not differ in response rates to depressive or psychotic subtypes.

These two studies have suggested that a serum valproate level of greater than 45 to 50 μg/mL is associated with an antimanic response. In the first study, the majority of patients randomized to valproate responded 1 to 4 days after achieving a serum level greater than or equal to 50 μg/mL. In the second study, patients who had achieved a serum level of greater than or equal to 45 μg/mL at day 5 were more likely to have a 20 percent or greater reduction in the manic syndrome and behavior–ideation subscales than those patients who had a serum level of less than 45 μg/mL. One study suggested a positive correlation between serum valproate level at day 3 and degree of manic improvement as measured by the Young Mania Rating Scale (YMRS). Finally, higher baseline plasma GABA levels may predict treatment response to valproate in acute mania.

A multicenter study of the divalproex ER preparation in acute mania was again highly positive compared with placebo for both mania and mixed patients and those with and without psychosis. Starting dose and rate of titration may have been too high because the side-effect profile was not diminished compared with the acute release studies. Serum levels of 71.3 to 85 μg/mL or higher appeared to be the therapeutic target, since those under this concentration show fewer good responses (smaller effect sizes

[0.4 to 0.8]) compared with those at higher blood levels (effect sizes 0.9 to 1.1). There was a hint of a positive blood level–response relationship with those above 95 μg/mL, showing effect sizes of 1.1 or the equivalent of –12 YMRS points, but it was not statistically dependable.

There is some evidence of a slightly lower bioavailability (AUC) with the ER preparation, which leads some to suggest that with a transition from the immediate release (IR) to ER preparation, slightly (20 percent) higher doses are required. The ER preparation was dosed on a once-daily basis in the morning, but clinically, one would recommend single bedtime dosing to enhance sleep and reduce side effects during awake time and increase compliance.

Because of valproate's robust effects in acute mania in adults, it was surprising that it failed to separate from placebo in a study of childhood- and adolescent-onset mania. Concomitant stimulant medication was allowed in that study, which may have confounded the results.

Given that there are many FDA-approved conventional mood-stabilizing agents for the treatment of acute mania (lithium, valproate, long-acting carbamazepine [Equetro], olanzapine, risperidone [Risperdal], quetiapine [Seroquel], aripiprazole [Abilify], and ziprasidone [Geodon]) and the extensive use of typical antipsychotics in acute mania (chlorpromazine [Thorazine] was FDA approved in 1970), controlled comparison studies provide important preliminary information regarding selection of an antimanic agent.

Although the typical antipsychotic agents have been extensively used for the treatment of acute mania, they represent a somewhat outmoded treatment because they are associated with more postmanic depressive symptom severity and breakthrough episodes of depression (vs. lithium) and their side-effect profile. Typical antipsychotics are associated with considerable side-effect liabilities, including neuroleptic-induced dysphoria, acute dystonia, extrapyramidal side effects, and, most importantly, long-term risk of tardive dyskinesia in some 20 to 40 percent of bipolar patients (even those treated intermittently). In one comparative study in psychotic mania, rapid loading of divalproex (20 mg/kg/day) was as effective as haloperidol and was associated with less need for adjunctive benztropine (Cogentin). In a larger, controlled, 3-week, double-blind study of 136 acutely manic patients treated with haloperidol or perazine (Taxilan), adjunctive valproate (N = 69, with a mean daily dose of 19.6 mg/kg and a mean serum level of 80.3 μg/mL) allowed for a 43 percent reduction of typical antipsychotic dosage compared to adjunctive placebo (N = 67) for which only a 14 percent reduction in neuroleptic dose was possible.

Table 34.34–2.

Controlled Studies of Valproate or Divalproex in Acute Mania

Study	Design (N)	Primary Outcome Variable	Results	Comment
Pope et al., 1991	3-week, randomized, db, pc, DSM-III-R mania, unresponsive to Li (*N* = 36)	% ↓ YMRS	VAL 54% ↓ vs. PLC 5% ↓ (*P* = 0.003), ITT (*P* = 0.013)	Response 1–4 days after achieving serum VAL >50 µg/mL
Bowden et al., 1994	3-week, multisite, randomized, db, pc, RDC confirmed mania (*N* = 179)	50% ↓ MRS	Mean dose/level DVPX: 2,000 mg/93.2 µg/mL) Li: 1,950 mg/1.2 mmol/L	All rapid-cycling patients (*N* = 8) randomized to DVPX; therefore, no comparative analysis possible
	DVPX vs. Li vs. PLC in 2:1:2 ratio. Dosage titrated to DVPX (150 µg/mL) and Li (1.5 mmol/L), except if adverse event was present		ITT: DVPX, 46% (*P* = 0.004 vs. PLC); Li, 49% (*P* = 0.025 vs. PLC); PLC 25%	48% of current Li responders reported past Li inefficacy or intolerance
Freeman et al., 1992	3-week, randomized, db, DSM-III-R mania (*N* = 27)	50% ↓ SADS-C mania	Li levels (0.8–1.4 mmol/L), mean DVPX level (98 ± 17 µg/mL)	Pretreatment depression significantly higher in DVPX responders vs. nonresponders (*P* <0.05)
	DVPX (titrated to 3,000 mg—lower dose for response or side effect) vs. Li (titrated to response, side effect, 1,800 mg q.d., or level 1.5 mmol/L)		Li 12/13 (92%) vs. DVPX 9/14 (64%)	Posttreatment mania scores higher in mixed vs. nonmixed manics treated with Li (*P* <0.05)
McElroy et al., 1996	1-week, randomized, sb, psychotic mania, rapid load DVPX (*N* = 21, 20 mg/kg/day) vs. HAL (*N* = 15, 0.2 mg/kg/day)	Mean change YMRS	DVPX 43% YMRS ↓; HAL 35% YMRS ↓	Equal efficacy Greater use of benztropine in HAL group
Hirschfeld et al., 1999	10-day, multisite, randomized, db, DSM-IV acute mania (YMRS ≥14) DVPX load (*N* = 20, 30 mg/kg/day on days 1 to 2, 20 mg/kg/day on days 3 to 10) vs. DVPX nonload (*N* =	Safety and tolerability DVPX load	84% DVPX load had level >50 µg/mL by day 3 vs. 30% DVPX	No difference in adverse event or YMRS change

	20) vs. Li (*N* = 19)		nonload	
Muller-Oerlinghausen et al., 2000	3-week, multisite, randomized, db, pc adjunct VAL (20 mg/kg/day) vs. PLC to neuroleptic (*N* = 136)	Mean dose ↓ neuroleptic at study end	Day 21 VAL level (80.3 ± 22.3 μg/mL)	Decreased neuroleptic burden in VAL patients
	Step-wise dose reduction of neuroleptic		43% ↓ of neuroleptic in adjunct VAL vs. 14% ↓ in adjunct PLC (*P* = 0.0007)	
Vasudev et al., 2000	4-week, randomized, db, CBZ vs. VAL (*N* = 30)	Mean change in YMRS	VAL (75.4 μg/mL) greater YMRS ↓ vs. CBZ (6.2 μg/mL) (*P* = 0.023)	Correlation between VAL level and ↓ YMRS at week 2 (Fig. 31.8e–2)
Tohen et al., 2002	3-week, multisite, randomized, db, parallel DSM-IV mania (YMRS ≥20) (*N* = 251)	Mean change YMRS	DVPX (mean dose 1,401 mg q.d., serum level 79.4 μg/mL), OLZ (mean dose 17.4 mg)	Weight gain 2.5 kg OLZ vs. 0.9 kg DVPX
	DVPX (*N* = 125, flexible dose of 500–2,500 mg q.d.) vs. OLZ (*N* = 123, 5–20 mg q.d.)	Response (50% ↓ YMRS)		Somnolence, dry mouth, ↑ appetite, tremor, neck rigidity more common with OLZ
		Remission (YMRS ≤12)	YMRS ↓ DVPX 10.4 points vs. OLZ 13.4 points (*P* = N.S.) Response and remission: DVPX 42.3% and 34.1% vs. OLZ 54.4% and 47.2% (*P* = 0.058; *P* = 0.039)	Nausea, diarrhea more common with DVPX
Zajecka et al.,	3-week, multisite, randomized,	Mean MRS	DVPX, mean	Weight gain,

2002	db, parallel DSM-IV mania (MRS ≥25) (*N* = 120)	change	dose of 2,115 mg; OLZ mean dose of 15 mg	somnolence, rhinitis, edema, ↑ cholesterol more common with OLZ
	DVPX (*N* = 43; loading 20 mg/kg/day) vs. OLZ (*N* = 57; start dose 10 mg with ↑ for persistent symptoms)	Response (50% ↓ MRS)	MRS ↓ DVPX 14.8 points vs. OLZ 17.2 points (*P* = N.S.) ITT: Response (DVPX 53%, OLZ 62%, *P* = N.S.)	↑ Monocytes, ↑ eosinophils, ↓ neutrophils, ↓ platelets more common with DVPX
Bowden et al., 2006	3-week, multisite, randomized db, pc, DSM-IV acute mania DVPX started at 25 mg/kg day 1 ↑ 500 mg on day 3 → 85–125 µg/mL (Dosing once daily in AM)	Response (50% ↓ YMRS)	ITT: Response DVPX 48% PLC 34% (*P* = 0.02) Mean final dose 3,057 mg per day (35.4 mg/kg) → mean serum level 95.9 µg/mL	On DVPX: somnolence, dizziness, complaints; weight gain: 1.8 kg DVPX 0.3 kg PLC Total cholesterol −13.5 mg/dL DVPX +2.5 mg/dL PLC

CBZ, carbamazepine (Tegretol); db, double blind; DVPX, divalproex sodium (Depakote); HAL, haloperidol (Haldol); HAMD, Hamilton Depression Scale; ITT, intent-to-treat analysis; Li, lithium (Eskalith); MRS, Mania Rating Scale; N.S., not significant; OLZ, olanzapine (Zyprexa); pc, placebo controlled; PLC, placebo; RDC, Research Diagnostic Criteria; SADS-C, Schedule for Affective Disorders and Schizophrenia-Change; sb, single blind; VAL, valproate; YMRS, Young Mania Rating Scale.

As reviewed previously, the post hoc analyses of comparison studies of lithium versus divalproex would suggest that depressive or irritable symptoms in mania are more effectively treated with divalproex. Other possible correlates of a relatively positive response are listed in Table 34.34–3. There is one large open trial in rapid-cycling patients of valproate prophylaxis as monotherapy in lithium-refractory patients (*N* = 43) or as add-on therapy (*N* = 101) with a greater than 70 percent antimanic or nonmixed response. The only comparison study of valproate and carbamazepine suggested greater efficacy and dose-dependent antimanic response with valproate.

Table 34.34–3.
Potential Correlates of Response to Valproate

	Strength of

	Association
Diagnosis-related factors	
Diagnosis of bipolar disorder [vs. schizoaffective or schizophrenia disorder]	+++
Dysphoric [or mixed mania]	+++
Bipolar II disorder	++
Clinical course variables	
Presence of rapid cycling	+
Decreasing or stable episode frequency	++
Later age of onset of illness (but see Fig. 34.34–2A,B)	+
Shorter duration of illness	+
Long and severe course of illness (see Figs. 34.34–2 and 34.34–3)	++
Comorbid psychiatric diagnosis	
Comorbid alcohol abuse	+++
Comorbid mental retardation	+
Comorbid anxiety and panic attacks	+++
Comorbid PTSD	+
Treated hypothyroidism	+
Marked sleep disturbance	+
Explosive dyscontrol and aggression	++
Associated neurological factors	
Presence of nonparoxysmal EEG abnormalities (not subsequently replicated)	±
Organic affective disorder	±
History of closed head trauma	+
Comorbid migraine	+++
Family history	
Positive family history of mood disorder	+
Treatment-related variables:	
Patients who are lithium naïve (especially BP-II)	+
Higher serum valproate concentrations at days 3–6 of treatment	
Higher serum valproate concentrations with more severe illness: cyclothymia <BP-II<BP-I	++
Brain imaging	
Low activity in the frontal and anterior cingulate and medial frontal gyrus	+
Plasma GABA	
Not low (near normal)	+

PTSD, posttraumatic stress disorder; EEG, electroencephalogram; BP, bipolar; GABA, γ-aminobutyric acid; +++, well documented in several open studies or one controlled study; ++, good support in literature; +, fair support; ±, weak support or equivocal.

Modified from West SA, Strakowski SM, Sax KW, McElroy SL, Keck PE Jr, McConville BJ. Phenomenology and comorbidity of adolescents hospitalized for the treatment of acute mania. *Biol Psychiatry*. 1996;39(6):458–460 and Post RM, Speer AM, Obrocea GV, Leverich GS. Acute and prophylactic effects of anticonvulsants in bipolar depression. *Clin Neuroscience Res*. 2002;2(3–4):228–251.

Two comparative studies of divalproex versus olanzapine were similar in design, with the exception of dose and titration. In one smaller study

that started with a rapid loading dose of divalproex (mean dose of 2,115 mg vs. olanzapine mean dose of 15 mg), the mean reduction in manic symptoms and response rates was not statistically different. In the larger-powered study that started with a nonloading and lower divalproex dose (mean dose of 1,401 mg vs. olanzapine mean dose of 17.4 mg), response and remission rates were significantly in favor of olanzapine. A 12-week comparative study in acute mania suggested the superiority of quetiapine to valproate for efficacy, but with a higher side-effect burden.

Acute Bipolar Depression

Conventional mood stabilizers have been less well studied in the depressive phase of bipolar disorder. Several open trials have clearly suggested antidepressant properties with valproate, generally at lower doses than studies in acute mania. In a 12-week open trial of 19 patients with bipolar II depression, divalproex sodium monotherapy (with a mean dose of 882 mg every night and a mean level of 80.7 μg/mL), 63 percent of the cohort responded to treatment as defined by a 50 percent reduction in the Hamilton Rating Scale for Depression.

As presented in Table 34.34–4, a controlled study of divalproex monotherapy in acute bipolar I or II depression did not statistically separate from placebo at the 8-week study end point. The small sample size (N = 45) does not allow definitive conclusions about acute antidepressant efficacy of divalproex. However, in another small randomized-controlled trial including 25 patients, valproate not only showed superior antidepressant effects compared with placebo, but showed even more clinically and statistically robust antianxiety effects in monotherapy of bipolar depression. A more recent double-blind placebo-controlled study showed marked superiority of valproate over placebo with a moderately large effect size of 0.81.

Current guidelines emphasize the importance of maximizing mood stabilizer or atypical antipsychotic treatment and revision before the commencement of unimodal antidepressant therapy. Consistent with this recommendation, one small retrospective review of bipolar patients requiring antidepressant augmentation found lower mean levels of valproate (54 μg/mL ± 24 μg/mL at the time of antidepressant intervention) compared with a diagnostically matched comparison group not needing antidepressants with valproate (73 μg/mL ± 13 μg/mL).

Prophylactic Mood Stabilization

There are six controlled studies that have evaluated the efficacy of valproate or divalproex in long-term mood stabilization (Table 34.34–5). These studies include one open randomization comparing valpromide versus lithium monotherapy; one open randomization comparing valproate

versus lithium monotherapy in a child, adolescent, and young adult population; one placebo-controlled, adjunctive divalproex trial to lithium-maintained bipolar I patients; one 52-week, parallel design comparing lithium, divalproex, and placebo monotherapy in recently recovered manic patients; and one controlled comparison study between divalproex and olanzapine monotherapy in those patients who had presented in an acute manic episode.

Several conclusions can be drawn regarding these studies. First, one study suggested better prophylaxis against manic episodes versus depressive episodes. Conversely, the largest controlled study to date did not find any difference between divalproex, lithium, or placebo in time to recurrence of a mood episode or, specifically, a manic episode. However, almost 200 patients, who were initially enrolled as acutely manic, were unable to fully recover and thus were not randomized into the prophylactic phase; this methodological limitation may have yielded a type II error in a cohort of only randomized mildly ill patients, thereby making it more difficult to show differences between active drugs and the low rate of manic relapse even on placebo. A secondary outcome measure, however, suggested greater depression prophylaxis with divalproex than placebo in the monotherapy analysis and in a subgroup of patients who were subsequently treated with adjunctive antidepressant therapy (i.e., sertraline [Zoloft] or paroxetine [Paxil]) added to divalproex versus placebo.

Table 34.34–4.
Controlled Studies of Valproate and Divalproex Sodium in Bipolar Disorder (Acute Depression)

Study	Design	Primary Outcome Variable	Results	Comment
Sachs et al., 2002	8-week, multisite, randomized, db, pc, DSM-IV outpatient bipolar depression (*N* = 45:25 bipolar I disorder, 18 bipolar II disorder, 2 bipolar disorder not otherwise specified) DVPX (250 mg PO q.h.s with 250 mg q.o.d. ↑ based on efficacy and tolerability)	50% improvement in HAMD and YMRS <10	Mean dose, 1,391 mg; mean serum level, 61.5 µg/mL ITT:DVPX response 9 of 21 (43%) vs. PLC 6 of 22 (27%) (*P* <0.4)	DVPX HAMD improvement was numerically superior to PLC at every measured time point of study
Davis et al., 2005	db, pc, DSM-IV output bipolar depression (*N* = 25) 500–2,500 mg/day		ITT: DVPX response 6 of 13 (46%) vs. PLC 3 of 12 (25%); >50% improvement	Moderately large effect size = .81
Ghaemi et al., 2007	db, pc, DSM-IV outpatient nonrefractory bipolar depression (*N* = 18); target dose to blood			

	level 70–90 ng/dL	on HDRS and score <9 MADRAS decreased 13.6 points on VPA and 1.4 points on PLC	
Muzina et al., 2008	Db, pc, 6 weeks *N* = 26 VPA *N* = 28 Placebo	MADRAS	Effect size = 0.68

DSM, Diagnostic and Statistical Manual of Mental Disorders; db, double blind; ran, randomized; DVPX, divalproex sodium; HAMD, Hamilton Depression Scale; YMRS, Young Mania Rating Scale; ITT, intent-to-treat analysis; PLC, placebo; PC, placebo controlled.

Table 34.34–5. Controlled Long-Term Evaluation of Divalproex Sodium in Bipolar Disorder

Study	Design (N)	Primary Outcome Variable	Results	Comment
Lambert et al., 1992	18-month, ran, open VAL 1,200 mg (*N* = 78) vs. Li (*N* = 72) in bipolar (*N* = 121) and unipolar (*N* = 29) patients Crossover for patients who deteriorated	Episodes per patient	VAL (0.51) vs. Li (0.61) Equally effective	Intolerance and deterioration VAL (5.1%) vs. Li (13.9%) ↑ Rate of response for those who crossed over to VAL (90%) vs. Li (75%)
Solomon et al., 1997	52-week, ran, db, pc, lithium-maintained, DSM-IV bipolar I disorder outpatients adjunctive DVPX (50–125 mg/day) vs. PLC (*N* = 12)	Relapse to any mood episode	0 of 5 (0%) adjunct DVPX vs. 5 of 7 (71%) adjunct PLC (χ^2 = 5.61, *P* = 0.014)	Adverse events DVPX + Li (5 of 5, 100%) vs. PLC + Li (3 of 7, 42%) (χ^2 = 3.93, *P* = 0.04)
Bowden et al., 2000	52-week, db, pc, maintenance trial of euthymic outpatients recovered from index mania within 3 months (*N* = 372) Ran to DVPX (day 30, mean level of 84.8 ± 29.9 µg/mL), Li	Primary outcome measure: time to recurrence of any mood episode Secondary outcome measure:	50% survival median time: DVPX (40 weeks), Li (24 weeks), PLC (28 weeks) (DVPX vs. Li χ^2 = 3.11, *P* = 0.06)	DVPX: ↑ sedation, infection, tinnitus (all vs. Li); ↓ WBC, platelets (all vs. PLC) Li: ↑ polyuria (vs. DVPX),

	(day 30, mean level of 1.0 ± 0.48 mmoL/L), or PLC at 2:1:1 ratio	time to manic episode, time to depressive episode		↑ WBC, uric acid, creatine, calcium (all vs. PLC) 199 patients failed to achieve recovery and were not ran
Tohen et al., 2002	47-week, multisite, ran, db, flexibly dosed DVPX (500–2,500 mg q.d.) vs. OLZ (5–20 mg q.d.) for DSM-IV confirmed acute mania (YMRS ≥20) (*N* = 251)	%/time remission (YMRS ≤12)	Median time to remission (DVPX: 62 days vs. OLZ: 14 days, *P* = 0.047);% remission (DVPX 45.5% vs. OLZ 56.8%, *P* = 0.098);% subsequent relapse (DVPX 56.5% vs. OLZ 47.4%, *P* = 0.416)	OLZ: ↑ sedation, ↑ appetite, weight gain, ↑ cholesterol, dry mouth, akathisia, liver function test abnormality DVPX: nausea, nervousness, rectal disorder, ↓ platelet count, ↓ albumin
Kafantaris et al., 1997	2-year, open, ran VAL vs. Li, 11–21 years of age (*N* = 26)	Relapse rate	↑ Rate relapse Li >VAL	↑ Compliance and ↓ adverse events with VAL vs. Li
Frankenburg and Zananini, 2002	6-month, ran, db, pc DSM-IV bipolar II disorder (currently euthymic) and borderline personality disorder (*N* = 30) DVPX (*N* = 20, titrated to 50–100 mg/L) vs. PLC (*N* = 10) × 6 months	Mean change SCL-90 scales of interpersonal sensitivity, anger and hostility, and depression and OAS-M	Mean dose DVPX (850 ± 249 mg/day) Greater ↓ in SCL-90 sensitivity, anger and hostility, and OAS-M with DVPX vs. PLC (*P* = 0.04, 0.01, 0.02)	
Geddes et al., 2010	Ran, open 24 months		54% of 110 an event on combo 59% of 110 an event on Lithium 69% of 110 an event on valproate	Li + VPA superior to VPA alone

db, double blind; DVPX, divalproex sodium (Depakote); ITT, intent-to-treat analysis; Li, lithium (Eskalith); OAS-M, Overt Aggression Scale, modified; OLZ, olanzapine (Zyprexa); pc, placebo controlled; PLC, placebo; ran, randomized; SCL-90, symptom checklist-90; VAL, valpromide or valproate; WBC, white

blood cell count; YMRS, Young Mania Rating Scale.

In the 47-week acute mania comparison study of divalproex versus olanzapine, the median time to remission (YMRS ≤12) took significantly longer for divalproex (62 days) than for olanzapine (14 days, $P = 0.047$). Similarly, there was greater improvement in the YMRS with olanzapine versus divalproex at weeks 2 to 15 and week 23; however, from weeks 30 to 47, there was no difference in the two groups. Once remitted, there was no difference in relapse rates between the two groups.

In a second 18-month maintenance study of lithium versus valproate, those patients who had deteriorated or who were intolerant of lithium treatment had a higher response rate with a subsequent crossover to valproate in comparison to those patients who were initially treated with valproate with a subsequent crossover to lithium.

In a maintenance study comparing lithium plus valproate versus lithium alone, the combination was more effective than lithium alone; however, the adverse event profile was higher with the combination treatment. In the largest practical clinical trial (using a randomized open design), Geddes and colleagues found clear superiority of lithium compared to valproate, with the combination marginally better than lithium alone. Despite an indication only in acute mania, and some evidence of inferiority compared to lithium, there is wider use of valproate than lithium in the United States in the maintenance phase of bipolar disorder, case examples of which are illustrated in Figures 34.34–2A,B and 34.34–3. The latest Cochrane analysis in 2013 concluded that limited evidence supported the efficacy of valproate in the maintenance treatment of bipolar disorder.

Valproate in combination with lithium was evaluated in adults with rapid-cycling disorder in preparation for a blind randomization to monotherapy. Disappointingly, only 25 percent of the evaluable sample responded to the combination acutely, and only 19 percent of the intact to treat (ITT) sample responded. After randomization, relapses were about 50 percent on either lithium or valproate monotherapy, suggesting a very low valproate monotherapy response rate (about 12 percent) in this difficult-to-treat but common subtype of rapid cycling (which make up 40 to 60 percent of outpatients in some series).

In a parallel study in children and adolescents with mania, some 43 percent responded acutely to the lithium plus valproate combination, but upon randomization to either monotherapy, some two-thirds failed to maintain their improvement. However, the great majority re-responded to the combination when it was reintroduced. Most of these youngsters also required stimulant comedication and several required an atypical antipsychotic as well. These systematic prospective data in a double-blind maintenance study suggest that, as in many adult patients that valproate monotherapy is very often insufficient, and combination treatment with

lithium or an atypical antipsychotic is typically required in children for sustained mood stabilization.

A recent randomized multicenter study of childhood mania, comparing valproate, lithium, or the atypical antipsychotic risperidone, supported this viewpoint. Geller and colleagues found that risperidone was superior to both lithium and valproate. The percents with CGI-BP improvement of 2 or less were risperidone, 68.5 percent; lithium, 35.6 percent; and valproate, 24.0 percent. However, risperidone had more side effects such as weight gain and prolactin increases. When nonresponders were rerandomized to the addition of one of the other two treatments, those receiving risperidone as part of the regimen again fared the best.

In a secondary prevention study of children at high risk for bipolar I illness by virtue of both the presence of (1) subthreshold symptoms or a bipolar not otherwise specified presentation and (2) a positive parental history of bipolar illness, valproate was no more effective than placebo. Whether other monotherapies would fare better in this early attempt at secondary prevention remains to be established. Thus, because of the documented inefficacy of valproate monotherapy in a large double-blind study in childhood and adolescent acute mania, its inferiority to risperidone in another study, and its lack of prevention of the disorder in children at high risk, the combination of valproate with lithium or with an atypical such as risperidone may be of critical importance in management of these children. The role of valproate in conjunction with the better tolerated atypical antipsychotics in long-term treatment also needs to be further explored.

Schizophrenia and Schizoaffective Disorder

There are no controlled studies of valproate in the treatment of schizoaffective disorder. Clinically, the treatment regimen of schizoaffective disorder usually includes a mood stabilizer, such as valproate, carbamazepine, or lithium, and an atypical antipsychotic. Recent prevalence data have suggested that the use of valproate in combination with atypical antipsychotics has substantially increased.

Although mood stabilizers are not identified as first-line agents in the treatment of schizophrenia, valproate may be useful as an accelerating agent to antipsychotic therapy. In a 28-day, double-blind, parallel-group, multicenter trial of 249 patients who were hospitalized with an acute psychotic exacerbation of schizophrenia, subjects were randomized to one of four treatments: (1) olanzapine monotherapy (15-mg fixed dose by day 15), (2) risperidone monotherapy (6-mg fixed dose by day 15), (3) olanzapine plus divalproex (15 mg/kg/day on day 1, titrated as clinically indicated to a maximum dose of 30 mg/kg/day by day 12), or (4) risperidone plus divalproex. The mean daily doses of divalproex were 2,364 mg in the olanzapine group and 2,259 mg in the risperidone combination

group, respectively. The primary outcome variable was percentage change (baseline to final evaluation) in the total score of the Positive and Negative Syndrome Scale (PANSS). Over the 28-day study period, improvement in the PANSS score was evident in the monotherapy and combination groups. However, statistically significant differences favoring the combination groups were noted as early as day 3 and maintained significance through day 21. At the final day-28 visit, there was no significant difference in PANNS total score reduction in the monotherapy groups versus the combination groups. The acceleration of treatment response in the combination group was associated with improvement in positive symptoms, but not in negative symptoms. Weight gain was not significantly different between the olanzapine–divalproex group (8.3 lb) versus the olanzapine group (7.7 lb); however, there was significantly more weight gain in the risperidone–divalproex group (7.5 lb) versus the risperidone group (4.2 lb).

This study does suggest that time course of treatment response of antipsychotic therapy (specifically, positive symptom severity) can be accelerated with divalproex; further study is encouraged to evaluate potential pharmacoeconomic and long-term benefits.

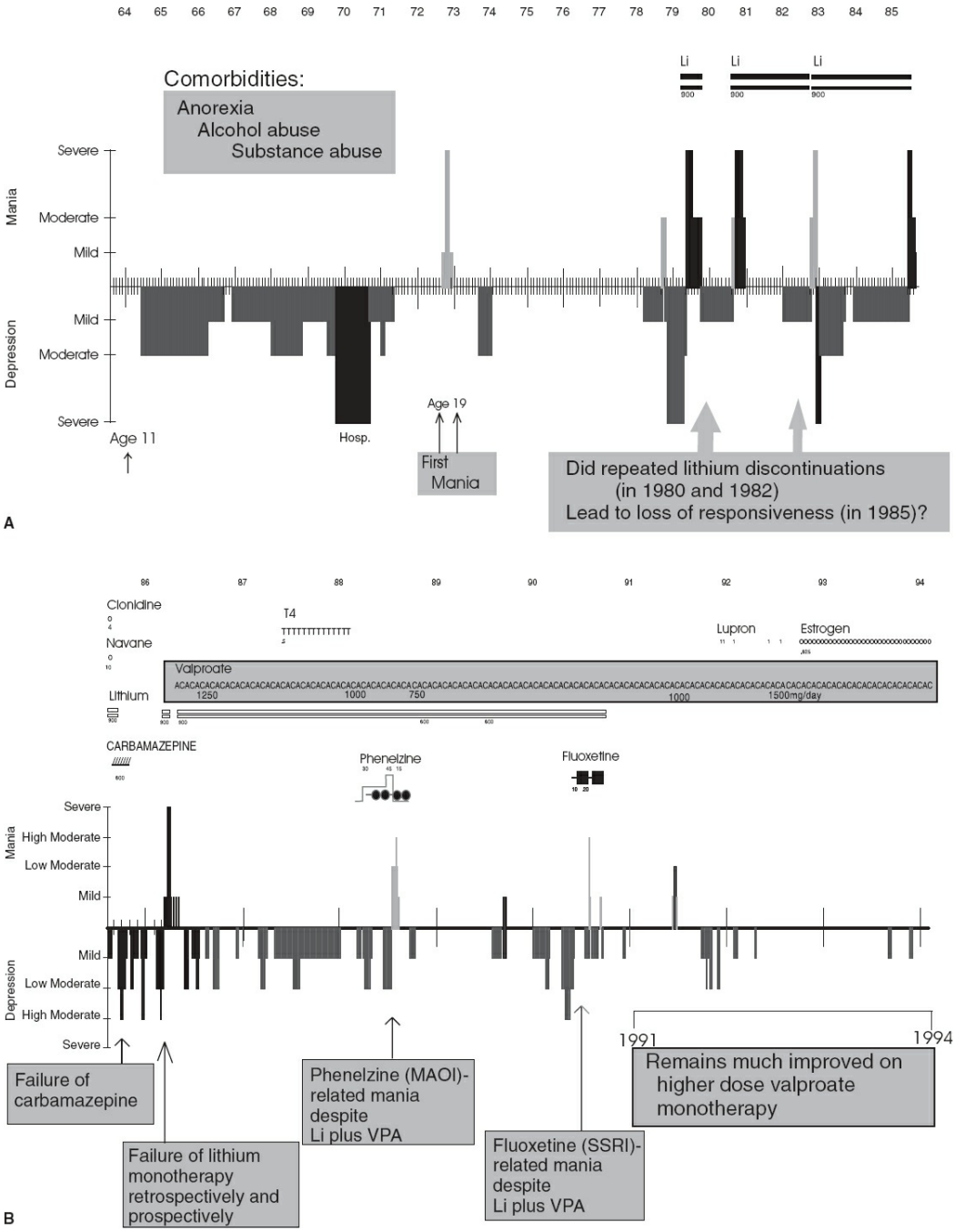

FIGURE 34.34–2. **A:** Early-onset illness (at age 11) with depression and multiple comorbidities and ultimate failure to respond to lithium in 1985. **B:** Good prophylactic response to valproate in this lithium and carbamazepine-nonresponder. She did well on the combination of lithium and valproate and then on valproate monotherapy since 1990. She had no further hospitalizations for mania (from 1986 to the present), but had occasional depressive recurrences. Note shift in time scale from (A). (From Post RM, Leverich GS. *Treatment of Bipolar Illness: A Casebook for Clinicians and Patients. Treatment of Bipolar Illness*. New York: W.W. Norton & Co.; 2008:198–200.)

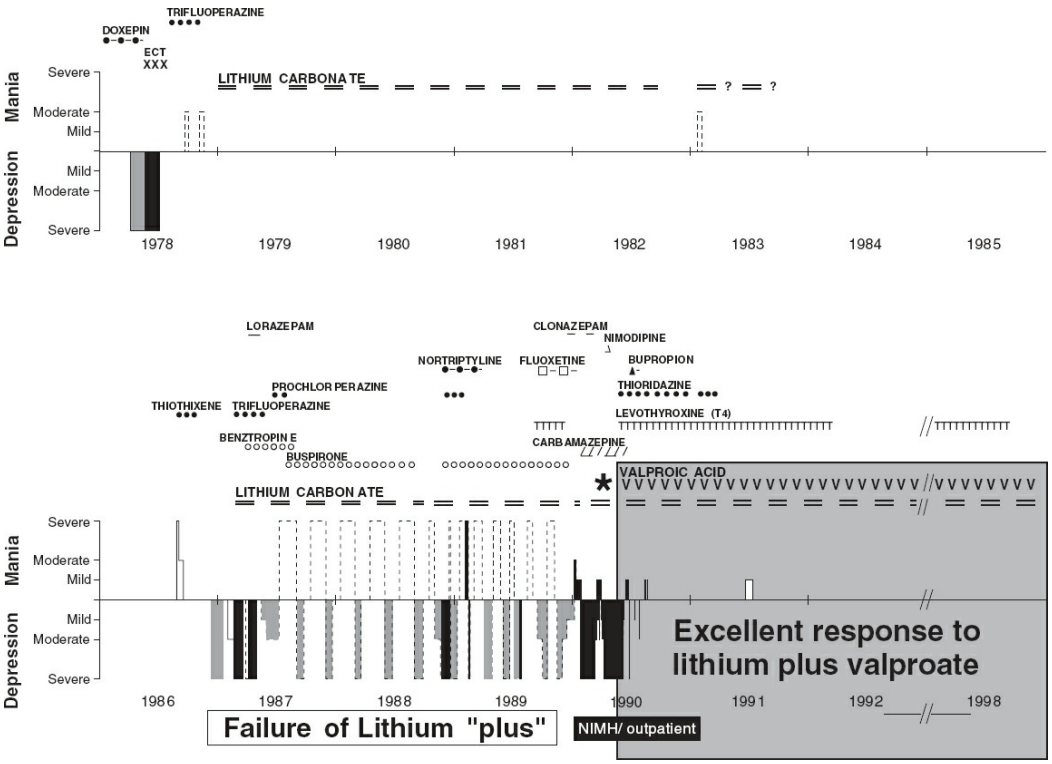

FIGURE 34.34–3. Valproate augmentation in a lithium-refractory bipolar patient with profound recurrent psychotic depression. Reemergence of the illness in 1986 was unresponsive to the reinstitution of lithium, buspirone, a variety of antidepressants, adjunctive antipsychotics, and minor tranquillizers. Carbamazepine, alone and in combination with lithium, was effective in reducing the amplitude of the manias, but had no impact on the patient's severe psychotic, delusional depressions. When carbamazepine was discontinued and valproate was added to lithium, the patient soon started to improve and reached a complete and sustained remission for the first time in 5 years. Her remission, illustrated in the figure from 1990 to 1998, persisted through 2008. ECT, electroconvulsive therapy.

Alcohol Withdrawal and Relapse Prevention

Divalproex, carbamazepine, and gabapentin have all been studied in the treatment of alcohol withdrawal. Although the intermediate (diazepam [Valium]) and long-acting benzodiazepines (lorazepam [Ativan] and clonazepam [Klonopin]) are the standard established agents for alcohol detoxification, comparative studies have shown that carbamazepine decreases global distress more than oxazepam and anxiety more than lorazepam.

In an open-label, 5-day study involving patients undergoing a standard protocol for alcohol detoxification, as needed dosing of lorazepam, with or without divalproex (500 mg by mouth three times a day), was given based on a Clinical Institute Withdrawal Assessment for Alcoholism, revised (CIWA-A-R) score. Although the CIWA-A-R score reduction was not significantly different, the cumulative lorazepam dosage was lower in the group that received divalproex. The treatment was well tolerated without significant side effects or hepatocellular enzyme elevation. In a second

study of 23 inpatients hospitalized for dysphoric mania complicated by substance use (primarily alcohol), treatment of the comorbidity with divalproex resulted in not only dysphoric symptom reduction (as measured by the YMRS and Hamilton Rating Scale for Depression), but also effective alcohol withdrawal and abstinence over the course of 20-week prospective follow-up. Mild hepatocellular enzyme elevation (not greater than three times normal) at the time of admission, presumably related to alcohol ingestion, decreased over the study period.

Alcohol relapse prevention (i.e., maintaining sobriety after alcohol detoxification) has also been evaluated with the mood-stabilizing anticonvulsants carbamazepine and valproate. In a 12-month, placebo-controlled study in 29 alcohol-dependent patients, carbamazepine treatment was associated with a decrease in the time to return to heavy drinking and a decrease in a number of drinking measures at 2- and 4-month follow-up. Recent data have suggested similar effectiveness with divalproex. Twenty-nine outpatients with alcohol dependence participated in a 12-week, double-blind, placebo-controlled, pilot trial of divalproex (mean divalproex level at week 12 was 88.2 μg/mL). There was a greater reduction in irritability in the divalproex treatment group, and a smaller percentage of divalproex-treated patients relapsed to heavy drinking (37 percent) compared with placebo-treated patients (63 percent; $P \leq 0.05$). There were no significant differences in side effects reported or abnormal laboratory values; in both treatment groups a reduction in hepatocellular enzymes (alanine aminotransferase [ALT], aspartate aminotransferase [AST], and γ-glutamyltransferase [GGT]) during the course of the study was observed. Given the extensive comorbidity of alcoholism with other axis I disorders, particularly bipolar I and II disorders, treatment with valproate may stabilize mood and help treat components of alcoholism (withdrawal and relapse prevention). These results have now been replicated and extended.

In a 24-week, double-blind, placebo-controlled study of bipolar patients with comorbid active alcoholism maintained on lithium and psychosocial treatment as usual, valproate showed greater effects than placebo on several measures of drinking. Of the 25 patients randomized to placebo, 68 percent returned to heavy drinking compared with 44 percent of the 27 randomized to divalproex ($P = 0.02$). Liver function tests, ALT, and AST did not differ, while GTP (γ-67P) was higher in those on placebo and was correlated with the amount of alcohol consumed per week. Of those who did return to heavy drinking (four or more drinks per day for women; five or more drinks per day for men), time to relapse was also significantly delayed for those on divalproex (93 days) compared with 62 days on placebo ($P = 0.02$).

Although other agents are effective in primary alcohol dependence (i.e., naltrexone [ReVia] and acamprosate [Campral], which are FDA approved,

and now topiramate in two large randomized-controlled trials, and also gabapentin), none have been specifically studied in bipolar patients with comorbid alcoholism with the exception of valproate, making it a particularly good agent for such dual diagnosis patients. It is noteworthy that in patients with bipolar illness, women are 7.35 times and men 2.8 times more likely than those in the general population to have problems with alcohol use and abuse. Greater amounts of depression and anxiety were correlates of the women who abused alcohol, suggesting that it was a form of self-medication, and that earlier and more effective intervention on these target symptoms potentially could have lessened the incidence of alcohol comorbidity.

Panic Disorder

Several open-label studies suggest that valproate may be effective in treating panic attacks. In one open-label study of lactate-induced panic attacks, valproate effectively blocked the reintroduction of panic attacks on lactate rechallenge in 10 of 12 patients; furthermore, 71 percent of patients had a greater than 50 percent reduction in panic attack frequency, with 43 percent achieving complete remission. These data are consistent with the marked decreases in anxiety when valproate is used in the treatment of bipolar depression. However, in a double-blind, placebo-controlled, crossover study, 35 patients with panic disorder responded equally well to valproate and placebo.

Posttraumatic Stress Disorder

Valproate has been investigated in several open trials (total N = 53) of patients with combat-related PTSD. The first open trial noted a decrease in avoidance and hyperarousal but no change in intrusive symptoms. The latter two open trials used contemporary diagnostic criteria and rating scales. In the first 8-week study, adjunctive valproate (mean dose of 1,365 mg, mean serum level of 80 μg/mL) was associated with a significant decrease in intrusion and hyperarousal but no change in avoidant symptoms. In the second 8-week open trial, which included a 2-week washout with a majority of patients on monotherapy, valproate (mean dose of 1,840 mg ± 270 mg, mean level of 69 μg/mL ± 29 μg/mL) was associated with a significant reduction in all symptom clusters of intrusion, avoidance, and hyperarousal.

Further clinical studies of valproate in the extinction of conditioned fear would appear appropriate in light of preclinical data that valproate specifically increases the promotor of a BDNF variant that is involved in the new learning that is necessary for extinction of conditioned fear. It is likely that the effects of valproate as a histone deacetylase inhibitor are involved in this effect.

Impulse Control Disorder

Several impulse control disorders, such as intermittent explosive disorder, pathological gambling, and kleptomania, share some clinical features with mania. Valproate has been investigated in several case reports and open trials with generally positive results, particularly if manic or mixed symptoms were present. The only placebo-controlled study of valproate in an impulse control disorder evaluated intermittent explosive disorder. One hundred and sixteen patients with intermittent explosive disorder were part of a 12-week, randomized, placebo-controlled study evaluating valproate in the treatment of impulsive aggression; in this study, only patients with impulsive aggression associated with borderline personality disorder responded to valproate. A second controlled, single-blind study randomized patients with pathological gambling to lithium (N = 23, mean dose of 795 mg ± 261.5 mg/day, mean level of 0.8 mEq/L ± 0.1 mEq/L) or valproate (N = 19, mean dose of 873 mg ± 280 mg/day, mean level of 72 μg/mL ± 14 μg/mL). Using the Yale-Brown Obsessive Compulsive Scale modified for pathological gambling as the primary outcome measure, those taking lithium and valproate showed significant improvement.

Highly significant effects of valproate compared with placebo in a crossover design were seen in children and adolescents with explosive temper and a disruptive behavioral disorder (oppositional defiant disorder or conduct disorder). These children were also severely mood-dysregulated and some may have met criteria for bipolar not otherwise specified, but they were not bipolar II or bipolar I. A positive family history of bipolar disorder in first-degree relatives was also not a prominent feature of this group. These findings are of considerable importance because they suggest the utility of valproate in children with explosive aggression (86 percent response on valproate vs. 25 percent on placebo), even in those who do not meet strict criteria for bipolar disorder. The fact that these irritable, hostile, and aggressive children show marked improvement over placebo, and that in another study of children with more classic bipolar presentations meeting DSM-IV criteria, valproate did not separate from placebo, suggests the usefulness of valproate in just this area of greatest diagnostic controversy; those who are the highly impaired by their irritability and dyscontrol but may or may not end up in the bipolar diagnostic spectrum.

Further concerted study in this area and in those with conduct disorder would be of great clinical importance, because the therapeutic approaches to these large numbers of children remain obscure and based on only a sparse literature.

Borderline Personality Disorder

Several open trials have suggested valproate's effectiveness in several symptom domains in patients with borderline personality disorder (i.e.,

mood lability, irritability). There have been three placebo-controlled studies generally supporting these preliminary observations. In the first, 10-week, placebo-controlled study, 12 patients treated with valproate (mean dose of 750 mg) had nonsignificant improvement in global symptom severity and functioning compared to patients treated with placebo. A second, 6-month, placebo-controlled study of 30 euthymic bipolar II disorder patients who also met criteria for borderline personality disorder were randomized to divalproex (mean dose of 850 mg) versus placebo. The primary outcome variable was the mean change in symptom checklist-90 (SCL-90) scales of interpersonal sensitivity, anger and hostility, and depression and the Overt Aggression Scale, modified (OAS-M). In a last-observation-carried-forward analysis, controlling for baseline severity, divalproex was more effective than placebo in reducing interpersonal sensitivity, anger and hostility, and aggression.

Another 12-week, randomized, placebo-controlled study evaluated the effectiveness of divalproex in the treatment of impulsive aggression in patients with borderline personality disorder (N = 96), intermittent explosive disorder (N = 116), and PTSD (N = 34). The primary outcome variable was the mean score reduction in the OAS-M obtained during the last 4 weeks of treatment. In the intent-to-treat analysis, there was no significant treatment difference in OAS-M scores between those patients treated with divalproex versus those treated with placebo. However, when the data were analyzed by diagnostic subgroup, there were significant treatment differences in patients with borderline personality disorder (χ^2 = 3.95; P = 0.047) but not in those with intermittent explosive disorder or PTSD. In the data set that could be evaluated (those who completed at least 21 days), patients with borderline personality disorder and impulsive aggression treated with divalproex versus placebo had a significant decrease in measures of aggression, verbal assault, assault against objects, and assault against others but not in assault against self. As well, there was a significant reduction in irritability, as measured by the OAS-M Irritability score, and severity of illness, as measured by the Clinical Global Impression (CGI) Scale. The mean dose of divalproex was 1,404 mg (with a range between 500 and 2,250 mg), with a corresponding serum level of 65.5 μg/mL (with a range between 0 and 147 μg/mL).

Behavioral Agitation of Dementia

Several open-trial investigations had suggested valproate's effectiveness in reducing behavioral agitation associated with dementia. The dose was generally lower than that used in other conditions, most likely reflecting an older patient population and a pharmacokinetic profile characterized by decreased protein-binding capacity. However, the one double-blind, placebo-controlled, 6-week study of higher doses of divalproex in the treatment of manic symptoms in elderly demented patients was suspended

prematurely because of a disproportionate rate of adverse events (primarily sedation and thrombocytopenia) and dropouts (noncompliance) reported with divalproex compared to placebo. At the time of study suspension, 85 patients had been randomized to placebo, and 87 had been randomized to divalproex (median dose of 1,000 mg daily or 18 mg/kg/day, mean weekly serum valproate levels ranging from 55.3 to 68.9 μg/mL). Although there was no significant difference in the primary outcome as measured by the Bech–Rafaelsen Mania Scale, there was a greater reduction in the Cohen–Mansfield Agitation total score (noted as early as week 2 and maintained throughout the 6-week period) and the verbal agitation subscale score at week 6. This study suggests that, although agitation can be reduced with valproate, many elderly patients may not be able to tolerate an aggressive dosing strategy (18 mg/kg/day).

THERAPEUTIC INDICATIONS

Neurological Conditions

Valproate is currently approved by the FDA for (1) monotherapy or adjunctive therapy of complex partial seizures that occur in isolation or in conjunction with other types of seizures, (2) monotherapy and adjunctive therapy of simple and complex absence seizures, and (3) adjunctive therapy for patients with multiple seizures that include absence seizures. Divalproex has an FDA indication in these same epilepsy syndromes but additionally has indications for (4) prophylaxis of migraine. The ER preparation is approved for seizures, acute mania, and migraine.

Psychiatric Conditions

Valproate and carbamazepine ER are currently the only mood-stabilizing anticonvulsants that are FDA approved for the treatment of acute mania associated with bipolar disorder. Based on extensive open studies and case series in treatment-refractory and rapid-cycling patients, valproate has also been widely used in the prophylactic treatment of recurrent mania and depression and has been recommended as a first-line option in many treatment guidelines. However, caution is urged in the use of valproate in women of childbearing age as emphasized below.

Valproate has been shown to be particularly effective in acute mania with depressive, dysphoric, and mixed symptoms, acute mania associated with a rapid-cycling course of illness specifier (four or more episodes in a 12-month period), and acute dysphoric mania complicated by substance abuse, particularly alcohol abuse and dependence (Table 34.34–3). Given its current neurological FDA indications, valproate may be particularly useful in bipolar disorder complicated by seizure disorder, electroencephalogram (EEG) abnormalities, head trauma, and migraine. Outside of formal DSM-IV-TR diagnoses, valproate appears to be effective

on target symptoms of irritability, aggressiveness, mood lability, and impulsivity.

PRECAUTIONS

Serious Adverse Reactions

Rare serious side effects—as noted in the FDA black box warnings—include hepatic failure, pancreatitis, and hyperammonemic encephalopathy associated with patients with urea cycle disorders. It is important to obtain baseline hepatic function tests before valproate treatment. However, baseline or serial laboratory monitoring does not necessarily predict severe hepatotoxicity or pancreatitis. Close clinical observation is far superior to routine clinical monitoring in this regard. Additional warnings include sedation in the elderly and thrombocytopenia.

Valproate is rated pregnancy category rating D, as noted by increased risk (1 to 4 percent) of neural tube defects, such as spina bifida (Table 34.34–6). Recent studies have also highlighted the notable occurrence of other teratological effects. The rate of any adverse outcome (a major congenital malformation or fetal death) was higher with valproate (20.3 percent) than with phenytoin (Dilantin) (10.7 percent), carbamazepine (8.2 percent), or lamotrigine (1.0 percent). The effects were not attributable to factors other than anticonvulsant exposure during pregnancy, and, notably, with valproate the adverse events were dose related.

In addition, behavioral teratological effects have also been reported. A syndrome of moderate to severe developmental delay has been identified, with a small percentage having severe retardation and an intelligence quotient (IQ) below 70. Thus, it would appear that avoiding pregnancy during valproate treatment or switching to other less problematic agents when planning for pregnancy is of very great importance. Although neural tube defects can be monitored for and detected with modern techniques (i.e., fetal ultrasound in weeks 18 to 20; α-fetoprotein; and if needed, amniotic fluid at 16 weeks), there is currently no way of predicting the behavioral deficits which, upon follow up of children exposed to valproate in utero, appear substantially more severe that with other anticonvulsants. For example, in those exposed to valproate, the IQ measured at 3 years of age averaged 9 points lower than those exposed to lamotrigine. Valproate's use in combination with other agents also appears to be an additional risk factor. Valproate has also been associated with a many fold increased risk for autism of childhood or lifetime onset. This clinical increased risk has been validated in animal models of fetal exposure to valproate.

Table 34.34–6.
Black Box Warnings and Other Warnings

More Serious Side

Effect	Management Considerations
Hepatotoxicity	Rare, idiosyncratic event Estimated risk 1:118,000 (adults) Greatest risk profile (polypharmacy, younger than 2 years of age, mental retardation) → 1:800
Pancreatitis	Rare, similar pattern to hepatotoxicity Incidence in clinical trials data is 2 of 2,416 (0.0008%) Postmarketing surveillance shows no increased incidence Relapse occurs with rechallenge Asymptomatic amylase not predictive
Hyperammonemia	Rare—more common in combination with carbamazepine (Tegretol) Associated with coarse (flapping) tremor and may respond to L-carnitine administration
Associated with urea cycle disorders	Discontinue valproate and protein intake Assess underlying urea cycle disorder Divalproex is contraindicated in patients with urea cycle disorders
Teratogenicity	Neural tube defect: 1–4% with valproate; high incidence of other congenital anomalies; developmental delay and IQ decreases prominent. Preconception education and folate-vitamin B complex supplementation and relative avoidance of valproate for women of childbearing potential
Somnolence in the elderly	Slower titration than conventional doses Regular monitoring of fluid and nutritional intake
Thrombocytopenia	Decrease dose if clinically symptomatic (i.e., bruising, bleeding gums) Thrombocytopenia more likely with valproate levels ≥110 µg/mL (women) and ≥135 µg/mL (men)

Adapted from *Physician's Desk Reference*. Oradell, NJ: Medical Economics Company; 2002.

Given the evidence of a highly elevated risk for neural tube and other teratological defects and for lower IQ, most neurologists attempt to avoid the use of valproate in women who wish to become pregnant. If the use of valproate is unavoidable because of failure of other agents, most would recommend the use of folic acid, and vitamins B6 and B12 in hope of ameliorating the known risks to the fetus, in addition to careful attention to the optimal forms of birth control. These cautions would appear even more appropriate to women of childbearing age who have bipolar disorder. The characteristics of mania itself, such as hypersexuality and poor judgment, make the risk of an unplanned or unwanted pregnancy even more likely. Thus avoidance of the use of valproate in these women would appear indicated, and in the absence of other alternative treatments, special emphasis on the importance of birth control and adjunctive use of folate, B6 and B12. It has been suggested that the analog of valproate valnoctamide has both anticonvulsant and antimanic efficacy and may have a lesser likelihood of congenital malformations based on studies in laboratory animals. This drug may have lesser potency as a histone deacetylase inhibitor which some have posited as the mechanism of valproate's teratogenic effects.

Valproate has been classified as compatible with breastfeeding, as case

reports have consistently reported low concentrations (1 to 10 percent of mother's serum level) of valproate in breast milk.

Several preclinical studies have suggested that valproate can stimulate human immunodeficiency virus 1 (HIV-1) and hepatitis C virus (HCV) replication. Although small clinical studies in HIV-1–infected and HCV-infected patients do not confirm these preclinical observations, close clinical monitoring is recommended.

In long-term toxicity studies in rats, valproate was associated with an increase in pulmonary adenomas and fibrosarcomas. There is no evidence to suggest a carcinogenic effect in humans, and to the contrary, there is considerable evidence of an antineoplastic effect in several tumor types, including chronic myelogenous leukemia (CML).

Minor Adverse Reactions

Common side effects include gastrointestinal (GI) distress, tremor, sedation, benign hepatic transaminase elevation, leukopenia, thrombocytopenia, hair loss, increased appetite, and weight gain (Table 34.34–7). Despite weight gain there are no increases in triglycerides or total cholesterol. Generally, most of the common side effects are dose related and can often be minimized by dose reduction, dose consolidation, or adjunctive pharmacotherapy. Valproate causes dose-related increases in homocysteine, and for this reason many would recommend routine cotreatment with folate to reduce this risk.

The relationship between valproate, polycystic ovaries, and polycystic ovarian syndrome (menstrual irregularity and clinical evidence of hyperandrogenism) remains controversial because patients with epilepsy and obesity appear to be at increased risk for the syndrome. Studies in women with bipolar disorder taking valproate or lithium have reported conflicting results. In one study of 22 outpatients, most patients reported menstrual irregularity, but none of them had evidence of significant hirsutism on physical examination or hormonal evidence consistent with polycystic ovarian syndrome. In another study of 32 bipolar outpatients, 7 patients currently on valproate and experiencing menstrual problems had evidence of hyperandrogenism (hirsutism and elevated testosterone).

Natalie Rasgon and colleagues again found a high incidence of menstrual irregularities in women with bipolar illness prior to treatment, and these increased further with valproate therapy. Hirsutism and significant elevation of serum testosterone were also not observed.

In contrast, data from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) group did find statistically significant testosterone increases on valproate and, rarely, hirsutism. Thus, it would appear that there is a clear pretreatment increased risk of the polycystic ovarian syndrome (menstrual irregularities, weight gain, increases in testosterone, even without polycystic ovaries themselves) and polycystic

ovaries in women with epilepsy, which is further exacerbated by valproate. However, in women with bipolar disorder, the polycystic ovarian syndrome at baseline pretreatment is relatively uncommon and exacerbation by valproate less apparent and more inconsistent.

Table 34.34–7.
Some Valproate Side Effects and Their Potential Management

Side Effect (Frequency)	Potential Treatments	Support	Comment
GI distress (O)	Switch to enteric coated preparation Add H_2 blocker Give with meals or all H.S.	++ + +	Effect may be additive with lithium
Tremor (O)	↓ dose Switch to ER formulation Add Propanolol	+ + ±	Additive with lithium?
Weight gain (O)	Early diet and exercise instructions Augment with topiramate, zonisamide	++ ++ ++	Additive with lithium?
Alopecia (hair loss) (UC)	Prophylaxis with zinc and selenium supplements	+	Straight hair may grow back curly
Thrombocytopenia (low platelets) (UC)	Discontinue valproate	++	Monitor for bleeding
PCO syndrome (UC)	Treatment with oral contraceptives will prevent PCOs Switch to lamotrigine	++ ++	Menstrual irregularities may precede use of VPA and may increase on VPA Testosterone not ↑ in 4 of 5 studies Hirsutism rare
Hepatic enzyme elevation; <3 times normal (O)	Monitor direction and rate of change	+	Caution patient to advise you if RUQ pain, fever, malaise, fatigue, cola-colored urine, or jaundice occurs. Risk (1:600) in children <2 years of age for fatal hepatitis.
Hepatitis (VR)	Discontinue VPA	++	
Pancreatitis (VR)	Discontinue VPA, monitor amylase	++	Patient to advise you if severe GI pain, nausea, vomiting occurs
Hypothermia (VR)	Discontinue VPA	+	Especially with topiramate
Metabolic:			
A. Asymptomatic ↑ ammonia (UC)	↓ Dose, add l-carnitine ↓ Dose, add l-carnitine Discontinue VPA	+ + ++	Is there a history of urea cycle abnormalities in patient or family?
B. Coarse, "flapping" tremor (UC)	Discontinue VPA, add l-carnitine, folate Add folate	++ ++ ++	EEG background slowing and high voltage theta–delta activity 75% occurs in first week of Rx VPA may lower folate
C. Encephalopathy (VR)			
D. ↑ homocysteine (C)			
E. ↓ carnitine and			

acylcarnitine (?)			
Teratogenicity:	Avoid pregnancy while on VPA. Use oral contraceptives, other devices.	+++	Avoid VPA in planned pregnancies
A. Neural tube defects; Spina bifida in 1–2% of exposed (UC)	Use folate ≥1 mg/day, vitamin B6 and B12 prophylactically in women of childbearing age	+++	Especially avoid VPA in combination with other anticonvulsants
B. Developmental delays (mild to severe C,D)	Avoid pregnancy	++	

O, occasional; UC, uncommon; VR, very rare; C, common; D, dose related; PCO, polycystic ovaries; PCOs, polycystic ovarian syndrome; H_2, histamine; ER, extended release; VPA, valproate; RUQ, right upper quadrant; GI, gastrointestinal; EEG, electroencephalogram;?, ambiguous or uncertain.

Previous studies in epilepsy have been unable to fully assess baseline endocrine profile before valproate treatment, impact of nonspecific weight gain as a risk factor for development of the polycystic ovarian syndrome, and the extent to which epilepsy and bipolar disorder are independent risk factors unrelated to medication exposure. All aspects of the syndrome are reversible in patients with epilepsy, if valproate is discontinued and lamotrigine is substituted. Menstrual history and weight should be followed carefully in women of reproductive potential who begin treatment with valproate. Since the polycystic ovarian syndrome can be mitigated by birth control pills, this side effect should be described but should not preclude women with bipolar illness from treatment with valproate. More pressing is the need to inform women of the major risks to the fetus should they become pregnant as discussed above.

DRUG INTERACTIONS

Valproate and Other Psychotropics

Valproate is commonly prescribed in combination with other psychotropic agents. These combination therapies are generally well tolerated and are often more effective than valproate alone. Outside of several specific concerns, the most likely caution is the additive side-effect burden of the two medications. For example, lithium and valproate are often used in combination. Additive side effects of tremor or cognitive dulling can often be treated with dose reduction. Additive problems with weight gain can be more daunting. Recent use of the combination of bupropion ER 150 to 300 mg with naltrexone 50 mg has produced dramatic weight loss in many individuals but this regimen has not been systematically studied for valproate or its combination with lithium. Additive GI problems and

dyspepsia can be treated with adjunctive agents histamine 2 [H_2] antagonists.

Typical antipsychotics and valproate are often prescribed together, especially in acute mania. Again, increased sedation and an increase in extrapyramidal side effects have been reported. More recently, atypical antipsychotics and valproate have been used together, with increased sedation (olanzapine) and weight gain (risperidone) reported as side effects. It is noteworthy, however, that valproate may prevent the increases in cholesterol induced by some atypical antipsychotics. This is thought to be attributable to valproate's influence on peroxisome proliferator-activated receptors (P-PARs), which increase β-oxidation and are transcription factors in steroid–thyroid hormone–retinoic receptor super family.

Valproate has also been used as an adjunctive mood stabilizer and anticonvulsant in patients treated with clozapine. In fact, for patients who have had a seizure related to high-dose clozapine, valproate is the preferred agent over carbamazepine, given its risk of agranulocytosis. The combination is generally well tolerated, with the exception of sedation and weight gain.

There are clinical data to suggest greater efficacy of the lamotrigine–valproate combination for treatment of complex-partial epilepsy compared with either drug alone and relatively good tolerability. Given valproate's efficacy in acute mania and lamotrigine's efficacy in acute bipolar depression, it is likely that the two will be used together in psychiatric practice. Given valproate's plasma protein–binding capacity and its metabolic inhibition of lamotrigine, free levels of lamotrigine are doubled when used in combination with valproate; to reduce the risk of serious rash, the initial dose, rate of titration, and final dose of lamotrigine should be reduced by approximately 50 percent of the standard monotherapy recommendations when combined with valproate.

Two case series suggesting efficacy of the valproate–carbamazepine combination for refractory rapid-cycling bipolar disorder have been reported. The pharmacokinetics of this combination warrants clinical anticipation and monitoring, however, adjunctive carbamazepine, as well as its cytochrome P450 3A4 autoinduction and heteroinduction, generally reduces levels of valproate, which are typically not consequential. However, valproate increases levels of free carbamazepine and its active metabolite by displacement from plasma protein. Valproate also increases levels of carbamazepine-10,11-epoxide by inhibiting its metabolism via epoxide hydroxylase to the inactive diol metabolite. Thus the increases in the carbamazepine-10,11 epoxide may not be evident if only carbamazepine levels alone are measured. Given these pharmacokinetic interactions, side effects of the combination generally are related to increased levels of carbamazepine and its active carbamazepine-10,11 epoxide, potentially

yielding diplopia, ataxia, sedation, and confusion in those already near their carbamazepine side-effect threshold. Lowering the carbamazepine dose in anticipation of adding valproate will generally avoid this difficulty (Table 34.34–8).

Antidepressants are often used in combination with valproate, although the efficacy of this approach has recently been questioned. There are case reports of the tricyclic antidepressant amitriptyline increasing valproate concentrations and, conversely, valproate increasing the levels of tricyclics. Additionally, adjunctive fluoxetine (Prozac) has been reported to increase serum levels of valproate.

Table 34.34–8.
Pharmacokinetic Interactions of Valproate

Metabolic Interactions		
VPA→↑ Drug	**Drug→↑VPA**	**Drug→↓VPA**
Benzodiazepines	ASA	CBZ
CBZ-E	Cimetidine	± Lamotrigine
Ethosuximide	Fluoxetine	Phenobarbital
Lamotrigine[a]	Felbamate	Phenytoin
Phenobarbital	Erythromycin	Rifampin
Phenytoin	Phenothiazines	Folate
TCAs	Risperidone	
Binding interactions	Nimodipine	
	Discontinuation of enzyme inducers	
VPA→↑Free Drug	**Drug→↓Free VPA**	
CBZ	ASA	
Diazepam	NSAIDs	
Phenytoin		
Tiagabine		
Tolbutamide		
Warfarin		
Zidovudine		

VPA, valproate; CBZ-E, carbamazepine-10,11-epoxide; TCAs, tricyclic antidepressants; ASA, aspirin; CBZ, carbamazepine; NSAIDs, nonsteroidal anti-inflammatory drugs.

[a]VPA doubles lamotrigine levels and increases risk of serious risk.

Valproate and Other Drugs

In contrast to carbamazepine, the primary pharmacokinetic liability of valproate is related to its plasma protein–binding capacity and the inhibition of other drugs metabolized by the liver. For example, valproate increases the free concentration of diazepam by plasma protein–binding site displacement and also has been shown to decrease the serum level of a diazepam metabolite *N*-desmethyldiazepam. Close clinical monitoring of this coadministration is indicated to avoid any potential for diazepam

toxicity. Similar observations (i.e., binding displacement or inhibition of metabolism, or both) have been noted with ethosuximide (Zarontin), phenobarbital (Barbita), chlorpromazine, and aspirin. Because valproate has been shown to inhibit the secondary phase of platelet aggregation, close clinical monitoring is indicated when using the drug with other drugs that affect coagulation, such as aspirin and warfarin (Coumadin).

VALPROATE AND LABORATORY INTERFERENCE

Valproate occasionally interferes with results of laboratory tests. As a fatty acid, serum levels, if measured, reflect an increase. Valproate metabolites may produce a false-positive result in urine analysis for ketones.

DOSAGE AND ADMINISTRATION

In the outpatient setting, one should obtain a baseline hepatic panel and complete blood cell count before starting a patient on divalproex, although this can be ordered concomitantly with or after starting medications in more urgent treatment indications for hypomania or full-blown mania. An appropriate starting dose is 250 mg three times per day, with increases of 250 to 500 mg every several days to a therapeutic threshold (serum level greater than 50 mg/L). Once a steady-state dosage has been achieved, consolidating to bedtime dosing maximizes treatment adherence. The divalproex-ER preparation also lends itself well to single nighttime dosing. Rapid loading of divalproex (20 mg/kg/day) in acute mania has produced therapeutic serum levels more quickly than standard titration and has generally been well tolerated. However, high-dose intravenous valproate (20 mg/kg) over 30 minutes was not effective. This adds to the suggestion that actions induced by chronic administration of the drug are necessary to its antimanic and other psychotropic effects. These findings, taken with those of no greater rapidity of onset of antimanic effects with aggressive titration of the ER preparation compared with less aggressive regimens, suggest that more modest dosing strategies are likely to be as effective, produce fewer side effects, and be more clinically prudent.

REFERENCES

Albanese MJ, Clodfelter RC Jr, Khantzian EJ. Divalproex sodium in substance abusers with mood disorder. *J Clin Psychiatry*. 2000;61(12):916–921.

Baetz M, Bowen RC. Efficacy of divalproex sodium in patients with panic disorder and mood instability who have not responded to conventional therapy. *Can J Psychiatry*. 1998;43(1):73–77.

BALANCE investigators and collaborators, Geddes JR, Goodwin GM, et al. Lithium plus valproate combination therapy versus monotherapy for relapse prevention in bipolar I disorder (BALANCE): a randomised open-label trial. *Lancet*. 2010;375(9712):385–395.

Banach R, Boskovic R, Einarson T, Koren G. Long-term developmental outcome of children of

women with epilepsy, unexposed or exposed prenatally to antiepileptic drugs: a meta-analysis of cohort studies. *Drug Saf*. 2010;33:73–79.

Bersudsky Y, Applebaum J, Gaiduk Y, et al. Valnoctamide as a valproate substitute with low teratogenic potential in mania: a double-blind, controlled, add-on clinical trial. *Bipolar Disord*. 2010;12(4):376–382.

*Bowden C, Gög˘üs A, Grunze H, Haggstrom L, Rybakowski J, Vieta E. A 12-week, open, randomized trial comparing sodium valproate to lithium in patients with bipolar I disorder suffering from a manic episode. *Int Clin Psychopharmacol*. 2008;23(5):254–262.

Bowden CL, Mosolov S, Hranov L, et al. Efficacy of valproate versus lithium in mania or mixed mania: a randomized, open 12-week trial. *Int Clin Psychopharmacol*. 2010;25(2):60–67.

Brady KT, Myrick H, Henderson S, Coffey SF. The use of divalproex in alcohol relapse prevention: a pilot study. *Drug Alcohol Depend*. 2002;67(3):323–330.

Bredy TW, Wu H, Crego C, Zellhoefer J, Sun YE, Barad M. Histone modifications around individual BDNF gene promoters in prefrontal cortex are associated with extinction of conditioned fear. *Learn Mem*. 2007;14(4):268–276.

Carr RB, Shrewsbury K. Hyperammonemia due to valproic acid in the psychiatric setting. *Am J Psychiatry*. 2007;164(7):1020–1027.

Chaudron LH, Jefferson JW. Mood stabilizers during breast-feeding: a review. *J Clin Psychiatry*. 2000;61(2):79–90.

Chen PS, Wang CC, Bortner CD, et al. Valproic acid and other histone deacetylase inhibitors induce microglial apoptosis and attenuate lipopolysaccharide-induced dopaminergic neurotoxicity. *Neuroscience*. 2007;149(1):203–212.

Christensen J, Grønborg TK, Sørensen MJ, et al. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA*. 2013;309(16):1696–1703.

Cipriani A, Reid K, Young AH, Macritchie K, Geddes J. Valproic acid, valproate and divalproex in the maintenance treatment of bipolar disorder. *Cochrane Database Syst Rev*. 2013;(10):CD003196.

Davis LL, Bartolucci AL, Petty F. Divalproex in the treatment of bipolar depression: a placebo-controlled study. *J Affect Disord*. 2005;85(3):259–266.

Donovan SJ, Stewart JW, Nunes EV, et al. Divalproex treatment for youth with explosive temper and mood lability: a double-blind, placebo-controlled crossover design. (Brief Report). *Am J Psychiatry*. 2000;157(5):818–820.

Findling RL, Frazier TW, Youngstrom EA, et al. Double-blind, placebo-controlled trial of divalproex monotherapy in the treatment of symptomatic youth at high risk for developing bipolar disorder. *J Clin Psychiatry*. 2007;68(5):781–788.

Geller B, Luby JL, Joshi P, et al. A randomized controlled trial of risperidone, lithium, or divalproex sodium for initial treatment of bipolar I disorder, manic or mixed phase, in children and adolescents. *Arch Gen Psychiatry*. 2012;69(5):515–528.

Ghaemi SN, Gilmer WS, Goldberg JF, et al. Divalproex in the treatment of acute bipolar depression: a preliminary double-blind, randomized, placebo-controlled pilot study. *J Clin Psychiatry*. 2007;68(12):1840–1844.

Langosch JM, Drieling T, Biedermann NC, et al. Efficacy of quetiapine monotherapy in rapid-cycling bipolar disorder in comparison with sodium valproate. *J Clin Psychopharmacol*.

2008;28(5):555–560.

Loscher W. Valproate: a reappraisal of its pharmacodynamic properties and mechanisms of action. *Prog Neurobiol*. 1999;58(1):31–59.

*Manji HK, Moore GJ, Chen G. Clinical and preclinical evidence for neurotrophic effects of mood stabilizers: Implications for the pathophysiology and treatment of manic-depressive illness. *Biol Psychiatry*. 2000;48(8):740–754.

O'Donovon C, Kusumaker V, Graves GR, Bird DC. Menstrual abnormalities and polycystic ovary syndrome in women taking valproate for bipolar mood disorder. *J Clin Psychiatry*. 2002;63(4):322–330.

Post RM, Leverich GS. *Treatment of Bipolar Illness: A Case Book for Clinicians and Patients*. New York: WW Norton; 2008.

Rasgon NL, Altshuler LL, Fairbanks L, et al. Reproductive function and risk for PCOS in women treated for bipolar disorder. *Bipolar Disord*. 2005;7(3):246–259.

Salloum IM, Cornelius JR, Daley DC, Kirisci L, Himmelhoch JM, Thase ME. Efficacy of valproate maintenance in patients with bipolar disorder and alcoholism: a double-blind placebo-controlled study. *Arch Gen Psychiatry*. 2005;62(1):37–45.

Smith LA, Cornelius VR, Azorin JM, et al. Valproate for the treatment of acute bipolar depression: systematic review and meta-analysis. *J Affect Disord*. 2010;122:1–9.

Thomas SV, Ajaykumar B, Sindhu K, Nair MK, George B, Sarma PS. Motor and mental development of infants exposed to antiepileptic drugs in utero. *Epilepsy Behav*. 2008;13(1):229–236.

Tohen M, Baker RW, Altshuler LL, et al. Olanzapine versus divalproex in the treatment of acute mania. *Am J Psychiatry*. 2002;159(6):1011–1017.

Winsberg ME, DeGolia SG, Strong CM, Ketter TA. Divalproex therapy in medication-naive and mood-stabilizer-naive bipolar II depression. *J Affect Disord*. 2001;67:207–212.

▲ 34.35 Electroconvulsive Therapy

JOAN PRUDIC, M.D., AND YUANLI DUAN, M.D., PH.D.

Electroconvulsive therapy (ECT), also known as electroshock therapy, is a standard psychiatric treatment in which electricity is delivered through the scalp to induce a therapeutic seizure. First conducted in 1938 by two Italian physicians, Ugo Cerletti and Lucio Bini, on patients with schizophrenia, ECT has become the oldest somatic treatment for many psychiatric conditions, including depression, bipolar disorder, schizophrenia, and catatonia. Following professional guidelines, ECT is generally reserved for patients who have failed to respond to pharmacotherapy or have adverse reactions to standard pharmacotherapy, as well as those with severe symptoms such as suicidality, psychosis, or grave functional impairment, resulting in a need for a rapid and definitive response. It is estimated that about one million patients worldwide and 100,000 in the United States received ECT annually. Contrary to the sometime negative depiction of ECT, its safety and efficacy has been much improved with modification of various aspects of the technique such as optimization of the stimulus energy, electrode placement, the use and optimization of general anesthesia, as well as improvement in cardiopulmonary management. The application of ECT is regulated by the FDA, and there are regulatory and professional guidelines for proper consent procedures including situations where the patient lacks capacity. The mechanism of action of ECT remains an active area of hypothesis and research, although a great deal of information has been obtained with the availability of newer research tools.

HISTORICAL OVERVIEW

Convulsive therapies for major psychiatric illnesses predate the modern therapeutic era with the use of camphor reported as early as the 16th century, and several accounts of camphor convulsive therapies from the late 1700s to the mid-1800s. With the success of malarial fever therapy for general paresis, interest in biologic therapies for psychiatric illnesses increased in the early 20th century. By the 1930s, insulin coma and psychosurgery had also been introduced.

Unaware of the history of camphor convulsive therapy, Meduna made the observation that the brains of epileptics had greater than normal numbers of glial cells, while those of schizophrenics had fewer, and hypothesized that there might be a biologic antagonism between convulsions and schizophrenia. Following animal experimentation, camphor was (again) selected as the appropriate agent to use for the therapeutic induction of seizures. In 1934, the first catatonic psychotic

patient was successfully treated using intramuscular injections of camphor in oil to produce therapeutic seizures. Thirteen of a series of 26 patients treated similarly were reported to have had some degree of remission of psychosis. In 1937, a conference on convulsive therapies and insulin coma was held in Switzerland and the use of convulsive treatments spread.

Although effective, camphor was relatively long acting and was soon replaced by pentylenetetrazol. This agent was also associated with problems, principally patient rejection of the noxious sensations associated with GABA antagonism while conscious, for example, increased heart rate, feelings of terror, and myoclonus preictally.

Bini and Cerletti were interested in the use of electricity to induce seizures, and, after a series of animal experiments and observation of the use of electricity commercially, were able to safely apply current across the heads of animals for this purpose. In 1938, the first electroconvulsive treatment (ECT) course was administered to a delusional and incoherent patient, who improved with one treatment and remitted after 11 ECT. Electrical induction of convulsive therapy could be made more reliable and shorter acting than chemically induced convulsive therapies, and, by the early 1940s, had replaced them. In 1940, the first use of ECT occurred in the United States.

Throughout the 1940s, ECT was a first-line therapy for a variety of major psychiatric illnesses. In the ensuing decades, efforts to improve its acceptability and safety profile continued. In 1940 A. E. Bennett developed curare for use in the prevention of ECT-associated fractures, which had previously resulted from the severe muscle contractions associated with the tonic phase of the seizure induction. By the 1950s brief general anesthesia with succinylcholine replaced curare. Other improvements in anesthesia, most notably oxygenation throughout the treatment session and cardiac and oxygen saturation monitoring became standard. Efficacy in disorders other than schizophrenia was demonstrated, most notably major depression.

In an effort to reduce the retrograde memory problems that persisted for some patients after the initial recovery period post-ECT, explorations of nondominant electrode placement and alternative, more efficient, waveforms were undertaken in subsequent decades. The practice of ECT also benefited from the introduction of controlled trials methodology, which demonstrated its safety and efficacy, and from refinements made in diagnostic systems and the process of informed consent. In the 1980s and 1990s efforts to assure uniformly high standards of practice were underway with the publication of recommendations for treatment delivery, education and training by professional organizations in the United States, England, Scandinavia, and Canada among others.

With the current widespread use of pharmacologic agents as first-line treatments for major psychiatric disorders, ECT is now more commonly

used for patients with resistance to those treatments, except in the case of life-threatening illness due to inanition, severe suicidal symptoms, or catatonia. While the failure of subconvulsive stimulation to induce remission of psychiatric illness and the effectiveness of chemical convulsive therapy suggested that the seizure was necessary and sufficient for therapeutic benefit with ECT, it is now known that there is a dose–response relationship with right unilateral ECT, and that bilateral ECT is likely to be ineffective with ultrabrief pulse widths. Work continues to explore the underlying mechanisms and biologic characteristics of effective ECT treatments with a focus on making the treatment focus on appropriate neural networks with a more efficient stimulus as a method of reducing cognitive side effects. With the growing understanding that depression is a chronic disease for many patients, more emphasis has been placed on continuation and maintenance treatments following an acute course of ECT. Work has been published demonstrating the efficacy of both psychotropics and ECT in reducing relapse post-ECT.

Utilization of ECT has diminished since the middle of the last century, but, because ECT remains the most effective treatment for major depression and a rapidly effective treatment for life-threatening psychiatric conditions, ECT, unlike its contemporaneous somatic therapies, such as insulin coma, remains in the active treatment portfolio of modern therapeutics. Its use has shifted from public to private institutions, and it is estimated that approximately 100,000 patients have received ECT annually over the last few decades in the United States (Table 34.35–1).

Table 34.35–1.
Milestones in the History of Convulsive Therapy

1500s	Paracelsus induces seizures by administering camphor (po) to treat psychiatric illness.
1785	First published report of the use of seizure induction to treat mania, again using camphor.
1934	Von Meduna begins the modern era of convulsive therapy using intramuscular injection of camphor for catatonic schizophrenia. Camphor is soon replaced with pentylenetetrazol.
1938	Cerletti and Bini, conduct the first electrical induction of a series of seizures in a catatonic patient and produce a successful treatment response.
1940	ECT is introduced to the United States.
A. E. Bennett develops curare for use as a muscle relaxant at ECT.	
1951	Introduction of succinylcholine.
1958	First controlled study of unilateral ECT.
1960	Ottosson demonstrates that attenuation of seizure expression with an anticonvulsant agent (lidocaine) reduces the efficacy of ECT. In concert with other work suggesting that subconvulsive treatment produces only weak clinical responses, the hypothesis that seizure activity is both necessary and sufficient for efficacy is upheld.
1960s	Randomized clinical trials of the efficacy of ECT vs. medications in the treatment of depression yield response rates that are significantly higher with ECT.
Comparisons of neuroleptics and ECT show that neuroleptic medication is superior for acute treatment, although ECT may be more effective long-term.	
1970	D'Elia develops what has become the most common electrode positioning for right unilateral

ECT.

1976 A constant current, brief pulse ECT device, the prototype for modern devices, is developed.

1978 The American Psychiatric Association publishes the first Task Force Report on ECT with the aim of establishing standards for consent and the technical and clinical aspects of the conduct of ECT.

Late 1970s to Early 1980s

Randomized controlled trials demonstrate ECT is more effective than sham treatment for major depression.

1985 The National Institutes of Health/National Institute of Mental Health Consensus Conference on ECT endorses a role for the use of ECT and advocates research and national standards of practice.

1987 The belief that the seizure in itself is sufficient for clinical response is challenged by a report that the combination of dosage just above seizure threshold and right unilateral electrode placement, while producing a seizure of sufficient duration, is ineffective.

1988 Randomized controlled clinical trials of ECT vs. lithium demonstrate them to be equally effective in mania.

2000 In controlled trials, the dose/response relationship for right unilateral ECT is validated; high-dose right unilateral and bilateral ECT show equal response rates in major depression, but right unilateral electrode placement is associated with fewer adverse cognitive effects.

Convulsive treatment is induced with magnetic stimulation by Lisanby and colleagues.

2001 The largest modern controlled trial of relapse prevention post-ECT with continuation pharmacotherapy demonstrates a significantly better outcome for combined treatment with a tricyclic antidepressant (nortriptyline) plus lithium compared to nortriptyline alone or placebo during the first 6 months post-ECT.

2006 In a large controlled trial, continuation ECT, given in a fixed schedule without concurrent psychotropic medications, is demonstrated to be effective, but a fixed schedule is only equivalent to combination lithium/nortriptyline.

2008 Ultrabrief right unilateral ECT is shown to be equivalent to brief pulse right unilateral and bilateral ECT in efficacy, but difficult to distinguish from a healthy never depressed control group in cognitive performance, including retrograde memory. In the same controlled trial, ultrabrief bilateral ECT at 150 percent above seizure threshold was ineffective, confirming earlier observations that the seizure itself is not sufficient for efficacy.

2013 Feasibility study of Focal Electrically Assisted Seizure Therapy published.

ELECTRICAL INDUCTION OF SEIZURES

Factors in the Generation and Measurement of Electrical Stimulation for ECT

ECT is a form of brain stimulation in which induction of a seizure occurs when an applied electrical stimulus creates an electrical field with flow of current in excitable tissue of the brain sufficient to depolarize cell membranes of neurons synchronously. Fundamentally, the generation and behavior of the electrical stimulus has been conceptualized by Ohm's law:

$$V \text{ (voltage)} = I \text{ (current)} \times R \text{ (resistance)}$$

Voltage is the electromotor force (volts) causing current to flow; current is the rate of flow of electrons (amperes); resistance is the relative ability to pass current (ohms). Ohm's law assumes a direct current circuit, but ECT devices function on alternating current. In this environment, it is necessary

to replace resistance with the term impedance (Z). Impedance involves capacitance and inductance as well as resistance. Since biologic tissues, being nonferrous, have low inductive properties, the inductance term is usually assumed to have little relevance to impedance during ECT. Capacitance (F) is the ability to store charge (farads). Neural tissue has relatively low capacitance.

(10^{-12} to 10^{-10} F), and the electrode skin interface (estimated 0.4 μF), is the source of relatively more of this term.

Although it is possible to trigger a seizure with a single DC pulse of long duration, voltage and current may vary over time in the electrical stimulus as clinically used. A variety of stimulus waveforms have also been used to deliver the stimulus, including sine wave, rectified partial sine wave, brief pulse square wave, and ultrabrief pulse square waves. These different waveforms are not of equal efficiency in depolarizing neural tissue. By examining mammalian neurons, it has become evident that fast rising current results in firing at a lower current intensity, and hence, lower dosage. The slow rise of sine wave results in accommodation and increases in the threshold for neuronal firing. This observation has led to a preference for brief pulse waveforms which "instantly" achieve peak intensity (on the order of microseconds) over the relatively slower rising sine waveform, typically reaching peak intensity in several milliseconds.

Chronaxie, and hence optimal pulse duration, for human neuronal tissue are on the order of 0.1 to 0.2 ms. Furthermore, once nerves have fired, they enter a refractory period, and continued stimulation will not elicit further response. This has led to a reexamination of the need for longer duration stimulation per phase where much of the stimulus is delivered when the neuronal tissue is refractory. Even brief pulse waveforms with pulse widths on the order of 0.5 to 2 ms, commonly used in modern devices, may be expected to deliver excessive energy in the refractory period, although it is an order magnitude lower than with sine wave stimulation. Finally, it has been observed that increasing the frequency of pulses is less efficient in eliciting depolarization than increasing train duration of pulses when using a brief pulse device. This phenomenon is probably also related to the complex relationship of frequency to depolarization, a U-shaped curve in primates, resulting in maximal efficiency in depolarization around 20 Hz. Duration of a pulse train has a more linear covariation with depolarization.

Once the electrical stimulus is generated, current is distributed inversely to the resistance of the tissue compartments traversed. There are three major compartments: scalp (about 200 Ω); skull (about 18,000 Ω); and brain tissue (about 200 Ω). Because current flow is greatest across the lowest resistance, it has been estimated that 80 percent of current is shunted through the scalp. The skull obviously provides the greatest component of resistance, and skull anatomy, for example, the position and

size of sutures, skull thickness, etc., greatly influences the proportion of remaining electrical dosage that enters brain tissue. Differences in scalp and skull anatomy are thought to be a major source of the wide variation among individuals in the amount of electrical dosage required to elicit a generalized seizure or seizure threshold. These differences have been observed to be of the order of 100- to 200-fold. It has been theorized and modeled that interindividual anatomic variation in tissue layer thickness and conductivity could be compensated by varying the amplitude of the stimulus to produce a more focal and uniform electric field exposure, but the degree of adjustment required has not been determined thus far. CSF volume also affects the distribution of the electrical stimulus, with higher volume of fluid being associated with a weaker more focal distribution and field strength.

Another influence on current shunting is interelectrode distance: The closer electrodes are placed on the skull, the more shunting occurs. It is thought that right unilateral electrode placements can result in the creation of a virtual electrode occurring in the area between the two electrodes and, hence, in reduced amounts of electrical dosage required for seizure elicitation. This increased shunting may also result in lower current density in brain tissue.

The quantity of the electric stimulus delivered has been expressed as energy (joules) traditionally, and, in the modern era, as charge (millicoulombs). Substantial evidence indicates that current density in neural tissue is the critical factor responsible for seizure production and the subsequent neurobiologic effects of the electrical stimulation. Although it is not presently possible to measure charge density directly in human neural tissue during ECT, dynamic impedance can be used to approximate charge density because of its relationship to current shunting during stimulus delivery. When electrode site preparation and positioning are standardized, anatomic factors become the major determinants of impedance. If these factors are also constant during the treatment, impedance can then reflect current density in brain tissue.

With the development of other forms of electrical and also magnetic brain stimulation, there has been revived interest understanding and exploring the impact of electrical stimulation on brain tissue. The various elements of brain have differing electrical properties and vary markedly in excitability with myelinated axons being most excitable, and neuronal cell bodies and dendrites least. Tissue response also varies with distance from electrical source, in this case, the stimulating electrode. Although direct recording of the effects of electrical stimulation is not available for ECT, except in animal models, functional neuroimaging and computational modeling are expected to expand the current understanding of the physiology of the treatment.

Devices Used to Deliver ECT

Although constant current, constant voltage, and constant energy devices have been in use, most modern devices are constant current as constant voltage devices have become outmoded. The range of constant current is in the order of 0.5 to 0.9 A in these devices. Dosage is manipulated by varying the time of exposure to a fixed current (total charge). With brief pulse devices, the most commonly used constant current devices, frequency of pulses, pulse width, and duration of pulse train are parameters that can be used to vary the stimulus dosage. Because impedance varies directly with voltage, particular attention must be paid to impedance in a constant current environment. Commonly, impedance is tested prior to stimulus delivery by passage of a low-amplitude, high-frequency test stimulus. If impedance exceeds manufacturer limits, treatment stimuli are not permitted by the device, although some devices are equipped with override features. If too high, impedance is usually lowered by improving skin electrode contact, although other parts of the circuit may also need examination and adjustment. By keeping impedance within safe guidelines, local tissue damage is avoided. When impedance is too low, voltage will also be very low, and may not be sufficient to drive enough current flow to brain tissue where resulting charge density may not be great enough to cause synchronous neuronal depolarization and the desired biologic effects. Low impedance is usually the result of current shunting in the scalp, and is corrected by interrupting any alternative current pathway at the skin surface, for example, saline used to prepare the electrode site runs together and forms a continuous pathway at the scalp.

Calculation of stimulus energy in joules requires information about interelectrode impedance, which is not readily available. Charge is commonly used to measure the intensity of the stimulus used in treatment with brief pulse devices. Charge is the quantity of electrons flowing through a conductor over a given time period (coulombs), and can easily be computed from the parameters of a brief pulse constant current device.

Constant voltage devices allow current to vary inversely with impedance. The typical device generates a sine wave stimulus and has a range of 70 to 170 V. With these devices, high impedance will result in low current flow, which in turn, theoretically, may be inadequate to induce sufficient charge in neuronal tissue to result in neuronal depolarization. Because the practitioner does not have access to crucial information about impedance when using a constant voltage device, and it may be difficult to assess current intensity in this setting, and judgments about the adequacy of parameters can be difficult to make. The use of these devices in the United States has become outmoded.

ECT device output in the United States has been capped at 100 J, which corresponds to a charge of 500 to 576 mC, depending on the particular

device specifications. With ultrabrief pulse width, this range is adequate to deliver right unilateral ECT to almost all patients. ECT devices in the European Union and elsewhere can have an extended range up to 200 J to allow delivery of adequate dosage with brief pulse right unilateral electrode placement, which is mandated in some EU nations.

MECHANISMS OF ACTION

The relative inaccessibility of the central nervous system has made understanding the pathophysiology of psychiatric diseases a continually evolving process. As a consequence, the mechanisms of action of ECT, the oldest somatic treatment currently in use in the field of psychiatry, continue to be an active area of hypothesis and research. Past and present, theories abound, and a wide range of data, particularly at the preclinical level, is available on the effects of electroconvulsive stimulation (ECS).

Proposed mechanisms have involved theories in almost every domain. These include psychological constructs; structural explanations, for example, neural plasticity; interpretations based on electrophysiology and neurophysiology; biochemical theories encompassing neurochemistry, including neuroendocrine phenomena; and other molecular rationales, including genetic and epigenetic mechanisms.

Among psychological theories, there is no evidence that fear, regression, or increased medical attention account for the efficacy or side effects of ECT. As many of the most recent advances in the field have reduced the adverse cognitive effects of ECT while maintaining efficacy, there is little reason to believe that memory loss has much mechanistic validity. Extensive reviews of the preclinical and clinical literature have refuted the assertion that brain damage occurs, let alone, accounts for ECT's effects. Furthermore, current anesthesia practice assures that hypoxia does not occur. Sham ECT (anesthesia alone) is ineffective; hence, anesthesia is unlikely to play a role in mechanism of action.

Distinguishing the relevance of observed phenomena with ECS constitutes a major challenge in the area of mechanisms. Considerations include questions about the relevance of animal data, much of which has been derived from rodents, although there has been recent interest in primate work. This work suffers from the current lack of adequate animal models of the psychiatric syndromes for which ECT is utilized. In preclinical work, ECS is often given in a time course that does not closely follow the treatment course given to patients, that is, two to three times weekly. For example, stimulation is administered once or several times in one day in preclinical investigations.

ECT has a very broad spectrum of efficacy and action, as will become evident from discussions of indications for its use, of efficacy in a variety of syndromes, and of its side effects. The treatment is often studied in comparison with antidepressants or to a lesser degree other psychotropics.

It is known that there is overlap in mechanistic effects, but ECT has more therapeutic effects than any studied medication. Coupled with the fact that ECT has more indications for treatment, the challenge of ascertaining critical similarities and differences is significant. Because of the perceived invasiveness of the treatment, the severity of illness among treated patients, and the necessity for anesthesia, little work is currently being done with sham ECT. There are limited data available from weaker forms of the treatment, for example, right unilateral ECT at low or modest dosage which is now known to be a less effective form of the treatment.

Biochemical Actions

There is a substantial body or work comparing ECS with psychotropic (and other) medications whose actions and mechanisms have been assayed in vitro and in vivo, most often preclinically, but also in human studies. Not surprisingly, the major areas of interest have been neurotransmitter and peptidergic systems. A primary focus has been on monoamine effects, many of which are shared by ECS in animals, ECT in humans and psychotropics, especially antidepressants. Although it appears that both ECS and antidepressants downregulate β-adrenergic receptors, human studies of broader changes in noradrenergic pathways have been characterized by inconsistent outcomes. There are also variable data on the effects of antidepressants and ECT on serotonergic systems: Antidepressants typically downregulate 5-hydroxytryptamine receptors ($5HT_2$), whereas ECT leads to increased density of $5HT_2$ receptors (in rodents) and decreased binding (in primates). Although serotonergic transmission may be enhanced by both antidepressants and ECT, the mechanisms appear distinct. ECT is more effective than antidepressant medications, so defining what possible mechanisms are responsible for this phenomenon are of pharmaceutical as well as heuristic interest. Currently, the critical differences are not known. Both animal and clinical studies indicate increased dopaminergic functioning following ECS, particularly D1 and D3 receptor function. Although this effect may be an underlying mechanism for ECT's therapeutic effects in Parkinson disease, it does not help clarify the underlying process for the benefits conferred by ECT on patients with schizophrenia and other psychotic illnesses, where dopaminergic antagonism, at least in some brain regions, has been presumed to be critical.

Another major area of interest has been the variety of different neurotransmitter systems, which may form the biochemical foundation for the anticonvulsant effects of ECS. Because of the efficacy of anticonvulsant medications in mania, there is also speculation that the anticonvulsant effects of ECT contribute to its efficacy in mania. Most of this work is preclinical, and indicates mixed GABAergic activity, upregulation of

adenosine receptors, and blocking of anticonvulsant effects with opiate antagonists. Currently, GABAergic effects in humans are being explored neurophysiologically. It has been difficult to duplicate anticonvulsant effects of opiate antagonists in humans. The proconvulsant effects of caffeine, aminophylline, and similar compounds are observable in humans and are thought to be due at least in part to alterations in adenosine transmission, even though these agents have other biochemical effects which may also contribute. With increased interest in the exploration of the glutaminergic system in the etiology and possible treatment for depression, preclinical studies demonstrate increased glutamate levels in hippocampus after ECS in a depressive rat model. Unfortunately, it is not currently possible to conclude that alterations in any transmitter system are associated with the efficacy of ECT in any syndrome for which it is used.

Most recent preclinical research focuses on growth factors and molecular mechanisms. It appears that BDNF and NGF are both increased by a single ECS and chronic ECS, whereas neurotrophin is decreased. There is interest in linking these observations to those made of marked neural plasticity in some brain areas, for example, hippocampus, due to ECS, as will be discussed below. Work continues with a wide variety of other neuropeptides, for example, somatostatin, CCK, precursors of opiates and thyroid hormones, etc. Other areas of interest concern the effects of ECS on second messenger systems, genetic expression, and transcription regulation. The work, focused largely on molecular phenomena related to cell proliferation—both neuronal and vascular, plasticity, and resiliency—including antiapoptotic effects, is expanding steadily, but its clinical and functional significance remains to be defined.

Hypotheses related to neuroendocrine mechanisms were generated early in the history of the use of ECT and have persisted to the present. Most hypotheses reason that stimulation of diencephalic structures, especially the hypothalamus, is involved and perhaps critical to the efficacy of the treatment. Experimental designs generated by these hypotheses have involved indirect quantification of function, specifically various peripheral neuroendocrine measures. Examples include CRH, ACTH and cortisol, both acutely and subacutely; studies of oxytocin and vasopressin; measures related to growth and thyroid hormones. Currently, there is no consistent evidence that acute or subacute changes in neuroendocrine factors impact efficacy or cognitive effects. Similarly, the involvement of immune system in the therapeutic effect of ECT has also been investigated. The available data suggest that a single session of ECT induces an acute, transient immune activation as measured by cellular immune function (lymphocyte subpopulations and NK cell activity) and levels of circulating cytokines, whereas repetitive ECT treatment results in long-term downregulation of immune activation. More prospective, controlled studies with larger sample sizes are needed to confirm this finding.

Neurophysiology

With advances in the area of brain imaging, neurophysiology has become a focus of interest in investigating the mechanisms of action of ECT. It is known that large increases in global cerebral blood flow (CBF) and cerebral metabolic rate (CMR) occur along with increased blood–brain barrier permeability during a generalized seizure, including those electrically induced. In the postictal period there is functional suppression with decreases in CBF and CMR as well as varying degrees of bioelectric suppression on EEG. Changes are also topographically distributed. These reductions have been correlated with treatment outcome.

The magnitude of reductions in CBF in specific prefrontal regions is related to efficacy of ECT, and responders are more likely to exhibit postictal suppression and develop slow wave activity on EEG than nonresponders. As techniques have been refined, the possibility of identifying specific functional networks and associating them to efficacy has improved. Related work in efield modeling has yielded information concerning the way seizures are generated with various electrode placement and electrical dosing techniques as well. For example, studies comparing effective forms of bifrontotemporal, bifrontal, and right unilateral electrode placements indicate that the prefrontal cortex, particularly on the right, the anterior cingulum and associated subcortical areas including the right medial thalamus, brainstem, and midbrain tegmentum are involved. Ineffective forms of ECT, for example, low dose right unilateral, trigger seizures from the motor cortex, and involve the prefrontal area less. Positron emission scan (PET) investigations point to these areas as well, particularly Brodmann area 25, the anterior cingulum, and the thalamus. These observations are important to clarifying which brain regions are critical to the efficacy of the treatment, as well as to understanding the dose–response relationship of electrical stimulation above seizure threshold to efficacy that has been observed for right unilateral ECT. The examination of the functional connectivity using functional MRI before and after bilateral ECT treatment revealed a significant increase in and around the left dorsolateral prefrontal cortical region (Brodmann areas 44, 45, and 46), where the average global functional connectivity was considerably decreased after ECT treatment, corresponding to the improvement in depressive symptoms. There has been expanding work using brain imaging technologies to investigate networks related to the cognitive effects of the treatment as well. Although preliminary, this work confirms the involvement of the medial temporal lobe in anterograde amnesia and points to prefrontal cortex involvement in retrograde amnesia.

Although the neurophysiologic inquiries are continuing, they are hindered by the limited localizing value and coarse resolution of currently

available methods. The degree of covariance among various regions and the spatial distribution and magnitude of association are yet to be defined. There is also a need to better characterize topographies and to determine whether normalization, abnormalization, or independence of observations applies. Currently, investigational samples are small and heterogeneous as well.

Genetic and Epigenetic Mechanisms

In addition to focus on cellular level changes relevant to the antidepressant effect of ECS, advances in genomics have provided new insights about the gene expression and epigenetic changes associated with ECS. The first studies of ECS-induced gene expression were published in the early 1990s. Further studies have revealed regulation of distinct neurotrophic signaling pathways (BDNF, Neuropeptide Y, vascular endothelial growth factor, fibroblast growth factor), synaptic vesicle proteins (synaptosomal-associated protein, *Syt 3 and Syn 1*) and immediate early genes (*c-fos, Egr 1, and Nrn 1*). BDNF has received extensive attention because of its involvement in the etiology of depression and response to antidepressant treatment. It has been shown that the BDNF expression in the hippocampi is elevated immediately (in the few hours) after acute ECS and decreased afterward in a temporal dependent manner, while BDNF expression was decreased in the ventral tegmental area. The expression patterns highlight the widespread action of ECS and the spatial and temporal diversity of expression changes for each gene type.

A few studies also attempted to examine possible epigenetic mechanisms, that is, the process of modifying gene expression independent from the primary DNA sequences in the therapeutic effect of ECT. Evidence indicates a strong effect of ECS on the different epigenetic mechanisms in the rodent brain. One of the earliest studies by Tsankova et al. in 2004, reported histone modification at gene promoter regions in hippocampi of rats undergoing acute and chronic ECS in response to electrical stimulation. Increased levels of histone H4 acetylation at the promoters of the *c-fos, BDNF,* and *CREB* genes were observed after acute and chronic seizures, which, in general, corresponded with increased mRNA levels of these genes. Demethylation at the *BDNF-IX* and *FGF-IB* promoter regions, increased DNA methylation at the *Arc* promoter, and demethylation at the *c-fos* promoter region were also reported. Generalizing this work to clinical populations remains to be carried out.

Neuroplasticity

Recently, there has been increased interest in structural changes in the brain associated with psychiatric syndromes and response to treatment. This has been particularly so for microscopic changes associated with ECS,

as well as, antidepressant and other medications. It is observed that the delayed behavioral response to common antidepressant medications in rodents coincides with the delayed upregulation of neurogenesis, at approximately 3 weeks. Similarly, the rapid antidepressive action of electroconvulsive shock is associated with a rapid stimulation of neurogenesis, producing immature cells demonstrated to be highly excitable compared to mature neurons. Synaptic plasticity, in addition to cellular plasticity (neurogenesis), for example, sprouting of mossy fibers, alteration of cytoskeletal structure, increased connectivity in performant pathways have been observed in animals after ECS. In addition to neurogenesis after ECS, microenvironment changes have been observed: for example, upregulation of activated microglia cells, the immune cells of the CNS, and increased proliferation of astrocytes, angiogenesis in certain layers of hippocampus, which may potentially modulate neurogenesis process after ECS. These reports have also galvanized controversy over various aspects of technical validity of the observations. It is unknown whether such changes occur clinically, and, if they do, what significance to efficacy and cognitive side effects might be discovered.

INDICATIONS

Although, historically, the first diagnostic indication for ECT was schizophrenia, practice has evolved in the past 75 years. Approximately 85 percent of patients receiving ECT in the United States currently have major depression as the diagnostic indication for its use. The remainder have schizoaffective disorders, mania, and schizophrenia with a few other diagnoses such as Parkinson disease completing the inventory of indications. The clinical literature demonstrating the short-term efficacy of ECT in the major syndromes for which it is used is considerable, and includes randomized controlled trials with sham treatment or with medications as alternatives, outcome data from studies involving various treatment techniques, as well as uncontrolled data and surveys of expert opinion.

Beyond diagnostic indications, there are other clinical considerations commonly brought to bear in the prescription of ECT. Historically, ECT was a first-line treatment for the major psychiatric syndromes, but, with the development of psychopharmacologic treatments, use as a primary treatment has decreased, and resistance to medications has become the most common clinical indication. None the less, when patients require rapid clinical response because of the severity and danger of psychiatric symptoms or medical problems resulting from psychiatric causes, ECT may be given without a prior drug trial. This is also the case when there has been prior optimal response with ECT or the patient prefers the treatment over medications (Table 34.35–2).

Table 34.35–2.
Indications for Use of Electroconvulsive Therapy

Diagnoses for which ECT may be indicated
Major diagnostic indications
Major depression, both unipolar and bipolar
Psychotic depression, in particular
Mania, including mixed episodes
Schizophrenia with acute exacerbation
Catatonic subtype, particularly
Schizoaffective disorder
Other diagnostic indications
Parkinson disease
Neuroleptic malignant disorder
Clinical Indications
Primary Use
Rapid definitive response required on medical or psychiatric grounds
Risks of alternative treatments outweighs benefits
Past history of poor response to psychotropics or good response to ECT
Patient preference
Secondary Use
Failure to respond to pharmacotherapy in the current episode
Intolerance of pharmacotherapy the current episode
Rapid definitive response necessitated by deterioration of patient's condition

Mood Disorders

Major Depression. ECT is the most effective treatment for major depression currently available. It is commonly used as the standard against which other treatments, including innovative treatments such as repetitive transcranial magnetic stimulation, are compared. No treatment has ever been found to be superior to ECT in the treatment of major depression in a controlled trial.

Evidence for ECT's efficacy derives from several different sources. The introduction of the treatment occurred prior to the development of clinical trials methodology in the post war era. In the 1940s, open clinical series and case studies cited response rates of 80 to 90 percent with the use of ECT as a first-line treatment. Decreased chronicity of illness, decreased morbidity, and the suggestion of decreased mortality were frequently cited as benefits of ECT. From the 1950s through the 1980s, ECT was subjected to the randomized clinical trials with sham ECT and psychotropic medications as comparators. In the sham versus real ECT trials, anesthesia alone was used for comparison groups because of a concern that there were benefits from its repeated use. ECT was shown to be consistently superior, and the further use of sham conditions has been abandoned in studies of major depression. From the 1950s onward, there was also interest in the relative efficacy of antidepressants and ECT. In these reports ECT was found to be 20 to 40 percent more effective than medications, and there has not been a study that found medication to be superior to ECT. Most of

this work was conducted in the earlier decades of this period, and is limited by the possibility that modern pharmacologic standards would dictate more aggressive dosage and duration of treatment in the medication conditions. In this type of study, adequate blinding is not possible, and concerns about bias cannot be completely dispelled even if clinical raters do not know the treatment conditions. Many of the antidepressants investigated in the past have been superceded by newer compounds, such as the specific serotonin reuptake inhibitors, and pharmacologically unique agents such as bupropion and mirtazapine. Unfortunately, there is little data to inform comparisons of ECT and modern pharmaceuticals.

Although in an earlier era prior to modern advances in defining the phenomenology of depression, it was thought that various clinical criteria conferred a negative prognosis to the use of ECT, more recent observations have found ECT to be effective for all subtypes of major depression including unipolar and bipolar depression, melancholia, and psychotic depression. There have been suggestive data that psychotically depressed patients have a superior response to ECT, whereas there is suboptimal response with antidepressants alone and the recommended regimens of adequate doses of both an antidepressant and an antipsychotic are seldom given, perhaps due to poor tolerance of the combination. ECT may be particularly useful for psychotically depressed patients because the syndrome is often accompanied by severe symptoms and greater risk of suicide.

ECT can also be used safely in all populations of patients, including those often considered to be at special risk: the elderly, the medically ill, pregnant women, and adolescents. Each of these groups may require specific considerations and extra preparation and/or precautions to be taken prior to treatment. ECT has even been used to treat preadolescent children, although rarely.

Dilemmas persist among experts in the field about when to offer ECT in the course of treatment for major depression. The hierarchy of factors that need to be considered has not been resolved, but it is important to gauge the risk of not terminating an episode of depression while giving sequential trials of antidepressants along with augmentation and combination medications against the risk of the principal side effects of ECT, namely persistent retrograde memory deficits of varying significance. With recent improvements in technique, these cognitive effects have been reduced to the point where, they can be difficult to detect. Even with these improvements, an acute course of ECT will often disrupt the routines of daily life because of the recovery requirements accompanying general anesthesia. Another important consideration relates to the fact that major depression is a subacute and often chronic disease with a need for continuation and maintenance treatment. For ECT, acute response is followed by the need to determine adequate maintenance treatments, most

often pharmacologic, although maintenance ECT is also an option. This circumstance can also affect the timing of the choice of ECT versus pharmacology.

Finally, the response rates cited early in the history of ECT must be adjusted to the changing make-up of patient populations receiving the treatment. It appears that prior medication nonresponse within an episode, extended lengths of episode, and comorbidity with other psychiatric conditions may lower the expected response rate. Comorbidity with personality disorder, that may independently contribute to mood changes, can be especially problematic since patients with personality disorder can also suffer major depression, and it is difficult to define what phenomenology belongs to each syndrome and to determine expectations. This does not automatically mean that ECT should be withheld from these populations. Response rates to alternative treatments may be even lower, while response to ECT in antidepressant nonresponders can be expected to be 50 to 70 percent.

Mania; Mixed States and Rapid Cycling. Historically, ECT was a first-line treatment for mania as well, but pharmacologic strategies, particularly lithium, anticonvulsants, and atypical neuroleptics, have become the mainstay of treatment for the manic episodes. In the modern era, the use of ECT is generally reserved for patients who are resistant or intolerant to the usual medication treatments, including those with mixed states or rapid cycling, or who have very severe symptomatology, for example, manic delirium. Organizations issuing guidelines for the use of ECT to treat mania have assigned it to second-line (WFSBP; APA; CANMAT) or third-line status (NICE) for acute episodes; only APA recommended its use as a maintenance treatment for bipolar disorder.

Compilations of the evidence for the use of ECT in mania have reported 75 to 80 percent response rates. The bulk of the evidence comprises uncontrolled case series, undoubtedly because of the great difficulty in conducting usual clinical trials in this population, due to uncontrolled behavior, difficulties in assuring capacity to consent, etc. Sham versus real ECT studies, with patients on neuroleptics in both conditions, demonstrated superior outcomes for ECT patients both in response rates and in lower neuroleptic requirements. Controlled retrospective comparisons have found ECT to be equivalent to lithium, and at least as efficacious as typical neuroleptics, if not more so. Prospective work demonstrated equivalence to lithium in studies where all patients received neuroleptics as well as the study treatments. One small study comparing ECT to lithium and neuroleptic in patients who failed lithium or neuroleptic found ECT to be noticeably superior to combined pharmacology. There have been no comparisons of ECT to anticonvulsants. Early work that suggested the need for increased frequency in treatment

schedule and increased numbers of sessions to complete a treatment course has not been born out in modern studies, which have used standard treatment schedules.

A prospective trial and a retrospective controlled review have found that patients in mixed states are responsive to ECT, although increased numbers of treatments may be required compared to pure manic or depressed states. Case reports have indicated that ECT is also potentially effective in managing rapid cycling, where finding effective pharmacologic approaches can be particularly complex and difficult. Response rates are lower than for acute manic states.

Schizophrenia

Although schizophrenia is the second most common diagnostic indication for ECT, there is a lack of expert consensus about its use and wide disparity in practice. This may be due to the current understanding of schizophrenia as a chronic illness and to the scarcity of evidence supporting the efficacy of ECT for chronic symptoms. Especially since the advent of atypical antipsychotic medications, pharmacologic treatment remains the first choice.

ECT was introduced for use in schizophrenia and the earliest reports, consisting of case series and comparisons with historical cohorts and nonsomatic treatments, were quite positive. As with other observations from this era, there was a lack of scientific methodology in diagnostic and outcome criteria, which would not conform to modern standards. Real versus sham ECT studies were carried out with chronic patients in the 1950s to 1960s, and more acutely ill samples on neuroleptics in the 1980s. The earlier studies showed no advantage for ECT over anesthesia, but the later studies showed ECT to be superior to anesthesia alone. The source of the disparate outcomes is not clear, but, in addition to differences in chronicity, the presence of neuroleptics in the modern samples may be critical. Because of the positive reports of efficacy in the first decades of use, ECT was also used as the standard against which to compare newer somatic therapies, particularly neuroleptics when they were first introduced. Although subject to the usual limitations of such early clinical trials, it soon became apparent that, at best, ECT was equivalent to neuroleptics in short-term efficacy. At this point, pharmacology superceded ECT for the treatment of schizophrenia. One observation from this era has remained intriguing and might merit further consideration. In long-term follow-up, patients who had received ECT had superior outcomes.

Equal attention has been paid to examining combination of ECT with neuroleptics versus treatment with either modality alone. The combination treatment is generally superior to monotherapies but the evidence for such observations is considered to be less than definitive due to relatively small samples and disparity in methodology among the studies. It appears that

lower doses of neuroleptic are needed in combination treatments. A recent study of combination treatment versus monotherapy in maintenance treatment of schizophrenia again showed promise, particularly in the first weeks of treatment. Although few of these studies reached full standards for a clinical trial, lacking random assignment and full blinding, recent reviews suggest that ECT may be useful for treatment-resistant patients and those who easily relapse on monotherapy. Thus far, most of the published observations on the use of ECT in schizophrenics resistant to neuroleptics are somewhat impressionistic as well. Recent reports confirm the safety of using ECT and atypical antipsychotics, with the clozapine as the oldest and most reported agent. Most professional bodies issuing guidelines for the treatment of schizophrenia reserve ECT for patients who are nonresponsive or intolerant to clozapine.

Since ECT does not have a broader therapeutic spectrum than neuroleptics, it is not surprising that good prognosis patients for pharmacologic treatment are also those that do best with ECT. Evidence to support the traditional idea that the presence of affective symptoms improves the likelihood of response to ECT is inconsistent, and there is no controlled evidence concerning its use for schizoaffective patients. Uncontrolled observational evidence suggests efficacy of ECT in schizoaffective illness, but response falls short of that obtained in straight-forward affective disorder, and resembles response rates obtained for schizophrenia.

One last area of interest is the impact of ECT on the movement disorders resulting from neuroleptic treatment of schizophrenia. Extrapyramidal symptoms induced by neuroleptics are reversible by pharmacologic intervention, but also respond to ECT. Furthermore, there is evidence from several case series that ECT may protect against the development of tardive dyskinesia in neuroleptic treated patients, independent of the observation that combination treatments with ECT and neuroleptic require lower doses of neuroleptic. This observation is consistent with preclinical observations of the effects of ECS on the dopaminergic system and the effects of ECT on patients with Parkinson disease.

Other Diagnostic Indications

In general, ECT treatment for other indications is supported by case material only. The use of ECT will be rare and only when standard treatments cannot be tolerated or fail.

Psychiatric. There is only very limited information about the impact of ECT on other DSM-IV axis I disorders, and no evidence for clinically significant efficacy without the presence of another primary diagnostic indication. There has been interest in anxiety disorders because of the

overlap in symptomatology and pharmacologic therapeutics between affective and anxiety disorders, but the only published data concern obsessive compulsive disorder (OCD). Prospective work demonstrated acute response in OCD, but patients soon relapsed, and ECT is not recommended for this disorder. Because of the not infrequent presence of symptoms of panic in depressed patients undergoing ECT, the question of the impact of the treatment on these symptoms has been raised, but there are no studies of the effect of ECT on panic disorder. Comorbidity with posttraumatic stress disorder (PTSD) and depression has provided an opportunity to observe the effects of ECT on PTSD. A minority of patients with PTSD may be vulnerable to increased symptoms related to the syndrome, but there are also case series where PTSD improved independent of the response of depression to ECT. Even though dysthymia is on the spectrum with major depression, there is no demonstrated efficacy. Patients with double depression (dysthymia and major depression) have similar response rates to those with episodic major depression alone, but are left with higher levels of mild depressive symptoms at the end of treatment, suggesting that dysthymia is more difficult to ameliorate with ECT even in the setting of comorbid major depression. The delirium that can accompany withdrawal and toxic states induced by some substances, that is, alcohol, phencyclidine and stimulants, can be treated with ECT, but the underlying substance abuse syndrome is not significantly affected. This use of ECT is largely of theoretical interest.

Another area of concern is the issue of comorbidity with personality disorders. Here most of the work concerns cluster B disorders, especially borderline personality disorder. A secondary area of interest should be cluster A disorders because of their hypothesized overlap with axis I anxiety disorders. In general, comorbidity with one or more personality disorders is associated with poorer outcomes, especially higher relapse rates. The challenge for clinicians is to discriminate between affective symptoms that result from the personality disorder and those that constitute an ongoing affective episode. ECT is not considered to be a treatment for any personality disorder.

Neurologic. ECT has been used rarely for the treatment of several neurologic disorders, usually in situations where they are comorbid with a psychiatric disorder that is an indication for ECT. The most common of these is catatonia, a syndrome characterized by stupor, mutism and a variety of motor symptoms, including posturing, waxy flexibility, and catalepsy. Schizophrenia and bipolar disorder are the most frequent underlying diagnoses, but with the increased occurrence of autism in modern populations, this illness is becoming a more common diagnosis with catatonia. One of the more commonly reported uses of ECT in adolescence is for patients with developmental disorder accompanied by

pharmacologically refractory dangerous behaviors thought to be on the catatonic spectrum. Currently, benzodiazepines are the primary treatment for catatonia along with a primary pharmacologic approach to the comorbid psychiatric disorder. When catatonia is resistant to benzodiazepines, ECT may be used. The data supporting this approach are largely uncontrolled, but quite positive. The syndrome of lethal catatonia is characterized by stupor or excitement, hyperthermia, autonomic dysregulation, rigidity, and changes in consciousness. ECT can be quite effective and, in fact, lifesaving. Neuroleptic malignant syndrome is an iatrogenically induced form of lethal catatonia and is often treated with dopaminergic strategies as well as sedatives. Outcomes with ECT are equivalent to those obtained pharmacologically, but ECT has the advantage of being a treatment for the comorbid psychiatric syndrome for which neuroleptics must be discontinued. There is evidence that NMS is less frequent with atypical antipsychotics, and the use of ECT for NMS may become less necessary as practice changes.

Open- and sham-controlled data have demonstrated that the motor symptoms of Parkinson disease, especially rigidity and bradykinesia, are lessened by ECT, particularly in the "off" phase of treatment that characterizes the phase of the disease treated with L-dopa. These effects are of 4- to 6-week duration generally, but there are reports that maintenance ECT may be helpful as well. ECT is seldom used as a primary treatment for Parkinson disease, and is more frequently used to treat the comorbid depression that can occur with the disorder. The combination of ECT with levodopa/carbidopa may require decreasing medication dosage in order to stave off dyskinesia or delirium as complications of combined treatment. With the increasing use of clozaril and other atypical neuroleptics to manage the side effects of levodopa/carbidopa, the use of ECT for Parkinson disease has diminished.

The anticonvulsant effects of ECT are well characterized, and ECT has been used in epilepsy, although effects are time limited and ECT has no demonstrated superiority over anticonvulsant medications. When such patients are treated, including when they are treated for comorbid psychiatric conditions more commonly treated with ECT, dosage of anticonvulsant medications is usually not changed, except when therapeutic seizures cannot be elicited. Recent reviews have suggested that ECT might be of use in intractable seizures or status epilepticus, particularly if the condition is secondary to a transient cause such as viral encephalitis.

There are reports of lessening of pain syndromes when comorbid depression is treated with ECT. ECT has also been used to mitigate other movement disorders, for example, Tourette syndrome. Historically, there have been cases of delirium of both medical and psychiatric cause that respond to ECT, although the treatment is very, very seldom used for this

reason currently.

The remainder of ECT use in neurologic syndromes is as treatment for comorbid affective and, less commonly, psychotic syndromes. There are case reports and series of its use in post stroke depression and mood disorders accompanying dementia, mental retardation, even head trauma and brain tumors. In all of these diagnoses, the risk of cognitive side effects of ECT can be greater, and modifications of the treatment may be needed, for example, to avoid skull defects found with trauma or to minimize the ictal blood pressure changes for patients with modest, stable brain tumors.

PATIENT MANAGEMENT

Pretreatment Evaluation

Evaluation of patients for treatment with ECT should be done as close to the first treatment session as possible. Both psychiatric and medical assessments are needed in order to conduct a risk–benefit analysis for the advisability of ECT treatment of an individual patient. In general there are no absolute medical contraindications to ECT. Although cochlear implant is labeled as a contraindication, ECT has been reported to be administered safely to the contralateral side of the implant, and the implants themselves were not damaged by ipsilateral ECT administration in cadavers.

The psychiatric component of the pretreatment evaluation is generally conducted by a practitioner privileged to administer ECT, but there are often contributions from the patient's attending psychiatrist if different from the ECT provider. Establishing the indications for ECT will include a careful diagnosis as well as consideration of other clinical factors contributing to the recommendation for ECT, for example, severe suicidal risk. There may be local requirements and regulations for second opinions, especially in special populations such as adolescents and children.

Past treatment with ECT should be reviewed not only for response, but also to identify technical aspects and tolerance of the treatment. Knowledge of prior ECT treatments, including use of specific ECT records, will guide recommendations for the current course. Pharmacologic treatments should be evaluated for number, adequacy, tolerance, and degree of response. Psychotropic history can be an indication for ECT as well as a basis for recommendations concerning post-ECT continuation treatment. Decisions regarding the adjustment of psychotropic medications should be based on the degree to which they may enhance response, for example, maintaining an antipsychotic; the possibility that they may interfere with ECT, for example, anticonvulsants used as psychotropics; the possibility that medications may contribute to adverse effects, for example, lithium; and the need to maintain medications for patient management while ECT takes effect.

As part of the pretreatment evaluation there should be documentation

of the target symptoms to facilitate evaluations during the treatment course and decisions concerning the termination of acute treatment. In addition to behavioral assessments, cognitive evaluation pretreatment is also critical. At minimum, those aspects of cognitive function most likely to be affected should be documented at baseline: orientation and memory, including recall of recent and remote events. Evaluation of language and visual motor and perceptual abilities can also be valuable as well. Finally, patients' subjective assessment of their cognitive and memory functioning should also be documented, particularly since the predominant diagnostic indication for ECT, major depression, is often accompanied by both subjective and objective memory problems.

The medical component of the pre-ECT evaluation focuses on the patient's ability to tolerate anesthesia as well as ECT, and should be conducted by an ECT practitioner as well as an anesthetist. The general medical evaluation will be similar to that conducted for patients undergoing outpatient surgery, and will focus on, although not be limited to, cardiac (e.g., congestive heart failure, symptoms of coronary artery disease), pulmonary (e.g., lower respiratory tract infection), and neurologic systems (e.g., acute stroke, increased intracranial pressure), those most likely to be impacted by ECT. The patient's past experience with anesthesia should be elicited, as well as family history of capacity to tolerate anesthesia. A complete list of medications including herbal and over-the-counter medications as well as medication allergies is also needed. Most providers document a complete blood count and electrolytes as well as electrocardiogram. In addition, a chest x-ray, thyroid function tests, a complete metabolic panel, and a pregnancy test in women of childbearing age may be needed. A brief dental examination for loose or missing teeth and the presence of bridges and dentures is recommended. Testing of plasma pseudocholinesterase or dibucaine number is done as indicated by patient and family history. The risk of fracture of the long bones or spine has been averted by the use of muscle relaxants. Hence, spinal x-rays are done as indicated by the patient's history of skeletal injury or disease. Brain imaging is recommended when a questionable neurocognitive history, focal findings on neurologic or funduscopic examinations, or altered neurocognitive examinations are present. The basic medical evaluation should reveal whether there may be a need for further specialty consultations, testing, or labs. Many institutions involve an internist routinely, especially when conditions that can be affected by or interact with ECT, for example, diabetes, will need management during the course of treatment.

Taken as a whole, the pretreatment evaluation will result in recommendations for the modification of medications the patient may be receiving, both on a medical and a psychiatric basis; the modification of ECT technique; the modification of anesthesia technique; and the

management of both psychiatric and medical conditions during the ECT course.

Consent

Written informed consent is standard in the practice of ECT, and involves a discussion of the risks and benefits of the treatment with the patient and/or legally sanctioned surrogate. Procedures for consent are locally regulated, particularly at the state level, and familiarity with these regulations is essential. Consent can include both the anesthetic procedures and the electrical stimulation, although institutions may provide separate consents for each of these aspects of treatment. Consent forms generally cover several different areas and can be lengthier than the type of document used for other comparable medical or surgical procedures. Descriptions of the procedure and expected benefits are detailed. Risks of the treatment including possible cognitive and medical adverse effects, including death, should be documented. The possibility of relapse and nonresponse should be included. Expected participation by or behavioral restrictions of the patient, for example, NPO for several hours before treatment, are spelled out. Alternatives to treatment, including no treatment, should be discussed. In addition to a generic description, patients should understand the individual indications and rationale for ECT. They should know who can answer any questions about the treatment that may occur during the course of ECT. It should also be clear that consent may be withdrawn during the course of treatment. The amount of information given should be tailored to each patient, recognizing that too much information can obscure the process just as too little. Guidelines and sample consent documents are available from professional organizations.

Most patients referred for ECT have capacity to consent and can give it voluntarily. Generally, capacity is assumed to be present unless there is strong evidence to the contrary. Where doubt exists as to capacity, an independent consultation may be prudent. It is important to weigh the protection of autonomy and the opportunity to receive effective treatment in the consent process. While patients with capacity should not be coerced to receive treatment, seriously ill patients who lack capacity should not be unnecessarily obstructed from receiving it, especially when ECT is being recommended to prevent severe adverse outcomes such as increased morbidity or even death. Although regulations vary by jurisdiction, most allow for involuntary treatment with surrogate consent and/or legal review, which can include formal judicial review. Some jurisdictions mandate not only judicial review, but also review by the department of health for patients who lack capacity and object to ECT.

Surrogates must proceed through the consent process in a similar way to patients with capacity. Family members and other significant associates or companions are encouraged to become educated about ECT. Audiovisual

aids and written material may be provided to patients, family, and others as indicated. Finally, it should be remembered that consent is a process throughout the course of ECT, and clinicians should maintain an open exchange about procedures, response, side effects, and prognosis. Professional organizations recommend separate consent be obtained for continuation or maintenance ECT as the indication has changed from acute treatment to relapse or recurrence prevention, and the risk–benefit assessment may have changed as well. With the increased use of tapering an acute course of ECT to continuation treatments, many hospitals have instituted consents that include a specific number of ECT sessions and duration of consent, regardless of whether they may be classified as acute or continuation ECT.

Parameters of the ECT Stimulus

Controlled trials in ECT over the past two decades have involved measuring the impact of varying aspects of ECT technique, most frequently electrode placement and stimulus dosing strategies, on response rates and adverse cognitive effects. Other aspects of the ECT stimulus include direction and amplitude of the current, pulse width and pulse frequency. Research from a prior era had led to a belief that seizure elicitation was necessary and sufficient for efficacy, because subconvulsive stimuli were associated with lower response rates in several studies. When attenuation of electrically induced seizures with lidocaine was also associated with reduced efficacy, adequate seizure duration was also postulated to be necessary for therapeutic response. These tenets have been revised in light of observations from modern studies, and current practice has been directly impacted.

A series of studies since the 1980s demonstrated that, depending on the combination of electrode placement and stimulus dosage, seizures could vary in eliciting clinical response from 20 to 70 percent. Recent studies, utilizing an ultrabrief pulse stimulus configured the waveform in a more physiologic way in order to minimize excess electrical dosage, similar to the substitution of brief pulse waveform for the original sine waveform. Right unilateral electrode placements were particularly affected by dosage, and a dose–response relationship exists up to a stimulus intensity of 8 to 12 times seizure threshold. At electrical dosage six times seizure threshold, right unilateral ECT is equivalent to bilateral ECT in clinical efficacy, but maintains an advantage in producing fewer adverse cognitive effects. For electrical dosage above six times threshold, the advantage in acute cognitive effects may be eroded, but there is evidence that an advantage in long-term cognitive effects remains. These findings make dosage relations with seizure threshold particularly critical for right unilateral ECT. Ultrabrief right unilateral ECT at six times seizure threshold has been shown to be equally effective to both brief pulse bilateral ECT and brief

pulse right unilateral ECT; ultrabrief bilateral ECT at 2.5 times seizure threshold was an ineffective treatment in at least one randomized trial. Manipulation of waveform in this manner had even more powerful effects on cognitive outcomes than had electrode placement with results statistically indistinguishable from those obtained in cognitive testing of age, gender, and IQ matched medically and psychiatrically healthy controls on measures of retrograde amnesia. Combined with results of neurophysiologic and brain imaging studies of the effects of ECT, the clinical impact of the interaction of electrode placement and dosage also suggests that the antidepressant effects of ECT are dependent on the anatomic specificity of the effects of current density, rather than on seizure elicitation alone. It was also found that speed of clinical response can be enhanced as stimulus intensity was increased above seizure threshold for both unilateral and bilateral electrode placements.

Although there has long been a theoretical interest in greater localization of the effects produced by ECS in brain tissue, the neurophysiologic and clinical findings related to the interaction of electrode placement and stimulus intensity have rekindled interest in concentrating current density in prefrontal areas by manipulating electrode placement.

Electrode Placement. Historically, most practitioners have used bifrontotemporal electrode placement because of its reliability in producing efficacy and its ease of use. This electrode placement is also associated with more short- and long-term adverse cognitive effects, and is more likely to produce delirium, which may require interrupting a course of ECT and perhaps even terminating it before optimal therapeutic effects have been obtained. Hence, when bifrontotemporal ECT is used, attention should be paid to restricting dosage to a moderately suprathreshold to attenuate adverse cognitive effects as much as possible. It should be emphasized that the combination of ultrabrief pulse and bifrontotemporal electrode placement has not been demonstrated to be consistently effective. Treatment with bilateral electrode placements, particularly bifrontal configuration, is more likely to manifest EEG seizure without motor seizure, and EEG monitoring can be particularly useful in detecting the occurrence of a seizure at treatment.

Newer electrode placements include bifrontal configuration and asymmetric placements. There are limitations to these strategies, imposed by the fact that the high impedance of the skull and scalp causes spreading of the electrical stimulus and restricts possibilities for localization of the stimulus. Bifrontal electrode placement, with positioning far enough laterally to minimize interference with impedance relations, has been investigated and there have been several demonstrations that bifrontal electrode placements are equally effective to bifrontotemporal and

adequately dosed right unilateral electrode configurations. Evidence of advantages in sparing of cognitive effects is preliminary, and adequately powered investigations with more extensive and sensitive cognitive batteries are needed. Seizure threshold is likely to be relatively higher with bifrontal ECT.

The relatively better cognitive side-effect profile of right unilateral ECT should encourage wider utilization now that the efficacy of this electrode placement can be assured with adequate dosing strategies. In contrast to bilateral ECT, dosage closer to 500 percent above seizure threshold for right unilateral ECT is more likely to assure efficacy. ECT devices in the United States are restricted to an output in the range of 504 to 576 mC. Approximately 90 percent of patients have seizure thresholds which can accommodate optimal dosing with brief pulse right unilateral ECT, and the combination of right unilateral electrode placement with ultrabrief pulse width extends the range of US devices so that most patients can be treated within these constraints. Individuals with exceptionally high seizure threshold may require bilateral electrode placements to remain within the device restrictions. Maximizing interelectrode distance by using the d'Elia placement for right unilateral ECT may also be optimal. Many other right unilateral placements have been described, but there is little work to support their use (Fig 34.35–1).

There has been some concern that left-handed patients may require different electrode placement from right-handed patients, especially if unilateral placement is desired. Even when handedness is lateralized to the left, the anatomic localization of language function in 70 percent of left-handed individuals is the same as in those who are right-handed. Furthermore, there is evidence for independent lateralization of affect, with the right hemisphere involved in sustaining depressed mood regardless of handedness. In addition, no differences in efficacy have been found to be related to handedness when examined in studies. Because of limited indications that affective function and efficacy of ECT are associated with handedness, it is not generally used to guide the choice of electrode placement.

Stimulus Waveform. The original ECT devices used in the 1930s generated the electrical stimulus as a sine wave, with a rather wide pulse width, approximately 8.33 ms (60 Hz). Modification of the waveform to a bipolar rectangular pulse, 0.5 to 2.0 ms wide, represented an advance: diminishing nonessential electrical dose, dramatically reducing adverse cognitive effects, but preserving efficacy. Modern ECT devices retain this capability. However, neurophysiologic observations suggest that the optimal pulse width for neuronal depolarization is actually in the range of 0.1 to 0.2 ms. Thus, brief pulse width is inefficient and delivers unnecessary stimulation. While reductions of pulse width to this extent may not be

practical with current engineering of ECT devices, several reports, including randomized controlled trial evidence, employing a pulse width of 0.3 ms (known as “ultrabrief” in contrast to “brief”) have demonstrated that right unilateral ECT, typically at six times seizure threshold, is equivalent in efficacy to both right unilateral and bilateral ECT using brief pulse widths.

In one report, patients treated with the ultrabrief pulse width showed significantly less impairment in acute, subacute, and long-term cognitive side effects of anterograde and retrograde memory than patients receiving standard pulse width stimulation, for both right unilateral and bifrontotemporal ECT. In fact, those who received ultrabrief right unilateral ECT did not deteriorate in any cognitive measure relative to pre-ECT performance. A smaller randomized trial confirmed the equivalent efficacy of ultrabrief and brief pulse right unilateral ECT. A third report comparing ultrabrief right unilateral and ultrabrief bifrontal ECT at 1.5 times seizure threshold found no change or improvements in cognitive performance, while efficacy was similar for the two pulse widths. In effectiveness, open-label prospective and retrospective studies, findings have been more mixed although predominantly positive for the efficacy of ultrabrief right unilateral ECT and sometimes even ultrabrief bilateral ECT, but the lower quality of these designs must be factored into judging the outcomes presented.

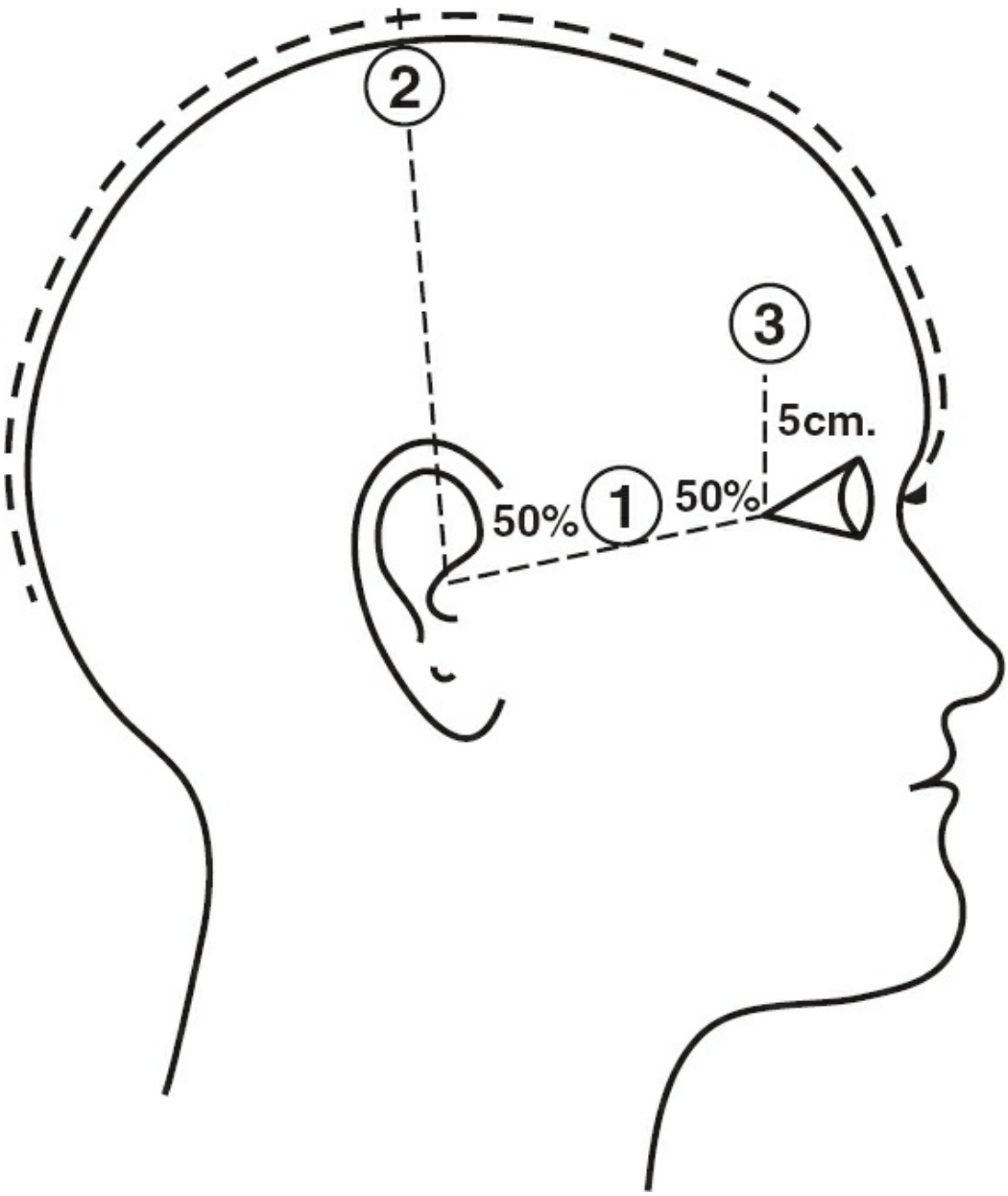

FIGURE 34.35–1. Electrode placements. Position #1 represents the frontotemporal position, used for both electrodes, one on each side of the head, in conducting traditional bilateral ECT. For right unilateral ECT, one electrode will be in the right frontotemporal position, and the other will be just to the right of the vertex at position #2. Position #3 represents the frontal placement, used for both electrodes in conducting bifrontal ECT.

Seizure Threshold and Dosing. As indicated above, the intensity of electrical stimulus, that is, the degree to which electrical stimulus exceeds seizure threshold, affects the efficacy of right unilateral ECT, the speed of clinical response regardless of electrode placement, and the cognitive side effects produced. The effects are in the same direction for each of these domains: the higher the stimulus dosage, the better the efficacy of right unilateral ECT up to 8 to 12 times seizure threshold, the quicker the clinical response by one or two treatments over a course, and the greater the cognitive side effects. These observations have led to the promulgation of

dosing methods to refine stimulus intensity.

Strategies for determining stimulus dose include dosing based on empirical determination of seizure threshold (titration), formula-based dosing, and fixed dosing. Empirical titration involves an approximation of seizure threshold in the first treatment session, followed by a calculation of optimal dosage based on the results obtained. The stimulus intensity estimated on this basis is then generally used for the remaining sessions. The first electrical stimulus in the titration is set low enough that a seizure results in only a minority of patients. Restimulations at intervals of 20 to 60 seconds are given in the same session using progressively greater amounts of charge until a seizure is produced.

The dosage at which a seizure is elicited is termed the seizure threshold, and has been found to vary at least 50-fold in patients undergoing ECT. Because of this wide variation in seizure threshold, dosing based on empirical titration is thought to be the most precise method of stimulus determination. Furthermore, it is safe for most patients when conducted with adequate physiologic monitoring and the use of anticholinergic agents as appropriate to prevent the brief periods of bradycardia that may accompany subconvulsive stimulation. In controlled trials where dosing was based on empirical titration, patients benefited clinically beginning with the first treatment session even though dosage was given near threshold, a theoretical concern with right unilateral electrode placement. Seizure threshold, as empirically determined, increases over the course of ECT, but returns to levels observed in the first treatment session within a week of ending treatment. Other measures of convulsant activity, such as seizure duration, may take up to 60 days to return to values observed in the first treatment session.

Formula-based dosing uses factors known to influence seizure threshold, such as age, electrode placement, stimulus pulse width, gender, and concomitant medications including anesthetics, to estimate adequate stimulus intensity. Seizure threshold is generally higher in men than women, with bilateral electrode placement than with unilateral electrode placement, and with increasing age. It is currently unresolved whether bifrontal electrode placement results in differences in seizure threshold when compared with bifrontotemporal placement. The basis for these observations has not been extensively investigated. The simplest formulae use age alone, but even the most detailed and elaborate formula can only account for about 40 percent of the variance in seizure threshold observed in patients. Hence, there has been limited success obtaining precision in seizure threshold approximation and dosing using formulae.

Fixed dosing occurs when the same dosage of electrical stimulus is applied to all patients receiving ECT. It is obviously the easiest technically and maximizes the chances of clinical response especially for right unilateral electrode placement. Unfortunately, this dosing strategy can

cause patients with the lowest seizure thresholds to receive up to 20 times an adequate electrical stimulus, especially with bilateral electrode placement. For low threshold patients, fixed dosing enhances the possibility of increased adverse cognitive side effects as a result. If fixed dosing is used with right unilateral electrode placement, the expected benefit in adverse cognitive effects could be lost. It is felt that fixed dosing may only be appropriate for patients with such serious medical conditions that even one subconvulsive stimulus, as is still a possibility even with formula-based dosing, must be assiduously avoided. Fixed dosing at the maximum output of the device should be avoided.

Seizure threshold and subsequent dosage determination can also be influenced by which stimulus parameters are adjusted. Comparisons of adjusting frequency versus stimulus train duration on brief pulse devices have indicated that longer duration is more efficient in eliciting seizures than increased frequency, although the effect sizes are small. Recent work has also shown that frequency does not exhibit a linear function in the depolarization, an observation that contributes to understanding this relationship. Efficiency in seizure elicitation can also be optimized by using narrower pulse widths in setting the parameters on brief pulse devices. This strategy has the additional effect of extending the range of devices which are capped in stimulus output. There has been recent interest in optimizing pulse amplitude as well with initial indications that further efficiencies may be obtained. Use of unidirectional stimuli is still in the early stages of investigation.

Anesthesia and Treatment Procedures. Since ECT is given under brief general anesthesia, the treatment is usually performed in a specific treatment area which generally meets standards for the performance of day surgery, including the capacity to manage medical emergencies and to recover patients postictally. The staffing commonly consists of a psychiatrist privileged to administer ECT, an anesthetist, and one or more nursing personnel. Although the anesthetist maintains responsibility for anesthesia procedures, particularly airway management, the treatment procedures are conducted collaboratively.

Before anesthesia is induced, appropriate physiologic monitoring is instituted and continued until the patient is adequately recovered. Cardiac monitoring and pulse oximetry are standard; monitoring end-tidal carbon dioxide is also standard at many institutions. Automated noninvasive blood pressure monitoring, set at a short frequency, is common. Monitoring of seizure activity can consist of observing both motor and EEG seizure duration. Motor seizure duration can be determined optimally using the "cuff method." A blood pressure cuff is applied to the forearm or just above the ankle and inflated to a level above the anticipated maximum systolic pressure (>250 mm Hg) to prevent anesthetics from affecting muscle

contraction in the distal limb. Motor movements resulting from seizure induction can then be directly observed. The cuff should be ipsilateral to electrode placement for unilateral ECT. Ictal movements may continue longer in the face and neck. Use of the cuff method may not be appropriate for patients with severe peripheral vascular conditions, especially in the lower limbs. Many ECT devices have incorporated technology to allow at least 2-lead EEG monitoring with a frontal–mastoid montage preferred. Care must be taken to prepare EEG sites adequately and to take into account artifacts, for example, ECG, movement.

Oxygenation is provided using 100 percent O_2 via mask, and administered under positive pressure when respirations have stopped secondary to anesthesia medications. Respiratory support is discontinued when there is return of spontaneous respiration adequate to maintain clinically acceptable hemoglobin saturation on room air. Preoxygenation may begin even earlier in the sequence of treatment procedures when medical conditions such as obesity or pulmonary disease pose a risk to respiratory function. In rare circumstances, such as third trimester pregnancy, patients may be intubated.

There is variation in the use of anesthetic medications. Not all practitioners employ anticholinergic medications pre-ECT, and, unfortunately, controlled studies comparing ECT with and without these agents did not include those patients at greatest risk for cardiovascular conditions. Anticholinergic medications such as atropine and glycopyrrolate are used to dry secretions that may interfere with respiration under anesthesia and to decrease the vagally mediated bradycardia and even asystole that may occur immediately after the electrical stimulus is administered. The cardiac effects of anticholinergic agents are most useful in the following clinical situations: during seizure threshold titration, when subconvulsive stimuli will not be accompanied by compensatory catecholaminergic output; in the presence of agents that promote bradycardia, for example, β-blockers; when cardiac conditions characterized by some degree of electrical blockade are pre-existent. Anticholinergic agents may be contraindicated when they might provoke pre-existing arrhythmias or threaten cardiac compensation through increases in workload from higher heart rate. Although glycopyrrolate does not cross the blood–brain barrier under normal circumstances and could have an advantage over atropine, controlled comparisons with atropine have not detected differences in cognitive adverse effects or postictal nausea, perhaps because ECT disrupts the blood–brain barrier and diminishes this advantage.

Ultrabrief general anesthetics are used to induce unconsciousness while muscle relaxants are in effect and during seizure elicitation. Methohexital, a short-acting barbiturate, has been the preferred agent in the United States. Alternative anesthetics include another short-acting barbiturate,

thiopental; and short-acting agents from a variety of classes, including propofol, etomidate, and ketamine. Thiopental has been associated with a higher incidence of postictal arrhythmias compared to methohexital, and its pharmacologic properties can be associated with long recovery periods when higher doses are used. Propofol reduces hemodynamic changes accompanying ECT and typically produces a rapid, smooth recovery. Its antiemetic properties can be useful when patients develop nausea or vomiting. This agent also attenuates seizure duration, but studies comparing it to more standard agents have not detected differences in clinical efficacy or cognitive recovery. Agents that attenuate seizure duration, generally are associated with slightly shorter emergence and recovery although a recent review concluded that the differences while statistically significant may not be of sufficient magnitude to affect the choice of agent. Numerous reports have explored the addition of short-acting opioids, for example, remifentanil, to standard agents. With remifentanil, less of the standard agent is required and seizure durations can be longer. Etomidate does not affect seizure threshold and is associated with fewer cardiac side effects, especially when compared to propofol. Etomidate can also produce myoclonus, and repeated use in animals has been associated with reductions in adrenal function. Ketamine, an NMDA receptor antagonist has been associated with longer seizures in animals given ECS. Given the potential involvement of glutamatergic system in the etiology of major depression and schizophrenia, as well as the potent independent antidepressant effect of ketamine, the theoretically neuroprotective effect against cognitive impairment associated with ECT, the use of ketamine in ECT treatment has resurged. Insufficient study has been conducted to make firm conclusions regarding the (1) possible synergistic antidepressant effect of combination of low dose of ketamine and ECT; (2) additional therapeutic effects and cognitive protective effects of using ketamine as first-line anesthetic induction agent in ECT treatment. Unfortunately, ketamine's sympathomimetic properties can result in hallucination and unusual states of consciousness as well as increases in heart rate and blood pressure. Until recently, both etomidate and ketamine have been reserved as alternative or adjunctive agents for situations where seizure elicitation is difficult. Greater attention to adequate dosing of these anesthetic agents has become increasingly important as the increased use of cognitively sparing techniques, such as ultrabrief right unilateral ECT have been accompanied by reports of patients regaining awareness and remembering the effects of succinylcholine before full respiratory capacity returns to baseline.

Before the development of muscle relaxants, musculoskeletal injury, particularly fracture of the long bones or vertebrae, was not uncommon, as it is today. Curare was developed to reduce the tonic–clonic movements produced by electrical induction of a generalized seizure, but its use has

been superceded by succinylcholine, a depolarizing agent with rapid onset and brief duration. Myalgia can be a side effect of succinylcholine. The standard treatment of these myalgias, which are often present in only the first few treatments, is the use of common analgesics. Although modification of movement is the usual goal of using muscle relaxants, complete paralysis may be desirable when there is a history of skeletal disease, such as osteoporosis or recent fracture. When the use of succinylcholine is contraindicated, as in pseudocholinesterase deficiency, nondepolarizing agents can be employed. Agents available in the US market are rather longer acting; even when reversal is produced with an anticholinesterase such as neostigmine, it may be at least an hour before recovery is sufficient for transfer from the ECT treatment area to a recovery area is safe.

Although complications at ECT have involved the cardiovascular system most commonly, there is controversy regarding the routine use of agents, such as β-blockers, which reduce the increases in hemodynamic activity that accompany ECT. Controlled studies of routine use do not demonstrate advantages for active intervention, perhaps because the patients who benefit are a minority of the entire population of ECT recipients, most of whom easily tolerate the observed increases in heart rate and blood pressure for a few minutes. There has also been some concern that the hypermetabolic cerebral state induced by the treatment requires a corresponding increase in cardiovascular functional level to maintain adequate oxygenation and nutrient supply to the CNS. The conservative position is to restrict the use of pharmacologic intervention to those conditions where the need is unambiguous, for example, aneurysm, sustained hypertension, and/or tachycardia.

Because the application of the electrical stimulus directly stimulates contraction of the jaw muscles, a bite block is placed in the mouth just before the delivery of the electrical stimulus to protect the teeth and oral soft tissues. The duration of the stimulus can be several seconds, but typically is only a few. The duration of the seizure is commonly less than 1 minute. The relationship of duration to efficacy has been questioned because of recent findings, and it is clear that duration alone is insufficient to insure clinical adequacy. Seizures of less than 15 seconds can occur close to seizure threshold and are often considered abortive. Seizures of similar duration can also be observed in the elderly, late in a treatment course, and with markedly suprathreshold dosage, but such short-duration seizures may not be associated with lack of adequacy. These circumstances highlight the usefulness of empirical titration techniques to inform the practitioner more exactly about seizure threshold and dosage, and to guide what actions, if any, to take when short seizures occur. Seizures longer than 3 minutes, with either motor or EEG monitoring, are considered prolonged and should be actively controlled using intravenous anticonvulsants, such

as the short-acting barbiturate used to induce anesthesia or benzodiazepines. The occurrence of long-duration seizures is unusual, except in the youngest patients, and the cause of such long duration should be investigated, especially if it was not easily terminated.

Patients are recovered with monitoring of cardiovascular and respiratory functioning until physiologically stable. If not awake and oriented at that time, further recovery is needed until the patient is alert and reoriented. Many facilities require that the patient be fully ambulatory before leaving the treatment and recovery areas. Bed rest may be needed for a period of time. Side effects, such as headache and nausea, are monitored and treated as needed. Headache frequently responds to analgesics, especially COX inhibitors; patients with migraine may require more specific antimigraine treatments. Postictal delirium with agitation, disorientation, and lack of full consciousness, occurs in a minority of patients. Decreasing environmental stimulation and securing protection may be adequate management, but not infrequently, patients will require intravenous pharmacologic intervention with sedatives and prophylaxis at subsequent treatments.

Treatment Course. Typically, ECT is given two to three times weekly on nonconsecutive days. The latter frequency is common in the United States. Twice weekly treatment is equally effective compared to thrice weekly treatment, can take longer to reach efficacy, and is associated with fewer acute cognitive side effects. Other frequencies of treatment are more problematic. Daily treatment is seldom done now, but, at one time, was believed to speed efficacy dramatically and was used in the most urgent situations, for example, severe mania. This treatment schedule was also characterized by a risk of prolonged severe cognitive side effects, especially when a course known as regressive ECT was used. In addition to being logistically difficult to obtain in many places, the risks appear to outweigh the benefits, and daily treatment is no longer advised. Another approach to increasing frequency was eliciting more than one generalized seizure per session under continuous anesthesia, known as multiple-monitored ECT (MMECT). Similar to daily ECT, the goal was to accelerate clinical response with the additional advantage of decreasing the number of times anesthesia is delivered. Although there have been both reports of uncontrolled observations and one controlled trial concerning MMECT, the literature is clear that severe adverse events accompany this treatment approach with only modest gains in efficiency of response. Of most concern are reports of cardiovascular morbidity and mortality, as well as serious adverse cognitive effects including prolonged confusion. Although there may be a place for the use of MMECT, it is only in extremely unusual circumstances, for example, intractable seizures, and should be restricted to double stimuli/session. At the other end of frequency, once weekly ECT is

associated with weaker clinical response and more suitable for continuation ECT schedules where the goal is to maintain a patient's status as responder or remitter—as opposed to acute treatment, which is used to induce response.

The number of treatment sessions in a course of ECT should be individualized, as patients vary in this requirement from a few sessions to more than 15, especially when changes in treatment technique have been made. The course length cannot be predicted, although the average is from 6 to 12 sessions. Fixed course lengths are not appropriate for these reasons. Patients should be evaluated every one to two treatments for clinical response and adverse side effects, particularly cognitive effects. These assessments are best done between treatment days when the acute effects of anesthesia and ECT will interfere less with evaluation. Acute treatment should end when there is remission of target symptoms or when further improvement has not occurred over 2 to 3 sessions with optimal technique.

When response is inadequate, modifications can be considered. Such modifications can include intensification of stimulus energy by increasing charge through manipulations of the various parameters of the stimulus or by changing electrode placement. Seizure threshold increases over the course of treatment, so higher stimulus intensity may be useful later in a treatment course if response is inadequate, particularly if response was more robust earlier in the course but waned. If any medications with anticonvulsant properties, including barbiturate anesthetics or propofol, are being used, discontinuation or decreasing dosage should be evaluated, particularly when the patient has high seizure threshold. At least 10 treatments with optimal electrode placement and dosing should be given before nonresponse is declared. When course length exceeds 15 to 20 sessions, especially with modifications of technique, reconsideration of the use of ECT, with careful balancing of the risk of an unresolved episode of illness against the risk of accumulating adverse effects, should be undertaken. Additional consultation, with reassessment of primary and comorbid diagnoses, may be advisable.

ECT is the only treatment for major psychiatric syndromes that is discontinued when remission occurs. Hence it is not surprising that a significant proportion of patients who relapse do so within the first few weeks of termination of ECT, and then require further ECT. Often additional ECT is given along with medications that will be used for continuation and maintenance of remission. Although not yet examined in a controlled trial, tapering the termination of ECT, by decreasing frequency of sessions while medications are stabilized and produce their effects, may be a more effective strategy for relapse prevention.

Continuation Treatment. Continuation treatment, that is, treatment for 6 months beyond remission of an acute episode of illness to prevent

relapse, is standard practice for the major syndromes that are somatically treated. It is recommended and routine post-ECT as well. Populations that receive ECT are particularly likely to be characterized by risk factors for relapse: medication resistance and the presence of psychotic features. At least half and upward of 85 percent of ECT responders in some samples will relapse without continuation treatment particularly in the first several weeks post treatment. Aggressive continuation treatment should be initiated as soon as remission is evident, although initiation of pharmacology during the ECT course itself has not been shown to be of additional benefit for depression.

Continuation treatment with medications is the predominant practice post-ECT. This has been based on studies conducted in the 1960s in Britain and on clinical experience. The class of agent advised matched that used in the absence of ECT: antidepressants for unipolar depression, antidepressants and mood stabilizers for bipolar depression, mood stabilizers with or without antipsychotics for mania, and antipsychotics for schizophrenia. Studies in the 1970s found little support for the use of lithium alone as relapse prevention following depression, but some support for lithium as maintenance treatment after 6 months of remission to prevent recurrence once remission is firmly established. The first randomized prospective controlled study of continuation treatment post-ECT in unipolar patients confirmed substantial relapse rates on placebo (84 percent), but only moderate improvement in relapse rate on the antidepressant nortriptyline alone (60 percent). The relapse rate in unipolar depression was significantly reduced in this study by the addition of lithium to nortriptyline (39 percent) with most of the relapse in this arm of the study occurred in the first 5 weeks. Randomization had been stratified by medication resistance and (mood congruent) psychosis, but mood congruent psychotic patients did not relapse more frequently. Medication dosage was in the range used for acute treatment. A prospective trial of combination antidepressant and antipsychotic continuation treatment post-ECT did not find an advantage for combination treatment over antidepressant alone in an elderly population, and the combination treatment was found to be associated with greater frequency of side effects. A recent randomized controlled trial found no advantage for beginning nortriptyline or venlafaxine during the course of ECT in preventing relapse. Observations of community practice post-ECT yield relapse rates similar to those found in the RCT assignments using antidepressants alone. Little is known about the use of second- and third-generation agents or anticonvulsant mood stabilizers in relapse prevention, although preliminary evidence suggests that the combination of lithium and antidepressant is more effective than other medication combinations in maintaining remission. There is also preliminary work indicating that combining medications with continuation ECT or cognitive therapy is more

useful than monotherapy. There are preliminary studies suggesting that relapse prevention is more effective when using a class of antidepressants that was not shown ineffective prior to ECT treatment. Modern practice suggests sparing use of antidepressants in bipolar patients, but in continuation treatment post-ECT, antiepileptic drugs or antipsychotics have not been shown to be all that effective in preventing relapse to bipolar depression. Little work in pharmacologic relapse prevention post-ECT for any other diagnosis than depression has been done in the modern era.

ECT was introduced before the psychopharmacologic era; thus, administering ECT after an index episode was not uncommon early in the history of ECT. With ECT populations becoming increasingly medication resistant and various pharmacologic strategies still characterized by relapse rates in the range of 40 percent, there is renewed interest in continuation ECT for patients who relapse on continuation pharmacology. The first randomized controlled comparison of continuation ECT in a fixed schedule with the combination of nortriptyline and lithium did not show an advantage for ECT and confirmed relapse rates observed in prior studies; relapse escalated when the schedule of administration dropped below once weekly. Other prospective data have confirmed these observations. Recent prospective, controlled, nonrandomized data as well as post hoc analyses from the RCT support the practice of tapering ECT after an acute course, combining ECT and pharmacology as continuation treatments, and using a symptom-driven, flexible schedule of ECT administration to improve relapse prevention. An RCT of this latter approach is nearing completion.

Recommendations from professional organizations have guided various aspects of the practice of continuation ECT although these bodies use differing thresholds for evidence when recommending strategies, a circumstance most notable when the National Institute for Health and Clinical Excellence in the United Kingdom issued guidelines in conflict with the Royal College of Psychiatrists as well as other professional bodies in the Commonwealth and globally. Suitable patients for continuation/maintenance ECT should have recurrent illness, demonstrated response to ECT, intolerance of or relapse on continuation medications, and ability to comply with continuation ECT protocols, particularly behavioral prescriptions, such as fasting prior to the procedure, and restrictions, such as refraining from important decision making for a period of time following each treatment. Generally, continuation ECT utilizes the electrode placement and dosage which produced remission. A taper of ECT from the acute phase to continuation phase has become common. Treatment frequency is gradually extended between treatment sessions, as tolerated, to monthly or even less often. It is thought that if patients remain well 2 months or more, they are out of episode, and the discontinuation of continuation treatment can be entertained. Continuation ECT is almost always given in an outpatient setting. Cognitive

functioning should be monitored as well as the status of psychiatric symptoms and medical fitness for each session. There has been one promising but uncontrolled pilot study of the use of cognitive behavioral therapy to augment somatic continuation treatments post-ECT.

Maintenance treatments are those given beyond the 6-month continuation treatment period and the purpose of these treatments is prevention of recurrence. Because patients receiving ECT in current practice are almost always suffering from highly recurrent illness, maintenance treatment can be entertained for most ECT responders who do not relapse with continuation treatments. There are now several controlled observations of the use of maintenance ECT up to 4 years with positive outcomes in relapse prevention for depression. Japanese investigators have published a case series on the use of ECT as a maintenance strategy for treatment-resistant schizophreniform and catatonic illness. Individual decisions about maintenance treatment will be informed by the number, frequency, and intensity of prior episodes; the ability to tolerate continued treatment; the ability to comply with treatment; the experience when attempts to taper treatment are made; and patient preference. Again the standard will be maintenance medications. Maintenance ECT is also available, especially for those patients undergoing continuation ECT when medications have proven inadequate. Maintenance ECT is similar in many aspects to continuation ECT, and it is performed at the longest intervals to sustain remission. There is no evidence that a lifetime maximum number of ECT applies.

INTERACTIONS OF MEDICATIONS AND ECT

ECT can interact with medications given for psychiatric as well as medical conditions. As such, it may be necessary to make changes in a patient's drug regimen when a course of ECT is undertaken. These changes can include tapering, withdrawing, augmenting, or adjusting the timing of administration of medications. Some psychotropics and many drugs for medical indications are continued. Medications that are already a part of the patient's treatment regimen and are expected to mitigate physiologic changes that accompany ECT or protect against adverse reactions to ECT or anesthesia are often given the morning of treatment. Examples are antihypertensives, inhalers for pulmonary conditions, and steroids, although anesthetists may wish to substitute IV steroids for oral steroids pre-ECT. It appears that stress doses of steroids may not be necessary at ECT. The administration of β-blockers pre-ECT should be carefully evaluated because of the propensity of this class of agents to enhance bradycardia or precipitate asystole; use of atropine pre-ECT should be considered in this circumstance. Labetalol crosses the blood–brain barrier and attenuates rCBF at ECT; both esmolol and labetalol are anticonvulsant at doses used in this setting. Lidocaine and its analogs can interfere with

seizure induction, and should be withheld. Other medications that should be held until after recovery are diuretics, hypoglycemics, including insulin, depending on an individual evaluation of the patient's needs, and long-acting cholinesterase inhibitors for glaucoma that may interfere with anesthesia recovery.

Medications that can result in severe adverse reactions include theophylline and lithium. Theophylline is a member of a class of agents that are known to increase seizure duration, and theophylline can result in status epilepticus, even at therapeutic levels, when administered during an ECT course. In fact, caffeine, a short-acting member of the same class of medications as theophylline, had been used to increase seizure duration when seizure length was thought to be a critical parameter to ECT efficacy. Inhaled bronchodilators and steroids, as well as other oral and inhaled agents, have replaced theophylline in modern pulmonary practice to a great extent. If theophylline is part of the therapeutic regimen and ECT is undertaken, consultation to evaluate the possibility of a substitute regimen is generally sought. Lithium combined with ECT has been shown to result in confusional states, serotonin syndrome, and prolonged and/or focal seizures. Although it is not neurotoxic in all patients and the toxicity of the combination may be dose related, discontinuation of lithium is frequently undertaken in an acute course of ECT. For continuation and maintenance ECT, doses are often withheld for 24 hours prior to treatment, allowing for washout, and markedly diminishing the chance of negative interaction. Another class of medications that can pose increased risk are β-lactam antibiotics and ciprofloxacin, which are proconvulsant, and have been associated with seizures in non-ECT patients and tardive seizures during a course of ECT.

Anticonvulsants can increase seizure threshold, decrease seizure expression, and may interfere with the efficacy of ECT, particularly in depressed patients. Interference with clinical outcome appears less marked in manic patients administered ECT with anticonvulsants, but only case report and observational data are available. When used for control of a seizure disorder, anticonvulsants are continued without alteration, unless poor seizure expression along with weak response occurs. In this circumstance, levels of drug may be reevaluated in consultation with neurology. If anticonvulsants are used for psychiatric indications, they are generally discontinued, particularly in bipolar depression, to prevent a negative effect on clinical response.

Benzodiazepines are also anticonvulsants, but more often used for psychiatric indications. There are reports of diminished efficacy when benzodiazepines are used in conjunction with ECT, especially right unilateral electrode placement. As a result, this class of medications is tapered to the least necessary dose or discontinued. If use of benzodiazepines is considered necessary for management of severe

behavioral syndromes, agents with shorter half-lives are preferable and should be withheld several hours prior to treatment. Use of flumazenil to reverse benzodiazepines has been reported, but requires further study and benzodiazepines may need to be administered during recovery to abort possible withdrawal. Little is known about the interaction of other sedative/hypnotics with ECT.

Anesthetists have been averse to allowing the use of monoamine oxidase inhibitors (MAOIs) concurrently with elective anesthesia, due to a potential for interference with management of cardiovascular abnormalities that may arise at treatment. ECT seldom results in a situation where the use of vasopressors, that would not be metabolized because of interference by the recent use of MAOIs, is necessary. Hence, the literature includes many reports of the safe performance of ECT in the presence of MAOIs. Current practice allows for consultation between the psychiatric and anesthetic practitioners as to the best recommendation for any particular individual or practice setting.

An augmenting effect for neuroleptics when combined with ECT in schizophrenia can be accepted with some confidence. As discussed previously, the combination treatment is more effective than ECT alone, and may be more effective than medications alone. A few older studies even suggest that the long-term outcome of schizophrenia may be enhanced by combination treatment. Whether the acute or long-term effects carry over to other psychotic conditions is not clear; there is evidence that they play no augmenting role in nonpsychotic depression. Most of the published data involve the use of typical neuroleptics, It appears that low-potency typical antipsychotics may be proconvulsant; haloperidol and fluphenazine appear to have little impact on parameters of seizure quality. The use of ECT combined with atypical neuroleptics appears to be safe. Clozapine and olanzapine have been observed to have proconvulsant properties and quetiapine, anticonvulsant properties. Atypical antipsychotics are reported to have U-shaped functions with respect to convulsant properties. Because of these circumstances, dosing of neuroleptics should generally be in the moderate range. The only antipsychotic that cannot be combined safely with ECT is reserpine.

Earlier work with tricyclics (TCAs) and MAOIs provided some evidence to recommend the use of antidepressants to augment ECT. A recent RCT found that both nortriptyline and venlafaxine increased the response rate to ECT by about 15 percent; nortriptyline was associated with diminished adverse cognitive effects, whereas venlafaxine was associated with increased adverse cognitive effects. The combination of ECT and antidepressants is common practice in many parts of the world. Increasingly, in the United States psychotropics, including antidepressants, are continued during a course of ECT because of expanding limitations placed by insurers on utilization of mental health resources. It appears that

TCAs are safe in noncardiac populations, and SSRIs in most populations although they can increase seizure duration. After a series of cases where asystole occurred at doses of venlafaxine greater than 225 mg with the concurrent use of atropine, it may be advisable to limit the dosage of venlafaxine if it is combined with ECT. The well-known proconvulsant effect of bupropion prompts caution in combining it with ECT in the absence of data documenting a clear benefit in doing so. It is unknown whether newer medications now used more frequently for depression have augmenting effects.

Finally, there has been limited interest in pharmacologic approaches to mitigating the cognitive side effects of ECT. Preliminary data using physostigmine, thyrotropin-releasing hormone, and naloxone have been published, but whether clinically significant effects could be obtained remains in doubt.

PHYSIOLOGY, ADVERSE EFFECTS, AND CONTRAINDICATIONS TO ECT

ECT is generally low risk and one of the safest procedures performed under general anesthesia. Estimated risk of serious complications, occurring in about 1/1,000 patients, is similar to that of general anesthesia for minor medical procedures. Risk of death, likewise in the range of that observed for such procedures, is about 1/10,000 patients in modern populations undergoing ECT. It has been argued that this rate is less than the spontaneous death rate in comparable community-dwelling populations. Some believe that these rates of serious complications and death are lower than those related to the use of antidepressant medications. Mandated reporting of deaths occurring during and up to 14 days post-ECT in Texas, revealed only one death related to the procedure in the first 5 years. Although cardiac complications are considered to be the most common cause of morbidity and mortality, this death was due to laryngospasm. Cardiac deaths in this population were unrelated to ECT.

Cardiovascular System

The ECT procedure routinely results in parasympathetic outflow during and immediately after the electrical stimulus. Unopposed by anticholinergic medications, sinus bradycardia and even brief, clinically insignificant, asystole can occur as a result. In patients over 50 years old, supraventricular ectopic beats are common, but clinically insignificant. In vulnerable patients, vagally related arrhythmias can occur such as atrial, junctional, and nodal rhythms as well as atrial flutter, and atrial fibrillation. Following parasympathetic stimulation, sympathetic output increases, mainly during the electrical stimulation and postictally when the adrenal gland also releases catecholamines. As a consequence, heart rate, blood pressure, and, perhaps, most importantly, rate pressure product increase.

These increases peak immediately postictally and drop off within minutes to pre-ECT values, although in those over 50 years, normalization takes up to 1 hour for blood pressure and longer for heart rate. There can be a short period of decreased heart rate at about 5 minutes post-ECT. Related abnormalities of cardiac rhythm can include bigeminy and trigeminy, which are generally benign. Repolarization abnormalities such as T-wave changes and ST segment depression can appear infrequently. Ventricular tachycardia, usually nonsustained, and fibrillation can occur rarely. About 55 percent of those with pre-existing cardiovascular abnormalities will manifest some abnormality in the course of ECT. About 10 percent of one sample manifested transient increase in troponins of uncertain clinical significance.

With judicious pretreatment evaluation of both the risks and benefits of ECT, cardiac consultation, and the stabilization of any underlying cardiac disease, many of the more cardiac serious complications can be avoided. About 90 percent of these patients will be able to complete a course of ECT safely. For example, ECT has been administered to patients with histories of pulmonary hypertension, subacute MI, abdominal aortic aneurysm, and Takatsubo syndrome. Congestive heart failure patients with ejection fractions over 30 percent have been able to tolerate ECT. In patients with myocardial infarct, a period of time must be allowed for healing, but the risk relates not only to the time elapsed but also to the severity of resulting cardiac compromise. As with arrhythmias and congestive heart failure, stress testing can be informative in deciding whether a patient can tolerate ECT. The presence of a pacemaker or device to abort arrhythmias is not a contraindication to ECT. ECT has been administered without complication to stable patients with transplanted hearts.

Respiratory System

Along with cardiac complications, adverse events related to pulmonary functions are a leading cause of morbidity and mortality associated with ECT. Most of the effects are associated with anesthesia procedures, rather than the physiology of ECT itself. The use of muscle relaxants leads to the cessation of breathing and oxygen delivered by positive pressure through a mask substitutes. Where patients have underlying pulmonary disease, exacerbation of asthma or COPD is possible, as well as excess secretions and aspiration. Rare cases of negative pressure pulmonary edema have been attributed to inspiration against an obstructed airway or mechanical irritation precipitating laryngospasm; it can also be neurogenically induced, for example, as a complication of status epilepticus. With proper preparation pre-ECT and careful anesthetic technique, these events are uncommon. Obstructive sleep apnea poses risk or may even be discovered during ECT anesthesia; it can generally be managed acutely with preoxygenation, and the use of nasal pharyngeal airway or laryngeal mask.

Another rarely observed complication is prolonged apnea, resulting from slow metabolism of succinylcholine, due to pseudocholinesterase deficiency. Supportive therapy suffices to manage this complication, and subsequent sessions of ECT can be given with a nondepolarizing agent.

CNS

Seizure induction results in a hypermetabolic state characterized by increased blood flow with significant elevation in oxygen and glucose consumption. Increased blood–brain barrier permeability has been hypothesized due to capillary hypertension. Postictally, metabolic suppression develops. These physiologic changes have not been demonstrated to produce structural brain changes or damage in well-designed and controlled animal and human studies or by brain imaging or analyses of cerebrospinal fluid. There is insufficient energy, even at maximal output of a modern brief pulse device, to cause brain tissue injury by temperature elevation.

ECT can be complicated by prolonged seizure duration, defined as greater than 3 minutes by motor or EEG measurement, in 1 to 2 percent of treatment sessions. The risk is greater in adolescents and young adults. There is a risk of structural brain injury along with cardiovascular and pulmonary complications if the seizure persists beyond 30 minutes, even when oxygenation is adequate. Termination is undertaken using either parenteral benzodiazepines or the short-acting barbiturate used for anesthesia. If this fails, a standard protocol for status epilepticus treatment may be instituted. Patients who experience prolonged seizures need to be reassessed for underlying etiology. Prolonged seizures are more common in the presence of proconvulsant medications such as theophylline or lithium, and possibly trazodone, or when more than one seizure is induced in a single session. They are also associated with medical conditions that decrease seizure threshold, such as electrolyte abnormalities and epilepsy. Nonconvulsive status should be suspected if there is lack of recovery from anesthesia or abnormalities in mental status persist. This state may be detectable by EEG monitoring, but can be present when the typical 2-lead EEG does not exhibit abnormalities. Tardive seizures are rare. The rate of onset of new seizure disorder after ECT does not exceed baseline rates in the general population. Because of its anticonvulsant properties, ECT can be given safely to patients with epilepsy.

Before the development of brain imaging, deaths and neurologic deterioration were reported when ECT was given to patients with occult brain tumors and brain metastases. Evaluation, including radiologic procedures, of patients with neurologic findings or abnormal mental status now most often leads to the identification and treatment of such diseases prior to any administration of ECT. It is possible for ECT to be given safely to patients with small, slow-growing tumors without associated increased

intracranial pressure, such as meningiomas. ECT has also been given to patients with healed skull and brain trauma, including craniotomy, as long as any skull defect is avoided. Patients may also experience transient CNS symptoms of almost any variety, for example, various cortical sensory abnormalities.

Depression not uncommonly complicates CNS disease such as stroke and dementia. If patients with stroke are neurologically stable several weeks after a cerebrovascular event, ECT may be given. Attention should be paid to the treatment and stabilization of the underlying cause of the stroke as well as to the current neurologic status. Patients with dementia receive ECT for complicating depression, but are at risk for greater adverse cognitive effects, such as delirium or prolonged confusion. There is interest in using the treatment for agitation complication dementia, but little supporting, controlled evidence. Treatment using techniques focused on mitigating cognitive effects, such as right unilateral electrode placement, specific tailoring of dosage, brief pulse waveform, and twice weekly frequency, are often employed.

Dental

Due to direct stimulation of jaw muscles during delivery of the electrical stimulus, patients bite down, sometimes strongly, and unstable teeth may be broken or dislodged completely. Also vulnerable are front teeth with crowns or other dental modifications such as laminates. The approach to this complication is routine prophylactic screening and may necessitate dental consultation and procedures prior to ECT. If teeth are injured during treatment, care must be taken to remove any broken fragments before respiration resumes, when they may obstruct the patient's airway.

Musculoskeletal

With the consistent use of muscle relaxants, the most serious of these adverse effects is only of historical interest. Before the development of curare and later succinylcholine, fracture of the long bones or spine occurred in up to 10 percent of patients as a consequence of unmodified movements associated with seizure induction and expression. The usual patient, who receives enough muscle relaxant to modify movement rather than fully abolish it, has no significant musculoskeletal complications. Those with skeletal disease, for example, osteoporosis, are more fully paralyzed for treatment. Although this complication has been almost completely abolished, elderly patients remain at risk for falls during a course of ECT and appropriate fall precaution protocols should be used.

More common in the modern era and, ironically, often an adverse effect of depolarizing agents, is myalgia. These muscle pains are a consequence of the fasciculations produced during depolarization and are most prominent

after the first treatment but seldom present after a few more sessions. Although myalgia can be mitigated by the use of curare along with succinylcholine, it can also be managed symptomatically with analgesics with the advantage of not prolonging or complicating the treatment session.

Those patients with temporomandibular joint problems may experience an exacerbation of pain, resulting from the unmodified contraction of jaw muscles during stimulus delivery. Spasm can also be induced. Prophylactic approaches may be advisable, such as the use of dental appliances already in the patient's possession or the administration of prophylactic analgesics such as nonsteroidal anti-inflammatory medications. Symptomatic post-ECT approaches may also suffice.

Neurocognitive

Cognitive side effects are a major limitation to the use of ECT, although in recent years the syndrome of postoperative cognitive dysfunction has been detected in patients undergoing medical procedures requiring either regional or general anesthesia with up to 10 to 20 percent incidence, particularly those over 60 years. With ECT, these effects have been best described for depressed populations. Most of the reports of cognitive function in schizophrenic and even autistic populations show stable cognitive function after a course of ECT. The phenomena are made more difficult to discriminate because depression itself is accompanied by impairments in attention, concentration, and learning. Post-ECT, with remission of depression, attention, and concentration are improved. Several factors influence the likelihood and extent to which impairments develop, but response to treatment is not a variable in the development of the cognitive adverse effects of ECT. Furthermore, functions such as reasoning, creativity, and the ability to form memories, are not affected in any enduring fashion. Nondeclarative memory, that is, procedural memory and priming, does not appear to be affected either. It is also important to note that objective measurement and subjective report of cognitive side effects do not correlate for many measures of subjective memory with the exception of global assessment.

The first factor in the development of adverse cognitive side effects is time point in treatment, which influences not only severity, but the type, of effect observed. Disorientation, diminished processing speed, decreased anterograde and retrograde memory, and errors in visual–spatial function and word finding are greatest immediately after a treatment session. These effects diminish fairly quickly once the treatment course ends. There is preliminary evidence that reactivating and reconsolidating a semantic memory may make it particularly vulnerable to impairment. Except for retrograde memory impairment, other cognitive effects generally return to pre-ECT baseline or even improve over baseline with right unilateral ECT.

Diminished processing speed persists when sine wave ECT has been administered. Retrograde memory improves more gradually, but spotty deficits may persist, with recent memories for public information more vulnerable than older memories for personal experiences. Retrograde memory deficits may persist several months after treatment ends. In a community setting, 12.4 percent of patients manifested marked and persistent retrograde amnesia, defined as deterioration in autobiographical memory of at least two standard deviations from pre-ECT baseline at 6 months post-ECT. The only significantly correlated technical factor was the number of bilateral ECT; female sex was the only demographic or clinical factor to achieve significance.

Table 34.35–3.
Treatment Factors Influencing the Degree of Adverse Cognitive Effects from ECT

Treatment factor	Technique Associated with Fewer Adverse Cognitive Effects
Electrode placement	Right unilateral
Stimulus waveform	Ultrabrief (and brief) pulse
Stimulus intensity	Threshold determination; dosage closer to threshold
Number of treatments	Fewest necessary to achieve remission or plateau of target symptoms
Frequency of treatment sessions	Less frequent, e.g., twice weekly for acute treatment
Number of seizures per session	One
Simultaneous use of psychotropic medications	Discontinue lithium, sedatives; reduce dose of neuroleptics
Dosage of anesthetic medications	Adjust dosage to produce light anesthesia; ketamine

The likelihood of developing adverse cognitive effects is influenced by technical factors in the treatment: bilateral electrode placement; inefficient stimulus waveforms, most particularly, sine waveform, but including brief pulse; markedly suprathreshold dosing; greater frequency and number of treatments; and larger doses of anesthesia are more likely to result in deficits during the acute and subacute course of treatment. Bilateral electrode placement is associated with greater likelihood of persistent retrograde memory deficits long term. There is mixed data on the relative cognitive profiles of bifrontal versus bifrontotemporal electrode placement. In recent decades, sine waveforms, which are no longer recommended, have been largely replaced by brief pulse and, most recently, ultrabrief pulse stimulation, and much has been learned about optimizing right unilateral electrode placement and restricting dosage while maintaining efficacy. The advantage of ultrabrief pulse stimulation over brief pulse stimulation in cognitive side-effect profile has become more apparent in the past decade of increasing use of ultrabrief pulse waveforms (Table 34.35–3).

Patient characteristics and clinical factors also impact the frequency and severity of adverse cognitive effects. Those with baseline neurologic disease, MRI abnormalities, and baseline impairments in global cognitive functioning are more vulnerable to developing deficits, as well as older and female patients. The concurrent use of certain medications such as lithium or medications with anticholinergic effects also increases the risk of adverse cognitive effects. There have been no well-established pharmacologic approaches to attenuating adverse cognitive effects, although there is interest in using NMDA antagonists, such as ketamine anesthesia, antiglucocorticoid agents such as mifepristone, and thyroid hormone.

Perhaps because subjective memory correlates best with mood and the majority of patients respond to ECT in systematic research, most patients studied in these settings report improved cognitive function post-ECT. It has been difficult to distinguish these patients from controls in the subjective assessment of memory months after ECT in randomized controlled trials; in community settings, which are characterized by inclusion of a broader range of patients and ECT techniques, subjective complaints of memory impairment are more frequent and correlate to the number of bilateral ECT received. Nonetheless, generally favorable attitudes are expressed by patients who have received ECT, especially compared to attitudes of the general public, and the majority of patients would have treatment again if necessary.

Other Adverse Effects

Postictal agitation has an undefined etiology, but can be associated with various factors, including electrode placement, electrical dose over seizure threshold, seizure duration, anesthetic dose, concomitant use of medications, and anxiety or agitation pre-ECT. Although ranging in severity, the more severe forms of agitation require emergent management with decreased stimulation, medications such as short-acting benzodiazepines, haldol and ECT anesthetics, and even restraints. Prophylaxis may include switching electrode placement to unilateral, attention to electrical dosing, increasing the dose of anesthetic agent pre- and post-ECT, modifying the dose of succinylcholine to prevent awakening while paralyzed or to decrease lactate generation, and alteration of medication regimens to address agitation by the use of neuroleptics.

ECT may be associated with a number of less serious, although nonetheless, uncomfortable side effects as well. Headaches and nausea occasionally occur postictally. Headache has been reported to occur in up to half of all patients, although this figure seems unrealistically high. The etiology of postictal headache is unknown, and ascertaining cause is complicated by the presence or withdrawal of medications and the fact that anxiety often accompanies target illnesses for which ECT is used, such as

depression. The throbbing quality of the headache suggests a vascular component, and migraineurs certainly seem more vulnerable to developing this side effect. Symptomatic management with analgesics, including nonsteroidal anti-inflammatory drugs, is often sufficient, but use of these agents prophylactically, for example, intravenous ketorolac, may be required. Migraine sufferers may need to use triptans posttreatment.

Nausea may accompany headache or occur alone in a minority of patients. Again the etiology is complex, including the use of anesthetic agents. Antiemetics, including dopamine antagonists such as compazine, can be employed symptomatically. The use of $5HT_3$ antagonists, such as ondansetron, has gained in use recently.

FUTURE DIRECTIONS

There is interest in continued refinements of ECT technique. Common themes in these approaches are as follows: focusing the treatment spatially to optimize dosing in brain areas associated with putative neural networks involved in depression and other psychopathologies that are indications for ECT, and to diminish dosing in areas associated with adverse cognitive effects; improving the efficiency of a noninvasive electrical stimulus in direction and amplitude, even to a subconvulsive level. Methodologies such as magnetic induction of convulsive treatment and focal electrically assisted seizure therapy are examples of innovation designs under investigation. This research is parallel to investigations in magnetic stimulation, for example, rTMS, and to the renaissance of invasive electrical techniques, for example, vagal nerve stimulation and deep brain stimulation.

REFERENCES

American Psychiatric Association. *The Practice of Electroconvulsive Therapy: Recommendations for Treatment, Training, and Privileging – A Task Force Report*. 2nd ed. Washington, DC: American Psychiatric Press; 2001.

Bergsholm P. Patients' perspectives on electroconvulsive therapy: a reevaluation of the review by Rose et al on memory loss after electroconvulsive therapy. *J ECT*. 2012;28(1):27–30.

Bouckaert F, Sienaert P, Obbels J, et al. ECT: its brain enabling effects: a review of electroconvulsive therapy-induced structural brain plasticity. *J ECT*. 2014;30(2):143–151.

Boylan LS, Haskett RF, Mulsant BH, et al. Determinants of seizure threshold in ECT: benzodiazepine use, anesthetic dosage, and other factors. *J ECT*. 2000;16(1):3–18.

Ghaziuddin N, Kutcher SP, Knapp P, et al; Work Group on Quality Issues; AACAP. Practice parameter for use of electroconvulsive therapy with adolescents. *J Am Acad Child Adolesc Psychiatry*. 2004;43(12):1521–1539.

Haskett RF, Loo C. Adjunctive psychotropic medications during electroconvulsive therapy in the treatment of depression, mania, and schizophrenia. *J ECT*. 2010;26(3):196–201.

*Kellner CH, Knapp RG, Petrides G, et al. Continuation electroconvulsive therapy vs

pharmacotherapy for relapse prevention in major depression: a multisite study from the consortium for research in electroconvulsive therapy (CORE). *Arch Gen Psychiatry*. 2006;63(12):1337–1344.

*Kellner CH, Knapp R, Husain M, et al. Bifrontal, bitemporal, and right unilateral electrode placement in ECT: a randomised trial. *B J Psychiatry*. 2010;196:226–234.

Mayur P, Byth K, Harris A. Autobiographical and subjective memory with right unilateral high-dose 0.3 millisecond ultrabrief-pulse and 1 millisecond brief-pulse electroconvulsive therapy: a double-blind randomized controlled trial. *J ECT*. 2013;29(4):277–282.

McCall WV, Reboussin D, Prudic J, et al. Poor health-related quality of life prior to ECT in depressed patients normalizes with sustained remission after ECT. *J Affect Disord*. 2013;147(1–3):107–111.

McCall WV, Reboussin DM, Weiner RD, Sackeim HA. Titrated moderately suprathreshold vs fixed high-dose right unilateral electroconvulsive therapy: acute antidepressant and cognitive effects. *Arch Gen Psychiatry*. 2000;57(5):438–444.

McClintock SM, Choi J, Deng ZD, Appelbaum LG, Krystal AD, Lisanby SH. Multifactorial determinants of the neurocognitive effects of electroconvulsive therapy. *J ECT*. 2014;30(2):165–176.

Østergaard SD, Bolwig TG, Petrides G. No causal association between electroconvulsive therapy and death: a summary of a report from the Danish Health and Medicines Authority covering 99,728 treatments. *J ECT*. 2014;30(4):263–264.

Perera TD, Luber B, Nobler MS, Prudic J, Anderson C, Sackeim HA. Seizure expression during electroconvulsive therapy: relationships with clinical outcome and cognitive side effects. *Neuropsychopharmacology*. 2004;29(4):813–825.

Petrides G, Malur C, Braga RJ, et al. Electroconvulsive therapy augmentation in clozapine-resistant schizophrenia: a prospective, randomized study. *Am J Psychiatry*. 2015;172(1):52–58.

*Prudic J, Haskett RF, McCall WV, et al. Pharmacological strategies in the prevention of relapse after electroconvulsive therapy. *J ECT*. 2013;29(1):3–12.

Prudic J. Strategies to minimize cognitive side effects with ECT: aspects of ECT technique. *J ECT*. 2008;24(1):46–51.

Prudic J, Olfson M, Marcus SC, Fuller RB, Sackeim HA. Effectiveness of electroconvulsive therapy in community settings. *Biol Psychiatry*. 2004;55(3):301–312.

Rasmussen KG. Do patients with personality disorders respond differentially to electroconvulsive therapy? a review of the literature and consideration of conceptual issues. *J ECT*. 2015;31(1):6–12.

Sackeim HA, Haskett RF, Mulsant BH, et al. Continuation pharmacotherapy in the prevention of relapse following electroconvulsive therapy: a randomized controlled trial. *JAMA*. 2001;285(10):1299–1307.

Sackeim HA, Long J, Luber B, et al. Physical properties and quantifications of the ECT stimulus: I. Basic principles. *Convuls Ther*. 1994;10(2):93–123.

*Sackeim HA, Prudic J, Devanand DP, et al. A prospective, randomized, double-blind comparison of bilateral and right unilateral electroconvulsive therapy at different stimulus intensities. *Arch Gen Psychiatry*. 2000;57(5):425–434.

Sackeim HA, Prudic J, Fuller R, Keilp J, Lavori PW, Olfson M. The cognitive effects of electroconvulsive therapy in community settings. *Neuropsychopharmacology*. 2007;32(1):244–254.

*Sackeim HA, Prudic J, Nobler MS, et al. Effects of pulse width and electrode placement on the efficacy and cognitive effects of electroconvulsive therapy. *Brain Stimul*. 2008;1:71–83.

Schoeyen HK, Kessler U, Andreassen OA, et al. Treatment resistant bipolar depression: a randomized controlled trial of electroconvulsive therapy versus algorithm-based pharmacological treatment. *Am J Psychiatry*. 2015;172(1):41–51.

Sundsted KK, Burton MC, Shah R, Lapid MI. Preanasthesia medical evaluation for electroconvulsive therapy: a review of the literature. *J ECT*. 2014;30(1):35–42.

*Tharyan P, Adams CE. Electroconvulsive therapy for schizophrenia. *Cochrane Database Syst Rev* 2002;CD000076.

Verwijk E, Comijs HC, Kok RM, Spaans HP, Stek ML, Scherder EJ. Neurocognitive effects after brief pulse and ultrabrief pulse unilateral electroconvulsive therapy for major depression: a review. *J Affect Disord*. 2012;140:233–243.

Zhuo C, Yu C. Functional neuroimaging changes subsequent to electroconvulsive therapy in unipolar depression: a review of the literature. *J ECT*. 2014;30(4):265–274.

▲ 34.36 Brain Stimulation Methods

STEFAN B. ROWNY, M.D., AND MELANIE LUCAS, PH.D.

INTRODUCTION

Brain stimulation in psychiatric practice and research uses electric current, or changing magnetic field to induce currents within the brain, to alter neuronal firing. There is a growing list of tools capable of eliciting such neuromodulation, each with a different spectrum of action and therapeutic utility in psychiatric disorders. These tools either apply electrical or magnetic fields transcranially or involve the surgical implantation of electrodes to deliver electrical currents to a cranial nerve or to the brain tissue directly. The transcranial techniques include cranial electrical stimulation (CES), electroconvulsive therapy (ECT), transcranial direct current stimulation (tDCS, also called direct current polarization), transcranial magnetic stimulation (TMS), and magnetic seizure therapy (MST). The surgical techniques include Implanted cortical stimulator (ICS), deep brain stimulation (DBS), and vagus nerve stimulation (VNS). These methods are briefly described in Table 34.36–1. ECT and DBS will be discussed in separate chapters.

The historical roots of modern brain stimulation in psychiatry can be traced to the introduction of ECT in 1938 by Ugo Cerletti and Lucio Bini. Since then, electrical methods for neuromodulation have been in continuous use in psychiatry. ECT, which uses alternating electrical current applied to the scalp to induce a seizure, has been substantially improved

over the intervening decades, both in refinements to the electrical stimulus characteristics and the placement of electrodes, as well as in advancements to the accompanying anesthesia, particularly with the use of muscle relaxants (e.g., succinylcholine, atracurium). In 1949 another alternating current brain stimulation method, CES, was developed in the former Soviet Union. It was originally called "Electrosleep" and as the name suggests was initially used to treat insomnia. This method uses much lower levels of the alternating current than ECT, and is thus subconvulsive. In 1985, nearly 50 years after the first use of ECT, Anthony and colleagues published on the first use of pulsed magnetic fields to stimulate the brain with a procedure called TMS. TMS was initially used in neurology for studies of nerve conduction, but quickly caught the attention of psychiatrists eager to explore other, less invasive alternatives to ECT. This nonconvulsive, well-tolerated stimulation method through TMS has been approved by the FDA for treatment of depression. It also continues to be under active study, with some promising results in the treatment of various other psychiatric disorders like bipolar and anxiety disorders (i.e., posttraumatic stress disorder and obsessive-compulsive disorder), schizophrenia as described by Lefauchear and colleagues in 2014, addiction disorders—particularly nicotine dependence, and even depersonalization disorder. MST, a convulsive treatment derived from application of more powerful magnetic stimulation, was first tested in an animal in 1998 and in a human in 2000. Since then, MST has been under investigation in nonhuman primates and in humans as a more focal means of inducing seizures in an attempt to retain the unparalleled efficacy of ECT but with fewer side effects.

Two more invasive brain stimulation methods, DBS and VNS, were introduced about a decade following the first trials of TMS. Both DBS and VNS were first approved by the U.S. Food and Drug and Administration (FDA) in 1997 with indications for treatment of neurological conditions: VNS was originally approved as adjunctive therapy in reducing the frequency of epileptic seizures in adults and children aged 12 years and over; while DBS was initially approved for the treatment of essential tremor and Parkinson tremor. Currently, indications for DBS include treatment of all symptoms of Parkinson disease, including tremor, slowness and stiffness, and involuntary movements induced by medications, as well as symptoms of OCD that are not adequately controlled with medications or other therapies. At present, indications for VNS include adult treatment-resistant depression and epilepsy.

Table 34.36–1.
Brain Stimulation Methods: Definitions

CES	**Cranial Electrical Stimulation** Weak alternating electrical current is applied via saline soaked felt-covered electrodes, usually to the earlobes or bifrontal scalp location.

cTMS **Controllable Pulse Shape Transcranial Magnetic Stimulation**
A transcranial magnetic stimulator with a modified circuit that can induce nearly rectangular pulses, in contrast to conventional TMS that induces cosine waveforms.

DBS **Deep Brain Stimulation**
A wire electrode with multiple stimulating sites that is surgically implanted into the brain and chronically administers electrical stimulation.

ECT **Electroconvulsive Therapy**
Electricity is applied transcranially via electrodes placed on the scalp to induce a seizure under anesthesia.

MST **Magnetic Seizure Therapy**
High intensities of transcranial magnetic stimulation are applied to induce focal seizures under anesthesia, as a less invasive form of ECT.

PAS **Paired Associative Stimulation**
Transcranial magnetic stimulation is paired with peripheral nerve stimulation, resulting in Hebbian plasticity, altering functioning of the corticospinal tract.

tDCS **Transcranial Direct Current Stimulation, also called Direct Current Polarization**
Weak direct electrical current is applied to the scalp via sponge electrodes to polarize underlying brain tissue.

TBS **Theta-Burst Stimulation**
High-frequency bursts (e.g., 3 pulses at 50 Hz) of transcranial magnetic stimulation are applied at a regular frequency (e.g., 5 bursts per second). TBS can be given continuously (cTBS) or intermittently (iTBS).

TES **Transcranial Electrical Stimulation**
Alternating electrical currents applied to the scalp.

TMS **Transcranial Magnetic Stimulation**
Rapidly alternating magnetic fields are applied to the scalp to induce small, focused electrical currents in superficial cortex. Remote sites may be activated transsynaptically.

TNS **Trigeminal Nerve Stimulation**
Mild electrical current stimulates branches of the largest cranial nerve—the trigeminal nerve—either externally or subcutaneously aimed at modulating activity in brain areas such as the anterior cingulate cortex, nucleus solitarius, locus coeruleus, vagus nerve, and the cerebral cortex.

VNS **Vagus Nerve Stimulation**
An implanted pacemaker administers electrical pulses to the vagus nerve in the neck, activating the vagal afferents.

Frequently, initial developments in brain stimulation have been in the field of neurology. For example, TMS, DBS, and VNS originated as tools to treat neurological conditions but psychiatrists quickly saw the potential for those tools in the treatment of psychiatric conditions and brain stimulation techniques in psychiatry are gaining momentum and continue to evolve. It is not clear, however, whether those technologies and doses that are designed to treat neurological indications can be directly co-opted for the purposes of treating mental illness.

THERAPEUTIC NEUROMODULATION: TREATING PSYCHIATRIC DISORDERS THROUGH BRAIN STIMULATION

Mechanism of Action

Electrical Stimulation—Common Pathway. The brain stimulation tools considered in this chapter generate either electrical or magnetic pulses. However, both of these share a common final pathway—they affect the neurons electrically. That electrical effect may either be through the direct application of electricity and the resultant electrical field that reaches neurons or through the indirect induction of electricity via magnetic stimulation. The direct forms of electrical stimulation are exemplified in either *transcranial* delivery, as with ECT, CES, TNS, and tDCS, or surgical *intracerebral* delivery, as in the case of DBS or direct cortical stimulation (epidural or subdural). The indirect forms of electrical stimulation include TMS and MST, which induce electric fields in the brain through the application of alternating magnetic fields (Fig. 34.36–1). Of note, both the epidural and intracerebral modalities are more focal than the transcranial application of electricity because electrodes are placed directly in the neuronal tissue, bypassing the impedance of the scalp and skull. Magnetic stimulation methods also bypass the impedance of the scalp and skull, hence they are also more focal when compared to transcranial direct application of electricity. Magnetic stimulation is considered an indirect method of electrical brain stimulation in that changing magnetic field from these devices induces electricity in the brain according to the principle first described by Faraday in a law that bears his name, and later incorporated into Maxwell's equations, which unify all of electromagnetism. The magnetic modalities achieve their enhanced spatial resolution (focality) noninvasively, in contrast to the intracerebral and epidural methods, therefore these modalities are in the center of intensive research as they offer a promise of an unparalleled degree of spatial specificity without the need for surgery.

All but one of the brain stimulation modalities described here act by stimulating neurons; the one exception is tDCS, which does not stimulate but rather it polarizes. In this sense, the "S" of tDCS is a misnomer. It is more accurate to conceptualize tDCS as exerting a polarizing effect aimed at altering the likelihood of neuronal firing.

The action of the subconvulsive modalities of stimulation relies on the effects of the repeated stimulation of the targeted neural circuitry. However, in the case of the convulsive modalities (ECT and MST), the action is dependent upon the seizure induced by the stimulation and the effects of repeated seizure induction on brain processes. This is not to say that the form of stimulation that triggers the seizure has no impact on outcome. Indeed, it is well replicated that electrode placement and electrical stimulus parameters have a profound impact on the efficacy and side effects of ECT. Whether the same will be true for MST is under active investigation. Importantly, MST has the capability of inducing seizure in a more localized manner and thus to target brain regions and their neuronal networks thought to be implicated in depression, while avoiding those

responsible for untoward side effects.

Classification of tools that act to stimulate the brain electrically, using direct (electrical) or indirect (magnetic) means.

Comparison of Brain Stimulation with Psychopharmacology. Psychopharmacology and brain stimulation differ in several clinically relevant respects. Pharmacological agents rely on molecular action at the synapse or intracellular targets. They bind to either presynaptic or postsynaptic receptors, or other intracellular targets in order to affect their function. In contrast, brain stimulation induces action potentials via electrical depolarization of axons independently of receptor binding. Thus, not only does brain stimulation have the potential to work when medications have failed but also may act synergistically with medications as their sites of action and mechanisms of action are different. Because the chief mechanism of action in all modalities of brain stimulation is direct electrical modulation of the brain's neuronal activity, this approach avoids the systemic side effects seen with medications that orally introduce exogenous compounds, which are then distributed throughout the body to sites that have little to do with the intended target. These compounds may be secondarily metabolized to other biologically active metabolites. In contrast, brain stimulation has no exogenous compounds that circulate systemically in the bloodstream, does not undergo clearance, is not protein bound, and has no known interactions with the metabolism or clearance of psychopharmacological agents.

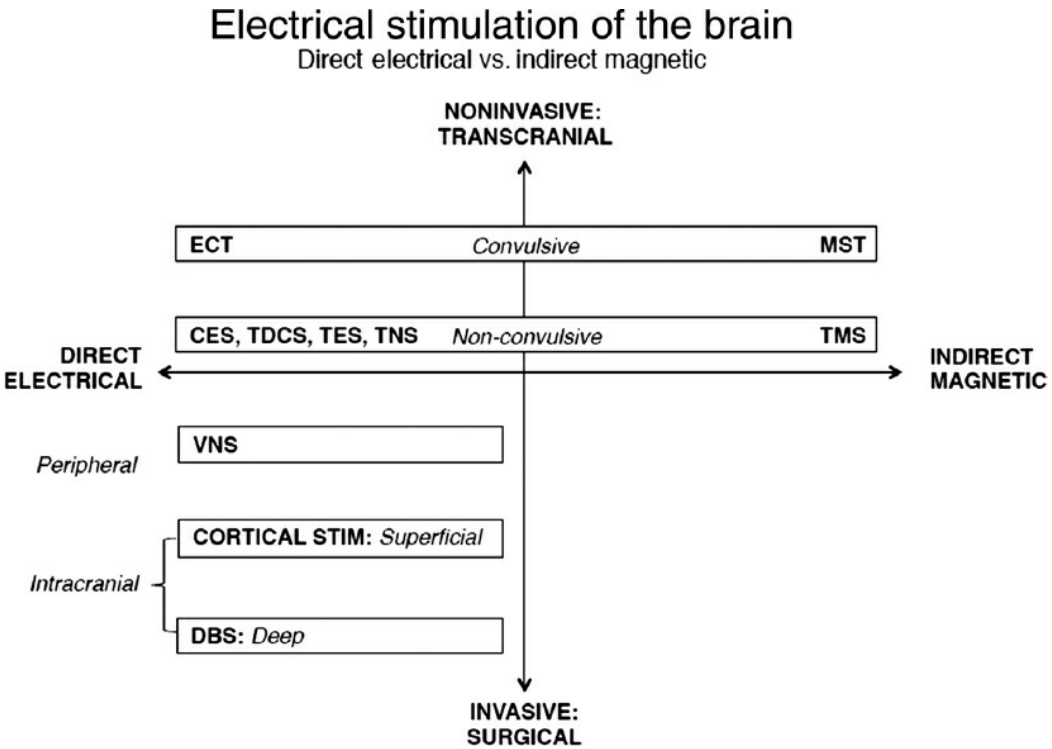

FIGURE 34.36–1. Classification of tools that act to stimulate the brain electrically using

direct (electrical) or indirect (magnetic) means.

Table 34.36–2.
Brain Stimulation versus Psychopharmacological Therapy

	Brain Stimulation	Psychopharmacology
Therapeutic agent	Electrical or electromagnetic	Neurochemical
Temporal characteristics	Phasic	Tonic
Delivery	Verifiable on electronic or clinical record	Dependent on patient compliance
Site of action	Axonal depolarization	Molecular target
Distribution	Focally applied to the brain	Systemic, throughout the body
Duration of availability	Milliseconds to seconds	Typical half-life of hours to weeks
Dosing parameters	Intensity, frequency, intertrain interval, number of pulses, site of stimulation	Milligrams, blood level
Metabolism	None	Hepatic
Clearance	None	Renal
Interactions with metabolism of other agents	None known	Possible

Brain stimulation is applied in a phasic fashion, which simply means that in contrast to medications that can remain biologically available in a patient's body for hours to weeks, or even months in the case of injectable depots, electricity can be turned on or off, with a flip of a switch. Most of the contemporary brain stimulation modalities that administer alternating electric current—either directly or via electromagnetic conduction (e.g., ECT, MST, DBS, VNS)—deliver pulses, or collections of pulses otherwise referred to as trains of pulses, with each pulse typically lasting 1 ms or less and trains of pulses lasting up to 30 seconds. The phasic nature of brain stimulation provides immediate washout that is not possible for pharmacological agents, which must be first metabolized and then cleared. While the transient on–off nature of brain stimulation has proved advantageous in research probing functional roles of certain brain areas, it poses a challenge to the induction of lasting effects that will persist beyond the period of the therapeutic stimulation (see Table 34.36–2).

Acute versus Prolonged Effects. Brain stimulation can have immediate or lasting effects (see Table 34.36–3). A single electrical pulse delivered at sufficient intensity can induce depolarization, trigger an action potential, release neurotransmitters at the synapse, and result in transsynaptic propagation with subsequent activation of a functional circuit. For example, brain stimulation applied to the hand area of the primary motor cortex may activate the corticospinal tract and induce a motor-evoked

potential in the contralateral hand. Such stimulation can result acutely in the induction of either a positive effect, as in the case of a muscle twitch or visualization of phosphenes, or a disruptive effect as in the case of visual masking. Although inducing a muscle twitch is an immediate, short-lasting effect, it allows one to examine the latency of this effect from the time of TMS single-pulse stimulation and thus to determine the nerve conduction delay, which may have value in the study of central motor pathways in clinical cases such multiple sclerosis and spinal injury.

Repetitive pulses delivered at fixed frequencies can exert even more powerful effects. Epstein and colleagues described in 1999 how repetitive TMS (rTMS) applied to the language-dominant hemisphere-induced arrest of the speech. Following termination of the stimulation, speech returned to normal. More recently, using interleaved functional neuroimaging with rTMS to examine network wide effects in a study with cats, investigators demonstrated behavioral impairments to visual attention with rTMS over the left intraparietal sulcus and interestingly also found downstream adaptive compensatory activity in areas of the visual attention network that were not stimulated.

Some of the more invasive brain stimulation modalities, such as DBS or VNS, are programmed to operate chronically, thus extending the acute action for as long as the stimulation is turned on. In the case of DBS, the pulses are typically given continuously at a high frequency, whereas in the case of VNS the pulses are given in trains lasting up to 30 seconds and typically repeated every 5 minutes. The less invasive modalities, such as rTMS, tDCS, CES, and even ECT, presumably require the induction of some form of neuroplasticity for their effect to become lasting.

Neuroplasticity. Persistent activation of a brain circuit can induce prolonged changes in the subsequent activities of that circuit. For example, repeated stimulation of the perforant pathway implicated in spatial memory learning and consolidation at high electrical frequency induces a lasting increase in the efficacy of that circuit; conversely, repeated stimulation with low frequencies depress it. These phenomena are termed long-term potentiation (LTP) and long-term depression (LTD), respectively, and are collectively known as neuroplasticity. Neuroplasticity is thought to occur through dynamic alterations in synaptic efficacy. Establishing evidence of longer-lasting neurochemical effects (i.e., LTP or LTD) from brain stimulation is crucial for establishing both the clinical utility and understanding the mechanism of action. Recent research in animal models offers evidence of lasting effects of high-frequency rTMS and there is evidence emerging in humans that rTMS can produce changes in synaptic efficacy that outlast stimulation induction. Initially described by Zhaoxiang Wang and colleagues, low-frequency rTMS (1 Hz or less) is believed to induce a sustained inhibitory effect on the excitability of the

motor cortex, whereas high-frequency rTMS (5 Hz or greater) produces facilitatory after-effects in excitability of the motor cortex. The work of Steve K. Esser, Reto Huber, and colleagues provided the first evidence, through coupling TMS with topographical EEG measurements of TMS-evoked potentials that these changes in excitability may come about as a result of an enhancement of a cortical potential. While such in vivo evidence is consistent with the hypothesis that TMS causes LTP without ex vivo experiments on excised human tissue, it is difficult to say with complete certainty whether these observations are true measurements of synaptic plasticity.

Table 34.36–3.
Acute and Prolonged Mechanisms of Action

Acute effects
Phasic activation of neural circuits
Observable motor responses (e.g., twitch)
Temporary disruption (e.g., speech arrest) or facilitation of ongoing processing (e.g., speeds reaction time)
Prolonged effects
Neuroplasticity
► Changes in synaptic efficacy akin to long-term potentiation or depression
► Alteration in neurotropic factors
► Modulation of cortical excitability
► Modulation of functional connectivity

Effect of Illness and Early and Late Brain Development on Brain's Response to Stimulation

Because brain stimulation exerts its acute effects by depolarization of neurons and over the longer term possibly through induction of lasting changes in synaptic efficacy, factors that influence neural transmission and plasticity may influence the brain's response to therapeutic stimulation. For example, investigators have found that infants and children have higher thresholds for activation of functional circuits in response to single pulses of TMS, presumably because myelination in their brains had not yet completed. Diseases of the central nervous system can also impact the response to stimulation. Fitzgerald and colleagues reported that patients with schizophrenia exhibited reduced plasticity in brain response to TMS when compared to healthy controls—specifically the patients failed to show inhibitory effects of low-frequency rTMS over motor cortex, possibly reflecting abnormalities in glutamergic functioning in schizophrenia.

Cortical atrophy due to normal aging or related to disease can reduce the total dosage of magnetic stimulation delivered because the strength of the magnetic field drops rapidly as a function of distance from the stimulating coil. Aging can also alter the brain's response to stimulation in a manner independent of the distance issue. Some investigators have

demonstrated an age-related differential response to rTMS over the motor cortex in older adults compared to younger subjects (45 years delineation of matched subjects), suggesting normal age-related changes in synaptic plasticity and the capacity for LTP and LTD that may reduce the efficacy of rTMS. These findings are in line with the generally accepted view that the efficacy of antidepressant treatment in the realm of pharmacotherapy is also highly age dependent (Levkovitz and Segal, 2001).

Relationship between Spatial Resolution and Invasiveness in Brain Stimulation

Each of the brain stimulation techniques can be separately classified with respect to the degree with which it can achieve precise focal targeting, that is, focality, and with respect to the degree of invasiveness, as seen in Figure 34.36–2. A high level of focality is important when brain stimulation is a tool being used to map regional brain functions and is also advantageous in clinical practice provided that a specific target is known or if stimulation of spatially adjacent regions might be counter therapeutic. On the other hand, if the target involves broad areas or is not precisely defined, then narrow focality may not offer benefits.

The least focal modalities are the noninvasive electrical stimulation methods such as tDCS, CES, and ECT. While ECT is more invasive by virtue of the need to use general anesthesia during the procedure and due to the induction of seizure, it is also nonfocal for the reasons it shares with the previous two methods—the electric current needs to traverse through the scalp and skull, which considerably impede the electric current, and which vary substantially from one individual to another.

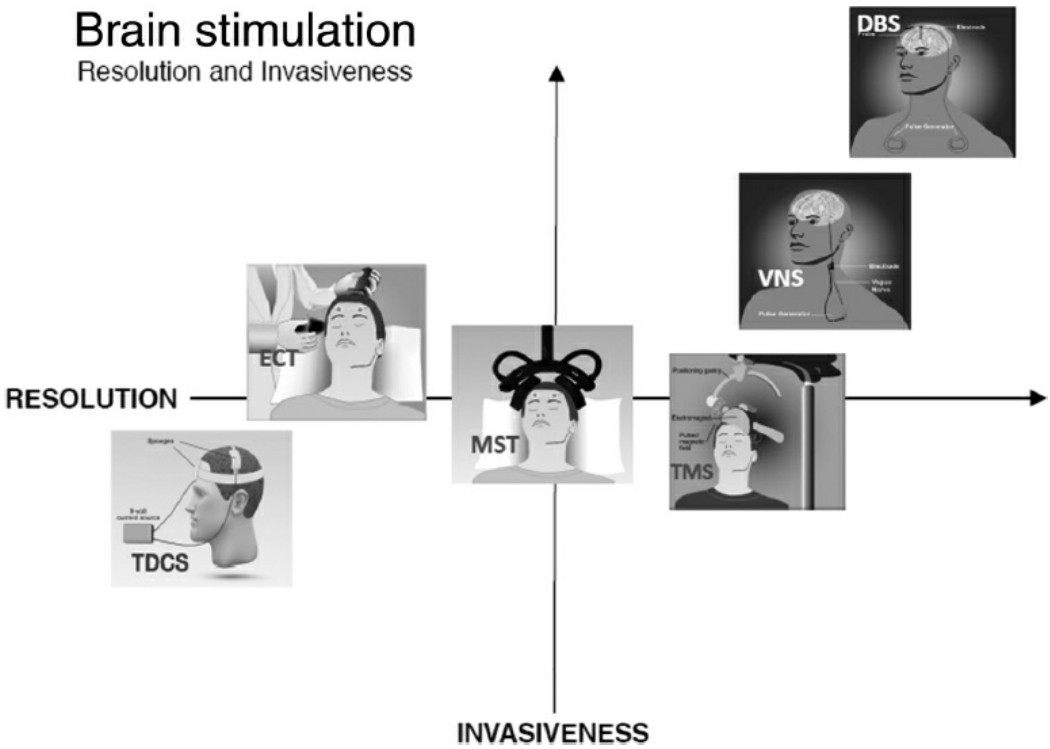

FIGURE 34.36–2. Relationship between spatial resolution and invasiveness of the brain stimulation tools. DBS, deep brain stimulation; ECT, electroconvulsive therapy; MST, magnetic seizure therapy; tDCS, transcranial direct current stimulation; TMS, transcranial magnetic stimulation; VNS, vagus nerve stimulation. (From National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services. Retrieved December 4, 2014 from http://www.nimh.nih.gov/news/image-library/index.shtml)

The most focal or precisely targeted method is DBS, where a thin lead usually containing a number of wires connecting electrodes placed close to its tip, is implanted directly into the targeted brain region. However, it is also the most invasive of all methods, requiring neurosurgical insertion of the stimulator lead, usually bilaterally, via a burr hole approximately the size of a quarter, which is drilled in the skull one for each of the leads. Somewhat less invasive but also less focal is VNS, which stimulates the vagal afferents on their way to the brain. The magnetic modalities like TMS and MST are considered more focal because the magnetic field passes through the scalp and skull without impedance, which is otherwise met by stimulation methods that directly administer electrical currents to the scalp. Magnetic flux has limited penetration (approximately between 2 and 4 cm), which is an important aspect of the stimulation's spatial resolution.

PSYCHIATRIC APPLICATIONS OF BRAIN STIMULATION

Dual Roles for Brain Stimulation: As a Probe of Brain Function in Health and Disease and as a Clinical Treatment of Psychiatric Disorders

Brain stimulation methods are used in psychiatric clinical and research practice along several parallel pathways. By virtue of their noninvasive nature, brain stimulation tools offer unique versatility as tools to examine normal brain function by transiently perturbing brain activity in creating temporary "virtual lesions" and monitoring the subsequent effects on cognition and behavior. Such investigation is particularly valuable for probing of brain areas implicated in the pathophysiology of psychiatric illness. Brain stimulation modalities are also used in the treatment of mental illness as novel therapeutic agents as shown in Figure 34.36–3.

Brain stimulation actively changes brain function and can induce measurable effects on brain activity that can be used to test causal relationship between brain and behavior. This is in contrast to neuroimaging or other methods to study brain function that passively acquire information about brain function and can only provide correlates of behavior. To demonstrate the utility of brain stimulation in testing causal relationship between brain and behavior, Pascual-Leone and colleagues in 1991 created a virtual lesion in the Broca area with rTMS, which consequently produced speech arrest in healthy participants. Using brain stimulation to investigate pathology, more recently investigators have compared the impact of rTMS on participants with autism spectrum

disorder to healthy matched volunteers to demonstrate differential effects of stimulation that suggest language-related neural network differences in those with autism spectrum disorder and neurotypical individuals.

In addition to examining brain–behavior relationships, the various stimulation modalities have the potential for delivering a therapeutic effect in the treatment of psychiatric disorders, primarily because brain stimulation modulates neuronal activity in a way that may be lasting, as reviewed in other parts of this chapter. The focal and temporal precision of some forms of brain stimulation poses a challenge of defining the disease-relevant anatomical and physiological targets for psychiatric disorders. The therapeutic use of neuromodulation in psychiatry will also require knowledge of the local and distributed effects of the various stimulation paradigms.

Integrating Brain Stimulation with Neuroimaging and Neurophysiology

The combination of brain stimulation and functional neuroimaging is a powerful research paradigm to explore and understand neural circuitry. TMS, DBS, and ECT have been successfully coupled with positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and electroencephalography (EEG) to investigate the modulatory effects of brain stimulation in both healthy and pathological populations. Commonly, physiological measures are performed sequentially *off-line* before and after brain stimulation to assess modulatory effects. However, for brain stimulation modalities such as TMS and tDCS, a simultaneous or *online approach* is now possible whereby stimulation and imaging occur in an interleaved manner. Whereas functional neuroimaging alone can indicate correlations between brain and behavior, simultaneous neuroimaging and brain stimulation can indicate causal brain–behavior relations.

With psychopathology increasingly conceptualized as dysfunction not only of one isolated brain area but rather on a network level, the combination of brain stimulation and neuroimaging is useful for examining of the dynamics of brain networks and the functional roles of network components, thus target sites for therapeutic stimulation. This is achieved by using brain stimulation to perturb or enhance one node of a network while monitoring activity in the entire brain. Results from such research allow elucidation of the functional connectivity of brain networks. For example, researchers have found that stimulation with TMS of the angular gyrus resulted in modulated activity to remotely connected nodes of the visuospatial attention network (enhanced activity in the contralateral V1 and V2), in addition to the behavioral effects of TMS. In another recent example, investigators delivered low-frequency rTMS to the intraparietal sulcus followed immediately with an attention fMRI task to demonstrate a decrement in activity in the targeted region and changes in activation of the corresponding contralateral region, as well as other regions in the visual

attention network. In this experiment, rTMS appears to have produced changes in activity not only at the site of stimulation, but also disrupted the interhemispheric balance of activity.

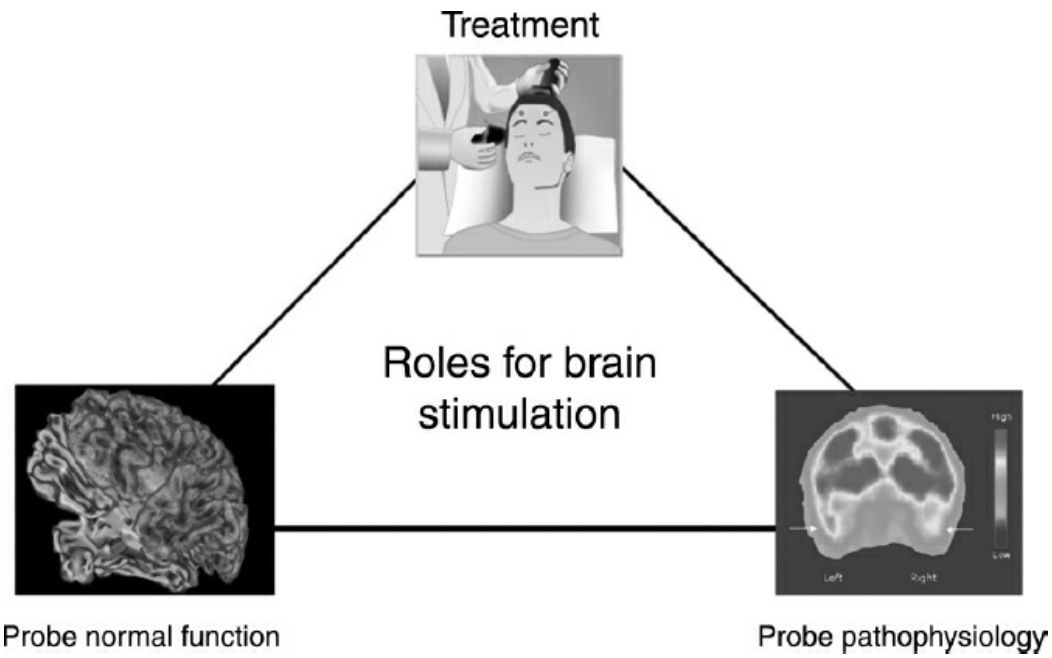

FIGURE 34.36–3. Research and clinical applications of brain stimulation in psychiatry. (From National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services. Retrieved December 4, 2014 from http://www.nimh.nih.gov/news/image-library/index.shtml)

Interleaved brain stimulation and neuroimaging also present utility in guiding and localizing stimulation with much improved accuracy rather than relying on gross structural anatomy as reference, which is susceptible to interindividual variability. In turn, accurately navigated TMS may reduce variability in clinical response.

These examples illustrate how combining brain stimulation with neuroimaging and physiology can inform exploration of mechanisms of action of various brain stimulation modalities and lead to effective screening for stimulation paradigms hypothesized to have therapeutic effects based on their ability to modulate activity of specific circuitry implicated in the illness. However, combining neuroimaging methods with brain stimulation is not without some limitations. For instance, the hemodynamic response as measured in fMRI is only an indirect measure of neural activity and neuroimaging with interleaved brain stimulation is susceptible to artifacts.

BRAIN STIMULATION TOOLS

Transcranial Magnetic Stimulation

Definition. TMS is the application of rapidly changing magnetic field

to the superficial layers of the cerebral cortex, which locally induces small electric currents, technically referred to as "eddy or Foucault currents." These are circular electric currents induced within a given conductor by a changing magnetic field in that conductor. Michael Faraday originally discovered this induction through his experiments in 1831 and later quantified in Maxwell's equations that encompass all of electromagnetism. Thus, TMS may be referred to as electrical stimulation without an electrode, as it uses magnetic fields to indirectly induce electrical pulses in the targeted cerebral tissue. TMS device delivers strong magnetic pulses via a coil that is held on the scalp. Since magnetic fields are unaffected by the electrical impedance of the scalp and skull, this method of stimulation enables the focal stimulation of smaller areas of the brain than is possible with other noninvasive devices which use either alternating (ECT, CES) or direct (tDCS) electrical current for primary stimulation. TMS is an example of noninvasive, stimulation of focal regions of the brain delivered at levels well below subseizure threshold and as such, can be used for research or therapeutically without the need for anesthesia. It is a relatively new treatment modality in the psychiatric practice, with the original device approved by the FDA in 2008 for treatment of depression for patients who have failed a trial of one, but not more than one, antidepressant medication like an SSRI. This initial indication has been since expanded to include another device and to treat adult patients with major depressive disorder (MDD) who have failed to benefit from any number of antidepressant medications.

TMS, when given one pulse at a time without regular rhythm and at low rates of delivery, is referred to as single-pulse TMS. When repeated at regular rhythmic intervals that usually range from 1 to 25 Hz, this stimulation method is described as rTMS, which can be further functionally divided into "low-frequency rTMS" when stimulation frequency is 1 Hz or less, or "high-frequency rTMS" when higher frequencies are used, usually 5 to 20 Hz. Theta-burst stimulation (TBS) refers to high-frequency bursts (e.g., 3 pulses at 50 Hz) applied at a regular frequency (e.g., 5 bursts per second). TBS can be given continuously (cTBS) or intermittently (iTBS). cTBS is reported to induce inhibitory effects while iTBS is reported to exert excitatory effects.

History. In 1985, Barker, Jalinus, and Freeston from the University of Sheffield in the United Kingdom reported the first use of TMS to stimulate the human brain. Early stimulators were capable of producing only low frequencies, typically less than 0.3 Hz, but it was the single-pulse TMS that in those early explorations proved to be very useful in the study of motor conduction and mapping basic brain processes. While these studies used large round coils that lacked the focality of today's technology, even this early work with low-frequency stimulation was promising enough in the

experimental treatment of depression and schizophrenia to encourage further work and development of TMS as a promising treatment modality in psychiatry.

With the subsequent development of devices capable of much higher repetitive rates up to 20 Hz—higher brain functions like language, mood, and memory became accessible to stimulation. In 1996, Pascual-Leone and colleagues reported that antidepressant properties of daily TMS stimulation for 5 days to the dorsolateral prefrontal cortex (DLPFC) might have a prolonged effect that maintains beyond the acute period of stimulation. Studies such as this motivated an intense interest in the potential therapeutic value of focally applied rTMS targeted to the neuronal circuitry underlying not only depression but also other psychiatric disorders like anxiety disorders, schizophrenia, and others.

Mechanisms of Action. By Faraday's law, the induced electromotive force in a closed loop is directly proportional to the time rate of change of magnetic flux through the loop. As applied in TMS, changing magnetic field induces electrical current in the circuitry of the superficial cortex underlying the TMS coil. This electrical current is commonly referred to as "eddy current" for its circulating motion in a plane perpendicular to the plane of the magnetic field (also known as Foucault currents named after the French physicist who described the phenomenon). At sufficient intensity, electrical currents across the cerebral cortex affect cell electrochemical gradients and produce local neuronal depolarization, which may result in a propagated action potential. For example, when the TMS coil is positioned over the hand area of the cerebral cortex's motor strip, changing magnetic field generated by the repetitive pulses induces local currents immediately below the site of stimulation that cause the neurons in area M1 to fire. In turn, this action potential propagates through the polysynaptic corticospinal tract and results in a twitch in the contralateral hand muscle, which can be used to study the excitability of the corticospinal system. In summary, TMS uses magnetic field to indirectly induce focal electrical currents in the brain, thereby triggering the firing of functional neuronal circuits that can lead to observable behavioral effects. This effect can be easily demonstrated by simply moving the TMS coil across the cortical representation of neighboring muscle groups (the homunculus) with a series of single TMS pulses.

Single TMS pulses can exert other effects when moved to different cortical areas. When positioned over the primary visual cortex (V1), scotomas, or "blind spots" are often elicited. This effect was referred to as "visual masking" or creating a "virtual lesion" and it illustrates that TMS can transiently disrupt functions.

Activation of motor neurons resulting in a muscle twitch or disruption of visual perception with single-pulse TMS represents examples of the

acute effects of TMS-induced neuronal depolarization as shown in Table 34.36–3. The effects of single TMS pulses are believed to be immediate and short lived. The muscle twitch as induced by TMS to area M1 is nearly instantaneous, with the hand movement occurring approximately 20 ms after the TMS pulse is applied. Both the motor-evoked potential and visual masking operate on a similar timescale measured in milliseconds. While producing such a motor-evoked potential does not require the patient to engage in activity, a single magnetic pulse does not typically affect cognition by itself. Rather, interference or improvement of performance in a cognitive task is observable when a train of magnetic pulses are delivered immediately preceding or during the time when the participant is engaged in a cognitive task, such as in visual masking.

Traditionally TMS coils were limited to direct stimulation of the cortical surface, activating neurons at 1.5 to 3.0 cm beneath the scalp rather than deeper subcortical limbic structures. However, recent development and research of coil designs for stimulation of deeper brain areas, such as the H-Coil, suggest TMS may stimulate subcortical structures up to 6 cm beneath the scalp. The cost of deeper stimulation, however, may be spatially wider electrical field distribution in the brain (i.e., reduced focality).

TMS can exert longer-lasting effects when the pulses are repeated at regular intervals in a process of rTMS, or when TMS pulses are paired with other forms of stimulation, such as electrical stimulation of a peripheral nerve (as in paired associative stimulation), or when TMS is coupled with audiovisual stimuli, as in the example of classical conditioning of the brain response to TMS. The mechanisms underlying these lasting effects of TMS have been described by various researchers and are thought to be related to neuroplasticity and alterations in synaptic efficacy.

Treatment of psychiatric disorders with rTMS has been informed by attempts to focally alter pathological cortical excitability whereby irregular excitability has been linked to a specific illness. For example, reduced activity in the left DLPFC has been implicated in several studies as a physiological correlate of affective disorders. To correct this, numerous studies have applied high-frequency rTMS, which purportedly increases excitability, to the left DLPFC in an attempt to normalize activity in this region. In a related approach, some investigators who implicated abnormal interhemispheric balance in activation between the right and left DLPFC to mood psychopathology, have applied low-frequency rTMS to the right DLPFC, in an attempt to normalize this balance with stimulation trains believed to be inhibitory.

The potential therapeutic role of developing rTMS to focally modulate cortical excitability has been validated in animal studies, which suggests that rTMS, like ECT and antidepressant medication, can exert effects on neurotransmitter release, receptor expression and gene expression.

However, animal models of the effects of rTMS do not perfectly replicate the effects of rTMS on the human brain due to size differences and consequently how they impact the strength and distribution of the induced electric fields.

Side Effects, Interactions with Medications and Other Risks. Administration of TMS is a noninvasive, relatively benign procedure when applied by a knowledgeable professional to a subject who has been properly evaluated. However, it is not entirely without risk. The most serious known risk of TMS is an unintended seizure, but this risk varies depending on stimulation parameters and patient history. Namely, single-pulse stimulation is less likely to result in a seizure compared to rTMS and variations in dosage—the combination of frequency, power, train duration, and intertrain interval—may also change the risk of unintended seizure. In addition, subject factors such as the presence of a neurological disorder (epilepsy or a focal brain lesion) or use of seizure lowering medications may also result in greater likelihood of an unintended seizure. The FDA's safety guidelines for avoiding seizure with rTMS are summarized in Table 34.36–4.

Basic Studies. TMS has helped advance the basic understanding of normal brain function, as well as the pathophysiology of psychiatric illness. The combination of noninvasiveness, high temporal resolution and particularly the ability to make use of frequency-dependent effects are significant advantages of TMS. Furthermore, the ability for TMS to be coupled with neuroimaging to explore the functional role of areas of activation seen with PET or fMRI imaging makes it a valuable tool for understanding brain function (see Fig. 34.36–4). For example, Luber and colleagues reported on the use of fMRI-guided stereotaxically applied rTMS to test the functional role of brain regions implicated in working memory, finding that rTMS applied to the precuneus facilitated working memory performance. Such an approach informs one's understanding of the brain networks involved in normal brain function. Investigators have also delivered high-intensity TMS to participants during functional magnetic resonance image acquisition to demonstrate that stimulation of the left motor cortex was correlated with distributed cortical and subcortical activity in the bilateral premotor cortex, bilateral primary somatosensory cortex, bilateral second somatosensory cortex, left medial pontine nuclei, bilateral cerebellum, deep cerebellar nuclei, thalamic nuclei, and the basal ganglia (caudate and putamen). From such correlated changes in activity researchers are able to postulate brain areas that form distributed networks.

Table 34.36–4.

Safety Guidelines for Repetitive Transcranial Magnetic Stimulation

	Intensity (% of Motor Threshold)													
Frequency (Hz)	**80–100**	**100**	**110**	**120**	**130**	**140**	**150**	**160**	**170**	**180**	**190**	**200**	**210**	**220**
1	>1,800	>1,880	>1,800	360	>50	>50	>50	>50	21	11	11	8	7	6
5	>10	>10	>10	>10	>10	7.6	5.2	3.6	2.6	2.4	1.6	1.4	1.6	1.2
10	>5	>5	>5	4.2	2.9	1.3	0.8	0.9	0.8	0.5	0.6	0.4	0.3	0.3
20	.05	2.05	1.6	1.0	0.55	0.35	0.25	0.25	0.15	0.2	0.25	0.2	0.1	0.1
25	1.28	1.28	0.84	0.4	0.24	0.2	0.24	0.2	0.12	0.08	0.12	0.12	0.08	0.08

This table gives the duration in seconds that could be administered at each combination of frequency and intensity (given in terms of percentage of the motor threshold) prior to the emergency of spread of excitation, a warning sign of seizure. Values preceded by > indicate that this is the longest duration tested.

Adapted from FDA's Maximum Safe Train Duration (seconds) Limits for Avoiding Seizure (2014) in Guidance for Industry and FDA Staff - Class II Special Controls Guidance Document: Repetitive Transcranial Magnetic Stimulation (rTMS) Systems; Wassermann EM. Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the International Workshop on the Safety of Repetitive Transcranial Magnetic Stimulation. *Electroencephalogr Clin Neurophysiol.* 1998;108(1):1–16; Rossi S, Hallett M, Rossini PM, Pascual-Leone A; Safety of TMS Consensus Group. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin Neurophysiol*. 2009;120(12):2008–2039.

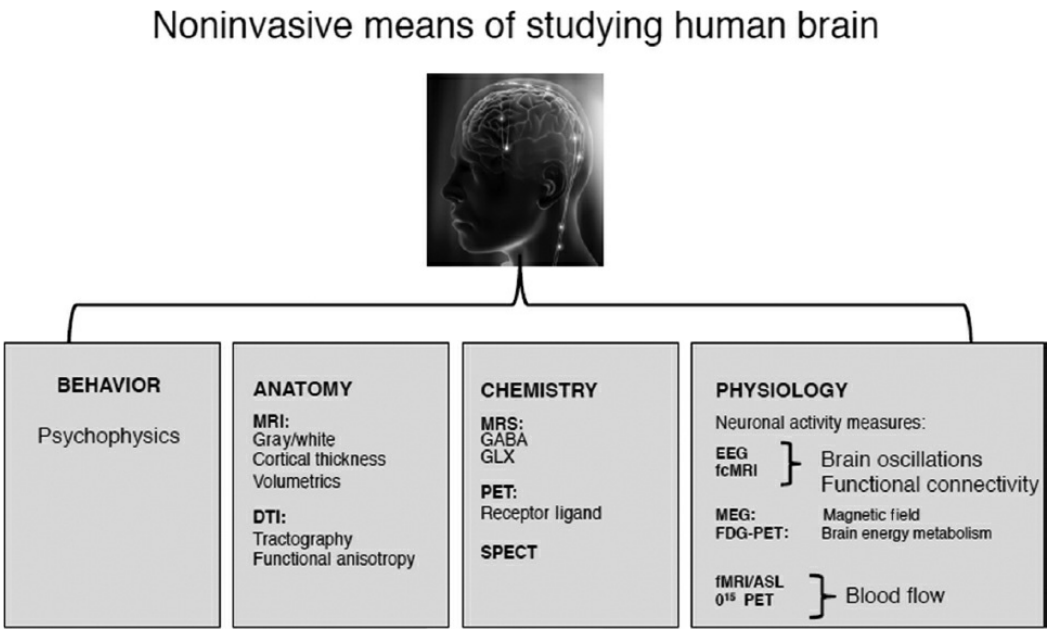

FIGURE 34.36–4. Noninvasive means for studying human brain function. The human brain can be *passively* studied in terms of its anatomy, chemistry, physiology, and behavior. However, only brain stimulation can *actively* change brain function, and thus can serve as a probe to test causal relationships. ASL, arterial spin labeling; DTI, diffusion tensor imaging; EEG, electroencephalography; ERP, event-related potentials; FA, fractional anisotropy; FDG, flurodeoxyglucose; fMRI, functional magnetic resonance imaging; MEG, magnetoencephalography; MRI, magnetic resonance imaging; MRS, magnetic resonance

spectroscopy; PET, positron emission tomography; SPECT, single photon emission computed tomography. (From Columbia Psychiatry [Online image]. (2013) Retrieved March 20, 2015 from http://columbiapsychiatry.org/disorders/brain-stimulation)

Clinical Studies. The use of rTMS in the treatment of MDD has been the most-studied therapeutic application of this technique. Evidence from a growing number of randomized controlled trials suggests a significant antidepressant effect with rTMS applied to the DLPFC. The largest study to date has been a multicenter trial conducted at four U.S. university hospital clinics reported by George and colleagues that included approximately 860 antidepressant medication-free patients with unipolar MDD who had not benefited from prior treatment. A total of 190 participants randomized to either active or sham treatment were included in the final analysis. Participants in the active treatment arm received daily left prefrontal rTMS for 3 weeks, with treatment standardized at 120 percent of the magnetic field intensity relative to the patient's resting motor threshold, at 10 pulses per second (10 Hz) for 4 seconds, and an intertrain interval of 26 seconds. The authors reported that TMS was effective in treating major depression with minimal side effects. Not all studies have been positive, however, and the average effect size in depression is moderate—less than with ECT, but comparable to that seen with traditional pharmacotherapy.

Use of TMS in the treatment of schizophrenia has yielded some promising results in experimental studies that used either low-frequency rTMS to inhibit temporal-parietal areas of the cortex linked to hyperactivity during auditory verbal hallucinations or high-frequency rTMS to the prefrontal cortex aimed at modulating dopamine hypofrontality linked to negative symptoms. In the first double-blind randomized trial of deep TMS for the negative symptoms of schizophrenia, investigators found statistically significant reduction in negative symptoms after treatment, although there was no statistically significant difference between the active treatment and sham groups. Ongoing research has also demonstrated some positive results of TMS as a treatment modality in anxiety disorders, including applications for obsessive-compulsive disorder, panic and Tourette disorders.

Current Status in Treatment Algorithms. In 2010, the APA included TMS within their Practice guidelines for the treatment of patients with MDD (3rd edition) recommending TMS as an initial treatment modality in the acute phase of depression. There are two FDA-approved TMS devices for use in adult patients with major depressive disorder: the NeuroStar TMS device manufactured by Neuronetics Inc. was approved in 2008 with the following FDA-approved indication: "for the treatment of MDD in adult patients who have failed to achieve satisfactory improvement from one prior antidepressant medication at or above the minimal effective dose and duration in the current episode"; and the Brainsway Deep TMS device that

excites or inhibits neurons deeper inside the brain by using an H-Coil was approved in 2013. The FDA-approved indication appears to be broader: "for the treatment of depressive episodes in adult patients suffering from MDD who failed to achieve satisfactory improvement from previous antidepressant medication treatment in the current episode." The Neuronetics device delivers 10-Hz frequency pulses for trains of 4-second duration with an intertrain interval of 26 seconds and a typical session duration of 37.5 minutes (total of 5 daily sessions of 3,000 pulses for 6 weeks), while the Brainsway device delivers pulses with 18-Hz frequency for 2 seconds with an intertrain interval of 20 seconds and total train duration of 20.2 minutes (5 daily sessions of 1,980 pulses for 4 weeks, followed by optional another 12 weeks of bi-weekly maintenance treatments). A treatment course with either device is applied to prefrontal cortex. rTMS treatments would most likely be considered as an alternative treatment modality to medications, prior to more invasive methods like ECT.

Table 34.36–5.
Transcranial Magnetic Stimulation Dosing Parameters

Scalp position, target localized method
Coil shape and orientation; number of coils
Circular; figure-eight; double-cone; H-coils; B-shaped; cloverleaf
Single pulse vs. paired pulse vs. repetitive vs. convulsive
Train: Frequency, duration, intensity, total pulses
Fast/high frequency at rates >1 Hz;
Slow/low frequency at rates <1 Hz;
Quadripulse repeated trains of four monophasic pulses separated by interstimulus interval
Intertrain interval
Number of trains/session
Number of session/day
Number of days/course
Taper/maintenance/reintroduction

Patient Selection. Patients who have failed a trial of one or more antidepressant medications are medication intolerant, and pregnant or postpartum patients avoiding use of medications may be appropriate candidates for TMS (Note: treatment of patients with Neuronetics device who failed two or more medication trials in the current episode, would be considered off-label based on its current FDA-approved indication). However, given the lower effect size of TMS, for urgent or severely refractory cases, ECT would remain the ultimate "gold" standard treatment.

Dosing. The major characteristics of TMS dosing are shown in Table 34.36–5. Each device's technical parameters, such as stimulation intensity, pulse frequency, intertrain interval, and total number of pulses

administered are the most basic adjustable elements of TMS dosing. In addition, coil shape and position relative to the target cortical area as well as the waveform shape and polarity, are also of important specific consideration (see Fig. 34.36–5). In general terms, the delivered intensity of rTMS is related to the distance of the coil from the target area of the cortex. Thus, increasing the distance decreases the intensity of the stimulation able to reach the brain. Such decreased intensity has been found to be negatively correlated with the antidepressant response and with the degree of brain activation as evidenced by functional imaging. It is important to note that despite the many aspects that go into each device's dosing characteristics, considerable individual variability in the response to TMS has been reported.

Future Directions and cTMS. TMS and other forms of magnetic stimulation hold a tremendous promise in psychiatric treatment due to their focality and noninvasiveness. However, much research is still needed to replicate preliminary findings with improved stimulators, optimize dosing, establish the patient characteristics that predict response, and examine the influence of concomitant medications on TMS effect. Research on the durability of clinical benefits of TMS is ongoing with some promising results. Using a large cohort of patients, investigators recently followed patients after their treatment for depression with TMS for 12 months posttreatment and found statistically and clinically meaningful durability of clinical benefits.

Other pursued directions are modification of pulse shapes to optimize stimulation as seen in controllable pulse shape TMS (cTMS), a device type that can generate rapid rate (<50 Hz) trains of mono/bi/polyphasic electrical pulses that are near rectangular in shape and vary in width, directionality, and amplitude. This work takes after the discovery in 1990s in the realm of ECT in which square or rectangular waves of variable width replaced the sinusoidal waveforms, generally now abandoned due to their inefficiency and creation of excessive cognitive side effects. In the first study examining cTMS use in human subjects, investigators found the cTMS device generated a broader range of pulse widths compared to conventional TMS devices, which these authors suggest could enable more accurate study of the pulse width dependence of cortical excitability.

Transcranial Direct Current Stimulation

Definition. tDCS is a noninvasive and nonconvulsive form of brain stimulation in which a very weak (0.5 to 2 mA) direct electrical current is applied using at least two surface scalp electrodes of which optimal placement continues to be investigated. This direct current (DC) form of electrical stimulation that can be distinguished from most other devices that stimulate with alternating currents in the brain such as CES, TMS,

ECT, MST, VNS, and DBS, does not directly result in neuronal firing by suprathreshold neuronal membrane depolarization, but rather tDCS elicits polarity-dependent alterations in the resting membrane potentials. These facilitating or inhibiting effects can last for up to 1 hour after stimulation and are thought to induce neuronal plasticity. The device is very portable and usually operated by 9V DC batteries.

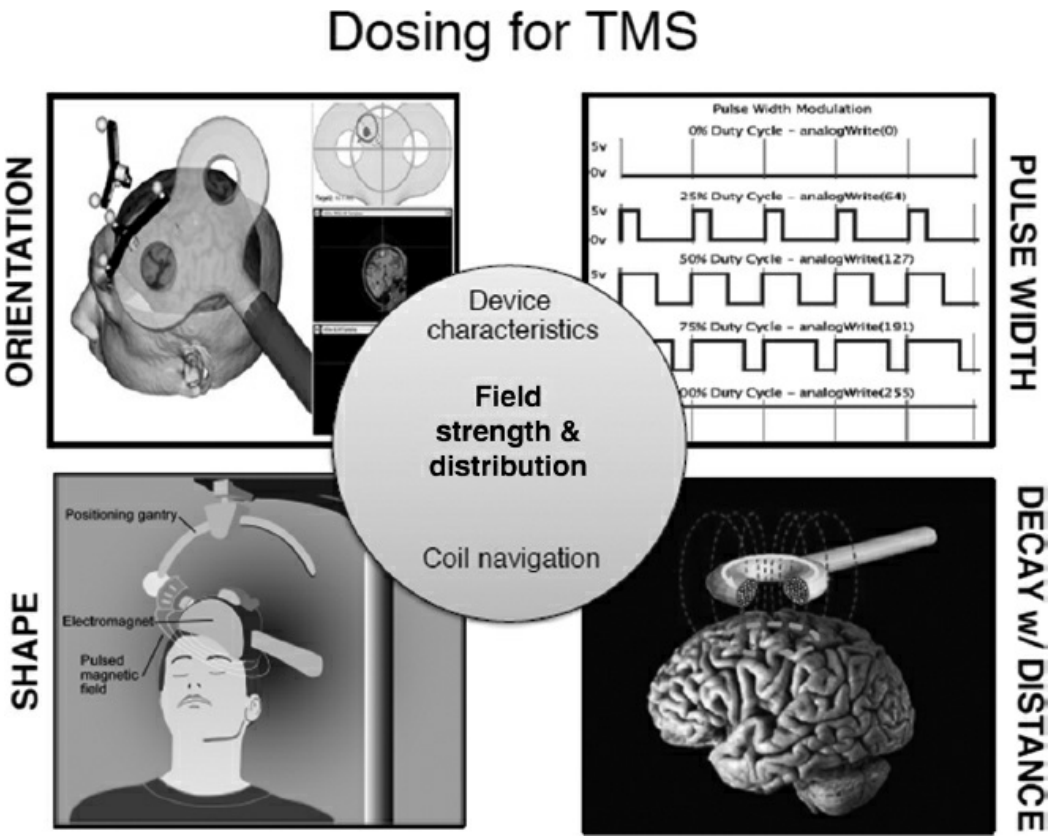

FIGURE 34.36–5. Transcranial magnetic stimulation dosing parameters. Images clockwise from top left: NITRC [Online image] (2015). (Retrieved March 16, 2015 from http://www.nitrc.org/project/list_screenshots.php?group_id=499&screenshot_id=374; Width Modulation [Online image] (2015). Retrieved March 16, 2015 from http://arduino.cc/en/Tutorial/PWM; Wasserman, E. via Wikimedia Commons [Online image]. Retrieved March 16, 2015 from http://commons.wikimedia.org/wiki/File%3ATranscranial_magnetic_stimulation.jpg; National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services. Retrieved December 4, 2014 from http://www.nimh.nih.gov/news/image-library/index.shtml)

History. There is history of ancient Egyptian, Roman, and Middle Eastern use of live electric catfish and torpedo fish as crude therapeutic interventions that involved use of DC stimulation for headaches and epilepsy. It was not until the 1960s that tDCS was empirically investigated. In Great Britain, Lippold and Redfearn (1964) reported that positive (*anodal*) currents induced elevated mood and motor activity and reduced anhedonia, whereas negative (*cathodal*) currents induced withdrawal and apathy. While subsequent attempts to replicate these findings failed, since

the 1980s there has been a growing body of evidence of positive effects from tDCS in cognitive and motor function. Of late, tDCS has experienced resurged interest due to its ease of application, small device size, low cost, and relative lack of risks and side effects as compared with treatments like ECT and TMS.

Side Effects. There are no known serious adverse effects of tDCS. It is well tolerated with low prevalence of mild adverse effects reported in the literature, which include mild tingling sensations, light itching sensation, skin irritation, mild burns, and headache.

Mechanism of Action. The exact mechanism of tDCS is assumed to be modification of the synaptic microenvironment—weak direct current shift of the resting membrane potential, either lowering or raising the threshold of activation—affecting the firing and conductance of neurons. Specifically, anodal electrodes excite the neuron shifting the resting membrane potential toward depolarization, while cathodal electrodes inhibit the neuron shifting potential toward hyperpolarization—effects which may be dependent on the subtype of cortical neuron being stimulated. Because tDCS works via polarization and does not affect action potential firing in cortical neurons, the term "transcranial direct current polarization" is favored by some investigators and both terms appear interchangeably in the literature.

Clinical Studies. Much of the experimental research on tDCS has been conducted with healthy controls demonstrating the effectiveness of this stimulation method in exciting and inhibiting brain activity. Preliminary research has also examined the effects of tDCS for neuropsychiatric disorders such as depression, schizophrenia, pain disorders, and epilepsy. While early research on the application of tDCS for depression produced mixed results, methodological improvements have reignited interest in research of tDCS as a therapeutic technique.

Current clinical psychiatric research is focusing on the potential effectiveness of tDCS in reducing depressive symptoms. A recent meta-analysis of seven randomized sham-controlled studies demonstrated significant antidepressant effects after left prefrontal anodal tDCS with a reduction of symptoms as measured by the Hamilton Rating Scale for Depression (HRSD) ranging from 14.6 percent to 80 percent in the active treatment group versus approximately 10 percent symptom reduction in a group receiving sham tDCS. In terms of the durability of this clinical response, four of the seven studies reported that antidepressant effects were sustained for 1 month after the end of treatment in the active DLPFC group. Further research is needed to replicate such findings.

Current Status in Treatment Algorithms. tDCS is not approved by FDA for psychiatric treatments and is thus considered an experimental research technology with much work needed to demonstrate its efficacy. Should sufficient evidence be attained and the device approved by the regulatory authorities, tDCS could represent an inexpensive and relatively safe alternative form of treatment to medications, prior to the more invasive stimulation methods like ECT or VNS. Within research, the typically applied stimulation protocols are two cephalic electrodes positioned below the cranial foraminae and fissures at 1- to 2-mA intensity, with stimulation up to 20 minutes per session.

Future Directions. Most of the current tDCS devices utilize large, saline-solution–soaked electrodes. Recent development of high-definition tDCS (HD-tDCS) devices that use small arrays of "high-definition electrodes" are under investigation for their increased focality compared to standard bipolar montages of traditional tDCS devices. Results of a HD-tDCS 4 × 1 montage that consists of once center electrode over the target and four surrounding return electrodes is promising in terms of increased excitation of a target area.

In general, current knowledge of tDCS safety, efficacy, indications, dose–response relationship, as well as predictors of response is limited and further exploration is needed.

Cranial Electrical Stimulation

Definition. CES, like tDCS, uses a weak (1 to 4 mA) current. However, with CES the current is alternating. It is traditionally applied via saline-soaked, felt-covered electrodes clipped onto the earlobes. Other placement strategies are also being currently investigated. For example, a study currently under way at Columbia University directed by Peter Bulow utilizes a cathode placed over the scalp area immediately in the vicinity of DLPFC and anode placed over right arm. The current of CES is in a similar range to that of tDCS with the major difference being that the current is alternating, creating either sinusoidal or modified square waves with a frequency range of 0.5 to 15,000 Hz, depending on the type of the stimulator.

History. CES was originally developed in 1949 in the former Soviet Union and used as an aid in inducing sleep. It was thus called initially "Electrosleep." In addition to the treatment for insomnia, it was also utilized as an aid in anesthesia, because of the observation that even when the current was reduced to a point of not causing excessive sleepiness, some patients seemed to benefit from the anesthetic effect of the stimulation. The device was brought to the United States and early research in America dates back to the 1960s. It was grandfathered-in by the FDA as

a Neurological Therapeutic Device for the treatment of insomnia, anxiety, and depression, and stress and the name CES, for Cranial Electrotherapy Stimulator, was suggested. In the 1970s CES was used by many US clinicians to treat substance abuse and substance abstinence syndrome in the belief that patients with a history of abuse were more susceptible to cross-addiction to psychotropic medications and because they were thought to be more resistant to the effects of such medications when compared to nonaddicted persons.

Mechanism of Action. The exact mechanism of action has not been elicited at this time and there is no agreement among researchers on the predominant mode of action. Previous hypotheses proposed that the stimulation with the alternating micro current affects the thalamic and hypothalamic brain tissue and facilitates the release of neurotransmitters. Claims have been made that through interaction with cell membranes, the stimulation produces changes in signal transduction associated with classical second messenger pathways including calcium channels and cyclic AMP. The most promising research to elucidate the mechanism of action comes from a neuroimaging study where researchers concurrently stimulated participants with CES. In this study, investigators found that CES at both 0.5 and 100 Hz (two commonly used frequencies in clinical practice) deactivated frontal and parietal area and significantly altered connectivity in the Default Mode Network. These researchers suggest CES may alter brain oscillation patterns, which is in line with previous EEG research of CES. While such research is promising, possible therapeutic effects are still not well understood.

Side Effects. It is believed that the CES stimulation is not harmful, primarily due to its low voltage power supply (9-V battery) and lack of any reported adverse event by the FDA. Local skin irritation, headache, nausea, as well as general feeling of dizziness have been reported. Device manufacturers do not advise the use of the device during pregnancy, in those with low blood pressure, or people who have arrhythmias or pacemakers.

Clinical Studies. While there is over 40 years of research on CES, many studies lack a sham comparator or control measure and many are open-label studies that lacked rigor. A recent Cochrane Review suggests that at the current time there are no high-quality studies comparing the use of CES with a sham for treatment of depression that warranted inclusion in the review, thus asserting that there is insufficient evidence to support the use of CES in treating acute depression.

In a pilot study using a small sample of 12 patients receiving CES for generalized anxiety disorder, investigators recently demonstrated that 6

weeks of CES at 0.5-Hz frequency and 300-μA intensity resulted in significant decreases in anxiety scores as measured by the Hamilton Rating Scale for Anxiety. While promising, much more research is needed.

Current Status in Treatment Algorithms, Patient Selection, and Dosing. Although the FDA approves a number of CES devices, the use of CES has not been studied sufficiently in the United States and it does not have a specific place in any algorithm of standard US psychiatric practice. Routinely, CES used for depression is delivered daily for a minimum of 3 weeks followed by a reduced schedule as needed for up to a year. It is reported to be generally well-tolerated, noninvasive treatment.

Future Directions. As with tDCS, basic issues of indications, patient selection, dose–response relationship, and efficacy are currently under active research and remain to be optimized.

Magnetic Seizure Therapy

Definition. MST is a novel form of a convulsive treatment, currently under development in several research institutions in the United States, Canada, England, Germany, and Australia. The treatment employs alternating magnetic field that crosses the scalp and the calvarium unaffected by their high electrical impedance, to in turn induce a more-localized electric current in the targeted regions of the cerebral cortex than is possible with ECT. The aim is to produce a seizure whose focus and patterns of spread may be controlled. Numerous publications characterize MST as an attempt to combine the potential benefits of unimpeded focality provided by magnetic fields, with the powerful antidepressant efficacy of ECT.

MST is a convulsive treatment, in many ways similar to ECT. It requires approximately the same preparation and infrastructure as ECT. However, MST is administered using a modified TMS device, one that can administer higher output than the conventional TMS devices, and thus relies on magnetic stimulation to induce a seizure, unlike external electrode electric stimulation in ECT. The MST procedure is performed under general anesthesia with a muscle relaxant. MST is at the stage of clinical trials, and is not presently FDA approved. Current studies concentrate on optimizing the efficacy through treatment applied to the putative depressive centers of the brain while avoiding those brain areas involved in memory formation and consolidation (medial temporal lobes) with the aim of lessening the acute cognitive side effects possible with ECT.

History. MST was first described as a convulsive treatment by Lisanby and colleagues in 2001. A 20-year-old inpatient with a medication-resistant depressive episode was reported to have been successfully treated in Bern,

Switzerland. The treatment resulted in a 50 percent reduction in depression scores, was well tolerated and had no adverse effect on Mini Mental State Exam Scores. Prior to this human treatment, magnetically induced seizures were first performed in rhesus monkeys in 1998 with a custom-modified TMS device, and later more extensive safety testing was completed with nonhuman primates.

The first trial of MST in the United States was performed at Columbia University in New York and reported fewer acute cognitive side effects with MST than with ECT in a within-subject, cross-over, blinded trial. The second center to perform MST in the United States was University of Texas, Southwestern Medical Center. Internationally, small human trials are either ongoing or have been recently conducted at centers in Canada, Australia, Wales, Scotland, and Germany.

Mechanism of Action. Induction of a seizure is hypothesized to be the underlying event responsible for the mechanisms of action of MST. As in ECT, these are not fully understood. However, due to its focality, MST appears to represent a tool better suited than ECT to study the mechanisms of action of convulsive therapy, through its potential of inducing seizures initiated in different regions of the brain. This focality of MST allows systematic investigation of the coupling between the site of seizure initiation and outcome. Furthermore, the quality of seizures induced by MST is believed to be distinct from those induced by ECT. MST-induced seizures show less impact on parasympathetic outflow, less generalization to hippocampus and deeper brain structures, and result in less serum prolactin surge.

Side Effects. Adverse effects from MST are largely connected to the risks associated with anesthesia and generalized seizure. In addition, MST magnetic coil produces a clicking noise that may potentially affect hearing. To mitigate that risk and prevent any cumulative damage both the patient and members of the treating team should wear earplugs.

Studies suggest MST results in less retrograde and anterograde amnesia than ECT. Specifically, investigators have demonstrated that there is significantly less acute memory disruption (i.e., anterograde amnesia on treatment days) with MST compared to right unilateral ECT. Though promising, this result should be replicated in a larger trial.

Currently, MST seizures are less well generalized and tend to produce less robust motor convulsions. For this reason, the dosage requirement of the muscle relaxant (usually succinylcholine) is lower than with conventional ECT. Shortening the period of paralysis could reduce the risks of the procedure by reducing the period of respiratory suppression. It should be noted however, that the attenuated robustness and generalization of the MST-induced seizures may be a function of cautious

dosing with new technology. Further studies are needed, both to in device development and efficacy dosing.

Clinical Studies. MST was used to treat a human patient for the first time in 2000 in Bern. A depressed patient received four convulsive sessions induced by a prototype MST device by Magstim (Magstim, Whitland, Wales, UK) using frequency of 40 Hz at 100 percent of that prototype machine's power intensity. Soon after, a second human treatment case was reported by the same center with both patients describing posttreatment relief of their depressive symptoms without experiencing any significant side effects. In 2003 after the first blinded, randomized, within-subject, crossover trial in 10 human patients, Lisanby and colleagues reported that there were less acute side effects of MST compared to ECT. A custom-modified Magstim device was used in this trial with the frequency increased to 50 Hz, a peak magnetic field of 1.2 T produced for 8 seconds, and a pulse width of 500 μs. One of the more important observations made in this trial was that although seizures were induced in all 10 patients, the seizure threshold was attained in several cases only at the very maximum of that device's power output. Since this experiment, there have been technological improvements such that MST devices are now capable of continuous stimulation at 100 Hz for up to 10 seconds (Magstim device) and even up to 250 Hz for a limited stimulation time (MagVenture device). It has been shown that antidepressant effects of convulsive therapy require a suprathreshold stimulation requiring at least doubling, and in some cases even six times more, the seizure-threshold power in order to achieve an optimal antidepressant result. Thus, the technological improvements in MST more reliably ensure induction of a seizure and may optimize clinical effects.

Convulsive treatment at 100 Hz was first used in 2004 in an ongoing study of rhesus monkeys, and starting in 2007, in human trials. Published results that compared MST to ECT in the nonhuman primates continue to show a superior safety profile of MST, even with the newer high-dose MST device prototypes. Pilot trials investigating the efficacy and safety of the 100-Hz MST in human patients demonstrate promising results. Most recently, investigators at Australia's Monash University and Alfred Psychiatry Research Centre reported on an open-label trial of 13 patients receiving three MST per week (a maximum of 18 treatments were provided) with a MagVenture device. These researchers found five out of the twelve patients who completed the study experienced a 50 percent reduction in depressive symptoms, and none of the 13 patients receiving treatment demonstrated changes in cognitive functioning after treatment.

A randomized controlled trial comparing the efficacy of MST to ECT in 20 patients with treatment-resistant depression was reported by Sarah Kayser and colleagues at the University Bonn. Six of ten patients receiving

MST in this study demonstrated a clinical response to treatment with 50 percent reduction in depressive symptoms as measured by the Montgomery Åsberg Depression Scale and the HRSD and there was no significant difference in efficacy or cognitive side effects when compared to the ten patients who received ECT. Notably, the patients in this study receiving MST were significantly quicker to recover and reorient immediately after treatment and did not experience the nausea, headache, or muscle pains that were evident in the ECT treatment group.

There is currently an ongoing double-blind human trial comparing MST to right unilateral ECT in treatment of severe depression in older adults at New York State Psychiatric Institute/Columbia University, an ongoing clinical trial of MST treating refractory depression, schizophrenia, and obsessive-compulsive disorder in Toronto, Canada, and also ongoing MST research in Australia at Monash University and Alfred Psychiatry Research Centre.

Current Status in Treatment Algorithms. No clinical algorithms exist for MST, as it is still an investigational protocol and treatments outside of research are not FDA approved. There are two MST devices being used within research: the MagPro MST device with Twin Coil and Dual Coil manufactured by MagVenture that delivers a biphasic waveform and pulse width of 370 μs allowing a maximal stimulation with 2,000 pulses at 100 pulses per second at 100 percent energy output (effective 70 to 75 percent energy due to power roll off at those levels) or 1,000 pulses at approximately 67 Hz at 100 percent energy output; and a MagStim MST device manufactured by MagStim in the United Kingdom which has a maximum capacity of 1,000 pulses at 100 Hz with a single round coil. As in ECT, where parameters such as electrode placement, pulse width, and dose are essential to maximize clinical benefit while minimizing risks of cognitive side effects, stimulation parameters, and coil placement are also important considerations in MST. Assuming the hypothesis that MST can approach the efficacy of ECT with fewer side effects is correct, this magnetically induced convulsive treatment will play an important role in referral for convulsive therapy. Furthermore, should future technological advances in delivery of MST, combined with the current improved understanding of the brain's circuitry implicated in depression, result in a treatment that is equally efficacious as ECT but with a more favorable side-effect profile, then it may replace the latter altogether. However, much more research is needed both in the realm of clinical treatment efficacy and technological device development.

Patient Selection. MST is investigated as a treatment alternative for those patients who have failed several adequate medication trials or whose symptom severity requires a fast treatment response, traditionally

requiring ECT treatment. Since general anesthesia is an integral part of the procedure, MST will most likely continue to be reserved for treatment-resistant patients after several classes of medications have been already tried.

Future Directions. MST is a novel treatment in early phases of clinical testing. Clinical treatment variables, including dosing, optimal coil placement, patient selection, and mechanisms of action are the topics of ongoing and future studies.

Vagus Nerve Stimulation

Definition. VNS is a direct, intermittent electrical stimulation of the left cervical vagus nerve via an implanted pulse generator. The stimulating electrode is wrapped around the left vagus nerve in the neck and is connected to the battery/generator usually implanted subcutaneously in the left chest wall. George and colleagues described VNS in a 2000 paper as an important new tool for brain research and therapy, where, as a result of external stimulation of the left vagus nerve, afferent signals are sent to the nucleus tractus solitarius and the connected limbic and cortical areas. Implantation surgery is completed under brief general anesthesia, usually performed as a day procedure, and involves two incisions—one in the chest for the battery/generator and another for the electrode in the left neck. Electrical parameters affecting the stimulation characteristics can be controlled by professional staff externally, with a programming wand. Patients have the ability to switch off the pulse generator by holding a magnet over the skin directly above the area where it is implanted.

History. In 2005, VNS was approved by the FDA for the adjunctive treatment of chronic, treatment-resistant depression. It had been previously approved for the treatment of resistant epilepsy. More than 50,000 patients worldwide have received VNS.

Mechanisms of Action. The majority of the fibers contained in the left vagus are afferents. It is estimated that as many as 80 percent of these fibers are upgoing afferents and thus chronic stimulation of these nerve fibers predominantly changes activity in the brainstem nuclei such as the nucleus of the tractus solitarius and other neighboring nuclei (e.g., Raphe) that alter serotonergic activity in cortical and limbic structures. Additionally, persistent stimulation of the vagal afferents is anticonvulsant, an effect that appears to be dependent upon the norepinephrine producing locus ceruleus.

Side Effects and Contraindications. Implantation of VNS is a minor procedure but carries a few risks. Besides the risk of perioperative

infection, the surgical implantation carries a small risk of vocal cord paralysis, bradycardia, or asystole. Stimulation also appears to be safe and generally well tolerated. The adverse events that have been reported are coughing, hoarseness dyspnea, reversible bradyarrhythmias, obstructive sleep apnea, neck pain, and most commonly, voice alteration. Long-term naturalistic study of VNS for major depression estimates less than 1 percent of patients have terminated VNS because of adverse events. Generally, there are no adverse cognitive side effects with VNS; to the contrary, research demonstrates performance on neurocognitive tests is significantly improved with VNS, likely due to the improvement in depression. However, modest cognitive impairment has been reported with higher stimulation parameters.

Outright contraindications to VNS include bilateral or left cervical vagotomy, use of short-wave diathermy and microwave diathermy, or therapeutic ultrasound diathermy. A special consideration may be that MRI in patients with implanted VNS device is not permitted, unless a special "send/receive" coil is utilized.

Clinical Studies. The first VNS study by Rush and colleagues in 30 treatment-resistant depressed patients reported in 2000, showed response rates of 40 percent after 10 weeks of open-label, adjunctive VNS treatment and initial remission rate of 17 percent. The response rate was sustained (46 percent) after an additional 9 months of VNS, while further significant increase in remission rates was observed, from 17 percent to 29 percent. Likewise, considerable improvements in function were also reported with long-term VNS treatment. Subsequently, in a report on an extended sample (n = 59), the short-term (3 months) and longer-term (24 months) effects of adjunctive VNS treatment revealed 10-week response rates of 31 percent (after 3 months), 44 percent (after 1 year), and 42 percent (after 2 years). Remission rates were 15, 27, and 22 percent, respectively, demonstrating continuous and stable improvement. By 2 years 81 percent were still receiving VNS. However, a large 10-week randomized, sham-controlled, blinded multicenter trial of VNS for treatment-resistant unipolar or bipolar depression later reported by Rush and colleagues failed to show a significant benefit on the primary outcomes between active and sham treatment after 10 weeks of treatment. However, an open-label long-term follow-up of the same cohort (n = 202), demonstrated increased response rates (27.2 percent response and 15.8 percent remission rates at 12 months), which was higher than response rates in a nonrandomized comparison group receiving treatment as usual (i.e., not VNS). The researchers did not attribute improvement in the VNS group to medication changes or a delayed placebo effect in such chronically ill patients.

Current Status in Treatment Algorithms. The FDA indicated VNS for the

adjunctive long-term treatment of chronic or recurrent depression in patients 18 years or older experiencing a major depressive episode in the setting of unipolar or bipolar disease who have not had an adequate response to four or more adequate antidepressant treatments. Consultation with another clinician experienced with treatment-resistant depression and VNS is recommended.

VNS treatment success rates are lower than those with ECT. Response to VNS is relatively slow—typically only 15 percent of patients respond in the first 3 months, increasing to approximately 30 percent, after 12 months. VNS may be worth considering, therefore, when patients have failed to respond to less invasive treatments, ECT was ineffective, or post-ECT relapse cannot be prevented with less invasive means. VNS might be helpful with longer-term relapse prevention, but results of controlled trials would be useful to guide practice.

Patient Selection. Stepwise approach to patient evaluation for VNS is outlined in Table 34.36–6. VNS is approved as an adjunctive long-term treatment for chronic or recurrent depressive episodes in adults with a major depressive episode that have not had a satisfactory response to four or more adequate antidepressant trials.

Consideration of what constitutes an "adequate" trial requires careful assessment of not only whether adequate dosages were prescribed for an adequate period of time but also whether the patient was reliably compliant. Optimally, such assessment should be substantiated by careful review of treatment records, pharmacy reports, and blood levels where appropriate. Failed trials should include different medication classes, including selective serotonin reuptake inhibitors (SSRIs) and selective serotonin/norepinephrine reuptake inhibitors (SNRIs), as well as tricyclic antidepressants and monoamine oxidase inhibitors (MAOIs) and various augmentation strategies with or without ECT. A second opinion from a psychopharmacologist specializing in treatment-resistant depression, knowledgeable about VNS, is advisable.

Table 34.36–6.
Evaluation for Vagus Nerve Stimulation

Does this patient meet the U.S. Food and Drug Administration labeling criteria?
Unipolar or bipolar disorder in a major depressive episode
Long-term major depressive episode (≥2 years)
Recurrent depression
Inadequate response to four or more adequate treatments
Have other causes of medication resistance been ruled out?
Medical, substance, personality disorders
Compliance with adequate doses and durations
Can the patient tolerate a slow-acting therapy?
Does the patient need acute response now?

Electroconvulsive therapy may be a better option for the acutely suicidal	
Get a second opinion	

Table 34.36–7.
Clinical Scenarios: Vagus Nerve Stimulation (VNS) or Electroconvulsive Therapy (ECT)

Acutely suicidal, in need of rapid response	ECT
Psychotic depression	ECT
Acutely depressed, medication resistant, never had convulsive therapy	ECT
Chronically depressed, able to tolerate slow onset of action	VNS
Had convulsive therapy before, but always relapses quickly despite adequate medication	VNS
Cannot be tapered off maintenance convulsive therapy, and it is not effective	VNS

Clinical scenarios when ECT might be favored over VNS, and vice versa, are outlined in Table 34.36–7. For example, VNS does not exert rapid antidepressant action and thus acutely suicidal patients and others in need of rapid response would be more appropriately treated with other strategies. VNS is indicated as a long-term treatment option only for those with chronic or recurrent course of illness.

ECT and TMS have been safely used in patients with an implant when the VNS generator was turned off during the treatments, which is necessary for ECT because of the anticonvulsant effects of VNS. It remains to be studied whether VNS could be useful in relapse prevention post ECT.

Dosing. There are varying stimulation parameters (i.e., current, pulse width, frequency, and duty cycle) that are determined by the treating doctor, which can be programmed into the pulse generator via telemetry. Optimum dose–response parameters are still being researched, although a large multicenter clinical trial did not find significant differences in efficacy in treating depression between "low" (output current of 0.25 mA, pulse width of 130 ms), "medium" (0.5 to 1.0 mA, pulse width 250 ms), or "high" (1.25 to 1.5 mA, pulse width 250 ms) at acute phases (8 to 22 weeks) or chronic phases of treatment (25 to 50 weeks).

The epilepsy literature suggests that there is a threshold current for efficacy. Because the adverse effects of VNS are known to be dose dependent, treatment parameters are often chosen to mitigate specific side effects. For example, lowering pulse width reduces neck pain, allowing patients to tolerate higher currents.

Future Directions. More research is required to establish the dose–response relationships for VNS. Future studies may explore optimal medication strategies to augment responses, test the potential role of VNS for long-term relapse prevention (e.g., after ECT), and study its mechanisms of action.

Implanted Cortical Stimulation

Definition. Cortical brain stimulation is a novel neurosurgical approach where electrodes are implanted directly over the surface of the cortex to provide electrical brain stimulation in a targeted superficial region. This approach is being studied for treatment of conditions like stroke, tinnitus, and treatment-resistant depression. Because of its novelty, different investigators use different names to describe it, ranging from implanted cortical stimulation (ICS), through cortical stimulation system, to cortical brain stimulation and direct cortical electrical stimulation (DCS). The stimulating electrode can be implanted relatively superficially in either epidural or subdural layers but the specific topographic placement of the stimulator varies depending on the condition being treated (see Fig. 34.36–6). For example, in psychiatric applications, investigators have reported on a trial involving an implanted cortical stimulator in the superficial brain cortex area overlying left dorsolateral prefrontal cortex. The electrodes are powered by a pacemaker-like device that is implanted in the upper chest area.

History. ICS is an invasive but reversible form of brain modulation and, like DBS and other less invasive neuromodulation methods, stems in part from the psychosurgical treatments for severe mental disorders performed during the 1940s and 1950s. These early forms were highly experimental and often resulted in significant adverse side effects due to physical surgery in the brain tissue. Today, with the advancement in the current understanding of the specific functional brain regions and their interconnections, as well as the relative ease of access to neuroanatomical and brain imaging studies, direct electrical stimulation is a safer way of either temporarily disrupting faulty function, enhance deficient functions, or otherwise reversibly modulate abnormal brain activity.

Trigeminal Nerve Stimulation

Definition. Research of trigeminal nerve stimulation (TNS) in humans has been ongoing since 2001 and has been most commonly studied as a therapeutic intervention for drug-resistant epilepsy. However, TNS has recently been also explored as an adjunct to pharmacotherapy in depression with positive results and clinical trials of TNS for ADHD are currently underway. At this time TNS is not approved for use outside of research however.

Mechanisms of Action. Due to the anatomy of the three bilateral afferent cutaneous branches of the trigeminal nerve with connections to the nucleus tractus solitarious, locus ceruelus, and raphe nuclei, noninvasive stimulation of the V1 branch of the trigeminal nerve has therapeutic

potential in the treatment of mood disorders and drug-resistant epilepsy.

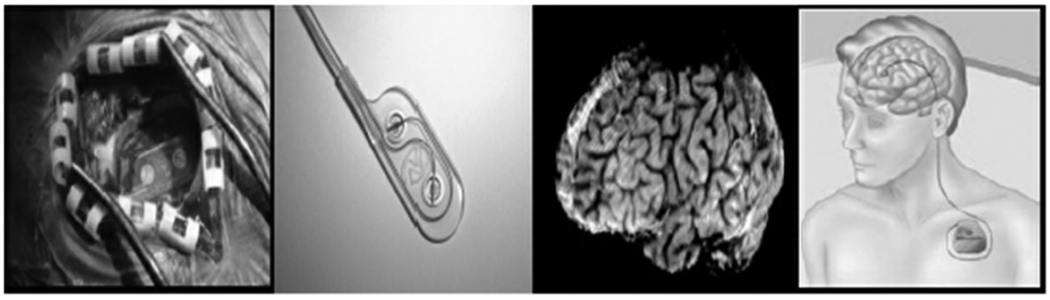

FIGURE 34.36–6. Implanted cortical stimulator.

Clinical Studies. Studies exploring psychiatric applications of TNS are ongoing, but results are not available as of this writing. Within epilepsy research, typical TNS treatment utilizes stimulation for at least 12 hours a day for 18 weeks. In a recent multicenter trial comparing high-frequency (experimental active treatment) and low-frequency (active control) external TNS (eTNS) for the treatment of drug-resistant epilepsy, investigators reported that response rates for high-frequency eTNS at 120 Hz with a pulse duration <250 μs and maximum off-time of 90 seconds were significantly greater (30.2 percent reduction in seizures) than low-frequency eTNS at 2 Hz with a duty cycle of 2 seconds on and 90 seconds off and 50-μs pulse duration (21.1 percent reduction in seizures).

Side Effects and Contraindications. TNS is generally well tolerated with low incidence of headache and skin irritation. Some patients report sensory experiences of tingling or vibration and there have been reports of significant increase in heart rate, but no significant changes to blood pressure.

Future Directions. The simplicity of administering eTNS by applying a gel electrode on the forehead is a unique advantage over surgically implanted devices. Thus far, research demonstrates that bilateral TNS is more effective than unilateral stimulation and high-frequency TNS is more efficacious compared to low-frequency stimulations. As clinical benefits may increase with treatment course, longer treatment length may increase therapeutic effects.

ACKNOWLEDGMENT

The authors would like to acknowledge the contribution to the previous edition of this chapter, written together with Dr. Sarah Hollingsworth Lisanby.

REFERENCES

American Psychiatric Association. *Practice Guideline for the Treatment of Patients with Major*

Depressive Disorder. 3rd ed. Arlington, VA: American Psychiatric Association; 2010.

Barker AT, Jalinous R, Freeston IL. Non-invasive magnetic stimulation of human motor cortex. *Lancet*. 1985;1(8437):1106–1107.

Esser SK, Huber R, Massimini M, Peterson MJ, Ferrarelli F, Tononi G. A direct demonstration of cortical LTP in humans: a combined TMS/EEG study. *Brain Res Bull*. 2006;69(1):86–94.

Fitzgerald PB, Brown TL, Marston NA, et al. Reduced plastic brain responses in schizophrenia: a transcranial magnetic stimulation study. *Schizophr Res*. 2004;71(1):17–26.

George MS, Lisanby SH, Avery D, et al. Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: a sham-controlled randomized trial. *Arch Gen Psychiatry*. 2010;67(5):507–516.

George MS, Sackeim HA, Rush AJ, et al. Vagus nerve stimulation: a new tool for brain research and therapy. *Biol Psychiatry*. 2000;47(4):287–295.

Kayser S, Bewernick BH, Grubert C, Hadrysiewicz BL, Axmacher N, Schlaepfer TE. Antidepressant effects of magnetic seizure therapy and electroconvulsive therapy in treatment-resistant depression. *J Psychiatr Res*. 2011;45(5):569–576.

Lisanby SH, Luber B, Schlaepfer TE, Sackeim HA. Safety and feasibility of magnetic seizure therapy (MST) in major depression: randomized within-subject comparison with electroconvulsive therapy. *Neuropsychopharmacology*. 2003;28(10):1852–1865.

Lisanby SH, Schlaepfer TE, Fisch HU, Sackeim HA. Magnetic seizure therapy of major depression. *Arch Gen Psychiatry*. 2001;58(3):303–305.

Luber B, Kinnunen LH, Rakitin BC, Ellsasser R, Stern Y, Lisanby SH. Facilitation of performance in a working memory task with rTMS stimulation of the precuneus: frequency- and time-dependent effects. *Brain Res*. 2007;1128(1):120–129.

Pascual-Leone A, Rubio B, Pallardó F, Catalá MD. Rapid-rate transcranial magnetic stimulation of left dorsolateral prefrontal cortex in drug-resistant depression. *Lancet*. 1996;348(9022):233–237.

Polster JD, Kayser S, Bewernick BH, Hurlemann R, Schlaepfer TE. Effects of electroconvulsive therapy and magnetic seizure therapy of acute memory retrieval. *J ECT*. 2015;31(1):13–19.

Rush AJ, George MS, Sackeim HA, et al. Vagus nerve stimulation (VNS) for treatment-resistant depressions: a multicenter study. *Biol Psychiatry*. 2000;47(4):276–286.

Rush AJ, Marangell LB, Sackeim HA, et al. Vagus nerve stimulation for treatment-resistant depression: a randomized, controlled acute phase trial. *Biol Psychiatry*. 2005;58(5):347–354.

Wang H, Wang X, Scheich H. LTD and LTP induced by transcranial magnetic stimulation in auditory cortex. *Neuroreport*. 1996;7(2):521–525.

Wassermann EM. Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the International Workshop on the Safety of Repetitive Transcranial Magnetic Stimulation, June 5–7, 1996. *Electroencephalogr Clin Neurophysiol*. 1998;108(1):1–16.

▲ 34.37 Neurosurgical Treatments

BENJAMIN D. GREENBERG, M.D., PH.D.

Over the last two decades, neurosurgical treatments for psychiatric illness have received increasing attention. Such methods are applied to small, highly selected groups of patients with disabling and otherwise treatment-intractable illness. Current techniques differ markedly from the original crude freehand "psychosurgery," especially prefrontal lobotomy of the mid-20th century, which saw wide and indiscriminate use, but including other surgeries done with handheld instruments. Modern techniques and practices have evolved considerably. Stereotactic ablation via craniotomy ("open procedures"), now rely on MRI guidance, vastly improving the accuracy of lesion placements. MRI-guided radiosurgery allows stereotactic ablation without craniotomy. Deep brain stimulation (DBS), while requiring craniotomy for device implantation, is relatively flexible, and stimulation can be stopped or modified if it causes adverse effects. In the United States, DBS for intractable OCD received FDA humanitarian approval in 2009, and is an approved treatment in the European Union. Other uses of DBS in psychiatric populations remain investigational. This chapter focuses on neurosurgery for otherwise intractable OCD or major depression. For work in Tourette syndrome, readers are directed to the 2015 review by Schrock and colleagues referenced below.

General standards of care for psychiatric neurosurgery include strict criteria for patient selection in a formalized practice, and in the case of Humanitarian Use Devices (HUDs), oversight of use by a local Institutional Review Board (IRB). DBS device implantation or other psychiatric neurosurgeries are not approved unless a multidisciplinary committee reaches consensus regarding the appropriateness for a given candidate, based on diagnostic evaluations and assessments of prior treatments. The informed consent process unfolds over time, in the context of a thorough medical and psychiatric evaluation. At this writing, surgical intervention is mainly reserved for patients with severe, incapacitating OCD who have exhausted conventional behavioral and medication therapies. As experience has been gained, the ethical context of psychiatric neurosurgery has increasingly resembled that surrounding any high-risk procedure or research, for example, neurosurgery for movement disorders.

A large body of open-label clinical data has supported the effectiveness and relative safety of modern neurosurgical interventions. Major centers continue to gather outcomes prospectively. Controlled data, although still limited, have begun to appear. There is also an increasing body of systematically collected data on adverse events, allowing risk/benefit

judgments to be better informed. While there are intriguing hypotheses, the mechanisms of action of psychiatric neurosurgeries remain to be established, a situation which parallels developments in functional neurosurgery for movement disorders. While these treatments are sometimes criticized for this reason, neurosurgical interventions are similar to other therapies insofar as the most relevant issues are safety and efficacy, rather than correction of still poorly understood pathophysiological processes.

Lesion procedures developed largely empirically, and only later became influenced by theories implicating cortico-limbic systems in disordered behavior. However, neuroscience tools increasingly permit both hypothesis generation and testing regarding potential mechanisms of action. Thus, psychiatric neurosurgery is now developing in a scientific context where translation of data between clinical results to cross-species anatomical, neuroimaging and physiological studies of neural networks involved promise to illuminate causal relationships between defined interventions in neurocircuitry and therapeutic actions. Such multidisciplinary work may inform other device-based techniques, both invasive and noninvasive. Brain stimulation using implantable devices that can respond selectively to patterns of brain activity putatively associated with symptoms (i.e., closed-loop or "responsive" DBS) has become a special focus of interest. A new method for ablative neurosurgery has also emerged which, like radiosurgery, does not require craniotomy: high intensity focused ultrasound (HIFU). Closed-loop DBS in psychiatry is very early in development and will not be discussed further here. HIFU is seeing applications across neurosurgery, but work in psychiatric illness is just beginning.

HISTORY

Trephination in prehistoric times was likely the earliest surgery for psychopathology. Neurosurgical treatment in psychiatry was first reported in 1891. These bilateral cortical excisions in demented and depressed patients were described as yielding mixed results. In 1935, John Fulton and Charles Jacobsen reported that chimpanzees showed less fear in the "experimental neurosis" paradigm after frontal cortical ablation (lobectomy). The animals retained an ability to perform complex tasks, but fear expression was reduced. Egaz Moniz, a Portuguese neurologist, attended Fulton's 1935 lecture. Prompted by these findings, Moniz pioneered prefrontal leukotomy in collaboration with his neurosurgeon Almeida Lima. First by using alcohol injections and subsequently mechanically using a leukotome (a kind of surgical knife), Moniz and Lima performed "psychosurgery" on 20 severely ill institutionalized patients; 14 were said to have exhibited worthwhile improvement. In an era of overflowing asylums and few effective treatments for debilitating illness,

surgery was at first enthusiastically embraced. Moniz won a Nobel Prize in 1949 for this contribution.

From the mid-1930s until phenothiazines emerged in the 1950s, these techniques proliferated globally. Walter Freeman, a neuropsychiatrist, was perhaps the most zealous promoter of psychosurgery in the United States. Initially, he worked closely with neurosurgeon James Watts. They developed freehand procedures to achieve prefrontal lobotomy (i.e., severing white matter connections between the prefrontal cortex and the rest of the brain). Freeman and Watts reported on their first 200 cases in 1942. Although the benefits of the surgery were highlighted, others acknowledged a significant rate of complications, such as frontal lobe syndromes (including apathy and disinhibition), seizures, and even deaths. At its peak, lobotomy was being performed on approximately 5,000 patients per year in the United States alone. A review of 10,365 prefrontal lobotomies performed from 1942 to 1954 in Britain concluded that while 70 percent showed improvement, adverse effects included 6 percent mortality, seizures in 1 percent, and disinhibition in 1.5 percent. There were widespread reports of blunted personality and socially inappropriate behavior.

In the late 1940s and early 1950s, recognition of these risks prompted attempts to develop modified stereotactic surgical procedures that might yield better results. For example, Spiegel and Wycis, who began stereotactic neurosurgery in humans, reported in the 1940s that dorsomedial thalamotomies improved obsessive-compulsive symptoms. However, with the introduction of chlorpromazine (Thorazine) in 1954, medical management of psychiatric illness became newly possible. Thus, despite the advent of stereotactic neurosurgical techniques and a continued high prevalence of severe, treatment-refractory psychiatric illness, psychiatric neurosurgery was all but abandoned in favor of nonsurgical therapies. Later, as the limitations of neuroleptics and electroconvulsive treatment became known, there was a renewed interest in psychiatric neurosurgery at a few academic medical centers in the 1960s. Refinement of stereotactic methods and emerging theories regarding the role of the limbic system in emotion shaped a newer age of psychiatric neurosurgery. The pioneers of these contemporary procedures faced harsh opposition from political groups, as well as the general population, based largely on the results of crude and indiscriminately applied operations of an earlier era. Political and social pressure dissuaded all but a handful of neurosurgeons and psychiatrists from offering such treatments. Despite these societal influences, in the late 1970s the authors of a US National Commission report found, to their own surprise, that available evidence indicated that many patients benefited after lesion procedures. They advocated continued systematic research and a national data registry. Although the registry never materialized, gradually, over nearly five decades, research and

education regarding modern neurosurgery led to its legitimacy as a treatment for severe, otherwise intractable psychiatric disorders, now with more refined practices, and surgical procedures. Indications, expectations of benefit, and risks have become better defined.

PATIENT SELECTION

Indications and Contraindications

Although limited reports have suggested efficacy across a broad range of psychiatric conditions, with research expanding rapidly, the best established indications for psychiatric neurosurgery remain otherwise intractable OCD or major depressive disorder (MDD). In evaluating candidates, several factors are considered:

1. *Primary diagnosis:* The patient must meet clinical criteria for the diagnostic indication, and this disorder should be a primary cause of the patient's debility and suffering.
2. *Severity:* The patient must have chronic, severe, and debilitating illness; duration of the primary illness must exceed 1 year and typically exceeds five years. Severity is gauged on standardized instruments (e.g., patients with OCD typically have Yale-Brown Obsessive-Compulsive Scale scores of 26 to 30; patients with major depression typically have Beck Depression Inventory scores of 30 or higher), while debility should be indicated by a low level of functioning (e.g., a Global Assessment of Functioning score of 50 or less) and a poor quality of life.
3. *Adequacy of previous treatment:* Patients must have already tried an exhaustive array of other available established treatments, delivered by competent practitioners, which are documented in detail. Assessment of adequacy of previous evidence-based psychotherapies is also highly important, and often requires interviews with treating psychotherapists as well as reviews of medical records. Medication trials must similarly be judged adequate; where appropriate, high-dose antidepressant trials (for OCD in particular) should have been attempted, with a duration of each medication trial no less than 2 months.
4. *Psychiatric comorbidity:* Appropriate treatment must also have been rendered for any comorbid psychiatric disorder. Strong relative contraindications include active psychoactive substance abuse or severe personality disorders which may interfere with adherence to presurgical treatment trials, or interfere with the exhaustive evaluation and consent processes, or with the recommended postsurgical treatment and follow-up.
5. *Medical comorbidity and surgical fitness:* Structural brain lesions or significant central nervous system injuries of clinical significance are strong contraindications. Medical conditions that increase neurosurgical risks (e.g., cardiopulmonary disease) and age over 70 years represents a

relative contraindication for lesion procedures, while DBS may be used, as it is for movement disorders, in somewhat older cases. A history of past seizures is a risk factor for perioperative seizures after lesion procedures and must factor into the overall risk-benefit assessment.

6. *Access to postoperative care:* The extensive evaluations and subsequent psychiatric neurosurgery mark the beginning of a new episode of care. It is crucial that patients have access to adequate postoperative treatment, including a psychiatrist (typically the referring physician) who will accept responsibility for managing the case after discharge. Arrangements for postoperative care (e.g., intensive behavior therapy) should be confirmed ahead of time. Importantly, after lesion procedures, care can generally be delivered in standard treatment settings without the need for highly specialized psychiatric neurosurgery teams. For deep brain stimulation, access to specialized teams is essential, over the long term. Once implanted, DBS patients require clinical monitoring and device adjustment, which can be intensive and time-consuming, especially early in treatment. Device monitoring and replacements may need to occur on a relatively urgent basis. The continuing costs incurred can be substantial, and adequacy of third-party reimbursement needs to be ensured in advance to the extent possible. After either lesion procedures or DBS, family or significant others may be needed to support and accompany patients to follow-up care, similar to the level of support that is usually necessary throughout the intensive evaluation process. Currently, DBS for intractable OCD is eligible for reimbursement under an FDA HDE (humanitarian device exemption) approval in the United States. The procedure has full approval in the European Union. Other psychiatric uses of DBS remain investigational, an important consideration for reimbursement.
7. *Informed consent:* Under no circumstances should psychiatric neurosurgery be performed on unwilling patients. The patient must be able and willing to render an informed consent. Consent monitoring, including the use of structured consent feedback tools, may be considered to help ensure that the consent process is adequate. In rare instances, these procedures are performed with assent of the patient and formal consent from a legal guardian. In this context, age less than 18 years also represents a relative contraindication.

Multidisciplinary Case Review

Assessing candidacy differs somewhat from center to center and between nations, as governed by law. Across centers, however, there is consensus that decisions regarding candidacy must be made after thorough review of the patient's case, taking into account the written record, current clinical examination, and formal referral from the patient's treating psychiatrist. In US expert centers, the decision is typically made by an interdisciplinary

committee; in Britain and Belgium, psychiatric neurosurgery requires additional formal review and approval from duly constituted national commissions. The importance of systematic review was highlighted by a study of 91 cases reviewed by the Flemish Advisory Board in Belgium. Surgery was not approved in a substantial proportion of cases on initial review, either because the primary psychiatric diagnosis appeared incorrect (10 percent of cases) or because more information was required to ascertain that diagnoses or conventional treatment trials were adequate (20 to 50 percent of cases). These findings indicate the particular importance of case review by a group with the appropriate breadth of expertise.

PREOPERATIVE WORK-UP

In addition to the above clinical evaluation and formalized review, the preoperative work-up for psychiatric neurosurgery includes physical and neurological examinations, standard presurgical blood and urine laboratory tests, as well as an ECG, plus specialized tests of brain structure and function. A comprehensive preoperative assessment should include brain MRI, psychometric testing using neuropsychological batteries that assess multiple domains, and in some cases an EEG.

POSTOPERATIVE CARE

Immediate postoperative care includes the usual medical and surgical management following any stereotactic neurosurgical procedure. Special attention is paid to signs or symptoms of potential surgical complications, including infection, hemorrhage, seizures, or altered mental status. For ablative procedures, a postoperative MRI should be obtained to document the placement and extent of lesions, timed to optimally visualize the evolution of the lesions; for DBS, imaging by CT assesses potential post-implantation hemorrhage. Over the longer term, intensive postoperative psychiatric treatment is recommended, since the efficacy of the intervention may rely on synergy between the neurosurgical intervention itself and enhanced response to pharmacological or behavioral therapies. Although dosages of psychotropic medications might be reduced during the immediate perioperative period, the medication regimen should be readjusted as tolerated postoperatively. Moreover, in the case of OCD, intensive behavior therapy should be continued or resumed early, e.g., as soon as the first postoperative month for DBS, or by 3 months for lesion procedures including gamma knife radiosurgery, where clinical benefit evolves more slowly. Longitudinal assessment of patient adherence to medication and behavioral treatments is essential so treating clinicians can intervene to optimize care if patients change their regimens on their own (as has happened in some cases postoperatively).

For DBS, as noted above, electrode implantation is followed by immediate postoperative imaging (e.g., computerized tomography; CT

scan) to assess for potential hemorrhage, and to evaluate general device placement. This is followed by a several-week recovery period to enable resolution of local edema and stabilization of other factors (e.g., pneumoencephaly) that might influence the response to stimulation. Postoperative imaging to more accurately visualize device locations (e.g., CT scan for fusion with preoperative MRI) may take place at that time. Then, systematic electrode survey and adjustment of stimulation parameters is performed before initial settings are determined. This is usually done in the outpatient setting, and is often a time-consuming process, lasting hours over 1 or more days. Ongoing protocols for DBS entail frequent follow-up, especially during the initial 6 months after implantation, to enable optimization of stimulation parameters, patient monitoring, and management including coordination of pharmacological and behavioral therapies.

DESIGN AND INTERPRETATION OF CLINICAL STUDIES

It is critical to appreciate the inherent challenges surrounding research in this domain. The patient populations in question are very challenging and can require intensive management over the months and years of a clinical trial. These populations can be very heterogeneous. High rates of comorbidity are the rule. The total numbers are small. It was recently estimated that under 1 percent of treatment-seeking OCD patients are eligible for neurosurgery, and proportions of appropriate patients with major depression might be similar. As a result, the major US psychiatric neurosurgical centers collectively treat a small number of cases yearly. Thus, statistical power is limited, as are estimates of efficacy and adverse effect rates. The major academic centers might have their own surgery of preference (sometimes more than one). Although methods are converging, historically centers have employed somewhat different measures for assessing treatment response. Patient selection criteria have also varied somewhat across centers, again especially for the earliest studies of a particular technique. Consequently, it is often difficult to contrast or combine data across these different groups. Moreover, determining the cause of long-term postoperative improvement can be confounded by exposure to other new therapies as they emerge, or even by the longitudinal course of illness.

Given this landscape, it is perhaps not surprising that there is that there is only one report of a sham-controlled trial for lesion procedures in a psychiatric population (intractable OCD), in a small sample. There have been more, but still relatively few, sham-controlled studies of DBS at several brain targets, for OCD and otherwise intractable depression. This research, all of which is quite recent, is discussed below.

LESION PROCEDURES

Although numerous approaches have been tried from the mid-20th century to the present, four lesion procedures have been the most used. All four entail bilateral lesions, performed using stereotactic methods (Figs. 34.37–1 and 34.37–2).

Subcaudate Tractotomy

Subcaudate tractotomy was introduced by Geoffrey Knight in Great Britain in 1964 as one of the first attempts to limit adverse effects by restricting lesion size compared to prefrontal lobotomy. By targeting the substantia innominata (just inferior to the head of the caudate nucleus), the goal was to interrupt white matter tracts connecting orbitofrontal cortex and subcortical structures. The surgery involved placement of radioactive yttrium-90 seeds at the desired centroid, yielding lesion volumes of approximately 2 cc on each side. Subcaudate tractotomy is used primarily in OCD and major depression. It is sometimes used as a procedure subsequent to anterior cingulotomy when the latter has been judged to have suboptimal outcome. This procedure is currently used by itself the least often among lesion methods described here, except in its continuing application as part of limbic leukotomy (see Limbic Leukotomy).

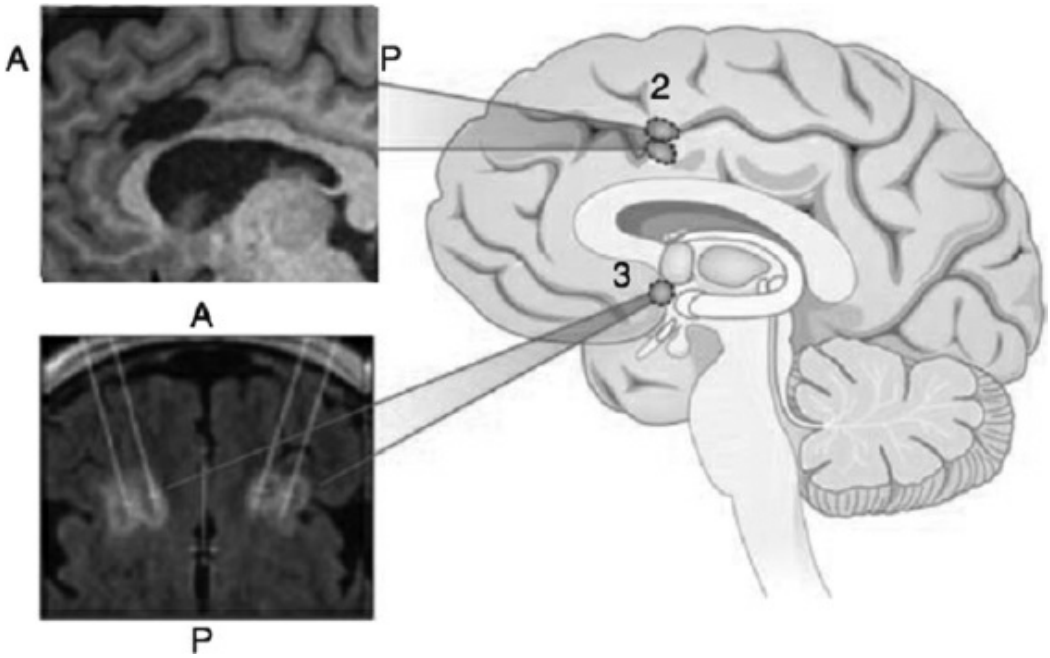

FIGURE 34.37–1. Sagittal schematic of target sites for surgical procedures linked to postoperative MRI images. Site 2: anterior cingulotomy linked to postoperative sagittal MRI (A, 1⁄4 anterior; P, 1⁄4 posterior). Site 3: subcaudate tractotomy linked to postoperative axial MRI. (From Greenberg BD, Rauch SL, Haber SN. Invasive circuitry-based neurotherapeutics: stereotactic ablation and deep brain stimulation for OCD. *Neuropsychopharmacology*. 2010;35[1]:317–336.)

Anterior Cingulotomy

Fulton was first to propose that the anterior cingulum as a target in

psychiatric illness. The first cingulotomies were performed for intractable pain (a practice which continues, along with newer research into DBS of the anterior cingulum for the same condition). It was noted that pain patients with comorbid anxiety or depressive conditions had better responses. H. Thomas Ballantine Jr. and colleagues later demonstrated the relative safety of anterior cingulotomy and studied its efficacy for a broad range of psychiatric indications. In the half-century since this work began in 1962, the group at Massachusetts General Hospital has performed over 1,000 anterior cingulotomies. The surgery is conducted under local anesthesia and two or three approximately 1-cc lesions are made on each side by thermocoagulation through bilateral burr holes. The target is within anterior cingulate cortex (Brodmann areas 24 and 32), at the margin of the white matter bundle known as the cingulum (Fig. 34.37–1). Originally determined by ventriculography, since 1991 anterior cingulotomy has been MRI-guided. Since the practice is to make the smallest possible lesion consistent with effectiveness, a subgroup of patients may return months after the first operation for a second procedure. This may be an extension of the first set of lesions, or addition of subcaudate tractotomy. There has also been investigational use of DBS at the ventral capsule/ventral striatum target after cingulotomy. Indications for anterior cingulotomy include major depression and OCD.

Limbic Leukotomy

Limbic leukotomy was introduced by Desmond Kelly and colleagues in England in 1973. The procedure combines subcaudate tractotomy and anterior cingulotomy. The lesions have typically been made via thermocoagulation or with a cryoprobe. Historically, the precise placement of the lesions was guided by intraoperative stimulation; pronounced autonomic responses were believed to designate the optimal lesion site. The indications for limbic leukotomy include major depression, OCD, and other severe anxiety disorders. There is also evidence that this procedure might be beneficial in cases of otherwise treatment-intractable repetitive self-injurious behaviors or severe tic disorders.

Anterior Capsulotomy

Anterior capsulotomy was first performed by Tailarach and colleagues in France. Later, Lars Leksell and colleagues further developed the procedure in Sweden. Anterior capsulotomy or its newer variant, Gamma Knife capsulotomy, are used in Scandinavia, the United States, Belgium, Brazil, and elsewhere. The procedure places lesions within the anterior limb of the internal capsule, which impinges on the adjacent ventral striatum, thereby interrupting fibers of passage between prefrontal cortex and subcortical nuclei including the dorsomedial thalamus (Fig. 34.37–2). The original

anterior capsulotomy procedure is performed using thermocoagulation via burr holes in the skull. For 20 years capsulotomy has also been performed using the Gamma Knife as an alternative which avoids craniotomy. Typically, gamma capsulotomy lesions are smaller than those induced by thermocapsulotomy, remaining within the ventral portion of the anterior capsule. Hence, the term "gamma ventral capsulotomy" is coming into use to describe this procedure. In contrast to thermocapsulotomy, gamma ventral capsulotomy may be performed as an outpatient procedure, with an overnight hospital stay usually the most that is required. The relative advantages and disadvantages of this radiosurgical approach are the focus of ongoing research, including a small controlled study of gamma ventral capsulotomy for OCD, the first of its kind for a lesion procedure in psychiatry. The rates of adverse effects appear lower for any anterior capsulotomy procedure when smaller volumes of tissue are lesioned. Anterior capsulotomy, while primarily used for OCD, has also seen use in major depression and other severe anxiety disorders.

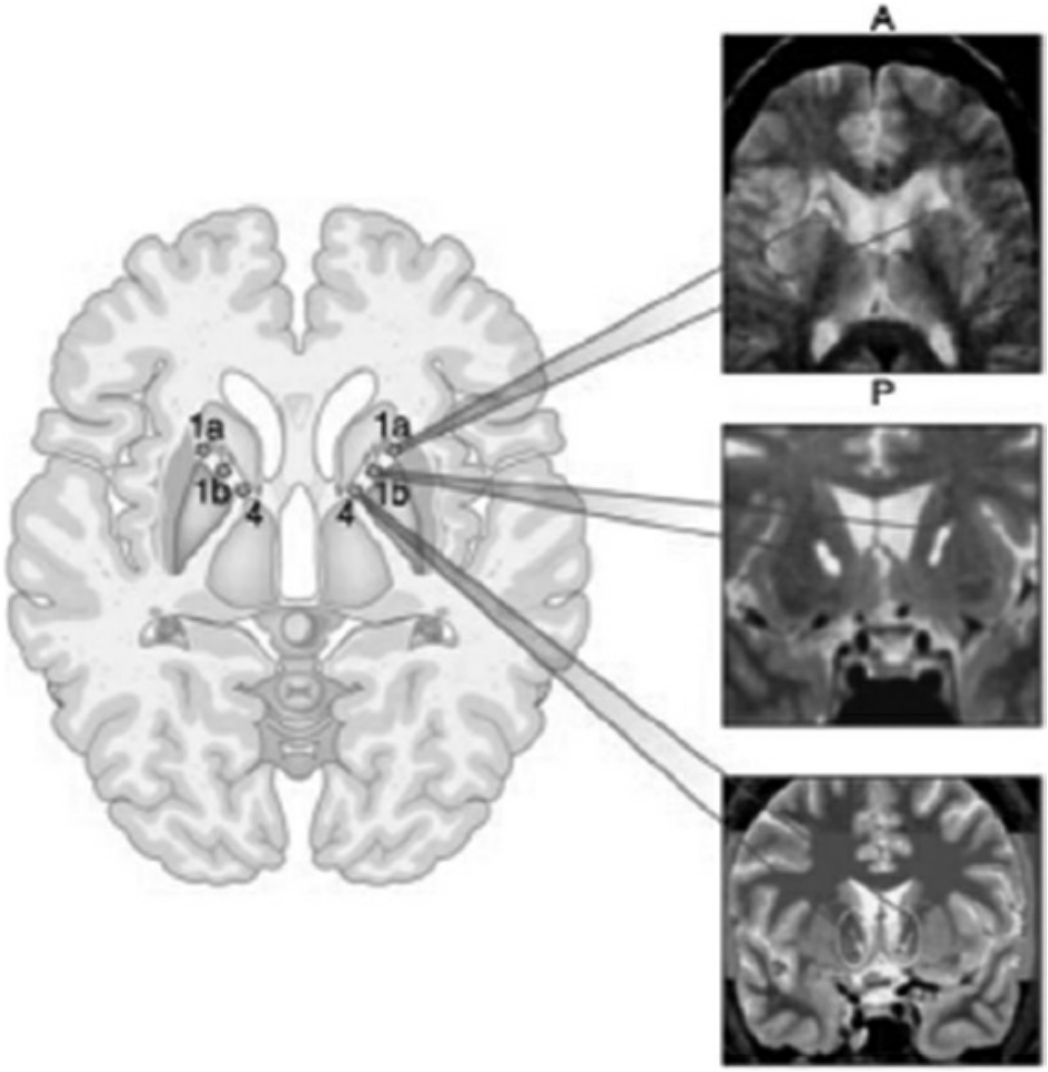

FIGURE 34.37–2. Axial schematic of target sites linked to postoperative MRI images.

Site 1a: anterior capsulotomy (thermocapsulotomy) linked to postoperative axial MRI (A, 1⁄4 anterior; P, 1⁄4 posterior). Site 1b: gamma ventral capsulotomy linked to a coronal MRI. Site 4: ventral capsule/ventral striatum (VC/VS) deep brain stimulation linked to postimplantation coronal CT scan fused with preoperative MRI. (From Greenberg BD, Rauch SL, Haber SN. Invasive circuitry-based neurotherapeutics: stereotactic ablation and deep brain stimulation for OCD. *Neuropsychopharmacology*. 2010;35[1]:317–336.)

Deep Brain Stimulation

DBS for psychiatric illness is not a new idea, although the devices, surgical techniques, and theoretical models of relevant neurocircuitry have all advanced. For example, in 1948, Pool used stimulation via a silver electrode in the caudate nucleus in an attempt to treat a woman with depression and anorexia. Modern DBS began 50 years later, when Nuttin and colleagues reported a small series of cases treated with otherwise intractable OCD who received DBS, initially based on the anterior capsulotomy target. The procedure involves placement of small-diameter brain "leads" (e.g., approximately 1.3 mm) with multiple electrode contacts into subcortical nuclei or specific white matter tracts. The surgeon drills burr holes in skull bone under local anesthesia and then places the leads, guided by multimodal imaging and stereotactic landmarking. Usually this is done bilaterally.

The subject is typically sedated but awake during surgery. Later, the "pacemaker" (also known as an implantable neurostimulator or pulse generator) is implanted subdermally (e.g., in the upper chest wall) and connected, via extension wires tunneled under the skin, to the brain leads. Since various combinations of electrodes can be activated, at adjustable polarity, intensity, and frequency, DBS permits a considerable range of ways one can modulate brain function. Thus, parameters can be optimized for individual patients. This might achieve improved efficacy and more favorable adverse effect profiles in comparison with ablation. The process of optimizing stimulation parameters for individual patients, typically performed by a specially trained psychiatrist in an outpatient setting, is quite time-consuming and requires attentive, long-term follow-up. In cases where no beneficial settings can be identified despite extensive efforts, the electrodes can be inactivated, and device components may be surgically removed. The surgical risks of device removal include infection and hemorrhage, as for implantation. There are important psychiatric risks of device removal, which patients may view as the end of a "last resort" episode of care. See Outcomes with DBS for further details.

TREATMENT OUTCOMES

For all ablative procedures, outcome cannot be fairly assessed for a considerable postoperative period, which could extend from 6 months to 2 years. While effects of DBS may develop sooner, e.g., within several months postoperatively, long-term follow-up remains essential to assess the

durability of responses. Early descriptions of ablation outcomes, all in open-label studies, used global measures such as the Pippard Postoperative Rating Scale, with ratings of (1) symptom-free, (2) much improved, (3) slightly improved, (4) unchanged, and (5) worse. Most studies considered meaningful benefit if a patient were symptom-free or much improved. Increasingly, reports employed disorder-specific severity measures, e.g., the Yale-Brown Obsessive-Compulsive Scale (YBOCS) for OCD, or specific depression scales. Increasingly, sham-controlled trials include both clinician- and patient-rated measures of symptom burden, functioning, and quality of life. Below we further review the four neuropsychiatric neurosurgical procedures by surgical approach.

Outcomes with Subcaudate Tractotomy

In 1975, results of subcaudate tractotomy were reported. Follow-up was 2.5 years on average. Optimal data, including structured interviews, were available for a subset of 134 of an initial 208 patients. Improvement was judged clinically significant in 68 percent of patients with MDD, 50 percent of patients with OCD, and 62.5 percent of those with other anxiety disorders. Individuals with schizophrenia, substance abuse, or personality disorders did poorly. Patients were usually ambulatory by the third postoperative day. Postoperative adverse events included transient headache and confusion or somnolence, which were typically transient (lasting 1 or more weeks). Longer-term adverse effects included mild undesirable personality changes in 6.7 percent, and seizures in 2.2 percent. Transient disinhibition syndromes were common. One patient died due to migration of an yttrium seed, and three others committed suicide.

In 1994, a large-scale review of 1,300 subcaudate tractotomy cases was conducted and concluded that the procedure enabled 40 to 60 percent of patients to lead normal or near normal lives. The suicide rate was 1 percent versus 15 percent in a comparison group similarly affected by with major affective disorders. There were no reported deaths due to other causes in this series. The seizure rate was comparable to that of previous reports. In contrast to the earlier report, the 1994 review did not find enduring personality changes. The investigators proposed predictors of positive response, including MDD diagnosis, sudden onset of symptoms, emergence of illness in the second half of life or peripartum, positive family history of psychiatric disorder, and previous positive response to electroconvulsive therapy (ECT). Physiological responses elicited during intraoperative electrical stimulation of the target area before it was lesioned were also thought to predict better responses. As noted above, subcaudate tractotomy is now rarely done as a single procedure.

Outcomes with Anterior Cingulotomy

In 1987, Ballantine and colleagues retrospectively reviewed 198 cases with mean follow-up after cingulotomy of 8.6 years. They judged there was significant improvement in 62 percent of patients with affective disorders, 56 percent with OCD, and 79 percent with other anxiety disorders. Most recently, a prospective report was published by the MGH group on 64 consecutive patients undergoing cingulotomy for intractable OCD between 1989 and 2009. Changes in symptom severity were assessed at the most recent follow-up using the YBOCS and the Beck Depression Inventory. A full OCD response was defined as a YBOCS score reduction of ≥35 percent, with improvements in the 25 to 34 percent range defined as partial responses. At a mean of 10.7 months, 35 percent of OCD had a full response, while 7 percent responded partially. Of the 64 patients, 30 had a second procedure (repeat cingulotomy or subcaudate tractotomy). At a mean of over 5 years (63.8 months), improvement rates climbed to 47 and 22 percent for full and partial responses, respectively. Of patients (24 of 64) who responded with a partial response at the mean follow-up of just under one year, 20 (83 percent) retained at least a partial response at the final follow-up, a mean of about 5 years later. Severity of comorbid major depression decreased modestly from baseline: Beck Depression Inventory scores improved by 17 percent at last follow-up.

Psychiatric adverse events included suicide in 2 of the 64 patients, 1 of whom with a prior history of both major depression and a previous suicide attempt. The second suicide occurred 8 days after a surgery which had no neurological complications, in a patient with comorbid bipolar depression. Other adverse events included: transient memory impairment (n = 5) which resolved over days to months; transient urinary retention (n = 2) lasting days, or worsening of pre-existing urinary incontinence (n = 1); postoperative abulia (n = 4) lasting days, in 1 case after initial cingulotomy and in 3 after subsequent limbic leukotomy; intraoperative (n = 3) or postoperative single seizures (n = 1); and recurrent seizures (n = 1) requiring anticonvulsant treatment. In 1 patient, a postoperative infection progressed into a subdural empyema requiring repeat surgery. There were no intracranial hemorrhages.

Outcomes with Limbic Leukotomy

In the 1970s, before the era of MRI-guided surgery, Kelly and colleagues published a series (n = 66) studied prospectively with a mean 16 months of follow-up. Improvement was judged significant in 89 percent of patients with OCD, 78 percent with major depression, and 66 percent with other anxiety disorders. In a separate report regarding 49 patients with OCD, they noted that 84 percent were improved at 20 months. Short-term side effects included headache, lethargy or apathy, confusion, and lack of sphincter control lasting from a few days to a few weeks. Following limbic leukotomy, it is common for postoperative confusion to last at least days,

and patients are often not discharged from the hospital in less than 1 week. There were no reported seizures and no deaths; however, enduring lethargy was present in 12 percent of cases, and one patient suffered severe memory loss due to improper lesion placement.

In 2002, Montoya and colleagues reported on 21 patients who underwent limbic leukotomy for OCD or major depression at MGH. Using stringent criteria, they reported 36 percent to 50 percent responders, respectively, at a mean follow-up point of 26 months. The adverse event profile was similar to that previously reported, with apathy, urinary incontinence, and short-term memory deficits infrequent and transient. Of note, in 2002, Price and colleagues published a series of five consecutive cases which suggested that limbic leukotomy might be beneficial for patients with severe self-mutilation, in the context of repetitive self-injurious, tic-like behaviors.

Outcomes with Anterior Capsulotomy

Thermocapsulotomy. As the name implies, in thermocapsulotomy, thermal lesions are made, usually via radiofrequency heating. This procedure involves craniotomy (hence its other name, "open capsulotomy"). The ablation initially involved a large portion of the dorsal-ventral extent of the anterior limb of the internal capsule. In the 1970s, Leksell and colleagues reported on their first 116 cases of thermocapsulotomy. In that early study they noted a favorable response in 50 percent of those with OCD and 48 percent of those with major depression. A subsequent review of the earlier experience with capsulotomy uncovered adequate data on 213 of 362 cases. Of these, 64 percent were deemed to have benefitted significantly on global scales. Short-term side effects included transient headache, incontinence, and confusion, the latter lasting up to 1 week. Patients were typically ambulatory in hours to days following the operation, with the length of hospital stay influenced by the duration of confusion. A subset of 24 patients, followed prospectively, had one case of intracranial hemorrhage and one episode of seizure. Fatigue was present in 29 percent, 17 percent had poor memory, and 8 percent exhibited lack of attention to appearance ("slovenliness"). Weight gain was a common, enduring side effect with a mean increase in body mass of about 10 percent.

A smaller 2008 study in 25 OCD patients who underwent thermocapsulotomy, gamma knife anterior capsulotomy, or both combined. It used disorder-specific rating scales and measures of adverse effects. Even though these patients had failed to respond to treatments unavailable to the earlier cohorts (thus their illnesses might be expected to be more refractory than in the early experience), effectiveness was in fact comparable to that in the early work: 48 percent response and 36 percent remission rates were seen on YBOCS OCD severity ratings. This series also

demonstrated a notable adverse effect burden. Forty percent were judged to have clinically significant frontal impairment (problems with executive functioning, apathy, or disinhibition). Most in this group had had repeated thermocapsulotomy and/or high-dose gamma knife radiation, which today would be avoided, resulting in larger lesions. Also in contrast to the earlier work, MRI was available. Interestingly, smaller lesions, less lateral in extent in the coronal plane and also posteriorly, were associated with not only fewer adverse effects, but also greater symptom reduction.

More recently a small prospective series of eight patients with severe chronic major depression received thermocapsulotomy for that illness. Lesions involved a large proportion of the dorsal-ventral extent of the capsule on MRI. At 24 to 36 months postoperatively, half of the group were responders having at least a 50 percent reduction in BDI scores. One patient developed a progressive vascular dementia with parkinsonism caused by autopsy-proven arteriolosclerosis. The response rate, in a group of patients receiving contemporary standard treatment, was similar to that reported for depression in the 1970s by Leksell.

Gamma Ventral Capsulotomy. In gamma knife radiosurgery, lesions are made without craniotomy, which means that the procedure can typically be performed on an outpatient basis. The Leksell Gamma Knife is used. Over 200 beams of radiation from cobalt sources are focused, using stereotactic methods and MRI guidance, on surgical targets in the brain, in intractable OCD. The most common approach has been to place two lesions ("shots") in the ventral portion of the anterior limb of the internal capsule using 4 mm collimators to guide the radiation. Lesions are either staged over two surgeries or combined in a single procedure. More recently, single lesions in the ventral quadrant of the capsule have been used in attempt to minimize total radiation doses and also to focus on ventral prefrontal cortico-basal ganglia-thalamic fibers, hypothesized as possibly the most relevant to therapeutic benefit in OCD.

Five-year outcome data illustrate the delayed nature of the response. Therapeutic benefit was achieved over 1 to 2 years and was essentially stable by 3 years. Full responses (a YBOCS OCD severity improvement by >35 percent) developed in over half of cases by 2 to 3 years after surgery. Partial responses averaged approximately 10 percent at 2 to 3 years postop.

Adverse events after the double-shot procedure (two lesions made bilaterally on the same day) included cerebral edema and headache (with onset delayed by months to years), small asymptomatic caudate infarctions, and possible exacerbation of pre-existing bipolar mania. One patient developed a mild frontal syndrome, including apathy, that resolved gradually over a period of years. Three patients developed radionecrotic brain cysts, the onset of which was delayed by 3 or more years. In two cases cysts were asymptomatic. In contrast, in the third case, cyst development

was progressive and associated with serious neurological sequelae: symptomatic increased intracranial pressure requiring shunt placement and ongoing neurosurgical management.

A controlled trial of gamma ventral capsulotomy carried out at the University of São Paulo, Brazil is unique in the psychiatric neurosurgery literature. Sixteen patients with rigorously defined intractable OCD were enrolled; eight received active GVC and eight received sham surgery in which they entered a specially constructed antechamber attached to the Leksell Gamma Knife. The groups were followed for 1 year, after which individuals in the sham group were offered actual surgery. There was no change in the sham group during the 1-year follow-up period. In contrast, two patients (25 percent) in the active GVC group showed reductions in OCD severity during the masked phase of 35 percent or greater on the YBOCS. Given the small sample, the effect did not reach statistical significance. In the group receiving active surgery first, five of the eight patients (62.5 percent) went on to respond during follow-up after the masked phase. Four of the eight patients who initially received the sham procedure underwent later active GVC. Two of these later responded, for an overall response rate of 58 percent after the active procedure. In this study, a longer masked sham period, which would have been optimal scientifically given the delayed nature of the response (see above) was not chosen given the ethical concern that severely ill individuals not undergo unduly long follow-up after a sham surgery.

Adverse events included manic episodes in two patients (17 percent), each of whom had a previous history of subclinical hypomanic symptoms. These individuals responded to addition of mood stabilizers to their medication regimens. One patient, who had no previous history of substance abuse, developed drug dependence 4 years after GVC. The most serious adverse event was related to substantial radiation-induced perilesional edema in one patient 8 months post-GVC. This was accompanied by delirium, confabulation, and visual hallucinations, all of which responded to corticosteroids within a few days. However, memory deficits and executive function impairment persisted for 5 months, after which the patient returned to his baseline level of neuropsychological functioning. At month 54 after surgery, the same patient developed a brain cyst of 6-mm diameter. This cyst was not associated with neurological or neuropsychological impairment.

Outcomes with DBS

Obsessive-Compulsive Disorder. Collaborative work in the 1990s led to the development of DBS at the ventral anterior limb of the internal capsule and adjacent ventral striatum (the VC/VS) for otherwise intractable OCD. The first implantation was performed by Nuttin and colleagues in Belgium. Stimulation target development was most directly informed by the prior

experience with anterior capsulotomy, both thermocapsulotomy and gamma knife ventral capsulotomy, which was itself based on thermocapsulotomy. The combined report focused on open-label data in 26 patients in total; at some study centers patients also underwent a masked crossover to DBS discontinuation from active stimulation. Across the four centers—Leuven/Antwerp, Butler Hospital/Brown University, the Cleveland Clinic, and the University of Florida—over 60 percent of patients showed therapeutic responses at last follow-up (a full response was defined as a 35 percent or greater reduction in YBOCS OCD severity). In the group as a whole, responses tended to stabilize after about 3 months of stimulation.

Cognitive functioning was not adversely affected by DBS. Adverse events included hemorrhage after implantation in two patients (~8 percent); in one of them apathy resulted, which resolved over 3 months. One patient (~4 percent) had a single seizure intraoperatively without recurrence. One patient developed a superficial surgical wound infection which responded to antibiotics. In two patients (~8 percent) a DBS system wire broke, requiring surgical replacement. Complications due to stimulation itself or stimulation interruption were the most common, including four (~16 percent) patients with worsened depression or suicidal thinking similar to events in their presurgical course of illness. In three of them, this worsened depression occurred during a blinded crossover to sham stimulation. Mild hypomania, judged as stimulation-related but nonserious, was seen in eight patients (31 percent). In all cases, hypomania was either reversible with a change in stimulation, or resolved over days without a DBS parameter change.

In a study of DBS at an anatomical target closely related to the VC/VS, the nucleus accumbens (NAcc), 16 patients underwent 8 months of open-label treatment which included (as did the VC/VS study) exposure-based behavior therapy and ongoing pharmacotherapy. The open phase was followed by a double-blind crossover phase with randomly assigned 2-week periods of active or sham stimulation, followed by open maintenance DBS. In the open phase, 56 percent of patients responded. In the double-blind, sham-controlled phase (n = 14), YBOCS scores were significantly lower for active versus sham stimulation. The most prominent stimulation-related adverse was elevated mood or hypomania in 50 percent of patients, judged nonserious in all cases. One patient (6 percent) had a surgical wound infection, while five (31 percent) reported forgetfulness and three (19 percent) reported word-finding difficulties. However, objective testing in this as in other studies of psychiatric DBS, found either a lack of cognitive deterioration, or improvement. This is unlike the situation in DBS for movement disorders, where stimulation may be associated with cognitive decline in Parkinson disease.

A controlled crossover study investigated effects of DBS at a different

target within the general cortical-basal ganglia-thalamic circuitry implicated in OCD, the subthalamic nucleus. Of a total group of 16 patients, half were randomized to active DBS and half to sham stimulation, after which each group crossed over to the other condition in double-masked fashion. There was a significant difference between active and sham DBS in favor of the active treatment. Adverse events included an intracerebral hemorrhage on device insertion in one patient (6 percent) leading to a permanent motor deficit in one hand, and infection leading to pulse generator removal in two patients (12 percent). One patient was described as having stimulation-related hypomania, and one with mania related to DBS. Both effects were reversible.

Other DBS stimulation targets within the general circuitry of interest include the inferior thalamic peduncle and ventral caudate nucleus, each used in OCD in small studies. In this respect, development of psychiatric DBS has followed, on a smaller scale, the path of stimulation for movement disorders, where several targets have been pursued. As in movement disorders, overlapping or converging effects of DBS at different anatomical sites on the neurocircuitry involved are likely and remain a focus of research.

Major Depression. One continuing line of research into DBS for major depression began with observations from functional neuroimaging pointing to the subgenual cingulate cortex as a node in circuits involved in the normal experience of sadness, symptoms of depressive illness, and responses to depression treatments. In the initial publication of its type, a small open-label study found that patients with unipolar depression refractory to four or more treatment trials improved after stimulation of white matter tracts adjacent to subgenual cingulate cortex. A larger prospective trial had similar findings. Another open-label study, which included patients with bipolar as well as unipolar depression, found benefit after subgenual DBS in the bipolar patients as well as those with unipolar depression, without induction of hypomania. A large industry-supported controlled trial of subgenual DBS, however, has not encouraged continuing commercial development of the method. The stimulation target continues to be refined based on updated hypotheses concerning which neural pathways, amendable to identification on diffusion MRI prospectively, might be most important to modulate.

Another line of research on DBS for depression was prompted by the VC/VS DBS OCD research discussed above, and on the reported antidepressant effects of anterior capsulotomy on which the VC/VC stimulation target was initially based. The OCD patients in both the lesion and DBS experience had high rates of comorbid depression, as expected in that population. Responders displayed rapid improvements in mood and motivation, which preceded reductions in core OCD symptoms. An open-

label study of DBS at the VC/VS target for depression without comorbid OCD enrolled patients meeting a very high threshold for treatment refractoriness and found clinically significant antidepressant responses in roughly half of the patients. Induction of transient, reversible mood elevation, occasionally reaching the diagnostic threshold for hypomania, was the most significant adverse effect of active stimulation. In contrast, an industry-led controlled feasibility trial was negative, due to low response rates in both the active and sham stimulation groups.

Current State of Psychiatric Neurosurgery

Despite progress and the promise of advances in conventional medication and behavioral therapies, there remains a small proportion of individuals with intractable OCD or major depression who would be judged as potential candidates for neurosurgical treatments. These remain severely affected and have inadequate responses to conscientiously applied standard care. As with any treatment, the potential risks and benefits of psychiatric neurosurgery must be weighed against the potential risks and benefits of foregoing these approaches in individuals likely to remain burdened with severe illness and very poor quality of life. The conclusion reached in previous editions of this volume, then based almost entirely on open-label lesion procedure outcomes, was that about half of carefully selected psychiatric patients may meaningfully benefit from neurosurgery over a period of years after such procedures, and that a subset might show outstanding improvement. Although data remain incomplete, and cases of attempted and completed suicide have continued to occur after psychiatric neurosurgery, these procedures are still judged as likely to reduce mortality, as evidenced by data on comparative suicide rates. Nonetheless, patients who undergo and fail to benefit from these “last resort” procedures are at particularly high risk for completed suicide.

There has since been only one controlled trial of a lesion procedure, gamma ventral capsulotomy for OCD, in a small sample. While benefit at the end of the 1-year masked phase was not significant, at longer follow-up improvement was seen after the active procedure, consistent with the previous open-label experience. A follow-up study of a different sample of patients who underwent anterior capsulotomy for OCD with thermal and/or gamma knife procedures suggested that lesion size and location matters to both efficacy and adverse effect burden, which, together with other small-scale open-label work, has led to modifications of lesion methods for ventral capsulotomy. Other functional or structural neuroimaging data suggest that subsets of individuals might be identified prospectively as most likely to benefit after anterior cingulotomy for depression or DBS for OCD. Such methods are still early in development.

As for DBS, at present, some controlled data support application of DBS for OCD at a series of anatomically related brain sites. A US multicenter

controlled study of VC/VS DBS for intractable OCD finished enrollment of the last patient in its sample in 2015, with results expected within a year. Based on open-label data supporting a probability of benefit and relative safety, DBS for OCD is available in the United States under a humanitarian (HDE) FDA approval. Small numbers of patients continue to receive that procedure, typically at expert centers. DBS for OCD has full regulatory approval in the European Union. For intractable depression, initial controlled DBS trials have been disappointing. Small-scale research continues to test refinements of DBS targeting, or novel brain sites for stimulation.

Given the current landscape, clinicians are faced with a number of issues when considering a particular clinical or research procedure for a given patient who might be judged a candidate for neurosurgery. The first, for OCD in particular, is consideration of a lesion procedure versus DBS. DBS may be acceptable to patients who would not choose to undergo lesion procedures, although the reverse is also true, given the long-term treatment context of DBS. Despite its advantages of reversibility and adjustability, DBS implantation carries risks, and requires that patients be treated by highly specialized teams willing and able to provide long-term care. The logistics and expenses involved, including inconsistent coverage by insurers, can represent significant barriers. In contrast, psychiatric care can be delivered in standard treatment settings after lesion procedures. However, although the relative risks of enduring adverse effects after psychiatric DBS remain to be clearly established, at this stage ablative methods appear to carry a greater potential for them. Rates of adverse outcomes are relatively low when modern lesion procedures are performed by highly experienced multidisciplinary teams, something also seen with DBS for movement disorders, where there has been the most experience. This represents a particularly strong rationale for referral of appropriate patients to expert psychiatric neurosurgery centers.

RESEARCH AND FUTURE DIRECTIONS

The best evidence on which one could base treatment recommendations would be adequately powered, long-term, prospective comparison studies using random assignment of individuals to different surgical methods. That is unlikely to become available. As noted above, the small number of appropriate patients, the experience and preferences of expert centers for particular methods, and the preferences of patients themselves for one method over another, are all impediments in this regard. In contrast, long-term follow-up studies, systematic reviews, and ongoing research pointing towards individualized predictors of response have occurred and should be encouraged.

An example is a 2015 systematic review, by Pepper et al., comparing outcomes of anterior capsulotomy to those of DBS for OCD. A total of 170

cases were identified across 20 studies. DBS targets were the VC/VS or the closely related NAcc. After DBS (n = 62) there was a 40 percent decrease in YBOCS scores, while after capsulotomy (n = 108) YBOCS OCD severity decreased 51 percent. There was a 9 percent higher likelihood of remission after capsulotomy than after DBS. The authors noted no overall difference in complication rates, although the pattern of complications was different (lasting adverse effects tend to be more common after capsulotomy). The follow-up period was very different: 19 months for DBS versus 61 months for capsulotomy, which may allow for further development of improvement after capsulotomy, and conversely might not allow assessment of how many DBS patients experienced device failures or were no longer receiving stimulation for a variety of reasons.

Another 2015 systematic review, by Brown and colleagues, also based on observational data, compared outcomes of anterior cingulotomy to those of anterior capsulotomy for OCD. Ten studies involving 193 total participants were included; the capsulotomy dataset was similar to that in the review above. Mean follow-up was 47 months for cingulotomy and 60 months for capsulotomy. At last follow-up, 41 percent were responders after cingulotomy, while 54 percent responded after capsulotomy. Serious or permanent adverse events, characterized as the number of events per procedure, was 5.2 percent across cingulotomy studies and 21.4 percent across capsulotomy studies, although in the latter, half of those events were attributable to very high-dose, multiple target gamma knife procedures (which no longer would be performed) across three patients. Excluding those, the rate of serious complications in the capsulotomy group dropped to 12.8 percent.

A key research direction is the use of increasingly high-resolution structural and functional imaging methods which can assess changes in the relevant networks, for example their connectivities, prospectively after lesion procedures to test and generate hypotheses about which brain network alterations are most relevant to therapeutic improvement and also to adverse effects. Such MRI-based methods are at present not easily applicable to DBS studies for reasons of safety during imaging. As is possible in studies of ablative methods, work in DBS can make use of multiple modalities of neuroimaging at baseline to test for morphometric, fiber tract, and local and network activity as individual predictors of outcomes. Both kinds of work take place in a context where the anatomy of fiber connections affected by lesions and by stimulation has become increasingly better defined, in both humans and in animal models (Fig. 34.37–3).

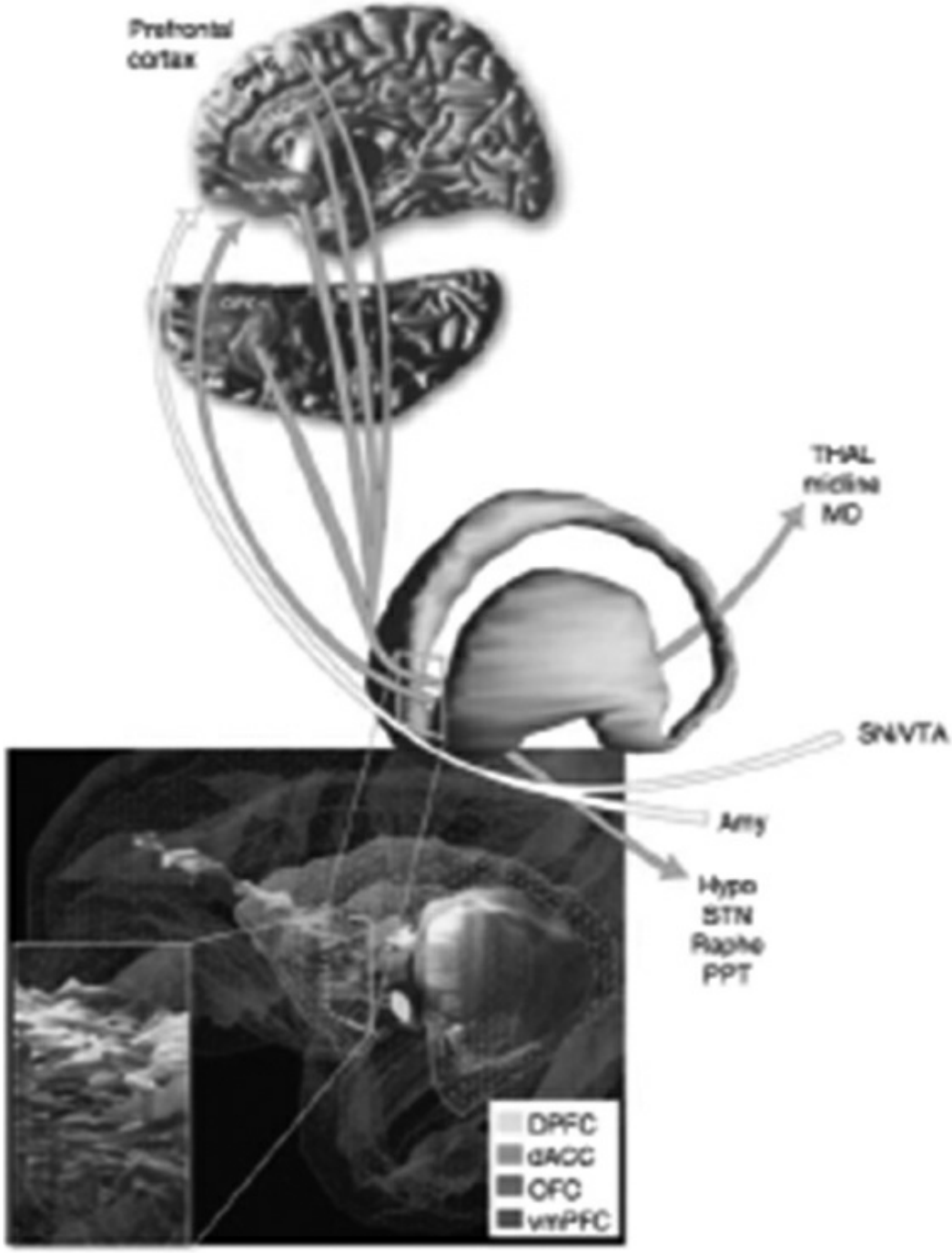

FIGURE 34.37–3. Schematic representation of reciprocal connections between prefrontal cortex regions and the ventral portion of the anterior limb of the internal capsule and the adjacent ventral striatum (VC/VS). Prefrontal cortex: VC/VS connections are color-coded to delineate pathways linking specific prefrontal regions. Other connections of the VC/VS represented: THAL, mediodorsal and midline nuclei of the thalamus; SN/VTA, substantia nigra and ventral tegmental area; Amy, amygdala; Hypo, hypothalamus; STN, subthalamic nucleus; Raphe, raphe nuclei; PPT, pedunculopontine tegmental nucleus. (From Greenberg BD, Rauch SL, Haber SN. Invasive circuitry-based neurotherapeutics: stereotactic ablation and deep brain stimulation for OCD. *Neuropsychopharmacology*. 2010;35[1]:317–336.)

In summary, current neurosurgical treatment continues to represent a viable option for appropriately selected patients with severe and treatment

refractory illnesses. Evidence is most robust for OCD, but there are no studies directly comparing one technique to another even for that diagnostic group. As evidence has accumulated, it has become increasingly clear that the patterns of risks, burdens, and benefits, while they overlap, are distinctly different for DBS versus ablation. Research intended to inform psychiatric neurosurgery is proceeding along a number of fronts. Neuroimaging techniques might be a potential aid in individual patient selection by predicting treatment response. Structural and/or functional neuroimaging might be used to tailor surgical interventions on an individual basis. Ablation or DBS might enhance the benefits of subsequent behavioral or medication therapies that proved ineffective before surgery. Both kinds of treatments are being studied in the context of an increasing focus on the neurocircuitry underlying both syndromic illness and dimensions of psychopathology. Insights gained from invasive methods have the potential to inform development of noninvasive, circuit-based treatments for psychiatric illness, such as transcranial magnetic or electrical stimulation.

REFERENCES

Ballantine HT, Bouckoms AJ, Thomas EK, Giriunas IE. Treatment of psychiatric illness by stereotactic cingulotomy. *Biol Psychiatry*. 1987;22:807.

Boccard SGJ, Fitzgerald JJ, Pereira EA, et al. Targeting the affective component of chronic pain: a case series of deep brain stimulation of the anterior cingulate cortex. *Neurosurgery*. 2014;74(6):628–637.

Bridges PK, Bartlett JR, Hale AS, Poynton AM, Malizia AL. Psychosurgery: stereotactic subcaudate tractotomy: an indispensible procedure. *Br J Psychiatry*. 1994;165:599.

Brown LT, Mikell CB, Youngerman BE, Zhang Y, McKhann GM, Sheth SA. Dorsal anterior cingulotomy and anterior capsulotomy for severe, refractory obsessive-compulsive disorder: a systematic review of observational studies. *J Neurosurg*. 2015:1–13.

Corkin S, Twitchell TE, Sullivan EV. Safety and efficacy of cingulotomy for pain and psychiatric disorders. In: Hitchcock ER, Ballantine HT, Myerson BA, eds. *Modern Concepts in Psychiatric Surgery*. Amsterdam: Elsevier; 1979:253.

Denys D, Mantione M, Figee M, et al. Deep brain stimulation of the nucleus accumbens for treatment-refractory obsessive-compulsive disorder. *Arch Gen Psychiatry*. 2010;67(10):1061–1068.

Dougherty DD, Rezai AR, Carpenter LL, et al. A randomized sham-controlled trial of deep brain stimulation of the ventral capsule/ventral striatum for chronic treatment-resistant depression. *Biol Psychiatry*. 2015;78(4):240–248.

Dunn LB, Holtzheimer PE, Hoop JG, Mayberg HS, Roberts LW, Appelbaum PS. Ethical issues in deep brain stimulation research for treatment-resistant depression: focus on risk and consent. *AJOB Neurosci*. 2011;2(1):29–36.

Fox MD, Buckner RL, Liu H, Chakravarty MM, Lozano AM, Pascual-Leone A. Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and

neurological diseases. *Proc Natl Acad Sci U S A*. 2014;111(41):E4367–E4375.

Freeman W, Watts JW. Prefrontal lobotomy in the treatment of mental disorders. *South Med J*. 1937;30:23.

Garnaat SL, Greenberg BD, Sibrava NJ, et al. Who qualifies for deep brain stimulation for OCD? Data from a naturalistic clinical sample. *J Neuropsychiatry Clin Neurosci*. 2014;26(1):81–86.

Gentil AF, Lopes AC, Dougherty DD, et al. Hoarding symptoms and prediction of poor response to limbic system surgery for treatment-refractory obsessive-compulsive disorder. *J Neurosurg*. 2014;121(1):123–130.

Greenberg BD, Gabriels LA, Malone DA, et al. Deep brain stimulation of the ventral internal capsule/ventral striatum for obsessive-compulsive disorder: worldwide experience. *Mol Psychiatry*. 2010;15(1):64–79.

Greenberg BD, Rauch SL, Haber SN. Invasive circuitry-based neurotherapeutics: stereotactic ablation and deep brain stimulation for OCD. *Neuropsychopharmacology*. 2010;35(1):317–336.

Holtzheimer PE, Kelley ME, Gross RE, et al. Subcallosal cingulate deep brain stimulation for treatment-resistant unipolar and bipolar depression. *Arch Gen Psychiatry*. 2012;69(2):150–158.

Jbabdi S, Lehman JF, Haber SN, Behrens TE. Human and monkey ventral prefrontal fibers use the same organizational principles to reach their targets: tracing versus tractography. *J Neurosci*. 2013;33(7):3190–3201.

Kelly D, Richardson A, Mitchell-Heggs N. Stereotactic limbic leucotomy: neurophysiologic aspects and operative technique. *Br J Psychiatry*. 1973;123:33.

Lehman JF, Greenberg BD, McIntyre CC, Rasmussen SA, Haber SN. Rules ventral prefrontal cortical axons use to reach their targets: implications for diffusion tensor imaging tractography and deep brain stimulation for psychiatric illness. *J Neurosci*. 2011;31(28):10392–10402.

Lopes AC, Greenberg BD, Canteras MM, et al. Gamma ventral capsulotomy for obsessive-compulsive disorder: a double-blind, randomized controlled trial. *JAMA Psychiatry*. 2014;79(1):1066–1076.

Mallet L, Polosan M, Jaafari N, et al. Subthalamic nucleus stimulation in severe obsessive–compulsive disorder. *N Eng J Med*. 2008;359(20):2121–2134.

Malone DA, Dougherty DD, Rezai AR, et al. Deep brain stimulation of the ventral capsule/ventral striatum for highly resistant depression. *Biol Psychiatry*. 2009;65(4):267–275.

Mayberg HS, Lozano AM, Voon V, McNeely HE, Seminowicz D. Deep brain stimulation for treatment-resistant depression. *Neuron*. 2005;45:651.

Moniz E. Prefrontal leucotomy in the treatment of mental disorders. *Am J Psychiatry*. 1937;93:1379.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Report and Recommendations: Psychosurgery. Department of Health and Human Services, Pub No (OS) 77-002. Washington, DC: US Government Printing Office; 1979.

Nuttin B, Cosyns P, Demeulemeester H, Gybels J, Meyerson B. Electrical stimulation in anterior limbs of internal capsules in patients with obsessive-compulsive disorder. *Lancet*. 1999;354:1526.

OCD-DBS Collaborative Group. Deep brain stimulation for psychiatric disorders. *Neurosurgery*.

2002;51(2):519.

Pepper J, Hariz M, Zrinzo L. Deep brain stimulation versus anterior capsulotomy for obsessive-compulsive disorder: a review of the literature. *J Neurosurg*. 2015;122(5):1028–1037.

Price BH, Baral I, Cosgrove GR, Rauch SL, Nierenberg AA. Improvement in severe self-mutilation following limbic leucotomy: a series of 5 consecutive cases. *J Clin Psychiatry*. 2001;62(12):925.

Rauch SL, Dougherty DD, Cosgrove GR, Cassem EH, Alpert NM. Cerebral metabolic correlates as potential predictors of response to anterior cingulotomy for obsessive compulsive disorder. *Biol Psychiatry*. 2001;50:659.

Rauch SL, Dougherty DD, Malone D, Rezai A, Friehs G. A functional neuroimaging investigation of deep brain stimulation in patients with obsessive-compulsive disorder. *J Neurosurg*. 2006;104:558.

Robertson M, Doran M, Trimble M, Lees AJ. The treatment of Gilles de la Tourette syndrome by limbic leucotomy. *J Neurol Neurosurg Psychiatry*. 1990;53(8):691–694.

Rück C, Karlsson A, Steele JD, et al. Capsulotomy for obsessive-compulsive disorder: long-term follow-up of 25 patients. *Arch Gen Psychiatry*. 2008;65(8):914–921.

Rylander G. Stereotactic radiosurgery in anxiety and obsessive compulsive states: psychiatric aspects. In: Hitchcock ER, Ballantine HT, Myerson BA, eds. *Modern Concepts in Psychiatric Surgery*. Amsterdam: Elsevier; 1979:235.

Schrock LE, Mink JE, Woods DW, et al. Tourette syndrome deep brain stimulation: a review and updated recommendations. *Mov Disord*. 2015;30(4):448–471.

Spofford CM, McLaughlin NCR, Penzel F, Rasmussen SA, Greenberg BD. OCD behavior therapy before and after gamma ventral capsulotomy: case report. *Neurocase*. 2012;20(1):42–45.

Synofzik M, Clausen J. The ethical differences between psychiatric and neurologic DBS: smaller than we think? *AJOB Neurosci*. 2011;2(1):37–39.

Van Laere K, Nuttin B, Gabriels L, et al. Metabolic imaging of anterior capsular stimulation in refractory obsessive compulsive disorder: a key role for the subgenual anterior cingulate and ventral striatum. *J Nucl Med*. 2006;47:740.

Yang JC, Papadimitriou G, Eckbo R, et al. Multi-tensor investigation of orbitofrontal cortex tracts affected in subcaudate tractotomy. *Brain Imaging Behav*. 2014;9(2):342–352.

▲ 34.38 Combination Pharmacotherapy

CHARLES DEBATTISTA, M.D., AND ALAN F. SCHATZBERG, M.D.

Historically, the combination of psychotropic medications for the treatment of major psychiatric disorders has been frowned upon. The practice of "polypharmacy" was often seen as evidence of poor clinical judgment. A patient who was on three or four psychotropics was one whom many clinicians and impartial observers would characterize as overmedicated and perhaps improperly diagnosed and cared for. The significant limitations of monotherapy with antidepressant, antipsychotic, and mood-stabilizing agents have become increasing apparent. For many

patients, monotherapy results in insufficient benefit. There are three primary reasons for combining medications in psychiatric practice. Perhaps the most common reason for polypharmacy is that many patients do not respond adequately to a single agent. For example, in acute trials with antidepressants, most unipolar depressed patients are either partial or nonresponders to monotherapy. Medications are thus frequently combined to enhance antidepressant response. A second, related reason for combining therapy is that not all aspects of a given illness may respond to a single agent. This may be particularly true in bipolar disorder where monotherapy with a single mood stabilizer has become "the exception rather than the rule." Treating the manic phase or the depressed phase may require different or combinations of medications. Finally, medications are sometimes added to counteract the side effects of the original medication. Thus, patients on typical antipsychotics are usually given an anticholinergic agent to mitigate extrapyramidal side effects.

In addition to the advantages of combination treatment, there can be important downsides to combining psychotropic agents. The most important of these are the risks of pharmacodynamic and pharmacokinetic drug interactions. Cytochrome P450 interactions can result in either substantially higher or lower serum levels which might increase side effects or render the one or more agents less effective. For example, the combination of paroxetine and imipramine can substantially raise the serum levels of imipramine and potentially increase antimuscarinic, hypotensive, cardiac conduction, or sedating effects of the imipramine. Likewise, pharmacodynamic interactions can be equally serious. For example, combination of an MAOI and a serotonergic agent may result in a potentially lethal serotonin syndrome. Thus, clinicians must carefully weigh the risks and benefits of all combination treatment.

COMBINATION STRATEGIES IN THE TREATMENT OF MAJOR DEPRESSION

Lithium–Antidepressant Combinations

Given the limited efficacy of antidepressant monotherapy in the treatment of major depression, augmenting agents or combination strategies are frequently employed in clinical practice to enhance response. Conventionally, augmentation is used to describe the addition of an agent that is not viewed as a primary antidepressant by itself (i.e., thyroid hormone). Combination has been used when a known antidepressant is added to another (i.e., bupropion [Wellbutrin] and fluoxetine [Prozac] combination).

Lithium (Eskalith) has been the best-studied augmenter to standard antidepressants in such patients. However, lithium augmentation has become less popular with clinicians in the past 15 years secondary to the potential toxicity of lithium, the requirement for laboratory monitoring,

and the availability of other augmenters. Nonetheless, lithium is still used by many clinicians to boost antidepressant response in partial and nonresponders to monotherapy with an antidepressant.

The early reports of lithium augmentation indicated that the addition of lithium to a TCA trial resulted in clinical improvement within 72 hours in patients who had failed to respond to a TCA alone. Since then, at least 50 studies have been completed on lithium augmentation, and the vast majority have suggested some acute benefit for the addition of lithium to antidepressant. However, many of these reports suggest maximum benefits may not be seen for weeks. Furthermore, the addition of lithium also appears to prevent relapses in unipolar patients more effectively than an antidepressant alone. Response often occurs at low doses (600 to 1,200 mg/day) and at relatively low serum levels. Current data would suggest that serum levels in the range of 0.5 to 0.8 mEq/L are effective in lithium augmentation of antidepressant effects.

Although most studies of lithium augmentation have employed TCAs, a number of studies have evaluated the efficacy of lithium augmentation of SSRIs. The most recent of these trials is the National Institute of Mental Health (NIMH)-funded Sequenced Treatment Alternatives for Resistant Depression (STAR*D) trial. In patients who had failed monotherapy with citalopram (Celexa) and an augmentation trial with either buspirone (BuSpar) or bupropion, only 15.9 percent of patients achieved a remission level response with the addition of lithium. This was a numerically but not statistically lower level of response than those patients who were augmented with thyroid hormone at that point in treatment. Since most lithium augmentation trials do not indicate the degree of treatment resistance in the study population, the STAR*D data may be the most relevant to date on the limitations of lithium augmentation in resistant-depressed patients. There may be certain characteristics in depressed patients that predict response to the addition of lithium to their antidepressant regimen. Depressed patients may be more likely to improve with lithium augmentation if they exhibit significant psychomotor retardation, have significant anorexia and weight loss, and have low serum cortisol levels.

Thyroid Supplement–Antidepressant Combinations

After lithium augmentation, thyroid augmentation of antidepressants remains perhaps the second most studied strategy in the treatment of resistant depression. To date, over 30 studies have been completed on thyroid augmentation in resistant depression. However, thyroid supplementation may be less consistently effective than lithium and has also faded in popularity with the advent of new augmentation options.

Trials that have compared thyroid to lithium augmentation directly, have generally suggested that thyroid augmentation is at least as effective

as lithium augmentation. The few head-to-head trials of thyroid to lithium augmentation have tended to favor thyroid augmentation.

Although most trials employing thyroid augmentation have involved the use of TCAs, a number of trials also suggest efficacy when thyroid supplements are added to SSRIs. A number of case reports and small studies have suggested that liothyronine (Cytomel) (triiodothyronine or T_3) at 25 to 50 mcg/mL/day, when added to various SSRIs, does potentiate its effects. The largest and most systematic thyroid augmentation of an SSRI trial to date is the STAR*D trial. In this trial, patients who had failed two other trials (monotherapy with citalopram and augmentation with bupropion or buspirone) were treated with triiodothyronine or lithium augmentation. About 25 percent of patients on T_3 augmentation of citalopram achieved a remission level response compared to about 15 percent of the lithium-augmented patients. Among the limitations of the STAR*D trial was that it was an open-label trial, and patients had the option of not participating in a given randomization. Not all placebo-controlled studies have shown the benefit of thyroid augmentation.

The question of whether T_3 is more effective than supplementation with levorotatory thyroxine (T_4) has not been adequately addressed in the literature. Most studies have employed T_3, and the only head-to-head study favored T_3 over T_4. Since the half-life of T_4 is relatively long and the study was short (2 weeks), it cannot be conclusively interpreted for the superiority of T_3. Indeed, there are positive studies employing T_4 in thyroid augmentation. Although side effects were not common in this trial, the use of supraphysiological doses of thyroxine for an extended period might be problematic for some patients.

Patient characteristics of responders to thyroid augmentation include females and those with a history of thyroid abnormalities. There is growing evidence that patients with mild thyroid abnormalities (slightly elevated TSH, normal T_3 and T_4) are less likely to respond to antidepressants alone but respond to the addition of a thyroid supplement. Women over the age of 50, perhaps because they are most susceptible to hypothyroidism, also appear to have high responses to thyroid augmentation. More recently, Lerer's group has reported that responders to T_3 could be identified by their having lower T_3 levels and by being carriers for a specific allele on an SNP for the deiodinase converting enzyme for T_3 to T_4. This finding is in need of replication but could explain why some patients appear to derive great benefit from thyroid augmentation. Cytomel from the onset in combination with antidepressants has in a larger scale trial proven not more effective than adding placebo.

Stimulant-Like Agent–Antidepressant Combinations

There are a variety of stimulants and stimulant-like agents that have been useful augmenters in depression. Amphetamines have been used since the early 1930s for the treatment of depression. In addition, dopamine agonists such as pramipexole (Mirapex), dopamine reuptake blockers including bupropion, nondopamine and wake-promoting agents such as modafinil (Provigil) have shown some promise as augmenters in the treatment of depression.

The use of stimulants as augmenting agents in the treatment of depression has been applied to special populations including medical and geriatric patients. Amphetamines tend to be used to target fatigue, hypersomnia, and cognitive difficulties in depressed patients. When combined with TCAs, methylphenidate and the amphetamines inhibit the cytochrome P450 CYP isoenzyme 2D6 and artificially increase TCA serum levels.

Although most amphetamine augmentation studies have involved TCAs, a few have shown benefits in combination with MAOIs and SSRIs. Pharmacodynamic interactions between the stimulants and MAOIs can be problematic. Most controlled trials with amphetamine in the treatment of major depression have not shown benefits as a monotherapy. By far, the largest and most comprehensive studies to evaluate a stimulant to augment standard antidepressants involved lisdexamfetamine combined with and SSRI or SRNI. Subjects who failed to respond to 8 weeks of antidepressant monotherapy were randomized to an additional 8 weeks of lisdexamfetamine or placebo combined with the antidepressant. The two phase III trials involving over 800 patients failed to show an advantage of adding the lisdexamfetamine over adding a placebo the antidepressant monotherapy.

There are now several dopamine agonists approved for the treatment of Parkinson disease including pramipexole. Preliminary work with pramipexole as an adjunctive agent to SSRIs has suggested that it may have a role in the treatment-resistant patients in doses of 1 mg/day (starting doses are lower). Ropinirole (Requip) is also actively being investigated as an adjunctive treatment in resistant depression. In addition, pramipexole and ropinirole clearly treat restless legs, including restless legs exacerbated by the use of SSRIs.

Modafinil and armodafinil are nonstimulants that are approved by the FDA for the treatment of excessive sleepiness associated with narcolepsy, obstructive sleep apnea, and shift work sleep disorder. The lack of a more significant abuse potential for modafinil and armodafinil have led to their application as adjunctive agents in the treatment of depression with the same symptom targets as the stimulants: fatigue and hypersomnia. Several studies have indicated that modafinil is well tolerated and appears to be useful in depressed partial responders to SSRIs. While meta-analyses of completed randomized trials with combination treatment with modafinil

and armodafinil have generally suggested efficacy to these drugs in bipolar and unipolar patients, recent registration trials of armodafinil for the treatment of bipolar depression failed to demonstrate adequate efficacy.

Bupropion has weak dopamine reuptake–blocking properties but may act primarily as a noradrenergic agent. It is among the most common augmenting agents used today in resistant depression. The combination of bupropion with SSRIs is simple and usually well tolerated. Studies completed to date have been retrospective or open. The largest of these open trials, the STAR*D trial, reported a remission rate of about 30 percent in patients who had failed to achieve remission on citalopram alone. As with many augmentation trials, it is not clear whether longer treatment with citalopram alone may have resulted in the improvement. In fact, the CO-MED trial demonstrated that starting with the combination of escitalopram and bupropion was no more effective than treating with escitalopram alone in achieving antidepressant remission or response. Nonetheless, the experience of US clinicians supports the use of the bupropion augmentation.

Antidepressant–Antipsychotic Combinations

Although antipsychotics have been used in combination with antidepressants for many years in the treatment of psychotic depression, the atypical agents also appear to be useful augmenters and adjunctive agents in the treatment of unipolar and bipolar depression. The pharmacology of the atypicals, including their 5-hydroxytryptamine type 2 ($5\text{-}HT_2$) antagonism, has long been thought to imbue them with antidepressant properties. In addition, they clearly help a number of symptom domains in depressed patients including sleep, appetite, and agitation. As of this writing, several second-generation antipsychotics have FDA approval in the adjunctive treatment of resistant depression. Thus, they are only a class of agents that are specifically FDA approved to augment the effects of standard antidepressants.

Olanzapine (Zyprexa) was the first of the second-generation antipsychotics to seek approval in the adjunctive treatment of depression. The combination of 10 to 20 mg/day of olanzapine with fluoxetine in the treatment of various types of depression including resistant depression. In 2007, the olanzapine-fluoxetine combination (Symbyax) was approved for the treatment of resistant depression. Combination pills, however, tend not be popular with clinicians. Many clinicians prefer the more flexible dosing options of combining two different agents. Since both olanzapine and fluoxetine are generically available, there is also a cost advantage of using two generic agents over one branded drug.

Aripiprazole (Abilify) has long been of interest in the treatment of depression because it is a partial agonist of dopamine and a $5\text{-}HT_{1A}$ agonist

in addition to its 5-HT_2 effects. Randomized controlled trials have demonstrated that the addition of aripiprazole to an SSRI or venlafaxine (Effexor) is more effective than the addition of placebo in relieving depressive symptoms. As a result, aripiprazole was FDA approved in the adjunctive treatment of major depression. The relative lack of weight gain, sedation, or metabolic effects in aripiprazole compared to other second-generation agents such olanzapine have made it among the most frequently prescribed antidepressant-augmenting agents. However, akathisia and some weight gain were noted in the aripiprazole augmentation trials.

Quetiapine (Seroquel), the first monotherapy approved for bipolar depression, also appears to be useful in many unipolar patients. Several of registration trials demonstrated that the addition of 150 to 300 mg/day of quetiapine to standard antidepressants improved antidepressant response. Thus, quetiapine was approved as adjunctive treatment for major depression in 2009. Other second-generation antipsychotics including lurasidone (Latuda), ziprasidone (Geodon), and risperidone (Risperdal) also appear effective in some open-label and controlled studies as adjunctive agents in unipolar depression. While lurasidone is approved as a treatment for bipolar depression and schizophrenia, it also appears has demonstrated efficacy in at least 4 registration trials of the treatment of depressive symptoms associated with schizophrenia. As with other combination strategies, the potential benefits of augmenting with a second-generation antipsychotic must be weighed against the possible side effects of weight gain, EPS, and metabolic effects.

Sex Steroid–Antidepressant Combinations

It is known that estrogen-deficiency states, such as the postpartum period, may increase the risk of depressive symptoms in some women. In addition, estrogen possesses some antidepressant properties including MAO inhibition and serotonin agonist properties that are dose and duration specific. However, estrogen supplementation has been inconsistently useful in the treatment of depressed women. Some studies found that estrogen had no significant antidepressant effect when used alone for perimenopausal depression. In addition, other studies have found estrogen to be an effective prophylaxis against recurrent affective illness in the postpartum period. However, there are few adequately controlled studies, and a number of studies have not shown substantial benefit.

The combination of estrogens with antidepressants has sometimes been useful. For example, there have been case reports of oral contraceptives augmenting the effects of antidepressants in patients with refractory depression. In addition, postmenopausal women may respond better to fluoxetine if they are on hormone replacement therapy than if they are not. Recent studies of estrogen augmentation of SSRIs in both pre- and

postmenopausal women have not shown a consistent benefit. However, 0.1 mg of transdermal estradiol (Alora) was suggested to accelerate SSRI response in some perimenopausal women. The potential benefits of estrogen augmentation must be balanced against the defined risks to long-term treatment with unopposed estrogen, and the benefits have not been clearly established in the treatment of depression. Overall, the data suggest maximal benefit during the perimenopausal phase.

Dehydroepiandrosterone (DHEA), a corticosteroid precursor of both androgens and estrogens, is used over the counter in many older patients with claims of enhanced mood, energy, and sexual vigor. There are a few controlled trials that indicate that DHEA may be efficacious in the treatment of major depression, minor depressive symptoms, and dysthymia. The safety of DHEA has been called into question because of its potential to enhance the growth of sex-steroid tumors and possible virilizing effects in women.

Other Antidepressant Combinations

A variety of other antidepressant combinations are used in the treatment of resistant depression with varying degrees of scientific support. Given the poor response of antidepressant monotherapy, the question of whether combining antidepressants from the onset of treatment would improve remission or response rates was evaluated a large multi-center trial. The CO-MED trial (2011) evaluated whether the combination of escitalopram and bupropion, or the combination of venlafaxine and mirtazapine from the onset of treatment was superior to monotherapy treatment with escitalopram or venlafaxine alone in achieving a remission or response. There appeared to be no benefit from starting with treatment with combination treatment versus monotherapy.

Trazodone (Desyrel) has been used primarily as a hypnotic to treat depression or emergent insomnia. Although some small open studies have suggested an augmenting role for the use of trazodone combined with SSRIs, there is little evidence of benefit beyond the hypnotic effects of the drug. Likewise, buspirone, a generically available anxiolytic, has been applied as an augmenter to SSRIs with inconsistent benefits. Double-blind trials for buspirone augmentation of SSRIs in the treatment of OCD have been negative, and only open studies in depression suggest some benefit of adding buspirone to an SSRI. The STAR*D study found that among patients who had failed to achieve remission to citalopram alone, about 30 percent of patients achieved remission with the addition of buspirone. This is the same rate found to remit with bupropion augmentation in the trial.

The β-blocker pindolol (Visken) has some support in the literature as an augmenter to SSRIs but little support among clinicians. There have been multiple controlled trials of pindolol augmentation with most being positive. The relative lack of use among clinicians may be related to the

perception that pindolol is not efficacious, to concerns about the implication of β-blockers in inducing depression, and some wariness of possible pharmacokinetic interactions between β-blockers and CYP 2D6 inhibiting SSRIs such as fluoxetine and paroxetine (Paxil). Mirtazapine, like trazodone, can treat depression-associated insomnia and may help stimulate appetite in depressed elderly patients. The combination of venlafaxine and mirtazapine in patients who had failed three other trials in the STAR*D trial, and the combination of venlafaxine and mirtazapine showed more of an improvement in depressive symptoms and was far better tolerated than was the MAOI tranylcypromine (Parnate). However, the overall response rates in the mirtazapine SNRI combination and tranylcypromine groups were not statistically different. Only about 10 percent of patients achieved remission.

COMBINATION STRATEGIES IN THE TREATMENT OF BIPOLAR DISORDER

Combination of Two Mood Stabilizers

Mood stabilizers typically refer to agents that are used to treat one or more poles of bipolar illness and might also have utility in the maintenance treatment of bipolar disorder. Mood stabilizers have historically included agents such as lithium, divalproex (Depakote), carbamazepine (Tegretol), and perhaps other anticonvulsants. Although agents such as atypical antipsychotics might also be appropriately referred to as mood stabilizers, in this section mood stabilizers will refer to lithium and anticonvulsant agents that have demonstrated utility in bipolar disorder. Approximately two-thirds of patients with bipolar affective disorder are treated with two more medications as maintenance therapy. In addition, atypical antipsychotics are now frequently added during acute psychotic episodes to antidepressants, and sedative/anxiolytics are used to manage sleep, anxiety, and agitation. Thus, monotherapy is the exception in the treatment of bipolar disorder rather than the standard of care.

Lithium is commonly combined with anticonvulsants such as valproate, carbamazepine, and lamotrigine in the treatment of patients with refractory mania, in maintenance treatment, and in the treatment of bipolar depression. There are few prospective data using such combinations; however, a number of reports have indicated that lithium–carbamazepine combinations are effective in patients who have failed to respond to these two agents given separately. Retrospective reviews have generally found the combination of lithium and carbamazepine to be useful and synergistic. There is also evidence that this combination may be particularly useful in rapid-cycling bipolar disorder. The combination of lithium and carbamazepine appears to be well tolerated. In addition, the combination may have a cumulative antithyroid effect. There is no evidence of increased neurotoxicity or blood dyscrasias with the combination.

Dosing schedules of the lithium and carbamazepine should parallel regimens used for prescribing each drug alone. Therapeutic serum levels of both drugs should be monitored and maintained.

Perhaps the most common mood stabilizer combination in recent years has been the addition of valproate (Depakote) to lithium in patients who have manic breakthroughs or are unresponsive to lithium monotherapy. There are still relatively few prospective studies that have examined the efficacy of this combination. One naturalistic study indicated that the addition of valproate to lithium is more likely to be effective if the lithium level is above 0.6 mEq/L. Open-label studies have suggested that the combination of valproate and lithium may be more effective than either agent alone in the treatment of rapid-cycling bipolar disorder. For example, one study reported that eight of nine rapid-cycling patients with bipolar disorder, most of who failed to respond to the combination of lithium and carbamazepine, responded when lithium was combined with valproate. However, other studies have not necessarily supported an advantage of addition valproate to lithium in the treatment of rapid-cycling bipolar disorder. Likewise, while the combination of valproate and lithium is commonly thought to be more effective in preventing cycles of mania and depression than monotherapy, there is little data to support this idea. The large trial to date to evaluate the efficacy of the valproate/lithium combination in the maintenance treatment of bipolar 1 disorder, the BALANCE study (2010), failed to show an advantage of the combination treatment over lithium alone. However, the combination was more effective than valproate monotherapy. The combination of lithium and valproate seems to be well tolerated, and the dosing should aim to achieve adequate serum levels for both drugs.

The combination of valproate with carbamazepine has fallen out of favor in the past decade. Unfortunately, only anecdotal data support the combination of valproate and carbamazepine in the treatment of resistant bipolar disorder. In addition, a CYP450-based interaction increases the risk of carbamazepine toxicity. Also, the two taken together may result in even greater leucopenia than either alone.

A variety of other anticonvulsants such as gabapentin (Neurontin), topiramate (Topamax), lamotrigine (Lamictal), tiagabine (Gabitril Filmtabs), levetiracetam (Keppra), and oxcarbazepine (Trileptal) have been combined with lithium and divalproex in the treatment of resistant bipolar patients. With the exception of lamotrigine, most of the reports detailing the benefits of combining standard mood stabilizers with these other anticonvulsants in the treatment of bipolar disorder have been anecdotal. For example, many open-label studies have supported a role for the treatment of bipolar disorder with gabapentin. At this time, no controlled trials have demonstrated a clear benefit of gabapentin in any phase of bipolar disorder as either an adjunct or a monotherapy. In one large-scale,

double-blind study, gabapentin was no more effective than placebo when added to lithium. In that study, the lithium dosage could be adjusted during the trial, and this may have mitigated against finding statistically significant differences. However, gabapentin may still benefit some symptoms in bipolar patients. For example, one small controlled trial has suggested that while gabapentin does not show acute benefits, it may augment the effect of standard mood stabilizers in preventing relapse in bipolar patients. Gabapentin is well tolerated in combination with most mood stabilizers and appears to benefit anxiety and agitation associated with a variety of conditions. A chemically related compound, pregabalin (Lyrica), has shown utility in the treatment of generalized anxiety but has not been extensively evaluated in the treatment of bipolar disorder.

Lamotrigine has shown benefits in the maintenance treatment of bipolar disorder and in some studies in the treatment of acute bipolar depression, but its efficacy in the treatment of acute mania has not been demonstrated. Open-label augmentation with lamotrigine for bipolar depression in the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) did not show significant benefits on the primary outcome measures when compared to either risperidone or inositol. However, the patients treated with adjunctive lamotrigine had lower depression ratings than patients treated either with adjunctive lamotrigine or inositol. The benefits of lamotrigine make it likely that it will be increasingly employed as an adjunctive treatment to lithium, an atypical antipsychotic, or valproate in the treatment of the depressed phase of the illness. There does not appear to be significant pharmacokinetic interactions between lamotrigine and lithium. However, a risk for interactions does exist in the combination of lamotrigine with either valproate or carbamazepine. Valproate doubles the serum level of lamotrigine, thus increasing the risk of rash, while carbamazepine, on the other hand, cuts the lamotrigine serum levels in half.

The addition of oxcarbazepine, tiagabine, levetiracetam, and topiramate to lithium or valproate has not been extensively studied in the treatment of bipolar disorder. Oxcarbazepine is thought to carry fewer risks of pharmacokinetic interactions than does its chemical cousin carbamazepine. Thus, oxcarbazepine may be easier to combine with standard mood stabilizers than carbamazepine. However, some enzyme induction still occurs with the drug. Earlier studies have supported the efficacy of oxcarbazepine in the treatment of mania and hypomania.

Topiramate has been studied primarily as an adjunctive agent in the treatment of bipolar disorder. Open trials suggest that the addition of topiramate to a standard mood stabilizer may help the depressed phase of the illness and improve symptoms in refractory patients. However, most controlled trials have failed to show a benefit of topiramate in bipolar disorder. The most common current rationale for combining topiramate

with a mood stabilizer is to mitigate the weight gain associated with standard therapy. Topiramate has been associated with weight loss averaging up to 5 kg in 8 weeks as adjunctive treatment in bipolar disorder.

Tiagabine has not been proven effective or practical, thus far, in the treatment of acute mania. The drug requires a slow titration secondary to potential side effects, including sedation and a risk of seizure, with loading doses. However, open trials of tiagabine as an add-on to a standard mood stabilizer in refractory bipolar disorder have reported benefits at low doses (1 to 8 mg/day).

Levetiracetam was approved as an adjunctive treatment for partial seizures in 2000. A number of small case series have suggested some benefit of adding 1,000 to 4,000 mg/day of levetiracetam to standard mood stabilizers in the treatment of resistant mania. Likewise, adjunctive treatment with zonisamide (Zonegran) may have benefits for both the manic and depressed phase of bipolar disorder in open-label studies. Like topiramate, zonisamide has been associated with weight loss.

Mood Stabilizer–Antipsychotic Combinations

Antipsychotics have long been used as adjunctive treatments in bipolar disorder. Historically, the antipsychotic was used to treat acute mania or the psychotic symptoms and agitation associated with mania. Most of the second-generation antipsychotics are now approved for the treatment of acute mania, and a number are now approved in the depressed and maintenance phase of bipolar disorder.

Most patients who are hospitalized with a manic episode will be treated with an antipsychotic soon after admission. The combination of lithium with typical antipsychotics has generally been well tolerated. However, pharmacodynamic interactions may increase the risk of EPS in patients treated with the combination. The pharmacodynamics risk of an EPS interaction appears greater with higher lithium serum levels.

Case reports and open studies have reported that adding atypical antipsychotics to lithium or valproate may augment antimanic effects. In addition, controlled trials employing risperidone and olanzapine have shown benefit in augmenting antimanic effects of mood stabilizers. In a pooled analysis of five large controlled trials of second-generation antipsychotics combined with lithium in the treatment of acute mania, there appeared to be about a 20 percent advantage of the combination treatment over monotherapy with lithium or the antipsychotic.

Given that the second-generation antipsychotics appear effective in preventing mania, the antipsychotic is commonly added to lithium or valproate when there are breakthrough symptoms on monotherapy. However, to date there is very little controlled data to evaluate the efficacy of the antipsychotic/mood stabilizer combination in prevention of bipolar cycling. Likewise, while several second-generation antipsychotics are now

approved in the treatment of acute bipolar depression, it is not clear if adding lurasidone, quetiapine, or olanzapine to a mood stabilizer increases the efficacy in any way. Adequately controlled studies on antipsychotic/mood stabilizer combination in the treatment of acute bipolar depression have not been done to date. Still, the combination of one of the approved antipsychotics with either lithium or lamotrigine makes intuitive sense. The utility of antidepressants or second mood stabilizers in bipolar depression is more questionable at this time.

The combination of lithium with the atypical antipsychotic clozapine (Clozaril) is occasionally used in the most treatment-resistant bipolar patients. A potential disadvantage of lithium combined with clozapine is that most of the known cases of NMS associated with clozapine have been associated with concurrent lithium use.

Carbamazepine–antipsychotic combinations may result in a pharmacokinetic interaction that reduces the serum levels of concurrent antipsychotics. A higher dose of the antipsychotic may be required.

Mood Stabilizer–Antidepressant Combinations

The treatment of bipolar depression remains a vexing problem. The mood stabilizers, even in combination, are often inadequate to prevent or treat the depressed phase of the illness alone. On the other hand, the addition of an antidepressant may increase the risk of a switch into mania or rapid cycling. It is even unclear whether the addition of an antidepressant has any benefit in most patients. The few controlled studies to date have not necessarily supported the benefit of adding an antidepressant to a mood stabilizer. For example, the STEP-BD did not find that the addition of antidepressant to lithium or valproate was more effective than adjunctive placebo in the treatment of bipolar depression. In fact, there was a nonsignificant trend favoring the placebo group in a number of secondary measures. However, the rates of switching were also not different between the groups.

TCAs may be more problematic than other classes in exacerbating the illness, but all effective antidepressants, including electroconvulsive therapy, have potential to switch a bipolar patient into a manic episode. Other SNRIs including venlafaxine may share a greater proclivity for inducing switches than bupropion. Bupropion may be less likely to induce rapid cycling or mixed states than either TCAs or venlafaxine, but there have been reported cases of bupropion-induced mixed states and manic episodes. In addition, the data on switch rates with bupropion is limited, although some studies including the Stanley Foundation study on antidepressant treatment in bipolar disorder have suggested that the switch rate into hypomania may be lower with bupropion than with some agents including SNRIs. The addition of antidepressants, including TCAs, to lithium does not appear to increase the risk of switching any more than

lithium alone. Thus, most antidepressants probably can be added to lithium without substantially increasing the risk of a switch. By extrapolation, it would be expected that the addition of an antidepressant to valproate or carbamazepine might also mitigate the risk of a switch to mania. However, prospective trials are lacking on antidepressant–anticonvulsant combinations in the treatment of bipolar depression.

An important question is whether the addition of a second mood stabilizer is more efficacious for the treatment of bipolar depression than the addition of an antidepressant. Although antidepressants may destabilize bipolar cycling in the long run, at least one acute study has suggested that adding an SSRI may be as useful as adding a second mood stabilizer. This finding may be more a reflection of the lack of efficacy of adding a second agent to a mood stabilizer rather than an indication that both antidepressants and a second mood stabilizer are equally effective.

COMBINATION STRATEGIES IN THE TREATMENT OF SCHIZOPHRENIA

Combining Two Antipsychotics

The combination of two or more antipsychotics in a schizophrenic patient who has inadequately responded to one agent is commonly attempted but has been rarely studied. An earlier onset of illness, a longer duration of illness, more frequent hospitalizations, and the presence of violence as a symptom of schizophrenia is more likely to be associated with treatment with a combination of antipsychotics. The augmentation strategy that has been perhaps most investigated in resistant schizophrenia has been the combination of clozapine with another antipsychotic. Since 40 to 70 percent of patients fail to achieve an adequate response to clozapine, clinicians have typically added a second antipsychotic. For example, a number of small studies have suggested a modest benefit of adding risperidone to clozapine-nonresponsive patients. Case reports and small open studies have also suggested a potential role for amisulpride (Solian) and aripiprazole in patients who do not respond to clozapine alone. A meta-analysis of 19 studies examining the combination of antipsychotics in the treatment of resistant schizophrenia concluded that there may be an advantage of combining antipsychotics in some resistant patients. However, the combination of antipsychotics has also been associated with more side effects including an increased risk of parkinsonism and elevated prolactin levels.

Risperidone has been combined with typical antipsychotics to produce some improvement in chronic schizophrenic patients studied prospectively. In addition, risperidone has been added to clozapine with benefit refractory schizophrenic patients. The primary risks of combining antipsychotics may be the increased risk of EPS or the risk of hypotension with low potency and atypical antipsychotics.

Antidepressant–Antipsychotic Combinations in Schizophrenia

Antidepressants are frequently combined with antipsychotics in the treatment of schizoaffective disorder and psychotic depression. In addition, antidepressants have shown benefit in the treatment of negative symptoms. Trazodone produces significant but modest improvements in negative symptoms compared to placebo when added to a standard antipsychotic. Likewise, the addition of a low-dose MAOI or moderate doses of nefazodone (Serzone) to a typical antipsychotic has been found to benefit negative symptoms. Controlled trials of adding SSRIs and mirtazapine to antipsychotics have been reported to benefit negative symptoms. Furthermore, the addition of mirtazapine has also been suggested to help antipsychotic-associated akathisia.

Mood Stabilizer–Antipsychotic Combinations in Schizophrenia

Some reports have suggested that lithium can augment standard antipsychotics in the treatment of schizophrenia. However, most controlled studies have not shown much benefit. In addition, the combination of lithium with standard antipsychotics may increase the risk of EPS.

Lamotrigine has been effective in some studies as an augmenter to clozapine in treatment-resistant schizophrenia. For example, in a double-blind study, 200 mg of lamotrigine was more effective than placebo in improving negative symptoms as well as general psychopathology in treatment-resistant schizophrenia patients who had not responded to clozapine alone. However, a more recent study of the addition of lamotrigine to other second-generation antipsychotics did not show an advantage of adding lamotrigine in the treatment of positive or negative symptoms of schizophrenia. In addition, lamotrigine has been observed to reduce alcohol cravings in at least one study when added to standard antipsychotics in the treatment of schizophrenics with alcohol dependence.

The addition of carbamazepine to standard antipsychotics has not been clearly beneficial. On the other hand, a clear disadvantage of adding carbamazepine to many antipsychotics is that carbamazepine can induce the metabolism of the concurrent antipsychotic and reduce serum levels. Oxcarbazepine is less likely than carbamazepine to have a pronounced effect on the CYP450 system.

Valproate has been added to typical and atypical antipsychotics in resistant schizophrenia and catatonia. The data suggest that there may be a role for the addition of valproate in controlling agitation and perhaps impacting some positive and negative symptoms in schizophrenia. In addition, valproate augmentation of risperidone and olanzapine has been associated with more rapid improvement in schizophrenia. However, most studies have not shown a significant advantage of adding valproate to an antipsychotic in the treatment of schizophrenia.

The use of newer anticonvulsants in combination with antipsychotics is largely untested in schizophrenic patients. The use of gabapentin has anecdotally been beneficial to treat anxiety and agitation in some schizophrenic patients. Likewise, there are anecdotal reports of tiagabine being used for agitation in combination with both typical and atypical antipsychotics in some patients. Depressed schizoaffective patients have sometimes benefited from the addition of lamotrigine. Valproate has been applied for the treatment of agitation in schizophrenia.

Other Augmentation Strategies in Schizophrenia

Abnormalities in glutamate transmission have been proposed to play a role in the pathophysiology of deficit symptoms in schizophrenia. As a result, studies have examined the role of glycine and *D*-cycloserine as augmenters to antipsychotics. Glycine is an amino acid that is an agonist of the NMDA receptor. Early open trials have suggested that glycine helps some schizophrenic patients but worsens symptoms in others. *D*-cycloserine, another agonist of the NMDA receptor, may improve negative symptoms and neuropsychological functioning in patients on standard antipsychotics but not clozapine.

Cyproheptadine (Periactin) is a general serotonin antagonist used primarily to treat migraines and allergies. Its serotonin antagonist properties have made it of interest to clinicians treating negative symptoms since all atypical antipsychotics are also 5-HT_2 antagonists. Cyproheptadine also has the advantage of being anticholinergic and can help with EPS. The problem with cyproheptadine is that most patients find it quite sedating, and 24 mg/day is a large dose for many patients. It is unclear whether smaller doses might work just as well.

FUTURE DIRECTIONS

The monotherapy of major depression, bipolar disorder, and schizophrenia is often insufficient to achieve adequate therapeutic response. Thus the combination of psychotropic agents is often attempted in order to improve outcome. Although augmentation strategies have been best demonstrated in the treatment of resistant depression, combination treatment is common in the treatment of bipolar disorder and increasingly common in the treatment of resistant schizophrenia. Studies such as STAR*D and STEP-BD have provided additional insights into the utility and limitations of combination pharmacotherapy. However, much remains to be learned about optimal strategies for combining agents in the treatment of resistant-psychiatric disorders.

REFERENCES

*BALANCE investigators and collaborators; Geddes JR, Goodwin GM, Rendell J, et al. Lithium

plus valproate combination therapy versus monotherapy for relapse prevention in bipolar I disorder (BALANCE): a randomised open-label trial. *Lancet*. 2010;375(9712):385–395.

Bauer M, Dopfmer S. Lithium augmentation in treatment-resistant depression: meta-analysis of placebo-controlled studies. *J Clin Psychopharmacol*. 1999;19:427–434.

Bowden CL, Singh V, Weisler R, et al. Lamotrigine vs. lamotrigine plus divalproex in randomized, placebo-controlled maintenance treatment for bipolar depression. *Acta Psychiatr Scand*. 2012;126(5):342–350.

Carpenter LL, Jocic Z, Hall JM, Rasmussen SA, Price LH. Mirtazapine augmentation in the treatment of refractory depression. *J Clin Psychiatry*. 1999;60:45–49.

Correll CU, Rummel-Kluge C, Corves C, Kane JM, Leucht S. Antipsychotic combinations vs monotherapy in schizophrenia: a meta-analysis of randomized controlled trials. *Schizophr Bull*. 2009;35(2):443–457.

Dé Montigny C, Grunberg F, Mayer A, Deschenes JP. Lithium induces rapid relief of depression in tricyclic antidepressant drug non-responders. *Br J Psychiatry*. 1981;138:252–256.

*Fava M, Rush AJ, Wisniewski SR, et al. A comparison of mirtazapine and nortriptyline following two consecutive failed medication treatments for depressed outpatients: a STAR*D report. *Am J Psychiatry*. 2006;163(7):1161–1172.

Geller B, Luby JL, Joshi P, et al. A randomized controlled trial of risperidone, lithium, or divalproex sodium for initial treatment of bipolar I disorder, manic or mixed phase, in children and adolescents. *Arch Gen Psychiatry*. 2012;69(5):515–528.

Goodwin FK, Prange AJ Jr, Post RM, Muscettola G, Lipton MA. Potentiation of antidepressant effects by L-triiodothyronine in tricyclic nonresponders. *Am J Psychiatry*. 1982;139:34–38.

Grosso G, Pajak A, Marventano S, et al. Role of omega-3 fatty acids in the treatment of depressive disorders: a comprehensive meta-analysis of randomized clinical trials. *PLoS One*. 2014;9(5):e96905.

*Joffe RT, Levitt AJ, Bagby RM, MacDonald C, Singer W. Predictors of response to lithium and triiodothyronine augmentation of antidepressants in tricyclic non-responders. *Br J Psychiatry*. 1993;163:574–578.

Johns CA, Thompson JW. Adjunctive treatments in schizophrenia: pharmacotherapies and electroconvulsive therapy. *Schizophr Bull*. 1995;21:607–619.

Juruena MF, Ottoni GL, Machado-Vieira R, et al. Bipolar I and II disorder residual symptoms: oxcarbazepine and carbamazepine as add-on treatment to lithium in a double-blind, randomized trial. *Prog Neuropsychopharmacol Biol Psychiatry*. 2009;33(1):94–99.

Keck PE Jr, McElroy SL, Vuckovic A, Friedman LM. Combined valproate and carbamazepine treatment of bipolar disorder. *J Neuropsychiatry Clin Neurosci*. 1992;4:319–322.

Loebel A, Cucchiaro J, Silva R, et al. Lurasidone as adjunctive therapy with lithium or valproate for the treatment of bipolar I depression: a randomized, double-blind, placebo-controlled study. *Am J Psychiatry*. 2014;171(2):169–177.

*Nierenberg AA, Ostacher MJ, Calabrese JR, et al. Treatment-resistant bipolar depression: a STEP-BD equipoise randomized effectiveness trial of antidepressant augmentation with lamotrigine, inositol, or risperidone. *Am J Psychiatry*. 2006;163(2):210–216.

Nierenberg AA, Price LH, Charney DS, Heninger GR. After lithium augmentation: a retrospective follow-up of patients with antidepressant-refractory depression. *J Affect Disord*. 1990;18:167–

175.
Ouyang WC, Hsu MC, Yeh IN, Kuo CC. Efficacy and safety of combination of risperidone and haloperidol with divalproate in patients with acute mania. *Int J Psychiatry Clin Pract*. 2012;16(3):178–188.
Palinkas LA, Barrett-Connor E. Estrogen use and depressive symptoms in postmenopausal women. *Obstet Gynecol*. 1992;80:30–36.
Peselow ED, Fieve RR, Difiglia C, Sanfilipo MP. Lithium prophylaxis of bipolar illness. The value of combination treatment. *Br J Psychiatry*. 1994;164:208–214.
Post RM, Leverich GS, Nolen WA, et al. A re-evaluation of the role of antidepressants in the treatment of bipolar depression: data from the Stanley Foundation Bipolar Network. *Bipolar Disord*. 2003;5(6):396–406.
*Sachs GS, Nierenberg AA, Calabrese JR, et al. Effectiveness of adjunctive antidepressant treatment for bipolar depression. *N Engl J Med*. 2007;356:1711–1722.
Sagud M, Vuksan-C´usa B, Zivkovic´ M, et al. Antipsychotics: to combine or not to combine? *Psychiatr Danub*. 2013;25(3):306–310.
Shelton RC, Tollefson GD, Tohen M, et al. A novel augmentation strategy for treating resistant major depression. *Am J Psychiatry*. 2001;158:131–134.
Szegedi A, Verweij P, van Duijnhoven W, Mackle M, Cazorla P, Fennema H. Meta-analyses of the efficacy of asenapine for acute schizophrenia: comparisons with placebo and other antipsychotics. *J Clin Psychiatry*. 2012;73(12):1533–1540.
Thase ME, Howland RH, Friedman ES. Treating antidepressant nonresponders with augmentation strategies: an overview. *J Clin Psychiatry*. 1998;59 Suppl 5:5–12.
*Trivedi MH, Fava M, Wisniewski SR, et al. Medication augmentation after the failure of SSRIs for depression. *N Engl J Med*. 2006;354(12):1243–1252.
Wolkowitz OM, Reus VI, Roberts E, et al. Dehydroepiandrosterone (DHEA) treatment of depression. *Biol Psychiatry*. 1997;41:311–318.
Woo YS, Bahk WM, Chung MY, et al. Aripiprazole plus divalproex for recently manic or mixed patients with bipolar I disorder: a 6-month, randomized, placebo-controlled, double-blind maintenance trial. *Hum Psychopharmacol*. 2011;26(8):543–553.
Young LT, Joffe RT, Robb JC, MacQueen GM, Marriott M, Patelis-Siotis I. Double-blind comparison of addition of a second mood stabilizer versus an antidepressant to an initial mood stabilizer for treatment of patients with bipolar depression. *Am J Psychiatry*. 2000;157:124.

▲ 34.39 Reproductive Hormonal Therapy: Theory and Practice

DAVID R. RUBINOW, M.D., AND PETER J. SCHMIDT, M.D.

The 19th-century medical literature contained several presumptions about the pathophysiology of mood disorders in women, largely based on anecdotal observations of reproductive endocrine dysfunction (e.g., amenorrhea) in psychiatrically ill women. These inferences, in turn, were translated into therapeutics. Thus, numerous reports also documented the beneficial effects on mood and behavior associated with medical or surgical

manipulations of a woman's reproductive function. In addition to their interest in the role of reproductive function in psychiatric disorders, medical researchers in the late 19th century also developed a general interest in glandular secretions that potentially could modify physiology. Early experiments in humans with preparations derived from animal glands including thyroid, adrenal, ovary, testes, and spleen provided the directions for the developing fields of endocrinology and immunology. In his seminal lecture in 1889, Charles-Édouard Brown-Séquard not only originated the formal study of endocrinology, he also reported the potential psychotropic effects of gonadal secretions. His description of the invigorating effects of testicular extracts resulted in a brief but widespread use of these compounds (as well as other interventions intended to increase the body's testicular or seminal fluid levels) in a variety of therapeutic settings. In addition to extracts of both testes and ovaries, investigators experimented with extracts of the thyroid and adrenal glands in psychiatric patients. This interest in the therapeutic use of thyroid and adrenal extracts enjoyed more enduring success with the recognition of their beneficial effects in the treatment of myxedema and Addison disease, respectively. However, the widespread use of gonadal extracts ultimately met with criticism due to some disappointing results and the realization that these extracts were not the "fountain of youth." Their subsequent clinical use was largely restricted to the treatment of menopause-related hot flushes (ovarian extracts). Nonetheless, the observation that ovarian extracts inhibited ovulation in rabbits was sufficient to maintain medical interest in the potential role of gonadal extracts for the control of human fertility. This process of promise and disappointment, passionate advocacy and evidence-based caution, is the story of both the birth and the ongoing existence of endocrinology. Many of the claims of the organotherapists were refuted by observations that most of the endocrine preparations were inactive when administered orally. Nonetheless, the so-called "practical men," the practitioners whose therapeutic zeal was not about to be extinguished by study results to the contrary, continued to claim efficacy of endocrine preparations for all manner of illnesses. An additional source of the pendular swing of practice, and one particularly noteworthy for reproductive preparations, was the recognition of previously undiscovered side effects, a phenomenon that is particularly relevant for the therapeutic use of estradiol, the primary active substance in ovarian extracts. Efforts to improve techniques of extraction and purification of biologically active substances from the gonads were fueled by the hope that factors regulating reproductive function would be identified. By the late 1920s and early 1930s, many gonadal steroids were isolated and characterized including estrone, estradiol, progesterone, and several androgens. Moreover, during the next 10 years chemists identified modifications of the steroids that could alter their absorption (i.e., acetylation) and potency (e.g., addition of

ethinyl group or removal of C-19 methyl groups). These findings initiated a resurgence in the medical use of gonadal steroids. Estrogen replacement therapy was used to treat menopausal symptoms in the 1950s, and oral contraceptives (i.e., mestranol and norethynodrel [Enovid]) were developed and first approved by the FDA in 1960. The use of gonadal steroids in women once again was widespread, and several papers were published reporting the therapeutic benefits of these compounds in involutional melancholia, premenstrual syndrome, and postpartum depression. However, their widespread usage was restricted—the pendular swing—after reports in the 1970s of increased rates of endometrial cancer secondary to unopposed estrogen replacement and increased rates of thrombosis and pulmonary emboli in women taking oral contraceptives. In the last two decades of the 20th century, popularity for gonadal steroid therapy resumed due in part to the reports of the enhanced safety of both hormone replacement therapy (HRT) (due to the protective effect on the endometrium of the addition of progesterone) and oral contraceptives (due to the lower doses of estradiol and new progestins) and to the reported beneficial/disease protective effects of gonadal steroids on several tissues including the musculoskeletal, cardiovascular, and central nervous systems. The discovery of several other factors involved in the control of reproduction also led to new drug development. The decapeptide GnRH was isolated and sequenced in the 1970s, and the observation that continuous GnRH infusion resulted in the downregulation of pituitary GnRH receptors led to the development of several GnRH agonists. These agonists were used to suppress reproductive endocrine function in a variety of medical conditions including hormone-dependent cancers and endometriosis. In combination with gonadal steroids, preparations of GnRH agonists provided physicians with a strategy to control reproductive function and regulate the exposure to specific gonadal steroids without resorting to surgery. Thus physicians could selectively eliminate and/or replace reproductive factors considered to be the potential source of a medical or psychiatric problem.

More recently, however, reports of increased risk of breast cancer and, particularly, the Women's Health Initiative demonstration of increased, not decreased, risks of both heart disease and dementia after hormone therapy have resulted once more in a mass exodus from reproductive hormonal therapy (HT). Subsequent findings suggest that the pendulum again may have swung too far (this time, in the direction of avoiding hormone therapy). Nonetheless, the conclusions that follow about the practical and potential applications of hormonal therapy in psychiatric disorders must be viewed in the evolving context in which HT now resides.

Reproductive steroids are powerful and complex neuroregulators. If one is to knowledgeably, responsibly, and creatively employ reproductive steroids in the psychotherapeutic armamentarium, some familiarity with

the structure and related complexity of activities of these compounds is required.

STRUCTURE AND ACTIONS

Gonadal steroids, like all steroid hormones, are derivatives of cholesterol, and all share a core four-ring structure. Despite the wide array of steroid hormones, the synthesis of individual steroids is performed by a much smaller number of enzymes with multiple sites of action in the hormone synthetic cascade. Indeed, during the 1980s the cloning of the cDNAs and genes of several steroidogenic enzymes led to the realization that multiple different steroids could serve as substrates for the same steroidogenic enzyme in most tissues. In the past two decades, there has been an explosion of findings that have demonstrated the enormous complexity of steroid synthesis. These findings include the identification of multiple isozymes for steroid synthetic enzymes (including 14 different isozymes of 17-beta hydroxysteroid dehydrogenase encoded by several different genes) and the presence of a range of alternate synthetic pathways leading to steroid production in both health and disease states (e.g., the "backdoor pathway" from progesterone to dihydrotestosterone without the usual androgen intermediates). In humans, plasma low-density lipoproteins are the major source of tissue cholesterol used for synthesizing steroids. The first step in steroid synthesis takes place in the mitochondria. The mechanism by which cholesterol is transported and deposited onto the outer mitochondrial membrane remains only partially understood. However, in steroidogenic cells, steroidogenic acute regulatory protein (STAR) and other members of the STAR-related lipid transfer family of proteins control the availability of cholesterol within the mitochondria, where it is cleaved in a series of steps by the P450 side chain cleavage enzyme (P450scc) to produce pregnenolone. Pregnenolone then serves as the precursor of the family of steroid hormones, individual members of which as stated above are generated through the actions of a relatively small number of enzymes with multiple sites of action (see Fig. 34.39–1). The resulting end products of this cascade are determined by the tissue in which the metabolism is occurring and the enzymes present in that tissue. For example, testosterone in both men and women may be the end product and act directly at the androgen receptor, or it may be reduced to a form with greater affinity for the androgen receptor (dihydrotestosterone), converted to a form with less affinity for the androgen receptor (androsterone—also a neurosteroid active at the GABA receptor), or aromatized to estradiol and act through the estrogen receptor. Recently, a new level of complexity has been introduced by the discovery of tissue-specific variations in the activity of steroidogenic enzymes. For example, 17-beta hydroxysteroid dehydrogenase exists in both oxidative and reductive forms: the type 2 isoform in the gut is oxidative and converts

estradiol to estrone (thus "inactivating" orally administered estradiol), while the type 1 isoform in the ovary and placenta is reductive and "activates" estradiol via its conversion from the less potent estrone. Further, a genetic variation in 5-alpha reductase type 1, an enzyme that converts testosterone to dihydrotestosterone and is localized in skin and the central nervous system, has been associated with enhanced hirsutism in women with polycystic ovarian syndrome. Thus polymorphic variants of steroidogenic enzymes can produce tissue-specific alterations in exposure to steroids with accompanying alterations in phenotype. Moreover, as an example of the multifunctional roles of steroidogenic enzymes, 5-alpha reductase also is the rate-limiting step in the conversion of progesterone to its neurosteroid metabolite allopregnanolone, which is active at the GABA receptor. Finally, given the structural similarities shared across different types of steroid hormones, steroids are classified by the particular receptor that mediates their actions rather than the structural elements of the hormone or the tissue of origin. By this classification, reproductive steroids are those that act via the androgen receptor, estrogen receptor, or progesterone receptor systems.

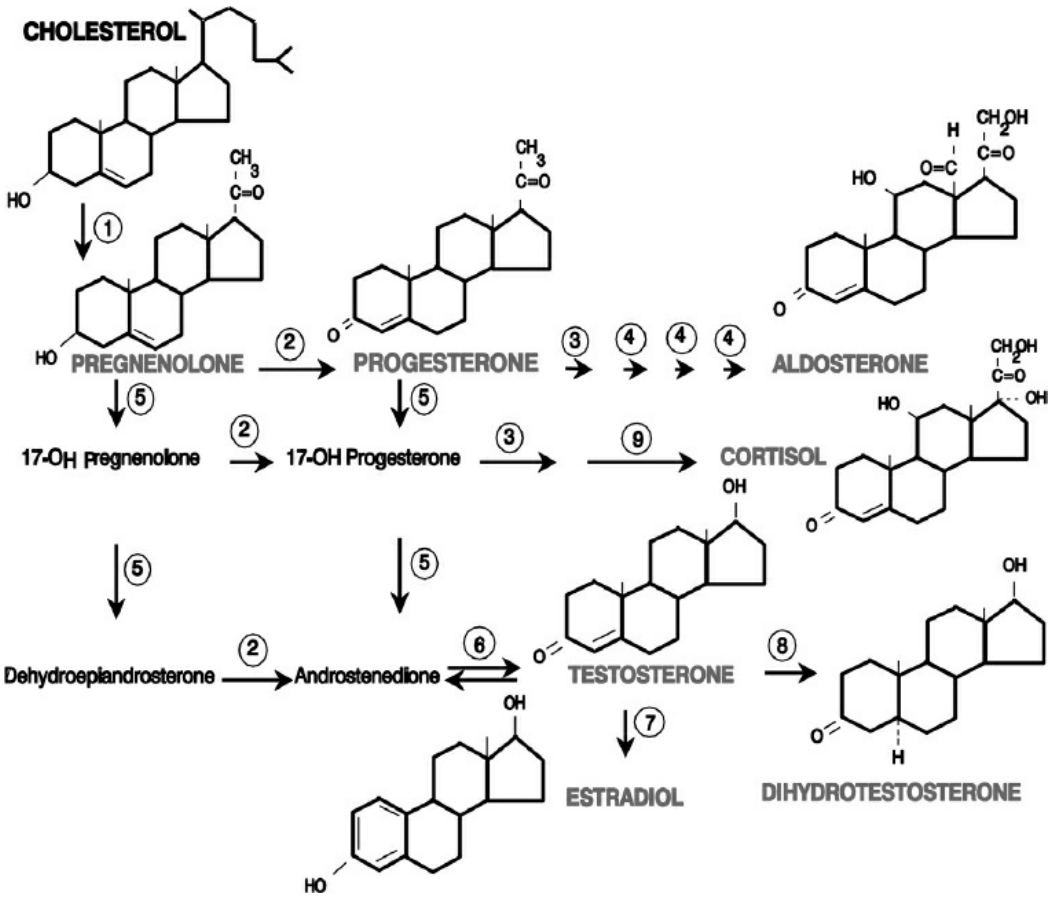

FIGURE 34.39–1. Synthetic pathways for steroid hormones. Circled numbers identify synthetic enzymes: 1, cytochrome P450 (CYP) 11A (P450 side chain cleavage enzyme); 2, 3β-hydroxysteroid dehydrogenase; 3, CYP21 (21-hydroxylase); 4, CYP11B2 (11β-hydroxylase, 18-hydroxylase, 18-oxidase); 5, CYP17 (17α-hydroxylase, 17,20-lyase); 6, 17β-

hydroxysteroid dehydrogenase (or oxidoreductase); 7, aromatase; 8, 5α-reductase; 9, CYP11B1 (11β-hydroxylase). (From Rubinow DR, Schmidt PJ. Androgens, brain, and behavior. *Am J Psychiatry*. 1996;153(8):974–984.)

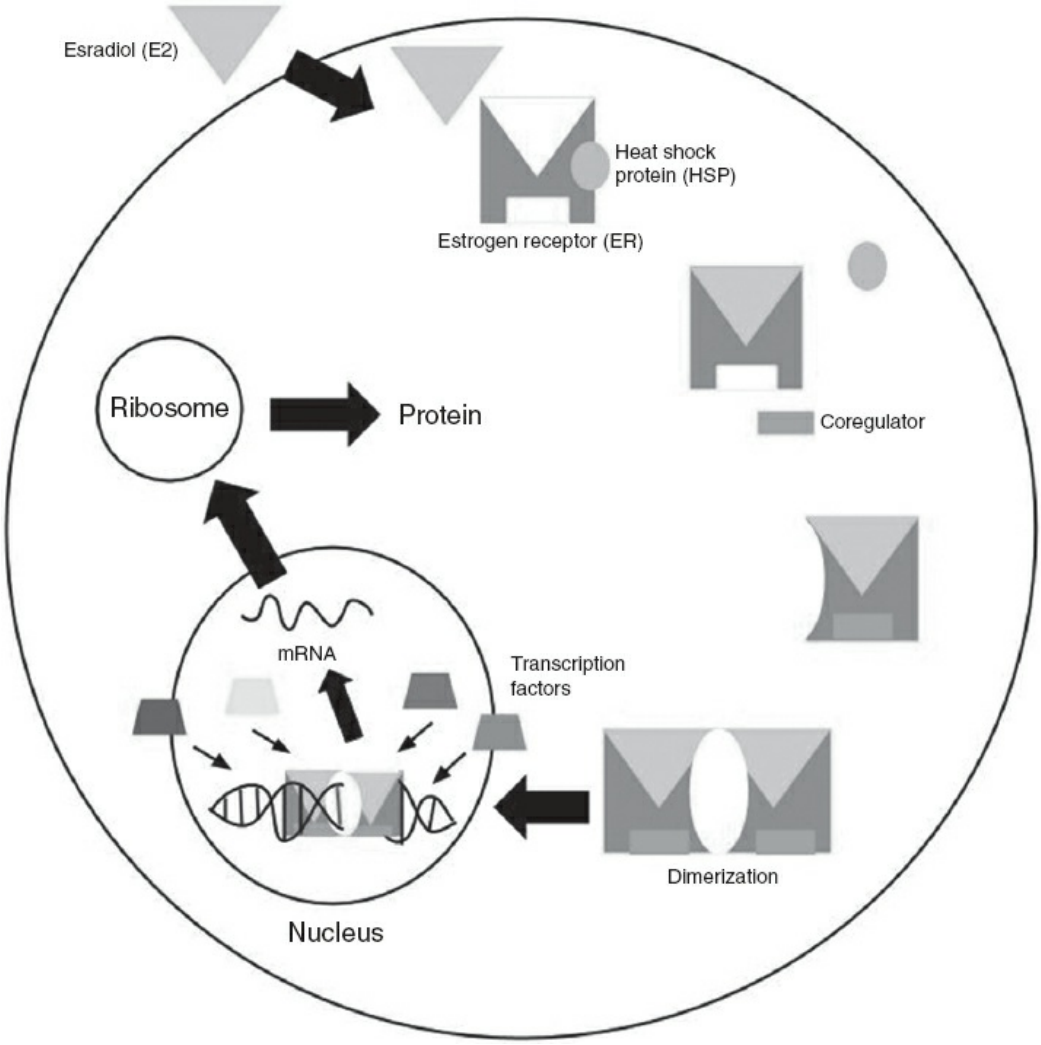

FIGURE 34.39–2. Genomic action of steroid hormones. Estradiol acts by binding (activating) its intracellular receptor, which dissociates from heat shock protein. The activated receptor undergoes a conformational change that facilitates the binding of a coregulator (in this case, a coactivator) with enzymatic properties. The receptor dimerizes, and the receptor/coregulator complex both binds the DNA hormone response element and recruits a variety of transcription factors. The DNA is then remodeled to facilitate transcription of mRNA, which is exported to the cytoplasmic ribosome for protein synthesis.

The classic genomic action of steroid hormones occurs after the steroid binds (activates) its intracellular receptor, which, after undergoing phosphorylation, binds (usually as a dimer) to a hormone response element on a gene and directs or modifies transcription of that gene (see Fig. 34.39–2). Note that in contrast to the membrane-bound receptors of neurotransmitters and neuropeptides, which act on the genome through a signal transduction cascade, the steroid hormone receptors are intracellular and are transcription factors that can directly bind to DNA. Further, once

the ligand–receptor complex binds to the hormone response element (which is located on or near the promoter of the gene), the result can be either the initiation or repression of transcription of messages coding for an enormous array of proteins, including synthetic and metabolic enzymes for neurotransmitters, neuropeptides, receptor proteins, transporters, and second messengers.

Alas, while accurate, this mechanistic description is glaringly inadequate, and it is in the missing complexity that the explanation for so many ostensibly idiosyncratic responses to hormones lies. First, an additional source of transcriptional variance is found in the different forms of the gonadal steroid receptors. For example, isoforms of the androgen and progesterone receptors exist that have different transcriptional effects, and two separate estrogen receptors, alpha and beta, have been identified, coded for by genes on the 6th and 14th chromosomes, respectively. These estrogen receptors have different distributions in the brain, different transcriptional actions, and even possess additional isoforms (e.g., insertional and deletional variants of ER beta). Additionally, recent studies have identified splice variants of steroid receptors present in some disease states, including a splice variant in the androgen receptor that lacks a ligand-binding domain but is constitutively active (and, therefore acts on the androgen receptor in the absence of steroid) and contributes to the development of hormone-refractory prostate cancer. Second, the activated steroid receptor regulates transcription by binding protein intermediaries called coregulators. Some coregulators stimulate transcription while others inhibit it. The activities of these coregulators are tissue specific, and different types of coregulators interact with specific isoforms of the steroid receptors. Thus, the relative expression of coregulators and receptor isoforms within different tissues could underlie, in part, opposite effects of the same hormone in different tissues. Approximately 350 of these coregulators have been identified, with coregulators usually part of assemblies consisting of up to eight different regulators. Since the coregulator function is altered by covalent modification of its constituent amino acid residues, it has been estimated that there are 10^{13} possible different functional states that can be created in response to hormone activation. Third, activated hormone receptors can influence the transcription of genes that do not possess hormone response elements through interactions with other cellular proteins called cointegrators, which translate hormone signals to oncogene and signal transduction–related genes. Fourth, hormones display acute nongenomic actions; that is, they can, in a receptor isoform–specific fashion, allosterically modify the activity of neurotransmitter gated ion channels, can directly bind and regulate ion channels, and can initiate signal transduction through binding membrane rather than cytoplasmic or nuclear receptors. Fifth, hormone receptors can be activated, even in the absence of hormones, by a variety of

neurotransmitters and growth factors. In this fashion environmental stressors may amplify the response to even minimal levels of a hormone signal. Sixth, hormones may influence each other's activity by competing for cofactors or by producing opposite effects on genes regulated through integrator proteins.

What are the practical implications of this biologic complexity? First, hormones of a given class (i.e., different hormonal preparations) are not going to produce similar effects on all tissues. For example, medroxyprogesterone differs from progesterone in at least two ways: first, the latter but not the former can be metabolized to neuroactive compounds called neurosteroids, which regulate $GABA_A$ receptor activity; and second, medroxyprogesterone inhibits the enzyme that converts pregnenolone to progesterone (i.e., 3-beta hydroxysteroid dehydrogenase) thereby further decreasing substrate for the production of endogenous neurosteroids. At a cellular level, it is unlikely that the equine estrogens used so frequently in HRT produce the same conformational changes in the ER as estradiol and hence would have a different profile of actions. Indeed, the genetic effects of conjugated equine estrogens are not identical to those of estradiol. Second, hormone plasma levels are only part of the story. Hormone effects are context specific and reflect the constituents of the milieu that bathes the cell as well as the specific constellation of expressed proteins within the cell. Depending on one's genetic background, then, one might expect that the same hormone level would elicit a range of possible effects. This has been clearly demonstrated in animal studies in which exogenous estradiol exerts 100-fold differences in effects on reproductive development across rodent strains, accounting for far more of the variance than the plasma level of estradiol achieved. Third, recognition of the critical nature of intracellular proteins and hormone signaling/transcription suggests a strategy for future drug development, namely designing drugs that will facilitate or disrupt binding interactions between the receptor and specific cofactors. This strategy has already borne fruit in the development of selective estrogen receptor modulators (SERMs) and selective androgen receptor modulators (SARMs) (discussed below).

WHY REPRODUCTIVE HORMONAL THERAPY?

It is biologically plausible that reproductive steroids might display psychotropic properties, as their pattern of neuroregulatory effects is highly convergent with those presumed to underlie psychotropic efficacy.

Indeed, wherever one finds a system believed to play a role in the etiology or treatment of depression, modulatory effects of gonadal steroids are observed as well. The neurotransmitter systems implicated in depression—serotonin, norepinephrine, dopamine, acetylcholine, GABA, glutamate—all are regulated by estradiol. Estradiol's regulation of the serotonergic system, for example, is extensive, involving serotonin

synthesis (tryptophan hydroxylase), receptors (transcripts, protein, and binding) (e.g., 5-HT1A, 2A), and transporter. Additionally, estrogen increases sensitivity to dopamine and cocaine and is believed to contribute to the increased vulnerability to substance abuse in women. Recent hypotheses about mood dysregulation center on deficits in neuroplasticity and neuronal/glial survival, with attendant observations that cell survival proteins like Bcl-2 or growth factors like BDNF are low in depression. Stress decreases BDNF in animal models, and both Bcl-2 and BDNF are increased by antidepressants and mood stabilizers, effects also seen after exposure to estradiol. Cell survival is regulated by the actions of several signal transduction systems, including ERK, Akt, WNT/beta catenin. In a brain region–specific fashion, estradiol increases ERK and cell survival in neurons (although it decreases both in glia), increases AKT and neuronal survival, increases trkA and phospho-CREB, and decreases GSK-3 beta. Similar effects are seen with antidepressants and mood stabilizers on ERK (increased) and GSK-3 beta (decreased). Finally, estrogen appears to play a major role in neuronal excitability, as it directly and acutely (in seconds to minutes) modulates both calcium and potassium ion channel activity.

Similarly, many areas of the brain that have been implicated in depression, including prefrontal cortex (PFC), amygdala, hippocampus, striatum, and thalamus contain estrogen receptors and are modulated by estradiol. For example, in the first demonstration of the effects of estradiol and progesterone on regional cerebral blood flow in humans, Berman et al., showed that both estradiol and progesterone increased cerebral blood flow in the dorsal lateral PFC, inferior parietal lobule, and posterior inferior temporal cortex compared with a hypogonadal condition produced by gonadotropin-releasing hormone agonists. More recent studies suggest that estradiol and progesterone regulate other brain regions in humans including the orbitofrontal cortex, amygdala, striatum, and midbrain—areas involved in the reward system. Further, regional functional connectivity—a better measure of network activity—is modulated both by the menstrual cycle and exogenous reproductive steroid administration, with the nature of the alterations both brain region specific and steroid dependent. Finally, the importance of gonadal steroids in affective regulation is suggested by studies demonstrating the role of estradiol in stress amplification, of ER beta in anxiety in female rats, and of ER alpha in arousal.

Progesterone is metabolized to the neurosteroid allopregnanolone, the levels of which have been shown to be decreased in depression and increased during treatment with a selective serotonin reuptake inhibitor (SSRI). Indeed, SSRIs increase the activity of the neurosteroid synthetic enzyme 3-alpha-hydroxysteroid oxidoreductase, and it has been hypothesized that the resultant increase in allopregnanolone may underlie the therapeutic efficacy of SSRIs in premenstrual dysphoria. Progesterone,

through allopregnanolone, modulates both the conformation and the ligand sensitivity of $GABA_A$ receptors, with a role for allopregnanolone implicated in both the genesis and treatment of anxiety (depending upon the time course of administration) and in the sensitivity of $GABA_A$ receptor to the anxiolytic benzodiazepines. Finally, allopregnanolone also exerts neuroprotective effects by reducing the expression of proapoptotic proteins and apoptotic DNA fragmentation, thereby reducing cell death and gliosis, which have been associated with depression. Neuroprotection may be mediated by immune regulation in depression, and both progesterone and allopregnanolone reduce the expression of the "proinflammatory" cytokine TNF-α, which is elevated in individuals with major depression. Thus, allopregnanolone (or progesterone) may exert antidepressant effects by reducing the physiologic impact of stress, promoting neuroprotection, and protecting against the immune activation and cytokine hypersecretion associated with major depressive disorder. As such, both progesterone and therapies that target progesterone metabolism might be expected to modulate mood.

TYPES OF REPRODUCTIVE HORMONE THERAPY IN AFFECTIVE ILLNESS

Reproductive hormone therapies serve two principal functions. First, they potentially may replace traditional pharmacotherapies if the latter have proved ineffective or are accompanied by unacceptable side effects. Second, some reproductive hormone therapies can also serve as short-term trials to inform patients and their health care providers about possible links between the secretion of a particular hormone and the presence or exacerbation of symptoms (e.g., GnRH agonist–induced ovarian suppression in menstrual cycle–modulated affective disorders, as described below).

Hormone Suppression in Premenstrual Dysphoric Disorder

A variety of hormonal therapies have been employed in premenstrual dysphoric disorder (PMDD). Currently the best evidence of efficacy exists for gonadotropin-releasing hormone agonists, compounds that reversibly suppress pituitary secretion of the gonadotropins FSH and LH and hence turn off ovarian steroid secretion. GnRH-induced ovarian suppression works in about two-thirds of women with PMDD, the same percentage (although not necessarily the same individuals) as are successfully treated by SSRIs. GnRH therapy should be considered for those women who do not respond to SSRIs, particularly if they have severe symptoms. While few predictors of therapeutic response to GnRH agonists exist, one report suggests that the presence of symptoms in the follicular phase predicts lack of efficacy (although the women in this report would not meet criteria for PMDD and more likely experienced a premenstrual exacerbation of a

chronic depressive illness). Because of the largely unknown risk of compromising fertility, and of nonreversible reduction in bone mass, this treatment should be reserved for those who do not intend to subsequently have children and should be administered only in consultation with a gynecologist and with full consideration of risks and benefits. Recently, concerns over the safety of long-term GnRH agonist–induced gonadal suppression in both fertility and bone health have been offset in part by reports of the return of normal reproductive function and bone mass in the majority of children with precocious puberty who were treated with these agents for several years. Nonetheless, prudent clinical practice would entail the use of GnRH agonists for not greater than 3 months without consultation with a woman's gynecologist and without consideration of the addition of bone-sparing adjunctive estradiol (which, if given continuously, would not be expected to reverse the efficacy of the GnRH agonist). Preliminary data suggest that while exposure to sequential hormone replacement (i.e., discontinuous or interrupted) does precipitate a recurrence of PMDD symptoms, continuous exposure does not. Consequently, continuous administration of oral contraception may show therapeutic efficacy in PMDD despite the inefficacy of discontinuous or sequential administration. Indeed, the shortened pill-free interval appears critical for the reported efficacy of an oral contraceptive approved for PMDD. If efficacy is demonstrated in ongoing trials, an alternative to SSRIs will be available. It should be noted that continuous oral contraception administration requires monitoring of the endometrium to avoid endometrial hyperplasia, but several oral contraceptives are approved to be administered continuously for 3 months at a time. One additional advantage to a therapeutic trial of GnRH agonist is that the results can determine the role of cyclical variation of steroid secretion in PMDD. If the patient does not respond to GnRH agonist–induced suppression (nonresponse is reported in about 30 percent of women with PMDD despite adequate ovarian suppression), you should look no further for a reproductive hormone etiology or intervention. Other hormones studied as potential treatments for PMDD include progesterone and estrogen. Despite early claims to the contrary, the majority of studies overwhelmingly refute the efficacy of progesterone in PMDD. While estradiol is described as superior to placebo in several studies, evidence to date must be regarded as preliminary and insufficient to support the recommended use of estradiol for the treatment of PMDD. Indeed, the doses of estradiol required to effectively treat PMDD suggest that in this condition estradiol is acting to inhibit ovulation and interfere with normal cyclic ovarian hormone secretion. No studies have examined the therapeutic potential of GnRH antagonists, which have recently become available in an oral preparation. Finally, preliminary evidence in a small sample of women with PMDD suggests that stabilizing neurosteroid levels across the menstrual cycle by

inhibiting 5-alpha reductase type 1 (the rate-limiting enzyme for the conversion of progesterone to allopregnanolone) with the medication dutasteride results in the elimination of PMDD symptoms. These findings hold great potential for understanding the role of neurosteroids in PMDD but must be viewed with caution until larger trials with longer treatment duration can confirm these results.

Although not unrelated to the treatment of women with PMDD, a trial of GnRH agonist–induced ovarian suppression can be undertaken in some women who present with a recurrent affective disorder and in whom the clinical presentation suggests a menstrual cycle recrudescence of affective symptoms. In this clinical scenario the results of the trial of ovarian suppression informs future decisions to include (or not) hormonal therapies to augment existing mood-stabilizing pharmacotherapy.

Estradiol Administration in Perimenopausal Depression and Postpartum Depression

Perimenopausal Depression

A 48-year-old gravida 2 para 2 married Caucasian female presented with an 8-month history of depressed mood, anhedonia, increased anxiety, tearfulness, avoidance of social activities, disturbed sleep, and decreased concentration. She did not feel that she was sad so much as, paradoxically, both more worried about even relatively trivial matters and less engaged by events in her life. She felt herself to be more emotionally reactive, finding herself becoming tearful in response to almost any emotionally laden situation. On questioning, the patient reported hot flushes, worse at night; however, daytime flushes were associated with marked anxiety, a racing heartbeat, and an enhanced sense of worry. Her menstrual cycle had been somewhat erratic in length for 2 years, and in the last year she had skipped several menses with periods of amenorrhea lasting 2 to 3 months. The patient had no past personal or family history of depression. Laboratory testing revealed plasma FSH levels ranging from 25 to 40 IU/L and estradiol levels of 50 to 60 pg/mL.

FSH levels greater than two standard deviations above the mean for women of reproductive age combined with irregular menstrual cycles and amenorrhea characterize the menopause transition and the first year postmenopause (i.e., the perimenopause).

Therapeutic trials of estradiol therapy (ET) have been conducted in perimenopausal and postmenopausal women with depression, dating back to 1934. Depression risk increases during the perimenopause, and depression is cited as a primary reason for resuming menopausal hormone therapy (HT). Community-based epidemiologic studies document a 1.5- to 3-fold greater risk of first onset and recurrent depressions in women during the perimenopause compared with those who are premenopausal (or who are several years postmenopausal). Observational studies report the emergence of depressive symptoms after the discontinuation of HT in 5 to 10 percent of women.

The role of estradiol (E2)—either declining or low levels—in the

precipitation of perimenopausal depression is unknown, largely due to the associational and indirect nature of the evidence linking ovarian function and depression and to the inability to accurately measure the low levels of plasma estradiol levels in some perimenopausal women. Correlative studies with plasma FSH or response to hormonal therapy do provide indirect evidence of the relevance of changes in reproductive hormones to mood disturbances. However, even prospective epidemiologic studies cannot test the estradiol withdrawal hypothesis, because perimenopausal changes in the secretion of several hormonal and metabolic factors could confound the effects of estradiol withdrawal. In other studies of the role of ovarian steroids in affective disturbance (e.g., PMDD and postpartum depression), changes in ovarian steroids (in the context of otherwise normal levels) have been shown to directly trigger depression, but only in a subset of women with histories of mood disorders linked to reproductive function. Recently, a study directly examined the hypothesis that declining ovarian function—estradiol "withdrawal"—underpins depression occurring during the perimenopause. The authors further sought to examine whether the response to experimentally induced estradiol withdrawal differed in women with a past perimenopausal depression compared with those without any past depression. Results showed that women with a past perimenopausal depression who were blindly withdrawn from ET experienced a significant increase in depression symptoms. Women with past perimenopausal depression who continued on estradiol and control women with no past history of depression remained asymptomatic. Thus, in women with past perimenopausal depression, the recurrence of depressive symptoms during blinded estradiol withdrawal suggests that normal changes in ovarian estradiol secretion can trigger an abnormal behavioral state in these susceptible women. These data are the first direct evidence that, in some women, estradiol withdrawal could contribute to the onset of depression during the perimenopause.

While several studies clearly suggested the beneficial effects of estrogen on mood in perimenopausal women, the antidepressant efficacy of estradiol remains unclear. A double-blind, placebo-controlled trial demonstrated that 3 weeks of estradiol significantly decreased depression ratings compared with placebo in perimenopausal women. A full or partial therapeutic response was seen in 80 percent of subjects on estradiol but in only 22 percent of those on placebo, consistent with the observed effect size ($d = 0.69$) in a previous meta-analysis of studies examining estrogen's effect on mood. The therapeutic response to estradiol was observed in women with major or minor depression and was observed irrespective of the antecedent presence of hot flushes, thus removing a major confound complicating interpretation of earlier studies (i.e., effects of estradiol on mood might merely be consequent to resolution of hot flushes and associated sleep disturbance). Efficacy of estradiol was similarly

demonstrated by other investigators in perimenopausal women with major depression but not in postmenopausal women. The small samples and short duration of treatment in trials of ET in perimenopausal depression notwithstanding, the best case for ET in the treatment of depression (either as monotherapy or as an adjunctive therapeutic agent) can be made in perimenopausal rather than postmenopausal women. It is further noteworthy that there are at present no clear predictors (including estradiol levels) of response to estradiol in women with perimenopausal depression.

While initial reports suggested beneficial effects of estradiol on both cognitive performance (e.g., verbal memory) and cognitive pathology (i.e., Alzheimer dementia), consistent, perhaps, with its neuroprotective effects, subsequent reports demonstrate a far more restricted impact on cognition if any (e.g., limited to women with perimenopausal symptoms). The timing of starting ET relative to the menopause has been proposed as one factor explaining the otherwise discordant observations (of the effects of ET on cognition) between the observational studies and the recent randomized controlled trials from the Women's Health Initiative (WHI). Thus, in several studies women who take ET during the late menopause transition or early postmenopause have a lower risk of dementia than those starting ET later. This construct of a reproductive stage-dependent differential "risk" for both the benefits of ET on cognitive performance and the adverse effects in young peri- and postmenopausal women compared with older postmenopausal women has recently been tested. Two primary prevention, randomized, controlled trials examining the effects of ET on cognitive performance in younger and older postmenopausal women have recently been completed. Preliminary results of both the Kronos Early Estrogen Prevention Study (KEEPS) and the Early versus Late Intervention Trial with Estradiol (ELITE) showed that the effects of ET on measures of cognitive performance (including verbal and working memory) did not differ from those of placebo, nor did young postmenopausal women show more beneficial effects of ET on cognitive performance than older women. Thus, neither the therapeutic nor the prophylactic benefits of ET on cognition in postmenopausal women are supported by currently available data. Further, beliefs about other indications for use of estrogen in the treatment of perimenopausal depression have been complicated by the HERS (Heart and Estrogen/Progestin Replacement Study) and WHI reports, which contradict the long-held view that estrogen protected against the development of heart disease in women. Since the cardio-protective effects previously reported were so substantial as to offset the increased risk of breast cancer after >4 years use of estradiol, particularly given the efficacy of estradiol in the prevention of osteoporosis, many physicians recommended HT as prophylaxis against future heart and bone diseases. With the demonstration that, to the contrary, estrogen combined with the progestin medroxy-progesterone may increase the risk of coronary

heart disease and stroke (during the first year of treatment), estradiol prophylaxis can no longer be recommended. However, estradiol treatment may still be indicated in women with perimenopausal symptoms (hot flushes, vaginal dryness) and depression, particularly in those who fail standard antidepressant treatment, who have a strong risk of osteoporosis, or who are intolerant of other acute treatments of perimenopausal symptoms. Indeed recent analyses of the WHI demonstrate that the adverse effects of ET on cardiovascular outcomes were largely confined to older women (and those at a greater distance from the menopause) compared with younger perimenopausal women, thus supporting the critical window or "gap" effect for the effects of estradiol on the cardiovasculature and providing some reassurance regarding the health effects of ET during the perimenopause.

Nausea is the most common side effect of ET administration. At conventional ET doses, this complaint seldom interferes with eating, and no weight loss has been reported. Breast engorgement, endometrial hyperplasia and bleeding are also common side effects of ET, which may enlarge pre-existing fibroid tumors of the uterus.

The relationship between estrogen, both endogenous and exogenous, and the development of endometrial carcinoma has been suggested by several different lines of investigation. Numerous retrospective case control studies published since 1975 have indicated that postmenopausal exposure to unopposed estrogens for more than 1 year results in a 2- to 12-fold increased relative risk for endometrial cancer. However, the addition of progesterone to ET decreases the risk of endometrial hyperplasia and endometrial cancer to equal or below that of women receiving no hormonal treatment. The optimal regimen to prevent hyperplasia and thus, inferentially, the risk of carcinoma, consists of 12 to 13 days of progestin treatment each month when estrogens are administered. There is an increased incidence of thromboembolism and stroke in women receiving ET and some SERMs; however, stroke is less likely in younger-aged perimenopausal women. Blood pressure, on average, appears to be unaffected by ET, although both increases and decreases have been reported. High doses of oral estrogens have been reported to elevate hepatocellular enzyme levels and, less commonly, cause cholestatic jaundice. The risk for gall stones and hepatocellular adenomas has been reported to be increased in association with oral contraceptive use, and although uncommon these complications may also occur with the use of more physiologic doses of estrogen. Most studies have suggested an increased relative risk of breast cancer after 4 or 5 years of use, similar to the risk expected if the onset of menopause is delayed for a comparable length of time; recent analyses, however, suggest that the increased risk is related to current use and the risk normalizes after cessation of HT. Additionally, it is recognized that ET (estrogen alone) is associated with

considerably less risk of breast cancer than combined estrogen and progestin formulations. Finally, the risk for breast cancer using hormone therapy for 5 years is associated with an increased "absolute risk" of only 0.67 percent, which some authors compare to the increased risk conferred by obesity and many other common exposures. For a more detailed review of the benefits and risks of menopausal ET, the reader is referred to recent updates from the Endocrine Society and the North American Menopause Society.

Estrogens may precipitate migraine headaches, and depression has also been reported to occur with the use of estrogens, but these reports are usually in the context of concurrent progestin use as in the oral contraception pill. If estradiol is to be used in the treatment of depression (see Table 34.39–1), then one should consider using the 50-microgram estradiol patch (transdermal estradiol [TDE]). The patch delivers estradiol, the form of estrogen that mediates most of its physiologic effects and does not cause an undesirable increase in the plasma proteins that mediate some of the adverse effects of estrogen. Levels achieved with this method are equivalent to levels in the early follicular phase and are slightly above the threshold necessary to preserve bone (60 pg/mL). In women without a uterus, estradiol may be given by itself and does not appear to have the risks associated with the combination of estradiol and progestin. Although ET does increase the risk of venous thromboembolism (VTE) and stroke, TDE and low-dose estrogen are not accompanied by an increased risk of VTE. If a progestin is needed, micronized progesterone tablets may be used, which deliver progesterone rather than medroxy-progesterone. Of note, progestins may differ markedly in their molecular effects, physiologic effects, and side effects. For example, second-generation progestins (e.g., norethisterone or levonorgestrel) are more androgenic than third- (e.g., desogestrel) or fourth-generation (e.g., drospirenone) progestins but, as a consequence, are more able to oppose the adverse (inducing) effects of estradiol on the liver and are less likely to result in VTE. Finally, individuals can show 20-fold differences in the blood level achieved with the same form of estrogen, so one must be prepared to change doses or preparations if necessary after soliciting reports from the patient about side effects, adverse effects, or lack of efficacy.

Table 34.39–1.
The Following Conditions Constitute Risk Factors or Contraindications* to Initiating Treatment with Estrogen

Women with a history of carcinoma of the breast*, or women with a family history of the following: premenopausal breast cancer or bilateral breast cancer in a first-degree relative; multiple family members (greater than three relatives) with postmenopausal breast cancer
Women with a history of uterine cancer*, endometriosis, ill-defined pelvic lesions*, particularly undiagnosed ovarian enlargement, undiagnosed vaginal bleeding*

(Active thrombophlebitis or thromboembolic disorders, or history of ischemic cardiac disease, pulmonary embolism, retinal thrombosis, or thrombophlebitis)*; (*risk factors for thromboembolic phenomena include cigarette smoking, varicose veins, prolonged periods of immobilization [including prolonged travel], and active heart disease*).

Known or suspected pregnancy*: Estrogens may cause fetal harm when administered to a pregnant women

Known hypersensitivity to estrogen or progestin (if prescribed) or the excipients (inactive compounds) contained within these medications. For example: Provera–calcium stearate, corn starch, lactose, mineral oil, sorbic acid, sucrose, talc; Alora–sorbitan monooleate, acrylic adhesive

History of cerebrovascular disease (stroke), epilepsy, hypertension, hypercalcemia

Cholecystitis or pancreatitis

Recurrent migraine headaches

Diabetes mellitus

History of familial hyperlipoproteinemia

Porphyria

Despite current concerns regarding ET and the requirement of adjunctive progestins for endometrial protection in women with an intact uterus, the specificity of actions of some of the newer selective estrogen receptor modulators (SERMS) and selective androgen receptor modulators (SARMS) (and hence the promise of separating therapeutic from toxic effects) suggests that estrogen receptor activators may occupy a useful niche in the future psychopharmacologic armamentarium. The recent introduction of tissue-selective estrogen complexes (TSECs), which combine estrogen with a SERM (some of which may be unable to enter the CNS), also holds promise for the delivery of individualized ET for those women who choose to employ hormonal therapies. Indeed, TSECs may have advantages over several SERMs (i.e., raloxifene and tamoxifen) since recent trials have documented that some TSECs (e.g., conjugated estrogen + bazedoxifene) are associated with a lower risk of VTE. These newer therapies could avoid the need for the addition of progestins (and the potential associated adverse effects) and could lead to therapies designed to optimize estrogen receptor activity. Further, there is emerging data that estradiol may be useful in the prevention or treatment of postpartum depression, a common psychiatric illness seen following approximately 10 percent of pregnancies.

Postpartum Depression. Despite the absence of endocrine abnormalities in this disorder, women with PPD show an enduring abnormal mood response to reproductive hormones. A previous study demonstrated that exposure of euthymic women with a history of postpartum depression to high doses of estradiol and progesterone followed by abrupt hormone withdrawal produced depression not seen in women lacking a history of postpartum depression. In parallel, supplemental reproductive steroids have been administered during the immediate postpartum in an attempt to prevent or treat postpartum

affective disorders. An open trial in women at risk for puerperal psychosis demonstrated that high-dose estrogen treatment resulted in a lower than expected 1-year recurrence rate (9 percent compared with an expected 35 to 60 percent without prophylaxis). Varying doses of estrogen (Premarin ranging in dose from 0.625 to 10 mg/day or intravenous estradiol 25 mg q8h) were administered immediately postpartum and then tapered over 4 weeks. It was suggested that estrogen administration could attenuate the rapid puerperal drop in estradiol levels, thereby reducing the negative impact of the postpartum "estrogen withdrawal state" on mood. In a follow-up study, investigators tested the suggestion that estradiol withdrawal caused postpartum depression in a double-blind, placebo-controlled study of estradiol in 61 women who developed major depression within 3 months of delivery. Eighty percent of the patients receiving estradiol (estrogen patch) experienced a significant reduction in depression severity after 3 months of treatment, compared with 31 percent of the placebo-treated group. Decreased mood symptoms on ET were observed in women regardless of concurrent antidepressant use, and estrogen's antidepressant effects were rapid and observed after 2 to 3 weeks of treatment. A similar rapid response to estradiol was also recently reported in an open label trial of sublingual estradiol, similar to the timing of the response to estradiol in perimenopausal depression. Double-blind, placebo-controlled trials of estradiol in the treatment of postpartum depression are ongoing and will help define the role of estradiol monotherapy in PPD.

Androgen Administration in Midlife Depression

Several investigations have suggested that the change in mood, cognition, and libido during the perimenopause/postmenopause may reflect deficiencies in androgens rather than (or in addition to) estrogens. Indeed, midlife in both men and women is characterized by a steady decline in the production of androgens such as dehydroepiandrosterone (DHEA) and androstenedione, which are mostly of adrenal origin yet contribute up to 50 percent to circulating androgen levels (testosterone, dihydrotestosterone) in women. Nonetheless, the definition of low testosterone levels is severely compromised by the current inability to accurately measure peripheral levels of testosterone in women or hypogonadal men (similar to the methodologic obstacles for measuring low estradiol levels). These assay limitations preclude the clear identification of men or women in whom androgen secretion is deficient (and in whom testosterone therapy would be indicated). Indeed, recent clinical practice guidelines recommend against making a diagnosis of androgen deficiency in women because of both the lack of a well-defined syndrome and the unavailability of reliable data correlating peripheral androgen levels with specific signs or symptoms. Moreover, the practice guidelines discourage the prescription of

either DHEA or testosterone therapies in women because there are only limited data supporting therapeutic benefits for the targeted symptoms and there are no long-term studies examining the risks of androgen therapy. That said, androgen therapy clearly benefits both mood and sexual function/libido in males who are clinically hypogonadal. Further, several investigators and a recent meta-analysis have shown that androgens improve a variety of elements of sexual function (desire, arousal, fantasies) when administered as sole treatment (ovariectomized women) or in combination with estradiol (postmenopausal women). Medical literature dating back to the isolation of testosterone in 1935 is replete with reports of the antidepressant efficacy of testosterone, androgen/anabolic steroids, and adrenal (and ovarian) androgens like DHEA. While most of these reports were uncontrolled, one study in ovariectomized women again demonstrated the superiority of testosterone to placebo in the mitigation of depressive symptoms. If androgens can improve mood, are they effective only in the context of clinically demonstrable androgen deficiency, the demonstration of which is severely limited, as noted above, by the inability to accurately measure low androgen levels? Further, even if shown to "improve well-being" (as was DHEA in older men and women), is androgen treatment effective in clinically significant affective disorders? A study that attempted to answer these questions entailed administration of the prevalent but weak androgen DHEA in a double-blind, placebo-controlled fashion to men and women with midlife-onset major and minor depression. Administration of 90 mg of DHEA for 3 weeks significantly decreased symptoms of depression compared to baseline and was significantly superior to effects seen with placebo. About 50 percent of subjects showed a beneficial effect, which was predicted neither by baseline DHEA (or its metabolites, testosterone and estradiol) nor by the change in hormonal measures following treatment. Nonetheless, this trial was small and DHEA was administered for only 6 weeks. Indeed, a recent meta-analysis concluded that in women with normal adrenal function the administration of DHEA is not indicated for complaints of either depressed mood or decreased sexual interest or arousal. This meta-analysis, however, does not answer the question of DHEA's efficacy in clinically significant depression since the majority of the reviewed studies were not sufficiently powered to evaluate DHEA's effects in depression. As should, by now, be clear, this, like other forms of hormone therapy, has attendant nonuniform benefits and risks. DHEA, for example, may improve bone mineral density with chronic administration, thus helping to prevent age-related osteopenia (although with the exception of small improvements in lumbar spine density in women, a recent meta-analysis found little evidence that DHEA had significant effects on bone density in women). It has been suggested, however, that DHEA administration, while not increasing total plasma testosterone levels in men, might be concentrated in the prostate, where

local conversion to testosterone could accelerate the growth of androgen-dependent tumors.

FUTURE DIRECTIONS

There is little question that reproductive hormones can regulate mood and that manipulations of reproductive hormones can, in some instances, dramatically alter the course and expression of affective illness. It is equally clear that these hormonal effects are not universal and appear in subgroups of individuals who appear differentially sensitive to these central actions of reproductive steroids. A major task for the future will be to identify the mediators of this differential sensitivity. Identification of the genetic or other predictors of vulnerability to reproductive endocrine–related mood disorders would assist the prediction of those who would respond to reproductive therapies and those who might be at risk for adverse effects. Indeed, the main constraint on the effective incorporation of reproductive hormonal modulators into the psychotherapeutic armamentarium as primary psychotropic agents or adjunctive therapies (e.g., in those who experience incomplete responses to other psychotropic agents) is the inability to predict the likelihood of adverse medical outcomes (e.g., breast cancer). Additionally, it is now clear that direct assays of sex steroids in populations with low levels of sex steroids (i.e., postmenopausal women, elderly men, and young children) are no longer acceptable evidence to confirm endocrine "deficiency" states, although even if accurate, single plasma hormone levels will not detect the pharmacodynamic differences across individuals that may result in marked phenotypic variation. The specificity and complexity of the actions of individual hormonal agents preclude generalization about the likely therapeutic or adverse outcome of exposure to a given hormone. Slight structural differences between the hormonal agents available in a given class can translate into dramatic differences in cellular and physiologic effects (and hence efficacy and risk), as can the timing of administration of these agents (i.e., their continuous or discontinuous administration). While generating current uncertainty, the specificity of actions of individual hormonal preparations, steroidogenic enzyme inhibitors, and steroid receptors antagonists creates extraordinary opportunity for the development of tissue-specific pharmacologic agents. The SERMs already available, for example, not only possess the beneficial effects of estradiol in some tissues (e.g., bone) but additionally block the potentially harmful effects of estradiol in other tissues (e.g., breast). Current knowledge, then, of tissue-specific coregulators will increasingly permit the design of reproductive agents that select and restrict the tissues affected and the actions elicited. In this fashion, the goal of psychopharmacologic precision will be greatly advanced.

REFERENCES

Bassuk SS, Manson JE. The timing hypothesis: do coronary risks of menopausal hormone therapy vary by age or time since menopause onset? *Metabolism*. 2016;65(5):794–803.

Bloch M, Daly RC, Rubinow DR. Endocrine factors in the etiology of postpartum depression. *Compr Psychiatry*. 2003;44(3):234–246.

Clarkson TB. Estrogen effects on arteries vary with stage of reproductive life and extent of subclinical atherosclerosis progression. *Menopause*. 2007;14(3):373–384.

Davis SR, Worsley R, Miller KK, Parish SJ, Santoro N. Androgens and female sexual function and dysfunction - findings from the fourth international consultation of sexual medicine. *J Sex Med*. 2016;13(2):168–178.

Dreher JC, Schmidt PJ, Kohn P, Furman D, Rubinow D, Berman KF. Menstrual cycle phase modulates reward-related neural function in women. *Proc Natl Acad Sci USA*. 2007;104(7):2465–2470.

*Freeman EW, Sammel MD, Lin H, Nelson DB. Associations of hormones and menopausal status with depressed mood in women with no history of depression. *Arch Gen Psychiatry*. 2006;63(4):375–382.

Lonard DM, O'Malley BW. Nuclear Receptor coregulators: modulators of pathology and therapeutic targets. *Nat Rev Endocrinol*. 2012;8(10):598–604.

Miller WL, Auchus RJ. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. *Endocr Rev*. 2011;32(1):81–151.

Nilsson S, Koehler KF, Gustafsson JA. Development of subtype-selective oestrogen receptor-based therapeutics. *Nat Rev Drug Discov*. 2011;10:778.

North American Menopause Society. The 2012 hormone therapy position statement of the North American Menopause Society. *Menopause*. 2012;19(3):257–271.

*Remage-Healey L. Frank Beach Award Winner: Steroids as neuromodulators of brain circuits and behavior. *Horm Behav*. 2014;66(3):552–560.

Rosner W, Hankinson SE, Sluss PM, Vesper HW, Wierman ME. Challenges to the measurement of estradiol: an endocrine society position statement. *J Clin Endocr Metab*. 2013;98(4):1376–1387.

Rossouw JE, Prentice RL, Manson JE, et al. Postmenopausal hormone therapy and risk of cardiovascular disease by age and years since menopause. *JAMA*. 2007;297(13):1465–1477.

Rubinow DR, Schmidt PJ. Androgens, brain and behavior. *Am J Psychiatry*. 1996;153(8):974–984.

*Rubinow DR, Schmidt PJ, Meltzer-Brody S, Harsh V. Gonadal hormones and behavior in women: concentration vs. context. In: Pfaff D ed. *Hormones, Brain and Behavior*. San Diego, CA: Academic Press; 2007.

Rubinow DR, Schmidt PJ, Meltzer-Brody S, Harsh V. Neurobiology of menstrual cycle-related mood disorders. In: Charney DS, Nestler EJ, Bunney BS, eds. *Neurobiologic Foundations of Mental Illness*. New York: Oxford University Press; 2007.

Rubinow DR, Johnson SL, Schmidt PJ, Girdler S, Gaynes B. Efficacy of estradiol in perimenopausal depression: so much promise and so few answers. *Depress Anxiety*. 2015;32(8):539–549.

Santen RJ, Kagan R, Altomare CJ, Komm B, Mirkin S, Taylor HS. Current and evolving approaches to individualizing estrogen receptor-based therapy for menopausal women. *J Clin*

Endocrinol Metab. 2014;99(3):733–747.

Schiller CE, Schmidt PJ, Rubinow DR. Allopregnanolone as a mediator of affective switching in reproductive mood disorders. *Psychopharmacology (Berl)*. 2014;231(17):3557–3567.

*Schmidt PJ, Ben Dor R, Martinez PE, et al. Effects of estradiol withdrawal on mood in women with past perimenopausal depression: a randomized clinical trial. *JAMA Psychiatry*. 2015;72(7):714–726.

Schmidt PJ, Daly RC, Bloch M, et al. Dehydroepiandrosterone monotherapy in midlife-onset major and minor depression. *Arch Gen Psychiatry*. 2005;62(2):154–162.

Schmidt PJ, Nieman L, Danaceau MA, et al. Estrogen replacement in perimenopause-related depression: a preliminary report. *Am J Obstet Gynecol*. 2000;183(2):414–420.

Schwartz TB. Henry Harrower and the turbulent beginnings of endocrinology. *Ann Intern Med*. 1999;131(9):702–706.

Soares CD, Almeida OP, Joffe H, Cohen LS. Efficacy of estradiol for the treatment of depressive disorders in perimenopausal women: a double-blind, randomized, placebo-controlled trial. *Arch Gen Psychiatry*. 2001;58(6):529–534.

Spearow JL, Doemeny P, Sera R, Leffler R, Barkley M. Genetic variation in susceptibility to endocrine disruption by estrogen in mice. *Science*. 1999;285(5431):1259–1261.

Taylor HS, Manson JE. Update in hormone therapy use in menopause. *J Clin Endocrinol Metab*. 2011;96(2):255–264.

Werner AA, Johns GA, Hoctor EF, Ault CC, Kohler LW, Weis MW. Involutional melancholia: probable etiology and treatment. *JAMA*. 1934;103:13–16.

Wierman ME, Arlt W, Basson R, et al. Androgen therapy in women: a reappraisal: an endocrine society clinical practice guideline. *J Clin Endocrinol Metab*. 2014;99(10):3489–3510.

Wisner KL, Sit DK, Moses-Kolko EL, et al. Transdermal estradiol treatment for postpartum depression: a pilot, randomized trial. *J Clin Psychopharmacol*. 2015;35(4):389–395.

*Writing Group for the Women's Health Initiative Investigators. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288(3):321–333.

▲ 34.40 Drug Treatment of Erectile Disorder

CARIDAD PONCE MARTINEZ, M.D., FADI HADDAD, M.D., AND ANITA H. CLAYTON, M.D.

INTRODUCTION

Erectile dysfunction (ED), also referred to as impotence, is defined as the inability to obtain or maintain an erection sufficient for sexual activity. The DSM-5 lists ED as erectile disorder (Table 34.40–1), while the 10th revision of *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) codes it as failure of genital response. A man with lifelong erectile disorder is not able to obtain an adequate erection since becoming sexually active after puberty. In acquired erectile disorder, a man

has disturbance of erectile function after a period of relatively normal sexual function. The disorder can be situational, in that it is specific to certain types of stimulation, situations, or partners, or it can be generalized. Severity of the disorder can be mild, moderate, or severe, depending on the level of distress the symptoms cause in the individual.

Table 34.40–1.
DSM-5 Diagnostic Criteria for Erectile Disorder

A. At least one of the three following symptoms must be experienced on almost all or all (approximately 75–100%) occasions of sexual activity (in identified situational contexts or, if generalized, in all contexts):
   1. Marked difficulty in obtaining an erection during sexual activity.
   2. Marked difficulty in maintaining an erection until the completion of sexual activity.
   3. Marked decrease in erectile rigidity.
B. The symptoms in Criterion A have persisted for a minimum duration of approximately 6 months.
C. The symptoms in Criterion A cause clinically significant distress in the individual.
D. The sexual dysfunction is not better explained by a nonsexual mental disorder or as a consequence of severe relationship distress or other significant stressors and is not attributable to the effects of a substance/medication or another medical condition.

Specify type: Lifelong or acquired type
Specify type: Generalized or situational type
Specify severity: Mild, moderate, or severe

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

The diagnostic criterion of duration of at least 6 months is new to the DSM-5, although no systematic research has validated this criterion. In fact, the International Society for Sexual Medicine's 2010 summary of recommendations on sexual dysfunctions in men establishes that a 3-month duration of symptoms is sufficient to make a diagnosis. In cases of trauma or surgically induced ED, the diagnosis can be made even sooner.

Erectile disorder is the most common type of sexual dysfunction in men, with an estimated prevalence in the US male population aged 20 years or older of 18.4 percent, equivalent to more than 18 million men. In the Massachusetts Male Aging Study, which surveyed approximately 1,700 men longitudinally, the prevalence of all ED was 39 percent in 40-year-old men and 67 percent in men aged 70 or older. After adjusting for age, the strongest correlates for ED were markers of cardiovascular disease—hypertension, diabetes, and heart disease. It is important to note that ED does not occur in all aging men. Instead, good health and the availability of a sexual partner are often stronger factors in determining satisfactory sexual function.

Normal erectile function involves psychological and biological processes. In general terms, erectile disorder can be classified as

psychogenic, organic, or mixed. Psychogenic factors such as performance anxiety, traumatic experiences, relationship problems, moral prohibitions, and psychosocial stressors can negatively impact erectile function. This type of erectile disorder tends to occur on a situational basis, has a sudden onset, and concurrently, men can experience normal erections with masturbation, at night, or upon awakening. Organic processes can be vascular (e.g., arterial insufficiency, venous leakage), neurological, endocrine (e.g., hypogonadism, hyperprolactinemia), substance induced, or due to a systemic illness. In these cases, ED tends to be more insidious in its onset and erections are deficient in most contexts (masturbation, nocturnal).

A vasculogenic etiology is the most common among organic etiologies of erectile disorder. Current evidence suggests a strong correlation of cardiovascular disease and ED, with a common pathophysiological basis of endothelial dysfunction, inflammation, and low testosterone. ED is also associated with coronary heart disease, stroke, and all-cause mortality. The presence of vasculogenic ED in young men can be a particularly early indicator of future cardiovascular disease, and these patients should receive an appropriate cardiovascular workup for risk stratification.

Most cases of ED have both biological and psychogenic features. Even in cases of clear organic etiology like perineal trauma, impaired sexual function can result in psychological consequences such as low self-esteem, low self-confidence, and even anxiety or depressed affect, which can further exacerbate the dysfunction. Involvement of sexual partners during evaluation and treatment is important in examining their expectations for diagnosis and treatment.

Risk factors for ED can be examined in accordance with etiology. Some neurological disorders (e.g., multiple sclerosis, stroke, Alzheimer disease) are associated with ED, and the risk of nerve injury leading to ED is very high in patients undergoing radical pelvic surgery. Endocrine risk factors include hypogonadism, hyperprolactinemia, and hypothyroidism. Some risk factors for vasculogenic ED are atherosclerosis, hypertension, hyperlipidemia, obesity, smoking, diabetes mellitus (DM), and pelvic irradiation. Among men with DM, 50 to 75 percent develop ED; the prevalence of ED in diabetics is three times that seen in nondiabetics. Many substances can cause or contribute to ED. The most common medications are antiandrogens, antihypertensives (thiazide diuretics, beta blockers, calcium channel blockers), antiarrhythmics (e.g., digoxin), and many psychotropic drugs (tricyclic antidepressants, selective serotonin reuptake inhibitors [SSRIs], phenothiazines, and butyrophenones). Use of recreational, illegal, or controlled substances such as marijuana, alcohol, cocaine, and opiates may also increase risk of impaired erectile function.

A comprehensive diagnostic interview is an important feature in an evaluation for erectile disorder. Information about psychiatric history,

medical and surgical history, sexual history, family history, and psychosocial history including habits and lifestyle should be obtained. Validated self-administered questionnaires can be used to aid the diagnostic process, and the most well known of these is the International Index of Erectile Function (IIEF). A detailed sexual history provides information about the extent of the disorder (duration, context, severity) and also helps to differentiate erectile disorder from other types of sexual dysfunction, such as hypoactive sexual desire disorder or premature ejaculation. The goal of such an assessment is to determine the presence of causes or risk factors for erectile disorder.

In addition to the interview, an evaluation for ED includes a general physical examination with particular emphasis on the nervous, vascular, and genital systems. Laboratory studies are used to determine endocrine causes of ED and also to identify associated comorbid conditions like hypercholesterolemia or DM. The International Society for Sexual Medicine recommends ordering tests to measure fasting glucose, HbA1c, lipids, and total testosterone levels. Supplementary tests such as Doppler ultrasonography or nocturnal penile tumescence are often reserved for specialists, and do not generally provide information that would alter the treatment course.

TREATMENT

Lifestyle Changes

There is a clear association between cardiovascular disease and increased risk of ED, therefore management of cardiovascular risk factors may improve sexual function. The adoption of a Mediterranean diet, weight loss, exercise, and smoking cessation are all beneficial. Eliminating illicit substance use may also improve erectile function. In men with obstructive sleep apnea (OSA), treatment with CPAP is associated with some improvement in sexual function. Finally, gastric bypass surgery can increase testosterone levels and therefore improve sexual function in men with morbid obesity.

Phosphodiesterase-5 Inhibitors

Nitric oxide (NO) mediates tumescence by increasing the levels of cGMP, which in turn leads to smooth muscle relaxation in the corpus cavernosum that allows the penis to dilate and admit blood.

PDE-5 is an enzyme that catabolizes cGMP. Phosphodiesterase-5 (PDE-5) inhibitors competitively inhibit the PDE-5 enzyme and increase intracavernosal cGMP thereby leading to an increase in both the number and duration of erections in men with ED.

Sufficient psychological and environmental cues that lead to sexual arousal and stimulation are required for PDE-5 inhibitors to work.

An evaluation of the underlying cause of ED is necessary prior to initiation of PDE-5 inhibitor treatment. These agents are contraindicated in men treated with nitrates and are to be used cautiously in men treated with alpha-adrenergic blockers.

Treatment response seems to be largely dependent on proper use and administration (Table 34.40–2).

Sildenafil. The efficacy of sildenafil in treating ED has been demonstrated by numerous clinical trials. A meta-analysis of 27 trials in 6,659 men with ED showed a marked difference in the percentage of successful sexual intercourse between sildenafil and placebo (57 vs. 21 percent, respectively).

Sildenafil was also found effective for treatment of ED in the setting of various medical conditions including DM, prostate cancer treatment and parkinsonism.

Sildenafil could also have psychological benefits in men with ED as shown by improvement in the mean Self-Esteem subscale score of the Self-Esteem and Relationship (SEAR) questionnaire (a psychometric instrument specific to ED).

Sildenafil should be taken 1 hour prior to a planned sexual encounter. The dose ranges between 25 and 100 mg with 50 mg being the usual dose. Lower dosing is recommended for use in geriatric men (older than 65 years), liver disease, and renal impairment. The duration of action is on average 4 hours.

Vardenafil. Vardenafil shares similar structure and pharmacokinetic profile with sildenafil. It is also efficacious in ED secondary to DM and nerve-sparing prostatectomy. An advantage is its availability in an oral disintegrating form with a more rapid onset of action. The dose range is 10 to 20 mg. Nothing is known about potential transfer of the product through oral or orogenital contact after use of the orally disintegrating formulation.

Tadalafil. Tadalafil is structurally distinct from sildenafil and vardenafil. There are no head-to-head comparisons with other PDE-5 inhibitors but it appears to be effective in the treatment of ED, with a longer duration of action.

Table 34.40–2.
Phosphodiesterase Type 5 (PDE-5) Inhibitors

Medication	Clinical Use	Dosing	Drug Interactions
Sildenafil	Use as needed for treatment of organic	Taken 1 hour before sex and effective for up to 4 hours Dose: 50 to 100 mg 25 mg is recommended in geriatric	Strong inhibitors of CYP3A4 (erythromycin, ketoconazole, itraconazole, grapefruit and its juice) elevate sildenafil

	or psychogenic ED	patients, patients with renal impairment, hepatic failure, use of potent CYP3A4 inhibitors and alpha-blocker therapy	concentrations
Vardenafil	Same as sildenafil	Taken 1 hour before sex and effective up to 4 hours Dose 10 to 20 mg 5 mg is recommended in geriatric patients, patients with renal impairment, hepatic failure, use of potent CYP3A4 inhibitors, and alpha-blocker therapy Oral disintegrating tablet (ODT) dose is 10 mg	Same as sildenafil
Tadalafil	Same as sildenafil	Duration of action is up to 36 hours Dosing: 10 to 20 mg as needed before sex **or** Daily treatment 2.5 to 5 mg once daily Dose adjustment is needed for renal or hepatic impairment	Same as sildenafil
Avanafil	Same as sildenafil	Onset of action is 15 to 30 minutes Dosing: 100 to 200 mg taken at least as 15 minutes before sexual activity 50 mg dose should be taken 30 minutes before sexual activity	Avanafil should be avoided with CYP3A4 inhibitors Grapefruit and grapefruit juice should be avoided within 24 hours of use

Avanafil. Avanafil is a new agent, currently approved in the United States and Europe for treatment of ED. Compared to other agents of the same class, it is more selective to PDE-5 and has a more rapid onset of action. Dosing ranges between 50 and 200 mg. It should be administered 30 minutes prior to sexual activity at 50 mg, while the 100 and 200 mg can be taken 15 minutes in advance.

ADVERSE EFFECTS AND PRECAUTIONS. All PDE-5 inhibitors share a similar side effect profile which is related to their vasodilatory properties. Common side effects include flushing, headache, and dyspepsia.

Sildenafil can cause “blue vision” in approximately 3 percent of men. This is caused by a cross-reactivity of the drug with the PDE-6 inhibitor present in the retina. Blue vision is transient and disappears spontaneously. This side effect has not been reported with other PDE-5 inhibitors.

Rare cases of nonarteritic anterior ischemic optic neuropathy were reported in men taking sildenafil and tadalafil.

Rare reports have indicated a risk of sudden unilateral hearing loss with sildenafil, vardenafil, and tadalafil. This side effect occurred within the first 24 hours of drug administration and is temporary in approximately one-third of the patients. No direct causal relationship has been demonstrated but the FDA requires labeling of all PDE-5 inhibitors to include this risk. It

would be recommended for providers to warn patients about this risk during the first 24 hours of treatment.

DRUG INTERACTIONS. All PDE-5 inhibitors are contraindicated in men on nitrates. Concurrent use with alpha-adrenergic antagonists causes symptomatic hypotension. Drugs that block the CYP3A4 metabolic pathway can prolong the half-life of sildenafil and thus should be avoided.

CHOICE OF DRUG. Current data indicates that all PDE-5 inhibitors are of equivalent efficacy and side-effect profile, therefore choice of agent is based on patient preference, cost, and desired duration of action.

CONCERN OVER RECREATIONAL USE. Sildenafil use leads to a marked reduction in the postejaculatory refractory time, thus, men may be capable of having a second erection in a shorter time frame after ingestion. This leads to recreational use of sildenafil. There are case reports of stroke in men taking high-normal doses (100 mg) of sildenafil.

Supplements Containing PDE-5 Inhibitors. One-third to one-half of over the counter natural supplements marketed to enhance sexual performance, contain prescription strength doses of sildenafil or tadalafil. The FDA has issued a warning against use of the above supplements for concern over severe hypotension in men on nitrates.

Testosterone

Testosterone regulates sexual desire and modulates almost every component involved in erectile function. There is significant overlap between hypogonadism and ED. The diagnosis of testosterone deficiency is made on the basis of the clinical symptoms of hypogonadism and consistently subnormal serum testosterone concentrations (below 8 nmol/L or 231 ng/dL) at 8 to 10 AM on three occasions.

Testosterone improves libido in men with low testosterone levels, but does not necessarily improve ED. As adequate testosterone levels are necessary for NO function, the addition of testosterone to PDE-5 inhibitors may improve erectile function in hypogonadal men who did not respond to initial treatment with PDE-5 inhibitors.

Testosterone is available in different formulations and preparations including oral tablets, transdermal gel/solution or patch, buccal application, subcutaneous implants, and intramuscular injections. Dosing varies depending on the formulation. Dosing is gradually titrated to restore the serum testosterone concentration to a normal range of 13.9 to 27.0 nmol/L or 400 to 700 ng/dL. The timing of serum testosterone measurement varies depending on the preparation used. Monitoring serum testosterone should be done 2 to 3 months after treatment initiation and

after changing the dose. Monitoring is then extended to every 6 to 12 months once the dose is stabilized.

Testosterone is contraindicated in men with a history of prostate cancer.

Common side effects of testosterone therapy include exacerbation of benign prostate hyperplasia, worsening of OSA, erythrocytosis, venous thromboembolism, and skin irritation in patients using the transdermal formulation. Unintentional transfer of transdermal testosterone products through contact should be avoided.

Clinical Vignette 1

Mr. B. is a 55-year-old man with medical history of obesity, hypertension, and hyperlipidemia, and psychiatric history of OCD who had been treated with a high dose SSRI and cognitive behavioral therapy for many years. In the past year, he developed generalized ED, yet there had been no change in his medications or in his long-term relationship. Medical workup revealed abnormal lipid profile, elevated HbA1C and low testosterone. Transdermal testosterone treatment was initiated and his dose was titrated over 3 months to achieve normal serum testosterone. There was a minimal response in terms of his ED with normalized testosterone levels. Tadalafil 5 mg daily was added to his regimen with good response. He was referred to his endocrinologist for more intensive management of his metabolic issues.

Intracavernous Injections

Alprostadil is a synthetic formulation of the endogenous prostaglandin E1, injected locally into the corpora cavernosa. An initial dose of 5 to 10 μg is recommended in vasculogenic males, and a smaller dose in the case of neurogenic ED (2.5 μg). The first intracavernosal agent to be approved by the U.S. Food and Drug Administration (FDA) for treatment of ED, alprostadil, acts by increasing levels of cyclic AMP within the corporal smooth muscle, causing muscular relaxation. This reaction lowers the vascular resistance inside the corpora cavernosa, increasing blood flow to the penis, and resulting in an erection. The erection can occur even in the absence of sexual stimulation. Alprostadil is considered efficacious, relatively safe, and fast acting (2 to 3 minutes), but high patient discontinuation rates indicate some concerns regarding long-term tolerability. Ideal candidates for treatment with intracavernous therapy include those who fail or have a contraindication to first-line treatments, such as those concomitantly using nitrates; patients who desire a rapid onset of erection, and patients with insufficient improvement with oral therapy. Possible adverse events include penile pain, priapism (rare), and penile fibrosis and curvature (uncommon, related to chronic injections). Intracavernous injection of alprostadil is contraindicated in patients with conditions that predispose them to priapism (e.g., sickle cell anemia, multiple myeloma, leukemia), severe penile fibrosis, vision problems that limit safe self-administration of injections, and use of monoamine oxidase

inhibitors which would limit use of phenylephrine if priapism occurred.

Alprostadil is also available for intraurethral administration (medicated urethral system for erection, MUSE), and the literature reports success rates of 43 to 69 percent. This form of administration is less invasive and has superior tolerability, although there are many reports indicating lower efficacy than intracavernous injection. Disadvantages include systemic side effects (hypotension, syncope), pain, and cost. Other medications often used in combination with alprostadil include papaverine, phentolamine, and vasoactive intestinal polypeptide. The combination of prostaglandin E, phentolamine, and papaverine (Trimix) has been found to be particularly effective.

Treatment Options without FDA Approval

Yohimbine. Yohimbine has been used for many decades in the treatment of ED. It is an alpha-2 adrenoceptor antagonist with central and peripheral effects, although its site of action in producing an erection is probably not peripheral. Placebo-controlled trials on patients with different types of erectile disorder have shown at most, a modest effect. Due to conflicting evidence, yohimbine is currently not recommended in the management of erectile disorder.

Melanocortin Receptor Agonists. The melanocortin-4 receptor modulates erectile function and sexual behavior, as supported by evidence in studies of mice, rats, and primates. In humans, bremelanotide (not yet approved for any indication) has been studied in the treatment of male and female sexual dysfunction, including ED. By acting on receptors in the hypothalamus, bremelanotide stimulates peripheral production of NO in the corpora cavernosa of the penis, leading to an erection. Interestingly, this mechanism of action allows initiation of erections even in the absence of visual sexual stimulation, and minimal external stimulation, unlike other pharmacological treatments. Additional advantages include low incidence of side effects, rapid onset of action, short half-life, and because it does not cause systemic vasodilation, it could be used with nitrates. Some phase 2 studies have reported efficacy and safety of bremelanotide as monotherapy, and in combination with sildenafil in the treatment of male erectile disorder.

Apomorphine. Apomorphine is a nonselective dopamine agonist that acts within the central nervous system to enhance natural signals that occur during sexual stimulation. In a sublingual preparation, it has been approved for use in the treatment of erectile disorder in Europe, but not the United States. Its rapid onset after administration (71 percent erections achieved within 20 minutes) allows for utilization on demand, usually at doses of 2 to 3 mg. Side effects may include mild to moderate nausea,

headache, dizziness, and in rare cases, vasovagal syncope. Apomorphine can be used with nitrates. In 28.5 to 55.0 percent of the cases, men are able to achieve an erection sufficient for intercourse. However, studies comparing its efficacy with that of sildenafil demonstrate significantly lower efficacy and patient satisfaction rates. Therefore, current use is limited to patients with mild/moderate or psychogenic erectile disorder.

Mechanical and Surgical Treatment

Vacuum constriction devices apply continuous negative pressure to the penis, drawing blood into the corpora cavernosa, which is then retained by an elastic band applied at the base of the penis. Efficacy can be as high as 90 percent regardless of the etiology of erectile disorder. Associated side effects include pain, numbness, bruising, and delayed ejaculation. Advantages include its use on demand and low cost of the device. One relative contraindication is anticoagulation therapy. This form of treatment can be beneficial for patients who have failed oral PDE-5 treatment and refuse intracavernosal injections.

Vascular surgery, including arterial bypass and venous ligation surgery, can improve vasculogenic erectile disorder. Given that such procedures target the etiology of the erectile disorder, in appropriate candidates, the surgery can be curative.

Surgical implantation of an inflatable or malleable (semi-rigid) penile prosthesis is an invasive and irreversible form of treatment, generally reserved for patients who have failed other types of treatment (third-line treatment). Most men prefer implantation of the inflatable three-piece device, as it provides a more “natural” erection. If successful, the rates of satisfaction are high, up to 70 percent. Adverse events include prosthesis infection (2 to 4 percent of cases), penile fibrosis, and mechanical failures. Mechanical failures can be as high as 50 percent by 10 years.

PSYCHOTROPIC-INDUCED ERECTILE DISORDER AND ITS MANAGEMENT

The most studied psychotropics with regards to impact on sexual function are antidepressants, given serotonin’s role in sexual function. SSRIs and serotonin-norepinephrine reuptake inhibitors (SNRIs) can negatively affect all stages of the male sexual response cycle. Nefazodone and mirtazapine have lower rates of sexual dysfunction than SSRIs. Bupropion, a dopamine and norepinephrine reuptake inhibitor; and buspirone, an agonist of 5-HT1A receptors, are not associated with sexual dysfunction and can be used as augmenting agents to treat antidepressant-induced sexual dysfunction, or as alternative agents to treat depression in men with ED. There is evidence supporting the use of PDE-5 inhibitors in men for treatment of ED induced by antidepressant medications.

Unlike antidepressants, the data on antipsychotic-induced sexual

dysfunction is scarce. Antipsychotics cause ED via dopaminergic antagonism within the tuberoinfundibular system, which causes elevation in plasma prolactin levels. Head-to-head comparisons are unavailable, but studies suggest that drugs that cause more hyperprolactinemia like risperidone and the conventional antipsychotics are more likely to cause ED, as compared to olanzapine, clozapine, quetiapine, and aripiprazole.

There is insufficient evidence regarding the treatment of antipsychotic-induced ED. Some authors have proposed dose reduction and switching to a prolactin-sparing second-generation antipsychotic as possible strategies. PDE-5 inhibitors may also be of benefit.

Clinical Vignette 2

Mr. J. is a 30-year-old man with no active medical problems and history of MDD whose initial treatment with a maximum dose of SSRI achieved a partial response. Risperidone 1 mg was added as an inexpensive augmenting agent resulting in a full remission of his depressive symptoms. Within a month he developed ED. No change in the relationship with his wife preceded the ED but he began to question his attractiveness to his wife and began avoiding sex. Despite his wife's supportive response to his ED, he developed anxiety associated with sexual activity and behavioral avoidance. He did not spontaneously report this side effect to his psychiatrist until questioned directly. The onset of ED seemed to correlate with the addition of risperidone to the high dose of SSRI. Interpersonal issues appeared to be the result of the ED rather than the cause. Sildenafil was added at 50 mg prior to sexual activity. At followup Mr. J. reported complete resolution of the ED and remained on the same medication regimen.

FUTURE DIRECTIONS

There are a variety of drugs currently under investigation for the treatment of ED. Many of these have novel therapeutic targets, including vasorelaxation (associated with the NO-cGMP pathway or independent from it), inhibition of vasoconstriction (e.g., Rho-kinase inhibitors), anti-inflammatory and anti-fibrotic therapies, regeneration products, and central nervous system targets.

Low-intensity extracorporeal shock wave (LI-ESW) therapy in the treatment of erectile disorder appears to function by recruiting endogenous stem cells and increasing expression of vascular endothelial growth factor, thereby stimulating angiogenesis. It is currently available for clinical use in countries like Israel and Canada, but not in the United States.

Additional avenues for treatment in the future include drug-eluting peripheral stent systems (Zotarolimus-Eluting Peripheral Stent System) in patients with internal pudendal artery stenosis, tissue engineering for the replacement or repair of corporal bodies, topical application of nanoparticles to achieve erection, and penile implants with drug delivery apparatus. Gene therapy uses transfected genetic material to correct molecular defects. The penis provides an ideal target tissue for gene therapy because of its external location and internal structure. Genes encoding vasoactive intestinal peptide, brain-derived neurotrophic factor,

and maxi-K (calcium-sensitive potassium) channels have been proposed targets, among others. To date, there has been a single phase I clinical trial in 21 men regarding the maxi-K channels, with encouraging results.

AUTHORS' RECOMMENDATIONS

The evaluation of male erectile disorder includes a thorough interview, with focus on medical, surgical, sexual, and psychosocial history. Vasculogenic ED is the most common type of organic erectile disorder, so a cardiovascular workup should be completed in patients with vascular risk factors.

- PDE-5 inhibitors should be used as first-line treatment for ED in men with normal testosterone levels.
- Testosterone replacement should be the first-line treatment in hypogonadal men with ED.
- In hypogonadal men who also have a concomitant vascular or metabolic disorder, testosterone treatment alone might be insufficient to restore erection; combination treatment with a PDE-5 inhibitor is recommended.
- Additional treatment options for men who do not respond to PDE-5 inhibitors (or have a contraindication to use) include intracavernous injections, intraurethral administration, and mechanical/surgical treatment.
- Augmentation with bupropion or buspirone may be an option in patients whose depression responded well to an SSRI/SNRI and are experiencing significant ED.
- In patients who did not fully respond to SSRI/SNRI treatment and have significant ED, switching to a medication with less ED risk (mirtazapine, buspirone, bupropion) may be the appropriate next step in treatment.
- The treatment for erectile disorder should be individualized, taking into account patient and partner's preferences, comorbid conditions, likely etiology of ED, and concomitant medications.

REFERENCES

*Andersson KE. Mechanisms of penile erection and basis for pharmacological treatment of erectile dysfunction. *Pharmacol Rev*. 2011;63(4):811–859.

Bella AJ, Brock GB. Intracavernous pharmacotherapy for erectile dysfunction. *Endocrine*. 2004;23(2–3):149–155.

Corona G, Rastrelli G, Burri A, Jannini EA, Maggi M. The safety and efficacy of Avanafil, a new 2nd generation PDE5i: comprehensive review and meta-analysis. *Expert Opin Drug Saf*. 2016;15(2):237–247.

Debruyne FM, Gittelman M, Sperling H, Börner M, Beneke M. Time to onset of action of vardenafil: a retrospective analysis of the pivotal trials for the orodispersible and film-coated

tablet formulations. *J Sex Med*. 2011;8(10):2912–2923.

Esposito K, Giugliano F, Di Palo C, et al. Effect of lifestyle changes on erectile dysfunction in obese men: a randomized controlled trial. *JAMA*. 2004;291(24):2978–2984.

Feldman HA, Goldstein I, Hatzichristou DG, Krane RJ, McKinlay JB. Impotence and its medical and psychosocial correlates: results of the Massachusetts Male Aging Study. *J Urol*. 1994;151(1):54–61.

Glina S, Cohen DJ, Vieira M. Diagnosis of erectile dysfunction. *Curr Opin Psychiatry*. 2014;27(6):394–399.

Goldstein I, Tseng LJ, Creanga D, Stecher V, Kaminetsky JC. Efficacy and safety of sildenafil by age in men with erectile dysfunction. *J Sex Med*. 2016;13(5):852–859.

Gratzke C, Angulo J, Chitaley K, et al. Anatomy, physiology, and pathophysiology of erectile dysfunction. *J Sex Med*. 2010;7:445–475.

Hatzichristou D, d'Anzeo G, Porst H, et al. Tadalafil 5 mg once daily for the treatment of erectile dysfunction during a 6-month observational study (EDATE): impact of patient characteristics and comorbidities. *BMC Urol*. 2015;15(1):111.

Hellstrom WJ. Clinical applications of centrally acting agents in male sexual dysfunction. *Int J Impot Res*. 2008;20(Suppl 1):S17–S23.

Kim KK, Kim DG, Ku YH, et al. Bilateral cerebral hemispheric infarction associated with sildenafil citrate (Viagra) use. *Eur J Neurol*. 2008;15(3):306–308.

Lue TF, Giuliano F, Montorsi F, et al. Summary of the recommendations on sexual dysfunctions in men. *J Sex Med*. 2004;1(1):6–23.

Maddox PT, Saunders J, Chandrasekhar SS. Sudden hearing loss from PDE-5 inhibitors: A possible cellular stress etiology. *Laryngoscope*. 2009;119(8):1586–1589.

Miner M, Nehra A, Jackson G, et al. All men with vasculogenic erectile dysfunction require a cardiovascular workup. *Am J Med*. 2014;127(3):174–182.

Moncada I, Martínez-Jabaloyas JM, Rodriguez-Vela L, et al. Emotional changes in men treated with sildenafil citrate for erectile dysfunction: a double-blind, placebo-controlled clinical trial. *J Sex Med*. 2009;6(12):3469–3477.

Montague DK, Jarow JP, Broderick GA, et al; Erectile Dysfunction Guideline Update Panel. Chapter 1: The management of erectile dysfunction: an AUA update. *J Urol*. 2005;174(1):230–239.

*Montorsi F, Adaikan G, Becher E, et al. Summary of the recommendations on sexual dysfunctions in men. *J Sex Med*. 2010;7(11):3572–3588.

Paduch DA, Polzer PK, Ni X, Basaria S. Testosterone replacement in androgen-deficient men with ejaculatory dysfunction: a randomized controlled trial. *J Clin Endocrinol Metab*. 2015;100:2956.

Sanford M. Vardenafil orodispersible tablet. *Drugs*. 2012;72:87–98.

Schover LR, Fouladi RT, Warneke CL, et al. The use of treatments for erectile dysfunction among survivors of prostate carcinoma. *Cancer*. 2002;95(11):2397–2407.

Selvin E, Burnett AL, Platz EA. Prevalence and risk factors for erectile dysfunction in the US. *Am J Med*. 2007;120(2):151–157.

Shabsigh R, Mattern A, REVITALISE Study Group. REVITALISE: a large observational study assessing the safety and effectiveness of vardenafil in men with erectile dysfunction and

metabolic syndrome. *Sex Med*. 2016;4(3):e135–144.

Shadiack AM, Sharma SD, Earle DC, Spana C, Hallam TJ. Melanocortins in the treatment of male and female sexual dysfunction. *Curr Top Med Chem*. 2007;7(11):1137–1144.

*Shamloul R, Ghanem H. Erectile dysfunction. *Lancet*. 2013;381(9861):153–165.

Stein MJ, Lin H, Wang R. New advances in erectile technology. *Ther Adv Urol*. 2014;6(1):15–24.

Tadalafil (Cialis) once a day for erectile dysfunction. *Med Lett Drug Ther*. 2008;50:27–28.

*Taylor MJ, Rudkin L, Bullemor-Day P, Lubin J, Chukwujekwu C, Hawton K. Strategies for managing sexual dysfunction induced by antidepressant medication. *Cochrane Database Syst Rev*. 2013;(5):CD003382.

*Tsertsvadze A, Fink HA, Yazdi F, et al. Oral phosphodiesterase-5 inhibitors and hormonal treatments for erectile dysfunction: a systematic review and meta-analysis. *Ann Intern Med*. 2009;151(9):650–651.

Wespes E, Amar E, Hatzichristou D, et al; EAU. EAU Guidelines on erectile dysfunction: an update. *Eur Urol*. 2006;49(5):806–815.

World Health Organization. *ICD-10 Version: 2010*. Chapter V, Section F50-F59. http://apps.who.int/classifications/icd10/browse/2010/en#. Accessed on December 11, 2016.

▲ 34.41 Vilazodone

LOURDES ILLA-SANCHEZ, M.D.

INTRODUCTION AND HISTORY

Vilazodone (VIIBRYD) is a novel dual-acting serotonergic antidepressant that combines the effect of an SSRI with 5HT1A-receptor partial agonist activity. It was approved by the FDA for the treatment of adults with major depressive disorder (MDD) on January 21, 2011, after nearly a decade of development. First developed by Merck and GlaxoSmithKline in 2001, development stopped in 2003 after several unsuccessful phase II trials. Development resumed in 2004 by Genaissance, later purchased by Clinical Data, Inc. A new drug application was submitted on March 22, 2010, and the FDA issued a new U.S. Patent on 11/16/10. Based on two successful phase III trials, vilazodone was approved by FDA in January of 2011. Clinical Data, Inc. was later purchased by Forest Laboratories, Inc. in February of 2011.

Chemistry

Vilazodone's chemical structure is 2-benzofurancarboxamide, 5-[4-[4-[5-cyano-1H-indol-3-yl]butyl]-1-piperazinyl]-, hydrochloride (1:1) (Fig. 34.41–1). Its molecular weight is 441.535. It was designed to have combined activity as a SSRI and an 5HT1A receptor agonist by combining indole-butyl-amine and chromenonyl-piperazine structural elements into a single molecular entity.

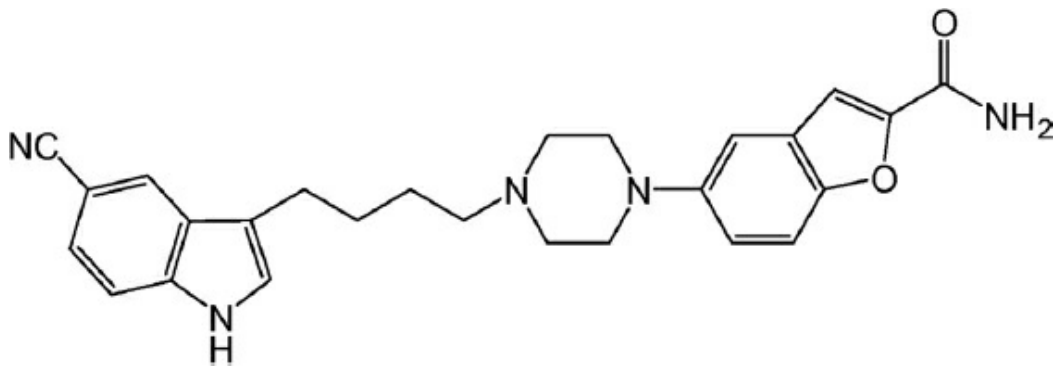

FIGURE 34.41–1. Chemical structure of vilazodone.

Pharmacological Actions

Pharmacokinetics. Vilazodone's activity is due primarily to the parent drug. Its pharmacokinetics are dose-proportional, and peak concentrations occur at 4 to 5 hours after administration. It has a terminal half-life of about 25 hours and achieves steady state in approximately 3 days. Vilazodone's absorption is largely dependent on food intake, with a reported bioavailability of 72 percent when administered with food. The drug is widely distributed and is 96 to 99 percent protein bound. Vilazodone is extensively metabolized, with only 1 and 2 percent recovered unchanged in urine and feces, respectively. It is primarily metabolized by CYP3A4, with minor contributions from CYP2C19 and CYP2D6.

Pharmacodynamics. Vilazodone exhibits a unique dual mechanism of action. Like SSRIs, it blocks the reuptake of serotonin, while simultaneously having partial agonist activity at the serotonin 5HT1A receptor. Due to its dual SSRI/partial 5HT1A mechanism of action, it is often referred to as a serotonin partial agonist/reuptake inhibitor (SPARI). In preclinical trials, vilazodone has been shown to bind with high affinity to the serotonin reuptake site, potently and selectively inhibit the reuptake of serotonin and bind selectively with high affinity to 5HT1A receptors. In animal models, adding 5HT1A partial agonism to SSRIs results in more rapid and robust elevations of 5HT levels than seen with SSRIs alone. While this mechanism has been thought to possibly result in rapid onset of antidepressant efficacy, this has not to date been confirmed in human clinical trials. Taken without food, vilazodone's blood concentrations can be reduced by as much as 50 percent when compared to the fed state, which may result in reduced effectiveness in some patients. For this reason, vilazodone should be taken with food.

Clinical Drug Studies

Design. Vilazodone's development program consisted of 24 phase I trials, 5 placebo-controlled phase II trials, and 3 phase III trials (2 double-

blind, placebo-controlled 8-week trials and 1 open-label 52-week trial). Despite important data provided by these studies, a few issues regarding study design deserve mentioning. The first relates to the phase II dosing strategies (Table 34.41–1). Only two of the phase II trials used a fixed dose titration schedule, with a maximum dose of 20 mg/day. These two trials showed a trend for a decrease in depressive symptoms at 20 mg/day, suggesting a possible treatment effect for the vilazodone 20 mg/day group. The other three phase II trials used flexible dosing strategies. While two of the studies used doses up to 100 mg/day, the flexible-dose designs limited their usefulness in providing dose-response information. The phase III trials design used vilazodone at 40 mg/day but did not include a 20 mg/day arm, despite the earlier phase II results. Additionally, three of the phase II trials included an active comparator. However, none of the active comparator drugs separated from placebo, indicating that the trials lacked assay sensitivity. Another issue involves the design of the only long-term phase III trial, which was an open-label study and lacked a placebo-control group. To date, there have been no long-term placebo-controlled studies on vilazodone's long-term efficacy and tolerability. Phase IV trials are needed to evaluate vilazodone's long-term effectiveness and safety profile. Results from a multicenter, randomized, double-blind, placebo-controlled, relapse prevention 28-week study are expected to be published in 2015. Finally, it has been suggested that vilazodone's potential advantages may include a rapid onset of action as well as decreased sexual dysfunction due to its 5HT1A partial agonism. However, the design of the phase III studies lacked an active comparator, which would have been necessary to support both of these assertions.

Table 34.41–1.
Summary of Phase II Trials

Study	Study Design	Dosing	Vilazodone Dose/Day	Placebo/Comparator	Study Length	*N*	Ages (years)	Stu Da
244	Multicenter, randomized, double-blind, placebo-controlled	Flexible dose	20–100 mg/day	Placebo fluoxetine 20 mg/day	8 weeks	410	18–65	2/(
245	Multicenter, randomized, double-blind, placebo-controlled	Flexible dose	10–20 mg 40–60 mg 80–100 mg	Placebo Fluoxetine 20 mg/day	8 weeks	481	18–70	3/(

246	Multicenter, randomized, double-blind, placebo-controlled	Fixed dose	10 mg 20 mg	Placebo citalopram 20 mg/day	8 weeks	616	18–70	12/
247	Multicenter, randomized, double-blind, placebo-controlled	Flexible dose	5–20 mg	Placebo	8 weeks	509	18–70	12/
248	Multicenter, randomized, double-blind, placebo- and active-controlled	Fixed dose	5 mg 10 mg 20 mg	Placebo	8 weeks	1,120	18–70	12/

Interpretation of Drug Studies. The FDA's approval of vilazodone was based on results from two phase III 8-week, multicenter, randomized, double-blind, placebo-controlled studies in adults with a major depressive episode (Table 34.41–2). The phase III trials had very similar designs as well as similar inclusion and exclusion criteria. In both studies, inclusion

criteria were a score of at least 22 on the 17-item Hamilton Depression Rating Scale (HAM-D-17) and at least 2 on item 1 (depressed mood). Exclusion criteria were serious psychiatric or medical comorbidity, history of inadequate response to two consecutive antidepressants, pregnant, lactating, or planning to become pregnant. Subjects receiving another antidepressant or excluded medication underwent a washout period. Participants were 18 to 65 years old in the first trial and 18 to 70 in the second. Patients randomized to vilazodone were titrated over 2 weeks to a dose of 40 mg/day. In the first trial only, subjects who did not tolerate 40 mg/day were allowed to reduce to 20 mg/day. The primary outcome measure was the Montgomery–Åsberg Depression Rating Scale (MADRS), and secondary outcomes were the HAM-D-17, Hamilton Rating Scale for Anxiety (HAMA), Clinical Global Impressions Severity of Illness scale (CGI-S) and Clinical Global Impressions Severity of Improvement scale (CGI-I). Sexual functioning was assessed in the first study using the Changes in Sexual Functioning Questionnaire (CSFQ) and in the second study using the Arizona Sexual Experiences Scale (ASEX). In both trials, treatment response was defined as a MADRS total score reduction from baseline to endpoint of at least 50 percent and remission as a MADRS total score less than 10 at end of treatment.

The majority of participants enrolled had moderate to severe depression at baseline. In both trials, vilazodone was superior to placebo in reducing depressive symptoms as measured by the mean change in MADRS total score. Vilazodone separated from placebo at week 1 in the first trial but not until week 6 in the second trial. Whether the separation at week 1 was clinically significant is unknown. It should be noted that the effect sizes in these trials were modest (0.23 and 0.30) and that the suggestions of early response were not supported by the results.

It has been postulated that vilazodone may be particularly effective in treating patients with MDD and symptoms of anxiety, given that its unique mechanism of action includes 5HT1A agonism, a property shared by the known anxiolytic, buspirone. A post hoc analysis of the two 8-week phase III trials classified patients as having anxious depression or nonanxious depression. Among anxious depressed patients, those treated with vilazodone showed significantly greater decreases in both depression and anxiety symptoms versus those on placebo. Although these results suggest a possible role for vilazodone in the treatment of depressed patients with prominent anxiety symptoms, active comparator studies are needed to determine whether vilazodone has any advantages over other antidepressants.

Table 34.41–2.
Summary of Phase III and Phase IV Clinical Trials

Study	Phase	Study Design	Intervention Arms	Study Length	*N*	Study Dates	Results
Rickels	III	Multicenter, randomized, double-blind, placebo-controlled	Vilazodone 40 mg placebo	8 weeks	410	2/06–5/07	► Vilazodone—significantly greater improvement vs. placebo on MADRS & HAMA and CGI ► Response: vilazodone 40%, placebo 28% ► Separation from placebo at week 1
Khan	III	Multicenter, randomized, double-blind, placebo-controlled	Vilazodone 40 mg/day placebo	8 weeks	481	3/08–3/09	► Vilazodone—significantly greater improvement on MADRS, HDRS-17, HARS, CGI vs. placebo ► Separation from placebo not until week 6 ► Response: vilazodone 44%, placebo 30% ► No significant differences in remission rates
Robinson	III	Multicenter, open-label	Vilazodone 40 mg/day	52 weeks	616	12/07–5/09	► MADRS: baseline—30; weeks 8–11; weeks 52–7 ► Common AEs: diarrhea, nausea, headache; most mild–moderate ► Vilazodone—no significant effect on BP,

							pulse, weight, labs, EKGs
Croft	IV	Multicenter, randomized, double-blind, placebo-controlled	Vilazodone 40 mg/day placebo	8 weeks	505	12/11–2/13	► Vilazodone—superior to placebo on MADRS, HAM-A, CGI
NCT01473381	IV	Multicenter, randomized, double-blind, placebo- and active-controlled	Vilazodone 20 mg Vilazodone 40 mg placebo Citalopram 40 mg (assay sensitivity)	10 weeks	1,120	12/11–6/13	► Results unpublished ► Initial findings suggest significantly greater improvement for all treatment arms compared with placebo in MADRS, CGI

Whereas both placebo-controlled phase III trials demonstrated significant differences in treatment response, neither showed a significant difference in remission rates. It has been proposed that the lack of remission may be attributed to the brief nature of the phase III trials and that longer duration of treatment may be necessary to observe remission, a finding frequently encountered in clinical practice. Results from a recent randomized, placebo-controlled phase IV trial represent the first available information regarding remission rates among depressed patients treated with vilazodone. Post hoc analysis of phase IV data found significantly higher remission rates in patients treated with vilazodone (34 percent) than those in the placebo group (22 percent). In this phase IV trial, significantly more patients on vilazodone (49 percent) had early improvement at 2 weeks compared to placebo (35 percent), supporting the possibility of an early response.

Head-to-head comparator studies are needed to compare vilazodone's efficacy and side-effect profile with that of other established antidepressant agents, as existing studies are limited to comparisons with placebo. A second phase IV trial comparing vilazodone with placebo included citalopram for assay sensitivity. Although results are yet unpublished, initial findings suggest that all active treatments showed significantly greater improvement in depression and global disease severity versus placebo. Additional studies are currently underway, evaluating the safety and effectiveness of vilazodone for the treatment of MDD disorder relative

to citalopram, and comparing the effects of vilazodone, paroxetine, and placebo on sexual functioning. Data on vilazodone's efficacy and tolerability compared to other antidepressants will help delineate vilazodone's role in the treatment of MDD and ultimately assist clinicians in selecting an antidepressant agent that will meet the needs of individual patients.

Effects on Specific Organs and Systems

Vilazodone was not associated with any significant changes in laboratory tests, blood pressure, heart rate, or weight during placebo-controlled trials.

Cardiac. Vilazodone was not associated with any clinically significant effects on electrocardiograms, arrhythmias, or QTc abnormalities. Four patients treated with vilazodone in phase III trials discontinued treatment because of palpitations, compared with no patients on placebo. However, in a study of vilazodone in doses up to 80 mg/day, no QTc abnormalities were observed.

Platelets. Vilazodone may affect platelets given that SSRI blockade of platelet transport proteins leads to decreased serotonin within platelets and slight inhibition of platelet aggregation. In clinical trials, however, no increased risk of bleeding was observed in patients treated with vilazodone versus placebo, and rates of bleeding events were similar in both groups. As with all SSRIs, the product labeling includes similar language, advising caution when coadministering vilazodone with drugs that potentiate the risk of bleeding.

Hyponatremia. Although no cases of hyponatremia were reported in the clinical trials of vilazodone, hyponatremia has occurred with treatment with SSRIs and SNRIs. It is potentially lethal, and risk factors include older age, concomitant use of diuretics, volume depletion, and other medications or medical conditions causing hyponatremia. The mechanism of action is thought to be due to the development of the syndrome of inappropriate antidiuretic hormone secretion (SIADH). In patients with symptomatic hyponatremia, vilazodone should be discontinued and appropriate medical intervention instituted.

Therapeutic Indications

To date, the only FDA-recommended indication for vilazodone is for the treatment of MDD in adults. Its efficacy was established in two 8-week, randomized, double-blind, placebo-controlled trials in adult patients with a diagnosis of MDD. However, given that several SSRIs and 5HT1A agonists such as buspirone are FDA approved for the treatment of anxiety disorders such as GAD, social anxiety disorder, panic disorder, and separation

anxiety disorder, as well as obsessive compulsive disorder, PTSD, premenstrual dysphoric disorder, and eating disorders, it is reasonable to think that vilazodone may likely also prove useful in treating these conditions. Several clinical trials are currently underway to evaluate vilazodone's safety, efficacy, and tolerability in GAD, PTSD, social anxiety disorder, and separation anxiety disorder.

Precautions and Adverse Reactions

Safety and Tolerability. The safety of vilazadone was evaluated using data from 2,177 patients who participated in phase III clinical studies, representing 552 patient-years of exposure. The rate of discontinuation due to adverse events in clinical trials is considered an indicator of tolerability. Vilazodone appears to be well tolerated overall. In pooled phase III data, discontinuation rates due to adverse effects were relatively low (7 percent vilazodone, 3 percent placebo). Rates of discontinuation due to lack of efficacy were not significantly different between the vilazodone and placebo groups. While no one single adverse event was responsible for discontinuation in >1 percent of the patients, the most common adverse event type leading to discontinuation were gastrointestinal complaints (2.3 percent vilazodone versus 0.2 percent placebo). Adverse events significantly more common in vilazodone patients compared with placebo were diarrhea, nausea, dizziness, insomnia, vomiting, abnormal dreams, gastroenteritis, restlessness or akathisia, palpitations, tremor, feeling jittery, and paresthesia; although in most cases symptoms were considered to be mild to moderate. Only 6.4 and 8.8 percent of participants in the two phase III trials reported a severe adverse event, compared to 5.6 and 5.4 percent in the placebo groups, respectively. In the vilazodone-treated group, symptoms in the first trial were diarrhea, fatigue, flatulence, gastroenteritis, palpitations, and paresthesias. In the second trial, symptoms reported were insomnia, nausea, vomiting, headache, and decreased libido.

Serious Adverse Events. There were no deaths in the vilazodone clinical trials in subjects taking vilazodone. Among all of the clinical trials, there were three deaths. These included a homicide in a phase I study and two suicides in a phase II study. Neither of the suicides was exposed to vilazodone. A total of 76 patients experienced a nonfatal serious adverse event in phase II and phase III trials, with similar rates for the vilazodone and placebo groups.

Gastrointestinal. The most commonly observed adverse events of vilazodone are gastrointestinal in nature. Fortunately, they tend to be mild or moderate in severity and appear to be transient in nature. In the placebo-controlled phase III trials, diarrhea was reported by 28 percent of

vilazodone-treated patients (versus 9 percent placebo) and nausea by 23 percent vilazodone-treated patients (versus 5 percent placebo). Slow dose titration over 2 weeks is recommended to minimize gastrointestinal adverse events.

Sexual Dysfunction. In the phase III placebo-controlled studies, 8 percent of vilazodone-treated patients versus 0.9 percent of placebo-treated patients reported at least one sexual adverse event, most commonly decreased libido (3.7 percent vilazodone, 0.2 percent placebo). Men tended to have a higher incidence of sexual adverse events than women, 15.5 versus 5.2 percent respectively. Pooled data from all phase III trials showed that 50 percent of men and 68 percent of women reported sexual dysfunction at baseline and that sexual function usually improved in both vilazodone and placebo groups, with over 91 percent of patients reporting stable or improved sexual function at the end of treatment. While data indicate that sexual dysfunction is a significant adverse event associated with vilazodone, it may be less severe since no patient treated with vilazodone discontinued treatment due to sexual dysfunction. However, whether vilazodone has a favorable profile regarding sexual dysfunction compared to other antidepressants will remain unknown until data from comparator studies is available.

Suicidality. Vilazodone carries the same warning as other SSRIs regarding suicidal thoughts and behaviors in children, adolescents, and young adults, based on data showing an increased risk of suicidal thinking and behavior (although no completed suicides) in children, adolescents, and young adults (ages 18 to 24) with MDD and other psychiatric disorders. In pooled phase III placebo-controlled vilazodone trials, suicidal ideation was experienced by one (0.2 percent) patient receiving vilazodone and three (0.7 percent) patients on placebo. Seven (1.2 percent) patients in the 52-week open-label study also experienced suicidal ideation. Suicidal behavior was identified by one (0.2 percent) vilazodone-treated patient and one (0.2 percent) patient in the placebo group in the pooled phase III placebo-controlled trials, and by two (0.3 percent) patients in the 52-week open-label trial. In a recent phase IV trial, among patients receiving vilazodone, there was one suicidal ideation and one suicide attempt. Patients treated with antidepressants should be monitored closely for clinical worsening or suicidality, especially upon drug initiation or the dose change.

Overdose. Limited data exist regarding overdose of vilazodone. Four patients and the child of a patient experienced an overdose of vilazodone at doses between 200 and 280 mg. Symptoms included serotonin syndrome, lethargy, restlessness, hallucinations, and disorientation. All patients

recovered.

Activation of Mania/Hypomania. As with all antidepressants, caution should be used when prescribing vilazodone to patients with a history or family history of bipolar disorder, mania, or hypomania. In clinical trials of vilazodone, six patients were identified as having probable treatment-emergent mania or hypomania, five receiving vilazodone and one placebo.

Discontinuation Syndrome. A discontinuation syndrome has been reported upon discontinuation of serotonergic antidepressants, especially when discontinued abruptly. While it is generally self-limiting, there have been reports of serious discontinuation symptoms. A gradual discontinuation of vilazodone is recommended to avoid a possible discontinuation syndrome.

Serotonin Syndrome. As with all SSRIs, the development of serotonin syndrome is possible with vilazodone, alone or when administered concurrently with other serotonergic drugs, including triptans, tricyclic antidepressants, fentanyl, lithium, tramadol, tryptophan, buspirone, and St. John's Wort, and with drugs that impair metabolism of serotonin, such as monoamine oxidase inhibitors (MAOIs). Two subjects in the vilazodone clinical trials were identified as having serotonin syndrome. The first was a patient in a phase II trial who received vilazodone 80 mg; the other was a patient in the 52-week, open-label, phase III study who took an overdose of approximately 240 mg of vilazodone. Patients should be monitored closely for possible serotonin syndrome. If symptoms occur, vilazodone and all other serotonergic agents should be discontinued immediately and supportive management initiated.

Nervous System. Although vilazodone has not been studied in patients with seizure disorders, the product label advises that caution be used when prescribing vilazodone to patients with seizure disorder. In addition, headache was reported as an adverse event in 20 percent of patients in the 52-week open-label study.

Sleep. In the two phase III trials, insomnia and abnormal dreams were reported by 6 and 4 percent of patients on vilazodone, versus 2 and 1 percent, respectively, of those receiving placebo. Decreased REM and increased slow-wave and delta-wave activity were reported in a single dose study of 10 healthy men.

Reproductive. Vilazodone is classified as Pregnancy Category C. There are no published data on vilazodone during pregnancy or during labor and delivery. In animal studies, however, vilazodone was not teratogenic in rats

or rabbits at doses of 48 and 17 times, respectively, the maximum recommended human dose (MRHD). Developmental toxicity was observed, but only at doses much higher than the MRHD. The FDA package insert carries the same warning regarding potential neonatal toxicity as is listed with other antidepressants. In the absence of clinical data on vilazodone, clinicians should consider potential risks associated with other serotonergic antidepressant agents. Although the overall rate of major malformations with SSRI exposure during pregnancy is consistent with that observed in the general population, data regarding the impact of antidepressant exposure during pregnancy on rates of miscarriage, preterm delivery, and low birth weights are mixed. Similarly, considerable debate exists regarding the association between SSRI use in pregnancy and persistent pulmonary hypertension (PPHH) of the newborn. However, it is estimated that if present, the absolute risk is very small. In fact, the FDA has concluded that it is premature to reach definitive conclusions regarding the risk. Another potential risk of prenatal antidepressant exposure is the development of poor neonatal adaptation syndrome (PNAS) as a result of third trimester exposure to serotonergic antidepressants. This syndrome is believed to be mediated by direct serotonergic toxicity or withdrawal effect, and the majority of cases are mild in severity. It is estimated that PNAS may occur in approximately 10 to 30 percent of neonates exposed to serotonergic antidepressants, although a population-based study of nearly 120,000 birth records failed to demonstrate an increased risk of PNAS in infants exposed to antidepressants. More research is needed in this area, and clinicians caring for a pregnant woman should carefully consider the risks and benefits of antidepressant medications as well as the risks of untreated depression to both the pregnant woman and the fetus.

Lactation. There are no published data regarding the excretion of vilazodone in breast milk. Animal studies have demonstrated that vilazodone is excreted into the milk of lactating rats. Risks and benefits should be carefully weighed and discussed with patients.

Fertility. There are no data on the effect of vilazodone on fertility in humans. In animal studies with rats, impairment of male fertility was observed with vilazodone at 30 times the MRHD, but not at 6 times the MRHD. No effect on fertility was observed in female rats.

Carcinogenicity. In mice studies, the incidence of hepatocellular carcinomas in males was increased at 16 times the MRHD but not at 5 times the MRHD. Similarly, the incidence of malignant mammary gland tumors was significantly increased at 16 times the MRHD but not at 1.8 times the MRHD. Vilazodone was not carcinogenic in rats at doses up to 36 times the MRHD.

Mutagenicity. According to the package insert, vilazodone was not mutagenic in the in vitro bacterial reverse mutation assay and was negative in the in vitro V79/HGRPT mammalian cell forward mutation assay. Vilazodone was clastogenic in two in vitro mammalian cell chromosome aberration assays but was negative for clastogenic activity in an in vivo rat bone marrow chromosome aberration assay and a micronucleus test. Vilazodone was also negative in an in vivo/in vitro unscheduled DNA synthesis assay in rats.

Drug Interactions

Since vilazodone is metabolized primarily by CYP3A4, 3A4 inhibitors or inducers can potentially affect its plasma concentration. If coadministered with a strong inhibitor of CYP3A4, it is recommended that the vilazodone dose be reduced to 20 mg daily. If coadministered with 3A4 inducers, the package insert states that the dose of vilazodone may be increased based on clinical response (not to exceed 80 mg/day). Due to the minor contributions from CYP2C19 and CYP2D6 in the metabolism of vilazodone, plasma levels of vilazodone are not expected to be affected by coadministration of inhibitors of these isoenzymes. The likelihood of vilazodone affecting other drugs appears to be low. A study of healthy adults found that vilazodone does not affect drugs metabolized by CYP1A2, 2C9, 2D6, or 3A4. While in vitro studies suggest that vilazodone may inhibit CYP2C8 substrates, this has not been tested in vivo.

Vilazodone's interaction with other highly protein-bound drugs has not been evaluated, but given its high protein binding, coadministration with other highly protein-bound drugs may result in increased free concentrations of the other drug. Coadministration of vilazodone with MAOIs is contraindicated, due to the risk of serotonin syndrome. At least 14 days should elapse between discontinuation of either an MAOI or vilazodone and initiation of the other. Due to potential for serotonin syndrome, caution is advised when coadministering vilazodone with other serotonergic drugs. It is also possible that concurrent use of vilazodone with NSAIDs, aspirin, or oral anticoagulants may potentiate the risk of bleeding.

Dosage and Administration

Vilazodone is supplied in 10-, 20-, and 40-mg tablets. A 30-day starter kit is available which contains (7) 10-mg tablets, (7) 20-mg tablets and (16) 40-mg tablets. Tablets should be stored at 77°F (25°C) with the following ranges permitted: 15°C to 30°C/59°F to 86°F. Vilazodone should be taken with food. The recommended dose for vilazodone is 40 mg daily, and should be titrated over 2 weeks, starting with 10 mg/day for 7 days, 20 mg/day for the next 7 days, and then 40 mg/day. No dose adjustments are

recommended for mild, moderate, or severe hepatic or renal impairment.

Geriatric Patients. No dose adjustment is recommended for use in geriatric patients. This is based on results from a study demonstrating similar pharmacokinetics in patients >65 years old and those of ages 24 to 55 years. Nonetheless, the product labeling cautions that greater sensitivity of some older individuals cannot be ruled out and that elderly patients may be at greater risk for hyponatremia.

Pediatric Patients. Vilazodone is not FDA approved for use in pediatric patients. A clinical trial is being conducted to evaluate the efficacy, safety, and tolerability of vilazodone relative to placebo in adolescents 12 to 17 years of age with MDD.

Future Directions

The need for effective new antidepressant treatments is evident given the high prevalence of depressive disorders and the relatively low response and even lower remission rates of existing antidepressant agents. Vilazodone is the first in a novel class of antidepressants, SPARI, named after its dual serotonin reuptake inhibitor and partial 5HT1A agonist mechanism of action. Although additional data are needed to determine vilazodone's relative efficacy and tolerability compared to other antidepressants, its unique mechanism of action has garnered considerable attention. There is growing interest in the role of the 5HT1A receptor for the treatment of depression and anxiety. More specifically, attention is being focused on 5HT1A's potential to indirectly modulate glutamate transmission. New clinical trials examining vilazodone's impact on glutamate levels in the anterior cingulate cortex may further expand vilazodone's role and widen its clinical applications.

References

Boinpally R, Gad N, Gupta S, Periclou A. Influence of CYP3A4 induction/inhibition on the pharmacokinetics of vilazodone in healthy subjects. *Clin Ther*. 2014;36(11):1638–1649.

Choi E, Zmarlicka M, Ehret MJ. Vilazodone: a novel antidepressant. *Am J Health Syst Pharm*. 2012;69:1551–1557.

*Citrome L. Vilazodone for major depressive disorder: a systematic review of the efficacy and safety profile for this newly approved antidepressant—what is the number needed to treat, number needed to harm and likelihood to be helped or harmed? *Int J Clin Pract*. 2012;66(4):356–368.

Citrome L, Gommoll CP, Tang X, Nunez R, Mathews M. Evaluating the efficacy of vilazodone in achieving remission in patients with major depressive disorder: post-hoc analyses of a phase IV trial. *Int Clin Psychopharmacol*. 2015;30(2):75–81.

Clayton AH, El Haddad S, Iluonakhamhe JP, Ponce MC, Schuck AE. Sexual dysfunction

associated with major depressive disorder and antidepressant treatment. *Expert Opin Drug Saf*. 2014;13(10):1361–1374.

Clayton AH, Kennedy SH, Edwards JB, Gallipoli S, Reed CR. The effect of vilazodone on sexual function during the treatment of major depressive disorder. *J Sex Med*. 2013;10:2465–2476.

Croft HA, Pomara N, Gommoll C, Chen D, Nunez R, Mathews M. Efficacy and safety of vilazodone in major depressive disorder: a randomized, double-blind, placebo-controlled trial. *J Clin Psychiatry*. 2014;75(11):e1291–e1298.

Dawson LA, Watson JM. Vilazodone: a 5-HT1A receptor agonist/serotonin transporter inhibitor for the treatment of affective disorders. *CNS Neurosci Ther*. 2009;15(2):107–117.

Deardorff WJ, Grossberg GT. A review of the clinical efficacy, safety and tolerability of the antidepressants vilazodone, levomilnacipran and vortioxetine. *Expert Opin Pharmacother*. 2014;15(17):2525–2542.

*Epstein RA, Moore KM, Bobo WV. Treatment of nonpsychotic major depression during pregnancy: patient safety and challenges. *Drug Healthc Patient Saf*. 2014;6:109–129.

Ferrando SJ, Owen JA, Levenson JL. Psychopharmacology. In: Hales RE, Yudofsky SC, Roberts LW, eds. *The American Psychiatric Publishing Textbook of Psychiatry*. 6th ed. Arlington, VA: American Psychiatric Association; 2014.

Frampton JE. Vilazodone: in major depressive disorder. *CNS Drugs*. 2011;25(7):615–627.

Guay DR. Vilazodone hydrochloride, a combined SSRI and 5-HT1A receptor agonist for major depressive disorder. *Consult Pharm*. 2012;27(12):857–867.

Jain R, Chen D, Edwards J, Mathews M. Early and sustained improvement with vilazodone in adult patients with major depressive disorder: post hoc analyses of two phase III trials. *Curr Med Res Opin*. 2014;30(2):263–270.

Jasiak NM, Bostwick JR. Risk of QT/QTc prolongation among newer non-SSRI antidepressants. *Ann Pharmacother*. 2014;48(12):1620–1628.

Khan A, Cutler AJ, Kajdasz DK, et al. A randomized, double-blind, placebo-controlled, 8-week study of vilazodone, a serotonergic agent for the treatment of major depressive disorder. *J Clin Psychiatry*. 2011;72(4):441–447.

Khan A, Sambunaris A, Edwards J, Ruth A, Robinson DS. Vilazodone in the treatment of major depressive disorder: efficacy across symptoms and severity of depression. *Int Clin Psychopharmacol*. 2014;29(2):86–92.

*Laughren TP, Gobburu J, Temple RJ, et al. Vilazodone: clinical basis for the US Food and Drug Administration's approval of a new antidepressant. *J Clin Psychiatry*. 2011;72(9):1166–1173.

Lindsey WT. Vilazodone for the treatment of depression. *Ann Pharmacother*. 2011;45(7–8):946–953.

Newport DJ, Stowe ZN. Psychopharmacology during pregnancy and lactation. In: Schatzberg AF, Nemeroff CB, eds. *Essentials of Clinical Psychopharmacology*, 3rd ed. Arlington, VA: American Psychiatric Publishing; 2013.

Pehrson AL, Sanchez C. Serotonergic modulation of glutamate neurotransmission as a strategy for treating depression and cognitive dysfunction. *CNS Spectr*. 2014;19:121–133.

Rickels K, Athanasiou M, Robinson DS, Gibertini M, Whalen H, Reed CR. Evidence for efficacy and tolerability of vilazodone in the treatment of major depressive disorder: a randomized, double-blind, placebo-controlled trial. *J Clin Psychiatry*. 2009;70(3):326–333.

Robinson DS, Kajdasz DK, Gallipoli S, Whalen H, Wamil A, Reed CR. A 1-year, open-label study assessing the safety and tolerability of vilazodone in patients with major depressive disorder. *J Clin Psychopharmacol*. 2011;31(5):643–646.

Schwartz TL, Siddiqui UA, Stahl SM. Vilazodone: a brief pharmacological and clinical review of the novel serotonin partial agonist and reuptake inhibitor. *Ther Adv Psychopharmacol*. 2011;1(3):81–87.

Singh M, Schwartz TL. Clinical utility of vilazodone for the treatment of adults with major depressive disorder and theoretical implications for future clinical use. *Neuropsychiatr Dis Treat*. 2012;8:123–130.

Spina E, Trifirò G, Caraci F. Clinically significant drug interactions with newer antidepressants. *CNS Drugs*. 2012;26(l):39–67.

*Stahl SM. Mechanism of action of the SPARI vilazodone: serotonin 1A partial agonist and reuptake inhibitor. *CNS Spectr*. 2014;19:105–109.

Stahl SM. *Stahl's Essential Psychopharmacology Online*. New York: Cambridge University Press; 2014, http://stahlonline.cambridge.org/essential_4th_toc.jsf.

Thase ME, Chen D, Edwards J, Ruth A. Efficacy of vilazodone on anxiety symptoms in patients with major depressive disorder. *Int Clin Psychopharmacol*. 2014;29:351–356.

*Wang SM, Han C, Lee SJ, Patkar AA, Masand PS, Pae CU. A review of current evidence for vilazodone in major depressive disorder. *Int J Psychiatry Clin Pract*. 2013;17:160–169.

35 Child Psychiatry

▲ 35.1 Introduction and Overview

CAROLY SUE PATAKI, M.D.

ADVANCES IN CHILD AND ADOLESCENT PSYCHIATRY

Over the last decade the field of child and adolescent psychiatry has demonstrated significant gains in knowledge of the neurotransmitters, neural circuits, and brain systems involved in a variety of childhood psychiatric symptoms and disorders. Efforts to link this knowledge to improved treatments for childhood psychiatric disorders have met with some success. Data from large randomized controlled studies have shown efficacy of medications, cognitive-behavioral therapy, and their combination, for certain disorders in youth including obsessive-compulsive disorder (OCD); several anxiety disorders: generalized anxiety disorder, separation anxiety disorder, social anxiety disorder; and major depressive disorder.

Genetic research, including genome-wide linkage scans seeking to identify gene loci, have focused on several childhood disorders, including attention-deficit/hyperactivity disorder (ADHD), autism spectrum disorders (ASDs), Tourette syndrome (TS), and OCD. Neuroimaging studies have elucidated the brain development of typically developing children, as well as in children with psychiatric disorders. For example, neuroimaging studies have revealed that whereas total brain volume remains virtually stable after the age of 5 years, there are two dynamic processes—synaptic pruning and myelination—that continue through adolescence and early adulthood and profoundly influence cognitive and social decision making and behavior. Furthermore, neuroimaging studies have led to the conclusion that girls and boys differ in the timing of maximal gray matter volume and its normal decline after puberty, in that the process occurs about 1 year earlier in boys than in girls. A variety of neuroimaging techniques, such as magnetic resonance imaging (MRI), functional MRI (fMRI), diffusion tensor imaging (DTI), computerized

tomography (CT), single photon emission computerized tomography (SPECT), and positron emission tomography (PET) have been increasingly employed to uncover both morphology and function of various parts of the brain to further understand childhood psychiatric disorders such as ADHD, mood disorders, OCD, and anxiety disorders, including posttraumatic stress disorder (PTSD).

The current *Diagnostic and Statistical Manual of Psychiatric Disorders*, 5th edition (DSM-5), published in 2013, has modified criteria for several childhood disorders in an effort to refine and match the criteria with available scientific evidence and expert consensus. It is well known that youth with the same psychiatric disorder, as defined by the DSM-5, may present with heterogeneous symptom constellations. Diagnostic categories have played an important role in providing explanations of symptoms to patients, communication between clinicians, and providing a method for researchers to devise systematic studies. A limitation of a categorical (disorder) approach, rather than a dimensional (symptom or neurobiological domain) approach in child psychiatric treatment research is that youth often present with overlapping symptoms of several disorders or present with multiple comorbid disorders. In fact, it is quite common for children to present with more than one psychiatric disorder. This clinical heterogeneity among childhood psychiatric disorders makes it a challenge to determine which constellation of symptoms will be most responsive to a given treatment.

In an effort to advance the *precision* of future psychiatric research, the National Institute of Mental Health (NIMH) developed the Research Domain Criteria (RDoC) project (2008) which focuses on observable behavior and neurobiological measures, rather than DSM-5 disorders. RDoC domains and examples of each include *negative valence domain* (fear, loss), *positive valence systems* (responsiveness to reward, habit), *cognitive systems* (attention, working memory), *systems for social processes* (social communication, attachment), *and arousal/modulatory systems* (arousal, sleep–wake). An example of how this model might be applied is in a study of impairing anxiety, a fear-induced startle response might be measured in people with different relevant genetic polymorphisms and a subjective level of distress could also be measured. In this way, a more specific constellation of variables could be associated with specific responses to either fear-inducing stimuli, or potential treatments. The goal of this project is to identify a more specific set of symptoms and traits that mediate the expression of behaviors and predict response to potential treatments.

An in-depth description of the clinical presentation, course, and treatment as well as updates on the scientific advances in childhood psychiatric disorders is presented in the following chapters.

CLASSIFICATION OF CHILDHOOD PSYCHIATRIC DISORDERS AND HIGHLIGHTS OF MODIFICATIONS IN DSM-5

The current version of the *Diagnostic and Statistical Manual of Mental Disorders*, 5th edition (DSM-5), provides a categorical classification reflecting evidence-based and expert consensus of current childhood disorders. Children and adolescents who meet criteria for the same psychiatric diagnoses often present with heterogeneous clinical pictures. DSM-5 defines a psychiatric (mental) disorder as *a syndrome characterized by clinically significant disturbance in an individual's cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning. Mental disorders are usually associated with significant distress or disability in social, occupational, or other important activities.*

The first edition of the World Health Organization's *International Statistical Classification of Diseases* to contain a section for mental illness was the sixth (ICD-6). The first formal draft of a US variant of this document, the DSM, was developed shortly thereafter, in 1952, by the Committee on Nomenclature and Statistics of the American Psychiatric Association.

Dennis Cantwell, M.D., along with many colleagues, was instrumental in advancing the classification of child psychiatric disorders through suggested revisions to the DSM in an effort to facilitate empirical investigation. These efforts were first realized in the development of the DSM-III in 1980. The DSM-III included explicit diagnostic criteria for each disorder, use of the multiaxial system, and a diagnostic focus on phenomenology rather than etiology. The DSM-5 does not include the multiaxial system, however, it allows for specifiers to denote important psychosocial factors.

Although both categorical and dimensional approaches have merit in describing psychiatric syndromes in youth, the DSM system uses only categorical criteria to identify psychiatric disorders. The overlap of behavioral and mood symptoms across multiple childhood and adolescent disorders made defining the boundaries challenging. In addition, it soon became evident that in more severely psychiatrically impaired youth, comorbidity is the rule rather than the exception.

Highlights of changes in childhood psychiatric diagnostic criteria in DSM-5 include the following:

Autism spectrum disorder is a new DSM-5 disorder characterized by core deficits in two areas: (1) social communication/social interaction, and (2) restricted pattern of interests, behaviors or activities. This new disorder included the previous five pervasive developmental disorders: autistic disorder, Asperger disorder, childhood disintegrative disorder, Rett disorder, and pervasive developmental disorder not otherwise specified.

Social (pragmatic) communication disorder—a new DSM-5 disorder characterized by persistent deficits in the social uses of verbal and nonverbal communication, however, without a restricted pattern of behaviors or interests.

Specific Learning disorder—new disorder which encompasses the previous disorders: reading disorder, mathematics disorder, disorder of written expression, and learning disorder not otherwise specified. Specifiers indicate which skills are affected.

Disruptive mood dysregulation disorder—a new DSM-5 disorder characterized by children and adolescents with persistent irritability and frequent episodes of severe tantrums.

There are several additional modifications to previous disorders which will be detailed in the corresponding chapter.

DEVELOPMENTAL PERSPECTIVE: IDENTIFYING THE EMERGENCE OF A CHILDHOOD PSYCHIATRIC DISORDER

Onset

Onset of a psychiatric disorder is a challenging concept in child and adolescent psychiatry because it is often difficult to pinpoint reliable landmarks for sufficient symptoms and impairment. Defining the onset of a disorder by a single impairing symptom might lead to the inclusion of a heterogeneous group of disorders, whereas inclusion of only a rigid set of symptoms might exclude important subtypes. Age at onset of a childhood or adolescent psychiatric disorder is a variable of considerable importance because the emergence of a known constellation of symptoms at a younger-than-expected age might reflect increased vulnerability genetically or environmentally. For example, an earlier-than-expected age of onset of major depressive disorders has been associated with an increased probability of depressive disorders in first-degree relatives and implies an increased genetic vulnerability.

Onset of Child and Adolescent Psychiatric Disorders

The investigation of age at onset of a psychiatric disorder can be applied to models of delayed onset of illnesses. For example, degeneration is a model of delayed onset of disease that characterizes conditions such as Alzheimer disease and Parkinson disease—that is, gradual loss of target neurons over time results in a predictable behavioral and neurological deficit. Developmental failure (i.e., a lack of physiological development within the nervous system at the appropriate time) is another model of delayed onset of a disease process.

A theory of delayed onset of a disorder, the two-hit model, hypothesizes that the emergence of a disorder is related to a “primary” genetic vulnerability triggered by exposure to a second event that may be

environmental. For example, the rate of emergence of bipolar disorders increases markedly during adolescence, is maximal in young adulthood, and then gradually diminishes. One model to explain the emergence of a disorder with this distribution is a genetic predisposition (first hit) leading to vulnerability for the disorder and a second exposure, possibly the process of puberty itself or an adverse life event (second hit), that results in the emergence of the disorder.

Genetic anticipation is defined as the emergence of more severe forms of a genetic disorder in successive generations. This mechanism has been shown to contribute to the severity of pathology in a number of genetic disorders that are caused by mutations of DNA triplet repeat nucleotide sequences. Increases in repeat patterns of three nucleotides (e.g., GAC GAC GAC) increase the length of DNA sequences and result in pathological genes. Triplet repeat expansion has been shown to be the mechanism of a number of genetic disorders, including Huntington disease and fragile X syndrome. Anticipation has been suggested as a possible explanation for increasingly severe psychopathology in successive generations in psychiatric disorders. The investigation of possible molecular bases of anticipation in psychiatric disorders, such as the increased density of genetic triplet repeat expansions shown in Huntington disease and fragile X mental retardation, is on the horizon.

Neuroimaging Studies in Child and Adolescent Psychiatry

Neuroimaging studies have provided evidence of the dynamic changes that occur in developing brains. fMRI, in studies by Swartz and colleagues (2012), has shown, for example, that children with ASDs have hyperaroused amygdalae leading to deficits in habituating to a variety of stimuli including sounds, sad faces, and certain tactile feelings. This may contribute to the typical sensory sensitivities in children. Singh and colleagues (2012) employed fMRI to show that adolescents with bipolar disorder have less activation in superior frontal brain regions after a negative experience, compared to healthy adolescents, which may be indicative of deficits in executive control over emotional decisions. Child and adolescent brains have been described by clinicians and researchers as "works in progress," with a series of predictable pattern of biological processes that extend into and beyond adolescence. The neuroimaging work of Giedd, Rapoport, Rosenberg, Castellanos, and others has revealed the typical wave of high production of gray matter among developmentally normal adolescents emerging just before puberty. Furthermore, investigations have documented the links between the pruning and myelination processes in various sites in the brain and the subtle increases in social and emotional problem-solving abilities that typically emerge in later adolescence. Other investigations have identified areas of the brain targeted in children who learn first and second languages early in life. fMRI

scans have determined the spatial relationship of language centers in children who have first and second languages. In such children, both language centers appear in the same cortical region whereas when a second language is acquired in adult life, the new language center is not represented in the same cortical region as the first language. MRI studies have also provided evidence that adolescents process emotional stimuli differently than adults. Using fMRI and presenting pictures of faces depicting various emotions, Yurgelun-Todd and colleagues showed that in young teens, who typically have difficulties with the task, the amygdala (a brain center for fear reactions) is activated during the task of identifying the emotions presented; whereas in older teens, who perform better on the task, the frontal lobes are more active in the process of identifying emotions in the stimuli. There were also additional findings indicating that language tasks correlated with activation of temporal lobes in young teens and of frontal lobes in older teens.

GENE X ENVIRONMENT INTERACTIONS: RISK AND PROTECTIVE FACTORS

There is a convergence of evidence regarding the interactional effects of genetic vulnerability and environmental events on the development of many psychiatric disorders. Much research has focused on the alleles of the serotonin transporter gene and sensitivity to adverse events and stress. The impact of age and developmental stage that a stressful event is experienced adds another factor in the equation, and finally, positive relational and environmental variables may increase resiliency. Risk factors include all variables that increase the probability that a given child or adolescent will develop psychopathology. Protective factors decrease the risk of developing psychopathology. Demographics, psychosocial factors, biological factors, genetics, family environment, and external environment interact to produce risk and protection from psychiatric disorders. Rarely can a single risk factor account for the entire variance between the emergence and inhibition of a psychiatric disorder. Thus, the study of psychiatric risk is complicated and multifactorial. Risk factors often studied for influence on the emergence of psychiatric disorders among children and adolescents include family history of psychiatric disorders, social class, intellectual function, adverse life events, temperamental factors, and the quality of family and peer relationships. Developmental delays, child maltreatment, and high levels of family conflict increase the risk of psychiatric disorder. Protective factors that have been identified and studied include individual temperamental predisposition, positive family relationship, and success experiences academically and socially.

INTERVENTIONS FOR PSYCHIATRIC DISORDERS IN YOUTH

Interventions to address psychiatric symptoms and disorders in youth include clinical, targeted, and universal interventions.

Clinical interventions provide treatment to a family with a child or adolescent identified as having a psychiatric problem. Psychosocial, psychopharmacological, and other environmental interventions can be compared with a placebo condition and with each other to assess efficacy. Approximately one-half of families who begin psychiatric treatment terminate prematurely, thus, the nature and delivery of the treatments also merit investigation.

Targeted interventions are designed for children identified as at increased risk of a psychiatric disorder but whose families are not seeking treatment. Children at risk may be identified either by an external factor, such as a family characteristic (e.g., offspring of a parent known to have a psychiatric disorder) or through a behavior (e.g., defiance or aggression in the classroom). A given child may be monitored or receive a psychosocial intervention on the basis of the risk factor—whether psychiatric disorder is already present or not; however, these interventions may not be welcomed by the family, or threshold for interventions may not be correlated with a clinically relevant situation.

Universal *interventions,* also termed *primary prevention,* are received by all children and families within a specific geographical distribution. They may occur throughout a targeted school or community or on a citywide, statewide, or national basis. This strategy obviously reaches more children than any other method of intervention, but questions arise about the benefit and need of the intervention for the population at large, as well as the cost–benefit ratio.

Study of the risk factors and protective factors is leading the field to develop the most effective combination of prevention strategies for children and adolescents. More effective treatments are being investigated in order to match the treatment to the individual. Future goals include mitigating risk factors, enhancing protective factors, and when prevention is not possible, to provide efficacious treatments.

References

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Arlington, VA: American Psychiatric Association; 2013.

Barr CL. Genetics of childhood disorders: XXII. ADHD, Part 6: The dopamine D4 receptor gene. *J Am Acad Child Adolesc Psychiatry*. 2001;40:118–121.

Blakemore SJ, Choudhury S. Development of the adolescent brain: implications for executive function and social cognition. *J Child Psychol Psychiatry*. 2006;47:296–312.

Blakemore SJ, den Ouden H, Choudhury S, Frith C. Adolescent development of the neural circuitry for thinking about intentions. *Soc Cogn Affect Neurosci*. 2007;2:130–139.

Court MH. A pharmacogenomics primer. *J Clin Pharmacol*. 2007;47:1087–1103.

Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med*. 2013;11:126.

Ernst M, Rumsey JM, eds. *Functional Neuroimaging in Child Psychiatry*. Cambridge: Cambridge University Press; 2000.

Franklin ME, Sapyta J, Freeman JB, et al. Cognitive behavior therapy augmentation of pharmacotherapy in pediatric obsessive-compulsive disorder: the Pediatric OCD Treatment Study II (POTS II) randomized controlled trial. *JAMA*. 2011;306:1224–1232.

Giedd JN, Rapoport JL. Structural MRI of pediatric brain development: what have we learned and where are we going? *Neuron*. 2010;67:728–734.

Gold AL, Brotman MA, Adleman NE, et al. Comparing brain morphometry across multiple childhood psychiatric disorders. *J Am Acad Child Adolesc Psychiatry*. 2016;55(12):1027–1037.

Goodyer IM. Genes, environments and depressions in young people. *Arch Dis Child*. 2015;100(11):1064–1069.

Kaufman J, Gelernter J, Hudziak JJ, Tyrka AR, Coplan JD. The Research Domain Criteria (RDoC) project and studies of risk and resilience in maltreated children. *J Am Acad Child Adolesc Psychiatry*. 2015;54:617–625.

Kavanaugh BC, Dupont-Frechette JA, Tellock PP, Maher ID, Haisley LD, Holler KA. Neurocognitive phenotypes in severe childhood psychiatric disorders. *Nerv Ment Dis*. 2016;204(10):770–777.

Krueger RF, Hopwood CJ, Wright AG, Markon KE. Challenges and strategies in helping the DSM become more dimensional and empirically based. *Curr Psychiatry Rep*. 2014;16:515.

Lenroot RK, Giedd JN. Annual Research Review: Developmental considerations of gene by environment interactions. *J Child Psychol Psychiatry*. 2011;52:429–441.

McMahon FJ, dePaula JR. Genetics and age at onset. In: Shulman KI, Tohen M, Kutcher SP, eds. *Mood Disorders across the Life Span*. New York: Wiley; 1996.

Molenaar D, Middeldorp C, van Beijsterveldt T, Boomsma DI. Analysis of behavioral and emotional problems in children highlights the role of genotype × environment interaction. *Child Dev*. 2015;86(6):1999–2016.

National Institute of Mental Health. *The National Institute of Mental Health Strategic Plan*. Bethesda, MD: National Institute of Mental Health; 2008 [http://www.nimh.nih.gob/about/strategic-planning-reports/index.html]

Nigg JT. Temperament and developmental psychopathology. *J Child Psychol Psychiatry*. 2006;47:395–422.

Piacentini J, Bennett S, Compton SN, et al. 24- and 36-week outcomes for the Child/Adolescent Anxiety Multimodal Study (CAMS). *J Am Acad Child Adolesc Psychiatry*. 2014;53:297–310.

Rauh VA, Margolis AE. Research review: environmental exposures, neurodevelopment, and child mental health - new paradigms for the study of brain and behavioral effects. *J Child Psychol Psychiatry*. 2016;57(7):775–793.

Rettew DC, Althoff RR, Dumenci L, Ayer L, Hudziak JJ. Latent profiles of temperament and their relations to psychopathology and wellness. *J Am Acad Child Adolesc Psychiatry*. 2008;47:273–281.

Singh MK, Chang KD, Kelley RG, et al. Reward processing in adolescents with bipolar I disorder. *J Am Acad Child Adolesc Psychiatry*. 2013;52:68–83.

Sowell ER, Thompson PM, Holmes CJ, Jernigan TL, Toga AW. In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nat Neurosci*. 1999;2:859–861.

Swartz J, Wiggins JL, Carrasco M, Lord C, Monk CS. Amygdala habituation and prefrontal functional connectivity in youth with autism spectrum disorders. *J Am Acad Child Adolesc Psychiatry*. 2013;52:84–93.

Teicher MH, Samson JA. Annual research review: enduring neurobiological effects of childhood abuse and neglect. *J Child Psychol Psychiatry*. 2016;57(3):241–266.

Thompson PM, Giedd JN, Woods RP, MacDonald D, Evans AC, Toga AW. Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature*. 2000;404:190–193.

*Treatment for Adolescents With Depression Study (TADS) Team; March J, Silva S, Curry J, et al. The Treatment for Adolescents With Depression Study (TADS): outcomes over 1 year of naturalistic follow-up. *Am J Psychiatry*. 2009;166:1141–1149.

*Uher R. Gene-environment interactions in severe mental illness. *Front Psychiatry*. 2014;5:48.

Walsh ND, Dalgleish T, Dunn VJ, et al. 5-HTTLPR-environment interplay and its effects on neural reactivity in adolescents. *Neuroimage*. 2012;63:1670–1680.

White TJ. A window into the neurobiology of childhood and adolescent psychopathology. *J Am Acad Child Adolesc Psychiatry*. 2013;52:9–11.

Yurgelun-Todd D. Emotional and cognitive changes during adolescence. *Curr Opin Neurobiol*. 2007;17:251–257.

Züchner S, Roberts ST, Speer MC, Beckham JC. Update on psychiatric genetics. *Genet Med*. 2007;9:332–340.

▲ 35.2 Normal Child Development

GABRIELLE SHAPIRO, M.D., ROBERT J. JAFFE, M.D., AND ALEXANDER KOLEVZON, M.D.

INTRODUCTION

Imagine an infant reciting Shakespeare, a toddler calculating trigonometry angles, or an 8-year-old driving to grade school each morning. The thought of children performing mundane adult tasks highlights the obvious: humans are born with a very limited skillset and develop a more complex one over time. In essence, development is the acquisition of new skills and the multiplex reactions to these additions. Whether it is the capacity to walk, talk, empathize, or think about the future, suddenly a capacity is there that previously was not. At no time in life is the explosion in newfound possibilities more pronounced than in early childhood, the focus of this section.

With a basic understanding of development, clinicians of every discipline can benefit their patients. Pathology can frequently be traced back to childhood, offering clues as to its inception and psychological sequelae. Often, adult patterns of symptom presentation or behavioral interaction with the environment mimic patterns from childhood.

Until the 1600s, children were viewed as “small adults” and expected to

work and socialize in an adult manner. Children were sources of cheap labor during the Industrial Revolution. This era's end, coupled with compulsory school laws in the United States, meant that by the early 20th century the vast majority of children were attending schools. John Dewey was essential in viewing schools as a vehicle through which childhood development could be fostered. G. Stanley Hall formally studied children and adolescents, and blazed a trail for the relatively recent surge in psychological theories by Sigmund Freud, Erik Erikson, Jean Piaget, and John Watson, to name a few.

Though the field remains ripe for further exploration, some critical concepts have emerged over the past 100 or so years. Normal development can be conceived of as what a typical child will go through in an average expectable environment. This section will highlight the ages when different changes are typically seen. It is imperative to understand that each new advance has a range of when its existence should be expected. Significant lags are thought of as developmental delays (with the exception of premature babies) whereas marked abnormalities in development are developmental disorders.

There is a debate as to whether development is a continuous or discontinuous process. What is clear is that the direction of the curve is not always positive, as there is a normal degree of progression and regression. It is normal during the acquisition phase of new skills for them to come and go. An example is a fluidly speaking 3-year-old who, upon the birth of a younger sibling, will revert to baby talk. This is distinct from a frank loss of a solidified skill, a symptom which should always prompt inquiry.

Current research offers evidence that the mechanism for growth comes from the interplay of genetics and the environment. Though the relative contribution of each remains an area of debate, this transactional model has become the established paradigm. Genetic background primes an individual and provides potential, but can be altered positively and negatively based on environmental stimulation. Identical twins reared in the same household do not develop identical personalities. There is evidence that children shape their environment and elicit different responses from their environment, affecting how they develop. Some aspects of development are more genetically determined, whereas the environment may play a greater role in others. When pertinent, this section will highlight relative contributions.

This chapter will cover nine areas of development from birth through adolescence. The domains of brain, language, cognitive, social, emotional, motor, sensory, gender, and psychosexual development will each be addressed. Developmental lines are also interdependent, maturing on their own yet influenced by one another. For example, language fosters advances with emotional and social development.

Finally, the scope of this chapter is on normal childhood development

only. Clinical correlates will be interspersed throughout to highlight the importance of understanding normal development in practice. There is also a comprehensive master table (Table 35.2–1) arranged by age on one axis and developmental domain on the other. This is intended as a quick reference to identify skills typically developed at a certain point in time or to follow one domain longitudinally.

BRAIN DEVELOPMENT

Unlike some other species, humans are born with an immature central nervous system (CNS). An infant is entirely dependent on others and is governed by innate reflexes, yet in time becomes a creature capable of complex cognitive, social, emotional capabilities and remarkable physical and verbal skills. How does one go from someone who cannot hold his own head up to an athlete who can sense his orientation in space and manipulate his body accordingly? This section will focus on the physical changes to the brain during infancy and childhood which allow for all the capacities described in the rest of the chapter.

Table 35.2–1.
Child Development Resource Table

Periods:		Gestational-Infancy: 0–2 Years			
	Theories	**0–2 Months**	**2–7 Months**	**7–12 Months**	**1–**
Emotional Development	(Integrated Theory) Acquisition of Basic Emotions	Stage 1: Infant has innate emotions and emotional responsiveness —emotions are available and easily stimulated at birth			Stag c e r
	Emotional skills	Touching evokes love, loud noises evoke fear. Restrictions evoke rage. The brain pathways for the forming of emotions begin to develop	Emotions begin to become self-regulated as the brain pathways of emotion grow	Self-regulation of emotion continues to develop	Shar e emb Dis oth
		Infants have a "social smile"	Begin to laugh		Begin er

	Emotional Behavior	in response to human voices. They display the ability to respond to emotions of others	and display more control over smiles. Anger is able to be shown as well	Ability to induce responsiveness increases	exp fe inf att ap lik alo
Social Development	John Bowlby Attachment Theory	Preattachment (0 to 6 weeks)	Attachment making (6 weeks to 8 months)	In month 8, clear-cut att begins to develop >> anxiety	
	Margaret Mahler Attachment Theory	Autistic Phase (0 to 6 weeks)	Develops a symbiotic relationship with the caregiver (6 weeks to 8 months), and the beginnings of differentiation	Differentiation continues	Rapp (18
	Robert Selman's Stages of Perspective Taking				

	Erik Erikson's Social Stages (Psychosexually Influenced)	Stage 1: Crisis: Basic Trust vs. Mistrust Virtue: Hope			C
	Social Behaviors	Infant acknowledges the caregiver as "different," recognizing caregiver's smell, voice, touch	Infant is able to differentiate self from caregiver and begins to manifest separation anxiety		Chil co car dev l ca a
Psychosexual Development	Freudian Psychosexual Stages	Oral Stage			
	Gender Identity				Gend and c
	Gender Role				Ster l
	Gender Relationship				Sex n v
	Psychosexual Skill	Ego directs sucking needs			Chi st obj lik 1.5 can se p

			oth clo hai as p 2.5 lac ov
Moral Development	Piaget's Stages of Moral Behavior Related to Cognitive Phases		
	Piaget's Two-Stage Theory of Moral Judgment (Cognitively Based)		
	Kohlberg's Six-Stage Theory of Moral Reasoning (Cognitively, Socially, and Emotionally Influenced)		
	Freud's Concept of Superego as Conscience		

	(Psychosexually Influenced)				
	Gilligan's Sex Theory of Moral Development (Cognitively and Psychosexually Influenced)	Agrees with K			
	Moral Skills	Sense of rightness is limited to the self >> Never having been hu isolated, baby knows that unresponsiveness to touch is wrong others. Cannot judge "right" and "wrong." The child must neg want			
Cognitive Development	Piaget's Stages of Cognitive Development	**Sensorimotor Stage:** Object permanence (8–9)			
	Cognitive Achievements	Recognizes faces (1st 2) Sleeps 16–18 hours, 2 naps	Eye-hand (4)	Hand/hand transfer (6–8) Self from caregiver (8)	Uses Follow Sleeps Names Know
Physical Development	Language Achievements	Cries (birth)	Coos (3–4)		First Combi
	Perceptual Achievements	All senses are active Can recognize mother's voice Can see patterns, shapes, and contrast	Eye tracking (2) Stereopsis (3) Follows 180 degrees (4) Monocular depth (6–7)	Near adult visual acuity (8)	
	Motor Achievements	Reflexes (birth)	Supports head (2–4) Rolls over in both directions Sits unsupported (6–7) Pat-a-cake (6–8)	Crawls (8–9) Pincer grasp (9) Stands (9–10) Cruises (10–12)	First w Po Beg s Eats/

There are two important and related concepts affecting development: *critical periods* and *plasticity*. Critical periods are windows of opportunity for growth in certain areas. This window offers an opportunity for tremendous gains, but is also vulnerable to disruption. For example, the first few years appear to be important for language, in particular the first year for a perceptual map of phonemes. Plasticity enables the ability of one part of the brain to compensate for another. An example is the primary auditory cortex in congenitally deaf individuals which assists in processing visual and somatosensory stimuli.

Early in life there are many changes occurring within the brain. Along with the rest of the body, the brain is growing substantially during the first few years. By age 5, it has tripled in size and is at 95 percent of its adult mass (1.2 kg for adult women and 1.3 kg for men). Growth occurs bottom-up in a caudocranial manner with the pons, medulla, and spinal cord ahead of the cerebellum, which is ahead of the cortex. This explains why infants rely on reflexes and instincts and lack coordination.

Histologically, the adult brain consists of billions of neurons configured into trillions of synapses. Early on it is a hotbed of activity. Neuronal *proliferation* and *differentiation* explode in utero and continue into the first year of life, particularly in the cortex. Though neurogenesis slows and was previously thought to halt, recent evidence has shown that two areas in adult brains possess the ability for new neurons, the subventricular zone and hippocampal dentate gyrus. These newly formed fetal and infantile neurons must get from their place of origin to their destination, which they accomplish in a task called *migration*. This journey is what lays down cortical columns and continues through the first 6 to 7 months of life, with the later neurons settling in the cortex. Once the neurons have reached their destination (and in certain cases even during migration), the cells experience dendritic and axonal outgrowth. This extension is particularly heavy during the first year, and continues over the following 18 months.

The settled and grown neurons rapidly interact with one another with an explosion of new synapses during the first 2 to 3 years of life. By age 2, the toddler has roughly twice the number of synapses an adult will have. This number gets reduced as the child's synapses are *pruned* from 18 months to age 20, with the majority occurring between 7 and 15 years. During this process the brain becomes more energy efficient, strengthening the connections that are needed while ridding itself (and the metabolic demand) of those that are not. Just which synapses are deemed necessary is a direct reflection of one's environment, a nature–nurture interaction which is fundamental to brain development.

To communicate with each other, the nerve cells need receptors and neurotransmitters. Serotonin appears first, followed by norepinephrine and dopamine, with acetylcholine and GABA produced later in fetal life.

Finally, the neurons become myelinated from birth to adulthood, a

process of insulation which increases transmission speed. In general, *myelination* occurs first in the spinal cord and lower brain, with the cerebellum fully myelinated by age 4, the corpus callosum at 8 to 12 years of age, and the cortex from age 10 into young adulthood.

The manner in which these brain changes occur is a result of complex interactions between genes and environment. Biologically, genes initiate the process of brain development through organizer genes and homeoboxes such as *Sonic hedgehog, Delta,* and the *Hox* genes among many others. The endocrine system is also involved with insulin-like growth factor (IGF) and thyroid hormones playing roles. Environmentally, illnesses and toxic exposures can alter development. Diet and nutrition is thought to have an effect, as industrialized nations demonstrate faster growth than do developing ones. For example, an increase in docosahexaenoic acid may be linked to greater visual acuity. By kindergarten and even earlier, some children from disadvantaged backgrounds have been shown to lag behind in language and vocabulary and may have heard 30 million fewer words in the home. Neglect in infancy has been linked to a significantly reduced midsagittal area of the corpus callosum. Clearly, genes alone do not account for brain development.

Genetic programing allows for brain development so that the infant can learn new skills. These capabilities facilitate interactions with the environment, which then precipitates further genetic modification. For example, motor skills afford one the chance to throw a ball and learn what happens to it. Repetition stimulates the same neural pathways, strengthening the synapses. Neurotrophins and prostaglandins affect synaptogenesis and long-term potentiation, a process that underlies learning and memory. Underuse of a neuronal pathway due to lack of exposure leads to its pruning and atrophy, mediated by genes. In short, genes lead to abilities which feedback and control other genes, affecting synapses and pathways. Thus, both the skills governed by genetics as well as the environment in which they evolve are integral to development.

LANGUAGE DEVELOPMENT

Language is so essential and omnipresent in everyday life, it is hard to imagine that it does not exist at birth. Though the ability to *communicate* is present from day one, language development is a process that takes years to develop. Speech delay is a common cause for clinical presentation, and language delays may be associated with later psychiatric comorbidities. Moreover, an undiagnosed speech delay is present in patients presenting for other reasons. Therefore, an understanding of normal language development is indispensable for all providers.

The development of language begins in utero, where a fetus has been shown to recognize its mother's voice. On delivery, an infant clearly lacks the ability to speak, but yet still requires a way to demonstrate needs. This

is achieved through facial expressions such as cries. An attuned parent can often tell what a baby needs (such as a feeding or diaper change) based on the particular type of cry. Though the exact timing of each step may differ, the overall process is quite constant. At about 3 to 4 months, typical infants make cooing (vowel) sounds followed by single syllable consonant babbling. These are learned by imitating others, and carry no meaning at this age. Repetitive vowel-consonant sounds (such as mama) may appear, but without meaning these are not considered first words. Meaningful words are produced at about 12 months, with single words intended to convey entire phrases. It is not unusual for words to come and go at this stage, and loss of a spoken word should not be cause for alarm. At 18 months, word combinations appear and about 25 percent of words are intelligible to strangers. At 24 months there are about 300 words, 50 percent of which are intelligible. By 30 months there are 500 words. At 36 months, 75 percent are intelligible to strangers.

Grammar develops significantly from months 20 to 36. As children learn the rules of speech, incorrect verb tenses are common ("I goed to school"). The same child may have used this correctly in the past, which likely reflected an imitative phrase rather than a novel one, and is considered normal. By age 4 most expressive language is developed and appropriately used. Nearly 100 percent of words should be intelligible to strangers, even though pronunciation mistakes may be present. Over the next 2 years children are able to connect sentences with one another logically. At ages 5 to 6, language centers slow their development, highlighting the importance of early language and the challenges of learning a second language later in life.

From a theoretical perspective Noam Chomsky emphasized the universality of linguistic structure and that all humans develop the ability to understand and produce language. This stands in contrast to a behavioral theory of language, championed by B. F. Skinner. Skinner believed that children imitate others and have their language reinforced through the environment. Factors such as large family size and lower social class correlate with later development and fewer words. Expanding on these environmental variables, the quality of verbal stimulation plays an important role. Hearing words from reading is more effective than just hearing spoken words; significantly more "unique words" are perceived through books, likely exposing children to words they would not otherwise encounter. Recent fMRI studies show that children with a higher reading exposure had greater activation in the left parietal-temporal-occipital association cortex, integrating the auditory with the visual and potentially stimulating imagery and narrative comprehension.

Clinical Correlate

A 24-month-old boy is referred from the pediatrician for severe behavioral outbursts. Upon

taking a history, you discover that he only knows about 25 words and cannot make word combinations. You astutely hypothesize that the tantrums are the manner in which this child is expressing himself, and aim to identify the reason for the speech delay. Your differential includes hearing and vision difficulties, and vision and hearing tests are ordered. Deficits in other areas are explored for signs of a broader developmental delay including Intellectual Disability and Autism Spectrum Disorder. Environmental contributors such as neglect or trauma must be ruled out. An auditory processing disorder is also a possibility. Finally, you conclude the delay is due to a speech or language learning disorder. A speech and language referral helps identify the nature of the delay and breaks speech down into further subcomponents: articulation (pronouncing sounds), syntax (the arrangement of words in a sentence), semantics (meaning of words), and pragmatics (appropriate social use of words). With speech therapy, the child makes linguistic gains and his behaviors subsequently improve as he now has other means to regulate and express himself.

COGNITIVE DEVELOPMENT

Cognition can be conceived of as what happens to sensory information once it reaches an individual. How this occurs is substantially more detailed and involves perception, understanding, memory, learning, reasoning, problem solving, and language. A discussion on cognitive development would be incomplete without mentioning *Jean Piaget*. He proposed four stages of cognitive development, with changes occurring through *assimilation* and *accommodation*. Assimilation, he argued, is taking in new information and fitting it in with an existing understanding, or *schema*. It is fitting new data into old ideas. This contrasts with accommodation, in which the new experience or sensory data causes the old schema to be altered so that it is in accord with this new information. For example, a child has an understanding of a tree, which is comprised of a trunk and branches. In the summer when leaves appear, the basic concept of tree remains the same but it has expanded to include this new knowledge. Later on the child encounters a bush and calls this a tree as well. The child is taught the differences between a bush and a tree and a new bush schema is formed. Though parts of Piaget's theories have come under question, he added tremendous insight into how children think differently from adults. This has consequences for everyone interacting with children, and is especially noteworthy for those in the educational and health care systems. Given his significance to the field, this section will be structured around Piagetian stages.

The first of Piaget's stages is *sensorimotor,* occurring between birth and approximately age 2. As the name implies, at this stage the infant is mostly limited to sensory information and simple motor responses. Within 2 months the infant can pay attention to faces and even recognize them. They can follow things with their eyes. By 4 months there is eye–hand coordination, as the infant can reach for something he sees. Recognizing familiar faces and important objects improves and extends to greater distances. At 6 months they show curiosity for things out of reach, and once within reach, they tend to put them in their mouths. Eye tracking further develops over the next few months, as the infant can even follow the path of

something as it falls. Picked up objects can be transferred from hand to hand smoothly. At 9 months, peek-a-boo is a particularly fun game, notably because at this age the infant is grappling with *object permanence*. To an infant lacking this skill, anything that cannot be perceived cannot exist. By the age of 1, there are clear signs that schemata are forming and are being used. A 12-month-old typically can follow directions and use things correctly, such as a cup. These schemata strengthen and broaden over the next 6 months and at 18 months they often clearly show knowledge of what a typical item is used for, such as a brush. As schemata are forming, a sense of humor can be seen when these expectations are playfully breached. Following simple instructions such as "sit down" implies an ability to understand language and know what to do with it. This implicates gains in memory and learning. By the age of 2 the individual can find hidden items, clearly demonstrating the achievement of object permanence. Make-believe games begin to take on a pleasurable role. Rudimentary organizational skills are demonstrated by sorting shapes and colors. Two-step instructions can be followed and four-block towers can be built.

The second stage Piaget described is *preoperational,* ranging from about 2 to 7 years of age. Symbolic play is a key feature of this time period. Pretend play and make believe, role playing, and symbolization (such as a stick representing an airplane) appear. A hallmark of this stage according to Piaget is that toddlers exhibit *egocentrism,* meaning they lack the ability to take another's point of view. Children in this stage lack the understanding of *conservation*. The classic example is if equal amounts of water are poured into a tall skinny glass and a short wide glass, those who have not mastered conservation yet will say the tall skinny glass has more water in it. This skill develops from ages 6 to 11 and may differ for various units of measurement.

Clinical Correlate

With egocentrism, children cannot take on another's point of view and are susceptible to misattribution. This normal stage in development leaves one vulnerable should a major stressor occur. For example, loss of a parent, divorce, or terrorist attacks demand explanation, and typical children with egocentrism are likely to blame themselves for the event. Though with age they may cognitively realize the situation was not their fault, the initial memory and experience was recorded in that way and guilt and self-blame may be present. Whether treating a child in real-time or an adult decades later, recognizing where the patient was developmentally at the time of tragedy is paramount for processing.

At age 3, children can copy a circle, complete three-to-four piece puzzles, and build a six-block tower. By age 4, they can understand the idea of same and different. They can start to understand time, begin to understand numbers and counting, draw two-to-four limbed bodies, name colors, and copy a square. At 5, they can count to 10, draw six-part bodies, copy a triangle, and start printing some letters.

The third Piagetian stage is *concrete operations,* from roughly ages 7 to

11. Order, rules, and organization are hallmarks of this time period. As the name implies, children at this age have difficulty with abstract and hypothetical thinking but are learning how to logically think about concrete things. Children implement inductive reasoning, generalizing from the specific (an example would be dad drives a red car, therefore all dads drive red cars). Deductive reasoning remains a challenge, requiring the child to understand something specific based on a generalized rule. Two other important cognitive skills emerge during this stage: *conservation and reversibility*. Children who have mastered conservation now realize that, as in the example above, there is the same amount of water in both glasses regardless of height. Additionally, children in this age group develop the capacity to see things from someone else's point of view, marking the end of egocentrism. This contrasts somewhat with theory of mind research. *Theory of mind* is the ability to recognize someone else as another entity with his/her own thoughts and feelings and emerges between ages 4 and 5. Learning *that* someone else thinks differently than oneself is one thing, anticipating *what* that person may be thinking or feeling is more complex and develops later.

This ability and others further develop in Piaget's final stage *formal operations* from age 12 and older. Here deductive reasoning develops, with many math and science implications. The ability to think abstractly appears, allowing for better planning and organization and problem-solving. Trial and error is no longer necessary to come to a conclusion, results can be anticipated. Though cognitive milestones have been discussed in the context of Piagetian stages, it is worth noting other developmental theories which help predict how and why children acquire the aforementioned skills. *Information Processing Theory* holds that children are goal-directed problem solvers who work to achieve certain ends. Improvements in memory, processing speed, planning, and logic help facilitate learning and solving which leads to more efficient approaches and greater knowledge. Change occurs continuously, gradually, and more subtly than in a stage-based paradigm as Piaget proposed. Similarly, *Lev Vygotsky* postulated that children continuously grow as social creatures heavily influenced by their social and cultural environments. The *dynamic systems theories* attempt to incorporate many factors viewing children as constantly evolving products of their past and present adapting to the context of their physical abilities and social environment.

Sleep wields tremendous influence on cognitive and physical development and has its own unique process. At birth, infants sleep 16 to 18 hours per day, including nighttime sleep and their two or more daytime naps. By age 1 there is typically just one nap, and daily totals are down to 15 hours. This is further reduced to 13 to 14 hours by age 2 and 12 hours by age 4, with naps typically ceasing. Common adult patterns of 7 to 10 hours are seen by ages 10 to 14. On polysomnography, rapid eye movement

(REM) sleep dominates the first year. By ages 3 to 4, deep non-REM sleep achieves a more adult-like share. It is worth noting that sleep can be highly individualized and influenced by culture. The importance of sleep cannot be overstated, with impacts on memory, learning, and physical growth. Recent evidence has shed light on the glymphatic system, which eliminates neurotoxic waste (proteins including beta-amyloid and metabolites) from the CNS during sleep. Finally, some sleep phenomena which in and of themselves are not pathological, including nightmares, are often reported, typically at age 5. Sleep terrors (also called night terrors) and sleepwalking occur in NREM stage 3 and are not uncommon in healthy children. Other normal sleep phenomena include hallucinations while falling asleep (hypnagogic hallucinations) and on waking (hypnopompic).

SOCIAL DEVELOPMENT

Social maturation is traditionally divided into four phases beginning with the attachment period and ending with late adolescence. There are many theories that attempt to explain social development (see Table 35.2–1).

Attachment theories explain the development of the sense of self. "Attachment is a deep and enduring emotional bond that connects one person to another across time and space" (Ainsworth, 1973; Bowlby, 1969). Infants have their primary needs met by a caretaker, typically a parent, and the emotional bond that forms between them becomes the foundation for the child's social, emotional and cognitive development. This *affective bond* provides the context for the "emerging pattern of dyadic regulation of emotion" (Stroufe, 2000).

The infant's behavior directed toward the caregiver becomes organized based on needs; thus, the quality of this attachment impacts their survival both biologically and psychologically. A healthy attachment supports the infant's exploration and mastery of the world. There are phases in the development of attachment that begin with *pre-attachment* at ages 0 to 6 weeks. From 6 weeks to 8 months, attachment progresses based on the responsiveness of the caretaker, and from 8 months to 2 years *reciprocal* relationships begin and the child organizes around their acquisition of mobility and the "secure base experience."

The historical leaders in attachment theory include John Bowlby, Harry Harlow, Rene Spitz, Mary Ainsworth, and Mary Main. Disorders of attachment are included in the *Diagnostic and Statistical Manual for Mental Disorders,* 5th Edition and their criteria will be reviewed in the chapter on abnormal child development.

Bowlby developed his theory by borrowing from evolutionary thinking and derived from the idea that the attachment of mother and infant was "due to a species-typical behavioral pattern to promote infant survival." He also incorporated ideas from Piagetian cognitive theory and the concept that "with increasing schema comes increasing complexity of thought" and

applied them to the intensifying connection of the infant's attachment behaviors.

Two researchers in the 1940s and 1950s provided the seeds for the idea that an infant has more than basic physical needs. Psychologically, the infant needs and craves human contact. This was demonstrated in the Harlow studies of rhesus monkeys preference for cloth monkeys and in Spitz's *hospitalism* research which confirmed a failure to thrive in infants who had their biological needs met but did not have a primary attachment figure with whom they had physical contact. Spitz described the concept of *anaclitic depression*, where separation from the caregiver leads to depression, withdrawal and loss of milestones.

Ainsworth's *Strange Situation Experiment* identified four types of infant attachment:

Avoidant—needs frequently not being met by the caretaker
Securely Attached—needs were generally met by the caretaker
Ambivalent—needs inconsistently met
Disorganized-Disoriented—correlated with a history of maltreatment.

Margaret Mahler's more complex theory proposes a sequence of attachment which, if successful, enables an individual to establish a secure sense of self. An infant who is securely attached and has established *basic trust* also experiences appropriate separation anxiety when away from the primary attachment figure.

Mary Main has developed an attachment interview for adults to assess the quality of their childhood attachments. She also described a new attachment classification to apply to Ainsworth's Strange Situation test, the disorganized/disoriented attachment classification.

Advances in biotechnology support the possibility that attachment begins in utero. Thomas Verny, M.D., in his book *The Secret Life of the Unborn Child* (1981), began the discussion of this and the idea of the "attachment pregnancy," and guides for parents to foster healthy in utero attachments have been popular recently.

The changing dynamics of gender roles for parents as well as the changing identity of the nuclear family have led to new research which has generally supported the concept that there are no great differences in quality of attachments in lesbian, homosexual, or transgender families with their children.

Successful secure attachments in infancy correspond to later psychosocial functioning and are characterized by better peer relationships, more cooperation with parents and teachers, greater security in unfamiliar social situations and in the exploration of new environments. A child's sense of self is generally more secure with positive early attachments, which are the foundation of emotional development. There are many consequences of abnormal attachments that will be discussed in the

sections on psychopathology including feeding problems, disruptive behavior disorders, and impaired ability to sustain lasting relationships.

Morality

There are societal and cultural norms that guide one's behaviors and actions. The internalization of these standards is a process that begins at birth when infants become aware that their behavior elicits a corresponding behavior in their caretaker, such as when a parent hugs a smiling happy infant. This cause-and-effect relationship is observed by infants and young children and becomes the basis for the rules that developing children assimilate and accommodate into their moral foundation. Children feel sadness and shame when they are punished or see that their behaviors bring displeasure to their parents or other important adult figures. Therefore behavior is shaped by the fear of punishment, and life can be explained and organized within the framework of cultural rules and a sense of fair play. Children tend to go from the fear of punishment to the desire to confess. As children enter adolescence and reach puberty, they develop guilt and the ability to be critical of their own behaviors and transgressions. In many cultures this is the age when children's lives become their responsibility and their conscience must be their guide.

There are many different theories of moral development but the most famous is that of *Lawrence Kohlberg,* which built on the foundations of Piaget's theory of cognitive development. Kohlberg's work was criticized for being too culturally biased toward the middle class male and, despite revisions over the decades, is still controversial. *Carol Gilligan* in 1982 offered an alternative theory, which corresponds to morality growing in parallel to empathy. Her theory had a gender basis, with males' moral development developing into a system dominated by "the ethics of justice and the assertion of rights" whereas "woman's ends in compassion and an ethic of caring."

EMOTIONAL DEVELOPMENT

The ability to acquire and express emotions and feelings is essential to adult functioning. Learning to identify, understand, express, communicate, and modulate affect is key to successful relationships. An infant's task is to learn to be emotionally responsive and not react to internal feelings and anxiety. This is achieved through the maturation and the development of coping skills. There are many theorists who have attempted to describe emotional maturation, with achievements highlighted in Table 35.2–1.

More recently, theories integrating the neurobiology of brain maturation with social interaction have focused on the interaction between genetics, environment, and behavioral patterns to explain emotional

development and maturation.

All of the theories to describe emotional maturation can basically be divided into four stages with basic milestones. Stage I is the acquisition of basic emotions that are learned through the interaction of the infant with others, including the use of mimicry. Mirror neurons, which are neurons that fire when one acts or observes an action, are proposed as being an essential pathway for this acquisition of a repertory of emotions by the infant. Stage II involves the development of self-conscious emotions, which are derived from approval or disapproval experienced and include shame, envy, embarrassment, pride, and guilt. The ability of emotional self-regulation is the task of Stage III, which continues into middle childhood. Stage IV is marked by true empathy, the ability to understand and share the feelings of another.

Temperament

Temperament has fascinated scientists since ancient times when Galen of Pergamum made his prophecy "that much of human behavior can be predicted according to an individual's temperament."

Stella Chess and Alexander Thomas made important contributions to the study of temperament in the 1970s. They defined three major clusters which are dependent on the perceived sense of ease with which children interact with their environment. The three different styles were assessed through the categorization of nine types of behavior that they defined as dimensions of temperament, including activity level, regularity/rhythmicity, approach or withdrawal, adaptability, threshold of responsiveness, intensity of reaction, quality of mood, distractibility, and attention span and persistence. Their longitudianl study sample showed the following distribution which was generalized to apply to the general population:

1 in 10 children had a *difficult* temperament
1 in 6 were categorized as *slow to warm up*
4 in 10 were characterized as having an *easy temperament*.

The *goodness of fit* between parent and child as described in their book, is an important factor in working with parent–child relational problems therapeutically. Jerome Kagan, inspired by Galen, also did work on temperament and described two main types of temperament: inhibited and uninhibited. He stated that "temperament refers to stable behavioral and emotional reactions that appear early and are influenced in part by genetic constitution." Behavioral geneticists are now exploring specific genetic polymorphisms that contribute to temperament. Some research cites polymorphisms in serotonin transporter and promoter region genes as having an effect on personality development. Additionally, polymorphisms in the dopamine, serotonin, and norepinephrine transporter genes affect

temperament.

MOTOR DEVELOPMENT

Perhaps more than any other developmental domain, gross and fine motor development is highly intertwined with progress in other areas. Motor development begins in utero, and is experienced by a pregnant mother as kicking. At birth, a premature CNS renders infants with little control over their actions. At this stage reflexes are prominent, such as the grasping, rooting, sucking, swallowing, startle, and Moro. These reflect CNS functioning and disappear at predictable time periods.

Cortical columns are being laid down during the first 6 to 7 months of life and synapses are exploding in number, allowing for greater complexity and control of actions. At 2 to 4 months of age, infants can support their own head and make smoother movements in their extremities, demonstrated by shaking their toys. Some infants may be able to roll over from their stomach to back, and by 6 months they can roll over in both directions. Also at 6 to 7 months, they can sit without support and may start to rock back and forth as if they are trying to crawl, which they will be doing over the next 2 to 4 months. A developing sensory system involving all modalities but particularly hearing, proprioception, and vision facilitate the emergence of new skills, and will continue to help them hone. Hand-to-hand transfer occurs at 7 to 8 months, first grossly then more finely, aided by a *pincer grasp*. This is when something is held between the thumb and index finger, first seen at about 9 months, and with practice develops strongly over the next year.

Clinical Correlate

Infants are unable to roll over from their stomachs to backs before 6 months, a primary reason the American Academy of Pediatrics has recommended babies be placed to sleep on their backs in order to reduce the chance of Sudden Infant Death Syndrome (SIDS).

A 9- to 10-month-old infant can stand with support, and may start walking while holding onto furniture, a practice called *cruising*. Walking typically is reported at 12 months with a normal range of 9 to 15 months. Highlighting the interdependent nature of development, the newfound ability to walk corresponds with Margaret Mahler's separation-individuation subphases of practicing and rapprochement.

By 18 months the now toddlers can drink from a cup, eat with a spoon, and help dress themselves. Athletic abilities emerge at age 2, with running, kicking, throwing, jumping, climbing, and copying straight lines. Simultaneous cognitive development is crucial as these skills require a coordinated effort involving attention and planning. A 3-year-old's athletic skills become more refined and the child can climb stairs and show off the skill of riding a tricycle. By the age of 4, movements continue to improve,

with catching skills and hopping abilities. Visuomotor coordination is needed and spurs development through activities such as sports and art, the latter aided by an ability to hold a pencil appropriately. These actions strengthen synapses in the developing cerebellum, which is fully myelinated by age 4, and in turn facilitates further coordination. This level of dexterity allows pouring and simple food cutting behavior to occur, which is much improved by age 5. At this time the child can engage in basic gymnastics such as somersaulting and can hop and begin to skip. Hand preference is typically established by age 5. The ability to tie shoelaces generally appears within the next year or 2.

Toilet training is influenced by culture, but myelination of pyramidal tracts allows for sphincter control by 18 months and children can start toilet training as early as age 2. They typically develop bowel control before bladder, and the vast majority of Western children are toilet trained by ages 3 to 4. Nighttime bladder control is typically achieved months to years later between ages 5 and 7.

As with all development, it is imperative to keep in mind that the listed ages should be considered normal ranges. There is a high degree of variability, and the amount of delay as well as pervasiveness of delays should inform clinical suspicion of abnormality. Typically a 3- to 5-month delay in the early milestones (sitting without support, standing holding on, first steps) is granted before there is concern, and a larger 10- to 15-month window for later complex achievements (such as climbing stairs, jumping). Family history is important, as gross motor development is considered to be highly influenced by genes. Environment and stimulation is of course important too, particularly in fine motor development such as utensil or pencil use and in sharpening gross motor sporting skills. Inquiring whether a skill has been lost or never developed is also a crucial piece of the clinical history.

SENSORY DEVELOPMENT

Imagine walking on a crowded city street at rush hour: neon lights flashing, sirens wailing, mysterious scents wafting, fellow pedestrians obliviously making contact. In a mere instant, the body is flooded with a barrage of stimuli. That the mind can perceive, organize, and make sense of it all to the point where such a moment would be considered unremarkable, is in itself quite remarkable.

Infants are born with all of their primary senses working in some capacity, some more developed than others. Within the first year, however, there is tremendous growth across the sensory modalities. As with other domains, there is a split into how important nature or nurture is in sensory development. For senses fairly functional at birth, genetics are playing a large but not exclusive role. For those areas that require more development after birth, any predisposition becomes dependent as well as modifiable via

environmental factors.

Hearing is functional and fairly developed by birth, and in fact there is evidence that even in the womb infants can distinguish their mother's voice. Despite this functionality, there is evidence that the softest sound an adult can hear would need to be amplified four times for an infant to perceive it. Children eventually reach adult levels of acuity as early as age 2. At birth, infants will turn their heads toward a sound in a phenomenon known as *auditory localization* and demonstrate a preference for voices. That infants can recognize music should come as no surprise, given the melodic nature of lullabies. Infants can discern consonant tones from dissonant ones, and can be seen bouncing up and down to the music's beat.

Given how much humans rely on vision, it is surprising that this sense is the least developed at birth. In the early months, infants require great contrasts in color to make distinctions. Their color-sensing retinal cones are immature and inefficient at capturing light. Their visual acuity is quite poor, with vision estimates ranging from 20/120 to 20/800 at birth. There is an explosion of improvements made as the cells that will make up the visual cortex proliferate, reach their destination, and then make and break connections. During the first 6 months, 18,000 new visual synapses have the potential to form. Infants have near-adult levels of visual acuity by 8 months, and by age 6 their vision is full.

A number of visual experiences appear to be present at birth. Research has shown that infants possess *perceptual constancy,* where an object is known to be a constant size or shape. Infants also possess the depth cue known as *optic expansion,* where an object getting closer is perceived to be approaching. They can also differentiate human faces. Eye tracking begins at 2 months, and infants are drawn to moving stimuli. Along with this newfound skill comes the ability to take in more of an object and process it. This phenomenon is particularly noteworthy when it comes to looking at faces, as 2-month-olds will scan a face much more comprehensively, and will spend significantly more time looking at eyes and mouths. This has implications for social, emotional, and language development.

At 3 to 4 months, depth perception is improved via *stereopsis,* where the quickly maturing visual cortex uses the slightly different images perceived by each eye to calculate depth. At 6 to 7 months, infants can demonstrate monocular depth, perceiving depth with one eye.

Infants (much like adults) have a preference for sweets, beginning in utero. An ingenious study by de Snoo (1937) demonstrated that when saccharin was injected into amniotic fluid, the fetus would drink more of it and excrete it via their mothers, helping to decrease high amniotic fluid volumes. Odor perception is inferred to take place in the womb as well, based on the fact that amniotic fluid has discernible scents and that this fluid comes into contact with olfactory receptors. After birth, infants show clear preferences for their mother's scent.

Touch is inferred to begin in utero, when a fetus may be seen grasping the umbilical cord, touching other parts of its developing body, and even sucking its thumb. Learning one's orientation in space helps developing proprioceptive skills. Initially after birth infants will explore new objects orally, putting everything in their mouths. They eventually develop stronger motor and visual skills and learn their environment through touch.

The human mind is able to integrate stimuli from various senses (and even multiple stimuli of just one sense) into a cohesive explanation through what is called *intermodal perception*. This ability allows one to successfully navigate the busy city street, all while listening to music.

GENDER IDENTITY

There are multiple theories of gender identity development, including biological, cognitive, developmental, and social learning theory. There are differences between gender identity, sexual identity, and sexual orientation. Gender development begins in utero. During the first trimester, fetuses with a Y chromosome have hormonal masculinization produced by the release of androgens, and feminine hormones and development are inhibited. The fetus with only X chromosomes develops ovaries at the end of this trimester. This is essentially how infants are born, with most of the development of secondary sex characteristics beginning in latency for girls and at preadolescence for boys.

A person's *gender identity* can be the same or different from their sexual assignment at birth.

The concept of gender identity is defined as the person's personal experience of whether they are male or female. A person comfortable with their gender assignment at birth is termed *cisgender*. People with gender identity disparity and discomfort can be gender non-conforming or *transgender,* and some societies define a third or other gender. Societal gender standards are typically binary and impact a developing child's social identity.

Most theories and research support that core gender identity is formed by age 3. Gender dysphoria is a marked incongruence between one's experienced or expressed gender and assigned gender. Gender dysphoria can be caused by attempting to reassign a child's gender and can lead to serious consequences, including depression and severe anxiety. It is thought that, between the ages of 6 and 8, gender identity is coalesced. However, recent evidence suggests that a percentage of people may later shift identities after puberty. Due to the historical changes in Western societal attitudes toward sexuality, research is being done on gender identity development and disorders as well as parenting styles such as genderneutral parenting that are helpful for the gender nonconforming child.

The *S.A.G.E. Test,* the sex and gender explorer is one of many new tests developed to assess gender identity.

Gender roles and gender relationships in society have changed drastically within the last century, as have sexuality and the acceptance of nonbinary sexuality; as social identities change with this new acceptance of different gender and sexual roles, so will parenting and child-rearing practices. The days of defining gender and sex roles in infant toys, clothes, and play are changing. The United States military now allows women in combat and men are staying home to care for children. Mental health professionals need to assist in maximizing a child's development by supporting an understanding and acceptance of new forms of family, gender, and sexual identities.

PSYCHOSEXUAL DEVELOPMENT

Sigmund Freud's contributions to psychiatry and understanding of the mind are the subjects of numerous textbooks, and even institutions in and of themselves. As this chapter is dedicated to development, this section will solely focus on a portion of Freud's theory, that of psychosexual development and its contributions to character formation, from his 1905 piece *Three Essays on the Theory of Sexuality*.

Freud postulated that children go through successive stages in childhood: oral, anal, phallic, latency, and genital. Their names correspond to the erogenous zone from which Freud believed *psychic energies* attach to (or cathect) at different times of life. Psychic energies are instinctual drives which motivate the individual, and at the time this theory was proposed that drive was sexual (aggressive motives would later be emphasized as well).

Unresolved conflicts can arise at any time during any phase, leading to a variety of personality types. In the *oral phase* (from birth until about age 1), an infant's main desire according to Freud is in breastfeeding (or bottle feeding—a warm caregiver is more important than the mode of feeding). Toward the end of the first year comes the emergence of teething and biting, aggressive tools. Undergratified suckling needs (such as neglect or emotional indifference) can lead to jealousy, envy, and a seeking and consuming nature. The typical example is the oral fixation, someone constantly chewing or smoking. A clinging dependency may also be seen and manifested by being unwilling or afraid to let go of someone or something for fear there will not be more. On the contrary, one whose needs have always been met may expect things will always work out and maintains an optimistic nature. This belief may lead to passivity and inactivity, on the other hand, and the potential for fewer coping skills when a need is not met. An infant's sexual (and aggressive) drives can be met with numerous responses by his caretaker, which can lead to a fear of losing the caretaker. To placate these fears, the baby will implement a

defense such as denial of aggression, which, while adaptive in infancy, can be destructive later in life.

In the *anal stage* (ages 1 to 3), the psychic energy has shifted and control becomes a central theme: control over one's autonomy, control over one's impulses, control over one's body and sphincters. Parental demands increase and compliance and rebellion with these requests are at odds here, contributing to the "terrible twos." Whereas the oral stage is a symbiotic one between the caregiver and the infant, the newly developed abilities to walk and talk make the anal stage marked by separateness, or the *dyad*. Frustration and aggression abound as the toddler is struggling to meet expectations and master control. Tolerating ambivalence toward others, learning to titrate aggression, and accepting that caregivers will not leave or withhold their love is crucial toward a successful resolution. It is in this stage where the anal-compulsive character originates, with the Freudian hallmark triad of pedantry, parsimony, and miserliness. A love of order and cleanliness embodies this person, who translates into the obsessive-compulsive personality type. The opposite is one who lives in disordered chaos.

The phallic phase (ages 3 to 6) is marked by the child's sexual desires for his/her opposite-sex parent. An obstacle to this achievement is the same-sex parent, who elicits murderous wishes from the child. In boys this is known as the *Oedipus complex,* in girls *Electra*. These forbidden thoughts evoke fears of bodily injury, notably castration anxiety in boys. Fear of injury is indeed typically seen during this time. These fears as well as impractical realities eventually force the child to give up his quest and alternatively identify with the same-sex parent. In doing so, he internalizes his parents' rules and mores, ending the phallic phase with the formation of a superego. There are now three entities involved (child plus mom and dad), and conflicts in this phase lead to issues in triadic relationships. Unresolved conflicts in this stage may lead to myriad traits, from self-evident Oedipal ones to difficulties with following rules.

Latency follows from resolution of the Oedipal complex until puberty. Freud believed this time was one of relative calmness sexually, though not entirely static. Psychic energy is invested in academic and social avenues as the child goes through grade school. This phase is dominated by school and peers, learning socialization and rules. Values and ideals are solidified as the newly formed superego evolves from an external presence to an internalized one.

Though outside the scope of child development, following latency, Freud described the genital phase, where the genitals (as in the phallic stage) again become the source of libidinal energy, but now the pubertal adolescent has biological potency. This newfound ability confers its own set of conflicts, and the teenager moves from within the family to outside the family.

It is important to bear in mind that these character types are simplifications. In addition, this theory is psychological and does not account for innate tendencies which were indeed noted by analysts as well. Finally, when conflicts occur they are from repeated behavior patterns between a child and the important caretakers. Though many are skeptical about the veracity of Freud's theories, his contributions are irrefutable. The impact of early experiences remains crucial to understanding an individual, as does the relationship to family members. The significance of an unconscious, that feelings occur outside of one's awareness, has enormous bearing on how one conceptualizes the mind and treats patients clinically.

CONCLUSION

In summary, development is the acquisition of new skills or abilities and the various responses both intra- and inter-personally. Humans are born with a genetic background, and the environment contributes to which pathways get further strengthened or pruned. Changes in the brain allow for new possibilities in the developing child. Small advances begin to accumulate and more possibilities arise. By the end of childhood, a person who started as mostly innate reflexes has become capable of writing essays about the future, performing complex athletic activities, establishing a sense of identity, and engaging in deep reciprocal relationships. However, although this section has focused on child development and all the significant achievements that are made during that period, it is important to note that development continues throughout the lifespan.

Acknowledgments

The authors would like to acknowledge Maureen Fulchiero Gordon, M.D., who authored the last edition of this chapter, and Jean Thomas, M.D., for their contributions to this chapter.

REFERENCES

Adler SA, Gerhardstein P, Rovee-Collier C. Levels-of-processing effects in infant memory. *Child Dev*. 1998;69:280–294.

Arbel S, Benjamin J, Golin M, Kremer I, Belmaker RH. Relation of shyness in grade school children to the genotype for the long form of the serotonin transporter promoter region polymorphism. *Am J Psychiatry*. 2003;160:671–676.

Bangoo RK, Leibnlluft E. Affective neurons and the state of normal and abnormal emotional regulation. *Child Adolesc Psychiatr Clin North Am*. 2002;11(3):19.

Berenson S. Effects on gender identity of prenatal androgens and genital appearance: evidence from girls with congenital adrenal hypertrophy. *J Clin Endocrinol Metab*. 2003;88(3):1102–1106.

*Brainard MS, Knudsen EI. Experience affects brain development. *Am J Psychiatry*. 1998;155:1000.

Cassidy J, Jones JD, Shaver PR. Contributions of attachment theory and research: a framework for future research, translation, and policy. *Dev Psychopathol*. 2013;25(4 Pt 2):1415–1434.

Coley RL. Children's socialization experiences and functioning in single-mother households: the importance of fathers and other men. *Child Dev*. 1998;69:219–230.

Courage ML, McCloy UR, Herzberg GR, Andrews W, Simmons BS. Visual acuity development and fatty acid composition of erythrocytes in full-term infants fed breast milk, commercial formula, or evaporated milk. *J Dev Behav Pediatr*. 1998;19:9–17.

Demiris Y, Meltzoff A. The robot in the crib: a developmental analysis of imitation skills in infants and robots. *Infant Child Dev*. 2008;17:43–53.

Eisenberg N, Fabes RA, Shepard S, Guthrie IK, Murphy BC. Parental reaction to children's negative emotions: longitudinal relation to quality of children's social functioning. *Child Dev*. 1999;70:513–534.

*Fisher SE, Lai CS, Monaco AP. Deciphering the genetic basis of speech and language. *Annu Rev Neurosci*. 2003;26:57–80.

Galliese V, Keysers C, Rizzolatti G. A unifying view of the basis of social cognition. *Trends Cogn Sci*. 2004;8:396–403.

Ghuman JK, Freund L, Reiss A, Serwint J, Folstein S. Early detection of social interactions problems: development of a social interaction instrument in young children. *J Dev Behav Pediatr*. 1998;19:411–419.

Giedd J. Brain development, IX: human brain growth. *Am J Psychiatry*. 1999;156(3):4.

Goodman JC, McDonough L, Brown NB. The role of semantic context and memory in the acquisition of novel nouns. *Child Dev*. 1998;69:1330–1344.

Huffman LC, Bryan YE, del Carmen R, Pedersen FA, Doussard-Roosevelt JA. Infant temperament and cardiac vagal tone: assessments at twelve weeks of age. *Child Dev*. 1998;69:624–635.

Kaufman AS. Generics of childhood disorders II: reports I and II. Genetics and intelligence. *J Am Acad Child Adolesc Psychiatry*. 1999;38:626–628.

Merriman J. Linking risk-taking behavior and peer influence in adolescence. *Neuropsychiatry Reviews*. 2008;9(1).

Muris P, Cor M, Merckelbach H, Sermon A, Zwakhalen S. Worry in normal children. *J Am Acad Child Adolesc Psychiatry*. 1998;37:703–710.

Newbury DF, Monaco AP. Talking genes—the molecular basis of language impairment. *Biologist*. 2002;49:255–260.

Pleak RR. Gender-variant children and transgender adolescents. *Child Adolesc Clin N Am*. 2011;20(4):xv–xx.

Plomin R, DeVries JC. The genetics of cognitive abilities and disabilities. *Sci Am*. 1998;278:62–69.

*Plomin R, Rutter M. Child development, molecular genetics, and what to do with genes once they are found. *Child Dev*. 1998;69:1223–1242.

*Rizzolatli G. The mirror neuron system and its function in humans. *Anat Embryol*. 2008;210:419–421.

*Robinson JL, Zahn-Waxler C, Emde RN. Patterns of development in early empathic behavior: environmental and child constitutional influences. *Soc Dev*. 1994;3:125–146.

Rutter JJ, Baumann MH, Waterhouse BD. Systemically administered cocaine alters stimulus-evoked responses of thalamic somatosensory neurons to perithreshold vibrissae stimulation. *Brain Res*. 1998;798:7–17.

Saudino K. Behavioral genetics and child temperament. *J Dev Behav Pediatr*. 2005;26:214–223.

Schore AN. Effects of a secure attachment relationship on right brain development, affect regulation, and infant mental health. *Infant Mental Health J*. 2001;22:7–66.

Stillwell BM, Galvin MR, Kopta SM, Padgett RJ. Moral volition: the fifth and final domain leading to an integrated theory of conscience understanding. *J Am Acad Child Adolesc Psychiatry*. 1998;37:202–210.

Tangney JP, Miller RS, Flicker L, Barlow DH. Are shame, guilt and embarrassment distinct emotions? *J Pers Soc Psychol*. 1996;70:1256–1269.

Tettamanti M, Buccino G, Sacculman MC, Gallese V, Danna M. Listening to action-related sentences activates fronto-parietal motor circuits. *J Cogn Neurosci*. 2005;17:273–281.

Zanh-Waxler C, Radke-Yarrow M, Wagner E, Chapman M. Development of concern for others. *Dev Psychol*. 1992;28:126–136.

▲ 35.3 Adolescent Development

ARDIS C. MARTIN, M.D., AND CAROLY SUE PATAKI, M.D.

Adolescence is a critical period characterized by neurobiological and physical maturation leading to enhanced psychological awareness and higher-level social, cognitive, and emotional interactions with peers and adults. The most striking social and behavioral changes in adolescence include (1) increased affiliation with a peer group, (2) striving for more autonomy from parents, (3) increasing ability to "mentalize," that is, understand what others are thinking and feeling, (4) an increase in risk-taking behaviors, and (5) greater sensation seeking. A convergence of data from neuroimaging studies, including the work of Blackmore and Mills (2014) have been able to identify certain regions of the adolescent brain specialized to produce increased mentalizing, and other regions of the brain that remain immature, but are adaptive to changing environmental demands through their "plasticity." The varying levels of maturation in diverse regions of the adolescent brain leads to a seeming paradox, for example, between an adolescent's increased ability for empathic social thought process and at the same time reckless risk-taking behaviors. This apparent "mismatch," that is, varying degrees of maturity across different regions of the adolescent brain has been referred to by some researchers as the "Dual System Model." The work of Strang, Chein, and Steinberg (2013) supports a view that increased risk taking in mid-adolescence may be, in part, due to an increased reactivity of the social and emotional regions of the brain in the context of less mature cognitive control brain development. The social and emotional reactivity system is believed to be localized

mainly in the ventral striatum and ventromedial prefrontal cortex, whereas the cognitive control system is believed to be localized in the lateral prefrontal, parietal, and anterior cingulate cortices. Emerging evidence from neuroimaging studies supports a correlation between increased adolescent social awareness and mentalizing abilities and activity in three specific brain regions including the medial prefrontal cortex (mPFC), the posterior superior temporal sulcus (pSTS), and the temporoparietal junction (TPJ). However, as older adolescents transition into young adulthood, there is a decrease during mentalizing tasks in the activation of the mPFC. Thus it appears that during the development of mentalizing abilities, the mPFC is a critical brain region, however, once the process is developed, in adulthood, this region no longer predominates, and other brain regions may take over this function. Giedd (2014), and other colleagues, using neuroimaging to track adolescent brain development find that gray matter in the brain, such as the PFC, does not grow larger after puberty, on the contrary, it matures by becoming more specialized, that is, by a process termed "pruning," which ultimately leads to a diminished size. Gray matter is at its peak volume in girls at approximately age 11 years, and in boys, at approximately 13 years. On the other hand, white matter, representing myelinated brain regions continues to increase throughout childhood and adolescence. The continued increase in white matter reflects increasing "connectivity" and continued ability for the brain to adapt to new cognitive demands throughout adolescence. Thus, the increasing and less specialized white matter growth allows for greater cognitive adaptive brain function, with a resulting delay in its maturation in adolescents. Consequently, adolescents experience a longer period to learn and respond to novel cognitive information and develop new responses, however, this delays the brain maturation guiding cognitive control and decision making.

DEFINITION

Adolescence is the period of maturation between childhood and adulthood heralded by the physiological signs and surging hormones of puberty. It is a time when peer relationships deepen, autonomy grows, and intellectual pursuits intensify and a sense of social belonging is sought. Adolescence is a time of accelerated social and psychosexual exploration, culminating in an integrated concept of "self." Adolescents can be viewed as neurobiological "works in progress," often consumed by academic, interpersonal, and emotional challenges, and exploring new territories using their talents, and experimenting with social identities.

WHAT IS NORMAL ADOLESCENCE?

The conceptualization of adolescence as a continual struggle, with its defining feature intense psychic pain and severe turmoil, was a historical notion. The hypothesis that most adolescents experience enduring severe

psychological distress, intense depressive symptoms, severe mood lability, and a disruption of their usual conduct was not borne out by epidemiological studies. Although certain psychiatric disorders, such as major depression, schizophrenia, and bipolar disorders, emerge with greater frequency in adolescents than in younger children, the majority of adolescents are free of psychiatric disorders. For up to approximately 75 percent of youth, adolescence is a period of successful adaptation to physical, cognitive, and emotional changes, largely continuous with their previous functioning. Thus, the majority of adolescents display reasonable autonomy in decision making, increasing independence in completing academic work, and successful peer relationships without significant alienation from their families or the emergence of conduct-disordered behaviors. Psychological maladjustment, self-loathing, disturbance of conduct, substance abuse, affective disorders, and other impairing psychiatric disorders emerge in approximately 20 percent of the adolescent population. Adolescent adjustment is, in general, continuous with previous psychological function; thus, psychologically disturbed children are at greater risk for psychiatric disorders during adolescence. Adolescents with psychiatric disorders are at increased risk for greater conflicts with families and for feeling alienated from their families. Although up to 60 percent of adolescents endorse occasional distress or a psychiatric symptom, this group of adolescents functions well academically and with peers and describes themselves as generally satisfied with their lives. Studies conducted since the 1980s indicate that most adolescents endorse more acceptance than rejection of their parents' values and lifestyles, and few report feelings of isolation or alienation from their families.

The developmentalist Erik Erikson characterized the normative task of adolescence as *identity versus role confusion*. In other words, adolescents, faced with the task of applying childhood talents, strengths, and experiences to their evolving physical and emotional selves, must be able successfully to change "who they are" so as to build a more mature identity. The integration of past experiences with current changes takes place in what Erikson termed *ego identity*. Adolescents express various aspects of their psychological selves by becoming fans of heroes or other well-known musical or political idols. Some adolescents appear consumed by their identification with a particular idol, whereas others are more moderate in their expression. Adolescents who feel accepted by a peer group are less likely to become consumed by adoration of an idol. Adolescents who are socially isolated, feel socially rejected, and become overly identified with an idol to the exclusion of all other activities are at higher risk for serious emotional problems and require psychiatric evaluation.

Adolescence is a time to search for ideals and social values that can be carried into adulthood. Erickson used the term *moratorium* to describe that interim period between the concrete thinking of childhood and a more

evolved complex abstract thinking and ethical development. He defined *identity crisis* as a normative task of adolescence in which adolescents pursue alternative behaviors and styles and then successfully mold these different experiences into a solid identity. A failure to do so results in *identity diffusion*, or role confusion, in which there is a lack of cohesive or confident sense of identity. Adolescence is the time to bond with peers, experiment with new beliefs and styles, fall in love for the first time, figure out who one wants to be and where one fits in, and explore creative ideas for future plans. The literature leads to the conclusion that the majority of adolescents go through this developmental process with basic optimism and good self-esteem, maintain good peer relationships, and sustain basically harmonious relationships with their families.

STAGES OF ADOLESCENCE

Early Adolescence

Early adolescence, from 12 to 14 years of age, is the period in which the most striking changes are noticed—physically, attitudinally, and behaviorally. Growth spurts often begin in these years for boys, whereas girls may have had rapid growth for 1 to 2 years prior. At this stage, many adolescents express a sense of identification with a particular peer group, often a same-sex group of peers join together, and may identify belonging to a peer group with markers including clothing, hair styles or affiliation with certain activities. Both boys and girls begin to identify themselves as distinct members of their families, and often critique usual family habits, insist on spending time with peers with less supervision, and may question previously accepted family values.

In the early stage of adolescence, there may be some degree of withdrawal from family activities with, increased interest in spending time with peers, greater awareness of style and appearance, a stated goal of becoming more independent, and claims of needing less parental involvement. A new awareness of sexuality may be displayed by increased modesty and embarrassment with their physical development or may exhibit itself in an increased interest in the opposite sex.

Early adolescents engage in subtle or overt displays of their growing desire for autonomy, sometimes with challenging behaviors toward authority figures, including teachers and school administrators, and exhibit disdain for rules. At this age, some adolescents begin to experiment with cigarettes, alcohol, and marijuana. The Youth Risk Behavior Surveillance System (YRBSS) surveyed all public and private schools from grades 9 to 12 throughout the United States (more information is available at http://www.cdc.gov/yrbs). The 2011 report indicated that 37.6 percent of ninth graders had tried smoking cigarettes and 6.8 percent of male ninth graders had smoked at least one cigarette every day consecutively for 1

month. Cigarette smoking that included more than 10 cigarettes per day was reported by 5.7 percent of girls and 9.4 percent of boys in the ninth grade. The YRBSS reported that 30.2 percent of ninth-grade girls and 29.3 percent of ninth-grade boys consumed at least one alcoholic drink in the month preceding the survey. Over the last 5 years, lifetime alcohol, that is, at least trying alcohol at least once, was almost 62 percent in both boys and girls in the ninth grade. Heavy drinking, defined as at least five drinks of alcohol in a row, was present in 13 percent of ninth-grade girls and 15 percent of ninth-grade boys. Marijuana was reported by 26.4 percent of ninth-grade girls and 34.9 percent of ninth-grade boys at least once, and current use of marijuana was present among 18 percent of ninth graders. Overall, the prevalence of current and binge drinking among 9th to 12th graders has decreased between 2009 and 2011 while the prevalence of one or more time use of marijuana and daily current use has increased over the same time frame. With regard to other illicit drug use (cocaine, heroin, hallucinogens, and amphetamines) among 9th to 12th grades, only the prevalence of ecstasy use has increased between 2009 and 2011.

During early adolescence, there is significant variation in the timing of acquiring new behaviors. For example, some 12 year olds identify with popular music or wearing "adolescent" clothing, whereas other adolescents are ot seeking to reinvent themselves with a more mature appearing image. Epidemiological surveys have not provided evidence for this notion that adolescence is characterized by overwhelming alienation, profound angst, and significant disruption of family relationships. In this age group, most adolescents spend at least twice as much time with their families as with their peers.

Suzanne, a 14-year-old freshman in high school concerned about being compared to her older sister, who was known as a troublemaker. Suzanne felt that her teachers didn't value her for being her "own" person and automatically assumed that she was a poor student, always looking for a way to avoid schoolwork. Suzanne secretly smoked cigarettes, occasionally drank alcohol, and had tried smoking marijuana, however, she found it extremely annoying that her parents were always checking up on her. She rationalized that her parents' intrusion was just because her older sister had been suspended several times for going to school intoxicated. Suzanne spent most of her time out of school in her room with her door closed, doing homework while watching TV, instant messaging to her friends on her computer, and text messaging, snap chatting, and using Instagram with her friends. Although Suzanne had enough friends, and enjoyed singing in the school choir, she found herself feeling anxious much of the time and fearing that she had forgotten to finish all of her homework or studied enough for a test. She found going to class more and more unpleasant, and decided that classes were a waste of time. Suzanne began to find the school rules increasingly irritating, especially being "locked" out of the classroom if she was a few minutes late. Suzanne felt that the teachers were too strict, and never gave her the benefit of the doubt. To avoid getting into trouble for being late, Suzanne began to cut classes then she began to cut the whole day of school if she thought she was late in the morning. Before long, she had fallen behind in her work but was too embarrassed to tell her parents. She told her friends, and they helped her get some of the schoolwork, but she had missed so much work that it was difficult to catch up. Since it was getting close to her first progress report and Suzanne could not talk to her parents, she finally went to her school counselor to say that she was not going to many classes because it was too stressful to get there on time and that she was in danger of failing. The school counselor convinced Suzanne to tell her parents about her academic problems, and, begrudgingly, she agreed to a meeting with the counselor and her parents. Suzanne's parents

were shocked to find out that she was so stressed, and near failing, since she consistently said that everything was fine as long as everyone left her alone. After the meeting, Suzanne realized that instead of trying to be so independent she was actually being self-destructive. Suzanne agreed to accept her parents' help with a ride to school in the morning, and despite some initial resistance tutoring sessions were started. Suzanne worked hard with her tutor, improved her grades, and began to leave her bedroom door open more often at home. She also turned to parents and sisters for help if she did not understand something and found that they were actually helpful. Suzanne ended up with three A's and one B for the first grading period and was proud of herself. Suzanne felt less stressed and more confident after the first grading period, she was more connected to her family, and looked forward to spending more time with friends in the remainder of her freshman year.

Middle Adolescence

During the middle phase of adolescence (roughly between the ages of 14 and 16 years), peer groups begin to shift to coed, and become more predominant in their importance. Concrete signs of greater autonomy are beginning to emerge, such as learning to drive, and in some states, obtain a driver's license by the age of 16 years. Along with a greater lifestyle of autonomy adolescents are also presented with increased responsibilities, especially in the area of social judgment. Conflict may arise when empathy and social awareness are paired with poor impulse control and decision making, potentially leading to health risks, car accidents, and alcohol, and drug use. Sexual behavior intensifies in this phase of development, initiating romantic relationships are more complicated, and self-esteem becomes a pivotal influence on positive and negative risk-taking behaviors.

The 2011 YRBSS reported more frequent cigarette use among white students (8.6 percent) and Hispanic students (5.8 percent) than black students (3.4 percent). Use of alcohol, that is, report of at least one alcoholic drink within 1 month of the survey time, was 42 percent for 10th graders and 47 percent for 11th graders. The prevalence of heavy drinking—five drinks or more in a row—increased with each increasing grade from 9th to 12th grade. Tenth graders reported rates of heavy drinking, 18.4 percent, which increased in 11th graders, 25.2 percent. Marijuana use was found to be somewhat higher among boys than girls across all high school grades, and the overall prevalence of current marijuana abuse was 26 percent among 11th graders. The prevalence of cocaine use was reported to be higher in boys than girls, and lifetime cocaine use was 7.5 percent among 11th graders and 8.5 percent among 11th-grade boys.

In this phase of development, adolescents tend to identify with a group of peers who become highly influential in their choices of activities, styles, music, idols, and role models. Adolescents' underestimation of the risks associated with a variety of recreational behaviors and their sense of "omnipotence," mixed with their drive to be autonomous, frequently cause some conflict with parental requests and expectations. These conflicts, however, for the most part, do not result in the destruction of meaningful relationships with parents or siblings. For most teens, the process of

defining themselves as unique and different from their families can be achieved while they still maintain alliances with family members.

Sean, a 16-year-old junior in high school, had just gotten his driver's license. He realized that he was lucky to have received his parents' old car as soon as he had turned age 16 years, but he was overwhelmed when his friends called constantly for rides. Sean drove his younger brother to school each morning, and this created tremendous friction between them since Sean was always ready early and his brother, Jake was always late. Sean would threaten to leave without him in the morning, and on some days he actually did. His parents would then become angry with him for leaving his brother since that meant that one of them would need to take him and inevitably be late to work. Sean was a confident driver but felt pressured by his parents and by his peers, and considered not driving to school for the rest of the year. When nobody hassled Sean, he would leave in time for school, find a parking place at the crowded lot, and make it to class on time. He drove himself to his karate class, where he functioned as an assistant teacher. Sean, however, felt unable to say no to giving his friends rides, even when they brought beer and smoked marijuana in his car. Sean would half-heartedly object, however, he had never been popular, and believed that he would lose the few friends he had, if he did not take risks. Sean did not feel comfortable telling his parents about this, and so he continued to allow his friends to bring beer into his car until he had a "close call" in which a police car was tailing him, but did not pull him over. Sean began to panic, and had to stop his car on the side of the road. Sean's friends congratulated him for not getting caught, however, they didn't understand how much stress he was under. After this, Sean told his parents that he no longer wanted to drive to school, and he told his friends that his parents had taken away the car. Although Sean did lose some of his old friends which disturbed him, he felt more in control of his life. Sean began to spend more time playing his guitar and played online interactive computer games. He went through a period of social isolation and sadness, but over time, he developed new friendships with peers who didn't use drugs or alcohol, and played the same online games. He looked forward to applying to college and was glad to leaving the social situation in his high school. Sean put energy into studying and taking challenging courses in high school, including multiple "advanced placement" courses, which he found stimulating, and he felt that his decision to shun drugs and alcohol, despite the loss of some peers, was the right decision for him.

Late Adolescence

Late adolescence (between the ages of 17 and 19 years) is a time when the adolescent cognitive brain processes, impulse control and decision making begin to catch up with the social brain and behavior. During this phase, exploration of academic pursuits, musical and artistic tastes, athletic participation, and social bonds leads an older teen toward greater definition of self and a sense of belonging to certain groups or subcultures within mainstream society. As Erickson aptly described, through this exploration, an adolescent gradually moves toward a solid "identity formation." Well-adjusted adolescents can be comfortable with current choices of activities, tastes, hobbies, and friendships yet remain open to the idea that their "identities" will continue to be refined during young adulthood. Peer interactions are primary, and some bridges may be formed with family members to increase communication and social repartee. This transition spans the completion of high school and at least the first year of college, where exploration of new relationships and interests, often in a new setting away from home, is initiated. Prior to finishing high school,

however, the use of cigarettes, alcohol, and marijuana is higher than in the lower high school grades. The 2011 YRBSS reported that 13.6 percent of 12th graders had had the occasion to drive while they had been drinking alcohol, and 27.4 percent of 12th graders had been passengers with drivers who had been drinking alcohol. The prevalence of current frequent cigarette use was reported to be 10.8 percent among 12th graders. Alcohol use within 1 month of the survey was reported by 48.4 percent of 12th graders and 51.2 percent of 12th-grade boys, with episodic heavy drinking of more than five consecutive drinks in a row occurring in almost one-third (31.5 percent) of 12th graders. Nationwide, current marijuana use was reported by 31.1 percent of 12th-grade boys and 24.7 percent of 12th-grade girls. The prevalence of having ever used ecstasy was reported by 12.6 percent of 12th-grade boys and 9.9 percent of 12th-grade girls.

Neuroimaging studies provide evidence suggesting that during late adolescence and even into adulthood, the prefrontal cortex neural networks continue to facilitate working memory, prospective memory (the ability to keep in mind an intention to do something in the near future while multitasking), perspective taking (understanding another person's point of view), and responding to others while keeping one's feelings in mind. This phase of adolescence is an exciting time for the development of finely tuned cognitive abilities, social skills, and enduring relationships.

Samuel, 18½ years old, was in his second semester of college. The novelty of being away from home was wearing off, and he was pleased that his relationships with college friends felt deeper and more important to him than his high school peers. Samuel had entered college without any idea about a potential career, and during his first semester he had taken what he termed "random" classes. In high school, Samuel had been especially talented in languages, history, and math. He loved to travel and never gave much thought to what it might be like to be settled in a career. Samuel's parents had encouraged him to explore interests during high school, however, they made it clear that when he got to college, he was expected to set a career goal as quickly as possible, so that his college years wouldn't be wasted. So far, Samuel had met a number of classmates who had entered college already focused on future careers, so he began to feel pressure to decide on a future plan. Samuel agonized about the possibilities, and wondered why so many peers seemed to be so sure of their future goals and plans, while he felt very confused. Samuel felt somewhat alienated by his lack of direction, and he began to have insomnia. When Samuel mentioned his insomnia to his friend Mia, she was surprised by his level of stress, and told him she viewed him as an open and social person who made friends easily, and got along well with everyone. Samuel felt that his versatility made it even more difficult to find a major where he fits best. Finally, Mia challenged Samuel's parents' expectation to figure out the rest of his life right now. Instead, she suggested that he take at least two more quarters of classes in the liberal arts curriculum before making a decision about which career was the best match for him. Samuel realized that Mia's approach to life was very reasonable, and free of the pressures to plan his future right now, make commitments early, and see them through. Samuel had been influenced to believe that quitting or changing can be a sign of weakness. Mia's approach promoted considered life decisions based on positive experiences rather than on a timeline that didn't seem to fit. After much thought, Samuel took Mia's advice and decided to allow himself another year to decide on a major in and what to pursue as a career.

COMPONENTS OF ADOLESCENCE

Physical Development

Physical growth, along with the onset of sexual development, is identified as the "hallmark" of adolescence. Puberty is the process by which adolescents develop physical and sexual maturity, along with reproductive ability. Pubertal development begins in response to a surge of gonadotropin-releasing hormone from the hypothalamus. The hypothalamus is influenced by a multitude of physiological and psychological factors, and the onset of puberty may be delayed or altered in situations of severe nutritional deprivation. Gonadotropin-releasing hormone acts on the pituitary gland to release the gonadotropin-luteinizing hormone and follicle-stimulating hormone. These two hormones, in turn, act to stimulate the production of a variety of androgens, testosterone, estradiol, and other peptides. Physical development during puberty is also regulated by two additional hypothalamic peptides, growth hormone–releasing hormone and somatostatin, both of which promote the release of growth hormone from the pituitary gland. Puberty unfolds over time in conjunction with the pulses and feedback loops of these gonadotropins. The first signs of the pubertal process are an increased rate of growth in both height and weight. This process begins in girls by approximately 10 years of age. By the age of 11 or 12 years, many girls noticeably tower over their male classmates, who do not experience a growth spurt, on average, until their 13th year. Even more striking are the differences in the development of secondary sexual characteristics in girls versus boys approximately 13 years old. By age 13 years, many girls have experienced menarche, and the majority have developed breasts and pubic hair. The initial appearance of pubic hair in both genders may occur before puberty, during adrenarche, which is dependent on hormones secreted by the adrenal glands, rather than by gonadal development. Some components of pubertal development are identical in boys and girls and include an increase in height, weight, muscle, and fat, whereas others, such as breast development or growth of facial hair, are specific to only one gender. There is wide variation in the normal range of onset and timing of pubertal development and its components. There is a set sequence, however, in the order in which pubertal development proceeds. Thus, secondary sexual characteristics in boys, such as increased length and width of the penis, for example, will occur after the release of androgens from developed enlarged testes.

Table 35.3–1.
Sexual Maturity Rating (SMR) of Girls and Boys

SMR	Girls	Boys
Stage 1	Preadolescent, papilla elevated No pubic hair	Penis, testes, scrotum preadolescent No pubic hair

Stage 2	Breast bud, small mound; areola diameter increased	Penis size same, testes, and scrotum enlarged, with scrotal skin reddened
	Sparse, long pubic hair, mainly along labia	Sparse, long pubic hair, mainly at the base of the penis
Stage 3	Breast and areola larger; no separation of contours	Penis elongated, with increased size of testes and scrotum
	Pubic hair darker and coarser; spread over pubic area	Pubic hair darker and coarser; spread over pubic area
Stage 4	Breast size increased	Penis increased in length and width
	Areola and papilla raised	Testes and scrotum larger
	Pubic hair coarse and thickened; covers less area than in adults, does not extend to thighs	Pubic hair coarse and thickened; covers less area than in adults, does not extend to thighs
Stage 5	Breasts resemble adult female breast; areola has recessed to breast contour	Penis, testes, scrotum appear mature
	Pubic hair increased in density, area; extends to thighs	Pubic hair increased in density; area extends to thighs

Sexual maturity rating (SMR), also referred to as *Tanner stages*, range from SMR 1 (prepuberty) to SMR 5 (adult). The SMR ratings include stages of genital maturity in boys and breast development in girls, as well as pubic hair development. Pubic hair development is a less reliable measure of sexual maturity, however, because it may emerge in response to androgens from the adrenal glands before the pubertal process has actually begun. Table 35.3–1 outlines SMRs for boys and girls.

The primary female sex characteristic is ovulation, or the release of eggs from ovarian follicles approximately once every 28 days. When adolescent girls reach SMR 3 to 4, ovarian follicles are producing enough estrogen to result in menarche, the onset of menstruation. When adolescent girls reach SMR 4 to 5, an ovarian follicle matures on a monthly basis and ovulation occurs. Estrogen and progesterone promote sexual maturation, including further development of fallopian tubes and breast development.

For adolescent boys, the primary sex characteristic is the development of sperm by the testes. In boys, sperm development occurs in response to follicle-stimulating hormone acting on the seminiferous tubules within the testes. The pubertal process in boys is marked by the growth of the testes stimulated by luteinizing hormone. There is no discrete sign of the formation of viable sperm, but it must be assumed that at the time that ejaculatory capacity occurs, viable sperm are likely to be present. An adolescent boy's ability to ejaculate generally emerges within 1 year of reaching SMR 2. Secondary sexual characteristics in boys include thickening of the skin, broadening of the shoulders, and the development of facial hair.

In general, the majority of adolescents adjust well to the changes in their physical growth and maturation. Inasmuch as overall identity formation is a gradual process, adolescents incorporate their developing

physical identities into their evolving adult selves.

Cognitive Maturation

Cognitive maturation in adolescence encompasses a wide range of expanded abilities that fall within the global category of "executive functions" of the brain. These include the transition from concrete thinking to more abstract thinking; an increased ability to draw logical conclusions in scientific pursuits, with peer interactions, and in social situations; and new abilities for self-observation and self-regulation. These result in an adolescent's capacity to develop hypotheses, express altruism, create idealistic scenarios, and develop hopes and dreams for the future. Adolescents acquire increased awareness of their intellectual, artistic, and athletic gifts and talents, yet it often takes many more years into young adulthood for them to establish a practical application for these abilities.

The cognitive developmental theorist Jean Piaget has been a central influence on the way in which clinicians and researchers understand the cognitive strategies of children throughout development and the changes that occur during adolescence. The central cognitive change that occurs gradually during adolescence is the shift from concrete thinking (*concrete operational thinking*, according to Piaget) to the ability to think abstractly (*formal operational thinking*, in Piaget's terminology). This evolution occurs, according to Piaget, as an adaptation to stimuli that demand an adolescent to produce hypothetical responses, as well as in response to the adolescent's expanded abilities to provide generalizations from specific situations. Recent magnetic resonance imaging (MRI) studies showing evidence of a preadolescent increase in cortical gray matter in the brain, possibly due to a wave of increased production of synapses, may be a physical correlate of the adolescent's ability for expanded abstract reasoning. The development of abstract thinking is not a sudden epiphany but, rather, a gradual process of expanding logical deductions beyond concrete experiences and achieving the capacity for idealistic and hypothetical thinking based on everyday life. Younger adolescents may be able to verbalize an understanding of complex abstract constructs and still react to situations in rigid and concrete ways. For example, a young adolescent may be able to verbalize an understanding of the link between drug and alcohol use and health risks but cannot imagine it happening to himself or herself. Similarly, adolescents often believe that they can "beat the odds," even when they understand the mathematical meaning of the probability of an event occurring. Adolescents often use an omnipotent belief system that reinforces their sense of immunity from danger, a false sense of immortality, even when confronted with logical risks. Some degree of child-like magical thinking continues to coexist with more mature abstract thinking in many adolescents. Despite the persistence of magical thinking into adolescence, adolescent cognition departs from that of

younger children with regard to an increased ability for self-observation and the development of strategies to promote strengths and compensate for weaknesses. For example, an adolescent who becomes aware of a weakness in completing academic work or tests during the time allotted in school may realize that he or she knows the material but cannot perform well in timed examinations and will make a request of the teacher to receive additional time on examinations. The ability to develop strategies for survival and use proactive methods of making them happen increases during the period of adolescence. One of the important cognitive tasks in adolescence is to identify and gravitate toward those pursuits that seem to match the adolescent's cognitive strengths in academic courses and in thinking about future aspirations. Piaget believed that cognitive adaptation in adolescence was profoundly influenced by social relationships and the dialogue between adolescents and peers, making social cognition an integral part of cognitive development in adolescence.

Socialization

Socialization in adolescence encompasses the ability to find acceptance in peer relationships and in turn the development of more mature social cognitions enabling the adolescent to accept others. It is common knowledge that, for early adolescents, the skills for developing a sense of belonging to a peer group, are of central importance to a sense of well-being. Being viewed as socially competent by peers is a critical component in building good self-esteem for most early adolescents. Peer influences are powerful and can foster positive social interactions, as well as apply a pressure in less socially accepted behaviors or even high-risk behavior. Belonging to a peer group is, in general, a sign of adaptation and a developmentally appropriate step in separating from parents and turning the focus of loyalty toward friends. Social cognition—the way in which adolescents use logic and reason in their assumptions about peers and in their peer interactions—has been explored by Robert Selman. (See Asher and Gottman, 1981.) After interviewing children and adolescents with respect to their interpretation of various social interactions, he identified stages of social cognition. In school-aged children and preadolescents, Selman described social cognitions as "fair weather cooperation." That is, the friendships are characteristically on-again/off-again, depending on the level of affiliation or disagreement in the last social interaction. Thus, children between the ages of 6 and 12 years are able to engage in exchanges of ideas and opinions and acknowledge feelings of peers, but the relationships often wax and wane in a discontinuous way on the basis of altercations and good times. Friendships deepen with repeated good times, but, for some school-aged children, a variety of peers are often interchangeable—that is, a companion is sought when a given child has free time, rather than out of a desire to spend time with a specific friend. As

adolescence ensues, friendships become more individualized, and personal secrets are likely shared with a friend rather than a family member. A comfort level is achieved with one or several early adolescent peers, and the group may "stick together," spending most free time together. In early adolescence, a blend of these two social modes may emerge, small "cliques" may arise, and, even within the cliques, competition, and jealousies regarding which dyads are "preferred" or higher ranked within the clique may result in some discontinuities in the relationships. In later adolescence, the peer group solidifies, leading to increased stability in the friendships and a greater mutuality in the quality of the interactions.

Moral Development

Moral reasoning among adolescents depends on the maturation of cognitive strategies for problem solving, societal conventions, exposure to familial attitudes, and personal values adopted. Morality is a set of values and beliefs about codes of behavior that conform to those shared by others in society. Adolescents, like younger children, tend to develop patterns of behaviors characteristic of their family and educational environments and by the imitation of specific peers and adults whom they admire. Moral development is not strictly tied to chronological age but, rather, is an outgrowth from cognitive development. Piaget described moral development as a gradual process parallel to cognitive development, with expanded abilities in differentiating the best interests for society from those of individuals occurring during late adolescence. The development of trends in moral judgment has been divided into levels by Lawrence Kohlberg, who posed "moral dilemmas" to children and adolescents and subsequently characterized their responses. According to Kohlberg, morality can be divided into preconventional, conventional, and postconventional. Before adolescence, decisions regarding actions that are loosely defined as "right" and "wrong" are dependent on the anticipated consequence to those actions. School-aged children, for example, denoted as using level 1 of preconventional morality, may define actions that result in a negative consequence or punishment as "wrong" simply due to the consequence that may ensue. The most important defining feature of "right" then becomes being able to "get away with it," and it only becomes "wrong" if there is punishment associated with it. Some young adolescents, still functioning within the category of preconventional morality but one level higher (level 2), may define "right" and "wrong" as synonymous with what is perceived as fair and reciprocal. Thus, the philosophy of "an eye for an eye, and a tooth for a tooth" may prevail in this stage of morality. During adolescence, conventional morality, level 3, becomes prevalent, with the view that there is a "right" and a "wrong" determined by society, often without the possibility of mitigating or extenuating circumstances. Adolescents espousing this level of morality are often rigid in their concept

of "good" and "bad" and unable to find a role for motivation in their judgment of behaviors. Thus, adolescents may often take the viewpoint of applying the harshest punishments that society can offer when asked to judge complex moral situations—for example, a poor man robbing a pharmacy to obtain a life-saving medication for his child. Later in adolescence, at level 4, a greater understanding of societal rules may be espoused within the context of conventional morality. It is generally not until late adolescence or early adulthood that the complexities of moral decision making are recognized and incorporated into their thinking. Postconventional morality (level 5), according to Kohlberg, implies the ability to interpret complex moral dilemmas in which the best moral decisions of an individual may come into conflict with the moral standards of society. At this level of thinking, adolescents are more flexible in their understanding of "right" and "wrong." The concept of *extenuating circumstances* is understood at this level of moral thinking, allowing exceptions to rules that are ordinarily followed. Kohlberg postulated a final, higher level of postconventional moral development that he labeled level 6, in which an individual is able to recognize a universal moral imperative that might take priority over standard societal moral values.

Although Kohlberg's and Piaget's notions of moral development focus on a unified theory of cognitive maturation for both sexes, Carol Gilligan emphasizes the social context of moral development leading to divergent patterns in moral development. Gilligan points out that, in women, compassion and the ethics of caring are dominant features of moral decision making, whereas for men, predominant features of moral judgments are related more to a perception of justice, rationality, and a sense of fairness.

Self-Esteem

Epidemiological surveys suggest that the majority of adolescents rise to the challenges of their physical and emotional development with optimism and solid self-esteem. Self-esteem is a conglomerate measure of one's sense of self-worth based on perceived success and achievements, as well as a perception of how much one is valued by peers, family members, teachers, and society in general. The most important correlates of good self-esteem are one's perception of positive physical appearance and high value to peers and family. Secondary features of self-esteem relate to academic achievement, athletic abilities, and special talents. Adolescent self-esteem is mediated, to a significant degree, by positive feedback received from a peer group and family members, and adolescents often seek out a peer group that offers acceptance, regardless of negative behaviors associated with that group. There are also developmental factors that influence adolescents' cognitions about themselves that affect global ratings of self-esteem in this age group. Studies comparing reported self-esteem across

the lifespan indicate that there are trends in self-esteem ratings that are age related. Starting in adolescence, there are also gender differences in reported self-esteem. A study investigating global measures of self-esteem in school-aged children—adolescents, 13 to 17 years old, and older teens to young adults, 18 to 22 years old—found group differences depending on age. Overall reported self-esteem was found to be highest in school-aged children as compared to adolescents and young adults, and boys and girls were identical in their self-reports. During adolescence, global reports of self-esteem decreased throughout, with adolescent girls experiencing a drop twice that of adolescent boys. Self-esteem, at a group level, continued to drop during the high school years, rose in young adulthood during the college years, and continued to rise in early to mid-adulthood. Girls continued to rate themselves with generally lower self-esteem into adulthood. The question arises as to how most adolescents surveyed report relative good adjustment and yet, at the same time, experienced a drop in self-esteem. These data appear to reflect the "inflated" self-esteem of school-aged children, who may be seeing themselves as better, smarter, and more capable than they really are. During adolescence, it is quite possible that even well-adjusted individuals realize that they are not as exceptional as they used to believe they were. Adolescents seem to be more aware of their shortcomings and, therefore, had lower self-reported scores on self-esteem measures. In young adulthood, however, and into mid-adulthood, along with expanded cognitive abilities for self-regulation, reported self-esteem increases and continues to increase well into mid-adult life.

ADOLESCENT BRAIN DEVELOPMENT: CORRELATIONS WITH EMOTIONAL, SOCIAL, AND COGNITIVE CHANGES

It is now widely accepted that emotional, social, and cognitive brain processes continue to develop throughout childhood and early and late adolescence. The structural and functional development of various regions of the adolescent brain interacts with the social environment to influence social sensitivity, cognitive decisions, and risk-taking behaviors.

The work of Blakemore and Mills (2014) and other researchers have consistently suggested that a network of specific brain regions is involved in social cognitive processes. For example, an adolescent's increasing ability to understand the mental states and thereby the behaviors of others, "mentalizing," also called "theory of mind," has been associated with a network of brain regions including the dorsal medial prefrontal cortex (dmPFC), the TPJ, pSTS, and the anterior temporal cortex (ATC). These brain gray matter regions have been shown to continue to develop in volume, cortical thickness, and surface area throughout adolescence before stabilizing in young adulthood. During adolescence the overall gray matter does not mature by growing larger, instead it is pruned to be more

specialized. However, the prefrontal cortex, parts of which are involved in mentalizing and other portions are involved with judgment, impulse control and long-range planning continue to undergo dynamic changes into the early 20s.

A variety of tasks related to mentalizing such as recognizing the meaning of facial expressions, have been shown to activate the mPFC, and in fact, this brain region has also been shown to become diminished in size and more specialized between late adolescence and adulthood. Social emotions such as embarrassment, shame, guilt, and pride require an understanding of what others may be thinking or feeling (mentalizing) about someone, whereas fear does not. In adolescence, increased sensitivity to peer opinions is a hallmark trait, the corresponding brain regions in adolescents may be activated. Functional MRI study investigating neural activation during tasks related to viewing or reading social emotional scenarios found that adolescents recruited the dmPFC and even more intensely than adults given the same tasks.

In a recent study of an online communication game requiring the subject to understand the mental state of the online player in order to make correct decisions, a mid-adolescent group made less mistakes than a mid-childhood group, but more mistakes than a young adult group. This result supports that notion that social development is still in play and not maximized even in late adolescence and that social sensitivity and social skills increase into young adulthood.

On the other hand, brain regions associated with risk taking and sensation seeking include the amygdala, nucleus accumbens, (NAcc) and the prefrontal cortex (PFF). A study investigating the time course of maturation of the brain regions associated with risk-taking behaviors found that relatively earlier maturation of the amygdala and/or the nucleus accumbens (NAcc) compared to the prefrontal cortex suggested a mismatch between these structures. The amygdala has been found to increase in volume from late childhood until mid to late adolescence with less growth after the age of 16 years. On the other hand, the prefrontal cortex showed little change in gray matter volume in childhood and continues to decline into young adulthood. More research is needed in order to draw conclusions about adolescent risk-taking behaviors and sensation seeking in view of these differentially developing brain regions.

In contrast to gray matter, white matter volumes increase throughout childhood and adolescence and reflect a developing integration or “connectivity” of brain function regions which are myelinated.

During the course of adolescence significant qualitative changes in thinking are demonstrated in greater ability to discriminate emotional cues, increased self-awareness, selective attention to information, and increased response inhibition. Perhaps the most dramatic changes in cognitive abilities are an increased capacity for abstract reasoning,

planning, and affect modulation. Over the past two decades, MRI and functional MRI (fMRI) have become important modalities for studying both the structure and the function of adolescent brains. Two areas of the brain consistently shown by MRI to evidence structural changes during adolescence are the prefrontal cortex and the parietal cortex. Neuroimaging studies indicate that the prefrontal cortex matures more slowly than other areas of the brain and that its development is correlated with increased capacity in adolescents to inhibit behavioral responses. The prefrontal cortex is known to mediate "executive function," that is, skills involving working memory, selective attention, and voluntary response inhibition. In addition to executive function, the prefrontal cortex is believed to be involved in social and cognitive abilities such as self-awareness and the development of an understanding of another person's perspective. Recent studies have provided evidence of a linear improvement in performance by age on certain tasks within executive function, such as those involving selective attention, working memory, and problem solving. This lends support for an increase in performance in these tasks as a function of the pruning and synaptic reorganization of the prefrontal cortex.

Social cognition—in particular, recognition of facial expressions of emotion—has been studied in healthy adolescents. The amygdala, a region of the brain known to be associated with emotion processing, was found by functional MRI to be activated in adolescents shown fearful facial expressions and in response to happy faces compared to neutral faces. Among adolescents, left amygdala activity when presented with fearful faces decreased in girls but not in boys. Female adolescents also demonstrated greater activation of the dorsolateral prefrontal cortex, whereas male adolescents did not show this pattern. It was hypothesized that the pattern of diminished amygdala and increased prefrontal activation in female adolescents reflects increased maturity in the regulation of emotional experiences in them compared to male adolescents. Structural and functional MRI studies implicate puberty as a period of synaptic reorganization in the prefrontal and parietal cortex, which may place the adolescent brain in a sensitized state with respect to the development of executive functions and social cognitive skills. Clearly the adolescent brain continues to develop structurally, and the relationships of these structural changes to behavioral and emotional responses warrant further investigation.

CURRENT ENVIRONMENTAL INFLUENCES AND ADOLESCENCE

Adolescent Sexual Behavior

Sexual interest and experimentation are critical components of adolescent development for boys and girls. Sexual experimentation in adolescents often begins with fantasy and masturbation in early adolescence, followed

by genital touching with the opposite sex or, in some cases, same-sex partners, oral sex with partners, and initiation of sexual intercourse at a later point in development. In one study of urban adolescents, at least one-third reported some form of genital sexual activity in the last year with a partner, and approximately 10 percent reported oral sexual behavior with a partner. The balance between healthy adolescent sexual experimentation and emotionally and physically safe sexual practices has been one of the major challenges for society in the new millennium.

Trends in adolescent sexual activity are monitored by the YRBSS, which measures the self-reported prevalence of health-risk behaviors, including sexual activity. Surveys obtained for the years 2010 to 2011 provide detailed information regarding sexual activity in adolescents and risks associated with these behaviors. Among high school students across the nation, 47.4 percent reported that they had had sexual intercourse during their life, and 33.7 percent of high school students reported that they were currently sexually active. Among those who were currently sexually active, 60.2 percent reported that they or their partner had used a condom during their last sexual intercourse. The prevalence of having had sexual intercourse at least once was higher among black males (66.9 percent) than Hispanic males (53 percent), followed by black females (53.6 percent) and Hispanic females (43.9 percent). The overall rate of having had sexual intercourse in high school among white students was 44.3 percent, with white females (44.5 percent) reporting it slightly more frequently than white males (44 percent). A total of 15.3 percent of students report that they had had sexual intercourse with at least four persons during their lifetime. Among the 33.7 percent of currently sexually active students nationwide, 18 percent reported that either they or their partner had used birth control pills as a contraceptive before their last sexual intercourse.

Eighty-four percent of high school students had received instruction that includes information about acquired immunodeficiency syndrome (AIDS) or HIV infection. Nationwide, 12.9 percent of students had been tested for HIV.

Factors That Influence Adolescent Sexual Behavior. Factors impacting sexual behavior in adolescents include peer groups, psychosocial factors, temperamental and personality traits, externalizing and internalizing psychiatric symptoms, gender, cultural and religious background, racial factors, family attitudes, and sexual education and prevention programs.

A recent longitudinal study of predictors of sexual initiation and high-risk sexual behaviors in young adolescents indicated that externalizing factors, specifically a history of violent delinquency in middle school and substance use, were associated with sexual initiation in early adolescence.

A survey of more than 400 female adolescents of varied racial backgrounds found that higher levels of impulsivity were associated with a

younger age at first experience of sexual intercourse; a higher number of sexual partners; sexual intercourse without the use of contraception, including condoms; and a history of sexually transmitted disease (chlamydia). The level of impulsivity was not found to be significantly associated with age, race, or education. Although this investigation focused only on girls, it is likely that similar results might be obtained for boys.

Historically, male adolescents have initiated sexual intercourse at a younger age than female adolescents, and recent report from the Guttmacher Institute (www.guttmacher.org) suggest that approximately 11 percent of girls and 14 percent of boys report age of first experience of sexual intercourse to be before age 15 years. In general, adolescents who initiate sexual intercourse at younger ages are also more likely to have a greater number of sexual partners.

There is a suggestion from surveys of urban adolescents that early initiation of sexual intercourse, before age 15 years, is more frequent among African-American boys than in Hispanic boys and more frequent in Hispanic boys than in white boys.

The additive effects of more highly educated families, social and religious youth groups, and school-based educational programs can be credited with a decline in high-risk sexual behavior among adolescents in the 1990s. Responsible sexual behavior among adolescents has been determined as one of the 10 leading health indicators for the next decade. One urban school–based intervention designed to promote sexual abstinence in middle schoolers and support safe sexual practices in those students who continue to be sexually active reported that successful abstinence was only possible in those students who had not initiated sexual intercourse by the time of the intervention.

Risk-Taking Behavior

Risk-taking behaviors are the product of cumulative contributions from brain development in particular regions of the adolescent brain and the social environment. These contributions include the possibility of overactive reward-related circuitry as in the ventral striatum, as well as the influences of peer acceptance. In experimental circumstances, adolescents may rate the potential reward of a risky action as inflated, however, at times adolescents also rate the relative risks of actions as inflated.

Reasonable risk taking is a necessary response to increasing autonomy in adolescence, leading to confidence in forming new relationships and competence in sports and in social situations. However, high-risk behaviors among this generation of adolescents are associated with serious negative consequences and can take many forms, including drug and alcohol use, unsafe sexual practices, self-injurious behaviors, and reckless driving. Alcohol is the most commonly used substance among adolescents and is associated with a multitude of coexisting high-risk behaviors, such as

motor vehicle accidents resulting in injury and death, academic deterioration, fighting, and criminal behaviors. According to the National Institutes of Health survey, heavy drinking (defined as at least five consecutive drinks at one time) was reported by almost one-quarter of high school seniors. But overall, alcohol use among teens is at a historically low level with declines in daily use by 10th and 12th graders (0.9 percent and 2.2 percent, respectively). Adolescent frequent and heavy use of alcohol was associated with future antisocial behavior, as well as increased levels of depression. Furthermore, even lower levels of alcohol use predicted similar problems as teens entered young adulthood. This implies the need for preventive interventions during adolescence for those teens who use alcohol, even if drinking behavior does not appear to be frequent or serious.

Illicit drugs, including marijuana, cocaine, heroin, hallucinogens, "ecstasy," and amphetamines, also pose a serious health risk among adolescents. A recent National Institute on Drug Abuse survey indicated that, in young adolescents, more than 8.5 percent used an illicit substance in a given month, and, in older adolescents, approximately 25 percent used illicit drugs in a given month. In younger adolescents, boys and girls had similar rates of drug use, whereas older adolescent boys had slightly higher rates. Although there are vast differences between adolescents who had experimented on rare occasions with marijuana and adolescents who engaged in daily drug use, the large numbers of adolescents engaged in these behaviors at every level is cause for public health concern. Furthermore, use of illicit substances is also associated with youth crimes and violent behaviors. In 2010, there was an all-time low in the number of violent crimes among adolescents. In the last few years, rates of violent crime victimization, including aggravated assault, robbery, rape, and homicide, according to Child Trends Data Bank, are in the range of 30 crimes per 1,000 adolescents per year. This rate is higher than the rate of crime victimization in adults. Male adolescents are more likely than female adolescents to be the victims of serious violent crimes, with older adolescents, 15 to 17 years old, being more likely to be victims than their younger adolescent counterparts, 12 to 14 years old. These grave statistics indicate a need to identify risk factors and to find protective factors for today's youth.

Identifying adolescents with high-risk behavior is not always feasible until the high-risk behaviors have resulted in academic or legal consequences. Once this occurs, a given adolescent can be identified as being at "high risk" for future difficulties and, optimally, should be monitored by a counselor through school, an outside therapist, and, in some cases, through the legal system. It is not always easy for an outsider, at first glance, to separate adolescents engaging in normal experimentation from those embroiled in high-risk behaviors. When there is evidence that an adolescent is engaged in frequent alcohol or drug use, reckless behavior,

or criminal activities, urgent psychosocial interventions are needed. In less obvious cases, adolescents who have ongoing academic difficulties, do not participate in any athletic or other extracurricular structured activities, and are socially isolated may also be in need of psychiatric evaluation.

A survey of high-risk behaviors in adolescents in a primary care medical clinic setting found that the presence of tattoos and body piercings was associated with a higher probability of high-risk behavior. In the general population, approximately 10 to 13 percent of adolescents have tattoos, compared to 5 percent of the total population, including all ages. Body piercings have also become more prevalent among mainstream adolescents over the last two decades. Of the more than 500 adolescents surveyed in this study, 13.2 percent reported at least one tattoo, and 26.9 percent reported at least one body piercing, other than in their ear lobe, at some point in their lives. Both tattoos and body piercings were more common in girls than in boys (16.6 vs. 8.1 percent for tattoos, 36.7 vs. 10.1 percent for body piercing). Adolescents who endorsed possession of at least one tattoo or body piercing were more likely to endorse the use of gateway drugs (cigarettes, alcohol, marijuana), as well as experience with hard drugs (cocaine, crystal methamphetamine, and "ecstasy"). Female adolescents who had tattoos or body piercings, especially younger adolescents, were more likely to endorse a history of suicidal ideation or behaviors. Male adolescents with tattoos and female adolescents with body piercings were more likely to report a history of involvement in physical fights, carrying a weapon, or bodily injury as a result of physical fighting. Adolescents with tattoos or piercings were more frequently sexually active than were adolescents without them and at greater risk for unsafe sexual behavior.

Adolescents may choose clothing, hairstyles, and/or body decorations as a means of fitting in with a selected peer group, or to attract attention and distinguish themselves as special or different from others. The presence of tattoos or body piercings does not necessarily indicate high-risk behaviors in adolescents, however, in certain peer groups it is associated with risky lifestyles. Conversely, many adolescents with serious high-risk behaviors do not possess tattoos or body piercings. Given the serious consequences of alcohol and drug use and violence among adolescents, inquiry into these behaviors should be an integral part of every health care evaluation.

Adolescent Use of the Internet and Social Media

The Internet and social media are ubiquitous among adolescents and serve as platforms to socialize, communicate, and share talents, as well as a myriad of personal information, affecting all aspects of life for today's adolescents. The "digital revolution" the explosion and integration of multiple technologies into daily life) among adolescents, as discussed by Giedd (2012), has highlighted the adaptive nature of the adolescent brain

in its ability to embrace and incorporate these technologies. According to Giedd, the amount of time it took for 50 million users to develop (pace of "penetration") with respect to older technologies such as television was 13 years. However, for newer technologies the pace of "penetration" is moving at record speed. For the Internet, it was 4 years, for Facebook it took about 3½ years, for Twitter about 3 years, for iPads 2 years, and for Google+ it took only 88 days to gain 50 million users!

Recent reports indicate that about 90 percent of children and adolescents use the Internet in myriad ways, including instant messaging, visiting social networking sites, blogging, chatting in chat rooms, posting photos, videos, and stories, and obtaining information by visiting numerous websites. Positive effects of the Internet identified for adolescents include exponentially increased communication with both friends and family and increased access to academic information, and familiarity with social and cultural habits of people around the world. Concerns raised regarding teens' access to a constant and overlapping barrage of stimuli, however, include the ubiquitous "multitasking" that seems to characterize adolescent practices in the United States. While adolescents may be under the impression that doing their homework while chatting with a friend, listening to music, instant messaging and viewing facebook seems like an efficient way to get more done in the same amount of time, however, the brain is actually rapidly shifting between tasks leading to interruption of attention to potentially vital tasks. For example, an MRI study of a simulated driving task with or without competing stimuli showed significant impairment in driving performance during competing conversation, in addition to a sizable reduction in activation in the parietal lobe.

Some adolescents use the internet to learn about health-related issues and for self-help. In this regard, several health-related, adolescent-friendly websites are available that provide information on a variety of psychological and medical topics. One such site is http://www.teenshealth.org, a US-based website, provides information about general and healthy living information for kids, teens, and parents, including information on mental health, sexual health, causes of illness, a question and answer section, and information on different coping skills. Mind Your Mind (http://mindyourmind.ca) is a Canadian website for youth by youth, which provides resources to help teens manage their stress and mental health issues and includes personal stories, music, and a blog. There are significant risks associated with Internet use, especially for adolescents who share personal information and communicate with strangers on line. According to the recent Third Youth Internet Safety Survey, a national telephone survey that sampled 1,600 Internet users between the ages of 10 and 17 years who used the Internet at least once per month for the previous 6 months, 11 percent of youth who used the Internet

were found to be targets of some type of online harassment within the previous year, increased from 2005 (9 percent). Other sources report that approximately 11 percent of youth online are the targets of unwanted sexual solicitation, decreased from 2005 (13 percent). More than two-thirds (66 percent) of the youth who were the victims of the harassment incidents knew the harasser in person before the events. There were a few more female adolescent harassers (48 percent) than male (43 percent). Thirty-four percent of targets reported an aggressive offline incident that involved either a telephone call or in-person contact. Those youth who were victimized had higher rates of harassing another youth online. Those youth who were the targets of adult harassment were often asked to send pictures of themselves or received a telephone call or a visit in person from the adult. Approximately 33 percent of youth who reported that they were harassed online experienced distress related to it. Internet safety programs urge youth to avoid sharing personal information and not to communicate with strangers online.

It is clear that adolescent culture in the United States is dominated by technology in a variety of ways and is a vital part of teen social interactions. Social media technologies have enabled adolescents to connect with a vastly larger scope of other teens than they could meet in daily life. Teens are not necessarily more or less social due to social media, however, the opportunities are broader to interact with more people using more diverse modalities.

The Effects of Online Gaming and Massively Multiplayer Online Role-Playing Games

Up to 90 percent of youth in the United States play video games, ranging from sports games to cooking games, to many games utilizing violent strategies in order to "win" or move to more advanced levels. The effects of violent videogames on adolescents are controversial. Some studies have documented that teens who routinely play violent videogames become "desensitized" to violence with repeated exposure, which has been demonstrated on MRI. On the other hand, there is no evidence of increasing real-world violent events historically during the explosion of video gaming use. In fact, over the last two decades in which the videogame industry has boomed, rates of violent juvenile crime have diminished. The explanation for this is not clear, however, what is clear is that adolescents do not have difficulty distinguishing video game violence from appropriate responses in real-world situations.

Massively multiplayer online role-playing games (MMORPGs) are highly sophisticated multiplayer universes in which players create an individualistic character (avatar). In this medium, millions of adolescents along with children and adults immerse themselves in a graphic virtual environment and can interact with each other through their avatars on a

daily basis. Researchers have questioned the potential dangers of such gaming among adolescents with regard to their social development, self-esteem, and aggressive behavior. One study investigated the social interactions that occurred within such games. The most popular of this genre of game is World of Warcraft, with City of Heroes, Ultima Online, Everquest, Lineage II, and RuneScape less often played. Based on an online questionnaire, it was reported by players that 76.2 percent of male and 74.7 percent of female players had made good friends within the context of the game, and that the vast majority of players were also playing with real-life friends or family. Another study investigating adolescents who played video games compared to those who did not play these games found no differences between gamers and nongamers in time spent with their parents or friends or in sports or leisure activities; however, gamers were found on average to spend less time with reading and on homework.

References

Blakemore SJ. Development of the social brain in adolescence. *J R Soc Med*. 2012;105(3):111–116.

*Blakemore SJ, Mills KL. Is adolescence a sensitive period for sociocultural processing? *Annu Rev Psychol*. 2014;65:187–207.

Caminis A, Henrich C, Ruchkin V, Schwab-Stone M, Martin A. Psychosocial predictors of sexual initiation and high-risk sexual behaviors in early adolescence. *Child Adolesc Psychiatry Ment Health*. 2007;1(1):14.

Child Trends. *Violent Crime Victimization*. 2013. Available at: http://www.childtrends.org/?indicators=violent-crime-victimization. Accessed December 2014.

Cole H, Griffiths MD. Social interactions in massively multiplayer online role-playing gamers. *Cyberpsychol Behav*. 2007;10(4):575–583.

Cookingham LM, Ryan GL. The impact of social media on the sexual and social wellness of adolescents. *J Pediatr Adolesc Gynecol*. 2015;28(1):2–5.

Coverdale GE, Long AF. Emotional wellbeing and mental health: an exploration into health promotion in young people and families. *Perspect Public Health*. 2015;135(1):27–36.

Cummings HM, Vandewater EA. Relation of adolescent video game play to time spent in other activities. *Arch Pediatr Adolesc Med*. 2007;161(7):684–689.

Douglas K, Poletti A. *Life Narraties and Youth Culture: Representation, Agency, and Participation (Studies in Childhood Youth)*. United Kingdom: Palgrave MacMillan; 2016.

*Eaton DK, Kann L, Kinchen S, et al. Youth risk behavior surveillance—United States, 2011. *MMWR Surveill Summ*. 2012;61(4):1–162.

Fox J, Warber KM. Romantic relationship development in the age of Facebook: an exploratory study of emerging adults' perceptions, motives, and behaviors. *Cyberpsychol Behav Soc Netw*. 2013;16(1):3–7.

*Gentile DA, Li D, Khoo A, Prot S, Anderson CA. Mediators and moderators of long-term effects of violent video games on aggressive behavior: practice, thinking, and action. *JAMA Pediatr*. 2014;168(5):450–457.

Giedd JN. The digital revolution and adolescent brain evolution. *J Adolesc Health*. 2012;51(2):101–105.

*Giedd JN, Raznahan A, Alexander-Bloch A, Schmitt E, Gogtay N, Rapoport JL. Child psychiatry branch of the national institute of mental health longitudinal structural magnetic resonance imaging study of human brain development. *Neuropsychopharmacology*. 2015;40(1):43–49.

Kohlberg L. *Essays on Moral Development: The Philosophy of Moral Development*. New York: Harper & Row; 1981.

*Mills KL, Goddings AL, Clasen LS, Giedd JN, Blakemore SJ. The developmental mismatch in structural brain maturation during adolescence. *Dev Neurosci*. 2014;36(3–4):147–160.

Mitchell KJ, Jones L, Finkelhor D, Wolak J. Trends in unwanted online experiences and sexting: Final Report. 2014. Available at: www.uh.edu/ccrc/pdf/Full Trends Report Feb 2014 with tables.pdf. Accessed December 2014.

Murty VP, Calabro F, Luna B. The role of experience in adolescent cognitive development: integration of executive, memory, and mesolimbic systems. *Neurosci Biobehav Rev*. 2016;70:46–58.

National Institutes of Health, National Institute on Drug Abuse. Monitoring the Future Survey, 2013. *Behavior and Social Environment Indicators*. Available at: http://www.childstats.gov. Accessed in 2014.

Oudekerk BA, Allen JP, Hessel ET, Molloy LE. The cascading development of autonomy and relatedness from adolescence to adulthood. *Child Dev*. 2015;86(2):472–485.

Owen DC, Armstrong ML, Koch JR, Roberts AE. College students with body art: well-being or high-risk behavior? *J Psychosoc Nurs Ment Health Serv*. 2013;51(10):20–28.

Paus T, Wong AP, Syme C, Pausova Z. Sex differences in the adolescent brain and body: findings from the saguenay youth study. *J Neurosci Res*. 2017;95:362–370.

*Pedersen W, von Soest T. Adolescent alcohol use and binge drinking: an 18-year trend study of prevalence and correlates. *Alcohol Alcohol*. 2015;50(2):219–25.

Puzzanchera C, Adams B, Sickmund M. *Juvenile Court Statistics 2008*. Pittsburgh, PA: National Center for Juvenile Justice; 2011.

Rousseau A, Beyens I, Eggermont S, Vandenbosch L. The dual role of media internalization in adolescent sexual behavior. *Arch Sex Behav*. 2016 Dec 16. [Epub ahead of print]. PMID: 27987087.

Sanders RA. Adolescent psychosocial, social, and cognitive development. *Pediatr Rev*. 2013;34(8):354–358.

Smyth JM. Beyond self-selection in video game play: an experimental examination of the consequences of massively multiplayer online role-playing game play. *Cyberpsychol Behav*. 2007;10(5):717–721.

Soliman A, De Sanctis V, Elalaily R. Nutrition and pubertal development. *Indian J Endocrinol Metab*. 2014;18(Suppl 1):S39–S47.

Soliman A, De Sanctis V, Elalaily R, Bedair S. Advances in pubertal growth and factors influencing it: can we increase pubertal growth? *Indian J Endocrinol Metab*. 2014;18(Suppl 1):S53–S62.

Spencer JM, Zimet GD, Aalsma MC, Orr DP. Self-esteem as a predictor of initiation of coitus in early adolescents. *Pediatrics*. 2002;109(4):581–584.

Steinberg L, Scott ES. Less guilty by reason of adolescence: Developmental immaturity, diminished responsibility, and the juvenile death penalty. *Am Psychol*. 2003;58(12):1009–1018.

Strang NM, Chein JM, Steinberg L. The value of the dual systems model of adolescent risk-taking. *Fron Hum Neurosci*. 2013;7:223.

Strenziok M, Kreuger F, Deshpande G, Lenroot RK, van der Meer E, Grafman J. Fronto-parietal regulation of media violence exposure in adolescents: a multi-method study. *Soc Cogn Affect Neurosci*. 2011;6(5):537–547.

*Sue DW, Sue D. *Counseling the Culturally Diverse: Theory and Practice*. 6th ed. New Jersey: Wiley; 2012.

Telzer EH, Fuligni AJ, Lieberman MD, Miernicki ME, Galvin A. The quality of adolescents' peer relationships modulates neural sensitivity to risk taking. *Soc Cogn Affect Neurosci*. 2015;10(3):389–398.

Van Rensburg EJ, Raubenheimer J. Does forgiveness mediate the impact of school bullying on adolescent mental health? *J Child Adolesc Ment Health*. 2015;27(1):25–39.

*Willoughby T, Good M, Adachi PJC, Hamza C, Tavernier R. Examining the link between adolescent brain development and risk taking from a social-developmental perspective. *Brain Cogn*. 2014;89:70–78.

Yonker LM, Scirica CV, Jethwani K, Kinane TB. "Friending" teens: systematic review of social media in adolescent and young adult health care. *J Med Internet Res*. 2015;17(1):e4.

Yousefi M, Karmaus W, Zhang H, et al. Relationships between age of puberty onset and height at 18 years in girls and boys. *World J Pediatr*. 2013;9(3):230–238.

36

Child Psychiatric Examination

▲ 36.1 Psychiatric Examination of the Infant, Child, and Adolescent

ROBERT A. KING, M.D., PRUDENCE W. FISHER, PH.D., MICHAEL H. BLOCH, M.D., ARMIN PAUL THIES, PH.D., AND MARY E. SCHWAB-STONE, M.D.

Core features of the clinical diagnostic assessment of the child are identification of the reasons for referral, evaluation of the nature and extent of the child's emotional or behavioral difficulties, and determination of factors in the child, family, and environment that cause, exacerbate, or potentially ameliorate those difficulties. Although information gathering and differential diagnosis are the primary aims of such assessment, the evaluation process should also facilitate the development of an alliance with the clinician around understanding the child's difficulties and pursuing the treatment recommendations that emerge from the evaluation.

Psychiatric assessment of the child or adolescent is undertaken for a variety of purposes, including outpatient office or clinic evaluation to determine the need for treatment; emergency department assessment of safety issues and need for hospitalization; consultation with pediatric, school, or legal services; as a component of family or other mode of treatment; and for research purposes. Although the structure and style of the assessment vary to some extent depending on the purpose and setting, certain elements are common to all of the assessments.

Because decisions about intervention follow from evaluation, clinical assessment may rightly be regarded as a cornerstone of child and adolescent psychiatry. Clinical evaluation of the child or adolescent is challenging and complex, in that the clinician must draw on multiple sources and a range of techniques for eliciting information on various aspects of the child's functioning. This information must then be integrated from multiple theoretical perspectives, using the clinician's knowledge of development and psychopathology to arrive at a diagnostic formulation that can guide treatment planning effectively.

The initial portion of this section focuses on the comprehensive clinical assessment of the child and adolescent; subsequent subsections discuss

more specialized types of assessment, such as the assessment of infants and adolescents, psychological testing, laboratory assessment, and clinical research assessment instruments.

CLINICAL INTERVIEW OF THE CHILD

The psychiatric assessment of the child requires a comprehensive approach that evaluates the child's developmental progress in various domains and positive adaptive capacities, as well as the presence of the pathognomonic symptoms of specific disorders. A developmental approach to the assessment of the child is essential because children differ from adults in certain key respects.

Distinctive Features of Childhood Psychopathology

First, psychiatric disturbances in children often consist of a lack of developmental progress in one or more domains rather than the presence of specific symptoms that are pathognomonic of adult disorders. For example, a nursery school child's failure to develop useful social language or interactions or a school-aged child's inability to meet the developmental expectation of separating from parents and settling into the school day may prompt the parent, school, or both to request an assessment.

Second, the child's developmental status may affect the clinical presentation of various syndromes. For example, in children, depression often presents with irritability and somatic complaints, whereas excessive guilt or depressive delusions are rare.

Third, development brings expectable periods in which distressing emotions or behavior may occur as part of a normal transition, for example, the separation anxiety of a child starting preschool or the oppositionality of the adolescent.

In many cases, clinical conditions represent severe forms of symptoms found in milder form in nonreferred children. Fears, tantrums, moodiness, or restlessness are relatively common in childhood and occur transiently at different stages. Assessment may be sought by concerned parents needing guidance on how to understand and manage these developmental manifestations. Thus, the clinician must judge whether the behavior is likely to resolve with time and without substantial deleterious effect on the child or family or whether, instead, the level of distress, compromised functioning, or symptom persistence indicates the need for clinical intervention.

To distinguish transient or normative difficulties from those that are more clinically worrisome, the evaluator must possess a solid knowledge of normal and abnormal child development. This developmental frame of reference includes an understanding of what behaviors can be expected normally in children of different ages, the time frame within which various

behaviors normally wax and wane, and the natural history of psychiatric disturbances at different stages in development, including knowledge of the ages at which particular syndromes are more or less likely to present.

Another difference in the adult and child psychiatric assessment is that many children coming to clinical attention have difficulties that cannot be neatly subsumed under the rubric of a single diagnostic label. Thus, comorbidity is usually not the exception but the rule in childhood disorders. Even in epidemiological studies of children and adolescents, as many as one half of those who meet diagnostic criteria for one disorder also meet criteria for at least another disorder. This high rate of comorbidity, found even in nonreferred populations, may have several sources. Many traditional nosological entities draw their definition from clinical experience with adults; although childhood analogues clearly exist, the appropriate descriptive boundaries may not be the same. Furthermore, impairing symptoms in childhood tend to have a ramifying effect by interfering with the ongoing acquisition of key developmental skills in multiple areas. Finally, regardless of whether the origins of a given form of childhood psychopathology lie in biological factors, family or social environment, or interactions among these realms, these pathogenic forces often produce symptomatic difficulties that cut across many diagnostic lines.

Distinctive Aspects of Child Assessment

Except for patients brought involuntarily for evaluation, most adults coming for psychiatric assessment acknowledge, at least implicitly, some degree of self-perceived distress and wish for help, no matter how ambivalent the wish may be. In contrast, for most children, the responsibility for initiating and following through with a psychiatric evaluation lies with the parents. In many cases, children's behavior is a greater source of distress to others—parents, teachers, and peers—than to the children. Other children, even while acknowledging distress, view the locus of the troubles as purely external, and hence see little need for evaluation, even if they comprehend the process. In any event, the lack of any stated wish for help or active role in seeking the evaluation colors the child assessment process from the start. As a result, explicit attention must be paid to assessing the child's view of the problem and to cultivating the child's engagement in the process.

A related aspect of the child psychiatric assessment involves the need for the clinician to assess and to accommodate to limitations in the child's ability to reflect, to conceptualize, and to report experiences and internal states. These capacities emerge only gradually with maturation and are influenced by normal and pathological factors. Thus, the clinician's communication must be attuned to the child's developmental level and may require play, stories, drawing, and other alternative modes of interaction in

addition to direct discourse.

Context and the Need for Multiple Informants

Even more so than for an adult, the child's functioning and psychological well-being are strongly related to the social context—the family, school, and community settings—in which the child lives and develops. Thus, to evaluate the nature and severity of the child's psychiatric status, information is needed about the child's relationship to, and functioning in, these diverse settings. The fit of the child and social environment can be appraised from what parents say about their relationships with the child and from the child's direct report, as well as from play and drawings; however, direct observation of the content and tone of parent–child interactions in the office setting also provides much useful additional data. In some cases, a home visit may provide invaluable information. [Here and throughout, "parent(s)" is used to refer to the child's primary care-giver(s), regardless of biological relationship.]

It is also important to gather information on the child's functioning in school, a task that often requires speaking with the child's teacher or guidance counselor, in addition to reviewing school reports and educational assessments. Here, too, direct classroom observation can provide invaluable in vivo information. For children involved with child welfare, juvenile justice, or other institutional systems, information from caseworkers, probation officers, and institutional caretakers is essential.

Special Types of Assessment

The main emphasis of this section is on the assessment of children and adolescents, especially those capable of participating verbally in the assessment process. Because developmental considerations shape the evaluation process, infants and toddlers, on one hand, and adolescents, on the other, require specific modifications of the general approach outlined in this section. Nonverbal or language-impaired children, such as those with pervasive developmental disorders or intellectual disabilities, also require specialized assessment techniques.

Although the prototype for clinical assessment of the child is the office consultation or clinic visit, evaluations occur in a variety of settings, including schools, pediatric wards, residential treatment centers, detention centers, and hospital emergency departments. Procedures necessarily vary with the setting and reasons for carrying out the assessment. The following general model for the outpatient evaluation can be adapted to fit more specific clinical situations. More-specialized types of evaluation are discussed later in a separate subsection.

Preliminary Contact

In the initial phone call to the clinician or clinic intake office, it is helpful to gather some basic information about the nature of the chief complaint; to set the time, date, and place of the first appointment; and, if relevant, to discuss issues of fee or insurance coverage.

During the evaluation, the clinician wants to see the parents alone, the child alone, and the parents and child together. The order in which these should be done and the corollary issue of who attends the first meeting should be discussed at the initial contact. This decision is usually made with reference to the child's age. When evaluating a young child, the parents are often seen first, without the child. This allows the clinician to obtain a history and to discuss how to prepare the child for the evaluation. However, the adolescent should usually be included in the initial interview so that the youngster does not identify the clinician as an agent of the parents, which may interfere with the adolescent's willingness to talk with the clinician and impair the formation of a treatment alliance. If children or adolescents are to attend the initial appointment, the clinician should discuss in the initial call how they are to be prepared for the appointment.

In the initial phone contact, the issue of which adults will attend should also be discussed. When the family is intact, the presence of both parents is optimal. When this is not the case, the clinician and parent should discuss whether a noncustodial parent, stepparent, or other adult should be present.

Parent Interview

In the initial parent interview, the clinician seeks to understand the parents' view of the problems that have led to the referral, including the nature of the child's current difficulties, the explicit and implicit reasons for the timing of the referral, and the effect of the child's difficulties on the parents, individually and as a couple, and the well-being and functioning of the family as a whole. A detailed history of the child's development and a review of family medical and psychiatric disorders are obtained. Finally, the clinician seeks to understand how the family is functioning in its community and cultural setting. This includes a careful appraisal of the strengths of the child and the family, in addition to the history of the presenting symptoms and areas of impairment.

Child's Current Difficulties and Reasons for Referral. In child psychiatry, referrals are usually initiated by the parent or parents rather than the child. Similarly, parents are key in determining whether a recommendation for treatment is followed. Over and above the need to gather information about the child's symptoms and functioning, the parent interview provides an opportunity to develop an alliance with the parents around the goals of identifying and helping the child's difficulties. Success of this process requires understanding the parents' expectations and concerns regarding

the evaluation and its outcome. How these issues are addressed depends on the nature of the child's problems and the parents' knowledge of and previous experience with psychiatric evaluations and treatment. Attention to the parents' view of the evaluation process allows the clinician to address unrealistic fears or excessive expectations about what can be accomplished in a given time and to reduce family stress when parental anxiety is too high.

In taking a history of the current problem, the clinician obtains a detailed picture of the parents' current emotional and behavioral concerns about the child. Parents should be asked to describe specific instances of the child's problematic behavior, frequency, intensity, duration, circumstances in which the behavior occurs, their responses, and the child's responses to it. The extent to which the symptoms cause functional impairment is ascertained by asking about the level of the child's distress, interference with social and academic activities, effect on ongoing development, and effect of the child's behavior on others. The clinician should also inquire directly about behaviors or symptoms not reported by the parents but known to be commonly associated with the presenting problem or with disorders that the presenting problem may suggest. Thus, one goal of the initial parent interview is to develop a detailed picture of the presenting problems, their course, their effect on functioning and on the family, and what has been tried to help alleviate them.

The clinician is also interested in the implicit and explicit reasons that the child has been brought for help at this particular moment. Often, the reason lies in an escalation of troublesome behavior or in the deleterious effect of a deficit emerging into sharper relief (e.g., a learning disability or attention deficit becoming more apparent as the child faces the increased demands of a new grade). In other cases, the child's behavior may not have changed but some change within the family or loss of compensatory support may have made it impossible for the parents to tolerate or to ignore previously bearable behavior. These *iatropic factors* (factors propelling families to seek clinical attention) are often complex. In surveys of the general child population, rates of diagnosable psychiatric disorder (defined as meeting diagnostic and impairment criteria) approach 20 percent, yet the number of children actively seeking or receiving clinical services is far lower. Perceived barriers to services alone do not account for this low rate of seeking help.

Throughout the history taking, the clinician strives to understand the meaning and function of the symptoms in relation to the factors in the child and environment that influence them. A given symptom (e.g., a period of anxiety, theft, or hallucinations) may have quite different meanings, functions, and clinical implications in different children. Thus, detailed information on precipitating circumstances, behavioral concomitants, and consequences and alleviating factors is needed, much as

is the case in taking a medical history. However, this information must be interpreted within the broader developmental and family contexts in which the symptoms occur; thus, while gathering history, the clinician simultaneously takes in information about these contextual elements. History taking and diagnostic formulation are not distinct processes; during the interview process, clinicians continually formulate and test tentative hypotheses that guide further questions and the diagnostic possibilities under consideration.

To understand the child's presenting problems, developmental status, and family and social context, the clinician must learn about the child's strengths, interests, talents, and areas of adequate or superior adjustment. Such information is essential for a full diagnostic understanding of the child and helps the clinician to appraise the child's overall adjustment and functioning. Moreover, in planning for treatment, factors that may help to counterbalance or remediate the child's vulnerabilities must be identified. Finally, the clinician's appreciation and discussion of the child's strengths and areas of good functioning provide support and reassurance at a time when the child's problems and limitations have been the parents' major focus of attention.

The parent interview should also cover the practical and administrative aspects of the diagnostic assessment, including discussion of fees, scheduling, confidentiality, and, if applicable, permission to obtain information from school personnel and other clinicians. If the child was not included in the initial interview, the clinician should discuss appropriate preparation of the child for a meeting.

Developmental History. The developmental history is a detailed accounting of the child's development across physical, cognitive, linguistic, social, and emotional domains. One task of this part of the parent interview is to ascertain the timing and sequence of developmental events from pregnancy to the present. Parents vary in their ability to provide precise information on the timing of specific milestones; however, an adequate sense of the course of development can often be obtained by having the parents compare the patient's progress in a particular domain with that of sibling or peers. Asking parents about the timing of changes in the family, such as moves or other memorable life events, may help to provide a chronological framework for remembering features of the child's development. Besides gathering descriptive information about the child's development, this part of the interview often allows one to tap the emotional aspects of the parent–child relationship by querying the parents' hopes, worries, expectations, and life circumstances in relation to different developmental events.

It is useful to conceptualize the child's past and current development in the following domains: cognitive and academic development; family

relationships; peer relationships; physical development and medical history; emotional development and temperament; development of conscience and values; interests, hobbies, talents, and avocations; and unusual circumstances.

COGNITIVE AND ACADEMIC DEVELOPMENT. The child's pattern of cognitive strengths and weakness should be traced from early childhood, including verbal and attentional skills. Although academic performance is important (including areas of particular strength and weakness), important information about aspects of emotional and social development also flows naturally from the discussion of how the child progressed from one school year to the next. Thus, in tracking cognitive development and academic success, one should ask about the child's ability to separate from parents and to attend school regularly, interpersonal relationships with peers and teachers, motivation to learn, ability to function independently, tolerance for frustration and delay of gratification, attitudes toward authority, and ability to accept criticism, among others. A grade-by-grade history of the schools attended can be obtained during this part of the interview, as well as the reasons for any changes. Whether child has ever received a special education classification should be queried and the supporting evaluations and resulting services and accommodations classification should be reviewed. When the child's behavior or progress at school is among the reasons for the psychiatric evaluation, one should obtain permission to communicate with the child's teachers, counselors, or other school personnel and to review the school records, including results of standardized tests.

FAMILY RELATIONSHIPS. The parent interview should include assessment of how the child relates to each family member and how the child fits into the overall family system. The child's reactions to major family changes should be noted (e.g., deaths; birth of siblings; marital separation, divorce, or remarriage; and changes in caretaking arrangements, custody, or visitation), as well as parental responses to those reactions. Compliance with family rules and standards should be queried, as should consequences when the child does not comply (i.e., the usual mode of discipline or limit setting) and the child's response to such interventions.

PEER RELATIONSHIPS. The clinician should gather information about how the child relates to peers, including the number of friends, preferences regarding age and gender of friends, and any major changes in peer group; the child's satisfaction with these relationships; their relative reciprocity and stability; activities and interests shared with peers; and parents' feelings about the child's close peers or lack thereof. The parents' perspective on the child's social skills and deficits should be assessed, including their sense about any difficulties the child has in this domain.

Episodes of bullying or being bullied, and the consequences thereof, should be queried. For adolescents, this part of the history includes issues such as the capacity for intimate relationships, romantic interests, sexual activity, and concerns over sexual orientation or gender identity.

PHYSICAL DEVELOPMENT AND MEDICAL HISTORY. Assessment of physical development includes fine and gross motor development, toilet training and its lapses, eating behavior and attitudes, and sleep patterns. Precocious development or delays in physical growth or pubertal maturation should be noted. In this context, parents should be asked systematically about medication, illnesses, hospitalizations, emergency room visits, operations, episodes of loss of consciousness, seizures, head injury, or other serious injuries, as well as the child's reactions to these and their effect on subsequent health and activities. Inquiry about tics, difficulties with hearing or vision, lead exposure, and other specific conditions may be suggested by the clinical presentation. When necessary, medical reports should be obtained and reviewed.

EMOTIONAL DEVELOPMENT AND TEMPERAMENT. Although some of these aspects of emotional development and temperament will have been covered in taking the history of the presenting problems, the clinician should fill in gaps with a systematic review. This domain of assessment includes the child's present and past mood and capacity for affect regulation, anxieties, and ability to adapt to changes and situations that are new, challenging, or frustrating. Prevailing mood as perceived by the parents should be assessed; the clinician should ask specifically about self-injury and suicidal ideation, gestures, or attempts, especially if the child has shown depressive symptoms, including irritability. Worries, fears, and other manifestations of anxiety should be covered, and the level of any distress and impairment should be explored. The generalized or specific nature of symptoms, situational evokers, historical precipitants, and effects on the child in terms of impairment or distress are relevant. It is also important to note unusual fears, excessive shyness or withdrawal, and obsessive or compulsive symptoms.

The child's capacity to regulate aggression falls within this domain and should be reviewed. The clinician should ask about circumstances in which the child becomes angry or aggressive, the mode of expressing anger (i.e., verbal, physical, or both), the effect on others, and the child's reaction to and processing of such feelings and behavior. In this context, the clinician seeks to evaluate how the child manages aggression and whether the child is too aggressive or too fearful of his or her anger or that of others. Specific aggressive symptoms (e.g., bullying, fire setting, vandalism, or cruelty to animals) should also be inquired after.

DEVELOPMENT OF CONSCIENCE AND VALUES. The clinician should assess the

development of conscience to gauge whether it is too harsh, lax, or overly focused on particular issues. The effectiveness of conscience in helping the child conform to family and societal expectations is important. Religious or ethical concerns and their concordance with those of the family can be addressed in this phase of history taking. The family's expectations with regard to the child's values and future life choices should be ascertained from the parents' perspective, and areas of potential discord should be noted.

INTERESTS, HOBBIES, TALENTS, AND AVOCATIONS. Although this line of inquiry is also pursued with the child, it is useful to obtain the parents' perspectives on the child's interests and activities and to assess the parents' approval, involvement, and support for them. Indications of parent–child conflict should be probed and may provide useful information about the relationship. The presenting difficulties may have affected the child's ability to focus on or to engage in previously enjoyable areas of interest or activity, such as sports or music.

The quantity and type of television programs, movies, screen time, and video games that a child is permitted provide important information on the child's interests and the quality of parental limit setting. Access to and use of the Internet, instant messaging, and social Web sites should be assessed, as well as the degree of parental supervision and monitoring.

UNUSUAL CIRCUMSTANCES. When present, the clinician should assess the child's experience of immigration or exposure to unusual or potentially traumatic circumstances, such as sexual or physical abuse, family or community violence, natural disaster, or armed conflict. If a history of such experience exists, the child's immediate and subsequent reactions and the nature of the response from parents or other adults should be assessed.

Family and Community Background. Children develop in the contexts of the family and community, and these provide opportunities and challenges that influence and shape the course of that development. The clinician should gain a picture of the child's functioning in these contexts and their influences on the child. Although the clinician learns a great deal about these areas from taking the history of the presenting problem and the developmental history, direct questioning may also be needed.

PARENTS. The clinician should assess the parents as individuals and as a couple, including strengths, weaknesses, and areas of conflict or difficulty. If the child resides with someone other than the parents (e.g., a relative or foster parent), the clinician should obtain information on the history and circumstances of that relationship, including how the child came to be in that adult's care. It is also necessary to determine what the child's previous experiences were with the parents and the nature of the bond with them.

The clinician should similarly assess the current caregivers and their parenting relationship with the child.

If the child is adopted, information is needed on the circumstances, the parents' feelings and expectations about the adoption, the child's age at adoption, the child's adjustment at the time and later, and how the child and family members currently view the fact of the adoption. The clinician should also determine how the adoption has been discussed with the child and any concerns or questions about the biological parents.

When taking the developmental history, it is often natural to inquire about the parents' feelings and involvement with the child at different stages. Parents' past attitudes and behavior are likely to shed light on their responses in the current situation; hence, one wants to know how they felt about past difficulties and what they have done to help and to support the child. The parents' individual views on the causes of the child's difficulties, and the extent to which they agree or disagree on this and other aspects of the child's care and management, affect how they respond to the clinician's recommendations. Information on the parents' backgrounds and histories with their families of origin helps the clinician to understand and work with their responses to their child.

The clinician should ask about the parents' ethnic and religious backgrounds. They may be a source of shared identification and values to be imparted to the child or a source of conflict between the parents or between an adolescent and the parents. If the clinician is not sufficiently familiar with the culture or language of the family to understand the cultural context of clinically relevant aspects of the child and family's situation, a respectful request for explanation should be made.

The parents' education and occupation may bear on how they understand and respond to the evaluation. How knowledgeable and worried they are may depend on their past experiences and familiarity with psychiatric issues. Financial resources and insurance coverage should be ascertained because this may be a source of stress or worry and may pose practical constraints on the treatment options available to the family.

The clinician should gain an understanding of how the family system functions, with whom the child lives, and the nature of relationships between household members and the child. Other issues to explore are boundaries and alliances within the family and the child's affinities and conflictual relationships with parents, siblings, and other household members. Family communication and problem solving should be assessed, including how issues of separation or disagreement are handled. The emotional tone of the family (e.g., tense, anxious, critical, competitive, warm, or supportive) can be ascertained by observing interactions and by direct inquiry. Stresses on the family as a whole or individual members (e.g., moves, migration, illness, accidents, job loss, or economic or legal difficulties) should also be addressed.

COMMUNITY. The clinician should understand the community or neighborhood in which the family lives and their feelings and identification with it. Information on the family's involvement with civic, community, and religious activities (and the extent to which the child participates) should be ascertained. The clinician also needs to know about issues of safety and adversity (e.g., poverty, poor housing, and high rates of crime or urban violence) in the neighborhood. In some communities, mental health resources are viewed positively; in others, their use is stigmatized, which may adversely affect the child. These issues should be assessed and understood because they may affect how and where the child receives any indicated treatment.

Family Medical and Psychiatric History. A careful family medical and psychiatric history is needed to identify any medical or psychiatric disorders with potential genetic or environmental implications for the child. Among those to be routinely considered are psychotic disorders, mood disorders, anxiety disorders, tic and obsessive-compulsive spectrum disorders, alcohol or substance use disorders, attention-deficit disorders, learning disorders, and antisocial personality disorder, as well as metabolic and neurological disorders. The effect of any family members' symptoms, hospitalizations, or incarcerations should be probed, with particular reference to the child.

Special Issues in the Parent Interview

INFORMANT DIFFERENCES. Although parents' concerns usually play a major role in initiating the evaluation, parents may perceive the child's problems differently and may disagree about their severity and the need for treatment. Parents' accounts of the child's difficulties and of various aspects of the developmental history may not agree completely with each other or with those of the child, teachers, or records of past events. To note these discrepancies is not to denigrate the reliability of parents as informants but rather to emphasize the necessity of multiple informants.

Informants provide differing accounts of children's problems for a number of reasons. Thus, discrepancies should be noted and regarded as potentially useful data regarding what they reflect about the clinical situation. Informants may interact with the child in different settings that vary in the extent to which the child expresses disturbing feelings and problematic behavior. This is commonly true for behaviors expressed most dramatically in settings that challenge the child in vulnerable areas (e.g., the school's demand that a child with attention-deficit/hyperactivity disorder [ADHD] sit still and pay attention). Information about such situation-specific symptoms is best provided by those who spend the most time with the child in that setting. These different perspectives on the

child's behavior may also have led to tension between informants who see the child discrepantly. (This is often true with divorced or separated parents, whose child may show quite different behavior in the context of each parent's home.) The clinician should be aware of these different perspectives and the feelings that they generate, so that the focus is maintained on evaluating and providing services for the child, with adults working together rather than at cross-purposes.

Another reason for discrepant reporting involves differences in how the informants evaluate the child's behavior. Observers may judge the same behavior by differing standards or expectations for many reasons. The clinician's appreciation of the informant's psychology or cultural background is useful for understanding and working with such differences. For example, when one parent is overly worried about the child having a serious disorder and the other is too ready to chalk up symptoms as something the child will "just grow out of," the clinician must navigate carefully to allow both parents to move to a reasonable middle ground that best serves the child's needs. In such instances, understanding each parent's defensive style helps the clinician to avoid appearing to collude with one parent and alienating the other.

Discrepant reporting also results from differences between informants in their willingness or ability to report their views to the clinician. For example, parents may be quick to report a behavior that they find disturbing or annoying, whereas a child may be unable to describe the problematic behavior verbally, may not appreciate the level of disturbance that it creates for others, or may not report it because of feelings of shame or fear of being blamed or punished.

Generally speaking, methodological studies and clinical experience suggest that parents are more likely than children to report disruptive or externalizing behaviors, such as restlessness, impulsiveness, opposition, or aggression. Conversely, children may be more likely to report anxious or depressive feelings and symptoms (including suicidal thoughts and acts) of which the parents may be unaware. Because children may be immature in their capacity to report information, such as frequency and duration of symptoms, parents are usually more accurate informants about chronological details. The child, however, may be the only source of information on events, such as sexual abuse, exposure to violence, or behavior that occurred in peer contexts (e.g., bullying). Generally, children become more reliable reporters of specific symptoms with age, with prepubertal children tending to be less reliable reporters of symptoms than their parents.

Mothers and fathers also differ in their experiences with their children, depending on the child's age and gender, shared interests, parental availability, and the respective temperaments and personalities of the parents and child. Although mothers have been reported to be more

reliable informants than fathers, in practice it is wisest, for gathering the fullest information possible and for fostering an alliance with the parents, to involve both parents from the outset. When appropriate, other informants, such as siblings, grandparents, or child care givers, may be called on to contribute information, but only after the implications for the family system and the child's specific relationships with household members are carefully thought out.

Parents and teachers may view the nature and severity of the child's difficulties differently because of the situation-specific nature of some presentations. When present at home and at school, symptoms such as hyperactivity and inattention may carry a different clinical significance and prognosis than those noted in only one (but not both) of the settings. Parents and teachers may also differ in their appraisals of the child's behaviors. For example, a stressed and preoccupied mother may be oblivious to her child's behavior, whereas an observant teacher with years of experience with similar-aged children may readily identify the child as distressed or disruptive.

For the clinician conducting the assessment, the issue is not which informant is right but how to understand the differences in viewpoint most effectively for the patient's benefit. Thus, eliciting and synthesizing information from multiple informants is a critical aspect of the clinical evaluation of children and adolescents.

INTERVIEW TECHNIQUES WITH PARENTS. The clinician should use a variety of interview techniques to elicit information from parents. The type and quality of information obtained are, in part, determined by the methods used and the clinician's skill and comfort in using them. A series of empirical studies confirm that systematic factual information and details about onset and timing of symptoms are generally most easily elicited by specific and direct questions, whereas an understanding of feelings and relationships in the family is usually gained most effectively by means of more open-ended, indirect questioning strategies.

In the initial interview, parents should have the opportunity to tell their story in their own way. For rapport to develop, they must feel that they have been able to express their concerns in their own words and were heard and empathically understood. An open-ended beginning also avoids narrowing down too quickly on a single aspect of the child's situation and conveys the clinician's interest in many different aspects of the child's situation. Subsequent probing can clarify details and can help to shape the clinician's understanding of diagnostic and potential treatment issues.

Parents are most likely to be nondefensive and comfortable about sharing personal information when they feel that the clinician is trustworthy, interested, understanding, and nonjudgmental. Clinicians must thus convey respect and empathy and avoid seeming aloof or

excessively familiar. They should discuss the child and family's situation in a natural way, using words and concepts that the parents can understand. The range of questions and the tone of the inquiry convey the clinician's interest in the whole child (including strengths, interests, and talents) and an appreciation that the child is not merely the repository of a set of symptoms.

EXPECTATIONS AND ANXIETIES. Parents may hold strong preconceived ideas about the nature of their child's problem and the kind of help needed. Because of the extensive coverage of issues relating to psychotropic medications in the popular press, many parents have drawn conclusions about the appropriateness of medications in general and in relation to their child's needs. For example, parents may arrive convinced that their child needs fluoxetine (Prozac) or that methylphenidate (Ritalin) is an addicting and inherently harmful drug.

Preconceptions or parental anxiety or conflict may also be played out around the issue of blame for the child's problems. In such situations, a parent may maneuver to have the clinician arbitrate or to agree that the teacher, school, or other parent is responsible for the child's behavior and distress. The clinician should be alert to such situations and should handle them without making either parent feel defensive or discounted. If ignored, such hidden or overt agendas can have a deleterious effect on the assessment process and the likelihood that treatment recommendations will be accepted or successfully implemented.

Parents may bring other entrenched, preconceived agendas to the clinical evaluation process. A parent may insist that the evaluation confirm that the child has been abused (e.g., by a former spouse) or does not have a psychiatric problem but a medical one. Alternatively, the parent may demand that the clinician agree that there is nothing wrong with the child. In such instances, clinicians must explain carefully and clearly that they will conduct a thorough assessment and provide thoughtful professional advice, adding that the assessment must be done openly, without preordained conclusions. This may not satisfy some parents, and some situations may simply not be workable.

Bringing a child for psychiatric evaluation is a charged and emotionally difficult step for nearly all parents, and it is only taken after much worry and disappointment that various accommodations, home remedies, or school-based adjustments have not sufficed. Children are the repositories of parents' dearest hopes (and often darkest fears) for themselves. Although a parent may rationally conclude that seeking consultation is reasonable and necessary, admitting that one's child (or oneself) needs help is always painful. Submitting one's child and one's parenting to professional scrutiny stirs many anxieties, as well as positive expectations of help, even in the most enlightened parents.

Parents often bring unspoken anxieties about the child's possible diagnosis or prognosis to the evaluation. These anxieties may be based in the parent's personal or family experience with psychiatric disturbance or may derive from accounts in the popular media. For example, is the child's uneven development a sign of autism or retardation? Do the facial grimaces mean that the child is developing a disabling tic disorder? Is the child's moodiness a harbinger of a serious mood disorder or suicide? Other parental anxieties may be guilty ones. Is failed parental rearing or transmitted heredity somehow to blame for the child's difficulties? Is the child fated to share the parent's (or other relative's) painful struggle with depression, anxiety, or psychosis? Does the parent's relationship to the child seem to recapitulate some distressing aspect of the parent's own early family experience?

The clinician must be aware of such parental anxieties and must remain alert for hints of them as the evaluation progresses; some parents are too anxious or ashamed to be forthcoming about them. By and large, it is best to have these concerns aired, even if they cannot be settled at the initial stage of the evaluation and perhaps cannot be settled for some time. When such anxieties are revealed or suggested indirectly, the clinician must express an interest in them with sensitivity. Respect and concern for these anxieties can be conveyed without dismissing them or providing premature or unfounded reassurance. The clinician may need to acknowledge the limits of knowledge about cause or prognosis in individual cases. Parents often seek psychiatric assessment after years of worry, having been given well-meaning (but inappropriate) reassurances that the child would grow out of it. Frank discussion of these issues may provide some relief to the parents, albeit perhaps not definitive answers to their questions.

PREPARATION OF THE CHILD. At some point in the initial interview, the clinician and parents should discuss the preparation of the child for his or her first appointment. Children have many reactions to a first visit to the psychiatric clinician, and parents must anticipate their child's possible reaction and prepare their response. Children may feel anxious because they cannot envision what will occur or because the evaluation may feel like a punishment. Some children fear that the assessment will reveal some profound or irremediable problem (this may be projected onto the clinician or parent with the reproach, "You think I'm crazy"). Many younger children fear shots or having to get undressed for a physical examination. In most cases, the child is probably aware of many of the parental or school concerns that have led to the evaluation and may feel a sense of relief (which can nevertheless coexist with anxiety) about the possibility of getting some help. The child's anxiety or defensiveness may be reduced if the purpose of the visit is put in supportive (rather than pejorative or accusatory) terms. The parents might be encouraged to tell the child

something like, “We know you’ve been having a hard time at school [or feeling upset and sad or having trouble with friends], so we are going to see Dr.—, who is used to helping kids with that sort of problem and can maybe figure out what would help.” For younger children, explanatory terms, such as a *feelings doctor* or a *talking doctor* who helps children with problems and worries, may be useful. The parents must give permission and encourage children to tell the clinician about their concerns, even those that might seem to involve private family matters not otherwise spoken about to others.

Child Interview

In the child interview, the clinician directly assesses the child’s psychological functioning, developmental status across different domains, and perceptions and experiences of the presenting problems. Like the parent interview, the child interview also provides an important opportunity to establish rapport. Such rapport reduces the level of stress experienced by the child during the assessment process, enhances the child’s engagement and willingness to share information, and facilitates the child’s cooperation with subsequent recommendations for intervention. Thus, from the outset, the clinician should convey empathy and interest in understanding and helping the child.

There has been debate regarding the reliability, validity, and clinical usefulness of various approaches to the child interview, but it is generally agreed that the direct child interview is an essential component of a comprehensive clinical child assessment. It provides observational data and diagnostic and historical information that are not accessible from other sources. In particular, information about feelings and moods, level of distress, and the child’s view of significant events is essential for a full diagnostic understanding of the child’s problems. For their own reasons, parents may not understand or fully appreciate the importance of these issues, or the child or adolescent may not have divulged them. The child may be guarding important information, such as secrets about abuse, bullying, sexual or antisocial activities, substance use, suicide attempts, pathological patterns of eating, or collusive family dynamics. Certain symptoms, such as compulsions, premonitory tic urges, obsessive or suicidal thoughts, or hallucinations, may be apparent only to the child. Other symptoms may escape or mislead the untrained eye; detection of tics or a thought disorder may require direct clinical observation.

There are many varieties of child interviews, depending on the nature of the presenting problem, the clinical setting, and the purpose of the assessment. Thus, the scope, focus, duration, setting, and predominant techniques used vary, depending on, for example, whether the assessment is an emergency department evaluation of a suicidal threat, a pediatric ward consultation of a regressed or uncooperative child, a forensic custody

evaluation, an investigation of alleged child abuse, or the evaluation of a child who is academically underachieving, anxious, or manifesting tics.

The clinician approaches the child interview with a developmentally based understanding of child and adolescent functioning and psychopathology. Within this framework, the clinician expects to obtain information relevant to the diagnostic questions raised by the presenting problems and so systematically reviews the major domains of symptomatology. The order and manner of conducting the child interview vary with the child's age and developmental status, the nature of the presenting problems, the setting and context, and the clinician's style. The constant elements in the interview process are the orienting framework chosen and the clinician's clinical skill in engaging and eliciting relevant data from children of varying ages and capacities with diverse problems and concerns. Some data may emerge spontaneously; some require questioning or other deliberate means of eliciting information.

Components of the Child Interview. The components of the child interview are often conceptualized under two broad headings: History taking and the mental status examination. *History taking* involves inquiry into the significant areas of the child's life and functioning, past and present, including the presenting problem. The *mental status examination* consists of an assessment and description of the child's appearance and functioning as manifested in the interview. Although these aspects of the child interview can be conceptually distinguished, they usually proceed concurrently. For example, the interviewer may ask a specific question, such as who lives at home with the child or how the child gets along with a sibling or a teacher. From a history-taking point of view, the child's response provides some information about these facts, as well as how the child feels about them. Although the child's response is thus of interest for the historical information it reveals, the child is also providing the clinician with data relevant to the mental status examination (e.g., how the child conceptualizes the response, how verbally fluent the child is, how readily the child engages, how confiding or suspicious the child is toward the interviewer). Similarly, while engaging a young child in play with figures or puppets, the clinician can simultaneously gather information for the mental status examination (e.g., speech, relatedness, and conceptual skills) and inferential data about the child's feelings, fantasies, and conflicts.

Several distinctive characteristics of clinical work with children bear on the manner of conducting the child interview and mental status examination and on interpreting the data that they yield. The focus of history taking and the mental status examination of the child is developmental, that is, it seeks to describe the child's current presentation in various domains and to compare the child's functioning in these areas with that expected for the child's age and phase of development. Another

distinctive aspect of the child interview involves the child's potential lability and propensity to react in more immature ways when tired, sick, anxious, or stressed by an unfamiliar situation. Although one can sometimes gather sufficient information for an assessment from a single interview, usually more than one interview is needed to assess the child's optimal or characteristic level of functioning.

Mental Status Examination. For the mental status examination, the child psychiatrist observes and assesses the following areas:

Physical appearance: Includes physical development, presence of minor congenital anomalies, style of hair and dress, cleanliness, and other indicators of the quality of self-care or parental attentiveness (or its lack) to the child's physical appearance and hygiene.

Manner of relating to examiner and parents: Includes ease of separation, how confiding or guarded the child seems toward the clinician, eagerness to please or to impress, flirtatious behavior, readiness to make excuses or to minimize, passive withholding, or open defiance. The child's behavior toward the examiner is compared to that observed or reported toward the parents or other adults. Reactions to the clinician's physical appearance, gender, or ethnicity (especially if different from that of the child) should be noted.

Affect: The child's predominant mood and range of emotion during the interview and appropriateness of mood.

Coping mechanisms: As manifested toward the clinician, expressed in play, or described by the child in relation to important others; age-appropriateness of the child's dependent longings, sexual interests and impulses, and aggressive feelings, with respect to intensity, object, and mode of expression; child's control or modulation of such urges (e.g., finding alternative or socially permissible means of satisfying them); strategies used for coping when anxious or frustrated.

Orientation to time, place, and person: Whether children are accurately aware of the date, where they are, and who they are. This can be evaluated by asking the child "What is the year? season? date? day? month?" or "Where are we?" (name of building, town, state, country).

Motor behavior, including activity level: Coordination, presence of unusual postures or motor patterns (e.g., tics, compulsions, or stereotypies, such as hand flapping or twirling).

Quality of thinking and perception: Includes presence of hallucinations, delusions, thought disorder, and flight of ideas and adequacy of hearing and vision.

Speech and language: Includes reading and writing; articulation, inflection, pitch, rhythm, and fluency of the child's speech; richness, limitation, or idiosyncratic aspects of vocabulary and syntax; presence of echolalia or persistent neologisms; misuse of pronouns and gender;

delayed or deviant social use of language; and lack of accompanying nonverbal communicative behavior (e.g., lack of facial expression or eye contact with interlocutor). When evidence of delays or problems in these realms exists, formal speech and language evaluation by a qualified specialist may be warranted.

Overall intelligence and fund of knowledge: Capacity for classification, abstraction, and inference appropriate to age; counting, alphabet, days of the week, months of the year, reading, writing, computational abilities; geographical and historical facts.

Attention, concentration, and impulsivity: Level of attention to an activity or topic of discussion, degree to which child jumps from activity to activity, need for physical limits, distractibility (e.g., by outside noises).

Memory: Immediate short- and long-term memory (e.g., repeat three items and recall at 5 and 30 minutes).

Neurological functioning: Fluctuating alertness, tremor, nystagmus, choreiform movements, abnormal gait, neurological soft signs, and cerebral dominance (preferred hand, foot, eye, or ear); testing for neurological soft signs can be incorporated in a game-like fashion (e.g., ball throwing, walking on a line, heel-to-toe walking, evidence of hand overflow when walking on outside of feet, and timing 20 repetitions of finger tapping, successive finger tapping, and alternating pronation–supination of the hand are relatively reliable components of a standardized exam for neurological soft signs).

Judgment and insight: Especially concerning the presenting problem; this is most usefully judged after the child has had the opportunity to develop some rapport with the clinician because the child's initial impulse may be to deny or to minimize the parents' presenting concerns.

Preferred modes of communication: Is the child open to talking directly about the presenting problems and the significant aspects of his or her life, or is the child more comfortable with indirect modes of expression, such as play or drawing?

Reliable detailed assessment of a child's speech and language capacities, intelligence, academic attainment, attentional and executive functioning, and memory requires standardized psychometric testing. Some simple tasks, such as those used in the Mini-Mental State Examination (as adapted for children by R. A. Ouvrier and colleagues), can be incorporated in a game-like fashion into the interview of the child to provide a rough screen of higher mental functions, such as orientation, attention, memory, language, and constructional ability (Table 36.1–1).

Although the clinician should usually evaluate all areas of the mental status assessment, the emphasis and detail given to the different elements vary with the presenting problem and the type of assessment. For example, in an agitated, potentially delirious child on a pediatric ward, particular attention is paid to the child's orientation, alertness, memory, verbal

coherence, neurological functioning, and the possible presence of hallucinations.

Child Interview Techniques. A range of techniques is used in the psychiatric interview of the child or adolescent. The choice and timing of techniques depend on the child's developmental, cognitive, and linguistic level; the emotional difficulty of the issue being addressed; and the degree of rapport between child and clinician. Being able to gauge the potential usefulness of a technique at a particular moment and feeling comfortable with shifting from one technique to another as the occasion demands are important acquired skills for those conducting child and adolescent clinical assessments.

PLAY TECHNIQUES. Children younger than 7 years of age have limited capacities to recount their feelings or interpersonal interactions verbally. For these younger children, as well as a number of older ones, play is a useful adjunct to direct questioning and discussion and is often a less challenging mode for the child. Some children find it easier to communicate in displacement; thus, imaginative play with puppets, small figures, or the interviewer can provide useful inferential material about the child's concerns, perceptions, and characteristic modes of regulating affects and impulses.

Table 36.1–1.
Mini-Mental State Examination

		Score	Points
Orientation			
1. What is the	Year?	—	1
	Season?	—	1
	Date?	—	1
	Day?	—	1
	Month?	—	1
2. Where are we?	Country	—	1
	State or territory	—	1
	Town or city	—	1
	Hospital or suburb	—	1
	Floor or address	—	1
Registration			
3. Name three objects, taking 1 second to say each. Then ask the patient all three after you have said them. (Tree, clock, boat.) Give one point for each correct answer. Repeat the answers until patient learns all three.		—	3
Attention and calculation			
4. Serial sevens. Give one point for each correct answer. Stop after five answers.		—	5
5. Spell WORLD backward.		—	5
Recall			

6. Ask for names of three objects learned in Q.3. Give one point for each correct answer.	—	3
Language		
7. Point to a pencil and a watch. Have the patient name them as you point.	—	2
8. Have the patient repeat "No ifs, ands, or buts."	—	1
9. Have the patient follow a three-stage command. "Take a piece of paper in your right hand. Fold the paper in half. Put the paper on the floor."	—	3
10. Have the patient read and obey the following: "CLOSE YOUR EYES." (Write it in large letters.)	—	1
11. Have the patient write a sentence of his or her choice. (The sentence should contain a subject and an object and should make sense. Ignore spelling errors when scoring.)	—	1
12. Have the patient copy the design printed below. (Give one point if all sides and angles are preserved and if the intersecting sides form diamond shape.)	—	1
Total		35

From Ouvrier RA, Goldsmith RF, Ouvrier S, Williams IC. The value of the Mini-Mental State Examination in childhood: a preliminary study. *J Child Neurol*. 1993;8(2):145.

The skilled interviewer can facilitate the child's engagement in play without prematurely introducing speculations or reactions that might distort or cut short the presentation of certain types of material. During the course of play, the clinician follows the sequences of play content, noting themes that emerge, points at which a child backs away from the story line or shifts to a new sequence or activity, and situations in which the child gets stuck or falls into a repetitive loop. To facilitate the play component of an interview, the interview room should have a supply of human and animal figures or dolls and appropriate props. These should be relatively simple because elaborate toys can become distractions rather than serve as vehicles for the expression of the child's fantasies. Stock figures (such as Barbie, Star Wars, or Disney characters) may impose specific story lines and thus may limit access to the child's concerns.

Not only is the content of the child's play a rich source of information, but careful observation of the form of play also provides important information for the mental status examination. During imaginative play, the clinician can observe the child's coordination and motor capacities, speech and language development, attention span, readiness to engage the interviewer, capacity for complex thought, and affective state. Absence of imaginative play or very limited, concrete, noninteractive play may indicate an autism spectrum disorder.

Games, such as cards or board games, are also useful for putting the child at ease and developing rapport. These provide opportunities to observe the degree of the child's engagement in and enjoyment of the

shared activity; how the child manages competition, including reactions to winning or losing; or whether the child is prone to cheating. Some play activities (e.g., throwing a ball back and forth or easy card games, such as War) are simple enough to permit ongoing conversation while helping to discharge tension and diminish the pressures of the interview situation. As with toys, elaborate games (e.g., chess) should be avoided; games that demand much cognitive energy and concentration usually preclude discussion of issues relevant to the assessment and may become a means of avoiding issues involved with the child's difficulties.

PROJECTIVE TECHNIQUES. In addition to imaginative play, projective techniques often help provide an indirect picture of concerns that the child may be reluctant or unable to report directly. These techniques can help the child to feel more comfortable with the clinician, are often experienced as fun, and may provide access to concerns that are important to the diagnostic formulation.

One commonly used technique is picture drawing. The child can be asked to draw a picture, leaving the choice of subject up to the child or, alternatively can be given a specific request (e.g., draw a person or the child's family doing something). When the picture is finished or nearly so, it is helpful to compliment the child's effort and to express interest in what is happening in the drawing. The child's elaboration provides information that may not be readily apparent from the drawing itself. The content and form of the drawing offer a window on the child's emotional concerns and aspects of intellectual and visuomotor development. For example, the relative size and placement or omission of family members in a family drawing may be important nonverbal indicators of the child's perceptions or feelings about the family. Aggressive or sexual themes may be reflected more readily in drawings than in words. Self-image may be indicated through depictions of the child as nonhuman, grotesque, inconsequential, or of the opposite gender. The clinician should become familiar with the developmental progression and norms for human figure details, such as limbs, joints, facial features, and clothing, which can provide a useful rough estimate of intellectual maturity. Various systems have been developed for systematically assessing the cognitive and emotional aspects of children's drawings. In addition, the child's behavior and speech while drawing may yield useful information (e.g., throwing the picture away unfinished and saying that it is no good).

Frequently used verbal projective techniques include asking what animal the child would most or least like to be or who the child would pick to take along to a deserted island. It is useful to ask what the child would do with three magic wishes; if elaboration is needed, the clinician can explain that the wishes could be to have anything, to have the world be different in any way, or to change oneself in any way. Children's responses are often

revealing. Some may needily or impulsively wish for material possessions, such as a video game or a billion dollars. Other responses may reveal longings to change distressing circumstances, such as "for my mom and dad to get back together again," "not to have tics anymore or to get teased about them," or "to have a dad who doesn't yell at me all the time." Still other children appear uncomfortable wishing for something for themselves, preferring instead seemingly altruistic responses, such as "no more poverty or wars." Children's responses can be used as the starting point for further exploration. For example, the child who wishes for "a big house and lots of money" may be asked who else would live there and what they would do. Children who wish for "no more fighting in the world" can be asked if there are some particular fights that they would especially like to stop.

Some interactive imaginative techniques useful to interviewers comfortable with them provide elements of playing a game, which may appeal to the child. The squiggle drawing game developed by Donald Winnicott consists of the clinician drawing a curvy line and asking the child to turn it into a picture of something; the child then draws a curvy line that the therapist elaborates, and so on, taking turns. The Despert Fables are a series of incomplete, affectively evocative stories that the child is asked to complete.

Asking the child to tell about a dream or a book, movie, or television program can provide information about the child's interests and preoccupations. (If the clinician is familiar with the plot, the distortions the child may introduce can be informative regarding the child's cognitive and emotional style.) Inquiring about what the child would like to do for a living when grown up can provide insight into the child's aspirations, values, and concerns, as well as those of the family.

DIRECT QUESTIONING. Direct questioning about the current problems and other aspects of the child's life requires tact, attention to the child's level of cognitive and linguistic development, and regard for the effect of the questions on the child's self-esteem. The goal of such discussion is to determine how children see their world and their functioning in it, including attachments and antipathies, pleasures and anxieties, and strengths and weakness, which may include the difficulties that led to the evaluation.

It is important to phrase questions in language that is appropriate to the child's developmental level and verbal capacities. Children may become confused or made more anxious by technical words or abstract or complicated questions. Closed questions often yield little information. For example, "Do you get along OK with your brother?" may bring forth a simple "yes" or "no," whereas open-ended or descriptive questions are more likely to elicit a richer picture of how a child thinks and feels (e.g., "What sort of guy is your brother?" or "Tell me about your last play-date

with Danny.”). Some usefully evocative questions include “What sorts of things make you mad (or make you feel afraid)?” “What is your saddest (or happiest) memory?” “What do you like to daydream about?” If it is important to establish the order or duration of events, school-aged children do better if the questions are anchored to markers in their lives (e.g., “Did that happen before or after your birthday?” or “Have you had trouble with that since school got out?”). Several empirical studies by Michael Rutter and colleagues have examined the effects of various interview styles on the information elicited from children and parents.

Although direct questions are often needed to initiate or to focus discussion, settling into a direct question format may constrict the interview process and give children the impression that the clinician just wants to hear the facts and is not so interested in the child’s feelings and experiences. Young children frequently respond in a manner that they believe to be socially desirable, whereas older children are often reluctant to acknowledge, even to themselves, feelings of sadness or vulnerability. Clinicians should convey, by demeanor and tone of voice, their interest in the child’s real feelings.

When discussion of feelings seems difficult for the child, it may be useful for the clinician to acknowledge the link between a situation and the child’s affective response (“Some kids might have felt mad if that happened to them; how about you?”). However, the clinician must avoid leading a child whom the clinician does not yet know well by assuming experiences or feelings. Probing a bland, equivocal response may prove useful; for example, in response to “Okay, I guess,” the clinician may say, “Sounds like you’re not quite sure.”

The assessment should convey respect for the child and should protect the child’s self-esteem. Many children feel ashamed or guilty about the difficulties that have brought them to the evaluation. When the clinician conveys a nonjudgmental understanding, using language that suggests that the difficulties are problems to be helped rather than transgressions, the child is more likely to be forthcoming. The child has come to the evaluation because of problems, but one should also talk with the child about strengths. For example, in inquiring about school, it may be useful to say, “Tell me about the parts of school you like best” before asking “Are there subjects [classmates, teachers] that give you trouble?” or “What are the rough spots at school?” It may be better to ask the socially isolated child “Do you sometimes have trouble making or keeping friends as much as you’d like?” and then “What do you think the trouble is?” rather than “Do you have any friends?” (especially if the truthful answer to this question is “no”).

Structure and Sequence of the Child Interview. When interviewing the child, the clinician has in mind an overall map of the functional and

historical areas to be covered; however, the format and sequence of any given interview are responsive to the individual child, the presenting problems, the unfolding pattern of the interview, and the clinician's personal style. For example, in every assessment, the clinician should ask about the cardinal symptoms of depression, anxiety, and disruptive behavior disorders, but when, how, and in what order this is done vary from one interview to another.

BEGINNING THE INTERVIEW. There are various ways to start the interview. The one chosen should be comfortable for clinician and child and should set the stage for a productive interchange. Many clinicians prefer not to address the presenting problem(s) immediately at the outset. Some clinicians, for example, remain quiet and let the child look about the room and begin playing or talking without prompting or guidance. This approach imposes a minimum amount of structure; the clinician can see how the child reacts to the materials present in the playroom and to the clinician. It also provides information about how the child handles the situational anxiety that is often evoked by the first meeting with the clinician. However, this unstructured approach may leave many children disoriented and uncomfortably uncertain as to the clinician's goals and intentions. Other clinicians are more direct in attempting to place the child at ease and begin by inviting the child to explore the available play materials ("What looks interesting?"). Others open the interview by inquiring about neutral or pleasurable topics ("Tell me about what you like to do for fun."). The clinician must read the child's cues regarding level of comfort and urgency in deciding when to address the problems that led to the evaluation. Some children prefer to first talk about nonthreatening topics, whereas others expect to get down to business and talk about their problems right away.

Although it may be useful to establish some initial rapport around neutral or interesting topics, most children understand that they have not come for a conventional social encounter. Hence, relatively early in the interview, it is useful to ask what the child's ideas are about the purpose of the meeting. It is usually best to do this in a matter-of-fact manner by asking, "What did your mom and dad tell you [or what do you know] about why they wanted you to come and see me today?" If the child has already met with the clinician together with the parents as part of an earlier contact, the clinician will have asked something similar at that time and may have heard about the child's view of the presenting problem and its history. If so, the clinician can follow up by saying, "Last time, mostly your mom and dad said what they thought about things. Today, we have a chance to talk more about how things look to you. Are there any things we talked about then that we should talk about today?" Asking what the child knows or thinks about the purpose of the evaluation shows the clinician's interest in the child's view of the situation and allows the clinician to hear

how the child describes or assumes responsibility for the problem (or repudiates it as someone else's concern). Asking for the child's view up front, before the child has become deeply engaged or focused on some aspect of the interview process, allows the clinician to address misapprehensions. It also provides an opportunity for the clinician to explain, in terms that can be understood and potentially acknowledged by the child, the issues (at least as the clinician understands them) that have prompted the assessment.

ASSESSMENT OF MAJOR REALMS OF FUNCTIONING. Throughout the assessment, the clinician elicits and assimilates information to develop a picture of the child's functioning in the realms of family, school, peers, and recreation. Discussion of the presenting problem reveals information, particularly with respect to impairments, in these domains; however, additional discussion of the child's experiences in these areas enhances the picture of the child's life and level of adaptation. It is often useful to begin exploration of each of these realms by asking for objective facts (e.g., "Who lives at home with you?" "Whom do you like to play with [or hang out with]?" "What school do you go to? What grade are you in?" or "What do you like to do for fun?"). This can be followed up with probes for more subjective views: "How do people in your family get along? Are there any problems at home? Things people fight about? Things you get in trouble for? Things people at home do that you don't like?" This approach is intended to open up discussion of the major realms of functioning, not simply to elicit rote responses.

It is often helpful to begin by inquiring about the child's recreational activities ("What do you like to do for fun?"). This allows the clinician to learn about the child's interests, preoccupations, and talents and may provide a less threatening or problematic starting point than the areas of peer and academic functioning. Family functioning can be approached by asking who lives at home, how the child gets along with each person, and how family members get along with each other. One can also inquire more specifically "What do you like to do with your dad [or other relative]?" or "What is your brother like?" In the case of nontraditional family structures or families of divorce or separation, the clinician may inquire about the history of the current arrangement and what sort of relationship is maintained with the noncustodial parent(s).

INQUIRY ABOUT PSYCHOPATHOLOGICAL SYMPTOMS. Just as a medical evaluation includes a review of key symptoms in the major systems, the child psychiatric evaluation should include a set of questions that screens for symptoms of common or major disorders. The family or child may not see a link between one behavior or area of distress and another and may not volunteer symptoms that seem unrelated to the presenting problem. Comorbidity is relatively common in children with psychiatric disorders.

Thus, the child psychiatric clinician must routinely question across the major areas of psychopathology. Using developmentally appropriate language, this inquiry should cover depressive symptoms (e.g., low self-esteem, anhedonia, and suicidal ideation or behavior), self-harm (e.g., cutting), excessive anxiety or fears, obsessions and compulsions, hallucinations and delusions, and antisocial behaviors. For example, one should ask "Do you feel sad or grumpy a lot of the time? Do you often feel like crying?" If the answer is positive, one can follow up concerning persistence ("Does that last for more than a few minutes?"), frequency ("Does that usually happen every week [or most days]?"), and context ("What sorts of things make you feel that way?"). To inquire about suicidal ideation, one can begin by asking, "Do you ever feel that life is so hard, you wish you weren't alive [or would be better off dead]?" following up with the more explicit "Have you ever tried to hurt yourself [or done anything to try to make yourself dead]?" Exposure to potentially traumatic experiences should be routinely covered, and developmentally appropriate language should be used (e.g., "Have any really scary things ever happened to you?").

MENTAL STATUS EXAMINATION. Throughout the clinical interview, the clinician observes and notes information relevant to the domains of the mental status examination. Some abnormalities or deficits may be evident through this process; however, other areas require direct examination, which can usually be incorporated in a playful way into the clinical interview. Although intensive developmental, psychometric, language, and neurological assessment usually cannot be subsumed within the clinical interview, a brief semistructured evaluation requires only approximately 5 to 10 minutes and provides useful information. This generally involves drawing, copying a design, and engaging in simple tests of orientation, recall, general knowledge, and language (e.g., naming three objects; repeating a phrase such as "No ifs, ands, or buts"; a three-stage command, such as "Take a piece of paper in your right hand, fold the paper in half, put the paper on the floor"; a writing sample [such as asking the child to write a sentence of his or her choice]; and a brief screen of neurological soft signs).

CLOSING THE INTERVIEW. In winding up the interview, the clinician can check for concerns that have not been covered and that also signal that the interview is coming to a close by asking, "We've talked about a lot of things; are there any other things that would be important for me to know about you?" Similarly, one might ask, "I've asked you a lot of questions; do you have any questions you want to ask me?" The clinician should also give the child a general idea of what will be done with all the information the child has shared and some sense of what happens next. For example, one might explain, "I'm going to think over what I've heard from you and your folks and meet again with them (and you) to talk about what might help with—.

Are there any particular things I should be sure to mention?" If the clinician already has a sense of what interventions might be helpful, this may be tentatively broached with the child, keeping in mind the parents' right to hear and decide on those recommendations.

SPECIAL ISSUES IN THE ASSESSMENT OF ADOLESCENTS

The clinical assessment of adolescents requires familiarity with the distinctive dynamics of adolescents and their characteristic patterns of managing conflicts, anxiety, and interactions with adults.

The adolescent's attitude toward the evaluation has important implications for how the assessment process unfolds. In contrast to the adult patient (who usually takes the responsibility and initiative for seeking help) and the child patient (who is simply brought by parents), many adolescents are cautiously ambiguous as to whether there is a problem with which they want help or whether they are merely passively complying with their parents' initiative.

One important initial decision concerns who is to be seen in the initial interview: the parents, the adolescent, or all together. Seeing adolescents alone first has the advantage of emphasizing their active participation in the process and underlines the clinician's interest in their perspective; however, it may mean that crucial elements of the history are not available to the assessor. Furthermore, some embattled youngsters may protest that they are being identified as the patient or bearer of a problem that they feel resides in their parents or in the family as a whole. The clinician may need to explain explicitly that the evaluation is not a judicial proceeding to assess blame but rather an attempt to understand what the problems are.

In some cases, when the adolescent balks at being seen at all, the clinician may need to meet with the parents alone to facilitate the youngster's participation in the evaluation. Regardless of when the adolescent is seen, the parents should, at some point, be seen alone to explain their concerns about the child, to obtain a full developmental and family history, and to hear about aspects of their spousal and parenting relationship that may be clinically significant but that may not be appropriate to discuss initially in the adolescent's presence.

Meeting with the adolescent and parents together is usually helpful. This provides an opportunity to explore how family members interact, their convergences or divergences of perspective, and their capacity to work together therapeutically. Such conjoint interviews usually require active structuring by the clinician, so that the purpose is framed as an occasion to talk together. Without such active intervention, the family session can easily deteriorate into an adversarial proceeding in which the parents complain about the adolescent, who, in turn, retreats into hostile or guarded silence.

Confidentiality

The assessment of adolescents often raises complex issues of confidentiality. Although the parents and the adolescent must know that the adolescent's communications are kept confidential, they must all understand from the onset that this does not extend to situations that pose a clear danger to the patient or others. For example, if the adolescent reveals significant suicidal ideation or a recent suicide attempt of which the parents are unaware, parents must be informed for the youngster's safety. Having made clear that the parents need to know, the task is then to negotiate whether the adolescent prefers to tell the parents directly or to have the clinician inform them. It is usually best to have the adolescent undertake this responsibility, but sometimes, after discussion with the youngster, the clinician must be the one to speak with the parents, preferably in the adolescent's presence. Less clear-cut situations in which the adolescent reveals problematic but not imminently dangerous antisocial or sexual behavior or substance use must be decided on an individual basis. Treatment may not be possible without some confidentiality; however, confidentiality must not become collusion.

Developmental Issues

Adolescents' usual strong push toward autonomy is often linked to a wariness of feeling vulnerable, dependent, or controlled. Tolerating feelings of ambivalence, guilt, or other painful affects or internal conflicts does not come easily to adolescents. In addition, adolescents are prone to externalize their concern or embarrassment over self-perceived shortcomings or difficulties. As a result, even adolescents who consciously desire help often approach a clinical assessment with the fear that revealing what they perceive as shameful weaknesses leaves them vulnerable to criticism, control, or regressive dependency. Fear of acknowledging vulnerability or dependency may result in bland denials of any problems or insistence that "I can handle it by myself" or "I don't need a shrink; I can talk about my problems with my friends." Externalizing their conflicted feelings about autonomy versus dependence, adolescents may fluctuate between grievances about parental overprotectiveness and complaints about perceived lack of care. Reversal of affect and counterphobic defenses may mask anxiety, shame, or other painful affects. The skilled clinician learns to look beyond surface behavior and to notice when an adolescent "doth protest too much."

Although many adolescents seem intolerant of self-reflection, an unrealistic belief in the omnipotence of thought often helps to support denial and avoidance. Even when confronted with a pattern of long-standing maladaptive behaviors, the adolescent may insist "I can stop any time I really want to" or "I don't do that any more" (i.e., not since

Saturday).

The lability of adolescents' moods and their truncated time perspective often make it useful to extend the evaluation over time to distinguish the transient issues from those that are more enduring.

Technical Issues in the Adolescent Interview

Interviewing adolescents requires tact, flexibility, a sense of humor, and clinicians who have come to terms with their own adolescent issues. Condescension, sarcasm, competitiveness, or excessive passivity on the part of the interviewer is likely to result in failure. However, although conveying a genuine interest in the youngster is essential, seductiveness, overfamiliarity, or the failure to maintain an adult role is likely to frighten the adolescent. The interviewer should resist the temptation to be too "with it" or too "cool"—nothing is more dated than last year's slang. Because adolescents can so easily feel defensive, criticized, or embarrassed, the interviewer must learn to talk frankly yet tactfully about areas of apparent difficulty. Many embattled adolescents see adults as either for them or against them. Hence, the clinician needs to convey a genuine empathic interest in the adolescent's view of a situation without implying collusive, uncritical acceptance of it or entering into an adversarial dispute.

The general observation that it is important to learn about youngsters' strengths and interests, as well as their problems and vulnerabilities, is especially true of adolescents. Without appearing squeamish or avoidant of the problem areas, the clinician should inquire about the adolescent's interests, hobbies, recreations, and friendships. Knowledge about the areas of accomplishment, talent, and gratification is required for a full picture that includes the youngster's adaptive and compensatory strengths and assets. Furthermore, such an inquiry helps to facilitate rapport and conveys an interest in the adolescent as a whole person rather than simply a bearer of problems. A too-exclusive focus on pathology puts many youngsters on the defensive, and an opportunity to talk about more neutral areas or those of greater mastery or confidence helps the adolescent to feel less on the spot. Instead of the interviewer straining to demonstrate familiarity with the latest band or winning team, letting adolescents teach the interviewer about the details of their interests lets the youngsters enjoy a sense of mastery, control, and a degree of parity with the adult interviewer that helps to put them at ease. Indeed, what Hilde Bruch called the *constructive use of ignorance* is often a useful technical element of the adolescent interview. This consists in being able to ask about some missing piece in the story or about some puzzling impasse with an ingenuous curiosity (rather than prosecutorial zeal) that invites the adolescent's self-reflection.

Table 36.1–2.

Developmental Growth Tasks of Adolescence

1. Gradual development as an independent individual
2. Mental evolution of a satisfying, realistic body image
3. Harnessing of appropriate control and expression of sexual drives
4. Expansion of relationships outside of the home
5. Implementation of a realistic plan to achieve social and economic stability
6. Transition from concrete to abstract conceptualization
7. Integration of a value system applicable to life events

Reprinted with permission from Gross RT, Duke PM. Adolescence. In: Levine MD, Carey WB, Crocker AC, Gross RT, eds. *Developmental Behavioral Pediatrics*. Philadelphia, PA: WB Saunders; 1983.

Although interviewing and taking the history of the adolescent cover the same general areas outlined previously, certain domains are of particular salience with the adolescent. These relate to the core developmental tasks of adolescence: developing increased autonomy and individuation from parents; taking over the tasks of physical self-care; developing a satisfying, realistic body image in the face of pubertal changes; mastering burgeoning sexual and aggressive impulses; consolidating an identity, including ethical values and vocational goals; developing a network of satisfying peer relations; and in later adolescence, developing a capacity for romantic intimacy with an appropriate partner (Tables 36.1–2 and 36.1–3).

In the realm of values, ideals, and aspirations, the clinician should explore the adolescent's values and the models whom the adolescent wishes to emulate or to reject. To what extent do these values complement or conflict with those of the parents? What expectations and aspirations (realistic or not) does the adolescent have concerning the future?

The world of peers and friends should also be explored with adolescents. With whom do they hang out? What crowd do they see themselves part of? How do they get along with peers, and what do they do for fun? Asking teenagers to describe a close friend provides a chance to learn how they think about people and relationships and to assess their capacity for empathy.

In this context, it is also important to inquire about the adolescent's use of the Internet and social media. How much time does the youngster spend on line or on a mobile device? What is the balance of in-person interactions with local acquaintances versus purely electronic ones? Has the youngster had any concerning experiences online, such as cyber-bullying or sexting?

The topic of friends leads naturally to the topic of romantic and sexual interests and relationships. Exploration of this area requires tact, and the clinician cannot always be sure of receiving a candid response. Nonetheless, as with questions concerning depression or substance use, the clinician's tone can convey the sense that the inquiry into these sensitive areas is motivated by a genuine and legitimate wish to understand the adolescent. One can begin by asking whether the teenager dates and

whether there is anyone of either sex with whom the youngster is close. If so, who are they, and what attracted the adolescent to them? Are the adolescent's interests reciprocated? How has the relationship gone? Does the teenager note any recurrent patterns in terms of the type of person to whom he or she is attracted or how such relationships have unfolded in the past? Have any of the teenager's relationships developed into sexual ones? Has the adolescent had other sexual experiences? (Here one must be open to hearing about possible episodes of sexual abuse or concerns about sexual orientation.)

Because of the high prevalence of depression, self-harm, suicidal ideation, and even suicide attempts in the general population, specific inquiry about these areas as outlined previously is particularly important. The clinician should also inquire matter-of-factly about whether the adolescent smokes, drinks, or uses drugs illicitly, and, if so, with what frequency, in what amounts, in what contexts, and with what consequences.

ASSESSMENT OF INFANTS AND TODDLERS

The assessment of infants and toddlers poses special challenges for the clinician. First, the evaluation of young children requires specific and detailed familiarity with the cognitive, linguistic, motor, and adaptive developmental landmarks from birth to 6 years of age, as well as the specialized techniques needed to assess these competencies. Second, although direct observation of the child is an important source of data, the preverbal or limited verbal capacities of the child severely restrict the historical material that can be obtained from the direct interview of the child. Third (and perhaps most important), infants and young children cannot be assessed, diagnosed, or treated apart from their caregiving environment, which includes not only the parents but also other members of the extended family and child care and other care providers.

Table 36.1–3.
Growth Tasks of Adolescence by Developmental Phase

Tasks	Early: 10–13 Years of Age	Middle: 14–16 Years of Age	Late: 17 Years of Age and Older
1. Independence	Emotionally breaks from parents, and prefers friends to family	Ambivalence about separation	Integration of independence issues
2. Body image	Adjustment to pubescent changes	Trying on different images to find real self	Integration of a satisfying body image with personality
3. Sexual drives	Sexual curiosity, occasional masturbation	Sexual experimentation, opposite sex viewed as sex object	Beginning of intimacy and caring
4. Relationships	Unisexual peer group,	Begin heterosexual peer	Individual relationships more

	adult crushes	group, multiple adult role models	important than peer group
5. Career plans	Vague and even unrealistic plans →		Specific goals and specific steps to implement them
6. Concept-ualization	Concrete thinking →		Ability to abstract
		Fascinated by new capacity for thinking	
7. Value system	Drop in superego, testing of moral system of parents	Self-centered →	Idealism, rigid concepts of right and wrong, other oriented, asceticism

Reprinted with permission from Gross RT, Duke PM. Adolescence. In: Levine MD, Carey WB, Crocker AC, Gross RT, eds. *Developmental Behavioral Pediatrics*. Philadelphia, PA: WB Saunders; 1983.

The developmental tasks of infancy include achieving physiological homeostasis; developing stability of state and learning to negotiate transitions between states; organizing basic physiological rhythms of wakefulness and feeding; coping with novel stimuli; perceiving coherent patterns of experience; mastering simple sensorimotor schema; establishing a specific and secure relationship with a caretaker; maintaining a positive affective tone; developing expressive and receptive channels of communication (including vocalizations and facial and gestural cues); and forming schemas of causality, intentionality, object permanence, and affect. The older infant or toddler must grapple with regulating and socializing bodily functions (e.g., eating, toileting, and sleeping), learning to concentrate and regulate shared attention, acquiring the capacity for imaginative play and more complex forms of imitation, dealing with the challenge of socialization (including development of more complex forms of social perception, communication, and empathy and negotiating conflicts of will with caregivers and peers), regulating negative and positive mood and aggression, and forming core gender identity. Development can proceed at a different pace in each area and may go awry or be delayed in any or all of them.

The goal of the developmental assessment of infants and toddlers is to depict their functional capacities in each of these areas, the relationships among the various domains of their development, their ability to adapt, and the range of coping strategies that they use in doing so, as well as the caregiving environment as it affects them, the quality of caregiving interactions, and the effect of their developmental status on the caregiver. As with older children, the purpose of the assessment is to define what, if any, interventions are necessary, while at the same time helping to form an alliance with the parents that facilitates implementation of any recommendations that may come out of the assessment. The assessments process itself, properly done, increases parents' sensitivity, awareness, and reflectiveness about their child's development.

The most common referral problems for infants and toddlers involve the following:

1. Developmental delays (e.g., specific delays, such as delayed, absent, or deviant motor or speech development, or general delays in multiple areas of development).
2. Psychophysiological dysregulation (fussy behavior, disturbances of sleep or feeding, rumination, eczema, self-stimulatory behaviors [e.g., head banging, rocking], and failure to thrive).
3. Behavioral disturbances (excessive tantrums, excessive negativism, aggression, and hyperactivity).
4. Disturbed social development (lack of apparent awareness of or interest in others, response to abuse, neglect, or multiple placements or repeated or prolonged separations).
5. Medical or genetic problems (e.g., prematurity and genetic syndromes), exposure to toxins (such as lead or fetal alcohol syndrome).
6. Exposure to environmental stressors (such as maternal depression, change of caretaker, domestic violence or other trauma, reaction to transitions, e.g., beginning preschool or out-of-home child care, parental marital separation).

Because so many aspects of early development are interdependent, the significance of difficulties in one functional domain cannot be understood without reference to the child's developmental status in other areas. For example, speech, language, and hearing difficulties profoundly affect personal and social development. Assessment of motor or adaptive competency that involves interaction with the caregiver or assessor is affected by disturbances in relatedness and hence yields findings that appear erratic or uneven. Infants and toddlers are particularly sensitive to environmental stresses, which may delay the appearance of phase-appropriate skills or may lead to transient regressions.

Technical Issues in the Development of the Infant and Young Child

The assessment of the young child usually requires several sessions. In addition to sessions spent with the parents alone to gather history and to present the results of the assessment, familiar adults must be present during the assessment of the child; this is necessary to place the child at ease in the assessor's presence and to permit direct observation of child and caregiver interaction. Furthermore, the behavior of infants and young children is highly state dependent (and sensitive to fatigue or illness), and one usually needs to see the child on more than one occasion to have a sufficient sample of characteristic behavior.

The general principles of history taking from the parents and assessment of the caregiving environment follow those outlined previously. However, particular emphasis is given to the parental perception of the

pregnancy with the child—whether the conception was planned or came at a good time in the life of the family and what the effect of the child's arrival was on the parents individually, on older siblings, and on the family as a whole. The history should explore parental (and grandparental) expectations for the child, the meaning that the child has for them, whom the infant reminds them of, and what they perceive as the child's best and worst traits.

The developmental history for the young child includes prenatal, perinatal, and postnatal complications, including maternal illness, medication and substance use, and neonatal status. In addition to the areas outlined previously, assessment of the child's behavioral organization should include the regulation of physiological state, alertness, and activity patterns. Children's ability to console themselves and to adapt to new situations and their response to separations are also key areas of evaluation. The history should also include the history of child care arrangements (nannies, relatives, out-of-home child care) and the child's reactions to new care environments or changes in these arrangements. (Many young children spend about as many waking hours in the care of a paid caregiver as they do with their parent(s), hence, in many cases, it may be very useful to hear the report of these extra-familial caretakers.)

The child's mastery of various gross motor, fine motor, language, adaptive, and personal-social milestones should be systematically assessed by history and in the direct evaluation of the child. In taking such a history, the assessor may be aided by a variety of screening instruments, such as the Vineland Adaptive Behavior Scales II. It is useful to inquire about the timing of the child's mastery of these milestones relative to those of any siblings; this also helps to gauge parental expectations and to learn more about sibling dynamics. The Vineland scales provide a semistructured interview format that can be easily woven into a developmental history.

Within the first few months of life, children begin to manifest distinctive and often enduring behavioral styles or temperament. Table 36.1–4 shows the nine temperamental categories and three distinctive constellations of temperament as defined by Stella Chess and Alexander Thomas. The goodness of fit, or degree of consonance between the child's temperament and the parents' expectations or demands of the child, is an important predictor of future development and parent–child interactions.

Particularly salient aspects of the parental interview include a detailed history of the parents' childhoods and history of care, out-of-home placement, separations, and experience in their families of origin. The effect of the child on each parent and their marital relationship should be assessed, as well as the appropriateness and degree of congruity in their respective expectations of the child.

Adequate caregiving for the young child requires a delicate balance among gratification, comfort, support, and frustration that is appropriate

to the child's developmental phase. It is important to hear (and to observe) which aspects of parenting and which features of the child's behavior the parent feels most comfortable with or most challenged by. The clinician should assess the parents' overall level of psychological functioning; the parental difficulties that adversely affect a young child's development do not always involve such gross failures as serious neglect or abuse but may involve problems more difficult to identify, such as persistent maternal or paternal depression. To help obtain a concrete picture of the parent, child, and their interaction, it is useful to ask for a detailed description of a day in the life of the child. In addition to providing data regarding the child's capacities, responses, and temporal organization, such a description provides clues to the parents' attunement and affective responses to the various elements of the child's behavior.

Table 36.1–4.
Temperament Categories

Activity level: The motor component present in a given child's functioning and the diurnal proportion of active and inactive periods.

Rhythmicity (regularity): The predictability or unpredictability in time of any function; it can be analyzed in relation to the sleep–wake cycle, hunger, feeding pattern, or elimination schedule.

Approach or withdrawal: The nature of the initial response to a new stimulus, be it a new food, a new toy, or a new person. Approach responses are positive, whether displayed by mood expression (e.g., smiling and verbalizations) or motor activity (e.g., swallowing a new food, reaching for a new toy, and active play). Withdrawal reactions are negative, whether displayed by mood expression (e.g., crying, fussing, grimacing, and verbalizations) or motor activity (e.g., moving away, spitting new food out, and pushing new toy away).

Adaptability: Responses to new or altered situations. One is not concerned with the nature of the initial responses, but with the ease with which they are modified in a desired direction.

Threshold of responsiveness: The intensity level of stimulation needed to evoke a discernible response, irrespective of the specific form of the response or the sensory modality affected. The behaviors used concern reactions to sensory stimuli, environmental objects, and social contacts.

Intensity of reaction: The energy level of response, irrespective of its quality or direction.

Quality of mood: The amount of pleasant, joyful, and friendly behavior (as contrasted with unpleasant, crying, and unfriendly behavior).

Distractibility: The effectiveness of extraneous environmental stimuli in interfering with or altering the direction of the ongoing behavior.

Attention span and persistence: *Attention span* concerns the length of time a particular activity is pursued by the child. *Persistence* refers to the continuation of an activity in the face of obstacles to the maintenance of the activity direction.

Qualitative analysis, supplemented by factor analyses, led to the formulation of, in addition, three constellations of temperament made up of various combinations of the individual categories, which appeared to have functional significance. These are as follows:

Easy temperament: The combination of biological regularity, approach tendencies to the new, quick adaptability to change, and a predominantly positive mood of mild or moderate intensity. This group comprised approximately 40 percent of the study population.

Difficult temperament: The opposite of easy temperament, namely, biological irregularity, withdrawal tendencies to the new, slow adaptability to change, and frequent negative emotional expressions of high intensity. This group comprised approximately 10 percent of the study population.

Slow-to-warm-up temperament: Withdrawal tendencies to the new, slow adaptability to change, and frequent negative emotional reactions of low intensity. Such individuals are often labeled shy. This group comprised approximately 15 percent of the study population.

Adapted from Chess S, Thomas A. Temperament and its clinical applications. In: Lewis M, ed. *Child and Adolescent Psychiatry*. 3rd ed. Baltimore, MD: Lippincott Williams & Wilkins; 2002:221.

Although much of the clinical observation is focused on the child's behavior and capacities, direct observation of the parent–child relationship is essential. Inviting the family to interact in the playroom as they usually would at home allows observation of their interaction in unstructured parent–child or family play; assigning various structured tasks may also be useful. Observing the child in interaction with each parent separately helps to reveal relationship-specific patterns of behavior. Among the dimensions of parental behavior to be assessed are the ability to engage the child (verbally and nonverbally); the quality and affective tone of this engagement; the parent's attunement to the child's cues; parental vigilance, protectiveness, and limit setting; the parent's facilitation or restriction of autonomous play by the child; thematic content; and the degree of mutual pleasure during play. Instruments and methodologies exist for formally assessing parental responsivity and parent–young child interaction, such as the Home Observation for Measurement of the Environment (HOME) and the Parent–Child Early Relational Assessment (PCERA). Some of these use standardized protocols or structured video recording and analysis. Some yield impressionistic global scores, whereas others provide detailed frame-by-frame coding of parent–child interactions. Although these methods are used primarily for research purposes, familiarity with the various facets of such instruments is clinically useful.

Observation and Examination of the Infant and Toddler

The direct examination of the young child draws on three levels of observation: The child's response to structured assessment items during formal testing, the child's response to other interactions with the assessor and the assessment environment, and the spontaneous interactions of the child and parent in a naturalistic setting. When observation of the child and parents takes place in the clinic, it is essential to remember that the unfamiliar setting may elicit behavior that is not characteristic of the home setting.

Structured Assessment Instruments. A variety of instruments exists for the structured assessment of infants and young children, and each has somewhat different goals, theoretical orientation, and psychometric properties (Table 36.1–5). These instruments do not yield diagnoses; they detail the child's development in various areas relative to a normative population. Some of these developmental tests are screeners designed to

help pediatricians or trained paraprofessionals identify children with significant motor, social, or language delays who need fuller evaluation, whereas others are complete batteries evaluating a full range of developmental competencies. For example, the Bayley Scale of Infant and Toddler Development III (Bayley-III), which is administered by a trained assessor, can be used to assess children 1 to 42 months of age; it includes scales yielding standard scores for cognitive abilities, gross and fine motor skills, receptive and expressive communication, and self-help competencies, as well as a behavior rating scale for assessing qualitative aspects of the child's behavior during the assessment.

Table 36.1–5.
Selected Formal Tests of Infant and Toddler Development

	Age	Domains[a]	Norm Sample	Reliability/Validity	Com
Neonatal Assessment Tests					
Graham/Rosenblith Behavioral Test for Neonates		M, TA, V, A, MT	Not applicable	Encouraging early findings	Curr s
Brazelton Neonatal Behavioral Assessment Scale—2	37–44 weeks GA	NI, B, PS	Not applicable	High interrater reliability, weak test-retest, poor predictor beyond first year	Mos u n
Infant/Toddler Development Tests					
Bayley Scales of Infant Development–III	1–42 months	Mental, M, B	Large, representative	Excellent for mental, adequate for psychomotor, varied for behavior	Mos u v ir
Mullen Scales of Early Learning	0–68 months	GM, FM, VR, RC, EC	Large, representative	Adequate to exceptional	Mos fo a p
Battelle Developmental Inventory	0–95 months	PS, Ad, GM, FM, EC, RC, Cg (and sub-domains)	Small, representative	Numerous psychometric concerns cast significant doubts (see text)	Use c d m fl
Griffiths Abilities of Babies	0–24 months	Lm, PS, HS, EH, P	Small, all from London	Mixed results	Mos E
Infant Psychological Development Scales	NA	See text	Not applicable	Adequate to exceptional reliability	Mos P ir
Screening Tests					

Battelle Developmental Inventory Screening Test	6–95 months	PS, Ad, GM, FM, EC, RC, Cg	Small, representative	Strong correlation with full BDI	Suff li fu
Bayley Infant Neurodevelopmental Screen	3–24 months	N, RC, EC/M, Cg	Large; representative	Excellent reliability; promising validity	
Birth to Three Developmental Scale	0–36 months	EC/RC, Cg, PS, M	Small; questionable representation	Strong interrater reliability, but little evidence of validity	
Child Development Inventory	15–72 months	GM, FM, EC, RC, Ad, PS, L, N	Small; from St. Paul, MN	Little evidence of reliability and validity	300 r
Denver Developmental Screening Test—II	0–72 months	GM, FM/Cg, PS, EC/RC	Large, all from Denver area	See text	See
Developmental Activities Screening Inventory—II	0–60 months	15 sensory and problemsolving areas	>200 disabled children	Little evidence of reliability; valid for severely delayed	Can w la a ir c
Developmental Indicators for the Assessment of Learning—Revised	24–72 months	M, AS, EC/RC, B	Large; representative	Acceptable reliability and validity	One s a
Diagnostic Inventory for Screening Children—4	0–60 months	FM, GM, RC, EC, AM, VM, Ad, PS	Small; all from southwest Ontario	Excellent reliability; limited validity	
Developmental Observation Checklist System	0–72 months	Cg, EC/RC, PS, FM/GM, Adj, PS&S	Adequate; representative	Sound reliability and concurrent validity	Pare o
Developmental Profile —II	0–114 months	M, Ad, PS, AS, EC/RC	Large	Adequate reliability and validity	Pare o
Early Screening Profile	24–72 months	Cg, EC/RC, M, Ad/PS, Ar	Adequate; representative	Good reliability; exceptional validity	One s a
Kent Infant Development Scales	0–12 months	Cg, M, EC/RC, Ad, PS	All from Northeast Ohio	Adequate reliability and validity	Prim p r

[a]A, auditory responsiveness; Adj, adjustment; Ad, adaptive, self-help or daily living; AM, auditory attention and memory; Ar, articulation; AS, academic or pre-academic skills; B, behavior; Cg, cognitive or problemsolving; EC, expressive communication/language; EH, eye-hand coordination; FM, fine motor; GM, gross motor; HS, hearing and speech; L, letters; Lm, locomotor; M, motor or physical; MT, muscle tone; N, numbers; Nl, neurological intactness; P, performance; PS, personal-social; PS and S, parental stress and support; RC, receptive communication/language; TA, tactile-adaptive; V, visual responsiveness; VM, visual attention and memory; VR, visual reception. A slash mark (/) indicates that multiple domains are assessed in the same scale or subtest.

From Gilliam WS, Mayes LC. Clinical assessment of infants and toddlers. In: Martin A, Volkmar FR, eds.

Lewis's Child and Adolescent Psychiatry: A Comprehensive Textbook. 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2007:309–322.

Although tests such as the Bayley-III may show good reliability and concurrent validity, their ability, when administered to children younger than 3 years, to predict later performance on intelligence quotient (IQ) assessments or later adaptive functioning is highly variable. Among the reasons for this weakness of prediction are sources of error, such as the rapid pace of development during this period and the child's susceptibility to state or environmental effects (e.g., fatigue and unfamiliar test setting), as well as the intervening effects of social and family environment and the heavy emphasis infant tests place on perceptual and motor skills that may have relatively little to do with information-processing capacities. However, very low scores on repeated measures are highly suggestive of continuing difficulties.

The mental status examination of the infant and young child can be organized using a schema such as that shown in Table 36.1–6.

Diagnostic Formulation of the Assessment of Infants and Young Children

Although intellectual disability and autism spectrum disorders can be at least tentatively diagnosed in infants and toddlers, the most widely used official diagnostic system—the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) and the tenth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10)—generally lack a developmental perspective and diagnostic criteria or categories suitable for infants and young children. The Zero to Three/National Center for Clinical Infant Program's *Diagnostic Classification of Mental Health and Development Disorder of Infancy and Early Childhood, Revised Edition* (DC:0–3R) provides a diagnostic schema for this age group that complements the DSM.

The appropriate outcome of a thorough clinical assessment of the young child thus extends far beyond an attempt at a specific categorical diagnosis. The diagnostic formulation should provide a detailed picture of the strengths and weaknesses in the child's social, emotional, cognitive, linguistic, motor, and perceptual development; the quality of the child's temperament, attachment, and regulatory capacities; hypotheses regarding the etiology of the child's difficulties; and potential exacerbating or risk factors and protective or ameliorative factors in the child and family. This formulation, which should be shared with the family in a comprehensible and usable form, should provide the basis for recommendations regarding intervention or further assessment.

SPECIAL TYPES OF ASSESSMENT

This section describes the comprehensive clinical assessment of the child.

Although the clinician must be aware of the many facets of history taking, interview, and observation described here, certain clinical settings or tasks may call for a narrower focus of inquiry or distinctive emphasis for the evaluation. For example, an urgent situation may require therapeutic intervention before a comprehensive assessment can be completed; in such cases, the immediate evaluation focuses on defining the crisis and selecting the most suitable immediate intervention. Thus, the emergency room or crisis center evaluation focuses on precipitants and extent of the current decompensation, predisposing vulnerabilities, the exacerbating factors affecting the child and family and the ameliorative resources available to the family in such situations, the risk of danger to self or others, and other issues germane to making an immediate disposition (including assessing the potential need for hospitalization). The assessment of suicidal ideation, threats, or attempts or aggressive outbursts is perhaps the most common clinical context for such urgent evaluations. Children exposed to acute trauma (e.g., natural catastrophe or violence) are often screened outside of the traditional clinic setting in a combination of emergency assessment and crisis intervention. The noncompliant, acutely agitated, or delirious child in the pediatric setting raises yet another set of assessment issues (Table 36.1–7).

Various specialized consultative purposes also require distinctive techniques or procedures or specific areas of inquiry. (The American Academy of Child and Adolescent Psychiatry has prepared valuable practice parameters providing guidance for assessment and treatment of children and adolescence in forensic settings, detention or correctional facilities, schools, community systems of care, and pediatric settings, as well as youngsters presenting with certain specific conditions, such as substance abuse, suicidal behaviors, aggression, exposure to trauma, etc.)

For example, forensic evaluations raise special issues of role definition, confidentiality, and privilege and require familiarity with the legal issues involved. Clinicians must be clear about whether they are serving as neutral impartial advocates for the best interests of the child at the behest of the court or as consultants to one party or another. Although the clinician treating or assessing the child for clinical purposes may have information valuable to a forensic evaluation, it is usually in the best interests of the therapy and the legal proceedings for the evaluator to be someone other than the therapist.

The assessment of children or adolescents who may have been physically or sexually abused is a common form of forensic evaluation. In addition to assessing the child's competency and credibility, the examiner may also be asked for recommendations regarding whether the child should testify in court (and, if so, how) and whether removal of the child or alleged perpetrator from the home is needed, as well as any therapeutic interventions that may be required. A forensic evaluation requires careful

documentation of parent and child interviews. The suggestibility of younger children raises the danger of false accusations and false denials and requires meticulous safeguards against repetitive, suggestive questioning.

Child custody evaluations also require careful definition of the assessor's role with all concerned. Specific attention must be given to the best interests of the child in terms of continuity and stability of care, the quality of parent–child attachment and attunement, the child's stated preferences, the mental health of each parent, and the level of parental conflict and its effect on the child. Consultations for the juvenile court regarding youthful offenders or children in need of services and probate or commitment hearings represent other forms of forensic evaluation. Other specialized forms of evaluation include medication consultations and consultation to day care centers, schools, or pediatric specialty clinics (e.g., pain or oncology clinics).

STANDARDIZED ASSESSMENT INSTRUMENTS FOR CHILDREN AND ADOLESCENTS

While psychosocial and environmental factors must inform clinical decision-making, the systematic assessment of symptoms, behavior and impairment is an essential part of clinical evaluation. Numerous tools—diagnostic interviews and rating scales—have been developed that elicit comprehensive information about psychiatric symptoms and diagnoses. The use of standardized instruments provides some protection against the risks to which traditional clinical approaches may be prone, for example, tendencies to make diagnoses too quickly, collecting information selectively that confirms initial impressions, ignoring contradictory information, and applying diagnostic criteria idiosyncratically. The relatively structured diagnostic measures are likely to yield more reliable diagnoses than those made from less systematic approaches, and these features in turn improve the likelihood of obtaining meaningful findings from research efforts, which depend on assessment of diagnoses in a consistent manner across a number of study participants.

Table 36.1–6.
Infant and Toddler Mental Status Examination by Anne L. Benham, M.D.

I. Appearance
   Size, level of nourishment, dress and hygiene, apparent maturity compared to age, dysmorphic features (e.g., facies, eye and ear shape and placement, epicanthal folds, and digits), abnormal head size, cutaneous lesions.
II. Apparent reaction to situation
   Note where and with whom the evaluation takes place.
   A. Initial reaction to setting and to strangers: Explores; freezes; cries; hides face; acts curious, excited, apathetic, or anxious (describe).
   B. Adaptation

1. Exploration: When and how child begins exploring faces, toys, stranger.
2. Reaction to transitions: From unstructured to structured activity; when examiner begins to play with infant; cleaning up; leaving.

III. Self-regulation

A. State regulation: An infant's state of consciousness ranges from deep sleep through alert stages to intense crying. Predominant state and range of states observed during session, patterns of transition (e.g., smooth vs. abrupt), capacity for being soothed and self-soothing, capacity for quiet alert state. (Some of these categories also apply to toddlers.)
B. Sensory regulation: Reaction to sounds, sights, smells, light and firm touch; hyperresponsiveness or hyporesponsiveness (if observed) and type of response, including apathy, withdrawal, avoidance, fearfulness, excitability, aggression, or marked behavioral change; excessive seeking of particular sensory input.
C. Unusual behaviors; mouthing after 1 year of age; head banging; smelling objects; spinning; twirling; hand flapping; finger flicking; rocking; toe walking; staring at lights or spinning objects; repetitive, perseverative, or bizarre verbalizations or behaviors with objects or people; hair pulling; ruminating; or breath holding.
D. Activity level: Overall level and variability (note that toddlers are often incorrectly called hyperactive). Describe behavior; for example, squirming constantly in parent's arms, sitting quietly on floor or in infant seat, constantly on the go, climbing on desk and cabinets, exploring the room, pausing to play with each of six to eight toys.
E. Attention span: Capacity to maintain attentiveness to an activity or interaction, longest and average length of sustained attention to a given toy or activity, distractibility. Infants: Visual fixing and following at 1 mo of age, tracking at 2–3 months of age, attention to own hands or feet and faces, duration of exploration of object with hands or mouth.
F. Frustration tolerance: Ability to persist in a difficult task, despite failure; capacity to delay reaction if easily frustrated—e.g., aggression, crying, tantrums, withdrawal, and avoidance.
G. Aggression: Modes of expression, degree of control of or preoccupation with aggression, appropriate assertiveness.

IV. Motor

Muscle tone and strength, mobility in different positions, unusual motor pattern (e.g., tics and seizure activity), intactness of cranial nerves (e.g., movement of face, mouth, tongue, and eyes, including feeding, swallowing, and gaze [note excessive drooling]).

A. Gross motor coordination. Infants: Pushing up, head control, rolling, sitting, standing. Toddlers: Walking, running, jumping, climbing, hopping, kicking, throwing and catching a ball. (It is useful to have something for the child to climb on, such as a chair.)
B. Fine motor coordination. Infants: Grasping and releasing, transferring from hand to hand, using pincer grasp, banging, throwing. Toddlers: Using pincer grasp, stacking, scribbling, cutting. Fine motor and visual-motor coordination can be screened by observing how the child handles puzzles, shape boxes, a ball and hammer toy, small cars, and toys with connecting parts.

V. Speech and language

A. Vocalization and speech production: Quality, rate, rhythm, intonation, articulation, volume.
B. Receptive language: Comprehension of others' speech as seen in verbal or behavioral response (e.g., follows commands), points in response to "where is" questions, understands prepositions and pronouns (include estimate of hearing, especially in child with language delay, e.g., response to loud sounds and voice; ability to localize sound).
C. Expressive language: Level of complexity (e.g., vocalization, jargon, number of single words, short phrases, and full sentences); overgeneralization (e.g., uses *kitty* to refer to all animals); pronoun use, including reversal; echolalia, immediate or delayed; unusual or bizarre verbalizations. Preverbal children: Communicative intent (e.g., vocalizations, babbling, imitation, and gestures, such as head shaking and pointing), caregiver's ability to understand infant's communication, child's effectiveness in communication.

VI. Thought

The usual categories for thought disorder almost never apply to young children. Primary process thinking, as evidenced in verbalizations or play, is expected in this age group. The line between fantasy and reality is often blurred. Bizarre ideation, perseveration, apparent loose associations, and the persistence of pronoun reversals, jargon, and echolalia in an older toddler or preschooler may be noted in a variety of psychiatric disorders, including pervasive developmental disorders.

A. Specific fears: Feared object, worry about being lost or separated from parent.
B. Dreams and nightmares: Content is sometimes obtainable in children 2–3 yrs of age. Child does not always perceive it as a dream; for example, "A monster came in the front door."
C. Dissociative state: Sudden episodes of withdrawal and inattention, eyes glazed, "tuned out," failure to track ongoing social interaction. Dissociative state may be difficult to differentiate from an absence seizure, depression, autism, or deafness. The context may be helpful (e.g., a child with a history of neglect freezes in a dissociative state as mother leaves the room). Neurological or audiological evaluation may be warranted.
D. Hallucinations: Extremely rare, except in the context of a toxic or medical disorder, then usually visual or tactile.

VII. Affect and mood

The assessment of mood and affect may be more difficult in young children because of limited language, lack of vocabulary for emotions, and use of withdrawal in response to a variety of emotions, from shyness and boredom to anxiety and depression.

A. Modes of expression: Facial, verbal, body tone and positioning.
B. Range of expressed emotions: Affect, especially in parent–child relationship.
C. Responsiveness: To situation, content of discussion, play, and interpersonal engagement.
D. Duration of emotional state: Need history or multiple observations.
E. Intensity of expressed emotions: Affect, especially in parent–child relationship.

VIII. Play

Play is a primary mode of information gathering for all sections of the Infant and Toddler Mental Status Examination. In young children, play is especially useful in the evaluation of the child's cognitive and symbolic functioning, relatedness, and expression of affect. Themes of play are helpful in assessing older toddlers. The management and expression of aggression are assessed in play as in other areas of behavior. Play may be with toys or with the child's own or another's body (e.g., peek-a-boo, roughhousing), verbal (e.g., sound imitation games between mother and infant), or interactional or solitary. It is important to note how the child's play varies with different familiar caregivers and with parents versus the examiner.

A. Structure of play (ages are approximate).
   1. Sensorimotor play
      a. 0–12 months of age: Mouthing, banging, dropping and throwing toys or other objects.
      b. 6–12 months of age: Exploring characteristics of objects (e.g., moving parts, poking, and pulling).
   2. Functional play
      a. 12–18 months of age: Child's use of objects shows understanding and exploration of their use or function (e.g., pushes car, touches comb to hair, and puts telephone to ear).
   3. Early symbolic play
      a. 18 months of age and older: Child pretends with increasing complexity; pretends with own body to eat or to sleep; child pretends with objects or other people (e.g., "feeds" mother); child uses one object to represent another (e.g., a block becomes a car); child pretends a sequence of activities (e.g., cooking and eating).
   4. Complex symbolic play
      a. 30 months of age and older: Child plans and acts out dramatic play sequences, uses

imaginary objects. Later, child incorporates others into play with assigned roles.

5. Imitation, turn taking, and problem solving as part of play.

B. Content of play. The toddler's choice and use of toys often reflect emotional themes. It is desirable to have on hand toys that tap different developmental and emotional domains. An overfull playroom may be overwhelming or overstimulating and may reduce meaningful observations. Young toddlers of both sexes often gravitate to dolls, dishes, animals, and moving toys, such as cars. The examiner's choice of specific materials may facilitate the expression of pertinent emotional themes. For example, a child traumatized by a dog bite may more likely re-enact the trauma if a dog and doll figures are available. The child's reaction to scary toys, such as sharks, dinosaurs, or guns, should be noted, especially if they are avoided or dominate the session. Does aggressive pretend play become real and physically hurtful? By 2.5–3 yrs of age, a child's animal or doll play can reveal important themes about family life, including reactions to separation, parent–child and sibling relationships, experiences at day care, quality of nurturance and discipline, and physical or sexual abuse. The examiner must use caution in interpreting play, viewing it as a possible combination of re-enactment, fears, and fantasy.

IX. Cognition

Using information from all previously mentioned areas, especially play, verbal and symbolic functioning, and problem solving, roughly assess the child's cognitive level in terms of developmental intactness, delays, or precocity.

X. Relatedness

A. To parents: How in tune do the child and parent seem? Does the child make and maintain eye, verbal, or physical contact? Is there active avoidance by child? Note the infant's level of comfort and relaxation being held, fed, molding into caregiver's body. Does the toddler move away from the caregiver and check back or bring toys to show, to put into his or her lap, to play with together or near the caregiver? Comment on physical or verbal affection, hostility, reaction to separation and reunion, and use of transitional objects (blanket, toy, caregiver's possession). Describe differences in relating if more than one caregiver is present.

B. To examiner: Young children normally show some hesitancy to engage with a stranger, especially after 6–8 months of age. Appropriate wariness in young children may result in a period of watching the examiner, staying physically close to a familiar caregiver before engaging, or showing some constriction of affect, vocalization, or play. After initial wariness, does the child relate? Does the child engage too soon or not at all? How does relatedness with a stranger compare to that with a parent? Is the child friendly versus indiscriminately attention seeking or guarded versus overanxious? Can the examiner engage the child in play or structured activities to a degree not seen with caregiver? Does the child show pleasure in successes if the examiner shows approval?

C. Attachment behaviors: Observe for showing affection, seeking comfort, asking for and accepting help, cooperating, exploring, controlling behavior, and responding to reunion. Describe age-related disturbances in these normative behaviors. Disturbances often are seen in abused and neglected children; for example, fearfulness, clinginess, overcompliance, hypervigilance, impulsive overactivity, and defiance; restricted or hyperactive and distractible exploratory behavior; and restricted or indiscriminate affection and comfort seeking.

From Thomas JM, Benham AL, Gean M, Luby J, Minde K, Turner S. Practice parameter for the psychiatric assessment of infants and toddlers (0–36 months). *J Am Acad Child Adolesc Psychiatry*. 1997;36(10 Suppl):21S–36S.

The advent of many of these tools can be traced to two related developments in psychiatric clinical thinking and research since 1980. First, the introduction of explicit, criterion-based diagnostic systems such

as the Research Diagnostic Criteria and the DSM-III followed by the DSM-III-R, the DSM-IV, most recently the DSM-5, as well as the research version of the ICD-10 facilitated the development of standardized interviews and questionnaires, which specify the relevant elements of a diagnosis. Second, in recent decades there has been a rapid expansion in psychiatric research, creating a need for tools for research assessment. Thus, working synergistically, major changes in the diagnostic system, the impetus provided by emerging research findings, and the increasingly apparent advantages of using standardized diagnostic tools in research (and, more recently, in clinical settings) are key developments that have presented the field with an array of psychiatric assessment techniques and thus with a corresponding growing body of research literature on their characteristics and utility.

Because the inventory of symptomatic and diagnostic tools for children and adolescents has rapidly expanded over recent years, it is important for those evaluating and treating youth to be familiar with the literature and the clinical and research relevance of these tools.

Table 36.1–7.
Differential Psychosocial Diagnosis of the Uncooperative Patient

Inadequate understanding of the nature of illness or injury, treatment details, or staff expectations and intentions
Anxiety and regression
Separation anxiety
Concerns over bodily integrity
Response to immobilization
Denial
Idiosyncratic sources of anxiety
Inadequate pain control
Depression
Exacerbation of pre-existing psychopathology, especially oppositional or conduct difficulties
Family issues
Parental anxiety or hostility toward staff
Parental psychopathology: Depression, psychosis, or factitious disorder by proxy
Prior family experiences with illness
Cultural issues
Developmental issues (especially adolescence)
Emphasis on autonomy, ambivalence regarding dependency
Emphasis on activity, peer involvement
Self-consciousness about body
The body as a means of expressing psychological issues
Mature decision to refuse treatment

From King RA, Lewis M. Differential psychosocial diagnosis of the uncooperative patient. *Child Adolesc Psychiatr Clin North Am.* 1994;3:531.

Issues in Developing Diagnostic Methods

The task of developing diagnostic tools in psychiatry is a complex and challenging one, given the complicated nature of the symptoms and their clusters and patterns over time. In addition, developing measures to assess psychopathology in children and adolescents is more challenging than the comparable task with adults. These challenges include the need to incorporate an understanding of aspects of development into the framing and wording of questions about symptoms and into the process of discriminating normative from pathological behaviors. Many behaviors only become "symptoms" when they occur at too high a frequency or at the wrong age (e.g., separation problems and oppositional behaviors). In addition, many symptoms reflect an underlying substrate of psychopathology with phenotypic manifestations that differ during different developmental stages. Moffit referred to these variations as a function of the "heterotypic continuity" of disorders. Antisocial behavior provides an illustrative example: what is difficult temperament in early months and years may progress to the characteristics of oppositional defiant disorder in middle childhood, then to conduct disorder in later childhood/adolescence, and eventually become manifest as adult antisocial personality disorder.

Another challenge in the development of structured interviews or scales is the necessity of obtaining information from both the child and another informant, usually the parent and at times the teacher, because developmental stage can limit the capacity of children to report on their own behaviors/symptoms. Investigators have repeatedly found poor agreement between different informants with respect to symptom presence and hence diagnosis; yet, as has long been realized by clinicians, there is general agreement that information obtained from any informant is diagnostically useful. There is no consensus, however, on how best to combine the information across informants or as to whether to combine the information at all. Some kinds of information may be uniquely accessible to one or another informant—for example, teachers may be better reporters of attention problems than parents, whereas parents are likely to be better reporters of eating behaviors and irritable mood. For alcohol or other substance use, youth consistently report more symptoms than parents do, and similar results have been found for sexual behavior and conduct disorder.

Characteristics of informants may influence how they report on symptoms and behavior. Parents who have anxiety disorders tend to report greater levels of psychopathology in their children than do independent observers. Mothers who are depressed tend to over-report behavior problems and symptoms in their children, even when contextual variables and child gender have been controlled for. Complicating this issue, children

of depressed mothers have been found to have more behavioral problems than children of nondepressed mothers. At younger ages, children often do not have the ability to compare themselves to others, which can affect their reports on oppositional behavior or hyperactivity.

Further challenges to the development of diagnostic measures include the problem of how to integrate concepts of social context and adaptive functioning into a symptom-based assessment. These factors are difficult to isolate conceptually and to operationalize without producing an overly long and complex measure.

Still another difficult aspect in developing structured diagnostic and symptom measures is the lack of a readily available gold standard with which to compare results from structured measures so as to modify and improve them. In the history of diagnostic methodological development, the psychometric properties of reliability and validity have been touted as the bases for determining what measure was best. It is important for users to understand these concepts and to be familiar with a proposed instrument's performance in these respects. Briefly, the reliability of an assessment method is often conceptualized as the extent to which the results of an assessment can be repeated or reproduced, whereas validity is conceptualized as the extent to which the instrument measures what it purports to measure.

There are various types of reliability. *Inter-rater reliability* involves the stability of an assessment when it is administered by different raters (in the case of psychiatric diagnostic interviews, by different interviewers). *Test–retest reliability* reflects the stability of results over time, that is, when an instrument is administered on two occasions, separated by a period of time in which the degree of real change in the condition being measured is expected to be trivial. To assess the test–retest reliability of child psychiatric diagnostic instruments, the measure is often administered twice, 1 to 2 weeks apart. The 1 to 2 week interval is chosen to minimize precise memory effects (i.e., the child or parent is not likely to remember, and thus repeat, exact responses to a great number of the questions previously asked). The relatively long duration of many child psychiatric conditions makes it likely that changes between the first and second administrations reflect unstable aspects of the measure rather than real changes in the disorder. For interviews, assessment of test–retest reliability is particularly complicated as a different interviewer administers the "retest" interview. This is necessary for two reasons. First, taking the same assessment from the same interviewer a second time may lead the respondent to leave out symptom reports, assuming that the interviewer already knows about the respondent's experience with an area of symptomatology or, conversely that the phenomena should be reported differently because the interviewer is going over the same issues a second time. Second, the interviewers themselves may recall the first interview and

this knowledge might lead them (consciously or unconsciously) to conduct the second interview differently than they would with a respondent they did not already know. Thus, test–retest reliability for diagnostic interviews will necessarily be somewhat limited by inter-rater reliability. In addition, discrepancies between the initial administration and the subsequent interview usually result from the informant reporting fewer symptoms at the second administration—the "attenuation effect." This has been found in all forms of repeated psychiatric inquiry and likely represents a systematic bias in reporting. *Internal consistency* is a third type of reliability and refers to the extent to which items designed to assess aspects of a known disorder (e.g., depression) are measuring the same underlying construct. In omnibus measures, the internal consistency of different diagnoses or domains may be assessed separately because these instruments by design assess a wide range of behaviors and symptoms rather than a single construct. Thus, the component items for each diagnosis would not be expected to correlate with the items from other diagnoses.

The validity of a measure is more difficult to assess than is its reliability. There are various ways to conceptualize and to assess validity. To assess what is termed *construct validity,* the psychiatric diagnoses or scores generated by a particular instrument are usually compared to those generated by an expert, or criterion, assessment. Statistical methods are then applied to quantify the degree of agreement about whether diagnoses are present or absent. In the methodological evaluation of the validity of an instrument, agreement is determined diagnosis by diagnosis, so that it can be revised in areas of poor agreement with the criterion. The essential challenge in determining the construct validity of a measure is to compare its performance to some standard that is known to be diagnostically correct. The problem in the field of psychiatric assessment is to determine that standard, because there is no laboratory test to provide objective evidence for the presence of a disorder. Pure clinical assessments of child psychiatric diagnoses are notorious for varying from one clinician to another; thus, the so-called "standard" is a moving target. Indeed, structured interviews and scales were developed in large part to improve the poor psychometric properties of nonstandard assessments. This situation has led to an approach that considers an instrument's performance within the nomological net of theory and findings about the condition being assessed. Stated briefly, this means that, to the degree that an instrument's findings converge with accumulated thinking and knowledge in the field, one may have relative confidence that, at least at this point, the instrument validly measures what it claims to assess. The major diagnostic assessments have shown considerable convergence in their evidence for reliability and validity; thus, in the complex process of instrument development, it is important to be aware when a new measure is falling below existing standards in these respects.

Despite—or perhaps because of—these various challenges, extensive research efforts have been directed toward refining psychiatric assessment methods in child and adolescent psychiatry. These endeavors have yielded a number of measures for use in research contexts; however, some of those described in the following discussion have also found appropriate adjunctive roles in clinical settings. The following discussion of methods first examines interview methods, which yield categorical diagnoses, and then reviews scales that cover symptom dimensions, often with respect to a specific form of psychopathology, for example, depressive symptoms.

Diagnostic Interviews

There are numerous "omnibus" diagnostic interviews for children and adolescents. Each of these is designed to be administered both to the parent (about the child/adolescent) and depending on his/her age, also to the young person. The exact content of the various diagnostic interviews may differ, but each covers a wide range of the more common psychiatric disorders of childhood and adolescence other than the autism spectrum disorders. Each version addresses the DSM diagnostic criteria (and often the ICD criteria as well), and updates for DSM-5 have either been prepared or are underway.

The omnibus diagnostic interviews differ from one another in several important respects, which should be taken into account when choosing among them. The first issue to consider is the purpose of the assessment (e.g., to identify a measure for research or to address a need for clinical treatment or triage? to establish diagnosis(es)? to identify co-morbidity? to measure severity?). Consideration of the purpose of the assessment also involves matters of the time frame of interest (current only? past year? ever? current and ever?). Practical issues also inform plans for implementation, including: how long it takes to administer the interview, whether the same interviewer must interview both parent and youth; qualifications for administrators/interviewers (is it necessary to have a clinician administer the assessment?); the standard training procedures and supervision; scoring procedures (are there specific rules or algorithms? is a computer program necessary; if so, what are the available reports? is scoring based on clinical judgment? is an expert editor/reviewer necessary? how transparent is the scoring? what data will the interview yield—diagnoses only or symptoms and scales as well?); infrastructure considerations (necessary staff, computer/analytic capabilities; office space to conduct private interviews); and, finally costs (including license fees, copying costs, storage). It is strongly recommended that before choosing an instrument, one obtain a copy of the interviews under consideration and critically review the contents to understand how symptoms and criteria are conceptualized by their authors.

Respondent-Based versus Interviewer-Based Interviews. Perhaps the most fundamental distinction between the different diagnostic interviews is the degree of structure they impose. In the past this was often characterized as "highly structured" versus "semi-structured," but more recently it has been conceptualized as the respondent-based versus the interviewer-based approach. The respondent-based interview (RBI) has been described by Shaffer and coworkers as a "precise script" to obtain clinical information from the informant. The RBI has precisely worded and ordered questions and equally structured response options; thus, the authors essentially transformed the diagnostic criteria into questions that should be easily understood by anyone. The RBI is totally dependent upon how clearly the questions are written and are understood by the respondent. A primary question that is answered positively leads to a branching sequence of nested questions that specify symptom frequency, intensity, duration, and degree of associated impairment. These questions help determine the significance of the reported behavioral-emotional state and aim to determine whether it reaches a certain clinical threshold. Alternatively, an initial negative response allows "skip out" of these specific questions. The relative advantages of the RBI include the potential for methodological development and also issues of economy and public health applicability. Because this type of interview is given in the same way each time, it is possible to study its performance and then modify it with a relative degree of confidence in the rationale for changes. Another advantage is the potential for relatively economical data gathering. Because the structure of this type of interview drives its administration, a range of administrators (e.g., including research assistants, college students and others with little psychiatric/medical background) can be trained in administration of an RBI, thus, substantially reducing research or clinical costs. The primary interviewer qualification is that one can read out loud (with expression). Furthermore, the use of a computerized version from which the administrator reads virtually eliminates administrator error, and diagnostic results can be made available quickly and efficiently because the computer assembles the information according to preprogrammed diagnostic algorithms. Even more cost-effective is self-administration using visual or auditory prompts or both, which bypass the need for and the expense of interviewers. The potential of RBI's as screening tools in community settings is another attractive aspect of this method.

Interviewer-based interviews (IBIs), however, use the skills of the clinically trained interviewer to make judgments about the presence and clinical significance of reported phenomena. Thus the IBI is totally dependent upon the interviewer's clinical skill and quality of training. IBIs contain some degree of structure, in that definitions are provided about what constitutes a symptom, but although questions are provided to prompt the clinician-administrator, it is expected that the interviewer will

pursue inquiry until clinical judgment is satisfied about whether a symptom is present. This allows for a degree of flexibility and responsiveness to the body of clinical data that are ordinarily used to determine the significance of reported emotions, moods, and behavior—for example, observation of affective state in response to a query or other observational data, such as eye contact, posture, and voice inflection.

Because the choice of interview type and then of the specific interview depend on the intended use, before selecting a measure, it is worthwhile to examine some of the extensive literature developed on these methodologies. Briefly, RBIs probably do not do well in diagnosing extreme and atypical phenomena (e.g., psychosis), because this method "assumes" that respondents (youth or parents) interpret the question as intended by the interview developer and also that the respondent has the capacity to reflect and to self-observe about the presence or absence of the symptom. IBIs are more flexible and responsive to the state of the respondent and hence could take less time to administer. For example, an astute clinician can usually assess manic symptomatology fairly quickly and efficiently in a previously high functioning youth. In practice, however, these interviews typically take as long as an RBI. In addition, with the IBI the careful clinician should catch most instances of the respondent misunderstanding the intent of the question, whereas the RBIs have little capacity to self-correct for this inevitable source of error. IBIs, although more costly than RBIs to administer, are particularly appropriate for clinically based research in which subtle diagnostic distinctions may be critical for defining relatively diagnostically "pure" samples and in which sample sizes are small enough to justify the expense of hiring interviewers trained in making these judgments.

RBIs, however, as noted previously, have many flexible and innovative applications in research and public health contexts. They have also found recent application in an increasing number of clinical contexts, for example, as efficient aids in collecting comprehensive symptom and diagnostic information at clinic or even inpatient intake. Because they systematically review symptoms across all domains, this type of screen sometimes uncovers significant symptomatology, such as suicidal tendency, that was not queried or not revealed on initial clinical interview. But, while the RBIs are, to some extent, mechanical and lack the personal element of the skilled clinical interview, it has been shown that many adolescents, in particular, appreciate the privacy of the computer-based methods (a familiar modality for many youth).

RBIs are particularly useful for epidemiological studies in which large sample sizes (i.e., in the thousands) require many interviewers, making the use of trained clinicians prohibitively expensive. These studies investigate issues of the prevalence of disorders, developmental patterns of psychopathology, psychosocial correlates, and risk and protective factors,

topics that often require large samples to provide enough statistical power to examine the study hypotheses. The ultimate choice of a measure requires careful weighing of the goals that it is intended to serve, the resources available, and the pragmatics of the situation.

RBIs and IBIs have been developed with the goal of reducing error in the collection of information relevant to making diagnoses. Error in the counting and combining of symptom data to determine a diagnosis can be minimized with RBIs or IBIs by the use of specified algorithms, usually computer applied, to establish whether sufficient symptoms are present for the diagnosis and whether they meet DSM criteria for intensity, frequency, and duration. Ultimately, the method that best serves the ends of error reduction, logistics, economy, (relative to the intended goal), and general appropriateness remains an informed judgment call. As opposed to the early 1980s (advent of DSM-III) when the development of structured and semi-structured instruments began, there now exists considerable theoretical and empirical information on which to base decisions about instrument choice.

Respondent-Based Interviews. The Diagnostic Interview Schedule for Children (DISC) is the most extensively developed and studied of the RBIs. Initial work on the DISC began in 1980, and since then, it has gone through numerous updates to maintain alignment with the DSM system in its iterations, as well as with methodological testing in both clinical and community settings. The DISC was specifically designed for use in large epidemiologic and community studies where it could be administered by "lay" (clinically untrained) interviewers who would read questions verbatim to parent (concerning children ages 6 to 18) and to child (ages 9 to 18) with most answers confined to "yes" and "no." Since it appeared, the DISC has also been used in clinical research studies, screening projects, clinical treatment settings and in juvenile justice settings. The current version, DISC-IV, is compatible with DSM-IV and ICD-10 (also the revised third edition of the DSM [DSM-III-R]) and covers three time frames—past 4 weeks, past year, and lifetime, although the accuracy of recall for the lifetime frame is questionable. An updated DISC-5, addressing DSM-5 as well as DSM-IV criteria, is in development. DISC diagnostic modules are self-contained, so that a user might select to administer only those modules of interest in a particular context. In the DISC, a uniform question structure has been applied in that for each criterion being assessed, an initial stem question is asked, which if endorsed, is then followed by contingent questions that elicit increasingly more specific types of information about characteristics of the potential symptom. This structure builds a great deal of efficiency into the interview. Given the complexity of information required for DSM diagnoses, administration of the whole interview using paper protocols is no longer practical, so the DISC is given

with the aid of a computer. Computer-assisted versions ensure fidelity to the intended structure of the questioning, as well as reducing lost time due to the interviewer's need to consider what the next question should be. There is a computerized version, the Voice-DISC, which can be self-administered. In addition to symptom questions, the DISC includes questions in each diagnostic section that inquire about associated impairment and perceived severity of impairment. These questions are asked if a respondent meets a certain base level of symptomatology for a given diagnosis—about half of what would be required for a DSM diagnosis.

An extensive literature is now available on the DISC and its associated technologies. Diagnoses are derived by application of a set of extensively developed computer algorithms that score the interview for informants separately and by combining information across informants. The DISC has few symptom skips so that it is possible to generate symptom and DSM criterion scale scores, further increasing its flexibility and general usefulness, especially in research settings. Of importance, respondent acceptability has been demonstrated to be quite good, despite some (but not an overwhelming) perception that it is a bit lengthy. A modified version, the Young Child DISC is administered to parents only and is used to assess childhood symptomatology down to age 4. The Young Adult DISC, with parallel self and "other informant" versions, is also available. A Spanish-language version has been developed and extensively tested. In addition, the DISC-IV has been translated into a number of other languages and has found various innovative applications around the world. Particularly exciting are those in which it can assist in clinical contexts where trained mental health clinicians are in short supply (e.g., its use in juvenile justice facilities, orphanages, and regions of the world with few mental health practitioners per capita). Of course, no RBI can substitute for thoughtful clinical practice, and the developers of this method have been clear that its use is adjunctive and facilitative, and not diagnostically definitive, when applied in clinical settings.

The Children's Interview for Psychiatric Syndromes (ChIPS) is a highly structured interview that was designed for use with lay interviewers to screen for current psychiatric disorders. The authors recommend that interviewers have at least a bachelor's degree in a field associated with mental health (e.g., psychology, social work, and psychiatric nursing). The first version of the ChIPS was specifically designed for interviewing children ages 6 to 12 years (unlike the DISC, which is not given to children under age 9); for later versions the age range was extended to include adolescents. Development of the parallel parent version (P-ChIPS) followed the child version. For each diagnosis, "cardinal" symptoms—those that are most frequently endorsed by youth who may have a particular disorder—are assessed first and, if not endorsed, the interviewer skips to the next diagnosis, making the interview briefer to administer than the DISC. The

interview is scored by the interviewer using a scoring booklet that provides rules for scoring DSM diagnoses.

The Diagnostic Interview for Children and Adolescents (DICA) and the Missouri Assessment for Genetics Interview for Children (MAGIC) both assess "lifetime" diagnoses and for some disorders, current diagnoses. The DICA had IBI and RBI versions (the latter being computer based) but only the RBI version is still in use. The MAGIC is a later version of the DICA, developed by the same team; although allowing for some interviewer discretion on probing, it was designed for administration by trained lay interviewers, with many specific probes and clearly specified procedures for clarifying responses and eliciting and coding symptoms, resembling an RBI more than IBI. Both the DICA and the MAGIC have parallel youth and parent versions and there are two youth versions, one for administration to children in the age range from 6 to 12 years and one for adolescents ages 13 to 17 years. The Parent and Youth forms are scored separately by well specified computer algorithms.

The Development and Well Being Assessment (DAWBA), which is widely used throughout Europe, although less so in the United States, assesses current diagnoses (the exact time frame varies by diagnosis). The DAWBA was designed to provide a shorter alternative to typical diagnostic interviews by including many "skips" and screening questions. It also makes use of a companion screening questionnaire, the Strengths and Difficulties Questionnaire (SDQ). The DAWBA employs a two-part procedure beginning with an RBI comprising both closed-ended, precisely specified questions (answered "yes" or "no" or using a choice of 2 to 3 relevant options) as well as open-ended questions (to which the answers are typed verbatim into an open text field by the interviewer or, in a self-administered version, directly by the respondent). The open-ended questions obtain supplementary information about the symptoms and accompanying functional impairment. Thus, like the DISC, and the MAGIC, the DAWBA does not require clinical training to administer. There are parent and youth versions and a teacher questionnaire. When the interviews and questionnaires are completed, all information is forwarded to a clinical rater who has been trained by the authors of the DAWBA. This clinician reviews and assigns diagnoses based on the transcripts of the interviewer notes and answers to the structured questions from all available informants (usually parent and child, and sometimes the teacher). Preliminary scoring yields "likelihood" estimates, and provisional diagnosis is done by computer. The scoring rules used for the DAWBA are not made available for review. The DAWBA and the SDQ have been translated into many languages.

Interviewer-Based Interviews. Various IBIs, developed to meet different research needs, have found use in clinical settings. Some

frequently used interviews include the Schedule for Affective Disorders and Schizophrenia for Children (K-SADS); the Child and Adolescent Psychiatric Assessment (CAPA), the Anxiety Disorders Interview Schedule for Children (ADIS-C); and the MINI International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). In addition, a youth version of the Structured Clinical Interview for DSM-5 (SCID), a widely used adult measure, is being planned. For each of the existing measures, a literature is available that defies summary in a single section. Following are some brief comments about the available IBIs.

The K-SADS, originally developed for use in studies of child/adolescent depression was based on the SADS, an adult research instrument, and was the first widely used IBI for children and adolescents. Unlike the SADS, the K-SADS has kept pace with changes to the DSM. There are a number of different versions of the K-SADS, reflecting the modification of this measure by researchers at different sites with differing specific interests and assessment philosophies. Most assess current and lifetime diagnoses. Some versions incorporate severity scales for the criteria assessed, allowing their use in the examination of change, whereas others use more categorical scoring. All versions of the K-SADS contain probes that are "suggested," but there are no questions that must be asked, because all versions of the K-SADS assume that the interviewer has been clinically trained and is experienced in assessing youth psychopathology. Nonetheless, many research studies use trained research assistants to administer the interview. Each version begins with an unstructured introductory module during which information about the child's presenting problem(s) and a short clinical history are obtained along with demographic information and information about current functioning. This is used to set the context for further inquiry. The interviewer then proceeds through a series of modules that assess different diagnoses. For most of the K-SADS versions, the same person interviews both the parent and the child and, when necessary, works to resolve differences; an exception is the Washington University K-SADS, which does not employ consensus ratings and uses different interviewers for each informant. For the other versions, the authors recommend that for younger children the parent be interviewed first, whereas for adolescents, the recommendation is to interview the youth first; in practice, however, many have found that it makes more clinical sense to interview the parent first, regardless of the young person's age. Consensus ratings for each diagnostic criterion are made, based on the interviewer's clinical judgment using information from both informants and from any other source (e.g., clinical records). Diagnoses are then assigned by the interviewer using a series of score sheets. A computer assisted version of the K-SADS is in preparation. An early childhood version (K-SADS-EC) (parent informant only) has been prepared and used in studies of preschool depression. (A web-based,

computer administered version of the K-SADS is currently undergoing testing. This interview will use a different format using completely specified questions and response options, thus becoming an RBI, rather than an IBI.)

The CAPA is a comprehensive measure that, in addition to psychiatric symptomatology and impairment, covers family, environmental, and psychosocial problems, life events, the effect of the child's problems on the family, and mental health service use. The time period assessed is last three months. Unlike the other interviews, the CAPA is not organized into obvious diagnostic modules. The CAPA contains a number of questions that should be asked exactly as written, with follow-up questions that are less specified. Unlike the K-SADS, the CAPA was designed for administration by nonclinicians who undergo extensive training in how to administer the interview. This is accomplished by providing detailed definitions and rules that govern the assessment of reported phenomena and determination of whether they should be rated as symptoms. Thus the onus is on the interviewer to assure that the questions are understood and the information provided is relevant. A clinically experienced supervisor reviews each CAPA interview before final ratings are made. Young Adult (YAPA) and Preschool Versions (PAPA) of the instrument are available. Like the DISC and MAGIC, the CAPA is scored according to well-specified computer algorithms.

The ADIS-C, derived from the A-DIS, an adult interview, was originally developed for use in studies of childhood anxiety. The time frame for inquiry is current and lifetime; the interview provides extensive coverage of anxiety disorders, and is particularly strong in this area, with less extensive coverage of other symptom areas. "Thermometers" are used to obtain severity, fear/avoidance, distress, and interference information and a clinician severity rating is made for each diagnosis (which makes possible tracking symptom change). The ADIS-C is designed to be administered by a clinician with experience in assessing child psychopathology. Like the K-SADS, the ADIS-C relies on a reconciliation procedure to sort out disagreements between parent and youth and the scoring is done by the clinician administering the interview, rather than by computer algorithm.

The MINI-KID a youth version of the MINI, a widely used research interview for adults, was designed to quickly assess the presence of current psychiatric disorders of youth. Like the MINI, the MINI-KID incorporates a branching structure with each diagnostic section containing screening questions, which if not endorsed, allows the interviewer to skip the rest of the section. Unlike the administration procedure for all of the other diagnostic interviews for youth, parents and children are interviewed together, with discrepancies between informants addressed by the clinician interviewer during the interview. For adolescents, sometimes only the adolescent is interviewed, depending on setting and user choice. Of all the

IBIs, the MINI-KID takes the least time to administer (the authors estimate half an hour), due to the many skips and the efficiency of combining parent/youth interviews. Diagnostic criteria are documented within each of the disorder sections and on a summary sheet by the clinician administering the interview.

Rating Scales

A large repertoire of questionnaires and rating scales have been developed to assess symptomatology and various aspects of behavior and adaptive functioning. Different measures have been designed for completion by parents, teachers, and children. These focus on a broad range of symptoms ("omnibus" measures) or on a more specific area (e.g., depression and anxiety). The development of each scale involves testing its psychometric properties, that is, the functioning of the scale with respect to the types of reliability, validity and matters of administering it to large and diverse samples to determine norms and clinically relevant cutoffs.

A number of specific aspects of scale functioning must be considered in selecting a scale in practice or research. Again, a basic issue is the nature of the intended use, that is, to address a research question or for aid in clinical assessment and planning. Intended use determines which domains of symptomatology and functioning must be covered and is tied to the nature of the sample or population with which it will be used. For example, the major scale requirements may vary for studies that assess clinical patients receiving an intervention designed to change symptom occurrence over time versus community populations in which the prevalence of a disorder(s) would be relatively stable or else likely to change in response to an imposed condition (e.g., intervention) or the changing developmental stage of a cohort tracked over time. In considering which scale to use for a particular purpose, the psychometric characteristics of various scales should be compared and factored into a decision, together with specific characteristics of the children or adolescents to be assessed. One obvious example is age. Given that developmental stage and not solely chronological age determines the young person's capacity to abstract at the level required by the particular scale's questions, the appropriateness of such questions to the youngster's activities and opportunities are all factors very important to accurate assessment. The young person's cultural context also may bear on symptom reporting for various reasons, such as acceptability of reporting certain types of symptoms, culture-specific theories of etiology, or expectations of what is normative in behavioral or emotional expression.

Thus, although a particular scale may appear to provide a convenient tool for clinical or research use, many features of its demonstrated functioning, of the young people themselves, and of their social contexts must be incorporated into the choice of a measure. Similarly, the critical

reader will want to consider such factors in interpreting findings from the research literature. Finally, there is the issue of cost, as many of these measures must be purchased, which can be prohibitively expensive in low resource settings.

Comprehensive (Omnibus) Rating Scales. Probably the most widely used comprehensive symptom rating scales are those that constitute the Achenbach System of Empirically Based Assessment (ASEBA) family of instruments (http://www.aseba.org). There are forms for school-age children—the Child Behavior Checklist for Ages 6 to 18 (CBCL/6 to 18), the Teacher Report Form (TRF/6 to 18), and the Youth Self Report (YSR/11 to 18); forms for younger children—the Child Behavior Checklist for Ages 1½-5 (CBCL/1½-5) and the Caretaker-Teacher Report Form (C-TRF); as well as forms for young and older adults to report on themselves. The ASEBA forms average 99 to 120 items and generate several syndrome scales. In addition, there are clinician-completed forms, including the Direct Observation Form (DOF), which is completed after observing classroom behavior. Important features of this system of measures include its empirical development and the large population sample on which it was normed, as well as testing on large samples of clinically referred children. The availability of age-based and gender-based norms provides a convenient way to screen for overall levels of symptomatology, and when scored, the CBCL yields a profile of problem areas and competencies. The CBCL is stable over relatively brief intervals (i.e., good test–retest reliability) but also reflects changes resulting from treatment. The CBCL, TRF, and YSR have been widely used in epidemiological and clinical research, in many clinical settings and in translated forms in numerous countries. These measures can be a useful clinical screening tool, as well as an efficient way for clinics to gather uniform information for characterizing the populations served as well as changes in those groups over time. Computerized data entry and scoring have greatly eased the use of the ASEBA measures and allow for comparisons of reports across informants.

The Behavior Assessment System for Children (BASC-3) is another widely used assessment system for evaluating aspects of behavior and personality in children and adolescents aged 2 to 21 years. It consists of parent and teacher rating forms, a Structured Developmental History form and a Student Observation System (for rating classroom behavior), and, for youngsters 8 to 21, a Self-Report of Personality (for thoughts and feelings). The parent and teacher measures have forms appropriate for ages 2 to 5 years, 6 to 11 years, and 12 to 21 years (the latter also has a youth form); each form has between 139 and 175 questions. A distinctive feature of the BASC is its emphasis on positive aspects of the child's adjustment, in addition to rating symptoms and problem behaviors. An additional feature of this system is its coverage of problems related to school functioning,

such as difficulties in social or study skills, and it has built-in validity checks. Norms have been established for children by age, gender, and clinical status. This system has been designed for application in clinical, educational, research, and program evaluation contexts. Use and interpretation of the BASC are eased by the availability of computer software that generates profiles and allows consideration of different informants' viewpoints.

The Conners Rating Scales-Revised is another family of instruments that are widely used in research and clinical settings. These include short and long forms of scales for parent administration CPRS-RS, CPRS-R-L, teacher administration (CTRS-R-S, CTRS-R-L), and self-report (CASS-S, CASS-L). These scales collect information on child oppositional behavior, cognitive difficulties, hyperactivity–impulsivity, anxiety and shyness, perfectionism, social problems, and psychosomatic complaints. The Conners scales have been developed and tested on samples of children and youth 3 to 17 years of age, for parent and teacher forms, and 12 to 17 years on the adolescent forms. These measures have shown excellent psychometric properties, and norms are available, as is computer software to aid in administration and scoring. The Conners scales have long been the standard instruments for the assessment of ADHD in children and adolescents, and their use for ADHD and Disruptive behaviors is described below.

In contrast to the omnibus scales described above, the SDQ, which has been gaining in popularity is very brief and is made freely available. The SDQ was designed to be used as a short screen to identify children (ages 2 to 17 years) with behavioral problems. The SDQ comprises 5 subscales, from a total of 25 items; it has cross informant forms (parent, teacher, and self) and has been translated into over 75 languages.

Other omnibus scales, which are available but have found less use in recent years include the Revised Behavior Problem Checklist (RBPC), the Louisville Behavior Checklist (LBC), and the Personality Inventory for Children (PIC). The RBPC focuses on symptoms of conduct disorder, socialized aggression, attention problems and immaturity, anxiety and withdrawal, with auxiliary scales for psychotic behavior and motor tension-excess. It was designed for use by parents and teachers of children and youth ages 5 to 17 years. Although generated using factor-analytic studies of clinical samples, it has also been normed on nonreferred children and youth. The LBC and the PIC are both comprehensive rating scales for parent report that do not have corresponding versions for other informants. The LBC consists of 19 scales that cover a wide range of behaviors and symptomatology. The PIC was developed on the model of the Minnesota Multiphasic Personality Inventory (MMPI) and also comprises a variety of component scales.

Disorder-Related (Domain Specific) Scales. Each of the scales described above was designed to collect information across a range of symptoms and competencies. More narrowly focused symptom rating scales have been developed to permit a more detailed quantitative assessment of specific symptom realms. These latter scales may be useful in quantifying the presenting severity of certain symptoms to establish a baseline against which response to a therapeutic intervention can be compared. A number of these scales are self-report measures. These have the advantage of providing an opportunity for the child to report on symptoms that might be difficult to conceptualize or verbalize. Sometimes parents or children (especially adolescents) are initially more willing to report an area of difficulty in the seemingly more anonymous context of a questionnaire than in a face-to-face interview. However, self-reports also carry the risk of the child misunderstanding questions or feeling compelled to provide socially desirable responses rather than acknowledge behaviors and feelings that the child believes might meet with disapproval. Fundamentally, the focused symptom approach may also ignore important interactions across domains of psychopathology, because of the focus on one particular area of symptomatology. With these potential limitations in mind, child self-report measures can provide useful adjunctive information in clinical and research settings.

SCALES FOR ASSESSING INTERNALIZING SYMPTOMATOLOGY'DEPRESSION. The Beck Depression Inventory (BDI) is a widely used scale for depressive symptoms in cognitive, behavioral, affective, and somatic domains. It has found wide use with adolescents. Because it is an "adult" measure, there is no corresponding parent version. The BDI was originally developed to measure the severity of depressive symptoms in patients who had been diagnosed with depression; it is also used to screen patients to identify depression and to monitor treatment outcomes. The BDI has been widely studied and used in clinical and nonclinical populations, with generally good psychometric properties. As with other scales, a fine-grained examination of the literature shows specific areas in which it falls short of ideal; for example, different cut-points are needed, depending on the population (inpatients vs. outpatients; male vs. female). Unlike the earlier version of the BDI, which did not provide complete coverage of the DSM depression criteria, the newest version, BDI-II, covers all of the depression symptom criteria in the DSM system and extends the time frame of inquiry to 2 weeks, again to reflect the DSM duration requirements for Major Depressive Episode. Nevertheless, varying degrees of stability that are inherent to psychopathology in adolescence must also be kept in mind in the overall determination of the functioning of any scale. Generally speaking, positive results from various psychometric studies, ease of use, scoring and interpretation have made the BDI an attractive instrument in

both clinical and research contexts.

The Children's Depression Inventory (CDI) is a self-report measure for assessing symptoms of depression in children 7 to 17 years of age. It shares a conceptual basis with the BDI and comprises five factors: dysphoric mood, acting out, loss of personal and social interest, self-deprecation, and vegetative symptoms. As with many measures of child psychopathology, various psychometric properties vary by characteristics of the sample used, especially patient versus community status, gender, and age. As noted for the BDI, participant characteristics play important roles in interpreting psychometric results because, for example, the capacity to conceptualize and report on symptoms improves with age. Hence, scales reaching into the younger school-age years are likely to show more variability in assessments of reliability and in examinations of agreement between child and adult informants. Levels of agreement with other measures related to depressive psychopathology and its known correlates have provided support for the validity of the CDI. It has been widely used across a range of populations; thus, as with the BDI, the availability of considerable information relevant to its performance with different groups of children and youth makes it attractive to those considering it for use in various research and clinical populations. Parent and teacher informant forms are available. Many useful results have emerged from studies using the CDI, further adding to its attractiveness as a measure of depressive symptomatology. One criticism of the scale is that adolescents may find it too "babyish."

The Center for Epidemiologic Studies-Depression Scale (CES-D) was developed first as a rating scale for depressive symptoms in epidemiological studies of adults, but it has also been extensively widely used with adolescents, and a modification (CES-DC) was developed specifically for use with younger individuals. Both are freely available. There is no corresponding parent measure. Psychometrically, the CES-DC functions sufficiently well in some domains for adolescents but, overall, does not hold up well for children. Both instruments are brief and easy to use, and, particularly in adolescents, some aspects of reliability and validity meet performance standards. However, for both measures, problems with discriminant validity pose limitations, suggesting that they may be most appropriately considered as screening tools for general psychopathology in adolescent populations.

The Mood and Feelings Questionnaire (MFQ) is another freely available brief self-report measure designed to screen for and assess core depressive symptoms for children and adolescents. Originally designed for use in epidemiological studies, the MFQ has also found use in clinical settings and screening programs. There are long (32 item) and short (13 item) versions of the scale, and parallel parent and youth reports. Although these questionnaires have demonstrated good psychometric properties and discrimination for depression appears better than for other self-report

measures, they are not as widely used as the other scales.

The Quick Inventory of Depressive Symptomatology (QIDS) and QIDS-A (a version for adolescents that includes an extra item assessing irritability) are short forms of the Inventory of Depressive Symptoms (IDS) and are confined to the symptoms of Major Depressive Episode. In addition to a self-completed questionnaire, there is a corresponding clinician rated version, as well as a structured interview guide; all are freely available. The QIDS has strong psychometric properties and is useful for tracking treatment response.

Two related scales, the Reynolds Adolescent Depression Scale (RADS) (for ages 11 to 20) and the Reynolds Child Depression Scale (RCDS) (for ages 7 to 13), are widely used measures to assess a wide range of symptoms associated with depressive disorder. Symptoms are rated on a 4-point Likert-type response scale that yields scores on four subscales (dysphoric mood, anhedonia/negative affect, negative self-evaluation, and somatic complaints). Both the RADS and the RCDS function very well in terms of reliability (both internal consistency and test–retest). The validity of the RADS is supported by relationships with other measures of depressive symptomatology and related constructs. It has been used with diverse populations, providing a broad base of evidence for its functioning, as well as normative comparison data. The strong psychometrics, basis in the highly specified symptomatology of the DSM system, and the linkage of the two scales for use in longitudinal studies or across a broad age range in a single study are advantageous features of these measures. The RADS and RCDS are often used as screening measures to identify depressed youth (there is no corresponding parent measure), and there is some evidence that these measures may be sensitive to clinical change, so they hold promise for evaluating treatment outcomes.

The Hamilton Rating Scale for Depression (HRDS), another freely available measure, differs from the aforementioned measures in being a clinician-rated scale. It is the most commonly used clinician rating scale for depression. Developed for assessing depressive symptoms in adults, the HRDS has also been used in studies of adolescents. The HRDS was initially developed to measure the severity of symptoms in patients who had a depressive illness and was widely used in psychopharmacological and other treatment studies evaluating outcomes of different treatments for depression. The original scale had 21 items, of which only the first 17 were scored; there are alternative versions with 23 items and also versions with fewer items. Not all of the versions have been psychometrically evaluated. Although internal consistency and inter-rater reliabilities have generally yielded acceptable to excellent results, specific training in its administration and the use of a structured or semistructured protocol for administering the HDRS improves the psychometric properties. The validity of the HRSD has not been as convincingly demonstrated in

adolescent populations as for adults. The main application has been in treatment outcome studies, in which findings have supported its sensitivity to the effects of certain psychopharmacologic agents, as well as to other specific therapeutic modalities.

The Children's Depression Rating Scale-Revised (CDRS) is a clinician-rated symptom inventory related to the HRDS but designed for the assessment of school-aged children. It uses clinician expertise by including their ratings of behavior observed during the interview and by having clinicians generate scores based on information from each parent and child considered alone, as well as a combined score that integrates information from both informants. Psychometrically, aspects of reliability have ranged from adequate to excellent, depending on the study and type of reliability considered. There is some support for validity, and the CDRS has found substantial use in clinical research generally and treatment outcome studies specifically. Overall, the unique features of the CDRS, clinician rating, and use of multiple informants have made it attractive for studies conducted in settings in which the use of clinicians is feasible and their skills in eliciting and integrating information are advantageous for the study design.

SCALES FOR ASSESSING INTERNALIZING SYMPTOMATOLOGY'ANXIETY. The Hamilton Anxiety Rating Scale (HARS) focuses on anxiety symptoms and, like its depressive symptom counterpart, the HDRS, is a freely available clinician-rated scale, which has been used with adolescents. The HARS is generally administered to both adolescent and parent to yield an overall rating. It covers both psychological and somatic components of anxiety (focusing heavily on the latter), with psychological symptoms showing better internal consistency than somatic ones. Other aspects of reliability are moderate to excellent, depending on the type of reliability examined, while validity is not strongly supported. The use of clinicians for the assessment and the use of multiple informants are potentially advantageous features, although more study is needed to elucidate the psychometric performance of the HARS and its utility as a treatment outcome measure.

The Pediatric Anxiety Rating Scale (PARS) is another clinician rating scale that is freely available. The PARS, modeled on the Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS), was developed for a study on child anxiety to target symptoms that the researchers thought were not adequately covered by the HARS. In administering the PARS, the clinician first reviews the child's specific symptoms of anxiety and then makes severity ratings on seven dimensions using a six-point scale. The psychometrics of the PARS are good and the scale is useful for tracking treatment response. However, administration can be lengthy.

The Multidimensional Anxiety Scale for Children (MASC) is an empirically derived self-rated scale for children and adolescents that

comprises the four major factors of anxiety, namely physical symptoms, social anxiety, harm avoidance, and separation anxiety. With respect to scale construction, the MASC maps onto the diagnosis of generalized anxiety disorder, and individual factors capture the specific diagnoses of separation anxiety disorder and social phobia. With respect to psychometrics, various types of reliability and validity have been moderate to excellent. It has attractive structural features in capturing and distinguishing particular types of anxiety symptoms, and some results indicate that it distinguishes the constructs of anxiety and depressive symptoms in young people and also that it is sensitive to treatment effects. The MASC has been translated into over 15 languages. The most recent form, MASC-2, has a corresponding parent version. Less is known about the psychometrics of the MASC-2 and its parent version than about some other measures, and further findings are anticipated on its statistical performance for specific research purposes.

The Screen for Child Anxiety Related Emotional Disorders (SCARED) is a freely available self-rated scale developed for children and adolescents ages 9 through 19 years, with a parallel form to be completed by parents. It maps onto the criteria for DSM-IV anxiety disorders of five types: somatic/panic, generalized anxiety disorder, separation anxiety disorder, social phobia, and school phobia. The overall scale and subscales show reliabilities comparable to those of other well-functioning scales, and there is evidence for convergent and discriminant validity of the SCARED with respect to other assessment measures and types of psychopathology. The SCARED has been shown to be sensitive to treatment effects. There are long (41 item) and short (5 item) versions of the scale, and the instrument has been translated into many languages.

The Revised Children's Manifest Anxiety Scale (RCMAS), a downward extension of the adult Manifest Anxiety Scale, was developed to assess trait anxiety in school-aged children, although it has also been used with adolescents. Unlike many of the other scales, it does not have a parent informant version. The RCMAS has three component subscales: physiological, worry/oversensitivity, and fear/concentration. It has been in use longer than the MASC and the SCARED, and thus experience has yielded a number of studies that document the psychometric properties of the RCMAS. Concurrent validity has been supported through comparisons with other measures of anxiety, although examination of discriminant validity has found that ADHD symptomatology may be rated for those who are primarily anxious. The many studies using the RCMAS have yielded numerous contributions to the understanding of assessment issues in anxiety disorders among the young and the role of anxiety in children and youth from various ethnic/cultural groups and diverse stressful social contexts. Some construct and other methodological limitations have been noted; many of which may have their origin in the inception of the RCMAS

prior to the newer, DSM-linked measures.

The State-Trait Anxiety Inventory for Children (STAIC), also based on an adult instrument, is another youth-only self-completed measure. The STAIC comprises two subscales, each of 20 items; one quantifies "state" with the other quantifying "trait" anxiety. The subscales can be used independently. The STAIC has demonstrated acceptable psychometric properties and has somewhat better specificity than sensitivity. It has been shown to be sensitive to change and useful in assessing response to stress.

SCALES FOR ASSESSING EXTERNALIZING SYMPTOMS'ADHD AND RELATED SYMPTOMS. Several scales developed by C. Keith Conners and colleagues are commonly used to assess and monitor ADHD and other externalizing disorders. The current is the Conners 3 (third edition), which has parent, teacher, and self-rated versions, all with good psychometric properties. There are long and short versions for each form, The Conners-3 assesses symptoms of ADHD, Conduct Disorder and Oppositional Defiant Disorder, with each symptom being rated on a 4-point scale and also includes the Conners ADHD index, and a Global Index and sophisticated computer software that aids in administration and scoring. Other commonly used Conners instruments include the Conners Abbreviated Symptom Questionnaire, Hyperactivity Index (CASQ-HI), the Conners ADHD Index and Inattention/Overactivity with Aggression (IOWA) Conners Scale, and the Conners, Loney, and Milich (CLAM) Scale. The CASQ-HI is a brief (10 item) checklist of the 10 most "medication-sensitive" hyperactivity items from the longer Conners' scales. In addition to assessing treatment response, the CASQ-HI has often been used to screen for ADHD; however, it has been criticized as potentially missing children whose primary problem is inattention. The Conners ADHD Index addresses this possible weakness in identifying inattentive symptomatology in that the 12 items that comprise the index were derived to discriminate between children with ADHD and without ADHD. The IOWA Conners Scale has two subscales: Inattention/Overactivity (IO) and Oppositional Defiant (OD). For each subscale, items were selected to have the least overlap with the other scale. In practice, however, the scales correlated moderately in clinical samples. The CLAM Scale combines the CASQ with the IOWA Conners. All of these scales are brief, widely applied, have demonstrated good psychometric properties, and have been used to identify children in need of intervention and to track treatment response. In addition, Conners has developed scales for preschool children and for identification of adults who have ADHD.

A second "family" of scales focusing primarily on ADHD has been developed by Swanson and colleagues, all of which are made freely available. The most widely used of these is the short version of the Swanson, Nolan, and Pelham (SNAP-IV) Rating Scale, which comprises a checklist of the DSM-III, DSM-III-R, and DSM-IV ADHD and ODD criteria

rated on a 4-point scale and allows comparison between the different diagnostic formulations. The scale can be completed by either parent or teacher. The longer SNAP-IV contains items for other DSM diagnoses as well, and a more abbreviated version of the SNAP-IV contains only the ADHD items. The Strengths and Weaknesses of ADHD-symptoms and Normal-behavior (SWAN) Scale includes the same DSM ADHD symptom types as the SNAP-IV, but each is written so that it can be considered as an area of strength or weakness (e.g., the SNAP-IV "often is distracted by extraneous stimuli" appears as "ignores extraneous stimuli" on the SWAN). The parent or teacher is asked to consider each symptom as a dimension and rate the strengths or weaknesses of the child using a 7-point metric that compares how the child does in comparison to other children, ranging from "far below average" to "far above average."

Finally, the Swanson Kotkin, Agler M-Flynn, and Pelham (SKAMP) rating scale was created to monitor the classroom behavior of a child, rather than DSM symptoms; it contains 10 items about typical target behaviors addressed by psychosocial interventions. The SKAMP is completed by the child's teacher and is incorporated in the full teacher version of the SNAP-IV. The main disadvantage of the Swanson family of scales is the limited psychometric or normative data on their functioning.

In addition to the Conners and Swanson scales, there are a number of other brief symptom checklists keyed to the DSM criteria that can be used to assess ADHD and other related behaviors. These include the ADD-H Comprehensive Teacher Rating Scale (ACTeRS), the Attention Deficit Disorders Evaluation Scale (ADDES; home and school versions), the Yale Children's Inventory (YCI; a parent rating scale designed for children with learning disabilities), the Vanderbilt AD/HD Diagnostic Teacher Rating Scale, the Vanderbilt AD/HD Diagnostic Parent Rating Scale, the ADHD Rating Scale, and the ADHD Symptom Checklist.

OTHER SCALES FOR EXTERNALIZING SYMPTOMATOLOGY. The Eyberg Child Behavior Inventory (ECBI) is a parent-completed rating scale, with its companion, the Sutter-Eyberg Student Behavior Inventory (SESBI-R), a teacher completed measure, was designed to assess a wide range of disruptive behavior symptoms in children aged 2 to 16 years. These easy to use scales have found wide use in treatment studies and in clinical settings and must be purchased. For each of the behaviors assessed (36 on the ECBI and 38 on the SESBI-R), the informant rates the frequency of the behavior on a 7-point scale (never to always) and then rates whether the behavior is "currently problematic." Two scores are generated—the Intensity Score (based on the frequencies) and the Problem Score (calculated using ratings on how problematic each behavior is). Normative data for these scores are provided. The scale scores have good internal consistency for gender, ethnic, and age groups, and although they are not tied to DSM diagnoses,

the ECBI has shown good discriminant and concurrent validity, strong temporal stability, and evidence of construct validity.

The Antisocial Process Screening Device (APSD) was designed to identify psychopathy and can be completed by either parents or teachers about children ages 6 to 13 years. This 20-item scale is based on Hare's Psychopathy Checklist-Revised. Items on the APSD address callousness, narcissism, and conduct problems that are related to poor impulse control. The scale has acceptable internal consistency, there is some evidence for construct validity, and norms are available for the scale. Sensitivity and specificity have not yet been established, and there are no recommended "cut offs" for diagnosing antisocial psychopathology.

The New York Teacher Rating Scale for Disruptive and Antisocial Behavior (NYTRS), which is freely available, includes symptoms of Oppositional Defiant Disorder and Conduct Disorder and deviant peer interactions as they occur in the school setting. The NYTRS has four subscales—defiance, peer problems, physical aggression, and delinquent aggression—and two composite scores (antisocial behavior and disruptive behavior) and also a global impairment score can be derived. The limited psychometric evaluation of the scale has shown promising results.

The Children's Aggression Scale (CAS), based on the Overt Aggression Scale (OAS) and the Buss-Durkee Hostility Index (BDHI), was developed with a specific focus on aggression, rather than a DSM diagnosis, as exhibited by youth ages 5 to 18 years. There are parallel parent and teacher versions (no self-report version), and CAS has shown good psychometric properties, and has been normed.

In addition, several other scales have been developed to monitor aggression and hostility. These include: The OAS, which is used widely in inpatient settings; modifications of the BDHI; the Vitiello Aggression Questionnaire (VAQ), which differentiates proactive from reaction aggression; and the Direct and Indirect Aggression Scale (DIAS). Unfortunately, normative data for most of these scales are not available for youth, although initial psychometric analyses look promising.

SCALES ASSESSING ADJUSTMENT AND IMPAIRMENT. Instruments also exist for assessing levels of adaptive functioning and impairment, globally or with respect to specific domains. The Children's Global Assessment Scale (CGAS) is an easy-to-use quantitative scale for the clinician's rating of overall impairment. The CGAS has good psychometric properties and has been translated into over 20 languages. A parent-completed version of the CGAS is also available. The Columbia Impairment Scale (CIS) and the Brief Impairment Scale (BIS) are both self-completed rating scales of impairment that tap four major areas of functioning: interpersonal relations, broad psychopathological domains, functioning in school or at a job, and use of leisure time with parent and youth (ages 9 and older)

versions. The WHO Disability Assessment Schedule (WHO-DAS) was adapted for youth as part of the DSM-5 revision process. This scale includes parent rated and youth rated questionnaires as well as a clinician rated components. The CGAS, CIS, BIS, and WHO-DAS are all free of charge.

The Vineland Adaptive Behavior Scales 2nd Edition (Vineland-II) provide a comprehensive age-normed framework for systematically recording and assessing the development of children and adults (multiple versions for birth to age 90) with respect to various realms of adaptive functioning. For youth, there are self-report, parent-report, and teacher rated forms. The Vineland has been widely used in both clinical and research settings.

Various structured and semi-structured formats for performing or recording portions of the child mental status examination have also been developed. Some, such as the child version of the Mini-Mental State Examination, contain specific standardized items for screening cognitive functions. Other instruments, such as the Mental Health Assessment Form, provide a format for organizing mental status data derived from a semistructured clinical interview.

Limitations of Structured Measures in Clinical Assessment

Interviews and scales, such as those described previously, can, at times, be adjuncts in clinical assessment if thoughtfully chosen and administered with sensitivity to their meaning in the overall process of the evaluation. For example, they can help to ensure a comprehensive review of symptoms or provide a standardized baseline measure of severity for later determination of intervention efficacy. They cannot, however, replace an individualized child psychiatric interview or be relied on as the sole basis for establishing diagnoses or planning treatment. Most standardized interview schedules focus on symptoms and behaviors, with little attention paid to coping mechanisms, situational context, and adaptive strengths that the clinical interview affords. Such factors may be as crucial to the clinical assessment and treatment planning as the presence or absence of a given pathognomonic symptom or categorical diagnosis. This limitation is acknowledged by the authors and indeed, for many of these instruments, the user is required to seek permission from the authors and sign a license or complete a qualification form before being allowed to administer the instrument to ensure proper professional use. Some, like the DISC, require that the user agree that the interview not be used in place of thorough clinical evaluation in making treatment decisions.

The researcher interested in selecting measures for a study protocol must review carefully the existing measures in the area of interest and consider the psychometric characteristics of potential measures as well as the appropriateness of the measures for the study sample and the goals of

the study. It is useful to begin by examining up-to-date critical reviews of these measures to evaluate the interviews/scales being considered and actually test them out on volunteers, or if not available, with clinically savvy colleagues who can role play in developmentally and symptomatically authentic ways. If the researcher is not experienced with biostatistics, it is highly recommended that the basic psychometric papers be reviewed by a statistician colleague. Combining information from these different perspectives will maximize the chances that the time and effort put into data collection will yield information that is interesting and ultimately useful to the fields of child psychopathology, treatment, psychosocial intervention, and ultimately prevention

In clinical settings there are many challenges that must be addressed when introducing standardized measures. Perhaps the most formidable of these has to do with perceptions and attitudes of clinical staff who are unfamiliar with how standardized instruments might be usefully included as part of clinical practice. Some clinicians may be unconvinced that adding additional procedures could improve their usual practice in evaluating patients and making clinical decisions because they have great faith in their own clinical judgment. Some worry that their patients and families will feel "put off" by having to complete questionnaires or structured assessments (although research suggests that this is not usually a problem) and that including specific instruments may constrain their inquiry. Others may be overwhelmed by and resentful of any additional paper work. Clinics often include standard measures as part of an intake process but sometimes the information gathered is not given to the evaluating clinician. It is important not to include measures that aren't perceived as valuable clinically. A different sort of problem can lie in the over-reliance on or "reification" of a measure, especially with less experienced clinicians, who may assume a score on an inventory is more "valid" than their clinical observations, even though none of these measures should be considered as perfect. As an example, a depressed child who has become withdrawn, has insomnia and problems concentrating, and isn't sleeping might not endorse symptoms on a depression inventory, but that doesn't mean the child isn't depressed. The child's clinician should take account of the child's affect, rate of speech and reports of how the child is interacting with peers and family, etc. before concluding that depression is not present. However, children and adolescents and their parents will often more readily admit symptoms on a self-completed questionnaire, so responses on the questionnaire shouldn't be ignored. Often self-completion forms can be clinically useful by giving a clinician a starting place with a non-communicative patient. In such instances, it is useful to provide the questionnaire itself to the clinician rather than just a score, so that the clinician can review it with the patient or parent. When using any of the measures in a clinical setting, it is important to consider carefully the procedures for generating the most

useful information from the measure and sharing that information with the clinical team.

PSYCHOLOGICAL AND NEUROPSYCHOLOGICAL ASSESSMENT

Psychological assessment using tests of personality and cognitive abilities can be a valuable part the child psychiatric evaluation. Psychological evaluation provides a unique perspective on the child's condition by generating information complementary to that obtained from history, observation, and interview. Combined with history and observation, such psychological data may clarify relationships among various etiological factors or may help direct the selection of appropriate intervention. Although partly inferential, the clinical information derived from projective tests of personality is less influenced by certain reporting biases (e.g., social desirability) than self-report. Psychological testing also provides data that are more quantitative than the qualitative information provided by the child's history, presenting complaints, and clinical presentation.

The quantitative character of test findings is more important in the assessment of cognitive abilities than in personality assessment, because the child's relative strengths and weaknesses in mental processing can be estimated only grossly from history and clinical observation. Such clinical estimates are also necessarily less accurate. The ability to discriminate meaningful differences among various domains of functioning helps clarify the particular pattern or syndrome evidenced by the child. The assessment of cognitive functioning also reveals cognitive impairments that may underlie or influence the psychological condition under investigation (e.g., whether poor social judgment in a child with conduct disorder also results from a nonverbal learning disability). Furthermore, the manner in which any cognitive impairments or strengths interact with the child's psychological disorder can be delineated. Findings from tests of personality also identify important dynamics that affect the child's clinical condition.

Referral for Psychometric Assessment

A referral for a psychometric assessment of the child is a request for a professional consultation from which an opinion regarding the referral questions can be expected. The referral for psychological examination is analogous to that made to a neurologist or radiologist, who uses his or her assessment methods and particular expertise to contribute potentially valuable information that can then be considered, evaluated, and integrated into the psychiatric evaluation. Referrals are, therefore, best worded in terms of the information that is needed, the issues to be resolved, or the decisions that must be made, rather than requesting that specific tests be administered.

Referral questions can be broad in scope, such as requesting a second

opinion regarding the child's condition from the unique perspective of psychological test findings. Requests for assistance delineating differential diagnoses or identifying idiosyncratic attributes of the child that may affect treatment are narrower referral questions that, nonetheless, still require relatively broader psychological examinations. In contrast, more narrowly focused questions, such as the adequacy of the child's reality testing, may permit more focused examinations. However, the extent to which psychological evaluations can be restricted to a narrowly defined referral question has a lower limit. Usually, a critical mass of data is required to interpret test findings properly.

Referrals for psychological examination should be made to professionals who use the appropriate instruments and can interpret the results. Just as psychiatrists may specialize in the treatment of adult, child, or geriatric patients or may develop a special expertise with certain disorders, so psychologists also develop specialized skills in assessing certain groups of children or syndromes. Just as in medicine, board certification in clinical psychology or pediatric neuropsychology is one method for identifying psychologists with specialty qualifications. The American Psychological Association directory or various directories of certifying boards are useful references in the search for an appropriate examiner.

Although it is still best to state a referral question, referrals are frequently made for a particular type of evaluation (e.g., psychological vs. neuropsychological). A request for a psychological evaluation is commonly understood by the psychologist to mean assessment of personality and intelligence, with only a screening of achievement or other cognitive functioning, unless a more thorough examination is specifically requested. In contrast, a request for a neuropsychological evaluation is assumed to mean comprehensive assessment of cognitive functioning, including intelligence, complex mental processing, attention, memory, language, perception, and motor functioning, with only a screening of social and emotional behavior.

The primary purpose of neuropsychological evaluations is not to detect brain damage. Rather, comprehensive assessment of cognitive functioning, attention, memory language, perception, and motor functioning permits the identification of deficits that cause academic, social, or developmental disabilities or that impede psychiatric treatment. Such assessments also identify the child's areas of strength and weaknesses and help provide the basis for prescribing rehabilitative interventions that may be delivered as special education, occupational therapy, physical therapy, or vocational or cognitive rehabilitation. By parsing the psychological and cognitive impairments that contribute to the child's functional disability, such assessments also help delineate the interaction between the child's psychological disorder and cognitive impairments. Although

neuropsychologists are familiar with the cognitive sequelae of various neurological conditions and can be particularly helpful in the assessment of children with neurological disorders, any child requiring a detailed assessment of cognitive functioning is appropriate for referral.

Psychological Test Batteries

Table 36.1–8 contains a list of commonly used psychological measures for various domains of functioning. A comprehensive examination samples all of the domains, whereas a more focused evaluation selects a subset. In the most general sense, a battery of tests consists merely of a collection of tests to measure a variety of mental functions. There are loosely defined batteries that are routinely administered in psychological examinations. For example, the psychological battery for a child often includes the Wechsler Intelligence Scale for Children-V (WISC-V) or the Kaufman Assessment Battery for Children-III (K-ABC-III) as the measure of intelligence. In addition, there may be a measure of achievement in basic linguistic skills and math, such as the Test of Achievement from the Woodcock-Johnson Psychoeducational Battery—Fourth Edition (W/JPB-IV)—or the Wechsler Individual Achievement Test—Third Edition (WIAT-III). A test of visuomotor integration or visual construction, such as the Beery-Buktenica Test of Visual Motor Integration (VMI) or the Rey-Osterreith Complex Figure, is also usually administered. Finally, a variety of personality tests and behavior rating scales are used.

Some published test batteries consist of measures of various mental functions that have been developed together and can usually be combined into composite or factor scores. For infants and preschool children, psychometric instruments primarily focus upon development, including motor, sensory, perceptual, and language development and the development of conceptual reasoning. Comprehensive developmental quotients are derived from many of these instruments. The Bayley Scale of Infant and Toddler Development III (Bayley-III) and the Mullen Scales of Early Learning are two notable examples. The Woodcock-Johnson Psychoeducational Battery—Fourth Edition (W/JPB-IV)—and the NEPSY-2 are appropriate for school-aged children and examine a variety of smaller domains of functioning.

Personality Measures

Personality tests for the young school-aged child consist primarily of projective tests of personality. Projective tests contrast with objective tests of personality in that projective tests use limited structure and ambiguous stimuli to which the child can respond in a virtually unlimited variety of ways. Perhaps the best-known example of a projective technique is the Rorschach Test. Other projective tests commonly used for the young,

school-aged child include the Robert's Thematic Apperception Test (RTAT) and the Children's Appreciation Test (CAT), which require that the child compose a story in response to a picture depicting a scene that is suitable for children. A variety of tests administered to this age group require the child to draw, such as the Draw-A-Person (DAP) test, the House-Tree-Person (HTP) test, and the Kinetic Family Drawing (KFD). (Methods exist for scoring children's drawings to assess cognitive and visual motor development, as well as affective content.) For the most part, the interpretation of responses to projective tests of personality is inferential; that is, the meaning is not self-evident in the response. The inferences are based on accumulated clinical experience reported in the literature, symbolic interpretation or analogous reasoning based on a theory (e.g., psychoanalytic theory), and, to a much lesser extent, empirically established norms. The notable exceptions are the scores derived from the RTAT and the Exner and Rorschach-Performance Assessment System (R-PAS) scoring systems for the Rorschach test, which have norms and have been extensively studied in empirical research.

Objective personality tests are usually more structured than are projective tests, in that the stimuli have limited ambiguity, and only a select number of responses are permitted. In most objective measures, individuals are asked whether various statements are true or false for them. Objective measures of personality for younger children are usually limited to behavior rating scales or structured reports by parents or caregivers. The Personality Inventory for Children—Second Edition (PIC-2) and the Behavior Assessment System for Children—Second Edition (BASC-2) are good examples of these measures. There are objective tests of personality available for adolescents, which are self-reports. The Minnesota Personality Inventory-Adolescent (MMPI-A), the Millon Adolescent Personality Inventory (MAPI), and the Millon Adolescent Clinical Inventory (MACI) are three well-known examples. These tests yield personality profiles based on the relative deviations of scales that measure various psychological attributes.

Validity and Reliability

Although the validity and reliability of individual tests, as measures of specific mental functions, are established with group data, the reliability and validity of findings for the individual child are also a concern for the examining psychologist. Test batteries provide some advantages in determining the reliability and validity of individual test results. An integrated battery contains some overlap in the testing of mental functions, permitting comparative assessment of the consistency of the child's performance across measures. In addition, if there is a theoretical or empirical basis for knowing that some mental abilities are associated with each other, then consistent findings across measures of those abilities

strengthen the confidence in the findings for the individual child. Examinations that focus on a few mental functions, by administering a limited number of tests, do not permit detection of transient performance factors that may cause spurious scores. In addition, no test is a pure measure of the target variable; therefore, performance on a single test may be impeded by impairment of a mental function necessary for performing the task but not the function that the test purports to measure. If previous psychometric evaluations have been performed, reliability of findings for a particular child can be estimated from the consistent of findings across examinations.

Finally, the assessment of children with motor, language, hearing, or visual impairments presents a special challenge. Every psychological measure requires a modality of perception, mental processing, and modality of expression. To assess children with impairments in any of these functions, tests are selected with forms of input and expression that avoid the obviously impaired functions. Children with pervasive disorders of mental processing require assessment methods appropriate for their level of functioning. In such cases, it is frequently more helpful to characterize the impairments in terms of criterion-referenced measures, which list abilities mastered, or in terms of the child's standing among children of the same diagnostic group.

Table 36.1–8.
Commonly Used Psychological and Neuropsychological Tests for Children and Adolescents

General Ability/Intelligence Tests	For Ages (years)
Differential Ability Scales, Second Edition (DAS-II)	2:6–17:11
Kaufman Assessment Battery for Children, 2nd ed. (KABC-2)	3:0–18:11
Stanford-Binet, Fifth Edition (SB-5)	2:0–89:11
Wechsler Scales	
Wechsler Preschool and Primary Scale of Intelligence, 4th ed. (WPPSI-IV)	2:6–7:7
Wechsler Intelligence Scale for Children, 5th ed. (WISC-V)	6:0–16:11
Wechsler Adult Intelligence Scale, 4th ed. (WAIS-IV)	16:0–90:11
Wechsler Abbreviated Scale of Intelligence™ 2nd ed. (WASI-II)	6:0–90:11
Woodcock Johnson Tests of Cognitive Abilities, 4th ed. (WJ-IV)	2–90+ years
Achievement Test Batteries	
Kaufman Test of Educational Achievement, 3rd ed. (KTEA-III)	4:0–25:11
Peabody Individual Achievement Test—Revised (PIAT-R)	4:0–50:11
Wechsler Individual Achievement Test, 3rd ed. (WIAT-III)	4:0–85:0
Wide Range Achievement Test 4 (WRAT4)	5–94 years
Woodcock Johnson Tests of Achievement, 4th ed. (WJ-IV)	2–90+ years
Neuropsychological Test Batteries	
Delis-Kaplan Executive Function System (DKEFS)	8–89 years
Halstead-Reitan Neuropsychological Test Batteries	5–8, 9–14, 15+ years

Luria-Nebraska Neuropsychological Battery—Children's Revision	8–12 years
Luria-Nebraska Neuropsychological Battery (LNNB)	15+ years
NEPSY: A developmental neuropsychological assessment, 2nd ed. (NEPSY-II)	3–16 years
Reitan-Indiana Neuropsychological Test Battery	5–8 years
Nonverbal Tests[a]	
Comprehensive Test of Nonverbal Intelligence, 2nd Ed. (C-TONI-II)	6:0–89:11
Leiter International Performance Scale Revised, 3rd ed. (Leiter-3)	3:0–75+
Naglieri Nonverbal Ability Test-Individual Administration, 2nd ed. (NNAT-2TM)	4–18 years
Test of Nonverbal Intelligence, 4th ed. (TONI-4)	6:0–89:11 years
Beery-Buktenica Test of Visual-Motor Integration (VMI)	2–18 years
Rey-Osterreith Complex Figure	9 years-adult
Adaptive Behavior	
AAMR Adaptive Behavior Scale-School, 2nd ed. (ABS-S2)	3–21 years
AAMR Adaptive Behavior Scales-Residential and Community, 2nd ed. (ABS-RC2)	–79 years
Adaptive Behavior Assessment System, 2nd ed. (ABAS-II)	0–89 years
Scales of Independent Behavior-Revised (SIB-R)	infancy-80+ years
Vineland Adaptive Behavior Scales, 3rd ed.	0–90 years
Vineland Social-Emotional Early Childhood Scales (SEEC)	0–5:11
Diagnostic Instruments	
Child and Adolescent Psychiatric Assessment (CAPA)	8–17
Children's Interview for Psychiatric Syndromes (ChIPS)	6–18
Diagnostic Interview for Children and Adolescents (DICA)	6–18
Diagnostic Interview Schedule for Children (DISC-IV)	6–17
Schedule for Affective Disorders and Schizophrenia for School-aged Children (K-SADS)	6–18
Structured Clinical Interview for DSM-IV Childhood Diagnoses (KID-SCID)	7–17
Projective Tests	
Children's Apperception Test	3–10
Draw-a-Person Technique	5–17
Kinetic Drawing System for Family and School	5–20
Rorschach Inkblot Test	5-adult
Roberts Apperception Technique for Children: 2	6–18
Rotter Incomplete Sentence Blank	Adolescent-adult
Thematic Apperception Test	6-adult

[a]The KABC-2, DAS-II, and SB5 also yield nonverbal composite scores.

From Tsatsanis K. Psychological and neuropsychological assessment of children. In: Martin A, Volkmar FR, eds. *Lewis's Child and Adolescent Psychiatry: A Comprehensive Textbook.* 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2007:357–371.

Interpretation of Psychometric Findings

The interpretation of psychometric findings is predicated on standardized administration of the tests. In addition, the proper interpretation of psychometric findings generally requires knowledge of parametric and nonparametric statistics, mastery of the literature concerning test performance and test construction, familiarity with the construction and psychometric properties of the specific tests employed, and knowledge of

the literature regarding those instruments.

The results of psychological examination are often best integrated in a concluding discussion of the findings as they pertain to the referral question. The results of a neuropsychological examination are organized around a discussion of various functions with a concluding integration. The interpretation of test results includes consideration of the immediate context in which the tests were administered (e.g., performance factors, such as fatigue, anxiety, rapport, level of cooperation and motivation of the child). There are generally three levels of analysis for the test scores: (1) pathognomonic symptoms, (2) levels of performance, and (3) syndromes or patterns. There is no direct interpretation for single scores. During syndrome analysis, scores are interpreted within the context of the results of other measures, which is the most powerful form of analysis in terms of the ability to extrapolate beyond the immediate tests data. This capacity to extrapolate is based on known characteristics of children evidencing the syndrome. Finally, when testing children, one must bring a developmental perspective to findings, so that future difficulties can be anticipated with the changes that occur as the child grows older.

Psychometric findings are communicated more easily if the clinician receiving them understands some basic principles of psychometrics. Most people cluster around some average score for most measures of mental abilities, and fewer individuals score high or low. The distribution approximates a binomial distribution in the form of an inverted U, known as the *normal curve*. Raw scores are frequently the number of correct responses during a test and are not usually reported. Percentile scores indicate the percentage of persons obtaining raw scores equal to or less than the raw score obtained by the individual. In the normal distribution, small changes in raw scores close to the average (where many people are grouped) result in relatively large shifts in the percentile rank, whereas large shifts in raw scores at the high and low ends of the distribution may change the person's percentile rank little.

To compensate for the variations in the rate of change of percentile scores, and to make performance levels comparable across tests, raw scores are converted to standard, or deviation scores that indicate the extent to which the individual deviates from the mean or average. By convention, the average standard score for most psychological tests is 100, and the standard deviation is set at 15 IQ points or standard score points with these psychometric properties. By the use of standard scores, performance levels for measures of specific mental abilities can be compared with IQ or with other test results. Standard scores also give the best indication of significant differences among scores, because the intervals between scores of the same magnitude are equal anywhere along the distribution of scores. Statistical significance occurs when the probability is low that a difference between scores could have occurred by chance, as the result of inherent

error variance in the test. By convention, a level of .05 probability is set as the standard for statistical significance. Although two scores may significantly differ, that does not mean that the difference is meaningful or has clinical significance. Many statistically significant differences occur frequently in the normal population. For example, a difference of 15 points between the Verbal and Performance IQ scores of the WISC-IV is statistically significant; however, a difference of this magnitude or greater occurs in approximately one-fourth of the normal subjects used to standardize the test. Of course, a difference of this magnitude could have clinical significance in a particular case, if the adverse effect of the difference was exacerbated by other factors.

Some tests for children allow calculating standard scores from age norms or grade norms, for example, the Woodcock-Johnson Psychoeducational Battery and the Kaufman Tests of Educational Achievement. The difference between these two methods of calculation illustrates the importance of understanding tests scores in relation to the normative reference group. A standard score based on age norms compares the performance of a child to that of other children the same age. A standard score based on grade norms compares the child's performance with that of other children in the same grade. If the child has been retained in grade to compensate for learning problems, the standard score based on grade norms is artificially inflated, masking the extent of the child's impairment. However, an achievement score based on grade norms indicates how well the child has mastered the curriculum at the grade level. In general, grade scores (grade and months) and age scores (years and months), as opposed to standard scores, should be interpreted cautiously, because their psychometric properties are poor, and the implications of the scores vary across levels. Clinical decisions should not be based upon these types of scores.

There has traditionally been a concern about reporting IQ scores, because people do not understand that the scores are not absolute and immutable. Reporting the standard error of measurement or confidence intervals helps others realize that scores have inherent error and vary to some extent over successive administrations of the test, even when the test is valid and reliable by present standards. Confidence intervals are the range of scores believed to contain the true score at a specified level of probability. The WISC-IV and WJPB-III scoring programs, for example, will calculate confidence intervals for standard scores, at 90 percent and 95 percent probability. Higher levels of confidence necessitate larger ranges of scores.

Report of the Psychometric Findings

Ordinarily, the child's parents are informed of the results of a psychological or neuropsychological evaluation. Because tests scores are easily

misinterpreted, reports of psychological test findings should not be given to the child or parents unless a qualified professional is present to explain the findings and to answer questions. It is helpful if the referring clinician indicates how much feedback is to be given. In many situations, especially if the findings could create distress for the child or family, the referring clinician may wish to provide the feedback. If a written report is desired, then a report can be written for the family that does not include confusing technical information and is designed to minimize adverse effects of sensitive findings. In other cases, the referring clinician may choose to have the psychologist provide the feedback. Alternatively, the referring clinician may decide to meet with the psychologist and family together, so that the psychologist can explain any technical aspects of the evaluation and its findings, whereas the referring clinician can place the findings within the context and perspective of the overall assessment and treatment recommendations.

PEDIATRIC AND OTHER ADJUNCTIVE EVALUATIONS

Pediatric Evaluation

Close collaboration with the child's pediatrician is usually an important component of any evaluation. The pediatrician is a valuable source of information about the child's developmental and medical history and can contribute a unique long-term perspective on the child and family. In addition to a general review of systems, the pediatrician should assess the child's current and past physical growth, including head circumference, height, weight, and pubertal status. Screening of hearing and vision, with more detailed testing as indicated, is important because deficits in these areas can manifest themselves in learning, language, and social delays or seeming inattention or opposition. History taking includes abnormal movements or absence spells suggesting seizures, abnormal postures, and tics. The physical examination notes difficulties in muscle tone and strength; asymmetrical, abnormal, or primitive reflexes; tremor; and gait or postural disturbances.

Table 36.1–9.
Clues to Developmental Disability from the Physical Examination

Area	Findings	Implication
Growth	Head circumference	Microcephaly correlates highly with cognitive deficits
		Macrocephaly may indicate central nervous system abnormalities (e.g., hydrocephalus) or storage disorder
	Linear growth	Small size may indicate genetic disorder or embryopathy (e.g., fetal alcohol syndrome) or endocrine abnormality (e.g., thyroid)

		Large size may indicate genetic disorder (e.g., cerebral gigantism or Sotos' syndrome)
Congenital anomalies	Organ defects	Storage disorders (e.g., mucopolysaccharidoses); may have enlarged organs
	Major anomalies	Major anomalies, such as heart defects, may be associated with developmental delays
	Minor anomalies	Minor anomalies, such as wide-set eyes, low hairline, or extra digits, may suggest a genetic syndrome
Skin	Hyperpigmentation	—
	Tumors or fibromas	Skin lesions may be seen in neurocutaneous syndromes, such as neurofibromatosis or tuberous sclerosis.
Eye findings	Lens dislocation Corneal clouding Cataracts Retinal pigmentation	Abnormalities inside the eye may suggest metabolic disorders or congenital infections.

Adapted from Levy SE, Hyman SH. Pediatric assessment of the child with developmental delay. *Pediatr Clin North Am.* 1993;40:465.

Especially in children with intellectual disability or autism spectrum disorder, a thorough history and physical examination can provide important clues to an underlying detectable medical cause. For children with developmental difficulties, a careful history of perinatal risk factors is taken; the pediatric examination includes careful evaluation for major or minor congenital anomalies, dysmorphic features, dermatological abnormalities (hyperpigmentations, tubers, and fibromas), or ophthalmological findings (lens abnormalities and retinal pigmentation) (Table 36.1–9). Loss of previously acquired developmental milestones or the presence of congenital anomalies, dysmorphic features, microcephaly or macrocephaly and other growth problems, abnormal skin pigmentation or lesions, and symptoms suggesting seizure activity or other neurological abnormality warrant more detailed genetic, metabolic, or neurological evaluation (Table 36.1–9).

Diagnostic Laboratory Tests

The clinical usefulness and cost-effectiveness of routine laboratory tests for children who present with psychiatric problems have not been thoroughly studied. Most guidelines for performing these tests for children have historically been developed using data from studies of adults. Adult studies generally suggest that routine laboratory tests are not clinically useful in typical psychiatric settings, such as outpatient clinics and most inpatient units, although laboratory screening tests may be of some use in psychiatric settings in which patients are at high risk of medical illness, such as the emergency room, substance abuse treatment centers, AIDS clinics, and geriatric clinics, as well as for those who have new-onset psychosis, depression, or dementia or when evaluating children with developmental disturbances, such as intellectual disability or autistic spectrum disorder.

Similarly, routine screening laboratory measures are more likely to yield clinically significant information when clinical symptoms of physical illness are present.

Evidence suggests that if anything, routine laboratory testing is less useful in pediatric compared to adult populations, where medical comorbidity that causes psychiatric manifestation is less common. The few studies of the use of routine laboratory tests in child psychiatry patient populations have yielded similar conclusions. One review of routine laboratory screenings (e.g., thyroid function tests, electroencephalogram [EEG], chest x-ray, chemistry panel, urinalysis, complete blood count, electrocardiogram [ECG]) in 100 consecutive adolescent inpatient admissions reported variable rates of abnormal values, depending on the specific test, but in only 1 of these 100 patients did the tests produce a change in diagnosis from functional to organic—and even then the diagnostic information proved to have little clinical relevance. Most laboratory abnormalities in this study were regarded as minor and did not indicate a need for clinical follow-up.

Similarly, more specialized diagnostic laboratory evaluations (EEG, computed tomography [CT], magnetic resonance imaging [MRI], or chromosomal analyses) also provide a relatively low yield of clinically useful information. In a study of 200 consecutive child psychiatric inpatients, these evaluations were performed only when clinically indicated; despite their judicious use, the tests provided clinically relevant information in only 7 patients (3.5 percent of the total patient sample) or in 7 of 136 tests (5.1 percent of all tests performed). Chromosomal analyses proved to be the most informative of all these tests, yielding new medical diagnoses in 5 of 32 (15.6 percent) selected children for whom they were performed. A study of 111 putatively high-risk inpatients with new-onset adolescent psychoses produced similar conclusions. In this population, routine endocrine and neuroimaging screening evaluations failed to provide any information of diagnostic usefulness in any patient (although inconsequential laboratory abnormalities were present in 15.4 percent of the neuroendocrine screens and 11 percent of the neuroimaging screening tests). More specialized neuroimaging technologies, such as positron emission tomography (PET), single photon emission computerized tomography (SPECT), and functional MRI (fMRI), currently have no routine clinical or diagnostic usefulness in child and adolescent psychiatric populations.

Two useful screening diagnostic tests in child psychiatric practice include (1) urine drug screening and (2) testing for sexually transmitted diseases in high risk patients. Substance use can cause a wide range of neuropsychiatric symptoms and mimic a large spectrum of psychiatric conditions. The high prevalence of substance abuse in adolescent populations has prompted recommendations that toxicology screens be

obtained for (1) all adolescents who have psychiatric symptoms or who exhibit acute behavioral changes; (2) high-risk adolescents, such as delinquents and runaways; and (3) adolescents who have recurrent accidents or unexplained somatic symptoms. Similarly, children and adolescents should be screened for risk factors associated with sexually transmitted diseases. Children and adolescents at risk due to either a history of sexual activity (or sexual abuse) or intravenous substance use should be evaluated for sexually transmitted diseases, including HIV infection and syphilis.

In contrast to routine screening, targeted laboratory testing in specific child psychiatric syndromes can be quite useful in (1) detecting the etiology of conditions or (2) uncovering treatable, high-risk medical comorbidities. For instance, children with intellectual disability or autism spectrum disorder may particularly benefit from targeted diagnostic testing. In general, the more severe the child's disability, the greater is the likelihood of an identifiable (albeit not necessarily treatable) cause. For example, one study found that a medical cause could be identified in 60 to 70 percent of children with severe intellectual disability versus only 35 to 55 percent of children with mild intellectual disability. Chromosomal abnormalities, especially trisomy 21 (Down syndrome) or fragile X syndrome, account for more than one half of cases of intellectual disability. Children with autism spectrum disorder have relatively high rates of intellectual disability, seizure disorder (20 to 40 percent), and tuberous sclerosis (3 to 9 percent). Most experts would therefore advocate a Wood's lamp examination to search for tubers, an EEG to exclude seizures, and high-resolution chromosome analysis to exclude known genetic syndromes in children who display features of autism spectrum disorder or intellectual disability. Identification of genetic causes of autism is now possible in more than 10 percent of affected children and, as current knowledge of specific genetic syndromes advances, high-resolution DNA sequencing is likely to increase one's ability to discern specific subtypes of neurodevelopmental disorders.

Lead testing is also indicated in cases of intellectual disability because chronic lead ingestion due to pica can cause substantial retardation; conversely, many children with retardation or pervasive developmental disabilities from other causes indulge in pica and may accumulate a substantial lead burden.

A comprehensive discussion of every diagnostic testing recommended for every psychiatric disorder that manifests in childhood is beyond the scope of this chapter. The authors refer the reader to disorder-specific chapters of this textbook for a more comprehensive discussion regarding diagnostic laboratory testing for specific psychiatric conditions.

Finally, periodic laboratory testing (as well as tracking height and weight) is also important when starting certain psychiatric medications or monitoring for potential adverse during ongoing administration.

DIAGNOSTIC FORMULATION AND RECOMMENDATIONS

In the diagnostic formulation, the clinician attempts to integrate the data gathered through the assessment process into a coherent account of the child's difficulties, the factors that appear to have predisposed the child to develop the problem, the concomitants and consequences of the problem, and the factors that maintain the problem or that might ameliorate it. It is not always possible to provide a complete formulation at the conclusion of the initial assessment; under such conditions, the best that the clinician may be able to provide is a differential diagnosis that can point toward the appropriate treatment options and the subsequent steps needed to establish a diagnosis.

One product of the assessment may be the assignment of a diagnosis of one or more categorical psychiatric disorders using the diagnostic schema of DSM-5 or ICD-10. Although these nosological systems have done much to standardize diagnostic practices and to improve the quality of research, one must be aware of their clinical limitations as applied to children. Many children experience sufficient symptoms, distress, or impairment to be regarded as psychiatrically disturbed and to warrant therapeutic intervention without meeting the full criteria of a categorical psychiatric disorder. Furthermore, as currently conceptualized, the DSM-5 classification is descriptive and atheoretical; it specifies neither presumed etiology nor treatment for the various diagnostic categories. If the assessment process is to provide a picture of the child that suffices to guide intervention, the resulting diagnostic formulation must go beyond simply assigning a categorical psychiatric diagnosis (if one is warranted). As noted in the introduction, comorbidity is often the rule rather than the exception in childhood, and many children have complex patterns of developmental and behavioral difficulties that are not adequately captured by a list of multiple categorical diagnoses. Children who share a given categorical diagnosis may differ in crucial respects that influence current severity, natural history, and treatment response. Examples of potential risk (or exacerbating) factors and protective (or ameliorative factors) include intelligence, compensatory skills, family and community resources and supports, and vulnerabilities, such as comorbid pathological personality traits, neurological impairments, or co-existing general medical conditions.

The clinician's diagnostic formulation should thus supplement the assignment of any formal categorical diagnoses by attempting to identify, as far as possible, the potential causes, predisposing factors, and current determinants of the child's difficulties. On the basis of this information and the clinician's expertise regarding the various interventions available for the diverse forms of childhood psychopathology and developmental difficulties, the clinician formulates appropriate recommendations for the treatment of the child's problems.

Communicating Findings and Recommendations

Communicating the findings and recommendations that result from the assessment to the parents and child is an essential part of the evaluation process and may require one or more sessions. Depending on the problem and the child's age and level of comprehension, presenting the findings usually involves meeting with the child and parents, separately or together.

Several principles must be observed to maximize the likelihood that the parents and child hear, understand, and experience the clinician's findings and recommendations as helpful. First, the clinician must communicate a sense of the child as a whole person, with strengths and abilities, as well as problems and vulnerabilities. This perspective conveys a sense of the clinician's appreciation and empathic understanding of the child, reduces defensiveness, and helps to mitigate the fear that there is only bad news.

The findings and recommendations are most useful to the family and likely to be implemented if the interpretive session is a dialogue rather than a lecture. The clinician must use language comprehensible to the parents and child, must avoid jargon, and must keep technical terms to a minimum. When technical terms are necessary, they should be explained clearly, so the parents and child understand them; terms that are merely descriptive to the clinician may hold frightening or negative connotations for the child or family. The clinician should allow ample opportunity for the parents and child to discuss the recommendations. This provides a chance to judge whether the clinician's formulation made sense to the family, to address any differences of opinion, and to explore the feasibility and acceptability of the recommendations offered. Remaining areas of uncertainty or ambiguity about the assessment or recommendations should be clearly noted and discussed.

If the assessment is a consultation requested by an agency, a school, or another clinician, the findings and recommendations should be communicated to the referring party in appropriate terms after being discussed with the parents and child and with their consent. If further evaluation or treatment is indicated but is best done by someone else, the clinician should offer to assist in appropriate referral.

References

Angold A, Erkanli A, Copeland W, Goodman R, Fisher PW, Costello EJ. Psychiatric diagnostic interviews for children and adolescents: a comparative study. *J Am Acad Child Adolesc Psychiatry*. 2012;51(5):506–517.

Angold A, Fisher PW. Interviewer-based interviews. In: Shaffer D, Lucas CP, Richters JE, eds. *Diagnostic Assessment in Child and Adolescent Psychopathology*. New York, NY: Guilford Press; 1999:34–64.

Bird HR, Canino GJ, Davies M, et al. The Brief Impairment Scale (BIS): a multidimensional scale of functional impairment for children and adolescents. *J Am Acad Child Adolesc Psychiatry*.

2005;44(7):699–707.

Bostic JQ, King RA. Clinical assessment of children and adolescents: content and structure. In: Martin A, Volkmar FR, eds. *Lewis Comprehensive Textbook of Child and Adolescent Psychiatry*. 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2007:323–344.

Canino GJ, Fisher PW, Alegria M, Bird HR. Assessing child impairment in functioning in different contexts: implications for use of services and the classification of psychiatric disorders. *Open J Med Psychol*. 2013;2(1):29–34.

*Fisher PW, Chin EM, Vidair HB. Use of structured interviews, rating scales, and observational methods in clinical settings. In: Tharper A, Pine DS, Leckman JF, Scott S, Snowling MJ, Taylor E, eds. *Rutter's Child and Adolescent Psychiatry*. 6th ed. Chichester: Wiley; 2015:419–435.

*Gilliam WS, Mayes LC. Clinical assessment of infants and toddlers. In: Martin A, Volkmar FR, eds. *Lewis's Comprehensive Textbook of Child and Adolescent Psychiatry*. 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2007:309.

Kaufman AS, Raiford SE, Coalson DL. *Intelligence testing with the WISC-V*. Hoboken, NJ: Wiley; 2016.

Leckman JF, Tayler E. Clinical assessment and diagnostic formulation. In: Tharper A, Pine DS, Leckman JF, Scott S, Snowling MJ, Taylor E, eds. *Rutter's Child and Adolescent Psychiatry*. 6th ed. Chichester: Wiley; 2015:419–435.

*Loeber R, Green SM, Lahey BB. Mental health professionals' perception of the utility of children, mothers, and teachers as informants on childhood psychopathology. *J Clin Child Psychology*. 1990;19:136–143.

Shaffer D, Fisher P, Lucas C. Respondent-based interviews. In: Shaffer D, Lucas CP, Richters JE, eds. *Diagnostic Assessment in Child and Adolescent Psychopathology*. New York, NY: Guilford Press; 1999:3–33.

*Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. NIMH Diagnostic Interview Schedule for Children, Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *J Am Acad Child Adolesc Psychiatry*. 2000;39(1):28–38.

Sparrow SS, Cicchetti DV, Saulnier CA. Vineland Adaptive Behavior Scales, Third Edition(Vineland-3). Pearson; 2006. Available at www.pearsonclinical.com.

Towbin KE. Physical examination and medical investigation. In: Tharper A, Pine DS, Leckman JF, Scott S, Snowling MJ, Taylor E, eds. *Rutter's Child and Adolescent Psychiatry*. 6th ed. Chichester: Wiley; 2015:449–459.

*Tsatsanis KD. Psychological and neuropsychological assessment of children. In: Martin A, Volkmar FR, eds. *Lewis's Comprehensive Textbook of Child and Adolescent Psychiatry*. 4th ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins; 2007:357–371.

▲ 36.2 Psychiatric Assessment of Preschool Children

HELEN LINK EGGER, M.D., AND CAROLY SUE PATAKI, M.D.

INTRODUCTION

The infant/preschool mental health field has made real strides in recent years in demonstrating that clinically significant psychiatric symptoms and disorders are already present and highly impairing in early childhood. Recent epidemiologic studies of children ages 2 to 5 years old have shown that the overall rate of psychiatric disorders in preschoolers is remarkably similar to overall rates reported for older children and adults. For example, the preschool rate of "serious emotional disturbance," defined as psychiatric disorder(s) that are associated with significant impairment, was 12.1, 9.1, and 7.1 percent in three different community preschool studies, compared with a median rate across community studies of 13.2 percent in older children and adolescents and 15.6 percent in adults. The growing recognition of the mental health needs of preschool children reveals a large gap in mental health services. Few clinicians have training in the assessment and treatment of young children. This section provides an overview of the components of comprehensive, empirically based psychiatric assessments of preschool children and their families.

DEVELOPMENTALLY APPROPRIATE CRITERIA FOR PSYCHIATRIC SYMPTOMS AND DISORDERS IN PRESCHOOL CHILDREN

Historically, young children's emotional and behavioral problems have been approached dimensionally, with cut-points set to characterize subsets of children at the extreme of the distribution of (1) normative behaviors and emotions and/or (2) temperament traits. In these models, extremes of problems or temperament traits have typically been regarded as precursors of, or risk factors for, later psychopathology rather than as psychiatric disorders similar to those identified at later stages of life. Although symptom checklists or temperament measures often provide clinically useful information about the young child, they do not have the symptom coverage (particularly for rare but highly concerning symptoms such as suicidality or stereotypies) nor details about the characteristics of symptoms (e.g., frequency, onset, duration) to enable the clinician to make a diagnosis. There have been two approaches to the diagnostic classifications of early childhood psychopathology (1) the *Diagnostic and Statistical Manual of Mental Disorders/International Statistical Classification of Diseases and Related Health Problems* (DSM/ICD) diagnostic classification, either unmodified or modified to be developmentally appropriate for young children, and (2) the Diagnostic

Classification: 0–3 (DC:0-3, DC:0-3R), an alternative system for classifying mental health disorders in infants and toddlers. These categorical approaches seek to identify clinically significant syndromes, characterized by severity, pervasiveness, persistence, and impairment, that are early-onset "disorders" rather than simply risk factors for later disorders.

Despite its lack of developmental specificity, the DSM has been a useful starting point for examining how psychopathology presents in early childhood, at least in children 2 years old or older. For example, the diagnostic criteria of the fifth edition of the DSM (DSM-5) work quite well as a means of advancing the current understanding of the presentation and course of autism spectrum disorders despite its lack of developmentally specific criteria. The dual approach of testing the applicability of current diagnostic criteria for identifying autism spectrum disorders in young children while exploring the validity and clinical utility of developmentally specific criteria and/or diagnostic algorithms for use with young children provides an exemplary road map for examining the validity of specific behavioral and emotional disorders in young children. Similar approaches are being taken in studies of early-onset attention-deficit/hyperactivity disorder (ADHD), anxiety disorders, depression, posttraumatic stress disorder, and disruptive behavior disorders in young children. In 2003, the Research Diagnostic Criteria–Preschool Age (RDC-PA) was published with the goal of facilitating further research on the diagnostic validity of psychiatric disorders in preschoolers by defining clearly specified, developmentally appropriate criteria for preschool psychopathology.

The RDC-PA explicitly sought to modify the DSM criteria to make them developmentally appropriate rather than to propose a new classification system. The DC:0-3 combines the RDC-PA "top-down" approach with a "bottom-up" approach that seeks to classify disorders in infants and toddlers that are not covered in the DSM. The original DC:0-3 was published in 1994. Despite its title, the DC:0-3 has commonly been used with children through age 5 years. The DC:0-3 classification system included alternative criteria for DSM-5 diagnoses (e.g., many of the Research Diagnostic Criteria [RDC] modifications have been incorporated in the revised edition), new diagnostic categories (e.g., regulation disorders of sensory processing), and a revised multiaxial system that places relationship characteristics and disorders on Axis II and the child's "functional emotional developmental level" on Axis V. The current DSM-5 no longer utilizes the multiaxial system, rather it employs specifiers, making the DC:0-3R even more important for preschool diagnoses. The revised version of DC:0-3 (DC:0-3R) was published in 2005 with a central goal of operationalizing the diagnostic criteria and categories so as to improve diagnostic reliability and facilitate research into the validity of these disorders. Despite its differences from the DSM, DC:0-3R, like the first version, was intended to supplement, not replace, the DSM/ICD

nosologies. The DC:0-3R explicitly provides a place on Axis 1 for coding DSM/ICD diagnoses.

What System Should a Clinician Use?

At this point, it is quite certain that the current diagnostic criteria defining preschool psychopathology are imperfect and will be revised and even supplanted as knowledge grows. For now, however, there is fairly convincing evidence that all these different approaches to classification contribute to identifying clinically significant psychopathology in early childhood. The clinician who seeks to provide mental health care for preschoolers and their families should be familiar with each of these approaches and use one or both of the diagnostic multiaxial nosologies for their assessments. Clinicians who use the DC:0-3 axes commonly may need to make a "cross-walk" to the DSM-5 or because DSM/ICD codes are often needed to obtain insurance reimbursement for services.

COMPREHENSIVE, EVIDENCE-BASED PSYCHIATRIC ASSESSMENTS OF PRESCHOOL CHILDREN

There are five key components of a comprehensive psychiatric assessment of a preschool child: (1) multiple sessions, (2) multiple informants, (3) a multidisciplinary approach, (4) multiple modes of assessment, and (5) a multiaxial diagnostic formulation. These components arise from the multidisciplinary, developmental, relationship-based, family-centered, and prevention-focused orientations that characterize the infant/preschool mental health field's approach to understanding psychopathology in early childhood.

Multiple Sessions

Despite the real reimbursement constraints that clinicians face in conducting multisession evaluations, these constraints cannot dictate the length of an evaluation. It is simply not possible to conduct an adequate psychiatric assessment of a young child in one session. Although it could be argued that this caveat is true for all of pediatric psychiatry, the rapid developmental, cognitive, emotional, and behavioral changes that characterize the preschool period, as well as the young child's physical and emotional dependency on his or her primary caregivers, make it imperative that information be gathered about the child across sessions and across time. For example, a tired 3-year-old child seen after a long day at preschool may look quite different when seen in the morning after he or she has had a good night's sleep. At some preschool clinics the clinicians choose to see the parent(s) without the child for one of the sessions so as to obtain a detailed history about the child's symptoms and the family history and current family structure and functioning. It is essential to obtain

information about current and past parental psychiatric and substance abuse problems and treatments, domestic violence in the home, and marital (or partner) difficulties, but it is not appropriate to discuss these issues in front of the child. Nor is it appropriate for the child to hear negative comments about himself or herself (e.g., "He is a bad kid") or to be privy to a parent's worst fears ("I am afraid that he will end up like my brother who killed himself"). These comments and fears are very important pieces of information about how the parent sees this child. Any parent of a preschooler also knows that it is quite difficult for young children to refrain from interrupting adult conversation long enough for the clinician to obtain the breadth and level of detail needed to understand the child's problems. However, even if the child is seemingly absorbed in play in the corner of the office, he or she will hear the adult conversation. An evaluation should include phone sessions with the child's teacher or child care provider (written consent to contact the teacher or child care provider should be obtained at the first session). A visit to the child's day care or preschool to observe the child's functioning with other adults and peers in an out-of-home setting may also be indicated.

Multiple Informants

The clinician must gather information from multiple informants. First and foremost, the child is an important informant. This is as true of a 2-year-old with limited expressive language as of a 5-year-old about to enter kindergarten. Information can be obtained by speaking with and/or playing with the child (using structured or unstructured approaches) and observing the child interacting with parents, with another clinician, and with teachers or child care providers or peers. It is not possible to assess or diagnose a young child based on parent report alone. Moreover, the clinician should make every effort to spend time alone with the child, without the parents present (severe separation anxiety is an example of a disorder that might make this very difficult).

Of course, an evaluation cannot be conducted without information from the child's parents. These may be the child's biological parents and/or adults who act as parents of the child such as stepparents, grandmothers, or foster parents. An evaluation should include a three-generation genogram identifying the important parental figures and family members (including psychiatric and substance abuse history for each member). Ideally, if there are two parents in the home, both of them will attend the evaluation sessions and fill out all measures and information separately. Often each parent provides a somewhat different description of the child's symptoms. This does not undermine the validity of these reports, but rather reflects the fact that the child shows different aspects of himself or herself within different relationships. This information is diagnostically important and may be useful in guiding treatment planning. The parents

will also provide information about themselves (e.g., current psychiatric symptoms and treatment, parental and marital stress). Information must also be obtained from the child's nonparental caregivers such as teachers or child care providers, nannies, and babysitters who spend a significant amount of time with the child. As described in more detail in the next section, other clinicians, such as the child's pediatrician, are also critical informants.

Multidisciplinary Assessment

A child psychiatrist conducting a preschool psychiatric assessment usually collaborates closely with other clinicians and care providers, including psychologists, primary care and developmental pediatricians, neurologists, occupational therapists, speech and language therapists, early interventionists, case managers, adult psychiatrists and psychotherapists who treat the disorders from which the parents may be suffering, and marital therapists. The child psychiatrist has an opportunity (if not obligation) to assume a leadership role among all of these care providers. At the evaluation stage, the child psychiatrist should integrate the assessments and recommendations made by other providers into his or her assessment and then facilitate treatment planning and coordination of communication among these providers. The child psychiatrist should never provide treatment for a young child, particularly treatment with a psychotropic medication, without having conducted a comprehensive assessment himself or herself.

Multiple Modes of Assessment

Because it is not possible to conduct a standard psychiatric interview with a young child about his or her emotions and behaviors, preschool mental health assessments must draw on information gathered using different approaches, including adult reports, observational assessments, structured or unstructured play/interactions with the child, and standardized cognitive, motor, language, and other developmental assessments. A physical exam (often conducted by the child's pediatrician) is essential. Laboratory measures, including genetic testing, may be indicated.

When possible, empirically tested measures with demonstrated reliability and validity should be used to assess child and family domains. Types of measures include symptom checklists, such as the Child Behavior Checklist 1½–5 or the Infant Toddler Social Emotional Assessment for children 12 to 36 months old, DSM-referenced rating scales such as the Early Childhood Inventory-4 (ages 3 to 6 years), and checklist measures of specific symptom clusters such as the ADHD Rating Scale or the Preschool Anxiety Scale. Teacher/Caregiver measures such as the Caregiver-Teacher Report Form, as well as those assessing social skills and peer relationships,

such as the Social Skills Rating System, are also useful. Similar types of rating scales are available to assess adult domains, such as psychopathology (e.g., Beck Depression Inventory), parental stress (e.g., Parenting Stress Index), or domestic violence (Conflict Tactic Scales).

In the future, structured diagnostic parent interviews for assessing preschool psychopathology will be the best way to obtain comprehensive psychiatric, developmental, and environmental information about the child and the family, but these measures are not yet user-friendly enough for use in clinical settings. The exception is the Autism Diagnostic Interview (ADI). The Preschool Age Psychiatric Assessment (PAPA) is currently the only comprehensive parent psychiatric interview for assessing psychiatric symptoms and disorders in children ages 2 to 5 years with demonstrated test–retest reliability and validity. A version of the PAPA for assessing nonpervasive developmental disorder (PPD) symptoms in children with autism and other PPDs (PAPA-PDD) was scheduled to be completed in 2008. The ePAPA, a web-based version of the PAPA (and PAPA-PDD) to be administered on a tablet personal computer, has recently been developed and will be available for use in clinical settings in the near future.

Direct "interviews" with the young child about his or her feelings and experiences are also an essential component of a comprehensive assessment. In most clinical practices, unstructured play is used to learn about the child's emotions. Structured child interviews, such as the Berkeley Puppet Interview (BPI) and the MacArthur Story-Stem Battery (MSSB), show significant promise for children 4 years old and older.

Observational assessments, particularly of parent–child interactions, play an important role in preschool assessments. Inspired by the success of the Autism Diagnostic Observational Schedule—a structured assessment for autism in which the examiner "presses" for PDD symptoms with specific tasks with the child—researchers have been developing a similar diagnostic observational measure for assessing disruptive behavior disorders in preschoolers (Disruptive Behavior-Diagnostic Observational Schedule [DB-DOS]). The same group is developing a similarly structured observational measure of emotional disorders for use with preschool children.

There are numerous standardized, developmentally specific measures for assessing the preschooler's status in cognitive, language, motor, and preacademic domains. Reviews of specific measures in this and the other areas mentioned in this section can be found in the textbooks listed in the reference section.

Multiaxial Assessment

A comprehensive psychiatric assessment in the preschool age group must be truly multiaxial, with each axis reflecting a developmental framework. Since DSM-5 no longer uses a multiaxial system, coding on DC:0-3 system reflects an understanding of the different presentations of psychiatric

symptoms from ages 2 to 5 years. Familiarity with the current consensus on developmentally appropriate criteria for psychiatric symptoms and disorders during the preschool period is essential. Valuable resources include the *Handbook of Infant Mental Health*, the *Handbook of Preschool Mental Health*, and the *Handbook of Infant, Toddler, and Preschool Mental Health Assessment*.

The child's symptoms across all axes must be understood within the child's current level of functioning across multiple developmental domains (e.g., cognitive functioning, expressive and receptive language, gross and fine motor skills, prekindergarten knowledge, social-emotional skills). These developmental assessments are important for (1) understanding the child's psychiatric symptoms within his or her developmental capacities (not simply the child's chronological age), (2) identifying developmental lags or deficits that can be effectively addressed (e.g., expressive language problems), and (3) ruling out (or ruling in) the presence of a developmental disorder such as autism. Because children 3 years old and younger (and in some states 5-years-old and younger) are eligible for early intervention services for developmental delays, it is critical that child psychiatrists identify these delays and provide appropriate treatment referrals.

Assessment of the context of the child's symptoms, as well as the child's environment within and outside of the family, is also vital. The pervasiveness of the child's symptoms, not only across settings, but also across relationships, will provide important information about the nature of the symptoms. For example, if a child has explosive tantrums that occur when the child is at home and at school and occur with the parent and with the child's day care provider, the child may well be showing a symptom of a behavioral disorder. The child who has similar tantrums only when separating from his or her mother may be experiencing a symptom of an anxiety disorder. The child who only has tantrums when being dropped off for a custody visitation with a divorced parent may be expressing distress that arises from the difficulty of moving back and forth between two parental homes. The child who only has tantrums with his or her mother but shows good emotional and behavioral regulation with all other adults may be showing signs of a relationship disturbance with the mother. The infrequent tantrums that occur only when a child is exhausted (e.g., after a long car trip traveling home from a visit with Grandma) may very well be developmentally normal. These examples highlight the importance of obtaining detailed examples of the child's behaviors with a focus on the characteristics of symptoms. Relevant questions include the following: What does the child do or say? How often does it happen? How long does it last? When did it start? What triggers the symptom? When, where, and with whom does the symptom occur? What helps? This descriptive information is not the same as asking the parent "Why does your child do that?" Often, one does not know "why," but detailed descriptions of specific

symptoms provide the information needed to rule in or rule out specific disorders. An important difference between DSM and DC:0-3R is that the DC:0-3R provides a place (Axis II) to record information about the nature and quality of the child's relationships with his or her primary caregivers and list relationship disorders. The DSM does not provide a place to record any relationship information included in the DC:0-3R text is the Parent–Infant Relationship Global Assessment of Functioning Scale and a relationship problems' checklist that provide a framework for formulating an understanding of the child's relationships with his or her primary caregivers.

Assessment of the child's environment is also essential. In the DC:0-3R systems, stressors and risk factors are listed on Axis IV. DC:0-3R includes a broad (but not exclusive) checklist of known stressors, along with a place to indicate the age of onset, duration, severity, and context of these stressors, which is meant to provide a guide to the assessment of this domain. Axis IV should reflect a careful history of the child's environment, including, to name a few, an understanding of (1) parental/family factors such as a family history of psychiatric and substance abuse disorders and treatment, including types and doses of psychotropic medications and types of psychosocial treatments (a history of postpartum depression should be obtained), family structure and functioning including a careful assessment for physical or emotional abuse or neglect and sexual abuse, parent–partner relationship functioning including an assessment of adult–adult domestic violence and emotional abuse, and sibling relationships; (2) pregnancy and birth factors, including premature birth, low birth weight, maternal smoking or use of other substances during pregnancy, neonatal complications, and maternal depression during or following the child's birth; (3) the child's school or day care experiences, including teacher–child ratios, frequency of changes in child care settings and providers, amount of time the child is away from parent(s), age when the child started out-of-home care, peer relationships, and relationships with nonparental child care providers; (4) community factors, including neighborhood safety and proximity of relatives; and (5) socioeconomic factors, including parental employment status, family's social economic status, and absence of health insurance. Major stressors (e.g., parental divorce, death of a parent, physical abuse, hospitalization, or natural disasters such as a house fire or hurricane) known to be potentially traumatic life events should be included. However, chronic stressors or those that may seem "minor" (e.g., change in day care, moving to a new house) should also be included on Axis IV. There is clear evidence that the accumulation of multiple stressors with varying intensities has as profound an impact as single traumatic events in early childhood, just as they do in later childhood and adolescence. Both acute and cumulative stressors significantly raise the risk that the child will have one or more psychiatric disorders and be impaired.

CLINICAL FORMULATION AND TREATMENT PLANNING

Clinical formulation is an integrative skill that involves the pulling together of multiple observations and sources of information about an individual within a general diagnostic scheme so as to guide decision making. It is worth emphasizing that a psychiatric nosology classifies disorders, not individuals. When a comprehensive psychiatric evaluation has been conducted with a preschool child and his or her family and the DSM and/or DC:0-3R axes have been completed, a clinical formulation about the child and a treatment plan addressing each of the child's problem areas should be developed. Clearly defined symptoms within the child must be identified, ideally using measures that are sensitive to change, so that the clinician will be able to assess the efficacy of the intervention and the outcome of treatment. Of course, the assessment may also identify targets for treatment beyond the child's mental health symptomatology. For example, a treatment plan might include referral of the child for speech or occupational therapy or referral of a parent for treatment of his or her depression.

The evidence is quite clear that approximately 10% of preschoolers have severe impairing psychiatric disorders that cause suffering and put them at risk of ongoing emotional and developmental problems. These children and their families deserve effective and safe treatment. Unfortunately, the evidence base for psychiatric treatments for preschoolers is very thin. Over the next decade there is every reason to hope that there will be major advances in the current understanding of the etiology and course of early-onset mental health problems, as well as the development and dissemination of effective treatments and preventive interventions. However, children and their families cannot wait for the research to catch up with this need. Clinicians who treat preschoolers must (1) know the current state of evidence on the assessment and treatment of preschool psychopathology, (2) know how to decide on and determine the efficacy of treatments that may not have been studied in young children, and (3) be committed to keeping up with and incorporating new findings into their practice. Clinicians and parents should together weigh the risk of the disorder(s) with the risk of treatments so that parents can make fully informed decisions about how to best care for their child. Single-case experimental designs (e.g., *n*-of-one trial) that seek empirically to determine the efficacy of treatments for preschool patients are the best ways to manage these risks. This is particularly vital when use of psychotropic medication is being considered. The guidelines for the psychopharmacological treatment of preschoolers recommend this empirically driven (and cautious) approach to treatment. The infant/preschool mental health field has always had an early intervention and prevention orientation. The hope is that the identification and treatment of preschool children with clinically significant psychiatric

symptoms and disorders will not only ease their suffering during early childhood, but also will contribute to ameliorating or even preventing the onset of new disorders or progression of these disorders into later childhood and adulthood.

REFERENCES

Angold A, Egger HL. Psychiatric diagnosis in preschool children. In: DelCarmen-Wiggins R, Carter A, eds. *Handbook of Infant, Toddler, and Preschool Mental Health Assessment*. New York: Oxford University Press; 2004.

Carpenter KLH, Sprechmann P, Calderbank R, Sapiro G, Egger HL. Quantifying risk for anxiety disorders in preschool children: a machine learning approach. *PLoS ONE*. 2016;11(11):e0165524.

Carter AS, Briggs-Gowan MJ, Jones SM, Little TD. The Infant-Toddler Social and Emotional Assessment (ITSEA): Factor structure, reliability, and validity. *J Abnorm Child Psychol*. 2003;31:495.

Costello EJ, Egger HL, Angold A. 10-year research update review: The epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *J Am Acad Child Adolesc Psychiatry*. 2005;44:972.

Egger HL, Angold A. Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *J Child Psychol Psychiatry*. 2006;47:313.

Egger HL, Erkanli A, Keeler G, Potts E, Walter B. The test–retest reliability of the Preschool Age Psychiatric Assessment (PAPA). *J Am Acad Child Adolesc Psychiatry*. 2006;45:538.

*Franz L, Angold A, Copelane W, Costello EJ, Towe-Goodman N, Egger H. Preschool anxiety disorders in pediatric primary care: prevalence and comorbidity. *J Am Acad Child Adolesc Psychiatry*. 2013;52:1294–1303.

Gleason MM, Egger HL, Emslie GJ, Greenhill LL, Kowatch RA. Psychopharmacological treatment for very young children: contexts and guidelines. *J Am Acad Child Adolesc Psychiatry*. 2007;46:1532.

Hamre Sveen T, Berg-Nielsen TS, Lydersen S, et al. Screening for persistent psychopathology in 4-year-old children. *Pediatrics*. 2016;138(4):e20151648.

*Luby J. *Handbook of Preschool Mental Health: Development, Disorders, and Treatment*. New York: Guilford Press; 2006.

Martin SE, Hunt JI, Mernick LR, et al. Temper loss and persistent irritability in preschoolers: implications for diagnosing disruptive mood dysregulation disorder in early childhood. *Child Psychiatry Hum Dev*. 2016. [Epub ahead of print]

Narrow WE, First MB, Sirovatka PJ, Regier DA. *Age and Gender Considerations in Psychiatric Diagnosis: A Research Agenda for DSM-V*. Arlington, VA: American Psychiatric Association; 2007.

Sim F, Haig C, O'Dowd J, et al. Development of a triage tool for neurodevelopmental risk in children aged 30 months. *Res Dev Disabil*. 2015;45–46:69-82.

Task Force on Research Diagnostic Criteria. Infancy and preschool: Research diagnostic criteria for infants and preschool children: The process and empirical support. *J Am Acad Child Adolesc Psychiatry*. 2003;42:1504.

Towe-Goodman NR, Franz L, Copeland W, ANgold A, Egger H. Perceived family impact of preschool anxiety disorders. *J Am Acad Child Adolesc Psychiatry*. 2014;53:437–446.

Tseng WL, Guyer AE, Briggs-Gowan MJ, et al. Behavior and emotion regulation modulation deficits in preschoolers at risk for bipolar disorder. *Depress Anxiety*. 2015;32:325–334.

*von Klitzing K, White LO, Otto Y, Fuchs S, Egger HL, Klein AM. Depressive comorbidity in preschool anxiety disorder. *J Child Pychol Psychiat*. 2014;55:1107–1116.

*Wichstrom L, Berg-Nielson TS, Angold A, Egger HL, Solheim E, Sveen TH. Prevalence of psychiatric disorders in preschoolers. *J Child Psychol Psychiat*. 2012;53:696–705.

Zeanah CH, Carter AS, Cohen J, et al. Diagnostic classification of mental health and developmental disorders of infancy and early childhood DC:0-5: selective reviews from a new nosology for early childhood psychopathology. *Infant Ment Health J*. 2016;37(5):471–475.

*Zero to Three: Diagnostic Classification. *0–3R: Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood: Revised Edition*. Washington, DC: Zero to Three Press; 2005.

Zucker N, Copeland W, Franz L, Carpenter L, Keeping L, Angold A, Egger H. Psychological and psychosocial impairment in preschoolers with selective eating. *Pediatrics*. 2015;136:e582–e590.

▲ 36.3 Neuropsychological and Cognitive Assessment of Children

MARTHA BATES JURA, PH.D., AND LORIE A. HUMPHREY, PH.D.

INTRODUCTION

This section describes cognitive and neuropsychological assessment as ways of understanding the referral issues of children and adolescents, and how the processing of information can affect a wide range of adaptive functions. Although these assessments might overlap, these approaches analyze behavior according to two different paradigms. Cognitive assessment is undertaken without reference to the possible neurobiological underpinnings of overt behavior, and it describes the patient very much as others might observe him or her in the world. Neuropsychological assessment is undertaken in the context of growing knowledge about brain–behavior relationships, and it has the additional possibility of describing the child in terms of unseen neural pathways. These approaches provide ways of conceptualizing how children integrate information (and their thinking, learning, and responding) at different levels. General cognitive assessment focuses on understanding behavior at a cognitive level and in descriptive terms. Focused neuropsychological assessment introduces the additional possibility of understanding behavior at neural levels and in neurobiological terms. However, regardless of the theoretical differences among the psychologists undertaking these evaluations, in a practical way it is the patient's referral issues that shape the assessment process and focus the interpretation of results.

BASICS OF PSYCHOLOGICAL ASSESSMENT

Psychological assessment involves more than testing. Although measurements are useful, testing involves more than scores.

The Testing Process

In addition to testing, assessment procedures include examining past records (medical examinations, prior testing, report cards), interviewing the client and his or her family (in structured and unstructured formats), obtaining information from home and school (and, sometimes, on-site observations), and getting rating scales filled out by the child's parents and teachers (regarding developmental, behavioral, emotional, and diagnostic issues). The diagnostic aspect of the process involves an attempt to determine the psychiatric and educational categories for which the client meets the criteria. Cognitive and neuropsychological testing are only aspects of an attempt to get a broad view of the way a child solves problems in the world, to understand his or her unique interaction with any diagnostic category, and to provide recommendations for interventions.

Psychologists collect information on instruments with a variety of psychometric properties. Behavior forms vary greatly in whether they involve noting the presence or absence of a symptom or include ways of rating and scoring the information. As controlled observations, test instruments involve not only ways of sampling behavior and measuring it, but also ways of noting the way in which the patient goes about solving problems. Although the present discussion about cognitive and neuropsychological assessment focuses on structured (standardized) testing, unstructured testing (such as projective instruments, whether scored or not) also can give a good sense of the cognitive organizing principles that patients bring to bear on a wide range of tasks and life. In fact, when looked at in basic ways, every cognitive test can have a projective aspect if the examiner is willing to analyze the client's response, and every projective test can have a cognitive aspect if the examiner is willing to see the organizing tendencies behind how the client integrates unstructured information.

It needs to be remembered that tests are simply formalized and standardized observations, and useful instruments provide information about both the basic skills and organizing principles that patients bring to understanding and dealing with their life. These principles can be described at neurobiological and/or cognitive levels and as they affect learning and functioning in school, as well as interacting with others.

Measurements in Testing

Although many techniques might be used to help understand a child client and his or her referral question, the emphasis here is on standardized

testing (based on regularized procedures as well as normative data). The point of testing is to create a way of comparing one individual to a population of such individuals, as well as the strengths and weaknesses within one individual. The psychologist will select an instrument that is valid (it measures what is intended) and reliable (it measures it consistently). The testing involves establishing a basal level (the level at which all items are passed) and a ceiling (the level at which no items are passed). The testing process involves converting a raw score to a standard score that can be compared to other scores along what is thought of as a normal distribution with predictable statistical properties. The standard deviation (SD) is a measure of dispersion around the mean; the further scores are from the mean and each other in terms of the SD, the more meaningful is the discrepancy. It is accepted that a measurement is an approximation and not exact. This approximation is recognized by the concept of standard error of measurement (SEM), which is the naturally occurring (random) error that takes place in the real world as one attempts to measure anything. The fact that measurements are not exact is also recognized by the concepts of confidence intervals (the probability that the true score falls within a range of scores) and statistical significance (the probability of finding a result by chance). Test scores are accompanied by confidence intervals; for example, a score of 103 might have a 95 percent chance of falling between 94 and 112. Test discrepancies can be calculated at different confidence levels; for example, a discrepancy might be selected because it has a 1 chance in 100 (0.01) rather than 5 chances in 100 (0.05) of occurring at random.

There are different ways of describing a child's placement in large populations, the most common of which are standard scores and percentiles. It should be noted that there are different kinds of scores and values within the normal distribution, and that they differ by mean and dispersion. (See Fig. 7.4–6 which depicts Z, T, and IQ scores, as well as rounded percentiles under the "bell" curve.) While standard scores (such as those with a mean of 100) can be informative, parents often respond well to the concept of a percentile. For example, a student who falls at the 68th percentile performs as well as or better than 68 percent of the children tested. In addition to percentiles and standard scores, many cognitive tests provide age equivalents or grade equivalents. These measures suggest how the child compares to others of his or her age or grade. Age scores note years and months (e.g., 4–6, for "4 years, 6 months"), whereas grade scores note basically monthly progress within the academic year (e.g., 5.2, for "fifth-grade, 2 months," or 2/10 into the school year).

Beyond Scores and Tests

It should be noted that the testing process involves more than scores. Although scores are important, how the patient goes about solving

cognitive problems also is carefully observed. The examiner is interested not only in test performance, but also in the client's reaction. It is important to the psychologist to note how the client arrives at right as well as wrong answers and to explore the client's cognitive strategies on tasks. In general, it is important to note whether the client responds in a deliberate or impulsive way. It also makes a difference to the examiner whether the child says "This is easy" as he or she fails or laments "Oh this is hard" as he or she succeeds. Both have to been seen and give the examiner a very different sense of how the client will experience school and life.

In addition, the testing process is not separate from the therapeutic process. Well-handled testing can become an extension of the treatment. Feedback about results and their relationship to the presenting problems can be provided as the evaluation unfolds. Often this discussion is among the first opportunities the parents have had to understand the patient's predicament in a sympathetic way and to begin a constructive process. It can be helpful for the family to comprehend the patient's difficulties as a natural outgrowth of problems with understanding and learning, so long as practical interventions and solutions are provided.

COGNITIVE AND NEUROPSYCHOLOGICAL ASSESSMENT

The present discussion describes both general cognitive assessment and focused neuropsychological assessment. Because the purposes of the evaluations tend to differ, so do the instruments and the way in which they are used. The general cognitive assessment tends to be a descriptive and practical event with an eye to the policies and possibilities in the outside world. As a result, the cognitive tests tend to be "comprehensive" instruments. Here they are presented according to the type of instrument administered: intellectual, achievement, and processing. The very factors that make them useful for general assessment limit them when it comes to fine tuning understanding of neurobiological functioning. The neuropsychological instruments tend to be more "precision" tests that attempt to assess very specific behaviors that represent neural constructs in an inner world. Here the instruments are organized according to the domains assessed: Memory, attention, executive functioning, and language, as well as visuoperceptual and sensory/motor functioning. Even when the results are explored at more descriptive levels, functioning within the domains is not seen as separate or independent of their neurobiological underpinnings.

Although the focus here is on the assessment of cognitive functioning and neuropsychological domains, depending on referral issues, the evaluation might also address behavioral, social, and emotional issues using a variety of techniques. In general, psychological assessments/reports might be identified based on the realm of concern—for example, psychoeducational assessment (diagnostic issues involving

academic difficulties) or psychodiagnostic evaluation (referral issues independent of school problems), or based on the professional undertaking the assessment (such as a neuropsychological assessment undertaken by a neuropsychologist). Psychological results might be approached from a descriptive or diagnostic point of view; that is, the clinician might simply try to convey how the patient is functioning and why, or might additionally try to characterize the patient's functioning in categorical terms. The terms might be legal, as in the case of the Individuals with Disabilities Education Act (IDEA) in relation to school systems, or clinical, as in the case of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) in mental health settings. In any event, psychological assessment of all kinds can add texture and detail to the understanding of a patient's referral complaints, especially around the possible effect of processing deficits. Specific information about processing problems (how the child perceives his or her experience and monitors his or her behavior, and therefore responds to his or her situation) can help adults to know better how to interact with the child, as well as plan appropriate educational and interpersonal interventions. Basic information about overall cognitive functioning (based on intellectual, achievement, and processing instruments), as well as neuropsychological functioning (based on assessment of the functional/neurobiological domains), as they are integrated with other assessment data addressing referral issues can help to clarify diagnosis, which in turn can help to clarify issues around medical, school, behavioral, and therapeutic possibilities.

In the context of a focus on general cognitive and neuropsychological assessment, what is of interest is that many of the same processing and learning factors that can affect academic functioning can also affect social/emotional functioning. These factors can include everything from selective attention, memory function, sequential processing, and planning ability to fluid reasoning and self-regulation. These factors can affect the ability to assimilate, track, interpret, and organize experience in school, with others, and in life.

In general psychological practice, the kind of cognitive assessment described here is particularly suited for the evaluation of learning issues as they affect school functioning. In the context of this section, social/emotional issues are seen from a "cognitive point of view" and how processing problems can affect social/emotional functioning. However, cognitive tests in the context of a full psychological evaluation would shed good light on a range of referral issues and provide a rich source of information about the patient's functioning. All the same, there are times when what is needed is a neuropsychological assessment (which uses diverse assessment and testing techniques) provided by a clinical neuropsychologist with advanced training and expertise in neuropsychology. This is true when a known neurological condition, such

as seizure disorder or head injury, is involved. This can also be true when diagnostic issues become particularly complex. Neuropsychological testing might be indicated or helpful when the identification of known neuropsychological circuits might contribute to diagnostic clarity; a systematic assessment of a functional hierarchy might allow for a complex skill to be broken down into its component parts; specific neuropsychological tests might help to clarify strengths and weaknesses in cases at the higher and lower intellectual ranges; or when serial testing is desired while reducing the possibility of practice effect.

In any case, basic neuropsychological concepts have suffused psychological testing, with commonly administered instruments allowing for the assessment of specific processes such as working memory or general constructs such as executive functioning (goal-directed or planned behavior). Furthermore, once the quantitative data are gathered, they mean nothing without a qualitative interpretation, as well as a conceptualization and presentation that helps to make sense of the child client to his or her caretakers.

It should be noted that tests are constantly updated. This is often undertaken to correct perceived shortcomings, to adjust for cohort changes, and to make instruments more relevant. Because of this, instruments are referred to in "generic form" in this text, without any notation that would identify the edition. However, in actual practice, it is important to know that the clinician has used the most up-to-date version of the test, unless there is a good reason not to. For example, some clinicians with specialized practices continue to use old editions (or subtests) of instruments because they find that particular instruments or tasks best meet the needs of their particular population or best answer a type of diagnostic question. (American Psychological Association guidelines require that basically the tests used to make interpretations or recommendations are not outdated or obsolete for the "current purpose.") Clinicians must select from among the available tests the one that seems to best meet the needs of any particular client. Tables 36.3–1 and 36.3–2 note the editions of tests that are current as the present text goes to press.

GENERAL COGNITIVE ASSESSMENT

Cognitive and neuropsychological assessment can be looked at from different perspectives as clinicians attempt to understand referral questions and individual children. Generally, cognitive testing can be addressed in terms of psychoeducational and psychodiagnostic evaluation. However, the combination of intellectual, achievement, and processing testing has commonly been given as part of psychoeducational assessment and to determine the cause of learning problems, as well as, at least in the past, to establish eligibility for services with public schools. The processing instruments may have clinical, as well as educational, applications. As do a

range of psychologists, neuropsychologists commonly administer the intellectual and achievement tests described. They might give some of the processing tests; however, in their exploration of brain–behavior relationships, they usually use more precise instruments.

Description of Cognitive Tests

Although psychologists use diverse tests, three kinds of cognitive testing are described in this subsection: intellectual, achievement, and processing instruments. Whereas some tests are labeled "processing" instruments, all of these tests can be seen in terms of a simple (simplistic) input–output model, with processing mediating the outcome. With various elaborations (the visual input might be text) or exceptions (the task might involve movement), most standardized tests involve the assessment of auditory or visual input and verbal or motor output. Generally, intellectual tests are thought of as measures of overall mental ability and processing tests are thought of as measures of discrete cognitive functions. However, they overlap in that they make varying demands on learning based on past exposure as it interacts with current cognitive strategies in situations that vary in terms of familiarity and novelty. Both kinds of tests can be seen in "processing" terms in that patterns of responding can emerge from the data that reflect not only on school but interpersonal and adaptive functioning. Achievement tests systematically assess past academic learning. Clinicians understand the results are affected by a wide variety of educational opportunities and impacted by a large variety of co-morbid conditions. Once disadvantageous conditions such as sensory disorder and intellectual disability, and, increasingly, lack of appropriate/adequate instruction have been ruled out as primary causes of academic difficulties, academic shortfall has commonly been attributed to processing problems. These can include everything from discrete cognitive functions such as phonological processing to more general mental constructs such as executive functioning. In fact, good basic processing, the ability to successfully use a wide variety of cognitive strategies in a wide variety of tasks, has commonly been thought to support intellectual level and academic success.

When it comes to learning problems, educational research has suggested just how complicated the identification of learning difficulties is, while recognizing the need to define them in order to provide services to children, particularly in school. Descriptions and definitions of learning deficits have recognized disparities in different aspects of functioning, and tried to parse out factors intrinsic to the child and specific to learning from aspects of their environment and experiences. Procedural definitions of learning disorders under DSM in the medical community and of learning disabilities under IDEA in the public schools have tried to grapple with this complexity. As a practical matter, both have generally focused on the concept of "unexpected underachievement."

When it comes to learning disability, what is at stake is the opportunity for a patient to obtain services through school at no cost to the family. Otherwise it falls to the family to obtain private educational therapy in the community. In schools, students might obtain targeted classroom instruction in areas of weakness, or some intervention through a student study team, or accommodations or modifications in class through informal 504 procedures. However, students are apt to receive more extensive and expensive services under IDEA, through more formal IEP (Individualized Education Program) procedures that identify and make recommendations around disabilities, including specific learning disability. (See Table 36.3–3 on education and law.)

Table 36.3–1. Cognitive Tests

Test	Age Range	Description
Intellectual tests		
Omnibus intellectual and ability tests		
Wechsler Preschool and Primary Scale of Intelligence–Fourth Edition (WPPSI-IV) Wechsler Intelligence Scale for Children–Fifth Edition (WISC-V) Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV) Stanford–Binet Intelligence Scales–Fifth Edition (SB5) Kaufman Assessment Battery for Children–Second Edition (KABC-II)	2 years, 6 months–7 years, 7 months 6 years–16 years, 11 months 16 years–90 years, 11 months 2–85+ years 3–18 years	The tests consist of several subtests that can be combined in different ways to understand basic intellectual functioning. The Wechsler tests yield a full-scale intelligence quotient (IQ) score, as well as overall scores within the verbal and nonverbal areas. There are ways of grouping subtests of the Wechsler instruments to distinguish between more verbal versus perceptual reasoning, as well as to measure working memory and processing speed. Because of its age range, the WISC is probably the most familiar of the Wechsler tests to child clinicians. The most recent WISC provides a full scale IQ, and five primary index scales in verbal comprehension, visual spatial, fluid reasoning, working memory, and processing speed. It provides additional ancillary and complementary index scales. The SB also provides full scale as well as verbal and nonverbal IQ scores. In addition, it provides factor indices in fluid reasoning, knowledge, quantitative reasoning, visual–spatial processing, and working memory. The KABC scales and their subtests (which minimize verbal instructions and responses) include measures of sequential and simultaneous processing, fluid reasoning and crystallized ability, and long-term retrieval. Many, if not most, cognitive tests (including intellectual, achievement, and processing) are structured in the same way (mean is 100, standard deviation is 15), and so results can be compared across instruments. This is particularly important in identifying

		learning problems.
Infant and child tests		
Gesell Developmental Observation-Revised (GDO-R) Bayley Scales of Infant and Toddler Development–Third Edition (Bayley III) Mullen Scales of Early Learning (MSEL)	2.5–9 years 1–42 months Birth–68 months	The Gesell is an "old" test that is still administered by medical professionals, as well as psychologists. Along with the Bayley III, the GDO-R is as useful for its opportunities for observations as it is for its scores. Both can be helpful in the identification of children at risk for developmental delay. The MSEL, with its five scales (Gross Motor, Visual Reception, Fine Motor, Expressive, and Receptive Language), is often valued in assessing readiness for school, as well as in identifying specific interventions for children with developmental disabilities.
Nonverbal and language-free tests		
Leiter International Performance Scale–Third Edition (Leiter-3) Test of Nonverbal Intelligence–Fourth Edition (TONI-4) Comprehensive Test of Nonverbal Intelligence, Second Edition (CTONI-2) Wechsler Nonverbal Scale of Ability (WNV) Raven's Progressive Matrices	3–75+ years 6 years–89 years, 11 months 6 years–89 years, 11 months 4 years–21 years, 11 months 5–17+ years	Although the nonverbal tests are often seen as language or culture free, strictly speaking, this is not always true because of representational content or language demands. However, some nonverbal tests (such as the TONI-4) depend more on abstract pattern recognition and can make fewer demands on language systems. The WNV provides batteries for younger (4–7:11) and older clients (8–21:11), and claims to be a general ability measure using nonverbal tests. Tests like Raven's Progressive Matrices (which actually come in several forms) are based on pattern recognition and thought of as "culturally reduced," if not unbiased or unloaded. Nonverbal tests can be useful in assessing individuals whose experience falls outside the "cultural norm."
Quick tests		
Kaufman Brief Intelligence Test–Second Edition (KBIT-2) Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II)	4–90 years 6 years–90 years, 11 months	Although the KBIT and WASI include verbal and nonverbal scores, they do not include as much information as more comprehensive intellectual tests. However, they provide an estimate of intellectual level in a fraction of the time.
Achievement tests		
Wechsler Individual Achievement Test–Third Edition (WIAT-III) Woodcock-Johnson Tests of Achievement-Fourth Edition (WJ-IV ACH) Kaufman Test of Educational Achievement–Third Edition (KTEA-3) Wide Range Achievement Test–Fourth Edition (WRAT4) and WRAT-Expanded (WRAT-E)	4 years–50 years, 11 months 2–80+ years 4 years–25 years, 11 months 5–94 years 4–24 years (Indiv.)	The WRAT has traditionally been seen as a screening instrument. The WRAT4 includes subtests on sentence comprehension, word reading, spelling, and math computation. An expanded version (WRAT-E) provides achievement and intellectual information (nonverbal reasoning). The KTEA assesses reading (decoding and comprehension), mathematics, spelling, and written expression. The WIAT and WJ-ACH also systematically address the basic areas that have been identified as relevant for learning disability.

Reading tests		
Gray-Oral Reading Tests–Fifth Edition (GORT-5) Woodcock Reading Mastery Tests–Third Edition (WRMT-III) Comprehensive Test of Phonological Processing-Second Edition (CTOPP-2)	6 years–23 years, 11 months 4 years 6 months–79 years, 11 months 4 years–24 years, 11 months	The WRMT (which includes a measure of listening comprehension) and GORT (which includes a measure of fluency) involve systematic approaches to assessing different aspects of reading skills. Because of the importance of phonological awareness in the development of reading skills, tests of phonological processing (such as the CTOPP) are often included in assessments in which reading problems are identified.
Processing tests		
General processing tests		
Woodcock-Johnson–Fourth Edition Tests of Cognitive Abilities (WJ-IV COG) Differential Ability Scales-Second Edition (DAS-II)	2–80+ years 2 years, 6 months–17 years, 11 months	The WJ-COG scores provide information about intellectual ability, as well as a variety of cognitive and clinical areas. The WJ-COG is co-normed with the WJ-ACH, and the new test of oral language (WJ-OL). The DAS provides a measure of general conceptual ability and a profile of cognitive strengths and weaknesses.
Specialized tests		
A Developmental Neuropsychological Assessment-Second Edition (NEPSY-II)	3 years–16 years, 11 months	Scored across six domains, the NEPSY provides scores in executive functioning/attention, language, sensorimotor, visuospatial, and memory/learning, and social perception.
Delis-Kaplan Executive Function System (D-KEFS)	8 years–89 years, 11 months	As a test of executive functions, nine stand-alone subtests assess cognitive flexibility, response inhibition, problem solving, and concept formation.
Visual–motor tests		
Bender Visual–Motor Gestalt Test–Second Edition (Bender-Gestalt II) Beery Developmental Test of Visual–Motor Integration–Sixth Edition, Revised (VMI-6) Test of Visual Perceptual Skills–Third Edition (TVPS-3)	4–85+ years 2 years–99 years, 11 months 4 years–18 years, 11 months	Both the Bender and VMI involve copying geometric figures. Because the Bender allows the student to organize the items on the page and the VMI asks the student to copy each figure in its own space, the two can be used together to assess organizational issues, as well as visual–motor integration. Both the current versions of the Bender and VMI provide ways of assessing perceptual separate from motor skills. The TVPS is motor free and involves making judgments about visual information—for example, a client might be asked to identify an item from a fragmented presentation.
Auditory–vocal tests		
Test of Auditory Processing Skills–Third Edition (TAPS-3) Test of Auditory Reasoning and Processing Skills (TARPS)	4 years–18 years, 11 months 5–14 years	The TAPS involves making judgments about auditory information. For example, the student might be asked to discriminate sounds or remember words or numbers under different constraints. The TARPS measures the “quality” and “quantity” of the client’s auditory thinking and reasoning.

Memory tests		
Wide Range Assessment of Memory and Learning–Second Edition (WRAML2)	5–90 years	Memory scales attempt systematically to assess memory skills within different spheres. The WRAML core battery consists of Verbal, Visual, and Attention/Concentration subtests. The CMS allows for the assessment of attention and working memory, verbal and visual memory, short-delay and long-delay memory, recall and recognition, and learning characteristics.
Children's Memory Scale (CMS)	5–16 years	
Social cognition		
Test of Problem Solving	6–12 years	Elementary and adolescent clients are asked to answer questions about pictures or scenarios to reveal social reasoning.
TOPS-3 Elementary	12–17 years	
TOPS-2 Adolescent	6 years–18 years, 11 months	
Test of Pragmatic Language-Second Edition (TOPL-2)		A test of language used to achieve goals, with information about six areas (setting, audience, topic, purpose, cues, and abstraction).

Publishers are constantly updating their tests, and these editions will be superseded by newer and better tests. Even now, there may be variations of some instruments that are supplemented, expanded, or newly normed.

Over the years, the basic definition of learning disability as "unexpected underachievement" has remained, but with new understanding and meaning. In this context it should be emphasized that there are two different systems for identifying learning problems: currently Specific Learning Disabilities in school settings and Specific Learning Disorders in clinic settings. For many years, learning disabilities and disorders were essentially interchangeable because of their focus on a severe or substantial (significant) discrepancy (typically one standard deviation or more) between ability and achievement. However, educational and medical definitions of learning deficits under IDEA and DSM, respectively, have changed.

In schools, current IDEA 2004 language forbids states to present a severe discrepancy model to local educational agencies as the only option, and indicates they must allow for other possibilities. Underachievement may be identified in several areas (focusing on reading, mathematics, and written expression). Basically, this can be based on failure to respond to research-based intervention as other students have, or the identification of a pattern of weaknesses in the context of strengths that indicates specific learning disability. Still, under this system, a severe discrepancy between ability and achievement might be identified as evidence of a pattern of strengths and weaknesses indicating the educational diagnosis. However, it can be seen there are potentially several ways of demonstrating learning disability in schools.

In clinics, under DSM-IV (1994) the discrepancy model remained essentially intact until DSM-5 in 2013, when the definition for learning disorder shifted and dropped the discrepancy between achievement and

ability. The new definition focuses on absolute academic deficit, and stipulates two basic requirements. The first requirement is that the learning difficulties have persisted for at least 6 months, "despite the provision of interventions that target those difficulties." The second requirement is that academic skills (in reading, written expression or mathematics) should be "substantially and quantifiably below those expected for the individual's chronological age," and cause significant interference with related activities, "as confirmed by individually administered standard achievement measures and comprehensive clinical assessment." The new definition also requires two additional factors: that the learning difficulties begin in the school-age years although they might not become manifest until later; and that the learning difficulties are not better explained by other issues, such as intellectual disability, other disorders, or inadequate instruction. The exact deficit level associated with the learning disorder might be left to the clinician's judgment, but the threshold is thought to be between 1 to 2.5 SD below the population mean (100), depending on the circumstances of the case. A 1 SD standard would fall at or below 85 (16th percentile); this might be sufficient for diagnosis "with converging evidence." A 1.5 SD standard would fall at or below 78 (7th percentile); this level is thought to provide "more diagnostic certainty." Specific learning disorder is an umbrella term, but impairments can be specified in the three academic areas, and can involve efficiency as well as accuracy. Severity of the disorder can be specified from mild to severe. The clinical standard for learning disorder might be seen as stricter (if not clearer) than the school standard for learning disability.

Table 36.3–2.
Neuropsychological Test List (Selected)

Memory and Learning

Verbal

- California Verbal Learning Test for Children (CVLT-C)
- Children's Auditory Verbal Learning Test
- Children's Memory Scale (CMS): Stories, Word Pairs, Word Lists
- NEPSY: Memory for Names, Narrative Memory, List Learning

Wide Range Assessment of Memory and Learning (WRAML): Story Memory, Sentence Memory, Verbal Learning

Nonverbal/Visual

- Benton Visual Recognition Test
- CMS: Dot Locations, Faces, Family Pictures
- NEPSY: Memory for Faces, Memory for Designs
- Rey-Osterrieth Complex Figure, delayed conditions
- WRAML: Design Memory Picture Memory, Visual Learning

Attention

Sustained Attention/Vigilance

Continuous Performance Tests (CPT): AX Version of the CPT; Connors Continuous Performance Test

(CCPT), Tests of Variables of Attention (TOVA)
NEPSY: Auditory Attention and Response Set, Visual Attention
Paced Auditory Serial Addition Test (PASAT)
Working Memory
Auditory Consonant Trigrams (ACT)
CMS: Numbers, Sequences
Wechsler Intelligence Scale for Children (WISC): Digit Span, Letter-Number Sequencing, Arithmetic, Spatial Span
WRAML: Finger Windows, Number/Letter Memory, Sentence Memory
Inhibition/Impulse Control
Delis-Kaplan Executive Functions System (D-KEFS) Color Word Interference (Inhibition)
Go/No Go Test
NEPSY: Auditory Attention and Response Set, Inhibition (Inhibition), Knock and Tap
Executive Functioning
Organization/problem solving/planning
Children's Category Test
D-KEFS: Sorting, Twenty Questions, Tower
NEPSY: Animal Sorting, Clocks
Rey-Osterrieth Complex Figure, Copy condition
Cognitive flexibility/Set shifting
Children's Color Trails
D-KEFS: Trails (Number-Letter Switching), Verbal Fluency (Category Switching), Design Fluency (Switching), Color-Word Interference (Inhibition/Switching)
Matching Familiar Figures Test
NEPSY: Inhibition (Shifting)
Trail-making Test A and B
Wisconsin Card Sorting Test
Fluency
Controlled Word Association Test
D-KEFS: Verbal Fluency, Design Fluency
NEPSY: Word Generation, Design Fluency
Language
Expressive
Boston Naming Test
Clinical Evaluation of Language Fundamentals (CELF): Word Structure, Recalling Sentences, Formulated Sentences, Word Classes-Expressive, Expressive Vocabulary, Sentence Assembly
Expressive One Word Picture Vocabulary Test
NEPSY: Sentence Repetition, Speeded Naming
Test of Language Competence (TOLC): Ambiguous Sentences, Oral Expression
Woodcock Johnson (WJ): Picture Vocabulary, Rapid Picture Naming
Receptive
CELF: Sentence Structure, Concepts and Following Directions, Word Classes-Receptive, Semantic Relationships, Understanding Spoken Paragraphs

NEPSY: Comprehension of Instructions
Peabody Picture Vocabulary Test
Token Test
TOLC: Listening Comprehension, Figurative Language
WJ: Story Memory, Understanding Directions, Oral Comprehension, Auditory Attention

Visuoperceptual/Visuomotor Functioning

Visuoperceptual

- Benton Facial Recognition
- Benton Judgment of Line Orientation
- Hooper Visual Organization Test
- Motor-Free Test of Visual Perception
- NEPSY: Arrows, Geometric Designs, Picture Puzzles, Route Finding
- Test of Visual Perceptual Skills

Wide Range Assessment of Visual–Motor Abilities (WRAVMA): Matching

Visuomotor functioning

- Beery Developmental Test of Visual–Motor Integration
- NEPSY: Block Construction, Design Copying
- Rey-Osterrieth Complex Figure, Copy
- WRAVMA: Drawing
- WISC: Block Design

Sensorimotor Functions

Sensory perceptual

Dean-Woodcock Sensory Motor Battery (DWSMB) (All sensory subtests)

- NEPSY: Finger Discrimination
- Reitan-Klove Sensory Perceptual Exam

Motor

- DWSMB: All motor subtests
- Grip Strength (Hand Dynamometer)
- Grooved Pegboard

NEPSY: Fingertip Tapping, Imitating Hand Positions, Manual Motor Sequences, Visuomotor Precision

- WRAVMA: Pegboard

The individual intellectual, achievement, and processing tests described in this section have been staples in assessing learning problems in the private as well as public sector over the years. However, their uses may be somewhat different in the present than in the past (and in the future) due to new educational and psychological research, and new legal and medical definitions. Despite recent changes in how specific learning disabilities and disorders are identified, the long legacy of these definitions of learning problems (and learning disability in particular) has impacted the way basic intellectual and academic tests have been developed over the years.

Table 36.3–3. Education and the Law: Students with Disabilities

Students with Disabilities

The federal government has protected the rights of students with disabilities since the 1970s under a series of public laws. In 1990, the 1975 Education for All Handicapped Children Act was renamed the Individuals with Disabilities Education Act (IDEA). Reauthorized in 1997 and in 2004, IDEA establishes the obligation of public school systems to provide a free appropriate public education (FAPE) in the least restrictive environment (LRE) to students with disabilities who “by reason thereof” (because of them) need special education and related services. What constitutes a free appropriate public education is determined on a case-by-case basis by Individualized Education Program (IEP) teams

and, if necessary, by administrative hearing officers and courts. The "least restrictive environment" refers to the right of students to be educated to the maximum extent appropriate in typical settings and alongside nondisabled peers. The issue of eligibility is established through a set of procedures established under IDEA, which involves an evaluation process and a team meeting, which culminates (if the student is found eligible as a student with a disability) in a written IEP. This series of events is sometimes referred to informally as the "IEP" process/team. A number of disabling conditions are specifically identified in IDEA law/regulations, including (in addition to sensory and orthopedic impairments) intellectual disability, speech or language impairments, serious emotional disturbance, autism, traumatic brain injury, other health impairments, and specific learning disability.

In general, three conditions need to be met for a student to be found eligible for special educational services: (1) The student needs to be found to have one of the several disabilities listed in IDEA, (2) the disability must adversely affect educational performance, and (3) the student must require special educational services, that is, "specially designed instruction." At this point, it should be emphasized that assessment is not identical with testing, and that in establishing eligibility for special educational services, schools "must" use a "variety of assessment tools and strategies to gather relevant, functional, developmental and academic information" about the student, including information provided by the parent(s). The evaluation process involves multiple forms of assessment of the disabilities for which the student might need/receive special educational services. However, the assessment process is usually focused on the issues of concern to the parents and school, and parents have an opportunity to approve the procedures to be used, subject to due process rights. School assessment usually involves some kind of testing, but other sources of evaluation (history, observations, ratings) are also important. The purpose of the assessments is to define not only academic but also functional and developmental needs that result from the child's disability. Not only must the school's assessment for disability employ a variety of techniques, but such assessments may be provided for reasons other than academic failure, and schools may not deny services because a disability is not reflected in typical ways (such as failing courses or grade retention).

Specific Learning Disabilities

According to language provided in 1997 and retained in IDEA 2004, a specific learning disability is defined as a disorder in "one or more of the basic psychological processes involved in understanding or in using language, spoken or written," that "may manifest itself in imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations." There are circumstances under which a learning disability cannot be designated—for example, when the academic problem has primarily been due to another problematic condition such as intellectual disability or emotional disturbance.

In the past under IDEA requirements, learning disabilities eligibility has been determined by establishing a severe discrepancy between intellectual ability and academic achievement in one of seven academic areas. The academic shortfall was explained by a processing problem (such as sequential visual memory, attention deficits, or an auditory processing problem), which was established by inclusion or exclusion. There have been a number of changes, and the academic areas now number eight, with slight changes in terminology, but with major changes in procedures.

Prior to 2004, the DSM criteria for learning disorders in mental health centers closely paralleled the IDEA procedures for identifying learning disabilities in schools. Both emphasized the discrepancy between intellectual ability and academic achievement. Interestingly, what have been regarded as exclusionary factors in public schools have been regarded as aggravating factors in medical centers. In research, some have increasingly been understood as co-morbid conditions associated with learning difficulties. All the same, despite differences in meeting criteria between public schools and medical offices, the correspondence between these two systems (learning disorders under DSM and learning disabilities under IDEA) may have made communication between professionals outside and inside of school systems somewhat easier. However, changes to IDEA in 2004 made communication around learning problems more complicated. Then in 2013, the DSM-5 was modified in an attempt to reduce the inconsistency between medical and educational handling of learning disorders and disabilities, respectively. All the same, the situation has remained complex.

While DSM-IV still described learning disorder in terms of discrepancy between ability and achievement, IDEA 2004 changed the way students in public schools might be identified with

learning disability in specific academic areas. The changes under IDEA came about because of the experience of professionals in school systems. In schools the discrepancy model has increasingly been experienced as a "wait-to-fail" approach, because students have often fallen behind, sometimes as much as two years, before being found qualified for services. Instead, based on educational research, policy makers have come to value a "response-to-intervention" (RTI) model, which calls for services to be provided for the student before the student is identified as learning disabled, and as soon as screening indicates they are lagging behind. This approach involves ongoing and systematic assessment (on curriculum-based measures), and intervention occurs in a number of steps (through a series of tiers). This takes place first in the context of general education and curriculum, and only subsequently through more targeted, intensive and individualized instruction, but still within the general education class placement. This approach has been adopted in the context of a general commitment to teaching reading (and with the knowledge of the importance of early intervention), but interest and expertise have been growing in other curricular areas and at higher grade levels. Under federal regulations, this RTI approach has been presented not just as an instructional technique, but potentially as part of the process of determining if a student has a specific learning disability.

As of the reauthorization of IDEA in 2004 (which came into effect in 2005, with regulations following in 2006) the procedures for determining eligibility around learning disability have changed.

When it comes to establishing criteria for specific learning disability, (1) a state must not require a severe discrepancy between academic achievement and intellectual ability in specific academic areas, and (2) the state must permit the use of a process based on scientific, research-based intervention. The IDEA regulations open up a third way of identifying specific learning disability, based on "other research-based procedures." This last option is not well understood, and seems merely to increase the discretion of the IEP team. The chosen options must be adopted in a manner consistent with other regulations/guidelines, including the procedures described below.

According to the current regulatory language (which can change), there are three components to being identified as learning disabled. (Please see http://idea.ed.gov, Part B, as well as www.greatschools.org). First, there are inclusionary criteria: The student shows underachievement (based on age and/or grade) in one of eight academic areas (oral expression, listening comprehension, written expression, basic reading skill, reading fluency skills, reading comprehension, mathematical calculation, and mathematics problem solving) when provided with appropriate learning experience and instruction. The second component contains the procedures for evaluating eligibility. Under the new regulations, two basic factors come into play in considering learning disability eligibility: (1) insufficient progress to meet grade/age standards when using a process based on scientific, research-based intervention, and (2) an option that considers patterns of strengths and weaknesses in performance and achievement relevant to the identification of a specific learning disability using appropriate specific and general assessment. This reference allows for many ways of determining that the student is not meeting expectations according to age or grade standards or intellectual development. The third component involves exclusionary criteria, such that the learning deficit is not primarily the result of visual, hearing, or motor disability, intellectual disability, emotional disturbance, cultural factors, environmental or economic disadvantage, or limited English proficiency. Lack of appropriate instruction in math and reading is also considered in the evaluation process; students who are academically weak because of school failures should not be labeled as having a disability.

Just what the new possibilities mean and how states and school systems will respond are in flux, creating opportunities and challenges in helping families to meet the needs of students.

Impacts on Assessment

Given the ambiguities in several interacting requirement systems, it is difficult to provide definitive information about how professionals should proceed in diagnosing learning problems in order to meet the needs of patients in the schools as well as the community. However, a number of observations can be made that are relevant to interpreting and meeting guidelines for specific learning disabilities in schools and specific learning disorders in clinics, and communicating between the two settings.

Under IDEA regulations pertaining to the identification of specific learning disability that might allow for special educational services, the first identification option (above) is understood as a reference to response-to-intervention (RTI). The evaluation process might involve data about low academic performance as well as monitored progress. Models for this option are not defined by regulation, but are formulated at state and local levels. Whether further testing using norm-referenced and individual instruments is needed as part of the eligibility decision around special education is a matter of debate. The second identification option regarding strengths and weaknesses (S/W) might be seen as framing alternative research-based and severe-discrepancy procedures. In the case of a discrepancy model, it is possible, for example, that professionals will need to demonstrate the diagnostic significance of any discrepancy/weakness and will be encouraged to identify any strengths that might facilitate intervention. While testing is not necessarily implicated, the S/W option would likely involve establishing a pattern of academic strengths and weaknesses between and within tests. There is some debate regarding whether contrasts within a cognitive test (such as high verbal comprehension versus low perceptual reasoning) instead of between or within achievement tests (such as high math and low reading) might be relevant. There are a great many ambiguities and uncertainties about how these regulations might be understood and implemented at state and local levels, and systems often fail to provide clear descriptions of the criteria they employ.

Further, communication between schools and clinics has been complicated not only by previous changes in the definition of learning disabilities under IDEA in 2004, but also by recent changes in the definition of learning disorders under DSM in 2013. Shifts in definition in both schools and clinics will have a variety of impacts on assessment.

Under DSM-5, individuals can be identified with specific learning disorder with impairments in reading, written expression, or math if they show persistent deficits "despite the provision of interventions that target those difficulties," and achievement "substantially and quantifiably below those expected for the individual's chronological age." This definition is very different from the learning disorder designation under the earlier DSM-IV in clinics (which mirrored the schools' ability/achievement discrepancy model) as well as somewhat different from the current IDEA 2004 learning disability definition in schools (which might involve patterns of discrepancies between different areas of achievement and/or between ability and achievement, or which might involve failure to respond to intervention). Currently school settings might offer more options for identifying learning disability than clinic settings do in identifying learning disorders, which seem focused on absolute deficit when looking at achievement and age. The main complications in practice are created by the requirement that deficits persist for 6 months despite targeted interventions. This stipulation introduces a possible ambiguity in the diagnosis, in that it would be commonly be the obligation of public schools to provide services and interventions although the diagnosis of learning disorder is made separately in clinic facilities.

Since the introduction of DSM-5 in 2013, professionals have tried to reconcile it with IDEA 2004, and to gauge the impact of these changes on practice. Tannock (2013) notes that the diagnosis of specific learning disorders in clinics is currently not necessarily synonymous with the designation of specific learning disabilities in schools. She suggests that not all children with the educational diagnosis of specific learning disability would meet criteria for the clinical diagnosis of specific learning disorder, while it is likely that a youngster with a DSM-5 diagnosis of specific learning disorder would meet IDEA/2004 criteria for the educational definition of specific learning disability. She notes two practice shifts as a result of changing definitions around learning difficulties. Some shifts will take place because of the overarching category of specific learning disorder under DSM-5 (with impairments in three academic spheres). Other shifts will take place because of the abandonment of the IQ-Achievement discrepancy criterion, along with the omission of cognitive processing problems in the diagnostic criteria. She notes that in practice, psychologists may shift from what she calls "assessment for diagnosis" to "assessment for intervention." Tannock notes a third and related shift, which calls for evidence of symptoms over time and the use of a wider array of data that might be called upon to characterize low academic achievement.

In assessing learning problems and seeking services through school systems, clinicians need to be aware of the different vocabularies and criteria that school and clinics use for identifying learning problems. In this context, it should be noted it is always possible that if the student is not found eligible for

special educational services under specific learning disability, they might be found eligible under one of the other designations, for example other health impairment (which can cover attention-deficit/hyperactivity disorder), speech and language impairment (which can impact aspects of academic learning), or autism (where learning problems can be associated with the diagnosis). Further, if the child does not qualify for special educational services under IDEA, they may qualify for support under Section 504 of the Rehabilitation Act of 1973.

If professionals wish to receive services for their child patients through schools, they should be aware of their instructional environment and how this might be related to assessment, as well as current state and local policies and procedures around disability eligibility in general, and learning disability in particular. In the past under the traditional discrepancy model, clinicians were advised to identify a processing problem associated with academic shortfall, in case it was a local requirement for specific learning disability. In the present given the various possible models that might be relevant, it might be wise to make it clear how diagnoses are impediments to progress in instruction, or shortfall is part of a pattern of strengths and weaknesses suggesting specific learning disability. It should be noted that IDEA is subject to re-authorization. At that time, the definition of learning disabilities and the procedures for identifying them can change. In fact, given the rhetoric around the last re-authorization, there may be some debate this next time about basic educational, definitional, and procedural issues. Up-to-date information can be obtained from http://www.copaa.org. It should also be noted that regardless of the model being used, it is ultimately up to the discretion of the IEP process/team to make findings around eligibility based on disability, and to recommend special educational services and placements. Any outside evaluation can be submitted to the IEP team; it must consider the results, but is not obligated to follow the recommendations. Once the special educational evaluation is underway, the opinion of outside professionals regarding learning issues becomes one data point among many in helping the IEP team decide whether the student is learning disabled and in need of special educational services.

It can be seen that changes within the school in providing instruction and evaluating learning disability will impact the way professionals see cognitive assessment that might be relevant to both community and school interventions. Regardless of the changes taking place, learning disability will still be understood in terms of “unexpected underachievement,” and this knowledge will guide the exploration of learning problems.

Provision of Services

At the end of the evaluation process, a decision about eligibility and services is made by the team that writes the IEP. The services might include special educational placements (such as resource support or special class assignment). The provisions might include classroom accommodations (adjustments in the way a student is taught and/or tested) or curricular modifications (adjustments in what the student is expected to learn). The IEP can also include referral to “related services,” such as counseling or language therapy to make it possible for the student to benefit from special education, or promote access to general education (curricula or peers). Furthermore, the IEP might include funding of services through an agency in the community (a non-public agency) or placement in a private special education school (a non-public school) to achieve FAPE.

In addition to IDEA and the IEP, there are also Section 504 protections and procedures established under the Rehabilitation Act of 1973, which is a philosophical spin-off of civil rights legislation of 1964 focused on issues of access. The Section 504 provisions cover individuals with a disability that “substantially limits” their ability to learn or to engage in another “major life activity.”

Although there is potentially some overlap and variability, the Section 504 procedures are potentially less formal than the IDEA and IEP process. The IEP involves an administrative meeting and is regarded as the most typical avenue for receiving special educational services such as resource teacher help, special class placement, and specialized interventions. In contrast, the Section 504 assessment can be very descriptive, and the evaluation process is less complicated and the meeting less formal than the IEP evaluation. However, the Section 504 process is designed basically to parallel the IEP process in terms of requirements for evaluation and placement, as well as a system of safeguards, for school children. Although it may be unclear that the Section 504 plan needs to be written, it would be prudent to do so to keep expectations clear. The interventions specified tend to

be accommodations/modifications that might benefit the student, such as seating toward the front of the room, extra time on tests, or adjustments in homework. However, services can also be required, depending on what is necessary in order to meet the student's needs as adequately as the needs of students without disabilities are met. The Section 504 procedure is a way of intervening with students with learning differences or problems who do not necessarily meet the strict criteria for disability under IDEA and its related regulations. Although some parents and schools prefer the Section 504 procedures because they can streamline getting help to a student, clinicians often prefer the IEP process because it includes more procedural protections than Section 504, and students found to be disabled through the IEP process are almost always covered under Section 504. As a result, clinicians have incentives for evaluating students to determine whether they have a learning disability and to pursue interventions (special services in addition to accommodations/modifications) through the IEP process. Although Section 504 procedures are thought of as less formal and interventions as less inclusive, like IDEA, Section 504's implementing regulations guarantee a free appropriate public education in the least restrictive environment—education that is provided to the maximum extent appropriate with nondisabled students. Affirmative services (such as speech-language therapy or even special educational placement) are also technically available through Section 504 procedures. In requesting school services, clinicians will need to consider whether pursuing interventions through 504 or IEP procedures would be "better" for their client.

How to Proceed

Whatever the disability in question, the starting point in receiving services for any student with school problems is to request an evaluation in writing from his or her home school. If the parent or guardian does not feel the evaluation team's recommendations will meet their student's needs, they can appeal the findings through a fair hearing (and opportunity for mediation) process. If they disagree with the district's assessment and the IEP team's findings, they may obtain an independent evaluation (before or after filing for fair hearing), which the district is required to consider and for which the district may be required to pay under certain circumstances. Independent evaluators generally have a right to observe students at school, and are most persuasive if they do so. Independent evaluators should seek teacher and school staff input orally and in writing. Practitioners who want their clients to receive services through the public school system should know the current state and local procedures that might apply to their assessment, eligibility, and services.

This discussion about educational law was prepared with the help of Maureen Graves, Attorney at Law. A practitioner with a client for whom legal issues will probably be important in obtaining special education services through the school district should consult an attorney or advocate (perhaps identified through the Council of Parent Attorneys and Advocates website, http://www.copaa.net). This discussion in no way constitutes legal advice.

Under the influence of definitions of learning disability that have emphasized the discrepancy between ability and achievement, it will be seen that many intellectual and academic instruments described here have been developed (perhaps co-normed) so as to facilitate this comparison. The tests assessing cognitive processing are quite varied, offering different ways of identifying and conceptualizing processing problems. Combined in different ways, the instruments described here can be useful in psychodiagnostic as well as psychoeducational evaluations. Because the focus of this subsection is not on learning disabilities or disorders but on cognitive assessment, a range of intellectual, achievement, and processing instruments and their uses are reviewed.

While intellectual and processing tests/subtests can be interpreted at "face value" they can also be assessed for "cognitive patterns" (perhaps in connection with projective tests) that might be relevant for social as well as

academic functioning. This enterprise depends on the same clinical judgment and skill (informed by research) that professionals bring to interpreting behavior in the history or observing the patient in the office. As a result, these intellectual, achievement, and processing tests are useful in the assessment of learning problems (such as SLD) as well as other complaints associated with disruptions in cognitive functioning and integration, such as Autism Spectrum Disorder (ASD) or Attention Deficit/Hyperactivity Disorder (ADHD).

Description of the General Applications of Cognitive Testing

Intellectual, achievement, and processing assessments are important in understanding learning and functional problems from a clinician's point of view. The instruments can provide useful information in exploring a wide range of referral issues, including psychoeducational and psychodiagnostic assessment, as well as neuropsychological assessment. However, each psychologist will select those instruments that meet the particular needs in any particular case. A full range of cognitive instruments and their uses are described in this section.

Intellectual Testing. Because of the importance of intellectual functioning as it affects the ability to learn and adapt in social and academic situations, intellectual testing is often part of a variety of psychological assessment batteries, including psychoeducational and neuropsychological evaluation, along with more general developmental and clinical evaluations.

BACKGROUND. Intellectual testing is usually undertaken to determine the patient's general intellectual level of functioning. The intelligence quotient (IQ) is valued because of its stability over time in the general population, as well as for the protective aspect of intelligence in a range of potentially disabling situations. It should be noted that, although intellectual level can be measured, it is hard actually to define intelligence.

Definition. Different theorists have had their own take on intelligence. Although, theoretically, Sattler saw intelligence as "multifaceted and hierarchically organized," practically, there is no agreement about what intelligence is, and this is reflected in the different ways in which it is measured. However, a commonsense definition of intelligence would include the ability to learn from and adapt to the environment. It would also include the ability to think abstractly (to use symbols and higher-level concepts) and to recognize patterns in diverse situations (on tests and in life). Although intelligence tests yield one IQ score (or several IQ or index scores), they are, in fact, devices for "sampling" many tasks in a variety of verbal and nonverbal areas. Intelligence tests differ from achievement tests

in that they are more process than content oriented. That is, for the most part, they put the student in a position in which he or she has to solve verbal or nonverbal problems in the present rather than use specific reading, mathematics, or writing skills learned in the past.

Stability. Although there is some disagreement, IQ scores tend to be relatively stable starting as young as 5 to 7 years of age. In general, the older the child is when tested and the smaller the interval between test administrations, the greater is the correlation between two IQ scores. Although using an IQ score can be useful as a way of assessing the client's basic trajectory through life, the prudent practitioner must be aware that there are a number of factors that can affect intellectual functioning and, thus, IQ scores. Factors associated with disorder and illness can suppress scores, particularly in psychiatric practice. These can include situational factors, such as lack of motivation, as well as transient factors, including inattention, depression, and psychosis. Some ongoing conditions (e.g., those that involve problems with relating, involvement, or cognition) can negatively affect learning and performance and potentially affect intellectual levels as well. In some cases, treatments and interventions implemented to improve cognition and development can improve functioning and, thus, IQ scores.

Protection. Despite conceptual and practical complications, high intelligence is associated with better prognosis in a wide range of psychiatric conditions; lower rates for behavior, conduct, and emotional problems in children; and lower rates of referral for psychiatric problems in adults. In the case of any kind of brain damage (neuronal death), intellectual level accounts for a great deal of variance in predicting outcome, with lower IQs associated with poorer outcomes and higher IQs associated with better outcomes.

ASSESSMENT. Although IQ is what is obtained with an IQ test, there are a variety of intellectual tests, as well as other ways of calculating intellectual level. There are a number of instruments from which to choose, and the psychologist must make his or her selection based on the specific characteristics of each test (e.g., normative sample and construction of the instrument) as they relate to the characteristics of the client (e.g., age and referral question). Once the test has been administered, the clinician must make interpretations based on the analysis of overall and subtest scores and their pattern in the context of the diagnostic process.

Sometimes, there are misinterpretations of scores at the bottom of the intellectual range. Distinctions must be made between level designations and diagnostic categories. Over the years professionals may have seen results from testing that suggest findings in the "intellectually deficient" or "mentally retarded" range. Over this period, even the description of

functioning within the mentally retarded range has not necessarily meant the client was diagnosed as mentally retarded. The diagnosis depended on other features of the case, including developmental course and adaptive functioning. The clinical terrain has shifted yet again with the change in DSM-5 diagnoses and language. The designation mental retardation has been replaced with intellectual disability, with less emphasis on IQ score alone. However, while intellectual level (IQ score) has been removed from the diagnostic criteria, they are still referred to in the text description. The term intellectual disability involves impairments in mental abilities that impact adaptive functioning in three domains (conceptual, social, practical). Severity of the disability is determined by adaptive level not intellectual score, because it is those impairments that determine the level of supports needed. It is hoped that the diagnostic change will encourage more comprehensive and complete assessment. However, it is important that measurement of intellectual level still be part of any individual's assessment, along with developmental course and adaptive deficits. Under DSM-5, intellectual disability is associated with intellectual levels two SD or more below the population mean, or an IQ score of about 70 or below. While in the past some agencies or regulations have referred to mental retardation as a population served or referenced, since Rosa's Law (2010) this designation has been changed to intellectual disability. In making determinations about low intellectual level, the psychologist must look at the pattern of strengths and weaknesses not only within the instrument, but also within the overall case presentation. For example, some IQ scores can be brought down by specific language disorders or perceptual–motor difficulties. One of the many reasons for looking for alternatives to the discrepancy model around specific learning deficits is that IQ can be seen as an outcome measure of the very same processes that are associated with academic shortfall.

Although the emphasis with intellectual testing has often been on overall intellectual "level," actually, IQ tests are very good for assessing processing, and contrasts can be made between not only composite scores but also subtest scores. The psychologist is very sensitive to the pattern of responding, with an eye to identifying strengths as well as weaknesses. For example, the psychologist might look at not only auditory–verbal versus visual–motor functioning, but also the complexity of demand for comprehension and expression on verbal tasks or the nature of any integrative problems within the perceptual sphere. There are many different ways of looking at processing issues within intellectual tests. Perhaps in part because of the way intelligence was initially understood and intellectual tests were historically constructed (particularly the Wechsler), there may have been a tendency to contrast two spheres of functioning. This might have involved comparing verbal versus nonverbal tasks or crystallized versus fluid reasoning. Or it might have involved

identifying language-based in contrast to nonverbal learning disorders. While these traditional ways of understanding intellectual functioning continue to be valid and useful, the understanding of "intelligence" itself has evolved, as have the instruments designed to detect it. For example, there has been an ongoing interest in concepts like crystallized versus fluid reasoning or intelligence. Crystallized intelligence has been associated with previously-learned information and fluid reasoning with on-the-spot problem solving. Traditionally these cognitive modes have been associated with verbal versus nonverbal functioning. However, over time, as the understanding of both intelligence and processing have evolved so have IQ tests. The changes in the Stanford–Binet and the Wechsler tests reflect developments in the understanding of intelligence in general, and crystallized and fluid reasoning in particular. As an example, the SB5 and the WISC-V both include verbal items in fluid reasoning options (e.g., verbal analogies or mental arithmetic, respectively). As a further example, on the SB5, the knowledge factor can be regarded as a measure of crystallized intelligence, and includes both verbal (e.g., vocabulary) and nonverbal tasks (e.g., procedural information and picture absurdities). In general, as understanding of these and other intellectual concepts has shifted, so have the ways instruments are designed to characterize intellectual functioning. As publishers have revised their intellectual instruments, there has been a tendency to produce more elaborate scoring systems and reporting options. While there are still ways of comparing things like basic verbal and nonverbal functioning, or crystallized versus fluid intelligence, how this might be done on both the Wechsler and the Stanford–Binet has changed. The scoring developments have served to give the clinician more flexibility in exploring basic processing and intellectual issues. Both the SB5 and WISC-V provide a full scale IQ, and additional scoring. The SB5 provides five factor scores that are comprised of both verbal and nonverbal subtests. The instrument also provides overall verbal and nonverbal IQ scores. The WISC-V provides five primary index scales, as well as eight ancillary and complementary scales. On both instruments there are ways for the clinician to assess things like verbal and nonverbal functioning, fluid and quantitative reasoning, visual–spatial processing, and working memory, but how this might be done is different on the two tests. Psychologists are sensitive to the pattern of responding within and among tasks and spheres, and the meaning of that pattern of strengths and weaknesses for functioning. Intellectual tests can be supplemented by a variety of instruments to further characterize any patient's presentation, in order to contribute to their understanding and diagnosis, and to help identify appropriate interventions. Going forward, "traditional" concepts (often dichotomies) will continue to be useful, but with "contemporary" understanding.

Comprehensive Intellectual Tests. The two best-known intellectual tests are probably the Wechsler Intelligence Scales and the Stanford–Binet Intelligence Scales (SB). The current editions of both are divided into separate subtests, and the data are analyzed in separate spheres. Each test has an overall mean of 100 (50th percentile) and a standard deviation of 15. Subtests have a mean of 10 and a standard deviation of 3. In the case of the Wechsler tests, three separate instruments (Wechsler Preschool and Primary Scale of Intelligence [WPPSI], Wechsler Intelligence Scale for Children [WISC], and Wechsler Adult Intelligence Scale [WAIS]) are designed for three different age groups over the lifespan. In the case of the SB, one instrument covers a lifetime. All instruments include full-scale IQ scores, as well as (broadly speaking) verbal and nonverbal composite scores. Both instruments have made attempts to assist in decision making regarding attentional problems. The SB provides an extended score range (at the top and bottom of the scale) that may make it more useful for individuals within the "gifted" and "impaired" ranges. The WISC anticipates providing a technical report with extended norms to differentiate highly gifted children. Both the SB and the Wechsler tests have been the "standard" individual intellectual tests for the school-age population. They allow for the assessment of overall functioning and verbal and nonverbal reasoning, as well as of some processing. Both instruments have taken care to relate to achievement tests to facilitate the recognition of learning disabilities according to a "traditional discrepancy model" in which estimated ability is compared to academic achievement. Such comparisons are still relevant to the identification of learning disabilities, where a discrepancy between ability and achievement might be evidence of a pattern of strengths and weaknesses suggesting learning disability (where another might be a pattern of academic strengths in one area in contrast to academic weaknesses in another). The new definition of learning disorder makes intellectual testing less immediately relevant, in that the emphasis is on absolute deficits in academic subjects (although intellectual level and processing may help to understand academic differences or difficulties). Intellectual testing is commonly a part of psychological evaluation as it provides a way of understanding a patient's overall level of functioning in school and in life.

Other Intellectual Tests. Sometimes, specific characteristics of clients, such as age, suggest the use of specific instruments.

1. *Infant and child tests*. Although very young clients can be "tested," these administrations are often as useful for their observations as their scores. Because of the emphasis on sensorimotor functioning and development of language skills of very young clients, the instruments (such as the Bayley Scales of Infant and Toddler Development) are not as useful

for predicting later IQ scores as they are for identifying individuals who are, or are not, at risk. The instruments do not become predictive in a more specific way until children are of school age.

2. *Nonverbal and language-free tests.* Although some tests seem especially suited for individuals who have difficulty in hearing or have language issues (such as the Leiter International Performance Scale), the nonverbal tests can have broader applications. Their emphasis is on cognitive problem solving based on visual information. Naglieri suggests that nonverbal tests can measure the same constructs as broader intellectual tests. In fact, the Wechsler Nonverbal Scale of Ability claims to be a test of general ability using nonverbal tasks, not a measure of "nonverbal intelligence." (It includes an Object Assembly subtest for younger and a Picture Arrangement subtest for older clients.) Although these nonverbal tests are useful, it is important to remember that they may not provide the whole picture, and the significance of these tests for the client's ability and functioning must be carefully assessed.

3. *Quick tests.* The commonly used tests in research and clinical practice are the Kaufman Brief Intelligence Test (KBIT) and the Wechsler Abbreviated Scale of Intelligence (WASI). Both the KBIT and the WASI include verbal and nonverbal tasks to assess crystallized and fluid thinking. The WASI and KBIT2 have been able to provide evidence of good correlations with established Wechsler tests. The available data suggest that the quick tests, in fact, yield very representative overall IQ scores. Although a clinician might not get all of the processing information made available by a full intellectual test, these correlation coefficients suggest that the quick tests can provide a very good estimate of the client's general level of functioning.

General Comment. In discussing intellectual testing, it is important to note what IQ tests are and what they are not. Although IQ scores are commonly used as estimates of ability, they are not measures of "native" or "raw" intelligence. Every client brings a long learning history, including opportunities, to the testing situation. The debate over whether tests are culture free or culture fair has raged over the years. Although some instruments (such as those that emphasize nonverbal functioning) may seem less biased or loaded, it must be acknowledged that probably no intellectual test is either culture free or fair. However, because IQ tests have, in large part, been designed to predict school success, they can be useful in the context of assessing learning capacity. Nonetheless, because of the way in which tests have been constructed and normed, caution has to be exercised in making conclusions in any case involving an unusual cultural or medical situation. Although IQ scores are useful as estimates of ability, they are, in fact, measures of functioning. There are a variety of transient factors that can be associated with alterations in IQ scores,

including problems with involvement, attention, and disorganization. As a result, the clinician has to be prepared for changes in intellectual level based on changes in medical status. For example, in some cases, clinicians have observed increases in IQ scores subsequent to medical treatment for attention-deficit/hyperactivity disorder (ADHD), partly due to increased efficiency on the instrument itself and partly due to increased learning over time in the environment and in school. Further, clinicians need to be aware of what IQ scores do and do not measure. They do not measure social intelligence or "street smarts"—the kind of savvy large portions of the population use every day to keep jobs and raise families. Nor do they reflect the creativity that artists of different kinds bring to their art and that entertains and informs people daily.

Considerations. Here are a number of factors professionals should keep in mind when they deal with intellectual test results.

1. The intellectual instruments have developed classification systems according to intellectual level. The Wechsler system, based on scores with a mean of 100 and a standard deviation of 15, has become widely recognized, although the labels have recently been adjusted for the WISC-V. Average scores fall between 80 and 89 (Low Average), 90 and 109 (Average), and 110 and 119 (High Average). New labels for the WISC-V levels appear hereafter in parentheses. Those scores above the average range between 120 and 129 have been seen as Superior (Very high) and those ≥130 have been seen as Very superior (Extremely high). Those scores below the average range between 70 and 79 have been labeled as "Borderline" (Very low). The scores ≤69 remain Extremely low. The Stanford–Binet test uses the same classification system for low average to high average scores (80 to 119). However the classification system has a wider range including scores from 40 to 160, from Moderately impaired or delayed to Very gifted or highly advanced. While the Wechsler system has been widely applied to a variety of instruments with a mean of 100 and standard deviation of 15, each instrument is free to specify their own classification system. Recently there has been some discussion about providing levels with more "neutral" and less "laden" labels.
2. The more variability seen on intellectual testing, the less do composite or summary scores mean. For example, on the WISC or SB, when the verbal and nonverbal scores (or other contributing indices or composites) differ significantly, the full-scale IQ is not informative. Similarly, when subtest scores differ significantly within a sphere, the emphasis should be on specific processing issues rather than overall functioning.
3. Sometimes a practitioner will ask what a patient's IQ is. It is clear from the multiplicity of the tests and scores available that there is no one IQ. Although IQ is what the client gets on an IQ test, it is prudent to be well

aware of which test was administered and which IQ is being quoted. The practitioner should be aware of disparities on different tests or in different areas; this pattern is often important in making decisions for a client.

4. In the context of a psychoeducational assessment, when scores differ significantly on intellectual testing, the full-scale IQ may not be the best estimate of "ability," and the clinician will make attempts to choose among the possible summary scores on a range of instruments to identify the one that best represents the student's learning potential.
5. What can be most helpful about the IQ score is the way in which it helps the practitioner to be aware of the client's general level of functioning. Each intellectual test uses its own language to indicate relatively low or high functioning. In school or practice, the designations at the higher end of the scale might include mildly gifted (starting at 115), moderately gifted (at 130), highly gifted (at 145), exceptionally gifted (at 160), and profoundly gifted (at 175 or 180). However, because of a ceiling effect on most newer intellectual test scores, there is some question about the utility of these traditional ranges. Furthermore, there is really no agreement about how to characterize or segment giftedness. In the context of this discussion, it is worth noting that the "old" Stanford–Binet L-M is useful in detecting these very high IQ scores. However, the SB5 has provided and the WISC-V anticipates providing methods of detecting and refining high scores in the gifted range.

FURTHER ASSESSMENT. The following is a list of patterns that suggest the need for further evaluation.

1. *There are significant discrepancies between verbal and nonverbal scores*. When verbal scores are significantly lower than the nonverbal scores, the possibility of a language disorder should be considered. In the case of low verbal scores, a speech and language assessment, to include measures of receptive and expressive language, as well as pragmatics (the factors affecting successful communication), might be indicated. It is particularly important to identify language-based learning disorders because of their potential effect on school (and relating), and because there are some very specific treatments that might be helpful. When nonverbal scores are significantly lower than verbal scores, further assessment regarding a variety of possibilities must be considered. More than the verbal tasks, the performance tasks (which might be timed) are vulnerable to ongoing and transient factors mediated by attention. In the case of particularly large discrepancies, the possibility of nonverbal learning disorder should also be considered. This is a pattern of cognitive assets and deficits that can be associated with peer, as well as academic, problems. The identification of a nonverbal learning disorder can be very helpful because it can lead to a wide variety of interventions, including those

around social skills.

2. *The full or composite score is above the 98th or below the 2nd percentile (130 or 70).* At the intellectual extremes, the examiner may consider whether further assessment regarding educational or adaptive needs might be in order. When a client has a very high intellectual level (e.g., moderately to profoundly gifted), the professional might refer the family to a psychologist specializing in very bright children or consult the website of the Davidson Institute for Talent Development, http://www.davidsongifted.org, for information about how to meet the needs of his or her client. In the case of individuals with very low intellectual level, further assessment can be obtained through the school system or the local public agency serving developmentally delayed clients. Professionals and parents may access the websites http://www.aaidd.org, http://www.thearc.org, http://www.autism-society.org, and http://www.eparent.com for information about services and products associated with intellectual and developmental disabilities.

3. *The intellectual score shifts downward in a significant way.* Because it is anticipated that intellectual scores will remain relatively stable over time, the mental health professional is advised to look for an explanation, either psychological or medical, for dramatic score changes. This is a time when the psychologist should look at raw, as well as standard, scores; a drop in a raw score is the test equivalent of losing functioning in everyday life. After some exploration, consideration of medical issues, as well as consultation with further medical personnel such as the pediatrician or neurologist, or further psychological professionals, such as a neuropsychologist, might be in order.

Achievement Testing. When achievement testing is undertaken along with intellectual and processing testing, the overall evaluation is commonly referred to as a psychoeducational assessment. The purpose of the assessment is to identify learning problems and usually to rule out other psychological factors that might be complicating learning. In addition, its purpose has been to establish eligibility for services through the public school system under learning disability. Clinicians communicating with school systems will need to be aware of the state and local guidelines for establishing learning disability, and, for example, how/whether/when the kinds of standard testing described here might be relevant under new identification models. However, practitioners will still value standardized (individual) achievement testing as an objective (normative) view of how the child stands nationally and for the ability to identify academic strengths and weaknesses. Of course, other information will be important—for example, the child's grades, academic interventions, and behavioral information.

BACKGROUND. Achievement testing is administered to determine a student's level of functioning in basic academic areas, including reading, mathematics, and writing. The factors assessed are more transient and circumscribed than those assessed in intellectual testing. Unlike intellectual testing, achievement testing is not necessarily expected to be stable over time because it measures the child's success in formal learning and is highly dependent on the home environment and the school curriculum.

Learning disability is commonly defined in terms of "unexpected underachievement"—that is, the child has the potential and opportunities to have learned more. The term learning disability is described in IDEA, which was reauthorized in 2004, with changes coming into effect in 2005 and regulations following in 2006. Since then, states and schools have been determining how they will respond to new regulations and procedures, and the role of achievement testing in schools is being seen in new ways. Table 36.3–3 provides further discussion of the federal laws that affect access to special services at the local level for students with disability.

ASSESSMENT. Psychologists commonly begin an assessment of academic achievement by giving a comprehensive test in order to get a sense of a client's areas of weakness in contrast to their strengths in reading, math, and writing. These findings are compared to school records. Because reading problems are a relatively common reason for referral and their causes are increasingly understood, where indicated, this comprehensive test is commonly followed by other tests of reading skills that separate out accuracy, fluency, and comprehension.

It needs to be noted that while specific learning disabilities (in schools) and specific learning disorders (in clinics) are commonly thought of in terms of underachievement in a specific academic area, learning difficulties can be more diffuse. They might not impact a particular subject in a significant way, but might impact a range of subjects in a moderate way. Or they might impact overall school performance in a drastic way. DSM-IV provided a way of recognizing this. Not only could learning problems be identified as a Learning Disorder in reading, mathematics, or written expression, but as learning disorder NOS. This allowed for the identification of a learning disorder across academic spheres and, arguably, in processing terms. For example, Learning Disorder, NOS might have been associated with executive dysfunction or nonverbal learning disorder. Under DSM-5, Specific Learning Disorder might be identified with impairments in reading, written expression, or mathematics. As part of the reconceptualization and reorganization of the diagnostic manual around neurodevelopmental disorders, learning problems associated with processing problems impacting general adaptive and social functioning rather than subject areas might be identified as Other Specified (or

Unspecified) Neurodevelopmental Disorder.

Comprehensive Achievement Tests. Each of the tests assesses a range of academic areas, so that they can be compared to each other within any individual or to academic achievement according to more external standards (age/grade expectations). The WIAT and the WJ-ACH allow for the systematic assessment of reading (basic skills, fluency, and comprehension), mathematics (calculation, fluency, and reasoning), and writing (spelling and brief to extensive composition), as well as other academic spheres. Each test has its strengths and weaknesses, and, often, most of one and portions of the other are included in an educational assessment. Both instruments are relevant for identifying specific learning disabilities whether under the (old) discrepancy model or the (new) patterns of strengths or weaknesses. They are also relevant in identifying specific learning disorders in reading, written expression, or mathematics. The WIAT is a companion test to the WISC. The WJ-ACH is a companion instrument to the Woodcock–Johnson Tests of Cognitive Abilities (WJ-COG), which evaluates processing skills and includes measures of intellectual ability, and the Woodcock–Johnson Tests of Oral Language (WJ-OL) which assesses specific oral language functions. As individual achievement tests, both the WIAT and WJ-ACH are widely used in the community and schools. An additional comprehensive test is the Kaufman Test of Educational Achievement (KTEA), which is co-normed with the Kaufman Assessment Battery for Children (KABC).

Other Achievement Tests. These and additional achievement tests are also available and can involve general or specific assessment.

1. *General achievement tests.* Some instruments focus on screening, such as the basic Wide Range Achievement Test. However, the fourth edition allows for the assessment of sentence comprehension, as well as reading, spelling, and math calculations. An expanded version provides even more information. There are other tests that may assess a range of academic areas in more depth. Some instruments focus on one area of achievement, usually reading.

2. *Specific reading tests.* Because of the importance of reading in school and in life, particular effort has gone into systematically assessing reading skills and the ability to make symbolic associations between sights and sounds, as well as the ability to decode words and understand text. In addition to reading subtests on major achievement instruments, there are specific reading tests, such as the Gray Oral Reading Tests (GORT) and the Woodcock Reading Mastery Tests (WRMT). Some instruments facilitate the pinpointing of the level and nature of the reading problems—for example, by presenting not only common words, but also nonsense words to assess decoding ability. In addition, because of the importance of

phonological awareness in reading, tests of phonological strengths and weaknesses, such as the Comprehensive Test of Phonological Processing (CTOPP), are also often given. It is up to the psychologist to select the battery of instruments that best meets the needs of a particular student in any particular situation. It should be noted that, because achievement tests assess academic functioning in slightly different ways, psychologists often administer a combination of achievement tests to any one client. For example, reading comprehension might be assessed by asking a student to read a passage and supply a missing word or by asking a student to answer questions about a paragraph. The format of the test can influence how well a student does. How the student does in different contexts may help to explain the nature of his or her academic problems.

General Comment. Under the response-to-intervention (RTI) approach as it unfolds, whether standardized/individual achievement testing undertaken within or outside the school will be useful or relevant may depend on whether the tests can fit into assessment protocols (usually involving repeated testing) and whether the results can be directly related to instructional models within the school (which need to be research based). The response-to-intervention model involves data-based decision making and includes screening, benchmarking, and progress monitoring. The assessment procedures might be called "curriculum-based measurement" (CBM) and tend to look like classroom activities but are undertaken in formal ways. CBM commonly involves criterion-referenced measures in contrast to the kind of norm-referenced measures described in this section.

Considerations. Here are a number of issues that mental health professionals should keep in mind in assessing achievement results:

1. Although estimates vary greatly, it can be realistically to conservatively estimated that between 15 and 30 percent of children with attention-deficit/hyperactivity disorder (ADHD) have a learning disability, and between 20 and 30 percent of children with a learning disability also have ADHD. As a result, when a clinician finds one condition, he or she should look for the other. Although they can co-occur, ADHD and learning disabilities are separate conditions requiring separate and specific recommendations—one potentially medical treatment, and the other educational intervention.
2. The vast majority of individuals with a learning disability have problems with reading (between 70 and 85 percent). Students who have learning issues in other academic areas commonly also have reading problems. Isolated learning disabilities in mathematics are much less frequent (perhaps 10 to 15 percent).
3. There are a variety of reasons why there have been objections to using a discrepancy formula (between ability and achievement) in assessing

learning disabilities. An important one is that it has been common for the identification of students with learning disabilities to be delayed into mid-elementary school, when earlier identification might have allowed for more preventive or effective work. In addition, there are other issues. One is whether it makes sense to identify a child with superior intelligence but average achievement as learning disabled, even when there is a processing problem. Another is whether it makes sense to exclude slow learners who might need special supports. It is hoped that the response-to-intervention model will avoid or sidestep some of these issues.

4. Although achievement tests commonly note grade- and age-equivalent scores, as well as standard scores, for a variety of reasons, the standard scores are more reliable than the age or grade equivalents. The concept of grade equivalent can be misleading. For example, it does not mean that a student functioning at the 6.5 level while in the fourth grade could perform sixth-grade work. The student has not been exposed to the same content as sixth-graders, nor would the fourth-grade test have included sixth-grade content. The score merely indicates that the typical sixth-grader would have received the same score, had sixth-graders taken the fourth-grade test. This same statement would also be true of age-equivalent scores. That is, it reflects performance in terms of the age level in the normative sample at which the average score is the same as the client's score.
5. Although all schools must provide services to individuals identified as disabled, there are other populations who, despite being at an advantage, may still need help but are not protected by federal laws and regulations. Because giftedness is not a disability, it is not defined and covered under IDEA. There is, however, a federal definition of "gifted and talented" referring to high-performance/achievement capability students who need services or activities "not ordinarily provided," and the federal government encourages their provision through state enrichment (Gifted and Talented Education [GATE]) programs.

Further Assessment. In practice, when any academic problem is extreme or does not respond to appropriate intervention, further evaluation of the problem is in order. Reading problems in particular may need very careful evaluation regarding the basic processes that might underlie the difficulty. A language specialist who deals with reading issues or a neuropsychologist who might shed light on the pathways available or impaired in a student can perform this assessment.

Processing Assessment. Processing problems can be addressed in a variety of situations and evaluations. One important one is in the context of more specific psychoeducational assessment in which processing problems

are identified in order to understand academic shortfall in basic reading, math, or writing skills. However, another important one is in the context of a more general psychodiagnostic evaluation in which processing issues are thought to be potentially relevant to more general social or diagnostic problems.

In the context of academic assessment, processing problems are identified to help explain academic problems or to help identify a student's learning style. In some cases the identification of processing problems has been necessary to meet criteria for learning disability. In the context of more clinical assessment, processing problems are identified to help explain problems with behavior regulation, cognitive confusion, and social judgment. In these cases the processing problems are handled in a functional and descriptive way.

Processing problems can also be assessed from a neuropsychological point of view (see the later discussion in the subsection Focused Neuropsychological Assessment) in which the processing issues are explored in an attempt to evaluate neurological integrity or to fine tune understanding of specific domains of cognitive functioning with the goal of understanding the child's referral problems and making appropriate recommendations. In this type of evaluation, instruments are chosen and administered in a systematic way so as to evaluate different systems in the brain and describe cognitive functioning in specific ways.

BACKGROUND. In the current psychological climate, it is understood that a child can have social learning as well as academic learning problems. Furthermore, it is understood that the same organizing principles that affect academic behavior can affect interpersonal interactions.

Educational Focus. The purpose of a psychoeducational assessment is to identify and explain learning disabilities and potentially rule out significant psychological factors as causes of school problems. In this context, processing issues are often evaluated as a supplement to achievement and intellectual testing in an attempt to find processing strengths and weaknesses, and to identify the processing factors associated with learning problems. The processing problems that might be evaluated as part of a psychoeducational assessment are diverse and not well defined. Even when schools have required the specification of a processing problem when finding a learning disability according to the discrepancy model, they have tended to include a wide range of processing possibilities (everything from attention to visual sequential memory), often without adequately determining what is in fact contributing to the immediate learning problem. New, more practical, and less theoretical approaches to identifying learning problems (such as RTI) seek to bypass definition and debate in order to provide service and supports to struggling students.

In educational evaluation, "processing disorders" is the term often used

by professionals to designate spheres of difficulty, such as specific perceptual inputs (visual, auditory) and expressive outputs (motor, vocal). The term "processing disorder" also includes difficulties in integrating information between two or more of these areas—for example, visual–motor dysfunction or sensory integration difficulties. The term may also refer to particular functions, such as memory or attention, or other factors, such as organizational issues. Such processing problems can interfere with academic learning and even social interaction.

Clinical Focus. The purpose of a psychodiagnostic assessment is to identify and explain the psychological factors associated with broad referral issues. When processing is explored in this context, tests are deployed to shed light on problematic psychological functioning and referral concerns. The emphasis is on the more integrative functions and control factors that might affect a diverse set of competencies, including general modulation and interpersonal interaction. In addition to standard processing instruments, tests of executive functioning and social cognition might be used, along with relevant rating scales in the context of history, in addition to projective tests interpreted from a variety of cognitive, social, and emotional vantage points. For example, tests or subtests might be selected to understand the interplay between "lower-level" and "higher-level" processing as a way of understanding problems with impulsivity in attention-deficit/hyperactivity disorder (ADHD) or social judgment in autism spectrum disorder (ASD). In this context, although a neuropsychologist might explore the neurological underpinnings of these behaviors and conditions, the more general clinician might "stand mute" regarding the neurological implications of the findings but try to apply research information and clinical understanding to relating test results to diagnostic issues. The practitioner is apt to recruit broad concepts of cognitive integration and functioning, as well as a range of cognitive and projective tests, to understand problematic behavior. In fact, in this context, every cognitive test is potentially seen as a projective test, and every projective test is seen as a cognitive test. Test scores in addition to observational data are enlisted to understand how the client makes sense (or does not) of his or her experience. These and other organizational issues might in turn affect how the client resolves everyday problems.

In the context of clinical assessment, processing issues might be identified in diverse testing, including general intellectual tests, as well as processing tests obviously designed for the specific purpose of assessing cognitive functioning in systematic ways. More important in this context, however, are the tests (and other instruments) that are designed to assess more integrative functions, particularly those structured and unstructured tests that might reveal the ability to make sense of things in general and executive functioning or social cognition in particular. For example, in this

context puzzle-like tasks from previous editions of intellectual instruments (such as the Object Assembly and Picture Arrangement from the "old" WISC-III and WAIS-III) might suggest that the same sequential/narrative problems that affect the ability to arrange cartoon pictures to tell a story that makes sense might also affect the ability to tell a Thematic Apperception Test (TAT) story, or the same holistic/simultaneous problems that affect the capacity to assemble a schematic drawing might also affect the capacity to make sense of Rorschach blots. In turn, these difficulties in making meaning on testing might be associated with "knowing what's going on" in life. For example, the child might not be able to understand social situations as they unfold, or to write a coherent story in school, or to process context in an academic or social situation.

ASSESSMENT. The assessment of processing variables can be conceptualized in terms of different models using omnibus or specialized tests.

Comprehensive Processing Tests. Comprehensive processing tests allow for the identification of areas of strength and weakness when it comes to processing issues. The WJ-COG allows for the extensive assessment of processing across a diverse range of areas and ages. Although it includes composite scores referring to intellectual ability, it is most often thought of as a processing test and is often administered in addition to a comprehensive intellectual test. The WJ-COG data are provided for individual subtests, as well as for informative composite scores. The Differential Ability Scales (DAS) provide a measure of general conceptual ability, as well as cluster scores in cognitive areas. It allows for the analysis of children's learning abilities and their strengths and weaknesses in cognitive processing. It has links with the WIAT that allow comparison of cognitive and achievement scores.

More-Specialized Tests. The NEPSY (a developmental neuropsychological assessment) is particularly useful as an instrument that allows for the systematic assessment of cognitive processing across a range of specific areas of functioning, including executive and attention, language, memory and learning, sensorimotor, visuospatial, and social perception (a relatively new domain). These areas of functioning are derived from neuropsychological constructs, and in the context of general psychoeducational or psychodiagnostic assessment they provide descriptive analyses and comparisons across domains. The NEPSY is only relevant to children and adolescents. The Delis-Kaplan Executive Function System (D-KEFS) is a systematic assessment of executive functioning across a wide range of ages into adulthood. The nine subtests allow for the evaluation of executive functioning at different cognitive levels (e.g., shifting set and verbal fluency, up to abstract thinking and deductive

reasoning). Analysis is made not only of the subtest scores, but also of the strategies the client uses for solving diverse problems. The data can be handled descriptively in terms of cognitive functioning, without comment on frontal lobe integrity.

Other Processing Tests. In addition to these more general tests, there are other instruments designed to provide more specific assessments according to different ways of understanding processing. Because they are comprised of subtests, separate tasks are often undertaken to illuminate different issues.

1. *Functions*. Processing understood in terms of functions refers to an input–output model dealing primarily with visual–motor or auditory–vocal systems. Testing involves the systematic presentation of auditory or visual stimuli, along with vocal or motor output. Such tests include the Test of Auditory Processing Skills (TAPS) and the Test of Visual Perceptual Skills (TVPS), as well as the Beery Developmental Test of Visual–Motor Integration (VMI). In addition, there may be other processing problems, perhaps having to do less with input–output models than pervasive functions, such as attention or memory. The NEPSY includes measures of attention. The Wide Range Assessment of Memory and Learning (WRAML) is helpful in the systematic assessment of what children recollect.
2. *Pragmatics*. As psychologists look at higher and higher levels of functioning, the application to more social spheres becomes apparent. The NEPSY's relatively new subtests (Theory of Mind and Affect Recognition) illuminate social perception. The same interest in "goal-directed behavior" that has informed the exploration of executive functioning has informed the study of social cognition as well. The Test of Problem Solving (TOPS) allows for the assessment of social judgment. Clients are asked questions about social situations presented in pictures or stories. The Test of Pragmatic Language (TOPL) allows for the assessment of pragmatic language—that is, language that is undertaken with a goal in mind.

General Comment. A wide variety of constructs at different cognitive levels have been used to understand processing patterns. Some have been alluded to, including visual versus auditory input and crystallized versus fluid intelligence, as well as nonverbal learning disorder and language-based learning disorder. However, there has been an increasing interest in the concepts of executive function (particularly around ADHD) and social cognition (especially around ASD). As has been noted, although there are tests specifically designed for their detection and description, information about integrative functions can be found in unexpected places, including everything from cognitive testing (puzzle-like tasks) to projective

instruments (with different demand characteristics).

It should be noted that, although assessing processing is commonly part of a psychoeducational assessment, the extent to which the direct remediation of processing problems can benefit learning disabilities varies, and such interventions should be viewed with caution. Some research cited by Kavale and Forness has been very disappointing. Actual treatment (if it does not involve medication for attention) is apt to be linked to the academic area affected by the processing difficulty. However, in contrast, recent studies linking the teaching of basic auditory processing and phonological skills in the context of reading programs are promising and leading to research-based intervention. In addition, simply describing a student's processing issues and how to work around them often helps teachers to instruct students and parents to manage and interact with their children.

There is also increasing information about interventions targeted toward executive functions and social cognition. Although there has been some success with the treatment of some executive functions, the identification of specific remediations for executive functioning, outside of interventions that look like organizational strategies and study skills, has remained problematic. Although there are some interventions targeting specific aspects of social cognition (such as recognition of facial expressions), again treatments for social cognition have often been an aspect of a larger event, such as social skills group. However, executive functioning and social cognition continue to be areas of interest, research and application, and, their identification can be crucial to the diagnostic process, and therefore ultimately to the intervention plan.

Considerations. These are some factors mental health professionals should keep in mind when considering processing problems:

1. The issue of the relationship between processing problems and academic difficulties has become controversial, especially around response-to-intervention, which so far has focused on reading. Some educators have maintained that the identification of a processing problem associated with academic shortfall is necessary to identify the core cognitive deficit predictive of learning disability. Other educators have maintained that even if present, it is not necessary to identify the actual processing deficit in order to prescribe educational remediation, and that the identification of learning disability cannot be separated from instruction.
2. Although it is always tempting to think in terms of organic versus functional, it needs to be remembered that these are interacting systems, and learning new skills can alter brain structure. This has been found in the case of taxi drivers in a major European city, as well as in individuals being taught to juggle.
3. It should be noted that the term "attentional problems" can be used

diagnostically or descriptively. The fact that clinicians detect problems with attention on a student's individual intellectual or processing tests (including computerized instruments) does not mean that the student has ADHD. Under DSM-5 this is a clinical diagnosis that involves specific criteria, is described behaviorally (inattention, hyperactivity, and impulsivity), and is determined by history (several symptoms by age 12 years) and current functioning (several symptoms seen in two settings). Systematic assessment can help to clarify issues around ADHD (and medication), but this will not be done with a single instrument and will require a pattern of responding on a selected group of tools along with interviews, observations, and, usually, rating scales from home and school.

4. It should be noted that learning (and emotional) problems can mimic attention deficit. In assessing ADHD, the clinician will need to rule out artifacts and to consider whether a child stops paying attention because he or she stops understanding (e.g., what he or she is reading or hearing) or has primary ADHD. In addition, the clinician will need to consider whether something like ongoing depression or situational anxiety is making it hard for the child to attend.
5. Broad constructs such as executive function can have relevance for a wide variety of diagnostic categories, such as ADHD and ASD. It is possible that about one third of the adults and children who have ADHD also have executive dysfunction (EDF), which can increase the difficulties facing these clients. As a result, as with ADHD and learning disorder, where there is a large overlap, when one sees ADHD one might look for EDF. The statistical overlap between ASD and EDF is harder to determine because of issues of definition, as well as of research results. Depending on how it is measured, EDF might be seen as ubiquitous in ASD. Research and clinical observations suggest that it would be useful to understand executive functioning in autism because of the usefulness of this information when interacting with affected individuals and making recommendations for them.

FURTHER ASSESSMENT. When processing problems are particularly severe or present a complicated pattern, further assessment may be in order. Neuropsychologists can fine tune information about processing in a variety of spheres.

Conclusions. A wide variety of intellectual, achievement, and processing tests have been explored as they apply to clarifying cognitive functioning in general and describing competencies in the academic and social spheres. There has been some discussion of learning disabilities and disorders, and some attempt to provide information so that psychiatrists might help patients receive services through the school as well as in the

community. However, the same processing issues that affect academic skills acquisition can affect behavioral control and social appropriateness. In addition to cognitive and neuropsychological testing, more projective testing might be undertaken to help identify the organizational issues that contribute to school or social problems, along with the more emotional factors that might interact with them. A battery of psychological tests provides a fairly systematic assessment of different kinds of input and output as they are affected by more and less structure.

COGNITIVE CONCLUSION

In general, when a systematic assessment of brain–behavior relationships is important to understanding referral issues, focused neuropsychological evaluation is indicated. When the clinical focus is served by a careful description of how cognitive functioning and processing issues are related to referral questions, the more general psychological assessment would be relevant and useful. However, every situation should be decided on a case-by-case basis, in consultation with the parents and involved professionals. It should be noted that general cognitive or psychological evaluation and focused neuropsychological assessment are complementary rather than competing approaches. They can be used as alternatives, depending on the needs of the case, or to supplement each other in complicated situations involving testing by one or more professional psychologists.

This subsection has focused on cognitive approaches to testing and psychological assessment in some general way, recognizing the complexity of the enterprise. The next subsection focuses on neuropsychological assessment, with a view to exploring brain as well as behavioral functioning. In this context, processing is seen in neurobiological and not just cognitive terms.

FOCUSED NEUROPSYCHOLOGICAL ASSESSMENT

Neuropsychology is dedicated to the study of brain–behavior relations and has matured into a clinical discipline for the diagnosis and characterization of brain function and dysfunction.

Assessment of Functioning. Neuropsychological assessment was originally developed for the assessment of *adult* patients and was not generally applied to child assessment until later. When it was used it was typically to assess children's *functioning* subsequent to acquired injuries or diseases affecting the brain. This functional assessment was important because the effect of trauma to the brain is highly variable across individuals, even when the precise location and size of the lesion is known. In these situations neuropsychological tests could provide specific functional information that took the child's age and developmental status

into account. This continues to be a prominent point of referral for pediatric neuropsychology because it discusses, not just the existence of the brain disruption, but its *meaning* in terms of the child's ability to function. This discussion of function is important both immediately after the injury and later on when intervention and development have had a chance to have their effect. This use of neuropsychological assessment is important with gross injury to the brain, but it is also valuable in situations in which sequelae are subtle and at risk of being attributed to psychological factors such as grief or poor motivation.

Technical Advances. Recent advances in neuroimaging have added to the use of neuropsychology in child assessment. One reason that neuropsychology was later in being applied to children was that technology for learning about *children's normal brain development* was not available. Given that neuropsychology studies the relationship between behavior *and the brain*, this gap in knowledge meant that inferences about brain function could not be applied to children. In contrast, neuropsychology *was* applied to adult populations based on the impression then held that brain development was finished by adulthood (thereby allowing for generalizations to be made regarding normal brain states and behavior). Because techniques such as positron emission tomography were prohibited in research with children, it was not until the introduction of functional magnetic resonance imaging (fMRI) in the early 1990s that full-scale research of child brain development could begin. Since that time there has been an unprecedented explosion of knowledge that has expanded the scientific understanding of child brain development exponentially every year up to the present.

Developments in Test Instruments. Other advances in pediatric neuropsychology include the introduction of tests specifically designed for use with children. These instruments assess similar behaviors as do their adult counterparts but use paradigms that are more engaging to children and better measure developmental transitions throughout childhood. These instruments are used in clinical assessments but are now also part of many research protocols examining childhood diseases and genetic conditions. Given its precision in measuring behavior, neuropsychological assessment is now involved, not just in the assessment of function after an injury, but also in the initial *diagnostic* processes. Another example of this specificity is the now-routine use of neuropsychological testing in genetic research of developmental disorders of childhood, given the precision that it adds to questions of endophenotype expression.

Neuropsychology's Application to Diagnosis and Treatment Planning. These new technologies have greatly enlarged the current understanding of

both normal and atypical brain development in children, affecting the current knowledge of childhood brain–behavior relationships, as well as diagnosis and treatment planning in pediatric groups. This increased understanding of typical as well as atypical brain development has made neuropsychological assessment useful, not just for children with acquired disorders, but also in cases of developmental disorders. In this context, the term “developmental disorders” is used in reference to a child who is not developing in step with peers but who for unknown reasons struggles greatly or fails to develop specific abilities. Examples are difficulty in learning to read in an otherwise intact child (referred to as the developmental disorder of dyslexia) and problems with developing social or self-regulatory skills (seen in, respectively, autism spectrum and attention-deficit disorders). These disorders stand in contrast to “acquired” disorders, in which a known event, such as injury or illness, has affected the child’s developmental trajectory.

EXECUTIVE FUNCTIONING AND THE PREFRONTAL CORTEX

The term "executive functioning" (EF) has become increasingly popular in the last few years, although many professionals remain unclear about its meaning. Despite the increasing familiarity of this term, many psychiatrists report gaps in their understanding, especially as the concept applies to the lives of their patients. This subsection reviews the history of the construct of EF, explains the factors it is thought to include, and then applies this information to an appreciation of its role in successful adult living.

Broadly described, EF refers to people's ability to coordinate behaviors so as to accomplish what they have set out to do (in other words, to reach their goals). Many of the abilities now included under the umbrella term of EF were recognized and researched before they were organized as part of this larger construct. For example attention, impulse control, and working memory are just three domains included in "executive functioning," but each of these preceded EF in the brain–behavior literature. It was not until later in the development of cognitive science that researchers and clinicians began to write of a larger, unified capacity to which many already-conceived abilities belonged. Furthermore, although many clinicians now recognize the multifactorial nature of EF, confusion continues to exist regarding how the specific abilities it includes were chosen, what they share in common, and why other behaviors (also seemingly important to goal attainment) are not part of this overarching concept. To answer these questions, it is helpful to understand some of the history that contributed to our current understanding of EF as well as the neurological underpinnings of the term and the specific behaviors included in it.

Although constructs included in the EF paradigm had been studied individually for some time, it was not until clinicians and researchers observed injured people who had *lost* this ability that they began to conceptualize it more specifically. The clinicians saw that although their patients (often the victims of severe head injuries) were still able to perform many basic acts independently, they now struggled to organize and calibrate their actions toward higher-level goals. Ultimately these observing physicians realized that the loss of this unnamed overarching ability was equally if not more handicapping than the absence of any of the individual functions making it up would have been.

One of the first and most celebrated of such cases was the story of Phineas Gage. Gage sustained a traumatic brain injury to his frontal lobes after an explosion sent a railroad tie through his head in 1842. He was immediately thrown onto his back convulsing, but after a brief period was reportedly "stunned . . . silent, but awake." Within a few minutes he was speaking and able to ride, sitting erect, in an ox car for the three-quarter-mile journey to the town hotel. An hour following the injury Gage was coherent, walking and talking, and telling the doctor and bystanders how he had been hurt. He survived an ensuing infection, and within 2 months, given the apparent preservation of so many necessary behaviors, was pronounced cured. Shortly thereafter, however, despite his intact hearing, sight in his right eye, and ongoing ability to speak, it became clear that he had changed a great deal. In contrast to his clear thinking and responsible personality prior to the injury, his behavior was now marked by profanity and impatience, as well as disrespect for others, and, most concerning and interesting, abandonment of his plans soon after he had created them. Ultimately, unable to maintain the consistent employment as a foreman at which he had earlier excelled, he worked at odd jobs and then as a circus attraction before his death in 1861.

This incident and subsequent clinical observations and research led to increasing understanding of the characteristics of executive functioning and its development. Three basic aspects of EF have been noted, including its association with the prefrontal cortex, its tendency to be late developing, and its involvement in mental planning.

1. *Prefrontal cortex*. The story of Phineas Gage, as well as that of many other survivors of brain injury or disease, illustrates one of the first characteristics of EF. The functions that are included in the EF construct are those subserved, at least in part, by the frontal (specifically *prefrontal*) cortices. In fact, the prefrontal cortex alone does not subserve executive functioning, but, rather, it works together with the basal ganglia and thalamus as part of several larger circuits that are now known to support this higher-level ability. These circuits include frontostriatal, frontoparietal, and frontocerebellar pathways. Because the extent of these neural connections were not well understood at the time that the concept of executive functioning was developed, in common parlance the prefrontal cortex is often cited as the "focus" of EF abilities. This abbreviation is helpful for people who are just learning about the EF construct and who may be confused as to why some behaviors are included and others are not, especially if they can all be logically described as being helpful in goal attainment. For example, why are set shifting and cognitive flexibility part of the EF paradigm whereas reading is not, especially given that, for some goals, reading seems like a more necessary skill? When facing these questions, it is helpful to stop and remember this first characteristic—namely, that unless a behavior can be mapped onto prefrontal cortex, it is not included in the EF construct.

2. *Later developing*. A second characteristic of EF follows from the first. The prefrontal cortex is one of the last cortical regions to develop, with full maturation peaking in adolescence and continuing into young adulthood. Because of the late development of this brain system, it would also be expected that EF will not be fully expressed until people have reached early adulthood. Thus, it is the later-developing cognitive functions—those arising in earnest in adolescence and reaching maturity in adulthood—that are part of the EF construct. One way of conceptualizing this point is to think about the abilities of a typical 10-year-old child versus those of a 25-year-old adult. One would notice no appreciable difference in their gait if one asked them to walk across the room, but one would likely see remarkable differences if the tasks suggested were those that reached maturity in late adolescence and early adulthood and that relied on mature executive functioning. For example, imagine these two if they were asked to sit through a 1-hour lecture, take out a mortgage, start a small business, or organize a family reunion. Although most 25-year-olds could competently manage these challenges (and thereby reach these higher-level "goals"), it is the rare 10-year-old who could anticipate and handle the multiple requirements of these "executive" tasks.

At this point it may be helpful to demonstrate which of the components of the examples described in the preceding paragraph are parts of the EF paradigm. The first request—sitting through a 1-hour lecture—obviously involves sustained attention, but, beyond that, also requires the *inhibition of distracters*. These two behaviors (attending to this, ignoring that) work together to allow us to *continue to focus* despite waning enthusiasm for the topic and competition for our attention such as a hungry stomach or tonight's football championship. The rest of the examples go beyond just sustained attention and include the need for insight to identify goals, imagination to create steps for reaching those goals, and anticipation of challenges that may arise and will need to be addressed, and then, of course, actually to get started on the work. Once the project has begun, one must be able to self-monitor using both verbal and visual reminders to stay on track, to identify unforeseen roadblocks and generate alternate plans for dealing with them, and then to efficiently switch to those new plans and not cling too long to the former blueprint.

3. *Mental planning*. The EF components necessary for the "25-year-old tasks" just listed almost all involve *thinking first*, followed by taking the appropriate action. Thus, a third aspect of executive functioning is that a great deal of it is accomplished mentally through planning, anticipating, generating solutions, evaluating them, and then making a choice of how to proceed before any actual action is taken. The ability to process ideas and actions mentally is crucial for mature executive functioning because it eliminates the time it would take actually to act out each potential step, making the process faster and more effective. Of course this mental planning must be followed by *action* if the goal is to be reached. But even after the action has begun, to remain effective executive functioning is necessary for monitoring the process mentally, thereby managing

challenges and interruptions that must be successfully overcome if work is to continue.

The characteristic of mental planning as part of executive functioning has its basis in the development of the EF system in general and the frontal lobes in particular. In this way further clarification is added to the understanding of the construct. Although *prefrontal* cortex does not fully mature until late adolescence and early adulthood, some regions of the frontal lobes are developed even at birth and show fast ascendance during childhood. These early-developing regions are predominantly involved in attainment of *goal-directed (motor) actions*. One of the earliest expressions of "goal-directed behavior" might be *seeing* the toy, *wanting* the toy, and then, finally, after weeks of practice and stretching of the body and hands, *actually being able to reach out for the toy and—joy—grab it!* Thus, development of deliberate and goal-directed movement can be seen as a first, concrete step toward the mental functions that serve goal attainment when the person is older. With maturation we become able, not just to move, but to *imagine* our movements. Alongside this development of foresight is the growing capacity to inhibit our movements until we decide whether we actually want to make them. The interaction of these two processes (thinking about doing it, then either doing it or choosing *not* to do it) creates an expertise for planning that stretches into adulthood and beyond.

This internalization of response represents another phase of prefrontal cortical development and another foundational step toward mature executive functioning. In other words, another way of understanding executive functioning is that it is our later-developing capacity to imagine, ponder, consider, and plan *eventual* movements that serve our *now-anticipated* goals. The anterior cortex is all about movement, motor development, and the translation of these skills into "doing." In this regard, executive functioning (also known as goal attainment) can ultimately be thought of as the ability to "do something mentally." By adulthood much of our "doing" occurs *only in our minds* as we imagine, anticipate, rehearse, and analyze potential movements (or actions). This *imagined doing,* however, has its roots in our *actual doing*—the movements we practiced and perfected when we were very young.

Thus, executive functioning is a multifactorial construct reflecting the behaviors that come increasingly "online" during adolescence and combine to address the challenges of adulthood. Although researchers do not fully agree regarding the specific list of behaviors that the EF construct should include, there is largely agreement regarding most of the neuropsychological domains, as identified through the guidelines described earlier, as well as factor analyses of competing potential properties of executive functioning. Typically, factors of response inhibition and execution, cognitive flexibility/set shifting, and planning and organization are included in models, as well as, less frequently, fluency and attention. Some professionals see these cognitive behaviors as providing the scaffolding for higher-level thinking (including abstraction).

FIGURE 36.3–1.

Previously, developmental disorders such as those mentioned earlier were often explained from psychological or educational perspectives, but they can now be reinterpreted from the standpoint of neurology (consistent with long-recognized similarities between symptoms shared by developmental and acquired learning disorders). Clarification of the neurological bases of these relatively common childhood disorders has demystified them and increased public awareness of their validity as diagnosable *and treatable* entities. With these discoveries has come increased funding for research on their treatment, yielding highly successful results and clear evidence regarding which interventions are most efficacious and appropriate for children struggling with these conditions.

Integration of Neuropsychological, Educational, and Psychological Paradigms in Testing. The effect on testing that these breakthroughs have made has also been significant. Most of this subsection has focused on test instruments that are foundational when assessing developmental differences in children. These measures (including, e.g., tests of IQ or academic achievement) are central when youngsters are not keeping up with their peers in one regard or another and are the backbone of testing regardless of whether the assessor is trained as a psychologist or a neuropsychologist or has a background in education. These instruments are pivotal because they measure the major paradigms of both education and psychology, paradigms that govern both diagnosis and the provision of services. These paradigms have been important, not just for the practical purpose of helping young patients receive support, but also because of the quality of the theories they represent and the background of extensive empirical research that has shaped them.

More recently, however, these established paradigms have been joined and affected by new information emerging from cognitive and

neuropsychology in concert with the aforementioned breakthroughs in neuroimaging. The integration of these findings has led to relatively rapid changes in educational law and in the instruments used to test children for learning disabilities. Earlier subsections went into detail about recent legal changes for how learning disabilities are identified. These changes are in large part the result of neuropsychological and cognitive research, including both testing and use of functional neuroimaging to clarify the neurological underpinnings of learning problems and the efficacy of early identification and specific interventions in treating them.

In addition to changes regarding the conceptualization of learning problems are actual changes in the tests used to assess children with unexpected delays. For example, in 2003 the Wechsler Intelligence Scale for Children (4th Edition) included changes in terminology to reflect cognitive and neuropsychological theory (e.g., the Freedom from Distractibility Index was changed to the Working Memory Index), as well as modifications of the subtests themselves, so as to allow for direct comparison between abilities known to reflect typical neural organization (e.g., the WISC-IV provided analogue tests of categorization, one mediated verbally, and the other nonverbally).

Applications of Neuropsychological Assessment. The functional and diagnostic assessment of children and adolescents often begins (and ends) with the kinds of evaluations described in the General Cognitive Assessment subsection. There are situations, however, in which the use of educational or psychological/cognitive testing alone is not able to clarify the diagnosis and determine the most appropriate treatment plan. In those cases the psychiatrist should consider additional neuropsychological testing. To help treating psychiatrists to know more about when to refer a patient for neuropsychological testing, the following list offers specific information about this approach to testing.

1. Neuropsychological tests are designed with the expressed intent of assessing brain function. Thus, if patients present with ADHD-like symptoms but have a history of brain damage (such as a traumatic brain injury or medical history including epilepsy or resection of brain tumor), neuropsychological assessment would be the testing of choice.

2. Children presenting with complexes of behaviors that could indicate more than one DSM diagnosis (e.g., ADHD *and/or* ASD) may benefit from the additional diagnostic clarity that neuropsychological assessment can bring. Because of the known correlations between specific neuropsychological tests and functioning in particular brain regions/circuits, neuropsychology can provide the psychiatrist with an indirect measure of brain function that can add clarity to diagnostic questions. For example, executive functioning profiles vary between

children with ADHD and autism spectrum disorders in some key respects, and so full assessment of executive functioning can add clarifying details when behavioral analysis alone has not been discriminating enough. Specifically, although executive functioning is typically impaired in both ADHD and people on the autism spectrum, those with ADHD tend to have more problems with impulse control and working memory, whereas core deficits in set shifting, cognitive flexibility, and planning are more representative in people with autism.

3. Neuropsychological tests are valuable when one aspect of a child's profile lowers a score that is being given to assess another function. For example, if a child works slowly on a reading fluency task, is the cause his or her reading problems or slow processing speed? Many neuropsychological tests employ a functional hierarchy of demands that allow several basic aspects of a complex skill to be measured independently to reduce confounding variables and more confidently examine the specific behavior initially intended.

For example, the Delis-Kaplan Executive Function System (D-KEFS) includes a version of a common neuropsychological instrument called the "Trailmaking Test" that measures set shifting and cognitive flexibility. On this test the child is shown a page with many circles, each of which contains either a number or a letter. The child must connect the dots as quickly as he or she can, alternating from a number to a letter to a number, in order (e.g., the child would connect "1" to "A" to "2" to B" and so on). The problem with this test as it was originally created was that, although it certainly *did* measure set shifting, it also measured many other abilities that could also lower the score. The D-KEFS version, however, provides four "control" conditions that provide information about other weaknesses that might also affect the set-shifting score. These additional subtests include a visual scanning subtest (showing a page of circles with numbers and letters and the demand to cross out all that contain the number "3"), two subtests that verify that the child can sequence numbers and letters separately (showing stimuli containing circles with numbers and letters and the demand to connect only the numbers and then only the letters), and a subtest assessing motor speed that requires the child to quickly draw a line over a dotted line to connect empty circles. If any of these conditions yields a very low score, the validity of interpreting a poor score on the final "switching" condition is in question, and interpretation is adjusted accordingly.

In addition, the neuropsychological battery contains many tests that appear to be very simple but are included because of the narrow range of abilities that are required for them to be done well. Examples of these are examination of sensory skills such as auditory processing in the right versus left ear, lateralized tests of motor development such as grip strength or pegboard placing, and the administration of timed, simple tests across several skills to highlight selected low scores that fit into known patterns

(such as the language-based weaknesses suggested by compromised rapid color naming alongside weak math automaticity).

4. Another application of neuropsychological testing occurs in instances in which cognitive testing does not demonstrate a clear score profile, not because of the diagnostic complexities addressed in the foregoing, but because of subtleties in score differences that are a result of the examinee's intelligence. These less-profound score distributions can occur when the examinee is very bright or when intelligence is very low.

A very bright examinee can "outsmart" tests that measure multiple abilities, making him or her a poor match for the normative sample to which he or she is being compared. Bright students with learning disorders have many unaffected domains of strength that can be applied to "multiple construct" tests to support the more discrete impairment. Although this will usually result in relatively lower-than-expected scores, in very gifted students the bottom of the range of such scores may be average to high average, making interpretation difficult. It is important to note, however, that the ability to apply a clever strategy to a test does not mean that the student can likewise use intellect to manage his or her symptoms in the larger contexts of life (which is the reason the child presented for assessment in the first place). Rather, the child's clever approach to testing means that there is a greater risk that all scores will cluster together at an above-average point and not demonstrate discrepancies that suggest learning problems. In these cases students are often not diagnosed at all or until much later, after life's demands have caused considerable disruption to their academic, social, and/or emotional success.

Conversely, when a student's intellectual functioning is very compromised, tests that do not include the functional hierarchies discussed earlier make it more likely that all scores will be low. This is a concern, given the importance of identifying relative cognitive strengths in these children because these islands of ability can be quite amenable to intervention and so will demonstrate more improvement with treatment than the other, very impaired skills. When these children are young they should receive intervention for both their areas of profound weakness and their relative strengths, but in adulthood it will likely be the isolated strengths (which, with specific early intervention, will develop on a higher trajectory than those that are more impaired) that provide the foundations for future participation in aspects of adult life.

5. A last benefit of neuropsychological tests that may be important to a psychiatrist is that many of these instruments have been constructed so as to control for "practice effects" (artificial increase of a score resulting because a test was given twice with too little time between the administrations). When practice effects are controlled tests can be readministered at short intervals, providing an effective way of assessing treatments such as medication that should show rapid improvement.

Although this approach needs to be undertaken cautiously (and results interpreted in light of slight improvements seen in nonclinical populations, as are known to occur with some medications), significant changes measured with retesting can add confidence to clinical observations obtained informally or with questionnaires.

One way of managing practice effects is to have two alternate forms of the same test, which allows them to be readministered after only short intervals. Many subtests from the D-KEFS provide these alternate forms, making them a good choice when disorders characterized by executive functioning deficits are part of the referral question. A second way of managing practice effects is with tests that are intentionally boring such as those measuring sustained attention (e.g., the Connors Continuous Performance Test [CCPT]). When, after a short interval, children with normal attention retake the CCPT they typically perform *worse* than they did the first time (contrary to what practice effects would suggest). The decrease in scores is attributed to the boring nature of this long, computerized test; the first time the child takes it there is some excitement about its novelty, but by the second time the child knows what he or she in for from the start, which affects the child's motivation. If a child does significantly better on the second administration, the change is attributed to something other than practice; if done in the course of a medication trial for ADHD, the assumption is that the medicine is helping performance more than the practice is hurting it.

There are several ways that retesting can be used to aid a psychiatrist in quantifying medication effects. For example, when a child who is already on medication for ADHD presents for testing, a 45-minute battery that measures behaviors that are sensitive to the medicine can be administered one morning without medication and a second morning with it. Alternatively, if at the close of testing ADHD is a diagnostic impression and the treating physician thinks a medication trial is appropriate, the child can return for brief reassessment with sensitive tests once the medicine is thought to be at therapeutic levels. Information obtained from retesting can provide data about the benefit of the medication and the appropriateness of the dosing, as well as increasing parents' and patients' confidence about the intervention.

Descriptions and Assessment of Typical Neuropsychological Domains. Typically, in addition to assessing IQ, academic achievement, and social and emotional functioning, neuropsychologists assess domains of memory, attention, executive functioning, language, visual perception, and sensory-motor development. Tests have been developed to examine specific aspects of these domains in isolation so as to increase diagnostic clarity. Although these domains are discussed as different constructs in this subsection, in

truth, they overlap with one another in many different ways. For example, the term *working memory* is often conceptualized as being an aspect of attention, as well as a necessary component of good planning (which is part of executive functioning). It is also a component of memory in that, when it is not well developed, it leads to the phenomenon of "forgetfulness." Many of the measures used to discuss one of these constructs will also be used to discuss others. In addition, neural substrates and pathways supporting their functions are often overlapping or, at least, parallel systems.

Thus, the interdependence of these domains is an assumption of this portion of this section. Each subsection provides an overview of the theory supporting the conceptualization of the domains, followed by a description of the ways in which a neuropsychologist might assess them.

MEMORY

Overview. To memorize something, a person must first pay attention to it. Assuming that this attention is intact, the next step is to "make" the memory, or *encode* it. Once the memory has been made, however, it must be available on demand to be of use. Knowing the name of France's first president is of no benefit if you cannot *retrieve* it during the history examination.

One metaphor for this two-step memory process is a filing cabinet. *Encoding*, then, is when a person puts information into the "filing drawer." Someone with a true amnesic disorder (such as Alzheimer disease) never gets the information into the drawer. No amount of cueing or reminders later on will help the person to recall the information because it never "got into the drawer" in the first place. This type of impairment can be seen in some children, most often those with seizure disorders that adversely affect the temporal lobes. For most children, however, the problem described as "poor memory" is actually a difficulty with retrieval. *Retrieval* is the ability to get information out of the "file drawer" once it has been put in. Poor retrieval is associated with problems of organization (the folders are missing labels) and is more often the issue when children are described as being "forgetful."

To differentiate between encoding and retrieval, children are asked to memorize material and then, 20 to 30 minutes later, to recall it. If they are unable to remember it spontaneously, the examiner does not know whether they have not encoded it or are having problems with retrieval. If the child can remember the material with cueing (e.g., "In the story I read you, was the boy's name Johnny or Sam?"), retrieval is implicated. For the child who cannot encode, however, cueing will not help.

Assessment. In assessing memory, several guidelines should be followed. Both visual and verbal memory tasks should be given. Visual memory tasks (such as learning the location of dots or memorizing faces) are usually

aided by the right hemisphere. In most people, verbal memory tasks (such as memorizing a shopping list or a story) are supported by the left hemisphere. In addition, material to be memorized should include rote tasks (such as word lists), as well as material that is presented in context (such as stories). Some memory tasks assess learning, or the child's ability to benefit from several presentations of the material. It is expected that, after three exposures to a picture of dots, the child's memory of it will be stronger than it was after the first exposure. If not, encoding may be implicated. A 20- to 30-minute delay should also be part of the memory assessment, and cues should be available to differentiate between encoding and retrieval difficulties.

Other terms in the neuropsychological literature appear to describe memory but are actually probably better classified as part of the attention system. These include *short-term memory* and *working memory*. These terms are discussed in the following subsection on attention.

ATTENTION

Overview. The attention literature is large and includes many different conceptualizations. The following illustration demonstrates some elements of good attention.

Suppose you arrive at a lecture hall, open your notebook, and rather than scanning the room indiscriminately, turn your attention to the instructor, who is just beginning to speak (*selective attention*). The lecture is mildly interesting, and you are able to pay attention for the full 20-minute presentation (*sustained attention* or *vigilance*). At the same time that you are listening to the instructor, you are taking handwritten notes incorporating headings and subheadings. It appears that you are able simultaneously to listen, write, and organize rather effortlessly, although you are probably shifting your attention among these competing tasks (*divided attention*). A fire engine goes by the lecture hall and you look up (*distraction*) but are then able to ignore the dimming noise of the siren (*inhibition*) and continue to listen to the lecture (again, *sustained attention*). Suddenly, the fire alarm rings, and you smell smoke. These distracters capture your full attention (*disengagement* from lecture), and their importance causes you to change your attention and behavior (*set shifting*) as you hurriedly head for the door.

A breakdown in any one of these areas can lead to a breakdown in attention. Some children are so caught up in their private thoughts that they have difficulty disengaging from their inner musings, not realizing, for instance, that their babysitter is asking them a question. Or, once they are involved in a task, they have trouble switching to something new, causing them to "get stuck" or perseverate.

The attention literature also discusses divided attention. This is the ability to pay attention to two things at the same time, although, in fact, a

shifting of attention between two things is probably occurring. Good divided attention is important in multitasking. Some authors use the terms *divided attention* and *working memory* synonymously.

Assessment. Assessment of attention requires a number of approaches. Children with attention problems exhibit them at home and at school whenever a task becomes less interesting to them. They function better when working one-on-one with a person or when working on a new activity because it is more stimulating. For this reason, the testing environment may not elicit the inattentive behavior (especially on the first day). To assess the child's attention "in real life" and across settings, attention questionnaires should be completed by both parents and teachers. Many researchers consider this aspect of the assessment of attention to be the most important. Some neuropsychological measures have been found to be sensitive to attention as well. Computerized measures of sustained attention that are designed to be long and boring can capture the loss of attention described here. In addition, specific kinds of performance patterns on these measures have been shown to differentiate different types of attention problems.

Assessment of *verbal* short-term memory might include the repetition of digits or of short sentences. Assessment of *visual* short-term memory can be achieved by having the child point to dots or circles on the page in the same order in which the examiner has just pointed to them. Working memory is usually assessed as the second part of a short-term memory test. It requires that the material that has been stored in short-term memory be manipulated in some way. Verbal working memory can be assessed by having the child repeat digits backward or by doing mathematics in his or her head. Saying the months of the year backward can also assess verbal working memory (as long as the child is able to give them in their usual order without difficulty). Having the child point to the dots on the page in the reverse order in which they are shown can assess visual–spatial working memory.

EXECUTIVE FUNCTIONING

Overview. Executive functioning could be considered to be the mature product of good attention. Although not developed in earnest until children reach adolescence, many aspects of executive functioning begin to appear when children are younger and, thus, can be measured. *Executive functioning* refers to the person's ability to organize his or her behaviors to perform a specific goal. Good executive functioning allows a person to identify problems, generate solutions, choose among them, follow through on the chosen strategy, and evaluate its effectiveness as the work progresses. Without good executive functioning, children who are bright have difficulty demonstrating their abilities. Their parents often report

school underachievement that cannot be explained by learning problems. The issue is not about "knowledge," but rather the application of that knowledge to everyday functioning.

Assessment. Assessment of executive functioning requires several tests, given its many facets. Good attention and working memory, already discussed, are crucial to goal-directed behavior. Inhibition can be tested by giving the child a task in which he or she must control an automatic response. For example, on tests using the Stroop paradigm, the child is shown a color word printed in ink of a different color (e.g., the word "BLUE" written in red ink). The automatic response is to just read the word "BLUE," but this task requires that the child, instead, name the color of the ink. After several items, inhibition appears to tire, and children with poor impulse control perform worse than their peers.

Fluency can be assessed by having the child generate category words under a time limit. For example, a child might be asked to name as many kinds of toys as he or she can in 1 minute. A variant of this task requires the child to create as many designs as he or she can in a 1-minute period, according to strict guidelines.

Cognitive flexibility is often tested with the Wisconsin Card Sorting Test (WCST), a measure of problem solving. On this test, the child is not told how to solve the puzzles; rather, he or she must use feedback that his or her attempts are "right" or "wrong" and is then expected to use this information to generate strategies. During the course of this test, the rules often change without warning, requiring that the child "regroup" and develop a new strategy. This measure generates information about the child's ability to initially figure out the task, his or her tendency to perseverate on wrong responses, and his or her ability to use feedback to generate new responses.

Planning is another aspect of executive functioning. Variants of a "tower" test are often used to assess this ability. On a tower test, the child is shown a picture with colored balls or disks stacked on top of one another on wooden pegs in a specific configuration. The child is told to move the balls or disks on the pegs on an actual model on the table to match the configuration shown in the picture. The child is instructed to move only one ball or disk at a time and to use as few moves as possible. To perform the task well, the child must first "hold back" and not make impulsive moves that may get him or her "cornered." The child must also visualize the first few steps of the problem. Thus, both impulse control and visual working memory are required to exhibit good planning on this rather entertaining test.

LANGUAGE

Overview. Human language organizes, supports, and communicates

knowledge, memories, and ideas. Beyond just allowing people to communicate with others, language organizes both thoughts and emotions, as well as helping people sequence their actions. Although traditionally discussed in terms of left hemisphere functioning, much of the human cortex is involved in various aspects of language. Communication includes both *speech*, the rapid and complex motor movements involved in talking, and *language*, the code used to express thoughts and ideas.

Linguists conceptualize language as being composed of four separate parts: *Phonemes*, defined as the smallest units of sound in a language; *morphemes*, the smallest units of meaning; *syntax* at the level of the sentence (e.g., use of direct or indirect pronouns); and *discourse*, the stringing together of sentences to create a narrative.

In considering language, perhaps the most common distinction made is between *expressive* and *receptive* language. Expressive language requires the production of language, including articulating clearly, finding the right word, and applying grammar and syntax to one's ideas, in addition to vocal fluency and voice tone (prosody). Receptive language involves the ability to comprehend and remember what is said.

Children with expressive language problems may appear to have little to say and considered to be shy. In fact, however, their difficulty may be with self-expression. Some children who are very talkative ("fluent") may also have difficulty with finding the word they want or organizing their sentences to make them clear. The paradox of a fluent child with an expressive language disorder may cause his or her problems to be overlooked.

Receptive language, or the ability to understand what is being said, represents another aspect of the language system. Children with poor receptive language have difficulty in processing information that is spoken to them and may have difficulty learning in the classroom or appear to be inattentive. Sometimes they appear to be oppositional because of their difficulty with understanding (and therefore *doing*) what they are told.

Secondary problems of children with language disorders include difficulties with social interactions and processing of emotions. Language is what humans use to interact and communicate their ideas to one another. When this ability is compromised, children may isolate or try to find less language-intense activities to occupy their time. Emotional problems may ensue from the child's difficulty with using language to express and, therefore, process his or her inner world.

Assessment. Assessment of language should include several measures meant to identify the child's specific language profile. Tests should assess all levels of language, including phonemes, single words, simple phrases, complex sentences, and conversation. Measures of both expressive and receptive language should be included. In the assessment of receptive

language, children are asked to distinguish between similar sounds and words, remember and repeat word lists and related strings of words, point to a picture that depicts a vocabulary word, and follow increasingly complex directions presented only once. In the assessment of expressive language, children are asked to perform tasks such as listing as many round objects as they can within a time limit, naming a depicted or described item, defining words or concepts, or creating a syntactically complex sentence according to strict guidelines.

In addition, the psychologist might explore *pragmatics*, which is the child's ability to participate in conversation and use social language. This involves interpreting nonverbal aspects of communication, as well as observing basic social rules, such as turn taking in conversation. Although neuropsychologists often evaluate pragmatics in addition to receptive and expressive language, they also work in concert with speech and language specialists when additional assessment is indicated.

VISUOPERCEPTUAL FUNCTIONING

Overview. There are several associated constructs in neuropsychology that reflect people's ability to make sense of what they see, to organize it, or to copy it. These abilities are referred to as *visuoperceptual–visuoconstructive abilities*.

Problems with *visuoperception* are distinct from problems with vision. A person with acute eyesight can struggle with perceptual difficulties, such as identifying which of several figures are exactly alike. Some children have difficulty seeing exactly where something is, and these children may have trouble localizing a point in space or judging the direction of a line. *Visuoconstruction* abilities allow a child to join parts to make a whole. These skills require the integration of the motor system with the visual system. Examples include the ability to put together blocks to form a design or to draw three lines to form a triangle.

Problems with visuoperceptual development have academic, as well as social, ramifications. Academic areas, such as mathematics, that are less reliant on verbal support are at risk. In addition, concepts such as time and monetary values may not be clearly understood. Students with these difficulties often exhibit a poor sense of direction, and problems with integrating complex visual arrays may lead to feelings of being overwhelmed. They may also have difficulty "reading between the lines," thereby making comprehension of less tangible reading concepts (such as theme) more elusive.

Social problems are also often seen in students with these delays. Many elements of good social interactions are nonverbal, including the ability to notice and interpret gestures, facial expression, body posture, and tone of voice. Students with visuoperceptual delays may be overreliant on verbal information and may not understand when people are being sarcastic or

when something is said in jest.

Assessment. Assessment of visual processing must address each of the specific elements of this system. The visuoperceptual abilities should be tested using tasks that do not require the child to use his or her hands to produce the response—for example, activities that require the child to identify designs that match or differ from the target, as well as measures of mental rotation (finding which design one is the same as the target, only rotated). Visuoconstruction tasks add the demand of integrating the hands and eyes in producing the response—for example, having the child draw copies of designs or use blocks to create a replica of a model.

SENSORY/MOTOR FUNCTIONING

Overview. The sensory/motor system is also assessed as part of the neuropsychological examination. Lateralized sensory or motor problems suggest neurological problems on the opposite side of the brain and are often correlated with cognitive processes localized to the right or left hemisphere. Tasks requiring the perception of visual/auditory fields or specific actions with right and/or left sides of the body are part of this domain. In addition, integration of perceptions or movements is also assessed.

Motor assessment is further broken down into the assessment of handedness and tests of large versus fine motor development, as well as the ability to plan motor responses (praxis).

Assessment. The sensory examination usually includes assessment of visual fields using clinical methods such as having the child look at the examiner's nose and then determine whether the assessor is moving the right or left extended hand. Similarly, assessment of bilateral auditory perception might include the assessor standing behind the child and rubbing his or her fingers near the child's right or left ear. Other perceptual tests might assess the ability to name unseen objects placed in the child's right or left hands. Finger agnosia is tested by touching a finger when the child's hand is hidden behind a screen and then having him or her indicate which finger was touched. Integration of perception might include having the child follow directions involving a picture that is shown. Age-based normative data are available for all of these tasks.

Both fine and large motor tests are usually assessed on both the right and left sides of the body. Lateralized fine motor tasks include quickly placing pegs in holes with each hand or squeezing a hand dynamometer with each hand to assess grip strength. Fingertip tapping is one way of testing motor sequencing, as do activities requiring the child to repeat sequences of movements from memory. Handedness is best assessed by having the child do a number of tasks with one hand (e.g., "Show me how you use this spoon," "Hand me the dime," and "Throw me the ball.") in

random order. Assessment of difficulties with motor planning can be done using pantomime.

Large motor testing involves having the child demonstrate gait while walking forward and backward, running, skipping, walking a straight line, and balancing on one foot. In cases in which the findings of motor screening are significant, the neuropsychologist might refer the child to an occupational or physical therapist for further, more specific, evaluation.

NEUROPSYCHOLOGICAL CONCLUSION

Neuropsychological testing is commonly undertaken according to several discrete domains that reflect areas of brain functioning. Typically, these include attention and executive functioning, memory, and language, as well as visuoperceptual and sensory/motor functioning.

Considerations. In considering neuropsychological issues, the following factors should be kept in mind:

1. After early brain injury, language and motor functioning are the most likely to benefit from "plasticity." Some research suggests that, with this process of reorganization, other functions (most notably, visuoperceptual abilities) may be "crowded out," yielding scores that are lower than expected.
2. Interventions for neurologically driven developmental delays have their most profound effect on younger children. Recent studies have shown that, in children with reading disabilities, bilateral representation of language identified with fMRI before intervention shifted to the left hemisphere by several orders of magnitude in every subject after only 80 hours of reading intervention. These changes in the brain were accompanied by improved reading skills. Thus, the philosophy of delaying intervention until a deficit is fully expressed may keep children from receiving the full benefit that early intervention provides.
3. Risk factors for reading disabilities include family history, early language delays, poor articulation, chronic ear infections, poor early rhyming abilities, inability to recite (not sing) the alphabet by the end of kindergarten, and early brain injury.
4. "Ambidexterity" (consistently using the right hand for some specific tasks and the left hand for other specific tasks) often runs in families in which several members are left-handed. In contrast, "ambiguous handedness" (or the use of either hand for the same task; sometimes writing with the right hand, sometimes writing with the left hand) can be a pathognomonic sign suggesting poor cerebral organization for specific behaviors.
5. ADHD more adversely affects abilities typically associated with right hemisphere functioning (such as fine motor skills and visuoperceptual abilities) and affects attention and executive functioning.

Psychostimulant medication has been shown to improve functioning in all of these domains in children with ADHD.

DIFFERENT APPROACHES

General cognitive assessment and focused neuropsychological assessment represent different paradigms: One assesses referral issues in a descriptive way without reference to brain-behavior relationships, and the other assesses them in a functional way from a brain-behavior point of view. When these two approaches are seen together, psychological assessment undertaken from a more processing standpoint allows opportunities to "zoom in" (to the neurobiological level) or "zoom out" (to the observational level). However, regardless of whether the psychologist is working from a more "macro" or "micro" level, ultimately it is his or her ability to integrate these and other test data with clinical observations, referral issues, and diagnostic possibilities that benefits the patient in a practical way in terms of understanding what is going wrong and making appropriate interventions.

REFERENCES

Baron IS. *Neuropsychological Evaluation of the Child*. New York: Oxford; 2004.

Denckla MB. A theory and model of executive function: A neuropsychological perspective. In: Lyon GR, Krasnegor NA, eds. *Attention, Memory, and Executive Function*. Baltimore, MD: Paul H. Brookes; 1996.

Dennis EL, Thompson PM. Typical and atypical brain development: a review of neuroimaging studies. *Dialogues Clin Neurosci*. 2013;15(3):359–384.

Fletcher JM, Lyon RG, Fuchs LS, Barnes MA. *Learning Disabilities: From Identification to Intervention*. New York: Guilford Press; 2007.

Fuster JM. *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe*. Philadelphia, PA: Lippincott-Raven; 1997.

Hale TS. A distributed effects perspective of dimensionally defined psychiatric disorders: and convergent versus core deficits in ADHD. *Front Psychiatry*. 2013;5:62.

Kavale KA, Forness SR. *The Nature of Learning Disabilities: Critical Elements of Diagnosis and Classification*. Hillsdale, NJ: Lawrence Erlbaum; 1995.

Kenworthy LE, Black DO, Wallace GL, Ahluvalia T, Wagner AE. Disorganization: The forgotten executive dysfunction in high-functioning autism (HFA) spectrum disorders. *Dev Neuropsychol*. 2005;28(3):809.

*Kolb B, Whishaw IQ. *Fundamentals of Human Neuropsychology*. New York: Freeman; 1996.

LeFever LE. *Adult Attention Deficit Disorder: Brain Mechanisms and Life Outcomes*. New York: New York Academy of Sciences; 2001.

Leung RC, Vogan VM, Powell TL, Anagnostou E, Taylor MJ. The role of executive functions in social impairment in autism spectrum disorder. *Child Neuropsychology*. 2016;22(3):336–344.

*Lezak MD, Howieson DB, Loring DW, Hannay HJ, Fischer JS. *Neuropsychological Assessment*. 4th ed. New York: Oxford; 2004.

Lichtenstein R, Klotz MB. Deciphering the federal regulations on identifying children with specific learning disabilities. *Communiqué (National Association of School Psychologists)*. 2007;36(3). Available at http://www.nasponline.org/publications/cq/mocq363regs.aspx.

Meltzer L. *Executive Function in Education: From Theory to Practice*. New York: Guilford Press; 2007.

O'Hearn K, Asato M, Ordaz S, Luna B. Neurodevelopmental and executive functioning in autism. *Dev Psychopathology*. 2008;10:1103.

Parke EM, Thaler NS, Etcoff LM, Allen DN. Intellectual profiles in children with ADHD and comorbid learning and motor disorders. *J Atten Disorder*. 2015;25.

Pennington B. *Diagnosing Learning Disorders: A Neuropsychological Framework*. 2nd ed. New York: The Guilford Press; 2008.

Poldrack RA, Congdon E, Triplett W, et al. A phenome-wide examination of neural and cognitive function. *Sci Data*. 2016;3:160110.

*Sattler JM. *Assessment of Children: Cognitive Applications*. San Diego, CA: Sattler; 2001.

Shaywitz S. *Overcoming Dyslexia: A New and Complete Science-Based Program for Reading Problems at Any Level*. New York: Knopf; 2003.

Simos PG, Fletcher JM, Bergman E, Breier JI, Foorman BR. Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology*. 2002;58:1203.

Sinzig J, Morsch D, Bruning N, Schmidt MH, Lehmkuhl G. Inhibition, flexibility, working memory and planning in autism spectrum disorders with and without comorbid ADHD-symptoms. *Child Adolesc Psychiatry Mental Health*. 2008;2:4.

Stiles J, Moses P, Passarotti A, Dick FK, Buxton R. Exploring developmental change in the neuronal bases of higher cognitive functions: The promise of functional magnetic resonance imaging. *Dev Neuropsychol*. 2003;24:641.

Tannock R. Specific Learning Disabilities in DSM-5: are the changes for better or worse? *Int J Res Learn Disabil*. 2013;1(2):2–30.

Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biol Psychiatry*. 2005;57:1336.

*Yeates KO, Ris MD, Taylor HG. *Pediatric Neuropsychology: Research, Theory, and Practice*. New York: Guilford Press; 2000.

▲ 36.4 Psychiatric Emergencies in Children and Adolescents

RUTH S. GERSON, M.D., AND FADI HADDAD, M.D.

INTRODUCTION

The past 20 years have seen a dramatic increase in the number of visits to emergency rooms by children and adolescents in psychiatric crisis, with estimates now reaching 2 to 5 percent of all visits, or over 800,000 visits per year. Unfortunately, appropriate assessment and treatment are not available to many youth presenting to the emergency department (ED) in psychiatric crisis. As an example, nationwide fewer than 50 percent of youth presenting to the ED with suicidal behavior receive a mental health evaluation, even though those who do not are at increased risk of a future suicide attempt. As the ED is the entry point to psychiatric care for an increasing number of youth, improving the care provided to these patients is imperative. This section will underscore the key components of the emergency evaluation of a child or adolescent in psychiatric crisis and will emphasize:

- Understanding the role of developmental, academic, and social context in assessing children in psychiatric crisis.
- Differential diagnosis of behavioral disturbances.
- Assessment of suicidality, aggression, new-onset psychotic symptoms, and risk-taking behaviors.
- Assessment of nonverbal children (including young children and those with autism or intellectual disability).
- Management of agitation in the ED.
- Legal issues in emergency child psychiatry, including assessing for child abuse.

EPIDEMIOLOGY

The emergency care system has unfortunately been slow to adapt to the increasing demand for child and adolescent psychiatric emergency care. Youth in psychiatric crisis are difficult for EDs to manage. Most children presenting to EDs in psychiatric crisis are seen in general medical or pediatric emergency rooms, or more rarely in adult psychiatric emergency services (PES). Most EDs lack the specialized staff (psychiatrists, child psychiatrists, social work, and nursing staff) necessary to effectively care for youth in psychiatric crisis, who also have longer length of stay and increased need for inpatient admission or transfer than other young patients. EDs also typically lack the physical resources to manage the high volume of child psychiatric patients safely. Agitated or suicidal youth are

often placed together in medical exam rooms full of potentially dangerous objects, in busy and noisy holding areas, or on stretchers in trauma bays with security or police watching over them. Youth that are seen in the adult PES instead are exposed to adult patients who are severely mentally ill, homeless, or intoxicated. This can be frightening and reinforces the stigma surrounding mental illness and psychiatric treatment. The PES also lack appropriate staff, as adult psychiatrists and nursing staff trained to work with adults may not know how to approach youth in a developmentally appropriate way. These factors can exacerbate a youth's anxiety or agitation, such that 15 percent of youth in the ED for psychiatric complaints end up restrained during their ED stay, putting themselves and others in danger.

Youth accessing emergency psychiatric care are increasingly presenting with severe symptoms including suicidality, severe aggression, mania, or psychosis, and many of these youth require inpatient hospitalization for safety and stabilization. But with the steady decline in access to inpatient child and adolescent psychiatry beds, young people are often left boarding in EDs for days, without meaningful treatment, awaiting inpatient admission.

Adding to the structural challenges of treating youth in psychiatric crisis in the ED, these patients are often clinically very complicated. ED clinicians are not only treating youth with known psychiatric illness presenting in behavioral crisis, they often also are treating youth with previously undiagnosed or undetected psychiatric illness including depression, anxiety, psychosis, posttraumatic stress disorder, substance use, autism, and attention-deficit hyperactivity disorder. Because of the shortage of outpatient clinic services to diagnose and treat these youth, psychiatric illness often goes undetected and untreated. It is not until symptoms have reached a breaking point that the child accesses care. Thus, by the time a youth presents to the ED, their untreated psychiatric symptoms may also have disrupted their school performance, social interactions, and family relationships. These secondary effects often prompt the ED visit, as in youth sent from school for aggressive outbursts or brought in by a foster family no longer willing to care for them due to difficult behaviors. These systems complexities make these youth even more difficult for ED clinicians to assess and treat.

Suicidal thoughts and severe aggression are also common precipitants for an ED visit, again often subsequent to a lengthy period of unrecognized or untreated psychiatric symptoms. Suicide is currently the fourth leading cause of death for 10- to 14-year olds and the third leading cause of death for 15- to 19-year olds, and as many as 10 percent of children attempt suicide sometime during their young lives. Studies indicate that 40 percent of youth who attempt suicide had been seen in the ED in the year before their death, suggesting that an opportunity for life-saving intervention at

that ED visit was missed. EDs are also increasingly seeing violent youth, who are often "frequent fliers" in the ED. Youth who frequently use the ED for mental health complaints are often there for threatening others or other criminal behaviors. They frequently have diagnoses of conduct disorder or other disruptive behavior disorders; are involved with the juvenile justice or child welfare systems; and have cognitive impairments that make it difficult for them to function in the community.

Though youth in psychiatric crisis are presenting to EDs in greater numbers with increasingly severe symptoms and complex systems involvement, clinicians in the ED lack training in how to evaluate and treat these youth. Pediatricians and emergency medicine physicians evaluating youth in psychiatric crisis are uncomfortable assessing suicidal versus nonsuicidal self-harm, violence risk, and complex presentations of psychiatric illness such as new-onset hallucinations or complex trauma reactions. Adult psychiatrists often lack the developmental perspective and understanding of psychosocial context necessary to inform a comprehensive assessment. This chapter will outline the key priorities, principles, and components of a comprehensive evaluation and risk assessment for a child or adolescent in psychiatric crisis.

PRIORITIES OF EMERGENCY PSYCHIATRIC EVALUATION FOR CHILDREN AND ADOLESCENTS

Safety

The psychiatric evaluation of a child or adolescent in the ED should have two main priorities. First, safety, which in the ED includes both safety of patients and staff while the patient is in the ED, and effective risk assessment to determine whether the child is safe to leave the emergency room or requires inpatient treatment. Ensuring safety in the ED starts with appropriate physical space. Children and teens presenting to the ED with suicidal ideation, aggressive agitation, or other acute psychiatric manifestations are unlikely to be safe in a standard medical exam room. The ED environment, whenever possible, should be modified for safety including: removing potentially dangerous items (sharps, cords, IV poles, and other objects that could be used for self-harm or as a weapon to harm others), decreasing overstimulation by avoiding bright lights and unnecessary noise, providing privacy while allowing the patient to be observed continuously, and enhancing patient comfort by giving the patient a comfortable place to lie or sit down as well as food or drink. The child and family should be given clear expectations as to the components of the evaluation, the waiting time, and the interventions and consequences that will be given should the patient become agitated or violent. Adequate trained personnel should be available to help if the patient becomes agitated, and staff should have a clear response plan for calming an

agitated or aggressive patient including with medication or restraint if needed (this is discussed further below). These modifications reduce the likelihood of aggression to self or others, and if such aggression does occur, mitigate the risk of serious injury to patient or staff. The second aspect of safety is a comprehensive and effective risk assessment to determine the child's safety to leave the ED or need for inpatient treatment. Such an assessment is difficult and often puts the ED team at odds with the parent, who may wish to take the patient home against medical advice or who may wish to hospitalize the patient when this is not clinically indicated. Assessing and managing risk of harm to self or others is discussed more at length below.

Evaluating Medical and Psychosocial Causes of Behavioral Disturbance

The second priority of the ED evaluation is ruling out and treating medical and social/systemic causes of behavioral disturbance. While children and adolescents generally lack the complex and chronic medical conditions of adult psychiatric patients, medical illness can nonetheless masquerade as psychiatric illness in subtle and confusing ways. Seizures (particularly temporal lobe seizures and absence seizures), infectious and autoimmune encephalitis, brain tumors, head trauma, and tropical diseases can manifest with hallucinations, mental status change, and behavioral dysregulation. Children may also present with behavioral changes related to undiagnosed chronic diseases such as porphyria, sickle cell anemia, metabolic disorders and other genetic syndromes. Drug intoxication and other self-poisoning (due to accidental ingestion or intentional acts meant as a suicide attempt or to get high) can also present with psychiatric symptoms. Red flags that are suggestive of a possible medical etiology include:

- Clouding of consciousness
- Waxing and waning mental status
- Disorientation
- Memory impairment
- Acute onset of psychosis in a child with no previous psychiatric history
- Abnormal vital signs

When a medical etiology is suspected, even if it is unlikely, an appropriate workup should be completed to ensure that a medical illness is not missed. Initial workup should include a full physical examination; diagnostic studies including a complete blood count with differential, liver function tests, chemistries including glucose and electrolytes, an electrocardiogram, urine drug screen, blood alcohol level, medication blood levels if applicable, neuroimaging when indicated, and in certain cases cerebrospinal fluid studies if encephalitis is a potential concern. Children with acute medical illness, who are severely intoxicated, or who have made

a suicide attempt by overdose or other method that has caused significant injury, will likely require medical admission prior to psychiatric hospitalization, and should be followed by the psychiatric consult–liason service during their medical care to avoid unnecessary delay in their psychiatric treatment.

Clinicians should also remember that normal developmental changes, such as puberty, and minor physical ailments such as ear infections, tooth pain, hunger, or fatigue can manifest with or trigger psychiatric or behavioral symptoms, particularly in young children and youth with intellectual or developmental disabilities (Table 36-4–1).

Table 36.4–1.
Common Medical Causes of Psychiatric Disturbance in Children and Adolescents

Fever
Nonconvulsive status epilepticus
Temporal lobe seizures
Absence seizure
Encephalitis
Central nervous system neoplasm or other occult malignancy
Cerebral systemic lupus erythematosus
Childhood confusional migraine
Head trauma or postconcussive syndrome
Porphyria
Sickle cell disease
Substance intoxication
Withdrawal from illicit substances
Anticholinergic delirium
Steroid psychosis or steroid-induced mania
Medication adverse effects such as akathisia from risperidone or irritability from levetiracetam

While medical illness can and does masquerade as psychiatric illness, what is even more common in children is psychosocial crisis manifesting as psychiatric disturbance. It is crucial to consider when a child presents with behavioral disturbance or even more classically psychiatric symptoms such as hallucinations or anxiety, that the child may be the “identified patient” in a distressed or pathological family or social system. Domestic violence and family conflict often manifests with aggressive behavior in a child; reports of auditory hallucinations may be a child’s mimicking the symptoms of a mentally ill parent; or the child may be a victim of physical or sexual abuse or of fabricated or exaggerated symptoms in cases of medical child abuse. Treating the child alone when the source of the symptoms are in the family or social environment not only puts the child at risk for adverse effects from unnecessary medication, but is also psychologically damaging as it teaches the child that the problem is him, not the toxic context in which he

finds himself. Alternatively, the child may have true psychiatric illness, but the emergency is not a worsening of symptoms per se but rather that the caretakers are for whatever reason less able to manage those symptoms. A parent may be weakened by illness, divorce, or financial stress, or may lose a key support or structure like an after-school program or grandparent willing to help care for the child; or the child may be removed from the parent and placed in a foster care home where the caretaker is not equipped to care for a child with mental illness. Here, again, the crisis is not psychiatric per se, but rather a breakdown of the caretaking system, which must be shored up if the child is to be discharged safely.

PRINCIPLES OF EMERGENCY PSYCHIATRIC CARE FOR CHILDREN AND ADOLESCENTS

Comprehensive Assessment of the Child

When a patient is in crisis, treatment must start immediately, so the initial assessment must be therapeutic in itself. Given the case complexity of youth presenting to the ED in psychiatric crisis, the first step in the evaluation is gathering essential information. This is more difficult than with adult patients, as information cannot be obtained from the patient alone. When a child or adolescent is brought to the ED, collateral from the parent or guardian is critical to give the facts of the situation that the child is unable to give due to cognitive and emotional immaturity. Collateral also provides crucial information about the child's life at home and at school, their strengths, stressors, medical issues and history (including allergies), history of psychiatric treatment (including response and compliance), developmental history, history of substance use, family history, legal history, social history including any prior trauma exposure, and academic history.

But the assessment should not be based exclusively on the parent or guardian's report, even if the child is very young or intellectually disabled. A careful, curious, and empathic interview of the child is imperative to understand their perspective, the reasons behind their behaviors, their interpretation of what is going on around them, their strengths and cognitive abilities, and the internalizing symptoms such as depression, anxiety, and trauma-related symptoms that often are missed by parents and other adults (Table 36.4–2).

Conducting the interview with an aim to truly understand and empathize with the child's perspective can be very therapeutic for the child. Most children have not had the experience of an adult truly listening to them, striving to understand their experience without judgment or reprimand, and trying to help them. This approach will be more effective than a "checklist" approach to the mental status examination. The interview with the child must include questioning about suicidal and

homicidal ideation, though with the recognition that children and adolescents often minimize or deny suicidal or homicidal ideation or intent, as they fear getting in trouble or being hospitalized. An empathic and nonjudgmental stance to these questions can increase the likelihood of an honest response and also serves to build rapport, even with hostile adolescents. Questions about hallucinations and delusions are also critical, though the clinician should be careful with phrasing these questions as young children or those with developmental delays may misinterpret the question (or their own internal experience) and falsely answer yes, while teens who are truly psychotic may deny such symptoms for fear of being labeled as "crazy." A therapeutic stance to the interview can help young patients to trust the process and the clinician, and starts to break the stigma around mental health treatment.

For very young or otherwise minimally-verbal or nonverbal children (such as those with autism or intellectual disability), the clinician should still attempt to engage the patient to complete a thorough mental status examination. A child with little expressive language may still be able to follow commands, engage in parallel or reciprocal play, and express needs through gestures. The clinician can also observe the child's fine and gross motor skills through drawing and playing, and can assess frustration tolerance, hyperactivity, distractibility, attachment to the parent or guardian (or evidence of fear of the parent), and anxiety with separation. All of these observations will lend diagnostic clarity and provide a therapeutic, calming experience for the child.

Table 36.4–2.
Components of a Comprehensive Assessment

Identifying information
Chief complaint
History of present illness
Past psychiatric history
Current medications and medication history
Past medical history
Substance use history
Trauma history
Social history including trauma and child welfare involvement
Developmental and educational history
Family history
Legal history
Mental status examination including cognitive screen
Medical clearance including physical if indicated

A therapeutic stance is equally important for interviewing parents, foster parents, and other caregivers, who are often exhausted from dealing with frightening and disruptive symptoms and behaviors without a good

understanding of their cause. Treating child psychiatric patients in the ED often feels like having two patients—the child and the family—and approaching both carefully and empathically is crucial to understand the situation and create buy-in for ongoing treatment. A parent's anxiety or agitation can be ameliorated by empathizing with how difficult a child's mental illness can be for a family, and acknowledging that the ED evaluation itself is a difficult process. Reducing the parent's anxiety or anger not only makes the ED evaluation easier, it can also help the parent to be more open to changing their own behavior toward the child and to obtaining treatment. These interviews with both the parent and the child, should be conducted in a quiet, private space free from interruption, so that everyone feels safe and respected when discussing their challenges and expressing painful feelings.

In many cases, a child will be brought to the ED by someone other than their parent (police, EMS, friends, or child welfare workers, for example). The parent should be contacted as quickly as possible, as their involvement will be crucial to the assessment and treatment planning. If the parent or guardian cannot be reached, most states allow clinicians in an emergency to initiate evaluation and treatment of the child, though ongoing attempts to reach the parent should be made. Occasionally, an adolescent will insist that the parent not be contacted. While many states allow adolescents to seek mental health services without the parent being involved or notified, these youth should be encouraged to include their parents in the evaluation, and often with support and empathy of the ED staff a rapprochement can be made between the parent and child.

The initial evaluation does not conclude with an interview with the patient and parent or guardian. A child's teachers and guidance counselors can report on behaviors at school and academic achievements that shed light on the child's cognitive abilities and social skills with peers. Outpatient providers (counselors, case workers, therapists, psychiatrists), and social services workers (child welfare workers, other case managers) can provide very helpful perspective on the child's recent symptoms and mental state, the family's strengths and vulnerabilities, and the resources and supports available to them in the community. The pediatrician can also give an important developmental perspective, having often worked with the child and family for many years, and can identify any concern about compliance, response to treatment, medical issues, physical traumas, and growth which could be a indicative of child abuse or neglect. Consent from the parent or guardian should ideally be obtained before contacting any of these individuals, but in potentially life-threatening situations such as suicidal ideation or concern for child abuse, the safety of the patient outweighs the risk of confidentiality, and the clinician can contact these individuals for information without violating the requirements of Health Insurance Portability and Accountability Act (HIPAA).

By integrating all the information into a comprehensive assessment, the ED clinician can begin to understand the child, family, and contextual factors that are contributing to the crisis at hand. Often the ED clinician will find that the descriptions of symptoms and behaviors and the explanations for such will vary between the child, the parent, and the other informants. These differences may stem from real differences in the child's symptoms and behaviors across different contexts (for example, at home or at school). Alternatively, a parent or other adult may not be aware of or attuned to the child's internal state, and thus may miss internalizing symptoms and over-emphasize externalizing behaviors. The child may have poor insight, may be guarded or minimizing of symptoms, or alternatively may be voicing a major stressor (such as abuse, domestic violence, or school bullying) that the adults are attempting to hide or of which they are unaware. Understanding these sources of variance and integrating them into a clear formulation is necessary to determine an appropriate and effective treatment plan.

Effective Psychotherapeutic and Psychopharmacologic Management of Agitation

Key to ensuring that the initial evaluation is therapeutic is effective management of agitation to avoid restraint and seclusion. Children and adolescents experience restraint and seclusion as frightening, and their perception of time can become distorted such that they perceive themselves as being restrained for much longer than it is in reality. After restraint children are angry at and frightened of staff, which will profoundly limit the staff's ability to fully assess and treat them. While 15 percent of youth in the ED for psychiatric complaints are restrained, much of this can be avoided through proactive engagement to prevent a crisis and by effective verbal and psychopharmacologic de-escalation strategies.

A safe environment is the first step to prevent agitation. Overcrowding, excessive noise, lack of privacy, long wait times, and a tense atmosphere can agitate even a normally quiet child. For a child with impulse control problems, psychosis, trauma history, or autism, such a setting can be catastrophic. A safe ED environment, as described above, will help avert a crisis. Staff who are vigilant to the early signs of agitation and who communicate effectively as a team can notice small shifts like changes in the patient's mannerisms or tone of voice, or their way of interacting with their family, other patients, and the surrounding milieu. When a child is showing these early signs of agitation, the ED staff should work as a team to deescalate the situation. First, the patient should be moved to a safe and quiet area, if they are not already there, and made as comfortable as possible by giving them personal space, separating them from other patients, and either separating them from or rejoining them with preferred family members based on the patient's preference. The patient should be

offered food, a drink, or basic entertainment (such as TV or a video game) to calm and distract them. One member of the staff should approach the patient in a nonthreatening way to listen to their complaint or concern, direct them away from the focus of their anger, and encourage them to verbalize their needs. If the child begins to explain their concern, even if they are still agitated or yelling, staff should empathize by repeating their statement or paraphrasing it back to ensure that the patient's view is understood correctly. As the child makes incremental steps to calm down, for example by talking to staff or by refraining from pacing, staff should praise these positive steps even if the child is still engaging in other negative behaviors (such as cursing or scowling). Offering a child choices, even small ones like which staff to speak to or where to sit, will help the child feel more in control and thus calm down. Most children in the ED are anxious about the process, worried about their parents' reaction and the possibility of hospitalization. Others may be paranoid, may have been abused and thus be hyper-reactive to perceived threats, or may be hungry or in pain. It is important to emphasize that staff wish to understand what happened to the patient to make him react in this way, and not to be judgmental or accusatory. At the same time, staff should be clear that they will work to keep the patient and all others safe and that violence will not be tolerated.

A child's agitation can also stem from a parent's anxiety, anger, or fear. Effective management of child psychiatric patients in the ED requires staff to contain the parent's anxiety, and redirect their response in a positive way. Parents can be anxious in the ED because of concerns about the child's symptoms or safety, fear that the child will be taken away from them (either by child protection or by hospitalization), or worry about missing work or affording the cost of the emergency evaluation or the ambulance that brought the child. They may be fearful about stigma, or they may have had their own negative prior experiences with psychiatric care. If the parent is upset about the child being in the ED, they may express this in a way that agitates the child, either by communicating their fears and anxieties or by blaming or berating the child for being brought to the hospital. The team must listen to and empathize with the parent, emphasize that the parent is not being blamed for the situation, and assure them that the staff will work with the family to ensure the child's safety. If a parent is agitated, it is most often helpful to separate the parent and the child while staff work to calm the parent, to avoid the child becoming agitated as well.

If the patient continues to be agitated and begins to escalate further, perhaps making threats or beginning to destroy ED property or attempt to elope, oral medications should be offered. If the patient is already in psychiatric treatment, carries a diagnosis and is on medication, this can guide the choice of medication for agitation. For example, an adolescent

with known anxiety disorder on a benzodiazepine might benefit from a small oral dose of the same or a different benzodiazepine, while a teen with schizophrenia may require an extra dose of their antipsychotic medication. Medications should be chosen carefully, as the medications most commonly used for agitation in adults can have significant adverse effects on children. Benzodiazepines are often disinhibiting and can cause paradoxical agitation, particularly in young children and those with developmental disabilities or intellectual impairments. Youth who are antipsychotic naive and are given high-potency neuroleptics for agitation, particularly agitation not related to psychosis or mania, are at increased risk for extrapyramidal symptoms. Often the antihistamine diphenhydramine (Benadryl) is the safest medication choice, particularly for young children (though in those with intellectual disability or autism diphenhydramine too can be disinhibiting) (Table 36.4–3).

Oral medication should always be offered before injectable medication (IM), particularly because many children (even ostensibly tough ones) are afraid of needles and can become more agitated when approached with an injection. Often, administering an injection will require a physical hold or restraint, so this should be a last resort for safety after all other interventions have failed. A child who initially refuses an oral medication may accept the oral medication if offered the choice between oral and injectable medication. This not only allows the patient to avoid the potentially traumatic experience of a restraint and IM, but also allows the child to make a choice and begin to demonstrate self-control.

If all of the above techniques to deescalate an agitated child have failed, and the child continues to refuse oral medication and engage in violent or dangerous behaviors, an injectable medication may be indicated for acute chemical restraint. As with oral *pro re nata* (PRN) dosing, if the child is already on an antipsychotic the same medication should be used if it is available in immediate-release injectable form, to avoid the greater risk of adverse effects inherent in concurrent use of two antipsychotics. If the child is not on a standing neuroleptic, the choice of medication should be based on the child's age, size, recent medication administration, medical conditions, and (if known) the psychological cause of their agitation. For younger children, diphenhydramine IM may be preferred as it is sedating and generally well tolerated, though total daily dose should be monitored as over-medication can lead to disinhibition, confusion, or delirium. For adolescents who are agitated due to anxiety, benzodiazepines such as lorazepam (Ativan) or sedating lower-potency antipsychotics such as chlorpromazine (Thorazine) may be effective, while those youth that are agitated due to mania or psychosis will likely require haloperidol or another high-potency antipsychotic. If high-potency antipsychotics are given, they should be given concurrently with diphenhydramine or benztropine (Cogentin) as prophylaxis for acute dystonic side effects, and

the patient should be monitored closely for such. Adding diphenhydramine to haloperidol (Haldol) or chlorpromazine can also magnify the sedating effect of the antipsychotic for a youth who is extremely agitated, particularly when due to mania or psychosis. In general, dosing of IM will be half of the oral dose used, such that for example 50 mg of oral diphenhydramine or chlorpromazine will correspond to 25 mg IM.

Table 36.4–3.
Medications for Agitation

Child (Age <12 or up to 90 lbs):		
Mild Agitation:	**Moderate Agitation**	**Severe Agitation**
Diphenhydramine 12.5–50 mg PO Or Risperidone PO 0.25–2 mg PO	Diphenhydramine 25–50 mg PO or IM Or Chlorpromazine 12.5 mg PO or IM Or Risperidone 0.25–2 mg PO (can be given with Benadryl 12.5–25 mg PO)	Chlorpromazine 12.5–25 mg IM Or Diphenhydramine 25–50 mg IM
Adolescent (Age >12 or >90 lbs):		
Mild Agitation	**Moderate Agitation**	**Severe Agitation**
Diphenhydramine 25–50 mg PO Or Risperidone 0.25–2 mg PO Or Chlorpromazine 25–50 mg PO	Risperidone 0.25–3 mg PO (can be given with Diphenhydramine 25–50 mg PO) Or Chlorpromazine 25–50 mg PO or IM	Haloperidol 1–5 mg IM with Lorazepam 1–2 mg IM (if lorazepam is contraindicated such as patient has intellectual disability or autism, give benztropine 1 mg IM with haloperidol instead) OR Olanzapine 2.5–5 mg IM (do not administer within 1h of benzodiazepine dosing)

In most cases, when injectable medication is required, a physical hold, restraint or seclusion is also necessary for safety. The choice of intervention (physical hold, restraint, or seclusion) should be made based on the regulatory standards of the Center for Medicare and Medicaid Services and The Joint Commission on Accreditation of Healthcare Organizations, state and local governmental and oversight agencies, and the policies of the hospital, as well as the individual patient history. For example, for a child who has been sexually abused, restraint may be deeply re-traumatizing such that seclusion should be chosen instead; a child who is engaging in self-injury may not be able to be safe in seclusion and thus restraint may be more therapeutic. The physician should ensure that the child is safe, physically comfortable, and medically stable in the restraint or seclusion, and should explain to the child the reason for the restraint or seclusion and that they will be released as soon as they can be safe. Regulation and hospital policy will dictate the allowed duration of restraint or seclusion,

but regardless of regulation the child should be removed from restraint or seclusion as soon as possible. Once the child is released, his safe behaviors should be praised and the physician should debrief with the patient to review what upsets the patient and how such episodes can be avoided. While restraints and seclusion can be traumatizing for children if executed poorly, with appropriate support, explanation and debriefing it can be an opportunity for the youth to better express his feelings and understand the roots of his behavioral dyscontrol.

Assessment of Developmental Factors in Psychiatric Emergencies

Children are not just little adults, physically or psychiatrically. When faced with a child in crisis, the interpretation and significance of the child's symptoms will vary greatly depending on the developmental status of the child. The developmental status includes not only the child's chronological or physical age but also any developmental delays, cognitive disabilities, genetic or metabolic syndromes, or deficits in speech, language or learning ability. The ED clinician may not have time for, or access to, a complete developmental history of the child during the initial evaluation, particularly if the biological parent is not present with the child. Nonetheless, by observing the child closely for physical signs of genetic or developmental syndromes such as fetal alcohol syndrome, and determining whether cognitive and emotional skills are appropriate given the child's physical age, the clinician can determine whether a developmental or cognitive delay may be present.

At the most basic level, development influences acute psychiatric symptoms in that symptoms can have different causes or meaning in children of different age and developmental stage. One important example of this is in the assessment of psychosis. A young child may present to the ED with "hallucinations" that are age-appropriate imaginary friends. An adolescent with significant intellectual disability may have a similar presentation, which would be concerning in a child of that age without intellectual disability but may still be normative in the presence of the disability. Sexual behaviors is another important example. Many children are brought to the ED for "sexual acting out," but the interpretation of sexual behavior and the threshold for what is normative differs greatly for a 5-year old or a 14-year old. A 5-year-old child who exposes his genitals to a peer and asks the peer to do the same is actually engaging in normal, common behavior for a child of that age. A similarly-aged child who touches a peer's or adult's genitals is still within the normal range of behaviors for that age, though the behavior may reflect poor impulse control or emotional immaturity. If a 14-year-old boy is sent to the ED from school after touching a peer's genitals without invitation to do so, this could be a predatory behavior, but it is important to assess whether a developmental or cognitive delay is present, as a 14-year old with autism or

intellectual disability may be closer to the 5-year old's developmental maturity than his 14-year-old body would suggest. Children with autism and attention-deficit/hyperactivity disorder (ADHD) are particularly likely to have disparities between their physical age and cognitive ability and their emotional and developmental maturity. Often children present to the ED with concerning behaviors or symptoms that look like psychosis or other acute illness, but a careful developmental history identifies undiagnosed autism or other developmental delay. Identifying developmental problems that are contributing to the current presentation is important to ensure that the child and family are set up with the appropriate services and treatment.

Assessment of Contextual Family and Social Factors Contributing to Psychiatric Emergencies

Children and adolescents present to the ED with a range of acute psychiatric complaints, from recent-onset psychosis or mania to longstanding obsessive-compulsive disorder (OCD) or autism. An important question for any child in the ED is why now—what has led the child to be brought today, and not yesterday or tomorrow, and what were the precipitating factors that contributed to this. Even when a child presents with an acute or new-onset illness such as psychosis or mania, it did not begin just minutes before their arrival to the ED; rather, someone became aware of the child's symptoms or something happened that was the tipping point for the child or family to seek help. Understanding these contextual factors are important to stabilize the crisis. For example, an adolescent may have been abusing opiates for weeks, but today stole money from his mother and became violent toward her. Stabilizing this crisis will require treatment for the child, but also support for his mother if she is to effectively care for him going forward. A young child sent from school for a tantrum may have longstanding ADHD, but the immediate triggers for his tantrum were frustration with schoolwork related to having been transitioned this week from a special education classroom to a mainstream one. His treatment may involve changes in his ADHD medications, but also increased academic supports. A teenage girl with autism is sent to the ED for head-banging and other self-injurious behaviors, but her symptoms are related not only to her autism but also to her caregiver's recent surgery that led to the girl staying with her aunt for three days. A safe discharge for her will be more dependent on psycho-education and support for her aunt than on additional medication for behavioral control.

The key contextual factors to assess are the child's home and family context; educational setting; and social factors. In the home and family sphere, the child's symptoms may be acutely worsened by shifts in the physical environment (moving into a different apartment or shelter, for example), changes in the family makeup (birth of sibling, remarriage,

divorce or death, or a move into foster care), or conflict or violence in the home. These factors may directly impact the child's mental health or may decrease the parent's ability to cope with symptoms, provide structure or supervision, or manage difficult behaviors. A parent may also be suffering from mental illness or substance abuse, or may be abusive, neglectful, or failing to protect the child from abuse or maltreatment. Understanding the child's home and family context can also highlight key strengths and supports for the child that may be useful in treatment and safety planning.

Multiple factors in a child's educational setting can precipitate an ED visit. Many children presenting to the ED are sent from school, often with a mandate for "psychiatric clearance" before the child will be allowed to return. When a child is sent from school, it is important to understand both the general academic context and learning needs of the child, to assess for unmet needs or mismatch, and also the immediate triggers for the child being sent to the ED. For example, for the child with ADHD who has a tantrum after being transitioned from special education to mainstream classes, it will be useful to know as well that the tantrum occurred in math, his worst subject, after he was being teased by peers for being a "retard." This child will have a different treatment plan than a similar boy with ADHD who has a meltdown during a fire drill, when his family recently became homeless after a fire destroyed their apartment building. Each of these boys can be discharged home but are likely to return to the ED with a similar outburst within days or weeks if the contextual factors are not addressed.

Social factors are also important to assess, particularly for adolescents. A child's social milieu can be a crucial support in a crisis, or a trigger for acute pathology. Self-injurious behavior, substance use, risky sexual behaviors, and suicidality can have strong peer effects among adolescents and pre-teens. Conflicts within friendships, romantic breakups, bullying and cyberbullying can precipitate suicidal thoughts and serious violence. Relationship violence and date rape is also unfortunately common among adolescents. Social isolation can be a marker for serious psychiatric illness, be it chronic (in the case of autistic or otherwise "odd" youth) or acute (for depressed, prodromal, or psychotic adolescents).

These contextual factors can be difficult to assess given the time limitations of the ED assessment. The ED clinician is unlikely to go in depth in discussion of each of these areas, but a brief inquiry into home/family, academic, and social life can quickly identify potential problems into which further investigation is needed. As with the other areas of the interview, the clinician's approach should be systematic but flexible, covering all important areas of functioning but focusing the interview based on the presenting concern.

SPECIFIC PSYCHIATRIC EMERGENCIES IN CHILDREN AND ADOLESCENTS

Suicidality and Self-Harm

Assessment. As noted above, suicide is a leading cause of death in adolescents, and is one of the most common psychiatric presentation of youth to the ED. When a child or adolescent presents to the ED after a suicide attempt, medical stabilization is the foremost step. After this, the ED clinician must determine the patient's suicide risk and whether inpatient hospitalization (including potentially 1:1 observation for safety) is required. Assessing suicide risk will include identifying risk factors such as prior suicide attempts or ideation; experience of hopelessness, anhedonia, or command auditory hallucinations to harm self; substance use; prior psychiatric diagnosis (depression, bipolar disorder, substance use and conduct disorders confer particular suicide risk); family history of suicide; and access to firearms and other means of self-harm. A careful assessment of suicidal thinking and behavior is also crucial. The ED clinician should ask about the ideation and planning leading up to the attempt, including the anticipated lethality of the attempt, if the child thought someone would discover or rescue them, if a suicide note or other preparation for death was made, and if the child experienced regret or ambivalence after the attempt and immediately called for help, or took steps to ensure no one would find out or be able to help him. The clinician should also assess current active or passive suicidal thoughts, plans, intent, future orientation, and reasons for living, though with recognition that youth may be motivated to minimize their current suicidal thoughts or intent for fear of being hospitalized or upsetting family members. The mental status examination can identify mental states that confer acute risk for suicide, which may be more difficult for a child to minimize or hide; these including hopelessness, preoccupation with death, severe depression, command hallucinations, and feelings of worthlessness.

It may not be immediately obvious when a child in the ED has made a suicide attempt. Hanging attempts and self-poisonings may have clear suicidal intent, but other youth may present with self-cutting, overdoses of prescription drugs, threats to stab self or jump from high places, attempts to drown in the bathtub, or other acts where the true suicidal intent is unclear. It is important not to dismiss "dramatic" suicidal gestures as related to personality disorder as might be done in adult patients, as in children and adolescents even suicidal gestures and parasuicidal behaviors often demonstrate real risk. Even if the child is repeatedly self-cutting with no current suicidal intent, the fact of engaging in repeated self-injury is a risk factor for later suicide attempts, and even youth who engage in suicidal behavior with ambivalence or for attention can inadvertently hurt themselves (e.g., youth who take overdoses of acetaminophen without realizing its potential lethality).

Management. The first decision in managing a suicidal patient is determining the need for inpatient hospitalization. If a child is on a high risk in one or more of the following ways, they likely require inpatient hospitalization for safety:

- Serious suicide attempt, with or without endorsing persistent or current suicidal thoughts
- Persistent, intense suicidal thoughts or intention
- Suicidal ideation/behavior in the context of severe depression, command auditory hallucinations to harm himself, hopelessness, or self-blame
- Suicidal ideation/behavior in the context of a profound psychosocial stressor, such as recent sexual trauma or disclosure of abuse, severe bullying, relationship loss
- Suicidal ideation/behavior in the context of poor supports (family rejection/stress, no outpatient treatment)

If the clinician is uncertain whether a child is safe for discharge, it is most prudent to hold the child for reassessment of suicidal ideation and for additional collateral history. If the child is not at high risk, a safety plan should still be completed prior to discharge. This involves helping the child identify triggers for suicidal feelings, coping strategies that the child can use if suicidal thoughts present, who the child can call on for support and what those individuals should do to help the child, and how to access professional help or return to the ED if needed. The safety plan should be shared with the parent or guardian. The child should also be set up with prompt (within a week) outpatient follow up for reassessment of suicide risk and to work to stabilize the psychological and social factors that led to the suicidal thoughts and behavior.

If a child presents with self-injurious behavior and the clinician is certain it is without suicidal intent, a safety plan can still be useful, modified to focus on urges to self-harm rather than on suicidal thoughts. It is also helpful to explore the function of the self-harm, which may include internal reinforcement such as relief of strong negative emotions or feelings of numbness, or external reinforcement such as attention from peers or family or release from social situations or conflicts with which the child is otherwise unable to cope. Youth with anxiety disorders, childhood trauma or neglect, and autism may engage in self-injurious behavior for this latter external reinforcement. The ED clinician should view such behavior not as manipulative but as representing an inability to cope or communicate distress and thus demonstrating need for treatment.

Aggression

Assessment. Youth can present to the ED with aggressive behaviors ranging from overblown tantrums to planned, intentional, even lethal

violence. The assessment of such youth starts with managing the child's behavior in the ED, using the techniques described above to deescalate and contain the youth's behavior with verbal interventions, medication, and restraint or seclusion when indicated. The child's response to these interventions will be telling for both diagnosis and risk assessment. An adolescent who is unable to respond to verbal deescalation and seems to be disorganized or paranoid may be psychotic or intoxicated. A child who is brought by police or school safety officers cursing and kicking but is able to speak calmly to the clinician once the uniformed officers have left, may be by contrast an impulsive child with ADHD or PTSD who became agitated due to teasing at school but with no true intent to harm anyone.

Once the initial agitation is addressed, the clinician must attempt to understand the aggressive behavior. Understanding from the child's perspective the triggers for their aggression can make a frightening experience into a therapeutic opportunity. The clinician should assume that there is always a reason for the aggressive behavior, even if the child cannot immediately articulate it. The reason may not be obvious to the clinician or even to the child at first, but empathic and patient questioning can uncover the triggers, both external (such as teasing or other provocative behavior from peers, being told "no," a disruption of routine, being touched by staff or police) or internal (frustration, anxiety, paranoia, command hallucinations, trauma flashbacks). Aggressive acts can be reactive (in response to a trigger or perceived insult) or instrumental (planned with intent to harm or to use aggression to get something). Reactive aggression may be triggered by events that appear benign to adults, like a school staff putting a hand on the child's shoulder to calm her, but which feel intrusive or threatening to the child due to their background and history. Children with autism, intellectual disability, and learning disabilities may become aggressive when frustrated because they lack the verbal skills to solve the problem and express themselves, or because they do not know any other way to ask for what they want.

Management. Once the initial agitation is addressed, the clinician's job is to determine whether the child's risk of future violence is sufficient to require inpatient admission, or if the child can be safely discharged. Risk for future violence depends in part on the nature of the violence, and on the presence or absence of underlying psychiatric illness. Youth with a range of psychiatric illness can be prone to reactive aggression, and the treatment for each must be targeted to the illness or symptoms present. Instrumental aggression is less common in mentally ill youth, but can be seen in conduct disorder and substance use. When aggression is reactive, the treatment involves identifying triggers, teaching skills to cope with those triggers without aggression, and addressing the core symptoms that predispose the youth to aggressive behaviors. When aggression is instrumental, treatment

involves setting clear consequences, often in collaboration with parents and law enforcement, while promoting prosocial behavior with positive reinforcements. The type of aggression and the diagnostic context in which it occurred also helps to determine treatment needs. For example, a girl with ADHD who strikes out impulsively when a peer teases her, or a boy with PTSD who becomes aggressive when told "no" by a teacher who resembles the stepfather that abused him but then feels remorseful, bears a chronically increased risk for future violence due to their psychiatric symptoms but they do not have intent or plan to harm anyone. They can likely be safely discharged, if set up with treatment for their illness and perhaps with some increased supports and monitoring at school. By comparison, a teen who brings a gun to school with intent to harm a girl who romantically rejected him, or who is hoarding knives in his room due to paranoid psychosis, remains at acute risk for violence and requires hospitalization (Table 36.4–4).

Determining whether inpatient hospitalization is indicated depends both on immediacy of risk for further aggression, but also on whether the risk for aggression is due to mental illness and will be helped by hospitalization. There are youth who are engaging in criminal behavior not due to psychiatric illness, and for them inpatient hospitalization may not be helpful; in such cases clinicians should collaborate with the parents and law enforcement to identify other supports to contain the youth's behavior.

Table 36.4–4.
Examples of Psychiatric Illnesses in Children That Can Lead to Aggression

Illness/Symptom	Treatment Focus
ADHD	Address impulsivity
ODD	Set clear limits, reinforce positive behaviors
Psychosis	Treat command hallucinations/paranoia that are triggering aggression
Mania	Stabilize mania to address impulsivity/irritability
Autism or intellectual disability	Promote verbal problem solving and other coping behaviors
Conduct disorder	Set clear consequences for antisocial behaviors and reinforce positive behaviors
Substance use	Increased supervision, motivation enhancement techniques to reduce use

Risky Behaviors and Substance Use

Assessment. Drug use, running away, truancy, risky sexual behaviors including prostitution, gang involvement, and other dangerous behaviors can prompt parents to bring a child to the ED desperate and looking for help. Such behaviors are uncommon in youth and can have serious and even fatal consequences; yet often the child is not immediately "dangerous to self or others." The ED clinician must assess whether the child's

behaviors are age-appropriate limit-testing, a reflection of underlying psychopathology, or a response to problems in the home or failures of supervision, and whether the risk of the behaviors is sufficient that hospitalization is indicated.

Determining safety is, as always, primary; this may involve assessing whether the youth is currently intoxicated with substances or at risk for withdrawal or whether a rape kit or STD testing and prophylaxis are indicated. Some risky behaviors are isolated incidents that the youth will learn from and not repeat. Youth in the ED for such behaviors will usually claim that this is the case, so collateral information is crucial to assess for repeat behaviors, a decline in functioning, or associated psychiatric symptoms. Risky behaviors can be related to a range of psychiatric illnesses, including depression, mania, post-traumatic stress disorder, ADHD, and conduct disorder, and can also be seen in intellectually disabled or autistic youth who may not comprehend the dangerousness of their actions. They can also reflect problems in the child's environment, such as a parent's inability to provide sufficient supervision due to illness or work schedule or the child's reaction to a divorce or other major stressor in the family. In cases of truancy the behavior may be a response to bullying, insufficient academic supports, or an inappropriate classroom placement for a child with special learning needs. Approaching the patient with an empathic and nonjudgmental stance is key to understanding why they are engaging in these behaviors and assessing their motivation to change.

Management. When risky behaviors put the child at immediate risk for harm, or when they are secondary to underlying psychopathology or problems in the environment, further intervention is needed. For a manic youth having promiscuous unprotected sex and abusing drugs, though this youth may not be suicidal or homicidal, he is endangering himself and others and will continue to do so until his mania is treated, so hospitalization is indicated. For an adolescent with intellectual disability who is running away and associating with gang members because his mother is medically ill and unable to supervise him, coordination with child protective services and law enforcement may be required to provide more support or place the child in a residential facility with greater supervision. Bolstering the parent's ability to relate to and engage the child will help them to prevent risky behaviors in the future. Often, the parent is so upset, angry, or anxious about the child's behavior that they are unable to effectively relate to the child, and the ED clinician must help the parent to contain their affect and set reasonable expectations. If the child and parent are unable to agree upon basic rules and behavioral expectations, then outpatient treatment has little chance for success and the ED clinician should coordinate with social work, case management, child protective

services, and any services currently involved with the child to bolster the family's capacity to keep the child safe.

Psychosis

Assessment. Children and adolescents are often sent to the ED from school or from outpatient providers due to concern for psychosis, particularly for command hallucinations to harm self or others. Yet chronic psychotic disorders like schizophrenia are extremely rare in children, instead usually developing insidiously in the late adolescent or young adult years. The ED clinician will often be in the position of discerning whether a child who is sent for "hearing voices" or being "internally preoccupied" is a child with early-onset schizophrenia, in need for inpatient hospitalization, or more likely, that this is a red herring. Young children can report "hearing voices" as a way to express their own thoughts (particularly thoughts that are shameful, uncomfortable, or that might get them in trouble), and older children and adolescents with intellectual disability or autism often do the same. Imaginary friends are also normal in children and again, adolescents with intellectual disability whose cognitive age is much lower than their chronological age may endorse imaginary friends much longer than their neurotypical peers. Children with autism and intellectual disability are also often sent to the ED for being "internally preoccupied" or "responding to internal stimuli," but this may be a symptom of their autism, not of psychosis. Often these children have never been diagnosed with autism or intellectual disability, if the family has not been consistent with medical providers or closely involved with the child's school. Children who have suffered a loss can report hearing or seeing their loved one in the context of grief. Anxiety disorders and OCD can present with thinking that borders on delusional, and post-traumatic flashbacks can look like the child is experiencing hallucinations. Depression and mania can be accompanied with either nihilistic and shameful or grandiose delusions, and at times with hallucinations berating or shaming the patient. An accurate diagnostic assessment is crucial as the treatment will vary greatly depending on the diagnosis and clinical formulation. Collateral information about the child's life, family, functioning, recent stressors, and prior traumas is important, particularly if the child is disorganized or cognitively limited and unable to give coherent answers.

If the clinician determines that the child is truly psychotic, organic causes must be ruled out. A complete physical exam, lab analysis including complete blood count, electrolytes, liver, kidney and thyroid functioning, full toxicology screening, and brain imaging are indicated, as well as other tests (lyme, erythrocyte sedimentation rate, rapid plasma reagin, lumbar puncture for cerebrospinal fluid analysis, etc.) depending on the patient's history and associated symptoms.

Management. Patients who are truly psychotic can be dangerous in the ED as they may act erratically and expose themselves or others to danger. These patients often require a 1:1 sitter for observation and close monitoring. Youth with psychosis can rarely be stabilized in a few hours in the ED and generally require inpatient hospitalization for stabilization. If the psychosis is secondary to substance intoxication, the symptoms may clear when the intoxication subsides, but with the new designer drugs such as synthetic cannabinoids (k2, spice) and substituted cathinones (bath salts), substance-induced psychosis is of much longer duration than with more traditional drugs and may require inpatient hospitalization for safety. If the psychotic symptoms are due to severe depression, mania or OCD, or if the ED clinician lacks the collateral information and time to observe the patient to make a definitive diagnosis, hospitalization may be the best option to keep the child safe and treat him carefully.

If the reported psychotic symptoms are in fact a manifestation of intellectual disability or immaturity, anxiety, grief, trauma, or a normal reaction to stress, the child can likely be discharged with a referral for outpatient treatment and reassurances for the child and family. Families in these cases are often deeply relieved and should be reminded that even though their child is not being diagnosed with schizophrenia, follow up psychiatric care is still important. Depending on the underlying psychopathology, the ED clinician may also consider starting treatment from the ED (for example, medication for anxiety or psychoeducation about trauma and flashbacks).

Youth with Autism and Intellectual Disability

Assessment. Children and adolescents with autism spectrum disorder (ASD) and/or intellectual disability (ID), particularly those with limited or minimal verbal communication skills, are often brought to the ED for behavioral symptoms including aggression, self-injury, inappropriate sexual behaviors, sleep cycle reversal, or running away. These cases can be difficult for the ED clinician, because youth with ASD or ID can be medically and psychiatrically complicated, and also they often have a difficult time in the ED due to the unpredictability and unfamiliarity of the ED, the noise, bright lights, and the need for blood draws and physical exams. Children with ASD or ID often are blocked from admission to inpatient psychiatric units due to their diagnosis, leaving them boarding in the ED for days awaiting an inpatient bed on a specialized unit.

When youth with ASD or ID present to the ED, the first step to keep them safe is to understand what works at home, including preferred toys and caregivers, communication tools (such as picture communication tools and sign language), and soothing techniques. Understanding what typically triggers negative behaviors at home will also help the ED clinician to avoid

outbursts in the ED. When a child with ASD or ID is brought to the ED for problematic behaviors, it is important first to understand if the behavior is longstanding or new. Longstanding behaviors such as self-stimulatory behaviors, impulsive aggression, or head-banging may be related to the underlying ASD/ID and indicate that the community treatments are insufficient or ineffective and should be reevaluated. New behaviors are often triggered by changes in the environment (changes in caregiver or home environment, school vacation, changes in school placement) or physical/medical problems (constipation, dental problems, ear infections, seasonal allergies, pain, fatigue, sedation due to polypharmacy, adverse effects from medication such as akathesia, onset of puberty). More rarely the youth with ASD/ID can have the new onset of a psychiatric disorder such as a mood or anxiety disorder, but environmental changes and medical problems should be ruled out before a new diagnosis is made.

Management. Keeping patients with ASD and ID safe in the ED requires removal of unnecessary stimuli (noise, light, needle sticks), clear communication (simple, concrete, and written or picture-based if possible) of what to expect from the evaluation, attention to physical comfort (pain, fatigue, hunger, constipation), and use of soothing activities or family members when possible. Medication is often used to sedate ASD and ID patients in the ED, however this is dangerous as they may have paradoxical reactions to medications such as diphenhydramine and benzodiazepines yet have little positive effect from antipsychotics. Behavioral interventions, while more complicated to execute, are much more likely to be effective.

Many youth with ASD/ID can be stabilized quickly if the medical or physical etiology of their distress is discovered and resolved. Effective pain control, dental care, management of insomnia and constipation, and resolution of polypharmacy and medication adverse effects can be accomplished in the ED and provide great relief to patients and families. Often when these problems are resolved, families discover that their child who had been increasingly aggressive or difficult to manage for a long time is now calm and comfortable.

As there are few inpatient units that accept youth with ASD or ID (particularly if they are aggressive or engage in self-injurious behaviors), the ED clinician should familiarize herself with community services available for people with intellectual and developmental disabilities. Often home-based and respite services are available to patients who qualify, though families may not know how to access these services.

Child Abuse or Neglect

Assessment. Unfortunately, it is all too common that youth who present to the ED for psychiatric complaints are found to be the victims of child abuse or neglect. Children with psychiatric and developmental

disorders and behavioral disturbances are more likely to be abused and neglected, and abuse and neglect can contribute to the development of a range of psychiatric illnesses. Other risk factors include poverty, single parent households, large family size, social isolation, parental substance use, and personal history of abuse for the parent. Assessing for abuse or maltreatment in the ED is very difficult, because children rarely admit to abuse, due to shame, guilt about betraying the family if the abuser is a family member, or fear of reprisal or removal from home. Clear evidence of abuse is rare. Children can present with medical symptoms such as unexplained physical injuries, signs or symptoms of sexually transmitted diseases, or blood in their underwear, and then the psychiatric clinician consults the pediatric team for further evaluation for abuse. Children who are being abused or maltreated can also present to the ED with a wide range of psychiatric symptoms. Behavioral or externalizing symptoms can manifest as aggression, stealing, hoarding food or eating from the garbage, running away, regression in toileting and other milestones, separation anxiety, or school refusal. Internalizing symptoms can include PTSD and depression but also panic attacks, irritability, inattention, and organizational problems that mimic ADHD, and withdrawal from family, social and academic activities. Adults who know the child has been abused or traumatized in some way often will not connect that maltreatment to the behaviors and symptoms presenting now. And very often, no one but the child and the abuser is aware that any abuse is happening.

The ED clinician must ask in a careful, thorough, nonjudgmental way about any abuse or maltreatment in the home, school, or elsewhere. The child should be interviewed separately from the parent, and should be reassured that he will be safe and protected no matter what he discloses. It is important to interview the child gently, starting with more subtle queries into the child's home life (such as who live in the home, who prepares meals, who helps with homework) and then moving to questions about who sleeps in bed with whom, who gets into arguments and how they are resolved, who yells the most at home, what punishments are for misbehavior, and other contextual questions before asking directly about abuse. Collaboration with the pediatrician if any physical injuries are present can also be helpful as these may be a marker for abuse, or the child may require a genital examination if there is concern for sexual abuse. When child abuse pediatrician consultation services are available, it can be helpful to obtain a specialized interview and physical exam to discover and document evidence of abuse. At the very least, when abuse is suspected, child protective services should be consulted immediately. It is not the ED clinician's job to determine whether abuse has definitely occurred (that is the job of child protective services). But if there is any concern, as a mandated reporter the ED clinician is obligated to call child protection. Parents should be notified that such a call is being made and why; even if

the parent is the perpetrator or failed to prevent abuse, they are often remorseful and looking for help, if the clinician is able to empathize with why the abuse has happened (such as a battered mother unable to prevent the abuse of her children, or an overwhelmed parent struggling to discipline a child with significant behavioral problems). The parent may need help as much as the child does, so maintaining an empathic alliance with the parent will give them and the child the best chance for recovery.

Management. Collaboration with child protective services will help to ensure the child's immediate physical safety, and this may require keeping the child in the hospital until an investigation is done and a safe place is found to which the child can go. The ED clinician can also play an important role in assessing for psychiatric sequelae of maltreatment, providing psychoeducation to the child and family (or foster family), and ensuring that the child receives appropriate psychiatric follow up care. If the child presented to the ED with behavioral problems, helping the family or caregiver to identify how these behaviors may be triggered by or related to the trauma or abuse can help avoid such behaviors (and repeat trips to the ED) in the future. Family therapy can be a very helpful referral in these cases. Anticipatory guidance can also be helpful for both the child and caregiver to understand what types of symptoms may come up in the future and how to access help if needed.

Eating Disorders

Assessment. Eating disorders are a source of significant morbidity and mortality in children and adolescents. When an eating disorder is discovered by a parent or teacher, the adult will often send the child to the ED for medical and psychiatric evaluation due to concern for the medical consequences of eating disorders as well as their association with suicidality and self-harm. The ED clinician may also discover an undiagnosed and unrecognized eating disorder in a child or adolescent who presents to the ED with dizziness, nausea, syncope, suicidal ideation, or self-injury. Younger children and boys may present to the ED with unexplained weight loss, and often no one has asked them about restricting, purging, or excessive exercise because clinicians are unused to seeing eating disorders in prepubertal patients or in males.

Table 36.4–5.
American Academy of Pediatrics Criteria for Medical Hospitalization for Eating Disorders

Anorexia nervosa
Weight ≤75% ideal body weight, or ongoing weight loss despite intensive management
Refusal to eat
Body fat <10%

	Orthostatic changes in pulse increase >20 bpm or blood pressure decrease >20 mm Hg
	Pulse <50
	Temperature < 96.0°F
	Arrhythmia
Bulimia Nervosa	
	Syncope
	Serum potassium concentration <3.2 mmol/L
	Serum chloride concentration <88 mmol/L
	Esophageal tears
	Cardiac arrhythmias including prolonged QTc
	Hypothermia
	Suicide risk
	Intractable vomiting
	Hematemesis
	Failure to respond to outpatient treatment

Management. Children and adolescents with eating disorders may require psychiatric hospitalization for safety due to suicidality or self-harm, or they may require medical admission for medical stabilization. Collaboration with pediatrics is crucial in all cases to determine the treatment plan and need for hospitalization. Assisting the family to identify resources for support and education is also important, as effective treatment of eating disorders is a long-term process that requires close family involvement for greatest success (Table 36.4–5).

Psychosocial Crisis

Assessment. Children are profoundly affected by a parent's state of mind, mental health, and relationships. Parental mental illness, divorce or severe conflict between parents, and in extreme cases medical child abuse (formerly called Munchausen's by proxy) can manifest with emotional and behavioral problems in the child. These stressors put the child in a desperate situation, and for a child with underlying psychiatric or behavioral symptoms, these family crises can be catastrophic. These situations can be difficult to assess. The parent is often unable to objectively reflect the stress on the child or even on the parent's own internal life and struggles. The child may lack the emotional maturity to describe clearly the problems at home, feel guilty, ashamed, protective of the parent, or afraid of being separated from the family. "Listening between the lines" is important, and when the given history does not make sense it can be an indication that hidden stressors at home may be at the root of the child's presenting problem. The child and parents should each be interviewed separately, and the clinician should approach the parent empathically and making it clear that their purpose is to help the child, not to be intrusive, judge the parents, or report them. If parents give deeply discrepant reports or perspectives on the child's symptoms, blame the other parent or the child, or belittle or berate the other parent or child, this

suggests significant conflict in the family that may be triggering the child's symptoms or behaviors or making it impossible for the family to effectively manage those symptoms or behaviors. When speaking with the child, vague or inconsistent statements that seem odd or unusual to the interviewer (such as, a child who states that the parent is sleeping most of the day and cannot be awoken) may serve as "red flags," and reassurance and specific questioning to the child may help them share the problems at home.

Medical child abuse can manifest with exaggeration or fabrication of a child's psychiatric symptoms. This is very rare and difficult to diagnose, but can manifest in the ED with hospital shopping, strange recurrences of symptoms despite good treatment, requests for specific medications, tests or procedures, and refusal to consent for the ED clinician to contact outpatient providers or other key collateral sources. The ED clinician should be cognizant that medical child abuse can present with psychiatric symptoms and should utilize tools for collateral information such as state prescription databases, pharmacy and insurance records, if such an etiology is suspected.

Management. Safety is, as always, the first priority; if a parent is severely mentally ill, the conflict in the home so extreme, or if the clinician is very suspicious that medical child abuse is occurring, then child protective services should be notified. The ED clinician may in extreme cases also need to refer the parent for their own emergency assessment. In less severe cases, with an empathic and strong rapport with the parents, the ED clinician may be able to help the parent to understand their role in the child's symptoms and their need for help and treatment. Getting the parent into individual or family treatment, and helping the child to voice and coping with the stress at home, will reduce the child's distress and the strain on the family.

LEGAL CONSIDERATIONS AND CONFIDENTIALITY

Clinicians working with youth in emergency psychiatric settings should be familiar with the laws and institutional regulations that govern emergency psychiatric treatment and hospitalization, and also those that apply to evaluation and treatment of minors. Many states allow for emergency psychiatric evaluation and treatment of a minor without a parent's consent, and several allow adolescents to access outpatient psychiatric care without the parent's knowledge or consent. If a child requires hospitalization, each state has statutes that dictate a standard for psychiatric hospitalization, generally differentiating between voluntary, involuntary, and emergency or diagnostic admissions. When working with minors, these laws generally dictate that a parent or guardian can hospitalize a child or adolescent under a "voluntary" admission, even without the minor's assent, and often provide a standard by which the decision to hospitalize a child or

adolescent can be made without the parent's consent. Clinicians working in the ED should know these statutes as they apply in their area, as well as the local and institutional policies for consent and hospitalization.

Clinicians should also be familiar about mandated reporting laws for cases of suspected physical or sexual abuse or other maltreatment. If a patient reports serious intent to harm another person, the clinician has a responsibility to warn this individual that supersedes the patient's right to confidentiality. Navigating the child's confidentiality vis-à-vis their parents can also be complicated. If the child is in imminent danger or a danger to others if the parent is not notified, the clinician should tell the parent, but first should attempt to understand the child's desire for secrecy. Often a child fears the parent will react with reprisal, anger, or distress, and may be able to share with the parent with the clinician's support and guidance.

CONCLUSION

Good ED care for children in psychiatric crisis means an accurate, developmentally-informed diagnosis and formulation, stabilization of the crisis, and a treatment plan that helps the child and family to overcome stigma and systems barriers to access the right care. The practical challenges of connecting a patient to appropriate treatment should not be underestimated either. If a patient requires inpatient hospitalization it can be difficult to find an open bed on a psychiatric unit for children and adolescents, but treatment on an adult unit or "boarding" in pediatrics should be avoided as both can be counter-therapeutic. If the child will be waiting in the ED for several days before an appropriate inpatient bed is available, the ED provider should begin treatment (therapy or medication) to the degree that is possible in the ED.

When a patient is safe for discharge, the ED clinician should make every effort to connect the patient to an outpatient provider, not just give the parent a phone number for a provider or for their insurance panel. Connecting to outpatient care can be a very difficult and lengthy process and some parents cannot navigate the system due to insurance issues or the parent's own cognitive or language difficulties. Providing the patient with an actual appointment for follow up has been shown to decrease future dangerous behaviors and so is crucial for safe discharge of high-risk patients. ED clinicians should also familiarize themselves with case management services; child welfare services; community supports such as family resource centers; social services within the school, shelter and family court systems; and respite and crisis services available through the state or city mental health, child welfare, or developmental disabilities services providers. These additional services can be crucial for avoiding a repeat presentation to the ED, particularly if psychosocial or family factors have influenced the current symptoms.

REFERENCES

Bridge JA, Marcus SC, Olfson M. Outpatient care of young people after emergency treatment of deliberate self-harm. *J Am Acad Child Adolesc Psychiatry*. 2012;51:213.

Case SD, Case BG, Olfson M, Linakis JG, Laskis EM. Length of stay of pediatric mental health emergency department visits in the United States. *J Am Acad Child Adolesc Psychiatry*. 2011;50:1110.

Chun T, Katz E, Duffy S, Gerson R. Challenges of managing pediatric mental health crises in the emergency department. *Child Adol Psych Clin N Am*. 2015;24:21.

Chun TH, Mace SE, Katz ER, et al. Evaluation and management of children and adolescents with acute mental health or behavioral problems. Part I: common clinical challenges of patients with mental health and/or behavioral emergencies. *Pediatrics*. 2016;138(3).

Chun TH, Mace SE, Katz ER, et al. Evaluation and management of children with acute mental health or behavioral problems. Part II: Recognition of clinically challenging mental health related conditions presenting with medical or uncertain symptoms. *Pediatrics*. 2016;138(3).

Dolan MA, Fein JA, Committee on Pediatric Emergency Medicine. Pediatric and adolescent mental health emergencies in the emergency medical service system. *Pediatrics*. 2011;127:e1356.

Grupp-Phelan J, Harman JS, Kelleher K. Trends in mental health and chronic condition visits by children presenting for care at U.S. emergency departments. *Public Health Rep*. 2007;122:55.

Haddad F, Gerson R, eds. *Helping Kids in Crisis: Managing Psychiatric Emergencies in Children and Adolescents*. Arlington, VA: American Psychiatric Publishing; 2014.

Jones L, Bellis MA, Wood S, et al. Prevalence and risk of violence against children with disabilities: a systematic review and meta-analysis of observational studies. *Lancet*. 2012;380(9845):899–907.

Mapelli E, Black T, Doan Q. Trends in pediatric emergency department utilization for mental health-related visits. *J Pediatrics*. 2015;167:905–910.

Regan K. Loosening restraints. *Nurs Spectr*. 2003;7(7):5–8.

Torio CM, Encinosa W, Berdahl T, McCormick MC, Simpson LA. Annual report on health care for children and youth in the United States: National estimates of cost, utilization and expenditures for children with mental health conditions. *Academic Pediatrics*. 2015;15(1):19–35.

Wharff EA, Ginnis KM, Ross AM. Family-based crisis intervention with suicidal adolescents in the emergency room. *Soc Work*. 2012;57:133.

37

Genetics in Child Psychiatry

ERIKA L. NURMI, M.D., PH.D., MEREDITH A. HANNAN, M.D., PH.D., AND JAMES T. MCCRACKEN, M.D.

INTRODUCTION

Before the completion of the Human Genome Project (HGP) in 2003, the role of genetics in medicine and psychiatry primarily focused on the identification of single-gene mutations that played causative roles in relatively rare Mendelian diseases such as cystic fibrosis, neurofibromatosis, and Tay–Sachs disease. However, classic genetic approaches for Mendelian disease failed when applied to the majority of common human diseases, including neuropsychiatric disorders. Common disorders, while often highly heritable, are complex in nature: caused by the presumed interaction of multiple genetic and environmental influences. While the HGP provided scientists with the 3 billion bases of primary DNA sequence comprising the human genome, assembling the puzzle pieces of complex genetic disease requires an understanding of human variation, DNA regulation, and other complexities that are still incomplete. Due to the efforts of large consortium projects such as ENCODE, the 1000 Genomes Project, and the International Hap Map Project, and fueled by continuing advances in technology, the amount of available genomic information has exploded over the past decade. The challenge remains to leverage these techniques and available data to advance the current understanding of genetic influences in psychiatric illness. Such knowledge, once in hand, should provide opportunities for improved diagnosis, treatment, and prevention. Despite the complexity of childhood-onset neuropsychiatric disorders, definite progress has been made which this chapter will review.

Psychiatric disorders are almost exclusively complex genetic disorders. Heritability and twin studies have consistently demonstrated that genetics play an important, and often major etiologic role in the development of psychiatric disease. Yet identifying susceptibility genes underlying psychiatric illness has proven more vexing than for complex diseases in other areas of medicine (e.g., hypertension, inflammatory bowel disease, diabetes, cancer). One barrier to these efforts is the phenotypic complexity of psychiatric illness, and the likelihood that most categories of psychiatric illness as classified in the DSM-5 are heterogeneous in nature. New research directions, such as the current NIH Research Domain Criteria

(RDoC), attempt to characterize psychiatric illness on dimensional and more objective neurobiological measures. Improved phenotyping is one approach hoped to lead to progress in understanding the genetics of complex illness. Interestingly, most childhood-onset neuropsychiatric disorders appear to be more highly heritable than similar adult-onset phenotypes, suggesting that research to identify genetic influences might be more fruitful in pediatric samples. Likewise, studies of genetic influences on treatment and prognosis may be more revealing in younger, more treatment-naive subjects. As more adult-onset psychiatric disorders are increasingly being conceptualized as neurodevelopmental in nature, with identifiable prodromal and subclinical childhood manifestations, the study of such early phenotypes may represent genetically subgroups more tractable for etiologic investigations.

Genetic information is also being applied in other areas to advance medicine and psychiatry. Knowledge of genetic background for treatment selection and the impact of genetic variation on drug response is referred to as *pharmacogenomics*. *Pharmacogenetics,* now often used synonymously, originally began with genetic determinants of drug metabolism and distribution. Such fields pursue the promise of "precision medicine"—the personalization of treatment based on genetics and other individual factors, which in other areas of medicine has yielded important, widely applied improvements in clinical care. Such efforts are also increasingly being examined in child psychiatry.

With the continued advances in technological, computational, and statistical methods in genetics, it is certain that genetic information will become even more important in understanding and diagnosing psychopathology, and in directing treatment recommendations. Psychiatrists need to appreciate the nature of the genetic discoveries as they emerge and to understand how to interpret and apply them in their practice; to do so requires basic knowledge of the research methods applied and the strengths and limitations of these findings. An update on current findings in the genetics of childhood-onset psychopathology forms a needed background. The goal of this chapter is to highlight progress in genetic research on the etiology of pediatric-onset mental illness as it points to future opportunities to inform clinical practice and treatment practices.

THE GENETIC BASIS OF PSYCHIATRIC DISEASE

DNA Variation

In each individual human genome there are millions of DNA variants, known as *polymorphisms*. Most changes appear to be neutral with respect to human illness but likely are critical for the diversity, robustness, and evolution of the human species. The earliest appreciated disease-associated

genetic variations were cytogenetically visible interchromosomal translocations that disrupt or delete genes or regulatory elements, thus contributing to disease. More recently, large *segmental duplications* (SDs) or *copy number variations* (CNVs), involving extensive duplicated, repeated, inverted, or deleted segments of DNA larger than 1 kilobase (KB) in size, are observed commonly in affected and unaffected individuals. When certain CNVs are observed repeatedly in a population of individuals with a disease, the CNVs can point to disease-causing loci in the genome, such as those CNVs producing the Prader–Willi syndrome (15q11–q13 paternally derived deletions), some cases of syndromic autism (15q11.2–q13 maternal duplications), and William syndrome (7q11.23 deletion). Whereas gross chromosomal derangements typically affect multiple genes and produce rare but characteristic syndromes, subtler variations are more likely to be involved in complex disease susceptibility. CNVs are likely to contribute to disease through changes in gene dosage or disruption of genes or regulatory elements in the genome.

Until recently, little attention had been paid to the role of CNVs in complex psychiatric illness. Surprisingly, CNVs were discovered to be common (about 2 to 12 percent of the genome) and often polymorphic, with individuals carrying from zero to many copies of a given segment of DNA. Small DNA stretches (less than 1 KB) that are deleted or inserted, collectively referred to as *"indels,"* are even more frequent, with more than 400,000 annotated indels in the human genome. CNVs are now appreciated to represent another genetic mechanism contributing to complex disease risk. When compared to typically developing controls, total CNV burden is elevated in individuals with autism and schizophrenia. Although not all CNVs are pathogenic, disease-associated CNVs are larger and more likely to encompass gene-rich regions of the genome.

The simplest and most common type of individual genetic variation is the *single-nucleotide polymorphism* (SNP). SNPs are changes of a single DNA base pair (e.g., adenine [A] to guanine [G]) that occur in approximately one per every 1,200 bases throughout the genome. Public databases catalog millions of SNPs and often include allele frequency and genotyping information because they are widely employed in genetic studies.

DNA Variation and Contribution to Disease Risk. Smaller indels and SNPs can influence disease through more targeted changes in individual genes. SNPs can directly alter protein structure through changes that alter amino acid sequence (nonsynonymous changes) in the gene product or distort start or stop signals. Even synonymous (silent) polymorphisms that do not alter amino acid–coding potential can influence RNA stability, splicing, targeting, editing, and efficacy of translation—many of the effects of these synonymous variants are unknown. Although noncoding changes

are often overlooked due to presumed lack of function, they can exert powerful effects on the regulation of gene expression and transcript processing, as illustrated by the noncoding risk alleles described in macular degeneration and Crohn disease.

Most investigators believe that complex disease susceptibility is driven by a combination of common alleles, each of which carries a small disease risk ("common variants for common disease" model). It is presumed that such alleles arose on an ancestral chromosome long ago and are now common in the population. Recent findings in ASD research have identified evidence that rare, highly penetrant (if present, highly likely to result in disease), loss of function (inactivating) mutations may produce similar disease phenotypes ("rare variants for common disease" model). In this scenario, rare alleles contribute to genetic risk within the context of genetic heterogeneity. Because of their rarity, such causative mutations are challenging to identify. Strides are being made on this front, however, by the introduction of next-generation sequencing (NGS) technology, in which entire genomes of individuals can be completely genotyped, and then scanned for genes that are disrupted and associated with a disease phenotype.

DNA Variation as Markers of Risk Alleles. Neutral polymorphisms, although not directly contributing to the disease phenotype, can serve as markers for functional risk alleles due to the fact that polymorphisms are not inherited independently but in large DNA blocks, a phenomenon known as *linkage disequilibrium* (LD). LD can persist for substantial distances across a chromosome and for numerous generations but decay as a function of the rate of recombination between homologous chromosomes. Therefore, it is not always necessary initially for an investigator to genotype a risk allele directly to find association with disease but merely to genotype one of often many markers "linked" with the risk allele. Although an associated marker may have no direct role in disease etiology, it is present on an ancestral portion of DNA that also contains the disease locus and can thereby pinpoint the genomic location of the disease allele.

Finding Risk Genes

The basic strategy behind studies aimed at disease gene discovery is to identify regions of the genome that are shared between affected individuals; the genes that underlie the disease must reside within these shared regions. The genetic tools best suited to risk mapping depends on the underlying genetic architecture of a given disorder. Common approaches applied to the search for psychiatric risk genes are discussed below, highlighting successful examples.

Linkage Analysis. Linkage analysis is a powerful tool for localizing

genes of single or major effect within a family of related individuals. Historically, hundreds of *microsatellite markers* (polymorphic short tandem repeats) were genotyped across the genome in large families or in many multiplex families in order to identify chromosomal regions that cosegregate with the disease phenotype. Broad loci spanning many megabases could be identified, and explored for candidate genes. With the advent of SNP microarrays, the genomic regions of interest can be further narrowed, and risk variants of large effect can be readily identified. Ultimately, the identified genes responsible for single-gene versions of common disorders (e.g., tuberous sclerosis in autism) may be distinct from common pathways to disease in the population at large. However, this approach may help to identify gene pathways that may be involved in a specific disease, and that information can help to guide treatment and understanding of the overall disease process.

Genome-wide linkage scans have been conducted with the aim of identifying gene loci for many psychiatric disorders. Linkage scans that have focused on a child proband are limited and include attention-deficit/hyperactivity disorder (ADHD), autism spectrum disorder (ASD), Tourette syndrome (TS), obsessive compulsive disorder (OCD), and the eating disorders, many of which share suggestive signals between two independent groups (Table 37–1). Nevertheless, despite more than a decade of linkage studies, relatively few confirmed findings have emerged using this approach, suggesting that the heterogeneity of these disorders is even greater than originally presumed and that even larger study samples, more precise phenotypes, and novel approaches are necessary to identify disorder-linked regions harboring risk genes.

Table 37–1.
Summary of Linkage Data

Disorder	Chromosomal Regions Supported
Attention-deficit/hyperactivity disorder	5p, 6q, 9q, 10q, 12q, 16p, 17p
Autism spectrum disorders (15 studies)	1q, 2q, 3q, 6q, 7q, 11p, 16p, 17q
Schizophrenia	1q, 2q, 6p, 8p, 13q, 18p, 22q
Major depressive disorder	1q, 4q, 7q, 8p, 11q, 12q, 13q, 15q
Obsessive compulsive disorder	1q, 3q, 6q, 7p, 9p, 15q
Tourette syndrome	2p, 4q, 8p, 19p
Bulimia nervosa	10p
Anorexia nervosa	1q, 2p, 13q

Association Mapping. Association mapping identifies regions of the genome shared between affected (but not necessarily related) individuals resulting from a distant ancestral founder who transmitted a risk allele to subsequent generations. Genotype data at selected markers are collected from cases and controls to compare allele frequencies between these

groups. A *family-based association design* (family-based association test, transmission disequilibrium test) examines the frequency of allele transmission to affected siblings, eliminating the problem of *population stratification* (spurious differences due to dissimilar genetic backgrounds in cases vs. controls). Alleles that are found more frequently in cases in a case-control design or preferentially transmitted to affected siblings in a family design implicate specific genes or genomic regions that may be involved in disease. Compared to linkage, association persists across much smaller distances, necessitating greater marker density but also enabling high-precision mapping.

Commonly, genetic studies of association rely on SNP markers because they are densely and regularly dispersed throughout the genome, robust to mutation, easily assayed by automated high-throughput genotyping, and more likely involved directly in disease susceptibility. However, their power is limited by reduced marker heterozygosity and informativeness. This drawback can be overcome by combining SNPs into *haplotypes*—discrete chromosomal fragments defined by a collection of individual SNP alleles. Whereas individual SNPs are biallelic, haplotypes are multiallelic and thus deliver enhanced genetic information by more accurately reflecting genotypic identity at a given locus. A disease-associated haplotype might represent an ancestral chromosome fragment harboring a disease allele.

Evolutionary studies of the genome sequence and structure have revealed that the genome exists as blocks of low recombination (haplotype blocks) flanked by recombination "hot spots." The haplotype alleles that a given individual carries can be identified by genotyping only a small number of key variants ("tag SNPs," or tSNPs) contained within the haplotype. It has been proposed that only five to seven SNPs could capture all the genetic information within an average gene, effectively reducing the number of genotypes that must be measured to study association at a gene locus.

CANDIDATE GENE STUDIES. Typically, studies of allelic association have employed markers residing within disease-relevant candidate genes. The choice of candidates is informed by existing knowledge of pathways or molecules thought to be involved in a specific disease. For example, because monoamine neurotransmission has long been thought to play a role in depression, common choices of candidate genes have included monoamine receptors and transporters. Once a candidate gene is selected, markers mapping within the transcriptional unit are genotyped in case and control populations to screen for deviations in allele frequency. The greatest limitation of candidate gene studies is the rudimentary understanding of disease mechanisms in psychiatry. Which are the correct candidates to test? Researchers have turned to genome-wide association studies (GWAS), to eliminate this bias and provide a comprehensive,

hypothesis-free screen of all genes.

Numerous candidate gene studies have been performed for psychiatric disorders, but very few have been replicated or supported by functional data. A summary of candidate genes most strongly supported as risk factors for specific disorders is shown in Table 37–2. Candidate gene studies are inherently poorly powered due to the low prior probability that any given gene is associated with disease, and therefore the true threshold for significance is very high. Future studies in psychiatric genetics will leverage the genome-wide association approach to thoroughly scan the genome for risk genes.

Table 37–2.
Summary of Replicated Candidate Gene Associations with Childhood Onset Psychiatric Disorders

Disorder	Candidate Genes
Attention-deficit/hyperactivity disorder	SLC6A3 (DAT1), DRD4, DRD5, 5HT1B, SNAP25, DBH, SLC6A4 (5HTT)
Autism spectrum disorders	NLGN3, NLGN4, CNTNAP2, NRXN1, MET, OXTR PDE4, SHANK3, PTEN, GABRB3
Schizophrenia	NRG1, DTNBP1, G72
Major depressive disorder	SLC6A4 (5HTT), BDNF, AVPR1B
Obsessive compulsive disorder	SLC1A1, 5HTT, HTR2A
Tourette syndrome	SLITRK1, DRD4, DRD2
Bulimia nervosa	GHRL
Anorexia nervosa	OPRD1, HTR1D, COMT, DRD2, SLC6A4 (5HTT)

GENOME-WIDE ASSOCIATION STUDIES. Over the last decade, microarray ("gene chip") technology has made ambitious GWAS possible; chips are available that can assay millions of SNPs simultaneously. Evaluating the approximately 10 million SNPs in the human genome is not a practical task and would be fraught with statistical complications. However, the number of SNP genotypes required can be reduced to a more manageable number by typing only tag SNPs within defined haplotype blocks. Currently, the consensus p value for significance in GWAS studies is set at 5×10^{-8}, correcting for the 50,000 predicted false positives that occur when a million tests are performed. While many psychiatric GWAS fail to meet this stringent level of significance, several large consortia have recently been successful by recruiting tens of thousands of participants (Table 37–3).

The most successful GWAS to date have examined schizophrenia in adults and provide a proof-of-concept that large enough sample sizes are indeed able to identify common risk variants of modest effect in psychiatric disease. Due to the cooperation and sharing between several multisite initiatives, tens of thousands of individuals with schizophrenia were analyzed in these studies. The largest psychiatric GWAS to date was

performed by the Schizophrenia Working Group of the Psychiatric Genetics Consortium (PGC), who conducted a meta-analysis of 49 independent datasets totaling 34,241 cases and 45,604 controls, as well as 1,235 parent affected–offspring trios. This well-powered meta-analysis produced 128 statistically independent associations between individual SNPs and schizophrenia diagnosis at 108 loci (including 25 previously reported associations and 83 novel associations). Of interest, at least 75 percent of these loci contained one or more protein-coding genes, many of which were involved in synaptic plasticity and dopamine (DRD2), glutamate (GR1A1, GRIN2A, GRM3), and calcium (CACNAC1C, CACNA11, CACNB2) signaling pathways.

To date, GWAS for other neuropsychiatric illnesses have resulted in limited success, likely due to insufficient sample sizes and substantial disease heterogeneity (especially in the case of major depression and anxiety disorders). Even so, progress has been made on a number of other psychiatric illnesses such as bipolar disorder and major depression, and with further large-scale collaborative efforts, findings are hoped to approximate those seen in schizophrenia GWAS. A total of 9 GWAS of bipolar disorder have yielded at least 17 independent genetic loci that have been replicated and reliably reproduced as risk alleles. In the case of more heterogeneous disorders, such as depression, mood disorders, panic disorders, and posttraumatic stress disorder, findings have been less robust; but the fact that gene loci that cross the significance threshold have been discovered lends credence to the genetic aspects of these disorders. Two large GWAS for major depression have identified four independent loci conferring risk for depression and response to antidepressant treatment, and with increased sample size this list is likely to grow. One GWAS found a single locus associated with mood disorders in general. A GWAS in panic disorder noted two loci associated with panic attacks, and a GWAS for posttraumatic stress disorder also unearthed two independent loci that conferred risk for the disorder.

Table 37–3.
Summary of Neurodevelopmental Disorder GWAS with SNPs That Reach Significance of $p < 5 \times 10^{-8}$

Psychiatric Disorder	Reference	Strongest SNP-Risk Allele	Genomic Region	Mapped Genes	*p* Value
Attention-deficit/hyperactivity disorder	Yang (2013)	rs1682874	2q37.1	NCL, ZBTB80SP2	6×10^{-9}
	Lesch (2008)	rs864643	3p22.1	MOBP	1×10^{-8}
		rs7995215	13q31.3	GPC6	1×10^{-8}
Autism spectrum disorder	Xia (2013)	rs6537825-A	1p13.2	TRIM33	3×10^{-8}
		rs926938	1p13.2	AMPD1	3×10^{-8}
	Anney (2010)	rs4141463	20p12.1	MACROD2	4×10^{-8}

	Wang (2009)	rs4307059	5p14.1	MSNP1, CDH9	2×10^{-10}
Bipolar disorder	Muhlesien (2014)	rs10994415-C	10q21.2	ANK3	7×10^{-11}
		rs12290811-A	11q14.1	TENM4	1×10^{-9}
		rs6550435	3p22.2	TRANK1	2×10^{-8}
	Winham (2013)	rs72977016	3q23	LOC102724068	7×10^{-9}
	Chen (2011)	rs9834970-C	3p22.2	HSPD1P6, TRANK1	1×10^{-12}
		rs2271892-G	2q11.2	LMAN2L	2×10^{-10}
		rs7618915-G	3p21.2	PPM1M	3×10^{-10}
		rs4948418-T	10q21.2	ANK3	4×10^{-10}
		rs4650608-T	1p31.1	RPL23P3, ELTD1	8×10^{-9}
	Cichon (2011)	rs1064395-A	19p13.1	NCAN	2×10^{-9}
	Jiang (2011)	rs420259	16p12.2	PALB2	9×10^{-9}
	Kerner (2011)	rs1039002	6q27	PDE10A	2×10^{-8}
		rs2727943	3p26.3	RPL23AP39	3×10^{-8}
	Sklar (2011)	rs4765913-A	12p13.33	CACNA1C	2×10^{-8}
		rs12576775-G	11q14.1	TENM4	4×10^{-8}
	Ferreira (2008)	rs10994336-T	10q21.2	ANK3	9×10^{-9}
	Baum (2007)	rs1012053-A	13q14.11	DGKH	2×10^{-8}
Conduct disorder	Dick (2010)	rs16891867-G	4p15.3	C1QTNF7	3×10^{-9}
		rs7950811-A	11q14.3	RPS3AP42, MTNR1B	1×10^{-8}
		rs11838918-C	13q31.1	LINC00331, HSPD1P8	1×10^{-8}
Schizophrenia	Ripke (2014)	rs115329265-A	6p22.1	TRNAI35	3×10^{-31}
	(108 associations, top5 noted)	rs1702294-C	1p21.3	MIR137HG	3×10^{-19}
		rs11191419-T	10q24.32	C10orf32-ASMT	6×10^{-19}
		rs2007044-G	12p13.33	CACNA1C	3×10^{-18}
		rs4129585-A	8q24.3	TSNARE1	2×10^{-15}
	Ripke (2013) (top 5 noted)	rs114002140-A	6p21.32	HLA-DRB9	9×10^{-14}
		rs7085104-A	10q24.32	C10orf32, AS3MT	4×10^{-13}
		rs6461049-T	7p22.3	MAD1L1	6×10^{-13}
		rs1198588-T	1p21.3	MIR137	2×10^{-12}
		rs1006737-A	12p13.33	CACNA1C	5×10^{-12}
	Aberg (2013)	rs1261117-T	18q21.2	TCF4	$3 \times 10^{-}$

					10
	Betcheva (2012)	rs7527939-C	1q32.2	HHAT	6×10^{-9}
	ISGC & WTCCC (2012)	rs2523722-G	6p22.1	TRIM26	1×10^{-16}
	Bergen (2012)	rs17693963	6p22.1	TRNAI25	3×10^{-11}
		rs11191580	10q24.33	NT5C2	2×10^{-9}
		rs12666575	7p22.3	MAD1L1	2×10^{-9}
	Shi (2011)	rs16887244	8p11.23	LSM1	1×10^{-10}
	Yue (2011)	rs1635	6p22.1	NKAPL, ZKSCAN4	7×10^{-12}
		rs11038167	11p11.2	TSPAN18	1×10^{-11}
	Ripke (2011)	rs2021722-C	6p22.1	TRIM26	2×10^{-12}
		rs1625579-T	1p21.3	MIR137	2×10^{-11}
	Stefansson (2009)	rs6932590-T	6p22.1	TRNAI25	1×10^{-12}
		rs3131296-G	6p21.32	NOTCH4	2×10^{-10}
		rs12807809-T	11q24.2	SPA17, NRGN	2×10^{-9}
		rs9960767-C	18q21.2	TCF4	4×10^{-9}
Tourette syndrome	Scharf (2012)	rs7868992-G	9q32	COL27A1	3×10^{-8}

While childhood-onset disorders may reflect greater genetic loading, the majority of GWAS for pediatric psychiatric disorders have resulted in only suggestive associations, most likely due to modest sample sizes. The majority of GWAS in ASD have failed to achieve genome-wide significance, highlighting the probable extreme heterogeneity underlying ASD. GWAS in anorexia nervosa have produced suggestive results, with a few genetic loci approaching the significance threshold. Other child focused GWAS, such as those examining more dimensional traits such as "behavioral disinhibition" and "preschool internalizing problems" have again shown several loci approaching genome-wide significance. GWAS in early-onset bipolar disorder have not yielded any promising results, likely due to small sample sizes and possible unreliable diagnostic criteria for the disorder itself. PGC Working Groups are currently collecting childhood-onset samples to facilitate large GWAS of the magnitude of those successful in adults.

Table 37–4.
Summary of Neurodevelopmental Cross-Disorder GWAS with SNPs That Reach Significance of $p < 5 \times 10^{-8}$

Psychiatric Disorders	Reference	Strongest SNP-Risk Allele	Genomic Region	Mapped Genes	*P* Value
Bipolar and major depressive disorders	Liu (2010)	rs1006737-A	12p13.33	CACNA1C	$3 \times 10^{-}$

					8
Bipolar disorder and schizophrenia	Ruderfer (2013)	rs1006737	12p13.3	CACNA1C	6×10^{-13}
		rs17693963	6p22.1	TRNAI25	3×10^{-11}
		rs9834970	3p22.2	HSPD1P6, TRANK1	1×10^{-10}
		rs10275045	7p22.3	MAD1L1	2×10^{-9}
		rs4356203	11p15.1	PIK3C2A	6×10^{-9}
		rs4650608	1p31.1	RPL23P3, ELTD1	8×10^{-9}
	Sleiman (2013)	rs4721295	7p22.3	MAD1L1	6×10^{-10}
		rs17693963	6p22.1	TRNAI25	2×10^{-9}
		rs6703335	1q43	SDCCAG8	4×10^{-9}
	Wang (2010)	rs11789399	9q33.1	TPT1P9, BRINP1	6×10^{-9}
Autism spectrum disorder, attention-deficit/hyperactivity disorder, bipolar disorder, major depressive disorder, and schizophrenia	Smoller (2013)	rs2021722	6p22.1	TRIM26	2×10^{-12}
		rs2535629	3p21.1	ITIH3	3×10^{-12}
		rs1625579	1p21.3	MIR137HG	2×10^{-12}
		rs12966547	18q21.2	MAP1LC3P, RNA5SP459	3×10^{-10}
		rs7914558	10q24.32	CNNM2	2×10^{-9}
		rs10994359-C	10q21.2	ANK3	3×10^{-9}
		rs9371601	6q25.2	SYNE1	4×10^{-9}
		rs1006737	12q13.3	CACNA1C	$5 \times 10^{-}$

rs10994397	10q21.2	ANK3	7×10^{-9}
rs8321	6p22.1	PPP1R11, ZNRD1	8×10^{-9}

Interestingly, cross-disorder GWAS have revealed loci that span diagnostic borders, demonstrating effects in multiple psychiatric categorizations (e.g., schizophrenia and bipolar disorder, or ASDs, ADHD, bipolar disorder, major depression, and schizophrenia; Table 37–4). These cross-phenotype risk genes imply a likely common underlying mechanism for certain aspects, or dimensions, of psychiatric disorders or the presence of resilience genes that confer protection against mental illness.

GWAS approaches are designed to detect common variants of small to moderate effect. As discussed previously, it is plausible that rare variants of larger effect may also be important in psychiatric disease. While some GWAS chips now include rare variants, the most comprehensive method of evaluating rare variation in disease risk is genome sequencing (discussed in the following section).

Comparative Genomic Hybridization and GWAS for Copy Number Variation Analysis. Comparative genomic hybridization (CGH) is a microarray-based technology that can detect CNVs in the human genome rapidly and accurately. *CGH arrays* are glass slides spotted with thousands of probes, from short oligonucleotides to hundred-megabase clone inserts, spanning the genome or focused on a genomic region of interest. Genomic DNA from a test subject and a reference sample are differentially labeled and then cohybridized to the array. Differences in hybridization signals between the test and reference sequences can be detected and decoded to reveal deletions or duplications relative to the reference sample. This method can be used to search for *de novo* (new, noninherited) CNVs in disease populations. CGH has also been used to identify polymorphic CNVs that can be tested for association with disease. Given that local SNPs are frequently in LD with CNVs, it is possible to test association at known CNVs by using nearby SNPs, bypassing the need to measure specific CNV genotypes directly. As an alternative to CGH, GWAS chips can be analyzed for fluorescent signal strength to similarly identify and quantitate CNVs, with increased signal across many adjacent SNPs pinpointing a deletion and increased signal reflecting a duplication.

CNV studies using these microarray-based methods have successfully revealed novel submicroscopic CNVs using advanced computational CNV-predicting algorithms present in populations with intellectual disability, ASD, and schizophrenia. CNVs have been detected in patients with

sporadic ASD at 10 times the frequency in controls, and many CNVs are significantly associated with ASD. A genome-wide linkage study using SNP arrays conducted by the Autism Genome Project Consortium similarly identified a high rate of CNVs in ASD. A screen of the Autism Genetic Research Exchange (AGRE) families by CGH identified a microdeletion and reciprocal duplication at 16p11.2 in approximately 1 percent of cases. Such findings again suggest the prospect of many identifiable subpopulations within larger, heterogeneous ASD samples.

CNV studies have also shed light on biological processes that are affected in schizophrenia. Inherited and *de novo* CNVs associated with schizophrenia more frequently contain genes involved in neurodevelopmental pathways, encoding proteins that function in synaptic architecture, calcium signaling, and glutamatergic neurotransmission. In general, schizophrenia-associated CNVs have large individual effect sizes but are extremely rare in the population, whereas associated SNPs have small individual effect sizes but are common in the population. The aggregate effect of both CNVs (*de novo* and inherited) and SNPs must be considered in order to fully understand genetic risk for psychiatric disease. In addition, smaller *de novo* CNVs have also been implicated in bipolar disorder and ADHD, but at a much lower rate and reduced overall genetic burden compared to ASD and schizophrenia. Table 37–5 summarizes recurrent CNV loci and their association with psychiatric illness.

Next-Generation Sequencing. While GWAS approaches are well suited to examining common genetic variation, genome sequencing studies are able to identify rare variation leading to disease. In the past decade, the advent of NGS, or massively parallel sequencing, made exome (protein-coding regions of the genome) and whole genome sequencing possible. This technology generates millions of overlapping short reads from a DNA template that are assembled and aligned using a reference genome. By comparison to the reference genome, software is used to call genotypes at millions of known and novel variants across the sequenced genome. Since healthy individuals have been shown to carry tens of thousands of novel variants and hundreds of damaging mutations, the signal-to-noise obstacle presented by NGS is staggering. One successful strategy limits the analysis to *de novo* variation by comparing proband to parent genomes. Three independent groups simultaneously published scans of *de novo* variation in ASD in 2012. Collectively, these studies revealed that highly disruptive (nonsense and splice-site) *de novo* mutations in brain-expressed genes are associated with ASD and carry large effects. These *de novo* variants are overwhelmingly paternal in origin (4:1 bias) and positively correlated with paternal age. Multiple independent observations of *de novo* variants within the same gene among unrelated probands and across studies reliably identified several risk genes, including genes involved in sodium transport

(*SCN1A*, *SCN2A*), chromatin remodeling (*CHD8*), and axon guidance (*NTNG1*).

Table 37–5.
Summary of Copy Number Variation Associated with Psychiatric Disorders

Locus	Copy Number Change	Disease Association
1q21.1	Deletion/duplication	SCZ, ID, ASD, ADHD
2p16.3	Deletion	SCZ, ID, ASD
3q29	Deletion	SCZ, ID, ASD
7q11.2	Duplication	SCZ, ID, ASD, ADHD, ANXIETY
7q36.3	Deletion/duplication	SCZ, ID, ASD
15q11.2	Deletion	SCZ, ID, ASD, ADHD, OCD
15q11–13	Duplication	SCZ, ID, ASD
16p11.2	Deletion/duplication	SCZ, ID, ASD, ADHD, ANXIETY, MOOD
16p13.11	Duplication	SCZ, ID, ASD, ADHD
17q12	Deletion	SCZ, ID, ASD
22q11.1	Deletion	SCZ, ID, ASD, ADHD, ANXIETY, MOOD, VCFS
Xq28	Deletion/duplication	ASD

SCZ, schizophrenia; ID, intellectual disability; ASD, autism spectrum disorder; ADHD, attention-deficit/hyperactivity disorder; VCFS, velocardiofacial syndrome.

Pathway Analysis. Individual common SNPs appear to confer only small effects on psychiatric phenotypes, and while they have larger effect sizes, disease-associated CNVs are rare in the population. Therefore, multiple risk alleles in combination are likely required for disease expression.

As the first successful studies identifying risk alleles in psychiatric illness have been published, common pathways impacting specific cellular or neurodevelopmental processes have emerged. Many studies now include pathway analyses to add functional relevance and validity to GWAS or NGS results. Using bioinformatic and gene ontological approaches (such as the DAVID or GO Gene Ontology databases), investigators can determine whether associated genes belong to particular gene networks or cellular processes above chance levels. Using these approaches, scientists can begin to build gene interaction networks based on a priori knowledge of gene function gleaned from the research literature. These data can inform the development of prediction models, generating testable hypotheses about how a particular risk allele might alter a biological process to drive disease risk. An example of how genes cluster in pathways and across diseases is noted in Table 37–6.

PHARMACOGENOMIC EFFECTS ON PSYCHIATRIC TREATMENT IN CHILDREN

Identifying genes that predict treatment response or adverse effects will

improve the safety, tolerability, and effectiveness of psychotropic medication. Thus far, most pharmacogenetic studies in adult psychopharmacology have been limited to investigations of single genes or small numbers of genes and frequently test only single variants within these genes for association with drug response. Research in child and adolescent samples is emerging, especially in the area of ADHD pharmacogenomics. With the increase in large-scale clinical trials in child psychiatry over the last decade, more pharmacogenetic studies are soon to follow. Results from both adult and child studies suggest that genetic testing may hold more promise and initial clinical relevance in the prediction of side effects over efficacy.

Genome-Wide Association Studies

Large-scale GWAS are a particular challenge for detecting pharmacogenomic phenotypes since there may be multiple response "subtypes" which, given the variability in typical outcome measures, may require the strength of repeated assessments throughout a controlled clinical trial in order to detect. No genome-wide significant results emerged from several GWAS conducted for pharmacogenomic predictors of treatment response in adult psychiatry: a meta-analysis of 2,245 individuals of European ancestry with depression (STAR*D, GENDEP, MARS), a depression treatment study of citalopram in 865 adults by the International SSRI Pharmacogenomics Consortium (ISPC), a schizophrenia treatment trial of 738 subjects receiving multiple antipsychotics (CATIE), and a multiple treatment arm study of 458 patients with bipolar disorder (STEP-BD). GWAS for adverse effects have similarly failed to identify compelling signals in adult samples, including antidepressant treatment-emergent suicidality and metabolic, cardiac and movement side effects of antipsychotics.

The few pharmacogenomic GWAS in pediatric populations have been more successful thus far. One success was found in a meta-analysis of a GWAS examining antipsychotic-induced weight gain in youth receiving several antipsychotics identifying a risk allele near the melanocortin 4 (MC4R) gene. When examined in a meta-analysis with three replication cohorts, a genome-wide significant result was reported. GWAS data generated by the OCD Collaborative Genetics Association Study (OCGAS) was used to examine self-reported positive response or nonresponse to serotonin reuptake inhibitors in 804 subjects with childhood-onset OCD; one SNP near DISP1, a gene involved in neurodevelopment, achieved genome-wide significance. Interestingly, enrichment pathway analysis of suggestive associations highlighted the potential roles of glutamatergic and serotonergic neurotransmitter systems. Two small GWAS of response (n = 187) to methylphenidate and stimulant-induced cardiovascular side effects (n = 140) in children with ADHD, however, failed to produce genome-wide

significant results.

Candidate Gene Studies

ADHD Pharmacogenomics. The pharmacogenomics of ADHD represents the area of most research in child psychiatry and reveals the promise, as well as the challenges, of this line of investigation. Studies of ADHD treatment response contain many limitations, including differing study designs, samples of low power to detect genetic effects, short observation periods, retrospective designs, differing medications and delivery systems, varying drug titration protocols, and different response characterizations.

Table 37–6.
Representative Gene Pathways Suggested in Child Psychiatric Illness

Gene	Description	Function	Diseases
CALCIUM SIGNALING PATHWAY			
CACNA1C	Cav1.2 L type calcium channel, voltage-dependent (LTCC), alpha 1C subunit	Calcium signaling in neurons, neuronal gene transcription	Bipolar disorder, schizophrenia
CACNB2	Cav1.2 L type calcium channel, voltage-dependent (LTCC), beta 2 subunit	Calcium signaling in neurons, neuronal gene transcription	Schizophrenia
NCAN	Neurocan, chondroitin sulfate proteoglycan	Calcium ion binding	Bipolar disorder
DOPAMINE SIGNALING PATHWAY			
COMT	Catechol-O-methyl transferase	Degradation of catecholamines (dopamine, epinephrine, norepinephrine)	Depression, schizophrenia
DTNBP1	Dystribrevin binding protein 1, dysbindin	Trafficking of DRD2	Schizophrenia
GLUTAMATE SIGNALING PATHWAY			
GRIN2A	Glutamate receptor, ionotropic, N-methyl D-aspartate 2A	Glutamate receptor activity	Schizophrenia
GRM1, GRM3	Glutamate receptor, metabotropic 1 and 3	Glutamate receptor activity	Schizophrenia
GRM5, GRM7	Glutamate receptor, metabotropic 5 and 7	Glutamate receptor activity	Attention-deficit/hyperactivity disorder
NRGN	Neurogranin	Glutamate signaling, postsynaptic trafficking	Schizophrenia
MAJOR HISTOCOMPATIBILITY COMPLEX/IMMUNE PATHWAY			
MHC	Multiple immune functioning proteins	Immune response	Schizophrenia
SEROTONIN SIGNALING PATHWAY			

SLC6A4/5HTT	Solute carrier family 6, member 4 (serotonin transporter)	Releases serotonin from synapse, neurogenesis	Depression
NEURONAL PROCESSES/NEUROGENESIS			
ANK3	Ankryn-3	Axon guidance, axonogenesis	Bipolar disorder
BDNF	Brain-derived neurotrophic factor	Glutamate receptor activity and trafficking, synaptogenesis	Depression
CNTNAP2	Contactin associated protein-like 2, neurexin family	Neuronal cell adhesion molecule	Autism, schizophrenia
DISC1	Disrupted in schizophrenia-1	Neurite outgrowth, cortical development	Schizophrenia
NRXN1	Neurexin 1	Synaptic maturation	Schizophrenia
ODZ4	Teneurin transmembrane protein 4	Neurite outgrowth, transcription regulation	Bipolar disorder
TCF4	Transcription factor 4	Neurogenesis, synaptic maturation, transcription regulation	Schizophrenia
VIPR1	Vasoactive intestinal polypeptide receptor 1	Neurotransmitter, Neuroendocrine hormone	Schizophrenia
ZNF804A	Zinc finger protein 804A	Putative role in early neurodevelopment	Bipolar disorder, schizophrenia

The DAT1 (*SLC6A3*) gene has been tested repeatedly as a predictor of stimulant response, and several reports have concluded effects from the common 9- or 10-repeat functional 3′-untranslated region (UTR) variable number tandem repeat (VNTR). The 9 allele has been shown to reduce the subjective effects of single-dose administration of d-amphetamine in studies of normal controls, in keeping with the hypothesis that the reduced expression of *SLC6A3* encoded by the 9 allele would be associated with dampening of effects of drugs acting on this target. However, numerous studies in ADHD have yielded varying conclusions where either the 9- or 10-allele predicts favorable treatment response, and a meta-analysis was unrevealing. Taken together, these mostly conflicting reports on *SLC6A3* variants and stimulant response in ADHD are difficult to interpret. Nevertheless, the hypothesized functional importance of *SLC6A3* variants strongly supports further efforts to relate genetic variation, in the context of other genetic or environmental factors, to treatment response.

The dopamine receptor DRD4 gene has also been examined as a moderator of ADHD treatment response. Consistent with the reduced in vitro agonist efficacy produced by the DRD4 exon 3 VNTR seven-repeat allele (*DRD4.7*), one group found that children with this variant showed reduced sensitivity to methylphenidate, as evidenced by a higher-dose requirement for optimal clinical benefit. This result was confirmed by a second pediatric study. The National Institute of Mental Health (NIMH) Preschool ADHD Treatment Study (PATS) sample showed a suggestive negative moderating effect of the DRD4.7 allele, yielding a −0.25 effect size

on methylphenidate response in one analysis. The PATS study also found two copies of the common 240-base pair variant in the DRD4 5′-UTR promoter region to be significantly associated with methylphenidate response, although the effect was not seen in other secondary analyses from the same sample. Illustrating possible epistatic interactions between genes, a small study noted that the *DRD4.7* allele was associated with reduced response to methylphenidate, but only in those individuals who were homozygous for the 5HTT (*SLC6A4*) short variant. Other reports, mostly in adult populations, have found no effect of *DRD4* variants.

SNPs in other catecholaminergic genes posited to play a role in stimulant action have received preliminary study for possible treatment-moderating effects, including adrenergic genes (norepinephrine transporter [*SLC6A2/NET1*] and α-2A adrenergic receptor [*ADRA2A*]). In an RCT study of 82 subjects with ADHD querying functional variation in 7 monoaminergic candidate genes, the serotonin transporter intron 2 (STin2) VNTR was associated with MPH response. Two *DRD4* VNTRs and the 5HTT-linked promoter region (LPR) variant showed gene by dose interactions on math performance. The *COMT* Val/Met SNP and 5HTT-LPR were associated with irritability and vegetative symptoms, respectively. Both dopaminergic (dopamine D2 receptor [*DRD2*]) and adrenergic (*ADRA2A*) variants were associated with response to dexmethylphenidate in the NIMH/CIDAR Translational Research to Enhance Cognitive Control (TRECC) RCT of children with ADHD. The TRECC study screened 120 SNPs capturing complete common variation across 23 candidate genes for ADHD treatment effects. This study was the first to examine pharmacogenomic moderators of guanfacine outcomes, reporting effects of dopamine D1 and D2 receptors on guanfacine treatment response. The TRECC study also identified multiple strong associations ($<10^{-8}$) with growth and cardiovascular side effects of both d-MPH and guanfacine treatment. Several of these effects were in agreement with prior literature showing association of noradrenergic genetic variation (*SLC6A2* and *ADRA2A*) with cardiovascular changes with MPH treatment. In 90 Caucasian children with ADHD, stimulant tolerability was associated with a *DRD1* variant.

Several genes in other systems have some early data. A prospective pediatric study supported involvement of the Val66Met functional polymorphism in brain-derived neurotrophic factor (BDNF). SNAP-25, involved in neurotransmitter release, moderated both methylphenidate response certain side effects, including tics, abnormal movements, and irritability in the NIMH PATS preschool study. A study of *SNAP-25* variants in adults with ADHD was negative. Two reports in children with ADHD found associations at a variant in the latrophilin 3 (*LPHN3*) gene, implicated by linkage in a South American population isolate. While an adult study also found association at this locus, the direction of effect was

reversed.

A potentially important area of pharmacogenetics, relates to genetic effects on drug metabolism and distribution. A combined analysis of four pediatric ADHD clinical trials of atomoxetine (Strattera) identified effects of CYP450 2D6 genotypes on efficacy and side effects. Subjects genotyped as poor 2D6 metabolizers showed significant gains in ADHD symptom reduction with treatment, presumably due to beneficial effects of higher plasma atomoxetine levels despite lower daily doses; although adverse events of anorexia and tremor were slightly greater in poor metabolizers, discontinuation rates did not differ. In a more recent study, 10/100 children receiving standard atomoxetine doses were selected for *CYP2D6* genotyping due to lack of response or prominent side effects. Eight out of 10 of these patients were revealed to have compromised CYP2D6 activity. While in many cases, dose adjustment resolved the issue, the authors relate that in a number of cases, willingness to continue therapy was adversely affected by the initial experience. These data underscore the potential clinical importance of knowledge of CYP genotypes and individual optimization of drug therapy. Drug transporters, such as the multidrug resistance protein (MDR1 encoded by *ABCB1*), can impact transport into the body from the gut and into brain from the blood. In 134 children with ADHD, those homozygous for an *ABCB1* functional variant had 9× greater risk of side effects.

In summary, an understanding of the role of these variants supported by candidate gene approaches will require larger studies, improved controls, objective measures of drug response, and analytic approaches considering potential interactions of multiple variants within and between stimulant target genes. A significant omission in ADHD pharmacogenomics is the absence of information regarding genetic effects on response to other stimulants, as well as second-line ADHD medications besides atomoxetine. Given community practice, which often incorporates a range of medications for ADHD, such knowledge would be important. Finally, incorporating gene by environment interactions can be fruitful when environmental risk factors are known. Two studies incorporating the risk of prenatal exposure to maternal smoking have found support for gene by environment interactions on MPH treatment response. Each study focused on a single candidate gene (*NET1, LPHN3*).

Pharmacogenomics in ASD. The variability in efficacy and side effects of pharmacotherapy targeting ASD symptoms has been widely noted. Given the lack of effective treatments for the core symptoms of ASD, only a few pharmacogenomics studies address genetic moderators of treatment response. Two very small studies have shown nominal associations of *SLC6A4* variants with SSRI response. In a study by the RUPP Autism Network, genetic association with response to MPH treatment for

attentional symptoms in the context of ASD was assessed at 34 variants across 10 monoaminergic loci selected due to their role encoding MPH drug targets. Seven variants (*DRD1*, *ADRA2A*, *COMT*, *DRD3, DRD4, SLC6A3*, *SLC6A4*) showed nominal association with treatment response, a number significantly above that expected by chance, suggesting that many of these associations are in fact true positives. Variants in dopamine D2 and D3 receptors moderated tolerability in the 14 children who discontinued the trial due to side effects.

The rise in use of the second-generation antipsychotics (SGAs) has been accompanied by greater research into their efficacy and safety in pediatric groups. Initial studies of risperidone treatment of severe behaviors associated with autism noted prominent weight gain as a drug-related side effect. Several studies have implicated monoaminergic variants in the moderation of SGA side effects in ASD. The most consistent result, seen in three pediatric populations, has been the association of the serotonin 2C receptor (*HTR2C*) with antipsychotic-induced weight gain (AIWG). The largest pharmacogenomic study in ASD combined the two NIMH RUPP trials of risperidone treatment of 225 children with ASD-associated aggression, and identified numerous gene variants contributing to AIWG, including HTR2C, other monoaminergic variants, and several candidates in the energy balance pathway.

Pharmacogenomics in Other Areas. Dopamine receptor (*DRD2, DRD3*); serotonin receptor (*HTR1A, HTR2A*); and transcription factor ZNF804A, a schizophrenia risk locus, have amassed some support in moderating SGA treatment response in adult schizophrenia. Studies of rare, serious side effects of antipsychotics have resulted in replicated association of *HLA-DQB1* with agranulocytosis and *DRD2*, *HTR2A*, *CYP2D6*, and heparan sulfate proteoglycan 2 (*HSPG2*) in tardive dyskinesia. Children and adolescents, especially those with ASD, appear to be more likely to develop withdrawal dyskinesias with antipsychotics. Studies in children would be valuable to explore risk for motor side effects. Such observations relating to these concerning adverse effects, if robust enough, could potentially inform clinical practice.

The variability in the response of pediatric depression to drug therapy and the heightened sensitivity to adverse events of activation and suicidal ideation/impulses points to another area ripe for pharmacogenetic investigation. In adults, serotonergic and glutamatergic candidate genes have been found to predict adverse events and dropout with antidepressants, and in some instances efficacy differences by genotype. Specific tests of genetic effects on antidepressant-induced mania have encouraged further research. Given the need to improve the benefits and safety of treatment matching for early-onset depression, adult research immediately suggests genetic hypotheses to test in younger depressed

populations.

While the field of psychiatric pharmacogenomics is only beginning, pharmacogenomics studies are likely to differ from categorical disease genetic studies in several ways. Effect sizes are often large, phenotypes can be quantitative and more easily measured (e.g., weight change), and longitudinal data can be leveraged to improve statistical power. Exciting preliminary studies using gene expression profiling suggest that other pharmacogenomic techniques may also facilitate the prediction of drug effect. These analyses seek to identify distinct clusters of transcription patterns associated with drug response. The use of such methods is sure to greatly increase one's ability to better stratify individuals by shared biological features, leading to greater homogeneity in treatment response.

CLINICAL CORRELATIONS

The application of genetics in the practice of child psychiatry is in its infancy, but is growing rapidly. As described below, genetic testing now has a recommended role in many clinical evaluations, and in a greater and greater number of treatment decisions. Given the expected expansion of knowledge in child psychiatric genetics, clinicians need to stay informed, as additional practice implications are sure to follow.

Genetic Testing in Current Child Psychiatric Practice

Published guidelines from the American Academy of Pediatrics and the American College of Medical Genetics both recommend CGH and GWAS analyses as a routine part of the evaluation of any child with suspected ASD and all children with intellectual disability, in an effort to determine molecular diagnoses. In addition, these clinical guidelines include the recommendation for genotyping for the fragile X triplet-repeat mutations and identified mutations in the Rett Syndrome *MECP2* gene. The guidelines note that girls with learning disabilities and suspected other developmental delays have been found to possess these mutations in the absence of full syndromic features. Children with macrocephaly should also be tested for mutations in the *PTEN* gene, given its role in growth pathways and association with head circumference. At present, estimates are that routine genetic testing in ASD as well as intellectual disability results in positive findings in as many as 20 percent of individuals.

Some centers have begun routinely obtaining high-resolution genomic sequencing for suspected Mendelian disorders in place of high-density microarray testing for some patients, hoping to identify more rare variant risk mutations not captured by arrays. Other centers are now collecting genome-wide sequence data for every patient entering their medical system. With the expansion of clinical testing taking place, more and more patients will present with already available genome sequence data,

challenging physicians to identify findings of clinical import. As a result, psychiatrists need to attempt to stay current with the research literature, and develop facility with the many searchable genetic databases to aid their interpretations of patient information.

Pharmacogenomic Testing in Current Clinical Practice

Pharmacogenomic testing has unquestionably entered the clinical practice of child psychiatry. As of mid-2015, the U.S. Food and Drug Administration (FDA) has required the inclusion of pharmacogenomic information of clinical importance for 165 drugs, 30 of which are psychotropics or commonly used and encountered in psychiatric treatment. The FDA provides a comprehensive list of these labels (http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenet

The majority of pharmacogenomic labeling pertains to specific drug safety concerns associated with specific genotypes and their corresponding phenotypes. Arguably the most compelling example is the FDA recommendation for *HLA-B*1502* allele testing prior to prescribing carbamazepine (Tegretol) in patients of Asian ancestry. The *HLA-B*1502* allele appears to be primarily restricted to Asian populations and is strongly associated with increased risk for Stevens–Johnson syndrome/toxic epidermal necrolysis, as it has been found to be present in greater than 90 percent of individuals who developed SJS associated with carbamazepine. A majority of other pharmacogenomic considerations relate to genetically mediated drug metabolism differences associated with other safety risks.

Reflecting the growing appreciation of the importance of individual genetic differences in pharmacokinetics and drug distribution, the FDA also approved the first pharmacogenetic test, the AmpliChip CYP450, in January 2005. Test results from this chip provide relatively comprehensive testing of well-defined, clinically relevant polymorphisms in two cytochrome P450 genes (*CYP2D6* and *CYP2C19*) that are known to be associated with large differences in the metabolism of many psychotropic medications. Polymorphisms in the genes encoding these enzymes can result in poor-metabolizer and ultrarapid-metabolizer phenotypes, producing extreme variability in circulating drug levels. An estimate of the US annual treatment cost of adverse antidepressant reaction in CYP2D6 poor metabolizers was $420 million in 1999; consequently, some authors advocate the economic viability of testing CYP genotypes prior to antidepressant therapy. Current guidelines do not endorse standard genetic testing for antidepressants, however, testing is recommended for prescribing of tricyclic antidepressants, as well as specific antidepressants from other classes. These guidelines, available at the Clinical Pharmacogenetics Implementation Consortium website, advise avoidance of tricyclic antidepressants in individuals with ultrametabolizer CYP2D6

status because of potential lack of efficacy and either avoidance or 50 percent dose reductions and monitoring of therapeutic drug levels in individuals with the "poor metabolizer" phenotype of CYP2D6 or CYP2C19. Indeed, genotype-tailored treatment is routinely performed in many other areas of medicine (tamoxifen, warfarin, antineoplastics), and will continue to expand in relation to the safe and effective use of psychotropic therapies. Besides the AmpliChip, a number of commercial laboratories have begun to offer a wider range of genetic tests ostensibly to inform psychiatric practice. Clinicians should also be aware that the supporting evidence for larger-scale pharmacogenetic testing is often limited, and lacks the wider replication required before such tests can be considered "evidence based."

FUTURE DIRECTIONS

As the adoption of clinical genetic testing expands, spurred by future technical and scientific progress, more and more genetic information will become available to clinicians. A number of publically available databases can be accessed by clinicians in their efforts to understand the possible role of genetics at the level of the individual patient. A list of these databases is provided below. As mentioned previously, a number of publically funded websites also include information on the clinical importance of pharmacogenetic differences with clinical import.

Newer scientific methods hold promise too. The field of epigenetics attempts to understand the impact of experience (e.g., environmental influences) on the regulation and expression of genes. Early findings have already shed light on the mechanisms by which early childhood adversity can lead to risk for adult mental disorders (posttraumatic stress disorder, mood disorders) via methylation of genes involved in the stress–response system. As these mechanisms are revealed, they are likely to point to novel prevention and treatment strategies that could be applied to youth, and potentially alter the trajectory of their development and adaptation. Such exciting prospects should encourage close attention to this important area of psychiatric research.

REFERENCES

Aberg KA, Liu Y, Bukszar J, et al. A comprehensive family-based replication study of schizophrenia genes. *JAMA Psychiatry*. 2013;70(6):573–581.

Addington AM, Rapoport JL. Annual research review: impact of advances in genetics in understanding developmental psychopathology. *J Child Psychol Psychiatry*. 2012;53(5):510–518.

Baker E, Jeste SS. Diagnosis and management of autism spectrum disorder in the era of genomics: rare disorders can pave the way for targeted treatments. *Pediatr Clin North Am*. 2015;62(3):607–618.

Burdett T, Hall PN, Hastings E, et al. *The NHGRI-EBI Catalog of Published Genome-Wide*

Associated Studies. Available at: www.ebi.ac.uk/gwas. Accessed April 2015.

Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*. 2013;381(9875):1371–1379.

Doherty JL, Owen MJ. Genomic insights into the overlap between psychiatric disorders: implications for research and clinical practice. *Genome Med*. 2014;6(4):29.

Escudero I, Johnstone M. Genetics of Schizophrenia. *Curr Psychiatry Rep*. 2014;16(11):502.

Fabbri C, Porcelli S, Serretti A. From pharmacogenetics to pharmacogenomics: the way toward the personalization of antidepressant treatment. *Can J Psychiatry*. 2014;59(2):62–75

Gejman PV, Sanders AR, Kendler KS. Genetics of schizophrenia: new findings and challenges. *Annu Rev Genomics Hum Genet*. 2011;12:121–144.

Goldman D, Domschke K. Making sense of deep sequencing. *Int J Neuropsychopharmacol*. 2014;17(10):1717–1725.

Gratten J, Wray NR, Keller MC, Visscher PM. Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nat Neurosci*. 2014;17(6):782–790.

Hamilton SP. The promise of psychiatric pharmacogenomics. *Biol Psychiatry*. 2015;77(1):29–35.

Hebebrand J, Scherag A, Schimmelman BG, Hinney A. Child and adolescent psychiatric genetics. *Eur Child Adolescent Psychiatry*. 2010;19(3):259–279

Kavanagh DH, Tansey KE, O'Donovan MC, Owen MJ. Schizophrenia genetics: emerging themes for a complex disorder. *Mol Psychiatry*. 2015;20(1):72–76.

Ku CS, Polychronakos C, Tan EK, et al. A new paradigm emerges from the study of de novo mutations in the context of neurodevelopmental disease. *Mol Psychiatry*. 2013;18(2):141–153.

Lohoff FW. Overview of the genetics of major depressive disorder. *Curr Psychiatry Rep*. 2010;12(6):539–546.

Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell*. 2012;148(6):1223–1241.

McMahon FJ, Akula N, Schulze TG, et al. Meta-analysis of genome-wide association data identifies a risk locus for major mood disorders on 3p21.1. *Nat Genet*. 2010;42(2):128–131.

Mehta S, Nurmi EL. Genetic pathways to autism spectrum disorders. *Neuropsychiatry*. 2013;3(2):193–207.

Pinto D, Delaby E, Merico D, et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Hum Genet*. 2014;94(5):677–694.

Polancyzk G, Zeni C, Genro JP, Roman T, Hutz MH, Rohde LA. Attention-deficit/hyperactivity disorder: advancing on pharmacogenomics. *Pharmacogenomics*. 2005;6(3):225–234

Rapoport J, Chavez A, Greenstein D, Addington A, Gogtay N. Autism spectrum disorders and childhood-onset schizophrenia: clinical and biological contributions to a relation revisited. *J Am Acad Child Adolesc Psychiatry*. 2009;48(1):10–18.

Ruderfer DM, Fanous AH, Ripke S, et al. Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Mol Psychiatry*. 2014;19(9):1017–1024.

Sakolsky DJ, McCracken JT, Nurmi EL. Genetics of pediatric anxiety disorders. *Child Adolesc Psychiatr Clin N Am*. 2012;21(3):479–500.

Schreiber M, Dorschner M, Tsuang D. Next-generation sequencing in schizophrenia and other neuropsychiatric disorders. *Am J Med Genet B Neuropsychiatr Genet*. 2013;162B(7):671–678.

Shinozaki G, Potash JB. New developments in the genetics of bipolar disorder. *Curr Psychiatry Rep*. 2014;16(11):493.

Singh S, Kumar A, Agarwal S, Phadke SR, Jaiswal Y. Genetic insight of schizophrenia: past and future perspectives. *Gene*. 2014;535(2):97–100.

Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet*. 2012;13(8):537–551.

38

Neuroimaging in Psychiatric Disorders of Childhood

FRANK P. MACMASTER, PH.D., MARIKO SEMBO, B.A., AND DAVID R. ROSENBERG, M.D.

INTRODUCTION

Over the last quarter century, there has been a revolution in the current understanding of the living human brain. The advent of brain imaging has allowed for the noninvasive assessment of brain structure, function, and chemistry. Until the development of brain imaging, the investigation of the neurobiological underpinnings of psychiatric illness in children was stalled compared to adults. Indeed, brain imaging has allowed for tremendous gains in pediatric psychiatric research. Evidence increasingly suggests that abnormal brain development underlies many childhood-onset neuropsychiatric disorders. The application of brain imaging to describe the biology of the illness can lead to improved diagnosis and innovative treatments. Moreover, the study of brain maturation may make it possible to find critical time periods during which brain abnormalities could lead to psychiatric disease. The research findings may also indicate potential "targets" for treatment intervention.

This chapter has three sections. First, core brain imaging methods are described. Second, special considerations (such as gender and development) are addressed. Third, relevant findings from brain imaging studies of psychiatric disorders of childhood are summarized. Disorders include attention-deficit/hyperactivity disorder (ADHD), depressive disorders such as major depressive disorder (MDD), bipolar and related disorders, such as bipolar I disorder (BP-I), anxiety disorders, obsessive-compulsive disorder (OCD), posttraumatic stress disorder (PTSD), Tourette syndrome (TS), childhood-onset schizophrenia, substance use disorders, feeding and eating disorders, autism spectrum disorder (ASD), and specific learning disorders including dyslexia. The purpose of this chapter is to provide an understanding of brain imaging in psychiatric disorders of childhood.

TECHNIQUES AND PRACTICAL CONSIDERATIONS

Neuroimaging Techniques

Twenty-five years ago, the study of the brain was largely limited to postmortem work. However, postmortem studies are not practical for developing an understanding of psychiatric disorders of childhood. Nearly all postmortem brains come from adults. Hence, they are confounded by chronic illness and exposure to treatment. The application of brain imaging techniques has overcome this obstruction to progress.

Morphology/Structure

Morphological studies include the examination of structure, volume, and shape. Early studies in psychiatry began with more of a "radiological" approach—that is, looking for the presence or absence of visible damage or lesions. An example of this is the link between magnetic resonance imaging (MRI) T2 hyperintensities in left frontal stroke and depression. These lesions can be observed with the naked eye. This approach was not adequate, as most psychiatric illnesses do not involve a visible lesion. Hence, the field focused on more quantitative methods.

A measure should be both valid and reliable. Measurement of biological structures is challenging, as many parameters change with development. In the human brain, not all aspects are MR-visible. This is critical when designing an experiment using MRI as the method. The effect must be of necessary scale to be detectable. However, the technology is continually improving and it may someday be possible to evaluate single cells in vivo using MRI-based technologies. In addition to the ability to be detected and measured, the structure must also be clinically of interest. In other words, it should be critically involved in a specific function associated with the psychiatric disorder in question.

The two principal steps in morphological MRI studies are data acquisition and quantification. With both of these steps, bias and error can diminish the validity of the study. Bias can obscure real effects or produce incorrect findings. Error is random variation. Both bias and error can reduce statistical power and result in erroneous findings. There is no single accepted method for reducing bias and error in each of the steps involved in MRI experiments. Furthermore, there is little consensus regarding how to properly perform anatomic quantification, with methods varying considerably. Due to inherent variation in regional brain volumes, it is accepted that large sample sizes are required to demonstrate even moderate effects between groups.

Volume and Shape. A "structure" should be distinct from its neighbors. Within the brain structure, there should be a common aspect to or resultant from the parts comprising the structure. This could be cell type, neurotransmitter receptor distribution, or function. For instance, *caudate* could be a descriptor because it has specific related functions and is contiguous spatially. When trying to understand the brain, you need to

know where functionally specific structures are located, what function they perform, and how the distributed processing of that function is organized. The volume of a brain structure is reflective of the size, shape, and density of its parts. It is reasonable to assume that functional changes are reflected physiologically with changes in cellular composition and/or activity. Consequently, by examining the volumetric properties of a structure one can make inferences about its functionality. Researchers may also study the shape of a brain structure and compare groups and conditions. The quantitative analysis of biomedical shape variation and shape change, while more complicated mathematically, offers useful information on two- and three-dimensional properties of structures that may be relevant to disease states. More recently, the investigation of cortical thickness has become increasingly common in the study of pediatric psychiatric populations. These approaches are typically automated and rely on established templates for the localization of differences.

MRI-based volumetric approaches are a viable tool in basic science studies. As a scientific method, it has served many purposes—mapping developmental trajectories, disease processes, and the effect of treatment. Thanks to its use of nonionizing radiation, MRI allows for repeated measures over time with a high degree of safety. MRI further avoids the physiological changes inherent in the death process and the effects of the lag between death and fixation processing that are part of postmortem studies. Furthermore, for psychiatric diseases of childhood, like OCD and ADHD, there are virtually no postmortem studies. Hence, investigation of psychiatric diseases of childhood has benefited from MRI approaches. It should be noted that MRI is obviously of lesser utility in children with braces or claustrophobia.

Diffusion Tensor Imaging. Diffusion tensor imaging (DTI) is based on the MR signal resulting from the restricted movement of water in brain tissue. DTI allows for the examination of the integrity and directional orientation of white matter tracts (e.g., superior longitudinal fasciculus) in the brain. The fundamental architecture of white matter has the axons running in parallel bundles or tracts. This forces water molecules to move along the main bearing of the tract. This oriented dispersal pattern is termed as anisotropic diffusion. The location, direction, and anisotropy of the white matter tracts can be measured.

Computed Tomography. Computed tomography (CT) was the first imaging method to make routine the direct, in vivo, and noninvasive evaluation of the living human brain. CT is typically used for the assessment of lesions, but also the capability to assess the volume of a structure. CT has limited resolution and contrast between gray and white matter. In practice, MRI has replaced quantitative CT in brain research due

to its improved resolution and tissue contrast. As stated earlier, the use of MRI also avoids the exposure to ionizing radiation associated with CT. This increase in risk (or perception of risk) renders CT's routine use for the study of pediatric conditions less appealing.

Chemistry

Brain imaging methods that evaluate brain chemistry may be more sensitive than structural imaging studies in the identification of localized abnormalities in psychiatric illness. As indicated above, the radiological hallmarks of psychiatric illness (i.e., differences from healthy controls or changes with treatment) are not typically visible to the naked eye. These differences are subtle and require very sensitive methods of detection.

Single-Photon Emission Computed Tomography and Positron Emission Tomography. Positron emission tomography (PET) and single-photon emission computed tomography (SPECT) methods have been critically informative regarding the pathogenesis and mechanisms of treatment response of several neuropsychiatric conditions in adults. These approaches permit the investigation of regional cerebral blood flow, glucose metabolism, and neurochemistry—including specific neurotransmitters, transporters, and receptors. These techniques have only been applied infrequently to pediatric psychiatric disorders. The trepidation arises from the fact that SPECT and PET use ligands labeled with radiation. This risk limits their role in the study of children and adolescents. Institutional review boards frequently disallow such studies in healthy children without neuropsychiatric illness. When deprived of a proper control group, discovering the pathology of childhood psychiatric disease using PET or SPECT is rendered problematic. These radiation concerns make longitudinal studies (e.g., treatment and developmental effects) not viable in children with or without disease. MR-based techniques are free of ionizing radiation and, as such, have rapidly become the preferred means for examining neuroanatomy, blood flow, and chemistry in children.

Magnetic Resonance Spectroscopy. Magnetic resonance spectroscopy (MRS) measures chemicals in the brain. Although many nuclei with odd-numbered protons or neutrons, including phosphorus (^{31}P), carbon (^{13}C), fluorine (^{19}F), sodium (^{23}Na), and proton (^{1}H) compounds, can be detected using MRS, ^{1}H-MRS is the most commonly used technique, providing measures of the following: *N*-acetylaspartate, a neuronal marker; creatine/phosphocreatine, involved in energy metabolism; choline compounds, membrane components; myo-inositol, a secondary messenger involved in signal transduction; and glutamate, glutamine, and γ-aminobutyric acid (GABA)—all of which occur in relatively high

concentrations in brain (Fig. 38–1). ^{31}P-MRS is the second most common approach in developmental populations. ^{31}P-MRS allows for the study of the products of membrane synthesis (phosphomonoesters [PMEs] like phosphoethanolamine and phosphocholine) and breakdown (phosphodiesters [PDEs] like glycerophosphoethanolamine and glycerophosphocholine), along with substrates of energy metabolism. The choice of MRS versus PET/SPECT measures of brain chemistry involves important tradeoffs. First, MRS does not have the power to detect chemicals occurring at lower concentrations, such as monoamines, neuropeptides, opiates, and so on. Second, MRS has relatively poor spatial and temporal resolution as compared to that of PET. However, with data acquisition and spectral editing programs becoming increasingly sophisticated and the increases in magnetic field strength, striking improvements in the spatial and temporal resolution of MRS are being fulfilled.

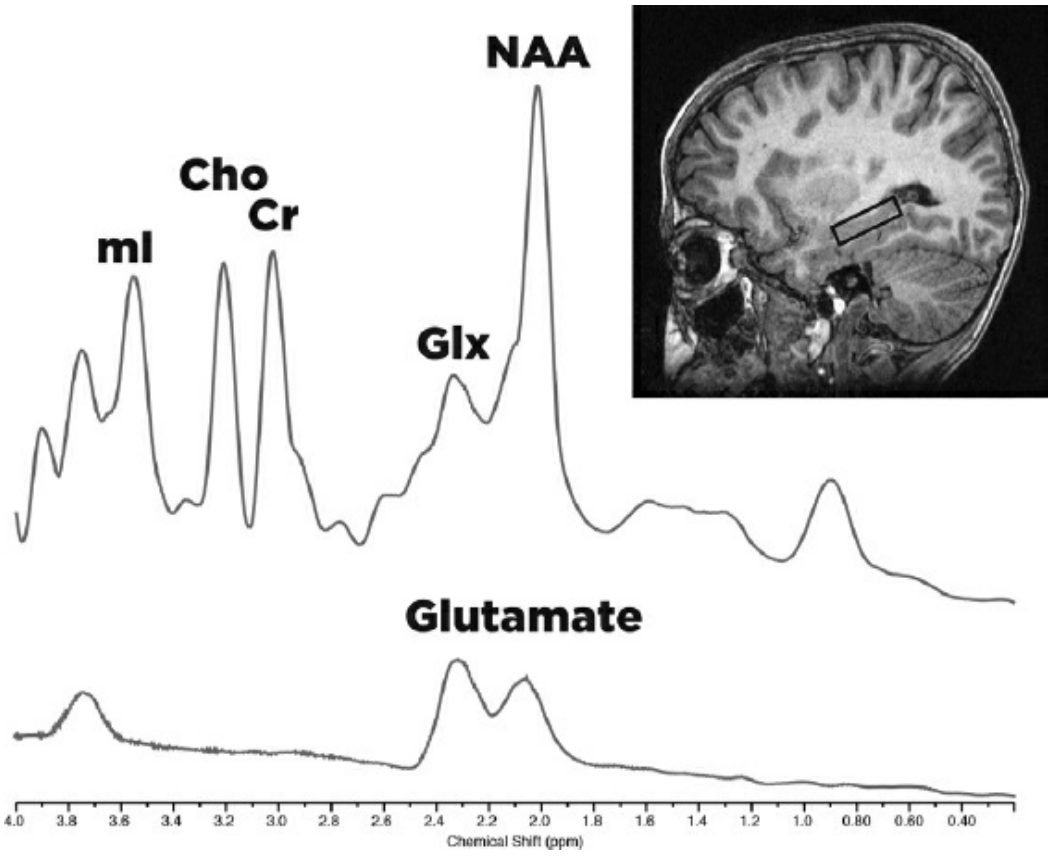

FIGURE 38–1. Sample proton magnetic resonance spectrum from the hippocampus. Spectra were obtained from a 6.0-mL volume of interest as shown on the magnetic resonance image. Cho, choline; CR, creatine/phosphocreatine; Glx, glutamate/glutamine/γ-aminobutyric acid; mI, myo-inositol; NAA, *N*-acetyl-aspartate. At higher field strengths, distinguishing glutamate from Glx is possible. Angulation of voxels to best accommodate anatomy (like the hippocampus) is also possible. (Figure courtesy of Frank MacMaster, Ph.D.)

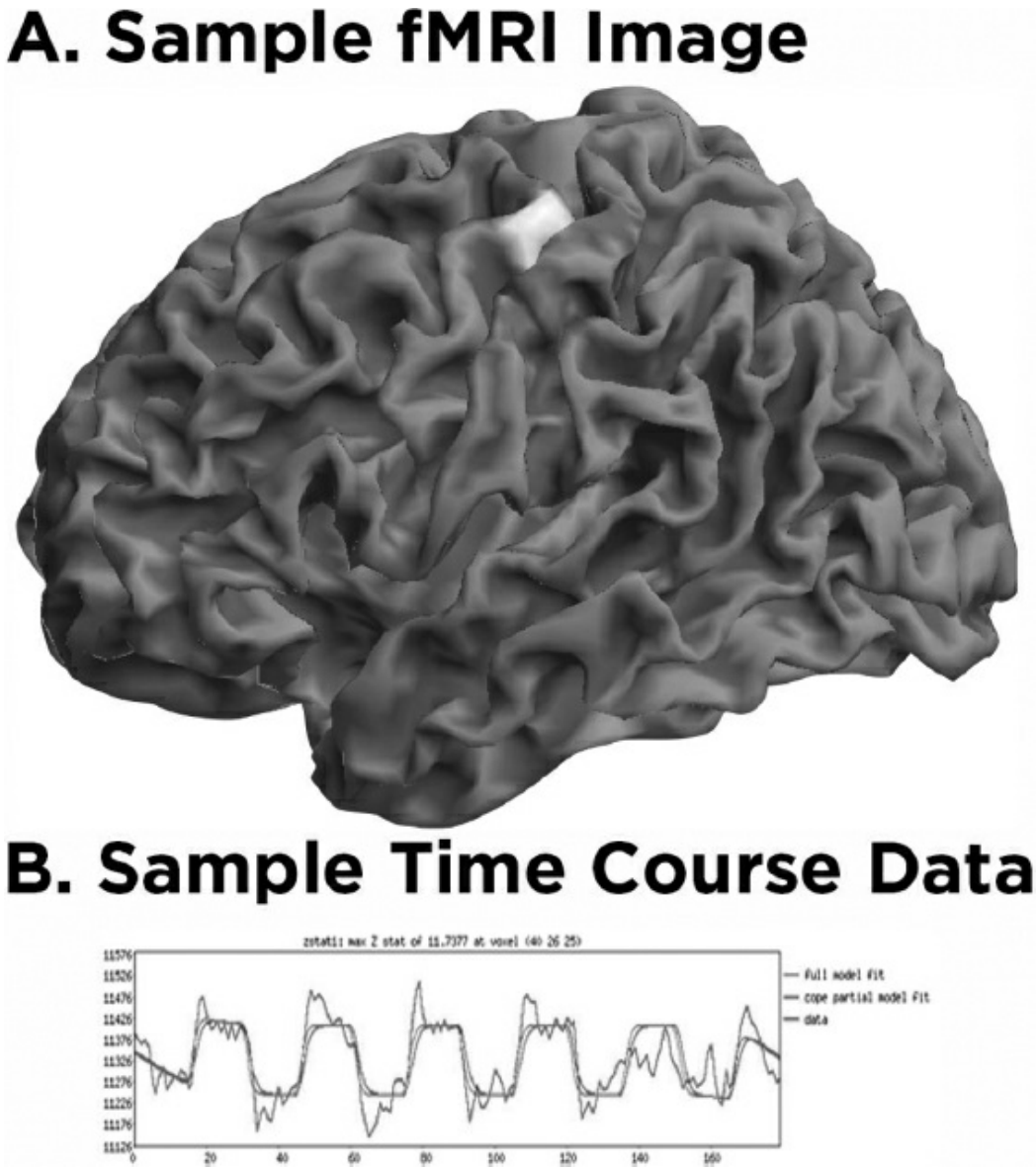

FIGURE 38–2. **A:** Functional magnetic resonance imaging (fMRI) data demonstrating activation during a simple finger-tapping task. **B:** Signal intensity from a pixel shows the correspondence with the task paradigm (not tapping/tapping). (Figure courtesy of Frank MacMaster, Ph.D.)

Functional Neuroimaging

The selection of functional magnetic resonance imaging (fMRI) versus PET/SPECT measures of cerebral blood flow also involves tradeoffs. fMRI delivers excellent spatial resolution (1 mm or less) along with potentially superior temporal resolution (about 1 second), compared to the PET/SPECT approaches. Second-to-second temporal resolution is now possible with event-related fMRI. Unlike PET/SPECT, fMRI does not provide an absolute measure of cerebral blood flow. Rather, fMRI can detects changes in blood flow, requiring the use of carefully designed neurobehavioral paradigms during imaging to generate activation signals (Fig. 38–2). New analytical techniques allow for the quantification of

"baseline" or "resting-state" blood flow using fMRI images. The study of so-called default mode networks offers an innovative method for determining differences in brain function between groups and conditions.

Critical Considerations

When acquiring and interpreting data from MR-based neuroimaging studies of children, a variety of factors merit consideration. A concise overview is presented below.

Technical Factors. The capacity of children to remain quiet and still while being scanned is of particular concern to investigators using any imaging technique. Motion artifacts are problematic in studies as they can often render data of such poor quality as to make it unusable. Remaining still can be challenging even for healthy children, let alone those suffering from neuropsychiatric and neurodevelopmental disorders. Children are often naturally curious about the scanning environment. Some children can become quite anxious in the small, confined space of the magnet. Even with earplugs and headphones, the loud noises inherent in certain specific imaging sequences can be disturbing. Pharmacological sedation has been used successfully in children who would be otherwise unable to tolerate neuroimaging in clinical settings, but in research studies, such procedures may entail unwanted side effects and alter the neurobiology being studied. For research studies, sedation and its potential risks are more ethically difficult to justify, especially in healthy control participants. Desensitization to the scanning environment with a "mock scanner" (Fig. 38–3) before the actual scan has been found to reduce subjective distress and heart rate, facilitating the successful completion of the study. Unfortunately, "mock scanners" can be costly and require considerable space in already crowded hospitals. Other means of desensitization, such as careful description of the scanner, videos of its use, and attention to the questions of patients and their parents, may also be helpful in reducing anxiety and habituating children to the scanning environment to facilitate successful MR-based study.

FIGURE 38–3. Child entering the mock scanner, which is designed to replicate the experience of a magnetic resonance scan, complete with the sounds that occur during the protocol to be used. (Figure courtesy of Frank MacMaster, Ph.D.)

Neuroanatomical Factors. Smaller body size and brain maturation are key technical considerations in performing MR-based neuroimaging research in children. Children may not fit well into head coils designed for adults, leading to increased magnetic field inhomogeneity and decreased signal-to-noise ratios. Greater variability of the cardiorespiratory cycle in children may also contribute to increased physiologic noise during fMRI. Head motion, which is more common in children, can influence functional connectivity measures as well. However, the brain is surrounded by less bony and muscular tissue in children than in adults, so that MR surface head coils lie closer to the brain regions being imaged, possibly enhancing signal detection.

Peak synaptic density within the brain occurs between 4 and 8 years of age. This peak is associated with greater metabolic and perfusion demands, which may lead to more robust increases in regional cerebral blood flow during neuronal activation and/or increased "noise" in fMRI experiments. Furthermore, the networking between brain regions is immature and may result in a weaker, more diffuse fMRI signal. Structurally, greater ratios of gray to white matter and less cerebrospinal fluid (CSF)—occupied free space around vasculature in children may lead to fundamental differences in the generation of MR signal in children versus adults. These factors must be taken into account when attempting to compare subjects of disparate

ages.

The interpretation of fMRI signals in children is further complicated by structural differences, because data processing typically involves the "warping" of study subjects' brains to a standard brain atlas derived mainly from adults. The development of standardized brain atlases for children of different ages offers a solution to this methodological problem. Still, brain volumes may vary during development even in same-aged children, so that the segmentation of brain into gray matter, white matter, and CSF may be necessary to generate individual descriptors of structural differences for use in statistical correction between study subjects. The use of study-specific templates for fMRI and voxel-based morphometry (VBM) is the most logical course. The volumetric MRI study of normal brain development in children and adolescents jointly sponsored by the National Institutes of Mental Health, Child and Human Development, and Neurologic Disorders and Stroke has recently been initiated and will generate a large normative database specific to pediatric structural maturation. Given the aforementioned variation in the developing brain and its effect on analysis methods, the importance of this resource cannot be overestimated.

Cognitive Factors. Another possible confound in the interpretation of certain fMRI paradigms in children relates to their ability to perform the cognitive tasks under investigation. Across development, children vary considerably in cognitive skill and their capacity to sustain attention and remain "on task." Potential tasks need to be vetted thoroughly in children and adolescents before robust interpretations of functional data can be made. Cognitive-based fMRI studies comparing children and adults should account for accuracy and reaction time to avoid falsely attributing performance differences to underlying neurobiology. Assessment of effort should also be considered because some children may guess or give up when a cognitive task is too difficult or too boring. In designing a cognitive paradigm for use in an fMRI study of children, tailoring the level of task difficulty to individual capability may help to ensure consistent performance across subjects at different developmental stages. The possibility that children may use different cognitive strategies on a given task can be examined by correlating task performance with neuropsychological measures. The correlation of activation patterns with performance rather than age may help to identify true functional differences based on brain maturity rather than structural- or performance-based differences, which can vary with age and between same-aged subjects during development.

Functional neuroimaging studies of executive functions, including working memory (the ability to respond to new situations on the basis of stored information) and neurobehavioral response inhibition (inhibition of

context-inappropriate responses), have demonstrated similar regional activation patterns (e.g., prefrontal cortex) in children and adults, with children recruiting larger volumes of activation with greater signal intensity. This suggests a less specific neural network and/or less efficient processing of cognitive stimuli in childhood. Not only may children display different activation strengths in response to stimuli, they may also show regional differences as well. On a task of spatial working memory, adults activated the anterior cingulate, a structure critically involved in arousal and mediating response conflict, whereas children did not. Lower accuracy on the task among children may be related to an incompletely developed anterior cingulate, resulting in a relatively reduced ability to modulate competing responses. In general, the functional connectivity between the anterior cingulate and other cortical and subcortical structures appears to increase with age, as demonstrated in a ^{1}H-MRS study in humans.

Opposite laterality of activation may also distinguish healthy children from healthy adults on tasks of response inhibition and working memory. Children may also recruit subcortical (e.g., basal ganglia, thalamus) and limbic (e.g., amygdala, hippocampus) structures to support incompletely developed prefrontal cortex (PFC) in mediating response inhibition. Moreover, whereas both healthy children and adults show altered activation of the amygdala in response to fearful faces as measured by fMRI, children but not adults show activation for neutral faces, as well as fearful faces, further suggesting that more specific neural processing of information develops with age.

NORMAL DEVELOPMENT

Psychiatric diseases of childhood can only be understood in the context of brain development. Brain size increases considerably in utero through 2 years of age, during which time synaptogenesis is most active. While total brain volume is comparatively static after 5 years of age, there are two dynamic processes that are ongoing through adolescence into young adulthood—synaptic pruning and myelination. Evidence suggests that increased myelination, connectivity, and changing functional activation patterns indicate that more efficient connections between brain regions are being made from childhood through adulthood.

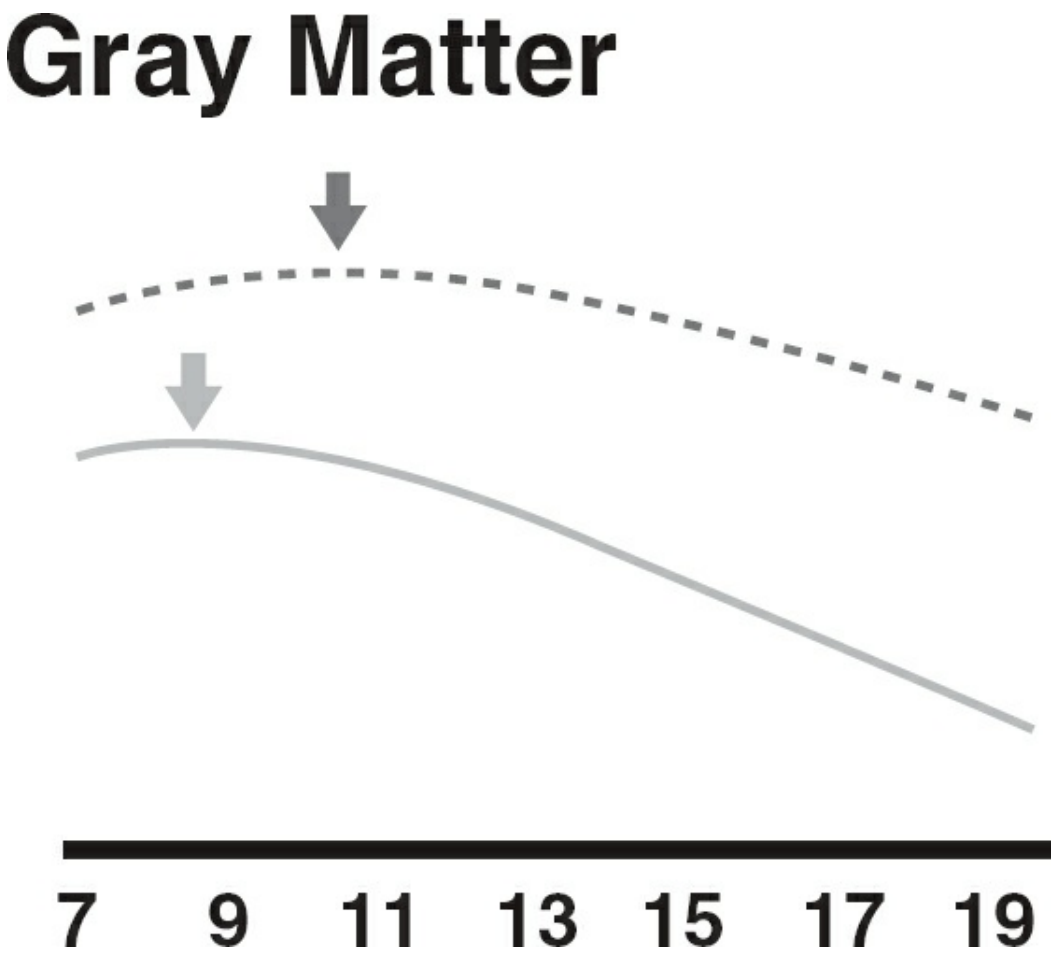

FIGURE 38–4. Schematic showing growth and aging changes in gray matter from childhood through adolescence. Gray mater volume reached a maximum by 9.5 years of age in females (*solid line*) and 10.5 years of age in males (*broken line*) and thereafter declined linearly. This is most pronounced in the frontal and temporal lobes. (Schematic adapted from Giedd JN, Raznahan A, Alexander-Bloch A. Child psychiatry branch of the National Institute of Mental Health longitudinal structural magnetic resonance imaging study of human brain development. *Neuropsychopharmacology*. 2015;40(1):43–49.)

FIGURE 38–5. Schematic adapted from positron emission tomography studies of glucose metabolism as a function of age in children through young adulthood. Note the steep rise through childhood followed by a sharp decline just as puberty and adolescence begins. (Schematic adapted from Chugani HT. Biological basis of emotions: brain systems and brain development. *Pediatrics*. 1998;102(5 Suppl E):1225–1229.)

Gray Matter

Gray matter volumes peak in mid-childhood to early adolescence, earlier in females than males. They then plateau briefly and show a decrease following puberty (Fig. 38–4). This pattern of gray matter maturation parallels changes in local cerebral metabolic rates for glucose in the cerebral cortex as a function of age (Fig. 38–5). These developmental changes most likely represent an increase in the efficiency of neural processing via the pruning of excessive synaptic connections. It is not clear what cellular changes (i.e., loss of neurons, dendrites, axons, and/or neuropil) are responsible for the reduction of gray matter. Indeed, it has recently been shown in 308 healthy children and adolescents that cortical thickness, as measured using MRI, can accurately predict an individual's age.

When considering the developmental basis of normal and abnormal cognitive/emotional function, it is important to note that it is heterochronous. That means that the timing, degree, and rate of cortical and subcortical gray matter change appear to be regionally specific. Gender also plays a significant role in brain development. In a large longitudinal MRI study of cortical volumes in 145 healthy children and adolescents (4 to 20 years of age), gray matter volume was found to peak and then decline after puberty in the parietal, frontal, and temporal lobes but on different time scales (Table 38–1). In contrast, gray matter in the occipital cortex

increased through early adulthood (approximately 20 years of age). With regard to gender differences, the peak for gray matter volume in the frontal and parietal lobes occurs nearly 1 year earlier in girls than boys. This seems to coincide with gender-specific differences in the timing of puberty and suggests the possible influence of gonadal hormones on brain development.

White Matter

The volume of white matter increases roughly linearly in the brain during childhood and adolescence. This is consistent with myelinization during this time (Fig. 38–6). Growth in cortical white matter appears to be regionally specific. The most dramatic increase in white matter occurs in the dorsolateral PFC between 5 and 17 years of age. Consistent with the progressive development of executive function through adolescence and into early adulthood, a segmentation analysis of MRI data found that gray matter reduction in the frontal cortex continued into young adulthood, after gray matter volumes had stabilized in other cortical regions. The "reduction" of frontal lobe gray matter was thus hypothesized to reflect the effects of an increase in myelination (and hence the relative amount of white matter). This interpretation is supported by the inverse relationship of cortical gray matter and brain growth—presumably due to white matter increase—in the frontal lobes. It has been suggested that myelinization may be as important to brain maturation as synaptic pruning. Indeed, the capacity for attention—a function largely attributable to the frontal lobe—has been shown to increase with larger areas of the anterior cingulate cortex over development, possibly the result of increased myelination. The corpus callosum (CC)—the white matter tract connecting the hemispheres—appears to demonstrate a particularly dynamic growth pattern (Fig. 38–7). Unlike the posterior–anterior progression of brain development described in much of the preclinical literature, CC growth occurs in a rostrocaudal wave between 3 and 15 years of age (Fig. 38–7). It may be that CC growth parallels synaptogenesis in associated cortical regions rather than pruning and myelinization. In a large study of 103 healthy participants (5 to 32 years of age), nonlinear development of major white matter pathways was observed. There were significant within-subject changes in the majority of children and early adolescents, with these changes being completed by late adolescence for projection and commissural tracts. In contrast, association tracts demonstrated post-adolescent within-subject maturation of both fractional anisotropy and mean diffusivity.

Table 38–1.
Peak Cortical Gray Matter Volumes by Age and Gender

Cortical Region	Girls (Age in years)	Boys (Age in years)
Parietal	10.2	11.8

Frontal	11.0	12.1
Temporal	16.7	16.5

Data summarized from Giedd JN, Raznahan A, Alexander-Bloch A. Child psychiatry branch of the National Institute of Mental Health longitudinal structural magnetic resonance imaging study of human brain development. *Neuropsychopharmacology*. 2015;40(1):43-49.

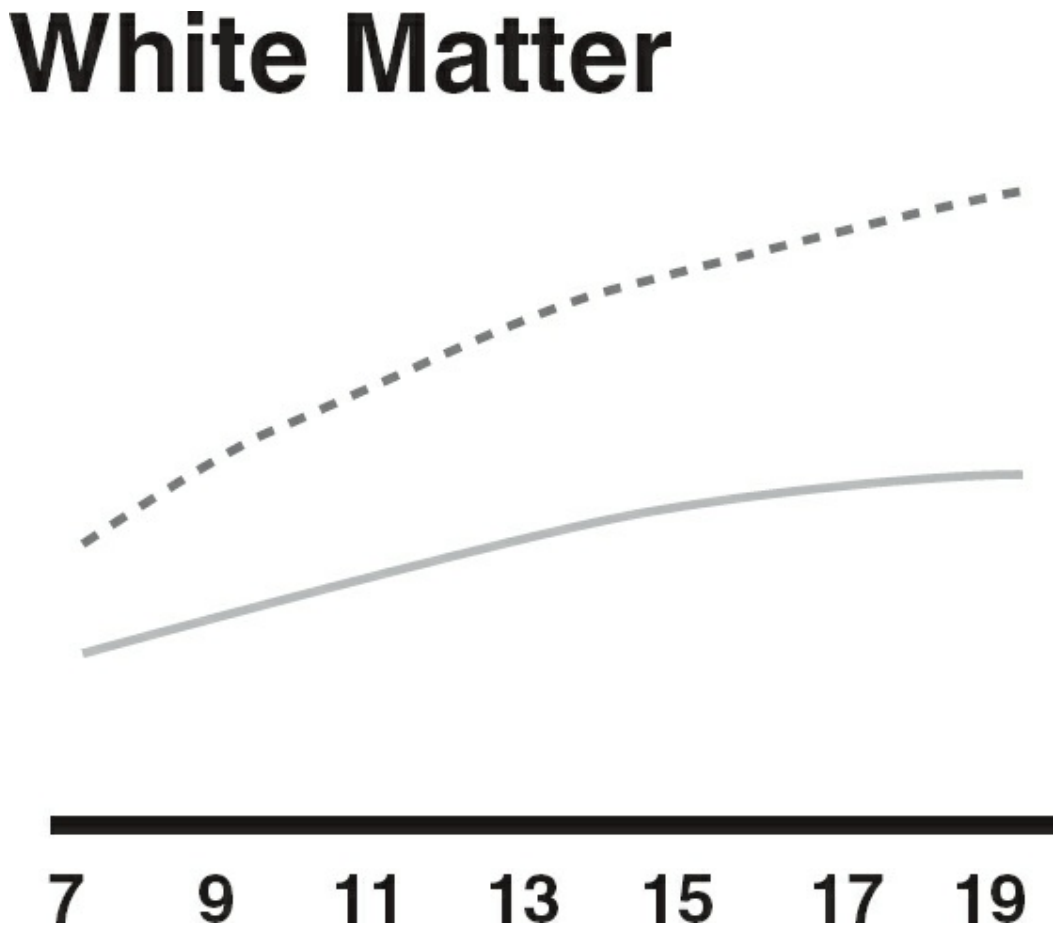

FIGURE 38–6. Schematic showing growth and aging changes in white matter from childhood through adolescence. White matter volume increased rapidly until 12 to 15 years of age then slowed to eventually plateau. Males (*dotted line*) females (*solid line*). (Schematic adapted from Giedd JN, Raznahan A, Alexander-Bloch A. Child psychiatry branch of the National Institute of Mental Health longitudinal structural magnetic resonance imaging study of human brain development. *Neuropsychopharmacology*. 2015;40(1):43–49.)

Gender Differences

Gender-specific and age-related differences in gray matter volumes have also been noted in the basal ganglia, hippocampus, and amygdala (Table 38–2). Although reduction of basal ganglia volumes has been demonstrated in both sexes through young adulthood, gray matter loss in this region is most dramatic in boys, especially in the caudate between the ages of 6 to 12 years. It is possible that this maturational difference is related to the preponderance of basal ganglia–associated psychiatric disease (e.g., OCD and ADHD) in boys during this critical developmental

window. In contrast to the basal ganglia and some areas of the cerebral cortex, the volume of the amygdala and hippocampus increases with age. Volumetric and functional abnormalities of these regions have been associated with depression and anxiety disorders, known to occur more commonly in females, particularly after the onset of puberty. The relatively smaller amygdala and larger hippocampal volumes that emerge in girls than in boys around the time of puberty may be particularly relevant in this regard. The preponderance of estrogen receptors in the amygdala and androgen receptors in the hippocampus suggests that gonadal hormones may influence age-specific changes in these regions and possibly the emergence of mood and anxiety disorders in adolescent girls. The pituitary gland also demonstrates a striking gender difference, with females developing a larger pituitary gland than males, particularly during adolescence. This may be associated with the differential timing of puberty in males and females.

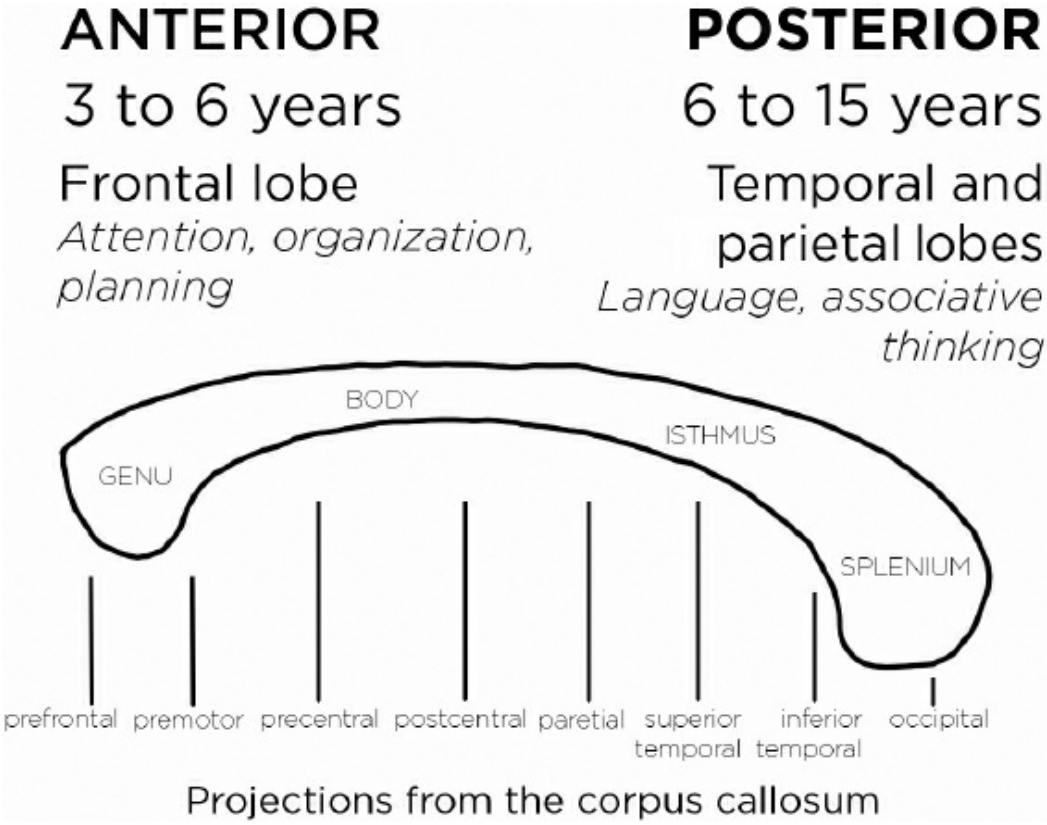

FIGURE 38–7. Schematic showing the corpus callosum, its subdivisions, and projections. (Schematic adapted from Thompson PM, Giedd HN, Woods RP, MacDonald D, Evans AC. Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature.* 2000;404(6774):190–193.)

Table 38–2.
Gender Differences in Brain Volumes

Region	Gender Difference (Age in years)

Total brain volume	10% larger in boys (4–18)
Cerebellum	8% larger in boys (4–18)
Caudate and putamen	More dramatic decrease in boys (preadolescent)
Amygdala	Larger in girls (postadolescent)
Hippocampus	Larger in boys (postadolescent)

Data summarized from Giedd JN, Raznahan A, Alexander-Bloch A. Child psychiatry branch of the National Institute of Mental Health longitudinal structural magnetic resonance imaging study of human brain development. *Neuropsychopharmacology*. 2015;40(1):43–49.

SPECIFIC NEUROPSYCHIATRIC DISORDERS

Attention-Deficit/Hyperactivity Disorder

The phenomenology, nosology, and empirically based treatment for ADHD have been well delineated. This makes pediatric ADHD a prime candidate for neurobiological studies of development. ADHD is also relatively consistent in expression across the lifespan. Therefore, the results of studies of the disorder in children, closer to illness onset, are applicable to adult subjects. Various neuroimaging strategies, including PET, fMRI, volumetric/anatomic MRI, and ^{1}H-MRS, have been used to better elucidate the pathogenesis and mechanisms of treatment response in ADHD.

Prefrontal Cortex. Core deficits in attention, arousal, and impulse control characterize children with ADHD. The PFC modulates these functions and is the seat of "executive" control. It is not surprising that the PFC has been implicated in the pathophysiology of ADHD. Smaller prefrontal cortical volumes, particularly of the superior and inferior frontal gyri, have been consistently noted in volumetric MRI studies of children with ADHD. However, one study found girls with ADHD to have normal frontal lobe volumes; this finding may be related to the lower prevalence of ADHD in females than males and an underlying gender difference in the pathophysiology of the disorder. Lower global cerebral metabolism in girls with pediatric ADHD than in male patients and normal controls has been noted, further suggesting gender-specific differences in the neurobiology of ADHD. Smaller PFC volumes correlate with impaired response inhibition in children with ADHD. A recent large-scale study of children with ADHD found that the structural maturation of the PFC lags behind typically developing children by 2 to 3 years. When considering the potential behavioral ramifications of such a lag in development, the implications for the symptoms of ADHD are clear. In addition, using phosphorus magnetic resonance spectroscopy (^{31}P-MRS), lower levels of membrane phospholipid (MPL) precursors were observed in older ADHD children as compared to healthy children. Lower MPL precursor levels suggest a possible underdevelopment of neuronal processes and synapses in ADHD. It was suggested that bottom-up developmental dysfunction, starting in the basal

ganglia, might result in ADHD (Fig. 38–8). During a response conflict task that typically involves the anterior cingulate (AC) in normal subjects, adult patients with ADHD failed to demonstrate a significant AC response. Of interest, both increased and decreased PFC activation during response inhibition has been found in different studies. These discrepant findings may be accounted for by the inverse relationship noted between symptom severity and perfusion in children with ADHD. Specifically, perfusion-based functional neuroimaging studies of "sicker" patients would start from a lower baseline than studies of less symptomatic patients. This may result in a reduction in signal when the PFC is involved.

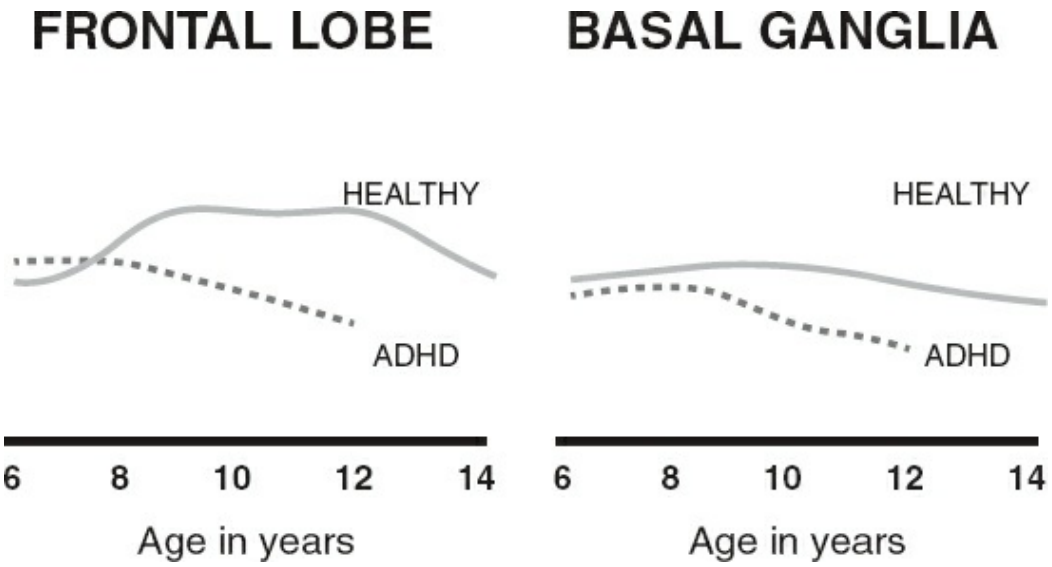

FIGURE 38–8. Schematic showing membrane phospholipid (MPL) precursor levels as associated with age in children with attention-deficit/hyperactivity disorder (ADHD) (*broken line*) and typically developing healthy children (*solid line*) in the frontal lobe and basal ganglia. (Schematic adapted from Stanley JA, Kipp H, Greisenegger E, et al. Evidence of developmental alterations in cortical and subcortical regions of children with attention-deficit/hyperactivity disorder: a multivoxel in vivo phosphorus 31 spectroscopy study. *Arch Gen Psychiatry.* 2008;65(12):1419–1428.)

Basal Ganglia. Another aspect of the frontal-striatal loop, the basal ganglia, has been repeatedly demonstrated to be abnormal (with regard to volumes and reversed asymmetry) in pediatric ADHD. The deficits in response inhibition and attention noted in children with ADHD have also been associated with smaller volumes and lower activation of the right and left caudate. The basal ganglia has been implicated in ADHD, TS, and OCD, and may account for the high degree of comorbidity among these conditions. ^{1}H-MRS studies have noted lower levels of the putative neuronal marker *N*-acetylaspartate (NAA) and greater levels of choline (Cho), a critical membrane component, in the striatum bilaterally in pediatric ADHD patients. More recently, ^{1}H-MRS studies of pediatric ADHD patients identified greater frontal-striatal glutamatergic levels in medication-free pediatric patients with ADHD 7 to 16 years of age as

compared to age- and sex-matched controls. Lower NAA levels suggestive of reduced neuronal viability may result from excess activity of glutamatergic (Glx) neurons, which can be toxic. Given evidence that changes in Cho, NAA, and Glx may be reversible in other neuropsychiatric conditions; longitudinal ^{1}H-MRS studies of these metabolites in ADHD are critical. As mentioned above, ADHD patients have lower PFC and basal ganglia MPL precursor levels than healthy children. In the basal ganglia, these deficits appear earlier, possibly indicating a bottom-up change in developmental trajectory. No significant differences were noted in the superior temporal region.

Cerebellum and Other Regions. Structural neuroimaging studies have demonstrated smaller cerebellar volumes in girls and boys with ADHD. This smaller volume may be localized to the vermis, a brain region involved in maintaining attention. In an fMRI study of motor function in pediatric ADHD patients, greater impairment in task performance was associated with lower cerebellar activation. Furthermore, motor impairment in ADHD patients was also associated with lower temporal lobe activation but greater thalamic activation. This suggests that complex network-wide dysfunctions might underlie the symptoms of ADHD. This hypothesis is supported by the findings of smaller anterior and poster aspects of the CC in ADHD patients.

Medication Effects on Brain Function in ADHD. Methylphenidate (MPH) is the first-line treatment for ADHD. Studies found that MPH changes frontal-striatal activation in pediatric patients with ADHD, with greater PFC and striatal activation as measured by fMRI demonstrated in ADHD patients. In addition, MPH was found to improve ADHD patient performance on the response inhibition task. Predrug activation during the task was greater among patients than with controls, suggesting that frontal cortical hyperactivity in ADHD may play a compensatory role that is further enhanced by MPH. In the striatum, predrug activation was lower in pediatric patients with ADHD than in pediatric controls. This deficit appeared to be corrected by MPH. MPH-treated healthy children also improved in their ability to inhibit inappropriate response. However, improved performance in healthy children was associated with a distinct pattern of frontal-striatal activation. These differences in MPH-related activation could reflect a fundamental abnormality of frontal-striatal dopamine systems in patients with pediatric ADHD as compared to healthy children. Positive correlation between ADHD symptom severity and increased midbrain accumulation of ^{18}F-dihydroxyphenylalanine (^{18}F-DOPA), a marker of dopamine synthesis, lends further support to the role of dysfunctional dopamine systems in ADHD. There is data suggesting a developmental shift in dysfunction of dopamine from the midbrain in

childhood and adolescence to frontal cortex during adulthood in ADHD. Interestingly in children with ADHD, responders to methylphenidate have greater gray matter concentrations compared to nonresponders in the caudate. This is similar to a larger posterior hippocampus predicting better response to antidepressant in adults with depression. The predictive power of imaging for treatment response has yet to be fully explored in child and adolescent psychiatry.

A recent meta-analysis of fMRI studies of the effect of stimulants on brain function in ADHD examined 14 data sets with 212 children with ADHD. Stimulants consistently enhanced right inferior frontal cortex and insula activation, with a smaller effect on putamen activation. This change in activation was present in those receiving methylphenidate alone. The inferior frontal cortex and insula play a role in cognitive control and an increase in activation may represent the positive clinical response found with treatment.

Depressive Disorders and Bipolar and Related Disorders

Biological studies of depressive disorders and bipolar and related disorders are providing increasingly converging evidence that alterations in medial temporal (e.g., amygdala, hippocampus) and prefrontal-striatal circuits may be involved in the symptomatology of patients with MDD and BP-I (Fig. 38–9). Although the literature is not entirely consistent, it appears that lesions to the left hemisphere are more associated with depression, whereas right-sided lesions tend to be more associated with mania.

Frontal Lobe and Striatum. In both adults and adolescents with MDD, smaller overall frontal lobe volumes have been among the more consistent findings in structural neuroimaging studies. In addition, adolescents with MDD were found to have significantly smaller frontal white matter volumes but larger frontal gray matter volumes than healthy comparison adolescents, even after controlling for age and whole brain. Cortical thickness in the middle frontal gyrus bilaterally and the left caudal anterior cingulate was found to be greater in adolescents with MDD as compared to healthy control participants. Replicating and extending this finding in adults, thicker caudal anterior cingulate was also found with first episode MDD. Left middle frontal gyrus was inversely correlated with age in MDD subjects. This may indicate an altered developmental path in youth with MDD. MRI and PET investigation in adults with MDD and BP-I have reported strikingly localized, left-sided smaller volume in the subgenual region of the prefrontal cortex (SGPFC). This smaller SGPFC was associated with lesser cerebral blood flow. These effects were most prominent in MDD and BP-I patients who had at least one first-degree relative with MDD or BP-I (familial MDD or BP-I). In patients with MDD or BP-I without family history of MDD or BP-I (nonfamilial MDD or BP-I)

no significant differences were observed in the left or right SGPFC. In an additional sample, a smaller left SGPFC volume was also noted in 30 young women with early-onset MDD (18 to 23 years of age) and 18 middle-aged women with recurrent MDD (24 to 52 years of age) than in age-matched controls (19 percent difference).

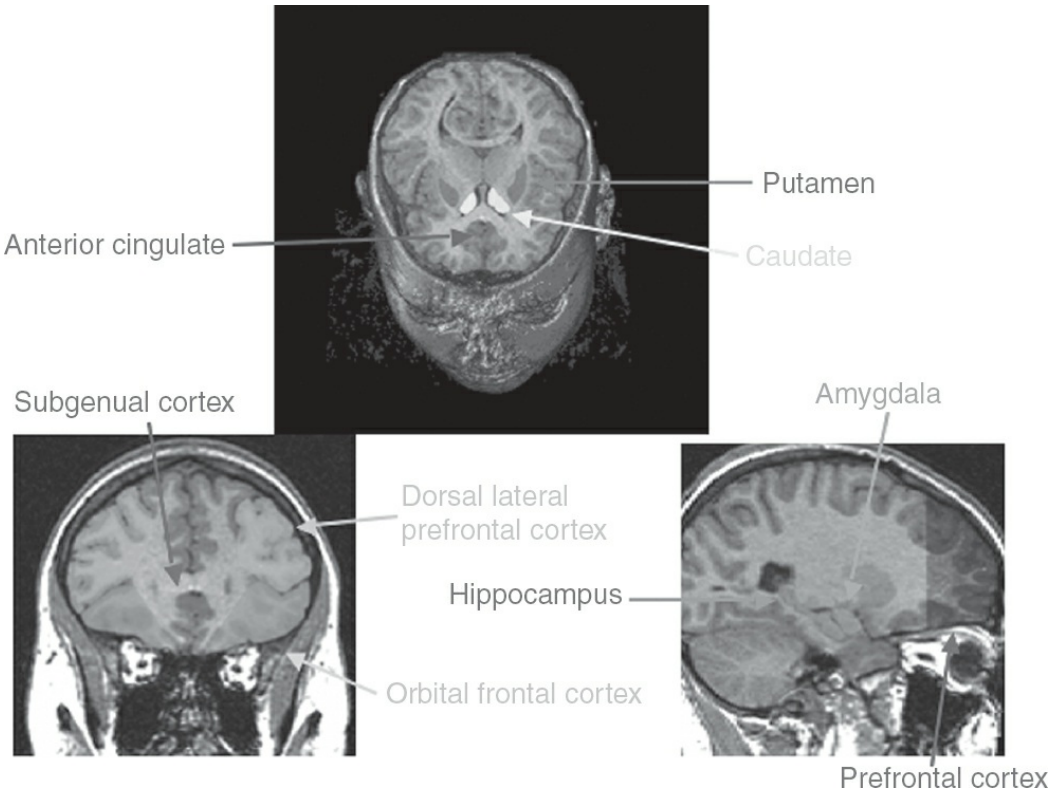

FIGURE 38–9. Brain regions implicated in the pathophysiology of major depressive disorder.

A study of 22 psychotropic-naive pediatric patients with MDD (9 to 17 years) compared to 22 case-matched controls found that patients with nonfamilial MDD had significantly larger left-sided but not right-sided prefrontal cortical volumes than patients with familial MDD and controls. In both the familial and nonfamilial pediatric MDD subgroups, left prefrontal volumetric abnormalities remained significant after controlling for severity of anxiety. This suggests that this region is integral to the pathogenesis of depression in children. Reduced left-sided prefrontal cortical volume was associated with increased severity of depression in familial but not nonfamilial patients with MDD. Longer illness duration also tended to be associated with reduced left prefrontal cortical volume in familial but not nonfamilial MDD patients. These results, coupled with investigations in adults, suggest that left-sided prefrontal cortical volume reduction in familial MDD patients may reflect degeneration of left prefrontal cortical volume, whereas increased left prefrontal cortical volume in psychotropic-naive pediatric patients with nonfamilial MDD

may reflect alteration in the developmental maturation of the prefrontal cortex.

Left PFC has been further implicated in pediatric MDD in two recent ^{1}H-MRS investigations. Increased left orbitofrontal choline levels were observed in 22 depressed adolescents as compared to 43 control adolescents. Increased orbital frontal choline levels were most prominent in adolescent patients with familial MDD as compared to both nonfamilial MDD patients and controls (Personal Communication, Dr. Perry Renshaw). Increased left but not right dorsolateral PFC choline concentrations were also observed in treatment-naive pediatric patients with MDD as compared to age- and sex-matched controls. Increased choline concentrations in pediatric MDD patients versus controls persisted even after controlling for severity of anxiety. Mean left dorsolateral prefrontal cortical choline concentrations were increased in familial MDD patients as compared to both nonfamilial MDD patients and healthy pediatric controls. Lower choline concentrations were also observed in the right anterior cingulate. Alterations in choline as measured by ^{1}H-MRS have been hypothesized to result from abnormalities in signal transduction involving phosphatidylcholine second messenger cascades and could reflect changes in glucose metabolic rates and/or neuroendocrine function. Increased left prefrontal choline levels may also be consistent with functional neuroimaging studies demonstrating reduced prefrontal cortical metabolism in adults with MDD. Functional neuroimaging studies have demonstrated that reduced prefrontal cortical activation in depressed adolescents and adults normalizes following treatment-induced symptom remission. Because choline alterations in adult patients with MDD appear to be reversible with treatment, the choline resonance may represent a potential target for treatment intervention in pediatric MDD, warranting longitudinal ^{1}H-MRS investigations in these patients. Finally, consistent with findings in adult patients with MDD, lower anterior cingulate Glx concentrations have been noted in pediatric MDD patients than in healthy controls. Glx was also associated with severity of illness in the MDD patients. In addition, distinct differences were noted in familial versus nonfamilial MDD patients. Subsequent investigation, using a more robust spectral fitting method, found that the lower anterior cingulate Glx in patients was due to glutamate but not glutamine. In 20 medication-free youth with MDD, lower anterior cingulate GABA was observed as compared to controls, with this difference being more pronounced in those with anhedonia. Recently, a small case series in adolescents with MDD replicated a finding in adults with MDD that showed an increase in DLPFC glutamate in response to repetitive transcranial magnetic stimulation treatment.

In contrast, recent ^{1}H-MRS investigation in ten medication-free children with BP-I versus ten healthy children found increased glutamate

and glutamine levels bilaterally in the frontal lobes and basal ganglia in children with BP-I. Choline levels did not differ in frontal lobes or basal ganglia in patients with BP-I or controls. Increased lipid levels were also found bilaterally in the frontal lobes of children with BP-I. In contrast, ^{1}H-MRS investigation of anterior cingulate cortex in pediatric patients with BP-I versus healthy pediatric controls found no significant differences in choline, glutamate, glutamine, GABA, or the neuronal marker NAA. Consistent with findings in adults, effective lithium treatment was, however, associated with reduction of the ^{1}H-MRS neurochemical marker myo-inositol in the anterior cingulate cortex of children with BP-I. A recent voxel-based morphometry study found that subjects with BP-I had reduced gray matter volume in the left DLPFC as compared to healthy controls.

In concert with the frontal lobe, the striatum plays an important role in executive and emotional processing. Evidence from adult imaging studies implicates the striatum in the pathophysiology of depressive disorders. A recent study noted that patients with MDD had significantly smaller right striatum (2.3 percent smaller) and caudate volumes (3.9 percent smaller) than healthy subjects. Furthermore, right striatum volumes were inversely correlated with increased severity of depression. These findings provide fresh evidence for abnormalities in the striatum of pediatric MDD and bolster the hypothesis of dysfunction in the cortico-limbic-striatal-pallidal-thalamic circuit in the pathophysiology of MDD.

Amygdala, Hippocampus, and Pituitary. Although prefrontal cortical involvement in pediatric MDD may be specific to depressive state, abnormalities of the amygdala in depressed children and adolescents may be more associated with severity of anxiety than with severity of depression. Ratios of left and right amygdala to hippocampal volume were found to be significantly larger in treatment-naive pediatric patients with MDD than in 23 age- and sex-matched controls. Furthermore, ratios of amygdala to hippocampal volume in pediatric patients with MDD were positively correlated with severity of anxiety but not with severity of depression or duration of depressive illness. Amygdala volumes were larger bilaterally in patients with MDD than in controls. In a magnetic resonance spectroscopic imaging ^{1}H-(MRSI) study of the amygdala in adolescents with depression, lower choline-to-creatine ratios were noted in the left, but not right, amygdala in patients than in controls. This change in cytosolic choline may be secondary to corticosteroid-neuroendocrine–driven changes, given that increases in cortisol can result in lower choline concentrations. Right and left hippocampal volumes were smaller in MDD patients. Not only are comorbid anxiety symptoms and disorders extremely common in patients with MDD, but also excessive fearfulness in adolescence has been found to predict the development of MDD in early adulthood. Smaller left hippocampal volumes have been noted in pediatric

patients with MDD with a familial history of depression than in controls. It may be that patients with a family history of depression are driving the smaller hippocampal volumes noted in previous reports. In a closer examination of the hippocampus using ^{1}H-MRSI, reduced left, but not right, NAA concentration was noted in MDD patients as compared to age- and sex-matched controls. This indicates that neuronal viability may be reduced in pediatric MDD.

In adolescents with MDD, at rest, the amygdala showed lower connectivity with the hippocampus, parahippocampus, and brain stem. This connectivity was inversely correlated with depressive symptoms. Use of emotional facial expressions as fMRI stimuli is common as a probe of affective processing in the amygdala in adults and children. An fMRI study of amygdala response to fearful faces in children with generalized anxiety disorder and panic disorder compared to children with MDD and healthy pediatric controls demonstrated altered activation of the amygdala in both patients with MDD and patients with generalized anxiety disorder/panic disorder, with the magnitude of signal change in amygdala activation being most associated with the severity of the child's rating of everyday anxiety. This increase in amygdala arousal to negative stimuli was recently replicated in 32 adolescents with MDD and 23 age and gender-matched controls.

It should be noted that greater amygdala volume is one of the more consistent findings in adults with BP-I. fMRI studies in adults with BP-I have also demonstrated increased amygdala activation in response to fearful facial affect. More recently, an event-related fMRI study reported that pediatric patients with BP-I demonstrated reduced memory for emotional faces and had increased neural activation in the striatum and anterior cingulate cortex when successfully encoding happy faces, and in the orbitofrontal cortex when successfully encoding angry faces. There were no between-group differences in neural activation during fearful face encoding. Consistent with adult BP-I studies, multiple fMRI studies reported increased amygdala activation in response to fearful faces. Differences across studies of BP-I may have to do with differences in disease state. Longitudinal MRI studies of amygdala volume and function in pediatric patients with MDD are clearly warranted to determine whether increased amygdala volume and altered activation level may serve as potential biomarkers for BP-I.

More recently, larger pituitary gland volumes were noted in adolescents with MDD than in case-matched controls. This was replicated in an independent treatment-naive sample. In the second sample, it was noted that the difference in volume appeared to be driven by nonfamilial MDD patients. Of interest, depressed patients with BP-I demonstrated larger pituitary gland volumes than controls. This was not consistent with a published adult study of BP-I and a pediatric sample. The main difference

between these reports is the disease state (depressed vs. mainly euthymic). This reinforces the need for studies of BP-I across the illness course and the need for further studies to differentiate MDD from BP-I.

Obsessive-Compulsive Disorder

There is a growing literature pointing to a dysfunction in ventral prefrontal-striatal-thalamic and medial temporal circuitry in OCD (Fig. 38–10). The phenomenology, nosology, and empirical treatment for childhood OCD have been well established, rendering it an ideal candidate for developmental neurobiological study. In comparison to depressive and bipolar and related disorders, OCD is more consistent in its expression across the lifespan. This allows investigators to study the illness close to onset, with the results remaining applicable to adult subjects. Studies in pediatric OCD have focused on treatment-naive populations. This minimizes potential confounds such as illness duration and medication exposure. As described in the following sections, evidence is converging that suggests a pharmacologically reversible, glutamatergically mediated thalamic-cortical-striatal dysfunction in pediatric OCD that may serve as a reliable pathophysiological and treatment response marker for the illness (Fig. 38–11).

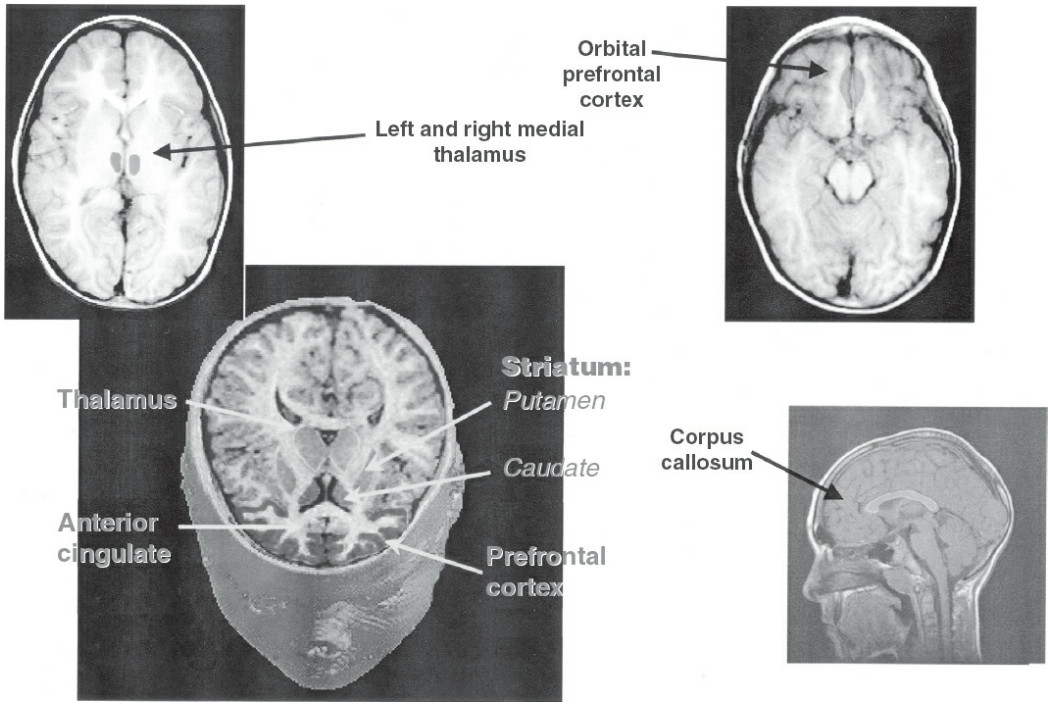

FIGURE 38–10. Brain regions implicated in the pathophysiology of obsessive-compulsive disorder. (Reprinted with permission from Rosenberg DR, MacMillan SN, Moore GJ. Brain anatomy and chemistry may predict treatment response in pediatric obsessive-compulsive disorder. *Int J Neuropsychopharmacol.* 2001;4(2):179–190.)

Basal Ganglia. Volumetric neuroimaging studies in pediatric patients

with OCD have demonstrated smaller basal ganglia volumes in patients that were associated with OCD symptom severity but not illness duration. In addition, smaller globus pallidus volumes were also noted in treatment-naive pediatric OCD patients than in controls. Recently however, 26 adolescents with OCD showed greater gray matter volume in the bilateral caudate as compared to healthy controls (N = 26). This may reflect fundamental differences in what properties are measured using voxel-based morphometry as compared to the manual tracing technique. Greater caudate glutamatergic concentrations, as measured by ^{1}H-MRS, were also observed in treatment-naive pediatric OCD patients than in age- and sex-matched healthy pediatric controls (Figs. 38–12 and 38–13). No significant differences in glutamatergic concentrations were observed between OCD patients and controls in occipital cortex, a region not implicated in the pathogenesis of OCD. After 12 weeks of treatment with the selective serotonin reuptake inhibitor (SSRI) paroxetine, but not with cognitive–behavioral therapy, caudate glutamatergic concentrations in OCD patients decreased to levels that did not differ significantly from those in controls (Fig. 38–14). Greater pretreatment caudate glutamatergic concentrations predicted a more robust response to paroxetine in OCD patients. More recent investigation suggested that caudate glutamatergic concentrations may persist even after medication discontinuation in pediatric OCD.

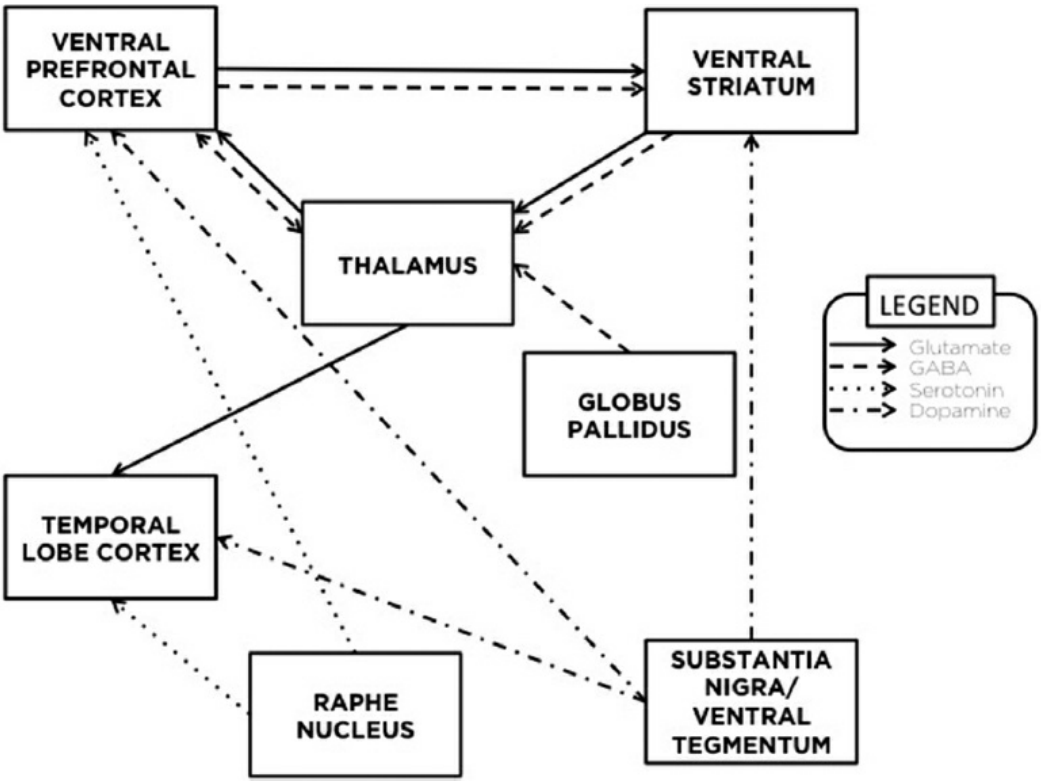

FIGURE 38–11. Schematic diagram showing selected aspects of corticostriatal connections in the neurodevelopment of obsessive-compulsive disorder. (Figure courtesy of Frank MacMaster, Ph.D.)

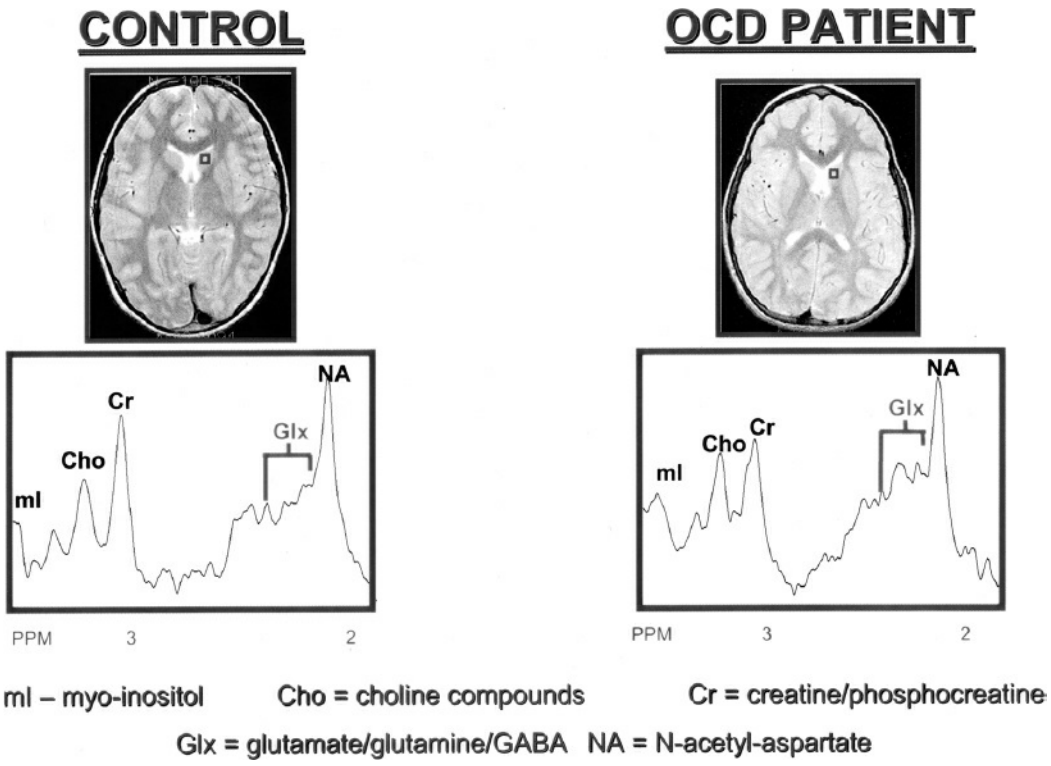

FIGURE 38–12. **Left:** Illustration of voxel placement in left caudate nucleus. **Right:** ^{1}H-Magnetic resonance spectroscopy of a 0.7-mL volume of interest centered in the left caudate in a 10-year-old healthy control and a 9-year-old treatment-naive patient with obsessive-compulsive disorder (OCD) as shown on the T1-weighted magnetic resonance images. GABA, γ-aminobutyric acid. (Reprinted with permission from Rosenberg DR, MacMillan SN, Moore GJ. Brain anatomy and chemistry may predict treatment response in paediatric obsessive-compulsive disorder. *Int J Neuropsychopharmacol.* 2001;4(2):179–190.)

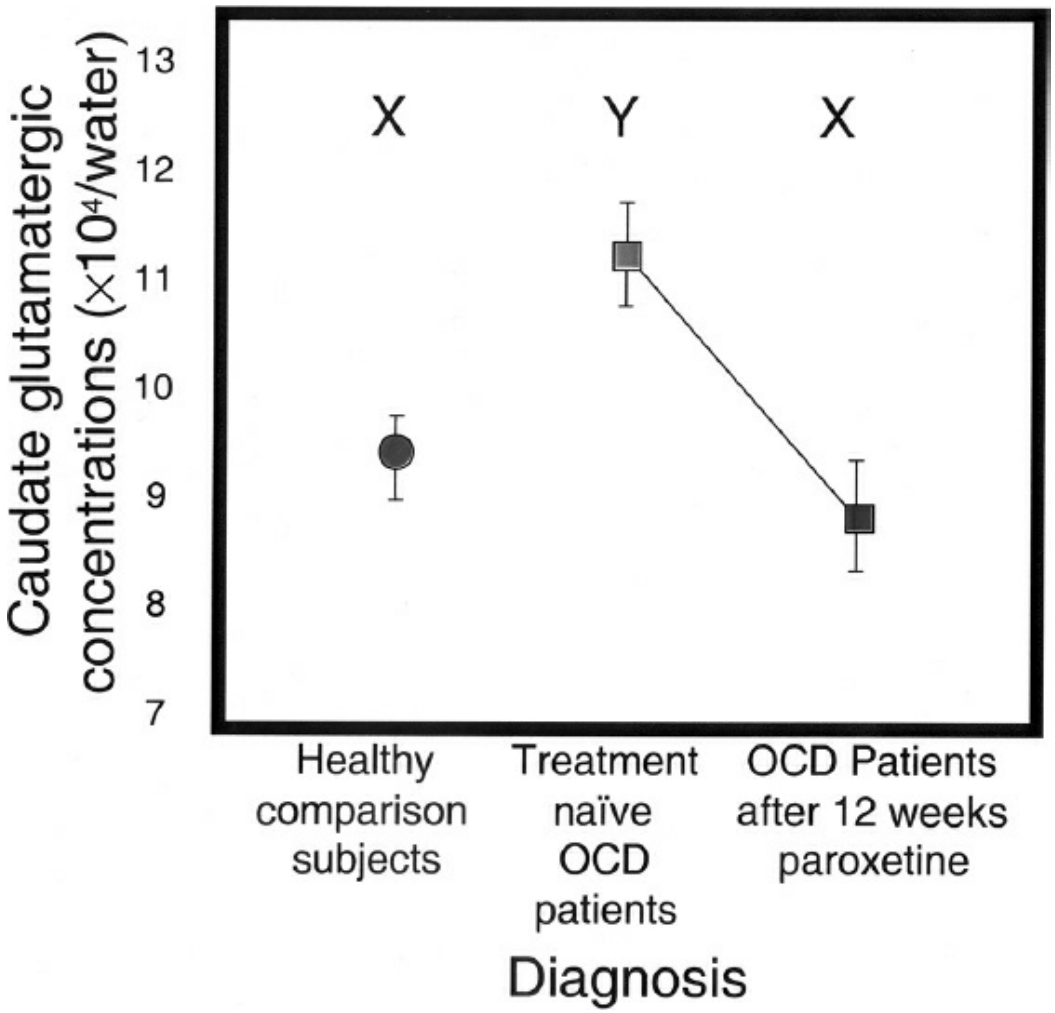

FIGURE 38–13. Caudate glutamatergic concentration by diagnostic and treatment condition. Groups not sharing the same letter are significantly different at $p < 0.05$. OCD, obsessive-compulsive disorder. (Adapted from Rosenberg DR, MacMaster FP, Keshavan MS, Fitzgerald KD, Stewart CM. Decrease in caudate glutamatergic concentrations in pediatric obsessive-compulsive disorder patients taking paroxetine. *J Am Acad Child Adolesc Psychiatry.* 2000;39(9):1096–1103.)

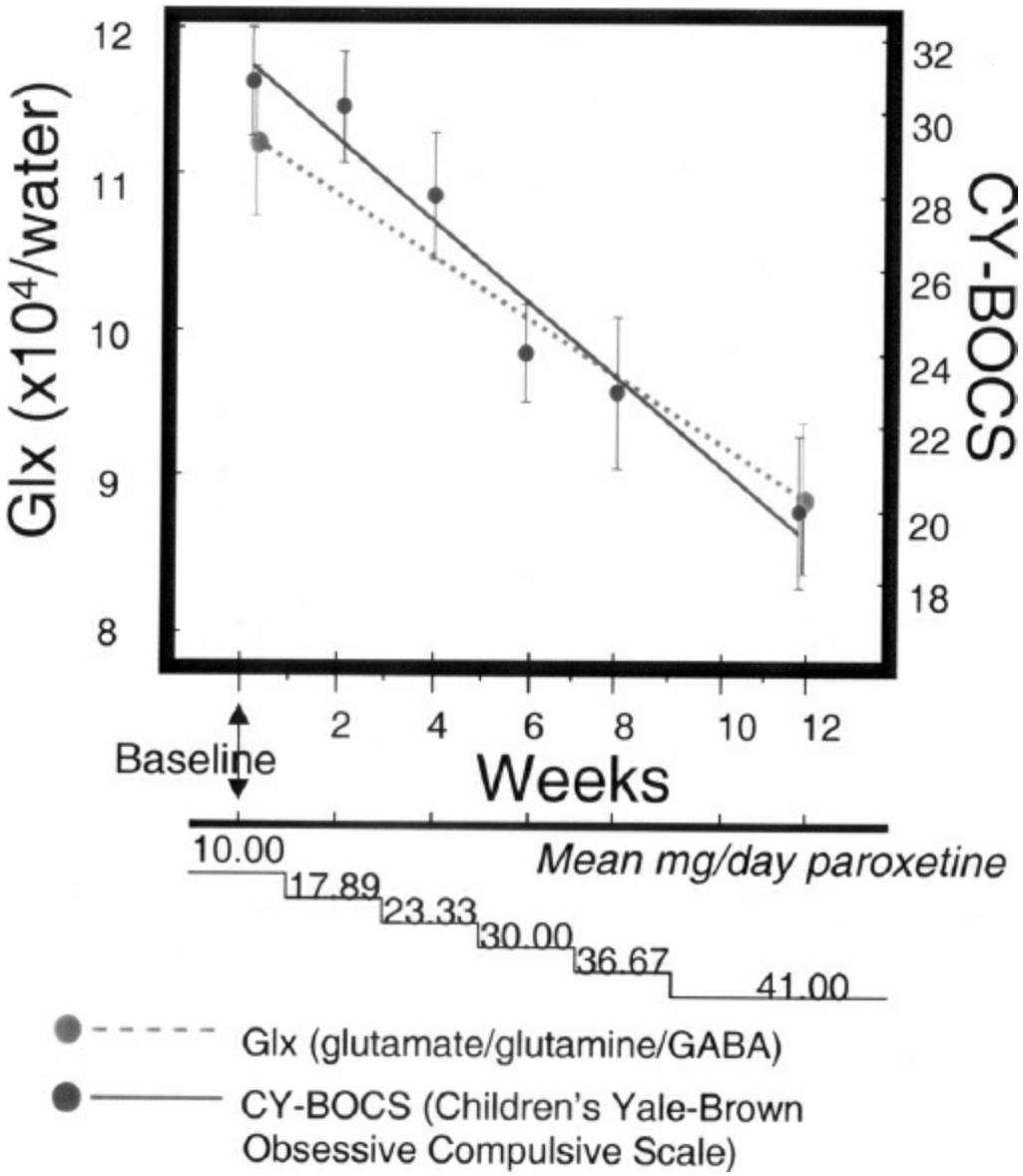

FIGURE 38–14. Decrease in left caudate glutamatergic concentrations associated with reduction in obsessive-compulsive score of the Children's Yale-Brown Obsessive Compulsive Scale. (Reprinted with permission from Rosenberg DR, Hanna GH. Genetic and imaging strategies in obsessive-compulsive disorder: Potential implications for treatment development. *Biol Psychiatry.* 2000;48(12):1210–1222.)

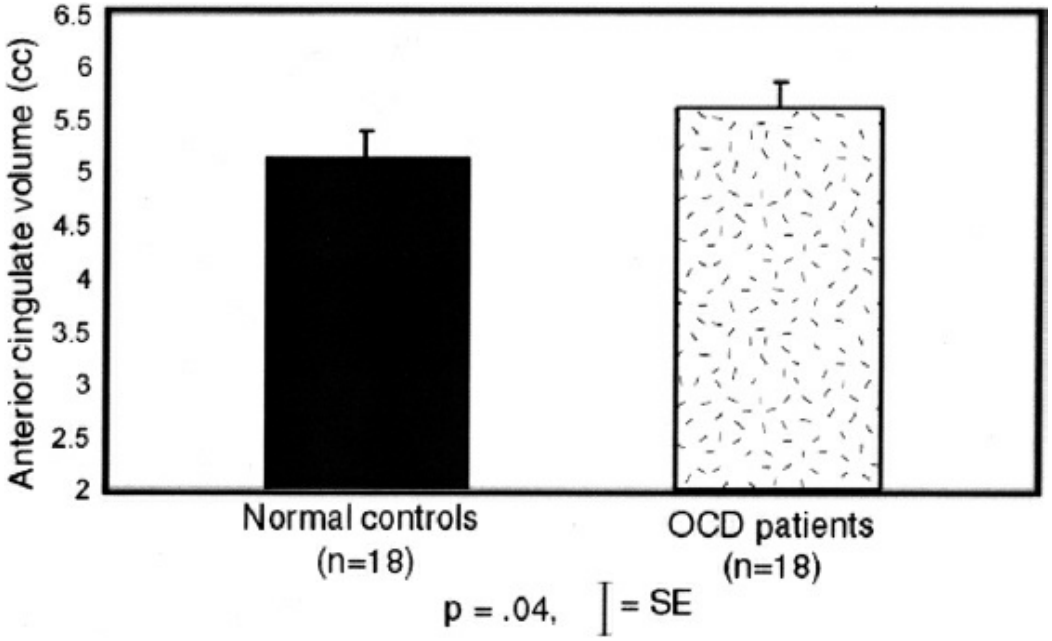

FIGURE 38–15. Anterior cingulate volume by group. OCD, obsessive-compulsive disorder. (Reprinted with permission from Rosenberg DR, Keshavan MS. Toward a neurodevelopmental model of obsessive-compulsive disorder. *Biol Psychiatry.* 1998;43(9):623–640.)

Prefrontal Cortex. MRI investigation in treatment-naive pediatric patients with OCD as compared to age- and sex-matched healthy controls demonstrated larger anterior cingulate volumes but not dorsolateral prefrontal cortical volume in the OCD patients (Fig. 38–15). These findings are consistent with findings of increased regional cerebral blood flow and metabolic activity in anterior cingulate cortex in adult and pediatric patients with OCD. Larger anterior cingulate volumes were associated with increased OCD symptom severity but not duration of illness. Pediatric OCD patients also failed to demonstrate the normal age-related growth in anterior cingulate volume seen in healthy controls (Fig. 38–16). Anterior cingulate volumes were inversely correlated with striatal volumes in the OCD patients. It should be noted that two independent studies noted larger anterior cingulate volumes in pediatric OCD, with the more recent study finding differences in the anterior cingulate gray but not white matter. Lower anterior cingulate Glx was also seen in OCD patients than in healthy controls. Of interest, the lower anterior cingulate Glx concentration is associated with the G/G genotype of the glutamate receptor gene polymorphism GRIN2B-rs1019385 in OCD patients. A recent fMRI study found greater activation and connectivity of the medial prefrontal cortex/dorsal anterior cingulate in pediatric OCD patients compared to controls. It was suggested that this might reflect an exaggerated concern for correct performance. Studies of the CC found a larger area and lower signal intensity of the genu of the CC in treatment-naive pediatric OCD patients than in controls. The genu is the region of the CC that connects left and

right ventral prefrontal cortical regions such as anterior cingulate cortex. This provides further evidence for the involvement of ventral prefrontal cortical regions, such as the anterior cingulate cortex, in the pathogenesis of OCD. In right prefrontal white matter, pediatric OCD patients demonstrated greater NAA and Choline concentrations when compared to healthy control participants. The greater NAA concentration correlated with OCD symptoms as well. Using the manual tracing technique, greater right white matter volume in the orbital PFC has also been noted. Indeed, a recent examination of frontal-striatal-thalamic white matter found greater age-related increases in fractional anisotropy in pediatric OCD patients as compared to controls. In the cingulum bundle, this correlated with OCD symptom severity.

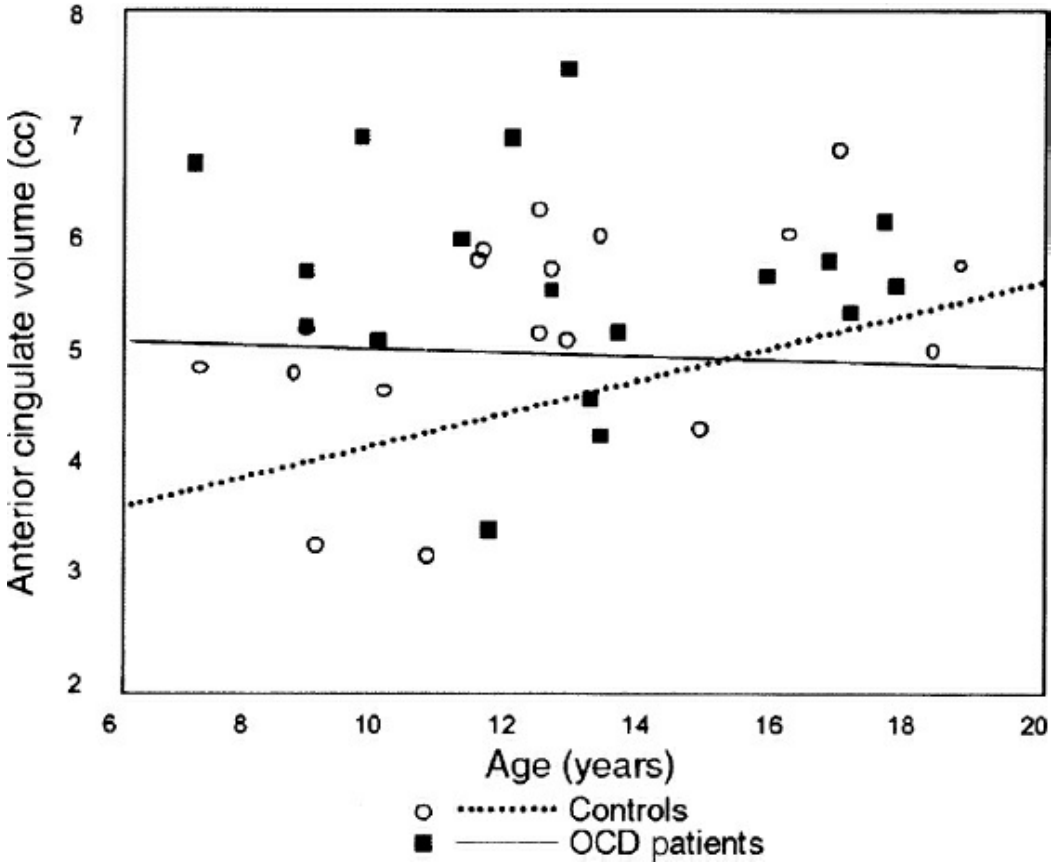

FIGURE 38–16. Anterior cingulate volume as a function of age for pediatric patients with obsessive-compulsive disorder (OCD) and normal controls. Opposite neurodevelopmental trajectories in the patients as compared with the healthy subjects may reflect an absence of early neuronal pruning in children with OCD. (Reprinted with permission from Rosenberg DR, Keshavan MS. Toward a neurodevelopmental model of obsessive-compulsive disorder. *Biol Psychiatry.* 1998;43(9):623–640.)

Thalamus. The thalamus acts as the relay station for neuronal input to the cortex and plays a critical role in the circuitry implicated in imaging studies of OCD. In an MRI study of treatment-naive pediatric patients with OCD as compared to age- and sex-matched healthy controls, OCD patients

were found to have larger thalamic volumes. After therapy with the SSRI paroxetine, thalamic volumes decreased in size to the point at which they no longer differed significantly from those of controls (Fig. 38–17). This reduction in thalamic volume was correlated robustly with a reduction in OCD symptom severity (Fig. 38–18). Indeed, a larger pretreatment thalamic volume predicted a better response to paroxetine treatment. When subdividing the thalamus into its component nuclei, the changes in volume were most prominent in medial rather than lateral thalamic regions. With cognitive–behavioral therapy (CBT), no significant changes in thalamic volume were observed before and after 12 weeks of treatment in treatment-naive pediatric OCD patients. This suggests that the reduction in thalamic volume may be specific to SSRI therapy as opposed to a more general treatment effect. In a closer examination of the thalamus in OCD, a ^{1}H-MRS study of treatment-naive pediatric OCD patients and age- and sex-matched pediatric controls was conducted. Lower levels of the neuronal marker NAA were found in OCD patients than in healthy controls. After 12 weeks of drug therapy with paroxetine, an increase in NAA in the medial but not lateral thalamus was seen. Thus, as enlarged thalamic volumes in OCD patients appear to be reduced toward normal after SSRI treatment, lower levels of NAA may be conversely increasing. This suggests a potential neurotrophic effect of SSRI treatment. In addition, significantly greater left and right medial thalamic Cho concentrations were observed in OCD patients than in both healthy control subjects and patients with MDD. Cho concentrations did not differ between patients with MDD and control subjects. This suggests that alterations in medial thalamic Cho differentiate patients with OCD from healthy control subjects and patients with MDD.

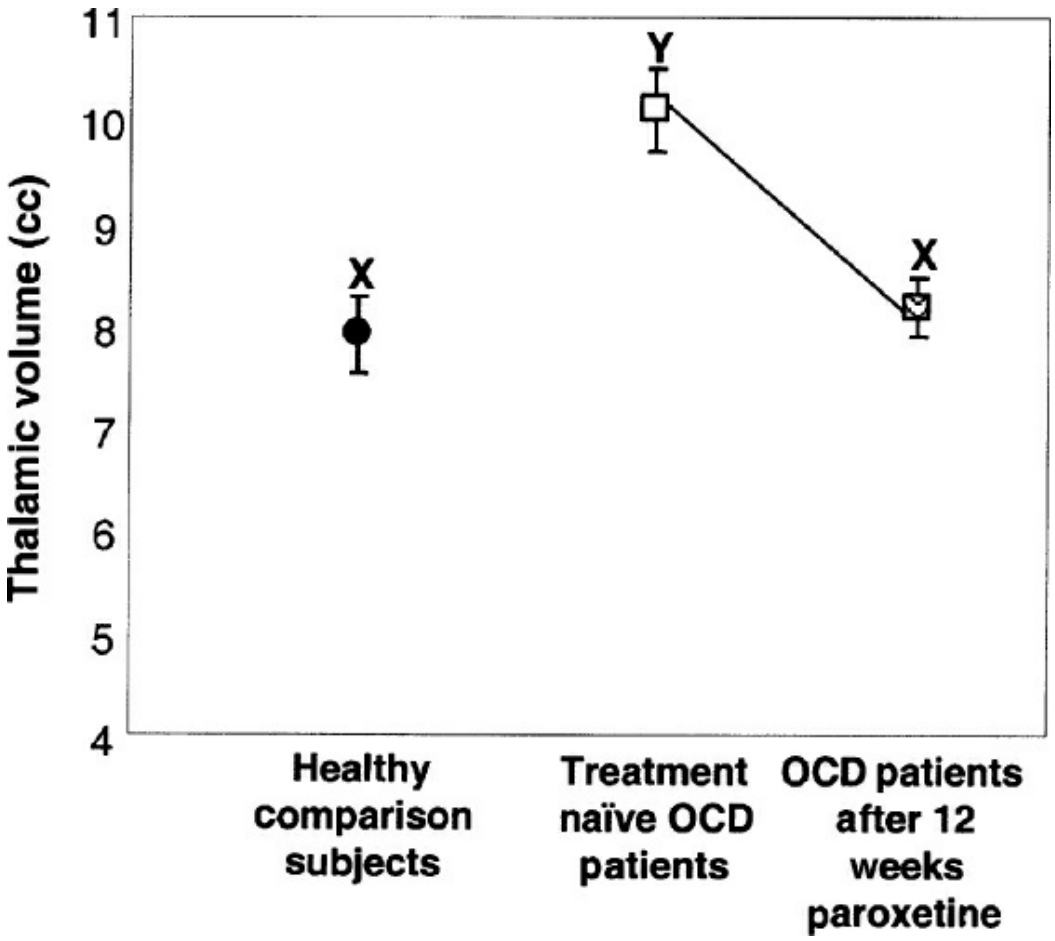

FIGURE 38–17. Data (presented as mean ± standard deviation) for three groups: Healthy controls, pretreatment obsessive-compulsive (OCD) patients, and posttreatment OCD patients. (Figure adapted from Gilbert AR, Moore GJ, Keshavan MS, Paulson LAD, Narula V. Decrease in thalamic volumes of pediatric patients with obsessive-compulsive disorder who are taking paroxetine. *Arch Gen Psychiatry.* 2000;57(5):449–456.)

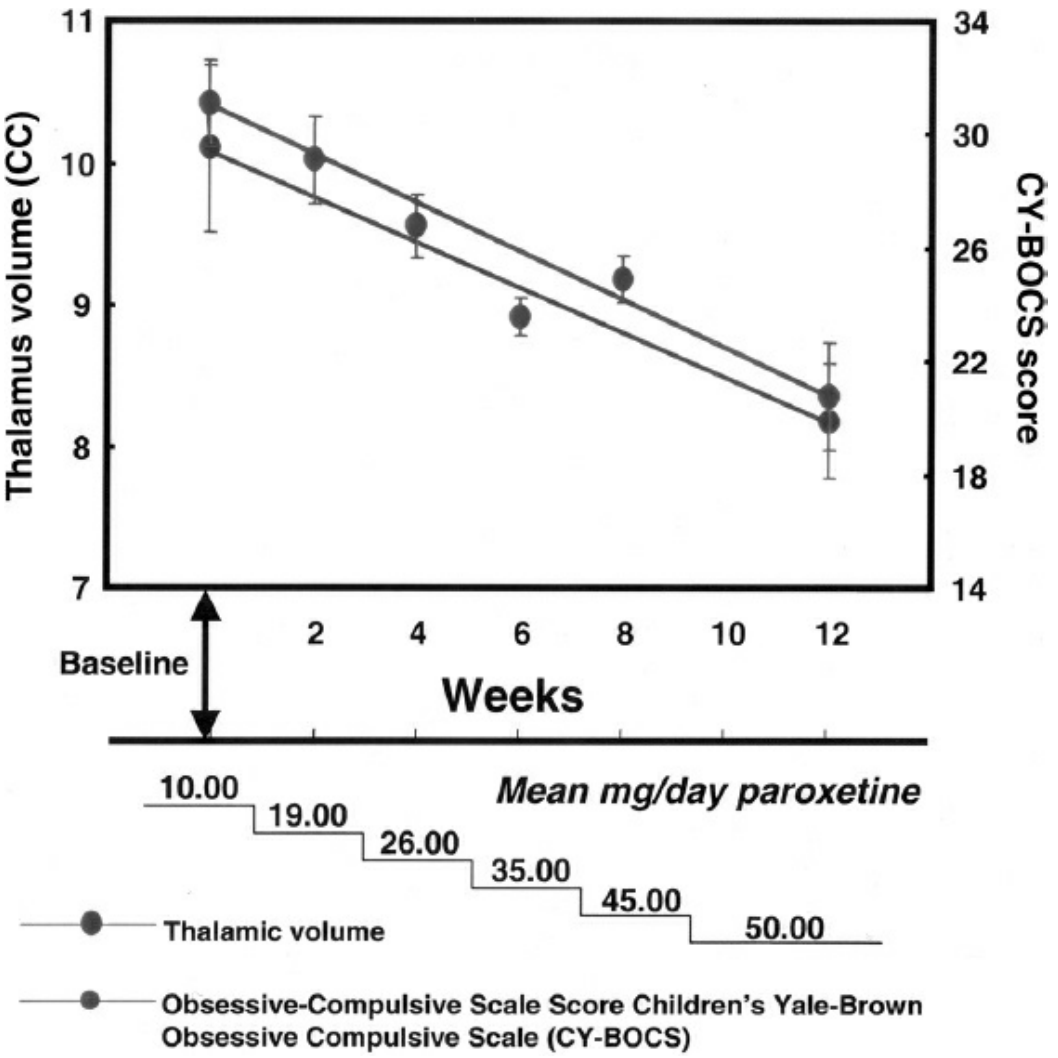

FIGURE 38–18. Decrease in thalamic volume associated with reduction in obsessive-compulsive score of the Children's Yale-Brown Obsessive-Compulsive Scale (CY-BOCS). (Reprinted with permission from Gilbert AR, Moore GJ, Keshavan MS, Paulson LAD, Narula V. Decrease in thalamic volumes of pediatric patients with obsessive-compulsive disorder who are taking paroxetine. *Arch Gen Psychiatry.* 2000;57(5):449–456.)

Amygdala. Given its role in fear conditioning, the amygdala is believed to be involved in the pathophysiology of OCD. The amygdala is also a target of the SSRIs, the mainstays of pharmacotherapy for OCD. In a recent study, amygdala volumes were evaluated for in vivo volumetric changes in 11 pediatric patients with OCD before and after 16 weeks of monotherapy with the SSRI paroxetine. Eleven healthy pediatric control subjects had baseline and follow-up scans, but none of these subjects received medication. OCD patients at baseline demonstrated significant asymmetry of the amygdala (left greater than right) as compared to healthy controls. The volume of the left amygdala decreased significantly following treatment in the OCD patients. This reduction in volume of the left amygdala in patients correlated significantly with higher paroxetine dose at the time of the follow-up scan and total cumulative paroxetine exposure between the scans. No changes in right or left amygdala volume were noted in healthy

controls over time. There are several possible mechanisms that may account for the paroxetine-mediated change in amygdala volume. Disease progression cannot be entirely ruled out, but this is unlikely, given the degree of change noted and the short time in which these changes occurred. The significant inverse relationship between left amygdala reductions and paroxetine treatment lends credence to the idea that the reversal of hypertrophy is due to treatment. Functionally, in most anxiety disorders, an overactive amygdala is often noted. In pediatric OCD however, amygdala activity is reduced in response to emotional stimuli. Caution is noted due to the small sample sizes involved. Further research is warranted.

Pediatric Autoimmune Neuropsychiatric Disorders Associated with Group A a-Hemolytic Streptococcal Infection

Whereas a smaller basal ganglia appears to characterize patients with early-onset OCD, a larger basal ganglia occurs in the subset of patients with pediatric autoimmune neuropsychiatric disorders associated with group A β-hemolytic streptococcal (GABHS) infection (PANDAS). In PANDAS, it is hypothesized that GABHS-associated antibodies cross-react with neurons in the basal ganglia. This causes the observed symptoms of OCD, as well as tics/TS and ADHD in susceptible children. This hypothesis is supported by evidence indicating larger basal ganglia volumes in some patients with OCD, tics, and/or ADHD that are positively correlated with the incidence of GABHS infection and presence of antibody titers. Furthermore, immunotherapy interventions such as plasmapheresis ameliorate OCD symptoms in some PANDAS patients with marked OCD symptoms. Recovery is also associated reduction in basal ganglia volumes. Of interest, thalamic volumes do not differ between PANDAS patients and controls. Although additional studies are needed, this suggests possible unique neurodevelopmental subtypes of OCD.

Anxiety Disorders

Compared with OCD, there have been relatively few neuroimaging studies of pediatric anxiety disorders.

Generalized Anxiety Disorder

In children with generalized anxiety disorder (GAD), regions associated with fear learning and extinction, mentalization, and amygdala regulation (Fig. 38–19) exhibited greater cortical thickness than controls (Fig. 38–20). These areas included the right inferior frontal gyrus, and the left inferior and middle temporal gyri. Furthermore, amygdala intrinsic functional connectivity is disrupted in children with GAD, affecting the amygdala's connections with the medial prefrontal cortex, insula, and

cerebellum. Functionally, greater left medial PFC and right ventrolateral PFC activation was observed in adolescents with GAD in response to a continuous performance task with emotional distractors. Disrupted connectivity arising from the amygdala was also seen. As discussed earlier (see Depressive and Bipolar and Related Disorders), the amygdala has extensive interconnections with cortical-limbic circuitry and plays a significant role in the processing of anxiety or fear. Indeed, fMRI studies in a combined sample of children with GAD and/or panic disorder compared to healthy children found exaggerated amygdala activation in response to fearful faces in patients with GAD. Furthermore, amygdala volume was significantly larger in pediatric patients with GAD than in healthy controls. In response to treatment with medication or cognitive behavior therapy, right ventrolateral PFC activation related to emotional faces increased. In addition to thicker cortices and larger amygdala volumes, subcortical regions are also affected in GAD. The right putamen was also found to be larger in children with GAD. Interestingly, maltreatment during childhood was associated with larger left thalamic volumes in children with GAD as well. Both the putamen and thalamus have been implicated in other anxiety disorders, but their precise role has yet to be delineated.

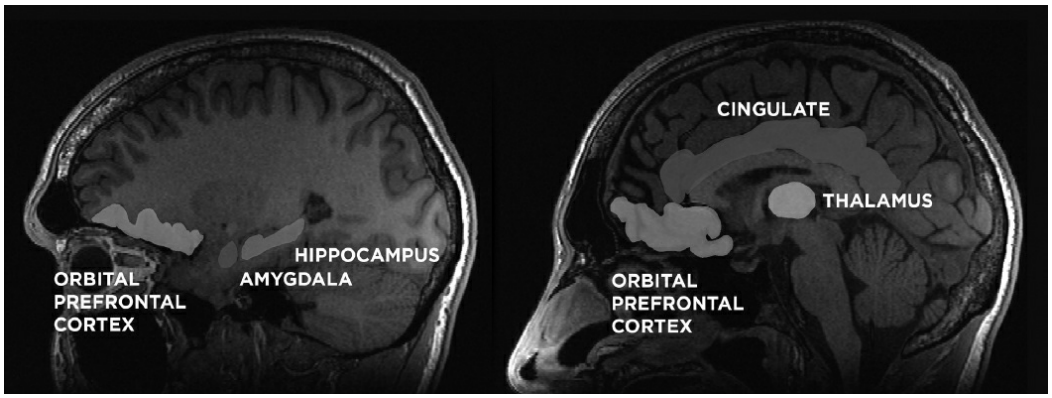

FIGURE 38–19. The thalamus, cingulate, amygdala, hippocampus, and the orbital frontal cortex (OFC) are activated in the emotion of fear. (Figure courtesy of Frank MacMaster, Ph.D.)

Posttraumatic Stress Disorder

The amygdala has been implicated in some studies of adult patients with PTSD, and recently hypertrophy of the amygdala in children with PTSD has been observed. In addition, smaller hippocampal volumes have been consistently found in adults with chronic PTSD, but not in children with PTSD. In children with PTSD, smaller whole-brain, prefrontal cortical, and right temporal lobe volumes are associated with an earlier age of onset and longer duration of abuse. More recently, a ^{1}H-MRS study demonstrated lower concentration of the neuronal marker NAA in the anterior cingulate cortex of maltreated children with PTSD than in that of controls.

Consistent with that finding, lower concentration of gray matter has been found in youth with PTSD and an increased activation in response to threat in the anterior cingulate cortex. Given the mixed evidence, discrete neurobiological profiles may be present in pediatric versus adult patients with PTSD. Longitudinal studies to characterize the developmental neurobiology of the illness are needed. Careful assessment for comorbid conditions—for example, anxiety and depression—are also necessary to discern unique neurobiological pathways among different anxiety disorders and depression.

Tourette Syndrome

TS is often comorbid with ADHD and OCD. Indeed, certain brain circuits are common in the pathogenesis of all three conditions.

Cortex. A recent large-scale volumetric MRI study of children and adults with TS found striking abnormalities of cortical development. In the parieto-occipital region, a greater volume was observed in children than in adults with TS. Age-associated reduction in parieto-occipital volumes occurred predominantly in females. In contrast, relatively larger parieto-occipital volumes were observed in older men with TS, suggesting that sexually dimorphic neurodevelopment may contribute to the greater prevalence of TS in males than in females. Larger dorsolateral PFC volumes were found in the younger TS patients but tended to decrease toward adulthood. Because tics often diminish with age, smaller dorsolateral prefrontal cortical volumes in older TS patients were hypothesized to reflect the relatively unusual persistence of tics in the adults participating in the study.

fMRI studies demonstrating increased activation of the PFC during tic suppression in adult TS patients suggest that larger prefrontal cortical volumes observed in children with TS may represent healthy, structural compensation to allow for tic suppression. Functional neuroimaging studies in healthy children and adults have implicated the orbital PFC in the processing of inhibitory control, such that the inverse correlation between orbital prefrontal volumes and tic severity in children and adults with TS likely reflects a fundamental deficit in inhibitory control allowing for the "release" of tics. In the motor and premotor cortices, increased volumes and functional activity appear to occur in adults but not children with TS, possibly reflecting the cumulative effect of years of repetitive, abnormal movements. More recently, it was noted that the normative developmental correlates of activity in frontostriatal circuits that subserve self-regulatory control are disturbed in persons with TS.

Basal Ganglia. Reduced basal ganglia volumes have been observed in both children and adults with TS. Volume reduction has been found to be

particularly prominent in the globus pallidus and putamen in patients with TS, implicating motor pathways within cortical-subcortical circuitry. This is consistent with the motor phenomena of tics. In support of the hypothesized role of gender in the neurobiology of TS, abnormal basal ganglia asymmetry has been demonstrated in boys but not girls with TS. Although volumetric neuroimaging studies of identical twins with TS have generated conflicting findings, one twin study found greater reduction of the caudate in the more severely affected twin, suggesting that the neurobiological underpinnings of TS expression may be influenced by environmental exposure, as well as by gender and genetic predisposition. Enlargement of thalamic volumes have also been noted in children and adults with TS as compared to healthy comparison subjects. Increased apparent diffusion coefficient (ADC) in corticostriatal projection pathways including left external capsule and left and right subcallosal fasciculus pathway have also been noted in children with TS. Based on the efficacy of dopamine receptor blockade in the treatment of TS, one recent PET study of children with TS found evidence for excessive dopamine decarboxylase activity in the left caudate and right midbrain, suggesting subcortical dysfunction of the dopamine system in pediatric patients with TS (Fig. 38–21).

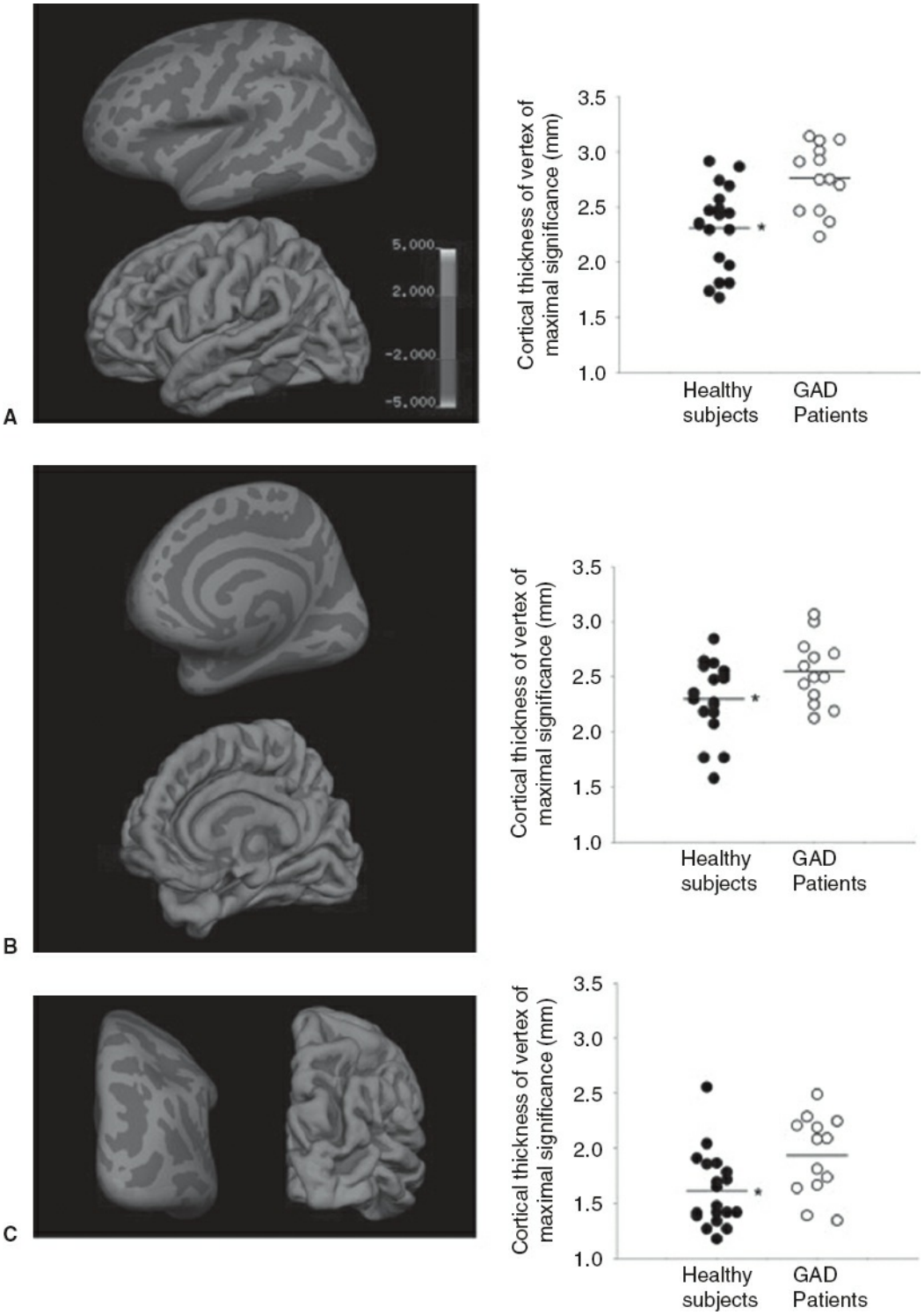

FIGURE 38–20. Cortical thickness in adolescents with generalized anxiety disorder (GAD) as compared to healthy controls. Warm colors (*red to yellow*) indicate greater cortical thickness in adolescents with GAD compared to healthy comparison subjects, and for each region, both cortical and inflated views are presented. Scatter plots are shown for the maximally significant vertex within each depicted region and healthy patients are represented by (•) while adolescents with GAD are represented by (°). The left temporal findings are shown in panel **A**, while the prefrontal differences and occipital findings are shown in panels **B** and **C**, respectively. Finally, an asterisk denotes a statistically significant difference between groups (p = 0.006 for the left temporal differences; p = 0.0004 for the frontal differences and p = 0.0001 for the occipital differences). (See color atlas.) (Reprinted with permission from Strawn JR, John Wegman C, Dominick KC, et al. Cortical surface

anatomy in pediatric patients with generalized anxiety disorder. *J Anxiety Disord.* 2014;28(7):717–723.)

Corpus Callosum. Conflicting results have been observed with regard to the corpus callosum in TS in youth, with reports of larger, smaller, and no difference in area. A study of monozygotic twins discordant for TS using both MRI and DTI methods failed to observe differences with regard to fiber-tracking to the corpus callosum. Fractional anisotropy was lower in the posterior aspects of the corpus callosum in the affected twin however, indicating a less-coherent level of organization.

Other Brain Regions. A recent structural MRI investigation evaluated the association between the size of the cavum septi pellucidi (CSP) and a diagnosis of TS. They found a smaller CSP in children with TS than in healthy pediatric controls. These findings were replicated in an independent sample of adult patients with TS as compared to healthy adults. An inverse association between CSP size and ADHD symptom severity was also noted. In addition, the overall volume of the hippocampus and amygdala was found to be significantly larger in the TS group than in controls. A surface analysis found that the larger volumes in the TS group were derived primarily from the head and medial surface of the hippocampus and the dorsal and ventral surfaces of the amygdala. The volumes of these subregions declined with age in the TS group but not in controls. Hence, the subregions were significantly larger in children with TS but significantly smaller in adults with TS than in controls. Clinically, volumes in these subregions correlated inversely with the severity of tic, OCD, and ADHD symptoms in patients. This suggests that the enlargement of the subregions may have a compensatory effect on tic-related symptoms.

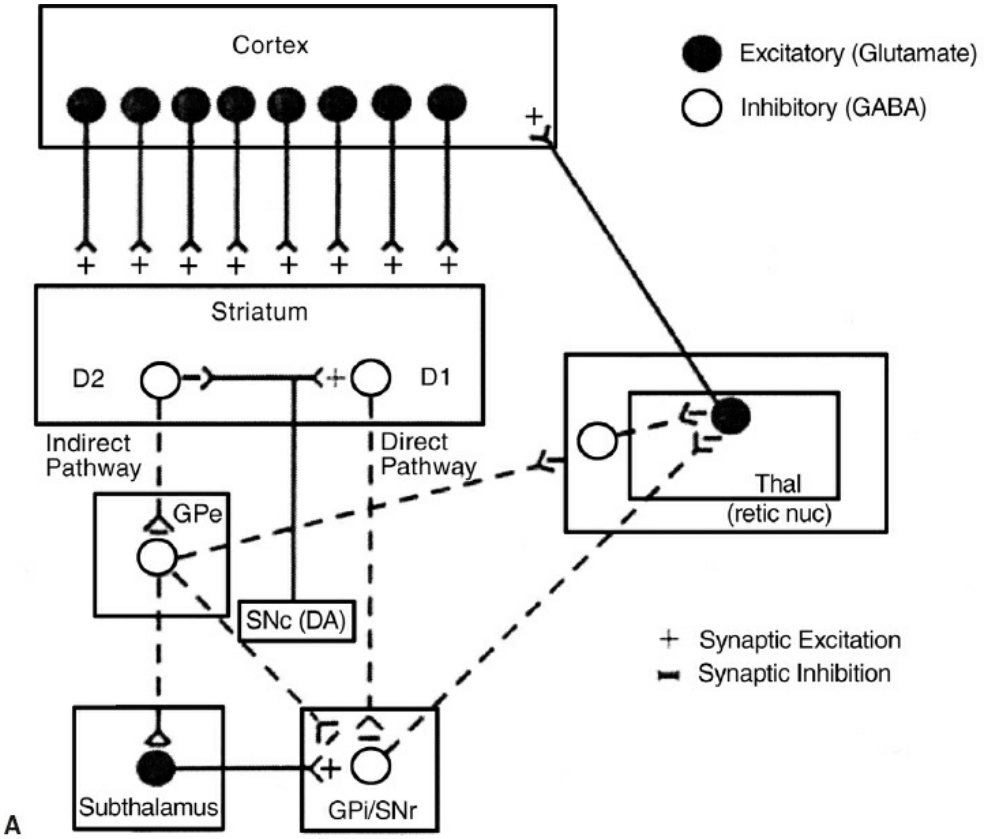
Cortex
Excitatory (Glutamate)
Inhibitory (GABA)
Striatum
D2
D1
Indirect Pathway
Direct Pathway
GPe
Thal
(retic nuc)
SNc (DA)
Synaptic Excitation
Synaptic Inhibition
Subthalamus
GPi/SNr
A

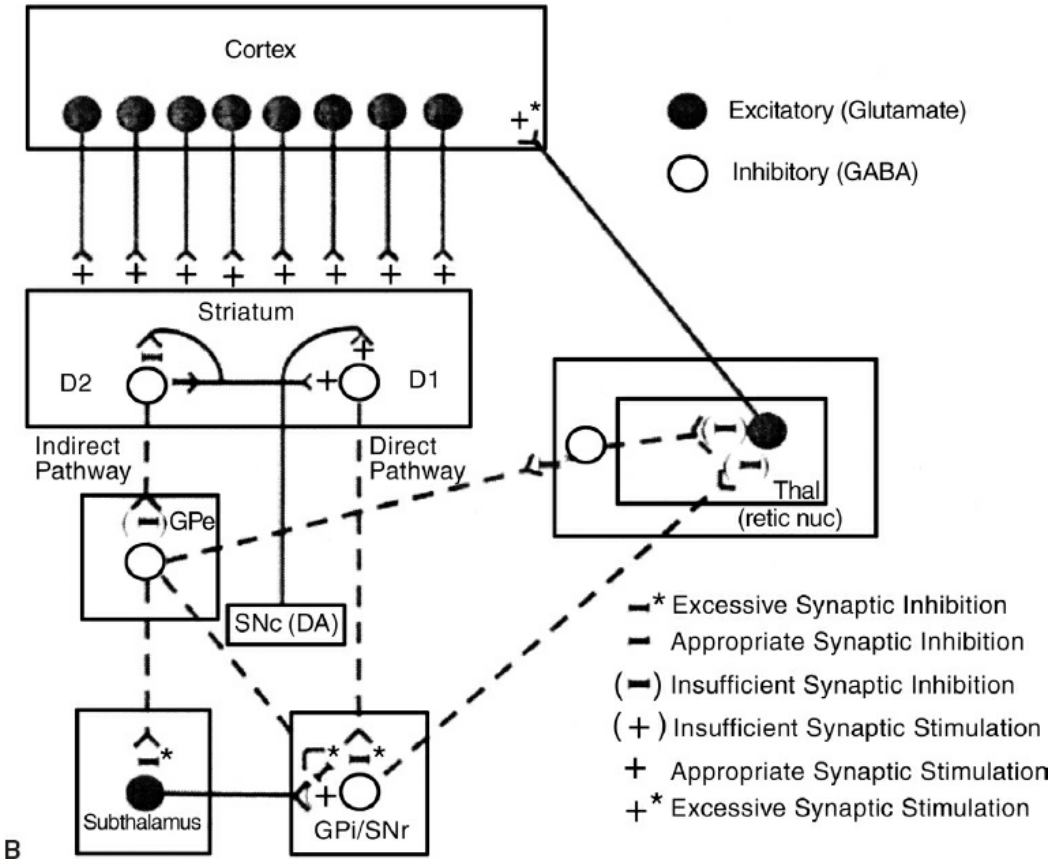
Cortex
Excitatory (Glutamate)
Inhibitory (GABA)
Striatum
D2
D1
Indirect Pathway
Direct Pathway
GPe
Thal
(retic nuc)
SNc (DA)
Excessive Synaptic Inhibition
Appropriate Synaptic Inhibition
Insufficient Synaptic Inhibition
Insufficient Synaptic Stimulation
Appropriate Synaptic Stimulation
Excessive Synaptic Stimulation
Subthalamus
GPi/SNr
B

FIGURE 38–21. Effects of postulated dopaminergic hyperinnervation of the striatum. Although some studies have reported normal integrity of striatal dopamine systems, preliminary neuroimaging and postmortem results suggest hyperinnervation of the striatum by dopaminergic neurons, which may contribute to the pathogenesis of tics. Excitatory (glutamatergic) projections are indicated by *solid lines* and inhibitory (GABAergic) projections by *dotted lines.* GABA, γ-aminobutyric acid; Gpe, external segment of the globus pallidus; Gpi/SNr, internal segment of the globus pallidus or the substantia nigra pars reticulata; retic nuc, reticular nuclei of the thalamus; SNc (DA), dopaminergic projections from the substantia nigra pars compacta; Thal, thalamus. (Reprinted with the permission of Cambridge University Press; Peterson BS, Thomas P. Tourette's syndrome: What are we really imaging? In: Ernst M, Ramsey JM, eds. *Functional Neuroimaging in Child Psychiatry*. New York: Cambridge University Press; 2000.)

Childhood-Onset Schizophrenia

Although it is a rare condition, childhood-onset schizophrenia has recently been the focus of considerable research investigation. It is believed that enhanced understanding of the developmental brain changes in early-onset schizophrenia might lead to an improved understanding of the neurobiology of adult schizophrenia. Indeed, using brain region data as input, machine-learning methods can correctly classify childhood-onset schizophrenia and healthy controls with 73.7% accuracy.

Total Brain Volume, White Matter, and Ventricles. Aside from disrupted hypothalamic–pituitary–adrenal (HPA) axis function in mood disorders, one of the most consistent findings in psychiatry is a reduction in total brain volume and an increase in the volume of the ventricles in adult patients with schizophrenia. In childhood-onset schizophrenia, the smaller total brain volume is even more dramatic than in adult patients. This suggests that childhood-onset schizophrenia may be a rare, more severe version of the disease. Smaller total brain volume in childhood-onset schizophrenia is accompanied by ventriculomegaly. The smaller amount of brain tissue in early-onset schizophrenia may result from a progressive loss of both gray and white matter in periventricular regions, including the posterior cingulate, caudate, and thalamus (Fig. 38–22).

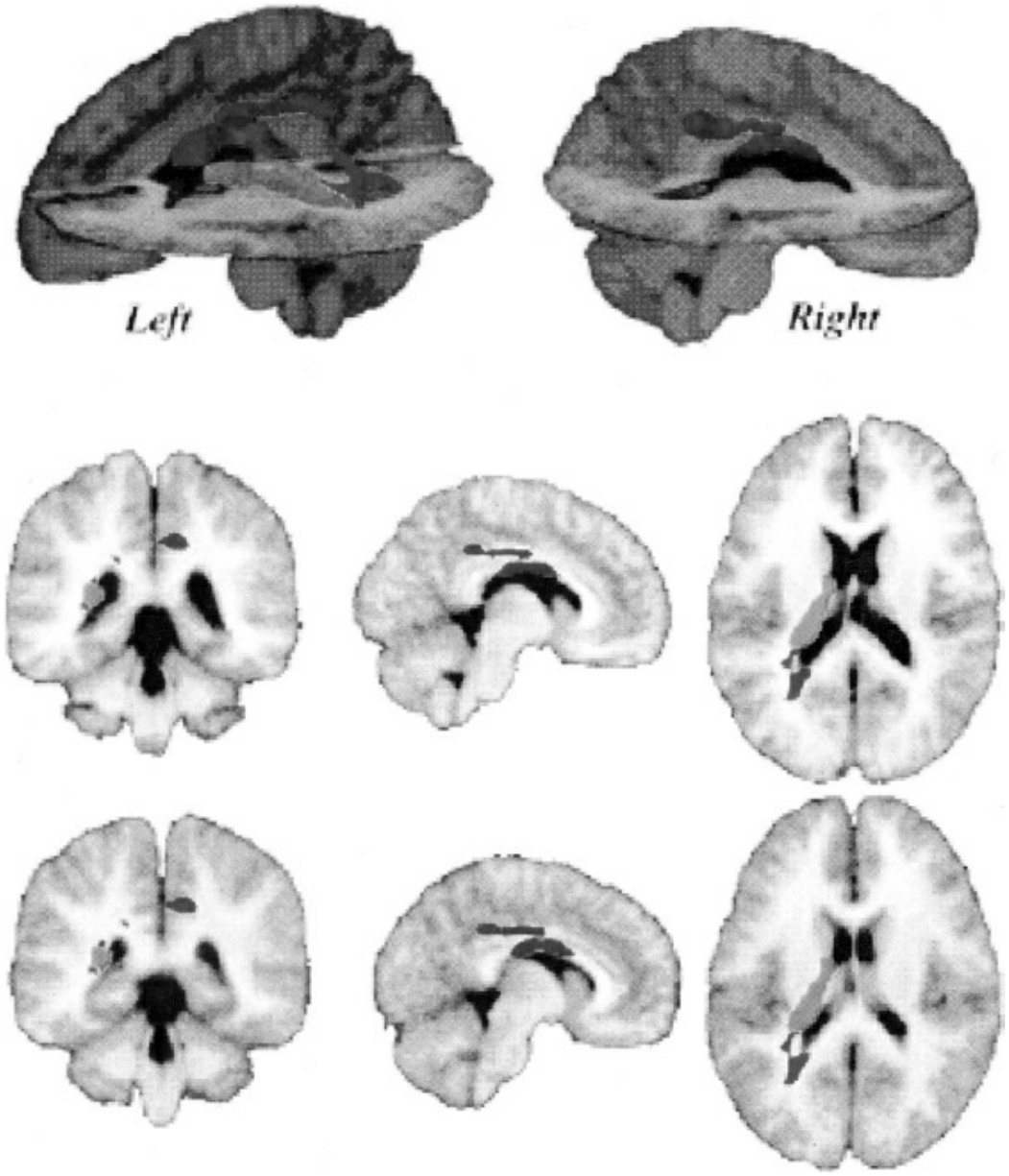

FIGURE 38–22. Top: Left and right hemisphere views showing the three-dimensional renderings of the significant clusters from the gray matter, white matter, and cerebrospinal fluid. Statistical parametric maps (SPMs) were mapped into the average brain space for nine patients with childhood-onset schizophrenia. **Middle:** Significant voxels form the gray matter (voxel threshold $p < 0.01$, corrected at $p = 0.05$), white matter (voxel threshold $p < 0.01$, corrected at $p = 0.05$), and CSF (voxel threshold $p < 0.01$, corrected at $p = 0.05$). SPMs mapped onto three orthogonal slices form the average brain in schizophrenia. **Bottom:** Over the same three slices of the average control brain. (Reprinted with permission from Sowell ER, Toga AW, Asarnow R. Brain abnormalities observed in childhood-onset schizophrenia: A review of the structural magnetic resonance imaging literature. *Ment Retard Dev Disabil Res Rev.* 2000;6(3):180–185.)

In young patients with childhood-onset schizophrenia, slower growth rates were observed in parietal white matter. However, the corpus callosum did not demonstrate any differences from controls with regard to volume and growth trajectory. Furthermore, patients with childhood-onset schizophrenia have demonstrated reduction in fractional anisotropy in the

cuneus when compared to healthy controls. This reduction was also observed in their healthy siblings, possibly indicating some shared genetic risk. This indicates regional effects on white matter with childhood-onset schizophrenia.

Prefrontal Cortex, Insula, and Temporal Lobe. Akin to the aforementioned finding in the cuneus, during a weather prediction fMRI task, siblings of childhood-onset schizophrenia patients demonstrated less activation of the frontal and striatal regions compared to controls. This indicates that impairments in cortical-striatal function may play a role in the liability of schizophrenia. It has been suggested that changes in frontal deficits are modulated in part by polymorphism of the catechol-o-methyltransferase (COMT) gene. Lower levels of the ^{1}H-MRS neuronal marker NAA and gray matter volumes in the frontal and anterior cingulate cortex are found in patients with childhood-onset schizophrenia and unaffected offspring of patients with schizophrenia at increased risk for developing schizophrenia as compared to controls. This suggests that an exaggerated pruning of neurons may be involved in the pathogenesis of schizophrenia. Controlling for medication exposure in neuroimaging studies is critical because typical neuroleptics have been shown to decrease frontal cortical activation, whereas atypical agents (e.g., risperidone) increase frontal activation. The insula plays a critical role in delineating self from non–self-stimuli. Given this function, it has been implicated in the positive symptoms associated with schizophrenia. In 98 patients with childhood-onset schizophrenia contrasted with 100 healthy control participants, smaller right and left insula volumes were noted in patients. Right insula volume correlated inversely with positive symptoms in patients with childhood-onset schizophrenia.

Consistent with studies of adult patients with schizophrenia, smaller volume of the temporal lobe has been reported in childhood-onset schizophrenia, although a contradictory report exists. Within the temporal lobe, smaller superior temporal gyrus volumes associated with increased symptom severity have been observed in patients with childhood-onset schizophrenia as compared to healthy controls. In adult patients with first-episode schizophrenia, volume reduction in the superior temporal gyrus was correlated with duration of pretreatment illness. Superior temporal gyrus volumes appear to increase toward normal after pharmacotherapy, suggesting that this region may be integral to symptom expression in schizophrenia. In a recent report, it was noted that patients with childhood-onset schizophrenia showed significant bilateral deficits (9 to 10 percent) in total hippocampal volume. This smaller volume in such patients remained consistent from age 9 to 26 years. Of interest, subregional maps showed heterogeneous changes, with a loss of hippocampal volume in both anterior and posterior ends, whereas the body of the hippocampus gained in

volume. This suggests that the subunits of the hippocampus are differentially affected in schizophrenia. Indeed, a reduction in the volume of the amygdala–hippocampal formation has also been demonstrated in adults with first-episode schizophrenia, in patients with childhood-onset schizophrenia, and in offspring and siblings of patients with schizophrenia. The reduction in amygdala–hippocampal formation in patients with childhood-onset schizophrenia and high-risk unaffected relatives suggests that it might play a critical role in the pathogenesis of schizophrenia. Earlier onset of illness in patients with traumatic perinatal histories has also been shown to be associated with more severe reduction in amygdala–hippocampal volume. Shape analysis revealed a striking inward deformation in the anterior hippocampus in patients with childhood-onset schizophrenia. These inward deformations were associated with positive symptoms. More recently it has been noted that the younger, healthy siblings of patients with childhood-onset schizophrenia showed significant gray matter deficits in the left prefrontal and bilateral temporal cortices and smaller deficits in the right prefrontal and inferior parietal cortices compared with controls. It may be that a loss in prefrontal and temporal gray matter in childhood-onset schizophrenia is a familial/trait marker of the illness.

Subcortical Regions. Increased caudate volume has been reported in childhood-onset schizophrenia, although this might be due to typical neuroleptic exposure. When typical neuroleptics were replaced with clozapine in patients with childhood-onset schizophrenia, caudate size decreased to normal within 2 years. However, caudate volume reduction has been reported in neuroleptic-naive adult patients with schizophrenia, so that a primary alteration in caudate volume cannot be ruled out. Recently, a shape analysis of the caudate in childhood-onset schizophrenia found an outward displacement in the dorsal aspect of the caudate body. One of the most consistent findings in psychotropic-naive young adults with schizophrenia and childhood-onset schizophrenia is decreased volume of the thalamus. Reduced thalamic volumes have also been found in the primary relatives of patients with schizophrenia.

Substance-Related and Addictive Disorders

Alcohol Use Disorder. Given the high prevalence of conditions such as binge drinking among adolescents, the need for research in this area is great. Consistent with adult studies, a neuroimaging study of alcohol use in adolescents found a significant volume reduction in bilateral hippocampi as compared to healthy pediatric controls. Hippocampal volume reduction correlated positively with age of onset and negatively with duration of alcohol use, suggesting that alcohol may interfere with normal growth of the hippocampus. Smaller prefrontal gray and white volumes have been

noted in adolescents who abuse alcohol. In youth at high risk for alcoholism, a smaller amygdala has been consistently noted. This is suggestive of a possible abnormality in the processing of fearful stimuli. An fMRI connectivity study found differences in resting state synchrony between the left nucleus accumbens and bilateral inferior frontal gyri and the left postcentral gyrus in youth at risk as compared to healthy controls. The results suggested that youth at risk had less segregation between the nucleus accumbens and frontal executive functions, and less integration with reward-related areas like the orbital frontal cortex. In a similar cohort, ^{31}P-MRS has been used to probe for metabolic abnormalities of frontal, occipital, and parietal cortex. In children with both a positive family history of alcoholism and a disruptive behavior disorder, decreased phosphodiester levels were observed, suggesting abnormalities in energy metabolism. These abnormalities were also associated with learning disability. In a longitudinal morphometric study, 40 adolescents were followed over three years. During that period, half initiated heavy drinking. At baseline, subjects who later started drinking heavily had smaller left cingulate, pars triangularis, and rostral anterior cingulate volume as compared to those who did not. Over time, reductions in left inferior and middle temporal gyri, left caudate and brain stem were noted as compared to substance naïve youth. This study suggests that there are pre-existing markers in future drinkers and a further volume reduction over time after alcohol abuse was initiated.

Cannabis Use Disorder. Cannabis using adolescents demonstrate thinner left and right supramarginal, left and right inferior parietal, right pars triangularis, left pars opercularis, left superior frontal, and left superior temporal cortical thickness compared to noncannabis using subjects. Greater cannabis use was also associated with thicker bilateral superior frontal gyri. In adolescents who are regular cannabis smokers, gray matter volume reduction was observed in the medial temporal cortex, temporal pole, parahippocampal gyrus, insula, and orbitofrontal cortex. These are brain regions rich in cannabinoid CB1 receptors and associated with motivational, emotional, and affective processing. In 17 chronic cannabis-using adolescents as compared to 17 healthy controls, reduction in anterior cingulate *N*-acetyl-aspartate, creatine, inositol, and glutamate were observed. Akin to what was observed in childhood onset schizophrenia, changes in prefrontal and insula cortical thickness were observed in adolescent cannabis users. In a resting state connectivity study, heavy cannabis use in adolescents was also associated with middle frontal gyri activity and inversely with the middle temporal gyrus. During an fMRI monetary incentive delay task, adolescent males who use cannabis showed striatal hyperactivity in anticipation of reward. A systematic review of diffusion weighted imaging studies of cannabis use in adolescence found

abnormal white matter microstructure in neocortical association pathways as well as in projection and thalamic pathways. One of the primary problems with studies of cannabis use in adolescents is the limited ability to quantify and control for exposure; with most studies relying on self-report of use with little data on the strength of cannabis used.

Feeding and Eating Disorders

Several studies have documented decreased global brain volume in children and adolescents with anorexia nervosa (AN) that may be somewhat reversible. However, decreased large-scale brain volume loss is probably the result of malnutrition and, is not, therefore, likely to be particularly helpful in elucidating the neurobiologic underpinnings of AN. In adolescents with AN, a lower concentration of frontal NAA have been noted than in controls. Of interest, a lower frontal phosphodiester level was reported as well, indicating a possible decrease in the fraction of membrane phospholipids. However, spectroscopy studies have yielded conflicting results overall. An fMRI study found greater activation in AN patients than in controls when they were viewing food-related stimuli. During working memory tasks, greater activation of temporal and parietal areas was noted in adolescents with AN that resolved with treatment. Functional studies of reward and punishment have also shown that adolescents with AN are especially responsive to losses compared to wins. Studies of bulimia nervosa have also noted reductions in frontal and temporal regions in affected adolescents compared to healthy subjects. Furthermore, fMRI data indicates that adolescents with bulimia demonstrate abnormal patterns of activation during tasks tapping into regulatory control and conflict resolution—with frontal and striatal regions failing to activate.

Autism Spectrum Disorders

The number of imaging studies investigating the underlying neurobiology of ASDs has increased fourfold since 2003 (Fig. 38–23). This tremendous progress occurred despite inherent limitations. First, there have been advances in scanning young children. Early studies into ASD were hampered by being conducted in adults or adolescents. Use of mock scanners and thoughtful approaches to the scanning procedure have allowed very young children to participate. Given that typical onset of ASD is by 3 years of age, studying mostly older patients places the research at an obvious disadvantage. Second, the majority of neuroimaging studies, especially fMRI studies, have been conducted in high-functioning patients with ASD. Differences in study populations, medication exposure, and methodological differences in image acquisition and analysis may have also contributed to the lack of reproducibility of findings.

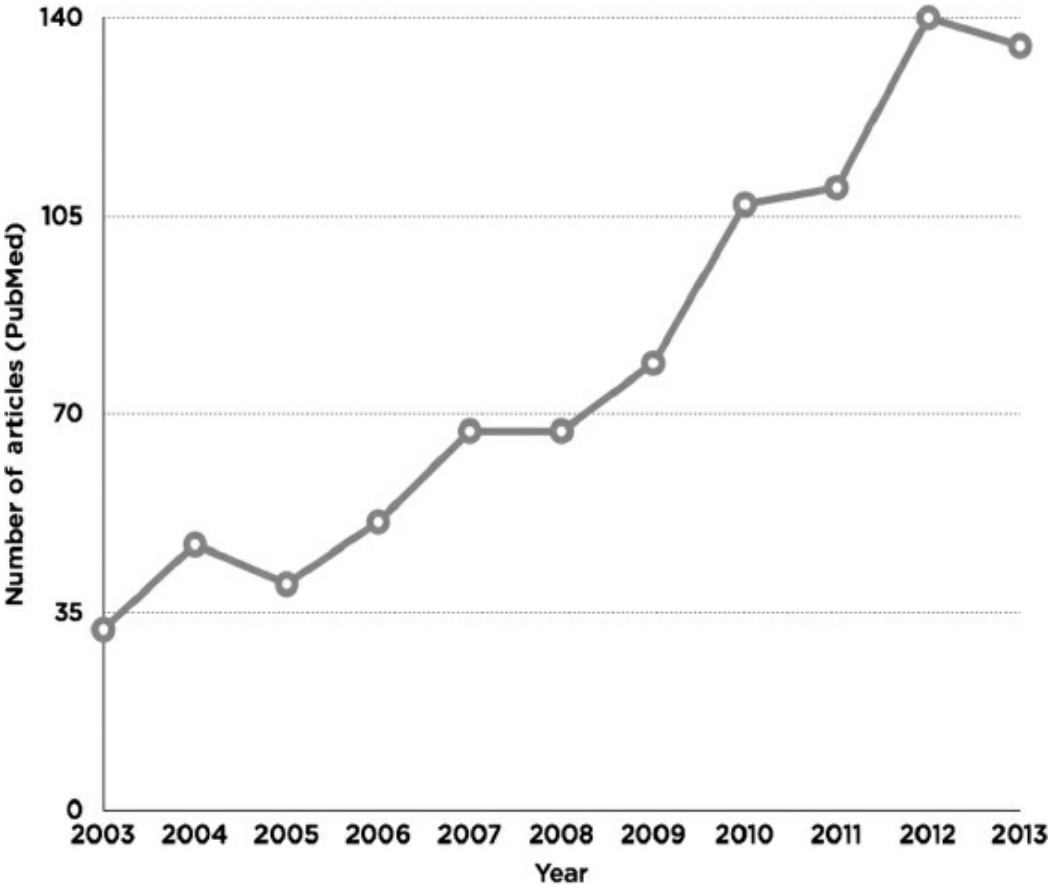

FIGURE 38–23. Graph showing increase in over time in the number of neuroimaging studies investigating the neurobiology of autism spectrum disorders. (Figure courtesy of Frank MacMaster, Ph.D.)

Global Brain Abnormalities. It is believed that alterations in global brain growth are involved in the pathogenesis of ASD. Greater total brain volume has been one of the more consistent findings in patients with ASD. When examining functional connectivity, it appears that in adolescents and adults with ASD, connectivity is reduced. In younger children with ASD, however, intrinsic functional connectivity is increased. These findings have resulted in a call for research onto the effect of puberty on the neurobiology of ASD. In a large-scale study (N = 717) of children 6 to 10 years of age, those with higher autistic traits (scores on the social responsiveness scale) demonstrated widespread areas of decreased gyrification. Whole-brain serotonin synthesis capacity has been measured in 2- to 15-year-old children with ASD with PET studies using the tracer α-[^{11}C]methyl-L-tryptophan (^{11}C-AMT) for PET. Children with ASD were compared to healthy siblings without ASD and to epileptic children without ASD because using PET to study the brains of healthy pediatric controls is rare. Sedation was required for all of the children with ASD or epilepsy and for many of the siblings without ASD who underwent PET scanning. Disruption in serotonin synthesis capacity was observed in children with ASD compared to their siblings and children with epilepsy. In normal

children, whole-brain serotonin synthesis capacity was greater than 200 percent of adult capacity until age 5 years, and then it declined to adult levels, whereas in children with ASD, serotonin synthesis capacity increased very gradually from age 2 to 15 years to levels 1 to 1.5 times adult normal values.

Regional Brain Alterations. In addition to global brain alterations in serotonin synthesis, comparisons of children with ASD who were otherwise healthy without seizures and their siblings without ASD showed asymmetries in serotonin synthesis in frontal cortex, thalamus, and cerebellum. Lower serotonin synthesis was observed in the left frontal cortex and thalamus in children with ASD (Fig. 38–24). Lesser accumulation of ^{18}F-DOPA in the anterior medial PFC was observed in medication-free children with ASD (Fig. 38–25). Lower levels of dorsolateral PFC phosphocreatine, as measured by ^{31}P-MRS, in high-functioning adolescent and young adult men with ASD suggest altered metabolism and increased membrane degradation in this region and was correlated with neuropsychological and language deficits. The temporal lobe plays a critical role in emotion processing. Hence, it is not surprising that pediatric ASD has been associated with lower metabolic rate, perfusion, and NAA in temporal regions. Volumetric studies of the amygdala and hippocampus have been inconsistent in ASD. In the cerebellum, the dentate nucleus has also been implicated in ASD. The dentate nucleus is involved in the regulation of language and motor output. Abnormal activation of the right dentate has been observed in patients with ASD when performing a language task. In response to sensory stimuli, along with the expected sensory areas, the hippocampus, amygdala, and orbital frontal cortex demonstrate hyperactivity in youth with ASD. Furthermore, greater brain serotonin synthesis has been reported in the right dentate nucleus. Reductions in the neuronal marker NAA in the dentate as measured by ^{1}H-MRS have also been reported.

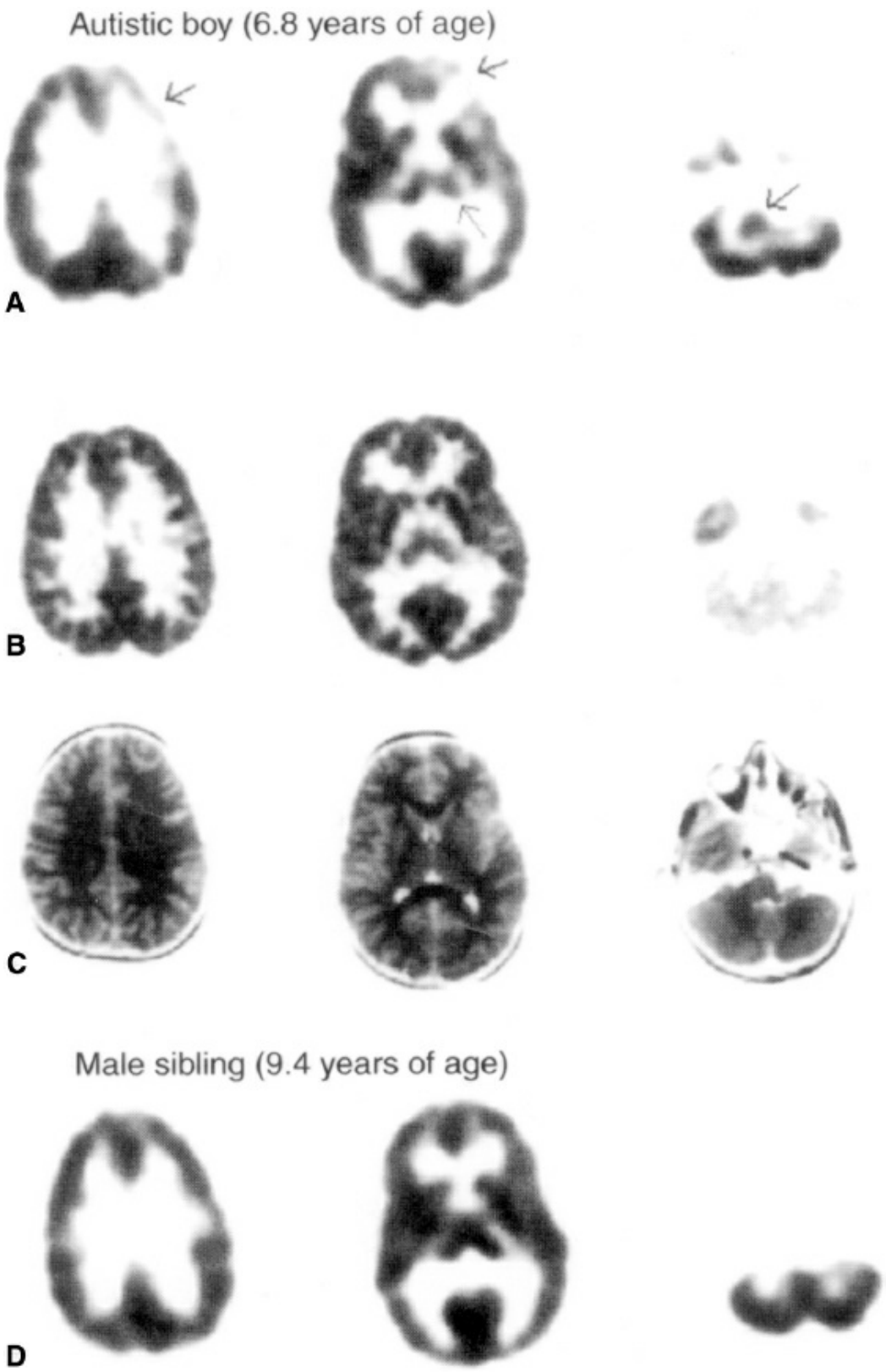

FIGURE 38–24. Scans from a boy with autism, age 6.8 years **(A–C),** and his brother, age 9.4 years **(D)**. (**A, D**) Positron emission tomography (PET) images using α-[^{11}C]methyltryptophan ([^{11}C]-AMT); (**B**) PET images using [^{18}F]fluorodeoxyglucose scan; (**C**) magnetic resonance imaging scan. *Arrows* denote decreased [^{11}C]-AMT accumulation in left frontal cortex and left thalamus and increased [^{11}C]-AMT accumulation in right dentate nucleus in the child with autism (the left-side image is the right side of the brain).

(Reprinted with the permission of Cambridge University Press; Chugani D. Autism. In: Ernst M, Ramsey JM, eds. *Functional Neuroimaging in Child Psychiatry*. New York: Cambridge University Press; 2000, 179.)

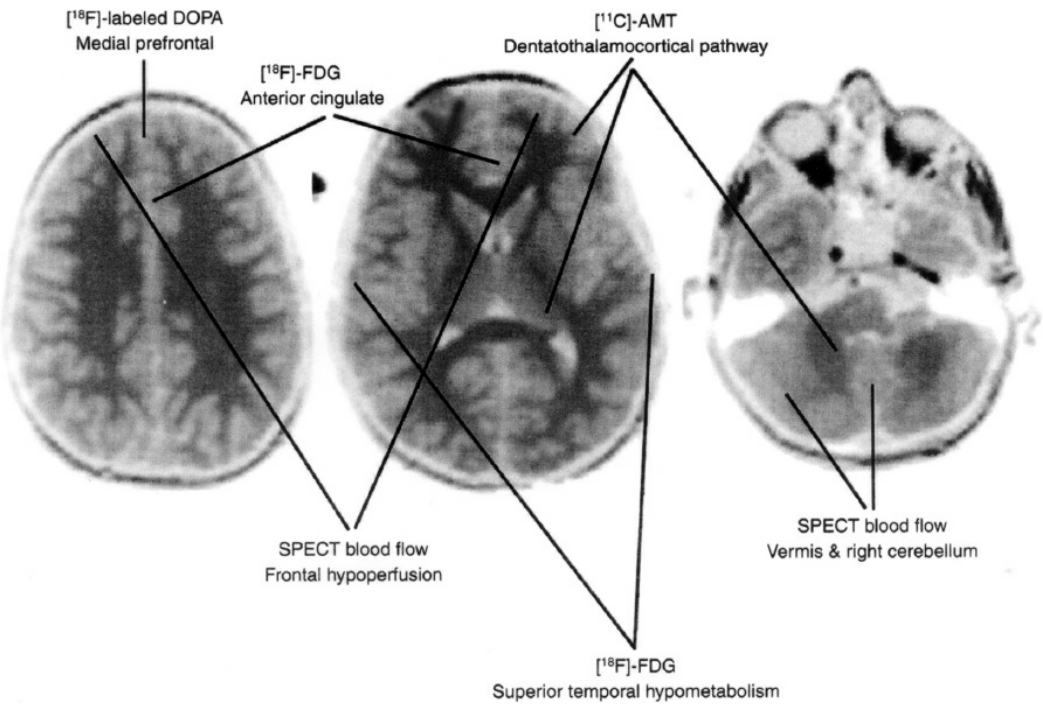

FIGURE 38–25. Brain regions identified as abnormal in individuals with autism by various functional imaging modalities. Frontal medial, prefrontal, temporal, and anterior cingulate cortical regions have been implicated with various functional tracers (AMT, α-methyltryptophan; FDG, fluorodeoxyglucose; DOPA, dihydroxyphenylalanine). SPECT, single-photon emission computed tomography. Abnormalities have also been reported in subcortical regions such as the thalamus and cerebellum. A convergence of the data from the various imaging modalities has not yet emerged to provide a unifying hypothesis of the brain mechanisms responsible for autism. (Reprinted with the permission of Cambridge University Press; Chugani D. Autism. In: Ernst M, Ramsey JM, eds. *Functional Neuroimaging in Child Psychiatry*. New York: Cambridge University Press; 2000:184.)

Specific Learning Disorders

The left hemisphere plays a critical role in the development of reading skills in healthy children. This function appears to be impaired in children with a reading disability, as suggested by lower functional activation in dorsal and ventral regions. Indeed, similar results have been noted in adults with a reading disability, with less functional connectivity between these circuit substrates.

More recently, bilateral prefrontal regions that link brain regions associated with phonological awareness and rapid naming also demonstrated abnormal functional connectivity in children with dyslexia. The planum temporale is associated with a number of auditory and language-related processes. In children with dyslexia, the planum temporale demonstrates a greater rightward asymmetry. Perhaps to compensate for this dysfunction, children and adults with a reading

disability "overactivate" the dorsolateral prefrontal cortex, including the inferior frontal gyrus. Adults with a reading disability also engage the parallel circuitry in the right hemisphere. Right hemisphere activation during reading is normal in children but typically diminishes into adulthood. This suggests a change in the lateralization that characterizes the neural substrates that mediate reading in individuals with persistent reading disability.

FUTURE DIRECTIONS

Although there are unique challenges to conducting neuroimaging studies in pediatric populations, the use of neuroimaging in children with psychiatric disease provides an exciting opportunity to explore the developmental abnormalities that lead to mental illness. Neuroimaging studies in childhood-onset neuropsychiatric disorders that use biomarkers as both moderator (present at baseline) and mediator (changing with treatment) may potentially result in an enhanced mechanistic understanding of the pathogenesis and maintenance of these illness, which could lead to better diagnostic and treatment approaches. Neuroimaging is also well matched with approaches based off the research domain criteria. MR-based technologies are particularly useful in youth because they entail no radiation exposure and can be successfully applied to longitudinal study. Although technological advances will undoubtedly continue to provide more precisely resolved images, investigators must pay special attention to developing consistency in methodological approaches across studies, along with larger sample sizes to account for the high variability of brain volume and function during childhood and adolescence. Exciting times lie ahead in the attempts to unravel the pathophysiology and mechanism of treatment response in various childhood-onset neuropsychiatric disorders.

REFERENCES

Ahmed F, Ras J, Seedat S. Volumetric structural magnetic resonance imaging findings in pediatric posttraumatic stress disorder and obsessive compulsive disorder: a systematic review. *Front Psychol.* 2012;3:568.

Chugani D. Autism. In: Ernst M, Ramsey JM, eds. *Functional Neuroimaging in Child Psychiatry*. New York: Cambridge University Press; 2000:179.

Chugani HT. Biological basis of emotions: brain systems and brain development. *Pediatrics.* 1998;102(5 Suppl E):1225–1229.

Croarkin PE, Wall CA, McClintock SM, Kozel FA, Husain MM, Sampson SM. The emerging role for repetitive transcranial magnetic stimulation in optimizing the treatment of adolescent depression. *J ECT.* 2010;26(4):323–329.

*Ecker C, Bookheimer SY, Murphy DG. Neuroimaging in autism spectrum disorder: brain structure and function across the lifespan. *Lancet Neurol.* 2015;14(11):1121–1134.

Eichele H, Plessen KJ. Neural plasticity in functional and anatomical MRI studies of children

with Tourette syndrome. *Behav Neurol*. 2013;27(1):33–45.

Ernst M, Rumsey JM. *Functional Neuroimaging in Child Psychiatry*. Cambridge University Press; 2000.

*Giedd JN, Raznahan A, Alexander-Bloch A. Child psychiatry branch of the National Institute of Mental Health longitudinal structural magnetic resonance imaging study of human brain development. *Neuropsychopharmacology*. 2015;40(1):43–49.

Gilbert AR, Moore GJ, Keshavan MS, et al. Decrease in thalamic volumes of pediatric patients with obsessive-compulsive disorder who are taking paroxetine. *Arch Gen Psychiatry*. 2000;57(5):449–456.

Jacobus J, Tapert SF. Effects of cannabis on the adolescent brain. *Curr Pharm Des*. 2014;20(13):2186–2193.

*Keshavan MS, Giedd J, Lau JY, Lewis DA, Paus T. Changes in the adolescent brain and the pathophysiology of psychotic disorders. *Lancet Psychiatry*. 2014;1(7):549–558.

*MacMaster FP. Translational neuroimaging research in pediatric obsessive-compulsive disorder. *Dialogues Clin Neurosci*. 2010;12(2):165–174.

MacMaster FP, Rosenberg DR, Kaufmann J. Neurobiology of early onset mood disorders. In: Martin A, Scahill L, Kratochvil CJ, eds. *Pediatric Psychopharmacology*. Oxford, United Kingdom: Oxford University Press; 2010:124–139.

Pavuluri M. Brain biomarkers of treatment for multi-domain dysfunction: pharmacological FMRI studies in pediatric mania. *Neuropsychopharmacology*. 2015;40(1):249–251.

Peterson BS, Thomas P. Tourette's syndrome: What are we really imaging? In: Ernst M, Ramsey JM, eds. *Functional Neuroimaging in Child Psychiatry*. New York: Cambridge University Press; 2000.

Rosenberg DR, Hanna GL. Genetic and imaging strategies in obsessive-compulsive disorder: potential implications for treatment development. *Biological psychiatry*. 2000;48(12):1210–1222.

Rosenberg DR, Keshavan MS. A.E. Bennett Research Award. Toward a neurodevelopmental model of obsessive–compulsive disorder. *Biological psychiatry*. 1998;43(9):623–640.

Rosenberg DR, MacMaster FP, Keshavan MS, Fitzgerald KD, Stewart CM, Moore GJ. Decrease in caudate glutamatergic concentrations in pediatric obsessive-compulsive disorder patients taking paroxetine. *J Am Acad Child Adolesc Psychiatry*. 2000;39(9):1096–1103.

Rosenberg DR, MacMillan SN, Moore GJ. Brain anatomy and chemistry may predict treatment response in paediatric obsessive–compulsive disorder. *Int J Neuropsychopharmacol*. 2001;4(2):179–190.

*Rubia K, Alegria AA, Brinson H. Brain abnormalities in attention-deficit hyperactivity disorder: a review. *Rev Neurol*. 2014;58 Suppl 1:S3–S16.

Sowell ER, Toga AW, Asarnow R. Brain abnormalities observed in childhood-onset schizophrenia: a review of the structural magnetic resonance imaging literature. *Ment Retard Dev Disabil Res Rev*. 2000;6(3):180–185.

Stanley JA, Kipp H, Greisenegger E, et al. Evidence of developmental alterations in cortical and subcortical regions of children with attention-deficit/hyperactivity disorder: a multivoxel in vivo phosphorus 31 spectroscopy study. *Arch Gen Psychiatry*. 2008;65(12):1419–1428.

Strawn JR, John Wegman C, Dominick KC, et al. Cortical surface anatomy in pediatric patients

with generalized anxiety disorder. *J Anxiety Disord*. 2014;28(7):717–723.

Thompson PM, Giedd JN, Woods RP, MacDonald D, Evans AC, Toga AW. Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature*. 2000;404(6774):190–193.

Val-Laillet D, Aarts E, Weber B, et al. Neuroimaging and neuromodulation approaches to study eating behavior and prevent and treat eating disorders and obesity. *Neuroimage Clin*. 2015;8:1–31.

Yang XR, Kirton A, Wilkes TC, et al. Glutamate alterations associated with transcranial magnetic stimulation in youth depression: a case series. *J ECT*. 2014;30(3):242–247.

39

Temperament: Risk and Protective Factors for Child Psychiatric Disorders

DAVID C. RETTEW, M.D.

Individuals differ in the types and intensity of their emotional and behavioral responses to various environments. Some people are quick to experience anger or anxiety, whereas others are more laid back. Some people seek out new and exciting experiences, whereas others prefer a predictable routine and the comfort of a few familiar friends. Some people are able to delay their impulses and overcome obstacles in the service of working toward a long-term goal, whereas others become quickly distracted or frustrated. These widely recognized human qualities refer to temperament or personality traits that, along with physical and cognitive dimensions, distinguish one person from the next.

Scientists have been observing and studying child temperament and personality development for many years. The bulk of this work has existed beyond the view of most mental health clinicians, under the notion that this area was distinct from clinical psychiatric disorders. More recently, however, accumulating research has demonstrated strong links between various temperamental traits and nearly all types of psychopathology. The prevailing question today no longer is *if* temperament is associated with psychopathology but *how* temperament interacts developmentally in shaping children toward and away from serious psychiatric disturbance. This chapter offers a brief overview of temperament, including its history, definition, core dimensions, stability, and underlying neurobiology. Following these sections is a summary of the known relations between specific temperamental traits and corresponding psychiatric disorders. Finally, the chapter concludes with a discussion of possible mechanisms through which temperamental traits operate to increase risk or, conversely, promote wellness and resiliency throughout development.

HISTORICAL ANTECEDENTS

When most students think of temperament, the two names that often arise are child psychiatrists Stella Chess and Alexander Thomas, who brought the concept of temperament and its role in developmental psychopathology into the spotlight more than 50 years ago. The foundations of temperament

theory, however, can be traced back in time much further. The ancient Greek physician Galen postulated the existence of four temperament types, which he called sanguine (extraverted and stable), choleric (extraverted and unstable), melancholic (introverted and stable), and phlegmatic (introverted and unstable), based on the relative concentrations of four "humors" (blood, yellow bile, black bile, and phlegm). Although this early conceptualization of temperament may appear somewhat comical by today's standards, a thought experiment that substitutes the four humors with modern neurotransmitters such as serotonin and norepinephrine reveals how much of this framework remains intact. Indeed, the very word "temperament," derived from the Latin *temperare*, which means "to mix," conveys this association between the level of important biological substrates in the brain and basic human behavioral and emotional predispositions.

Fast forward to the 20th century and the dawn of modern psychiatry: it is worth noting that Sigmund Freud made few qualitative distinctions between the origins of normal personality traits and the factors responsible for psychiatric symptoms, although the nature of those forces were certainly not described as temperament-based. As the tenets of behaviorism and learning theory gained prominence, there again were few distinguishing features between the postulated mechanisms underlying personality and those that caused the formation of psychiatric disorders.

It was in this climate during the 1950s that Chess and Thomas began their work on temperament. Based on interviewing parents and observing children across many years, they proposed nine basic dimensions of temperament, as shown in Table 39–1, and three temperament types—easy, slow to warm up, and, most famously, difficult. These basic elements of temperament were hypothesized to be substantially heritable rather than based on learning or unconscious conflicts. It was this heritable quality of temperament that ran against the prevailing wisdom of both the psychoanalytic and the behavioral schools of thought, and the work of Chess and Thomas was subjected to intense criticism.

Through the 1980s and continuing today, a medical model of psychiatric illness began to emerge. Under this framework, disorders like depression or panic disorder arise like other medical illnesses such as diabetes or the flu. Genetic variations and environmental pathogens are viewed as important etiological factors in this model, and the disorders are conceptualized as binary entities that a person does or does not possess. While such an approach acknowledges the influence of more innate dispositional factors, this disease model of psychiatric disorders also divorces symptoms from an individual's temperament. A panic disorder, for example, is an illness that "comes out of nowhere" rather than being related to a temperamental proclivity toward experiencing fear.

Table 39–1.
Thomas and Chess Classification of Temperament

Nine temperament dimensions
Activity level
Rhythmicity
Approach–withdrawal
Adaptability
Response threshold
Response intensity
Mood
Distractibility
Attention span/persistence
Three temperament types
Easy
Difficult
Slow to warm up

Thus, researchers and clinicians interested in temperament continue to challenge and be challenged by the changing models and theories of psychiatric disease. The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), while specifically acknowledging the dimensionality of most emotional-behavioral symptoms, retained a disease model of psychopathology in its operational structure. The implications of this decision will be discussed later in this chapter after additional background information on temperament is provided.

DEFINITION OF TEMPERAMENT

Despite the long history of temperament investigation, the precise definition of temperament remains elusive. One prominent temperament researcher, Mary Rothbart, described temperament as "constitutionally based individual differences in reactivity and self-regulation." Although the debate continues, there exists some consensus on key components of a temperamental trait. First, there needs to be a heritable basis to the trait. This quality does not mean that temperament is purely genetic, but rather that there is at least some genetic influence on the phenotypic expression of that trait. Second, a temperament dimension also needs to be relatively stable across time. Similar to the point about genetic influence, stability does not mean inflexibility or, certainly, destiny, but rather that the trait at one time point can predict future time points even many years into the future. Finally, a temperament dimension is expressed early in life, with its rudiments evident as early as infancy. This latter characteristic distinguishes temperament from other important human qualities, such as life satisfaction or generosity, which incorporate higher cognitive functions. These aspects of temperamental traits are summarized in Table 39–2.

MAJOR DIMENSIONS

The core dimensions of temperament continue to be a source of discussion. As previously described, Chess and Thomas hypothesized nine core temperament dimensions and three temperamental types, with the latter referring to a constellation of traits that often coexist. Subsequent testing using different samples and more sophisticated statistical modeling have supported many of these dimensions, whereas others were found in need of modification.

Currently, there exist a number of alternate, even competing, frameworks that have received some empirical support for their reliability and validity. Most temperament taxonomies are structured around basic incentive response systems of approach and withdrawal often merged or crossed with emotional valence (propensities toward positive and negative emotions). Across these various taxonomies, however, there exist named dimensions that overlap somewhat but not completely with a corresponding dimension in another theoretical model. The temperament trait of novelty seeking as described by Cloninger is similar but not identical to a dimension referred to by Rothbart and colleagues as extraversion or surgency. Adding additional complexity is a dimension's relative place along a hierarchical organization from lower- to higher-order traits. Activity level, for example, exists as a component of the surgency dimension in Rothbart's framework but functions as a broad higher-order dimension in the Buss and Plomin model. Furthermore, differences of opinion exist between the conceptualization of temperament traits as quantitative factors that are distributed normally along a continuum versus more binary yes/no designations. All of these variations between the different models of temperament that exist today provide a rich environment for further research and discussion but can be challenging when trying to compare findings across studies.

Table 39–2.
Definition of Temperament

Individual differences in basic reactivity and self-regulation that include the following elements:

1. Heritability: temperamental traits are often passed on from one generation to the next through both genetic and environmental mechanisms.
2. Early appearance: the antecedents if not the full traits are evident in infancy and toddlerhood.
3. Stability: although there is certainly the potential for change, temperamental traits are moderately stable across many years.

Clearly, there is a tremendous amount of overlap between the core dimensions of temperament and personality, and a full discussion of this topic is well beyond the scope of this chapter. Briefly, it is probably fair to say most temperament and personality theorists view temperament as one of the core foundations of personality, with the latter being further developed through additional life experiences, cultural influences, and

cognitive abilities. Because these influences act on children and adults alike, however, a clear boundary between a temperament and a personality trait is obscure and somewhat akin to pinpointing exactly when someone is, for example, an adult. At the same time that some prominent personality researchers such as Costa and McCrae argued that their personality dimensions can largely be seen as temperament, child researchers found evidence that the "Big Five" taxonomy of personality is largely applicable in youth samples. These efforts suggest that at least some of the historical distinction between these two constructs was semantic and based on the age of the sample being studied. These conceptual challenges notwithstanding, several temperament dimensions are fairly consistent across different nomenclatures and constitute the highest level of temperamental traits. These traits are depicted in Table 39–3.

Withdrawal and Propensity to Experience Negative Emotions

Children differ in the threshold required to illicit a negative emotion such as sadness, fear, or anger, the intensity of the emotion when triggered, and the duration of that emotional experience. This trait is found ubiquitously among the different taxonomies and thus has been described using a number of different labels, including negative affectivity, emotionality, and neuroticism. Some researchers have described a similar trait that reflects an underlying sensitivity to signals of punishment and have referred to this dimension as harm avoidance. Gray described individual differences in the behavioral inhibition system (BIS), which responds with fear to signals of punishment or nonreward. Some evidence supports greater separation between the tendency to experience anger and frustration and the tendency to experience fear or anxiety. Nonetheless, components of this dimension often include elements such as worry, shyness, irritability, and difficulty being soothed once aroused. This trait also figured heavily in the Chess and Thomas designation of the "difficult" child.

Table 39–3.
Current Dimensions of Temperament and Personality

Developer/Name	Withdrawal and Propensity to Negative Emotions	Approach Behavior and Positive Emotions	Activity	Sociability	Regulatory Abil
Rothbart	Negative affectivity	Surgency			Effortful Contro
Cloninger	Harm avoidance	Novelty seeking		Reward dependence	Persistence
"Big Five"	Negative emotionality	Extraversion		Agreeableness	Conscientiousne

EAS	Emotionality	Activity		Sociability
Kagan	Behavioral inhibition/disinhibition			Behavioral inhibition/disinhibition

EAS, emotionality, activity, sociability.

Approach and Positive Emotions

The other temperament dimension that exists throughout various taxonomic structures refers to the degree to which a person actively engages the world and other people. This trait often also encompasses a tendency to experience positive emotions such as joy and excitement. Additional components of this dimension according to the various frameworks include elements of dominance, preference for high stimulus activities, and approach behavior. Especially when considered in a more social context, this trait is often referred to as extraversion, whereas from a slightly broader perspective it has been labeled novelty or sensation seeking or the behavioral activating system (BAS).

Emotional Regulation and Constraint

The third dimension common to most temperamental systems relates to emotional and attentional regulation especially in the face of obstacles or in the pursuit of a long-term goal. As a separate dimension, it has been referred to as effortful control, persistence, or, under the "Big Five" framework, conscientiousness. Children with lower levels of this trait are often described as impulsive, distractible, and disorderly. Parallels to more philosophical terms encompassing a person's degree of voluntary control such as "will" are apparent and distinguish this dimension from behavior that is restrained by anxiety in children with high negative affectivity or low approach.

Other Dimensions

A number of other traits have been classified as broad-level traits under some frameworks. In others, however, the dimension is either defined as a facet of one of the three principal dimensions previously described.

Social Affiliation. Although aspects of social affiliation can be subsumed under broad conceptualizations of extraversion, other researchers have viewed a person's tendency toward prosocial behavior as a distinct higher-order trait. Elements of social affiliation are also found in terms such as shyness, although merging these concepts has been seen as problematic by some because it equates the *desire* for social contact with the anxiety or *discomfort* associated with social encounters. As an aspect of temperament, this dimension has been called sociability and reward

dependence, among others, whereas under a personality "Big Five" taxonomy it is related to the trait of agreeableness. In some systems, this dimension encompasses elements such as empathy and the desire and ability to share emotional states with others, as well as with dependency and compassion.

Activity Level. A person's level of motor activity is similarly viewed as a component of extraversion in some models, whereas in others, such as that proposed by Buss and Plomin, it is given additional standing in itself. Its ability to be readily observed by others, genetic influence, and emergence as a distinguishing feature very early conform to some of the core requirements of a temperamental trait.

Temperament Types

As mentioned, an alternative approach to understanding temperament focuses on distinguishing groups of individuals who may possess similar levels of a variety of traits. This emphasis on temperament types should not be considered to be a true substitute to a variable-centered focus but rather a complementary approach, in that these person-centered classifications also rely on specific dimensions to define various classes or types. One potential advantage of using temperament types in research is that this lends itself well to the possibility that the effect of one particular trait depends on the presence of other traits. For example, two people might possess similar levels of novelty seeking but different levels of effortful control. When coupled with high effortful control, novelty seeking might be a very adaptive, whereas without effortful control this same quality could lead to behavioral problems. Understanding these interrelations might provide more meaningful insight than considering each dimension in isolation.

The Difficult Child. Probably the most famous temperamental type is the "difficult" child designation of Chess and Thomas, which was distinguished from the "easy" and "slow to warm up" groups (and notably from children who didn't fit any of the three types). The "difficult" category referred to a cluster of about 10 percent of young children who possessed the particular combination of low rhythmicity, low response threshold and adaptability, and high response intensity. Although not hypothesized to reflect psychopathology per se, it was this group of children in longitudinal studies who were most likely to have poor outcomes and greater psychiatric disturbance.

Behavioral Inhibition and Disinhibition. Another well-studied temperament that referred to people rather than specific dimensions was behavioral inhibition (BI), as well as, more recently, behavioral

disinhibition (BD). BI was defined by the Harvard psychologist Jerome Kagan as referring to a minority of young children—again around 10 to 15 percent—who consistently responded to new situations and people with reticence and fear. He, along with many colleagues, later defined a group of children with the exact opposite response to novelty and referred to them as possessing behavioral disinhibition. As will be discussed in more detail, children who manifest BI as toddlers have been shown to be at risk for psychiatric disorders later in life, in particular social anxiety disorder.

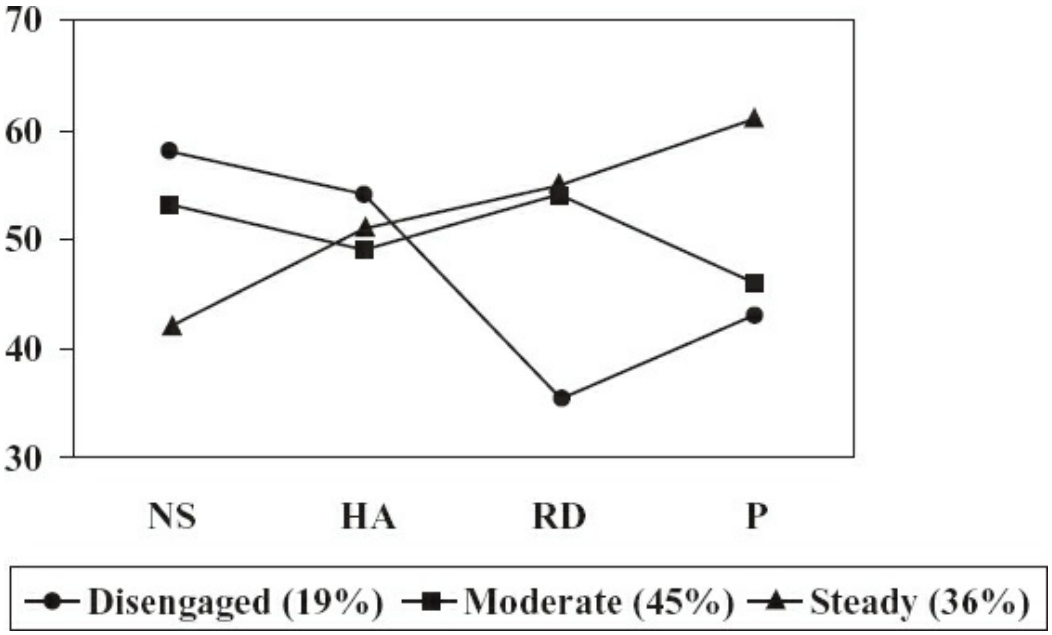

FIGURE 39–1. Temperament profiles as determined though latent profiles analysis. The "moderate" group has average levels across all dimensions, whereas the "steady" group is characterized by low novelty seeking (NS) and high persistence (P). The "disengaged" group has moderately high levels of both novelty seeking and harm avoidance (HA), with low levels of reward dependence (RD). The groups have been found to be related to levels of psychopathology and wellness. (Adapted from Rettew DC, Althoff RR, Dumenci L, Ayer L, Hudziak JJ. Latent profiles of temperament and their relations to psychopathology and wellness. *J Am Acad Child Adolesc Psychiatry*. 2008;47[3]:273–281.)

New Approaches to Temperament Types. Statistical procedures such as cluster analyses and latent profiles analyses can also be used to identify groups of children who manifest similar levels of traits across several dimensions. These analyses can be performed without preconceived theories of how different traits *should* exist together. A latent profiles analysis of a group of children of age to 18 years revealed the presence of three temperament types (Fig. 39–1). The first, as might be expected, contained people who did not have extreme levels of any dimension and was labeled the "moderate" group. A second group was characterized by low novelty seeking and high persistence and was named the "steady" group. Finally, the least common group contained those who possessed high novelty seeking, high harm avoidance, low persistence, and, most

prominently, low reward dependence (low sociability and attachment to others). This last group was labeled the "disengaged" group. As will be discussed, these groups had strong associations with both positive and negative outcomes. Similar efforts by other investigators and using other samples have also tended to find evidence for a relatively small (often around 3 to 6) number of groups distinguishable by their unique temperamental profile.

FEATURES OF TEMPERAMENT

Temperamental traits can be observed in infancy and, according to some, even before birth. Although they contribute to an individual's unique human characteristics, these traits have been widely studied in animals ranging from nonhuman primates to dogs to guppies. Through further understanding of the phenomenology of temperamental traits, it may be possible to advance the knowledge of where temperament and psychopathology share common characteristics and where there are areas of divergence.

Sex Differences

A review and meta-analysis of sex differences using Cloninger's temperament framework showed higher levels of reward dependence and harm avoidance in women than in men, whereas novelty seeking and persistence were found to be roughly equal. Of interest, another meta-analysis of child studies using Rothbart's temperament structure found slightly different results with corresponding dimensions. Specifically, effortful control was found to be higher in girls, surgency was higher in boys, and negative affectivity was roughly equal. These results suggest the possibility that gender differences in temperament may be age-related.

Continuity

By definition, temperamental traits are at least moderately stable across time, and most research supports this assertion. This quality does not mean, however, that a person's temperamental profile should be viewed as unchangeable or as someone's "destiny." Studies that measure how strongly a level of a temperament trait at one time point predicts the same trait in the future routinely find modest but statistically significant correlations even up to decades later. Longitudinal assessments of behavioral inhibition have shown that, while most children do not retain that designation years later, the presence of BI at follow-up can be predicted by initial BI status. In general, increasing stability of a temperament trait has been related to (1) higher baseline age of temperament assessment, (2) shorter time intervals between assessments, and (3) individuals who present initially with more extreme levels of the

trait.

An interesting offshoot of the temperament stability question relates to the possibility that physiological "footprints" of traits persist even in the absence of observable behavior. In one study, adults who at the age of 2 years were classified as being behaviorally inhibited were scanned using functional magnetic resonance imaging (fMRI). When shown pictures of novel faces, the previously inhibited adults showed higher amygdalar activity than adults who were not inhibited as toddlers, despite the fact that these subjects did not report any increased anxiety. This intriguing result and others like it suggest the possibility that the physiology underlying a temperamental trait can linger despite the discrete behavioral correlates being lacking. One mechanism through which this may occur is that other brain areas and circuits, such as those involved in executive functioning, might mature with development and keep a person's behavior more regulated despite enduring internal feelings of discomfort.

NEUROBIOLOGY AND ETIOLOGY

There has been an explosion of research into the neurobiological underpinnings of temperament and personality. Although a thorough review of this literature is beyond the scope of this chapter, many important findings are particularly relevant for understanding the boundaries between temperament and psychopathology. For many years, research on the neurobiology of psychiatric disorders existed in parallel with similar efforts in temperament and personality. As these independent efforts began to reveal similar findings, however, there have been increased efforts to assess temperament and psychopathology concurrently.

Genetic Influences

The genetic influence on temperamental traits, albeit at a magnitude less than many had expected, has been well established through a large number of behavioral and molecular genetic studies. Twin and adoption studies across all ages have generally shown that genetic factors explain about 20 to 60 percent of the variability in temperament scores within study populations, depending on age and the particular dimension being examined. These heritability coefficients tend to be highest in twin studies when temperament is assessed by a questionnaire completed by a single informant (such as a parent rating both twins) and lower in adoption designs or studies that measure temperament using observational protocols. Furthermore, this genetic influence has often been found to be mainly of the additive type, in which multiple genes exert small, independent effects. Nonadditive genetic influence, such as in genetic dominance or epistasis, have also been found to a lesser degree in some studies. Although most studies usually do not find major sex differences in

the *magnitude* of genetic factors, it is not uncommon to find evidence that different genes contribute to a particular dimensions between males and females. One counterintuitive finding that has emerged is that the heritability of some traits appears to be fairly low in the neonatal and early infancy period—a time in which some might expect to find the largest genetic influence. Thus, environmental factors play important roles even in infancy, and new genetic and environmental factors come online to contribute to both temperament stability and change across development. These results are paralleled in the anecdotal reports of many parents who marvel at how different their children's temperaments are from each other despite the children sharing half of their genes and being raised in the same household. Indeed, studies have documented only modest correlations between siblings. Although statistically significant, the moderate level of temperamental differences between siblings suggests that the effects of genetic factors are limited and that individual environmental experiences may play a large role in shaping traits over time.

The question of exactly *which* genes contribute to a particular trait is left for molecular genetic studies that attempt to link temperamental traits to chromosomal regions or known alleles of specific genes. Perhaps the most celebrated, albeit inconsistent, finding in molecular genetic studies of temperament has been the purported link between the dopamine 4 receptor gene and the trait of novelty seeking. As shown in Table 39–4, however, a number of polymorphisms on many other genes have also been linked to specific traits. As is the case of genetic studies of many complex diseases, however, each individual gene typically only accounts for a very small percentage (often about 1 percent) of that phenotype. Nevertheless, these studies give clues about some of the underlying neurotransmitters and pathways that may be related to temperament, which in turn can inform comparative studies on corresponding psychiatric disorders.

Environmental Factors

As mentioned, the other major source of influence for temperamental traits comes from environmental factors. These experiences have been documented to have stronger influences on temperament than many have assumed. From a behavioral genetic perspective, environmental factors come in two main types. The first are those events that have a similar effect on all members of a family and are called shared or common environmental events. Socioeconomic status, parenting practices, and school quality are all potential shared environmental factors to the degree that their effect on siblings in a family is equal and in the same direction. Nonshared environmental factors, by contrast, are environmental events that are not shared by members of the same family and can contribute to making siblings different from each other. Examples here could include a child's peer group or a powerful positive or adverse experience.

Table 39–4.
Selected Candidate Genes for Temperamental Traits

Gene	Description	Temperamental Trait
DAT1	Dopamine transporter	Novelty seeking
DRD4	Dopamine 4 receptor	Novelty seeking
5-HT transporter	Serotonin transporter	Harm avoidance
5-HT-2a	Serotonin 2 receptor	Neuroticism
DRD2	Dopamine 2 receptor	Novelty seeking
MAO_A	Monoamine oxidase A	Neuroticism; reward dependence
NPY	Neuropeptide Y	Trait anxiety
β1-adrenergic receptor	β1-adrenergic receptor	Extraversion
BDNF	Brain-derived neurotrophic factor	Trait anxiety
RGS2	Regulator G protein signaling 2	Behavioral inhibition

Many of the same twin and adoption studies that have shown evidence for genetic influence also show unequivocal evidence for environmental influence, particularly of the unshared type. At the same time, there has been much less support for robust shared environmental effects. This pattern does not necessarily mean that things such as parenting or poverty are unimportant, but rather that they may not affect people uniformly to make them more alike within a family.

Genetic–Environmental Interplay

The old nature-versus-nurture debates assumed that genetic and environmental effects work through mutually exclusive pathways. A child inherits a particular gene or experiences a certain event, and those factors affect behavior directly. Although such a mechanism can occur in various psychiatric and nonpsychiatric diseases, it has become increasingly appreciated that genes and environments have bidirectional influences. Environmental events can control gene expression. Conversely, genetically influenced tendencies can cause certain environments to be more likely to occur (gene–environment correlations), and those environments can, in turn, affect behavior. Finally, the effect of a particular gene can depend on the presence of a particular environment and vice versa (gene–environment interactions). Understanding these complex relations has become a major focus of neuroscience research and has brought with it a new perspective on human development that crosses old boundaries.

Neurotransmitters and Brain Circuitry

Whether behavioral influences are genetic, environmental, or an intricate combination of the two, their ultimate effect is to shape specific brain regions and their connections to other areas. Advances in the knowledge of brain function and its plasticity from environmental factors are rapidly rendering obsolete terms such as “biological” to describe some behaviors

but not others. Nevertheless, there is still much to learn about the specific pathways and neurochemical substrates involved in temperament. Many of the specifics remain unknown, and current models and explanations are acknowledged to be vast oversimplifications.

The trait of negative affectivity and related terms such as harm avoidance and behavioral inhibition have often focused on the amygdala and its multiple afferent and efferent pathways. In a study combining genetic and neuroimaging techniques, researchers found significant associations between reduced amygdala/anterior cingulate connectivity and higher levels of harm avoidance. Moreover, this reduced connectivity was also related to having the short allele of the serotonin transporter gene. Amygdalar hyperreactivity has also been postulated to be one of the core features underlying the trait of behavioral inhibition. Ascending pathways from the reticular activating system to the medial septal area and modulated by the orbital frontal cortex have also been related to passive avoidance and behavioral extinction. Related to these functions, dysregulation of the hypothalamic–pituitary–adrenal (HPA) axis has also been found to be associated with this dimension.

Animal studies have implicated ascending serotoninergic pathways from the dorsal raphe to the basal ganglia, substantia nigra, nucleus accumbens, and amygdala. These pathways can be modulated by γ-aminobutyric acid (GABA), as well as by dopaminergic pathways from nigrostriatal regions. Dopamine pathways have frequently been found to play an important role in the approach behavior and spontaneous activity related to the trait of novelty seeking or extraversion. The mutually interacting systems of harm avoidance and novelty seeking may be related to the typical negative correlation between these two traits. Some brain laterality has also been reported in EEG studies, with left frontal activation being related to more positive emotions and approach, and right frontal activation being related to more negative emotions and withdrawal.

Social affiliation and related traits such as reward dependence are hypothesized to be related to noradrenergic pathways from the locus ceruleus, as well as serotoninergic projections from the median raphe, which innervate a number of areas in the cortex, thalamus, and hippocampus. Neuroactive hormones such as oxytocin may also play important roles in social behavior, as illustrated by the marked behavioral differences between two species of voles that correspond to alterations in the distribution of oxytocin receptors in brain reward centers. Dimensions such as persistence or effortful control are thought to be mediated through many areas of the anterior cortex that modify the function of subcortical regions. Many of these pathways also involve dopaminergic transmission. As previously described, the anterior cingulate gyrus is connected both to limbic structures such as the amygdala and to areas of the lateral cortex, thereby providing a link between key emotion processing and executive

function regions.

In describing candidate genes, neurotransmitters, and neuropathways related to temperament, it becomes quickly apparent that the discussion involves all of the "usual suspects" in psychiatric illness. Realization of this repeated finding has prompted more-systematic efforts to delineate the boundaries between temperamental traits and psychiatric illness in the hope of identifying potential targets of therapeutic intervention.

TEMPERAMENT AND PSYCHOPATHOLOGY

Despite the current conceptualization of psychiatric illness in DSM-5 as binary entities that exist in all-or-none form, there is now overwhelming evidence that most all symptoms of psychopathology exist on a continuum. This spectrum ranges from a large segment of the population who present with few or no difficulties in a particular area to a gradually diminishing number of individuals with higher degrees of symptomatology and finally to a small number of people with extreme levels. Having both a quantitative and dichotomous perspective on a particular disease is not unique to psychiatry. Although all people have a measurable blood pressure, for example, individuals with a pressure greater than 140/90 mm Hg are considered to have the diagnosis of hypertension.

As the dimensional aspects of psychiatric disorders have become increasing appreciated and measured, it is not surprising that people have begun to hypothesize that temperament or personality dimensions may underlie this wide distribution of symptom severity. This realization may have been somewhat inevitable, given hypothesized temperament dimensions such as mood, activity level, and attention span, all of which have descriptive similarities to specific psychiatric conditions. Indeed, when the range of psychiatric instruments is extended to include those on the opposite end of the spectrum (e.g., going beyond the distinction between children with attention problems and no attention problems to being also able to distinguish children with average attentional skills from children with excellent abilities), the distribution of the scores begins to approximate the bell-shaped curve of many temperament traits. This understanding has led to the speculation that what is currently classified as psychiatric disease may actually represent merely the extreme tail of normally distributed traits that exist in the entire population.

To distinguish "true" psychiatric illness from subthreshold or developmentally expected levels of behavior, clinical teaching has dictated that one of most important considerations is *impairment*, that is, whether the behaviors, thoughts, or emotions interfere with a person's daily life and functioning. While theoretically appealing, however, research has demonstrated that this important benchmark also tends to exist along a continuum. People who are severely depressed, for example, are quite impaired, whereas those who are a little depressed are somewhat impaired,

and so on, with no convenient cut-off. Complicating matters further in the field of child psychiatry is the fact that impairment is often not defined by the child but rather by those caring for the youngster. These caretakers may differ in their definition of impairment based on their own strengths, weaknesses, and expectations. In comparing two young boys with exactly the same level of motor activity and impulsivity, for example, one could be viewed as extremely impaired in the context of more orderly parents and teachers with high expectations of the child, whereas the other might be considered not impaired by caretakers with a different set of traits and goals.

As psychiatry continues to accept these strong associations between temperament and psychopathology, several mechanistic hypotheses have been proposed to explain this link. These models are summarized in Table 39–5.

The two most investigated potential mechanisms have been described as the *spectrum model* and the *risk or vulnerability model*. The spectrum perspective postulates that psychiatric symptoms and temperament dimensions lie on a common continuum with no clear qualitative distinctions. The difference between a trait and a disorder, therefore, is simply a matter of degree or magnitude. Implicit in this theory is that there also is common underlying neurobiology between temperament and psychopathology. The factors that determine a child's level of normal developing activity, for example, are the same as those that underlie another child with hyperactivity, albeit perhaps more amplified.

The risk or vulnerability model of temperament/psychopathology relations considers temperamental traits and psychiatric disorder to be qualitatively different entities but with temperament acting as a risk or protective factor for future development. Although it is easy to differentiate conceptually between a risk factor such as sunlight exposure and a disease such melanoma, the description of something like shyness being a risk factor for social anxiety disorder lacks these salient points of obvious divergence. Proponents of this mechanism, however, point out that there are many examples of normally distributed variables that mask a number of very different underlying processes. Intelligence quotient, for example, often shows a normal distribution in large samples. To assume, however, that there is a common quantitative mechanism underlying this distribution would miss, among other things, the small numbers of individuals whose score is related to a distinct and qualitatively different mechanism, such as the trisomy of chromosome 21 that leads to Down syndrome. In addition, there is also a hybrid of the spectrum and risk models that proposes true phenotypic differences between temperament and psychopathology but some common underlying features, such as shared genetics, that are then acted on by other factors, such as specific adverse or positive experiences to create important points of divergence.

Table 39–5.
Possible Mechanisms of Temperament/Psychopathology Associations

Model	Description	Example
Spectrum/continuum	Most psychiatric disorders are best conceptualized as existing on a continuum with temperament or personality traits.	Activity level is a temperamental trait, and people with attention-deficit/hyperactivity disorder are at the end of that spectrum.
Risk/vulnerability	Temperament is different from psychopathology but can increase the risk of its occurrence.	People with high levels of novelty seeking are at increased risk for substance abuse.
Common factor	Temperament and psychopathology share common etiological factors, which are then acted on by other factors that lead some people to develop a disorder.	People with posttraumatic stress disorder share similar genes with temperamentally anxious people but have experienced a severe trauma.
Pathoplastic	Temperament and psychopathology are unique but have bidirectional influences on each other.	The long-term prognosis of depression is better for individuals with high effortful control.
Scar	The pathophysiology of psychiatric disorders can alter temperament.	The plaques and tangles associated with Alzheimer disease cause personality changes that are observed prior to other symptoms.

These three models tend to be the most dominant in the literature when it comes to temperament/psychopathology interactions. However, two other models, both of which preserve the uniqueness of pathophysiological processes, deserve mention. In the pathoplastic model, temperament and psychopathology are distinct but can mutually influence one another. Temperament can increase the chance that a child will suffer a major depression, for example, and that depressive episode can, in turn, exert effects on the child's personality going forward. Finally, the "scar effect" emphasizes the separation of personality and psychopathology but states that psychiatric disorders can modify personality. An example of this effect may be Alzheimer disease. Although personality change is common with this condition and often precedes more specific Alzheimer symptoms, it is thought that the early-developing pathology of plaques and tangles affects personality, rather than the other way around.

Scientists have thus turned to modern research designs and tools to help disentangle these complex relations among temperament, personality, and psychopathology. In so doing, there has been a realization that what were previously thought to be clear boundaries might be much more muddled. What appears to be a continuous trait, like intelligence, may develop from a number of small, discrete causes. Conversely, what appears to be a binary yes/no illness might have more embedded gradations when examined closely. In Huntington disease, for example, the number of

trinucleotide repeats in the huntingtin gene affects the onset and course of illness. It is also important to keep in mind that the various mechanisms are not mutually exclusive. Within a particular trait–disorder pair, there may not be a single mechanism at work across different individuals and perhaps even within individuals. Some individuals with obsessive-compulsive disorder may have strong temperamental correlates, whereas the symptoms of other individuals might be the result of an autoimmune process.

As is clear from these examples, improved understanding of the way illness is conceptualized has important clinical implications. From a research standpoint, it is vital to ensure that any vertical boundaries, should they exist, between temperamental traits and psychiatric illness be delineated because all results downstream of that categorization are likely to be affected. If children who just barely meet DSM-defined ADHD have a brain physiology that is nearly identical to children who just miss the diagnoses, then using the latter group as controls in a functional MRI study, despite their lack of diagnosis, is unlikely to yield fruitful results. Thus, the next section is devoted to describing the current state of knowledge with regard to the existence and mechanism of temperament/psychopathology relations (Table 39–6).

ADHD and Disruptive Behavior Disorders

Attention-deficit/hyperactivity disorder (ADHD) is a prime example of a disorder that at least on the surface appears to be largely in the upward range of temperamental traits related to activity level and attention span. Indeed, it is likely that much of the public controversy surrounding this diagnosis is related to the perception that what is being diagnosed as illness is normal temperament, albeit at somewhat higher levels than are typically observed. As might be expected, low levels of effortful control and persistence have been found, in addition to higher levels of novelty seeking and extraversion, especially among children who meet criteria for the hyperactive-impulsive or combined subtype of ADHD. As mentioned, molecular genetic studies of novelty seeking and ADHD have converged on similar candidate genes, often related to dopaminergic transmission. Furthermore, a recent study of 8,500 twin pairs found evidence for a continuum model of ADHD with nearly identical heritability estimates (around 0.60) for those with diagnostic versus subthreshold levels of behaviors. Another recent study subtyped a group of children with ADHD according to temperament and identified three groups that they labeled mild, surgent (with high levels of approach behavior), and irritable (with higher emotionality and poorer regulatory skills). These temperament-based groups, compared to traditional DSM-based presentations, were shown to provide better cohesion with regard to stability over time, functional brain connectivity, and cardiac physiological parameters.

Table 39–6.
Associations between Temperament Traits and Various Categories of Psychiatric Disorders

	Negative Affectivity	Extraversion	Effortful Control
Attention-deficit/hyperactivity disorder	High	High	Low
Other disruptive behavior disorders	High or low	High	Low
Major depression	High	Low	Low
Anxiety disorders	High	Low	Low
Substance use	High	High	Low
Autism	High	Low	Low
Eating disorders	High	—	High (anorexia) or low (bulimia)

Adapted from Rettew DC. Child Temperament: New Thinking About the Boundary Between Traits and Illness. New York: W. W. Norton & Company; 2013.

Other disruptive or externalizing disorders include oppositional defiant disorder (ODD), conduct disorder, and, in adults, antisocial personality disorder. All of these entities are rather heterogeneous, making the task of linking specific temperamental traits with these disorders difficult. Aggression, for example, is often divided into more instrumental and reactive subtypes. In the former, sometimes known as type 1 psychopathy, aggressive acts can be planned in advance and without provocation. These behaviors are often associated with individuals who tend to lack empathy and possess what are termed callous and unemotional traits, now incorporated into the conduct disorder diagnosis in DSM-5. Physiologically, many of these individuals have been shown to have an underaroused autonomic nervous system that craves additional stimulation. In reactive or type 2 psychopathy, individuals quickly perceive threat from others and have difficulty modulating their emotional responses. This latter type is more strongly linked to deficits in emotional regulation, as well as to increased negative affectivity. Oppositionality has similarly been found to break into three dimensions, labelled irritable, headstrong, and hurtful. These dimensions show different comorbidity patterns with regard to other DSM disorders.

Mood and Anxiety Disorders

Often collectively referred to as internalizing disorders, mood and anxiety disorders are some of the most common psychiatric conditions in childhood and adolescence. Although they are considered to be separate categories of illness according to current clinical nomenclature, there is good empirical evidence to support their being closely linked. Consistent with this hypothesis is the finding of similar temperamental correlates

common to both mood and anxiety disorders. By far the strongest association is with the dimension referred to as negative affectivity, neuroticism, or harm avoidance. Major depression has also been associated with the combination of high negative affectivity and low approach. Bipolar disorder, especially the purported "broad" phenotype characterized by extreme irritability, explosiveness, and mood lability and with a lack of clear episodes, euphoria, or grandiosity, has been shown to be related to a somewhat unique temperamental combination of high novelty seeking *and* high harm avoidance. This "approach plus avoidance" pattern stands in contrast to the general tendency of novelty seeking and harm avoidance being inversely related to each other.

Especially in the consideration of more generalized and chronic internalizing disorders such as generalized anxiety disorder (GAD), dysthymia, and generalized social anxiety disorder, there again has been speculation that many mood and anxiety disorders are best conceptualized as the high end of a spectrum of negative affectivity. Direct analyses of this hypothesis, however, suggests that such a model is oversimplified. As shown in Figure 39–2, many people who possess even very high levels of harm avoidance do not meet criteria for GAD, whereas some individuals with moderate levels of harm avoidance do. On the genotype level, behavioral genetic studies indicate that approximately half of the genes that contribute to neuroticism are the same genes that contribute to internalizing disorders, whereas half do not. Thus, despite the obvious content overlap between descriptions of anxious temperament and descriptions of anxiety and depressive disorders, the majority of evidence appears to point to temperament dimensions having more complex relations with these disorders rather than simply being a different point along a continuum of a single trait. The cyclic and episodic nature of other types of mood disorders, such as recurrent major depression or panic disorder, further suggests that elements other than temperament need to be incorporated into the current theoretical models.

Low effortful control or persistence has further been associated with internalizing disorders. Indeed, temperamental dimensions relating to regulatory control appear to be a common correlate to most psychiatric disorders, with other dimensions related more to the specific *type* of psychopathology. As mentioned, the construct of behavioral inhibition developed by Kagan and colleagues likely combines elements of high negative affectivity, low approach, or surgency, and perhaps low effortful control. Longitudinally, however, the strongest association has been between BI and generalized social anxiety disorder with a meta-analysis revealing that 43 percent of young BI children meeting criteria for social anxiety disorder later in development.

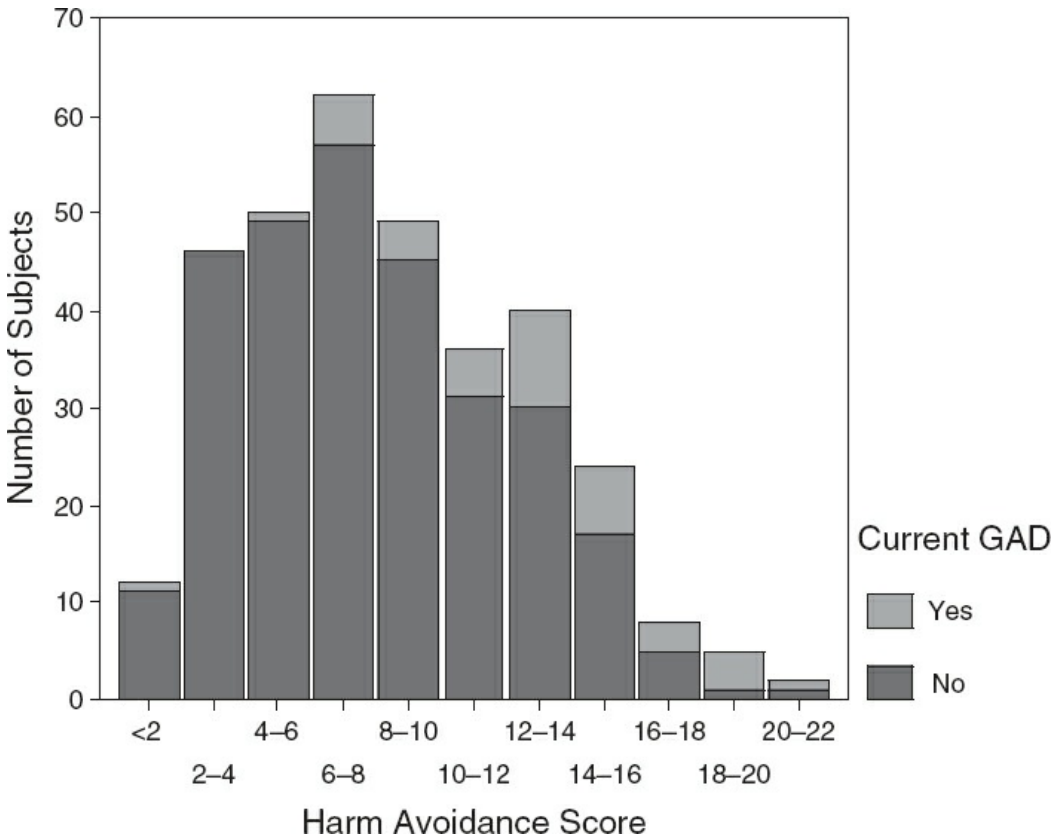

FIGURE 39–2. Distribution of harm avoidance (HA) scores in children and adolescence. Subjects whose results are given also have generalized anxiety disorder (GAD), according to the criteria of the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV). The pattern reveals substantial but not complete overlap between harm avoidance and GAD, with a few subjects with high HA without GAD, some individuals with moderate HA with GAD, and very few individuals with low HA with GAD. (From Rettew DC, Doyle AC, Kwan M, Stanger C, Hudziak JJ. Exploring the boundary between temperament and generalized anxiety disorder: a receiver operating characteristic analysis. *J Anxiety Disord*. 2006;20[7]:931–945.)

Autistic Spectrum Disorder

To many clinicians observing a nonverbal child with stereotypical behaviors, autism would seem to be one group of disorders that truly is more binary in nature. However, accumulating evidence has demonstrated that many of the core symptom clusters of autism, such as social cognition, communication, and affiliation, as well as the stereotypical behaviors themselves, exist along a continuum when measured quantitatively. Temperament dimensions such as low sociability or reward dependence, high negative affectivity, and low extraversion have all been associated with autism. At the same time, it is important to remember that the finding of a continuous distribution of behaviors does not prove that a common neurobiological mechanism underlies the entire spectrum. Indeed, recent studies showing focal areas of cytoarchitectural disorganization in the neocortex of individuals with autism suggest that, at least for some individuals with autism, a scar mechanism may be at work in which the

pathophysiology of the disorder drives the temperamental differences that are also observed. Further support for the possibility that qualitative differences in brain structure and function exist between individuals who meet criteria for autism and those who possess some of the same traits (albeit at lower levels) comes from the neuroimaging literature, although much more work is needed to explore this intriguing hypothesis.

Other Disorders

Some types of disorders are defined by behaviors that are more binary in nature and thus do not lend themselves as easily to spectrum hypotheses with temperamental traits. Discrete actions such as a binging and purging in eating disorders, hallucinations in psychotic disorders, and the taking of illicit drugs in substance use disorders may fall into this area. Many of these entities also tend to have onsets in adolescence or later and thus have most of the research directed at personality traits rather than temperament per se.

In substance use disorders, the wide variety of substances that are abused, each with unique physiological effects, might be sought after by individuals with different temperamental profiles. Different agents also vary in their legality, which itself could lead to unique associations with particular traits. Alcohol use, for example, is legal for adults but not children and thus may have different temperamental associations by age. One follow-up study of temperament found indications that adult alcohol abuse was associated with being both temperamentally undercontrolled and overcontrolled as a child. Novelty or sensation seeking, especially when defined apart from positive emotionality, has also been associated with increased substance use. Once again, dimensions such as effortful control are related to lower amounts of substance use disorders. An interesting caveat in personality research in substance use acknowledges the shifting cultural perceptions of particular substances over time. Cigarette use, for example, has become viewed as increasingly deviant over the last 50 years across many segments of Western culture. One might expect, consequently, that its use is less associated with traits such as agreeableness over time or compared to cultures in which its acceptance level is greater.

Whereas most individuals with eating disorders regardless of specific diagnosis share a preoccupation with weight and self-image, there also appear to be substantial differences that generally fall along diagnostic lines, particularly between the restricting subtype of anorexia nervosa (AN) and bulimia nervosa (BN), and these may reflect underlying differences in temperament and personality. In particular, low effortful control and persistence has been found with BN but not in AN, in which it has been found in some cases to be higher than in controls. High neuroticism or harm avoidance has also been found for both disorders and parallels the frequent finding of anxiety disorder that can co-occur or precede eating

disorders.

Temperamental associations with psychotic disorders have been less studied, although parallels especially with the negative symptoms of schizophrenia would appear to be a fertile ground for investigation. High harm avoidance has been found in both patients with schizophrenia and their relatives, whereas high novelty seeking may be associated with more violent behavior. While positive symptoms such as hallucinations might be assumed to be relatively rare in the population and be present in more of an all-or-none form, recent research has suggested that lower-level forms of psychotic thinking do exist and are more prevalent than previously thought.

A Case Example

Previous "diathesis–stress" models of psychopathology that include temperament often frame temperamental traits as the diathesis component, which is then acted upon by various stressors, such as parenting behavior, traumatic events, and so on to produce psychiatric symptoms. While data described in this chapter reveal that such a model is likely overly simplistic when it comes to how temperament is conceptualized, certain aspects of this model can still be useful. One of the principal advances from older conceptualizations of the diathesis–stress model is the full appreciation that the stressors often do not act on the diathesis randomly but rather are probabilistically more likely to occur because of those predispositions. As previously discussed in the section on temperament development, these processes have been called genotype–environment correlations and are readily applicable to the understanding of how risk factors like certain temperamental traits contribute toward the creation of environments such as parental behaviors and peer groups that accentuate those initial predispositions. A case example may illustrate this mechanism best.

Kayla was a child who from birth appeared to be more fearful and easily upset than other infants. She cried frequently and forcefully, and as a toddler she was more reluctant to leave her mother's side to play with other children and to explore new situations. Her behavior was a challenge to her mother, who herself struggled with significant anxiety and shyness. Her mother was not prone to seek out activities such as play dates with other parents due to her own social anxiety, and when she did, Kayla's behavior made them frequently go badly. Hoping to avoid another embarrassing scene, Kayla's mother became increasingly overprotective in her efforts to shield her daughter from situations that might provoke discomfort. By the time kindergarten started, Kayla had had significantly fewer opportunities to interact with other children and tended to keep to herself. Throughout elementary school, she avoided team sports and other structured social activities, preferring other activities such as learning to play the piano. She did manage to make some lasting friendships, although they tended to be with other children who were also quiet and reserved. Kayla became an easy target of teasing by others, and as an adolescent she became increasingly self-critical about her demeanor. At the age of 14 years, she saw a psychiatrist for the first time and was diagnosed with social anxiety disorder and major depression.

This illustration demonstrates how insidious the progression can be from temperamental trait to psychiatric disorder and how difficult it can be to ascribe the causation of symptoms as purely genetic or environmental. The case also illustrates the role of other factors in shaping a temperamental trajectory toward more maladaptive outcomes.

OTHER CONCEPTS IN TEMPERAMENT–PSYCHOPATHOLOGY RELATIONS

Parenting Behavior

The fact that some parenting behavior is evoked by children does not mean that it is either unimportant or inevitable. Neither does it mean that suboptimal parenting is the fault of the children. Rather, a more transactional view appreciates the bidirectional influences of children and parents that together can create important home environments with important developmental effects. Research has shown, for example, that parents of anxious children are more likely to teach avoidant rather than prosocial solutions to problems and may facilitate the perception of threat in ambiguous situations. As shown in the case example, these parents can become overprotective, which limits the crucial exposure opportunities that help children to overcome their anxieties. Lack of parental warmth, poor supervision, harsh and inconsistent discipline, and obviously abuse are all important parental behaviors that have been found to be related to various types of psychopathology and serve as critical intermediaries between temperament and psychopathology.

Cognitive Factors

Although some cognitively laden attributes, such as attentional skills, are often included under the temperament umbrella, other aspects, such as intelligence, processing speed, and memory, are generally considered outside of the temperament domain. Nevertheless, these cognitive processes likely also play important moderating roles. Children with relatively diminished attention spans but above-average intelligence, for example, may be able to use that intellectual ability to counteract their attention deficits and prevent the clear impairment required for a diagnosis of ADHD. Other cognitive functions, such as the attributional style that a person applies to interpret themselves and the world around them, also likely shape and are shaped by temperament. These types of appraisals and attributions have been found to be important in a variety of both internalizing and externalizing disorders, from antisocial disorders to mood and anxiety.

Trauma and Adverse Events

Clearly, both single-episode traumas, such as natural disasters and car accidents, and more chronic stressors, such as some forms of abuse and

abject poverty, can propel temperamentally at-risk children toward full-blown psychiatric symptoms. At the same time, it is becoming increasingly clear that one of the strongest predictors of more severe symptoms after a stressful event is poorer functioning prior to that event. Thus, a temperamentally aggressive child might be more likely to develop antisocial behaviors in the face of significant stressors, and a more anxious child might develop posttraumatic stress disorder after a single traumatic event.

Peer Groups and Activities

A powerful behavioral influence, especially for older children and adolescents, comes from their peers and the activities that are done together. Similar to parenting behaviors, it is very likely that peer groups have a bidirectional influence on temperament and personality traits. A high-novelty-seeking child, for example, is likely to have similar-minded friends, who may push each other to ever riskier and more dangerous activities. If many of these group experiences occur within the context of a supervised team or structured activity, there can be additional opportunities to channel temperamental tendencies productively and to provide opportunities to overcome more maladaptive propensities. Research has shown, for example, that adolescents on sports teams are less likely to smoke cigarettes. Given that participation in sports teams appears to have its strongest environmental influence early in life, a possible intervention target emerges for future research.

Temperament and Resiliency

Just as some temperamental traits can mix with the aforementioned variables to generate negative momentum toward psychopathology, other traits can serve as protective factors in the face of adversity and indeed may be important not only for a lack of psychopathology, but also for true mental health and wellness. The previously mentioned study using latent profiles analysis found that compared to a group of children with moderate levels of most temperament traits, the “steady” group, characterized by high levels of persistence and low levels of novelty seeking, performed better at school by parent and teacher report, engaged in more activities, and functioned better socially. Just like the at-risk traits, these dimensions are likely operating with a number of other factors through the same mechanisms of gene–environment correlations and interactions. The Project Competence study, which followed children into adulthood, found robust predictions between child temperament/personality and adult functioning even after controlling for a number of baseline and demographic characteristics. In general, the temperament dimensions did not interact significantly with adversity, again suggesting a wider effect of

temperament than simply guarding against the effects of psychosocial stressors.

Goodness of Fit

Although it is easy to slip into the notion that some temperament traits are good and others are bad, it is important to remember that all traits possess both positive and negative qualities. High novelty seeking, for example, contains generally positive elements of being energetic and spontaneous in addition to being prone toward extravagance and impulsivity. Thomas and Chess popularized the notion of goodness of fit decades ago and it remains a cornerstone of child development theory. Simply put, the idea behind goodness of fit is that particular temperamental traits are not positive or negative per se but become so in an environmental context. A fearful temperament, for example, might be very adaptive in more dangerous settings but becomes less useful when the environment is safer and more predictable.

Research in animals and humans has supported this important hypothesis, although not necessarily at extreme levels. In a study by Rettew and colleagues, certain child temperament traits were related to behavior problems only in the presence of certain parental traits (Fig. 39–3). High child novelty seeking showed no relation to attention problems in families with a low-novelty-seeking mother, but did show a high degree of attention problems when paired with a high-novelty-seeking mother. Indeed, the much maligned short-repeat allele of the serotonin transporter, which has been related to the trait harm avoidance, may be related to environmental sensitivity, thus conveying adaptive *benefit* in some families while having a more negative effect in others.

CLINICAL IMPLICATIONS

It is clear at this point that the benefits of improved understanding regarding precisely how temperamental predispositions push children toward or away from more pathological outcomes have much more than academic implications and may guide major efforts in early identification and intervention. If negative cycles of genetic–environmental interplay leading to disabling symptoms can be identified, they potentially can be stopped, perhaps even reversed. At the same time, more positive cycles can be cultivated and encouraged. Returning to the case example may be worthwhile.

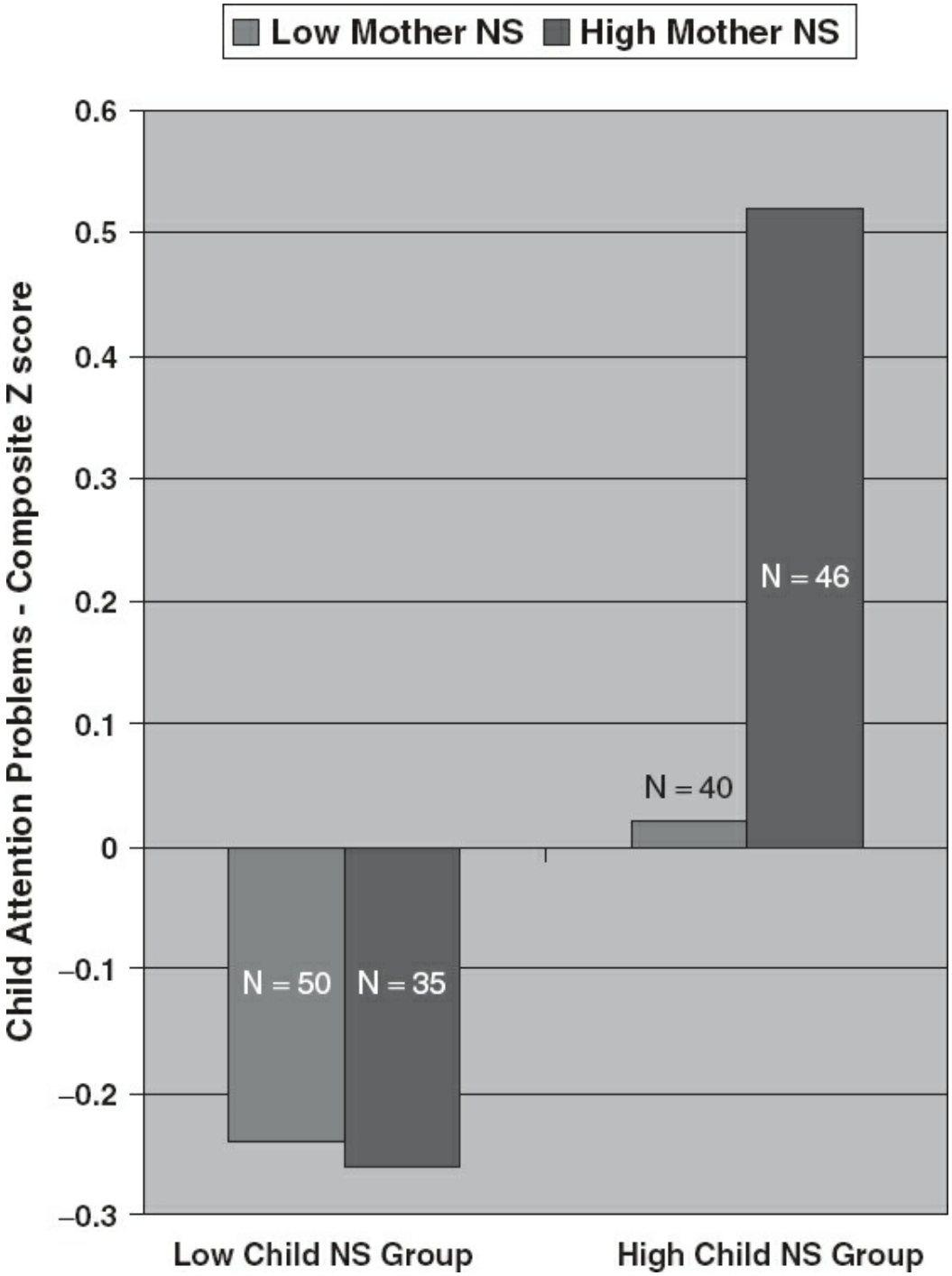

FIGURE 39–3. Relations between child and mother novelty seeking (NS) and attention problems. Children with low NS show few attention problems regardless of maternal NS. In high-NS children, however, a higher degree of attention problems is found in families with a high-NS mother but not in families with a low-NS mother. Results support the construct of goodness of fit in temperament/psychopathology associations. (From Rettew DC, Stanger C, McKee L, Doyle A, Hudziak JJ. Interactions between child and parent temperament and child behavior problems. *Compr Psychiatry*. 2006;47[5]:412–420.)

After diagnosing Kayla with major depression and social anxiety disorder, her psychiatrist recommended cognitive–behavioral therapy to help Kayla overcome her dysfunctional beliefs and provide motivation for her to learn and apply new skills. He also recognized similar symptoms in Kayla's mother and recommended her for psychiatric evaluation, which resulted

in her being prescribed a selective serotonin reuptake inhibitor (SSRI) and receiving individual treatment. Over time, Kayla learned to identify and challenge many of the negative thoughts that surrounded her behavior. Through the use of incentives, she was able to take small risks in approaching others socially. Her therapist recommended that she use her musical skills to an advantage, such using her piano playing as a mechanism through which other people would be encouraged to approach *her*. Her initial gains propelled her to take additional risks, and she joined the school orchestra. Meanwhile, her mother found that the decrease of her own anxiety allowed her to implement the therapist's suggestions more easily. Although Kayla never became a social butterfly, her depressive symptoms resolved, and she graduated from high school with confidence in her abilities and comfort with who she was.

Clinically oriented scientists are working hard to harness the energy that causes temperamentally predisposed tendencies to snowball into full-blown psychiatric syndromes and create pathways that not only will lead out of pathology but also into wellness. Given the complexity of these interactions, it is exceedingly likely that the specific relations between a particular trait and a particular disorder will vary considerably among individuals. Nevertheless, there is tremendous promise that further study of temperament will not only dramatically alter some of the fundamental definitions of psychiatric disorders but also illuminate mechanisms that can improve the lives of children.

References

*Buss AH, Plomin R. *Temperament: Early Developing Personality Traits*. Hillsdale, NJ: Lawrence Erlbaum; 1984.

Clauss JA, Blackford JU. Behavioral inhibition and risk for developing social anxiety disorder: a meta-analytic study. *J Am Acad Chi Adolsc Psychiatry*. 2012;51:1066–1075.

Cloninger CR, Svrakic DM, Przybeck TR. A psychobiological model of temperament and character. *Arch Gen Psychiatry*. 1993;50:975–990.

Else-Quest NM, Hyde JS, Goldsmith HH, Van Hulle CA. Gender differences in temperament: a meta-analysis. *Psychol Bull*. 2006;132:33–72.

Gray JA. *The Psychology of Fear and Stress*. New York: Cambridge University Press; 1987.

Hettema JM, Neale MC, Myers JM, Prescott CA, Kendler KS. A population-based twin study of the relationship between neuroticism and internalizing disorders. *Am J Psychiatry*. 2006;163:857–864.

Johnson W, Hicks BM, McGue M, Iacono WG. Most of the girls are alright, but some aren't: personality trajectory groups from ages 14 to 24 and some associations with outcome. *J Pers Soc Psychol*. 2007;93:266–284.

Kagan J. *Galen's Prophecy*. Boulder, CO: Westview Press; 1994.

Karalunas SL, Fair D, Musser ED, Aykes K, Iyer SP, Nigg JT. Subtyping attention-deficit/hyperactivity disorder using temperament dimensions: toward biologically based nosologic criteria. *JAMA Psychiatry*. 2014;71:1015–1024.

Kaufman J, Yang BZ, Douglas-Palumberi H, Houshyar S, Lipschitz D. Social supports and serotonin transporter gene moderate depression in maltreated children. *Proc Natl Acad Sci U S A*. 2004;101:17316–17321.

Larrson H, Anckarsater H, Råstam M, Chang Z, Lichtenstein P. Childhood attention-deficit hyperactivity disorder as an extreme of a continuous trait: a quantitative genetic study of 8,500 twin pairs. *J Child Psychol Psychiat*. 2012;53:73–80.

McCrae RR, Costa PT Jr, Ostendorf F, Angleitner A, Hrebickova M. Nature over nurture: temperament, personality, and life span development. *J Pers Soc Psychol*. 2000;78:173–186.

Munafo MR, Clark TG, Moore LR, Payne E, Walton R. Genetic polymorphisms and personality in healthy adults: a systematic review and meta-analysis. *Mol Psychiatry*. 2003;8:471–484.

Oldehinkel AJ, Hartman CA, De Winter AF, Veenstra R, Ormel J. Temperament profiles associated with internalizing and externalizing problems in preadolescence. *Dev Psychopathol*. 2004;16:421–440.

*Rettew DC. *Child Temperament: New Thinking about the Boundary between Traits and Illness*. New York: W. W. Norton & Company; 2013.

Rettew DC, Althoff RR, Dumenci L, Ayer L, Hudziak JJ. Latent profiles of temperament and their relations to psychopathology and wellness. *J Am Acad Child Adolesc Psychiatry*. 2008;47(3):273–281.

Rettew DC, Doyle AC, Kwan M, Stanger C, Hudziak JJ. Exploring the boundary between temperament and generalized anxiety disorder: a receiver operating characteristic analysis. *J Anxiety Disord*. 2006;20:931–945.

Rettew DC, Stanger C, McKee L, Doyle A, Hudziak JJ. Interactions between child and parent temperament and child behavior problems. *Compr Psychiatry*. 2006;47:412–420.

Roberts BW, DelVecchio WF. The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. *J Pers Soc Psychol*. 2000;126:3–25.

Ronald A, Sieradzka D, Cardno AG, Haworth CM, McGuire P, Freeman D. Characterization of psychotic experiences in adolescence using the Specific Psychotic Experiences Questionnaire: findings from a study of 5000 16-year-old twins. *Schizoph Bull*. 2014;40:868–877.

*Rothbart MK, Ahadi SA, Evans DE. Temperament and personality: origins and outcomes. *J Pers Soc Psychol*. 2000;78:122–135.

*Rutter M, Dunn J, Plomin R, Simonoff E, Pickles A. Integrating nature and nurture: implications of person–environment correlations and interactions for developmental psychopathology. *Dev Psychopathol*. 1997;9:335–364.

Shiner R, Caspi A. Personality differences in childhood and adolescence: measurement, development, and consequences. *J Child Psychol Psychiatry*. 2003;44:2–32.

Slagt M, Dubas JS, Dekovic´ M, van Aken MAG. Differences in sensitivity to parenting depending on child temperament: a meta-analysis. *Psych Bull*. 2016;142:1068–1110.

Stergiakouli E, Martin J, Hamshere ML, et al. Shared genetic influences between attention-deficit/hyperactivity disorder (ADHD) traits in children and clinical ADHD. *J Am Acad Child Adolesc Psychiatry*. 2015;54:322–327.

Stringaris A, Goodman R. Three dimensions of oppositionality in youth. *J Child Psychol Psychiatry*. 2009;50:216–223.

Stoner R, Chow ML, Boyle MP, Sunkin SM, Mouton PR, Roy S, et al. Patches of disorganization in the neocortex of children with autism. *New Eng J Medicine*. 2014;370:1209–1219.

Sylvester DM, Barch DM, Harms MP, et al. Early childhood behavioral inhibition predicts cortical thickness in adulthood. *J Am Acad Child Adolesc Psychiatry*. 2016;55:122–129.

*Thomas A, Chess S. *Temperament and Development*. New York: Bruner/Mazel; 1977.

Whittle S, Yucel M, Fornito A, et al. Neuroanatomical correlates of temperament in early adolescents. *J Am Acad Child Adolesc Psychiatry*. 2008;47:682–693.

40

Intellectual Disability

BRYAN H. KING, M.D., KAREN E. TOTH, PH.D., NINA DE LACY, M.D., AND DAN DOHERTY, M.D., PH.D.

HISTORICAL OVERVIEW

While centuries of experience have demonstrated that stigma will not disappear with a name change, the field continues to wrestle with how to capture the condition now referred to as "intellectual disability" or "intellectual developmental disorder." For decades, the field has also worked hard to encourage "people first" language. Individuals should not be defined by their illnesses or conditions—for example, John is not a "leper," or a "schizophrenic," or "retarded." Instead, people have illnesses or conditions—for example, John has schizophrenia. Although these lexical and grammatical sensitivities may seem trivial for some, there is no faster way to lose credibility in some settings than to characterize one's patient as "retarded."

Why such heightened sensitivity in this field? While it is well recognized that intellectual gifts have been distributed on a continuum for as long as mankind has walked the earth, the understanding and treatment of persons with intellectual disability has moved like a pendulum between extremes. At one end, persons with intellectual disability have been exalted, and considered "Les enfants du Bon Dieu" ("children of the Good God"). Certain media may convey the message that persons with intellectual or developmental disability are somehow blessed with simpler, more straightforward understandings of basic human truths. And at the other end of the pendulum swing, such persons have also been vilified. The word "cretin" (congenital hypothyroidism), takes its origin from "Christian or Christ-like," even as its modern dictionary definition includes descriptors like "stupid, vulgar, and insensitive." In the introduction to his text on "mental defectives" a century ago, Barr (1904), the chief physician at the Pennsylvania training school for the feeble-minded, observed that: "The terms idiot and imbecile indiscriminately applied for centuries to a despised and neglected class, have naturally, through the mere force of association, become synonymous with and even expressive of opprobrium and reproach. So closely welded has been this association of word and idea, that even the advance of the nineteenth century with all its revelation of true conditions has failed to sunder them" (p. 17).

Evidence for the recognition and treatment of intellectual disability

dates to the earliest of medical writings. Hippocrates described microcephaly and craniostenosis, and Galen actively explored causes of cognitive disability. In the Middle Ages, Avicenna proposed treatments for meningitis and hydrocephalus, and even defined levels of intellectual function.

Esquirol (1838) is credited as the first medical writer to define the disorder and was careful to distinguish intellectual disability from mental illness or disease. He focused on congenital onset in noting that: "Idiocy is not a disease, but a condition in which the intellectual faculties are never manifested; or have never been developed sufficiently to enable the idiot to acquire such an amount of knowledge as persons of his own age, and placed in similar circumstances with himself, are capable of receiving. Idiocy commences with life, or at that age which precedes the development of the intellectual and effective faculties, which are from the first what they are doomed to be during the whole period of existence."

The modern history for the field of intellectual disability begins in the late 18th and early 19th centuries. At that time, Jean-Marc Itard attempted a natural experiment to educate Victor, a "wild-child" discovered in the forests of Aveyron, France. Although Itard himself judged his work with Victor a failure, this famous experiment marked the important consideration that persons with disabilities might be educable. Itard's efforts generated interest in educational and other interventions for persons with disabilities. In some ways, Itard and those that followed provided a bridge from beliefs that individuals with intellectual disability were subhuman (even suggested by Linnaeus to be a separate species in the genus Homo with names like: Juvenis Lupinus Hessensis—a young man found in Hesse among wolves; or Juvenis Hannoverianus—a young man found in Hanover), to a more enlightened appreciation of their humanity and their potential.

Even so, in the preface to his text, Barr (1904) still straddles this divide as he categorically writes off the severely affected, while advocating for the less impaired:

"In endeavoring to emphasize the utter hopelessness of cure, and also the needless waste of energy in attempting to teach an idiot, I have sought to make clear the possibilities that may be attained in the training of the imbecile, the urgent need of preventing the backward child from degenerating into imbecility, and of safeguarding the absolutely irresponsible ammoral imbecile from crime and its penalty."

Howe and Wilbur are credited with introducing the enlightened approach to educating individuals with intellectual disability to the United States with the establishment of training schools like those that were appearing in Europe.

By the turn of the century, there were 30 such institutional settings in 21 states, both public and private, and some reserved for women, and for

individuals with epilepsy. Many of these institutions had residents numbering in the hundreds and gradually the focus shifted away from education and more toward custodial care.

During their peak usage, from 1950 until the late 1960s, institutions were home to approximately 1 per 1,000 Americans.

Since the mid-1960s, many fewer persons with intellectual disability have been institutionalized in the United States. This move toward "deinstitutionalization" has arisen from many sources. First, the overcrowding and neglect common in many large institutions came to light in various exposés during the late 1960s. Blatt and Kaplan's photo story, "Christmas in Purgatory" (reproduced in Look magazine), Geraldo Rivera's television reports, and Robert F. Kennedy's attacks on the Rome and Willowbrook institutions in New York all shocked the nation.

Deinstitutionalization

Other forces also led to deinstitutionalization. Perhaps the most important of these forces was the philosophy of "normalization," the idea that individuals with intellectual disability were entitled to a more normal lifestyle, including a culturally normal rhythm to one's day (school or work, leisure time, sleep), week (weekdays and weekends), and year (vacations, holidays). In 1972, Wolfensberger extended the idea of normalization to the service delivery system itself, calling on all residences, schools, and other services for persons with intellectual disabilities to be as normative as possible. Parent and professional advocacy groups also fought hard for legislative and legal victories to decrease the size of large institutions.

Over the past 50 years, changes in US residential institutions have been dramatic. Several states have closed their institutions, and remaining facilities have become much smaller. Typically, individuals remaining in large or even medium-sized institutions have the most severe impairments; many of these residents also have severe behavior problems, or motor or sensory disabilities in addition to their severe-profound intellectual disability. Overall, from 1967 to 1994, the total institutional population—defined as state-operated institutions with 16 or more residents—decreased by almost two-thirds, from 194,650 to 65,818. In California, the number of persons served by the Department of Developmental Services increased 61.6 percent between January 1992 and December 2001. During this same period however, the State developmental center population decreased by 44.4 percent. In many states, Medicaid funding waivers (the ability to access Medicaid dollars in circumstances in which family income criteria or other eligibility criteria are waived) are in place explicitly to support children with intellectual disabilities who meet "institutional level of care" standards. These so-called "deeming waivers," or "Katie Beckett waivers" are meant to ensure that Medicaid support can follow an individual out of an institutional setting if such a move can be done safely and cost

effectively. Katie Beckett was a young lady in Iowa who became ventilator-dependent following a catastrophic CNS infection, and whose parents discovered that if they took her home from the skilled nursing facility, she would lose her Medicaid coverage. As the costs associated with institutionalization are invariably in the tens of thousands for any individual (ranging from a low of $104,025 per year in Arkansas to a high of $375,000 in Tennessee in 2009), the cost effectiveness criterion is rarely a limiting factor, but clinicians may be called upon to attest to the level-of-care criterion for their patients who are seeking support under this waiver.

School. Schooling is the other major service for persons with intellectual disability. Following the move to "school" all American children that began in the mid-1800s, teachers and administrators soon discovered that a subset of children were having difficulties in performing school lessons. As a result, classes for "problem children" began in many cities and towns in the late 1800s. In 1894, the first classes specifically for children with intellectual disability were formed in Providence, Rhode Island. In rapid succession, special education classes were begun in Springfield, MA (1894), Chicago (1898), Boston (1899), New York (1901), Philadelphia (1901), and Los Angeles (1902).

Local school districts have historically held an ambivalent attitude toward special education classes. Often housed within the least desirable buildings and rooms, these classes were frequently furnished with scarce equipment and materials, and had few specially trained teachers. By 1968, Lloyd Dunn declared that most children with intellectual disability could be "mainstreamed" in classes with age-mates; he questioned the need for segregated special education classes for most children with intellectual disability. Dunn's article took note of the Supreme Court's ruling that "separate could not be equal" in educational settings, that research had not shown that children with mild intellectual disability learn better in special education versus regular education classes, and that educational techniques had advanced sufficiently to allow for the effective schooling of most children with retardation alongside other children.

In both education and living settings, then, the post-1970 era witnessed the strong influence of mainstreaming, community living, and normalization. On the whole, such movements have proven beneficial as persons with intellectual disability increasingly take their rightful place within modern society. Many professionals do, however, question whether normalization has sometimes gone too far. Not every child with intellectual or other disabilities may be able to be schooled optimally with typically developing age-mates, nor might every adult be able to live independently in the community. Indeed, some families are advocating for specialized schools or residential settings for children with autism even as state agencies fear a return to institutional approaches. These issues continue to

be hotly debated among professionals, families, and organizations associated with persons with intellectual disability.

Research

Like service delivery, the modern history of research also begins in the mid-19th century. In 1838, Esquirol differentiated intellectual disability from mental illness, and proposed several levels of impairment. Later in the century, various different classification systems were proposed. Probably the most well-known was a system of "ethnic classifications" proposed by J. Langdon Down in 1866—hence the term "Mongolism" for persons with Down syndrome, the type of intellectual disability he first identified.

In addition to classifying persons with intellectual disability, genetics plays an important historic role in intellectual disabilities research. Unfortunately, early in this century, that role was mainly negative. Spurred by the rediscovery of Mendel's genetic findings, various workers attempted to show that intellectual disability was inherited. Most notorious were Dugdale's (1877) study of the Jukes, and Goddard's (1913) study of the Kallikaks. In both cases, multi-generational families were used to argue that intellectual disability was inherited and that sterilization was necessary for the "preservation of the race." Such eugenic scares led to court decrees and sterilization laws in California and other states during the 1920s.

Also playing an important role in intellectual disability research was the new science of psychological testing. Henry Goddard, the research director at the Vineland Training School in New Jersey, was the first to use Binet and Simon's IQ tests in the United States. After testing Vineland residents over a several-year period, Goddard concluded that "the vast majority of feeble-minded children are not changing and are not improving in their intelligence levels," a finding that another leading worker, Walter Fernald, called "the most significant … and the most discouraging that we have ever known."

Since early in this century, research has progressed in both the behavioral and biomedical fields. Behaviorally, recent work has allowed for better diagnoses and classification. Nearly a century ago, psychologists invented tests of motor, nonverbal intelligence, achievement, adaptive behavior, and other skills. Following similar work with typically developing children, psychologists have recently learned much about the development of cognitive, linguistic, social, and adaptive skills in persons with intellectual disability. In addition, many studies now examine the presence of psychiatric disorders in children and adults with intellectual disability, and families, schools, group homes, and workshops have all received research attention. Much of this work has occurred from the 1960s until today, a time of strong, often federally supported, progress in intellectual disability behavioral research.

But arguably the most important progress has occurred in the

biomedical and genetic fields. As early as 1934, the Norwegian physician, Folling, hypothesized that several of his patients with retardation were unable to metabolize phenylalanine. Such insights later led to a dietary treatment of this inborn error of metabolism; the phenylalanine-free diet, though not perfect, has significantly limited PKU as a cause of intellectual disability in most industrialized countries. In addition, PKU can now be screened at birth through the Guthrie test, after which dietary treatments can be instituted. PKU remains one of the major success stories in intellectual disability even as new treatments for this condition, for example tetrahydrobiopterin, are still being developed.

Modern genetics constitutes another success story in intellectual disability research. In 1959, Lejeune, Gautier, and Turpin discovered that most cases of Down syndrome were due to trisomy of the 21st chromosome. Fragile X syndrome, the second most common genetic cause of intellectual disability, was discovered in 1969, and its cytogenetic diagnosis was refined in the mid-1970s. With the banding techniques of the 1970s and the molecular genetic techniques of the past two decades, exact causes are being described for many of the nearly 1,500 different genetic disorders associated with intellectual disability. The genetic—and, increasingly, the behavioral—characteristics of many of these disorders are now known (see below).

Organizations

The field of intellectual disability is also a history of organizations. The first such organization was the Association of Medical Officers of American Institutions for Idiotic and Feeble-Minded Persons, begun in 1876. As the first professional group devoted exclusively to persons with intellectual disability, the Association of Medical Officers was the forerunner to the American Association on Intellectual and Developmental Disabilities (AAIDD, formerly AAMR), the field's most influential organization. Through its various journals and committees, AAIDD has been at the forefront of research, policy, and legislative advances for persons with intellectual disability.

In addition to AAIDD, other groups have also played important roles, particularly concerning changes in federal policy toward people with disabilities. The Council for Exceptional Children (CEC) was begun in 1922, and continues to champion the profession of special education and the education of all children with disabilities. The Arc, formerly the National Association for Retarded Citizens (NARC) is the main parent-led advocacy group. Arc was particularly influential in the passage of the Education for All Handicapped Children Act of 1975 (PL 94–142), the federal law that for the first time mandated the "free, appropriate public education" of all children with disabilities throughout the United States. Each of these groups was also influential in the passage of the Americans with

Disabilities Act, federal legislation that took effect in 1992, and outlawed discrimination of persons with disabilities. The National Association for the Dually Diagnosed, now simply NADD, was established in 1983, specifically in recognition of the lack of attention by the professional community to the needs of individuals with intellectual disability who also experience psychiatric illness. Over the ensuing years the organization has grown to be the principal association in North America to advance mental wellness for persons with developmental disabilities through the promotion of excellence in mental health care.

Federal agencies constitute an additional category of influential organizations. Begun during the presidency of John F. Kennedy (himself a sibling to a woman with intellectual disability), the National Institutes of Child Health and Human Development (NICHD) have long supported disability research. Other federal programs, including the nationally distributed network of nearly 70 University Centers for Excellence in Developmental Disabilities Education, Research, and Service (UCEDDs), and NICHD-funded Mental Retardation Research Centers (MRRCs), have also played important roles in research and policy. Taken together, professional, parent, and federal organizations provide an important infrastructure to the intellectual disability field.

Before discussing definitional issues, it is worth revisiting terminology. In different countries throughout the world, different terms are used for intellectual disability. In Europe, the term "intellectual disabilities" has long been used, and a major international journal is the Journal of Intellectual Disabilities Research. In Great Britain, alongside "intellectual disability," the term "learning disability" (or learning disabilities) is often used—a term with a very different meaning in the United States.

DEFINITION OF INTELLECTUAL DISABILITY

Over the years, the American Psychiatric Association (APA) has adopted definitions of intellectual disability from the American Association on Intellectual and Developmental Disabilities (AAIDD), previously the American Association on Mental Retardation (AAMR). While the definitions in the third and fourth editions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM; American Psychiatric Association) are very similar, DSM-5 (American Psychiatric Association, 2013) includes substantive changes to the definition of intellectual disability. For the first time in DSM, for example, the term "intellectual disability" replaces "mental retardation," and the language of DSM-5 is more closely aligned with that of the AAIDD. (The term proposed for ICD-11 [2015] is "intellectual developmental disorders.") Further, severity levels are now classified by adaptive rather than intellectual function or IQ scores. This section reviews the changes (see Table 40–1 for summary).

Intellectual and Adaptive Functioning

Intellectual Functioning. Standardized intelligence tests that meet appropriate psychometric criteria for reliability and validity are commonly used to identify intellectual disability. These instruments are summarized in Table 40–2, along with each test's age range and cognitive domains.

Both in terms of the testing situation itself and in the choice of an appropriate measure, administering such tests to people with intellectual disability is quite challenging. As people with intellectual disability have increased risks of comorbid psychiatric or behavioral disorders, examiners need to ensure that difficulties such as hyperactivity or poor frustration tolerance do not impede optimal test performance. Even in persons without behavioral challenges, certain personality characteristics may interfere with testing. Many people with intellectual disability, for example, look to others for solutions to difficult problems, and are quick to acquiesce, or become easily discouraged by failure. To minimize these problems, examiners can begin with easier tasks to provide success and incentives or rewards for effort.

Table 40–1.
A Comparison of DSM-IV and DSM-5 Diagnostic Criteria for Intellectual Disability

DSM-IV Diagnostic Criteria		
A. Intellectual Function	**B. Adaptive Function**	**C. Onset**
Significantly subaverage intellectual functioning: an IQ of approximately 70 or below on an individually administered intelligence test	Concurrent deficits or impairment in present adaptive functioning (i.e., the person's effectiveness in meeting the standards expected for his or her age or cultural group) in at least two of the following skill areas: communication, self-care, home living, social/interpersonal skills, use of community resources, self-direction, functional academic skills, work, leisure, health, and safety	The onset is before age 18 years.
DSM-5 Diagnostic Criteria		
A. Intellectual Function	**B. Adaptive Function**	**C. Onset**
Deficits in intellectual function, such as reasoning, problem solving, planning, abstract thinking, judgment, academic learning, and learning from experience, confirmed by both clinical assessment and individualized, standardized intelligence testing	Deficits in adaptive functioning that result in failure to meet developmental and socio-cultural standards for personal independence and social responsibility. Without ongoing support, the adaptive deficits limit functioning in one or more activities of daily life, such as communication, social participation, and independent living across multiple environments, such as home, school, work, and community.	Onset of intellectual and adaptive deficits during the developmental period

Data from American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th

ed. Washington, DC: American Psychiatric Association; 1994; American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. Washington, DC: American Psychiatric Association; 2013.

For individuals who are nonverbal, there are several measures that can be selected that do not rely on expressive language (see Table 40–2). People with certain genetic syndromes (e.g., Williams syndrome, Down syndrome) also present unique test challenges, as many show syndrome-specific profiles of cognitive strength and weakness that are not easily described by a global IQ score (see phenotype section). Persons with intellectual disability and co-occurring sensory or motor deficits also require adaptation to routine test procedures. Finally, when testing individuals from some minority groups, examiners need to be sensitive to language and cultural issues, and use a multi-method test approach, especially when diagnosing persons with mild intellectual disability.

Adaptive Functioning. Edgar Doll (1935) was the first to develop a formal definition and measure of adaptive behavior. Two decades later, the AAMR (now AAIDD) officially included deficits in adaptive behavior in its definition of intellectual disability. Since that time, deficits in adaptive behavior have been formally included in all definitions of intellectual disability. However, until DSM-5, IQ score alone was used to classify individuals in terms of *severity* of intellectual disability. Table 40–3 illustrates the substantive nature of the changes in severity classification from DSM-IV to DSM-5.

Table 40–2.
Measures for Assessing Intellectual Ability

Measure	Age Range (years)	Domains Assessed
Wechsler Preschool and Primary Scale of Intelligence, 4th edition (WPPSI-IV) (Wechsler, 2012)	2.6 to 7–3	Verbal IQ, performance IQ, full scale IQ
Wechsler Intelligence Test for Children, 5th edition (WISC-V) (Wechsler, 2014)	6 to 16–11	Verbal IQ, performance IQ, full scale IQ
Wechsler Adult Intelligence Scale, 4th edition (WAIS-IV) (Wechsler, 2008)	16 to 90	Verbal IQ, performance IQ, full scale IQ
Stanford–Binet Intelligence Scales, 5th edition (Roid, 2003)	2 to 85+	Verbal, quantitative, abstract/visual, short-term memory, composite score
Kaufman Assessment Battery for Children, 2nd edition (K-ABC-II) (Kaufman & Kaufman, 2004)	3 to 18	Sequential and simultaneous processing, mental processing composite
Kaufman Adolescent and Adult Intelligence Test (KAIT) (Kaufman & Kaufman, 1993)	11 to 85+	Crystallized and fluid scales, composite IQ

Differential Ability Scales, 2nd edition (DAS-II) (Elliott, 2006)	2–6 to 17–11	Verbal, nonverbal reasoning, spatial abilities, general conceptual ability
Raven's Coloured Progressive Matrices (CPM) (Raven, 1998, 2003)	5 to elderly	Figural reasoning; group test
Test of Nonverbal Intelligence (TONI)-2 (Brown, Sherbenou, & Johnsen, 1990)	5–85	Abstract problem-solving: simple matching, analogies, classification, intersections, and progressions
Leiter International Performance Scale, Revised (Leiter-R) (Roid & Miller, 1997)	5–0 to 85–11	Reasoning ability, similarities, differences, relationships
Universal Nonverbal Intelligence Test (UNIT) (Bracken & McCallum, 1998)	5–0 to 7–11	Reasoning, memory, symbolic, and nonsymbolic tasks

Table 40–3.
DSM-IV Severity Criteria for Intellectual Disability (APA, 1994)

Severity Criteria				
Mild	**Moderate**	**Severe**	**Profound**	**Unspecified**
IQ level 50–55 to approximately 70	IQ level 35–40 to 50–55	IQ level 20–25 to 35–40	IQ level below 20–25	When there is a strong presumption of mental retardation but the person's intelligence is untestable by standard tests.

DSM-5 Severity Criteria for Intellectual Disability (APA, 2013)

Severity Criteria	Conceptual Domain	Social Domain	Practical Domain
Mild	For preschool children, there may be no obvious conceptual differences. For school-age children and adults, there are difficulties in learning academic skills involving reading, writing, arithmetic, time, or money, with support needed in one or more areas to meet age-related expectations. In adults, abstract thinking, executive function (i.e., planning, strategizing, priority setting, and cognitive flexibility), and short-term memory, as well as functional use of academic skills (e.g., reading, money	Compared with typically developing age-mates, the individual is immature in social interactions. For example, there may be difficulty in accurately perceiving peers' social cues. Communication, conversation, and language are more concrete or immature than expected for age. There may be difficulties regulating emotion and behavior in age-appropriate fashion; these difficulties are noticed by peers in social situations. There is limited understanding of risk in social situations; social judgment is immature for age, and the person is at risk of being	The individual may function age-appropriately in personal care. Individuals need some support with complex daily living tasks in comparison to peers. In adulthood, supports typically involve grocery shopping, transportation, home and childcare organizing, nutritious food preparation, and banking and money management. Recreational skills resemble those of age-mates, although judgment related to well-being and organization around recreation requires support. In adulthood, competitive employment is often seen in jobs that do not emphasize conceptual skills. Individuals generally need support to make health care–related and legal decisions, and to learn to perform a skilled vocation competently. Support is typically needed to raise a family.

	management), are impaired. There is a somewhat concrete approach to problems and solutions compared with age-mates.	manipulated by others (gullibility).	
Moderate	All through development, the individual's conceptual skills lag markedly behind those of peers. For preschoolers, language and pre-academic skills develop slowly. For school-age children, progress in reading, writing, mathematics, and understanding of time and money occurs slowly across the school years and is markedly limited compared with that of peers. For adults, academic skill development is typically at an elementary level, and support is required for all use of academic skills in work and personal life. Ongoing assistance on a daily basis is needed to complete conceptual tasks of day-to-day life, and others may take over these responsibilities fully for the individual.	The individual shows marked differences from peers in social and communicative behavior across development. Spoken language is typically a primary tool for social communication but is much less complex than that of peers. Capacity for relationships is evident in ties to family and friends, and the individual may have successful friendships and sometimes romantic relations in adulthood. However, individuals may not perceive or interpret social cues accurately. Social judgment and decision-making abilities are limited, and caretakers must assist with life decisions. Friendships with typically developing peers are often affected by communication or social limitations. Significant social and communicative support is needed in work settings for success.	The individual can care for personal needs involving eating, dressing, elimination, and hygiene as an adult, although an extended period of teaching and time is needed for the individual to become independent in these areas, and reminders may be needed. Similarly, participation in all household tasks can be achieved by adulthood, although extended teaching and ongoing supports are needed for adult-level performance. Independent employment in jobs that require limited conceptual and communication skills can be achieved, but considerable support from coworkers and others is needed to manage social expectations, job complexities, and responsibilities such as scheduling, transportation, health benefits, and money management. A variety of recreational skills can be developed but require supports and learning opportunities over an extended period. Maladaptive behavior is present in a significant minority and causes social problems.
Severe	Attainment of conceptual skills is limited. The individual generally has little understanding of written language or of concepts involving	Spoken language is quite limited in vocabulary and grammar. Speech may be single words or phrases and may be supplemented through augmentative means. Communication is	The individual requires support for all activities of daily living (meals, dressing, bathing, and elimination), requires supervision at all times, and cannot make responsible decisions regarding well-being of self or others. In adulthood,

	numbers, quantity, time, and money. Caretakers provide extensive supports for problem solving throughout life.	focused on the here and now. Language is used for social communication more than explication. Individuals understand simple speech/gestures. Relationships with familiar others are a source of pleasure and help.	tasks at home, recreation, and work require ongoing support/assistance. Skill acquisition in all domains involves long-term teaching and support. Maladaptive behavior, including self-injury, is present in a significant minority.
Profound	Conceptual skills generally involve the physical world rather than symbolic processes. The individual may use objects in goal-directed fashion for self-care, work, and recreation. Certain visuospatial skills, such as matching and sorting based on physical characteristics, may be acquired. However, motor and sensory impairments may prevent functional use of objects.	The individual has very limited understanding of symbolic communication in speech or gesture. He or she may understand some simple instructions or gestures. The individual expresses his or her own desires and emotions largely through nonverbal, nonsymbolic communication. The individual enjoys relationships with familiar others, and initiates and responds to social interactions through gestural and emotional cues. Sensory and physical impairments may prevent many social activities.	The individual is dependent on others for all aspects of daily physical care, health, and safety, although he/she may participate in some of these activities as well. Individuals without severe physical impairments may assist with daily work tasks at home, like carrying dishes to the table. Simple actions with objects may be the basis of participation in some vocational activities with high levels of support. Recreational activities may involve music, movies, walks, or participating in water activities, all wit h support. Physical and sensory impairments are frequent barriers to home, recreational, and vocational activities. Maladaptive behavior is present in a significant minority.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Table 40–4.
Common Measures of Adaptive Behavior

Measure	Age Range	Domains Tested
Vineland Adaptive Behavior Scales, 2nd edition (Sparrow et al., 2005)	Birth to 18–11	Communication: receptive, expressive, written; daily living skills: personal, domestic, community socialization: interpersonal relations, play & leisure, coping skills; motor skills: fine, gross
Adaptive Behavior Assessment System, 2nd edition (ABAS-II); (Harrison and Oakland, 2003)	0–89	Assesses conceptual, social, and practical aspects of adaptive behavior, as well as 10 adaptive skill areas.

Scales of Independent Behavior, Revised (SIB-R) (Bruininks, Woodcock, Weatherman, & Hill, 1996)	Birth to 80+	Motor: fine, gross
		Social interaction & communication: social interaction, language comprehension, language expression
		Personal living skills: eating & meal preparation, toileting, dressing, community living skills: time & punctuality, money & value, work
Diagnostic Adaptive Behavior Scale (DABS) (AAIDD release 2015)	4 to 21	Focuses on the critical "cut-off area" for the purpose of ruling in or out a diagnosis of ID or related disability.
Street Survival Skills Questionnaire (SSSQ) (Linkenhoker & McCarron, 1993)	9 to 40+	Administered to person; basic concepts, functional signs, tools, health

Though meanings vary, adaptive behavior is typically viewed as the performance of behaviors required for social and personal sufficiency. Further, adaptive behavior is an inherently developmental and social construct. Adaptive behavior changes as children grow into adolescence and adulthood, and demands for social adaptation are also defined by expectations from others—from one's family, society, and culture. Adaptive skills typically change across various settings; one's adaptive performance on the job or at school may differ from one's performance with friends or at home. Measurements of adaptive behavior, then, need to have a developmental orientation, to be socially and culturally sensitive, and to represent the many settings in which people live, work, and play.

In contrast to Doll's time, many instruments now exist that measure adaptive behavior across multiple domains (e.g., community skills, personal grooming). Table 40–4 summarizes commonly used measures of adaptive behavior, their age ranges, and various domains. Although most of these measures are administered as interviews with parents or other care providers, some measures are administered directly to persons with intellectual disability. Central to all measures of adaptive function is the idea that adaptive behavior is measured by performance as opposed to ability. If people with intellectual disability are able to (possess the skills to) perform certain behaviors, but do not regularly do so or do not do so without prompts, reminders, and assistance, then they necessarily have compromised adaptive functioning.

Severity of Intellectual Disability

There are remarkably variable strengths and competencies of people with intellectual disability. DSM-5 (and similar to AAIDD 2002) provides a classification system based on adaptive functioning across three domains: conceptual, social, and practical. Table 40–3 summarizes the classification

system in DSM-IV (by IQ score) and the DSM-5 system based on characteristics across the three domains and associated with mild, moderate, severe, and profound intellectual disability. As people with intellectual disability can also be classified by the etiology of their conditions, later sections detail how adaptive functioning changes across diverse genetic and other etiological groups.

Mild Intellectual Disability . Mild intellectual disability constitutes the largest group of people with intellectual disability, possibly as many as 85 percent of all persons with ID. These individuals appear similar to unaffected individuals and often blend into the general population in the years before and after formal schooling. Many achieve academic skills at the sixth grade level or higher, and some graduate from high school. As adults, many of these individuals hold jobs, marry, and raise families—yet at times they may appear slow or need extra help negotiating life's problems and tasks.

Robert was a full term infant, the last of three children born to his 38-year-old mother, a high school music teacher, and 40-year-old father, a science teacher. Pregnancy was unremarkable, and Robert's two older sisters were healthy and developing nicely. The family lived in a rural town in the Midwest.

Robert was an extremely fussy newborn, and had extended periods of crying that the pediatrician described as "colic." When Robert was 2 months old, his parents were told that he had supravalvular aortic stenosis, a case that warranted monitoring but no surgeries. Although Robert became less fussy over time, he was a picky eater, refusing solid foods. Robert's parents also noted that he was more "high-strung" than his siblings, often quick to cry or cringe when his sisters played too loudly.

Milestones were slightly delayed, with Robert sitting unassisted at 10 months and walking at 18 months. Language was also delayed, and although his first words appeared at 20 months, Robert had always made his wants and needs known. Although his parents were concerned that he was delayed compared to his sisters, they were reassured by the explanation that boys often had slight delays, and that he was a lively, social boy who would quickly catch up.

When Robert was 3, his parents insisted on a developmental evaluation, which revealed modest delays in cognitive, language, and motor functioning. His developmental quotient was 74 (average range is roughly 85 to 115). He was described as friendly and engaging, a real "charmer," with a cute face that endeared him to many. Robert was enrolled in a special kindergarten, and he remained in a combination of special education and mainstream (general education) classes throughout his academic career. Like the rest of his family, Robert enjoyed listening to music and singing, and he took an active interest in the piano.

At age 7 years, the school psychologist evaluated Robert and felt that he fit a "learning disability" profile. Robert had an overall IQ of 66, but with close to average functioning in short-term memory and expressive language, and pronounced deficits in visual-spatial functioning. He struggled with writing tasks and arithmetic, but loved science and music, and was amazingly conversant to anyone who would listen to him. Indeed, his parents feared he was "too friendly," as well as too active, and with transient, intense interests in unusual items such as vacuum cleaners.

As he entered adolescence, Robert became increasingly anxious, so much so that he occasionally rubbed his hands or rocked, and he "fretted" about day-to-day issues and what would happen next. His long-standing sensitivities to loud sounds seemed to wane slightly, but he developed fears of storm clouds and dogs, and refused to ride on elevators. He became tearful and upset after one of his older sisters left for college, and worried about her health and her ability to watch the weather at college. Although Robert suffered from nightmares, and

would occasionally pace with worry and complain of stomachaches, he attended school, and had a small group of friends in the Special Olympics bowling league. Further, he enjoyed singing with the high school chorus and he was delighted to be frequently selected to play the piano at school concerts.

When Robert was 17 years old, his parents happened to watch a television documentary on Williams syndrome. They were overwhelmed by the similarities between Robert and the people described in the program. They later described the experience like a "jolt." They had always accepted Robert, quirks and all, and had stopped pushing their doctors for reasons "why" when Robert was a preschooler. Nevertheless, they immediately called the informational number offered in the show, and within 2 months, they had the genetic tests done that confirmed their strong suspicion that Robert had Williams syndrome.

While Robert's day-to-day life did not change dramatically after his diagnosis, his parents reported a big difference in Robert's outlook. He met new friends with Williams syndrome at a conference; he applied to go to a summer music camp for young adults with Williams syndrome; and he reported that he felt less alone. Robert's parents reported a mixture of feelings at having such a late diagnosis—disappointment in their doctors, relief in finally knowing, and twinges of guilt. They were energized by having a new community of Williams syndrome families to share their feelings and worries with; like their son, they too felt less alone.

Elisabeth is a 29-year-old widowed Caucasian mother. Elisabeth was the youngest of three children and born to her parents after an unremarkable pregnancy and delivery. Early development was similarly without distinction as she met her developmental milestones as expected. However, in the context of her enrollment in a Head Start program, some cognitive delays were noted. More strikingly, it became clear that Elisabeth had some visual problems evidenced by her falling a lot and stumbling into things. An examination revealed that Elisabeth had significant visual problems prompting patching and other interventions. Elisabeth was routed into special education classes from kindergarten, and although she was described as somewhat of a loner, never really playing with any of her many cousins, behavioral concerns did not really surface until puberty. At that time, she became more withdrawn, into a "fantasy world." Her family recalls that imaginary friends took on new importance such that Elisabeth would carry on animated conversations with them. She would talk to posters on the wall—even kiss them on occasion. Elisabeth began to have increasing difficulties in school around this time as well. Her mother recalls that Elisabeth seemed always to have a level of discomfort with sexuality and, after menarche, Elisabeth would refuse to attend to personal hygiene. Elisabeth was also incapable of eating in the cafeteria with the rest of the students. Her sister recalls that Elisabeth would invariably "start eating like a pig," and because of these behaviors she would be placed in a separate lunchroom by herself.

At the age of 17, a psychological evaluation was done to determine the level of Elisabeth's intellectual ability. The narrative at that time described how Elisabeth had been coded as having mild intellectual disability but that there were other factors that were of concern, notably frequent absences from school and the need to build in extra incentives to get her to attend to tasks. Elisabeth was described in the evaluation at that time as being a cooperative test subject, appearing "healthy in all respects, but she was quiet, somewhat shy and seemed to have difficulty articulating," which was attributed to word-finding difficulties.

Following her graduation from high school Elisabeth found herself in and out of counseling, she became estranged from her family, and at 24 years of age presented to a mental health clinic where she was described as a young woman "with developmental disability who had recently given birth (2 months earlier) to a son." She was described as experiencing symptoms of depression related to the suicide of her husband approximately a year earlier. The story that Elisabeth told at that time was described as chaotic and confusing, including issues relating to her relationship with her parents, sisters and other individuals in her life.

With the birth of Raymond, Elisabeth's first child, the family recalls how they took Elisabeth back into the home to support her and the new grandchild and that they were very concerned about the behaviors they witnessed, including Elisabeth's absolute refusal to change diapers or to wash the baby's genitalia, claiming that such behavior would constitute sexual abuse. They observed her force-feeding the child, and inappropriately dressing him for the

weather. Ultimately, her father overheard a conversation from an adjacent room in which Elisabeth said that she was going to do the same thing that the child's father did (commit suicide) but that she would take Raymond with her. This prompted a call to the authorities and was the proximal event that led to the removal of Raymond from Elisabeth's custody.

Over the next several years, her family describes numerous and volatile interactions with Elisabeth that were consistently characterized by irritability on Elisabeth's part. The family is quite clear in being unable to identify discreet periods of time during which Elisabeth's behavior cycled up or down into specific depressive or manic episodes. The most prominent recollection is of stable and durable preoccupations with fantasy, misinterpretations and even paranoid accusations, recalling recently how she refused to let her brother turn on the lights because she knew that the CIA would be observing the family if the lights were on.

During a recent involuntary hospitalization, the psychiatrist noted that Elisabeth made sporadic eye contact, that her speech was slurred, and that she smiled or giggled inappropriately as though responding to some private thoughts or perhaps "voices." While she denied hallucinations, she was convinced that both her mother and sister intended to kill her if they got the chance, that both had murdered others or had people do the killing for them, and that they were conspiring to make sure that Elisabeth never got her son back. During this admission, Elisabeth was started on risperidone, her first medication trial, at 1 mg per day and subsequently the dose was increased to 3 mg. Elisabeth agreed that the new medicine helped her to think clearly and it was observed that the delusional thinking about her neighbors and the Mental Health Center gradually disappeared during the course of her hospitalization.

Currently, Elisabeth lives on her own in an apartment with close monitoring from community supports. Elisabeth has been cooperative in participating in various programs that have been offered including weekly swimming opportunities as well as participation in an art studio. Staff report that they have not been aware of any changes in Elisabeth's sleep or appetite, no significant changes in her weight, no changes in mood, and no psychotic symptoms.

Historically, people with mild intellectual disability were thought to show relatively few clear-cut organic causes for their delay. However, in recent years there has been an increase in the number of people with genetic syndromes who function in the mild range of intellectual disability. Examples include most people with Prader–Willi syndrome (PWS), as well as some males and most females with Fragile X syndrome.

A more striking finding is that more people with mild intellectual disability come from minority groups and low socio-economic (SES) backgrounds than would be expected from their percentages in the general population. This overrepresentation of minority groups has been used to criticize standardized intelligence tests that are used for placement in special education classroom settings, as well as to highlight the importance of environmental-cultural influences on intellectual development and assessment.

Moderate Intellectual Disability. Moderate intellectual disability is seen in approximately 10 percent of the population with ID, and includes people with more impaired cognitive and adaptive functioning. People with moderate intellectual disability are typically diagnosed in their preschool years, and some show a clear organic cause for their delay. Many persons with Down syndrome, the most common chromosomal cause of ID, often function in this range, as do many adolescents and adults with Fragile X

syndrome. Most children with moderate intellectual disability require special education services, achieving academic skills at the second to third grade level. Need for supportive services continues throughout the lifespan, and with proper supports, many live, work, and thrive in their local communities. In a study by Ross and colleagues, 20 percent of persons with IQs from 40 to 49 lived independently, whereas 60 percent were considered partially dependent and 20 percent totally dependent on others. In a similar way, some individuals in this range are employed in the competitive job market, and need minimal job supervision, whereas others require more extensive supervision on the job, and may work in sheltered workshops or other, more segregated settings.

Bryan is a 17-year-old sophomore with moderate intellectual disability associated with Down syndrome. Bryan had significant cardiac problems in infancy, and eventually required surgery to repair a ventricular septal defect. Although his energy level improved significantly thereafter, he still had some delays in motor and other developmental milestones. Nevertheless, throughout his education, Bryan has been placed in general education classrooms with same-aged peers and with the support of classroom aides. Bryan's IQ is 53, he has some difficulties with articulation, and his language mainly consists of two-to-four-word sentences. Recently, he has made progress in reading and can read 50 sight words. Bryan has had some minor problems with managing his anxiety over the years, and can become somewhat stubborn at times, but he is well-liked by his teachers and classmates, and his parents boast that Bryan has been a charming, pleasant child from infancy. In addition to attending classes at the local high school, Bryan also spends some time learning job skills in the community. He enjoys his position as the assistant manager for his high school's cross-country team and actively participates in Special Olympics.

Severe Intellectual Disability. Severe intellectual disability occurs in about 3 to 4 percent of the population of people with intellectual disability. Individuals at this level often have one or more organic causes for their delay, and many show concurrent motor, ambulatory, and neurological problems, as well as poorly developed communication skills. Most persons with severe intellectual disability require close supervision and specialized care throughout their lives. Some individuals learn to perform simple tasks or routines that facilitate their self-care or their ability to perform in a sheltered workshop or pre-workshop type of setting.

Harrison is a 16-year-old resident at a school for children with severe disabilities. His mother raised concerns early on when at the age of 2.5 Harrison lost acquired language skills. He was seen by a pediatric neurologist, who elected to introduce carbamazepine in the presence of an abnormal EEG. His mother recalls that the carbamazepine not only appeared to improve Harrison's overall function but also seemed to have a positive effect on Harrison's sleep and activity level, both of which were problematic. In fact, the hyperactivity and impulsivity later prompted a trial of methylphenidate and Harrison's mother remembers that while there may have been some initial improvement on this drug, over time Harrison became "an animal" on the medication, resulting in its discontinuation. Indeed, methylphenidate is the only medication that his mother lists when given the opportunity to identify drugs that have been tried in the past that were clearly bad.

On the other side of the equation, his mother lists carbamazepine and other anti-seizure medications as having been consistently helpful over the years. Staff at Harrison's school note

that carbamazepine is unique among the anticonvulsants that have been tried in that there seemed to be benefit with respect to reducing seizure frequency but also significant behavioral benefit. Over the years, the dose of carbamazepine has varied and has been periodically discontinued in favor of trials of newer anticonvulsants because of concerns about Harrison's borderline white blood cell counts. However, because of the unequivocal behavioral benefit that staff and family attribute to the carbamazepine, it is being used again, albeit at doses that may be lower than have been utilized in the past. His mother's recollection is that a number of medication trials have been explored over the years and that each of the medications has conferred only temporary benefit. In addition to carbamazepine and methylphenidate, previous medication trials have included melatonin, clonidine, sertraline, aripiprazole, risperidone, and diphenhydramine.

The behavioral difficulties that have prompted medication changes include agitation, hyperactivity, and mood instability in the form of giddiness, and more recently, inappropriate sexual advances towards female staff. There have also been significant aggressive outbursts in recent years. In addition to the periods of heightened irritability and hyperactivity, there have been periods of diminished activity, and loss of appetite such that Harrison's weight has occasionally dropped into the 120s in spite of his 6′3″ frame. Teachers and Harrison's mother report that during these times his interests also diminish, including his participation in activities that he otherwise enjoys. Observers note that Harrison seems to want to eat but just cannot muster the interest or enthusiasm when food is placed in front of him.

On examination Harrison comes quite willingly into the office setting accompanied by staff and family. He is a tall and slender young man dressed in sweat pants that are loose on his thin frame. Harrison extends his hand to acknowledge the presence of the examiner and does make eye contact on occasion. His psychomotor activity is remarkable for fairly constant stereotyped hand movements and occasionally some vocalizations. Harrison takes a piece of paper from his mother when he sees it, and proceeds to flick it in front of his eyes and to tear small strips from it and place them in his mouth. He readily surrendered these bits of paper upon request. Harrison's affect was full with very clear and bright smiles, especially when interacting with his mother. Later during the visit, he seemed to become more agitated and would occasionally place his hand to his mouth in a self-biting gesture or place his face into the palm of his hand while vocalizing and rocking. Harrison followed simple requests and was able to sign "thank you" after having been given some crackers from the staff who accompanied him.

When Harrison's mother was given a description of bipolar disorder she quite readily recalled all of those symptoms in her older brother, reporting, for example, that he had periods of elevated mood noteworthy for reckless and impulsive behaviors. At other times he would withdraw, become more aloof, and had expressed suicidal thoughts at various times.

Lithium was subsequently initiated, and resulted in a significant improvement in Harrison's irritability, sleep disturbance, and hyperactivity. He remains on a relatively low dose of carbamazepine for control of his epilepsy. His antipsychotic is gradually being tapered toward discontinuation.

Profound Intellectual Disability. Profound intellectual disability affects relatively few individuals (1 to 2 percent) with ID, and involves pervasive deficits in cognitive, motor, and communicative functioning. Impairments in sensory-motor functioning are often seen from early childhood on, and most individuals require extensive training to complete even the most rudimentary aspects of self-care such as eating and toileting. The vast majority of people with profound intellectual disability have identifiable organic causes for their delay, and most require total supervision and care throughout life.

Annalise was born following a pregnancy that was unremarkable—her mother's first. There were significant feeding difficulties in the first few weeks of life, and failure to thrive was considered early on. Annalise was quite hypotonic and her milestones were delayed. She did

not walk until approximately 4 years of age. While she was enrolled in an early intervention program, Annalise's progress was variable and she never used spoken language. She has and uses a few signs to make her needs known, and more recently has begun to use an assistive communication device. Her family reports that Annalise's medical history is extraordinarily complicated. Approximately 2 years ago, Annalise had a life-threatening episode of congestive heart failure requiring intubation and aggressive treatment that now includes a beta-blocker, a diuretic and digitalis.

At that time, it was discovered that Annalise had hypothyroidism, so she has also required thyroid supplementation. Approximately a year ago, the diagnosis of diabetes was made and Annalise is insulin dependent. Her family also reported significant problems with what appears to have been a milk allergy, and wonder whether that may have contributed to some of Annalise's difficulties in gaining weight in her early years. However, when Annalise was approximately 4 or 5 years of age, she began to become completely preoccupied with food, consuming everything in sight, with a particular preference for carbohydrates. Genetic testing for PWS was negative, but subsequent testing revealed a deletion on chromosome 1.

Annalise's history is remarkable for her never having been particularly attentive to the thoughts or feelings of others, literally walking over other children when she was able to do so. The family reports that Annalise has many circumscribed interests, including particular vignettes from Disney videos, which, if she were allowed, she would view repeatedly. She also has had a preoccupation with sorting and stacking objects, and at one point would tantrum if she weren't able to organize stones or pebbles at the local beach. When Annalise was 6 years of age, a methylphenidate trial was initiated. The dose was pushed up to 5 mg three times daily (she was unable to swallow a sustained release capsule intact) and her family reports that over the course of the year that Annalise received that drug, she was less emotionally available and seemed to be more ritual-bound and preoccupied with her limited interests. The family has not observed prominent stereotypic behaviors, merely the preoccupation with objects and with sorting and arranging. Because of Annalise's preoccupation with food, the family has had to lock their refrigerator and cabinets. Her family hopes that medication may help to allay some of Annalise's anxiety and obsessionality that continue to present significant barriers for her full participation in activities, particularly in the community.

These four levels—mild, moderate, severe, and profound—have long-been used to classify severity of intellectual disability, yet have come under attack because of their focus on deficits instead of competencies. How people with intellectual disability are defined and classified—based on strengths, impairments, or some combination of both—has profound implications for studies on the epidemiology of this condition.

PREVALENCE

According to DSM-5, 1 percent of the population has intellectual disability. Indeed, a meta-analysis of 52 peer-reviewed papers has estimated the global prevalence of ID to be 1,037 per 100,000 population. Rates are higher in males. A recent epidemiological study based on a national register in Sweden found that the prevalence of ID in older adults is increasing as a result of medical and social advances, and that approximately 0.5 percent of the population above 55 years of age has ID.

Etiology of Intellectual Disability

As noted by Esquirol in 1845, intellectual disability is not a disease in and of itself, but the developmental consequence of a variety of pathogenic

processes. Intellectual disability results from brain dysfunction, generally due to abnormal brain development or brain injury resulting from genetic and environmental causes (Table 40–5). Many genetic causes of intellectual disability have been known for decades, while modern genetic techniques have identified hundreds of new causes of intellectual disability.

Genetic causes of intellectual disability can result in multiple malformation conditions (e.g., Down syndrome) or what appears to be "non-syndromic" intellectual disability. That said, most genetic causes of intellectual disability are associated with characteristic clinical or neurobehavioral features, if the patients are examined thoroughly. Genetic disorders with intellectual disability display a wide variety of inheritance patterns and genetic mechanisms including chromosome rearrangements (often de novo), autosomal recessive, autosomal dominant (inherited and de novo), X-linked, triplet repeat expansion, and imprinting. The important role for de novo mutations has only recently been appreciated. Knowledge of these genetic mechanisms can aid with diagnosis and is essential for recurrence risk counseling. More and more, the genetic diagnosis influences treatments to implement (e.g., specific melatonin/propranolol treatment for sleep disruption in patients with Smith–Magenis syndrome) and treatments to avoid (e.g., valproate in ornithine transcarbamylase deficiency). In addition, some disorders are associated with health implications for family members (e.g., Fragile X–associated tremor/ataxia syndrome and premature ovarian failure in adult relatives of individuals with Fragile X).

With advances in medicine generally and in molecular genetics in particular, new causes of intellectual disability, or the genetic etiologies of formerly unspecified syndromes, are identified each year. For example, a national database of genetic disorders, *Online Mendelian Inheritance in Man,* lists >2,500 genetic conditions associated with intellectual disability. The most common causes of intellectual disability include Down syndrome, Fragile X syndrome, prematurity, and fetal alcohol syndrome. Variability in patient cohorts and methods for defining etiologies make it difficult to estimate the proportion of patients accounted for by each cause. For example, persons drawn from institutional settings, where severe and profound intellectual disability and visible congenital anomalies are more likely to be found, more commonly have an identifiable chromosomal abnormality than do persons with mild disability residing in community settings.

Table 40–5.
Etiological Classification of Intellectual Disability

Etiology	Examples	Estimated Frequency (%)

A. Prenatal Causes		
1. Genetic disorders		4–28
Chromosomal aberrations	Down syndrome	
Monogenic mutations	Tuberous sclerosis; phenylketonuria and other metabolic disorders; fragile X syndrome	
Multifactorial	"Familial" mental retardation	
Malformation syndromes due to microdeletions	Prader–Willi, Williams, Angelman syndromes	
2. Congenital malformations		7–17
Malformations of the central nervous system	Neural tube defects	
Multiple malformation syndromes	Cornelia de Lange syndrome	
3. Exposure		5–13
Maternal infections	Congenital rubella, HIV	
Teratogens	Fetal alcohol syndrome	
Toxemia/placental insufficiency	Prematurity	
Other	Radiation, trauma	
B. Perinatal Causes		2–10
1. Infections	Meningitis	
2. Delivery problems	Asphyxia	
3. Other	Hyperbilirubinemia	
B. Postnatal Causes		3–12
1. Infections	Encephalitis	
2. Toxins	Lead poisoning	
3. Other postnatal causes	Traumas, brain tumors	
4. Psychosocial problems	Poverty, psychotic illness	
D. Unknown Causes		30–50

Understanding the cause of intellectual disability in an individual person is essential for providing prognostic information, guiding treatment, and educating families about recurrence risk. For instance, distinguishing between a one-time injury such as hypoxia and a genetic neurodegenerative disorder such as a lysosomal storage disorder makes a huge difference in treatment, emotional impact on the family, how they plan their child's life, and how they approach having additional children. A specific diagnosis can also reduce family stress by providing an "answer" for why their child has a significant disability and allowing them to connect with other families affected by the same condition. More and more, an accurate diagnosis will guide more appropriate treatments to maximize neurocognitive function, prevent decline, and direct specific medical screenings or treatments that optimize the patient's health and adaptive function.

In the era of rapidly expanding molecular genetic diagnosis, each provider must decide how to manage the explosion in genetic information and the impact on diagnosis and treatment based on their level of interest, educational background, and access to other subspecialists. At a minimum, psychiatric providers need to know when to suspect an underlying genetic condition and how to ensure adequate evaluation and management (through reference materials and/or consultants). For instance, ornithine transcarbamylase deficiency results from a defect in the urea cycle that can cause encephalopathy due to elevated ammonia levels. Patients can present at any age, and the best way to improve behavioral issues in a patient with undiagnosed or inadequately treated ornithine transcarbamylase deficiency, is to recognize the possibility of the condition, diagnose it using biochemical (plasma ammonia level) and genetic testing (*OTC* gene sequencing), and provide dietary and medication treatment to lower ammonia levels and maintain nutrition. Valproate, haloperidol and other medications should also be avoided, since they have been associated with hyperammonemic crises in these patients. Indeed, the decompensation associated with administration of these medications in an undiagnosed patient can prompt workup and accurate diagnosis of ornithine transcarbamylase deficiency by the vigilant psychiatrist. Genetic diagnosis is now directly influencing treatment choices in a variety of diseases, most prominently cancer, and it is only a matter of time before gene-specific treatments become widespread in psychiatry.

Medical Diagnostic Approach to the Patient with Intellectual Disability

Clinical History and Examination. The elements of a comprehensive clinical evaluation for the patient with intellectual disability are summarized in Table 40–6. As for any patient, a good history is essential, and in many cases collateral information is extremely useful. Family members may be able to fill in important aspects of the family pedigree, providing information about relatives ranging from learning problems to physical anomalies, developmental disabilities, and even response to treatment.

The clinical examination should be performed with particular attention to the presence of minor anomalies, aberrant growth, and physical development (Table 40–6). Increasingly, the examination will look for behavioral symptoms that may comport with a known behavioral phenotype (see later discussion) and help in diagnosis.

Genetic Evaluation. Given the incredible advances in DNA sequencing technologies and other methods for identifying mutations across the entire human genome, clinical genetic testing in patients with intellectual disability is very much a moving target. Currently, patients without an obvious syndrome or extrinsic cause should be evaluated for chromosomal

abnormalities and Fragile X (by chromosomal microarray and Fragile X repeat length testing respectively), since these are the most common causes of "nonsyndromic" intellectual disability. Genetic disorders should still be considered in patients who do not follow the expected course for a presumed environmental cause. Unless a specific aneuploidy (such as Trisomy 21) or a balanced translocation is suspected, a karyotype is no longer the appropriate test when chromosome microarray is available. Beyond these basic tests, expert clinical genetics consultation may be the most cost-effective next step, rather than extensive laboratory testing and imaging. Genetic consultation can help guide diagnostic testing as well as providing counseling about prognosis, surveillance for complications, and recurrence risk.

Table 40–6.
Elements of the Clinical Evaluation of the Patient with Intellectual Disability

Clinical History
Prenatal exposures (alcohol, infections, etc.)
Birth history
CNS infection, traumatic brain injury
Developmental regression
Focal neurological findings, hypertonia, hypotonia, or ataxia
Extra-CNS malformations and medical issues
Visual impairment or hearing loss
Family Pedigree (Three-Generation)
Relatives with learning problems, psychiatric disorders, intellectual disability, abnormal brain imaging findings, neurological or degenerative disorders
Consanguinity
Family resemblance (reduced between family members with and without some genetic conditions)
Physical Examination
Assessment of major and minor physical anomalies including facial features and skin findings
Growth and physical development
Head circumference compared to norms (congenital versus "acquired" microcephaly/macrocephaly)
Growth trajectory (comparison with earlier measures)
Description of facial features (micrognathia, hypertelorism, thin upper lip)
Complete neurological examination
Skin examination for café-au-lait and other birthmarks
Documentation of behavioral phenotype
Adjunct Diagnostics
Hearing evaluation
Ophthalmology evaluation of eye structure, eye movements, and visual function
Psychometric assessments
Laboratory Testing (When Indicated)
Creatine kinase (to evaluate for muscular dystrophies)
7-dehydrocholesterol testing if Smith–Lemli–Opitz syndrome is suspected
Metabolic studies for lysosomal, peroxisomal, and mitochondrial disorders
Urine organic acids, plasma amino acids
Congenital disorders of glycosylation testing
Genetic Testing (When Indicated)
Chromosomal microarray

Fragile X testing
DNA sequencing (single gene, gene panel, whole exome)
DNA methylation testing if Prader–Willi or Angelman syndromes are suspected
Imaging
Structural brain magnetic resonance imaging [MRI]
Functional brain MRI [fMRI] (still largely research-based)
Diffusion tensor brain MRI [DTI]
Other Studies
Muscle biopsy for suspected mitochondrial disease (or muscular dystrophy)

When a specific disorder is suspected, a single gene test may be the best option; however, as genetic testing becomes more widely available, the spectrum of presentations associated with each gene has broadened, so it can be difficult to determine which single gene test to perform. Fortuitously, DNA sequencing has become much more efficient, so panels of genes, or even most of the coding regions in the entire genome (the "exome"), can now be sequenced in a single test. As the cost of sequencing decreased and human sequence variation becomes better characterized, genome-wide sequencing methods will become first-line genetic tests, despite the challenges with interpreting the significance of many DNA variants. It is important to remember that currently available sequencing tests do not detect repeat expansions or imprinting disorders.

Neuroimaging. Neuroimaging is clinically indicated for individuals with intellectual disability in the presence of microcephaly or macrocephaly, abnormal head circumference, seizures, focal neurological signs, and/or progressive cognitive impairment or loss of psychomotor skills. MRI is generally superior to CT unless acute hemorrhage or calcification are suspected. MRI has been useful in elucidating a diagnosis in around 5 percent of people with non-progressive intellectual disability. More commonly, in ~50 percent of patients, subtle and nonspecific structural findings indicating altered neurodevelopment are present such as ventriculomegaly, dysplasia of the corpus callosum, minor abnormalities of white matter signal, cortical gyration, and cerebellar vermis. The American Academies of Neurology and Pediatrics and the American College of Medical Genetics periodically publish updated practice guidelines for the use of brain imaging in the evaluation of patients with intellectual disability. Magnetic resonance spectroscopy is lower-yield and less commonly employed, although it can yield a diagnosis when a metabolic disorder is suspected despite otherwise normal genetic and metabolic work-up. Diffusion tensor imaging holds promise for identifying disorders associated with defects in major axon pathways that do not affect brain morphology; however, it is not yet in wide clinical use. To date, large studies utilizing structural MRI or quantitative methods such as those performed in ADHD or autism have not been performed in intellectual disability, and little has been done in the way of functional MRI (fMRI)

research. Advanced imaging techniques may serve to better elucidate the neurobiology of intellectual disability, resulting in improved diagnostic precision and treatment monitoring in the future.

Electroencephalography. Electroencephalographic investigations, such as waking/sleep video-EEG-polygraphic studies may be useful in patients with a clinical history of seizures and those with severe language deficits and/or specific genetic syndromes such as Angelman syndrome. EEGs are often requested in patients with known cortical malformations associated with high seizure risk (e.g., polymicrogyria) when acute behavioral or cognitive changes have occurred.

Metabolic Testing

EVALUATION FOR INBORN ERRORS OF METABOLISM. The Inborn Errors of Metabolism (IEMs) or biochemical genetic conditions are a large and fast growing group of disorders that are for the most part individually quite rare, but collectively affect 1 to 1.5 percent of the population. These disorders originate from disruptions to genes involved in biochemical and metabolic pathways, many of which are important in brain development and/or function. Generally, symptoms and findings arise from either the accumulation of substances that prove toxic in excess or the inability to synthesize enough of a required product. Well-known IEMs causing intellectual disability include phenylketonuria (PKU), Lesch–Nyhan syndrome, and the Lysosomal and Glycogen Storage diseases. It is important for the practicing clinician to retain the IEMs in the differential diagnosis, since many require specialized treatment to prevent morbidity, progression, and mortality, as well as to optimize development. Frequently, treatment involves supplementation with a deficient enzyme or metabolic intermediate with or without dietary restriction. It should be noted that each IEM may have variant ages of onset such as “infantile,” “juvenile,” or “adult” forms, with concomitantly different mixes of medical and neuropsychiatric symptoms. Later onset forms may manifest in less medically severe findings and a presentation more dominated by neuropsychiatric symptoms. Specific features in each disorder vary, but the child psychiatrist should consider the IEMs in the differential diagnosis of every individual with intellectual disability.

While guidelines published by the American Academies of Pediatrics and Neurology do not recommend routine metabolic testing for all persons with intellectual disability, several studies suggest that IEMs may account for up to 2.5 percent of patients after screening by clinicians. More recent studies have shown that even after broad-based metabolic screening, a second line of testing identified IEMs in 14 percent of patients, many of which were treatable. A careful history should be taken to make sure the

patient received usual newborn screening tests (particularly in the case of immigrants or where the patient lacked usual perinatal care) and whether the testing followed current recommendations (especially in older individuals). IEMs should be particularly considered in the patient with intellectual disability who presents with catatonia, mania, seizures, encephalopathy, and/or visual hallucinations, when symptoms appear exacerbated by physiological stress and illness, occur with abdominal pain or vomiting, or functional regression (Table 40–7).

The authors recommend the following screening labs for patients with intellectual disability when an IEM is suspected: complete blood count, complete metabolic panel (electrolytes, BUN, creatinine, glucose, transaminases), lactate, pyruvate, urine organic acids, plasma amino acids, ammonia, plasma acylcarnitines, and total and free carnitine. Testing during a symptom exacerbation is recommended, since assays may be normal during periods of relative metabolic stability. Referral to a specialist biochemical geneticist is also recommended for evaluation and/or management.

GENETIC INTELLECTUAL DISABILITY SYNDROMES AND BEHAVIORAL PHENOTYPES

Remarkable genetic advances have sparked renewed research interest in the behavioral features of persons with fragile X, Prader–Willi, and other genetic syndromes. In their reviews in the 1970s of behavioral aspects of chromosomal disorders, S. Kessler and R. Moos, for example, discussed only five syndromes. Since that time, the number of disorders for which a "behavioral phenotype" has been suggested has grown considerably. A good working definition of a "behavioral phenotype" is that proposed by Dykens, namely, that of a syndrome in which the probability of the expression of certain behaviors or constellations of behaviors is greater than what would otherwise be expected. Stereotyped hand movements, for example, have been described in many contexts, but a particular form of stereotyped hand wringing is almost invariably seen in Rett syndrome. Hyperphagia and obesity are nonspecific symptoms, but without aggressive dietary intervention, the probability of an individual with PWS becoming obese is nearly 100 percent.

Table 40–7.
Selected Clinical Findings and Laboratory Abnormalities Which Increase Suspicion for Underlying Metabolic Disorder

Growth abnormality
Recurrent, unexplained illness
Exacerbation of symptoms by physiological stress
Seizures
Ataxia

Loss of psychomotor skills
Hypotonia
"Coarse" appearance
Eye abnormalities (cataracts, ophthalmoplegia, corneal clouding, retinal abnormality)
Recurrent somnolence/coma
Abnormal sexual differentiation
Arachnodactyly
Hepatosplenomegaly
Metabolic/lactic acidosis
Hyperuricemia
Hyperammonemia
Low cholesterol
Structural hair abnormalities
Unexplained deafness
Bone abnormalities (dysostosis; occipital horns; punctate calcifications)
Skin abnormalities (angiokeratoma; "orange-peel" skin, ichthyosis)
Brain imaging abnormalities cortical and/or cerebellar atrophy, evidence of stroke, abnormal white matter signal, lactate peak on MRS

Adapted from Curry CJ, Stevenson RE, Aughton D, Byrne J, Carey JC. Evaluation of mental retardation: recommendations of a consensus conference. *Am J Med Genet*. 1997;72:468.

Table 40–8provides a representative sample of syndromes for which behavioral phenotypes have been described—some anecdotally, others with controlled studies. Although the study of behavioral phenotypes is only at its beginning stages, this line of work holds much promise for both clinicians and researchers. Identifying syndrome-specific behavioral patterns helps clinicians to zero in on some diagnostic possibilities over others, and these behavioral profiles also guide intervention and treatment. From a research standpoint, phenotypic work advances the understanding of gene–brain–behavior relationships. To illustrate these advantages, the following subsections review behavioral and genetic features for fragile X, Prader–Willi, Down, and Williams syndromes. Although not representative of all syndromes, these examples amply demonstrate the promise of phenotypic research for both intervention and science.

Examples of Behavioral Phenotypes in Intellectual Disability Syndromes

Fragile X Syndrome. Fragile X syndrome, the most common inherited cause of intellectual disability, results in a wide range of learning and behavioral problems, with males being more often and more severely affected than females. The fragile X gene (*FMR1*) represents a relatively newly identified type of human disease, caused by an amplification (or excessive repetition) of trinucleotide sequences (CGG) that make up DNA. Greater than a certain threshold number of these triplet repeats (about 200), people are fully affected with the syndrome. Numbers of repeats between the normal threshold of 50 and less than 200 are termed "premutations." As many as 1 in 259 women in the general population may carry the premutation, with 1 in 4,000 males and 1 in 8,000 females being

fully affected with the syndrome.

Table 40–8.
Phenotypes of Selected Conditions Associated with Intellectual Disability

As genetic testing has become more widespread, the known clinical presentation for each genetic disorder has expanded beyond the classic presentations described in the early literature. More and more, disorders are being grouped by their shared molecular pathophysiology, e.g., Cornelia de Lange syndrome as a cohesinopathy (defects in the cohesion DNA-binding complex involved in DNA repair, chromosome segregation, and gene regulation) and Neurofibromatosis 1 as a RASopathy (defects in the RAS/MAPK network of genes involved in cell cycle regulation and differentiation).

Disorder	Cause	Inheritance	Prevalence[a]	Recurrence Risk	Testing
Chromosomal					
Cri-du-chat syndrome	5p15.2 deletion	Dominant deletion (usually de novo)	1/20,000–50,000	Very low	Chromoson microarr
Down syndrome	Trisomy 21	Sporadic	1/2,500–1/32 births based on maternal age	Low, unless due to translocation	Karyotype
Smith–Magenis syndrome	17p11.2 deletion	Dominant deletion (usually de novo)	1/25,000	Very low	Chromoson microarr

22q11 Deletion syndrome (Velocardiofacial/ DiGeorge)	22q11 deletion	Dominant deletion (de novo or inherited)	1/4,000	Low if de novo, up to 50% if parent affected	Chromoson microarr
Williams syndrome	7q11–23 deletion	Dominant deletion (usually de novo)	1/7,500	Low if de novo, up to 50% if parent affected	Chromoson microarr
Single gene					
Fragile X syndrome	*FMR1* trinucleotide repeat expansion	X-linked with anticipation	1/4,000 males, 1/8,000 females	Up to 50%	FXS testing length
Rett syndrome	Altered gene expression due to loss of *MECP2* function	X-linked, 99% de novo	1/8,500 females (rare in males)	Low	*MECP2* gen sequenci

Cornelia de Lange syndrome (Cohesinopathy)	*NIPBL, SMC1A, SMC3, RAD21, HDAC8*		1/40,000–100,000		Panel seque exome
Tuberous sclerosis complex 1 and 2	Loss of *TSC1* or *TSC2* function	Autosomal dominant	1/10,000		*TSC1*/*TSC2* sequenci
Neurofibromatosis type[a] (RASopathy)	Loss of *NF1* function	Autosomal dominant (de novo or inherited)	1/2,500–4,000	Low if de novo, 50% if parent affected	*NF1* sequer
Imprinting					
Prader–Willi syndrome (PWS)	Lack of paternal-specific imprinting of 15q11.2-q13	Imprinting (90% sporadic)	1/10,000–30,000	Low, except with paternal mutation	PWS/AS me testing

Angelman syndrome (AS)	Loss of *UBE3A* function via mutation or lack of maternal-specific imprinting of 15q11.2-q13	Imprinting (90% sporadic)	1/12,000–24,000	Low, except with maternal mutation	PWS/AS me testing a sequenci
Metabolic					
Smith–Lemli–Opitz	Loss of *DHCR7* function, defect in cholesterol metabolism	Autosomal recessive	1/20,000–40,000	25%	7-dehydroc level, *DH* sequenci
Phenylketonuria (prototypical biochemical genetic disorder)	Accumulation of phenylalanine	Autosomal recessive	1/10,000	25%	Newborn s plasma a
Mitochondrial	Defects in electron transport chain	Autosomal recessive,	1/5,000	Up to 25%	Lactate, py plasma a

		autosomal dominant, X-linked or maternal			acids, MI biopsy
Lesch–Nyhan syndrome	Loss of *HPRT1* function, accumulation of uric acid	X-linked	1/380,000	50% of sons	Urine urate 24-hr uri *HPRT1* se
Ornithine Transcarbamylase Deficiency	Loss of *OTC* function and ammonia accumulation	X-linked	1/14,000–77,000	Carrier mothers: 50% of sons; Affected fathers: 50% of daughters are carriers	Ammonia le sequenci
Lysosomal storage diseases (diverse group of >50 disorders)	Accumulation of various macromolecules	Autosomal recessive, X-linked	1/5,000–10,000	25%	Enzymatic a sequenci
Toxin					
Fetal alcohol syndrome	Alcohol neurotoxicity	NA	1/3,000 in western countries, 1/300 with fetal alcohol effects	NA	Clinical diag

[a]Estimated prevalence in mixed European populations.

ADHD, attention-deficit/hyperactivity disorder; ID, Intellectual Disability; IQ Intelligence Quotient.

Many of the behavioral features in people with fragile X syndrome are a function of the number of CGG repeats and a cascade of subsequent effects. The number of repeats relates to the degree that the *FMR1* gene is methylated and subsequently silenced, or blocked from transcription and

translation. This, in turn, affects the amount of *FMR1* protein (FMRP) produced, and the range of clinical expression in fragile X syndrome is associated with the amount of FMRP. Fully affected, fully methylated males have no FMRP, whereas high-functioning males, although rare, have some FMRP. In affected females, only one X chromosome carries the *FMR1* gene, whereas the other produces some FMRP. Because all females have one X chromosome that is randomly inactivated in all cells, the degree of involvement in females with fragile X is related to the ratio of normally active X chromosomes to the total active plus inactive X chromosomes. About half of fully affected females thus show mild to moderate intellectual disability, whereas half have average IQ but show learning difficulties and/or social affective impairments. Both males and females with the full syndrome exhibit similar neurocognitive profiles, including deficits in auditory short-term memory, processing of sequential information, and attention, with relative strengths in processing simultaneous information, long-term memory, and distinguishing between one's own and others' representations of the world.

Although fragile X syndrome is a genetic disorder, there is also evidence for environmental contributors to IQ and other specific neurocognitive deficits, including factors relating to the home environment (enrichment opportunities and socioeconomic status). All of these factors taken together —gender, variation in the *FMR1* repeats, and environment—contribute to the significant within-syndrome variability found in fragile X syndrome.

Fragile X Syndrome and Autism. It has been reported that many fully affected males with fragile X syndrome show autistic-like symptoms such as language delay, echolalia, stereotypies, self-injurious behavior, perseveration, poor eye contact, and tactile defensiveness. It is not surprising that early workers wondered whether fragile X syndrome would reveal a common genetic cause of autism.

This flurry of research faded as new studies emerged suggesting that instead of autism per se, many males with fragile X syndrome showed a willingness to interact with others coupled with social and performance anxiety and mutual gaze aversion. Controlled studies and meta-analyses now suggest that only about 5 percent of males with fragile X syndrome have autistic disorder. Instead, the majority of affected males can be placed on a spectrum of social anxiety, shyness, avoidance, and gaze aversion. Although some of these boys have anxiety disorders or autism features, others may simply show slow-to-warm temperament styles, including shyness or social withdrawal.

In addition to these difficulties, hyperarousal, hyperactivity, and attention deficits are seen in the vast majority of boys and girls with fragile X syndrome. Attention-deficit/hyperactivity disorder (ADHD) symptoms are higher among affected boys relative to control subjects, and girls with

fragile X syndrome have lower prevalence rates of ADHD than boys with fragile X syndrome. Among adults, problems in attending and in sustaining effort have been found in the neuropsychological profiles of women with expanded *FMR1* repeats, which may contribute to problems in math, abstract reasoning, and planning.

Fragile X syndrome thus involves vulnerabilities toward shyness, gaze aversion, social anxiety, avoidant disorders, schizotypal disorder, ADHD, PDD-NOS, and, more rarely, autistic disorder. These difficulties vary in severity but are typically found in persons across the IQ spectrum, from those with moderate intellectual disability to those with mild learning disabilities.

Individuals with premutation repeat lengths (typically family members of affected individuals) are also at risk for psychiatric and medical issues including social dysfunction, primarily shyness, gaze aversion, and social anxiety and the adult-onset fragile X–associated tremor/ataxia syndrome (FXTAS). FXTAS is characterized by progressive intention tremor and gait ataxia, with associated dementia and global brain atrophy resulting in significant morbidity and mortality.

Prader–Willi Syndrome. First identified in 1956, PWS affects about 1 in 15,000 births and is best known for its food-related characteristics. Whereas babies with PWS invariably show hypotonia and pronounced feeding-sucking difficulties, young children between 2 and 6 years of age develop hyperphagia and food-seeking behavior such as food foraging and hoarding. Hyperphagia is likely associated with a hypothalamic abnormality resulting in a lack of satiety. Food preoccupations are lifelong, and without prolonged dietary management, affected individuals invariably become obese. Indeed, complications of obesity remain the leading cause of death in this syndrome.

PWS is the first known human disease to show the effects of genomic imprinting, in which genes are modified and expressed differently depending on whether they are inherited from the mother or the father. About 70 percent of PWS cases are caused by a paternally derived deletion on the long arm of chromosome 15. Remaining cases are attributed to maternal uniparental disomy (UPD) of chromosome 15, in which both members of the chromosome 15 pair come from the mother. In either case, there is absence of the paternally derived contribution to this specific region of the genome. When missing information in this same region of chromosome 15 is maternally derived, it results in a completely different and more severe developmental disorder, Angelman syndrome.

Although people with PWS invariably obsess about food, a remarkably high proportion of persons also show nonfood obsessions and compulsive behaviors. These nonfood symptoms include skin picking, hoarding, needing to tell, ask, or say things, and having concerns with symmetry,

exactness, ordering, arranging, cleanliness, and sameness in daily routine. Often these symptoms are associated with distress or adaptive impairment, suggesting marked risks of obsessive-compulsive disorder in this population. Indeed, obsessive-compulsive disorder is many more times likely in people with PWS than in the general population of people with intellectual disability.

In addition, even as compared to others with intellectual disability, children and adults with PWS show high rates of temper tantrums, aggression, stubbornness, underactivity, excessive daytime sleepiness, and emotional lability. Coupled with food seeking, these impulsive behaviors often lead people with PWS to need more restrictive levels of care than would be predicted by their mild levels of intellectual disability.

Psychopathology and cognitive profiles may be associated with an individual's genetic variant of PWS. Thus, research has shown behavioral differences between PWS due to paternal deletion and maternal UPD. Cases with deletions may show lower IQ, especially verbal IQ, and more frequent or severe problem behaviors, such as skin picking, hoarding, temper tantrums, overeating, and social withdrawal. Although a dampening of symptom severity is suggested in many UPD cases, there are occasional cases of more severe problems in UPD, primarily autistic-like features and relatively low IQ. In addition, persons with UPD may do better than their counterparts with deletions on verbal tasks but more poorly on visual-spatial tasks, including solving jigsaw puzzles. Indeed, although many persons with PWS are remarkably skilled in solving puzzles, this unusual strength seems to be limited to those with paternal deletions. The mechanisms that explain such differences in maladaptive behavior and cognition across genetic subtypes are unknown.

Many people with PWS, then, are at increased risk for obsessive-compulsive, impulse control, and affective disorders. Yet even those who do not meet diagnostic criteria for these psychiatric disorders often show significant maladaptive behaviors that interfere with optimal adaptive functioning.

Down Syndrome. Occurring in approximately 1.2 in 1,000 live births, Down syndrome (DS) is the most common chromosomal abnormality leading to intellectual disability. It is most often the result of nondisjunction of chromosome 21. The syndrome has been the subject of considerable investigation.

More than 80 clinical features characterize the phenotype of DS, including cognitive impairments, hypotonia, facial dysmorphisms, congenital heart disease, and other anomalies. Studies of children with DS have shown distinctive cognitive and linguistic profiles. Notable strengths have been highlighted in visual versus auditory processing. Language impairment can be extensive, with particular difficulty in expressive

language, grammar, and pronunciation.

In contrast to the relative overall weakness in the cognitive arena, persons with DS are often described as being particularly socially adept. However, studies of adaptive functioning suggest that these skills are not uniform, with communication scores trailing behind measures of daily living skills and socialization. The stereotype of a "Down personality" as being happy, good-tempered, affectionate, placid, and stubborn has been difficult to verify.

Compared to others with intellectual disability, persons with DS appear to suffer less often and less seriously from psychopathology. Adults with DS appear to be somewhat less prone to psychiatric disturbance than controls. This trend seems to extend to children and adolescents as well, in whom rates of psychiatric and behavioral problems are greater than in the general population but appreciably less than in other groups with intellectual disability. Commonly noted problems include difficulties with attention, impulsivity, hyperactivity, and aggression. In contrast to these problems, depression seems to be less common among children and adolescents than expected norms, although some youth appear to become increasingly withdrawn in their adolescent years. For many, this shift toward increased withdrawal seems gradual and subtle, and it may be an early harbinger of a later onset of depression or dementia in the adult years. Autism and pervasive developmental disorders appear to be relatively rare.

Williams Syndrome. First identified in 1961, Williams syndrome is caused by a microdeletion on chromosome 7 that includes the gene for elastin, a protein that provides strength and elasticity to certain tissues such as the heart, skin, blood vessels, and lungs. The condition affects 1 in 7,500 individuals, and persons with Williams syndrome often show hyperacusis, hypercalcemia, neuromusculoskeletal and renal abnormalities, and characteristic facial features described as elfin-like, cute, and appealing. People with Williams syndrome also typically show cardiovascular disease, especially supravalvular aortic stenosis, and these problems are likely associated with elastin insufficiency.

Apart from these medical issues, recent attention has focused on the interesting, possibly unique, cognitive-linguistic profile shown by many persons with this syndrome. Children with Williams syndrome show relative strengths in verbal short-term memory and language. Indeed, for many years it was thought that these children might even perform at chronological-age levels on a variety of linguistic tasks. Although age-appropriate performance in language has now been found in only small percentages of children with Williams syndrome, these children nevertheless show relative strengths in language, as well as in auditory processing and in some areas of music.

In contrast, many children with Williams syndrome perform especially

poorly on a variety of visuospatial tasks. Mervis and her colleagues found that 47 of 50 children with Williams syndrome showed the pattern of weakness in visuospatial tasks and a relative strength in various measures of language (vocabulary, grammar, and short-term auditory memory). Although not every child with Williams syndrome shows linguistic strengths and visuospatial weaknesses, the large majority present with such a pattern.

Studies have yet to fully examine the personality or psychiatric features of people with Williams syndrome. Early descriptions of people with Williams syndrome hinted at a "classic" Williams syndrome personality, described as pleasant, unusually friendly, affectionate, loquacious, engaging, and interpersonally sensitive and charming. Such qualities may change over the course of development, with adults being more withdrawn and less overly friendly than children.

Use of the Reiss Personality Profiles has shown that, relative to controls, adolescents and adults with Williams syndrome are more likely to initiate interactions with others (87 percent of sample), to enjoy social activities (83 percent), to be kind-spirited (100 percent) and caring (94 percent), and to empathize with others' positive feelings (75 percent) or when others are in pain (87 percent). At the same time, however, these subjects did not fare well in making or keeping friends and were often dangerously indiscriminate in their relating to others. Indeed, although sociability in Williams syndrome has generally been viewed as a strength, these features also seem to reflect the type of social disinhibition that is characteristic of people who are anxious, impulsive, and overly aroused. It is not surprising that salient problems in Williams syndrome include hyperactivity, inattentiveness, and a proneness for ADHD, which may diminish with age.

Recently, attention has begun to focus on anxieties and fears in this population. Generalized anxiety, worry, and perseverative thinking are commonly seen in Williams syndrome, and people with the syndrome appear to show unusually high levels of fears and phobias. Relative to suitably matched controls, fears in persons with Williams syndrome are more frequent, wider ranging, and more severe, and they may also be associated with impaired social-adaptive adjustment. In one study, Dykens compared fears in 120 people with Williams syndrome (age 6 to 48 years) to those with intellectual disability of mixed etiologies. Only two fears—getting a vaccination shot and going to the dentist—were mentioned by more than 50 percent of the group with mixed causes of intellectual disability. In contrast, 50 different fears were mentioned by more than 60 percent of subjects with Williams syndromes. Such fears ran the gamut. Some involved interpersonal issues like being teased, getting punished, or getting into arguments with others. Others involved such physical issues as getting injections, being in a fire or getting burned, or getting stung by a bee. Still others related to these children's hyperacusis or clumsiness (loud

noises such as sirens; falling from high places; thunderstorms). Thus, although not every person with Williams syndrome shows any or all of these fears, the vast majority do appear to be overly fearful compared to most others with intellectual disability.

Smith–Magenis Syndrome

Bryan was a 19-year-old young man with moderate intellectual disability and a history of psychiatric issues including impulsivity, aggressive behaviors, and a striking history of self-injurious behaviors such as self-hitting, skin-picking and pulling out fingernails and toenails. He had been given a diagnosis of autism. Bryan had restricted language abilities but got his basic needs met using a small set of phrases and gestures, though his speech was hard to understand and muffled. He lived at home among a devoted family including his parents and elder brother, and participated in a transitional school program, where he was segregated in a classroom for children deemed "lower functioning" with "behavioral needs." Bryan and his parents visited clinic with the aim of finding a better medication regimen that might be able to support his adaptive functioning, articulating their "dream" that their son might one day live independently or participate more easily in social activities.

At evaluation, Bryan's medication regimen consisted of extended-release dexmethylphenidate, trazodone, buspirone, risperidone, and sertraline, with as-needed lorazepam and additional risperidone for agitation and aggression. These medications were not fully controlling his psychiatric symptoms. A careful history was taken which elicited a history of infantile feeding difficulties, hypotonia and failure to thrive, and a striking pattern of lifelong sleep disturbance. Bryan fell asleep successfully, but woke every night around 2 to 3 a.m., did not go back to sleep, and was sleepy in the afternoons. He was noted to have somewhat heavy and coarse features with brachycephaly and prognathism. Accordingly, first line genetic testing for intellectual disability consisting of chromosome array and Fragile X was sent, revealing a 3.5-Mb interstitial deletion of chromosome 17p11.2, consistent with a diagnosis of Smith–Magenis Syndrome (SMS). SMS is a genetic syndrome with a prevalence of 1:15,000–1:25,000, with disruptions to the *RAI1* gene (located within the deletion) accounting for the majority of features. A number of important studies support the presence of distinctive disruption to melatonergic control of the sleep cycle in SMS and the use of melatonin and beta-blockers to normalize sleep, with resultant improvement in behavioral and cognitive function. Bryan was started on melatonin 3 mg in the evening and propranolol 10 mg in the morning. Over 2 months, these were increased to final doses of 5 mg and 20 mg respectively. His overnight sleep period increased from 4 hours to 8 hours and by the conclusion of the titration, his self-injury was reduced by 90 percent and aggressive behaviors stopped. He began to exhibit new learning, such as starting to search for and complete crossword puzzles, and he displayed an increased interest in social interaction and improved attentiveness. The authors were able to eliminate the use of risperidone, trazodone and all as-needed medications. Importantly, standard-of-care recommended medical screenings for SMS were completed, revealing that Bryan was deaf. He received bilateral hearing aids and speech therapy and began to speak more. A year later, Bryan's adaptive function was significantly improved and he was able to transition to a supported group living situation with two male peers, and began work at a part time job.

Smith–Lemli–Opitz Syndrome

Brianna was a 13-year-old young woman carrying a diagnosis of intellectual disability and Smith–Lemli–Opitz syndrome (SLOS). SLOS is a congenital syndrome characterized by growth retardation, microcephaly and multiple malformations including a distinctive facial appearance. Most affected individuals have moderate to severe intellectual disability but some have normal development and it is increasingly thought that the disorder and carriers may be much more common than previously understood. In SLOS, mutations in the *DHCR7* gene

result in derangement of cholesterol metabolism with elevated serum 7-dehydrocholesterol (7DHC) and low serum cholesterol. Affected individuals almost always display a variety of structural midline and para-midline brain findings, most commonly abnormalities of the corpus callosum. Sterol levels display strong correlation with MRI brain findings and the degree of somatic malformations, supporting a link between altered cholesterol/sterol metabolism, deficient brain myelination, and changes in neurotransmitter function that may underlie the neuropsychiatric phenotype.

SLOS is therefore a specific etiology resulting in intellectual disability, as well as a disorder that may be simultaneously conceptualized as being a biochemical genetic disorder, multiple congenital anomaly syndrome, and/or brain malformation. As with many metabolic disorders, it is thought that dietary supplementation of the deficient substrate, in this case cholesterol, may support ongoing brain development. Brianna received cholesterol supplementation since diagnosis and ongoing dietary optimization, as well as support for learning via an Individualized Educational Plan at school. She received low-dose trazodone to support sleep fragmentation; however, as she entered her teenage years shortly after puberty, Brianna began to exhibit increasing challenges with cognitive control and behavioral dysregulation, becoming much more aggressive towards caregivers and peers and with self-injurious behaviors. Simultaneously, she talked of visual and auditory hallucinations, particularly centered around seeing, hearing, and interacting with a person called "Cassie." These phenomena waxed and waned somewhat on a multiweek or monthly basis, but generally periods of increased aggressive behaviors coincided with what were interpreted as the perceptual or psychotic phenomena.

Olanzapine was started to address aggressive behaviors, self-injury, and presumed psychosis and was titrated to 10 mg twice per day. Brianna's aggressive behaviors declined somewhat but the perceptual phenomena appeared unresponsive and she gained some 40 pounds. With concern for treatment resistance and weight gain, she was switched to haloperidol 5 mg twice per day. With the use of this agent Brianna became more irritable with continuing hallucinations. At this point the treating child psychiatrist sought consultation with a developmental neuropsychiatrist and biochemical geneticist at a specialist center. Laboratory studies revealed that 7DHC levels were elevated. In accordance with best practice guidelines for SLOS, both trazodone and haloperidol, which are contraindicated in this disorder due to their link with 7DHC elevation, were stopped. Brianna was started on valproic acid, which proved effective in reducing her behaviors and addressing psychosis. Melatonin was prescribed for sleep initiation difficulties with good effect.

DEVELOPMENTAL CONSIDERATIONS IN CHILDREN WITH INTELLECTUAL DISABILITY

Throughout the 20th century, children with intellectual disability have been an ongoing topic of developmental research. Small groups of researchers have examined development of various Piagetian cognitive domains, language abilities, social skills, and even morality, but besides providing information about basic milestones or natural histories, these studies aid in understanding the underlying processes of development. Over the past decade, researchers have begun to specify the ordering of development (sequences), the ways in which levels in certain areas relate to levels in other areas (cross-domain relations), and the points at which the speed of development (rate) changes. Such analyses are now being applied to children with different etiologies of intellectual disability.

In addition, developmental analyses have extended beyond studies of children themselves. Many developmental psychologists examine mother–child interactions, family behaviors in families of different-aged children, and the child's interactions with peers, neighbors, and schools. This

section, then, examines both child-related and environment-related aspects of development in children with intellectual disability.

Child-Related Aspects of Development

Overview. Children with intellectual disability traverse the same sequences in their development as do children with typical development, but in the setting of certain forms of intellectual disability there are strengths and weaknesses not generally found in typically developing children. Children with intellectual disability also show certain ages and tasks of particularly slowed development, which seem related to specific types of intellectual disability, although the reasons for such slowings are the subject of ongoing research.

When examining child–environment interactions, it is found that mothers act "the same and different" compared to mothers of typically developing children at similar levels of functioning. Mothers of children with intellectual disability are similar in that they shorten their speech, focus on key words, and in other ways structure the environment for their child's development. They seem, however, to be much more didactic in their interactions. Families of children with intellectual disability also differ widely one from another. When the child with intellectual disability shows less maladaptive behavior and when the parent's style of coping is more problem-focused, the result seems to be lessened levels of parent and family stress.

Sequences. Typically developing children show a specific, possibly universal, ordering to their development. For example, in Piagetian cognitive development, children proceed from sensorimotor, to preoperational, to concrete operational, to formal operational thought. Even within these four larger stages, smaller orderings hold: thus, within sensorimotor development, normal infants proceed in order through Piaget's six sub-stages in each of several sub-domains.

Do children with intellectual disability also follow a "similar sequence" to their developments? For almost 30 years, similar sequences have been found for a variety of children over many tasks: in almost twenty Piagetian domains, in symbolic play, and in linguistic grammar and pragmatics. Similar sequences even hold for those children with genetic or other organic causes for their intellectual disability. The only possible exceptions include some children with uncontrollable seizures (where accurate testing is difficult), and some children with autism, who may show different orderings due to their particular disabilities on certain social tasks. For almost all children and for a wide variety of behaviors, development for children with intellectual disability proceeds along similar sequences to those found in typically developing children.

Cross-Domain Relations. If children with intellectual disability proceed in the same developmental ordering as do typically developing children, might they also show the even or flat profiles that groups of typically developing children show from one domain to another? Such a "similar structure" to development was proposed by Zigler in the late 1960s.

Contrary to similar sequences—which hold for all children with intellectual disability—cross-domain structures may differ based on which group one considers. With a few exceptions, children with cultural-familial intellectual disability do show even or near-even performance across various intellectual domains.

In contrast, children with different organic forms of intellectual disability show specific intellectual strengths and weaknesses. For example, three separate groups show weaknesses in sequential (or bit-by-bit, serial) processing compared to simultaneous (i.e., gestalt, holistic) processing. Such sequential deficits are found in boys with Fragile X syndrome, children with PWS, and children with Smith–Magenis syndrome. Children with Down syndrome may have particular weaknesses in expressive versus receptive communication, and specific problems in grammatical abilities. Children with Williams syndrome show extra deficits in visuospatial processing skills, and some subset of these children show heightened abilities in language. Different etiologies of intellectual disability, then, show different characteristic intellectual strengths and weaknesses.

Developmental Rates. A further issue concerns children's rates of development. By definition, children with intellectual disability develop at slower rates than typically developing children. As noted in the Prevalence section, non-specific or mixed groups show reasonably steady rates of development, as shown by stable IQ after the early childhood years. Children with different types of intellectual disability, however, may have periods of accelerated or slowed development.

These periods of accelerated or slowed development seem to be related to either age or task difficulties. For example, boys with Fragile X syndrome show stable developmental rates until the age of about 9 or 10, at which time development slows. Similarly in Down syndrome, both cross-sectional and small-scale longitudinal studies show that these children slow over the 6 to 11 year period. Throughout this age-span, children with Down syndrome slow in their development of linguistic grammar, intellectual skills, and adaptive behavior. In both Fragile X and Down syndromes, it remains unknown why such age-related slowings occur, whether all children with these disorders show such slowings, or if slowings are seen across all domains.

A second type of slowing relates to pronounced difficulties mastering a specific developmental task. These can occur at a variety of ages, whenever children with a particular type of disability reach a specific developmental

task. Consider young children with Down syndrome (who also show age-related slowings from 6 to 11 years). These children show difficulties in acquiring certain tasks of infant intelligence, even after accounting for their already-slower rates of development. Later, toddlers show increasing delays in expressive as opposed to receptive language skills. As a result of both age- and task-related slowings, then, children with Down syndrome generally show their highest IQ scores during the first year of life, after which IQ declines.

Contextual Aspects of Development

Mother–Child Interactions. Following research on interactions between typically developing children and their mothers, mother–child interaction studies in intellectual disability began during the mid-1970s. This research has shown that mothers of children with intellectual disability are both similar and different in their interactions than mothers of typically developing children of the same mental age (MA). Specifically, the two groups of mothers behave similarly on what might be called the structural aspects of behavior. In language, for example, mothers shorten their speech, simplify their vocabularies, raise their pitch, repeat key words, and generally provide the "motherese" typical of mothers of typically developing children. As a result, mothers of typically developing 2-year-olds act almost identically to mothers of children with intellectual disability who are at 2 years MA.

However, compared to mothers of typically developing, MA-matched children, mothers of children with intellectual disability are oftentimes more didactic. They more often initiate the topic of interaction, take longer speaking turns, talk at the same time as their child, and in other ways control the interaction. As a result, the child is allowed fewer opportunities to initiate and control the conversation.

It remains unclear why such stylistic differences occur. The prevailing view holds that mothers of children with intellectual disability are worried about their child's development: in their zeal to teach and instruct their children, these mothers might inadvertently be using a style of interaction that is not the most effective. Although only a few studies exist to date, young children with intellectual disability seem to develop faster when their mothers are less intrusive. One study examined mother-child interactions at 23 months and child outcomes at 36 months. Compared to mothers who were constantly redirecting their child's attention, mothers who simply commented on the object at which the child was looking had higher receptive language ages at 36 months. These studies suggest that the most effective style for mothers might be one that responds to the child's interests and initiatives.

Family Reactions. Over the past two decades, family studies in

intellectual disability have changed enormously, from a predominantly negative to a more balanced perspective. Before the early 1980s, families of children were perceived to be families in crisis. Mothers were examined for their "mourning" reactions, couples for divorce, and mothers, fathers, and unaffected siblings for the presence of depression and other forms of psychopathology.

Since the early 1980s, however, family researchers began to change how they conceptualized families of children with intellectual disability. From an earlier focus on family pathology, these families came to be thought of as families facing increased stress. Children with intellectual disability might therefore add stress to the family system, but now researchers understood that this stress could result in negative or positive adaptations. This changing view implied a more normative view of these families: just as every family faces a variety of stressors in life (e.g., illness, loss of job, etc.), so too do families of children with intellectual disability.

The move from a negative to a stress-coping perspective also has other implications. First, it allows for a change from a predominantly "between-group" research strategy. Previously, most studies compared families of children with disabilities to families of typically developing, nondisabled children of the same chronological age. Recently, however, the between-group focus has been complemented by within-group studies. Given that families of children with intellectual disability vary in their ability to cope, what causes better adjustment in one family and worse adjustment in another? What child, parent, or family characteristics foster better adaptation?

To date, most work has examined child characteristics such as the child's age and degree of impairment. For both variables, findings have been inconsistent. Some studies find it more stressful to parent older children with intellectual disability, and some suggest that families may experience more stress when the child begins puberty (11 to 15 years) and again when the child reaches early adulthood (20 to 21 years). However, other studies do not find such relations between increased family stress and the child's age. Similarly, some studies find that it is more difficult to parent children with more severe levels of disability, whereas other studies do not.

Another characteristic that has been examined is the child's degree of maladaptive behavior. So far, strong connections have been found between child maladaptive behavior and family stress in families of children with mixed etiologies, with PWS, and with 5p- (cri-du-chat) syndrome. Less clear is why such connections occur. Although most researchers feel that child maladaptive behavior leads to increases in familial stress, increased familial stress may also elicit more child behavior problems.

In studies of parent coping style, parents who more actively, constructively attempt to manage their child do better than those parents

who adopt a "palliative" coping style, one that either dwells on or ignores parental emotions. This difference in coping style may help buffer parents and families from the increased stresses of parenting a child with intellectual disability.

Psychopathology in Intellectual Disability

Developmental disability is a significant risk factor for psychopathology in general, and, as may be inferred from the description of causes above, this increased risk may derive from biological vulnerabilities as well as risks that accrue from the environment. Moreover, individuals with IQs below 70 have a two- to fivefold higher rate of psychiatric disorders overall compared with typically developing persons. A host of explanations, highlighted in Table 40–9, have been put forward in efforts to account for this added risk. Included are developmental experiences with which these individuals must contend, including perceived rejection from peers and frustration from parents. The repeated experiences of failure or difficulty achieving what appears to come naturally to typically developing peers, the recapitulation of experiences of failure with each developmental stage, and less well-developed or perhaps less supportive peer groups all take a toll on ego development. In addition to specific biological vulnerabilities, the common comorbidity of intellectual disability with physical illness, for example, epilepsy, may also increase the risk of mental illness.

Moreover, the treatments for epilepsy and other medical conditions may carry some "behavioral toxicity" which can increase the likelihood of diagnosed mental illness. Phenobarbital was widely reported to increase the risk of motoric hyperactivity and disinhibition in children and in individuals with developmental disorders, and phenytoin may cause "cognitive toxicity." While neither of these agents is used as frequently today, similar concerns may obtain for essentially any of the medications that are used for the management of epilepsy or other chronic medical conditions.

Table 40–9.
Possible Contributions to Increased Vulnerability to Mental Disorders in Persons with Intellectual Disability

Neuropathological process(es) responsible for intellectual disability may also cause or increase the risk for mental illness.
There may be an increased likelihood of loss and separation, particularly in out-of-home placements.
Communication deficits may predispose to emotional/behavioral disturbance.
There is vulnerability to exploitation and abuse by others.
There may be inadequate coping skills.
Family stress may be heightened in the presence of a child with developmental disability.
There is a risk of a limited network of social relationships and repertoire of social skills.
There is a risk of reduced opportunities for the development and exercise of recreational and occupational skills.

There is a possible adverse effect of the disability on self-esteem, and possible dysmorphology.

In general, individuals with developmental disabilities may experience the full range of mental illnesses as occur in the general population. A meta-analysis of epidemiological studies of psychiatric disorders in individuals with ID revealed higher rates of psychiatric disorders in children with vs. without ID, and higher rates in children and adults with more severe forms of intellectual disability vs. milder forms. However, evidence of a greater prevalence of psychiatric disorders in adults with *mild* disability as compared to adults in the general population is less clear. Other studies have shown that, in children, there is evidence that the developmental *course* of psychopathology is similar from ages 6 to 18 for children with and without ID. For instance, in general, internalizing behaviors increase with age and externalizing behaviors decrease with age. Specific risk factors for psychiatric disorders in children with intellectual impairments include both family and child characteristics: moderate vs. mild levels of intellectual impairment, adaptive behavior deficits, poor socialization, language impairments, living with one biological parent, and low family SES.

In 2007, the National Association for the Dually Diagnosed (intellectual disability and mental illness), in collaboration with a number of other professional organizations, published a manual to assist clinicians in psychiatric diagnosis with this population. The DM-ID provides a number of explicit examples of adapted diagnostic criteria for each diagnosis and is currently under revision. Additional specific psychiatric disorders as they appear in individuals with ID are reviewed below.

Attention-Deficit/Hyperactivity Disorder. The rates of ADHD in intellectual disability are estimated to between 9 and 18 percent. Although considerable research has been done on the treatment of ADHD, virtually all of these studies have specifically excluded children with ID. Moreover, the diagnosis of ADHD is based upon developmental considerations, namely, hyperactivity, impulsivity, and inattention significantly greater than that expected for a given developmental age, and thus the threshold for diagnosis in persons with severe to profound ID should be elevated.

For persons with intellectual disability, the diagnosis of ADHD is qualified as being excessive for an individual's *mental age*. In the context of profound intellectual disability, attention span, distractibility, or "on-task" behavior are predictably quite variable. Individuals given the diagnosis of ADHD in this context should, in comparison to their peers with similar levels of intellectual impairment, exhibit unusually short attention span, excessive psychomotor activity level, remarkable impulsivity, and so on. In many cases, the clinician will encounter a situation in which an individual does not evidence remarkable psychomotor activity or attentional

difficulties, but may be unusually impulsive. In these situations one should entertain the diagnosis of an impulse control disorder. Such a diagnosis, for example, might be appropriate for an individual who inexplicably strikes out at a peer in the absence of any identifiable environmental stressor.

Impulse Control and Stereotypic Movement Disorders: Self-Injury and Aggression. Every effort should be made to determine and treat the underlying cause of self-injurious behavior (SIB) or aggression, not merely to suppress the behavior. However, in some cases, for example, stereotypic movement disorder with SIB, the symptom is essentially the diagnosis. Aggression and self-injurious behavior are common in ID and increase with greater severity of cognitive impairment. Self-injurious behavior typically takes the form of chronic, repetitive, and frequently stereotyped behavior causing trauma. It occurs in the context of specific genetic syndromes, for example, self-biting in Lesch–Nyhan Syndrome and fingernail pulling in Smith–Magenis Syndrome, but also and more commonly in persons with unknown or nonspecific causes of intellectual disability. DSM-5 added specifiers to the diagnosis of stereotypic movement disorder for these situations.

In a review of psychiatric consultations in an institutional setting, it was reported that SIB was specifically cited as a reason for referral in 36 percent of the 251 cases examined. However, SIB as a presenting symptom is generally unhelpful in predicting the ultimate psychiatric diagnosis.

As SIB and aggression are nonspecific symptoms, one must take into account the presence or absence of a variety of factors to arrive at a presumptive diagnosis. Among these factors are the chronicity of the behavior, what function it serves, whether it is invariant in its topography (e.g., hitting only the right ear suggesting an ear infection), whether it is situational, whether it occurs in concert with regression from a previous level of function, and whether there are associated neurovegetative signs that correlate with its onset.

Oppositional Defiant Disorder/Conduct Disorder . The DSM-5 diagnosis of oppositional defiant disorder or conduct disorder also requires comparisons to others of similar mental age. Further, both diagnoses assume some degree of deliberateness on the part of patients (e.g., disobedience motivated by spite or resentment), which can be very difficult to discern in nonverbal subjects with profound cognitive deficits. As such, there is a lower limit in terms of severity of ID below which the reliability and utility of these diagnoses becomes suspect and an unspecified disruptive or impulse control disorder is likely to be utilized for diagnostic purposes.

Anxiety Disorders. Specific anxiety disorders, for example, separation

anxiety, obsessive-compulsive disorder, panic disorder, generalized anxiety disorder, and so on, rely heavily on an individual's ability to describe the subjective symptoms of anxiety. Patients who are clearly avoidant, who exhibit autonomic arousal in the face of stimuli that most of their peers would not find aversive, who present with other features of anxiety but who cannot articulate their subjective states, might be given a diagnosis of unspecified anxiety disorder.

In some cases in which individuals engage in behavior that appears compulsive or driven, and for whom such behavior seems ego-alien, the diagnosis of unspecified obsessive-compulsive or related disorder might be considered. Oftentimes these patients engage in self-restraint (securing their extremities in their clothing) or cling to their parents or care providers, seemingly to prevent self-injurious behaviors. Simple repetitive behaviors or the insistence on sameness that can accompany autism spectrum disorders should not be given a separate OCD diagnosis. However, it may be useful to organize treatment with an additional diagnosis beyond "autism" if, for example, the focus of treatment is a sleep disturbance, impulsive aggression, anxiety, or the like.

Although common, anxiety disorders appear to be underdiagnosed in persons with ID. The wide variability in prevalence rates, from 1 to 25 percent, is attributed to difficulty in making a diagnosis. Moreover, an individual with ID may not be able to identify subjective anxiety as an underlying cause of distress, and a patient's aggression or agitation may be suggestive of a disorder of impulse control rather than reflective of underlying anxiety. Indeed, common symptoms of anxiety in the population with intellectual disability include aggression, agitation, compulsive or repetitive behaviors, self-injury, and insomnia. Panic may be expressed as agitation, screaming, crying or clinging, which may even appear as delusional or paranoid behavior. Phobias may also occur in this population and may even be more common in persons with developmental disabilities. Experienced clinicians have noted that persons with developmental disabilities are at high risk for abuse, thus placing this population at a greater risk for posttraumatic stress disorder (PTSD). With data suggesting that it may be seen in nearly 8 percent of the general population, PTSD is an important diagnosis to consider in individuals with intellectual disability.

Eating Disorders. Because of the reliance of diagnostic criteria upon subjective experiences, the diagnoses of anorexia nervosa and bulimia are effectively precluded for individuals with severe or profound intellectual disability. For example, one would be hard pressed to identify in nonverbal patients such things as classic distortions in body image, or guilt feelings associated with bingeing. Food refusal or self-induced vomiting would have to be viewed as atypical eating disorders if such symptoms were to occur in

the absence of other diagnosable disorders (e.g., depression, rumination, etc.). Pica is perhaps the most common eating disorder among persons with intellectual disability.

Organic Mental Disorders. One could argue that by definition everyone with intellectual disability has some "organic" cerebral dysfunction, and thus any psychiatric illness should be regarded as organic or due to a general medical condition. In his classic study of psychiatric illness in a sample of institutionalized patients with Down syndrome in the 1970s, Menolascino appropriately argued that psychiatric nosology did not have to be reinvented to accommodate individuals with a "tissue diagnosis." Moreover, he reasoned that patients with so called "dual diagnoses" of mental illness and Down syndrome need to be distinguished from those with Down syndrome alone. Thus, the application of the diagnoses of organic mental syndromes and disorders is best approached as if patients do not have intellectual disability. The same principle should apply to personality disorders. The diagnosis of "organic" personality disorder is probably best reserved for individuals whose pre-existing personality was altered in a pathological way by some additional cerebral insult. In essence, this category has more clinical utility if reserved for patients whose intellectual disability is acquired after a period of typical development, usually secondary to trauma experienced in childhood or early adolescence.

Psychosis. Patients with developmental disorders are at an increased risk for schizophrenia (approximately 1 percent prevalence rate in persons with ID), bipolar disorder and other mental illnesses that may include symptoms of thought disorder and hallucinations. The diagnosis of schizophrenia requires that a patient relate the experience of delusions or hallucinations. Thus, in individuals with profound ID with limited communicative ability, the diagnosis of classic schizophrenia is arguably impossible. Nonetheless, some individuals display presumptive evidence of response to hallucinations (e.g., striking or shouting at empty space, throwing imaginary peers from furniture) or adopt catatonic postures that can appear to be psychotic in origin. In these cases the diagnosis of psychosis should be considered if these signs exist in the absence of sufficient evidence of a supervening mood disorder.

Mood Disorders. Even in profound intellectual disability, the diagnosis of mood disorders is fairly straightforward. Generally, a change in mood from baseline is obvious (recent onset lability, tearfulness, mood elevation, irritability). If coupled with changes in interests, activity level, sleep, appetite, or sexual behavior, of sufficient duration and causing sufficient impairment in adaptive function, the diagnoses of mania or of depression can be made in nonverbal patients.

Mood disorders are not uncommon in persons with ID, although it is unclear from studies to date whether or not mood disorders present at higher rates than in the general population. As noted above, learning problems, social skills deficits and low self-esteem are often associated with developmental disabilities and represent risk factors for the development of mood disorders. There are no striking differences between the expression of mood disorders in persons functioning in the mild and moderate range of ID and their typically developing peers. Differences may emerge among persons with severe to profound disability, but potential equivalents of mood disturbance may be recognizable, including irritability, crying, problems in sleep or appetite regulation, agitation, mood lability, social withdrawal, and isolation. Aggression or self-injurious behavior may also be seen as behavioral manifestations of dysphoria in persons regardless of developmental level.

Other Disorders. The diagnosis of Tourette syndrome is difficult in persons with profound ID. Frequently, these individuals display stereotyped or other movements and it can be challenging to discriminate intentional from unintentional movements or sounds, or to distinguish vocal tics from spontaneous, stereotyped, or echolalic vocalizations in individuals frequently incapable of functional speech. The diagnosis of stereotyped movement disorder might also be considered in such circumstances.

Since elimination disorders require a mental age of 4 or 5 years in order to be considered, the diagnoses of functional encopresis or enuresis are seldom made in the context of severe ID. In some instances, individuals appear to lose previously acquired skills, for example, urinary continence, but such losses typically do not occur in isolation, suggesting alternate diagnoses (e.g., delirium, depression, etc.).

Somatoform disorders, depersonalization disorders, and sexual disorders are less frequently diagnosed in the context of ID, though certainly not precluded. Sleep disorders ultimately require the subjective input of the patient regarding the adequacy of rest, occurrence of nightmares, and so on. Given the frequent history of abuse reported for people with ID as a group, one should not overlook the possibility of PTSD when sleep disturbance is a presenting problem.

Provisional Diagnoses. In many cases individuals with ID may not clearly fall into a single diagnostic category. Comorbidity is common. Additionally, some individuals may have psychiatric symptoms that significantly interfere with habilitative function, but which do not allow for a clear distinction between certain diagnoses. It may be very difficult to distinguish between an impulse control disorder, perhaps characterized by an individual who engages in impulsive aggressive acts, versus an anxiety

disorder, suggested by an individual who strikes out in the context of a stressor that would go unnoticed by most people. The clinician should always make a best effort at a working diagnosis, and be prepared to make modifications as indicated by data gathered through collateral sources and from increasing familiarity with a particular patient.

Approach to Maladaptive Behavior

As with psychiatry in general, there is little specificity that can be attached to a given symptom. Persons with ID will typically be referred for evaluation because of self-injurious, aggressive, impulsive, or hyperactive behavior. Because of the lack of diagnostic specificity for each of these symptoms, a diagnostic decision tree cannot be constructed. Instead, it is perhaps more useful to ask a series of questions about the expression of a particular behavior. If the behavior is of recent onset, one is more likely to consider an acute medical or psychiatric etiology. If the behavior is highly situational, occurring primarily in the context of the stress of task demands, the likelihood of a psychosis or mood disorder is probably reduced. If attempts are made to avoid the behavior by self-restraint, the inference of some ego-dystonic features may be tenable. Assessing the sum of these and collateral data will lead the clinician to a presumptive diagnosis that will form the basis for a treatment plan.

Diagnostic Instruments, Rating Scales

A number of diagnostic and behavioral rating scales have been developed for the population with intellectual disability. While neither diagnostic nor treatment decisions should be based entirely upon the results of a given instrument, the systematic collection of data through such scales is a very useful adjunct to the clinical diagnostic or monitoring process. Among the most common general behavior rating scales are the Aberrant Behavior Checklist (ABC) and the Developmental Behavior Checklist (DBC). The Behavior Problem Inventory (BPI) is particularly useful for self-injurious, aggressive, and stereotyped behaviors. The Diagnostic Assessment of the Severely Handicapped (DASH), the Reiss Screen for Maladaptive Behavior, and more recently the Reiss Scale for Children's Dual Diagnosis are all instruments whose primary utility is to poll for the presence of mental disorders.

TREATMENT

Prevention

Preventive treatment strategies largely focus on specific medical interventions to reduce the prevalence of intellectual disability, and upon mitigating associated complications, for example, treating associated mental disorders. The merits of primary prevention are obvious, and the

successes enjoyed with PKU should continue to provide powerful incentive for the ongoing collaborations of basic scientists and clinicians. Recent evidence suggests that newborn metabolic screening is extremely successful in reducing the incidence and severity of some intellectual disability syndromes. On the other hand, the impact of more recent programs is less clear. For example, although folic acid supplementation appears to significantly reduce the risk of neural tube defects, and may also reduce the risk of autism, compliance with recommendations to increase dietary folate appears to be disturbingly low relative to knowledge that supplementation is beneficial. It also appears that the prevalence of Trisomy 21 is likely to remain unchanged despite the availability of prenatal diagnostic programs. A recent review of the medical literature concluded that while the prevalence of intellectual disability caused by specific medical conditions has decreased significantly over the past 50 years due to medical interventions, only a small dent has been made in the prevalence of ID overall.

Influence of Etiological Factors on Treatment

As suggested above, the approach to treatment begins with the diagnostic process. In some cases, the underlying cause of intellectual disability may be particularly important in considering treatments. For example, in intellectual disability associated with phenylketonuria, a number of attempts have been made to minimize or to attenuate hyperactivity and impulsivity by dietary modification. More recently, as animal models of this disorder are developed and explored, it appears that dopaminergic abnormalities may play a particularly salient role in the expression of maladaptive behaviors and perhaps in the design of specific treatments. Thus, early intervention with dopamine agonists may ultimately play a unique role in the treatment of behavioral and cognitive disturbance associated with PKU.

Abnormalities in serotonergic function have been reported in the context of Down syndrome. This hyposerotonemia led to trials of serotonin replacement with mixed results, but serotonergic drugs may yet hold particular relevance for persons with Down syndrome. In the context of PWS, serotonergic abnormalities are also being reported as responses to serotonergic agents. Taken together, the identification of the underlying cause of intellectual disability has become increasingly important in considering biological treatments. In the same way, the diagnosis of mental disorders or syndromes in persons with ID will guide and influence treatment strategies.

TREATMENT OF MENTAL DISORDERS

Experience over the past decade has made it clear that intellectual

disability is a multidisciplinary problem, and optimal treatment is multimodal. Biological, psychological, social, and developmental dimensions should all be considered when designing a treatment plan for an individual with intellectual disability. Further, an integrative treatment approach should focus not only on ameliorating symptoms, but on restoring mental well-being, and caring for the family system as well. Typically, a treatment plan will include attention to psychoeducational, psychotherapeutic, and psychopharmacological interventions.

Psychotherapy

One should never assume that a person with ID cannot benefit from psychotherapeutic intervention simply because of his or her impairment in intellectual functioning. That being said, there are a number of potential barriers to effective psychotherapeutic treatment approaches, including limited verbal ability, limited cognitive flexibility, psychological-mindedness, and limited self-efficacy. In order to overcome these barriers, certain modifications in approach are beneficial. It is important, for example, that an active therapeutic stance be employed, as well as concrete and supportive interventions and careful attention to the language abilities and developmental level of the patient in treatment. When these types of alterations are made, many patients with ID clearly are able to benefit.

Specific psychotherapeutic approaches that have been shown to be effective include behavioral (in particular, applied behavior analysis models), cognitive–behavioral, psycho-educational, and skills training (e.g., coping skills, social skills). Behavioral therapies are demonstrably effective in managing many maladaptive behaviors, particularly aggression and self-injury, in persons with ID. Interventions that should be considered include environmental modifications, caregiver education, and applied behavior analysis (ABA). There is a large evidence base for ABA, and recent attention has been paid to dealing with low-base-rate behaviors, improving generalization, and incorporating choice-making to improve outcomes. Typically, a behavioral assessment begins with a "functional analysis" of behavior or detailed examination of the variables that reinforce or maintain particular behaviors. One considers the antecedent events and consequences of a behavior in question, and typically must test hypotheses by systematically manipulating key variables. Matson and colleagues have shown that a checklist, called the Questions about Behavioral Function (QABF), can be used successfully to derive clear behavioral functions for most individuals (84 percent) across target behaviors of self-injury, aggression, and stereotypy. Further, when the results of the functional analysis informed the behavior program, subjects improved significantly more than those whose behavior programs were more generic. While screening instruments may thus be quite useful in beginning to identify reinforcing variables for maladaptive behavior, the development and

implementation of an effective behavior program is best conducted by a behavioral psychologist.

In addition to behavior therapies, there is some evidence for use of cognitive–behavioral therapy to treat depression in individuals with mild intellectual disability. Skills training approaches, such as teaching individuals with ID brief relaxation techniques, may be useful in reducing anxiety even in the context of moderate to severe intellectual disability. Parent skills training approaches have been shown to improve developmental, behavioral, and family outcomes. Supportive groups for parents and siblings may also be of particular benefit. Family therapy can be particularly useful and provides a setting for education, for support, and for consolidation of behavioral treatment and other interventions.

An important area of current research is to examine not only efficacy but service delivery. A recent review of a range of psychosocial interventions (behavior analytic, cognitive training and rehabilitation, and parent training interventions) delivered by non-specialist care providers shows that interventions delivered in this way do produce good outcomes compared to no treatment or treatment as usual. This finding is critical given the scarcity of specialists in rural and low-resource areas and countries.

Pharmacotherapy

There is a relative paucity of controlled clinical drug trials that include persons with intellectual disability, particularly individuals with the most severe cognitive impairment. As a result, clinicians must generally extrapolate from the general population with regard to pharmacotherapy in persons with ID. Fortunately, there is little reason to suspect that drugs discriminate on the basis of intellectual ability in terms of their mechanisms of action.

However, because of drug–drug interactions that may affect the availability or effectiveness of concurrent medications, individuals with intellectual disability may require relatively unique dosing strategies. In institutionalized populations, for example, where 30 to 40 percent of persons have epilepsy and as many as 70 percent may have some other medical condition of significance, drug interactions become an increasingly important consideration. At the same time, studies are also making it clear that genetic copy number variants associated with both loss of function and gain of function may be associated with some intellectual disability syndromes. As these data come to inform more personalized or targeted treatment approaches, it is likely to be the case that some medications will emerge as being specifically beneficial or specifically contraindicated in certain ID syndromes relative to the general population.

Antidepressants. Special considerations in the use of antidepressant

medication in ID include the common medical comorbidities as noted previously. Tricyclic antidepressants in particular (e.g., clomipramine) must be used with the knowledge that the risk of lowering seizure threshold is relevant. Although this risk in the general population is in the order of 1 in 1,000, in individuals with intellectual disability this risk may increase to closer to 1 in 5. Cardiac anomalies are common in some intellectual disability syndromes, and the anticholinergic side effects of some medications may be particularly significant in persons with Down syndrome. Individuals with intellectual disability may require lower levels of antidepressant drugs than their normally developing peers, and disinhibition has been described even with typical antidepressant doses of the serotonin reuptake inhibitors.

There are some individuals—particularly a subgroup within the pervasive developmental disorder spectrum—who appear to exhibit extraordinary sensitivity to drugs of this and other classes. This sensitivity, when it occurs, is manifest by disinhibition, and may present as an increase in target symptoms at doses of medication that might otherwise appear quite standard. In this context, individuals may be converted to responders with significant dose reductions, for example, reducing a starting dose of fluoxetine (Prozac) or citalopram (Celexa) from 10 to 2 mg, daily. Moreover, it is increasingly the norm that initiation doses for selective serotonin reuptake inhibitors (SSRIs) in this population are lower, and the rate of dose increase slower, than for the general population.

A subgroup of persons with SIB also exhibit self-restraining behaviors, for example, binding their extremities in their clothing or seeking out helmets to wear for self-protection. Such behavior might indicate an ego-dystonic quality of self-injury and suggests the possibility of SIB as a compulsion. As noted earlier for PWS, obsessive-compulsive disorder occurs more commonly than in the general population, and compulsive skin picking may improve along with other compulsive behaviors with an SSRI.

Given these findings, trials of serotonin reuptake inhibitors in clinical practice are increasingly common among patients with SIB. Favorable experiences have been reported for fluoxetine, paroxetine, sertraline, trazodone, and clomipramine in this regard but all of these reports involve small numbers of subjects. In addition, of these agents, only clomipramine has been shown to be useful in a well-controlled study. Due to its effect of lowering seizure threshold, clomipramine is generally not a first-line treatment for compulsive SIB in individuals frequently comorbid for epilepsy.

Anticonvulsants. Data on the use of anticonvulsant medications for indications other than epilepsy in this population are limited. However, considerable experience suggests that, as for the population in general,

some anticonvulsant drugs may improve cyclical mood disorders and impulsive aggression. Valproate has probably been studied most in this regard. Current prevalence surveys suggest that across all indications, carbamazepine is the most widely prescribed anticonvulsant for persons with intellectual disability.

Crawford and colleagues performed an open-label but randomized, parallel-group, multicenter, add-on study comparing gabapentin with lamotrigine in 109 patients with treatment-resistant focal epilepsy and intellectual disability. Subjects in both the gabapentin and lamotrigine groups experienced a reduction in seizure frequency by nearly half. Both drugs were equally well tolerated. Gabapentin may have had an edge with respect to some behavioral indices, denoted by the investigators as communication, cooperation, and restlessness, but lamotrigine also was credited with improving challenging behavior. Thus, both medications may hold promise for persons with intellectual disability and treatment-resistant epilepsy, and their effects on behavior should be the focus of additional study.

Of interest, there is some experience with a small number of persons with PWS treated with the anticonvulsant topiramate whose skin-picking behavior improved. The mechanism underlying this effect is unknown, and it might even involve peripheral effects of the drug, for example, on sensation, or perhaps even rate of wound healing.

Conversely, SIB and aggression can also dramatically improve with the withdrawal of phenobarbital. The potential behavioral "toxicity" of this and any drug should not be overlooked in designing treatment strategies for persons with intellectual disability. Moreover, medical comorbidity is the rule in individuals with developmental disabilities, and the importance of identifying and treating underlying medical problems or refining that treatment cannot be overstated.

Anxiolytics. Although benzodiazepines are commonly prescribed in treatment for anxiety in the general population, there are unique concerns when they are used in the context of developmental disorders, particularly regarding the possibility of increased confusion, cognitive impairment, unsteadiness, and paradoxical excitement. Barron and Sandman reviewed disinhibition associated with benzodiazepines that occurred in 35 to 68 percent of an institutionalized intellectually disabled population as compared to controls. Nevertheless, alprazolam, clonazepam, and lorazepam are widely used for the treatment of acute anxiety, particularly that associated with procedures. Their use should be considered in the absence of evidence of previous abnormal responses to these agents.

On the other hand, recent findings from animal modeling of haploinsufficiency of the SCN1A gene (Dravet Syndrome) by Catterall and colleagues suggest that activation of GABA(A) receptors with low dose

clonazepam improves behavioral and cognitive impairments in mice. There is growing interest in this mechanism in genetic conditions associated with ID.

Buspirone is another serotonergic agent that has been reported to be of benefit in some persons with developmental disorders. Ratey and colleagues reported the use of buspirone in an open trial in persons with diagnosed anxiety disorders manifested by aggressive and self-injurious behaviors. Typical doses at which patients responded were on the order of 15 to 45 mg per day. In another open trial by King and Davanzo, it appeared that nonresponders to buspirone remained so, even as the medication dose was increased from the levels used by Ratey to 60 mg daily. A study by Verhoeven and Tuinier also explored buspirone in the treatment of anxiety in persons with intellectual disability, suggesting that this agent should be among those considered in treating this population. Moreover, advantages include a relatively benign side-effect profile, specifically the absence of common motor or cognitive toxicities, at doses that are used to treat anxiety.

Antipsychotics. Antipsychotic medications have long been used, arguably indiscriminately, in persons with intellectual disability. These are the most widely prescribed class of psychotropic medication—even more commonly prescribed than anticonvulsant drugs for persons with intellectual disability. Given this ample experience in persons with intellectual disability, such that in residential or institutional settings as many as 50 percent of individuals may be treated with this class of drug, the side effects are well known.

Individuals with intellectual disability appear to be at greater risk of developing tardive dyskinesia than the general population, with recorded rates ranging from 18 to greater than 30 percent following chronic receipt of first-generation antipsychotics. On the other hand, spontaneous abnormal involuntary movements are not uncommon in this population, and this may represent a confound in interpreting rates of neuroleptic-induced tardive dyskinesia. The ascendance of atypical neuroleptic drugs and their use in persons with intellectual disability have reduced the risk of tardive dyskinesia and extrapyramidal symptoms, as well as cognitive toxicity. Considerable evidence is accumulating to support the effectiveness of risperidone in treating severe disruptive behavior in children and adolescents with cognitive disability. Large multisite trials have found that at daily doses ranging from 0.02 to 0.06 mg/kg, children and adolescents with various disruptive behavior disorders, the majority of whom also had attention-deficit/hyperactivity disorder, improved significantly. Sedation and weight gain were the most common side effects. Similar studies exist for aripiprazole although both of the controlled trials were sponsored by the manufacturer. Studies reporting the effectiveness of clozapine,

olanzapine, quetiapine, and sulpiride, through open trials involving small numbers of subjects, also support the use of these agents in individuals with intellectual disability.

There is evidence to support the use of dopamine antagonists in SIB and aggression both in theory and in practice. For thioridazine, in particular, ample experience exists to indicate that both SIB and aggression may improve. Typical doses average less than 300 mg per day, with daily doses as low as 15 mg per day being reported for some children. A "black box warning" about QT prolongation induced by thioridazine and the availability of the atypical neuroleptics have dramatically reduced the use of thioridazine in this population.

Although suggested by some, there is no convincing evidence that the mechanism of action of antipsychotics in SIB or aggression is merely to suppress behavior generally through a nonspecific sedating effect. Such an outcome is clearly undesirable in an individual with pre-existing cognitive impairment, and many of the earliest reports specifically note the absence of sedation.

Psychostimulants. Despite reports of paradoxical responses to stimulant medications in persons with intellectual disability, with higher-than-expected rates of emergent motor tics and emotional lability, a growing body of literature supports the use of stimulant drugs for the treatment of ADHD in the context of intellectual disability. A problem awaiting additional study concerns the likelihood of a response to stimulants in relation to the severity of cognitive disability. Data are mixed with respect to whether IQ negatively or positively predicts treatment response.

Atomoxetine is a nonstimulant that is used for similar target symptoms in the general population. There are no studies specifically examining its effectiveness in ID, but growing evidence suggesting that in autism spectrum disorders, atomoxetine may be useful for target symptoms of inattention, hyperactivity, and noncompliance/oppositionality.

Opioid Antagonists. Some individuals who self-injure appear to have altered pain sensitivity, as was suggested nearly a century ago by Ireland. This observation, coupled with data indicating that opioid antagonists can attenuate stereotypy and self-injury in animal models and data indicating that opioids may modify the function of dopaminergic systems, has fueled interest in this class of drugs. Naltrexone is the opioid antagonist most widely used for SIB, but the literature is mixed. Typical daily doses range from 0.5 to 2.0 mg/kg in children and up to 200 mg in adults. The relatively long half-life of naltrexone, 72 hours in brain, must be kept in mind in designing titration strategies for this drug. On balance, naltrexone appears to be well tolerated in persons with developmental disorders, with

sedation as the side effect most likely to be observed.

Nootropics. Perhaps the "holy grail" of psychopharmacology in intellectual disability would be drugs that positively affect cognition, or nootropic drugs. There is no such drug that has yet to demonstrably improve cognition in ID, but the growing number of strategies that are emerging that "rescue" animal models of specific genetic disorders may redefine this class of therapeutics for the population with ID. Even as some of these novel treatments move in to clinical trials, the question as to whether there are particular developmental windows for therapeutic effectiveness remains important to address.

Other Drugs. There is ample evidence to support an antiaggressive effect of lithium salts in the context of intellectual disability. Similarly, there is evidence to suggest that in the setting of cyclical mood disturbance, lithium may also be helpful. The side effects associated with this agent, particularly the cognitive dulling that is occasionally reported by some recipients without intellectual disability, should raise the threshold for the use of this agent in the setting of cognitive disability.

Beta-adrenergic antagonists have also been reported to be of use in the population with developmental disorders. It is not clear, however, whether the mechanism is central or peripheral. Ratey and colleagues were led to explore this class of drugs based on their perception that many individuals with developmental disabilities appear to have a low frustration tolerance for ambient stimuli. At doses of propranolol upward of 1,000 mg per day, the precise mechanism of action of this drug arguably becomes nonspecific, but lower doses of this and other beta-adrenergic antagonists have also been reported to be beneficial.

A growing appreciation exists for the potential therapeutic utility of drugs acting at glutamate receptor subtypes. Glutamatergic and dopaminergic interactions in the neostriatum are the focus of research on the pathogenesis of a host of neuropsychiatric illnesses, including schizophrenia, obsessive-compulsive disorder, SIB, and aggression. Dextromethorphan, an antitussive agent that also binds to the *N*-methyl-D-aspartate (NMDA) receptor, was reported by Welch and Sovner to have markedly attenuated SIB in a 25-year-old individual with congenital rubella syndrome. The patient was maintained on dextromethorphan for 16 months with sustained benefit. It is surprising and perhaps ominous that no follow-up studies have been reported with this drug in SIB. Lamotrigine is an anticonvulsant drug that also appears to antagonize glutamate by reducing its release. A recent case report suggested that lamotrigine may also be useful in reducing self-injury in the context of a stereotyped movement disorder. A larger study in which lamotrigine was added to the anticonvulsant regimen of children with ID suggested that those with

autism experienced gains that could not readily be attributable to better seizure control alone. A recent double-blind, placebo-controlled trial of lamotrigine in autism, however, was negative.

Amantadine is another drug with affinity at the NMDA receptor. An open trial of the use of amantadine in children with various developmental disabilities and disruptive behaviors was quite promising. A subsequent small multisite, controlled trial in children with autism yielded mixed results. The closely related drug memantine has also shown promise in open trials for agitation in the context of autism.

Sleep disturbance is a little-studied but often significant problem for some persons with intellectual disability. Jan prescribed melatonin openly to children with moderate to severe intellectual disability of mixed etiologies. The children ranged in age from 1 to 11 years. All but two of the children were characterized as dramatic responders based on measures of hours of sleep, nocturnal awakenings, delayed sleep onset, and early-morning awakening. Melatonin was administered as a single 3-mg dose 1 to 2 hours before target bedtime each night. Pillar and colleagues also administered a single 3-mg dose and monitored sleep efficiency, as well as activity (wrist actigraph), in their sample of five children (3.5 to 13 years old) with severe intellectual disability. Sleep was significantly improved in the subjects receiving melatonin, and two of the subjects were still receiving melatonin up to 18 months after the completion of the study. This preliminary evidence suggests that melatonin deserves broader consideration for the treatment of children with intellectual disability and disturbed circadian rhythm of sleep.

Services and Supports

In contrast to even 20 years ago, individuals with ID can now access a wide variety of services. These services can be found throughout the United States and last throughout the lifespan; see Table 40–10 for a listing of some of these services.

Table 40–10.
Service Needs and Resources for Families of Children at Different Ages

	Needs	Resources
Ages 0–3 years		
Child	Evaluation: physical, motor, cognitive, linguistic, social-emotional; early intervention services	Multidisciplinary evaluation, which results in an individualized family service plan (IFSP), with child and family receiving either center- or home-based early intervention services for set number of hours per week
Mother	Emotional support, caretaking behaviors	Support groups by disability, region, and etiology; part of early intervention evaluation, intervention, and IFSP
Family	Support, financial	Support groups; depending on the problem, state

	assistance, information	developmental disabilities or insurance payment for some services; hospitals, agencies, groups
Ages 3–21 years		
Child	Evaluation, referral, and individualized education program (IEP)	School system: involves legal process of evaluation and placement (notification, hearings, appeals if necessary); information on "transition" to adult services as child nears age 21 years (and school services end).
Family	Information, financial assistance, support	Local and national groups; departments in some states, including respite care, camps, art (Very Special Arts) or athletic activities (Special Olympics), scholarships for adolescents with some disabilities (deafness, blindness)
Older than 21 years		
Offspring	Residential services, work	Run by state developmental disability departments; parents and offspring have major say concerning whether residential or work placements are appropriate.
Family	Support, information, guardianship issues	Continuation of many of the services provided during the school years; particularly for individuals with severe disabilities, provisions for residential and work status after parents can no longer serve as offspring's legal guardians

From Hodapp RM. *Development and Disabilities*. New York: Cambridge University Press; 1998:196.

Early intervention and educational services became widespread following the passage of PL 94–142 (the Education for all Handicapped Children Act of 1975); this legislation has recently been extended and modified by IDEA, the Individuals with Disabilities Education Act of 1990, with amendments in 1997 and 2004.

Early Intervention. Early intervention serves young children from birth until a child's third birthday. Such services are supported by the state, with a different "lead agency" coordinating services in each state. Services vary widely from state to state. For instance, in some states, birth to 3 (B-3) care may begin with early intervention specialists providing services, such as speech and language therapy, in the child's home, while in other states B-3 care can also entail center-based services. Once the child is preschool age, most services are center-based and include both one-on-one therapies and small group education services. B-3 and preschool services focus on multiple domains of functioning, including cognitive development, language, motor, social, and adaptive functioning. With the passage of Part H of PL 99–457 (the Education of the Handicapped Amendments of 1986) early intervention services have increasingly emphasized the family's role. Agencies are required to develop an Individualized Family Service Plan (IFSP) for each child and family, focusing on family strengths and needs in caring for their young child with intellectual disability.

School. From ages 3 to 21, specialized education services are provided through local school districts. The Individuals with Disabilities Education Act 2004 requires that public school systems serve all children with

disabilities, and outlines the legal provisions to do so. This process provides for notice that the child is being considered for special services; a formal, scheduled hearing at which the parent(s) and school personnel discuss what is the most appropriate school placement and specialized services; the right to present a defense if parents disagree with school personnel; a written decision given to the parents notifying them of the school's decision; and the parents' right to appeal this decision, first to the school itself, then to district and state educational authorities, and eventually to the court system. This entire process results in an Individualized Education Program (or Plan), called an IEP, which assesses the child's academic performance, provides annual goals and instructional objectives, describes services that the child will receive, and includes starting dates of services and when and how the child's progress will be re-evaluated. Although receiving services for one's child can be complicated, educational services have become a right—not a privilege—for parents of children with intellectual and developmental disabilities.

A further element of education law involves the right to a public education "within the least restrictive environment." This phrase has led to many debates within the special education community about inclusion. Essentially, this debate centers on whether all children should be educated in classes composed primarily of typically developing children. Inclusion advocates conceptualize education alongside typically developing children as every child's right, while proponents of a continuum of care argue that the child's individual educational needs should be paramount.

Special educators have long advocated "non-categorical" approaches to children with ID. Such approaches do not consider the cause of the child's disability, and even knowing about the etiology of the disability has been considered harmful, possibly leading to increased stigmatization.

Over the past few years, however, it has become apparent that the etiology of the child's intellectual disability does influence behavior. As discussed above in reference to Prader–Willi, Fragile X, Williams, and Down syndromes, individuals with each syndrome do differ from others with intellectual disability in maladaptive behavior and psychopathology, but also in relative strengths (or weaknesses) in language versus other abilities and across cognitive domains. Such etiology-related profiles may eventually lead to etiology-related interventions.

Consider Down syndrome. Most children with Down syndrome show particular difficulties in linguistic grammar, expressive language, and articulation, even as their abilities in visual short-term memory appear relatively strong. Thus, when asked to recall a series of hand movements on the Kaufman Assessment Battery for Children (K-ABC), these children perform better than when recalling a series of spoken numbers or words. Using this visual-over-auditory profile, various researchers have become interested in teaching children with Down syndrome to read. In short, one

can use a strength to bootstrap a weakness, using the visual modality as a "way into" language.

Similar suggestions apply to other syndromes. Children with Williams syndrome generally show higher language than visuospatial abilities, while children with either Prader–Willi or Fragile X syndrome show better abilities in simultaneous (i.e., Gestalt) processing compared to sequential (step-by-step, serial) processing. In each instance, educators and test-constructors have already devised ways to help children with a particular learning style, and such specific, tailored educational approaches might prove particularly effective as well for children with different intellectual disability syndromes.

Although such "etiology-based" educational interventions hold promise, such approaches will not necessarily prove effective for every child with a particular syndrome. Not all children with Down syndrome have particular difficulties in language; not all children with Williams syndrome demonstrate especially high language or low visuospatial skills. While etiology-related interventions may be helpful, they need to be considered in light of the child's own specific strengths and weaknesses.

Further, most etiology-related interventions have adopted the approach of "playing to strengths" as opposed to ameliorating weaknesses. To give a few examples, in Down syndrome, visual short-term memory is relatively more advanced compared to auditory short-term memory in the teen than in the grade-school child. Similarly in Williams syndrome, more pronounced patterns of high language and weak visuospatial skills are seen in older as opposed to younger children. In this sense, playing to strengths may reinforce already strong areas, but hopefully in ways that help children to learn a wide variety of information and skills. Despite the ultimate nature of etiology-related interventions, as more information about etiology-related strengths and weaknesses in a variety of syndromes becomes available, such information should gradually become incorporated into more targeted, more effective programs of intervention.

Adult Services. Once an individual with intellectual disability turns 21, service delivery comes under the auspices of each state's department of developmental disabilities. In contrast to educational services, adult services are not federally mandated. States thus vary widely in their services and how they offer them. Services are often spread out across various locations, and qualifying and getting services can sometimes be difficult or impossible.

During the adult years, residential and vocational services predominate. Residential services run the gamut from institutional to community-based, and from constant to no supervision. More restrictive placements involve large or small institutional residences, which mainly serve the lowest functioning, often multiply-impaired individuals. Group homes for 4 to 8

residents are probably the most common residential option; these homes typically exist within residential neighborhoods and provide round-the-clock staff to ensure safety and consistent programming. Supervised living arrangements usually consist of apartments that are shared by a few persons with ID; these residents receive visits from program staff, who often instruct residents in cooking, cleaning, and shopping. Unsupervised apartments, where individuals live by themselves or with a roommate under little supervision, mainly succeed with clients who have lesser intellectual impairments and fewer behavior problems.

Vocational services also range from most to least restrictive. Most restrictive are sheltered workshops, in which individuals with ID assemble or test products next to others with disabilities. Supported employment provides a job coach or other specially trained individual to work with the client to find a job, to learn how to take public transportation to get to a job, and to acquire the technical and social skills needed to handle the job. Competitive employment is the least restrictive job option for adults with intellectual disability. Competitive employment involves a job with the same pay levels and benefits that a nondisabled person would have, but also with the same pressures and responsibilities for work performance. As in unsupervised apartment living, competitive employment is an option for the highest functioning and least behaviorally disordered individuals.

In considering working and living arrangements for any young adult with intellectual disability, one should also consider transitional services. There is a federal mandate requiring youth with disabilities age 16 and older to have an IEP with a postsecondary transition plan. This plan typically includes teaching vocational and adaptive skills in support of long-term life goals. Parents and schools must work together to determine the most appropriate long-term goals for each young adult.

Supports

In addition to services provided by early intervention, school, and adult services, individuals with intellectual disability and their families can also benefit from a variety of other supports. These run the gamut from short- or longer-term respite care (allowing families a break from full-time care), to summer camps, and to sports activity programs such as Special Olympics (see Fig. 40–1). In addition, at several major universities, specialty clinics exist for behavioral and medical management of certain types of intellectual disability.

A wide variety of organized groups are also helpful for affected individuals and their families. Parent groups are available for parents of children with disabilities and for parents of children with specific types of disability. Sibling groups and networks are also available, and the federal government fosters service and advocacy through the University Affiliated Programs (UAP's) in almost every state.

Legal Issues in Intellectual Disability. The Americans With Disabilities Act of 1990 spells out four aims: (1) to provide a clear and comprehensive national mandate for the elimination of discrimination against individuals with disabilities; (2) to provide clear, strong, consistent, enforceable standards addressing discrimination against individuals with disabilities; (3) to ensure that the Federal Government plays a central role in enforcing the standards established in the Act on behalf of individuals with disabilities; (4) and to invoke the sweep of congressional authority, including the power to enforce the 14th Amendment and to regulate commerce, in order to address the major areas of discrimination faced day-to-day by people with disabilities. The ADA represents a major turning point in the assurance of the rights of persons with disabilities who, even 50 years ago were frequently institutionalized and compulsorily sterilized in misguided efforts to protect society from this population. While the pendulum has shifted repeatedly from attitudes of asylum to punishment of persons with disabilities over the centuries, the prevailing paradigms in developed countries are those of community integration and support. Disabilities law over the past three decades reflects these values as articulated in Section 504 of the Rehabilitation Act of 1973, and subsequently the Individuals With Disabilities Education Act (IDEA) of 1990 (originally Pub. L.-94–142 of 1975), with amendments in 1997 and 2004, and the ADA. Collectively, these laws ensure that all children with disabilities receive a free appropriate public education in the least restrictive environment, and that reasonable accommodation is made for the employment of adults. Clinicians working with this population should be aware of their personal obligation to provide an accommodating environment, but should also familiarize themselves with local guardianship, informed consent, and abuse reporting laws in order to best serve this population.

FIGURE 40–1. Athletes competing in activities sponsored by the Special Olympics, whose athletes take the following oath: "Let me win. But if I cannot win, let me be brave in the attempt." The Special Olympics was founded in 1968 by Eunice Kennedy Shriver. The mission of the organization is to provide year-round sports training and athletic competition in a variety of Olympic-type sports for individuals with intellectual disability by giving them continuing opportunities to develop physical fitness, demonstrate courage, experience joy, and participate in a sharing of gifts, skills, and friendship with their families, other Special Olympics athletes, and the community.

References

Abbeduto L, Brady N, Kover ST. Language development and fragile X syndrome: profiles, syndrome-specificity, and within-syndrome differences. *Ment Retard Dev Disabil Res Rev*. 2007;13:36–46.

Aït-Yahya-Graison E, Aubert J, Dauphinot L, et al. Classification of human chromosome 21 gene-expression variations in Down syndrome: impact on disease phenotypes. *Am J Hum Genet*. 2007;81:475–491.

*Aman MG, Alvarez N, Benefield W, Crimson ML, Green G. Expert Consensus Guideline Series: treatment of psychiatric and behavioral problems in mental retardation. *Am J Ment Retard*. 2000;105(3):159–226.

American Association on Intellectual and Developmental Disabilities. *Mental Retardation: Definition, Classification, and Systems of Supports*. Washington, DC: Author; 2002. Atkins v. Virginia. 536 U.S. 304 (2002).

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders. Fifth Edition (DSM-5)*. Washington, DC: American Psychiatric Association; 2013.

Battaglia A, Carey JC. Diagnostic evaluation of developmental delay/mental retardation: an overview. *Am J Med Genet C Semin Med Genet*. 2003;117C(1):3–14.

Brosco JP, Mattingly M, Sanders LM. Impact of specific medical interventions on reducing the prevalence of mental retardation. *Arch Pediatric Adolesc Med*. 2006;160:302–309.

Carter EW, Hughes C. Increasing social interaction among adolescents with intellectual

disabilities and their general education peers: effective interventions. *Res Pract Persons Severe Disabil.* 2005;30(4):179.

Crawford DC, Acuna JM, Sherman SL. FMR1 and the fragile X syndrome: human genome epidemiology review. *Genet Med.* 2001;3:359–371.

Decobert F, Grabar S, Merzoug V, et al. Unexplained mental retardation: is brain MRI useful? *Pediatr Radiol.* 2005;35(6):587–596.

De Leersnyder H. Smith-Magenis syndrome. *Handb Clin Neurol.* 2013;111:295–296.

de Ruiter KP, Dekker MC, Verhulst FC, Koot HM. Developmental course of psychopathology in youths with and without intellectual disabilities. *J Child Psychol Psychiatry.* 2007;48(5):498–507.

Engbers HM, Berger R, van Hasselt P, et al. Yield of additional metabolic studies in neurodevelopmental disorders. *Ann Neurol.* 2008;64(2):212–217.

*Fletcher R, Loschen E, Stavrakaki C, First M, eds. *Diagnostic Manual—Intellectual Disability: A Textbook of Diagnosis of Mental Disorders in Persons with Intellectual Disability*. Kingston, NY: NADD Press; 2007.

Gothelf D, Furfaro JA, Penniman LC, Glover GH, Reiss AL. The contribution of novel brain imaging techniques to understanding the neurobiology of mental retardation and developmental disabilities. *Ment Retard Dev Disabil Res Rev.* 2005;11:331–339.

Grey IM, Hastings RP. Evidence-based practices in intellectual disability and behaviour disorders. *Curr Opin Psychiatry*. 2005;18:469–475.

*Harris JC. *Developmental Neuropsychiatry*. New York: Oxford University Press; 1995.

*Hodapp RM. *Development and Disabilities: Mental, Motor, and Sensory Impairments*. New York: Cambridge University Press; 1998.

Hodapp RM, Ricci LA. Behavioural phenotypes and educational practice: the unrealized connection. In: O'Brien G, Udwin O, eds. *Behavioural Phenotypes in Clinical Practice*. London: MacKeith Press; 2002:137.

*Hodapp RM, Zigler E. Past, present, and future issues in the developmental approach to mental retardation. In: Cicchetti D, Cohen D, eds. *Manual of Developmental Psychopathology. Vol. 2. Risk, Disorder, and Adaptation*. New York: Wiley; 1995:299.

Jarrold C, Baddeley AD, Hewes AK, Phillips C. A longitudinal assessment of diverging verbal and non-verbal abilities in the Williams syndrome phenotype. *Cortex.* 2001;37:423–431.

Koskentausta T, Livanainen M, Almqvist F. Risk factors for psychiatric disturbance in children with intellectual disability. *J Intellect Disabil Res.* 2007;51(1):43–53.

Lee RW, Conley SK, Gropman A, Porter FD, Baker EH. Brain magnetic resonance imaging findings in Smith-Lemli-Opitz syndrome. *Am J Med Genet A.* 2013;161A(10):2407–2419.

Low-Functioning Autism and Nonsyndromic Intellectual Disability. Magnetic resonance imaging (MRI) findings. *J Child Neurol.* 2015;30(12):1658–1663.

Nirje B. The normalization principle and its human management implications [1969]. Reprinted: Rosen M, Clark GR, Kivitz M, eds. *A History of Mental Retardation: Collected Papers*. Vol. 1. Baltimore, MD: University Park Press; 1976:361.

*Reiss S, Aman MG, eds. Psychotropic Medication and Developmental Disabilities: The International Consensus Handbook. Columbus, OH: OSU Nisonger Center; 1998.

Scheerenberger R. A *History of Mental Retardation*. Baltimore, MD: Brookes; 1983.

Sedel F, Baumann N, Turpin JC, Lyon-Caen O, Saudubray JM, Cohen D. Psychiatric manifestations revealing inborn errors of metabolism in adolescents and adults. *J Inherit Metab Dis*. 2007;30(5):631–641.

Sparrow SS, Balla D, Cicchetti DV. *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service; 1984.

*Szymanski L, King BH. Practice parameters for the assessment and treatment of children, adolescents, and adults with mental retardation and comorbid mental disorders. *J Am Acad Child Adolesc Psychiatry*. 1999;38(12 Suppl):5S–31S.

41

Specific Learning Disorder

ROSEMARY TANNOCK, PH.D.

INTRODUCTION

Humans are not born to read, write, spell, or perform arithmetic calculations: These are skills that have to be learned. In contrast to oral language skills, which humans have possessed for over 100,000 years, cultural symbols used to represent specific numbers and speech sounds were invented only a few thousand years ago.[1]

These arbitrary cultural symbols, which are the foundational codes for reading, writing, and arithmetic, are not acquired "naturally" by simple exposure; each human being has to learn these symbols as well as how they map onto specific speech sounds, specific quantities, and numerals. The process of developing these literacy and numeracy skills reorganizes and changes each individual's brain. This is because the human brain did not evolve to accommodate the recent human invention of symbols, which are far too recent to have exerted any evolutionary pressure on brain evolution. Thus, during the human developmental period, existing functions and networks in the brain are "recycled" for the new uses. For example, learning to read involves the creation and automation of an efficient interface between vision and spoken language ("visual word form area" in the left visual cortex); learning arithmetic involves creating connections with nearby areas of the parietal cortex that code for space and eye movement.

Multiple skills are involved in learning to process one's cultural symbol systems. These include the awareness that spoken language can be segmented into smaller units (words, syllables, phonemes) and mapped onto written visual symbols (e.g., letters, characters, numerals); the ability to rapidly identify letters, letter clusters, and words (orthographic awareness); the ability to extract meaning from this written language; and the ability to transcode meaning between symbol systems (e.g., to understand that the symbol "3," the quantity of three dots (***), the spoken word "three" (/th/ /r/ /ee/) and the written word "three"; all represent "threeness" (see Table 41–1 for definitions of key terms related to reading and number sense).

For a surprisingly large subgroup of children (at least 6 percent), this multidimensional developmental task of learning to process and use one's

cultural symbols present major lifelong challenges that are unexpected, given their abilities in other developmental domains. The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) uses the term "Specific Learning Disorder" (SLD) to describe persistent and impairing problems in acquiring and using the cultural symbols that are required for reading, writing, and arithmetic.

Significance of SLD for Psychiatrists

At first glance, the diagnostic category of SLD may appear of lesser importance to psychiatrists and other clinicians in their quest to understand and help individuals suffering from the burden and stigma of mental health problems. However, SLD has high comorbidity rates with other common psychiatric disorders, ranging from ADHD (most common), anxiety, and mood disorders, suicide, to substance abuse (least common). Thus clinicians need to consider the presence of SLD when assessing for these other psychiatric disorders. Also, the burgeoning concern about health literacy and health numeracy in the medical literature highlights the need for clinicians to understand the medical implications of SLD. Health literacy is defined as the ability to obtain, process, and understand basic health information and services needed to make appropriate health decisions; health numeracy is defined as the individual-level skills needed to understand and effectively use quantitative health information (e.g., dose, quantity, probability) to guide health behavior and decision making. Basic reading skills (e.g., letter–sound knowledge and word recognition), reading comprehension, and basic numeracy skills are core attributes that underlie health literacy and numeracy. Poor health literacy and numeracy are independent predictors of health disparities, worse health outcomes, and mortality.

Recent surveys of the readability of patient education materials and medical research consent forms typically indicate that many of these materials have higher-than-acceptable levels of readability for the general public. For example, the American Academy of Pediatrics reported that at least half of their materials manifest unacceptably high readability levels. Few studies have investigated the readability of patient education materials in psychiatry, which is of particular concern, given the low educational attainment of many patients with mental health illnesses. Low numeracy skills constrain patient choice, reduce medication compliance, impair risk communication, and affect medical outcomes. SLD tends to run in families, and so patients and their families may have great difficulty accessing patient education materials and comprehending dosage instructions or nonpharmacological interventions. Clinicians in the workplace may also have SLD themselves (e.g., almost 2 percent of medical students declare a diagnosis of SLD in reading on entry to medical school), which may pose additional challenges in the effective performance of their clinical duties.

Thus an understanding of SLD and knowledge of effective strategies to enhance the clinical performance of clinical trainees and certified clinicians with SLD is of key importance for clinical educators.

To understand SLD, it is necessary to understand how the neuroplasticity of the human brain makes it possible to learn to read, write, and calculate, how the acquisition of literacy and numeracy transforms the brain, and how disturbances in brain development disrupt the normal acquisition of reading, writing, and arithmetic. Two of the most engaging accounts of the history and neuroscience of reading are by neuroscientists, Maryanne Wolf (*Proust and the Squid, 2007*) and Stanislas Dehaene (*Reading in the Brain*, 2009). Likewise, fascinating accounts of the earliest invention of numerals and how the human brain learns to understand and manipulate numbers, are to be found in Stanislas Dehaene's neuroscientific account of number sense (*The Number Sense*: *How the Mind Creates Mathematics*. Rev Upd, 2011), and in a mathematician's (Amir Aczel) personal account of his journey to find the first "zero"—the keystone of our current number system—eventually located on an inscribed stone in an old shed in Cambodia (*Finding Zero*, 2015).

Table 41–1.
Definition of Key Terms Used in Relation to Specific Learning Disorder

Term	Definition
Alphabetic principle	The insight that words are made up of sounds and each sound can be signified by a symbol.
Phoneme	Smallest sound units in a word: The smallest unit of speech that distinguishes one word from another (e.g., the word "peach" has three phonemes: /p/ /ea/ /ch/; the phoneme /p/ distinguishes the word "peach" from "teach").
Phonemic awareness	An explicit awareness that there are discrete speech sounds (phonemes) in speech. It begins to emerge between ages of 2 to 4 years, takes several years to develop, and is crucial to success in reading and spelling. For example, the child must be able to isolate the phoneme /p/ in words such as "pet," "peach," and "pear" to fully understand that the grapheme "p" represents this sound and to associate the sound with this letter when attempting to decode a word like "port."
Phonological awareness	An explicit awareness that spoken language comprises discrete units, ranging from words and syllables to smaller subsyllabic units of onsets and rimes, and phonemes. It includes phonemic awareness.
Phonological coding	Translating the letters or spelling patterns of a written word into speech patterns to identify the word and gain access to its meaning.
Phonics	An instructional approach for teaching children the systematic relationship between letters and sounds and how to use that system (alphabetic principle) to read words. In other words, instruction in phonics requires the child to have some level of phonemic awareness.
Onset-rime	Parts of monosyllabic words in spoken language: These units are smaller than syllables but larger than phonemes. An "onset" is the initial consonant sound of a syllable (the onset of "peach" is "p-"). The rime is the part of the syllable that contains the vowel and all that follows it (the rime of "peach" is "-each").
Orthography	The representation of the sounds of a language by written or printed symbols.

	Patterns link graphemes (letter) to sounds (phonemes). Orthography describes or defines the set of symbols (e.g., graphemes, diacritics) used and the rules about how to write these symbols.
Orthographic transparency	Deep orthographies are writing systems (such as English) that do not have a one-to-one correspondence between the spoken phoneme and the written grapheme. A "shallow" orthography means that the correspondences between letters and sounds (graphemes/phonemes) in the writing system are close to one-to-one (e.g., Finnish, German).

DEFINITION

SLD differs from most other mental disorders recognized by DSM-5, in that not only is it a clinical condition recognized and defined in the medical field, but also it is a discrete category of special educational needs (a.k.a., "exceptionalities") recognized and defined in the field of education. One major challenge for clinicians, educators, parents, and researchers alike, is that the purpose of classification differs in the medical/mental health and educational systems, as do their underlying concepts and definitions of SLD. Thus clinicians who diagnose SLD must understand both the medical and educational definitions, as well as their different implications for accessing school-based supports and intervention services.

Medical/Psychiatric Definition

According to DSM-5, SLD is a Neurodevelopmental Disorder that involves marked and persistent difficulties in learning and using one's cultural symbol systems (e.g., alphabetic letters, characters, Arabic numerals) that are required for skilled reading, writing, and arithmetic. SLD disrupts the normal pattern of learning these essential academic skills and is not simply a consequence of an intellectual disability, sensory deficits, other mental or neurological disorders, lack of opportunity of learning or proficiency in the language of instruction, or inadequate educational instruction. Difficulties learning keystone (foundational) academic skills of reading, writing, and arithmetic may also impede learning in other academic subjects (i.e., history, geography, science) but those problems are attributable to difficulties learning the basic underlying academic skills.

Neurodevelopmental disorders are complex and heterogeneous conditions arising from perturbations of the central nervous system, typically during the pre- and perinatal period, which interfere with the developmental of language and speech, motor skills, attention, behavior, memory, learning, or other neurological functions. Neurodevelopmental disorders are typically heritable, show early onset, persist through the lifespan, and impair academic, social, occupational, and mental health outcomes. DSM-5 delineates six neurodevelopmental disorders, many of which often co-occur: Intellectual Disabilities, Communication Disorder, Autism Spectrum Disorder, Attention-Deficit/Hyperactivity Disorder (ADHD), SLD, and Motor Disorders. The neurodevelopmental disorder of

learning (i.e., SLD) represents the extreme bottom end of normal distributions of culturally acquired skills, rather than a discrete categorical entity.

Like other neurodevelopmental disorders, SLD is understood to be a heterogeneous condition, in which the learning difficulties may be either selective (e.g., occur only with learning one specific aspect of reading such as word decoding) or broad (e.g., occur in learning to decode words, spell, understand what one has just read, as well as in arithmetic calculation), and frequently coexists with other neurodevelopmental (e.g., with Communication Disorders, ADHD) or psychiatric traits or disorders (e.g., with anxiety, anxiety disorders). Moreover, SLD is understood to be a heritable, chronic condition of childhood that typically persists into adulthood, albeit with changes in the ways the learning difficulties manifest.

Medical practitioners must understand that a medical diagnosis of SLD, using the DSM or ICD classification systems, does not necessarily mandate eligibility for special assistance at school.

Federal/Educational Definition

In the field of education, the term "Learning Disabilities" is used to delineate a specific category of "special learning needs," that mandates the provision of special education and other services to permit more equitable access to the curriculum and promote their learning.

In the United States, the conceptualization and definition of "Learning Disabilities" has been strongly influenced by the federal definition adopted by the U.S. Office of Education in 1969, operationalized by the U.S. Office of Education initially in 1977 (Individuals with Disabilities Education Act [IDEA]), and reauthorized in 2006 to determine eligibility for special education services that must be provided by all public schools. The 1977 IDEA definition of a specific learning disability refers to:

"A disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations. The term includes such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, and developmental dysphasia."

Note that the federal definition of Learning Disabilities includes oral or spoken language impairments, which are classified as Communication Disorders in psychiatric taxonomies and differentiated from SLD. Moreover, the educational definition incorporates notions of presumed etiology (i.e., deficits in underlying psychological processes), whereas psychiatric classification does not.

Identification of specific learning needs in the educational system mandates eligibility for special education services. Due in part to

differences in the purpose of identification, the educational system may identify some children as having special educational needs who would not meet the clinical threshold for a diagnosis of SLD in the medical system.

One substantive problem with the educational definition is that standards of identification of SLD differ across school boards, postsecondary institutions and across countries, so that an individual identified as SLD by one school board, or in one region of the country, or in one country, may not be identified in another. Moreover, as both educational laws and psychiatric taxonomies change over time, different groups of individuals are identified as having SLD or Learning Disabilities, depending on the specific set of criteria applied. Thus, the terms Specific Learning Disorder and Learning Disabilities are not synonymous. To date, there is no consensus definition of this condition: Current assessment and diagnostic practices in both medical and educational systems are varied, often opaque and discriminatory, and therefore vulnerable to legal challenges.

HISTORY AND COMPARATIVE NOSOLOGY

History

The term "Learning Disability" (credited to Samuel Kirk, 1962) was first recognized as a formal category in the field of special education in the 1960s, although the first clinical case of a child's unexplained difficulties learning to read had been described much earlier in the United Kingdom in 1896. In 1963, parents, educators, and policy specialists in the United States, agreed on the term "Learning Disabilities" to refer to a large group of individuals who had unexpected and significant difficulty learning to read, write or do arithmetic, despite average or above intellectual ability.

Kirk defined learning disability as:

> A retardation, disorder, or delayed development in one or more of the processes of speech, language, reading, writing, arithmetic, or other school subject resulting from a psychological handicap caused by a possible cerebral dysfunction and/or emotional or behavioral disturbances. It is not the result of mental retardation. Sensory deprivation, or cultural and instructional factors (Samuel Kirk, *Educating Exceptional Children*, 1962, p. 34).

Early terminology, which focused primarily on difficulties learning to read, included terms such as, reading blindness, word blindness, congenital word blindness, strephosymbolia, specific reading disability, specific developmental disability, unexpected reading failure, specific reading retardation, and garden-variety poor readers, and "dyslexia" (from the Greek *dys* and Latin *lexia*). Descriptions of difficulties learning arithmetic emerged later. For instance, Gerstmann (1957) described an "isolated disability to perform simple or complex arithmetical operations and an impairment of orientation in the sequence of numbers and their fractions" (Gerstmann syndrome). The term "dyscalculia" (from the Greek *dys* and

Latin *calculia*) was used by Cohn (1968) to refer to a "failure to recognize numbers or manipulate them in an advanced culture," and later by Kosc (1974), a Bratislavian, as "a disorder of the special abilities for mathematics without a simultaneous defect in general mental abilities." The first descriptions of *developmental* disorders of written language in the 1960s were based on case studies of isolated impairments of written expression, with other forms of verbal behavior remaining intact. Since these early descriptions involved visual–motor integration, the term *dysgraphia* (literally "difficulty with writing") was the first diagnostic label associated with disorders of written expression. Although it is sometimes used as a synonym for disorder of written expression, the term *dysgraphia* is used more correctly to describe difficulties with the mechanics of handwriting. By contrast, SLD in written expression emphasizes difficulties with accurate and fluent generation and composition of written text.

Terminology and conceptualizations of SLD differ widely among professionals from different disciplines or different countries; they also differ between professionals and the lay public. For instance, in the field of research, the term "dyslexia" is typically understood to refer primarily to problems in decoding single words and is associated with phonological deficits, but lay understanding differs and is broader (e.g., often incorporates concepts such as sensory, perceptual, motor, and emotional difficulties). Also, in Europe, the term "dyscalculia" is used to refer to developmental problems in basic numerical processing, such as automatic or implicit processing of quantities or numbers. By contrast, the term is used rarely in North America, which instead uses terms such as "mathematical disabilities" or "arithmetic difficulties" that implicate a broad array of problems with arithmetic facts, arithmetical procedures, and mathematical reasoning.

Nonetheless, early perspectives of Learning Disabilities shared some commonalities: (a) recognition that impaired academic skills (reading, arithmetic) occur in the context of average or even superior intellectual abilities; and (b) a focus on impairments in specific component skills (e.g., word reading, calculation) rather than viewing an academic domain as a unitary construct. Other important insights were that these disabilities were believed to be congenital, heritable, and manifest more commonly in males than females. However, whether or not learning disabilities could occur either as an isolated domain-specific deficit or be accompanied by deficits in other academic domains remained controversial, as it does today.

The advent of IQ tests in the early 1900s (e.g., Stanford-Binet) permitted direct evaluation of the observed difference in a person's academic skills and intellectual ability. In 1920, the neurologist Samuel Orton postulated a neurological basis for developmental reading difficulties, suggesting that delayed development of specialization of the left

hemisphere for language was potentially causal. However, it was not until the 1970s that an adequate knowledge base was established, which localized the difficulties at the single-word level and pinpointed the cause in most cases as language-based deficits in phonological processing. The 1940s, known as the era of the "brain-damaged child," incorporated a broader array of cognitive deficits into the category of Learning Disabilities, including perceptual, perceptual–motor, and attention difficulties. However, it was not until the 1960s that Learning Disabilities emerged as a formal category and the notion of a discrepancy between IQ and achievement prevailed as one of its defining features. The concept was broadened to include perceptual as well as linguistic processing deficits as underlying learning disabilities.

As previously mentioned, in the 1970s in the United States, the conceptualization of Learning Disabilities was consolidated in the federal definition (U.S. Office of Education, 1977) and in the definitions of professional organizations (e.g., National Joint Committee on Learning Disabilities, 1978), and achieved official federal status as an eligible category for direct services (Public Law 94–192). The notion of underlying deficits in cognitive processing as causal factors was instantiated in the conceptualization of Learning Disabilities and the condition included conditions such as perceptual disabilities brain injury, minimal brain dysfunction, dyslexia, and developmental dysphasia (thereby continuing its earlier conceptualizations), but excluded learning problems resulting from visual, hearing, or motor disabilities, mental retardation, emotional disturbance, or environmental, cultural, or economic disadvantage. Thus this prevailing legal definition restricts the concept of Learning Disabilities to a language-based disorder, and does not include non–language-based learning difficulties, such as dyscalculia, as defined by some researchers. In other words, speech and oral language problems are included in the federal category, but are coded separately as Communication Disorders in psychiatric (ICD, DSM) taxonomies. Thus, the federal/legal and medical categories of Learning Disabilities differ in the range of learning difficulties they encompass as well as in the imputed boundaries between disabilities/disorders.

Comparative Nosology

The first psychiatric/medical descriptions of SLD in reading or mathematics occurred in DSM-III (1980), with SLD in written expression added in DSM-III-R (1987), under the general category of "Academic Skills Disorder" on axis II. These learning difficulties were classified into three separate disorders: *developmental reading disorder, developmental arithmetic disorder,* and *developmental expressive writing disorder,* respectively. By contrast, DSM-IV-TR coded learning disorders on axis I, as does DSM-5. Both of these DSM versions dropped the "developmental"

descriptor that is used by other clinicians and researchers to distinguish between developmental versus acquired learning problems (e.g., as a result of traumatic brain injury, dementia, stroke). There was no compelling logic for the coding shift from axis II to axis I, but the shift may have served to render these impairing learning disorders more salient for diagnosticians, whose clinical practice focuses primarily on other types of mental health problems. The 2013 revision of DSM (DSM-5) no longer recognizes subtypes of SLD distinguished in previous DSM versions. Instead it recognizes an umbrella category of SLD and then uses specifiers for clinicians to use to detail more precisely the range of academic skills that are impaired *at the time of assessment*, thereby recognizing the possibility of changes in manifestation of SLD with development or educational instruction. Specifiers refer to three major academic domains that may be affected (i.e., reading, writing, mathematics) as well as the type of skills impaired. For instance, *SLD in reading* may manifest as difficulties in accurately decoding single words (a.k.a., dyslexia), fluency in reading text accurately, or reading comprehension; *SLD in written expression* may manifest as impairments in spelling words accurately, in accurate use of grammar and punctuation, or in the clarity or organization of written expression; *SLD in mathematics* may manifest as impairments in basic number sense, memorization or retrieval of arithmetic facts, accurate or fluent calculation, or accurate math reasoning.

The DSM-5 conceptualization of "SLD in reading," "SLD in written expression," and "SLD in mathematics"—although comparable with—is not identical to other terms used primarily in the fields of education and in educational and clinical psychology (e.g., "reading disabilities," "developmental dyslexia," "dyslexia," "developmental expressive writing disorder," "written language disabilities," "dysgraphia," "spelling disorder," "math disabilities," "specific disorder of arithmetic skills," "dyscalculia").

The DSM-5 nosology of SLD also differs from that of ICD-10, which differentiates even more specific subtypes of SLD than does any version of DSM. For instance, although ICD-10 distinguishes three subtypes that appear to be similar to those of DSM-IV-TR ("specific reading disorder," "specific spelling disorder," "specific disorder of arithmetic disorder," as well as "mixed disorder of scholastic skills"), they differ in their requirement for the restriction of impairment to that specific academic domain. For instance, the categorical subtype of Specific Reading Disorder excludes individuals who also manifest impairments in arithmetic; Specific Spelling Disorder excludes those who also manifest a reading disorder; individuals who manifest impairments in more than one academic domain are given the diagnosis of Mixed Disorder of Scholastic Skills.

Ongoing concerns about the DSM (and ICD) approach to nosology (e.g., Casey et al., 2013; Greenberg, 2012) have spurred a recent initiative by the U.S. National Institute of Mental Health to focus on alterations in the brain

and its substrates that might signal the location and source of human mental stresses or woes (RDoC: Research Domain Criteria Project; Cuthbert & Insel, 2013). However, this approach is still in the exploratory research phase and has yet to yield clinically relevant syndromes or traits.

EPIDEMIOLOGY

SLD is one of the most frequently diagnosed Neurodevelopmental Disorders. A recent epidemiological study suggested that the lifetime prevalence of SLD is around 10 percent, with the prevalence being greatest among children with special health care needs (28 percent) compared to typically developing children (5.4 percent). However, prevalence rates vary as a function of the definition, cut-offs, and instrumentation used. Moreover, most estimates are based on prevalence of a single manifestation of SLD (e.g., in reading or mathematics) without considering the co-occurrence of SLD across academic domains. Nonetheless, SLD in the domain of reading (particularly at the word level) is the most common and well-studied manifestation, affecting at least 90 percent of all individuals identified as having SLD. It is estimated that two of five children in special education have been placed because of reading difficulties. However, children showing a deficit in one academic domain (e.g., reading) are highly likely to show deficits in other domains as well (e.g., spelling, written expression, mathematics), either concurrently or emerging later, as demands for efficient and fluent use of basic academic skills increase.

To date there are two studies of prevalence of SLD based on DSM-5 criteria. One, conducted in a German school population, found that SLD across several academic domains (and subdomains, such as word reading, fluency, and comprehension) occurred as frequently as an isolated disorder in any one domain. The rates of multiple-domain SLD were high (particularly for co-occurrence of arithmetic and spelling difficulties) even when applying the stricter criteria of achievement scores of 1.25 to 1.5 standard deviations below the mean for age. No sex differences were observed when considering SLD across multiple academic domains, however, more girls manifest problems in arithmetic and more boys had spelling problems and combined reading and spelling problems compared to girls. The second study, conducted in Brazil, a low- to middle-income country, reported that almost 8 percent of the representative sample (about 1,600 second- to sixth-grade children from four geographic areas) manifest SLD across the academic domains (global impairment), 7.5 percent in reading, 6 percent in arithmetic, and about 5 percent in written expression. Moreover, SLD with global impairment was significantly associated with ADHD; SLD in arithmetic was associated with both ADHD and Anxiety Disorders.

Although more cross-cultural epidemiological studies of the prevalence of SLD based on full DSM-5 criteria are needed, the practical implications

of the existing studies are that assessment for SLD should include all three academic domains as well as the subskills. Clinicians are advised to specify the diagnosis of SLD by providing additional information about the learning domains that are affected at the time of assessment (i.e., implementing Specifiers for SLD listed in DSM-5, as described in the later subsection on diagnosis).

ETIOLOGY

Different levels of explanation must be considered in understanding causal influences on learning academic skills. Converging evidence from cognitive, neuroimaging, twin, family, behavioral, and molecular genetic studies indicates that SLD is a multifactorial neurodevelopmental disorder arising from the complex interplay of biological and environmental risk factors that shape the development of brain systems underlying academic learning. Currently, a scientific schism is discernible in etiological research on SLD, in terms of the existence of two contrasting hypotheses: A generalist versus specific risk hypothesis. According to the generalist genes hypothesis, individual differences in academic achievement across the school years from entry to graduation are heritable and attributable to a set of genes that underlie these heritable traits. In other words, the same genes underlie both learning abilities and learning disabilities across various academic domains. By contrast, the specific risk hypothesis presumes that discrete domains of learning disabilities exist (e.g., dyslexia, dyscalculia) with different etiology.

One major challenge to understanding the etiology of SLD is that the research describing its genetics, neurological substrate, cognitive and environmental factors have not only been based primarily on the specific risk hypothesis, focusing on the etiology of specific manifestations of SLD (i.e., in reading, mathematics, or written expression), but also this "disorder-specific" research has been conducted independently of one another, despite the common co-occurrence of various manifestations of SLD across the lifespan. Thus studies continue to focus on genetic, cognitive, and environmental factors associated with specific manifestations of SLD (e.g., on the clinical phenotypic categories of dyslexia or dyscalculia) rather than on the broader category of SLD, often aggregating findings across a wide range of ages, rather than testing for developmental changes in etiological factors. Another challenge is that neither the clinical category of SLD nor its previously delineated subtypes are clear-cut disorders, distinct from normal development. Rather SLD and its impact on subskills, such as word decoding, reading comprehension, spelling, number sense, arithmetic computation, represents the lower tails of normally distributed traits of these learning skills. However, most research have used a categorical approach in sample selection.

Accordingly, based on the research trends, the following sections

summarize findings on etiological factors underlying specific manifestations of SLD. However, readers should be aware of the emerging evidence to support the generalist risk hypothesis, and the likelihood that multiple and interacting genetic factors, brain differences, cognitive deficits, and environmental stressors underlie SLD, each of small effect size.

Genetic Factors

Behavioral Genetics. Family and twin studies indicate that most dimensions of cognition (general intelligence, learning to read, spell, do arithmetic) are moderately heritable: About 50 percent of variance in a given population is attributable to genetic influences, with increases as a function of increasing age. For example, heritability estimates of individual differences in reading skills increase from about 20 percent of variance at age 6 to about 80 percent at age 12. Increases in heritability estimates with age most likely reflect a reduction of environmental influences once children enter formal education and increasing correlations between genotype and environment as with increasing maturity and independence, individuals tend to select environments that fit their abilities (genotype). A recent meta-analysis, based on 5,330 monozygotic and 7,084 dizygotic twin pairs in unselected samples (i.e., typically developing), estimated heritability of academic achievement at 73 percent for reading words, 44 percent for spelling, 49 percent for reading comprehension, 57 percent for mathematics, and 645 for language. Moreover, a common genetic factor has been demonstrated among spelling, reading fluency, and reading comprehension, with genetic influences accounting for around 80 percent of the covariation between core academic skills, such as reading, spelling, and arithmetic, with many of the same genes affecting the learning of various academic subjects independent of intelligence.

Importantly, twin studies involving large samples with diverse abilities (e.g., including typical readers and those with dyslexia) reveal common genetic influences across the range of ability and disability, suggesting that a particular set of genes affects academic learning in both typically developing samples and those with SLD. Also, despite the paucity of behavioral genetic research on SLD in mathematics, family studies indicate that the familial prevalence of dyscalculia is about 10 times higher than expected by chance and that genetic factors contribute to the observed covariance between reading and mathematics deficits.

It is important to note that the preceding findings do not mean that all individuals with SLD inherited the condition since twin studies also reveal that both common and unique environmental influences play a role, as discussed below. However, from a clinical perspective, the high heritability estimates for SLD highlight the importance of probing family history in the assessment for SLD, and ensuring that parents and patients with SLD can

read and comprehend any written reports or information brochures provided to them.

Molecular Genetics. To date, most molecular genetic studies have focused on genes implicated in reading ability and disability, particularly those associated with word reading and spelling difficulties (a.k.a., dyslexia). Several candidate genes have been implicated (e.g., DYX1C1, DCDC2, ROBO1, KIAA0319, and the MRPL19/C2ORF3 locus), with some being replicated in many studies. One problem, is that most of the candidate risk SNPs identified so far for dyslexia have been located in noncoding regions of the genome, which poses a challenge for functional investigations aimed at gaining insights into the molecular mechanisms underlying dyslexia. This is because the biological role of much noncoding DNA is poorly understood, making it hard to design appropriate assays. However, DCDC2 is of interest because of its role in axon guidance and ciliary function: Cilia play a key role in coordinated neuronal migration and placement of interneurons and projection neurons. However, genome-wide association studies have failed to implicate these genes in SLD in reading (dyslexia). Instead, a recent genome-wide scan identified novel candidate genes, one of which (an intronic variant within RBFOX2) is highly expressed in the fetal brain with a role in CNS development. Rare variants and copy number variations have also been reported, but only apply to a minority of cases with SLD in reading. Moreover, recent studies report genetic interactions between two DCDC2 risk haplotypes (READ1 alleles) and the risk haplotype KIAA0319, which influence both reading and oral language, as well as intelligence, with both protective and deleterious alleles. Findings from these studies are consistent with the generalist genes hypothesis and also underscore the close association between Communication Disorders and SLD, which are conceptualized and coded as distinct disorders in DSM and ICD systems.

Molecular genetic studies of SLD in mathematics are sparse and findings inconsistent. For example, two genome-wide association studies failed to identify any significant associations and a third reported a significant genetic component underlying mathematical abilities but did not identify specific risk factors. Another genome-wide association study identified one marker (myosin-18B), that was not only associated with mathematical ability in two independent cohorts selected for dyslexia, but also was associated with brain alterations (reduced depth of right intraparietal sulcus that has been linked with numerical processing) in adults with low mathematical performance. However, this association was not replicated by another group of researchers using independent samples.

As noted above, genetic studies of SLD in written expression have focused on the manifestation of spelling difficulties. A literature search revealed no molecular genetic studies to date on other manifestations of

written expression difficulties.

Neurological Substrate

SLD in Reading. Most research has focused on the neurobiological underpinnings of specific categorical manifestations of SLD, particularly dyslexia. Structural and functional neuroimaging studies have revealed brain regions that are most consistently involved in reading words and which are altered in dyslexia. These regions are typically lateralized to the left hemisphere and include inferior frontal, superior and middle temporal, and temporoparietal regions. Experienced readers also involve an area of the left fusiform gyrus, known as the *visual word form area*, which is activated for orthographic (print) processing. This reading network typically develops with age as children gain automaticity in specific skills required for reading.

Meta-analytic studies have revealed patterns of structural and functional differences in children, adolescents, and adults with SLD in reading words (dyslexia). The most common difference is reduced activation in *left* temporal, parietal, and fusiform (visual word form area) regions. Network analysis has revealed abnormalities in several white matter tracts (e.g., arcuate fasciculus, superior longitudinal fasciculus, inferior fronto-occipital fasciculus, optic radiation, and callosal fibers). Variations in findings may reflect maturational and instructional effects, heterogeneity of SLD in reading, as well as methodological differences.

One challenge has been to ascertain whether the observed brain differences reflect underlying neurobiological etiology or the consequence of years of altered and dramatically reduced reading experience. One approach to dissociating the cause and consequence of dyslexia has been to compare youngsters with dyslexia to “ability-matched” groups (who are at the same reading level but chronologically much younger) as well as age-matched groups. Studies using this approach have found that children with dyslexia exhibit reduced left parietal and occipitotemporal activations compared to both ability- and age-matched groups, suggesting that these differences may be related to the cause of dyslexia, whereas left prefrontal activity tracked ability level. Another approach, in which prereading children at high risk of dyslexia (familial dyslexia) are followed longitudinally, has revealed altered neurophysiological responses to auditory stimuli (e.g., mismatched negativity) that shows potential as an early endophenotype of dyslexia. Some of these biomarkers enhance or even outperform behavioral measures in predicting long-term reading outcomes: But, at present, none of these biomarkers can be used to rule in or rule out the diagnosis of SLD in reading.

SLD in Mathematics. Evidence for the neurological base of mathematical cognition and SLD in mathematics is sparse in comparison

to that for reading and SLD in reading. However, converging evidence from structural and functional imaging studies indicate the *right* intraparietal sulcus as a major (but not sole) locus of brain differences and cognitive deficits in children and adults with SLD in mathematics. Alterations in other interconnected brain regions have also been observed, including posterior regions (bilateral intraparietal sulcus and fusiform gyrus) and prefrontal cortex (premotor cortex and middle frontal gyrus). Also, network analysis reveals multiple dysfunctional white matter circuits, suggesting a neurodevelopmental disconnection condition.

Cognitive Factors

According to the generalist genes hypothesis, shared genetic factors will lead to shared cognitive difficulties across the various manifestations of SLD (and other neurodevelopmental disorders). One domain-general cognitive deficit is executive functioning, which includes the control and coordination of short-term, working- and longer-term memory systems, encoding and retrieval systems, attention, and planning. Working memory, which refers to the ability to temporarily hold and manipulate auditory–verbal and visual–spatial information, while learning new information or problem solving, is believed to be impaired in most individuals with SLD. Thus, individuals with SLD have difficulties suppressing irrelevant information from working memory, as well as simultaneously holding and processing information, which manifests as difficulties updating what they know or have just learned. In other words, the more restricted working memory capacity associated with SLD makes it more difficult for these individuals to suppress irrelevant information and update new knowledge as they learn.

Working memory deficits will impair learning to read new words, which requires the ability to remember letter–sound combinations, their order, as well as the many special rules (e.g., “silent” letters, such as “g” in gnash or height, “e” at the end that changes the vowel sound but itself remains silent, as in “site”). Working memory deficits have an even greater impact on reading comprehension and writing: Reading comprehension requires updating information in working memory to monitor the meaning of each sentence and each paragraph, extract the main ideas, and construct mental images of the emerging situations and events; writing requires the ability to plan, organize, compose and edit, while keeping in mind the writing goals, content, audience, as well as spelling, vocabulary, and grammar. Poor working memory will also impair mathematical abilities, including the ability to retrieve and hold math facts for use in computation, follow arithmetic procedures in calculation, and in math reasoning. A recent meta-analysis indicates the domain-general nature of working memory deficits across the various manifestations and severity of SLD, but with some indication of more severe verbal working memory deficits in

individuals with SLD in reading and more severe deficits in visual–spatial working memory in those with SLD in mathematics.

Slow processing speed is another domain-general factor that has a negative impact on learning. For instance, rapid automatized naming, which refers to the ability to rapidly name familiar alphanumeric (letters, numbers) and nonalphanumeric (colors, objects) stimuli, is impaired in a substantial proportion of individuals with SLD. Although some research has shown distinct rapid naming profiles in students with SLD in reading versus SLD in mathematics (impairment in alphanumeric in the former and in nonalphanumeric stimuli in the later), a recent meta-analysis concluded that rapid naming of nonalphanumeric stimuli is impaired in both manifestations of SLD, whereas alphanumeric naming may be more specific to SLD in reading. However, rapid naming deficits are not unique to SLD: They are also shared with other neurodevelopmental disorders (e.g., ADHD, Autism Spectrum Disorders, Communication Disorders).

Language-based factors, although often linked in the older literature exclusively with SLD in reading (particularly its manifestation as "dyslexia"), are associated with all manifestations of SLD. For instance, oral language difficulties interfere with all aspects of reading, writing, and mathematics, although in the latter domain, the impact is primarily of those aspects of mathematics involving verbal processing. Impairment in a specific language-related skill, phonological processing, is widely recognized as the core cognitive deficit in SLD in word reading. However, some evidence suggests that early phonological deficits may have deleterious effects subsequently on learning arithmetic. For example, one study reported that over 50 percent of children with phonologically based reading difficulties but no apparent difficulties in learning basic arithmetic at age 5 years, manifest learning difficulties in mathematics as well as continued problems in reading at age 7.

An ongoing controversy in the field of SLD is whether SLD in Reading single words (dyslexia) is associated with a unique and potentially diagnostic set of cognitive deficits. However, the conclusions from a large-scale investigation of two population-based samples (US cross-sectional study; international longitudinal study) are that cognitive predictors believed to play a causal role in dyslexia (e.g., phonological processing, rapid automatized naming) are neither necessary nor sufficient for the diagnosis of SLD in reading. Moreover, the investigators concluded that the relationship between the cognitive measures and reading skill is probabilistic not deterministic or diagnostic. Also it remains unclear whether cognitive factors are a cause, consequence, or correlate of SLD. Hence, DSM-5 does not include cognitive processing deficits in the diagnostic criteria for SLD.

Environmental Factors

Twin studies indicate that shared and nonshared environmental factors contribute up to about 40 percent or 50 percent of the variance in academic skill development, and to the development of SLD, but the nature of these environmental factors remains underinvestigated. Candidate risk factors include low socioeconomic status, the prenatal environment, and the postnatal learning environment provides by parents.

Prenatal Environment. Prenatal exposure to toxic substances, such as alcohol, tobacco smoke, lead, and other environmental toxins have been linked to neurodevelopmental outcomes. Prenatal exposure to tobacco smoke remains high despite the overall worldwide reduction in smoking among adults: A recent study reveals that maternal smoking rates during pregnancy range from about 4 to 19 percent in European countries. Known risks of maternal smoking and alcohol intake during pregnancy include premature birth, low birth weight, as well as fetal alcohol syndrome for prenatal alcohol exposure. However, findings for the outcomes in terms of subsequent academic achievement in the offspring are inconsistent, perhaps due to methodological shortcomings of the studies in terms of exposure assessment, measurement of academic outcomes, and inadequate control or consideration of confounding factors. Nonetheless a recent review of 30 studies concluded that both prenatal exposure to the ingredients in tobacco smoke and alcohol can result in a spectrum of adverse developmental outcomes, ranging in severity from intellectual disorders to more subtle negative impact on academic achievement, memory, verbal abilities, and poor executive functioning.

More robust evidence comes from recent studies that postdate the above review. For example, a recent large-scale longitudinal study focusing specifically on reading outcomes of prenatal exposure to tobacco smoke, found that high prenatal exposure to nicotine tobacco smoke (>17 mg nicotine per day) was negatively associated with reading performance, particularly in decoding single words and spelling, with greatest effects on children with poorest phonological skills—suggesting a possible gene X environment effect. Another longitudinal epidemiological study conducted in the United Kingdom examined the health and academic trajectories of children born in the late 1950s—an era in which there was a high rate of maternal smoking during pregnancy (about 30 percent)—from the prenatal period through to age 16. The research revealed that low birth weight, prenatal exposure to tobacco smoke, and poor school-aged health are negatively related to academic achievement in reading and math, suggesting a cumulative effect of early adversities. Also, meta-analysis of the academic outcomes associated with children born very preterm (<32 weeks) or with very low birth weight (<1,500 g) reveals negative effects on IQ scores (by an average of 11 IQ points), and poor academic performance, with greater frequency of SLD in mathematics compared to SLD in reading.

About two-thirds of children born very prematurely or with very low birth weight have special educational needs, especially for SLD in mathematics.

Family Environment. Parents of children with SLD may well have SLD themselves and so may be less able to engage in "code-focused" or "meaning-focused" activities with their children who have SLD: These activities refer to parent–child engagement in rhyming and counting games, teaching letter names and sounds, letter–sound pairs, teaching numbers and how to count (code-focused); and reading stories aloud in an interactive way or playing board games that involve symbol recognitions and counting (meaning-focused). However, although parents of children with SLD report being more stressed than parent of children without SLD, there is no evidence of poorer family functioning in families in which someone has SLD. Moreover, although the home literacy and numeracy environment, including parental SLD predict children's growth in literacy and numeracy abilities in preschool and kindergarten years, by school age, genetic influences as well as basic skills knowledge (e.g., letter and number knowledge) tend to predominate.

Gene X Environment Interaction. The complex interplay between genetic and nongenetic influences has garnered relatively little empirical attention, given the general consensus that SLD arises from gene-by-environment interactions. However, some recent studies provide preliminary evidence for such interactions in SLD in single word reading (a.k.a., dyslexia). For example, one study reports gene X environment effects between adverse prenatal factors (e.g., maternal smoking during pregnancy, low birth weight, socioeconomic status) and a genetic marker associated with dyslexia (DYX1C1-1259C/G) in 165 nuclear families in which at least one member had dyslexia. Findings revealed that the interaction of common "C" allele with lower birth weight was associated with lower reading scores and the interaction between the rare "G" allele and maternal smoking during pregnancy was associated with poorer short-term and working memory scores.

The clinical implications of findings from current research on etiological factors of SLD include the importance of including in an assessment for SLD the family history of SLD, the individual's prenatal and birth history, and for children with suspected SLD, an assay of their home literacy and numeracy environment. Clinicians are advised to be alert to the increased risk for SLD associated with prematurity, low birth weight, and prenatal exposure to tobacco smoke, alcohol, or other drugs.

CLINICAL FEATURES, ASSOCIATED PROBLEMS, AND DIAGNOSIS

Clinical Features

The learning difficulties manifest as an array of observable and persistent behaviors or symptoms that change across the lifespan, many of which are summarized in Table 41–2. These include (but are not restricted to) poor word recognition and difficulties sounding out words, frequently guessing words, reading slowly and hesitantly; difficulty understanding what is read (may get the basic facts, but not the deeper meaning); marked difficulties spelling words correctly; poor sentence and paragraph structure in written work, with numerous punctuation and grammatical errors; difficulties understanding numbers and their relative magnitude and relationships; persistent use of primitive counting strategies (e.g., counting on fingers); difficulties recalling number facts and getting lost in the midst of arithmetic computation; and difficulties applying basic arithmetic concepts to solve quantitative problems.

These observable symptoms are not simply the consequence of an intellectual disorder, sensory deficits or neurological disorders, lack of proficiency in the language of instruction, inadequate educational instruction, or lack of opportunity to engage in full-time formal schooling.

Table 41–2.
Common Risk Factors, Signs and Symptoms of Specific Learning Disorder

Preschool

- Lack of interest in playing games with language sounds (e.g., repetition, rhyming)
- Trouble learning nursery rhymes
- Problems counting correctly
- Frequent use of baby talk and mispronunciation of words
- Failure to recognize letters in own name
- Trouble remembering names of letters, numbers, or days of the week

Kindergarten

- Unable to recognize and write letters, write own name, or use invented spelling
- Trouble breaking down spoken words into syllables (e.g., “cowboy” into *cow* and *boy*)
- Trouble recognizing words that rhyme (e.g., “cat,” “bat,” “hat”)
- Trouble connecting letters with their sounds (e.g., letter “b” makes the sound /b/)
- Unable to recognize phonemes (e.g., does not know which of the set of words {“dog,” “man,” “car”} starts with same sound as “cat”)
- Unable to count in sequence beyond “4” (e.g., “1,2,3,4…um 7,11”)

Primary Grades (Grades 1–3)

- Receives reports of “not doing well in school”
- Continues to have problems recognizing and manipulating phonemes
- Unable to read common one-syllable words, such as “mat” or “top”
- Problems in connecting sounds and letters (e.g., “big” for “got”)
- Unable to recognize common irregularly spelled words (e.g., “said,” “two”)
- Difficulty in sequencing numbers and letters
- Complains that reading is hard, and avoids or refuses to do it
- Difficulty learning simple math facts of 10
- Counting is often inaccurate above about 5 or 6 items
- Unable to identify the relevant numerical information or generate a math equation needed to solve a simple arithmetic word problem

Middle Grades (Grades 4–6)

- ► Mispronounces or skips parts of long, multisyllable words (e.g., says “conible” for “convertible” and “aminal” for “animal”)
- ► Confuses words that sound alike (e.g., “tornado” for “volcano”)
- ► Trouble remembering dates, names, telephone numbers
- ► Trouble reading small function words (e.g., “that,” “the,” “an,” “in”)

Middle Grades (Grades 4–6) Continued

- ► Gets the first part of word correct, then guesses (e.g., reads “clover” as “clock”)
- ► Poor comprehension with or without slow, effortful, and inaccurate reading
- ► Trouble completing homework or tests on time
- ► Fear of (refuses) to read aloud, avoids reading
- ► Terrible spelling; unable to recognize correct spelling from plausible alternatives (e.g., necessary/necessery; friend/frend)
- ► Continues to rely on finger counting for simple arithmetic
- ► Difficulties mastering basic math facts (e.g., cannot remember that 3 + 2 = 5)
- ► Shows high levels of anxiety around any activity requiring basic arithmetic

High School, College, Work

- ► Slow, effortful reading of single words and connected text
- ► Trouble pronouncing multisyllable words
- ► Frequent need to reread material to understand or get the main point
- ► Trouble making inferences from written text
- ► Avoids activities that demand reading (reading for pleasure, reading instructions)
- ► Terrible spelling; same word misspelled in different ways within one piece of work
- ► Excessive problems generating written text: short, immature, not representative of the student’s knowledge or creative ideas
- ► Avoids writing activities; written expression is very short, immature, and disorganized
- ► Great difficulties in using arithmetic in functional activities (e.g., calculating tip percentage, calculating amount of change required after a purchase, monitoring bank account)

General

- ► First-degree relative (parents, sibling) with current or history of SLD
- ► Slow learning to talk
- ► Anxious and embarrassed when asked to read or work out a problem in public (e.g., in class, workplace)
- ► Avoidance or refusal to read, write or do arithmetic
- ► Reluctance to go to school

Signs and Symptoms of SLD in Reading. The most common and earliest manifestation of SLD in reading, particularly in the English language, is inaccurate, effortful and slow reading at the word level (a.k.a., “dyslexia”). Common signs of this difficulty include reluctance and resistance in reading aloud; reads hesitantly, slowly, and inaccurately, often with visible effort and discomfort (e.g., sighing, squirming, becoming flushed); guesses based on the sound of the first letter; may recognize a few words by sight, but is unable to “sound out” unfamiliar words; or an inability to reconstruct the word having sounded out its component parts (e.g., sounds out /b/ /a/ /n/ /k/, but cannot reconstitute to “bank”); adolescents and adults may read accurately but so slowly and with such effort that reading is not functional for everyday use, and so may avoid reading.

Lack of automaticity in letter–sound mapping not only interferes with

reading at the word level, but will also make it difficult for the person to understand what has just been read at the level of the sentence, paragraph, or entire text. However, some individuals with SLD in reading at the word level may have very little difficulty in comprehending connected text (sentence, paragraph, entire text) if the need to decode words is bypassed by having the text read aloud at a reasonable pace (e.g., by using text–speech software).

These observable behaviors are due to an underlying difficulty in learning how the letters (graphemes) and speech sounds (phonemes) of one's language map on to one another and being able to do so automatically. The general consensus in the scientific community is that the problems with letter–sound conversion arise from an underlying deficit in processing phonemes—the elementary constituents of speech sounds.

Common features of SLD in reading comprehension include reading text aloud with a flat intonation and no modulation to acknowledge punctuation marks (e.g., no pauses at commas or periods; no change in pitch to represent a question mark or exclamation point); may read the text accurately and fluently—often too fast—but is unable to extract and use the inherent meaning (e.g., may not grasp the sequence of events, relationships among events or people, may not be able to draw inferences from text or process the deeper meaning of the text); tends to simply reread the text without using effective strategies before restarting to read (e.g., does not predict, scan for or review unfamiliar words), during reading (e.g., does not make notes or self-question to check one's understanding; does not create visual images of the scenarios), and after reading (e.g., does not check the accuracy of understanding; does not make a synopsis).

Over time, these challenges in reading comprehension tend to result in a reduction in or avoidance of reading print (e.g., newspapers, novels, nonfiction, lengthy texts regardless of format such as digital or paper). In turn, restrictions in the amount of reading will have a negative impact on vocabulary growth, so that the individual's range of vocabulary use and understanding is noticeably reduced.

Signs and Symptoms of SLD in Written Expression . One of the most common clinical features is noticeable and persistent poor spelling for age —often described as "atrocious spelling" (e.g., often adds, omits, or substitutes vowels and consonants; may spell phonetically but incorrectly, such as "pikchr" for "picture"). Sentences are often very short or incomplete and contain grammatical errors, omission of words, and little or no use of punctuation, as well as spelling errors. Also written paragraphs are often short, poorly constructed and organized (e.g., does use a starting sentence to introduce the topic; includes several topics in one paragraph without a logical order; does not maintain the correct verb tense, incorrect use of pronouns) and their intended meaning is unclear. A common

complaint from teachers (and parents) is of "procrastination": The youngster fails to start writing, writes very little, and with great effort and reluctance. In general, the person's written work does not accurately reflect that individual's knowledge, creativity, or depth of understanding of an issue: The person's spoken or visual presentations are often very much better than the written one. Also, essential written activities, such as note taking, pose major challenges. This is because note taking requires the physical act of writing or typing notes to occur simultaneously with a complex web of other cognitive processes involved in listening, comprehending, retaining key information while continuing to process incoming and new information, and summarizing key points in a legible and organized format.

Signs and Symptoms of SLD in Mathematics. One common sign that persists across the lifespan is the sole reliance on finger counting for simple addition and subtraction that is inappropriate for age. Note that typically, children start to switch from finger counting to more efficient calculation strategies, such as retrieval of math facts from memory, by about age 8 or 9 (e.g., see the interview of an accomplished author and Defence correspondent, describe his lifelong struggles with SLD in mathematics, including his use of finger counting under the boardroom table! http://www.mathematicalbrain.com/qt/paul.mov).

Other common features include a poor sense of numbers in terms of their magnitude and relationship; inaccurate counting and calculation; failure to attend to key mathematical symbols (e.g., symbols for addition, subtraction, multiplication, division); difficulty with identifying the previous or next item in a sequence (e.g., difficulty identifying the name of previous or next day, month, or number, despite being able to rote list those items); difficulty understanding the concept and use of "zero" or decimal points; and "gets lost" in the middle of multidigit calculation and may switch procedures or apply them incorrectly (e.g., in the problem 57 – 38, may start to subtract 7 from 8 instead of 8 from 7, then either subtracts 3 from 5 without considering the trading procedure, or may switch to adding 5 and 3). Many struggle with understanding the "language" of mathematics, especially in *word problems*, that are emphasized in current curricula across all grade levels. Word problems consist of two interwoven semiotic layers: A narrative about some event or sequence of events in the real world, and a latent web of mathematical relations typically expressed in words, and related by a question involving concepts of quantity such as "more," "more than," "less than," "altogether." Thus, many individuals with SLD in reading have great difficulty with the language-based demands of word problems. Moreover, older children, adolescents, and adults manifest difficulty understanding and using money, telling the time, using calculators, and remembering numerical PIN codes.

Math anxiety is another common feature that persists across the lifespan. Often described as "a feeling of tension, apprehension, or fear that interferes with math performance," it may manifest as sweaty palms, a racing pulse, breaking out in a cold sweat, or choking feeling in the back of the throat, when required to perform mathematics in everyday life (e.g., answer math questions in class; complete math homework, counting out money to buy something, splitting a restaurant bill or calculating the tip, entering specific numbers for cash withdrawal or deposit at an ATM machine).

General Signs and Symptoms of SLD. In school-aged children, the typical presenting problem concerns difficulty in learning at school ("not doing well at school—and we don't know why"). The puzzling picture for parents and teachers is that the child appears intelligent and often succeeds or even excels in other aspects of the curriculum (e.g., oral discussions, art, drama, sports) that do not depend on the core academic skills.

In addition to the domain-specific features summarized above, SLD imposes a great emotional toll. For example, other common features include embarrassment and low self-esteem; and extreme discomfort or even refusal when asked to read, write, or spell. Children and adolescents may be reluctant or refuse to go to school; may report being exposed to daily criticism, taunting, or bullying; they may look anxious, stressed, sad, depressed, or express despair or suicidal ideation.

Associated Problems

SLD is commonly, but not invariably, associated with a developmental history of problems in speech and language, attention, and motor skills. Moreover, SLD commonly, but not invariably, co-occurs with other neurodevelopmental (e.g., Communication Disorders, ADHD) and mental health disorders (e.g., Anxiety and Mood Disorders). These associated problems are clinically important because the diagnosis of SLD cannot be confirmed until the child has started school and been exposed to formal academic instruction, but Communication Disorders in particular are likely to be apparent and diagnosed earlier and so indicate a risk for later SLD. Likewise, a developmental and family history of SLD may indicate a risk for later anxiety and mood disorders in adolescence and adulthood.

Speech and Language Impairments. Oral language constitutes an essential foundation for abilities in written language. Thus it is not surprising that a developmental history of Communication Disorders (Speech Sound Disorder, Specific Language Impairments, alone or in combination) in preschool years is a common precursor of SLD, particularly for poor skills in reading words and reading comprehension, spelling, arithmetic fact retrieval, and calculation. (Speech sound disorder,

defined as problems in producing sounds on one's language accurately and intelligibly, whereas language impairments refer to problems in developing the structural aspects of language, such as grammar [syntax] and vocabulary [semantics]). For example, one longitudinal study of a community-based sample of children identified with pervasive speech/language disorders in kindergarten were found to have an estimated three- to sixfold greater risk for SLD (all manifestations alone or in combination) in young adulthood compared to typically developing youngsters. Moreover, according to a recent population-based study, the cumulative risk of SLD in written language by age 19 is about 64 percent in youngsters with speech/language impairments compared to 18 percent in those without.

Attention-Deficit/Hyperactivity Disorder. Epidemiological and clinical studies indicate that the overlap between SLD and ADHD is greater than expected by chance alone. For example, around 20 to 45 percent of children with ADHD also meet diagnostic criteria for SLD, but the strongest link appears to be between the inattention dimension of ADHD and SLD, particularly SLD in reading and writing. Inattention in the classroom (even at subclinical levels), as rated by teachers, is a robust predictor of concurrent and subsequent poor academic outcome, as well as a poor response to evidence-based academic intervention. Importantly, there is little robust evidence that medication treatment for ADHD (including stimulant and nonstimulant medications) have any substantial effect on SLD, despite its positive effects on the behavioral symptoms of ADHD and academic productivity (i.e., completing more arithmetic problems in a specified time frame). By contrast, preliminary evidence suggests that intensive educational intervention may improve (but not normalize) the behavioral problems. The clinical implications of these findings are that assessment for SLD should include assessment for ADHD and vice versa: Youngsters with SLD who have coexisting inattention or ADHD require multimodal intervention including intensive educational intervention, since medication for ADHD will not address the SLD and intervention for SLD will be less effective if the inattention remains unaddressed.

Emotional Problems, Anxiety and Mood Disorders, Suicide Risk. Individuals with SLD are highly likely to manifest concurrent social, emotional, and mental health problems. For example, they are only too aware of their academic difficulties, leading to low self-esteem and these negative self-perceptions are very difficult to change and contribute to their problems in social and academic functioning. Avoidance of or reluctance to engage in activities requiring the academic skills is common in children, adolescents, and adults. Also, recent meta-analyses confirm that the

majority of youngsters with SLD (about 70 percent) have higher rates of anxiety and depressive symptoms compared to non-SLD peers. Episodes of severe anxiety or anxiety disorders (including somatic complaints or panic attacks) are common across the lifespan and accompany both the circumscribed and broader expression of learning difficulties. Moreover, SLD is associated with increased risk for suicidal ideation and suicide attempts in children, adolescents, and adults. The co-occurrence of internalizing symptoms and disorders is clinically significant whereas SLD is treated primarily in the field of education, these internalizing symptoms often remain unaddressed. Thus psychiatrists and other mental health practitioners play an important role in identifying, monitoring, and treating the concurrent internalizing problems.

Criteria for Diagnosis

There are no biomarkers to date for SLD. Thus, SLD is a clinical diagnosis, which is best made by an experienced clinician, based on a comprehensive assessment with careful clinical synthesis of information from multiple sources, including developmental, medical, educational, and family history; school reports; teacher observations, including response to classroom instruction and other intervention provided at school; as well as direct standardized psychometric assessment of the individual's academic skills—including reading (accuracy, fluency, comprehension), spelling, written expression, and arithmetic (number sense, counting, algorithmic computation). The DSM-5 diagnostic criteria for SLD are presented in Table 41–3.

This diagnostic category is indicated by using one overarching term "Specific Learning Disorder," which incorporates difficulties in learning academic skills, such as reading, writing, and mathematics, which were classified as separate disorders in previous versions of DSM and in ICD-10. The generic term reflects the evidence from behavior–genetic studies show that learning difficulties in the various academic domains share genetic variance, and from clinical and epidemiological studies that individuals showing a deficit in one academic domain frequently show deficits in other domains as well, although deficits isolated to one academic domain may occur. Using this overarching term, diagnosticians can then specify the details of the areas of learning difficulty and their severity, as manifest at the time of assessment.

Table 41–3.
DSM-5 Diagnostic Criteria for Specific Learning Disorder, with ICD-9-CM Codes (ICD-10 codes are shown Parenthetically)

Diagnostic Criteria
A. Difficulties learning and using academic skills, as indicated by the presence of at least one of the

following symptoms that have persisted for at least 6 months, despite the provision of intervention that targets those difficulties:

1. Inaccurate or slow and effortful word reading (e.g., reads single words aloud incorrectly or slowly and hesitantly, frequently guesses words, difficulties sounding out words)
2. Difficulty understanding the meaning of what is read (e.g., may read text accurately but not understand the sequence, relationships, inferences, or deeper meanings of what is read)
3. Difficulties with spelling (e.g., may add, omit, or substitute vowels or consonants)
4. Difficulties with written expression (e.g., makes multiple grammatical or punctuation errors within sentences; employs poor paragraph organization; written expression of ideas lacks clarity)
5. Difficulties mastering number sense, number facts or calculation (e.g., poor understanding of numbers, their magnitude, and relationships; counts on fingers to add single-digit numbers instead of recalling the math fact like peers; gets lost in the midst of arithmetic computation and may switch procedures)
6. Difficulties with mathematical reasoning (e.g., severe difficulty applying mathematical concepts, facts, or procedures to solve quantitative problems)

B. The affected academic skills are substantially and quantifiably below those expected for the individual's chronological age, and cause significant interference with academic or occupational performance, or activities of daily living, as confirmed by individually administered standardized achievement measures and comprehensive clinical assessment. For individuals aged 17 years and older, a documented history of impairing learning difficulties may be substituted for the standardized assessment

C. The learning difficulties begin during school-age years but may not become fully manifest until the demands for those affected academic skills exceed the individual's limited capacities (e.g., as in timed tests, reading or writing lengthy complex reports for a tight deadline, excessively heavy academic loads)

D. The learning difficulties are not better accounted for by Intellectual Disabilities, uncorrected visual or auditory acuity, other mental or neurological disorders, psychosocial adversity, lack of proficiency in the language of academic instruction, or inadequate educational instruction.

Note: The four diagnostic criteria are to be met based on a clinical synthesis of the individual's history (developmental, medical, family, educational), school reports, and psychoeducational assessment.

Coding Note: Specify all academic domains and subskills that are impaired. When more than one domain is impaired, each one should be coded individually according to the following specifiers.

Specify if:

315.00 (F81.0) With impairment in reading:

Word reading accuracy

Reading rate or fluency

Reading comprehension

Note: *Dyslexia* is an alternative term used to refer to a pattern of learning difficulties characterized by problems with accurate or fluent word recognition, poor decoding, and poor spelling abilities. If dyslexia is used to specify this particular pattern of difficulties, it is important also to specify any additional difficulties that are present, such as difficulties with reading comprehension or math reasoning.

315.2 (F81.81) With impairment in written expression:

Spelling accuracy

Grammar and punctuation accuracy

Clarity or organization of written expression

315.1 (F81.2) With impairment in mathematics:

Number sense

Memorization of arithmetic facts

Accurate or fluent calculation

Accurate math reasoning

Note: *Dyscalculia* is an alternative term used to refer to a pattern of difficulties characterized by problems processing numerical information, learning arithmetic facts, and performing accurate or

fluent calculations. If dyscalculia is used to specify this particular pattern of mathematic difficulties, it is important to also specify any additional difficulties that are present, such as difficulties with math reasoning or word reasoning accuracy.

Specify current severity:

Mild: Some difficulties learning skills in one or two academic domains, but may of mild enough severity that the individual may be able to compensate or function well when provided with appropriate accommodations or support services, especially during the school years.

Moderate: Marked difficulties learning skills in one or more academic domains, so that the individual is unlikely to become proficient without some intervals of intensive and specialized teaching during the school years. Some accommodations or supportive services at least part of the day at school, workplace, or at home may be needed to complete activities accurately and efficiently.

Severe: Severe difficulties learning skills, affecting several academic domains, so that the individual is unlikely to learn those skills without ongoing intensive individualized and specialized teaching for most of the school years. Even with an array of appropriate accommodations or services at home, school, or workplace, the individual may not be able to complete all activities efficiently.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

The manifestation of learning difficulties *at the time of assessment* is coded by adding *Specifiers* to the generic diagnosis of SLD to provide more precise information about the domains that are affected. Specifiers are deficits in reading, deficits in writing, deficits in mathematics, each of which can be further specified by detailed descriptions (e.g., SLD with impairment in reading [currently manifests as problems in word reading accuracy, reading fluency, and comprehension] and impairment in mathematics [currently manifests as problems with memorization of math facts, as well as in accuracy of calculation]). Use of the terms "dyslexia" and "dyscalculia" to refer to problems in learning to read or do arithmetic, respectively, is permitted according to clinician and patient preference, but detailed coding of the manifestation of SLD should still be completed. Severity ratings (mild, moderate, severe), which reflect the dimensional nature of SLD, permit further description by considering the range of learning difficulties and impairments, as well as the likely need for accommodations and support to best promote ongoing learning. DSM-5 specifies a set of four diagnostic criteria, all of which must be met to confirm a clinical diagnosis of SLD.

Criterion A. This criterion lists a range of observable descriptive behaviors or symptoms (Criterion A1 to A6, Table 41–3) which may be observed, probed during the clinical assessment, and confirmed by evidence from the school reports, teacher observations or ratings, previous psychoeducational assessment reports, and current performance scores on standardized achievement tests. At least one of these symptoms must be present. Moreover, Criterion A also requires evidence that the symptoms have persisted for at least 6 months despite the provision of intervention that targets each of the affected academic domains. For children, the term

"*persistent*" is confirmed by evidence of restricted progress in mastering the affected academic skill, over at least 6 months, despite the provision of intervention that targets that affected skill. For older adolescents and adults, persistent difficulties refer to ongoing difficulties with literacy or numeracy skills that first manifest during the school years, as confirmed by school reports, evaluated portfolios of past and current work, or previous assessments.

The term "*intervention*" refers to extra help at school or home, small group or individualized instruction, using any instructional approaches available at the school that focus on the affected skill. DSM-5 does not require the provision of intensive evidence-based (research-based) intervention—as specified by a current educational model in the United States (Response-to-Intervention [RTI]). RTI is a multitiered approach to academic and behavioral intervention designed to provide early, systematic, and appropriately intensive assistance to children who are at risk or already underperforming academically or behaviorally, compared to grade-appropriate standards. To date the RTI approach, while laudable in its goals, not only lacks a strong evidence base but also is not consistently available within the United States and is rarely available outside of the United States.

Criterion B. This comprises two components: (i) academic achievement that is well below average and (ii) impairment resulting from the learning difficulties. In terms of low academic achievement, the DSM-5 applies the "average-person standard" that has been applied in disability discrimination laws for years. An individual's scores on standardized, individually administered, culturally sensitive academic achievement tests are compared to the population mean for chronological age, not to IQ or to achievement scores of subsets of the population, such as peers attending university, medical, or law school. Examples of commonly used standard achievement tests are shown in Table 41–4.

DSM-5 does not specify what cut-off scores to use for determining low academic achievement within the diagnostic criterion itself, but provides guidelines within the text. One suggested guideline for below-average academic performance is to require evidence that scores on standardized academic achievement tests fall at least 1.5 standard deviations below the general population mean for age: This cut-off score is equivalent to a standard score >78 or >7th percentile. Standardized tests vary in the number and composition of items relevant to age, as well as in the particular aspect of reading, arithmetic, spelling, written expression. Moreover, it is important to recall that standardized testing occurs in rather artificial situations (quiet space, minimal noise or other distraction, use of a standardized procedure that often precludes repetition of a question by the examiner), performance may be influenced by factors other

than academic skill level, such as motivation and effort, familiarity with the testing situation, anxiety, hunger, and sleep status, etc. Thus, greater diagnostic certainty is established by evidence of low scores on more than one test or subtest.

A second guideline suggests that more lenient cut-off scores may be used (1.0 to 1.5 standard deviations below mean for age) when supported by other converging evidence of learning difficulties, such as academic history, other test scores, school reports, teacher observations, etc.

For individuals aged 17 and older, a documented history of long-standing and impairing learning difficulties (based on self-reported history of difficulties learning to read, spell, do arithmetic or written work; previous standardized testing; school reports; previous systematic clinical assessments for SLD) may be substituted for the standardized assessment. The goal of this substitution is to avoid the time and expense needed for unnecessary repeated standardized assessments, based on the evidence that childhood-onset SLD typically persists into adulthood.

Identification of SLD in individuals in postsecondary education poses challenges. DSM-5 does not preclude the possibility that an individual with high IQ, giftedness, or attending postsecondary education may have SLD. Evidence required for greater diagnostic certainty in this situation includes previous or current reports (from school, significant others), self-reported history of marked learning difficulties, manifestation of self-report of an extraordinary effort required or dependence on the provision of support and accommodations needed to attain/maintain performance; previous evaluated portfolios of academic work; previous clinical assessments indicating SLD; family history of SLD; or current evidence of low academic scores (e.g., below the population mean for age). Extensive testing is rarely required: Research evidence indicates that tests of fluent word reading, reading comprehension, word spelling, number sense, and fluency of arithmetic computation, suffice.

In the absence of available culturally appropriate and normed standardized academic achievement tests (as occurs in many low- to middle-income countries, or non–English-speaking countries), converging evidence of markedly low academic performance must be derived from careful clinical review of scores on locally available tests, school reports, and clinical assessment. (See Tables 41–1 to 41–4 for a description of commonly used tests.)

Moreover, Criterion B emphasizes impairment, in requiring that the learning difficulties cause significant impairment in academic, occupational, social functioning, or in activities in daily living that requires the academic skills. Thus, low academic scores per se are not sufficient to meet the specifications of Criterion B: Evidence of associated impairment is required. This component is in line with the standard of US “Americans with Disabilities Act (ADA)” for determining a disability.

Criterion C. This third criterion requires evidence that the symptoms of SLD manifest in the school years—most reliably interpreted as during the elementary (primary) school years. Symptoms of SLD may manifest in preschool years, but a diagnosis cannot be confirmed until the individual has received formal instruction in the core academic skills. However, a diagnosis of SLD can be made for the first time in adolescence of adulthood. This is because the manifest learning difficulties may have been overlooked, not assessed, or attributed solely to concurrent emotional behavioral difficulties. Also, some individuals (particularly those with higher adaptive and intellectual functioning, or in highly supportive environments) may have coped with the demands for the core academic skills during the school years, albeit by putting forth extraordinary effort, or by using compensatory skills, that no longer suffice in later years, when high demands on accurate and fluent academic skills (e.g., reading or writing lengthy complex reports for a tight deadline, excessively heavy academic or occupational loads) start to exceed the individual's limited capacities and compensatory mechanisms. In such cases, greater diagnostic certainty is established by converging evidence from comprehensive clinical assessment, and the ability to rule out loss of previously acquired academic skills as a result of a major medical event or condition.

Table 41–4.
Examples of Commonly Used Standardized Tests of Reading, Written Expression, Arithmetic and Related Skills

Name of Standardized Test	Reading Skills
Comprehensive Test of Phonological Processing (CTOPP); standardized. Ages: 5 years to young adulthood	Basic reading skills (e.g., phonological awareness, phonological coding, working memory, rapid naming)
Woodcock Johnson-III Achievement battery	Letter-Word Identification, Word Attack and Reading Fluency, Reading Comprehension
Test of Word Reading Efficiency (TOWRE); Ages: 5 years to young adulthood	Single-word reading efficiency (e.g., speedy and accurate naming of single words and pseudowords)
Gray Oral Reading Test; Nelson–Denny Reading Test; Gray Silent Reading Test	Reading fluency and comprehension (e.g., rate of reading connected text for meaning)
	Written Expression Skills
Wechsler Individual Achievement test (WIAT-II) Ages: 5–19 years	Spelling, written expression
Test of Written Language (TOWL; 3rd edition) Ages: 7.5–17.9 years	1. Elicited writing, including vocabulary, spelling, capitalization and punctuation, logical sentences, sentence combining 2. Spontaneous writing combining
Test of Early Written Language (TEWL; 2nd edition)	Basic, global, and contextual writing

Ages: 3–11 years	
Test of Written Spelling (TOWS; 4th edition)	Spelling of phonetically regular and nonregular words
Ages 6–18 years	
Woodcock–Johnson Psychoeducational Battery	Dictation, proofing, writing fluency, writing samples
Ages: 5 years–adult	
	Mathematics Skills
Woodcock–Johnson Achievement Battery-III	Math fluency, calculation, quantitative concepts, applied problems
Test of Early Mathematical Abilities —3rd edition Age: 3–0 to 8–11	Numbering skills, number-comparison facility, numeral literacy, mastery of number facts, calculation skills, and understanding of concepts
Test of Mathematical Abilities Age: 8–0 to 18–11	Mathematical symbols and concepts; computation; mathematics in everyday life; word problems; attitude toward math
Teacher Academic Attainment Scale (TAAS)	Subjective rating scale of student academic achievement (7 items)
Math Anxiety Questionnaire	Math anxiety
Abbreviated Math Anxiety Scale	1. Learning anxiety: anxiety related to learning math 2. Testing: anxiety related to being tested in math
Subjective Numeracy Scale	Subjective rating of one's health numeracy
Test of Functional Health Literacy in Adults (s-TOFHLA)	Objective measure of health literacy
Rapid Estimate of Adult/Adolescent Literacy in Medicine [REALM])	Objective measure of health literacy

Criterion D. The fourth criterion refers to a range of exclusionary conditions. These include intellectual disability, other major mental or neurological conditions, uncorrected visual or auditory acuity, severe psychosocial adversity, lack of proficiency in the language of instruction (for adults, lack of proficiency in the language of the workplace or social community), lack of or inadequate educational instruction, as may occur for children excluded from education for various reasons (e.g., because they are girls; living in war zones, no schools, no teachers, etc.). These exclusionary criteria are evaluated by means of the comprehensive clinical assessment.

As of 2012, UNICEF reports that about 58 million children are excluded from primary school education and 63 million from lower secondary school education (http://www.uis.unesco.org/Education/Pages/out-of-school-children.aspx). Certainly it makes no sense to make a diagnosis of SLD in immigrant children who previously have been excluded from education. The issue of excluding from the SLD category those children whose low academic scores may be due to instructional factors (inadequate instruction; limited opportunity to learn) is important, but has been challenging for clinicians to operationalize. One component of Criterion A attempts to address this challenge, by requiring evidence that symptoms of learning difficulties have persisted for at least 6 months despite the

provision of intervention that targets the affected academic skills. As noted previously, the RTI approach was not adopted by DSM-5 due to insufficient evidence and lack of universal availability of RTI approaches. However, in many countries, the quality of teacher training has improved over recent years (especially in teaching children to read), affording some assurance that a student's limited progress is not due to inadequate instruction and so SLD is a possibility.

Identification of SLD in individuals whose first language (mother tongue) differs from the language of instruction is challenging. Greater diagnostic certainty is accorded by evidence of the following: The individual has a history of language delay or disability in the native language; the individual has had difficulty developing literacy or numeracy skills in the native language (assuming adequate instruction in the native language); there is a family history of SLD in parents, siblings, or other close relatives, the individual is not making adequate progress over 6 months compared to other nonnative speakers, despite the provision of targeted intervention.

Does Information from Tests of Intellectual Capacity or Cognitive Processing Assist in the Diagnosis of SLD or Guide Educational Treatment?

Intelligence testing (and IQ profiling using various subtests) with the requirement for a significant discrepancy between intellectual and academic achievement has been the primary and often sole criterion for the diagnosis of SLD in previous versions of the DSM, continues to be used internationally to determine eligibility for special education services and accommodations, and figures strongly in the understanding of SLD by various stakeholder groups (e.g., teachers, psychologists, parents and parent organizations, lobby groups and researchers). Moreover, this "discrepancy" model has become ingrained in clinical training and practice, despite the several decades of research evidence against its validity for diagnosis or utility in guiding educational instruction. For example, discrepancy scores are unreliable, and children with learning difficulties who do and do not meet the IQ-achievement discrepancy do not differ in symptomatology, associated cognitive deficits, or response to intervention. In other words, the purported differences between those with discrepant and nondiscrepant IQ-achievement scores have not proven to be clinically or educationally meaningful. Thus, a measure of intelligence is not required for a diagnosis of SLD (unless there are concerns about the individual's intellectual ability and adaptive behavior).

Accordingly, based on the available scientific evidence, DSM-5 eliminated the requirement for a significant discrepancy between the IQ score and an achievement score for the diagnosis of SLD. The DSM-5 does not prohibit the inclusion of intellectual assessment in the diagnostic assessment of SLD: Occasionally information from an IQ test might be

helpful in changing a teacher or parent's overall view of the child and their misunderstanding of the child's intellectual ability. However, when formulating a diagnosis of SLD, clinicians should also keep in mind the DSM-5 changes to the conceptualization and criteria for Intellectual Disability: Specifically, poor adaptive and intellectual functioning, with the latter suggested by IQ score less than 70 ± 5 standard points. This means that individuals with an IQ score of greater than 70 ± 5 standard points and adequate adaptive functioning, could potentially—but not necessarily—meet criteria for SLD, since the differential diagnosis of Intellectual Disability can be ruled out.

In recent years, testing of cognitive processes (e.g., phonological processing, rapid automatized naming, auditory and visual processing, processing speed) and executive functions (e.g., working memory, inhibitory control, cognitive flexibility, sustained attention) has been incorporated increasingly in assessment for SLD, in addition to IQ testing. The goal of such testing is to document patterns of cognitive strengths and weaknesses to inform the development of differentiated and more effective educational instruction and intervention. However, it is unclear whether the relationship between SLD and these processes or executive functions is causal or correlational. There is little or no robust evidence that training these cognitive processes and executive functions will boost academic learning particularly in students with SLD. Also, recent meta-analyses conclude that effect sizes of the magnitude of the relationship between children's cognitive characteristics and response to intervention were small and unable to improve clinical prediction from measures of academic skills per se. Moreover, there is little evidence the cognitive profiles are valuable for planning intervention (i.e., little evidence for Aptitude X Treatment interactions). Thus current recommendations are for assessment and intervention for SLD to focus on educational rather than cognitive factors—a position that is understandably threatening to clinicians and other professionals trained in the psychometric tradition and who might lack expertise in curricular and pedagogical practices. Hence the DSM-5 criterion B does require the use of individually administered standardized tests of academic achievement in reading, writing, and arithmetic, where possible.

DIFFERENTIAL DIAGNOSIS

SLD can be distinguished from normal variation (individual differences) in learning academic skills due to other internal and external factors by the severity of the learning difficulties, the specificity of the difficulties in the presence of intact intellectual abilities, and their persistence in the presence of adequate educational opportunity and receipt of the same instruction as the peer group, targeted home or school-based intervention, competency in the language of instruction (even when different from one's

mother tongue).

A diagnosis of SLD is not usually made before the child is about 7 years old because evidence of failure to learn to read, write, or do arithmetic is required. Difficulties in learning the foundational academic skills may occur for a variety of reasons (e.g., lack of adequate exposure to print, limited education) and are associated with numerous mental health conditions, including mental retardation, developmental coordination disorder, communication disorders, ADHD, conduct disorder, hearing or visual problems, and prematurity. The most common distinguishing feature of SLD in reading is a persistent impairment in phonological processing and spelling; that of SLD in mathematics is a persistent impairment in number facts. In a child with visual or hearing deficits, the learning problems must be in excess of the difficulty typically associated with these sensory disorders. In a child with learning difficulties in the presence of low intellectual abilities (i.e., IQ score of 70 ± 5), it is important to assess adaptive behavior to be able to rule out Intellectual Disabilities.

SLD is distinct from Intellectual Disabilities: Individuals diagnosed with intellectual disability have reduced cognitive capacity that has a global impact on learning and adaptive functioning in everyday life. By contrast, individuals with a diagnosis of SLD have at least average intellectual capacity but exhibit significant difficulties in one or more *academic* areas, while coping well or even excelling in other areas of academic, sporting, or artistic achievement. Moreover, SLD is distinguished from "learning difficulties," which is a broader term that encompasses difficulties learning as a consequence of a range of causes, including intellectual disabilities, physical or sensory deficits (e.g., hearing or visual impairments), emotional or behavioral difficulties, inadequate environmental experiences, or appropriate educational opportunities.

SLD in Reading, Mathematics, or Written Expression is frequently associated with ADHD, and a concurrent diagnosis is not only permissible, but is also essential. An explicit statement should be made concerning the presence of both clinical conditions to ensure that the child can access the relevant and effective interventions. Moreover, even if the individual does not meet full criteria for ADHD but does manifest persistent symptoms of inattention (particularly in the school setting), clinicians should report these problems, given the increased risk conferred for poor response to intervention if inattention is not targeted in the treatment plan.

It might be possible to determine inadequate instruction as the primary explanation for reading problems through detailed interview with parents and teachers. Probes for the number of schools attended, absenteeism (voluntary due to illness or vacation; required due to class or school suspensions and expulsions), and type of reading instruction provided at the school(s) in the primary grades may be informative. However, differential diagnosis may be complicated by parental history or current

manifestation of SLD, given that parents with SLD would be less likely to be able to provide the necessary early reading or numerical experiences for the child.

COURSE AND PROGNOSIS

Converging evidence indicates that SLD is persistent and does not remit with increasing age in the absence of effective intervention. SLD tends to be a persistent and chronic condition that remains with the individual his or her entire life, albeit with developmental changes in its manifestation.

Onset, recognition, and diagnosis of SLD usually occurs during the first few years of elementary school when children are expected to learn to read, write, spell, and do basic arithmetic. Precursor signs (e.g., delayed onset of talking; difficulties with counting or rhyming; problems with fine motor skills required for writing) commonly occur in preschool years before the start of formal schooling. Moreover, preschoolers with poor phonological short-term memory (STM) skills and specific language impairment (SLI) are highly likely to meet diagnostic criteria for SLD in reading at school age. However, referral and formal diagnosis rarely occurs until a couple of years later in second grade, when the children have failed to respond to formal classroom instruction in reading.

SLD in elementary school-aged children (grades 1 to 3) typically manifests first as marked difficulty learning to connect letters with the sounds they make (particularly in English-speaking children), to recognize and read common one-syllable words (e.g., mat, top) or common but irregularly spelled words (e.g., said, two). Reading aloud is slow, inaccurate, and effortful. Some children may also (or instead) show problems in learning to count, remember math facts, and perform simple single-digit arithmetic. Recent evidence suggested that it is possible to estimate an individual child's risk for SLD in reading at school age from the following factors measured in preschool years (age 3.5 years): Familial history of reading problems, poor letter and object naming speeds, and poor phonological awareness.

In the middle grades (grades 4 to 6), children with SLD may not only continue to show problems with word reading and basic number sense, but also may manifest problems with spelling, writing, and understanding word problems in mathematics. Others may gradually be able to sound out words, but their reading is slow and effortful, they have difficulty understanding what they have just read, and also problems start to manifest in their written work (e.g., poor spelling, poor punctuation, words left out of sentences, short disconnected sentences, illegible and disorganized writing). Some youngsters continue to count on their fingers since they cannot remember math facts. It is not uncommon, even at this young age for children to dread going to school and to try and avoid reading, writing, or arithmetic.

By adolescence, some youngsters may show improved word recognition and the ability to sound out words, but reading remains slow and effortful, often with poor comprehension, and so they cannot keep up with the increasing demands for reading complex material. They often need to reread material to get even the main point and have great difficulty making inferences from written text. Written work is markedly impoverished: numerous spelling errors, illegible writing, with poor sentence and paragraph structure. Those who showed problems in basic number sense at a younger age, may use long and inefficient strategies to work out even simple arithmetic, such as "7 × 8" (e.g., draw 8 groups of 7 items, and then counting all of the items). Ongoing dread of and failure in school may result in high-school dropout, anxiety, and depression.

Longitudinal investigation of a 1990 US cohort of children, reveals that individuals with SLD in word reading (a.k.a., dyslexia) show an achievement gap in reading that is discernible upon school entry in grade 1 (about age 6) and that persists throughout childhood and adolescence. It remains unknown whether early reading intervention (i.e., kindergarten, grade 1) will eliminate the observed persistent difference in growth trajectories in reading. Likewise, a longitudinal investigation of a large sample of kindergarten children recruited in 1998 to 1999 and followed through 8th grade, revealed that children with SLD in mathematics enter kindergarten with lower skill levels and the achievement gap persists through grade 8.

Young adults with persistent problems in both reading accuracy and fluency appear to develop compensatory mechanisms to cope with reading, such as greater reliance on memory-based rather than analytic strategies to assist in word identification. By adulthood, many individuals try to hide the fact that they cannot read, spell, or do arithmetic by avoiding both leisure and work-related activities that require these skills, or use alternative approaches to access print (e.g., text-to-speech or speech-to-text software, audiobooks, audiovisual media).

The clinical course of SLD is often complicated further by comorbidity with other mental health problems, such as ADHD or conduct disorder, as well as with low self-esteem, anxiety, depression, and increased risk for suicidal ideation and suicide attempts. The causal relationships between SLD and the various comorbid disorders are unknown and longitudinal epidemiological studies suggest that the relationships are complex.

On a more positive note, a recent study reported that adults with SLD in reading (a.k.a., "dyslexia") are overrepresented among entrepreneurs and manifest higher levels of visual problem-solving skills compared to peers with average reading abilities. This subgroup of individuals attributed their success to an admixture of high self-motivation, educational accommodations, and a supportive social network, particularly parents.

Juan P, a 20-year-old Hispanic college student, presented at the primary care clinic in his home city, with complaints of increasing problems sleeping, inability to concentrate or manage his college schedule, failing some courses, and feeling "down" most days. Clinical history indicated a normal birth, slowness in learning to talk, childhood problems with concentration, hyperactivity and angry outbursts, as well as struggles learning to read, at school. Family history was positive for ADHD (older brother) and SLD (father, now deceased, was reported to have had "dyslexia" and dropped out of school early). He had been diagnosed with ADHD, Predominantly Combined Type, at age 9, but did not meet criteria for SLD, since the discrepancy between his intellectual capacity and reading achievement was deemed not to be "substantial" as required by DSM-IV and his local school board.

During the clinical interview, JP reported that since starting college a year ago, he had been unable to remain focused while reading or listening to lectures, felt overwhelmed, and was unable to complete written assignments on time. He complained of feeling restless, agitated, exhausted, and often in a "blue mood." He denied any substance use or suicidal thoughts. He stopped taking medication for ADHD when he was aged 12, and had not sought registration with Student Disability Services at his college, for help with his ADHD, because he thought he could get through college by himself. Current psychoeducational assessment confirmed that JP's reading fluency and comprehension, spelling and written expression, were substantially and quantifiably below that expected for age and was causing significant impairment.

Diagnostic formulation: *Attention-Deficit/Hyperactivity Disorder*, with predominantly inattentive presentation, of mild to moderate severity; *Specific Learning Disorder*, with impairment in reading (fluency, comprehension) and written expression (spelling, organization of written expression), all currently of moderate severity; depressive symptoms but did not meet criteria for a mood disorder.

TREATMENT

Treatment focuses almost exclusively on academic intervention and rarely includes mental health services despite the fact that students with SLD experience higher levels of inattention (and ADHD), anxiety, and depression than their non-SLD peers. Moreover, treatment of SLD is typically carried out in education settings or specialized psychology clinics, and so is not the primary responsibility of psychiatrists or primary care providers.

The primary role of medical health care providers is that of coordinator, which, as a first step, includes the provision of psychoeducation about the disorder and its longer-term implications. Thus, physicians need to be able to explain what Specific Learning Disorder is and refer parents (or the adult patient) to organizations, websites, or reading materials that will help (see the References for some useful websites). Also, parents may need help in understanding their child's needs and rights under the Individuals with Disabilities Education Act, including possible need for accommodations. Given that Specific Learning Disorder tends to run in families, physicians may need to modify how any numeric health information is communicated (prevalence rates, risk ratios, dosage, etc.), as well as to check that the individual or caregivers are able to comprehend any relevant printed or website material.

A second critical step is to establish a monitoring plan with the patient and family, to permit early detection and management of commonly associated mental health problems, such as ADHDs, anxiety, and

depression, as well as to ascertain whether recommended accommodations have been put in place along with intensive, individualized, explicit instruction. Moreover, the need for transitional planning is essential to help individuals with SLD through major developmental and life transitions (e.g., transition from elementary to secondary school; from secondary school to the workforce or postsecondary education; marriage, parenthood, etc.) and facilitate access to services.

One of the major problems is that the standard instruction in regular classrooms is too little, too general, and nonsystematic. For example, one large-scale study found that children receiving special education over a 1-year period advanced in reading by only 0.04 standard deviations per year, which means that it would take at least 8 years to raise these students' scores from the 5h to the 9th percentile. By contrast, students with SLD in reading who received 1 to 2 hours daily of intense and focused intervention for a total of 35 to 70 hours of instruction showed dramatic gains in performance on standardized measures of both decoding and comprehension, which were maintained for several years after intervention. The instruction programs were very explicit and visible, conducted in small, interactive groups of children matched for reading abilities, which promoted engaged and repeated reading of connected text presented in a way that controlled task difficulty, and included instruction in basic reading abilities (e.g., teaching phoneme-to-grapheme correspondences), as well as the application of metacognitive strategies for decoding unfamiliar words and comprehending connected text. Of note, preliminary evidence suggests that intense and focused instruction may in fact alter the aberrant brain activation profiles observed in children with SLD. Thus, clinicians need to be aware that recommendations for placement in special education may not be sufficient to improve reading or arithmetic skills. Rather, they need to advocate for the provision of intense, focused, and explicit instruction in each of the affected academic domains.

Also, given the importance of literacy and numeracy skills in health, clinicians are advised to screen for and monitor health literacy and numeracy in all patients (and parents of pediatric patients), regardless of their referral problem. Low literacy or numeracy cannot be inferred reliably on the basis of the patient's education level, intelligence, or other observable behaviors. Reliable short screening forms can be used to assess health literacy and numeracy in adolescents and adults and to guide health communication.

References

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

Ashkenazi S, Black JM, Abrams DA, Hoeft F, Menon V. Neurobiological underpinnings of math and reading learning disabilities. *J Learn Disabil*. 2013;46(6):549–569.

Bishop DV. The interface between genetics and psychology: lessons from developmental dyslexia. *Proc Biol Sci*. 2015;282(1806):20143139.

*Butterworth B, Kovas Y. Understanding neurocognitive developmental disorders can improve education for all. *Science*. 2013;340(6130):300–305.

*Callaway E. Dyscalculia: number games. *Nature*. 2013;493(7431):150–153.

Cho K, Frijters JC, Zhang H, Miller LL, Gruen JR. Prenatal exposure to nicotine and impaired reading performance. *J Pediatr*. 2013;162(4):713–718.

de Zeeuw EL, de Geus EJC, Boomsma DI. Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement. *Trends Neurosci Educ*. 2015;4(3):69–76.

Donker M, Kroesbergen E, Slot E, Van Viersen S, De Bree E. Alphanumeric and non-alphanumeric rapid automatized naming in children with reading and/or spelling difficulties and mathematical difficulties. *Learn Indiv Differ*. 2016;47:80–87.

*Ferrer E, Shaywitz BA, Holahan JM, Marchione KE, Michaels R, Shaywitz SE. Achievement gap in reading is present as early as first grade and persists through adolescence. *J Pediatr*. 2015;167(5):1121–1125.

Fortes IS, Paula CS, Oliveira MC, Bordin IA, de Jesus Mari J, Rohde LA. A cross-sectional study to assess the prevalence of DSM-5 specific learning disorders in representative school samples from the second to sixth grade in Brazil. *Eur Child Adolesc Psychiatry*. 2016;25(2):195–207.

*Hulme C, Snowling MJ. The interface between spoken and written language: developmental disorders. *Philos Trans R Soc Lond B Biol Sci*. 2013;369(1634):20120395.

Katusic SK, Colligan RC, Weaver AL, Barbaresi WJ. The forgotten learning disability: epidemiology of written-language disorder in a population-based birth cohort (1976–1982), Rochester, Minnesota. *Pediatrics*. 2009;123(5):1306–1313.

*Kaufmann L, von Aster M. The diagnosis and management of dyscalculia. *Dtsch Ärztebl Int*. 2012;109(45):767–777.

Kere J. The molecular genetics and neurobiology of developmental dyslexia as model of a complex phenotype. *Biochem Biophys Res Commun*. 2014;452(2):236–243.

Klassen RM, Tze VM, Hannok W. Internalizing problems of adults with learning disabilities: a meta-analysis. *J Learn Disabil*. 2013;46(4):317–327.

Kucian K, von Aster M. Developmental dyscalculia. *Eur J Pediatr*. 2015;174(1):1–13.

Locke R, Scallan S, Mann R, Alexander G. Clinicians with dyslexia: a systematic review of effects and strategies. *Clin Teach*. 2015;12(6):394–398.

Lyytinen H, Erskine J, Hämäläinen J, Torppa M, Ronimus M. Dyslexia—early identification and prevention: highlights from the Jyväskylä longitudinal study of dyslexia. *Curr Dev Disord Rep*. 2015;2(4):330–338.

Mascheretti S, Bureau A, Battaglia M, et al. An assessment of gene-by-environment interactions in developmental dyslexia-related phenotypes. *Genes Brain Behav*. 2013;12(1):47–55.

Moll K, Kunze S, Neuhoff N, Bruder J, Schulte-Körne G. Specific learning disorder: prevalence and gender differences. *PLoS One*. 2014;9(7):e103537.

Nelson JM, Harwood H. Learning disabilities and anxiety: a meta-analysis. *J Learn Disabil*. 2011;44(1):3–17.

Nelson JM, Harwood HR. A meta-analysis of parent and teacher reports of depression among

students with learning disabilities: evidence for the importance of multi-informant assessment. *Psychol Sch*. 2011;48(4):371–384.

Niklas F, Schneider W. Casting the die before the die is cast: the importance of the home numeracy environment for preschool children. *Eur J Psychol Educ*. 2014;29(3):327–345.

Peng P, Fuchs D. A meta-analysis of working memory deficits in children with learning difficulties: is there a difference between verbal domain and numerical domain? *J Learn Disabil*. 2016;49(1):3–20.

Pennington BF, Santerre-Lemmon L, Rosenberg J, et al. Individual prediction of dyslexia by single versus multiple deficit models. *J Abnorm Psychol*. 2012;121(1):212–224.

*Peterson RL, Pennington BF. Developmental dyslexia. *Annu Rev Clin Psychol*. 2015;11:283–307.

Plomin R, Kovas Y. Generalist genes and learning disabilities. *Psychol Bull*. 2005;131(4):592–617.

Polanska K., Jurewicz J, Hanke W. Smoking and alcohol drinking during pregnancy as the risk factors for poor child neurodevelopment—A review of epidemiological studies. *Int J Occup Med Environ Health*. 2015;28(3):419–443.

Rimfeld K, Kovas Y, Dale PS, Plomin R. Pleiotropy across academic subjects at the end of compulsory education. *Sci Rep*. 2015;5:11713.

Sexton CC, Gelhorn HL, Bell JA, Classi PM. The co-occurrence of reading disorder and ADHD: epidemiology, treatment, psychosocial impact, and economic burden. *J Learn Disabil*. 2012;45(6):538–564.

Simms V, Cragg L, Gilmore C, Marlow N, Johnson S. Mathematics difficulties in children born very preterm: current research and future directions. *Arch Dis Child Fetal Neonatal Ed*. 2013;98(5):F457–F463.

Smedberg J, Lupattelli A, Mårdby AC, Nordeng H. Characteristics of women who continue smoking during pregnancy: a cross-sectional study of pregnant women and new mothers in 15 European countries. *BMC Pregnancy Childbirth*. 2014;14(1):213.

Snowling MJ, Melby-Lervåg M. Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychol Bull*. 2016;142(5):498–545.

Stoeckel RE, Colligan RC, Barbaresi WJ, Weaver AL, Killian JM, Katusic SK. Early speech-language impairment and risk for written language disorder: A population-based study. *J Dev Behav Pediatr*. 2013;34(1):38–44.

Swanson E Vaughn S. Learning disabilities: academic and mental health needs. In: Holt MK, Grills AE, eds. *Critical Issues in School-based Mental Health: Evidence-Based Research, Practice, and Interventions*. New York: Routledge; 2015.

Tannock R. Specific learning disabilities in DSM-5: are the changes for better or worse? *Int J Res Learn Disabil*. 2013;1(2):2–30.

Tannock R. Provision of evidence-based intervention is not part of the DSM-5 diagnostic criteria for specific learning disorder. *Eur Child Adolesc Psychiatry*. 2016;25(2):209–210.

[1]Reading was invented approximately 5400 years ago; symbolic arithmetic is more recent since the Arabic notation was not available even a thousand year ago.

42

Developmental Coordination Disorder

CAROLY SUE PATAKI, M.D., AND WENDY G. MITCHELL, M.D.

Developmental coordination disorder (DCD) is a neuromotor disability emerging in the early developmental period which can affect both fine and gross motor skills, interfering with daily activities. Children with DCD have motor performance which is typically slower, less accurate, and more variable than peers of the same age. DCD is estimated to occur in 5 to 6 percent of children aged 5 to 11 years, and 1.8 percent of children 7 years of age exhibit severe DCD. DCD affects boys more often than girls, most recent population studies report ratios of male:female of about 2:1. Children exhibiting DCD often show impairment in gross motor skills such as jumping, hopping, and running, fine motor skills including handwriting, tying their shoelaces, or using utensils, and balance. Motor coordination disorder is comorbid up to 50 percent of the time with attention deficit/hyperactivity disorder (ADHD). Other disorders that may co-occur with DCD are communication disorders and specific learning disorders. Delayed motor milestones may be the first sign of DCD, in the absence of intellectual disability. Without intervention, DCD often continues to cause impairment throughout childhood and adolescence, leading to significant consequences physically, educationally, socially, and psychiatrically.

DEFINITION

The essential feature of DCD according to DSM-5 is impaired motor skills leading to deficits in coordination compared to children of the same chronological age and intellectual level, interfering with activities of daily living, social activities, and school productivity. Key symptoms are marked delays in achieving motor milestones, such as walking, crawling, and sitting; constantly dropping objects; inability to catch or throw a ball, and poor handwriting. Children with DCD perform motor coordination tasks slower and less accurately than those peers of the same chronological age and intellectual capacity. Specific manifestations of DCD are detailed in Table 42–1.

HISTORY

In the early 1900s, childhood motor clumsiness was termed motor *deficiency syndrome*. This syndrome was characterized by awkward

voluntary actions, excessive tendon reflexes, mild hypertonicity, and neurological overflow movement. The term *clumsy child syndrome* first appeared in the literature in the 1930s to describe a set of symptoms associated with clumsy motor behaviors of unknown etiology, not associated with a known neurological disorder. The observation was made that many children with motor coordination difficulties also exhibited ADHD, communication disorders, and specific learning disorders. During the 1940s, remedial programs were established in an attempt to improve academic learning deficits through treatment of coordination difficulties observed in children.

By the 1960s, the *clumsy child syndrome* was officially distinguished from other categories of learning disorders. From this period to the mid-1980s, significant investigations were conducted into perceptual–sensory causes for motor impairments, as well as social and emotional problems associated with such impairments. Children with clumsy behavior were noted to have deficits, rather than solely delayed fine and gross motor coordination skills.

COMPARATIVE NOSOLOGY

Many terms have historically been applied to the description of *clumsiness* in a child. Terms used in the past that overlap with a description of DCD include congenital maladroitness, choreiform syndrome, psychomotor syndrome, developmental apraxia, developmental dyspraxia, congenital clumsiness, developmental agnosia, and developmental dyspraxia–dysgnosia. It is only appropriate to make a diagnosis of DCD after any suspicion of a neurological disorder is ruled out. In the DSM-5, DCD is not diagnosed if symptoms are attributable to intellectual disability or a neurological condition such as cerebral palsy, muscular dystrophy, or a degenerative disorder. Although, currently, this syndrome is still occasionally referred to as clumsy child syndrome, DCD and specific developmental disorder of motor function are more appropriate terms.

Table 42–1.
Manifestations of Developmental Coordination Disorder

Gross motor manifestations
Preschool age
Delays in reaching motor milestones, such as sitting, crawling, and walking
Balance problems: Falling, getting bruised frequently, and poor toddling
Abnormal gait
Knocking over objects, bumping into things, and destructiveness
Primary-school age
Difficulty with riding bikes, skipping, hopping, running, jumping, and doing somersaults
Awkward or abnormal gait
Older

Poor at sports, throwing, catching, kicking, and hitting a ball
Fine motor manifestations
Preschool age
Difficulty learning dressing skills (tying, fastening, zipping, and buttoning)
Difficulty learning feeding skills (handling knife, fork, or spoon)
Primary-school age
Difficulty assembling jigsaw pieces, using scissors, building with blocks, drawing, or tracing
Older
Difficulty with grooming (putting on makeup, blow-drying hair, and doing nails)
Messy or illegible writing
Difficulty using hand tools, sewing, and playing piano

DCD has been included as a diagnosis in DSM-III-R and appears in essentially the same form in DSM-5. The tenth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) also includes the disorder, but the slightly more formalized definition requires an intelligence quotient of at least 70 and performance at least two standard deviations below the mean on a standardized test of motor coordination. In 1994, an international Consensus meeting, held in London, Ontario agreed on the term DCD to describe children with a heterogeneous set of motor impairments that caused impairment in performing motor tasks compared to same age peers.

Although DSM-5 and ICD-10 view the disorder as a unitary undifferentiated syndrome, several investigators have suggested that there may actually be several distinct subtypes. Supporting that view is the individual variation in symptomatology, only some of which is clearly age related. One recent attempt to reclassify DCD proposed subtypes, such as ataxic syndrome (marked by unsteadiness and mild tremors), hypotonic syndrome (characterized by below-average resting muscle tone), tension syndrome (manifested by constant and excessively high levels of muscle tone), dyspraxic syndrome (identified by the inability to link a series of submovements necessary to execute smooth motor movements), manual–graphic syndromes (two separate manual syndromes, one marked by inadequate dexterity but adequate ability to reproduce written symbols and the other characterized by adequate dexterity but inability to reproduce written symbols accurately using fine motor movements), visual–perceptual clumsiness (ranging from ocular–motor problems involving inefficient movement of eye muscles to severe retinal deterioration), and mixed syndromes (a combination of symptoms already presented under the aegis of the previous six classifications). It is always important to keep in mind that significant ataxia, hypotonia, or any asymmetry between left and right side muscle function may be an indicator of a neurological disorder and requires a neurological assessment.

EPIDEMIOLOGY

The prevalence of DCD has been estimated to be approximately 5 to 6 percent among school-aged children. DCD has been noted in Europe, Asia, Africa, Australia, and North America. Clinical studies of children with DCD have historically found greater rates in boys than girls, of up to 7:1. More recent population studies have found less disparity, and report rates of approximately 2:1, male to female.

ETIOLOGY

A meta-analysis of performance deficits in DCD by Wilson and colleagues (2013) found that several clusters of deficits in learning motor skills emerged, suggesting that brain networks including parietocerebellar and cerebellar regions are involved. The clusters of motor skills dysfunction include "internal modeling of action" (predictive control of motor skills), rhythmic coordination and timing of motor movements, executive function, and dynamic control of posture and gait. The result of insufficient internal models, due to deficits in neural pathways in DCD leads to a lack of fluency and efficiency in coordinated movement. That is, the sequence of a complex motor movement in children with DCD is not properly or efficiently ordered, leading to awkward motions. Poor ability to coordinate a sequence of motor movements in DCD may be attributable to dysfunction in the corticocerebellar brain region. In terms of executive function deficits, impaired attention, working memory, and inhibition may contribute to poor action planning, and exacerbate overall poor motor coordination. The exact etiology of DCD is not yet clear, however, it is likely to include heritable factors, with potential prenatal and prenatal developmental contributions. There is much heterogeneity in the expression of DCD. Motor control is dependent on numerous skills, including perception of objects in space, regulation of muscle activity, and balance. Children with DCD appear to be less proficient and less consistent in regulating muscle activity in heterogeneous ways. Some children with DCD are poor in manual skills, while exhibiting proficiency in balance, and for others, the opposite is true.

Developmental Coordination Disorder and ADHD

Epidemiologic surveys have documented the comorbid occurrence of DCD in 30 to 50 percent of children diagnosed with ADHD. A recent study of children comorbid for ADHD and DCD in the Netherlands by Fliers et al., in 2007 confirmed previous research demonstrating a consistent increased risk of each of these disorders based on the presence of the other one. This study also investigated the association between the various types of motor coordination problems and inattentive versus hyperactive or impulsive subtypes of ADHD. In this study, data suggested that inattentive rather than hyperactive or impulsive symptoms were related to higher measures

of motor coordination symptoms on the Developmental Coordination Disorder Questionnaire (DCD-Q) that was filled out by parents. This finding lends support for hypotheses that deficits in information processing are contributing problems underlying both ADHD and DCD. Although parents report that motor coordination problems often diminish with age, half of adolescents comorbid for ADHD and DCD continue to be impaired by coordination problems. One possible commonality between DCD and ADHD is a deficit in cerebellar activity. The coexistence of DCD and ADHD may reflect a shared heritable disturbance of the dopamine pathway. These two disorders may also share dysfunction in the basal ganglia and the prefrontal cortex.

Developmental Coordination Disorder and Premature Birth

Evidence indicates biological and cognitive-developmental correlates of DCD. The biological–organic factors associated with the disorder include various birth and perinatal problems, such as low birth weight, prematurity, hypoxia, and neonatal malnutrition. A study by Marlow and colleagues investigated motor and "executive" cognitive function in 241 children who were born severely prematurely, at less than or equal to 25 complete weeks of gestation. These children were assessed at a median age of 6 years and compared with 160 classmates born at full term. Children underwent assessment of cognitive and neuromotor skills, including the Kaufman Assessment Battery for Children (K-ABC) and the NEPSY (a teacher-based assessment of academic achievement). Motor function was assessed by clinicians using the following items from the Movement Assessment Battery for Children (MABC): Heel-toe walking (15 steps), 1-leg balance (more than 20 seconds), and time to post 12 coins. Results from this study indicate that compared to classmates, children born severely prematurely and in the absence of cerebral palsy exhibited high prevalence of impairment in gross motor function at early school age, in addition to impaired visual–spatial, perceptual–motor attention, and executive function. In each of the above impairments in motor skills or executive functions, at least half of the deficits were not accounted for by limitation of the overall cognitive score. Thus, motor and executive function disturbances in extremely prematurely born children are not due to compromised intellectual function but more specifically associated with perceptual deficits. Children who were born extremely prematurely were observed to need more time for tasks compared to full-term classmates, and they were noted to have greater difficulty inhibiting associated movements and maintaining secure postural basis during activities. As many as 7 percent of extremely prematurely born children were unable to balance on one leg for 20 seconds (compared to 1 percent of classmates). This study showed that extremely prematurely born children did not differ significantly from classmates on observed off-task motor behavior (17

percent vs. 25 percent), mirror movements (32 percent vs. 46 percent), and overflow movements (45 percent vs. 69 percent). Some investigators have postulated a continuum of reproductive casualty. There are also suggestions of incomplete cerebral dominance or lateralization, neurochemical abnormalities in the brain, and structural lesions in the parietal lobes.

Developmental Coordination Disorder and Psychiatric Comorbidities

DCD frequently occurs concurrently with speech–language disorders, and specific learning disorders, including reading disorders and dyslexia. A study showed that more than half of children with severe dyslexia exhibited motor coordination impairments severe enough to warrant intervention. Reading disorders, dyslexia, and speech and language disorders have been hypothesized to be associated with cerebellar involvement as has DCD. One longitudinal study of children with language impairments revealed a continuing correlation between peg-moving performance and language performance. Another study found that fine motor skills, response speed, and upper limb speed were associated with language disturbances, whereas gross motor coordination and visual–motor control were not. Furthermore, evidence indicates that articulation problems are related to motor impairment, although the motor impairment is not necessarily dyspraxia.

DIAGNOSIS AND CLINICAL FEATURES

In 2012, The European Academy of Childhood Disability (EACD) published an interdisciplinary clinical guideline for the assessment, diagnosis, and treatment of children with DCD. The essential feature of DCD is poor skill in daily motor activities for a given child's age and intellectual level. Impairments may be observed in gross motor activities, such as running, kicking, or catching a ball, as well as in fine motor coordination, resulting in poor ability to write letters neatly, to drink from a cup, or to use a knife and fork well. Children with DCD exhibit a heterogeneous pattern of impairments and thus function more competently in some motor task areas than others.

Assessment Instruments for Developmental Coordination Disorder

The Developmental Coordination Disorder Questionnaire. The DCD-Q is a 17-item parent questionnaire developed in Canada and validated to identify motor coordination problems in children from 8 years to 14½ years by Wilson et al. (2000) validated in children 4 to 8 years of age. For each item, parents rate the degree of coordination in their child compared to others of the same age on a 5-point scale. By adding the scores for each item, a total DCD-Q score can be calculated. In the study validating this instruments for young children, a score of less than or equal to 53 (15th

percentile) was used as the criterion for DCD. The 17 items on the DCD-Q are:

1. Throw ball
2. Catch ball
3. Hit ball
4. Jump over
5. Run and stop
6. Plan activity
7. Writing speed
8. Writing legibly
9. Pencil pressure
0. Cutting
1. Team sports
2. Avoids sports
3. Rides bicycle
4. Learns skills
5. Clumsy
6. Awkward
7. Fatigues easily

The Movement Assessment Battery for Children. The The MABC-second edition, was introduced by Barnett, Henderson, and Sugden in 2007, and provides an indication of gross and fine motor skills for children 4 years to 16 years of age. The MABC has two parts: a checklist and a performance test. On the MABC checklist, teachers rate the motor performance of children on 48 items reflecting everyday activities in the lives of children from 4 to 12 years. The checklist is divided into 4 sections with 12 items in each section. On the performance test, there are four age-related item sets, each consisting of eight items that measure manual dexterity, ball skills, and static and dynamic balance. The checklist is a reliable instrument within sections and overall. Since the nature and severity of problems in DCD is very broad, an instrument meant to screen children with DCD must be able to measure a wide variety of motor skills.

Bruininks–Oseretsky Test of Motor Proficiency for Identifying Children with Motor Impairment. The Bruininks–Oseretsky Test of Motor Proficiency (BOTMP) assesses motor skills in children from 4½ to 14½ years old and provides a comprehensive index of motor proficiency as well as individual measures of gross and fine motor skills. In addition to the long form (LF), there is also a short form (SF), which consists of 14 items from the long form that have been validated. A recent study of the BOTMP in the assessment of preschoolers investigated the correlation between the long form and the short form of the instrument. Although the correlation between the two forms were high, short form total scores were found to be

significantly higher than long form scores, identifying less children as having motor impairment. Therefore, although the long form has been validated for young children, the short form is questionable in its ability to identify motor impairment in very young children.

Additional tests include the Frostig Movement Skills Test Battery, the Slosson Drawing Coordination Test, the Southern California Sensory Integration Test, and the Stott Test of Motor Impairment, which provide a comparative measure of a child's abilities.

Symptoms of DCD may be broken down into the following categories: dyspraxia, synkinesia, hypotonia, hypertonia, tremors, impersistence, and asymmetries. *Dyspraxia* describes a child's difficulty in exhibiting sequenced, coordinated motor movements when presented with an oral request. Movements identified as *synkinesia* are best described as unintentional muscle movements, also termed *motor overflow*. Mirror movements, such as finger twitching on the opposite hand when the child is asked to perform a finger opposition task or facial grimaces when the child is asked to make hand movements, are examples of synkinetic motor actions.

Two categories that describe muscle tone abnormalities are *hypotonia* and *hypertonia*. Hypotonia can be observed in all parts of the body and is characterized by a flaccid or sleepy quality in the child's facial expression and general muscle tone. Children with this condition may exhibit poor posture and obesity. Children with hypotonia often appear to be floppy and weak, but, on actual muscle strength testing, they may have normal muscle strength. *Hypertonia* describes generally high levels of muscle tone in a child. For example, a child with hypertonic movement cannot finish drawing a line at a given point, instead overshooting the intended stopping point. Another example is a child who tends to throw a ball too hard and inaccurately at short distances.

The movement category of *tremors* is marked by irregular unsteadiness in motor movements observed in tasks requiring walking, drawing, or both. For example, legibly formed letters might reveal, on closer inspection, rapid tremors in the actual lines themselves. Unsteadiness might be exhibited in the child's gait as high-amplitude fluctuations in leg muscle movement.

The category of *impersistence* refers to the child's inability to sustain and maintain various body postures for reasonable periods of time. Asking a child to stick out his or her tongue is an example of a task used often to investigate impersistent motor movements. In this example, the clinician would keep track of how long the child could maintain this tongue position.

The final movement category, *asymmetries*, describes motor behaviors that affect only one side of the body. This category includes unusual flexions or muscle weakness, or both, on one side of the body or general muscle weakness in the limbs on one side. These asymmetrical muscle

movements can be observed by requiring a child to move laterally in a rapid manner ("like a basketball player on defense, not crossing your feet"). If the child executes the movement more fluidly on one side than the other, it might indicate an asymmetry. A child who is observed to have asymmetrical motor movements is likely to be referred for a magnetic resonance imaging (MRI) study to rule out an organic cause for this symptom. If the neurological work-up is negative, then this symptom may be considered as part of a DCD.

Specific manifestations vary across individuals; thus, there is no typically clumsy child. In some instances, manifestations may be quite specific; one case report describes a child who could sew and do jigsaw puzzles without difficulty but could not write neatly.

The most common way to identify DCD is through the observation of delays in achieving motor milestones (e.g., sitting, standing, crawling, and walking). Gross motor coordination manifestations may be seen in difficulties with activities, such as sitting, walking, running, balancing, throwing, and kicking. Fine motor coordination manifestations may be identified through difficulties in performing tasks, such as using utensils, tying laces, buttoning or zipping clothing, drawing, and, especially, printing or writing. During motor tasks, children with DCD may exhibit overflow movements in unsupported limbs, mirror movements, slight abnormalities of reflex, or poor ability to make movements smooth.

Some manifestations of DCD are age dependent. Young children show symptoms, such as delayed development, in early motor milestones and delayed acquisition of self-help skills, such as self-dressing (i.e., tying ties, fastening snaps, zipping zippers, buttoning, and unbuttoning) or self-feeding (handling a knife, fork, or spoon). Preschool-aged children may also have balance problems, awkward gait, and a history of frequently bumping into objects, falling down, and getting bruised. Problems with balance, hopping, jumping, catching and bouncing a ball, and tracing appear to be among the more common symptoms of school-aged children with DCD. A common sign of DCD among school-aged children is avoidance of sports activities because of poor performance or teasing by classmates.

Table 42–2.
DSM-5 Diagnostic Criteria for Developmental Coordination Disorder

A. The acquisition and execution of coordinated motor skills is substantially below that expected given the individual's chronological age and opportunity for skill learning and use. Difficulties are manifested as clumsiness (e.g., dropping or bumping into objects) as well as slowness and inaccuracy of performance of motor skills (e.g., catching an object, using scissors or cutlery, handwriting, riding a bike, or participating in sports).
B. The motor skills deficit in Criterion A significantly and persistently interferes with activities of daily living appropriate to chronological age (e.g., self-care and self-maintenance) and impacts

academic/school productivity, prevocational and vocational activities, leisure, and play.

C. Onset of symptoms is in the early developmental period.

D. The motor skills deficits are not better explained by intellectual disability (intellectual developmental disorder or visual impairment and are not attributable to a neurological condition affecting movement (e.g., cerebral palsy, muscular dystrophy, degenerative disorder).

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

In addition, children with DCD frequently have comorbid communication and learning difficulties. Among the more common of these difficulties are expressive language disorder, mixed receptive–expressive language disorder, phonological disorder, reading disorder, disorder of written expression, and mathematics disorder.

Secondary complications for children with DCD often involve rejection from peers who are critical of their poor abilities in sports activities and other tasks. Over time, children who are chronically "picked last" for sports teams and who are unable to keep up physically with their peers may become demoralized or depressed. The DSM-5 diagnostic criteria for DCD are listed in Table 42–2. The ICD-10 diagnostic criteria for specific developmental disorder of motor function are presented in Table 42–3.

Differential Diagnosis

Informal observation by parents, teachers, and clinicians of gross motor coordination (e.g., hopping, jumping, and standing on one foot), fine motor coordination (e.g., finger tapping and shoelace tying), and hand–eye coordination (e.g., catching a ball and copying letters) is the most common initial screen for identification of coordination difficulties. However, given the vast normal variation in motor skills, it is often not easy to make a diagnosis of DCD. When coordination difficulties interfere with academic achievement or activities of daily living, however, formal testing and evaluation of these skills are recommended (Table 42–4).

Table 42–3.
ICD-10 Diagnostic Criteria for Specific Developmental Disorder of Motor Function

A. The score on a standardized test of fine or gross motor coordination is at least two standard deviations below the level expected for the child's chronological age.

B. The disturbance described in Criterion A significantly interferes with academic achievement or with activities of daily living.

C. There is no diagnosable neurological disorder.

D. Most commonly used exclusion clause. IQ is below 70 on an individually administered standardized test.

From World Health Organization. *The ICD-10 Classification of Mental and Behavioral Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization; 1993, with permission.

Table 42–4.
Motor Signs and Symptoms Indicating Referral for Comprehensive Neurological Evaluation

Hypertonia	Chorea
Hypotonia	Weakness
Ataxia	Asymmetry
Tremor	

Ruling Out Neurological Disorders. Children with developmental motor disorders progress in gross and fine motor landmarks, although at a slower rate than peers. They do not regress. Thus, a child who previously walked, ran, and climbed at age-appropriate levels, but who no longer does these activities well should be evaluated for neurological disorders. A child whose balance has deteriorated needs further testing to determine the cause of the change. Similarly, a child whose fine motor development was at age expectation but who stopped using a hand, switched apparent hand dominance, or lost previous abilities should be assessed for other medical conditions. Even a child whose initial motor landmarks were delayed should be carefully investigated if he or she loses previously established abilities.

Family history is important in the consideration of neurodegenerative conditions such as spinocerebellar degeneration, Huntington disease, peripheral neuropathies, or other neuromuscular disorders. Further investigation should be pursued even if the positive family history is of onset in adulthood. Many dominantly inherited neurodegenerative conditions show "anticipation," presenting in subsequent generations at earlier ages, even early childhood.

A physical examination, including neurological examination, should be performed. Skin stigmata, for example, may suggest a neurocutaneous syndrome (i.e., multiple café-au-lait spots, axillary or inguinal freckles suggesting NF1, depigmented macular lesions, sebaceous nevi, facial adenoma sebacea suggesting tuberous sclerosis). Excessive diffuse skin pigmentation, particularly in nonsun-exposed areas, is a clue to possible corticotropin excess, seen in boys in the early phases of adrenoleukodystrophy, at a time that the only other complaints may be loss of adept motor skills and behavior problems. Dysmorphic features, even subtle, may suggest an etiology. The presence of scoliosis, or a dimple, bulge, or hair tuft over the spine, or limb asymmetries may suggest that the motor findings are due to a structural spinal cord abnormality. Examination of the ocular fundi is sometimes challenging in small children, but is particularly important if loss of motor landmarks is accompanied by other symptoms such as headache or vomiting.

Screening evaluation of cranial nerve function in young children is most easily assessed by observation of eye movements, facial movements, listening to quality of speech, and watching the child swallow. For example, a child who cannot easily swallow water from an open cup or who coughs or gags when doing so may need to have gag reflex formally tested. The extremities should be evaluated for muscle bulk, tone, and symmetry of length and girth. Asymmetric extremity size suggests either structural brain or spinal cord lesion. In young children, strength is best assessed by observation of the child performing specific activities, such as rising from the floor, climbing onto a stool, reaching overhead for a toy, jumping, and hopping. Diffuse or symmetric muscle atrophy or weakness may suggest a neuromuscular disorder (i.e., muscular dystrophy, congenital myopathy, and neuropathy), malnutrition, or systemic illness. For example, boys with early manifestations of Duchenne muscular dystrophy may appear motor delayed and maladroit well before there are obvious complaints of weakness. Although testing of tendon reflexes may yield somewhat inconsistent results in young children, consistent asymmetries, complete absence of reflexes, or frank hyperreflexia, clonus, or extensor plantar responses mandate further investigation. Gait should be observed in multiple ways: Watch the child walking, running, tandem walking, skipping, and walking on heels and toes. In addition to watching the gait pattern, the examiner should pay attention to posture and movement of upper extremities as the child walks or runs. Abnormalities such as asymmetric arm swing or dystonic posturing of one or both hands are sometimes only appreciated with activity. Balance can be assessed both watching the child perform ordinary activities such as rising from the floor, walking, or turning, and by more structured tasks such as standing on one foot. Frank ataxia, as opposed to immaturity and delay in acquiring skills, should mandate additional investigations. Many different genetic conditions present with truncal ataxia, with or without other deficits. Initial presentation may be relatively subtle. Fine motor abilities can be observed both directly (i.e., finger to thumb tap, thumb to sequential finger tap) and by watching the child during age-appropriate activities. For example, watch the child pick up a toy, color, write, use a hand-held computer game, tablet-type computer, drink from a cup, or use a spoon. Look for asymmetries, dystonic posturing, or tremor. Look for tremor or dystonic posturing of the outstretched arms.

Poor motor coordination frequently coexists with learning disorders and disruptive behavior disorders, such as ADHD, particularly in children with high levels of impulsivity. A child may meet criteria for DCD and ADHD. The tendency for children with DCD to be accident-prone, to fall down, to hurt themselves, and to break objects may, at times, be construed as destructiveness. In the school setting, a child's motor difficulty with holding a pencil and writing legibly may result in great frustration for the

child and the teacher. For children whose impaired writing ability and poor production are inadvertently interpreted as lack of motivation, demoralization is common, and an evaluation for special educational services may be warranted. Children with DCD are at higher risk for poor social skills and are often victimized by bullies.

Jimmy was referred for evaluation to determine his need for adaptive physical education (PE) age 9 years of age. Jimmy was failing PE, was unable to run a mile, and he could not throw or catch the ball, and was generally slow motorically. His classmates laughed because he looked so "funny" while running, and he always seemed to be getting hit with the ball instead of catching it. Jimmy was always picked last for team sport games and was beginning to be bullied by his classmates for his poor physical skills. Jimmy already had an accommodation for more time on tests due to his previously diagnosed ADHD. He was being treated with Vyvanse which was working well. A developmental history revealed that Jimmy was born prematurely at 31 weeks, and his motor milestones were all delayed, although verbal milestones were right on time. Currently, Jimmy was unable to tie his shoes correctly, and he struggled to hand write letters correctly, but his parents were hoping that he would outgrow those problems. Upon questioning about Jimmy's current motor skill, it became apparent that not only was he having coordination problems, but that he was now being rejected by peers and had become quite socially isolated. At home, Jimmy's parents had accepted the fact that he was just a clumsy boy and noted that Jimmy constantly spilled his drinks and was awkward when he used a fork. His food often fell off a fork or spoon before it reached his mouth, and he had great difficulty using a knife and a fork. A comprehensive assessment of fine and gross motor skills yielded the following results: Jimmy was able to hop, but he could not skip without briefly stopping after each step. He could stand with both feet together but could not stand on tiptoe. Jimmy held a ball bounced to him at chest level, and he could not catch a ball bounced to him on the ground from a distance of 15 ft. Jimmy's agility and coordination were measured with the BOTMP, which revealed functioning levels commensurate with those of an average 7-year-old child.

Jimmy was referred to a neurologist for a comprehensive evaluation to rule out neurological disorders. On neurological examination, Jimmy was found to have no signs of a specific neurological disorder. Neurological examination was negative for conditions that could account for his poor motor skills. Based on the history and finding on the BOTMP, Jimmy was given a diagnosis of DCD. Jimmy's symptoms included gross and fine motor impairments, poor balance, and mild hypotonia. Based on an individualized educational plan (IEP) meeting with his school, it became evident that Jimmy required an adaptive physical education program and other specific learning disorders. In addition to his known ADHD, Jimmy was found to have a writing disability. A treatment plan was developed that included twice weekly with an occupational therapist who used perceptual–motor exercises to improve his fine motor skills, targeting particularly writing and use of utensils. Jimmy was also enrolled in a treatment program using motor imagery training to improve clumsiness and was also referred to a social skills group to help him deal more effectively with his peers.

Jimmy was relieved to be receiving help, especially for his handwriting and physical education class, which he had been close to failing. Over a period of 4 months of treatment, Jimmy showed some improvement in the legibility of his handwriting, although he remained a slow writer. He was responsive to the stimulant medication and his academic achievement improved as his attention span increased. Jimmy was receiving more praise from his teachers than ever before, and his classmates were teasing him less. As he began to feel better about himself, he began to try to play sports informally with his peers, although not competitively. He was given a modified physical education program in school, and he was not required to play on teams, but he practiced throwing and catching a ball and was allowed to walk his mile at a slower time. Jimmy continued to exhibit some impairment in fine motor skills over the year, and he was cooperative with the therapeutic interventions, showed some continual improvement, and was demonstrating improved peer relationships.

COURSE AND PROGNOSIS

Children with DCD represent a heterogeneous group with a variety of impairments in gross and fine motor behaviors needed for everyday living as well as in static balance. The profile of specific deficits for a given child is individual and has been shown in recent studies to be modified positively by specific motor skills interventions. The outcome of motor coordination disorder is influenced by the severity of the impairment, the age of initiation and duration of therapeutic interventions, and the presence of comorbid disorders. A recent study of preschool children with DCD at 7 to 8 years of age revealed that those children who had significant impairment at 5 years of age continued to show impairment 1.5 years later. Longer-term studies have shown that motor coordination improves over the long term in at least 50 percent of children when they reach adolescence. A 2-year follow-up study of children of primary school age who had failed a motor test battery revealed motor performance that, although improved, was still inferior to that of controls. The children with motor impairments also showed poorer educational performance than controls.

A 10-year follow-up of clumsy children and their age-matched, sex-matched peers was done in children over a period of 8 to 9 years. When first identified, the clumsy children were approximately 6 years of age; at follow-up, they ranged from 15 to 17 years of age. Follow-up testing indicated that the clumsy children were still significantly below the controls in most motor skills (e.g., dexterity, ball skills, and balance, cutting with scissors, catching, and walking backward) and in global ratings by physical education teachers. The clumsy adolescents showed higher rates of dysgraphesthesia, dysdiadochikinesis, and motor slowness. In addition, the educational achievements of the clumsy adolescents were significantly poorer, and they had significantly lower perceived social competence than their nonclumsy peers. That finding substantiates reports of continuing secondary problems, including peer problems, low self-esteem, and academic struggles in children and adolescents with significant coordination difficulties.

A long-term follow-up of children with DCD and those with ADHD and DCD at 22 years of age suggests that there are psychosocial ramifications into adulthood. Psychosocial adjustment difficulties present in the young adults with only DCD and those who were comorbid for ADHD included unemployment, use of substances, continued difficulties with reading, and legal problems.

DCD tends to improve in most individuals over time, and those with more severe impairment at a younger age, as well as those comorbid for additional psychiatric disorders, have the poorest prognoses.

TREATMENT

Interventions for developmental motor coordination can be divided into three major categories: (1) *Task-oriented* approaches such as Neuromotor

Task Training (NTT) and Cognitive Orientation to Occupational Performance (CO-OP); (2) *Process-oriented* approaches such as Sensory Integration (NI) and kinesthetic training; and (3) *traditional physical and occupational therapy* such as perceptual–motor training, and therapist-guided parent or teacher intervention. A meta-analysis of interventions to treat DCD in children by Smits-Engelsman and colleagues (2013) found that the strongest evidence of effective treatment for DCD was for task-oriented approaches and traditional motor training therapies, with process-oriented therapy yielding only weak efficacy.

Task-Oriented Approaches

Task-oriented therapies focus on learning particular motor skills with a focus on the aspects of the movements that pose the greatest difficulties for the child. NTT is a hybrid of practicing specific problematic specific motor skills with the addition of cognitive behavioral strategies to increase motor planning, reduce anxiety, and increase incentive for trying. Cognitive Orientation to Occupational Performance (CO-OP) is another form of task-oriented therapy which directs a child to form a mental model of how a particular set of movements will be accomplished, and then review with the therapist how well the mental image fit the actual motor task trials. CO-OP consists predominantly of cognitive behavioral problem solving approaches, and utilizes cognitive behavioral strategies proposed by Donald Meichenbaum.

Computer gaming with virtual physical motor tasks has been found to benefit children with DCD often without explicit instructions or learning. A study of children with DCD playing a computer game designed in which children need to catch a ball virtually on the screen, showed that they used visual cues to help them virtually catch the ball without even realizing it. These children were able to improve their computer score implicitly by practicing the game and to make use of a perceptual cue without having to be specifically instructed on how to improve. This has implications for treatment in that, in addition to providing direct instruction to children with motor coordination problems, improvement can also occur with practice through environmental cues that are not necessarily identified.

A newer task-oriented strategy is Nintendo Wii Fit Training (Wii training). The Nintendo Wii is a gaming system that combines motion sensor technology in handheld controls and a balance board provides both aspects of biofeedback and virtual reality to the player. In this way, an interactive videogame can serve as a therapeutic tool for motor skills improvement.

A study of 56 children between 6 years and 10 years with DCD, compared NTT (n = 37) to Wii training (n = 19) by Ferguson and colleagues (2013). Children in the NTT group were provided two sessions per week (45 to 60 minutes) for 9 weeks in which they played games such as soccer, tag,

and netball under the guidance of therapists who facilitated learning of motor skills, provided positive feedback, and manipulated the environment when necessary. Children in the Wii group, participated in 30 minutes of gaming three times a week for 6 weeks under the supervision and guidance of therapists who encouraged, praised, and motivated the children. Games included sports such as cycling, soccer, skateboarding, and skiing. Games required that the children shift their weight from side to side, and jump over virtual objects, and some games required twisting of arm movements. The results found mean motor performance scores in both groups improved during the course of the study. The NTT group showed greater improvement in overall motor performance, functional strength and cardiorespiratory fitness, while the Wii group showed significant improvement in anaerobic performance. Thus, this study suggests that both NTT and Wii training can provide valuable therapeutic gains.

Overall recommendations for first-line interventions for children with DCD include task-oriented therapies of various types. Younger children can benefit from NTT, for problems such as handwriting, that is, writing letters, even before they learn to read. Older children and those with verbal skills may be able to benefit from CO-OP approaches which incorporate more verbal planning. Nintendo Wii and other gaming interventions can be beneficial for children with DCD of all ages.

Process-Oriented Approaches

There has been only a weak suggestion from studies that process-oriented approaches are effective therapies for DCD, however, they are often used in conjunction with other strategies. Process-oriented interventions aim to target neural structures such as the cerebellum for improving balance, or sensory processes, such as proprioception in children with DCD. Interventions aim to enhance and remediate sensory input processing, for example, sensory integration therapy (SIT). SIT approach was introduced in the work of Ayres in the late 1970s who focused remediation on enhancing sensory information processing in order to achieve more accurate motor execution. Sensory integration programs are often delivered as a related service to children with DCD under the direction of occupational therapists. These programs prescribe specific physical therapies that attempt to modify the motor and sensory functioning of these children. These therapies are based on the notion that sensory integration disorders interfere with body movements and body awareness. Activities are designed to target the tactile, vestibular, and proprioceptive systems. A sensory integration program that addresses disorders affecting the proprioceptive system (bumping into walls and inability to button, to skip, or to write) might involve using scooter boards to improve balance and body awareness. Multisensory modalities, including visual, auditory, and tactile materials, are utilized in this form of treatment. Motor imagery

training for clumsiness in children, have received attention recently, based on results from controlled investigations of this modality compared to perceptual–motor interventions. A recent study used a successful motor imagery intervention through CD-ROM, which contained six main training components for children with coordination difficulties. The components included visual imagery exercises involving predictive timing, relaxation and mental preparation, visual modeling of fundamental motor skills, mental rehearsal of skills from an external perspective, mental rehearsal of skills from an internal perspective, and overt practice. This model of treatment is based on the notion that deficits in a child's ability to internally represent the movement tasks contribute to poor motor coordination. The ability for visual imagery of objects is not typically impaired in children with DCD. The previously mentioned training exercises were designed to improve the timing and accuracy of a given child's ability to visualize motor tasks, as well as to reduce muscle tension while imagining completing the given motor task. Finally, the children were trained in mentally rehearsing the task before actually practicing it.

Traditional Physical and Occupational Therapy

Traditional physical and occupational therapies target actual activities of daily function and seek to improve the execution of these motor behaviors. A variety of cognitive motor approaches have been utilized to improve the planning and execution of motor skills. One cognitive motor approach to improving functional skills by Sugden and Hendersen has been updated and labeled ecological intervention, which places emphasis on the actual control of movement and is done in a real-life setting.

Traditional perceptual–motor training consists of a combination of gross motor, fine motor, and perceptual–motor activities related to a child's actual daily motor tasks, such as hopping, skipping, jumping, climbing and moving through hoops, climbing ladders, and jumping on a trampoline, to practice competency. Specific tasks addressing fine motor skills, for example, handwriting training, includes activities to improve grip of a pencil, wrist rotation, and posture. The results of the previously mentioned study were extremely notable in that the mental imagery treatment was as efficacious as traditional perceptual–motor training exercises in improving a given child's coordination.

An example of functional skill intervention would be those that address impairment in handwriting. It is estimated that between 10 and 30 percent of children exhibit persistent impairment in handwriting, a skill that remains a critical means of communication even in the face of the pervasive use of computers. Handwriting impairment is associated with academic skills disorders. Handwriting treatment approaches may include perceptual–motor, visual–motor, motor control, individualized interventions or exercises, and supplementary handwriting instruction.

Few controlled studies have examined which techniques are most efficacious, and a survey of 50 pediatric occupational therapists found that 90 percent of them advocated an eclectic approach to interventions for handwriting improvement. A study of a school-based occupational therapy intervention for school-aged children with poor handwriting legibility showed efficacy in improved legibility of letters. Children with handwriting difficulties in this study received direct traditional perceptual instruction on writing letters for approximately 30 minutes per week for approximately 7 months. This consisted of a variety of handwriting interventions, such as using vibration or resistance to writing to increase appropriate grip and writing on a chalkboard or a vertical surface to increase arm strength and stability while writing. At the end of this period, children made significant gains in handwriting legibility, compared to a control group that did not receive the intervention. Handwriting speed in the children who received the treatment increased from 32 to 37 letters per minute. Although this was a trend toward speed, this did not represent a statistically significant increase in speed. Some students in the treatment group wrote their letters with a greater degree of deliberate care, and, although this increased legibility, it may have actually diminished their speed. Overall, the increase in legibility applied to letters more than numbers but did not significantly increase writing speed.

References

Adams IL, Lust JM, Wilson PH, Steenbergen B. Compromised motor control in children with DCD: a deficit in the internal model? A systematic review. *Neurosci Biobehav Rev*. 2014;47:225–244.

*Blank R, Smits-Engelsman B, Polatajko H, Wilson P. European Academy for Childhood Disability: recommendations on the definition, diagnosis and intervention of developmental coordination disorder (long version). *Dev Med Child Neurol*. 2012;54(1):54–93.

Caçola P. Physical and mental health of children with developmental coordination disorder. *Front Public Health*. 2016;4:224.

Camden C, Wilson B, Kirby A, Sugden D, Missiuna C. Best practice principles for management of children with developmental coordination disorder (DCD): results of a scoping review. *Child Care Health Dev*. 2015;41(1):147–159.

Case-Smith J. Effectiveness of school-based occupational therapy intervention on handwriting. *Am J Occup Ther*. 2002;56(1):17–25.

Crawford SG, Wilson BN, Dewey D. Identifying developmental coordination disorder: consistency between tests. *Phys Occup Ther Pediatr*. 2001;20(2–3):29–50.

Dewey D, Kaplan BJ, Crawford SG, Wilson BN. Developmental coordination disorder: associated problems in attention, learning, and psychosocial adjustment. *Hum Mov Sci*. 2002;21(5–6):905–918.

Dunford C, Street E, O'Connell H, Kelly J, Sibert JR. Are referrals to occupational therapy for developmental coordination disorder appropriate? *Arch Dis Child*. 2004;89(2):143–147.

Feder KP, Majnemer A. Handwriting development, competency and intervention. *Dev Med Child Neurol*. 2007;49(4):312–317.

Ferguson GD, Aertssen WF, Rameckers EA, Jelsma J, Smits-Engelsman BC. Physical fitness in children with developmental coordination disorder: measurement matters. *Res Dev Disabil*. 2014;35(5):1087–1097.

*Ferguson GD, Jelsma D, Jelsma J, Smits-Engelsman BC. The efficacy of two task-oriented interventions for children with Developmental Coordination Disorder: Neuromotor Task Training and Nintendo Wii Fit Training. *Res Dev Disabil*. 2013;34(9):2449–2461.

Fliers E, Rommelse N, Vermeulen SH, et al. Motor coordination problems in children and adolescents with ADHD rated by parents and teachers: effects of age and gender. *J Neural Transm*. 2008;115(2):211–220.

Gillberg C, Kadesjo B. Why bother about clumsiness? The implications of having developmental coordination disorder (DCD). *Neural Plast*. 2003;10(1–2):59–68.

Gueze RH, Jongmans MJ, Schoemaker MM, Smits-Engelsman BC. Clinical and research diagnostic criteria for developmental coordination disorder: a review and discussion. *Hum Mov Sci*. 2001;20(1–2):7–47.

Hay JA, Hawes R, Faught BE. Evaluation of a screening instrument for developmental coordination disorder. *J Adolesc Health*. 2004;34(4):308–313.

Holsti L, Grunau RV, Whitfield MF. Developmental coordination disorder in extremely low birth weight children at nine years. *J Dev Behav Pediatr*. 2002;23(1):9–15.

Miller LT, Polatajko HJ, Missiuna C, Mandich AD, Mcnab JJ. A pilot trial of a cognitive treatment for children with developmental coordination disorder. *Hum Mov Sci*. 2001;20(1–2):183–210.

Noten M, Wilson P, Ruddock S, Steenbergen B. Mild impairments of motor imagery skills in children with DCD. *Res Dev Disabil*. 2014;35(5):1152–1159.

Noordstar JJ, van der Net J, Voerman L, Helders PJ, Jongmans MJ. The effect of an integrated perceived competence and motor intervention in children with developmental coordination disorder. *Res Dev Disabil*. 2016;60:162–175.

Novak I. New evidence in developmental coordination disorder (DCD). *Phys Occup Ther Pediatr*. 2013;33(2):170–173.

O'Hare A, Khalid S. The association of abnormal cerebellar function in children with developmental coordination disorder and reading difficulties. *Dyslexia*. 2002;8(4):234–248.

Pless M, Carlsson M, Sundelin C, Persson K. Preschool children with developmental coordination disorder: a short-term follow-up of motor status at seven to eight years of age. *Acta Pediatr*. 2002;91:521–528.

Prunty MM, Barnett AL, Wilmul K, Plumb MS. The impact of handwriting difficulties on compositional quality in children with developmental coordination disorder. *Br J Occup Ther*. 2016;79:591–597.

*Rasmussen P, Gillberg C. Natural outcome of ADHD with developmental coordination disorder at age 22 years: a controlled, longitudinal, community-based study. *J Am Acad Child Adolesc Psychiatry*. 2000;39:1424–1431.

Rivard L, Camden C, Pollock N, Missiuna C. Knowledge to practice in developmental coordination disorder: utility of an evidence-based online module for physical therapists. *Phys*

Occup Ther Pediatr. 2015;35(2):178–194.

Schoemaker MM, Flapper B, Verheij NP, Wilson BN, Reinders-Messelink HA, de Kloet A. Evaluation of the developmental coordination disorder questionnaire as a screening instrument. *Dev Med Child Neurol*. 2006;48(8):668–673.

*Smits-Engelsman BC, Blank R, Van Der Kaay AC, et al. Efficacy of interventions to improve motor performance in children with developmental coordination disorder: a combined systematic review and meta-analysis. *Dev Med Child Neurol*. 2013;55(3):229–237.

Sugden DA. Current approaches to intervention in children with developmental coordination disorder. *Dev Med Child Neurol*. 2007;49:467–471.

*Sugden DA, Chambers ME. Intervention in children with developmental coordination disorder: the role of parents and teachers. *Br J Educ Psychol*. 2003;73:545–561.

Tsai CL, Wu SK, Huang CH. Static balance in children with developmental coordination disorder. *Hum Move Sci*. 2008;27(1):142–153.

Venetsanou F, Kambas A, Aggeloussis N, Serbezis V, Taxildaris K. Use of the Bruininks-Oseretsky Test of Motor Proficiency for identifying children with motor impairment. *Dev Med Child Neurol*. 2007;49(11):846–848.

Visser J. Developmental coordination disorder: a review of research on subtypes and comorbidities. *Hum Mov Sci*. 2003;22(4–5):479–493.

Wade MG, Kazeck M. Developmental coordination disorder and its cause: the road less travelled [published online ahead of print November 19, 2016]. *Hum Mov Sci*. doi: 10.1016/j.humov.2016.08.004

Williams J, Kashuk SR, Wilson PH, Thorpe G, Egan GF. White matter alterations in adults with probable developmental coordination disorder: an MRI diffusion tensor imaging study. *Neuroreport*. 2017;28(2):87–92.

*Wilson PH, Ruddock S, Smits-Engelsman B, Polatajko H, Blank R. Understanding performance deficits in developmental coordination disorder: a meta-analysis of recent research. *Dev Med Child Neurol*. 2013;55:217–228.

Wilson PH, Thomas PR, Maruff P. Motor imagery training ameliorates motor clumsiness in children. *J Child Neurol*. 2002;17(7):491–498.

Woodward S, Swinth Y. Multisensory approach to handwriting remediation: perceptions of school-based occupational therapists. *Am J Occup Ther*. 2002;56(3):305–312.

*Zwicker JG, Missiuna C, Harris SR, Boyd LA. Developmental coordination disorder: a review and update. *Eur J Paediatr Neurol*. 2012;16(6):573–581.

43 Communication Disorders

▲ 43.1 Language Disorder

EMIKO KOYAMA, PH.D. AND JOSEPH H. BEITCHMAN, M.D.

INTRODUCTION

Communication disorders are among the most common problems exhibited by young children. Nonetheless, communication disorders often go unrecognized, particularly when they cooccur with other emotional or behavioral problems. Thus, child psychiatrists, psychologists, and other clinicians need to be alert to the potentially damaging impact of communication problems on a child's overall well-being. Proficiency in verbal communication is critical to success in virtually all human activities and is therefore highly valued in society. Individuals with communication disorders may be unable to participate fully and competently in everyday interpersonal, learning, and occupational situations. Limited participation in such key life activities may further restrict growth in communication and related domains.

To develop effective communication, children must master language, the conventional code used to express ideas about the world and speech, the complex and rapid motor movements that translate such ideas into spoken words. Children need to also learn how to use language and speech appropriately in a given social context. The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) recognizes five diagnostic categories for communication disorders. One category concerns a disorder in language (language disorder), two categories relate to disorders in speech (speech sound disorder and childhood-onset fluency disorder), one category is associated with language use (social [pragmatic] communication disorder), and the final category covers communication disorder not otherwise specified.

DEFINITION AND COMPARATIVE NOSOLOGY

Individuals with language disorders show persistent impairments in the

expression and reception (production and comprehension) of language. These impairments will manifest across different modalities including speaking, writing, and sign language. The expressive difficulties may include limitations in vocabulary size, sentence complexity, and coherence in discourse. These expressive problems, however, are of less overall concern than the receptive difficulties. Children with receptive problems may manifest deficits in basic auditory processing skills, such as auditory discrimination, detection of rapid sound changes, association of sounds and symbols, and memory for sound sequences. They may be slow to learn word meanings, especially those that are uncommon or abstract. They may not understand the questions or comments of others and may misinterpret complex directions and explanations. They may have difficulty following conversations or learning from lectures. Because of comprehension failures, individuals with language disorders may give illogical or inappropriate responses. Furthermore, their comprehension problems may go unrecognized and instead may be misattributed to inattention, stubbornness, or other psychological or behavioral problems. Because of these problems, individuals with a language disorder would often have limitations in academic achievement, occupational success, and other social participation.

The definition of language disorder in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) combines elements of both the expressive and mixed receptive-expressive language disorders of the revised fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR) and the 10th edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10). Both the DSM-IV-TR and the ICD-10 conceptualized the expressive and receptive subtypes as being mutually exclusive language disorders, characterized by significant language problems, despite normal levels of nonverbal intelligence. DSM-IV-TR, however, differentiated between the developmental and acquired subtypes, whereas the ICD-10 only recognizes the developmental type, and acquired language problems such as acquired aphasia with epilepsy are listed under separate categories.

EPIDEMIOLOGY

The prevalence estimates of language disorder differ greatly from roughly 3 to 15 percent. Receptive problems are thought to be somewhat less common than expressive problems. The variation in prevalence estimates is most likely due to differences in language measures used to define the disorder, the age of the children surveyed, and the nature of the sample tested (community or clinic referred).

Language disorders are more common in boys than in girls. Various studies have yielded sex ratios as high as 5:1 and as low as roughly 2:1.

Familial aggregation is characteristic of language disorders. A variety of

methods (aggregation, pedigree, twin, adoption, and linkage studies) have provided clear evidence of genetic influences on language disorder. Complementary work also points to environmental input as a key factor that likely contributes to familial aggregation.

Approximately 10 to 15 percent of children younger than 3 years of age show delayed onset and slow initial progress in language development. Because of the large range of normal variability in the language production skills of young children, it is not yet clear whether an initial delay consistently warrants a diagnosis of language disorder. The majority (50 to 80 percent) of these children eventually acquire language skills within the normal range, but some show persistent language difficulties into adulthood. Children who continue to exhibit language problems after age 4 are more likely to experience long-term language deficits which may persist into adulthood, although the specific language profiles may change over time.

ETIOLOGY

There are no known causes for language disorder. Various factors have been shown to correlate with language disorder, but none has yet been established as causal. Language disorder most likely has multiple determinants, including various biological and environmental factors. Different factors or combinations may be implicated in some cases and not others, as suggested by the heterogeneity in the population. At least seven general types of possible causal factors have been investigated: (1) brain abnormalities, (2) genetic factors, (3) environmental influences, (4) neurodevelopmental immaturity, (5) general cognitive factors, (6) specific linguistic factors, and (7) auditory processes. Each of the proposed factors enjoys some empirical support, but none has gained widespread acceptance as a plausible account for all cases of language disorder. Moreover, the factors are not mutually exclusive and may interact in complex ways.

Structural or functional brain abnormalities can now be studied in a sophisticated fashion with advanced neuroimaging techniques. Such abnormalities have long been considered possible causal factors in language disorder, because (1) language characteristics of language disorder resemble those of acquired language disorders that result from known neurological insults, and (2) children with language disorders show elevated rates of soft neurological signs (e.g., mixed cerebral dominance, attention problems, or fine motor difficulties). Neuroimaging studies of children with language disorders have demonstrated structural and functional abnormalities, often in brain regions known to be crucial for language processing (e.g., left hemisphere or certain subcortical structures). Nonetheless, many questions remain. Both the language disorder phenotype and the techniques used to examine the brain abnormalities differ greatly among the studies, making it difficult to draw

conclusions about how brain abnormalities contribute to language disorder. Furthermore, because of the important role that experience plays in brain development, it is not clear whether some of the abnormalities noted are the cause or the result of language disorder.

Genetic factors have now been clearly implicated in language disorder, but it is far less clear just what these factors are and how they operate. Loci on chromosomes 13, 16, and 19 have been tentatively linked to aspects of language disorder in several studies. Although these regions are unlikely to be the sole determinants of genetic influences on language, their investigation may yield important insights. For example, some recent linkage studies have suggested that chromosome 16q may be associated with phonological short-term memory and chromosome 19q with expressive language. Work also continues to refine or broaden the language disorder phenotype to accurately identify possible subtypes and modes of transmission. An interesting possibility suggested by several lines of evidence is that autism and language disorder may share some relevant genetic characteristics. Genetic factors may also be important in understanding the origin of the brain abnormalities noted earlier.

Various environmental factors may also be critical, at least in some cases of language disorder. Children from families of low socioeconomic status (SES) are at elevated risk for language disorder. Low SES, in turn, correlates with other known risk factors, such as single-parent family status, large family size, lower parental education, pregnancy and birth complications, and lack of appropriate learning materials in the home, and abuse and neglect. Children from low SES families are also at risk for malnutrition, other health problems, increased exposure to known toxins (e.g., lead), and limited educational and social opportunities. The amount and quality of linguistic input available to children have also been shown to vary dramatically across families, with some children hearing as few as 700 utterances a day and others as many as 11,000! Furthermore, and not surprisingly, these variations in input were strongly related to language development as measured by standardized tests. In adverse conditions, supportive extended family and neighbors may act as protective factors assisting the children's language development. Further work is needed to clarify whether or how various environmental and input factors may interact with one another and, perhaps, with biological factors to cause language disorder.

Neurodevelopmental immaturity may also underlie language disorder. Proponents argue that language development generally follows a normal sequence and course but at a slower rate; that is, it is delayed rather than deviant. Delayed maturation of neural structures or slow myelinization of neural pathways might also account for associated characteristics that are common in children with a language disorder, such as attention and motor problems and other behavioral difficulties. Furthermore, such

neurodevelopmental immaturities may occur more often in boys than in girls. A related view is that children with language disorder simply have language skills that fall at the lower end of the normal distribution. The only reason that these children are perceived as handicapped is because language skills are so essential for success in a literate society. When other less crucial abilities, such as musical talent, fall in the lower end of the normal distribution, children are not perceived as handicapped, because musical proficiency is not expected of all members of society.

A variety of general cognitive factors have also been proposed to account for language disorder. Speed of cognitive processing is one such factor. Children with a language disorder typically perform slower than their peers without the disorder on a variety of linguistic (e.g., naming or grammatical judgments) and nonlinguistic (e.g., fine motor or visual-spatial) tasks. This slower processing speed might somehow interfere with language learning. Some research studies, however, failed to show significant correlations between processing speed and language achievement, casting doubt on this suggestion. Others propose that phonological working memory, the immediate maintenance and manipulation of verbal information, is a key cognitive factor in language disorder. Children with a language disorder show deficits in phonological working memory, which have been linked to deficits in vocabulary learning and reading proficiency. However, strong causal inferences about the role of phonological working memory are not yet possible based on the available correlational evidence.

In contrast to the neurodevelopmental immaturities and general cognitive delay perspectives, that imply that children with a language disorder experience an overall delay but nevertheless follow a typical course of language development, some researchers have argued that children with language disorders may show specific lingering deficits in certain aspects of language that a young child with otherwise matched language ability may not show. For example, children with language disorders have particular difficulty in learning certain grammatical features (e.g., verb tense markers or pronouns in English) to the extent that their use of these features is even poorer than that of younger children at the same level of language development. Much-needed cross-linguistic work in languages with varied grammars is helping theorists hone in on the commonalities and differences in the grammatical features that pose particular problems for children with language disorders who are learning different languages. At this point, however, these linguistic theories are primarily descriptive rather than explanatory.

Auditory problems have also been investigated as possible causes of language disorders. One concern is otitis media, an inflammation of the middle ear that results from ear infections or allergies. Otitis media is often accompanied by mild hearing loss, which is usually temporary if the

underlying infection is treated. Repeated episodes of otitis media and associated hearing loss are common in young children during the earliest stages of language development. Although otitis media alone is not sufficient to cause language disorder, it may contribute in combination with other risk factors. A second concern centers on auditory perception of sounds. Many (but not all) children with language disorders have difficulty perceiving small sound differences and rapid auditory transitions, which are aspects of the acoustic signal thought to be important in understanding speech. It has proven difficult, however, to develop convincing theoretical and empirical links between auditory perceptual skills and the various language difficulties shown by children with language disorder.

DIAGNOSIS AND CLINICAL FEATURES

The hallmark of language disorder is impairment in receptive and expressive language development. The rate of language growth is slower in children with language disorder than children with typical language development. The receptive and expressive language characteristics of affected children may be largely similar to those of younger children who are acquiring language in a typical fashion. Certain language features, such as the grammatical marking of tense, aspect, number, gender, or case may be especially difficult for children with language disorders to acquire. They may have difficulty learning the correct use of auxiliary or modal verbs (e.g., is, were, could), plurals (e.g., cats, geese), verb tense endings (e.g., -ing, -s, -ed), and pronouns (e.g., he, she, they).

Comprehension often varies with the complexity of the language. For example, words for familiar, concrete objects, actions, or relations may be comprehended (e.g., *wagon*, *apple*, *walk*, *give*, *in*, *over*), but those with less familiar, abstract meanings may be unknown (e.g., *neuron*, *freedom*, *receive*, *pretend*, *between*, *when*). Similarly, sentences with simple grammatical forms may be understood (e.g., "The boy is smiling." "The phone rang."), whereas those with more complex structures may be misinterpreted (e.g., "The car that rammed the bus was towed to the garage." "Before she finalized the arrangements, she consulted with her lawyer."). Complex or lengthy instructions, questions, and explanations may not be comprehended. When children with language disorders misconstrue the language and intentions of others, they may inadvertently make inappropriate responses or comments during conversations. Their comprehension failures may also contribute to confusion or inaccuracy in telling stories, relating events, and giving explanations, problems that may be further exacerbated by their expressive language limitations. When their comprehension difficulties are not recognized, children with language disorders may be mistakenly thought to be inattentive, lazy, or uncooperative.

The specific manifestations of language disorder vary with age. Some

children with language disorder eventually acquire language skills within the normal range, but the overall prognosis is poorer for children with severe receptive language problems than for children with just expressive language difficulties. The expressive and receptive features adversely affect personal interactions and academic achievement because spoken language is poorly expressed and understood.

Amy was a sociable, active 5-year-old who was diagnosed with a language disorder. She often played with Lisa, her kindergarten classmate. During pretend play one day, each girl told the story of Little Red Riding Hood to her doll. Lisa's story began: "Little Red Riding Hood was taking a basket of food to her grandmother who was sick. A bad wolf stopped Red Riding Hood in the forest. He tried to get the basket away from her but she wouldn't give it to him."

By contrast, Amy's story illustrated her marked difficulties in verbal expression: "Riding Hood going to grandma house. Her taking food. Bad wolf in a bed. Riding Hood say, what big ears, grandma? Hear you, dear. What big eyes, grandma? See you, dear. What big mouth, grandma?

Eat you all up!"

Amy's story contained many features characteristic of children with language disorders at her age, including short, incomplete sentences, simple sentence structures, omission of grammatical function words (e.g., is and the) and inflectional endings (e.g., possessive -'s and present tense verb -s), problems in question formation, and incorrect use of pronouns (e.g., her for she).

Nevertheless, Amy performed as well as Lisa did in understanding the details and plot of the Riding Hood tale, as long as she was tested with methods that did not involve verbal responding. Amy also demonstrated adequate comprehension skills in her kindergarten classroom, where she readily followed the teacher's complex, multistep verbal instructions (e.g., "After you write your name in the top left corner of your paper, get your crayons and scissors, put your library books under your chair, and line up at the back of the room.").

Julio was a quiet, sullen 8-year-old whose language problems were no longer obvious in casual, social conversations. His speech now rarely contained the incomplete sentences and grammatical errors that were so evident when he was younger. His expressive problems, however, still surfaced in tasks involving elaborate or abstract uses of language, like those required in much of his third-grade academic work. An example was Julio's explanation of the outcome of a recent science experiment: "The teacher had stuff in some jars. He poured it, and it got pink. The other thing made it white." Although each sentence was grammatical, his explanation as a whole was difficult to follow, because key ideas and details were omitted or poorly explained. Julio also showed problems in word finding, that is, in using specific words for the concepts and actions he was describing. Instead, he relied on vague and nonspecific terms, such as thing, stuff, and got.

In early elementary grades, Julio had kept pace with his classmates in reading, writing, and other academic skills. By third grade, however, the increasing demands for written work began to negatively affect his overall academic standing. His written work was characterized by problems similar to those noted in his oral expression, such as poor organization and lack of specificity. Classmates also began to tease him about his difficulties, and he reacted quite aggressively, sometimes to the point of fighting. Nonetheless, Julio continued to show relatively good comprehension of spoken language, including classroom lectures concerning abstract concepts. He also comprehended sentences that were grammatically and conceptually complex (e.g., "The car the truck hit had hubcaps that were stolen. Had it been possible, she would have notified us by mail or by phone").

Susan was a pleasant 2-year-old who did not yet use any spoken words. She made her needs known with vocalizations and simple gestures (e.g., showing or pointing) like those typically used by younger children. She seemed to understand the names for a few familiar people and

objects (e.g., *mommy*, *daddy*, *cat*, *bottle*, and *cookie*). However, compared to other children her age, she had a small comprehension vocabulary and showed limited understanding of simple verbal directions (e.g., "Get your doll." "Close your eyes."). Nonetheless, her hearing was normal, and her motor and play skills were developing as expected for her age. She showed interest in her environment and in the activities of the other children at her day care.

Min was a shy, reserved 5-year-old who grew up in a bilingual home. Min's parents and older siblings spoke English and Cantonese proficiently. Her grandparents, who lived in the same house, spoke only Cantonese. Min began to understand and speak both languages much later than her older siblings had. Throughout her preschool years, Min continued to develop slowly in comprehension and production. At the start of kindergarten, Min understood fewer English words for objects, actions, and relations than her classmates did. She showed difficulties in following classroom instructions, particularly those that involved words for concepts of time (e.g., *tomorrow*, *before*, or *day*) and space (e.g., *behind*, *next to*, or *under*). It was also hard for Min to match one of several pictures to a syntactically complex sentence that she had heard (e.g., "It was not the train she was waiting for." "Because he had already completed his work, he was not kept after school."). Min occasionally tried to speak with other children. However, these conversations usually broke down because she misinterpreted what others said or could not express her own thoughts clearly. Consequently, her classmates generally ignored her, preferring instead to play with more verbally competent peers. Min's infrequent interactions further limited her opportunities to learn and to practice her already weak language skills. Min also showed limited receptive and expressive skills in Cantonese, as revealed by an assessment conducted with the assistance of a Cantonese interpreter. Nonetheless, her nonverbal cognitive and motor skills were within the normal range for her age. She showed no difficulties in solving spatial, numerical, conceptual, or analogical problems, provided they were presented nonverbally.

Fred received a diagnosis of language disorder when he was a preschooler. By 8 years of age, he had also received the comorbid diagnoses of specific learning disorder and attention-deficit/hyperactivity disorder (ADHD). This combination of language, learning, and attention problems made it virtually impossible for Fred to succeed in school. His comprehension and attention difficulties limited his ability to understand and to learn important information from classroom instructions, discussions, and lectures. Because the unlearned information was critical to the understanding of future academic lessons, Fred fell further behind his classmates. He was also disadvantaged by the fact that he could read only a few familiar words. This meant that he could not learn other academic information that his classmates acquired by reading textbooks, library books, newspapers, and other written materials. His poor reading ability also limited his opportunities to learn the new vocabulary, complex sentence forms, and sophisticated ideas that other children absorbed from reading. Despite his academic problems, however, Fred continued to show nonverbal intellectual skills within the low normal range, although his scores were somewhat lower than those he had earned as a preschooler.

The disorders of speech, learning, and development may cooccur with language disorder. The two most frequently cooccurring disorders are speech sound disorder and specific learning disorder. Speech sound disorder, which is characterized by difficulty in correct production of speech sounds, occurs in roughly 40 to 60 percent of preschoolers with language disorder. Specific learning disorder, especially difficulties in reading, occurs in approximately 50 to 90 percent of school-age children with language disorder and most frequently for children with receptive language problems. Difficulties in learning to read may be apparent early or may not surface until later in the learning process when knowledge of

sophisticated vocabulary and sentence structures is required to support reading comprehension. Because of this high comorbidity between language disorders and reading problems, some researchers have theorized that language impairment and reading difficulties may be different manifestations of the same underlying problems that manifest itself at different ages. Genetic studies also suggest that families with known risk of language disorders also tend to be at higher risk for reading problems and vice versa. However, other researchers suggest that reading and language problems should be considered separate disorders, as children with language problems manifest problems that go beyond phonological awareness and include syntax and semantics. Future studies in molecular genetics may help clarify this debate. A few children with language disorders also experience stuttering and developmental coordination disorder.

The risk for concomitant psychiatric disorders is also high for children with language disorder, especially when receptive language is impaired. Estimates vary, but comorbid psychiatric disorders may affect as much as 70 percent of school-age children with language disorder who have severe receptive language problems. The most common comorbid condition is ADHD, occurring in 30 to 60 percent of cases. Internalizing disorders are also noted but not as often as ADHD. Boys are at particular risk for externalizing disorders. Behavior difficulties tend to occur most frequently in children who have concomitant language and reading problems. Girls with language disorders may be withdrawn and quiet, making detection of language disorders more difficult.

Most children with language disorder do not show frank neurological symptoms (e.g., tremor, paralysis, or seizures). As a group, however, children with language disorders show elevated rates of mild to moderate abnormalities (e.g., in reflexes or fine motor function) on careful neurological examination. They are also more likely than typically developing children to show mild to moderate abnormalities in measures of brain function or structure, such as electroencephalographic (EEG) readings or magnetic resonance imaging (MRI) and functional MRI (fMRI) scans. Furthermore, the number of abnormal neurological signs may be associated with the severity of language difficulties.

PATHOLOGY AND LABORATORY EXAMINATION

There are no laboratory tests that can provide a diagnosis of language disorder. Neuroimaging techniques, such as MRI and fMRI, sometimes reveal structural or functional brain abnormalities associated with language disorder. However, such abnormalities have neither been found in all cases nor have their possible etiological or diagnostic values been determined.

DIFFERENTIAL DIAGNOSIS

Standardized evaluations of expressive and receptive language and nonverbal intellectual functioning are the key elements for differential diagnosis of language disorder. Receptive language testing is particularly important because comprehension deficits may be less obvious than expressive difficulties. Receptive and expressive language performance must fall substantially below the range of normal language performance expected for a child's age. In addition, the severity of the receptive and expressive language problems must hinder academic performance or social communication. Speech–language pathologists often use observational techniques and analyses of spontaneous language to assess the functional impact of language difficulties and to complement information obtained from standardized language tests. A thorough functional assessment of language abilities may also be substituted when appropriate standardized tests are not available (e.g., if a child speaks a language for which tests have not yet been developed).

A significant discrepancy between receptive language and nonverbal abilities precludes the alternate diagnoses of selective mutism and intellectual disabilities. In selective mutism, receptive language and nonverbal performances usually are similar in level and markedly higher than expressive language skills. In intellectual disabilities, performance is impaired substantially in all three areas of receptive language, expressive language, and nonverbal abilities, relative to age-level expectations.

Deficits in receptive language may also be evident in cases of hearing impairment, environmental deprivation, or autism spectrum disorder. A hearing test should be part of the assessment of all cases of suspected language disorder. Collection of a careful history can usually identify severe environmental deprivation. Observations of disturbed social interaction, limited communicative intent, and unusual repetitive behaviors suggest autism spectrum disorder. A concurrent diagnosis of language disorder may be made when receptive and expressive language deficits exceed those expected for the degree of intellectual disabilities, hearing impairment, or environmental deprivation.

Differential diagnosis requires attention to other comorbidity issues. Speech sound disorder is a possible comorbid diagnosis when difficulties in speech sound production accompany language disorder. School-age children identified with either language or specific learning disorder should be screened for difficulties in the other area because the two disorders frequently cooccur. Similarly, children with psychiatric diagnoses, particularly ADHD, should be screened for language disorder because it is known that a high proportion of children referred for psychiatric assessment shows previously undetected comorbid language disorders. Sometimes a child shows expressive language that is so difficult to follow and so disorganized that the practitioner questions whether it should be

considered disordered thinking symptomatic of schizophrenia.

Illustrative of this concern is the case study presented here with a brief excerpt from a conversation with Max, a highly anxious 9-year-old with language disorder who had just visited the dentist. The stream of language in this case study raises serious concerns about a possible diagnosis of schizophrenia. However, a child with language disorder may produce language this disorganized, especially in situations of heightened anxiety. Diagnosis can usually be clarified by reassessment when the child is less anxious and by careful attention to the presence or absence of comorbid symptoms.

What do you suppose you'd like to do today?

Taking a dentist appointment ...

[You're taking a dentist appointment.]

Yeah, he has a little cavity ...

[He has a little cavity.]

... and I'm going to take it from him cause he's stuck there he sure doesn't like it but he knows there is too far (mumbled)...

... too far and so he was losing a lot of hair so he needs a new way ... sounds like he's bald ... so he has everything (mumbling) a guy who's not recep ... teaching he will have to it's important like he will or he won't.

Acquired language problems may result from a variety of neurological insults, including (1) cerebral vascular accidents, such as strokes; (2) infections, such as encephalitis or meningitis; (3) severe respiratory problems that result in intraventricular hemorrhage or anoxia; (4) near drownings; (5) closed head injuries from falls, motor vehicle accidents, or other blows to the head; (6) open head injuries, such as those from gunshot wounds; (7) toxins, such as lead or household poisons; (8) damage from tumors or radiation treatment; or (9) seizure activity. Landau–Kleffner syndrome (acquired aphasia with epilepsy) is a rare condition marked by an abrupt loss of language skills, which is often associated with the onset of seizures or abnormal EEG activity.

Acquired language problems vary considerably in their language symptoms as a function of the nature and severity of the underlying neurological condition. Acquired disorders also show distinctive language characteristics as well as some features shared with language disorder. For example, closed head injury, which often occurs in motor vehicle accidents or falls, typically results in sparing of basic skills in vocabulary and grammar (after initial recovery stages), whereas more complex skills, particularly in discourse, remain markedly impaired. Comprehension for lengthy or complex discourse may be limited. Likewise, expression in discourse may be disrupted by disorganized, tangential, inappropriate, and illogical comments, even if basic vocabulary and sentence structures are used correctly. Communication effectiveness may be further hindered by accompanying deficits in memory, attention, affect, behavior, and motor skills.

Landau–Kleffner syndrome also shows a unique pattern. Children with Landau–Kleffner syndrome seem to be developing normally until around ages of 3 to 7 years, when they start to have trouble with expressive and

receptive language with no apparent reason. The decline in language ability may be sudden or gradual. Hearing loss is usually suspected by parents; however, laboratory tests disconfirm this suspicion. In about 80 percent of children with Landau–Kleffner syndrome, one or more episodes of epileptic seizures occur, typically during sleep. Behavioral problems such as hyperactivity, aggressiveness, or depression may cooccur.

Elliott, 4 years of age, was a cheerful, talkative preschooler who suddenly experienced a series of seizures for no apparent reason. His language skills, which had been excellent, simultaneously showed a precipitous decline. He no longer understood words, commands, and sentences that had posed no difficulty for him before the seizures. He also lost his expressive skills and began using shorter, less sophisticated sentences, which became increasingly incoherent as he omitted or misused important words. He also developed significant memory and attention problems. Nonetheless, he continued to show nonverbal intellectual skills that were grossly within the normal range. Although Elliott's seizures were eventually controlled with medication, he experienced persistent problems in comprehension and production of language, which never fully resolved.

Course and prognosis are also variable for children with acquired language problems. Some spontaneous recovery generally occurs rapidly within days or weeks of the neurological insult, but additional improvements may show a protracted course over months or even years. Approximately two-thirds of children with acquired disorders eventually recover language functioning that is broadly within the normal range, although subtle deficits may persist. Those who do not attain normal language functioning show persistent language and cognitive deficits of varying severity. Prognosis depends on many factors, including the severity, extent, and location of brain damage, the specific etiology, the age of onset, the extent of preinjury language development, and the existence of other preinjury disorders.

Language disorders also occur in children from diverse linguistic and cultural backgrounds. A difficult diagnostic challenge is to differentiate language errors that are a normal part of acquiring a new language from those that are indicative of disorder. A valid assessment usually requires the cooperation of a well-trained informant from the same cultural and linguistic background as the child. The informant assists in collecting and interpreting history information, assessing the child's skills in the other language, and discussing assessment results and recommendations with the family.

COURSE AND PROGNOSIS

Language disorder is characterized by considerable variability in severity, course, and outcome. Individuals who had mild to moderate language disorders are less likely to suffer from significant long-term mental health problems than individuals who had severe language disorders. Youngsters with early receptive language delays show higher rates of persisting

disorder and less favorable long-term outcomes than those with early expressive-only delays. Some, however, do eventually attain language functioning within the normal range. Better outcomes are generally associated with higher levels of early gestural communication, receptive language, and nonverbal intelligence.

Children with persistent language disorders continue to develop language comprehension and production skills during the preschool and school-age years. However, the rate of development may be quite slow, causing them to fall further behind peers who are developing language at a normal rate. Overall, the communicative prognosis for children with persistent language disorders continues to be poorer than for those with expressive-only disorders.

Children with language disorders are also at risk for academic problems once they enter school. A negative spiral of poor academic performance, low self-esteem, reduced motivation, and social isolation may stem from the language comprehension difficulties, particularly when they are coupled with attention or reading problems. Fortunately, some progress has been made in using performance on early language-related tasks to predict which kindergarten children may be at risk for later reading difficulties. Moreover, interventions to improve phonological awareness, a key skill for early reading, have proven successful with children with language disorders, at least in the short term.

Many adolescents and young adults with a history of language disorder continue to show problems in communicative effectiveness, academic performance, and social functioning. These problems may limit vocational choices and social participation. Individuals with language disorder are also at heightened risk of being bullied or victimized in general, including sexual assault or abuse for girls and women. Serious psychiatric problems, such as social anxiety disorder or antisocial personality disorder, may also develop in some individuals. The societal consequences are significant, as evidenced by the high rate of language difficulties in adults who live in poverty, commit crimes, incur psychiatric illnesses, and achieve limited literacy.

TREATMENT

Experts disagree about the timing of intervention for language disorder. Some advocate a wait-and-see attitude for young children with early expressive delays (late talkers), because most acquire normal language functioning during the preschool years. They argue that intervention should be provided only to those with persistent expressive delays at 4 or 5 years of age. Others suggest that early intervention might prevent or minimize later language, academic, and behavioral difficulties. Unfortunately, few relevant data are available on the efficacy of early- versus late-language intervention.

Direct and mediated interventions have been shown to promote short-term language growth in children with language disorders. In direct interventions, the primary intervention agent is a professional (usually a speech–language pathologist) who works directly with the child to improve language functioning. In mediated interventions, the primary intervention agents are parents, teachers, or paraprofessionals, who have been taught by the speech–language professional to deliver the intervention to the child. Mediators are encouraged to incorporate language facilitation techniques into daily activities and routines, thereby providing regular, naturalistic opportunities for children to use newly learned language features.

Further research on treatment efficacy is needed to address important outstanding issues. For example, studies are required to assess (1) the ways in which interventions might be tailored to address the specific language needs of individual children, (2) the long-term merits of early intervention for children with expressive delays, (3) the extent to which the benefits of intervention achieved in laboratory studies can also be demonstrated in routine community-based interventions, and (4) what the potential harms of early screening or intervention for language delays may be, including the risk involved in false-positive and false-negative results, labeling, parent and child anxiety, and unnecessary evaluation and intervention. One recent study of routine early interventions suggests that systematic efforts may be required to improve their effectiveness.

Children with language disorders are generally recommended for intervention to address their receptive and expressive language needs. Various language intervention techniques and programs have been developed, but some have not been rigorously evaluated to demonstrate their efficacy. The available efficacy literature, however, suggests that language intervention can foster short-term progress toward specific communicative goals, a conclusion that extends to various intervention techniques and programs. Unfortunately, there has been limited study thus far of the possible long-term benefits of language intervention in limiting or reversing the negative impacts of language disorder on linguistic, academic, and psychosocial outcomes.

Interventionists have, however, begun to focus on long-term, broad-based outcomes. Preschoolers with language disorders are receiving interventions designed to promote not only oral language outcomes but also social communication and early literacy skills. Kindergartners with language disorders are being taught certain key prereading skills, which may prevent the early reading failure that is linked to subsequent academic and psychosocial problems. More attention is being given to enabling peers, parents, and teachers to learn methods for fostering frequent, natural, and optimal communicative interactions with children who have language disorders. In schools, there are efforts to coordinate the various types of interventions that may be required by a child with a language

disorder (e.g., language intervention, special reading help, special education support, or behavioral interventions). Despite some initial promising results from these efforts, much remains to be learned concerning optimal goals, techniques, and programs for language intervention.

Language disorders may also necessitate modifications to other interventions that rely heavily on language as a medium (e.g., psychotherapy or cognitive behavioral modification). Comprehension can be facilitated by simplification of vocabulary, sentence structure, questions, and verbal instructions and by repetition of verbal information as required. Compensation for expressive difficulties can be achieved by using nonverbal response modes to supplement verbal responses. Ambiguous or unclear responses should also be clarified. Assessment procedures that depend heavily on language (e.g., projective tests or diagnostic interviews) may require similar modifications.

References

Ahmed ST, Lombardino LJ, Leonard CM. Specific language impairment: definitions, causal mechanisms, and neurobiological factors. *J Med Speech Lang Pathol*. 2001;9(1):1–15.

Bartlett CW, Flax JF, Logue MW, et al. A major susceptibility locus for specific language impairment is located on 13q21. *Am J Hum Genet*. 2002;71:45–55.

Baydar N, Küntay AC, Yagmurlu B, et al. "It takes a village" to support the vocabulary development of children with multiple risk factors. *Dev Psychol*. 2014;50(4):1014–1025.

*Beitchman JH, Brownlie EB. Language disorders in children and adolescents. In: Wedding D, ed. *Advances in Psychotherapy: Evidence-based Practice*. Cambridge, MA: Hogrefe; 2014.

Beitchman JH, Brownlie EB, Bao L. Age 31 mental health outcomes of childhood language and speech disorders. *J Am Acad Child Adolesc Psychiatry*. 2014;53(10):1102–1110.

Bishop DV. How does the brain learn language? Insights from the study of children with and without language impairment. *Dev Med Child Neurol*. 2000;42:133–142.

Bishop DV, Snowling MJ. Developmental dyslexia and specific language impairment: same or different? *Psychol Bull*. 2004;130(6):858–886.

*Brownlie E, Bao L, Beitchman J. Childhood language disorder and social anxiety in early adulthood. *J Abnorm Child Psych*. 2016;44:1061–1070.

*Buschmann A, Jooss B, Rupp A, et al. Children with developmental language delay at 24 months of age: results of a diagnostic work-up. *Dev Med Child Neurol*. 2008;50(3):223–229.

Cohen NJ, Vallance DD, Barwick M, et al. The interface between ADHD and language impairment: an examination of language, achievement, and cognitive processing. *J Child Psychol Psychiatry*. 2000;41:353–362.

Dale PS, Price TS, Bishop DV, Plomin R. Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *J Speech Lang Hear Res*. 2003;46(3):544–560.

Evans JL, Viele K, Kass RE, Tang F. Grammatical morphology and perception of synthetic and natural speech in children with specific language impairments. *J Speech Lang Hear Res*.

2002;45:494–504.

Falcaro M, Pickles A, Newbury DF, et al. Genetic and phenotypic effects of phonological short-term memory and grammatical morphology in specific language impairment. *Gen Brain Behav*. 2008;7(4):393–402.

Fujiki M, Brinton B, Clarke D. Emotion regulation in children with specific language impairment. *Lang Speech Hear Serv Sch*. 2002;33:102–111.

Gillon GT. The efficacy of phonological awareness intervention for children with spoken language impairment. *Lang Speech Hear Serv Sch*. 2000;31:126–141.

Lahey M, Edwards J, Munson B. Is processing speed related to severity of language impairment? *J Speech Lang Hear Res*. 2001;44:1354–1361.

Law J, Rush R, Schoon I, Parsons S. Modeling developmental language difficulties from school entry into adulthood: literacy, mental health, and employment outcomes. *J Speech Lang Hear Res*. 2009;52:1401–1416.

Leonard LB. *Children with Specific Language Impairment*. Cambridge, MA: Bradford; 1998.

*Mayes AK, Reilly S, Morgan AT. Neural correlates of childhood language disorder: a systematic review. *Dev Med Child Neurol*. 2015;57:706–717.

McArthur GM, Bishop DV. Which people with specific language impairment have auditory processing deficits? *Cogn Neuropsychol*. 2004;21:79–94.

McGregor KK, Oleson J, Bahnsen A, Duff D. Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *Int J Lang Commun Disord*. 2013;48(3):307–319.

Paul R. *Language Disorders from Infancy through Adolescence: Assessment & Intervention*. 2nd ed. St. Louis, MO: Mosby; 2001.

Prior M, Bavin E, Cini E, Eadie P, Reilly S. Relationships between language impairment, temperament, behavioural adjustment and maternal factors in a community sample of preschool children. *Int J Lang Commun Disord*. 2011;46(4):489–494.

Robinson RO, Baird G, Robinson G, Simonoff E. Landau-Kleffner syndrome: course and correlates with outcome. *Dev Med Child Neurol*. 2001;43:243–247.

Spinath FM, Price TS, Dale PS, Plomin R. The genetic and environmental origins of language disability and ability. *Child Dev*. 2004;75:445–454.

Tomblin JB, Hafeman LL, O'Brien M. Autism and autism risk in siblings of children with specific language impairment. *Int J Lang Commun Disord*. 2003;38:235–250.

*Tomblin JB, Zhang X, Buckwalter P, O'Brien M. The stability of primary language disorder: four years after kindergarten diagnosis. *J Speech Lang Hear Res*. 2003;46:1283–1296.

*Toppelberg CO, Shapiro T. Language disorders: a 10-year research update review. *J Am Acad Child Adolesc Psychiatry*. 2000;39(2):143–152.

Young AR, Beitchman JH, Johnson C, et al. Young adult academic outcomes in a longitudinal sample of early identified language impaired and control children. *J Child Psychol Psychiatry*. 2002;43:635–645.

▲ 43.2 Social (Pragmatic) Communication Disorder

INTRODUCTION

Social (pragmatic) communication disorder is a new diagnostic category introduced in DSM-5. The main characteristic of social (pragmatic) communication disorder is the persistent deficit in the social use of nonverbal and verbal communication and difficulty using the linguistic context to make inferences and to resolve ambiguities. The symptoms include deficits in understanding and applying social rules concerning verbal and nonverbal communication, difficulty adapting one's language according to the needs of the listener or situation, problems following the rules of conversations (such as turn taking) and storytelling, and difficulty understanding ambiguous and nonliteral language through integrating the linguistic context and prior knowledge. Social (pragmatic) communication disorder has significant overlaps in symptoms with autistic spectrum disorder and language disorder. However, unlike individuals with autism spectrum disorder, individuals with social (pragmatic) communication disorder do not exhibit clinically significant restricted interests or repetitive behaviors. Although social (pragmatic) communication disorder can cooccur with other communication disorders such as language disorder, the social communication and pragmatic deficits must exceed what would be expected from structural language deficits alone. The diagnosis of social (pragmatic) communication disorder cannot be made along with autism spectrum disorder.

HISTORY

Although social (pragmatic) communication disorder was not included until DSM-5, there has been a long-standing debate regarding the diagnostic status of children with deficient social communication and pragmatic skills. In fact, children who have social communication and pragmatic deficits have been documented in the speech and language literature for over 30 years. In the early 1980s the term semantic-pragmatic deficit syndrome was introduced to characterize children who demonstrated problems with verbosity, word-finding deficits, unusual word choice, and poor topic maintenance. A similar term introduced in the late 1980s was semantic-pragmatic disorder, which was thought to occupy the diagnostic space between autism spectrum disorder and specific language impairment. Both of these terms reflect the idea that semantic (i.e., the meaning of an expression) and pragmatic (i.e., the use of language in context) difficulties tend to cooccur. In the 1990s, however, research started to emerge showing that semantic deficits did not necessarily cooccur with pragmatic deficits, and that semantic abilities did not reliably differentiate between typically developing and clinical groups. As a result, the term pragmatic language impairment was coined and became the

generally accepted term used in the speech and language literature.

The addition of social (pragmatic) communication disorder to the DSM-5 is closely associated with the changes made to the diagnostic criteria concerning autism spectrum disorder. In DSM-IV-TR, the umbrella term, pervasive developmental disorders, was used to cover autistic disorder, Asperger syndrome, and pervasive developmental disorder—not otherwise specified. The pervasive developmental disorder—not otherwise specified was applied to individuals who were on the autism spectrum but did not meet the full criteria for either autistic disorder or Asperger syndrome. Therefore, children or adults who showed core deficits in one area such as social deficits, but did not meet the diagnostic criteria for other symptoms such as restrictive, stereotypical, and repetitive behaviors would have been given the diagnosis of pervasive developmental disorder—not otherwise specified. Research has shown that more than half of the children diagnosed with pervasive developmental disorder—not otherwise specified were given the diagnosis due to the fewer restrictive and repetitive behaviors they exhibited.

In DSM-5, the diagnostic category of autism spectrum disorder replaced the pervasive developmental disorders of DSM-IV-TR. The major reason for this change was due to the lack of reliable and meaningful differentiation in the subclassification under pervasive developmental disorders, even among experts. One of the consequences of this change was that the pervasive developmental disorder—not otherwise specified was removed as a diagnostic classification, and that the diagnosis of individuals who showed social interaction deficits but mild or no restrictive and repetitive behaviors were given the new diagnosis of social (pragmatic) communication disorder under the diagnostic classification of communication disorder in the section of neurodevelopmental disorders.

Social (pragmatic) communication disorder is thought to largely align with the pragmatic language impairment discussed in the research literature. However, various definitions of pragmatic language impairment exist, and the traditional definitions do not include impairments in nonverbal communication. Therefore, precaution needs to be taken when making inferences from the previous literature regarding the social (pragmatic) communication disorder.

DIAGNOSTIC FEATURES

Individuals with social (pragmatic) communication disorder have difficulties with pragmatics—the rules that govern social conventions of verbal and nonverbal language that allow people to relate to one another in socially appropriate ways. While most children pick up these social rules intuitively, children with social (pragmatic) communication disorder may find it difficult to learn the basic rules of social communication, such as when to start talking, when to listen, how to phrase a question, stay on

topic and know when the talk is over. As a result of deficits in pragmatics, individuals with social (pragmatic) communication disorder may struggle with holding conversations, may interrupt frequently, monopolize the conversation, or not talk at all. They may say things that are inappropriate or unrelated to the conversation at hand. This can make it difficult for individuals with social (pragmatic) communication disorder to develop meaningful social relationships, and communicate well with others at school or in the workplace. The symptoms of social (pragmatic) communication disorder include the following:

1. Problems using nonverbal and verbal language such as greeting and sharing information, in ways appropriate for the social context.
2. Impairment in changing communication to match the context or the need of the listener such as speaking differently in the classroom than on the playground, or talking to a child than to an adult.
3. Difficulties following social conventions of conversation and storytelling, such as taking turns in conversations, rephrasing when misunderstood, and knowing how to use verbal and nonverbal signals to regulate the interaction.
4. Difficulty making inferences when something is not explicitly stated, and difficulty understanding the meaning of nonliteral or ambiguous language, such as idioms, humor, and metaphors, and when there are multiple meanings depending on the context.

These deficits experienced by individuals with social (pragmatic) communication disorder impede effective communication, social participation, academic achievement, or occupational performance. Difficulties with social interactions are manifest early in life, although the impairments may not become fully apparent until social communication requirements become more demanding. The symptoms of social (pragmatic) communication disorder should not be attributable to another medical or neurological condition. The deficit in social communication should not be solely attributable to problems in the mechanics of language as is the case for individuals with language disorder. The deficits should not be better accounted for by intellectual disability, global developmental delay, autism spectrum disorder, or any other mental disorders.

As is the case with autism spectrum disorder, individuals with social (pragmatic) communication disorder may show little interest in social interactions. They may have difficulty adapting language to different listeners (for example, talking the same way to an adult and to a child), and to different situations (for example, talking the same way in a classroom as on the playground). They may have trouble making inferences and understanding things that are implied and not stated explicitly. They may not give background information when they are speaking to an unfamiliar person. They may have trouble properly greeting people, making requests,

or getting attention. They tend to be overly literal and not understand riddles or sarcasm. They may have difficulty understanding nonverbal communication such as reading others' facial expressions. Due to the difficulties in making inferences, understanding the nuances and social subtleties, some children may have trouble with reading comprehension, despite being able to develop phonemic awareness. As a result, their reading issues may not be noticed until mid-elementary school, when children move from learning to read to reading to learn.

Some experts suggest that social (pragmatic) communication disorder usually never occurs by itself. Language disorder is the most commonly occurring disorder associated with social (pragmatic) communication disorder.

Attention-deficit/hyperactivity disorder is a common comorbidity associated with social (pragmatic) communication disorder. The hyperactivity and inattention problems can aggravate the problems of pragmatic competence. In addition, a lack of social pragmatic skills may lead to children becoming frustrated and acting out.

Jonny was a 6-year-old boy who often spoke out of context and had great difficulty playing with other children. He would make associations with a word or phrase and continue to perseverate. He would respond inappropriately to questions, for example, in the midst of a totally unrelated discussion, he might shout: "The toilet is cracked and white stuff is coming out." One of his favorite statements was, "what a lovely bouquet," which he would repeat, usually without reference to anything that was happening at the time. People were puzzled by him and found it difficult to communicate with him. While he appeared able to make grammatically correct sentences, and had no evidence of problems with articulation or speech, he seemed to have little appreciation of the meaning of his statements or that his comments were often totally out of context.

N is a 17-year-old youth with a long history of physical confrontations with other youth. He presented with a flat affect and no facial expression. He was prone to misinterpret others' intentions and found himself easily provoked. Although he said he could discern others' emotions, his history of misjudging others' intentions contradicted this belief. His emotional reactions were extremely difficult to assess because of the absence of facial expression. He has had difficulty making friends and staying in relationships. He does not get along with peers his own age and seeks friendship with individuals 4 years older than him. He has difficulties in social situations and would like to change his interactions. He plans on ignoring the "sly friends" and only engaging in physical altercations if they provoke him first even if they are joking. There was no evidence of symptoms suggestive of an autism spectrum disorder, no history of current or past repetitive, restricted interests or behaviors.

As this vignette demonstrates, individuals with social (pragmatic) communication disorder may show signs of deficits in nonverbal communication despite relatively intact structural language.

DIFFERENTIAL DIAGNOSIS

Differential diagnosis should take into consideration the relationship between social (pragmatic) communication disorder and other psychiatric

disorders, especially autism spectrum disorder and language disorder. Many children referred to clinics for psychiatric or psychological assessment have been found to show some form of social communication or pragmatic language deficits. The rate of comorbidity is especially high for children with language disorder, attention-deficit/hyperactivity disorder, conduct disorder, and traumatic brain injury. Differential diagnosis can be challenging due to overlaps in symptoms among various psychiatric disorders and lack of well-validated assessment tools with high reliability.

Social communication is one of the core deficits of autism spectrum disorder. Children with autism spectrum disorder may show adequate structural language abilities, yet have difficulty with pragmatic language such as overly formal speech, verbosity, and turn taking in conversations. For children with autism spectrum disorder, these pragmatic language and social communication difficulties are part of a larger social deficit which includes restricted interests and repetitive behaviors. The past and/or present symptoms of clinically significant specialized interests and restricted, repetitive behaviors can help differentiate autism spectrum disorder from social (pragmatic) communication disorder.

Because a diagnosis of autism spectrum disorder can be given to individuals with just a history of fixed interests and repetitive behaviors, individuals who only show significant social and pragmatic language deficits but no present behavioral patterns characteristic of autism spectrum disorder may still be given a diagnosis of autism spectrum disorder. Complicating the differential diagnosis further, stereotyped language, which was considered one of the symptoms of pragmatic language impairment, is now considered part of the repetitive and restrictive behaviors required for giving a diagnosis for autism spectrum disorder. Finally, some children may show an increase in repetitive behaviors and specialized interests as they grow older, further complicating the differential diagnosis.

Children with language disorder show deficits in word knowledge, grammar, and discourse. As discourse includes conversational exchange and narrative, deficits in these areas tend to manifest in social communication problems. For children with language disorders, however, social and pragmatic language deficits are thought to be a result of, or secondary to, the structural language impairment. A strong association between language results and measures of social deficits would support the diagnosis of language disorder.

Some studies have shown group differences in symptom profiles of social communication and pragmatic deficits, although making differential diagnosis based on these symptom profiles may prove difficult. For example, children with attention-deficit/hyperactivity disorder as a group may show problems with inappropriate initiation (such as talking about a

topic that no one is interested in). One study examining children with William syndrome and children with autism spectrum disorder showed that children with William syndrome tend to have fewer problems with stereotyped language, nonverbal communication, social relations, and coherence compared to children with autism spectrum disorder. An observational study comparing children with social (pragmatic) communication disorder and children with language disorder has found that, as a group, children with social (pragmatic) communication disorder tend to have more problems with turn taking in conversations, provide no response or inappropriate responses to the communication at hand, and make little use of nonverbal signals compared to a group of children with language disorder. More work is needed examining the sensitivity and specificity of these social and pragmatic language profiles for the purpose of differential diagnosis, and how the risk factors associated with different clinical populations affect the course and prognosis of social and pragmatic language development.

There is overlap in symptoms between social (pragmatic) communication disorder and social anxiety disorder. Some children with severe social anxiety disorder may be misdiagnosed with autism spectrum disorder due to their apparent uneasiness in social situations and lack of eye contact. The timing of the onset of symptoms may help to differentiate between these two disorders. Children with social anxiety disorder have typically developing social communication skills, yet do not utilize them due to anxiety associated with social interactions. Children with social (pragmatic) communication disorder, on the other hand, would have never had effective social communication.

ETIOLOGY

There is no published research examining the heritability and environmental factors associated with social (pragmatic) communication disorder using the DSM-5 criteria. There is accumulating evidence, however, which suggests that a family history of autism spectrum disorder, language disorder, or specific learning disorder may increase the risk of having social communication and pragmatic language deficits. Studies are also emerging, associating the impairments in pragmatic language with some common gene variants. These gene variants, however, are not specific to pragmatic language impairments but are also found to be associated with other neurodevelopmental and psychiatric disorders including autism spectrum disorder, intellectual disorders, and language disorder. More research is needed examining the familial aggregation and genetic factors associated with social (pragmatic) communication disorder. Although the causes of social (pragmatic) communication disorder still remains unknown, several theories have been proposed regarding the underlying deficits related to social (pragmatic) communication disorder. The social

inference theory suggests that individuals with social (pragmatic) communication disorder, much like individuals with autism spectrum disorder and conduct disorder, may have deficits in social cognition, especially in theory of mind (ToM), which involves making inferences about a person's mental state and using that information to understand and predict the other person's intentions and behaviors. It is often theorized that this type of understanding and representation of other people's mental states precedes the development of social communication. However, it is also possible that an understanding of others' mental states develops as a result of social communication. A third possibility is that individuals with social communication problems have a core deficit in relating to others, which may affect both social communication and forming mental representations of others' intentions. The weak central coherence theory posits that social communication deficits may be associated with a more general processing problem of integrating different sources of information. This theory suggests that social communication deficits occur as a result of individuals with social communication problems narrowly focusing on inappropriate or less crucial aspects of the communication while ignoring the bigger picture, which in turn makes it difficult for them to use linguistic and social context to derive meaning. The executive dysfunction theory suggests that individuals with pragmatic language and communication issues have a more general issue with executive functioning. Executive functions are associated with (but not limited to) the activity of the prefrontal regions of frontal lobes of the brain, which is associated with high-level cognitive functioning such as impulse control, behavior monitoring, flexibility of thinking, abstract thinking, concept formation, problem solving, and planning. The executive dysfunction theory helps explain the great heterogeneity in pragmatic language difficulties experienced by individuals with social and pragmatic difficulties. While executive dysfunction has been implicated in psychiatric disorders such as autism spectrum disorder and attention-deficit/hyperactivity disorder which is associated with pragmatic deficits, more research is needed to examine the hypothesized causal mechanism between high-level processing and social communication or pragmatic language deficits among individuals with social (pragmatic) communication disorder.

EPIDEMIOLOGY

Although there are no reliable prevalence estimates for social (pragmatic) communication disorder, it is acknowledged by clinicians that children with social communication and pragmatic language deficits are one of the fastest growing groups being referred for speech and language interventions. How demographic factors such as age, gender, and socioeconomic backgrounds may correlate with the prevalence estimates

also remains to be investigated. One study in the United Kingdom found that 22 percent of the children attending a language clinic had pragmatic language impairment. However, due to the differences in diagnostic criteria, it is unclear what percentage of these children will also meet the diagnostic criteria for social (pragmatic) communication disorder.

Another study examining a sample of children with autism spectrum disorder found that roughly 1.5 percent of the sample showed signs of social communication deficits without meeting the criteria for fixed interest and repetitive activities and behaviors. Another estimate showed that over 50 percent of the children who were given the diagnosis of pervasive developmental disorder—not otherwise specified showed signs of social communication deficits without clinically significant specialized interest and repetitive behaviors. However, neither of these studies used the DSM-5 criteria for social (pragmatic) communication disorder and therefore the prevalence rate is likely to differ from that of social (pragmatic) communication disorder.

A field trial comparing the prevalence rate of autism spectrum disorder using the DSM-5 criteria and pervasive developmental disorders using the DSM-IV criteria showed a decrease in the prevalence of autism spectrum disorder in one site. Nevertheless, when the prevalence rate of autism spectrum disorder was combined with the prevalence rate of social (pragmatic) communication disorder, the prevalence rate was not significantly different from that of pervasive developmental disorder of DSM-IV. This suggests that the decrease, if any, in the estimated prevalence of autism spectrum disorder may be accounted for by the inclusion of social (pragmatic) communication disorder in DSM-5.

DEVELOPMENT AND COURSE

The diagnosis of social (pragmatic) communication disorder is rarely given before 4 years of age, as an adequate development of speech and structural language is necessary before making a diagnosis of social (pragmatic) communication disorder. By 4 or 5 years of age, most children should have developed sufficient social communication and language skills to make a diagnosis. Some milder forms of social (pragmatic) communication disorder may not be manifest until adolescence, when the social communication and language demands intensify. Although social (pragmatic) communication disorder requires a relatively intact structural language and impaired social communication, there is great heterogeneity in the severity and patterns of manifestation among those who are diagnosed with social (pragmatic) communication disorder. As a group, children with autism spectrum disorder tend to have more severe social communication and pragmatic language deficits than children with social (pragmatic) communication disorder. Children with language disorder, in turn, tend to show milder symptoms of social and pragmatic language

deficits than children with social (pragmatic) communication disorder. The outcome of social (pragmatic) communication disorder differs greatly. Some children's social and pragmatic skills improve considerably while others experience social deficits into adulthood. In general, studies suggest that children with pragmatic language difficulties tend to perform better in academics compared to children with structural language problems. Nevertheless, children with social communication deficits may struggle with written expression and reading comprehension. In addition, there is some evidence suggesting that pragmatic language difficulties in elementary school years may be associated with emotional and behavioral difficulties, higher rates of referral for psychiatric and psychological assessments, and longer duration attending special education. Social (pragmatic) communication disorder may continue into adulthood, making it difficult to form and maintain close social relationships.

Social communication problems rarely occur in isolation, absence of other psychiatric disorders. Cooccurring disorders, as well as other risk factors, are likely to affect the developmental course and prognosis of social (pragmatic) communication disorder.

TREATMENT

Social (pragmatic) deficits rarely occur without other cooccurring behavioral or emotional difficulties. These behavioral and emotional problems may serve as risk factors for further social communication difficulties, and social difficulties in turn may accentuate the cooccurring behavioral and emotional problems.

Several treatment options have been proposed for children with social communication and pragmatic language difficulties who also present behavioral/emotional problems. Several behavioral interventions and cognitive-behavioral treatment programs have been proposed, involving multiple collaborators that include speech–language therapists, mental health professionals, family members, teachers, and peers.

Functional communication training is a behavioral technique that has commonly been used to treat problem behaviors. After a functional analysis is conducted to identify the environmental trigger that evokes the problem behavior in various settings, socially acceptable communicative responses are taught, reinforced, and subsequently applied in different social settings. Cognitive-behavioral therapy is also a popular intervention technique used to modify maladaptive cognitive patterns, emotional responses, and behaviors.

High-quality, randomized controlled trials of social communication interventions specifically targeting social (pragmatic) communication disorder are very rare, however, and more research is required to examine the effectiveness of these treatments. One randomized controlled trial examined the social outcomes of children with social communication

deficits (both with symptoms of social (pragmatic) communication disorder and autism spectrum disorder) who received intensive, individualized, multifaceted interventions as compared to children who received treatment as usual. The intervention focused on developing (1) social interaction skills and social understanding, (2) verbal and nonverbal pragmatic skills, and (3) language processing skills. The children received up to 20 sessions from specialized speech–language professionals and trained assistants. The children who received the experimental intervention showed significant improvement in blind ratings of communication skills, teacher ratings of learning behaviors, and parental ratings of social and pragmatic skills. There were, however, no significant changes in the standardized assessment measures of social communication, which may be due to the lack of sensitivity of these measures to detect change.

Didactic training in social skills on its own, however, has met with little meaningful social changes. This is likely due to the fact that individuals with social (pragmatic) communication disorder do not generally have impaired knowledge of social norms and appropriate behaviors. Interventions targeting underlying deficits, developing strategies to minimize distracting information or make explicit important verbal and nonverbal cues may be more beneficial.

References

*Adams C, Lockton E, Freed J, et al. The Social Communication Intervention Project: a randomized controlled trial of the effectiveness of speech and language therapy for school-age children who have pragmatic and social communication problems with or without autism spectrum disorder. *Int J Lang Commun Disord*. 2012;47(3):233–244.

Alwell M, Cobb B. Social and communicative interventions and transition outcomes for youth with disabilities: a systematic review. *Career Dev Transit Except Individ*. 2009;32(2):94–107.

Gibson J, Adams C, Lockton E, Green J. Social communication disorder outside autism? A diagnostic classification approach to delineating pragmatic language impairment, high functioning autism and specific language impairment. *J Child Psychol Psychiatry*. 2013;54(11):1186–1197.

Lord C, Bishop SL. Recent advances in autism research as reflected in DSM-5 criteria for autism spectrum disorder. *Annu Rev Clin Psychol*. 2015;11:53–70.

*Martin I, McDonald S. Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain Lang*. 2003;85(3):451–466.

McTear MF. Pragmatic disorders: a question of direction. *Br J Disord Commun*. 1985;20(2):119–127.

Miller M, Young GS, Hutman T, Johnson S, Schwichtenberg AJ, Ozonoff S. Early pragmatic language difficulties in siblings of children with autism: implications for DSM-5 social communication disorder? *J Child Psychol Psychiatry*. 2015;56(7):774–781.

Murphy SM, Faulkner DM, Farley LR. The behaviour of young children with social communication disorders during dyadic interaction with peers. *J Abnorm Child Psychol*.

2014;42(2):277–289.
*Norbury CF. Practitioner review: Social (pragmatic) communication disorder conceptualization, evidence and clinical implications. *J Child Psychol Psychiatry*. 2014;55(3):204–216.
Oliver BR, Barker ED, Mandy WP, Skuse DH, Maughan B. Social cognition and conduct problems: a developmental approach. *J Am Acad Child Adolesc Psychiatry*. 2011;50(4):385–394.
*Swineford LB, Thurm A, Baird G, Wetherby AM, Swedo S. Social (pragmatic) communication disorder: a research review of this new DSM-5 diagnostic category. *J Neurodev Disord*. 2014;6(1):41.

▲ 43.3 Speech Sound Disorder

EMIKO KOYAMA, PH.D., AND JOSEPH H. BEITCHMAN, M.D.

INTRODUCTION

Children with a speech sound disorder do not correctly produce the speech sounds that are expected for their age, developmental stage, and dialect. Sounds may be omitted (e.g., saying *hou* for *house* or *dum* for *drum*), substituted for other sounds (e.g., saying *tup* for *cup* or *yewo* for *yellow*), or distorted because they are produced in an atypical way (e.g., saying /sh/ with too much air escaping from the sides of the mouth or producing /s/ or /z/ with the tongue protruded). Often, the speech sound errors occur in systematic patterns. For example, consonants that should involve a steady airflow instead have an interrupted airflow (e.g., *pat* for *pass* or *bacuum* for *vacuum*) or consonants that should appear at the end of words are instead omitted (e.g., *be* for *bed* or *ro* for *rose*). In severe cases of speech sound disorder, speech may be completely incomprehensible to others.

As with other aspects of language, children must gradually learn the regularities of phonology. Not surprisingly, then, speech sound disorder is similar in many ways to other communication disorders. The affected communication skills generally develop in a typical sequence, but with a delayed onset and slow rate of growth.

Linguistic and motor learning, as well as their integration and coordination, is essential for speech sound development. Linguistically, children acquire knowledge about the individual sounds that comprise a particular language and their permissible contrasts and combinations. Children also induce the regularities governing syllable and word stress, and the intonation patterns (melody) of speech. In other words, linguistic learning presumably involves gradual and implicit acquisition of a complex rule system that underlies the representation, perception, and production of speech. It is not yet clear whether such rules explain or merely describe the nature of phonological knowledge.

Speech sound development also involves considerable motor learning.

Speech production entails the quick, accurate, and coordinated actions of the many muscles that support breathing, voicing, and articulation of speech sounds. Eventually, adults learn to control the complex coordination of these fine motor activities well enough to enable the production of intelligible speech at rates as high as 250 words per minute.

Speech sound development normally proceeds in a systematic fashion. The first words that children use generally consist of simple sequences of consonants (C) and vowels (V) that comprise syllables of the forms CV (e.g., bye), VC (e.g., up), or CVC (e.g., dog) and short repetitive sequences of the form CVCV (e.g., dada). More complex syllable and word shapes are gradually mastered, such as those with multiple, nonrepetitive syllables and consonant cluster sequences, such as /thr-/, /kl-/, /-rst/, or /-nt/. Vowels are generally thought to be acquired early, and most studies suggest that vowel production is relatively accurate by 2 years of age. As such, numerous vowel errors made beyond age 2 years may be early warning signs of impaired speech. Early words generally include consonant sounds designated as the *early eight* (/m/, /b/, /y/, /n/, /w/, /d/, /p/, and /h/), which are also commonly used during the babbling stage from about 6 months of age. As development continues, the *middle eight* consonant sounds (/t/, /ng/, /k/, /g/, /f/, /v/, /ch/, and /j/) are mastered gradually, followed by the *late eight* (/sh/, /th/ [two variants], /s/, /z/, /l/, /r/, and /zh/). Children may not achieve full mastery of the late eight consonant sounds until 7 or 8 years of age. Throughout the course of typical development, children may produce speech sounds incorrectly and inconsistently. Three types of errors are common: omissions (e.g., *ha* for *hat*), substitutions (e.g., *fin* for *thin*), or distortions (e.g., a slurpy-sounding /s/).

Speech sound errors may also be described in terms of phonological processes or rules that encompass a pattern of errors. For example, young children often use a phonological process of final consonant deletion in which they regularly omit consonants at the end of words. This deletion process is used often by young children but usually disappears from their speech by 3 years of age. Similarly, there are other common processes that gradually disappear during normal development, including reduplication (repetition of a syllable to form a multisyllabic word, e.g., *bahbah* for *bottle*), unstressed syllable deletion (omission of a syllable that is not stressed, e.g., *nana* for *banana*), stopping (use of an interrupted airflow for sounds that should have a continuous airflow, e.g., *tee* for *see*), consonant harmony (e.g., using a sound similar to one that occurs elsewhere in a word, e.g., *tat* for *cat*), and consonant cluster reduction (omitting or simplifying sounds in a consonant cluster, e.g., *top* for *stop*). These phonological processes all occur regularly in the speech of young children but all are usually overcome before 5 years of age.

Even if a child mispronounces words, it may not affect the overall

comprehension of the child's utterances. If a child's spoken word or phrase cannot be understood, the child's speech is said to be *unintelligible*. Intelligibility is influenced not only by the produced speech sounds but also by other factors such as pragmatic context and the complexity of the sentence. In normal speech development, roughly half of children's speech is considered intelligible at about 2 years of age, roughly 75 percent is intelligible by 3 years of age, and by age 4 years, nearly all of the child's spoken language becomes intelligible. If less than two-thirds of the child's utterances are intelligible by age 4, clinical attention may be warranted.

COMPARATIVE NOSOLOGY

Speech sound disorder in the 5th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) largely corresponds to phonological disorder in the revised 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR). The designation of *speech sound disorder* is typically reserved for speech sound problems with no known cause, which are thought to reflect difficulties in learning the regularities of phonology (the sound system of a language). The 10th edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) provides the category of *specific speech articulation disorder* for speech problems of unknown etiology and designates a different category for speech impairment arising from sensory, structural, or neurological conditions. In the research literature, terms such as (developmental) phonological disorder, (functional) articulation disorder, developmental verbal dyspraxia, and childhood apraxia of speech are commonly used in association with speech sound disorder.

EPIDEMIOLOGY

The prevalence of speech sound disorder is at least 3 percent in preschoolers, 2 percent in 6- to 7-year-old children, and 0.5 percent in 17-year-old children. Estimates however vary widely in the research literature, reflecting the diverse criteria and samples that have been used across studies. Slightly more than half show problems only in speech sound production; the remainder also had language difficulties. Overall, speech sound disorder occurs far more frequently than speech impairment with known structural or neurological origins.

Speech sound production difficulties occur more often in boys. The literature includes estimates for sex ratios as high as 4:1 and as low as 1.5:1. Speech sound disorder also shows high rates of familial aggregation. Approximately 20 to 40 percent of children with speech sound disorder have at least one immediate family member who is also affected. There is evidence in favor of genetic contributions to familial aggregation, but environmental factors may also play a role. Familial risk factors such as low maternal education have also been identified as a potential risk factor.

ETIOLOGY

Speech sound disorder has no known cause. Proposals regarding causation typically focus on many of the same factors considered as possible causes for language disorders, including subtle brain abnormalities, genetic influences, environmental factors, neurodevelopmental immaturity, linguistic factors, and auditory processes. In addition, speech motor control and speech mechanism factors have also been entertained as possible causes.

Speech sound disorder most likely has multiple causal determinants. Although some evidence supports each proposed determinant, no single one is implicated in all cases of speech sound disorder. The idea that brain abnormalities may be responsible for speech sound disorders is supported primarily by indirect evidence, which includes reports of associated soft neurological signs at higher than normal rates and findings of structural or functional brain abnormalities in children with related disorders, such as language disorder and reading disorder. The evidence in support of genetic factors comes from a variety of methods, including pedigree, twin, adoption, and linkage studies. Recent studies have found susceptibility of speech sound disorder genes in chromosomes 3, 6, and 15, which are also thought to be candidate regions for dyslexia. The advancement of work on genetic influences will probably stimulate corresponding advances in the understanding of environmental determinants. At present, however, no conclusive evidence links specific environmental factors to speech sound disorders, although it is reasonable to assume that such factors contribute in as-yet unspecified ways. Neurodevelopmental immaturity is suggested by the fact that children with speech sound disorder usually follow the same sequence in speech sound development as children without the disorder, albeit at a slower rate. Additional indirect support comes from findings of soft neurological signs and subtle deficits in motor skills. Children with speech sound disorders may also have linguistic deficits that interfere with efficient learning of the regularities of phonology (e.g., permissible sounds, combinations, or positions of occurrence) from the available language input. Difficulties with auditory discrimination and perception or temporary hearing losses may also contribute to some cases of speech sound disorder. Children with speech sound disorder also show elevated rates of subtle abnormalities in the structure and function of the speech mechanism, but it is not yet known whether these abnormalities are causes or just correlates of the disorder.

DIAGNOSIS AND CLINICAL FEATURES

A diagnosis of speech sound disorder is given when a person experiences persistent difficulties in speech sound production due to different degrees of impairment in phonological knowledge and coordination of movements

involved in speech. The speech sound production is impaired for what is expected based on the speaker's age and dialect, and cannot be attributed to medical or neurological conditions such as cerebral palsy, cleft palate, traumatic brain injury or hearing impairment. A further criterion is that the speech sound difficulties must interfere with social communication, academic achievement, or occupational success. Speech sound disorder ranges in severity from mild, in which errors occur on one or two sounds and speech intelligibility is hardly affected, to severe, in which errors occur on most sounds and speech intelligibility is significantly impaired.

A delayed onset and slow rate of speech sound development are typical of most children with speech sound disorders. However, the course and sequence of development usually resemble those of typically developing children. Simple syllable and word shapes are gradually elaborated to include more complex shapes. The rate of acquisition may be slow, but the early eight consonant sounds usually develop first and are followed, in turn, by the middle eight and then the late eight consonant sounds. The errors in speech sound production resemble those that occur in normal development, but the relative distribution of omissions, substitutions, and distortions may be somewhat different. Children with speech sound disorder generally use phonological processes that also occur in normal development but may use these processes well beyond the age at which they normally disappear. Children with speech sound disorder also sometimes use unusual or idiosyncratic phonological processes that are not observed in normal development.

Ramon was a talkative, likable 3-year-old with virtually unintelligible speech, despite his normal hearing and language comprehension skills. Ramon's level of expressive language development could not be determined because his speech was so difficult to understand. The rhythm and melody of his speech, however, suggested that he was trying to produce multiword utterances, as would be expected at his age. Ramon produced only a few vowels (/ee/, /ah/, and /oo/), some early developing consonants (/m/, /n/, /d/, /t/, /p/, /b/, /h/, and /w/), and limited syllable shapes (V, CV, and CVCV). This reduced sound repertoire made many of his spoken words indistinguishable from one another (e.g., he said *bahbah* for *bottle*, *baby*, and *bubble*, and he used *nee* for *knee*, *need*, and *Anita* [his sister]). Moreover, he consistently omitted consonant sounds at the end of words and in consonant cluster sequences (e.g., /tr-/, /st-/, /-nt/, and /-mp/). On occasion, Ramon reacted with frustration and tantrums to his difficulties in making his needs understood.

Kent was a pleasant, cooperative 5-year-old boy, who had been diagnosed with a speech sound disorder when he was a preschooler. His hearing and language comprehension skills were within normal limits. However, he showed some mild expressive language problems in the use of certain grammatical features (e.g., pronouns, auxiliary verbs, and past-tense word endings) and in the formulation of complex sentences. He correctly produced all vowel sounds and most of the early developing consonants but was inconsistent in his attempts to produce later-developing consonants (e.g., /r/, /l/, /s/, /z/, /sh/, /th/, and /ch/). Sometimes, he omitted them; sometimes, he substituted other sounds for them (e.g., /w/ for /r/ or /f/ for /th/); occasionally, he even produced them correctly. Kent had particular problems in correctly producing consonant cluster sequences and multisyllabic words. Cluster sequences had omitted or incorrect sounds (e.g., *blue* might be produced as *bue* or *bwue*, and *hearts* might be

said as *hots* or *hars*). Multisyllabic words had syllables omitted (e.g., *efant* for *elephant* and *getti* for *spaghetti*) and sounds mispronounced or even transposed (e.g., *aminal* for *animal* and *lemon* for *melon*). Strangers were unable to understand approximately 80 percent of Kent's speech. Fortunately, Kent often spoke more slowly and clearly than usual when he was asked to repeat something, as he often was.

Natasha was a shy, reserved 8-year-old girl, with a history of significant speech delay. During her preschool and early school years, she had overcome many of her earlier speech errors. A few late-developing sounds (/r/, /l/, and /th/), however, continued to pose a challenge for her. Natasha often substituted /f/ or /d/ for /th/ and produced /w/ for /r/ and /l/. Overall, her speech was easily understood despite these minor errors. Nonetheless, she was often reluctant to speak in front of others because of the teasing she received from her classmates about her speech.

Other language, learning, and developmental disorders may cooccur with speech sound disorder. Language disorder, especially deficits in expressive language, frequently cooccur with speech sound disorder. Children who have both speech sound disorder and language disorder are at increased risk for other problems, such as reading and other learning disorders, attention-deficit hyperactivity disorder, and perhaps, developmental coordination disorder. Children with isolated speech sound disorders are at a relatively small risk for severe academic, behavioral, or psychiatric disorders. Children who continue to display severe speech errors by the time they enter school may, however, be at increased risk of developing reading and spelling problems.

PATHOLOGY AND LABORATORY EXAMINATION

There are no laboratory tests that can provide a diagnosis of speech sound disorder. When necessary, specific laboratory tests may help identify and document an underlying neurological condition that is responsible for apraxia or dysarthria.

DIFFERENTIAL DIAGNOSIS

The most important task in diagnosis of speech sound disorder is to judge if speech sound production is inappropriate for the age and dialect of the speaker. The judgment may be challenging because the speech of children with speech sound disorder often resembles that of younger children who are developing normally. The same speech errors that may indicate disorder in a 6-year-old child may also indicate typical development in a 4-year-old child. Thus, it is critical to compare carefully a particular child's pattern of speech sound errors to the normal patterns of speech sound development in children.

One must also carefully consider the dialect spoken by the child. Sound differences that may be quite acceptable in one dialect may suggest disorder in another. For example, the production of *caht* for *cart* is acceptable for speakers from Boston but not for those from most other

parts of North America.

One of the challenges facing clinicians is how to diagnose speech sound disorders in bilingual children. Depending on such factors as the child's first language background and the timing of acquisition of the second language, even typically developing bilingual children may exhibit more frequent and atypical phonological error patterns compared to their same-age monolingual peers. Research is needed to establish the typical and atypical speech patterns of children acquiring a second language. Meanwhile, an assessment of speech sound production skills in English and the child's other language, accompanied by a careful history, may lead to a correct diagnosis.

The diagnostic assessment should also include a hearing test and an oral examination. The hearing test identifies temporary or permanent hearing difficulties that may require medical or other interventions to enhance the prognosis for the development of understandable speech. The results from the oral examination may help clarify whether the speech sound production problems are developmental, structural, or neurological in origin. For example, children with neurological conditions may be slow or inaccurate in performing various speech and nonspeech movements (e.g., sequencing syllables, as in *putuku*; protruding the tongue; or rounding the lips) or may show signs of muscle weakness or paralysis. Drooling, feeding, or swallowing problems, suggestive of dysarthria, may also be observed or reported in the history. During the oral examination, the diagnostician may also identify structural abnormalities in the mouth, such as tumors or submucous clefts of the palate, which may need to be considered in overall management of the child.

Selective mutism is an anxiety disorder that is characterized by the consistent failure to speak in specific social settings. Children with speech sound disorder may not speak due to embarrassment about their speech impairments. However, children with selective mutism typically exhibit normal speech in settings where they feel "safe," such as with family or close friends.

Screening for possible cooccurring communication disorders is also important for children diagnosed with speech sound disorders. When intellectual developmental disorder is suspected, a cognitive assessment may help confirm or rule out this diagnosis. In cases in which speech difficulties are in excess of those expected for a particular degree of intellectual disability, a diagnosis of cooccurring speech sound disorder may be made.

A practitioner should make a referral to a speech–language pathologist for assessment of speech sound production skills when (1) a 2 year old uses little or no speech, (2) a 3 year old is poorly understood by familiar members of the family, (3) a 4 year old is poorly understood by peers and strangers, (4) a child of any age seems to be withdrawing from social

interactions as a result of teasing or frustration associated with speech difficulties, (5) a parent or teacher is concerned about a child's speech, or (6) a child (or adult) experiences a sudden deterioration in speech production skills. The last may be the first warning sign of a developing neurological condition and therefore also warrants referral for neurological examination.

COURSE AND PROGNOSIS

Diagnosis may occur as early as 2 years of age in severe cases of speech sound disorder, which may be accompanied by cooccurring language involvement. However, diagnosis around 4 years of age is more usual. Speech sound disorders show considerable variability in severity, course, and outcome. In approximately 75 percent of mild-to-moderate cases, speech sound production skills normalize by 6 years of age, with or without intervention. Thus far, no specific predictors of normalization have been identified from among the various speech, hearing, oral, cognitive, linguistic, or psychosocial variables that have been investigated. For children whose speech has not normalized by 6 years of age, speech sound development proceeds in spurts and plateaus until 8 or 9 years of age. If normalized speech has not been achieved by then, residual articulation problems often persist into adulthood unless successfully treated. These residual errors usually involve distortions of later-developing sounds, such as /l/, /r/, /s/, /z/, /sh/, /ch/, and /th/. Distortions of the sounds /s/, /z/, /sh/, and /ch/ are termed *lisps*. Residual errors often have little effect on speech intelligibility but may influence others to erroneously perceive a speaker as lazy, immature, or of low intelligence.

Children with mild-to-moderate speech sound disorders are at relatively little risk for adverse long-term outcomes. Although speech typically normalizes by 8 to 9 years of age, subtle deficits in the speed and accuracy of articulation during complex speech tasks (e.g., rapid pronunciation of multisyllabic words) may remain detectable into adulthood but have little functional impact. Overall, language, cognitive, academic, psychosocial, and psychiatric outcomes are usually within the normal range for adults with a history of mild-to-moderate speech sound disorder.

Children with severe speech sound disorders that persist to school age are at a greater risk for poor outcomes. They may not develop adequate phonological awareness skills (perception, identification, and mental manipulation of speech sounds), which, in turn, may lead to problems in learning to read and spell. Parents and teachers should monitor carefully the reading and spelling development of school-age children with severe, persistent speech sound problems. If required, reading and spelling performance can then be fostered through explicit instruction in phonological awareness. It is not yet clear whether severe speech sound disorders are associated with increased risk for adverse psychosocial or

psychiatric outcomes.

Children with cooccurring language disorders are, however, known to have elevated risk for adverse long-term outcomes in communicative, academic, psychosocial, or psychiatric functioning. Those with receptive difficulties are at the greatest risk.

TREATMENT

Treatment is typically recommended for children with moderate-to-severe speech sound disorders. A speech–language pathologist may provide the treatment directly to the child or indirectly through paraprofessionals, parents, teachers, peers, or other volunteers who are taught how to provide treatment for the child. Although there are many intervention techniques, styles, and programs that have been advocated, two main approaches to treatment are discernible.

The *traditional approach* is typically used with children who produce substitution or distortion errors for just a few sounds. In this approach, the child participates in activities that highlight the auditory and motor characteristics of a particular sound. The child then practices the production of the sound, while the clinician provides feedback and cues concerning the correct placement and movement of the articulators. Treatment is discontinued once the child produces the new sound accurately in spontaneous speech.

A *phonological approach* is usually chosen for children with patterns of multiple speech sound errors that are described by phonological processes (e.g., final consonant deletion or consonant cluster reduction). In this approach, the child participates in activities that demonstrate the communicative value of correctly using certain classes of sounds, such as final consonants or consonant clusters. The child then practices the production of various final consonants or clusters in meaningful communicative situations until the targeted groups of sounds are incorporated into the child's phonological system. Treatment is discontinued once the child uses the new classes of sounds correctly and consistently in everyday speech.

Traditional and phonological approaches have proven effective in accelerating the rate of speech sound development. However, not all variations of these approaches have yet been subjected to rigorous investigations of their efficacy. Additional research is also needed to compare the relative merits of various techniques for children with different profiles of speech errors, although some work of this nature has begun to appear in the literature. Overall, treatment success depends on whether the new speech skills become established as automatic patterns of production. A key factor in achieving this goal may be the extent to which a child or adolescent is motivated to change existing speech patterns.

In addition to speech sound intervention, children with cooccurring

language disorder may need interventions to promote language or fluency. Few studies have systematically evaluated the efficacy of such joint interventions. It is possible that speech production gains may occur as an additional benefit of language intervention that did not focus specifically on speech sound production.

REFERENCES

*Baker E, McLeod S. Evidence-based practice for children with speech sound disorders: Part 1 narrative review. *Lang Speech Hear Serv Sch*. 2011;42(2):102–139.

Bauman-Waengler J. *Articulatory and Phonological Impairments: A Clinical Focus*. 2nd ed. Needham Heights, MA: Allyn & Bacon; 2004.

*Beitchman JH, Brownlie EB, Bao L. Age 31 mental health outcomes of childhood language and speech disorders. *J Am Acad Child Adolesc Psychiatry*. 2014;53(10):1102–1110.

Beitchman J, Jiang H, Koyama E, et al. Models and determinants of vocabulary growth from kindergarten to adulthood. *J Child Psychol Psychiatry*. 2008;49(6):626–634.

*Bowen C. *Children's Speech Sound Disorders*. 2nd ed. Somerset, NJ: Wiley-Blackwell; 2014.

Campbell TF, Dollaghan CA, Rockette HE, et al. Risk factors for speech delay of unknown origin in 3-year-old children. *Child Dev*. 2003;74:346–357.

Dodd B, Bradford A. A comparison of three therapy methods for children with different types of developmental phonological disorder. *Int J Lang Commun Disord*. 2000;35:189–209.

Felsenfeld S. Finding susceptibility genes for developmental disorders of speech: the long and winding road. *J Commun Disord*. 2002;35:329–345.

Gierut JA. Treatment efficacy: functional phonological disorders in children. *J Speech Lang Hear Res*. 1998;41:S85–S100.

Hesketh A, Adams C, Nightingale C, Hall R. Phonological awareness therapy and articulatory training approaches for children with phonological disorders: a comparative outcome study. *Int J Lang Commun Disord*. 2000;35:337–354.

Johnson CA, Weston AD, Bain BA. An objective and time-efficient method for determining severity of childhood speech delay. *Am J Speech Lang Pathol*. 2004;13:55–65.

Johnson CJ. Getting started in evidence-based practice for childhood speech-language disorders. *Am J Speech Lang Pathol*. 2006;15:20–35.

Kent RD. Research on speech motor control and its disorders: a review and prospective. *J Commun Disord*. 2000;33(5):391–427.

Klein ES. Phonological/traditional approaches to articulation therapy: a retrospective group comparison. *Lang Speech Hear Serv Sch*. 1996;27:314.

*Law J, Garrett Z, Nye C. Speech and language therapy interventions for children with primary speech and language delay or disorder. *Cochrane Database Syst Rev*. 2003;CD004110.

Lewis BA, Freebairn LA, Hansen A, Taylor HG, Iyengar S, Shriberg LD. Family pedigrees of children with suspected childhood apraxia of speech. *J Commun Disord*. 2004;37:157–175.

MacDermot KD, Bonora E, Sykes N, et al. Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am J Hum Genet*. 2005;76:1074–1080.

Morrisette ML, Gierut JA. Lexical organization and phonological change in treatment. *J Speech Lang Hear Res*. 2002;45:143–159.

Munson B, Bjorum EM, Windsor J. Acoustic and perceptual correlates of stress in nonwords produced by children with suspected developmental apraxia of speech and children with phonological disorder. *J Speech Lang Hear Res*. 2003;46:189–202.

Rice ML, Smith SD, Gayán, J. Convergent genetic linkage and associations to language, speech, and reading measures in families of probands with Specific Language Impairment. *J Neurodev Disord*. 2009;1(4):264–282.

Rvachew S, Chiang PY, Evans N. Characteristics of speech errors produced by children with and without delayed phonological awareness skills. *Lang Speech Hear Serv Sch*. 2007;38(1):60–71.

Rvachew S, Nowak M. The effect of target-selection strategy on phonological learning. *J Speech Lang Hear Res*. 2001;44:610–623.

Schick JH, Kundtz AM, Taylor HG, et al. Evidence of linkage for speech impairment with chromosome 7 markers. *Am Soc Human Gen*. 2001.

Shriberg LD, Austin D, Lewis BA, McSweeny JL, Wilson DL. The speech disorders classification system (SDCS): extensions and lifespan reference data. *J Speech Lang Hear Res*. 1997;40:723–740.

Smith SD. Genes, language development, and language disorders. *Ment Retard Dev Disabil Res Rev*. 2007;13:96–105.

Tyler AA, Lewis KE, Haskill A, Tolbert LC. Efficacy and cross-domain effects of a morphosyntax and a phonology intervention. *Lang Speech Hear Serv Sch*. 2002;33:52–66.

Tyler AA, Lewis KE, Welch CM. Predictors of phonological change following intervention. *Am J Speech Lang Pathol*. 2003;12:289–298.

*Verdon S, McLeod S, Wong S. Reconceptualizing practice with multilingual children with speech sound disorders: people, practicalities and policy. *Int J Lang Commun Disord*. 2015;50(1):48–62. doi:10.1111/1460-6984.12112

Yavas M, Goldstein B. Phonological assessment and treatment of bilingual speakers. *Am J Speech Lang Pathol*. 1998;7:49–60.

Yorkston KM. Treatment efficacy: dysarthria. *J Speech Hear Res*. 1996;39:S46–S57.

▲ 43.4 Stuttering

ROBERT M. KROLL, PH.D., AND JOSEPH H. BEITCHMAN, M.D.

DEFINITION

Stuttering is a condition in which the flow of speech or fluency is disrupted by involuntary speech motor events. Stuttering may include a variety of specific dysfluency types, such as sound and syllable repetitions, sound prolongations, dysrhythmic phonations, and complete blocking or unusual pauses between sounds and syllables of words. More severe forms of stuttering are frequently accompanied by a variety of struggle behaviors, commonly referred to as *accessory*, or *secondary*, features of the disorder. These struggle behaviors may be observed at the respiratory, phonatory, or articulatory level of the speech mechanism and may manifest themselves in

disordered breathing, glottal fry (abnormal voice quality), lip pursing, and tongue clicking. Additional behaviors, such as eye blinks, facial grimacing, head jerks, and abnormal body movements may be observed before or during instances of disrupted speech. In addition to the core and secondary physical symptoms of stuttering, the intended communication is often altered or distorted by linguistic avoidance due to anticipatory sound or word fears. Linguistic avoidance is a commonly observed feature of stuttering, which results in word substitution and circumlocution as the individual approaches the feared sound or word during the communicative attempt. Furthermore, it has been reported that varying levels of both utterance length and temporal complexity function to modulate stuttering frequency in adult stuttering speakers. Similarly, situational avoidance is often characteristic of chronic stuttering, and the type of communicative situation often influences the amount and severity of stuttering. Stuttering patterns are extremely variable, ranging from mild, occasional dysfluencies in speech to severe stuttering that renders communication all but impossible for the affected individual and often results in serious psychological difficulty and social withdrawal. Moreover, the person who stutters may exhibit widely fluctuating periods of fluency and dysfluency during the course of a day, depending on the communicative situation, anxiety level, fatigue level, and many other associated variables. It is to be noted that stuttering severity is not solely a measure of the frequency of observed stuttering instances, but is usually a function of the communicative hardship that it places on the speaker. In other words, an individual employed in an office setting who exhibits a stuttering rate of 15 percent may feel less handicapped than an airline pilot who is 4 percent dysfluent but whose fluent communication is critical. Most recently, treatment efficacy studies of stuttering have begun to utilize measures of functional communication and quality of life to supplement traditional data gathered from counts of stuttering frequency.

HISTORY AND COMPARATIVE NOSOLOGY

Psychiatric syndromes can be defined from the point of view of the observer or the patient. Depressed individuals may not look depressed, but if they report subjective symptoms sufficient to meet criteria, a diagnosis of a depressive disorder can be made. Likewise, stuttering can be defined from the point of view of the listener or the speaker. The speaker's sense of subjective distress or loss of control of the speech apparatus reflects an approach to nosology that focuses on the speaker's experience. The *Diagnostic and Statistical Manual of Mental Disorders* (DSM), on the other hand, focuses on the observer's point of view—that is, on the observable behaviors.

In both DSM-III and DSM-III-R, stuttering was placed in the category of disorders usually first evident in infancy, childhood, or adolescence. In

both of these systems, stuttering was coded on axis I, along with clinical psychiatric syndromes. This remains controversial, as stuttering is not commonly viewed as a psychiatric disorder among speech–language professionals, and speech–language pathologists have argued that it should not be included in DSM at all.

In DSM-IV-TR, stuttering is categorized as a communication disorder. This is consistent with the speech–language pathology classification, in which all speech and language disorders are kept within the same category. However, it remains to be shown that it belongs in this category; aside from the similarity of causing communicative difficulties, stuttering may differ fundamentally from the other communication disorders in this category. The 10th revision of the *International Statistical Classification of Disease and Related Health Problems* (ICD-10) separates stuttering (stammering) from the speech and language disorders and places it in the category of other behavioral and emotional disorders with onset usually occurring in childhood and adolescence. The diagnostic criteria in ICD-10 are similar to the DSM-IV-TR diagnostic criteria, but a few differences are notable.

DSM-IV-TR requires that the disorder interfere with academic or occupational achievement or social communication, whereas the ICD-10 diagnosis applies only if it is severe enough to cause marked disturbances in fluency. ICD-10 requires a minimum duration of 3 months; DSM-IV-TR has no duration criterion. DSM-IV-TR further specifies that the disturbance must exceed what is appropriate for the individual's age, and, if a speech–motor or sensory deficit is present, the speech difficulties must exceed those usually associated with these problems. ICD-10 has an additional category for cluttering, which is not accorded separate status in DSM-IV-TR.

According to the DSM-5, stuttering is defined by a disturbance in the normal fluency and patterning of speech that are inappropriate for the individual's age and language skills, persist over time and are characterized by frequent and marked occurrences of one (or more) of the following: sound and syllable repetitions; sound prolongations of consonants as well as vowels; broken words (e.g., pauses within a word); audible or silent blocking (filled or unfilled pauses in speech); circumlocutions (word substitutions to avoid problematic words); words produced with an excess of physical tension; monosyllabic whole-word repetitions (e.g., "I-I-I see him").

The disturbance causes anxiety about speaking or limitations in effective communication, social participation, or academic or occupational performance, individually or in any combination.

The onset of symptoms is in the early developmental period. The disturbance is not attributable to a speech–motor or sensory deficit, dysfluency associated with neurological insult or another medical condition, or better explained by another mental disorder.

EPIDEMIOLOGY

Surveys conducted chiefly in the United States and Europe indicate that the prevalence of stuttering in the general population is approximately 1 percent. Although data regarding incidence vary widely, most specialists would agree that the likelihood of an individual ever exhibiting some form of stuttering is between 3 and 4 percent. Of those who begin to stutter, the majority stop by adulthood. The estimated percentages of spontaneous recovery in stuttering, again, vary widely because of differing definitions of stuttering and normal nonfluency, age groups studied, and data collection techniques. Several studies have estimated spontaneous recovery to be as high as 80 percent. Stuttering is more prevalent among boys than girls, with the ratio of boys to girls among US schoolchildren being approximately 3:1. Current studies have challenged the above estimates of prevalence, incidence, and gender ratios. Among other factors, it has been postulated that most of the risk for stuttering onset is over by age 5. Even in its earliest stages, stuttering has been shown to be both qualitatively as well as quantitatively different from normal dysfluency.

A high familial incidence is indicated by the presence of stuttering among the relatives of approximately 50 percent of individuals who stutter. Additionally, a high degree of concordance of stuttering exists in identical twins. Individuals who stutter appear to have normal developmental histories, except for speech and language. Numerous studies have demonstrated that young stuttering children are somewhat delayed in early language and articulation development; however, studies of older stuttering children appear to demonstrate that language skills are once again age appropriate. With regard to personality and social and emotional adjustment, children who stutter are often regarded by their parents or primary caregivers as nervous and prone to specific fears. They are often described as overly sensitive to their environment. Adults who stutter tend to be somewhat more anxious and emotionally reactive than their nonstuttering peers. For the most part, studies relating intellectual ability to stuttering are inconclusive; however, there is a significantly higher prevalence of stuttering in the developmentally delayed population.

ETIOLOGY

There has been growing agreement by researchers and clinicians that stuttering is a multifactorial disorder. Specifying the level of causation can facilitate understanding the etiology of stuttering.

The first of these levels pertains to underlying causation, which refers to the genetic, neurophysiological, and psychological factors that may predispose a child to fluency disruption. Research in these areas has been rather cyclic since the 1940s, and various models implicating psychological, learning, and neurophysiological variables in stuttering have been

presented.

One early psychoanalytic theory of stuttering postulated that disrupted speech represents part of the individual's psychological history, such as disturbed early parent–child interaction. Empirical support for this psychoanalytic model is largely lacking, and studies investigating personality factors associated with stuttering have not revealed anything suggesting emotional disturbance or neurosis as a possible cause.

Later psychological models of stuttering have drawn on the principles of classical and operant conditioning to explain stuttering. Several behaviorally oriented researchers offered a model of operant conditioning in which the early dysfluencies of the child were subject to an intermittent schedule of reinforcement, resulting in incipient stuttering. The two-factor theory of stuttering posited that fluency failure resulting from negative emotion becomes associated with speech-related stimuli through a process of classical conditioning. The individual's response to the fluency failure, the repetitions, sound prolongations, and attempts to terminate the instance of disruption are instrumentally conditioned. Perhaps the best known of modern psychological theories of stuttering is the diagnosogenic (semantogenic) model. This theory postulates that stuttering is caused by parents' diagnoses of the developing child's normal dysfluencies as stuttering.

The large proportion of stuttering individuals who have relatives with a similar speech disorder and the high concordance for stuttering in identical twins have long suggested that biological inheritance plays some part in stuttering. Current models of genetic transmission that can account for familial patterns of stuttering provide strong evidence for the biological inheritance of stuttering. Moreover, research utilizing the most current genotyping technologies has confirmed the evidence for a strong genetic factor in stuttering. However, these genetic models point to a complex interaction between heredity and environment. Most recent studies have pointed out several possible specific candidate genes. However, it is too early to declare that specific genes underlying stuttering in the general population have in fact been identified. Moreover, at this point in time, it remains unclear as to which specific combinations of motor, linguistic, temperament, or other aspects of speech are being transmitted.

The theory of cerebral dominance, proposed in the late 1920s, was one of the earliest neurophysiological models used to explain the cause of stuttering. This theory postulated that stuttering resulted from the lack of an appropriately developed hemispheric dominance of speech and language function.

The theory of incomplete or atypical cerebral lateralization for speech has received varying amounts of support, but not until relatively recently did advances in electrophysiological techniques and brain imaging begin to yield more definite data on cortical and subcortical processing of speech in

stuttering. Much of the earlier research into the nature of these hypothesized brain processes was based largely on the use of behavioral observations and electromyographic measures. Relevant observations included the effects of auditory delayed feedback on the speech fluency of nonstuttering and stuttering speakers; differences in dichotic listening and dual task interference performance between stuttering and nonstuttering speakers; and articulatory discoordination and oral kinesthetic deficiencies in stuttering individuals. These observations, among others, represented indirect evidence of differences in neural processing between stuttering and nonstuttering individuals.

Most recently, several brain-imaging techniques used to measure regional cerebral blood flow have been applied to the investigation of stuttering. Data from these initial studies using single-photon emission computed tomography (SPECT) and positron emission tomography (PET) point to differential activation patterns in stuttering during both oral and silent reading and even immediately after treatment for stuttering. Findings from these studies have reported a general overactivation in several regions of cortical and subcortical motor systems in individuals who stutter. Recent functional imaging studies have reported increased activation in the right frontal-lateral cortex and the medial-anterior cingulate during silent and oral reading in stuttering individuals. Moreover, the neural activation in stuttering adults is characterized by a more generalized bilateral recruitment of cortical and subcortical neural resources involved in both language formulation and speech production.

Data coming from studies investigating effects of treatment seem to suggest that behavioral treatment initially results in an even further increase in recruitment of resources, including cerebellar activity, after an intensive 3-week treatment period, followed by a significant reduction in the overall amount of activation after 1 year. At the 1-year mark, the overall activation pattern observed in the stuttering speakers is very similar to that observed in the normally fluent speakers. This finding has been interpreted as signaling increased automaticity in the application of learned speech motor skills.

A recent study attempted to establish whether disconnections between speech-related cortical areas were the structural basis for stuttering. The investigation reported that stuttering subjects showed disturbed patterns of activation in the premotor and motor cortices, particularly below the laryngeal and tongue representations in the left sensorimotor cortex. Further evidence pointing to anomalous anatomical cortical structures in stuttering has been reported in a study using volumetric MRI scans. The study found that anomalous anatomical configurations of perisylvian cortical speech–language areas are more common in adults with stuttering. In addition to several other findings, the authors report that both the right and left planum temporale were significantly larger in the stuttering

subjects compared to the controls, and planar asymmetry was reduced, implicating speech and language laterality. The authors point out that the left planum temporale is composed of auditory association cortex, which is important in higher-order processing of linguistic information and is structurally larger than its homologous counterpart in healthy adults, thus the resulting leftward asymmetry. Researchers have speculated that the left planum temporale may be the neuroanatomical substrate for language, and there is some evidence of a relationship between planar asymmetry and language laterality. Additional structural anomalies have been reported in recent studies examining the neuroanatomical differences in speech-related cortex between stuttering and nonstuttering individuals using voxel-based morphometry. Results revealed significant differences in localized gray and white matter densities of left and right hemisphere regions involved in auditory processing and speech production.

The past decade has seen an exponential growth in the numbers of peer-reviewed articles investigating both functional and structural brain differences in stuttering individuals. While there appears to be much in the way of unclear or conflicting evidence, there have been some consistent findings.

First, several studies have demonstrated that people who stutter demonstrate greater activity in the right hemisphere cortex (gray matter) than nonstuttering individuals. Such overactivation is observed particularly in an area of the right frontal lobe which is homologous to Broca's area of the left hemisphere. Structural studies have further demonstrated that white matter differences between stuttering and fluent speakers, located primarily in the right primary auditory and frontal cortices, were observed contralateral to the gray matter differences. The implication from these studies is that structural deficits are compensated for by the contralateral region through the formation of new white matter pathways. In a recent study of stuttering children, reduced gray matter volumes in the bilateral-inferior frontal gyri and left putamen were observed. These stuttering children also demonstrated increased gray matter volumes in right ventral premotor and superior-temporal gyri relative to fluent children. These findings are interpreted as reflecting developmental factors resulting in a deficit in speech motor learning. A number of studies have also reported that stuttering subjects show significantly increased activation in the cerebellum, which is critical for integrating sensory and motor functions early in the acquisition of motor skills.

A second etiological level of stuttering involves the precipitating causes —that is, when, where, and under what conditions stuttering is most likely to occur. At this level, concern is with the loci of stuttering and the specific speaking situations affecting fluency. Linguistic factors associated with stuttering have led some theorists to view it as a higher-level cognitive dysfunction. The vast majority of stuttering occurs mostly on the initial

sounds of words or syllables. Fluency-facilitating conditions include choral reading, repeated reading of the same material, whispering, and speaking under conditions of altered auditory feedback or masking noise. Based on these and other phenomenological features of stuttering, new theories have implicated higher-level linguistic processes as the fundamental causal elements for speech disruption in stuttering. These processes involve the planning or adjustments of linguistic plans via internal monitoring loops. One such theory, the covert repair hypothesis, emphasizes the role of internal monitoring in detecting and correcting phonological errors in speech planning and production. The theory uses elaborate models of internal monitoring, self-repair, and priming in attempts to explain both the type of dysfluency and the conditions that facilitate speaking.

A third level of causation and an area of active investigation concerns the motor speech and vocal tract dynamics characteristic of stuttering. The vocal tract events generating disrupted speech patterns have been the focal point of much of the physiological research in stuttering since the early 1980s. It is reasoned that increased knowledge of vocal tract events and the component behaviors of stuttering will ultimately have a direct impact on the assessment, measurement, and modification of these behaviors. Much information is available on the respiratory, laryngeal, and articulatory dynamics of stuttering. Physiological studies indicating differences between stutterers and nonstutterers are closely scrutinized because they may in fact be dealing only with peripheral events surrounding stuttering, not with underlying causation. The above argument notwithstanding, individuals who stutter, as a group, exhibit relatively longer laryngeal, respiratory, and articulatory reaction times. It has been speculated that the speech motor systems of people who stutter are limited in their skill or ability to prepare and perform the motor actions that are required to implement the various demands imposed by linguistic, cognitive, emotional, and motor aspects of speech.

The final level of causation involves the environmental factors associated with the onset and development of stuttering. Much information has been collected on personality traits of young children who stutter and their parents, as well as the psychodynamic, sociological, and conditioning factors influencing stuttering and their interaction with linguistic development. Considerable inconsistency exists in data obtained from research on the personality of stutterers. There is little (if any) conclusive evidence that stutterers, as a group, exhibit a definable character structure or predictable set of personality traits. This inconsistency is further evidenced in a series of recent articles investigating temperament, emotional regulation, and emotional reactivity in young children who stutter. Some of these studies suggest that greater emotional reactivity experienced by preschool children who stutter, together with their relative inability to regulate their emotions and control their attention, may lead to

fluency breakdown. Other studies have been unable to distinguish between the children who stutter and those who do not, on these variables. Moreover, there appears to be much overlap in overall adjustment profiles between stutterers and nonstutterers, with lower social adjustment and self-esteem scores characterizing individuals who stutter. These traits are most often explained by the impact of the stuttering itself on these individuals and their environment.

DIAGNOSIS AND CLINICAL FEATURES

Stuttering is typically characterized by involuntary disruptions in the fluency of verbal expression. These disruptions may be audible or silent and are usually characterized by repetitions or prolongations of sounds or syllables, or by tense pauses or the complete cessation or blocking of speech sounds.

Accessory, or secondary, features frequently accompany instances of stuttering. These secondary behaviors involve both the speech mechanism and related or unrelated body structures. Such behaviors may include eye blinks, facial grimacing, lip and tongue tremors, tics, vocal fry, head jerks, abnormal movements of the limbs, and breathing irregularities. Linguistic and situational avoidance behavior is often observable with confirmed stuttering. The affected speaker will often scan the intended utterance to predict stuttering and then attempt to avoid such disruption by substituting words, revising phrases, circumlocutions, or stopping the entire communicative process. Situational avoidance is often reported, such as speaking in front of a group or on the telephone. Most adult stutterers report specific sounds, words, or situations that increase stuttering. Most of this information comes from studies carried out on chronic adult stutterers. Children who stutter have not been studied extensively, but it is generally accepted by clinicians that stuttering in its earlier forms is more often restricted to simpler types of dysfluency, such as repetitions and prolongations, and less often involves the more severe struggle and avoidance behaviors. Definitions of *stuttering* vary widely and are usually based on listeners' observations of the disrupted speech pattern. However, some authors emphasize that the affected individual's sensations of the stuttering moment are crucial to any complete definition. Many forms of stuttering are so subtle that they are undetectable by the listener and are only perceived by the speaker.

Although stuttering varies markedly in terms of frequency and severity, both within and among individuals who stutter, some fairly predictable events or situations are commonly observed that either facilitate fluency or result in a deterioration of the stuttering (e.g., choral reading, singing, whispering, speaking in the presence of masking noise or under conditions of delayed auditory feedback, speaking when alone). Conditions commonly observed to increase stuttering include talking on the telephone, talking to

groups, making introductions and saying one's own name, talking to authority figures, and attempting to communicate during stressful situations such as job interviews or formal presentations.

Although severe forms of stuttering are most often observed in adults, children and adolescents can be similarly affected. Stuttering can have a significant impact on the individual's quality of life, causing significant social, emotional, educational, and vocational problems. The preschool and young school-age stutterer frequently develops social and school-related anxiety. The adolescent who stutters tends to withdraw from normal social interaction. The young adult stutterer often reports feelings of restricted academic and career choices and often forgoes the opportunity to develop specific talents and career paths.

Adam was a pleasant and likable 8-year-old boy in the second grade. He had a younger brother and older sister whose speech developed normally. Adam's speech was also developing normally until around the age of 3½ at which time he began to show numerous repetitions of words, syllables, and sounds. Adam's pediatrician felt that the dysfluent speech was only a phase that Adam would soon outgrow. Adam's speech grew progressively worse and he began to show signs of struggle as he tried to force his words out on many occasions. The family remained convinced that the physician's predictions would in fact be borne out and that Adam would eventually grow out of the problem. By the time Adam was in the second grade, his speech was characterized by severe stuttering taking the form of silent blocks lasting up to 20 seconds. These blocks were associated with generalized upper body tension and breathing abnormality just prior to the instance of disruption. In addition to being visibly frustrated by his inability to communicate at school and with his friends, Adam became withdrawn and socially isolated. He would not participate in class and would often refuse to go to school if he knew that he would be called on during a certain activity. Adam's mother also reported that there had been several instances of bullying and teasing by two of Adam's classmates. On several occasions, Adam returned home from school crying because of the remarks made concerning his speech. At this point, Adam's mother consented to having the school speech–language pathologist perform a full assessment and begin therapy sessions in order to provide Adam with the relief and support that he had been lacking.

The case study of Adam illustrates two commonly observed scenarios with regard to children's early speech development. First, childhood dysfluent speech patterns are often mistakenly interpreted as normal developmental phenomena. Although it is true that many young children will pass through periods of disrupted speech patterns and will spontaneously recover from these, a significant minority, estimated as high as 30 percent, of these children will continue to exhibit disrupted speech patterns, and if left unattended, they will develop confirmed stuttering. Recent studies point to the importance of early clinical intervention and, at the very least, referral for assessment by a speech–language pathologist to determine whether treatment is indicated. The case study highlights a second frequently observed social impact on children who stutter, teasing and bullying. Recent studies have demonstrated that more than 80 percent of children who stutter have at some point experienced bullying, usually in the form of teasing or name-calling. Children who stutter will often

experience feelings of shame and reduced self-esteem as they are identified by their peers as different and vulnerable. They will often be excluded from their peer group activities and become increasingly withdrawn and socially isolated. Because they feel inadequate as communicators, these feelings of inadequacy quite often evolve into generalized feelings of low self-worth.

Susan is a 28-year-old bilingual speaker of English and Bengali. At the time of her assessment, her speech was characterized by severe amounts of hard glottal attacks, sound prolongations, and laryngeal blocks. Secondary behaviors observed included eye blinks and lip tremor. She reported that she regularly used avoidance strategies such as starters, fillers, and word switching to avoid stuttering in public, whenever possible. Susan also reported that she avoided speaking with unfamiliar listeners and speaking in group situations. Susan reported that she had participated in treatment at the age of 8 and 10 with some success in structured settings, but her gains had been difficult to maintain.

Based on her assessment, it was recommended that she participate in an adult group intensive treatment program focusing on fluency-shaping strategies. Susan was judged to be an excellent candidate for this model of treatment based on the following criteria: First, she demonstrated through her assessment interview and reading sample that she possessed the intellectual capacity to manage the rigorous demands of the intensive nature of the program. Second, she demonstrated through her oral mechanical examination and trial therapy probe that she was capable of managing the physical demands of altering her motor movement patterns. Third, she evidenced the emotional readiness to benefit from group therapy and to address the cognitive restructuring requirements of the program. Fourth, she evidenced a high level of motivation to successfully complete treatment and participate in a long-term maintenance program.

Susan's treatment program included 3 weeks of daily clinic sessions lasting for close to 5 hours per day. She further dedicated between 2 and 4 hours to the completion of home assignments each evening.

In addition to successfully acquiring the motor skills required for fluent speech, Susan responded well to the cognitive restructuring goals of the program. She took to heart the suggestion that optimal success in treatment cannot be achieved without a genuine acceptance of one's self as a person who stutters. She gained a level of comfort and openness about stuttering by completing a number of assignments that required her to practice with a variety of partners outside the group. This afforded her the opportunity to discuss her speech condition and therapy program with a number of family members, friends, and coworkers.

Susan is currently in the maintenance phase of her treatment program. She has taken many positive steps to creating a lifestyle that supports her goal of long-term fluency maintenance. Susan carefully completes her daily schedule of maintenance practice. When she neglects her practice, she realizes the negative effect on her speech but manages these ups and downs with a positive attitude. She thoughtfully selects short-term goals at each follow-up meeting that will support her long-term goal of successful maintenance. Most recently, she revealed that she has been accepted to a graduate program in speech–language pathology.

DIFFERENTIAL DIAGNOSIS

The ability to produce fluent speech is a skill that develops as the young child grows, and it depends on such factors as motor coordination and timing and linguistic and cognitive development, as well as emotional development and maturity. Fluent speech production results when the motor skills to produce a desired message occur sequentially and synchronously in a smooth manner. This requires a fully developed system composed of respiratory, phonatory, resonatory, and articulatory components acting interdependently to produce fluent speech. Given that

these systems are not fully developed at 3 or 4 years of age, preschool children quite commonly exhibit dysfluency in speech during their communication efforts. The factors that distinguish normal nonfluency from incipient stuttering involve both the type and frequency of the observed disrupted speech patterns. Children who stutter exhibit predominantly part-word repetitions, voiced or unvoiced sound prolongations, and broken or fragmented words, whereas the normally nonfluent child is more likely to demonstrate whole word and phrase repetitions and revisions. Several researchers have emphasized the degree of speech fragmentation as the distinguishing factor between stuttering and normal nonfluency. The child who is developing normal speech patterns usually does not evidence more than 10 percent of any type of dysfluency in speech, whereas the stuttering child often exhibits higher amounts of speech disruption. Another distinguishing factor is the presence or absence of struggle during speech attempts. The child who is stuttering often exhibits difficulty in initiating or maintaining breath flow or voicing during speech attempts. The normally developing child is more likely to evidence easy, effortless speech, even if it is intermittently marked by repetitions of whole words or phrases. Unlike the stuttering child, the normally developing child produces speech, voice, and airflow in a continuous, uninterrupted fashion.

The discussion thus far has centered on developmental stuttering, the disorder of fluency most commonly observed in children and adults. It should be noted that there are other types of speech dysfluencies that may be observed in clinic populations, two of which will be mentioned here.

Acquired neurogenic stuttering is a condition that can appear for the first time in adulthood and is usually associated with a variety of temporary, progressive, and nonprogressive neurological disorders, such as traumatic brain injuries, strokes, and extrapyramidal diseases. In addition to adult cases of neurogenic stuttering, there have also been reports in which childhood stuttering worsened or reappeared with the onset of the neurological disorder. Several features of neurogenic stuttering help to distinguish it from developmental stuttering: general annoyance with the problem but no apparent anxiety, instances of dysfluency occurring in any part of the utterance, lack of secondary symptoms, stuttering on small grammatical words and substantive words with approximately equal frequency, and reduced variability of stuttering across different communicative situations.

Cluttering is another disorder of fluency characterized by rapid speech, reduced intelligibility, and frequent telescoping of words. Cluttering is often marked by grammatical and syntactic errors and is not typically associated with any specific sound or word fears or even limited awareness of the speech disorder. Cluttering has often been associated with a central language imbalance, and some authors have suggested that cluttering exists

when an individual presents with one or more impairments in each of five broad communicative dimensions reflecting cognitive, linguistic, pragmatic, speech, and motor abilities. One of the features differentiating cluttering from developmental stuttering is self-monitoring and general awareness. Whereas the person with developmental stuttering is usually hypersensitive to the sound of his or her own speech, the person who clutters is often only vaguely aware of the aberrant speech pattern. Prognosis for improvement with treatment for cluttering is usually guarded, in large part because of the individual's general lack of awareness of the speech pattern.

COURSE AND PROGNOSIS

The onset of stuttering can occur anytime between the ages of 2 and 7, peaking at approximately age 4. Most of the risk for the onset of stuttering is usually over by the age of 5. The stuttering fluctuates in severity initially, so that the course of development is not steady. Initially, days or weeks may go by with little or no speech difficulty, only to be followed by a sharp increase in the amount and type of stuttering behavior. The stuttering pattern changes with time. Despite fluctuations in severity, young children who begin to demonstrate signs of true stuttering do not continue to be dysfluent in the same manner in which they began. As children begin to demonstrate more struggle and fragmentation of words and syllables, the stuttering pattern becomes more complex. Moreover, although the stuttering is usually episodic in the beginning, it becomes more chronic as it progresses. In its most advanced form, speech disruptions are accompanied by secondary behaviors such as those previously described. Complete recovery from stuttering is rare after puberty. As many as 80 percent of stuttering children experience spontaneous recovery, with girls more likely to become spontaneously fluent than boys. Investigations of prognosis in therapy for the older, confirmed stutterer have focused primarily on the influence of personality factors in treatment; subjects with less favorable speech attitudes were less likely to obtain long-term benefit from behaviorally oriented speech therapy for stuttering.

TREATMENT

The vast majority of treatment techniques for stuttering have been conducted on adults. Speech–language pathologists have traditionally practiced two distinct forms of therapy. The first deals directly with the psychological aspects of the problem—that is, the attitudes, feelings, and emotions displayed by the individual who stutters. Treatment procedures use various types of psychotherapy directed at self-acceptance, attitude change, and avoidance or anxiety reduction. The second major speech therapy approach involves modification of the stuttering response to fluent-sounding speech by the application of systematic steps and rules of

speech mechanics. In this framework, therapy procedures attempt to reconstruct the respiratory, phonatory, and articulatory gestures used in speech production, either by stuttering modification or fluency-shaping techniques. Therapy formats for adults who stutter have been intensified, and several major stuttering treatment programs now offer full-time treatment approaches administered in blocks of several weeks. Current research and clinical efforts are directed at maintenance of fluent speech in the posttreatment environment. A significant number of individuals have been reported to exhibit either partial regression or even total relapse of stuttering without continued practice of speech skills after treatment.

Varying definitions of stuttering and of successful treatment have largely hindered measures of treatment efficacy. Current therapies report success rates ranging from 60 to more than 90 percent when the sole dependent variable is perceived fluent-sounding speech immediately after treatment. However, thorough treatment outcome measures for stuttering should include three elements: (1) a significant positive change in speech output is produced, (2) the change is generalized across speaking situations, and (3) the change is maintained over time.

Far fewer studies have been done on the effects of speech therapy for young stuttering children. Clinicians tend to be somewhat divided on the issue of direct therapeutic intervention for the preschooler who stutters. Many advocate an indirect approach, incorporating parent counseling and environmental manipulation to treat the dysfluent child. This approach incorporates such techniques as parental modeling of relaxed, unhurried speech and reduction of communicative stressors in the home environment. A second approach trains parents of very young children to apply behavioral techniques of verbal contingencies to stuttering and to verbally reinforce stutter-free speech in the child's everyday speaking environment. If it is determined that these indirect approaches are not sufficient to fully address the stuttering, direct techniques of speech modification or fluency shaping, such as respiratory control and gentle phonatory onsets, are taught in a manner similar to those used with older populations.

References

Adams MR. A clinical strategy for differentiating the normally nonfluent child and the incipient stutterer. *J Fluency Disord*. 1977;2:141–148.

Ambrose NG, Yairi E. Normative disfluency data for early childhood stuttering. *J Speech Lang Hear Res*. 1999:42(4):895–909.

Ambrose NG, Yairi E, Loucks, TM, et al. Relation of motor, linguistic and temperament factors in epidemiologic subtypes of persistent and recovered stuttering: Initial findings. *J Fluency Disord*. 2015;45:12–26.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

Baxter S, Johnson M, Blank L, et al. The state of the art in non-pharmacological interventions for developmental stuttering. Part 1: a systematic review of effectiveness. [Review] *Int J Lang Commun Disord*. 2015;50(5):676–718.

Beal DS, Gracco VL, Brettschneider J, Kroll RM, De Nil LF. A voxel-based morphometry (VBM) analysis of regional grey and white matter volume abnormalities within the speech production network of children who stutter. *Cortex*. 2013;49:2151–2161.

Beal DS, Gracco VL, Lafaille SJ, De Nil LF. Voxel-based morphometry of auditory and speech-related cortex in stutterers. *Neuroreport*. 2007;18:1257–1260.

Bloodstein O, Ratner NB. *A Handbook on Stuttering*. 6th ed. Clifton Park, NY: Thomas Delmar Learning; 2008.

Chang SE, Zhu DC. Neural network connectivity differences in children who stutter. *Brain*. 2013;136(Pt 12):3709–3726.

Chang SE, Zhu DC, Choo AL, Angstadt M. White matter neuroanatomical differences in young children who stutter. *Brain*. 2015;138(Pt 3):694–711.

Cykowski MD, Kochunov PV, Ingham RJ, et al. Perisylvian sulcal morphology and cerebral asymmetry patterns in adults who stutter. *Cereb Cortex*. 2008;18(3):571–583.

Daly DA, Burnett ML. Cluttering: Traditional views and new perspectives. In: Curlee RF, ed. *Stuttering and Related Disorders of Fluency*. 2nd ed. New York: Thieme; 1999:222–254.

De Nil LF, Beal DS, Lafaille SJ, Kroll RM, Crawley AP, Gracco VL. The effects of simulated stuttering and prolonged speech on the neural activation patterns of stuttering and nonstuttering adults. *Brain Lang*. 2008;107(2):114–123.

De Nil LF, Kroll RM, Kapur S, Houle S. A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. *J Speech Lang Hear Res*. 2000;43:1038–1053.

Devinsky O, Morrell MJ, Vogt BA. Contributions of anterior cingulate cortex to behaviour. *Brain*. 1995;118(Pt 1):279–306.

Foundas AL, Bollich AM, Corey DM, Hurley M, Heilman KM. Anomalous anatomy of speech-language areas in adults with persistent developmental stuttering. *Neurology*. 2001;57:207–215.

Fox PT, Ingham RJ, Ingham JC, et al. A PET study of the neural systems of stuttering. *Nature*. 1996;382:158–161.

*Guitar B, McCauley RJ. *Treatment of Stuttering: Established and Emerging Approaches*. Baltimore, MD: Lippincott Williams & Wilkins; 2010.

Helm-Estabrooks N. Stuttering associated with acquired neurological disorders. In: Curlee RF, ed. *Stuttering and Related Disorders of Fluency*. 2nd ed. New York: Thieme; 1999:255–268.

*Howell P, Au-Yeung J, Pilgrim L. Utterance rate and linguistic properties as determinants of lexical dysfluencies in children who stutter. *J Acoust Soc Am*. 1999;105:481–490.

Ingham RJ, Cordes AK. Self-measurement and evaluating stuttering treatment efficacy. In: Curlee RF, Siegel GM, eds. *Nature and Treatment of Stuttering: New Directions*. 2nd ed. Boston, MA: Allyn & Bacon; 1997:413–437.

Johnson W, et al. *The Onset of Stuttering: Research Findings and Implications*. Minneapolis: University of Minnesota Press; 1959.

Kang C, Drayna D. A role for inherited metabolic deficits in persistent developmental stuttering.

Mol Genet Metab. 2012;107(3):276–280.

Kolk H, Postma A. Stuttering as a covert repair phenomenon. In: Curlee RF, Siegel GM, eds. *Nature and Treatment of Stuttering: New Directions*. 2nd ed. Boston, MA: Allyn & Bacon; 1997:182–203.

*Kroll RM, De Nil LF, Kapur S, Houle S. A positron emission tomography investigation of post-treatment brain activation in stutterers. In: Hulstign W, Peters HMF, eds. *Speech Motor Production and Fluency Disorders. Proceedings of the Third International Conference on Speech Motor Production and Fluency Disorders*. Amsterdam: Elsevier; 1997:307–320.

Kroll RM, Scott-Sulsky L. The Fluency Plus Program: An integration of fluency shaping and cognitive restructuring procedures for adolescents and adults who stutter. In: Guitar B, McCauley RJ, eds. *Treatment of Stuttering: Established and Emerging Interventions*. Baltimore, MD: Lippincott Williams & Wilkins; 2010:448.

Kronfeld-Duenias V, Amir O, Ezrati-Vinacour R, et al. The frontal aslant tract underlies speech fluency in persistent developmental stuttering. *Brain Struct Funct*. 2006;221(1):365–381.

Namasivayam AK, van Lieshout P. Speech motor skill and stuttering. *J Mot Behav*. 2011;43(6):477–489.

Neef NE, Anwander A, Friederici AD. The Neurobiological grounding of Persistent Stuttering: from structure to function. [Review]. *Curr Neurol Neurosci Rep*. 2015;15(9):63.

Ntourou K, Conture EG, Walden TA. Emotional reactivity and regulation in preschool-age children who stutter. *J Fluency Disord*. 2013;38(3):260–274.

Onslow M, Packman A. The Lidcombe program of early stuttering intervention. In: Ratner NB, Healey EC, eds. *Stuttering Research and Practice: Bridging the Gap*. Mahwah, NJ: Lawrence Erlbaum; 1999:193–210.

Pardo JV, Pardo PJ, Janer KW, Raichle ME. The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proc Natl Acad Sci USA*. 1990;87:256–259.

Perkins WH. Stuttering: Why science hasn't solved it. In: Curlee RF, Siegel GM, eds. *Nature and Treatment of Stuttering: New Directions*. 2nd ed. Boston, MA: Allyn & Bacon; 1997:218–235.

Sommer M, Koch MA, Paulus W, Weiller C, Büchel C. Disconnection of speech-relevant brain areas in persistent developmental stuttering. *Lancet*. 2002;360:380–383.

Starkweather CW. Learning and its role in stuttering development. In: Curlee RF, Siegel GM, eds. *Nature and Treatment of Stuttering: New Directions*. 2nd ed. Boston, MA: Allyn & Bacon; 1997:79–98.

Van Lieshout PHHM, Hulstijn W, Peters HFM. Searching for the weak link in the speech production chain of people who stutter: a motor skill approach. In: Maassen B, Kent R, Peters HFM, van Lieshout P, Hulstijn W, eds. *Speech Motor Control in Normal and Disordered Speech*. New York: Oxford University Press; 2004:313–356.

Webster RL. *The Precision Fluency Shaping Program: Speech Reconstruction for Stutterers, Volumes I and II*. Roanoke, VA: Communications Development Corporation; 1975.

*Yairi E, Ambrose N. Epidemiology of stuttering: 21st century advances. *J Fluency Disord*. 2013;38(2):66–87.

Yairi EH, Seery CH. *Stuttering: Foundations and Clinical Applications*. 2nd ed. Boston, MA: Pearson; 2015.

▲ 43.5 Unspecified Communication Disorder

TIM BRESSMANN, PH.D., AND JOSEPH H. BEITCHMAN, M.D.

INTRODUCTION

The section on Communication Disorders in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) focuses mainly on developmental speech and language disorders, which is consistent with previous versions. The category unspecified communication disorder is open ended and may capture other notable aspects of a patient's symptomatology as they relate to verbal communication.

In order to fill the diagnostic category unspecified communication disorder with life and illustrate its possible application, this chapter, similar to previous editions of the *Comprehensive Textbook of Psychiatry,* will focus on voice disorders. Beyond the linguistic information that is conveyed through a speaker's voice and speech, the voice quality gives important information about the speaker's emotional and physical status. The patient's voice quality and quantity can be an important clinical sign in the anamnesis of a range of psychiatric conditions, and psychiatric and voice symptoms can sometimes interface in interesting ways. Voice disorders can equally be an outlet or a cause for stress and emotional turmoil. The boundary between structural and functional (psychological) etiologies is often blurred.

The voice is part of the speech production mechanism. Speech is a complex physiological phenomenon that relies on the coordination of multiple anatomical systems. Airflow from the lungs causes the vocal folds to vibrate periodically. This vibration generates a raw sound that is amplified and modulated in the pharynx. The velopharyngeal sphincter functions like a valve that can direct the airflow orally to produce oral speech sounds. If the velum is lowered, speech sounds will be nasal. The movement of the lips, jaw, and tongue creates differentiated, meaningful speech sounds. Accordingly, the subsystems of speech can be broken down as follows: respiration (airflow from the lungs), phonation (sound generation in the larynx), resonance (shaping of the sound quality in the vocal tract), oral–nasal balance (direction of airflow into the oral or nasal direction at the level of the velopharyngeal sphincter), articulation (modulation of the sound stream into consonant and vowel sounds with the tongue, jaw, and lips), and suprasegmentalia (speech rhythm, loudness, and intonation).

DEFINITION

In the DSM-5, unspecified communication disorders are described as

"presentations in which symptoms characteristic of communication disorder that cause clinically significant distress or impairment in social, occupational, or other important areas of functioning predominate but do not meet the full criteria for communication disorder or for any of the disorders in the neurodevelopmental disorders diagnostic class." The wording of this descriptor is somewhat unfortunate because it implies that unspecified communication disorders do not reach a threshold of criteria to qualify as proper communication disorders. The DSM-5 definition goes on to state that the category is applicable "in situations in which the clinician chooses not to specify the reason that the criteria are not met for communication disorder or for a specific neurodevelopmental disorder, and includes presentations in which there is insufficient information to make a more specific diagnosis." In reality, the DSM-5 categories are not broad enough to capture the full range of communication disorders. Important and frequent acquired communication disorders such as aphasia (acquired language disorder), dysarthria (neurogenic motor speech disorder) as well as structurally related speech disorders related to head and neck cancer, to name but a few, are omitted in the DSM-5. The clinician may make a diagnosis of an unspecified communication disorder not because of insufficient information but simply because of lack of a more appropriate diagnostic category.

Examples of disorders that lend themselves to the diagnosis of unspecified communication disorder include voice disorders but also structurally related disorders such as oncological pathologies (laryngectomy or glossectomy). Because the latter occur only rarely during childhood and are rather self-evidently related to surgical resections, this section is focused on voice disorders. Voices are as varied and as individual as faces. What is considered acceptable as a normal voice quality, loudness, and pitch may vary markedly between languages or in different cultures. The clinician will therefore try to determine whether a client's voice is appropriate (rather than "normal") with regard to the speaker's physical makeup, culture, and language. A voice may be considered disordered when the speaker experiences marked functional limitations, such as a voice that does not project well or gives out. The vocal quality may attract negative attention and may be perceived as unpleasant by the speaker or the environment. Speaking may be uncomfortable or painful.

HISTORY

The beginnings of modern vocology can be dated to the late 19th century. Much of the research during this time was conducted by singing teachers. The early 20th century saw a rapid improvement in laryngeal imaging methods and a more thorough understanding of the vocal mechanism. During this time, researchers also observed that many voice disorders are the result of chronic hyperfunctional misuse, and that the voice can be a

target of psychosomatic processes. These findings roughly coincided with the development of early psychoanalysis. As a result, functional voice disorders were generally interpreted as a manifestation of a suppressed personal conflict, resulting in a conversion dysphonia. However, in recent decades, voice scientists and speech–language pathologists have begun to differentiate more carefully between true psychogenic dysphonias and the functional voice disorders that are related to excessive muscle tension. Patients with muscle tension dysphonia will experience voice problems chiefly because of poor musculoskeletal alignment. The muscle tension dysphonia can be treated on its own merit, and a psychodynamic interpretation of the disorder is no longer regarded necessary or helpful.

COMPARATIVE NOSOLOGY

Human communication disorders are as complex and as varied as psychiatric conditions. The DSM-5 continues to recognize the relevance of communication disorders to the psychiatric profession. However, it is neither the aim nor the scope of the DSM-5 to provide the level of detail and differentiation that a speech–language pathologist would bring to the diagnosis and treatment of communication disorders. There are manifold communication disorders that cannot meet the relatively narrow criteria for the types of developmental communication disorders laid out by the DSM-5. Consequently, the DSM-5 category unspecified communication disorder is a diagnostic wild card.

EPIDEMIOLOGY

There are few reliable epidemiological estimates for the prevalence of voice disorders, and the available figures vary considerably between studies. In part, this is due to uncertainty about the concept of the healthy voice: The acceptability of different vocal qualities and quantities varies considerably in different cultures, languages, and on different socioeconomic levels. Voice quality is an auditory–perceptual impression, so what is considered to be or not to be pathological is very much in the ear of the beholder. The estimated prevalence for pediatric voice disorders ranges from 6 to 23 percent. The prevalence of adult voice disorders has been estimated between 3 and 10 percent. The available epidemiological reports vary in group sizes and compositions as well as in methodology. It is also difficult to estimate how many of the speakers who have been identified as having a pathological voice quality will require treatment. The need for behavioral, surgical, or pharmaceutical intervention can probably be estimated to be closer to the lower, more conservative end of the prevalence estimates.

The occurrence of voice disorders varies across the lifespan. In children, boys tend to have more voice disorders, probably due to more aggressive vocal behavior and boisterous vocal play. By adolescence, the prevalence is roughly balanced between the sexes. In adults, voice disorders are more

frequent among men than women. However, women seek help more often and sooner than men. The highest prevalence for voice disorders in adults is found in the age group of 45 to 64 years. Overall, most voice disorders are benign in nature (vocal nodules, polyps, edemas, functional dysphonias).

ETIOLOGY

The precise etiology of a voice disorder is often difficult to determine and needs to be evaluated based on the nature of the pathology, the patterns of voice use, the emotional and physical states, and the underlying or accompanying medical conditions of the individual patient. Sudden onsets tend to denote acute processes, which are mostly reversible, and slow progressive onsets tend to denote longer-standing vocal trauma or degeneration that may require more comprehensive behavioral interventions, often in combination with phonosurgery. Voice disorders are characterized by deterioration in vocal quality and quantity. A lesion on the free margin of the vocal fold results in incomplete vocal fold closure and a breathy, rough, or hoarse voice quality with a decreased loudness and pitch range. A referral to the otolaryngologist is necessary if a persistent change in vocal quality and quantity is observed. The otolaryngological examination always comprises a visual inspection of the larynx in order to rule out potential malignant lesions, such as squamous cell carcinoma.

Voice disorders can be congenital or acquired. The relatively rare congenital disorders of the larynx include laryngomalacia (softness of the cartilages of the larynx and the epiglottis), subglottal stenosis or laryngeal webbing between the vocal folds, and rare laryngeal clefts. The primary concern in infants with these pathologies is to keep the airway patent and ensure the patient can breathe.

The majority of voice disorders are acquired. Acquired voice disorders present as a hen-or-egg problem: It is difficult to determine whether abusive vocal behavior has led to a vocal pathology, or whether a pre-existing pathology has triggered a vicious cycle of maladaptive behavior. Consequently, voice therapy and vocal re-education are paramount in all vocal pathologies.

In both children and adolescents, acquired voice disorders can be caused by a number of factors. The most important factor that causes injury to the vocal folds is vocal misuse or abuse (vocal hyperfunction) through excessive yelling, singing, talking, or abusive behaviors such as excessive throat clearing or aggressive vocal play. Voice disorders related to vocal misuse and abuse are often labeled as hyperfunctional voice disorders. Acute and traumatic vocal abuse and misuse can lead to lesions in the vocal fold, such as acute inflammatory (usually bilateral) vocal nodules, acute vocal fold polyps, or acute vocal hemorrhage. If these conditions are the consequence of a singular event (e.g., a spectator sports

event), they are usually easily reversible by vocal rest and vocal hygiene. Protracted vocal abuse requires more comprehensive behavioral and possibly surgical interventions.

Long-standing voice misuse or abuse may result in chronic, fibrotic vocal nodules or chronic, indurated polyps. The delicate and supple superficial layer of the vocal fold (Reinke's space) is particularly prone to irritation and swelling, resulting in chronic inflammatory edema of the vocal folds (Reinke's edema, chronic laryngitis, polypoid degeneration). Granulomas or contact ulcers on the arytenoid cartilages may be a consequence of low-pitched, strained phonation. Many of these pathologies will be accompanied by excessive muscle tension, which is often visible and palpable in the extrinsic strap muscles of the larynx. This tension leads to an elevated position of the larynx during phonation. Muscle tension dysphonia can be a symptom that accompanies an organic pathology, or it can be a voice disorder by itself.

Excessive muscle tension is also often found in psychogenic dysphonias. Patients with psychogenic voice disorders can present with severe dysphonia or even complete aphonia (loss of voice) although the larynx is free of any visible pathology. The symptoms in the psychogenic dysphonias can be variable: Patients may claim a total loss of voice and present with only whispered or breathy phonation, or they may present with an excessively rough and hoarse voice quality, often accompanied by pain and discomfort when speaking. In some psychogenic dysphonias, the ventricular folds (false folds) are adducted together with the true vocal folds. The ventricular folds are situated above the true vocal folds. This mixed glottal–ventricular phonation can be visualized endoscopically and results in a rough, low-pitched voice quality (comparable to the singing technique used by Louis Armstrong). The onset of the psychogenic dysphonias is mostly sudden, sometimes related to a bout of cold or laryngitis in combination with a high level of life stress or emotional upheaval. According to reports, patients with psychogenic dysphonias are predominantly female, but men and children may also be affected. A specific psychogenic dysphonia of the age of adolescence is the postmutational falsetto, in which the prepuberty pitch and voice quality are retained by the patient. The result is an age-inappropriate high-pitched, shrill voice quality. It is more obvious in men than in women since the male voice drops about an octave (12 semitones) in puberty, compared with about a terce (three semitones) in females. This voice disorder is called postmutational falsetto (in men) or juvenile voice (in women). It is idiopathic but may be associated with resistance to adolescence and adulthood or a feminine self-identification (in men). It is often possible to elicit a more physiological pitch from these speakers during nonspeech maneuvers such as coughing, throat-clearing, and laughing. The discrepancy between the pitch of the speaking voice and the laryngeal pitch

during nonspeech tasks can be an indicator of a psychogenic dysphonia or of postmutational falsetto.

The complete refusal to vocalize or speak is called selective mutism. This disorder is predominately observed in younger children. These children acquire normal speech and language but at some point refuse to use it because of emotional or psychosocial distress. Most children with selective mutism limit their interaction to a small number of immediate family members with whom they have close relationships. The incidence of prolonged mutism (i.e., longer than 6 months) has been estimated at less than 1 child in 1,000. Girls are more prone to this disorder than boys. The DSM-5 classifies selective mutism as an anxiety disorder.

A disorder that is difficult to categorize in children as well as in adults is idiopathic paradoxical vocal fold movement. Patients with this disorder are frequently misdiagnosed with either asthma or with a panic disorder. The patients experience spontaneous respiratory distress, often related to strenuous physical activity, in a fashion that is similar to an asthma attack. However, no swelling of the bronchi and the trachea occurs in an episode of idiopathic vocal fold movement. The patient experiences isolated involuntary laryngospasms with a complete closure of the vocal folds. Consequently, asthma medication will not provide the necessary relief. This disorder will instead respond to behavioral therapy in most cases.

Viral infection with the human papilloma virus can affect the respiratory tract in children and adults and lead to multiple lesions. In order to prevent the spreading of the papillomatosis into the trachea and lungs, the lesions are surgically removed. Although this treatment is effective for the viral infection, it may result in severe scarring of the larynx and the vocal folds. An incomplete or irregular vocal fold closure with a hoarse and breathy voice may result.

Neurogenic voice disorders are related to lesions or degenerative processes in the central nervous system. Peripheral vocal fold paralyses can be caused by strokes, degenerative disorders, mechanical or surgical trauma, or viral infections. The peripheral paralysis can be uni- or bilateral, and the location of the lesion along the nerve will determine the position of the affected fold and the consequences for breathing, swallowing, and speech. Lesions in the central nervous system have a more generalized impact on the whole speech motor system on all levels: respiration, phonation, resonance, articulation, and prosody. Depending on the localization of the lesion along the central motor pathways, a spastic, hyperkinetic, hypokinetic, ataxic, or pseudobulbar movement disorder will result. The associated motor speech disorders are labeled accordingly (as spastic, hyperkinetic, and so forth dysarthrias) and roughly reflect the nature of the underlying movement disorder. Apraxia of speech is a motor speech disorder that is characterized by deficits in the muscle coordination and sequencing of speech gestures. This also affects the coordination of the

laryngeal voicing and devoicing for voiced and unvoiced speech sounds (e.g., to differentiate /f/ and /v/).

Spasmodic dysphonia is an insidious neurogenic voice disorder that is considered to be a focal dystonia (like torticollis or blepharospasm), possibly caused by small lesions in the basal ganglia. In spasmodic dysphonia, voicing is interrupted by adductor (adductor spasmodic dysphonia) or abductor spasms (abductor spasmodic dysphonia). The voice quality in spasmodic dysphonia is unsteady and often effortful, marked by a chaotic pattern of vocal fold adduction and abduction. Visceral voice production during coughing or throat clearing is often unimpaired. However, during conscious, purposeful speech, the spasms become obvious. Spasmodic dysphonia is very rare in children and adolescents but occasional cases have been reported.

An aggravating factor for all voice disorders can be insufficient hydration of the vocal folds. The vocal folds can be dehydrated by insufficient intake of fluids, active and passive smoking, or exposure to irritating or toxic chemical fumes in the work or home environment. Dehydration makes the vocal fold less flexible and more vulnerable for mechanical trauma caused by loud singing, shouting, aggressive throat clearing, and so forth. Medications can also contribute to vocal fold dehydration. Bronchodilators can influence airflow and air passage. Diuretics, corticosteroids, and decongestants can reduce the fluid level in the vocal fold. Antihistamines, antitussives, and antirefluxiva can reduce pharyngeal and laryngeal secretions. Long-term hormonal therapies (especially with testosterone) can change the vocal fold structure and permanently lower the pitch. Acid reflux from the esophagus can lead to esophagitis (gastroesophageal reflux disease, GERD) with ulcerations or granulomas on the posterior aspect of the vocal folds or the arytenoid cartilages and a resulting voice disorder. Reflux esophagitis with irritation of the larynx may also be the result of self-induced vomiting in bulimic patients. A number of psychotropic drugs can cause xerostomia and have dehydrating effects on the vocal folds, including tetracyclic, tricyclic, antipsychotic, and anxiolytic drugs. Usually, the level of oral secretions is a good predictor of the laryngeal secretions and the lubrication of the vocal folds. If marked xerostomia is observed, the patient should be counseled about hydration by a speech–language pathologist before he or she develops voice problems.

The perceived subjective impairment resulting from a voice disorder is highly variable across patients. Voice disorders may cause anxiety and depression and may impact the patient's work and social interactions. Anxiety about the restitution of the premorbid voice quality is usually highest among professional voice users, such as singers and actors. Professional voice users are often very emotional about even minor changes in their vocal quality and range. Voice problems in professional voice users

may sometimes be related to unrealistic or unhealthy artistic demands on the voice, such as prolonged singing or shouting with a deliberately rough and loud voice. In this sense, the vocal folds of a professional voice user are like the knees of a professional basketball player: Irrespective of the talent, the body has individual physiological limits for every organ. Some rock music singers have a seemingly endless vocal stamina and never sustain vocal injury, even though they abuse their vocal folds every night on stage. Other singers hit the physiological limits of their larynx a lot sooner, despite a correct singing technique.

A population with unique voice therapy needs are transgender patients who strive to adjust their voice quality and speech to their new gender role. In female-to-male transgender patients, the male hormone therapy will lead to a thickening of the vocal folds which will lower the speaking pitch. The vocal tract size will still be female and the patient will have to work with a speech–language pathologist on developing convincingly male-speaking mannerisms. The transition is more difficult for a male-to-female transsexual patient. The patient must learn to raise the speaking pitch and develop more feminine speaking habits. However, the larger male vocal tract size makes it more difficult for many male-to-female transgender patients to convincingly portray the new female personality. For the transgender patient's satisfaction with the treatment outcomes and quality of life, the voice change can be more relevant than the physical appearance. Therefore, it is important to ensure that the patient is referred to a speech–language pathologist who is competent in voice therapy for gender-reassignment patients.

A final etiology that needs to be mentioned in this section is velocardiofacial syndrome (22q11 microdeletion syndrome). This congenital syndrome is related to a microdeletion on the long arm of chromosome 22 and has highly variable symptoms. As a result, patients are sometimes not identified at birth but only later when symptoms become manifest. Many of the patients show various developmental delays, especially with regard to speech and language development. In childhood, attention-deficit/hyperactivity disorder, anxiety disorders, phobias, or emotion regulation disorders are considerably more frequent than in the general population. In adulthood, schizophrenia and schizoaffective disorders are frequently observed. As part of the syndrome, many patients present with cleft palate, which leads to hypernasality (too much air and sound emitted through the nose during speech). The cleft palate in this population is often isolated so that it does not affect the lip and the alveolar ridge and is not visible in the patient's face. It may even be submucous (hidden under a lining of mucosa) and hidden to a casual oral inspection. Patients with velocardiofacial syndrome are a prime example of the importance of interdisciplinary treatment of complex disorders. The members of the craniofacial treatment team need to be aware of the

patients' risk for mental health problems during the lifespan. On the other hand, a psychiatrist who notices hypernasal speech in a patient who has not been followed by a craniofacial team should make the appropriate referral.

DIAGNOSIS AND CLINICAL FEATURES

Any unexplained change in voice quality, especially when it involves hoarseness or pain and persists longer than 2 weeks, should be examined by an otolaryngologist. The diagnosis of voice disorders relies on a number of different components. An important component is the perceptual evaluation of the voice disorder by an experienced otolaryngologist or speech–language pathologist. This evaluation assesses the patient's voice quality and quantity. Accepted general descriptors of voice quality are breathiness (resulting from incomplete closure or hypoadduction of the vocal folds) and roughness (irregular vocal fold vibration resulting from lesions on the vocal fold or hyperadduction of the vocal folds). *Hoarseness* is defined as a mixture of roughness and breathiness. Additional descriptors include strain or asthenia. Vocal quantity can be described in terms of the pitch and dynamic range of the voice. If pitch is inappropriate, a speaker's voice is too high or too low for his or her age, maturation, and sex. If the dynamic range is inappropriate, a patient may speak too loudly or too softly or may have difficulties raising his or her voice. Associated perceptual assessments concern the patient's oral–nasal balance, which may make speech sound hypernasal or hyponasal, or may add a muffled or twangy quality (however, it is important to note that hyper- or hyponasality are not voice disorders per se). Finally, suprasegmental features of speech (prosody, speech melody) are assessed. Pathological prosodic patterns include monotonous, unmodulated voice versus exaggerated inflections, excessively slow or fast speech, and inappropriate or uneconomical breaks for inhalation during speech. Quantitative instrumental measures of voice quality, voice range, or airflow can supplement the perceptual diagnosis. This is combined with a detailed interview that assesses the voice usage and medical history together with health practices, social and occupational circumstances, and psychological considerations. In pediatric patients, developmental and family information will be included.

PATHOLOGY AND LABORATORY EXAMINATION

Most lesions of the vocal folds are benign and do not require detailed histological analysis. However, suspicious lesions should be biopsied and analyzed in order to rule out malignancies. The malignant laryngeal tumors are usually of the squamous cell carcinoma type. Irrespective of the clinical presentation of the voice disorder, the vocal folds should always be examined visually by an experienced otolaryngologist in order to correctly assess or rule out any organic causes of the problem. This can be done with a simple mirror examination, but ideally, a flexible endoscopy or a

videostroboscopy should be performed. The videostroboscopic examination allows the physician to slow down the display of the vocal fold motion so that the patterns of vocal fold vibration can be observed. High-speed video kymography allows the physician to film vocal fold motion in real time, with a temporal resolution of several thousand exposures per second. Although a complete and comprehensive assessment of a patient's vocal structures and voice quality is crucial to determine the correct treatment, many otolaryngologists tend to be content with only superficial investigations, especially of pediatric clients. Malignancies of the pediatric larynx are extremely rare, while on the other hand a videostroboscopic examination of a small child can be very tedious. However, this wait-and-see attitude may sometimes delay appropriate treatment. Voice disorders in pediatric and adolescent populations have to be taken just as seriously as adult voice disorders.

DIFFERENTIAL DIAGNOSIS

Today, most voice centers use videostroboscopy or flexible endoscopy. With these methods, organic voice disorders can be readily detected. However, the challenge in voice diagnosis is to identify the behavioral components of the voice disorder. In particular, the differential diagnosis between hyperfunctional voice disorders and psychogenic voice disorders can often be difficult for otolaryngologists and speech–language pathologists. As a result, the necessary referral to psychiatric services may be delayed. Psychogenic voice disorders often have a characteristic onset. They come about suddenly, often related to a minor viral infection. The discrepancy between visceral vocalizations such as yawning and voice production in speech is often informative. If voice can be produced during coughing, throat clearing, humming, yawning, and sighing but not in speech, then it is likely that the patient can be helped back to a normal voice quality by a short course of voice therapy. After a normal voice has been restored, the patient should be referred for a psychiatric follow-up. On the other hand, if the patient presents with a total loss of voice but does not cooperate with the therapist's efforts to restore voice production, a different disorder to consider in pediatric patients is selective mutism.

One of the most difficult diagnoses for an otolaryngologist or a speech–language pathologist is spasmodic dysphonia, which is often diagnosed by exclusion of other applicable pathologies. In particular, psychogenic and neurogenic voice disorders must be ruled out. Since not every clinician is aware of spasmodic dysphonia, many patients undergo prolonged diagnostic and therapeutic odysseys before the correct diagnosis is finally made. The voice breaks in spasmodic dysphonia may sometimes resemble the tonic blocks of a person who stutters. While the onset of spasmodic dysphonia is in adulthood, stuttering begins during childhood. Another difficult diagnosis is that of idiopathic paradoxical vocal fold motion, which

can very much resemble an asthmatic attack. However, the airway constriction in these patients will be restricted to the level of the larynx and will not include the bronchi and the trachea. Finally, patients with velocardiofacial (22q11 microdeletion) syndrome may often present with hypernasality as well as a variety of psychiatric disorders throughout their lifespan. Therefore, it is important to ensure that any psychiatric patient with hypernasality of unclear etiology is referred to a craniofacial team for multidisciplinary assessment and genetic testing.

COURSE AND PROGNOSIS

As voice disorders may be caused by many different conditions, there is a high degree of variability in the long-term prognoses and treatment success. Acute organic processes that are associated with a onetime event of vocal abuse usually have a good prognosis. Examples for such pathologies are acute nodules or polyps. Vocal nodules in children are sometimes not removed by surgery as the vocal fold expands during puberty and the nodules tend to resolve spontaneously. Organic pathologies that are related to long-standing vocal abuse will require more comprehensive therapeutic interventions that also address the associated abusive behaviors. Examples for such pathologies are chronic nodules, chronic polyps, and Reinke's edema. If there is persistent voice abuse combined with aggravating factors such as smoking or alcohol abuse, the long-term prognosis is poorer. Lesions such as nodules may recur quickly after a surgical resection if the maladaptive behaviors cannot be modified. Voice disorders with a pronounced psychogenic component have to be addressed with a combined effort by the speech–language pathologist and the psychiatrist. Although the voice disorder can usually be managed effectively in speech therapy, the underlying emotional or psychosocial factors that have contributed to the voice disorder have to be evaluated by a psychiatrist.

TREATMENT

Intervention in voice disorders is dependent on the nature of the problem. Organic pathologies are typically managed surgically with laser surgery or microsurgery, followed by a course of voice therapy. In functional voice disorders, the speech–language pathologist will counsel the patient about the nature of the disorder and commence a course of vocal re-education, aimed at modifying or eliminating vocal abuse. If psychological factors contribute to the genesis or the persistence of the voice disorder, a referral to psychiatric services will be necessary.

There are many different approaches to voice therapy. Most of these fall within five main streams of voice therapy: symptomatic, physiological, psychogenic, hygienic, and holistic. This distinction is largely academic because the boundaries between the different approaches are blurred. Most voice therapists will move between these different schools of thought and

pursue a more personal eclectic approach to achieve optimum outcomes for their patients. *Symptomatic* voice therapy uses interventions and voice exercises that are geared toward the voice symptoms of the individual patient. *Physiological* voice therapy draws on concepts from physical therapy such as respiration or posture exercises to achieve the foundations for a physiological voice production. *Psychogenic* approaches to voice therapy draw on the cognitive behavioral approach to behavior modification in order to change the way in which a patient reacts to life stresses with a change in voice quality. This type of therapy also uses relaxation techniques such as progressive muscle relaxation. *Hygienic* voice therapy is a supplement rather than a standalone therapy approach. The basic concept here is a maximum hydration of the vocal folds. Hydration results in clearer, healthier secretions and well-lubricated, soft vocal folds that are less prone to mechanical trauma. Finally, a number of schools of voice therapy are *holistic,* in that they try to achieve goals on multiple levels through a fixed program of exercises. The *vocal function exercises* treat the vocal muscles like other muscles in the body and provide a number of "stretching" exercises to increase the vocal range and stamina. The *accent method* uses a combination of controlled respiration, chanting, and dancing to elicit relaxed phonation. *Resonant voice therapy* aims at maximizing oral resonance and achieving good voice projection through a combined program of vocalizing and increasing self-awareness.

Marvin was a boy of 4 years of age whose speech, language, and overall development had been within normal range. At the age of 4 years, he suddenly became aphonic and started using a whispered voice for all verbal communication. After a while, the family grew concerned and asked the family doctor for advice. The family doctor referred Marvin to an otolaryngologist who did not find any visible pathology on his vocal folds. A speech–language pathologist and a clinical psychologist were also consulted but neither was able to coax Marvin into using his voice again. The ensuing months were a period of anguish for the family until the boy started speaking normally again. It was only years later that Marvin felt that he was able to share what had prompted him to use the whisper voice: His older brother had told him one night that he was talking too much and that this was dangerous because there was only a limited amount of voice for his lifetime. He had advised Marvin not to waste his voice so that he would not run out of it in a few years, so Marvin had come up with the solution of whispering instead.

Steven was a 17-year-old teenager who was referred to the speech–language pathologist because of his high-pitched voice. Steven had experienced teasing and negative comments about his voice in the classroom and in other social situations. He was very motivated to change the pitch of his voice. The patient was of normal height and built for his age. An otolaryngological mirror examination had ruled out the presence of a vocal fold web (patients with a laryngeal web often have a high pitch because the vibrating edge of their vocal folds is shorter). Upon palpation, the larynx was normal in size and had a normal movement range when Steven was asked to swallow (laryngeal elevation) and to yawn (laryngeal lowering). An acoustic measurement revealed a fundamental frequency of around 230 Hz, which was twice as high as the appropriate pitch for a male speaker. When the patient was asked to throat-clear and to hum, the fundamental frequency dropped to an appropriate range. The speech–language pathologist used gentle vocalization exercises such as chanting in combination with manual circumlaryngeal massage in order to help the patient relax his larynx and lower his

fundamental frequency. Steven required only six sessions with the speech–language pathologist to change the pitch of his voice. However, he was apprehensive about using the new voice with his classmates at school. He was worried that the change was quite startling and that he would be ridiculed for seeing a speech–language pathologist. Steven decided that he would spend the Christmas holidays with his family and avoid seeing his friends during these days. When he went back to school in the new year, he told his classmates that he had been sick with flu and that the associated cough had done some damage to his voice box and changed his voice.

The patient described here was very motivated to change his voice to an age-appropriate pitch. However, some patients may initially resist the speech–language pathologist's efforts. Speaking with a different fundamental frequency changes a physical activity that is deeply engrained, and that is as difficult to change as one's posture or gait. Patients may sometimes complain that the new way of phonating feels uncomfortable, hurts their larynx, or even affects their breathing. These initial complaints are usually overcome, and postmutational falsetto should be resolved with a short course of voice therapy. The effect on the patient's appearance in social situations can be stark, so it is important to help the patient develop strategies to gently introduce the environment to the new voice.

References

Adler RK, Hirsch S, Mordaunt M. *Voice and Communication Therapy for the Transgender/Transsexual Client: A Comprehensive Clinical Guide*. 3rd ed. San Diego: Plural; 2012.

Baker J. The role of psychogenic and psychosocial factors in the development of functional voice disorders. *Int J Speech Lang Pathol*. 2008;10(4):210–230.

Boone DR, McFarlane SC, von Berg SL, Zraick RI. *The Voice and Voice Therapy*. 9th ed. Boston, MD: Pearson, Allyn & Bacon; 2014.

*Butcher P, Elias A, Cavalli L. *Understanding and Treating Psychogenic Voice Disorder: A CBT Framework*. Chichester: John Wiley; 2007.

Connor NP, Cohen SB, Theis SM, Thibeault SL, Heatley DG, Bless DM. Attitudes of children with dysphonia. *J Voice*. 2008;22(2):197–209.

D'haeseleer E, Behlau M, Bruneel L, et al. Factors involved in vocal fatigue: a pilot study. *Folia Phoniatr Logop*. 2016;68(3):112–118.

Dietrich M, Verdolini-Abbott K, Gartner-Schmidt J, Rosen CA. The frequency of perceived stress, anxiety, and depression in patients with common pathologies affecting voice. *J Voice*. 2008;22(4):472–488.

Gallena SJK, Solomon NP, Johnson AT, Vossoughi J, Tian W. The effect of exercise on respiratory resistance in athletes with and without paradoxical vocal fold motion disorder. *Am J Speech Lang Pathol*. 2015;24:470–479.

Gothelf D, Frisch A, Michaelovsky E, Weizman A, Shprintzen RJ. Velo-cardio-facial syndrome. *J Ment Health Res Intellect Disabil*. 2009;2(2):149–167.

Gundel H, Busch R, Ceballos-Baumann A, Seifert E. Psychiatric comorbidity in patients with spasmodic dysphonia: a controlled study. *J Neurol Neurosurg Psychiatry*. 2007;78(12):1398–1400.

Koufman JA, Block C. Differential diagnosis of paradoxical vocal fold movement. *Am J Speech Lang Pathol*. 2008;17(4):327–334.

Lang C, Nir Z, Gothelf A, et al. The outcome of children with selective mutism following cognitive behavioral intervention: a follow-up study. *Eur J Pediatr*. 2016;175(4):481–487.

Miller T, Deary V, Patterson J. Improving access to psychological therapies in voice disorders: a cognitive behavioural therapy model. *Curr Opin Otolaryngol Head Neck Surg*. 2014;22(3):201–205.

Roy N, Holt KI, Redmond S, Muntz H. Behavioral characteristics of children with vocal fold nodules. *J Voice*. 2007;21(2):157–168.

Roy N, Merrill RM, Gray SD, Smith EM. Voice disorders in the general population: prevalence, risk factors, and occupational impact. *Laryngoscope*. 2005;115(11):1988–1995.

*Stemple JC, Fry L. *Voice Therapy. Clinical Studies*. 3rd ed. San Diego: Plural; 2010.

*Stemple JC, Roy N, Klaben B. *Clinical Voice Pathology: Theory and Management*. 5th ed. San Diego: Plural; 2014.

Verduyckt I, Remacle M, Morsomme D. On the accuracy of adults' auditory perception of normophonic and dysphonic children's personality. *Logoped Phoniatr Vocol*. 2015;40(3):136–147.

Viana AG, Beidel DC, Rabian B. Selective mutism: a review and integration of the last 15 years. *Clin Psychol Rev*. 2009;29(1):57–67.

Whurr R, Lorch M. Review of differential diagnosis and management of spasmodic dysphonia. *Curr Opin Otolaryngol Head Neck Surg*. 2016;24:203–207.

44

Autism Spectrum Disorder and Social Communication Disorder

FRED ROBERT VOLKMAR, M.D., AMI KLIN, PH.D., ROBERT T. SCHULTZ, PH.D., AND MATTHEW W. STATE, M.D., PH.D.

OVERVIEW

Autism spectrum disorder (ASD) includes a range of conditions (the "spectrum") characterized by early onset of severe delay and deviance in the development of social and communicative and unusual response to the environment (repetitive behaviors, difficulties with change, and unusual sensitivity to the inanimate environment). In DSM-5, criteria are included for ASD (which essentially corresponds to disorders previously grouped within the pervasive developmental disorder class); DSM-5 also includes a new category of social communication condition. The DSM-5 approach is a marked change from that employed in DSM-IV.

AUTISM SPECTRUM DISORDER

Definition

This condition encompasses a number of concepts, all previously subsumed under the overarching pervasive developmental disorder category. Of these conditions the DSM-IV autistic disorder, also known as childhood autism, infantile autism, and early infantile autism, is, by far, the best known and, in important respects, the definition proposed in DSM-5 probably is closest to this older concept. In DSM-5, the previously used triad of impairments in social interaction, communication, and repetitive and restricted interests/behaviors was collapsed into two domains: social communication and restricted and repetitive behaviors. The four criteria included in the social communication group all must be met (see Table 44–1), while in the restricted and repetitive behaviors domain, two of four criteria must be met. The latter category includes, for the first time, a criterion reflecting sensory difficulties (over- or undersensitivity to the environment). The four symptoms must be present in early childhood. Because of concerns (noted subsequently) that a large number of individuals might lose service eligibility, an exception rule allows individuals with Asperger disorder and pervasive developmental disorder-not otherwise specified (PDD-NOS) and

higher-functioning autism to retain their diagnosis, thus effectively maintaining the DSM-IV system for individuals who already had a diagnosis so that, effectively, two systems will remain in use for some time.

Other changes in the DSM-5 approach involved the overarching decision to abandon use of not otherwise specified categories and the multiaxial approach to diagnosis that had been in place since DSM-III. In DSM-5 a series of specifiers are provided, for example, if a known etiological factor is present. A second specifier allows for description of required levels of support and impact on levels of functioning for the two criteria domains. These severity specifiers range from levels 1 to 3, indicating a need for support, substantial support, or very substantial support, respectively. It is made explicit that severity specifiers do not correspond directly to qualification for or need for services because they do not account for a person's profile of abilities or an individual hierarchy of intervention objectives. A third specifier can be used to indicate the presence of associated intellectual disability and a fourth to indicate whether language impairment is present. A final specifier allows for indication of the presence of catatonia (Table 44–2).

The changes introduced with DSM-5 have been the source of considerable controversy. Clinically the work group sought to preserve the strengths of DSM-IV while improving on its limitations—notably the controversies surrounding the validity of categories like Asperger disorder and PDD-NOS in DSM-IV and also to improve the applicability of the criteria to infants and very young children. In contrast to DSM-IV, where a large international field trial was undertaken in conjunction with the ICD-10 revision, the decision was made to analyze a large body of data collected from standard interviews of parents and an individual assessment specifically designed for individuals with autism. Although the original DSM-5 report suggested 10 percent of cases might "lose" their diagnosis a series of studies, using a range of different samples, suggested that a significant number of cases would lose their diagnoses—these estimates varied depending on the nature of the sample but suggested that more cognitively able individuals with autism (as diagnoses in DSM-IV) and also Asperger and PDD-NOS (as diagnosed in DSM-IV) were at risk—with perhaps 80 percent or more of such cases losing their diagnoses. Young children with autism were also noted to be at risk. A more recent study using an epidemiologically representative sample suggests that overall, perhaps 20 percent of cases might be at risk of losing their diagnoses more. The utility of the various proposed modifiers remains to be established but concern has been made over the reliability of ratings of severity since in children with autism needs can be dramatically different depending on the environmental context within which observation is made.

History

In 1943, Leo Kanner described 11 cases of what he termed autistic disturbances of affective contact. In these cases, there was a congenital "inability to relate" to people in usual ways. Kanner also noted unusual responses to the environment, which could include both stereotyped motor mannerisms and resistance to change or insistence on sameness, as well as unusual aspects of the child's communication skills, such as pronoun reversal and a tendency to echo language (echolalia). Kanner's phenomenologically based description of this condition remains worthy of review. However, at the same time, he was mistaken about some aspects of the condition, and his early misconceptions took some time to clarify. For example, Kanner believed that the condition was not associated with mental retardation because the children were not dysmorphic and looked "intelligent" and did well on some parts of intelligence (IQ) tests. As time went on, it became clear that many function in the intellectual deficient range if the entire test is administered. However, consistent with Kanner's original observation, it continues to be the case that marked scatter in skills —with occasional areas of ability ("splinter skills") or even unusual ability ("savant skills")—is often observed. In his original paper, Kanner mentioned that parents of affected children were unusually well educated or successful in their occupations. This led to the notion that autism might somehow result from pathological patterns of care. Considerable evidence suggests this is not the case and his impression resulted from ascertainment bias. There is no particular difference in terms of social class distribution of autism if factors that control for case ascertainment are controlled and if potential effects of lack of parental awareness are figured in. It is also the case that parents of autistic children do not exhibit specific deficits in parenting or other aspects of child care and are not likely to have other psychiatric disabilities at an increased rate, with the exception of anxiety/mood disorders and a range of developmental difficulties. Finally, Kanner's use of the term autism was reminiscent of Eugen Bleuler's earlier use of the term to describe the quality of self-centered thinking seen in schizophrenia. Although Kanner originally believed that autism and schizophrenia were unrelated, his use of this word became a source of confusion for many years. There was an assumption of continuity based on severity—that is, because autism was so severe, it was assumed to somehow be on a continuum with schizophrenia.

Table 44–1.
Autism Spectrum Disorder

Diagnostic Criteria
A. Persistent deficits in social communication and social interaction across multiple contexts, as manifested by the following, currently or by history (examples are illustrative, not exhaustive, see text):

1. Deficits in social–emotional reciprocity, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, or affect; to failure to initiate or respond to social interactions.
2. Deficits in nonverbal communicative behaviors used for social interaction, ranging, for example, from poorly integrated verbal and nonverbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and nonverbal communication.
3. Deficits in developing, maintaining, and understanding relationships, ranging, for example, from difficulties adjusting behavior to suit various social contexts; to difficulties in sharing imaginative play or in making friends; to absence of interest in peers.

Specify current severity:

Severity is based on social communication impairments and restricted repetitive patterns of behavior (see Table 44–2).

B. Restricted, repetitive patterns of behavior, interests, or activities, as manifested by at least two of the following, currently or by history (examples are illustrative, not exhaustive; see text):
   1. Stereotyped or repetitive motor movements, use of objects, or speech (e.g., simple motor stereotypies, lining up toys or flipping objects, echolalia, idiosyncratic phrases).
   2. Insistence on sameness, inflexible adherence to routines, or ritualized patterns or verbal–nonverbal behavior (e.g., extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take same route or eat food every day).
   3. Highly restricted, fixated interests that are abnormal in intensity or focus (e.g., strong attachment to or preoccupation with unusual objects, excessively circumscribed or perseverative interest).
   4. Hyper- or hyporeactivity to sensory input or unusual interests in sensory aspects of the environment (e.g., apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement).

Specify current severity:

Severity is based on social communication impairments and restricted, repetitive patterns of behavior (see Table 44–2).

C. Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities, or may be masked by learned strategies in later life).
D. Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning.
E. These disturbances are not better explained by intellectual disability (intellectual developmental disorder) or global developmental delay. Intellectual disability and autism spectrum disorder frequently co-occur; to make comorbid diagnoses of autism spectrum disorder and intellectual disability, social communication should be below that expected for general developmental level.

Note: Individuals with a well-established DSM-IV diagnosis of autistic disorder, Asperger disorder, or pervasive developmental disorder not otherwise specified should be given the diagnosis of autism spectrum disorder. Individuals who have marked deficits in social communication, but whose symptoms do not otherwise meet criteria for autism spectrum disorder, should be evaluated for social (pragmatic) communication disorder.

Specify if*:*

With or without accompanying intellectual impairment

With or without accompanying language impairment

Associated with a known medical or genetic condition or environmental factor (Coding note: Use additional code to identify the associated medical or genetic condition.)

Associated with another neurodevelopmental, mental, or behavioral disorder (**Coding note:** Use additional code[s] to identify the associated neurodevelopmental, mental, or behavioral disorder[s].)

With catatonia (refer to the criteria for catatonia associated with another mental disorder, pp. 119–120, for definition) (**Coding note:** Use additional code 293.89 [F06.1] catatonia associated with

autism spectrum disorder to indicate the presence of the comorbid catatonia.)

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Kanner's original report emphasized the centrality of deficits in social relatedness, as well as unusual behaviors in the definition of the condition. During the 1960s, there was much confusion about the nature of autism and its etiology but by the 1970s, a considerable body of evidence began to accumulate suggesting the neurobiological basis of the disorder. This included high rates of seizures as children were followed over time, and the observation of persistent neurological signs and symptoms. Follow-up studies of large groups of "psychotic" children also revealed a bimodal distribution of onset of this poorly defined set of cases. The early-onset group, noted by Israel Kolvin, Michael Rutter, and others, appeared to have many of the characteristics described by Kanner. Children with later-onset "psychosis" (after age 5 or 6) seemed to resemble those with schizophrenia in that they had delusions, hallucinations, and so forth. Family history also discriminated the late-onset groups with higher rates of schizophrenia in family members. A landmark in classification occurred in 1978 when Rutter synthesized Kanner's work and subsequent research in a new definition that served, in many respects, as the basis for that used in DSM-III. In DSM-III, the condition was first recognized and placed in a new class of disorders—the pervasive developmental disorders.

Table 44–2.
Severity Levels for Autism Spectrum Disorder

Severity Level	Social Communication	Restricted, Repetitive Behaviors
Level 3		
"Requiring very substantial support"	Severe deficits in verbal and nonverbal social communication skills cause severe impairments in functioning, very limited initiation of social interactions, and minimal response to social overtures from others. For example, a person with few words of intelligible speech who rarely initiates interaction and, when he or she does, makes unusual approaches to meet needs only and responds to only very direct social approaches.	Inflexibility of behavior, extreme difficulty coping with change, or other restricted/repetitive behaviors markedly interfere with functioning in all spheres. Great distress/difficulty changing focus or action.
Level 2		
"Requiring substantial support"	Marked deficits in verbal and nonverbal social communication skills; social impairments apparent even with supports in place; limited initiation of social interactions; and reduced or abnormal responses to social overtures from others. For example, a person who speaks simple sentences, whose interaction is limited to narrow special interests, and how has markedly odd nonverbal communication.	Inflexibility of behavior, difficulty coping with change, or other restricted/repetitive behaviors appear frequently enough to be obvious to the casual observer and interfere

		with functioning in a variety of contexts. Distress and/or difficulty changing focus or action.
Level 1 "Requiring support"	Without supports in place, deficits in social communication cause noticeable impairments. Difficulty initiating social interactions, and clear examples of atypical or unsuccessful response to social overtures of others. May appear to have decreased interest in social interactions. For example, a person who is able to speak in full sentences and engages in communication but whose to-and-fro conversation with others fails, and whose attempts to make friends are odd and typically unsuccessful.	Inflexibility of behavior causes significant interference with functioning in one or more contexts. Difficulty switching between activities. Problems of organization and planning hamper independence.

Comparative Nosology

In the first and second editions of DSM (DSM-I and DSM-II), autism was not officially recognized; rather, it was viewed as being on some continuum with schizophrenia. It is now recognized that childhood schizophrenia, particularly very early-onset childhood schizophrenia, is quite rare.

Although the definition of autism in DSM-III was based largely on Rutter's synthesis of Kanner's original description and subsequent work, both the name of the disorder "infantile autism" and the criteria proposed reflected a certain lack of developmental orientation (e.g., criteria for the condition were monothetic and most appropriate to younger children with autism); a category, residual infantile autism, was available for children, adolescents, and adults who had once met the criteria for the condition but no longer did so. The latter concept was problematic in several respects since older children and adults still exhibited significant difficulties and needed services.

In DSM-III-R, there was more attention to developmental concerns and a set of 16 very detailed criteria were provided, grouped in the traditional three categories of social disturbance, communicative disturbance, and restricted and repetitive behaviors. An additional change in DSM-III-R was the multiaxial placement of autism and other pervasive developmental disorders. In DSM-III-R, the condition was moved to axis II, and only the two diagnoses of autistic disorder and PDD-NOS were included. It quickly became apparent that the diagnostic concept had been unintentionally broadened.

The term pervasive developmental disorder was the topic of some debate. It was originally coined in 1980 as a new term for the class of disorders to which autism was assigned. The term was chosen to reflect the fact that multiple areas of functioning are affected in autism and related conditions. However, some investigators and clinicians objected to it since it seems to imply a greater degree of severity than is always present.

However, it is clear that, by the time of DSM-III-R, the term had come into general usage and was used as a category for the 10th revision of *International Statistical Classification of Diseases and Related Health Problems* (ICD-10).

In ICD-9, the category of psychoses with onset in childhood included the subgroup of infantile autism, among others. The use of the term psychosis in ICD-9 perpetuated the mistaken impression that autism was related to adult schizophrenia. In ICD-10, the pervasive developmental disorder category includes childhood autism, atypical autism (in which there is a failure to meet the behavioral or onset criteria), Rett syndrome, other childhood disintegrative disorder, overactive disorder with mental retardation and stereotyped movements, Asperger syndrome, other pervasive developmental disorder, and pervasive developmental disorder unspecified. In contrast to DSM-IV-TR, there was an even greater attempt to differentiate subgroups of pervasive developmental disorders. However, the validity of these conditions has been the topic of some debate, and definitive data are lacking. Despite the concerns about such distinctions, the working group responsible for DSM-IV decided, partly on the basis of a large international field trial, to try to achieve as much consistency with ICD-10 as possible. In the end, this was largely achieved, although there were some differences in terms of emphasis.

For DSM-IV, potential new criteria, as well as DSM-III and DSM-III-R and pending ICD-10 criteria, were evaluated in an international, multisite study including nearly 1,000 cases evaluated by more than 100 raters at 20 some sites around the world. In DSM-IV and DSM-IV-TR, autism was defined on the basis of behavioral features and age of onset (age of onset must be before 3 years). Behavioral difficulties must include some feature of social disturbance, communicative disturbance, and restricted interests or repetitive behaviors. The definitions of autism in both ICD-10 and DSM-IV-TR were conceptually identical. There was an explicit attempt to avoid the over inclusiveness of DSM-III-R criteria for autistic disorder. In addition, the disorder was returned to axis I. The number and detail of criteria were also reduced in length. Interrater reliability was assessed for autism and related conditions and was generally good to excellent, particularly for experienced clinicians. The DSM-IV-TR provided updated text for autism and related disorders particularly for Asperger disorder. As noted previously DSM-5 takes a markedly different approach.

Epidemiology

Prevalence. The first epidemiological study of autism was conducted by Victor Lotter in 1966, who reported a prevalence rate of 4.5 in 10,000 children among the entire 8- to 10-year-old population of Middlesex, a county northwest of London. Since then many epidemiological studies have appeared (Table 44–3). The median prevalence rate of these studies is 8.7

in 10,000. However, rates tend to be higher in studies with smaller samples and also tend to be higher more recently. If one excludes smaller and older studies, the median rate is slightly higher—approximately 9.5 in 10,000. Variability among studies reflects methodological issues, such as sample size, syndrome definition, and study design. Increased awareness of the condition on the part of health professionals and educators may also have contributed to increased case detection. Possible reasons for the increased rate in more recent studies involve the adoption of definitions of autism more applicable to the entire range of syndrome expression, as well as use of smaller target populations. Some studies have focused on the broader pervasive developmental disorder spectrum rather than on autism more narrowly defined. It is reasonable to say that approximately 1 in every 800 to 1,000 children may have strictly defined autism, with a larger number (1 in 150) exhibiting some features of the condition. The issue of whether autism is increasing in frequency—that is, if there is truly a secular change in rates of autism—has been the object of much debate. Unfortunately, the data available to address this issue are limited. There are marked variations in prevalence, even for studies conducted at the same point in time. It is clear that at least some of the apparent increases may be accounted for by other factors including diagnostic substitution; prospective surveillance data are needed to address this issue. It does appear likely that the new DSM-5 approach will significantly reduce diagnostic rates for newly diagnosed cases; cases with "well-established" previous diagnoses are automatically allowed to keep them.

Sex Ratio. Studies based on both clinical and epidemiological samples have repeatedly noted a higher incidence of autism in boys than in girls, with ratios reported averaging around 3.5 or 4.0 to 1. This ratio varies in relation to intellectual functioning. Some studies have reported male to female ratios of up to 6.0 or higher to 1 in individuals with autism without mental retardation, whereas ratios within the moderately to severely mentally retarded range have been reported to be as low as 1.5 to 1. It is still unclear why females are underrepresented in the normal IQ range. One possibility is that males have a lower threshold for brain dysfunction than females, or, conversely, that more severe brain damage would be required to cause autism in a female. According to this hypothesis, when the person with autism is a girl, she is more likely to be severely cognitively impaired.

Social Class. Although a few early studies supported Kanner's impression of an association between autism and upper socioeconomic status, most epidemiological studies published in the 1980s and 1990s have failed to reveal such association. In addition to the bias for more educated and successful parents to seek referral or the autism label for intervention purposes ("diagnostic substitution"), it seems likely that families from

disadvantaged backgrounds are still underrepresented in clinically referred samples. Outreach initiatives are still needed to provide equal access to diagnostic and intervention services to children from all socioeconomic backgrounds. Autism clearly is seen in all social classes and in all countries.

Etiology

Psychosocial Theories. Kanner's speculation that emotional factors might be involved in the pathogenesis of autism led others to conclude that the condition was always caused by the experience of a "refrigerator" mother who was not responsive to the child's emotional needs. In this view, intensive psychotherapy was needed for mother and child, or, sometimes, the child was removed from the family in an attempt to remediate the presumed deficits. Unfortunately, there was no evidence that such efforts were efficacious. A generation of parents were traumatized by the experience of being blamed for their child's condition.

Biological Theories. As children with autism were followed, various factors suggested a biological basis of the condition. These included the high rates of mental retardation and seizure disorders and the recognition that various medical and genetic conditions were sometimes associated with autism. The current consensus is that autism is a behavioral syndrome caused by one or more factors acting on the central nervous system (CNS). Although the underlying biological abnormalities of autism are unknown, efforts are now under way to delineate more precisely testable neuropathological mechanisms.

Table 44–3.
Selected Epidemiological Studies of Autism

Study, Year	Country	Size of Target Rate Population	Criteria Used Based on	Prevalence (per 10,000)
Lotter, 1966	UK	78,000	Kanner, 1943	4.5
Wing et al., 1976	UK	25,000	Kanner, 1943	4.8
Gillberg et al., 1984	Sweden	128,600	DSM-III, 1980	4.0
Cialdella and Mamelle, 1989	France	135,180	DSM-III, 1980	4.5
Gillberg et al., 1991	Sweden	8,106	DSM-III-R, 1987	9.5
Fombonne et al., 1997	France	325,347	ICD-10, 1992	5.3
Baird et al., 2002	UK	16,235	ICD-10, 1992	30.8
Chakrabarti and Fombonne, 2001	UK	15,500	ICD-10, 1992	16.8
Croen et al., 2002	USA	4,950,333	Full syndrome Autism	11.0
Fombonne et al., 2006	Canada	27,749	DSM-IV	21.9
Isaksen et al., 2012	Norway	31,015	ICD-10	14

Genetic Factors. Family and twin studies demonstrate that genetics plays a central role in the etiology of ASD. Approximately 60 percent of monozygotic twins share a diagnosis of strictly defined autism while 80 to 90 percent are concordant for the broader spectrum. When compared to the dizygotic concordance rate, the cumulative data from these types of studies suggest that genes account for a greater share of the liability for ASD than does the environment, though both contribute substantially to population risk.

For several decades, despite a strong consensus that genes play a significant role in autism, defining the specific nature of that risk proved to be quite challenging. However, over the past several years, with the availability of large patient and family cohorts, rapidly advancing genomic technologies, and a wholesale move toward large-scale collaborative studies, there has been very rapid progress in illuminating the genetic architecture of autism and related conditions.

Several important themes have emerged: first, the majority of progress so far has been a consequence of a focus on mutations that are rare in the population, are often new (or de novo) in the affected individual, and carry relatively large biology effects; second, while there is substantial evidence that polymorphisms that are common in the population convey risk, unlike disorders such as schizophrenia and bipolar disorder, so far, no specific common genetic variants have been reproducibly associated with ASD using genome-wide association strategies, the current gold-standard methodology for this type of analysis. This is almost certainly a consequence of still underpowered patient cohorts and the small individual effects conveyed by these types of genetic variations. Third, while recent studies have underscored an exceptionally high degree of genetic heterogeneity underlying ASD, so far, the genes that have been discovered appear to converge on a much smaller number of biological processes. To date, multiple studies have highlighted the role of, for example, chromatin modification, synaptic proteins, and targets of fragile X mental retardation protein (FMRP) as playing important roles in pathophysiology and underscored the importance of mid-fetal human cortical development. Fourth, there is incontrovertible evidence that rare, de novo variations in the submicroscopic structure of chromosomes, also called copy number variations (CNVs), carry risk for ASD. Moreover, it has also been repeatedly shown that the identical CNV in different individuals may increase the liability for a very broad range of neurodevelopmental disorders, ranging from schizophrenia, to epilepsy, intellectual disability, or specific language impairment. These findings raise important questions about the specificity of genetic risks underlying ASD. Finally, while the relative contribution of genes is thought to be quite large, influences other than variation in the deoxyribonucleic acid (DNA) sequence are likely to turn out to be

important as well, including epigenetic mechanisms and environmental factors. Indeed, one of the most important outcomes of the successful search for autism transcripts will be the light it will shed on nongenetic, potentially modifiable risks.

Other Medical Conditions and Autism. ASDs have also been associated with other conditions with a strong genetic component, for example fragile X syndrome and tuberous sclerosis. In fragile X, a marker X chromosome is noted to be "fragile" in that, when grown in a special medium deficient in folic acid, there is a site that commonly fractures. The fragile X mutation typically involves a triplet repeat of cytosine–guanine–guanine in the regulatory domain of the gene *FMR1*, which may amplify with succeeding generations. Physical signs of the condition include characteristic facies, enlarged testicles, associated mental retardation, and autistic features. Behavioral difficulties include attention problems, impulsivity, and anxiety. Recent data suggest that between 20 and 30 percent of boys with fragile X meet diagnostic criteria for autism spectrum conditions, and between 1 and 8 percent of boys with idiopathic autism will be found to carry a fragile X mutation.

Given the marked increase in risk for ASD among those with abnormal expression of the fragile X gene and recent findings suggesting that many genes carrying risk for "typical" ASD are targets of the FMR protein (FMRP), recent advances that have begun to illuminate the biology of this molecule may turn out to have important implications for understanding more common causes of autism. FMRP is an RNA-binding protein that has been found to interact with multiple transcripts and to repress gene translation, to contribute to neuronal plasticity through involvement in the process of long-term depression (LTD) at the synapse, and to help regulate the function of group I metabotropic glutamate receptors (mGluR). Loss of FMRP leads to increased mGluR5 signaling activity and adversely affects synaptic plasticity. Such findings have led to novel and unanticipated avenues for potential intervention: In animal models of fragile X syndrome, use of glutamate receptor antagonists (such as MPEP [2-methyl-6-(phenylethynyl) pyridine] or lithium [Eskalith]) can rescue some of the phenotypic features resulting from the loss of FMRP expression. Whether these insights will lead to improved intervention for fragile X, with or without autistic features, is not yet known.

Tuberous sclerosis (TS) is characterized by abnormal tissue growth, or benign tumors (hamartomas), which affect various organ systems. This autosomal dominant disorder is associated with a range of phenotypes, including intellectual disability and seizure disorder. Studies of individuals with autism find TS in between 0.4 and 2.8 percent of cases, a significant elevation over the rate of TS in the general population.

The mutations that lead to TS are found in the genes encoding the TSC1

(Tuberous Sclerosis Complex 1) and TSC2 (Tuberous Sclerosis Complex 2) proteins which are members of the mTOR (mammalian target of rapamycin)/Akt (a.k.a., Protein Kinase B) intracellular signaling pathway, involved in multiple cellular processes including proliferation, growth, cell survival, and mobility. Interestingly, other ASD-risk genes, such as PTEN (phosphatase and tensin homolog) are also members of this pathway. The specific manner in which mTOR/Akt pathway dysregulation leads to ASD is not fully understood. However, as is the case with FMRP, restoring the normal function of the signaling pathway in model systems carrying ASD-related mutations has led to restoration of function, even in adult animals.

PRE-, PERI-, AND NEONATAL FACTORS. Several studies have shown increased rates of pre-, peri-, and neonatal complications in children with autism. Much of the difference relates to observation that something unusual is noted about the child at birth. This may reflect the operation of genetic as well as perinatal factors. The genetic predisposition to autism may interact with perinatal factors in producing the syndrome.

OTHER ETIOLOGIES. Various reports have associated autism with a host of other conditions. To date, many of these reports are in the form of single case reports or case series rather than controlled studies and their significance is difficult to interpret. Moreover, controlling for potential confounds when investigating the prevalence of ASD in neurodevelopmental syndromes accompanied by significant intellectual impairment, can be challenging. Several associations have been of particular interest; for example, autism associated with neurofibromatosis and congenital rubella. In the former case, there have been several lines of recent compelling biological evidence supporting this observation. In the latter case, it has become clear that when children with congenital rubella, who were initially thought to have autism, were followed over time, their "autistic-like" features tended to diminish. Also, such children exhibit a range of sensory deficits and intellectual disability, both of which complicate the diagnostic process.

There has been much interest in recent years regarding whether the reported increase in the prevalence of autism might be the result of environmental or other factors. There clearly is some role for environmental contributions although precise mechanisms remain to be specified. The combination of advances in the genetics and neurobiology with well-designed prospective epidemiological studies will help clarify which specific environmental risks are most relevant and, importantly, modifiable. Of note, to date, despite multiple focused investigations, there is no evidence that vaccinations play a role in the environmental liability for ASD.

Postmortem and Neuroimaging Studies

NEUROIMAGING AND BRAIN DEVELOPMENT. ASD is a neurodevelopmental condition that emerges early in life. Much has been learned about ASD from postmortem, model system, and neuroimaging studies over the last few decades. However, of these, MRI studies have provided the largest contribution to the current understanding of attributes of brain structure and function associated with autism. This section will review the main themes of this research, focusing on neuroanatomical studies using structural MRI (sMRI) and brain functioning using functional MRI (fMRI).

Neuroimaging as a tool and as an approach to understanding brain-based disorders emerged on the research landscape in the 1990s, with the arrival of fMRI to complement sMRI, which became a useful research tool about a decade earlier. Imaging in autism research, however, did not gain traction until the late 1990s, after which it accelerated rapidly, as it did in all subfields of neuropsychiatry. The dominant research approach has been the case-control design, which considers individuals with autism as categorically distinct from other groups, including typically developing controls (TDCs). Thus, studies focus on describing how the brain is different in autism than TDCs, at what ages those differences emerge, and how specific brain features map onto to phenotypic elements of autism, genetic risk variants, and outcome measurements. Recent data, however, demonstrate that ASD is extremely heterogeneous, both in its core features, with severity ranging from the boundary of typical development to profoundly impaired, and with respect to many frequently co-occurring conditions such as intellectual disability, epilepsy, anxiety, and attention-deficit hyperactivity disorder (ADHD). This is relevant to understanding the neurobiology of autism research literature because often studies do not replicate upon follow-up, and thus they cannot provide reliable information. This is inherent both to the heterogeneity that accompanies the current conceptualization of the essence of ASD, as well as to the practice of publishing results based on inadequate sample size, for example, where case and control groups have less than 20 or 30 individuals each. Inadequate sample size is an acknowledged problem across all of neuroscience because it leads to failed replications, undermining the ability to accumulate a stable base of knowledge. Thus, the review of studies herein is confined to research where there have been several studies arriving at similar conclusions, with a bias away from describing studies based on small samples.

DEVELOPMENTAL NEUROANATOMY. The observation that individuals with autism tend to have enlarged heads dates back to the origin of the diagnosis itself. Although subsequent studies of head circumference confirmed this observation, it did not receive much attention until the advent of MRI, which provided the excellent soft tissue contrast needed for more precise quantitation of brain morphology. The first MRI study from the middle of

the 1990s, found total tissue volume of individuals in the autism to be about 6 percent larger than the control group. Subsequent studies suggested that the size increase includes both gray and white matter. Courchesne et al. were the first to report that megalocephaly was observable as young as 2 years of age (but was not present at birth). This has been replicated and extended several times. Most recently Hazlett et al. showed that while gray matter volume and cortical surface area is enlarged in ASD, cortical thickness is not, which is an important observation because it constrains the possible biological mechanisms that could account for these size differences.

The Infant Brain Imaging Study (IBIS) promises to bring greater precision to the current understanding of the earliest signs of developmental brain differences. IBIS is a multisite, national study collecting data prospectively on at-risk baby siblings of youth with autism, with the first brain MRI between 3 and 6 months of age, repeated at regular intervals up until 24 months of age. The first IBIS results showed that axonal fibers had a blunted development in ASD between 6 and 24 months of age, with results clustering in brain regions implicated by fMRI study as part of the "social brain." These data suggest that autism begins to develop neurobiologically before the end of the first year of life in those who go on to a formal diagnosis at older ages (e.g., age 3). This is consistent with emerging data from very careful behavioral testing which also shows significant deficits in the first year of life for those who go onto develop autism. Another IBIS study found white matter microstructural abnormalities at 7 months of age in the group who went on to have ASD, in fiber bundles associated with visual orienting and with the ability to disengage attention. These data help pinpoint brain circuits that appear to underlie a "sticky attention," a phenomenon in autism that causes significant behavioral management issues, and undermines learning. Both studies confirm the brain connectivity is an early feature that goes awry in autism. A third IBIS study has focused more explicitly at how the white matter's individual axonal fibers are organized in autism using diffusion tensor imaging (DTI). This study found significant fiber network inefficiencies in the brains of toddlers with ASD at 24 months of age. Thus, a developing theme is that white matter shows aberrant organizational properties in the first year of life, which accentuate with time, and subserve core features of ASD.

While most studies have not tried to localize areas of the brain that may be disproportionately enlarged, those that have often find frontal and temporal regions most enlarged. In addition, Dawson et al. discovered early overgrowth of the amygdala, a structure vital to emotional processing and memory, in a sample with an average age of 4 years. Subsequent studies have verified and extended these findings showing associations between degree of overgrowth and severity of clinical symptoms. However, findings

are confined to preschool ages, as increases are noted later in childhood and adolescence. Several MRI studies have suggested that the brain's white matter is disproportionately enlarged in autism, which is viewed as supportive of more recent conceptualizations of autism as involving disrupted connectivity. However, not all white matter fiber tracts are enlarged. For example, numerous studies have found decreased cross-sectional width of the corpus callosum. Moreover, there is evidence that later myelinating areas in particular, contribute the most to the increase in brain volume in ASD relative to TD controls (i.e., frontotemporal and frontoparietal regions). Allometric models of cerebral development suggest that doubling the number of neuronal processing units (the gray matter "minicolumn") supports a fourfold increase in neuronal connectivity, potentially generating a fourfold increase in white matter volume. However, even though white matter enlargement seems a bit disproportionate compared to gray matter enlargement in ASD, it may be that the white matter enlargement is less than would be expected under model assumptions. The implication is that the bigger brains in autism are actually less well connected than smaller brains among TDCs.

The neurodevelopmental trajectory of ASD was refined by subsequent studies that found that ASD increases in head size were observable only below age 12, after which group differences disappeared. This observation has been replicated many times, but the finding is not without opposing results. The single largest and best controlled study to date, albeit on head size (an imperfect proxy for brain size), found significant head circumference enlargement regardless of age. This study is remarkable for its very large sample size (n = 1,889 individuals with ASD from young childhood to young adulthood), and for controlling known confounds such as age, body size, and IQ measurements, as well as familial effects on head size, since both parents and an unaffected sibling were also included in the analyses (total n = 7,225). In summary, the brain growth and volume literature suggests that brain overgrowth begins in the first few years of life in ASD, and that by adolescence brain size is much more typical. This abnormal pattern of growth across infancy and childhood is presumed to have very important implications for the functional properties of the brain and to likely perturb functional organization of systems that support typical social–emotional and cognitive skills. This atypical growth pattern also provides clues as to the possible cellular–genetic bases of autism, especially mechanisms that govern neural numbers.

Consistent with this summary, a postmortem study of brains of those with ASD recently pinpointed a possible mechanism for brain volume differences in autism. This study found good evidence for early deficits in neural pruning in autism. While relatively small by MRI standards, the 20 brains analyzed (about half with ASD) represents one of the larger published postmortem studies. Samples came from individuals who died

between 2 and 20 years of age. Synaptic pruning after the first year of life is well known as a normal developmental process, which is presumed to yield improved computational efficiencies. While typical controls lost about 40 percent of their dendritic spines (that part of a neuron that receives signal transmissions) across the age span, those with ASD lost half that much. The investigators also used a mouse model of autism to study possible causes of this pruning failure. They found that mTOR, a key protein in brain function and development, was hyperactive, which impaired the ability to prune synapses. Overactivity of the mTOR pathway led to blockage of autophagy processes, reduction in pruning and resulted in mice with ASD-like social behaviors. Significantly, the investigators were able to correct the pruning deficits and ASD-like behaviors in mice through use of an immunosuppressive drug, rapamycin, which inhibits mTOR activity. While this is unlikely to become a standard treatment in humans, owing to its very serious side effects, mTOR is currently being tested in clinical trials for a special subgroup of ASD (those with tuberous sclerosis), where the risk has been judged appropriate. Finally, the investigators also found in the human postmortem tissue proteins that reflect deficient autophagy and synaptic pruning malfunctioning. Collectively, these data are consistent with growing evidence of neural connectivity abnormalities in autism, especially regional overconnectivity. Given that a great many of the known ASD genetic risk variants interact with the mTOR pathway, deficits in synaptic pruning provide plausible mechanisms contributing to the development and maintenance of autism. Lack of pruning would also lead to an abundance of synapses, which would create the opportunity for seizures, a common comorbidity of ASD.

Recently attention has focused on neuroanatomical changes later in adolescence and adulthood. One particularly relevant study matched a very large number of cases and controls (age 7 to 25 years) on age, gender, and IQ, and found significant, widely distributed reductions in cortical thickness with age in ASD relative to TDCs. These declines were concentrated in the same areas of the brain shown in other studies as exhibiting hypertrophy early in life—frontal and temporal lobe regions, suggesting a unique inverted U-shaped developmental growth trajectory. Prefrontal cortex and the anterior temporal lobe also showed significantly reduced surface area with age (an age by group interaction). These results are consistent with several other recent studies that showed accelerating cortical thinning with age, including in the temporal lobe. Collectively, these data raise the concern that there might be undetected declines in cognitive and social–emotional functioning as individuals with ASD move into adulthood.

FUNCTIONAL IMAGING. fMRI studies of ASD have painted a complex portrait of differences in brain functioning in autism. While MRI studies of

neuroanatomy focus on measurements of a trait of a person, fMRI study results are influenced both by enduring personal traits (e.g., some people are very good at recognizing faces or others are not), and also mental and emotional states at the time of the assessment. Because the MRI experience can vary between individuals (e.g., stress provoking for some, sleep provoking for others), measurement signal can be influenced by state factors. The signal analyzed from fMRI comes from fluctuations in the magnetic field with different blood oxygen levels that accompany changes in neural activity. While the measurement is in this way indirect, fMRI has advantages over other neuroimaging techniques because of its ability to localize activity very precisely to different parts of the brain, with resolution in most studies of about 3 mm^3 (voxel size). Although this is precise by many standards, it is important to recognize that an average cortical gray matter voxel contains about 630,000 neurons.

fMRI studies can be of different types. Some are task based, and have focused hypotheses about regions of the brain (e.g., the amygdala) that are hypo- or hyperactive in autism. This type of study, based on a specific task (e.g., showing emotionally engaging images versus neutral images with the participant providing a judgment that is recorded by a button press), has expectations that specific brain "modules" will be differentially engaged has dominated the literature particularly the early literature. This approach has been successfully used to test a range of theories about the primary deficits in autism. The social, communication, and behavioral problems that occur with ASD suggest that the syndrome affects a functionally diverse and widely distributed set of neural systems. At one point or another, disturbance in nearly every neural system in the brain has been proposed to be a fundamental mechanism causing autism. These theories are typically derived from beliefs about the most salient behavioral and psychological features of the disorder, sometimes with explicit consideration of developmental course. For example, those who emphasize difficulty with complex information processing as the principal characteristic of autism, postulate widespread cortical abnormalities sparing early sensory and emotional processes as the neural basis of autism. Such cortical abnormalities, when present from early in life, would impair learning and contribute to developmental features of autism. On the other hand, those who focus on the social reward and saliency deficits and their role in social difficulties, often highlight the limbic system and reward system circuitry in the pathogenesis of autism. Less sensitivity to social reinforcement and drive for social engagement would curtail social development. Theory-driven research is able to articulate specific hypotheses and to pursue programmatic research that accumulates knowledge across time. Table 44–4 summarizes the knowledge about brain functioning that has been gained with this type of approach. Replicated studies that support a role of specific brain areas whose functions are

atypical in autism are represented in this table.

Table 44–4.
Brain Regions of Potential Importance in Autism

Brain Area	Function
Amygdala	Emotional arousal, emotion perception, and emotional learning
Extended amygdala, ventral striatum, and nucleus accumbens	Social reward circuitry
Fusiform gyrus (especially lateral aspect of the right fusiform)	Face (person) recognition
Posterior, superior temporal sulcus	Interpreting nonverbal communications, action observation, and theory of mind
Orbital prefrontal cortex	Emotional learning
Medial prefrontal cortex	Social cognition (interpreting what others might be thinking and feeling)

Table 44–4reflects the literature and the strong interest in the field in localizing specific functions related to social ability to discrete locations in the brain. Perhaps the best studied process and brain region involves face perception and fusiform gyrus (see Fig. 44–1). The fusiform has a specific role in person identity from visual images. Numerous studies show that it is often under active in autism, paralleling difficulties on psychological tests of face recognition. The fusiform is functionally synchronized with the limbic system, and in particular the amygdala. The amygdala, along with the reward system, is believed to play a role in directing a person's attention to socially relevant things in the environment such as people and their social interactions. Both the amygdala and the reward pathway have been shown to be hypoactive in ASD, though the amygdala's social responsiveness is confounded by its role in signaling anxiety and fear. There has recently been considerable speculation regarding the relationship between ASD symptoms and a constellation of brain areas referred to as the mirror neuron system (MNS). MNS has proven responsive to both the perception and implementation of specific body movements. The posterior part of the superior temporal sulcus has been highlighted in many fMRI studies of social perception (especially recognizing nonverbal cues, such as facial expressions, gestures, and eye gaze direction), as well as part of larger systems for representing the intentions of others (i.e., for generating a theory of mind of others). The hub of theory of mind activity, however, is generally localized to the medial prefrontal cortex. Orbital prefrontal cortex represents the value of experiences and of rewards and as such it plays a key role in emotional learning.

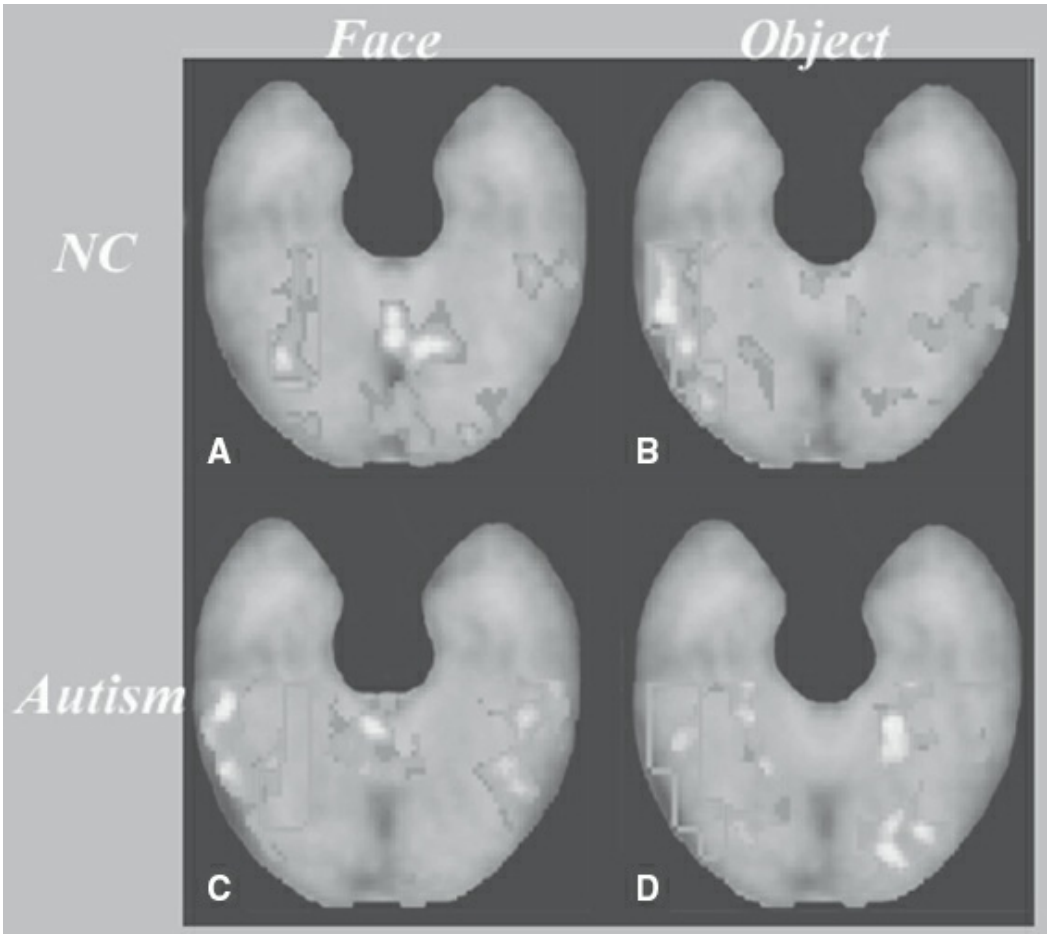

FIGURE 44–1. Composite functional magnetic resonance image comparing a group of 14 normal controls (NCs) and 14 individuals with autism or Asperger syndrome (autism), matched on age and intelligence quotient. Composite images show areas preferentially activated by faces (**A,C**) and nonface objects (**B,D**). Note that the fusiform "face area" has been defined by a green box in **A**. Among individuals with autism or Asperger syndrome, there is significant underactivation of the fusiform gyrus during face perception and compensatory activation of a region lateral to the fusiform. This more lateral region was most strongly activated by the object task among the NCs. Hypoactivation of the fusiform face area is the best replicated functional neuroimaging finding in the literature. (See color atlas.) (Adapted from Schultz RT, Gauthier I, Klin A, et al. Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. *Arch Gen Psychiatry*. 2000;57:331–340.)

In the past 5 years, there has been increased attention given to task-free fMRI studies, most notably what has been termed resting state fMRI (rsfMRI). With this approach, participants are not required to engage in any specific mental activity. Rather fMRI signal is gathered continuously during a rest period, generally of about 5 minutes long and often with more than one rest session. This approach has been most valuable in the context of functional connectivity studies of brain systems, an approach which examines the cross correlation of disparate regional brain activity fluctuations. In essence, the approach searches for synchronous fMRI activity in order to create networks of regions that appear to be working together in a time-locked fashion. This approach has yielded numerous

study findings, but many have not been consistently found across studies, and these therefore await larger studies. Those that are consistent often identify brain system differences associated with ASD that can be supported by lesion studies, or task-based fMRI (e.g., as described in Table 44–4). rsfMRI is an important component of a major NIH initiative—the Human Connectome Project. This multicenter program of research is collecting imaging data on 1,200 participants. The Human Connectome and related endeavors will allow connectivity matrices derived from graph theory on resting state and diffusion data from individuals with ASD to be compared to several other groups, including, for example, typical development, Alzheimer's, and ADHD. These publicly available resources will generate the large samples needed to make gains in characterizing neural circuits with multimodal imaging in basic cognitive neuroscience and in ASD. This is significant given other lines of evidence support the general theory of underconnectivity of the brains of individuals with ASD. This approach, along with increased implementation of multimodal imaging studies (where sMRI, fMRI, and other techniques are combined to address hypotheses about the neural architecture of ASD), promise to allow important advances in the understanding of ASD in the coming years.

Neurochemistry. Beginning in 1961, a number of studies have reported that approximately one-third of children with autism have increased peripheral levels of the neurotransmitter serotonin. This has been the focus of much research, but the significance of this finding remains unclear, as the finding is not specific to autism, and the relation of peripheral levels to central levels of serotonin is unclear.

Another line of work has focused on other neurotransmitters such as dopamine. A hyperdopaminergic functioning of the brain might explain the overactivity and stereotyped movements seen in autism. This would be consistent with the general observation that administration of stimulants, which increase levels of dopamine, sometimes worsens behavioral functioning in autism. Studies of cerebrospinal fluid (CSF) of dopamine metabolites and catecholamine metabolites have been inconsistent. It is clear that agents that block dopamine receptors are effective in reducing the stereotyped and hyperactive behaviors of many autistic children.

The endogenous opioids have been investigated, given the possibility that these compounds—enkephalins and endorphins—might lead to social withdrawal and unusual sensitivities to the environment. This was the rationale for use of opioid antagonists such as naltrexone (Depade) in treating children with autism. Although it does appear that there may be a modest effect of these agents on the high levels of activity and agitation overall, results have been disappointing. More recently, work has centered on potential use of agents like oxytocin that may impact attachment and social brain systems more directly.

Immune Theories of Autism. Some work has suggested a possible role of immunological factors in autism. There has been a suggestion that maternal antibodies directed against the fetus may be produced in utero. There also have been reports of autism associated with viral infections. Considerable controversies have arisen over the question of whether exposure to the measles–mumps–rubella (MMR) immunization might be a causative factor. This assumption rests largely on case reports that link the apparent onset of autism with the immunization; however, it has long been known that a small number of children with autism present with a developmental regression, and the bulk of the available evidence does not support a causative association with immunization. The potential negative effect of an increase in children not immune to measles is considerable.

Diagnosis and Clinical Features

In DSM-5 a diagnosis of autism required three essential social communicative features and at least two restricted/repetitive features. The social communicative deficits are in the areas of social interaction across settings, deficits in social–emotional reciprocity, and deficits in nonverbal communicative behaviors used in social interaction, for example, difficulties with eye contact and so forth. All three features must be present. In addition at least two of four restricted and repetitive behaviors must be present. These four features include: (1) stereotyped motor mannerisms; (2) restricted interests and insistence on sameness including difficulties with change and overreliance on routine; (3) highly restricted interests abnormal in focus or intensity, for example, overattachment to an unusual object; and (4) hypo- or hypersensitivity to sensory input. The symptoms must be present in early development (although their severity may change over time) and symptoms must cause clinically significant impairment. As noted previously, a variety of additional codes can be used to indicate the presence of specific features such as intellectual delay and catatonia. The symptoms cannot be better explained by intellectual difficulty. As noted earlier, individuals with previously "well-established" DSM-IV diagnosis of autistic disorder, Asperger disorder, or PDD-NOS should be given the diagnosis of ASD. Finally, for persons with major problems in social communication but who do not meet criteria for autism the new category of social communication disorder can be considered (see below).

DSM-5 provides a number of specific developmental/behavioral examples but does note that the clinician is not limited to these. Clearly there is a marked range of syndrome expression in ASD over both age and intellectual and communicative level. For the lowest cognitive level, often the child is typically largely or entirely mute, is isolated from social interaction, and makes few social overtures. At the next level, the child may accept social interaction passively but does not seek it out. At this level, some spontaneous language may be observed. Among the higher-

functioning and somewhat older children, the child's social style is different, in that he or she may be interested in social interaction but cannot sustain it in typical ways. The social style of such individuals has been termed "active but odd," in that they often have difficulty regulating social interaction once it has commenced. This last group may pose the greatest challenge diagnostically in use of DSM-5. Given the range of syndrome expression and variations in presentation with age, there is considerable potential for misdiagnosis. Evaluation of the child with autism should include a detailed history, comprehensive medical and developmental examination, and psychological testing. Practice guidelines for screening, assessment, and treatment are available.

Age of Onset. The onset of autism is by definition early in life. As shown in Figure 44–2, parents typically become concerned between the ages of 12 and 18 months as language fails to develop. Although there may be concern that the child is deaf, the parents also note the child may respond quite dramatically to sounds in the inanimate environment; occasionally, parents report, in retrospect, that the child was "too good," made few demands, and had little interest in social interaction. This is in stark contrast to normally developing infants for whom the human face and social interaction are among the most interesting and salient features in the world. Occasionally, parents report that the child seemed to develop some language, and, then, his or her language either plateaued or was lost. Such a history is reported in perhaps 20 to 25 percent of cases. Often, detailed review of early developmental history or of videotapes of the child at a younger age will reveal some earlier abnormality, and true regressive autism (i.e., with a dramatic onset of symptoms after previously normal development) is clearly much less common. The DSM-IV diagnosis of childhood disintegrative disorder applied in instances where a marked and profound developmental regression and development of autistic like behaviors occurred after a clear period of normal development. This rare condition is of considerable interest, for example, it might represent the action of specific genetic or other mechanisms. It is unclear how many such children who now come to diagnosis will receive an ASD diagnosis, since the mean age of onset is approximately 4 years. Similarly individuals who exhibit a pattern of social difficulties more consistent with the DSM-IV Asperger diagnosis may present problems given the emphasis on early onset and their higher-cognitive levels. In such cases social interest is often very present although pragmatic skills are typically weak while vocabulary and other aspects of language may be quite advanced. Table 44–5 summarizes some of the features important in differentiating autism from other disorders in the first years of life.

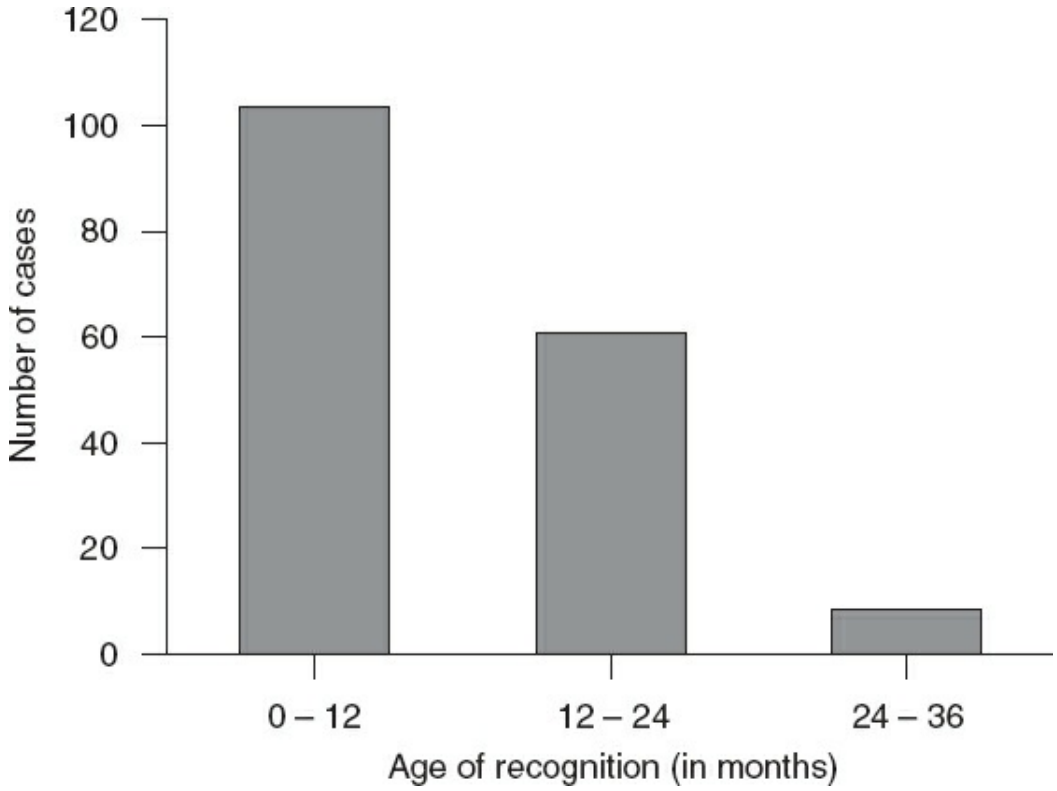

FIGURE 44–2. Age of recognition in autism in a series of 174 consecutive cases.

Impairments in Social Communication. Normally developing infants have a marked interest in social interaction and the social environment from birth. This predisposition is an important foundation for development of other skills. In infants and young children with autism, the human face holds little interest; disturbances are seen in the development of joint attention, attachment, and other aspects of social interaction. For example, the child may not engage in the usual games of infancy, may have difficulties with imitation, and may lack usual play skills. These deficits are highly distinctive and are not just due to associated developmental delay.

Table 44–5.
Symptoms Differentiating Infants and Toddlers with Autism Form Typical and Developmentally Delayed Peers

	Social Interaction	Communication	Stereotypical and Repetitive Behaviors
First year	Limited ability to anticipate being picked up	Poor response to name	Excessive mouthing
	Low frequency of looking at people	Infrequent looking at objects held by others	Aversion to social touch
	Little interest in interactive games		
	Little affection toward family members		
	Content to be alone		
Second	Abnormal eye contact	Low frequency of verbal or	Hand/finger mannerisms

and third years		nonverbal communication	
	Limited social referencing	Failure to share interest (e.g., by pointing, giving, showing)	Inappropriate object use
	Limited interest in other children		Repetitive interests/play
	Limited social smile		Unusual sensory behavior (hyper/hyposensitivity to sounds, texture, taste, or visual stimuli)
	Low frequency of looking at people	Failure to respond to communicative gestures (showing, pointing, giving)	Poor response to name
	Limited sharing of affect/enjoyment	Use of other's body as tool (takes hand to get object without making eye contact as if hand is agent)	
	Limited range of facial expressions	Unusual vocalizations	
	Unusual attachment object (hard not soft, the specific object less important than object less important than category of object)		
	Little interest in interactive games		
	Limited functional play		
	No pretend play		
	Limited motor imitation		

Adapted from Chawarska K, Volkmar FR. Autism in infancy and early childhood. In: Volkmar F, Klin A, Paul R, Cohen D, eds. *Handbook of Autism and Pervasive Developmental Disorders*. 3rd ed. Hoboken, NJ: John Wiley & Sons; 2005:230.

Social interest may increase over time. There is often a developmental progression, with younger and more impaired individuals being avoidant or aloof from interaction, whereas somewhat older or more advanced individuals are willing to passively accept interaction but do not seek it out. Among the most able individuals with autism, there is often social interest, but the person has difficulty in managing the complexities of social interaction; this often leads to an unusual or eccentric-appearing social style.

In terms of communication as many as 30 to 40 percent of individuals with autism never use language for communication (e.g., they never speak or only echo or repeat statements of others). Delays in the acquisition of language are the most frequent presenting complaint of parents. Usual patterns of language acquisition are observed (e.g., children with autism may not babble or may take the parent's hand to obtain a desired object

without making eye contact, as if the hand, rather than the person, is obtaining the item). In contrast to the child with a language disorder, there is no apparent motivation to engage in communication or attempt to communicate via nonverbal means.

When individuals with autism do speak, their language is remarkable in various ways. They may echo speech (echolalia). Speech tends to be less flexible, for example, there is no appreciation that change in perspective or speaker requires pronoun change; this leads to pronoun reversal. Speech may be nonreciprocal in nature, for example, the child produces language that is not meant as communication. Although the syntax and morphology of language are relatively spared, vocabulary and semantic skills may be slow to develop and aspects of the social uses of language (pragmatics) are particularly challenging. Thus, humor and sarcasm may be a source of confusion, as the person with autism may fail to appreciate the speaker's communicative intent, resulting in an overly literal interpretation of the utterance. Although typically developing babies acquire the prosody of their mother language during the first year of life, for more able individuals with autism, this aspect of language may not develop, and often, intonation is monotonic and robot like. Nonverbal behavior, as used in social interaction, is very impaired with individuals oblivious to clear social cues, for example, talking at great length about topics of interest to them but not their conversational partner.

Deficits in play may include a failure to develop usual patterns of symbolic–imaginative play. The autistic child may explore nonfunctional aspects of play materials (e.g., taste or smell) or use aspects of materials for self-stimulation (spinning the tires on a toy truck).

Restricted Activities and Interests. Children with ASD often have difficulty tolerating change and variation in routine. For example, an attempt to alter the sequence of some activity may be met with what appears to be catastrophic distress on the part of the child. Parents may report that the child insists that they engage in activities in very particular ways, and changes may elicit great opposition or upset. The child may develop an interest in a repetitive activity, for example, collecting strings and using them for self-stimulation, memorizing numbers, or repeating certain words or phrases. In younger children, attachments to objects, when they occur, differ from usual transitional objects in that the objects chosen tend to be hard rather than soft, and, often, it is the class of object, rather than the particular object, that is important. For example, the child may insist on carrying a certain kind of magazine around with him or her. Stereotyped movements may include toe walking, finger flicking, body rocking, and other mannerisms; these are engaged in as a source of pleasure or self-soothing. The child may be preoccupied with spinning objects, for example, he or she may spend long periods watching a ceiling

fan rotate. This group of behaviors tends to be one of the last to develop and may be minimal until the child is around three years of age.

Associated Features. In contrast to Kanner's early belief that children with autism had good cognitive potential, approximately 31 percent have intellectual disability and 23 percent have IQs in the borderline, below average range. The profile on psychological testing is marked by significant deficits in abstract reasoning, verbal concept formation, and integration skills and on tasks requiring a degree of social understanding. For example, on the Wechsler scales, weaknesses are usually obtained on the similarities and comprehension subtests, while relative strengths are usually observed in tasks involving rote learning and memory skills. Visual–spatial problem solving is often a strength, particularly if the task can be completed piecemeal, that is, without having to infer the context, or gestalt, of the task. The typical preference for rote and sequential, rather than conceptual and integrative tasks usually carries the implication that individuals with autism fail to see "the trees from the leaves," a difficulty that cuts across functioning modalities—from cognitive testing to language and communication, to style of social interaction. Given that verbal deficits are common in autism, performance on the Wechsler scales is also usually characterized by higher perceptual reasoning than verbal comprehension scores, particularly in the individuals scoring within the range of intellectual disability. However, the opposite pattern is sometimes seen in individuals with above average or gifted intelligence.

Several influential psychological theories of the social dysfunction in autism have been proposed. One hypothesis posits that there is a lack of a central drive for coherence, with the consequent focus on dissociated fragments rather than integrated "wholes," leading to a fragmentary and overly concrete experience of the world. Another hypothesis posits deficits in executive functioning, defined as the capacity for abstracting rules, inhibiting irrelevant responses, shifting attention and profiting from feedback, and maintaining a focus on multiple aspects of information in decision making. As executive functions are thought to be mediated by frontal brain areas, this hypothesis has historically highlighted some similarities between autism and conditions resulting from acquired frontal lobe lesions. Yet another influential hypothesis posits that autism results from the child's inability to attribute mental states (such as beliefs and intentions) to others. Devoid of such a "theory of mind," individuals with autism are presumed to be unable to infer the thoughts and motivations of others, thus failing to predict their behavior and to adjust accordingly. This results in lack of reciprocity in communication and social contact, and great challenges in successfully navigating social situations that require contextual understanding of implicit social mores and of nonliteral speech such as humor and irony. In the past decade or so, eye-tracking technology

has been used to study social–visual engagement in a highly quantitative fashion, with a focus on abnormalities of skills such as preferential visual orientation to salient aspects of social stimuli that typically facilitates social engagement (Fig. 44–3). This line of research has documented significant impairments in social orientation and motivation skills, skills that are both developmentally early emerging and evolutionarily highly conserved. Examples include preferential orientation to biological motion (or movement of living beings) and eye gaze, both of which are essential components of joint or shared attention, which in turn is a key component of reciprocal social engagement and communication. These findings suggest that social development is likely derailed very early on in the lives of children with autism, leading to a cascade of developmental events that have an impact on subsequent social cognitive milestones and on brain specialization. More recently, investigators have capitalized on the high recurrence rates of autism within sibships, following prospectively and longitudinally the developmental trajectories of large cohorts of younger siblings of children with autism. One to two in five among the younger siblings either meets criteria for autism or exhibits subthreshold levels of autism-related symptoms. Using eye tracking, EEG, sMRI, fMRI, and DTI methods, prodromal markers of autism have already been identified in the first year of life of children later diagnosed with autism.

FIGURE 44–3. Eye tracking data from a typically developing young adult (*top line*) and a young, high-functioning man with autism (*bottom line*). The man with autism focuses almost exclusively on the mouth of the person speaking, avoiding the upper part of the face and not noticing the emotional responses of the listener. (From Klin A, Jones W, Schultz R,

Volkmar F, Cohen D. Defining and quantifying the social phenotype in autism. *Am J Psychiatry*. 2002;159(6):895–908, with permission.)

Although all of these hypotheses have informed clinical practice and guided a surge of research into the neuroscience of social dysfunction in autism, challenges remain in tying these primarily descriptive findings to specific mechanisms of pathogenesis and pathobiology. Critically, there is still a vast gap between advances in molecular genetics research, which has implicated hundreds of genetic variants in the etiology of autism, and the still limited understanding of the developmental unfolding of autism symptoms in the first 2 years of life. The next generation of studies is likely to advance gene–brain–behavior research by probing specific brain systems known to be affected in autism via coregistration measurements of genetic and epigenetic liabilities and of their instantiation as behavioral or brain endophenotypes.

One of the most fascinating cognitive phenomenon in autism is the presence of so-called islets of special abilities or splinter skills, which are preserved or very highly developed skills in specific areas that contrast with the child's overall deficits in cognitive functioning. It is not unusual, for example, for children with autism to have great facility in decoding letters and numbers, at times precociously (hyperlexia), although comprehension of what is read is much impaired. Perhaps as many as 10 percent of individuals with autism exhibit a form of "savant" skill, that is, high, sometimes prodigious performance on a specific skill in the presence of mild or moderate intellectual disability. This fascinating phenomenon usually relates to a narrow range of capacities—memorizing lists or trivial information, calendar calculation, visual–spatial skills such as drawing, musical skills involving a perfect pitch or playing a piece of music after hearing it only once. Interestingly, individuals with autism represent a disproportionate majority of all savant persons. Individuals with normative or gifted IQs often display highly circumscribed interests, typically expressed in memorization of vast amounts of factual information. Such narrow interests can interfere with learning more generally and can also result in a one-sided, nonreciprocal style of communication with others.

MOTOR AND SENSORY ABNORMALITIES. One of the prototypical aspects of autism is the display of motor stereotypies such as hand flapping, body rocking, finger twiddling, or waving in front of the eyes, as well as other repetitive and purposeless motor mannerisms such as toe walking and the assumption of odd postures. Stereotypies are also seen in other conditions, particularly severe intellectual disability and sensory impairments such as blindness and deafness, but are clearly more common in autism, often emerging around the ages of 3 or 4 although more subtle expressions have been reported as early as the second year of life. They often decrease in adolescence and adulthood, and are less common in individuals with

normative levels of intellectual functioning. Some children with autism may copy other people's motor movements (echopraxia) without necessarily learning the purpose of that movement. This is, to some extent, the motor equivalent of echolalia. Normative motor skills are usually preserved relative to the child's intellectual level, although this may not be the case over time, as the social aspects of motor activity become more important as the child becomes older. In fact, older and higher-functioning individuals are often perceived as motorically "clumsy" or awkward. Over activity is common in autism, particularly in the preschool years. At the same time, a lack of curiosity in the environment and passivity or hypoactivity, may also be observed in children with autism. The two may alternate for the same child in different settings or in relation to different activities. Task engagement and attention are important considerations in educational programming.

Both hypo- and hypersensitivity to sensory stimuli are typical of children with autism as they are in other developmental disorders. Children with autism may be acutely sensitive to some sounds (hyperacusis); they may, for example, cover their ears when hearing a dog bark or the noise of a vacuum cleaner. Other children may appear oblivious to some loud noises or to people calling them, and yet may be fascinated by the faint ticking of a wristwatch or the sound of crumpling paper. Bright lights may be distressing, although some children are fascinated by light stimulation and may, for example, move an object back and forth in front of their eyes. There may be extreme sensitivity to touch (tactile defensiveness), including major reactions to specific fabrics or social or affectional touch, and yet many children appear insensitive to pain and may not cry after a severe injury. There is often a fascination with certain sensory stimuli, such as spinning objects or toy parts that can spin, and some children enjoy vestibular sensations such as twirling, engaging in this action without apparently becoming dizzy.

Physical Characteristics

With autism, there is a higher incidence of minor physical anomalies such as ear malformations. Such anomalies may reflect the embryological period in which the factors responsible for autism act, for example, ears are formed at around the same time as older regions of the brain develop. However, most young children with autism are often described as attractive and do not exhibit any forms of stigmata. Physical attractiveness may diminish in adolescence and adulthood, as the various disabilities and unusual behaviors have an impact on physical appearance. A significant number of toddlers with autism develop macrocephaly.

Pathology and Laboratory Examination

Electroencephalography and Seizure Disorders. A variety of electroencephalographic (EEG) abnormalities may be seen in autism, including diffuse and focal spikes, paroxysmal spike-and-wave patterns, multifocal spike activity, and a mixed discharge. The incidence of EEG abnormalities in autism (in the absence of a clinical seizure disorder) ranges from 10 to 83 percent and depends on the number of recordings and the nature of the sample obtained (Fig. 44–4). The incidence of abnormal EEGs is significantly higher in mentally retarded autistic individuals. The finding of high incidence of EEG abnormalities and seizure disorders in autism was one of the first compelling pieces of evidence supporting a biological basis for the condition.

Both early and more recent studies on auditory brainstem–evoked potentials in autism indicate that if children with underlying neurological conditions are excluded and age and gender are controlled, there is no evidence for abnormalities in the auditory brainstem pathways. However, abnormalities of cognitive potentials, particularly the auditory P300 (which represents the brain's processing of sensory stimuli), have been shown to be abnormal in autism. This presumably reflects abnormalities in higher auditory processing and neural pathways. A growing body of research uses EEG to explore issues like recognition of faces with characteristic patterns observed in individuals with autism as compared to TDCs.

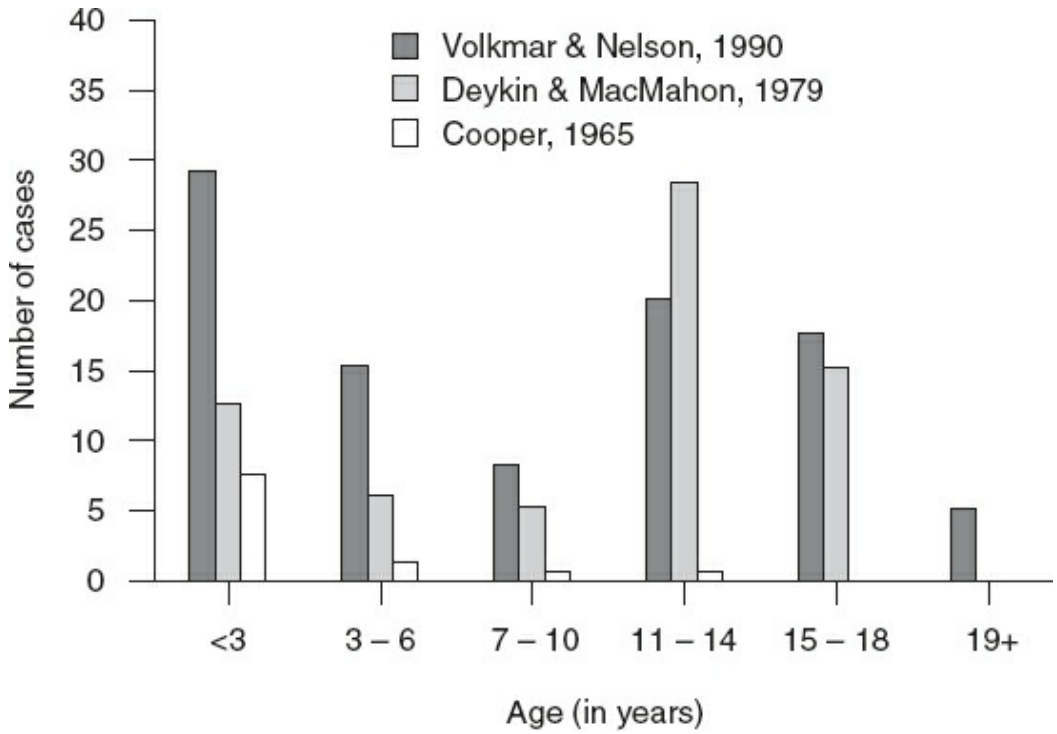

FIGURE 44–4. Rates of first seizure (incidence rate per 1,000) in samples of individuals with autism. (Data from cases with autism from Volkmar F, Nelson D. Seizure disorders in autism. *J Am Acad Child Adolesc Psychiatry*. 1990;29:127–129; and from cases with autism/pervasive developmental disorder from Deykin EY, MacMahon B. The incidence of seizures among children with autistic symptoms. *Am J Psychiatry*. 1979;136:1310–1312; as compared to rates in a normative sample from Cooper JE. Epilepsy in longitudinal survey of

5,000 children. *Br Med J*. 1965;1:1020–1022.)

The development of epilepsy in autism occurs in approximately 10 to 25 percent of the autistic population by young adulthood, based on the occurrence of readily identifiable major motor seizures. Although there is evidence of seizure onset at all ages, early childhood and adolescence have been reported to be the peak periods. Lower-functioning individuals are at increased risk. The onset of seizures may be associated with deterioration.

Neuroimaging Studies. As noted previously, a growing body of work has emerged on the neuroanatomical and neurophysiological substrates of autism using neuroimaging and other techniques. However, neuroimaging studies are not routinely needed as part of the evaluation of a child with potential autism in the absence of a specific indication.

Laboratory Studies. As indicated by history and examination various laboratory tests may be relevant as part of the assessment process. These include screening for conditions known to be more frequently associated with autism, for example, fragile X and tuberous sclerosis. Increasingly sophisticated genetic testing is also now available; indications are summarized in practice guidelines provided in the reference list. It should be noted that knowledge in this area is evolving rapidly and recommendations are frequently updated.

Differential Diagnosis

ASD must be differentiated from a range of other problems (some of which are, at times, associated with it). These include neurodevelopmental disorders such as intellectual deficiency or language–communication disorders. Visual and hearing impairments must also be considered. A variety of useful and well-developed instruments are available to aid in diagnosis—these include both parent interviews and direct assessment of the individual. It should be noted that some of these require extensive training and results should not replace clinical judgment. Selected instruments are summarized in Table 44–6.

In mild to moderate intellectual deficiency (mental retardation), social skills are usually consistent with abilities in the areas of cognition and communication. In individuals with severe and profound retardation, the frequency of both autism and autistic-like behaviors increases. A common source of diagnostic error is the confusion of stereotypies in such individuals. In higher-functioning individuals, there is potential confusion relative to those personality disorders that involve social isolation.

Table 44–6.
Summary of Selected Assessment Instruments for Autism Spectrum

Disorder

Scale (See Legend)	Uses	Age Range	Method of Administration	Population Studied	Scale Characteristics	References
ABC	Screening	Children	Teacher/parent	AD	57 items, scale 1–4	Krug et al., 1980
CARS II	Screening	3 to adult	Clinician rated	AD	15 items, severity ratings	Schopler & van Bourgondian, 2010
M-Chat	Screening	Toddlers	Parent rated	AD	23 items, yes/no	Robins et al., 2001
CSBS-DP-IT-Checklist	Screening	Toddlers	Parent rated	AD	24 items	Wetherby et al., 2008
ASQ	Screening	Child/adult	Parent rated	AD/AspD	40 items, yes/no	Berument et al., 1999
AQ	Screening	Child/adult	Self or parent rated	AspD	50 items, scale 0–3	Baron-Cohen et al., 2001
CAST	Screening	4–11 years	Parent rated	AspD	37 items, yes/no	Scott et al., 2002
ASDS	Screening	5–18 years	Parent or teacher rated	AspD	50 items, yes/no	Myles et al., 2000
GADS	Screening	3–22 years	Parent or teacher rating	AspD	32 items, scale 0–3	Gilliam, 2001
ASDI	Screening	Child/adult	Interview + clinician rated	AspD	50 items, yes/no	Gillberg et al., 2001
SRS	Screening	4–18 years	Parent or teacher rating	AspD	65 items, scale 1–4	Constantino et al., 2003
ADI	Diagnostic	Child/adult	Interview + clinician rated	AD/AspD	See text	Lord et al., 2003
DISCO	Diagnostic	Child/adult	Interview + clinician rated	AD/AspD	See text	Wing et al., 2002
ADOS	Diagnostic	Child/adult	Semi-structured interactive session	AD/AspD	See test	Lord et al., 1994

ABC, Autism Behavior Checklist; AD, autism disorder; ADI, Autism Diagnostic Interview-Revised; ADOS, Autism Diagnostic Observation Schedule; AQ, Autism Quotient; ASDI, Asperger Syndrome Diagnostic Interview; ASDS, Asperger Syndrome Diagnostic Scale; AspD, Asperger disorder; ASQ, Autism Screening Questionnaire; CARS, Childhood Autism Rating Scale; CAST, Childhood Autism Screening Test; M-CHAT, Checklist for Autism in Toddlers; CSBS-DP-IT-Checklist, Communication and Symbolic Behavior Scales Developmental Profile Infant-Toddler Checklist; DISCO, Diagnostic Interview for Social and Communication Disorders; GADS, Gilliam Asperger Disorder Scale; Parent, primary caregiver; SRS, Social Responsiveness Scales. These instruments may need to be revised to provide evidence of validity for DSM-5 ASD and supplement but DO NOT REPLACE clinical diagnosis.

Adapted from Volkmar F, Siegel M, Woodbury-Smith M, et al. Practice parameters for the assessment

and treatment of children and adolescents with autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry*. 2014;53(2):237–257.

In obtaining a history, the use of videotapes or other aids to memory may be helpful. It also may help if the parent is asked to recall the child at certain particular dates (e.g., third birthday). In most cases of autism, there is not a history of normal development, although in a small group of cases (20 to 25 percent), parents may report an apparent period of normal, or near-normal, development before either developmental stagnation or developmental regression. A careful review is indicated in these cases. Many times, some abnormality in development or behavior will have antedated parental concern. A history of clearly normal development raises the possibility of selective mutism, Rett syndrome, the DSM-IV childhood disintegrative disorder, language disorder, schizophrenia, and degenerative CNS disorders. Typically, children who have experienced severe neglect exhibit delayed or deviant social skills, but other features of autism are not present and the social deficits remit with appropriate care. In childhood-onset schizophrenia, there usually has been a long period (many years) of normal or near-normal development before the onset of characteristic delusions, hallucinations, and so forth. In some instances, the diagnosis can be clarified with certainty only over time.

John was the second of two children born to middle-class parents after normal pregnancy, labor, and delivery. As an infant, John appeared undemanding and relatively placid; motor development proceeded appropriately, but language development was delayed. Although his parents indicated that they were first concerned about his development when he was 18 months old and still not speaking, in retrospect, they noted that, in comparison to their previous child, he had seemed relatively uninterested in social interaction and the social games of infancy. Stranger anxiety had never really developed, and John did not exhibit differential attachment behaviors toward his parents. Their pediatrician initially reassured John's parents that he was a "late talker," but they continued to be concerned. Although John seemed to respond to some unusual sounds, the pediatrician obtained a hearing test when John was 24 months old. Levels of hearing appeared adequate for development of speech, and John was referred for developmental evaluation. At 24 months, motor skills were age appropriate, and John exhibited some nonverbal problem–solving skills close to age level. His language and social development, however, were severely delayed, and he was noted to be resistant to changes in routine and unusually sensitive to aspects of the inanimate environment. His play skills were quite limited, and he used play materials in unusual and idiosyncratic ways. His older sister had a history of some learning difficulties, but the family history was otherwise negative. A comprehensive medical evaluation revealed a normal EEG and computed tomography (CT) scan; genetic screening and chromosome analysis were normal as well.

John was enrolled in a special education program, in which he gradually began to speak. His speech was characterized by echolalia, extreme literalness, a monotonic voice quality, and pronoun reversal. He rarely used language in interaction and remained quite isolated. By school age, John had developed some evidence of differential attachments to family members; he also had developed a number of self-stimulatory behaviors and engaged in occasional periods of head banging. Extreme sensitivity to change continued. Intelligence testing revealed marked scatter, with a full-scale IQ in the moderately retarded range. As an adolescent, John's behavioral functioning deteriorated, and he developed a seizure disorder. Now an adult, he lives in a group home and attends a sheltered workshop. He has a rather passive interactional style but exhibits occasional outbursts of aggression and self-abuse.

Course and Prognosis

Autism is generally a lifelong disability, with most individuals needing significant family and community support. Although increasing claims for cure are reported (and indeed more individuals appear to be entering college and the work force), even in the most able individual, some social oddities persist. Changes in outcome likely reflect early diagnosis and intervention preventing or minimizing some of the barriers that ASD presents for typical learning. The first outcome studies were rather pessimistic regarding long-term outcome with perhaps 5 percent of cases achieving independence and self-sufficiency and two-thirds or more need high levels of intensive support. Over time these numbers have changed (as have diagnostic and intervention practices) but it is clear that more individuals are able to achieve adult self-sufficiency.

In other instances some support is needed in adulthood. Even in more cognitively challenged individuals, most will exhibit some improvement in social relatedness, communication, and self-help skills, and a major long-term goal should be used to facilitate acquisition of important adaptive skills, with the aim of maximizing the potential for independence.

Various factors predict long-term outcome. These include, in particular, the presence of some communicative speech by the age of 5 or 6 and higher nonverbal–intellectual level. Overall cognitive ability in the normal range is a positive prognostic sign. The relationship of severity of autistic symptomatology in early childhood to adult symptoms is somewhat less clear. With early intervention, some children make significant gains, whereas other children provided with such intervention may not. Several groups have now shown the potential for observable brain changes in response to treatment using EEG or fMRI response to social stimuli.

Gains in social compliance and communication are often made during elementary school years, particularly if structured, individualized, and intensive interventions are in place. During adolescence, some autistic children may exhibit behavioral deterioration; for a minority among these, the decline in language and social skills may be associated with the onset of a seizure disorder. Various interactional styles can be observed, ranging from aloof to passive to eccentric (e.g., children who make attempts to initiate contact with others but do so in a very awkward or rigid fashion). These styles are related to the developmental level. Depressive and anxiety symptoms may appear in higher-functioning adolescents, who become painfully aware of their inability to form friendships despite a desire to do so and begin to experience the cumulative effect of years of failed contact with others and teasing by peers.

Treatment

The goals of treatment for children with autism are to reduce disruptive behaviors and promote learning, particularly in the areas of language acquisition, communication, and self-help skills. These goals are best achieved based on a comprehensive assessment to determine a profile of strengths and needs, and an individualized and structured program is implemented by professionals experienced in working with children who are autistic. A recent review of several treatment studies highlighted the importance of intensive and early intervention in individual outcome. Treatment goals should be updated regularly, as different issues should be prioritized as a function of the child's rate and profile of progress and age (e.g., vocational training and independent living skills should be aggressively promoted in adolescence). Additionally, psychopharmacological treatments may be beneficial when specific indications are present.

Educational Approaches. Autistic children require intensive and highly structured special education from as early as the child is able to tolerate a school routine. Given the challenges involved in teaching children with autism, a classroom setting with a low student to teacher ratio is usually essential. For the more impaired children, a typical hierarchy of priorities should include the ability to: (1) tolerate individual adult guidance in performing tasks; (2) consistently follow a daily routine; (3) develop communication intent and communication means; and (4) move from associative to conceptual learning. Learning should take place in an environment that minimizes distractions (e.g., nearby windows, richly decorated walls). For the more challenged children, individual workstations and forceful adult intrusion may be necessary. The use of highly predictable and consistent routines is necessary to eventually promote the child's own internal sense of order, scheduling, and organization of experiences, thus promoting more systematic learning. Children with autism must be helped to generalize skills, for example, using new words acquired at home as well as in school. Both the ability to better predict impending activities and the ability to better express one's wishes and protests are usually beneficial in reducing frustration and associated disruptive behaviors. A number of well-established evidence-based treatment models are now available. These include approaches relying on applied behavior analysis (ABA), models using a more developmental approach (The Early Start Denver Model [ESDM]), hybrid approaches in which developmental models inform behavioral treatments as in pivotal response therapy, and eclectic approaches drawing on a range of intervention methods.

The focus of speech–language–communication therapy is on the use of words for the purpose of meaningful communication. Children with autism may acquire a considerable vocabulary that is dissociated from the act of

communication. Therefore, language acquisition should go hand in hand with the promotion of the child's intent to communicate with others. For this purpose, the focus of vocabulary expansion should be on words that are relevant to everyday life. Children who do not vocalize should be engaged in programs focused on alternative forms of communication, including signs, communication boards, or other forms of augmentative communication. The use of nonvocal forms of communication should not preclude the simultaneous use of words in children whose vocalizations are just emerging and for whom vocal communication is a realistic goal. For older or higher-functioning children, the core of the educational program should be an intensive focus on social and communication skills training. Positive actions in frequently troublesome situations may have to be rehearsed and scripted. Concrete social and communication skills—including eye gaze, voice modulation, gestural communication, posture, proximity, greeting behaviors, rules of conversation, and social expectations—may have to be taught in a very explicit fashion. The setting for the social and communication skills therapy may have to alternate between small group instruction (in which appropriate behaviors can be practiced and supportive feedback can be gained) and naturalistic settings (in which the newly acquired skill can be put to practice or additional problematic behaviors can be identified for practice in the small setting). Successful techniques used for this purpose include modeling of behaviors by an instructor, self-observation, role playing, and the use of individualized social stories. The advent of numerous computer interventions and applications adaptable to the individual's situation and levels of functioning are now available.

Behavior Therapy. This treatment, based on learning theory principles, uses behavior-modification techniques to establish desired behaviors and eliminate problem behaviors. Most educational programs for children with autism use behavioral management techniques, although they vary in terms of how integrated these procedures are in the comprehensive educational program. Such techniques can be particularly useful in the management of disruptive behaviors that impact on the child's ability to engage in remedial educational intervention. After a functional analysis of the target behavior is performed and patterns of reinforcement are identified, techniques such as shaping, prompting, and extinction are used to promote the desired behavioral alternative, which is then reinforced by increasingly mature rewards. Behavior therapy is also used for the facilitation of learning, including the promotion of early cognitive skills such as categorization and elicitation of vocalization and speech. It is clear that many children, although not all, will significantly benefit from behavior therapy; other forms of teaching can also be used particularly for more intellectually able children.

Psychotherapy. With the shift from a psychogenic to a biological understanding of autism, it became clear that psychodynamic psychotherapy and unstructured play therapies, in general, were not appropriate in the treatment of young children. Individual psychotherapy may be appropriate for higher-functioning individuals who may present with anxiety and depressive symptoms as they grow older and become more aware of their differences and difficulties relating to others. In these cases, psychotherapy should focus on rather explicit problem-solving skills rather than being insight oriented, with the goal to promote better adjustment and self-satisfaction. New approaches that include modified CBT specific to individuals with ASD should be considered.

Psychopharmacology. No pharmacological agent has proven curative, but certain medications may be of benefit with regard to specific symptoms although work is now ongoing on agents that might have this potential, for example, oxytocin. Such target symptoms may include self-injury, aggression, stereotyped movements, and over activity. Because individuals with autism are often enrolled in educational and behavioral programs, it may be possible to have staff participate (e.g., by collecting behavioral data that can be used to monitor effectiveness of the drug). As with all medications, the potential benefits and side effects should be considered (e.g., with the major tranquilizers, there is potential for over sedation, which can be mistaken for a positive therapeutic response). Before beginning drug treatment, baseline laboratory studies (tests of liver and renal function, electrocardiogram [ECG], urinalysis, blood count, weight, and blood pressure) should be conducted. Typically, an examination for abnormal movements is performed (typically using the Abnormal Involuntary Movement Scales [AIMS]).

MAJOR TRANQUILIZERS. Major tranquilizers have been the most extensively studied agents in autism. At relatively low doses, they may decrease stereotyped behaviors and agitation and may help the individual profit from remedial programming. Some data suggest that the combination of these agents with behavior therapy is more effective than either treatment alone. Side effects may limit the usefulness of these agents (e.g., sedation, withdrawal, and tardive dyskinesia). The higher-potency neuroleptics have often been used given the lower likelihood of sedation, although these agents do carry increased risk for dystonias. Recent interest has centered on the atypical neuroleptics, and these agents appear to offer considerable potential promise. A large double-blind, placebo-controlled study of one of these agents has shown significant benefit in reducing agitation and other problematic behavioral difficulties and risperidone (Risperdal) has now been approved for use in autism.

SELECTIVE SEROTONIN REUPTAKE INHIBITORS. Several reports suggested the potential usefulness in autism and related conditions of the selective serotonin reuptake inhibitors (SSRIs). The potential use of such agents was suggested by the observation of higher peripheral serotonin levels in autism. In several controlled studies, these medications have proven superior to placebo relative to symptoms such as repetitive behaviors and impulsive aggression. There is a suggestion that seizure disorders may be exacerbated in some cases. Occasionally, children become agitated on these agents.

CLONIDINE. Given the apparently high levels of arousal in autism, use of clonidine (Catapres) was suggested. This agent, an α_2-adrenergic agonist, reduces noradrenergic activity. There may be a modest effect on activity levels, although sedation and hypotension can present serious side effects.

STIMULANTS. Individuals with autism, particularly school age children, may exhibit significant problems with attention that respond to traditional stimulant and related medications. Unfortunately the level of side effects is higher than in the regular ADD population.

OTHER AGENTS. Various other agents have been reported to sometimes be effective in treating at least some of the symptoms of autism. Such evidence is usually based on single case reports, and controlled studies are often lacking. Lithium is generally not clearly helpful unless there is a personal or family history of bipolar illness. Although the available data are limited, it appears that, in general, stimulants exacerbate behavioral difficulties in autism, probably acting through their effect on dopamine. β-Blockers have been used in several open trials concerned with reducing aggressive, self-injurious, and impulsive behaviors. These agents may act to decrease arousal levels, and positive responses have been noted, although the potential for serious side effects has limited their more general use.

Parents often are interested in alternative treatments, including diet and vitamin therapy, among others. Although the focus of much interest, substantive data are lacking. Usually, a single case report or uncontrolled study is reported, and tremendous enthusiasm follows, but controlled studies do not substantiate the claims made. This was recently observed for secretin, a gut peptide, which, in a series of double-blind, controlled studies, failed to offer significant improvement over placebo. Various somatic treatments have not proven clinically useful. In general, pursuit of unproven treatments should be avoided, particularly if they are pursued at the expense of educational and behavioral interventions that are known to be efficacious. In some instances, treatments that pose actual danger to the child have been proposed and, clearly, should be avoided. Given the relatively poor prognosis associated with autism, parents are easily attracted by treatments that propose a rapid cure. Parents should be

encouraged to pursue treatments with known efficacy.

SOCIAL (PRAGMATIC) COMMUNICATION DISORDER

Definition

This condition, newly recognized in DSM-5, is characterized by persistent difficulties in the use of social and nonverbal communication. The problems must be obvious in social contacts, include difficulties in matching the communication to the context or conversational partner, problems in following or conducting a conversation, and in understanding implicit meanings. These difficulties must appear in the developmental period, cannot be due to another condition, and must be the source of some functional limitation in effective social communication.

REFERENCES

Asperger H. Autistic psychopathy in childhood. In: Frith U, translated and editor. *Autism and Asperger Syndrome*. Cambridge, MA: Cambridge University Press; 1991:37.

Button KS, Ioannidis JP, Mokrysz C, et al. Power failure: Why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013;14(5):365–376.

Chawarska K, Macari S, Volkmar FR, Kim SH, Shic F. ASD in infants and toddlers. In: Volkmar FR, Paul R, Pelphrey K, eds. *Handbook of Autism and Pervasive Developmental Disorders*. 4th ed. Hoboken, NJ: John Wiley & Sons; 2014.

*Chevallier C, Kohls G, Troiani V, Brodkin ES, Schultz RT. The social motivation theory of autism. *Trends Cogn Sci*. 2012;16(4):231–239.

Courchesne E, Pierce K, Schumann CM, et al. White matter microstructure and atypical visual orienting in month-olds at risk for autism. *Am J Psychiatry*. 2013;170(8):899–908.

Hazlett HC, Poe MD, Gerig G, et al. Early brain overgrowth in autism associated with an increase in cortical surface area before age 2 years. *Arch Gen Psychiatry*. 2011;68(5):467–476.

Heller T. Dementia infantilis. *Zeitschrift Erforschung Behandlung Jugenlichen Schwachsinns*. 1908;2:141–165.

*Hermelin B. *Bright Splinters of the Mind: A Personal Story of Research with Autistic Savants*. London: Jessica Kingsley; 2001.

*Herrington JD, Schultz RT. Neuroimaging of developmental disorders. In: Shenton ME, Turetsky BI, eds. *Understanding Neuropsychiatric Disorders: Insights From Neuroimaging*. Cambridge: Cambridge University; 2011.

Howlin P. Outcomes in adults with autism spectrum disorders. In: Volkmar FR, Paul R, Pelphrey K, eds. *Handbook of Autism*. 4th ed. Hoboken, NJ: John Wiley & Sons; 2014.

Huerta M, Bishop SL, Duncan A, Hus V, Lord C. Application of DSM-5 criteria for autism spectrum disorder to three samples of children with DSM-IV diagnoses of pervasive developmental disorders. *Am J Psychiatry*. 2012;169(10):1056–1064.

Kanner L. Autistic disturbances of affective contact. *Nerv Child*. 1943;2:217–250.

Klin A, Jones W, Schultz R, Volkmar F, Cohen D. Defining and quantifying the social phenotype in autism. *Am J Psychiatry*. 2002;159:895–908.

Lord C. Follow-up of two-year-olds referred for possible autism. *J Child Psychol Psychiatry*. 1996;36:1065–1076.

Lord C, Corsello C, Grzadzinski R. Diagnostic instruments in autistic spectrum disorders. In: Volkmar FR, Rogers SJ, Paul R, Pelphrey K, eds. *Handbook of Autism*. 4th ed. Hoboken, NJ: John Wiley & Sons; 2014.

Lyall K, Schmidt RJ, Hertz-Picciotto I. Environmental factors in the preconception and prenatal periods in relation to risk for ASD. In: Volkmar FR, Paul R, Pelphrey K, eds. *Handbook of Autism*. 4th ed. Hoboken, NJ: John Wiley & Sons; 2014.

McCracken JT, McGough J, Shah B, et al; Research Units on Pediatric Psychopharmacology Autism Network. Risperidone in children with autism and serious behavioral problems. *N Engl J Med*. 2002;347:314–321.

McPartland JC, Pelphrey KA. The implications of social neuroscience for social disability. *J Autism Dev Dis*. 2012;42(6):1256–1262.

National Research Center. *Educating Young Children with Autism*. Washington, DC: National Academy Press; 2001.

Powers MD, Palmieri MJ, Egan SM, Rohrer JL, Nulty EC, Forte S. Behavioral assessment of individuals with autism: current practice and future directions. In: Volkmar FR, Paul R, Pelphrey K, eds. *Handbook of Autism*. 4th ed. Hoboken, NJ: John Wiley & Sons; 2014.

Presmanes-Hill A, Zuckerman KE, Fombonne E. Epidemiology of autism spectrum disorders. In: Volkmar FR, Paul R, Pelphrey K, eds. *Handbook of Autism*. 4th ed. Hoboken, NJ: John Wiley & Sons; 2014.

Rutter M, Bailey A, Bolton P, Le Couteur A. Autism and known medical conditions: myth and substance. *J Child Psychol Psychiatry*. 1994;35(2):311–322.

Rutter M, Thapar A. Genetics of autism spectrum disorders. In: Volkmar FR, Paul R, Pelphrey K. *Handbook of Autism*. 4th ed. Hoboken, NJ: John Wiley & Sons; 2014.

*Schaefer GB, Mendelsohn NJ. Genetics evaluation for the etiologic diagnosis of autism spectrum disorders. [erratum appears in *Genet Med*. 2008;10(6):464.] *Genet Med*. 2008;10(1):4–12.

Smith IC, Reichow B, Volkmar FR. The effects of DSM-5 criteria on number of individuals diagnosed with autism spectrum disorder: a systematic review. *Journal of Autism and Developmental Disorders*. 2015;45(8):2541–2552.

Tang G, Gudsnuk K, Kuo SH, et al. Loss of mTOR-dependent macroautophagy causes autistic-like synaptic pruning deficits. *Neuron*. 2014;83(5):1131–1143.

Volkmar F. Autism and the pervasive developmental disorders. In: Lewis M, ed. *Child and Adolescent Psychiatry: A Comprehensive Approach*. 2nd ed. Baltimore, MD: Williams and Wilkins; 2002:489–502.

Volkmar FR. A brief history of autism. In: Fatemi S, ed. *The Molecular Basis of Autism*. New York, NY: Springer; 2015:3–14.

*Volkmar F, Siegel M, Woodbury-Smith M, King B, McCracken J, State M; American Academy of Child and Adolescent Psychiatry (AACAP) Committee on Quality Issues (CQI). Practice parameters for the assessment and treatment of children and adolescents with autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry*. 2014;53(2):237–257.

45 Attention-Deficit Disorders

▲ 45.1 Attention-Deficit/Hyperactivity Disorder

ALMA M. SPANIARDI, M.D., LAURENCE L. GREENHILL, M.D., AND LILY I. HECHTMAN, M.D.

INTRODUCTION

Attention-deficit/hyperactivity disorder (ADHD) is the most common childhood behavioral disorder diagnosed in outpatient settings in the United States. It is a chronic disorder which can cause impairment into adolescence and adulthood. In 2007, Polanczyk and colleagues estimated worldwide prevalence of ADHD in children and adolescents to be 5.29 percent using a comprehensive meta-analysis including hundreds of articles and more than 100,000 patients. Similarly, a meta-analysis from 2012 by Willicut using full DSM-IV diagnostic criteria found the prevalence of ADHD to be 5.9 to 7.1 percent. It has been suggested that the prevalence of ADHD is higher in males and younger children. The National Survey of Children's Health (NSCH), which showed that 11 percent of US school-aged children had been diagnosed with ADHD by 2011, an increase of 42 percent since the previous survey in 2003. It has been suggested that the increasing rates of diagnosis and treatment may be due to increased awareness and access to care as opposed to an increase in the number of children in the population meeting criteria for ADHD.

HISTORY

ADHD first was described by George Still in 1902, who wrote about children who were restless, impulsive, and inattentive, with intense affective responses and conduct problems. After the influenza pandemic and the epidemic of encephalitis lethargica in 1919 to 1920, children who survived the flu frequently developed severe behavior problems and are now thought to have suffered organic brain damage. For that reason, the childhood condition was termed "minimal brain damage syndrome," even though brain damage could not be proven. In the early 1960s, the condition was renamed "minimal brain dysfunction." Not only did these terms

postulate unproven etiological mechanisms, they were also stigmatizing. The ninth revision of the *International Classification of Diseases and Related Health Problems* (ICD-9) and the second edition of the DSM adopted the same descriptive term for the condition—hyperkinetic syndrome of childhood. The 1980 version of the APA classificatory system, DSM-III, renamed this diagnosis attention-deficit disorder (ADD) due to research suggesting that the main disability was the difficulty of maintaining sustained attention and impulsivity. Much of the diagnostic criteria for ADHD has remained the same between DSM-IV in 1994, the text revision, DSM-IV-TR in 2000, and the most recent version, DSM-5, released in May 2013.

DIAGNOSIS AND DEFINITION

The fifth edition of the American Psychiatric Association's (APA) *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5), published in May 2013, describes ADHD as a pattern of inattentive and/or hyperactive–impulsive behavior inconsistent with developmental level which interferes with functioning in social, educational, or work settings. Symptoms are divided into two categories: inattention and hyperactivity and impulsivity. There are five main diagnostic criteria: (1) onset before age 12 years old; (2) duration greater than 6 months; (3) children must have at least six symptoms from the inattention and/or hyperactive/impulsive symptom list, while older adolescents and adults must have at least five; (4) several symptoms must be present in two or more settings and interfere with functioning; and (5) symptoms that do not occur exclusively during the course of schizophrenia or other psychotic disorder and are not better accounted for by another mental disorder, such as depression. DSM-5 has no exclusion criteria for those with Autism Spectrum Disorder (ASD), allowing both Autism and ADHD to be diagnosed simultaneously in the same patient. Possible specifiers for the disorder in DSM-5 include combined presentation (criterion met for both inattention and hyperactivity–impulsivity), predominantly inattentive presentation or predominantly hyperactive/impulsive presentation. Other specifiers describe current severity of the disorder being mild (few to no symptoms in excess of those required), moderate or severe (many symptoms in excess of those required), or to specify if the disorder is in partial remission (fewer than full criteria met in past 6 months, but still result in impairment).

The inattention symptom criteria list includes: (a) Often fails to give close attention to details or makes careless mistakes in schoolwork, at work, or during other activities; (b) Often has difficulty sustaining attention in tasks or play activities; (c) Often does not seem to listen when spoken to directly; (d) Often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace; (e) Often has

difficulty organizing tasks and activities; (f) Often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort; (g) Often loses things necessary for tasks or activities; (h) Often easily distracted by extraneous stimuli (or unrelated thoughts for adolescents and adults); and (i) Often forgetful in daily activities.

Hyperactivity and impulsivity symptoms criteria include: (a) Often fidgets with or taps hands or feet or squirms in seat; (b) Often leaves seat in situations when remaining seated is expected; (c) Often runs about or climbs in situations where it is inappropriate; (d) Often unable to play or engage in leisure activities quietly; (e) Often "on the go," acting as if "driven by a motor"; (f) Often talks excessively; (g) Often blurts out an answer before a question has been completed; (h) Often has difficulty waiting his or her turn; and (i) Often interrupts or intrudes on others.

In the DSM-5 enhanced item descriptors were added to diagnostic criteria to demonstrate symptoms across the lifespan. For example, motor hyperactivity appears far less frequently in affected adolescents and adults, who continue to have attentional and organizational problems, poor anger management, job instability, and problems with social relationships and self-esteem. All 18 of the DSM-IV-TR symptoms of ADHD were retained in the DSM-5, but enhanced elaborators of each were added to better describe difficulties experienced by adolescents and adults. For instance, the inattention criteria "Often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort" now includes specific examples of this symptom "(preparing reports, completing forms, reviewing lengthy papers)" to be applied to older adolescents and adults in addition to "(schoolwork or homework)" which is primarily specific to children.

Concerns about Changes in DSM-5

DSM-5 addresses limitations in the prior edition. There is a concern that the decision in DSM-5 to increase the age of onset allowed for diagnosis from 7 years old to 12 years old may inflate the prevalence of positive diagnoses in adults. The onset criterion of age 7 used in three prior editions of the DSM lacked empirical validation. The aim of the 5-year extension in upper age limit for diagnosis was to improve diagnosis of adults, while still maintaining ADHD as a neurodevelopmental disorder with childhood onset. However, there is concern that the prevalence of ADHD will increase as a result, especially in older children, adolescents, and adults (McKeown et al., 2015). The prevalence rate may also increase in individuals with inattentive subtype, as inattention has been found to have a later onset. A recent study by Matte et al. has suggested a 27 percent increase in the expected prevalence of ADHD among young adults when comparing DSM-IV to DSM-5 criteria while Polanczyk et al. found that the age adjustment to 12 would have minimal impact.

COMPARATIVE NOSOLOGY

Clinicians in Europe use a very different diagnostic description of ADHD than that in the DSM-5: The ICD-10 diagnosis of "hyperkinetic disorder (HD)." The differences include the extent to which medication should be used and the algorithm used to diagnose the disorder. The DSM-5 diagnosis of ADHD and the ICD-10 diagnosis of HD are based on the same 18 symptoms, but the two systems use different decision rules. Fewer children are classified with ADHD using ICD-10 than in DSM-IV. For example, investigators in the United Kingdom applied an algorithm to the 579 children with DSM-IV ADHD Combined Subtype who had been recruited into the National Institute of Mental Health (NIMH) Multimodal Treatment Study of Attention-Deficit/Hyperactivity Disorder (MTA trial) in order to generate an ICD-10 diagnosis of HD. Only 25 percent of the MTA sample continued to meet the diagnostic criteria for HD. This suggests that cultural differences in referral practices and clinician preferences play a role in determining international differences in diagnoses of ADHD and HD.

ETIOLOGY

ADHD is a complex disorder of multifactorial etiology, thought to have multiple genetic and environmental variables of small effect acting together to produce vulnerability to ADHD. Although the etiology of ADHD yet has to be determined, there is a growing consensus that the condition involves functional and anatomical dysfunction in the brain's cortico-basal ganglia-thalamo-cortical circuitry. These areas support the regulation of attentional resources, the programming of complex motor behaviors, and the learning of responses to reinforcement. Theories involving these areas have been examined in series involving neurobiological studies of healthy humans, humans with ADHD, and animal models.

Genetics

Twin, sibling, adoption, and family studies all suggest a strong genetic component in the development of hyperactivity, inattention, and impulsivity, with heritability index estimated to range from 0.6 to 0.98. Heritability estimates include both genetic influences, as well as the effects of gene–environment interaction. ADHD-like behaviors have been reported in rare genetic disorders such as fragile X syndrome, neurofibromatosis type 1, DiGeorge syndrome, tuberous sclerosis, Turner syndrome, Williams syndrome, Angelman and Prader-Willi.

Twin Studies. Twin studies suggest that 76 percent of the variance in the transmission of ADHD is attributable to genetics, making ADHD one of the most heritable psychiatric disorders. Monozygotic twins have been

found to be more concordant for ADHD symptoms of hyperactivity, inattention, and impulsivity than are same-sex dizygotic twins. The concordance rate among monozygotic twins ranges from 59 to 92 percent, whereas the concordance rate in dizygotic twins ranges from 29 to 42 percent.

Sibling and Half-Sibling Studies. MRI studies have found reduced volumes in right prefrontal gray matter and left occipital gray and white matter in both ADHD subjects and their unaffected siblings. An early study compared the incidence of ADHD symptoms in 29 full-sibling and 22 half-sibling pairs for evidence of minimal brain dysfunction. Each pair had been raised together by a common mother. More than half (10) of the 19 full-sib pairs were concordant for ADHD, compared with only 2 of the 22 half-sib pairs, a significant difference, supporting the genetic influence on the transmission of ADHD in families.

Adoption Studies. Studies have found that the biological relatives of hyperactive children are more likely to be hyperactive themselves. This relationship is not seen in adoptive relatives of hyperactive children. Adoptive relatives of children with ADHD have risk of ADHD similar to relatives of control children, while the biological relatives of children with ADHD are more likely to have ADHD or associated disorders. In a study of international adoptees aged 10 to 15 years, Van den Oord and colleagues estimated that genes accounted for 47 percent of the variance of inattention scores on the Child Behavior Checklist. Eighteen percent of biological relatives of adopted-away children with ADHD were found to be at risk for the disorder versus 6 percent of the child's adopted relatives.

Family Studies. Family studies have shown that first-degree relatives of children with ADHD are two to eight times more likely to have the disorder. First-degree relatives of children with ADHD have a 20 to 25 percent risk for ADHD, compared with 4 to 5 percent for relatives of controls. If a parent has ADHD, 50 percent of his or her offspring are likely to have that condition. The increased incidence of ADHD is also extended to second-degree relatives.

Mode of Inheritance. Several hypotheses of ADHD's mode of transmission have been advanced. Omenn explored the possibility of sex-linked transmission because of the male preponderance in the condition. Other theories include a polygenetic inheritance model, including those of Morrison and Stewart, but limitations in sample size made this hard to prove. Rhee and colleagues postulated a polygenetic multiple threshold model after analyzing sex differences in an Australian twin and sibling-pair study. However, differences in fit between genetic models were modest, as

is true for comparisons of multifactorial and single-gene inheritance. This led to the suggestion that symptoms of ADHD may be caused by several interacting genes of modest effect. This hypothesis is consistent with ADHD's high population prevalence and high concordance in monozygotic twins, but modest recurrence risk for first-degree relatives. However, diagnostic uncertainty impedes progress in developing genetic models of inheritance.

Molecular Genetic Studies. There has been increasing interest in attempting to identify the specific "risk" genes that may be implicated in the development of ADHD. This search has yielded generally inconsistent results due to complexity of the trait and heterogeneity across studies. Candidate genes have been chosen for study based on theoretical considerations related to their function and the majority of studies have concentrated on dopamine-related genes given the implication of dopamine dysfunction in ADHD. Candidate genes discovered so far are responsible for less than 5 percent of genetic variation in ADHD. Genome-wide association studies (GWAS) of ADHD have not yet revealed any genome-wide significant association to common variation, indicating that the effect sizes of risk variants are likely very small.

The dopamine transporter gene (DAT1) codes for a sodium-dependent dopamine transporter protein that regulates reuptake of dopamine from the synaptic cleft. Cook et al. were the first to show an association between ADHD and DAT1 with 480–base-pair allele. Several studies have indicated that DAT1 might be more strongly associated with the hyperactive/impulsive or combined subtypes of ADHD. Another early gene studied was the dopamine type D2 receptor gene (DRD2). The gene is not specific to ADHD and is not believed to be a primary cause of the disorder. The dopamine 4 receptor seven-repeat allele gene (DRD4) has been linked to novelty seeking and connected to the impulsivity seen in ADHD. DRD4 has also been related to the inattentive subtype of ADHD. The synaptosomal-associated protein 25 gene (SNAP25) codes for a protein involved in neurotransmitter regulation and axonal growth. McGough et al. found an association between SNAP25 and sensitivity to methylphenidate (MPH) adverse events in preschoolers with ADHD. Genes that code for dopamine β-hydroxylase (DBH), catechol-O-methyltransferase (COMT), dopamine 5 receptor gene, norepinephrine transporter gene (SLC6A2), serotonin transporter gene (SLC6A4), α-2A adrenergic receptor (ADRA2A), thyroid receptor B gene, androgen receptors and factors in immune function and regulation have also been reported to correlate with ADHD symptoms. Subsequent studies and meta-analysis have not sufficiently supported the association between ADHD and these genes.

Neuroanatomical Aspects

Investigators have postulated the presence of an anterior attention network for promoting multiple aspects of attention, including focusing, executing, sustaining, and shifting functions. The attention networks are widespread through the cortex of the brain and are closely related to the default mode network described by Castellanos and others. MRI studies suggest that there are structural abnormalities in the brains of children with ADHD, including differences in the frontostriatal areas, temporoparietal lobes, cerebellum, basal ganglia, corpus callosum, amygdala, hippocampus, and thalamus. A meta-analysis of voxel-based morphometric (VBM) studies by Nakao and colleagues of children and adults with ADHD found global reductions in gray matter volumes, especially the right lentiform nucleus extending to the caudate nucleus, with slightly larger gray matter volumes in the left posterior cingulate cortex. These areas were demonstrated to normalize with age as well as with stimulant medication treatment. Developmentally, children with ADHD have a 3- to 5-year delay in cortical thickness maturation, particularly in the frontal and temporal brain regions. Diffusion tensor imaging (DTI) studies show white matter tract abnormalities and dysfunctional connectivity in ADHD during rest and while engaging in cognitive tasks.

Deficits in neural activity within frontostriatal and frontoparietal circuits are consistent findings in ADHD. fMRI studies using cognitive tasks to target the neural circuits show significantly reduced frontal lobe activity in subjects with ADHD compared to controls. Meta-analysis shows hypoactivation in the frontoparietal and ventral attention networks in children with ADHD, areas which control goal-directed executive processes and decision making. Hart and colleagues found that right dorsolateral prefrontal cortex activation was reduced in medication-naive patients, but normal in long-term stimulant-medicated patients. Electroencephalography (EEG) studies have reported an elevated theta/beta ratio (TBR) in subjects with ADHD versus controls, however, increased theta band activity has also been associated with substance use and bipolar disorder. Recent meta-analysis determined that excess theta and elevated TBR was not a reliable diagnostic measure for ADHD, but could have prognostic value for treatment outcome.

Neurotransmitters in ADHD

The nonspecific catecholamine hypothesis of ADHD posits that the psychophysiology of ADHD symptomatology emerges from an imbalance among various neurotransmitters, including norepinephrine, epinephrine, and dopamine. Human studies investigating the catecholamine hypothesis of ADHD produced conflicting results, possibly due to the difficulty of measuring brain levels of these neurotransmitters in peripherally obtained samples.

Dysfunction in brain norepinephrine systems, particularly a lack of

inhibition of locus coeruleus neurons, could explain the inattention, cognitive deficits, and higher levels of gross motor activity seen in children with ADHD. Randomized clinical trials (RCTs) have suggested that the norepinephrine reuptake inhibitors (NRIs), tricyclic antidepressants (TCAs) and atomoxetine, are effective in reducing ADHD symptoms by restoring a more normal ratio of epinephrine and norepinephrine. Molecular genetic studies have targeted genes that code for dopamine receptors and PET scans in adults with ADHD have provided imaging data that correlate stimulant-driven increases in dopamine concentration. There is weak evidence for serotonin's role in ADHD. Medications that exclusively affect serotonin function, such as the selective serotonin reuptake inhibitors (SSRIs), have no efficacy in the treatment of children with ADHD. For that reason, serotonin is thought to play a secondary role in the pathology of ADHD.

Environmental Factors

ADHD has been associated with numerous environmental factors, although it is difficult to conclusively determine causality. In utero risk factors include prenatal exposure to alcohol, nicotine, recreational drugs, glucocorticoids, the toxins hexachlorobenzene and polychlorinated biphenyls (PCBs), poor maternal diet and maternal stress during pregnancy. Perinatal ADHD risks include prematurity, low birth weight, and obstetric complications. ADHD has been associated with childhood illnesses such as meningitis, encephalitis, cardiac disease, thyroid disease, epilepsy, head trauma, and autoimmune and metabolic disorders. High blood lead levels have also been linked with ADHD diagnosis, however, this is likely not a major factor in the disorder. Psychosocial adversity including poverty, childhood maltreatment, negative parenting, family discord, and bullying have all been associated with ADHD. It has been difficult for investigators working with children affected by adversity to determine whether their ADHD symptoms reflect a response to negative parenting, a harsh environment, a genetically influenced biological problem, or some interaction among these factors.

Dietary Influences. Dietary factors as a cause and possible treatment of ADHD are controversial. There have been reports of a connection between ADHD and nutritional deficiencies of iron, zinc, and essential fatty acids (omega-3 and omega-6), but the beneficial effects of supplementation are inconclusive. Some studies have indicated that a small subset of children may have sensitivity to certain foods, additives, excess sugar or nutritional deficiencies, causing behavioral effects. There has been some support for elimination diets to treat ADHD, such as the Feingold (additive and salicylate-free) diet, oligoantigenic (hypoallergenic) diet, and ketogenic diet. In 2011, the U.S. Food and Drug Administration determined that

existing data do not support a causal link between additives and behavioral disturbance in children.

DIAGNOSIS AND CLINICAL FEATURES

The diagnosis of ADHD requires the integration of data from multiple sources, including clinical interview, direct observation, behavioral rating scales, and the subjective observations of parents and teachers. Most children with ADHD are referred for care because of impairment in academic, family, and/or peer relationship functioning. Symptoms of impulsivity, overactivity, and inattention drive this impairment across the lifespan.

Clinical Features

DSM-5 Diagnostic Criteria. DSM-5 diagnostic criteria for ADHD defines three types of presentations based on two criteria lists of nine symptoms of inattention and nine symptoms of hyperactivity/impulsivity. Children with six or more symptoms of inattention and fewer than six symptoms of hyperactivity–impulsivity are diagnosed with ADHD predominantly inattentive presentation, while those with six or more symptoms of hyperactivity/impulsivity and fewer than six symptoms of inattention meet criteria for ADHD predominantly hyperactive/impulsive presentation. ADHD combined presentation is diagnosed if full criteria from both lists are met. Older adolescents and adults require only five symptoms from either (or both) of the two criteria lists. For formal diagnosis of ADHD it is necessary for symptoms be present prior to age 12, persist for at least 6 months, have a negative impact on social, academic, or occupational activities and be present in two or more settings. The diagnosis is not appropriate if the symptoms occur exclusively during the course of another primary psychotic, mood, anxiety, or substance or personality disorder, but ADHD may coexist with pervasive developmental disorder. DSM-5 has included expanded item descriptors of symptoms to better represent the experience of adolescents and adults of ADHD across the lifespan. There are specifiers available to indicate severity of the disorder as mild, moderate, or severe.

Hyperactivity. Hyperactivity describes the inappropriate excessive motor activity observed in ADHD, including restlessness, fidgeting, and appearing to be “driven by a motor.” These behaviors are often first noticed when the child is a toddler, but can be normative before age 4. The level of gross motor activity usually decreases with age, however, the inner sense of restlessness may continue into young adult life. Children with ADHD display sustained levels of high activity in different settings such as home, school, and playgroup. Activity increases occur during sleep as well.

Hyperactivity may be absent in novel situations or when child is engaged in an activity they find particularly engrossing such as playing videogames.

Impulsivity. Impulsive actions occur without forethought about consequences and may be associated with desire for immediate rewards or to avoid delay. Impulsivity might be seen when the child engages in dangerous activities, yells out in class, or interrupts or intrudes on others during games or conversations. Impulsive behavior might also result in trouble with parents, teachers, or other children, including verbal or physical fights. Children with ADHD can demonstrate overly quick and error-prone performance on standardized tasks. Symptoms of impulsivity can persist into adulthood, even after hyperactive symptoms have diminished.

Attentional Difficulties. Inattention includes difficulty sustaining focus, trouble maintaining organization, and being easily distracted by extraneous stimuli. The symptoms of inattention often become prominent in elementary school when the child is approximately 8 to 9 years old and symptoms can be lifelong. Attentional difficulties are most often seen in routine settings in which the youth with ADHD must sit and carry out tasks that involve repetition under conditions of low levels of reinforcement and external motivation. Sluggish Cognitive Tempo (SCT) is a type of attentional difficulty that has been closely associated with the inattentive subtype of ADHD. SCT describes a constellation of symptoms such as daydreaming, staring, tendency toward confusion, mental fogginess, sleepiness, apathy, and physical hypoactivity. There is growing evidence suggesting that SCT is a condition distinct from ADHD, however, SCT is currently not considered a separate disorder or diagnostic subtype.

Associated Factors. Children with ADHD might have areas of impairment that are not listed under the DSM-5 symptom criteria covered exactly by the 18 symptom exemplars of hyperactivity, impulsivity, or inattention. Children and adolescents with ADHD often show difficulty with time management and do not develop an internal sense of pace in planning tasks. This poor sense of time leads to problems in estimating the actual difficulty of waiting in line, planning how much time a task requires, or even knowing when to come home when out playing with other children. Children with ADHD often lack persistence and can become bored playing interactive games with peers. They also find it difficult to delay gratification.

Neuropsychological testing has found that individuals with ADHD demonstrate impaired executive functioning including abnormal response inhibition. On any given measure of executive function, less than half of children with ADHD have been found to be impaired. Executive

functioning problems also occur in other psychiatric disorders of childhood, such as depression, and are not specific to ADHD. It has been shown that academic functioning is more strongly affected by an impulsive need to get through tests quickly, a deficit closely linked to poor behavioral inhibition, rather than poor executive functioning.

ADHD is often associated with problems in emotion regulation, resulting in temper outbursts, mood lability, and reactivity. Moods can change dramatically with no obvious connection with what is going on in the environment. Social skills are often significantly impaired in children with ADHD compared to controls. Individuals with ADHD may have problems accurately interpreting nonverbal social cues and have trouble cooperating with other children and following rules in games. Peers identify these individuals as intrusive, bossy, and insensitive to the needs of others. Children with ADHD often overreact and respond to frustration in social situations which can lead to verbal or physical aggression. This is a strong stimulus for peer rejection, which has been shown to be a reliable long-term negative predictor of development, particularly in adolescence.

MEDICAL AND LABORATORY EXAMINATION

Children with ADHD should be given a medical evaluation, including a thorough history and physical examination, to rule out physical disorders that can mimic the symptoms of ADHD. The medical history should cover the prenatal, perinatal, toddler, and preschool phases of development. Inquiries should be made about pregnancy complications, such as maternal illness (eclampsia, diabetes), maternal smoking, alcohol, or illicit drug use. The perinatal period should identify the presence of labor problems, delivery complications, prematurity, jaundice, and low birth weight. Feeding and sleeping routines during the perinatal and postnatal periods need to be described. Developmental milestones, illnesses, injuries, and symptomatology need to be reviewed. A detailed family, medical, and psychiatric history should be completed.

The medical examination can identify physical problems (such as arrhythmia, tachycardia, or hypertension) that can increase the risk of serious adverse events should a stimulant medication treatment be initiated. Height and weight at baseline should be recorded. Physical examination of a child with ADHD may reveal nonlocalizing neurological signs, such as poor coordination, dysrhythmias, and overflow movements. Medical problems that might aid in the differential diagnosis, such as petit mal epilepsy, hearing and vision difficulty, thyroid dysfunction, and hypoglycemia, all need to be ruled out. There are no blood, genetic, or neuroimaging tests available to diagnose ADHD.

RATING SCALES

Many standardized behavioral rating scales for ADHD have been

developed. Rating scales should not be used independently to diagnose ADHD, but can be particularly helpful in establishing the presence of core ADHD symptoms in more than one setting. ADHD-specific scales have been found to have greater than 90 percent sensitivity and specificity if used in the correct population. Commonly used rating scales for ADHD include Conners Rating Scales for parents (CPRS-R) and teachers (CTRS-R), Conners Wells Adolescent Self-Report Scale, Vanderbilt ADHD Diagnostic Parent and Teacher Scales, Swanson, Nolan and Pelham (SNAP-IV), and Brown ADD Rating Scales for Children, Adolescents, and Adults. Only the Conners Comprehensive Behavior Rating Scales and the ADHD Rating Scale IV have been validated in preschool-aged children.

DIFFERENTIAL DIAGNOSIS

A clinician must make a differential diagnosis when judging whether a patient meets criteria for ADHD. The distinction between differential diagnosis and comorbidity is not well defined in the DSM-5. If the impairing condition is best accounted for by another disorder, that other disorder rules out ADHD. For children with intellectual disability, inattention may be a natural accompaniment of the intellectual limitations, so the diagnosis of ADHD should only be applied if that symptom is excessive for the mental age of the child. Children from chaotic, disorganized, or inadequate home environments may also show symptoms of overactivity or inattention. The symptoms of children with oppositional defiant behavior must be distinguished from the reluctance and resistance that children with ADHD exhibit when they avoid effortful mental tasks. Symptoms of impulsivity, gross motor hyperactivity, or inattention may appear during the use of medications, such as the akathisia that accompanies second-generation antipsychotics or from bronchodilators. Such children should not receive a diagnosis of ADHD, but one of Other Substance-Related Disorder.

COURSE AND PROGNOSIS

Parents often notice very high levels of gross motor activity when the child with ADHD is a toddler, just when the child has learned to walk independent of the parent's help. However, the energy, oppositionality, and curiosity of toddlers can be confused with the excessive, almost random motion of older children with ADHD, so that one must be cautious when applying the ADHD diagnosis to a preschooler. Usually, the ADHD diagnosis is first applied in primary school, during grades 1 to 6, when adjustment to the sedentary learning style is compromised. Thus the school-aged child often experiences academic failure and peer rejection. Behavior problems at school and at home are not infrequent with oppositional defiant disorder (ODD) being a common comorbidity.

In childhood 65 to 75 percent of children with ADHD have one or more

comorbid conditions, which in addition to ODD, can include conduct disorder (CD), anxiety and mood disorders, tics or Tourette syndrome, learning disorders, and autistic spectrum disorders. Many of these comorbidities continue into adolescence.

In adolescents, the external overactivity lessens but the internal restlessness does not. The adolescent with ADHD continues to have academic and social problems, which erode his/her self-esteem and at times promote the gravitation to a negative peer group. Thus, adolescents with ADHD who are untreated, compared to peers with no mental disorder, have a threefold increase in substance use and abuse, more frequent delinquent acts and trouble with the law, and increased rates of car accidents when they begin to drive. Adolescents with ADHD also start sexual intercourse earlier with a higher number of sexual partners and greater frequency of unprotected sex with corresponding higher rates of sexually transmitted diseases and unwanted pregnancy. The rate of depression and anxiety also increases in adolescence because of the many functional problems.

Approximately 60 percent of children with ADHD continue to be impaired well into adult life with prevalence estimates suggesting that 4 percent of adults may suffer from ADHD. In adulthood, hyperactivity may be expressed by inner feelings of restlessness, inability to relax and excess talkativeness, fidgeting, and inability to remain seated for long periods of time, for example, meetings, symposia, church. Impulsivity may be expressed in impatience, interrupting others, acting without thinking, ending or starting ventures on impulse, for example, jobs, relationships, and at times sensation-seeking behavior.

Inattention is seen in distractibility, disorganization, chronic lateness, or missing appointments. The distraction and disorganization often results in careless mistakes and forgetfulness. Procrastination is another important problem. Even when things are started, they are often not finished and so deadlines are thus missed.

The inattention and distractibility also make it appear that the adult with ADHD is not interested and not listening to what is being said. This can strain relationships at home and at work. Some adults with ADHD also suffer from labile moods and “short fuses.” These characteristics coupled with the problems described above result in more frequent losses of jobs and relationships. The consequences of this are lower job status and lower income and a sense of underachievement or failure on the part of the adult with ADHD. The inattentiveness and impulsivity also increases the likelihood of driving accidents and infractions, other types of accidents are also more prevalent. Generally, adults with ADHD may have an unhealthy lifestyle, with smoking, alcohol, and drug abuse risky sexual behavior, and chronic sleep problems frequently seen.

In adulthood, comorbidity is the rule, with more than 75 percent of

adults with ADHD having at least one comorbid condition. Anxiety, depression, substance abuse, and some personality disorders are the most common comorbidities. Gambling and other addictions are also frequently seen.

It should be noted that even though the male/female ratio in children with ADHD can be 3–4:1, in adulthood it is 1:1. It is possible that many girls who tend to be less disruptive are not identified or referred in childhood but have significant impairment as adults and seek help then.

TREATMENT OF ADHD

The efficacy of both medication and psychosocial interventions for the treatment of ADHD has been well established. The preferences of families and availability of services largely guide the choice of treatment.

Stimulants

Stimulants are considered first-line agents to treat ADHD symptoms based on decades of proven safety and efficacy. Therapeutic effects of stimulants include a decrease of hyperactivity, impulsivity, and inattention, as well as improvement in associated behaviors. The two groups of stimulant medication that have received U.S. Food and Drug Administration (FDA) approval for the treatment of youth with ADHD are amphetamines and MPHs. Many stimulants are available in both generic and branded versions. Stimulants are classified with drugs that can be abused or cause dependency under the category of schedule II controlled substances by the United States Drug Enforcement Agency (DEA) (Table 45.1–1).

Pharmacology and Pharmacokinetics. Both MPH and amphetamine can be described as psychostimulants, which refers to their ability to increase central nervous system activity in brain regions and increase blood pressure and pulse in the periphery. Although the exact mechanism of action is not completely understood, stimulants increase the concentration of dopamine and norepinephrine in the synaptic cleft. MPH appears to act primarily as a dopamine–norepinephrine reuptake inhibitor, while amphetamines block the reuptake of both norepinephrine and dopamine as well as facilitate dopamine release through reverse transport. Berridge et al. showed—during in vivo microdialysis in rodents—that low-dose MPH substantially increased norepinephrine and dopamine reflux within the prefrontal cortex (PFC) while enhancing PFC-dependent cognitive functioning.

Table 45.1–1.
Stimulant Drugs, Doses, and Pharmacodynamics

| | | |

Medication	Forms	Pediatric Dose, Start/Typical	Effect (hours)	Duration of Action (hours)
Dexmethylphenidate				
Short-acting	2.5-, 5-, 10-mg tabs	2.5 mg AM/10 mg b.i.d.	1–4	5–6
Long-acting	5-, 10-, 15-, 20-, 25-, 30-, 35-, 40-mg caps[a]	5 mg AM/10 mg AM	1–4	12
Dextroamphetamine				
Short-acting	2.5-, 5-, 7.5-, 10-, 15-, 20-, 30-mg tabs, 5 mg/5 mL[b] oral solution	5 mg b.i.d./10 mg b.i.d.	1–3	4–6
Long-acting	5-, 10-, 15-mg caps[c]	5 mg AM/15 mg AM	1–4	6–8
	20-, 30-, 40-, 50-, 60-, 70-mg caps[d]	30 mg AM/70 mg AM	2–3	up to 12
Methylphenidate				
Short-acting	5-, 10-, 20-mg tabs[e]	5 mg b.i.d./10 mg t.i.d.	1–3	3–5
	2.5-, 5-, 10-mg chewable tablets[e]	5 mg b.i.d./10 mg t.i.d.	1–2	3–5
	5 mg/5 mL, 10 mg/5 mL[e] oral solution	5 mg b.i.d./10 mg t.i.d.	3–3	3–5
Immediate-acting	10-, 20-mg[f]	10 mg b.i.d./30 mg AM	3	3–8
Long-acting	10-, 15-, 20-, 30-, 40-, 50-, 60-mg tabs/caps[g]	10 mg AM/30 mg AM	5	8–12
	18-, 27-, 36-, 54-mg tabs[h]	18 mg AM/36 mg AM	8	10–12
	10-, 15-, 20-, 30-mg transdermal patch[i]	10-mg patch on 9 hours, off 15 hours 30-mg patch on 9 hours, off 15 hours	7–9	10–12
	5 mg/5 mL oral solution[j]	20 mg AM	5	10–12
	20-, 30-, 40-mg chewable tabs			up to 13
Mixed salts of amphetamine				
Short-acting	5-, 7.5-, 10-, 12.5-, 15-, 20-, 30-mg tabs[k]	5 mg b.i.d./10 mg b.i.d.	1–3	4–6
Long-acting	5-, 10-, 15-, 20-, 25-, 30-mg caps[k]	5 mg AM/30 mg AM	1–4	8–10
	3.1-, 6.3-, 9.4-, 12.5-, 15.7-, 18.8-mg orally disintegrating tabs[l]	6.3 mg AM/18.8 mg AM	5	8–10
	2.5 mg/mL oral suspension[m]	2.5 mg AM/20 mg AM	3.9	up to 13

b.i.d., twice daily; t.i.d., thrice daily.

[a]Focalin, Focalin XR; [b]Dexedrine, Zenzedi, ProCentra solution; [c]Dexedrine Spansule; [d]Vyvanse (Lisdexamfetamine); [e]Ritalin, Methylin chewable, Methylin oral suspension; [f]Metadate ER; [g]Metadate CD, Ritalin LA, Aptensio XR; [h]Concerta; [i]Daytrana. The four patch sizes deliver 10-, 15-, 20-, and 30- mg over 9 hours, respectively, supplied in sealed trays containing 10 to 30 patches in individual pouches.; [j]Quillichew ER, Quillivant XR oral solution; [k]Adderall and Adderall XR; [l]Adzenys XR-ODT; [m]Dyanavel XR.

FORMULATIONS

The racemic form of MPH is d,l-threo-methylphenidate, with d-threo-methylphenidate (d-MPH) being the more active isomer. Both racemic MPH (including Ritalin, Metadate, and Methylin) and the monoisomer d-MPH (Focalin) are commercial products. MPH is available in immediate-release tablet, chewable and liquid formulations.

Extended-release MPH is available in many different preparations. It is manufactured using beaded technology, consisting of both immediate- and delayed-release beads which are equivalent to two doses of immediate-release medication given 4 hours apart. The beads can be sprinkled on food for children who have difficulty with swallowing pills. MPH is also available as a transdermal patch (Daytrana), which has onset of action of 60 minutes and can be worn up to 9 hours per day. MPH utilizing an osmotic-release oral system (OROS-MPH or Concerta) releases the stimulant via an osmotic pump in an ascending dose curve over a 10- to 12-hour period. This allows for a single-dose product that is consistent with the effect of a three times per day dose of immediate-release MPH hydrochloride. The newly available MPH hydrochloride extended-release liquid suspension (Quillivant XR) has demonstrated efficacy in reducing ADHD symptoms for up to 12 hours, is well tolerated and is especially useful in children who have difficulty swallowing pills.

Amphetamine is manufactured as immediate-release formulations, including mixed amphetamine salts, which is a mixture of d- and l-amphetamine, or in the dextro isomer. With regard to just the dextro isomers, dextroamphetamine (d-AMP) is 1.5 times more potent, milligram for milligram, than is d-MPH. Another popular product, mixed salts of amphetamine (MAS), is marketed as Adderall, and combines four different salts of amphetamine, including both d-AMP and l-AMP. One small study showed that there are no efficacy differences between d-AMP and MAS.

AMP is also available in extended-release beaded formulations, as the isomer d-amphetamine sulfate (Zenzedi) and as a liquid (Procentra). Lisdexamfetamine dimesylate (Vyvanse) is a therapeutically inactive prodrug that is converted by enzymatic hydrolysis to inactive l-lysine and active d-amphetamine predominantly by red blood cells, with a smaller amount converted in the gastrointestinal tract. This formulation was developed to decrease the potential for abuse, as lisdexamfetamine used intranasally or intravenously has similar effects to oral administration.

Although absorption of MPH products is generally not affected by foods, some long-duration preparations of d-AMP show decreased concentrations when taken with high-fat meals. Whereas MPH is metabolized in the gut wall and plasma by esterases, the metabolism of d-AMP and MAS takes place in the liver. Variants in the humans CES-1 gene coding for the hydrolyzing enzyme have led to interindividual variations in MPH metabolism and plasma level.

Amphetamine product package inserts approved by the FDA warn that urinary excretion of AMP metabolites can be increased by foods of acidic content (vitamin C and orange juice), thus lowering plasma concentration. Conversely, foods that shift urinary pH higher reduce excretion and thus increase the plasma concentration of d-AMP or MAS.

Efficacy. Efficacy studies of stimulants consistently show high numbers of children who are responders (Storebø et al., 2015). The MTA study found that 77 percent responded to MPH, and an additional 10 percent of MPH nonresponders responded to d-AMP. Teacher rating scales showed significant reductions in ADHD symptoms for stimulants compared with placebo, with effect size ranging from 0.8 to 1.0 standard deviation (SD) on behavioral measures and 0.5 SD for cognitive measures for those children randomized to active drug. Placebo response rates in controlled studies have tended to be low, generally less than 20 percent of those randomized to that condition. The Preschool ADHD Treatment Study (PATS) determined that preschool-aged children require a lower optimal doses of IR methylphenidate, experienced much higher rates of adverse events, do not respond as robustly showing a lower effect size, and had slower clearance of the drug than school-aged children (Vitiello et al., 2015).

Amphetamines and MPHs both share the ability to reduce gross motor overactivity, increase sustained attention on tasks, and reduce impulsive responses on laboratory measures. These characteristics have been shown in more than 200 short-duration, double-blind, placebo-controlled clinical trials involving patients with ADHD across a wide age range, including preschoolers, school-aged children, adolescents, and adults. Some individuals respond preferentially to MPH, whereas others respond preferentially to the amphetamines. The majority of evidence indicates that MPH and amphetamine are equally efficacious.

Cognitive and Behavioral Effects. Greater effects due to stimulants are found to occur in behavioral domains rather than in cognitive areas. Psychostimulant medications have been shown to reduce gross motor overactivity, out-of-seat behavior, calling-out in the classroom, disruptiveness, impulsive behavior, and noncompliance while improving self-perception, academic productivity, peer nominations of the child and mother–child interaction. Stimulants have been shown to improve vigilance and reaction time and reduce variability, short-term memory, and learning of verbal and nonverbal material in children with ADHD. The MTA Titration Trial showed a linear dose–response curve for behavioral measures between 0.5 and 1.0 mg/kg/day. Although earlier studies showed no long-term academic benefit from stimulant treatment, recent controlled trials revealed improvements in school-based productivity and accuracy in

children with ADHD consistently treated with stimulants.

Adverse Effects

SHORT-TERM ADVERSE EFFECTS. Common side effects of the stimulant medications include decreased appetite, weight loss, delayed onset of sleep, headaches, stomachaches, and small increases in blood pressure and pulse. Most of the short-term side effects can often be dealt with by changing the time or dose of medication, or both.

Infrequent side effects include motor tics, the unmasking of Tourette syndrome, and late withdrawal effects, sometimes called rebound. Rebound presents the sudden return of more severe overactivity, impulsivity, and inattention in the late afternoon or evening than present at baseline off medication. Attempts to demonstrate their presence in full-day, laboratory classrooms have been unsuccessful, so these effects may be generated as the child interacts with task demands in his or her natural environment. Long-acting psychostimulant preparations or a small dose of immediate-release medication may be given before onset of the rebound to reduce this problem.

Stimulants rarely produce visual and tactile hallucinations, priapism, choreiform movements, and self-directed behavior, such as lip licking, nail biting, and rubbing of sores. These problems are first treated by discontinuing the medication until the adverse event disappears.

Two large recent studies in children and one in adults found that ADHD medications are not associated with an increased risk of severe cardiovascular events. General cardiac side effects of stimulant drugs are increase in heart rate of approximately 1 to 2 beats/minute and increase in systolic and diastolic blood pressure of 3 to 4 mm Hg. Stimulants have not been found to induce clinically significant changes in electrocardiographic parameters. The evidence suggests a potential for severe adverse cardiac events in some children with certain forms of congenital heart disease or arrhythmias. Routine electrocardiograms are not recommended prior to initiating stimulants in the absence of concerning physical examination findings or personal or family cardiac history.

Long-Term Adverse Events

EFFECTS ON GROWTH. The use of stimulants in children has been associated with growth suppression in some studies, but methodological issues prevent a clear conclusion to be drawn. Two large-scale NIMH-supported randomized, clinical trials have shown that MPH can reduce the rate of height and weight gains in developing children with ADHD. This suppression effect, which may be dose related in some children, has been more pronounced in those children treated with d-AMP and occurs predominately in the first year of treatment. The reductions in linear

growth have been minimal (approximately 1 cm in height and 2.5 kg in weight during the first year of treatment with stimulants). There is growing evidence of that this decrease in growth velocity diminishes after the first year treatment. The effects on growth are thought to be secondary to appetite suppression and reduction of caloric intake. Small, single-site studies are discordant concerning whether psychostimulants suppress human growth hormone (HGH) or effect prolactin levels. Some researchers have suggested that the growth deficit may be associated with the disorder itself, rather than stimulant medication. Clinicians are recommended to regularly monitor height and weight of children with ADHD treated chronically with psychostimulant medication. It is also encouraged to institute drug holidays during summer vacations to increase the probability of growth rebound.

SUBSTANCE ABUSE. There is no evidence to suggest that the use of stimulants in young people increases the risk for later substance use disorders (SUD). A meta-analytic study by Wilens and colleagues suggests that stimulant treatment in childhood has a protective effect, bringing the risk down to the level of the general population. Other studies have revealed that stimulant therapy neither increases nor decreases the risk for subsequent SUD in children and adolescents with ADHD. Clinicians should take into consideration if family members are currently abusing prescription medication before prescribing stimulants for ADHD.

Stimulant medication can be misused or diverted for the purpose of academic performance enhancement or euphoric effects, with immediate-release stimulants more often misused compared to extended-release formulations. A systematic review of 21 studies indicated prevalence of psychostimulant misuse in the year prior to the study ranged from 5 to 9 percent in grade and high school–aged children and 5 to 35 percent in college-aged individuals. Close monitoring, pills counts, or switch to nonstimulant medication should be employed by the clinician if misuse or diversion is suspected.

Summary of Stimulant Medication. Continuous treatment of children with ADHD with psychostimulants is both effective and safe over the first 2 years. Long-acting preparations provide the convenience of once-daily dosing, reduce stigma associated with taking medication at school and can decrease the potential for abuse. Short-acting forms are useful for younger children who often require smaller doses than is available in extended-release preparations. Many adverse events can be handled by adjustment of dose and/or the time of administration of the medication. Long-term adverse events such as growth rate suppression concern parents but may be managed by instituting drug holidays to permit growth rebound.

Nonstimulant Medication in the Treatment of Children with ADHD

Whereas psychostimulants constitute the first-line treatment for children with ADHD, other medications need to be used for nonresponders or those who experience moderate to severe adverse events. Some nonstimulant medications, such as atomoxetine and long-acting clonidine or guanfacine, have been approved by the FDA for the treatment of ADHD in children and adolescents. Other nonstimulant medications are used off-label but are less efficacious than the stimulants in treating ADHD symptoms. Advantages of nonstimulant medications include a longer duration of action and the absence of psychostimulant effects such as behavioral rebound or delay of sleep onset.

Atomoxetine HCl. Atomoxetine (ATX), an NRI marketed under the name Strattera, is an approved alternative to stimulants. It may be considered first line in individuals where parents have a personal objection to stimulants, or the child has comorbid anxiety symptoms or severe stimulant side effects. Atomoxetine is well absorbed after oral administration and is minimally affected by food. Maximum plasma concentrations are reached approximately 1 to 2 hours after ingestion and typical duration of action is 10 to 12 hours. Ninety-eight percent of circulating ATX is bound to plasma protein, mainly albumin. ATX is dosed by weight, with RCTs showing no benefit from exceeding total daily doses of 1.4 mg/kg. The time to therapeutic effect is in the range of 1 to 4 weeks, with full response at 6 to 12 weeks.

ATX has been shown to reduce ADHD behaviors in children, adolescents, and adults with a medium effect sizes of 0.5 to 0.7. Common adverse events associated with ATX involve abdominal discomfort, nausea, decrease in appetite, sedation, daytime sleepiness, headache, dizziness, vertigo, irritability, and mood swings. Daytime sleepiness and gastrointestinal tract problems may be diminished by titration to full dose over 1 week in twice-daily dosing. Minor increases in blood pressure and pulse have been reported and thus should be regularly monitored. Rare adverse events in youth include priapism, disturbances in sexual functioning, increased suicidality and liver toxicity.

ATX is metabolized by the cytochrome P450 (CYP) 2D6 hepatic enzyme system. A fraction of the population are poor metabolizers of CYP 2D6–metabolized drugs. Such individuals may experience elevated plasma concentrations and extended plasma half-life, therefore very slow initial titration may be required to prevent excessive dosing of these patients. Patients taking other medications metabolized by CYP 2D6, such as fluoxetine (Prozac), paroxetine (Paxil), and quinidine (Cardioquin), might show signs of competitive inhibition, namely increased plasma levels of both ATX and the competing drug.

α -Adrenergic Agonists. The α-adrenergic agonist agents were initially approved for the management of hypertension in adults. The long-acting α-adrenergic agents guanfacine ER (Intuniv) and clonidine ER (Kapvay) have been FDA approved as monotherapy and adjunctive treatment of ADHD in children 6 to 17 years old. The immediate-release clonidine (Catapres) and guanfacine (Tenex) have not been approved to treat ADHD, although they are commonly used off-label. The supposed mechanism of action of the α-adrenergic agents involves stimulating α (alpha) receptors in the PFC leading to improvement inattention, impulsivity, and hyperactivity. The α-adrenergic agonists are clearly less efficacious than the psychostimulants with several meta-analyses suggesting a moderate effect size. However, a small multisite, double-blind, RCT showed that children and adolescents comorbid for ADHD and Tourette syndrome benefited from clonidine alone or in combination with MPH, in both cases significantly better than placebo. Guanfacine has proven efficacy in reducing oppositional symptoms in children with comorbid oppositional behaviors and ADHD.

Common side effects of the α-adrenergic agents are sedation, fatigue, somnolence, headache, hypotension, syncope, and bradycardia. There have been rare reports of hallucinations in children treated with guanfacine. Clonidine is very sedating and has been used to treat sleep difficulties of children with ADHD, as well as to offset the delay in sleep onset caused by stimulant medications. It is initiated at low doses, 0.05 mg at bedtime. Guanfacine is a more selective α-adrenergic agonist and is slightly less sedating than clonidine. It is used in higher total daily doses than clonidine, between 1 and 3 mg. The α-adrenergic agonists need to be discontinued very slowly to prevent rebound adrenergic overdrive, with hypertension, agitation, fever, headache, tachycardia, chest pain, sleep disturbance, nausea, and vomiting. This can create a problem if a family goes on holiday and the parents forget the child's medication.

Tricyclic Antidepressants. TCAs have data supporting their use in ADHD, but are used infrequently as they have a much more serious side-effect profile. Spencer reviewed 29 studies of children and adults treated with TCAs (imipramine, desipramine, and nortriptyline) and found significant reductions in ADHD symptoms in 22 of the studies. There are many adverse events associated with TCAs, including cardiovascular adverse events, such as slowing of cardiac conduction and increasing the electrocardiographic PR and QRS intervals indicating increased risk of cardiac arrhythmias and heart block. There have been reported sudden deaths in patients on stable therapeutic doses of TCAs. For that reason, electrocardiograms at baseline and after each dose adjustment are necessary for routine monitoring. Other side effects commonly experienced with TCAs include lowering of seizure threshold and cholinergic side

effects, such as constipation, dry mouth, or blurred vision. When discontinued, TCAs should be tapered over several weeks.

Bupropion. Bupropion is an antidepressant with dopaminergic and noradrenergic properties that has shown modest efficacy in ADHD, but is less effective than stimulants or atomoxetine. It is available in extended-release preparations (Wellbutrin SR, Wellbutrin XL) that permit once-daily dosing. It might be more effective in adolescents and is not recommended for children. Starting dose for young adolescents is 75 mg twice a day to a maximum of 200 to 300 mg/day. The doses can be increased once weekly until improvement is noted. Bupropion usually takes 4 to 6 weeks to show effect.

Side effects of bupropion include fatigue, dry mouth, insomnia, headaches, nausea, vomiting, constipation, tremor, and skin rash. Seizures will reliably occur if the adolescent's dose exceeds 400 mg/day. This medication is FDA approved for the treatment of depression and smoking cessation, but not ADHD, in individuals older than 18 years of age. Bupropion is considered a second-line, off-label treatment for ADHD.

Other Alternative Medications. Anticonvulsants, including carbamazepine (Tegretol), lithium (Eskalith), SSRIs have not been shown to be effective for the treatment of ADHD. Second-generation antipsychotic medications such as risperidone have been shown to be effective for the treatment of aggression when combined with parent training and MPH, but not when used as the sole therapeutic agent. Modafinil (Provigil), an antinarcoleptic stimulant agent, was shown to be effective in three double-blind, placebo-controlled RCT, but the manufacturer withdrew its application to the FDA because of a severe Stevens–Johnson skin rash that might have occurred in one patient and reported visual hallucinations in registration trials. In summary, then, these medications might be useful in treating some comorbid conditions that children with ADHD often have.

Summary of Medication Treatments. Most children with ADHD respond to stimulant medication, so it continues to be the first-line treatment choice for reducing the symptoms of ADHD. However, if stimulants are ineffective, practice parameters first suggest turning to the serotonin–norepinephrine reuptake inhibitor atomoxetine. Alternatively, one can combine a stimulant medication with an α-2 agonist when the youth shows only partial response to a stimulant. If the child shows severe adverse events on stimulants, off-label medications can be tried. One must keep in mind that bupropion is less effective than either the TCAs or stimulants and carries a risk of seizures at high doses. Clonidine and guanfacine are sedating and have significant cardiovascular side effects. For all these drugs, careful monitoring and slow, cautious discontinuation

are essential. TCAs are not routinely used for the treatment of ADHD due to the numerous side effects and high toxicity.

Treatment of ADHD Comorbidities

Tic Disorder. Stimulants alone or in combination with clonidine, guanfacine, and atomoxetine appear to reduce ADHD symptoms in individuals with both ADHD and Tourette syndrome or tic disorders. Although stimulants have not been shown to worsen tics in most patients with tic disorders, they may exacerbate tics in some individual cases. In these instances, decreasing the dose of stimulants or treatment with α-agonists or atomoxetine may be an alternative. Cognitive behavioral therapy (CBT) with the use of opposing muscle contractions to address tics has also shown some efficacy and can be considered to address problems with tics.

Seizure Disorder. Children with seizure disorders may also suffer from ADHD. Although there are concerns about psychostimulants, particularly MPH, lowering the seizure threshold, experiencing a first seizure when starting a psychostimulant is extremely rare. Studies of EEG epileptiform activity, seizure rates, or interactions between antiepileptic drugs (AEDs) and psychostimulants have not shown any additional seizure risk. On the contrary, studies have shown that children with ADHD and a seizure disorder that is adequately controlled on an AED show reduced ADHD symptoms and no change in seizure frequency.

Aggression, Oppositional Defiant Disorder, and Conduct Disorder. Individuals with ADHD and comorbid Aggression, ODD and/or CD require comprehensive intervention which includes pharmacotherapy and psychosocial treatments. Stimulants have been shown in randomized, controlled trials to improve ADHD, aggression, ODD and CD symptoms. α-Blockers (guanfacine and clonidine) alone or as adjuncts to stimulants have also been shown to decrease ADHD, aggression, and oppositional symptoms. Stimulants also reduced negative social interactions and covert antisocial behavior (stealing and vandalism but not cheating). However, one cannot solely rely on medication to treat children with ADHD and these comorbid conditions. A comprehensive treatment approach which may include parent training, social skills training, and anger management for the child as well as an appropriate academic program and support may also be needed. Treating Aggression, ODD, and CD early and effectively is critical as long-term consequences of these disorders include later drug and alcohol abuse and antisocial behavior with involvement of the criminal justice system.

Anxiety Disorder. It had been suggested that children with ADHD and

comorbid anxiety may be less responsive to and experience more side effects with stimulant medication (Coughlin et al., 2015). The large Multimodal Study of ADHD (MTA) did not find this to be the case. However, children with ADHD and anxiety disorder seemed to benefit significantly from psychosocial treatment as well as medication. Thus for children with ADHD and comorbid anxiety, slow and careful titration of stimulant medication for the ADHD symptoms is recommended and psychosocial treatments such as parent training, social skills training, and/or CBT for the child combined with an appropriate academic program and academic support is suggested. However, should all these measures prove inadequate, addition of SSRIs (fluoxetine [Prozac] or sertraline [Zoloft]) to address the anxiety can be added.

Mood Disorder. In community samples, major depressive disorder (MDD) in youth with ADHD is more than five times higher than in youths without ADHD. Depressive disorders in youth with ADHD usually occur several years after the onset of ADHD and may be due to the cumulative effects of ADHD-related impairments and negative environmental circumstances they need to cope with. Youths with depression and ADHD have greater levels of psychosocial impairment than those with either condition alone. For individuals with both disorders, depressions start earlier, last longer, are more likely to recur and have higher rates of suicides and hospitalizations than those who have depression alone.

A combination of pharmacotherapy and psychosocial interventions is recommended for patients who have both ADHD and major depression. Generally, as per the Texas Children's Medication Algorithm Project (CMAP), it is suggested that the clinician begin with the most effective medication for the most impairing condition for the patient. Only after the full effect of this medication is determined for both conditions, should a second medication be added to deal with residual symptoms of the second condition. Thus, in some cases of the combination of ADHD and mood disorders both stimulants and SSRIs may need to be used.

Psychosocial interventions both for the depression, for example, CBT, and the ADHD, for example, organizational, study skills remediation, and appropriate academic and/or vocational program may all be needed.

The question of ADHD and Bipolar Disorder is being resolved with the introduction of the clinical entity of Disruptive Mood Dysregulation introduced in DSM-5. As a result, mainly classical symptoms of Bipolar Disorder such as severe elation of mood, grandiosity, racing thoughts, and hypersexuality are used to confer the bipolar diagnosis. In these circumstances, the Bipolar Disorder needs to be stabilized with mood stabilizers and/or antipsychotics before stimulant treatment can be introduced.

Developmental Disorders. ADHD symptoms of inattention and hyperactivity are seen in children with mental retardation and ASD. Stimulants have been shown to be beneficial in both groups, particularly in those children with IQs greater than 50. A recent randomized, controlled trial found that MPH was more effective than placebo in reducing symptoms of ADHD in children with autism, while few effects were seen on ASD symptoms.

Some studies suggest that these populations may be more susceptible to adverse events of psychostimulants. While other studies do not suggest this. However, if severe side effects occur, for example, severe agitation, irritability, fearfulness, or stereotyped behavior, a low dose of second-generation antipsychotic medication may need to be added.

There are few randomized, double-blind controlled trials with other medications for patients with ADHD and ASD. Atomoxetine has been shown to have some positive effects on ADHD symptoms with little effect on ASD. The side effects associated with Atomoxetine appear to be less problematic than with stimulants. There are some clinical suggestions that some anticonvulsant medications and guanfacine may also be beneficial for the ADHD symptoms in these populations but no well-controlled trials have been reported to date. Although psychosocial interventions are also suggested, again data on the efficacy of these approaches is also lacking.

Psychosocial Treatment of Children with ADHD

Psychosocial treatment of children with ADHD refers to nonmedication treatment. This type of treatment includes different modalities, such as psychoeducation, academic organization skill teaching and remediation, parent training, behavior modification, CBT, social skills training, and individual therapy. Of these modalities, parent training, intensive behavior modification, social skills training, and academic organizational skills have shown efficacy for children with ADHD in controlled trials.

Among nonpharmacological treatments, direct contingency management uses systematic manipulation of punishments and rewards, often in specialized settings. Although it can produce impressive effects on behavior and performance in the classroom, its effects tend not to generalize or be maintained outside of the settings in which they are applied. In addition, published outcomes are achieved through single-case designs. In the field, behavior therapy often is carried out in consultation with teachers and parents regarding in-school and home management. Published trials often depend on parent and teacher ratings of improvement but do not include independent evaluators. However, a recent review showed that even with blinded assessments, significant effects persisted for improving parenting and decreasing childhood conduct problems. Although controlled studies of systematic combinations of direct contingency management plus intensive behavior therapy have yielded

significant improvement, effect sizes of clinical change are smaller than those found with stimulant medication.

Intensive behavioral interventions are based on a number of principles. These begin with psychoeducation about the course, risk factors, and long-term outcomes of ADHD. Second, the parents are encouraged to attend more carefully to their child's behavior, particularly when the child complies. Third, the parents are trained to use time out effectively. Fourth, the parents are instructed in establishing a contingency management or token economy system at home. Then the parents learn how to manage noncompliant behaviors in public settings. Finally, advances in prosocial behavior in school are supported by use of a daily report card.

It is crucial to evaluate the parents and family for dysfunction related to the child's ADHD. Parental ADHD may interfere with behavioral modification programs, indicating that treatment of the affected parent may be necessary before the child's intervention can be successful. Other dysfunctions might be present in the family as well, such as marital problems, substance abuse, or parental depression.

Behavior modification has been successfully applied to the classroom, with a meta-analysis of 70 studies showing an effect size of 0.6 SD compared to an attention control condition. In contrast, CBT has not been shown to be effective (Vidal et al., 2015). Practice parameters published for child and adolescent psychiatrists strongly suggest that behavior therapy in the form of behavioral parent training and behavioral classroom management be added to psychopharmacological treatment if the child is shown to be a partial responder to stimulant medications or exhibits a comorbid psychiatric disorder whose impairment does not respond to the medication.

SUMMARY AND CONCLUSIONS

ADHD is a prevalent condition which begins in childhood but often continues into adolescence and adulthood. It is accompanied by significant academic, occupational, social, and emotional impairment. The patients with ADHD are often comorbid with other conditions. These comorbidities include oppositional defiant, conduct, learning, and anxiety disorders in childhood, and anxiety, depression, and substance use and abuse in adolescence and adulthood.

Treatment must include the ADHD and the various comorbidities. Thus, pharmacotherapy and psychosocial treatment as well as appropriate academic programs and support are needed for optimal outcome. It must be stressed that ADHD is a chronic condition for which ongoing long-term monitoring and treatment is required to optimize functioning.

SPECIFIC CASE OF CHILDHOOD ADHD

Ben is an 8.5-year-old African-American adopted boy who lives with his parents and 11-year-old brother. He attends third-grade regular education in his local public school and received special services including speech, physical, and occupational therapy. Although his drug screen at birth was negative, Ben's biological mother was said to have used crack and methadone during her pregnancy. No information was available about his biological father. Ben had suffered several upper respiratory infections at age 2 years requiring several rounds of antibiotics. Developmentally, some motor delays were reported, with crawling by 1 year and walking by age 2 years. He was speaking one word by his first birthday, but by age 3½ years his expressive language was behind his peers in fluency. He demonstrated difficulties following multipart directions, even though his hearing tested in the normal range.

Ben has had difficulty with daydreaming and wandering about since kindergarten. His teachers complained that Ben did "not seem to listen," had "poor concentration," and was constantly "out of his seat." Ben had never had close friends and other children seemed to avoid him because he could not stick with an activity for very long and he was always "in their space." Ben's adoptive parents were drained from the nightly battles over homework and discouraged by his lack of focus and "spaciness." His inattention and impulsivity put him in danger and caused injury. He had run into traffic on several occasions, even after being hit by a truck with resulting injuries.

Ben's parents and his third-grade teachers completed the SNAP Rating Scale. Clinician assessment revealed seven symptoms of inattention. His mother also endorsed all nine symptoms of hyperactivity/impulsivity. These symptoms were moderately to severely impairing across multiple settings, including home, peer relations, personal safety, and during athletics.

Ben was started on extended-release MPH at 18 mg daily which was later increased to 27 mg daily. His teachers reported improvement in attention and hyperactive behavior at school. Homework was no longer such a struggle and he was noted to play quietly with his toys in the evenings.

References

American Academy of Pediatrics; Subcommittee on Attention-Deficit/Hyperactivity Disorder and Committee on Quality Improvement. Clinical practice guideline: treatment of the school-aged child with attention-deficit/hyperactivity disorder. *Pediatrics*. 2001;108(4):1033–1044.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

Arns M, Conners CK, Kraemer HC. A decade of EEG theta/beta ratio research in ADHD: a meta-analysis. *J Atten Disord*. 2013;17(5):374–383.

Barkley RA. Sluggish cognitive tempo (concentration deficit disorder?): current status, future directions, and a plea to change the name. *J Abnorm Child Psychol*. 2014;42(1):117–125.

Castellanos FX, Margulies DS, Kelly C, et al. Cingulate–precuneus interaction; a new locus of dysfunction in adult attention-deficit/hyperactivity disorder. *Biol Psychiatry*. 2008;63(3):332–337.

Coughlin CG, Cohen SC, Mulqueen JM, Ferracioli-Oda E, Stuckelman ZD, Bloch MH. Meta-analysis: reduced risk of anxiety with psychostimulant treatment in children with attention-deficit/hyperactivity disorder. *J Child Adolesc Psychopharmacol*. 2015;25(8):611–617.

Daley D, Van der Oord S, Ferrin M, et al; European ADHD Guidelines Group. Behavioral interventions in attention-deficit/hyperactivity disorder: a meta-analysis of randomized controlled trials across multiple outcome domains. *J Am Acad Child Adolesc Psychiatry*. 2014;53(8):835–847, 847.e1–e5.

Daviss WB. A review of co-morbid depression in pediatric ADHD: etiologies, phenomenology,

and treatment. *J Child Adolesc Psychopharmacol*. 2008;18(6):565–571.

Faraone SV, Khan SA. Candidate gene studies of attention-deficit/hyperactivity disorder. *J Clin Psychiatry*. 2006;67(Suppl 8):13–20.

Gizer IR, Ficks C, Waldman ID. Candidate gene studies of ADHD: a meta-analytic review. *Hum Genet*. 2009;126(1):51–90.

Greenhill L, Kollins S, Abikoff H, et al. Efficacy and safety of immediate-release methylphenidate treatment for preschoolers with ADHD. *J Am Acad of Child Adolesc Psychiatry*. 2006;45(11):1284–1293.

Hart H, Radua J, Nakao T, Mataix-Cols D, Rubia K. Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects. *JAMA psychiatry*. 2013;70(2):185–198.

Hazell P. Drug therapy for attention-deficit/hyperactivity disorder-like symptoms in autistic disorder. *J Paediatr Child Health*. 2007;43(1-2):19–24.

*Hechtman L. Assessment and diagnosis of attention-deficit/hyperactivity disorder. *Child Adolesc Psychiatr Clin N Am*. 2000;9(3):481–498.

Hechtman L. Developmental neurobiological and psychosocial aspects of hyperactivity, impulsivity, and attention. In: Lewis M, ed. *Child and Adolescent Psychiatry; A Comprehensive Textbook*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2002:366–387.

Kooij SJ, Bejerot S, Blackwell A, et al. European consensus statement on diagnosis and treatment of adult ADHD: the european network adult ADHD. *BMC Psychiatry*. 2010;10(1):67.

Matte B, Anselmi L, Salum GA, et al. ADHD in DSM-5: a field trial in a large, representative sample of 18- to 19-year-old adults. *Psychol Med*. 2015;45(2):361–373.

McGough J, McCracken J, Swanson J, et al. Pharmacogenetics of methylphenidate response in preschoolers with ADHD. *J Am Acad Child Adolesc Psychiatry*. 2006;45(11):1314–1322.

McKeown RE, Holbrook JR, Danielson ML, Cuffe SP, Wolraich ML, Visser SN. The impact of case definition on attention-deficit/hyperactivity disorder prevalence estimates in community-based samples of school-aged children. *J Am Acad Child Adolesc Psychiatry*. 2015;54(1):53–61.

Nakao T, Radua J, Rubia K, Mataix-Cols D. Gray matter volume abnormalities in ADHD: voxel-based meta-analysis exploring the effects of age and stimulant medication. *Am J Psychiatry*. 2011;168(11):1154–1163.

Nigg JT, Lewis K, Edinger T, Falk M. Meta-analysis of attention-deficit/hyperactivity disorder or attention-deficit/hyperactivity disorder symptoms, restriction diet, and synthetic food color additives. *J Am Acad Child Adolesc Psychiatry*. 2012;51(1):86–97.

*Pliszka S. AACAP Work Group on Quality Issue. Practice parameters for the assessment and treatment of childhood attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2007;45(6):642.

Polanczyk G, Caspi A, Houts R, Kollins SH, Rohde LA, Moffitt TE. Implications of extending the ADHD age-of-onset criterion to age 12: results from a prospectively studied birth cohort. *J Am Acad Child Adolesc Psychiatry*. 2010;49(3):210–216.

Polanczyk G, de Lima MS, Horta BL, Biederman J, Rhode LA. The worldwide prevalence of

ADHD: a systematic review and metaregression analysis. *Am J Psychiatry*. 2007;164(6):942–948.

Pringsheim T, Steeves T. Pharmacological treatment for Attention Deficit Hyperactivity Disorder (ADHD) in children with comorbid tic disorders. *Cochrane Database Syst Rev*. 2011;(4):CD007990.

Storebø OJ, Krogh HB, Ramstad E, et al. Methylphenidate for attention-deficit/hyperactivity disorder in children and adolescents: Cochrane systematic review with meta-analyses and trial sequential analyses of randomised clinical trials. *BMJ*. 2015;351:h5203.

The MTA Cooperative Group. Moderators and mediators of treatment response for children with attention-deficit/hyperactivity disorder: the multimodal treatment study of children with attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry*. 1999;56(12):1088–1096.

Vidal R, Castells J, Richarte V, et al. Group therapy for adolescents with attention-deficit/hyperactivity disorder: a randomized controlled trial. *J Am Acad Child Adolesc Psychiatry*. 2015;54(4):275–282.

Visser SN, Danielson ML, Bitsko RH, et al. Trends in the parent-report of health care provider-diagnosed and medicated attention-deficit/hyperactivity disorder: United States, 2003–2011. *J Am Acad Child Adolesc Psychiatry*. 2014;53(1):34–46.e2.

Vitiello B, Lazzaretto D, Yershova K, et al. Pharmacotherapy of the preschool ADHD treatment study (PATS) children growing up. *Am Acad Child Adolesc Psychiatry*. 2015;57(7):550–556.

Willcutt EG. The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review. *Neurotherapeutics*. 2012;9(3):490–499.

▲ 45.2 Adult Manifestations of Attention-Deficit/Hyperactivity Disorder

JAMES J. MCGOUGH, M.D., M.S.

Although attention-deficit/hyperactivity disorder (ADHD) was originally conceptualized as a developmental delay of childhood, it is now defined as a lifespan neurodevelopmental disorder and adult manifestations are widely recognized. Longitudinal studies suggest that a majority of school-age children diagnosed with ADHD continue to have clinically meaningful symptoms and impairment into young adulthood. Numerous lines of epidemiological evidence reveal that prevalence rates of adult ADHD are among the highest of all psychiatric disorders. Clinical investigations indicate that adults with ADHD have patterns of psychiatric comorbidity and neurocognitive deficits similar to children, but suffer with a much broader array of functional impairments. Both structural and functional brain imaging studies demonstrate distinct abnormalities in groups of adults with ADHD compared to those without, and genetics studies conducted in adults have replicated pediatric findings. Adult ADHD is associated with significant medical, occupational, economic, and social impacts, and has emerged as a major public health concern.

DEFINITION

ADHD was classified as a Neurodevelopmental Disorder in the 5th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), reflecting the disorder's initial emergence in childhood and common cooccurrence with other early-onset brain-based syndromes. In a departure from previous editions, DSM-5 fully acknowledged the persistence of ADHD across the lifespan and modified diagnostic criteria for older adolescents and adults to improve symptom sensitivity to developmental changes with increasing age. While criteria in younger patients require evidence of six or more inattentive or hyperactive/impulsive symptoms of either inattention or hyperactivity/impulsivity, this threshold was reduced to five symptoms in either domain for diagnosis of patients aged 17 years and older. This modification recognized that the population prevalence of ADHD symptoms decreases with increased age, but adults with as few as four symptoms suffer significant clinical impairments compared with similarly aged peers. As with younger individuals, symptoms must have been present for a minimum of 6 months and exceed what is expected for a given patient's expected developmental level. Additionally, there must be some evidence that symptoms occurred prior to age 12 and have manifested in multiple life settings, such as home, school, work, social relationships, or personal activities. Symptoms must currently lead to clinically meaningful impairment in social, academic, or occupational functioning. Symptoms cannot occur exclusively during an episode of schizophrenia or other psychotic illness, and cannot be better attributed to other psychiatric conditions such as mood, anxiety, dissociative, or substance use disorders.

DSM-5 eliminated ADHD subtypes. Instead, *specifiers* describe the overall clinical presentation evidenced in the prior 6 months. Adults with at least five symptoms of inattention and hyperactivity/impulsivity exhibit the combined presentation. Those with five or more inattentive but fewer than five hyperactive/impulsive symptoms exhibit the inattentive presentation. Those with a minimum of five hyperactive/impulsive but fewer than five inattentive symptoms exhibit the hyperactive/impulsive presentation.

Other specifiers describe disorder severity. *Mild ADHD* is specified when the minimal number of symptoms required for diagnosis occur and functional difficulties are at the threshold for clinical significance. *Severe ADHD* is specified when there are many symptoms in excess of the required threshold, or in the presence of marked functional difficulties in one or more settings. Moderate ADHD is specified when the number of symptoms and associated degree of impairment falls between "mild" and "severe."

DSM-5 further permitted the diagnosis of "other specified ADHD" when an individual exhibits significant impairment due to ADHD symptoms, but fails to meet full diagnostic criteria. In this instance, the clinician should

specify which criteria are not met, for example, "other specified ADHD with insufficient symptom number." Clinicians may also assign the diagnosis "unspecified ADHD" if they choose not to specify which criteria are lacking.

HISTORY

For the first half of the 20th century, the condition now recognized as ADHD was variably referred to as minimal brain damage (MBD), minimal brain dysfunction, or hyperkinesis. While these terms conveyed some expectation that the disorder arose from underlying brain pathology, the syndrome was principally viewed as a developmental delay that was most often outgrown during adolescence. This position was maintained in the second edition of the DSM (DSM-II), published in 1968, which placed the diagnostic category named *hyperkinetic reaction of childhood* within the grouping "Behavioral Disorders of Childhood and Adolescence." Unique among editions of the DSM, DSM-II further implied that core features of short attention span, restlessness, and hyperactivity were a reaction to unsupportive environments and/or inadequate parenting.

In the 1970s, concurrent with development of the third edition of the DSM (DSM-III), seminal work by Paul Wender and colleagues at the University of Utah identified a group of adults with persistent childhood ADHD symptoms. Wender asserted that symptoms under consideration for inclusion in the DSM revision were developmentally inappropriate for older patients. He proposed the Wender Utah Criteria for Adult ADHD, empirically derived from extensive experience with older patients, as an alternative to the DSM. Wender Utah Criteria required (1) retrospective diagnosis of childhood ADHD made on the basis of patient interview and, when available, confirmation by a third-party informant; (2) evidence of ongoing impairment in adulthood from hyperactive and impulsive symptoms adulthood; and (3) the presence of additional symptoms felt to be typical of adults with the disorder, including mood lability, irritability, impulsivity, disorganization, and poor stress tolerance. The diagnosis could not be made in the patients suffering from severe depression, psychosis, or severe personality disorders.

DSM-III, published in 1980, continued to define attention deficit disorder as a childhood condition. However, with some deference to work by Wender and others, DSM-III allowed the diagnosis of attention deficit disorder, residual type, for adults previously diagnosed as children who had ongoing evidence of some symptoms. This category inherently implied that adults could not meet full diagnostic criteria. Text of the revised third edition (DSM-III-R) allowed that ADHD might continue through adulthood in up to 30 percent of affected children.

Throughout the 1980s and 1990s, and in the absence of clearly applicable DSM criteria, Wender Utah Criteria were used widely in clinical practice as well as research studies of adult medication response, brain

imaging, and genetics. Concurrently, longitudinal follow-up studies of adults previously diagnosed with childhood hyperactivity reported ADHD persistence in 40 to 60 percent of cases. These converging lines of evidence influenced subsequent DSM conceptualizations of the disorder and increased recognition of ADHD in adults. Although the Wender Utah Criteria had tremendous impact in promoting initial acknowledgment and study of adult ADHD, their use was largely abandoned as DSM criteria were increasingly applied outside of pediatric age groups.

In the 1990s, emerging research reports and several best-selling lay publications spurred widespread public awareness of adult ADHD. These studies and concomitant increases in public demand for clinical services catalyzed growing interest in the phenomenology, causes, and potential treatments of adult ADHD. Interestingly, although the fourth edition (text revision) of the DSM (DSM-IV-TR) provided some further leeway for adult diagnosis, it continued to maintain that for most cases ADHD symptoms resolved by late adolescence.

Revisions in DSM-5 incorporated more than two decades of research and clinical experience in adults. Enumerated ADHD symptoms remained unchanged from those previously identified in field trials of school-age children, but text clarifications were added to improve recognition of developmental changes that occur with increasing age. The age of onset criterion from previous editions of the DSM was modified from requiring evidence of symptoms and impairment prior to age 7 to requiring some recognition of symptoms alone prior to age 12. In response to studies that demonstrated an overall decline in symptom prevalence rates in older populations as well as associated functional impairments seen in adults with as few as four ADHD symptoms, only five symptoms in either the inattentive or hyperactive/impulsive domain were required for individuals aged 17 and older. Finally, descriptions of typical impairments associated with ADHD were broadened to include the range of difficulties frequently seen in adults, notably problems with work, personal activities, social relationships, and general health maintenance.

Comparative Nosology

In the 10th revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10), the entity defined as ADHD is subsumed under the category hyperkinetic disorders. Similar to DSM-5, ICD-10 requires impaired attention and overactivity that are developmentally inappropriate for a person's age. Unlike the DSM, both symptoms of inattention and overactivity are required, making it more difficult to apply the diagnosis to adults who generally exhibit fewer problems with hyperactivity. Also unlike DSM-5, but similar to previous versions of the DSM, ICD-10 retains the requirement that ADHD symptoms must have been generally present before 6 years of age. ICD-10

divides ADHD into subtypes based on the presence or absence of conduct disorder.

Epidemiology

Epidemiological studies demonstrated that ADHD is the most common behavioral disorder of childhood, with prevalence estimates ranging from 5 to 10 percent of school-age children. Results from the National Comorbidity Survey Replication estimated a 4.4 percent prevalence of ADHD among adults in the United States. International studies found slightly lower but comparable rates. These epidemiological studies are consistent with findings from longitudinal reports that describe persistence of ADHD in over 70 percent of adolescents and 50 percent of young adults. These converging reports confirm that significant numbers of adults suffer from ADHD and related impairments. Unlike other common psychiatric conditions such as mood and anxiety disorders that are typically episodic, adults with ADHD are apt to experience chronic and unremitting symptoms that require ongoing management.

Etiology

ADHD is best viewed as a neurodevelopmental disorder that arises from interplay of genetic and environmental factors. Twin and family genetic studies demonstrated that the mean heritability of ADHD exceeds 75 percent, approaching the estimated heritability of height and exceeding that seen in other psychiatric disorders such as schizophrenia and anxiety. Approximately one-third of adults with ADHD have at least one child with the disorder, and similarly, one-third of children with ADHD have at least one affected parent. Evidence suggests that ADHD is a complex behavioral phenotype resulting from multiple genes, each contributing additive risk for the condition. Early molecular genetic studies identified numerous candidate genes, particularly ones involved in brain catecholaminergic activity, with demonstrated associations with ADHD. Meta-analyses of these studies confirmed positive associations for several candidates, but none predicted substantial ADHD risk. Similarly, initial genome-wide association studies (GWAS) of ADHD failed to identify genomic regions with major risk polymorphisms. However, a related GWAS investigation found several regions likely to harbor common risk alleles for ADHD, depression, bipolar disorder, autism spectrum disorder, and schizophrenia. Other studies suggested individuals with ADHD have greater numbers of rare copy number variants (CNVs) throughout their genomes, suggesting that the sum of individual mutations better predicts disorder occurrence than specific mutations within given sets of genes. Ongoing investigations will further elucidate the genetic, environmental, and gene-by-environment interactions that affect ADHD risk.

Structural and functional brain imaging studies consistently demonstrated significant differences between adults with and without ADHD. It is unclear if these differences are primary causes of ADHD or secondary effects. Comparisons using computerized tomography (CT) and magnetic resonance imaging (MRI) found smaller total brain volumes, particularly in frontal and cerebellar regions. Positron emission tomography (PET) studies of adult parents of ADHD-affected children revealed decreased glucose metabolism in premotor and superior prefrontal cortices. Several studies utilizing single-photon emission tomography (SPECT) demonstrated increased dopamine transporter (DAT)-binding densities in the striatal regions. This is a particularly interesting finding given that methylphenidate (Ritalin) blocks DAT activity, possibly leading to normalized DAT functioning, and is consistent with reports implicating the DAT gene with ADHD. In more recent investigations that examined dysfunctional patterns and altered connectivity across multiple neuronal networks, adults with ADHD exhibited decreased activation in frontoparietal systems and overactivation in the visual, dorsal attention, and default networks.

Several environmental variables contribute increased risk for development of childhood ADHD, notably maternal use of nicotine or alcohol during pregnancy, birth prior to completion of 32 weeks' gestation, and increased serum lead levels. Adverse factors such as paternal incarceration, parental mental illness, financial stress, foster care, and child abuse increase risk for more severe adult outcomes, while higher intelligence and consistent structure at home and school decrease risk. Adults with concomitant executive functioning deficits generally exhibit more severe difficulties than those without. Nonetheless, while up to 50 percent of diagnosed children fail to manifest adult ADHD, the factors that contribute to remission or persistence of ADHD are not identified. Further elucidation of specific risk factors for adult persistence might facilitate development of preventative strategies and improved long-term outcomes.

Table 45.2–1.
DSM-5 Symptoms and Adult Correlates

Inattention		Hyperactivity and Impulsivity	
Younger Individuals	**Older Individuals**	**Younger Individuals**	**Older Individuals**
Fail to notice small details, make careless errors	Careless, inaccurate work	Frequently fidget, tap hands or feet, squirm while seated	Often fidgety or restless
Difficulty maintaining attention	Problems staying focused on boring work, mind often wanders, trouble with extensive reading	Frequently get out of seat or have trouble staying seated	Can't sit long at meetings or other activities where seating is required

Seem not to listen even when spoken to directly	Trouble retaining what people say even when addressed directly	Often run or climb at inappropriate times	Often feel internally restless or unsettled
Trouble finishing work or following through on instructions	Sidetracked easily, problems following directions or finishing tasks	Have trouble playing quietly; often too loud	Often louder than necessary or appropriate
Problems with organization	Can't prioritize, procrastinate, messy, miss deadlines, and bill payments	Often "on the go," as if "driven by a motor"	Difficulty unwinding or relaxing; always needing to be busy
Avoid activities that require sustained concentration	Delay and avoid work and other activities that are boring or repetitive	Get in trouble for excessive talking	Talk too much in social settings, don't tolerate silence
Lose or misplace things frequently	Often misplace things needed at home, work, or other activities	Often blurt out answers before questions are finished	Finish others' sentences or answers before questions are completed
Distracted easily when interrupted	Easily distracted by surrounding noises, interruptions, or extraneous thoughts	Have difficulties when waiting for turn or standing in line	Frustrated when needing to wait
Frequently forgetful	Fail to remember appointments, deadlines, and other obligations	Frequently interrupt or intrude	Interrupt others' conversations and activities, socially intrusive

Adapted from American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

DIAGNOSIS AND CLINICAL FEATURES

Although parents and teachers generally serve as primary informants in the assessment of pediatric ADHD, adult patients themselves are usually the primary sources of information in their own assessments. Unfortunately, research showed that young adults are frequently poor reporters of ADHD symptoms and associated difficulties, although reliability is much improved when patients reach their late 20s. Studies also revealed that both adults with ADHD and parents of these adult children exhibited poor retrospective recall of problems related to childhood ADHD, often reporting ages of onset up to 4 years later than documented in prospectively collected data. As with children, optimal assessment of adult ADHD depends on the clinical synthesis of all available information, including reports of childhood and current difficulties, symptom reports from parents and significant others, school and occupational reports, and a review of pertinent medical records.

Consideration of an ADHD diagnosis should be given to any adult with long-standing difficulties from inattention or distractibility. A single DSM symptom, "Is often distracted by extraneous stimuli," correctly differentiated 97 percent of ADHD-affected adults compared with 98 percent of community controls. One additional symptom, "Avoids tasks requiring sustained mental effort," correctly identified 89 percent of ADHD cases when compared with other clinical cases without ADHD.

As with children and adolescents, diagnosis of ADHD in adults is based on DSM-5 criteria. DSM-5 acknowledged developmental changes observed in adults by requiring only five inattentive and/or hyperactive/impulsive symptoms for diagnosis, compared with six for younger individuals. ADHD symptoms in DSM-5 are identical to DSM-IV and were initially derived from field trials conducted in school-age children. However, text descriptions in DSM-5 provided adult correlates for these symptoms and allow greater applicability to older patients, as summarized in Table 45.2–1.

ADHD assessment requires careful consideration of potentially comorbid psychopathology as well as a broad range of functional impairments. Common cooccurring conditions include lifetime histories of disruptive behavior and antisocial activities, nicotine dependence, substance use disorders, anxiety disorders, depression, risky behavior, and poor health maintenance. Typical impairments include difficulties in academic, occupational, social, and family functioning.

Diagnosis is based principally on information derived from a clinical interview with the adult patient. A self-administered questionnaire, the Adult ADHD Self-Report (ASRS), developed under the auspices of the World Health Organization, is useful for screening. Table 45.2–2 summarizes additional rating scales and structured diagnostic interviews used in adult assessment. Some clinicians remain highly skeptical of the validity of adult ADHD and have a general view that adults seeking the diagnosis are malingering or drug seeking. Clinical corroboration of self-reported ADHD symptoms and related difficulties with past records, third-party informants, and an overall life course consistent with the natural history of the disorder, remains the best approach for establishing the adult diagnosis.

Laboratory Studies

There are insufficient data to support the use of neuropsychological tests, computerized tests of attention (CPTs), or brain imaging modalities in the diagnosis of adult ADHD. Neuropsychiatric testing can be useful in differentiating ADHD from specific learning disabilities and other cognitive impairments, particularly with a predominance of inattentive symptoms or history of school failure. Nonetheless, neuropsychiatric testing has no demonstrated value in assessment of ADHD itself. Research studies using

electroencephalography (EEG), functional magnetic imaging (fMRI), PET, and other imaging modalities reveal distinct features in groups of adults with ADHD compared with nonclinical controls. However, these findings have poor predictive value in assigning diagnoses to individual patients and no evidence supports their utility in clinical settings. As with children and adolescents, adult ADHD remains a clinical diagnosis derived through careful history taking, adherence to clinical criteria, and differential diagnosis of other psychopathology.

Table 45.2–2.
Rating Scales and Diagnostic Instruments for Adult DSM Attention-Deficit/Hyperactivity Disorder (ADHD)

Rating Scales	Structured Interviews
Barkley Adult ADHD Rating Scale-IV (BAARS-IV)	Brown Adult ADHD Diagnostic Interview
Brown Adult Attention Disorders Scale	Conners' Adult ADHD Diagnostic Interview (CAADID)
Conners' Adult ADHD Rating Scales (CAARS): self and observer reports	Diagnostic Interview for ADHD in Adults (DIVA)
World Health Organization Adult ADHD Self-Report Scale (ASRS)	Adult ADHD Clinical Diagnostic Scale (ACDS)

Differential Diagnosis

Clinicians should consider a diagnosis of ADHD in any adult patient complaining of pervasive difficulties with inattention, hyperactivity, or impulsivity. Several ADHD symptoms, such as problems with sustained attention, motor restlessness, increased speech, and acting "on the go," are common in other psychiatric disorders. Clinicians must distinguish ADHD from other psychiatric syndromes while remaining sensitive to a high probability of coexisting psychopathology.

Depressive disorders are differentiated from ADHD in that problems with concentration and motor restlessness occur only during periods of dysphoric mood, anhedonia, or increased irritability. The occurrence of discrete episodes of somatovegetative symptoms, such as appetite changes and alterations in sleep patterns, further confirms a diagnosis of depression. Patients with both ADHD and depression will report ongoing difficulties with core ADHD symptoms that persist outside of discrete mood episodes. Similarly, patients with bipolar I disorder report racing thoughts, rapid speech, and impulsive behaviors only in the context of sustained periods of elevated or irritable mood. Somewhat more controversial is the overlap of ADHD with hypomania and rapidly cycling bipolar disorder. A clear past history of discrete depressive episodes or strong family history might assist in clarifying the diagnosis, but in many cases patients will meet criteria for both conditions. Anxiety disorders are distinguished from ADHD by subjective worry that is either generalized

and difficult to control, or specific to a particular situation or stimulus. Patients with substance use disorders and ADHD should have histories of ADHD symptoms that precede initiation of substance use. Patients with histories of early school difficulties and childhood conduct disorder are likely to be comorbid for both ADHD and substance use. Adults with ADHD and a childhood history of conduct disorder frequently exhibit antisocial personality disorders.

Adults with histories of school difficulties and inattentive symptoms in the absence of hyperactivity or impulsivity should be evaluated for specific learning disabilities or other executive functioning deficits. This generally requires a battery of neurocognitive tests to measure academic aptitude and achievement. Patients with significantly idiosyncratic interests and impaired social functioning might also be diagnosed with autism spectrum disorder.

Although considerable research has addressed the occurrence of significant psychiatric comorbidity in adults diagnosed with ADHD, far less attention has been paid to prevalence rates of ADHD cooccurring with other adult psychiatric disorders. Emerging evidence suggests that up to 16 percent of adults currently suffering from major depression have childhood histories of ADHD, and that in approximately 13 percent of adults with current mood disorders, 10 percent with current anxiety disorders, and up to 25 percent with current substance use disorders meet ADHD criteria. In view of its high overall adult prevalence and associated developmental risk for comorbid psychopathology, all adult psychiatric assessments should include routine diagnostic consideration of ADHD.

A 42-year-old woman obtained an evaluation for distractibility and general nervousness. The patient had spent the past 16 years at home caring for her children. Her youngest child recently began elementary school, and she obtained part-time work as an assistant manager in a large retail department store. She was excited about the position, particularly the flexible work hours and the opportunity to earn extra income for the family. However, she reported particular difficulty remaining focused on her job, had general trouble organizing her time, and was in danger of losing her job due to her failure meeting several deadlines. She was tense, irritable, had trouble relaxing, and described difficulty falling asleep because she was worried about what she needed to do the next day. She complained that her "mind goes completely blank" when overstressed. She denied depressed mood, loss of pleasure in her usual activities, change of appetite, or suicidal preoccupations.

The patient reported that she was considered a bright student, particularly in elementary school. She liked school, but was generally shy and uncomfortable in unfamiliar social situations. Beginning in middle school and throughout college she felt increasingly bored, easily distracted, and particularly frustrated with long reading assignments. She graduated with a 3.0 grade point average and subsequently worked for 2 years as a retail clerk. She became pregnant soon after her marriage and left her job to raise her family. She described her marriage as "typical," generally stable and satisfying with occasional marital tension due to financial stresses and child-rearing. Her middle child, a 12-year-old boy, had been diagnosed with ADHD and successfully managed with stimulant therapy and behavioral techniques learned through parent training.

On a self-completed ACDS Rating Scale, the patient endorsed six

inattentive and two hyperactive/impulsive symptoms. An ADHD-IV Rating Scale completed by her husband reflected a similar symptom pattern. Clinical synthesis of all available information supported an ADHD diagnosis with a retrospectively determined age of onset in middle school, a history of impairment at school, work, and home, and a comorbid diagnosis of generalized anxiety disorder.

The treating clinician outlined several options for medication therapy, specifically an initial stimulant trial with subsequent selective serotonin reuptake inhibitor (SSRI) treatment for residual anxiety versus atomoxetine (Strattera) monotherapy. The patient agreed to a medication trial with main target symptoms of distractibility, nervousness, and sleep disturbance.

On stable medication, she reported improved focus at work, better general organization, less general anxiety, and improved sleep. She was followed with monthly medication visits for 3 months and then visit frequency was reduced to every 3 months. After 6 months of medication treatment, the patient was pleased to report that she recently finished reading a novel, her first since college. She resolved her initial difficulties at work and had been given a commendable employment evaluation. She was feeling less stressed at home, but remained troubled by what she viewed as her inadequacies and inability to manage all her responsibilities. In response, her physician referred her for a brief course of cognitive-behavioral therapy (CBT) to address residual anxiety symptoms and organizational skills.

COURSE AND PROGNOSIS

The prevalence of ADHD symptoms declines with age. Many children with the combined presentation of ADHD exhibit fewer hyperactive-impulsive symptoms by mid-adolescence and subsequently exhibit the inattentive presentation as adults. Given the age-related decline of ADHD symptoms in the general population, adults can remain significantly impaired relative to same-age peers with fewer symptoms than are required for diagnosis in childhood. This suggests that reported rates of ADHD remission are more a function of syndrome definition than true syndrome remission. Long-term outcome studies of ADHD-affected children are limited, and there are insufficient data to draw conclusions regarding the developmental course of ADHD from early to middle adulthood. Between 30 and 45 percent of ADHD children do not suffer ADHD-related impairments as adults. Nonetheless, many adults with ADHD continue to struggle with the consequences of the disorder throughout their lives.

Adults with ADHD demonstrate significant rates of cooccurring psychopathology and limitations in adaptive functioning when compared with unaffected peers. Adults with ADHD exhibit higher rates of childhood conduct disorder with a subset having antisocial personality disorder.

Significantly higher rates of substance use disorders are consistently reported compared with unaffected adults. There is no clear evidence that adults with ADHD prefer one type of recreational drug compared with another, but those who develop drug dependency have a more difficult time achieving sobriety than individuals without ADHD. Similarly, adults with ADHD initiate nicotine use at earlier ages, demonstrate higher rates of lifetime nicotine addiction, and have more difficulty in maintaining smoking cessation. Adults with ADHD suffer high rates of lifetime mood and anxiety disorders. The differential impact of comorbid psychopathology on long-term ADHD outcome has not been adequately studied.

ADHD confers significant risk for lifetime functional impairments. Lower self-esteem, poorer lifetime academic achievement, less successful work histories, and more failures in long-term relationships are evident. Adults with ADHD often make poor choices regarding personal health, including participation in high-risk activities, unprotected sex, and unsafe driving. Conversely, some of these succeed in making lifestyle adjustments that minimize evidence of impairment. Individuals find occupations for which they are well suited, can live alone or find suitable close relationships, or can utilize other talents that compensate for ADHD-related difficulties. Additional research is required to identify predictive factors that mitigate these impairments.

A 26-year-old man obtained an evaluation for severe distractibility after being involved in a minor traffic accident. In early grade school he had particular difficulties making careless mistakes on tests, forgetting his assignments, and sitting still in class. Although he was considered intelligent, his grades were average and he had no aspirations for college. After high school he had several short-term jobs in the construction industry, which were enjoyable because he spent most of his time outdoors. In the past year he had obtained part-time work as a wedding videographer and found himself increasingly interested in making and editing videos.

During his clinical interview, the patient described ongoing difficulties maintaining his attention on uninteresting tasks, easy distractibility, problems with time management, generally feeling restless, and frustration when required to sit for extended periods. He endorsed six inattentive and five hyperactive/impulsive symptoms on a DSM-5 ADHD checklist of current symptoms. The patient was diagnosed with the combined presentation of adult ADHD.

After discussing the nature of ADHD with his physician and receiving additional educational materials, the patient agreed to a medication trial. A once-daily extended-release preparation of stimulant medication was prescribed in escalating doses over several weeks to determine the optimal dose that maximized symptom reduction and minimized side effects. At his follow-up visit, the patient admitted to robust improvement in his ability to focus and surprise at his increased motivation to attend to the technical details of video editing. After several months of treatment, he enrolled in a community college to pursue formal coursework in video production.

TREATMENT

Good long-term clinical care is contingent on adequate patient education about the nature of the disorder and effects of treatment. Once a diagnosis

is made, clinicians should provide sufficient information to assist adult patients in choosing treatment options that meet individual patient needs. Given the degree of misinformation about ADHD available from the Internet and other lay sources, patients are well served when physicians guide them to appropriate reference materials. Many patients are relieved to understand that they have a biologically based condition that is shared by a significant number of other adults. National organizations such as Children and Adults with ADD (CHADD, www.chadd.org) or the National Alliance for the Mentally Ill (NAMI, www.nami.org) provide access to local support networks and reliable educational materials.

Pharmacotherapy remains a treatment cornerstone. Medications found to be useful in the treatment of childhood ADHD are generally shown to be effective in adults. FDA-approved medications for adult ADHD include stimulant and nonstimulant medications, as summarized in Table 45.2–3. Other medications used off-label in adults with ADHD include immediate- and extended-release forms of the alpha-2-agonists, guanfacine and clonidine; tricyclic antidepressants, buproprion; and the wake-promoting agents, modafinil and armodafinil. Stimulant medications are the most frequently prescribed. The greatest clinical effect from stimulants is seen with robust dosing. A particular difficulty in clinical management of adults is the frequent necessity for multiple daily dosing, even with extended-release formulations, in order to provide adequate relief from ADHD symptoms throughout the day. Many clinicians remain reluctant to prescribe stimulant medication due to concerns about abuse, misuse, and illicit diversion. Several studies have reported widespread diversion of prescribed stimulants among college students, although in most cases the medications are used to enhance academic performance as opposed to recreational use. Several lines of evidence suggest that prodrug and certain extended-release stimulant formulations are less likely to create abuse liability than older immediate-release forms. Nonstimulant medications remain an option in patients at high risk for abuse or misuse of ADHD stimulants.

Table 45.2–3.
FDA-Approved Medications for Adult Attention-Deficit/Hyperactivity Disorder (ADHD)

	Brand Name	Typical Single Dose	Approved Maximum Dose (mg/day)
Stimulants			
Methylphenidate	Ritalin	5–20 mg	60
	Metadate CD	20–60 mg ER	60
	Ritalin LA	100–40 mg ER	60
	Concerta	18–72 mg ER	72
Dexmethylphenidate	Focalin	2.5–10 mg	20

	Focalin XR	10–40 mg ER	40
Dextroamphetamine	Dexedrine	5–30 mg	60
	Dextrostat	5–30 mg	60
	Dexedrine Spansule	5–40 mg SR	60
Lisdexamfetamine	Vyvanse	20–70 mg	70
Mixed amphetamine salts	Adderall	5–30 mg	40
	Adderall XR	5–30 mg ER	60
Nonstimulants			
Atomoxetine	Strattera	40–100 mg	100

ER, extended release; SR, sustained release.

Although many adults have significant lifetime histories of other psychiatric disorders, there have been few treatment studies for comorbid ADHD. Current evidence suggests that clinicians should first initiate therapy for the condition (i.e., ADHD vs. other psychopathology) that is causing the most distress. One example is sequential pharmacotherapy with either a stimulant or SSRI for cooccurring ADHD and anxiety or depression. A particular vexing problem is whether stimulant medications should be provided to patients with known histories of substance use disorder. It is essential that physicians maintain proper documentation of patients' diagnostic assessments and prescription medication records. Some clinicians are comfortable initiating ADHD pharmacotherapy in the presence of intermittent alcohol or cannabis use, but are advised against stimulant pharmacotherapy in those patients with chronic use or use of other psychoactive agents such as opioids or methamphetamine.

There is growing support for certain nonmedication treatments in adult ADHD. The strongest evidence exists for CBT, which has been administered successfully both in individual and group settings. Well-controlled blinded investigations of omega-3 fatty acid supplementation have also yielded small but significant positive effects on ADHD symptoms. Other nonmedication approaches, such as ADHD coaching, mindfulness training, meditation, yoga, cognitive remediation programs, EEG neurofeedback, and use of digital technologies such as electronic scheduling and reminder systems remain popular, despite a lack of sufficient scientific data to support their general use. Individuals are best served when psychosocial interventions are individually tailored to their own specific difficulties. For example, patients who suffer from marital or relationship difficulties might be referred to couples therapy. The emphasis in this treatment should be the couple's relationship problems, not the individual's ADHD. Parents might benefit from participation in a parenting course, particularly if they themselves have children with ADHD. Insight-oriented psychotherapy might be useful for some patients with significant self-esteem or interpersonal difficulties.

Proper management of adult ADHD is also likely to include appropriate

educational planning and consideration of academic and occupational accommodations. Occupational counseling and vocational training can be better options for some individuals than insistence on attending a 4-year college. Adults who attend college should be referred to their school's office for students with disabilities. The Americans with Disabilities Act (ADA) has particular implications for adults with ADHD. Regulations implementing the ADA require that *reasonable* accommodations be made in areas of education and employment for individuals with documented disabilities. Colleges and universities develop their own guidelines for provision of special accommodations based on Federal mandates. At higher-grade levels and for standardized testing at the college level and beyond, it is generally necessary to conduct formal neuropsychological testing to document ADHD-related learning impediments prior to granting accommodations. In occupational settings, the ADA requires that reasonable accommodations be made for persons with disabilities, but does not require that anyone be hired for a position who is not otherwise qualified. It is essential that thresholds for disability be determined in reference to population or societal norms, and not speculations about the individual's potential capabilities. *Reasonable* accommodations should have minimal impact on employer expectations, job performance, and coworkers.

A 19-year-old woman was evaluated at a drug rehabilitation center for methamphetamine dependence. In early grade school she was extremely hyperactive, reckless, and prone to impulsive judgments. Her father suffered from alcohol dependence and was physically abusive to her and her mother. She began smoking cigarettes at age 11 and using alcohol and marijuana at age 12. At age 13, she was raped by a high school senior that she had been dating. Two years later, she quit high school and ran away from home. She began smoking crystal methamphetamine the previous year, and was court ordered for drug evaluation after an arrest for prostitution.

A careful history of early school experiences revealed significant problems with distractibility, completing assignments, sitting still in class, and general irritability. On a DSM checklist assessing childhood symptoms, she endorsed six inattentive and eight hyperactive/impulsive symptoms. Currently, she endorsed nine inattentive and nine hyperactive/impulsive symptoms, general dysphoria, fatigue, anhedonia, and suicidal preoccupation.

The patient enrolled in a long-term, court-funded residential treatment program for drugs and alcohol. After several weeks her mood symptoms diminished, but she continued to suffer from marked inattention, distractibility, and restlessness. After careful consideration, her physician initiated nonstimulant medication treatment for ADHD. At 1-month follow-up, she described good response to treatment and had begun studying to obtain a general equivalency diploma (GED). At discharge, she had resolved to stay in outpatient drug treatment and obtain part-time work. Her physician continued to monitor her on a monthly basis with intermittent, unannounced drug testing to monitor ongoing sobriety.

FUTURE DIRECTIONS

Ongoing advances in neuroscience are apt to increase acknowledgment of adult ADHD as a biological syndrome. Changes in clinical practice increasingly reflect a shift from viewing adult ADHD as a unique and rare

phenomenon to routine consideration of ADHD both alone and as potentially comorbid with other traditional adult psychopathologies. There will be ongoing interest in developing medication treatments with less perceived potential for abuse and longer durations of effects. Concurrently, there will be sustained efforts to develop effective nonmedication therapies. Although revisions made to ADHD criteria for DSM-5 better reflect the developmental aspects of the behavioral phenotype, ongoing research will increasingly focus on identification and modification of underlying dysregulation in brain neural networks. Such discoveries would raise the possibility of early identification of risk and treatments that target specific dysregulated systems in individual patients.

REFERENCES

Adler L, Spencer T. *The Adult ADHD Clinical Diagnostic Scale (ACDS) V 1.2*. New York: New York University School of Medicine; 2004.

*Asherson P, Buitlelaar J, Faraone SV, Rohde LA. Adult attention-deficit hyperactivity disorder: key conceptual issues. *Lancet Psychiatry*. 2016;3:568–578.

Barkley RA. *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment*. 4th ed. New York: Guilford Press; 2014.

Barkley RA, Murphy KR, Fischer M. *The Science of ADHD in Adults: Clinic-Referred Adults and Children Grown Up*. New York: Guilford Press; 2007.

Brown TE. *Brown Attention-Deficit Disorders Scales, Manual*. San Antonio, TX: Psychological Corp; 1996.

Conners CK, Erhardt D, Sparrow EP. *Conners' Adult ADHD Rating Scales, Technical Manual*. North Tonawanda, NY: Multi-Health Systems; 1999.

Cross-Disorder Group of the Psychiatric Genomics Consortium; Lee SH, Ripke S, Neale BM, et al. Genetic relationships between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet*. 2013;45:984–994.

Cubillo A, Rubia K. Structural and functional brain imaging in adult attention-deficit/hyperactivity disorder. *Expert Rev Neurother*. 2010;10:603–620.

Doshi JA, Hodgkins P, Kahle J, et al. Economic impact of childhood and adult attention-deficit/hyperactivity disorder in the United States. *J Am Acad Child Adolesc Psychiatry*. 2012;51:990–1002.

DuPaul GJ, Power TJ, Anastopoulos AD, Reid R. *ADHD Rating Scale-IV: Checklists, Norms, and Clinical Interpretation*. New York: Guilford Press; 1998.

Faraone SV, Mick E. Molecular genetics of attention deficit hyperactivity disorder. *Psychiatr Clin North Am*. 2010;33:159–180.

Fayyad J, De Graaf R, Kessler R, et al. Cross-national prevalence and correlates of adult attention-deficit/hyperactivity disorder. *Br J Psychiatry*. 2007;190:402–409.

Kessler RC, Adler L, Ames M, et al. The World Health Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the general population. *Psychol Med*. 2005;35:245–256.

Kessler RC, Adler L, Barkley R, et al. The prevalence and correlates of adult ADHD in the United

States: results from the National Comorbidity Survey Replication. *Am J Psychiatry*. 2006;163:716–723.

*Kooij SJ, Bejerot S, Blackwell A, et al. European consensus statement on diagnosis and treatment of adult ADHD: The European Network Adult ADHD. *BMC Psychiatry*. 2010;10:67.

Manos MJ. Psychosocial therapy in the treatment of adults with attention-deficit/hyperactivity disorder. *Postgrad Med*. 2013;125:51–64.

*McGough JJ. *ADHD*. New York: Oxford University Press; 2014.

*McGough J. Treatment controversies in adult ADHD. *Am J Psychiatry*. 2016;173:960–966.

Mongia M, Hechtman L. Cognitive behavior therapy for adults with attention-deficit/hyperactivity disorder: a review of recent randomized controlled trials. *Curr Psychiatry Rep*. 2012;14:561–567.

Moriyama TS, Polanczyk GV, Terzi FS, Faria KM, Rohde LA. Psychopharmacology and psychotherapy for the treatment of adults with ADHD—a systematic review of available meta-analyses. *CNS Spectr*. 2013;18(6):296–306.

Nigg JT. Attention-deficit/hyperactivity disorder and adverse health outcomes. *Clin Psychol Rev*. 2013;33:215–228.

Polanczyk G, Caspi A, Houts R, Kollins SH, Rohde LA, Moffitt TE. Implications of extending the ADHD age-of-onset criterion to 12: results from a prospectively studied birth cohort. *J Am Acad Child Adolesc Psychiatry*. 2010;49:210–216.

*Purper-Ouakil D, Ramoz N, Lepagnol-Bestel AM, Gorwood P, Simonneau M. Neurobiology of attention deficit/hyperactivity disorder. *Pediatr Res*. 2011;69:69R–76R.

Rabiner DL. Stimulant prescription cautions: addressing misuse, diversion and malingering. *Curr Psychiatry Rep*. 2013;15:375.

Ramos Olazagasti MA, Klein RG, Mannuzza S, et al. Does childhood attention-deficit/hyperactivity disorder predict risk-taking and medical illnesses in adulthood? *J Am Acad Child Adolesc Psychiatry*. 2013;52:153–162.

Solanto MV, Wasserstein J, Marks DJ, Mitchell KJ. Diagnosis of ADHD in adults: what is the appropriate DSM-5 symptom threshold for hyperactivity-impulsivity? *J Atten Disord*. 2012;16:631–634.

Sonuga-Barke EJ, Brandeis D, Cortese S, et al; European ADHD Guidelines Group. Nonpharmacological interventions for ADHD: systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *Am J Psychiatry*. 2013;170:275–289.

Weyandt LL, Dupaul GJ. ADHD in college students: developmental findings. *Dev Disabil Res Rev*. 2008;14:311–319.

Weyandt L, Swentosky A, Gudmundsdottir BG. Neuroimaging and ADHD: fMRI, PET, DTI findings, and methodological limitations. *Dev Neuropsychol*. 2013;38:211–225.

Wilens TE, Morrison NR, Prince J. An update on the pharmacotherapy of attention-deficit/hyperactivity disorder in adults. *Expert Rev Neurother*. 2011;11:1443–1465.

46

Disruptive Behavior Disorders in Children and Adolescents

DANIEL F. CONNOR, M.D.

Oppositional defiant disorder (ODD) and conduct disorder (CD) are grouped as the disruptive behavior disorders and included in the DSM-5 chapter on disruptive, impulse control, and CDs. These conditions involve individual problems in emotional and/or behavioral regulation that either bring the individual into significant conflict with authority figures (ODD and CD) or violate the rights of others and/or societal norms (CD). Although ODD frequently overlaps with CD and their symptoms show high correlation, these two categorical disorders are relatively distinct dimensions of psychopathology, with some separate correlates, sequelae, and prognoses. Consistent with DSM-5, in this chapter ODD and CD are considered separately.

When behaviors and emotions are considered dimensionally rather than categorically, these disorders are associated with an externalizing spectrum that includes the personality dimensions of disinhibition, low constraint, and negative emotionality. The disruptive behavior disorders generally have an onset in childhood or adolescence, and are usually more common in males than females.

Considered together, these disorders are one of the most frequent reasons for mental health referral, accounting for between a third-and-half of all child and adolescent psychiatry clinic referrals. These disorders are compelling to understand and treat because they are common and because they are associated with high rates of morbidity, high rates of associated psychiatric illness and psychopathology, and are very costly for society, being associated with higher rates of public expenditure than for youngsters with other neuropsychiatric conditions. These disorders are frequently comorbid with or linked to increased risk for the development of other psychiatric disorders such as depression, bipolar disorders, anxiety, attention-deficit/hyperactivity disorder (ADHD), substance use, and/or antisocial personality disorders. Many societal institutions that serve children and adolescents including the family—the educational system, juvenile justice systems, community and public health systems, primary care systems, and mental health systems—experience the impact of

youngsters with early-onset disruptive behavior disorders.

OPPOSITIONAL DEFIANT DISORDER

Definition

The essential features of ODD are a frequent and persisting pattern of angry/irritable mood, argumentative/defiant behavior, or vindictiveness that is clearly more frequent, more intense, and more persistent across the child's development than is typically observed in individuals of similar age and developmental level. Argumentative and defiant behaviors are expressed by persistent stubbornness, resistance to directions, and unwillingness to compromise, give in, or negotiate with adults or peers. Defiance may also include deliberate or persistent testing of limits, usually by ignoring requests or commands, arguing, and failing to accept blame for misdeeds. Irritability can manifest as anger or aggression and can be directed at adults or peers. Temper outbursts and aggression generally involve the expression of verbal aggression, usually without the more serious physical aggression seen in CD. Symptoms of the disorder are typically more evident in interactions with adults or peers whom the individual knows well. Usually individuals with this disorder do not regard themselves as oppositional or defiant, but justify their behavior as a response to unreasonable demands or circumstances. Symptoms generally begin early in life and are usually more prevalent and severe at home in the child's interactions with the parents. Only with more severe and pervasive ODD and later in the course of illness, does the disorder tend to generalize out of the home with angry/irritable and argumentative/defiant symptoms expressed toward authority figures at school or in the community.

ODD is not recognized if symptoms only occur during the course of a psychotic, bipolar, and depressive or substance-use disorder. In these cases, the symptoms of ODD are felt to be part of the underlying disorder. Disruptive mood dysregulation disorder (DMDD) shares symptoms of chronic negative mood and temper outbursts with ODD. However, symptoms are much more severe, frequent, and chronic in individuals meeting criteria for DMDD than in those with ODD. In individuals who meet diagnostic criteria for both ODD and DMDD, only the more severe disorder is diagnosed, and a diagnosis of ODD is not given.

CD and ODD are both related to disruptive behaviors and externalizing symptoms that bring the individual into conflict with parents or other authority figures. The symptoms of ODD are generally less severe than those of CD and typically do not involve overt aggressive behaviors toward others or a pattern of deceit or theft. Furthermore, ODD includes problems with emotional regulation (irritability, anger) that are not found in the diagnostic criteria for CD. When criteria are met for both ODD and CD, both diagnoses can be given.

History

ODD emerged as a distinct diagnosis separate from CD for the first time in the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III). Previously, the Group for the Advancement of Psychiatry had suggested the concept of childhood oppositional personality in 1966, emphasizing the symptoms of disobedience and pervasive negativism. Early conceptualizations of the disorder emphasized the passive-aggressive nature of the oppositional symptoms. With the introduction of DSM-III in 1980, the idea of an oppositional disorder was first introduced and links to the notion of a passive-aggressive personality were retained. Subsequent editions of the DSM added a defiant label and expanded the ODD symptom list from two of five possible items (DSM-III), to five of nine items (DSM-III-R), and finally to four of eight items (DSM-IV, DSM-IV-TR, and DSM-5). As evidence mounted that early-onset oppositional behaviors were linked to increased risk for later antisocial behaviors, ODD became grouped with CD under the disruptive behavior disorders and links to passive-aggressive personality traits were dropped. Limited field trials supported the diagnostic criteria for ODD in DSM-III-R. More extensive field trials and secondary data analyses were performed in support of ODD in DSM-IV. Evidence from the DSM-IV field trials indicated that the diagnostic criteria of four of eight possible ODD items possessed adequate internal consistency, positive predictive power, and negative predictive power. Four of eight symptoms remain required for the diagnosis of ODD in DSM-5. Subsequent research supports the reliability and validity of the ODD category as separate from the more severe diagnosis of CD, and this distinction is retained in DSM-5.

Although the eight ODD symptoms found in DSM-5 are the same as for DSM-IV-TR, they are now grouped into three categories (i.e., angry/irritable mood, argumentative/defiant behavior, and vindictiveness). Recent research has shown that outcomes may differ depending on which grouping of ODD emotional/behavioral symptoms is most predominant. For example, the angry/irritable mood category predicts an increased risk for internalizing disorders, the argumentative/defiant behavior category predicts risk for disruptive behavior disorders, and the vindictiveness category may be associated with risk for instrumental aggressive behaviors. Thus, in DSM-5, ODD is now conceptualized as a disorder with both emotional and behavioral dimensions and contained in a new chapter on the disruptive, impulse control and CDs.

Comparative Nosology

Compared with DSM-IV-TR, the grouping of DSM-5 ODD symptoms into three categories suggests that subgroups of ODD children exist and can be identified each with differing risk for the development of associated

psychopathology. This may enhance differential provision of clinical treatments.

The DSM-5 and the 10th edition of the *International Classification of Diseases* (ICD-10) include similar symptoms in the criteria lists for making the diagnosis of ODD and CD. However, while DSM-5 treats each disorder separately and both disorders may be diagnosed in an individual, if criteria for both disorders are present, ICD-10 considers ODD jointly with CD. In DSM-5, ODD is specified as a separate disorder from CD with four of eight symptoms in three categories required for a diagnosis (Table 46–1). The symptom criteria for ODD are substantially different from the symptom set for CD, supporting the idea that they are separate disorders. However, the notion that ODD might be related as a less severe developmental precursor to CD is built into the DSM-5 criteria by grouping together argumentative/defiant behaviors as a subcategory of ODD symptoms that predicts increased risk for disruptive behavior disorders including CD.

ICD-10 includes a 23-item symptom list very similar to the combined symptom lists of DSM-5 ODD and CD, but divides them into 15 more severe items appropriate for the diagnosis of CD and eight less severe items equivalent to the DSM ODD symptoms (Table 46–2). A diagnosis of CD is made similarly in both systems using three or more of the more severe items in ICD-10. However, unlike the DSM system, ICD-10 treats ODD as a subtype of CD and requires a total of four or more symptoms from the full list of 23 items to be met, but where no more than two items come from the list of 15 more severe conduct symptoms. This results in an ICD-10 diagnosis of CD and ODD subtypes. As a result, all children who meet DSM-5 criteria for ODD also meet ICD-10 criteria for the disorder. However, not all children who meet an ICD-10 diagnosis of CD and ODD subtype (e.g., four ODD + CD symptoms combined total) will meet DSM-5 ODD diagnostic criteria (e.g., four ODD symptoms and no symptoms of CD required). Recent studies have shown that the DSM criteria may exclude from the diagnosis, children who meet criteria for an ODD diagnosis by ICD-10 criteria and that these omitted children are substantially impaired. Many of these omitted children may be assigned a DSM-5 diagnosis of other specified or unspecified disruptive, impulse control and CD. As a consequence, population prevalence rates in epidemiological studies of ODD may differ because of the different diagnostic systems utilized with the DSM system at risk for underestimating the true rate of ODD compared to ICD-10.

Epidemiology

The prevalence of disruptive behavior disorders varies according to age, gender, socioeconomic status, neighborhood, and degree of urbanicity. Evidence suggests that the population prevalence of ODD may be higher than CD during childhood and remains constant or declines somewhat

during adolescence (when CD prevalence begins to rise). The disorder appears more prevalent in males than females (1.4:1) prior to adolescence. The male preponderance disappears in adolescents and in adults with ODD. The criteria used to make the diagnosis also cause prevalence to vary. For example, studies have shown that the prevalence of ODD decreased as more stringent diagnostic criteria were applied from 1980 to 2013 across DSM-III to DSM-5. In nonreferred epidemiological samples, the prevalence of ODD varies between 4.5 and 15.4 percent in males and between 1.5 and 15.6 percent in females aged 7 to 21 years old. DSM-5 cites an average overall prevalence estimate of 3.3 percent for childhood ODD. Recent studies in adults have reported a 12-month ODD prevalence of 1 percent and a lifetime prevalence of ODD of 8.5 percent in the United States. Good data on the prevalence of ODD in the preschool population are presently lacking. ODD appears to be more prevalent among children and adolescents from families of low socioeconomic status living in inner city disadvantaged neighborhoods.

Etiology

Presently, there is little scientific investigation into the origin of ODD as a separate and discrete disorder. More is known about the etiology of CD and the general categories of disruptive behavior disorders, delinquency, and antisocial behavior than is understood about the specific causes of ODD. Most theories favor a cumulative risk model arising out of a complex interplay between biological and environmental risk and protective factors as the child interacts with caregivers, authority figures, and the environment across development. As risk factors accumulate and begin to outweigh protective factors, risk for the initiation of ODD increases. Most experts agree that single cause or main effects models are inadequate to explain the complexity inherent in the origin of ODD. Vulnerability gene by environmental interactions and dynamic, interactive models examining reciprocal interactions between the child and their caregivers over time as development unfolds appear best suited to eventually provide a clearer understanding of the etiology of ODD. Understanding the etiology of ODD is important because as risk aggregates, the impact of prevention interventions and treatments may diminish.

Biological. Biological investigations into the disruptive behavior disorders have generally combined CD and ODD together or focused on larger categories of disruptive behaviors such as conduct problems, aggression, delinquency, criminality, or antisocial behavior. When specific diagnoses are investigated, more studies have investigated the biological correlates of CD than those of ODD. Few studies have specifically examined the biological etiology of ODD. Neurobiological markers associated with CD such as lower heart rate and skin conductance reactivity to environmental

stimuli, reduced basal cortisol reactivity, and functional imaging abnormalities identified in the prefrontal cortex and amygdala are also correlated with ODD. It is presently unclear if there are neurobiological markers specific to ODD.

Table 46–1.
ICD-10 Diagnostic Criteria for Research

F91 Conduct Disorders

G1. There is a repetitive and persistent pattern of behavior, in which either the basic rights of others or major age appropriate societal norms or rules are violated, lasting at least 6 months, during which some of the following symptoms are present (see individual subcategories for rules or numbers of symptoms).

Note: The symptoms in 11, 13, 15, 16, 20, 21, and 23 need only have occurred once for the criterion to be fulfilled.

The individual:

1. Has unusually frequent or severe temper tantrums for his or her developmental level;
2. Often argues with adults;
3. Often actively refuses adults' requests or defies rules;
4. Often, apparently deliberately, does things that annoy other people;
5. Often blames others for his or her own mistakes or misbehavior;
6. Is often "touchy" or easily annoyed by others;
7. Is often angry or resentful;
8. Is often spiteful or vindictive;
9. Often lies or breaks promises to obtain goods or favors or to avoid obligations;
10. Frequently initiates physical fights (this does not include fights with siblings);
11. Has used a weapon that can cause serious physical harm to others (e.g., bat, brick, broken bottle, knife, gun);
12. Often stays out after dark despite parental prohibition (beginning before 13 years of age);
13. Exhibits physical cruelty to other people (e.g., ties up, cuts, or burns a victim);
14. Exhibits physical cruelty to animals;
15. Deliberately destroys the property of others (other than by fire setting);
16. Deliberately sets fires with a risk or intention of causing serious damage;
17. Steals objects of nontrivial value without confronting the victim, either within the home or outside (e.g., shoplifting, burglary, forgery);
18. Is frequently truant from school, beginning before 13 years of age;
19. Has run away from parental or parental surrogate home at least twice or has run away once for more than a single night (this does not include leaving to avoid physical or sexual abuse);
20. Commits a crime involving confrontation with the victim (including purse snatching, extortion, mugging);
21. Forces another person into sexual activity;
22. Frequently bullies others (e.g., deliberate infliction of pain or hurt including persistent intimidation, tormenting, or molestation);
23. Breaks into someone else's house, building, or car.

G2. The disorder does not meet the criteria for dissocial personality disorder, schizophrenia, manic episode, depressive episode, pervasive developmental disorders, or hyperkinetic disorder. (If criteria for emotional disorder are met, the diagnosis should be mixed disorder of conduct and emotions.)

It is recommended that the age of onset be specified:
Childhood-onset type: Onset of at least one conduct problem before the age of 10 years;
Adolescent-onset type: No conduct problems before the age of 10 years.

Specifications for possible subdivisions

In addition to these categorizations, it is recommended that cases be described in terms of their scores on three dimensions of disturbance:

Hyperactivity (inattentive, restless behavior);

Emotional disturbance (anxiety, depression, obsessionality, hypochondriasis); and

Severity of conduct disorder:

Mild: Few if any conduct problems are in excess of those required to make the diagnosis, and conduct problems cause only minor harm to others;

Moderate: The number of conduct problems and the effects on others are intermediate between "mild and severe";

Severe: There are many conduct problems in excess of those required to make the diagnosis, or the conduct problems cause considerable harm to others—e.g., severe physical injury, vandalism, or theft.

F91.0 Conduct Disorder Confined to the Family Context

The general criteria for conduct disorder (F91) must be met.

Three or more of the symptoms listed for F91 criterion G1 must be present, with at least three from items 9–23.

At least one of the symptoms from items 9–23 must have been present for at least 6 months.

Conduct disturbance must be limited to the family context.

F91.1 Unsocialized Conduct Disorder

The general criteria for conduct disorder (F91) must be met.

Three or more of the symptoms listed for F91 criterion G1 must be present, with at least three from items 9–23.

At least one of the symptoms from items 9–23 must have been present for at least 6 months.

There must be definitely poor relationships with the individual's peer group, as shown by isolation, rejection, or unpopularity and by a lack of lasting close reciprocal friendships.

F91.2 Socialized Conduct Disorder

The general criteria for conduct disorder (F91) must be met.

Three or more of the symptoms listed for F91 criterion G1 must be present, with at least three from items 9–23.

At least one of the symptoms from items 9–23 must have been present for at least 6 months.

Conduct disturbance must include settings outside the home or family context.

Peer relationships are within normal limits.

F91.3 Oppositional Defiant Disorder

The general criteria for conduct disorder (F91) must be met.

Four or more of the symptoms listed for F91 criterion G1 must be present, but with no more than two symptoms from items 9–23.

The symptoms in criterion B must be maladaptive and inconsistent with the developmental level.

At least four of the symptoms must have been present for at least 6 months.

F91.8 Other Conduct Disorders

F91.9 Conduct Disorder, Unspecified

Childhood:

- ► Behavioral disorder, NOS
- ► Conduct disorder, NOS

From World Health Organization. *The ICD-10 Classification of Mental and Behavioral Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization; 1992.

Table 46–2.

DSM-5 Diagnostic Criteria for Conduct Disorder

A. A repetitive and persistent pattern of behavior in which the basic rights of others or major age-appropriate societal norms or rules are violated, as manifested by the presence of three (or more) of the following 15 criteria in the past 12 months from any of the categories below, with at least one criterion present in the past 6 months.

Aggression to people or animals

1. Often bullies, threatens, or intimidates others.
2. Often initiates physical fights.
3. Has used a weapon that can cause serious physical harm to others (e.g., a bat, brick, broken bottle, knife, or gun).
4. Has been physically cruel to people.
5. Has been physically cruel to animals.
6. Has stolen while confronting a victim (e.g., mugging, purse snatching, extortion, armed robbery).
7. Has forced someone into sexual activity.

Destruction of property

8. Has deliberately engaged in fire setting with the intention of causing serious damage.
9. Has deliberately destroyed others' property (other than by fire setting).

Deceitfulness or theft

10. Has broken into someone else's house, building, or car.
11. Often lies to obtain goods or favors or to avoid obligations (i.e., "cons" others).
12. Has stolen items of nontrivial value without confronting a victim (e.g., shoplifting, but without breaking and entering; forgery).

Serious violations of rules

13. Often stays out at night despite parental prohibitions, beginning before age 13 years.
14. Has run away from home overnight at least twice while living in parental or parental surrogate home (or once without returning for a lengthy period).
15. Is often truant from school, beginning before age 13 years.

B. The disturbance in behavior causes clinically significant impairment in social, academic, or occupational functioning.

C. If the individual is age 18 years or older, criteria are not met for antisocial personality disorder.

Specify whether:

Childhood-onset type: Individuals show at least one symptom characteristic of conduct disorder prior to age 10 years.

Adolescent-onset type: Individuals show no symptom characteristic of conduct disorder prior to age 10 years.

Unspecified-onset type: Criteria for a diagnosis of conduct disorder are met, but there is not enough information available to determine whether the onset of the first symptom was before or after age 10 years.

Specify if:

With limited prosocial emotions: To qualify for this specifier, an individual must have displayed at least two of the following characteristics persistently over at least 12 months and in multiple relationships and settings. These characteristics reflect the individual's typical pattern of interpersonal and emotional functioning over this period and not just occasional occurrences in some situations. Thus, to assess the criteria for the specifier, multiple information sources are necessary. In addition to the individual's self-report, it is necessary to consider reports by others who have known the individual for extended periods of time (e.g., parents, teachers, coworkers, extended family members, peers).

Lack of remorse or guilt: Does not feel bad or guilty when he or she does something wrong (exclude remorse when expressed only when caught and/or facing punishment). The individual shows a general lack of concern about the negative consequences of his or her actions. For example, the

individual is not remorseful after hurting someone or does not care about the consequences of breaking rules.

Callous—lack of empathy: Disregards and is unconcerned about the feelings of others. The individual is described as cold and uncaring. The person appears more concerned about the effects of his or her actions on himself or herself, rather than their effects on others, even when they result in substantial harm to others.

Unconcerned about performance: Does not show concern about poor/problematic performance at school, at work, or in other important activities. The individual does not put forth the effort necessary to perform well, even when expectations are clear, and typically blames others for his or her poor performance.

Shallow or deficient affect: Does not express feelings or show emotions to others, except in ways that seem shallow, insincere, or superficial (e.g., actions contradict the emotion displayed; can turn emotions "on" or "off" quickly) or when emotional expressions are used for gain (e.g., emotions displayed to manipulate or intimidate others).

Specify current severity:

Mild: Few if any conduct problems in excess of those required to make the diagnosis are present and conduct problems cause only minor harm to others (e.g., lying, truancy, staying out after dark without permission, other rule breaking).

Moderate: The number of conduct problems and the effect on others are intermediate between those specified in "mild" and those in "severe" (e.g., stealing without confronting a victim, vandalism).

Severe: Many conduct problems in excess of those required to make the diagnosis are present, or conduct problems cause considerable harm to others (e.g., forced sex, physical cruelty, use of a weapon, stealing while confronting a victim, breaking and entering).

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Studies have estimated the genetic and environmental contributions to oppositional defiant behavior from mother, father, and teacher reports using rating scales in twin samples. These studies find support for additive genetic, unique environmental, and shared environmental influences on oppositional behavior and for both boys and girls. These studies generally support the idea that vulnerability to oppositional and defiant behaviors has a genetic basis that interacts with the child's environment in complicated interactional ways across development to result in the expression of ODD as a disorder. This research supports the clinical observation that oppositional and disruptive behaviors frequently cluster in siblings and biological relatives of the ODD child proband.

Temperament is defined as the characteristic style of emotional and behavioral response of an individual in a variety of differing situations and to a variety of differing environmental stimuli. Temperament is thought to be a constitutional facet of child development that may be observed early in life and generally persists over time and across situations. As the child grows older, temperamental characteristics shade imperceptibly into personality traits. Individual differences in infant and child temperament have long been recognized. Temperament has heritable and biologically based, as well as environmentally mediated aspects. Temperamental characteristics are presumably derived from a combination of genetic,

intrauterine, central nervous system, and postnatal environmental factors. Although not rigid and fixed, an individual's temperament style is likely to be relatively constant over time, especially when interplay between the individual and his or her environment is relatively stable across development.

A cluster of temperament dimensions described as the "difficult temperament" has been shown to correlate with concurrent and future behavioral adjustment in children of varying ages and social and ethnic groups. Although many definitions exist, the difficult temperament is generally characterized by infant or young child qualities of overactivity, undercontrol, high intensity of responses, inattention, predominantly negative mood, and low adaptability to new situations. Cross-sectional studies document a relationship between difficult temperament and generalized psychopathology, including oppositional defiant and conduct problems in childhood, adolescence, and adulthood. Longitudinal studies assessing the relationship between difficult temperament initially assessed early in the preschool years and behavior problems later in life are also largely consistent in finding a relationship with increased risk for generalized psychopathology, including externalizing behavioral problems such as oppositional defiant and CD, delinquency, aggression, and antisocial behaviors. This risk appears for both boys and girls, but may be stronger for boys. It should be noted that difficult temperament by itself is a relatively weak predictor of the development of oppositional defiant and conduct problems. When a negative child temperament is combined with family dysfunction, marital conflict, poverty, upbringing in a high-crime neighborhood, or parental psychopathology, prediction for disruptive behavior disorder is stronger.

Other temperament dimensions with relevance for oppositional defiant and disruptive behaviors have been described. One is novelty or sensation seeking. These children are described as outgoing, actively seeking stimulation, impulsive, and lacking in anxiety and fear. Because they are driven to seek stimulation and lack fear, they may have a more difficult time learning from experience and inhibiting their behaviors than children who do not exhibit these temperamental qualities. These children may appear undercontrolled, defiant, and oppositional. Other temperamental dimensions associated with risk for development of ODD include low levels of prosocial behavior and impulsivity and a lack of persistence. Temperamentally undercontrolled children at age 3 years have been found at age 18 years to have personality traits characterized by impulsivity, danger seeking, and aggression, and at higher risk for more aggression, antisocial behaviors, conduct problems, and criminal involvement compared to youth low on these temperament dimensions. The temperament dimension of novelty seeking may have a genetic basis. In molecular genetic investigations, polymorphisms in dopamine D_2 and D_4

receptors are significantly associated with personality traits of novelty seeking in population-based studies.

Psychological. A very influential theory in the causation of oppositional defiant and conduct problems in children is coercion theory. Coercion theory posits that parents and children in families with behavior-disordered children behave in such a way that oppositional and conduct-problem behaviors are behaviorally reinforced. This theory integrates contextual variables, parent variables, and child variables into an integrated understanding of oppositional and antisocial children that has the support of much empirical evidence. Coercion theory focuses on learning theory, specifically negative reinforcement of antisocial behavior. A behavior is negatively reinforced when it allows an individual to escape from a noxious, aversive, anxious, or unpleasant situation or task. In coercion theory, the negative behavior of one member of a parent–child dyad serves to terminate the ongoing negative behavior of the other, thereby negatively reinforcing the first member's original negative behavior. Over time, the child learns that aversive behavior "works" to solve social problems and may begin to generalize learned aversive and oppositional interpersonal behavior out of the home to interactions with authority figures at school and in the community.

Social learning theory posits that children learn by modeling the behaviors of caregivers and important others in their lives. To the extent that parents demonstrate coercive, aggressive, angry, and oppositional defiant behaviors and attitudes as a means of social problem solving, their children may learn that hostile behavior is an acceptable means of interacting with others.

Although the quality and pattern of an infant's attachment to a caregiver contributes to the development of individual personality and risk for psychopathology later in life, a link between attachment and oppositional defiant behaviors has yet to be conclusively demonstrated. The behavioral manifestations of insecure attachment, especially anxious–avoidant attachment, may be similar to the clinical characteristics of disruptive behavior disorders. In these cases, oppositional behavior is seen as a child's attempt to gain attention from an unresponsive parent.

Sociological. Ecological and environmental variables have long been noted to be associated with risk for the development of disruptive behavior disorders, including ODD and CD. Family characteristics, including poor family cohesion, family discord, inconsistent discipline practices, poor parental monitoring of the child's whereabouts, family stress, and poverty, all contribute to risk for child disruptive behaviors. Poverty appears to be a risk factor, although it does not appear to be a proximal or direct risk factor. Rather, low socioeconomic status is associated with disruptive

behavior disorders through family interactions and stress. Poor parents may have less time, patience, or energy to invest in parental monitoring and involvement with their children because of the increased daily hassles and stress associated with poverty. For similar reasons, parental psychopathology such as maternal depression is associated with child disruptive behaviors. Studies of the relation between maternal depressive symptomatology and the development of externalizing behavior problems in children find that maternal–child interaction quality partially mediates this association, even when the effects of social class are taken into account. Depression may adversely affect parenting ability such that increased negative interactions including irritability or hostility may occur between the child and caregivers, or parental monitoring and supervision of offspring may break down.

Diagnosis and Clinical Features

ODD is a disorder of emotional/behavioral regulation. A pattern of angry/irritable mood, argumentative/defiant behavior or vindictiveness lasting at least 6 months and occurring during interaction with at least one individual who is not a sibling, are the essential features of the disorder. A minimum of four symptoms total from any of the three ODD categories are required for diagnosis. The first ODD category is angry/irritable mood consisting of the symptoms—often losing temper, is often touchy or easily annoyed, and is often angry or resentful. The second category is argumentative/defiant behavior and encompasses the symptoms often argues with authority figures including parents and adults, often defies or refuses to comply with requests from authority figures or with rules, often deliberately annoys others, and often blames others for his/her mistakes. The final ODD category is vindictiveness including the symptom has been spiteful/vindictive at least twice in the past 6 months. The symptoms of ODD cause distress in the individual or others in his or her social environment and create disturbance in at least one domain of child functioning, including social, family, academic, or occupational functioning. Problems with angry/irritable mood include loss of temper, being easily annoyed, and frequent angry and resentful moods. Argumentative and defiant behaviors include arguing with adults, actively refusing to comply with requests or rules, and deliberately annoying others during which the individual often externalizes responsibility for his or her actions onto others in the environment. Vindictiveness refers to behaviors that suggest a petty desire to harm or humiliate others and/or that the individual seeks revenge on others perceived to have wronged him or her. Unlike in CD, in ODD the basic rights of others or age-appropriate societal rules are not violated.

To meet the threshold for clinical diagnosis, oppositional and defiant behaviors must be more severe than would be expected for the child's

developmental age. Criteria for distinguishing a behavior that is symptomatic from a behavior that is within normal limits for age and development include persistence and severity of symptoms. For children younger than five years of age, ODD behaviors should occur on most days for a period of at least 6 months. For older individuals, the behaviors should occur at a minimum frequency of once per week for a period of at least 6 months. For example, defiant behavior and temper tantrums are usual for a 2 year old and oppositional behavior is normative in early adolescence. For a diagnosis of ODD to be given, the behaviors are more extreme than is found for others of the same developmental age and cause daily impairment in functioning.

The diagnosis is not limited to a particular age group, but most commonly emerges between the ages of 4 and 8 years old. Studies have shown that ODD suggests a more circumscribed disturbance of lesser severity than other disruptive behavior diagnoses such as CD or ADHD. Problem behaviors are usually evident at home in relationships with those most familiar to the child and may not appear upon clinical examination in the office setting. In more severe cases, ODD symptoms may occur outside the home, usually in situations where authority figures set limits or place demands on the child. The ODD child generally does not see his or her behaviors as a problem and generally places blame on others for any problems. Thus, the clinical evaluation should include multiple reporters such as parents or teachers in addition to the child.

William is a 7 year old whose mother complained that he was frequently out of control at home, especially when asked to do simple chores such as helping to clean up his room or when limits were placed on him such as asking him to turn off the television at dinnertime. His mother reported that she frequently gave in to him in order to prevent more serious escalations of his temper, but no matter how many times he got "his way," his behaviors did not seem to improve around the house. Occasionally she reported getting "so mad" that she spanked him. Then his behavior would improve, but "only for a little while." He becomes easily frustrated and argumentative with his younger siblings. His parents divorced when William was 3 years old. Mother is a single parent with two younger siblings in the home. She does not work and supports the family on welfare. Social and parenting supports for the family are few. Mother reports that she divorced William's father because "he drank too much." She worries that William is "just like his dad ... he never listened either." After the divorce, mother was treated for an episode of depression. Although William appears to be progressing at school in the first grade, he struggles with reading and has few friends. Teachers describe him as somewhat withdrawn in the classroom. However, with peers he is described as bossy and frequently interrupting of others, often leading to conflicts that require teacher intervention. William is in good physical health, with normal vision, hearing, and physical development. Mother described him as an irritable infant who was frequently demanding of her attention.

Pathology and Laboratory Examination

No laboratory investigations are available for the diagnosis of ODD. The disorder is diagnosed on criteria established by clinical examination and by history. Several validated rating scales are available for oppositional and

defiant behaviors. These document symptom severity relative to same age and gender peers. Rating scales may assist in treatment planning by identifying specific targets for interventions and by documenting baseline disorder severity and subsequent response to treatment. Rating scales to measure the degree of functional impairment caused by symptoms are also available. These scales may assist the clinician in establishing that the oppositional and defiant symptoms exhibited by a particular child exceed those of children of the same gender and of similar developmental age, cause daily impairment, and thus help in establishing that a diagnosis of ODD exists.

Differential Diagnosis

ODD must be distinguished from disruptive behaviors that occur as part of normal development. Generally, ODD is distinguished from normal oppositional behaviors by its severity, impairment, and persistence across development. Because of the need to distinguish the disorder from disruptive behaviors occurring normally as part of development, the reliability and validity of ODD has not yet been established for preschool children. ODD as a diagnosis generally becomes clear around 4 to 8 years old. New-onset oppositional and defiant behaviors in early adolescence may be due to the process of normal separation or individuation. Defiant and disruptive behaviors may occur after a stressor such as a traumatic event or the loss of a loved one. If the disturbance lasts less than 6 months, adjustment reaction is the more appropriate diagnosis. In persons with intellectual disability, a diagnosis of ODD is only given if the oppositional behaviors are markedly greater than is commonly observed among individuals of comparable age, gender, and severity of cognitive disability. Irritable, defiant, and oppositional behaviors are common associated features of mood disorders such as depression, dysthymia, or anxiety disorders such as obsessive-compulsive disorder or separation anxiety disorder in young children. Bipolar disorder presenting in those younger than 18 years is often characterized by severe aggression, disruptive behaviors, defiance of authority, oppositionality, and irritability. Oppositional behaviors may be a feature of psychotic illness in youngsters. As such, ODD is not diagnosed separately if the symptoms occur only during the course of a mood or psychotic disorder. Refusal to follow the directions of authority figures needs to be distinguished from disorders of hearing and language. Children with hearing loss or receptive language disorders may look oppositional on the basis of a failure to hear or understand spoken commands.

In children, the diagnosis of ADHD may frequently be accompanied by a diagnosis of ODD. Children and adolescents with ADHD are often disruptive as a result of the core ADHD symptoms of inattention to tasks requiring sustained vigilance and deficits in impulse control. The willful

defiance of children with ODD must be distinguished from the poor compliance of children with ADHD.

The disruptive behaviors of individuals with ODD are of a less severe nature than those youngsters with CD. ODD typically does not involve violation of the basic rights of others, violation of age-appropriate social norms, severe aggression toward people or animals, recurrent destruction of property, or a pattern of covert theft or deceit. In DSM-5 ODD is identified as a distinct diagnosis from CD. Research has shown that the three categories of ODD (i.e., angry irritable mood, argumentative/defiant behavior, and/or vindictiveness) identify symptom clusters that overlap yet are substantially separate from CD. Longitudinal studies have established that the different ODD categories identify differing prognoses and risks across development. Thus, if criteria for both disorders are met, both diagnoses may be assigned to an individual.

DMDD shares with ODD, symptoms of negative mood, depression, irritability, and temper outbursts. However, as conceptualized in DSM-5, DMDD is a more severe disorder than ODD. In DMDD the severity, frequency, and persistence of aggression, irritability, and temper outbursts lead to more severe symptoms, greater impairment, and a worse prognosis than for ODD. Thus, if criteria for both DMDD and ODD are met in an individual, DMDD is diagnosed as the more severe condition and ODD is not diagnosed.

Comorbid Conditions

Children and adolescents with ODD appear to have higher rates of comorbidity with other psychiatric disorders relative to psychiatric comparison subjects. In addition, significantly greater family dysfunction and social impairments are found in youths with ODD compared to psychiatric controls. Population-based studies find that in cases of ODD, 14 percent of children have comorbid ADHD, 14 percent have accompanying anxiety disorders, and 9 percent have depressive disorders. Learning impairments and language disorders are also common in ODD youngsters. In studies ascertaining comorbidity in clinically referred samples, rates are much higher. The major comorbidity issue in clinically referred ODD children is with ADHD. Studies have shown that 60 to 80 percent of ODD youngsters assessed in clinics may meet criteria for ADHD, and about 50 percent of ADHD elementary school children will meet diagnostic criteria for ODD. Approximately 15 to 20 percent of clinically referred children and adolescents with ODD have comorbid depression, bipolar disorder, anxiety disorders, or language disorders. Concurrent substance abuse and cigarette use should be considered especially in teenagers with oppositional defiant disorder and/or ADHD.

Course and Prognosis

The course and prognosis of ODD varies by the age of onset of ODD, by symptom severity, with the presence of comorbid psychiatric disorders, and by the predominant ODD symptom cluster present. ODD often precedes the development of CD. An earlier age of onset of the disorder and more severe baseline ODD symptoms convey a poorer prognosis in terms of risk for progression to CD. The combination of early-onset ADHD and ODD also confers more risk of progression to CD. Studies have demonstrated that children with an earlier age of onset of ODD have a threefold increase in risk for the eventual development of CD, compared with children who meet criteria for ODD for the first time only later in the elementary school years or beginning in adolescence. About 30 percent of ODD children will eventually develop CD. Although the diagnosis of ODD is relatively stable over time, about two-thirds of children initially meeting diagnostic criteria will eventually exit from the diagnosis over a 3-year follow-up. Predictors of desistance from ODD remain unclear. Just as ODD confers risk for the eventual development of CD, CD confers risk for the development of antisocial personality disorder in adulthood. Longitudinal studies have shown that about 10 percent of ODD youngsters will grow up to meet antisocial personality disorder diagnostic criteria in adulthood. Preschool children who develop early-onset ODD are likely to exhibit additional psychiatric disorders later in childhood including ADHD, anxiety disorders, and mood disorders. Since longitudinal studies investigating the course and prognosis of ODD have been largely completed in male samples, it remains unclear if risk for progression to CDs is also true for females with early-onset ODD.

The course and prognosis of ODD also varies depending on which type of ODD symptoms are most predominant. The presence of argumentative, defiant, and vindictive symptoms carry most risk for the eventual development of CD. A preponderance of angry and irritable mood symptoms carries increased risk for the development of later emotional disorders such as anxiety and/or depression.

Treatment

Most experts consider primary and secondary prevention of at-risk children and families to be fundamental aspects of the treatment for ODD and other disruptive behavior disorders. Interventions are delivered in schools, in the community, and in the child's home. Key elements in prevention approaches to disruptive behavior disorders include treatment early in the child's life, preferably during preschool for at-risk youngsters, sustained treatment durations of longer than 2 consecutive years, intensive treatment occurring several times weekly, and simultaneously targeting multiple domains of the child's environment for therapeutic interventions including the child, parents, family, and school. Regular therapeutic home visitation to high-risk families as a preventive intervention has produced

positive results in preschool children at risk for ODD and disruptive behavior disorders.

Clinical approaches to the treatment of preschool and school-aged children with oppositional and defiant behavior problems in the family have typically been based on a psychosocial learning-based parent training model of behavioral intervention. Although many different types of parent management training (PMT) exist, the underlying rationale for all programs is based on the general view that child oppositional problem behaviors are inadvertently learned, developed, and sustained in the home by chronic and maladaptive parent–child interactions. During PMT, parents are trained to act as behavior change agents in the home. The child may or may not be directly involved in therapy. Therapy emphasizes the behavioral principles of a priori agreement on the most important oppositional behaviors to be targeted by therapy, developing a hierarchy of problem behaviors to be targeted for change, helping parents to learn not to inadvertently reinforce their child's maladaptive behaviors (for example, with attention), consistency of consequences for the child's defiant behavior across different caregivers and sustained across time, and consistent praise and reinforcers for the child when they demonstrate prosocial behavior. PMT can be implemented individually or in parenting groups. Older ODD children and adolescents may benefit from social problem-solving skills training. This type of therapy involves the child directly as a patient and teaches prosocial skills involved in social conflict resolution based on the principles of cognitive-behavioral therapy.

There exists little scientific evidence to suggest that ODD responds directly to medications. An exception is when ODD occurs in the context of ADHD. Psychopharmacological intervention studies with stimulants or noradrenergic agents, as monotherapy or in combination, have shown significant reductions in oppositional defiant symptoms when these occur in ADHD children.

CONDUCT DISORDER

Definition

The essential characteristic of CD is a repetitive and persistent pattern of behavior in which the basic rights of others or major age-appropriate societal norms or rules are repeatedly violated beginning in childhood or adolescence. These behaviors fall into four main groupings: aggression to people or animals, destruction of property, deceitfulness or theft, and serious rule violations. To meet criteria for diagnosis at least 3 of 15 criterion behaviors must be present during the past 12 months, with at least one behavior present in the past 6 months (Tables 46–1 and 46–2). The disturbance in behavior causes clinically significant impairment in social, academic or in occupational functioning. Severity is classified as mild,

moderate, or severe based on symptom frequency, severity, and harm to others. With the exception of requiring that the criteria for adult antisocial personality disorder are not present if the individual is aged 18 years or more, there are no exclusionary criteria. Youth with any diagnosable neuropsychiatric illness such as posttraumatic stress disorder (PTSD), depression, bipolar disorder, an organic condition such as a seizure disorder or head trauma, or schizophrenia can be additionally diagnosed with CD, if their behavior over the past year meets inclusion and exclusion criteria. As currently defined, CD is a very broad and heterogeneous diagnostic category that encompasses many diverse types of behaviors and vulnerabilities to different kinds of psychopathology.

Subtypes and Specifiers of Conduct Disorder

In hopes of reducing the heterogeneity inherent in the diagnostic criteria for CD, subtypes are recognized. DSM-5 recognizes three subtypes of CD: childhood-onset type, adolescent-onset type, and unspecified-onset type. A specifier identifies CD with or without limited prosocial emotions. Subtyping CD by age of onset in DSM-5 is retained from the DSM-IV-TR criteria.

Childhood-onset type refers to the presence of one criterion behavior prior to the age of 10 years. It has been known for some time that age of onset of antisocial behavior has important implications. Research finds that a relatively small group of children (generally boys) with onset of aggressive and antisocial behaviors in childhood (as contrasted with onset in adolescence) display many features suggestive of chronic psychopathology and neuropsychiatric disease. This early-onset CD subgroup has an especially poor prognosis over the lifespan. Compared to the later-onset group, children with early-onset CDs display more neuropsychological deficits, comorbid sustained attentional deficits, failure to inhibit impulsive responding, chronic learning failure and academic underachievement, more violent antisocial behaviors, aggressive and antisocial behaviors generally committed alone and not part of a larger social group such as a gang, a family history characterized by antisocial activities, and a higher risk for a more persistent and chronic course of antisocial activity across development. In epidemiological samples, this early-onset group accounts for a disproportionate amount of illegal activities, despite being only a tiny fraction of the total population of all youngsters with conduct-problem behaviors. For example, research has shown that the early-onset, persistent subtype of CD is characteristic of about 5 percent of males in the population.

In the adolescent-onset type, all the constituent behaviors for diagnosis begin after 10 years of age. The adolescent-onset group is generally comprised of girls as well as boys, generally expresses less violence in their antisocial activities, and may desist from conduct-disordered behaviors

when they achieve economic, psychological, and educational levels of adult independence in society. Although not all studies agree, research generally finds a steep decline in aggressive behavior (but not in nonaggressive antisocial behaviors) when CD begins at a later age of onset in adolescence. Youths with earlier ages of onset of conduct problems generally engage in more frequent and more variegated conduct problems, especially physical aggression, than youths with a later adolescent age of onset of conduct problems.

CD unspecified onset is used when criteria for a diagnosis of CD are satisfied but the age of onset of the first CD symptoms cannot be reliably determined.

DSM-5 includes a CD specifier assessing the presence of limited prosocial emotions. Characteristics include a lack of remorse or guilt when doing something wrong, a general lack of concern about the negative consequences of one's antisocial behaviors, a lack of empathy causing a general unconcern for the feelings of others even when they are harmed, little concern about poor performance at school, work, or in other important activities, a tendency to externalize blame for one's own poor performance onto others, and a shallow, insincere, or superficial affect. Researchers have identified these characteristics as callous–unemotional (CU) traits, a core dimension of psychopathic traits in youth. CU traits appear to be moderately heritable and help predict a subset of CD youth who tend to show more persistent CD symptoms, higher levels of instrumental and proactive aggression, deficits in the cognitive processing of facial emotional expressions of fear and distress in others, and more serious adult antisocial outcomes including elevated risk for adult antisocial personality disorder.

History

The clinical criteria for CD describe a type of individual with personality traits and behavior patterns that psychiatrists have long recognized and struggled to describe and to help. Indeed, descriptions of youths with conduct, delinquent, and antisocial symptoms antedate the very existence of the field of child and adolescent psychiatry. As early as 1837, Prichard used the term *moral insanity* to describe young adult patients with recurrent patterns of antisocial behavior who were not overtly psychotic. Benjamin Rush, the father of American psychiatry, described cases exhibiting "innate, preternatural moral depravity" and speculated that these patients had a defect in the "moral faculties of the mind." In the 19th century, scientific thought emphasized the inherent traits and physical defects (i.e., phrenology) that appeared to mark the antisocial individual.

Beginning in the early 20th century, a shift in thinking occurred as clinicians, social workers, pediatricians, welfare workers, and academics began to consider the impact of the environment and adverse individual or

social conditions on the young juvenile offender. Concern over keeping the youthful antisocial offender out of jail, combined with hope that advances in the new psychoanalytic and psychodynamic therapies that were just being introduced into America in the first two decades of the new century might undo the psychological damage that early childhood adverse experiences might cause, lead to new optimism in the treatment of youngsters with antisocial behaviors. This shift in thinking combined with hope for new treatments led to the beginning of the child guidance movement in the United States with the opening of the Juvenile Psychopathic Institute in Chicago in 1909, and the Judge Baker Clinic in Boston in 1917. The initial purpose of these clinics was to advise the courts on the psychology of juveniles under arrest. In 1923, Karl Menninger brought together a group of psychiatrists concerned with youthful offenders and started the American Orthopsychiatric Association. Thus, psychiatry's concerns for youths with conduct problems and antisocial behaviors have been intimately intertwined with shaping psychiatry's own development as a medical discipline over the past 100 years.

Today, the thinking about the etiology of CD has evolved into considerations about the interaction between vulnerability genes and adverse environmental rearing experiences in the shaping of the individual across development. This evolution in thinking is reflected in changes describing antisocial individuals in the DSM system. The diagnosis of sociopathic personality with antisocial reaction for adults was first codified in 1952 in DSM-I. In 1953, the American Academy of Child and Adolescent Psychiatry was founded. The new discipline was reflected in DSM-II published in 1968 that contained for the first time, a section on behavior disorders of childhood and adolescence. The diagnosis for children with conduct problems became unsocialized aggressive reaction of childhood (or adolescence). Descriptions of the new diagnosis were more pejorative than scientific, describing such youths as quarrelsome, vengeful, and disobedient. The diagnosis of CD first appeared in 1980 with the publication of DSM-III and contained four subtypes: aggressive, nonaggressive, socialized, and undersocialized. In DSM-III-R, published in 1987, three new subtypes of CD were introduced as solitary aggressive type, group type, and undifferentiated type. To meet the diagnosis of CD, 3 of 13 criteria present for the past 6 months were required. DSM-IV (1994) and DSM-IV-TR (2000) expanded the symptom criteria list to 15 items grouped into four categories including aggression to people and animals, destruction of property, deceitfulness or theft, and serious violation of rules, and required that three items be present over the past 12 months with at least one criterion present over the past 6 months. The subtypes were changed to a childhood-onset type, with onset of at least one criterion characteristic of CD prior to age 10 years, and an adolescent-onset type, with absence of any CD criteria prior to age 10 years. There also is an

unspecified-onset subtype for the diagnosis of CD when criteria are met but the age of onset cannot be determined. DSM-5 retained the age of onset subtypes and added a specifier identifying the presence of limited prosocial emotions as described above.

Comparative Nosology

Diagnostic criteria for CD in DSM-5 appear in Table 46–2 and for ICD-10 appear in Table 46–1. The two diagnostic systems have both similarities and differences. Although the criterion lists for DSM-5, ODD and CD contain similar items as the criterion lists for CD in ICD-10, the symptom lists are combined in ICD-10 and are separate in DSM-5. Furthermore ICD-10 lists 8 symptoms of CD as mild and 15 symptoms as severe. The differential weighting of symptoms influences diagnosis in ICD-10 where mild symptoms are more consistent with CD, ODD subtype (discussed above under ODD), and more severe symptoms are used to diagnose other subtypes of CD. This is consistent with the ICD-10 approach of combining the diagnoses of ODD and CD, while DSM-5 separates them into two distinct diagnoses. Another difference between the two diagnostic systems is the identification of CD subtypes. As mentioned above, DSM-5 identifies three subtypes of CD based on the age of onset and a limited prosocial emotions specifier. The current ICD system identifies four CD subtypes, including CD confined to the family context, unsocialized CD, socialized CD, and CD, ODD subtype. ICD-10 also identifies a CD other subtype and a CD unspecified category, similar to the other and unspecified disruptive, impulse control, and CD designations in DSM-5. As a consequence, the two diagnostic systems each identify a slightly different group of youngsters with CD. The clinical and research implications of these differences are presently somewhat unclear.

Epidemiology

The prevalence of CD as a categorical diagnosis is to be distinguished from the prevalence of delinquency, conduct problems, and or antisocial behaviors. The prevalence of CD varies with the type of diagnostic criteria used, sample characteristics, age, and gender. Past modifications between DSM-III and DSM-III-R made the diagnostic criteria more stringent and reduced the diagnostic prevalence of the disorder. Few changes in diagnostic criteria for CD occurred between DSM-III-R and DSM-5, resulting in similar diagnostic prevalence rates across these DSM editions. One-year DSM-5 prevalence rates of CD range between 2 percent and more than 10 percent with a median of 4 percent in the general population. Rates are higher in clinically referred youths and in juvenile detention populations. For adults older than 18 years, the estimated lifetime prevalence of CD in the United States is 9.5 percent, including 12.0 percent

for males and 7.1 percent for females. Studies assessing lifetime prevalence rates using retrospective recall methodologies of adults suggest that the median age of onset of CD in the United States is 11.6 years. Age of onset generally occurs between mid-childhood and mid-adolescence. Onset after adolescence is rare. The prevalence of CD is much greater in boys than girls. Lifetime prevalence of CD is associated with younger age, male gender, and low educational attainment, being separated or divorced as an adult, and residing in urban settings.

Etiology

In order to rationally guide prevention and treatment efforts for conduct-disordered youth, it is important to understand etiologic and causal mechanisms. Identification of causal mechanisms would facilitate understanding of why CD develops in some at-risk youth while in other at-risk youths it does not. However, causality in the development of CD and by extension antisocial behavior, aggression, and violence is very difficult to demonstrate. Most research in this area has focused on revealing correlates of antisocial behaviors, delinquency, conduct problems, and CD. Correlates have many different ways of fitting into causal models and knowing the correlation does not inform as to which causal explanation fits best. For example, correlations do not tell us what factors are specific for the development of CD and what factors are generalized risks for the development of global psychopathology, one aspect of which may be CD. Similarly, correlations do not inform as to what constitutes a direct risk factor for the development of CD and what constitutes a risk factor that is indirect and mediated through other mechanisms. Most correlational research has focused on single variables and sought to determine their significance in the development of CD. These are called univariate or main effects models. However, causal mechanisms for CD and other types of antisocial behaviors are multifaceted, complex, and transactional (i.e., multiple risk and protective factors that feed forward and backward to influence one another over the course of a child's development). The field is rapidly moving away from single variable, main effects perspectives in the etiology of CD and toward viewpoints that attempt to integrate person specific (genetic and developmental factors), family specific, and environmental factors (shared and nonshared) into more ecologically valid models of etiology.

Biological

HERITABLE FACTORS. Research has recognized that antisocial, aggressive, and conduct problems cluster in families to a greater extent than pure chance would suggest. Familial clustering can reflect shared environmental factors (i.e., factors that all siblings in a family share such as socioeconomic

status), genetic factors, or the interaction between genetic vulnerability factors and adverse environmental experiences (i.e., gene–environmental correlations and interactions). Shared environmental factors known to be associated with aggregation of conduct problems and antisocial behaviors in families include growing up in a family with antisocial parents, having a father with an externalizing behavior disorder, parental psychological distress including maternal depression, exposure to paternal alcoholism, residing in an urban environment, and low socioeconomic status.

The evidence for genetic influence on conduct problems utilizes twin and adoption methodologies. In adoption methodologies, biological children of antisocial parents who are adopted away at birth to nonantisocial adoptive parents are compared on rates of conduct problems in later childhood or adulthood to rates of antisocial behaviors in adopted away offspring of nonantisocial parents. These studies are consistent in finding a genetic influence for antisocial behavior, aggression, and adult crime. Other studies consistently document statistically significant associations between antisocial personality disorder and crime in biological parents and similar disorders in adopted away offspring. Identical twin studies demonstrate a substantial genetic component for both criminality and antisocial personality disorder in adults. Behavioral genetic research finds that genetic influences account for about 40 percent of the variation in antisocial behavior among individuals in the population and that the magnitude of genetic influences on CD are around 50 percent. Shared environmental factors account for about 11 percent of the variance and the remainder is accounted for by aspects of the environment unique to an individual (e.g., physical or sexual abuse). Significantly larger estimates of genetic effects are found for more severe conduct and antisocial problems in contrast to less severe behaviors and in antisocial behavior using clinic referred, as opposed to community, samples.

Examples of gene by environmental interactive effects in the genesis of CD, are studies investigating the gene for the enzyme monoamine oxidase type A (MAOA). The MAOA enzyme is involved in the metabolism of several central nervous system monoamine neurotransmitters important in the etiology of diverse psychiatric disorders. In boys who experienced sexual or physical abuse during development (a nonshared and individual unique environmental factor) polymorphisms in the MAOA gene explained variance in the eventual development of CD. Abused boys with a gene variant for low MAOA activity were twice as likely to develop CD as those with a gene variant coding for high MAOA activity. In boys without abuse, variations in MAOA were not related to the eventual development of CD. These studies demonstrate the importance of examining both environmental risk factors and genetic vulnerability factors and their interactions in the genesis of a heterogeneous condition such as CD.

PHYSIOLOGICAL FACTORS. Studies have shown that individuals with CD, antisocial behaviors, and aggression often demonstrate many signs of physiological underarousal and decreased sensitivity to stress. Arousal theory hypothesizes that low levels of arousal are related to conduct problems in two ways. First, reward processing theory suggests that low physiological arousal to environmental stimulation is a constant and unpleasant internal state for antisocial individuals. Thus, more sensory input is needed by such individuals for the maintenance of a euthymic affect. As such, risk-taking activity including aggression and conduct-problem behaviors are considered a mode of sensation seeking through which the individual seeks to normalize or optimize arousal levels and emotional state. Next, fearlessness theory postulates that low levels of physiological arousal are markers of decreased sensitivity to fear. Lack of fear can explain increased risk for conduct problems because such children may not generate anticipatory anxiety in response to inappropriate behaviors and/or may not appropriately respond with anxiety to environmental cues of punishment. Since socialization and learning to refrain from inappropriate behaviors are based on aversive conditioning (i.e., that inappropriate behaviors will be punished resulting in the child experiencing a state of internal negative affect; the child learns to avoid negative internal affect by refraining from inappropriate behaviors), lack of fear sensitivity may result in decreased individual responsivity to environmental cues of sanction and increased risk for conduct-problem behaviors. Research has shown low resting heart rates and heart rate reactivity to stressful situations in fearless individuals. A lack of fear and anxiety in childhood to cues of social punishment may contribute to disturbed fear conditioning, the preferential use of aggression to solve interpersonal conflicts, and a lack of moral and conscience development in antisocial individuals. Low resting heart rate has been shown to be a predictor of violence independent of other social risk factors in longitudinal studies. However, many individuals with a low resting heart rate do not become conduct disordered. When physiological factors such as low resting heart rate are combined with environmental risk factors such as having a teenage mother, being economically disadvantaged, or residing in a family where parents' divorce, the risk for antisocial behavior becomes much greater. Other methodologies to measure physiological arousal include the resting electroencephalogram and skin conductance activity. These studies also report a correlation between measures of low physiological arousal and increased risk for antisocial behavior.

NEUROENDOCRINOLOGIC FACTORS. The hypothalamic–pituitary–adrenal (HPA) axis is a major stress-regulating system in the body. The final product of the HPA axis is the release of the hormone cortisol into the bloodstream. Cortisol has many effects on biological arousal systems,

preparing the organism to respond to environmental threats. Consistent with arousal theory, studies have found low basal cortisol levels to be correlated with conduct problem and antisocial behaviors in clinically referred, at-risk and in nonreferred community samples of children and adolescents. Studies have also reported blunted cortisol reactivity in response to stress in antisocial youngsters. Longitudinal studies have assessed the effects of low cortisol levels at age 10 to 12 years on predicting aggression at age 15 to 17 years. Other investigations have found that children meeting diagnostic criteria for disruptive behavior disorders had lower cortisol levels than control children when exposed to frustrating or provoking events.

Since males are more physically aggressive than females in animal species as well as humans, attention has focused on testosterone, the most important male sex hormone. Although many preclinical studies in rodent and mammal species demonstrate that the presence of testosterone is a prerequisite for the expression of aggressive behaviors, human studies are equivocal. Human studies implicating testosterone in aggression are more robust when investigating postpubertal as compared with prepubertal youngsters.

NEUROTRANSMITTER FACTORS. There is support for a relationship between lowered catecholaminergic activity (DA and NE) in the CNS and increased risk for aggressive behavior, ODD, and CD. Catechol-*O*-methyltransferace (COMT) and MAOA are enzymes important in the metabolism of catecholamines. Lower levels of COMT and MAOA activity are associated with increased risk for disruptive behavior disorders. Genetic studies have reported that variations in the structure of the gene encoding the dopamine D4 receptor (DRD4) are associated with increased risk for aggressive behavior as well as attentional deficits and increased hyperactive–impulsive behavior. Altered CNS dopamine and norepinephrine activity suggest hyposensitivity to reward and diminished salience of environmental cues of punishment for inappropriate behaviors associated with deficits in reward and punishment processing associated with disruptive behavior disorders.

Serotonin is a ubiquitous central nervous system neurotransmitter that plays an important role in stabilizing information processing in neural systems, resulting in coherent behavioral, emotional, and cognitive outputs. Disturbed serotonin neurotransmission is associated with disinhibited, aggressive, and impulsive behavior in humans, perhaps by increasing sensitivity to cues that elicit irritation and aggression and by diminishing sensitivity to stimuli that signal punishment or consequence for one's actions. Many studies suggest that serotonin plays a role in aggression directed toward others and directed toward the self (e.g., suicide). In adults there is compelling evidence of an inverse relationship

between peripheral and central measures of serotonin and antisocial behaviors. In children with disruptive behavior disorders, several (but not all) studies provide evidence of a role for serotonin in the modulation of conduct-problem behaviors in children and adolescents. Age may be an important developmental factor in explaining differences between serotonin study results in adults and in children. Conduct-disordered behaviors may be associated with developmental abnormalities of the serotonin system. Decreased sensitivity of the serotonin 1A receptor (5-HT_{1A}) and antagonism of the 5-HT_{1B} receptor are associated with increased impulsive aggression in animal and human studies.

NEUROANATOMIC FACTORS. Brain imaging research has consistently implicated several central nervous system anatomic regions in the expression of impulsive aggression and antisocial behaviors. Techniques using positron emission tomography (PET), magnetic resonance imaging (MRI), and functional MRI (fMRI) in violent adult offender samples, including murderers, violent schizophrenics, drug-abusing sociopaths, and community subjects with antisocial personality disorder, have consistently found anatomic or functional deficits in anterior brain regions, particularly in the prefrontal cortex. These findings suggest that frontal lobe dysfunction may be associated with deficits in the top-down regulation of behavior and with increased risk for generalized antisocial and violent behavior. Findings from clinical case studies involving brain injury implicate the ventromedial and orbitofrontal subregions of the frontal cortex. Impaired functioning in the orbitofrontal cortex may be associated with compromised affective decision making in ODD and CD. Imaging studies of impulsively violent offenders with executive cognitive functioning deficits suggest involvement of the dorsolateral prefrontal cortex. Other studies have examined subcortical regions of interest. Emerging evidence implicates amygdala underfunctioning associated with psychopathic personality traits.

NEUROTOXIC FACTORS. Maternal prenatal smoking has been associated with adverse behavioral events in offspring. These include increased rates of antisocial behaviors including delinquency, impulsivity, CD, truancy, and attentional difficulties. Individuals with substantial prenatal exposure to alcohol have been found later to more frequently engage in impulsive aggression and antisocial behavior relative to comparison groups. There is evidence of a linear alcohol dose–response effect. The greater the amount of maternal alcohol consumed during fetal development, including even moderate drinking, the higher the risk for offspring conduct problems. Postnatal lead exposure may have a small effect on later risk for conduct problems. Birth complications such as fetal anoxia during labor are associated with risk for the subsequent development of conduct problems,

especially in disorganized families with few resources. These associations continue to retain their statistical significance despite controlling for multiple potentially confounding factors.

Psychological

COGNITIVE FACTORS. Cognitive control theory posits that the optimal and adaptive regulation of thought, emotion, and behavior to changing environmental demands occurs via intact prefrontal cortical functioning. Mental processes such as planning, sequencing appropriate actions, working memory, and inhibition of inappropriate emotional or behavioral responses allow for flexibility in adaptation to changing environmental demands and optimizes individual emotional and behavioral responses. These mental processes dependent on intact prefrontal cortical functioning are known as executive cognitive functions (ECFs). Because the prefrontal cortex matures with age, prefrontal cortex control over behavior increases across development from childhood to young adulthood when neuronal myelination is thought to be complete.

ECF deficits are associated with conduct-problem behavior independently of their correlation with ADHD. For example, boys with disruptive behavioral disorders show diminished response flexibility to changing parameters of reward testing. In the laboratory, they continue a previously rewarded response that is changed and now punished. In other words, ECF impairment in these individuals leads to response perseveration and diminished flexibility in responding to changing environmental circumstances. This is thought to reflect diminished top-down cognitive control over emotions and behaviors. Overall, in investigations of school-aged children, adolescents, and adults with a variety of antisocial and externalizing behaviors, the effect size of a variety of ECF measures for the groups with conduct-problem behaviors ranges between 0.39 and 0.62 (a moderate effect size).

Other problems in information processing are also correlated with CD. A consistent finding has been poor academic performance, especially delays in reading, in youngsters with CD. Many studies report a correlation between impaired verbal ability and conduct problems. On tests of intelligence, many conduct-disordered youth have a diminished verbal intelligence quotient (IQ) relative to performance IQ. Delays in speech and language development are not uncommon in youngsters with CD.

Youngsters with CD frequently fail to appropriately attend to and accurately encode social cues from others in the environment. For example, delinquent and aggressive boys demonstrate a bias to attribute hostile intentions to others, even when the expressed cues are neutral and nonthreatening. Boys with disruptive behavior disorders have problems encoding social cues and generating nonaggressive responses. These boys

often feel confident in their ability to carry out an aggressive response. Both boys and girls with CD demonstrate more affective empathy deficits toward others than those without CD. Since empathy involves feelings that are congruent with another's situation, deficits in empathy may increase the risk for aggressive responding. Research reports stronger findings for empathy deficits in conduct-disordered adolescents than in younger children.

PARENT–CHILD FACTORS. Many studies find that poor parenting practices are related to disruptive behaviors in offspring. Aspects of child rearing that have been consistently associated with children's disruptive behaviors include the degree of parental involvement with the child, poor monitoring of the child's whereabouts, not knowing what peers the child is associating with, coercive family process, and harsh and inconsistent discipline practices. Physically aggressive punishment of the child has been linked to child aggression. Low parental warmth and lack of involvement with the child has been linked to child oppositional and defiant behaviors. The effects of parenting practices on child behaviors are complex and reciprocal. Children influence parental responses to their own behaviors as much as parents influence behaviors in their offspring. Parents may interact with sons differently than daughters. The past decade has seen an increase in the complexity of psychological models of parent–child interactions and a growing recognition that contextual and genetic variables must be accounted for in any attempt to characterize the complicated relationships between parents and offspring behaviors.

TRAUMATIC FACTORS. Childhood physical abuse, sexual abuse, and a history of neglect are associated with an increased risk for the development of a wide variety of both internalizing symptoms, such as depression and anxiety, and externalizing behavior problems, such as aggression, delinquency, and conduct problems in children and adolescents. No specific psychiatric syndrome has been linked to childhood experiences of maltreatment. Rather, abuse appears to be a risk factor in the development of many different and varied psychological and behavioral problems in youth. It is also important to note that many childhood abuse victims do not later demonstrate many of these sequelae. Childhood physical abuse may be a particular risk factor for the later development of conduct problems and aggressive behavior. This association is found for both community and clinically referred youth. Investigations of the effects of different forms of early abuse on later development of crime and violence in nonreferred community samples found that physical abuse and neglect early in childhood are significantly associated with arrest for violent crime in adulthood. This association retains its statistical significance even after studies control for the confounding effects of age, gender, and race.

Childhood physical abuse, childhood neglect, or the combination of childhood physical abuse and neglect appear especially associated with increased arrest rates for violent offenses in adolescence and adulthood. Sexual abuse in childhood may also be associated with the development of aggressive behavior and conduct problems in childhood and adolescence. However, childhood sexual abuse is associated with a wide range of developmental outcomes, including children who are asymptomatic, and aggression is one of many symptom patterns that may occur in sexually maltreated youth. Characteristics of the abuse such as severity, duration, frequency, relationship between victim and perpetrator, and violence associated with sexual abuse have a bearing on outcome including risk for CD.

PARENTAL FACTORS. Children of parents who demonstrate psychopathology have higher rates of emotional and behavioral difficulties compared to children of parents who have no psychiatric difficulties. Early studies of children of alcoholics found an increase in the rates of externalizing behavior problems in offspring, including ADHD, CD, and ODD. Later studies either failed to replicate an association between parental alcoholism and child conduct problems or found an increased risk for generalized psychopathology in offspring that is not specific for aggression and conduct problems in particular. Children of opiate-dependent parents demonstrate increased rates of conduct problems in addition to increased rates of generalized psychopathology. Maternal depression has been linked to an increased risk of conduct problems in offspring in some studies. In studies of depression in multiple familial generations, a threefold elevated risk for conduct problems and disruptive behavior disorders was found for children of depressed parents compared to children of nondepressed parents. Parental antisocial personality disorder and criminality, especially in the father, appears to be a strong and more specific risk factor for the development of child antisocial behavior, aggression, and conduct problems. Children who have both parents criminally involved are at very high risk for the development of aggressive and antisocial behavior.

Sociological. Extrafamilial risk factors for aggression include those that occur in the neighborhood, community, or in association with a deviant peer group. Larger trends in society such as poverty, the easy availability of firearms, and the effects of media violence are also important.

PEER FACTORS. The association between peer factors and aggression encompasses two distinct yet interrelated phenomena: first, early rejection of the aggressive child by peers and then later association with a deviant peer group. In the early school years, aggression toward peers is highly

associated with peer social rejection and peer social group victimization of the aggressive child. The aggressive child is disliked and singled out by the social peer group and then excluded from participation in the peer group's activities. The antisocial child then begins to associate with a deviant peer group. Both rejection and victimization appear associated with increased risk for subsequent conduct problems, aggression, and antisocial behaviors in the aggressive child and adolescent. Indeed, the extent to which peer relationships are organized around deviance and antisocial behaviors is a strong predictor of continued antisocial behavior across adolescence and into young adulthood.

SOCIAL FACTORS. General social deprivation includes such factors as poverty, low socioeconomic class, unemployment, poor housing, residing in violent neighborhoods, and overcrowded living conditions. There exists a relationship between indices of general social deprivation and increased rates of aggression, crime, and antisocial behavior in youngsters. Low social class and poverty may be related to serious criminal offenses. On the level of the individual family and child, the effects of general environmental social deprivation on youth conduct problems may be mediated by individual parenting factors discussed above. As the number of stressors grows in an individual family and as economic resources become increasingly constrained, the ability of parents to supervise and monitor their offspring in a manner conducive to the promotion of prosocial behavior may become increasingly strained and limited. Ultimately, the effects of social deprivation on risk for youth CD may be mediated by the availability and skills of individual parents.

Cumulative Risk Factors. The importance of a cumulative risk model for onset and maintenance of conduct-problem behaviors in youth is widely supported. Risk factors do not occur in isolation from one another, are frequently interactive and reciprocal, and are often multiple and chronic in a child's life. Since most biological, psychological, and sociological risk factors are nonspecific and exert their effects on risk for antisocial behaviors indirectly, the specific type of risk factor appears less important for the development of antisocial behaviors than the total number of risk factors present, their chronicity, and their interactions across child development. Cumulative effects of multiple risks (i.e., susceptibility genes for antisocial behavior and violence, parental psychopathology, low socioeconomic status, adverse life events, poor parenting practices, heritable risk factors for neuropsychiatric illness, antisocial personality, or psychopathy) have been demonstrated to have a far more serious impact on developmental outcomes in children than any specific type of risk factor. Two models of cumulative and interactive effects of risk factors on antisocial propensity and on aggressive behavior are illustrated in Figures

46–1 and 46–2.

Protective Factors. Protective factors are variables that exert a buffering effect in youth at high risk for CD. Protective factors are more than the absence of risk factors. Protective factors confer some specific resiliency for prosocial behaviors in the face of risk. Protective factors can be considered under differing domains such as child factors, family factors, and extrafamilial factors. Protective factors in CD are outlined in Table 46–3. It should be noted that many of these protective factors are general and diminish risk for many different types of psychopathology, not specifically just for CD.

Risk and protective factors interact with one another in different ways to determine child behavioral outcomes. Multiplicative interaction occurs when the presence of one factor multiplies the effect of another factor. Potentiating interaction occurs when the presence of one factor potentates the effect of another. Catalytic transactional interactions occur when factors affect child behavioral outcomes only in the presence of other factors, but do not independently alter behavior. In the compensatory model, stressors, risks, and vulnerability factors combine additively. In the challenge model, competency and resilience are enhanced as long as stress does not exceed a threshold. Once the threshold is exceeded, competency is overwhelmed and breaks down. In the protective model, buffering factors modulate the impact of stressors, for example, by improving individual coping skills, adaptation, and the building of competence. Some have suggested that the interaction of risk and protective factors is a balance between the power of the person and the power of his or her social and physical environment, with the adaptational outcome depending on the relative balance at specific periods of development.

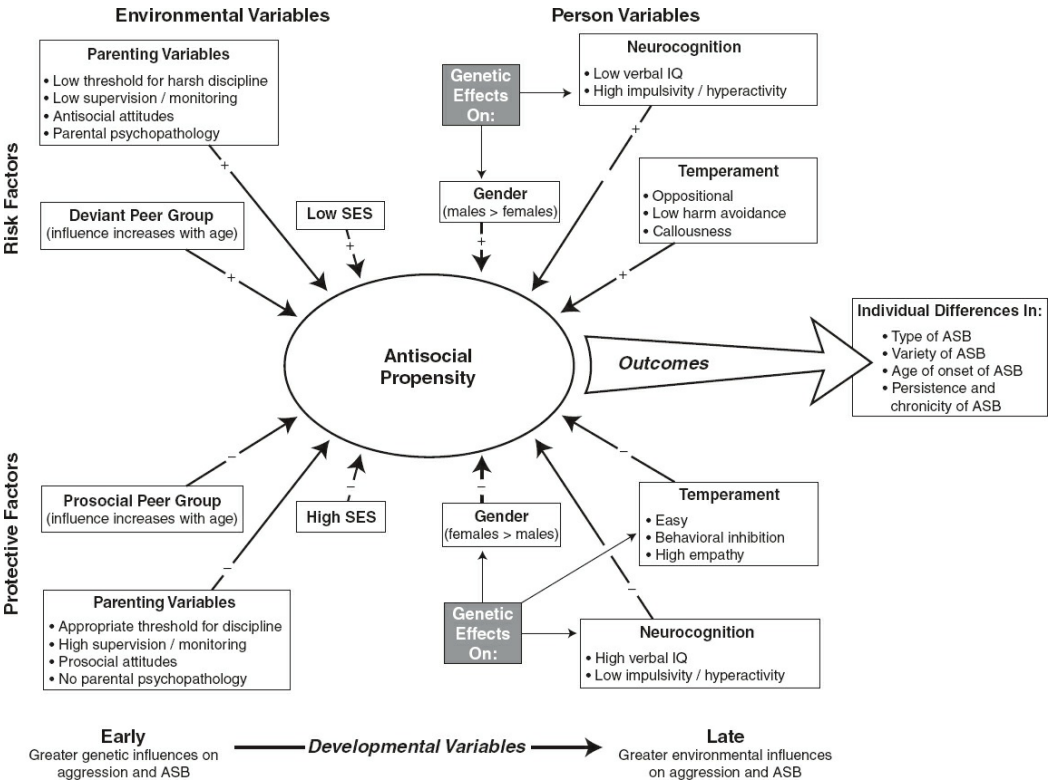

FIGURE 46–1. The interaction between selected risk and protective factors is illustrated across environmental and person variables as development proceeds. The resulting outcomes of complicated interactions between these different variables are individual differences in the expression of conduct problems and antisocial behaviors. IQ, intelligence quotient; ASB, antisocial behavior; SES, socioeconomic status. (From Connor DF. *Aggression and Antisocial Behavior in Children and Adolescents: Research and Treatment*. New York: Guilford; 2002, with permission.)

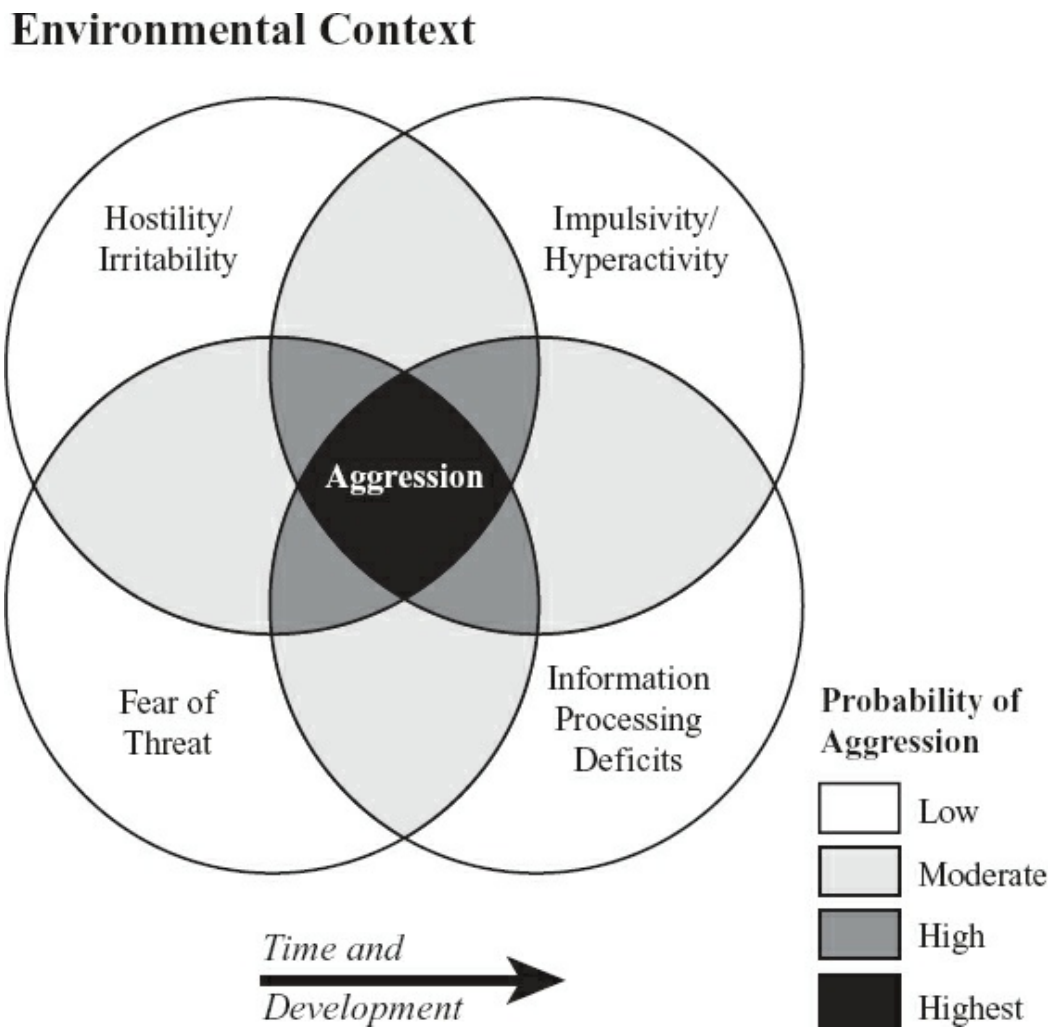

FIGURE 46–2. A descriptive interactive model of the relationship between four neuropsychological domains commonly found in children and adolescents with conduct and other psychiatric disorders and the probability of aggressive responding to environmental cues. Negative emotions such as hostility or irritability interact with impulsive disinhibition, exaggerated fears of threat, and/or information processing deficits to increase risk for aggression.

Diagnosis and Clinical Features

The diagnostic criteria for CD appears in Tables 46–1 and 46–2. The essential characteristics and associated features of CD have been discussed above. Assessment must involve multiple informants including parents and the youngster. Research has shown that self-report involving covert and overt conduct-disordered symptoms is reliable if done in a therapeutic manner, free of coercion or judgmental attitudes, and in a setting in which honesty is valued. The use of rating scales is encouraged and often used to document the severity and pervasiveness of CD symptoms. An important issue for clinicians in assessment is the comorbidity that frequently accompanies a diagnosis of CD. Assessment must evaluate the presence of comorbid psychiatric conditions such as ADHD, depression, bipolar disorder, substance-use disorders, developmental delay, learning

disabilities (e.g., reading, speech and language delays), or psychotic disorders that frequently accompany a diagnosis of CD.

Table 46–3.
Protective Factors

Child Factors	Family Factors	Extrafamilial Factors
Female gender	Good parent–child relations	External supports
Easy temperament	Highly educated parents	Prosocial friendships
Higher-intelligence quotient (IQ)	High-functioning parents	Availability of opportunities
Internal locus of control	Intact family structure	High socioeconomic status
High self-esteem		Safe neighborhoods
Academic competence		Safe schools
Social competence		
Competence in activities		
Absence of susceptibility genes		

Kyle is a 16 year old who was psychiatrically evaluated while he was serving a 3-year sentence in a juvenile correctional facility for property destruction and arson. He reported becoming angry with his father over a limit being set while he was mowing the family lawn. To retaliate against his father, he flipped a lit cigarette into his father's car that was parked in the driveway. Flammable objects placed on the seat quickly caught fire and destroyed the car. When asked about his judgment in doing this, Kyle replied, "How was I supposed to know there was stuff in the front seat that would burn?" This was his third arrest and second term in a correctional facility. His first arrest occurred at age 14 years for possession of drugs with intent to sell and his second arrest at age 16 for disorderly conduct. His family functioning was characterized by conflict between his parents and between his father and himself. His mother had a history of depression and his father a history of alcoholism and a poor work record with frequent periods of unemployment. Early-onset difficulties were noted when Kyle began to attend school and a diagnosis of ADHD was made when he was 6 years old. Several trials of stimulants were prescribed but discontinued because of parental fears that they would predispose him to later substance abuse. Problems with school performance, especially poor reading skills, were noted beginning in the third grade. By grade six, Kyle was frequently skipping school and he dropped out of school permanently in the ninth grade. He reported that he "liked prison" because "the other kids here are just like me."

Pathology and Laboratory Examination

No laboratory tests are useful or required for making the diagnosis of CD. All youngsters meeting diagnostic criteria for the disorder merit a complete physical examination to rule out any associated physical trauma associated with a high-risk lifestyle. Particular attention should be paid to the possibility of drug abuse and risk for sexually transmitted diseases. In females with CD, early-onset sexual activity may lead to the possibility of early pregnancy. Because of the association with learning disabilities, the presence of academic difficulties requires psychoeducational testing in youngsters with CD.

Differential Diagnosis and Comorbid Disorders

CD is a heterogeneous condition, and youngsters meeting diagnostic criteria for the disorder often meet criteria for other psychiatric diagnoses. Isolated instances of conduct-problem behaviors that do not cause substantial daily impairment need to be differentiated from a pattern of rule breaking and violations of the rights of others that may constitute a disorder. Early-onset CD needs to be differentiated from ADHD that can be marked by hyperactivity, impulsivity, risk taking, and problems with affect regulation. When ADHD and CD co-occur, the prognosis is generally worse than the prognosis for either disorder alone. Although some studies support the conceptualization of ADHD and CD as a single disorder in light of their common co-occurrence, other studies support the conclusion that they are separate but related disorders. There appears to be a reciprocal relationship between CD and substance-use disorders. The presence of one disorder is highly associated with the occurrence of the other disorder. In clinically referred adolescents, bipolar disorder and depression may be comorbid with CD. In youngsters with conduct problems and academic and school difficulties, the possibility of comorbid learning disabilities or developmental delay needs to be evaluated. The risk for development of psychiatric comorbidity in youngsters meeting criteria for CD appears highest in those with the early-onset CD subtype.

CD and ODD share with intermittent explosive disorder (IED) symptoms that bring the individual into conflict with adults and authority figures such as teachers and parents. However, IED does not include the nonaggressive features of CD and contains no element of premeditation in order to achieve some contingent goal such as money or social dominance. The individual meeting criteria for ODD generally exhibits behaviors that are less severe than the individual with conduct-disordered behaviors and ODD behaviors do not include overt aggression toward people or animals, destruction of property, or a pattern of theft and deceit. While ODD generally precedes CD in the course of development, many children with ODD do not go on to meet criteria for CD later in adolescence. In DSM-5, ODD and CD are conceptualized as separate disorders yet with many overlapping features. When diagnostic criteria are met for both ODD and CD, both diagnoses may be assigned.

Course and Prognosis

Longitudinal studies have shown that the risk for CD unfolds across development in an orderly fashion, with the onset of less severe behaviors generally preceding the onset of more serious aggressive and antisocial behaviors. This orderly progression is illustrated in Figure 46–3, demonstrating a normative and a deviant developmental pathway. In the normative pathway, there is a gradual decrease in oppositional, defiant, and aggressive behaviors during childhood, with a transient increase in delinquent behaviors in adolescence, eventually leading to a prosocial

outcome in young adulthood. In the deviant pathway, the gradual accumulation of risk factors leads to a steady worsening of aggressive and oppositional behaviors in childhood and increasingly serious acts of antisocial behavior leading to CD in adolescence.

Researchers have found support for other developmental pathways to serious CD in boys. The "overt pathway" progresses from minor acts of aggression to physical fighting and then violence. The "covert pathway" before age 15 progresses from minor covert behaviors to more serious acts of property damage such as fire setting and vandalism. The "authority conflict pathway" before age 12 progresses from stubborn behavior to defiance of authority such as truancy, running away, and staying out late at night. These pathways represent different developmental lines, and children may progress to CD along more than one pathway simultaneously. Comorbid psychiatric disorders begin to emerge with early-onset CD in childhood. These may first include anxiety or depression and then later progress to substance-use disorders. As noted above, comorbidity with ADHD portends a more serious course with risk for more severe symptoms and an early age of onset of antisocial behaviors. Once established, conduct problems and antisocial behaviors generally persist over time, with 50 to 90 percent of youngsters still meeting criteria for diagnosis after 4 years. Studies show that the severity of ODD in childhood is significantly associated with adolescent CD independent of childhood ADHD. Current research suggests that about 30 percent of boys with ODD will progress to CD, and about 40 percent of conduct-disordered adolescents will progress to eventual antisocial personality disorder. This implicates about 10 percent of ODD boys being at risk for antisocial personality disorder in adulthood. At present, it is unclear if a similar picture holds for girls meeting childhood criteria for ODD, and more research on risks for CD in girls is needed.

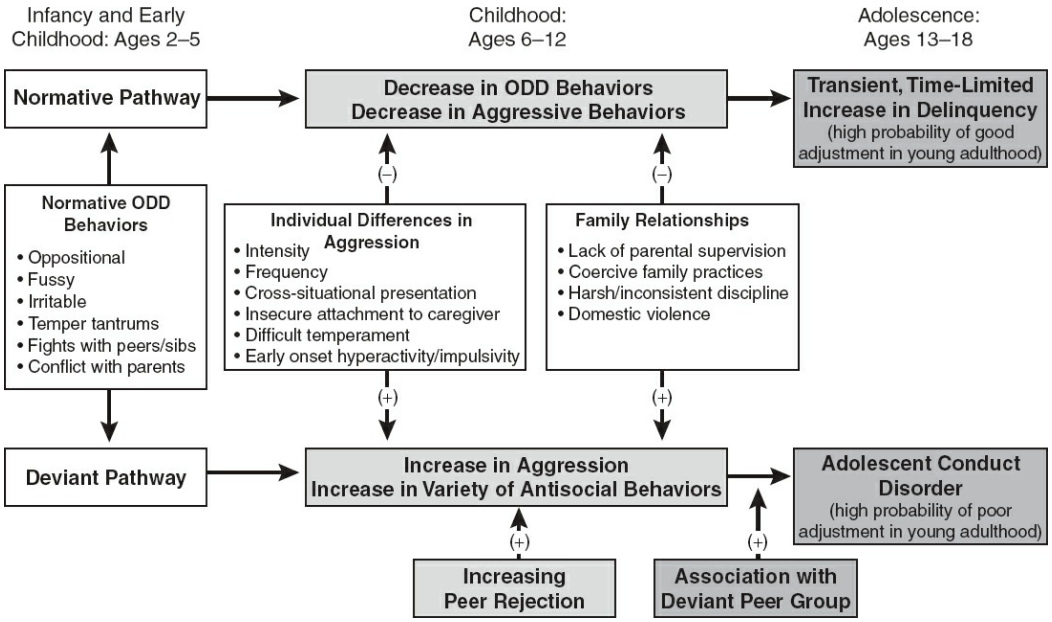

FIGURE 46–3. The developmental pathway leading from childhood oppositional and defiant behaviors to adolescent conduct disorder illustrating the gradual accumulation of risk factors. ODD, oppositional defiant disorder. (From Connor DF. *Aggression and Antisocial Behavior in Children and Adolescents: Research and Treatment*. New York: Guilford; 2002, with permission.)

Treatment

Reviews of treatment research conducted in the 1960s, 1970s, and 1980s documented a lack of treatment efficacy for aggressive children with antisocial behaviors and conduct problems. Many clinicians wondered whether it was at all possible to intervene effectively for children and adolescents diagnosed with CD. Psychiatric interventions largely based on psychoanalytic principles and psychological treatments given singly and sequentially in isolation did not seem to work for these youngsters. Early research was yet to elucidate the importance of developmental psychopathology in the etiology of early-onset CD, nor did clinicians appreciate the need for multimodal interventions applied concurrently as opposed to single interventions applied in isolation in the treatment of aggressive and antisocial youngsters. Although most contemporary clinicians would still agree that effective treatments are still lacking for CD and other forms of aggressive and antisocial behaviors in youths, clinicians now view such behaviors as partly, but not robustly, modifiable.

Most clinicians regard primary prevention as the gold standard in CD intervention. Results of programs specifically designed to prevent the onset of childhood aggression and antisocial behaviors show that they generally are effective, with long-term benefits in reducing antisocial behaviors, delinquency, CD, and substance-use initiation in at-risk youngsters over follow-up periods lasting several years. However, not all secondary prevention programs for conduct and antisocial problems have proved successful. Successful programs typically include: (1) multimodal

interventions that simultaneously target the child and family, parent supports, teacher involvement, and early childhood education (interventions with only a single focus appear much less effective); (2) intensive interventions generally delivered on a daily to weekly basis; (3) sufficient duration of intervention, generally 2 to 5 years or longer; (4) use of specific interventions empirically shown to be effective in ameliorating known psychosocial mechanisms that increase risk for conduct problems (e.g., coercive family process, harsh and inconsistent parental discipline, failure of parents to monitor the whereabouts of older children); (5) child and parent interventions that emphasize skill building, problem solving, and coping skills; (6) interventions that begin early in a child's life (between 0 and 6 years old); (7) individual case management; and (8) intensive collaboration among community, juvenile justice, school, family, and mental health professionals. These programs are expensive and generally require governmental financial support. Primary prevention programs that have these characteristics appear to reduce aggression and conduct problems during child development. For children and adolescents who demonstrate aggressive and violent behavior in the school setting, another approach is to emphasize violence prevention in the school. Studies reveal moderate violence prevention effects of various school programs.

Three well-studied mediators of early-onset aggression and conduct problems include coercive family processes, problematic parent practices, and deficient social skills. The first mechanism describes how parents and children train each other through the mechanism of mutual negative reinforcement to become oppositional and aggressive as a way of escaping aversive interpersonal situations. This learning may then generalize out of the home as the child develops and facilitate a developmental trajectory toward oppositional or CD. Other parenting practices associated with the development of antisocial behaviors include lack of parental monitoring of the child's whereabouts and friendships, low nurturing and engagement in the child's daily life, and harsh and inconsistent discipline. Research has described specific skill deficits from which aggressive children misinterpret neutral environmental stimuli as threatening and then react aggressively. As a result of this research, psychosocial treatments emphasize PMT and individual skill-building approaches. PMT teaches consistent parenting, positive and less harsh discipline practices, monitoring of the child, and positive feedback for the child. A variety of programs have been developed and tested against control conditions. In general, these parent-training programs appear effective in decreasing young children's oppositional, defiant, noncompliant, and aggressive behaviors. Several family therapy programs have been found to be promising, such as functional family therapy. Skills training approaches focus on the individual child. These programs emphasize social skills, problem-solving techniques, and anger

management strategies in reducing child aggression and conduct problems, and individual therapy approaches based on the principles of cognitive-behavioral therapy. Multifocused psychosocial programs for child conduct problems combine PMT, structural family therapies, and child skill-building treatments. The overall effect size of multifocused psychosocial treatment programs on preventing the worsening of conduct problems and aggression in at-risk youngsters is medium to large. Many of these programs demonstrate lasting effects after treatment is completed. Enhanced social skill building and conflict resolution skills in aggressive families may function as a mediator of treatment.

Multi-Systemic Treatment (MST) and Multidimensional Treatment Foster Care (MTFC) are programs that have been developed to treat older children and adolescents with established antisocial and aggressive behavior patterns necessitating juvenile justice involvement. These programs are designed as alternative placements to incarceration or psychiatric hospitalization and encompass a variety of psychoeducational, community, and mental health services delivered simultaneously to the adolescent and their family. MST and MTFC have demonstrated effectiveness in the treatment of severely disturbed families and violent adolescents, reporting decreased arrest rates ranging from 25 to 70 percent over 1- to 4-year outcomes. As such, these programs appear promising in the treatment of severely aggressive adolescents with chronic juvenile justice histories.

Psychopharmacologic treatments of CD are adjunctive to comprehensive psychosocial, community, and psychoeducational interventions. Pharmacological treatments generally do not focus on CD per se but on associated impulsivity, affective lability, on negative emotions such as fear and irritability, and on explosive aggression that may drive CD symptom frequency and severity. Psychopharmacological interventions are generally palliative and not curative in CD. Given high rates of diagnostic comorbidity with ADHD and affective disorders, medications targeting associated disorders may improve symptom severity, reduce daily impairment, and thus improve overall functioning of the conduct-disordered youngster. Meta-analyses have supported typical and atypical antipsychotics, mood stabilizers, and lithium (Eskalith) as being moderately effective in the treatment of explosive aggression in youngsters with CD. When antisocial symptoms arise in the context of ongoing ADHD, studies support the effectiveness of stimulants, atypical antipsychotics, and alpha-2A agonists such as guanfacine extended release.

FUTURE DIRECTIONS

There is much heterogeneity inherent in the diagnosis of CD. Since a DSM-5 diagnosis of CD requires the presence of any 3 of 15 possible symptoms, a grand total of 32,647 distinct symptom profiles meet criteria for a diagnosis

of CD. If one purpose of medical and psychiatric diagnosis is to facilitate understanding of a disorder's etiology and pathophysiology and thus arrive at a rational and scientifically based treatment plan, then there exists ongoing professional concern that CD fails as a useful and informative clinical diagnosis. DSM-5 provides for subtypes based on the age of onset, and a specifier identifying the presence of limited prosocial emotions, and there is evidence supporting the clinical utility of these subtypes in predicting the longitudinal course of CD. However, identification of more precise CD subtypes would be helpful in making the diagnosis more useful for clinicians. Other classification schemes with research support include distinctions between overt and covert antisocial symptoms, reactive, affective, defensive, impulsive (RADI) aggression, and predatory, instrumental, proactive (PIP) aggression; socialized versus unsocialized types of conduct problems; and the severity of relational aggression especially among females. Indeed, empirical support for these distinctions suggests that a more multidimensional classification system is needed to help best characterize CD.

REFERENCES

American Academy of Child and Adolescent Psychiatry. Practice parameter for the assessment and treatment of children and adolescents with oppositional defiant disorder. *J Am Acad Child Adolesc Psychiatry*. 2007;46:126–141.

Bakker MJ, Greven CU, Buitelaar JK, Glennon JC. Practitioner review: psychological treatments for children and adolescents with conduct disorder problems—a systematic review and meta-analysis [published online ahead of print August 8, 2016]. *J Child Psychol Psychiatry*. doi: 10.1111/jcpp.12590

Blair RJ, Veroude K, Buitelaar JK. Neuro-cognitive system dysfunction and symptom sets: a review of fMRI studies in youth with conduct problems [published online ahead of print October 26, 2016]. *NeurosciBiobehav Rev*. doi: 10.1016/j.neubiorev.2016.10.022

Bloomquist ML, Schnell SV. *Helping Children with Aggression and Conduct Problems: Best Practices for Intervention*. New York: Guilford; 2002.

Buitelaar JK, Smeets KC, Herpers P, Scheepers F, Glennon J, Rommelse NN. Conduct disorders. *Eur Child Adolesc Psychiatry*. 2013;22(Suppl 1):S49–S54.

*Caspi A, McClay J, Moffitt TE, et al. Role of genotype in the cycle of violence in maltreated children. *Science*. 2002;297:851–854.

Colman I, Murray J, Abbott RA, et al. Outcomes of conduct problems in adolescence: 40 year follow-up of national cohort. *BMJ*. 2009;338:a2981.

Comer JS, Chow C, Chan PT, Cooper-Vince C, Wilson LA. Psychosocial treatment efficacy for disruptive behavior problems in very young children: a meta-analytic examination. *J Am Acad Child Adolesc Psychiatry*. 2013;52:26–36.

Connor DF. *Aggression and Antisocial Behavior in Children and Adolescents: Research and Treatment*. New York: Guilford; 2002.

Connor DF, Carlson GA, Chang KD, et al. Juvenile maladaptive aggression: a review of

prevention, treatment, and service configuration and a proposed research agenda. *J Clin Psychiatry*. 2006;67:808–820.

Dishion TJ, Racer KH. Development of adult antisocial behavior. In: Beauchaine TP, Hinshaw SP, eds. *Child and Adolescent Psychopathology*. 2nd ed. Hoboken, NJ: John Wiley & Sons; 2013:453.

Eyberg SM, Nelson MM, Boggs SR. Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *J Clin Child Adolesc Psychol*. 2008;37:215–237.

*Fairchild G, van Goozen SH, Calder AJ, Goodyer IM. Research review: evaluating and reformulating the developmental taxonomic theory of antisocial behavior. *J Child Psychol Psychiatry*. 2013;54:924–940.

Foster EM, Jones DE; The Conduct Problems Prevention Research Group. The high costs of aggression: public expenditures resulting from conduct disorder. *Am J Pub Health*. 2005;95:1767–1772.

*Granic I, Patterson GR. Toward a comprehensive model of antisocial development: a dynamic systems approach. *Psychol Rev*. 2006;113:101–131.

Hambly JL, Khan S, McDermott B, Bor W, Haywood A. Pharmacotherapy of conduct disorder: challenges, options and future directions. *J Psychopharmacol*. 2016;30(10):967–975. doi: 10.1177/0269881116658985. Review.

Hudziak JJ, Derks EM, Althoff RR, Copeland W, Boomsma DI. The genetic and environmental contributions to oppositional defiant behavior: a multi-informant twin study. *J Am Acad Child Adolesc Psychiatry*. 2005;44:907–914.

Keenan K, Wakschlag LS, Danis B, et al. Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. *J Am Acad Child Adolesc Psychiatry*. 2007;46:457–468.

Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62:617–627.

Loeber R, Burke JD, Lahey BB, Winters A, Zera M. Oppositional defiant and conduct disorder: a review of the past 10 years, Part I. *J Am Acad Child Adolesc Psychiatry*. 2000;39:1468–1484.

Lovett BJ, Sheffield RA. Affective empathy deficits in aggressive children and adolescents: a critical review. *Clin Psychol Rev*. 2007;27:1–13.

*Matthys W, Vanderschuren LJ, Schutter DJ. The neurobiology of oppositional defiant disorder and conduct disorder: altered functioning in three mental domains. *Dev Psychopathol*. 2013;193–207.

Moffitt TE, Arseneault L, Jaffee SR, et al. Research review: DSM-V conduct disorder: research needs for an evidence base. *J Child Psychol Psychiatry*. 2008;49:3–33.

Moffitt TE, Caspi A, Rutter M, Silva PA. *Sex Differences in Antisocial Behaviour: Conduct Disorder, Delinquency, and Violence in the Dunedin Longitudinal Study*. Cambridge, UK: Cambridge University Press; 2001.

Nock MK, Kazdin AE, Hiripi E, Kessler RC. Prevalence, subtypes, and correlates of DSM-IV conduct disorder in the National Comorbidity Survey Replication. *Psychol Med*. 2006;36:699–710.

*Odgers CL, Caspi A, Broadbent JM, et al. Prediction of differential adult health burden by

conduct problem subtypes in males. *Arch Gen Psychiatry*. 2007;64:476–484.

Rowe R, Maughan B, Pickles A, Costello EJ, Angold A. The relationship between DSM-IV oppositional defiant disorder and conduct disorder: findings from the Great Smoky Mountains Study. *J Child Psychol Psychiatry*. 2002;43:365–373.

Rubia K, Halari R, Smith AB, et al. Dissociated functional brain abnormalities of inhibition in boys with pure conduct disorder and in boys with pure attention deficit hyperactivity disorder. *Am J Psychiatry*. 2008;165:889–897.

Van Goozen SH, Fairchild G, Snoek H, Harold GT. The evidence for a neurobiological model of childhood antisocial behavior. *Psychol Bull*. 2007;133:149–182.

Viding E, Fontaine NM, McCrory EJ. Antisocial behavior in children with and without callous-unemotional traits. *J R Soc Med*. 2012;105:195–200.

Waldman ID, Lahey BB. Oppositional defiant disorder, conduct disorder, and juvenile delinquency. In: Beauchaine TP, Hinshaw SP, eds. *Child and Adolescent Psychopathology*. 2nd ed. Hoboken, NJ: John Wiley & Sons; 2013:411.

Whelan YM, Stringaris A, Maughan B, Barker ED. Developmental continuity of oppositional defiant disorder subdimensions at ages 8, 10, and 13 years and their distinct psychiatric outcomes at age 16 years. *J Am Acad Child Adolesc Psychiatry*. 2013;52:961–969.

47

Feeding and Eating Disorders of Infancy and Early Childhood

IRENE CHATOOR, M.D., AND LOREDANA LUCARELLI, PSY.D.

INTRODUCTION

The term *feeding disorder* is commonly used to emphasize the dyadic nature of eating problems in infants and young children. The term describes a variety of conditions including food refusal, food selectivity, eating too little, food avoidance, fear of feeding, pica, and rumination. It is estimated that in the general population 25 to 45 percent of infants and young children have various feeding problems, and that severe feeding disorders associated with poor weight gain occur in 1 to 5 percent of pediatric hospital admissions and between 4 and 14 percent in ambulatory care. However, an exact evaluation of the prevalence or incidence rates of feeding disorders, and of particular types of feeding disorders, has not been published up to now because of the variability of definitions used and the lack of a diagnostic classification of feeding disorders shared by clinicians and researchers. Overall, several studies have found that early feeding problems not only cause distress to the young child but affect the whole family and often continue into later childhood and adolescence.

Definition

Criteria for Feeding Disorder of Infancy or Early Childhood in DSM-IV have been criticized for being too restrictive and not capturing many young children with feeding difficulties, and an alternative system in DC: 0–3R (2005) has often been used. This system includes six subtypes of feeding behavior disorder proposed by Chatoor (2002). Avoidant/Restrictive Food Intake Disorder (ARFID) in DSM-5 has replaced Feeding Disorder of Infancy or Early Childhood in DSM-IV. It has expanded the earlier more limited diagnostic criteria by introducing three main categories of feeding/eating problems, and in the text addresses the three remaining feeding disorders from DC: 0–3R. In addition, the ARFID criteria can be applied to individuals of all ages, consistent with the recognition of the continuity of psychopathology across the age range from early childhood to adulthood.

The new criteria of ARFID in DSM-5 describe three variants of an eating

or feeding disturbance that is characterized by, for example, apparent lack or interest in eating or food; avoidance based on the sensory characteristics of food; concern about aversive consequences of eating. These new DSM-5 criteria of ARFID with the accompanying text map on directly on three of the DC: 0–3R subtypes. The new subtype "apparent lack of interest in food or eating," has been described in DC: 0–3R as infantile anorexia (IA); the second variant "avoidance based on the sensory characteristics of food" describes children who select or avoid food in relation to the appearance, taste, texture, smell, or temperature of food and has been defined in DC: 0–3R as Sensory Food Aversions; the third variant "concern about aversive consequences of eating" is described in DC: 0–3R as Feeding Disorder Associated with Insults to the Gastrointestinal Tract and has been described by Chatoor as Posttraumatic Feeding Disorder (2002).

In addition, in the text there is allowance to diagnose ARFID in infants and children with the remaining three feeding disorders described in DC: 0–3R, those who have difficulty with state regulation, those who may not engage with the caretaker during feeding, or those whose feeding problems are associated with a medical condition when the disturbance of intake is beyond what can be explained by the physical condition.

To delineate feeding disorders from less severe feeding difficulties, DSM-5 has expanded the criteria of impairment to affect one or more of four areas of health or functioning:

- Significant weight loss (or failure to achieve expected weight gain or faltering growth in children)
- Significant nutritional deficiency
- Dependence on enteral feedings or oral nutritional supplements
- Marked interference with psychosocial functioning

History and Comparative Nosology

Prior to the inclusion of the diagnostic category of Feeding Disorder of Infancy or Early Childhood in the DSM-IV, there were no nationally defined criteria for the diagnosis of feeding disorders. In the absence of a recognized standard, different authors have used several labels for feeding problems: Picky eaters, choosy eaters, selective eaters, problem eaters, food refusal, feeding resistance, food phobia, food aversion, and dysphagia. However, none of these diagnostic labels has been operationalized. The field of feeding disorders has been further complicated by the fact that feeding disorder and failure to thrive (FTT) have been used interchangeably. The term *failure to thrive* refers to inadequate weight gain, based on standard growth charts. Historically, FTT has been dichotomized along organic and nonorganic axes. Whereas the former attributes growth failure to a medically diagnosable cause, the latter has commonly been thought to reflect maternal deprivation or neglect. A third

category was later added and represents a mixture of organic and environmental etiological factors of growth failure.

Concerns have been expressed about the usefulness of the label FTT as a "catch-all" diagnosis for all feeding disorders. First, the overlap between FTT and feeding disorders is confusing because not all infants with FTT have feeding disorders, and not all infants with feeding disorders have FTT. Second, Goldbloom pointed out that the term FTT is purely descriptive and does not address the process that caused the infant to grow inadequately. Although many authors agree that FTT should be viewed as a symptom rather than a diagnosis, FTT continues to be used as a diagnostic label particularly in the pediatric and psychological literature.

As described above, because of the limitations of the diagnostic criteria of Feeding Disorder of Infancy or Early Childhood in DSM-IV, an alternative classification of feeding disorders has been adopted by DC: 0–3R, which served as a template for the new diagnostic category of ARFID in DSM-5.

The original diagnostic criteria were published by Chatoor in 2002 and were modified with the help of the 2003 Task Force for Research Diagnostic Criteria for Infants and Preschool Children. They were further developed with the help of a work group that was supported by the American Psychiatric Association, and the revised criteria were published in 2007 in *Age and Gender Considerations in Psychiatric Diagnosis*. As explained above, the diagnostic criteria for these six feeding disorder subtypes were also included in the 2005 edition of the *Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood: Revised Edition* (DC: 0–3R).

The six feeding disorder subtypes are as follows: (1) feeding disorder of state regulation, (2) feeding disorder of caregiver–infant reciprocity, (3) infantile anorexia, (4) sensory food aversions, (5) feeding disorder associated with a concurrent medical condition, and (6) feeding disorder associated with insults to the gastrointestinal tract. The first four feeding disorders tend to start during specific developmental periods and are arranged accordingly, whereas posttraumatic feeding disorder and feeding disorder associated with a concurrent medical condition can occur at various stages of the child's development.

Since recently published research on feeding disorders has frequently used the classification in DC: 0–3R, the authors will discuss the various feeding disorders as listed in DSM-5 and will use the descriptions of the feeding disorders in DC: 0–3R as subtitles.

SPECIFIC FEEDING DISORDER SUBTYPES

Avoidant/Restrictive Food Intake Disorder

Subtype: apparent lack of interest in eating or food (DSM-5)

Infantile Anorexia (DC: 0–3R)

Diagnosis and Clinical Features. IA is characterized by poor appetite, lack of interest in eating, and poor growth in an otherwise active and playful child. It was first described by Chatoor in the 80s and was adopted in DC: 0–3R (2005) as one of the subtypes of Feeding Behavior Disorder. It usually begins in the first 3 years of life, most commonly between 9 and 18 months of age, when infants learn to walk, begin to talk, and are transitioned to spoon and self-feeding. The children are active and playful and appear to be interested in everything else but food and eating. They resist going into the highchair; they refuse to open their mouth to be fed and throw the food and feeding utensils. The parents become increasingly worried because of the children's poor food intake and growth, and they often resort to distracting, coaxing, and bribing to get the children to eat. Some parents become so anxious that they try to feed the children whenever they are willing to open their mouth, and some become so desperate that they resort to force feeding their children. Consequently, there is increasing conflict between the children and their parents, and mealtimes become miserable for the whole family.

Differential Diagnosis. IA needs to be differentiated from other feeding disorders that are characterized by food refusal. ARFID subtype: "concern about aversive consequences," usually has a more sudden onset of the food refusal following a traumatic event to the oropharynx or gastroesophageal tract, such as choking, severe gagging, or vomiting, or the insertion of feeding or endotracheal tubes. The infant with this feeding disorder may refuse all food, or only solid food, or only the bottle, depending on what mode of feeding the infant associates with the traumatic experience(s). Infants with ARFID subtype: "avoidance based on the sensory characteristics of food" or sensory food aversions (DC: 0–3R) consistently refuse specific foods but eat well if offered preferred foods.

Etiology and Course. Toddlers with IA have been found to have a higher level of physiological arousal and more difficulty down regulating their arousal. This may explain why they have so little awareness of hunger because they seem very excited by their play and enjoy talking, but they have difficulty turning away from these pleasurable activities. Studies from different centers have found significant correlations among difficult child temperament, irregular feeding and sleeping patterns, negative and willful behaviors by the toddlers, and mother–child conflict during feeding. At the same time, mothers' insecure attachment to their parents and mothers' drive for thinness and bulimia, anxiety, and depression also correlated significantly with mother–child conflict during feeding. Finally, mother–child conflict during feeding correlated strongly with the child's weight: the

more conflict between mother and child, the lower was the child's weight. However, correlations need to be used with caution to explain etiology, although the feeding histories of children with IA suggest that the lack of appetite and food refusal by the toddlers evoke severe anxiety and feelings of helplessness in the parents, who then try to engage the children in distractions and do anything to get the children to eat. The more insecure the parents feel, the more they allow the child to control the feeding situation, and the more ineffective the parents become in setting appropriate limits to the child's demanding behaviors. Consequently, parents and children become trapped in maladaptive interactional patterns.

Research further revealed that toddlers with IA on average performed in the normal range of cognitive development. However, the Mental Developmental Index (MDI) score of a control group of healthy eaters was 11 points higher than that of the group of toddlers with IA. Further analyses showed that not weight but mother–toddler interactions during feeding and play seemed to contribute significantly to these differences. Mother–toddler conflict and struggle for control during feeding and maternal intrusiveness during play were associated with lower cognitive performance of the toddlers.

A longitudinal study from Italy by Lucarelli and colleagues of children with IA who received only minimal intervention at the time of diagnosis, revealed some interesting findings. About 70 percent of the children with IA seen at follow-up during school age, showed ongoing eating problems and higher scores in anxiety, moodiness, aggressive behaviors, and somatic complaints when compared to the controls. These longitudinal data also showed significant correlations between continuing maternal eating and psychopathology, especially depression and anxiety, and internalizing problems of the children. This study on the natural course of IA highlights the importance of early diagnosis and specific interventions for children diagnosed with IA.

A follow-up study of 20 children at 6 months to 2 years at the time of diagnosis, who had completed a specific treatment described later, demonstrated that 17 children had changed considerably. The conflict in the parent–child relationship over the child's food intake was resolved. The children had learned to recognize hunger and increased their food intake, which had led to weight gain and improved growth. However, the parents reported that many of the children continued to lose their appetite when they were excited, if they had houseguests, if they had a birthday party, or if they traveled. Once they calmed themselves, they would regain their appetite and usually make up for the lost meals. There were three children who continued to struggle with their parents over their poor eating. In two cases, day care seemed to be too stimulating to allow these children to relax enough for eating. In the third case, the home situation was very

overcrowded and disorganized.

Clinical experience suggests that older children who have been untreated and who present for the first time with the diagnosis of IA during school age are usually thin and short for their age. However, their head size and their intellectual development appear to be normal for their chronological age. This was confirmed in a follow-up study by Chatoor and colleagues of children who were treated as toddlers for IA. It demonstrated that two-thirds of the children had outgrown their feeding disorder whereas one-third did not. Regardless of their malnutrition at the time of diagnosis or at follow-up, there was no correlation between their weight and intellectual functioning. Indeed, one-third of the children functioned in the superior range.

Treatment. This subsection describes a treatment that is based on a transactional model for the development of IA. According to this model, the infant's difficult temperament characteristics of emotional intensity, stubbornness, lack of hunger cues, irregular feeding, and sleeping patterns evoke conflicts over issues of control, autonomy, and dependency in a vulnerable parent who is anxious and insecure in regard to limit setting. The treatment addresses the infant's temperament and the parents' anxiety in regard to limit setting to facilitate internal regulation of eating of the infant.

1. The infant's temperament is discussed with the parents to help them understand that the infant's curiosity and emotional intensity interfere with the infant's awareness of hunger. Because the infant is very strong-willed and has learned that food refusal is a powerful tool with which to control the parents' attention, efforts by the parents to distract or entertain the infant to get him or her to eat may help temporarily but only perpetuate the problem by getting the infant more excited.
2. The parents' source of anxiety in regard to limit setting is explored. Some parents have been sensitized by previous losses of pregnancies, by infertility problems, or by the harshness and/or emotional unavailability of their parents. Understanding the parents' sensitivity in regard to limit setting is critical to helping them to be more effective in this area with their infant.
3. After this ground work, the parents are trained in respect to changing their behavior so as to facilitate internal regulation of eating by their infant. To boost the infant's biological rhythms of hunger and satiety, the parents are encouraged to feed the infant only at regular 3- to 4-hour intervals and not to offer anything but water in between these regular feedings. They are helped to understand that it is important to encourage the infant's self-feeding by commenting on the infant's feeding skills ("you get the spoon in your mouth all by yourself"), but they are not to comment on how much or how little the infant eats. It is helpful to give

the infant only a few bites at a time and after a while change to another food to keep the infant interested in eating and engaged with the parent without the parent becoming intrusive or distracting the infant. In addition, this helps the infant to continue eating until full and to learn to regulate his or her food intake by his or her internal signals of hunger and satiety. Eating should not be a performance for the parents. There should be no distractions during feedings, and no television or games, so that the infant is kept focused on eating. If the infant engages in inappropriate behaviors during feeding (climbing out of the high chair, throwing food or feeding utensils), the parents need to learn to use the "time-out" procedure. The time-out is used to accomplish two goals: To help the infant to stop behaviors that interfere with eating and to learn self-calming when upset because he or she cannot have his or her way.

This three-step intervention with the parents is best accomplished in double sessions, lasting 2 hours each, with the first two sessions being held within a few weeks of each other. The intensity of this brief intervention facilitates a close therapeutic alliance between the therapist and the parents and enables many parents to make major shifts in their interactions with their infants. Many of these parents need only a few follow-up sessions to deal with some specific questions that may arise. Frequently, after the infant has changed his or her eating pattern, the parents ask for help concerning the sleeping problems that many of these children have as well. However, there are some families who are more deeply entrenched in maladaptive interactional patterns with their infants, and some infants are so severely malnourished that more intensive interventions or hospitalization may be required. In particular, considering the associations between the severity of IA and risk factors such as the high mother–child interactional conflict, difficult child temperament, emotional and eating dysregulation in both, the children and of their mothers, the therapeutic approach should have a dual focus on the child and on the parents.

Rachel was 18 months old when she was referred by her pediatrician because of food refusal and malnutrition. The mother looked exhausted as she described how Rachel had never been a big eater from birth but that she stayed on the 25th percentile for weight and height on the growth chart until she was about 10 months of age. Her refusal to be fed by spoon became so strong that her intake of solid food became less and less. She depended primarily on bottles with milk for her nutrition. In spite of all of the parents' efforts to distract Rachel during feeding, to offer different foods, and to feed her whenever she would accept anything (including at night) and their desperate efforts to force food into her mouth, Rachel's food refusal had gotten worse, and she had fallen below the 5th percentile on the growth chart for weight and height. The mother admitted that trying to feed Rachel had taken over her entire life and that she felt exhausted and envious when she watched her friends' children eating without any effort on the part of the parents.

The observation of mother and infant during feeding and play revealed that Rachel was a thin little girl who seemed "all eyes," observing everything in the room, including the pictures on the wall and her image in the mirror. However, she paid little attention to the food. She ate some finger food, but refused to open her mouth and arched her back when her mother tried to

feed her by spoon. After a few minutes she appeared bored sitting in the high chair, threw the food, and indicated by grunting that she wanted to be taken out of the chair. Once she had her freedom, she ran around the room entertaining herself by exploring the furniture while her mother sat in her chair looking anxious and defeated.

The treatment addressed Rachel's special temperament characteristics, her intense curiosity and emotional intensity, her strong will, and her poor awareness of hunger, which seemed to have led to her increasing food refusal. The parents were helped to understand that they had a child with special temperament characteristics that required special parenting with special emphasis on limit setting. Specific instructions were discussed on how to handle mealtime, how to deal with Rachel's food refusal, and how to use "time-out" with a young child.

Although the mother struggled initially with how to set limits to Rachel's inappropriate behaviors—for example, stopping her from throwing food and keeping her in the high chair—the parents supported each other in implementing the new feeding instructions and were pleased with the change in Rachel's behavior. Once Rachel realized that her screaming did not change the rules, she stayed in the high chair without protest and became interested in feeding herself with the spoon. Within a few weeks, she seemed more aware of hunger, and her food intake increased accordingly. When seen for follow-up 1 year later, there was no longer conflict over Rachel's eating. Rachel had gained weight and grown in length, which brought her back to the 10th percentile on the growth chart.

Avoidant/Restrictive Food Intake Disorder

Subtype: avoidance based on the sensory characteristics of food (DSM-5)

Sensory Food Aversions (DC: 0–3R)

Diagnosis and Clinical Features. Sensory food aversions are common. However, whereas some children refuse to eat only a few specific foods, others may refuse most foods and take in only a limited diet of a few foods. The diagnosis of a feeding disorder should only be made if the food selectivity results in dietary deficiencies, and/or has led to oral motor and speech delay, and/or in older children is associated with social anxiety around eating and has led to marked interference of psychosocial functioning.

Within this disorder, food refusal is related to the taste, texture, temperature, or smell of particular foods. When the infant is fed specific foods, the infant's aversive reactions range from grimacing, to spitting the food out and wiping the tongue or mouth, to gagging and vomiting. After an initial aversive reaction, the infants usually refuse to continue eating that specific food. In addition, infants tend to generalize and refuse to even try foods of similar color, appearance, or smell. Occasionally, sensory food aversions become apparent in the newborn period when the infant is switched from breast milk to formula or from one formula to another. However, most commonly, infants begin to show aversive reactions when introduced to baby foods with different flavors and different textures or in the second year of life, during the introduction of different table foods. After the initial aversive reactions described earlier, some infants show such intense food refusal that they become limited to eating only a small variety of foods. They may refuse to eat whole food categories, such as

vegetables, fruits, or meats. In extreme cases they even refuse to eat a preferred food if it touched a "feared" food on the plate or if the food is not prepared by a specific restaurant or company. Children who refuse to eat vegetables and fruits usually develop vitamin deficiencies, and children who refuse to drink milk or eat meats may develop protein, zinc, and/or iron deficiencies. In addition, children who refuse to eat foods that require more chewing, such as meats, hard vegetables, or fruits, present with delay in oral motor development, which may be associated with articulation difficulties.

In addition to their sensory food aversions, many of these children experience hypersensitivities in other sensory areas as well. Parents frequently report that these children do not like to get their hands "messy," become distressed when asked to walk on grass or sand, and do not like to wear socks or shoes, certain fabrics, or labels on clothing. Many of these children are also hypersensitive to smells, and some may have difficulty with loud sounds.

Differential Diagnosis. This feeding disorder needs to be differentiated from oppositional food refusal that is common during the toddler years. Toddlers who like to exercise control over their parents may state that they do not want to eat a certain food and demand something else to eat. However, the food choices of oppositional toddlers are inconsistent and vary according to the mood of the toddler, whereas toddlers with sensory food aversions are usually consistent in their food choices, and they become distressed when the parents insist that they eat a certain food to which they are averse. Sensory food aversions also need to be differentiated from ARFID subtype: concern about aversive consequences of eating (DSM-5) or described as posttraumatic feeding disorder by Chatoor, which is characterized by the child's refusal to eat all solid food or to drink from the bottle, depending on what mode of eating the child associates with the traumatic event(s). However, some children with sensory food aversions may develop a posttraumatic feeding disorder if they have a strong aversive reaction to a food resulting in gagging or vomiting and/or the parents continue to insist that the child eat a food to which the child is averse. These children may become fearful of eating altogether.

Etiology and Course. There are no empirical studies that have examined the etiology of sensory food aversions in infants and toddlers. However, research with preschool and older children and with adults has related taste sensitivities to the bitter substance propylthiouracil, to strong food preferences, and to the number of taste buds on the individual's tongue (those with strong taste sensitivities, the "supertasters," have significantly more taste buds than "nontasters"). A study by Cooke and colleagues that examined food neophobia (the fear to try new foods, which

is one of the characteristics of sensory food aversions) in a large sample of twins aged 8 to 11 years estimated that 78 percent was heritable and another 22 percent of the variance was explained by nonshared environmental factors. Some studies have documented that parents and their biological children can share taste sensitivities, and various models of heritability—for example, incomplete penetrance, a two-locus model, and multilocus and multiallele models—have been suggested. In 2003, Kim and colleagues reported that specific polymorphisms on the taste receptor protein gene family Tas2r on chromosome 7 differentiated supertasters from nontasters.

In addition, certain aspects of the eating environment have been linked to the development of food preferences. For example, parents with strong taste sensitivities may offer a restricted range of food to their children, and limited exposure to a variety of foods may enhance the toddlers' food selectivity. On the other hand, if parents try to coax children or become forceful to get their children to try new foods, the children become oppositional and increasingly anxious and scared to try new foods. This pattern of forcing children to eat certain foods without adequate support to manage their anxiety may be a risk factor in sensory food aversions outcome. In support of this, longitudinal research by Nicholls and colleagues has shown that selective eating is associated with anxiety and obsessive-compulsive symptoms, social limitations, and school difficulties.

A longitudinal study of children's food preferences by Skinner and colleagues, who followed children from 2 to 3 years to 8 years of age, found that the number of foods liked at 8 years was predicted by the number liked at 4 years of age. Children were more likely to accept newly tasted foods between 2 and 4 years of age than during school-age years. School-age children with sensory food aversions may experience social anxiety and try to avoid social activities (parties, sleepovers, summer camp) that involve eating. Parents of children with sensory food aversions often report that they were unable to eat a variety of foods as young children and that they learned to eat new foods primarily during their adolescent or young adult years, when they wanted to fit in with their peers.

Treatment. Research by Birch and colleagues indicated that repeated exposures to new foods enhance young infants' acceptance not only of that food but also of similar foods. However, toddlers required more than 10 exposures of the same new foods until acceptance was reached. In addition, it is very difficult to get toddlers to try new foods if they are reluctant to do so. Birch and colleagues demonstrated that coercive techniques, such as threatening children to have them remain seated at the table until they finished eating everything on their plate or depriving them of certain privileges unless they tried the new food, have a significant negative effect. On the other hand, toddlers are very responsive to modeling by their

parents. They are more willing to try a new food if they observe their parents eating the food but not offering it to them.

No empirical studies have systematically examined how best to treat infants and young children with sensory food aversions, but clinical experience suggests that if infants or toddlers show strong aversive reactions to a new food, such as by spitting the food out or gagging or vomiting, the parents are best advised to give up on having the child eat this food. On the other hand, if the infant just grimaces, the parents may expose the infant to the new food at another time and pair a small amount of the aversive food with a preferred food. A gradual increase of the aversive food over several meals may allow the infant to get used to the new food.

Modeling eating new foods by the parents and waiting until the toddler asks to try a new food generally is more effective than putting new foods on the toddler's plate and requesting that the toddler try them. If the parents stay neutral as to whether the toddler likes the requested food, the toddler remains neutral as well and does not appear to become afraid of trying new foods. Once children fear to try new foods, their diet becomes more and more limited, and, by 3 years of age, most young children are not swayed by what their parents eat. Occasionally, preschool children may be willing to try new foods in a preschool setting, but more often they become anxious in social situations and try to avoid eating with others. On the other hand, preschool children are rich in phantasies and feeding the baby dolls with playfood can help to explore her fears and encourage them to first feed the baby doll and then eat themselves new foods.

John was brought for an assessment by the multidisciplinary feeding disorders team because he refused to eat vegetables, fruits, and meats. His parents had started to give him a multivitamin daily, but they were worried that he was missing other essential nutrients in his limited diet, and they were concerned that his articulation was so poor that nobody outside of the family could understand what he was saying. At the age of 2½ years, his daily diet consisted of Cheerios or pancakes in the morning and peanut butter sandwiches, some cheese (which had to be "American"), pasta, cake, crackers, ice cream, and milk during the rest of the day. His feeding problems started when he was introduced to baby food. He accepted baby food sweetened with fruit but grimaced when given baby food with vegetables and spat it out. He gagged when introduced to baby food with lumps of meat, and he refused to accept any baby food after that. At 9 months of age, John was offered soft table foods, which he ate only if he could finger-feed himself. He had another episode of gagging when introduced to green beans at 15 months of age, and consequently he refused to touch any green vegetables and fruits. He never tried any meats, and he became very distressed when his parents put any food on his plate that he did not want. When his parents insisted that he should at least try one bite of a new food before he could leave the table, John cried throughout the whole meal and ate nothing, not even his preferred foods. The parents had strong disagreements about how to handle John's food refusal. The father had similar food dislikes as a child and wanted to leave him alone. The mother was very concerned about John's health because of his limited diet, and she tried everything—bribing, coaxing, distracting, and threatening—just to get John to try new foods. Both parents experienced mealtime as very stressful and felt that John was getting increasingly rigid about his food likes and dislikes.

The observation of feeding revealed intense conflict between John and his mother regarding

the various foods that he refused to eat. The nutritional assessment showed that although John was taking in adequate calories, he was not getting enough zinc and iron from his diet and, without the multivitamin, his diet was deficient in vitamins. The oral motor assessment revealed that John used an immature chewing pattern that was felt to be due to his lack of practice with chewing crunchy and chewy foods. This oral motor delay was felt to account for his articulation problems.

The parents were instructed to obtain vitamins complete with zinc and iron to compensate for the nutrients missing in John's diet. They were also offered speech therapy for John to help with his articulation problems. In addition, the parents were given a few parent training sessions to help them to emotionally neutralize mealtime. After the mother's anxiety about John's health was alleviated, she was able to offer John his preferred foods without attempting to get him to eat new foods. The parents were instructed to eat their regular diet in front of John without offering him any of their food, and to give him only a small piece of their food if he were to ask for it. They were not to praise him if he liked the food nor become disappointed if he no longer wanted to have it. After she had stopped offering John any new food for about 2 months, John surprised his mother by asking her to have some of her peach. He liked it and asked for more on another day. Gradually, he began to eat soft fruits without the peel, and after several months, he tried chicken nuggets. His diet increased in variety very slowly, and when he entered preschool, he tried a number of foods at school that he had not touched at home. His diet became more complete. However, John was not able to eat the variety of foods the rest of his family enjoyed.

Avoidant/Restrictive Food Intake Disorder

Subtype: concern about aversive consequences of eating (DSM-5)
Feeding Disorder Associated with Insults to the Gastrointestinal Tract (DC: 0–3R)
Posttraumatic Feeding Disorder (Chatoor, 2007)

Clinical Features. This feeding disorder may present with total food refusal, refusal to swallow solid food, or refusal to drink from the bottle, depending on what mode of feeding the child associates with the traumatic event. The onset of the food refusal is frequently rather sudden and follows traumatic experiences that involve the oropharynx or gastrointestinal tract, such as severe gagging, choking, vomiting, insertion of feeding or endotracheal tubes, or force-feeding. It is estimated that posttraumatic feeding disorder may affect 4 percent of infants with gastroesophageal reflux who do not have a history of esophageal surgery and up to 40 percent of children who undergo esophageal surgery. Frequently, the parents are not aware that the event was so frightening to the child that it triggered the food refusal because other children who may undergo the same experience will not necessarily develop a posttraumatic feeding disorder. It appears that children who develop a posttraumatic feeding disorder are more prone to anxiety and/or are more sensitive to pain than the average child. Older children who refused to eat solid food after an incident of gagging or choking reported that they were afraid that the food would get stuck in their throat and choke them to death. Infants and young children express this fear through their behavior by crying in anticipation of being fed or when seeing the high chair, the bottle, or the spoon. Some

infants or children are brave enough to put the food in the mouth but cannot bring themselves to swallow it. They keep the food in their cheeks and/or spit it out. In severe cases, the infants refuse to feed all together. These infants are in acute danger of dehydration because the fear of eating appears to override any awareness of hunger or thirst.

Often the food refusal of these children is so intense that the parents react with severe anxiety to their child's food refusal. They worry about their child's nutrition. Frequently, they try to coax and distract the child, they offer various types of food, and they try to feed the child day and night without success. However, some very young infants who are afraid to drink from the bottle when they are awake may drink when they are asleep and not aware of what they are doing. However, if they wake up and see the bottle, they usually push it away and cry. In general, the food refusal of these infants poses an acute threat to their health, and they require acute intervention.

Differential Diagnosis. This feeding disorder needs to be differentiated from IA and from food refusal because of sensory food aversions. IA is characterized by an inconsistent pattern of food refusal, depending on the mood of the child. Anorectic children are not afraid to eat and can eat all types of food if they want to eat. Sensory food aversions usually involve foods with a certain taste, texture, temperature, and/or smell. Usually, the food refusal is more selective and not as global as seen in a posttraumatic feeding disorder. However, sometimes children with sensory food aversions develop a posttraumatic feeding disorder if they experience a strong aversive reaction or if they are force-fed aversive foods that trigger gagging and vomiting.

Etiology and Course. The observation of infants with a posttraumatic feeding disorder reveals that their food refusal seems to be caused by fear. However, no empirical studies have been done to explain why only some infants with distressing experiences to the oropharynx react with a posttraumatic feeding disorder. There are also no systematic longitudinal data on the course of this feeding disorder. However, individual case studies indicate that many infants and children get locked into their fears of eating. Some drink milk from the bottle and eat only pureed food until they reach school age, when the social embarrassment of their eating behavior causes the parents to seek help. In extreme cases, when the infants refuse all food, gastrostomy feedings have to be implemented. To survive, some of these children may depend on gastrostomy feedings for years.

A recent study by Lucarelli and colleagues showed that children with a posttraumatic feeding disorder were in the clinical range of the aggressive behavior dimension, which is mainly characterized by angry moods, temper

tantrums, stubbornness, child screams a lot, is easily frustrated, and wants a lot of attention. Taking into account the history of traumatic feeding experiences, the emotional dysregulation in these children may further contribute to their ongoing feeding problems. Moreover, this study found that their mothers had significantly higher scores in anxiety symptoms, and some of them received a diagnosis of anxiety disorder. These data suggest reciprocal influences between the dysregulation in emotional functioning of these children and the anxiety symptoms of their mothers. Actually, maternal anxiety may compromise their parenting capacity to moderate the impact of the child's traumatic experience and consequent emotional dysregulation.

Treatment. Because of the complexity of many of these cases, particularly those that result from trauma inflicted by unavoidable medical procedures (e.g., intubation, suctioning), a multidisciplinary team consisting of a pediatrician or gastroenterologist, nutritionist, occupational therapist, and a psychiatrist or psychologist is best equipped to meet the needs of these infants. In cases of total food refusal, the infant is in acute medical danger of dehydration and malnutrition, and intravenous fluids need to be given. It is important to assess whether the infant might be coaxed into accepting any form of food or whether nasogastric tube feedings need to be instituted. Unfortunately, the insertion of feeding tubes can intensify the fear of feeding, and, if this is the case, the infant is better served by a gastrostomy tube. In any case, the infant's nutritional requirements need to be addressed prior to the implementation of the psychiatric treatment.

The psychiatric treatment of patients who exhibit total or partial food refusal centers on the desensitization of the infant to the fear of eating. Outpatient treatment can only be carried out if the infant accepts bottle feedings or receives adequate nutrition through a nasogastric or gastrostomy tube. It is best if the child is given the tube feedings at night and that during the day there are scheduled oral feedings at 3- to 4-hour intervals with no feedings in between. If the child feeds from the bottle, the bottle feedings should follow the oral feedings in the high chair, and there should be no bottle feedings at other times, to make the child hungry at mealtimes. Behavioral techniques of positive reinforcement, praise and social interactions of food acceptance, and preferred toys or favorite videos for distraction are instituted in the initial management to overcome the child's fear of eating. Encouragement of self-feeding is frequently helpful for these infants to gain mastery over the anticipatory anxiety of eating. When the infant begins to accept food from the spoon, it is important to proceed slowly when advancing the texture of food to avoid any incident of gagging that may set back the infant. The behavioral manipulation of the infant's eating frequently leads to external regulation of eating in response

to the reinforcers. Once the infant has overcome the fear of eating, it is important to phase out the external reinforcers and to encourage eating according to internal signals of hunger and satiety. The same principles described with regard to facilitating internal regulation of eating in children with IA should also be applied in this final stage of the treatment.

Inpatient treatment becomes necessary if the child is unable to overcome the fear of eating and remains dependent on bottle or tube feedings. Inpatient treatment involves two major components: Appetite manipulation and contingency management. For children who have feeding tubes, manipulating calories and ensuring hydration through water or Pedialyte are used to allow the child to become hungry. Positive and negative reinforcements for food acceptance and food refusal are the behavioral interventions most commonly used. However, these techniques need to be adjusted to the individual child and require a skilled team to help the child to overcome the resistance to feeding and to learn to eat in response to hunger and fullness. Counseling and psychological intervention for parents is necessary to help them cope with the child's traumatic and stressful experiences and support their parenting capacities during the interventions for their child.

Robert was 7 months old when he was referred by a gastroenterologist for an urgent psychiatric evaluation because of severe crying when offered the bottle and refusal to drink unless he was almost fully asleep and seemed unaware of his feeding. The mother reported that Robert was an 8-lb, full-term baby who had a vigorous suck and was feeding well from birth. However, after a few weeks he started to spit up during and after feedings, and by 3 months he developed projectile vomiting. In spite of the vomiting, he gained weight until he was about 4 months of age, when he was referred to a gastroenterologist and diagnosed with gastroesophageal reflux. Several medications were tried, but the vomiting continued, and by 6 months of age Robert started to arch himself during feedings as if he were in pain. He would cry and refuse to continue with the feeding. Soon afterward, the mother noticed that he would start to cry at the sight of the bottle, and he refused to accept the bottle totally, as long as he was awake. To stave off starvation, the mother relied on feeding him during the night when he was asleep. His food intake greatly diminished, and he started to lose weight. The mother tried to introduce solid food, but Robert seemed to have difficulty moving the food back to his throat and spat most of it out. At the time Robert was referred, his mother was exhausted and frightened because of Robert's continued weight loss. A recent examination by the gastroenterologist had shown that the reflux was markedly improved and that Robert's ongoing feeding difficulties could not be explained by his medical condition.

The observation of mother and infant during feeding revealed that Robert became distressed the moment his mother positioned him for feeding, and he started to cry and arch himself as soon as he saw the bottle. His mother put the bottle away and tried to comfort him, but it took more than 10 minutes until he could settle in her arms. When the mother got the bottle and tried to feed him again, he cried even louder and would not be calmed until the bottle was out of his sight. Because it was obvious that Robert was too frightened to feed from the bottle, the mother tried to feed him in his infant seat with the spoon, but he batted the spoon away and did not want to have anything put in his mouth.

As the first step of treatment, the parents were made aware that infants as young as Robert can make associations between pain, vomiting, or other distress and feeding and consequently become frightened of feedings. The parents were instructed to get Robert on a schedule of regular naps and to feed him from the bottle only when he appeared almost asleep to prevent any further activation of his fears while looking at the bottle. After a successful attempt at the

hospital to feed Robert in the high chair with a spoon in his hand and lots of encouragement to put it in his mouth, the mother was able to use another spoon and put some food in Robert's mouth. Consequently, she was encouraged not to focus on the amount of food she was able to get into him, but to look at the feedings as practicing sessions to help Robert overcome his fear of feeding and to develop oral motor skills that would allow him to handle solid food.

In the first 2 weeks, Robert would cry and refuse to continue feeding if he woke up and saw the bottle. However, gradually he seemed to realize that it was safe to drink from the bottle. The mother reported that after 4 weeks, she could even start the bottle-feeding when Robert was still awake. Robert also progressed quickly from accepting very little food from the tip of the spoon to eating up to several spoonfuls. After 3 weeks, Robert was finishing a whole baby jar at one meal. When Robert was 9 months old, the mother called to report that Robert was feeding better and starting to gain weight. At 10 months of age, Robert was taking the bottle when awake, and he had learned to use the spoon and feed himself.

Avoidant/Restrictive Food Intake Disorder (DSM-5)

Subtype: Very young infants may present as being too sleepy, distressed, or agitated to feed

Feeding Disorder of State Regulation (DC: 0–3R)

Clinical Features. This feeding disorder begins in the postnatal period and is characterized by irregular, poor feedings and inadequate food intake by the infant. Infants with this feeding disorder exhibit difficulties in state regulation, which interfere with their ability to feed effectively. They have difficulty reaching and maintaining a state of calm alertness necessary for feeding. Some are too irritable and cry excessively and cannot calm themselves for feeding. Others are too sleepy and cannot wake up or stay awake long enough to feed adequately. The parent may be anxious or depressed or may present with more severe psychopathology. Mother–infant interactions during feeding are characterized by irritability and/or sleepiness of the infant, maternal tension, and poor engagement between mother and infant.

Feeding disorder of state regulation can be seen in infants without any organic problems, but it is frequently associated with prematurity or dysmaturity and gastrointestinal, cardiac, or pulmonary disease.

Differential Diagnosis. This feeding disorder needs to be differentiated from pure organic disorders that can interfere with the regulation of the state of the infant. As described earlier, organic problems can trigger the feeding disorder and should always be addressed first. The diagnosis of feeding disorder of state regulation should be considered if treatment of the organic problem—for example, cardiac or pulmonary disease—cannot resolve the infant's difficulty in state regulation.

Etiology and Course. There are no empirical studies that have explained why some infants have more difficulty in regulation of state than others. However, the foundation for the regulation of feedings, which is so closely linked to the regulation of state, is further developed in the first

months of life. Infants with feeding problems during the early months trigger anxiety in their mothers and frequently continue to have difficulty in self-regulation of their food intake during the transition to self-feeding in the second year of life. In a longitudinal study from Sweden, 70 percent of infants with feeding difficulties and food refusal in the first year of life were observed to have eating problems during childhood.

Treatment. Treatment needs to be individualized and can be directed toward the infant, toward the mother, and toward the mother–infant interaction. If the infant is too sleepy to feed and difficult to arouse, gentle infant massage to stimulate the baby may help to improve the infant's alertness and facilitate better feeding. If the infant is too irritable and fidgety to feed, reducing stimulation before and during feedings by taking the child to the quietest room in the house and swaddling the baby during feedings may help to settle the infant and facilitate better feedings. However, if the infant tires easily and is unable to take in adequate calories to grow in spite of all efforts to improve the feedings, nasogastric tube feedings may have to be considered to supplement oral feedings. On the other hand, if the mother is exhausted, overly anxious, or depressed, her difficulties have to be addressed to enable her to be more effective with her infant.

Videotaping the feedings and observing the tape together with the mother is frequently helpful to heighten the mother's awareness of the infant's cues and to explore alternative methods that she can use to facilitate the infant's regulation of state and ability to feed.

Thomas was 3 months old when he was referred for a psychiatric evaluation because of his feeding difficulties and poor weight gain since birth. His parents were college educated, and both had pursued their professional careers until Thomas was born. Although Thomas was full-term and weighed 7 lb at birth, he had difficulty drinking from the breast. When he was 4 weeks old, his mother had reluctantly switched him to bottle feedings because he was losing weight. Although his intake improved somewhat on bottle feedings, he gained weight very slowly and was still less than 8 lb at 3 months of age. His mother appeared tired and described how Thomas would drink only 1, 2, or 3 oz at a time, wiggle and cry, and refuse to continue with the feeding. After a few hours, however, he might cry again as if he were hungry. However, she could not settle him to feed, and he would continue to cry inconsolably. The mother described that she would attempt to feed him on an average of 10 to 15 times in a 24-hour period, that Thomas would cry a lot during the day and at night, and that everybody in the family was getting very little sleep.

The observation of mother–infant interactions during feeding and play revealed that Thomas was a very alert and wiggly baby who had difficulty settling in his mother's arms. While drinking from the bottle he would kick his feet and move around with his arms, and soon the nipple of the bottle would slip out of his mouth. This upset him, and he started crying. His mother appeared anxious and tried to restart him by changing his position in various ways, but this only agitated him more. After repeated unsuccessful attempts to continue the feeding, mother and baby appeared exhausted, and the mother gave up.

The assessment revealed that Thomas was a very active and excitable baby who had difficulty keeping calm during feedings. After reviewing the videotape with the mother, the therapist explored ways in which the mother could better facilitate calming during feedings.

Using a quiet corner in the house, swaddling Thomas in a blanket, and singing to him before starting the feeding were the most useful suggestions. Thomas stayed calm during feedings, was able to drink larger amounts of milk, and waited longer between feedings. This in turn relieved the mother's anxiety and helped both to have calmer interactions.

Avoidant/Restrictive Food Intake Disorder (DSM-5)

Subtype: Infants and young children may not engage with the primary caregiver during feeding

Feeding Disorder of Caregiver–Infant Reciprocity (DC: 0–3R)

Clinical Features. This feeding disorder is characterized by a lack of engagement between mother and infant, leading to inadequate food intake and growth failure of the infant. The onset of the growth failure is usually between 2 and 8 months of age. The infant shows lack of age-appropriate social responsivity—for example, lack of smiling response, lack of vocal reciprocity, lack of anticipatory reaching out when about to be picked up, and lack of molding and cuddling when held. The mother frequently presents with acute or chronic depression, drug or alcohol abuse, and/or high psychosocial stress, which appear to interfere with her care of the infant. Mother–infant interactions are characterized by a lack of mutual engagement and lack of pleasure in their relationship. The mother frequently appears detached and noncontingent in her responses to the infant's cues.

Differential Diagnosis. This feeding disorder needs to be differentiated from organic conditions that lead to a lack of weight gain and weakness of the infant. However, mother and infant usually show better mutual engagement, and the infant responds more readily to the examiner.

Etiology and Course. This feeding disorder has been referred to in the literature as "maternal deprivation" or "deprivation syndrome." It has been postulated that lack of emotional nurture and infrequent feedings lead to the growth failure and developmental delays of these infants. Characteristically, infants with this feeding disorder become engaging and gain weight when admitted to the hospital and are in the care of a nurturing nurse.

This feeding disorder has also been referred to as "nonorganic failure to thrive," and in the DSM-III, this disorder was defined as "Reactive Attachment Disorder of Infancy" associated with FTT. However, in the DSM-IV the same diagnosis was changed to encompass only problems in relatedness without growth failure. However, several studies of infants with the diagnosis of "nonorganic failure to thrive" have demonstrated very high rates of insecure attachment, ranging from 50 to 90 percent of the infants studied. However, not all infants with insecure attachment are failing to

thrive. It appears that only a severe degree of a disturbance in the attachment process of the infant to the caretaker, or what some authors describe as "nonattachment," results in the severe disturbances in relatedness and growth failure seen in infants with this feeding disorder.

Children with a feeding disorder of caregiver–infant reciprocity are at high risk for delayed cognitive and disturbed emotional development. When such children are followed up to school age, deficits in cognitive performance, behavioral organization, ego control, and ego resiliency and behavioral symptoms have been reported.

Treatment. Depending on the severity of the growth deficiency and the evidence of neglect, various interventions have been suggested. In cases of no evidence of deprivational behavior on the part of the mother, or if the parents have sought medical care in the past and have some support in the extended family or the community, an outpatient approach may be safe. However, if the growth deficiency is more severe, if there is serious hygiene neglect, if the mother is abusing drugs or alcohol, or if the mother lives a chaotic lifestyle, immediate hospitalization is indicated. During the hospitalization, it is critical to assign a warm and nurturing nurse to engage the infant during feedings. While the nutritional, emotional, and developmental rehabilitation of the infant takes place, a more in-depth evaluation of the mother and the mother–infant relationship needs to be addressed as well.

Many of the mothers have experienced unsatisfactory relationships with their caretakers when they were growing up, and consequently they are distrustful and avoidant of professionals. It is important to identify any positive behavior the infant shows toward the mother to see whether the mother has any potential to engage in a mutually satisfying relationship with her infant. In addition, the mother's support system and her ability to become engaged with a therapist need to be explored prior to returning the infant to her care. In some situations of severe neglect and associated abuse, Protective Services need to be involved, and the infant has to be placed in foster care.

Discharge from the hospital is a critical time when all services need to be in place to assure follow-through of the treatment plan. The treatment plan should be individualized and may range from home visits by a health care professional, to day care for the infant, to mother–infant psychotherapy, and to family therapy for the parents.

Anna was 8 months old when she was admitted to the hospital through the emergency room because of a respiratory tract infection and severe "failure to thrive." After she had recovered from the infection and a medical workup for her malnutrition was found to be negative, a psychiatric consultation was requested. After repeated unsuccessful attempts, the psychiatrist finally reached the parents and set up an appointment to obtain more history and to observe mother and infant during feeding. The parents explained that the father had difficulty

obtaining time off from his construction job and that the mother did not have any means of transportation to the hospital. The couple was not married because the father had not gotten his divorce from his first wife, but they lived together with Anna in one room in a house that they shared with nine other members of the father's extended family.

The mother, who became pregnant with Anna during her last year in high school, described Anna as a quiet baby and denied any feeding problems. However, when the mother temporarily left the room, the father turned to the therapist and complained that the mother did not give enough attention to Anna and that she would talk on the phone or read from a magazine while feeding the baby.

The observation of mother and infant during feeding revealed that the mother looked sad and preoccupied while she held the baby loosely on her lap. The infant lay in a "surrender position" with her arms raised besides her head, and she looked away while drinking from the bottle. However, when she looked at her mother, the mother looked away as if she could not stand to have eye contact with her infant.

Further evaluation of the mother revealed that she felt very lonely and depressed. She was unable to finish high school because of the pregnancy with Anna. Her family had abandoned her for becoming involved with a married man. She described the father's family as hostile toward her. The baby's father worked long hours and was hardly around. The telephone calls the mother made to her girlfriends from high school were the only relief in her unhappy life.

During the hospitalization, Anna gained weight, became more engaged, and rewarded her favorite nurse with frequent smiles. After the social worker had arranged for transportation for the mother, she came regularly for her appointments with the social worker and spent more time with Anna. While nurturing the mother and trying to help her work on her situation with her family and the father's family, the social worker also spent time with mother and infant during play. She would point out how the baby was making progress, how she was trying to get her mother's attention, and how she had that special smile for her mother.

After 4 weeks in the hospital, Anna seemed stronger, and her mother seemed committed to continue working with the social worker. The hospital staff felt comfortable about discharging Anna. After Anna had returned home with her mother, a visiting nurse came to the home twice a week to check on mother and infant, and the mother visited weekly with Anna at the hospital to see the social worker. After 6 months, the visiting nurse came less frequently, but the mother and Anna stayed in treatment with the social worker for another year, when both seemed engaged and comfortable with each other.

Avoidant/Restrictive Food Intake Disorder

ARFID can be diagnosed if the feeding disturbance may persist after being triggered by a medical condition and following resolution of the medical condition.

Feeding Disorder Associated with Concurrent Medical Condition (DC: 0–3R)

Clinical Features. Although respiratory problems or cardiac disease that may tire out the infant during feeding are easily recognized, some medical conditions are not readily diagnosed, and food refusal may be the leading symptom. For example, food allergies can be difficult to diagnose in infants, and "silent" gastroesophageal reflux without vomiting can be overlooked because vomiting is usually the leading symptom that alerts pediatricians to this condition. Typically, infants with reflux can drink 1 to 2 oz until the reflux becomes activated. However, once reflux occurs the infants experience distress and may arch, cry, and refuse to continue feeding. Some infants can be calmed and resume feeding, but others may become increasingly agitated while their caretakers continue to offer the

bottle or food. Infants with respiratory distress may feed for a while and take a few ounces until they tire and stop feeding. In general, the feeding difficulties of these infants lead to inadequate food intake, failure to gain weight, or loss of weight.

Differential Diagnosis. This feeding disorder needs to be differentiated from other feeding disorders characterized by food refusal, such as IA, sensory food aversions, and posttraumatic feeding disorder. What helps in understanding this feeding disorder is the observation of feeding, which reveals that the infant usually initiates feeding without difficulty and that mother and child engage in pleasant reciprocal interactions until the infant experiences distress and stops feeding.

Etiology and Course. FTT of infants and young children was initially believed to be caused by either organic or nonorganic causes (e.g., maternal deprivation). In 1981, Homer and Ludwig described a third category of FTT, which is caused by a combination of various organic and nonorganic problems. Since then, it has been accepted that organic conditions can be complicated by psychological difficulties and lead to severe feeding problems and growth failure. However, although many young infants suffer from gastroesophageal reflux, only less than one-third of them are reported to present with feeding difficulties and growth failure. Clinically, these infants appear particularly alert and sensitive. However, no empirical studies have been done to explore why some infants develop severe reactions to their reflux, whereas others hardly seem bothered by it.

Only a few studies have been reported on the follow-up of these infants. Fortunately, with therapy, up to 80 percent of children with gastroesophageal reflux improve within 12 months. For those with more serious feeding difficulties and inability to take in adequate amounts of food for growth, modern technology allows for gastric tube feedings over a long time to safeguard their survival. However, the regulation of oral feedings is often very difficult, and even after the medical condition has been successfully treated, it may take time until the child learns to eat independently.

Treatment. Because of the interaction of organic and psychological factors contributing to the severe feeding difficulties of these children, collaboration between the pediatrician or pediatric specialist and the psychiatrist is critical. Optimal medical treatment of the child's illness is necessary before psychological interventions can be successful. Direct observation of infants with their primary caregiver during feeding is most helpful to monitor how well the medical condition is responding to the treatment. In addition, the observations during feeding can guide the therapist in helping the parents to develop more appropriate strategies that

allow infants to calm instead of escalating their distress during feeding. However, in situations in which the medical illness cannot be adequately treated and the infant continues to experience distress during feedings, supplemental nutrition through nasogastric or gastrostomy tubes must be considered. Then the therapist must work with the parents to maintain the infant's oral feeding skills while most of the nutrition is given via tube feedings. The parents need to learn to feed the infant to the point of discomfort but not beyond it. Reviewing the videotape of the infant's feeding and discussing the infant's cues with the parents is usually helpful in sorting out how to best feed the infant. In general, these are very difficult cases that require individualized attention by an experienced multidisciplinary team.

Sarah was 6 months old when she was referred by her pediatrician to the Multidisciplinary Feeding Disorders Team because of severe food refusal and poor growth. The mother reported that since birth Sarah occasionally "spat up" a little during and after feedings but during the last 6 weeks she had become increasingly difficult to feed. Usually she took 1 or 2 oz of her formula and then started to arch and pushed the nipple out from her mouth. If her mother tried to continue feeding Sarah, she cried and became increasingly agitated until her mother picked her up and comforted her by holding her on the mother's shoulder and walking her. Sometimes after 10 to 20 minutes, her mother could resume feeding Sarah, but often, Sarah started to cry again after suckling from the bottle for a few minutes only. In spite of her mother's efforts to feed Sarah every 2 hours, Sarah drank only half of the formula that she was supposed to take in for her age, and in recent weeks, she had started to lose weight.

The observation of feeding revealed pleasant mother–infant interactions at the beginning of the feeding, but after a few minutes of suckling, Sarah pushed the nipple of the bottle out of her mouth, arched her back, and started to cry when her mother tried to get her back to feeding. The harder her mother tried, the more agitated Sarah became, and mother and baby looked very distressed and gave up feeding. These feeding interactions raised concerns that Sarah experienced gastroesophageal reflux that caused her pain and interfered with her feeding. Further medical testing confirmed the diagnosis of moderate gastroesophageal reflux, which was somewhat surprising to the pediatrician and the gastroenterologist because Sarah had not shown any vomiting, the most common symptom of reflux.

Sarah was treated with medications for reflux, and her food intake doubled quickly. However, if she drank more than 3 or 4 oz at a time, the symptoms recurred. However, once the mother understood the reason for Sarah's food refusal, she learned to ease Sarah's feeding difficulties by thickening the formula and feeding her smaller amounts at more frequent intervals. Gradually, Sarah tolerated larger amounts of formula at each feeding, and at 15 months of age, she was found to have outgrown the gastroesophageal reflux.

Rumination Disorder

Clinical Features. Rumination has been defined as repetitive regurgitation of gastric contents that are subsequently rechewed and then reswallowed or spit out. The regurgitation occurs very soon after a meal and tends to persist for 1 to 2 hours. The sequence of events leading to regurgitation has been documented by using combined multichannel intraluminal impedance and manometry. These events include an increase in intra-abdominal pressure identified in the stomach, which initiates the reflux, followed by peristaltic fore contractions, which clear the

intraesophageal content. Rumination has been described in infants, in individuals with development disabilities, and in children, adolescents, and adults with psychiatric or neurological disorders. The major proposed change for rumination disorder in the *Diagnostic and Statistical Manual for Mental Disorders,* Fifth Edition is its relocation from section, titled "Disorders Usually First Diagnosed in Infancy, Childhood, or Adolescence," to the newly proposed section, "Feeding and Eating Disorders." This change emphasizes that this disorder occurs across the age range, including adulthood. Minor changes have been introduced to the DSM-5 phrasing of diagnostic criteria for rumination disorder to improve clinical utility by ensuring applicability across the age range and removing some ambiguity inherent in the DSM-IV phrasing.

Some infants can be observed to put the thumb or the whole hand in the mouth, to suck the tongue rhythmically, or to arch the back to initiate regurgitation. However, some infants ruminate when left alone, after they have been placed in the crib to sleep. They are frequently found in a puddle of vomitus, which is then falsely attributed to some organic problem. The latter type of rumination is most commonly observed in infants who have received little emotional stimulation and who have learned to stimulate and soothe themselves through rumination. Although initially most of the regurgitated food is vomited, some infants gradually learn to hold more of the food in the mouth, to rechew and to reswallow it. "Experienced" ruminators are able to bring up the food through tongue movements and reswallow it without losing any of it. Their rumination can only be discovered by observing the movements of their cheeks and the foul oral odor. Social functioning may be significantly impaired due to the recognized social undesirability of this behavior.

Differential Diagnosis. The term *regurgitation* implies that the behavior is volitional, distinguishing it from vomiting or gastroesophageal reflux. The regurgitation is not associated with nausea and is not explained by an illness or medical condition. Individuals may cough, engage in tongue or abdominal muscle contractions or movements, or place fingers in the mouth to facilitate food being brought back up. Because in young children reflux and rumination frequently coexist, it is important to observe the infant in different situations to see the initiation and the rhythmic oral movements of the ruminatory activity. As in pica, the diagnosis of rumination disorder should be given in the context of another mental disorder only if the symptoms require additional clinical attention.

Epidemiology. Rumination is believed to be a rare disorder. It appears to occur more often in males than in females and also in individuals with mental retardation. There are a few reports of children, adolescents, and adults who were diagnosed with rumination disorder, and some had

associated eating disorders.

Etiology and Course. Various mechanisms of etiology have been proposed. Several authors have attributed rumination to separations or a disturbed mother–infant relationship—for example, lack of responsiveness or neglect by the mother secondary to stressful family relations or life events. Others have considered rumination as a symptom of gastroesophageal reflux. Several authors have postulated that rumination is a learned behavior that is maintained by special attention to the regurgitation. In more recent years, opiate receptor insensitivity or reduced endorphinergic transmission have been implicated in rumination.

A biopsychosocial model has been proposed by which rumination is seen along a continuum in which an infant may have gastrointestinal pathology and little psychopathology at one end of the spectrum, or the opposite, in which an infant may have no organic pathology and severe psychopathology in the mother–infant relationship at the other end of the spectrum. Frequently, vomiting secondary to gastroesophageal reflux or vomiting associated with an acute illness precedes the beginning of rumination. It appears that at some point the infant learns to initiate vomiting and to rechew the food to achieve relief of tension, to self-soothe, and/or to self-stimulate. It seems that in situations in which the infant fails to elicit or loses either caring attention or tension-relieving responses from the mother, rumination becomes a means of self-regulation. Once the infant has experienced rumination as a means of self-soothing or self-stimulation, the rumination develops into a habit that is difficult to break.

The onset of rumination is frequently in infancy, most commonly in the first year of life. However, rumination may start at any point in life. It is believed that in some individuals the disorder remits spontaneously. However, some infants suffer serious complications, such as electrolyte imbalance, dehydration, and weight loss, and in earlier reports mortality rates as high as 25 percent have been described.

Treatment. Depending on the theories of the etiology of rumination, different treatments have been suggested. These range from mechanical restraints as described in the early literature and surgical interventions (Nissen fundoplication) to prevent reflux, to behavioral and psychodynamic treatments.

Based on the assumption that rumination is a learned habit that is reinforced by the increased attention that is given to the regurgitation, unlearning by behavioral techniques has been suggested. These treatments include the use of aversive taste stimuli (lemon juice or hot sauce), oral hygiene, differential reinforcement of incompatible behaviors, and electric shock. Proponents of the assumption that rumination is based on an unsatisfactory mother–infant relationship have suggested a psychodynamic

approach. A mother substitute who feeds the infant while the mother is helped through psychotherapy has been proposed.

After a thorough evaluation that should include the assessment of the infant and the mother, the therapist needs to individualize the treatment. In severe situations, if the infant is malnourished and continues to lose most of the food through rumination, a jejunal tube may have to be inserted before any psychosocial treatment can begin. Frequently, a combination of psychodynamic and behavioral interventions to enhance the mother–infant relationship in general and to treat the symptom of rumination in particular is most effective.

Justin was 9 months old when he was referred by a gastroenterologist for a psychiatric evaluation because of concerns that he continued to vomit because of rumination. Justin was born full-term and had developed nicely until he was about 6 weeks old, when he began to vomit increasing amounts of milk during and after feedings. He was diagnosed with gastroesophageal reflux, which was treated with thickened feedings and medication. Justin responded well to the treatment; he stopped vomiting almost completely and gained weight adequately. Because Justin was doing so well, his mother decided to go back to work when Justin was 8 months old. She transitioned his care to a young woman who would come to the house during the mother's working hours. Justin started to vomit soon after his mother left the house. The vomiting seemed to increase from day to day in frequency and in intensity, and after 2 weeks of the mother's return to work, Justin vomited several times daily and was losing weight. He was seen by a gastroenterologist and, during the barium swallow, it was noticed that Justin put his hand in his mouth, which seemed to trigger vomiting. Justin was put back on medication for gastroesophageal reflux, but he continued to vomit with increasing frequency, which led to the psychiatric consultation.

Observation of mother and infant during feeding revealed that as soon as Justin finished feeding, he put his hand in his mouth and vomited. When his mother restricted his hand, Justin moved his tongue back and forth in a rhythmic manner until he vomited again. This happened repeatedly, and Justin continued the rhythmic tongue movements even when he could not bring up any more milk.

Because of his poor nutritional state and moderate dehydration, Justin was admitted to the hospital, and a nasojejunal tube was inserted for feedings. When Justin was awake, a special nurse or the parents played with him and tried to distract him whenever he attempted to put his hand in his mouth or thrust his tongue rhythmically. Justin became increasingly engaged, and his ruminatory activity decreased accordingly. After 1 week in the hospital, small feedings were started, but Justin tried to ruminate again, and the oral feedings had to be stopped. At this point, the mother decided to stop working and take Justin home to continue the treatment there. The mother started small feedings, played with Justin after feedings, and was able to keep him from ruminating. After 4 weeks of slow increments in his feedings, Justin was able to take all of his feedings by mouth without ruminating, and the nasojejunal tube could be removed.

Pica

Clinical Features. Young children with this disorder typically eat plaster, paper, paint, cloth, hair, insects, animal droppings, sand, pebbles, and dirt. Many of the children engage in other oral activities, such as thumb sucking or nail biting, which they seem to use for relief of tension and for self-soothing. In some younger patients, as well as those with neurodevelopmental or learning disorders, some clinicians regard pica as a

form of self-soothing behavior. The pica may lead to anemia, diarrhea/constipation, worm infestation, toxoplasmosis, lead poisoning, and malnutrition. In some cases it may lead to intestinal obstruction as the result of hairball tumors. In addition to heterogeneity of substances consumed, pica's associated behavioral features can differ considerably among individuals (e.g., its compulsive nature and relationship to emotional arousal or need for oral stimulation).

The major proposed change for pica disorder in the *Diagnostic and Statistical Manual for Mental Disorders*, Fifth Edition is its relocation from the section, titled "Disorders Usually First Diagnosed in Infancy, Childhood, or Adolescence," to the newly proposed section, "Feeding and Eating Disorders." This change emphasizes that this disorder occur across the age range, including adulthood. In particular, changes to phrasing of diagnostic criteria in DSM-5 are intended to guide clinicians in distinguishing eating behaviors that warrant a diagnosis of pica from behaviors that are developmentally normal, culturally supported or socially normative, or that support a diagnosis of a different mental disorder.

Differential Diagnosis. Before the age of 18 to 24 months, mouthing and sometimes eating of nonnutritive substances can be relatively common and should only be diagnosed as pica if it is persistent and inappropriate for the developmental level of the child. In DSM-5, 2 years is suggested as minimum age for a diagnosis of pica; at younger ages, mouthing different kinds of objects is considered developmentally normal. Consumption of substances consistent with local cultural practices or social norms would not warrant a pica diagnosis. Pica may be comorbid with a number of mental disorders, which are more likely to come to clinical attention than pica itself, e.g., intellectual development disorder, autism spectrum disorder, schizophrenia, and obsessive compulsive disorder. In addition, pica can co-occur with hair pulling disorder and skin picking disorder when hair or skin is ingested. When nonnutritive substances are ingested as a means to suppress appetite in anorexia nervosa, a pica diagnosis is not warranted. Some individuals with ARFID, particularly those whose selective or restricted intake is based on sensory aspects of food, may also present with comorbid pica.

Epidemiology. A survey of a large clinic population with a wide range of ethnic backgrounds revealed that 75 percent of 12-month-old infants and 15 percent of 2- to 3-year-old toddlers were reported by their parents to put nonnutritive substances in their mouth. However, the true incidence of pica is not known. Available data suggest that in the US pica is rare in healthy children older than 2 years. However, a study reported that 33.9 percent of children in Detroit who were in treatment for sickle cell anemia had pica eating. Pica has also been reported as prevalent in children elsewhere,

including in some school-age populations in Africa (as high as 77 percent). Among individuals with mental retardation, the prevalence of the disorder appears to increase with the severity of the retardation. The prevalence of pica among institutionalized individuals with mental retardation has been estimated to range from 10 to 33 percent.

Etiology and Course. Organic, psychodynamic, socioeconomic, and cultural factors have been implicated in the etiology of pica. Inadequate intake of iron and calcium has been shown in rats to induce pica and has been considered to lead to abnormal cravings and pica in humans. Low plasma levels of iron and zinc have been found to be associated with pica, and it has been suggested that hypozincemia with low iron levels may be a possible cause rather than a consequence of pica. Psychosocial factors—for example, poverty, maternal neglect and abuse, and disorganized family situations with inadequate nurturing and supervision of the children—have been implicated in the etiology of pica. The cultural acceptance of pica in certain population groups has also been considered to play a role in its etiology. Some authors have proposed a multifactorial model of etiology by which organic, familial, socioeconomic, and cultural factors interact with each other.

In most instances, the disorder is believed to last for several months and then remit. However, it may occasionally continue into adolescence or, less frequently, into adulthood. Several authors have pointed to the seriousness of the developmental effect of the disorder. The younger children showed delays in their speech and psychosocial development. Half of the adolescents were observed to suffer from depression and/or personality disorders, engaged in other forms of disturbed oral activities, such as thumb sucking and nail biting, and abused tobacco, alcohol, or drugs. More recently, some authors pointed to a strong relationship of pica during early childhood and bulimia nervosa during the adolescent years.

Treatment. A variety of treatments have been proposed. These include physical restraints, aversive and nonaversive behavior therapy, the teaching of children to exchange inedible items for edibles, environmental enrichment, and individual and group therapy for the children and their mothers. It is important to conduct a comprehensive evaluation of the child, the caretakers, and the child's environment to develop an individualized treatment plan. This should include helping the adults to become aware of the short- and long-term dangers of pica, providing a lead-free and childproof environment, facilitating a more satisfying parent–child relationship, and applying behavioral techniques to extinguish the maladaptive behavior of pica.

Susan was 3 years old when her mother took her to the pediatrician because of abdominal pain

and lack of appetite. The mother complained that Susan put everything in her mouth but did not want to eat regular food. The pediatrician observed that Susan looked pale, thin, and withdrawn. She sucked her thumb and quietly looked down while her mother reported that Susan liked to chew on newspapers and put plaster in her mouth.

The medical examination revealed that Susan was anemic and suffered from lead poisoning. She was admitted to the hospital for treatment, and a psychiatric consultation was obtained.

Further exploration of the history and the observation of mother and child during feeding and play revealed that the mother was overwhelmed with having to care for three young children and had little affection for Susan. The mother was unmarried and lived with her three children and five other family members in a three-bedroom apartment in an old housing project. Her 4-year-old son was hyperactive and demanded almost constant supervision. Her 18-month-old infant was an engaging and active little girl, whereas Susan would sit quietly, rock herself, suck her thumb, or chew on newspaper.

The treatment plan included the involvement of social services and protective services to remove any lead paint from the walls in the apartment and to look for better living arrangements for the family. The mother was helped to enroll Susan and her brother in a preschool program that provided them with more structure and stimulation and gave the mother a few hours of relief every day. In addition, Susan was seen with her mother and her younger sister in play therapy to help the mother understand the different temperaments of the two children and to make her more responsive to Susan's weak attempts to engage her. Once the mother felt more supported and less overwhelmed by her situation, she became more empathic and understanding of Susan. When Susan put something in her mouth, the mother was able to engage her in a play activity rather than scream at her and scold her for whatever she was doing. Over the period of 1 year, the relationship between Susan and her mother gradually improved, and Susan seemed less in need of putting her thumb or inedible things in her mouth.

References

Ammaniti M, Lucarelli L, Cimino S, D'Olimpio F, Chatoor I. Maternal psychopathology and child risk factors in infantile anorexia. *Int J Eat Disord*. 2010;43(3):233–240.

Ammaniti M, Lucarelli L, Cimino S, D'Olimpio F, Chatoor I. Feeding disorders of infancy: a longitudinal study to middle childhood. *Int J Eat Disord*. 2012;45(2):272–280.

Banerjee SP, Bhandari RP, Rosenberg DR. Use of low-dose selective serotonin reuptake inhibitors for severe, refractory choking phobia in childhood. *J Dev Behav Pediatr*. 2005;26(2):123–127.

Benoit D, Wang EE, Zlotkin SH. Discontinuation of enterostomy tube feeding by behavioral treatment in early childhood: a randomized controlled trial. *J Pediatr*. 2000;137(4):498–503.

Benton D. Role of parents in the determination of the food preferences of children and the development of obesity. *Int J Obes Relat Metab Disord*. 2004;28(7):858–869.

*Birch LL. Development of food preferences. *Annu Rev Nutr*. 1999;19:41–62.

Bryant-Waugh R, Markham L, Kreipe RE, Walsh BT. Feeding and eating disorders in childhood. *Int J Eat Disord*. 2010;43(2):98–111.

Chatoor I. Feeding disorders in infants and toddlers: diagnosis and treatment. *Child Adolesc Psychiatr Clin N Am*. 2002;11(2):163–183.

*Chatoor I. Diagnostic classification of feeding disorders. In: *Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood: Revised Edition (DC: 0–3R)*. Washington, DC: Zero to Three Press; 2005.

*Chatoor I, Ammaniti M. A classification of feeding disorders in infancy and early childhood. In: Narrow WE, First MB, Sirovatka PJ, Regier DA, eds. *Age and Gender Considerations in*

Psychiatric Diagnosis: A Research Agenda for DSM-5. Arlington, VA: American Psychiatric Association; 2007.

*Chatoor I, Conley C, Dickson L. Food-refusal after an incident of choking: a posttraumatic eating disorder. *J Am Acad Child Adolesc Psychiatry*. 1988;27(1):105–110.

Chatoor I, Dickson L, Einhorn A. Rumination: etiology and treatment. *Pediatric Ann*. 1984;13(12):924–929.

Chatoor I, Ganiban J, Harrison J, Hirsch R. Observation of feeding in the diagnosis of posttraumatic feeding disorder of infancy. *J Am Acad Child Adolesc Psychiatry*. 2001;40(5):595–602.

Chatoor I, Ganiban J, Hirsch R, Borman-Spurrell E, Mrazek DA. Maternal characteristics and toddler temperament in infantile anorexia. *J Am Acad Child Adolesc Psychiatry*. 2000;39(6):743–751.

Chatoor I, Hirsch R, Persinger M. Facilitating internal regulation of eating: a treatment model for infantile anorexia. *Inf Young Children*. 1997;9:12–22.

Chatoor I, Surles J, Ganiban J, Beker L, Paez LM, Kerzner B. Failure to thrive and cognitive development in toddlers with infantile anorexia. *Pediatrics*. 2004;113(5):e440–e447.

Chial HJ, Camilleri M, Williams DE, Litzinger K, Perrault J. Rumination syndrome in children and adolescents: diagnosis, treatment, and prognosis. *Pediatrics*. 2003;111(1):158–162.

Cooke LJ, Haworth CM, Wardle J. Genetic and environmental influences on children's food neophobia. *Am J Clin Nutr*. 2007;86(2):428–433.

Dellert SF, Hyams JS, Treem WR, Geertsma MA. Feeding resistance and gastroesophageal reflux in infancy. *J Pediatr Gastroenterol Nutr*. 1993;17(1):66–71.

Essick GK, Chopra A, Guest S, McGlone F. Lingual tactile acuity, taste perception, and the density and diameter of fungiform papillae in female subjects. *Physiol Behav*. 2003;80(2–3):289–302.

Fraiberg S, Adelson E, Shapiro V. Ghosts in the nursery. A psychoanalytic approach to the problems of impaired infant-mother relationships. *J Am Acad Child Psychiatry*. 1975;14(3):387–421.

Gaddini E. *A Psychoanalytic Theory of Infantile Experience: Conceptual and Clinical Reflections*. London, UK: Taylor & Francis e-Library; 2005.

Green AD, Alioto A, Mousa H, Di Lorenzo C. Severe pediatric rumination syndrome: successful interdisciplinary inpatient management. *J Pediatr Gastroenterol Nutr*. 2011;52(4):414–418.

Hartmann AS, Becker AE, Hampton C, Bryant-Waugh R. Pica and rumination disorder in DSM-5. *Psychiat Ann*. 2012;42:426–430.

Kern L, Starosta K, Adelman BE. Reducing pica by teaching children to exchange inedible items for edibles. *Behav Modif*. 2006;30(2):135–158.

Lee WS, Beattie RM, Meadows N, Walker-Smith JA. Gastro-oesophageal reflux: clinical profiles and outcome. *J Paediatr Child Health*. 1999;35(6):568–571.

Linscheid TR. Behavioral treatments for pediatric feeding disorders. *Behav Modif*. 2006;30(1):6–23.

Lucarelli L, Cimino S, D'Olimpio F, Ammaniti M. Feeding disorders of early childhood: an empirical study of diagnostic subtypes. *Int J Eat Disord*. 2013;46(2):147–155.

Lucarelli L, Cimino S, Zavaroni D, Ammaniti M. Infantile post-traumatic feeding disorder and

infantile anorexia: a clinical evaluation on maternal and child psychopathological risk factors. *Infant Ment Health J*. 2010;32(Suppl):7.

Lucarelli L, Speranza AM. The relationship between maternal attachment/psychopathology and the severity of infantile anorexia. *Infant Ment Health J*. 2014;35(Suppl):106.

Lyons-Ruth K, Zeanah CH, Benoit D. Disorders and risk for disorders during infancy and toddlerhood. In: Mash EJ, Barkley RA, eds. *Child Psychopathology*. New York: Guilford Press; 2006.

*Marchi M, Cohen P. Early childhood eating behaviors and adolescent eating disorders. *J Am Acad Child Adolesc Psychiatry*. 1990;29(1):112–117.

McNally RJ. Choking phobia: a review of the literature. *Compr Psychiatry*. 1994;35(1):83–89.

Nelson SP, Chen EH, Syniar GM, Christoffel KK. One-year follow-up of symptoms of gastroesophageal reflux during infancy. Pediatric Practice Research Group. *Pediatrics*. 1998;102(6):E67.

Nicholls D, Christie D, Randall L, Lask B. Selective eating: symptom, disorder or normal variant. *Clin Child Psychol Psychiatry*. 2001;6:257–270.

Oelschlager BK, Chan MM, Eubanks TR, Pope CE 2nd, Pellegrini CA. Effective treatment of rumination with Nissen fundoplication. *J Gastrointest Surg*. 2002;6(4):638–644.

Olden KW. Rumination. *Curr Treat Options Gastroenterol*. 2001;4(4):351–358.

Singhi S, Ravishanker R, Singhi P, Nath R. Low plasma zinc and iron in pica. *Indian J Pediatr*. 2003;70(2):139–143.

Skinner JD, Carruth BR, Wendy B, Ziegler PJ. Children's food preferences: a longitudinal analysis. *J Am Diet Assoc*. 2002;102(11):1638–1647.

Tutuian R, Castell DO. Rumination documented by using combined multichannel intraluminal impedance and manometry. *Clin Gastroenterol Hepatol*. 2004;2(4):340–343.

48

Tic Disorders

RAHIL JUMMANI, M.D., AND BARBARA J. COFFEY, M.D.

INTRODUCTION

Perhaps the best known of the tic disorders is Tourette disorder (TD), also known as Tourette syndrome, named after the French neurologist Georges Gilles de la Tourette, who in 1885, was among the first to recognize and describe the disorder and its comorbid conditions. Since the descriptions of Gilles de la Tourette, a spectrum of tic disorders has come to be recognized and codified in psychiatric nomenclature. No longer is Tourette disorder viewed as a rare, bizarre psychological entity with primarily disturbing manifestations, such as the shouting of obscenities, screeching, or barking. The comorbid psychiatric disorders that frequently occur in individuals with TD have become more widely recognized in recent years; it is these disorders that often may warrant treatment more than the tics themselves. Although the underlying etiologies of these disorders continue to be elusive, substantial progress has been made in the last few decades through the use of innovative neuroimaging techniques, genomics studies, and clinical science. Tic disorders have emerged as complex, heterogeneous, neurodevelopmental disorders that are the manifestation of an intricate interplay of genes and environment. Understanding of the etiology, genetics, neuroanatomy, and epidemiology of tic disorders has been significantly enhanced. Novel treatment approaches have also emerged, providing alternatives to the established pharmacological management of tics. Specific behavioral interventions have recently been established as first-line treatments for tic disorders. Deep brain stimulation, botulinum toxin injection, targeted surgery, and stem cell transplantation are examples of promising new treatments. The future holds tremendous possibilities as new insights into the etiology, pathophysiology, and treatment of tic disorders are rapidly developing. This chapter provides a review and update of the history, comparative nosology, epidemiology, etiology, diagnosis and clinical features, pathology, differential diagnosis, course and prognosis, and clinical treatment of tic disorders, including Tourette disorder.

DEFINITIONS

Tics are rapid, recurring motor movements (motor tics) or vocalizations

(phonic or vocal tics) that are nonrhythmic, involuntary or semivoluntary, and sudden in onset. Tics typically involve one muscle or a group of muscles and may be characterized by their anatomical location, number, frequency, duration, and complexity (Table 48–1). They can be classified as *simple* (involving one muscle, muscle group, or sound) or *complex* (slower, more purposeful movements involving coordinated muscle groups or multiple sounds). Simple motor tics cause a brief, jerking movement (*clonic tic*); a briefly sustained abnormal posture (*dystonic tic*); or an isometric contraction (*tonic tic*). Examples of simple motor clonic tics are eye blinking, shoulder shrugging, and head turning. Dystonic tics include sustained eye closure (blepharospasm), ocular deviations, teeth grinding (bruxism), mouth opening, twisting and tilting of the neck (torticollis), and shoulder rotation. Tonic tics are typically manifested by tensing of abdominal or limb muscles. Complex motor tics may include touching objects, jumping, rotating, or imitating others' gestures (*echopraxia*). Examples of simple vocal tics are throat clearing, coughing, and sniffing; repeating syllables, words, or one's phrases (*palilalia*) or repeating others' words or phrases (*echolalia*) are examples of complex vocal tics. Less commonly occurring tics include the involuntary vocalization of obscenities (*coprolalia*) and the involuntary production of obscene gestures (*copropraxia*).

Tics are often emitted in bouts of one or several tics separated by variable periods of no tic activity. Tics change in frequency, anatomical location, number, complexity, and severity over minutes, hours, and days, as well as over longer periods of time. Tics may persist during sleep. Simple motor tics are the most common initial presentation of a tic disorder, and eye blinking is the most common initial tic.

Tics are experienced as irresistible but can be suppressed for varying periods of time. Some patients describe suppressing tics for several hours, for example, while at work or school, and then experiencing a bout of tics. There is usually a reduction in the frequency and severity of tics during concentration on mental or physical tasks such as reading. Stress, excitement, boredom, fatigue, and exposure to heat can exacerbate tics. Tic frequency may also increase during relaxation following a period of stress, or sometimes when an individual is actively attempting to suppress tics. Some children, and occasionally adults, are unaware of their tics.

Many individuals with TD characterize their tics as a voluntary response to an uncomfortable feeling that precedes them. These feelings, or premonitory sensations, have been described by the majority of individuals with Tourette disorder; they can be localized or general, and physical or mental, in nature. These premonitory sensations distinguish tic disorders from other movement disorders. An example is a burning feeling in the eye before an eye blink. However, there are often nonlocalizable and more generalized premonitory phenomena, such as urges, sensations, or anxiety.

Many individuals report a "just right phenomenon," that to relieve the uncomfortable urge they must repeat a particular movement or sound or pattern of movements or sounds, until "it feels good" or "it feels just right."

Table 48–1.
Examples of Tics

Motor		Vocal	
Simple	**Complex**	**Simple**	**Complex**
Eye blinking	Touching objects or self	Throat clearing	Syllables or words
Nose twitching	Hitting or kicking	Coughing	Phrases
Shoulder shrugging	Hand gestures	Sniffing	Swearing
Sustained eye closure	Throwing	Blowing	Repeating one's or others' words
Tensing of abdominal or limb muscles	Squatting or jumping	Barking	Changes in the rate, rhythm, and volume of speech

Complex tics may be difficult to distinguish from compulsive behaviors. This is further complicated by the frequent presentation of obsessive-compulsive disorder (OCD) symptoms with tics. The difference between some complex tics and compulsions is subtle. It is important to determine whether the behavior is being pursued as a strategy to prevent or relieve anxiety rather than simply to satisfy an urge. If anxiety relief or prevention is the case, a compulsion rather than a tic is likely. Studies by Miguel and coworkers demonstrated that more cognitions and autonomic anxiety and fewer sensory phenomena are reported with repetitive behaviors in individuals with OCD without concomitant tics, whereas more sensory phenomena, such as the "just right phenomenon," and fewer cognitions are associated with repetitive behaviors in individuals with TD.

The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) and the tenth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) provide diagnostic criteria in common use for tic disorders. According to the DSM-5 (with the equivalent ICD-10 diagnoses in parentheses), the tic disorders include provisional tic disorder (transient tic disorder in ICD-10), persistent motor or vocal tic disorder (chronic motor or vocal tic disorder in ICD-10), Tourette disorder (combined vocal and multiple motor tic disorder [de la Tourette] in ICD-10), and other specified tic disorder and unspecified tic disorder (tic disorder not otherwise specified in ICD-10).

By DSM-5 criteria, a provisional tic disorder is characterized by the presence of single or multiple motor and/or vocal tics for less than 1 year. There has been a change in DSM-5 from the text revision of the fourth edition (DSM-IV-TR), the previous edition of the *Diagnostic and Statistical Manual of Mental Disorders*. The requirement of the presence of tics for more than a 4-week period is no longer stipulated in DSM-5.

A persistent motor or vocal tic disorder is diagnosed when either motor or vocal tics, but not both, are present for more than 1 year. For this category of tic disorders, DSM-5 provides specifiers to indicate whether motor or vocal tics are present: *With motor tics only* and *With vocal tics only*. There has been a significant change in DSM-5 from DSM-IV-TR here as well. The absence of a tic-free period of more than 3 consecutive months is no longer required. The DSM-5 criteria are now more consistent with ICD-10 criteria.

Tourette disorder (combined vocal and multiple motor tic disorder [de la Tourette] in ICD-10) is defined by the presence of both multiple motor and one or more vocal tics for more than 1 year. Again, the DSM-5, departing from DSM-IV-TR and more aligned with ICD-10, no longer mandates that there may not be a tic-free period of more than 3 consecutive months.

According to the DSM-5, other specified tic disorder is a tic disorder that does not meet criteria for any of the other tic disorders, and the clinician specifies the reason criteria are not met. An unspecified tic disorder, on the other hand, is a condition that also does not meet criteria for any of the other tic disorders, but for which the clinician does not elaborate a specific reason. Included are presentations for which information to make a more specific diagnosis is lacking.

In addition to the foregoing criteria that the DSM-5 and ICD-10 share for the diagnosis of a specific tic disorder, the DSM-5 specifies a few common criteria that must be met to diagnose provisional tic disorder, persistent motor or vocal tic disorder, or Tourette disorder. Onset of tics must be before age 18 years, and a substance or medical condition must not be causally implicated. Shifting from DSM-IV-TR, DSM-5 no longer states in its diagnostic criteria for the tic disorders that tics occur many times a day, nearly every day. Rather, DSM-5 notes that tics wax and wane in frequency. The changes in the latest edition reflect information about the tic disorders that has accumulated over the years from both research trials and clinical experience.

HISTORY

In 1825, the French physician Jean-Marc Gaspard Itard provided the first clinical description of Tourette syndrome in a French noblewoman, the Marquise de Dampierre, who displayed body tics, barking sounds, and uncontrollable utterances of obscenities. In his description, Itard also noted the presence of comorbid obsessive thoughts. In 1885, Georges Edouard Albert Brutus Gilles de la Tourette, a French neurologist, described eight patients, in addition to the patient originally reported by Itard, with childhood-onset tics and coexisting emotional symptoms and behavioral problems, such as inattention, hyperactivity, impulsivity, obsessions, and compulsions. Gilles de la Tourette coined the term

coprolalia for vocal tics that manifested as uncontrollable utterance of obscenities. Furthermore, he and his mentor, Jean-Martin Charcot, viewed the disease as progressive and believed they had identified a hereditary degenerative disorder.

Psychoanalytic views of Tourette disorder, introduced in the 1920s and 1930s, thrived in the early twentieth century, and interpreted tics as expressions of repressed childhood sexual conflict. In the 1940s, Margaret Mahler attributed tic symptoms to infantilization, parental overindulgence, and overprotection, as well as sexual conflict. Serge Lebovici, the leading French child psychoanalyst during the 1950s and 1960s, asserted that neurotic mothers were the cause of tics in their children.

As more biologically based theories of the etiology and pathophysiology of tic disorders emerged, sedatives, stimulants, carbon dioxide inhalants, and other agents came to be used in the treatment of tics, in conjunction with, or as alternatives to, psychoanalytic psychotherapy. In the 1960s and 1970s, as the scientific era of the understanding of Tourette syndrome began with the work of Arthur and Elaine Shapiro, the first randomized clinical trials using neuroleptic medications were published. The perception of Tourette and tic disorders as rare, bizarre, psychological conditions slowly began to change. Advances in the last 50 years have led to the use of the α-adrenergic agonists and second-generation neuroleptics as the primary psychopharmacological agents for the treatment of tics. Comorbid psychiatric disorders have come to be more widely recognized and treated. Over the last three decades, the Tourette Syndrome Association (TSA) has effectively promoted clinical awareness of the condition and supported research into its causes and treatment. Behavioral interventions have come to the forefront. The quest for a more comprehensive understanding of the underlying etiology and pathophysiology of the tic disorders continues.

COMPARATIVE NOSOLOGY

The nosology of the tic disorders has depended, since their recognition, on the presence and classification of their primary symptom—tics. The two predominant current classification systems of tic disorders, the DSM-5 and the ICD-10 criteria, are based primarily, as described earlier, on the type of tics present (motor, vocal, or both) and the duration of symptoms (greater or less than 1 year). As new information about the etiology and pathogenesis of tic disorders emerges, a new nosology based on an etiological or pathogenic understanding may replace the current symptom-based taxonomy. Over the years, with newer iterations of the DSM, individual criteria have changed slightly as advances in understanding of the prevalence and course of tic disorders has emerged. For example, age of onset of Tourette disorder has varied from before age 15 years to before age 21 years. The most recent version, DSM-5, has instituted changes to the

criteria that reflect the current understanding of the neurodevelopmental and temporal nature of tics, as tic disorders are now contained in the section on Neurodevelopmental Disorders. Tics are now recognized to wax and wane, and restrictions on the time frame of tic-free periods have been removed.

EPIDEMIOLOGY

Provisional tic disorder, which can encompass both motor and vocal symptoms, is the most common of the tic disorders; the prevalence is estimated as up to 20 percent of school-age children. A methodologically rigorous study of more than 800 schoolchildren ages 4 through 16 years in the Basque region, however, indicated a prevalence of approximately 6.5 percent for all tic disorders. The vast majority of tics were mild in severity and duration. The prevalence of persistent tic disorders is between 2 and 4 percent. Prevalence estimates for Tourette disorder also vary. As in all population-based studies, results may be influenced by a variety of factors, including criteria used for diagnosis, methods of sampling and inquiry, gender and age distributions in the study population, and size of the sample. Once considered an extremely rare disorder, current lifetime rate estimates vary from 1 to 10 per 1,000 children in European and Asian populations, and a common prevalence figure for individuals meeting criteria for Tourette disorder is 1 per 200 for the full spectrum of chronic tic disorders. A population-based, epidemiological study of lifetime prevalence of Tourette disorder in Israel estimated a point prevalence of 3.1 to 5.5 per 10,000 in 16- and 17-year-olds. Prevalence rates have been estimated to be as high as 4 percent in one community-based study, and 3 percent in a survey of middle school youth in the United Kingdom. In a school-based community survey, Kurlan and colleagues reported prevalence rates of tic disorders of 23 percent in special education classes and as high as 8 percent for TD, as compared to 19 percent for tic disorders and about 3 percent for Tourette disorder in regular classrooms. The 2007 US National Survey of Children's Health (NSCH) surveyed parents of children and adolescents aged 6 to 17 years. The estimated prevalence of a lifetime diagnosis of Tourette disorder by parent report was 3 per 1000. Most recently, in a meta-analysis of 13 studies of children Knight and colleagues reported a prevalence of Tourette disorder at 0.77 percent (95 percent confidence interval, 0.39 to 1.51 percent).

Typically beginning between ages 3 and 8 years in the vast majority of patients, Tourette disorder manifests by 11 years of age. The average age of onset is age 6 years. Thought to be lifelong, limited and somewhat contradictory information exists on the longitudinal course of the disorder. Recent studies suggest that whereas tics may often persist into adulthood, tic severity often declines significantly in adolescence. Only 20 percent or fewer of those with Tourette disorder continue to experience significant

functional impairment by age 20 years. In the majority of cases, tics in adults represent the persistence or recurrence of childhood-onset tics, but in rare instances, patients may have their first occurrence of tics or an exacerbation of tics during adulthood. Studies in both children and adults suggest that males are at least three to four times more likely than females to manifest Tourette disorder.

Comorbidity with psychiatric disorders, especially attention-deficit/hyperactivity disorder (ADHD) and OCD, is common in both clinically referred individuals and community samples. In addition, mood disorders and anxiety disorders, such as separation anxiety disorder and generalized anxiety disorder, are not uncommon. In 2007, the US National Survey of Children's Health (NSCH) surveyed parents of children and adolescents aged 6 to 17 years, 79 percent had at least one comorbid mental health or neurodevelopmental condition. ADHD was the most frequent co-occurring condition, and the prevalence of comorbid ADHD with Tourette disorder was reported to be as high as 60 to 70 percent. In addition, approximately 50 percent of patients with Tourette disorder have prominent obsessive-compulsive symptoms.

Jeffrey was a 5-year-old boy whose parents sought evaluation for recent onset of tics. Jeffrey developed onset of eye-blinking tics at age 4 years and subsequently began repeatedly to push out his hands and legs in front of him, kicking and stomping his feet. These symptoms persisted for about 3 months and remitted. The movements returned about 1 year later, somewhat abruptly, according to his parents. At that time, Jeffrey developed eye blinking and nose twitching. Sometimes, he hopped while walking. He also developed vocal tics around that time, including clicking, throat clearing, sniffing, and coughing. Approximately 1 month prior to evaluation, Jeffrey developed truncal tics that occasionally began to disrupt his gait.

Jeffrey's parents were also concerned about his problems paying attention in school in the last year. Hyperactive since preschool, he had a very short attention span and was often distracted. He could be impulsive, often unaware that he was doing things that could be dangerous. He often interrupted his parents when they were speaking. He was restless and fidgety in school, and his teachers were concerned that he might disrupt his classmates.

Jeffrey was the product of a pregnancy complicated by intrauterine growth retardation. When amniotic fluid was found to be at low volume at 36 weeks, delivery was induced. The delivery was uncomplicated, and birth weight was 5 pounds, 10 ounces. Except for a minor delay in the acquisition of speech, developmental milestones were within normal limits.

Currently in kindergarten, Jeffrey had attended a well-regarded public school since he was 2 years of age. He lived with his parents and 2-year-old sister, had many friends, and enjoyed play dates.

Jeffrey had frequent upper respiratory infections. There were no known allergies and no major medical problems. There was no known family history of Tourette disorder, tics, or OCD.

MENTAL STATUS EXAMINATION

Jeffrey was a healthy-looking 5-year-old boy, well-groomed and dressed. He was initially very shy, clinging to his mother. He appeared to be quite active and restless. Multiple tics were noted, including nose twitches, finger snapping, and a cough. Jeffrey was somewhat oppositional when asked to do a drawing.

IMPRESSION

In summary, Jeffrey was a 5-year-old boy with multiple motor and vocal tics for short episodes, including a bout lasting 3 months at age 4 years, and just over 2 months currently, consistent with provisional tic disorder. Tourette disorder could not be entirely ruled out, but there was

no diathesis on the basis of family history. Jeffrey also appeared to meet criteria for ADHD, combined presentation, which may have predisposed him to the development of tics. The ADHD symptoms were causing impairment in his academic functioning.

TREATMENT PLAN

Parent and teacher ratings of ADHD symptoms were recommended at initial evaluation, and at follow-up after treatment was initiated. Treatment of ADHD was recommended, starting with a cautious trial of a methylphenidate-based stimulant. The plan was to add a low dose of an α-adrenergic agonist such as guanfacine or clonidine if tics increased significantly. Consultation and psychoeducation were recommended for Jeffrey's teacher.

ETIOLOGY AND PATHOPHYSIOLOGY

The etiology and pathophysiology of Tourette disorder and other tic disorders remain unknown. Considerable data accumulated during the last decade from neuroanatomical and neurochemical studies of TD point to a diffuse process in the brain involving corticostriatothalamicortical (CSTC) pathways in the frontal lobes and basal ganglia, including the striatum (caudate nucleus and putamen). Several neurotransmitters and neuromodulators have been implicated, including dopamine, serotonin, and endogenous opioids. Although a specific animal model has not been identified, some parallels may exist with horses with crib biting, which display spontaneous motor tics accompanied by vocalizations, dogs with acral lick syndrome and mouse models with repetitive movements and stereotypies.

A host of biochemical, imaging, neurophysiologic, and genetic studies support the concept that TD is an inherited, neurodevelopmental disorder. Twin studies have reported 89 to 94 percent concordance for Tourette disorder. However, environmental factors have also been implicated. In one study of 16 pairs of monozygotic twins, low birth weight was a strong predictor of more severe tics. Several gestational and perinatal insults, including hypoxia–ischemia, exposure to androgens, heat, and infectious agents, have been implicated in the pathophysiology of tic disorders. The effect of prenatal ischemic events may be to alter dopamine signaling, heralding the development of tics. Delivery complications, especially forceps deliveries, and fetal exposure to high levels of coffee, cigarettes, or alcohol were associated with male gender and OCD in a sample of TD probands.

Androgen exposure is speculated to play a role in the development of tics, given the higher prevalence of Tourette disorder in males. Studies in both children and adults suggest that males are at least three to four times more likely than females to manifest Tourette disorder. X-linked genomic transmission has also been considered, but this appears less likely in the context of frequent male-to-male transmission within families. The role of stress, given the strong association of tics with anxiety, is also compelling. Tourette disorder patients have been found to have significantly higher levels of cerebrospinal fluid (CSF) corticotrophin-releasing factor (CRF)

than normal controls or patients with OCD who do not have a tic disorder, demonstrating a heightened reactivity of the hypothalamic–pituitary–adrenal axis. Increased reactivity of the noradrenergic sympathetic system has also been shown in patients with Tourette disorder. A dopamine-mediated pathway in the hypothalamus that modulates sweating in response to changes in temperature has also been linked with tic activity. Increases in ambient and core body temperatures have been shown to transiently increase tics in some individuals. A stress diathesis model modulated by the hypothalamic–pituitary–adrenal axis, the associated stress-related neurotransmitters and hormones, and their targets may represent the final common pathway that may explain the variable expression of the tic disorders.

Pediatric Autoimmune Neuropsychiatric Disorders Associated with Streptococcal Infection

A postinfectious mechanism for the development of tics and OCD has been postulated since the late 1800s. A post-streptococcal infection autoimmune etiology has long been well established for Sydenham's chorea, the movement disorder that has been the model for the development of Pediatric Autoimmune Neuropsychiatric Disorders Associated with Streptococcus (PANDAS) hypothesis. In addition to chorea, some Sydenham's patients display motor and phonic tics, as well as obsessions, compulsions, hyperactivity, impulsivity, and inattention. In 1998, Swedo and colleagues first described a putative subgroup of 50 children with PANDAS with OCD and/or tic disorders with acute symptom onset and/or exacerbation precipitated by infection with group A β-hemolytic Streptococcus (GABHS). Although GABHS is considered the initial autoimmune response-inciting event in children with PANDAS, viruses, other bacteria, and even noninfectious immunological responses are thought to be capable of precipitating later symptom exacerbations. Case reports have also implicated Lyme disease and Mycoplasma pneumonia in the development of infection-triggered tics. By molecular mimicry—the postulated autoimmune mechanism—antibodies are formed against the inciting agent and subsequently attack neurons because of some structural similarity.

Specific diagnostic criteria for PANDAS have been described, including the presence of OCD and/or tic symptoms, prepubertal onset, episodic course with sudden, explosive onset and/or exacerbations and remissions, and a temporal relationship of GABHS infection and neuropsychiatric symptoms. A case-control study of 144 children with OCD, Tourette disorder, or another tic disorder conducted by Mell and colleagues provided the strongest evidence for the involvement of GABHS in the onset of tic disorders and OCD. Patients with OCD, Tourette disorder, or another tic disorder were found to be more likely than controls to have had a

streptococcal infection in the 3 months prior to the onset of symptoms. Risk was greatest among those with multiple streptococcal infections in the preceding 12 months. Others have found no more than a chance association between GABHS infections and tic symptom exacerbations. A case-control study by Perrin and colleagues found little evidence of increased tics, obsessions, or compulsions after well-documented and treated GABHS infections.

Several groups have reported increased antistreptococcal antibody titers in individuals with OCD and tic disorders, but several other groups have not found this association. A double-blind, placebo-controlled, crossover study of prophylactic oral penicillin to attempt to reduce PANDAS recurrences failed to show differences between drug and placebo. More recently, in two prospective case-control studies of children with PANDAS Kurlan and colleagues reported that Group A B-hemolytic streptococcus infection was not the only antecedent associated with exacerbations of tics and OCD symptoms in the cases; similarly, Leckman and colleagues found there was no evidence for a temporal association between GABHS infections and tic/OC symptom exacerbations in children who met published PANDAS diagnostic criteria.

Treatments based on a postulated autoimmune mechanism have been nonspecific, and the results have been inconsistent. Immune therapies such as intravenous immunoglobulin (IVIG), plasmapheresis, corticosteroids, and plasma exchange have been explored with inconsistent results. Immunomodulatory treatments remain experimental, and current recommendations are for use only in research protocols.

The link between infectious etiology and tic disorders/OCD remains a work in progress. Building on previous work, most recent investigation has shifted its focus on acute, explosive onset OCD symptoms, which Swedo and colleagues have described as Pediatric Acute Onset Neuropsychiatric Syndrome (PANS). Anxiety, emotional lability, developmental and behavioral regression, deterioration in schoolwork and handwriting, and excessive urination are among the secondary characteristics.

Other new areas of research include investigation of proinflammatory cytokines elevation, as well as decreased numbers of regulatory T cells in some Tourette disorder patients. Clinical implications suggest vigilance in evaluation of tic or OCD patients with pharyngitis or streptococcus exposure, a low threshold for obtaining streptococcus throat culture, anti-streptolysin O [ASO] and anti-DNAse B titers, and vigorous treatment with antibiotics when streptococcal infection is confirmed. Antibiotic treatment has been shown to be remarkably effective in some unambiguous cases of sudden onset or exacerbation of tics with associated laboratory documented streptococcal infection.

Neuroimaging Studies

Structural MRI studies in Tourette disorder demonstrate volume changes in brain regions associated with CSTC circuits. Some studies have shown a loss of the normal basal ganglia asymmetry, as well as decreased basal ganglia volume in TD patients. In ten pairs of monozygotic twins, the right caudate was smaller in the more severely afflicted individuals, supporting evidence for the role of environmental components. The largest study of basal ganglia volume with 154 patients with Tourette disorder and 130 controls, conducted by Peterson and colleagues, found a significant decrease in the volume of the caudate nucleus in both children and adults. Larger dorsal prefrontal regions in children, but significantly smaller volumes in adults, were also reported by this group. Dorsal prefrontal volume was inversely related to tic severity, such that patients with greater prefrontal volumes exhibited decreased tic severity. Larger prefrontal volumes in children with TD may be a manifestation of synaptic plasticity associated with the constant need to suppress tics in social contexts. Adults with persistent tics may represent a subsection of the overall Tourette disorder population who do not generate a neuroplastic, compensatory response in the prefrontal cortex, leading to increased severity of the disorder and its continuity into adulthood.

Larger parieto-occipital region volumes and smaller inferior occipital volumes have also been observed in individuals with Tourette disorder when compared with controls. Bloch and colleagues found an inverse correlation between caudate volume in childhood and tic severity in late adolescence and early adulthood. Smaller caudate volumes were associated with more significant later tic activity, as well as with OCD symptoms. Another more recent study reported increased volume of the putamen bilaterally in boys with Tourette disorder.

Decrease in corpus callosum size in children and an increase in corpus callosum size in adults with Tourette disorder have also been noted. Corpus callosum size correlates inversely with prefrontal volume and is positively related to tic severity. These findings may indicate age-dependent changes in the white matter tracts in the disorder. Alternatively, decreased corpus callosum size, like increased dorsolateral prefrontal volume, may represent a compensatory mechanism children with Tourette disorder develop to facilitate suppression of tics. Adults with TD tend to show the opposite pattern, exhibiting relatively decreased dorsal prefrontal and increased corpus callosum volumes that may be associated with the persistence of symptoms into adulthood.

Several studies consistently implicate primary cortical dysfunction in Tourette disorder. Cortical thinning has been documented in frontal and parietal lobes, especially the sensorimotor cortex; the degree of thinning may correlate with tic severity. Increased sensorimotor cortical volume has been correlated with strength of premonitory sensations in a study of adults with TD. Evaluation of subgroups of Tourette disorder patients has

shown a relationship between symptoms and location of cortical thinning. While most have found thinner cortex or decreased volume in several different cortical areas, a few studies have reported overall brain volume decreases in Tourette disorder subjects.

Larger left thalamic volumes were found in 18 patients with Tourette disorder compared to 16 control subjects in one study. Increased gray-matter volumes bilaterally in the ventral putamen and decrease in gray matter in the left hippocampal gyrus have also been shown. Peterson and team demonstrated significantly increased overall hippocampus and amygdala volumes in TD patients, with a decline in volume with age in the TD group, but not in healthy controls. In children and adults with Tourette disorder, these volumes were inversely correlated with the severity of tics, OCD symptoms, and ADHD symptoms. An explanation for these findings is that the sub-region plays a compensatory and neuromodulatory role with tic, OCD, and ADHD symptoms. In addition, cerebral hyperintensities—bright patches in brain parenchyma on T_2-weighted MRI that are indicative of neuroinflammatory or neurodegenerative processes—have been detected at increased rates in children and adolescents with Tourette disorder, OCD, or ADHD in subcortical regions (primarily the basal ganglia and thalamus). Taken together, volumetric studies suggest that cortical and subcortical impairment, including striatal anomalies, and disruption of the temporolimbic pathways of the CSTC circuits are implicated in the pathophysiology of Tourette disorder.

Functional neuroimaging studies using state-dependent cerebral blood flow, oxygenation, and energy (glucose) metabolism parameters as a measure of brain activity (e.g., functional MRI [fMRI], positron emission tomography [PET], and single photon emission tomography [SPECT]) suggest altered activity (increased or decreased) in various areas of the brain (e.g., frontal and orbital cortex, striatum, putamen) in TD patients compared to controls, or across brain regions within TD patients. An fMRI study by Bohlhalter and colleagues that analyzed activity two seconds before and after tic onset indicated that paralimbic and sensory association areas are critical to tic generation. Other fMRI studies in individuals with TD indicate that greater ability of basal ganglia structures to suppress cortical activity may be related to decreased tic severity, supporting findings from PET and SPECT studies of the role of the basal ganglia in Tourette disorder. Functional neuroimaging studies suggest that voluntary tic suppression involves deactivation of the putamen and globus pallidus, coupled with partial activation of regions of the prefrontal cortex and caudate nucleus.

Event-related PET scanning has connected increased or decreased activity in many brain areas with the presence of tics, including the prefrontal cortex, premotor and primary motor cortex, anterior cingulate cortex, putamen, and caudate. fMRI scanning has demonstrated increased

neural activity in individuals with Tourette disorder compared to controls throughout various motor areas, including the sensorimotor cortex, putamen, pallidum, and substantia nigra; this increased activity correlates with symptom severity. On the other hand, weaker activity has been demonstrated in the TD patients' motor control pathways, including caudate and anterior cingulate cortex. More severe tics have been observed in those with progressively less activity in these regions, suggesting disruption in these areas results in failure to control tics or premonitory urges. Other fMRI studies have also shown age-related abnormalities in frontostriatal and frontoparietal connections, implicating abnormal maturational changes in Tourette disorder patients, strengthening the understanding that tics are the result of neurodevelopmental deficits.

Abnormal dopamine neurotransmission that contributes to the pathophysiology of Tourette disorder is supported by increased levels of dopaminergic innervation of the striatum of TD patients in SPECT studies. A study in monozygotic twins concordant for Tourette disorder but discordant for severity demonstrated abnormalities in D_2 receptors in the caudate area in the more severely afflicted twin. These neuroimaging findings are consistent with results of well-established tic suppression treatments with neuroleptic drugs that block dopamine receptors. Furthermore, TD patients have been shown to release more dopamine in response to amphetamine at dopaminergic synapses than healthy controls, and dopamine-releasing drugs may precipitate or exacerbate tics. Midbrain dopaminergic neurons play a central role in motor control and attentional processes through direct connections to the striatum and cortex, respectively. Glutamatergic, serotonergic, cholinergic, noradrenergic, and opioid systems may also be implicated, given that they are present in CSTC circuits and interact with the dopaminergic system.

Taken together, structural and functional neuroanatomical findings support the understanding that Tourette disorder results from abnormal activity in diverse cortical and subcortical brain regions. Active suppression of tics may require activation of dorsolateral prefrontal circuits that increase overall activity in the caudate nucleus, thereby inhibiting activity in the putamen.

Genetic Findings

Genetic data have been derived historically from family pedigree and twin studies. Tourette disorder, persistent tic disorders, and OCD cluster in families. Twin studies have shown a high concordance rate for tic disorders among monozygotic (MZ) pairs (ranging from about 50 to 90 percent), in contrast to a relatively low rate among dizygotic twins (often around 20 percent). Other studies have shown that first-degree relatives of TD patients have a higher risk of Tourette disorder, persistent tics, and OCD

than normal controls. Historically, genetic analyses were thought to be consistent with an autosomal dominant mode (with incomplete penetrance) of inheritance; some investigators argued that a semidominant, semirecessive pattern could also exist. However, older studies were confounded by relatively high rates of bilineal transmission with apparent genetic contributions from both paternal and maternal lines.

More recently, investigators have proposed multiple gene loci in combination with environmental factors. The precise type or severity of tic disorder may be different from one generation to another and may be modified by nongenetic factors. A systematic genome scan of 76 affected sibling pair families with 110 sib pairs conducted by the Tourette Syndrome Association International Consortium for Genetics (TSAICG) demonstrated that 4q, 8p, and four other regions on chromosomes 1, 10, 13, and 19 were regions of interest. Efforts to confirm and extend these findings, however, have led to the identification of new regions and a failure to replicate the original findings.

Specific TD populations have been studied in South Africa, Costa Rica, and French-speaking Canada. Regions near the centromere of chromosome 2, as well as on 6p, 8q, 11q, 14q, 20q, 21q, and X, have been implicated. A linkage analysis study of a large pedigree of multigenerational TD patients in the United Kingdom is suggestive of linkage at loci on chromosomes 5, 10, and 13 in at least one of the families. The largest genetic linkage study, undertaken by the Tourette Syndrome Association International Consortium for Genetics (TSAICG), consisting more than 2,000 individuals, using a whole-genome scan with 390 microsatellite markers demonstrated strong evidence of linkage for a region on chromosome 2p.

Cytogenetic abnormalities in Tourette disorder families have also been studied. A study by the TSAICG reported disruption of the contactin a–associated protein 2 gene on chromosome 7, coding for a membrane protein important for signal conduction along myelinated neurons at the nodes of Ranvier, in TD patients. Affected individuals were found to have alterations in chromosomes 2 and 7. Other Tourette disorder patients have been described with de novo duplication of the long arm of chromosome 7, a balanced translocation between chromosomes 6 and 8 with a translocation breakpoint on chromosome 8q, and an 18q21.1–q22.2 inversion. Abelson and colleagues identified a candidate gene, SLITRK1, on chromosome 13q31.1 in some individuals with Tourette disorder, with expression of the gene product in the basal ganglia and deep layers of cortex, brain regions implicated in the pathophysiology of Tourette disorder. Most recently, Ercan-Sencicek and colleagues reported a rare functional mutation in the *HDC* gene encoding l-histidine decarboxylase, the rate-limiting enzyme in histamine biosynthesis, in linkage analysis in a two-generation Tourette disorder pedigree.

Overall, current understanding of Tourette disorder and tic disorders

implicates multiple vulnerability genes that likely determine the form and severity of tics in the wide variety of afflicted individuals and their families. Candidate genes have been identified in individuals with Tourette disorder, including the dopamine receptor genes, the dopamine transporter genes, several noradrenergic genes, and some serotonergic genes. However, the recently published TSAICG genome-wide association study (GWAS) reported that no genes reached threshold significance.

DIAGNOSIS AND CLINICAL FEATURES

A comprehensive, detailed history with observations derived from multiple sources is the cornerstone of clinical evaluation of tics and comorbid neurodevelopmental and psychiatric disorders. Neurological, laboratory, and imaging examinations in patients with tic disorders are usually normal. Because tic disorders are diagnosed on the basis of history, reliable sources of information are essential. This is particularly important because in many patients tics may be suppressed during the initial examination. It is critical to determine the degree to which tics interfere with emotional, social, family, and academic functioning, because treatment strategies are most effective when the severity and impact of tics are taken into consideration. Diagnostic evaluation is closely connected with the first steps of treatment.

Historical inquiry should include detailed medical and developmental information, medication history (including substances of abuse or recreation), educational and occupational data, social and interpersonal history, and a thorough family pedigree covering at least three generations. A careful, descriptive, longitudinal assessment of the tic disorder is important, including onset, course, current severity, presence of premonitory sensations and the capacity for tic suppression, overall burden, and treatments to date. The physical examination should include height, weight, presence or absence of dysmorphic features, posture, gait, reflexes, and a systematic rating of current abnormal movements. The Abnormal Involuntary Movement Scale (AIMS) is a systematic assessment procedure (and rating form) that can be adapted for use with children. Videotaped standardized interviews and evaluation procedures are used in some research studies; these could also be used in the clinical setting. However, the absence of observed tics on initial examination does not preclude the diagnosis of a tic disorder. Unusual histories, co-occurring mental status changes, or the presence of seizures always warrant referrals for thorough neurological evaluation.

Psychiatric evaluation should include a formal assessment of the behavioral and emotional problems known clinically to cluster with tic disorders, including ADHD, OCD, anxiety disorders, mood disorders, and manifestations of impaired or dysregulated affect (e.g., aggression). The use of structured or semistructured interviews, such as the Diagnostic Interview Schedule for Children (DISC) or the Children's Schedule for

Affective Disorders and Schizophrenia (K-SADS), can improve classification and evaluation of psychiatric comorbidity. Standardized rating scales developed specifically for the tic disorder population have improved diagnostic reliability in research studies and can also be helpful in clinical care. The Yale Global Tic Severity Scale (Y-GTSS) and the Tourette Syndrome Symptom List (TSSL) rate tics, compulsions, and other associated features. Specific rating scales for OCD (the Children's Yale-Brown Obsessive Compulsive Scale [C-YBOCS]) and ADHD (the Iowa Conners parent and teacher rating scales, or the Swanson, Nolan, and Pelham [SNAP] scale) can also be used. Because rating instruments can provide quantifiable baseline data, such as frequency and intensity of tics, these data can also be used to measure treatment response (Table 48–2).

Auxiliary data from outside sources is essential. Pediatric and medical records may document developmental and medical history, adequacy of medication trials, hospitalizations, and laboratory findings. Review of educational records is advised for children because many TD patients manifest tics or associated problems in school settings. Report cards can document academic performance. Direct phone contact with teachers may provide data about attentional functioning, social and emotional competencies, and any behavioral difficulties. Neuropsychological or speech and language testing may be indicated for patients with impairments in school or occupational functioning. Identified areas of strength and weakness are subsequently conveyed to appropriate personnel for inclusion into educational or vocational planning.

Table 48–2.
Rating Scales

Domain	Type	Reliability and Validity	Sensitive to Change
Tics			
Yale Global Tic Severity Scale (YGTSS)	Clinician	Excellent	Yes
Abnormal Involuntary Movement Scale (AIMS)	Clinician	Excellent	Yes
Tic Symptom Self-Report	Parent/self	Good	Yes
ADHD			
Swanson, Nolan, and Pelham-IV (SNAP-IV)	Parent/teacher	Excellent	Yes
Abbreviated Conners Questionnaire	Parent/teacher	Excellent	Yes
OCD			
Yale-Brown Obsessive Compulsive Scale (YBOCS) and Children's Yale-Brown Obsessive Compulsive Scale (CYBOCS)	Clinician	Excellent	Yes
National Institute of Mental Health Global	Clinician	Excellent	Yes
Mood and Anxiety Disorders			
Children's Depression Inventory (CDI)	Clinician	Good	Yes
Children's Depression Rating Scale (CDRS)	Self	Good	Yes
Multidimensional Anxiety Scale for Children (MASC)	Self	Good	Yes

General			
Achenbach Child Behavior Checklist (CBCL)	Parent/teacher	Excellent	No
Clinical Global Impressions Scale (CGI)	Clinician	Excellent	Yes

PATHOLOGY AND LABORATORY EXAMINATION

In general, laboratory tests are indicated only to rule out other disorders when such conditions are highly implicated in the differential diagnosis of a given patient. For example, specific tests are available for the definitive diagnosis of Huntington disease, Wilson disease, and neuroacanthocytosis.

DIFFERENTIAL DIAGNOSIS

Differential diagnosis of tic disorders can be challenging (Table 48–3). Medical and family histories, physical and neurological examinations, and baseline laboratory testing enable clinicians to rule out potential causes of de novo tics such as infection, trauma, substance use, or toxin ingestion. Diagnostic hallmarks of tic disorders include the waxing-and-waning natural history, onset in childhood or adolescence, and repetitive, rapid, nonrhythmic phenomenology, and involuntary features. In the absence of at least some history of simple tics, the diagnosis of a tic disorder may be questioned. Furthermore, the presence of vocal tics is helpful in ruling out other disorders because phonic manifestations are rare in other neurological conditions. Tics must be distinguished from other types of abnormal movements: myoclonus, tremors, chorea, athetosis, dystonias, akathisia, and compulsive rituals.

Diagnostic considerations among the movement disorders include Sydenham's chorea, poststreptococcal neurological complications such as choreiform, writhing, and truncal movements; Huntington disease, an autosomal dominant disorder presenting with chorea and dementia and with an onset typically in the fourth or fifth decade; Parkinson disease, typically a late-life disorder, characterized by flattened facies, bradykinesia, rigidity, cogwheeling, and "pill-rolling" resting tremor; Wilson disease, an autosomal recessive disorder of copper metabolism that presents with liver abnormalities and neurological symptoms: early dementia, mood disorders or psychosis, asterixis (a flapping tremor of the hands); parkinsonism; structural diseases such as hemiballismus associated with sub-thalamic nucleus impairment; and neuroacanthocytosis, a rare movement disorder marked by progressive muscle weakness and atrophy, progressive cognitive loss, chorea, acanthocytosis (spiked red blood cells), and, sometimes, facial tics. In psychiatric settings, patients who have been exposed to neuroleptics are at risk for neuroleptic-related dyskinesias, including tardive and withdrawal dyskinesias. It is important to remember that dyskinesias may also be observed in tic disorder patients secondary to neuroleptic therapy but may be easily overlooked.

Table 48–3.
Differential Diagnosis of Abnormal Movements

Descriptive Terms	Observable Movement Pattern(s)
Akathisia	Motor restlessness (unpleasant need to move), usually in the lower extremities
Athetosis	Slow, irregular, writhing movements, usually in the hands and fingers but often involving the neck
Ballismus	Intermittent, coarse, large-amplitude, jerking, shaking, flinging movements
Chorea	Irregular, spasmodic movements, usually affecting the limbs or face
Compulsion	Repetitive behavior that is more goal directed and may be undertaken to prevent or reduce anxiety or distress
Dyskinesia	Choreiform or dystonic movement that is involuntary, stereotyped, and not suppressible
Dystonia	Sustained, tonic contraction that progresses to abnormal postures
Myoclonus	Sudden, brief, clonic, shock-like jerks, usually involving the limbs
Periodic movements of sleep	Periodic dorsiflexion of the foot and flexion of the knee, occurring during sleep
Spasm	Stereotypic, prolonged movement involving the same group of muscles
Stereotypy	Repetitive, usually meaningless, gestures, habits, or automatisms
Tic	Stereotyped, rapid, recurring motor movement (motor tic) or vocalization (phonic or vocal tic) that is nonrhythmic, involuntary or semivoluntary, and sudden in onset
Tremor	Unintentional, somewhat rhythmic, muscle movement involving oscillations of one or more parts of the body

COURSE AND PROGNOSIS

The typical course of Tourette disorder is characterized by onset of simple motor tics of the eyes, face, head, or neck at about age 6 or 7 years, followed by a rostral-caudal progression of motor tics over several years to involve the shoulders, trunk, and extremities. Eye blinking is the most common initial tic. Vocal tics typically start at ages 8 to 9 years; more-complex tics often begin later, as do obsessive-compulsive symptoms, at about ages 11 to 12 years. As many as half of clinically referred Tourette disorder patients may show signs of ADHD, such as hyperactivity, impulsivity, and distractibility, prior to the onset of tics. Tics range from mild to severe and tend to stabilize over time; most children evaluated even in specialty clinics have mild to moderate tics. Some patients have lengthy periods during which most or all manifestations attenuate or remit. Recent studies indicate that tic severity improves significantly in most patients by early to mid-adolescence.

Social functioning may be impaired because of rejection by peers or anxiety about having tics in social situations. Tics may impair academic functioning by their disruptive impact in class or interference with homework completion; occupational functioning may be impaired by inability to use a computer or complete work tasks. Persistent tic symptoms may cause considerable distress and can lead to social isolation and

personality changes. In severe cases, tics may directly interfere with daily activities. For example, violent head and neck tics may cause physical complications, such as compressive cervical myelopathy.

Comorbid Psychiatric Disorders

The relationship between Tourette disorder and frequently observed comorbid psychiatric disorders has not yet been disentangled. Obsessive-compulsive symptoms, developmentally inappropriate motoric hyperactivity and inattention, anxiety, and aggressive dyscontrol have often been described in patients with Tourette disorder. Motor, vocal, behavioral, cognitive, and emotional dysfunction may represent manifestations of an underlying central core disinhibition.

A majority of clinically referred individuals with Tourette disorder meet full or sub-threshold diagnostic criteria for ADHD, OCD, or both at some time during the course of their illness. Interestingly, these findings are also reported in community samples. ADHD and OCD are more likely to interfere with overall functioning and academic and work performance than the tics. A recent investigation of functional impairment in youth with Tourette disorder conducted by Storch and colleagues found that non–tic-related problems caused more dysfunction than the tics per se, most often either ADHD or OCD symptoms.

ADHD is diagnosed by history of an enduring pattern of developmentally inappropriate inattention and/or hyperactivity and impulsive behavior; there is a bidirectional relationship between tics and ADHD. Some investigators have reported that 50 to 75 percent of Tourette disorder patients also meet criteria for ADHD, whereas about 10 to 30 percent of ADHD patients also meet criteria for a tic disorder. Co-occurring Tourette disorder and ADHD are often associated with academic difficulties, peer rejection, family conflict, and disruptive behaviors such as explosive outbursts, overt aggression, low frustration tolerance, and poor treatment adherence. The nature of the relationship between ADHD and Tourette disorder is not firmly established, but it has been clearly demonstrated that ADHD predicts greater disability than that associated with tic disorders alone, including an earlier diagnosis of Tourette disorder, and a much higher rate of anger control problems and oppositional defiant disorder, sleep difficulties, learning disability, mood disorders, social skills deficits, sexually inappropriate behavior, and self-injurious behavior.

OCD is characterized by obsessions—recurrent, intense, intrusive ideas, thoughts, impulses, or images that cause anxiety or distress, or compulsions—repetitive behaviors or mental acts sought to prevent or ward off anxiety or distress caused by an obsession. Additionally, there appears to be a bidirectional relationship between Tourette disorder and OCD in most patient cohorts. From 20 to 40 percent of Tourette disorder patients meet full criteria for OCD, and up to 90 percent have been reported to have

sub-threshold symptoms such as repetitive counting, touching, or symmetry needs.

Family studies provide the most compelling evidence for the association of tic disorders with OCD. First-degree relatives of TD probands have higher frequencies of TD, as well as OCD, and first-degree relatives of OCD probands have higher frequencies of OCD and tic disorders than relatives of control subjects. Studies of OCD probands indicate that the earlier the onset of obsessions or compulsions in the proband, the higher the genetic loading in family members for both OCD and tics.

In addition to ADHD and OCD, mood and anxiety disorders have been described in clinically referred and community samples of Tourette disorder patients, including major depression, bipolar disorder, generalized anxiety disorder, separation anxiety disorder, panic disorder, and simple and social phobias. Coffey and colleagues reported that comorbid mood disorders were the most robust predictors of psychiatric hospitalization and illness severity in a cohort of Tourette disorder patients. Tic severity was only marginally significant as a predictor of psychiatric hospitalization. More recently, in a controlled study of older adolescents with TD, Gorman and colleagues found that, independent of comorbidity with ADHD, TD patients had significantly higher rates of depression and poorer overall outcome than healthy controls.

Whether mood symptoms are primary (i.e., related to the underlying pathophysiology of Tourette disorder) or secondary to the demoralization and/or impairment related to having a chronic illness remains to be clarified. Furthermore, a substantial minority of youth with Tourette disorder evaluated in clinical settings may meet criteria for intermittent explosive disorder manifested by rage attacks or oppositional defiant disorder. Budman and colleagues reported that children with TD plus ADHD and OCD are more likely to experience explosive outbursts, and the number of outbursts is positively correlated with the number of comorbid conditions. Tic type and severity are not linked to the presence of explosive outbursts.

Learning and academic problems are common in children with Tourette disorder; in a database of 5,450 patients with Tourette disorder, 1,235 (22.7 percent) had learning disabilities. Young patients with Tourette disorder may experience problems with executive functions, such as planning, persistence, organizational skills, and social problem solving. Patients with Tourette disorder have been consistently shown to have difficulties with procedural learning, fine motor control, motor inhibition, nonverbal memory, and visual motor integration. In addition, handwriting may be poor.

Many clinically referred children with Tourette disorder are reported to have difficulty with sensory-perceptual functioning, such as extreme sensitivity to touch or sound. Investigators have hypothesized that

dysregulation of central sensory integration and/or sensory gating problems may result in these sensitivities, but more research is needed.

Mark was a 13-year-old boy referred by his pediatrician for consultation for tics and anxiety. Mark developed the onset of chewing his collar at around age 4 or 5 years. Between ages 5 and 6 years, he began to make sniffing and snorting sounds, and at age 6 or 7 years, he developed eye blinking and head and neck movements. Mark was subsequently referred at age 8 years to a pediatric neurologist, who diagnosed Tourette disorder. No treatment was initiated at that time because Mark's tics were not causing any distress or impairment. His parents reported that since age 10 years, the severity of Mark's tics had diminished. He was currently repeating words and making guttural sounds and bird noises at times.

Parents reported that Mark had a long history of anxiety. As a very young child, he was anxious about separation for preschool. He was described as a child who was slow to warm up, with a cautious temperament, waiting until he was familiar with a situation to enter into a group. Over the years, teachers commented that they would like to see more class participation from him. Parents reported that Mark often preferred to stay at home rather than to go out or take initiative to see friends on play dates.

Mark reported that he did not like to speak in public because he was afraid that what he would say would "not be perfect." Mark described ruminating on "things he should have said or done differently." Mark always worried about performance, about his mother dying, and about bad things happening. Mark had slept with his parents until age 10 years because of a history of nightmares.

Mark had had concerns about contamination, and he used Purell® and washed his hands frequently. Mark hoarded objects, and preferred to do things a certain number of times; in his backpack he kept 20 books, 20 blue pencils, and 20 black pencils. In the past, he used to sharpen pencils excessively to get them to a "just right" point.

Mark was the product of a full-term, uncomplicated pregnancy; developmental milestones were described as entirely within normal limits. Currently in eighth grade, he had attended private school since kindergarten and was an A student. Mark enjoyed sports and spending time with friends.

There was no family history of Tourette disorder or OCD. Mark's father had been treated for anxiety. A paternal uncle had a long-standing history of eye blinking. Mark had no history of hospitalization, head injury, seizures, allergies, or other major medical problems.

MENTAL STATUS EXAMINATION

Mark was a slender, peripubertal boy in age-appropriate casual clothing. He was very restless. There were multiple motor and vocal tics, including eye blinking, shoulder shrugging, facial grimacing, mouth opening, coughing, and grunting. Mark showed a range of affect and was pleasant and cooperative. Mark reported that his tics did not really cause him any problems but he felt worried a lot.

IMPRESSION

In summary, Mark was a 13-year-old peripubertal boy with a long history of multiple motor and vocal tics consistent with Tourette disorder. Given the contamination concerns, obsessional worries and ruminations, and the tendency to count and arrange, Mark also met criteria for OCD. In addition, there were symptoms suggestive of social anxiety disorder, given his reluctance to speak up in class, his slow-to-warm behavior, and a history of shyness. The excessive worry about performance was suggestive of a generalized anxiety disorder.

TREATMENT PLAN

Treatment of Mark's anxiety was indicated because it appeared to be pervasive and was causing distress. The recommendation was to start cognitive-behavioral therapy to address both his social anxiety and obsessive-compulsive behaviors. Given that Mark was neither in distress nor impaired by his tics and was probably past his lifetime peak severity, simple monitoring of the tics was recommended. It was likely the tics would decrease once the anxiety was treated.

TREATMENT

Individuals with mild and even some moderate tics may not need treatment at all, if the tics are not causing significant distress or impairment. Simple, active monitoring may be all that is necessary for some individuals with mild tics. Indications for treatment of tics include physical pain or discomfort, interference with social interactions, and impairment in any aspect of educational or occupational functioning. The first step in treatment is proper education of the patient, family members, teachers, and others who interact with the patient. Parents, educators, and physicians must work as partners to advocate for the optimal school environment for children with tic disorders. National and local support groups, such as the Tourette Association of America (http://www.tourette.org), can provide additional information and can serve as a valuable resource for patients and their families.

Nonpharmacological treatments have been developed that are as successful as pharmacotherapy. In fact, the first-line treatment for mild to moderate tics is cognitive-behavioral intervention. Medications are usually considered only when symptoms are significant, unable to be addressed by psychotherapeutic interventions, and interfere with academic or job performance, peer relationships, social interactions, or activities of daily living. Pharmacotherapy for tic disorders is individualized and tailored specifically to the needs of the patient; the most distressing and/or impairing symptoms are targeted first. The goal of treatment is not to eliminate tics, but to relieve tic-related discomfort or embarrassment and to achieve a degree of control of tics that allows the patient to function as normally as possible.

General Guidelines

Behavioral intervention is the cornerstone of effective treatment for tic disorders. Treatment goals should be to reduce symptoms, support adaptive functioning and strengths, and enhance developmental progress. Self-esteem at all times should be supported, given that most patients with tic disorders will have at least some sense of personal and emotional vulnerability. Early on in the diagnostic process, education as to phenomenology and natural history of the tic disorders is essential.

Clarification that the patient's symptoms are not voluntary in most situations eases psychological burdens for both patients and families. This is especially important because families observe that tics, at times, can be suppressed. Struggles over whether the symptoms are tics or "negative behavior" are pointless and should be curtailed; a more practical goal involves personal management of, and responsibility for, symptoms and behavior, regardless of their origin.

Containment is another cornerstone of treatment. Even with the use of

effective medications, it is rare for tics to remit completely. Use of a "tic room" or a "time out" area provides an opportunity to contain problematic tics or compulsions and to "de-stimulate." Because emotional conflicts and stress frequently increase symptom intensity and frequency, time-limited withdrawal from stressful situations can be beneficial. A balanced, healthy diet and exercise regimen may contribute to stress reduction and overall well-being. Caffeine and excess amounts of sugar should be minimized because they might exacerbate tics in some patients.

Psychoeducational Approaches

Educating patients and their families is essential. Patients and their parents frequently experience relief at the time of diagnosis. There should be ongoing opportunities for discussion of patients' and families' questions and concerns. Referral to the Tourette Association of America and to its local chapters is recommended for all newly diagnosed patients and their families to maintain continuing education and support.

Behavioral Therapies

Behavioral treatment is a primary intervention for many tic disorder patients. Simple behavioral approaches can be used with the majority of Tourette disorder patients to "de-stimulate" and contain the symptoms. Specific paradigms are indicated for obsessions, compulsions, and, more recently, tics. Relaxation techniques, such as deep breathing, guided imagery, and use of relaxation tapes, can be useful for the anxious TD patient.

Habit-reversal therapy or training (HRT) forms the essence of therapeutic intervention in the cognitive behavioral treatment known as comprehensive behavioral intervention for tics (CBIT) in youth and adults. The superiority of CBIT over supportive therapy and psychoeducation alone is also well established; CBIT is now considered first-line treatment for mild to moderate tics when intervention is warranted. The primary components of CBIT include awareness training, competing response training, relaxation training, contingency management, social support, and relapse prevention. Awareness training teaches patients to identify the premonitory urges or sensations that typically precede tics, as well as to recognize the situations in which tics are likely to occur. The second component of CBIT, teaching the patient a competing response to the targeted tic, has been developed to prevent the emergence of the tic. Parents are taught to provide appropriate support and encouragement to the child in practicing the technique at home. Tic substitution (substituting a more socially tolerable tic for a less socially tolerable one) can be useful.

Behavioral interventions are also critical for the management of oppositional, defiant, or explosive outbursts, which are not uncommon in

clinically referred youth with tic disorders. Scahill and colleagues demonstrated that parent management training is helpful in improving disruptive and explosive behavior problems in children with tic disorders. In children and adolescents, cognitive-behavioral paradigms are also the cornerstone of treatment of mood and anxiety disorders, which often co-occur in clinically referred individuals with Tourette disorder.

Psychotherapy

Individual, group, and family therapies are supportive adjuncts to behavioral interventions and pharmacotherapy. Individual supportive therapy is indicated for patients having difficulty adjusting to the disorder, to their peers or family, or with school or occupational functioning. It is particularly useful when there is evidence of moderate stress or anxiety or another comorbid psychiatric disorder responsive to this form of psychotherapy (e.g., mild depression). It can help to restore or maintain self-esteem and promote mastery and coping. Furthermore, supportive psychotherapeutic interventions may facilitate the normal developmental tasks of friendship development, school adjustment, and day-to-day stress management.

Family therapy is extremely useful in dealing with the many complex interpersonal and family issues that arise for children with Tourette disorder. At a minimum, both parents should be involved in the initial evaluation. Ideally, family members can be sources of support and nurturance. Ongoing family therapy is indicated when family development (i.e., growth and maturation) has slowed or halted because of the focus on the patient with the tic disorder. It may also help with maladaptive reactions from or to siblings or with specific symptoms that are affecting the entire family.

Group therapy is another important adjunct treatment. Support and education for parents and family members of patients with Tourette disorder can be found through the Tourette Association, which has created a well-organized network of state and community support groups. Informal, unstructured activity can be arranged through local chapters of the Tourette Association. Formal, structured support is also possible through social skills groups.

School and Occupational Interventions

Learning problems and classroom difficulties occur commonly in tic disorder patients. Specific developmental disorders and ADHD- or OCD-related symptoms may interfere with academic performance.

Clinicians should be available to provide consultation, guidance, and education to teachers or employers of patients with tic disorders. Useful special educational interventions include creating moderate, task-oriented

structure in the classroom, preferential seating, one-to-one support, and individualized educational or work plans. Flexibility is a key element in educational and occupational intervention. To promote tic control, there may be a need for optional "time outs" from class or work settings. Time limits may have to be extended or eliminated for exams. Some patients may also benefit from writing aides, such as laptops. Laptop and tablet computers may allow patients with handwriting difficulties to take notes in class without disturbing others. Adaptive, as contrasted to regular, physical education can be particularly helpful for children with Tourette disorder. This form of physical education can be arranged through special education departments in local schools. Programs can be tailored to each child's needs, strengths, and weaknesses.

Specific workplace interventions include structured tasks, organization of tasks into smaller units, flexible time limits, and ample physical space.

Medications

The decision to treat tics and/or comorbid symptoms with medication should be based on a comprehensive evaluation. Patients with symptoms that significantly interfere with adaptation to family, school or work life, peer relationships, or developmental progress should be treated, especially if psychotherapeutic options are either unavailable or not optimally beneficial. Patients who suffer significant emotional distress or inadvertently cause significant distress to significant others in their lives may also be the candidates for treatment, even when symptoms are relatively mild. Most patients with very mild symptoms need only monitoring, education, guidance, and support. Moderate to severe symptoms usually should be treated aggressively.

The clinical workup before initiating pharmacological treatment should include a complete physical examination and neurological screening, psychiatric evaluation, vital signs including height, weight, and body mass index (BMI), an AIMS screening, basic laboratory screening (including a complete blood count with differential, urinalysis, and a blood chemistry screen including liver and thyroid tests, fasting blood glucose, and metabolic screen including cholesterol and triglycerides), and an electrocardiogram (ECG).

Medication should be initiated at the lowest possible dose, and dose titration should be gradual in general. Most maximum doses will be low (compared to doses needed for other indications of these same medications) in Tourette disorder patients. When possible, a single medication (monotherapy) should be used initially. At times more than one agent will need to be prescribed simultaneously (targeted combined pharmacotherapy) when monotherapy has not been effective or has resulted in limiting adverse effects or when target symptoms include more than one disorder.

Determining an adequate duration for a medication trial for patients with tics can be quite challenging because the natural history of Tourette disorder is characterized by waxing and waning of symptoms over time. A 1-month baseline observation period before treatment is initiated is recommended, if feasible. In general, it is best to wait at least several weeks after changing a dose before making any conclusions about therapeutic response. External stressors must be taken into account at all times as potential determinants of increased symptomatology during medication trials. Comorbid conditions, such as ADHD and OCD, which significantly affect the individual should be targeted aggressively because it is well established that these conditions are often more impairing than tics and that they respond to treatment.

α-Adrenergic Agonists. The α-adrenergic agents are recommended as first-line pharmacological treatment for most patients with mild to moderate Tourette disorder, especially in those with comorbid ADHD. Clonidine (Catapres), an α_2-adrenergic (presynaptic) agonist, has been used for more than 20 years for the treatment of Tourette disorder. When given in low doses, it is postulated that presynaptic noradrenergic effects mediate the observed clinical improvement in motor and vocal tics. Clonidine activates presynaptic autoreceptors in the locus coeruleus to reduce norepinephrine release and turnover in the cerebral cortex. In addition, clonidine also reduces the disinhibition, impulsivity, inattention, and hyperactivity often present in young Tourette disorder patients.

Clonidine should generally be started at 0.025 mg daily to twice daily for prepubertal children and increased by 0.025 mg every 5 to 14 days. Prepubertal patients will generally need dosing three or four times daily. Adolescents or adults can be started on 0.05 mg daily and increased by 0.05-mg increments to twice-daily dosing. The total daily dose typically is 0.05 mg to 0.45 mg (i.e., up to 8.0 μg/kg). Common side effects include sedation, headache, dry mouth, stomachache, mid-sleep awakening, and irritability. Sedation, the most common and sometimes limiting side effect, usually abates over several weeks; when it does not, dose reduction may be helpful. Hypotensive effects are minimal in this dose range, but blood pressure, pulse, and ECG should be monitored at baseline and follow-up visits. More concerning is the risk of rebound increases in blood pressure, anxiety, and tics if abrupt discontinuation of clonidine occurs. A transdermal form (Catapres TTS-1, TTS-2) is available and may be an alternative when patients require very frequent dosing and/or cannot tolerate the sedative effects. A long acting oral formulation is now also available (Kapvay) and designed to be administered twice daily.

Guanfacine (Tenex), an alternative α_2-adrenergic agonist, may also be efficacious for hyperactivity, impulsivity, and tics. Guanfacine activates postsynaptic prefrontal α-adrenergic cortical receptors. Prepubertal

patients are typically started on 0.25 mg daily and increased by 0.25-mg increments every 5 to 14 days given twice daily; older patients and adults are typically started on 0.5 mg daily and increased by 0.5-mg increments to about 3.0 mg daily. Adverse effects, generally less significant and more tolerable than with clonidine, include sedation, mid-sleep awakening, constipation, and hypotension. A long acting, once daily form (Intuniv) is now also available.

Neuroleptic Agents. Since the late 1960s when haloperidol was first introduced as a treatment for Tourette disorder patients, haloperidol (Haldol) and pimozide (Orap) have been the only Food and Drug Administration–labeled, formally approved agents for the treatment of Tourette disorder. These typical neuroleptic medications block postsynaptic D_2 (dopamine) receptors in the basal ganglia, decreasing dopaminergic input from the substantia nigra and ventral tegmentum and thus reducing tics. These medications are generally effective in tic reduction; however, considerable variation exists in clinicians' and patients' thresholds for neuroleptic use. The decision to use a neuroleptic may be guided by both symptom severity and quality-of-life concerns. Significant side effects, including acute dystonic reactions, oculogyric crises, torticollis, parkinsonism, akathisia, tardive dyskinesia, weight gain, sedation, dry mouth, blurred vision, galactorrhea, gynecomastia, constipation, urinary retention, social phobia, loss of drive/energy/personality, and electrocardiographic changes including tachycardia and QTc prolongation, limit their use to second- or third-line treatments. In general, neuroleptic agents are not recommended for mild tics. Of the typical neuroleptics, haloperidol and pimozide have been the most investigated, with double-blind, placebo-controlled studies supporting their use.

The newer, atypical neuroleptics, or second-generation antipsychotic agents, are recommended as first-line treatment when a neuroleptic is indicated for moderate to severe tics. These agents, which block both D_2 dopamine and 5-HT_2 serotonin receptors, have the potential advantage of fewer extrapyramidal side effects (EPSs) (acute dystonic reactions, oculogyric crises, torticollis, parkinsonism, akathisia, tardive dyskinesia). Their dual pharmacological activity appears to protect against the neurological adverse effects associated with typical neuroleptic agents. Among them, risperidone (Risperdal) has been found to be efficacious for tics. Multiple trials, including randomized, controlled trials, support risperidone use. It has been shown to be superior to placebo and equally effective as pimozide and clonidine for tic reduction. Doses are typically in the 1- to 3-mg range. Common side effects include appetite increase, weight gain, lipid and glucose metabolism abnormalities, and sedation. Sleep disturbance, social phobia, galactorrhea, gynecomastia, and erectile

dysfunction have also been reported.

Additional atypical options with some evidence base for efficacy in Tourette disorder include ziprasidone (Geodon) and olanzapine (Zyprexa). Results of one placebo-controlled trial of ziprasidone indicated that it was as efficacious as risperidone for tic reduction. Side effects, including weight gain, are similar. Ziprasidone carries the added risk of altered cardiac conduction, especially QTc prolongation. An ECG has been recommended before initiating treatment, during dose adjustments, on reaching the maintenance dose, and annually thereafter. Evidence for effectiveness of olanzapine is supported by open-label and case-report studies.

Aripiprazole (Abilify) is a novel atypical neuroleptic with partial agonist–antagonist effects on the dopamine and serotonin system. Several open-label trials as well as a placebo-controlled study have been published on its effectiveness and tolerability in Tourette disorder. The most common adverse effects include weight gain, akathisia, and sedation.

Fasting glucose and a lipid panel are recommended before initiating treatment, as well as on reaching the maintenance dose and at regular intervals thereafter, with all atypical neuroleptics. Weight, waist circumference, and BMI should also be monitored. An ECG may be helpful, especially when there is a known individual or family cardiac history.

Neuroleptics should be initiated at the lowest possible doses and increased very gradually. Risperidone is typically started at 0.125 to 0.25 mg daily to twice daily and titrated upward over weeks to about 1 to 2 mg daily in prepubertal children; adolescents are typically started on 0.25 mg daily to twice daily. For haloperidol, children are typically started at 0.125 to 0.25 mg daily to twice daily and increased by 0.125- to 0.25-mg increments weekly or biweekly. Adults are usually started on 0.5 mg with weekly or biweekly increases of 0.5 mg. Pimozide is an alternative D_2--blocking agent, which may cause fewer side effects than haloperidol. However, because pimozide also has calcium-channel–blocking properties, normal cardiac function (ascertained from cardiac assessment and an ECG) is a prerequisite to its use. Pimozide has been shown to prolong QTc intervals, and the risk for cardiac conduction abnormalities may increase when pimozide is combined with drugs that inhibit the cytochrome P450 3A4 isozyme. An ECG is also recommended during dose adjustments, on reaching the maintenance dose, and annually thereafter. Approximately half as potent as haloperidol, pimozide should be started at 0.5 mg to 1.0 mg for children and at about 1.0 mg for adolescents or adults. Most patients require less than 10 mg daily.

Common adverse effects of neuroleptics include weight gain, sedation, muscle cramps, and stiffness. Dose reduction or switch to an alternative drug is the best way to address these side effects. Fortunately, EPSs do not appear to be common sequelae from neuroleptic use at usual doses in tic disorder patients; nevertheless, clinicians should be alert to the possibility

of these reactions and consider appropriate interventions (e.g., dose reduction, anti-Parkinson agents,) on a case-by-case basis. EPSs may be less common with the atypical agents than with the typicals. Tardive dyskinesia, although probably not common, has been described in patients with Tourette disorder and may be difficult to differentiate from tics.

Weight gain is likely with both typical and atypical neuroleptics, but the risk appears to be greater with use of the atypical agents such as olanzapine and risperidone. The possibility of weight gain should be addressed proactively in all youth who are candidates for neuroleptics. Education regarding the need for a healthy diet and regular exercise at the outset of treatment is helpful; at times nutritional consultation may be indicated.

Tiapride and sulpiride are D_2-receptor–blocking neuroleptics used in tic treatment in Europe but not available in the United States. Tetrabenazine, a nonneuroleptic dopamine antagonist that acts by inhibiting central vesicular monoamine transporter type 2 (VMAT2), may be useful for the treatment of tics, but it is currently only approved as an investigational drug and more study is needed.

Selective Serotonin Reuptake Inhibitors. Serotonergic agents such as fluoxetine (Prozac), paroxetine (Paxil), fluvoxamine (Luvox), citalopram (Celexa), escitalopram (Lexapro), and sertraline (Zoloft) are first-line agents for pharmacological treatment of significant OCD, anxiety disorders, and mood symptoms. All of the selective serotonin reuptake inhibitors (SSRIs) except citalopram and escitalopram have been studied in placebo-controlled investigations in children with OCD. Although there is no evidence to support the use of the SSRIs for the suppression of tics, their use may nonetheless be warranted. Because many Tourette disorder patients have obsessions or compulsions and neuroleptics and the α_2-agonists do not effectively alter these specific symptoms, the SSRIs may be beneficial in some patients. In addition, complex motor tics such as repetitive squatting and touching may be very similar to compulsive rituals. Furthermore, stress and anxiety are known to exacerbate tics, and the use of the SSRIs to reduce anxiety may indirectly serve to reduce tics.

Fluoxetine has received the most study in Tourette disorder, but efficacy likely is relatively comparable among serotonergic agents. Choice typically rests with consideration of the side-effect profiles of the agents. Often, higher doses of the SSRIs are required for the treatment of OCD symptoms than for the treatment of anxiety disorders or depression, and many patients with OCD plus a tic disorder respond less well to SSRIs. The addition of small doses of a neuroleptic or an atypical neuroleptic, which increases the response to the SSRIs, may be required.

Mild OCD symptoms often respond to cognitive behavioral therapy alone, but when obsessions or compulsions are of moderate to severe intensity or impair adaptive functioning, SSRIs are potentially indicated.

Fluoxetine typically is started at 2.5 to 5.0 mg daily for children and at 5.0 to 10 mg for adolescents or adults. Dose ranges are not well established for children and adolescents with TD. It appears that most children will respond to 10 to 20 mg daily; adults may require 20 to 80 mg daily. Fluoxetine in combination with a neuroleptic can be useful for both tics and obsessive-compulsive symptoms; neuroleptic doses may require reduction or adjustment because of the cytochrome P450 enzyme–inhibiting properties of fluoxetine. Common side effects include activation, insomnia, headaches, anorexia, and weight loss.

Sertraline, fluvoxamine, citalopram, and escitalopram may also be beneficial, but experience with them is limited. There are more data available on the use of clomipramine (Anafranil) in OCD, but adverse effects may be more problematic. These medications may represent alternatives for patients who experience limiting side effects on fluoxetine. Current recommendations for the treatment of comorbid tics and OCD are to use the appropriate agent indicated for each symptom. For example, one may combine an α-agonist or atypical neuroleptic with an SSRI to target significant tics and OCD symptoms.

Tricyclic Antidepressants. Clomipramine should be started at 10 to 25 mg daily for children and 25 to 50 mg for adolescents or adults when used for the treatment of OCD. Total dose ranges are typically 100 to 300 mg daily; blood levels may be obtained, but no therapeutic range has been established. Common side effects are those seen typically with other tricyclic antidepressants, including sedation, dry mouth, constipation, blurring of vision, weight gain, tachycardia, and ECG changes.

Some studies of desipramine (Norpramin) have suggested potential efficacy for treatment of tics as well as ADHD symptoms in patients with comorbid ADHD. Doses are not established, but may parallel doses used in ADHD without tics. Careful cardiovascular monitoring must take place, including blood pressure, pulse, and baseline and follow-up ECGs (Table 48–4).

Benzodiazepines. Benzodiazepines have not been well investigated in tic treatment, but may hold some promise (e.g., clonazepam [Klonopin]). Some studies suggest a reduction in tics independent of anxiolytic effects. This effect may also result from a reduction in anxiety that, in turn, could reduce tic frequency. In addition, benzodiazepines can be used to treat comorbid anxiety symptomatology or panic attacks that can also occur in Tourette disorder patients. Doses are not established, but are often quite low (e.g., 0.125 mg daily to twice daily for clonazepam). Adverse effects of particular concern in the juvenile population are sedation, disinhibition, cognitive impairment, and dependence.

Stimulants. Stimulants are first-line agents for the treatment of ADHD. There was controversy in the last decade about the role of stimulants in Tourette disorder and whether their use causes or exacerbates tics; however, recent studies have demonstrated that TD patients with significant ADHD symptoms may be candidates for stimulants. In the first systematic investigation of methylphenidate use, by Gadow and colleagues, behavior improved and tics did not worsen at moderate doses of 0.1 to 0.5 mg/kg methylphenidate. When group data are analyzed, there is no significant increase in tics when stimulants are used in patients with tic disorders; however, individual patients may experience at least a transient increase in tics. In some studies, stimulant use is associated with a decrease in tics. Stimulants are equally effective in improving ADHD symptoms with or without comorbid tics. A randomized, controlled study conducted by the Tourette Syndrome Study Group of clonidine and methylphenidate in ADHD and chronic tics reported that the combination and each drug alone was more effective than placebo in reducing both tics and ADHD symptoms. Increase in tics did not occur in the methylphenidate group to any extent more significantly than in the placebo group. Finally, in a recent meta-analysis examining pharmacotherapy for youth with comorbid ADHD and tics, methylphenidate, α_2 agonists, desipramine, and atomoxetine (Strattera) were effective for ADHD symptoms. Supra-therapeutic doses of dextroamphetamine appeared to worsen tics. There was no short-term evidence that methylphenidate exacerbated tics. A judicious approach to the treatment of ADHD in the context of a tic disorder is to start with a methylphenidate-based stimulant, in low doses, and gradually increase, monitoring for any exacerbation of tics. Common adverse effects of stimulants include insomnia, appetite suppression, and weight loss.

Nonstimulant ADHD Treatment: Atomoxetine. Atomoxetine, a selective norepinephrine reuptake inhibitor, is approved for treatment of ADHD in children, adolescents, and adults. It has also been studied in youth with ADHD and chronic tics. One recent study reported that atomoxetine improved both ADHD symptoms and tics in children and adolescents with comorbid ADHD and Tourette disorder. It is dosed based on body weight and titrated from a daily starting dose of 0.5 mg/kg to 1 to 1.5 mg/kg. Therapeutic response, unlike stimulants, takes a few weeks. Common adverse effects include headache, nausea, fatigue, and stomach aches.

Table 48–4.
Pharmacotherapies for Tourette and Tic Disorders

Class/Medication	Typical Range (mg)	Starting Dose (mg)	Maximum Dose[a] (mg)	Common Side Effects

Neuroleptics				
Haloperidol	0.25–5.0	0.25–0.5	5.0	Sedation, weight gain, dysphoria, extrapyramidal effects, dry mouth, blurred vision, constipation, urinary retention
Pimozide	1.0–10	0.5–1.0	15–20	Sedation, weight gain, dysphoria, extrapyramidal effects, dry mouth, blurred vision, constipation, urinary retention, ECG changes (QTc prolongation)
Risperidone	1.0–3.0	0.25–0.5	4–6	Weight gain, sedation, lipid and glucose metabolism abnormalities, extrapyramidal effects
Ziprasidone	20–80	5–10	120–160	Weight gain, sedation, lipid and glucose metabolism abnormalities, extrapyramidal effects, ECG changes (QTc prolongation)
Aripiprazole	2.5–15	1.25–2.5	20	Akathisia, sedation, agitation, weight gain, lipid and glucose metabolism abnormalities, extrapyramidal effects
Partial α_2-adrenergic agonists				
Clonidine	0.05–0.45	0.025–0.05	0.45	Sedation, headaches, insomnia, stomachaches, hypotension, dry mouth, midsleep awakening, irritability
Guanfacine	1.0–3.0	0.25–0.5	3–4	Sedation, headaches, hypotension, midsleep awakening, constipation
Tricyclic antidepressants				
Desipramine	25–300	10–25	250–300	Sedation, dry mouth, weight gain, tachycardia, constipation, blurred vision, ECG changes
Clomipramine[b]	50–200	10–25	250–300	Sedation, dry mouth, weight gain, tachycardia, constipation, blurred vision, ECG changes
Selective serotonin reuptake inhibitors[b]				
Fluoxetine	5.0–40	2.5–10	60–80	Insomnia, activation, anxiety, anorexia, weight loss, headaches, gastrointestinal distress, sexual dysfunction; suicidality
Sertraline	25–200	12.5–25	200–300	Insomnia, activation, headaches, gastrointestinal distress, sexual dysfunction; suicidality
Paroxetine	10–40	5–10	50–60	Insomnia, activation, headaches, gastrointestinal distress, sexual dysfunction; suicidality
Fluvoxamine	50–300	12.5–25	300	Insomnia, activation, headaches, gastrointestinal distress, sexual dysfunction; suicidality
Citalopram	5–40	2.5–5	60	Insomnia, activation, headaches, gastrointestinal distress, sexual dysfunction; suicidality
Escitalopram	2.5–20	2.5	20	Insomnia, activation, headaches, gastrointestinal distress, sexual dysfunction; suicidality

ECG, electrocardiogram.

[a]These are usual maximum doses for adults.

[b]These medications have been found to be effective in patients with obsessive-compulsive disorder.

Other Alternatives. Many other medical treatments have been used in Tourette disorder patients. Immune therapies such as IVIG, plasmapheresis, and corticosteroids have been explored in patients with

PANDAS/PANS. These remain experimental treatments under study. Opioid antagonists (e.g., naltrexone) may be an alternative for patients with self-injurious behaviors. The dose range for naltrexone may be approximately 25 to 75 mg per day. Nicotine, both in gum and in the transdermal form, has also been studied in Tourette disorder. Results have been equivocal. In a double-blind, placebo-controlled trial of 61 patients with Tourette disorder at doses up to 7.5 mg daily, mecamylamine (Inversine), a nicotine antagonist, was found to be no more effective than placebo.

Baclofen, a GABAergic muscle relaxant, has been studied in one large open-label and one controlled trial. In the controlled trial, there was no difference between drug and placebo in total tic score, but baclofen was superior to placebo with regard to tic-related impairment on the Yale Global Tic Severity Scale. A placebo controlled study of omega-3 fatty acids derived from fish oil revealed no improvement in tic measures, but clinically meaningful improvement in tic-related impairment compared to placebo.

The use of botulinum toxin to suppress tics by injection into discrete muscles involved in tic generation has been examined in several studies. Botulinum toxin blocks acetylcholine release at the neuromuscular junction and produces a temporary, reversible diminution of muscle activity, which may last weeks to months for dystonic tics. Open-label and placebo-controlled trials have demonstrated efficacy for isolated tics. Botulinum toxin may be the most useful agent for the treatment of solitary, focal dystonias. There are reports of suppression of ear-wiggling tics and laryngeal tics. Adverse effects include soreness at the injection site and muscle weakness.

Targeted combined pharmacotherapy—the specific combination of two or more medications at one time—has evolved as a treatment option. Ideally, this approach should be used only after adequate trials of monotherapy have failed. However, many patients with Tourette disorder also have comorbid ADHD or OCD, and often both, and frequently monotherapy addresses one problem but not the remaining disorders. Judicious use of SSRIs and neuroleptics for Tourette disorder and comorbid OCD is one example of this approach; in some patients with OCD and tics, this approach may be more effective than monotherapy. Another example is the combination of methylphenidate and clonidine to target tics and ADHD symptoms. In general, for comorbid tic disorders and OCD or ADHD, each condition should be treated with its respective first-line option. A randomized, controlled trial of pramipexole, a mixed dopamine agonist with presynaptic autoreceptor effects, failed to show efficacy.

Surgical Treatment

Initial reports on surgical interventions for Tourette disorder appeared in

the 1960s. Stereotaxic neurosurgical techniques have since been used in a few cases of severely impaired adults. Bilateral capsulotomy was associated with at least 50 percent improvement in both motor and vocal tics in 12 patients who had Tourette disorder, with no severe adverse effects. In 1999, deep brain stimulation (DBS) was introduced for the treatment of intractable tics. Several patients have since undergone bilateral thalamic stimulation with promising results. However, given the significant medical risks for these procedures, there has been limited experience with use of these techniques in youth, and they have generally been reserved for severe and refractory tics that significantly interfere with function and persist into adulthood.

Transcranial magnetic stimulation involves generating a brief, powerful magnetic field by a small coil positioned over the skull, inducing an electrical current in the brain. This noninvasive approach has been demonstrated to effect long-term changes in cortical excitability, which may be abnormal in tic disorders. It is an exciting potential future treatment for tics. Neural stem cell transplantation is another proposed treatment for tics, and animal studies are underway.

Jack was a 14-year-old boy referred by his mother for management of his Tourette disorder. Jack developed the onset of skipping and clapping episodes when he was approximately age 7 years. This was followed by stretching movements of his mouth and nose accompanied by facial grimaces. His mother reported that Jack's tics remitted entirely until just before his 12th birthday when he began to "give the finger" for no apparent reason, and began to curse using the "F" word. Children often teased him, and he was in significant distress about this.

Jack was the product of a full-term, uncomplicated pregnancy; developmental milestones were described as within normal limits. Jack's parents divorced when he was 2 years of age. Jack attended private school from pre-kindergarten through, at the time of referral, the 9th grade. He had difficulty with mathematics. He had a group of friends, including a best friend, and enjoyed music and video games.

Jack's maternal grandmother had OCD, and maternal great uncles had eye-blinking tics. Jack had no history of head injury or seizures. Electroencephalographic and computed tomography scans were within normal limits. Past medication included adequate trials of clonidine, guanfacine, risperidone, and ziprasidone. Medications at the time of referral included haloperidol, 4 mg, which led to excessive sedation, including somnolence in school.

MENTAL STATUS EXAMINATION

Jack was a moderately overweight adolescent boy. Of note were multiple motor tics, including pelvic thrusting, hand movements, outward leg twisting, head and neck thrusting, eye blinking, abnormal inhalation, chin hitting, shoulder licking, and genitalia scratching. In addition, there was copropraxia in the form of "giving the finger." There was also a change in his vocal inflection intermittently and a tendency to yawn on occasion.

IMPRESSION

In summary, Jack was a 14-year-old adolescent referred for management of his Tourette disorder. His clinical picture was consistent with moderately severe Tourette disorder. He had been treated with a variety of tic medications that were ineffective or intolerable.

TREATMENT PLAN

The recommendation was for a trial of aripiprazole, which is not formally approved for treatment of Tourette disorder. In addition, a trial of comprehensive behavioral intervention for tics was recommended. Jack and his parents were also counseled that at age 14 it was

possible he was seeking treatment currently at his lifetime peak tic severity, and tic symptoms might begin to attenuate over the next few years.

FUTURE DIRECTIONS

The French neurologist Georges Gilles de la Tourette in 1885 was among the first to recognize and describe the neurodevelopmental disorder that now bears his name, as well as its psychiatric comorbid conditions. Since the first case descriptions of Gilles de la Tourette, a spectrum of tic disorders has come to be recognized and codified in psychiatric nomenclature. Tic disorders are complex neurodevelopmental disorders with heterogeneous phenotypes characterized by core disinhibition mediated by the CSTC tracts. Tic disorders are genetically determined in most cases. There is a growing appreciation and understanding of the role of psychiatric comorbidity in the diverse clinical presentation, and it is often the neurobehavioral, occupational, and psychosocial problems that co-occur in clinically referred individuals that warrant treatment rather than the tics per se. Treatment strategies should be multimodal in nature and include cognitive-behavioral, individual and family intervention, educational monitoring and, where indicated, pharmacotherapy. Future investigation is needed to better understand the phenomenology, developmental psychopathology, and novel approaches to the treatment of tics. Tremendous progress has been made in the last few decades through the use of innovative neuroimaging techniques, genetic studies, and clinical interventions. Understanding of the etiology, genetics, neuroanatomy, and epidemiology of tic disorders has been significantly enhanced. Novel treatment approaches have also emerged, providing alternatives to the pharmacological management of tics. Targeted behavioral intervention has been proven to be as effective as medication and is now first-line treatment. Deep brain stimulation, botulinum toxin, targeted surgery, and stem cell transplantation are examples of promising new treatments. The future holds tremendous promise as new insights into the etiology, pathophysiology, and treatment of tic disorders are rapidly developing.

REFERENCES

Abelson JF, Kwan KY, O'Roak BJ, Baek DY, Stillman AA. Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science*. 2005;310:317–320.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

Bloch MH, Leckman JF, Zhu H, Peterson BS. Caudate volumes in childhood predict symptom severity in adults of Tourette syndrome. *Neurology*. 2005;65:1253–1258.

Bohlhalter S, Goldfine A, Matteson S, Garraux G, Hanakawa T. Neural correlates of tic generation in Tourette syndrome: an event-related functional MRI study. *Brain*. 2006;129:2029–2037.

Budman CL, Bruun RD, Park KS, Lesser M, Olson M. Explosive outbursts in children with Tourette's disorder. *J Am Acad Child Adolesc Psychiatry*. 2000;39:1270–1276.

Castellanos FX, Giedd JN, Elia J, Marsh WL, Ritchie GF. Controlled stimulant treatment of ADHD and comorbid Tourette's syndrome: effects of stimulant and dose. *J Am Acad Child Adolesc Psychiatry*. 1997;36(5):589–596.

Coffey BJ, Biederman J, Geller DA, Spencer TJ, Kim GS. Distinguishing illness severity in children and adolescents with Tourette's disorder. *J Am Acad Child Adolesc Psychiatry*. 2000;39:556–561.

Ercan-Sencicek AG, Stillman AA, Ghosh AK, et al. L-histidine decarboxylase and Tourette's syndrome. *N Engl J Med*. 2010;362:1901–1908.

Erenberg G. The relationship between Tourette syndrome, attention deficit hyperactivity disorder, and stimulant medication: a critical review. *Semin Pediatr Neurol*. 2006;12:217–221.

George MS, Sallee FR, Nahas Z, Oliver NC, Wassermann EM. Transcranial magnetic stimulation (TMS) as a research tool in Tourette syndrome and related disorders. *Adv Neurol*. 2001;85:225–235.

Gorman DA, Thompson N, Plessen KJ, Robertson MM, Leckman JF, Peterson BS. Psychosocial outcome and psychiatric comorbidity in older adolescents with Tourette syndrome: controlled study. *Br J Psychiatry*. 2010;197(1):36–44.

Hardesty DE, Sackheim HA. Deep brain stimulation in movement and psychiatric disorders. *Biol Psychiatry*. 2007;61:831–835.

Houghton DC, Capriotto MR, Conelea CA, Woods DW. Sensory phenomena in Tourette syndrome: their role in symptom formation and treatment. *Curr Dev Disord Rep*. 2014;1:245–251.

Kawikova I, Leckman JF, Kronig H, Katsovich L, Bessen DE. Decreased number of regulatory T cells suggests impaired immune tolerance in children with Tourette syndrome: a preliminary study. *Biol Psychiatry*. 2007;61:273–278.

Knight T, Steeves T, Day L, Lowerison M, Jette N, Pringsheim T. Prevalence of Tic disorders: a systemic review and meta-analysis. *Pediatr Neurol*. 2012;47:77–90.

Leckman JF, King RA, Gilbert DL, et al. Streptococcal upper respiratory tract infections and exacerbations of tic and obsessive-compulsive symptoms: a prospective longitudinal study. *J Am Acad Child Adolesc Psychiatry*. 2011;50:108–118.e3.

*Leckman JF, Zhang H, Vitale A, Lahnin F, Lynch K. Course of tic severity in Tourette syndrome: the first two decades. *Pediatrics*. 1998;102:14–19.

Mell LK, Davis RL, Owens D. Association between streptococcal infection and obsessive-compulsive disorder, Tourette's syndrome, and tic disorder. *Pediatrics*. 2005;116:56–60.

Miguel EC, Baer L, Coffey BJ, Rauch SL, Savage CR. Phenomenological differences appearing with repetitive behaviors in obsessive-compulsive disorder and Gilles de la Tourette's syndrome. *Br J Psychiatry*. 1997;170:140–145.

Perrin EM, Murphy ML, Casey JR, Pichichero ME, Runyan DK. Does group A beta-hemolytic streptococcal infection increase risk for behavioral and neuropsychiatric symptoms in children? *Arch Pediatr Adolesc Med*. 2004;158:848–856.

Peterson BS, Choi HA, Hao X, Amat JA, Zhu H. Morphologic features of the amygdala and hippocampus in children and adults with Tourette syndrome. *Arch Gen Psychiatry*. 2007;64:1281–1291.

Piacentini J, Woods DW, Scahill L, et al. Behavior therapy for children with Tourette disorder: a

randomized controlled trial. *JAMA*. 2010;303:1929–1937.

Sallee FR, Nesbitt L, Jackson C, Sine L, Sethuraman G. Relative efficacy of haloperidol and pimozide in children and adolescents with Tourette's disorder. *Am J Psychiatry*. 1997;154:1057–1062.

*Scahill L, Erenberg G, Berlin CM Jr, Budman C, Coffey BJ, Tourette Syndrome Association Medical Advisory Board. Practice Committee: Contemporary assessment and pharmacotherapy of Tourette syndrome. *NeuroRx*. 2006;3:192–206.

Scahill L, Sukhodolsky DG, Bearss K, Findley D, Hamrin V. Randomized trial of parent management training in children with tic disorders and disruptive behavior. *J Child Neurol*. 2006;21:650–656.

Singer HS, Wendlandt J, Krieger M, Giuliano J. Baclofen treatment in Tourette syndrome: a double-blind, placebo-controlled, crossover trial. *Neurology*. 2001;56:599–604.

Spencer TJ, Sallee FR, Gilbert DL, Dunn DW, McCracken JT. Atomoxetine treatment of ADHD in children with comorbid Tourette syndrome. *J Atten Disord*. 2008;11:470–481.

Storch EA, Lack CW, Simons LE, Goodman WK, Murphy TK, Geffken GR. A measure of functional impairment in youth with Tourette's syndrome. *J Pediatr Psychol*. 2007;32:950–959.

*Swain JE, Scahill L, Lombroso PJ, King RA, Leckman JF. Tourette syndrome and tic disorders: a decade of progress. *J Am Acad Child Adolesc Psychiatry*. 2007;46:947–968.

*The Tourette Syndrome Association International Consortium for Genetics. Genome scan for Tourette disorder in affected-sibling-pair and multigenerational families. *Am J Hum Genet*. 2007;80:265–272.

The Tourette's Syndrome Study Group. Treatment of ADHD in children with tics: a randomized controlled trial. *Neurology*. 2002;58:527–536.

49

Elimination Disorders

EDWIN JENS MIKKELSEN, M.D.

ENURESIS

The disorder of enuresis has been recognized for centuries. However, substantial progress in the understanding of the disorder did not begin to occur until there were effective treatments. The first effective treatment was the application of the bell-and-pad method of conditioning, which was subsequently followed by pharmacological treatment in the form of imipramine (Tofranil) and more recently desmopressin acetate (DDAVP). Each of these advances has led to etiological research that might explain the efficacy of the particular intervention. The amount of interest in the disorder, as assessed by the number of published articles, has been relatively constant over the last several years. At the beginning of this decade, a literature search using the key terms *enuresis and children* (0 to 18 years), with a time frame of the prior 10 years, yielded a total of 822 citations. A recent search performed with the same parameters produced 785 citations. This chapter reviews the accumulated clinical knowledge concerning enuresis, as well as the emerging research that may lead to a greater understanding of the disorder.

Definition

Enuresis is defined as the involuntary or intentional voiding of urine. The severity is determined by the frequency of urination; quantity is not a diagnostic consideration per se. Quantity can become a factor in treatment decisions if a child emits only small quantities of urine; in actual practice, quantity usually does not figure heavily into the treatment plan. Frequency, however, can be important in planning a hierarchy of treatment approaches. The major diagnostic qualifier relates to age of onset. Finally, the definition precludes a physical cause for the disorder.

The length of time before continence is considered established varies in the literature between 6 months and 1 year. The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) requires a chronological age of 5 years, or alternately, a developmental level equivalent to that age. In practical clinical terms, children who have never maintained continence for more than 1 year are referred to as having *primary enuresis*, whereas children who have achieved continence for 1

year or longer and then resume wetting are referred to as having *secondary enuresis*. The DSM-5 requires a frequency of at least 2 episodes a week for 3 consecutive months. However, it allows an exception for distress or impairment in functions. This means that an individual can receive the diagnosis if they are deemed to experience significant distress and do not meet the frequency criteria.

The disorder is also further defined according to the timing of the episodes throughout the day. Episodes occurring only at night are referred to as *nocturnal*; daytime wetting is labeled *diurnal*. The term "nocturnal and diurnal" refers to a combination of these two subtypes. Most children display only nocturnal enuresis, but some manifest diurnal or nocturnal and diurnal patterns.

History

A scholarly review of the history of enuresis found references dating back about 3,550 years to the Papyrus Ebers. The term derives from the Greek word *enourein*, which means "to void urine." In retrospect, many of the treatment approaches that were used in the past now appear almost sadistic in nature. Although some of these interventions may have been misguided but well-intentioned therapies (as in the use of leeching for other disorders), they may also reflect the commonly held view that the disorder is somehow under the (direct or indirect) control of the child and that punishment would lead to its cessation.

A confluence of somewhat disparate observations contributed to the evolution of the current understanding of enuresis. Although each of these is discussed in more detail in the following sections, a brief reference to them here will provide an overview of the various lines of research and their relative contributions to the present understanding of the disorder. The long-standing observation that enuresis tends to run in families evolved into the observation that a positive family history was a positive prognostic indicator for response to pharmacological interventions and, ultimately, to sophisticated research with large pedigrees pointing toward specific genetic loci and modes of transmission. The observation that imipramine was an effective treatment for enuresis ushered in the modern era of research because the efficacy of imipramine suggested biological contributions to the etiology of the disorder. Imipramine has now been largely supplanted by DDAVP as a pharmacological treatment modality, and related research into its mechanism of action has contributed interesting insights into the pathophysiology of the disorder. Juxtaposed with the research into pharmacological treatment and organic contributions to the disorder has been the impressive literature, which now dates back several decades, indicating that behavioral treatment in the form of the bell-and-pad method of conditioning is equally effective as the two leading pharmacological interventions for the disorder. The

observation that the wetting episodes usually occur at night also led to decades of research investigating the relation between sleep architecture and the occurrence of enuretic events. Although none of these lines of research has provided a parsimonious explanation, each has provided important contributions to the evolving understanding of the disorder. The accumulating literature with regard to enuresis increasingly suggests that a single unified theory of enuresis that applies to all children will not be identified.

Comparative Nosology

The variety of differing nosological and diagnostic schemas seen with many disorders does not exist for enuresis. The behavior is concrete and definite: The child wets or does not. That criterion makes it possible to integrate research data from different countries, cultures, and treatment centers with a greater reliability than is possible for many other disorders. The efficacy of a given treatment is reported in terms of its effect on the frequency of enuretic events. The primary differences with regard to diagnosis concern the frequency required to make a diagnosis of enuresis as a pathological state and the period of continence necessary to separate primary from secondary enuresis, with a duration of 6 months occasionally being used instead of 1 year. Only slight differences exist between the tenth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) and the DSM-5. The ICD-10 states that enuresis "would not ordinarily be diagnosed in a child under the age of 5 years or with a mental age under 4 years." This is slightly different from the DSM-5 cutoff of 5 years for chronological and mental age. However, Tables 49–1 and 49–2 indicate the same 5-year cutoff for chronological and mental age as is described in the DSM-5. The ICD-10 also notes a change in required frequency of enuresis after 7 years of age (at least twice a month before and once a month after). Recent research has suggested there may be a meaningful clinical distinction between monosymptomatic and polysymptomatic enuresis. The latter term refers to those children who not only have nocturnal enuresis, but also have multiple episodes of wetting, daytime wetting, soiling, and signs of bladder hyperactivity.

Table 49–1.
ICD-10 Diagnostic Criteria for Nonorganic Enuresis

A. The child's chronological and mental age is at least 5 years.
B. Involuntary or intentional voiding of urine into bed or clothes occurs at least twice a month in children younger than 7 years of age and at least once a month in children 7 years of age or older.
C. The enuresis is not a consequence of epileptic attacks or of neurological incontinence and is not a direct consequence of structural abnormalities of the urinary tract or any other nonpsychiatric medical condition.
D. There is no evidence of any other psychiatric disorder that meets the criteria for other ICD-10

categories.

E. Duration of the disorder is at least 3 months.

Reprinted with permission from World Health Organization. *The ICD-10 Classification of Mental and Behavioral Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization; 1993.

Epidemiology

The first comprehensive data concerning the incidence of enuresis came from the Isle of Wight Study. They found that 15.2 percent of boys (7 years of age) were enuretic less than once a week, whereas 6.7 percent were wetting at least once a week or more. In 7-year-old girls, the frequency was 12.2 percent for those wetting less than once a week and 3.3 percent with a frequency of once a week or more. By 9 and 10 years of age, 6.1 percent of the boys were wetting less often than once a week, and 2.9 percent were wetting once a week or more. The corresponding figures for 9- and 10-year-old girls were 3.5 percent wetting less than once a week and 2.9 percent wetting once a week or more. The frequency for boys decreased considerably by 14 years of age, at which time 1.9 percent were wetting less than once a week, and 1.1 percent were wetting at a frequency of once a week or more. The rates also decreased for girls by 14 years of age, with 1.2 percent wetting less than once a week and only .5 percent wetting once a week or more.

Table 49–2.
ICD-10 Diagnostic Criteria for Nonorganic Encopresis

A. The child repeatedly passes feces in places that are inappropriate for the purpose (e.g., clothing or floor), involuntary or intentionally. (The disorder may involve overflow incontinence secondary to functional fecal retention.)
B. The child's chronological and mental age is at least 4 years of age.
C. There is at least one encopretic event per month.
D. Duration of the disorder is at least 6 months.
E. There is no organic condition that constitutes a sufficient cause for the encopretic events.

Reprinted with permission from World Health Organization. *The ICD-10 Classification of Mental and Behavioral Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization; 1993.

A New Zealand study involving 8-year-old children found the prevalence of nocturnal enuresis to be 7.4 percent, with 3.3 percent labeled primary and 4.1 percent labeled secondary. A similar large Scandinavian study involving 7-year-old children reported a prevalence of 9.8 percent, with the highest prevalence (6.4 percent) involving children with nocturnal enuresis, 1.8 percent with diurnal enuresis, and 1.6 percent with mixed nocturnal and diurnal enuresis. A report from Australia described an incidence of 5.1 percent for children in the range from 5 to 12 years of age. A review of the published epidemiological literature concerning enuresis

indicates remarkably consistent findings from different countries and cultures, even though the research methods are somewhat variable. A standard consensus format that defines uniform reporting parameters would greatly enhance the epidemiological literature base.

The epidemiological literature is consistent in confirming that enuresis is a self-limiting disorder that will eventually spontaneously resolve. Common point prevalence figures are in the range of 7 percent of boys and 3 percent of girls at 5 years of age, dropping to 3 percent of boys and 2 percent of girls by 10 years of age. Only 1 percent of boys still wet at age 18 years of age, and still fewer girls wet at this age. A spontaneous remission rate of between 5 and 10 percent per year after 5 years of age is usually cited. Secondary enuresis may occur at any time, but most commonly it begins between 5 and 8 years of age.

Etiology

Inclusion in the DSM-5 (as well as previous editions) definition of enuresis of children who wet intentionally with those whose enuresis is involuntary can be problematic. The vast majority of enuretic children do not wet intentionally or even on a subconsciously motivated basis. Increasingly, the research is pointing toward causal factors that may involve irregularities in physiological processes. Clearly, children who wet intentionally are in a different phenomenological grouping than those who do so on an involuntary basis, even if they do meet the other diagnostic requirements. The most likely explanations for voluntary intentional wetting are an oppositional defiant disorder (ODD) or a psychotic disorder. There are also a small number of children who originally have enuretic events involuntarily and subsequently manifest the behavior on a voluntary learned basis as well.

There is a correlation between enuresis and psychological disturbance that increases with age. This association is even more significant for enuresis, which persists into adolescence. Children living in socially disadvantaged situations and experiencing psychosocial stress have a greater frequency of enuresis than those who are not. The type and range of behavioral disturbance seen in children with enuresis are broad, and no marker can reliably differentiate between behaviorally disturbed and nondisturbed enuretic children. Thus, the associated behavioral disturbances are nonspecific and may represent a coincidental or secondary relationship rather than a causal correlation. Further supporting a secondary correlation are the repeated findings that children with enuresis have significantly more developmental delays and minor neurological dysfunction than nonenuretic children, compared to controls and other children attending a psychiatric clinic. Reports of delayed bone development in children with enuresis have also been linked to a possible general delay of maturation as a common underlying factor. However, new

studies refute this finding, and no clear conclusion can be drawn at this time. The most recent research into a general developmental delay contributing to enuresis has used evoked potentials and electromyographic recording of startle eye-blink performance. Separate investigations using these methods have pointed to delayed maturation of brainstem functions.

Another approach to this etiological question is represented by a large, longitudinal, population-based study in the United Kingdom, which compared the intelligence (as assessed by the Wechsler Intelligence Scale for Children–Third Edition [WISC-III]) of controls with children who still exhibited bedwetting in the age range of 7 years and 6 months to 9 years and 3 months. The authors reported that the bedwetting was associated with lower intelligence quotient (IQ) even after children with an IQ of less than 70 were excluded. However, this finding was primarily related to lower scores on the Performance section of the WISC-III. This observation led the authors to conclude that enuresis may be related to deficits in the maturation of the central nervous system.

One comprehensive effort reviewed the psychodynamic literature concerning enuresis and encopresis to define the concepts conveyed in that literature and then to determine empirically how often those concepts were verified and supported by the analysis of clinical material. The literature review yielded 24 generalizations that were consistently applied to enuretic and encopretic children. In subsequent statistical analysis of the clinical material, only 2 of the 24 generalizations reached statistical significance. Another recent study in this regard used a standardized behavioral rating scale, the Eyberg Child Behavior Inventory (ECBI), which was applied to children with primary nocturnal enuresis (PNE) (N = 92), a nonclinical control sample (N = 92), and a clinical sample (N = 92). The results indicated that the children with PNE had higher scores than the nonclinical control sample but significantly lower scores than the clinical sample. However, the mean score for the children with PNE was actually lower than the clinical cutoff score that had been determined for the ECBI.

Related to the association between behavioral disturbance and PNE are studies that address the frequency of diagnostic comorbidity. The most commonly reported comorbid diagnosis is attention-deficit/hyperactivity disorder (ADHD). In general, studies have reported that the co-occurrence of ADHD and PNE is approximately 30 percent greater than the frequency that would be expected from chance alone. A large contemporary study found that 32 percent of children diagnosed with ADHD also met diagnostic criteria for enuresis, whereas only 14 percent of a control group met the criteria. This finding was statistically significant at the $p < 0.001$ level. The rates of primary and secondary enuresis were not significantly different in the two groups. This study then investigated the degree of psychopathology and impairment in the ADHD children (with and without comorbid enuresis) and found no significant differences. Similar findings

were reported for the rates of psychopathology when the control children with enuresis were compared to those without the condition. Thus, the diagnosis of enuresis did not appear to increase the risk of psychopathology for the controls or the degree of impairment in the children who also had a diagnosis of ADHD. However, there was a higher incidence of learning disabilities and impaired academic performance in the control children with enuresis when compared to the control children without enuresis. This difference in the incidence of learning disabilities was not found when the ADHD children with enuresis were compared to those without the comorbid disorder. Thus, the co-occurrence of enuresis did not appear to increase the risk of learning disabilities in children with ADHD.

Another investigation used a family pedigree methodology to study the transmission of enuresis in the relatives of index family members who were diagnosed with PNE and in control probands who were diagnosed with ADHD as compared to those without this diagnosis. The results indicated that the two disorders were transmitted independent of one another.

Enuresis has been noted to occur statistically more frequently in the premorbid histories of children who develop early adolescent bipolar I disorder as compared to controls. An increased incidence of enuresis and other childhood psychiatric disorders has also been found in the history of adults with bipolar disorder and is associated with earlier onset of the disorder.

The association between behavioral disturbance and enuresis is stronger with secondary enuresis. This has been demonstrated by controlled studies and anecdotal reports of an association between the onset of enuresis and the loss of a parent through divorce or death. However, it should be noted that a large study from the Netherlands reported similar rates of psychopathology in children with primary enuresis as compared to those with secondary enuresis. The prevalence of psychodynamic factors in children with secondary enuresis would intuitively make sense because they have, by definition, demonstrated the physiological competence to maintain continence for a prolonged period of time before losing that ability. The most specific study of this association indicated that exposure to four or more stressful life events in a year and delayed acquisition of nocturnal continence were risk factors for the development of secondary enuresis. An association between the occurrence of traumatic events and the development of secondary enuresis has also been demonstrated.

There has been one innovative attempt to explore the possible relation between physiological and psychological factors. Specifically, it was postulated that children with enuresis who did not present with a concomitant behavioral disturbance would be those with dysfunctional or abnormal bladders, whereas those with a behavioral disturbance would be more apt to have normal bladders. The study found the converse to be true:

The group with behavioral disturbance had more dysfunctional bladders with lower volumes. The behaviorally disturbed group also had a greater number of developmental delays. A subsequent investigation reported no difference in bladder size among controls, children with PNE, and those who had been diagnosed with PNE but who were now continent. It also was demonstrated by two independent investigations that fluid loading can precipitate enuresis in children who have no prior history of enuretic events. The most recent investigations into this issue have used ultrasound techniques, which investigate both the volume capacity of the bladder and the thickness of the bladder wall in both controls and children with enuresis. This research suggested three subtypes: Large-capacity bladder with thick wall; normal-capacity bladder with normal wall thickness; and low-capacity bladder with thick wall. The children with enuresis were then assessed with regard to their response to a treatment with DDAVP. The results indicated that the children with the two abnormal bladder subtypes were less responsive to treatment than those with the normal bladder physiological characteristics.

The observation that enuresis tends to occur in family members has been known for decades. A large Scandinavian study involving more than 3,000 children found that a child's risk for developing enuresis was 5.2 times greater if the mother also had the disorder and 7.1 times greater if the father had the disorder, lending further support to a genetic influence. Within the last few years, there have been tremendous advances in the study of the genetic transmission of enuresis, which suggests that genetic contributions are likely to be significant in families with multigenerational histories of enuresis. A Scandinavian group identified 11 families that manifested PNE over three generations in a pattern that was suggestive of an autosomal-dominant inheritance, with penetrance greater than 90 percent. Genetic linkage studies with these families indicated markers on chromosome 13q. Similarly designed studies have also implicated chromosomes 4p, 8, 12q, and 22, suggesting genetic heterogeneity. A large multigenerational study that specifically investigated chromosome 22 found that only 39.3 percent of families had linkage to genetic markers on this chromosome. The genetic modes of transmission that have been discussed include sporadic, autosomal recessive, and autosomal dominant. Linkage studies have also suggested the possibility of mutation in the aquaporin-2 (AQP2) water channel locus in families with a dominant mode of transmission of primary enuresis, in that AQP2 appears to be essential for concentrating urine. Although genetic linkage studies continue to provide important information with regard to the etiology of enuresis, it increasingly appears that the genetic contributions are heterogeneous in nature and vary substantially among pedigrees that contain multiple generations of individuals with enuresis. Thus, it is likely that there is more than 1 mode of genetic transmission that contributes to the disorder; and

different mechanisms may be operating in different family pedigrees.

One particularly important etiological hypothesis was that urinary tract obstruction frequently caused enuresis. The significance of that belief revolves around the related implications for surgery to correct the obstruction. Extensive review of the relevant literature found no basis for concluding that bladder neck repair or urethral dilation was a reasonable or effective treatment for enuresis. That generalization does not apply to patients with specific anatomical or pathophysiological findings.

Research into the association between sleep physiology and enuretic events has evolved steadily since the 1970s. The research can be seen as evolving through the following four stages. During the first phase of the research, it was postulated that enuretic events were dream equivalents that occurred in deep sleep. The second major theory grew out of the observation that enuretic events originated in delta sleep and followed arousal signals. Thus, enuresis was seen as a disorder of arousal, which led to the hypothesis that enuretic children who were not psychologically disturbed did not produce arousal signals, whereas behaviorally disordered children did but failed to respond to them. The third phase of sleep research is represented by the two largest controlled studies, both of which found that, when the time of night is also considered, enuretic events occur in all sleep stages in proportion to the amount of time spent sleeping in each stage. A fourth phase of investigation has attempted to couple cystometry with sleep studies in an attempt to identify subtypes. These investigations have yet to produce any definitive results that can be replicated.

A long-debated issue has been whether children with nocturnal enuresis are more difficult to awaken from sleep than other children. Earlier literature had suggested that anecdotal reports from the parents of children with enuresis that they were difficult to arouse stemmed from their lack of experience in attempting to arouse their other children, who did not have enuresis, from a deep sleep. However, studies have suggested that children with enuresis may, in fact, be more difficult to arouse than control children. Recent reports have also found that obstructive sleep apnea may play a role in the development of PNE in some children.

The efficacy of DDAVP as a treatment for enuresis has led to the theory that some children lack the ability to concentrate urine produced during the night and thus cannot reduce urine volume, and manifest enuretic episodes as a result. That hypothesis has led to studies investigating the circadian variation of plasma atrial natriuretic peptide (ANP). Children with enuresis have not consistently been found to differ from controls with regard to ANP. In one study, 14 of 55 children with enuresis had lower ANP concentrations than controls, and 9 of the 14 children had an excellent response to treatment with DDAVP. An investigation into the concentration of vasopressin in plasma and urine in 18 children with PNE

and 20 matched controls found the concentration of vasopressin in plasma at 8:00 AM was significantly lower in the enuretic group. Total 24-hour urinary vasopressin excretion was also lower in the enuretic group but not significantly so.

Another hypothesis that has been investigated is that endogenous arginine vasopressin (AVP) levels would be different in children with enuresis who respond to DDAVP as opposed to those who do not. This would suggest the existence of a specific subgroup of children with enuresis based on their endogenous AVP levels and their positive response to DDAVP. A study involving a small number of subjects supported this hypothesis by indicating that AVP levels in the morning differentiated between controls and children with enuresis, as well as between DDAVP responders and nonresponders. However, a large study that also controlled for water intake over 72 hours and that sampled AVP more frequently found that only 14 of the 55 children with enuresis had a significant decrease in AVP when compared to controls. Of these, 9 had a significant positive response to DDAVP. Further complicating this line of research is the finding that AVP is not secreted in a continuous manner over 24 hours, but rather is secreted in a pulsatile pattern, suggesting that definitive studies of this variable require frequent sampling. Investigations that have adopted this methodology have yielded conflicting results, in that one reported no differences between DDAVP responders and nonresponders, whereas another study with a small number of subjects that used hourly measurements found that in the 11:00 PM to 4:00 AM time period, children with enuresis ($N = 9$) had significantly lower AVP levels than did controls ($N = 8$).

There have also been related studies into the role of urine osmolality. One hypothesis is that a decreased ability to concentrate urine could contribute to an increase in the volume of urine. In general, these studies indicated that children with enuresis do have lower urinary specific gravities (as compared to controls) but not to a level that reaches statistical significance. It has also been reported that the levels of AVP can be increased in controls and children with enuresis by fluid restriction and, furthermore, that plasma osmolality can effect AVP secretion. In this regard, a comparison of controls, DDAVP nonresponders, and DDAVP responders indicated that fluid deprivation produced an increase in AVP levels in all groups, but the response was less robust in the DDAVP responders than in the other groups.

Another variation on this line of research is that refractory enuresis persisting into adolescence and adulthood may be secondary to a relative insensitivity to AVP, which can be positively affected by DDAVP. These observations contribute to the hypothesis that the primary pathophysiology of enuresis may occur at the receptor level, and this may also explain why some children do not respond to DDAVP.

Further complicating the research into the mechanism of action of DDAVP are two independent investigations that suggested that there is a central nervous system mechanism that may account for its efficacy in some children with enuresis.

Thus, as the foregoing review suggests, what initially appeared to be an intuitive explanation for the mechanism of action of DDAVP has not been supported by a number of studies, and a precise mechanism that applies to all children who respond to DDAVP has not been identified.

There have also been investigations that have implicated tubular mechanisms with regard to the urinary excretion of potassium and sodium by children with enuresis. Although this research has primarily focused on the possible mechanism of action of DDAVP, it has also led to a reanalysis of imipramine's mechanism of action. Imipramine can produce a decrease in the excretion of sodium and potassium, which then produces a reduction in osmolal clearance and a decrease in urinary output. It has also been reported that imipramine can affect the secretion of nocturnal urinary antidiuretic hormone (ADH) and that this may account for its efficacy.

Diagnosis and Clinical Features

The concreteness and simplicity of enuresis make the diagnosis relatively easy. The DSM-5 continues the same age criteria that were used in the DSM-IV and DSM-IV-TR by specifying that the diagnosis is not made in a child whose chronological or mental age is younger than 5 years of age. The stated rationale is that, by this age, continence can be expected to have occurred. As noted by the reference to a mental age of 5 years, children with developmental disabilities may have a greater chronological age. The wetting must occur at least twice per week for at least 3 consecutive months, or, if less frequent, it must produce significant distress or functional impairment. Physical causes, such as a bladder infection, must be excluded. Qualifiers to the diagnosis indicate whether it is primary or secondary enuresis. As previously noted, *primary enuresis* refers to those children who have never maintained continence for any length of time, whereas *secondary enuresis* refers to those who had achieved continence for at least 6 to 12 months and then resumed wetting. The other qualifiers refer to the timing of the enuretic event during the day. Although most children exhibit only nocturnal enuresis, some have daytime (diurnal) patterns or a combined nocturnal and diurnal pattern.

The inclusion of children who wet intentionally with those who wet involuntarily under the same diagnostic heading can be problematic because that distinction has significant clinical implications. Children who are wetting intentionally almost certainly are doing so as a manifestation of a psychological disturbance. Although there is an association between behavioral disturbance and involuntary enuresis, the nature of the correlation is relatively nonspecific. Thus, although the coexistence of other

behavioral problems is not a diagnostic issue per se, it is of clinical importance.

John was a 10-year-old boy who had consistently wet the bed throughout his life. He averaged three to four nocturnal episodes per week, and his longest period of continence was a few weeks. Thus, John warranted a diagnosis of primary enuresis. John's father had been enuretic until 12 years of age, as were his father and a paternal uncle. Accordingly, the family had always assumed that this was a genetic hereditary disorder that John could not control. In keeping with this perception, John's parents had never adopted a punitive approach to the bedwetting. The wetting was handled in a matter-of-fact manner, with John washing his sheets every morning and showering before going to school.

The family sought consultation at this time because John was increasingly being asked to sleep overnight at the homes of friends and also wanted to attend an overnight Boy Scout camp. These concerns led the family to meet with their pediatrician, who suggested a trial of the bell-and-pad method of conditioning. This approach was not effective, in part because the family lived in cramped quarters and the alarm that was meant to awaken John also woke his two brothers who shared the bedroom with him.

At this point, the family's pediatrician suggested that they meet with a child psychiatrist with whom she worked. The child psychiatrist described the clinical literature on DDAVP. However, when John's father indicated that his health insurance did not cover prescriptions and that the family was having difficulty covering their household expenses on his income, which was only moderately above minimum wage, the psychiatrist felt compelled to discuss the relative costs of DDAVP and imipramine, which was less expensive than DDAVP. However, the psychiatrist also indicated that John would require a baseline electrocardiogram (ECG) (which his insurance could cover).

Ultimately, the family decided that John would need medication only for brief, discrete periods (overnights at friends and 1 to 2 weeks of summer camp) because they had worked out an acceptable plan for responding to the wetting on a day-to-day basis. Furthermore, they reasoned that, within approximately 2 years, John would experience a spontaneous remission of his enuresis, given that this was the family pattern. Thus, they decided that, because DDAVP has a rapid onset of action, they would pay out of pocket for the relatively small amounts of DDAVP that would be needed for these discrete periods over the next 2 years. The pediatrician advised the family of the Federal Drug Administration (FDA) alert regarding the risk of hyponatremia and seizures. They decided to use the oral form of DDAVP because the FDA identified the nasal spray as being more likely to produce hyponatremia and indicated that this form of DDAVP was no longer indicated for childhood enuresis. They were also advised to (1) limit John's fluid intake on the nights when DDAVP was administered, (2) discontinue its use if John developed the flu or another illness that could affect his fluid balance, and (3) discontinue the medication and notify their physician if symptoms of nausea, vomiting, or headache developed.

Pathology and Laboratory Examination

Because urinary tract infections can produce enuresis, a urinalysis should be part of every evaluation. Using radiographic procedures with contrast media to detect an anatomical or physiological cause for the enuresis is more problematic because the procedures are invasive and painful and the diagnostic yield is low. A large study carried out in a pediatric primary care setting found a 3.7 percent incidence of obstructive lesions in children with enuresis. Others have reported similar findings. The literature cited earlier in the section on etiology concerning the use of noninvasive ultrasound techniques to establish bladder capacity and wall thickness have not been replicated to a degree that would warrant their use on a routine basis.

However, they may be of use for children who are refractory to standard treatment.

Differential Diagnosis

As suggested previously, the primary differential diagnosis is a urinary tract infection. This is especially true for girls, who are more prone than boys to develop urinary tract infections. A urinary tract infection should be the first consideration for a girl who has been continent for a considerable period of time and has recently begun wetting. Although the diagnostic yield is smaller with boys, a urinalysis should still be carried out.

Enuresis can result from anatomical malformations or obstructive lesions, but the percentages are relatively low. If the history and interviews with the child suggest that the enuresis is intentional, then it is almost certainly related to an underlying psychological disturbance. The relation between psychological disturbance and involuntary enuresis is less clear. The coexistence of another behavioral disturbance should be noted and attended to clinically. The most common comorbid psychiatric disorder is ADHD.

It has also been documented that enuresis has a negative effect on self-image and that enuretic children have more negative feelings about their disorder as compared to children who have other chronic illnesses.

There have also been case reports linking the onset of enuresis in children to the administration of both the selective serotonin reuptake inhibitor (SSRI) antidepressants and the atypical antipsychotic agent risperidone (Risperdal). Thus, the possibility of enuresis as a side effect to these medications should be considered if a chronological correlation exists.

Secondary enuresis has also been identified as an important symptom of the abrupt onset of Type 1 Diabetes in children and is related to the polydipsia and polyuria that accompany that disorder.

Course and Prognosis

The natural history of enuresis is significant because it is a self-limited disorder that can spontaneously remit. The fact that the diagnosis is not made until 5 years of age takes into account children who have delays in toilet training that are not outside the accepted range (between 2 and 5 years of age). The prevalence of enuresis is relatively high, between 5 and 7 years of age, and then drops off substantially. The vast majority of enuretic children experience a spontaneous resolution of the problem at some time, and only a few remain enuretic into adulthood. By 14 years of age, only 1.1 percent of boys wet once a week or more. The peak age for secondary enuresis is between 5 and 8 years of age.

Treatment

The methods for the treatment of enuresis that have been empirically proven to be effective are primarily behavioral and pharmacological. Psychotherapy may be useful for ameliorating some of the associated behavioral problems that can be seen with enuresis, especially secondary enuresis. A particularly common clinical scenario for secondary enuresis is the development of wetting after the loss of a parent through death or divorce. In these patients, psychotherapy is the primary treatment modality. A review of the efficacy of psychotherapy for primary enuresis found a 20 percent success rate, which is probably not significantly greater than would be expected by the rate of spontaneous remissions and random chance.

Behavior Therapy. A comprehensive review of several studies determined the success rate for behavioral interventions as 75 percent. Recent studies have yielded comparable response rates. The primary behavioral intervention is the bell-and-pad method of conditioning. A pad is placed on the bed, with a wire running to a bell. When the child wets, the moisture completes a circuit in the pad, ringing the bell and waking the child. With repeated use, the child learns to awaken before wetting occurs. Several studies have indicated that concomitant behavioral disturbance reduces the likelihood of success of behavioral interventions.

Recent behavioral studies have focused on embellishment of the standard method and comparisons with different methods. For example, no difference was found between an audio alarm and a vibratory alarm. A study that compared the basic bell-and-pad method with methods in which the bell and pad were coupled with more response contingencies for wetting versus not wetting found no statistical difference other than a slightly lower relapse rate with the adjunctive use of reward contingencies. One innovative strategy involved replacing the bell and pad with an alarm clock timed to go off after 2 to 3 hours of sleep, when maximum bladder capacity would be expected. The authors reported responses that were equal to results with the standard bell-and-pad method.

The relationship between bladder capacity and response to behavioral treatment has also been investigated. One study found that bladder capacity did not affect outcome with the bell-and-pad method of conditioning, whereas another found that children with smaller bladder capacities tended to do slightly better when retention-control training was linked to the bell-and-pad method. A meticulous investigation into which treatment variables might relate to the efficacy of the bell-and-pad method found that the inability to be awoken by the alarm correlated significantly with failure, as did low functional bladder capacity.

There is some evidence to support the use of biofeedback for children

who have small bladder capacities and unstable detrusors and who have been refractory to prior treatment. The newest advance in this behavioral methodology has been the use of an external ultrasonic monitor that is attached to a waistband. The monitor then signals the alarm when the bladder reaches maximal capacity. Clinical research with this device has yielded success rates similar to those with the basic bell and pad and has also reported increases in bladder capacity. The bell-and-pad approach has also been coupled with the use of DDAVP for children who have been refractory to prior treatment owing to family dysfunction or a concomitant behavioral disorder, or both. Although the bell and pad is the most intensively studied method of behavioral intervention, it is not the only type of behavioral treatment. A recent comprehensive review of this subject also noted other behavioral methods, such as evening fluid restriction, nighttime toileting, rewards, overlearning, and retention-control training.

Pharmacotherapy. The initial description of imipramine's efficacy for enuresis was followed by more than 40 double-blind studies that confirmed its therapeutic value. This marked the advent of the pharmacological era of the treatment of enuresis, with imipramine being the dominant pharmacological entity used for more than two decades. Although imipramine has now been largely replaced by DDAVP, there are still reasons to include a discussion of its use in a comprehensive review of enuresis. Imipramine is still a consideration for children who are refractory to other forms of treatment. There is also a significant cost differential between imipramine and DDAVP, so that imipramine may well be the first pharmacological choice for those families who do not have any form of insurance and who have limited financial resources.

Later studies with imipramine focused on the relation between response and concentration of imipramine in the blood. Although one study found no connection between positive response and imipramine concentration alone or in combination with its metabolite desipramine (Norpramin), three studies found a significant correlation between positive response and total concentration of imipramine and desipramine. More specifically, optimal response was found at combined concentrations (imipramine plus desipramine) of greater than 60 ng/mL in one study and of greater than 80 ng/mL in another. The most recent study also found that efficacy was related to increasing dose but noted tremendous variation in serum concentrations among individual children receiving the same dose. The adverse effect of dry mouth also correlated with concentration in blood, providing clinicians with a crude index of blood concentration in children who are extremely reluctant to have their blood drawn.

Clinically, when one encounters a history of a lack of response to imipramine, one must ascertain the specific dose used because nonresponse is frequently related to dose: Primary care physicians often

prescribe only 25 to 50 mg per day. However, low-dose responders do exist, and it is reasonable to start treatment at a daily dose of 25 mg, titrating up by 25 mg every 4 to 7 days. This enables identification of the low-dose responders because the response to imipramine usually occurs within 1 to 2 days. The majority of children exhibit a positive response in the range from 75 to 125 mg. The standard maximal limit for dose is 5 mg/kg body weight, and ECG monitoring is recommended at doses greater than 3.5 mg/kg. A pretreatment baseline ECG is also recommended. Blood concentration values may also be clinically useful if higher doses fail to elicit a response.

There is also a possible risk of overdose by the child with enuresis. Children can engage in magical thinking, and there have been reports of children who believed that, if taking a few pills would stop the wetting for a night, then consuming the whole bottle would make it disappear forever. Thus, parents should be advised to control the medication carefully and to store it in a secure place. Severe overdoses may require treatment with physostigmine (Antilirium), whereas more moderate overdoses may be treated by the symptomatic management of arrhythmias and seizures.

Because almost all children at some point experience spontaneous remission of the enuresis, it does not make sense to prescribe medication for years without checking for a spontaneous remission. One solution to that problem is to taper the imipramine dose at 3-month intervals. Should enuresis reappear as imipramine dose is tapered, it can simply be returned to the optimal level for another 3-month period. Clinically, one often finds that more children remain dry after 3 to 6 months of treatment than can be expected from spontaneous remission alone. In general, treatment with imipramine should be viewed as necessary until spontaneous remission occurs.

Three clinical subtypes of enuretic response to imipramine exist: True nonresponders, true responders who maintain a response over time, and transient responders who respond for a few weeks and then resume wetting. Increasing the dose for transient responders usually recaptures the response but only for another few weeks, and ultimately one reaches a point at which the dose cannot be increased further. However, the initial transient response can be replicated after a medication-free interval; thus, one can use the medication for especially important, time-limited social events, such as summer camp.

As indicated previously, DDAVP has now surpassed imipramine as the first-line pharmacological approach for enuresis. A thorough review done as early as 1993 revealed 18 randomized, controlled trials, including a total of 689 subjects. Many subjects had been refractory to previous treatment for enuresis. DDAVP reduced enuretic frequency by 10 to 91 percent in these studies. Positive prognostic factors were age (older than 9 years of age) and lower pretreatment frequency of enuretic events. In most studies, wetting resumed after the medication was discontinued. Follow-up studies

indicated that 5.7 percent of children remained dry after cessation of medication, but many of those cases could have been due to spontaneous remission. The most common reported adverse effects were mild abdominal pain, epistaxis, headache, and nasal stuffiness. Several subsequent studies attempted to find factors that might predict or relate to an increased likelihood of a positive response to DDAVP. No one factor appears to predict a positive response, but larger functional bladder capacity and a positive family history of enuresis appear to be positively correlated with treatment success. Factors that are not reliably related to treatment outcome are urine osmolality, plasma osmolality, ADH secretion, urine production, and nocturnal AVP concentration. However, individual studies can be found that link these factors with a positive outcome.

A large multicenter follow-up treatment study of DDAVP from Sweden (N = 399) used a 4-week baseline observation phase, a 6-week dose titration period (range from 20 to 40 μg), and a 1-year follow-up period. The study design also included an assessment for the rate of spontaneous remission by means of a 1-week medication-free period every 3 months. Sixty-one percent (N = 245) of the children met the requirement of at least a 50 percent reduction in frequency during the dose titration period that was required to enter the 1-year follow-up phase of the study. The average frequency of enuretic events in the baseline observation period was 5.3 events per week, and this declined to 0.8 nights by the end of the year-long follow-up period. As with other studies, older age was a positive prognostic sign. Of interest was the observation that 77 children experienced a complete remission of their enuresis within the first 6 months of the active treatment phase. There has been one report of an attempt to increase the rate of spontaneous remission by aggressively titrating the dose until cessation of wetting is obtained and then maintaining this dose for 4 to 6 weeks before beginning to decrease the dose by 10 μg per month until it is discontinued. The authors reported that, by using this strategy, complete remission was achieved for 71 percent of the children. The median length of follow-up was 18 months, and the mean dose of DDAVP was 20 μg.

Studies have also specifically investigated the efficacy of DDAVP for children who were refractory to conditioning treatment and imipramine. One of these involving 52 children (dose range from 20 to 40 μg intranasally) found complete cessation of wetting in 53 percent, a partial response in 19 percent, and minimal or no response in 28 percent. A long-term, follow-up study (mean length of follow-up of 13 months) of children treated with DDAVP found no hematological or hormonal adverse effects.

Treatment with DDAVP has been compared to the bell-and-pad method of conditioning. The difference in response rate was not statistically significant; 86 percent improved with the bell and pad, and 70 percent improved with DDAVP. Another investigation comparing the combined use of the bell-and-pad method and DDAVP with placebo indicated that the

combination achieved significantly more dry nights. Similar results were found in a study comparing alarm monotherapy to alarm treatment coupled with DDAVP. The combination proved superior, especially for children with severe, almost nightly, wetting.

The development of oral tablets greatly simplified the administration of DDAVP. The nasal spray was difficult for some children to use, and this could (in some cases) contribute to variable administration. A controlled multicenter study found no difference in response between oral DDAVP and the traditional nasal spray. A dose of 400 μg produced more dry nights than did a 200-μg dose.

A significant dose–effect correlation was found in a placebo-controlled study that used 200-, 400-, and 600-μg oral doses. In a separate follow-up study (7 years), the rate of spontaneous remission at 2 and 7 years was reported as greater than would be expected solely on the natural history of the disorder. As with other treatment modalities, consistent findings in the clinical literature concerning positive responses to DDAVP are older age, larger bladder capacity, and lower pretreatment frequency of enuresis. The newest innovation in the use of DDAVP has been the development of a sublingual desmopressin oral lyophilisate (MELT). A recent randomized, controlled study that compared this formulation to the oral tablets found it to be equally effective. The children also expressed a greater preference for the MELT. This increase in preference was especially pronounced in the younger age group (5 to 11 years of age).

As the use of the DDAVP nasal preparation became more common, there were an increasing number of case reports of hyponatremic seizures with prolonged use. A review of 14 articles that reported data on serum sodium concentration in DDAVP-treated patients and 11 articles that reported individuals who developed altered consciousness or a seizure while taking DDAVP for enuresis concluded that hyponatremia is a potential adverse effect of DDAVP treatment. The authors also noted that excess fluid intake was a contributing factor in 6 of the 11 case reports. They concluded that patients receiving DDAVP for nocturnal enuresis should be advised not to consume more than 8 oz of fluid on nights when DDAVP is administered. The data on hyponatremia also suggest that periodic monitoring of serum sodium concentration would be prudent. The number of reported cases of hyponatremia and seizures, including two deaths, prompted the FDA to issue an alert concerning DDAVP in December 2007. The alert identified 61 postmarketing incidents of hyponatremia-related seizures that were related to the use of DDAVP. Sodium levels obtained at the time of the seizure ranged from 104 to 130 mEq/L. A particular concern was that two of the individuals who experienced seizures with hyponatremia died. The intranasal formulation of DDAVP was implicated in 36 of the cases, and 25 of these were individuals younger than the age of 17 years, most of whom were receiving DDAVP for treatment of nocturnal

enuresis. The FDA alert concluded that the intranasal formulation of DDAVP was no longer indicated for the treatment of PNE. The alert also suggested that treatment with the tablet form should be halted during illnesses that could produce disturbances in fluid dynamics and/or produce electrolyte imbalance. The alert also recommended that fluids should be restricted from 1 hour before to 8 hours after the administration of the tablets. The latter recommendation is consistent with earlier published reports that linked the amount of fluid intake to the risk of acute hyponatremia.

Although imipramine and DDAVP have been the dominant pharmacological treatment modalities over the last several years, there continue to be reports of success with anticholinergic agents. The majority of these reports relate to oxybutynin, both as a first-line intervention and for individuals who are refractory to other interventions.

General Treatment Considerations. The results of the studies reviewed, in conjunction with the natural history of enuresis, can help the clinician to develop a decision tree tailored to a given child. The first decision, naturally, is whether to treat at all. That decision is affected by the frequency of the wetting, the age of the child, and the amount of psychological distress that the wetting causes to the child. Enuresis has a negative effect on the self-esteem of these children, and self-esteem improves after successful treatment. One may choose not to treat a child who is wetting infrequently and who is at an age at which a spontaneous remission may be expected.

If one does elect to treat, it makes sense first to try the bell-and-pad method of conditioning because its efficacy has been repeatedly demonstrated and behavioral intervention is considered less invasive than pharmacological approaches. A large follow-up study that compared observation only with imipramine, DDAVP, and the bell and pad indicated that, although all of the active treatments were significantly superior to observation alone, the positive response was much more likely to be sustained after treatment cessation with the bell and pad than with either pharmacological treatment. Despite these apparent advantages, one large population-based investigation revealed that only a relatively small number of enuretic children (38 percent) saw a physician, and, of that number, only 3 percent were treated with the bell-and-pad method of conditioning, whereas more than one third received pharmacological treatment. A more geographically restricted survey of primary care physicians yielded significantly different responses, in that 80 percent of physicians indicated that they recommended the bell-and-pad method of conditioning. The most recent comparative study of the long-term efficacy of DDAVP versus the enuretic alarm involved 142 children who were randomized to either form of treatment for 6 months, with determination of success rates at 6 and 12

months. Interestingly, only 5.2 percent of the DDAVP group withdrew after randomization, as compared to 30.7 percent of those assigned to the enuretic alarm group. The results for those who completed 6 months of treatment (defined as at least 90 percent in wet nights) indicated that 76.8 percent of the DDAVP group was deemed to be successful, as compared to 61.8 percent of the alarm group. The success rates at 12 months were comparable with 77.8 percent of the DDAVP group meeting success criteria, as compared to 75 percent of the alarm group. However, when the large drop-out rate after the assignment to the alarm group was taken into account, the success rates were adjusted to 68.8 percent for the DDAVP group, and 46.2 percent for the alarm group. Thus, this study essentially confirmed the results of numerous prior studies, in that the two treatment modalities were found to be equally effective, but the ability to implement the bell-and-pad alarm system was too difficult for many families to successfully implement.

When the bell-and-pad method of treatment is unsuccessful or not feasible, it makes sense to proceed to pharmacological treatment, assuming that the enuresis is severe enough to warrant pharmacological intervention. For decades, imipramine was the first choice for pharmacological treatment. Although precise figures are not available for its relative use in the treatment of PNE, DDAVP appears to have surpassed imipramine as the most popular pharmacological treatment.

Only recently, with the effect of managed care, have articles appeared that address the financial cost of the various treatments for enuresis. The bell-and-pad method of conditioning appears to be the most cost-effective treatment. It is as effective as the pharmacological approaches, is more likely to lead to sustained improvement after the cessation of active treatment, and requires only the initial expense of the apparatus and replacement pads. Recent suggested retail prices for the basic apparatus ranged from $6.00 to $196.00, depending on the brand chosen. Other expenses to be considered are the time of the professional who counsels the parents on its use and the parental time involved. Imipramine is available generically, so its actual cost is quite modest. Related costs are the physician's time and the cost of a baseline ECG, periodic monitoring of ECG, and periodic blood concentration determinations.

The nasal preparation of DDAVP has been available in a generic form for some time. However, as noted earlier, the recent alert by the FDA indicates that the nasal preparation of DDAVP is no longer indicated for PNE. The oral preparation is now available generically but is still more expensive than imipramine. This cost differential has narrowed with the availability of generic DDAVP. A recent comparison of the retail price for 30 of the 0.2 mg DDAVP tablets ranged from $28.00 to $113.00, with the average being in the $30.00 to $100.00 range. The cost for 60 of the 50 mg imipramine tablets ranged from $17.00 to $33.00, with the average falling

in the mid $20.00 range. Thus, for some families without insurance, that covered DDAVP cost would still be a differential. The newer sublingual formulation has been in Phase III clinical trials for several years and has not yet been approved by the FDA. However, it is available in Canada and the United Kingdom.

ENCOPRESIS

Encopresis does not receive as much research attention as does enuresis. A literature search at the beginning of this decade using the key words *encopresis and children* yielded 165 citations. A recent search using the same parameters yielded 118 citations. Thus, although this disorder does not generate as much research interest as enuresis, the rate of new publications has been relatively steady over the last two decades.

Definition

The DSM-5 criteria are virtually unchanged from those of the DSM-IV and DSM-IV-TR and define *encopresis* with four related criteria: (1) the repeated inappropriate passage of feces, usually involuntary, (2) occurrence at least once a month for at least 3 months, (3) a chronological or mental age of 4 years of age, and (4) exclusion of a substance or medical condition as a cause.

For several years, the clinical and research literature concerning encopresis has distinguished between retentive and nonretentive encopresis. The DSM-5 recognizes this distinction and lists two subtypes: With constipation and overflow incontinence, and without constipation and overflow incontinence. The subtype with constipation and overflow incontinence corresponds to retentive encopresis. The subtype without constipation and overflow incontinence corresponds to nonretentive encopresis.

Some clinicians refer to a *primary type* in which the person has never established fecal continence and a *secondary type* in which the disturbance develops after a period of established fecal continence.

History

The history of encopresis is not as richly detailed as is that of enuresis, possibly because encopresis affects fewer children, and the events are less frequent. Unless the feces are particularly fluid, it may also be easier for children to hide the symptoms. A high frequency of encopresis and enuresis was found in children who were separated from their parents in World War II, thus documenting the effect of environmental factors.

The more contemporary history of encopresis is noteworthy for the establishment of a treatment methodology that involves cathartics to re-establish bowel dynamics (coupled with education), which has proven to be

quite successful. An innovative use of biofeedback technology showed much initial promise after its introduction in the 1980s, but it now appears that it is not more effective than the older method. Constipation is a common precursor to encopresis, and recent research has also focused on effective treatment for this disorder.

Comparative Nosology

As noted above, the DSM-5 has adopted the widely held clinical distinction between retentive and nonretentive encopresis, although the terminology differs slightly. The current DSM-5 definition includes voluntary and intentional encopresis that is obviously a distinction of major clinical importance. This statement appears to minimize the importance of a rather substantial subgroup of children with nonretentive encopresis who are encopretic because they cannot adequately control the anal sphincter or do not recognize the need to defecate in time.

The most comprehensive categorization of encopresis distinguishes three types of patients: (1) intentional—children who do have bowel control but intentionally deposit feces in inappropriate places for psychological reasons; (2) involuntary—those who cannot adequately control the sphincter or lack an awareness of the process, or both; and (3) involuntary—those whose soiling is related to excessive fluid from retentive overflow (75 percent of this category), diarrhea, or anxiety. Thus, this classification schema incorporates the retentive and nonretentive dichotomy and the distinction between voluntary and involuntary incontinence.

The ICD-10 diagnostic classification for encopresis essentially follows this schema and allows notation of the three subtypes. The primary differences with the DSM-5 criteria are the ability to code for three subtypes rather than two and a duration of 6 months at a frequency of at least once a month, as opposed to the DSM-5 duration of 3 months at a frequency of at least once a month. The ICD-10 also states that if encopresis coexists with enuresis, then "the coding of the encopresis should have precedence." This statement is problematic in terms of using diagnostic databases for research, given that it is important to know how many children are polysymptomatic.

Epidemiology

Two large epidemiological studies have yielded consistent results. The Isle of Wight Study reported a prevalence of 1.3 percent for boys and .3 percent for girls (10 to 12 years of age) soiling at least once a month. A large study involving 8,863 children found a 1.5 percent prevalence rate for boys (7 and 8 years of age). In that study, the ratio of boys to girls was more than 3:1. A significant correlation between encopresis and enuresis has also been found.

Etiology

The dramatic nature of fecal soiling has contributed to psychodynamic speculation. The term *chronic neurotic encopresis* has been used to describe the small number of children whose encopresis is predominantly psychological. The clinical characteristics described in these children are (1) a distant father and neurotic mother, (2) early and often harsh bowel training, and (3) a history of neurological delay. The one systematic study that attempted to verify the clinical features usually attributed to enuretic and encopretic children in the literature by analyzing actual case material could not substantiate the existence of 20 of the 22 factors developed through the literature review. As with enuresis, the distinction between primary and secondary encopresis relates to the issue of associated psychopathology. A study involving 63 boys with encopresis found that boys with primary encopresis were more likely to have developmental delays and associated enuresis, whereas those with secondary encopresis were more likely to have experienced higher levels of psychosocial stressors and to be diagnosed with conduct disorder. Another large study examined several psychological and physiological variables in children with chronic constipation as compared to children with encopresis unrelated to constipation. The group with constipation had significantly longer colonic transit times, whereas both groups had relatively high frequencies of abnormal defecation dynamics. A subsequent independent investigation of colonic transit time (CTT) in children with chronic constipation found that a CTT greater than 100 hours was related to a poor outcome at 1-year follow-up.

A large population-based study in the United Kingdom involving 8,242 children (7 to 8 years of age) investigated the association of soiling with a variety of psychological disturbances. The children with encopresis were divided into those who soiled frequently (at least once a week) or occasionally (defined as less than once a week). There was also a large control group. Both of the symptomatic groups had significantly higher rates of behavioral and psychological disturbance than the controls. The rates of disturbance were also higher in the group that soiled frequently than in the occasional-soiling group. However, there was no particular behavioral profile that correlated with encopresis. The authors noted higher rates of attentional, obsessive-compulsive, and oppositional presentations in the symptomatic children than in the controls.

A premorbid history of encopresis has been found to occur more frequently in individuals with prepubertal and early adolescent bipolar disorder phenotype (PEA-BP) than in controls. A recent large, retrospective study (in a military health-system database) involved 32,773 children who were identified as having ADHD. The results indicated that 4.1 percent of those children had constipation, as opposed to 1.5 percent of control

children without ADHD. The corresponding rates for fecal incontinence were 0.9 percent in the children with ADHD, as compared to 0.15 percent of the controls. Both of these results were statistically significant with high levels of confidence, leading the authors to conclude that ADHD was a risk factor for fecal incontinence.

A recent large birth cohort study involving 358 children also found higher rates of both enuresis and encopresis among children with ADHD.

No correlation between encopresis and social class was reported in two independent studies. The issue of neurological competence as a contribution to encopresis has been investigated. One study found no abnormalities of anal-rectal motor or sensory function but did find that a significant number of encopretic boys had abnormal anal-rectal expulsion dynamics. A sophisticated physiological study using a triple-lumen catheter attached to a hydraulic manometry infusion system found that age of onset, frequency, and duration of encopresis were all highly correlated with the amount of anal sphincter spasm that occurred when defecation was attempted. The amount of pain reported with bowel movements correlated positively with frequency of encopresis and inversely with maximum squeeze pressure. An investigation involving the concentrations of gastrointestinal (GI) hormones in controls and children with retentive encopresis indicated that the latter had higher concentrations of pancreatic polypeptide after a meal and less of a motilin response. However, the authors noted that these abnormalities could be the result of, or the cause of, the chronic constipation.

An extensive series of studies concerning physiological factors and the interplay of physiological and psychological factors has been carried out. The principal physiological findings were (1) 56 percent of children with retentive encopresis could not defecate rectal balloons (research technique used to assess sphincter competency), and most had abnormal anal sphincter contractions; (2) only 14 percent of those who could not defecate the balloons were found to have responded to treatment at 1-year follow-up, compared to a 64 percent success rate for those who could; (3) at 1-year follow-up, 70 percent of children who could relax the anal sphincter at the time of initial evaluation were improved, as opposed to 13 percent of those who could not; and (4) none of the children who initially presented with an abdominal fecal mass was improved at 1-year follow-up, regardless of other factors.

One subset of these studies physiologically compared constipated children to controls during the act of bearing down. All of the control children displayed decreased anal sphincter activity when bearing down, compared to 58 percent of the constipated children who could defecate a balloon and 7 percent of constipated children who could not defecate the balloon. The constipated children who could not defecate the balloon were also significantly less likely to respond to laxative treatment. Nonrelaxation

of the anal sphincter was found in 75 percent of encopretic children, as opposed to 13 percent of a control group in a study, which used anal electromyography. This study also reported lower pressures for the encopretic children, at rest and on squeezing; this observation has also been described by other investigators.

An attempt was also made to assess the relative effect of physiological and psychological factors. Specifically, anal-rectal manometric and electromyographic assessments were correlated with behavioral and social competence profiles as related to treatment outcome. Psychological variables were not found to be predictive of outcome, but physiological variables were.

Diagnosis and Clinical Features

The rather objective and concrete nature of the encopretic event makes diagnosis relatively easy and uncomplicated from a phenomenological perspective if the child meets the frequency, duration, and age requirements set forth in the DSM-5. The involuntary versus intentional dichotomy is not a requirement or reason for subclassification according to the DSM-5, but it has important clinical implications. Similarly, the retentive versus nonretentive schema has clinical usefulness. Differentiating between primary and secondary encopresis is important, given that one might expect medical or psychological factors to play a greater role in secondary encopresis.

Henry was an 11-year-old boy with almost-daily encopresis and a number of associated behaviors, including hiding the feces around the house. He resided in a specialized foster care setting, having been removed from his biological parents at 7 years of age because of physical and sexual abuse. Both parents were involved with substance abuse, and his early history was not well documented. However, the parents did indicate that he had never exhibited sustained bowel continence for several months. Henry had also been enuretic until 6 years of age, but this had resolved to an occasional nocturnal episode every 4 to 6 months.

Henry also qualified for a diagnosis of ODD. Although he had experienced physical and sexual abuse, he did not have flashbacks or other symptoms that would meet the criteria for posttraumatic stress disorder (PTSD). Henry also had ADHD and was being effectively treated with 10 mg of methylphenidate (Ritalin) twice a day.

The foster family resided in an urban area that had access to a nationally recognized children's hospital. The ambulatory care department had a specialized behavioral encopresis program that coupled the bowel training method with a psychoeducational component and psychotherapy. The psychiatric consultant to the specialized foster care program doubted that this program would be successful for Henry because he had so much associated psychopathology and the feces were often deposited around the house in a symbolic manner. In addition, the encopresis was not of the retentive-overflow type, and the feces were always well formed. However, because no apparent harm could come from the referral, the consulting child psychiatrist agreed to it. Much to the surprise of the consultant, the several-week outpatient bowel training course coupled with the psychoeducational component and psychotherapy resulted in a complete cessation of the encopresis. On one of her visits to the home, Henry proudly showed his case manager a diagram of the functioning of the digestive system that was part of the psychoeducational program. In retrospect, it appeared that, although there were symbolic aspects to Henry's encopretic behavior, the soiling was ego-dystonic, and he was highly motivated to change the behavior, although this motivation could

not be prospectively detected by the treatment team because of his oppositional defiant manner of responding to adults.

Pathology and Laboratory Examination

The extensive, detailed physiological studies cited previously suggest that physiological factors are important in many children with encopresis, given that they correlate with treatment outcome at 6-month and 1-year follow-up. However, that does not imply that a physiological workup that is this thorough is required for every child who presents clinically. The clinician needs to rule out definite physiological illnesses, such as Hirschsprung disease. Beyond that, clinicians might consider more detailed physiological investigations, such as those used in the experiments described previously, for children who prove refractory to conventional treatment. That type of specialized testing could also be difficult to obtain outside of an academic research center. A simple plain abdominal roentgenogram may aid in the diagnosis and confirmation of retentive encopresis. In one study, 78 percent of children who met the diagnostic criteria for encopresis also met roentgenographic criteria, whereas 22 percent did not.

Children with a condition referred to as *slow transit constipation* may present with intractable encopresis. Although the etiology of this disorder is not fully understood, some of these children appear to have abnormal innervation of the colon. Clinical signs of this disorder include severe constipation within the first year of life and refractory encopresis at an early age. Colonic transit studies are necessary to confirm the diagnosis.

Psychological testing can be useful for those children who exhibit concomitant behavioral problems, or for those who are intentionally encopretic. Finally, clinicians should not overlook the importance of a detailed history obtained from the parents and the child that documents the time course, frequency, and circumstances surrounding the encopresis.

Differential Diagnosis

Encopresis can be a symptom of other processes or can be a syndrome in itself, as described previously. The medical illnesses that need to be ruled out include stenosis of the rectum or anus, endocrine abnormalities, smooth muscle disease, and Hirschsprung's disease.

The child's psychological profile should be considered because children with intellectual deficits or pervasive developmental delay may have episodes of soiling related to those processes. Children with impulse-control disorders or attention-deficit disorder may at times have episodes of soiling related to lack of attention. Encopresis has been reported as one of a number of symptoms that can be seen following sexual abuse. A large retrospective study investigated the hypothesis that the association between encopresis and sexual abuse was of sufficient significance that

encopresis could be classified as an indicator of sexual abuse. The three study groups (4 to 12 years of age) included 466 children with documented sexual abuse, 429 children with externalizing psychiatric disorders, and 641 controls. A history of sexual abuse had been ruled out in the latter two groups. The rates of fecal soiling were similar in both the abuse (10.3 percent) and psychiatric groups (10.5 percent). The soiling rate in the control group was much smaller (2 percent). Thus, the authors concluded that the symptom of fecal soiling could not, in and of itself, be used as a predictive indicator of sexual abuse. A thorough history and psychiatric assessment should make it possible to identify those children, as well as those children who soil on an intentional-oppositional basis. Extreme stress may also provoke encopretic episodes in children who are otherwise functioning well.

Course and Prognosis

As with enuresis, continued maturation provides increasing numbers of spontaneous remissions over time. However, the rate of spontaneous remission is not as clearly defined as that for enuresis. That trend should be factored into the assessment of the efficacy of any long-term intervention. The results from relatively simple behavioral interventions involving educational, behavioral, and physiological components are striking (78 percent), suggesting that the disorder responds to treatment in most cases.

Treatment

The first clinical approach to encopresis is primarily behavioral, with educational and physiological components. That method entails initial educational interventions with the child and family around bowel function. The process is meant to correct misconceptions that the family may have and to dissipate anxiety that may develop around the soiling. The physiological component is instituted next with an initial bowel catharsis, followed by daily doses of laxatives or mineral oil. The behavioral element of the program entails daily timed intervals on the toilet, with success being rewarded. Follow-up studies have reported a 78 percent success rate for this method. The physiological research findings concerning abnormal defecation dynamics in children with encopresis led to interest in biofeedback training as an adjunctive treatment. Although this treatment received positive reports from uncontrolled studies, a review of controlled studies indicated that biofeedback training did not provide additional benefit over conventional treatment. A large, long-term, controlled outcome study that compared conventional treatment with biofeedback training plus conventional treatment indicated that the positive results for both groups were high and quite similar. Specifically, the conventional method yielded an 86 percent improvement rate, as compared to 87

percent for the combined group. This suggests that the combination treatment was no better than the conventional treatment alone. However, it was also found that length of time at follow-up was a powerful factor because it significantly correlated with improvement for the group as a whole at the $p < 0.01$ level. This indicates that the passage of time is an important variable to consider in any investigation of treatment approaches.

A recent systematic meta-analysis of the literature with regard to behavioral and/or cognitive treatments for fecal incontinence concluded that treatment with biofeedback did not improve the results obtained from the traditional treatment.

One element of the aforementioned physiological studies, namely, the ability to defecate a rectal balloon, has been shown to correlate with treatment outcome. A study involving 139 children with constipation and encopresis and 20 controls found that all of the controls were able to defecate the balloon, as compared to 47 percent of the encopretic group.

One year after beginning treatment, 51 percent of the children who could defecate the balloon at baseline were considered to be recovered, but only 34 percent of those who could not defecate the balloon were improved. Although the group comparisons were considered significant, the balloon test alone could not predict response to treatment on an individual basis.

Psychotherapy may be of use for children who have concomitant behavioral problems and those with psychodynamic factors that appear to be especially strong and are contributing to the disorder. This would also be true for those who began soiling after a stressful or traumatic event. Naturally, psychotherapeutic approaches are of primary importance for children who are intentionally soiling or depositing feces in inappropriate places.

Pharmacological treatment for encopresis has not been extensively studied or reported. There are six case reports (encompassing 15 individual patients) that have described imipramine as effective for encopresis. Twelve of the 15 patients were boys. The therapeutic effect had relatively rapid onset (as it does with enuresis), occurring within a few days to 2 weeks, usually at doses of 25 to 75 mg. Although the subtype of encopresis is not clearly delineated in all of the case reports, one would anticipate that, if imipramine were to be helpful, it would be in those with nonretentive encopresis. There is one well-controlled study that substantiated the efficacy of the prokinetic-propulsive agent cisapride (Propulsid). This agent is no longer commercially available in the United States, having been removed from the market owing to concerns about side effects.

The relatively high success rate of the combined educational, physiological, and psychological approach described makes it a reasonable first approach once other diagnoses have been ruled out. Psychotherapeutic and pharmacological approaches could be considered for children who

prove refractory to that form of treatment.

REFERENCES

Baeyens D, Roeyers H, Demeyers I, Verté S, Hoebeke P, Vande Walle J. Attention-deficit/hyperactivity disorder (ADHD) as a risk factor for persistent nocturnal enuresis in children: A two-year follow-up study. *Acta Paediatr*. 2005;94(11):1619–1625.

Benninga MA, Voskuijl WP, Akkerhius GW, Taminiau JA, Buller HA. Colonic transit times and behaviour profiles in children with defecation disorders. *Arch Dis Child*. 2004;89:13–16.

*Borowitz SM, Cox DJ, Sutphen JL, Kovatchev B. Treatment of childhood encopresis: A randomized trial comparing three treatment protocols. *J Pediatr Gastroenterol Nutr*. 2002;34:378–384.

Brazzelli M, Griffiths PV, Cody JD, Tappin D. Behavioral and cognitive interventions with or without other treatments for the management of faecal incontinence in children. *Cochrane Database Syst Rev*. 2011;(2):CD002240.

Butler RJ, Holland P, Gasson S, Norfolk S, Houghton L, Penney M. Exploring potential mechanisms in alarm treatment for primary nocturnal enuresis. *Scand J Urol Nephrol*. 2007;41(5):407–413.

El-Mitwalli A, Bediwy AS, Zaher AA, Belal T, Saleh AB. Sleep apnea in children with refractory monosymptomatic nocturnal enuresis. *Nat Sci Sleep*. 2014;6:37–42.

Freeman KA, Riley A, Duke DC, Fu R. Systematic review and meta-analysis of behavioral interventions for fecal incontinence with constipation. *J Pediatr Psychol*. 2014;39(8):887–902.

Freitag CM, Rühling D, Seifen S, Pukrop R, von Gontard A. Neurophysiology of nocturnal enuresis: Evoked potentials and prepulse inhibition of the startle reflex. *Dev Med Child Neurol*. 2006;48(4):278–284.

Fritz G, Rockney R, Bernet W, Arnold V, Beitchman J, Work Group on Quality Issues, AACAP. Practice parameter for the assessment and treatment of children and adolescents with enuresis. *J Am Acad Child Adolesc Psychiatry*. 2004;43(12):1540–1550.

Glazener CM, Evans JH, Peto RD. Alarm interventions for nocturnal enuresis in children. *Cochrane Database Syst Rev*. 2005;(2):CD002911.

Hergüner S, Kilingaslan A, Gürker I, Tüzün U. Serotonin-selective reuptake inhibitor-induced enuresis in three pediatric cases. *J Child Adolesc Psychopharmacol*. 2007;17(3):367–369.

Hergüner S, Mukaddes NM. Risperidone-induced enuresis in two children with autistic disorder. *J Child Adolesc Psychopharmacol*. 2007;17(4):527–530.

Joinson C, Heron J, Butler U, von Gontard A. Psychological differences between children with and without soiling problems. Avon longitudinal study of parents and children study team. *Pediatrics*. 2006;117(5):1575–1584.

Joinson C, Heron J, Butler R, von Gontard A, Butler U. A United Kingdom population-based study of intellectual capacities in children with and without soiling, daytime wetting, and bedwetting. *Pediatrics*. 2007;120(2):e308–e316.

*Joinson C, Heron J, Emond A, Butler R. Psychological problems in children with bedwetting and combined (day and night) wetting: A UK population-based study. *J Pediatr Psychol*. 2007;32(5):605–616.

Klages T, Geller B, Tillman R, Bolhofner K, Zimerman B. Controlled study of encopresis and enuresis in children with a prepubertal and early adolescent bipolar-I disorder phenotype. *J Am Acad Child Adolesc Psychiatry*. 2005;44(10):1050–1057.

Landgraf JM, Abidari J, Cilento BG Jr, Cooper CS, Schulman SL, Ortenberg J. Coping, commitment, and attitude: quantifying the everyday burden of enuresis on children and their families. *Pediatrics*. 2004;113:334–344.

Lottmann H, Froeling F, Alloussi S, El-Radhl AS, Rittig S. A randomized comparison of oral desmopressin in lyophilisate (MELT) and tablet formulations in children and adolescents with primary nocturnal enuresis. *Int J Clin Pract*. 2007;61(9):1454–1460.

McKeown C, Hisle-Gorman E, Eide M, Gorman GH, Nylund CM. Association of constipation and fecal incontinence with attention-deficit/hyperactivity disorder. *Pediatrics*. 2013;132(5):e1210–e1215.

Mellon MW, Natchev BE, Katusic SK, et al. Incidence of enuresis and encopresis among children with attention-deficit/hyperactivity disorder in a population-based birth cohort. *Acad Pediatr*. 2013;13(4):322–327.

Mellon MW, Whiteside SP, Friedrich WN. The relevance of fecal soiling as an indicator of child sexual abuse: a preliminary analysis. *J Dev Behav Pediatr*. 2006;27(1):25–32.

Mikkelsen EJ. Enuresis and encopresis: ten years of progress. *J Am Acad Adolesc Psychiatry*. 2001;40:1146–1158.

Önol FF, Guzel R, Tahra A, Kaya C, Boylu U. Comparison of long-term efficacy of desmopressin lyophilisate and enuretic alarm for monosymptomatic enuresis and assessment of predictive factors for success: a randomized prospective trial. *J Urol*. 2015;193(2):655–661.

Roben JH, Sze M, Knoers NV, Eggert P, Deen P, Müller D. Relief of nocturnal enuresis by desmopressin is kidney and vasopressin type 2 receptor independent. *J Am Soc Nephrol*. 2007;18(5):1534–1539.

Robson WL, Leung AK, Norgaard JP. The comparative safety of oral versus intranasal desmopressin for the treatment of children with nocturnal enuresis. *J Urol*. 2007;78(1):24–30.

Schulz-Juergensen S, Rieger M, Schaefer J, Neusuess A, Eggert P. Effect of 1-desamino-8-D-arginine vasopressin on prepulse inhibition of startle supports a central etiology of primary monosymptomatic enuresis. *J Pediatr*. 2007;151(6):571–574.

Touchette E, Petit D, Faquet J, Tremblay RE, Boivin M, Montplaisir JY. Bed-wetting and its association with developmental milestones in early childhood. *Arch Pediatr Adolesc Med*. 2005;159(12):1129–1134.

Van Hoecke E, De Fruyt F, De Clercq B, Hoebeke P, Vande Walle J. Internalizing and externalizing problem behavior in children with nocturnal and diurnal enuresis: A five-factor model perspective. *J Pediatr Psychol*. 2006;31(5):460–468.

*von Gontard A, Freitag CM, Seifen S, Pukrop R, Rühling D. Neuromotor development in nocturnal enuresis. *Dev Med Child Neurol*. 2006;48(9):744–750.

von Gontard A, Schaumburg H, Hollmann E, Eiberg H, Rittig S. The genetics of enuresis: a review. *J Urol*. 2001;166:2438–2443.

50

Other Disorders of Infancy, Childhood, and Adolescence

▲ 50.1 Reactive Attachment Disorder and Disinhibited Social Engagement Disorder

MARGARET J. HAUCK, PH.D., MARY MARGARET GLEASON, M.D., AND CHARLES H. ZEANAH, M.D.

Reactive attachment disorder (RAD) of infancy and early childhood first appeared in the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III). The DSM-IV-TR delineated two subtypes—an emotionally withdrawn/inhibited pattern and an indiscriminately social/disinhibited pattern. Research, conducted primarily in the past 10 years, has demonstrated that both subtypes emerged only after an infant or very young child had experienced a significantly inadequate caregiving environment and yielded evidence that the subtypes are distinctive, with different presentations, neurobiology, and course of development. Based on a growing body of literature, DSM-5 describes two disorders—reactive attachment disorder (RAD), which is evident when children show the emotionally withdrawn/inhibited pattern of behavior, and the disinhibited social engagement disorder (DSED), which describes children who demonstrate the indiscriminately social/disinhibited pattern. Together, these disorders are often designated "disorders of attachment," though controversy remains about whether DSED should be considered an attachment disorder.

The normal development of the infant and young child's attachment system, first described by John Bowlby, is the trajectory to which behavior indicating disordered attachment is compared (Table 50.1–1). In the context of a grossly inadequate caregiving environment, the child's attachment does not develop normally. Descriptions of the phenotypes associated with attachment disorders emerged in case reports and descriptive studies in the mid-20th century, but systematic research appeared only in the past two decades. Recent studies of disturbances of attachment in institutionalized and maltreated infants and young children

have provided broad support for the phenotypic patterns of the two disorders, begun to characterize the disorders in later childhood and adolescence, and started to uncover neurobiological correlates of each disorder.

DEFINITIONS AND CONTEXT

RAD, as defined in DSM-5, describes young children who show no consistent evidence of selective attachment behaviors towards an adult caregiver. Instead of approaching adults for comfort, support, nurturance, and support, they fail to approach, show distress without seeking or responding to comfort, and display emotion regulation difficulties, especially absence or minimal positive emotional displays and sometimes unexplained outbursts of irritability, fear, or sadness. They rarely engage in socially reciprocal interactions with adults. DSED describes young children who exhibit lack of expected reticence about approaching and engaging with unfamiliar people. They also display physically or verbally intrusive behavior and violation of social boundaries. Both diagnoses (RAD and DSED) also require that the children's clinical problems are preceded by and presumably explained by grossly insufficient care that limits opportunities to form stable attachments, such as that which occurs in social neglect/deprivation, repeated changes of primary caregivers, or rearing in institutions.

Attachment

Attachment, as described by Bowlby in his influential theory, refers to a biobehavioral system which coordinates the balance between the need for seeking safety in proximity to a caregiver or set of caregivers and the tendency to explore and seek novelty and mastery in infancy and early childhood. Bowlby argued that human infants were motivated by an attachment behavior control system to seek proximity to a small number of identified caregivers. He inferred that the reason for seeking proximity in order to attain an internal feeling of security. This internal sense of security is reinforced by repeated experiences of the caregiver's emotional and physical availability in times of need.

Beginning around 7 to 9 months of age, infants are thought to develop internal representations of their relationships with important caregivers. These representations, or internal working models, form the basis for the intense emotional bond between infants and their primary caregivers and behavior in later relationships. The normal developmental progression of attachment, captured in attachment-related behaviors, is summarized in Table 50.1–1. The commonly accepted method for classifying patterns of attachment is the Strange Situation Procedure, which involves an examination of the balance between the child's tendency to explore and

tendency to seek comfort while interacting with an attachment figure and an unfamiliar adult (the "stranger"). Individual differences in the balance between attachment and exploratory behavior and how those are expressed are used to classify the infant's attachment to this caregiver into one of four types of attachment: secure, in which children express distress directly, seek comfort actively, and obtain soothing readily from attachment figures; avoidant, in which children minimize signs of distress and show no comfort seeking efforts despite physiological arousal; resistant, in which children show significant distress to separation but fail to obtain comfort from their attachment figure and cannot be soothed; or disorganized, in which children show some mixture of the secure, avoidant or resistant, or anomalous behaviors that preclude classification as another pattern. These patterns of attachment represent risk or protective factors rather than diagnoses.

Table 50.1–1.
Developmental Features of Attachment in Early Childhood

	Birth–2 Months	2–7 Months	7–12 Months	12–18 Months	Older than 18 Months
Phase of attachment	Limited discrimination	Discrimination with limited preference	Preferred attachment	Secure base	Goal-oriented partnership
Characteristics	Physical attributes of babyishness attract caregivers, but infant expressions of preference are limited to olfactory and auditory realms and require experimental conditions to be demonstrated	Infant differentiates among different interactive partners; infant may seem more comfortable with primary caregiver but is willing to interact readily with other social partners	Clearly expressed preference for a small number of caregiving adults; separation protest and stranger wariness are normative	Use of the attachment figures as a secure base from which to venture out confidently and to explore the world and a safe haven to return to in times of danger; proximity to caregiver promotes feeling of security in infant	Cooperation with caregiver despite frequent conflicts; goal is to balance autonomous functioning with reliance on caregiver for help when needed; verbal relatedness is important, including attention to emotional integration and fluency of discourse

Early Environment

RAD and DSED derive from adversity in the early caregiving environment, involving extremes of social and emotional deprivation (although the child's basic physical survival needs may well have been met). The importance of early experience on infant development was recognized as far back as the 13th century when the Holy Roman Emperor Frederick II reportedly ordered that a group of infants be chosen who would be given simple physical care but with whom no one would be allowed to speak. The experiment was intended to see what language the children would eventually use. According to the Franciscan monk whose chronicle recounted the events, the experiment ended with the premature death of the children, who "could not live without clappings of the hands, and gestures, and gladness of countenance, and blandishments."

The relatively common practice of institutionalizing infants in the United States in the first half of the 13th century unwittingly provided more evidence of the detrimental effects of what has been called "maternal deprivation." Pediatricians had long argued against the practice of institutionalization, particularly because of the high mortality rates related to failure to thrive, but other health care providers did not take notice of the effects of institutionalization on social and emotional development until René Spitz in the 1940s and James and Joyce Robertson in the 1950s conducted more rigorous studies (including control groups). These studies also led Spitz to propose a diagnostic entity that he called *anaclitic depression*, which foreshadowed the DSM-III criteria for RAD. Anaclitic depression, which included mood and social symptoms as well as immunological and growth problems, seemed to follow prolonged separation of infants from their caregivers and was most severe if the infants were old enough to have developed a preference for that caregiver.

A study by Barbara Tizard identified the two sets of behaviors that inform the current DSM-5 diagnoses as well as the *International Statistical Classification of Diseases and Related Health Problems, 10th revision* (ICD-10) classifications of attachment disorders. The children in the study were in institutions, but with a relatively low caregiver-to-child ratio, mixed age groups, and toys and books to enrich the environment. However, the caregivers were explicitly instructed not to form attachments to the children. Although 8 of the 26 children followed in the study managed to form secure attachments despite the constraints of the environment, 8 were found to be emotionally withdrawn and socially unresponsive (similar to children with RAD), and 10 were indiscriminately social with unfamiliar adults (as seen in children with DSED).

The eventual recognition of the clinically detrimental consequences of institutional rearing led to a marked decrease in state-run orphanages in the United States and the United Kingdom. Nevertheless, institutional care remains widely used as an intervention for orphaned, abandoned, and maltreated young children in Asia, Central and South America, and Europe,

particularly Eastern Europe.

Recognition of the scope and consequences of the problem of child maltreatment in the United States did not begin until the latter part of the 20th century, spurred initially by C. Henry Kempe's classic 1962 article on the *battered child syndrome*. Since the early 1960s, there has been a great deal of research on the developmental effects of maltreatment on children. This body of work includes substantial consequences of serious neglect, including serious disturbances of attachment.

COMPARATIVE NOSOLOGY

The criteria for RAD have evolved markedly since the diagnosis was first introduced in the DSM-III. This early version of the disorder description included growth failure and lack of social responsivity as central features. The diagnosis had to start by 8 months of age (the age at which preferred attachment to a restricted set of caregivers is usually just beginning to be evident) and could not result from a diagnosable medical condition. Gross neglect of the infant's physical and emotional needs had to be evident, and, as with later criteria, the presence of autism or mental retardation precluded the diagnosis. Pertinent behaviors cited in the criteria included poor tone, weak cry, excessive sleep, lack of interest in the environment, and weak rooting and grasping when feeding.

Although failure to thrive is a potentially important problem in infant mental health, it is not equivalent to an attachment disorder. The restrictive age range criterion and the requirement for growth failure made the DSM-III definition difficult to apply in most cases. Furthermore, it is rarely possible to rule out autistic disorder or mental retardation in severely malnourished infants younger than 1 year of age. The link between failure to thrive and RAD was dropped in the revised third edition of the DSM (DSM-III-R), and the age of onset was changed to the first 5 years of life. The two subtypes of the disorder—inhibited and disinhibited—were introduced with the DSM-III-R and persisted in the DSM-IV-TR. The criteria continued to include an emphasis on the etiological link between the disorder and evidence of "pathogenic care" at the hands of the young child's primary caregivers and the exclusion of children whose symptoms might be accounted for by cognitive delay or one of the pervasive developmental disorders.

With the publication of the DSM-5, the two subtypes of RAD have been separated into two different diagnoses: RAD and DSED. This separation brings the DSM nosology more in parallel with the ICD-10. The ICD-10 includes the Reactive Attachment Disorder of Childhood (RAD), which is similar to RAD in DSM-5 and Disinhibited Attachment Disorder of Childhood (DAD), which is similar to DSED in DSM-5. Although differences among these approaches can be identified, common features include aberrant social behavior that is cross-contextual, insufficient care

(neglect) as a necessary etiological factor, and the two clinical pictures of RAD (emotionally withdrawn/inhibited behavior) and DSED (indiscriminately social/disinhibited).

Recent research examining the criteria has affirmed that RAD and DSED may be reliably identified in high-risk children. Studies have also demonstrated that these diagnoses are rare in clinical samples, unless children have histories of severe neglect. Children with RAD also seem to be quite responsive to improvements in the caregiving environment. For example, in a randomized, controlled trial of foster care as an alternative to institutional care in Bucharest, Romania, children with diagnoses of RAD responded within months to placement in families with decreases in signs of RAD to levels comparable to never-institutionalized children in the community. However, children with DSED continue to show behaviors consistent with the diagnosis after family placement, typically showing a reduction but not an elimination of their signs of DSED. In addition, studies of children who have been internationally adopted from institutions show that these children do not manifest RAD but do demonstrate persistent signs of DSED. Signs of RAD, which reflects a lack of attachment, diminish when children are placed in families and have the opportunity to form an attachment to foster or adoptive parents. However, signs of DSED persist in a minority of children even after they form attachments to foster or adoptive parents, suggesting that the core deficit in this disorder is not the lack of an attachment relationship.

EPIDEMIOLOGY

Very little data exist on the prevalence of either RAD or DSED. RAD is rarely seen in adequate caregiving environments and always following a history of severe deprivation. DSED appears to be more prevalent, though it is seen in a minority of children even if they live in inadequate caregiving environments. Preliminary estimates are that for children who are known to have been neglected or raised in institutional settings in the first few years of life, prevalence is roughly 10 to 20 percent.

ETIOLOGY

Research has established that the risk of RAD and DSED increases in the context of grossly neglectful care. Less clear is what degree of neglect is necessary for the disorders to occur. Still, even in environments characterized by severe deprivation, most children do not develop signs of RAD or DSED. Thus, more research on vulnerability factors and the critical elements of neglectful care responsible for these disorders is needed.

The longer children live in conditions of deprivation, the more likely they are to develop an attachment disorder. A critical age may exist, as research has found attachment disorders to be very rare in children adopted before 6 months of age, and signs of DSED have not been reported

in children whose deprivation begins only after 2 years of age. Children with lower IQs and stunted growth may also have attachment disorders because deprivation compromises both growth and IQ; it is difficult to know whether these problems precede or are merely additional consequences of deprivation. Studies are also beginning to identify possible genetic polymorphisms as risk factors for attachment disorders.

Preliminary research suggests that children who have experienced serious deprivation may show differences in brain structure and functioning compared to nondeprived children, and many of these children may also show signs of attachment disorders. For example, researchers have found reductions in gray and white matter volumes and reduced electrical activity in higher frequencies and increased electrical activities in lower frequencies in children who have experienced institutional deprivation. Importantly, reduced white matter volume mediated the effect of institutional rearing on EEG alpha power, meaning that reductions in myelinated circuitry likely account for the reduced EEG power. If children were placed with families before 24 months of age, the differences in electrical activity between deprived and nondeprived children were no longer seen at 8 years of age. These findings have been demonstrated in children with histories of institutional deprivation and are not specific to those with attachment disorders, however.

Difficulties with inhibitory control also have been shown in children with indiscriminant social behaviors, implicating ventral frontostriatal circuitry, though findings are not consistent across studies. Children adopted from institutions also showed reduced amygdala discrimination between mothers and strangers on fMRI, and these reductions were moderately associated with the children showing indiscriminant behaviors. Thus, preliminary research is demonstrating that early deprivation has an impact on brain development, although this relationship has been shown primarily in children with the behaviors consistent with a diagnosis of DSED and does need replication.

Although research on what causes caregivers to be neglectful is limited, caregiver characteristics, from severe personality disturbance to psychopathology, appear to be important. Parental mental retardation and poor basic parenting skills, particularly in the context of little social support, may also increase risk of neglect, though neglect sufficiently severe to contribute to RAD and DSED is rare in parents even with these challenges. Infants who are cared for by multiple caregivers in succession, as sometimes happens in the foster care system in the United States, are also at increased risk of attachment disorders. The traumatic experience of prolonged separation from a caregiver in early childhood, as documented by Spitz, the Robertsons, and Bowlby, is enough to cause attachment disturbance. Repeated, prolonged separations typically cause more severe symptoms, and, when these separations are accompanied by neglect,

attachment disorders appear to be even more likely to occur.

DIAGNOSIS AND CLINICAL FEATURES

RAD as described in DSM-5 is characterized not only by an inhibition of positive emotional and social engagement but also by the lack of selective or discriminating attachment behaviors directed at an adult caregiver. At times when young children should be motivated to seek closeness or comfort for distress or protection from danger, children with RAD do not seek or even readily respond to adults' efforts to comfort, reassure or protect them. Instead, they may appear withdrawn and unresponsive or they may show distress with no anticipation of comforting. Emotion regulation abnormalities include reduced or absent expressions of positive emotions and unexplained displays of irritability, sadness, or anxiety. Children have histories of severe neglect, and the signs typically remit within weeks of being placed in adequate caregiving environments. Occasionally, children may be seen with RAD that has persisted following placement into foster care. This is often because de facto neglect continues in the foster home because the child's withdrawn and negative behavior is off putting for adult caregivers who fail to override the children's miscues and avoid interacting with them. The diagnosis also requires the child to have experienced, a pattern of extremes of insufficient care, operationalized as social neglect/deprivation, repeated changes of primary caregivers that limit opportunities to form stable attachments (e.g., frequent changes in foster care), and/or rearing in group settings that severely limit opportunities to form selective attachments (e.g., institutions with high child-to-caregiver ratios).

The disturbed attachment behaviors must be present before the chronological age of 5 years and the child must have a developmental age of at least 9 months. This final criterion ensures that children have reached the point where they are cognitively able to form attachments; children in whom RAD is being considered must have a cognitive age of at least 9 months. Finally, the child must not meet criteria for autism spectrum disorder (ASD).

DSED, as described in DSM-5, is characterized by lack of stranger wariness which ordinarily becomes apparent between 7 and 9 months of age. The disinhibition in DSED is social, leading to behaviors in which a child actively approaches and interacts with unfamiliar adults. Though sometimes referred to as indiscriminate friendliness, the behavior is usually experienced as intrusive and may include overly familiar verbal or physical behavior that goes beyond culturally sanctioned and with age-appropriate social boundaries. Children with DSED often fail to track their caregivers whereabouts, even in unfamiliar settings and may even be willing to go off with an unfamiliar adult without hesitation. As in RAD, the child with DSED must have experienced a level of insufficient care such as

occurs in social neglect or repeated changes in placement. DSED criteria also stipulate that the disordered behaviors are not limited to the kinds of behavioral impulsivity that occurs in attention-deficit/hyperactivity disorder (ADHD), but rather is socially disinhibited behavior. Again, as in RAD, the care is presumed to be responsible for the behavior and the child must have a developmental age of at least 9 months.

Diagnosis of RAD or DSED requires a multisession evaluation, particularly when the caregiver's behavior is not known to be abusive or when the child's history does not include multiple placements. Ascertaining whether the child has less deviant behavior in the presence of the clinician or another trusted adult is important and may take time. This comparison may be especially helpful in distinguishing highly social children, for example, from those with DSED or somewhat behaviorally inhibited children from those with RAD. The child's relationship with the primary caregiver can be adequately assessed with a combination of careful history, and observations of interaction during free play, structured tasks that stress the dyad, and a brief separation and reunion. Videotaping the assessment is a useful way to document social responses, and the introduction of a clinically trained person whom neither partner has met is often informative. Children who meet criteria for this disorder may look quite different, depending on their developmental stage.

An 11-month-old girl recently placed in foster care was referred by state child protective services along with her biological and foster families to assist with long-term case management. Her parents both had chronic and serious substance abuse histories that included violence in their relationship and grossly inadequate care of the 11-month-old and her older siblings. Foster care placement occurred when police were called to the home during an episode of violence, found squalid conditions, including narcotic pain pills strewn on the floor.

The foster mother reported that when the child first arrived, she looked frightened much of the time, cried sporadically and was not soothed by her and was unresponsive to social overtures. It was weeks before she smiled. Her appetite increased over the first few days and she began to stuff herself to the point that the foster mother, with directions from her pediatrician, to limit what she offered the child at meals and snacks. For the initial couple of weeks after placement, the child exhibited little interest in exploring and mostly rocked quietly by herself. This pattern was suggestive of RAD.

The clinicians encouraged the foster mother to interact with the child despite the girl's lack of interest and to respond to distress as if the girl had requested it in order to teach her that she could be comforted when distressed. Gradually, the girl seemed to relax more, to smile—at first fleetingly but then more frequently—and to accept comfort for distress and then gradually to seek comfort for distress. By 1 month following placement, the foster mother reported that the girl seemed to be showing attachment behaviors towards her. She also reported feeling closer to the girl herself, who had first been hard for foster mom to connect to. The stuffing behavior slowly dissipated after this point. The girl became somewhat clingy to her foster mother and was quick to show distress to even minor changes in routine, but she began to seem much affectively brighter and to have much better emotion regulation.

A 45-month-old boy was referred by his adoptive parents because of hyperactivity and disruptive behavior at school. He had been placed with the family at 30 months and adopted at 36 months of age, after living most of his life in a Romanian orphanage, in which he received care from a rotating shift of caregivers. Although he had been below the 5th percentile for

height and weight on arrival, he quickly approached the 10th percentile in his new home. However, both of his adoptive parents were frustrated by their inability to "reach him." They had initially worried about a hearing disturbance, although testing and his capacity to engage many adults and children verbally suggested otherwise. He approached and engaged socially with complete strangers and more than once had followed strangers willingly. Teachers liked him, although they found him to be attention seeking and sometimes disruptive. At times, they thought he was appealing and even charming, but at other times they felt that his feelings were shallow. He was frequently injured because of seemingly reckless behavior, although he had an extremely high tolerance for pain. He was diagnosed with disinhibited social engagement disorder. Intensive intervention focused on decreasing his self-endangering behavior, enhancing his ability to read emotional expressions, and describe social situations and rewarding him for socially appropriate reticence. Target behaviors improved over several months, though tendencies to engage with unfamiliar adults persisted at a somewhat reduced and more manageable level.

DIFFERENTIAL DIAGNOSIS

Since RAD and DSED differ significantly in their clinical presentations, each have different diagnoses that need to be examined when differential diagnoses are made. For RAD: ASD, and depressive disorders; for DSED: ADHD. Also, developmental delay and intellectual disability should be distinguished from both diagnoses.

Depressive Disorders

Children with RAD and with depression both show reduced or absent positive affect and social withdrawal. In a longitudinal study of institutionalized or previously institutionalized young children, there was a moderate-to-high association between RAD and signs of depression in young children, but the majority of children who were diagnosed with a depressive disorder did not meet criteria for RAD. In general, children with depression ought to turn to caregivers for comfort and show behaviors that are indicative of an attachment. The neurovegetative signs of depression and the excessive guilt associated with depression are not characteristic of RAD. Only if these children had a history of social neglect and did not seek comfort would the diagnosis of RAD be warranted.

Autism Spectrum Disorder

Children with both ASD and RAD show impaired social behaviors, especially social withdrawal and minimal reciprocity. However, children with ASD typically also show atypical language development, while language skills in RAD may be delayed but are not stereotyped or otherwise deviant. Additionally, children with ASD show restricted and repetitive behaviors which are not found in children with RAD. Finally, children with ASD do show attachment behaviors with a limited number of caregivers some of the time and very rarely have the history of social neglect or maltreatment that is seen in children with RAD. The differential diagnosis of RAD and ASD may be challenging because quasi-autistic behaviors have

been described in formerly institutionalized, adopted children. However, the course of quasi-autism seems to be different from ASD in that at least partial remission may occur after several years of familial placement.

Attention-Deficit/Hyperactivity Disorder

Young children who are hyperactive and impulsive may also appear socially disinhibited. Furthermore, deprivation is known to increase the risk of both inattention or hyperactivity and DSED. The social disinhibition of DSED, however, is not necessarily accompanied by inattention or hyperactivity. However, children with DSED can also meet criteria for ADHD, requiring careful assessment of the potential differential diagnosis as well as comorbid conditions.

Intellectual Disability (Intellectual Developmental Disorder)

The adverse caregiving environment necessary for the manifestation of RAD or DSED also may include environmental deprivation that can lead to developmental delays. However, in longitudinal studies of children from institutions, RAD and DSED were only moderately associated with cognitive delays. In terms of differential diagnosis, children with intellectual disability typically show social and emotional skill levels that are relatively consistent with their cognitive levels. These children are also capable of forming stable attachments, as long as their developmental ages are greater than 9 months, the age at which attachments to caregivers should start to form.

COURSE AND PROGNOSIS

There is limited research addressing the natural course of RAD. Among children in institutions, the diagnosis appears relatively stable; young children who met criteria for RAD at one point in time continued to show the behaviors at other times. However, children who are placed into families typically do not show signs of RAD on follow-up (i.e., 1 year later). This would suggest a relatively good prognosis for amelioration of signs of RAD once they are in a stable caregiving environment, though whether they are at increased risk of insecure or atypical attachments, or other psychiatric disorders seems less clear.

More longitudinal research following children who meet the criteria for DSED is available. These studies suggest that the diagnosis is moderately stable from year to year in young children. Studies that have followed children into later childhood find that those with over-friendly and attention seeking behaviors at the age of 4 tend to continue to show these behaviors at the age of 8. Moreover, at the age of 16, these children demonstrated superficial and conflicted relationships with peers. Thus, the behaviors consistent with a diagnosis of DSED appear somewhat resistant

to change once established.

TREATMENT

The first consideration in the assessment of children exposed to grossly inadequate care is the child's safety. Child maltreatment is associated with significant morbidity, and mortality is not uncommon, particularly in children younger than 48 months of age. Early involvement of child protective services is often warranted, and assessment of parental fitness may be a necessary part of the evaluation. In some cases, placement of the child may be necessary, and reunification of parent and child may not be warranted. Children in this situation often have not had appropriate medical care, and medical referral is almost always indicated. Unfortunately, the foster care system and the family court system sometimes lead to multiple placements in the first years of life and thus increase the likelihood of an attachment disorder. The clinician may play a crucial role in staying involved with the child who is placed in foster care, providing expert testimony in court and individual or family treatment.

Once the child is in a safe placement, full attention can be paid to psychosocial intervention, which must often be tailored to the individual. For children who exhibit RAD, providing a primary attachment figure is the goal. Caregivers who foster secure attachments in infants and young children are those who are sensitively responsive, value the child as an individual, and place the needs of the child ahead of their own. These also are the characteristics needed for assisting the child with RAD. For children who are establishing attachments in adoptive or foster families and who exhibit the high levels of indiscriminate behavior seen in DSED, the clinical consensus is to advise parents to limit their unnecessary contact with adults outside the family for several months if possible. The rationale is that children may need time to consolidate attachment relationships within their family before they experience other relationships. Although this approach has face validity, it has not been evaluated formally.

Additional interventions should be those that reduce child or adult (caregiver) barriers to attachment. Child barriers include various miscues about their needs, including aggressive and oppositional behavior, emotional dysregulation, and withdrawn behavior. Adult barriers include distractions from the newly developing relationship, such as here-and-now crises, problematic relationship histories—including unresolved losses and trauma—and lack of psychological commitment to the child. Some attachment-specific treatments are available, such as Circle of Security, which weaves together psychoeducation with psychotherapy, or Attachment and Biobehavioral Catch-Up that is also manualized and more prescriptive about which behaviors are coached. Both treatments rely heavily on review of videotaped parent–child interaction. Other interventions that involve some form of caregiver–child dyadic therapy are

attachment informed rather than attachment derived, such as child–parent psychotherapy. These approaches typically weave together developmental training and guidance with an active attempt to address pressing instrumental issues (e.g., poor housing and inadequate medical care). The complexity of these approaches reflects the complexity of the clinical problem. Long-term interventions are necessary in these cases, and psychiatric treatment should be bolstered by early intervention programs and ongoing medical care for the child.

In general, research following children with RAD suggests that the symptoms of the disorder disappear with placement in a family that can provide a positive caregiving environment. This research additionally does find that these children can continue to show insecure or disorganized attachments, indicating that treatment to improve attachments with the foster or adoptive caregivers is important. However, research finds that some children with a diagnosis of DSED continue to show indiscriminant behavior, even after being placed in enhanced caregiving environments. Children with DSED may even form secure attachments, indicating that more than enhanced caregiving may be necessary to significantly reduce the signs of DSED. It has been suggested that intervention to target social cognitive abnormalities that may underlie the indiscriminate social behaviors shown in DSED would be a possible way to improve these behaviors, but research is needed in this area.

References

Ainsworth MD, Blehar MS, Waters E, Wall S. *Patterns of Attachment: A Psychological Study of the Strange Situation*. Hillsdale, NJ: Lawrence Erlbaum; 1985.

Bakermans-Kranenburg MJ, Steele H, Zeanah CH, et al. Attachment and emotional development in institutional care: characteristics and catch-up. In: McCall RB, van IJzendoorn MH, Juffer F, Groza VK, Groark CJ, eds. *Children without Permanent Parental Care: Research, Practice, and Policy*. Monographs of the Society for Research in Child Development, Serial No. 301, 2011;76:62–91.

Boris NW, Hinshaw-Fuselier SS, Smyke AT, Scheeringa MS, Heller SS. Comparing criteria for attachment disorders: establishing reliability and validity in high-risk samples. *J Am Acad Child Adolesc Psychiatry*. 2004;43:568–577.

Bowlby J. *Maternal Care and Child Health*. Geneva: World Health Organization; 1951.

Bowlby J. *Attachment and Loss*. 2nd ed. Vol. 1–3. New York, NY: Basic Books; 1982.

Chisholm K. A three-year follow-up of attachment and indiscriminate friendliness in children adopted from Romanian orphanages. *Child Dev*. 1998;69:1092–1106.

Drury SS, Gleason MM, Theall KP, et al. Genetic sensitivity to the caregiving context: the influence of 5httlpr and BDNF val66met on indiscriminate social behavior. *Physiol Behavior*. 2012;106(5):728–735.

*Gleason MM, Fox NA, Drury S, et al. The validity of evidence-derived criteria for reactive attachment disorder: indiscriminately social/disinhibited and emotionally

withdrawn/inhibited types. *J Am Acad Child Adolesc Psychiatry*. 2011;50:216–231.

Gleason MM, Fox NA, Drury SS, Smyke AT, Nelson CA 3rd, Zeanah CH. Indiscriminate behaviors in young children with a history of institutional care. *Pediatrics*. 2014;133:e657–e665.

Hodges J, Tizard B. Social and family relationships of ex-institutional adolescents. *J Child Psychol Psychiatry*. 1989;30:77–97.

Hoffman KT, Marvin RS, Cooper G, Powell B. Changing toddlers' and preschoolers' attachment classifications: the circle of security intervention. *J Consult Clin Psychol*. 2006;74:1017–1026.

Lawler JM, Hostinar CE, Milner S, Gunnar MR. Disinhibited social engagement in post-institutionalized children: differentiating normal from atypical behavior. *Dev Psychopathol*. 2014;26:451–464.

Lyons-Ruth K, Bureau J-F, Riley CD, Atlas-Corbett AF. Socially indiscriminate attachment behavior in the strange situation: convergent and discriminant validity in relation to caregiving risk, later behavior problems, and attachment insecurity. *Dev and Psychopath*. 2009;21:355–372.

O'Connor TG, Marvin RS, Rutter M, Olrick JT, Britner PA. English and Romanian Adoptees Study Team. Child–parent attachment following early institutional deprivation. *Dev Psychopathol*. 2003 Winter;15:19–38.

Olsavsky AK, Telzer EH, Shapiro M, et al. Indiscriminate amygdala response to mothers and strangers after early maternal deprivation. *Biol Psychiatry*. 2013;74:853–860.

Pears KC, Bruce J, Fisher PA, Kim HK. Indiscriminate friendliness in maltreated foster children. *Child Maltreat*. 2010;15:64–75.

Rutter M, Colvert E, Kreppner J, et al. Early adolescent outcomes for institutionally-deprived and non-deprived adoptees. I: Disinhibited attachment. *J Child Psychol Psychiatry*. 2007;48:17–30.

*Rutter M, Kreppner J, Sonuga-Barke E. Emanuel Miller Lecture: attachment insecurity, disinhibited attachment, and attachment disorders: where do research findings leave the concepts? *J Child Psychol Psychiatry*. 2009;50:529–543.

Smyke AT, Dumitrescu A, Zeanah CH. Disturbances of attachment in young children: i. The continuum of care taking casualty. *J Am Acad Child Adolesc Psychiatry*. 2002;41:972–982.

*Smyke AT, Zeanah CH, Gleason MM, et al. A randomized controlled trial comparing foster care and institutional care for children with signs of reactive attachment disorder. *Am J Psychiatry*. 2012;169:508–514.

Soares I, Belsky J, Mesquita AR, Osorio A, Sampaio A. Why do only some institutionalized children become indiscriminately friendly? Insights from the study of Williams Syndrome. *Child Dev Perspectives*. 2013;7:187–192.

Spitz RA. Anaclitic depression; an inquiry into the genesis of psychiatric conditions in early childhood. *Psychoanal Stud Child*. 1946;2:313–342.

Tizard B, Hodges J. The effect of early institutional rearing on the development of eight-year-old children. *J Child Psychol Psychiatry*. 1978;19:99–118.

Tizard B, Rees J. The effect of early institutional rearing on the behaviour problems and affectional relationships of four-year-old children. *J Child Psychol Psychiatry*. 1975;16:61–73.

Zeanah CH, Berlin LJ, Boris NW. Practitioner review: clinical applications of attachment theory

and research for infants and young children. *J Child Psychol Psychiatry*. 2011;52:819–833.

Zeanah CH, Chesher T, Boris NW; the AACAP Work Group on Quality Issues. Practice parameters for the assessment and treatment of children and adolescents with reactive attachment disorder of infancy and early childhood and disinhibited social engagement disorder. *J Am Acad Child Adolesc Psychiatry*. 2016;55(11):990–1003.

*Zeanah CH, Gleason MM. Annual research review: attachment disorders in early childhood–clinical presentation, causes, correlates, and treatment. *J Child Psychol Psychiatry*. 2015;56:207–222.

Zeanah CH, Scheeringa M, Boris NW, Heller SS, Smyke AT, Trapani J. Reactive attachment disorder in maltreated toddlers. *Child Abuse Negl*. 2004;28:877–888.

Zeanah CH, Smyke AT, Dumitrescu A. Disturbances of attachment in young children: II. Indiscriminate behavior and institutional care. *J Am Acad Child Adolesc Psychiatry*. 2002;41:983–989.

Zeanah CH, Smyke AT, Koga SF, Carlson E; Bucharest Early Intervention Project Core Group Core Group. Attachment in institutionalized and community children in Romania. *Child Dev*. 2005;76:1015–1028.

▲ 50.2 Stereotypic Movement Disorders in Children

ROBERT LLOYD DOYLE, M.D., D.D.S.

INTRODUCTION

A stereotypy is defined as a frequent, almost mechanical repetition of the same posture, movement, or form of speech (as in schizophrenia). This rather broad definition encompasses a range of motor phenomena from statuesque catalepsy as occurs in schizophrenia to dizzying degrees of spinning seen in autism. A stereotypy can be as benign as eye blinking or can be quite dangerous, as in the case of intense head banging resulting in traumatic brain injury. Although the term "stereotypy" may be applied to many types of behaviors, including normal repetitive behaviors of infancy, abnormal repetitive behaviors in autistic individuals, dyskinesias caused by psychotropic medications, and symbolic gestures associated with schizophrenia, to name a few, this section focuses on stereotypic movement disorders of childhood as described in the text revision of the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) and the tenth edition of the *International Classification of Diseases and Related Health Problems* (ICD-10).

HISTORY

The first mention of purposeless, repetitive behaviors in the clinical literature dates back 300 years to Nehemiah Grew, who wrote, "We see also Mad people . . . run upon some one action . . . without variation." Possibly the first description of a dopamine-driven stereotypy should be attributed

to Samuel Gee, who conducted an experiment in which he injected a large dose of apomorphine into a dog. A few minutes after the animal vomited, it began to pace around the room in a "curiously persistent methodical manner." In the early 1900s, Emil Kraepelin linked stereotypic behaviors to dementia praecox. Of note, Kraepelin's careful observations of stereotypic behaviors were made before the development of antipsychotic medications, and so the behaviors were likely related to a pathological brain state rather than to medication-induced dyskinesia. Several seminal papers published in the 1940s set the stage for the following 50 years of research on the subject.

Landmark research on stereotypic behaviors began with Lauretta Bender and Paul Schilder in the early 1940s. Their contribution differentiated impulsions from obsessions and compulsions based on the premise that impulsions were associated with positive affect, whereas obsessions were notably negative. They also speculated that impulsions could have connections to the primitive motor stereotypic behaviors of infancy, and that impulsions might become obsessions and compulsions in adolescence. In 1949, Reginald S. Lourie theorized that the feedback from infantile rhythmic behaviors fostered normal motor development. For instance, infants will assume the four-point crawling position on hands and knees, then rock front to back as if rehearsing the more complex coordinated movement in which only the arm and leg on one side of the body moves forward to provide locomotion. This simple repetitive rocking in place could serve as the basis for a repertoire of more-coordinated movements needed to crawl. Thus, according to this theory, repetitive behaviors occur in a normally developing child, but these recede as the child masters more-complex, functional behaviors. These behaviors become stereotyped if they persist beyond the developmental purpose they served, as seen in persons who are blind or have autism or mental retardation.

ETIOLOGY

Several theories have been proposed to explain the emergence and maintenance of stereotypic behaviors. As just described, stereotypic movements may be the precursor to mastery of more-sophisticated motor tasks. This path toward more complex motor behaviors may veer in a pathological direction based on the interplay among abnormal development, behavioral reinforcement, and psychological conflict. In Down syndrome, for instance, stereotypic behaviors are similar to those seen in children without a diagnosis of mental retardation but differ in that in Down syndrome they appear at a later-than-expected chronological age and do not disappear. Therefore, developmental factors could contribute to the manifestation of stereotypic behaviors.

Behavioral factors might maintain the stereotypic behavior through

positive or negative reinforcement. Careful behavioral analysis can identify social reinforcers and then aim to eliminate them. From a psychodynamic standpoint, stereotypies are considered instinctual, autoerotic, and self-soothing or gratifying behaviors. The psychodynamic factor implies that a stereotypic behavior, such as nail biting, might be a reaction to intense feelings about one or more conflicts. Biologically based theories have replaced most psychodynamic formulations explaining stereotypic behaviors.

Early observations in animals and humans indicated that environmental factors might modify the expression of stereotypic behaviors. For example, R. J. Gallagher and Gershon Berkson showed that even short periods of normal social interaction reduced the expression of stereotypic behaviors in institutionalized infants with profound retardation. Most studies show that substituting alternative activities for stereotypic behaviors effectively reduces their expression.

Even though enriched environments usually diminish stereotypic behaviors, they do not eliminate them. The predictable persistence of stereotyped behaviors in childhood autism and other conditions, despite environmental enhancement, indicates that organic factors might play a role. One organic theory proposes that endogenous opioid compounds are released in persons with autism as they engage in stereotypic behaviors. The pleasurable effect of the endorphin release could reinforce the repetitive maladaptive behaviors. Although studies using the opioid antagonists naloxone (Narcan) and naltrexone (Trexan) resulted in modest improvements in stereotypic behaviors, these have not been a definitive cure for stereotypic behaviors in individuals with autism. Therefore, an organic-based explanation is not entirely satisfactory either. Perhaps an interplay between experiential effects on the brain and physiological processes within the brain might best account for stereotypic behaviors. Bender and Schilder theorized that a common connection must tie all sorts of stereotypy together; however, studies fail to support the existence of an encompassing stereotypic factor.

Nevertheless, research has pointed to several factors that might act independently or in concert to bring about the expression of a stereotypic movement. These include genetic factors, as seen in nail biting and Lesch–Nyhan syndrome, and neurobiological factors, as evidenced by the many compounds that can precipitate stereotypic or self-injurious behaviors (SIBs). For example, dopamine D_1 supersensitivity due to chronically low levels of dopamine has been described in animal models of SIBs. Of interest, postmortem tissue from patients with Lesch–Nyhan syndrome appears to be dopamine deficient as well. Furthermore, dopamine antagonists have successfully treated nonsyndromal SIBs and at least one case of Lesch–Nyhan syndrome. Dopamine has also been implicated in the emergence or exacerbation of certain repetitive behaviors related to

stimulant use. Psychostimulants work by blocking the reuptake of dopamine, thereby causing a relative increase in the availability of dopamine in the synaptic cleft. Stimulant-induced repetitive behaviors include bruxism, skin picking, lip biting, compulsions, nail biting, perseveration, tics, and various stereotypies. Other neurotransmitters, such as serotonin, opioids, and γ-aminobutyric acid, may be involved as well.

EPIDEMIOLOGY

Many stereotypic behaviors emerge in early childhood and are considered part of normal development. These rhythmic, repetitive behaviors appear and progress in a cephalocaudal direction and can be correlated with chronological age. An estimated 80 percent of normal children display rhythmic activities that disappear by age 4 years. Although determining the prevalence rate in the general population is difficult, studies estimate that some 3 to 15 percent of children without disabilities engage in head banging. About 20 percent of children exhibit repetitive body rocking. If these behaviors continue beyond the normal developmental age range for such behaviors, they might evolve into a stereotypic movement disorder. Consider thumb sucking. Endearing in infants, thumb sucking easily becomes a reason for teasing in the school-age child. In an extreme case, an adolescent boy with profoundly severe mentally retardation sucked his thumb until a large portion of soft tissue eroded away and exposed the underlying bone. Finger and thumb sucking occurs more often in girls than boys, and more in whites than in African Americans.

Bruxism—the habitual grinding of teeth—occurs in approximately 56 percent of normally developing infants. Even if chronic bruxism wears down enamel, this rarely results in long-term problems in young children because deciduous teeth will be replaced by the permanent dentition. Bruxism in adolescents and adults can cause major dental pathology ranging from periodontal disease to temporomandibular joint dysfunction.

Sean was a burly but affable dental student with severe bruxism. He had no personal or family history of psychiatric illness. He was the eldest child of a large, upper–socioeconomic status family. Sean had always maintained a respectable grade point average, and he continued to do so in dental school. By age 23 years, Sean had worn the cusps of his posterior teeth completely away due to his chronic bruxism, which began in childhood. Although bruxism is often caused by malocclusion between maxillary and mandibular dentitions, this could not explain his persistent problem. Sean's posterior teeth had been worn down to flat occlusal surfaces without any cusps to cause interferences in his bite. His case was so impressive that his professors requested permission to take photographs and make models of his teeth to teach future generations of dental students. Ironically, his bruxism worsened as the stress of dental school caused him to chronically clench his teeth together. His dental health was compromised by broad areas of exposed dentin in places once covered by enamel. Sean was at risk of losing his molar teeth, as well as for developing painful temporomandibular joint dysfunction.

Treatment took a multimodal approach. Sean was advised to monitor himself for clenching and grinding during the day. Whenever he became aware of these behaviors, he was instructed to simply separate his upper and lower teeth. To prevent neuromuscular programming, his dentist recommended that he stop chewing gum. These interventions helped during the day;

however, nocturnal bruxing continued. Sean was not aware of grinding his teeth in his sleep, but he did notice tenderness around his jaw in the morning. On exam his dentist noted crepitus —the sound of bone rubbing against bone—with mandibular movement. Because crepitus is a sentinel sign of temporomandibular joint dysfunction, the dentist fitted Sean with a custom night guard that served two purposes: It prevented further eroding of enamel and it curtailed the chronic temporomandibular joint inflammation. Sean also took ibuprofen to relieve inflammation and pain. His dentist offered a prescription for clonazepam (Klonopin) to help him deal with the pressures of his professional education. Sean, however, did not want to become addicted to a benzodiazepine, so he chose to try other measures first. Buspirone (BuSpar) was also offered as an option because of its reduced abuse potential. Fortunately, his dental condition stabilized fairly quickly, and he set his sights on having his lost tooth structure restored with ceramic crowns.

Nail biting occurs quite commonly in childhood and frequently continues into adulthood. Occasionally, nail biting progresses to the point of causing dermatological or even dental problems. It does not seem to be associated with the child's age, level of intelligence, or socioeconomic status, but nail biting seems to have a familial basis. Toenail biting, however, does appear to be much more prevalent in girls. An estimated 50 percent of children exhibit nail biting. It can be seen as early as 4 years of age, with peak prevalence between age 10 and 18 years. According to one study, 4.5 percent of adults continue to bite their nails. Adults with onychophagia may not have comorbid psychiatric illness, and children with nail biting do not appear to be more anxious than peers who do not bite their nails. In most adult cases, nail biting seems to cause distress rather than result from stress.

As mentioned earlier, genetics seems to play a factor in the expression of some stereotypic behaviors. In twin studies, children born to parents who were nail biters were nearly three times as likely to be nail biters as control subjects. The concordance rate for monozygotic twins was 66 percent versus 34 percent in dizygotic twins. Onychophagia is the most severe type of nail biting. Monozygotic twins with onychophagia had a 75 percent concordance rate, whereas dizygotic twins only had an 18 percent concordance.

Of interest, stereotyped behaviors often occur in blind children. Elisa Fazzi and others studied a group of congenitally blind children. In their group of 11 boys and 15 girls with and without neurodevelopmental disabilities, 73 percent showed stereotypic behaviors. The most frequently encountered stereotypic behaviors were body rocking (30.7 percent), repetitive handling of objects (30.7 percent), and eye pressing along with eye poking (30.7 percent). These researchers also found that the stereotypic behaviors could be decreased by encouraging appropriate adaptive behaviors and increased by restricting the environment, reducing sensory stimulation, and limiting mobility.

Stereotypic behaviors are more common among children with a history of parental neglect, but they may be present in children who do not have emotional problems. An estimated 20 percent of individuals with severe

mental retardation exhibit stereotypic behaviors. Alfred A. Baumeister and Rex L. Forehand found that two-thirds of a population institutionalized with mental retardation exhibited stereotyped behaviors, and J. R. Dura and colleagues later reported that 34 percent of nonambulatory individuals with profound retardation who were institutionalized exhibited at least one stereotypic behavior. Despite the need for longitudinal studies, evidence from cross-sectional studies suggests that developmental age and experiential history, rather than chronological age, are the relevant mediating variables for stereotypic movements. Stereotypic movements frequently peak in adolescence and then diminish thereafter.

The prevalence of SIBs varies from between 2 and 3 percent for children and adolescents in the community to about 10 to 15 percent for adults with severe or profound retardation living in residential facilities. In individuals with severely retardation living in institutions, stereotyped movements increase with chronological age; however, the opposite is true of children without disabilities. In individuals with mental retardation, the most common stereotypic movements are body rocking, head rolling, and repetitive hand movements, and self-hitting, head banging, and self-biting are among the most common forms of SIBs. The prevalence of head banging ranges between 3 and 15 percent in normally developing children. Approximately three times as many boys as girls exhibit the behavior. Head banging usually emerges between 5 and 11 months of age, but only 5 percent of these cases persist beyond 3 years of age.

Per the DSM-5, stereotypic movements may occur with or without SIBs. A study by Gal and colleagues monitored SIB and stereotypy in typically developing children as well as those who were intellectually impaired, visually impaired, hearing-impaired, or autistic. In this sample of 90 subjects with at least one SIB, only three did not exhibit a co-occurring stereotypy. Interestingly, the rank order of prevalence of SIB and stereotypy was same across all five groups.

Given that the strongest predictor for the presence and severity of SIB was other repetitive behavior, an earlier model proposed by Guess and Carr in which SIB was thought to evolve from stereotyped movements makes some sense. SIBs, such as head banging, and stereotypy occur in typically developing children, but the SIB emerges later in development than stereotypy. In individuals with intellectual disabilities, SIB and repetitive behaviors appear temporally related; however, stereotypic behaviors still precede SIB.

Of note, Cornelia de Lange syndrome (mutations including the NIPBL gene in region 5p13, SMC1L1 in the Xp11.2 region, SMC3 gene in region 10q25), Fragile X syndrome (mutation at Xq27.3), Smith–Magenis (3.7 Mb interstitial deletion in chromosome 17p11.2), Cri-du-Chat syndrome (deletion in chromosome 5p15), and Rett syndrome (mutation in the methyl CpG–binding protein 2) are genetic disorders that present with

various forms of stereotypic movements and SIB. Since these genetic defects are quite diverse, yet each gives rise to stereotypies with and without self-injury in the same individual, finding treatments for stereotypic behaviors remains a challenge.

Nevertheless, theories for stereotypic movements abound, and research continues to reveal new information regarding the etiology of these behaviors. An array of new research techniques will allow genetic information to be linked with neuroimaging results and facilitate understanding of the phenotypic expression of stereotypic behaviors.

NOSOLOGY

In many cases, the origin of stereotypic movements seems contextual to certain psychiatric illnesses or medical conditions in which they occur. The evolution of the DSM has tried to reflect such insights in its various iterations. Recent versions of the DSM had reserved the diagnosis of stereotypic movement disorder for cases in which the stereotypic behavior is not a core symptom of another psychiatric disorder or the expected manifestation of a medical condition. Therefore, the repetitive facial movements due to either prolonged neuroleptic use or rapid discontinuation of this class of medication would be described as stereotypic movements, but these would not have met DSM-IV-TR criteria for stereotypic movement disorder. Although this continues to be the case in DSM-5, the current criteria requires the clinician to specify whether stereotypic movement is associated with a known medical or genetic condition, neurodevelopmental disorder, or environmental factor. For instance, a person with Down syndrome and stereotypic movements would be coded as (307.3) stereotypic movement disorder associated with Down syndrome. In the case of autism spectrum disorder, a comorbid diagnosis of stereotypic movement disorder should be made whenever the stereotypic behaviors cause injury or appear sufficiently severe to become a focus of treatment.

Also, the DSM-5 has placed more emphasis on culture-related issues. Although stereotypic movements, with or without self-injury, manifest in all races and cultures, the attitudes of a particular culture may delay diagnosis. Clinicians need to be aware of cultural context and remain sensitive as they explain the diagnosis to patients and their families who come from a different country or culture. This will help patients accept and maintain appropriate treatment for the condition.

At first, the DSM criteria were somewhat broad. The DSM-III provided criteria for a diagnosis of atypical stereotypic movement disorder for conditions such as rocking, head banging, and repetitive hand movements characterized by quick, rhythmic hand rotations or tentative voluntary movements that usually involved the arms or fingers. The DSM-III stipulated that such movements differed from tics in that they were

voluntary but not spasmodic or distressing. In the DSM-III, atypical stereotypic behaviors were considered often to occur in patients with extreme psychosocial deprivation, mental retardation, and pervasive developmental disorders; however, another mental disorder did not have to be present.

In the revised version of the DSM-III (DSM-III-R), atypical stereotypic movement disorder was renamed stereotypy/habit disorder. The criteria for stereotypy/habit disorder became broader, so that intentional, repetitive, nonfunctional behaviors such as nail biting, nose picking, skin picking, and object mouthing were included. The DSM-III-R also specified a severity criterion, which required the behavior to cause injury or markedly interfere with normal activities to be considered a clinically significant "disorder." The severity criterion allowed clinicians to determine whether the stereotypic movement was part of normal development, as is often seen with transient nose picking, or a more pathological picture, in which chronic nose picking causes significant tissue trauma or social stigmatization. A stereotypy or habit described by DSM-III-R criteria could include more intense forms of normal motor patterns seen in childhood, or they might be associated with some type of psychiatric or organic pathology. An additional criterion, however, excluded the diagnosis if the patient met criteria for pervasive developmental disorder or tic disorder. This exclusion criterion recognized and eliminated the redundancy of making the diagnosis of stereotypy/habit disorder in these disorders, which have repetitive behaviors as core symptoms.

Another name change occurred with the publication of the DSM-IV. In an effort to be compatible with the ICD-10, the DSM-IV renamed the disorder stereotypic movement disorder. Because the intentional nature of a stereotypic movement might be impossible to determine in certain cases —for example, individuals with severe mental retardation with limited ability to communicate, the clinician is no longer obligated to prove that the movement was intentional. The DSM-IV also drew boundaries between stereotypic movement disorders and several other disorders, including pervasive developmental disorder, tic disorder, obsessive-compulsive disorder (OCD), trichotillomania, and mental retardation. As in the DSM-III-R, the diagnosis is not made if the stereotypic behavior is attributable to pervasive developmental disorder or tic disorder. Given the similarity between tics and compulsions, the DSM-IV excluded stereotypic behaviors that were compulsions related to OCD. The DSM-IV also stipulated that the stereotypic behavior could not be due to the physiological effects of a substance or a general medical condition. A diagnosis of stereotypic movement disorder would not be made for repetitive hair pulling if DSM-IV criteria for trichotillomania were met. In addition, a diagnosis of stereotypic movement disorder with SIB could be made in persons with mental retardation if the SIB was significant enough to become a focus of

therapy. Subsequent versions of the DSM—the DSM-IV-R and the DSM-IV-TR—used those criteria established by the DSM-IV.

In general, the DSM-IV-TR and DSM-5 as well as ICD-10 criteria for stereotypic movement disorder describe the diverse and unrelated stereotyped movements that are not part of any other recognized neurological or psychiatric condition. According to the DSM-5, a stereotypic movement disorder is a repetitive, seemingly driven, and nonfunctional motor behavior. It lists hand shaking or waving, body rocking, head banging, mouthing of objects, self-biting, picking at skin or body orifices, and hitting one's body as examples. The DSM-5 definitions excludes behaviors with a clear purpose or adaptive function, but, as acknowledged in the DSM-IV-TR, the driving force behind a stereotyped movement may be difficult to determine in cases in which the person has limited ability to communicate internal states. What is more, a stereotypic behavior may provide a signal that an undiagnosed neurodevelopmental disorder exists, especially in children younger than 3 years of age.

The DSM-5 specifies that for a repetitive motor behavior to be considered a stereotypic movement disorder, it must interfere with social, academic, or other activities and may result in self-injury. In severe cases, the behavior can cause death, but even less severe cases might result in disabling or disfiguring impairments.

Consider the case of Ruth, a middle-aged woman, was admitted to the inpatient unit at a university-based teaching hospital due to unrelenting head scratching. Her outpatient psychiatrist had tried the full array of pharmacological interventions, including combinations of anxiolytic, antidepressant, and antipsychotic medications, but the stereotypic scratching progressed to the point that she required surgical intervention. Ruth had scratched through her scalp and penetrated the skull, exposing the dura mater.

In Ruth's case, a complete DSM-5 diagnosis would be Stereotypic Movement Disorder with Self-Injurious Behavior, Severe. Per the DSM-5, the diagnosis is specified as with or without SIB. Note, however, that the disorder should be specified as "with self-injurious behaviors," if injury would occur in cases that protective measures were not present to prevent injury. Ruth obviously displayed severe symptoms, given that continuous monitoring and protective measures were required to prevent serious injury. Moderate symptoms require protective measures and behavioral modification; whereas, mild symptoms can be easily suppressed with sensory stimulus or distraction. Clinicians should also specify whether the stereotypic movement is associated with a known medical or genetic condition, neurodevelopmental disorder, or environmental factor. An additional code identifying the associated medical/genetic condition or neurodevelopmental disorder aides aids in billing and communicating with third party payers.

The ICD-10 diagnostic criteria for Stereotypic Movement Disorders

appear more succinct. Per the ICD-10, the stereotyped movements must either cause physical injury or markedly interfere with normal activities. The duration of symptoms lasts at least 1 month. In contrast to DSM-5, no other mental or behavioral disorder can be present, except for mental retardation.

DIFFERENTIAL DIAGNOSIS

The outward appearance of a complex tic or a compulsion may closely resemble stereotypic movements, and this may lead to diagnostic difficulties. James W. Bodfish, T. W. Crawford, and colleagues found high rates of stereotyped behavior in adults with comorbid mental retardation and compulsion disorder. Compared with a noncompulsive matched control group, their sample showed an increased prevalence of stereotypy (32 percent), an increased prevalence of self-injury (40 percent), and an increased prevalence of stereotypy plus self-injury (37 percent) in those with compulsive disorder. A review by J. L. Rappoport reported that 20 percent of patients with OCD also exhibit motor tics, which complicates the diagnosis further. Nevertheless, some simple caveats can help to sort out stereotypic movements from other repetitive movements.

Generally speaking, a stereotypic movement is a rhythmic, repetitive motor behavior. It usually is simpler than other forms of perseverative behaviors. Although a stereotypic behavior appears to be purposeless, it can also serve to alleviate anxiety or another aversive state. Tics ordinarily have an involuntary quality even though tics may be suppressed for limited periods of time and they are not rhythmic. Compulsions appear more complex and ritualized. A stereotypic movement differs from a complex tic and a repeating compulsion in that it is not performed as a preventative measure, nor is it motivated by a sense of incompletion. In addition, complex tics and compulsive behaviors typically are not rhythmic.

With regard to autism, a recent study by Fiona Zandt and colleagues suggested that compulsions coexist with stereotypic behaviors in children with autism spectrum disorder. They found that children with OCD had more obsessions and compulsions than children with autism spectrum disorder; however, both the OCD group and the group with autistic stereotyped behaviors exhibited more obsessions and compulsions than a comparison group with a typical developmental trajectory. Similarly, J. M. Ringman and J. Jankovic reported on a series of 12 patients with autism spectrum disorders and stereotypic movements. These patients were referred to a specialized movement disorder clinic for the evaluation of tics. Tics were clearly present in seven of the patients, and six of these patients met criteria for both Asperger syndrome and Tourette syndrome. Even though the diagnosis of stereotypic movement disorder excludes repetitive movements that may be core symptoms of other DSM-5 disorders, the literature suggests and this author observes that stereotypic behaviors in

autistic spectrum disorders may be comorbid with OCD or tic disorder.

The DSM-IV-TR specifies that the stereotypic behavior not be due to the direct physiological effects of a substance or a general medical condition. This criterion would exclude tardive dyskinesias as in chronic neuroleptic exposure or skin picking secondary to stimulant use. Although tremors associated with Parkinson syndrome and Parkinson disease are abnormal and stereotypic, neither would fit the current criteria for stereotypic movement disorder because they are due to the physiological effects of neuroleptic medication in Parkinson syndrome and a medical condition in the case of Parkinson disease.

Mental retardation is a common feature of a number of genetic disorders. If the stereotypic or SIB occurs in the context of intellectual disability, the behavior no longer must be severe enough to become a focus of the treatment to meet diagnostic criteria for DSM-5. Repetitive motor behaviors noted in a person with intellectual disabilities would be correctly recorded per the DSM-5 as stereotypic movement disorder associated with intellectual developmental disability. In Lesch–Nyhan syndrome, the degree of mental retardation does not seem to predict the magnitude of self-mutilation. Some patients with mild mental retardation may display extreme forms of self-injury. Notably, stereotypic movements occur in approximately two-thirds of individuals with intellectual disabilities living in residential facilities.

In addition, the DSM-5 criteria no longer require that the repetitive movement must be present for 4 weeks or more. Typically, stereotypic movements peak in adolescence, but they have no typical age of onset. Stereotypic movements in persons with severe or profound mental retardation may last for years; however, an abrupt onset of stereotypic behaviors may signal the presence of an underlying acute medical condition in nonverbal individuals with severe mental retardation. Clinicians should be mindful of clues that such individuals may be experiencing pain, distress, anxiety, or mood shifts and then promptly treat any identified problems. Stressful events also may precipitate stereotypic behaviors. However, stereotypic movements may emerge and become a focus of treatment in individuals with mental retardation who are institutionalized in nonstimulating environments.

ICD-10

The ICD-10 defines stereotypic movements in childhood in a fashion similar to the DSM-5, since one of the goals of DSM-5 aimed at providing compatibility between the two diagnostic systems. As in the DSM, physical injury or marked impairment in normal activities must be present for the diagnosis. This criterion attempts to sort out stereotypic movement disorder from stereotypic movements that are benign or are typical of normal development. The duration criterion provides another dimension to

the impairment and ensures that the stereotypic movement is more than a passing novelty. In addition, history of another psychiatric or behavioral disorder should not be present and account for the stereotypic movement.

DIAGNOSIS

No clinical tests are available to make the diagnosis of stereotypic movement disorder. Nevertheless, the history and physical exam may be revealing. Dental and physical exams may provide evidence of tissue trauma that would raise suspicions for stereotypic behaviors, especially the self-injurious type. Anemia may be a sign of self-induced rectal bleeding. The delicate architecture of the eye is particularly vulnerable to repetitive trauma from behaviors such as head banging and eye poking. Sometimes a subtle physical sign clues the clinician to stereotypic behaviors and movements that go unobserved by others.

Rating scales assist in diagnosing and tracking response to treatment in many psychiatric conditions; however, rating scales for stereotypic behaviors have not been so helpful. A number of rating scales have been developed, but none provides optimal psychometric properties across the spectrum of stereotypic behaviors. Even commonly used instruments for tracking drug-induced stereotypic behaviors, such as the Angus Involuntary Movement Scale (AIMS) and the Barnes Scale, focus on involuntary movements rather than the wide range seen with stereotypic movements. Berkson and Davenport developed a stereotypy checklist to track stereotyped behaviors mainly due to social deprivation. Unfortunately, experiments intended to test the reliability and validity of the Berkson and Davenport Stereotypy Checklist have shown variable results.

Recently, video and computer-assisted technologies have been used to study tardive dyskinesia. These technologies may elucidate the characteristics of other stereotypic behaviors such as that developed for use in monitoring attention-deficit/hyperactivity disorder (ADHD). These motion-detecting devices connected to sophisticated computer programs can give detailed descriptions of the randomness, as well as the range of movements on a second-by-second basis. This type of technology might prove valuable in research or for clinical applications related to stereotypic behaviors.

Eventually, research may lead to feasible diagnostic tests using genetic and/or neuroimaging techniques to assist in the diagnosis and treatment of stereotypic behaviors. In the meantime, a comprehensive clinical exam remains the standard for making an accurate diagnosis.

INJURIOUS BEHAVIORS

Stereotypic movements can manifest in a variety of injurious behaviors that put patients and caregivers at risk. The ramifications of head banging range

from hematomas and head injuries to retinal detachment and blindness. However, even seemingly benign stereotypic behaviors could cause cumulative damage over time. For example, hair loss could be a consequence of chronic head rolling. In addition, parents and caregivers are at risk of accidental bites and bruises as they attempt to protect children and many adults who engage in stereotyped self-injury. Even children who typically do not exhibit self-injury can be hazardous to others. A child being held may surprise a parent or caregiver with spontaneous swift head bobbing that results in a broken nose or blackened eye. Stereotypic movements, with or without potential for self-injury, may cause major management problems, which hamper habilitation and stifle social adaptation.

Onychotillomania refers to compulsive picking or tearing at the nails. It can produce ridges and a central depression on the nail from chronic rubbing by the nail of another finger. Most commonly, the thumbnail of the dominant hand produces the trauma to the nail plate of the second or third finger. Onylysis refers to separation of the nail from the nail bed. Onychotillomania and onycholysis may be manifestations of rare conditions involving self-mutilation like Lesch–Nyhan syndrome and congenital insensitivity to pain.

Appropriate attention to the patient's history will help to sort out other stereotypic or SIBs. For example, a stereotypic behavior may have been present in a patient with schizophrenia prior to any neuroleptic medication exposure. A careful history of symptom onset can prevent a misdiagnosis of medication-induced stereotypy. Secondary gain in factitious disorders might be the driving force in some instances. In other cases, the inability to control anger or anxiety may lead to repetitive cutting as seen with patients with borderline personality disorder.

TREATMENT

A variety of treatment options from psychotherapy to pharmacotherapy have been used to treat stereotypic and SIBs. Even surgery and electroconvulsive treatments have been used in refractory cases. Behavioral treatments have been used with some success, and these include positive reward techniques such as differential reinforcement of behaviors other than the targeted maladaptive behavior (DRO) and differential reinforcement of behavior incompatible with the targeted maladaptive behavior (DRI). Rewards may be immediate or delayed, as in token economies, but they must be individualized so that they produce the maximum motivation. Rewards also may lose their novelty over time; therefore, changing rewards may provide a fresh incentive for motivation. Punishment can range from misting water into the face to applying mild electric shocks to the patient. Although these techniques might seem mean-spirited, they should be employed if self-injurious or stereotypic behaviors

have high potential for severe harm. Of course, the clinician should document a thoughtful risk versus benefits analysis and discuss this with the guardian or family before using such methods. Other forms of punishment include positive practice, removal of reinforcement, and restitution. A study by Jonathan M. Miller and others used behavioral modification techniques of habit reversal and differential reinforcement of other behaviors in an effort to suppress stereotypic behaviors in 12 nonautistic children between the ages of 6 and 14 years. After a mean follow-up time of 12.1 months, motor stereotypies were significantly reduced on the Stereotypy Linear Analog Scale and the Stereotypy Severity Scale. These results correlated with the subject's motivation and the number of sessions attended. The Child Global Assessment Scale also improved, but no correlations were seen with level of motivation or number of sessions.

Pharmacological approaches have been a mainstay treatment of stereotypic behaviors. An estimated 50 percent of persons with mental retardation living in residential or institutional settings are prescribed high doses of antipsychotic medication, but this statistic does not give an accurate account of the percentage of these patients who have a known psychotic illness. About 75 percent of patients receiving psychotropic medications develop some form of motor dysfunction. In the early days of treatment, movement disorders occurred at even higher rates due to the misconception that motor dysfunction was the definitive sign that a patient had achieved an adequate dose of medication to treat psychosis. Even though antipsychotic agents are used to treat stereotypic movements, they also can cause repetitive, purposeless movements. For instance, akathisia secondary to neuroleptic medication may appear clinically as a shifting from foot to foot, crossing and uncrossing of legs, pacing, hand wringing, or rocking, among other manifestations.

The newer generation of atypical antipsychotic medications does not pose the typical risk for motor dysfunction seen in older drugs. To date, risperidone (Risperdal) and aripiprazole (Abilify) have been approved by the U.S. Food and Drug Administration (FDA) for self-injurious and disruptive behaviors in children with autism. For the double-blinded, placebo-controlled clinical trial, the dose ranged from 0.5 to 3.5 mg per day in the 101 children studied. In a placebo-controlled study of chlorpromazine (Thorazine) used in doses ranging from 2 to 3 mg/kg per day, four adolescents and young adults with profound retardation exhibited a 35 percent reduction in stereotypic behaviors. Unfortunately, the improvement in stereotypic behaviors was offset by a cognitive decline. Several double-blinded, placebo-controlled studies used haloperidol (Haldol) in daily doses of between 0.5 and 4 mg to treat stereotypic behaviors in youth with autism. In these studies, haloperidol significantly reduced stereotypic behaviors without causing notable cognitive

impairment.

A variety of other agents have been tried, but these have fallen short due to limitations of study size or design. Although lithium carbonate (Eskalith) has been studied for SIBs in patients with mental retardation, the results of the few studies are inconsistent. Fenfluramine (Ponderax, Pondimin) reportedly reduced stereotypic behaviors and other problem behaviors in individuals with autism; however, the results also varied across studies. In addition, fenfluramine has been associated with cardiovascular and pulmonary problems. Several other serotonergic agents have been tried, including fluoxetine (Prozac), sertraline (Zoloft), citalopram (Celexa), trazodone (Desyrel), and buspirone (Buspar), but methodological issues related to the size or design of each study limited their usefulness.

Clomipramine (Anafranil) and desipramine (Norpramin) have been studied for their effect on stereotypic behaviors in autism and onychophagia. Gordon and coworkers showed that clomipramine was superior to desipramine, but both were better than placebo in reducing stereotypic behaviors associated with autism. Clomipramine was better than desipramine in treating severe onychophagia, but the dropout rate was greater than that seen in other studies of its use in OCD and trichotillomania. Henrietta L. Leonard and colleagues concluded that clomipramine was not the optimal treatment for severe onychophagia.

An endogenous opioid theory for SIBs has prompted the use of opioid antagonists. The original study of naloxone (Narcan) successfully reduced SIBs in two young adults with profound retardation. SIBs with low frequency of occurrence responded best to the intervention at low doses. Subsequently, several other studies showed that individuals responded variably to naloxone or naltrexone. However, these two agents may substantially decrease stereotypic, self-injurious, and hyperactive behaviors in some patients.

Clearly, more controlled studies are needed to evaluate these and other therapies based on behavioral and pharmacological interventions. Recent technological advances in genetic analysis, neuroimaging, and movement monitoring will promote this goal.

Although stereotypic behaviors manifest in many ways, diagnostic criteria for a stereotypic movement disorder in the DSM-5 and the ICD-10 describe repetitive, rhythmic movements that are not due to another psychiatric disorder, medical condition, or medication effect. Criteria for impairment help to clarify whether a stereotypic movement is an inconsequential habit or part of normal development. The careful clinician should use a thorough history and exam to identify the presence and possible causes of stereotypic behaviors. Once the clinician accurately characterizes the stereotypic movement, behavioral therapy with or without medication intervention may be applied. Although the literature on the treatment of stereotypic behaviors is limited, it does provide a starting

point for developing treatment plans that will likely be subject to change as warranted.

References

Bodfish JW, Crawford TW, Powell SB, Parker DE, Golden RN, Lewis MH. Compulsions in adults with mental retardation: prevalence, phenomenology, and comorbidity with stereotypy and self-injury. *Am J Mental Retard*. 1995;100:183–192.

Casey DE. Neuroleptic drug-induced extrapyramidal syndromes and tardive dyskinesia. *Schizophr Res*. 1991;4:109–120.

Fowles RE, Cloward TV, Yowell RL. Endocardial fibrosis associated with fenfluramine-phentermine. *N Engl J Med*. 1998;338(18):1316.

Gal E, Dyck MJ, Passmore A. The relationship between stereotyped movements and self-injurious behavior in children with developmental or sensory disabilities. *Res Dev Disabil*. 2009;30:342–352.

Gedye A. Anatomy of self-injurious, stereotypic, and aggressive movements: evidence for involuntary explanation. *J Clin Psychol*. 1992;48:766–778.

Gordon CT, State RC, Nelson JE, Hamburger SD, Rapoport JL. A double-blind comparison of clomipramine, desipramine, and placebo in the treatment of autistic disorder. *Arch Gen Psychiatry*. 1993;50(6):441–447.

Guess D, Carr E. Emergence and maintenance of stereotypy and self-injury. *Am J Ment Retard*. 1991;96:299–319; discussion 321–344.

Hewlett WA. Obsessive-compulsive disorder. In: Ebert MH, Loosen PT, Nurcombe B, eds. *Current Diagnosis and Treatment in Psychiatry*. New York: Lange Medical Books/McGraw-Hill; 2000:351.

Kolmen BK, Feldman HM, Handen BL, Janosky JE. Naltrexone in young autistic children: replication study and learning measures. *J Am Acad Child Adolesc Psychiatry*. 1997;36:1570–1578.

Lewis MH, Baumeister AA. Stereotyped mannerisms in mentally retarded persons: animal models and theoretical analyses. In: Ellis NR, ed. *International Review of Research on Mental Retardation*. New York: Academic Press; 1982.

McCracken JT, McGough J, Shah B, et al; Research Units on Pediatric Psychopharmacology Autism Network. Risperidone in children with autism and serious behavioral problems. *N Engl J Med*. 2002;347(5):314–321.

Miller JM, Singer HS, Bridges DD, Waranch HR. Behavioral therapy for treatment of stereotypic movements in nonautistic children. *J Child Neurol*. 2006;21(2):119–125.

Muehlmann AM, Lewis MH. Abnormal repetitive behaviours: shared phenomenology and pathophysiology. *J Intellect Disabil Res*. 2012;5(56):427–440.

Petty J, Allen D, Oliver C. Relationship among challenging, repetitive, and communicative behaviors in children with severe intellectual disabilities. *Am J Intellect Dev Disabil*. 2009;114:356–368.

Posey D, McDougle CJ. The pharmacotherapy of target symptoms associated with autistic disorder and other pervasive developmental disorders. *Harv Rev Psychiatry*. 2000;8(2):45–63.

Ratey JJ. *Mental Retardation: Developing Pharmacotherapies*. Washington, DC: American Psychiatric Press; 1991.

Richman DM, Lindauer SE. Longitudinal assessment of stereotypic, proto-injurious, and self-injurious behavior exhibited by young children with developmental delays. *Am J Ment Retard*. 2005;110:439–450.

Ringman JM, Jankovic J. Occurrence of tics in Asperger's syndrome and autistic disorder. *J Child Neurol*. 2000;15(6):394–400.

Sallustro F, Atwell CW. Body rocking, head banging and head rolling in normal children. *J Pediatr*. 1978;93:704–708.

Sprague RL, Newell KM. *Stereotyped Movements: Brain and Behavior Relationships*. Washington, DC: American Psychological Association; 1996:211.

Tanguay PE. Pervasive developmental disorders: a 10-year review. *J Am Acad Child Adolesc Psychiatry*. 2000;39(9):1079–1095.

Teicher MH, Polcari A, Foley M, et al. Methylphenidate blood levels and therapeutic response in children with attention-deficit hyperactivity disorder: 1. Effects in different dosing regimens. *J Child Adolesc Psychopharmacol*. 2006;16(4):416–431.

Thompson T, Gray DB. *Destructive Behavior in Developmental Disabilities*. Thousand Oaks, CA: Sage; 1994.

Tripp G, Luk SL, Schaughency EA, Singh R. DSM-IV and ICD-10: a comparison of the correlates of ADHD and hyperkinetic disorder. *J Am Acad Child Adolesc Psychiatry*. 1999;38(2):156–164.

Troster H. Prevalence and functions of stereotyped behaviors in nonhandicapped children in residential care. *J Abnorm Child Psychol*. 1994;22:79–97.

Volkmar F. *Autism and Pervasive Developmental Disorders*. Cambridge: Cambridge University Press; 1999.

Winchel RM, Stanley M. Self-injurious behavior: a review of the behavior and biology of self-mutilation. *Am J Psychiatry*. 1991;148:306–317.

Zandt F, Prior M, Kyrios M. Repetitive behaviour in children with high functioning autism and obsessive compulsive disorder. *J Autism Dev Disord*. 2007;37(2):251–259.

51

Mood Disorders in Children and Adolescents

▲ 51.1 Depressive Disorders and Suicide

KAREN DINEEN WAGNER, M.D., PH.D., AND DAVID A. BRENT, M.D.

DEPRESSIVE DISORDERS

Introduction

Depressive disorders in children and adolescents have a significant adverse effect on a youth's emotional, social, and cognitive development. Early onset of major depression is associated with more lifetime depressive episodes and suicide attempts, greater psychiatric and medical comorbidity, and poorer quality of life than later age onset of depression. Recognition of these disorders is essential so that children and adolescents may receive effective evidence-based treatments.

Definition

The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) includes in the category of depressive disorders the following: disruptive mood dysregulation disorder, major depressive disorder, persistent depressive disorder, premenstrual dysphoric disorder, substance/medication-induced depressive disorder, depressive disorder due to another medical condition, other specified depressive disorder, and unspecified depressive disorder. To be in the category of depressive disorders, there cannot be a history of ever having had a manic or hypomanic episode. There are coded specifiers for the depressive episode for severity (mild, moderate, severe, with psychotic features) and course (in partial remission, in full remission). The uncoded specifiers for the current episode include with anxious distress, with mixed features, with melancholic features, with atypical features, with mood-congruent psychotic features, with catatonia, with peripartum onset, and with seasonal pattern.

Disruptive mood dysregulation disorder in children is characterized by recurrent verbal or behavioral temporal outbursts (>3 or more times per

week) that are excessive and age inappropriate. In between these episodes, mood is irritable or angry most of the day nearly every day. These symptoms occur in at least two settings (school, home, or with peers). The duration of symptoms must be more than 12 months, without a 3 months symptom-free period in this time frame. Age of onset of the symptoms must be before 10 years old, and the diagnosis should not be made before age 6 years or after age 18 years.

Criteria for major depressive disorder in children and adolescents are at least one major depressive episode in which there has been at least a 2-week period of either depressed or irritable mood or a loss of interest or pleasure. In addition, at least four of the following symptoms are required: weight loss, weight gain, failure to make expected weight gain, or an increase or decrease in appetite; insomnia or hypersomnia; psychomotor agitation or retardation; fatigue or loss of energy; feelings of worthlessness or excessive or inappropriate guilt; diminished ability to think or concentrate or indecisiveness; and recurrent thoughts of death, recurrent suicidal ideation, suicide attempt, or suicide plan. The symptoms must cause impairment in functioning or cause clinically significant distress.

Persistent depressive disorder in children and adolescents is a chronically depressed or irritable mood for at least 1 year. In addition, there must be at least two of the following symptoms: poor appetite or overeating; insomnia or hypersomnia; low energy or fatigue; low self-esteem; poor concentration or difficulty making decisions; and feelings of hopelessness. During the 1-year period, the youth also has not been without depressive symptoms for more than 2 months at a time.

Premenstrual dysphoric disorder requires at least five symptoms in the final week before the onset of menses which start to improve a few days after the onset of menses and are either minimal or absent in the week following menses. One or more of the following symptoms must be present: affective lability, irritability/anger/interpersonal conflict, depressed mood/hopelessness/self-deprecating thoughts, or anxiety/tension/keyed up or on edge. In addition at least one of the following symptoms must be present: decreased interest, concentration difficulties, fatigue/low energy, appetite changes, hypersomnia/insomnia, feeling overwhelmed/out-of-control, and physical symptoms. The symptoms must be severe enough to interfere with functioning or cause clinically significant distress.

A diagnosis of other specified depressive disorder is used when full criteria for a depressive disorder are not met but the symptoms cause significant distress or impairment. The specific reason for not meeting full criteria is identified. The diagnosis of unspecified depressive disorder is used when the reasons for not meeting full criteria are not identified.

History

Although depression in children and adolescents is now recognized as a

common and serious illness, it was not until the 1980s that studies demonstrated that children do develop depressive disorders that can be diagnosed in a standardized manner. It was originally thought that children and adolescents developed a "masked" form of depression with predominantly behavioral symptoms, but subsequent studies have shown that children and adolescents have symptoms similar to those of adults with depression.

Comparative Nosology

The DSM-5 and DSM-IV-TR require five symptoms, one of which must be depressed or irritable mood or a loss of interest or pleasure over a 2-week period, for a diagnosis of major depressive disorder in children and adolescents. For a diagnosis of depressive episode, the tenth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) requires at least four symptoms, of which at least two of the following must be present: depressed mood, loss of interest or pleasure, and decreased energy or increased fatigability. Unlike DSM-5 and DSM-IV-TR, which allow specific depressive criteria for children and adolescents (irritable mood, failure to make expected weight gain), the ICD-10 has no specific criteria for depression in youth compared to adults. Similarly, DSM-5 persistent depressive disorder and DSM-IV-TR dysthymic disorder criteria allow a 1-year period of symptom duration for children and adolescents, compared to 2 years for adults. However, there are no differences in the minimum 2-year duration criterion for dysthymia in ICD-10.

The major differences between DSM-5 and DSM-IV-TR criteria for major depressive disorder and ICD-10 criteria for depressive episode are number of symptoms required, type and number of core symptoms, and allowing for mood to be irritable in children and adolescents. A study was conducted to determine whether there would be differences between the two diagnostic classification systems in diagnosing children with a depressive disorder. It was found that more children met DSM-IV-TR criteria for major depressive disorder than ICD-10 Diagnostic Criteria for Research (ICD-10-DCR) criteria for a depressive episode. Of the 33 children who had a diagnosis of major depressive disorder according to DSM-IV-TR criteria, 8 (24 percent) did not have a depressive episode according to ICD-10-DCR criteria because these children did not meet the criterion of at least two core symptoms. Using DSM-IV-TR criteria as the gold standard, the sensitivity for major depressive disorder with the ICD-10-DCR was 76 percent and the specificity was 99 percent. Therefore, use of the ICD-10 may miss the diagnosis of children who meet criteria for major depressive disorder.

Disruptive mood dysregulation disorder is a new diagnosis in DSM-5 depressive disorders to address issues related to potential overdiagnosis of

bipolar disorder in children. This diagnosis is not included in ICD-10. Premenstrual dysphoric disorder is a diagnosis in DSM-5 but had been in the DSM-IV-TR appendix for further study. Premenstrual dysphoric disorder diagnosis is not contained in the ICD-10.

Epidemiology

The prevalence of major depressive disorder in children is approximately 2 percent. In adolescence, there is a substantial increase in the prevalence of major depression, with prevalence rates of 4 to 8 percent. In children, the gender ratio is 1:1, whereas with puberty, adolescent girls are two to three times more likely to have an episode of major depression than adolescent boys. By age 18 years, the cumulative incidence of major depression is 20 percent. Children with a family history of affective disorders are about three times more likely to develop a major depressive disorder than children with a first-degree relative without a lifetime affective disorder.

The prevalence rate for persistent depressive disorder in children ranges from 0.6 to 4.6 percent, and in adolescents it ranges from 1.6 to 8.0 percent. Estimates of the prevalence rate of disruptive mood dysregulation disorder are up to 3 percent. It is more commonly diagnosed in boys compared to girls.

Etiology

Genetic. The most predictive factor for a child or adolescent to develop major depressive disorder is a high family loading for major depressive disorder. There is an increased risk of depression in relatives of depressed children (odds ratio = 3.98 compared to healthy controls) and among offspring of depressed adults (odds ratio = 2.12 compared with healthy controls). Younger age of onset of depression in parents increases the risk of depression in children. Parental major depression is associated with greater severity of illness, longer episode duration, more impairment, and more recurrences in depressed offspring. Twin studies have demonstrated that depression is moderately heritable, with rates about 31 to 42 percent.

Neurobiology. Neuroendocrine studies have evaluated the hypothalamic–pituitary–adrenal (HPA), hypothalamic growth hormone (HGH), hypothalamic–pituitary–thyroid (HPT), and hypothalamic–pituitary–gonadal (HPG) axes in youth to determine whether there are abnormalities related to depression. HPA-axis abnormalities have been detected in depressed children and adolescents with some evidence of nonsuppression in stimulation studies. Growth hormone–axis abnormalities have been found in prepubertal depressed children who have a blunted growth hormone response to stimulation with growth hormone-

releasing hormone. There have been inconsistent findings with regard to abnormalities in the HPT axis in children and adolescents with depression. The majority of studies of corticotropin-releasing hormone have found no abnormalities in depressed youth.

Abnormalities in the serotonergic system have been examined in depressed youth. Serotonin challenge agents that cause the release of prolactin have shown greater prolactin release in depressed children compared to controls.

Neuroimaging studies of depressed youth have found smaller left subgenual prefrontal cortex in adolescent girls, smaller frontal white matter volumes, larger frontal gray matter volumes, and larger lateral ventricular volumes. Amygdala functioning has been shown to be disrupted in depressed children. Disconnectivity in frontolimbic circuitry has been associated with depression in adolescents. Larger pituitary- and amygdala-to-hippocampal volume ratios have been found in depressed youth compared to healthy controls. Depressed children and adolescents have been found to have smaller left and right amygdala volume than healthy controls. Increased pituitary gland volume has been found in treatment-naive youth with major depressive disorder, particularly in males with nonfamilial major depressive disorder. Magnetic resonance spectroscopic imaging has found an increase in choline in the left dorsolateral prefrontal cortex and orbitofrontal cortex and increased choline/creatinine and choline/*N*-acetylaspartate ratios in the left anterior medial frontal lobe in depressed youth.

Sleep abnormalities have been examined in depressed youth, and the findings from polysomnography have been mixed. Some studies found no differences between children with major depressive disorder and controls; whereas other studies have found shorter rapid eye movement (REM) latencies, longer sleep latency, higher number of REM periods, and increased REM time for depressed children. In studies of depressed adolescents, shorter REM latency, more REM sleep, more REM density, and longer sleep latency have been reported. Subjective reports by youth of sleep disturbance (sleep quality, waking in the middle of night, morning tiredness) were not confirmed by electroencephalogram (EEG). Of interest, the depressed children who rated their sleep most poorly had better sleep by EEG measures than those who rated their sleep more positively. In a study using polysomnography as a measure of recurrence of depression in youth, those youth who had a later recurrence had decreased sleep efficiency and delayed sleep onset at baseline sleep polysomnography.

Environmental. Adoption studies have supported the role of environment more than genetic factors in the development of pediatric depression. Family factors, including parental substance abuse, poor marital adjustment, parent–child conflict, and lack of family cohesion, are

risk factors for depression in children and adolescents. Life stressors, such as abuse, neglect, death of a parent, and psychosocial adversities, may contribute to depression in youth. Cigarette smoking, physical illness, and functional impairment have been associated with youth depression. Recurrence of major depression may also be affected by environmental stressors.

There may be gene–environment interactions with regard to pediatric depression. Twin studies of life events have shown that the effect of life events on depression is greater in youth with higher genetic liability. Family conflict and family adversity have also been demonstrated to interact with genetic liability and contribute to depression in youth.

Cognition. Depressed children and adolescents may have a negative view of themselves, the world, and the future. They may possess a negative attributional style in which they view themselves as the cause of negative outcomes and view positive outcomes as externally derived. Low self-esteem, feelings of worthlessness, and perceived incompetence are some outcomes of these cognitive distortions. Cognitions about self-worth and self-efficacy have been found to be mediators of stressful life events in the development of depression. The combination of stressful life events and negative self-concept has been shown to be predictive of depression in youth.

Diagnosis and Clinical Features

It is important for the clinician to conduct a comprehensive evaluation of a child to ensure the accuracy of the primary diagnosis, as well as of comorbid diagnoses. The clinician should interview the child and parent separately so that each can express their concerns. Teenagers are often reluctant to disclose their feelings or issues related to peers in the presence of their parents. Similarly, parents may be reticent to discuss sensitive topics with their children present. Other sources of information may be obtained from discussion with school personnel, from school records, or from medical records.

Children with the diagnosis of disruptive mood dysregulation disorder are persistently irritable or angry most of the day nearly every day and have severe recurrent (>3 times per week) temper outbursts, either verbal or behavioral. These temper outbursts are developmentally inappropriate and are out of proportion to the situation. For example, a child who is chronically irritable at school may throw a chair at a peer during one of these temper outbursts. To make this diagnosis, the symptoms must be present for at least 12 months (no period >3 months without symptoms in this time frame) and be present at least two settings (school, home, or with peers). The diagnosis cannot be given to children under the age of 6 years, and the onset of symptoms must be before age 10 years. Disruptive mood

dysregulation disorder cannot co-occur with oppositional defiant disorder, bipolar disorder, or intermittent explosive disorder. If a child has symptoms that meet criteria for both disruptive mood dysregulation disorder and oppositional defiant disorder, then the diagnosis of disruptive mood dysregulation disorder should be given. If the child has ever had a manic or hypomanic episode, then the diagnosis of disruptive mood dysregulation disorder should not be given.

Major depression in children and adolescents is defined using DSM-5 criteria of at least 2-week duration of depressed or irritable mood and/or loss of interest or pleasure, along with at least four of the following symptoms: Weight or appetite change, sleep disturbance, psychomotor retardation or agitation, fatigue or loss of energy, feelings of worthlessness or guilt, diminished concentration, and suicidal ideation, intent, or plan. The symptoms must cause impairment in the child's functioning with regard to, for example, social or school performance or peer relationships, which is critical to the diagnosis in youth. Impairment helps to distinguish these symptoms from a passing phase of childhood or adolescence.

Youth with major depressive disorder often display irritable mood rather than dysphoria. Typically they are not aware of the extent of their irritability or its effects on their interaction with other people. Teenagers who have some insight into their irritability might say that everything makes them angry whether it is important or not. Loss of pleasure or interest may be noted by the youth's withdrawal from school or sports activities and from friends. Sleep disturbance is common in depressed youth, particularly difficulty falling asleep. Weight loss or failure to gain weight is more common than significant weight gain. Youth with depression are often fatigued and may nap after school. Diminished concentration may be manifested by declining school grades. A child often describes a feeling of worthlessness as no one liking him or her, including peers, teachers, and family members. Suicide attempts and psychotic features are more common in depressed adolescents than in depressed children.

Depressed youth typically do not view themselves as having depression, because their mood is usually irritable rather than depressed. Parents often do not recognize individual symptoms of an adolescent's depression, but they recognize the burden on the family in dealing with an extremely irritable adolescent. Children and adolescents with depression are more likely to be brought for evaluation because of declining school grades, substance abuse, suicide attempt, or a change in behavior.

For a diagnosis of persistent depressive disorder, a child or adolescent must have depressed or irritable mood for at least 1 year. In addition, the youth must have two of the following symptoms: poor appetite or overeating, insomnia or hypersomnia, low energy or fatigue, low self-esteem, poor concentration or difficulty making decisions, and feeling of

hopelessness. Youth with persistent depressive disorder have fewer melancholic and neurovegetative symptoms than youth with major depressive disorder. Although the depressive symptoms are milder in persistent depressive disorder than in major depressive disorder, the chronicity of persistent depressive disorder results in significant psychosocial impairment (e.g., relationships with friends and family).

Depressive disorders in youth are often comorbid with other psychiatric conditions. Rates of comorbidity range from 40 to 90 percent of children and adolescents with a depressive disorder, with up to 50 percent of depressed youths having two or more comorbid disorders. Anxiety disorders often precede the onset of depression and are a very common comorbid disorder. Major depression and persistent depressive disorder also frequently co-occur in youth. Other common comorbidities are attention-deficit/hyperactivity disorder, oppositional defiant disorder, conduct disorder, and substance use disorder. It is important for clinicians to identify comorbid disorders that are associated with depressive disorders in youth, since these comorbidities may adversely affect the course and treatment outcomes of depressive disorders.

Laboratory Examinations/Assessment Instrument

No laboratory findings are diagnostic of depressive disorders in youth. If a child presents with symptoms of depression and hypothyroidism (e.g., dry skin, feeling cold, and memory difficulties), then thyroid function tests should be obtained. If hypothyroidism was found, the diagnosis would be depressive disorder due to another medical condition.

The Children's Depression Rating Scale-Revised (CDRS-R) is a 17-item, clinician-administered instrument that assesses symptoms of depression in children and adolescents. The scale assesses cognitive, somatic, affective, and psychomotor symptoms. The items are rated for severity on a scale of 1 to 7 for some items and 1 to 5 for other items for a total score of 113. Typically a score of 45 or greater is considered a marker of clinically significant depression. The CDRS-R is useful for obtaining baseline symptomatology prior to treatment and for monitoring treatment outcome.

Given that youth with bipolar disorder may present with a depressive episode, it is important to ask about symptoms of mania or hypomania. Some diagnostic checklists for mania have been shown to discriminate bipolar from nonbipolar disorder in youth of ages 5 to 18 years old, including the Mood Disorder Questionnaire-Adolescent (MDQ-A), the Parent General Behavior Inventory 10-item short form (PGBI-SF10), and the Parent Young Mania Rating Scale (P-YMRS). These checklists are completed by the parents about their child's behavior. Adolescent self-report measures about symptoms of mania have demonstrated low validity in identifying bipolar disorder.

Differential Diagnosis

Depressive symptoms may be related to stressors and fulfill criteria for an adjustment disorder with depressed mood. A child may have experienced the death of a significant person in his or her life, and the depressive symptoms may be related to bereavement. Depression or irritability may be attributable to substance use or withdrawal, which would be considered a substance-induced depressive disorder.

When a child presents with depressive symptoms, it is important for the clinician to consider bipolar disorder within the differential diagnosis. The clinician should ask about a past or present history of manic or hypomanic symptoms.

There is substantial symptom overlap among various psychiatric disorders, which might complicate the differential diagnosis. Irritability and sleep disturbance may be found in depression, persistent depressive disorder, bipolar disorder, posttraumatic stress disorder, and generalized anxiety disorder. Appetite changes may be associated with eating disorders. Children with attention-deficit/hyperactivity disorder who are failing in school and having difficulty with peer relationships may exhibit symptoms such as depressed mood and low self-esteem. Anxiety disorders and oppositional defiant disorders may have some symptom features of depression. Children with pervasive developmental disorder may present with depressive symptoms but not fulfill diagnostic criteria for major depression.

Medical conditions may manifest with symptoms of a depressive episode and should be diagnosed as depressive disorder due to a general medical condition if the mood disorder is attributable to the consequences of the medical condition. Neurological disorders (e.g., seizure disorder and multiple sclerosis), endocrinological abnormalities (hypothyroidism, hypopituitarism, Cushing disease), infections (e.g., pneumonia, AIDS, and mononucleosis), autoimmune diseases, and cancer are some medical conditions to consider within the differential diagnosis. Medications such as anti-inflammatory agents, antibiotics, steroids, stimulants, and oral contraceptives may lead to depressive symptoms.

Course and Prognosis

Disruptive mood dysregulation disorder onset must be before age 10 years old. The symptoms are predominantly limited to childhood. Of children who were diagnosed with disruptive mood dysregulation disorder, only 19 percent continued to meet criteria for the disorder at 12 and 24 months follow-up. Children with disruptive mood dysregulation disorder are at risk to develop depression and anxiety disorders in adulthood.

The mean length of an episode of major depression in children and adolescents is 8 months. Recovery rates from an index episode of

depression are high in youth, even independent of treatment. More than 90 percent of depressed children and adolescents typically recover from an episode of major depression within 1 to 2 years.

Relapse rates, with relapse being defined as a return of symptoms in the index episode, are between 34 and 50 percent, with the greatest likelihood of relapse in the first 6 months to 1 year after treatment discontinuation. Recurrence rates (onset of a new episode of depression) are high, with rates ranging from 20 to 60 percent after 2 years of remission and up to 70 percent at 5-year follow-up.

Risk factors for relapse or recurrence include severity of illness, suicidality, comorbid disorders, residual symptoms, psychosocial stressors, family conflict, family history of recurrent depression, negative cognitive style, low socioeconomic status, early onset of parental mood disorder, baseline social difficulties, and history of sexual abuse.

Youth who present with an episode of major depression are at risk of the development of bipolar disorder. Overall estimates are that 20 to 40 percent of youth with a major depressive episode are at risk of subsequent development of bipolar disorder. In a longitudinal study of prepubertal children diagnosed with major depression, the rate of bipolar I disorder was 33 percent and that of bipolar spectrum disorders (I, II, not otherwise specified [NOS]) was 48 percent by young adulthood. Risk factors for switch to bipolar disorder are prepubertal onset of major depression, loaded or multigenerational family history for mood disorders, psychotic features, and hypomania when treated with antidepressant medications.

There is increased recognition of the continuity between depression in youth and depression in adulthood. In a follow-up study of depressed school-aged children in community and clinic samples, 18 percent continued to have major depression into adulthood. In a clinic sample, adolescent-onset major depressive disorder increased the risk twofold of major depressive disorder in adulthood. In a community sample, 53 percent of adolescents with a history of major depressive disorder developed a new major depressive disorder in young adulthood.

Early age of onset of depression increases the risk of anxiety disorders, substance abuse, psychiatric comorbidity, suicide attempts, psychiatric and medical hospitalization, impairment in relationships with parents, siblings, and friends, more lifetime depressive episodes, nicotine dependence, increased adulthood body mass index, educational underachievement, unemployment, early parenthood, and worse overall quality of life.

The duration of persistent depressive disorder ranges from 2 to 3 years in clinically referred and community-based youth. After recovery after an episode of persistent depressive disorder, the estimated relapse rate is 45 percent. Youth with persistent depressive disorder are at increased risk of psychosocial difficulties, substance use disorder, episodes of major depression, and suicidality. The risk of suicide attempts in youths with

persistent depressive disorder increases with comorbid conduct disorder or substance use disorder.

Children and adolescents with both major depressive disorder and persistent depressive disorder have a worse course of illness than those with either disorder alone. These youth have more-severe depression, longer depressive episodes, more comorbid disorders, more suicidality, and greater psychosocial impairment.

Pharmacological Treatment

There are no FDA-approved medications to treat disruptive mood dysregulation disorder. Since this is a new childhood diagnosis, there are no studies of medication treatments for this disorder. Pharmacological treatment recommendations range from stimulants to treat aggression, mood stabilizers such as atypical antipsychotics and valproate to treat mood dysregulation, or selective serotonin uptake inhibitors to treat depressive symptoms.

Selective serotonin reuptake inhibitors (SSRIs) are the only antidepressants that have demonstrated efficacy in double-blind placebo-controlled trials for children and adolescents with major depression. Two SSRIs have FDA approval for acute and maintenance treatment of major depression in children and adolescents: fluoxetine (>8 years old) and escitalopram (>12 years old).

SSRIs that have also demonstrated efficacy in controlled studies of depressed youth are citalopram and sertraline. Other studies of SSRIs failed to demonstrate superiority of medication to placebo on the primary outcome measure, including citalopram, paroxetine, and escitalopram. A meta-analysis of SSRI trials in youth showed superiority of SSRIs to placebo, with an average response rate of 60 percent for antidepressants and 49 percent for placebo.

Other classes of antidepressants have not shown statistically significant improvement with antidepressant medication compared to placebo, including venlafaxine, nefazodone, and mirtazapine and selegiline transdermal system. In a failed study, duloxetine and fluoxetine were not superior to placebo. Tricyclic antidepressants have not been shown to be effective for the treatment of depressed children and adolescents. There is some speculation that the early maturation of the serotonergic system compared to the noradrenergic system may account for the lack of efficacy in noradrenergic antidepressants compared to serotonergic agents in youth. There are no controlled trials of bupropion in the treatment of major depression in children and adolescents.

There has been much discussion about why controlled studies of antidepressants in children and adolescents with major depression have not shown efficacy, that is, statistically significant separation from placebo. One explanation is the high rate of placebo response in these studies—up to

59 percent. Other explanations are methodological issues in the antidepressant trials, including number of sites, experience of site personnel, patient recruitment source, inclusion and exclusion criteria, duration of treatment trial, dosing, and outcome measures.

Adverse Events. Common side effects of SSRIs are nausea, abdominal pain, headaches, dry mouth, insomnia, somnolence, and sexual dysfunction. Impulsivity, agitation, irritability, and behavioral activation may occur, particularly with young children. Potential serious side effects include serotonin syndrome, increased bleeding, and amotivational syndrome. Common side effects of venlafaxine are anorexia, abdominal pain, insomnia, somnolence, dizziness, and elevated blood pressure. Common side effects of mirtazapine are increased appetite, weight gain, and somnolence. Common side effects of duloxetine are nausea, headache, abdominal pain, and somnolence.

FDA Warning and Suicidality. In a combined analysis conducted by the FDA of 24 short-term, placebo-controlled trials of antidepressant medications in child and adolescent major depressive disorder, obsessive-compulsive disorder, or other psychiatric disorders, the risk of suicidality (suicidal thinking and behavior) was 4 percent—twice the placebo rate of 2 percent. There were no suicides in any of the clinical trials. The FDA directed manufacturers to add a black box warning to the health professional label of antidepressant medications to describe the increased risk of suicidal thoughts and behavior in children and adolescents being treated with antidepressants, as well as to recommend the need for close monitoring of patients on antidepressant medications.

In a later meta-analysis of 27 trials of pediatric major depression, the risk of suicidality (suicidal ideation and attempts) was 3 percent in the antidepressant group and 2 percent in youth who received placebo. The pooled absolute rate of antidepressant response was 61 percent, and the placebo response was 50 percent. The number needed to treat (NNT) (number needed to treat to get one positive response to medication) was 10, and the number needed to harm (NNH) (number needed to get one event of suicidality) was 112. Based on this finding, the investigators concluded that the benefits of antidepressants in youths outweigh the potential risk from suicidal ideation or attempts.

There have been recent reports of an increase in suicide rate in youth following the black box warning on antidepressants and a significant decrease in the rates of diagnosis and treatment of pediatric depression.

A medication guide that addresses antidepressant medications, depression, and suicidality has been approved by the FDA. The guide states that depression and other serious mental illnesses are the most important causes of suicidal thoughts and actions and that antidepressant medication

may increase suicidal thoughts or actions in some children and teenagers when medication is first started. The patient or family member is encouraged to contact the health care provider if there is a change in mood or thoughts.

Informed Consent. Prior to prescribing antidepressant medication to a child or adolescent, an informed consent process should occur with the youth and the parent. Elements of informed consent are the purpose of treatment, description of the treatment process, benefits and risks of the proposed treatment, and alternative treatments. Because there is minimal information about the long-term effects of antidepressant medication in children, parents should be informed that there may be unknown risks of these medications.

Medication Treatment Algorithm. An evidence-based consensus algorithm for the treatment of major depression in children and adolescents was reported in which four stages of medication treatment were identified. It was recommended that if a child failed to respond to treatment in one stage, then the clinician should move to treatment in the next stage. The first stage of treatment is with an SSRI. The second stage of treatment is with an alternative SSRI. The third stage of treatment is selection of a different class of antidepressant medication (venlafaxine, bupropion, mirtazapine, duloxetine). If a child fails to respond to these agents, then it was recommended that there be reassessment of the diagnosis and treatment consultation.

In the Treatment of Adolescents with SSRI-Resistant Depression (TORDIA) study, it was found that there was no difference in response rates for those adolescents switched to venlafaxine or another SSRI (48 vs. 47 percent) after treatment failure with an SSRI. Therefore, if an adolescent does not respond to the initial antidepressant, there is approximately a 50 percent chance of responding to the second antidepressant.

Table 51.1–1.
Dosing of Antidepressants in Children and Adolescents

	Typical Starting Dose (mg)		
Medication	**Child**	**Adolescent**	**Target Dose (mg)**
Citalopram	5–10	10	20–40
Escitalopram[a]	5	10	10–20
Fluoxetine[b]	5–10	10	20–40
Paroxetine	5–10	10	20–40
Sertraline	25	50	100–200
Bupropion	50 twice daily	50 twice daily	100–200
Duloxetine	20	20	60–120
Mirtazapine	15	15	30–45

Venlafaxine	37.5	37.5	150–225

[a]Approved by the U.S. Food and Drug Administration for the treatment of major depression for children ages 12 years and above.

[b]Approved by the U.S. Food and Drug Administration for the treatment of major depression for children ages 7 years and above.

Lower initial starting doses of antidepressants are typically used in children and adolescents than in adults. However, the dose range tends to be similar to that used with adults. Antidepressant doses in children and adolescents are shown in Table 51.1–1. Doses are usually increased by 3 weeks if there is no symptom improvement. If there is no improvement with additional dose adjustment after 8 to 10 weeks, then switching to another treatment is recommended. Antidepressants should be continued for 6 to 12 months after symptom remission. At the time of discontinuation of an antidepressant, the dose should be tapered slowly, usually over a 2- to 3-month period.

Psychotherapy

The two most-studied treatment modalities for depression in adolescents are cognitive-behavior therapy (CBT) and interpersonal psychotherapy (IPT). There are no published controlled trials of these treatments in children with major depression.

Cognitive-Behavior Therapy. CBT focuses on cognitive distortions and behavioral deficits associated with depression in youth. Components of CBT for children and adolescents include psychoeducation and mood monitoring, pleasant activity scheduling and behavioral activation, and cognitive restructuring. Relaxation training and improvement in social and problem solving skills may be included.

The efficacy of CBT was assessed in adolescents with major depression or dysthymia. The youth were randomly assigned to adolescent group CBT, adolescent group CBT with a separate parent group, or wait-list control. The CBT groups showed higher depression recovery rates (67 percent) than the wait-list group (48 percent) and greater reduction in self-reported depression. The addition of a parent group did not affect outcome.

CBT (12 to 16 sessions) was compared to nondirective supportive psychotherapy and systemic behavior therapy in a randomized trial. CBT was significantly superior to the other treatments, with greater reduction in depression, more rapid symptom relief, and a higher remission rate—65 percent, compared to 39 percent for nondirective supportive therapy and 38 percent for systemic behavior family therapy. However, in a 2-year follow-up study, there were no significant differences in outcome among the psychotherapies. Eighty percent of the adolescents had recovered from depression, but 30 percent had a recurrence of depression.

Computerized CBT in the form of an interactive fantasy game (SPARX) has been developed for adolescent depression. A study comparing SPARX to treatment as usual for adolescents with depressive symptoms showed similar treatment response rates.

The Treatment for Adolescents with Depression Study (TADS) was the largest randomized, controlled trial to compare the efficacy of CBT, fluoxetine, CBT plus fluoxetine, and placebo in adolescent outpatients with a diagnosis of major depression. Response rates (much or very much improved) for CBT (43 percent) were not statistically significant compared to those for placebo (35 percent). Response rates to fluoxetine plus CBT (71 percent) and fluoxetine alone (61 percent) were significantly superior to those to CBT alone and to placebo. Fluoxetine with CBT showed the greatest reduction in suicidal thinking. Although remission rates at the end of 12 weeks of treatment were highest in the combination group (37 percent) compared to the other treatment groups (fluoxetine, 23 percent; CBT, 16 percent; placebo, 17 percent), these overall remission rates were low. The long-term effectiveness of these treatments was assessed over 36 weeks. It was found that for adolescents with moderate to severe depression, treatment with fluoxetine alone or the combination of fluoxetine with CBT accelerated the improvement of depression compared to CBT alone. The addition of CBT to fluoxetine decreased persistent suicidal ideation and treatment-emergent suicidal events.

Although the TADS study showed some benefit to the combination of CBT plus fluoxetine, other trials failed to show a significant advantage of adding CBT to medication for depressed adolescents. In one study, for adolescents with moderate to severe major depression, there was no difference in outcome at 28 weeks between youth who received combination CBT plus an SSRI and youth who received an SSRI with routine clinical care. In addition, there was no evidence of a protective effect of CBT on suicidal thinking or action. In another study, it was found that adding CBT to SSRI treatment in a pediatric primary care setting showed no significant improvement in depression compared to primarily SSRI medication.

In the TORDIA study, adolescents who did not respond to initial treatment with an SSRI had a 55 percent response to CBT added to an alternate medication compared to 41 percent for those treated with an alternate medication only.

Interpersonal Psychotherapy. IPT focuses on the relationship between depressed symptoms and interpersonal relationships. Four areas of focus are loss, interpersonal disputes, role transition, and interpersonal deficits. IPT has been modified for depressed adolescents (IPT-A) and has added a focus on developmental issues such as separation from parents, authority figures, peer pressures, and dyadic relationships. The adolescent and

parents are educated about the association between depressive symptoms and events in interpersonal relationships. The adolescent learns skills to improve interpersonal relationships that may help to relieve depression.

The efficacy of IPT-A was assessed in a 12-week study of 48 adolescents with major depression who were randomly assigned to IPT-A or clinical monitoring. IPT-A led to greater decrease in depressive symptoms, improvement in social functioning, and improved problem solving compared to clinical monitoring.

The effectiveness of IPT-A in school based health clinics was evaluated in depressed adolescents who were randomly assigned to IPT-A or treatment as usual for 16 weeks. The IPT-A training was done with existing staff at the health care clinics in high school. The adolescents who received IPT-A had greater symptom reduction and improvement in overall functioning. The greatest treatment effects occurred in older and more severely depressed adolescents.

There is also a group format (IPT-AG) that allows for peer support for depressed adolescents who have similar interpersonal problems. Adolescents can receive feedback from their peers about their interpersonal interactions and strategies to improve them. The efficacy of IPT-AG was compared to individual IPT-A, and there was similar improvement in depression for both formats of IPT.

Psychotherapies. In a meta-analysis of psychotherapy for depression in children and adolescents, the overall effect size was small (0.34). CBT was not superior to other psychotherapy treatments. It was also found that the treatment effect diminished over time to the point of no effect at 1-year follow-up. Another meta-analysis of randomized, controlled trials found that psychotherapy, particularly CBT and IPT, was more effective than no-treatment, waiting-list, or attention-placebo approaches at posttreatment for youth with depression. However, the effectiveness of psychotherapy was no longer significant at 6-month or later follow-up. The investigators concluded that psychotherapy appears to be helpful for treating depressed youth in the short term but that more long-term effective treatment strategies are needed.

Other Treatment Components

Treatment of depressed children and adolescents should also include psychoeducation, supportive management, family involvement, and school involvement.

Psychoeducation involves educating the patient and parents about depression, including causes, symptoms, course, and treatment options, including benefits and risks of treatment. Depression should be explained as an illness not a character flaw of the child or parent. The parent and child should be informed about the importance of adhering to treatment to

ensure an optimal outcome.

Supportive management includes assisting the child with problem solving, improving coping skills, and fostering involvement in treatment.

It is important to involve parents in the treatment process. Parents can provide important information about the child's symptoms over the course of treatment. Parents may also need guidance in managing their depressed child at home.

With regard to school involvement, children with depression may have academic difficulties that require school remediation or accommodation in the child's school schedule or program.

Clinical Issues Affecting Treatment Response

When a depressed child fails to respond to evidence-based treatment, there may be clinical issues that are affecting treatment outcome. The diagnosis should be reassessed because there might be symptom overlap among disorders that could lead to misdiagnosis. Comorbid disorders might have been unrecognized and the depression treatment response adversely affected by lack of treatment for the comorbid disorder. Psychosocial issues such as abuse, family conflict, parental psychopathology, peer difficulties, and family discord might be exacerbating the symptoms of depression. Lack of compliance with treatment will affect treatment outcome. Some youth do not take medication as prescribed and are reluctant to inform the clinician. Similarly, some youth do not perform behavioral exercises or assignments that might be a component of psychotherapy treatment. With regard to medication management, it is important for parents to participate in monitoring their child's medication compliance.

Treatment Recommendations

The American Academy of Child and Adolescent Psychiatry published treatment recommendations for children and adolescents with depressive disorders. For children with uncomplicated or brief depression or with mild psychosocial impairment, psychoeducation, supportive treatment, and case management are generally adequate initial treatment. For children who do not response to supportive psychotherapy after 4 to 6 weeks, a trial of specific psychotherapy and/or antidepressants is indicated. Similarly, for children with moderate to severe depression, chronic or recurrent depression, significant psychosocial impairment, suicidality, agitation, or psychosis, it is recommended that psychopharmacological treatment or specific types of therapy be initiated. CBT or IPT alone may be effective for moderate depression. However, more severe depression typically requires treatment with antidepressants. For children who do not respond to either form of treatment alone, combination with psychotherapy or antidepressants may be required.

SUICIDAL BEHAVIOR IN CHILDREN AND ADOLESCENTS

Introduction

It is appropriate to discuss child and adolescent suicidal behavior together with depression because depression is the most significant psychiatric risk factor for suicidal behavior, and because suicidal behavior is a symptom of depression. Nevertheless, the vast majority of depressed individuals neither complete nor attempt suicide, and many children and adolescents who engage in suicidal behavior do not have a mood disorder. Moreover, even in those depressed individuals who also manifest suicidal behavior, it is still an open question as to whether the treatment of depression is sufficient to reduce suicidal risk.

Definitions

Suicidal phenomena occur on a spectrum ranging from suicidal thoughts to suicide. The category of *suicidal thoughts* or *suicidal ideation* can vary from passive thoughts of death (e.g., maybe I would be better off dead), to active suicidal ideation (e.g., I would like to kill myself), to suicidal ideation with a *plan and intent* (e.g., I plan to kill myself with my hunting rifle when my parents go out tonight). *Suicidal threats* are the verbalization of suicidal ideation. A *suicide attempt* can be defined as "self-destructive behavior with inferred or explicit intent to die." This definition differentiates suicidal behavior from "*non-suicidal self-injurious behavior (NSSI)*," which involves intentional self-injury such as cutting for motivations other than to die, most commonly, to relieve negative affect. *Suicide* is a suicide attempt that results in death. *Autoerotic asphyxia* can superficially resemble suicide by hanging. In autoerotic asphyxia, the individual partially cuts off carotid circulation using a neck ligature simultaneous with masturbation to heighten the intensity of sexual climax. The presence of a ligature, sexual paraphernalia, and evidence of sexual activity support the latter condition over suicide.

Epidemiology

The Youth Risk Behavior Survey for 2014, which gathered data on risk behaviors using self-report, found that, in the past year, 13.6 percent of youth had a suicidal plan, 8.0 percent made a suicide attempt, and 2.7 percent made an attempt requiring medical attention. Lifetime rates of suicidal behavior using data obtained by interview rather than self-report yield lower estimates of lifetime suicidal ideation with plan and attempts, both around 4 percent. Deaths by suicide are much more rare, occurring in 8.26/100,000 of adolescents 15 to 19 and 1.87/100,000 of adolescents aged 10 to 14. The rates of suicidal ideation, attempts, and deaths by suicide increase with age across childhood and adolescence. The increase in

ideation is explained by an increased incidence of depression that occurs during adolescence. The increased rate of attempts and suicide deaths in adolescence relative to childhood are explained by higher rates of depression and alcohol and substance abuse, as well as by a greater ability to plan and execute a potentially lethal suicide attempt.

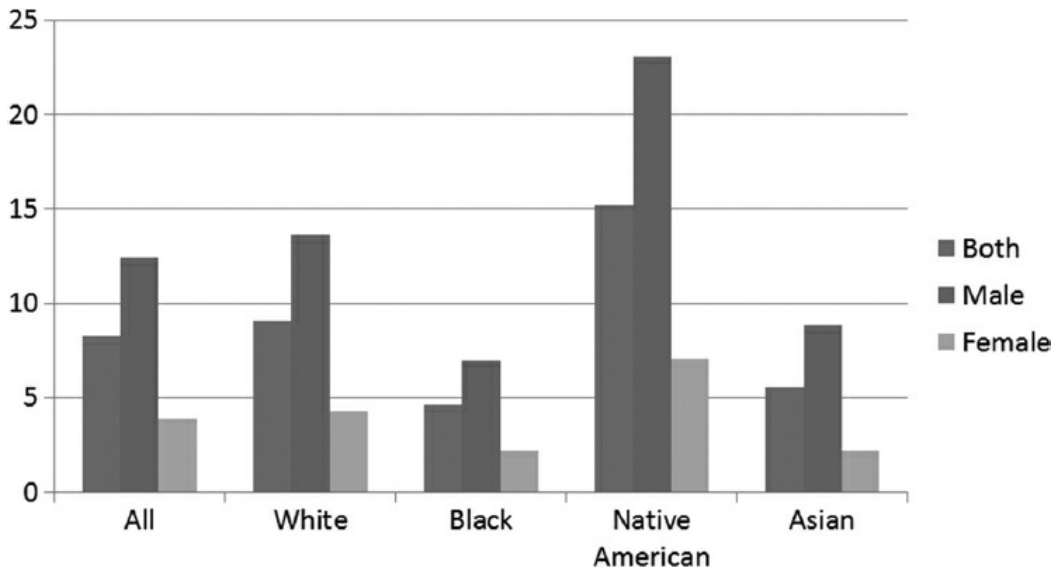

FIGURE 51.1–1. US suicide rate in 15 to 19 year olds (per 100,000) by sex and race (2013). (Accessed Injury Prevention & Control: Data & Statistics (WISQARS) http://www.cdc.gov/injury/wisqars/index.html. Updated July 13, 2015.)

Suicidal ideation and attempts are more common in female than male adolescents after puberty, following a similar pattern to depression. However, suicide deaths are over twice as common in males than females aged 10 to 14, and over three times as common in males than females aged 15 to 19. The higher rate of suicide in late adolescent males relative to females may be due to higher rates of substance abuse and conduct disorder and a tendency to use more lethal, and therefore more irreversible, means of suicide.

The rate of suicide is particularly high in the United States among Native Americans, who show a suicide rate about twice that of the overall rate, 15.2/100,000 for 15 to 19 year-olds in 2013 (see Fig. 51.1–1). This pattern of high suicidal risk among indigenous peoples is found worldwide and may be related to erosion of traditional cultural values, increased alcoholism, legal difficulties, and unemployment. Whites have a higher rate of suicide than African-American youth, although the rate of suicide among African-American boys has climbed over the years to around 75 percent of the rate in white males.

After peaking in the 1990s, the suicide rate in adolescents declined until 2004, when it increased for the first time in a decade. Since 2007, there has been a gradual increase in the suicide rate both in males and females (see Fig. 51.1–2). The decline in the rate during the 1990s has been attributed to

more restrictive gun control laws and to increased prescription of SSRIs through correlative studies, although it is difficult to identify causal factors. The use of SSRIs among adolescents has declined since 2004, which corresponds in time with the FDA's black box warning on antidepressants. Some pharmacoepidemiological studies have shown a relationship between the decline in use of antidepressants and an increase in the rate of suicide and suicide attempts, but none of these studies are definitive.

The most common method for adolescent suicide attempts is overdose, with the second-most-common method being wrist cutting. Although suicide attempts in adolescents are frequently of relatively low intent, ingestion of acetaminophen is associated with a significant risk of liver toxicity and fatality without prompt medical attempt. Much more rarely, adolescents engage in highly lethal suicide attempts, using methods more commonly found in completed suicide. The most common method in the United States for suicide fatality is by firearms, followed in frequency by hanging, jumping, carbon monoxide poisoning, and taking an overdose.

Risk Factors, Assessment, and Etiology

Because completed suicide is so rare and most of the risk factors for completed suicide overlap with those for suicidal ideation and attempt, accurate prediction of suicide completion is not possible. However, the identification of individuals at risk of suicidal behavior is still worthwhile because most of the risk factors are associated with significant functional impairment and therefore require treatment.

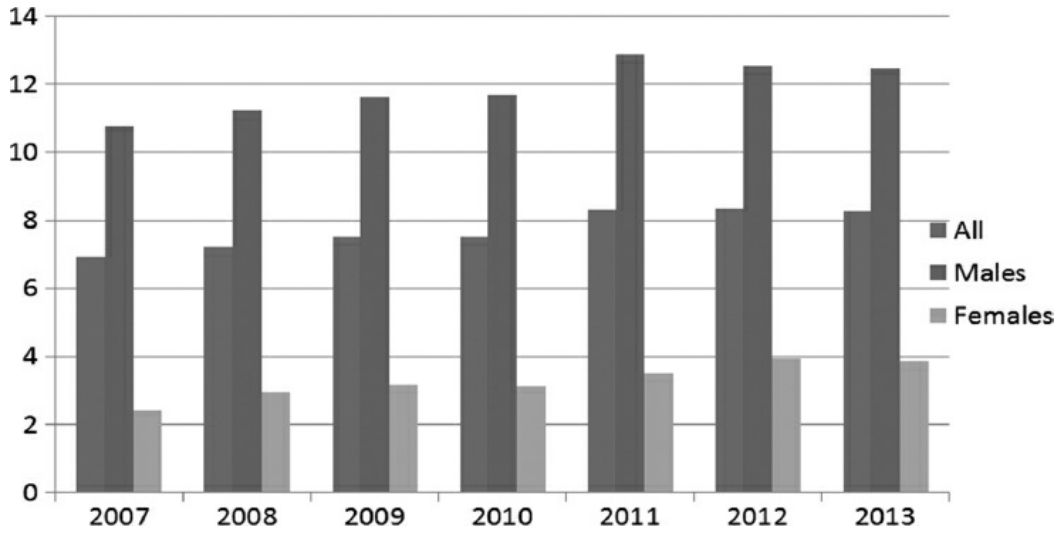

FIGURE 51.1–2. US rates of suicide (per 100,000) in youth 15 to 19: overall by sex. (Accessed Injury Prevention & Control: Data & Statistics (WISQARS) http://www.cdc.gov/injury/wisqars/index.html. Updated July 13, 2015.)

The best predictor of future suicidal behavior is current and recent past suicidal behavior. The more serious the previous episode of suicidal

ideation (e.g., with a plan) or behavior (high suicidal intent), the greater is the likelihood of a suicide attempt. Suicidal behavior and ideation are assessed across five domains: intensity and intent, lethality, precipitant, motivation, and availability of lethal agents.

Intensity and Intent. Suicide ideation is assessed both with respect to *intensity* (i.e., how frequent and how compelling) and *severity* (from passive ideation to ideation with intent and a plan). *Suicidal intent* refers to the degree to which the individual with suicidal ideation intends to carry out a suicidal act; when the individual has made an attempt, intent refers to the degree to which the individual wished to die. The worst-ever level of suicidal ideation is also predictive of future attempts and suicide deaths. Suicidal ideation is a strong predictor of future suicidal behavior. In one study, 33.4 percent of those with suicidal ideation developed a plan within a year, and 60.8 percent of those with a suicidal plan made a suicide attempt within a year. The risk of attempt is lower in those with ideation without a plan, about 20 percent in 1 year.

An adolescent who has made a suicide attempt has a 15 to 30 percent chance of a reattempt within 1 year, with the highest risk of reattempt in the first 3 months after the attempt. Suicide attempters are at a 10- to 60-fold increased risk of completed suicide, with those who make attempts of high lethality and intent at highest risk. Suicidal intent is assessed not only based on the suicidal individual's self-report, but also based upon his or her actual suicidal behavior, such as the degree of preparatory behavior and timing of the event to minimize discovery, which often belie the suicidal individual's self-report. Those who die by suicide are also much more likely than attempters to have confided to a friend about their suicidal intent. This is reflective of a more stable and intense ideation in those who die by suicide versus attempters, but it also reflects a simultaneous wish to live and wish to die on the part of even the most suicidal individuals. NSSI has emerged as a potent predictor of suicidal behavior, in some studies showing a stronger association with future suicide attempts than a previous history of a suicide attempt. Finally, adolescents who express regret for having survived a suicide attempt are at increased risk of recurrent and fatal suicidal behavior.

Motivation. Motivation refers to the reason why the individual engaged in suicidal behavior. An understanding of motivation is key to the development of an effective treatment plan. About one-third of adolescents who attempt suicide do so with a wish to die. The remainder of adolescent suicide attempters report that they wanted to escape an intolerable situation, could not tolerate their psychological pain and distress, wanted to express hostility, or to get attention and support.

Precipitant. Parent–child conflict, difficulties in peer and romantic relationships, and disciplinary problems are the most common precipitants for suicidal behavior, and to the extent that these precipitants are likely to recur, the suicidal individual continues to be at increased risk of a reattempt. In younger children and adolescents, precipitants are most commonly parent–child conflict, whereas older adolescents are more likely to report romantic or peer difficulties as precipitants. Part of therapeutic intervention should be to help the patient cope differently, without engaging in suicidal behavior, when faced with a similar precipitant in the future.

Lethality. Lethality refers to the medical dangerousness of the suicide attempt. Intent and lethality are not always highly correlated, especially in children and young adolescents. Therefore, an attempt of low lethality should not be dismissed as a "gesture," because this attempt could still convey high suicidal intent. Conversely, impulsive attempts without clear high intent can sometimes result in highly lethal or even fatal attempts, such as in overdose of acetaminophen, or in younger adolescents without clear psychopathology, for whom the availability of a loaded gun is by far the most significant risk factor for suicide. However, those individuals whose attempts resemble suicide fatalities with regard to method—for example, with the use of hanging or firearms—are much more likely to eventually die by suicide.

Availability of Lethal Agents. In the United States, approximately 60 percent of youth suicides involve firearms, although hanging is the leading cause of death in many other countries and has been increasing as a method of choice in American youth as well. Case–control studies, cross-country comparisons, and quasi-experimental studies consistently show an association between the availability of guns, particularly handguns, and risk of suicide. Guns are two to ten times more likely to be found in the homes of suicide victims than in those of demographically, and even psychiatrically, matched controls. The only risk factor that differentiated between youth who died by suicide and living control youth, all of whom had no clear evidence of psychiatric disorder, is the availability of a loaded gun in the homes of the suicides. Restriction of access to guns has been associated with regional and national declines in firearm-related and overall suicide rates. Restricting the availability of acetaminophen in the United Kingdom and in France by requiring a prescription and dispensing a limited number wrapped in bubble pack to prevent easy ingestion of large quantities have been shown to favorably affect the rate of overdose and the case fatality per overdose. Other studies have suggested that the addition of barriers to a bridge that is commonly used for suicide might have reduced the suicide rate from that site. However, not all studies of method

restriction show a consistent effect, and in some studies there is evidence for partial or complete method substitution.

Other salient risk factors and correlates of adolescent suicidal behavior include mental disorder, health risk behaviors, chronic medical conditions, psychological characteristics, familial–environmental factors, exposure to suicidal behavior, and biological risk factors.

Mental Disorder. The rate of psychiatric disorder in suicide and attempted suicide is 90 and 80 percent, respectively. Greater chronicity, severity, and complexity (e.g., comorbidity) are all characterizations of psychiatric illness in those individuals with suicidal ideation, attempts, and completions. Because the majority of psychiatrically ill individuals never engage in suicidal behavior or die by suicide, the treatment of psychiatric disorder is likely to be necessary but may not be sufficient to prevent suicidal behavior.

Mood disorders are strongly associated with suicidal ideation, attempts, and suicides. Approximately 60 percent of adolescent suicide victims had a mood disorder at the time of death. Bipolar disorder conveys a particularly high risk of suicidal attempts and completions, especially when the patient is experiencing rapid cycling or a mixed state. About one-third of clinically referred youth with bipolar disorder report having made at least one suicide attempt. Among the symptoms of depression most closely associated with suicidal ideation and behavior are insomnia, agitation, and irritability. Longitudinal follow-up of one study of adolescent-onset depression found a 10-year rate of suicide of 7.7 percent versus no suicides among the controls.

Anxiety disorders—specifically social anxiety, panic disorder, and post-traumatic stress disorder—have all been reported to be associated with suicidal ideation and attempts, although most studies find that anxiety most strongly contributes to suicidal risk when comorbid with mood disorder. Suicide and attempted suicide are also commonly associated with alcohol and substance abuse disorders, especially when these conditions are comorbid with mood disorders. Alcohol abuse differentiates ideators and attempters, possibly by reducing restraint against engaging in suicidal behavior in an individual already contemplating suicide. In one large study, the prototypical older adolescent male who died by suicide was triply comorbid with mood disorder, conduct disorder, and substance abuse. The risk of suicide and suicidal behavior among incarcerated youth is very high.

Other conditions, such as schizophrenia and eating disorders, also have a high lifetime risk of suicide attempt and suicide fatality, but in adolescence is usually associated with suicide attempts, with the increased mortality risk emerging during young adulthood.

Health Risk Behaviors. Suicidal behavior most commonly occurs as

part of a constellation of other health risk behaviors, such as NSSI, binge drinking, abnormal eating behavior (either restricting or bingeing), weapon carrying, and having unprotected sex. Therefore, it is important to assess for the range of health risk behaviors when assessing suicidal risk. Those who engage in self-poisoning not only have an increased risk from dying by suicide but also have a 10-fold increased risk of accidental death as well. Studies have shown that common risk factors for these health risk behaviors include impulsivity and poor problem-solving skills, a weak connection between parent and child and/or a conflictual parent–child relationship, low degree of parental supervision and monitoring, poor connection between child and school, and association with deviant peers.

Medical Conditions. Some specific chronic illnesses show higher rates of depression and suicidal behavior, particularly in illnesses that involve the central nervous system (e.g., migraine and epilepsy), or inflammation (e.g., inflammatory bowel disorders and asthma). There appear to be three components explaining the increased risk of depression and suicidal behavior in chronic medical disorder. First, there is a general effect of chronic illness that results in pain, disability, and disrupts the ability to enjoy oneself and achieve mastery. Second, there are specific pathophysiological aspects of the diseases that increase the risk of depression and suicidal ideation and behavior, as noted above. Third, there may be aspects of treatment that increase the risk of depression and suicidal ideation, such as the use of some antiepileptic drugs, interferon, and steroids.

Psychological Factors. Psychological characteristics that differentiate adolescent suicide attempters from psychiatrically disorder comparison groups include impulsivity, impulsive aggression (the tendency to respond to frustration or provocation with hostility or aggression), and cognitive bias and attraction to suicide.

Most adolescent suicide attempts are impulsive, so it is not surprising that impulsivity is a strong correlate of suicidal behavior. Neurocognitive testing has demonstrated that adolescent suicide attempters make impulsive decisions and show a failure to learn from consequences. In addition, impulsive aggression, as measured by self-report, assessment of Cluster B traits, and neurocognitive testing, is associated with suicidal behavior. Impulsive aggression also predicts the onset of mood disorder, which in turn predicts the familial transmission of suicidal behavior. Disordered emotion regulation, characterized as having a lower threshold for dysregulation, is showing more extreme and a more prolonged response to provocation. Emotion dysregulation has been demonstrated with self-report, neurocognitive, and fMRI studies as a key contributor to suicidal risk.

There also is evidence that suicide attempters are more likely to show a self-identification with death and suicidal behavior that is detectable on an Implicit Association Test, and that this increased self-identification with suicide is a robust predictor of future suicidal behavior.

Same-sex attraction has been consistently shown to be a risk factor for suicidal behavior in youth, in part mediated by family and peer rejection, and an increased rate of past maltreatment and sexual exploitation, which in turn accounts for the increased rates of mood, substance abuse, and suicidal behavior.

Familial and Environmental Factors. The parents of adolescent suicide attempters and completers show much higher rates of depression, substance abuse, and assaultive behavior than do the parents of psychiatric and normal controls. Adoption, twin, and family studies also show that the suicidal behavior is transmitted within families independent of the transmission of psychiatric disorder. For example, monozygotic twins show a higher concordance for suicidal behavior than do dizygotic twins, even after controlling for the concordance of psychiatric disorder. The first-degree relatives of adolescent attempters and completers show a two- to sixfold increase in the risk of suicidal behavior, even after controlling for the increased risk of familial psychopathology. Studies show that the offspring of depressed suicide attempters are at much higher risk of making an attempt than are the offspring of depressed, nonattempting parents. In addition, greater family loading is related to a greater risk and an earlier age of onset of suicidal behavior. The familial transmission of suicidal behavior is mediated by the familial transmission of impulsive aggressive traits.

However, in addition to evidence that suicidal behavior is heritable, there is also substantial evidence for a role for familial–environmental factors. Parent–child discord, perceived lack of support, and high levels of criticism and hostility are associated with child suicide attempt. Conversely, expressing high parent–child warmth, having high levels of parental supervision, and spending time together (e.g., eating dinner together and engaging in leisure time together) are protective against suicidal behavior. Some studies indicate that having a moral objection to suicide is also protective against suicidal behavior.

Parental psychiatric hospitalization and parental loss, the latter especially true in youth who lose their parent prior to the age of 12, are both associated with an increased risk of suicide in offspring. Suicide among either family or peers does not seem to increase the risk of imitation, but a suicide attempt in a friend is prospectively associated with an increased risk of an adolescent attempt. Although it is difficult to differentiate the effects of assortative friendships from imitation, assortative friendships alone would not explain the *timing* of the relationship between peer

attempts.

The most clinically significant familial–environmental risk factor for suicidal behavior is maltreatment, namely physical and sexual abuse, and emotional and physical neglect, with the strongest associations being reported between sexual abuse and suicidal behavior. The population-attributable risk of sexual abuse for suicidal behavior has been estimated as slightly less than 20 percent, meaning that, in theory, one in five cases of suicide attempt could be eliminated with the elimination of sexual abuse. In reality, the relationship between abuse and suicidal behavior is more complicated because abuse occurs within the social context of socioeconomic strain, family disruption, parental history of abuse, parental depression, substance abuse, suicidal behavior, and low parental monitoring and supervision. The likelihood of suicidal behavior is greater with greater severity of abuse—for example, with vaginal or anal penetration, as compared to molestation, the latter of which is not associated with an increased risk of suicidal behavior.

Youth at risk of suicidal behavior are not well connected to school, work, or family. Youth who are engaged in bullying, are victimized by other peers, or both are at increased risk of suicidal ideation and behavior. Students at risk of dropping out of school, as characterized by poor academic performance and attendance, are also at high suicidal risk. Youth disconnected from all major support systems are at very high risk of completed suicide, although it is likely that other factors such as depression and substance abuse that contribute to suicidal risk also contribute to the disruption of these connections. Conversely, youth who have a strong connection to school, even in the face of other risk factors, experience some protection from suicidal risk. A deviant peer group is associated with increased risk of suicidal behavior, in part because such a group encourages behavior that may lead to precipitants for suicidal behavior, such as a disciplinary or legal crisis. Conversely, youth otherwise at high risk, namely those with a history of abuse, are protected against suicidal risk by having a prosocial peer group.

Exposure to Suicidal Behavior. Approximately 5 to 10 percent of all youth suicides occur in time-space structures, implying at least the possibility that one suicide increases the risk of suicidal behavior in those who are exposed. Studies of the effect of media publicity about suicide have consistently shown that such publicity is followed by an increase in suicide and suicidal behavior, most markedly in the young. The increase is proportional to the amount of publicity, the notoriety of the individual who committed suicide, and the extent to which the report romanticized the suicide as compared to presenting information that describes the relationship between suicide and psychiatric illness. In some instances, specificity of method choice in imitation has also been demonstrated.

Biological Factors. On neurocognitive tasks, adolescent suicide attempters show a bias toward suicidal preoccupation, impaired reward learning, and a tendency toward more risky decision making. Neuroimaging studies of adolescent depressed suicide attempters, compared to depressed nonattempters, and healthy controls show alterations of attention and salience networks that are not evinced during "cold" cognitive tasks, but only emerge with emotional stimuli (e.g., exposure to angry faces). Overall, these findings are consistent with a profile of impulsive aggression, a tendency to risk taking, and poor emotion regulation that are known clinical characteristics of adolescent suicide attempters. Alteration in HPA axis and in levels of inflammatory cytokines may also be related to risk of suicidal behavior. Postmortem studies in adolescents who died by suicide have demonstrated the greatest degree of alteration in the prefrontal cortex and hippocampus, areas associated with known deficits for suicidal individuals, namely emotion regulation and problem solving. These brain studies have shown alteration in 5-HT_{2A} binding, decreased activity of protein kinase A and C, downregulation in cAMP response element-binding (CREB) protein, and decreased activity of brain-derived neurotrophic factor (BDNF), the latter two of which are involved in signaling and protection from the neurotoxic effects of stress.

Course and Prognosis

Suicidal behavior is highly recurrent. Around 15 to 30 percent of adolescents reattempt suicide with 1 year of the initial attempt, and the risk of a completed suicide is around 0.5 to 1 percent per year. Factors that discriminate between those who die by suicide and those hospitalized for suicidal ideation or an attempt include male sex, availability of a gun, mood disorder comorbid with substance abuse or conduct disorder, diagnosis of bipolar disorder, and high suicidal intent. Among those adolescents who engage in self-poisoning, prospective studies show that the hazard of death by suicide and accidents is elevated more than 40-fold, with a mean time a suicidal death was 1.6 years after presentation in the emergency department. Predictors of suicide in this sample of youth who took overdoses included male sex, older age, history of multiple overdoses, depression, and previous psychiatric care.

Assessment of Imminent Suicidal Risk

Often, the clinician is called upon to assess a patient for imminent suicidal risk, meaning whether the person is likely to engage in suicidal behavior in the next 24 to 48 hours. Suicide occurs because of an imbalance between distress and restraint. Factors that increase distress and that decrease restraint increase imminent suicidal risk. High distress is found in patients with very high anxiety, hopelessness a mixed bipolar state, agitation,

psychosis, or emotional dysregulation. Insomnia is a particularly lethal contributor, because it tends to both increase distress and decrease restraint. Factors that decrease restraint include availability of a lethal means of suicide, impulsivity, alcohol and substance abuse, or feeling that people will not care or will be relieved by the person's suicide. Acute crisis intervention to decrease suicidal risk aims to decrease distress and augment restraint.

Treatment

Although the treatment of depression is an important component of reducing suicidal risk, at least in depressed patients, adolescent depressed patients who are also suicidal probably require additional targeted interventions. Suicidal ideation is highly correlated with the severity of depression, and clinical trials of child and adolescent depression have shown a concomitant decrease in suicidal ideation along with reduction in depression. In comparing clinical trials of fluoxetine in those under and over the age of 24, those depressed individuals over the age of 24 showed improvement in both depression and suicidal ideation, whereas in those under the age of 24, depression improved just as much as it did in older subjects, but suicidal ideation did not improve to the same degree. This may be due to the fact that, although depression is salient to suicidal behavior across the life span, it accounts for less of the variance in suicidal ideation and behavior in younger people compared to mid-life and elderly adults. There is some evidence that use of antidepressants is associated with a slight increase in the risk of suicidal events (defined as either new onset or increase in ideation, or an attempt). These events are almost all increases in suicidal ideation, with relatively few attempts and no suicide completions, so that the clinical significance of these events is unclear. Predictors of the development of these events include high levels of baseline suicidal ideation and irritability. In one meta-analysis of the benefits and risks of antidepressants in child and adolescent depression found an increased rate of spontaneous events of 7 for every 1,000 patients treated and that 11 times more individuals benefited from the medication than developed a suicidal event, which, in the view of many but not all clinicians and families constitutes an acceptable risk–benefit ratio.

In contrast, pharmacoepidemiological studies consistently show that increases in sales or prescriptions of antidepressants are related to a lower suicide rate both in cross-sectional and longitudinal studies. The apparent discrepancy between the higher rate of suicidal events in clinical trials and the protective effect of antidepressants in pharmacoepidemiological trials may be attributed to the following differences between randomized clinical trials and pharmacoepidemiological studies: outcome (suicidal events in trials vs. suicide in pharmacoepidemiological studies), patient selection (those with high suicidal risk excluded from clinical trials), sample size

(much larger in pharmacoepidemiological studies), and time frame (much shorter in clinical trials). Therefore, pharmacoepidemiological studies may be more representative of the "real world," and much better able to detect effects on suicide and suicidal behavior, even though clinical trials are conducted with more rigorous experimental control, albeit in high selected patient samples studied over relative brief periods of time.

There are no definitive psychosocial or pharmacological treatments for suicidal behavior. In clinical trials of the treatment for depression, suicidal ideation declines with the treatment of depression, and in some, but not all studies, the combination of antidepressant and CBT produces the most rapid decline in suicidal ideation. While there are no definitive studies for the treatment of suicidal behavior, there are four promising treatment approaches. Attachment-based family therapy, which increases supportive interaction between parent and adolescent, has been shown to reduce depression and suicidal ideation in one clinical trial. An integrated CBT with a strong substance abuse and family treatment component was shown to reduce suicide attempts more so than treatment as usual. Mentalization, which teaches youth to conceptualize actions in the forms of thoughts and feelings, has been shown to reduce self-harm (which includes both suicide attempts and NSSI) more so than treatment as usual. A program to promote more positive parent–child interactions, when added to treatment as usual, has been shown to reduce suicidal risk related to usual care alone. Finally, an adaptation of Dialectic Behavior Therapy (DBT), which teaches emotion regulation and interpersonal effectiveness skills, was shown to reduce self-harm behaviors to a much greater extent than the usual treatment. All of these treatment approaches await replication with larger samples by groups not associated with the development of the treatment before they can be considered "evidence-based." There have been no studies of pharmacological treatments for youth with suicidal behavior, although meta-analyses in adults suggest that lithium may be efficacious in preventing suicide in individuals with bipolar, and even unipolar, depression.

Clinical guidelines for the management of suicidal youth support treating the related psychiatric disorder and addressing psychosocial risk factors for suicidal behavior, such as family discord, emotional dysregulation, and dysfunctional problem solving. Best practice recommends the development of a safety plan, that is, *a plan shared by the family, patient, and clinician that can be utilized to avoid precipitants for recurrent suicidal behavior and to cope with suicidal ideation if it reoccurs*. The development of a safety plan involves a review of the sequence of emotions, behaviors, thoughts, and events leading up to the attempt and identification of the approaches that are simplest and most likely to be successful for coping with ideation and suicidal urges in the short term, such as relaxation, avoidance (leaving the room if an argument

appears to be imminent), seeking support at home or from a friend, and calling the clinician or his or her designatee. Written and verbal no-suicide contracts have been shown not to reduce the risk of subsequent suicidal behavior and are not recommended. Current guidelines recommend that the restrictiveness of the care of the suicidal individual be proportional to the assessed suicidal risk, including high suicidal ideation with inability to formulate and/or adhere to a safety plan. However, there has never been an empirical evaluation of the therapeutic benefit of psychiatric hospitalization. Post-discharge from psychiatric hospital is a very high risk period for suicide and suicidal behavior, and some evidence suggests that better post-discharge assertive follow-up, availability of 24-hour crisis services, multidisciplinary reviews of patient suicides, and availability of dual-diagnosis services all can contribute to reductions in regional post-discharge patient suicides.

There have been relatively few evaluations of programs intended to prevent suicidal behavior in adolescents. The Signs of Suicide (SOS), which combines screening and identification of at-risk students with education about depression and suicide, was shown, in a randomized clinical trial, result in greater knowledge and more adaptive attitudes toward depression and suicide, and a lower rate of attempts among youth exposed to the program. In a very large, multicountry cluster-randomized clinical trial based conducted in Europe, a program similar to SOS was compared to gatekeeper training, screening and referral, and a psychoeducational control intervention. Only the youth mental health awareness intervention showed a decrement in both severe suicidal ideation and suicide attempts. One universal preventive intervention delivered to first-graders in poor, urban neighborhoods, "The Good Behavior Game," has been shown in one sample, to be associated with a reduction in suicidal ideation and behavior two decades later.

In light of the paucity of empirical studies, it is clear that treatment studies of suicidal individuals are greatly needed, and in addition, it is important to include suicidal individuals in clinical trials that focus on depression. Current guidelines recommend that care for the suicidal adolescent include an assessment of the proper level of care to safely deliver the treatment, the development and continual updating of a safety plan, securing or removing lethal agents from the home, treating the underlying psychiatric disorders, and addressing the psychosocial issues, such as family discord and poor emotional regulation skills that have been found to be salient to suicidal risk.

REFERENCES

*Birmaher B, Brent D, Bernet W, et al; AACAP Work Group on Quality Issues. Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. *J Am Acad Child Adolesc Psychiatry*. 2007;46:1503–1526.

*Brent D, Emslie G, Clarke G, et al. Switching to another SSRI or to venlafaxine with or without cognitive behavioral therapy for adolescents with SSRI-resistant depression: the TORDIA randomized controlled trial. *JAMA*. 2008;299(8):901–913.

Brent D, Poling K, Goldstein T. *The Treatment of Depressed and Suicidal Adolescents: A Clinician's Guide*. New York: Guilford Press; 2011.

Bridge JA, Goldstein TR, Brent DA. Adolescent suicide and suicidal behavior. *J Child Psychol Psychiatry*. 2006;47(3–4):372–394.

*Bridge JA, Iyengar S, Salary CB, et al. Clinical response and risk for reported suicidal ideation and suicide attempts in pediatric antidepressant treatment: a meta-analysis of randomized controlled trials. *JAMA*. 2007;297(15):1683–1696.

Bridge JA, McBee-Strayer SM, Cannon EA, et al. Impaired decision making in adolescent suicide attempters. *J Am Acad Child Adolesc Psychiatry*. 2012;51(4):394–403.

Copeland WE, Angold A, Costello EJ, Egger H. Prevalence, comorbidity, and correlates of DSM-5 proposed disruptive mood dysregulation disorder. *Am J Psychiatry*. 2013;170(2):173–179.

Copeland WE, Shanahan L, Egger H, Angold A, Costello EJ. Adult diagnostic and functional outcomes of DSM-5 disruptive mood dysregulation disorder. *Am J Psychiatry*. 2014;171(6):668–674.

*Finkelstein Y, Macdonald EM, Hollands S, et al; Canadian Drug Safety and Effectiveness Research Network (CDSERN). Risk of suicide following deliberate self-poisoning. *JAMA Psychiatry*. 2015;72(6):570–575.

Gibbons RD, Brown CH, Hur K, Davis J, Mann JJ. Suicidal thoughts and behavior with antidepressant treatment: reanalysis of the randomized placebo-controlled studies of fluoxetine and venlafaxine. *Arch Gen Psychiatry*. 2012;69(6):580–587.

Hughes CW, Emslie GJ, Crismon ML, et al; Texas Consensus Conference Panel on Medication Treatment of Childhood Major Depressive Disorder. Texas Children's Medication Algorithm Project. Update from Texas Consensus Conference Panel on Medication Treatment of Childhood Major Depressive Disorder. *J Am Acad Child Adolesc Psychiatry*. 2007;46(6):667–686.

Kennard BD, Emslie GJ, Mayes TL, Hughes JL. Relapse and recurrence in pediatric depression. *Child Adolesc Psychiatr Clin N Am*. 2006;15(4):1057–1079, xi.

Klomek AB, Mufson L. Interpersonal psychotherapy for depressed adolescents. *Child Adolesc Psychiatr Clin N Am*. 2006;15(4):959–975, ix.

Lu CY, Zhang F, Lakoma MD, et al. Changes in antidepressant use by young people and suicidal behavior after FDA warnings and media coverage: quasi-experimental study. *BMJ*. 2014;348:g3596.

March JS, Silva S, Petrycki S, et al. The Treatment for Adolescents With Depression Study (TADS). Long-term effectiveness and safety outcomes. *Arch Gen Psychiatry*. 2007;64(10):1132–1143.

Marshal MP, Dietz LJ, Friedman MS, et al. Suicidality and depression disparities between sexual minority and heterosexual youth: a meta-analytic review. *J Adolesc Health*. 2011;49(2):115–123.

Melvin GA, Dudley AL, Gordon MS, Ford S, Taffe J, Tonge BJ. What happens to depressed adolescents? A follow-up study into early adulthood. *J Affect Disord*. 2013;151:298–305.

Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MF. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. *BMJ*. 2012;344:e2598.

Nock MK, Banaji MR. Prediction of suicide ideation and attempts among adolescents using a brief performance-based test. *J Consult Clin Psychol*. 2007;75(5):707–715.

Nock MK, Green JG, Hwang I, et al. Prevalence, correlates, and treatment of lifetime suicidal behavior among adolescents: results from the national comorbidity survey replication adolescent supplement. *JAMA Psychiatry*. 2013;70(3):300–310.

Ougrin D, Tranah T, Stahl D, et al. Therapeutic interventions for suicide attempts and self-harm in adolescents: systematic review and meta-analysis. *J Am Acad Child Adol Psychiatry*. 2015;54(2):77–102.

Roy AK, Lopes V, Klein RG. Disruptive mood dysregulation disorder: a new diagnostic approach to chronic irritability in youth. *Am J Psychiatry*. 2014;171:918–924.

Thapar A, Rice F. Twin studies in pediatric depression. *Child Adolesc Psychiatr Clin N Am*. 2006;15:869–881, viii.

Vasa RA, Carlino AR, Pine DS. Pharmacotherapy of depressed children and adolescents: current issues and potential directions. *Biol Psychiatry*. 2006;59(11):1021–1028.

Wagner KD. Management of treatment-resistant childhood disorders. In: Nemeroff CB, ed. *Management of Treatment-Resistant Major Psychiatric Disorders*. New York: Oxford University Press; 2012:319–344.

Wagner KD, Pliszka SR. Treatment of childhood and adolescent disorders. In: Schatzberg AF, Nemeroff CB, eds. *Essentials of Clinical Psychopharmacology*. 3rd ed. Arlington, VA: American Psychiatric Publishing; 2013:693–750.

*Wasserman D, Hoven CW, Wasserman C, et al. School-based suicide prevention programmes. The SEYLE cluster-randomised, controlled trial. *Lancet*. 2015;385(9977):1536–1544.

Wilcox HC, Kellam SG, Brown CH, et al. The impact of two universal randomized first- and second-grade classroom interventions on young adult suicide ideation and attempts. *Drug Alcohol Depend*. 2008;95 Suppl 1:S60–S73.

Youngstrom E, Meyers O, Demeter C, et al. Comparing diagnostic checklists for pediatric bipolar disorder in academic and community mental health settings. *Bipolar Disord*. 2005;7:507–517.

Zalsman G, Oquendo MA, Greenhill L, et al. Neurobiology of depression in children and adolescents. *Child Adolesc Psychiatr Clin N Am*. 2006;15:843–868, vii–viii.

Zisook S, Lesser I, Stewart JW, et al. Effect of age at onset on the course of major depressive disorder. *Am J Psychiatry*. 2007;164:1539–1546.

▲ 51.2 Early-Onset Bipolar Disorder

GABRIELLE A. CARLSON, M.D., STEPHANIE E. MEYER, PH.D., AND CAROLY SUE PATAKI, M.D.

INTRODUCTION

The frequency with which early-onset bipolar I disorder (BP) occurs in children is controversial compared to onsets after puberty which more

closely resemble classic adult BP. It is especially challenging to identify a manic episode in a prepubertal child, partly due to the frequent absence of distinct mood episodes, as well as the difficulty distinguishing it from severe attention-deficit/hyperactivity disorder (ADHD), which is may be comorbid with mania in children. Relying purely on symptoms of mania rather than discrete episodes has caused a significant increase in the diagnosis of bipolar spectrum disorders in children. There have been concerns over the last 15 to 20 years about whether children who appear to meet symptom criteria for mania grow up to be adults with bipolar disorder.

The confusion between severe ADHD and mania evolved from a series of diagnostic criteria changes in earlier versions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM), in which *emotion dysregulation* was removed as a core feature of ADHD, and the demarcation of an *episode* of mania became less well defined. DSM-5 now mandates that in mania, both "an abnormal and persistently elevated, expansive or irritable mood" *and* "persistently increased goal-directed activity or energy" are present. DSM-5 further clarifies an episode by specifying that changes in mood and behavior must vary to a *significant degree* and represent an *observable change* from usual behavior. In an effort to provide a diagnosis for youth who are chronically (vs. episodically) irritable, explosive, and equally disabled, but without other symptoms of mania, a new diagnosis "disruptive mood dysregulation disorder" was added to DSM-5 to identify children whose mood is ". . . persistently irritable or angry most of the day, nearly every day, and is observable to others" (DSM-5).

Definition

The episodic nature of mania and depression has been the classic hallmark of bipolar disorder. The DSM-5 defines an episode as a "distinct period" different from the person's "usual self" during which specific symptoms co-occur. For mania, these include an "abnormally and persistently elevated, expansive, or irritable mood" *and* "persistently increased goal-directed activity or energy," accompanied by at least three additional symptoms (four if mood is irritable), including inflated self-esteem or grandiosity, decreased need for sleep, pressured speech, flight of ideas, distractibility, increased activity in work or socially, or psychomotor agitation, and excessive involvement in activities with a high potential for painful consequences. To meet full criteria for mania, symptoms must last at least 1 week, or 4 days for hypomania. The mood disturbance is "sufficiently severe to cause marked social or occupational functioning" or hospitalization and represents a change from prior functioning.

Episodes of depression, which must last a minimum of 2 weeks, are characterized in children by a sad, empty, or irritable mood along with

other symptoms of feeling worthless, preoccupation with death or suicide, agitation/feeling slowed down, and changes in appetite, sleep, and energy level. The diagnosis of dysthymic disorder, now called persistent depressive disorder, which is required to last 2 years in adults, has only a 1-year requirement in children.

In DSM-5, the coexistence of "three manic symptoms" within a depressive episode is enough to meet criteria for the specifier "with mixed features." Likewise, within a full manic episode, the presence of three depressive symptoms also merits the specifier "with mixed features." In fact, when a full manic episode and a full major depressive episode coexist, according to DSM-5, a diagnosis of manic episode with the specifier "with mixed features" should be made, due to the severity and impairment associated with mania.

HISTORY

Recognition of manic depression in child and adolescent psychiatry began with the dissemination of Kraepelin's monograph *Manic Depressive Insanity and Paranoia* after 1921. Child psychiatrists searched for the same condition in their patients but found mostly adolescent depression, which was highly familial. Mania was very rare, even in teens. Brief attempts to broaden the concept of manic depression for younger children were discouraged by a literature review conducted by Anthony and Scott in 1960 that used very stringent criteria to identify manic depression in youth. They found only 3 of 60 cases reported prior to puberty. However, the tension between restricting the diagnosis to classical presentations versus altering the criteria to fit children continues to the present.

In the 1960s, evidence of lithium's efficacy for treating adults with manic depression led to interest in identifying lithium-responsive clinical profiles in children. Early studies suggested that lithium response was most promising in early-onset cases resembling classic manic depression, whereas there was no support for the use of lithium in treating children with broader behavioral dysregulation. When standard symptom criteria became available in the 1970s for describing mania more precisely in adults, Weinberg and Brumback, two pediatric neurologists, adapted these criteria for use with children. The concept of a "manic-depressive variant syndrome of childhood" was subsequently introduced in 1979 and characterized children between 6 and 16 years of age who were hyperreactive to seemingly minor events and were also hyperactive, had disruptions in interpersonal relationships, and often had a family history of manic depression.

By the 1980s, there was increasing acceptance of the diagnosis of manic depression, or bipolar disorder as it became known to the field, in youth. Research efforts initially focused on describing the phenomenology of pediatric bipolar disorder and established standard assessment techniques.

In more recent years, there has been a surge of genetic, high-risk, and treatment studies, as well as investigations of neurobiological, behavioral, and physiological markers. However, decades later, there continues to be disagreement regarding the definition and application of specific criteria for bipolar disorder in children, thus significantly slowing these efforts and complicating the interpretation of findings.

Epidemiology

Prevalence rates of bipolar disorder vary depending on the whether bipolar I, II, or spectrum conditions are included, the age studied, the definitions and methods used to make the diagnosis, and the type of sample interviewed. Furthermore, rates of BP can vary depending on whether shared symptoms of ADHD and BP are counted toward each condition, or if shared symptoms must have exacerbated in order to be counted toward a manic episode. Rates of reported lifetime mania in youth have varied from 0.1 percent by Stringaris and colleagues, to 1.7 percent in adolescents, and 2.5 percent for bipolar I and II lifetime disorders among adolescents, by Merikangas and colleagues. In a meta-analysis of studies worldwide, Van Meter and colleagues reported an overall rate of 1.2 percent for bipolar I disorder. The sample most reliably addressed adolescents as the information on children was scanty. Prevalence rates of subthreshold symptoms of bipolar disorder range from 5.7 percent among adolescents in the Oregon Adolescent Depression Study to at least 10 percent in the mid-Missouri sample, and between 2.4 and 6 percent in adult samples. Bipolar spectrum disorder rates from the recent VanMeter meta-analysis were as high as percent. While tightly defined bipolar spectrum disorder appears to have an up to 40 percent rate of conversion to BP I or II, manic symptoms alone are less often associated with such a conversion as seen in both the Longitudinal Assessment of Manic Symptom Study (LAMS), where mania rarely developed, and the Oregon Adolescent Depression Study, where it never developed.

The diagnosis (vs. the actual condition) of bipolar disorder among youth in the community has increased dramatically over the last decade. A study of outpatient physician visits in the United States found that juvenile-onset cases of bipolar disorder comprised 42 percent of outpatient mental health-related visits in 2003, an increase from 6.67 percent of cases in the early 1990s. Another study, using data from the National Hospital Discharge Survey, revealed increases in the rate of bipolar disorder diagnoses among psychiatrically hospitalized children from 10 percent in 1996 to 34.11 percent in 2004, and from 10.24 percent to 25.86 percent among psychiatrically hospitalized adolescents during that same period. The increased rate of bipolar disorder appeared to reflect a shift in diagnostic practices such that childhood mood lability and aggression, which previously might have been attributed to conduct problems, ADHD or

impaired family relationships, has increasingly being considered by some clinicians as indicative of bipolar disorder. The psychopharmacology clinic at Massachusetts General Hospital (MGH), however, using the same assessments for the last 15 years, reported a stable rate of youth bipolar disorder in their clinic of 16 percent, over this time period. Although one cannot generalize from one clinic to the next, the consistency of the MGH rate over time suggests that the actual number of children with bipolar disorder has probably not increased.

ETIOLOGY, MECHANISMS, AND RISK FACTORS

Estimates from adult twin studies suggest that genetic influences explain approximately 60 to 93 percent of the variance in bipolar disorder, whereas shared and unique environmental factors account for 30 to 40 percent and 10 to 21 percent, respectively. There are no "narrow phenotype" twin studies of early-onset bipolar disorder. However, two large-scale twin studies using the *Child Behavior Checklist-Juvenile Bipolar Disorder/Dysregulation Profile* (CBCL-JBD) as a proxy measure of broad phenotype BP suggested that genetic influences account for considerable variance in scores, although shared environmental factors may play a significant role in the specific presentation course of this phenotype. Other data from these studies also suggested that most children evaluated with the CBCL-JBD do not have traditionally defined BP I or II. Environmental risk factors including sexual, physical abuse, or neglect, harsh parenting, poor social support, and prenatal exposure to substances may interact with genetic factors resulting in earlier onset of BP disorder. In turn, earlier onset of first BP episode and exposure to abuse have been found in the Course and Outcome of Bipolar Youth (COBY) study follow-up to be prevalent in the group with the worst outcome.

Elevated rates of bipolar disorder have been found among relatives of children with narrow phenotype bipolar disorder as compared to adult-onset cases. However, there does not appear to be an increased risk of narrow phenotype bipolar disorder within families of children with broad phenotype bipolar disorder/severe mood dysregulation (SMD), suggesting that broad and narrow BP may be etiologically distinct. Moreover, symptoms of affective dysregulation and attention problems appear to segregate together in certain families. Whether this symptom profile represents a particularly virulent form of early-onset bipolar disorder or a marker of severe but nonspecific psychopathology has yet to be determined.

Rates of depression, anxiety, and behavioral disorders (especially ADHD) are consistently elevated in offspring of parents with BP and are considerably higher than rates of bipolar I disorder. Rates of bipolar disorder I among high-risk youth, however (4 to 6 percent in most studies), are significantly higher than in offspring of well parents, for whom rates of

BP are negligible. Bipolar spectrum conditions occur in up to 38 percent of offspring of bipolar parents. Rates of childhood BP disorders depend on how broadly the bipolar spectrum is defined, and also the clinical setting in which parents and children are evaluated. Parents in tertiary care clinics appear to have offspring with more psychopathology than parents in community samples, suggesting gene–environment interactions.

BRAIN STRUCTURE AND FUNCTION

Youth with bipolar disorder exhibit impairments in global social functioning, and display particular difficulties with the identification of emotional faces. Whitney and colleagues have documented that even youth at risk of BP, such as offspring of bipolar parents, demonstrate impaired social reciprocity. Significant brain networks that process facial expressions include the amygdala and the prefrontal cortex. Garrett and colleagues have shown that a combination of deficient prefrontal cortex regulation and heightened amygdala response to emotional stimuli occurs in youth with bipolar disorder.

Early-onset BP may reflect earlier evidence of neurocognitive deficits, but with a potential for improvement with treatment, however, this remains to be documented.

Similar to studies of adults, results from both structural and functional magnetic resonance imaging (MRI) studies indicate that early-onset bipolar disorder is associated with abnormalities in cortical and subcortical brain regions involved in the regulation of emotion. Two replicated morphological findings are the presence of white matter hyperintensities and decreased amygdala volume. In functional MRI (fMRI) studies, it appears that heightened activation of the amygdala during emotional challenges may interfere with attentional capacities in children with BP, increasing the probability that they will misinterpret the incoming stimuli (facial expression) and respond in an inappropriate manner.

Findings from studies using magnetic resonance spectroscopy, although preliminary, provide further support for the presence of cortical and subcortical abnormalities in pediatric bipolar disorder. Studies of youth meeting criteria for early-onset bipolar disorder have preliminary evidence of decreased *N*-acetylaspartate (a marker of neuronal density) in the prefrontal cortex and elevated myo-inositol levels (an indicator of decreased neuronal homeostasis) in the anterior cingulate cortex. Moreover, there is evidence to suggest that abnormal findings normalize with the introduction of medication treatment.

Assessment

The American Academy of Child and Adolescent Psychiatry (AACAP) Practice Parameter advocates (1) screening patients for mania and bipolar

depression, both current and past; (2) applying DSM criteria (currently DSM-5); (3) using the designation “bipolar NOS” for children with manic symptoms with durations of less than 4 days (called “other specified bipolar disorder in DSM-5); (4) examining for comorbidities; and (5) using caution when diagnosing preschool children.

Making an accurate diagnosis of bipolar disorder in children requires systematic and thoughtful evaluation procedures that go well beyond the simple application of DSM-5 criteria for mania and depression. Screening measures, completed by multiple informants (i.e., parent, teacher, child), are useful in alerting the clinician to current symptoms of mania, depression, and other possible comorbidities. Interviews with caregiver(s) and child and review of other diagnostic information are required. It is most useful to interview caregivers first because they are likely to provide the most reliable information, and this material can be presented within the child interview to address any discrepant responses. Caregiver interviews should include an overview of current symptoms, developmental history, and longitudinal course.

Although the most pertinent information regarding symptoms of mania and depression may come from the parent interview, the child interview is critical for eliciting important information regarding environmental factors contributing to mood variability, as well as features of other conditions that may be confused, or co-occur, with mania, including autism spectrum disorder, language or thought disorder, psychosis, anxiety, suicidal behavior, physical/sexual abuse, and substance use.

Structured and semistructured diagnostic interviews are used in research settings to elicit symptoms of bipolar disorder and other childhood conditions, and they may also be useful in clinical settings to confirm a diagnosis. The Schedule for Affective Disorders and Schizophrenia in Children (K-SADS) is the most commonly used interview and has a number of versions, including the K-SADS-PL (present and lifetime version) and a version developed at Washington University, the WASH-U-K-SADS. Clinical judgment is required to determine whether a respondent is under- or over-reporting symptoms, misinterpreting developmentally appropriate behaviors, or misunderstanding questions about clinical behaviors. To this end, it is critical that clinicians probe beyond affirmative responses to elicit examples of specific behaviors, as well as the context in which behaviors are occurring.

Several rating scales have been developed to measure the severity of mania. These include an interview measure, the Young Mania Rating Scale (Y-MRS), which is used in both adult and youth medication trials, and parent rating forms for the child, including the Child Mania Rating Scale, the General Behavior Inventory, and Parent Young Mania Rating Scale.

Collateral information, such as teacher input, is helpful in determining whether manic-like symptoms are occurring across settings. Although there

are discrepancies between parent and teacher reports, home and school observation can be useful. Outside assessments, including neuropsychological, occupational therapy, psychoeducational, speech and language assessments, and medical records, can also be extremely helpful in identifying developmental, cognitive, and medical factors that may contribute to emotional dysregulation and complicate episodes of mania and depression. It may also be useful, in certain cases, to obtain a sleep-deprived encephalogram (EEG), a sleep EEG for sleep apnea, a thorough neurological assessment, or laboratory tests. Family history of mood disorder, bipolar and unipolar, and comorbid psychiatric disorders including oppositional defiant disorder, ADHD, learning disorders, and autism spectrum disorders should be assessed as part of the parent interview. If there is a positive family history for bipolar disorder, it is important to probe for details regarding clinical presentation, comorbid conditions, and response to treatment.

Diagnostic Issues Relevant to Children

Defining a mood episode in children, especially younger children, is complicated by the fact that a "usual self," developmentally, is a moving target. Thus, a previously docile infant who enters the "terrible two's" may be developmentally normative, and yet approximate criteria for mania, if this is defined as a persistent mood state different from the child's usual self. Children stressed by school may have observable persistent mood "episodes" which subside over the summer. Geller and colleagues proposed a definitional framework in which the term "episode" describes the length of time that an individual exhibits all criteria for mania/hypomania or depression. The term "cycle," then, refers to shifts in mood that occur throughout each day of the episode, which many consider mood lability occurring within the context of a single episode. Studies that used this framework reported that between 78 and 99 percent of children diagnosed with bipolar disorder experienced daily cycling within an episode duration ranging from 1 to 4 years. Indeed, the boundaries of an episode are blurred both in their onset and offset. In adult studies, in which a typical manic episode lasts for 4 to 8 months, a euthymic period of 8 weeks constitutes remission. Using the above framework, it is unclear how long to construe a remission if most of an individual's childhood has been considered an episode.

Efforts have been made to adapt adult definitions of the cardinal symptoms of mania for young children, taking into account developmental differences. However, children may have difficulties with the concept of elation, or increased self-esteem, which requires an ability to self-assess, a skill absent in young children. Finally, silliness and bravado, which some clinicians construe as childhood manifestations of euphoria and grandiosity, may also be seen in children with ADHD, autism spectrum

disorders, or social skills deficits.

What has emerged in attempts to apply DSM symptom criteria to children are descriptions of young people whose course does not resemble the episodic nature of bipolar disorder in late adolescents or adults. According to McClellan and colleagues, "outbursts of mood lability, irritability, reckless behavior, and aggression," which have come to represent the modal presentation of very early-onset bipolar disorder in the community, may represent a subtype of bipolar disorder or may characterize mood dysregulation that has no particular diagnostic specificity.

Leibenluft and colleagues at the National Institute of Mental Health (NIMH) have been exploring bipolar subtypes in an effort to distinguish between *narrow phenotype bipolar disorder* and a possible broad bipolar phenotype that they termed SMD. The narrow phenotype BP calls for direct application of the former DSM-IV-TR (and the current DSM-5) criteria for BP without modification, including the presence of clear manic/hypomanic episodes characterized by euphoric mood and/or grandiosity. SMD describes children exhibiting chronic irritability with extreme reactivity to negative stimuli and symptoms of hyperarousal. Up to 80 percent of children with SMD met DSM-IV-TR criteria for both ADHD and oppositional defiant disorder. Longitudinal follow-up of youth with chronic, but nonepisodic irritability and mood dysregulation, from community samples and clinical samples suggests that these youth are not destined for bipolar disorders, rather they are at risk of new-onset unipolar depressive disorders and anxiety disorders. The conceptualization of SMD has evolved into the new DSM-5 *Disruptive Mood Dysregulation Disorder* (DMDD); however, in contrast to SMD, DMDD does not include symptoms of hyperarousal, and it cannot coexist with oppositional defiant disorder, intermittent explosive disorder, or bipolar disorder. Other investigators have focused on symptoms rather than episodes. Geller and colleagues define *pediatric and early adolescent bipolar disorder* (PEABD) as requiring two of three cardinal symptoms of mania (euphoria, grandiosity, and/or irritability) but with reinterpretations of euphoria and grandiosity which is described in the Washington University version of the K-SADS. A third early-onset bipolar phenotype, used by Biederman, Wozniak, and colleagues at MGH, characterizes children who manifest severe, explosive irritability and other symptoms of mania, with or without accompanying grandiosity and euphoric mood.

Even where there is agreement regarding definitional parameters, there may be varying ways in which the criteria are applied or interpreted. In a study comparing diagnostic practices regarding pediatric mania in the United Kingdom and the United States, 73 psychiatrists from England and 85 psychiatrists in the United States were enlisted to formulate provisional diagnoses in five case vignettes, each describing a child with "possible

bipolar disorder." Only one case was considered to be a "classic" case of mania in a young adolescent, whereas four were chosen to capture the complexities of diagnosing bipolar disorder in children. In the first instance, a young woman developed a depression at age 10 years after the loss of her grandmother, which remitted with treatment. She later developed an acute, severe mania, which was clearly different from her otherwise unremarkable personality. In the second case, an 11-year-old girl with a long history of ADHD symptoms gradually became increasingly defiant and irritable, disregarding schoolwork, believing it was unnecessary, and became interested in pornographic Web sites. Her parents, however, not her teachers, complained of the more manic-like features. In the first case, there was good agreement on a diagnosis of mania. In the second, US clinicians diagnosed mania and ADHD; UK clinicians diagnosed ADHD and behavior disorder. In other vignettes as well, US psychiatrists were significantly more likely to diagnose mania in children than were UK psychiatrists, whereas UK psychiatrists were more likely to view these children as exhibiting attention-deficit disorder, autism spectrum disorders, and adjustment disorders. The authors concluded that "results illustrate the current confusion that exists surrounding the diagnosis of prepubertal mania, where similar symptoms may be interpreted to have different diagnostic meanings."

Comorbidities

Children and adolescents meeting criteria for BP nearly always present with co-occurring conditions. Symptoms of mania may overlap with those of concurrent ADHD and oppositional defiant/conduct disorder, depression, and anxiety, precisely the conditions with which co-occurrence is highest. ADHD is the most common comorbid diagnosis among children with BP, with published rates of 70 to 95 percent among preadolescent youth, and 50 percent among adolescents. The frequency of comorbid diagnoses is influenced by how different interviews manage symptom overlap between mania and the comorbid disorder. If the interview or investigator deems that a given symptom counts toward all disorders (as does the WASH-U-KSADS) or interviews from Biederman and colleagues, rates of comorbid ADHD and ODD will be higher than if symptoms counted toward mania are only those presenting an intensity significantly beyond a usual baseline. Thus, counting any overlapping symptoms toward all diagnoses, the Treatment of Early Mania Study found that 92.8 percent of the youth with BP had comorbid ADHD in their index episode.

The LAMS study identified ADHD, bipolar spectrum disorders, and comorbidity in a multisite sample of 707 children who were rated by their parents as having elevated symptoms of mania. The LAMS study found that 76.2 percent of their sample met criteria for ADHD, 22.9 percent met criteria for bipolar spectrum disorders, 16.5 percent had both disorders,

and 17.5 percent of the sample had neither disorder. The majority of children who met criteria for BP I disorder (69 percent) also had ADHD. The high rate of comorbid ADHD in the entire sample was best explained by independent rates of these disorders in youth, rather than a hypothesis that ADHD is a risk factor for bipolar disorder in youth.

Rates of conduct disorder and oppositional defiant disorder are reported to be as high as 79 percent in youth with BP. The combination of externalizing disorders and bipolar disorder appears to run in families and is associated with significant functional morbidity. Comorbid anxiety disorders have been reported in up to 77 percent of children with narrow phenotype bipolar disorder, with particularly high rates of generalized anxiety disorder, separation anxiety, and specific phobia. There is some evidence that, in a number of cases, anxiety symptoms may intensify during episodes of depression and mania, resolving when mood symptoms are effectively treated. Finally, features of autism spectrum disorders appear to be quite common among youth with narrow phenotype bipolar disorder, with one study reporting that approximately 20 percent of their sample met full criteria for a comorbid autism spectrum condition.

As with adults, substance and alcohol abuse are common comorbidities in adolescents with BP. ADHD and conduct disorder are both risk factors for the development of substance use problems. In an epidemiological sample, young bipolar patients with substance use problems had conduct disorder rates three times higher (52 percent) than those without substance abuse (14.8 percent; $p < 0.01$). By comparison, young subjects without either mania or substance problems had conduct disorder rates of only 7.75 percent. Substance abuse in bipolar adults may be more related to childhood conduct disorder than uncomplicated bipolar disorder. Nevertheless, substance abuse and bipolar disorder clearly exacerbate each other, and youth with bipolar disorder are at high risk of substance abuse.

For externalizing disorders comorbid with BP there appears to be a developmental shift over time. Rates are highest in children (greater than 75 percent), lower in adolescence (approximately 50 percent, especially if both depression and mania have onset after puberty), and lowest, but still present, in adults, in whom rate of lifetime ADHD has been reported to be approximately 20 percent. Similarly, conduct disorder/antisocial personality co-occur in adults at double the rates in those with bipolar onset before age 30 years (33 percent) versus after age 30 years (16 percent).

Differential Diagnosis

The five most common clinical conditions to be distinguished from bipolar disorder are ones with which it is most often comorbid. Severe ADHD and oppositional defiant disorder in young children are the ones most often confused with mania. The reason for emphasis on episodes, euphoria, and

grandiosity in making the diagnosis of mania emerges because ADHD is, by contrast, a chronic (rather than episodic) disorder for which elation and grandiosity are not considered symptoms. The clearer these symptoms and behaviors are, the easier the differential diagnosis. If severe irritability and explosiveness become the hallmark of early-onset mania, the distinction from severe ADHD with oppositional defiant disorder is extremely difficult. The new "disruptive mood dysregulation disorder" which identifies youth with temper outbursts and chronic, severe irritability, cannot be diagnosed in the presence of manic symptoms lasting more than a day. This exclusion was included in order to help separate children with non-episodic irritability from those with mania. Another difficult diagnostic issue is the development of conduct disorder in older children with ADHD. Children with conduct disorder often are very irritable and oppositional, stay out late in defiance of their parents, and thus, may appear to need less sleep, engage in early sexual behavior and drug use (pleasure-seeking without regard for consequences), and are hedonistic in their self-interest, caring little for the rights of others.

A third difficult diagnostic distinction is between an agitated, anxious depression and a mixed episode of mania. This complicated differential diagnosis is not unique to children and adolescents. In adult psychiatry, the trend has been to err on the side of making a diagnosis of bipolar disorder and avoiding medications to treat the depression and anxiety in order to diminish the potential risk of exacerbating mania. It is necessary to keep an open mind about this issue, and if treatments for mania are not helpful, a diagnosis of depression and anxiety alone should be reconsidered.

Substance abuse not only may render a teen depressed, irritable, emotionally labile, and euphoric, it may also precipitate a psychosis that may be indistinguishable from mania. In the short term, it is often impossible to determine whether the psychopathology is primary or substance induced.

Finally, especially in adolescence, first-episode schizophrenia poses an important differential diagnosis. For many years, schizophrenia was considered to be the only psychosis that occurred in adolescents. It is now understood that depression with psychosis (often a harbinger of bipolar disorder) and mania with psychosis are equally likely to occur. In the clearest cases, teens developing schizophrenia will have an insidious onset to psychosis, long-standing cognitive difficulties, and prominent and bizarre delusions. However, that may not be the case, and only careful follow-up over several years will allow the true diagnosis to emerge.

Jake is a 14-year-old male, 9th grader who was brought to the emergency room due to confusion, agitation, and bizarre behavior. Jake has had behavioral and academic problems since kindergarten. In the first grade, Jake began to have problems paying attention, he was aggressive with his peers, and seemed unable to learn letters and numbers at the same rate as his classmates. An individualized Educational Plan was developed for Jake and he was

provided a resource room in which he received individualized help each day with reading, writing skills, and math. He progressed slowly and by the time he had reached the 5th grade he was only scoring at <25 percentile on academic achievement tests, and had begun to express a dislike for school. During the emergency room interview, Jake disclosed that he had been sexually abused from ages 7 to 13 years by his mother's boyfriend, and that his mother never knew. By the time Jake entered middle school, he had developed moderate anxiety and depression, intermittently expressing suicidal ideation especially when he was frustrated about his school work and felt rejected by his peers. According to his parents, Jake started to become more irritable and erratic about a month prior to the emergency room visit. Over the course of the month, Jake became excessively religious, reading books all night, with increased activity and talking, praying in school, and threatening his teacher on several occasions. He was belligerent with the police when they were called by his school, and he was brought to the hospital in hand cuffs. When Jake arrived at the emergency room, he was confused, incoherent and posturing, expressed his believe that he was God, and also heard the voice of God, was getting messages from the TV, and exhibited paranoid thoughts that creatures from space would kill him.

During his hospitalization he was agitated, hyperverbal, paranoid, and continued to feel that he had special powers. He was started on olanzapine, and was titrated up to a dose of 20 mg/day. At the end of his month-long hospitalization, he was calmer and less paranoid, was no longer hearing voices and was willing to go home with his parents. His discharge diagnosis was bipolar I disorder, manic with psychotic features. After discharge, Jake returned home, he remained on his medication, and attempted to go back to school. He was placed on home instruction because he could not function. Even with a one-to-one tutor, he could concentrate for only brief periods of time. Jake said he felt emotionally flat, almost numb. It was concluded that he was depressed.

When Jake was evaluated psychiatrically 6 months after his hospital discharge, he was found to be sad, and lonely, socially isolated, introverted, suspicious, and he was experiencing illusions. He thought that he could read minds, had restricted emotions; and expressed depersonalization. He was not having auditory hallucinations, and no formal thought disorder was present. He was being treated with lithium, 1,500 mg per day along with olanzapine. At his 6 months evaluation, he was assessed to be in partial remission. Over the next year, Jake seemed to get back to "normal," and he went on to finish the 10th and 11th grades. By 24 months, after his initial hospitalization, however, Jake had another manic episode. He believed he was a saint from God and he had to stop crime; he felt powerful, slept less than 3 hours per night, talked a lot, could not concentrate, and had racing thoughts. Jake became hostile, irritable, and was suspicious of his mother, believing that she might be trying to poison him. Jake's olanzapine was increased to 30 mg/day, and his lithium increased to 1,800 mg/day. Over the next 2 months his mania resolved, but he became even less social, had trouble concentrating, and spent much of his time alone. He reported that he had problems with his memory.

By 48 months, after his initial hospitalization, Jake remained on lithium but experienced another manic episode with referential delusions and paranoia about his mother. After this episode of mania, Jake, now 18 years of age, was demoralized and no longer in school. He had managed to get his GED in between his manic episodes but he spent all day at home. He had very limited psychosocial functioning (watched TV and went shopping with his mother for food), and was not working. Jake's mood was flat, his affect blunted, he had concise speech but mild poverty of content. He was well groomed and had, mild deficits in face and body expression; marked poverty of speech, and long latency of response.

Jake's sister reported that he took medications reliably but noted that he had become more aggressive with each episode, and she was worried that he would never be able to hold a job or live on his own.

Course

There have been four areas of investigation into the outcome of bipolar disorder: (1) the duration and relapse rate of manic and depressive

episodes, (2) percent time symptomatic over longitudinal follow-up, (3) functional outcome, and (4) diagnostic stability (i.e., how likely is a youth to retain the diagnosis of bipolar disorder).

The median duration of an episode of mania in adults is 3 to 6 months. Mixed episodes, which last longer, have durations of about a year. In young people, where episodes are often "mixed," the duration is measured in years. However, in comparing episodes reported in research studies, it is important to pay attention to whether studies describe the index episode as mania, depression or do not specify the polarity of the episode. Depressive episodes are almost always longer than manic episodes.

Duration of episode (or time to remission), when remission is defined as 8 weeks of either no symptoms or subthreshold symptoms, varies widely in young people. In Geller's sample of children with a mean age at intake of 11 years, the index episode lasted 3.6 (±2.5) years.

The Course and Outcome of Bipolar Illness in Youth study (Birmaher et al., 2008, 2014) in which bipolar youth ages 7 to 17 years were combined from 3 different sites reported a median episode duration of 78 weeks. Purely adolescent-onset samples fare somewhat better. The median duration of index episode from one hospitalized sample (mean age 16 years) was 11 weeks; the mean in another was 29 ± 31 weeks. In older adolescents (mean age 17 years), median episode duration was 30 weeks. Outpatient teens with mania (average age 14 years) from India had a median episode duration of only 8.7 weeks. The reasons for these differences in duration are not clear.

Among those children and adolescents who achieve remission, rates of relapse range from 27 percent among community-based adolescents at 2-year follow-up to 70 percent among child and adolescent outpatients at 4-year follow-up.

General risk factors for longer episode duration include nonadherence with pharmacological treatment, rapid cycling, psychotic symptoms, earlier age of onset, low socioeconomic status, comorbid anxiety, ADHD, and/or disruptive behavior disorders. Alcohol use disorders and exposure to antidepressants have been associated with increased risk of recurrence; involvement in psychotherapy appears to have a protective effect. Notably, males diagnosed with pediatric bipolar disorder appear to have a less severe course with greater chance of recovery from the initial episode than females. Moreover, compared to children with bipolar I or II disorder, those with *unspecified* bipolar disorder endured a more chronic course of illness with persistent subthreshold symptoms and delayed treatment response.

Of 367 outpatient youths participating in the COBY study who were interviewed approximately every 6 months over an average period of 8 years, there were four distinct outcome groups based on a latent class growth analysis of percent time euthymic who were (I) "predominantly

euthymic" (24 percent of sample), (II) "moderately euthymic" (34.6 percent/ of sample), (III) "ill with improving course" (19.1 percent of sample), and (IV) "predominantly ill" (22.3 percent of the sample). Group I was found to be euthymic, on average for 84 percent of the 8 years; group II was euthymic for 47.3 percent of the 8 years study period; group III was euthymic for 42.8 percent of the 8 year period, and group IV was euthymic for 11.5 percent of the 8-year follow-up period. Factors related to a better course, with longer periods of euthymia included (1) older age at onset of symptoms and first episode, (2) less severe depression, (3) less severe manic or hypomanic symptoms, (4) less frequent subsyndromal episodes, (5) fewer suicide attempts, (6) lower rates of sexual abuse, and (7) less family history of bipolar disorder or substance use disorders. The youth in the groups with greater periods of euthymia had a less complicated course overall, with lower rates of comorbid ADHD and anxiety disorders, and lower rates of all psychiatric disorders in family members. In fact, almost half of the group with the most time euthymic did not experience another episode during the 8-year follow-up. However, although the other half of group I was also euthymic for about 84 percent of the 8-year study period, they did continue to have new mood episodes during the follow-up period.

Comprehensive investigation of the stability of diagnosis of bipolar disorder in adults is uncommon and has yet to be addressed in any large scale studies of children and adolescents. The questions regarding BP in youth are for whom will the cycles be mainly depressive, or manic, how often will they reoccur, and for whom will BP completely and forever remit? This is important to keep in mind to determine how long a medication (mood stabilizer or antipsychotic) should be administered during and after resolution of an episode, and whether there are observable predictors of longitudinal course. In a 10-year follow-up of early episode 15 to 50 years old patients ever diagnosed with bipolar disorder by Carlson and colleagues, diagnostic consistency occurred in only half the sample. Much of the diagnostic instability was predicated on change in course of illness. BPI became clearer over time in 17 percent (as people with psychotic depression developed mania for example), but in the rest, even though BPI seemed obvious early, other conditions emerged. The most powerful predictors of diagnostic instability (i.e., not maintaining a BPI diagnosis) were having any childhood psychopathology (OR 2.8, CI 1.4 to 5.6), and specifically having a childhood externalizing disorder (OR 4.7, CI 2.5 to 8.9). The more acute and uncomplicated the diagnosis at onset, the more stable the diagnosis; however, poor premorbid adjustment in terms of childhood psychopathology and nonacute onset predicted diagnosis change. Prognostic implications were substantial as the consistently diagnosed bipolar group had significantly better functioning on Global Assessment Functioning (GAF) (p <0.001), spent less time symptomatic over their first 48 months of illness (26 percent vs. 47 percent) (p <0.001)

and much less time psychotic (7 percent vs. 31 percent) (p <0.001). Jake, the case previously described, represents either a youth with poor-outcome bipolar disorder, or one who has developed negative symptoms and is better considered schizoaffective.

Functional outcome outside of symptom status has not yet been addressed in longer-term longitudinal studies. The Child Global Assessment Scale score in Geller and colleagues' 8-year follow-up of early-onset mania was 33.5 (SD 10.6) for those with the disorder at age 18 or older. Another 1-year follow-up study reported functional recovery in only 39 percent of the sample. ADHD comorbidity and poor medication adherence was associated with poor function. However, it appeared that functional outcome was also associated with syndromal status.

TREATMENT

Current evidence-base addresses both medication and psychosocial intervention with considerably more data for the former in the treatment of mania. Industry-sponsored clinical trials have addressed the relative efficacy of atypical antipsychotics with less robust data for anticonvulsants and some data supporting lithium. Trials of medication for bipolar depression are much smaller and to date, only one regimen has been shown to be more effective than placebo. Psychosocial interventions are beneficial in addition to medication not replacements for it. If a child has SMD or if rages are the behaviors being targeted, their frequency, intensity, number, and duration need to be monitored. Medication management also requires obtaining baseline laboratory tests relevant to the medication used and monitoring side effects.

Practice parameters and clinical guidelines, such as the AACAP Practice Parameters are consistent in advocating the importance of medication.

Pharmacotherapy

As part of the *Best Pharmaceutical Act for Children,* newly developed medications to be used in children or adolescents must be tested for safety and effectiveness in that sample. U.S. Food and Drug Administration (FDA) labeling requirements allow fewer studies to be done if the condition —in this case mania—is the same condition that exists in adults, such that similar measures to test safety and efficacy are valid in that age group. A consensus conference convened in 2002 established that youth between ages 10 and 17 years can experience the same condition. The same BP disorder in younger children was less well established, hence the cutoff at age 10 years. This is important because it means that the medications and doses used for adults are being studied in older children and adolescents.

As of 2015, lithium, chlorpromazine, olanzapine, quetiapine, risperidone, aripiprazole, and asenapine have FDA approval in adolescents.

Risperidone, aripiprazole, quetiapine and asenapine have approval for children aged 10 to 17 years of age. There has been a positive randomized clinical trial in youth for ziprasidone but no FDA approval yet as of this writing. Although medication treatment for mania is the same, youth and adult treatment responses are not identical. Until recently, although lithium had been approved for adolescent mania, there were no decent-sized, placebo-controlled trials to support this "grandfathered" approval, although open trials of lithium suggested response rates of 50 to 60 percent. Four of the atypical antipsychotic medication trials completed for pediatric bipolar disorder, were in response to an FDA written request. Olanzapine was tested in adolescents, flexibly dosed, because the written request had been issued before the FDA decision to include children down to age 10 years. With doses ranging between 10 and 20 mg, 44.8 percent of youth on olanzapine showed at least a 50 percent reduction in symptoms compared to a placebo response rate of 18.5 percent. This was significant ($p < 0.001$). The mean decrease from the baseline Y-MRS entry score of 33 was 17.7 points and the effect size was 0.84.

Other atypical antipsychotic trials for risperidone, aripiprazole, and quetiapine were required to use low- and high-dose alternatives. Both high and low doses of risperidone and aripiprazole were significantly effective as measured by a 50 percent reduction in Y-MRS entry scores and a decrease of Y-MRS scores of between 14 and 18 points compared to placebo responses of 8 to 10 points. In general, 45 to 65 percent of subjects improved on active medication versus 18 to 37 percent on placebo. The absolute treatment benefit (percentage of drug responders minus placebo responders) averaged 19 to 38 percent, and the number needed to treat (NNT) in successful trials in which drug was better than placebo varied from about 2.6 (high-dose aripiprazole and high-dose risperidone) to around 5 (low-dose aripiprazole and high-dose quetiapine). Side effects were usually greater for the high-dose condition without a correspondingly higher improvement.

Several anticonvulsants and lithium have FDA treatment indication for mania in adults. Data for bipolar youth is less robust. An industry-sponsored trial with divalproex ER requested by the FDA, which was only 4 weeks in duration (the duration of adult trials), was negative; drug and placebo responses 24 and 23 percent, respectively. Published trials of two other anticonvulsants in youth, topiramate and oxcarbazepine, have not been found to be significantly better than placebo.

Atypical antipsychotics outperform divalproex in clinical trials. Additionally, where it was measured, the time to response for atypicals was shorter than for divalproex. Blood levels were adequate in both the industry-sponsored trial and in the ("Treatment of Early Age Mania" study [TEAM]. In the latter study, poor response likely had due to the >90 percent rate of comorbid ADHD.) As noted below, comorbid ADHD is a

predictor of poor treatment response.

There have been small trials of lithium in children for many years but none addressing acute mania with a placebo control until recently. The TEAM study, which randomized children to risperidone, lithium or divalproex reported a lithium response rate of 35.6 percent. There was no placebo arm but placebo responses in industry sponsored trials are about 25 to 30 percent. Thus, lithium would not have beaten placebo in that study. However, the Collaborative Lithium Trials (COLT) found a 4.5 point difference on the Y-MRS scale at week 8 between lithium and placebo-treated participants (p = 0.031). Even so, rates of response (≥50 percent drop in Y-MRS) did not differ significantly between lithium and placebo groups, respectively (37.7 percent vs. 28.6 percent, Robb et al., 2014).

Correll and colleagues provide a meta-analysis with relative effect sizes of different classes of medication. Atypical antipsychotics for mania/mixed mania show a larger effect size than mood stabilizers (effect size [ES] 0.65, confidence interval [CI] 0.53 to 0.78 compared to ES 0.20 CI 0.02 to 0.39) and a 50 percent reduction in Y-MRS entry score is seen in 50 to 70 percent of subjects compared to 20 to 37 percent on placebo. The NNT in successful trials where drug was better than placebo varied from 3.8 to 5.3.

The presence of ADHD complicates and attenuates treatment response for youth with BP. Attempts to treat both conditions include one study in which mixed amphetamine salts was added to divalproex decreasing ADHD symptoms in that group. In another study, children with bipolar disorder stabilized on a thymoleptic received varying doses of methylphenidate (MPH) and placebo. ADHD rating scale scores were significantly improved compared to placebo in the MPH-treated youth. MPH added to aripiprazole did not destabilize youth in a cross-over study (though ADHD was not significantly improved, either). Finally, a small open label study of atomoxetine added to a mood-stabilized regimen for BP and ADHD and found atomoxetine was effective in treating the ADHD, although 2 subjects out of 12 reported worsening mood at some point during the 8 week study.

Given the level of severity and impairment with which most children and teens begin a bipolar treatment study, the improvement, when it occurs, is noteworthy, but it often is not sufficient. Few data are available to predict treatment response. Use of more than one medication has increasingly become acceptable for treating mania/bipolar disorder in both adults and young people. In youth, there are three types of studies: (1) comparing two medications to one—for example, combined divalproex and quetiapine versus quetiapine alone; (2) adding one to another if the first drug does not work—for example, adding risperidone to lithium or divalproex; and (3) starting two together—for example, lithium and divalproex—and discontinuing one.

Although the types of adverse effects of medication are similar in children and adolescents, the frequency and severity may differ. For

instance, weight gain with metabolic syndrome especially in medication-naive patients and cases of polycystic ovaries may be of particular concern in young people. Similarly, disinhibited behavioral responses to many types of central nervous system medications seem to occur more often in children and perhaps slightly less so in teens.

Prevention of Relapse

Children and adolescents with bipolar disorder appear to have higher relapse rates and more episodes over the course of a year than adults. Regarding relapse prevention for mania, there is some suggestion that aripiprazole in one study and lamotrigine in another by Findling increase the time to relapse for mania. Beyond these trials, it is prudent to assume that medications that stabilized a child with BP should be continued for at least an entire school year, while a child is in school. The AACAP Practice Parameter suggested 12 to 14 months. In cases in which it has taken a long time to stabilize the child and in which there is clear evidence from symptom monitoring that treatment has been effective, it may be wise to continue the treatment until the child has improved function, academically, socially, and within the family. The decision to continue medication treatment or not should be assessed annually. In cases in which treatment response is more tenuous and side effects more troublesome, longer maintenance is more difficult to support.

Treatment of Bipolar Depression

Treatment of bipolar depression in young people has received relatively little attention. Even in adults, only three medications (olanzapine/fluoxetine combined [OFC], quetiapine, and lurasidone) have FDA-approved indications for bipolar depression. In youths, there have been two open trials—one of lithium monotherapy in hospitalized adolescents with bipolar depression and one using lamotrigine alone or with other medications in outpatient adolescents. Symptoms declined in both studies. In the absence of placebo-controlled studies, and given the high placebo response in depression studies in youth in general, the treatment utility of these medications in youth remains to be demonstrated. Moreover, the adolescent samples studied had considerably more subthreshold manic symptoms (Y-MRS scores were 12 or fewer) than in the adults, whose Y-MRS scores were 2 or fewer. Thus, it is not even possible to extrapolate intelligently from adult data.

The one positive study of bipolar depression used the OFC. This regimen was superior to placebo (effect size of 0.46) and has been approved by the FDA for the acute treatment of bipolar I depression in patients 10 to 17 years of age.

There is a modest, if anecdotal, literature on the efficacy of

electroconvulsive therapy (ECT) in mood disorders in teens. The reader is referred to the AACAP's ECT Practice Parameter on its indications and contraindications.

Treatment of "Broad Phenotype" Bipolar Disorder or Explosive Outbursts

Treatment of youth diagnosed with the "broad phenotype bipolar disorder"/SMD by some clinicians, unspecified bipolar disorder by others, and mania with comorbid ADHD by still others, is currently dependent on consensus recommendations. Most clinicians recommend stabilizing the mood disorder first and then addressing the comorbid disorder. ADHD is the only condition that has been systematically studied in children with mania or unspecified bipolar disorder. A recent meta-analysis found that comorbid ADHD with BP in youth, lowered treatment responsiveness to the bipolar disorder treatment, which suggests that the ADHD may need to be addressed first. Although there have been concerns among some clinicians that using stimulant medications for ADHD in children with definite or possible mania may exacerbate or precipitate mania, an increasing number of studies are substantiating the utility of treating comorbid ADHD either before or after mood stabilization. It is important to note that treating ADHD with therapeutic doses of stimulants is quite different from the effects of abusing cocaine and other stimulant drugs which may truly precipitate mania and psychosis.

Psychosocial Treatment

Psychosocial interventions (1) target the recognition of future episodes before they progress too far, (2) enhance treatment adherence, and (3) address environmental stressors that act as possible precipitants to further episodes. There are four models of psychosocial treatment studied in children and adolescents with early-onset bipolar disorder. These include multifamily and individual family psychoeducation, family-focused therapy for adolescents, child and family-focused CBT for younger children, and collaborative problem solving. Each emphasizes the importance of psychoeducation and destigmatization, as well as increasing parental collaboration by reducing parent self-blame for "causing the disorder." Other components include helping parents to distinguish between normative developmental behavior and bipolar symptoms, taking proactive steps to decrease the risk of relapse, developing tools for effectively managing emotional arousal, improving family communication skills, and teaching adaptive problem solving strategies.

FUTURE DIRECTIONS

Developments in neurosciences, neurobiology, genetics, and neuroimaging, as well as longer-term follow-up studies, will undoubtedly help to advance

the understanding of this complex and disabling disorder, in particular its relationship to other psychiatric disorders such as ADHD, conduct disorder, developmental disorders, and psychotic disorders. It is hoped that such advances can shed light on the etiology and underlying pathophysiology, including the identification of dysfunctional brain circuits that may underlie pediatric mania.

REFERENCES

Althoff RR, Faraone SV, Rettew DC, Morley CP, Hudziak JJ. Family, twin, adoption, and molecular genetic studies of juvenile bipolar disorder. *Bipolar Disord*. 2005;7:598–609.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Arlington, VA: American Psychiatric Association; 2013.

Arnold LE, Demeter C, Mount K, et al. Pediatric bipolar spectrum disorder and ADHD: comparison and comorbidity in the LAMS sample. *Bipolar Disord*. 2011;13:509–521.

Benarous X, Rafiln M, Bodeau N, Dhossche D, Cohen D, Consoli A. Adverse Childhood Experiences Among Inpatient Youths with Severe and Early-Onset Psychiatric Disorders: Prevalence and Clinical Correlates [published online ahed of print March 22, 2016]. *Child Psychiatry Hum Dev*. DOI: 10.1007/s10578-016-0637-4

Biederman J, Mick E, Faraone SV, Spencer T, Wilens TE, Wozniak J. Pediatric mania: a developmental subtype of bipolar disorder? *Biol Psychiatry*. 2000;48:458–466.

Birmaher B, Axelson D, Goldstein B, et al. Four-year longitudinal course of children and adolescents with bipolar spectrum disorders: the Course and Outcome of Bipolar Youth (COBY) study. *Am J Psychiatry*. 2009;166:795–804.

Birmaher B, Gill MK, Axelson D, et al. Longitudinal trajectories of and associated baseline predictors in youths with bipolar spectrum disorders. *Am J Psychiatry*. 2014;171:990–999.

Blader J, Carlson GA. Increased rates of bipolar disorder diagnoses among U.S. child, adolescent, and adult inpatients, 1996–2004. *Biol Psychiatry*. 2007;62:107–114.

Carlson GA, Glovinsky I. The concept of bipolar disorder in children: a history of the bipolar controversy. *Child Adolesc Psychiatr Clin N Am*. 2009;18:257–271, vii.

Carlson GA, Klein DN. How to understand divergent views on bipolar disorder in youth. *Annu Rev Clin Psychol*. 2014;10:529–551.

Carlson GA, Kotov R, Chang SW, Ruggero C, Bromet EJ. Early determinants of four year clinical outcomes in bipolar disorder with psychosis. *Bipolar Disord*. 2012;14:19–30.

Carlson GA, Palaki C. Understanding early age of onset: a review of the last 5 years. *Citrr Psychiatry Rep*. 2016;18:114.

Correll CU, Sheridan EM, DelBello MP. Antipsychotic and mood stabilizer efficacy and tolerability in pediatric and adult patients with bipolar I mania: a comparative analysis of acute, randomized, placebo-controlled trials. *Bipolar Disord*. 2010;12(2):116–141

Detke HC, DelBello MP, Landry J, Usher RW. Olanzapine/Fluoxetine combination in children and adolescents with bipolar I depression: a randomized, double-blind, placebo-controlled trial. *J Am Acad Child Adolesc Psychiatry*. 2015;54(3):217–224.

Dubicka B, Carlson GA, Vail A, Harrington R. Prepubertal mania: diagnostic differences between U.S. and UK clinicians. *Eur Child Adolesc Psychiatry*. 2008;17:153–161.

Faraone SV, Biederman J, Wozniak J, Mundy E, Mennin D, O'Donnell D. Is comorbidity with ADHD a marker for juvenile-onset mania? *J Am Acad Child Adolesc Psychiatry*. 1997;36:1046–1055.

Findling RL, Cavus I, Pappadopolous E, et al. Efficacy, long-term safety, and tolerability of ziprasidone in children and adolescents with bipolar disorder. *J Child Adolesc Psychopharmacol*. 2013;23:545–557.

Findling RL, Chang K, Robb A, et al. Adjunctive maintenance lamotrigine for pediatric bipolar i disorder: a placebo-controlled, randomized withdrawal study. *J Am Acad Child Adolesc Psychiatry*. 2015;54:1020–1031.

Findling RL, Pathak S, Earley WR, Liu S, DelBello M. Safety, tolerability, and efficacy of quetiapine in youth with schizophrenia or bipolar I disorder: a 26-week open-label continuation study. *J Child Adolesc Psychopharmacol*. 2013;23(7):490–501.

Findling RL, Robb A, McNamara NK, et al. Lithium in the acute treatment of bipolar I disorder: a double-blind, placebo-controlled study. *Pediatrics*. 2015;136:885–894.

Findling RL, Short EJ, McNamara NK, et al. Methylphenidate treatment of children and adolescents with bipolar disorder and attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2007;46:1445–1453.

Findling RL, Youngstrom EA, McNamara NK, et al. Double-blind, randomized placebo-controlled long-term maintenance study of aripiprazole in children with bipolar disorder. *J Clin Psychiatry*. 2012;73:57–63.

Garrett AS, Reiss AL, Howe ME, et al. Abnormal amygdala and prefrontal cortex activation to facial expressions in pediatric bipolar disorder. *J Am Acad Child Adolesc Psychiatry*. 2012;51:821–831.

Geller B, Luby JL, Joshi P, et al. A randomized controlled trial of risperidone, lithium or divalproex sodium for initial treatment of bipolar I disorder, manic or mixed manic phase, in children and adolescents. *Arch Gen Psychiatry*. 2012;69:515–528.

Geller B, Tillman R, Bolhofner K, Zimerman B. Child bipolar I disorder: prospective continuity with adult bipolar I disorder; characteristics of second and third episodes; predictors of 8-year outcome. *Arch Gen Psychiatry*. 2008;65:1125–1133.

Goodwin F, Jamison KR. *Manic-Depressive Illness*. Vol. II. New York: Oxford University Press; 2007: Chapters 4, 6, 23.

Hafeman DM, Merranko J, Axelson D, et al. Toward the definition of a bipolar prodrome: dimensional predictors of bipolar spectrum disorders in at-risk youths. *Am J Psychiatry*. 2016;173:695–704.

Hudziak JJ, Althoff RR, Derks EM, Faraone SV, Boomsma DI. Prevalence and genetic architecture of Child Behavior Checklist-juvenile bipolar disorder. *Biol Psychiatry*. 2005;58:562–568.

Leibenluft E. Severe mood dysregulation irritability and the diagnostic boundaries of bipolar disorder in youths. *Am J Psychiatry*. 2011;168:129–142.

McClellan J, Kowatch RA, Findling RL; Work Group on Quality Issues. Practice parameter for the assessment and treatment of children and adolescents with bipolar disorder. *J Am Acad Child Adolesc Psychiatry*. 2007;46:107–125.

Merikangas KR, Cui L, Kattan G, Carlson GA, Youngstrom EA, Angst J. Mania with and without

depression in a community sample of US adolescents. *Arch Gen Psychiatry*. 2012;69:943–951.

Pataki C, Carlson GA. The comorbidity of ADHD and bipolar disorder: any less confusion? *Curr Psychiatry Rep*. 2013;15:372–378.

Preisig M, Strippoli MP, Castelao E, et al. The specificity of the familial aggregation of early-onset bipolar disorder: a controlled 10-year follow-up study of offspring of parents with mood disorders. *J Affect Disord*. 2016;190:26–33.

Serafini G, Pompili M, Borgwardt S, et al. Brain changes in early-onset bipolar and unipolar depressive disorders: a systematic review in children and adolescents. *Eur Child Adolesc Psychiatry*. 2014;23:1023–1041.

Stringaris A. What we can all learn from the Treatment of Early Age Mania (TEAM) trial. *J Am Acad Child Adolesc Psychiatry*. 2012;51:861–863.

Stringaris A, Santosh P, Leibenluft E, Goodman R. Youth meeting symptom and impairment criteria for mania-like episodes lasting less than four days: an epidemiological enquiry. *J Child Psychol Psychiatry*. 2010;51:31–38.

van Hulzcn KJ, Scholz CJ, Franke B, et al; PGC ADHDWorking Group; PGC Bipolar Disorder Working Group. Genetic overlap between attention-deficit/hyperactivity disorder and bipolar disorder: evidence from genome-wide association studymeta-analysis [published online ahead of pint Otober 18, 2016]. *Biol Psychiatry*. doi: 10.1016/j.biopsych.20l6.08.040

Van Meter AR, Moreira AL, Youngstrom EA. Meta-analysis of epidemiologic studies of pediatric bipolar disorder. *J Clin Psychiatry*. 2011;72:1250–1256.

Whitney J, Howe M, Shoemaker V, et al. Socio-emotional processing and functioning of youth at high risk for bipolar disorder. *J Affect Disord*. 2013;148:112–117.

Youngstrom E, Meyers O, Youngstrom JK, Calabrese JR, Findling RL. Diagnostic and measurement issues in the assessment of pediatric bipolar disorder: implications for understanding mood disorder across the life cycle. *Dev Psychopathol*. 2006;18:989–1021.

52

Anxiety Disorders in Children

▲ 52.1 Obsessive-Compulsive Disorder in Childhood

ADAM B. LEWIN, PH.D., AND JOHN PIACENTINI, PH.D.

Pediatric obsessive-compulsive disorder (OCD) is a chronic neuropsychiatric condition associated with broad impairments in functioning. Over the past two decades, increased attention has been focused on OCD in childhood, given heightened public and professional awareness of substantial upward revision of the prevalence of this disorder in youth. Advances in functional neuroimaging, genetics, neurosurgery, neuropsychological, and psychoneuroimmunology have contributed to a better understanding of etiological factors contributing to the pathogenesis of OCD in childhood. Concurrently, treatment-outcome research has produced multiple controlled trials of first-line treatments of OCD in childhood, cognitive-behavioral therapy (CBT) with exposure and response prevention (E/RP), and serotonin reuptake inhibitors (SRIs). More recently, studies have extended support for family-based E/RP to very young children with OCD. Nevertheless, despite promising interventions OCD remains a chronic and disabling disease affecting children and families; marked social, academic, and familial dysfunction are common sequelae. For individuals who do not benefit from available treatment, further refinement of treatment approaches remains paramount in order to diminish suffering and impairment while increasing the likelihood that youth with OCD can maintain relatively normal lives. Because many youth do not receive treatment due to lack of access to child psychiatrists and psychologists trained in evidence-based intervention, improving availability of care is another key goal.

DEFINITIONS

OCD is characterized by the presence of recurrent obsessions or compulsions that are distressing, time consuming, or debilitating. Obsessions are defined as recurrent, persistent, and intrusive thoughts, ideas, impulses, or mental images that are experienced by the individual as

distressing and as a product of one's own mind (to differentiate them from thought insertion). These thoughts or ideas are usually described as generating (or associated with) feelings of anxiety, tension, disgust, or aversion. Common obsessional fears are listed in Table 52.1–1. Compulsions are purposeful, repetitive behaviors, or mental acts that typically function to neutralize, prevent, or reduce anxiety. Common compulsive behaviors are also listed in Table 52.1–1. Often, compulsions are performed in response to an obsessive thought or according to some other rigidly applied rules. Although compulsions function to mitigate distress or anxiety or prevent some dreaded event, they are clearly excessive or are not realistically connected with the triggering stimulus (e.g., tapping three times to prevent illness after someone in the room coughs). Although obsessions and compulsions are frequently ego dystonic and are often recognized by children as senseless or exaggerated, such insight is not required for the diagnosis of OCD in childhood. Otherwise, symptoms to OCD in childhood resemble those seen in adults.

HISTORY

Early examples of childhood onset OCD, consisting of case descriptions, date back to the late 19th century. In 1895, Sigmund Freud described an 11-year-old child with obsessions. Subsequently, the French physician Pierre Janet provided the first description consistent with OCD in childhood in his 1903 report on a 5-year-old boy. In the first half of the 20th century, descriptions of clinical samples and generalizations emerged. In 1942, L. Berman extracted clinical information on 62 children with possible obsessive-compulsive neurosis out of a case registry of 3,050 cases. In 1955, Louise Despert described more than 60 cases of children with "obsessive-compulsive neurosis," noting a male preponderance, children's recognition that their symptoms are abnormal and unwanted, and the tendency to attempt to conceal their symptoms from treatment professionals and others. Leo Kanner's 1962 textbook of child psychiatry contained descriptions of childhood OCD and the consequences of social isolation and family accommodation.

In 1965, Lewis Judd provided descriptive criteria for the diagnosis of OCD in children that closely resembles the current diagnostic criteria, including the requirement of significant impairment in functioning. Reports as recent as 1973 implicated parental withdrawal of attention as a critical causal factor in childhood OCD. However, by the 1980s, the notion that OCD was a neurobiological disorder with possible childhood onset became more commonplace. Further, findings from the Epidemiologic Catchment Area (ECA) study that indicate that most adults with OCD report an onset of the disorder during adolescence prompted increased awareness of the disorder in childhood as well as research in this area. Programmatic work in the 1980s under the direction of Judith Rapoport

and her group at the National Institute of Mental Health provided the basis for the first systematic studies of epidemiology, phenomenology, and treatment of childhood- and adolescent-onset OCD as well as fundamentally dismissing any notion that the emergence of OCD in youth was unusual or unexpected.

Table 52.1–1.
Common Symptoms of Pediatric Obsessive-Compulsive Disorder Obsessions

Obsessions
Concerns with dirt, germ exposure, fears of illness
Fears of harm befalling self or others
Need for symmetry, order, exactness
Need to save or hoard
Excessive moral, religious, or sexual concerns
Compulsions
Cleaning rituals
Checking
Repeating behaviors
Ordering or arranging
Confessing, reassurance seeking

COMPARATIVE NOSOLOGY

Diagnostic criteria for OCD have remained relatively constant over the past three editions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM). The DSM-5 published in 2013 defines the following features of OCD: (1) recurrent obsessions or compulsions; (2) obsessions or compulsions cause marked distress, are time consuming (take more than 1 hour per day) or cause significant impairment in functioning; and (3) obsessions and compulsions not due to the presence of another mental disorder, substance, or other medical condition. The DSM-5 also includes specifiers to code degree of insight and to indicate whether the individual has a history of tic disorder. The tenth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) provides a description of OCD that is consistent with DSM-5.

EPIDEMIOLOGY

Although estimates vary, studies suggest lifetime prevalence for OCD in children between 1 and 3 percent indicating that OCD is more common than previously thought and as common in youth as in adults. Sample age may explain some of the variability in these estimates. Older adolescents yield prevalence rates between 0.53 and 4.0 percent, whereas estimates for school-aged children remain lower. In fact, a large prevalence study from the United Kingdom showed an exponential rise in rates across age groups

(5 to 7, 8 to 10, 11, 12, and 13 years). Another recent community sample found a prevalence rate of 3 percent for "clinical" OCD, defined as symptoms severe enough to cause distress and interference in functioning. The most commonly accepted 6-month prevalence rate is between .5 and 1 percent of the general pediatric population. Of note, rates of OCD in late adolescence make it more common than panic disorder, schizophrenia, and bipolar disorder.

Mean age of onset is between ages 6 and 11. Onset age appears bimodal and varies by gender with a male preponderance in cases of prepubertal onset with an initial peak of incidence around puberty and subsequently in young adulthood. The gender distribution becomes roughly equivalent during adolescence. It is currently believed that up to 80 percent of adult OCD cases have an initial onset during childhood. Children with an early onset of OCD (below age 7) have also been shown more likely to be male and to have a family history of OCD in comparison to children with later onsets, suggesting that genetic factors may be more likely to play a role in the development of early onset cases. Referral bias may account for the underrepresentation of minorities and youngsters from lower socioeconomic strata in clinical samples of OCD youth, as this pattern has not been replicated in epidemiological studies with adult samples.

The occurrence of OCD in childhood may be somewhat higher than reported given secrecy that children often display in reporting embarrassing thoughts and behaviors. Additionally, limited insight into the abnormality of their symptoms (and linking OCD symptoms to potential impairment), parent inability to recognize OCD symptoms, and the lack of awareness about the availability of efficacious treatments may contribute to underdiagnosis. Therefore, it is reasonable to assume that many childhood cases of OCD remain undetected.

Table 52.1–2.
Comorbid Diagnoses in Pediatric Obsessive-Compulsive Disorder

Diagnosis	Frequency (%)
Anxiety disorders	26–70
Mood disorders	11–62
Tic disorders	11–26
Attention-deficit/hyperactivity disorder	16–20
Disruptive behavior disorders	9–19
Autism spectrum disorders	3–7

Comorbidity

In childhood, psychiatric disorders co-occurring with OCD are the norm rather than the exception. For example, in an analysis of 70 clinical cases of youth with OCD, 84 percent carried at least one additional Axis I

psychiatric diagnosis. It has been estimated that as many as 50 percent of pediatric patients with OCD have two or more co-occurring diagnoses (Table 52.1–2). The most common comorbid disorders include other anxiety disorders, depressive disorders, tic disorders including Tourette disorder, attention-deficit/hyperactivity disorder (ADHD), and disruptive behavior disorders.

ETIOLOGY

Neurobiological Factors

Although the cause of OCD remains unknown, multiple neurochemical, neuroanatomical, and immunological etiologic theories have been proposed. Clinical descriptions of OCD associated with Sydenham chorea (SC), tic disorders, traumatic brain injury, encephalitis, and Huntington chorea have helped to solidify the broad concept of the disorder as brain based and presumably reflecting impaired striatal functioning. Although the neurocircuitry and neurochemistry involved in OCD symptomology have been better clarified through the use of in vivo brain techniques and evaluation of responses to psychoactive medications, there continues to be a need to further refine etiological models of the condition.

Neurochemical. Dysfunction of the central serotoninergic neurotransmitter system continues to be suspected in OCD symptomology. Primary support for this premise is observed efficacy of selective serotonin reuptake inhibitors (SSRIs) and altered sensitivity to acute administration of 5-hydroxytryptamine (5-HT) agonists, as well as from results of preliminary studies of 5-HT synthesis. Platelet [3H]imipramine binding has been observed to be reduced in young patients with OCD; conversely, those OCD patients with higher [3H]imipramine binding at baseline differentiated responders versus nonresponders to SSRI treatment.

Involvement of the dopaminergic neurotransmitter system has also been suggested. This may be especially relevant to childhood OCD given the increased rate of comorbid tic disorders. Clinically, administration of psychostimulants appears to worsen obsessions and compulsions. Conversely, the coadministration of dopamine antagonists with SSRIs may augment the beneficial effects of these agents. However, direct assessment of the dopamine system remains unstudied in children, including a lack of controlled administration of dopamine antagonists. Further, animal studies suggest that excessive glutamatergic output from the cortex to the striatum is associated with repetitive movements. Human studies in children with OCD have reported correlations between striatal glutamate with OCD symptoms and the drop in striatal glutamate with SSRI treatment. In sum, available neurochemical models suggest the possibility of several neurotransmitter abnormalities associated with OCD.

Neurostructural. The pathogenesis of pediatric OCD has been characterized by abnormalities in the prefrontal–basal ganglia circuits, specifically frontal–striatal–thalamic–cortical (FSTC) circuit disruption. This premise is supported by imaging studies of OCD, co-occurrence with other disorders with presumed basal ganglia disruption, as well as neuropsychological studies. Volumetric analyses using both magnetic resonance imaging (MRI) and computed tomography (CT) approaches have revealed that treatment-naive children and adults with OCD have smaller basal ganglia segments than adults. For example, a recent study found reduced globus pallidus volume and gray matter volumetric abnormalities in the anterior cingulated gyrus in pediatric patients with OCD compared to healthy controls. More recently, abnormalities in white matter coherence have been found in child OCD, with greater white matter irregularities linked to increased OCD severity differences compared to healthy controls.

Treatment-naive children appear to also have large thalamic volumes. Thus, it has been hypothesized that children with OCD display abnormalities in normal developmental remodeling of neuronal structures affecting integration, response inhibition, gating/cognitive control, and processing. Positron emission tomography (PET) suggests dysfunction in FSTC circuitry, with studies showing increased glucose metabolism in orbitofrontal cortex, ventral striatum, right caudate nucleus, and anterior cingulate gyrus, at rest and in response to symptom provocation. Orbitofrontal and caudate metabolic rates are highly correlated, suggesting that interconnecting loops are freed from usual inhibitory influences and reverberate. Disruption of affective/reward processing loops and cognitive control loops have been linked to OCD, and differential development of these systems may be expressed in OCD course and presentation. Recent studies in child OCD suggest that differential development of FSTC circuits and various interconnectivity (e.g., between behavioral inhibition) may explain differential symptom expression, for example, a finding suggested hypoconnectivity of cognitive control circuity in early-onset cases.

The observations that both medication and behavioral therapies reduce orbitofrontal and caudate blood flow seem to validate the importance of these brain regions in OCD pathogenesis. Paradoxically, lower baseline orbitofrontal glucose metabolism predicts better response to pharmacotherapy. Imaging also suggests the frontal–limbic–basal ganglia system may mediate OCD symptom expression, demonstrating a wider interaction of cognitive and emotional circuits. Specifically, symptom provocation appears to activate paralimbic regions as well as insular, anterior temporal, and amygdalar regions.

Neuroimmunology. Significant attention has been focused on the triggering of OCD-like symptoms following group A beta-hemolytic

streptococcal infections (GABHS), similar to OCD symptoms observed in SC. These infection-triggered cases of OCD have been coined pediatric autoimmune neuropsychiatric disorders associated with streptococcus (PANDAS) and are believed to represent an autoimmune process related, if not identical to, that associated with rheumatic fever (RF) and SC. (It is noteworthy that patients with RF and SC both display higher frequencies of obsessions and compulsions versus comparison groups, and up to one third of patients with SC meet research criteria for the diagnosis of OCD.) Immune activation leads to the inflammation of the basal ganglia and disruption of the cortical–striatal–thalamo–cortical (CSTC) function. This acquired basal ganglic dysfunction may result in choreiform movements, tics, obsessions, compulsions, and hyperactivity. PANDAS are characterized by abrupt childhood onset and episodic or sawtoothed course. Relationships between the size of the basal ganglia nuclei and OCD symptom severity have been observed using MRI. Similarly, small sample studies of immune-modulating treatments (intravenous immunoglobulin G [IgG] and plasmapheresis) have been associated with improvements in OCD patients with suspected PANDAS. However, there is a lack of well-controlled studies; these interventions remain experimental, and efforts to establish markers of PANDAS have been inconclusive. In addition, the proportion of child OCD cases potentially attributable to streptococcal infection remains unknown. Although additional research into PANDAS appears warranted, the association between GABHS infection and OCD remains debated. In light of the current literature, evidence-based psychopharmacological and behavioral (e.g., CBT with E/RP) approaches remain first-line options for treating OCD symptoms of suspected immunological origin. Experimental, high-risk interventions should be used cautiously given safety concerns and lack of clear evidence in support of their efficacy.

More recently, the term PANS (pediatric acute-onset neuropsychiatric syndrome) was coined to characterize cases with similar clinical presentations as PANDAS but without link or presumed cause with GAS infection. Similar to PANDAS, PANS requires abrupt/dramatic/overnight onset of OCD symptoms, and severe and acute onset of at least two of the following: anxiety, emotional lability, irritability/aggression, developmental regression, academic deterioration, sensory sensitivity, somatic symptoms (including sleep disturbance, enuresis or urinary frequency). Similar to PANDAS, diagnosis is not straightforward and even putative PANS casesness does prohibit potential benefits of CBT or SSRIs nor does it necessitate the need for experimental interventions. Given the lack of definitive tests for PANS/PANDAS, evaluation requires careful multifaceted assessment of medical and psychiatric histories. Parents should be cautioned regarding pursuit of PANS/PANDAS diagnosis/treatment without first exhausting evidence-based OCD

treatments.

Neuroendocrine. Observations of OCD onset during puberty, pregnancy, and during the postpartum period have led to speculation that hormones may influence the emergence or severity of OCD symptoms. Although transitory, antiandrogen treatment has appeared to show benefits in some refractory individuals. The effects of adrenarche and menarche on relevant brain circuits that underlie OCD symptoms need further study.

Family Genetic, Behavioral, and Cognitive Factors

Family Genetic. OCD is a heritable disorder—the familial nature in most cases of OCD is well recognized. Monozygotic twins have a higher OCD concordance rate (0.52) compared to dizygotic (0.21). A meta-analysis of studies exploring familial patterns of psychopathology in relatives of subjects with OCD has documented a fourfold increased frequency of OCD in first-degree relatives of OCD probands (8.2 vs. 2.0 from control families). Rates of OCD among parents of youngsters with OCD range from 17 to 19 percent. A 2005 review of twin studies suggests 27 to 65 percent heritability rates for OCD (including nonpediatric onset). In addition, subclinical features of OCD (usually defined as typical symptom constellations without the level of interference required for formal diagnosis) are higher among family members with OCD (e.g., reported in 52 percent of parents). In fact, relatives of probands with early onset OCD exhibit even higher rates of OCD and tic symptoms. Analysis of family patterns of OCD finds support for subclinical OCD as an intermediate phenotype that shares similar genetic risk factors.

The relationship between OCD and tic disorders, including Tourette syndrome, has emerged from family studies. Tic disorders and some cases of OCD appear to share some susceptibility factors, leading to arguments that OCD may be an alternative expression of the risk factors underlying tic disorders. Some suggest an autosomal dominant model of inheritance with incomplete penetrance. In other words, the phenotypic expression varies in expression from OCD symptoms to tic disorders.

Data from twin and family studies suggest that OCD may be transmitted in families, but specific genes for the disorder have not been identified. Findings from the multisite Obsessive Compulsive Genetic Association Study (OCGAS), the largest such study to date (1,065 families and 1,406 patients with OCD), failed to identify single nucleotide polymorphisms at genome-wide significance. Investigators posit larger samples that are needed to identify genes associated with OCD. Preliminary findings suggest OCD is polygenetic and likely arises from a large number of diverse genes. Candidate gene studies have yet to produce replicable, significant genes culpable in OCD. The neuronal glutamate transporter gene (SLC1A1) has the strongest support, with some findings implicating the serotonin

promoter variant (5-HTTLPR). Definitive copy number variants associated with OCD have not been identified. To date, OCD-specific genetic markers remain to be found.

Behavioral Factors. Hobart Mower's two-factor conditioning theory provides a behavioral explanation for OCD symptoms. Via classical conditioning, a neutral stimulus (or event) becomes conditioned to elicit distress because of its association with an unrelated aversive (or feared) stimulus. Fear can be conditioned to both mental events (e.g., blasphemous thoughts) and physical stimuli or behaviors (e.g., a "contaminated" object, items perceived as dangerous, driving, eating). Subsequent to the acquisition of the conditioned fear, rituals develop to reduce or avoid distress. These rituals (or compulsive behaviors) persist and become excessive because they temporarily ameliorate the distress associated with obsessive thoughts via operant conditioning (specifically, a negative reinforcement paradigm). In other words, the escape-from/reduction-of anxiety associated with performing a compulsive behavior or ritual results in an increase likelihood of the ritual being used in future instances of anxiety.

However, specific traumatic events responsible for the conditioned response are rarely identified. Additionally, behavioral theories do not adequately explain the acquisition of obsessive fears. Regardless, these behavioral theories are useful for understanding the maintenance of many OCD symptoms and provide the foundation for behavioral intervention.

Cognitive Factors. The cognitive perspective of the etiology of OCD posits that OCD results from the interpretation of intrusive thoughts rather than the frequency or content of these cognitions. If intrusive thoughts are interpreted in such a way that an individual perceives responsibility for causing or failing to prevent harm, obsessional patterns develop. Attempts to neutralize intrusive thoughts (or obsessions) via motor or cognitive rituals, avoidance, and reassurance-seeking behavior prevent the disconfirmation of the patient's fears and facilitate the proliferation of anxiety. The most common cognitive misinterpretations include thought-action fusion or the tendency to equate thought with action (e.g., thinking something bad is the same as doing it), and exaggerated responsibility for potential untoward consequences of inaction (e.g., "If I don't wash my hands and someone gets sick as a result, then it's my fault.").

Similar to behavioral theory, cognitive theory does not adequately explain the etiology of symptoms. It is likely that cognitive misinterpretations and reinforcement theory (see above) maintain and expand OCD symptoms in individuals with biological propensity for OCD development.

DIAGNOSIS AND CLINICAL FEATURES

Presenting problems among children and adolescents with OCD generally resemble the adult presentation. However, compared to adults, children are more likely to deny impairment due to obsessions and compulsions. Children may also be somewhat less likely to consider their OCD symptoms excessive, perhaps due to lower levels of cognitive awareness. Children do not need to recognize their symptoms as senseless or unrealistic for diagnosis to be assigned; the level of insight varies considerably likely related to a child's developmental level. Unlike most adults, children might not report symptoms to be ego dystonic. However, children are typically willing to discuss distress from family conflict arising from their rituals (e.g., chronically running late, excessive need for reassurance, time burden on parents, disruption of siblings functioning) or interference in schoolwork resulting in dropping grades. Overall, most children are quite open in describing their symptoms and seek assistance in better managing the illness.

Although comorbidity is the rule rather than the exception in childhood OCD, the content of the obsessions and compulsions cannot be better explained by a co-occurring Axis I disorder. Examples include self-deprecating ruminations associated with depression, obsessive concern with appearance in body dysmorphic disorder, hair pulling in trichotillomania, and fears of harm befalling a loved one in separation anxiety disorder. Similarly, the symptoms also are required not to be substance induced or due to the influence of a medical condition. The range of obsessions and compulsions experienced in childhood is similar to that of adults with OCD, although children are more likely to engage in compulsive reassurance seeking and involve family members in their rituals. Most children with OCD have both obsessions and compulsions; pure obsessional illness is rare in this age group. In fact, most childhood cases of OCD involve multiple obsessions and compulsions (mean number of four current compulsions and four current obsessions in one report). Compulsions in the absence of anxiety-related or otherwise distress--inducing obsessions may occur in (mostly younger) children, who often describe their rituals as being performed in response to an irresistible urge or an otherwise vague sensation (the "just-right" phenomena). The absence of reported obsessions in these cases may be due to underdeveloped awareness or ability to articulate these thoughts or result of a misdiagnosed tic disorder. Specific types of symptoms (e.g., touching/tapping/rubbing) may change with age, but the number of symptoms generally remains constant.

The most common obsessions in childhood include fears of contamination (e.g., dirt, exposure to germs, toxins, contracting an illness), fears of harm befalling self or others (e.g., becoming aggressive, failing to prevent harm), need for symmetry or exactness, hoarding or saving

concerns, and excessive moralization or religiosity. Obsessions also may manifest as excessive guilt, self-doubt, or other forbidden thoughts. Common compulsions involve cleaning/washing rituals, checking, repeating behaviors (touching, tapping, rubbing), counting, ordering/arranging, avoidance, doubt/indecision, confessing, seeking reassurance, and slowness. The content and type of obsessive and compulsive symptoms may change over time. OCD is also stress sensitive, and many children experience acute symptom exacerbations during times of psychosocial stress or change (e.g., attending a new school, moving to a new home, death or a family member).

Similar to adult OCD, family members are commonly drawn into the child's OCD symptoms through provision of reassurance to the child, modification of family routines as a result of the child's symptoms, or even actual participation in the child's compulsive rituals. Such accommodation can negatively impact the family environment and lead to increased levels of family conflict and negative emotional reactions toward the affected child. Accommodation has also been associated with increased feelings of depression and anxiety on the part of the accommodating family member (typically mothers).

Given the typically strange and senseless nature of many OCD symptoms, many children and adolescents will attempt to camouflage their rituals. It is not uncommon for children with OCD to be able to inhibit their rituals for short periods of time to avoid detection (e.g., in class). Not infrequently, teachers, peers, and even parents remain unaware of the child's OCD for months or years, only becoming aware of the problem after the child can no longer control the symptoms or becomes too overwhelmed to cope.

DIFFERENTIAL DIAGNOSIS

Developmentally Normative Rituals and Superstitions

Benign habits, transitory superstitions, or mildly excessive thoughts are not uncommon during childhood. In certain phases of early childhood, rituals or superstitions may be utilized by a child to manage new or stressful situations (first day of preschool, separating from a parent) or achieving a sense of mastery. For example, a child may require a particular bedtime ritual or demand to wear a particular item of clothing. These normative rituals usually fade by the time of entry into formal schooling and are rarely associated with distress, significant interference with activities, or overlap in content with typical obsessions or compulsions. The absence of functional impairment differentiates mild, age-appropriate superstitions and habits from obsessive-compulsive symptoms.

Other Anxiety Disorders

Despite the presence of preoccupation and worry about issues related to their fears, children with other anxiety disorders typically do not engage in compulsive behaviors to reduce their anxiety. The content of the worry can also be useful in differentiating syndromes. Pediatric patients with generalized anxiety disorder (GAD) present with heightened anxiety relating to their health and safety, safety of others, and their future. The content of these symptoms, although exaggerated, is grounded in everyday matters (finances, health, grades, safety) as opposed to the unusual or even bizarre content of obsessive worries. Behavioral analogs to compulsions are rare besides the requests for reassurance from others to reduce anxiety. Patients with social anxiety have fears of negative evaluation and extreme scrutiny by others. Again, fears are by definition excessive, but typically not bizarre. Rituals are rare, although avoidance of feared situations is common. Similarly, the content of fears expressed by children with separation anxiety disorder are, by definition, confined to themes of separation or fears of harm to self or loved ones. Besides avoidance of separation, other behavioral features or rituals are absent from presentation.

Tic Disorders and Tourette Syndrome

Children and adolescents with tic disorders typically display phonic or motor tics that occur frequently and may appear behaviorally similar to compulsive actions, particularly more complex tics. Tics are differentiated from compulsions is usually being avolitional and are not preceded by precipitating obsessional thoughts. The content of tics is unlikely to be as purposeful as common compulsions (e.g., washing and arranging). It is noteworthy that there is substantial overlap between tic disorders and OCD in childhood, symptoms of each are not uncommon and both diagnoses may be warranted.

Autism Spectrum Disorders

The term autism spectrum disorder (ASD) comprises of children from a wide range of intellectual (and adaptive) functioning. Pervasive and stereotypic behaviors (including pragmatic language) are characteristic of ASD. In contrast with OCD, perseverations lack an association with anxiety; indeed, ASD perseverations or stereotyped behaviors may contain sought-after or pleasurable topics. Similarly, the stereotyped behavior may be elicited by excitement or frustration or may serve to manage stimulation and are generally not associated with preceding intrusive thought or impulse. Obsessions may be described but they are generally better characterized as idiosyncratic interests accompanied by marked perseveration and no awareness of the disinterest of others (or the excessiveness of their focus). Examples include ordering, arranging,

sequencing, and preservative speech. It is not uncommon for parents to report a child appearing "driven" to engage in his or her "compulsions," especially younger children. Although OCD can overlap with ASD (6 to 20 percent of patients with OCD are estimated to have ASD traits), distinction can usually be made by establishing the developmental context of the behaviors.

Psychotic Disorders

The differential diagnosis of OCD from incipient psychosis in youth can be challenging. The fixed nature of bizarre or unusual obsessions can have delusional qualities and resemble schizophrenic symptomology. Although OCD-related obsessions and compulsions may appear bizarre, insight is usually preserved and common themes are often evident. Children with OCD are typically aware that symptoms are products of their own mind (i.e., not thought insertion). OCD can usually be distinguished from schizophrenia by the absence of other symptoms, including formal thought disorder and hallucinations, and by the ability to engage in reality testing. In severe cases of OCD, reality testing may be challenged, and the boundary between obsession and delusion may be difficult to distinguish. OCD and psychotic disorders, though rarely, do co-occur.

Hypochondriasis

If the distressing thoughts are exclusively related to fears of having a serious medical condition, based on misinterpretation of bodily symptoms, hypochondriasis should be considered instead of OCD. Behaviorally, persistent need for medical evaluation and reassurance is not uncommon. However, if the preoccupation with illness is accompanied by rituals (e.g., washing, avoidance, checking based on the concern about contracting, exacerbating, or spreading illness to others), then a diagnosis of OCD may be appropriate.

COURSE AND PROGNOSIS

OCD with onset in childhood is a chronic condition characterized by substantial morbidity. Follow-up studies suggest that 43 to 68 percent of youth diagnoses with OCD continue to meet diagnostic criteria for the disorder up to 14 years after identification. Worse outcomes at follow-up were predicted by poor initial treatment response, a lifetime history of tic disorder, and a parental Axis I psychiatric disorder. The literature lacks long-term controlled treatment studies, and most follow-up data are from extant treatment trials following only a few months post-treatment. Further, this literature is complicated by multiple confounding variables, including treatment type, comorbidity, age or age of onset, and access to follow-up care.

The largest long-term follow-up study to date was based on 142 youth initially treated for OCD at the Maudsley Hospital and recontacted a mean of 9 years after intake. At follow-up, 41 percent of subjects still met diagnostic criteria for OCD with 4 percent reporting subclinical OCD symptoms. Although many reported functioning reasonably well, 50 percent were currently receiving treatment and 70 percent were diagnosed with any psychiatric disorder (following OCD, GAD, major depressive disorder [MDD], and tic disorders were most common). One other long-term retrospective follow-up study reported similar findings. Overall, despite reasonable success with acute treatments of pediatric OCD, the limited available long-term data suggest considerable room for improvement with regard to the durability of these interventions.

Even with remission following treatment, children with a history of OCD or subclinical OCD symptoms should continue to be monitored for relapse in symptoms. Likewise, relatives (siblings) of patients with OCD should be monitored closely and considered for treatment with the emergence of subclinical symptoms, even in the absence of more severe or impairing symptoms.

TREATMENT

Currently, there are two approaches supported in the research literature for pediatric OCD: CBT based on E/RP and pharmacotherapy with SRIs. A meta-analysis of randomized controlled trials (RCTs) for pediatric OCD yielded effect sizes of 1.45 (95 percent CI = 0.68 to 2.22) for CBT and 0.48 for pharmacotherapy (95 percent CI = 0.36 to 0.61), although methodological issues complicate direct comparison of efficacy for these two treatment approaches. Combined intervention with both CBT and SSRI might yield benefits beyond use of either of these treatments alone. As a result, the American Academy of Child and Adolescent Psychiatry treatment guidelines consider CBT or CBT + SSRI medication the first-line treatment of pediatric OCD. It is important to note, however, that behavioral approaches are overwhelmingly preferential to parents (75 percent prefer CBT only compared to CBT + SSRI with the disparity increasing among parents of children under age 9). Notably, a large multisite controlled study found that for youth who did not respond to an initial course of CBT with E/RP, extending E/RP was as effective as SSRI augmentation. This suggests that extending E/RP may be an alternative to SSRI augmentation in some cases. A recent Cochrane meta-analysis suggested CBT generally out performs SSRIs for pediatric OCD; Combined SSRI + CBT is favorable to SSRI alone. For those who had failed an SSRI, adding CBT was superior to continuing the SSRI. However, those who failed CBT did as well with continued CBT as with switching to an SSRI.

Cognitive-Behavioral Therapy with Exposure and Response Prevention

The efficacy of CBT in children has been demonstrated in numerous open trials and at least six controlled trials. Indeed, treatment response rates in the extant CBT trials are quite high, ranging from 57 to 90 percent. Unlike pharmacotherapies (for which relapse is not uncommon when medication is discontinued), treatment gains from CBT commonly endure more favorably after therapy is completed in contrast to medication discontinuation. In contrast, CBT with E/RP for OCD is distinguished from other "talk therapies" with no supported efficacy in the extant literature, including role-play–based, supportive, insight-oriented, relaxation, psychoanalytic, and psychodynamic therapies. Several treatment manuals outlining CBT with E/RP for children and adolescents are available. They consist of exposure (placing the patient in situations that elicit anxiety related to their obsessions), response prevention (deterring the compulsive or ritualistic behaviors that may serve to reduce or avoid anxiety from occurring), and cognitive therapy (training the patient to identify and reframe anxiety-provoking cognitions). The first step of treatment is psychoeducation about OCD and E/RP. It is important for the child and parents to understand the rationale for exposure to fear and restraint from anxiety-reducing compensatory behaviors. Next, a "fear hierarchy" is developed rank ordering symptoms. Following this, patients are systematically exposed to the feared situations (least feared first) while being instructed not to engage in their ritualistic behavior. Over repeated exposures, the associated anxiety dissipates through the process of autonomic habituation. In addition, when the feared consequences of not ritualizing fail to occur, the patient's heightened expectations of harm disappear, reducing anxiety even further.

Format and intensity of CBT with E/RP are variable. Studies suggest that weekly outpatient CBT for child OCD is equivalent in efficacy to intensive CBT. It is recommended that treatment of child OCD be conducted in vivo and in the context of the family in order to promote generalization. There is preliminary support for telehealth-based E/RP approaches but more studies are needed.

In addition to E/RP, adjunctive cognitive strategies are utilized with pediatric patients. However, younger children and those with impaired cognitive functioning receive fewer and less sophisticated cognitive components to treatment due to developmental and cognitive limitations. A child's age, cognitive functioning, and insight into the nature of his or her OCD is paramount in determining the direction of treatment, as introduction of cognitive components of therapy depends on the child's developmental level and insight.

Additionally, children's rituals frequently involve parental and family participation, and more severe family dysfunction has been shown to predict worse child treatment outcome. Moreover, E/RP requires significant out-of-session practice for mastery and extinction of OCD

symptoms. Many children are unwilling or unable to complete the necessary "homework" from therapy on their own (perhaps due to anxiety, lack of motivation, distraction). Thus, substantial parental involvement appears warranted, especially for younger children. Family treatment should include significant psychoeducation about OCD as well as mechanisms of E/RP. Parents may benefit from observing techniques presented by a skilled therapist so that they can assist their child with exercises in a supportive, noncoercive approach. Additionally, this can help parents to correctly extinguish ritualistic behavior and disengage from OCD rather than scold/react/blame. Teaching family members experiencing significant OCD-related conflict or discord skills to enhance affect modulation and problem-solving has been shown to enhance child treatment outcomes. Studies suggest family based CBT is effective for youth with OCD symptoms as young as age 3.

Pharmacotherapy

The efficacy of pharmacotherapy for OCD in pediatric populations has been demonstrated in several controlled trials with SRIs and SSRIs. At present, the U.S. Food and Drug Administration (FDA) approval has been granted for four medications for pediatric OCD use: clomipramine, fluoxetine, sertraline, and fluvoxamine, although the minimum approved age varies by compound. Table 52.1–3 lists medications and dose ranges for pediatric OCD.

Table 52.1–3.
Medications and Dose Ranges for Pediatric Obsessive-Compulsive Disorder

Medication	Typical Dosage (mg/day)
Fluoxetine (Prozac)	10–80[a]
Sertraline (Zoloft)	50–200[a]
Fluvoxamine (Luvox)	50–300[a]
Paroxetine (Paxil)[b]	
Citalopram (Celexa)	10–60[a]
Escitalopram (Lexapro)	5–20
Clomipramine (Anafranil)	50–200[a]

[a]Mean doses as per the American Academy of Clinical Child and Adolescent Psychiatry 2012 Practice Parameters, based on randomized controlled trials.

[b]Paroxetine is not recommended for use in children/adolescents.

Clomipramine, a tricyclic antidepressant, is the most researched SRI in the treatment of pediatric OCD. In a double-blind, 8-week, placebo-controlled study, 60 percent of pediatric patients showed significant improvement. Patients treated with clomipramine reported a 37 percent

mean reduction in OCD symptoms compared to 8 percent for the placebo group (with an effect size approaching 1.0). In another 10-week controlled trial, a significant difference was found between clomipramine and placebo, with 75 percent of pediatric patients showing at least moderate improvement. Other research found that clomipramine was superior to the noradrenergic reuptake inhibiting tricyclic antidepressant desipramine. This crossover trial found that 64 percent of patients who initially received clomipramine during their first treatment showed relapse of OCD symptoms during desipramine treatment. Overall, a 2003 meta-analysis of pharmacotherapy trials in children identified clomipramine to be significantly superior over SSRIs, which did not differ among one another, in reducing OCD symptoms. Nevertheless, the risk profile, adverse effects, and required electrocardiogram and blood-level monitoring associated with tricyclic antidepressants (e.g., antiadrenergic, anticholinergic, and antihistaminergic adverse effects) are of concern with clomipramine. Thus, clomipramine is relegated to a second- or third-line treatment choice for children and adolescents and should not be used until sufficient trials of CBT with E/RP, two or more SSRIs, and combination CBT with SSRI therapy have been attempted.

Selective Serotonin Reuptake Inhibitors

In a 20-week, double-blind, placebo-controlled trial of the SSRI fluoxetine in children and adolescents with OCD, 44 percent reductions in OCD symptoms were reported. A 13-week controlled trial also demonstrated the efficacy of fluoxetine, with 55 percent of patients treated with fluoxetine rated as much or very much improved. Another 16-week, placebo-controlled trial of fluoxetine in children reported that 57 percent of patients demonstrated significant improvement. Several open trials also present favorable findings for the use of fluoxetine for pediatric OCD.

Data also support the use of the SSRI sertraline for the treatment of pediatric OCD. A 12-week, multicenter, randomized, placebo-controlled trial of children and adolescents with OCD found that 42 percent of patients receiving sertraline were rated as much or very much improved. Further, in a 52-week, open-label extension of the previous study, 71 percent of children (aged 6 to 12) and 61 percent of adolescents (aged 13 to 18) demonstrated 25 percent decreases in OCD symptoms and were rated as much or very much improved. Another multisite, randomized controlled trial found that 21.4 percent of pediatric patients entered clinical remission following a 12-week course of sertraline.

A 10-week, multicenter, placebo-controlled trial of fluvoxamine also yielded a 42 percent response rate along with evidence for the safety and 1-year durability of this treatment. Citalopram and escitalopram are also used for pediatric OCD, with a putative mechanism of action similar to sertraline and fluoxetine, but RCT data is scant. One pilot randomized

study of only 29 youth found no group differences in efficacy or safety between fixed dose (20 mg) citalopram and fluoxetine at 6 weeks. Although paroxetine has also been studied for use in pediatric OCD, an FDA advisory recommending that paroxetine not be used in the treatment of MDD, due to increased risk of suicidal ideation, has largely curtailed the use of this treatment in children and adolescents.

Overall, clinically significant reductions in OCD symptomology have been documented in children and adolescents using SSRIs with a relatively minimal side-effect profile (as compared to clomipramine). The most frequent SSRI side effects include nausea, insomnia, headache, and asthenia. During clinical trials, the inability to tolerate treatment due to side effects was reported in 15 percent or less of cases. Based on these findings, SSRIs are the consensus first-line medication for pediatric OCD. Although there are very limited controlled comparisons between these medications in children, research suggests that the SSRIs are equally efficacious in children and the specific choice should be based on the patient's medical history, concomitant medications, and the adverse events. Sertraline and fluoxetine have been studied more extensively in youth, suggesting possible initial options. Poor clinical response to one SSRI is not necessarily predictive of failure with other SSRIs, suggesting adequate trials of multiple SSRIs may be indicated before augmentation. Additionally, significant clinical response is unlikely within the first few weeks of taking an SSRI; generally, 10 to 12 weeks at adequate dosage is necessary to fully evaluate the efficacy of the medication.

In 2005, the FDA issued a black box warning for antidepressant medications to alert health care providers to an increased risk of suicidality (suicidal thinking and behavior) in children and adolescents being treated with these agents. In light of this warning, increased attention to symptoms of depression among pediatric patients being treated with antidepressant medications appears warranted. This should include formal assessment of suicidal ideation prior to and throughout the course of pharmacological treatment. An implication of this warning, combined with strong research support for CBT with E/RP, suggests attempting an adequate trial of the CBT prior to pharmacological intervention. Notably, a recent large-scale Scandinavian study suggested that for CBT nonresponders continued CBT was as effective in comparison to SSRI augmentation.

Other Treatments

In pediatric cases that are unresponsive to CBT and trials with multiple SSRIs, pharmacological augmentation may be considered. There is growing support for SSRI augmentation with atypical antipsychotics for refractory OCD in adult patients. For example, an 8-week, double-blind, placebo-controlled trial was performed to determine the efficacy of risperidone augmentation in adults with treatment-resistant OCD found a 40 percent

response rate. A preliminary open label study of aripiprazole (mean dose 4.75 mg/day) 16 children (non-responders to CBT and at least two SSRIs) with OCD (80% with comorbidity) suggested benefit after 12 weeks. However, replication is needed in a larger, randomized-controlled design. Other data suggest aripiprazole augmentation of SSRIs may be beneficial for treatment resistant youth (failed multiple SSRIs and CBT). Additionally, two recent meta-analyses of off-label use of neuroleptics (including haloperidol, risperidone, olanzapine, and quetiapine) in adults with OCD suggest that initial findings from open trials, controlled trials, and case series are promising despite a need for substantial further research, especially in youth. Lithium carbonate and buspirone augmentation have generally not been efficacious augmentation strategies for treatment-refractory OCD in adults but may ameliorate depressive symptoms. Thus, augmentation may also be the case in children with comorbid conditions, for example, a tic disorder may require the addition of an α_2-agonist or neuroleptics.

In sum, unlike the adult literature, there is a paucity of empirical data supporting augmentation strategies for pediatric OCD. Additionally, significant undesirable side effects should be considered prior to prescribing neuroleptics and benzodiazepines in children.

Ace is a 15-year-old male of European ancestry from a southeastern suburban community. He lives with his parents and two younger sisters. Ace originally presented to his primary care physician in the fall following an exacerbation in obsessive-compulsive symptoms. Parents described their son as a worrier, even as a young child. He would worry about his own performance, health, and safety. He was always questioning adults to ascertain, “Am I doing this right?” or “You’re not angry at me?” In addition to questioning, Ace would always check over his homework, chores, and personal details (e.g., appearance) very carefully. However, approximately 6 months ago, more unusual and atypical behaviors developed. Ace would rewrite homework assignments six or seven times, shredding papers if he had to erase the slightest mistake. He would change his clothes several times per day, stating that items “got dirty,” when there was no evidence for any need of cleaning. His parents noticed a pile of his previously favorite, sports jerseys and expensive athletic shoes stashed in the attic. When questioned, Ace claimed they were “contaminated” and could not be worn again, despite any indication of soiling. He became late for school and activities, locking himself in his room for long periods before leaving the house. He was evasive when questioned about this behavior, but stated, “There are things I need to do to make sure things will be okay.” He latter clarified that he would repeat a sequence of prayers and sayings in his head until “they felt just right” and he could subsequently go out of the house. Whereas Ace’s ninth grade teacher had tolerated tardies and late assignments given his history of being a stellar student, his high school teachers did not accept late work and reprimanded him for being late. Consequently, his grades were plummeting. At the same time, his extracurricular interests appeared to be diminished, and rituals replaced the time he had spent on age-appropriate activities. Fights were erupting between Ace and his parents because Ace’s rituals made his mother late for work, and both parents grew tired of their son’s increased need for reassurance.

There was no history of OCD in Ace’s family. However, his father had a history of untreated “anxiety” and Ace’s 7-year-old brother had been diagnosed with Tourette syndrome. Ace’s problems had never been mentioned to a health care professional before. His parents felt they were just idiosyncratic aspects of his personality. Moreover, he had been functioning at a high level and there was no evidence of distress, except when his parents were not available to provide reassurance. They presented for treatment given Ace’s increased distress and

impairment. It was recommended by his pediatrician that Ace begin 20 mg of fluoxetine. Ace's parents met with a child and adolescent psychiatrist for additional guidance. At that visit, Ace's parents discussed reluctance to start medication, especially given the warnings that they had heard that "these drugs can make my son suicidal," and had read that many children can show remission with E/RP alone. Consequently, the psychiatrist referred the family to a colleague in his practice—a board certified child and adolescent psychologist with expertise in child OCD. Ace and his family presented for initial consolation with the clinical psychologist who agreed CBT with E/RP was indicated and she reviewed the treatment model with the family. Ace and his family participated in weekly sessions for about 12 weeks, focusing on exposure of his feared thoughts and prevention of ritualistic behavior in a systematic method. Therapy was family-based, emphasizing family-based in vivo practice of E/RP, generalization of exercises outside of session, and minimizing disruption to school. A hierarchy of feared symptoms was developed at visit two, and in vivo E/RP commenced at visit three.

After about 3 months of CBT with E/RP, Ace reported an improved ability to manage his obsessive thoughts and to resist rituals. He reported decreased anxiety as well as urges to ritualize. His parents reported an improved understanding of OCD as well as how to react to his reassurance seeking (and other compulsive behaviors) in order to facilitate his recovery. The family understood the algorithm for intervening with future obsessions or compulsions as well as the potential need to resume CBT if symptoms resurged. The family chose to remain with the psychologist for a few additional sessions to address family conflict that had developed surrounding the OCD behaviors.

References

AACAP. Practice parameter for the assessment and treatment of children and adolescents with obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry*. 2012;51:98–113.

Ardic UA, Ercan ES, Kutlu A, Yuce D, Ipci M, Inci SB. Successful treatment response with aripiprazole augmentation of SSRIs in refractory obsessive–compulsive disorder in childhood. *Child Psychiatry Hum Dev*. 2016. doi:10.1007/s10578-016-0694–8

Anticevic A, Hu S, Zhang S, et al. Global resting-state functional magnetic resonance imaging analysis identifies frontal cortex, striatal, and cerebellar dysconnectivity in obsessive-compulsive disorder. *Biol Psychiatry*. 2014;75(8):595–605.

Browne HA, Hansen SN, Buxbaum JD, et al. Familial clustering of tic disorders and obsessive-compulsive disorder. *JAMA Psychiatry*. 2015;72(4):359–366.

Douglass HM, Moffitt TE, Dar R, McGee R, Silva P. Obsessive-compulsive disorder in a birth cohort of 18-year-olds: prevalence and predictors. *J Am Acad Child Adolesc Psychiatry*. 1995;34:1424–1431.

Ercan ES, Ardic UA, Ercan E, Yuce D, Durak Sl. A promising preliminary study of aripiprazole for treatment-resistant childhood obsessive-compulsive disorder. *J Child Adolesc Psychopharmacol*. 2015;25(7):580–584.

Fitzgerald KD, Liu Y, Reamer EN, Taylor SF, Welsh RC. Atypical frontal-striatal-thalamic circuit white matter development in pediatric obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry*. 2014;53(11):1225–1233.

Fitzgerald KD, Welsh RC, Stern ER, et al. Developmental alterations of frontal-striatal-thalamic connectivity in obsessive compulsive disorder. *J Am Acad Child Adolesc Psychiatry*. 2011;50(9):938–948.e3.

Franklin ME, Sapyta J, Freeman JB, et al. Cognitive behavior therapy augmentation of pharmacotherapy in pediatric obsessive-compulsive disorder: the pediatric OCD treatment study II (POTS II) randomized controlled trial. *JAMA*. 2011;306(11):1224–1232.

Garcia AM, Sapyta JJ, Moore PS, et al. Predictors and moderators of treatment outcome in the pediatric obsessive compulsive treatment study (POTS I). *J Am Acad Child Adolesc Psychiatry*. 2010;49(10):1024–1033; quiz 1086.

Grant PJ, Joseph LA, Farmer CA, et al. 12-week, placebo-controlled trial of add-on riluzole in the treatment of childhood-onset obsessive-compulsive disorder. *Neuropsychopharmacology*. 2014;39(6):1453–1459.

Ivarsson T, Skarphedinsson G, Kornør H, et al; Accreditation Task Force of The Canadian Institute for Obsessive Compulsive Disorders. The place of and evidence for serotonin reuptake inhibitors (SRIs) for obsessive compulsive disorder (OCD) in children and adolescents: views based on a systematic review and meta-analysis. *Psychiatry Res*. 2015;227(1):93–103.

Kircanski K, Peris TS, Piacentini JC. Cognitive-behavioral therapy for obsessive-compulsive disorder in children and adolescents. *Child Adolesc Psychiatr Clin North Am*. 2011;20(2):239–254.

Lewin AB, McGuire JF, Murphy TK, Storch EA. The importance of considering parent's preferences when planning treatment for their children—the case of childhood obsessive-compulsive disorder. *J Child Psychol Psychiatry*. 2014;55(12):1314–1316.

Lewin AB, Park JM, Jones AM, et al. Family-based exposure and response prevention therapy for preschool-aged children with obsessive-compulsive disorder: a pilot randomized controlled trial. *Behav Res Ther*. 2014;56:30–38.

Lewin AB, Wu MS, McGuire JF, Storch EA. Cognitive behavior therapy for obsessive-compulsive and related disorders. *Psychiatr Clin North Am*. 2014;37(3):415–445.

Micali M, Heyman I, Perez M, et al. Long-term outcomes of obsessive–compulsive disorder: follow-up of 142 children and adolescents. *Br J Psychiatry*. 2010;197(2):128–134.

Murphy TK, Gerardi DM, Leckman JF. Pediatric acute-onset neuropsychiatric syndrome. *Psychiatr Clin North Am*. 2014;37(3):353–374.

Pediatric OCD Treatment Study Team [POTS]. Cognitive-behavior therapy, sertraline, and their combination for children and adolescents with obsessive-compulsive disorder: the Pediatric OCD Treatment Study (POTS) randomized controlled trial. *JAMA*. 2004;292:1969–1976.

Piacentini J, Bergman RL, Chang S, et al. Controlled comparison of family cognitive behavioral therapy and psychoeducation/relaxation training for child obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry*. 2011;50(11):1149–1161.

Skarphedinsson G, Weidle B, Thomsen PH, et al. Continued cognitive-behavior therapy versus sertraline for children and adolescents with obsessive-compulsive disorder that were non-responders to cognitive-behavior therapy: a randomized controlled trial. *Eur Child Adolesc Psychiatry*. 2015;24(5):591–602.

Stewart SE, Hezel D, Stachon AC. Assessment and medication management of paediatric obsessive-compulsive disorder. *Drugs*. 2012;72(7):881–893.

Stewart SE, Yu D, Scharf JM, et al. Genome-wide association study of obsessive-compulsive disorder. *Mol Psychiatry*. 2013;18(7):788–798.

Torp NC, Dahl K, Skarphedinsson G, et al. Effectiveness of cognitive behavior treatment for pediatric obsessive-compulsive disorder: acute outcomes from the Nordic Long-term OCD Treatment Study (NordLOTS). *Behav Res Ther*. 2015;64:15–23.

Torp NC, Dahl K, Skarphedinsson G, et al. Predictors associated with improved cognitive-

behavioral therapy outcome in pediatric obsessive-compulsive disorder. J Am Acad Child Adolesc Psychiatry. 2015;54(3):200–207.e1.

van den Heuvel OA, Veltman DJ, Groenewegen HJ, et al. Frontal-striatal dysfunction during planning in obsessive-compulsive disorder. *Arch Gen Psychiatry*. 2005;62(3):301–309.

Varigonda AL, Jakubovski E, Bloch MH. Systematic review and meta-analysis: early treatment responses of selective serotonin reuptake inhibitors and clomipramine in pediatric obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry*. 2016;55(10):851–859.e2.

Watson HJ, Rees CS. Meta-analysis of randomized, controlled treatment trials for pediatric obsessive-compulsive disorder. *J Child Psychol Psychiatry*. 2008;49:489–498.

▲ 52.2 Posttraumatic Stress Disorder in Children and Adolescents

JUDITH A. COHEN, M.D.

Two-thirds of American youth America experience at least one potentially traumatic event (PTE, e.g., child abuse, domestic or community violence, disaster, accidents, medical trauma, war, or multiple/complex trauma). A third experience multiple traumas, increasing risk of serious mental, and medical health problems. Up to 10 percent develop full-blown posttraumatic stress disorder (PTSD); many more suffer from subsyndrome PTSD symptoms with equivalent functional impairments. New criteria for PTSD in DSM-5 increase attention to developmental and cultural manifestations of trauma symptoms. Genetic and psychobiological factors also influence the development of PTSD. Recognition of the frequency with which children develop PTSD symptoms has led to improved strategies for assessing and treating traumatized children.

DEFINITION

PTSD is defined by developing characteristic features after exposure to trauma, that is, events involving threat to life, safety or physical integrity, experienced or witnessed directly by the child, or affecting someone close to the child (criterion A). The clinical presentation varies and not all children manifest emotional fear, helplessness, or horror as previously required. Children ages 7 years or older must have one intrusion symptom such as recurrent trauma-related memories or dreams, or intense psychological of physiological response to trauma reminders (criterion B); one symptom of persistent avoidance of trauma-related stimuli (criterion C); two symptoms of negative alterations in cognitions such as persistent negative trauma-related beliefs or persistent posttrauma negative emotional state (criterion D); and two arousal symptoms (criterion E). The revised criteria include developmentally sensitive diagnostic criteria for children 6 years and younger. These children must have one intrusion symptom (criterion B); one symptom of avoidance or negative alteration in

cognitions (criterion C); and two arousal symptoms (criterion D). The duration of symptoms must be more than 1 month, and the disturbance must cause clinically significant distress or functional impairment.

HISTORY

Descriptive accounts of children traumatized by the London blitz during World War II, of the kidnapping and subsequent burial of children in their school bus, and of children exposed to a fatal schoolyard sniper attack were published by early child trauma pioneers. Subsequent empirical studies of children exposed to sexual abuse, peer suicide, cancer, burns, and terrorist attacks confirmed that increased levels of trauma exposure and lack of adequate social support strongly predict the emergence of PTSD and other psychological symptoms. The September 11, 2001 terrorist attacks, the Asian tsunami of 2004, and Hurricane Katrina have documented that peritraumatic panic, delayed evacuation, and perceived life threat to self or family members are significant risk factors for later PTSD development. In the past 20 years, a number of empirically rigorous randomized controlled trials (RCTs) have documented the superiority of trauma-focused psychosocial treatments over usual care for pediatric PTSD. Few pharmacological RCTs have examined the treatment of pediatric PTSD.

COMPARATIVE NOSOLOGY

PTSD was first officially recognized in 1980, in DSM-III. DSM-III-R was the first version to include representative items for childhood PTSD. Importantly for children, this version increased the number of avoidance or numbing symptoms needed to meet diagnostic criteria from one to three. DSM-IV further elucidated childhood variants of symptomatology. The development of an alternative algorithm for diagnosing PTSD in preschool children and growing research led to the new diagnostic criteria in DSM-5 for children younger than 7 years old. Research documenting that many do not respond with horror, fear, or helplessness supported the removal of this criterion in DSM-5. Other research led to including a new cluster for alterations of mood and cognitions in DSM-5.

EPIDEMIOLOGY

Surveys document PTSD prevalence rates in the general youth population of 8 to 9 percent with another 14.1 percent reaching subclinical levels. More relevant are prevalence rates among trauma-exposed children, as by definition PTSD is only possible in this population. Nonreferred samples of adolescents exposed to various forms of trauma have documented that 25 to over 90 percent develop PTSD, depending on the trauma experienced, time since exposure, and assessment method used. Chronic traumas, such as ongoing child abuse, or traumas that result in disruption of family and

the broader social support network, such as war or refugee experiences, tend to result in greater risk of PTSD than acute traumas, such as natural disasters. Exposure to multiple traumatic events is associated with an increase in PTSD. Female gender also tends to predict greater risk of developing PTSD as well as other mood and anxiety disorders. Furthermore, the above rates only include children meeting full diagnostic criteria; several studies have indicated that the majority of trauma-exposed children develop PTSD symptoms severe enough to interfere with functioning, even in the absence of a PTSD diagnosis.

ETIOLOGY

PTSD is one of the few DSM disorders that require the presence of an etiological agent (a serious traumatic event). Exposure factors also predict outcome with severity and frequency of trauma exposure, exposure to interpersonal violence, greater perceived life threat, death of a family member, greater peritraumatic panic, and delayed evacuation (if during a mass trauma experience) all predicting higher risk of developing PTSD. Nevertheless, not all children exposed to such events develop PTSD, so other etiological factors must also be considered.

Biological Factors

Female gender is a risk factor for developing PTSD. Children with pre-existing anxiety disorders are at increased risk to develop PTSD following trauma exposure, and a history of depression in first-degree relatives also predicts increased vulnerability to developing PTSD. Prior trauma exposure is a risk factor for developing PTSD upon exposure to a new trauma, with greater number of traumas predicting greater risk of disorder. Recent studies have demonstrated that genetic risk factors interact with other risk and protective factors, such as social support, in determining whether trauma-exposed children will develop PTSD, or will be resilient despite traumatic experiences. Numerous psychobiological abnormalities have been found in children with PTSD, including complex alterations in hypothalamic–pituitary–adrenal (HPA) axis functioning, increased excretion of adrenergic and dopaminergic metabolites, smaller intracranial volume and corpus callosum area, as well as lower intelligence quotient (IQ) and memory deficits. In addition to these alterations, adults with PTSD have evidence of amygdala hyperactivity, medial prefrontal cortex underactivity, and decreased hippocampal volume. However, because of the paucity of prospective studies, it is unclear whether these findings are indicative of etiological (causal) factors in the development of PTSD or the effects of PTSD on children's developing brains.

Psychological Factors

Classical and operant conditioning are involved in the development of PTSD. Extreme stressors typically produce negative emotions (terror, anger, sadness) as well as physiological arousal from the activation of the sympathetic nervous system (the "fight, flight, or freeze" response). Classical conditioning occurs when neutral cues that were present at the time of the trauma become associated or "paired" with the trauma to the extent that they too elicit these intense psychological and physiological reactions. Operant conditioning occurs when the child learns to avoid trauma-related cues in order to reduce the likelihood of experiencing these distressing reactions. As avoidance behaviors are repeatedly reinforced, they may also generalize. New responses can be acquired through pairing of exposure and inhibitory learning.

Children's maladaptive cognitions are also relevant to PTSD development. Shame, self-blame, believing that one is undeserving of being loved, protected or being treated well, or believing that others cannot be trusted are common maladaptive cognitions among children who experience interpersonal trauma. Such children may have difficulty differentiating between what was done to them (the traumatic experiences) and who they are (their identity). This has important implications for how such children's trauma symptoms are manifested (i.e., in assessing complex PTSD symptoms as described below).

Social Factors

Many pretraumatic social factors including lower socioeconomic status, economic deprivation, family dysfunction, parental separation or death, lower intelligence, minority racial or ethnic status, and family psychiatric history have been found to predict higher risk of developing PTSD. Absence of a strong social support network has been found to increase the risk of developing PTSD. Parental psychopathology has also been implicated as a risk factor in children developing PTSD following exposure to a variety of traumatic events. Thus, there is growing evidence that social factors may have a role in the etiology of this disorder. However, a recent study demonstrated that twins who had been sexually abused were at significantly greater risk of developing a variety of adverse outcomes than their nonabused twins, despite shared genetics and family environment. Social etiological factors should therefore not be overestimated.

DIAGNOSIS AND CLINICAL FEATURES

Exposure to a Traumatic Event

To qualify for a diagnosis of PTSD, the child must have been exposed to a traumatic event that consisted of direct personal experience; witnessing an event in person; learning about the trauma of a close family member or close friend (in the case of death, the event must have been violent or

accidental); or experiencing repeated or extreme exposure to aversive details of the traumatic events (but not only through media descriptions). A number of qualifying traumatic events are listed in DSM-5, which specifies that in the case of child abuse, there does not have to be threat physical violence or injury (e.g., genital fondling without penetration would qualify).

Re-Experiencing

The child must have at least one persistent intrusion symptoms related to the traumatic event. Children may have recurrent, involuntary, and intrusive thoughts, memories, images, or body sensations reminiscent of the traumatic event. In young children, this may take the form of repetitive play in which traumatic themes are present (e.g., sexually abused children may repeatedly re-enact inappropriate touching in their play). Children may have recurrent distressing dreams about the event, or in the case of younger children, there may be frequent frightening dreams without any recognizable content. Children may have dissociative reactions; these may include periods when they act or feel as if the event were reoccurring in the present moment (a flashback). In young children, trauma-specific re-enactment may occur (e.g., sexually abused children may insert objects into their own or other children's private parts or otherwise recreate abusive acts perpetrated against them). Children may experience intense psychological distress or physiological reactivity when exposed to triggering events, people, or situations that remind them of the original traumatic event (e.g., children who witnessed a drive-by shooting may become agitated or have difficulty breathing when they hear thunder or a backfiring car). These reminders may be highly idiosyncratic to individual children.

The above criteria also apply to diagnose PTSD for children 6 years or younger, but in these children, spontaneous and intrusive memories might not appear distressing. It may not be possible to ascertain that the frightening content of dreams is related to the traumatic event.

Avoidance and Numbing

The child must experience persistent avoidance of stimuli associated with the traumatic event and numbing of general responsiveness, as indicated by at one of the following symptoms. Children may avoid or make active efforts to avoid memories, thoughts or feelings about or closely associated with the traumatic event. They may avoid or make efforts to avoid external reminders (e.g., discussions about the traumatic event or people, places, situations, or activities that remind them of the trauma). For example, after the sudden death of a loved one, children may refuse to say the name of the deceased or insist that others not do so.

In order to diagnose PTSD for children 6 years and younger, one or

more symptom representing *either* persistent avoidance of trauma-related stimuli *or* negative alterations in cognitions and mood associated with the traumatic event must be present.

Negative Alterations in Cognition and Mood

Children may be unable to recall one or more important aspect of the traumatic event (dissociative amnesia). Children may have persistent overly negative beliefs about themselves (e.g., "No one loves me"); others (e.g., "I can't trust anyone") or the world (e.g., "Nothing is safe"); or maladaptive cognitions about why the trauma happened (e.g., "I should have known he would hurt me") leading to distorted self- or other-blame. Children may develop negative emotional states such as shame, guilt, anger, sadness, or fear that persist after the traumatic event. Children may have diminished interest or participation in their normal activities, for example, not wanting to play with friends (social withdrawal) or participating less in the classroom. They may feel detached or estranged from others (e.g., feel that no one understands them; affiliate with deviant peers) or may experience diminished capacity to feel emotions, especially warm interpersonal emotions (e.g., children may say, "I don't feel like I love anyone anymore.").

In order to diagnose PTSD for children 6 years and younger, one or more symptom representing *either* persistent avoidance of trauma-related stimuli *or* negative alterations in cognitions and mood associated with the traumatic event must be present.

Arousal and Reactivity

The child must also experience persistent symptoms of arousal and reactivity that were not present prior to the traumatic event, as evidenced by at least two of the following. Children may have increased irritability and angry outbursts, with little or no provocation; these may be expressed as aggression toward others (and are often misdiagnosed in children as externalizing behavior disorders). The may have reckless or self-destructive behaviors (e.g., self-injury, substance abuse, reckless driving, or unprotected sexual activity). They may exhibit hypervigilance (e.g., children may ask their parents to check several times every night to be sure that all the doors are locked) or exaggerated startle reaction (e.g., jumping whenever the telephone rings or a pencil drops in the classroom). Children may have difficulty concentrating or completing tasks, which may impair academic performance. Children may have trouble falling or staying asleep, or may have restless sleep, whether or not they are having trauma-related dreams.

The above criteria also apply for children 6 years and younger except that these children do not typically exhibit reckless or self-destructive

behaviors.

Duration, Types, and Clinical Features

These symptoms must be present for at least 1 month and must be accompanied by significant distress or impairment in an important area of functioning. PTSD with dissociative symptoms should be specified if the child meets the above criteria and experiences persistent or recurrent depersonalization or derealization symptoms that are not attributable to substance abuse. If the onset of symptoms does not occur until more than 6 months after the traumatic event, the diagnosis is PTSD with delayed expression. "Complex PTSD," "complex trauma," and developmental trauma disorder have been suggested to represent the complexity of symptoms existing in children experiencing repeated, interpersonal traumas involving attachment disruption. Such children have severe dysregulation of affect, behavior, cognitive, attachment, perceptual, and biological domains that result in marked functional impairment. It is not clear whether the new DSM-5 PTSD criteria will more adequately describe these complex problems or whether a new diagnosis will be required.

PATHOLOGY AND LABORATORY EXAMINATION

Although several alterations in laboratory, neuropsychological, and neuroimaging tests have been reported in children with PTSD, there is currently no laboratory or other test that can reliably distinguish children with PTSD from other children on an individual basis.

DIFFERENTIAL DIAGNOSIS

Assessment of pediatric PTSD is complex and beyond the scope of this article. PTSD symptoms overlap and co-occur with several other psychiatric conditions. Acute stress disorder (ASD) occurs and is resolved within the first month after traumatic exposure. Children with less significant trauma-related adjustment symptoms may be diagnosed with an adjustment disorder. Several diagnostic criteria of PTSD overlap with those of anxiety disorders and major depressive disorder. For example, recurrent intrusive thoughts are a hallmark of obsessive–compulsive disorder, while avoidance of social situations, social withdrawal, and tantrums is characteristic of social phobia. The specific content of these thoughts or fears may clarify that they are not trauma focused. However, because children frequently do not disclose traumatic experiences unless specifically asked, it is important for the assessing clinician to inquire directly about traumatic exposure whenever children present with anxiety-related symptoms.

In a similar manner, irritability, feelings of emptiness, decreased interest in normal activities, sleep disturbance, psychomotor agitation, inappropriate guilt, and decreased concentration are all symptoms shared

by PTSD and major depressive episode. Distinguishing between these two disorders may be particularly difficult in children with avoidance symptoms that prevent them from discussing the traumatic event. Careful history taking is essential in determining whether the child's symptoms are trauma related and to what extent the child meets criteria for either or both diagnoses. Children with traumatic bereavement are particularly likely to meet partial or full criteria for both disorders.

Children and adolescents presenting with substance use disorder may have a past history of trauma that is either covert (never disclosed) or minimized by the child or assessing clinician. Accurately identifying PTSD in such children may be critical to their optimal treatment, as standard substance use interventions do not contain important elements of trauma-focused treatment. If such children are using drugs or alcohol to self-medicate trauma-related symptoms, it is unlikely that their substance use will improve without adequately addressing these symptoms.

Additionally, PTSD may coexist with or be misdiagnosed as behavioral disorders such as attention-deficit/hyperactivity disorder (ADHD) or conduct disorder. Particularly in younger children who are less able to verbalize their anxiety, PTSD-related irritability, sleep difficulties, and difficulty concentrating may present as classic ADHD. Eliciting a history of trauma exposure and evaluating the timing of these symptoms may be important in distinguishing these disorders.

PTSD can mimic medical conditions such as hyperthyroidism, seizure disorder, caffeinism, migraine, asthma, or side effects from prescription or nonprescription medications.

Symptoms of complex PTSD (severe affective, behavioral, relationship, and cognitive dysregulation) may resemble those of bipolar disorder. Particularly if such children are avoidant of disclosing a trauma history, they will likely be diagnosed with bipolar disorder rather than PTSD. Since first-line treatments of these disorders are very different (medications for the former; psychotherapy for the latter), it is crucial to always inquire about children's trauma histories and to assess their symptoms from a trauma-informed neurodevelopmental perspective.

Psychotic symptoms may occur in up to 20 percent of children with PTSD; flashbacks or other intrusive phenomena must be distinguished from true hallucinations or substance-induced disorders.

COURSE AND PROGNOSIS

The natural course of childhood PTSD has not been adequately determined. Based on the available studies, it appears that while PTSD symptoms spontaneously remit in some children, they persist for years or even decades in many others. There is also growing evidence that trauma-focused treatment significantly alters the natural course of pediatric PTSD. Studies have documented continued remission of PTSD 2 years

posttreatment in children who received trauma-focused cognitive-behavioral therapy (TF-CBT).

The prognosis for children with untreated PTSD has become a matter of increasing concern since research has documented the numerous psychobiological abnormalities associated with this disorder, and the severity of some of these (decreased intracranial volume, decreased corpus callosum area, and lowered IQ) is directly correlated with the duration of symptoms. Growing evidence that childhood sexual abuse is predictive of adult suicidality as well as depression and suicidality in the offspring of such victims also suggests a poor prognosis for unresolved PTSD in childhood. This emphasizes the importance of identifying symptomatic children and providing effective interventions early in the course of this disorder.

Treatment

Children with PTSD receive many treatments, but the effectiveness of most of these interventions has not been evaluated. To date trauma-focused psychotherapies have the most empirical evidence of efficacy for treating pediatric PTSD. Although several selective serotonin reuptake inhibitors (SSRIs) have been shown to effectively treat adult PTSD symptoms, two RCTs have failed to show superiority to placebo in traumatized children. Similarly, although open trials of adrenergic agents have had promising results, two RCTs have failed to show superiority of propranolol to placebo in treating pediatric PTSD.

Psychotherapeutic Treatments

Cognitive-Behavioral Therapies. Currently TF-CBTs have the strongest evidence for resolving child PTSD and associated problems. TF-CBT has been found in 15 RCT to be superior to other active treatments or comparison conditions in decreasing PTSD as well as depressive, anxiety, cognitive, behavioral symptoms, and functional impairment in traumatized children. TF-CBT has been tested in youth who have experienced sexual trauma, domestic violence, terrorism, disaster, war, multiple and complex trauma, in multicultural cohorts in the United States, Europe, Africa, and Australia. TF-CBT is typically provided over 12 to 25 treatment sessions. Its core components are summarized by the acronym PRACTICE: *Psychoeducation* about the child's traumatic experiences and responses; *Parental or caregiver inclusion* in treatment, including a parenting skills component; *Relaxation skills,* which are individualized for each child and parent; *Affective* modulation skills; *Cognitive* coping techniques (connections between thoughts, feelings, and behaviors are taught and practiced); *Trauma narrative* and cognitive processing, consisting of carefully calibrated efforts to encourage the child to describe increasing

details about the traumatic event and to identify and process inaccurate or unhelpful thoughts related to the traumatic event; In vivo mastery of trauma reminders; *Conjoint* child–parents sessions in which optimal communication is emphasized and practiced; and *Enhancing safety* and future developmental trajectory, in which the child develops and practices effective strategies for maintaining personal safety. A group CBT model for school settings, Cognitive Behavioral Intervention for Trauma in Schools, has been tested in schools and found to be effective in reducing PTSD and depressive symptoms. Other CBT-based treatments such as the UCLA Trauma/Grief Program for Adolescents and a child adaptation of Narrative Exposure Therapy have also been tested with positive results among refugee or war-exposed children.

Eye movement desensitization and reprocessing (EMDR) utilizes exposure and cognitive reprocessing interventions with bilateral eye movements. EMDR has been adapted for children and tested in two well-designed studies. A recent study comparing EMDR to TF-CBT found that both treatments were equally effective and efficient at improving children's PTSD symptoms, with TF-CBT being superior for improving comorbid psychiatric symptoms.

Psychodynamic and Attachment-Based Therapies. Child–parent psychotherapy (CPP) is an attachment-based treatment of trauma-exposed infants and preschool children and their primary caretakers. CPP has been tested in one RCT for preschoolers exposed to marital violence and found to be superior to case management in improving these young children's PTSD symptoms and in their mothers' PTSD symptoms. CPP includes the following components: Providers begin by providing developmental guidance or psychoeducation about the child's behavior and how it may be understood in light of normal development and changed through simple behavioral interventions. If needed, additional interventions are added to help the parent to understand the child's behaviors and enhance child–parent attunement. These may include behavior management, cognitive reframing and processing, developing a joint narrative of the child and parent's traumatic experiences, and interpreting child or parent behaviors, emotions, and past experiences.

Psychopharmacological Treatments

RCTs of sertraline (Zoloft), D-cycloserine, and propranolol (Inderol) have failed to show significant differences from placebo in improving pediatric PTSD. While open treatment trials have demonstrated clinical improvement in child PTSD symptoms following use of adrenergic blocking agents (guanfacine [Tenex], clonidine [Catapres]), a dopamine antagonist (risperidone [Risperdal]), or an opiate (morphine [Roxanol]), none of these medications have been tested in well-designed double-blind, randomized

trials. Pharmacological treatments thus lack evidence for treating pediatric PTSD. Children who have comorbid psychiatric conditions for which effective pharmacotherapy is established should receive the appropriate psychopharmacological treatment of those disorders; it is not yet known whether such treatment will also impact favorably on coexisting PTSD symptoms.

Biological Treatments

Since dysregulation of the HPA axis is generally found in both children and adults with PTSD, use of a corticotropin-releasing factor (CRF) antagonist has been suggested as a potential intervention for preventing this disorder from developing or in reversing PTSD symptoms once they are present. This intervention has not yet been tested in children or adults.

C was a 15-year-old girl referred to treatment following hospitalization for a suicide attempt. From 6 to 12 years old, C had been sexually abused by paternal grandfather. C disclosed the abuse to her parents at 12 years old, who were supportive; however, a week later grandfather committed suicide. The following year C came out to her parents and close friends who were supportive. Parents thought C was doing well until the suicide attempt. In the hospital C disclosed the following: Since coming out, C experienced worsening physical, sexual, and verbal abuse in school related to her gay identity, but kept this from parents. During the current school year, C was repeatedly sexually assaulted by three male peers who threatened to kill her younger brother if she disclosed the abuse. Guilt (she "froze" during these episodes and did not say "no" or fight back) and negative self-cognitions led to the suicide attempt.

During the initial evaluation, C reported intrusive thoughts and nightmares about sexual abuse and grandfather's gory suicide. C tried unsuccessfully to push these images out of her mind. She acknowledged maladaptive cognitions including that she "deserved to be raped for killing my grandfather" and that she was "a terrible person." She endorsed sleep disturbance, increased startle, irritability, and sadness. She scored 58 on the UCLA PTSD Reaction Index (RI), consistent with very severe PTSD, and 20 on the Beck Depression Inventory (BDI), consistent with major depressive episode.

C and her parents attended 23 sessions of TF-CBT. Twenty of these were individual sessions (half of the session was spent with C and half with parents); three sessions were conjoint child–parent sessions in which the family met together. During the stabilization phase (11 sessions), C and parents learned about trauma reminders (e.g., the threat of sexual assault would serve as a trauma reminder for C in light of her history of sexual abuse), trauma impact (e.g., fight, flight, freeze, and trauma impact on affect, behavior, biology, cognitions, relationships and perceptions), as well as a variety of coping skills for trauma reminders. C and her parents reported that this helped them to make sense of C's response to the rapists as a trauma response and was "a huge relief" for C. C preferred using visualization and the cognitive triangle (changing thoughts in order to change feelings and behaviors). She carried a notebook with her and practiced changing her thoughts whenever she had a negative feeling. She found that this was very helpful and when paired with visualization, enabled her to reduce her distress level significantly in most situations. C's parents were very upset at the school for failing to protect C from harassment and physical assaults at school, and with C's agreement, investigated other schools that C could attend.

During the trauma narrative phase (7 sessions), C developed a detailed description and cognitively processed her personal trauma experiences. Through this process her understanding and cognitions about these experiences changed substantially. She identified how her grandfather manipulated her into worrying about her parents' response if she disclosed, and contributed to her self-blame for his suicide. Her new cognitions were that only the perpetrators were responsible for the abuse; grandfather chose to kill himself rather than face up to prison and that she did not deserve to be hurt for any reason. During this process, C

also made the decision that she wanted to go to a different school because "they didn't protect me and other gay kids the way they should have."

During the integration and consolidation phase (5 sessions), C shared her narrative with parents and engaged in safety planning with them. This included how C could protect herself from bullying going forward, and how parents could also advocate for her. C made the transition to her new school successfully and found this to be a more supportive environment. C and parents addressed healthy sexuality together and in individual sessions and C reflected on her future hopes and plans. At the end of treatment, C's RI score was 15 and BDI was 10 (both normal).

FUTURE DIRECTIONS

1. Evaluating DSM-5 diagnostic criteria for children ages 7 to 17 years, particularly those with complex trauma to determine whether the new criteria adequately describe the complex difficulties of these youth, or an additional diagnosis is needed.
2. Developing and testing effective treatments for children who have PTSD in the presence of comorbid psychiatric disorders.
3. Evaluating the relative efficacy of pharmacological treatments for pediatric PTSD, alone and in combination with psychotherapy.
4. Developing a biomarker for pediatric PTSD and other effective neurobiological tools for assessing pediatric PTSD.
5. Developing effective dissemination strategies for training community therapists in evidence-supported treatments for children with PTSD, including making these treatments usual care by embedding evidence-based treatments into routine graduate school training for counseling and social work students.
6. Developing and testing the use of social media applications for screening and treating pediatric PTSD. Given the large number of children who have PTSD, some will never reach mental health providers but most have access to electronic media.

REFERENCES

Caspi A, McClay J, Moffitt TE, et al. Role of genotype in the cycle of violence in maltreated children. *Science*. 2002;297:851–854.

*Cohen JA, Mannarino AP, Deblinger E. *Trauma-Focused CBT: Treatment Applications*. New York: Guilford; 2012.

*DeBellis MD, Zisk A. The biological effects of childhood trauma. *Child Adolesc Psychiatr Clin North Am*. 2014;23(2):185–222.

Foa EB, Keane TM, Friedman MJ, Cohen JA. Effective treatments for PTSD. *Treatment Guidelines from the International Society for Traumatic Stress Studies*. 2nd ed. New York: Guilford; 2008.

*Lieberman AF, Van Horn P. *Child-Parent Psychotherapy with Infants and Young Children*. New York: Guilford; 2008.

Saunders, BE, Adams, ZW. Epidemiology of traumatic experiences in childhood. *Child Adolesc Psychiatr Clin North Am*. 2014;23(2):167–184, vii.

Scheeringa MS, Zeanah CA, Cohen JA. PTSD in children and adolescents: towards an empirically based algorithma. *Depress Anxiety*. 2011;28(9):770–782.

Stein B, Jaycox LH, Kataoka S, et al. A mental health intervention for schoolchildren exposed to violence: a randomized controlled trial. *JAMA*. 2003;290:603–611.

▲ 52.3 Separation Anxiety, Generalized Anxiety, and Social Anxiety

COURTNEY P. KEETON, PH.D., JESSICA LEE SCHLEIDER, M.A., AND JOHN T. WALKUP, M.D.

Fear and worries are very common throughout childhood. Developmentally normal anxiety is short lived and nonimpairing and includes fears about the dark, monsters, or strangers for younger children and worries about friends, school, or decision making for older children and adolescents. Despite the common occurrence of transient fears and worries, a number of children experience persistent and impairing levels of anxiety that warrant a diagnosis of an anxiety disorder and clinical attention.

The three most common pediatric anxiety disorders—separation anxiety, generalized anxiety, and social anxiety—share the same underlying construct of anxiety, have significant symptom overlap, are highly comorbid with one another, and show similar familial relationships and continuity with adult anxiety and depressive disorders. For these reasons, the three disorders are often grouped together as they are in this section. The three anxiety disorders are described separately here; however, the reader should note that each etiology and treatment section is informative with regard to separation anxiety, generalized anxiety, and social anxiety.

Anxiety disorders comprise the most prevalent class of mental health problems among youth, affecting up to 1 in 5 individuals by the time they reach adulthood. Children with anxiety disorders have an early onset and a persistent or recurring course with high levels of short- and long-term impairment. In addition to being comorbid with one another, anxiety disorders are often comorbid with nonanxiety disorders. Fortunately, there are highly effective psychosocial and pharmacological treatments for these disorders. Current challenges include identification of treatment methods to optimize outcomes and address co-occurring emotional and behavioral problems, and dissemination and utilization of evidence-based treatment.

SEPARATION ANXIETY DISORDER

Definition

Separation anxiety disorder is characterized by excessive and developmentally inappropriate fear of being away from an attachment figure, avoidance of such separations, and irrational beliefs about the consequences of being apart from a caregiver. Attachment figures are usually parents, but may also be siblings, grandparents, or other caregivers

or, for adults, children or spouses. Whereas developmentally appropriate separation anxiety typically presents around age 6 months and declines between ages 2 and 3 years, children with separation anxiety disorder have either persistent and worsening or new onset separation anxiety in the school-age years (i.e., ages 6 to 12 years). This section differentiates between developmentally normal separation anxiety in infancy and toddlerhood and separation anxiety disorder in school-age children.

Comparative Nosology

Separation anxiety disorder appeared initially in the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III) and was classified as an anxiety disorder of childhood or adolescence. Although it was subsequently categorized in the DSM-IV as a "disorder usually first diagnosed in infancy, childhood, or adolescence," separation anxiety is currently classified in the DSM-5 as an anxiety disorder. Since first introduced the diagnostic criteria for separation anxiety disorder have been modified in four ways, along with modified wording in DSM-5 to capture symptom presentation in adulthood. Unlike earlier versions, the DSM-5 does not require onset prior to age 18 and recognizes that some adults report onset after age 18 years. The minimum duration was increased from 2 weeks in DSM-III to 4 weeks in DSM-IV and subsequent versions, and DSM-5 specifies at least 6 months in adults to avoid overdiagnosis of transient fears. Also, evidence of serious distress or impairment was added for diagnosis starting in DSM-IV. Finally, recurrent excessive distress about (1) anticipated or (2) actual separation was collapsed into one symptom in DSM-IV, and thus the symptomatic threshold for separation anxiety disorder was reduced from three of nine symptoms in DSM-III-R to three of eight symptoms in DSM-IV and currently in DSM-5.

Epidemiology

Epidemiological studies show the lifetime prevalence rate of childhood and adolescent separation anxiety disorder is at least 4 percent, with peak onset between 7 and 9 years of age. Separation anxiety disorder is more common in childhood than in the preschool years or adolescence and is the most common anxiety disorder in children under the age of 12. Although the peak age of onset is in middle childhood, preschool children can be diagnosed with separation anxiety disorder (range from 0.5 to 10.5 percent) and can be distinguished from children with developmentally normal separation anxiety. Children with separation anxiety symptoms who do not meet diagnostic criteria are very common, with reported rates as high as 50 percent. Community studies suggest that girls are more likely than boys to meet criteria even in the preschool years, but clinical studies find the disorder to be equally common in girls and boys.

Etiology

Although the exact cause of separation anxiety is unknown, genetic, early temperamental, and family or environmental variables have all been associated with early-onset anxiety disorders. Although genetic factors likely play a significant role, early experiences, many of which are family-based, may play an important role in the onset of childhood anxiety disorders. Methods to sort the various contributions of genetics, temperamental, and biological vulnerabilities and family or environment are often confounded. However, it is likely that genetically based, temperamental, or biological vulnerabilities predispose affected children and their parents to interact in ways that either protect children from or predispose children to developing an anxiety disorder. For example, an 8-year-old child with separation anxiety born to a mother with an anxiety disorder may have a very different outcome if the mother was symptomatic or successfully treated during the child's early years. Although in both situations the child and the mother share genetic and temperamental and biological vulnerabilities, the effect of the child's anxiety on the affected (and treated) parent and the affected (and treated) parent's impact on the child may account for significant difference in both short- and long-term outcomes.

Twin studies provide perhaps the greatest evidence for the role of biological and environmental factors (shared and nonshared) in the etiology of childhood anxiety disorders. Although historically both genetic and environmental factors have been implicated in the etiology of childhood anxiety disorders, a large well-designed study of 6-year-old twins suggests significant genetic influence for the separation anxiety phenotype. The heritability estimate for separation anxiety was high, 73 percent, with the remaining variance attributed to nonshared environment, suggesting genetic effects on separation anxiety are substantial and more significant than environmental effects. Another investigation of the role of genetic and shared environment effects in a sample of twins and siblings over time reported that the influence of genetic effects increased and shared environmental effects decreased with age. Finally, newly emerging data suggest that the psychobiological trait of CO_2 hypersensitivity may increase risk for separation anxiety (and supports a proposed genetic association between separation anxiety and later-onset panic disorder).

Behavioral inhibition (BI) to the unfamiliar, characterized by fear, withdrawal, and avoidance of unfamiliar situations or things, is considered a temperament vulnerability in infants and toddlers that predicts anxiety disorder outcomes in later childhood. BI is usually assessed using observational methods in laboratory settings for children aged 9 months through 6 years. Approximately 10 to 15 percent of children are thought to display BI and longitudinal outcomes, suggesting young children with BI

are more likely than noninhibited children to exhibit clinical levels of multiple anxiety disorders including separation anxiety upon long-term follow-up. In cross-sectional studies, BI is associated with a range of anxiety symptoms and disorders in youth, and it has traditionally been considered a vulnerability that lacks specificity for any childhood anxiety disorder. However, longitudinal data found BI to predict *social anxiety disorder* in middle childhood, but not other anxiety disorders.

Parental psychopathology is also highly associated with childhood anxiety disorders. A variety of mechanisms are possible; parents and children can share the same genetic or temperamental vulnerability or parent–child interactive effects could further shape or elaborate both parental and child anxiety. Regardless of the mechanism, children of parents with an anxiety disorder are up to five times more likely to be diagnosed with an anxiety disorder. More specifically, parental panic disorder and major depression are associated with an increased risk for separation anxiety disorder in children.

Longitudinal studies examining the role of infant–mother attachment in the development of childhood anxiety often used the Ainsworth Strange Situation to assess attachment and then follow children overtime. In multiple studies children with an anxious or resistant attachment during infancy were more likely than securely attached infants to develop anxiety disorders, including separation anxiety and related behaviors (e.g., school avoidance). Although infant temperament or maternal psychopathology may affect attachment behavior, cross-sectional assessment of attachment status appears to be the more important factor in determining the presence of anxiety disorders at longer-term outcome.

Specific parenting behaviors such as high levels of parental intrusiveness, overprotection, and psychologically controlling behavior in the child's daily activities may also contribute to a child's risk for developing anxiety. Such anxiety-enhancing parenting approaches are thought to promote dependency and reduce the development of autonomy. Controlling and overprotective parenting may contribute to the development of anxiety by preventing the child from having a range of experiences that foster broad and complex coping skills. Furthermore, parental overprotection may communicate to the child that the world is a dangerous place, introducing a chronic and pervasive response bias. Parents may also support anxious behavior by encouraging anxious cognitions and avoidance in their children or children may mimic anxious cognitions and behaviors modeled by their parents.

Diagnosis and Clinical Features

The defining feature of separation anxiety is excessive and persistent worry about harm befalling an attachment figure (e.g., parent getting into car accident, spouse becoming ill or dying) or oneself (e.g., getting lost, being

kidnapped, having an accident) upon separation. Symptoms may include avoidant behaviors such as refusal to go to school or work, sleep alone, or be without the attachment figure, in order to minimize separation and to reduce distress. Other common symptoms include nightmares with references to separation, and physical symptoms such as stomachaches or nausea in response to possible or actual separation. DSM-5 requires three of eight symptoms, a duration of 4 weeks in children and adolescents, and corresponding functional impairment. The diagnosis can be made in adults who are typically excessively concerned about their spouses or children and symptomatic for at least 6 months.

At home, children with separation anxiety display clingy behavior and shadow their caregiver. They often are reluctant to be in any room alone and may not want to sleep alone in their bedroom. These children often request a parent to fall asleep with them or find their way to their parent's bed at night. Also, they are uncomfortable staying with babysitters and may request to call parents frequently. In public, children with separation anxiety stay close to their caregiver and prefer to sit on their lap or be in physical contact. They are reluctant or refuse to attend birthday parties, sleepovers, camps, and extracurricular activities if the attachment figure is not present. Somatic complaints especially gastrointestinal (e.g., stomachaches, nausea) and cardiovascular (e.g., pounding heartbeat, light-headedness) are commonly reported, particularly before other anxiety-provoking separations. Children may show their separation anxiety distress by tantruming or becoming panicky or oppositional when anticipating separation.

At school, the child may make repeated requests to call their parents to check in or visit the school nurse with physical complaints. When children experience anxiety at school they may become restless and fidgety and be so preoccupied that they are unable to attend to classroom work. Separation anxiety disorder is also a common cause of school refusal and may result in frequent tardiness or high numbers of school absences.

Younger children with separation anxiety disorder tend to report a greater number of symptoms and more nightmares with separation themes than adolescents. Some older children and adolescents with separation anxiety may not present obvious impairment or distress, as their parents and school find ways to accommodate the child's fears to decrease their distress.

As separation anxiety not only affects the child, disruption of family life or parent stress may result from accommodating and managing the child's anxious and avoidant behavior. It can be helpful in the assessment to attempt to quantify the amount of accommodation and to ascertain what the anxiety severity would be if the parents did not offer ongoing accommodation around the child's anxiety.

Anna is a 7-year-old referred for outpatient treatment by her pediatrician. She exhibited a pattern of clingy behavior with her mother and expressed worry about something bad happening to her mother. She expressed worry that her mother would die or that she would be kidnapped. Anna had a history of shyness and difficulty separating from her mother. There is a family history of maternal panic disorder and paternal depression.

Anna often becomes irritable on school mornings or complains of feeling like she needs to vomit. When it is time to leave, Anna hides in her closet or under her bed. She refuses to ride the school bus and insists that her mother drive her to school. Her mother escorts Anna into school, and many times Anna cries and begs her not to leave. Anna does well academically and gets along with peers; however, she occasionally asks to call home or visits with the nurse with stomach aches, hoping to be sent home.

At home Anna will not be in any room alone and requires that her mother is in sight or sound. Anna refuses to shower, brush her teeth, or fall asleep without her mother present. As a result, Anna routinely sleeps with her parents in their bed. She often complains of nightmares of being kidnapped or seeing the "boogie man" near her mom.

Anna will not participate in activities or social events without a parent present. She has attended birthday parties or dance class only when accompanied by her mother. When Anna's parents make arrangements to leave her with a sitter or relative, she cries in distress and begs them not to go. Her parents call home frequently to check in with her. In public, Anna stays close to her mom and frequently sits on her lap when at restaurants or appointments.

Differential Diagnosis

Separation anxiety disorder can be distinguished from developmentally normal separation anxiety based on its onset or increasing severity in the school-age years, the duration of symptoms, and the degree of impairment or distress or need for parental accommodation.

Separation anxiety also should be differentiated from other emotional and behavioral disorders. Children with separation anxiety, generalized anxiety, and obsessive-compulsive disorder (OCD) all have concerns about harm. Children with separation anxiety worry about events that lead to being apart from an attachment figure. Children with generalized anxiety are more fearful of the actual event such as a dog bite, lightning striking, or getting sick than the impact on separation. Children with OCD have repetitive intrusive and senseless thoughts of harm befalling them or their loved ones if they do not complete a ritual. For example, a child with OCD might fear they will die if they do not ritualistically wash their hand four times. Similarly, children with separation anxiety may have social avoidance (e.g., parties or school) similar to children with social anxiety disorder, generalized anxiety disorder, OCD, or even posttraumatic stress disorder (PTSD), but the rationale for the avoidance will differ for each of the disorders. Children with separation anxiety may be described as panicking upon separation or anticipation of separation, but the panic is distinguished from spontaneous panic attacks that characterize panic disorder.

Finally, children with separation anxiety may exhibit defiance, irritable or labile mood, and explosive behavior when presented with a situation that involves separation. These behaviors or moods should be distinguished from oppositional defiant disorder, which involves more general patterns of

argumentativeness and misbehavior, and from bipolar disorder, which has prominent elevation in mood, decreased need for sleep, grandiosity, and increased goal-directed activities.

Course and Prognosis

The course of separation anxiety is not consistent for all children. Some children with separation anxiety appear to spontaneously remit; some remain anxious but are no longer prominently avoidant of separation triggers, and some truly evolve from separation anxiety into other anxiety disorders or depression. Multiple well-done prospective studies document that the majority of children with separation anxiety appear to improve with time. Some indeed appear to spontaneously remit, while others as they get older make avoidant choices consistent with the separation anxiety and may not appear distressed or obviously avoidant. For others with separation anxiety their symptoms evolve into more adult patterns of anxiety or depression. Prospective, longitudinal studies including both community samples and children at familial risk found that those meeting DSM-IV criteria for separation anxiety disorder were at an increased risk of developing subsequent panic disorder with agoraphobia, specific phobia, generalized anxiety disorder, OCD, pain disorder, and alcohol dependence, even after controlling for the risk of parental psychopathology. On the contrary, a recent meta-analysis reported consistent evidence that separation anxiety disorder poses increased risk for all anxiety disorders, but evidence was weak regarding risk for depression and substance use.

Treatment of Separation Anxiety Disorder and Other Anxiety Disorders

Both cognitive-behavioral therapy (CBT) and antidepressant medications appear to be effective for the short- and longer-term management of youth with separation anxiety and other anxiety disorders (e.g., generalized anxiety disorder, social anxiety disorder). Immediately following treatment with CBT or a selective serotonin reuptake inhibitor (SSRI), over 60 percent of anxious youth show clinically significant improvement. In the longer-term (between 2 and 19 years posttreatment), at least half of treated youth maintain initial gains (ranging from 48 to 93 percent). Recent research advances, including completion of the largest double-blind, placebo-controlled trial for childhood anxiety disorders (the "Child/Adolescent Anxiety Multimodal Study [CAMS]") have begun to address questions about the relative efficacy of CBT and SSRI treatment, predictors and moderators of treatment response, the utility of specific CBT components, and efficacy of CBT for very young children.

The CAMS compared the relative efficacy of CBT, sertraline, combination therapy (CBT plus sertraline), and placebo for youth with generalized anxiety disorder, separation anxiety disorder, and/or social

anxiety disorder. All active treatments were superior to placebo, and combination therapy led to significantly higher response rates than did CBT or sertraline (81 percent vs. 60 percent vs. 55 percent, respectively), supporting additive benefits of these two treatment modalities. Remission rates (loss of principal diagnosis, loss of all anxiety diagnoses) were lower than response rates, ranging from 20 to 68 percent, and followed the same patterns as response rates based on treatment condition. Factors associated with higher response and remission rates in the CAMS included absence of comorbid internalizing disorders, absence of social anxiety disorder at baseline, lower baseline anxiety severity, lower caregiver strain, nonminority status, and younger age. Additional variables with a pattern on inconsistent findings related to outcome in prior (and perhaps less well-powered) studies, including parental psychopathology, youth gender, and socioeconomic status, were examined with the CAMS data and found not to predict remission status.

New research also has emerged on effective treatments for anxiety in very young children. Specifically, controlled trials using developmentally sensitive downward extensions of CBT treatments with preschoolers diagnosed with anxiety disorders suggest that anxiety disorders are treatable very early in life. CBT for preschool-age youth is differentiated from CBT for school-age youth and adolescents by its greater emphasis on parent anxiety management and modification of parenting practices believed to maintain child anxiety. In addition to psychosocial approaches, an algorithm proposed by the Preschool Psychopharmacology Working Group suggests prescribing fluoxetine to preschool-age children (as low as 5 to 8 mg/day) if symptoms persist after CBT treatment.

Psychotherapy. Of the existing psychosocial treatments for youth anxiety, CBT has the greatest empirical support and is considered the treatment of choice. Manualized CBT programs for child anxiety such as Coping Cat, Coping Koala, SET-C, and FRIENDS, generally consist of 10 to 16 1-hour sessions delivered in child-focused individual, family, or group formats. CBT is based on principles of classical, operant, and social learning and generally include the following components: (1) psychoeducation: Corrective information about the nature of anxiety and feared stimuli; (2) somatic management techniques: Feelings identification, diaphragmatic breathing, and progressive muscle relaxation to target autonomic arousal and related physiological responses; (3) cognitive restructuring: Identification of maladaptive thoughts and teaching realistic, coping-focused thinking; (4) exposure techniques: Graduated, systematic, and controlled exposure to feared situations and stimuli; and (5) relapse prevention: Methods to consolidate and generalize treatment gains over time. Cognitive restructuring and exposure have been validated as essential components by empirical data showing accelerated

improvement in symptom severity and overall functioning following their introduction in treatment.

Over 30 randomized trials have demonstrated the efficacy of CBT for youth with separation anxiety disorder, generalized anxiety disorder, and social anxiety disorder. Meta-analytic data show that CBT is equally effective for children and adolescents, when delivered in individual or group formats, and when compared to passive (no treatment control) or active control groups. CBT for anxiety yields a mean effect size of $d = 0.66$, and acute gains are sustained for follow-up periods of 2 years or longer. This moderate-to-large effect size finding and durability of treatment effects over time supports the conclusion that CBT is an effective treatment for anxiety disorders with children and adolescents.

Family involvement in CBT is always recommended, but studies of family-focused treatments have yielded equivocal results. Some studies suggest additional benefit and some report no additional benefit above and beyond that of individual CBT. One study comparing standard Coping Cat treatment to a separation anxiety-specific CBT protocol with a parent intervention component yielded no statistically significant differences in outcomes. However, as family-focused treatment studies vary greatly and target different aspects of treatment (e.g., psychoeducation about anxiety, support for the child's participation in CBT, addressing parental anxiety directly, addressing "anxiety enhancing" parenting behaviors, or basic contingency management for child avoidance behavior), it is difficult to draw conclusions about the effectiveness of family treatment for child anxiety.

In recent years, a handful of research groups have successfully adapted and modified parent–child interaction therapy (PCIT), originally developed to treat externalizing problems in young children, to treat various early anxiety problems. These adapted PCIT protocols target child symptoms by reshaping parent–child interaction patterns associated with the maintenance of child anxiety symptoms. This goal is achieved through the use of live, unobtrusive parent coaching, delivered through a bug-in-the-ear receiver from a therapist situated behind a one-way mirror. Recently, the first randomized controlled trial (RCT) was conducted to test the efficacy of a 9-week, PCIT treatment for separation anxiety in children ages 4 through 8. At posttreatment, 73 percent of children assigned to PCIT no longer met criteria for a diagnosis of SAD, and results were largely maintained at a 3-month follow-up visit. Preliminary efficacy studies of another PCIT adaptation, the Coaching Approach behavior and Leading by Modeling (CALM) Program, designed to target the full range of anxiety disorders affecting young children (generalized anxiety disorder and social anxiety disorder, in addition to separation anxiety disorder) suggest the promise and feasibility of this approach.

Although evidence strongly supports CBT as an efficacious treatment

for childhood anxiety disorders, disseminating this specialized form of psychotherapy is an ongoing challenge. Also, service utilization data show that far fewer treatment sessions are completed in practice as compared to protocol-based treatment trials for anxiety. Thus, recent years have spawned the development of computer-assisted treatments such as BRAVE-ONLINE and Camp-Cope-A-Lot, which combine computer-based sessions with online and phone contact with a CBT clinician. Preliminary trials of these programs, including one RCT, suggest that computer-assisted and in-person CBT produce statistically similar clinical improvements. Given the initial promising findings, and the capacity of computer-based CBT to increase treatment accessibility, these programs merit continued investigation.

Another newly emerging computer-based treatment for child anxiety disorders is Attention Bias Modification (ABM; also called Cognitive Bias Modification). This approach draws from cognitive theories of anxiety, specifically relating to attentional biases. Evidence indicates that individuals who experience anxiety disorders attend more to threat-related stimuli, leading them to grow hyper-vigilant to perceived threats. Over time, experiencing this hypervigilance exacerbates anxiety levels, which in turn increases attention to threat and produces a maladaptive feedback loop. ABM aims to intercept this cycle by using a dot-probe paradigm to implicitly modify biases toward threat in anxious individuals. A recent meta-review of ABM trials for pediatric anxiety found that, in eight of ten existing efficacy trials (including three RCTs), participants experienced significant improvements in anxiety postintervention. While these findings are promising, further research is needed to clarify the optimal frequency and duration of ABM, the most effective intervention setting, and whether ABM is a useful adjunct to traditional CBT.

Pharmacology. SSRIs are considered first-line medications for this population due to the favorable treatment response (effect sizes greater than 0.8) and low side-effect profile identified in a number of single-site and multisite randomized, controlled trials. Agents that have demonstrated efficacy include fluvoxamine (Luvox), fluoxetine (Prozac), sertraline (Zoloft), paroxetine (Paxil), duloxetine (Cymbalta), and the tricyclic antidepressant imipramine (Tofranil). Despite demonstrated efficacy of SSRIs for anxiety in RCTs, no SSRIs are approved by the U.S. Food and Drug Administration (FDA) for use with children and adolescents. However, the serotonin-norepinephrine reuptake inhibitor duloxetine received an FDA indication for generalized anxiety disorder in children and adolescents ages 7 to 17 years.

Open-label follow-up studies of SSRIs suggest that youth who benefit acutely continue on a positive trajectory for up to 6 months. Further, since anxiety disorders tend to run a chronic course with associated impairment,

longer treatment with medication may be justified to support consolidation of initial gains and prevent return of anxiety symptoms.

How long to treat children with medication to maximally reduce the chance for return of symptoms upon medication discontinuation is not known. Therefore, making the decision when to discontinue medication is often determined pragmatically. Treatment for 6 to 12 months after remission in symptoms may provide extended benefit without undue burden and may offer an opportunity to reduce the chance for a return of symptoms after medication discontinuation. However, longer-term treatment may be necessary if children experience a return of symptoms when medication doses are missed or during known trigger points such as the start of school in the fall or after the winter holidays (separation anxiety), grade reports or final examinations (generalized anxiety), or group activities at school or camp (social). Such tenuous symptom control suggests that continued treatment, dose increases, or more intensive psychosocial treatments be provided to stabilize the patient before considering medication discontinuation. In addition, the timing of the discontinuation trial is critical to prevent the child from being exposed to a return of symptoms at a time when anxiety triggers are present or when family or medical support are not available (i.e., start of school for separation anxiety or participation in performance activities for generalized or social anxiety).

Benzodiazepines have not been extensively studied for the acute or chronic management of anxiety in children, and while potentially useful, they may lead to both psychological and physical dependence and thus complicate treatment longer term. Given the efficacy of SSRIs and other antidepressants for childhood anxiety, benzodiazepines should not be considered as first-line treatment. If used at all, benzodiazepines may be considered for short-term augmentation of antidepressants early in the course of treatment for some highly anxious children.

There is little evidence supporting the use of other medications such as antihistamines, α-adrenergic agonists, or antipsychotics for the treatment of childhood anxiety disorders; however, it is not uncommon to see patients clinically who are treated with such agents to address problems with insomnia related to anxiety. Combining other medications with SSRIs to address anxiety-related insomnia should only occur when antidepressants have been maximally adjusted upward and anxiety and insomnia persist. Then the use of sedating antidepressants is preferred to the use of antihistamines, α-adrenergic agonists, or antipsychotics.

GENERALIZED ANXIETY DISORDER

Definition

In contrast to separation and social anxiety, which generally have one

anxiety triggering experience (i.e., separation and social humiliation, respectively), generalized anxiety disorder is characterized by excessive and uncontrollable worry about a number of anxiety-provoking events. The most common symptoms among affected youth are poor concentration, restlessness, and irritability, but other symptoms include fatigue, sleep disturbance, or muscle tension. As with other anxiety disorders, generalized anxiety must cause marked distress, interfere with social, emotional, and educational functioning, and is not simply a reaction to recent stressors.

Comparative Nosology

Prior to DSM-IV, a diagnosis of generalized anxiety was reserved for individuals 18 years or older, and children with multiple fears and worries could be diagnosed with overanxious disorder. Overanxious disorder was characterized by excessive or unrealistic worry for a period of at least 6 months. Due to the poor reliability and validity of this diagnostic category and symptomatic overlap and continuity with adult generalized anxiety, overanxious disorder was incorporated into the DSM-IV generalized anxiety disorder diagnosis. Research comparing cases diagnosed by DSM-III-R and DSM-IV criteria found that the change in terminology did not affect the characteristics of cases. Criteria remained unchanged for DSM-5.

Epidemiology

Generalized anxiety disorder is a common diagnosis in children, with a reported lifetime prevalence of 9.0 percent. Community samples of preschool children report a prevalence range of 0.5 to 6.5 percent for overanxious disorder or generalized anxiety disorder. Based on data from two community epidemiology studies, the 3-month prevalence rate for generalized anxiety disorder among children aged 9 to 12 is 1.4 percent. The prevalence for adolescents aged 13 to 18 ranges between 2.3 and 5.0 percent. Whereas rates of separation anxiety decline with age, rates of generalized anxiety increase with age, with a mean age of onset of 8 to 9 years, and median age of onset of 12 years. Girls are twice as likely as boys to meet diagnostic criteria.

Pediatric generalized anxiety disorder is commonly associated with other anxiety disorders and mood disorders. Data from the NCS-A and GSMS found that up to 2/3 of affected youth met criteria for another anxiety disorder, and about ½ met criteria for a mood disorder. Preschoolers with generalized anxiety have the highest rates of comorbidity compared to preschoolers with other anxiety disorders. Few significant differences in rates and patterns of comorbidity across age and gender have been identified; younger children are more likely than adolescents to have separation anxiety as a comorbid diagnosis, and males are more likely than

females to have an externalizing disorder as a co-occurring diagnosis.

Etiology

Similar to separation anxiety disorder, both genetic and environment factors contribute to the development of generalized anxiety disorder. However, heritability estimates are lower for generalized disorder (approximately 30 percent), suggesting that environmental factors, both shared (e.g., uterine environment in the case of twins, exposure to common parenting in the case of siblings) and nonshared (e.g., personal traumatic experience, peer relationships), may have a more prominent etiological role. Twin studies show that generalized anxiety may be better explained by additive genetic factors and nonshared environmental factors compared to shared environment factors. Low heritability estimates in twin studies and high familial rates of generalized anxiety suggest that increased rates in families may be more related to environmental factors. For example, parental anxiety, parenting approaches, or exposure to stressful life events may negatively affect childhood anxiety outcomes. Similarly, childhood anxiety may also impact parental or sibling anxiety as the child's anxiety may change the larger home environment.

Anxiety sensitivity is a temperamental or trait-like factor under genetic influence that is considered a risk factor for anxiety in general. It is characterized by a tendency to interpret physiological symptoms as a sign that something harmful will result. Such factors, while not specific to anxiety, may predispose children in certain environments to become symptomatically more anxious and impaired. Similarly, an attentional bias toward threatening cues and a tendency to interpret ambiguous situations as threatening are associated with children with current anxiety disorder. Although such relationships cannot be considered causal, it is possible that such attentional biases coupled with environmental variables such as parenting influences (e.g., hearing information from a parent about the potential dangers of a situation) may ultimately shape anxiety-related beliefs and behaviors. Anxiety sensitivity has been related specifically to panic in adolescents and otherwise has been linked to anxiety in general.

Diagnosis and Clinical Features

The hallmark feature of generalized anxiety disorder is excessive worry that is difficult to control (Table 52.3–1). The threshold for diagnosis differs slightly for adults and children; adults are required to experience three of six associated symptoms, whereas children are required to have only one associated symptom. A diagnosis also requires the child to experience excessive worry more days than not for 6 months, the worry and associated symptoms must interfere with daily functioning, and the anxiety is not triggered by recent stressful events.

Table 52.3–1.
DSM-5 Diagnostic Criteria for Generalized Anxiety Disorder

A. Excessive anxiety and worry (apprehensive expectation), occurring more days than not for at least 6 months, about a number of events or activities (such as work or school performance).
B. The individual finds it difficult to control the worry.
C. The anxiety and worry are associated with three (or more) of the following six symptoms (with at least some symptoms having been present for more days than not for the past 6 months):
Note: Only one item is required in children.
   1. Restlessness or feeling keyed up or on edge
   2. Being easily fatigued
   3. Difficulty concentrating or mind going blank
   4. Irritability
   5. Muscle tension
D. Sleep disturbance (difficulty falling or staying asleep, or restless, unsatisfying sleep)
E. The anxiety, worry, or physical symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
F. The disturbance is not attributable to the physiological effects of a substance (e.g., a drug of abuse, a medication) or another medical condition (e.g., hyperthyroidism).
G. The disturbance is not better explained by another mental disorder (e.g., anxiety or worry about having panic attacks in panic disorder, negative evaluation in social anxiety disorder [social phobia], contamination or other obsessions in obsessive-compulsive disorder, separation from attachment figures in separation anxiety disorder, reminders of traumatic events in posttraumatic stress disorder, gaining weight in anorexia nervosa, physical complaints in somatic symptom disorder, perceived appearance flaws in body dysmorphic disorder, having a serious illness in illness anxiety disorder, or the content of delusional beliefs in schizophrenia or delusional disorder).

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

The vast majority of pediatric patients with generalized anxiety (75 to 80 percent) endorse physical symptoms. Common physical symptoms include headaches, tension, restlessness, gastrointestinal distress, and heart palpitations. Cognitive characteristics of generalized anxiety include negative thinking errors such as catastrophizing (i.e., expecting the worse possible outcome) or overestimation of the likelihood that something undesirable will occur. Generalized worries cover a range of everyday issues, such as being on time, upcoming activities, failure of loved ones or friends to meet basic expectations, or whether unexpected events (e.g., inclement weather) will change daily plans or schedules. Children and adolescents with generalized anxiety are often described as perfectionistic and overly sensitive. They are plagued by negative self-image, need for reassurance, and insecurities about competence, performance, and expectations about the future.

Andrew is a 13-year-old referred for outpatient treatment by his pediatrician for frequent headaches thought to be associated with anxiety. Andrew was born via cesarean section with no reported perinatal or postnatal complications. As an infant he was difficult to soothe.

Medical history was unremarkable. Andrew was reported to be sociable and well-liked by his peers. Academically, he was an honor roll student enrolled in advanced classes. He had no psychiatric treatment history. Family history was significant for paternal social anxiety and maternal depression. Andrew's two siblings, one older and one younger, were reported to be high functioning with no psychiatric history. Clinical interview revealed no concerns regarding externalizing behavior. Worries were reported about a range of situations including school, the future, athletic performance, family matters, and health and safety.

During the interview, Andrew was engaged and cooperative. He described himself as "stressed out" because of school. He was an A student, but endorsed worries related to examinations, grades, and whether his upcoming high school courses would prepare him adequately for college. He also played point guard on the school basketball team and acknowledged worries about his speed and height compared to other players and questioned whether he was good enough to make the high school team. In addition, he expressed concern about family-related issues such as finances, parental conflict, and the potential for his parents to get divorced. Andrew acknowledged having headaches and stomach upset at least 4 days per week. Upon questioning, he reported that he cannot stop worrying and that worries keep him awake at night about 5 days per week. When given rating scales to complete, Andrew asked several questions about the directions and was concerned about completing the forms correctly.

Andrew's parents described him as a "worry wart," saying he has always been anxious. They reported that he often has health concerns. Most recently, he stopped using an over-the-counter face wash containing salicylic acid and expressed concern about accidentally swallowing the product and being poisoned. Also, his recent trip to the pediatrician was prompted by Andrew's questioning about whether he needed a computed tomography (CT) scan due to his headaches. His parents report that he is frequently irritated and says that he is "stressed out" when questioned about school or basketball. He often seeks reassurance from his parents, but they find their attempts at doing so are unhelpful over the long term.

Differential Diagnosis

Worry associated with generalized anxiety has multiple targets that include worries regarding everyday life experience. Therefore, it may be more difficult to differentiate generalized anxiety from normative anxiety experienced by all individuals. However, the overall amount of anxiety, poor anxiety control, the lack of amelioration with appropriate reassurance and support and longer-term impairment associated with generalized anxiety differentiate it from normal worry.

Generalized anxiety can be differentiated from emotional disorders based on the greater number of anxiety triggers and the specific nature of the worries. Children with generalized anxiety are concerned with past and future experiences, physical health, weather, environmental changes, and unexpected events. They may appear to have similar concerns to children with separation and social anxiety but differ in important ways. For example, children with generalized anxiety may avoid going to school, not because of separation or social concerns but because of academic performance worries. That said, it is not uncommon for younger children with generalized anxiety disorder to also be diagnosed with co-occurring separation anxiety and older children with co-occurring social anxiety disorder. Individuals with generalized anxiety and perfectionism may superficially resemble those with OCD. For example, a child with generalized anxiety and concern about academic performance may

routinely check over homework for accuracy or study material repeatedly for the purpose of improving the work product or getting it "perfect." In contrast children with OCD know the work product is good or perfect, but experience doubt that their perception is accurate and then recheck for the purpose of rechecking. Also, individuals with generalized anxiety may have recurrent uncontrollable worries that could be mistaken for obsessions; however, worries of generalized anxiety disorder do not appear senseless, intrusive, or unwanted, rather they are described as powerful bona fide concerns and are usually tied to events or situations that would stimulate normal concerns. For example, normal children may worry about bad weather as it happens, generalized anxiety disorder children will worry that bad weather may happen on a cloudless day, and children with OCD may worry that if they do not count to three in sets of three then a loved one may die in a weather-related incident.

Ruminations (i.e., recurrent depressive cognitions) are characteristic of depression and tend to regard experiences negatively and hopelessly. For example, repetitive thoughts such as "No one will like me no matter what I do" can be readily distinguished from "I need to be perfect for my teachers and parents" by identifying the depressive (ruminative) quality of the former and the perfectionistic (generalized anxiety) quality of the latter.

Course and Prognosis

Prospective and cross-sectional studies of youth suggest that GAD is an episodic or recurrent disorder, and that prognosis is dependent on comorbidity and severity. Individuals with a comorbid diagnosis tend to have higher ratings of severity and functional impairment. Those with greater severity and psychosocial impairment at baseline are more likely to retain a diagnosis over time.

Like other anxiety disorders, generalized anxiety disorder is considered a risk factor for the development of additional psychiatric problems over time. Specifically, generalized anxiety may lead to the subsequent development of anxiety disorders, substance use, depression, or behavior problems. The New York Child Longitudinal Study investigated the associations between anxiety disorders and later psychiatric disorders in 9- to 18-year-old children and adolescents. DSM-III overanxious disorder was related to subsequent major depression, social anxiety disorder, and generalized anxiety. Whereas children with separation anxiety symptoms were less likely than other children to begin drinking alcohol were and did so later than others, those with generalized anxiety symptoms were more likely than other children to begin drinking were and did so earlier. Results from the GSMS found that DSM-IV generalized anxiety was associated only with later conduct disorder, but did not predict anxiety or other psychiatric problems.

Treatment of Generalized Anxiety Disorder and Other Anxiety Disorders

Similar to the previous section on separation anxiety, both CBT and SSRI treatment have demonstrated efficacy for pediatric generalized anxiety disorder. Preliminary trials also support PCIT adaptations as potentially effective treatments for very young children.

Psychotherapy. Although there are no randomized controlled CBT trials specifically for generalized anxiety, the basic concepts of CBT as introduced in the section on separation anxiety disorder apply to generalized anxiety disorder, but may require modifications to address the unique symptom presentation. For example, perfectionism is often a highly valued trait. Making perfectionism a negative symptom and using exposure exercises to tolerate the discomfort and distress associated with making mistakes may be more challenging than addressing more clearly impairing symptoms. It is also important to elicit physiological signs (i.e., heart pounding, sweating) in the controlled context of the exposure so that, in the process of achieving effective coping, the child learns there is a component of self-control in reducing somatic symptoms. In addition, the exposure must elicit relevant anxious cognitions (i.e., "If this isn't done just right, then I won't get the best grade.") so that the child learns to effectively challenge and reframe thoughts about specific anxiety-provoking situations.

Children with generalized anxiety disorder often have specific worries that are difficult or unsafe to challenge in reality and require greater use of imaginal exposures. Unrealistic worries, for example, about health, safety, death, or dying, may be difficult or impossible to enact through an in vivo exposure. A child with excessive worries about contracting a disease, for example, would be treated via imaginal exposure whereby detailed events and feelings are identified first. Next, problem solving, relaxation exercises, and cognitive restructuring are used to help the child interrupt the cycle of worrying, learn to tolerate discomfort associated with the anxiety, and develop realistic thought patterns. Parental involvement in treatment is essential to modify ineffective reassurance statements ("Don't worry. You won't get sick.") with realistic notions of uncertainty ("Chances are low that you will get sick, but there is no way to know for sure.") and by encouraging children's awareness of their self-coping ("What do you think would happen if you did get sick? But what else could happen?") rather than offering reassuring solutions ("If you did get sick, we would get you medicine to make you better.").

Pharmacology. Published trials specifically targeting generalized anxiety included one effectiveness trial with sertraline and two trials—one positive and one negative—with venlafaxine (Effexor), which were combined and published together. More recently duloxetine was found to

be effective in a large, industry-sponsored, randomized, placebo-controlled trial of children and adolescents 7 to 17 years old with generalized anxiety disorder leading to the first FDA indication of any medication for a non-OC, non-PTSD anxiety disorder in children and adolescents.

Although adults with generalized anxiety disorder may be treated acutely with benzodiazepines, there are few data to support the efficacy of this class of medications for youth. For example, a small double-blind, placebo-controlled randomized trial of alprazolam (Xanax) for the treatment of pediatric overanxious and avoidant disorders, demonstrated no benefit of medication over placebo. As noted previously in the section on separation anxiety, the efficacy of antidepressant medication combined with the lack of efficacy for benzodiazepines and the risk for physical and psychological dependence make it difficult to consider benzodiazepines as an important treatment option for childhood anxiety disorders.

Buspirone (BuSpar), a nonbenzodiazepine, nonantidepressant anxiolytic, is effective in studies of adults with generalized anxiety disorder and was evaluated in a large, multisite, industry-sponsored, RCT. Although the trial was not published, a report of the negative outcome of the study is available on the FDA website.

SOCIAL ANXIETY DISORDER

Definition

Social anxiety disorder, also called social phobia, is characterized by excessive, persistent, and impairing anxiety about potential scrutiny or embarrassment in social or performance situations. The worry may be restricted to a specific performance situation such as fear of being negatively evaluated during an athletic or musical performance. However, social anxiety disorder is more commonly generalized such that worry about being scrutinized or embarrassed occurs across a range of activities and situations.

Comparative Nosology

Social anxiety disorder (social phobia) first appeared in DSM-III, with minimal alterations in subsequent editions. One impactful change from DSM-III to DSM-IV was the removal of the diagnosis of avoidant disorder. Avoidant disorder was classified as an anxiety disorder of childhood or adolescence, and characterized by fear and avoidance of strangers resulting in significant impairment in peer relationships. Concern about the overlap between the two disorders and continuity between childhood avoidant disorder and adult social anxiety disorder resulted in the removal of avoidant disorder as a diagnostic category in DSM-IV. In DSM-5, there is no longer a requirement that individuals over age 18 must recognize their fear as excessive and unreasonable; instead, the anxiety must be out of

proportion to the actual threat posed by the event after taking into account sociocultural factors. Whereas the 6-month duration was limited in DSM-IV to individuals under age 18, this criterion now applies to all ages. Finally, unlike DSM-IV, there is no longer a specifier of "generalized" but there is a specifier for "performance only" in DSM-5.

Epidemiology

Social anxiety is a common anxiety disorder, with a cumulative incidence of 11 percent. Prevalence rates have increased over time, likely because of the removal of the avoidant disorder diagnostic category from DSM-III to DSM-IV; that is, youth who were once characterized with avoidant disorder are now diagnosed with social anxiety disorder, resulting in increased prevalence rates. According to data from the National Comorbidity Survey Replication Adolescent Supplement and the Early Developmental Stages of Psychopathology Study (EDSP), the 12-month prevalence for adolescents is 5.2 to 8.2 percent. The lifetime prevalence of social anxiety is 9.5 percent for females and 4.9 percent for males. Like generalized anxiety, the prevalence of social anxiety disorder increases with age, possibly reflecting changes in cognitive and emotional development (i.e., the capacity for self-evaluation and self-consciousness).

Age of onset for social anxiety disorder is during adolescence. The average age of onset in epidemiological studies is between 10 and 16.6 years and varies depending on the sample studied and the strategy for determining retrospective reports for age of onset. According to the EDSP which involved a community sample of 3,021 individuals aged 14 to 24, the incidence of DSM-IV social anxiety was greatest among those aged 10 to 19 years.

Etiology

Similar to both separation anxiety disorder and generalized anxiety disorder, factors that contribute to the development of social anxiety disorder include genetics, temperament, and family variables. Twin studies are useful in distinguishing between environmental and genetic influences. There are many discrepancies in the literature regarding the contributions of genetic and environmental influences because of differences in measurement (i.e., rating scale vs. psychodiagnostic interviews), informants (i.e., self vs. parent reports), and criterion (i.e., symptoms vs. severity vs. clinical diagnosis). However, the literature is consistent in showing that anxiety disorders are heritable and that both shared and nonshared environmental factors also play a role. Concordance rates for anxiety disorders including social anxiety disorder are higher for monozygotic twin than dizygotic twins (34 vs. 17 percent). Familial transmission is also evident by studies showing a high concordance rate of

social anxiety among first-degree relatives. Overall, studies investigating genetic transmission find evidence for an inherited propensity toward anxiety in general rather than for a specific type of anxiety.

As previously mentioned, BI is a temperamental construct characterized by increased physiological reactivity and avoidant behavior in the face of novel or challenging situations. Inhibited children have greater sympathetic activation than uninhibited children, evidenced by higher and more stable heart rates and greater heart-rate acceleration under stressful and novel conditions. Many studies support the existence of a specific link between BI and social anxiety. Both longitudinal and retrospective studies show that toddlers with BI assessed by observational or questionnaire methods are more likely than uninhibited children to exhibit social anxiety as adolescents. In a recent prospective longitudinal study, 215 children assessed for BI were assessed for psychiatric outcomes in a 5-year follow-up at a mean age of 9.6 years. Structured diagnostic interviews revealed that BI present at age 4 or 6, but not at age 21 months, predicted social anxiety. BI was not associated with other specific anxiety, mood, or behavior disorders. In another study, adults from a clinical sample provided retrospective reports of childhood inhibition, and those with the highest levels of inhibition also had the highest levels of clinical social anxiety disorder. It is important to note that less than 50 percent of toddlers and young children characterized as behaviorally inhibited go on to develop social anxiety disorder. Similarly, children who are considered shy or have shy temperaments most likely do not develop social anxiety disorder. Thus, BI is considered a manifestation of a biological predisposition that serves as a risk factor for anxiety.

Environmental variables such as parent–infant attachment and parenting behaviors also are associated with the expression of social anxiety disorder. As stated in previous sections, retrospective and observational research shows links between insecure infant attachment and the development of anxiety disorders. One study examined 172 adolescents 16 years after participating in the Ainsworth Strange Situation test at age 12 months and confirmed that infants with insecure attachments were more likely to meet criteria for anxiety, including social anxiety disorder. Adult parents with social anxiety disorder may sponsor social isolation by discouraging opportunities for social engagement. They also provide information that has the potential to heighten anxious arousal.

Diagnosis and Clinical Features

The defining feature of social anxiety is excessive fear about being scrutinized in social or performance situations (Table 52.3–2). Although the same criteria apply to children and adults, some provisions are indicated for children. First, social anxiety cannot be limited to adult settings and must occur in peer contexts. Second, in children, the anxiety

evoked by the social situation may involve crying, tantrums, freezing, or shrinking from social situations with unfamiliar people. Common fears related to social anxiety include public speaking, participating in class, initiating or participating in conversations, attending social events, and eating or writing in front of others. Fear of public speaking is considered the most commonly feared social situation; however, test anxiety is commonly feared among adolescents.

Some differences have been noted with respect to the presentation of social anxiety among different age groups. Adolescents endorse anxiety across a broader range of situations and more avoidance compared to children; however, both groups experience similar levels of distress. In addition, adolescents endorse more cognitive symptoms of social anxiety that likely reflect their greater cognitive maturity compared to children. Poor social skills, however, are evident across all youth with social anxiety disorder.

Table 52.3–2.
DSM-5 Diagnostic Criteria for Social Anxiety Disorder (Social Phobia)

A. Marked fear or anxiety about one or more social situations in which the individual is exposed to possible scrutiny by others. Examples include social interactions (e.g., having a conversation, meeting unfamiliar people), being observed (e.g., eating or drinking), and performing in front of others (e.g., giving a speech).
   Note: In children, the anxiety must occur in peer settings and not just during interactions with adults.
B. The individual fears that he or she will act in a way or show anxiety symptoms that will be negatively evaluated (i.e., will be humiliating or embarrassing; will lead to rejection or offend others).
C. The social situations almost always provoke fear or anxiety.
   Note: In children, the fear or anxiety may be expressed by crying, tantrums, freezing, clinging, shrinking, or failing to speak in social situations.
D. The social situations are avoided or endured with intense fear or anxiety.
E. The fear or anxiety is out of proportion to the actual threat posed by the social situation and to the sociocultural context.
F. The fear, anxiety, or avoidance is persistent, typically lasting for 6 months or more.
G. The fear, anxiety, or avoidance causes clinically significant distress or impairment in social, occupational, or other important areas of functioning.
H. The fear, anxiety, or avoidance is not attributable to the physiological effects of a substance (e.g., a drug of abuse, a medication) or another medical condition.
I. The fear, anxiety, or avoidance is not better explained by the symptoms of another mental disorder, such as panic disorder, body dysmorphic disorder, or autism spectrum disorder.
J. If another medical condition (e.g., Parkinson disease, obesity, disfigurement from burns or injury) is present, the fear, anxiety, or avoidance is clearly unrelated or is excessive.

Specify if:
Performance only: If the fear is restricted to speaking or performing in public.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Infants and toddlers cannot meet DSM-5 anxiety diagnoses guidelines primarily due to developmental issues; however, many very young children exhibit higher than normal levels of social fear and associated distress. For example, some children have difficulties playing with new children, refuse to participate in playgroups and classes, play alone at parties or on the playground, and refuse to eat in restaurants or attend public events because of severe distress. Some pilot data reveal that children aged 18 months to 5 years can be diagnosed with a modified version of social anxiety disorder and are more likely than their nonanxious counterparts to display high levels of shyness, inhibition, and to have anxious parents. Many preschool-age children meet DSM-5 criteria for selective mutism which involves restricted speech in social situations due to high levels of fear and anxiety.

Jessica is an 11-year-old in the sixth grade at a public elementary school. She was referred for outpatient treatment by her school counselor for anxiety. Jessica was accompanied by her mother, who is a single parent.

During the assessment, Jessica was compliant but timid. She gave succinct responses to interviewer questions and frequently turned to her mother who often answered on Jessica's behalf. Her medication history was significant for irritable bowel syndrome. Psychiatric history was unremarkable. There is a history of maternal shyness and social anxiety, and Jessica's mother is presently prescribed an SSRI. Jessica was described as "painfully shy" as a very young child.

Jessica endorsed feelings of loneliness and reported having few friends. She reported that, in school, she does not answer questions or do other things that could draw attention to herself (e.g., leave class to use the bathroom). Recently, she refused to go to school on a day she was scheduled to give a presentation in one of her classes. In addition, she admitted to purposefully leaving her gym shoes at home on several occasions because she does not like to change in front of her peers in the locker room at school. Jessica is a Girl Scout but said she does not like the meetings or events and feels that the other girls do not like her. She also endorsed worries that other girls will make fun of her clothes. Jessica does not talk during the meetings and often asks her mother if she can skip. Jessica has attended a couple of birthday parties but typically does not initiate conversations and has been reluctant to fully engage in some of the birthday activities (e.g., bowling, hitting a piñata).

According to her mother, Jessica is shy in public social situations. She rarely orders her own food in restaurants and does not respond when asked a question by a store clerk. At home, Jessica does not answer the phone unless she recognizes the incoming call number identifies a familiar relative.

Differential Diagnosis

Social anxiety disorder can be distinguished from other disorders based on the predominant feature that fears center around the potential for embarrassment or receiving negative criticism. Although social anxiety and separation anxiety may both involve avoidance of social situations, the fear with separation anxiety is associated with concern about being apart from a caregiver and is not associated with potential humiliation or embarrassment. The generalized type of social anxiety is differentiated from generalized anxiety disorder on the basis of the individual feeling humiliated or embarrassed in public rather than a more global worry. Also,

children with generalized anxiety disorder may worry about a host of social issues (e.g., friendships, performance, how others view them), but worries are not limited to social or performance situations. Although individuals with social anxiety disorder may experience panic in the face of challenging situations, the fear and associated somatic and cognitive symptoms occur in response to specific situations and are not unexpected or uncued as they are in panic disorder.

Social anxiety is also distinguished from selective mutism, which is characterized explicitly by a failure to speak in situations where speech is expected and has a much younger age of onset relative to social anxiety disorder. Selective mutism is considered by some professionals to be a developmental variant of social anxiety.

Finally, it is important to distinguish social anxiety disorder from developmentally normal fears. For example, it is not uncommon for children to be anxious about speaking in front of the class. However, normative anxiety does not result in significant impairment in functioning or clinically significant distress as with social anxiety disorder.

Course and Prognosis

Social anxiety disorder that emerges by the early teen years is associated with progressive worsening or persistence of symptoms through early adulthood. Late adolescence is characterized by a waxing and waning course during which symptoms are exacerbated with social stressors. Factors associated with persistence of symptoms include early onset, more anxiety cognitions, severe avoidance and impairment, co-occurring panic, parental social anxiety and depression, BI, and harm avoidance.

Few individuals experience spontaneous remission; data from the prospective-longitudinal Early Developmental Stages of Psychopathology study suggest a natural remission rate of 15.1 percent. Individuals with social anxiety disorder withdraw from or avoid situations, limiting opportunities for optimal social, academic, and emotional development. Research shows that children with social anxiety disorder are less socially skilled than nondisordered counterparts, limiting their capacity to form intimate relationships with friends, family members, or partners. Symptoms and impairment at school may result in academic underachievement, including failure to graduate high school or pursue college. Disabilities in educational and social functioning may lead to demoralization and other emotional problems. Longitudinal studies show, for example, that social anxiety disorder is consistently linked with increased risk for subsequent depression. Further, individuals with comorbid social anxiety and depression are at increased risk for suicidal ideation and suicide attempts.

Several longitudinal studies find that social anxiety disorder is associated with the development of substance use problems. For example,

one prospective study based on a subsample of 816 individuals from the Oregon Adolescent Depression Project found that social anxiety disorder is a unique risk factor for the subsequent onset of alcohol and cannabis dependence. Specifically, a lifetime history of social anxiety disorder at age 17 was associated with 4.5 times greater odds of alcohol dependence and 6.5 times greater odds of cannabis dependence by age 30. Also, results from a different prospective study showed that youth with social anxiety disorder were significantly more likely 4 years later to smoke cigarettes and have a higher rate of nicotine dependence than a control group.

Treatment of Social Anxiety Disorder

As noted in previous sections on separation anxiety disorder and generalized anxiety disorder, effective treatments for social anxiety disorder include CBT and SSRIs. There is one comparative treatment trial specifically focused on youth with primary social anxiety disorder. Children and adolescents aged 7 to 17 were randomly assigned to Social Effectiveness Therapy for Children (SET-C, a behavioral approach), fluoxetine, or placebo and results revealed both active treatments to be superior to placebo in reducing social distress and behavioral avoidance and increasing general functioning. SET-C, however, was superior to fluoxetine on these outcomes and was the only treatment superior to placebo in terms of improving social skills, decreasing anxiety in specific social interactions, and enhancing ratings of social competence. Although both treatments are considered effective, SET-C provided added benefit by enhancing social skills.

Across psychosocial and medication trials for separation anxiety, generalized anxiety, and social anxiety, one-third to one-half of children continue to have clinically significant anxiety problems. Although most studies have not been adequately powered to detect factors that influence treatment outcomes, results of the CAMS suggested that a principal diagnosis of social anxiety disorder was associated with less favorable outcomes. Whereas Coping Cat CBT + sertraline yielded superior outcomes across all diagnoses and treatment conditions, stand-alone sertraline produced slightly better outcomes for individuals with social anxiety compared to stand-alone CBT. These findings paired with investigations of SET-C suggest that cognitive-behavioral treatments for social anxiety may need to include social skills training and exposure tasks with age-matched peers in order to produce optimal gains.

Psychotherapy. In addition to the SET-C comparative trial, several wait-list controlled trials of individual CBT with or without family therapy and group-based CBT have been conducted with youths with primary social phobia and show favorable results. Treatment is based on skill development, such as increasing coping self-talk, adopting relaxation

exercises, and using problem-solving approaches. In vivo exposures for social phobia focus on increasing comfort in situations involving unfamiliar people and peers.

Technological advances are also being used to develop innovative, more accessible treatment modalities for social anxiety disorder. One such modality uses virtual reality technology to facilitate graded, in vivo exposures. Within this approach, youth are immersed in virtual environments where they interact with clinician-operated avatars to practice socially challenging tasks (e.g., greeting peers, answering questions, assertiveness skills). The preliminary acceptability and feasibility of this approach has been demonstrated, and further research will assess its efficacy as a stand-alone treatment and adjunct to traditional CBT.

Pharmacology. There have been three studies that have assessed the efficacy of antidepressants in social anxiety disorder—two industry sponsored trials and one federally funded study. Venlafaxine extended-release preparation was found to be more effective than placebo in a large sample of children and adolescents (n = 293) ages 8 to 17 years old. Paroxetine was found to be more effective than placebo in a large sample (n = 322) of children and adolescents ages 8 to 17. Finally the only federally funded study that assessed an SSRI, fluoxetine, and found it to be more effective than placebo was the SET-C study described above.

Overall, the best evidence supports the use of SSRIs in the treatment of childhood anxiety disorders. The value of the SSRIs probably extends across the class so that medication choice can be based on a number of factors including evidence of efficacy in high-quality clinical trials, half-life, risk for drug–drug interactions, any subtle differences in side-effect profiles, and patient, family, and provider preference. The recent FDA indication for duloxetine may soon make it the medication of choice for generalized anxiety and given the high comorbidity of generalized anxiety with separation and social anxiety disorders it may become the first-line treatment for the triad of pediatric anxiety disorders. The TCAs, which can be effective, have not been consistently able to separate from placebo in larger clinical trials, so should be considered only as an alternative for those who have not responded to SSRIs or NSRIs. The value of TCAs in ADHD may make them a second-line choice for children with anxiety disorders and ADHD.

SAFETY. The SSRIs as a class share a number of important side effects that include physical symptoms such as stomachache and headache, and can also be associated with changes in behavior such as activation/disinhibition, insomnia, apathy, manic reactions and the potential for suicidal behavior. Activation/disinhibition (and similar or

related symptoms such as akathisia) are quite common in the pediatric population with 10 to 15 percent difference between SSRI and placebo groups. Activation often occurs early in the course of treatment—days after the initial dose or first meaningful dose increase (e.g., sertraline 25 or 50 mg; fluoxetine 10 or 20 mg), and goes away with medication discontinuation. It appears to be more common in younger children and those whose anxiety co-occurs with symptoms of ADHD, autism spectrum disorders or other early developmental psychopathology. The SSRIs impact sleep in two ways: (1) activation can make it more difficult to fall asleep and (2) SSRIs disrupt sleep architecture, hence users may recover from anxiety, may experience restful sleep, but may sleep more lightly and experience middle of the night awakenings. At higher doses of the SSRIs apathy may occur, leaving patients relatively free of anxiety but unmotivated or feeling disconnected from living.

Perhaps of greatest concern to prescribers of SSRIs in anxious children is the induction of a manic reaction that may herald a new and more severe course of illness. The concern regarding a manic switch is probably greatest when the anxiety disorder is complicated by problems with impulse control or aggressive behavior suggestive of bipolar disorder. Thankfully, manic reactions in the acute phase, large scale, randomized, controlled trials of the SSRIs are exceedingly rare (<1 percent). And even in longer-term follow-up studies the rates are low. Differentiating a manic reaction from the activation side effect can be challenging but an important distinction. Misdiagnosing an activation side effect as a manic reaction may shift the prescriber away from the antidepressants to the use of antipsychotics and mood stabilizer, neither of which are particularly effective for the anxiety disorders. Manic reactions can be differentiated by their presentation, time course of onset, and response to medication discontinuation. Manic symptoms include elevated and expansive mood, grandiosity, decreased need for sleep, and increased goal-directed activities, and symptoms are distinct from the irritability and impulsivity that is commonly seen with activation. Activation occurs early in the course of treatment; manic reactions occur later. With medication discontinuation activation symptoms decrease in relationship to blood level whereas manic symptoms may persist and require medication treatment to reduce symptoms.

Finally, the risk for suicidal behavior emerging or worsening during antidepressant treatment is probably the public's greatest concern. Recent meta-analysis suggests that the suicidal behavior risk difference between the antidepressant and placebo in high-quality clinical trial is <1 percent with numbers needed to harm of >100 and the number needed to treat in high-quality RCTs in depression (NNT = 3–4 in NIMH funded trials), OCD (NNT = 5) and the triad of anxiety disorders (NNT = 3). These data, the observed reduction in antidepressant use with the advent of a black box warning, and the subsequent spike in youth suicide gave rise to concerns

that the black box warning may have had unintended negative consequences. While the black box warning on antidepressant has not been changed, prescribers, but perhaps not the general public, are regaining comfort in the use of the antidepressants.

The side-effect profile of the NSRIs are similar to the SSRIs but now include symptoms consistent with norepinephrine reuptake including nausea, and the potential for initial worsening of anxiety (not activation per se), and for heart rate and blood pressure increases. The side-effect profile of the tricyclic antidepressants are well known and include prominent anticholinergic effects, dry mouth, constipation and urinary retention, and noradrenergic effects on cardiac function, elevated heart rate and orthostatic blood pressure changes, and the potential for prolongation of important cardiac intervals including the QTc.

FUTURE DIRECTIONS

Issues that warrant further clarification pertaining to treatment of youth with separation anxiety, generalized anxiety, or social anxiety include identification of moderators of treatment outcome, identification of critical elements of psychotherapy that contribute to outcome, questions of treatment sequencing and whether youth who fail to respond to either medication or CBT as their first treatment will respond to the other treatment, strategies for enhancing outcomes and addressing treatment resistance, and elucidating long-term outcomes particularly following medication treatment. A follow-up study of the CAMS sample, now underway, should provide information about long-term diagnostic and functional outcomes following receipt of evidence-based treatment, as well as patterns of posttreatment service utilization.

REFERENCES

Beidel DC, Turner SM, Sallee FR, Ammerman RT, Crosby LA. SET-C versus fluoxetine in the treatment of childhood social anxiety disorder. *J Am Acad Child Adolesc Psychiatry*. 2007;46:1622–1632.

Biederman J, Petty CR, Hirshfeld-Becker DR, Henin A. Developmental trajectories of anxiety disorders in offspring at high risk for panic disorder and major depression. *Psychiatry Res*. 2007;153:245–252.

Bittner A, Egger HL, Erkanli A, Costello EJ, Foley DL. What do childhood anxiety disorders predict? *J Child Psychol Psychiatry*. 2007;48:1174–1183.

Bolton D, Eley TC, O'Connor TG, Perrin S, Rabe-Hesketh S. Prevalence and genetic and environmental influences on anxiety disorders in 6-year-old twins. *Psychol Med*. 2006;36:335–344.

Buckner JD, Schmidt NB, Lang AR, Small JW, Schlauch RC. Specificity of social anxiety disorder as a risk factor for alcohol and cannabis dependence. *J Psychiatr Res*. 2008;42:230–239.

Compton SN, March JS, Brent D, Albano AM, Weersing VR. Cognitive-behavioral psychotherapy

for anxiety and depressive disorders in children and adolescents: an evidence-based medicine review. *J Am Acad Child Adolesc Psychiatry*. 2004;43:930–959.

*Connolly SD, Bernstein GA, Bernet W, Bukstein O, Arnold V. Practice parameter for the assessment and treatment of children and adolescents with anxiety disorders. *J Am Acad Child Adolesc Psychiatry*. 2007;46:267–283.

Costello EJ, Egger HL, Angold A. The developmental epidemiology of anxiety disorders: phenomenology, prevalence, and comorbidity. *Child Adolesc Psychiatr Clin North Am*. 2005;14:631–648.

Ginsburg GS, Becker EM, Keeton CP, et al. Naturalistic follow-up of youths treated for pediatric anxiety disorders. *JAMA Psychiatry*. 2014;71:310–318.

Gleason MM, Egger HL, Emslie GJ, Greenhill LJ, Kowatch RA. Psychopharmacological treatment for very young children: contexts and guidelines. *J Am Acad Child Adolesc Psychiatry*. 2007;46:1532–1572.

Gosch EA, Flannery-Schroeder E, Mauro CF, Compton SN. Principles of cognitive-behavioral therapy for anxiety disorders in children. *J Cogn Psychother*. 2006;20:247.

Hirshfeld-Becker DR, Biederman J, Henin A, Faraone SV, Davis S. Behavioral inhibition in preschool children at risk is a specific predictor of middle childhood social anxiety: a five-year follow-up. *J Dev Behav Pediatr*. 2007;28:225–233.

Kendall PC, Compton SN, Walkup JT, et al. Clinical characteristics of anxiety disordered youth. *J Anxiety Disord*. 2010;24:360–365.

Kendall PC, Suveg C. Treating anxiety disorders in youth. In: Kendall PC, ed. *Child and Adolescent Therapy: Cognitive-Behavioral Procedures*. 3rd ed. New York: Guilford; 2003:375.

Kessler RC, Avenevoli S, Costello EJ, et al. Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Arch Gen Psychiatry*. 2012;69:372–380.

Kessler R, Petukhova M, Sampson N, Zaslavsky A, Wittchen H. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int J Methods Psychiatr Res*. 2012;21:169–184.

Manassis K, Hudson JL, Webb A, Albano AM. Beyond behavioral inhibition: etiological factors in childhood anxiety. *Cogn Behav Pract*. 2004;11:3.

Manassis K, Lee TC, Bennett K, et al. Types of parental involvement in CBT with anxious youth: a preliminary meta-analysis. *J Consulting Clin Psychol*. 2014;82:1163–1172.

*Morris TL, March, JS. *Anxiety Disorders in Children and Adolescents*. 2nd ed. New York: Guilford Press; 2004.

Negreiros J, Miller LD. The role of parenting in childhood anxiety: etiological factors and treatment implications. *Clin Psychol (New York)*. 2014;21:3.

*Reynolds S, Wilson C, Austin J, Hooper L. Effects of psychotherapy for anxiety in children and adolescents: a meta-analytic review. *Clin Psychol Rev*. 2012;32:251–262.

Rooksby M, Elouafkaoui P, Humphris G, Clarkson J, Freeman R. Internet-assisted delivery of cognitive behavioural therapy (CBT) for childhood anxiety: systematic review and meta-analysis. *J Anxiety Disord*. 2015;29:83–92.

Rynn M, Puliafico A, Heleniak C, Rikhi P, Ghalib K, Vidair H. Advances in pharmacotherapy for pediatric anxiety disorders. *Depress Anxiety*. 2011;28:76–87.

*Strawn JR, Welge JA, Wehry AM, Keeshin B, Rynn MA. Efficacy and tolerability of antidepressants in pediatric anxiety disorders: a systematic review and meta-analysis. *Depress Anxiety*. 2015;32(3):149–157.

Vasey MW, Dadds MR. *The Developmental Psychopathology of Anxiety*. New York: Oxford University Press; 2001.

Velting ON, Setzer NJ, Albano AM. Update on and advances in assessment and cognitive-behavioral treatment of anxiety disorders in children and adolescents. *Prof Psychol Res Pract*. 2004;35:42.

*Walkup JT, Albano AM, Piacentini J, et al. Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *N Engl J Med*. 2008;359(26):2753–2766.

Warren SL, Umylny P, Aron E, Simmens SJ. Toddler anxiety disorders: a pilot study. *J Am Acad Child Adolesc Psychiatry*. 2006;45:859–866.

Wittchen HU, Fehm L. Epidemiology and natural course of social fears and social anxiety disorder. *Acta Psychiatr Scand*. 2003;108:4–18.

Yap MH, Pilkington PD, Ryan SM, Jorm AF. Parental factors associated with depression and anxiety in young people: a systematic review and meta-analysis. *J Affect Disord*. 2014;156:8–23.

▲ 52.4 Selective Mutism

R. LINDSEY BERGMAN, PH.D., AND ARACELI GONZALEZ, PH.D.

INTRODUCTION

Selective mutism (SM) is a childhood behavioral disorder characterized by a persistent failure to speak in one or more specific social situations despite speaking in other situations. Although the syndrome was recognized as long ago as 1877, the disorder (earlier called elective mutism) did not appear in the *Diagnostic and Statistical Manual of Mental Disorders* until its third edition (DSM-III) in 1980. Current conceptualizations of SM emphasize the relationship between considerable social anxiety and the failure or extreme reluctance to speak. In addition, researchers continue to investigate the presence of speech or language anomalies as well as oppositional symptoms among children with SM. Although there has been increased interest related to SM recently, in comparison with many other childhood emotional or behavioral disorders, there is a relative lack of consensus on many issues related to the epidemiology, phenomenology, and treatment of the disorder.

DEFINITION

The DSM-5 defines SM as a consistent failure to speak in specific social situations where speaking is expected (e.g., to peers at school), despite speaking regularly in other situations. DSM-5 also specifies that in order

for the diagnosis to be assigned, the lack of speech must interfere with achievement or communication, must last at least 1 month (beyond the first month of school), and must not be due to a lack of knowledge or comfort with the language required in a given situation. Finally, DSM-5 specifies that the lack of speech cannot be better accounted for by a communication, developmental, or psychotic disorder.

HISTORY

In the 19th century, the term *aphasia voluntaria* was first noted in the psychiatric literature and was used to describe children who voluntarily chose not to speak in certain situations. Many other terms including voluntary mutism, speech phobia, suppressed speech, functional mutism, and speech avoidance have also been used to describe children who meet the current criteria for SM. The term *elective mutism* was first used in 1934 and appeared in DSM-III in 1980. Some believed that descriptors such as "voluntary" and "elective" that appeared in these early terms suggested a volitional choice to refrain from speaking, and possibly, an oppositional attitude. This attitude was also reflected in the DSM-III definition of elective mutism, which contained a reference to a "refusal" to speak in the definition of SM. In DSM-III-R, elective mutism was renamed SM, perhaps reflecting shifting conceptualizations of the disorder such that attention was more focused on social anxiety as opposed to oppositional or volitional aspects of the nonspeaking behavior.

COMPARATIVE NOSOLOGY

Consistent with the shifting conceptualization described above, SM was reclassified as an anxiety disorder in the current edition of the DSM (DSM-5). In the DSM-IV, SM was included in the section Disorders Usually First Diagnosed in Infancy, Childhood, or Adolescence. In the DSM-5, disorders within that category were reorganized into different categories and SM is now included as an Anxiety Disorder. This reclassification of SM in the Anxiety Disorders section of the DSM-5 is a logical fit given growing evidence for the phenomenological similarities between SM and social anxiety (see Etiology section). However, the nature of relationship between SM and social anxiety is complex and requires further research as discussed below. Other than the relocation of the SM diagnosis from Disorders Usually First Diagnosed in Infancy, Childhood, or Adolescence to the Anxiety Disorders section, the diagnostic criteria for SM are virtually identical in the DSM-IV and -5. Interestingly, a diagnosis of SM does not require the presence of fear or anxiety.

The disorder was not renamed in the tenth edition of the *International Statistical Classifications of Diseases and Related Health Problems* (ICD-10), and the current version describes *elective mutism* as a disorder characterized by a "marked, emotionally determined selectivity in

speaking," such that the child demonstrates a language competence in some situations but fails to speak in other situations. The disorder is also described as associated with social anxiety, withdrawal, sensitivity, or resistance.

EPIDEMIOLOGY

SM is often assumed to be rare. However, until recently, estimates of the occurrence of SM were based on scant evidence that was often inaccurate. Only a handful of community-based prevalence studies have been published. These studies, all but one conducted in Europe, have yielded widely disparate findings due to different sampling strategies and definitions of SM. A United Kingdom study identified 0.69 percent of 4- to 5-year-olds as completely mute (i.e., not speaking at all) at school. However, this rate dropped to 0.08 percent after approximately 8 months of school. Another survey, also in the United Kingdom, identified only 2 of 3,300 (0.06 percent) 7-year-olds as selectively mute at 36 months of age. More recent investigations have focused on the school environment to investigate the prevalence of SM. Two Scandinavian surveys found teacher-reported rates of DSM-IV-TR SM of 0.18 and 1.9 percent, respectively, and another study with Irish primary school children ages 4 to 12 years also found a prevalence rate of 0.18.

One US community prevalence study surveyed kindergarten, first-, and second-grade classrooms in an urban public school district and reported a DSM-IV-TR SM prevalence rate of 0.71 percent (16 of 2,256). A study in Israel yielded comparable results. These recent studies provide evidence that SM may be more common than is frequently recognized. In fact, it is likely that the point prevalence rate for SM is comparable to rates of other childhood psychiatric disorders, including obsessive–compulsive disorder (0.5 to 1 percent), and major depressive disorder (0.4 to 3 percent). It has also been reported that immigrant and bilingual or language minority children are at higher risk of SM, and prevalence of the disorder may be three times higher in these populations.

Overall prevalence does not seem to vary markedly by sex or race/ethnicity; however, there is evidence that slightly more girls than boys seem to be affected by SM, with an average reported female to male ratio of 1.2 to 1. There is some evidence to suggest that this gender difference may become more pronounced among older children. Although it is commonly reported that children with SM show significant symptoms from the age at which they spoke fluently, with the age of onset typically reported from 2 to 5 years old, SM is often not noticed until the time of entry into school.

ETIOLOGY

The conceptualization of SM has undergone a considerable shift over the past several years. At this point in time, evidence suggests that SM results

from multiple developmental pathways involving the complex interplay of genetic, neurobiological, temperament, and behavioral (e.g., avoidance) variables. Previously, it was often suggested that SM was related to a variety of conditions including oppositionality, trauma, or family neuroses. However, current conceptualizations view the disorder as closely related to, or even a developmental variant of, social phobia. Strong support for this position is derived from studies reporting that over 95 percent of children with SM also met DSM-III-R criteria for social phobia. Additional evidence comes from the finding that pharmacological agents used to treat adult social phobia are similarly effective in the treatment for SM. Last, family history data reveal an elevated rate of social phobia among parents of children with SM, and severity of children's SM symptoms may predict parents' diagnosis of generalized social phobia. These findings suggest a familial relationship between SM and generalized social phobia and provide further support that SM may be a developmental precursor to social phobia. Research regarding genetic risk of SM is in its infancy. One recent study examining genetic susceptibility factors for SM in families with social anxiety-related traits found evidence of increased risk of SM in individuals with a genetic variation in CNTNAP2, a gene implicated in various forms of language impairment.

Researchers and clinicians have also noted the resemblance between SM children and children described as "behaviorally inhibited." In fact, reluctance to speak is often considered to be a sensitive index of behavioral inhibition. Consistent with this is the finding that parents of children with SM rate their children as not responding well to new situations or handling transitions well, both widely considered to be fundamental elements of behavioral inhibition. Since there is evidence that the presence of behaviorally inhibited temperament is associated with increased risk of future anxiety disorders, it is possible that behavioral inhibition is an etiological factor in the development of SM. By contrast, a small proportion of children with SM do not exhibit social phobia with regard to nonverbal communication and will engage in social interaction using gestures, facial expressions, body language, and other socially engaged behavior. Such cases raise questions about the nature of the relationship between SM and social phobia, suggesting that some individuals may experience anxiety specifically related to expressive language and not to social situations more generally. Some researchers have investigated the notion that the "nonspeaking behavior" of children with SM essentially should be seen as an emotion regulation strategy aimed to reduce anxiety around social interaction.

As mentioned previously, oppositionality and recent trauma are often cited as etiological factors related to SM; however, studies utilizing systematic diagnostic procedures do not provide support for these relationships. In comparing children who met criteria for only social

phobia to those with both SM and social phobia and those with no disorder, researchers found no significant differences on parental reports of oppositionality. Furthermore, mean scores for ratings of oppositionality were in the normal range for children with SM. Thus, findings suggest that the majority of children with SM do not present with clinically significant oppositional behaviors. However, 29 percent of children with SM were rated as meeting criteria for oppositional defiant disorder (ODD) by clinicians. Thus, a subset of children with SM may exhibit comorbid ODD.

Somewhat more confusing are conflicting findings related to the prevalence of communication disorders among children with SM and the association between developmental delay and SM. Several studies have examined possible deficits in auditory or speech and language processing in children with SM. In one study, children with SM were compared to children with an anxiety disorder and normal controls on measures of language, nonverbal working memory, and social anxiety. When compared to children with other anxiety disorders and normal controls, children with SM had lower scores on language measures and deficits on tasks requiring them to remember verbal information. In another study, researchers found that children with SM did not exhibit deficits in visual memory, but they did show significantly poorer verbal memory than controls. Thus, children with SM may exhibit meaningful deficits in speech and language processing.

Although some children with SM may experience language delay or dysfunction and deficits in speech and language processing, it is noteworthy that among researchers it is not suggested that these problems supersede social anxiety as a major contributing factor toward nonspeaking. Furthermore, although language delay or speech disorder may be more common among children with SM, these problems could be a result, rather than a cause, of the disorder. Finally, speech and language problems associated with SM are likely to be subtle and are exclusionary criteria for the diagnosis of SM when they cannot be ruled out as the cause of the lack of speech. Ultimately, however, it is important to note that many in the field believe that neurodevelopmental and language anomalies occur more frequently among children with SM than among children with other anxiety related disorders.

DIAGNOSIS AND CLINICAL FEATURES

The primary symptom exhibited by children with SM is the failure to speak in some situations where speaking is expected despite speaking in other situations. The situation in which the child has no difficulty speaking is almost always in the home environment. In fact, families of children with SM often report that the child seems to be “overly” talkative at home. However, in severe cases, there may be some members of the household to which the child with SM does not speak. If a child who is presumed to have

SM does not speak at all at home, there is reason to seriously question the diagnosis of SM.

Several studies indicate that the situation in which speaking is most commonly avoided is at school. Within this common pattern (speech at home, failure to speak at school), there is large individual variation in terms of speaking patterns among children with SM. In the school environment, children with SM more commonly fail to speak to adults (e.g., teachers) than to peers, although a lack of speech with peers is not uncommon. However, many children with SM have individually distinct and stable patterns regarding to whom they do, and do not, speak. For instance, there are often particular extended family members with which the child does not speak and others to whom they consistently speak. Some children with SM have little or no difficulty speaking with strangers, whereas others place strangers at the top of their hierarchy of people with whom it is difficult to speak.

The label of SM can be used somewhat broadly to describe children who are consistently reluctant to speak or display low frequency of speech. Although some argue that children who speak inconsistently, reluctantly, or with low volume do not, in fact, have SM, these speaking behaviors are relevant for the diagnosis of SM because children who completely fail to speak in some situations often speak inconsistently, with low-frequency or volume, or reluctantly in others. For example, some children with SM may fail to speak at all in some situations, while in others they do not speak in a "normal" voice but whisper fairly readily. Also fairly common is the child who does not speak at all in some situations, but in other situations speaks in a "disguised" or odd voice that is mechanical, gruff, or babyish.

Given that SM appears to be so closely related to social phobia, it is not surprising that children with SM also have difficulty with nonverbal but socially relevant tasks such as eating in front of others, having their picture taken, and communicating nonverbally. When social anxiety symptoms not related to speech are present and meet criteria for social anxiety disorder, the additional diagnosis should be given. Even among those who are not also diagnosed with social anxiety disorder, children with SM are typically described as shy and self-conscious. However, it is important to note that not all children with SM are shy; although not typical, there are children with SM who present as outgoing, friendly and eager to interact but only interact nonverbally. This presentation has led some researchers to view SM as a type of specific phobia of speaking.

The following case exemplifies a common presentation of a child with SM.

James is a 5-year-old Chinese American boy who lives with his biological mother, father, and older sister. James's parents reported that he speaks normally at home but that he started to have difficulty speaking to adults and children at school upon entry to preschool. He continued to have difficulty speaking throughout his preschool year and has ongoing difficulty in

kindergarten. At home, his parents report that he is extremely talkative and outgoing when he is around immediate family members and extended family members close to his age. At school, James does not speak to any of the adults but answers classmates with one-word answers and speaks only in a whisper. By his parents' report, James also exhibits extreme social anxiety, to the point of "freezing" in certain situations when attention was focused on him. He often watches other children playing from a distance, does not join in group activities, and does not participate in circle time at school. James's parents also report that he hides behind them when strangers are around and that he dislikes having his picture taken. At the time of his evaluation, James had not received prior treatment for SM or any other emotional or behavioral disorders. James was born in the United States and speaks fluent English and limited Mandarin. His parents speak to him in both English and Mandarin and he mostly responds to them in English. According to his parents, he developed normally with no delays and appears to have above-average intelligence. They also reported that James enjoys riding on his bicycle, singing, and art activities.

During initial evaluation, James hid behind his parents and failed to respond in any way to the intake clinician. James's parents reported that this behavior is typical of him when in a novel situation with new people, but that he might answer with one-word answers softly, communicate nonverbally, and make eye contact with people once he gets more comfortable with them. James's parents reported that he speaks spontaneously and fluently when he is at home with his family. They also reported that he can be quite animated, is a lot of fun, and is very talkative at home. James received diagnoses of SM (moderately severe) and social phobia (moderately severe). Behavioral treatment was recommended.

Behavioral treatment was initiated following the intake evaluation. Initially, the therapist played nonverbal games and engaged in nonverbal activities with James to help him feel comfortable in the session (e.g., coloring, blowing bubbles, hiding pennies, and giving nonverbal clues). Next, James and his mother were asked to create a list of easy, medium, and hard speaking situations and a list of small, medium, and large rewards. These lists then became the basis for assignments (and reinforcement) for speaking tasks that gradually increased in difficulty. In general, sessions included time with James and his mother together to review past and future assignments as well as time with James and the therapist alone.

When treatment began, James did not communicate at all verbally with the therapist. The therapist gradually increased rapport (via unstructured play) while having James try increasingly difficult tasks, such as first whispering to his mother with the therapist in the corner, then nodding yes or no in response to the therapist's questions, pointing, using a puppet to "talk for him," whispering to his mother while facing the therapist, and then subsequently responding verbally to the therapist. A successful technique in working with James was having him whisper to his mother while playing games with the therapist and then having his mother move a few inches away each time he was asked to respond to the therapist. As the whisper became more audible, the therapist was able to hear his speech and responded directly to the whisper (without waiting for his mother to tell the therapist what James said). James became more comfortable with the therapist hearing his voice, and he was slowly able to direct his speech toward the therapist. After three sessions, James consistently and spontaneously talked to the therapist. James received stickers for completing each speaking assignment, and after filling up "sticker charts," he received rewards (i.e., small toy, special activity with parent, or treat from the reward list).

James was also given assignments that involved both his teacher and classmates. These were implemented in a gradual fashion and included waving to the teacher, playing a game with his parent first in an empty classroom then while his teacher was in the classroom, leaving a phone message for the teacher, playing an audiotape of him saying "hello" to the teacher, whispering to his parent loudly enough for his teacher to hear, whispering "hello" to the teacher, and speaking "hello" to the teacher in a regular voice. Assignments also involved classmates and included playing games with his parent and one classmate, asking classmates questions about their favorite things (e.g., color, animal, and food), playing games with classmates without his parent, and play dates at home. After approximately 12 sessions, James spoke a complete sentence to a group of three classmates, spoke to his teacher directly, and answered a question aloud in front of the entire class.

During the last few sessions, which were tapered over several weeks, James's mother took an increasingly primary role in assigning and following up on speaking assignments. James's

mother had him continue practicing talking assignments with extended family members and unfamiliar adults and children. James entered the first grade, and after only a few days, began regularly speaking to his teacher (both privately and in front of peers) and to most peers in class. Following completion of therapy, James's mother continued to monitor his speaking behaviors and to promote speaking in new situations by encouraging (and rewarding) James's gradual successes with novel people and situations.

DIFFERENTIAL DIAGNOSIS

Although children with SM can be misdiagnosed as having autism due to their lack of speech in the clinic setting, this is a readily avoided diagnostic error. The language deficits of a child with autism are not dependent on the situation, whereas the lack of speech associated with SM is by definition limited to certain situations. Similarly, the diagnostic criteria requiring that the child speak normally in some situations can often rule out a severe communication disorder.

Although many perceive the child with SM to be oppositional, distinguishing SM from primary ODD is fairly straightforward. Most importantly, the child with SM is unlikely to engage in significant oppositional behaviors outside of the context of verbal communication. If a child who appears to have SM (failure to speak in certain situations despite speaking in other situations) also exhibits oppositional behaviors in a variety of situations, both diagnoses are given. For children whose failure to speak does not appear to follow a consistent pattern with regard to situation and targeted individuals, appears primarily motivated by instrumental gains, and occurs in the absence of any other anxiety symptoms, it is possible that ODD rather than SM is the correct diagnosis.

As discussed above, current understanding of SM emphasizes that social anxiety underlies most selectively mute behavior. The resulting conceptual and diagnostic overlap between SM and social phobia complicates the clinician's determination of whether a child with SM also meets criteria for the presence of comorbid social phobia. One of the most successful strategies for the clinician attempting to make this determination is to direct queries toward the presence of social phobia symptoms that do not involve speaking. For example, the presence of fear and avoidance of eating in public, having pictures taken, using public restrooms, or writing in front of others provide strong evidence for an additional diagnosis of social phobia in a child who meets criteria for SM.

The DSM-5 criteria state that a child should not be diagnosed with SM if the failure to speak is due to a lack of knowledge of, or comfort with, the spoken language required. Researchers in the area of bilingual children with SM report that immigrant children and those learning a second language in their new country may be overlooked or misdiagnosed. Thus, the diagnosis of immigrant and language minority children can be difficult, and it is important to understand normal bilingual child development to diagnose SM properly. Many children learning a second language may go

through a nonverbal or "silent period" due to the realization that their native language is not understood by classmates and teachers and that their skills in the new language are not sufficient. To diagnose SM in bilingual children and children who are acquiring a second language, it is important to know that bilingual children with SM exhibit a failure to speak in both their native and newly acquired languages, in several unfamiliar settings, and for an extended period of time (>6 months). Although a child learning a new language may not feel completely comfortable speaking a new language for several months most children show no progression of speech as they gain mastery over their new second language, whereas a child with SM typically does not.

COURSE AND PROGNOSIS

SM appears to have an insidious onset, with most parents reporting failure to speak as a problem that was evident from the time their child began talking. However, commonly, the lack of speech outside of the home is not considered to be a problem until school entry. Although SM is generally considered to be a disorder that is difficult to successfully treat, at the same time it is believed to spontaneously remit over time in a large majority of cases. Based in part on one study that found a large reduction in SM over a 6-month period (from beginning to end of the school year), it was believed that the disorder spontaneously remitted over time in a large number of cases and that in these cases, and that SM should be described as transient rather than persistent. This distinction has been refined somewhat based on a school-based study that revealed that although close to 30 percent of children with SM showed some improvement from the beginning to the end of a school year, as a group, the SM children remained significantly more symptomatic than matched unaffected peers.

In a controlled long-term outcome study (mean follow-up period of 13 years), young adults with SM in childhood were compared to young adults with anxiety disorders in childhood and young adults with no psychiatric diagnosis in childhood. Researchers found that 18 percent of patients with SM in childhood showed only slight improvement in young adulthood, while the rest of the patients with childhood SM were found to be either markedly improved or totally improved in young adulthood. They reported a complete remission rate of 58 percent. Unfortunately, these data did not include comparisons related to variation in course with or without treatment. In the only controlled treatment study that assessed youngsters 1 year following successful behavioral treatment reported that children maintained or improved upon initial gains when reassessed.

Although it seems intuitively useful to discriminate between transient and persistent SM and to assume that persistent mutism may be related to more severe impairment, this distinction requires further investigation. Similarly, there is no information related to ongoing difficulties that might

result from even relatively short-term or transient SM. Given that that there is little information regarding lasting consequences related to early disturbance in verbal communication and that nonspeaking behavior is often conceptualized as a learned behavior that is reinforced considerably over time, it makes sense to target treatment toward younger, recent onset youngsters. Additional support for the importance of early intervention (vs. delaying treatment to establish if SM will resolve spontaneously) comes from recent findings that age of onset, duration of SM, and age at treatment all predict treatment effect size. Specifically, young age and shorter duration of SM were associated with larger treatment effects.

TREATMENT

Unfortunately, much of the published data on the treatment of SM consists of single case studies and case series of markedly varying quality and studies based on retrospective reports or records. Indeed, several additional studies reporting SM treatment outcome rely on small samples or single cases and do not identify diagnostic procedures, assessment or outcome methods, number of treatment sessions, or details of the treatment method. In spite of these limitations, the conviction that behavioral techniques are an essential component of treatment for SM is widespread. Recently, two small randomized controlled trials of behavioral treatment for SM were completed and support the efficacy of such interventions compared to waitlist control. Generally, these and other published reports in the literature describing treatment efforts provide a solid basis for the development of behavioral treatments that include parent and school involvement in treatment. In addition to behavioral strategies, a small but growing literature provides support for the efficacy of pharmacotherapeutic intervention (primarily the selective serotonin reuptake inhibitors [SSRIs]). In addition to a growing effort aimed at empirical validation of treatment for SM, recent efforts have also resulted in better instruments for the assessment of specific symptoms of SM. Instruments such as the Selective Mutism Questionnaire, a parent self-report measure of the affected child's lack of speech, and the School Speech Questionnaire, the companion teacher measure are commonly used in assessment of symptoms in both research and clinical settings.

Behavioral Treatment

Multiple published reports describe the successful utilization of relatively standard behavioral techniques such as contingency management, stimulus fading, systematic desensitization, negative reinforcement, and shaping for the treatment of SM. A combination of behavioral techniques is probably the most common and successful treatment approach. Behavioral approaches typically utilize a hierarchy or rank-ordered list of situations in

which the child has difficulty speaking. Then, the child is guided to systematically engage in speaking-related behaviors (e.g., mouthing speech, making sounds, and whispering) in increasingly more difficult situations. Over repeated successful attempts, the associated anxiety dissipates through the process of autonomic habituation. In addition, when the feared consequences of speaking fail to occur, patients' heightened expectations of harm disappear, reducing anxiety even further. Typically, the child receives positive reinforcement (rewards) following attempts to engage in speaking-related behaviors in order to shape the child's behavior. There are also several reports of treatments utilizing somewhat unique self-modeling behavioral interventions including audio and video editing techniques. Using these techniques, audio or videotapes are edited to show the child speaking in settings where the child usually does not speak. Then, these tapes are repeatedly played for the child so that the child habituates hearing him- or herself speak in these settings and no longer feels uncomfortable.

A number of considerations serve to complicate the behavioral treatment of SM and to distinguish it from treatment for social phobia, a disorder to which it is closely related. These factors include the child's failure to speak (even to the therapist), the young age of most SM patients, and the need for significant school involvement in almost all cases. The young age of children with SM necessitates considerably more parental involvement in treatment than is typically recommended for other anxiety disorders.

Although not without exception, most children with SM do not initially speak to the treating clinician. Obviously, a treatment approach for SM needs to consider the initial communication difficulties that are likely to be present and to include strategies for facilitating verbal communication with the therapist. The initial lack of verbal communication with the therapist is assumed to reflect the general anxiety underlying failure to talk in situations outside the home and is addressed accordingly. To decrease this anxiety and promote speaking in the treatment setting, traditional anxiety-reducing behavioral techniques (shaping, gradual exposure, reinforcement) are first introduced within the session.

The SM child's failure to speak represents a somewhat unique challenge for others in the child's life, particularly parents and school personnel. Although difficulty in differentiating oppositional behavior from anxiety-related avoidance characterizes many child anxiety disorders, when failure to speak is involved, parents and others are particularly likely to attribute avoidance to oppositional motivations and to get frustrated and angry. In contrast, other reactions to nonspeaking include the tendency to accommodate the child's lack of speech and speak "for" the SM child, thus decreasing the necessity of independent communication. This tendency, along with other ways that nonspeaking may be reinforced, also needs to be

addressed in the treatment of the disorder.

Nearly all forms of youth anxiety involve school-related impairment to some extent. However, unlike SM, they do not typically revolve as extensively around the school environment. For this reason, treatment that integrates school participation as a central component of the intervention is necessary for the effective treatment of SM. Most importantly, close rapport with school personnel that includes regular communication and support must be established to ensure the success of behavioral interventions in the school setting. Furthermore, treatment programs should include teacher education and training in the behavioral methods utilized for treatment. The utilization of teachers in the treatment of SM has been highlighted as a particularly cost-effective treatment approach.

Other Psychosocial Therapies

Although the most common accounts of successful psychosocial treatment for SM involve behavior therapy, there are descriptions in the literature of successful treatment for SM with play therapy, family therapy, psychodynamic therapy, and group therapy. Although most of these descriptions are flawed by their small sample sizes and failure to adequately describe the diagnostic or treatment intervention processes, it is noteworthy that descriptions of play, family, and group therapy all contain considerable emphasis on the use of behavioral interventions. In general, traditional psychotherapy aimed at addressing underlying intrapsychic conflicts or other interventions (including play therapy, art therapy, supportive therapy, and family therapy) not containing the behavioral elements described above have not been found to be effective for SM.

Speech and Language Therapy

As noted by the DSM-5, the criteria for SM include a caveat that the symptoms cannot be primarily due to a speech or language disorder. However, it is not uncommon for children with SM to also have some degree of speech or language difficulty that may exacerbate speech-related anxiety or self-consciousness. Although there is a lack of research on the efficacy of multimodal approaches, the combination of cognitive-behavioral techniques and speech or language therapy can be quite effective for the treatment of children with comorbid SM and speech or language difficulties. In addition, even in cases without significant contribution of speech or language dysfunction, the small supportive group setting of auxiliary speech services found in many schools may provide a beneficial environment in which to implement the graded behavioral intervention approach described earlier.

Psychopharmacological Treatments

Currently, there are no approved U.S. Food and Drug Administration (FDA) medications for the treatment of SM. However, several small studies suggest that SSRIs may be effective treatments for SM in children. This is not surprising given the central role of anxiety in the phenomenology of the disorder. Results from a single double-blind, placebo-controlled trial of fluoxetine in children with SM indicated significant improvement on parent ratings of mutism and anxiety. Results from open trials and case studies have also supported the efficacy of various SSRI treatments. In addition, a large multisite study of fluvoxamine for the treatment of separation anxiety disorder, social phobia, and generalized anxiety disorder in children included several children with SM who appeared to benefit from this treatment. These results, along with the mounting evidence that SM is closely related to social phobia, have begun to establish SSRIs as an important treatment option for children with SM. However, given concerns about medication use in young children, medication is not a first option for many families. In practice, behavioral therapy, when practical and available, should be considered as the first-line intervention strategy. In treatment resistant cases, combination treatment (behavioral therapy plus medication) appears to be the strategy of choice.

References

*Bergman RL, Gonzalez A, Piacentini J, Keller ML. Integrated Behavior Therapy for Selective Mutism: a randomized controlled pilot study. *Behav Res Ther*. 2013;51(10):680–689.

*Bergman RL, Keller ML, Piacentini J, Bergman AJ. The development and psychometric properties of the selective mutism questionnaire. *J Clin Child Adolesc Psychol*. 2008;37:456–464.

Bergman RL, Piacentini J, McCracken JT. Prevalence and description of selective mutism in a school-based sample. *J Am Acad Child Adolesc Psychiatry*. 2002;41(8):938–946.

Black B, Uhde TW. Treatment of elective mutism with fluoxetine: a double blind, placebo-controlled study. *J Am Acad Child Adolesc Psychiatry*. 1994;33:1000–1006.

*Carlson JS, Mitchell AD, Segool N. The current state of empirical support for the pharmacological treatment of selective mutism. *School Psychol Quart*. 2008;23:354.

Chavira DA, Shipon-Blum E, Hitchcock C, Cohan S, Stein MB. Selective mutism and social anxiety disorder: all in the family? *J Am Acad Child Adolesc Psychiatry*. 2007;46(11):1464–1472.

Cohan SL, Chavira DA, Stein MB. Practitioner review: psychosocial interventions for children with selective mutism: a critical evaluation of the literature from 1990–2005. *J Child Psychol Psychiatry*. 2006;47(11):1085–1097.

Dummit ES 3rd, Klein RG, Tancer NK, Asche B, Martin J. Fluoxetine treatment of children with selective mutism: an open trial. *J Am Acad Child Adolesc Psychiatry*. 1996;35(5):615–621.

Dummit ES 3rd, Klein RG, Tancer NK, Asche B, Martin J, Fairbanks JA. Systematic assessment of 50 children with selective mutism. *J Am Acad Child Adolesc Psychiatry*. 1997;36:653–660.

Holmbeck GN, Lavigne JV. Combining self-modeling and stimulus fading in the treatment of an

electively mute child. *Psychotherapy*. 1992;29(4):661.

Letamendi AM, Chavira DA, Hitchcock CA, Roesch SC, Shipon-Blum E, Stein MB. Selective Mutism Questionnaire: measurement structure and validity. *J Child Psychol Psychiatry*. 2008;47:1197–1204.

Manassis K, Tannock R, Garland EJ, Minde K, McInnes A, Clark S. The sounds of silence: language, cognition, and anxiety in selective mutism. *J Am Acad Child Adolesc Psychiatry*. 2007;46(9):1187–1195.

*Muris P, Ollendick TH. Children who are Anxious in Silence: A Review on Selective Mutism, the New Anxiety Disorder in DSM-5. *Clin Child Fam Psychol Rev*. 2015;18:151–169.

Oerbeck B, Stein MB, Wentzel-Larsen T, Langsrud Ø, Kristensen H. A randomized controlled trial of a home and school-based intervention for selective mutism—defocused communication and behavioural techniques. *Child Adolesc Ment Health*. 2014;19(3):192–198.

Omdal H, Galloway D. Could selective mutism be re-conceptualised as a specific phobia of expressive speech? An exploratory post-hoc study. *Child Adolesc Ment Health*. 2008;13:74–81.

Piacentini J, Bergman RL. Developmental issues in the cognitive treatment of childhood anxiety disorders. *J Cogn Psychother*. 2001;15:165–182.

Research Unit on Pediatric Psychopharmacology Anxiety Study Group. Fluvoxamine for the treatment of anxiety disorders in children and adolescents. *N Engl J Med*. 2001;344:1279–1285.

*Scott S, Beidel DC. Selective mutism: an update and suggestions for future research. *Curr Psychiatry Rep*. 2011;13(4):251–257.

Sharkey L, McNicholas F. Selective Mutism: a prevalence study of primary school children in the republic of Ireland. *Ir J Psychol Med*. 2012;29(01):36–40.

Sharp WG, Sherman C, Gross AM. Selective mutism and anxiety: a review of the current conceptualization of the disorder. *J Anxiety Disorder*. 2007;21:568–579.

Stein MB, Yang B-Z, Chavira DA, Hitchcock CA, Sung SC, Shipon-Blum E, et al. A common genetic variant in the neurexin superfamily member CNTNAP2 is associated with increased risk for selective mutism and social anxiety-related traits. *Biol Psychiatry*. 2011;69:825–831.

Steinhausen HC, Juzi C. Elective mutism: an analysis of 100 cases. *J Am Acad Child Adolesc Psychiatry*. 1996;35:606–614.

Steinhausen H-S, Wachter M, Laimbock K, Metzke CW. A long-term outcome study of selective mutism in childhood. *J Child Psychol Psychiatry*. 2006;47(7):751–756.

Stone BP, Kratochwill TR. Treatment of selective mutism: a best-evidence synthesis. *School Psychol Quart*. 2002;17(2):168.

Toppelberg CO, Tabors P, Coggins A, Lum K, Burger C. Differential diagnosis of selective mutism in bilingual children. *J Am Acad Child Adolesc Psychiatry*. 2005;44:592.

Vecchio J, Kearney C. Selective mutism in children: a synopsis of characteristics, assessment, and treatment. *Behav Ther*. 2006;29(5):102.

Yeganeh R, Beidel DC, Turner SM. Selective mutism: more than social anxiety? *Depress Anxiety*. 2006;23:117–123.

53

Early-Onset Psychotic Disorders

LINMARIE SIKICH, M.D., AND TARA CHANDRASEKHAR, M.D.

Positive psychotic symptoms such as hallucinations and delusions may be present in a number of child psychiatric disorders including the primary psychotic disorders, schizophrenia, and schizoaffective disorders, as well as more common disorders such as depression, bipolar affective disorder, severe anxiety, and posttraumatic stress disorder (PTSD). In addition, transient psychotic-like experiences including sleep-related hallucinations and imaginary friends may occur in young children with a mental age of less than 8 years who have no mental health problems. The diagnostic distinctions between these disorders are primarily related to the duration and context of the hallucinations and the presence of specific co-occurring symptoms.

Children and adolescents, as well as their families, who have primary psychotic disorders and affective disorders with psychotic features, are often devastated by these illnesses. These illnesses are frequently characterized by recurrent crises, interruption of normal developmental processes, a deteriorating course, and treatment resistance. Further, the care of affected youth is associated with very high societal costs since these youth often require extraordinary care from their parents, frequent hospitalizations, and multiple long-term interventions from mental health, educational, and social service agencies. However, many believe that if effective treatments can be identified and implemented early in the course of illness, youth may prove more responsive and resilient than adults with psychosis. In addition, many researchers believe that it may be easier to identify the neurobiological processes involved in the pathogenesis of schizophrenia and related disorders in affected youth because they have a more severe but clearly developmental form of the disorder that appears more strongly associated with genetic factors.

DEFINITION

Positive psychotic symptoms consist of hallucinations, delusions, and bizarre or disorganized thinking and behavior including catatonia. Negative psychotic symptoms consist of flattened affect, apathy, avolition (difficulty making choices), poverty of speech, and poverty of thought content. The diagnoses of psychotic disorders in children and adolescents are generally

made using the same criteria as in adults. Youth with schizophrenia, like adults with the disorder, are required to demonstrate positive or negative symptoms for at least 6 months and to have significant social or occupational dysfunction or failure to achieve expected levels of social development. In addition, although not yet formally part of the fifth edition revised text version of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) criteria, schizophrenia is associated with marked neurocognitive declines in general cognition, executive functioning, social cognition, and memory. Indeed these symptoms, not positive symptoms, seem most closely related to long-term functional outcome.

Early-onset schizophrenia is defined as schizophrenia with onset of clear positive symptoms prior to 18 years of age. Schizophrenia with onset between ages 13 and 18 is also frequently referred to as adolescent onset. Onset of schizophrenia prior to age 13 is referred to as very early onset or childhood onset. Schizophrenia spectrum disorders include schizoaffective disorder and schizophreniform disorder. Schizoaffective disorder is defined by the prominence of psychotic symptoms, presence of significant periods of co-occurring major depression or mania, and psychotic symptoms present for at least 2 weeks in the absence of any mood symptoms during the lifetime duration of the illness. Schizophreniform disorder is defined by its short duration of psychotic symptoms (1 to 6 months) and generally evolves into either schizophrenia or schizoaffective disorder. Major depression with psychotic features and bipolar affective disorder meet all diagnostic criteria for the affective disorder but also include positive psychotic symptoms that are present only during periods of affective disturbance. Key distinguishing features of these disorders and other disorders with psychotic symptoms presenting during childhood are presented in Table 53–1.

Judith Rapoport and colleagues have defined a developmental disorder associated with transient episodes of psychosis during periods of stress called multidimensionally impaired (MDI) disorder, which is not yet recognized by DSM-5. Diagnosis of MDI disorder requires poor reality testing, daily affective lability, poor social skills, information processing deficits, and absence of a formal thought disorder. MDI disorder is frequently confused with early-onset schizophrenia, appears to have a poor prognosis, and carries increased risk of the development of bipolar disorder, type I. This should be distinguished from multiple complex developmental disorder (MCDD), which is characterized by more impaired social behavior (similar to the severe social deficits seen in autism spectrum disorder), greater evidence of formal thought disorder, high levels of anxiety or disruptive behaviors, and onset before age 5. Disruptive mood dysregulation disorder, which is new to DSM-5, is characterized by severe, frequent verbal and/or behavioral outbursts and chronic, persistent irritability with intact reality testing, and lack of hallucinations or formal

thought disorder. Disruptive mood dysregulation disorder is distinct from bipolar disorder and carries an increased risk of unipolar depression and/or anxiety disorders. Table 53–2 presents key characteristics of disorders that have overlapping features to assist with differentiating complex clinical presentations. These latter three disorders have great clinical significance but their biological roots and relationships with one another are poorly characterized at this time.

HISTORY

Schizophrenia has been reported to occur in children since the diagnosis was introduced by Emil Kraepelin. From the early 1900s to 1980s, the term childhood schizophrenia or prepubertal schizophrenia was used to refer to a range of severe psychiatric disorders presenting in youth including autism, leading to uncertainty about the diagnosis of research subjects during this period. In the 1970s, Israel Kolvin and colleagues conducted seminal work to distinguish between primary psychotic disorders and autistic disorder. Beginning with the ninth revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-9) and DSM-III, autism has been recognized as a distinct disorder and schizophrenia has been diagnosed using identical criteria in youth as adults. Most available research has supported this decision, although there appear to be a small group of individuals who have both autism and schizophrenia.

Table 53–1.
Disorders Presenting with Psychotic Symptoms in Children and Adolescents

Disorder	Distinguishing Features
Early-onset psychosis (EOP)	Any of a number of disorders in which positive psychotic symptoms are a prominent feature.
Early-onset schizophrenia (EOS)	Schizophrenia with onset prior to age 18. ► Involves a change from usual functioning, typically within 1 to 5 years of diagnosis, often has negative and neurocognitive symptoms rather than just positive symptoms. ► Affective symptoms may occur, but are not present for a substantial part of the illness. ► If patient also has an autism spectrum disorder, positive symptoms must include hallucinations or delusions. ► Impact on functioning may include failure to achieve developmentally expected social functioning or deterioration in social or occupational functioning. ► DSM-5 requires symptoms be present for 6 months, ICD-10 requires symptoms for only 1 month.
Very early-onset schizophrenia (VEOS)	► VEOS and COS are used interchangeably. ► Schizophrenia with onset prior to age 13, almost always after 5.

Childhood onset schizophrenia (COS)	► Onset prior to age 3 should raise strong suspicion of autism spectrum disorder.
Schizoaffective disorder (SA)	► During a substantial portion of the psychotic illness, the affected individual also meets criteria for a major depressive episode, manic episode or mixed manic/depressive episode. ► There are at least 2 weeks during the lifetime duration of the illness which psychotic symptoms are present in the absence of any affective symptoms.
Schizophreniform disorder	► DSM-5 diagnosis reflecting schizophrenia criteria for at least 1 month but fewer than 6 months. ► Not included in ICD-10. ► Generally evolves into schizophrenia or schizoaffective disorders, but may less frequently evolve into bipolar disorder.
Early-onset schizophrenia spectrum disorders (EOSS)	► Includes EOS, schizoaffective disorder and schizophreniform disorder with onset prior to age 18 and VEOS or COS. ► At this point there is not clear evidence that the treatment or prognosis of early-onset schizoaffective disorder is significantly different from EOS.
Multidimensionally impaired disorder (MDI)	► A developmental disorder beginning during childhood characterized by transient hallucinations, daily affective lability, poor social skills, and extreme emotional reactivity. ► Distinguished from VEOS by absence of negative symptoms and a formal thought disorder and from pediatric bipolar disorder by absence of decreased need for sleep, grandiosity, hypersexuality, and euphoria.
Multiple complex developmental disorder (MCDD)	► A developmental disorder with some characteristics of autism spectrum disorder (social disinterest, detachment, avoidance or withdrawal) as well as impaired affective regulation and thought disorder. ► Differentiated from early-onset schizophrenia by greater evidence of autism spectrum disorder and onset before age 5.
Major depression with psychotic features	► Youth describes clear depressed mood versus numbness or apathy. ► Psychotic symptoms are only present if mood symptoms are present. ► Psychotic features are usually mood congruent. ► Formal thought disorder is not usually present.
Bipolar disorder with psychotic features	► Psychotic symptoms are only present if mood symptoms are present. ► There is decreased need for sleep rather than insomnia. ► Negative symptoms are absent.
Substance-induced psychosis	► Psychotic symptoms are limited to periods of substance intoxication or withdrawal. ► Psychotic symptoms did not precede substance use. ► Psychotic symptoms generally resolve with 7 to 30 days of abstinence.
Posttraumatic stress disorder (PTSD)–associated psychosis	► Positive psychotic symptoms relate directly to trauma, and are not associated with other areas of life or random events. ► There are other symptoms of PTSD.
Anxiety-related hallucinations	► Positive psychotic symptoms are limited to periods of high anxiety. ► There are not negative symptoms or signs of a formal thought disorder.

COMPARATIVE NOSOLOGY

The diagnostic criteria for schizophrenia and schizoaffective defined by the ICD-10 are similar to those of DSM-5. The only difference is that ICD-10

requires a total duration criterion of only 1 month, compared to the 6 months' duration specified by DSM-5. Jorge Armenteros and colleagues found a high rate of diagnostic agreement between DSM-III-R, DSM-IV-TR, and ICD-10 in hospitalized psychotic adolescents. Similarly there is good agreement between the diagnostic criteria for major depression with psychotic features and bipolar affective disorder with psychotic features between ICD-10 and DSM-IV-TR. It is highly likely that the same level of agreement will be found with increased use of DSM-5 criteria.

EPIDEMIOLOGY

There are very few studies of the population prevalence of positive psychotic symptoms or schizophrenia in children and adolescents. The diagnostic challenges associated with making these diagnoses are likely to lead to marked underreporting in population-based surveys such as the Smoky Mountain Survey. In clinical samples, which are likely to over-report, positive psychotic symptoms seem to occur in a significant minority of children. A meta-analysis of 19 population studies described a median prevalence of psychotic symptoms of 17 percent among children aged 9 to 12 years and 7.5 percent among adolescents aged 13 to 18 years, indicating that the prevalence of psychosis appears to decrease with age. In one outpatient pediatric mood and anxiety clinic, 2,031 children 5 to 21 years (mean age 14.3 years) were evaluated using a structured diagnostic interview. Nearly 10 percent had definite or possible psychotic symptoms. Among those with psychotic symptoms, 62 percent had depressive disorders, 24 percent had bipolar disorder, and 14 percent had schizophrenia or schizoaffective disorder. Inpatient samples of children with depression have reported that as many as 60 percent will have psychotic symptoms. However, few depressed youth are hospitalized, so this is likely a significant overestimate. Clinical and research samples of youth with bipolar disorder have shown similarly high rates of psychosis.

Table 53–2.
Clinical Features Differentiating Complex Disorders of Childhood That Have Overlapping Characteristics

	Early-Onset Schizophrenia	Multidimensionally Impaired	Multiple Complex Developmental Disorder	Disruptive Mood Dysregulation Disorder
Cognitive deficits	+	+ (informational processing)	–	–
Delusions	+	+/–	+	–
Developmental delay	+	+	+	+/–
Emotional lability	+/–	+	+	+

			(affective dysregulation, anxiety)	(temper outbursts)
Hallucinations	+	+ (typically transient and stress induced)	+	–
Irritability	+/–	+	+	+
Social deficits	+ (social cognition deficits)	+ (impaired interpersonal skills, despite desire for friendships)	+ (social withdrawal, detachment, ambivalence)	+ (outbursts and irritability impair ability to initiate or sustain friendships)
Thought disorder	+	–	+	–
Long-term outcome	Schizophrenia	Increased risk of bipolar disorder, type I	Not yet well characterized	Increased risk of unipolar depressive disorder or anxiety disorders

Less information is available about the incidence of schizophrenia, although it appears very unusual for children, especially those less than 10 years old, to experience schizophrenia. In one area in Germany, only 3 of 100,000 children between 6 and 10 years old who presented for treatment were diagnosed with schizophrenia. Based on this clinical sample, the incidence of schizophrenia appears to increase by about fourfold between 10 and 14 years to 13 of 100,000 and by another four- to fivefold between 14 and 18 years to 20 to 55 of 100,000. At least 5 percent of individuals with schizophrenia become ill prior to age 14 years and up to 40 percent have been estimated to become ill prior to 18 years of age.

Given the greater prevalence in youth with major depression and bipolar affective disorder, but the pervasive presence of psychosis in youth with schizophrenia spectrum disorders, it is likely that as many or more cases of pediatric psychosis will be related to affective disorders as to schizophrenia. Furthermore, in prepubertal children, affective psychoses are considerably more likely than schizophrenia spectrum disorders. It appears that early-onset schizophrenia and affective psychoses are more common in individuals with family histories of the respective disorders. Early-onset schizophrenia appears slightly more common in boys than girls, whereas major depression is much more common in female adolescents than in male adolescents. Both genders appear equally affected by bipolar disorder. A clear relationship with onset of puberty has not been demonstrated.

ETIOLOGY

Schizophrenia is a heterogeneous disorder of unknown etiology. Similarly, the etiology of positive psychotic symptoms, which occur in many disorders, is unknown. However, research on schizophrenia consistently

suggests that there are subtle differences in neuronal connectivity, especially in the hippocampus and parietal, temporal, and frontal cortices. These differences appear to lead to less efficient cortical processing as demonstrated by functional magnetic resonance imaging (fMRI). Further, these differences seem to involve disruption of normal neurodevelopmental processes, particularly those that involve maturation of the frontal cortex and its connections during the second and third decades of life. In adults, alterations in the morphology of specific subclasses of γ-aminobutyric acid (GABA)ergic interneurons and of the volume and synaptic density of pyramidal neurons that receive thalamic afferents have been repeatedly demonstrated. These differences suggest aberrations in the normal developmental processes of synaptic refinement and pruning. Some researchers have speculated that multiple disruptions or "hits" to these normal neurodevelopmental processes are required for schizophrenia to develop. Such disruptions might result from genetic vulnerabilities; intrauterine adversities as a result of famine, infection, or inflammation; or various environmental exposures later in life including severe psychological stress including trauma and exposure to drugs of abuse, particularly cannabis.

Early-onset schizophrenia appears to be associated with more profound neurodevelopmental disruptions than adult onset schizophrenia. Frequently individuals with early-onset schizophrenia have social and attentional difficulties during early childhood. Many even have evidence of autism spectrum symptoms. Structural imaging studies have repeatedly demonstrated an exaggeration of the normal pattern of cortical volume loss that occurs throughout adolescence, with excessive and progressive reductions being especially prominent in the cingulate, temporal, and frontal cortices. In addition, studies of prodromal individuals, many of whom are adolescents, have shown regional gray matter volume reductions are present before overt positive symptoms and progress during the first few years of the illness particularly in cingulate, temporal, and frontal cortices.

Genetic vulnerabilities also appear more prominent in early-onset schizophrenia. Youth with early-onset schizophrenia are at least twice as likely as individuals with adult onset schizophrenia to have a parent with the disorder and morbid risk of schizophrenia for parents of individuals with early-onset schizophrenia ranges from 14 to 25 percent. Further, individuals with early-onset schizophrenia are significantly more likely than those with adult onset schizophrenia to have chromosomal abnormalities and DNA copy number variants. Genetic abnormalities suggest involvement of immunological and neurodevelopmental functions.

DIAGNOSIS AND CLINICAL FEATURES

All types of early-onset psychosis other than MDI and MCDD are diagnosed

using the same criteria as the adult forms of the disorder. However, the characteristics of specific symptoms often show important developmental characteristics that may be confusing for clinicians who infrequently work with pediatric patients. Hallucinations are frequently multimodal, with the affected youth seeing, hearing, and/or feeling the hallucinated being. Children will frequently name the hallucinated being, often using stereotypes such as “an angel,” “the devil,” or “a monster” or will refer to it on the basis of a physical characteristic such as “the red sweater guy” or “the man with no skin.” Youth very seldom complain about their hallucinations if not asked directly. Family members often are not aware of hallucinations but may endorse that the child asks about being called when no one has said anything. Delusions are seldom highly complex or systematic prior to age 16. Instead youth often report vague symptoms like “everyone is out to get me” or “people are laughing at me.” Often reports from collateral sources are needed to determine whether such reports are delusional or reality based. Frequently, youth with early-onset schizophrenia present with nonspecific symptoms such as poor attention or increased aggression.

Onset of early-onset schizophrenia is often insidious, occurring over a few years. In addition, premorbid symptoms such as being odd and socially isolated or having multiple (but generally mild) developmental delays in cognitive, motor, sensory, and social functioning are common and appear more severe than in adults with adult onset schizophrenia. Thus, when assessing youth for early-onset schizophrenia, it is important to ask about changes in functioning and behavior that have occurred over the past 3 to 4 years, to specifically inquire about hallucinations and delusions using language that the child understands, and to encourage the child to describe his or her experiences in detail using his or her own words. One should ascertain whether the child is in control of the phenomena, which is not seen with true hallucinations, or whether he or she can only moderate his or her own response to the phenomena. It is also crucial to determine the frequency and persistence of positive symptoms and associated negative symptoms. Finally, often, but not always, early-onset schizophrenia is associated with a formal thought disorder.

Clinicians are frequently reluctant to diagnose early-onset schizophrenia even when there is clear evidence of the disorder. Although the reasons for this are unclear, unfamiliarity with the developmental variations in symptoms, concerns about over-labeling, likely poor outcome, and stigma are frequently involved. However, failure to make the diagnosis when indicated may reduce access to appropriately intense interventions and relevant psychoeducational resources, increase the likelihood of treatment nonadherence, and result in hopelessness if the youth does not robustly respond to treatment. Further, longer duration of untreated psychosis has been consistently associated with poor prognosis in adult

onset schizophrenia.

Jack was first brought to a psychiatrist at age 12 when his parents were divorcing. He was obviously upset, was doing less well in school over the past year, and had gone on a hunger strike for the past month because he thought it would make his parents get back together. His parents noted that he had always been somewhat socially awkward, had a limited number of friends, and seldom had extended conversations. He was treated with fluoxetine (Prozac) and psychotherapy. After about 2 months, he was eating normally and seemed less preoccupied with his parents' divorce. However, he would make frequent statements that he needed to protect his younger brother, because he did not trust his parents to take care of the brother. He also said he was trying to find a house where he, his brother, and his father could live together.

A year later, after his mother remarried, he became markedly more irritable and was frequently noncompliant with her requests. His mother reported that there were episodes in which he would suddenly become extremely upset with little or no provocation. For example, she described an episode when they were out walking and he suddenly jumped on her back, saying that she looked like a prostitute. On another occasion, when his stepfather had instructed his brother to wash a glass that had been left out in the living room, Jack became very upset and said he planned to beat his stepfather to death with a golf club while his stepfather was sleeping. Jack's band instructor noted that he frequently seemed preoccupied and sometimes seemed not to know what to do next. Between these episodes, Jack was frequently pleasant and engaged superficially with the family, but had very seldom seen his friends or left the house. He experimented with marijuana by himself three or four times and said it made him feel calmer.

When asked about the incidents with his mother and stepfather, Jack seemed minimally concerned and did not provide an explanation. During his psychiatric evaluation, Jack reported that he often felt extremely angry without really knowing why. He acknowledged some periods of confusion. When asked directly about ever hearing anyone or anything talking to him, he said that his legs often talked to him and told him what to do. He denied significant mood symptoms and described his mood in nondescript ways. He had some poverty of content. At times his speech was disorganized and hard to follow. Organic workup including MRI and electroencephalogram (EEG) was noncontributory. Family history was significant for a great uncle who had been odd, socially isolated, and had been hospitalized multiple times at the state psychiatric hospital.

Jack was started on an atypical antipsychotic and showed a remarkable improvement in his oppositionality and irritability. He also became more engaged with family members and his grandfather noted he was having conversations with Jack for the first time ever. However, Jack continued to have increasing trouble in school, particularly as he entered high school. Although his positive symptoms were well controlled, he was failing all but one class. It was unclear to what extent his difficulties in school were related to sedation from the medication or cognitive limitations. Cognitive testing showed he was now functioning in the low average to borderline range and had particular difficulty with abstract thought and problem solving. Reading achievement was grade appropriate but he was significantly behind in math. Several meetings with the schoolteachers were held and an individualized education plan was developed. However, he continued to have difficulty. Various medication changes were tried without significant change in his presentation. When the antipsychotic dose was lowered to try to reduce possible sedation, he became more irritable and had thoughts of beating his mother to death while she slept because she had removed his video games for 2 days as a punishment. His school program was changed from a college preparatory program to an occupational preparation program. He felt much less stressed by this but was unable to find and maintain work at a part-time job as required for graduation. He did not seem particularly concerned by this and had no specific plans for the future. His family was considering applying for disability so that he could maintain his insurance after he turned 18 and was trying to figure out how to more effectively structure his time. Use of clozapine (Clozaril) was also discussed given the persistence of negative symptoms and vulnerability to positive symptoms when his antipsychotic dosage was reduced.

FIGURE 53–1. Features of chromosome 22q11 deletion syndrome (velo-cardio-facial syndrome [VCFS]). **A:** Demonstrates fluorescent in situ hybridization. One bright probe marks the end of each of the two chromosome 22s. Another probe, which binds to a marker within the typical large deletion in VCFS, is present on the upper chromosome 22 (noted with the larger *gray arrow*) but not on the lower chromosome 22 (noted with a smaller *white arrow*). **B:** Illustrates the facial features of three youth with VCFS. Note the elongated face, tubular nose with a wide nasal bridge, bulbous tip, and small lateral nares. Frequently individuals with the 22q11 deletion have low facial tone, small mouths, learning disabilities, or mild mental retardation and a variety of congenital malformations, especially of the heart and palate. A history of feeding problems in infancy is common.

PATHOLOGY AND LABORATORY EXAMINATION

The physical examination of youth suspected of having early-onset schizophrenia should include a careful neurological examination. If there is a focal neurological finding or impairment of consciousness, structural MRI is indicated. However, the volumetric, spectroscopic, and diffusion tensor imaging abnormalities consistently reported to show differences in group comparisons within research studies are not of sufficiently great magnitude to be used for diagnostic purposes. If there is any indication of episodic psychosis in the absence of mood symptoms or excessive sedation or confusion following psychotic behaviors, an EEG should be obtained to rule out a seizure disorder. Exposure to illicit substances should be assessed with a urine toxicology screen. Blood work should include screening for thyroid abnormalities, electrolytes, complete blood count, liver function tests, ceruloplasmin, and urinary copper (using a penicillamine test in children) for Wilson disease, heavy metals, and serum organic and amino acids to rule out a metabolic disorder. Antinuclear antibodies and serum and cerebrospinal fluid (CSF) antibodies to the *N*-

methyl-D-aspartate (NMDA) receptor should be considered to rule out lupus or anti-NMDA receptor encephalitis, especially if the onset of psychotic symptoms is very sudden or if subtle neurological complaints are present. Genetic testing including fluorescent in situ-hybridization (FISH) or evolving genetic techniques for chromosome 22q11 deletion syndrome (also known as velo-cardio-facial syndrome [VCFS]) should also be considered, particularly if the youth has a dysmorphic appearance, any congenital abnormalities or learning disabilities or other cognitive limitations. Approximately 25 percent of individuals with VCFS become psychotic during adolescence, with most experiencing schizophrenia. Since the incidence of VCFS is estimated at 1 in 4,000 live births, as many as 10 percent of individuals with schizophrenia may have VCFS (Fig. 53–1). Projective psychological testing may be useful in confirming suspected cognitive disorganization. However, failure to demonstrate psychotic disorganization during projective testing is not sufficient to reject the diagnosis. In addition, neurocognitive evaluation including intelligence quotient (IQ), auditory and visual memory, working memory, problem solving, processing speed, and attention measures may be useful in helping adapt academic expectations for individuals with early-onset schizophrenia who often are expected to perform at the same level as prior to onset of their illness despite significant changes in their capabilities.

DIFFERENTIAL DIAGNOSIS

Misdiagnosis of early-onset schizophrenia is a major concern. Factors that may lead to misdiagnosis include children's inherent difficulties in describing psychotic symptoms that are outside their normal experience, overinterpretation of transient psychotic symptoms, and the high prevalence of positive psychotic symptoms in child psychiatric disorders other than early-onset schizophrenia. However, failure to recognize early-onset schizophrenia when present may slow implementation of appropriate treatments and worsen the youth's long-term prognosis.

Evaluation of a child or adolescent who may have psychotic symptoms should always include separate interviews of the child and his/her guardian with a focus on clarifying the child's developmental trajectory, changes in functioning, internal experiences elicited in a developmentally sensitive way, past exposure to trauma, and environmental stressors. Family history, including information about extended psychiatric hospitalizations and suicides, should also be obtained. The clinician can usually make an accurate diagnosis with attention to co-occurring symptoms and their time course and strict adherence DSM-5 criteria (Table 53–3). However, affected youth and their families should be educated that the diagnosis will be re-evaluated over time as additional information becomes available.

Table 53–3.
DSM-5 Criteria for Schizophrenia

A. Characteristic symptoms: At least two of the following are needed for a period of at least 1 month: Delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior, or negative symptoms. The duration of these symptoms may be less than 1 month if they have been alleviated by treatment.
B. Social/occupational dysfunction: The level of social, occupational, and self-care functioning has markedly deteriorated below the highest previously obtained level. In children and adolescents, this includes the failure to achieve expected levels of social development.
C. Duration: The disturbances must be present for a period of at least 6 months. This includes an active phase of overt psychotic symptoms (Criterion A) with or without a prodromal or residual phase. A prodromal phase involves the deterioration in functioning before the onset of psychotic symptoms, whereas the residual phase follows the active phase. Symptoms characteristic of both prodromal and residual phases include marked social isolation, deterioration in occupational functioning, peculiar behavior such as food hoarding, poor hygiene, blunted or inappropriate affect, disordered thought processes (tangentiality, circumferentiality, poverty of speech, or speech content), odd beliefs or perceptions, and anergia.
D. Exclusion of mood disorder: Schizoaffective disorder and mood disorder with psychotic features must be ruled out. If mood episodes are present, their duration must be brief relative to the course of the schizophrenic symptoms. This exclusion criterion is especially pertinent for adolescents with bipolar disorder, as manic episodes in this age group frequently include schizophrenic-like symptoms at onset. This needs to be systematically reassessed in patients, as continued follow-up may be the only accurate method for distinguishing the two disorders.
E. Substance/general medical condition exclusion: The symptoms are not due to either the direct effects of a substance or a medical condition.
F. Relationship to a pervasive developmental disorder: If there is a history of autism, or another pervasive developmental disorder, prominent delusions or hallucinations (for at least 1 month unless successfully treated) are required to make a diagnosis of schizophrenia.

Reprinted with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, (Copyright © 2013). American Psychiatric Association. All Rights Reserved.

Schizoaffective Disorder

Like early-onset schizophrenia, schizoaffective disorder is characterized by persistent positive and/or negative psychotic symptoms and functional decline or failure to realize expected developmental progress. However, during a substantial period of the active and residual phases of illness, the individual also experiences prominent mood symptoms that meet criteria for a major depressive episode, manic episode, or mixed manic episode. Psychotic symptoms must be present in the absence of any mood symptoms for at least 2 weeks during the lifetime duration of the illness. If an individual is in the midst of their only psychotic episode and has had significant mood symptoms throughout, it may not be possible to distinguish between schizoaffective disorder and an affective psychosis. However, the presence of negative symptoms would increase the likelihood of a schizoaffective disorder. Early-onset schizophrenia with a secondary major depression may be distinguished longitudinally because the depressive symptoms are not present for a substantial period of the illness.

Major Depression with Psychotic Features

Individuals with major depression are generally able to describe their mood vividly rather than simply label it as depressed. In contrast, individuals with early-onset schizophrenia may say they are depressed but instead describe a state of apathy and emotional numbing that reflects their experience of negative symptoms. Insomnia, failure to enjoy activities, social withdrawal, lack of motivation, cognitive slowing, agitation, suicidality, and psychomotor retardation may be seen in both disorders. Likewise weight loss and decreased food intake may be seen in either disorder. However, in early-onset schizophrenia these symptoms are usually the result of delusional beliefs about the food or one's weight. Although mood-incongruent hallucinations and delusions may occur in major depression, they are not common.

Bipolar Affective Disorder with Psychotic Features

It can be extremely difficult to distinguish a mixed manic state from early-onset schizophrenia or schizoaffective disorder because hallucinations, delusions, irritability, and agitation are common to all three disorders. The primary distinguishing factors are the presence of decreased need for sleep and hypersexuality in bipolar disorder versus negative symptoms in the schizophrenia spectrum disorders. Youth with bipolar disorder, especially those who require hospitalization, appear to have greater numbers of varied psychotic symptoms than adults hospitalized with bipolar disorder.

Substance-Induced Psychosis

In substance-induced psychosis, positive symptoms are limited to periods of substance intoxication or withdrawal and are not better accounted for by a diagnosis of early-onset schizophrenia, schizoaffective disorder, or affective disorder. DSM-5 notes that such symptoms should resolve within 1 month of discontinuing the substance, although in clinical practice most substance-induced psychoses in youth resolve much more quickly (e.g., 7 to 10 days). If positive psychotic symptoms have been present prior to onset of substance use or during periods of sustained abstinence from substances, one of the other psychotic diagnoses is more appropriate. In addition, one is much more likely to see negative symptoms in schizophrenia spectrum disorders than in substance-induced psychosis, although decline in functioning is common in both. Comorbid substance use is common in older adolescents with early-onset schizophrenia. Substance use, particularly cannabis use, appears to be both a precipitating and exacerbating factor. Nonetheless, many youth with early-onset schizophrenia appear to self-medicate with drugs of abuse.

Autism Spectrum Disorders

Autism spectrum disorders and early-onset schizophrenia are distinguished by their developmental course and the presence of hallucinations and formal thought disorder. Although autism spectrum disorders may not be diagnosed until late childhood, onset of symptoms prior to 3 years of age is required to diagnose an autism spectrum disorder. Such abnormalities may be present in individuals with early-onset schizophrenia and schizoaffective disorder, but additional symptoms should present later in life that lead to the psychotic diagnosis. Onset of early-onset schizophrenia prior to age 10 is extremely rare. Further, although individuals with autism spectrum disorders may be extremely concrete and rigid in their thinking and verbal responses, they do not exhibit a formal thought disorder. However, it is also possible for individuals with autism spectrum disorders to develop a primary psychotic disorder. In this case, the individual must demonstrate clear hallucinations or delusions and negative symptoms alone cannot be used to satisfy diagnostic criteria. Some researchers feel the risk of schizophrenia is increased among individuals with autism spectrum disorders, although results of a small case study are frequently used to dispute this contention.

Anxiety Disorders and Posttraumatic Stress Disorder

Transient hallucinations may occur in youth with severe anxiety disorders. However, these youth seldom experience negative symptoms and are less likely to experience delusions or a formal thought disorder. In individuals with PTSD, the psychotic symptoms are generally limited to reliving events that occurred during the trauma and to the perpetrator of the trauma. Data regarding the long-term course of individuals with these symptoms is lacking. Some studies report frequent personality disorders, whereas a recent study found that a large proportion of individuals reporting hallucinations at age 11 but not appearing to have a primary psychotic disorder had developed one when reassessed as adults.

Multidimensionally Impaired Disorder

MDI disorder is distinguished from early-onset schizophrenia by the absence of formal thought disorder, the transient and reactive nature of the hallucinations, daily affective lability, and less profound social difficulties. Reductions in brain volume appear less profound than in early-onset schizophrenia. Long-term outcomes are not yet clear, though there may be an increased risk of bipolar disorder, type I.

Multiple Complex Developmental Disorder

MCDD presents with more prominent features of autism spectrum disorder than seen in early-onset schizophrenia. In addition, youth have greater incidence of formal thought disorder (hallucinations, delusions, magical

thinking), anxiety and affective dysregulation, and primitive or violent episodes of behavioral disorganization. Initially criteria for MCDD were conceived within the framework of autism spectrum disorder, though more studies are needed to characterize the incidence and long-term outcomes of this disorder.

Organic Disorders

The list of potential organic etiological agents of early-onset schizophrenia is extensive. Entities that must be considered include (1) delirium, (2) seizure disorders, (3) central nervous system lesions (e.g., brain tumors, congenital malformations, and head trauma), (4) metabolic disorders (e.g., endocrinopathies and Wilson disease), (5) autoimmune disorders (e.g., anti-NMDA receptor encephalitis and lupus), (6) neurodegenerative disorders (e.g., Huntington chorea and lipid storage disorders), (7) infectious diseases (e.g., encephalitis, meningitis, or human immunodeficiency virus [HIV]–related syndromes), and (8) toxic encephalopathies (both substances of abuse and medications). Individuals should be specifically asked about the use of solvents, phencyclidine, hallucinogens, cannabis, stimulants (including street-acquired amphetamines), synthetic or designer drugs, corticosteroids, and anticholinergic agents. Also, it is important to examine for comorbid medical or developmental disorders that will likely impact the type and extent of psychosocial treatments required. For example, there is an increased risk of psychotic disorders associated with VCFS, which is also associated with significant learning disabilities.

Emily, a 16-year-old girl with no previous psychiatric history, was brought to the emergency room due to a sudden behavioral change and suicidal ideation. A previously cheerful and accommodating girl, for the past 2 weeks she had been increasingly irritable, moody and easily frustrated. The day prior to presentation she began endorsing suicidal ideation and expressing a belief that her thoughts were controlled by an external presence (whom she referred to as "the man"). She also reported that her eyes had been replaced by cameras that allowed her to see ghosts. She was admitted to an inpatient psychiatric unit, where she was agitated, pacing and continuing to endorse hallucinations. Her physical examination was notable for ataxia and movement disorder (semirepetitive orofacial and limb movements). Routine labs (electrolyte, blood count, liver enzymes, thyroid function, and urine drug screen) were unrevealing. An EEG was grossly normal and her brain MRI showed no evidence of mass lesion or midline shift. She was treated with risperidone in escalating doses with little benefit.

On hospital day 7 she developed increasing fluctuations in her blood pressure and pulse, and then experienced a generalized tonic–clonic seizure. She was transferred to the intensive care unit due to worsening dysautonomia and central hypoventilation. Lumbar puncture results showed protein, glucose, neutrophils, and lymphocytes within the normal range. Though CSF anti-NMDA antibodies were negative, serum anti-NMDA receptor antibody panel was positive, increasing the suspicion of anti-NMDA receptor encephalitis as a cause of her symptoms. Imaging for an underlying tumor revealed an ovarian teratoma, which was surgically removed. She was concurrently treated with corticosteroids and experienced a gradual improvement in her psychotic symptoms and remission of her movement disorder. At a follow-up visit 3 months later, she had returned to school part-time. She denied psychotic symptoms but reported residual fatigue and impaired concentration.

COURSE AND PROGNOSIS

Subtle developmental symptoms are frequently present throughout childhood in individuals who have early-onset schizophrenia. The onset of overt psychotic symptoms is frequently insidious, particularly in individuals younger than 15 years. Afterward, approximately equal proportions of youth have insidious and subacute onset, with rapid onset over the course of several days being rare. Rapid recovery from the initial episode appears rarer than in adult onset schizophrenia. In contrast, many youth with early-onset schizophrenia show improvement but continue to experience significant positive symptoms despite medication adherence for 2 to 5 years. The course tends to be chronic rather than episodic as is reported in adult onset schizophrenia. Further, there appears to be significant decline in cognitive functioning during the 4-year period surrounding onset of illness. This cognitive decline appears to reflect both biological brain processes and interruption of normal development and learning. A recent study found that 30 percent of individuals with early-onset schizophrenia attempted suicide over the first 10 years of illness with one of six of those who attempted completing suicide. Suicidality was associated with more depressive symptoms and fewer negative symptoms at first episode.

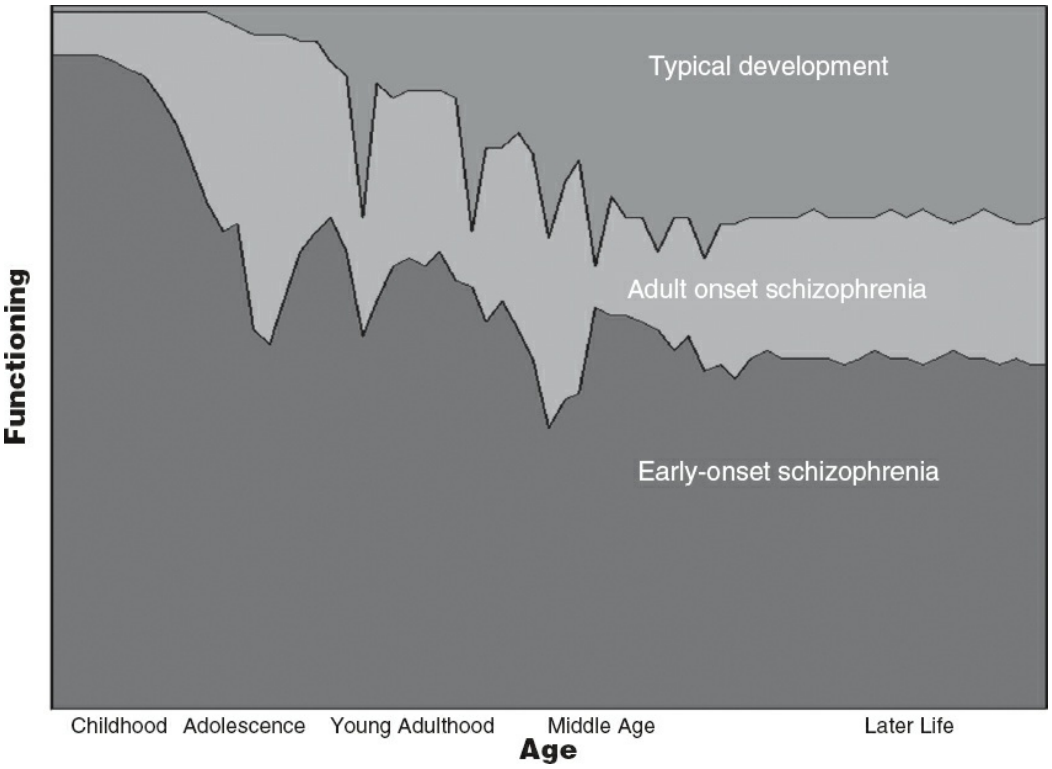

FIGURE 53–2. Comparative course of early-onset schizophrenia (EOS) and adult onset schizophrenia. Note the lower premorbid functioning of EOS, more insidious onset, and more chronic course with more limited periods of recovery.

Long-term prognosis in early-onset schizophrenia also appears worse

than in adult onset schizophrenia. A meta-analysis of 320 studies of outcome in adult onset schizophrenia found that about 40 percent of individuals had good outcomes, whereas about 60 percent had a chronic disabling course. In contrast, a number of studies have found that very few individuals with early-onset schizophrenia (5 to 25 percent) experience good outcomes and that 75 to 80 percent have poor outcomes with about half of those having extremely poor outcome. Few live independently as adults (7 to 29 percent), participate in full-time education, or are employed in the open market (17 to 28 percent). Another study found that fewer (25 percent) individuals with psychosis complete high school compared to age matched peers (84 percent). Treatment resistance may also be more common, although this has not been adequately studied. Figure 53–2 schematically illustrates the differences in course between early-onset schizophrenia and adult onset schizophrenia.

TREATMENT

The treatment of early-onset schizophrenia is multimodal and should include pharmacotherapy, family and patient psychoeducation and psychotherapy, and interventions in the community to support ongoing education and transition to employment when appropriate. The evidence is strongest for pharmacotherapy. A variety of traditional antipsychotics have demonstrated efficacy for reducing psychotic symptoms. A series of studies have compared various atypical antipsychotics to placebo in adolescents with early-onset schizophrenia and have demonstrated considerable antipsychotic efficacy, frequently reducing the Positive and Negative Syndrome Scale (PANSS) by about 20 percent. In these studies, olanzapine (Zyprexa) has been demonstrated to be associated with the greatest acute weight gain and lipid changes. Risperidone (Risperdal) seems to have more moderate acute weight changes and aripiprazole (Abilify) more limited acute weight changes. Extrapyramidal symptoms appear infrequent, but do occur. Only a few comparative trials have been done. These studies have generally found that the various antipsychotics differ in their side-effect profiles but have few differences in efficacy. Specifically, midpotency traditional agents appear to have relatively little impact on weight, lipids, and glucose tolerance, but are modestly more likely to cause moderate or severe extrapyramidal symptoms including akathisia. Significantly, these symptoms are generally well controlled with the addition of anticholinergic medications or propranolol. There is also a strong body of work indicating that individuals with treatment-resistant early-onset schizophrenia respond significantly better to clozapine than to haloperidol (Haldol) or olanzapine.

Compared to the literature in adult schizophrenia, few studies of psychosocial interventions have been conducted thus far. However, both cognitive remediation therapy (CRT) and supported employment

interventions show promise in improving the functioning of affected youth. A recent study found that clinically stabilized youth who received CRT experienced significant improvements in verbal memory and executive functions, and modest improvement in daily living, adaptive functioning, and reduced family burden compared with controls. Most of these changes were maintained after 3 months. Similarly a small Australian study of supported employment (Individual Placement and Support [IPS]) found that youth engaged in IPS gained significantly more jobs (23 vs. 4), earned significantly more money ($2,432 vs. $0), and worked for a longer time period (45 vs. 19 days). An overall positive vocational outcome (return to work or education) was achieved by 85 percent of the subjects, compared to 6 percent in the control group. Therapeutic interventions must focus on functional improvement in addition to symptomatic improvement in order to facilitate recovery. Targets of treatment include reduction in self-stigma and caregiver burden and restoration of self-esteem, self-efficacy, hope, and optimism.

FUTURE DIRECTIONS

Improving the recognition and timely diagnosis of early-onset psychosis will be important in reducing the disability, personal, and societal costs associated with these debilitating illnesses. It will be important to further characterize the long-term course of early-onset schizophrenia and other early-onset psychotic disorders, particularly with regard to neurocognitive and biological outcomes. In addition, it is essential that psychosocial treatments such as cognitive remediation, social skills training, and supported education and employment programs are rigorously evaluated and manualized where appropriate. More effective and safe medication strategies are also required to provide the most effective and safest treatment for these highly vulnerable individuals. This will require both more comparative trials and more studies of long-term outcomes. Development of additional psychopharmacological treatments, particularly those that are more focused on the neurodevelopmental processes that appear so prominent in early-onset schizophrenia, is also essential. Finally, early-onset schizophrenia also can serve as a more homogenous model for understanding the pathophysiology of schizophrenia and identifying key neurodevelopmental pathways in which to intervene.

REFERENCES

Armando M, Pontillo M, Vicari S. Psychosocial interventions for very early and early-onset schizophrenia: a review of treatment efficacy. *Curr Opin Psychiatry*. 2015;28(4):312–323.

Clemmensen L, Vernal DL, Steinhausen HC. A systematic review of the long-term outcome of early onset schizophrenia. *BMC Psychiatry*. 2012;12:150.

de Bruin EI, de Nijs PF, Verheij F, Hartman CA, Ferdinand RF. Multiple complex developmental

disorder delineated from PDD-NOS. *J Autism Dev Disord*. 2007;37(6):1181–1191.

Driver DI, Gogtay N, Rapoport JL. Childhood onset schizophrenia and early onset schizophrenia spectrum disorders. *Child Adolesc Psychiatric Clin North Am*. 2013;22:539–555.

Fraguas D, de Castro MJ, Medina O, et al. Does diagnostic classification of early-onset psychosis change over follow-up? *Child Psychiatry Hum Dev*. 2008;39(2):137–145.

Jiang L, Xu Y, Zhu XT et al. Local-to-remote cortical connectivity in early-and adulthood-onset schizophrenia. *Transl Psychiatry*. 2015;5:e566.

Kelleher I, Connor D, Clarke MC, Devlin N, Harley M, Cannon M. Prevalence of psychotic symptoms in childhood and adolescence: a systematic review and meta-analysis of population-based studies. *Psychol Med*. 2012;42:1857–1863.

Killackey E, Alvarez-Jiminez M, Allot K, Bendall S, McGorry P. Community rehabilitation and psychosocial interventions for psychotic disorders in youth. *Child Adolesc Psychiatric Clin North Am*. 2013;22:745–758.

Kumar A, Datta SS, Wright SD, Furtado VA, Russell PS. Atypical antipsychotics for psychosis in adolescents. *Cochrane Database Syst Rev*. 2013;10:CD009582.

Kumra S, Jacobsen LK, Lenane M, et al. "Multidimensionally impaired disorder": is it a variant of very early-onset schizophrenia? *J Am Acad Child Adolesc Psychiatry*. 1998;37(1):91–99.

Leeson VC, Harrison I, Ron MA, Barnes TR, Joyce EM. The effect of cannabis use and cognitive reserve on age at onset and psychosis outcomes in first-episode schizophrenia. *Schizophr Bull*. 2012;38(4):873–880.

Lewandowski KE. Cognitive remediation for the treatment of cognitive dysfunction in the early course of psychosis. *Harv Rev Psychiatry*. 2016;24(2):164–172.

Liu JJ, Norman RM, Manchanda R, De Luca V. Admixture analysis of age at onset in schizophrenia: evidence of three subgroups in a first-episode sample. *Gen Hosp Psychiatry*. 2013;35(6):664–667.

Pavuluri MN, Herbener ES, Sweeney JA. Psychotic symptoms in pediatric bipolar disorder. *J Affect Disord*. 2004;80(1):19–28.

Puig O, Penadés R, Baeza I, et al. Cognitive remediation therapy in adolescents with early-onset schizophrenia: a randomized control trial. *J Am Acad Child Adolesc Psychiatry*. 2014;53(8):859–868.

Puig O, Penadés R, Baeza I, et al. Processing speed and executive functions predict real-world everyday living skills in adolescents with early-onset schizophrenia. *Eur Child Adolesc Psychiatry*. 2012;21(6):315–326.

Rapoport JL, Giedd JN, Gogtay N. The neurodevelopmental model of schizophrenia: update 2012. *Molec Psychiatry*. 2012;17(12):1228–1238.

Reichert A, Krelker S, Mehler-Wex C, Warnke A. The psychopathological and psychosocial outcome of early-onset schizophrenia: preliminary data of a 13-year follow-up. *Child Adolesc Psychiatry Ment Health*. 2008;2(1):6.

Remschmidt H, Martin M, Fesichhaker C, et al. Forty-two-years later: the outcome of childhood-onset schizophrenia. *J Neural Transm*. 2007;114(4):505–512.

Rund BR, Barder HE, Evensen J, et al. Neurocognition and duration of psychosis: a 10-year follow-up of first-episode patients. *Schizophr Bull*. 2016;42(1):87–95.

Rund BR, Moe L, Sollien T, et al. The psychosis project: outcome and cost-effectiveness of a

psychoeducational treatment programme for schizophrenic adolescents. *Acta Psychiatr Scand*. 1994;89:211–218.

Sanchez-Gistau V, Baeza I, Arango C, et al. Predictors of suicide attempt in early-onset, first-episode psychoses: a longitudinal 24-month follow-up study. *J Clin Psychiatry*. 2013;74(1):59–66.

Schneider M, Debbané M, Bassett AS, et al. Psychiatric disorders from childhood to adulthood in 22q11.2 deletion syndrome: results from the International Consortium on Brain and Behavior in 22q11.2 Deletion Syndrome. *Am J Psychiatry*. 2014;171(6):627–639.

Schothorst PF, Emck C, van Engeland H. Characteristics of early psychosis. *Compr Psychiatry*. 2006;47(6):438–442.

Sikich L, Frazier JA, McClellan J, et al. Double-blind comparison of first- and second-generation antipsychotics in early-onset schizophrenia spectrum disorders (TEOSS) study. *Am J Psychiatry*. 2008;165:1420–1431.

Tillman R, Geller B, Klagen T, Corrigan M, Bohofner K, Zimerman B. Psychotic phenomena in 257 young children and adolescents with bipolar I disorder: delusions and hallucinations (benign and pathological). *Bipolar Disord*. 2008;10:45–55.

Ventura J, Subotnik KL, Gitlin MJ, et al. Negative symptoms and functioning during the first year after a recent onset of schizophrenia and 8 years later. *Schizophr Res*. 2015;161(2–3):407–413.

Walsh T, McClellan JM, McCarthy SE, et al. Rare structural variants disrupt multiple neurodevelopmental pathways in schizophrenia. *Science*. 2008;320(5875):539–543.

World Health Organization. *The ICD-10 Classification of Mental Health and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva: World Health Organization; 1992.

54

Child Psychiatry: Psychiatric Treatment

▲ 54.1 Individual Psychodynamic Psychotherapy (Child Psychiatry)

DAVID L. KAYE, M.D., AND ZHANNA ELBERG, M.D.

INTRODUCTION AND DEFINITION

Psychodynamic psychotherapy (PDP), rooted in psychoanalytic theory and used interchangeably with psychoanalytic psychotherapy by most practitioners, has a long and storied tradition in child and adolescent psychiatry both in the United States and internationally. Although for many years the backbone of child psychiatric theory and practice, it has been under siege for much of the past four decades from different quarters. The ascendancy of neurobiological approaches and understanding, third-party payers' demands for briefer and less expensive treatments, and the public's demand for empirical support all contributed to the diminished influence of psychoanalysis. Despite this, psychoanalysis has remained a vibrant, and still the most comprehensive, framework for understanding psychotherapy and the human mind. In the last decade, increasing empirical support for PDP with adults and children has energized the approach. The groundbreaking review by Shedler, *The Efficacy of Psychodynamic Psychotherapy,* published in American Psychologists in 2010 was a landmark in establishing the efficacy of PDP with adults. Since then, Leichsenring has published a recent systematic review (2015) that concludes that PDP is efficacious for a wide range of adult mental disorders. In 2012, the American Academy of Child and Adolescent Psychiatry published the first ever practice parameter for PDP with children. Two volumes of the *Child and Adolescent Psychiatric Clinics of North America* (January, April 2013) were devoted to PDP. Evidence-based, manualized approaches such as interpersonal therapy (IPT) and mentalization-based therapy (MBT) are based solidly on PDP principles. Further, there has been a softening of the boundaries between psychotherapy treatments with inclusion of aspects of different therapeutic approaches now commonplace in child therapy.

Although the place of psychoanalytic psychotherapy within child and adolescent psychiatry has declined from the hegemony it once held, it has remained a primary framework for many practitioners and a psychotherapeutic foundation of many, if not most, residency training programs. In the past it was often seen as a stand-alone treatment of child psychiatric disorders. In current clinical work, practitioners work pragmatically, flexibly, and in concert with other approaches including family therapy, behavior therapy, cognitive-behavior therapy, psychopharmacology, and school interventions. Regardless of approach, the understanding of patients stems from an overarching psychoanalytic framework. Psychoanalytic concepts also continue to be applied broadly to clinical work in residential or other intensive treatment milieus, social services, police and juvenile justice settings, and schools. In looking over his career Leon Eisenberg, a long-time critic of psychoanalysis, observed that:

> Where I erred was in failing to appreciate the powerful and lasting contribution psychoanalysis made to psychiatry by teaching trainees to listen to patients and to try to understand their distress, rather than merely to classify them by some diagnostic algorithm, or snow them with drugs, or lock them away, or release them to homelessness. Basing care on what is unique to the individual patient is central to clinical competence, whether in medicine or surgery or psychiatry.

Far from a single and monolithic set of ideas, psychoanalysis encompasses a wide diversity. Disagreement exists about the nature of the mind, development, psychopathology, and treatment. Despite this, psychodynamically oriented practitioners share the following precepts:

1. The primary emphasis and focus of PDP is the individual's subjective, inner experience of self and others. A primary goal in treatment is self-acceptance and self-understanding through the instrument of a relationship. The primary "tool" of psychoanalytically informed therapy is the relationship between the doctor and the patient.
2. Unconscious mentation is critical to understanding patient's subjective experiences and relationship patterns. Unconscious affects, motivations, and processes (including defense mechanisms) are accepted as critical determinants of normal and abnormal human behavior.
3. Behavior, emotion, and relationship patterns can be understood through the lens of psychic determinism. Human behavior has a logic of its own, happens for psychological "reasons," and is understandable through following a link of multiple impulses, affects, defenses, and their associated relational experiences. Often these links are unconscious and the work of therapy includes greater awareness of the impact of defenses and warded-off emotions and wishes.
4. The child is father to the man. Early experience in a child's life is critical in initiating important and enduring patterns in perceiving, thinking, feeling, and behaving that are carried forward in life via internalized

beliefs, working models, assumptions, expectations, and attitudes (i.e., transference).

5. Treatment critically focuses on transference. Focusing on the here-and-now transference reactions of the patient are central to understanding the patient and promoting growth in the psychoanalytic process.
6. Treatment goals are focused on psychological growth broadly, not just symptom amelioration. For example, in child psychotherapy goals often include a return of the child to a healthy developmental pathway, including greater self-acceptance, self-awareness, and self-expression.
7. Emotional barriers ("resistance") to psychological growth or resolution of symptoms can be expected in the course of treatment.

HISTORY

The first psychoanalytic reference to child cases was Sigmund Freud's case of "Little Hans" in 1909. Although Freud was skeptical that children could be analyzed, in this case he provided supervision to Hans' physician-father, who successfully analyzed the 5-year-old boy's phobia of horses. A few years later Sandor Ferenczi reported his attempt to directly treat a child, but found that the child quickly became bored with the analytic process and wanted to return to his toys to play. It was not until a series of women psychoanalysts who entered the scene that it became appreciated that play was the central medium of expression for children, and should be incorporated into work with those with emotional disturbance. Hermine Hug-Hellmuth, a teacher by prior training, was the first psychoanalyst to report the use of play in the treatment of children in 1921. She believed that treatment of children was primarily educational and limited in its goals.

By far the major figures in the early development of psychoanalytic treatment of children were Anna Freud and Melanie Klein. Klein began to develop a technique for working with children while in Berlin in the early 1920s, and in 1926, moved to London where she continued her work. She saw children as a cauldron of passionate drives (both aggressive and sexual) with rigid, harsh superegos. Her technique relied on equally powerful and direct interpretations of unconscious impulses. She appraised play as the equivalent of adult-free association and eschewed any role for parents in the treatment. Anna Freud, also a teacher by training, began her psychoanalytic training in the early 1920s and shortly thereafter formed a child psychoanalysis study group in Vienna. She viewed children as immature and undeveloped, necessitating a gentler and educative approach that was careful not to interpret too deeply for fear of overwhelming the child. She recognized work with the parents as difficult but necessary to further the development of the child and also to support treatment with the child. She saw the goals of treatment as being to return the child to his/her developmental path, and not to effectuate a definitive or complete cure. In 1927 Anna Freud wrote the first book on child analysis, *Introduction to the*

Technique of Child Analysis. The differences between Klein and Freud led to a confrontational Symposium on Child Analysis in May 1927. These "controversial discussions" brought points of disagreement into sharp focus, and this schism has persisted to the present day, albeit lessening in the recent past. The Kleinian school became dominant in the United Kingdom and South America, while the Freudian school held sway in North America.

Into this fray entered the analysts of the "independent school" in Great Britain in the 1940s, including Donald Winnicott, Wilfred Bion, and others associated with the object relations perspective. Starting in the Kleinian camp, they attempted to bridge the divide with the Freudians. Winnicott contributed enormously to the analytic understanding of child development and treatment, remaining perhaps the most well-recognized influence in the field. His rich body of work included an emphasis on the importance of the early mother–child relationship; the role of play, mirroring, and the establishment of the "holding environment" in healthy development and in analysis; the meaning of the transitional object; and the true versus false self.

In North America an additional perspective was emerging, beginning with Kohut's work in the 1970s on the development of the self. This evolution has continued to the present with major contributions from interpersonal, relational, and intersubjective theorists. Harry Stack Sullivan, Robert Stolorow, Stephen Mitchell, Joseph Lichtenberg, James Fosshage, and Irwin Hoffman among others have been major figures representing these perspectives. All of these authors have written primarily about clinical work with adults, with a few current authors applying the concepts to work with children (e.g., Neil Altman). These theorists have also been informed by convergences with infant and attachment researchers such as Daniel Stern, John Bowlby, Mary Main, and Mary Ainsworth. Rather than emphasizing the role of drives, defenses, resistance, conflict, and other traditional concepts these authors assert that the basic building blocks of mental life are relations with others upon which major psychic structures are built. This reflects a shift in the foreground from an intrapsychic focus to a relational field model, from objectivism to constructivism. Consequently there has been a shift in emphasis in treatment from *interpretation* to the *"co-creation of developmentally needed relational experiences"* (Jay Greenberg and Mitchell). Recently there has been a convergence of attachment, self-psychological, and relational perspectives. This is embodied by the work of Peter Fonagy, Mary Target, Efrain Bleiberg and others who have written about the central importance of *mentalization,* or self-reflective function. This ability to read and understand the feelings, thoughts, and impulses associated with human behaviors is fundamentally dependent upon the establishment of secure attachments. In the past several years this has led to MBT, a

manualized therapy approach that has been evaluated empirically, primarily in adolescents. Even more recently, neuroscience-based understandings have begun to be integrated into PDP with books such as Siegel's *The Developing Mind* and Delgado, Strawn, and Pedapati's *Contemporary Psychodynamic Psychotherapy for Children and Adolescents: Integrating Intersubjectivity and Neuroscience*.

While clinical work with infants is beyond the scope of this chapter, a great deal of parent–infant psychotherapy has been based on psychodynamic principles. Psychodynamic theory has influenced the field beginning with Selma Fraiberg and early programs developed by Alicia Lieberman, up to present-day evidence-based treatments such as parent–child interaction therapy.

It should also be noted that the term "psychodynamic" was originally used to describe the ever-changing balance of mental and neurophysiological energy intrinsic to man's inner life. This was contrasted with an approach that emphasized behavior, descriptive phenomenology, or anatomic structure. In this sense, psychodynamic and psychoanalytic are not synonymous. In the United States there have been a number of influential writers who would be considered within this broad psychodynamic tradition, even though they were not psychoanalysts. These psychodynamic therapists (e.g., Frederick Allen, Clark Moustakas, Ann Jernberg, John Meeks, Virginia Axline, and more currently Charles Schaefer, Garry Landreth) have also had a major place in the history of child and adolescent psychotherapy and continue to influence current practice.

THEORETICAL ISSUES IN PSYCHODYNAMIC WORK WITH CHILDREN

Beginning from its earliest days child psychodynamic psychotherapists have recognized that there are developmental differences between child and adult patients. While the overarching theoretical issues in PDP reviewed above hold, there are fundamental differences in work with children and adolescents as compared with adults.

Children are not "little adults." Child patients differ in significant ways from adults. These differences are crucial in informing practice. Firstly, children's limitations in language, cognition, and experience must be taken into account to engage and sustain children in treatment. Because of cognitive and emotional limitations, their understanding of the problem or the need for treatment is likely to be incomplete or inaccurate. Children tend to externalize problematic issues and have little ability to understand how therapy might help them now, let alone in the future. Further they have limited capacity for tolerating pain and anxiety, creating another barrier for keeping child patients in treatment. Meeting with a stranger who will meet and talk intimately with the child is certainly extraordinary, and may be confusing and threatening. For all these reasons the child

therapist must *actively* woo, educate, and work with the child to create the secure space needed to establish a therapeutic alliance.

Another major difference is that children do not come to treatment at their own request. They are brought by their parents, often for reasons that are more of concern to the adults rather than to the child. Further, the stigma associated with treatment is often difficult for children (especially adolescents, who are struggling to feel "normal" and "like other kids"). The reasons for treatment must be explored and explained in a way that is understandable and meaningful to the child. The rationale that engages the child may be different from the reasons the parents bring the child in.

A 7-year-old boy was brought for treatment by his working-class parents who were concerned about his depressive irritability and argumentativeness, which had created great tension in the family. When seen individually, the therapist asked the child about his understanding of why they were meeting. The boy responded "I don't know." The therapist then asked why he thought the parents had brought him, and he again avoided the question. Later in the interview the therapist returned to the issue while the child played out a puppet scene of different animals fighting. The therapist explored how the animals felt and the child said "sad." As the play developed, the therapist explored whether the child might also be sad. This seemed to resonate and the child nodded his head "yes" and went back to playing with the puppets. As the session closed, the therapist suggested that in meeting together they might come to understand "why you feel sad and what we can do to help you feel better. How does that sound to you?" The boy looked up shyly and nodded his approval.

The child psychotherapist must recognize the critical and central role of the parents or caretakers in the child's life. At a minimum the therapist must be able to sustain enough of a relationship with the parent(s) or caretaker(s) to support the treatment of the child. The parents must be engaged sufficiently to bring the child to treatment, pay for the treatment, and withstand their own emotional reactions that may arise as the child changes. Further, a positive alliance with the parents can sustain a child patient through the usual feelings of reluctance to coming to treatment. Ideally the parents can be assisted in fostering sturdier attachment, boundaries, and mentalization. This all requires the psychodynamic child therapist to develop a solid working relationship with the parents while at the same time preserving a sense of confidentiality and trust with the child.

A 10-year-old girl was seen in twice-weekly PDP. As the child left her sessions, her mother each time asked "how is she doing?" The therapist made a separate appointment with the parents to discuss these concerns further, while also exploring how the child felt about this meeting. The girl had concerns about "what they are going to say about me," but assented with support and understanding from the therapist. After discussion with the parents it was agreed that the therapist would meet with the parents once every 3 weeks to discuss the child's progress and current functioning, provide support and guidance to the parents, and monitor the parents' anxiety. Once this plan was initiated, the mother's anxiety seemed to lessen and she no longer waited expectantly after the sessions with her daughter.

The child patient is dependent on the parents materially, physically,

and emotionally. This has profound implications and has led to the conviction, even among the most traditional child psychoanalysts, that therapeutic work with the parents is essential and necessary in working psychodynamically with children. Furthermore, the ongoing interactions with caretakers have a continuing effect on the growth and progress of children in therapy. Because a child's personality is unfolding and shaped strongly by continuing developmental experiences, the ongoing parent–child interactions have a "double-barreled" effect. They not only affect the child's developing internal world of expectations, assumptions, and beliefs about self and others, but then in turn impact on the child's exposure to future developmental experiences. By inhibiting or encouraging exposure to growth-promoting developmental experiences, the ongoing parent–child interaction exerts a multiplier effect on the child. This observation strongly influences the overall aim of PDP with children, which, as initially articulated by A. Freud, is to restore the child to the path of normal development. Once back on track, normal development can then take over and provide the experiences needed for internal growth. From a psychological perspective, the child's dependency also impacts the individual treatment directly. Adolescents, especially during the early teen years, are struggling mightily to separate from their parents and establish an independent sense of self (i.e., identity). As a result, looking to and depending on another adult, like a therapist, can be fraught with complication and ambivalence. This can interfere with the establishment of a therapeutic relationship even when the adolescent recognizes a need for help.

Children and adolescents evoke stronger emotional responses in psychotherapists attempting to help them. As is known by anyone who has spent any time working in child mental health settings, children evoke strong feelings in the therapists who work with them. This often results in desires to protect and rescue the child, especially when therapists observe here-and-now pathogenic interactions and traumatic experiences enacted by parents. At other times the child's behaviors can be exasperating, infuriating, and frightening. All of this results in feelings within the therapist that are generally much higher in intensity than those generated in adult work. This places a premium on the therapists' recognition and containment of their own emotional responses. While these reactions may be helpful clues in understanding the child's world, they can make for highly challenging work.

Play is the central medium for communicating with younger children. Since its initial recognition by early psychoanalysts, play has been recognized as the critical in communicating with children in PDP. This is especially true for younger children whose language skills have developed to a limited degree and for whom play is their "natural" language. Older children are increasingly comfortable talking about their own inner states

and by 9 or 10 may find playing "babyish." Whether or not to use play materials with an individual child must be determined on a case-by-case basis. Psychodynamic psychotherapists will encourage play by having a number of toys visible and accessible to the child. This helps to engage the child and conveys that the therapist is willing to meet him or her on "their turf" and create a safe space for talking about their problems. Much of the "talking" is done "in the displacement" of the play as opposed to the direct discussion that occurs in therapy with older adolescents and adults. Most therapists recommend a limited number of toys that allow for the symbolic expression of a wide array of affects and themes (e.g., doll house, dolls, puppets, animals, blocks). Art materials (drawing, painting, clay) and techniques (e.g., Winnicott Squiggle Game), storytelling (e.g., Gardner Mutual Storytelling Technique), and role playing are often incorporated into dynamic therapy. Therapists also may utilize structured games (e.g., checkers, chess) or therapeutic games (e.g., The Talking, Feeling, and Doing Game) in some circumstances.

As Winnicott has pointed out, play exists "in the space between" the inner world and external life, *between* the patient and the analyst. Play requires active participation from both the patient and the analyst. It is aimless and nonpurposeful; there are no social consequences. As such, play allows for the safe exploration of affects, conflicts, areas of concern, ideals, roles, and wishes. This accords play a self-healing effect in and of itself. Play is a forum for problem solving and experimentation. Play also offers a window into the inner life of the child. This allows for a deepening of the understanding of the child and may be the basis for targeted interventions. As Winnicott has written, when play is not possible then the work of the therapy is to help the child to be able to play.

A 4-year-old boy from a poor urban family was brought for treatment because of severe aggression and destructiveness of long-standing duration. In addition to an evaluation for medication the child was seen in twice-weekly psychoanalytically oriented psychotherapy. The beginning sessions were marked by the repeated need to set limits and contain the child's aggressive behaviors. Two months into treatment he began to pump himself up, roar, and announce that he was "the incredible hulk." He would then proceed to stomp around the play therapy room, at times attempting to hit the therapist and kick the door. The therapist then said: "You know you can't really *be* the hulk. You can *feel* like the hulk ... you can *pretend* that you are the hulk. Then maybe we can play this hulk thing together." After a number of similar exchanges the child gradually became able to contain his behavior and play at being the hulk, without becoming the hulk.

Transference in children and adolescents is more complex than with adults. Much controversy has surrounded the concept of transference in work with children. Children's internal representational world is still being actively formed and altered by ongoing current interactions with parents and others. As such, transference reactions, while organized to a degree, are still in flux. Nevertheless, transference reactions occur ubiquitously in

PDP as children enact the beliefs, assumptions, and expectations they are developing. These reactions are here-and-now examples of the prototypical interactions that are interfering with their growth and development. Interpretations of transference phenomena primarily address this here-and-now aspect ("perhaps the nervous feeling you have is because of some anger you are feeling") and much less so, address genetic roots (i.e., the analyst would less often make interpretations about the child's reaction being connected to previous experiences with parents). PDP focuses its greatest attention on these transference reactions, seeing this as the most powerful arena to help children gain flexibility and greater freedom in their understanding of themselves and their interactions with others. Psychodynamic work with children is made more complex because the child's transference reactions intersect with the parent's emotional needs and vulnerabilities. This can lead to divided loyalties between the child's desire for growth and the parent's needs. Countertransference reactions are seen as invaluable clues to understanding the transference reactions and warded-off self-experiences of the child.

Barriers to change exist within the child patient (i.e., resistance) and within the family system. Resistance is a concept that originates in classical theory to describe the inevitable intrapsychic barriers to growth. Traditionally, this is thought to derive from conflicts between the patient's ego, identity, and superego. Subsequent theorists have conceptualized this differently. Some have understood these barriers as attempts to master traumatic experience. Others have seen the terror and dread of changing and giving up familiar ways of seeing oneself and others as central. Still, other contemporary theorists discard the term entirely as implying an adversarial and passive attitude that resides within the patient. Instead, they account for these impasses as the child's understandable emotional response to the absence of needed relational experiences and empathic failures by the therapist.

While resistance is typically seen as intrapsychic, there are also systemic barriers to change. As the child becomes less symptomatic and grows psychologically, the parent's self-concept, parent–child or marital relationship may become challenged. This, in turn, can create an additional impediment to the child's progress and requires the therapist to attend and respond to the parents' anxiety about change.

An 8-year-old boy was seen in individual PDP for separation anxiety disorder. As the child's symptoms lessened and he developed greater confidence, he began to experiment with desires to move out from the orbit of his family and expressed interest in playing on a soccer team. This, in turn, stimulated the mother's worries about the child's physical and emotional safety. She turned to her husband for increasing support and empathic understanding, revealing a long-standing issue in the marriage. The parents were not able to negotiate this in a more satisfactory way. These systemic implications impinged on the child, and created a check on his progress.

There is an increasing focus on attachment, mentalization, and self-regulatory issues versus oedipal issues. Classical theorists have highlighted the central importance of oedipal issues in understanding both psychopathology and treatment. Early case descriptions appear to explain clinical phenomena almost entirely through this lens and treatment appears to hinge on adequately addressing through interpretation these oedipal issues consistently. Triadic relationship patterns and issues of competition, sharing, sexuality, and rivalry are paramount. In contrast, contemporary theorists have instead placed the primary focus of understanding psychopathology and treatment to early parent–infant interaction, "basic trust," attachment, separation-individuation concerns, self-reflective process (mentalization), and self-regulation. While not disavowing a place for triadic issues, this has prompted a major shift in emphasis to dyadic issues.

An 8-year-old boy is brought for treatment because of his multiple fears and poor academic performance. Toward the end of the psychotherapy, after the child had made much progress, his father attempted suicide. While the original symptoms could be understood as a reflection of oedipal issues (e.g., castration fear and guilt over rivalry with the father, leading to academic inhibition and fearfulness), contemporary theorists would also emphasize the boy's anxiety over separation and loss due to his father's own attachment issues and tenuous commitment to living. Consideration would also be made of the security of the boy's attachment with his mother and his sense of a "home base."

While interpretation was seen in classical viewpoints as the central determinant of how psychodynamic psychotherapy works ("therapeutic action"), alternative explanations are now in ascendancy. Classical analysts have emphasized the central role of interpretation and "making the unconscious conscious" in the treatment of children. Interpretations proceed from that closest to consciousness to those mental contents furthest away (i.e., unconscious). This is especially important with children. These interpretations may be directed toward defenses, wishes or intentions, or transference phenomena. Interpretations are traditionally made in a specified order, from the manifest level to what is unconscious: first affect or behavior, then defense, then impulse. In other words, the analyst first would interpret the affect or behavior ("You appear to be feeling angry right now"), followed by the defense ("perhaps you are angry because you were disappointed that I did not remember your birthday"), and finally, the underlying impulse ("perhaps you were really hoping that someone would remember your birthday"). Timing of interpretations must be aligned with the patient's readiness and the state of the therapeutic alliance. Theorists differ in what they see as the primary impulses or fundamental motivators. Classically oriented analysts follow Freud's original drive theories and see sexuality and aggression as primary; contemporary analysts might emphasize self-cohesion, self-regulation,

attachment, mastery, and assertion as the most powerful organizers of motivation. Contemporary theorists hence would see therapeutic action as stemming from the provision of needed relationship experiences that allow for the development of a cohesive and vital self. Fonagy, Bateman, and others have attempted to bridge these views, attachment theory, and cognitive neuroscience by highlighting the goal of mentalization. Mentalization, or the self-reflective capacity, enables the child to perceive and process mental activity in the self and others. This allows the child to develop a model of the inner thoughts and feelings that underlie human behavior. Optimally, this capacity develops out of the child's early attachment with parents who relate to the child as an intentional being with underlying thoughts, feelings, and intentions. As a result of these experiences, the child is then able to understand others' behaviors and inner states. These theorists posit that this is the centerpiece of psychoanalytic work with children (and adults) and success depends on the degree to which this develops.

PRACTICAL MATTERS

Psychodynamic psychotherapy with children and adolescents can take differing forms and the principles can be applied in work with individuals, groups, and families. Classically, child psychoanalysis involved seeing the child four or more times per week, lasting over a period of years. In this form, psychoanalysis is rarely used currently, although there is some research to suggest it may have a place with particularly disturbed children (see Research section). PDP occurring once, or perhaps twice, per week over a period of months to years, is much more common in both private practice and public clinic settings, although it is increasingly difficult because of payment issues and the hectic pace of family life in the United States. The nationalized health systems in the United Kingdom and Canada continue to support such treatments widely.

While typically open-ended in length, there have been recent adaptations of brief, time-limited PDP to work with children. For example, IPT has been applied to work with adolescents and likely will be adapted to work with children. Psychoanalytic principles have also been widely applied to group and family therapies. As noted previously, psychoanalytically informed principles often are integrated into all treatment efforts with the child. Owen Lewis has delineated such a model of integrated treatment in which psychodynamic principles fundamentally inform the gamut of treatment interventions (e.g., behavior therapy, psychopharmacology, family therapy, and so forth). The more recent review by Malberg and Mayes (The Contemporary Psychodynamic Developmental Perspective) elaborates on this evolution.

Perhaps, even more than in work with adults, the therapeutic frame and setting are critical in working with children and adolescents. Child

therapy requires a dedicated office that can serve as a concrete reminder of safety, security, and predictability for the child. Using the same office for every session fosters these goals. Some analysts create special cubby holes or boxes for children to keep their particular toys, drawings, and so forth. This also serves to personalize the space for the child. The office should be neither so small that the child feels closed in, nor so large that the child is overwhelmed and overstimulated. An explicit therapeutic contract should be discussed with the child and parents. Children should not be forced to attend therapy although reluctance is not uncommon. Assent (vs. consent) is a more appropriate goal. It is generally recommended that therapy occur at the same time, on the same day of the week. This also strengthens the predictable frame necessary for successful work. All these ingredients contribute to establishing what Winnicott referred to as the "holding environment."

Confidentiality is a particularly complex issue with children. On the one hand, there is a clear need to maintain a sense of safety and security in order for the child to reveal his/her deepest worries, hopes, thoughts, and desires; on the other hand, therapists have a legal and ethical obligation to safeguard the child's welfare. When issues of safety (e.g., suicidality) arise, parents must be informed. At other times, the involvement of child-protective authorities may be needed, which in some situations may end the treatment. In any event, the limits and extent of confidentiality should be carefully explained to both the child and parents at the beginning of treatment.

Working with parents is essential in psychotherapy with children and adolescents. As already noted, parents are decisive players in PDP with children. Although there are times that the parents need to honor the necessary boundaries of treatment with the child, it is seldom appropriate that the parents are excluded from the treatment. Typically, parents are included from the beginning to obtain the history and to begin to establish a working relationship that can sustain the treatment of the child. At the conclusion of the evaluation period, it is helpful to explicitly review the assessment and establish a treatment plan with the parents. The analyst must reach concordance with the parents on the nature of this treatment contract from the beginning and may need to revisit this throughout the course of treatment. Regular meetings with the parents (with or without the child) are routine parts of PDP. These ongoing meetings serve to maintain the working relationship, provide further information about the child, and create a forum for parent guidance or exploration of parental psychological issues that impede the child's growth. When it becomes apparent that the parent(s) are in need of their own treatment then referral is generally made to another therapist. Providing PDP to two members of the same family contemporaneously is discouraged due to the inherent conflict of interest and encroachment on confidentiality that may undercut

the alliance with one or both persons.

CLINICAL CONSIDERATIONS

The aims of psychoanalytic psychotherapy are ambitious. Psychoanalytic psychotherapy aims to go well beyond the relief of symptoms and/or greater coping or adaptation. With adults, the goals have been to seek a "cure," broadly improving the quality of life (i.e., "to work and to love") and inoculating the patient from future stressors. In contrast, A. Freud first articulated the aims of child psychoanalysis as a return to the path of normal development. Subsequent analysts have emphasized the aims of increased mentalization, self-understanding, self-regulation, personality integration, self-expressiveness, self-esteem, flexibility, increased frustration tolerance, and capacity to play. It is assumed that the reaching of these goals in a "good enough" fashion will result in a diminution of symptoms as well as improved functioning in school, in the community, with peers, and family members.

The indications for psychodynamic psychotherapy generally include children with internalizing disorders. Indications for PDP have historically focused on the presence of certain strengths, or loose categories of problems generally seen as due to intrapsychic conflict. Typically, indications have included those children with neurotic level disorders and sufficiently good internal and external resources such that the gains will be sustainable following termination of treatment. This situation rarely exists in clinical practice. Further, research has indicated that this group of children may benefit as well or better from less intensive treatment efforts involving the school, family, or community resources. At the same time children with more severe disturbances, often subsumed under the category of borderline conditions, may require an intensive effort such as psychoanalysis for effective results to be achieved (see section on Research and Evaluation). Another way to look at indications is to consider the goals of the treatment. For example, if the goal is to understand one's feelings, thoughts, and the connections between this inner world and the outer world of behavior (mentalization or self-reflective function) then a PDP is likely to be helpful. Because PDP is now generally integrated with other treatment approaches, it is not necessary for this treatment to be all encompassing but rather to aid in forwarding critical developmental processes.

There are some children for whom psychodynamic psychotherapy is not likely to help or is contraindicated. Contraindications include children with externalizing problems, especially when there are "callous-unemotional" features or when there is no comorbid internalizing disorder or an inability to restrain impulses consistently enough to play. Another group of children for whom there is little to justify this treatment would be in children with more than mild mental retardation, classical autism and

many autism spectrum disorders, and traumatic brain injury (and other neurological conditions) when receptive language abilities are substantially compromised. These may be relative contraindications as there may be focal psychodynamic goals even in some of these cases. Again, treatment for many of these individuals will be multimodal and other treatment modalities may be more critical (i.e., school interventions for children with autism-spectrum disorders).

AN OVERVIEW OF THE PHASES OF TREATMENT

Beginning Treatment: Focusing on the Establishment of a Working Alliance

This requires not only creating a therapeutic ambiance for the child but also establishing a supportive, working relationship with the parents. The initial goal with the child is to create a playful, confidential environment in which he/she and the therapist can explore the world of the child's thoughts, feelings, and behaviors. In the beginning, some exploration and understanding of why the child is meeting with the therapist is useful as well as an explicit clarification of the treatment frame and contract ("we will be meeting for 45 minutes every Tuesday at 4:00. This time will just be for us. We'll try really hard to meet every week"). It is also helpful for the therapist to explain to the child that he/she is "someone who meets with kids to help them with their feelings," "to see if we can understand you," "to help you feel better," "to help you understand how you feel and why you do some of the things that cause you trouble," and so forth. The gist of these comments is to explain the role of the therapist and to convey a steadfastly hopeful attitude to the child (and parents) through what will be inevitable ups and downs. For the playroom to acquire a "specialness" to the child, these messages must be communicated repeatedly. The ground rules also should be explained ("we will talk and play together during the time we are together. Anything is okay as long as no one gets hurt and nothing gets broken.") Of course once the child begins to express him/herself, they will see what the therapist is "made of." This is the true acid test for the child. Only over time can the child get the sense that the therapy is not a punishment, that the therapist has hope for the child, the child can feel understood, and problems can be solved. For the most part, in the initial phase the therapist keeps his/her comments and interpretations focused on the displacement of the play rather than making direct comments about the child's real life situation. This initial launching typically takes a few weeks (in the healthiest children) to a number of months to achieve a secure working relationship.

Mid-Phase Treatment

Begins when the child has established a clear working alliance. The child has a familiarity with the frame of treatment, even though emotional

upheavals may continue in the sessions. During this phase, the primary emphasis is on the "work" of therapy, that is, the slow and repetitive recognition of feelings, defenses, and their relationship to behaviors. Over this time the same theme in a play session may occur for months at a time. The therapist continues to create with the child a sense of hope and comfort, as increasing recognition and understanding of the child's inner world accrues. It is in this phase that interpretation of prototypic problematic patterns and transference reactions occurs. The focus on the here-and-now reactions of the child with the therapist is especially fruitful, although much work continues to be done in the play displacement. Gradually in this phase, the child becomes more able to connect this displacement with his/her own life. Reconstructive statements linking current emotional or behavioral patterns with previous life experiences are generally not emphasized in work with children.

A 10 year old is in PDP three times per week because of severe obsessive compulsive disorder. In most sessions, beginning from the fourth month of treatment and lasting until the eighth month, the child would crawl under a table, conducting conversation from this womb-like space they had created. The analyst commented to the child, who had been born 2 months prematurely, "It feels like a really good place to go to when *you* want, and to be able to come out when *you* want." This interpretation emphasized the current emotional experience of the child, rather than linking this in some direct way to the child's real life experience (e.g., "you want to climb back in your mother's womb and are angry that you can't"). Connecting the here-and-now with the past must be timed carefully so as not to be experienced as confusing and overwhelming to the child.

Ending Treatment

Optimally, treatment ends when the child is free from symptoms, has made sufficient progress in psychological growth and maturity, and is back on track developmentally. In the real world, termination generally comes about for other reasons. Frequently, termination occurs when partial improvement has been made and practical considerations (insurance, finances, transportation, moves, graduation) become paramount. Termination provides an opportunity for the child to focus on issues of separation, loss, and mourning. Paulina Kernberg has elegantly stated:

> "As the child approaches the end of her treatment, the same issues that brought her to therapy once again come to the fore: (1) Do they (the parents) care? (2) Does the therapist care? (3) Do I care for you (the therapist)? (4) Do I care for them (the parents)? (5) Can I care for myself? (6) Will you remember me? (7) Will I remember you?"

In the process of ending, it is important to include the child actively in setting a final date. Generally 6 weeks to 3 months is sufficient time to allow the child and family to meaningfully concentrate on these issues. Setting the date so that it does not coincide with other real-life endings (e.g., graduation, moving dates) permits the child and family to experience the fullest impact of the termination. During the ending phase children

frequently regress, with a recurrence of symptoms and defenses in the face of this stressor. Denial of the importance of the relationship by the child, parent, and therapist is not uncommon and should be understood as reactions to the impending separation and loss. In the course of ending, the child should have an opportunity to explore and claim the gamut of feelings that inevitably arise: anxiety, guilt, anger, sadness, satisfaction, ambivalence, and so forth. This frees the child from developing defensive patterns to face the inevitable future losses of life.

After 8 months of weekly psychotherapy an 11-year-old girl's symptoms of perfectionism and severe recurrent abdominal pain had improved sufficiently that the child came in and stated that "I don't have pain! I don't need to come here anymore." She later added that "maybe I need help with boys." This led to a series of discussions about termination, which resulted in the setting of a final date 6 weeks hence. In the first meeting after the termination date was set, she revealed that she was able to comfortably talk with the therapist, unlike with her parents with whom she felt she could not be open. The next week she became irritated by a comment of the therapist and said "If I had known you would say that I would have left weeks ago!" She continued to talk about other separations in her life, especially the death of her grandmother. In the final sessions she alternated between denial and recognition of her emotional reactions about the upcoming termination. She continued to discuss other endings, moves, friends leaving. In the last session she said: "Okay. I have something to say. I think I'll miss you. You helped me." To which the therapist responded: "Well we have worked well together. And I will miss you too." She left the session knowing that she could make future appointments if the need arose.

RESEARCH AND EVALUATION

Although virtually nonexistent 30 years ago and lagging behind that with adults, research into the efficacy of psychoanalytic psychotherapy with children has been growing. Because of psychoanalysis' emphasis on the fundamental uniqueness of each individual, single case studies have been preferred traditionally. These have typically been reported as detailed narratives without the systematic or objectified approach typically taken in reports of behavioral or psychopharmacological treatments. With the ascendancy of the randomized controlled trial (RCT), this traditional psychoanalytic approach has been widely seen as lacking. While there remains persistent criticism within psychoanalytic circles of the overvaluation of the RCT, there is greater acceptance of the need for RCTs involving PDP. RCTs are seen as necessary, but not sufficient, to establish a treatment as "evidence-based." Patient values and preferences, long-term outcomes, and clinician's observations also are held as important factors as well (see Palmer, 2013). Its typical length, ambitious aims, the complexity of operationalizing goals, and the difficulty in manualizing the approach have been serious barriers to applying the RCT model to psychoanalytic psychotherapy. Despite this, there have been increasing efforts to evaluate PDP with adults as well as children and adolescents. A more recent problem is that the evolution of psychotherapy has occurred in such a way that psychodynamic principles are increasingly integrated into more focal

or modular treatments. For example, IPT is clearly rooted in psychodynamic principles and has been adapted in adolescents. Many other branded therapies, for example, parent–child interaction therapy, mentalization-based treatment, adolescent-focused therapy, attachment-based family therapy, and floor time, among others, have strong psychodynamic foundations. Further, the emphasis on the therapeutic alliance that began with psychoanalysis now is accepted as a key component for virtually all psychotherapy. Many of these treatment approaches have been investigated empirically although, are not generally included in reviews of the traditional research base for PDP. Much of the literature on PDP has been conducted on small numbers of real-world patients with multiple diagnoses and with broad definitions of PDP. While the largest studies have been observational, there have been a limited number of RCTs and some quasi-experimental design studies. The largest empirical study of psychoanalytic therapy with children has been carried out at the Anna Freud Centre (formerly known as the Hampstead Clinic) by Fonagy, Target, and others there. Recently there have been several systematic reviews of the literature, the most recent by Midgley (2011), Abbas (2013), and Palmer (2013). Following extensive searches and rigorous inclusion criteria, they identified over 30 primary studies that involved over 1,200 patients. This represents a significant increase in studies from 10 years ago but still must be considered limited, especially when compared to studies of other psychiatric treatments. These systematic reviews come to generally, though not completely, concordant conclusions:

1. There is some evidence to support the effectiveness of psychoanalytic psychotherapy of children and adolescents.
2. Children with internalizing disorders tend to respond more favorably than those with externalizing disorders; further, externalizing disorders require greater intensity of treatment to be effective. In particular, there is evidence to support PDP for children with depression, GAD, separation anxiety, anorexia nervosa, traumatized children (especially younger ones); conversely there is little evidence to support its use in OCD.
3. Support and involvement of parents in treatment is essential and may highlight the narrowing distinction between family and individual psychodynamic approaches.
4. Younger children respond more favorably to this treatment than older children and adolescents, but sufficient frequency of sessions must be achieved to produce positive outcomes. For adolescents length of time in treatment appears relevant, although there seems to be little benefit from increased frequency of sessions.
5. Children with conduct disorders appear least likely to respond; children

with psychotic disorders, receptive language disorders, and intellectual disability (i.e., mental retardation) are also unlikely to respond to this form of treatment. Children with autism generally do not respond well, although there is evidence that there may be a place in the broader autism spectrum group for approaches influenced by psychodynamic principles.

6. The typical "good patient" (i.e., "neurotic," higher functioning, less disturbed child) responds as well to low-intensity treatment as high-intensity psychoanalysis. As such, from a cost-effectiveness perspective, it is hard to justify psychoanalysis for this population.
7. Conversely, the more disturbed child (e.g., what have been referred to in the past as "borderline," multiple complex developmental disorder, multidimensionally impaired) may require high-frequency treatment to produce any positive results; lesser forms of treatment appear to offer little benefit and may even worsen the outcome.
8. Some studies indicate that patterns of improvement may reveal a "sleeper effect," whereby positive effects continue to accrue after treatment has ended.

Of note, recent national guidelines have included PDP as recommended treatment for two conditions. The NICE guidelines now include PDP as an evidence-based treatment for depression in children and adolescents. In 2015, the AACAP Practice Parameter for Eating Disorders included adolescent-focused therapy as a recommended treatment in cases of anorexia nervosa when family-based treatment was not available or acceptable to the patient and family.

While the empirical data supporting psychoanalytic psychotherapy with children is growing, larger-scale studies for more targeted populations and improved case studies continue to be needed. Recently, a number of studies have been initiated to better address the research gaps. In the best example, Goodyer has embarked on a large RCT (540 adolescent subjects) testing the efficacy of PDP in the prevention of relapse for moderate to severe depression. The other two arms of the study include cognitive-behavior therapy and specialist clinical care (a nonspecific supportive approach). This study is also using magnetic resonance imaging to identify neural mechanisms that may predict response to targeted treatments in individual patients. Other studies are underway, targeting children with severe anxiety, bulimia, personality pathology, and social anxiety. All of these studies will be welcome additions to support the use of PDP. At the same time, psychodynamic approaches will continue to be incorporated into other forms of therapy as the field continues to strive to find "what works for whom."

FUTURE DIRECTIONS

Psychoanalytically informed psychotherapy has a long tradition in child and adolescent psychiatry. Many practitioners over many decades have found this orientation extremely helpful for large numbers of child and adolescent patients. While it has remained vital intellectually and clinically, there is better empirical support for PDP with adults and additional research support for its effectiveness with children and adolescents is needed. Larger studies are underway and by the next edition of this textbook, this approach may be on much stronger empirical ground, clarifying the indications and contraindications for this approach (especially understanding those patients or clinical problems for which this technique is especially or uniquely relevant), its mechanism of action, and critical areas of technique. Increasingly elements of psychodynamic thinking and practice have been incorporated into other brands or modalities of psychotherapy. This trend is expected to become more prevalent as the field moves to more targeted, individualized therapies.

References

Abbass AA, Rabung S, Leichsenring F, Refseth JS, Midgley N. Psychodynamic psychotherapy for children and adolescents: a meta-analysis of short-term psychodynamic models. *J Am Acad Child Adolesc Psychiatry*. 2013;52(8):863–875.

Allen F. *Psychotherapy with Children*. New York: WW Norton & Co.; 1942.

Altman N, Briggs R, Frankel J, Gensler D, Pantone P. *Relational Child Psychotherapy*. New York: Other Press; 2002.

*Bateman A, Fonagy P. Mentalization-based treatment. *Psychoanal Inq*. 2013;33:595–613.

*Bromfield R. *Doing Child and Adolescent Therapy: Adapting Psychodynamic Treatment to Contemporary Practice*. 2nd ed. New Jersey, NJ: John Wiley and Sons, Inc.; 2007.

*Delgado SV, Strawn JR, Pedapati EV. *Contemporary Psychodynamic Psychotherapy for Children and Adolescents: Integrating Intersubjectivity and Neuroscience*. New York: Springer-Verlag Berlin Heidelberg; 2015.

Eisenberg L. The past 50 years of child and adolescent psychiatry: a personal memoir. *J Am Acad Child Adolesc Psychiatry*. 2001;40(7):743–748.

Fonagy P, Cottrell D, Phillips J, Bevington D, Glaser D, Allison E. *What Works for Whom? A Critical Review of Treatments for Children and Adolescents*. 2nd ed. New York: The Guilford Press; 2014.

Fonagy P, Target M. Predictors of outcome in child psychoanalysis: a retrospective study of 763 cases at the Anna Freud Centre. *J Am Psychoanal Assoc*. 1996;44:27–77.

Freud A. *The Psychoanalytic Treatment of Children*. London: Imago Publishing; 1946.

Goodyer IM, Tsancheva S, Byford S, et al. Improving mood with psychoanalytic and cognitive therapies (IMPACT): a pragmatic effectiveness superiority trial to investigate whether specialised psychological treatment reduces the risk for relapse in adolescents with moderate to severe unipolar depression: study protocol for a randomised controlled trial. *Trials*. 2011;12:175.

Greenberg JR, Mitchell SA. *Object Relations in Psychoanalytic Theory*. Cambridge,

Massachusetts: Harvard University Press; 1983.

Hagan CC, Graham JM, Widmer B, et al. Magnetic resonance imaging of a randomized controlled trial investigating predictors of recovery following psychological treatment in adolescents with moderate to severe unipolar depression: study protocol for Magnetic Resonance—Improving Mood with Psychoanalytic and Cognitive Therapies (MR-IMPACT). *BMC Psychiatry*. 2013;13:247.

Kernberg PF, Chazan SE. *Children with Conduct Disorders: A Psychotherapy Manual*. New York: Basic Books; 1991.

*Kernberg PF, Ritvo R, Keable H; American Academy of Child and Adolescent Psychiatry (AACAP) Committee on Quality Issues (CQI). Practice parameter of psychodynamic psychotherapy with children. *J Am Acad Child Adolesc Psychiatry*. 2012;51(5):541–557.

Klein M. *The Psychoanalysis of Children*. London: Hogarth Press; 1932.

Lanyado M, Horne A. *The Handbook of Child and Adolescent Psychotherapy: Psychoanalytic Approaches*. 2nd ed. London: Routledge; 2009.

Leichsenring F, Leweke F, Klein S, Steinert C. The empirical status of psychodynamic psychotherapy—an update: Bambi's alive and kicking. *Psychother Psychosom*. 2015;84:129–148.

Lewis O. Integrated psychodynamic psychotherapy with children. *Child and Adolescent Clinics of North America*. 1997;6(1):53–68.

Lichtenberg JD, Lachmann FM, Fosshage JL. *Self and Motivational Systems: Toward a Theory of Psychoanalytic Technique*. Hillsdale, NJ: Analytic Press; 1992.

Lock J, La Via MC; American Academy of Child and Adolescent Psychiatry (AACAP) Committee on Quality Issues (CQI). Practice parameter for the assessment and treatment of children and adolescents with eating disorders. *J Am Acad Child Adolesc Psychiatry*. 2015;54(5):412–425.

*Meeks JE, Bernet W. *The Fragile Alliance: An Orientation to Psychotherapy of the Adolescent*. 5th ed. Malabar, FL: Krieger Publishing Company; 2001.

Midgley N. The 'inseparable bond between cure and research': clinical case study as a method of psychoanalytic inquiry. *J Child Psychother*. 2006;32(2):122–147.

*Midgley N, Kennedy E. Psychodynamic psychotherapy for children and adolescents: a critical review of the evidence base. *J Child Psychother*. 2011;37(3):232–260.

*Palmer R, Nascimento LN, Fonagy P. The state of the evidence base for psychodynamic psychotherapy for children and adolescents. *Child Adolesc Psychiatr Clin N Am*. 2013;22:149–214.

Sandler J, Kennedy H, Tyson R. *The Technique of Child Analysis: Discussions with Anna Freud*. London: Hogarth Press; 1980.

Shedler J. The efficacy of psychodynamic psychotherapy. *Am Psychol*. 2010;65(2):98–109.

Von Hug-Hellmuth H. On the technique of child-analysis. *Int J Psychoanal*. 1921;2:287–305.

Weiss S. Vicissitudes of termination: transferences and countertransferences. In: Schmukler AG, ed. *Saying Goodbye: A Casebook of Termination in Child and Adolescent Analysis and Therapy*. Hillsdale, NJ: The Analytic Press; 1991:265–284.

*Winnicott DW. *Playing and Reality*. London: Tavistock Publications; 1971:54.

▲ 54.2 Brief Psychotherapies for Childhood and Adolescence

ANTHONY L. ROSTAIN, M.D., AND MARTIN FRANKLIN, PH.D.

INTRODUCTION

Improvements in the current understanding of developmental psychopathology in children and adolescents have prompted a shift from nonspecific interventions toward more problem-focused treatments that target specific diagnostic categories and, within diagnoses, symptoms. Indeed, in the past several decades a diverse range of empirically supported cognitive-behavioral and pharmacological therapies have been developed that cover a broad range of childhood-onset disorders. Herein, first a brief overview of the theoretical rationale for cognitive-behavioral therapy (CBT) treatment of youth with psychiatric disorders is provided, combined treatments are discussed briefly, and then the extant literature on short-term psychosocial treatments for internalizing and externalizing disorders is summarized. Though still widely used in clinical practice, psychodynamic approaches, "play therapy," and other therapies that have not evaluated empirically via randomized controlled trials (RCTs) are not reviewed. Interpersonal psychotherapy (IPT), a short-term psychosocial approach to depression that was developed for adults but has now been extended down the developmental spectrum for use with adolescents, is a notable exception to this rule. Experimental interventions that are not best classified as psychotherapies, such as response inhibition training, attentional bias training, or other such procedures are not discussed. Their role in augmenting short-term treatments such as pharmacotherapy and psychotherapy needs to be further explicated, but at present they do not constitute "treatments" in the truest sense, nor does their collective evidence base justify placing them on equal footing with these other empirically validated interventions.

PROTOCOLS AND TECHNIQUES

Cognitive-Behavioral Therapy for Anxiety Disorders

Traditional treatment models in child psychology and psychiatry were guided by theories positing that psychological problems have their origin in historical relationship problems that must be uncovered and processed in therapy. CBT clinicians do not adhere to this view, believing instead that more direct methods should be employed to reduce symptoms, such as identification and correction of mistaken beliefs, confrontation with feared and/or avoided stimuli in a therapeutic environment, and/or by teaching the patient a set of adaptive coping skills for specific symptoms that are associated with distress and impairment. Befitting its deep roots in learning theory, there is also an inherent belief in CBT that change is

measurable, and thus a strong interest in assessing changes in symptoms over time in order to evaluate progress in treatment and to identify when modifications to the chosen approach need to be made. There is strong emphasis that the application of treatment procedures brings about changes in behavior, cognition and, down the line, emotion, rather than the positive interaction with the treatment provider. That said, CBT is still thought to be a collaborative process in which the therapist–patient alliance also plays a role in treatment success, and these nonspecific factors must therefore also receive attention in CBT supervision and training. With children especially, CBT should contain elements of fun and enjoyment in order to foster this important alliance and provide positive reinforcement for effort both during and between sessions.

Notably, the CBT approach, though rooted in cognitive and behavioral theories about the nature of psychopathology, is procedurally heterogeneous. CBT can involve working with parents, changing social, peer, and other environmental influences, as well as the use of behavioral contingencies and cognitive processing. Further, considerable tailoring to the individual disorder and associated symptoms is necessary for successful treatment, as is a full appreciation of the developmental and family factors likely to influence symptoms and hence require attention in therapy. Accordingly, CBT protocols vary, sometimes considerably, in their emphasis on behavioral versus cognitive mediators of change, in part depending on the nature of the disorder being treated and also in response to the many contextual factors discussed above. Nevertheless, an emphasis on the fundamental tenet that behavior and cognition are the horses whereas emotion is the cart cuts across CBT protocols for myriad disorders and problems.

Psychosocial Approaches for Depression

Just as with anxiety disorders, CBT programs for depression in children and adolescents vary considerably with respect to procedural emphasis but share some core theoretical and therapeutic assumptions. For example, a shared theoretical assumption is that depressed mood can be reduced by helping patients to identify and challenge faulty thinking, reduce certain behaviors such as avoidance and increase others such as engaging in pleasant, mood-enhancing activities, and improving interpersonal interactions and conflict resolution skills. Here again, protocols vary in the degree to which they emphasize one therapeutic component over another in order to affect change in mood symptoms and to improve psychosocial functioning but, according to experts such as Curry and Reinecke, successful CBT for mood disorders in youth should include providing psychoeducation about the nature of depression, assisting adolescents in monitoring moods; identifying and modifying maladaptive cognitions; and improving problem solving, affect regulation, and social skills.

IPT is a form of treatment for depression that posits that adolescents' depressive symptoms can be reduced and interpersonal functioning improved by focusing treatment on one of four problem areas: grief, role disputes, role transitions, or interpersonal deficits. The treatment is present focused and the development of skills to manage problems in these identified areas is a clear goal of treatment but, befitting its roots in psychodynamic theory, the importance of the therapeutic relationship is also strongly emphasized; a manual that is based on IPT is used to guide treatment with adolescents as conducted with adults. Unlike most psychodynamically oriented treatments, IPT is a brief psychotherapy that has been delivered in 12 sessions conducted over 12 to 16 weeks in school-based clinics for depressed adolescents and in an even shorter form when modified for use in a depression prevention study that was conducted by Horowitz and colleagues.

Internalizing Disorder: CBT Case Example

As a specific example of how the CBT model might be applied, take the case (composite) of a 9-year-old girl with separation anxiety difficulties that have resulted in occasional school refusal behavior. A diagnostic work-up determines that the primary diagnosis is separation anxiety disorder, with a secondary comorbid diagnosis of generalized anxiety disorder (GAD). The review of specific symptoms indicates that the child is intensely fearful of separation from her parents although more so from her mother than her father, is worried that something bad might happen to her parents or to herself during separations, reports increased muscle tension and uncontrollable worry that are associated with nausea, headaches when thinking about being separated from parents or when separated. Accordingly, the child will beg, plead, and cry to prevent her parents from leaving her at home with family members or with sitters. When she believes that something at school is likely to be stressful or unpredictable (e.g., performances, substitute teachers), she refuses to attend school, a behavior that her parents have decided for now not to challenge. Both parents report that the school refusal, the refusal to allow parents to leave her, and her increasing sense of withdrawal from activities in which she might be separated from parents (e.g., travel team soccer) are sources of significant tension in the home that is now spilling over to affect the marriage and the other younger siblings.

The CBT therapist in this case would first need to determine the parents' and patient's motivations for change, their recent history of how they have been managing separation issues, and their current understanding of the nature of the problem and the associated symptoms. If this discussion leads the therapist to believe that CBT targeting separation anxiety primarily (as opposed to primarily targeting GAD or family processes in this case) is appropriate for reducing these specific symptoms as well as overall family distress, he or she would likely work within the framework of one of the established protocols for separation anxiety such as Kendall's Coping Cat program. A detailed review of what happens in the family currently around separation, such as who drops off the child at school, what is said in the morning about compliance with parental requests and what is the child's typical response to requests to get

ready to go to school, how sitters are selected and how the child responds when the plan for the parents to go out without her are conveyed. The child's physical, emotional, and behavioral reactions must also be surveyed so that the therapist can begin to consider which elements to target first, which situations evoke the most distress, etc. It is also imperative to explain that the physical sensations of anxiety often mirror those of illness, and that differentiating between the two will assist the family in knowing when to proceed with the original plan and when the child is physically ill for other reasons.

The therapist must also assess the child and parents' motivation for change as well as the various things that the child finds reinforcing, which will then be used as part of the reward system that will help sustain motivation during the most anxiety-provoking aspects of treatment. A symptom hierarchy will then be developed that will reflect the key variables that influence anxiety around separation, such as duration of separation, degree of separation, which parent is instituting the separation, whether the event is parent or child oriented (e.g., child's soccer vs. parents' night out), and the rungs on the hierarchical ladder will be assigned a rating on the "Fear Thermometer" from 0 (no anxiety) to 10 (maximum anxiety). The reward system will be worked out in detail with both parent and child, so that accomplishments (e.g., staying at home with a sitter without calling parents' cellphone) and effort (e.g., monitoring progress between session) have a clear "point value" that can then be depicted on a sticker chart and reimbursed by approved items from the reward system.

Planned exposure to feared situations will play a key role in treatment, as will discussion of the nature of cognitions likely arise in anticipation of and during separation. Although CBT protocols vary greatly with respect to the degree of emphasis on cognition and the procedures used to identify and change mistaken or inaccurate beliefs, the essential ingredients all involve helping the child to recognize anxious thinking when it occurs, label it as such, treat these cognitions as testable hypotheses rather than as facts (e.g., "If I'm anxious now, I will be worse in an hour"), and then perform the experiments that will allow the child to see whether the anxious thinking is realistic or not. The development of alternative "coping thoughts" is then used to help the child manage anxiety in anticipation of or in the context of exposure, allowing the child to gain from distance from these unpleasant cognitions and continue to confront the feared situations without engaging in avoidant behaviors. In their seminal paper on emotional processing theory and its application to understanding the efficacy of exposure-based treatments, Foa and Kozak suggested that by doing so, exposure will provide corrective information (e.g., "My mom left me for the day but she came back, and I was able to handle it"), which in turn weakens the fear structure and results in between-session habituation that is characteristic of successful CBT for anxiety disorders. Some CBT

protocols will also include anxiety management training (AMT) such as relaxation exercises, but multiple RCTs in anxiety disorders for youth that have not included AMT have been found efficacious. Moreover, a dismantling study in adult panic disorder conducted by Schmidt and colleagues indicated that inclusion of breathing retraining in an exposure-based treatment actually compromised outcome rather than enhancing it, perhaps because the use of anxiety management strategies to prevent anxiety from escalating is inconsistent with the panic protocol's more general message, which is that anxiety and panic are simply false alarms, are not inherently dangerous, and are best ignored.

Treatment progresses as the child chooses exposures from the fear hierarchy and succeeds at conquering the lower and middle range situations, which in her case will likely involve permitting her parents to leave for specified amounts of time, minimizing the amount of delaying she does in preparing for school by rewarding her for on-time behavior (e.g., lunch packed and dressed by 8:15 AM on school days can be charted, as can 5 consecutive days of "on-time" behavior as defined above), and reducing avoidant behavior such as declining invitations to visit at friends' houses. Parents will be instructed by the therapist in how best to proceed when arguing and tantruming behavior does occur. They will be vigorously encouraged to follow the reward system faithfully and to provide social reinforcement (e.g., praise) for the child's efforts to confront feared situations using the new coping strategies of exposure and challenging anxious thinking.

Most protocols for anxiety in youth including separation anxiety recommend 12 to 20 weekly treatment sessions, the first several of which are devoted to psychoeducation and skill building, followed by exposure and cognitive training, then by the development of a tailored relapse prevention program to help the child and family maintain gains made during the core treatment phase. The efficacy data reviewed below are quite encouraging with respect to the degree and durability of improvement associated with CBT, and the current challenges facing CBT researchers involve determining which patients are likely to respond well to initial treatment, which are likely not to respond or to have clinically significant residual symptoms at the end of a full course, and what course of treatment to pursue next with those who fail to respond completely.

Cognitive-Behavioral Therapies for Disruptive Behavior Disorders (ADHD, ODD, and CD)

At the present time, the most empirically validated psychosocial treatments for disruptive behavior disorders (DBDs) in children and adolescents are based on CBT techniques adapted to the specific context of interest (e.g., family, school, residential treatment facility). These interventions include

carefully delineating and measuring positive and negative behaviors, identifying key antecedents and contingencies reinforcing these behaviors, developing a plan for increasing positive behaviors and extinguishing negative behaviors, finding appropriate rewards and consequences for the child's behaviors (including token economy systems), and closely monitoring the patient's responses to the plan. While differing in their emphasis, these therapies are directed at altering the behaviors of key adults in the child's world (e.g., parents, teachers, caregivers) so as to affect changes in the child's or adolescent's behavior. In addition, cognitive techniques are often employed to teach the child alternative responses to typically stress-inducing situations such as being directed to comply with instructions from an authority figure, being told "No" to a request or demand, not having wishes/desires gratified immediately (i.e., being frustrated), or having to share toys or activity space with peers.

As with treatment for internalizing disorders, DBD treatments are usually administered on a weekly basis for anywhere from 12 to 30 sessions. Initial sessions include delineating basic principles, teaching observation skills to parents and/or self-observation skills to children and adolescents, measuring baseline rates of key positive and negative behaviors, setting initial target goals (e.g., 25 percent fewer tantrums, 50 percent greater compliance with directives), developing rewards and consequences schedules, practicing the administering of rewards and consequences, and monitoring impact of the schedules on target behaviors. In most treatment approaches, instruction is provided in recognizing and managing one's own negative affects/behaviors/cognitions, and in developing alternative strategies to typical ineffective reactions to stress- or anger-inducing situations. For example, typical parental emotional reactions to oppositional defiance on the part of a child are highlighted and discussed in the context of a broader framework of spiraling negative interaction cycles leading to parental aggression or to parental relinquishing of authority, either of which are reinforcing the child's noncompliance. Patient-focused approaches provide tools for developing alternative responses to challenging or frustrating situations. Opportunities for practicing alternative responses through role modeling and rehearsal are provided during sessions, and assignments are given to carry out between meetings. Subsequent sessions include reviewing homework, obtaining feedback from participants about their ability to implement behavior management strategies, identifying barriers to implementation of the plan, and designing modifications to increase success during the intervening time between visits. Periodically, participants are asked to measure progress made toward desired treatment goals, to reinforce successful strategies where possible and if appropriate, to move on to other target behaviors. The greatest advantages of CBT for DBD are its adaptability to a variety of diagnoses, problem behaviors, and contexts; its

emphasis on practical and measurable outcomes; its inclusion of multiple participants (i.e., patients, families, teachers, and significant others); and the availability of a wide variety of treatment manuals.

Combined Treatments. A comprehensive review of the evidence base for pharmacotherapies for pediatric disorders is beyond the scope of this chapter, but there is a wealth of evidence supporting the efficacy of pharmacotherapy for many pediatric–psychiatric conditions, and that combined treatments have also proven efficacious, especially with DBDs. Careful inspection of the associated outcome data also reveals that although active medication has been found consistently superior to placebo virtually across the board, response to the available pharmacotherapies has been neither universal nor complete. Moreover, some children and adolescents have unfavorable side effects to these treatments that preclude their use altogether or compromise outcomes by limiting the doses that can be used. Accordingly, there is certainly still room for treatment development within pharmacotherapy, and for considering combined treatment approaches that include pharmacotherapy plus time-limited psychosocial interventions of established efficacy. If the evidence bases for pediatric disorders were truly comprehensive, selecting an appropriate treatment for a child or adolescent from among the many possible options and combinations would be reasonably straightforward. In the complex world of research and clinical practice, however, choices are rarely so evident. This is perhaps more so the case with those disorders in which few RCTs have been conducted, and hence the evidence base is insufficiently developed to provide confident conclusions about which treatments are most likely to be effective for a given child or, more specifically, whether combined treatment is necessarily superior to monotherapy with CBT or with medication. Few studies published to date have compared the relative and combined efficacy of CBT and pharmacotherapy for anxious youth, for example, and thus pediatric mental health practitioners are left without a substantive evidence base to inform their treatment recommendations regarding this critical question. In light of this reality, initial decisions to select monotherapy with CBT or pharmacotherapy or combined treatment are influenced by primary diagnosis, symptom severity, comorbidity, family factors, and readiness for change. Combined treatments should also be considered if an attempt at monotherapy with medication or CBT proves only minimally efficacious; in such cases, a second treatment can be introduced to address symptoms that are thought to be limiting efficacy, such as adding an SSRI for an anxious child whose comorbid depression negatively affects his ability to participate actively in CBT.

Developmental Considerations in Time-Limited Treatments

The treatment manuals for youth that have been evaluated empirically

typically run across the developmental spectrum, although the majority of the studies conducted thus far have involved children between the ages of 8 and 13 years. In the view of some in the field and perhaps best articulated by Westen and colleagues, manual-based treatments have acquired the reputation of being somewhat rigid. However, in contrast to this characterization, those who regularly conduct RCTs involving treatment manuals actually encourage therapists working with this framework to remain flexible with the application of the session goals, as life must be "breathed" into the manual to better fit the needs and functioning of the patient. The debate about the advantages and disadvantages of using manualized treatments in clinical practice has been the topic of many papers and books in the last decade. The authors are confident that this will lead researchers as a field to be both responsive to the evidence base for clinical interventions as well as to the practical realities of the clinical contexts in which most children and adolescents who seek treatment for psychiatric problems will receive their clinical care.

With respect to the flexible application of core treatment principles, it is essential that therapists using CBT protocols are fully cognizant of developmentally related differences across the targeted age ranges in physiological functioning, emotional vulnerability, social and peer pressures, and comorbid conditions, as well as any other changes that these children may be experiencing. The nature and extent of family involvement in treatment will also necessarily vary depending on the age and developmental level of the child. This important treatment consideration is explained below.

Experts such as Kazdin and Kendall recommend that the structure and process of manualized treatments should differ somewhat for children and adolescents, respectively, although the general format and goals may well be the same across development. Even within more truncated age ranges, developmental factors still have to be taken into account as treatment is conducted. Perhaps this is best illustrated around the issue of parent involvement in treatment. Indeed, parental involvement is likely to vary based on developmental factors on the following dimensions: (a) amount of time a parent/parents will be present in each treatment session; (b) the role parents will play in assisting patients with implementation of treatment techniques such as self-monitoring, incentive plans to enhance motivation, and use of core techniques at home and in school; (c) the role of parents in interacting with relevant school personnel regarding the child's efforts to combat specific symptoms with treatment techniques in the school setting; (d) degree of emphasis on privacy and respect for the importance of individuation; and (e) parental assistance in helping patients negotiate the increasing importance of peers and dealing with concerns that others might discover the patient is in treatment. For example, a 10 year old might benefit from more formal parental assistance with self-monitoring tasks

and with the development of an incentive plan to promote motivation to use the treatment strategies in "high-risk" situations at home; doing so might necessitate more time in session with the parent present to ensure that all parties are in agreement and that the parent's role is made clear. Treatment for the typical 17 year old should take into consideration the adolescent's need for individuation and privacy, and provide them control over the amount of disclosure about treatment that they are comfortable with while also helping parents feel like they remain "in the loop." As a guiding principle, it is important to note that the treatment protocol is designed to allow for what Kendall calls "flexibility within fidelity." This means that while the general principles that serve as the foundation for the treatment and specific procedures that flow from these core principles must be adhered to, the application of the treatment for a given patient must at the same time be sufficiently flexible to accommodate that patient's developmental level, symptom picture, family context, and learning style.

Review of the Evidence Bases for Time-Limited Psychosocial Treatments

Anxiety Disorders and Related Conditions. Pediatric anxiety disorders have received quite a bit of attention from clinical researchers in the last 25 years, and the efficacy of several CBT protocols for youth has now been established. As reviewed above, many of these protocols contain common elements, such as exposure to feared situations, contingency management, positive reinforcement, shaping, modeling, and hierarchy development. Moreover, these common principles and the similarity of presentation across anxiety disorders led many investigators to conduct anxiety disorder treatment studies that included multiple anxiety disorders rather than any individual diagnostic entity (e.g., social anxiety disorder), although there have been some diagnosis-specific treatment developments.

One of the best researched multicomponent CBT packages for anxiety disorders is the protocol known as the Coping Cat program, which was developed by Kendall for the treatment of separation anxiety disorder, social anxiety disorder, and GAD (formerly overanxious disorder). The Coping Cat protocol has been the subject of multiple treatment studies and the stepping stone for many others in the last 15 years. The first randomized comparison of CBT for pediatric anxiety disorders compared the Coping Cat protocol to a waitlist control condition. Forty-seven children ages 9 to 13 participated, and results indicated that those who received CBT had superior self-reported, parent-reported, and observational outcomes compared to those who received WL. Moreover, treatment gains were maintained long term. A second RCT from the same laboratory replicated and extended these findings, this time with a considerably larger sample ($N = 94$).

The Coping Cat protocol was directly compared to sertraline and combined treatment in the context of the multisite Child Anxiety

Management Study (CAMS), an NIMH-sponsored, multisite, placebo-controlled evaluation of the relative and combined efficacy of CBT and sertraline for separation, social, and GAD. With 488 participants of ages 7 to 17 with either social anxiety, separation anxiety, or GAD randomized to one of four conditions, CAMS has provided extremely valuable information about CBT, pharmacotherapy, combined treatment, and moderators and mediators of response to these conditions. Baseline data indicated that the sample had moderate to severe anxiety and was highly comorbid; anxiety symptom severity and comorbidity patterns did not differ by race or gender. Major depressive disorder was a rule out, which limits the degree to which conclusions can be drawn about this population. Acute treatment outcomes indicated that all three of the active conditions were superior to pill placebo control (PBO) on both continuous outcome measures as well as dichotomous CGI-I responder rates, which were 81, 60, 55, and 24 percent for COMB, CBT, SER, and PBO, respectively. Adverse events between SER and PBO did not differ, which indicates that the side-effect profile does not contraindicate the use of SER as one of the first-line options for pediatric anxiety; at the same time, adverse events in CBT were lower than those reported in SER, which thus offers a potential reason to start with CBT if available.

Accordingly, all three of these treatments can be offered to families of anxious youth, which allows clinicians to take into consideration family preferences and burdens, treatment availability, and associated costs. The search for moderators, baseline characteristics that associated with differential response to conditions, or nonspecific predictors, yielded a few key findings: low anxiety severity and caregiver strain predicted, but did not moderate, anxiety severity at posttreatment. With respect to moderators, having a principal diagnosis of social phobia was associated with less favorable outcome to CBT or PBO, but not to the conditions that included medication (COMB or SER). Accordingly, the authors recommend that medication may be more helpful or necessary for those with a principal diagnosis of social phobia. Importantly, African-American participants did not differ from Caucasians with respect to treatment response, remission, or relapse. Follow-up analyses are presented below in the section on maintenance of gains.

Since the completion of Kendall's initial randomized study, a number of other investigators from around the world have examined the efficacy of CBT for anxiety disorders in youth using the same sampling frame as Kendall and colleagues. Barrett completed an RCT comparing a family-based, group CBT intervention with group CBT alone and WL for 60 youth ages 7 to 14 with social, separation, or overanxious disorder. Results indicated that the CBT interventions were superior to WL though not different from one another. Flannery-Schroeder and Kendall also examined the relative efficacy of individual and group CBT compared to WL, albeit in

a smaller sample (*N* = 45), and found that individual was superior to group CBT on some measures but that both were clearly superior to WL. Group CBT has received empirical support in several other studies that employed mixed anxiety disorder samples, although the feasibility of running groups versus conducting individual therapy outside the context of academic medical settings in large population centers is highly dependent on adequate patient flow.

Following on from Barrett's study of CBT with and without a family component, several investigators have directly compared the two modalities in youth with anxiety disorders. Though the answer might initially appear obvious, the findings from these studies have proven to be equivocal. Some of these studies have found that the family component was associated with superior outcomes relative to individual CBT, whereas other studies failed to find an additive effect. The inconsistent findings are puzzling given the central role that family factors play in the development and maintenance of anxiety in youth. Low statistical power in some of these studies might also underlie the failure to differentiate these two approaches; it may also be the case that there is enough family contact in the "individual" child-focused protocols to promote family support for and involvement in the child's attempt to confront fear-provoking situations rather than to avoid them routinely. In light of the importance of knowing whether formal parental involvement above and beyond what is already incorporated into child-focused CBT is necessary, Kendall and colleagues (2008) published the largest study of its kind in 161 youth ages 7 to 14 with separation anxiety, social anxiety, or generalized anxiety disorder. Family-based CBT was compared to individual CBT and to a psychoeducational control condition that constituted a more active comparison group than had typically been employed since the majority of other studies used a WL condition. Results again provided evidence for the efficacy of both forms of CBT, although the control condition did yield statistically significant and clinically meaningful reductions in anxious symptoms as well. In the direct comparison of the CBT protocols, the two failed to differentiate from one another on most outcome measures. The large sample size allowed the investigators to identify several putative moderators of response; notably, although comorbidity did not differentially affect outcome by condition, participants whose mother and father were diagnosed with anxiety disorders themselves fared better when assigned to FCBT than ICBT. These findings further underscore the efficacy of CBT, and suggest that more intensive family involvement in CBT may not be needed for all patients and families.

The studies described above included at least a subset of participants diagnosed with social anxiety disorder (social phobia) or the previous diagnostic term, avoidant disorder. However, individual treatment protocols for social anxiety disorder in youth have also been developed, and

their efficacy evaluated in RCTs. Findings from these studies support the use of social phobia-specific treatments delivered both in the clinical and school context, although they also indicate that there is room for further improvement and hence treatment development. Based on meta-analyses of adult social phobia treatment studies, Beidel and colleagues concluded that exposure in some form may be a key ingredient in CBT programs for social anxiety disorder in youth, and that treatment interventions which combine social skills building, as well as anxiety reduction, may be necessary for optimal outcome. Following on from an initial encouraging test of such a program for socially anxious youth in which treatment gains were maintained for up to 5 years, Beidel and colleagues in 2007 completed an RCT in 122 youth ages 7 to 17 with social anxiety disorder, comparing social effectiveness therapy (SET) to fluoxetine (FLU) and to a PBO. Both of the active treatments were superior to PBO, and differences favoring SET over FLU were found. Initial responders' relapse rates following FLU and SET discontinuation were relatively low at 1-year follow-up (17 and 10 percent for FLU and SET, respectively), suggesting the durability of improvements over time.

In addition to studies that have specifically targeted social anxiety disorder, several other RCTs have been conducted examining the efficacy of CBT for other anxiety disorders in youth. Silverman and colleagues, studying treatments for specific phobias, examined the relative benefits of an operant-based contingency management treatment and a cognitive-based self-control treatment, in comparison to an education support control group. Graduated in vivo exposure was used in both the self-control and the contingency management conditions but not in the education/support condition. Phobic children (N = 81) between 6 and 16 years of age and their parents were evaluated using child, parent, and clinician measures. Although all three conditions were found to impart improvement in the child's functioning as measured by the reports of children, parents, and clinicians, clinically significant improvements were noted only in the two cognitive-behavioral treatment conditions. Specifically, on a measure of clinical distress at posttreatment, 80 percent of the participants in the self-control and 80 percent of the participants in the contingency management conditions reported very little or no distress compared to only 25 percent in the education/support condition. Notably, 88 percent of the participants in the self-control condition no longer met diagnostic criteria at posttreatment, compared to 55 percent in the contingency management and 56 percent in the education/support condition. This study underscores the importance of using active comparison groups in RCTs for anxiety disorders in youth, who may be responsive to less active forms of intervention than adults with anxiety disorders have proven to be.

In an example of a CBT intervention tailored to a particular end of the

developmental spectrum, Pincus and colleagues (2010), conducted the first RCT of any treatment for panic disorder in adolescents, randomizing 26 adolescents to either 11 weeks of panic control therapy for adolescents (PCT-A) or a self-monitoring control condition. PCT-A was superior to self-monitoring on IE-rated panic symptoms as well as self-rated anxiety, depression, and anxiety sensitivity. Posttreatment discussions with participants led the investigative team to develop a more intensive protocol in light of feedback that the treatment could have been delivered more efficiently.

In a particularly novel innovation, Öst and colleagues evaluated the effects of a CBT program referred to as "one-session treatment" (OST) with phobic children and adolescents ages 7 to 17 years. The treatment was designed to be maximally effective in one session, with patients progressing to the top of the symptom hierarchy by the end of the session, which lasted 3 or more hours. As was the case in several of the studies discussed above, children were randomized to CBT conditions in this study (conducted either in the presence or absence of a parent) or to WL. Results indicated that the active treatments produced significant gains immediately after treatment and at 1-year follow-up. WL was associated with no change at posttreatment. Here again, no differences were found between CBT conditions with or without a parent present in this comparison as well. Even with the relatively short treatment, CBT gains in this study were found to be comparable to other, longer treatments with more sessions, and perhaps superior to some other published results on specific phobias. These findings were confirmed in a large, multinational (US, Sweden) study of OST versus educational support and a waitlist control in which Ollendick and colleagues (2009), randomized 196 phobic youth ages 7 to 16 to one of the three conditions: OST, education/support, or waitlist. Findings indicated a clear and durable advantage for OST over the other two conditions; education/support was also superior to WL. A subsequent randomized study was conducted by Ollendick and colleagues (2014), in which the efficacy of OST with and without parent involvement was examined in 97 youth ages 6 to 15. Contrary to expectations, there were no differences between the two conditions, with each yielding response rates (loss of diagnosis) of approximately 50 percent.

Studies of CBT for school phobic children have also been conducted, although it is clear that school phobia is far from a simple condition and that such samples are diagnostically heterogeneous, with separation anxiety and social anxiety being quite common. King and colleagues found that 4 weeks of CBT was superior to a waitlist control: school attendance was higher in the treated than untreated children (89 vs. 60 percent), and ratings of anxiety also showed meaningful differences. In contrast, Last and colleagues found no significant difference in improvement rate between CBT and standardized educational support treatment: mean percentages of

school attendance were 67 and 60 percent, respectively. The more stringent criterion of 95 percent school attendance was met by 22 percent of the CBT group and 21 percent of those receiving educational support. Bernstein and colleagues extended this work by comparing augmentation of CBT with imipramine to augmentation with pill placebo in a school refusing sample, and found that the children who received combined treatment had superior outcome. Bernstein and colleagues' sample also included children who were diagnosed with primary depression, so the conflicting evidence likely driven by methodological heterogeneity confounds recommendation of treatment approaches for school-refusing children.

Depression. RCTs of CBT for depression in youth have been conducted for years, and collectively these studies indicate that CBT is a valuable treatment, either alone or in combination with pharmacotherapy (for a comprehensive review, see Curry, 2014). Randomized studies that have specifically examined CBT in clinical samples of children diagnosed with depression (as opposed to epidemiological or school-based samples) have been fewer, however, and their results indicate that although the majority of the samples meet criteria for clinical improvement, response is usually partial and relapse is common. The very nature of depression itself makes studying the efficacy of its treatments extremely complex, since mood disorders are known to be less stable than anxiety disorders usually are. Diagnostic heterogeneity within depression and other mood disorders also contributes to difficulty comparing across samples and drawing firm conclusions based on studies with different inclusion and exclusion criteria.

It is in this context that the Treatment for Adolescents with Depression Study (TADS) was initiated and conducted. TADS was a multisite, placebo-controlled examination of the relative and combined efficacy of a CBT package against FLU, and represents the largest study of its kind (N = 439). Reflecting the extant outcome literature and expert opinion about essential and optional procedures to include in a CBT protocol, the CBT intervention selected for TADS was conducted in 15 sessions over the course of 12 weeks, and was largely skills-focused. The required elements in the CBT protocol were delivered in the first 6 weeks, and included psychoeducation about the nature of depression, goal setting, mood monitoring, increasing pleasurable activities, social problem-solving, and cognitive restructuring designed to identify and challenge mistaken beliefs. At week 7, the therapist and the adolescent chose modules for use that reflected their mutual perspective as to what other problems were central. Modules on problems in social engagement, communication, negotiation, compromise, and assertion were available. TADS also included at least two parent-only sessions and, depending on the specifics of the case, conjoint sessions to focus on specific parent or adolescent concerns. TADS' findings were convergent with the prediction that combined treatment would be superior to both of the

monotherapies and to PBO. FLU also appeared to be superior to CBT on many of the outcome measures, and CBT failed to separate from PBO on most outcome measures. The CBT-alone condition did not fare as well with respect to response rate (43 percent) as it had in other published CBT studies, although it did enhance FLU-alone outcomes. Sample severity and comorbidity might have influenced CBT-alone outcomes, although the authors also note that modifications to the manual made to ease cross-site dissemination might also have played a role in producing this unexpected outcome. In terms of prevention of depression in adolescents with a prior history of depression, Brent and colleagues (2015), found an efficacious and durable effect for a group-based CBT program compared to treatment as usual, with positive outcomes still evident after 6 years postintervention.

Disruptive Behavior Disorders. There is ample evidence from randomized trials that CBT for DBD is moderately effective in achieving improvements in targeted outcomes. As seen in treatment studies of depression, the heterogeneous nature of DBD, the complexity of variables that affect treatment response, and the wide variation in treatment study designs make it difficult to compare studies and draw firm conclusions about which specific form of CBT is most effective for any given patient. The largest and most comprehensive randomized control trial is the Multimodal Treatment Study of Children with ADHD (MTA). The study enrolled 579 children with ADHD between the ages of 7 and 9, of which 40 percent had comorbid oppositional defiant disorder (ODD) and another 14 percent met criteria for conduct disorder. Subjects were randomly assigned to receive one of four treatments over 14 months: medication management only, intensive behavioral treatment, combined medication and behavioral treatment, and standard community care. The behavioral treatment consisted of 27 group-administered and 8 individual sessions of behavioral parent training, a school-based intervention including consultation with the subject's teacher and the assignment of a paraprofessional aide trained in behavior management, and an 8-week summer treatment program.

The results of the study showed improvements in ADHD symptoms in all four groups, with the greatest improvement in the medication only and the combined groups. Behavioral treatment compared favorably with community care which consisted of medication management in two-thirds of the cases. When outcomes such as reductions in oppositional and aggressive behaviors were examined, combination treatment was superior to the other groups, and the children in this group required lower doses of medication to attain equivalent or superior outcomes. Thus, while there is no question that stimulant medication is effective in reducing the symptoms of DBD, behavioral interventions added additional benefits and had a clear impact on oppositional and aggressive behaviors as well as

anxiety symptoms cooccurring with ADHD.

With respect to other outcomes, behavioral treatments showed the most significant impact on reductions in parenting stress and increases in child social behavior and peer acceptance. Moreover, behavioral treatment of children with comorbid anxiety disorder was more effective than combined, medication, or community care.

At 36 months, 22 months after the end of the study, all groups showed substantial improvement over baseline but there were *no* differences found across the original treatment groups. Symptoms remained above those of a control group with half the children still meeting criteria for ADHD and almost one-third meeting criteria for either ODD or CD. Thus for the majority of children in the study, CBT for DBD, either alone or in combination with medication, showed moderate efficacy, although the beneficial effects were not consistent. As might be expected, poorer outcomes were seen in children with greater symptom severity at the outset, along with more adverse family situations (e.g., maternal depression, lower socioeconomic status) and lower intelligence. The complex interplay of moderating and mediating variables on long-term outcome measures is well summarized in reviews by Owens and colleagues (2003), and by Swanson and colleagues (2008).

Recent reviews of other published studies of CBT for DBD have supported the central findings of the MTA study. In addition to behavioral parent training, school-based classroom interventions (such as daily report cards and contingency management programs) have been shown to improve both classroom deportment and academic performance of ADHD children with effect sizes in the range of 1.44 (although better results were observed in behavior compared with grades). Academic interventions such as task and instructional modification, homework assistance, peer tutoring, computer-assisted instruction, and organizational skills training all seem to benefit children with ADHD. Moreover, programs emphasizing family–school partnerships through behavioral consultation, family involvement in homework, and daily report cards have shown promising results in enhancing school functioning and reducing behavior problems. Behavioral peer interventions have also been shown to improve social skills and reduce peer rejection. Large behavioral effects on both core symptoms and functional measures (e.g., following rules, sports skills, self-esteem) have been demonstrated by summer treatment programs, such as those that were incorporated in the MTA study. Last but not least, treatment sequencing for ADHD has been studied. Preliminary evidence suggests that beginning with behavioral management is preferable to medication because it generates greater parental involvement and results in better treatment outcomes.

CONCLUSION

The above review reflects the prevailing belief that CBT is a valuable addition to the treatment armamentarium for childhood disorders, either alone or in combination with pharmacotherapy depending on the specific disorder, case severity and complexity, and availability of expertise in its delivery. The latter point is however a major stumbling block, as the number of pediatric mental health professionals trained in and competent to deliver CBT to children and adolescents is nowhere near enough to meet the clinical need if the various published practice parameters were actually to be followed. The movement from CBT efficacy studies, which emphasize internal controls and are necessarily restrictive in both patient and therapist sampling, toward effectiveness studies that are more generally inclusive may help to bridge the gap between science and clinical practice and thereby promote dissemination. On the other hand, it is also imperative to conduct CBT research that incorporates state-of-the-art measurement techniques so that the underlying neurobiological factors associated with change can be better understood especially given that placebo responding is not uncommon in childhood psychiatric disorders. Stepped care models may also help reduce the burden on centers of CBT expertise by finding creative ways to disseminate CBT principles through other means, such as websites, CD-ROMs, paraprofessionals, bibliotherapies, and other methods.

Although the data are generally encouraging, CBT is by no means a universal panacea, and a return to the laboratory is imperative if the field is to make progress in refining and improving the currently available CBT programs for youth with psychiatric illnesses. It is also important to conduct larger studies that will allow for the identification of moderators and mediators of response to CBT, which can then be used to inform treatment recommendations. Recent examples include the study by Kendall and colleagues, who found that family-based CBT was more effective for anxious youth if both parents had been diagnosed with an anxiety disorder, as well as one by March and colleagues, who found that pediatric OCD patients with comorbid tics responded less favorably to pharmacotherapy (sertraline) alone than they did to CBT or combined treatment. Improvements in treatment outcome are likely to be seen if clinicians have this kind of knowledge prior to making their treatment recommendations to patients and families. Clinical researchers must remain cognizant of the imperative to design and conduct studies with sufficient statistical power to identify these kinds of influences on outcome, some of which are likely to be somewhat subtle and yet clinically meaningful; this approach is highly convergent with the broad goals of precision medicine, wherein research-based algorithms are used to identify those more or less likely respond to a given treatment regimen. Despite continuing limitations in the current knowledge base, great strides are being made in establishing the evidence base for time-limited structured

psychotherapy for children and adolescents, and in applying these treatments to patients and families across the entire diagnostic spectrum. Over time, the effectiveness and acceptability of these approaches should become more universal.

REFERENCES

*Antshel KM, Barkley R. Psychosocial interventions in attention deficit hyperactivity disorder. *Child Adolesc Psychiatr Clin N Am*. 2008;17:421–437.

Barrett PM. Evaluation of cognitive-behavioral group treatments for childhood anxiety disorders. *J Clin Child Psychol*. 1998;27:459–468.

Beidel DC, Turner SM, Sallee FR, Ammerman RT, Crosby LA, Pathak S. SET-C versus fluoxetine in the treatment of childhood social phobia. *J Am Acad Child Adolesc Psychiatry*. 2007;46:1622–1632.

Bernstein GA, Borchardt CM, Perwien AR, et al. Imipramine plus cognitive-behavioral therapy in the treatment of school refusal. *J Am Acad Child Adolesc Psychiatry*. 2000;39(3):276–283.

Brent DA, Brunwasser SM, Hollon SD, et al. Effect of a cognitive-behavioral prevention program on depression 6 years after implementation among at-risk adolescents: a randomized clinical trial. *JAMA Psychiatry*. 2015;72(11):1110–1118.

Compton SN, Peris TS, Almirall D, et al. Predictors and moderators of treatment response in childhood anxiety disorders: results from the CAMS trial. *J Consult Clin Psychol*. 2014;82(2):212–224.

Curry JF. Future directions in research on psychotherapy for adolescent depression. *J Clin Child Adolesc Psychol*. 2014;43(3):510–526.

Curry JF, Reinecke MA. Modular therapy for adolescents with major depression. In: Reinecke MA, Dattilio FM, Freeman A, eds. *Cognitive Therapy with Children and Adolescents: A Casebook for Clinical Practice*. 2nd ed. New York: Guilford; 2003:95–127.

*Evans SW, Owens JS, Bunford N. Evidence-based psychosocial treatments for children and adolescents with attention-deficit/hyperactivity disorder. *J Clin Child Adolesc Psychol*. 2014;43(4):527–551.

*Fabiano GA, Schatz NK, Aloe AM, Chacko A, Chronis-Tuscano A. A systematic review of meta-analyses of psychosocial treatment for attention-deficit/hyperactivity disorder. *Clin Child Fam Psychol Rev*. 2015;18(1):77–97.

Flannery-Schroeder EC, Kendall PC. Group and individual cognitive-behavioral treatments for youth with anxiety disorders: a randomized clinical trial. *Cognit Ther Res*. 2000;24:251–278.

Foa EB, Kozak MJ. Emotional processing of fear: exposure to corrective information. *Psychol Bull*. 1986;99:20–35.

*Gordon-Hollingsworth AT, Becker EM, Ginsburg GS, et al. Anxiety disorders in Caucasian and African American children: a comparison of clinical characteristics, treatment process variables, and treatment outcomes. *Child Psychiatry Hum Dev*. 2015;46(5):643–655.

Horowitz JL, Garber J, Ciesla JA, Young JF, Mufson L. Prevention of depressive symptoms in adolescents: a randomized trial of cognitive-behavioral and interpersonal prevention programs. *J Consult Clin Psychol*. 2007;75:693–706.

Kazdin AE. Cognitive-behavior modification. In: Wiener JM, Dulcan MK, eds. *Textbook of Child*

and Adolescent Psychiatry. 3rd ed. Washington, DC: American Psychiatric Publishing; 2004:985–1006.

Kendall PC. Treating anxiety disorders in children: results of a randomized clinical trial. *J Consult Clin Psychol*. 1994;62:100–110.

Kendall PC, Chu B, Gifford A, Hayes C, Nauta M. Breathing life into a manual: Flexibility and creativity with manual-based treatments. *Cogn Behav Pract*. 1998;5:177–198.

Kendall PC, Chu BC. Retrospective self-reports of therapist flexibility in a manual-based treatment for youths with anxiety disorders. *J Clin Child Psychol*. 2000;29(2):209–220.

Kendall PC, Flannery-Schroeder E, Panichelli-Mindel SM, Southam-Gerow M, Henin A, Warman M. Therapy for youths with anxiety disorders: a second randomized clinical trial. *J Consult Clin Psychol*. 1997;65:366–380.

Kendall PC, Hudson JL, Gosch E, Flannery-Schroeder E, Suveg C. Cognitive-behavioral therapy for anxiety disordered youth: A randomized clinical trial evaluating child and family modalities. *J Consult Clin Psychol*. 2008;76:282–297.

King NJ, Tonge BJ, Heyne D, et al. Cognitive-behavioral treatment of school-refusing children: a controlled evaluation. *J Am Acad Child Adolesc Psychiatry*. 1998;37:395–403.

Last CG, Hansen C, Franco N. Cognitive-behavioral treatment of school phobia. *J Am Acad Child Adolesc Psychiatry*. 1998;37:404–411.

Lewinsohn PM, Clarke GN. Psychosocial treatments for adolescent depression. *Clin Psychol Rev*. 1999;19:329–342.

March JS, Franklin ME, Leonard H, et al. Tics moderate treatment outcome with sertraline but not cognitive-behavior therapy in pediatric obsessive-compulsive disorder. *Biol Psychiatry*. 2007;61(3):344–347.

March J, Silva S, Petrycki S, et al; Treatment for Adolescents With Depression Study (TADS) Team. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents With Depression Study (TADS) randomized controlled trial. *JAMA*. 2004;292:807–820.

MTA Cooperative Group. A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. Multimodal Treatment Study of Children with ADHD. *Arch Gen Psychiatry*. 1999;56:1073–1086.

MTA Cooperative Group. National Institute of Mental Health Multimodal Treatment Study of ADHD follow-up: 24-month outcomes of treatment strategies for attention-deficit/hyperactivity disorder. *Pediatrics*. 2004;113:754–761.

Mufson L, Dorta KP, Wickramaratne P, Nomura Y, Olfson M, Weissman MM. A randomized effectiveness trial of interpersonal psychotherapy for depressed adolescents. *Arch Gen Psychiatry*. 2004;61:577–584.

Mufson L, Moreau D, Weissman MM, Klerman GL. *Interpersonal Therapy for Depressed Adolescents*. New York: Guildford Press; 1993.

Ollendick TH, Halldorsdottir T, Fraire MG, et al. Specific phobias in youth: a randomized controlled trial comparing one-session treatment to a parent-augmented one-session treatment. *Behav Ther*. 2015;46(2):141–155.

Ollendick TH, Ost LG, Reuterskiöld L, et al. One-session treatment of specific phobias in youth: a randomized clinical trial in the United States and Sweden. *J Consult Clin Psychol*.

2009;77(3):504–516.

Öst LG, Svensson L, Hellström K, Lindwall R. One-session treatment of specific phobias in youths: a randomized clinical trial. *J Consult Clin Psychol.* 2001;69:814–824.

Owens EB, Hinshaw SP, Kraemer HC, et al. Which treatment for whom for ADHD? Moderators of treatment response in the MTA. *J Consult Clin Psychol.* 2003;71:540–552.

*Pelham WE Jr, Fabiano GA. Evidence-based psychosocial treatments for attention-deficit/hyperactivity disorder. *J Clin Child Adolesc Psychol.* 2008;37:184–214.

Pelham WE Jr, Fabiano GA, Waxmonsky JG, et al. Treatment sequencing for childhood ADHD: a Multiple-Randomization Study of Adaptive Medication and Behavioral Interventions. *J Clin Child Adolesc Psychol.* 2016;45:396–415.

Pincus DB, May JE, Whitton SW, Mattis SG, Barlow DH. Cognitive-behavioral treatment of panic disorder in adolescence. *J Clin Child Adolesc Psychol.* 2010;39(5):638–649.

Power TJ, Mautone JA, Soffer SL, et al. A family-school intervention for children with ADHD: results of a randomized clinical trial. *J Consult Clin Psychol.* 2012;80(4):611–623.

Reinecke MA, Ryan NE, DuBois DL. Cognitive-behavioral therapy of depression and depressive symptoms during adolescence: a review and meta-analysis. *J Am Acad Child Adolesc Psychiatry*. 1998;37:26–34.

Silverman WK, Kurtines WM, Ginsburg GS, Weems CF, Lumpkin PW, Carmichael DH. Treating anxiety disorders in children with group cognitive-behavioral therapy: a randomized clinical trial. *J Consult Clin Psychol.* 1999;67:995–1003.

Swanson J, Arnold LE, Kraemer H, et al. Evidence, interpretation, and qualification from multiple reports of long-term outcomes in the Multimodal Treatment study of Children with ADHD (MTA): part I: executive summary. *J Atten Disord.* 2008;12(1):4–14.

Walkup JT, Albano AM, Piacentini J, et al. Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *N Engl J Med.* 2008;359(26):2753–2766.

▲ 54.3 Cognitive-Behavioral Psychotherapy for Children and Adolescents

ANNE MARIE ALBANO, PH.D., AND SANDRA S. PIMENTEL, PH.D.

INTRODUCTION AND DEFINITION

The terms *cognitive-behavioral therapy, cognitive therapy,* and *behavior therapy* are often used interchangeably. *Cognitive-behavioral therapy* (CBT) is a form of psychotherapy that combines behavior therapy and cognitive therapy into a broad approach composed of empirically based methods built on the fundamentals of social learning theory and experimental psychology. It is the treatment of choice for various mental health problems due to its efficacy. CBT is widely taught in various mental health training programs and is a mandated component of psychiatric residency training in the United States. CBT is a mandated first-line treatment in the United Kingdom for specific mental health problems such

as anxiety, eating disorders, and depression.

HISTORY

CBT is best viewed as a multifaceted theoretical approach with a strong allegiance to the experimental method and a resultant set of techniques applied within a CBT formulation. The cognitive-behavioral approach is the convergence of distinct movements in the study and science of human behavior, including classical conditioning, operant conditioning, and the cognitive revolution. In the early 1920s, Ivan Pavlov's work demonstrated the processes of classical (otherwise known as respondent) conditioning and established the stimulus–response learning paradigm. Essentially, classical conditioning is a form of learning in which a neutral stimulus is repeatedly paired with a stimulus that elicits a reflexive response. In his classic experiment, Pavlov paired the neutral stimulus of a light with food, a stimulus that naturally produces salivation. Eventually, presentation of the neutral stimulus (light) by itself elicited the conditioned response of salivation. This process is illustrated in Table 54.3–1. This work greatly influenced John B. Watson, who applied Pavlovian classical conditioning principles to understanding the origins of psychological disorders in humans. Known as the "father of behaviorism," Watson promoted the tenet that psychology should be concerned only with directly observable stimuli and responses and rejected any notion of unconscious and dynamic processes (i.e., mentalistic constructs) underlying behavior, including cognitive processes. Early laboratory research with animals demonstrated the acquisition of the fear response through classical conditioning learning trials.

Table 54.3–1.
The Process of Classical Conditioning

Before conditioning	
Food →	Salivation
(Unconditioned stimulus)	(Unconditioned/reflexive response)
During conditioning	
Light + Food →	Salivation
(Neutral stimulus) + (Unconditioned stimulus)	(Unconditioned response)
After conditioning	
Light →	Salivation
(Conditioned stimulus)	(Conditioned response)

Watson and colleagues demonstrated that fear reactions are also conditioned in humans who previously did not evidence such a reaction. In their classic experiment, Watson and Rosalie Rayner conditioned a fear reaction to a white rat in a young child named Albert. Initially, Albert played happily with the rat and showed no evidence of fear or avoidance of

this animal. Next, the experimenters followed the introduction of the rat with a loud noise made in close proximity to the child, causing him to startle. The pairing of the loud noise (aversive stimulus) and the rat (neutral stimulus) in a sequence of trials resulted in Albert developing fear of the rat (conditioned response). Subsequently, not only did Albert become fearful of the presence of the rat, eventually his fear generalized to other white, furry stimuli, such as cotton balls. Inasmuch as this experiment demonstrated the process of Pavlovian learning principles in the etiology of psychological disorders, Mary Cover Jones demonstrated that one could eliminate human fear conditions in children through a similar learning paradigm that would foreshadow later methods of behavioral treatments for anxiety disorders. With Watson as her dissertation mentor, Jones used an approach involving the techniques of modeling and graduated exposure to treat 3-year-old Peter's fear of rabbits. Initially, Peter was exposed to observing other children playing happily with a rabbit (i.e., modeling), and, subsequently, while Peter was eating his favorite food, the rabbit was shown to him during a time when he was in a calm state. Gradually, the rabbit was moved closer to Peter until he was able to pet and play with it without experiencing fear (exposure). Although many will recognize Jones' work as straightforward examples of the basic behavioral procedures of modeling and graduated exposure, her approach to treating Peter was revolutionary at the time. In fact, it was decades before the techniques employed in Peter's treatment were systematically studied and identified as valuable therapeutic procedures.

At the same time that Pavlov and Watson were experimenting with classical conditioning paradigms, Edward Thorndike of Columbia University was beginning pioneering work in understanding the processes of operant (instrumental) conditioning. Thorndike demonstrated that by changing the consequences of behavior, the occurrence of the behavior is either increased or decreased—that is, providing a reward after the occurrence of a certain behavior increases the probability of that behavior occurring again, whereas withholding the reward, over time, results in the decrease of the behavior. These principles are widely used today in therapy and educational settings, such as in parent education and training. A common example is that of a young child who throws a temper tantrum in response to being denied a treat. If the parent provides attention to the child through admonishing, coaxing, or giving in and providing the treat, the tantrum behavior is likely to occur again at the next instance of the child's being initially denied the treat. Note that the reward in this case is parental attention, regardless of whether it is the positive attention of coaxing or giving in or the negative attention of admonishment. In any case, the child gets the parent's attention. However, if the parent refrains from giving attention and ignores the child's crying and whining, the tantrum will stop and subsequent tantrums are unlikely to occur when the

child is denied or frustrated. The child will learn that tantrum behavior does not result in positive outcomes. This type of learning is central to parent management training (PMT), an empirically-based treatment approach, and similar parent training programs. Change in child behavior (decrease in tantrums) is facilitated and enhanced through the simultaneous pairing of positive child behavior with parental attention, so parents are also taught to give attention and praise when their child behaves in acceptable and desired ways (e.g., "I like the way you are playing nicely with your sister"). The procedures of systematically ignoring undesirable behavior while systematically reinforcing desired behavior are known together as *differential reinforcement of other behavior*. Throughout the 1950s and 1960s the principles of operant conditioning were greatly expanded and researched by Harvard psychologist B. F. Skinner and his students; eventually, this work moved from the laboratory analogue studies of learning processes with animals to applied evaluations of the utility of operant procedures in the treatment of adult psychiatric inpatients. For example, Ted Ayllon and Nate Azrin introduced the first comprehensive token economy program at Anna State Hospital in Illinois. Despite initial resistance and skepticism from the hospital staff, the token economy resulted in significant changes in targeted behavior of inpatient adults with severe mental illnesses and disabilities. Token economy systems flourished in psychiatric settings and training schools and were adapted for widespread use in programs serving youth, such as hospitals, residential facilities, and schools.

The emergence of behavior therapy as a distinct field of study began in the 1950s, in part due to advances in the understanding of both classical and operant learning processes, and also because of a growing discontent with traditional psychoanalysis. A major goal of psychoanalytic treatment is for the patient to develop insight into his or her problems. *Insight* refers to the emotional and intellectual understanding and acceptance of one's problems and is a crucial step in achieving therapeutic benefit, and gaining insight is associated with uncovering the unconscious conflicts and motivations underlying the problematic behavior. The assumption of unconscious conflicts is a major premise of psychoanalysis that is rejected by the behavioral approach. The focus on achieving insight, along with the often inordinate length of time spent in psychoanalytic therapy, contributed to discontent with the approach. South African psychiatrist Joseph Wolpe, dissatisfied with his patients' progress in psychoanalysis, was among the founders of contemporary behavior therapy and moved its applications from experimental settings and institutions to more mainstream psychotherapy contexts. In the 1950s, Wolpe developed procedures of systematic desensitization and assertive training, directly stemming from Cover Jones' work with children in the 1920s. In each procedure, phobic and anxious patients are taught to replace fear reactions

with deep relaxation and proactive communication behaviors respectively. These procedures provided a more rapid therapeutic response than traditional analysis and proved to be more cost-and time-effective to patients. Very quickly, Wolpe's procedures were extended to individual and group therapy contexts and adapted for use in children and adults. Worldwide, behavior therapy was developed, advanced, and promoted by pioneers such as Wolpe, Cyril Franks, Hans Eysenck, Stanley Rachman, and Arnold Lazarus. As a field, behavior therapy took shape through the systematic study of behavioral principles and the testing of its methods in single-case and large research trials. In 1966, Franks became the first president of the newly formed Association for Advancement of Behavior Therapy (now the Association for Behavioral and Cognitive Therapies), the flagship professional organization whose missions remains that of promoting the development and dissemination of behavior and CBT.

Radical behaviorists, steeped in the tradition of Watson, eschewed the study of cognition, in that thought processes were viewed as mentalistic constructs not readily observable or verifiable by others. From a purely behavioral perspective, directly addressing inappropriate or abnormal behavior will change any upsetting or inappropriate thoughts and feelings associated with the behavior, and, hence, direct access to thoughts is not necessary. During the 1960s however, the cognitive revolution began with the work of Albert Bandura, Aaron T. Beck, and Albert Ellis. The cognitive therapist takes a different view, espousing that, if one changes dysfunctional thoughts, change in functional behaviors will follow. The cognitive therapist assumes that thoughts are made observable through methods such as direct interviewing, paper-and-pencil self-report scales, journal or diary entries, and listing of thoughts. At Stanford University, Bandura integrated the principles of classical and operant conditioning with observational learning to formulate the tenets of social learning theory. Through observational learning, the process of observing the actions and reactions of others can effect a change in the behavior of an individual. In essence, observing the actions of others and their resulting consequences, without the individual needing direct experience with a situation, can shape the emotional and behavior repertoire of an individual. Of course, risk and protective variables interact with observational learning to shape behavior such that the temperamentally inhibited child may learn to respond with anxiety and withdrawal when watching the distress of a peer who is teased by the class bully, whereas the child with a more rambunctious disposition may focus on the behavior of the bully and learn to respond in similarly aggressive ways. This theory places critical importance on the role of cognitions in the development and maintenance of behavior, both normal and abnormal, in that cognitions involve thoughts, images, beliefs, expectations, memories, and attributions, among other simple and complex cognitive processes. In the bully example, the

inhibited child may think, “oh no, that can happen to me and embarrass me.” This type of thinking leads to anxious withdrawal and fear reactions. Conversely, the more extroverted child may think, “It is fun making this kid cry and I want to do it next,” leading to imitation of the aggressive behavior. The response of the environment to either child’s response then becomes involved in the complexities of this learning paradigm to further influence the observational learning by either rewarding/accepting or punishing/extinguishing the particular response. The process of observational learning is the basis for concern expressed by parents and other concerned parties over violence depicted in television, movies, and video games and one of the reasons for parental ratings found in the media and music industry.

The cognitive revolution continued with the work of psychiatrist Aaron T. Beck at the University of Pennsylvania, who observed within the course of his psychoanalytic treatment of patients that certain problems were evident in the thoughts of depressed individuals and that traditional psychoanalysis did not allow for the direct confrontation of these thoughts. Beck’s study of depressed patients revealed a certain negative triad of thoughts related to pessimistic views of the self, the world, and the future. Through the Socratic method of questioning the veracity of these thoughts, Beck was able to assist patients with gathering evidence to refute the thought and engage emotional and behavioral resources to counter the depression and increase healthy functioning. Cognitive therapy emerged from this work and became a mainstay in treatment of a wide range of emotional disorders, including depression, anxiety, personality disorders, and, more recently, psychoses. Concurrent with the development of Beck’s approach, psychologist Albert Ellis took note of the irrational beliefs guiding his patients and undermining their ability to proactively manage activities and stressors in everyday life. Ellis developed rational emotive therapy, which seeks to challenge certain core irrational beliefs such as “I must be loved by everyone” and “I must be perfect in my performance/relationships.” According to Ellis, irrational beliefs are conditioned cognitive responses that serve to render an individual vulnerable to disappointment and failure and engender a sense of helplessness and eventual psychopathology. Both Beck’s and Ellis’ therapies integrate behavior principles and processes, such as behavioral exposure and behavior activation, to activate and challenge negative, irrational thinking. From these roots, the cognitive approach, integrated with behavior therapy in forming contemporary CBT, was expanded and adapted to meet the needs of children, adolescents, adults, and families to address a wide range of psychological and behavior problems.

THEORETICAL ISSUES

The value of CBT as an evidence-based and front-line treatment, is ever

expanding and evolving. Since its inception and integration of the behavioral and cognitive movements, treatment packages such as self-instructional training, problem-solving therapy, harm reduction, behavioral family therapy, and behavioral couples therapy have been advanced. Problem specific treatments have emerged and include: the Coping Cat for childhood anxiety disorders, the Adolescent Coping with Depression Course, CBT of Childhood OCD: It's Only a False Alarm, The Incredible Years and Parent–Child Interaction Therapy for disruptive behavior in young children, Multisystemic Therapy for serious child and adolescent disruptive behavior and substance abuse, and the Maudsley Approach for adolescent eating disorders. All CBT protocols are based on the principles of social learning theory and several basic theoretical assumptions of CBT (Table 54.3–2). First, the cognitive-behavioral therapist does not assume that an underlying, unconscious process or drive is fueling the behavior of any individual. Rather, the cognitive-behavioral therapist assumes that all behavior, whether adaptive and pathological, is subject to conditioning principles, although there may be limits on learning due to constitutional constraints (e.g., intellectual capacity, physical limitations) or contextual constraints (e.g., dysfunctional home, depressed community settings). As such, the therapist works with the patient (and with the parents in child cases) to examine samples of behavior across contexts to uncover the antecedents and consequences of behavior. The *ABCs of therapy* refer to the antecedents–behavior–consequence chains that serve to establish and perpetuate behavioral patterns. Rather than assume that an individual is predisposed to behaving in a certain way due to traits such as aggressiveness, introversion, or obsessiveness, the cognitive-behavioral therapist takes a more interactionist approach and assumes that the individual is interacting with the environment in developing response patterns to discrete and specific stimuli that may bear resemblance to prior similar stimuli or experiences, thus shaping apparently trait-like responses. Through the process of CBT, triggers to behavior and resultant consequences are identified and dealt with within the treatment plan. The cognitive-behavioral therapist, based on empirical findings, applies procedures steeped in the principles of learning to change or modify the patient's behavior. Through a process of collaborative empiricism, the therapist and patient (including parents or caregivers in the case of dependent youth) specify clearly defined goals for treatment and establish a time frame for achieving subgoals and overall goals. Therapist and patient will then engage in a continual process of evaluating whether goals are achieved, and if not, examining impediments to change including resistance throughout the course of therapy. CBT is applied flexibly to meet each individual patient's needs. That is, techniques are adapted to be developmentally appropriate for a given age or cognitive level and otherwise adapted to meet the needs of an individual patient.

Table 54.3–2.
Assumptions of Cognitive-Behavioral Therapy

Cognitive-behavioral therapy, relative to traditional psychotherapies, concentrates on behavior (this includes verbal behavior and covert behavior, such as thoughts, beliefs, memories, attributions, and images), rather than on some presumed unconscious underlying cause.
Cognitive-behavioral therapy assumes that maladaptive behaviors are acquired through learning in the same way that any behavior is learned.
Psychological principles, especially learning principles, are highly effective in changing and modifying maladaptive behavior.
Cognitive-behavioral therapy involves setting specific, clearly defined, and observable treatment goals.
Cognitive-behavioral therapy rejects classic trait theory—that is, the notion that an individual has an inherent predisposition to similar behavior in a wide range of situations.
The cognitive-behavioral therapist adapts the method of treatment to a patient's problem.
Cognitive-behavioral therapy is focused on the here and now.
Cognitive-behavioral therapists place great value on the empirical testing of their techniques and protocols.

Contrary to popular belief, CBT does not ignore historical information. The patient's history is essential for understanding the original setting events and learned reactions comprising the etiology of the problem. However, the work of CBT is focused on the present maintaining factors that prevent healthy functioning and allow the persistence of the patient's problem. Thus, CBT is focused on the here and now, with emphasis on understanding the current reinforcement paradigm, allowing persistence of the problem and teaching new skills to promote adaptive behavior. Finally, scientific rigor, whether on a large-group scale or within the single-case approach, is highly valued by the cognitive-behavioral therapist. It is expected that the cognitive-behavioral therapist will be current with the empirical literature and apply the scientific method to his or her practice of therapy. Thus, within the clinical practice, each patient is regarded as an individual case study in which the therapist and patient again collaborate to test hypotheses of behavioral change and refine their techniques or approach to maximize the potential for therapeutic benefit.

TECHNIQUES

Techniques are not the essence of CBT, nor are they sufficient for the resolution of any clinical problem. A thorough understanding of the empirical bases and theoretical principles of CBT, along with appropriate training and experience, is a prerequisite to the development of competency to formulate a sound treatment plan and deliver CBT to persons in need. In addition, CBT is a very action-oriented approach focused on the here and now. It requires a directive rather than passive stance. Therapists must be ready to teach, model, guide, and reinforce the patient in working toward their goals. Cognitive-behavioral therapists are very active in treatment and often conduct sessions in vivo—that is, in the situations in which the child or adolescent may be experiencing distress.

The cognitive-behavioral therapist should expect to make a trip to the school, home, or other settings in which the child has difficulty, as appropriate, with careful planning to maintain confidentiality while promoting therapeutic benefit. Moreover, certain disorders, such as the phobias and obsessive-compulsive disorder (OCD), require the therapist to model the handling or touching of phobic or feared objects such as dogs, insects, and potentially "contaminated" objects such as doorknobs or toilet seats. Cognitive-behavioral therapists take care to devise these in vivo experiences to be maximally beneficial to the child while minimizing risk.

Table 54.3–3.
Comparison of Adolescent and Parent Reports in the Assessment of Social Anxiety Disorder

Interview Question[a]	Adolescent Report	Parent Report
Do you worry about what other people think of you or that you may do something to embarrass or humiliate yourself in front of others and then they'll think badly of you? Do you think you worry about this social anxiety more than other kids your age?	Yes, I worry about this sometimes. But I don't think it's any more than anyone else. I don't think so.	He is always worried about what he says and often says he does not know what to say. He fights with us about this. We try to push him to go out and talk to people, but he's too shy. He won't speak up and he is anxious that others will think he's stupid or awkward. This is different than our other kids who are quite social.
Do you think you have the same number, fewer, or more friends than other kids your age?	The same number as other kids.	Fewer. Actually, we don't think he has any friends that we're aware of and hasn't had since elementary school.
How often do you see your friends outside of school? What kinds of things do you like to do with friends?	Well, not much, maybe once a month. There's not much to do and nothing's really going on. And I have a lot of homework.	He doesn't see anyone outside of school. He just stays in his rooms and plays videogames or is on the computer. He doesn't go to any parties or to school events. We feel like we are constantly nagging him to go.
Do you call or text your friends to do things, or do you wait for them to call or text you?	I text them back sometimes.	He doesn't seem to call or text anyone. We see our other kids always texting or on social media. Not him. One classmate called to get homework and even that was hard for him.

[a]Questions are modified appropriately when addressed to parent.

Behavioral Assessment

A classic phrase found in most reviews of CBT is that "behavioral assessment is the cornerstone of (cognitive) behavior therapy." In essence,

CBT is dependent on a carefully constructed and thorough functional analysis of the patient's presenting problem. The functional analysis involves the use of appropriate assessment methodologies to uncover the relationships and conditions under which the patient developed and has maintained his or her problematic behavior (the ABCs defined earlier). Assessment in CBT involves multiple methods, multiple informants, and multiple time periods. Assessment involves multiple methods to capture the most useful and valid information concerning the patient and his or her functioning across a variety of contexts. Thus, diagnostic interviews, self-report questionnaires, diary forms, and behavioral observation techniques can all be used, given time and cost restraints. The use of multiple methods and multiple informants allows for data to be captured without relying on any one form of assessment or individual that may be subject to biases or limited in scope. For example, although children generally are able to report internal states with accuracy, younger children tend to report what is within their immediate experience, and, as such, even recent events may be distorted or inaccurately represented. In addition, the nature of some disorders may interact with youth report on any particular method of assessment. Adolescents with social phobia in general have difficulty admitting to this form of anxiety and tend to minimize or deny symptoms. In contrast, parents of youth with social phobia are able to identify and report the behavioral avoidance associated with the syndrome. Consider, for example the comparison of responses to questions from a diagnostic interview by a 15-year-old boy with social phobia and that of his parents (Table 54-3–3). Although the youth denies a lack of friendships and minimizes his social anxiety, the parents' report establishes his isolation due to avoidance of social interaction and activities with peers. In this case, the youth may be defensive or unable to admit to the anxiety due to self-presentational demands; however, the parents are readily able to identify the specific behavioral contexts (after-school activities, telephone use) in which behavioral avoidance is apparent. This information, taken within the broader context of a diagnostic interview, will be compiled to establish a diagnosis and specify goals for treatment (e.g., engage in social activity outside of home and school; engage with friends on the phone or texting).

Behavioral assessment often involves obtaining samples of behavior through behavioral approach tasks. The behavioral approach task can be individually tailored for each patient or standardized for a particular disorder. Standardized behavioral tests for youth with social phobia involve giving a speech in front of a camera or a small, live audience. Youth with separation anxiety disorder may be asked to stay in the clinician's office while the parent is sent on an errand. The behavioral approach task allows the clinician to assess the behavioral limits of the patient. If the child refuses to attempt the task, important information about his or her initial reactivity and ability to manage anxiety is accessed. No child or adolescent

is ever forced to attempt a behavioral approach task; rather, he or she is allowed to avoid or quit the task at any point. The behavioral approach task also provides information on how long the patient can stay in the situation and what methods make the situation better or worse for the patient. For example, children who fear the dark may be better able to sit in a darkened room with the use of a flashlight or with another person with them. These safety signals provide information about what coping strategies the child is using and whether the strategy needs to be refined or removed through treatment. The behavioral approach task also allows for access of specific cognitive, physiological, and behavioral content by having the patient list his or her top six thoughts that occurred during the task, taking measures such as heart rate or respiration, and recording the length of time engaged in the task or how close the patient can be to a feared stimulus. Overall, the behavioral approach task provides both qualitative and quantitative information to establish baseline functioning and evaluate the progress of treatment. Sample behavioral approach task situations for anxiety disorders are given in Table 54.3–4.

Table 54.3–4.
Sample Behavioral Approach Tasks for Child and Adolescent Anxiety Disorders and OCD

Social anxiety disorder: Reading aloud in front of the clinician and office staff, initiating a conversation with an unfamiliar person, asking for assistance from a store clerk, asking for directions to the restroom.

Selective mutism: Whispering to the parent in front of the therapist, speaking at a low volume/whispering to the therapist, asking a question in a normal voice to the therapist, speaking on the phone with a relative, asking the teacher a question, calling and inviting a friend to get together.

Separation anxiety disorder: Staying alone in the therapy room for 10 minutes, staying alone in a darkened room for 10 minutes, parents leaving the child at the clinician's office to run an errand.

Generalized anxiety disorder: Purposefully making mistakes on a homework assignment, being late for an appointment, reading an upsetting story online or watching an upsetting news clip about crime or world events.

Animal/insect phobia: Being in the same room with a caged animal, staying in the room with the animal on a leash, touching the animal, allowing the animal to roam free in the room, allowing the animal to interact freely with self.

Injections/doctor phobia: Entering an examination room, looking at the medical instruments, holding a syringe, allowing self to be examined by a physician, allowing needle to rest against skin.

Darkness phobia: Sitting in a dimly lit room with another person, sitting in the dimly room alone for 10 minutes, sitting in the room for an unknown period of time, sitting in complete darkness while holding a flashlight, sitting in complete darkness.

Obsessive-compulsive disorder: Touching "contaminated" objects (e.g., doorknobs, shoe soles), not washing hands, shaking hands with various office staff, saying "forbidden" words or thoughts aloud (e.g., "God is dead"), using "unlucky" or "bad" numbers on purpose.

The use of multiple informants involves obtaining information from the child, parents, teachers, and, potentially, peers, other relatives, and former caregivers or therapists. The use of multiple informants addresses issues of

responder bias while accessing the opinion of individuals who have experience with the child over a sustained period and in multiple contexts. Information gained from parents and teachers, for example, may reveal the occurrence of behaviors that are not likely to present in the therapeutic environment. However, reports from others may be limited by an inability to appreciate the antecedents or consequences of behavior and the inability to understand the child's internal dialogue or ideas. Finally, assessment occurs throughout the course of CBT. The initial assessment is used to establish a baseline diagnostic picture and establish the limits of functioning and identify behavioral excesses, strengths, and deficits. The goals for treatment are defined through data accessed during the assessment, and specific treatment protocols or techniques are chosen based on the assessment results. Assessment continues throughout the course of treatment to track progress, continue to gather data in specific domains (e.g., daily ratings of anxiety or depression, number of social contacts, diary forms to elicit though patterns and beliefs), and inform the therapist and patient about treatment response.

Psychoeducation

A core component of CBT for just about any disorder or problem is psychoeducation for the child or adolescent and parents. Psychoeducation typically begins in the first treatment session and is carried on throughout therapy. Developmentally appropriate information is delivered to the youth and parents concerning the nature of the disorder in question (e.g., addressing the proposed etiological pathways and maintenance of depression) and the process and procedures of CBT. It is important for patients and parents to learn that certain emotions may be "charged up" or "cooled down," given changes in the environment (being teased by a bully, getting an A+ on a difficult examination, canceling a picnic due to rain), physical state (e.g., hunger, fatigue), time of day or year (school holidays, final exams) and other internal and external factors. Individuals may be born with a predisposition or vulnerability to experience certain emotions more strongly and readily than others, and over time, by interacting with the wider environment, a person can develop a style of thinking or psychological bias to interpret things in a more emotional manner (the "negative filter"). Although these biases exist, children and parents are further taught that individuals can learn strategies to manage their emotional reactions to a large degree, either with or without the assistance of medications. Emotions are broken down into three components—the physical (e.g., tension, irritability, physical symptoms of anxiety and depression), cognitive (e.g., automatic thoughts, beliefs, cognitive distortions), and behavioral (e.g., actions, aggression, withdrawal, avoidance). The relevant cognitive-behavioral techniques and principles tied to each component are offered as the ways in which the child will learn

to recognize the triggers of his or her emotional reactions and use appropriate coping strategies effectively to manage the emotion and situation. Parents are also given specific information to address any gaps in their understanding and expectations for normative child development.

Psychoeducation also involves the introduction and training of the child and parents in self-monitoring. Structured diary forms are often used to guide the monitoring of triggers to emotional states (both positive and negative) and rating of the intensity of emotion (e.g., from 0 to 10). Monitoring sets the stage for later tracking of cognitive and behavioral responses that become the focus of other treatment components. A sample monitoring form for an adolescent with depression is shown in Tables 54.3–5 and 54.3–6.

Somatic Management Techniques

The vegetative and physiological symptoms accompanying anxiety and depression are addressed through structured relaxation and exercise assignments. Deep diaphragmatic breathing and progressive muscle relaxation training form the mainstay of techniques used to dampen anxious or aggressive arousal and increase a child or adolescent's ability to control his or her unhelpful or inappropriate responding. Mindfulness exercises may also be introduced to teach children and adolescents strategies for nonjudgmental awareness of their bodies. Prescribed exercise regimens involving dancing, walking, bicycling, playing team sports, and engaging in similar physical activities are geared toward behavioral activation to reduce fatigue and increasing stamina and energy in the depressed adolescent. Nutritional counseling and sleep hygiene methods are also used as indicated to address the concomitant problems of weight control and sleep disturbance accompanying the internalizing disorders.

Table 54.3–5.
Daily Diary for an Adolescent with Major Depressive Disorder during the Early Phases of Treatment

Column A What was happening? (trigger)	Column B Mood Rating (0 = best, 10 = most depressed ever)	Column C What was I thinking?	Column D What did I do?
Parents were arguing.	8	This is miserable. I hate it here at home. Why don't they stop?	Went to my room and cried. Didn't eat dinner and stayed there all night.
Had an essay due in English class.	7	I can't think of anything to write about. There's nothing interesting to write. I can't do this.	Listened to music and lay in bed. Tried to write but couldn't. Watched some TV and then went to sleep.

Friend called to go to the movies.	5	I don't feel like going anywhere. I'm no fun and I'm too tired.	Told my sister to make up an excuse for me. Didn't talk to my friend.

Cognitive Interventions

The goals of cognitive interventions are to build on psychoeducation and provide corrective information concerning the presenting problem while systematically uncovering and refuting the automatic thoughts and schema that drive the youth's emotional distress. Automatic thoughts are discrete, rapid, and often context-specific thoughts that vulnerable individuals treat as fact about themselves, the world, and their future. All individuals have automatic thoughts each day. Typical automatic thoughts for youth with anxiety and depression are listed in Table 54.3–7. Cognitive restructuring involves identifying these thoughts through the use of diary forms, employment of Socratic questioning in the therapy room, and then developing rational responses based on actual evidence for the thoughts to assist with their emergence in everyday life. In effect, patients are taught to treat the thoughts as hypotheses, not facts. The therapist then guides the patient in planning home-based behavioral experiments to test whether the thoughts are true through the process of gathering evidence to either accept or refute them. Dispute handles, also shown in Table 54.3–7, are questions used to examine the evidence for the thought and guide the child or adolescent toward considering alternatives to his or her negative biases.

Depending on the age and cognitive developmental level of the child, cognitive interventions can be simplified or rendered more complex. Young children and those who are more concrete in thought processes can be presented with a simple cue to "stop and think" when experiencing strong emotions such as fear or disappointment. A picture of a stop sign or favorite cartoon character holding up a stop sign may suffice to teach the child to resist the negative emotion and engage in proactive coping. Emotive imagery is a technique by which the child's favorite superhero or character provides "guidance" in managing the difficult situation. For example, a 7-year-old boy with separation anxiety disorder might be taught to imagine how Spider-Man's parent was late in picking him up from school. The therapist would coach the child to imagine Spider-Man handling each situation and then have the child imagine himself in each situation, with Spider-Man "standing by" in the image and assisting the child with coping strategies, such as deep breathing and thinking realistic, helpful thoughts. Older children are often provided acronyms to remember a series of steps in cognitive restructuring: Feeling frightened? Expecting bad things to happen? Actions and attitudes to help me cope with this situation. Results on how the situation turned out and how I handled it and reinforce myself for trying. Homework in cognitive restructuring involves

the use of additional columns in the diary for (Tables 54.3–5 and 54.3–6). These forms are reviewed in session, whereby the therapist and patient evaluate each step of the restructuring process and work through plans to overcome challenges and reinforce success experiences.

Table 54.3–6.
Daily Diary with Cognitive Restructuring Component for an Adolescent with Depression

Column E Which dispute handles can I use?	Column F What is my Rational Response?	Column G New mood rating based on evidence and realistic thoughts	Column H What is my action plan?
How often do they argue? Is this the worst thing that could happen? What are they usually like together?	Mom and dad are loving to each other, and all couples disagree and argue at times.	3	Focus on what they are usually like together. Sit with the family for dinner and, if the time is right, let them know I was upset hearing their argument.
How many times have I written an essay? How likely is it that I wont be able to think of anything if I try? Do I have to do this all at once?	I've prepared for class and read all the material, so I do have ideas to use in this essay. I still have 4 days to get this finished and that gives me plenty of time to work on this.	3	I will review my notes from class and make an outline for the essay. After a short break for a snack, I will begin to prepare the content of the essay. If I get this done tonight, I can rest and finish tomorrow.
How do I know it won't be fun to go to the movies if I don't try? Did I ever enjoy going to the movies? Is it possible that I'm tired because I haven't gotten out?	My friend called me because we have had fun together, so I will likely also have fun at the movies.	2	Call my friend back, apologize for not being available, and tell him I'd like to get together for a movie or to do something else that's fun.

Table 54.3–7.
Typical Cognitive Distortions and Dispute Handles for Youth with Anxiety and Depression

Cognitive Distortions
I'm not as smart as other kids.
I don't know what to say to others.
I'm not as attractive as others.

No one wants to hang out with me.
They think I'm a geek because I was nervous giving that oral report.
What if I do something stupid?
Nothing will work out for me.
This is the worst situation anyone could ever be in.
I know I'm going to fail.
If I don't ace this test there's no point in trying anymore.
I should have done better.

Dispute Handles

Am I 100% sure that ____ will happen?
Do I have evidence that ____ will happen?
How many times has ____ happened before?
Is ____ really so important that my whole future depends on its outcome?
What is the likelihood that ____?
Does ____'s opinion reflect that of everyone else?
Do I have a crystal ball?
Am I responsible for this entire conversation?
What is the worst that could happen?
What can I do to cope/handle the situation?

Exposure/Behavior Change Techniques

In the anxiety disorders, behavioral exposures are used to provide systematic success experiences while managing difficult situations and refining cognitive restructuring skills. Anxiety is partly maintained by avoidance of or escape from challenging situations, and thus the child does not gain experience to refute anxious thoughts and learn that the feared negative outcomes are unlikely to occur or can be managed. An child with separation anxiety disorder tries to avoid, at all costs, the experience of being away from loved ones for fear that something terrible will happen to result in permanent separation. By avoiding separation situations, the child maintains a focus on the catastrophic fear and does not learn that the feared negative outcome is highly unlikely. Moreover, avoidance prevents the child from learning that, when adverse or less-than-desirable or expected events occur, such as a parent returning home late from work, the event is independent of the child's fear and avoidance behavior. Exposure directly addresses the fear and avoidance partly through the process of habituation, by which the anxiety response naturally dissipates during the exposure. Exposures also provide opportunity for inhibitory learning, in which the child can learn that catastrophic feared events did not occur and new beliefs about potential outcomes come to inhibit previous beliefs. The child concurrently gains experience with coping and interacting with the feared stimulus or situation. Exposure is usually conducted in a stepwise manner of introducing increasingly challenging situations. Flooding, which is not recommended for young children or those with more severe anxiety levels, involves having the patient confront a situation that provokes the

highest levels of anxiety and stay in the situation until the habituation occurs. Exposure may be imaginal or in vivo but is typically arranged hierarchically for children, from least to most challenging. Role plays with the cognitive-behavioral therapist may be included in preparation for exposures; however, they do not substitute for exposures. During imaginal exposure, the therapist may construct scripts about anxiety-provoking situations and ask the child to listen to the scene as if it were really happening. As the child's subjective anxiety decreases, scenes of greater intensity are presented in succession. In contrast, in vivo exposure involves having the child actually engage in the feared situation. Again, when the child reports tolerable levels of anxiety, the situation may be enhanced to produce a greater challenge to the child and require a new level of cognitive and behavioral coping. In contrast to situational exposures, interoceptive exposure involves inducing the sensations of panic through certain exercises and activities that cause physiological responding, such as tachypnea, tachycardia, palpitations, diaphoresis, and paresthesias. Again, these exposures are designed to break the fear of internal sensations of anxiety and increase tolerance for normative changes in body functioning. Recent experimental studies indicate that habituation may not be necessary to reduce anxiety, but rather, exposure may assist in the learned tolerance of expected anxiety reactions that would previously have resulted in avoidance. For example, learning to tolerate the discomfort of dental procedures, the arousal accompanying asking someone for a date, and other situations that are inherently challenging to anyone, can result in improved functioning and quality of life, and reduced anxiety overall. In all types of exposures, children are encouraged to use their somatic management skills, as well as cognitive restructuring. Use of safety signals may be allowed for less challenging exposures, but systematic removal may be necessary for increasing exposure potency and testing potential feared outcomes.

Behavioral activation tasks are structured experiences planned for youth with major depression or dysthymia. The reinforcement theory of depression asserts that low rates of contingent reinforcement contribute to depressed mood and anhedonia and set up a cycle by which withdrawal behavior occurs and, thus, the youth is even less likely to engage in or have access to reinforcing events. For example, withdrawal from team sports, procrastination in homework assignments, and loss of interest in self-care may occur over time in depressed adolescents. By withdrawing from the team, social reinforcement and contact with peers comes to halt. Withdrawal also may have transactional effects on the teen as peers may disengage and taper contact in response, thereby further reinforcing potential distorted beliefs about himself or herself. Procrastination leads to a sense of dread and failure in school activity, and the drop in self-care compounds the youth's negative self-image. The result is a systematic

decrease is positive reinforcement and corresponding increase in negative feelings about the self. In CBT, the therapist surveys the youth for activities that may bring social reinforcement, success (self-efficacy or accomplishment), or pleasant feelings (e.g., relaxation). Typically, monitoring of mood occurs concurrently with the monitoring of these activities to provide feedback to the youth that engaging in such activity that the youth is likely to attempt, setting up parameters for engaging in the activity, and self-monitoring and processing the outcome of the event. Planned activities increase in type and number as treatment progresses, with the goal that the youth will assume the responsibility of planning these activities and that they will again become a natural and fixed part of his or her behavioral repertoire.

Parent Involvement in Treatment

Parents and family members are often inadvertently drawn into the cycle of negative affect accompanying the internalizing disorders. Family members may provide reassurance or reinforce escape and avoidance as a way of calming the anxious child or protecting the child from becoming upset and anxious. In depression, parents may take over tasks the adolescent is unable to complete due to anhedonia and vegetative symptoms, such as doing the teenager's homework or other compulsory acts. The intention of the parents is to prevent the child or adolescent from becoming even more anxious or depressed and, thus, to allay the distress experienced by their child. Unfortunately, although well intentioned, these parent behaviors are incorporated into vicious cycles of continued anxiety or withdrawal. In CBT, parents are provided a model of negative affect and education about the proposed etiological mechanisms and maintenance of these states. The therapist and parents collaborate to examine maintaining variables in the child's environment and family system, to develop plans to change parental behaviors when necessary, and to rearrange contingencies in the immediate environment to foster the child's recovery. For example, in selective mutism, usually another child or family member will speak for the child and reinforce the child's mutism. The parents of adolescents with social phobia order for their adolescent in restaurants, speak to teachers on the adolescent's behalf, and have fewer expectations for their adolescent's social life. Parents of depressed youth may fear worsening the child's depression by challenging or encouraging participation in activities and thus prefer to "leave well enough alone." Parents are often taught to be good CBT "coaches" for their child. In essence, parents are taught to ask questions (use the Socratic method) to guide their child in arriving at a coping/action plan rather than providing an answer for the child. Much like a good coach, parents are taught to provide opportunities to practice the skills learned in CBT while refraining from judging the child. Reinforcing and praising the child's effort at coping, no matter how small, is essential

for parents to learn in CBT. Parents are also taught systematic skills to ignore certain child behaviors (e.g., whining, refusal to speak, irritability) and to reinforce with positive attention those behaviors that approximate the goal behaviors. Parents are also trained in the use of contingency management procedures that may involve setting specific goals and contracts, such as "If Jonathan sleeps alone is his bed until 7 a.m., he gets 15 minutes of extra iPad time after school." The goal of contingency management procedures is to rearrange the reinforcement paradigms in the child's environment such that the child receives contingent positive reinforcement for appropriate behavior—in some cases, this may be stickers, small treats, or special privileges—whereas reinforcement is removed or withheld when inappropriate behavior occurs. These procedures also provide structured and, if necessary initially, scripted opportunities for parents to attend to positive behaviors and to effectively communicate this reinforcement. Cognitive-behavioral therapists do not advocate the use of punishment, which has been found to be ineffective in maintaining long-term behavioral change, in addition to being ethically unacceptable.

There are a few guidelines and caveats when planning the level and type of involvement of parents in the child or adolescent's therapy. First, the greater the distress and impairment for the child, the more likely it is that the parents are needed to reinforce the treatment program in the home environment. This is especially indicated in cases involving school refusal behavior, in which the parents must rearrange contingencies in the home and actively work in cooperation with the school and therapist to promote attendance. Second, externalizing disorders or other complex comorbid conditions tend to fare better in CBT with parent involvement. Third, younger children have been found to respond better in acute treatment when parents are actively incorporated into the treatment plan. Fourth, the developmental goals of adolescence involve forming independent self-identity and behavioral skills; thus, parent involvement should take this and other developmental issues into consideration. Finally, in cases in which parental psychopathology or marital or family discord is present, involvement of the parents may hinder rather than help the child's progress in treatment. In these cases, involvement of the unaffected parent or minimal parent involvement may be indicated, with referrals given as appropriate for concurrent martial or family therapy, parent management training, and individual parent therapy.

Relapse Prevention

As treatment draws toward a close, the therapist begins the process of transferring the responsibility for continued progress from the therapist to the child. That is, specific strategies and techniques are used to consolidate gains, reinforce coping skills, generalize gains to a wide range of contexts,

and bolster the child's sense of self-efficacy. In the initial phase of relapse prevention, treatment sessions are moved to a biweekly and then monthly schedule to allow the child time to practice skills on his or her own and decrease reliance on the therapist. By identifying upcoming events or changes in the child's life (e.g., starting school, moving to a new neighborhood, facing developmental challenges calling for increased independence), the child is encouraged to develop a coping plan for each event and to plan his or her self-guided exposures in advance of the event. Hence, if a child is changing to a new school, plans to visit the school and meet the teacher or some peers are encouraged and cognitive restructuring is completed for any negative thoughts or expectations that may occur. Similarly, booster sessions can be planned for anticipated tough times (e.g., start of school, before annual examinations) to reinforce coping skills and assist with transitions. Throughout treatment, therapists can work with parents and children to incorporate technology to optimize coping. For example, deep breathing or mindfulness exercises can be recorded into smartphones or coping reminders can be set for a specific date or time to encourage regular practice or situation-specific coping. In addition, smartphone applications and websites are available that provide age-appropriate CBT support (e.g., Kids Feelings App; Camp Cope-A-Lot). Finally, the child's diaries, workbooks, and journals from treatment become excellent relapse prevention aids for the child to refer to after the end of treatment. Photos taken during exposures or behavioral activation tasks that illustrate the child's accomplishments are pasted into these journals, bookmarked on the child or adolescent's smartphone, or put on posters in the child's room to reinforce gains. Finally, the Coping Cat program for anxious youth advocates the development of "commercials" by which the therapist and child produce a recording that incorporates all of the skills acquired during treatment. The child is the "star" of the recording and is shown as if teaching others how to overcome a problem. Commercials can also be produced in the form of comic strips, short books, poems, and other creative method that inspires the child. Whatever the form the commercial takes, the child can refer to it for relapse prevention purposes and to reinforce gains. A resource for parents' "Do's and Don'ts" can also be generated for the parent component of relapse prevention. Table 54.3–8 presents a summary of the CBT components reviewed here.

CLINICAL ISSUES

Indications

CBT is indicated for the treatment of the range of disorders of childhood and adolescence, including anxiety and mood disorders, attention-deficit/hyperactivity disorder (ADHD), oppositional defiant disorder, conduct and oppositional disorder, habit disorders, OCD, PTSD, substance

abuse, eating and feeding disorders, and parent–child discord, in addition to other conditions and behavioral disturbances. CBT is generally accepted as the first-line treatment for the range of disorders in childhood, with combination treatment of CBT and medication preferred for more intractable or complex cases.

Table 54.3–8.
Summary of the Main Components of Cognitive-Behavioral Therapy

Component	Focus/Goals	Associated Techniques
Psychoeducation	Provide corrective and accurate information about the nature of the disorder or problem in question.	Didactic instruction, self-monitoring (diaries, journals), assigned reading
Somatic management techniques		
Anxiety-related	Target autonomic nervous system arousal and related physiological symptoms.	Breathing retraining (deep diaphragmatic breathing), relaxation training (progressive muscle/cue-controlled/applied relaxation), meditation, exercise
Depression	Plan behavioral activation tasks involving increase in activity to counteract negative, vegetative symptoms.	Exercise and related physical activities
Sleep hygiene	Restructure sleep cycle through scheduling of awake activities and self-soothing activities to promote healthy and restful sleep.	Set tasks and guidelines geared toward decreasing arousal and promoting relaxation prior to bedtime; decrease/extinction of arousal activities such as exercise or use of stimulants (caffeine) in close proximity to bedtime.
Cognitive restructuring	Identify maladaptive (automatic thoughts, schema, irrational beliefs) thoughts, beliefs, and images and challenge thoughts with realistic, coping-focused thinking; in children, uncover "unhelpful self-talk" and replace	Monitor thought processes through the use of diaries or questionnaires to reveal thinking styles, identification and labeling of automatic thoughts; teach rational disputation of automatic thoughts using realistic evidence and problem-solving steps.

	with coping-focused self-talk ("coaching").	
Exposure/behavior change techniques	Systematic exposure to avoided situation(s) to provide experience with using cognitive restructuring and other deficient skills (e.g., social skills, problem-solving skills, conflict resolution, affect regulation).	Imaginal exposure, in vivo exposure, gradual or massed (more intense) exposure to anxiety-provoking situations; in depression, use of behavioral activation tasks or behavioral experiments to counter negative and self-defeating behaviors and gather evidence to refute depressive ideation.
Reinforcement techniques	The systematic application or withdrawal of stimuli to either increase desired behavior or decrease undesirable behavior.	Patient earns reinforcement through behavioral contracts, level programs, or parent-administered contingency management programs; reinforcers can include tangible items (stickers, money, food), social reinforcers (praise, positive attention, elevation in status), or desired activities (free time, movies, extended curfews); time out from positive reinforcement is the removal of the individual from a reinforcing situation for a set period; punishment, which is the application of a noxious or undesired stimulus, does not result in long-term behavioral change and is not recommended for use by child behavior therapists.
Positive reinforcement	Provide a desirable stimulus (praise, attention, tangible reward) after a desired behavior to increase the probability of the behavior occurring again.	
Negative reinforcement	Remove an undesirable stimulus after the production of a desired behavior to increase the probability of the behavior occurring again.	
Punishment	Apply a noxious stimulus after an undesirable behavior to decrease the probability of it occurring again.	
Relapse prevention	Focus on consolidating cognitive-	Fading of sessions (from weekly to biweekly), role reversal (patient acts as therapist for a session), planned booster sessions.

	behavioral therapy skills and generalizing treatment gains over time; decrease reliance on therapist and others (e.g., parents, spouse) for managing problems.

Limitations

Child or adolescent-focused CBT may not be the treatment of choice in situations involving extensive parent–child discord or marital problems. In such cases, as noted previously, the preferred course of treatment would involve family or marital therapy. The level of training and experience of the therapist typically the limit the effectiveness of CBT and, despite the widespread dissemination of CBT manuals and workshops, partially trained or inadequately trained therapists are no substitute for specialists in CBT. Hence, families and referral sources must be fully informed as to the degree of expertise in CBT of the therapist and adjust expectations accordingly. Similarly, the motivation and adherence to the treatment plan demonstrated by the patient and family may limit CBT. The therapeutic relationship, contrary to popular belief, is of critical importance to the cognitive-behavioral therapist. Thus, the cognitive-behavioral therapist will work to establish trust, positive regard, openness, and collaboration with the patient and family, no matter how rapid a course of treatment is planned, to facilitate treatment progress.

Complications

Incomplete assessment and failure to uncover fully the antecedents and consequences of behavior may result in complications in therapy. It is not unusual for younger children or those who have successfully avoided phobic stimuli to underestimate their true level of fear. Hence, an exposure may be too complex or overwhelming for a child, requiring the therapist to break the situation down into smaller, more manageable parts. Therapists who allow a child to escape or prematurely end an exposure session before adequate habituation and processing of the exposure can reinforce the anxiety and escape. Thus, it is essential that enough time be allotted for an exposure session and that the parameters of the exposure to be clearly and fully examined with the child or adolescent. CBT does not purport to teach positive thinking, but rather to teach optimistic, realistic, and coping-focused thinking. As such, exposure to less-than-ideal situations and outcomes is a part of treatment.

Contradictions

CBT is a flexible set of techniques and approaches that has been effectively applied to a wide range of child and adolescent disorders. There are no known contraindications for the use of CBT in young people.

SOCIAL ANXIETY DISORDER

Michael was a 16-year-old boy from an intact middle-class family enrolled in the 11th grade of a public school. His parents sought evaluation and treatment for him for a long-standing history of "shyness" and anxiety in social situations. Evaluation revealed social anxiety disorder as the primary disorder. Michael was motivated for treatment and reported that he wished to make friends and feel more comfortable with other people and in social situations. Consequently, Michael was treated with cognitive-behavioral group treatment for adolescents with social phobia—a 16-session course of treatment combining psychoeducation, cognitive restructuring, social skills, behavioral exposure, relapse prevention, and four sessions of parent involvement. As treatment progressed, Michael gradually increased his visibility at school-sponsored social events such that he caught the attention of a female classmate and was asked to attend the prom with her. The therapists designed several prom exposures, by which the various things that could happen at a dance or on a date were present to Michael and the group, including being offered alcohol, dancing, having an argument with the date, and having other girls ask him to dance. As it turned out, Michael's date did, in fact, ignore him and leave him at the dance. Michael was prepared for this less-than-desired outcome by the group experience, and asked other girls to dance and interacted with other peers. He considered the evening a success and subsequently went on to other social events with friends made at the prom. In this case study, the importance of exposing youth to the range of potential outcomes for a given social situation was crucial to Michael's trust in the therapy and therapists and belief in the credibility of treatment. Moreover, he was appropriately prepared through behavioral exposure and practice to handle what could have been an awkward and discouraging situation.

SELECTIVE MUTISM

Anna was a 5-year-old girl from an intact middle-class family referred by her kindergarten teacher due to selective mutism. Anna was described as always having been a shy child who was behaviorally inhibited. Although she would speak and interact with immediate family members, she would withdraw and become quiet around peer and strangers. When she started preschool, she did not speak to her teachers or any school personnel and also began to withdraw from speaking around extended family members at holidays and parties. She also experienced distress about separation from her mother, who would generally respond for her and make requests for her. The school had also allowed her to nod and use nonverbal cues to communicate her preferences, and she was not required to read aloud or speak in front of her peers. She occasionally played and engaged with peers on the playground and was otherwise pleasant and friendly. During evaluation, it was revealed that Anna enjoyed eating Skittles as a treat. Treatment was designed to provide an incentive for speaking through the delivery of a contingent reinforcer of high value, the Skittles. Hence, Skittles became available only in the school and therapist's office and, temporarily, were not available in her home. The therapist enacted a process of graduated shaping of communication behaviors—first nonverbal and then vocal noises—and trained the teachers to do the same. Skittles were kept in full view of Anna at all times during the initial phase of treatment, and when she was "caught" gazing at the package, the therapist or teacher would prompt her for acknowledgement that she wanted the treat. Pointing, looking, and nodding in their direction resulted in receiving two Skittles. While in session the therapist would also leave the room while she and her mother were left to play; if the therapist returned and she continued to vocalize or speak, she would receive additional reinforcement via Skittles. Next, the therapist and parent spent time "making noise" with objects (e.g., rattling keys, pressing telephone buttons) and then with their mouths (e.g., animal sounds, silly sounds), and Anna was prompted to join in and make similar sounds to receive the Skittles. Finally, prompts to ask for the Skittles were enacted, first with an "Ssss,"

then the full word, and finally the full sentence/question (e.g., "Can I have a Skittle?"). This shaping procedure took three extended-length sessions and collaboration, along with telephone consultation with her school teachers. In addition to treat requests, Anna was increasingly reinforced for additional situations on her fear and avoidance hierarchy: making eye contact, saying hello to neighbors, ordering her own food in a restaurant, purchasing her own Skittles from a drug store, singing in the park, etc. The treat was faded over time—that is, delivered on a variable ratio schedule of every three to eight times that she spoke, to promote further speaking and decrease the association with the treat. As a part of their relapse prevention plan, Anna's parents were reminded to allow or prompt Anna to speak for herself in various pre-identified possible situations.

OBSESSIVE-COMPULSIVE DISORDER

Joanne was a 13-year-old girl with a family history of anxiety and mood disorders. Her parents brought her to treatment due to recurrent obsessions involving contamination and germs, and corresponding compulsions by which she had her parents overcook her food, checked her food for contaminants, and washed her hands repeatedly to the point of their becoming raw and bleeding. Evaluation revealed a fear that, unless her parents checked her food for bugs or germs, the meal was contaminated and she would get sick and die. Hence, after cooking it, her parents had to pull apart her food physically and examine it to her satisfaction, which could take upward of 1 hour for each meal. This caused much distress and family discord. In addition, Joanne washed her hands repeatedly after opening a door, using a cell phone, tying her shoe laces or putting on or taking off her shoes, or touching any object that she deemed dirty. First, Joanne, her parents, and the therapist mapped her OCD and constructed a hierarchy based on her compulsions, rituals, and avoided situations (from least upsetting, checking food prepared by her mother, to the most upsetting, touching something that was wet and slimy and then touching her mouth). The process of exposure and response prevention was introduced in combination with a contingency reinforcement plan for Joanne's completion of "brave challenges." Systematically, the therapist had Joanne first engage in imaginal exposure to a scene (e.g., "You watch dad take a bite of his medium rare burger, then you imagine taking a bite of a medium rare burger with the red juice oozing out. You realize it is not super well done and they did not check it for germs.") until her anxiety dropped to minimal levels. This usually occurred after 20 to 30 minutes. Next, in vivo exposures were gradually introduced, by which foods were introduced with "contaminants" in them (e.g., putting pieces of uncooked rice into the burger to mimic "grit"), phasing out the parents' checking of meals as well as her own checking behaviors. As treatment progressed, Joanne learned that her feared consequence of becoming sick was not likely to occur. Washing rituals were addressed by having her touch items with various substances coating them and then touching her face and mouth. At first, she was asked to delay washing her hands over increasing time increments, and then to resist for larger blocks of time. Importantly, it was critical to address some of Joanne's mental ritualizing of thoughts regarding when she would be next able to wash her hands. Treatment occurred over a 14-session program, with some sessions including extended time to allow for exposure and response prevention to occur. Her parents were taught to assist her with these exposures in the home and refrain from engaging in her rituals or accommodating them (e.g., removal of antibacterial sanitizing gels). They were also provided coaching scripts so as not to engage in disputes over meals. Relapse prevention plans were enacted to expand her range of food choices and situational contexts (cafeterias, food stands, restaurants) for exposure. By the end of treatment, Joanne was eating without the need for checking and with minimal to no anxiety, and her handwashing had decreased to within normative range of daily washing.

Goals of Treatment

In the abstract, the overall goals of CBT are to increase an individual's freedom of choice—that is, to free the individual from confines of a disorder which removes the opportunity to make decisions and choices for one's self

out of personal preference and which, instead, determines an individual's choices as a result of fear, depression, anger, or other emotional constraints. The goals of CBT are symptom relief, education about the nature of one's symptoms, vulnerabilities, and patterns of behavior, and the teaching of adaptive, coping-focused, realistic, and optimistic behavior (cognitive and overt) to maximize the person's potential, prepare him or her to meet the challenges of everyday life, and improve his or her quality of life.

Ethical Issues

CBT is a safe, tolerable, and effective treatment for a range of behavioral health problems, with increased efficacy linked in some clinical studies to more seasoned and well-supervised therapists. The main risks of CBT may lie within exposure-based treatments, which are associated with temporary emotional discomfort when confronting feared situations. Contrary to a myth that exposure may result in adverse reactions (e.g., unrelenting anxiety, psychosis, breach of therapeutic alliance), any discomfort and boundary issues in conducting exposure can be skillfully and carefully managed by the experienced therapist. When conducting exposure in community (in vivo) settings, such as schools, parks, and playgrounds, therapists need to be mindful of confidentiality concerns and take precautions to minimize compromising the patient's right to privacy.

Research and Evaluation

CBT has demonstrated efficacy in the treatment of childhood internalizing and externalizing disorders through a number of controlled, randomized trials, including combined trials of medication and CBT. Long-term follow-up of youth treated for anxiety, OCD, or depression suggests maintenance of treatment gains, in addition to further improvement over time. It is incumbent on clinical trial researchers in CBT to test this approach in youth presenting with variations in cognitive and developmental functioning, comorbidity patterns, and various environments adversities to test the effectiveness or therapy programs in mental health service settings. Further research is needed to evaluate treatment algorithms, discontinuation procedures, and pretreatment factors that predict the maintenance or relapse of treatments over the long term. The critical issue faced by today's clinical scientists involves moving treatments from the laboratory to the community. This rests on the collaboration and cooperation of clinical scientists, clinicians, patient advocates, and consumers for critical feedback to flow in multiple directions to adapt and adopt evidence-based treatments in community settings.

References

Barrios BA, Hartmann DP. Fears and anxieties. In: Mash EJ, Terdal LG, eds. *Behavioral Assessment of Childhood Disorders*. 3rd ed. New York: Guilford Press; 1997:230.

Barrios BA, O'Dell SL. Fears and anxieties. In: Mash EJ, Barkley RA, eds. *Treatment of Childhood Disorders*. 2nd ed. New York: Guilford Press; 1998:249.

CATS Consortium. Implementing CBT for traumatized children and adolescents after September 11: lessons learned from the Child and Adolescent Trauma Treatments and Services (CATS) Project. *J Clin Child Adolesc Psychol*. 2007;36(4):581–592.

Cavallo DA, Cooney JL, Duhig AM, et al. Combining cognitive behavioral therapy with contingency management for smoking cessation in adolescent smokers: a preliminary comparison of two different CBT formats. *Am J Addict*. 2007;16(6):468–474.

Comer JC, Puliafico AC, Aschenbrand SG, et al. A pilot feasibility evaluation of the CALM Program for anxiety disorders in early childhood. *J Anxiety Disord*. 2012;26(1):40–49.

David-Ferdon C, Kaslow NJ. Evidence-based psychosocial treatments for child and adolescent depression. *J Clin Child Adolesc Psychol*. 2008;37(1):62–104.

Freeman JB, Chaote-Summers ML, Moore PS, et al. Cognitive behavioral treatment for young children with obsessive compulsive disorder. *Bio Psychiatry*. 2007;61(3):337–343.

Freeman JB, Garca AM, Coyne L, et al. Early childhood OCD: preliminary findings from a family-based cognitive-behavioral approach. *J Am Acad Child Adolesc Psychiatry*. 2008;47(5):593–602.

Goldfried MR, Davison GC. *Clinical Behavior Therapy*. New York: Wiley; 1994.

Hayward C, Varady S, Albano AM, Thienemann M, Henderson L, Schatzberg AF. Cognitive behavioral group therapy for female socially phobic adolescents: results of a pilot study. *J Am Acad Child Adolesc Psychiatry*. 2000;39:721–726.

Hirshfeld-Becker DR, Masek B, Henin A, et al. Cognitive behavioral intervention with young anxious children. *Harv Rev Psychiatry*. 2008;16(2):113–125.

Jablonka O, Sarubbi A, Rapp AM, Albano AM. Cognitive behavior therapy for the anxiety triad. *Child Adolesc Psychiatr Clin N Am*. 2012;21(3):541–553.

Jones MC. A laboratory study of fear: the case of Peter. *Pedagog Semin*. 1924;31:308.

Jones MC. The elimination of children's fears. *J Exp Psychol*. 1924;7:383.

Kazdin AE, Weisz JR. Identifying and developing empirically supported child and adolescent treatments. *J Consult Clin Psych*. 1998;66:19–36.

Kearney CA, Albano AM. *When Children Refuse School: A Therapist's Manual*. San Antonio, TX: Psychological Corporation; 2000.

Kendall PC, ed. *Child and Adolescent Therapy: Cognitive Behavioral Procedures*. 4th ed. New York: Guilford Press; 2011.

Kuehn BM. Cognitive-behavioral therapy shows promise for children with mental illness. *JAMA*. 2007;297(5):453–455.

Mash EJ, Barkley RA, eds. *Child Psychopathology*. 3rd ed. New York: Guilford Press; 2014.

Pavlov IP. *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. London: Oxford University Press; 1927.

Pavlov IP. *Lectures on Conditioned Reflexes*. Vol. 1. London: Lawrence Wishart; 1928.

Pediatric OCD Treatment Study (POTS) Team. Cognitive-behavior therapy, sertraline, and their combination for children and adolescents with obsessive-compulsive disorder: the Pediatric

OCD Treatment Study (POTS) randomized controlled trial. *JAMA*. 2004;292(16):1969–1976.

Scheeringa MS, Salloum A, Arnberger RA, Weems CF, Amaya-Jackson L, Cohen JA. Feasibility and effectiveness of cognitive behavioral therapy for posttraumatic stress disorder in preschool children: two case reports. *J Trauma Stress*. 2007;20(4):631–636.

Shooshtary MH, Panaghi L, Moghadam JA. Outcome of cognitive behavioral therapy in adolescents after natural disaster. *J Adolesc Health*. 2008;42(5):466–472.

Silverman WK, Pina AA, Viswesvaran C. Evidence-based psychosocial treatments for phobic and anxiety disorders in children and adolescents. *J Clin Child Adolesc Psychol*. 2008;37:105–130.

Storch EA, Merlo LJ, Larson MJ, et al. Impact of comorbidity on cognitive-behavioral response in pediatric obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry*. 2008;47(5):583–592.

Szigethy E, Kenney E, Carpenter J, et al. Cognitive-behavioral therapy for adolescents with inflammatory bowel disease and subsyndromal depression. *J Am Acad Child Adolesc Psychiatry*. 2007;46(10):1290–1298.

Tsiros MD, Sinn N, Brennan L, et al. Cognitive behavioral therapy improves diet and body composition in overweight and obese adolescents. *Am J Clin Nutr*. 2008;87(5):1134–1140.

Van den Akker EL, Puiman PJ, Groen M, Timman R, Jongejan MY, Trijsburg W. A cognitive behavioral therapy program for overweight children. *J Pediatr*. 2007;15(3):280–283.

Walkup J, Albano AM, Piacentini J, et al. Cognitive-behavioral therapy, sertraline, or a combination in childhood anxiety. *N Engl J Med*. 2008;359(26):2753–2766.

Watson JB. Psychology as the behaviorist views it. *Psychol Rev*. 1916;20:158.

Watson JB, Rayner R. Conditioned emotional reactions. *J Exp Psychol*. 1920;3:1.

▲ 54.4 Group Psychotherapy

MARGO L. THIENEMANN, M.D., SONAL RANA, M.D., AND NICOLE QUITERIO, M.D.

DEFINITION

Group therapy is the counseling, guidance, or training of children and adolescents involving interactions of three or more members that has the potential for reciprocal influence. The literature describing group therapy for children has, until relatively recently, consisted of anecdotal reports—rich in description but thin in providing scientific evidence to substantiate its assertions. Empirical support for group therapy for children and adolescents is growing. In 1997, Matthew J. Hoag and Gary M. Burlingame published a meta-analysis of the 56 English-language studies of child group therapy that used randomization or matching in their experimental designs. Compared to wait-list or placebo control, group therapies helped to a moderate degree (effect size 0.61, range –0.04 to 2.99).

Group interventions for children and adolescents can be categorized by the goals of the group, therapeutic processes used, duration, and population served, and may be referred to as *psychotherapy groups, psychoeducational groups, support groups,* and *self-help groups*. Delivering therapy via groups may be better than individual therapy for addressing specific illnesses, or may more effectively address treatment issues such as motivation for change or need for social reinforcement of learning.

HISTORY

Therapists have been describing group treatments for children for approximately 100 years. Early in the 20th century, adolescents met for psychodrama groups in Vienna. In the 1930s, *activity group therapy* developed. In these groups, children played among therapists who used nonverbal, noninterpretive interactive strategies to effect ego repair. After World War I and during the Depression, youth needed constructive supervision, although resources and manpower were short, and so delinquent adolescents were treated in groups. After World War II, understaffed military hospitals in the United States and England used groups to deal with psychiatric casualties. The 1960s saw the emergence of encounter groups, the community mental health movement, antiestablishment groups, and an increased tolerance in society for talking openly about one's feelings.

Today, groups are used frequently and ubiquitously. Psychotherapy, support, psychoeducational, and self-help groups may be found in all human service facilities, including schools, hospitals, and clinics and on the

Internet. For example, in an outpatient clinic in Leon, Nicaragua, 80 percent of all outpatient visits are in a group format.

Multiple environmental factors have shaped these historical trends. Pollock and Kymissis, discussing the future of adolescent group therapy, hypothesized that theoretical and clinical developments, training opportunities, scientific research, economics, social change, and public and professional attitudes all influence the type, quality, and availability of group therapy. Psychoanalytic theory informed early groups, which aimed to use group interactions to recapitulate family interactions. Developmental theory, considering the peer group's role in identity formation, pointed to using the group for adolescents' therapy. The development of learning theory and cognitive-behavioral therapy has led to problem-focused, time-limited group interventions.

Societal attitude change and economic factors have increased the appeal of group interventions. According to Moffitt, the current health care climate of high demand and limited resources makes the group therapy format cost effective and fiscally responsible. Due to the developmental tasks of adolescence, which include identity formation and self-concept through peer relationships, peer approval, and belongingness, group modalities are perhaps even more useful for pre-adolescent and adolescent children (paraphrased from Miffitt, 2014). Countering these currents, psychiatric training in psychotherapies, including group therapy, has received relatively less emphasis because of growth of reliance on and teaching about pharmacological treatments.

Other factors serve as potential obstacles to the use of group psychotherapies. Some group interventions developers fear that, because of the difficulty of manualizing and researching psychodynamic group therapies, potentially effective modalities might not make it into clinical practice. Managed care, because of group therapy's potential cost-effectiveness, initially provided incentive for increasing the use of outpatient group therapy. This has not borne out with reimbursement for inpatient or outpatient groups. An increased societal comfort with self-disclosure may increase the likelihood that a patient will participate in a group, but until more research demonstrates an advantage for group therapy, therapists and patients may reject group therapy, considering it diluted and less effective. Finally, Pollock and Kymissis noted that few training programs include adequate training in group psychotherapy. Perhaps, as has been the case for pharmacological, individual, or other modalities, research will demonstrate that specific group therapies for specific problems administered by specifically trained individuals help, and the interventions may gain adequate support for dissemination.

THEORETICAL ISSUES

Therapeutic Power of the Group

Gathering individuals together for therapy may help them more than working with them alone. A number of factors, many described by Irving Yalom, make psychological treatment in groups especially effective. Yalom listed instillation of hope, universality, imparting of information, altruism, corrective recapitulation of the primary family group, development of socializing techniques, imitative behavior, catharsis, the addressing of existential factors, cohesiveness, and interpersonal learning as factors leading to change. In children and adolescents, the use of the group to address developmental issues also may be an advantage.

Hope. In groups, children can meet other children with similar pathology. As they see peers overcome obstacles, they can develop hope that they, too, are capable of helping themselves get better. By observing others addressing problems actively, individuals may expect that their participation in therapy will help them. A parent of an unmotivated anxious child remarked that seeing other children tackling their anxieties gave her hope that her child would be able to do so in the future. Similarly, in multifamily groups, parents often remark that seeing other children improve behavior-wise gave them hope that their families could also begin to function more healthily.

Universality. Children with psychiatric disorders frequently feel they are weird, crazy, or losers. Getting to know worthwhile, admirable individuals with similar issues can lessen the pain and hit to one's self-esteem that can come with having a mental health disorder. Adolescents are in the midst of exploring and consolidating their identities during the time that a number of major psychiatric illnesses begin to manifest themselves. Having the opportunity to interact with others with similar troubles can help teenagers constructively to integrate having a psychiatric illness into their sense of self. In a recent pilot study by Straub et al., reviewing the most positive evaluations of their time-limited cognitive-behavioral group for teen outpatients, the most positive elements identified by participants were being with other adolescents and feeling comfortable in the group. By seeing others in the group, teens can come to see a disorder as only one of many aspects of themselves.

Imparting Information. Children learn together at school, on teams, and in families. In group therapy, children can learn about the course and characteristics of their disorder. In addition to presenting live examples, the group format is suited to teaching via leaders' instruction, educational videos, and structured learning activities. When the child teaches and helps other members, his or her learning is reinforced. Following the Coping Cat anxiety protocol model, members of one adolescent group treatment for obsessive-compulsive disorder (OCD) ended the group by making a

didactic video. The group members delighted in their educational video, which taught about etiology, symptoms, treatments and impact of OCD. They watched it repeatedly and shared it with their families.

Altruism. Supporting others in the group is gratifying. Altruistic actions can boost one's sense of well-being and self-esteem and help to defend against upsetting feelings. Children and adolescents in group therapy often make thoughtful suggestions to other members. One study of adolescents participating in cognitive-behavioral group therapy reported that most participants tried to help their peers to design behavioral interventions in each session.

Corrective Recapitulation of the Primary Family Group. Through examining psychodynamic processes and addressing treatment resistances, the group format can provide a wealth of opportunities for addressing family conflicts. In a group setting, children develop transference relationships with leaders and peers. Providing a safe, constructive atmosphere in which maladaptive interpersonal patterns are addressed and reworked may provide group members a chance to work through dynamic conflicts originating in the family.

Development of Socializing Techniques. Yalom stated that all therapy groups help patients develop social skills. Social deficits, social dysfunction, and isolation are prominent symptoms of children with psychiatric disorders, including those with pervasive developmental disorders, learning disorders, attention-deficit disorder and other externalizing disorders, and social anxiety and depressive disorders. The group allows patients to practice new skills and receive feedback from professionals and peers in real time. In one controlled group therapy study of girls without friends, girls in the treatment group did learn to approach peers and begin conversations. Reinforcement and criticism from peers may be the most meaningful feedback possible for adolescents.

Imitative Behavior. Seeing how others behave greatly influences what children and adolescents do. Children in groups observe other children working on solving problems similar to theirs. Group leaders often demonstrate strategies, like social problem solving and conflict resolution, as ways to deal with these. Many group protocols call for role playing and rehearsal of skills that leaders demonstrate. Sexually abused children practice acting assertively, and impulsive children practice entering an ongoing conversation successfully. In the pilot study by Straub et al. examining a CBT group for teens with depression, participants reported that they felt understood by and comfortable with the group leaders, and rated their advice as helpful. Imitating effective behaviors modeled by

leaders who children respect and peers with whom children may identify effectively teaches functional behaviors.

Catharsis. Group therapy can allow one to express one's distress in the presence of individuals who can truly understand one's suffering. Although catharsis provides some relief, catharsis alone is insufficient for change. Members of one adolescent trichotillomania treatment group noted that talking at support groups had helped initially but left them feeling dejected after hearing about others' experiences. They felt more hopeful when they could learn new tools with which to address their behaviors.

Existential Factors. In the trichotillomania group just mentioned, one member openly expressed hopelessness and refused to address her hair pulling. The other group members challenged her stance, urging her to take responsibility for attempting to change her behavior. Unwilling to take charge at that time, she left the group. Issues of personal responsibility, will to change, loneliness, and demoralization frequently had emerged in the context of group therapy, which provided a rich format in which to discuss these universal existential concerns. Adults in one substance abuse group rated the group's ability to address existential issues as its most powerful component.

Cohesiveness. The experience of being part of something, being accepted, sharing and working together feels good and can improve children's motivation to work in therapy. When children and adolescents are secretive about their disorders, they may have a sense they do not belong anywhere. In a group, they can enjoy a sense of closeness when they can be open about their troubles, are accepted, and are understood. A study by McConachie et al. investigating group therapy for anxiety in children with autism spectrum disorder found that the most helpful strategies were learning alternative coping strategies, sharing experiences, and feeling accepted.

Interpersonal Learning. By working in the here and now of relationships in the group sessions, individuals have an opportunity to address dysfunctional interpersonal difficulties. The group can be a safe place to develop and practice healthy ways of interacting. For example, in a mentalization-based therapy group, teens are often encouraged to discuss when they find other participants' behaviors confusing or distressing.

Developmental Factors

Each stage of development has its tasks. Younger children's egocentric focus differs from preadolescents' yearning for same-sex soul mates and from adolescents' empathic, individuating, and affiliative capacities.

Understanding the developmental needs and capacities of a set of children or adolescents is necessary for designing and enacting a successful group therapy intervention. A parent-only, group cognitive-behavioral intervention for children with anxiety disorders has been used to take advantage of the working relationship between school-age children and their parents (and of the benefits of parents working together at their common stages of parenting). Aronson delightfully discussed ways in which adolescent developmental issues are addressed in a paper, "Where the Wild Things Are." In it, he likened the journey of Maurice Sendak's Max in the book of that name—trying on his new identity, affiliating with peers, individuating, and returning with a new relationship to his family—to psychological processes occurring in an adolescent group therapy. Whether the group is a cluster of same-age peers or contains a range of ages, the group interaction can foster stage-appropriate maturation.

Other Factors

Most geographic areas have an inadequate number of psychotherapists, and most individuals have little surplus money for psychotherapy. The group format uses therapist resources in a low-cost way to provide care that otherwise might not be available due to lack of resources.

TECHNIQUES

Gains made in symptom resolution and psychosocial development in groups result from general therapeutic factors and from learning specific information and skills. Before discussing specific diagnosis- and problem-focused interventions, this subsection reviews general technical issues.

Selection of Patients

Number. A typical group contains 6 to 10 members—a size large enough to allow for a rich exchange but small enough to prevent members from being neglected. Having an adequate number at the outset ensures against depletion due to dropouts and absences. Sometimes, goals call for a smaller or larger group. In one treatment study, aggressive young boys practiced coping with being teased. Because of the intensity of the intervention and potential for strong emotions and behavioral acting out, the interventions were done in triads. On the other hand, Meezan and O'Keefe studied a large-group format used to address abusive and neglectful caregivers and their children. Two or more members of at least two generations of each family met with five to seven other families and four therapists. More change (including a reduced incidence of child abuse and improved child behavior) occurred in the families in the intervention group than in a control group using an individual family treatment. The authors believed that families improved more because they met families

who had "been there" and other families who motivated them and encouraged them to change. Families felt that compared to feedback from therapists, feedback from other families was easier to take and helped them the most.

Describing an inpatient community group of 18 severely disturbed and conduct-disordered boys between the ages of 11 and 16 years, another author remarked that although maintaining structure and creating a safe environment presents a challenge with a large group of disturbed children, the larger group was powerful and a good choice for reluctant participants and those limited in their capacity to share their thoughts and feelings.

Age and Sex. Participants' age defines the nature of the interactions and structure of the intervention. Some investigators have included children of different ages and have found that members benefit from the interesting modeling and teaching opportunities provided from the family-like grouping. Most often, groups are composed of children at similar developmental levels, (e.g., 8 to 12 years and 12 to 18 years). When patients are of elementary school age, mixed-sex groups are often appropriate, especially for the younger grades. For this age group, games and activities can be fun while they facilitate interactions and learning. As the children age, they can tolerate a format increasingly focused on verbal interaction. Children in preadolescence prefer same-sex peers and can be distracted and inhibited by members of the opposite sex. Adolescents enjoy and can use mixed-sex groups again. Mixing high school–aged adolescents with younger ones may expose the younger children to topics with which they are not comfortable.

Excluding Patients. Certain individuals in group therapy may undermine the process sufficiently that, although the individual may benefit from participating, his or her detriment to the group calls for exclusion. Particularly in adolescent groups, individuals who actively use drugs may disrupt the group by showing up under the influence or reveling in escapades. When this does not serve the group's goals, these individuals should be excluded. Other fragile children and adolescents may experience the group as so stressful as to contraindicate their participation. Most group protocols exclude actively suicidal or psychotic individuals, who may monopolize the leader's attention, upset other members, and require the leader to leave the session to appropriately triage the patient. On the other hand, crises and decompensations are part of the territory of treating mentally ill children and adolescents. Witnessing the therapist compassionately and competently handle a troubled peer may be reassuring to a group member. Including an unwilling "captive" member in a group could drag the group down in a mutinous way or, skillfully handled, highlight issues of motivation and personal responsibility for each

member.

Investigating the importance of carefully selecting members (considering such factors as level of acting out or ego development), Johnson and Gold sorted 34 latency-aged boys into four groups based on a measured leadership quality score. Groups met weekly for 20 sessions. All group-treated boys improved in popularity ratings and leadership. The authors remarked that the therapist's ability to match techniques to the characteristics of the group members was more important than group composition.

Parents

Parents spend many more hours with children than therapists can and have the opportunity to influence their children by teaching, shaping, and modeling. A growing literature supports including parents in group interventions, often in parallel parent groups and sometimes as a sole intervention. A University of Michigan group tested the effects of a child-only, child-plus-mother, and waiting-list intervention for children who had witnessed intimate partner violence. The children participated in a 10-week group therapy aimed at increasing their knowledge about families and family violence, managing their emotions, and managing interpersonal conflict. The separate mothers' group taught the mothers about the influence of violence on children's development and worked to improve parenting and disciplinary skills. The child-plus-mother therapy was the most effective in improving externalizing behaviors, with gains continuing after the group ended.

When sexually abused children aged 3 to 12 (n = 617) were treated using group therapy with a concurrent parent group, it was found that children who participated in the group improved significantly in their symptoms, which included somatic complaints, interpersonal difficulties, psychological state, and behavioral problems. These children were less likely to be revictimized or themselves become sexual offenders, particularly when the parents also attended a group. Parents are an important resource for the reinforcement of skills and psychoeducation taught in groups.

Setting

Group therapies for children typically take place in the clinic or school. Advantages of clinic-based groups include confidentiality and coordination of care. The more available school setting offers the specific advantages of convenience for the children, opportunities for developing friendships with schoolmates who are accessible, and potential for doing primary prevention work. Some social skill group therapies based in clinics have improved socializing with group members, but results have failed to generalize to

school peers. Conducting social skills groups in the schools increases the chance for generalizability of skills. Unfortunately, with declining school budgets, schools are becoming less able to provide such care.

Therapist

The effective group psychotherapist understands child development, psychodynamic concepts, and characteristics of the child-targeted psychopathology. In addition, the therapist needs knowledge of factors particular to groups and the specific intervention curriculum. The therapist must foster a climate of security, tolerance, and acceptance for differences in feelings and behaviors; be able to control the tensions and anxiety level in individual group members within acceptable limits; and control group phenomena such as bullying, scapegoating, monopolizing, and instigating. Groups benefit from therapists who are flexible, playful, can regress without becoming a peer, have an ability to provide structure in an authoritative (not authoritarian) manner, and provide a good co-therapy relationship. The group therapist must be aware of and contend with his or her countertransference to the group and its members. As with all countertransference reactions, the therapists' anxiety and resistance can be used to further understand group process. By analyzing one's countertransference, one can alert oneself to possible enactments. It is important to remember that group resistance may reflect the therapist's own anxiety.

Co-Therapist Relationship

Having more than one therapist expands the capacity for processing multiple members' reactions and input and more efficiently administrates the group. Having a co-therapist also provides back-up in case a leader is absent. The co-therapist relationship has been likened to a marriage or sibling relationship: It takes a cooperative, complementary partnership that must withstand multiple potential stressors. Often, co-therapists are individuals with unequal training and experience. The pair must withstand group members responding to them as a parental duo and/or mentally splitting them into good and bad therapists. Throughout the therapy session, co-therapists must negotiate the dips and turns of members' (and their own) emotions in front of the group. Witnessing co-therapists problem solving together provides the group's child and adolescent members with an opportunity for modeling and recapitulation of the family. The co-therapist relationship requires a respectful, egalitarian stance in which therapists are open to processing their interaction and their reactions to the group. Unfortunately, due to declining reimbursements, co-therapy is becoming less feasible in community settings.

Tools

Creative group therapy tools range from simple games to high-tech video interventions and can facilitate the therapeutic process. One author described a low-tech approach—a variation of pass the parcel, in which a package is passed around the group for each member to remove a layer of wrapping and answer the question written on it.

Heiman and Ettin used a variety of interesting techniques in a therapy group for sexually abused girls. They began with a puppet-making exercise included to promote disclosure and build group cohesion. Girls built puppets in dyads and used them to introduce each other and as a supportive "friend." Girls named the group in the second session, a task that began their working together as a group and set the stage for them to take ownership and define the goals of the group. Other tools included a board game that asked questions about specifics of abuse, a coloring book about girls' rights and bodies, a collage project about girls at different ages used to facilitate a discussion of appropriate sexual behavior at different ages, letter writing to perpetrators to express strong feelings, and throwing into the trash drawings and writings symbolizing things they would like to put into the past.

A group for sexually abused boys made a group flag, which required creativity and compromises that led to group solidarity. This same group rolled a "feelings cube," which directed them to describe a situation causing a certain feeling, act the feeling out, or answer questions posed by group leaders and members. They made and photographed clay figures portraying their abuse. Each of these tools was chosen to help the girls and boys confront and process difficult issues.

Videos can facilitate learning and discussion in groups for children. Prevention films for sexually abused children can be used to teach self-protection skills. Films of social vignettes can provide a take-off point for discussion and role playing for social skills groups for children with attention-deficit disorder. The film *The Touching Tree* provides psychoeducation and models exposure/response prevention training in the context of a drama about a boy with OCD. Watching the video stories together is an efficient teaching medium and provides a shared experience from which discussion can flow in a lively way.

Games can also serve to introduce levity that promotes camaraderie and communication. In dialectical behavioral therapy groups, games are often used to enforce mindfulness techniques. In one activity, members choose a favorite from a bowl of mixed candies. Then they are asked to pass the favorite to the person on their left and mindfully eat the new piece of candy. In a mentalization-based therapy technique, teens are encouraged to "play doctor" using a stethoscope and dressing in white coats to empower themselves in discussing what they think may be going on in the minds and

hearts of other family or group members. Similarly, playing freeze tag where participants "freeze" showing how they look when feeling a certain emotion can help teens understand the concept of the "opaque mind" and illustrates that they can only guess at what goes on in another mind. Role play is yet another effective tool to promote perspective taking and communication.

In the era of YouTube, creating original video recordings is an exciting developmentally appropriate and available tool. For instance, to consolidate knowledge of information and skills acquired in one anxiety group, members chose to make an instructional video. In one report, therapists introduced a continuously recording video camera to a group in order to spark group interest when it was waning. They observed that the camera's eye acted as both an additional group member and as a transference object.

Food, obviously symbolically and physically nurturing, may be included in group therapies for practical and therapeutic reasons. Children coming to the clinic after school may be hungry, and feeding them may help to improve their attention. Some treatment protocols use a snack break to practice less structured socializing for socially phobic or impulsive children. Others use food as reinforcement for specific behaviors. The group therapist must decide whether the distraction caused by food is a reasonable cost for the gained therapeutic benefit.

Length of Intervention

The rate of improvement is greatest in early therapy sessions. With increased emphasis on cost containment and measuring outcome, most recent group therapies are scheduled to last only between 8 and 12 sessions. Therapeutic factors may also call for time-limited work. Formally dividing time into phases like introductory/group forming, working, and consolidation/termination phases may help with focus and momentum. This is not to say that longer interventions are not beneficial. Some youth may take longer to benefit and some benefit by reinforcement and taking a relative leadership role having acquired skills and knowledge.

Group Rules

In order for individuals to feel safe in taking the risks necessary for change and for therapists and group members to serve multiple individuals' needs, groups need rules. Developing a set of rules is usually a group activity early in the intervention. Fairly universal rules include those regarding confidentiality, personal safety, and attendance, not coming to the group while intoxicated, and respecting others. While, at first glance, prohibiting socializing with other group members outside of sessions may seem unnecessary, it has proven consistently to be an important rule. For

example, after members of a structured adolescent cognitive-behavioral group met with each other outside of the group, one member left the group, uncomfortable with feeling pursued by the other. Members of a trichotillomania group requested permission to use instant messaging on the Internet to seek support from each other when they had urges to pull. Although most felt that this was helpful, some eventually desisted, feeling that boundaries were impinged on.

Family Involvement

Families may be involved in group therapies for children and adolescents in a variety of ways: they may be included with the children in the session (multifamily group), in parallel parent groups, with families included in parts of each session, certain sessions, or not included at all. Childhood anxiety disorder researchers have examined the role of family involvement in treating anxious children. One study compared outcomes of group therapy in three conditions: parent-and-child parallel group therapies, a child-only intervention, and a parent-only intervention. Its authors found that, although all groups improved, those children in the parent-and-child intervention group used new coping strategies more actively outside of the group. Barrett and coworkers included parents in group therapy for anxiety disorders because the parents were at increased risk for having anxiety disorders and because the therapists believed that family interactions play a major role in the development of children's anxiety. They randomly assigned 79 children to group cognitive-behavioral therapy, group cognitive-behavioral therapy with family involvement, and wait-list control. The family component encouraged parents to reinforce courageous behavior and ignore excessive complaining and anxious behavior and it taught them communication and problem-solving skills to deal successfully with their children's emotional upsets and anxieties. After treatment and 1 year later, those in the combined-treatment group had a better outcome than those in the child-only condition (95 percent diagnosis-free vs. 70 percent).

Similarly, in a recent multi-center randomized controlled trial by Scholten et al., parent involvement for children with chronic illnesses did reveal additional positive effects on parent-reported internalizing, child-reported externalizing, information seeking, and social competence, when compared to the child-only arm. Parental involvement also appeared to contribute to effect persistence of the group intervention.

Process in the Structured Group

Group therapies increasingly use structured and manualized treatment protocols. Critics and skeptics of manualized treatment fear that therapy from a manual will be stiff and "cookie-cutter." Kendall, who developed the

most-researched manual for treating childhood anxiety disorders, the Coping Cat, noted that the therapist cannot be a "cookbook" therapist but must have a sound understanding of the theory behind the intervention and adequate flexibility and creativity to use the tools in a way that suits the needs of each patient. Manuals and outlines of therapy interventions inform and guide but do not control the group therapist. The therapist will lose effectiveness and opportunities to help if he or she ignores psychodynamic forces and other resistances to treatment. In an article on group therapy for sexual abuse patients, Heiman and Ettin stated that group dynamics and collective process do not disappear even within a highly structured psychoeducational group. If group members' affective reactions are not monitored and the collective dynamics are not acknowledged, then the group curriculum may have little impact. Especially in a group of adolescents, joining is crucial to promote learning. In some cases, the group process can act to disrupt the viability and stability of the group. For example, if a group member decides that a game or activity is silly, it may be difficult to have the rest of the group participate. At times, it may be necessary to pull such an individual out of the group to explore their reasons for disagreement or isolation and encourage rejoining. Group processes can be harnessed to maximize therapeutic impact. By responding to the messages underlying children's behaviors, rather than suppressing them, an opportunity exists to respond and transform the function of children's actions and maximize the efficacy of structured activities.

CLINICAL ISSUES

Indications

Group therapy should be considered if the child or adolescent has a condition for which there is evidence that group therapy is effective. This subsection provides descriptions of empirically supported group therapy interventions for specific pediatric populations. Rather than being an exhaustive review of all reported studies, it presents examples of treatments with support to describe the treatment elements and factors that made group therapy the choice. For some children, group therapy alone may be sufficient, whereas others require concomitant pharmacological, individual, or family intervention.

After the Events of 9/11/2001. The nation was stunned by the 9/11 terrorist attacks. At least 6,000 children lost a parent, and many others developed posttraumatic stress disorder (PTSD). Group psychotherapy served as a place where children could put their experiences into words and play, find support, and begin self-repair. Scheidlinger and Kahn made several observations during their group work with children after 9/11. They

used play, drawings, and discussion questions with elementary school–aged children to help them to express their feelings and fears. Play was especially important for traumatized children aged 4 to 8 years, helping them to describe their experiences. Expressing feelings and thoughts in the group setting allowed other group members to share their experiences and relate to others, a significant step toward building trust and unity within the group. The clinicians observed that traumatic events can continue to affect processes of social and mental development, and they emphasized the importance of making effective interventions like group therapy available in events with global impact.

Haen described a group therapy intervention for Connecticut children directly affected by 9/11. The group therapy consisted of two phases: an acute stage and a working-through phase. Many children felt alone and feared for their safety after experiencing the sudden loss of loved ones. It was imperative to provide a safe environment to bring together clients with similar experiences to minimize the sense of isolation they often felt. The first phase of group therapy aimed to develop trust, safety, and security and included work toward psychoeducation and connecting to one's support system.

The second phase in Haen's group therapy was designed to provide an open forum to address and discuss distortions and misconceptions developed by the children after 9/11. When children blamed themselves for the loss of their parent(s), work focused on correcting cognitive distortions. To facilitate the second phase of group therapy, Haen used play, art, and other nonverbal tools to communicate feelings and thoughts. In one example, drama therapy allowed group members to role-play and alter endings to traumatic events. In some cases, they were able to prevent imagined disasters. These tools helped participants to accept and cope with losses and improved the perception of security and safety for young clients.

Attention-Deficit/Hyperactivity Disorder. Although the Multimodal Treatment Study demonstrated that stimulant medication is the treatment of choice for uncomplicated attention-deficit/hyperactivity disorder (ADHD), children with additional difficulties require additional psychosocial interventions to optimize function. Hinshaw and coworkers designed a cognitive-behavioral therapy intervention for boys with hyperactivity, in which peers were a necessary component for practicing social skills. An initial study compared outcomes for a group cognitive-behavioral therapy intervention for 21 boys between the ages of 8 and 13 years treated with methylphenidate (Ritalin) or placebo. The boys were first taught problem-solving strategies: define the goal or problem, generate multiple solutions, consider the outcome of the various solutions, and choose and operationalize a behavior. Working in groups of three, the subjects applied these problem-solving strategies, first to a series of

academic and fine-motor tasks, and then to interpersonal problem vignettes. In addition, the boys practiced maintaining self-control while actually being teased by peers in the group. After this intervention, both groups improved in self-control and in the intensity of their response to teasing.

In a second study, 24 boys with hyperactivity and ADHD participated in daily 80-minute cognitive-behavioral training groups in the context of a 5-week summer enrichment program at the University of California, Los Angeles. Following a scripted protocol, therapists instructed participants in problem-solving skills, self-talk strategies, and discussion about medication issues. One subset of subjects practiced recognizing signs of their own anger, rehearsed strategies for dealing with provocation, and then experienced provocation in the form of teasing from group peers and group leaders. The other subset talked about perspective taking without actively rehearsing specific strategies or undergoing the teasing stress inoculation. All boys were randomized in a double-blind manner to receive placebo or methylphenidate. Those boys trained by rehearsing self-control with peers demonstrated significant improvements in self-control and engaged in purposeful coping strategies. Post hoc analysis suggested that those with more experience seeing peers model coping behaviors in the group benefited most. The medication influenced outcome ratings of self-control intensity but not frequency of reaction to teasing.

Pfiffner and McBurnett randomized 27 children with ADHD to receive 8-week group social skills training, social skills training with a parent group and teacher feedback, or wait-list control. The group format was selected so children could practice newly learned social skills and receive social reinforcement from peers. The therapy included having the children report on their behavior (measured as points earned) during the week, pool their points together for a group pizza party, role-play specific social skills with immediate peer feedback on performance, and share fun activities. Outcome results after several months supported the utility of social skills training, with or without parents'and teachers' involvement. The authors noted that strong group reinforcement in a stimulating, developmentally appropriate program led to significant improvements in social skills and problem behavior.

Waxmonsky and coworkers designed a group cognitive-behavioral therapy, along with a parent-training intervention to manage both behavioral and emotional dysregulation elementary school boys treated with psychostimulants. The intervention consisted of nine parallel 105-minute child and parent groups that were run simultaneously. After 9 weeks, the boys showed improvements in mood symptoms. Their parents reported that they were parenting more effectively.

Aggression/Conduct Disorder. Children who are aggressive and have

conduct disorder have social difficulties, especially in frustrating circumstances. Concerned that parent training for oppositional defiant and conduct disorder did not help all families or generalize to school, Wester-Stratton and coworkers designed a social skills and problem-solving training protocol (Incredible Years: Dinosaur Curriculum) for children between the ages of 4 and 8 years. The investigators used the group as a supervised place to practice social skills. Ninety-nine children with the diagnosis of Oppositional Defiant or Conduct Disorder were randomly assigned to a control condition or a group intervention that included directly taught social skills, problem solving, and anger management. Groups of five or six children met weekly for approximately 6 months. Children in the treatment condition watched videotaped vignettes of children coping with stressful situations. They discussed the vignettes and collaborated to practice a repertoire of acceptable solutions and coping skills. They also discussed feelings they had about troubles caused by their own behaviors. To enhance motivation, attention, and learning and provide a developmentally appropriate milieu, therapists used puppets, cue cards, coloring books, small reinforcements, group activities, and games. The treatment group demonstrated significant clinical improvement in conduct problems as rated by parents and teachers to a greater extent than did controls. Those gains were persisted at 1-year follow-up. The authors noted that many of the children made their first good friends in the group therapy and postulated that intervention might be even more effective if delivered in the subjects' actual peer setting.

In response to a paucity of resources for treating conduct disorder in South Africa and in hopes that group cognitive-behavioral treatment could help children and adolescents with conduct disorder to respond more appropriately to social situations, a group of investigators designed and studied a 12-week intervention program. An occupational therapist and a psychology student were trained and supervised by an experienced adolescent group therapist. The sessions included: (1) an overview of rules, expectations, and rewards; (2) teaching problem solving/decision making; (3) impulse control, response gap thinking; (4) assertiveness training; (5) requesting behavior change; (6) reward session; (7) empathy, perspective taking; (8) sex differences, interactions, relationships; (9) sex education video, relationships; (10) norms, values, role models; (11) giving compliments; and (12) closure. Parents of treated and control subjects were invited to six sessions of parent behavior management training. The treatment group showed significant improvement in defiance, physical aggression, and delinquent aggression posttreatment and at 6 months, whereas the control group demonstrated improvement only in defiance and a remarkable increase in truancy. The authors noted that this brief intervention using community workers showed promise as a sustainable intervention.

Anxiety. Withdrawal, social isolation, and difficulty with social skills are significant troubles for most anxious children and adolescents. An opportunity for direct exposure to peers, practicing social interaction, and receiving social support and reinforcement make group therapy a theoretically attractive treatment modality for anxious youth.

Cognitive-behavioral therapy for anxious children and adolescents has empirical support. Several manualized treatments have been researched and are readily available. Most include psychoeducation about the somatic, cognitive, and behavioral components of anxiety and build coping skills (relaxation, cognitive restructuring, active approach to feared situations, problem solving). Most use graded exposure to anxiety-provoking situations, homework, and reinforcements to achieve symptomatic change.

In 1998, Barrett reported results of a cognitive-behavioral therapy group for anxious children and their families. The manual-based treatment focused on providing children and families with skills in managing anxiety and avoidance, problem solving, and mutual family support. Posttreatment and 12 months later, 65 percent of children in the child-only group, 85 percent of those in the group with family involvement, and 25 percent of the wait-list control group were free of diagnosis.

Using a different cognitive-behavioral protocol with a strong emphasis on contingency management, Silverman and coworkers also compared a group intervention for anxious children and their parents to a wait-list control group. They considered disadvantages of group therapy—dampening of the therapist–child relationship, overly taxed therapist resources, possible negative modeling—along with possible advantages. Sixty-four percent of the treated subjects versus 13 percent of the wait-listed subjects were free of their primary diagnosis posttreatment. At 6 months, 79 percent of the treated group no longer had their primary anxiety diagnosis. This investigation's finding elevated group cognitive-behavioral therapy to "probably efficacious," according to the American Psychological Association's task force, and took it closer to being a "well-established" therapy, in that it used a manual, had clearly specified subject characteristics, and had been successful with more than one set of investigators.

Building on evidence supporting individual cognitive-behavioral therapy efficacy in treating childhood anxiety, and predicting that group therapy would lead to gains in social competence, friendships, social activities, and improvements in social isolation and loneliness, Flannery-Schroeder and Kendall treated children between the ages of 8 and 14 years with individual or group manualized cognitive-behavioral therapy. Posttreatment, 73 percent of individually treated, 50 percent of group-treated, and 8 percent of wait-listed children no longer met diagnostic criteria for their primary anxiety disorder. Anxious distress and coping improved. The hypothesis that social functioning would improve more in

the group was not supported. The authors noted that the protocol did not specifically address social functioning and that only a small proportion of the sample endorsed problems in social functioning at outset, each of which may explain this finding.

Beidel and coworkers compared a combined 12-week group and individual treatment for latency-age social anxiety disorder with an active control condition—the Test Busters study skills protocol. The treatment condition consisted of child and parent education about social anxiety, social skills training, peer generalization experiences, and in vivo exposure. In the group social-skills training component, children learned about greetings, starting conversations, maintaining conversations, listening and remembering skills, skills for joining groups, positive assertion, negative assertion, and telephone skills. One topic was covered each week, using instruction, modeling, behavioral rehearsal, corrective feedback, and homework. Immediately after the social skills training session, the anxious children joined a group of not-anxious peers for a 90-minute fun, unstructured group activity (such as bowling or in-line skating). During another session, each week, children met individually with the therapist to accomplish exposure activities, practicing one task, such as reading aloud in front of others, per week.

The control group met once individually and once weekly in a group to learn and practice study and test-taking skills. At the end of the group, 67 percent of those in the treatment condition and 5 percent in the study skills condition no longer met diagnostic criteria for social anxiety disorder. Gains were maintained at 6 months. In this treatment, investigators thoughtfully combined two forms of group intervention and an individual session. The first group session allowed for economic use of resources for teaching and an opportunity to practice skills in a structured session. The second, unstructured group included non-affected peers and provided an opportunity to expand on and generalize the skills in a more challenging situation. The individual exposure session allowed for customized, individual, symptom-driven intervention.

In another exploration of the benefit and duration of early intervention for anxious children, an in-school group treatment study compared cognitive-behavioral group therapy with and without parental inclusion with a waiting list control group. In this sub-clinical sample, identified by a questionnaire or by teacher identification, children in the group treatments sustained improvement in their function and anxiety symptoms more than the control group over the 12 months examined. This study by Bernstein and colleagues advanced knowledge in three ways: by including a control group throughout the study, by serving as an early intervention for anxious children and by bringing the intervention to the children in their schools.

Recently, a group piloted a school-based group therapy targeting at-risk adolescents with sub-clinical social anxiety and depression. They chose

students who had been victimized by peers: physically, by rejection or by damaging their reputations. After 3 individual and 10 group sessions based on Interpersonal Therapy for Depression, social anxiety, depression and peer victimization had each improved.

Bereavement. Universality may be an especially powerful therapeutic factor in groups gathered to deal with loss. Pfeffer and coworkers used manual-based group therapy to assist children after the suicide of a parent or sibling, with the goal of decreasing depression and anxiety. Families of children between the ages of 6 and 15 years were contacted within 3 weeks of the suicide death of a parent or sibling and were randomly assigned to receive or not receive intervention. Children with psychiatric diagnoses were excluded. The intervention consisted of ten 1.5-hour weekly group sessions of two to five children within 3 years of age of each other, with simultaneous parent groups. The children's curriculum consisted of offering psychoeducation about death and its permanence, identifying feelings of grief, defining what suicide is, and discussing why people commit suicide. Children's groups discussed preventing suicidal urges and presented problem-solving skills. The therapy facilitated expressions of grief and promoted identification with the positive aspects of the deceased. Parents learned about childhood bereavement, learned to help their children express their grief and discuss the suicide, learned to identify morbid reactions in the children, and learned to promote children's healthy functioning. The parents' grieving was also supported. Children in the therapy developed new supportive relationships. They showed greater reductions in anxiety and depression than did the nonintervention children.

Depression. Group therapy is among the most empirically supported therapies for adolescent depression. A large study compared group cognitive-behavioral therapy, group cognitive-behavioral therapy with families, and a wait-list control group and three intensities of follow-up care for depressed adolescents. One hundred and twenty-three adolescents with major depression were randomly assigned to 16 sessions of group cognitive-behavioral therapy, group cognitive-behavioral therapy plus nine sessions of parent group, or a wait list. Mixed-gender groups of up to 10 adolescents met for 2 hours twice weekly for 8 weeks. The therapy used a manual and workbooks and taught skills of mood monitoring, improving social skills, increasing pleasant activities, decreasing anxiety and depressogenic thoughts, and improving communication and conflict resolution. In the parent condition, parents reviewed the adolescent course content and met once per week and twice jointly with the adolescents. Posttreatment, 67 percent of treated and 48 percent of wait-listed subjects were diagnosis free. Neither parental involvement nor booster sessions

made a significant difference in measured outcome.

Hypothesizing that treating depressed Latino youth in groups would be successful, researchers recently compared group and individual cognitive-behavioral therapy (CBT) and interpersonal therapies (IPT). Authors considered that Puerto Rican cultural value placed on *personalismo* (preference for personal contact in social situations) and *familismo* (placing the interest of the family over the individual) would be served in the group therapy format. Analysis of outcome favored CBT over IPT, which may have resulted from its increased structure and directiveness (consistent with *respeto*) and found results to be similar in the individual and group formats.

The next study described addresses incorporation of new referrals into an ongoing group treatment. This is important because difficulty gathering a cohort of patients at once can impede the initiation of group work. Wood and coworkers studied a group format for treating adolescents who deliberately harm themselves. Sixty-three adolescents who had intentionally injured themselves were randomly assigned to either group therapy along with routine care or routine care alone. The group format was chosen for clinical and administrative reasons: because poor peer relationships and impaired problem solving were thought to be amenable to group intervention, and because group therapy was hoped to decrease use of individual therapy and train of a team of therapists for dealing with self-harm. The authors advanced the field in writing a manual describing necessary complex psychotherapeutic techniques. The manualized therapy —a developmental group psychotherapy—included a variety of techniques, including problem solving, cognitive-behavioral therapy, positive corrective therapeutic relationships, dialectical behavior therapy, and psychodynamic group therapy. After assessment, six acute sessions were followed by long-term group therapy, which was open-ended. Group members could join the groups any time. Themes of acute sessions included relationships, school problems and peer relationships, family problems, anger management, depression and self-harm, hopelessness, and feelings about the future. While the group therapy did not reduce levels of depression or suicidal thinking, participation in the group (the median number of sessions was 8; the range was 0 to 19) did lead to a change in the incidence of deliberate self-harm and the use of other therapy resources. Those who had more group sessions were less likely to harm themselves again, whereas those who had more individual sessions were more likely to repeat self-harm. This study is interesting not only for the clinical outcome, but also for a model for running a continuous group, adding newcomers after a six-session lead-in.

Stallard and coworkers designed a clinical and cost-effective study of classroom-based CBT for depression in high-risk adolescents. Of the 5,030 students that were asked to participate, 1,064 adolescents aged 12 to 16

years old were considered high-risk, and were randomized to a CBT group, the usual school curriculum: Personal, Social, and Health Education (PSHE) delivered by teachers, or an attention control group (the usual school PSHE with two additional facilitators). This study examined whether these preventative interventions for depressive symptoms could be both clinically effective and cost-effective. All three groups showed decreased depression scores with no significant differences between groups. Because it required trained therapists, classroom-based CBT was not shown to be cost-effective compared to the teacher-administered groups. The study did suggest that a preventive school-based group intervention could decrease adolescents' degree of depressive symptoms.

Internet-based interventions may be useful to decrease social isolation, which may contribute to depression. Van der Zanden et al. studied a treatment that involved groups of six or fewer people aged 16 to 25 in 90-minute chat-room therapy sessions. Participants were randomly assigned to this online cognitive-behavioral group course known as Master Your Mood (Grip op Je Dip) and a wait-list control group. The therapy focused on restructuring thinking patterns by recognizing one's unproductive, unrealistic thoughts and transforming them to more helpful, realistic thoughts. Performance of daily activities and mood monitoring was also done. Reductions in depression and anxiety and an increased sense of mastery were seen after intervention and at 3-month and 6-month follow-up. A functional MRI study found alterations in strengths of brain area connectivity after only 5 sessions of group psychotherapy for depressed adolescents.

Borderline Personality Disorder/Self-Harm. Few empirically validated treatments for self-injurious behaviors and suicidality exist. In a study of brief group therapy, Green and coworkers randomized adolescents that had had 2 episodes of self-harm in the past 12 months into two arms: group therapy intervention with routine care vs. routine care. The manualized intervention incorporated cognitive-behavioral therapy and dialectical behavioral therapy. Therapists with 3 years of experience administered the therapy over 6 weekly sessions. There were no differences in terms of frequency of self-harm and severity of self-harm between the intervention and routine care in this brief trial. However, in a longer intervention, Fleischhaker and coworkers conducted a pilot study of 12 adolescent females who were engaging in non-suicidal self-injurious behaviors. Each female completed Dialectical Behavioral Therapy for Adolescents (DBT-A) over a period of 16 to 24 weeks. Prior to therapy, 8 of 12 adolescents had attempted suicide at least once. Results were examined 4 weeks post-DBT-A and at a 1-year follow-up. No suicide attempts were reported during DBT-A or at the 1-year follow-up.

Substance Abuse. Many interventions for substance abuse are in use, many with little empirical support. Kaminer and coworkers compared two different group therapies in a randomized study of substance-abusing adolescents. Subjects assigned to a cognitive-behavioral group therapy and those assigned to a psychoeducational group therapy each improved. In discussing this outcome, the authors noted that group therapy for this population may be useful, in part because of exposure to the type of peer who could be a negative influence under other circumstances (a credible peer), in the setting in which each has a goal of not using psychoactive substances. This provided the adolescent with positive modeling, positive reinforcement for efforts, and an alternative peer group within the group.

The mutual self-help group, Alcoholics Anonymous (AA), has had 2 million recovering alcoholic members attending meetings in over 100,000 groups since 1935. In adults, participation in AA improves outcome and reduces health care costs. Many adolescents are referred to AA for continuing care. In an attempt to assess AA's effects on adolescent substance use, Kelly and colleagues followed 160 adolescents at intervals for 8 years (84 percent) after an inpatient stay for substance abuse. They found that participation was high initially and dropped consistency over the 8 years, with those whose lifetime frequency of alcohol intoxication and whose perception that they had a problem with alcohol attending most. Those who attended treatment most frequently during the first year had significantly better outcomes over the 8 year duration of the study. Those who perceived that they could use alcohol in moderation attended fewer groups and were consistently less abstinent. The data suggested that in adolescents with substance abuse, attendance at one or two AA meetings per week increased abstinence and that attending at a frequency of three times per week was associated with *complete* abstinence. The authors speculated that the mutual self-help group model succeeds because it may help maintain motivation, provide a new social context of sober peers, enhance coping skills and improve self-efficacy.

Motivational interviewing (MI) has been typically used in individualized treatments, but D'Amico and coworkers developed a group MI intervention called *Free Talk*. They compared substance abuse of youth in a teen court setting who were randomized to receive either *Free Talk* or usual care. *Free Talk* was administered by psychology doctoral graduate students and focused on MI strategies that included pros and cons of continued alcohol and other drug use vs. cutting back or quitting, used willingness and confidence rulers to determine if teens wanted to change or not change their drug use, and supported teens where they were in terms of their drug use. While *Free Talk* teens reported higher quality and satisfaction ratings, alcohol and drug use and delinquency were not significantly different at 3-month and 12-month follow-up.

Internet Addiction, Online Multiplayer Video Gaming. With the Internet being one of the most commonly used methods of communication and entertainment, Internet "addictions" are emerging among adolescents.

There have been growing concerns that online video game play may be addictive and invite people to escape from their problems or daily responsibilities. Snodgrass and coworkers sought to explore some of these concerns using the online game World of Warcraft (WoW), which may be played by groups of people regularly, simultaneously, internationally, and all night. They also investigated online gamers' motivation, looking at the factors of achievement, social connectivity, and immersion, factors that tend to foster problematic play in the online worlds. "Achievement" refers to competitiveness drawing players deeper into game, where they spend more time online than they might wish. "Social connectiveness" refers to getting drawn into communities or associations that also demand more of their time. "Immersive play" refers to players who not only temporarily escape but more persistently or maladaptively avoid real-world problems. Researchers conducted 30 formal interviews and 252 people participated in a web survey. Results showed that subjects' online play could be a source of distress in their lives. They also showed that players with low levels of offline success (i.e., good job, high income, having a good social network of family and friends) may attempt to compensate by playing excessively and players with high levels of offline success are less likely to overplay. Players who experienced dissociation (losing track of the real world immersed in the virtual world) were more likely to have problematic play. Playing multiperson WoW to be socially connected correlated with such dissociation and problematic play. While WoW group interaction can be seen as restorative for some individuals, it can be problematic when players who dissociate too fully, losing themselves in the virtual world of WoW, use it as a primary social outlet and thereby risk more problematic behaviors both online and offline.

In another online survey in the United States and Germany aimed at exploring group online interaction, 523 respondents were asked questions regarding their "character attachment," defined as internalization and psychological merging of the player's and character's mind, to explain their prosocial and antisocial gaming motivations. Bowman and coworkers showed that players who were strongly attached to their character had stronger motivation to use their social skills. Less attached players tended to care less about negative consequences such as developing a negative reputation, losing other players' confidences, and difficulty forming offline relationships. More attached players has desires to form more good relations with others.

Du used a cognitive-behavioral group therapy to treat Internet addiction in adolescents. Fifty-six adolescents aged 12 to 17 were randomly assigned to an active treatment group or a control group. Group CBT

involved 6 to 10 students participating in an eight-session group CBT run by two child and adolescent psychiatrists. Each session was of 1.5 to 2 hours. Topics included recognizing and controlling one's feelings, healthy communication between parents and children, techniques for dealing with relationships and content via the Internet, techniques to control one's impulses, techniques for recognizing when behavior is becoming addictive, and ways to stop addictive behavior. A group parent session was conducted simultaneously with the adolescent group CBT. Parents learned how to recognize teen's feelings, how to promote good communication, problem-solving principles and techniques, controlling one's own feelings and behavior, and managing adolescents with Internet addiction. Subjects had higher levels of anxiety at baseline compared to the control group. After the intervention, and 6 months later, the intervention group had more control of their time and used time more efficiently. They had fewer emotional symptoms, anxiety, conduct problems, and attention-deficit problems and better peer relationships.

Eating Disorders. Heinicke and colleagues completed a study of manualized group therapy for girls who identified themselves as having disordered eating, using a self-help manual, My Body, My Life: Body Image Program for Adolescent Girls. The intervention consisted of six weekly group sessions facilitated by a trained therapist, and postintervention sessions at 2-month and 6-month follow-up. Individuals were asked to read the manual and complete the activities each session. Seventy-three adolescent girls aged 12 to 18 years old were randomized to the intervention group or a delayed treatment. The 28 girls who completed the intervention showed improvements in body dissatisfaction, disordered eating, and depression, post-intervention and also at 2-month and 6-month follow-ups.

Gender Identity Issues. Choosing group therapy as a way to address shame and isolation, Rosenberg used supportive group therapy as a component of treatment for children with gender identity issues. In a case report, she described her regimen of individual plus group treatment for working with 12 children and their parents. All but one child had comorbid symptoms of anxiety, inattention, and social isolation. All of the children had been taunted and derided. The child group component to treatment was proposed to help self-esteem through the experience of universality. In the parent group, parents expressed relief to talk openly, for the first time, about their children. They shared their experiences and fears that their children might be harmed physically, emotionally, or by sexually transmitted illnesses. They were coached on ways to interact assertively with school personnel, friends, families, and neighbors and on how to access resources. The author reported that the children made significant

clinical improvements in comorbid symptoms and began to express acceptance of their biological gender.

Menvielle and Tuerk also reported benefits of a group for parents of non–gender-conforming children. In the monthly parent support group, therapists focused on child themes and parent themes. Discussions focused on reducing the polarity of gender-based interests (girl stuff vs. boy stuff), communicating acceptance of the child's thoughts and feelings about gender and interests, setting limits on behaviors that invite trouble, facilitating peer relationships, dealing with bullying, and discussing the possibility that the child will be homosexual as an adult. Parent-focused themes included helping parents to grieve, developing coping mechanisms for dealing with embarrassing situations, and learning to intervene when necessary in social and school situations. The authors noted that the group helped to destigmatize the nonconformity for parents and that humor helped the group to build cohesiveness and reduce tension, elicited compassion, and provided insights.

Posttraumatic Stress Disorder/Sexual Abuse. Group therapies are widely used to treat individuals who have been traumatized by molestation or other events. In an Indonesian community plagued by poverty and a decade of armed violence between religious groups, Tol and colleagues introduced a 15-session school-based intervention administered in schools by volunteers with no formal mental health training. The manualized cognitive-behavioral therapy protocol included 18 group sessions, with an individual pull-out session to establish a stimulus hierarchy at week 10. Elements, in order of presentation, included overview, information gathering, group rules, relaxation training, anger management through problem solving, examination of self-talk and cognitive restructuring around PTSD, introduction of a narrative about the trauma, imaginal and in vivo exposure, confrontation of PTSD-induced dysfunctional beliefs, summary, and graduation party. Subjects treated with trauma-processing activities, cooperative play, and creative-expressive elements improved and maintained over 6 months in measurements of posttraumatic stress disorder symptoms and hope, in the setting of continued environmental conflict and stress. Girls improved more than boys did. Authors concluded that a classroom-based intervention can be an appropriate approach to target large numbers of children with stress-related symptoms, in which the start dates but not the psychometric assessment dates were staggered. The authors hoped that such time-limited manualized treatment might help in situations of severe natural or human traumas.

British investigators Trowell and coworkers randomly assigned 71 symptomatic, sexually abused girls between the ages of 6 and 14 to participate in either 30 sessions of manualized, focused, weekly individual psychoanalytic therapy or 18 sessions of manualized psychoeducational

and psychotherapeutic group psychotherapy (hours spent in therapy were comparable). In the individual psychoanalytic therapy condition, play materials were provided, topics to be addressed were designated in the manual, but the topics were addressed as they arose, and each was covered. The group topics were prearranged, with notebooks, task sheets, and play material. Time was also allowed to address the relationships between the girls and their relationships with the co-therapists, and this was linked to past and current relationships, losses, and disruptions. Outcome was measured posttreatment at 1 and 2 years. Individually treated subjects had significantly greater improvements in PTSD symptoms. Each treatment led to improvement in general functioning and in anxiety and depressive symptoms.

In the specific context of child sexual abuse, group psychotherapy is considered the "treatment of choice" because of its impact on negative effects of sexual abuse: isolation, stigma, and self-esteem. The group format can foster a sense of belonging among peers and help overcome the element of secrecy, which is a common component of child sexual abuse. Group and individual treatments of sexually abused children and adolescents have now been found to be equally efficacious, finding that the adjusted expense of individual therapy is 64 percent more than group treatment formats.

Deblinger and coworkers performed a randomized comparison study of group supportive versus group cognitive-behavioral therapy for young children who had been sexually abused and their mothers. The interventions aimed to improve the adjustment of these young children by helping them to learn to communicate about their feelings, identify inappropriate touching, and learn abuse response skills. In the children's support group, material was presented in a didactic format. The children's cognitive-behavioral therapy format was interactive, using a workbook, role play, behavioral rehearsal, and joint parent–child exercises. In groups parallel to the children's, the parents were randomly assigned to cognitive-behavioral group therapy or support groups. In the cognitive-behavioral therapy condition, parents were assisted in coping with their emotional reactions to the abuse and their children, taught about effective communication with their children in general and with regard to sexual abuse, and provided behavior management skills to assist them in handling their children's difficult behaviors. The support group was run with a manual guiding the therapist with skills of active listening, unconditional positive regard, reflecting feelings, and empathy and provided information about group dynamics.

The magnitudes of effect sizes for changes posttreatment and at 3 months for children and mothers who participated in the cognitive-behavioral therapy groups were large on most measures. For the support group subjects, effect sizes were medium. Parents in cognitive-behavioral

therapy expressed greater satisfaction with the therapy. Children in the cognitive-behavioral groups (which used interactive behavioral methods) demonstrated greater knowledge and retention of body safety skills.

Improving Adherence to Medical Regimens in Chronic Illness. Parents of children with chronic illnesses must deal with many demands: attending to their sick children, interacting with the health care system, their employers, while caring for the rest of their families and their own emotions. In New Zealand, investigators developed and tested a Strong Parents–Strong Children 6-week group intervention to help parents to manage these stresses and learn practical skills for accessing services. Parents learned to plan time for themselves, address catastrophizing cognitive distortions, and use relaxation techniques. They practiced strategies for obtaining and sharing information from relevant health professionals and maintaining family life and social contacts with the goal of handing off aspects of illness management to the child when developmentally appropriate. Parents who participated in the group perceived themselves as more effective and more able to manage stress and perceived the family as more cohesive, less anxious, and better able to cope than did those on the waiting list.

For some adolescents, adherence to life-saving medical regimens is poor. Some preliminary studies suggest that treating adolescents in groups together with parent or "treatment buddy" involvement may help. Viner and colleagues compared a measure of diabetes control (hemoglobin A1c) in adolescents who elected group therapy with that in adolescents who were assessed but declined therapy. Intervention consisted of motivational enhancement therapy, analysis of strategies that help (solution-focused therapy), and a parallel parent group. Those treated in group therapy improved in glucose control (decrease of hemoglobin A1c of 1.5 percent). In a pilot of a group intervention involving "treatment buddies" and adolescents aimed at improving adherence to antiviral medication therapy, nearly all participants increased adherence to medication, as well as other important health behaviors.

Social Skills. Group treatment settings provide obvious opportunities for learning and practicing social skills. Children and adolescents with various disorders may have impaired ability to perceive important aspects of social intercourse, deficient knowledge of what to do to interact effectively, and/or ineffective means of being with others. Social skills have improved using therapy in groups for children and adolescents with autism spectrum disorder, aggression, or ADHD.

A German group recently reported results of an 11-month intervention for youth with communication difficulties using the manualized Frankfurt Social Skills Training (KONTAKT). Seventeen subjects aged 9 to 20 attended weekly (for children) and biweekly (adolescents) sessions in

groups of 4 to 7 participants. The protocol encouraged flexibility within structured sessions. Flexible elements included social interaction games, role play, group discussion, group discussions, pantomime games, and computer-based affect recognition games focused on recognition of emotions on human faces and eyes. In this pilot study, the youth showing the most improvement included those with higher IQs and higher initial levels of communication difficulties.

Fifty-two peer-nominated, socially rejected, inner-city, fourth-grade boys participated in a controlled study of social relations training program by Lochman and coworkers. Over 7 months, this program combined 26 individual sessions with 8 small-group sessions to focus on social problem solving, positive play training, group entry skill training, and dealing effectively with strong negative feelings. For example, in learning how to enter a group, children were trained to assess the general climate of the group, analyze the leadership structure, and match their style of entry to the ongoing play and activities of the group. Only the children who were aggressive showed significant treatment effects in reduced aggression, reduced social rejection, and improved peer prosocial behavior at the end of the intervention and at 1-year follow-up. This study illustrates the successful use of a group component of an intervention chosen for its use as a practicing ground.

Targeting social perception, knowledge of social rules, and peer group behavior in hospitalized boys with ADHD and conduct disorders, Kolko and coworkers compared outcomes of group social skills training with social activity group participation. Fifty-six boys were assessed by self-report on assertiveness and loneliness, by staff reports on social problems and social acceptance by peers, and by peers on popularity. The subjects were rated on 15 role-plays for competence and on behavioral observations that detailed social involvement, gaze, physical space, appropriate voice loudness and inflection, using openers and compliments, making assertive requests, responses to provocations, and appropriate play.

The manualized treatment groups of four to five boys met for three 1-hour sessions per week for 5 weeks. In each session, group rules were read, the last lesson was reviewed, and questions were asked to assess retention. Skills training was accomplished by reviewing and rating videotaped role plays of situations handled correctly and incorrectly. Incorrectly played situations were correctly role-played, discussed, and then rehearsed in the group. Points earned in the unit's token economy were recorded in the group. Outside the group, a "Secret Agent" awarded stars when the group members were caught appropriately putting the skill learned into action.

The social activity groups of five to seven boys met for semistructured group activities and games designed to demonstrate specific concepts likely to promote improved relationships, such as cooperation, friendship, empathy, and telling the truth. These sessions also began with a rule

review, followed by a description of the planned activity and its underlying concept, and then the activity. Leaders provided the boys with feedback about their overall participation and awarded points.

Outcome was assessed posttreatment and at 1-year follow-up. Those in the social skills training group showed significant improvement in loneliness, in staff-rated peer relationships, and in social competence as measured by the role-play assessment. Children in the social skills training group also showed improvement in almost all of the behavioral observation measures, whereas those in those social activity groups showed none. Both groups improved in assertiveness and social initiations. At 1 year, teacher reports revealed more improvement for the social skills group than for the social activity group. Providing a format with opportunities for rehearsal and both therapist and peer feedback likely led to better results.

Limitations of Group Therapy

When nonbehavioral approaches such as nondirective and psychodynamic models are used, meta-analysis found that group therapy demonstrated a smaller effect size. In group therapy, the therapist is constrained in his or her ability to attend to specific individual troubles and has less flexibility in a crisis. Gathering sufficient group members may prove difficult for individual practitioners.

Complications

A patient's ability to discuss emotional issues and be vulnerable in practicing new skills in a group is dependent on his or her feeling safe and respected. When confidentiality is breached, when occurrences of disrespect are inadequately processed, or when individuals feel physically unsafe or threatened by interaction outside of the group, the group is acting counter-therapeutically. Adequate therapist training and supervision, careful screening of patients, and close attention to the group rules and process may guard against these occurrences.

Goals of Treatment

Each group therapy aims to ameliorate symptoms and promote psychosocial growth. The rationale for choosing to provide therapy in a group format must be that group therapy, in particular, either has evidence to support its choice or has qualities that make it theoretically suited to address the patient's individual problems.

ETHICAL ISSUES

As implied previously, group therapy, although possibly cost-effective, should not be prescribed with that sole consideration in mind. It is properly used as one tool of many in the armamentarium of psychotherapists when

evidence supports its use or, lacking other more empirically supported therapy, theory provides a rationale.

RESEARCH AND EVALUATION

As delineated by the Task Force on Promotion and Dissemination of Psychological Procedures, those treatments possessing empirical support should be identified, developed, disseminated, and used. According to its criteria, most traditional group therapies fell into the experimental category. A bustling movement is growing to develop well-established treatments. Doubtless, group therapies for children and adolescents will be among those whose use will be supported.

REFERENCES

Aronson S. Where the wild things are: the power and challenge of adolescent group work. *Mt Sinai J Med*. 2004;71:174–180.

Billow RM. The therapist's anxiety and resistance to group therapy. *Int J Group Psychother*. 2001;51(2):225–242.

Bowman ND, Schultheiss D, Schumann C. "I'm attached, and I'm a good guy/gal!": how character attachment influences pro- and anti-social motivations to play massively multiplayer online role-playing games. *Cyberpsychol Behav Soc Netw*. 2012;15(3):169–174.

Clarke GN, Rohde P, Lewinsohn PM, Hops H, Seeley JR. Cognitive-behavioral treatment of adolescent depression: efficacy of acute group treatment and booster sessions. *J Am Acad Child Adolesc Psychiatry*. 1999;38:272–279.

D'Amico EJ, Hunter SB, Miles JN, Ewing B, Osilla KC. A randomized controlled trial of a group motivational interviewing intervention for adolescents with a first time alcohol or drug offense. *J Subst Abuse Treat*. 2013;45:400–408.

Deblinger E, Stauffer LB, Steer RA. Comparative efficacies of supportive and cognitive behavioral group therapies for young children who have been sexually abused and their nonoffending mothers. *Child Maltreat*. 2001;6:332–343.

Du Y, Wenqing J, Vance A. Longer term effect of randomized, controlled group cognitive behavioral therapy for internet addiction in adolescent students in Shanghai. *Aust and New Zea J Psych*. 2010;44:129–134.

Flesichhaker C, Bohme R, Sixt B, Bruck C, Schneider C, Schulz E. Dialectical behavioral therapy for adolescents (DBT-A): a clinical trial for patients with suicidal and self-injurious behavior and borderline symptoms with one-year follow-up. *Child Adol Psych Ment Health*. 2011;5:3.

*Graham-Bermann SA, Lynch S, Banyard V, Halabu H. Community-based intervention for children exposed to intimate partner violence: an efficacy trial. *J Consult Clin Psychol*. 2007;75:199–209.

Green JM, Wood AJ, Kerfoot MJ, et al. Group therapy for adolescents with repeated self harm: randomized controlled trial with economic evaluation. *BMJ*. 2011;342:d682.

Haen C. Rebuilding security: group therapy with children affected by September 11. *Int J Group Psychother*. 2005;55:391–414.

*Heiman ML, Ettin MF. Harnessing the power of the group for latency-aged sexual abuse

victims. *Int J Group Psychother*. 2001;51:265–282.

Heinicke BE, Paxton SJ, McLean SA, Wertheim EH. Internet-delivered targeted group intervention for body dissatisfaction and disordered eating in adolescent girls: a randomized controlled trial. *J Abnorm Child Psychol*. 2007;35:379–391.

Jerram H, Raeburn J, Stewart A. The strong parents–strong children programme: parental support in serious and chronic child illness. *N Z Med J*. 2005;118:U1700.

Johnson DL, Gold SR. An empirical approach to issues of selection and evaluation in group therapy. *Int J Group Psychother*. 1971;21:456–469.

Kessler D, Lewis G, Kaur S, et al. Therapist-delivered internet psychotherapy for depression in primary care: a randomized controlled trial. *Lancet*. 2009;374:628–634.

La Greca AM, Ehrenreich-May J, Mufson L, Chan S. Preventing adolescent social anxiety and depression and reducing peer victimization: intervention development and open trial. *Child Youth Care Forum*. 2016;45(6):905–926.

Liang-Jen Wang, Shing-Fang Lu, Mian-Yoon Chong, et al. A family-oriented therapy program for youths with substance abuse: long-term outcomes related to relapse and academic or social status. *Neuropsychiatr Dis Treat*. 2016;12:699–706.

McConachie H, McLaughlin E, Grahame V, et al. Group therapy for anxiety in children with autism spectrum disorder. *Autism*. 2014;8:723–732.

Menveille EJ, Tuerk C. A support group for parents of gender-nonconforming boys. *J Am Acad Child Adolesc Psychiatry*. 2002;41:1010–1013.

Miffitt LA. State of the science: group therapy interventions for sexually abused children. *Arch Psychiatr Nurs*. 2014;28:174–179.

Nowicka P, Savoye M, Fisher PA. Which psychological method is most effective for group treatment? *Int J Pediatr Obes*. 2011;6:70–73.

*Ollendick TH, King NJ. Empirically supported treatments for children and adolescents. In: Kendall PC, ed. *Child and Adolescent Psychotherapy*. New York: Guilford Press; 2001.

Pfeffer CR, Jiang H, Kakuma T, Hwang J, Metsch M. Group intervention for children bereaved by the suicide of a relative. *J Am Acad Child Adolesc Psychiatry*. 2002;41:505–513.

*Pollock KM, Kymissis P. The future of adolescent group therapy: an analysis of historical trends and current momentum. *J Child Adolesc Group Ther*. 2001;11:3–11.

Rosenberg M. Children with gender identity issues and their parents in individual and group treatment. *J Am Acad Child Adolesc Psychiatry*. 2001;41:619–621.

Scholten L, Willemen AM, Last BF, Maurice-Stam H, et al. Efficacy of psychosocial group intervention for children with chronic illness and their parents. *Pediatrics*. 2013;131:1196–1203.

Silverman WK, Kurtines WM, Ginsburg GS, Weems CF, Lumpkin PW. Treating anxiety disorders in children with group cognitive-behavioral therapy: a randomized clinical trial. *J Consult Clin Psychol*. 1999;67:995–1003.

Stallard P, Phillips R, Montgomery AA, et al. A cluster randomized controlled trial to determine the clinical effectiveness and cost-effectiveness of classroom-based cognitive-behavioral therapy (CBT) in reducing symptoms of depression in high-risk adolescents. *Health Tech Assess*. 2013;17(47):vii–xvii, 1–109.

Straub J, Metzger CD, Plener PL, Koelch MG, Groen G, Abler B. Successful group psychotherapy

of depression in adolescents alters fronto-limbic resting-state connectivity. *J Affect Disord*. 2017;209:135–139.

Straub J, Sproeber N, Plener PL, Fegert JM, Bonenberger M, Koelch MG. A brief cognitive-behavioural group therapy programme for the treatment of depression in adolescent outpatients: a pilot study. *Child Adolesc Psychiatry Ment Health*. 2014;8:9.

Thienemann M, Moore P, Tompkins K. A parent-only group intervention for children with anxiety disorders: a multiple-baseline design study. *J Am Acad Child Adolesc Psychiatry*. 2006;45:37–46.

Tol WA, Komproe IH, Susanty D, Jordans MJ, Macy RD, De Jong JT. School-based mental health intervention for children affected by political violence in Indonesia: a cluster randomized trial. *JAMA*. 2008;300:655–662.

van der Zanden RA, Kramer JJ, Cuijpers P. Effectiveness of an online group course for adolescents and young adults with depressive symptoms: study protocol for a randomized controlled trial. *Trials*. 2011;12:196.

Webster-Stratton C, Reid J, Hammond M. Social skills and problem-solving training for children with early-onset conduct problems: who benefits? *J Child Psychol Psychiatry*. 2001;42:943–952.

Wood A, Trainor G, Rothwell J, Moore A, Harrington R. Randomized trial of group therapy for repeated deliberate self-harm in adolescents. *J Am Acad Child Adolesc Psychiatry*. 2001;40:1246–1253.

*Yalom ID. *Inpatient Group Psychotherapy*. New York: Basic Books; 1983.

▲ 54.5 Family Therapy

JOHN SARGENT, M.D., AND NEHA SHARMA, D.O.

Family therapy represents an organized, well-developed, and variable set of interactions with family members that addresses mental health and substance use problems that affect them. It also has been used to assist with recognized relationship problems, including communication difficulties and difficulties occurring within spouse and parent–child relationships. (In this section, "parent" refers to any adult with custody of the child and responsibility for the child's well-being and development.)

These interventions take advantage of the importance of family relationships to individuals and the willingness of individuals to alter their behavior and thinking to help family members. Family therapy theory and practice are based on a range of ideas and ideologies (Table 54.5–1) and oriented to respond to a wide range of problems. The therapist's responsibility is to appreciate the nature of the problem, develop a shared appreciation of the problem with family members, and collaborate with the family to develop the best overall adaptation of family members, including ameliorating the manifestations and effects of the problem.

Family interventions address problems that have neurophysiological

and genetic elements and have developed in relation to unique environmental occurrences. Interactional responses often naturally occur that reinforce the problem. Family members are also emotionally affected by their relatives' difficulties, leading to a further mismatch between family members' needs and others' behavior. A child may be identified as the patient while the family struggles with a poor marital relationship or with parental mental health concerns. Family therapy is much more interested in the solutions to problems and improvement in the overall well-being and functioning of the family than in the etiology of the problem. This leads to a future-oriented focus of treatment and also suggests that assessment is carried out in a manner that is engaging and therapeutic. Often, family members are frustrated and frightened, and often their interactions are unproductive and repetitive, reinforcing and maintaining the problems. Therapy occurs through the interventions of the therapist and the relationships he or she develops with family members and with the family as a whole. The therapist's behaviors are intentional and carefully chosen and carried out to facilitate engagement, mutual care, support, and a resolution of problems that the family is experiencing. The therapist also appreciates that improving family relationships and interactions often leads to improved overall functioning, which can become mutually reinforcing and self-perpetuating.

Table 54.5–1.
Family Tasks as Child Development Unfolds

Infancy
Ensuring parental availability
Ensuring resources to enable parental caretaking
Providing support to parents when stressed
Providing education about infants and parenting
Encouraging balance between jobs of adult and parent, especially during times of transition
Preschool Children
Enabling parents to respond to separation/individuation efforts of the child
Encouraging parental consistency in discipline and limit setting
Supporting positive parent–child connections based on individual nature of the child and the child's competencies
Encouraging play, use of fantasy, and creativity in the interactions between parents and child
Encouraging child's relating ability with other children and with other caretaking adults
School-Age Children
Ensuring parental consistency in understanding, accepting, and being available for the child
Developing rule-based behavioral expectations and consequences for the child
Encouraging the child's competent engagement with school, friendships, and community activities
Gradually assisting the family in developing a consistent, positive, and hopeful appreciation of the child
Encouraging the child's independent perception of and participation in family life
Adolescence
Assisting the parents with adolescent problems that require rules, consequences, and consistency
Encouraging parental figures in assuming a supervisory role with children, with parents willing to allow for experimentation and learning, although also establishing essential rules

Encouraging the adolescent's voice and role in family life, as well as his or her competency and involvement in extrafamilial experiences
Assisting the parents in encouraging self-direction, self-definition, integrity, and values in adolescent children
Assisting parents with refocusing on important aspects of their lives as they separate from their children

This section focuses on the use of family assessment, family interventions, and family therapy to address the problems that families have with their children's development (Table 54.5–1), emotions, cognitions, and behavior. It is based on the premise that the practice of child and adolescent psychiatry always involves interacting with and understanding family interactions and functioning, and always involves engaging the family in the treatment and monitoring process (Table 54.5–2). The section includes case examples demonstrating the range of applications of family therapy, as well as the various forms that family treatment may take. Tables reviewing these significant models of family therapy (Table 54.5–3), the theoretical foundations of family therapy (Table 54.5–4), specific treatment approaches for common child mental health problems (Table 54.5–5), and unique family treatment approaches to differences in family composition, family history, and the culture and ethnicity of the family (Tables 54.5–6 and 54.5–7) are included.

Table 54.5–2.
Assumptions of Family Treatment

Family therapy is an integral part of the practice of child and adolescent psychiatry.
Child and adolescent emotional and behavioral difficulties always affect family life and relationships.
Family relationships always affect child development, emotional status, and behavioral capacity.
Adaptation of the child and adolescent with psychopathology to the family is always a goal of treatment and requires change on the part of the family, accommodation by the child when possible, and consideration by treating professionals.
The family can always be a source of understanding, support, and healing for children and adolescents.
The family should agree to, participate in, and assist in directing all mental health interventions at all levels of care.
Family involvement in all treatment should be defined and structured to achieve family goals and facilitate achieving goals for the child with psychiatric difficulties.
Various therapeutic interventions and classes of interventions can be used to achieve these goals.

Over the last 20 years, manualized family therapy treatments have been developed and identified as evidence-based treatments through randomized, controlled trials. These treatment approaches have been developed for specific child and adolescent disorders, such as traumatic stress, substance abuse, anorexia nervosa, and conduct disorder.

PROCESS OF FAMILY THERAPY

Engagement

The key to successful family treatment is successful engagement of the therapist with the family and with family members. This process requires that the therapist possess the skills of connecting with more than one person at the same time and making each person feel understood, comfortable, and cared about. Engagement conveys the therapist's hope for the family's successful future and his or her belief that with effort, patience, and mutual care and concern changes can occur and that a successful future can be realized. Family therapy in whatever fashion is always a creative act, an act of vision. The therapist's connection with the family and its members is based on his or her assertion of the meaning and value of this family. Through the acceptance inherent in engagement, the therapist works to create a context of safety that promotes greater flexibility in the interactions of family members and greater possibility that the situation can be accepted and improved. In addition to warmth, curiosity, hope, and vision, the therapist must be able to hold and act on opposite ideas: that the family possesses strengths and has difficulties, that the family is doing the best it can and that the family needs to change, and that the family and its situation are unique and also resemble other families in their connections to each other, their concern for each other, and their wish that every other family member have the best life possible.

The therapist also engages with the family as an entity. By acting in a manner that reflects the therapist's belief in the integrity of the family, family members can begin to relax, see each other differently, and behave in a manner that reflects their interest in membership in their family and their commitment to its well-being. The therapist asks about the family's history, culture, rituals, and beliefs; respects what he or she hears; and treats the family as an entity in which he or she believes in and cares about. This helps the family members to come together, work together, and change the ways in which they treat each other. In doing this, family members begin to understand that relationships can improve, conflicts can be dissolved, and mutual understanding and mutual support can develop. Through successful engagement with individual family members and with the family as a whole, the therapist creates a solid foundation for family change and symptomatic improvement.

Table 54.5–3.
Varieties of Family Therapy

Psychodynamic

Assumptions and Major Concepts

- Family relationships are based on previous parent–child relationships of each parent.
- Difficulties arise from developmental arrest, current interactions and projections, and current stressors.
- Improvements develop through the process of family members gaining insight into problematic relationships from the past.
- Therapy consists in shared exploration of past relationships and current distortions and role

confusions.
Therapy also encourages the development of mutual appreciation of the needs and limits of family members.

Major Theoreticians and Practitioners
Nathan Ackerman
James Framo
Ivan Boszamenyi-Nagy

Experiential/Existential

Assumptions and Major Concepts
The family is a developing organism, and the goal of therapy is to encourage the growth of the family and its members through the process of therapy.
This process is best supported by encouraging expressiveness, openness, and communication within the family throughout the therapeutic encounter.
The family is responsible for its solutions; the therapist is a facilitator who encourages growth and confronts stagnation.

Major Theoreticians and Practitioners
Carl Whitaker
Virginia Satir

Intergenerational/Bowenian

Assumptions and Major Concepts
The family is an emotional relationship system, with the goal of relationships to promote engagement, foster differentiation, and avoid fusion.
In conflictual situations, engaging a third party (triangulation) stabilizes the family system.
Individuals marry partners at the same level of differentiation, leading, at low levels of differentiation in married couples, to severe personal and relationship difficulties.
Therapy is aimed at clarifying current relationships, promoting individuation, resolving current conflictual interactions, and working through intergenerational and introjected difficulties and patterns of behavior.

Major Theoretician and Practitioner
Murray Bowen

Structural

Assumptions and Major Concepts
Focused on current organization of the family as demonstrated through repeated patterns of interaction.
Particularly attentive to how the family's organization promotes or inhibits task performance, especially child rearing.
Analysis of interactions allows the therapist to determine the nature (distance vs. closeness) and the flexibility of family relationships in relation to task performance.
The therapist promotes interactions in sessions to determine family relationships and then intervenes in those interactions to introduce greater flexibility, alternative patterns and structures, and more effective task performance.
Enmeshment (ineffective closeness) and disengagement (excessive distance), as well as rapid fluctuations between extremes, are manifestations of dysfunctional family relationships when associated with significant symptoms in a family member.
This form of family therapy was developed to respond to child and adolescent emotional and behavioral problems and is particularly effective in those situations.
The therapist is both the director of therapeutic action and participant in the process of therapy.

Major Theoretician and Practitioner
Salvador Minuchin

Systemic/Strategic

Assumptions and Major Concepts
Symptomatic behavior is embedded within the patterns of communication and relationships in a family.
Ineffective attempts at solutions to problematic behavior become part of the pattern also and

thus part of the problem.
The therapist is responsible for developing and prescribing tasks and behaviors that disrupt usual patterns of interaction while strengthening other, more productive patterns.
The therapist maintains a stance of neutrality and distance with respect to the family as whole and individual members.
Specific methods of questioning designed to elicit information about family relationships (circular questioning) are used; posing hypotheses, reframing, externalizing blame, and the use of positive connotation are common techniques.

Major Theoreticians and Practitioners
Jay Haley
Cloé Madanes
Mara Selvini-Palazzoli and colleagues (Milan group)

Cognitive-behavioral

Assumptions and Major Concepts
Concepts of learning theory—conditional contingency, reinforcement, extinction—explain problematic behavior and identify therapeutic interventions.
Sessions are entirely problem- and solution-focused in the present.
Skills training is emphasized throughout treatment; homework, record keeping, and between-sessions tasks are used regularly.
Therapy is aimed at changing both cognitions and behavior and uses new cognitions among family members to direct and evaluate behaviors.
Particularly useful for managing mild to moderate disruptive behaviors in children and adolescents, focusing treatment on parent skills training.

Major Theoreticians and Practitioners
Gerald Patterson
James Alexander

Table 54.5–4. The Theoretical Foundations of Family Therapy

General Systems Theory

The whole is greater than the sum of the parts.
Actions that affect one part of a system affect the whole.
Each part of a system is in a constantly changing relationship with all other parts.

Cybernetics

Interaction among family members is circular and continuous, with each behavior being the effect of one behavior and the cause of another.
Circular causality encourages attention to pattern of interactions and feedback between behaviors and relationships.

Communication Theory

All behavior is communication (both actions and words); behaviors define the nature of the relationships among family members.

Structural Theory

Relationships are structured in a hierarchical, potentially developmentally oriented fashion.
The structure of a family can only be determined by noting the patterns of interaction among family members.
Structures are functional or adequate insofar as they provide for the achievement of family goals.

Social Constructionist Theory

Each individual constructs or makes sense of his or her experience from his or her perspective.
Therefore, many different interpretations of a set of circumstances may be simultaneously true.
Families, therefore, must and do develop shared constructions of their situation.
The constructions may become unsatisfactory when problems or changes require interaction with systems (health care, school, extended family, community agency, etc.) that have alternative

constructions of events.

Psychodynamic Theory

Early experience is formative in defining who people are.

The capacity for projection and projective identification can seriously inhibit the flexibility, resilience, and growth of relationships.

Attachment Theory

Affect regulation occurs within close relationships.

Secure attachment facilitates social relationships and the ability to know and communicate about mental states.

Safety

Throughout the engagement and assessment process the therapist's first responsibility is to ensure the safety of the child, the family members, and others in the community, as well as his or her safety. In some families aggressiveness is a part of their interactions, emotions may run high when disappointments, hurts, and conflicts have occurred, and family relationships may feel and at times be unsafe. Thus the therapist must be attentive to signs of the possibility of violence, aggressiveness, abuse, and suicidal behavior. If there is concern about safety, consideration of hospitalization or involvement of human services or the justice system must be considered. The therapist's role is to work with the family to decide together on the level of care necessary to promote the safety of all while also beginning treatment in a manner that ensures the safety and connection of all. The therapist creates a safe space where family members can speak to and listen to each other and where hurts and disappointments can be spoken about and reconnections can be achieved. The therapist can establish ground rules for sessions and, if concerned, meet with family members individually to ensure a commitment to self-control and mutual respect in treatment and ultimately throughout family life. The quality of the therapist's engagement with the family and his or her presence, integrity, and his belief in his or her vision of successful resolution of problems are crucial in leading the family to the changes needed to ensure nonviolent family life.

Assessment

Family therapy is based on an ongoing assessment of the child's problem and family life. It is important for the therapist to recognize that the assessment is part of the therapeutic process. Assessment is completed in an interactive manner that is based on the therapist's interest in the family's strengths and individual capacities and interests, as well as the therapist's professional responsibility to appreciate the family's problems. The therapist, connected to the family in the treatment system, elicits a discussion about the questions the family has about their child and the concerns and issues the family experiences in their life.

One can view the process of assessment in child-oriented family therapy

as the iterative development of an imagined video of the child and family at home, their relationship with the child's school, and their presence in their community. This understanding appreciates not only the child's symptoms and abilities, but also the interactions within which those symptoms and strengths are embedded. Thus, symptoms—either those that are spontaneous or those that are reactive—are viewed within the interactions that precede them and those that follow them. One way to understand the patterns in which symptoms are embedded is to encourage interaction among family members in sessions and to be attentive to spontaneous interactions as they occur in sessions. Understanding these patterns provides the therapist with many possibilities for discussion and intervention.

The family and the therapist can also decide together other questions to be answered through the treatment process. Intellectual concerns, educational abilities, and vocational aptitudes all can be evaluated if questions exist. Diagnostic questions and the use of psychopharmacology also can be discussed in a collaborative fashion with the family. The guiding issues are always: What are your questions? What are your goals? What of these alternatives do you want to do? The key to understanding this process is to appreciate several key principles:

1. Assessment results are impressions and not judgments.
2. All assessments of the family are hypotheses based on limited data and thus are subject to change.
3. Because interactions can change and are largely dependent on context, the therapist holds the belief that if the context changes, the interactions can change.
4. Any intervention represents a question: will this intervention be successful in achieving the goal? If not, a different intervention is required.

The process of developing a shared opinion of both the child and the family, including an appreciation of strengths and difficulties among members of the treatment system, is another opportunity for therapeutic intervention. There may be opportunities to discuss and resolve disagreements between caregivers about the child's capacities and problems. There may be opportunities, depending on developmental ability, for the child to participate in this discussion. The therapist supports the child's voice and expects the child to participate responsibly. As this process proceeds, the family's ability to negotiate effectively can be supported and reinforced by the therapist.

Table 54.5–5.
Family-Oriented Treatment Approaches for Specific Clinical Problems

Ensure appropriate diagnostic evaluation, accurate diagnosis, and adequate antipsychotic treatment.
Psychoeducational treatment model is very helpful, providing education about the disorder, outlining family roles in treatment, and providing guidelines about how to respond to crisis and specific problems.
Efforts at reducing blaming and intensity of affective expression within the family are often helpful.
Ensuring family collaboration with the treatment approach, especially concerning communication with professionals, medication, vocational training, and work and placement, is extremely helpful.
Parent support is an important aspect of treatment.

Depression

Ensure that all involved recognize and accept the child's depressed mood.
The therapeutic orientation should include attention to building attachments and connections, reducing personalized criticism, and building support among family members.
Building individual competence is essential.
Cognitive-behavioral strategies are most effective.
Setting limits to deal with acting out, personal irritability, and poor academic performance is often helpful.
Special attention should be paid to parental psychiatric problems.
Psychopharmacology should be integrated within the therapeutic approach.

Anxiety Disorders and Phobias

Collaboration in treatment reduces the power of the child's fears and vulnerability.
Cognitive-behavioral approaches with the family are helpful.
Agreement about the expectations for the child, especially regarding fears, is necessary.
Distraction, relaxation, social skills assistance, and focus on building peer relationships all are helpful for the entire family to focus on.
Recognition of temperament and the possibility of recurrence of worries during stressful periods often aid long-term outcome for the family.

Disruptive Disorders

Attention-Deficit/Hyperactivity Disorder (ADHD)
Agreement about diagnosis is essential.
Education concerning the various aspects of ADHD is helpful.
Family involvement in measuring and monitoring the results of behavioral interventions and psychopharmacology is helpful.
Parent support groups can help long-term adaptation.
Specific plans for impulsivity, reactivity, and aggressiveness are helpful.
Independence, self-regulation, self-organization, and personal responsibility are goals of treatment.
Special attention should be directed to developing the strengths, competencies, and self-esteem of the ADHD child and the family as a whole.
Conduct Disorder
Collaboration among parents and all involved professionals and agencies is essential.
Responsibility and accountability for the child are primary goals of treatment.
Parameters of treatment must be explicitly stated.
Behavioral approaches highlighting consistency, expectable consequences, and predictable responses to transgressions are the cornerstone of treatment.
Plans must be made for responses to violence and other serious problems.
Impulsivity in treatment should be avoided; treatment must be consistent, continuous, developmentally oriented, and focused on building successful adaptation of the child, as well as on reduced behavior problems.
Medication is to be used for specific target symptoms, and its use and effectiveness must be supported by all participants in treatment.

Substance Abuse

Agreement about avoiding substance use among all involved adults in necessary.
Family tolerance for affective expression usually needs to be expanded.
The family must also attend to comorbid problems, depression, learning problems, ADHD, and so forth.

The role of consequences and limits for substance use must be expressed clearly.
Parental substance use also must be addressed in treatment.
Family participation in mutual support groups (Narcotics Anonymous, Alcoholics Anonymous) is often very helpful.
Peer problems, legal difficulties, and relapses should be addressed directly with the family.

Eating Disorders

Structural techniques build family cohesiveness and a consistent approach to the disturbed eating behavior.
Simultaneous support for individuation and connection is necessary.
Cognitive-behavioral techniques assist in establishing consistency in family relationships, as well as in responses to eating disorder symptoms.
Family therapy offers an opportunity for dealing with gender issues, as well as concerns associated with control and perfectionism.
The age of the eating-disordered individual defines how treatment is structured; early adolescents respond best to family-oriented treatment; older adolescents and young adults respond better to family therapy approaches that they arrange and manage in conjunction with simultaneous individual therapy.

Abuse (Physical or Sexual)

Safety of the child is paramount.
Parents should collaborate with the therapist and social service agency staff.
A clear role for parents in the treatment process should be outlined, with specific expectations and consequences defined.
Parent training and skills building are helpful, especially if the unique contributions of the child to the family situation are highlighted and the parents have the opportunity to practice responses and receive praise and support when they demonstrate competence.
Honesty, clarity, and directness often reduce defensiveness and evasiveness on the part of family members.
As much as possible, the child should have a developmentally appropriate, active role in treatment.
Encouragement of verbal expression can be helpful in affect regulation and maintenance of consistent responses to behaviors among family members.

Academic or Learning Problems

The child's learning or academic problem must be defined to the satisfaction of parents and school professionals.
Parents assist with developing questions to be addressed through psychological, psychiatric, and educational evaluations.
Roles and responsibilities of parents, school professionals, mental health professionals, and the child should be clearly identified.
Specific educational and school-based goals should be set; these should be development of academic competency, self-esteem, and greater connection between the child and school.
Steady progress through school marks development and achievement and builds greater maturity.
Other difficulties influencing learning (e.g., depression, anxiety, ADHD, and other disorders) need to be addressed simultaneously.

Table 54.5–6. Variations in Family Composition

Single-Parent Families

Recognize the strength, responsibility, burden, and stresses of the custodial parent.
Support the independent competence of both the parent and the child.
Encourage the parent's use of alternative caretakers and other modes of personal connection and support.
Encourage the child's exploration of his or her relationship with the noncustodial parent.
Assist with transitions that may occur with new relationships that the parent might develop.

Divorced Families

- Encourage parental collaboration in divorcing and parenting processes.
- Assist the independent emotional reactions of each family member to the separation and divorce process, including the recognition that this is a highly painful experience for children.
- Encourage the development of effective separate relationships between each parent and each child.
- Build strengths of sibling relationships.
- Assist the competency and expression of each child.
- Assist parents in collaborating in parenting tasks.

Grandparent- or Relative-Headed Families

- Recognize the effort and strength of the child-rearing adult.
- Encourage the child's emotional adaptation to not living with his or her biological parents.
- Encourage effective limit setting and adult affiliation with the child in these families.
- Emphasize the child's competency and extrafamilial engagement and accomplishment.
- Support consistency in involvement of biological parents to the greatest extent possible.

Adoptive Families

- Recognize that the child's problems may be associated with preadoptive phenomena.
- Support with sensitivity the parents' emotional reaction to their child's difficulties.
- Possible special attention to the child's background and special attributes.
- Build connections among family members while dealing with the losses and separations associated with adoption.

Families with Same-Sex Parents

- Similar strategy as adoptive families
- Support parents to feel empowered and effective in their parental role.
- Guide parents to differentiate between their personal childhood experience and their child's experience.
- Build connections in the community such that family members feel welcomed and accepted.

If other agencies or systems are involved with the family, especially in response to the child's problem, their representatives also will need to participate in the agreements about strengths and problems. This can be difficult, especially if the family has a history of conflict with these systems. The family may believe that a system with which they must be involved, such as the child's school, is treating the child ineffectively. The family will need the therapist's encouragement and potentially the therapist's coaching to be able to come to a collaborative approach with school personnel so that the child's education proceeds successfully. Similar negotiations will need to occur with the child protective system, the juvenile justice system, after-school programs, coaches, extended family, and church leaders, depending on who is involved with the child and what the child's needs, abilities, and problems are. At times the therapist will need to convene meetings of the family and those involved. The therapist, with the family's permission, may need to speak individually with the teacher, probation officer, and so on. These individuals may need to change their impressions of the child or family, and at times they may need to change the way in which they interact with the family. Supporting these new impressions is part of the therapist's responsibility for the success of his or her interventions. These changes become part of the therapeutic effort and goals for the treatment.

Table 54.5–7.
Family Therapy Approaches: The Effects of Cultural Differences

Cultural Difference and Family Worldview

Worldview defines the relationship of a family to people, institutions, and objects.

Cultural experience affects a family's worldview.

Race, religion, and traditional position within the power structure also affect a family's worldview.

These factors influence family relationships, family roles, and expectations of parenting, child rearing, and child development.

Examples of Cultural Considerations in Family Treatment

White families: focus on individuals, use verbal expression, and distinguish between physical and mental factors.

African-American families: focus on actions and behavior; realize that extended family is important, as are religion and church; also realize that history of oppression and racism affects relationships with institutions.

Hispanic families: acknowledge bilingual background with intrafamilial differences in acculturation and language use; religion and extended family are important; respect for the family's heritage and background is very helpful in ensuring engagement in treatment.

Asian-American families: recognize specific national origin; family, privacy, and saving face are central to family life; recognize intrafamilial differences in acculturation and language use.

Attending to Cultural Issues in Family Therapy

Basic Values

Any family is a microculture.

The family's culture is necessarily a synthesis and amalgam of cultures.

Problems develop when the family experiences conflict internally between the cultures of the extended family or with external cultures, institutions, or systems.

The family is responsible for defining itself and its culture and for developing the most successful adaptation with other cultures.

The therapist's responsibility is to support and facilitate that adaptation, recognizing (1) that there are many ways to be a family and raise children that are healthy and (2) that there may be issues that raise ethical concerns for the therapist, and in those situations further knowledge of the family's culture and consultation or supervision may be helpful.

Interactive Approaches

The therapist must assume a stance of respect for and acceptance of the family as it is.

The therapist should ask about the meaning of requesting mental health assistance for the family.

Honest curiosity expressed in open questioning about how the family operates and how the family's cultural background influences its operation is essential.

The therapist should refrain from judging the family, but rather should ask about how it achieves its goal.

The therapist should avoid behaviors based on assumptions about a particular cultural background.

The therapist should inquire about child-rearing practices in the family and responses to academic, emotional, behavioral, and social problems in the child as they reflect cultural beliefs and cultural expectations.

Clinical Approaches

Specific responses to the child's symptoms and assistance with parenting skill need to be embedded within the therapist's and the family's understanding of cultural beliefs and expectations.

Negotiation of the therapist's role and behavior should be based on the family's sense of their culture's expectations of professional behavior.

The therapist can help the family to understand their cultural background as a method of helping the family to define and understand themselves and clarify their approach to therapy.

Explorations of culture assist the family in assuming direction and responsibility for treatment.

Discussions about respective cultural backgrounds and what those backgrounds mean can assist in conflict resolution and enhance the development of the children.

The therapist can assist the family in making accommodating measures with schools and other

institutions based on cultural considerations.

Engaging Families in All Levels of Care

Family members should be actively involved in decisions about treatment programs, levels of care and care delivered, and inception and discontinuation of treatment. This expectation helps the parents or caretakers develop monitoring, supervisory, and organizational skills that they might be lacking. Because the parents possess decision-making power, they must learn to understand their child, make the child's behavior make sense, and begin the process of accepting the child's differences. This can accentuate feelings of support and understanding, which may have been undermined by their responses to symptoms and their consequences. This is especially true in the case of families with children with disruptive disorders, in which difficulties in school or with the justice system may have felt hurtful and shaming.

The journey to get out-of-home services often results in parents feeling ineffective and disempowered. The purpose of family intervention is to empower parents to conduct effective parenting. Thus, family involvement is essential for children being treated in treatment programs such as hospitalization, day treatment, or residential treatment. For the home environment to reinforce positive behaviors and not reinforce negative behaviors, it is necessary for program clinicians and staff to be attentive to the interactive sequences of which the symptoms were a part and intervene with the family to alter those sequences so that the home environment is different when the child is discharged or the program ends. If discipline is problematic, parents can be consulted when program rules are broken and consequences are required. Parents can be informed by the child of successful experiences so that their praise is added to that of the staff. These successful experiences can translate into hope for change for parents. Parents can be expected to meet regularly to have a meaningful voice in treatment review, learning to develop realistic expectations of their child based on how the child is able to function rather than based on everyone's wishes about how he or she might function. It can be difficult at times to share power with parents, but with patience the staff can help the parents to feel empowered throughout the treatment program. If passes home are part of the treatment program, activities can be planned for the entire family that emphasize practice with new interactions at home prior to discharge.

Ensuring a Balance between Support and Limits in Parent–Child Interactions

One of the most challenging aspects of child-focused family therapy is developing an effective balance between, on the one hand, enhancing parent–child warmth, support, and the interactions that reflect secure

parent–child attachment and mutual respect, and on the other, working to establish effective rules, supervision, and consequences for misbehavior and rule breaking. This often requires skill, focused attention, and persistence on the part of the therapist. Because parental warmth forms the foundation of the mutual respect needed for effective discipline, the therapist's first task is to build that warmth. This often occurs at a time when parents feel betrayed, helpless, angry, and victimized. At this point it is the therapist's belief in the parents, in their love for their child, and in the possibility that they might help their child grow up safely and successfully that matters most. The therapist models warmth in his or her relationship with the parents. Compassionate understanding characterizes these interactions. Interactions that caringly convey the therapist's belief that the parents care for their child and believe in the child's future characterize the therapy. This must occur even if the child expresses skepticism or disbelief about the parents caring for him or her. As trust develops between the therapist and the parents, who know that the therapist believes that they care truly about their children, a wide range of interventions is possible. These include establishing parent interactions with the child aimed at understanding the child's disagreements with his or her parents in a nondefensive way and in a dispassionate fashion while establishing expectations for their child's behavior as well as consistent consequences for failing to meet expectations. In this process, the therapist helps the parents to deal with their disappointments with their child so that consequences are not vindictive and helps the parents appreciate that their child could want their approval and could wish to earn their regard. In this approach, parents understand and support their child's emotions, disappointment, frustrations, and anger while also holding firm expectations, limits, and consequences. In a sense, parents are feeling one way—frustrated and angry—and acting another way—calm and firm—while allowing the child to act one way—frustrated and angry—and expecting the child to feel another way—accepted and respected even while adhering to parental limits and appreciating parental involvement and concern.

Parents often want advice about establishing discipline and limits, especially in situations in which the child has a disruptive behavior disorder and defies parental rules and limits. This often represents a challenging therapeutic concern. Often in these situations there is a difference in approach between parental figures. Everyone is exhausted, hurt, frustrated, and angry. The parents will frequently claim that consequences do not work. Disrespect and defeat characterize parent–child interactions. In these situations the therapist might choose to meet alone with parents to help them to begin to work together, as well as to develop strategies to reinstitute effective limits. The therapist can help parents to prioritize their expectations and plan to establish one rule at a time. In the process the therapist can also help the parents to develop realistic

expectations of their efforts with their child. The parents can receive validation of the experience of frustration and their sense of being betrayed by their child. As long as the parents hold on to their experience of being hurt by their child, it is very unlikely that they will notice small attempts at improvement on their child's part. They also will have difficulty holding to their expectations if the child becomes angry or has a temper outburst when a limit is imposed. In such frustrating situations, guided by the therapist, a parent can gain support from his/her partner that could result in decreased emotional intensity and improved ability to maintain positive expectations. Another frequent occurrence that can be discussed with the parents alone is the possibility that the child will act in a totally unaffected way when the limit is imposed, as if to say that the parents have no effect on the child. By discussing these possible scenarios, the therapist can help the parents prepare to institute appropriate limits in a planned and controlled manner. In these sessions the parents can practice talking with their child and maintaining firmness even if the child is particularly challenging. The therapist may also choose to meet alone with the child to see whether she or he has concerns that the therapy might resolve. Defiant, disruptive children and adolescents often want to be trusted and understood, and to have autonomy.

The therapist can help the child to receive parental approval if the child performs acceptably and gains autonomy in relation to his or her ability to act responsibly. When both parents and child are prepared, a session with all of them together can be held outlining the rules, expectations, and consequences. The child has the opportunity to ask for what he or she wants, and the therapist can help the parents to promise the child the desired autonomy if the child abides by the rules and accepts the consequences. The consequences should be simple, easy to undergo, brief, and have some relationship to the transgression. The key to consequences is that they are meaningful but not onerous. Losing access to a cell phone or computer for one evening is a meaningful but transient consequence. If the parents and child form a committed relationship with the therapist, the therapist can then, with patience and time, bring them back together in an initially tenuous but gradually caring manner. Building understanding and trust, establishing a realistic plan for limits, and noticing small successes provides a pattern of improvement that all can recognize and believe in.

FAMILY THERAPY FOR SPECIFIC PROBLEMS

Attention-Deficit/Hyperactivity Disorder

Families with a child with attention-deficit/hyperactivity disorder (ADHD) need to learn about the disorder, understand appropriate treatment, and support the development and functioning of their child. This requires that parents agree about the presence of the problem, agree to seek the help of a

physician, and agree to have their child take medication if prescribed. It is also important for the parents to learn to be consistent with behavioral expectations and to appreciate and reinforce the child's strengths. Often, prior to diagnosis, the child may have experienced failure at school and rejection by peers and marked negative interactions at home. Family therapy after diagnosis of ADHD may be needed to help the family assist the child with these previous difficulties, and also to help the child to be more successful at school and with friends and to feel consistently accepted at home. Comorbidity with learning disability is common, and the family may need assistance with ensuring appropriate education for the child. Transitions, including moving and changing schools, can be challenging. The family may also need assistance during the child's adolescence, particularly with impulse control, exposure to drugs and alcohol, and self-management as the child becomes a young adult. A chronic illness framework often is extremely helpful for families with children with ADHD.

Jason was a 7-year-old boy who attended the second grade and who had been noted by his teacher to be extremely active in class and to have significant difficulty concentrating and following class rules. He especially had difficulty during unstructured time in school, both in the cafeteria at lunchtime and on the playground at recess.

When his teacher checked with his previous teachers from kindergarten and first grade, they stated that Jason had difficulties with his attention and activity level in both of those grades. Despite these concerns he was quite bright and had been able successfully to complete his grades and pass to the second grade. Jason's teacher invited his mother to come to school for a meeting after having sent home reports concerning his overactive behavior and his inability to finish his work.

At the meeting with Jason's mother, his teacher spoke about her concerns and suggested that Jason should undergo a psychiatric evaluation and might need medication to help him concentrate and manage his activity. His mother reported that she was also concerned about how active he was at home and how difficult it was to help him get ready for school in the morning or finish his homework in the afternoon.

Jason lived at home with his mother, stepfather, and 4-year-old stepsister. His parents had divorced when Jason was 3 years old, and his mother remarried his current stepfather 1 year prior to Jason turning 7 years old. Jason saw his biological father on a regular basis and spent alternate weekends at his father's house. His father had also remarried, and Jason had an 18-month-old half-brother who lived with his father and stepmother.

Jason was seen by a child psychiatrist shortly after his mother and teacher met. The child psychiatrist believed that Jason was a child with ADHD and suggested to the mother that a trial of stimulant medication might be helpful for him. The mother was given rating scales to complete after Jason had started on the medication, along with written information about both ADHD and the stimulant that was prescribed for him.

The child psychiatrist also recommended that there be a meeting of all four adults involved in Jason's life—his mother, stepfather, stepmother, and biological father—so that they could talk together about Jason's difficulties and ensure consistent treatment. The goals of the meeting were: to encourage regular use of the stimulant medication and to help everyone to understand the need for clear and focused directions, focused attention, follow-through on homework, and consistent discipline in both households.

The family attended this meeting. The psychiatrist noted that Jason's biological father was concerned about the use of medication and asked about the risks and side effects of the medicine. Jason's stepmother was concerned that there was no way successfully to help Jason without medication. As the family talked together with the child psychiatrist, asking questions and gathering information, they came to the conclusion that the best approach was to institute

a trial of stimulants. The trial was begun, and Jason's teacher reported to the family that he was much more able to focus his attention and manage his behavior in school. He was noted to have less difficulty during unstructured times in school and was also able to do better academically than he had previously.

The child psychiatrist also suggested that the family come back for a second meeting, after Jason had been on medication for 6 weeks, to review everyone's impression of Jason's progress. The child psychiatrist recognized that the biological father and stepmother were now comfortable using medication for Jason. They agreed to follow through with the medicine and to help Jason manage his behavior better in their home.

Three years later, as Jason continued to do well in school on stimulants, during a routine medication review with the child psychiatrist, Jason's mother reported that his father was getting divorced for a second time. Jason tearfully admitted that he was very upset by this and that he was going to miss his times with his stepmother and younger half-brother.

The child psychiatrist asked to speak to the father, who agreed to come for a session with Jason and talk with him about what was happening in his family. Jason's father agreed to help him have continued phone contact with this stepmother as the father and stepmother proceeded with their plans for divorce.

Six years later, Jason was old enough to obtain a driving permit. He had continued to take stimulants daily and had done well in school. He had grown up without significant emotional and behavior troubles and was eager to learn how to drive. Jason's father and mother had continued to meet with the child psychiatrist every 4 months to review Jason's medication and to check in with how everyone was doing. They were concerned about Jason driving. Jason was sure that he would be able to drive perfectly well even though he was known to be more impulsive and more disorganized than other teens his age. The child psychiatrist helped the family to discuss the privilege of driving. They decided that they would let him drive with one of them in the car and not let him drive after dark for the first several months that he was driving. They decided not to allow him to drive with peers until he had driven successfully for several months. This negotiation was difficult for Jason, but Jason's father and mother both felt very supported by the child psychiatrist. They were comfortable with their decision that Jason could learn to drive as long as he was monitored quite closely and supervised very carefully.

Oppositional Defiant Disorder and Conduct Disorder

The primary goal of family therapy for oppositional defiant disorder and conduct disorder in one or more children is the establishment of a warm and mutually respectful relationship between parents and child, which forms the foundation of effective limits, appropriate supervision, and responsible prosocial behavior by the child. Often the family situation is complicated by any or several of a variety of family risk factors. These include parental stress, parental psychopathology or substance abuse, criminality and incarceration, poverty, marked parental preoccupation with time-consuming work, or preoccupation with other children. Overcrowding, living in a violent community, and domestic violence also occur frequently. Negative coercive cycles often occur in the interactions between parent and child, in which the child is demanding and the parent becomes controlling, leading the child to seek attention through misbehavior and leading the parent to believe that only direct control will make the child behave appropriately. The child fails to learn self-control and self-direction, and the parent becomes gradually and insidiously disempowered and negative toward the child. As the child grows, the parent increasingly becomes hopeless about the child and begins to believe

that the child will fail to grow up successfully. This results in parents feeling helpless and re-emphasizes that improvement must come from outside the family. Therefore, the parent hopes that something else will lead the child to develop responsibility and seeks placement of the child in residential treatment or in the juvenile justice system. This perspective localizes the problem in the misbehaving child, leaving the child feeling blamed and even more disconnected from his or her family.

Two forms of family therapy deal specifically with these issues. Parent–child interaction therapy is specifically designed to treat families with young (preschool) children with oppositional defiant disorder. Multisystemic therapy has demonstrated effectiveness for families with adolescents with delinquency who are involved with the juvenile justice system. Parent–child interaction therapy focuses on building the positive aspects of the relationship between the parent and young child. Through the therapist's praise, the parent relaxes and becomes more accepting and more openly caring toward the child. The parent receives the therapist's praise and care and passes that on to the child, leading to closeness, the experience of being understood and cared for, and warmth and attunement in the parent–child relationship. This also leads to improved cooperation between parent and child, greater compliance on the child's part, and, on the part of the child, much greater interest in the parent's praise and positive regard. Parent–child interaction therapy has been shown in more than 30 clinical trials to be effective for families with preschool children (between ages of 2 and 6 years) with oppositional defiant disorder. It has also frequently been requested by child protective services to improve parent–child relationships in situations in which mild physical abuse has occurred.

Multisystemic therapy has been shown to be effective in reducing recidivism and delinquency for children and adolescents in treatment trials and through follow-up. Trained therapists with very small caseloads help the parents to manage the stresses in their lives and supervise and set limits on their child. The therapist is available to the family consistently and the treatment is intensive, with three or four contacts with the family weekly over 4 months. The therapist builds family cohesion and encourages the child to identify and earn privileges by being responsive to parental direction. Rules are defined and consequences are specified early in treatment. Parents are expected to be available, observant, and consistent in monitoring their children. For this to occur, other difficulties that the family experiences need to be addressed through the therapy. This treatment is best suited to moderate delinquency and works best early in the course of delinquency.

Another form of family treatment for conduct disorder is parent management training. In this therapy, parents attend sessions without the child and learn and practice techniques of setting limits and imposing

consequences. Most important in this treatment is helping the parents plan and practice how they will withstand the child's demanding and defiant behavior and how they will respond if the child's behavior becomes worse before it gets better (as is expected). In this treatment the therapist supports a united front among adult authority figures and builds mutual support among these authorities. Parental self-control is another goal of these parent-coaching sessions. The therapist needs to ensure that the parents learn the skills necessary to set effective limits and manage their lives at the same time.

Isaac was a 4½-year-old boy who was brought to a child mental health clinic for evaluation of defiant, challenging, and difficult behavior. He had been referred by his pediatrician after his mother discussed his defiant and difficult behavior with the pediatrician during a regular checkup. The history also revealed that Isaac had been expelled from his preschool and had to be looked after by a neighbor in her house because of aggressive and difficult behavior at school.

When the clinician met with Isaac and his mother, Isaac appeared to be a well-behaved, quiet child who was able to focus his attention and play easily with the interviewer. However, it was noted that his mother was extremely negative about him, and was frustrated, angry, withdrawn, and sad. During the interview the clinician noted that Isaac's mother was exhausted, overwhelmed, and observably sad and irritable.

The more comfortably Isaac played with the clinician in his mother's presence, the more upset and withdrawn she became. She reported that Isaac's father had left before he was born and that she had an older, 8-year-old boy with developmental disability and cerebral palsy. That child was in special education classes. She lived alone with the children. She stated that she was estranged from her mother and that she worked as a cashier in a grocery store. The family lived in a small apartment in an impoverished area of the city, and the mother reported that at times it was difficult to have enough food for her, Isaac, and his brother to eat.

The clinician was struck by the poignancy of Isaac's mother's situation, by her isolation and lack of support, and by her level of exhaustion and sense of being overwhelmed. The therapist was also touched by Isaac's attention to his mother and his tendency to watch her and comfort her when she became sad during the interview. The clinician was attentive to and supportive of the mother, and during the interview she began to calm down and admitted to being depressed, sad, and lonely.

Isaac and his mother arrived for their second session on time and both seemed eager to talk. The mother had noticed that Isaac was a little bit better behaved at home, and she seemed more relaxed and comfortable.

The therapist continued to support the mother and encouraged her to engage with Isaac in a game during the session. The therapist was warm, caring, and praising of the mother's efforts with her son. The therapist also recommended that Isaac and his mother have some regular time to enjoy each other's company during the week and suggested 15 to 20 minutes every day, during which their neighbor could look after Isaac's older brother. As Isaac received more attention from his mother at home, their negative interactions decreased. He seemed to relax and become less quick to anger, both at his babysitter's house and at his mother's house.

The clinician also encouraged the mother to reconnect with her siblings and her mother, who were in the community, and to reach out to them for support and understanding. The mother noticed that she was more relaxed and less stressed and that she was more able to experience enjoyable times with her sister. She also began to communicate in a more open way with her mother.

As the sisters became more involved with Isaac's family, there were more family gatherings and more opportunities for the mother to experience some relief from her caretaking obligations. Over a period of 3 months of weekly sessions and 2 months of biweekly sessions, Isaac's mother became much more relaxed, was able to obtain babysitters for her children, and got more involved in her older son's special education programs. She became more assertive

about obtaining a raise at work and socialized with her colleagues. She spent more time with her family and her neighbors. She also resumed attending church, and Isaac became involved in the church-sponsored preschool kindergarten group.

The therapist stayed in touch with Isaac and his mother and heard from her that Isaac had successfully begun kindergarten during the next academic year. His behavior was much improved and his curiosity, desire to play and learn, and attentiveness to adults had helped him to learn and grow in kindergarten.

Joseph was referred to the intensive in-home treatment team after his discharge from the juvenile detention facility in his community. He was 15½ years old and had been placed in the detention facility after his arrest for vandalism and breaking curfew. The vandalism involved breaking windows in his school, and well as painting graffiti on several businesses in his downtown neighborhood. Joseph lived with his mother, stepfather, 12- and 19-year-old brothers, and 17-year-old sister.

When the in-home treatment therapist went to meet Joseph and his family, the mother was at home with the younger brother. Joseph was not at home. He had told his mother that he needed to visit his friends after school and that he would be home in time for the appointment. He was late and had not informed his mother.

The history the therapist obtained was that Joseph had been a very good student and a fine boy until the age of 9, when his grandfather suddenly died and his mother moved the family into the grandmother's house to look after her. Joseph had been close to his grandfather and had reacted to his death with sadness and withdrawal from the family.

His mother noted that he began to act up at that time and said that he had started hanging around the "wrong kids" at the age of 11 and had begun truancy when in middle school. Joseph had regularly received reprimands from the school administration, mostly for fighting, defying teachers' authority, and misbehavior during recess or lunch period. Joseph's mother said that at this point she believed that Joseph was outside of her control and she had hoped that 8 weeks in the juvenile detention facility would help him to learn to manage his behavior. Because he had not come home for the therapy appointment, she was worried that perhaps things were just as bad as they were prior to his arrest and stay in the detention center.

The therapist noted that the family lived in a cramped apartment in a rundown building that needed maintenance. When the therapist asked Joseph's mother about her husband, she stated that he also was supposed to have come home for the therapy session. She said that he often stopped at the bar on his way home to hang out with his friends and drink. Joseph's mother was increasingly worried about her 12-year-old son, who seemed to be following in Joseph's footsteps, and reported that her 17-year-old daughter would likely be moving out of the home with her boyfriend when she had her baby in approximately 3 months.

Joseph's 19-year-old brother was in and out of the house, and his mother was never sure whether she could expect him home on a specific evening. She thought that Joseph's older brother was having difficulties with substance use, and she reported that there were times when she smoked marijuana or drank just to get away from her concerns and feel better for a little while. The therapist asked Joseph's mother about the family's economic situation and learned that the stepfather had a regular job as an auto mechanic but was not paid well and that the mother worked part time in a department store as a cashier.

The family had Children's Health Insurance for the children. This managed Medicaid program would provide financial support for the intensive home-based treatment program. During the first meeting the therapist worked to understand the family's situation and recognized Joseph's mother's affection for her children and the dismay that she was experiencing because of their difficulties. The therapist spent 90 minutes getting to know Joseph's mother and her youngest son. During this first meeting, Joseph's sister came into the house during their conversation and told her mother that she had seen Joseph hanging out on the street with friends that he used to get in trouble with before he was in the detention center.

The therapist proceeded with the assessment and commented on the strengths of the family, including their strong connections with each other, and on the hopes of the family for the futures of the younger brother and the sister. The therapist developed a deep respect for Joseph's mother and closed the first session by saying that she would come back the next day and meet with the mother again. The therapist wrote a message to Joseph, saying that she had

been there and that she had heard about him and knew that she would very much enjoy meeting him and invited him to be present next evening for her visit.

Over the next 6 weeks, the therapist visited Joseph and his family in their home three to four times a week. She invited Joseph's mother to call any time if she had concerns, and she helped Joseph's mother to think about her priorities for their treatment process. Joseph's mother said that her goals included helping Joseph stay out of detention and resume school; repairing her relationship with the children's stepfather and encouraging him to take care of his chronic alcohol use; improving the family's financial situation by obtaining a full-time job; and deciding with the stepfather whether to continue to allow Joseph's older brother to come and go in the household.

Mother and stepfather set up rules and consequences for Joseph with three expectations: (1) that he would attend school all day, each school day, (2) that he would attend an afternoon program at the local Boys and Girls Club, where he could receive tutoring to complete his homework, as well as engage in basketball and video games with other young people at the center, and (3) that Joseph would return home after the Boys and Girls Club for dinner and that he would have access to his cell phone and his video game set in the evening if he followed through with the plan.

Joseph initially strongly resisted the rules and expectations, telling his mother that "she had no right telling him what to do," that he was grown, and that "he could make up his own mind and make his own decisions." The therapist helped Joseph's mother reiterate to her son that she would be happy for him to have greater freedom if he demonstrated greater responsibility and was able to comply with the house rules. The therapist also visited Joseph's school and talked to his counselor and teachers and encouraged them to be warm, engaging, and supportive of Joseph, to help him know that they wanted him to be in school every day and that they wanted to get to know him and learn about him. The therapist helped Joseph to write a biography of his grandfather, which Joseph researched by talking with other community leaders and neighbors who had known him. The therapist also encouraged Joseph to resume his interest in basketball and baseball and to learn as much as he could about the history of his family and the history of his family's role in his community.

As adults who expressed interest in him and his ideas surrounded Joseph, he began significantly to improve his mood, decrease his level of anger, and increase his hopes for the future and his interest in school and his after-school program. The intensive treatment program continued to work with Joseph and his family for 3 months and encouraged the adults in Joseph's life to express more responsibility and interest in him and to offer him more opportunities to be a part of his school and a part of his neighborhood. The therapist also encouraged the mother and stepfather to resume attending their church, which the family did, and the therapist talked to the minister about being particularly welcoming for Joseph, his older sister, and his younger brother. The therapist also talked with Joseph's stepfather about his alcohol use. The stepfather was receptive to decreasing his alcohol use and being home more in the evenings and paying more attention to his family and his wife.

Together with consistent, nearly daily contact with the family, involvement with several important adults, and child-serving institutions in Joseph's community, the therapist, by creatively using Joseph's interest in his family and in his grandfather and identifying Joseph's interest in telling the stories of his family, was able to improve Joseph's school attendance, sense of himself, and the family's overall sense of well-being in the 12 weeks of contact. When the program ceased, Joseph had several adults involved in his life, his mother had a full-time job that felt rewarding, her relationship with her daughter who was about to have her first child had significantly improved, and Joseph's younger brother had resumed his academic progress and his athletic activities.

A family therapist at the local mental health center was able to meet Joseph and his family for the last 3 weeks of the intensive program once a week at their clinic. That therapist continued to meet with Joseph and his family over the next several months on a weekly or bimonthly basis.

At the close of treatment, 1 year after Joseph had been discharged from the detention center, he was at home consistently, was attending school, had caught up academically, and had resumed regular activities with peers who were interested in sports and academic success. Joseph's mother and stepfather had decided to remain married and had improved their relationship.

Family Therapy for Depression

The core of treating depression in children is the consistent building of competence and connections for the child. The problem in accomplishing these goals is that the depressed child has problems with concentration, which directly interferes with competence, especially at school, and marked irritability, which greatly interferes with successful relationships with family, peers, and important adults. Furthermore, the depression creates a marked lack of motivation, which leads the child to not believe in or pursue treatment. Substance abuse is often a comorbid problem, which may leave the child feeling better in the short term but with a significant long-term problem in addition to the symptoms of depression. The longer depression persists, the more likely it is that such a child will develop significant and persistent secondary problems, including academic failure, legal difficulties, and further alienation from parents, who may also be experiencing stress.

In the face of these difficulties, family treatment offers several advantages. The parents can be asked to ensure that the child participate in treatment, even if the child would rather not, as long as the child's problems persist. Suicidal ideation, plans, and attempts can occur at this point and require specific focused attention by the therapist and the parents. Parental warmth and encouragement can be achieved by linking the child's problems to the experience of depression. Depression can be defined as a phenomenon that has taken over their child, and which is painful and hurtful while draining the child's energy and motivation. The therapist can reinforce and support the parents caring for and commitment to their child. As the parents express their care for the child, the therapist can build the child's trust in the parents.

The therapist can then encourage the parents to expect the child to participate in school, both in academics and in activities while enlisting school personnel to engage with the child and provide remedial support if needed. The therapist encourages parents to be available and attentive to the child's accomplishments and supportive of the child's emotional experience. Therapeutic support for building on the child's interests, strengths, and abilities is also important, as are efforts to build relationships with prosocial peers while also limiting substance use. Prescription of selective serotonin reuptake inhibitors (SSRIs) while monitoring suicidality can promote this process for the child and family if needed.

Lydia was a 14-year-old girl who was brought to the emergency room of her local hospital after ingesting an overdose of 15 acetaminophen tablets. Lydia had taken the overdose after an extremely difficult day in which she was teased by friends at school, was criticized by her teacher for not having completed her homework, and then engaged in an argument with her

mother about her wish to go out with friends that evening. Lydia challenged her mother and threatened to leave the house. Her mother was worried that Lydia might be meeting a teenage boy and got very angry at Lydia and called her a demeaning name. During the argument Lydia rushed into the bathroom and found the bottle of acetaminophen tablets and took the overdose. She then became quite frightened and told her mother that she had taken the pills. Her mother called an ambulance, and Lydia was brought to the emergency room.

Lydia maintained consciousness and talked openly with the clinician in the emergency room about her increasing distress over the last 12 months. She had become increasingly involved with peers who were in trouble, and her performance in school had declined. She also reported that she had stopped participating in her usual activities at school, including quitting the cheerleading squad. She had been sexually active with a 16-year-old boy on one occasion and reported that the boy had stopped going out with her after they had had sex, and he had talked about the experience in an embarrassing way with other young people at her school. She had previously been a good student and popular in her class but had recently become more irritable, more isolated, and less interested in academics and her friends.

Lydia's mother reported that her daughter had changed from a happy, upbeat teenager to a surly, irritable child who was very difficult to communicate with. Lydia lived with her mother and her three younger brothers, who were 9, 7, and 5 years old, respectively. Their father had left the family to work in another community and was sending money home to the mother. He was away for several months at a time doing construction work and was only in contact with the family by phone once a week. Lydia had been very close to her father, and his recent occupational absence had been distressing for her. She had felt isolated and disconnected from her family.

Lydia's relationship with her mother had always been ambivalent; she had been very helpful but not very emotionally close to her mother and had become more quick to anger and argumentative with her mother over the last several months.

Lydia was admitted to the adolescent medical unit for monitoring because of the possibility of toxicity from the acetaminophen. While in the hospital, she and her mother were seen by a child psychiatry consultant, who invited Lydia to talk openly with her mother about why she had made the suicide attempt and what had happened that day that had so upset her. The consultant also encouraged Lydia to talk about missing her father and her recent sadness. Her mother, while being somewhat defensive, was supported by the clinician to listen to her daughter's difficulties and concerns.

Lydia was able to be open with her mother about the problems she was having in school, the difficulties that she was having with peers, and the fact that school suddenly had become very difficult for her. The clinician held daily sessions in the medical hospital over the next 3 days with Lydia and her mother to enhance the closeness and supportiveness of their relationship. He encouraged Lydia's mother to listen to her daughter in a nonjudgmental and nonreactive fashion. As the mother felt the support of the clinician and his encouragement to be clear with Lydia about her concern about the impulsive suicide attempt, her mother was also able to be more understanding and available to her daughter during these meetings. The clinician also arranged a phone session with Lydia, her mother, and her father.

After 3 days of medical hospitalization Lydia was able to apologize to her mother for taking the overdose and to ask her for enhanced involvement and understanding at home. The family was then referred to an outpatient clinician at the local mental health center, where they were seen 3 days after discharge. Lydia reported improved mood and increased attention from her mother at home and had returned to school. The new therapist asked permission to contact Lydia's teachers to arrange more support at school. She strongly praised the enhanced closeness between Lydia and her mother.

Lydia was not initially placed on psychopharmacology. She had noted improved mood at home and had managed her behavior with less irritability and impulsivity in the initial week following her hospitalization. Future sessions were held weekly and were divided, with half the session for Lydia individually and half of the time with Lydia and her mother. The therapist focused with Lydia on her life at school, her schoolwork, and her peer relationships. The meetings between Lydia and her mother led to enhanced understanding, improved communication, and greater mutual support. Regular phone calls involved Lydia's father.

With the enhanced family support and Lydia's increasing competence as a teenager, her mood improved dramatically and her relationship with her mother grew closer. The telephone

contacts between Lydia and her father re-engaged them and restored the warmth they had shared.

After 3 months of outpatient treatment Lydia was doing well and came for sessions once every 2 weeks. The therapist noted that Lydia's mother was sad and missed her husband and offered the mother the opportunity to come for individual meetings to talk about this and how she could carry on at home while her husband was away.

Traumatic Stress in Children

Trauma-focused cognitive-behavioral therapy for children has been identified as an evidence-based treatment for traumatic stress in children. It is important to recognize that this treatment has an essential family component. Family education, emotional support for the parent who may be distancing themselves due to the child's traumatic experience, and family involvement in supporting the child while developing and sharing the trauma narrative are essential elements in this treatment. This often requires that parents accept the reality that they could not protect their child from harm or that one of both of them might actually have perpetrated the harm. The challenge of helping the parents acknowledge their feelings and the realities of their child's experience of exposure to or injury by trauma often can be quite challenging. The therapist has two goals: to support and care for the parents, and to assist the child in understanding and dealing with what has happened to him or her. This can be horrific and is often frightening or infuriating for the therapist. Therapists in this situation often need an opportunity to discuss cases with colleagues or with a supervisor. Attention to the safety of all participants provides the foundation for engagement, and engagement builds the trust needed to share the painful, shameful, and sad aspects of trauma and its aftermath. These realities are present regardless of the source and nature of the trauma. Focusing on a successful resumption of family life founded on the belief that the trauma has led to closer relationships and greater resilience is often healing in itself. The recognition of family resiliency is then coupled with acceptance of the unpleasant emotions that are normal responses to trauma among them all. A step-by-step therapeutic approach including the family is extremely helpful. This approach leads from providing safety to ensuring that basic needs are met. It continues with education for understanding, behavioral competence, emotional support, and narrative creation and validation. The trauma becomes a remembered but not excessively arousing experience. As with family treatment of other child problems, the treatment can lead not only to resolution of traumatic stress, but also to an enhanced awareness of family strength among all members.

Monica was 6 years old and was referred by her pediatrician to a child psychiatrist for evaluation after she had been involved in a serious motor vehicle accident. One month prior to referral Monica had been riding in the car with her mother. She had been seated in a booster

chair in the back seat. Their car was hit by another car that was driving very fast. Only Monica and her mother had been in the car, and Monica's mother was seriously injured in the accident. When the ambulance came to pick Monica up they noticed that she was crying for her mother and that she was shaking in the back seat, although she did not seem to have been physically injured in the accident.

Monica's mother had lost consciousness in the front seat and needed to be taken to the emergency room in a helicopter. The police had removed Monica from the back seat, and one of the policemen held her on his lap while she was crying out for her mother. They found the mother's pocketbook and called Monica's father at work. As Monica's mother was being flown to the emergency room in the helicopter, Monica's father came to the scene of the accident and picked up his daughter.

Monica had previously been a healthy girl attending first grade. She had two younger sisters ages, 4 and 2 years, respectively. Her father was a lawyer in a firm in the community and her mother was a part-time bookkeeper for a local physician. The family had significant resources, and the children attended either private school or private day care.

Monica was slow to warm up and approached new situations with fearfulness and reticence. Now, 3 months into the school year, she had established friends and had a very positive relationship with her teacher. When her father brought her home and got a neighbor to stay with his three daughters while he went to the hospital, everyone noted that Monica was extremely upset.

During the mother's hospitalization, which lasted for 2 weeks, she required abdominal surgery to remove her spleen and was found to have a fractured tibia and fibula in one leg and required a cast and crutches. Monica had been able to talk with her mother by phone several times in the hospital beginning on the fourth day after the accident, and near the end of her mother's hospitalization, Monica had been able to visit her in the waiting area of the hospital.

When Monica's mother came home she learned that Monica had been crying every morning on the way to school and also while at school. Monica also had become increasingly withdrawn at home from her younger sisters and at school from her peers and from school activities. Because of this withdrawal and problems with separation, Monica's pediatrician recommended that she and her family see the child psychiatrist.

The child psychiatrist obtained the history just described and noted that Monica's mother was also having nightmares and flashbacks of the accident and was increasingly hypervigilant. She had become frightened to drive the car. Monica's father was very concerned about his wife and about Monica and attended the initial session with them. He stated that he knew that his wife and his oldest daughter were both quiet and shy and were very much alike and very close to each other.

The child psychiatrist determined that Monica's father was deeply invested in his marriage and in wanting to help Monica and his wife recover from the physical and psychological effects of the automobile accident. During the conversation Monica's mother stated that she felt extremely guilty for having the accident and was frightened both about what had happened to her and also the possibility that the accident could have hurt her daughter. She began crying in the session and it was obvious that her distress deeply affected both Monica and her father.

The child psychiatrist talked with the family about the fact that they had a normal reaction to an extraordinarily frightening and traumatic event. He also explained that the trauma for Monica was somewhat different than the trauma for her mother. Monica's trauma involved not only the accident, but also her fear about her mother and her experience of being separated from her mother at the scene of the accident and during her mother's hospitalization.

The psychiatrist contrasted that with the mother's fears from her injury and treatment, her worries about her daughter, and her guilt from the car accident. The child psychiatrist also spoke about the strengths he recognized in the family: the father's intense connection to the family, his dedication to be helpful to his wife and his daughter, and their impressive closeness and love for each other. As he said this, the family relaxed. The therapist suggested that there would need to be a realignment of the family's energy now that the threat to survival was over. He suggested that the family begin to orient themselves to resuming their lives as a family while also recognizing their fragile emotional state.

The child psychiatrist recommended a family-based course of therapy over the next 8 to 10 weeks in which the mother, father, and daughter would work on reconnecting with their future lives while integrating their experience of trauma, injury, and anxiety into their past. The

psychiatrist proposed homework for the family; he suggested that the father spend time alone with Monica, talking with her about the fact that her mother was better now and that it was her job to resume being a big sister to her littler sisters, a student at school, and a regular 6-year-old girl. The therapist also suggested that mother and father talk together to help the mother to accept the accident and also to appreciate that because of her appropriate attention to her child's safety, her child was in the booster seat in the back seat and was not injured. They could discuss the fact that the mother had taken very good care of Monica and that Monica was having trouble because she continued to be frightened that her mother was not all right.

Over several sessions, Monica's mother was able increasingly to reassure her daughter that she was okay and encouraged her daughter to re-engage with her school, friends, and enjoyment of activities and fun as a child. She also gradually reassumed her role as an effective and caring mother for her three daughters. The father stayed actively involved with all of the children and maintained his belief in his wife's ability to recover and resume her role as his wife and their children's mother.

It became apparent to the therapist and Monica's parents that they had too rigidly adopted complementary roles as breadwinner and caretaker and that some readjustment of their partnership would need to be addressed in the therapy and recovery process. Monica's mother was grateful to her husband for being more flexible and responsive to her and their children's needs. He continued to be invested in his career, but with less intensity and less single-mindedness.

After 4 weeks of treatment Monica was increasingly acting as she had prior to the accident, as a successful student, a confident big sister, and someone who could more easily separate from her mother and the family home. She had fewer episodes of tearfulness and was able to sleep through the night.

After 8 weeks Monica's teacher reported to the parents that she had in fact resumed previous functioning and had caught up academically and appeared to all to have recovered. Her mother was less frightened and was sleeping well and had resumed work and driving. She and Monica could discuss the accident openly without crying.

Treatment of Anorexia and Substance Abuse

Family therapy–based treatment models have demonstrated effectiveness for both anorexia nervosa and adolescent substance abuse. In both of these problems the child is involved in an engaging, self-reinforcing, and dangerous behavior. This behavior often is pursued in an extreme fashion and ultimately leads to failure to achieve other developmental skills. The lack of these skills also often renders the child anxious and further narrows the child's orientation to his or her life. There may be genetic predisposition to either anorexia or substance abuse. Despite this, the child will need to learn to manage the impulse to diet or the impulse to use substances in a dangerous manner. In the majority of these situations, the parents will bring the child for treatment and will be concerned about their inability to lead their child out of his or her self-destructing behavior patterns. Two evidence-based family treatments have been developed for these problems, and their effectiveness have been documented through randomized, controlled treatment trials. These treatments are multidimensional family therapy for adolescent substance abuse and family-based treatment for anorexia. Both of these treatments assign responsibility to the parents for reducing the intensity of the symptomatic behaviors. In family-based treatment for anorexia the parents are given the responsibility for helping their child gain weight with the therapist's

support. As the parents take on this responsibility, the therapist helps them to be successful in attaining the child's attention and the child's compliance. The child gains greater support and autonomy as he or she participates in treatment. The therapist also supports the child's interest in and competence with activities and with peers. If treatment is initiated early in the course of the disorder, it can be effective and relapse is rare. Family treatment for substance abuse in adolescents is similar. However, often school failure, legal difficulties, a history of poor supervision, and the influence of negative peers complicate the clinical situation. The treatment manual for multidimensional family therapy identifies ways of addressing these challenges while simultaneously providing the supervision, support, and limits necessary to reduce substance use. Once again the keys to treatment effectiveness are opportunities to behave in prosocial ways with intense family support for competency and success associated with consistent unwavering expectations and reasonable but not excessive consequences.

Joan was 14 years old and lived with her parents and 12-year-old sister. She was referred to a child psychiatrist by her pediatrician, who had recently evaluated her for significant weight loss and dramatically reduced food intake.

Joan's parents were middle class, and both worked long hours. They had become concerned about her dieting and weight loss but had been unable to have it cease. When Joan and her mother met with the child psychiatrist, Joan stated very clearly that her weight was a major concern to her, that she still felt that she needed to diet and lose weight, and that under no circumstances was she going to eat what the pediatrician and nutritionist had recommended. Joan's father was a busy executive who traveled twice monthly and who had grown increasingly distant from his daughters over the last year. Joan's mother stated her concern that her husband was not as involved with his daughters as he should be and that now his sole interaction with Joan concerned her eating. He yelled at her, trying to get her to eat. Joan's mother had thought that her daughter was experiencing anorexia nervosa, and the therapist concurred. The therapist strongly suggested that Joan's mother and father come to a session later that week to discuss a plan to help Joan with her anorexia. The therapist received a phone call from the father stating that he needed to be out of town on business, but the therapist stated firmly that he should postpone his trip and attend the session with his wife and daughter.

The therapist's first efforts in the meeting were to emphasize to both parents the need for them to work together and to assist Joan with rules that would encourage her to follow through with the food recommendations of her pediatrician and her nutritionist.

As the therapist helped Joan's parents plan how they would expect Joan to eat the required amount and gain weight as suggested by her pediatrician, the therapist also suggested that Joan finish her meals before she left the table and pursued other activities or completed her homework. Joan was upset and frightened at the prospect that she would gain weight. The therapist helped the parents to be understanding but firm and to reiterate their expectation of consistent weight gain. The parents were also advised to be explicit about caloric expectations and meal planning.

As Joan gained weight, she gained more freedom to choose which foods the family had in the house, but she was still required to eat the required amount of calories. Her rate of weight gain was consistently 1 lb per week, and she went from 80% of ideal body weight to 95%. As Joan began gaining weight, the therapist noted increasing distress in her parents' relationship. The therapist suggested private couple sessions for the parents to discuss their relationship. Individual sessions were held with Joan to assist her with assertiveness, peer relationships, and ways of decreasing her perfectionism with regard to her schoolwork.

Joan participated in her individual sessions eagerly and was soon noted to be spending more time with friends and less time preoccupied with food, dieting, and her weight. When she reached 100% of ideal body weight her menses resumed. The therapist was also meeting regularly with Joan's parents. Joan's mother expressed her concern that her husband was more involved in work than with the family and that this was not all right with her. She became assertive with him, and they began to negotiate a different approach to family and work. They also were advised to spend some time alone together. Joan's father responded well to the couple sessions, and her parent's marriage improved and became more satisfying for both of them.

Over the 8 months that therapy continued, Joan resumed normal adolescence, family relationships improved significantly, and Joan's father became more engaged with his daughters.

Stephanie was a 15-year-old girl who was referred to the substance abuse treatment program because her parents were concerned about her alcohol and marijuana use. She lived with her stepmother and biological father. Prior to this school year she had been a successful student and had played in her high school band.

Stephanie's biological mother had significant trouble with alcohol and drugs when Stephanie was a preschooler. After failed attempts at treatment, Stephanie's father had separated from her mother and gained full custody of his daughter. Stephanie knew that her mother was in another state and believed that she was still having trouble with alcohol use. She had not had contact with her mother for 12 months prior to her referral.

Stephanie's father had remarried when Stephanie was 6 years old, and Stephanie had two younger half-siblings, a 7-year-old brother and a 4-year-old sister. This year Stephanie had increasingly felt disconnected from her stepmother and her younger half-siblings. Her relationship with her father had begun to deteriorate as she became more defiant of rules and more involved with peers. She was home less frequently and was less involved with the family activities. Stephanie's stepmother had been called by her school when she had been found somnolent and disoriented at school.

Stephanie's stepmother and father took her to the emergency room, where she was found to be intoxicated with alcohol. She said she had gotten the alcohol when she and some friends had gone to a friend's house during lunch period. Stephanie's father and stepmother had been concerned that Stephanie might be using substances frequently and in a problematic way. Over the last several months she had been coming in late after curfew, staying over at friends' houses frequently, and had been less involved with her previous friends and more involved with a group of friends that the family did not know very well. Stephanie had also dropped out of band and stopped playing her instrument. Her academic performance had decreased significantly.

Stephanie agreed that she was using drugs and alcohol but steadfastly stated that she did not have trouble with them and that she was only doing what her friends did. When Stephanie, her father, and her stepmother met with the family therapist at the drug treatment center it became apparent to all that Stephanie's problems were quite significant and were seriously jeopardizing her physical safety as well as her developmental progress.

In the first session Stephanie tearfully admitted to involvement with unprotected sex, especially when she was intoxicated. With guilt and shame she admitted that she did not want to be doing what she was doing but had found it very difficult to refuse substances or affection and physical intimacy when she was with peers.

The family therapist suggested to Stephanie's father that she was scared and lost rather than defiant and delinquent. Stephanie's father, with some difficulty, acknowledged that this might be true. The family met with the therapist three times a week at the program while Stephanie participated in the adolescent after-school drug treatment groups. The therapist helped Stephanie's parents to institute a curfew and ensure adult supervision after dinner and on weekends. Stephanie reluctantly, but in fact with relief, accepted these rules and adhered.

Regular drug testing was part of Stephanie's program, and over the first month of treatment she was consistently free of evidence of drug and alcohol use. The therapist also strongly recommended that Stephanie attend the adolescent 12-step program meetings with the teens in the treatment group. The therapist also spoke with Stephanie's counselor and teacher at

school. They were advised to strongly encourage Stephanie to rejoin the band, resume her musical practice, and participate in tutoring and additional homework to improve her academic standing. Stephanie responded to this with some initial reluctance, but ultimately with the support and encouragement of her teachers, school counselor, therapist, and band leader she was able, after a rocky period of a couple of weeks, to be more involved in school and more involved with the friends she had distanced herself from after she had gotten involved with her substance-abusing friends.

During this time Stephanie was also seen by a child psychiatrist. The child psychiatrist reported that Stephanie was sad, but he did not believe that she warranted antidepressant medication. The child psychiatrist told Stephanie, her father, and her stepmother that he was very pleased with the progress that they were making in the intensive substance abuse treatment program.

Stephanie's mood continued to brighten. She was encouraged to have evening and weekend times with her father and stepmother. Her stepmother was advised to spend time talking with Stephanie about her friends and her activities. She was encouraged to avoid disciplining Stephanie but to inform her father about any concerns. Stephanie responded to this with improved connections with both her father and her stepmother. She also spent more time with her half-siblings.

Three months into treatment, at the time that Stephanie's discharge from the intensive treatment program was planned, her urine tested positive for alcohol. Her parents also discovered that she had come home after curfew apparently intoxicated after attending a party without parental permission.

The therapist helped the family to recognize that Stephanie had had a difficult weekend and continued to have difficulty resisting the attraction of alcohol and the encouragement of peers to use substances. The therapist worked intensively for the next 2 weeks to reconnect Stephanie with her father and stepmother and to re-establish more-intensive supervision for Stephanie. The therapist encouraged the father and stepmother to maintain hopefulness about Stephanie's success in the future. Stephanie asked her parents to be understanding and forgiving of her episode of alcohol abuse. Her father and stepmother realized that they could not yet expect Stephanie to be able to make mature decisions when exposed to substances and that they would need to provide intensive supervision.

The therapist also spoke with Stephanie's father and stepmother about the difficulties they were having as a blended family and worked with them to understand the challenges of that experience. They worked on building one family of five, rather than two families, one with four members and another with three members. At the same time the therapist spoke with Stephanie about her experience of abandonment and hurt in her relationship with her mother and about her worries about her mother's well-being. Stephanie talked about her fear that she would grow up to be like her mother. As Stephanie held these conversations with her therapist she was able to be increasingly mature and thoughtful about herself and her previous experience of neglect and abandonment. She also was encouraged to speak about this with her father in a session.

The involvement of the entire family in treatment, the appreciation of the family's history and experiences, Stephanie's involvement in a teen substance use program, and the active involvement of her school had all been elements in treatment. Individual therapist support to help her to understand and accept her difficult relationship with her mother was also essential. All of this helped the family to bond together more openly and to be increasingly successful in the developmental challenges that they were facing.

Family Treatment for Divorcing and Blended Families

Family transitions occurring as a result of parental separation and divorce and remarriage are disruptive to children. If parents are able to be congenial with each other and to amicably resolve issues related to the divorce (including financial issues, grieving the end of the marriage, and child care and custody issues), the children, though they still grieve the end

of their family and accommodate the changes to family life with sadness, will usually do so without significant relationship problems and persistent behavior problems. However, if significant resentment, anger, and conflicts continue between the former partners, the effect on the children can be significant and persistent. Problems with academic performance and mood and behavior problems can develop. The conflicting parents often do not manage these difficulties successfully, and family treatment can be necessary. The first step in this treatment process is to identify and support the concern that each parent has for the children and for resolving the children's problems. Parents may be motivated by different concerns—for example, one by a child's academic difficulties and the other by the child's sadness. In any event, bringing together antagonistic parents is an important element in treatment. The family therapist should work extremely hard to help the parents to understand that their children need both of them and need them to work together in parenting for the children's well-being. This process may be difficult and time consuming, but it provides the best foundation for the children's successful outcome. As the therapist helps the biological parents work together, he or she can also help the children express their concerns with each parent individually in separate therapy sessions.

Several aspects of post-divorce life can be problematic for the children. Often a parent can be very involved in his or her life, including work and relationship building. This can lead the children to be unsupervised and to further engage in problematic behaviors. To address this, both parents (if possible) will need to supervise the children and to reiterate their belief in the children's strengths and capacity to be successful.

When stepparents are involved in problematic post-divorce situations, matters can worsen. Stepparents may decide that enhanced discipline is needed and may attempt to act to deliver that. As the stepparent works to impose discipline, a child often becomes indignant and defiant. The more forceful and demanding the stepparent becomes, the more defiant the child becomes. Family therapy can reorganize this situation and lead both biological parents to take a central role in the child's life and success, and to address challenging behaviors. Family therapy can also help stepparents to take a peripheral but very important role as supporter and ally for both the child and biological parent. The stepparent can express concerns and worries about the child, but rather than to the child directly, these concerns can be expressed to the biological parent, who can act on them in concert (if possible) with the biological parental partner. The therapist can also encourage the stepparent to develop a warm, accepting, and caring relationship with the child based on mutual interest. This decreases the child's antagonism toward the stepparent and may significantly improve the mood of the family while also enhancing trust and attachment between the child and the biological parent.

As time passes, the child can adapt to the changing family and maintain a supportive and engaged relationship with both biological parents and accept stepparents and step-siblings. The therapist can also help the new partners build their relationship while they manage successful relationships with all children and with former spouses.

Kevin was an 11-year-old boy who lived with his mother, maternal grandmother, 17-year-old sister, and 14-year-old high-functioning autistic brother. Kevin was referred to see a child psychiatrist because his mother was concerned about his excessive worries. The whole family had been living with maternal grandmother since Kevin's parents separated and divorced when he was a year old. Post-divorce, the parents had a long history of custody battles. At the time of the initial visit, parents had shared legal custody but the mother had physical custody. Their divorce agreement specified that if there is a disagreement regarding medical and academic needs and decisions, the mother's decision supersedes father's preference. The father's consent to provide psychiatric treatment was received, though he did not see a reason for an evaluation. Kevin and his brother would stay with their father every other weekend and have dinner with him at least once a week.

Kevin's mother wanted him to be engaged in treatment because she believes that he suffered during the custody battles. He always had difficulty academically but his performance deteriorated when his sister refused visitation with their father 3 years ago. When he was anxious, he would be short of breath and would pace around the house and scream. This would occur specifically when he was struggling with his homework and when he was supposed to visit his father. This behavior would take place every 3 to 4 days. Nightmares would wake him up shaking in the middle of the night.

His father was invited to participate in the evaluation and treatment. Initially, his involvement was limited, but when he realized that some of Kevin's difficulties were related to his upcoming wedding, he became more engaged.

Kevin and his father were encouraged to engage in conversations to communicate his difficulties with "his growing family." Kevin was able to feel safe in therapy to inform his father of his grievances including his concerns about stepmother taking over his mother's role. He worried about his mother being alone while also grieving the permanency of their divorce. With the support and involvement of his father's fiancé, his mother and his father were able to hear his concerns, to encourage him to speak-up, and to compromise when necessary. They were able to validate the loss of his original family but also to reinforce that they all had a special relationship with him individually. He was reassured that these relationships were going to remain unchanged and that one relationship cannot take place of the other. The family members were able to discuss their expectations of him at both households. They agreed to keep the rules and these expectations consistent so that it would be less confusing for Kevin. As all relationships improved, Kevin's anxiety and difficulty in school was resolved.

Family Therapy for Families with Children with Developmental or Medical Challenges

Families with children with developmental delay, autistic spectrum disorders, or chronic medical illness face unique challenges. Family therapy can help those families who are having difficulty managing emotions, behavior, and illness treatment or achieving the optimal development of the affected child or siblings. These situations often are extraordinarily sad and difficult to deal with. There can be financial challenges, career advancement can be negatively affected, illness or handicap care routines can be onerous, and opportunities for parental intimacy and family enjoyment can be threatened. Parental burden and stress can be extreme,

and at times the expected positive effects of medical treatment may not occur. Marital relationships may be threatened, and siblings may not receive adequate attention and supervision. Parental isolation may also seriously jeopardize family well-being. Family therapy can address these difficulties by offering the parents, the affected child, and siblings a safe environment in which to discuss reactions to the child's problems and to collaborate on making the situation the best that it can be. Parental cooperation and parental mutual understanding are goals. All those involved can be guided by the therapist to learn about the child's condition, current and potential treatments, and prognosis. The parents, the child, and siblings can be encouraged to ask questions and to be assertive with the medical care team. The family and child can be encouraged to recognize the areas of the child's life that are not affected by the illness or handicap and the child's strengths and abilities. The therapist can also recognize the courage and care of all in dealing with the challenges of the child's condition. As the family recognizes their commitment to the affected child, they can come to see themselves as heroic and resilient, and the child can note his or her courage and competency regardless of the illness or handicap. The therapist can help the family work with the school and teachers and with extended family members to understand and support the affected child while also helping the child engage teachers, neighbors, peers, and extended family in an assertive, positive, and confident manner. Finally, the child's active participation in his or her care can be noted and recognized as a sign of competence and mastery that will help the child throughout life. The family therapist can also support the family if the child's condition fails to improve or deteriorates. The therapist elicits and ratifies legitimate responses of sadness and powerlessness and accepts his or her powerlessness over the illness while staying involved with and attuned to family members' sadness and grief.

Dave is a 9-year-old boy who lives with his parents and 1-year-old brother. He had been undergoing treated for high-functioning autism spectrum disorder, attention-deficit/hyperactivity disorder, and anxiety disorder for 2 years. His treatment included a low-dose atypical antipsychotic, a low-dose selective serotonin reuptake inhibitor, a stimulant, and an alpha agonist. Additionally, he was receiving individual therapy every 2 weeks to increase his coping skills. The family had been referred for family therapy but, due to both parent's full-time jobs and other commitments, they had not engaged in that treatment. The family self-referred to a different child psychiatrist because they felt that Dave's treatment had become stalled.

The family's chief complaint was Dave's explosive behavior around limit setting at home and social difficulties with his peers. His father struggled to follow through with consequences, whereas his mother was always able to do it. This inconsistency had resulted in confusion for Dave and also increased conflict between the parents. Dave's mother felt that she was not being supported by her husband. Dave's father felt that his wife was very strict with Dave and would attempt to compensate for this strictness. He felt that Dave's developmental delay made school so difficult for him that he deserved to have a less demanding environment at home. As the parents talked with the therapist they realized the need to negotiate expectations and consequences for Dave. They also recognized the need to support each other more and

redevelop the warmth in their relationship. Although Dave remained difficult and challenging, he gradually improved and became much less explosive.

FUTURE DIRECTIONS

Family therapy has evolved from a therapy developed in opposition to common psychiatric care to a basic orientation to the treatment of child mental health challenges. All situations require the assessment of the involvement and education of family members. Some situations require the specific and monitored engagement of family members in the treatment plan. In many instances, treatment of the child requires the involvement of the family in the manner outlined in the manual of an evidence-based model of family intervention. Knowledge of these treatments and skills implementing them are necessary elements in child mental health care. For many child mental health problems, the problem leads to intuitive family interactions that are counterproductive and often worsen the situation. Family therapy stops these negative cycles and can introduce successful, self-reinforcing, positive family interactions that promote family life and optimal child development. These interventions can help the child's situation to improve, and the family's experience of warmth, care, and effectiveness can emerge and persist.

REFERENCES

Alexander J, Barton C, Gordon D. *Blueprints for Violence Prevention, Book 3: Functional Family Therapy*. Boulder, CO: Center for the Study and Prevention of Violence; 1998.

Anderson H, Gerhart D, eds. *Collaborative Therapy: Relationships and Conversations that Make a Difference*. New York: Routledge; 2006.

Baldwin S, Christian S, Berkeljon A, Shadish W. The effects of family therapies for adolescent delinquency and substance abuse: a meta-analysis. *J Marital Fam Ther*. 2012;38(1):281–304.

Barkley R. *Defiant Children. A Clinician's Manual for Assessment and Parent Training*. 2nd ed. New York: Guilford; 1997.

Boyd-Franklin N. *Black Families in Therapy: Understanding the African American Experience*. 2nd ed. New York: Guilford; 2006.

Carter B, McGoldrick M. *The Expanded Family Life Cycle: Individual, Family, and Social Perspectives*. 3rd ed. Boston, MA: Allyn & Bacon; 2005.

Combrinck-Graham L. *Children in Families at Risk: Maintaining the Connections*. New York: Guilford; 1995.

*Combrinck-Graham L. *Children in Family Contexts: Perspectives on Treatment*. 2nd ed. New York: Guilford; 2006.

*Diamond G, Josephson A. Family-based treatment research: a 10-year update. *J Am Acad Child Adolesc Psychiatry*. 2005;44:872–887.

Diamond G, Rosson J, Levy S. Attachment-based family therapy: a review of the empirical support. *Family Process*. 2016;55(3):595–610.

Handen B, Johnson C, Butter EM, et al. Use of a direct observational measure in a trial of

risperidone and parent training in children and pervasive development disorders. *J Dev Phys Disabil*. 2013;25(3):355–371.

Hembree-Kigin T, McNeil C, Eyeberg S. *Parent–Child Interaction Therapy*. New York: Springer; 1995.

Henggeler S, Sheidow A. Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. *J Marital Fam Ther*. 2012;38(1):30–58.

Henggeler S, Schoenwald S, Borduin C, Rowland M. *Multisystemic Treatment of Antisocial Behavior in Children and Adolescents*. New York: Guilford; 1998.

Henggeler S, Schoenwald S, Rowland M, Cunningham P. *Serious Emotional Disturbance in Children and Adolescents: Multisystemic Therapy*. New York: Guilford; 2002.

Henggeler SW, Schaeffer CM. Multisystemic therapy: clinical overview, outcomes and implementation research. *Family Proces*. 2016;55(3):514–528.

Jewell T, Blessitt E, Stewart C, Simic M, Eisler I. Family therapy for child and adolescent eating disorders: a critical review. *Family Process*. 2016;55:577–594.

Josephson AM; AACAP Work Group on Quality Issues. Practice parameter for the assessment of the family. *J Am Acad Child Adolesc Psychiatry*. 2007;46:922–937.

Kazdin H, Weisz J. *Evidence-Based Psychotherapies for Children and Adolescents*. New York: Guilford; 2003.

Liddle H. *Multidimensional Family Therapy for Adolescent Cannabis Users*. Rockville, MD: Center for Substance Abuse Treatment/Substance Abuse and Mental Health Services Administration; 2003.

Liddle HA. Multidimensional family therapy: evidence base for transdiagnostic treatment outcomes, change mechanisms, and implementation in community settings. *Family Process*. 2016;55:558–578.

Lock J, Le Grange D, Agras W, Moye A, Bryson S, Jo B. Randomized clinical trial comparing family-based treatment to adolescent-focused individual therapy for adolescents with anorexia nervosa. *Arch Gen Psychiatry*. 2010;67(10):1025–1032.

*Madsen W. *Collaborative Therapy with Multi-Stressed Families: From Old Problems to New Futures*. New York: Guilford; 2003.

McFarlane WR. Family interventions for schizophrenia and the psychoses: a review. *Family Process*. 2016;55:460–482.

McGoldrick M, Gerson R, Petry S. *Genograms: Assessment and Intervention*. 3rd ed. New York: W. W. Norton; 2008.

*McGoldrick M, Giordamo J, Garcia-Preto N. *Ethnicity and Family Therapy* 3rd ed. New York: Guilford; 2005.

McGoldrick M, Hardy K. *Re-visioning Family Therapy: Race, Culture and Gender in Clinical Practice*. 2nd ed. New York: Guilford; 2008.

Miklowitz DJ, Chung B. Family-focused therapy for bipolar disorder: reflections on 30 years of research. *Family Process*. 2016;55(3):483–499.

Miklowitz DJ, Schneck CD, George EL, et al. Pharmacotherapy and family-focused treatment for adolescents with bipolar I and II disorders: a 2-year randomized trial. *Am J Psychiatry*. 2014;171(6):658–667.

Minuchin P, Colapinto J, Minuchim S. *Working with Families of the Poor*. 2nd ed. New York:

Guilford; 2006.

Minuchin S, Lee W, Simon G. *Mastering Family Therapy: Journeys of Growth and Transformation*. 2nd ed. New York: Wiley; 2006.

*Nichols M, Schwartz R. *Family Therapy: Concepts and Methods*. 8th ed. Boston, MA: Allyn & Bacon; 2007.

Pineda J, Dadds M. Family intervention for adolescents with suicidal behavior: a randomized controlled trial and mediation analysis. *J Am Acad Child Psychiatry*. 2013;52(8):851–862.

Robbins MS, Alexander JF, Turner CW, Hollimon A. Evolution of functional family therapy as an evidence-based practice for adolescents with disruptive behavior problems. *Fam Process*. 2016;55(3):543–557.

Sargent J. Family therapy in child and adolescent psychiatry. *Child Adolesc Psychiatr Clin North Am*. 1997;6:151.

Sholevar C, Schwoeri L. *Textbook of Family and Couples Therapy: Clinical Applications*. Washington, DC: American Psychiatric Publishing; 2003.

Waldron H, Brody J. Functional family therapy for adolescent substance use disorders. In: Weisz JR, Kazdin AE, eds. *Evidence-Based Psychotherapies for Children and Adolescents*. 2nd ed. New York: Guilford Press; 2010:401–415.

*Walsh F. *Normal Family Processes: Growing Diversity and Complexity*. 3rd ed. New York: Guilford; 2006.

Walsh F. *Strengthening Family Resilience*. 2nd ed. New York: Guilford; 2006.

White M, Epston D. *Narrative Means to Therapeutic Ends*. New York: Guilford; 1990.

*White M. *Maps of Narrative Practice*. New York: W. W. Norton; 2007.

▲ 54.6 Pediatric Psychopharmacology

TIMOTHY E. WILENS, M.D., AND CHRISTOPHER J. KRATOCHVIL, M.D.

Psychotropic medications play a crucial role in the treatment of children and adolescents with psychiatric disorders, and the past two decades have seen significant growth in the use of these agents. Fortunately, this growth has been coupled with a tremendous expansion in the pediatric pharmacotherapy database. There have been substantial advances in pediatric research due to initiatives such as the doubling of research funding at the National Institutes of Health in the 1990s, the Food and Drug Administration Modernization Act (FDAMA) of 1997 which provided 6 months of additional market exclusivity for companies who conducted pediatric clinical trials, and the Pediatric Research Equity Act (PREA) which required pharmaceutical companies to conduct pediatric studies of medications likely to be used by children and adolescents. Numerous federally and privately funded clinical trials have, as a result, systematically evaluated the safety and efficacy of a variety of psychotropic medications, while comparator trials have helped to clarify the potential role of these agents in treatment.

In the last two decades, dozens of randomized, placebo-controlled pediatric trials have provided safety and efficacy data which led to Food and Drug Administration (FDA) approvals for multiple single and combined agents for the treatment of youth with major depressive disorder (MDD), bipolar disorder, schizophrenia, attention-deficit/hyperactivity disorder (ADHD), autism, enuresis, generalized anxiety disorder (GAD), and obsessive compulsive disorder (OCD). Additionally, several trials conducted by the National Institute of Mental Health (NIMH) in MDD, ADHD, anxiety disorders, and OCD have compared pharmacotherapy, psychotherapy, and their combinations. The expansion in controlled clinical trials has provided clinicians in practice with much-needed information regarding the role of pharmacotherapy in the treatment of mentally ill children and adolescents, and facilitated the development of practice parameters such as those put forth by the American Academy of Child & Adolescent Psychiatry (AACAP) and the American Academy of Pediatrics (AAP). This chapter will highlight data from practice guidelines, reviews, and key clinical trials including large multisite NIH and pharmaceutically funded studies that have resulted in important new findings and/or FDA approval of pediatric psychopharmacology.

GENERAL ISSUES IN PEDIATRIC PSYCHOPHARMACOLOGY

Role of Pharmacotherapy

While psychotropic medications often play an integral part in the treatment of children and adolescents with psychiatric disorders, they are usually used in combination with other modalities. Education about diagnosis, course of illness, prognosis, and available treatment options is important. Additionally, depending on the individual, a treatment plan may warrant: individual, group, and/or family therapy; educational interventions; parent training; substance use disorder treatment; out-of-home placement; or any of a number of additional interventions.

Baseline Assessments and Rating Scales

A thorough and developmentally sensitive assessment is essential prior to the initiation of pharmacotherapy. An appropriate treatment plan is predicated on a comprehensive and accurate diagnostic evaluation, with consideration of potential diagnoses beyond the primary disorder. The majority of children and adolescents presenting with a psychiatric disorder also suffer from comorbid medical, psychosocial, cognitive, and/or psychiatric conditions that can significantly contribute to the youth's impairment in functioning. Failure to address comorbidities may limit the benefits of pharmacotherapy or any therapeutic intervention.

Broad- and narrow-band rating scales can be useful in the screening and diagnosis of psychiatric disorders. Standardized measures generally provide normative data to allow clinicians to compare their patient's symptoms to those of children of similar age and same gender. Such instruments available through AACAP, AAP, schoolpsychiatry.org, and proprietary operations are available for the majority of major psychiatric disorders, and can be valuable tools for establishing baseline target symptoms and severity, as well as longitudinal tracking of changes in symptoms during the course of pharmacotherapy.

Developmental Considerations

Not unlike the importance of employing a developmental perspective in making a psychiatric diagnosis, when prescribing psychotropic medications to children, clinicians need to carefully consider developmental issues regarding the pharmacokinetics of medications as well. Unfortunately, compared to the generally large adult database, there is a more modest amount of data in grade school children and adolescents, and limited data in preschool children. This is particularly important when one considers that the data in adults cannot necessarily be simply extrapolated to a younger population. Changes in age can be accompanied by differences in hepatic metabolism, renal function, volume of distribution, and fat content, along with variations in tolerability and efficacy. Additionally, the patient's perspective on need for treatment, acceptability of pharmacotherapy, and adherence to medication may change.

DISRUPTIVE BEHAVIOR DISORDERS

Attention-Deficit/Hyperactivity Disorder (ADHD)

ADHD is one of the most commonly presenting pediatric neurobehavioral disorders. ADHD affects 6 to 9 percent of children and adolescents worldwide, and persists into adulthood in roughly 50 percent of cases. It is also one of the best studied, and is unique in that, rather than relying on information extrapolated from adult data, a pediatric focus has historically guided ADHD psychopharmacology trials. The efficacy of the stimulants for the treatment of ADHD in children was well established by the 1970s, and stimulant use as well as data supporting its safety and efficacy has grown exponentially since that time.

Research has been particularly helpful in clarifying the role of pharmacotherapy in the treatment of ADHD in children, including studies enrolling preschoolers as young as 3 years of age. Concurrently, major strides in developing delivery mechanisms have occurred, resulting in extended-release stimulant medication preparations with improved ability to sustain treatment effects throughout the day and lower potential abuse liability. Similarly, there have been successes in the development of novel medications for ADHD.

Pharmacotherapies for ADHD. The stimulants are highly efficacious medications, with more than 6 decades of randomized controlled trials (RCTs) documenting their safety and efficacy. Studies have generally demonstrated that 65 to 75 percent of patients treated with stimulants respond, with similar rates for methylphenidate and amphetamine, in immediate-release as well as extended-release preparations. The effect size for stimulants is one of the largest of the psychotropic medications, averaging 1.0. The rapid onset of action, overall tolerability and safety, extensive database, and robust treatment effects have led to extensive use of these medications.

Table 54.6–1.
FDA-Approved ADHD Pharmacotherapy

Medication	Trade Name	Mode of Delivery	FDA Approved Max/Day	Strengths Available (mg)
STIMULANTS				
Methylphenidate	Concerta	OROS delivery system: 18% immediate-release outer coating and 82% gradually released osmotically; designed to replicate t.i.d. immediate-release	54 mg in children; 72 mg in adolescents 13–17 y.o.	18, 27, 36, 54
	Daytrana	Patch worn up to 9 hours per day, gradually releasing	30	10, 15, 20, 30

		methylphenidate		
	Metadate CD	Beaded delivery system: 30% immediate release and 70% 3 hours later	60	10, 20, 30, 40, 50, 60
	Metadate ER	Methylphenidate gradually released from wax matrix	60	10, 20
	Methylin	Immediate release	60	5, 10, 20
	Methylin ER	Methylphenidate gradually released from wax matrix	60	10, 20
	Ritalin	Immediate release	60	5, 10, 20
	Ritalin LA	Beaded delivery system: 50% immediate release and 50% 4 hours later	60	10, 20, 30, 40
	Ritalin SR	Methylphenidate gradually released from wax matrix	60	20
	Quillivant XR	Once daily (up to 12 hours), extended-release liquid formulation of methylphenidate HCl	60	25 mg/5 cc
	QuilliChew ER	Methylphenidate chewable tablet	60	20, 30, 40
D-methylphenidate	Focalin	Immediate release	20	2.5, 5, 10
	Focalin XR	Beaded delivery system: 50% immediate release and 50% 4 hours later	20	5, 10, 15, 20
Amphetamine	Adzenys XR	Extended release amphetmaine-based orally disintegrating tablet	18.8 mg in children; 12.5 mg in adolescents	3.1, 6.3, 9.4, 12.5, 15.7, 18.8
	Dexedrine	Immediate release	40	5
	Dexedrine Spansule	Beaded delivery system, initial dose released immediately and remainder gradually released	40	5, 10, 15
	DextroStat	Immediate release	40	5, 10
	Dyanavel XR	Extended release amphetamine-based liquid suspension	20	2.5 mg/cc
	Evekeo	Immediate release	60	5, 10
Mixed amphetamine salts	Adderall	Immediate release	40	5, 7.5, 10, 12.5, 15, 20, 30
	Adderall XR	Beaded delivery system: 50% immediate release and 50% 4 hours later	30	5, 10, 15, 20, 25, 30
Lisdexamfetamine	Vyvanse	D-Amphetamine with lysine attached, activated in GI tract when lysine is cleaved	70	20, 30, 40, 50, 60, 70
NONSTIMULANTS				
Atomoxetine	Strattera	Immediate release: generally dosed QD, but can be dosed BID	Lesser of 1.4 mg/kg or 100 mg	10, 18, 25, 40, 60, 80,

				100
Guanfacine	Tenex	Immediate release	4	1
	Intuniv	Extended release	4	1, 2, 3, 4
Clonidine	Catapres	Immediate release	0.4	0.1, 0.2, 0.3
	Catapres Patch	Extended release	0.4	0.1, 0.2, 0.3
	Kapvay	Extended release	0.4	0.1, 0.2

Note: Not all stimulant doses and preparations will be available at all pharmacies. Often, a less frequently used stimulant preparation may need to be ordered in by your pharmacist. Some of the stimulants are generic (available usually less expensively for either private full pay or for co-pays).

ATOMOXETINE (STRATTERA). The first nonstimulant approved for the treatment of ADHD, atomoxetine, works by blocking presynaptic uptake at noradrenergic neurons. While atomoxetine has been proven efficacious in several controlled trials, its efficacy appears to be less than that of the stimulants. One meta-analysis reported atomoxetine's effect size as 0.62, while the stimulants were from 0.91 (immediate release) to 0.95 (sustained release).

ALPHA-AGONISTS. The extended-release preparations of clonidine (Kapvay) and guanfacine (Intuniv) have multiple, recent controlled trials demonstrating their efficacy as monotherapy in the treatment of ADHD in children and adolescents. The alpha-agonists may be useful for core symptoms of ADHD as well as associated impulsivity, sleep, and tic disorders. Typically the extended-release clonidine is given twice daily, and the extended-release guanfacine is given once daily either at night or in the morning with equal efficacy through the day. Extended-release alpha-agonists have been shown safe, effective, and FDA approved as monotherapy as well as in combination with stimulants. Interestingly, in stimulant partial responders, alpha-agonists appear to heighten ADHD response (effect sizes of 0.4 in addition to partial response to stimulants).

Several additional medications have demonstrated benefit in controlled trials, but have not been approved by the FDA for the treatment of ADHD.

BUPROPION (WELLBUTRIN, ZYBAN). This is a nonstimulant that acts via dopamine and norepinephrine. Data is very limited, and efficacy appears to be less than that of stimulants, but it could be a consideration in the presence of co-occurring mood disorders, substance abuse, or smoking.

MODAFINIL (PROVIGIL). Efficacy has been demonstrated in three double-blind, placebo-controlled studies in pediatric ADHD using doses higher than those typically utilized for narcolepsy, but was not approved by the FDA for the treatment of ADHD due to safety concerns (rash; erythema multiforme). Data in adolescents are lacking. Modafinil may be particularly useful for enhancing attention, motivation, and general arousal.

TRICYCLIC ANTIDEPRESSANTS. Imipramine (Tofranil), nortriptyline (Pamelor), and desipramine (Norpramin) have all demonstrated efficacy in the treatment of ADHD. The tricyclic antidepressants may be particularly useful for all-day coverage of ADHD and for comorbid ADHD. However, monitoring requirements, side-effect burden, and overdose concerns limit their use.

Select Key Studies

MULTIMODAL TREATMENT STUDY OF CHILDREN WITH ADHD (MTA). The MTA study enrolled 579 children with ADHD, aged 7 to 9.9 years. The children were randomized to either 14 months of a manualized medication management program, an intensive psychosocial treatment, a combination of the two, or community treatment. All four groups improved, with the greatest improvement in ADHD symptoms occurring in the medication alone or combined medication/psychosocial treatment groups. Combined treatment did not yield significantly greater benefits than medication management alone for core ADHD symptoms. Modest advantages were identified for certain non-ADHD symptoms and other functional outcomes.

NEW YORK–MONTREAL LONG-TERM ADHD TREATMENT STUDY. This 2-year study enrolled 133 children (aged 7 to 9 years) with ADHD who had responded to short-term methylphenidate, and randomized them to: methylphenidate alone; methylphenidate plus psychosocial treatment (parent training and counseling, social skills training, psychotherapy, and educational assistance); or methylphenidate with a psychosocial attention control treatment. Congruent with the MTA findings, combination treatment did not lead to superior functioning compared to methylphenidate alone, and all treatment groups demonstrated significant improvement that continued over 2 years. Subsequent follow-up of this sample out to 8 years has been reported including both effectiveness and safety in multiple domains.

PRESCHOOL ADHD TREATMENT STUDY (PATS). Despite current FDA warnings against the use of methylphenidate in children under the age of 6, significant numbers of young children with ADHD are treated with this medication. PATS, funded by the NIMH, enrolled 303 moderately to severely impaired preschoolers with ADHD. Less than 10 percent of the children responded to an initial course of parent training alone, and ultimately 165 were initiated on pharmacotherapy. Mean optimal dose of immediate-release methylphenidate, dosed t.i.d., was 14 mg/day. This dose, 0.7 mg/kg/day, is lower than the mean optimal dose of 1.0 mg/kg/day in the MTA study. While effective, the effect size was smaller than that found in school-aged children in the MTA study, perhaps due at least in part to the conservative dosing. Parents spontaneously reported moderate to severe adverse effects in 30 percent of the preschoolers,

including emotional outbursts, initial insomnia, repetitive behaviors/thoughts, decreased appetite, and irritability. Eleven percent of the children discontinued treatment because of intolerable methylphenidate adverse effects, compared to <1 percent of school-age children in the MTA study.

Clinical Implications and the Role of Pharmacotherapy. The combined data from the two comparator studies provides a direct comparison of well-delivered pharmacotherapy, intensive behavioral therapy, and a combination of the two. The significant improvement in ADHD symptoms, which were most dramatic in the arms with well-articulated pharmacotherapy, established pharmacological treatment for ADHD as superior to behavioral therapy alone. This role was reinforced in PATS, which demonstrated limited improvement in moderately to severely ill preschoolers with ADHD when only parent training was implemented. These data provide strong support for the role of medication as a fundamental consideration in a comprehensive treatment plan for children with ADHD. The expansion of findings to include the nonstimulants highlights the role of these agents alone or in combination with stimulants in the armamentarium for ADHD.

Oppositional Defiant Disorder

Another commonly encountered mental health disorder in children and adolescents, oppositional defiant disorder (ODD) generally presents with a complex etiology. It is often comorbid with other psychiatric disorders, such as ADHD, and can precede the development of conduct disorder, substance abuse, and severe delinquent behavior.

Pharmacotherapies for ODD. Studies addressing the use of medication for ODD per se are limited, although trials of extended release mixed amphetamine salts (Adderall XR) demonstrated improvements in ODD symptoms in youth with comorbid ADHD as well as those without ADHD. Several trials of stimulants, both amphetamines as well as methylphenidate, have demonstrated improvements of co-occurring oppositional behaviors within the context of ADHD. While an initial trial with atomoxetine suggested a similar benefit in ODD symptoms in youth with ADHD, subsequent work was not able to replicate that finding. Similarly, studies with extended-release guanfacine (Intuniv) and tricyclic antidepressants (nortriptyline) have shown improvement in ODD symptoms associated with ADHD. Several studies of typical and atypical antipsychotic agents have demonstrated benefit as adjuncts to the treatment of aggression in children with mental retardation and pervasive developmental disorders.

Clinical Implications and the Role of Pharmacotherapy. The AACAP ODD practice parameters recommend a multimodal treatment approach incorporating individual and family psychosocial interventions, with a specific emphasis on using one of the empirically tested parent training interventions. Medication often serves as an adjunct to the overall treatment program. In this role, pharmacotherapy is generally used either to address specific ODD symptoms or to treat comorbid conditions. Treatment of comorbid disorders (e.g., mood or ADHD) may also improve ODD symptoms.

Conduct Disorder

Conduct disorder has only been studied in a small number of controlled trials that generally targeted the aggressive behaviors that are frequently present.

Pharmacotherapies for Conduct Disorder. The limited number of studies of mood stabilizers such as divalproex sodium (Depakote) and lithium carbonate (Eskalith, Lithobid), antipsychotics, and the stimulants have demonstrated some promise. Atypical antipsychotics are among the most commonly prescribed medications for acute as well as chronic irritability and aggression. There does appear to be a difference in response to pharmacotherapy based on the type of aggression however. While the aggressive symptoms associated with dysphoric conduct problems may be responsive to pharmacotherapy, the predatory aggression is much less likely to improve with medication.

Clinical Implications and the Role of Pharmacotherapy. In light of the wide range of diagnoses that may be complicated by aggressive and oppositional behaviors, it is generally recommended that, when present, comorbid disorders should be the primary target of pharmacotherapy.

MOOD DISORDERS

Major Depressive Disorder

Mood changes are normal in childhood and adolescence; however, prolonged sadness, anhedonia, irritability, and potentially self-harmful or suicidal behaviors are not, and occur in up to 10 percent of youth. A thorough assessment of biological and psychosocial factors that may result in mood symptoms is warranted.

Pharmacotherapies for Major Depression. With multiple RCTs lending support, fluoxetine (Prozac) and escitalopram (Lexapro) are the antidepressants with the greatest amount of data supporting their safety and efficacy in the pediatric population—and both are FDA approved for

MDD in pediatric groups. There is also support for the use of medication in conjunction with cognitive-behavioral therapy (CBT) for initial outcome and to reduce the risk of relapse. In a recent RCT feasibility study, 46 youths (aged 11 to 18 years) who had responded to 12 weeks of fluoxetine treatment were treated with 6 months of continued medication management with and without relapse prevention CBT. The addition of CBT substantially reduced the risk of relapse (eightfold).

Other SSRIs including citalopram (Celexa) and sertraline (Zoloft) have some support in the literature as well, but to a more limited extent. Bupropion has open data supporting its use. Tricyclic antidepressants, however, have repeatedly demonstrated benefits for the treatment of adults with MDD, but have no controlled pediatric trial data to date showing benefit. Monoamine oxidase inhibitors have historically had very limited data surrounding pediatric use, which, combined with concerns regarding dietary issues and hypertension, have likely led to the limited use in this population.

SUICIDALITY. Over a decade ago, an FDA advisory Committee reviewed results of a meta-analysis of 24 controlled clinical trials of 9 antidepressants, which included approximately 4,400 pediatric patients. While there were no completed suicides, the cumulative risk for suicidality (suicidal thinking or behavior) when adverse events were simply collected by spontaneous report was approximately 4 percent on active medication vs. approximately 2 percent on placebo. This resulted in the FDA issuing a "black box" warning that any antidepressant use in children and adolescents was accompanied by an increased risk for suicide, and in mandatory distribution of a medication guide with each pediatric antidepressant prescription filled. Also, for the initial 3 years, the FDA had guidelines in place for the frequency of monitoring when initiating antidepressants. Importantly, in the Treatment for Adolescents with Depression Study (TADS), suicidality data was systematically collected, and questions regarding suicidality were posed to the youth prior to and during participation of the study, rather than relying on spontaneous report. When data was collected in this manner, almost 30 percent acknowledged recent thoughts or behaviors related to self-harm at the start of the study. All four treatment groups (CBT, fluoxetine, their combination, and pill placebo) showed diminished suicidality during the study with systematic data collection, although fluoxetine demonstrated the smallest reduction.

Select Key Studies

THE TREATMENT FOR ADOLESCENTS WITH DEPRESSION STUDY (TADS) . In this NIMH-sponsored study, 439 adolescents (aged 12 to 17 years) with moderate to severe depression were randomly assigned to one of four treatments: fluoxetine with medical management, weekly CBT, their combination, and

pill placebo. Adolescents who received fluoxetine and combination therapy had significant improvements at 12 weeks, while those receiving CBT alone did not separate from placebo. This demonstrated the key role of pharmacotherapy in the treatment of MDD in adolescents; however, when parameters beyond the primary outcome measures were examined, the combination treatment was most successful acutely. Interestingly, by the end of 9 months of treatment, combination, fluoxetine, and CBT responses were all virtually identical. A subsequent study led by Emslie demonstrated that continuation treatment with fluoxetine was superior to placebo in preventing relapse in children and adolescents with MDD.

THE TREATMENT OF SSRI-RESISTANT MDD IN ADOLESCENTS (TORDIA). This trial enrolled a more chronically depressed and treatment resistant sample than TADS. TORDIA is an RCT of 334 adolescents (12 to 18 years) with MDD that had not responded to a 2-month initial trial with an SSRI. Subjects received 12 weeks of: a second, different SSRI (paroxetine [Paxil], citalopram, or fluoxetine); a different SSRI plus CBT; venlafaxine (Effexor); or venlafaxine plus CBT. The two arms with CBT plus medication demonstrated a higher response rate (54.8 percent) than a medication switch alone (40.5 percent), with no difference in response rate between venlafaxine and a second SSRI. TORDIA demonstrated that, for adolescents with depression who do not respond to an initial SSRI, a switch to another antidepressant, combined with CBT, should be considered. Even if CBT is not feasible, those simply changing medications had a 40.5 percent improvement.

Clinical Implications and the Role of Pharmacotherapy. For managing "mild," uncomplicated pediatric depression, the AACAP practice guidelines recommend consideration of education, support, and case management. For children and adolescents who present with a more complex MDD, or have a limited response to supportive psychotherapy, a more specific psychotherapy and/or an antidepressant is indicated. Severe MDD and/or chronic recurring MDD will generally require pharmacotherapy, either alone or in combination with psychotherapy.

Dysthymia

Dysthymia is a persistent, long-term mood disorder of less intensity than MDD, with a mean episode length of 3 to 4 years. While this is a relatively common disorder, particularly in adolescents, there are no controlled trials for the treatment of dysthymia in youth. Based on the available adult literature, efficacious treatments for MDD may also be effective for dysthymia.

Pharmacotherapy for Dysthymia. In the absence of any controlled pharmacotherapy trials for dysthymia, when clinical presentation indicates

that initiation of an antidepressant is warranted, selecting an antidepressant with support for MDD appears prudent.

Clinical Implications and the Role of Pharmacotherapy. While the AACAP Practice Parameters do not provide specific guidance on the role of pharmacotherapy for dysthymic disorder, their general guidance for MDD would appear to be applicable. In light of the duration of symptoms required to meet diagnostic criteria for dysthymia (1 year), it is likely that many of the youths with dysthymia would require specific types of psychotherapy and/or pharmacotherapy as recommended by the guidelines for those who do not respond to supportive psychotherapy or have a complicated MDD.

Bipolar Disorder

Once considered rare, childhood-onset bipolar disorder is now diagnosed with much greater frequency (up to 4 percent) in children and adolescents. While controversy continues regarding the diagnostic criteria for pediatric bipolar disorder, it is clear that a number of youth suffer from explosive dysregulated mood, emotional lability, and manic symptomatology leading to significant impairment in functioning and, potentially, hospitalization. These children and adolescents frequently demonstrate rapid fluctuations in mood and behavior (ultrarapid or ultradian cycling) rather than clearly delineated manic and depressive phases. Bipolarity in youth is commonly accompanied by comorbid ADHD and conduct and anxiety disorders, as well as substance use disorders in adolescents. Juvenile bipolar disorder is a highly remitting and relapsing disorder with some estimates indicating that the disorder will persist in three-quarters of children and adolescents into adulthood.

Pharmacotherapies for Bipolar Disorder. Fortunately, despite a presentation that may be quite different to that seen in adults, several pharmacotherapies effective for adults with bipolar disorder have demonstrated benefits in youth. These pediatric mood symptoms are generally treated with atypical antipsychotics, which are both FDA approved and now considered first-line treatment.

ATYPICAL ANTIPSYCHOTICS. Olanzapine, risperidone, aripiprazole, and quetiapine are FDA approved for bipolar disorder in children and/or adolescents, with trials of other atypicals underway. The atypical antipsychotics typically manifest their action within 2 weeks and maintain effectiveness over time. Mood stabilizers are often utilized as second line treatments or used in conjunction with atypical antipsychotics. In addition to demonstrating improved manic symptoms, these agents may assist in the prevention of recurrent manic episodes and the depressive subphase of

the bipolarity. While monotherapy is certainly desired, combination pharmacotherapy is often required—such as an atypical antipsychotic and mood stabilizer, or an antimanic agent with treatment for comorbid ADHD, anxiety, and/or OCD. Additionally, once a youth with bipolar disorder responds to a pharmacotherapy, most will require ongoing treatment with medication to prevent relapse. Current practice parameters suggest the regimen needed to stabilize the acute mania should be maintained for 12 to 24 months.

LITHIUM. Lithium is FDA approved for bipolar disorder in youth 12 years of age and older, although this approval was based primarily on the adult literature with limited controlled clinical trial data supporting its use. Open-label and controlled trials of lithium with and without concomitant antipsychotics have been positive, although the effect sizes and onset of action appear less robust than atypical antipsychotics.

ANTICONVULSANTS. Early reports derived from case series and open studies showed the usefulness of the antiseizure medications divalproex and carbamazepine (Tegretol). However, controlled studies with either divalproex or oxcarbazepine (Trileptal), alone or in combination with other medications, failed to demonstrate efficacy of the anticonvulsants for the treatment of acute mania. In addition, their teratogenic and metabolic effects have limited their usefulness. The role of anticonvulsants in combination with atypical antipsychotics has yet to be further described.

Select Key Studies. Results from multisite studies of olanzapine for juvenile bipolar disorder were among the first data published demonstrating efficacy of any medication for bipolar disorder in this age group. For instance, in a 3-week RCT in 161 children and adolescents aged 13 to 17 years, olanzapine significantly improved symptoms of mania compared to placebo. Similar results were reported with quetiapine monotherapy in 277 patients aged 10 to 17 years. Additionally, a 4-week RCT of flexible-dose ziprasidone demonstrated improvement at week 1, suggesting early control of the symptoms. In most of the studies of atypical antipsychotics, weight gain has exceeded that noted in adult studies. Moreover, changes in prolactin, fasting glucose, fasting total cholesterol, uric acid, and hepatic enzymes have been reported in youth receiving atypical antipsychotics, necessitating close monitoring (at least bi-yearly to yearly).

Clinical Implications and the Role of Pharmacotherapy. The AACAP practice parameters describe pharmacotherapy as the primary treatment for well-defined DSM-IV-TR mania. Adjunctive treatment with behavioral and psychosocial therapies may be useful to address disruptive symptoms and their sequela, which often impact family and community. Due to the

difficulty of stabilizing active bipolar and frequent comorbid symptoms, atypical antipsychotics are recommended as first line treatments, although it is often necessary to employ combined pharmacotherapeutic strategies in these youth. For instance, multiple studies have shown that stimulants can be used in youth with stable bipolar disorder to treat comorbid ADHD.

ANXIETY DISORDERS

Obsessive-Compulsive Disorder (OCD)

A number of youth (1 to 2 percent) suffer from OCD characterized by rituals and obsessions with impairment that often continue into adulthood. In fact, one-third to one-half of adults with OCD are able to identify the onset of symptoms in childhood or adolescence.

Pharmacotherapy for OCD. A number of SSRIs have demonstrated efficacy and been approved by the FDA for the treatment of children and adolescents with OCD. These medications include fluvoxamine (Luvox), fluoxetine, clomipramine (Anafranil), and sertraline. It is noteworthy that serotonin-based medications are the only class of antidepressants to date to demonstrate efficacy in the treatment of OCD. The atypical antipsychotics and benzodiazepines may have an adjunctive role in the treatment of refractory OCD, but there is no support for these medications as a first-line treatment for pediatric OCD.

For cases of acute, sudden-onset OCD, particularly with a corresponding positive test for strep, treatment may increase in complexity. If a clinical diagnosis of PANS/PANDAS is appropriate, treatment efforts may require a two-sided approach. The first aim is to treat the underlying, aggravating infection with antibiotics or intravenous immunoglobulin (IVIG). Second, CBT and/or SSRIs in lower doses than normally recommended for children with OCD may prove beneficial in addressing the residual OCD, tics, anxiety, or accompanying emotional lability.

Select Key Studies: The Pediatric OCD Treatment Study (POTS). Both short-term OCD-specific CBT and the SSRIs have demonstrated efficacy in the treatment of pediatric OCD. POTS was the first study to examine the relative and combined efficacy of these two approaches, randomizing 112 patients with OCD (aged 7 to 17 years) to treatment with CBT alone, medical management with sertraline alone, the combination of the two, and pill placebo. All three active treatments were superior to placebo in reducing OCD symptoms, although clinical remission rates for combined treatment was 53.6 percent, for CBT alone 39.3 percent, for sertraline alone 21.4 percent, and for placebo only 3.6 percent.

The POTS (II) RCT assessed 124 pediatric outpatients with OCD (aged 7 to 17 years) assigned to 1 of 3 treatment strategies, all of which included

SSRIs. The positive response rates for medication management only (30.0 percent) and medication management plus instructions in CBT (34.0 percent) were significantly lower than combined medication management plus CBT (68.6 percent). These findings indicate that truncated versions of CBT are not efficacious comparatively, and that efforts should focus on making the full CBT protocol available more widely in community settings.

Clinical Implications and the Role of Pharmacotherapy. The empirical literature clearly supports the role of pharmacotherapy for the acute and maintenance treatment of pediatric OCD. In light of the POTS I/II findings, recommendations would be for children and adolescents with OCD to begin treatment with either CBT alone or in combination with an SSRI.

Posttraumatic Stress Disorder

Approximately one in four children will experience a significant traumatic event prior to reaching adulthood, with a minority who develop subsequent acute as well as chronic mental health problems. Risk factors for developing posttraumatic stress disorder (PTSD) following exposure to trauma include: prior trauma exposure, degree of exposure to index trauma, pre-existing psychiatric disorders, female gender, parental psychopathology, and limited social support.

Pharmacotherapy for PTSD. There are no randomized clinical trials published demonstrating efficacy of any pharmacotherapy in treating PTSD in children. SSRIs have certainly demonstrated efficacy in the treatment of adults with PTSD, but their efficacy remains dubious in pediatric populations with PTSD. Other medications that have demonstrated at least some evidence of support in open-label pediatric clinical trials include: atypical antipsychotics, alpha- and beta-adrenergic blocking agents, non-SSRI antidepressants, and mood stabilizers.

Clinical Implications and the Role of Pharmacotherapy. The AACAP PTSD practice parameters recommend that trauma-focused psychotherapies be considered first-line treatment for children and adolescents with PTSD, with SSRIs considered a potential treatment option (acceptable based on emerging evidence or clinical opinion) when used in conjunction with psychotherapy. New research on the role of "prophylactic" very early intervention with beta blockers and alpha agonists to protect against the long-term development of PTSD after trauma are under investigation.

"Other" Anxiety Disorders (Separation Anxiety Disorder, Generalized Anxiety Disorder, Social Phobia, Specific Phobia, Panic Disorder, Agoraphobia without Panic Disorder)

The anxiety disorders as a group are the most common of the pediatric psychiatric disorders. Youth with an anxiety disorder are at increased risk of additional anxiety disorders, MDD, and substance abuse, with an increased risk of anxiety and depressive disorders in adulthood as well.

Pharmacotherapy for Anxiety. SSRIs are considered the medication of choice in the treatment of children and adolescents with anxiety disorders when the symptoms are moderate to severe, if the impairment makes participation in psychotherapy difficult, or if psychotherapy alone leads to only a partial response. Several RCTs of SSRIs have demonstrated benefits in the short-term treatment of GAD, social phobia, separation anxiety disorder, and selective mutism with social phobia. The long-term role of SSRIs has not been studied, so clinicians might consider a trial off medication after a year of stable symptom improvement. There is no evidence of greater efficacy of one SSRI over another at this time.

Key Studies: The Child/Adolescent Anxiety Multimodal Treatment Study (CAMS). This trial helped to delineate the appropriate role for pharmacotherapy in anxiety disorders by comparing sertraline, CBT, their combination, and placebo in youths with separation anxiety disorder, social phobia, and GAD. In the trial, 488 children aged 7 to 17 years were randomized to one of the four treatment options for a 12-week period. Following this acute exposure, responders attended 6 monthly booster sessions in their treatment arm, with sertraline and combination groups continuing on their medication throughout this period. At both weeks 24 and 36, the majority (>80 percent) of acute responders in all three treatment conditions maintained positive response. The combination group maintained superiority on some but not all outcomes, whereas the monotherapies were indistinguishable across all outcomes studied, leading to the conclusion that all 3 CAMS treatments are viable alternatives.

Another significant study in pediatric anxiety pharmacotherapy was conducted by an RCT of 128 children, aged 6 to 17 years, with social phobia, separation anxiety disorder, or GAD, all of whom had failed to improve with 3 weeks of a psychosocial intervention. Following 8 weeks on fluvoxamine dosed up to 300 mg per day or placebo, fluvoxamine was found to significantly improve outcomes compared to placebo on both the Pediatric Anxiety Rating Scale and the Clinical Global Impressions-Improvement Scale.

A recent systematic review and meta-analysis investigated the efficacy, safety, and tolerability of SSRIs and SSNRIs for the treatment of pediatric anxiety disorders. Nine randomized, double-blind, placebo-controlled trials involving 1,673 patients (and 6 medications) were included. Analyses demonstrated efficacy for all individual agents with an overall modest effect size compared to placebo. The authors suggested that treatment selection

by providers should reflect tolerability, half-life, and cost, rather than short-term efficacy.

Clinical Implications and the Role of Pharmacotherapy. The AACAP Anxiety Practice Parameters recommend initial treatment of anxiety disorders in mildly ill children with psychotherapy. Exposure-based CBT has received the most empirical support of the psychotherapies in the treatment of pediatric anxiety. The addition of a medication may be in response to a need for acute symptom reduction in the moderately to severely anxious child, a comorbidity requiring pharmacotherapy, or a partial response to psychotherapy.

SCHIZOPHRENIA

While the onset of schizophrenia is rare prior to 13 years of age, numbers steadily increase throughout the adolescent years, and by 18 years of age 12 to 33 percent of patients who will ultimately develop schizophrenia will have symptoms.

Pharmacotherapy for Schizophrenia

Following several decades of practice based primarily upon extrapolation of data from research in adults with schizophrenia, RCTs of aripiprazole and risperidone have demonstrated the safety and efficacy of these medications in the treatment of schizophrenia in adolescents aged 13 to 17 years. The studies were sufficient to warrant FDA approval for these drugs in this teen population. At this point the atypical antipsychotics, particularly aripiprazole and risperidone, would be considered first-line medications for adolescents with schizophrenia. Studies involving clozapine (Clozaril) in treatment-refractory early-onset schizophrenia-spectrum disorders have demonstrated superiority of clozapine compared to haloperidol (Haldol) and olanzapine, with regard to reduction in negative symptoms and to a certain degree with overall response. Based on the results of one RCT, the use of omega-3/fish oils may help delay the onset and/or reduce the ultimate expression of psychotic symptoms in high-risk children.

Clinical Implications and the Role of Pharmacotherapy

Psychopharmacology plays a critical role in the treatment of schizophrenia, with the atypical antipsychotics considered among the first-line treatments. Fish oil may be useful in adolescents at high risk to develop fulminant psychotic symptoms. Psychosocial interventions are considered adjunctive to the medication, but are strongly recommended.

SUBSTANCE USE DISORDERS

The lifetime prevalence of drug abuse or dependence increases rapidly

throughout adolescents, jumping from 3.3 percent in 15-year-olds to 9.8 percent in 17- to 19-year-olds. The primary goal of treatment in adolescents with substance use disorder (SUD) is reducing substance use (harm reduction) or ultimately abstinence from substance use. Up to 90 percent of adolescents with SUD have comorbid psychopathology.

Pharmacotherapy for SUD

Guidance on the use of pharmacotherapy for SUD in adolescents is based primarily on controlled trials in adults. Medications targeting alcohol-related cravings, such as naltrexone (ReVia), acamprosate (Campral), and ondansetron (Zofran), have been described in a limited manner with generally open trials in adolescents. Opioid replacement such as buprenorphine (Suboxone) is FDA approved in adolescents older than 16 years and may be considered in opioid addicted youth who have failed a trial of abstinence. These medications, along with aversive agents such as disulfiram (Antabuse), are recommended to be considered for use in adolescents who fail or refuse psychotherapy. Pharmacotherapy has been shown to be useful for the treatment of the comorbid psychiatric disorders such as mood disorders, but is generally used as adjunct to SUD-based psychotherapies.

Clinical Implications and the Role of Pharmacotherapy

A number of psychotherapies including family-based treatments and individual cognitive-behavioral therapies with motivational enhancement have solid empirical support for the treatment of adolescents with SUD, and are recommended as first line for SUD. Practice parameters describe a role for medication in the management of craving, withdrawal, and for aversion therapy, but this is based primarily upon expert opinion and open studies. Treatment of comorbid psychiatric disorders should be considered concurrent with issues such as understanding the comorbid disorder, the latency to initiate treatment, selection of agents, and monitoring on a case-by-case basis. Whereas pharmacotherapy of psychotic and mood disorders should be undertaken in parallel with the onset of addiction treatment, ADHD and/or anxiety treatment should be sequenced after some stabilization in substance use.

AUTISM SPECTRUM DISORDERS (ASD)

Pharmacotherapy for Autistic Spectrum Disorders (ASD)

ASDs, including autistic disorder and the older DSM-IV Asperger disorder and pervasive developmental disorder PDD-NOS, are lifelong neurodevelopmental disorders characterized by core features of impaired socialization, communication, behavioral problems, and associated features of anxiety, irritability, aggression, and ADHD-like symptoms. Recent

prevalence rates have been increasing with more recent studies conducted in Asia, Europe, and North America reporting an average of around 1 percent in the general population. Though emerging evidence suggests efficacy of psychopharmacological interventions in improving core features of ASD such as ritualistic behaviors, there is definitive evidence for the therapeutic role of psychotropic intervention for the associated features of ASD, such as irritability and disruptive and aggressive behaviors. Risperidone and aripiprazole have been approved by the FDA for the treatment of irritability associated with autistic disorder in children and adolescents. With the advent of DSM-5, comorbidity with ADHD and other disorders may be better identified and treated within ASD.

Key Studies

Research Units on Pediatric Psychopharmacology Autism Network: Risperidone Trial. This 8-week RCT of risperidone in 101 children with autistic disorder (mean age 8.8 ± 2.7 years) showed that risperidone was superior to placebo in the short-term treatment of irritability, hyperactivity and stereotypic/repetitive behaviors, although there was a lack of response in social and communication deficits. Common adverse effects of increased appetite, fatigue, drowsiness, dizziness, and drooling were reported. Significant improvements in irritability, hyperactivity, and repetitive behaviors observed in the acute treatment with risperidone were maintained in the continuation phase of the trial without an increase in the mean daily dose of risperidone (1.8 ± 0.7 mg; range 0.5 to 3.5).

In an 8-week RCT of 218 children and adolescents (aged 6 to 17 years) with a diagnosis of autistic disorder, participants randomized to aripiprazole (5, 10, or 15 mg/day) had significantly greater improvements in irritability than placebo (but only 5 mg dose was statistically significant due to high placebo response). The most commonly reported adverse events associated with aripiprazole treatment were sedation, EPS, and weight gain.

Research Units on Pediatric Psychopharmacology Autism Network: Methylphenidate Trial. This RCT examined the efficacy and safety of low (0.125 mg/kg/day), medium (0.25 mg/kg/day), and high (0.125 mg/kg/day) doses of methylphenidate in 72 children (mean age 7.5 ± 2.2 years) with moderate to severe hyperactivity in addition to ASD. Of 66 subjects who tolerated the trial doses of methylphenidate, the medication was significantly superior to placebo in improving ADHD symptoms with a small to medium magnitude of response (effect size range 0.20 to 0.54). The medication was associated with decreased appetite, sleep difficulties, and emotional outbursts, and 18 percent discontinued the trial due to inability to tolerate methylphenidate.

Clinical Implications and the Role of Pharmacotherapy

As demonstrated in the methylphenidate trial, compared to a typically developing population, response to psychotropic medications in youth with ASD may be less robust with higher rates of adverse effects to both medication and placebo. Thus, it is advisable to be conservative in dosing during the initiation and titration of psychotropic medications in this population. SSRIs including clomipramine and fluoxetine appear to be modestly effective in improving overall global functioning, including symptoms of repetitive behaviors, anxiety, irritability, and hyperactivity in children with ASD. Mood dysregulation in the form of irritability, aggression, temper tantrums, and self-injurious behaviors is commonly associated with ASD. Atypical antipsychotics are not only effective in treating the ASD associated features of irritability and hyperactivity and core feature of ritualistic behaviors, but additionally appear useful and are suggested to be superior to first-generation antipsychotics in improving the other ASD core features of impaired socialization and communication. Stimulants are the most widely prescribed drug for target symptoms of ADHD symptoms in youth with ASD, although the rate and magnitude of response to stimulants in ASD children is less than observed in children with ADHD. Atomoxetine or low dose of stimulants is recommended for the treatment of ADHD in children with ASD.

SLEEP DISORDERS

Insomnia

Insomnia is common in the general pediatric population, and is certainly more common in youth with psychiatric disorders due to the disorders themselves as well as a sequela of the medications prescribed to treat the primary disorder.

Pharmacotherapy for Insomnia. There are no FDA-approved sleep agents for children. After aggressive sleep hygiene and education, controlled trials have demonstrated improved sleep latency outcomes with melatonin. Data regarding other agents used chronically are largely derived from expert opinion or case series, including benzodiazepines, clonidine, mirtazapine (Remeron), imipramine, trazodone (Desyrel), and, less frequently, sedating atypical antipsychotics.

Clinical Implications and the Role of Pharmacotherapy. Pharmacotherapy should be considered only after a thorough assessment for the etiology of the sleep disruption (e.g., sleep hygiene, potential iatrogenic causes, sleep apnea, insomnia as a sequela of a psychiatric disorder). For pharmacologically induced sleep disruption, consider alternative preparations, classes of medication, or medication

discontinuation. In the case of psychopathology related sleep issues or necessary medication-related sleep disruption, use of medication should be considered.

Enuresis

While a specific etiology for enuresis is not commonly identified, a careful assessment is necessary in consideration of potential urologic, developmental, psychosocial, and/or sleep-related causes. If the history, physical examination, and urinalysis are unremarkable and do not suggest a specific etiology, supportive approaches combined with conditioning is considered the first-line treatment. Supportive interventions include education as well as recommendations against punishing the child for enuretic episodes. Behavioral conditioning methods with portable battery-operated alarms have been repeatedly demonstrated to be highly effective.

Pharmacotherapy for Enuresis. Two medications have demonstrated efficacy in the treatment of enuresis.

DESMOPRESSIN (DDAVP). Desmopressin is a synthetic analog of vasopressin, which decreases urine production. It is generally given orally, dosed initially at 0.2 mg at bedtime, titrated up to 0.4 to 0.6 mg as needed. Success rates of 10 to 65 percent have been reported. If successful, a course of treatment is recommended for 3 to 6 months.

IMIPRAMINE. Imipramine is FDA approved for enuresis, and has been used for decades despite an unknown mechanism of action, with several studies demonstrating 40 to 60 percent effectiveness. It is given in a single bedtime dose of 1.0 to 2.5 mg/kg/day. If effective, treatment should be continued for 4 to 6 months. In light of the risk of cardiac arrhythmias with tricyclic antidepressants, an ECG can be considered prior to treatment and periodically during treatment. Serum levels are not recommended unless there is no effect despite dosing the imipramine at 2.5 mg/kg/day.

Clinical Implications and the Role of Pharmacotherapy. If supportive approaches and behavioral conditioning with a urine alarm are ineffective or impractical, medication may be considered as a treatment option. Medication may also be useful intermittently to allow youth to attend overnight activities such as camp or sleepovers without embarrassment.

Tic Disorders

Tic disorders are common in youth, occurring in 15 percent of the general population, and are more common in youth with other psychiatric disorders. Most tic disorders do not affect functioning adversely, are transient, abate with age and do not require intervention. Tic disorders

may include motor and/or vocal tics (the combination referred to as Tourette syndrome) and may be transient or chronic, simple or complex. Most tic disorders present *de novo,* but some may be exacerbated by stress or with certain classes of medication.

Pharmacotherapy for Tic Disorders. The antipsychotic pimozide is FDA approved for managing tics. Based on controlled clinical trials, the current armamentarium for tic and Tourette disorders includes the alpha-agonists guanfacine and clonidine, as well as the typical (e.g., pimozide [Orap] or haldol) and atypical antipsychotics (risperidone, aripiprazole). Multisite studies have demonstrated the efficacy and tolerability (e.g., non–tic exacerbation) of combination methylphenidate and clonidine in treating children with ADHD and comorbid tics. Specific psychotherapies are being tested to assist the child in focusing on symptoms of tics to help abate tic expression.

Clinical Implications and the Role of Pharmacotherapy. Tic and Tourette symptoms should be evaluated carefully to establish the level of resulting impairment. For youth with impairment related to their tics, treatment with pharmacotherapy should be considered. Give the high rates of disabling anxiety, OCD, and ADHD, consideration should be given to treating the comorbidity with careful selection of non–tic exacerbating agents. For instance, ancillary treatment of co-occurring anxiety may reduce the severity of tics and reduce the medication burden.

ENDOCRINE AND METABOLIC EFFECTS OF PSYCHOTROPIC MEDICATIONS

Accompanying the expanding use of psychotropic medications in the pediatric population is a growing awareness and understanding of the potential for endocrine and metabolic effects caused by many of these agents. Psychostimulants, antipsychotics, and mood stabilizers are some of the common classes of medication that have demonstrated these types of adverse effects.

Antipsychotics

As discussed previously, antipsychotics are used for a wide range of psychiatric disorders, including mood, disruptive behavioral, tic, anxiety, and psychotic disorders. The use of antipsychotics, particularly atypical antipsychotics, has grown tremendously over the past decade, and clinicians are faced with managing the metabolic sequela of these medications.

Weight Gain and Metabolic Issues. Children and adolescents appear to be susceptible to significant weight gain and metabolic adverse effects of the atypical antipsychotics, perhaps even more so than adults. While all of

the atypical antipsychotics may promote weight gain, particularly in patients who are treatment-naïve, some seem to be more likely than others. Correll has suggested the propensity to promote weight gain and the subsequent development of a metabolic syndrome in the pediatric population to be: clozapine = olanzapine >> risperidone ≥ quetiapine > ziprasidone ≥ aripiprazole. Body mass index (BMI) is a way to express body weight adjusted for height, and is calculated as:

$$\text{BMI} = (\text{weight in kg})/(\text{height in meters})^2$$

Up-to-date BMI tables and charts containing normal values for age are readily available from the Centers for Disease Control (www.cdc.gov/growthcharts).

Some patients who gain significant weight will develop lipid abnormalities, including elevations in triglycerides, total cholesterol, and low-density lipoprotein-cholesterol levels. Additionally, weight gain and obesity is associated with the metabolic syndrome, which includes abdominal obesity, dyslipidemia, glucose intolerance, hypertension, and insulin resistance. Clinical strategies for addressing this issue include education, reinforcement of healthy lifestyle behaviors, selecting a medication with lower likelihood of adverse effects, and intervening if weight gain and its sequela do become problematic. Monitoring of BMI, height, weight, blood pressure, waist circumference, fasting blood glucose, and fasting serum lipids should begin at baseline prior to initiating medication and continue throughout treatment with antipsychotics. For patients who do require ongoing treatment despite metabolic abnormalities, there is evidence that metformin (Glucophage) might be beneficial.

When assessing the risk/benefit relationship of initiating these medications, it is important to realize that not all children will develop these metabolic problems, and that those of normal BMI may be less prone to develop difficulties. Additionally, the clinician must factor in the fact that the greater the impairment in a child's functioning, the greater the risk of not treating.

Prolactin. Elevations in prolactin can occur with atypical antipsychotics as well as the conventional neuroleptics, although the degree of elevation varies by medication and dose. The relative potency of antipsychotic medications in inducing hyperprolactinemia, based on pediatric and adult data, is roughly: risperidone > haloperidol > olanzapine > ziprasidone > lurasidone > asenapine > clozapine > quetiapine > aripiprazole. Fortunately, while the elevation in serum prolactin is common, particularly in children and adolescents, it is often asymptomatic and frequently will decrease over time even despite continuation of

medication. Potential signs and symptoms of hyperprolactinemia include: amenorrhea or oligomenorrhea in women of reproductive age, breast enlargement/engorgement, galactorrhea, decreased libido, erectile dysfunction, and failure to enter or progress through puberty.

Stimulants and Atomoxetine

Both the stimulants and atomoxetine appear to cause a mild growth delays in certain patients; however, data suggest that height and weight delays largely attenuate by 2 to 3 years into treatment. While the effect on final adult height seems to be negligible, data following patients into adulthood is sparse. In light of this, the AACAP practice parameters recommend measuring height and weight and plotting the information on a growth chart, prior to and throughout treatment. For those patients who do experience more significant weight or height delays, clinicians may consider caloric supplementation, breaks from medication, alternative agents, and/or combinations (e.g., Periactin).

Mood Stabilizers

Lithium. Data is limited regarding lithium's potential for adverse effects on thyroid and renal functioning for children and adolescents specifically, although the potential is well documented for adults. Considering the possibility for chronic use in a child who responds to lithium treatment, monitoring for thyroid and renal dysfunction over time is essential.

Valproate. Valproate use has been associated with weight gain as well as polycystic ovary syndrome, and may be related to teratogenic effects in pregnancy. While the specific mechanism and risk factors for polycystic ovary syndrome are unclear, clinicians should monitor for weight gain, menstrual dysfunction, hirsutism, and acne in female patients taking valproate. Combining concerns of polycystic ovary disease and teratogenicity, many clinicians avoid valproate in adolescent females.

PSYCHOPHARMACOLOGICAL TREATMENT FOR VERY YOUNG CHILDREN

The Preschool Psychopharmacology Working Group, a multidisciplinary collaboration of early childhood professionals, released recommendations to guide clinicians considering psychopharmacological treatment in young children. In light of the limited evidence base for the use of most psychotropic medications in children under 6 years of age, the group recommended that a trial of evidence-supported psychosocial treatments precede psychopharmacological treatments, and concluded that medication was not indicated for preschoolers with only mild or single-context symptomatology or impairment. If medication is ultimately indicated, the

group made the following recommendations based upon the limited data available.

ADHD

In light of data from PATS and several smaller methylphenidate studies, as well as the lack of randomized controlled trials of amphetamines in preschoolers, methylphenidate is recommended as the first-line psychopharmacological treatment for preschool ADHD.

Disruptive Behavioral Disorders

Risperidone is recommended as the first-line medication for young children with disruptive behavioral disorders and concomitant severe aggression without co-occurring ADHD.

Major Depressive Disorder

Fluoxetine is recommended as the first-line medication for preschoolers with major depressive disorder.

Bipolar Disorder

Risperidone is recommended as the first-line medication, as it has the most data available on effectiveness and tolerability in preschool-age children with bipolar disorder.

Anxiety

Fluoxetine is the first-choice medication for preschool anxiety disorders in general, although for PTSD the working group did not recommend the use of medication and for OCD there was insufficient evidence to suggest one was more likely to be efficacious than the other between fluoxetine, fluvoxamine, and sertraline.

Autistic Spectrum Disorders

Risperidone is approved by the FDA for treating children as young as 5 with autism and severe aggression and irritability.

Primary Sleep Disorders

If the sleep disturbance is significant enough that the well-being and daytime functioning of the child and/or caregiver is compromised, and if behavioral treatment is unsuccessful, then melatonin 1 to 3 mg could be considered for use as the first-line pharmacotherapy.

References

Abikoff H, Hechtman L, Klein RG, et al. Symptomatic improvement in children with ADHD

treated with long-term methylphenidate and multimodal psychosocial treatment. *J Am Acad Child Adolesc Psychiatry*. 2004;43:802–811.

Amminger GP, Schafer MR, Papageorgiou K, et al. Long-chain omega-3 fatty acids for indicated prevention of psychotic disorders: a randomized, placebo-controlled trial. *Arch Gen Psychiatry*. 2010;67(2):146–154.

Birmaher B, Brent D; Work Group on Quality Issues, Bernet W, et al. Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. *J Am Acad Child Adolesc Psychiatry*. 2007;46(11):1503–1526.

Brent D, Emslie G, Clarke G, et al. Switching to another SSRI or to Venlafaxine with or without cognitive behavioral therapy for adolescents with SSRI-resistant depression: the TORDIA randomized controlled trial. *JAMA*. 2008;299:901–913.

*Correll CU, Carlson HE. Endocrine and metabolic adverse effects of psychotropic medications in children and adolescents. *J Am Acad Child Adolesc Psychiatry*. 2006;45:771–791.

Emslie GJ, Kratochvil C, Vitiello B, et al; Columbia Suicidality Classification Group; TADS Team. Treatment for adolescents with depression study (TADS): safety results. *J Am Acad Child Adolesc Psychiatry*. 2006;45:1440–1455.

*Franklin ME, Sapyta J, Freeman JB, et al. Cognitive behavior therapy augmentation of pharmacotherapy in pediatric obsessive-compulsive disorder: the Pediatric OCD Treatment Study II (II) randomized controlled trial. *JAMA*. 2011;306(11):1224–1232.

Findling RL, Correll CU, Nyilas M, et al. Aripiprazole for the treatment of pediatric bipolar I disorder: a 30-week, randomized, placebo-controlled study. *Bipolar Disord*. 2013;15(2):138–149.

Gleason MM, Egger HL, Emslie GJ, et al. Psychopharmacological treatment for very young children: contexts and guidelines. *J Am Acad Child Adolesc Psychiatry*. 2007;46:1532–1572.

*Greenhill L, Kollins S, Abikoff H, et al. Efficacy and safety of immediate-release methylphenidate treatment for preschoolers with ADHD. *J Am Acad Child Adolesc Psychiatry*. 2006;45:1284–1293.

Kratochvil CJ, Vitiello B, Walkup J, et al. Selective serotonin reuptake inhibitors in pediatric depression: is the balance between benefits and risks favorable? *J Child Adolesc Psychopharmacol*. 2006;16:11–24.

March J, Silva S, Petrycki S, et al. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: treatment for adolescents with depression study (TADS) randomized controlled trial. *JAMA*. 2004;292:807–820.

*March JS, Silva S, Petrycki S, et al. The treatment for adolescents with depression study (TADS): long-term effectiveness and safety outcomes. *Arch Gen Psychiatry*. 2007;64:1132–1143.

McClellan J, Kowatch R, Findling RL; Work Group on Quality Issues. Practice parameter for the assessment and treatment of children and adolescents with bipolar disorder. *J Am Acad Child Adolesc Psychiatry*. 2007;46:107–125.

McCracken JT, McGough J, Shah B, et al; Research Units on Pediatric Psychopharmacology Autism Network. Risperidone in children with autism and serious behavioral problems. *N Engl J Med*. 2002;347(5):314–321.

McDougle CJ, Scahill L, Aman MG, et al. Risperidone for the core symptom domains of autism:

results from the study by the autism network of the research units on pediatric psychopharmacology. *Am J Psychiatry*. 2005;162(6):1142–1148.

Pappadopulos E, Macintyre li JC, Crismon ML, et al. Treatment recommendations for the use of antipsychotics for aggressive youth (TRAAY). Part II. *J Am Acad Child Adolesc Psychiatry*. 2003;42:145–161.

Piacentini J, Bennett S, Compton SN, et al. 24- and 36-week outcomes for the Child/Adolescent Anxiety Multimodal Study (CAMS). *J Am Acad Child Adolesc Psychiatry*. 2014;53(3):297–310.

Pliszka S; AACAP Work Group on Quality Issues. Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2007;46:894–921.

Research Units on Pediatric Psychopharmacology Autism Network. Randomized, controlled crossover trial of methylphenidate in pervasive developmental disorders with hyperactivity. *Arch Gen Psychiatry*. 2005;62:1266–1274.

Spencer TJ, Abikoff HB, Connor DF, et al. Efficacy and safety of mixed amphetamine salts extended release (Adderall XR) in the management of oppositional defiant disorder with or without comorbid attention-deficit/hyperactivity disorder in school-aged children and adolescents: a 4-week, multicenter, randomized, double-blind, parallel-group, placebo-controlled, forced-dose-escalation study. *Clin Ther*. 2006;28:402–418.

Steiner H, Remsing L; Work Group on Quality Issues. Practice parameter for the assessment and treatment of children and adolescents with oppositional defiant disorder. *J Am Acad Child Adolesc Psychiatry*. 2007;46:126–141.

Strawn JR, Welge JA, Wehry AM, Keeshin B, Rynn MA. Efficacy and tolerability of antidepressants in pediatric anxiety disorders: a systematic review and meta-analysis. *Depress Anxiety*. 2015;32:149–157.

The Pediatric OCD Treatment Study (POTS) Team. Cognitive-behavioral therapy, sertraline, and their combination for children and adolescents with obsessive-compulsive disorder: the pediatric OCD treatment study (POTS) randomized controlled trial. *JAMA*. 2004;292:1969–1976.

Vitiello B, Emslie G, Clarke G, et al. Long-term outcome of adolescent depression initially resistant to selective serotonin reuptake inhibitor treatment: a follow-up study of the TORDIA sample. *J Clin Psychiatry*. 2011;72(3):388–396.

Volkmar F, Siegel M, Woodbury-Smith M, et al. Practice parameter for the assessment and treatment of children and adolescents with autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry*. 2014;53(2):237–257.

Walkup JT, Albano AM, Piacentini J, et al. Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *N Engl J Med*. 2008;359(26):2753–2766.

Wehry AM, Beesdo-Baum K, Hennelly MM, Connolly SD, Strawn JR. Assessment and treatment of anxiety disorders in children and adolescents. *Curr Psychiatry Rep*. 2015;17(7):52.

Wilens TE, Bukstein O, Brams M, et al. A controlled trial of extended-release guanfacine and psychostimulants for attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2012;51:74–85. e2.

*A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. The MTA Cooperative Group. Multimodal Treatment Study of Children with ADHD.

Arch Gen Psychiatry. 1999;56:1073–1086.

▲ 54.7 Inpatient Psychiatric, Partial Hospital, and Residential Treatment for Children and Adolescents

DANA KOBER, M.D., AND ANDRéS MARTIN, M.D.

Psychiatric inpatient beds rose significantly after the influenza epidemic of the 1920s. This was due to the need for long-term custody of children with mental and behavioral problems post encephalitis. The number of inpatient beds reached a peak in the 1950s, prior to the deinstitutionalization movement of the 1960s. Deinstitutionalization was propagated by poor conditions of some institutions as well as changes in treatment approaches with the advent of psychopharmacological interventions.

The number of inpatient and residential pediatric beds has continued to decrease since then, particularly since the 1980s. A nationally representative study of community hospitals conducted by Case and colleagues showed that inpatient lengths of stay continued to decrease through the end of the twentieth century, going from a median of 12.2 days in 1990 to 4.5 days in 2000, a change of 63 percent. Meagher and colleagues looked at one inpatient setting in Massachusetts which also demonstrated a significant decline in length of stay, ranging from a mean of 44.5 days in 1991 and 10.7 days in 2008. Though the mean length of stay is different in these two studies, they both show the steady decline of the amount of time children spend in the hospital. The latter study also showed significantly increased number of psychotropic medications at admission and discharge, lower GAF at intake, and increased number of prior hospitalizations for patients admitted in 2008 versus 1991. There has been a concurrent increase in the number of psychiatric inpatient admissions for children and adolescents, as noted in an analysis of National Hospital Discharge Survey data between 1996 and 2007 by Blader and colleagues. One conclusion may be that the outpatient, community-based, systems-of-care approach is working for many children and adolescents; however, children and adolescents admitted currently are sicker than those admitted in the 1990s.

The effects of institutions closing their doors, a shift to managed care and other third-party funders, the privatization of Medicaid, and changing paradigms of effective psychiatric treatment options combined to result in a shift in the way in which inpatient psychiatric hospitalization is conceptualized and used, particularly in the United States: Emphasis has shifted to acute stabilization, diagnostic assessment, implementation of multidisciplinary treatment interventions, and return to the community and its available supports. The focus of inpatient psychiatric hospitalization

has shifted more toward assessment and stabilization and less toward long-term treatment, although there is recognition that for a minority of children, inpatient and highly structured care remain the preferred treatment approach.

Amanda B. Geller and Thomas M. Wickizer argued that by the end of the 20th century the reduction of inpatient services had gone too far, with lengths of stay becoming unrealistically short, leading to a rebound in readmission rates and a "revolving door" approach to inpatient and crisis care. In many states, there is a clear shortage of inpatient and residential treatment centers, resulting in youth being seen in emergency departments, waiting for appropriate treatment in juvenile detention centers, or simply floundering untreated in schools and the community at large. Emergency departments across the nation particularly have had to deal with the burden of this shifting ecology of services.

Although there are significantly fewer inpatient beds since the onset of deinstitutionalization in the 1960s, inpatient psychiatric treatment continues to be necessary and essential. Due to improvements in psychopharmacological, behavioral, and community and in-home treatment options, theoretically fewer hospitalizations would be necessary. However, rates of admission and total inpatient days increased significantly between 1996 and 2007. Nonetheless, a national shortage of inpatient beds and outpatient treatment options remains. As a temporary measure, some hospitals have adapted to "board" children on general pediatric wards, or in the emergency department itself, while they await a psychiatric bed or until they are stabilized on that unit. Wharff and colleagues studied a large urban pediatric emergency department in which one-third of patients needing inpatient psychiatric admission were boarded in the emergency department or pediatric floor. Odds of boarding were increased for patients with diagnoses of autism, intellectual delay, and/or developmental delay, as well as children with severe suicidal ideation. There is a very recent trend to provide beds for children with autism and developmental delays in a separate but attached unit, as the needs of these children are different than the typical population in inpatient units. Children who do not meet strict inpatient admission criteria may be referred to crisis stabilization programs that provide short-term treatment while longer-term services are arranged.

Table 54.7–1.
Differences among Levels of Care

Level of Treatment	Purpose	Length of Stay	Benefits	Example of Reason for Admission
Short-term inpatient	Acute stabilization	Days to weeks	Improvement in symptoms and functioning	Acute risk of harming self or others

Intermediate-term inpatient	Arrangement for appropriate placement	Weeks to months	Safe environment; ongoing therapy	Child in need of longer stay; may be awaiting residential placement
Partial hospital program (PHP) or intensive outpatient program (IOP)	Transition from, or prevention of, inpatient hospitalization	Weeks to months	Maintenance at home with structured therapeutic and school day programs; PHPs typically include a school component, whereas IOPs provide after-school care	Difficulty attending school; step down from inpatient stay
Residential	Long-term treatment	6 months to several years	Long-term, consistent, safe therapeutic environment	Unable to maintain safe behavior at home, often after extensive community-based services have been tried and exhausted

LEVELS OF CARE

Stabilization of the child's acute symptoms, coordination of family and relevant provider meetings, and rapid behavioral and pharmacological intervention occur during inpatient hospitalization to return the child to a safe environment as soon as realistically feasible. Partial hospitalization programs (PHPs) are used when the child is safe enough to stay at home yet needs structured therapeutic day supports, either as a temporary treatment or as a transition into a therapeutic school or residential program. Residential programs care for the child around the clock over a longer period of time than an inpatient hospitalization and have a different staffing approach. Table 54.7–1 provides comparative highlights of the key features of inpatient, partial hospital, and residential programs.

INPATIENT TREATMENT

Physical Structure, Staffing Organization, Daily Routine

Inpatient hospitalization is used for stabilization and return to the community with attention to setting up support structures prior to discharge. The physical structure and daily schedule of inpatient units share some characteristics with the outside environment of the child while also having unique characteristics relevant to the goals of hospitalization and specific aspects of safety. Inpatient units can be part of larger general or children's hospitals, in which case they may be better suited to address the needs of medically challenging or complicated children (e.g., the child with brittle diabetes or uncontrollable seizures, or the adolescent who has taken a serious overdose) or to pursue complex medical workups (e.g., children requiring lumbar taps or 24-hour video encephalography). By contrast, "free-standing" inpatient units in more naturalistic or child-friendly environments may be better suited for the needs of children

requiring longer-term hospitalization. Regardless of their possible affiliation with an overarching hospital structure, the principles of inpatient and milieu-based care are similar across settings.

Child and adolescent inpatient units can vary, depending on the conditions of admission and the purpose of the unit. Most units tend to be locked for safety purposes, given that one of the more common reasons for admission is to manage aggression or acute risk of harm to self or others and given the unique developmental needs of children. Open or unlocked units typically exclude patients who are at acute risk to harm themselves or others and are usually geared for older adolescents or adults. Such units tend to provide treatment over a longer stay following acute hospitalization or in cases in which acute stabilization is not necessary. For example, an adolescent struggling with long-term depression who is not suicidal but who is not improving with aggressive outpatient treatment may benefit from the comprehensive assessment available in an unlocked unit.

The schedule on an inpatient unit is often similar to a typical school day schedule. As depicted in Table 54.7–2, the therapeutic day includes schoolwork, recreation groups, group therapy, supervised free time, and time for individual and family meetings. Among the unique features of inpatient, partial hospital, and residential care is the role of milieu-based, group-wide activities such as those highlighted in the table. As part of these activities, children benefit and learn from (and are routinely challenged by) the presence and demands of their peers and staff members. One goal of treatment is for the child to successfully navigate daily routine tasks that imitate their routine outside of the hospital.

Staff hierarchy structures can vary from unit to unit. Some follow a traditionally vertical, medical hierarchical organization, placing physicians at the top. However, child units tend to be unique in having a more horizontal working hierarchy, with each staff person's assessment observations of the child being important and valued, seeking to maximize collaboration around the child's needs. Inpatient mental health services for children require especially high staff-to-patient ratios, which poses financial and organizational challenges. Thus, even as horizontal hierarchies are preferred, clear lines of accountability and reporting are still necessary for the smooth functioning of the unit. During an inpatient hospitalization, 24-hour/day observation demonstrates the child's behaviors and interactions. Nurses and milieu counselors are especially positioned to observe parent–child and child–peer interactions. The observation of challenging behaviors may take some time, given that it is not uncommon for children to have a quiescent ("honeymoon") period early in their hospitalization.

Table 54.7–2.

Typical Inpatient Day Schedule

7:30–8:00 AM	Wake up, morning activities of daily living (ADLs), such as washing face and brushing teeth Assistance with ADLs may be provided by staff.
8:00–8:45 AM	Breakfast and morning medications
8:45–9:00 AM	Community morning meeting
9:00–11:00 AM	School
11:00 AM–12:00 PM	School or recreational therapy
12:00–12:30 PM	Lunch
12:30–1:00 PM	Free time
1:00–2:00 PM	Individual therapy
2:00–3:00 PM	Outside play
3:00–4:00 PM	Group therapy or coping skills group
4:00–5:00 PM	Family visiting time. Rotating activity, e.g., music therapy, yoga, pet therapy
5:00–6:00 PM	Family time and free time: board games, reading, video games
6:00–6:30 PM	Dinner and evening medications
6:30–7:30 PM	Showers and ADLs
7:30–8:30 PM	Movie time for kids; parent education group for families
8:30–9:00 PM	Bedtime routines

Multidisciplinary team rounds typically occur daily, in which representatives from each discipline come together to discuss each patient's problems and treatment planning. In such meetings, staff can also work out the tensions of the unit hierarchy or examine how the dynamics of the children and their families may reciprocally affect that of the staff. Handover of information between shifts provides staff with a framework that helps alleviate frustrations, improves communication across different staff members and disciplines, and is critical to the smooth functioning of an inpatient unit. Daily multidisciplinary team rounds are among the best places to anchor such efforts and help to solidify the team. Coordination is further enhanced through efficient written communications (increasingly through an electronic medical record), through shift-to-shift and cross-discipline in-person sign-outs, and through the availability of ad hoc "mini-team" meetings in which to discuss cases, to problem solve, and to deal with the inevitable differences of opinion inherent in treating seriously ill children.

The physical structure of a psychiatric unit is generally different from that of medical inpatient units. There often is a multipurpose common room, similar to a living room, used for group discussions, watching TV, playing video games, or engaging in arts and crafts. This room may also be transformed into the dining area, for which the tables double as work areas during recreational therapy time. Couches, chairs, and tables can be moved to the perimeter of the room to make floor space for yoga or other activities. Some units have a secure outside playground, providing another

important outlet for physical expression and exercise. Bedrooms typically have beds for one to three children; time spent in the community room is encouraged. Inpatient units are designed with attention to safety considerations: Beds and furniture are secured to the ground; curtains are attached with the use of Velcro; mirrors and windows are plasticized and not breakable into shards; if there is a seclusion room, its walls and doors are soft or padded.

Management of Aggression, and the Use of Restraint and Seclusion Revisited

Aggression is perhaps the most common reason for inpatient hospitalization in children. As such, inpatient units are charged with dealing with safe ways of de-escalation and of stabilization of aggression and violence. Seclusion into contained spaces (such as a "quiet room") and mechanical restraint through the tying of limbs to bed railings with Velcro or leather straps have been two historical means of dealing with acute violence and the imminent risk of injury directed at self or others. The use of restraint or seclusion had been considered a therapeutic intervention and conceptualized as an external means to help children regain their internal locus of control. However, following reports of deaths due to asphyxiation during restraints and physical holds in the early 1990s, there have been concerted and federally supported efforts to reduce the use of such restrictive interventions. Indeed, restraint and seclusion are better conceptualized as treatment failures, reflecting more about the failed efforts leading to them than about the patients whose aggression they seek to contain.

Abolishing the use of restraint and seclusion cannot be accomplished unilaterally or through a top-down directive. Units in which staff members are not provided with alternative therapeutic means to prevent, approach, and contain aggression and violence are unlikely to eliminate restraint or seclusion use or to do so in a manner that maintains patient and staff safety. A number of successful approaches, such as Collaborative Problem Solving, have been introduced to help safely to reduce or eliminate the use of such practices. Even as approaches can vary, they all have certain core features, including six core elements that are deemed essential for the systematic reduction of restrictive interventions in inpatient settings: leadership support of organizational change, use of data to inform practice, workforce development, use of seclusion and restraint prevention tools, inclusion of patients and families, and rigorous debriefing following restrictive events.

Use of antipsychotics in response to aggression in inpatient units, as well as the outpatient setting, has increased. Additionally, the more severe the behavior or aggression, the more likely medication is changed by the outpatient provider after discharge, most often being the start of a new medication.

Admission and Discharge Criteria

Treating the child in the least restrictive environment should be a guiding principle of any level of care. Short inpatient stays provide a safe, containing environment for a child who is acutely at risk of harm to self or others. At the same time, as early as the process of assessment and stabilization begins, support services necessary for discharge should be explored. Stated alternatively: inpatient care should not be viewed as an endpoint, but rather as a site for stabilization and a likely way station toward longer-term treatment options, be they outpatient and home-based approaches or restrictive settings such as subacute or residential placements.

Typical admission criteria for inpatient hospitalization include the following:

1. A psychiatric condition that places the child or adolescent at imminent risk of harm to self or others
2. Psychiatric illness that poses medical complications
3. Loss of judgment or impulse control that results in dangerous, threatening, or grossly disorganized behavior toward self or others
4. Failure of less restrictive treatment options

Because one of the roles of inpatient hospitalization is to protect the child from harming self or others due to dangerousness or aggression, it follows that a discharge criterion is that the child no longer poses such a risk. In addition, the home environment or contributing factors to the child's admission must be molded enough to prevent immediate readmission. Limitations to prompt discharge from the hospital may include inadequate level of supervision and care at home or wait lists for outpatient, community-based, or subacute or longer-term placements. At other times, it is the family work or psychosocial rather than narrowly defined psychiatric reasons in the identified patient that limit discharge feasibility from inpatient or residential settings. Indeed, it is common for inpatient settings to have a bimodal length of stay, with children admitted for psychiatric stabilization and shorter lengths of stay sharing a milieu with a minority of children whose adverse psychosocial circumstances and lack of adequate discharge options translate into lengthier hospitalizations (e.g., over 3 weeks). In these circumstances, it is not uncommon for inpatient settings to serve the dual role of an acute stabilization unit and de facto residential placement, with the ensuing logistical, clinical, and administrative challenges.

Assessment

During an inpatient admission, several assessments occur. Assessment is not only individualized to the child, but it also encompasses the input of

multiple disciplines. This is not unlike the parable of the five blind men defining an elephant based on the part with which they are in contact. In assessing a child through such a multidisciplinary approach, there is an inherent risk of fragmentation of efforts that needs to be carefully avoided. At the same time, it is helpful to delegate, in a coordinated manner, parts of the assessment to those team members best suited for the task.

Assessment includes the current presentation and reason for admission, as well as a review of past treatments, including treatment failures and successes. Information regarding the child's past treatments may be evaluated by obtaining prior records and speaking with the child's outpatient providers. Observation of the child in groups and interacting with peers is a particularly helpful aspect of inpatient hospitalization. Provision of inpatient care also allows clinicians and staff to observe the external and internal deficits that may contribute to the child's presentation. There is an initial expedited assessment of the patient's strengths, weaknesses, needs, and precipitants for decompensation or aggression. The child's family dynamics are assessed in family meetings that may occur with a social worker or the primary clinician. After the initial assessment is complete, all team members come together, typically during team rounds, to discuss each part of the assessment. This process is refined iteratively throughout the length of the hospital stay.

Table 54.7–3 outlines the roles of various inpatient staff members in assessing the child within the contexts of the family, school, and community. For didactic purposes, the roles are outlined in a somewhat artificial delineation of labor. However, it is important to prevent the fragmentation of care that can take place in a staff-rich inpatient setting. The major strength of an inpatient's number of participants can devolve into its greatest weakness when communication falters or when faulty assumptions are made about the responsibilities of others.

Treatment Planning

Inpatient treatment interventions are individualized to the child's strengths and weaknesses and the family's capabilities and identified problems. It is helpful to co-construct the treatment goals with the family because they will be most involved in implementing the treatment once the child is discharged with appropriate support services in place for assistance. Several different modalities of treatment are examined and possibly implemented during the inpatient stay, including psychoeducation, individual therapy, psychopharmacotherapy, and family work. During an inpatient admission, the child's outpatient or community treaters are contacted. This serves several purposes. First, collateral information aids in understanding the longitudinal progression of the child's deterioration and resultant admission. Second, preventing replication of previously failed treatment efforts saves time and frustration. Details of the dosing of

medication, duration of treatment, likelihood of compliance, and reason for discontinuation are all to be obtained from the provider or from past records. Finally, outpatient clinicians are able to provide information about diagnostic conundrums and family dynamics. Typically the child will return to the same provider(s); therefore, collaboration about treatment is essential for continuity after discharge.

Table 54.7–3.
Role and Responsibility Highlights for Multidisciplinary Team Members

Team Member	Role/Responsibility	Example
Patient	Participate in therapeutic programming. Ask for help.	Go to groups, individual therapy, recreational therapy. Be able to tell staff when not feeling safe.
Family	Provide information about problems warranting treatment. Participate in treatment planning and behavioral strategies.	Engage in family meetings and parent education groups. Able to visit child during hospitalization.
Child Protective Services	May be legal guardian, temporary custodian, additional service provider, or involved in investigation. When serving in loco parentis, may provide background information and be responsible for authorizing necessary consent.	Attend family meetings. Provide medication history and consent. Arrange appropriate placement postdischarge. Fund and help to identify necessary services.
Home school	Can provide information about child's daily school functioning.	Have phone contact with unit school staff. Provide Individual Education Plan and past testing.
Psychiatrist	Provide diagnostic assessment. Treat child with therapy and medication and overall coordination of care. Engage family.	Admit patient. Perform physical examination. Order appropriate testing. Engage child in therapy. Assess and manage medications.
Social worker	Engage family and arrange follow-up services.	Conduct family meetings. Obtain genogram. Apply for community supports.
Psychologist	Provide diagnostic assessment and therapy.	Provide individual therapy with child. Conduct psychological testing.
Nurse	Provide assessment throughout the day of the child's mental status and attention to daily needs.	Administer medication, including emergency medication when necessary.
Milieu counselor or psychiatric aide	Manage aggression by providing external structure and daily demonstration of coping skills.	De-escalate tense situations. Provide constant assessment of child's current condition, including mental status and current needs. Help child to learn coping skills and how to solve problems between peers.
Unit teachers	Assess each child and teach to his or her level.	Provide daily teaching to children of

	May perform educational testing or screeners.	different levels. Work with individual children or in small groups. May provide recommendations to home school.
Therapeutic recreation therapists	Provide daily activities that build sense of accomplishment and help in understanding feelings and problem solving.	Provide specific games, crafts, or groups tailored toward a given lesson or goal.
Occupational therapists	Assess child's need for assistive devices or alternative means of sensory stimulation or reduction.	Provide alternative means of alleviating anxiety in a socially appropriate manner.
Consulting pediatrician	Provide medical needs not within expertise of psychiatrist.	Consult for treatment of moderate to severe medical condition or of complex subspecialized care.

An individualized treatment plan matches the child's and family's needs. For instance, a complex treatment that cannot be sustained outside of the hospital would set up the family and child to fail. Family meetings aide in facilitation of communication between child and family members, as well as between the family and the treatment team. Some treatment approaches do not translate across environments and practitioners.

In obtaining consent, it is important to ascertain the patient's and family's expectations of the particular treatment modality. Institutional transference or negative thoughts about the suggested treatment setting may or may not be easily corrected. Unrealistic hopes may result in an initially strong response, later leading to apparent failure and disappointment when such expectations are not met. Thus, equally important is communication of facts about the treatment, including adverse reactions, side effects, and time to effect. For example, a side effect of a behavioral treatment might include an extinction burst, in which behavior worsens before improvement. Common side effects of many medications are sedation and weight gain. Appropriate laboratory tests should be obtained prior to starting medications. Risks and benefits must be weighed for each treatment, and the reality of compliance with the treatment must be considered. For example, if the family does not have the appropriate skills to comply with a certain behavioral treatment, they are unlikely to be successful in supporting the child in the treatment. This can contribute to further feelings of frustration and helplessness, as well as to possible negative feelings toward psychiatric treatment more generally.

Support services, such as home-based therapeutic interventions, may be part of the discharge plan if available and appropriate. Treatment recommendations may also include recommendations for further testing or educational recommendations to the child's school.

PARTIAL HOSPITAL

The structure of a partial hospital program (PHP) or intensive outpatient program (IOP) is loosely based on inpatient programs. With current behavioral and individual assessments and interventions, families are included as part of the assessment and locus of treatment. Programs can include a school component (typically called a PHP) or may be an after-school program (often referred to as an IOP or an extended day program). Most programs have a staff structure similar to an inpatient unit, in that many PHPs are an extension of an inpatient program, and many of the children "step down" into this less restrictive setting of care after their initial stabilization. Part of the assessment occurs prior to the PHP admission to determine whether the PHP is an appropriate treatment option. If school is part of the PHP, the patient's school district is typically involved in authorizing attendance to the PHP as counting toward temporary school placement. The daily routine of a PHP differs from an inpatient unit schedule in that the child wakes up and gets ready at home and then is transported to the facility. At that point the therapeutic school day begins, and group and individual therapy also take place. The child then goes home in the late afternoon, to return again the following day. If the child continues to attend a regular school, therapeutic programming occurs after school in the afternoon.

Children are typically admitted to a PHP or IOP program because of a behaviorally related lack of school successes or consistent problems at home not responsive to outpatient treatment efforts. These children are not necessarily at acute risk of harming themselves or others. An example might be a 13-year-old girl who is depressed and school avoidant but is not suicidal or engaging in self-injurious behaviors. For such patients, once-weekly visits to their outpatient clinicians may not suffice, given the severity of their problems, but inpatient admission may not be warranted either. PHP can also be a useful step down from inpatient hospitalization.

A partial program is also useful for observing the child and for making educational assessments and recommendations. Psychological testing and educational testing can be performed, and recommendations may be given to the regular school. Partial hospitalization programs may be a transitional placement while the child awaits a more permanent therapeutic school placement. Assessment and treatment are similar to those in inpatient programs, including contact with outside providers and school personnel. There is also emphasis on making accommodations so that the child can succeed at school. Such accommodations may include a different school placement or changes within the child's current school. Discharge from a PHP or IOP program depend on the preparation of the child and family to be able to continue care at home with adequate outpatient services.

RESIDENTIAL TREATMENT CENTERS

Much consideration occurs prior to the decision to place a child in a

residential treatment center (RTC). “Residential” can have a variety of meanings, from residential placement of a child with no other home option, to a voluntary placement sought by families to provide 24 hour-care, to housing youth in the juvenile justice system, or providing therapeutic treatment with the goal of return to the community. Additionally, there are group homes, which are not the subject of this chapter. Therapeutic boarding schools have also been used as synonyms. The history of identifying residential stays as treatment began with the establishment of the American Association of Children’s Residential Centers in 1954. Even at that time, there was variation in the definition of residential treatment, though most centered around a psychodynamic orientation, requiring isolation of the child from the environment as a necessity for treatment. This trend continued through the 1960s and 1970s until the beginning of deinstitutionalization. In the 1980s, with the advent of managed care, intensive outpatient treatment became more economically advantageous as compared to residential treatment. Residential treatment also declined due to improvements in psychopharmacology and advocacy for community-based, wrap-around, family-centered outpatient care, partly due to the more current philosophy that children are better assessed in the context of the environment in which maladaptive behaviors occur. Despite the best efforts, some children will still require residential placement, whether temporary or longer term.

Children in need of residential treatment tend to have either severe psychiatric and/or developmental disorders, preventing them from being managed through community and family supports alone, or have such circumstances that prevent them from living either at home or in therapeutic foster care. The parent–child relationship may be damaged to such an extent that the child’s behavior becomes consistently unmanageable when at home. The parent(s) may not have the capacity due to severe mental illness or substance abuse, and there may be no suitable, less restrictive alternative.

RTC staffing differs from that in inpatient and PHP settings. For example, the psychiatrist on staff often serves in a consultant role, thereby lessening psychiatrist involvement in treatment. There are also fewer nurses per child, but the number of day-to-day residential staff is increased. Rather than a vertical hierarchy, the RTC relies on a horizontal stratification by nature of the interaction of the child with the childcare staff. A popular RTC model divides up children into smaller “cottages,” with a master of the cottage or “houseparent” residing with them. Attempts are made at RTCs to make the environment as natural as possible, which helps in the comfort of the child, as well as in creating the best background for generalizing skills to be used after discharge. Lengths of stay have decreased significantly since the 1950s, with further evidence since the early 2000s of 3- to 6-month lengths of stay having efficacy. The use of

cognitive-behavioral and dialectical-behavioral approaches has replaced psychodynamic orientation. Reduction in seclusion and restraint has also occurred as in inpatient hospitals.

Typically, children are stable enough to not need an inpatient hospitalization; therefore, residential programming focuses on individual coping skills, relationship building, and basic living skills, including self-care. There is a routine schedule as in inpatient units, but with more individualized free time. Focus is placed on deficient adaptive and social skills, as well as on improving impulsivity by having the children learn to better tolerate frustration. A school program may be on the RTC grounds, or the child may be transported out to a separate therapeutic school or prevocational program.

There has been more push in the past decade for integrating the family into treatment, as the ultimate goal for many children is eventual return to their homes of origin. Continued family contact preserves bonds necessary for success after discharge. Parent training and family therapy are an integral aspect of residential treatment and prepare the family for the child's return home. Rather than designating family visits as a reward for desired behavior, it is considered necessary to have ongoing family involvement to achieve the end goal of successful reunification of the family and success in outpatient treatment. For children whose families are unable to care for them due to severe dysfunction, preparations are made for discharge to a group home or to a foster placement when appropriate.

COMMUNITY SUPPORTS AND IN-HOME SERVICES

A change in admission philosophy and inpatient treatment goals has caused mental health care providers and administrators to think differently about systems of treatment. Since the turn of this century, there have been improvements in treatment system options. Preadmission function of the family and the child and different systems of family, school, and community are crucial to consider in both assessment and treatment considerations for postdischarge recommendations. With shorter hospital lengths of stay and scarce inpatient resources and their associated high costs, it is vital to establish the tools and techniques that families will need to maintain the child's behavior and safety after discharge because frequent admissions may repeat the trauma of separation. All efforts are made to have community supports in place at time of discharge. These can include in-home services, extended day services, PHPs or intensive IOPs, and mentors or case managers. Communication with outpatient providers is key to continued stabilization postdischarge.

Intensive in-home treatment is a relatively new modality of administering child psychiatric assessment and treatments. Earlier conceptualizations of in-home services had been government-funded and followed a visiting nurse model, excluding other mental health specialties

such as psychology or psychiatry. However, in the last two decades, shifts have occurred to include more mental health care services that can be paid for by insurance, whether private or state, and the creation of mental health teams to include social workers, psychologists, and psychiatrists.

The treatment team typically delivers environment-relevant assessment and recommendations several times a week in the family's home. Problem solving can occur in vivo and in an ecologically naturalistic setting. This is a significant improvement consistent with the philosophy of "meeting families where they are," both in literal and figurative terms. These types of services play a large role in the prevention of acute destabilizations that would cause inpatient hospitalizations. One of the first home-based services described and studied was multisystemic therapy (MST). Originally developed for youth with juvenile justice involvement or substance use disorders, the model has been adapted to the specific needs of children with psychiatric illness and complex family and psychosocial situations.

FUTURE DIRECTIONS

Advances in psychiatric treatment options, the availability of a range of community services, and a general shift in philosophy have gradually changed the conceptualization of psychiatric inpatient hospitalizations for children and adolescents from one of longer-term treatment to one of multidisciplinary diagnostic and needs assessment and acute stabilization. Although lengths of stay and philosophies about seclusion and restraint have changed over the past decade, basic dynamics of staff, patients, and families remain. Partial hospital programs and in-home services contribute to the capability for children to have shorter inpatient stays and possible reunification after inpatient or residential stays.

REFERENCES

*Blader JC. Acute inpatient care for psychiatric disorders in the United States, 1996 through 2007. *Arch Gen Psychaitry*. 2011;68:1276–1283.

Blader JC. Pharmacotherapy and postdischarge outcomes of child inpatients admitted for aggressive behavior. *J Clin Psychopharmacol*. 2006;26:419–425.

Case BG, Olfson M, Marcus SC, Siegel C. Trends in the inpatient mental health treatment of children and adolescents in US community hospitals between 1990 and 2000. *Arch Gen Psychiatry*. 2007;64:89–96.

Costello EJ, Copeland W, Cowell A, Keeler G. Service costs of caring for adolescents with mental illness in a rural community, 1993–2000. *Am J Psychiatry*. 2007;164:36–42.

Cotton NS. *Lessons from the Lion's Den: Therapeutic Management of Children in Psychiatric Hospitals and Treatment Centers*. San Francisco: Jossey-Bass; 1993.

Curie C. SAMHSA's commitment to eliminating the use of seclusion and restraint. *Psychiatr Serv*. 2005;56:1139–1140.

Darwish A, Salmon G, Ahuja A, Steed L. The community intensive therapy team: development and philosophy of a new service. *Clin Child Psychol Psychiatry*. 2006;11:591–605.

Donovan A, Siegel L, Zera G, Plant R, Martin A. Child & adolescent psychiatry: seclusion and restraint reform: an initiative by a child and adolescent psychiatric hospital. *Psychiatric Serv*. 2003;54:958–959.

Drell M. Innovations: child & adolescent psychiatry: "sweet are the uses of adversity": a transition program for children discharged from an inpatient unit. *Psychiatr Serv*. 2006;57:31–33.

Enns MW, Cox BJ, Inayatulla M. Personality predictors of outcome for adolescents hospitalized for suicidal ideation. *J Am Acad Child Adolesc Psychiatry*. 2003;42:720–727.

*Geller JL, Biebel K. The premature demise of public child and adolescent inpatient psychiatric beds—Part I: overview and current conditions. *Psychiatr Q*. 2006;77:251–271.

*Geller JL, Biebel K. The premature demise of public child and adolescent inpatient psychiatric beds—Part II: challenges and implications. *Psychiatr Q*. 2006 Winter;77:273–291.

Greene RW, Ablon JS, Hassuk B, Regan KM, Martin A. Innovations: child and adolescent psychiatry: use of collaborative problem solving to reduce seclusion and restraint in child and adolescent inpatient units. *Psychiatr Serv*. 2006;57:610–612.

Henggeler SW, Schoewald SK, Rowland MD, Cunningham PB. *Serious Emotional Disturbance in Children and Adolescents: Multisystemic Therapy*. New York: Guilford; 2002.

LeBel J, Goldstein R. The economic cost of using restraint and the value added by restraint reduction or elimination. *Psychiatr Serv*. 2005;56:1109–1114.

LeBel J, Stromberg N, Duckworth K, et al. Child and adolescent inpatient restraint reduction: a state initiative to promote strength-based care. *J Am Acad Child Adolesc Psychiatry*. 2004;43:37–45.

*Leichtman M. Residential treatment of children and adolescents: past, present, and future. *Am J Orthopsychiatry*. 2006;76:285–294.

Libby AM, Cuellar A, Snowden LR, Orton HD. Substitution in a Medicaid mental health carve-out: services and costs. *J Health Care Finance*. 2002 Summer;28:11–23.

Meagher SM, Rajan A, Wyshak G, Goldstein J. Changing trends in inpatient care for psychiatrically hospitalized youth: 1991–2008. *Psychiatr Q*. 2013;84:159–168.

Miller L, Riddle MA, Pruitt D, Zachik A, dosReis S. Antipsychotic treatment patterns and aggressive behavior among adolescents in residential facilities. *J Behav Health Serv Res*. 2013;40:97–110.

National Association of State Mental Health Program Directors Technical Assistance Center. Six Core Strategies to Reduce Seclusion and Restraint Planning Tool. 2005. Available at http://www.nasmhpd.org/ntac.cfm.

Pottick KJ, Bilder S, Vander Stoep A, Warner LA, Alvarez MF. US patterns of mental health service utilization for transition-age youth and young adults. *J Behav Health Serv Res*. 2008;35:373–389.

Stanton A, Schwartz M. *The Mental Hospital: A Study of Institutional Participation in Psychiatric Illness and Treatment*. New York: Basic Books; 1954.

Stefan S. *Successful Seclusion and Restraint Reduction Programs as Quality Indicators for Psychiatric Services*. 2006. Available at http://www.medscapecom/viewarticle/528949_1.

Wharff EA, Ginnis KB, Ross AM, Blood EA. Predictors of psychiatric boarding in the pediatric emergency department: implications for emergency care. *Pediatr Emerg Care*. 2011;27:483–489.

Wickizer TM, Lessler D, Boyd-Wickizer J. Effects of health care cost-containment programs on patterns of care and readmissions among children and adolescents. *Am J Public Health*. 1999;89(9):1353–1358.

*Woolston JL. The administration of hospital-based services. *Child Adolesc Psychiatr Clin N Am*. 2002;11:43–65, vi.

Woolston JL, Adnopoz JA, Berkowitz SJ. *IICAPS: A Home-Based Psychiatric Treatment for Children and Adolescents*. New Haven, CT: Yale University Press; 2007.

▲ 54.8 Community-Based Treatment

ANDRéS J. PUMARIEGA, M.D.

INTRODUCTION AND DEFINITION

Community-based treatment for children and adolescents has undergone significant transformation over the past 30 years. The traditional community mental health model has largely yielded to the community-systems-of-care philosophy and model. This model is based on an individualized, intensive, and flexible approach to service delivery for children and families within the context of their homes and communities as an alternative to treatment in out-of-home settings. This model also focuses on systems issues that affect care, including interagency coordination, service access and utilization, financing, and clinical and cost-effectiveness of mental health and support services provided to children and their families. Psychiatry, which was central in the traditional model, has only recently re-engaged itself as a discipline in this new model. At the same time, as the model has been scaled up to encompass regional and statewide implementation, inadequate fidelity and resource allocation continues to contribute to lack of access to community-based services and the excessive use of residential services, many of lower quality and therapeutic value. More recently, in response to the cry by pediatric primary care physicians for greater access to child mental health services, the collaborative and integrated behavioral health models have begun to emerge to form more effective alliances with this important sector of care. Psychiatrists are meeting the challenge in reaffirming their unique value in these models and in integrating their developing clinical and scientific knowledge and skills base.

The increasing need for and utilization of child mental health services in the United States has spurred their rapid growth. Recent studies suggest that approximately 20 percent of children and adolescents have a diagnosable mental, emotional, or behavioral disorder, with 5 percent for

serious mental illness and emotional disturbance. A number of morbidities have been associated with childhood emotional disturbance and mental illness, including suicide, homicide, substance abuse, child abuse, teenage pregnancy, school dropout, youth crime, and associated institutionalization and incarceration. However, less than 1 percent of children in the United States receive mental health treatment in hospital or residential settings, with another 5 percent receiving treatment in outpatient or community-based settings; greater than two-thirds of children in need still receive insufficient or no mental health services whatsoever.

In contrast, funds dedicated to the financing of public mental health services for children have grown significantly since the 1970s. Medicaid, the public insurance program for the poor and disabled, funds a significant proportion of child mental health services in the United States. In June 2013, over 28 million children were enrolled in Medicaid and another 5.7 million were enrolled in CHIP. Mental health and substance abuse services is 11.5 percent of the over 600 billion dollars of overall Medicaid costs. The federal government contributes from 50 to 75 percent of costs, depending upon the state. Additionally, 53 percent of children receiving Supplemental Security Income (SSI) disability benefits have psychiatric disabilities (a total of 640,000). The total public costs of child behavioral health care come to $247 billion annually.

The main cost containment approach to reduce these expenditures has been to control inpatient expenditures, which has more recently led to a major shortage of child and adolescent acute inpatient beds in the US. A disproportionate level of these expenditures is still dedicated to residential treatment, which is viewed with increasing skepticism, given their relatively high cost, as well as limited documentation of outcome. In addition, there has been increasing documentation of poor practices and abuses in residential treatment as a result of underfunding and decreasing standards and quality of care. As a result of the growth in expenditures, the resources available to fund community-based child mental health and human services are increasingly limited. However, the great majority of these children and youth are poor, underserved children of ethnic minority backgrounds. Children from these populations experience higher levels of stressors, such as poverty, discrimination, immigration, acculturation stress, and exposure to violence and trauma, and they are likely to have higher levels of need for services.

Minority populations have also been traditionally underserved by mental health, health, and human services, in terms of both overall access to services and the cultural competence of services. The cost of serving these populations of children and adolescents is in contrast to the high cost of the psychosocial morbidity with which they contend, including lost productivity and the costs of welfare dependency and institutionalization. The United States continues to have the highest youth incarceration rates

in the world, and the highest levels of out-of-home or community residential placement for youth. A number of health and human service agencies (schools, social welfare agencies, child protective agencies, juvenile justice, and public health) have shouldered the increasing burden of psychosocial morbidity experienced by these children and youth. These agencies typically address pieces of the service system puzzle, with little to no coordination with other agencies serving the same youth resulting in increasing care fragmentation.

These trends have increased pressures on public child mental health and social service agencies to demonstrate improved clinical and cost-effectiveness, increasingly turning to managed care approaches to finance and organize mental health and social services. The majority of children covered by Medicaid are under managed care plans, including for behavioral health services. However, most managed care approaches were developed with private sector populations in mind, relying on a priori benefit restrictions based on actuarial data on middle-class populations with adequate social supports, who are traditionally lower users of services. When applied to public child mental health services, these approaches deprive children at high risk of effective intervention and preventive services. They also contribute to fragmentation of care and burden shifting to other child-serving agencies and systems, significantly increasing morbidity and acute need. Some Medicaid entities have successfully implemented community-systems-of-care approaches and integrated them with managed care methods, but many have had difficulty integrating them. Over the last 20 years, there has been an exponential growth of class action lawsuits against state Medicaid, educational, child welfare, and juvenile justice agencies failing to meet federally required service mandates, particularly around their high representation of children with significant mental health needs. The most egregious of such lawsuits involve the practice of parents having to give up custody of their children to obtain necessary mental health services.

HISTORY

From its earliest origins, children's mental health services have emphasized a community orientation. In the United States, these services began in the 1890s in response to the perceived need for counseling juvenile offenders rather than incarcerating them with adults. At that time, much like today, the United States was undergoing rapid sociocultural changes due to immigration, industrialization, and urbanization, and these social strains contributed to marked increases in juvenile crime and status offenses. The earliest juvenile courts in Chicago and Boston established clinics that comprised the first child mental health services in the nation. Their success led the Commonwealth Foundation to commission a study in the 1920s that recommended the development of child guidance clinics throughout

the United States staffed with interdisciplinary teams of professionals that could serve children and their families. They provided low-cost services oriented to the needs of the child and the family, with treatment modalities evolving to include individual psychodynamic psychotherapy, family therapy, crisis intervention, and even day treatment programs. They served as models for the community mental health centers that, with some alterations, emerged in the 1960s, and many have survived to this day.

The "medicalization" of psychiatry, starting in the late 1970s, served to move child and adolescent psychiatry toward the more hospital-based, tertiary care model. This left the child guidance clinics, and the community mental health centers that followed them, without significant psychiatric input. Many of the children previously receiving psychiatric services in these clinics were then served in free-standing psychiatric hospitals, residential facilities, or detention facilities, often in the custody of child welfare or juvenile justice agencies.

The modern era of community-based systems of care for children was ushered in by the publication of Jane Knitzer's groundbreaking book *Unclaimed Children,* which exposed the aforementioned consequences of neglecting the provision of community-based mental health services for children and their families. Her advocacy and that of others led to the development of the Child and Adolescent Service System Program (CASSP), which assisted all 50 states in the development of an infrastructure for publicly funded, community-based services. The CASSP initiative was supported by the conceptual work of Stroul and Friedman, who coined the term "community-based system of care for seriously emotionally disturbed children" and enunciated the principles behind such systems of care. Stroul and Friedman's work spurred the development of many innovative community-based treatment modalities, as well as a number of model demonstration programs in different parts of the United States that embody the use of these modalities within the context of an organized interagency system of care. In the early 1990s, the Robert Woods Johnson Foundation established eight pilot demonstration community-systems-of-care programs in different parts of the country, which demonstrated the viability of the systems-of-care approach and demonstrated cost savings, as well as less restrictive levels of care. Starting in 1994, the Center for Mental Health Services (CMHS) of the Substance Abuse and Mental Health Services Administration established the Comprehensive Community Mental Health Services for Children and their Families Program. This program has funded 173 community-systems-of-care sites throughout the nation in widely diverse communities. The *Surgeon General's Report on Mental Health,* published in 1999, focused its recommendations for children's mental health around the systems-of-care model and its benefits for service system reform and individualized care. More recently, given the focus on school violence and the threat of

terrorism in the United States, there has been additional focus by the federal government on these two aspects of community children's mental health. The Safe Schools Health Students program has funded 240 schools and communities to develop school-based preventive and clinical programs oriented at decreasing school violence and substance abuse, and the National Child Traumatic Stress Initiative funded many sites to adapt, implement, and evaluate promising intervention approaches addressing child traumatic stress in communities. More recently, as a result of the continued school shootings in the United States, most exemplified by the Sandy Hook Elementary shooting in Connecticut, the Federal government prioritized an additional $235 million investment in programs to help identify mental health concerns early, improve access to mental health services, and support safer school environments, including $160 million for the Department of Health and Human Services and $75 million for the Department of Education.

THEORETICAL ISSUES

The CASSP initiative set forth the principles inherent in community-based systems of care. The key principles include access to a comprehensive array of services, treatment individualized to the child's needs, treatment in the least restrictive environment possible (with full utilization of the resources of the family and the community), full participation of families as partners in services planning and delivery, interagency coordination, the use of case management for services coordination, no ejection or rejection from services due to lack of "treatability" or "cooperation" with interventions, early identification and intervention, smooth transition of youth into the adult service system, effective advocacy efforts, and nondiscriminating, culturally sensitive services.

The wraparound approach is also critical for systems-of-care interventions and services. This approach is based on the principle of individualization of services, and proposes that an individualized service plan be "wrapped around" the child, including clinical, support, and nontraditional interventions. A wraparound approach is greatly facilitated by the use of noncategorical and flexible funding for services, which enables a case manager and family to purchase whatever services the child and family needs at that time (including living needs, educational, social support, and even staff security services). Wraparound plans can build a level of service intensity rivaling that of inpatient or residential settings without removing a child from his or her home, especially using systematic wraparound protocols that have measured fidelity.

Cultural competence is an integral element of community-based systems of care. It implies that practitioners in systems of care develop the necessary attitudes, skill, and knowledge base to serve minority and culturally diverse children and families in their communities, as well as

having service systems that develop policies and practices to remove barriers for access to services and make these more responsive to the needs and values of diverse communities. The community-based systems-of-care approach is congruent with the cultural values of ethnic minority populations, which emphasize strong extended family involvement in the life and upbringing of children and the use of natural community resources first in dealing with the emotional and physical problems of family members. These factors have been shown to be protective from some of the morbidities associated with emotional disturbance, such as substance abuse and suicidality. National guidelines for cultural competence for managed Medicaid services have been published by the CMHS, and the Practice Parameter adopted by the American Academy of Child and Adolescent Psychiatry (AACAP) operationalizes elements of cultural competence at both the clinical and systems levels.

Another important principle inherent in this approach is the prioritization of services to "seriously emotionally disturbed children." This designation includes the presence of an Axis I diagnosis under the criteria of the *Diagnostic and Statistical Manual of Mental Disorders* of the American Psychiatric Association but places equal emphasis on the child's inability to function in at least one of his or her life domains (e.g., school, home, socially with peers). This reflects the difficulty many children and their families have in accessing necessary services when the diagnosis was not "severe enough," particularly when the child was diagnosed with a disruptive disorder (although many children with disruptive behavioral disorders have comorbid serious mental illness when evaluated systematically). These problems are more pronounced in the diagnostic assessment of ethnic minority youth with serious emotional disturbance, with serious underestimation of comorbid affective and substance abuse disorders by clinicians. Other studies indicate that the level of care received by children is only partially accounted for by their clinical diagnosis, whereas the level of function and psychosocial stressors are stronger predictors.

The movement to managed care has resulted in the application of community-based systems-of-care principles to these approaches. In fact, these principles embody a more clinically based managed-care philosophy, that is, active care management becomes the primary modality for utilization management rather than arbitrary benefits restriction. In combination with the principles of care in the least restrictive environment, systems of care result in children actually receiving higher-quality services with fewer disruptions in their lives and development at a lower cost than with more restrictive, higher levels of care. The AACAP has developed guidelines for states using managed care for children's behavioral health services funded by Medicaid, integrating community-based systems-of-care principles into their contracts with private managed care providers or into

their policies and procedures for public managed services. It also has developed (in conjunction with the American Association of Community Psychiatry) an evidence-based level of care determination tool, the Child and Adolescent Service Intensity Instrument (CASII). It uses level of function and contextual variables as the main variables for level-of-care determination, with recommended levels of care based on treatment intensity using systems-of-care and wraparound approaches. The CASII has demonstrated significant inter-rater reliability and validity with mental health, child welfare, and juvenile justice populations, and has contributed to significant systems-of-care reform in the many states, particularly the state of Hawaii.

More recently, the Federal Affordable Care Act that has promoted health care reform has promoted the establishment of Medical Homes, where individuals (including children and youth) receive their primary care and also the coordination of specialty care. This has facilitated a movement in both general and child-and-adolescent psychiatry toward collaborative and integrated behavioral health models, to enhance collaboration with pediatric primary care physicians in Medical Homes and to enhance access to entry-level community child psychiatric services. The full implementation of these models calls for systematic screening of behavioral health needs served by the Medical Home and a stepped model of care, where either: (a) the primary care physician directly delivers psychiatric care in collaboration with a care manager, (b) the primary care physician refers the child for stabilization by the child mental health service or specialist and receives the child back once stabilized, or (c) the primary care physician and child mental health service collaborate longitudinally in the child's care. The primary care physician is trained and/or supported by child psychiatric and mental health consultation in screening, assessment, and entry-level behavioral health services as delivered within the primary care setting. There is both ready access to child psychiatric consultation (direct and indirect) and case management support to enable the primary care physician to deliver these services within the medical home and assist with specialty referrals and continuity tracking of the child's course of treatment and outcomes. In integrated models, there is embedded case management and therapist support in the Medical Home, as well as access to regular child psychiatric consultation on the panel of patients served by the Medical Home.

TECHNIQUES

A community-based treatment approach to services begins with access to services facilitated through multiple portals of entry, including direct referral or referral through child service agencies (such as the school or juvenile justice, child welfare, or public health agencies). A triage system should direct children and families to services and providers appropriate to

their special needs. A full interdisciplinary team, including a child and adolescent psychiatrist, should be involved in the development of triage procedures and protocols. Comprehensive mental health assessments should be made by licensed providers who are credentialed and trained to evaluate children and families, in consultation with a child and adolescent psychiatrist. Such assessments should include clinical/diagnostic, functional, developmental, family, and cultural/community evaluations of the needs *and* strengths of the child and family. Whenever possible, standardized measures should be used that are reliable, valid, clinically useful, culturally competent to the population being served, and useful for programmatic evaluation.

Children and families who require time-limited, less intensive services should have an initial care plan developed by a licensed mental health professional in consultation with a child and adolescent psychiatrist. The initial care plan for children requiring more extensive treatment should be developed by an interdisciplinary team, including all relevant child mental health and child service professionals. The parent(s), other relevant family members, and the child, if appropriate, should be members of this team. The care plan should identify problem areas and strengths that are relevant to the child and family's functioning, as well as community-based interventions to address these needs. The patient and his or her family should be equal participants with the clinical team in the development and implementation of the plan of care. In some programs, special "child and family teams" that are composed primarily of the child, family, and other natural supports in the community develop the outlines of care plans, and these are then negotiated by case managers with interdisciplinary teams of professionals and agencies. Plans of care should be individualized to the needs of the child and family, including attention to cultural issues, and should follow the child continuously throughout all levels of care, being amended as needed. The plan of care should promote continuous coordination and integration of all applicable services and agencies that have responsibilities for providing services to the child and family (such as primary health care, juvenile justice, schools, and social services).

Treatment and other services should be specific to the unique needs of the child and his or her family and use approaches that maintain or reintegrate the child into the family and community. Child and adolescent psychiatrists, psychologists, social workers, psychiatric nurses, and other members of the team should participate in the development of service protocols to ensure that the special developmental, physical, mental, and emotional needs of the child are addressed. The system should have as many diverse therapeutic modalities with children and families as possible, particularly those modalities that have demonstrated effectiveness with seriously emotionally disturbed children. These include Medical Homes, traditional and intensive outpatient services, nonresidential community-

based interventions (including home-based, day treatment, and school-based services), flexible support services, community-based residential services (such as therapeutic foster homes and group homes), and brief, acute inpatient services when needed. These interventions are aided by the increasing effectiveness of cognitive, behavioral, family-based, and psychopharmacological therapies, which should be integrated into all community-based interventions and levels of care.

Case management services are essential in community-based treatment, including assessment of problem areas, measurement of functioning level, determination of needs, linkage with necessary resources, care coordination, advocacy, and monitoring of services provided and their outcomes. The child and family should be assigned a case manager who has the necessary skills to ensure seamless coordination of care across all levels, different providers, and the full time course of illness/disturbance. The frequency and intensity of case management services are proportional to the clinical and psychosocial needs of the child and family. The patient/family should be participants in the case management process, with shared responsibility with case managers for using and coordinating services. Case management protocols should facilitate the coordination of multiagency, multisystem interventions, integrating services from the various providers and agencies (including primary care and health providers) and the development of wraparound plans that are truly individualized for the child and family's needs. These protocols should include access to flexible noncategorical funds to enable the purchase of flexible services needed for an effective wraparound plan.

ROLES AND FUNCTIONS OF PSYCHIATRISTS

Given the complexity of these new systems of care, the skills and roles that psychiatrists must display go far beyond the circumscribed clinical roles to which they are usually limited. These roles include serving as a front-line clinician, clinical consultant and collaborator to other professionals, clinical team leader, administrative leader in delivery organization/system, quality assurance/improvement consultant, consultant to interagency teams, and outcome evaluator/researcher in systems of care. Child and adolescent psychiatrists, given their broad biopsychosocial and developmental perspectives, should be best able to integrate and coordinate community-based treatment delivered by multiple professionals with diverse skills. However, given the emphasis on tertiary care models in the training of many child and adolescent psychiatrists, they often have not developed all the skills needed to serve as effective members and leaders of these new systems. Three publications from the American Academy of Child and Adolescent Psychiatry, *Guidelines for Training Child and Adolescent Psychiatrists in Community-Based Systems of Care,* the Practice Parameters for Child Mental Health in Systems of Care, and *Best*

Principles for Integration of Child Psychiatry into the Pediatric Health Home provide guidance on the skills, knowledge, attitudes, training experiences, and practices that psychiatrists need to function effectively within community-based systems of care.

CLINICAL ISSUES

Indications

Community-based treatment approaches are indicated for the treatment of the majority of children with moderate to serious emotional disturbance or mental illness, which is often associated with comorbidities and multiple psychosocial stressors and needs. It is most indicated for the prevention of out-of-home placement of such children and for the treatment of children served by multiple agencies with limited family resources. It can be contraindicated when the child's family is too fragmented or inaccessible. It often has to be modified with more intensive and restrictive treatment elements when the child faces imminent danger from whatever source (suicide, homicide, abuse, victimization, severe substance abuse), dangerous symptomatology, or serious decrease in function (such as acute psychotic symptoms or delirium). However, community-based treatment can be delivered in these instances as long as staff and physical security and safety needs are addressed through more intensive wraparound services (such as one-to-one staffing or on-site medical services). These services can be de-escalated when such symptoms or conditions are resolved (usually after brief stabilization) or the family is able to again function to an extent to which it can safely and effectively participate in the child's care.

Limitations

The community-based treatment approach can suffer from limitations when there is a lack of appropriately trained child mental health professionals who are conversant in evidence-based psychotherapies with children and youth and when there is lack of an adequate and accessible continuum of care with sufficient options for appropriate level of care intensity to meet the child and family's clinical and functional needs. Additionally, lack of adequate services coordination and integration can lead to significant care fragmentation. This can occur when psychiatric and nonpsychiatric services are segmented and there is lack of timely care coordination services, leaving children for inadequate periods of time in inappropriate levels of care intensity (either too low when they are developing more acute needs, or too high after no longer needed).

Complications

The misapplication of the community-based treatment approach can lead to children and youth and their families being underserved. This can result

from lack of access to necessary clinical and support services, lack of access to more intensive services, excessive pharmacological treatment to compensate for lack of evidence-based psychotherapy services, and lack of timely care management services.

Contraindications

Community-based treatment services are contraindicated for children who suffer from acute symptomatology. These include acute suicidality with imminent danger to self, psychosis or mania leading to risk to self or others or inability to be adequately protected/cared for, or substance abuse that leads to risk to self or others. Additionally, the child who is at risk of victimization in their home or community environment requires out-of-home placement. In all these instances, the community-based treatment approach calls for continued coordination with services in the home community and all possible efforts to return the child back to their family or to a normalized community setting with adequate supports.

Jose was a 14-year-old Latino who lived with his mother and older brother. He came to the attention of juvenile justice authorities for truancy from school, curfew violations, defacing a neighbor's property with graffiti, and stealing some money from another neighbor. A forensic psychiatrist was asked to evaluate him, and much to the psychiatrist's surprise, he found Jose to have significant anxiety symptoms that prevented him from attending school regularly, and the psychiatrist referred him to the local partial hospital program for evaluation by a child and adolescent psychiatrist and comprehensive treatment. He also appealed to the juvenile court judge that Jose be placed on probation with the expectation of attending treatment. On admission to the day treatment program and evaluation of the youth and his family, it became evident that Jose had not only a significant anxiety disorder, but also a learning disability that had not been adequately addressed and poor parenting from his mother. She seemed overwhelmed by raising her two teenage sons in the absence of the father that had abandoned them and as a result of her chronic depression. The staff of the partial hospital treatment program called for an interdisciplinary team meeting of the school, juvenile justice, and public child and adult mental health agencies to develop a coordinated, comprehensive wraparound treatment plan for the family. The plan included continued day treatment for Jose along with cognitive-behavioral therapy and pharmacotherapy for his anxiety and disruptive behavior disorders and intervention for his learning disability. Parenting intervention and treatment for depression was recommended for the mother, but it became clear that this would have to be delivered in the home (because she would never come to the partial program) and by someone of the same ethnic background for it to have any chance of reaching this poorly motivated parent. A highly skilled Latina home-based family therapist was located in one of the local agencies, and she readily developed a close relationship with the mother. Through taking a maternal, supportive approach, this therapist was able to motivate Jose's mother not only to reassert her role as a parent, but also to resume treatment for her depression, which she had lapsed from for years. Jose was able transition back to his school and refrained from further offenses and illegal activities, and he and his brother were able successfully to follow their mothers' more effective parenting guidance.

Goals of Treatment

The goals of community-based treatment usually include the following: (1) to restore the child's function within the family, at school, with peers, and

in the community; (2) to maintain or reintegrate the child within the family and community and avert separation or disruption of the community tenure; (3) to individualize interventions to the unique clinical and psychosocial needs of the child and family; and (4) maximally to use the strengths and resources of the family and the community on behalf of the child.

ETHICAL ISSUES

Ethical issues can arise in community-based treatment when the child experiences serious symptomatology in the face of family disorganization or risk of harm within the family and home. The ethical conflict arises from the inherent tension between the child's needs for safety and security and the goals of retaining the child within their family or community, avoiding restrictive care, and avoiding disconnection from the child's natural supports. However, a home in which there is a high level of conflict or abuse can be more restrictive for a child than some out-of-home programs due to the degree to which the child's function and activities are constrained. On the other hand, the child may require a more structured or restrictive environment than the family can provide, such as when the child is experiencing suicidal, violent, or severe runaway behaviors. Interdisciplinary teams are usually well equipped to face such ethical dilemmas, given their representation of multiple stakeholders; chief among these is the voice of family and child advocates.

Another inherent clinical ethical dilemma arises when considering whether to place a child on a pharmacological regimen versus intensive community-based or psychosocial interventions for disruptive or dangerous behaviors or symptoms. In these instances, the tension is between the biases toward less "medicalized" psychosocial interventions over pharmacological interventions. The short- and long-term risks versus benefits (and evidence base) of any psychopharmacological intervention for a child should be carefully weighed. However, many community-based interventions can be adjunctively aided by targeted pharmacological interventions, which can further aid in the reduction of restrictiveness and integration into the community. The role of the psychiatrist in addressing these ethical dilemmas is critical. The utilization of some community-based treatments exclusively to reduce costs without considering the needs of the child and/or family can be another serious ethical breach in the pursuit of community-based treatment. Unfortunately, the financial incentives to provide less intensive care inherent in most managed care approaches can contribute to such dilemmas, potentially endangering the child if they place the child in greater danger or risk. Community-based approaches need to be pursued within organizational and financial structures that do not have arbitrary benefit limits and use on-site case management to avoid such situations.

RESEARCH AND EVALUATION

Research in child mental health services has particularly focused on the effectiveness of community-based treatment. Research in this area has included the evaluation of community-based residential approaches, nonresidential intensive community-based approaches, and the clinical, functional, and cost outcomes of community-based systems of care, as well as factors affecting the need and utilization of mental health services in different populations of children.

Community-Based Interventions

The Surgeon General reports on mental health and on youth violence point to research evidence supporting the effectiveness of a number of community-based interventions for children and youth, such as intensive case management, therapeutic foster care, partial hospitalization, and intensive in-home wraparound interventions (particularly multisystemic therapy). Other community-based interventions that show promise include school-based interventions, mentoring programs, family support and education programs, wilderness programs, crisis mobile outreach teams, partial hospitalization, time-limited hospitalization with coordinated community services, and family support services. Studies involving these modalities have demonstrated significantly better outcomes than traditional outpatient or residential services. These include reduced levels of externalizing and internalizing symptoms, improved family functioning, reduced utilization of more restrictive services, diversion from juvenile justice involvement or detention, and improved cost-effectiveness. In addition, a number of in-home preventive parent–child interventions with infants and young children have demonstrated long-term effectiveness in reducing child abuse, as well as adverse behavioral and academic outcomes (such as the Elmira Project by Olds and coworkers in upstate New York and the child abuse prevention projects in New Orleans by Boris Zeanah and colleagues). On the other hand, there is a lack of research evidence or actual lack of effectiveness for a number of more restrictive interventions, such as inpatient hospitalization in isolation of community services, long-term residential treatment, therapeutic group homes, and boot camps.

The wraparound model of care embodies this approach toward multimodal community-based interventions. This model emphasizes aggressive outreach, the use of least restrictive treatment options, and individualized, flexible, and unconditional child- and family-centered services for children imminently at risk or already in out-of-home placements. A number of studies of wraparound programs in different communities across the nation with diverse populations of at-risk children and families have reported positive outcomes in terms of reduction of externalizing behavioral problems, level of function, reduction in out-of-

home placement, improved family management skills and function, and consumer/family satisfaction. Multisystemic therapy (MST) combines home-based, wraparound, and cognitive-behavioral interventions within a systemic context. It has been evaluated with youth at risk of detention/incarceration and at risk of psychiatric or substance abuse hospitalization, with significant results in reducing out-of-home placement, externalizing problem behaviors, rates of recidivism, and costs of treatment. A more recent community-based intervention—Cognitive-Behavioral Intervention for Traumatic Stress—has demonstrated effectiveness in addressing traumatic stress among children of diverse cultural backgrounds in school settings.

Community Systems Evaluations

Outcome data on early system-of-care demonstration projects in Vermont, New York, Ventura County, California, and the eight Robert Woods Johnson pilot sites demonstrated reduced hospitalization rates and days, decreased out-of-home placement and less restrictive placements, significantly lower incidence of negative behaviors placing the child at risk for out-of-home placement, and significantly lower rates of overall problem behaviors. The Robert Woods Johnson pilot sites also demonstrated significantly lower costs using a managed subcapitation approach for funding flexible services for children with serious emotional disturbances. Studies of the Civilian Health and Medical Program of the Uniformed Services (CHAMPUS)–funded Fort Bragg Demonstration Project and the Stark County CMHS system-of-care site compared their outcomes to those of usual service sites. These programs achieved significantly lower restrictiveness of care, higher family and consumer satisfaction, increased access to services, and greater funding spent in less restrictive services. However, clinical and functional outcomes and costs were not significantly different overall. Critics of these studies point to their short-term nature, and more recent studies have supported that significant clinical and functional outcomes from system-of-care programs are found on longitudinal evaluation.

The Comprehensive Community Mental Health Services for Children and their Families Program has included both a centrally designed evaluation, with required systematic data collection on all children served, and locally administered evaluations. Two main designs have been used: a longitudinal design with multiple assessment baselines at program entry and every 6 months until 36 months (and after program exit in newer sites), and two sets of comparison studies of system of care demonstration and control community sites. Children and youth in the system-of-care programs have had significant results in improved child behavior, improved child functioning, fewer contacts with law enforcement, increased stability in living situation, improvements in child and family

strengths and resources, reduced caregiver strain, and improved family function. The most recent cumulative results from the national evaluation reported in 2005 indicate that the program had reduced inpatient and residential utilization by 25 percent, decreased juvenile probation involvement by 30 percent, decreased detention or other secure facility use by over 50 percent, reduced the number of children with suicide attempts by over two-thirds, increased regular school attendance by youth by more than 10 percent, and significantly reduced or stabilized behavioral problems in 90 percent of children served. Results have also demonstrated reduced cost of care within these programs when indirect cost offsets are considered, such as child welfare, juvenile justice, educational, and general health costs.

In addition, some of the current and past CMHS-funded sites have begun to evaluate the added effect of the integration of evidence-based psychotherapeutic interventions within the system-of-care model. Some system-of-care programs have begun to integrate evidence-based manualized therapies (such as parent–child interactive therapy) and community-based interventions (such as multisystemic therapy), demonstrating enhanced clinical outcomes. The Hawaii children's system of care has pursued the most ambitious integration of evidence-based interventions, creating a menu and an algorithm for the application of such interventions, with significant improved functional and clinical outcomes and decreased level of care intensity of services over time. However, there are major challenges in integrating evidence-based interventions into a community system of care, including maintenance of fidelity, need for added training resources, attitudinal changes by clinicians, and the overall complexity of the enterprise.

Services for Minority Youth

Services research literature with underserved minority youth has been sparse, and initially was focused on documenting racial differences in services provided. African American children tend to remain in foster care for longer periods of time and to have more foster care placements than white children. Studies have shown culturally diverse children to be underrepresented in mental health institutions and overrepresented in child welfare and juvenile justice settings and placements compared to nonminority youth, even when they are equally psychiatrically impaired. Ethnic and racial differences in the diagnoses of culturally diverse adolescents have been identified by some investigators, including overdiagnosis of conduct disorder and psychoses and underdiagnosis of affective, personality, and substance abuse disorders. Significantly lower rates for overall services utilization, and specifically for the treatment of depression among culturally diverse groups, have been documented. Some studies have found a significant relationship between race and the service

needs of racial ethnic minorities. Studies have also shown higher retention in treatment of minority youth if the clinician is of a similar ethnic/racial background.

Collaborative and Integrated Behavioral Health in Primary Care

Most of the research evaluations so far for collaborative and integrated behavioral health models have focused on their application to adult populations. There are a number of pilot collaborative care programs that have been developed which have demonstrated significant successes. For example, Massachusetts has successfully demonstrated the Massachusetts Child Psychiatry Access Project (MCPAP), a program that provides timely telephonic psychiatric and clinical guidance to primary care providers (PCPs) treating children with mental health problems. The program allows enrolled primary care physicians to get assistance for any child in their care. The program has enrolled most primary care practices, representing an estimated 95 percent of all youth in the state, and has high rates of PCP participation, with over 5,000 patients served from 2008 to 2009. Primary care physicians who use the program have reported significantly higher ratings of their ability to serve children with mental health problems. Similar statewide programs have been developed in New York State and Washington State, with similar results. A few studies have begun to demonstrate effectiveness of collaborative and integrated care models. For example, Kolko and colleagues studied children and their caregivers who participated from 8 pediatric practices and were cluster-randomized to Doctors Office Collaborative Care (DOCC; n = 160) or enhanced usual care (EUC; n = 161). In DOCC, a care manager delivered a personalized, evidence-based intervention. DOCC was associated with significantly higher rates of treatment initiation (99.4 percent vs. 54.2 percent) and completion (76.6 percent vs. 11.6 percent), improvement in behavioral problems and parental stress (P, 0.05–0.001), remission in behavior and internalizing problems, treatment response, consumer satisfaction, and greater perceived practice change and efficacy.

FUTURE RESEARCH NEEDS

There are still many areas of significant deficit in the evidence base for children's mental health services and systems of care. Areas in which further research is needed include the effectiveness of culturally competent services and culturally specific interventions, the effectiveness of school mental health services, the effectiveness of mental health services integrated within the juvenile justice and child welfare systems, the effect of different cost containment and reimbursement mechanisms, the effectiveness of a systems-of-care approach in combination with evidence-based interventions (both psychological and pharmacological), and the

integration of child mental health and primary care services. The application of population health and mental health services research findings in service system policy and planning are the ultimate objectives and value of such research.

References

Abram KM, Paskar LD, Washburn JJ, Teplin LA, Zwecker NA, Azores-Gococo NM; Office of Juvenile Justice and Delinquency Prevention; US Dept of Justice and United States of America. Perceived barriers to mental health services among detained youth. *OJJDP Juvenile Justice Bulletin*. 2015. Available at https://www.ojjdp.gov/pubs/248522.pdf. Accessed January 22, 2017.

Asarnow JR, Rozenman M, Wiblin J, Zeltzer L. Integrated medical-behavioral care compared with usual primary care for child and adolescent behavioral health: a meta-analysis. *JAMA pediatrics*. 2015;169(10):929–937.

Bookman-Frazee L, Garland AF, Taylor R, Zoffness R. Therapist's attitudes toward psychotherapeutic strategies in community based psychotherapy with children with disruptive behavior problems. *Adm Policy Ment Health*. 2009;36:1–12.

Bruns EJ, Suter JC, Leverentz-Brady KM. Relations between program and system variables and fidelity to the wraparound process for children and families. *Psychiatr Serv*. 2006;57:1586–1593.

Bruns EJ., Pullmann MD, Sather A, Brinson RD, Ramey M. Effectiveness of wraparound versus case management for children and adolescents: results of a randomized study. *Adm Policy Ment Health*. 2015;42(3):309–322.

*Burns BJ, Hoagwood K. *Community Treatment for Youth: Evidence-Based Interventions for Severe Emotional and Behavioral Disorders*. Oxford: Oxford University Press; 2002.

*Chorpita BF, Becker KD, Daleiden EL. Understanding the common elements of evidence-based practice: misconceptions and clinical examples. *J Am Acad Child Adolesc Psychiatry*. 2007;46:647–652.

Cohen R, Ventura AB. The status of care for at risk youth prior to the 1990's. In: *Transforming Children's Mental Health Policy into Practice: Lessons from Virginia and Other States' Experiences Creating and Sustaining Comprehensive Systems of Care*. Lanham: Lexington Books; 2017:3–16.

Daleiden EL, Chorpita BF, Donkervoet C, Arensdorf AM, Brogan M. Getting better at getting them better: health outcomes and evidence-based practice within a system of care. *J Am Acad Child Adolesc Psychiatry*. 2006;45:749–756.

Fallon T Jr, Pumariega AJ, Sowers W, et al. A level of care instrument for children's systems of care: construction, reliability, and validity. *J Child Family Stud*. 2006;15:140–152.

Franco E, Soler RE, McBride M. Introducing and evaluating parent–child interaction therapy in a system of care. *Child Adolesc Psychiatr Clin N Am*. 2005;14:351–366, x.

Holt W. *The Massachusetts Child Psychiatry Access Project: Supporting Mental Health Treatment in Primary Care. Case Study 32: High-Performing Health Care Organization*. Case Study 32. Boston, MA: Commonwealth Fund; 2010:18.

Kataoka SH, Stein BD, Jaycox LH, et al. A school-based mental health program for traumatized

Latino immigrant children. *J Am Acad Child Adolesc Psychiatry*. 2003;42:311–318.

McMahon T, Hamlett N, Olezeki C, Van Deusen T, Harris N, Flanagan D. Children, adolescents, and young adults in the publically funded system of care. In: Jacobs S, Steiner J, eds. *Yale Textbook of Public Psychiatry*. New York: Oxford Press; 2016;133–154.

Pumariega AJ. Residential treatment for youth: introduction and a cautionary tale. *Am J Orthopsychiatry*. 2006;76:281–284.

Pumariega AJ, Atkins DL, Rogers K, et al. Mental health and incarcerated youth. II: service utilization in incarcerated youth. *J Child Family Stud*. 1999;8:205–215.

Pumariega AJ, Diamond J, England M, Fallon T, Hanson G. *Guidelines for Training towards Community-Based Systems of Care for Children with Serious Emotional Disturbances*. Washington, DC: American Academy of Child and Adolescent Psychiatry; 1996.

Pumariega AJ, Fallon T. Psychopharmacology in systems of care. In: Pumariega AJ, Winters NC, eds. *Handbook of Child and Adolescent Systems of Care: The New Community Psychiatry*. San Francisco: Jossey Bass; 2003:120–148.

*Pumariega AJ, Rothe E, Mian A, Carlisle L, Toppelberg C, Harris T, et al; American Academy of Child and Adolescent Psychiatry (AACAP) Committee on Quality Issues (CQI). Practice Parameter for Cultural Competence in Child and Adolescent Psychiatric Practice. *J Am Acad Child Adolesc Psychiatry*. 2013;52(10):1101–1115.

Rowland MD, Halliday-Boykins CA, Henggeler SW, et al. A randomized trial of multisystemic therapy with Hawaii's Felix class youths. *J Emot Behav Disord*. 2005;13:13–23.

Stambaugh LF, Mustillo SA, Burns BJ, et al. Outcomes from wraparound and multisystemic therapy in a center for mental health services system-of-care demonstration site. *J Emot Behav Disord*. 2007;15:143–155.

*Stroul B, Friedman R. *A System of Care for Severely Emotionally Disturbed Children and Youth*. Washington, DC: Georgetown University Child Development Center/CASSP Technical Assistance Center; 1986.

US Department of Health and Human Services. *Mental Health: A Report of the Surgeon General*. Washington, DC: Center for Mental Health Services/Substance Abuse and Mental Health Services Administration/US Department of Health and Human Services; 1999.

US Department of Health and Human Services. *Culture, Race, and Ethnicity. A Supplement to: Mental Health: A Report of the Surgeon General*. Washington, DC: Center for Mental Health Services/Substance Abuse and Mental Health Services Administration/US Department of Health and Human Services; 2001.

US Department of Health and Human Services. Evaluation Findings: annual report to congress 2010. *Substance Abuse and Mental Health Services Administration, Child, Adolescent and Family Branch*. Washington, DC: Center for Mental Health Services/US Department of Health and Human Services; 2010.

Vaughan T, Pumariega AJ, Klaehn R. Systems of care under legal mandates. In: Pumariega AJ, Winters NC, eds. *Handbook of Child and Adolescent Systems of Care: The New Community Psychiatry*. San Francisco: Jossey Bass; 2003:414.

Winters NC, Chenven M, Pumariega AJ. Managed care and systems of care. In: Pumariega AJ, Winters NC, eds. *Handbook of Child and Adolescent Systems of Care: The New Community Psychiatry*. San Francisco: Jossey Bass; 2003:380.

*Winters NC, Pumariega AJ. Work group on community-based systems of care, committee on community psychiatry, work group on quality issues. Practice parameters for child mental health in systems of care. *J Am Acad Child Adolesc Psychiatry*. 2007;46:284–299.

▲ 54.9 The Treatment of Adolescents

STEVEN C. SCHLOZMAN, M.D., AND EUGENE V. BERESIN, M.D.

DEFINING ADOLESCENCE

Any discussion of the treatment of adolescents must begin with a definition of adolescence. Although most developmental theorists view adolescence as roughly coinciding with the onset of puberty, it is important to note that the exact criteria for both the onset and the conclusion of adolescent development are complex and multidetermined. Biological, psychological, social, and cultural paradigms all play a role in the conceptualization of what exactly constitutes adolescent development.

From a biological perspective, adolescence begins with the onset of puberty. It is important to note that puberty is not limited to the development of secondary sexual characteristics. In addition, neurodevelopment research suggests that during puberty the brain undergoes a number of important changes. These changes are most likely necessary for the emerging adolescent to move away from what Piaget characterized as a relatively rule-bound and concrete understanding of the world toward an increasing capacity for abstract thought. Neuron migration, changes in myelination, pruning of gray matter, and increases in white matter all lead to changes in frontal lobe activity and have been implicated as playing major roles in long-observed cognitive changes. However, there is also growing evidence for incomplete frontal lobe myelination leading to decreased capacities for complicated problem-solving activities in the presence of hotly affective environments. Thus, despite substantial development of brain regions responsible for executive function, the relatively poor myelination of these regions explains some of the classically observed lapses in judgment and some of the impulsivity that characterizes adolescent behavior. In this fashion, Piaget's observations have been validated and also challenged by a growing understanding of adolescent neurobiology.

Psychologically, adolescence has typically been viewed as a period of separation and individuation coupled with an increasing desire to form a coherent and consistent sense of self. These tasks, roughly characterized as *identity formation* by theorists such as Erik Erikson, Anna Freud, and Peter Blos, are central to the experience of the adolescent as he or she continues along the developmental trajectory. According to Erikson's epigenetic model, adolescence is determined by the quest for identity

formation at the risk of *role diffusion*. The onset of important advances such as the development of sexual and romantic attractions, the formation of political and ethical opinions, and the willingness to differ in opinion from both peers and adult caregivers are all key elements of these psychological changes.

In addition, G. Noam and others have proposed that during early adolescence, group identity and cohesion plays a much more important role than individual identity formation. Noting that belonging to a group and conforming to the norms of that group are important aspects of the early adolescent years, these scholars theorize that adolescents first experiment with separation and individuation by embracing the rules and regulations of their peer group. In "Eriksonian terms," this developmental challenge could be coined *inclusion at the risk of exclusion*. To this end, as they mature, healthy adolescents move away from all of the rules of any one peer group and instead choose those aspects of many different groups that best suit their developing sense of self. The normal early adolescent emphasis on popularity thus grows to also include the later emphasis on being comfortable that one can maintain a coherent identity without relying on the restrictions and rules of any single peer group.

Social norms and pressures contribute to adolescent development in many ways. Ideals of beauty, power, success, masculinity, and femininity are to a large extent socially determined. In modern industrial Western society, the adolescent's individual and group identities are significantly influenced by mobility and the increasing role of media as a means of information, entertainment, and communication. New technologies such as cellular telephones, broadband connections, and the Internet have expanded the adolescent's circle of relationships and generated instant access to groups of individuals, movies, music, news broadcasts, streaming video, and advertisements. Although group norms and role models previously were confined to relationships in real time, adolescents today are barraged with detailed and graphic information about the intimate lives of celebrities and other public figures in the news. Current events, often dramatic and traumatic events, are instantly available to the adolescent. The rapid transmission of information plays a major role in both defining the world for the adolescent and providing a constant stream of data through which the adolescent may become immersed and influenced in powerful ways. Although the media is no substitute for life experience in real time, the power of the media to create a social milieu with figures for identification, imitation, and role modeling has grown exponentially in recent years. For example, studies have demonstrated that the media has a powerful effect on adolescent body image and self-esteem and has complex manifestations across various ethnic, racial, and cultural groups. All of these factors have complicated the tasks of separation and identity formation in many ways that are only beginning to be understood, and

these social changes have led to worries regarding addictive Internet use, which seems to start most commonly during adolescence. These issues are addressed later in this section.

Finally, as for all developmental stages, adolescence exists within a specific cultural milieu. Although virtually every culture embraces rituals intended to recognize both the onset of adolescence and the transition to adulthood, the conceptualization of what characterizes adolescence varies in terms of cultural norms. What one culture might consider a normal development is often somewhat startling taken out of cultural context. As the world becomes increasingly multicultural, these considerations are especially important. For example, the Asian youth who begins to question his parents' values is often viewed very differently in Western versus Eastern contexts. Although someone with a Western outlook might view this as a healthy and normal emotional development, the youth's parents might see it as particularly pathological and dangerous. Cultural differences such as this are obviously generalizations, but keeping these potential differences in mind is extremely important to any clinical consideration of adolescent populations.

CONTROVERSIES IN CONCEPTUALIZING NORMALITY AND PATHOLOGY IN ADOLESCENCE

The adolescent's quests for identity and separation have led both clinicians and popular culture to view adolescence as a normally tumultuous period. Although most adolescents navigate the transition from childhood to adulthood without serious incident, there are times of upheaval that are experienced as real crises. These periods may be obvious sources of concern to significant figures in the teenager's life, such as parents, siblings, teachers, or other adults, or they may be concealed and go unnoticed. For most adolescents, the movement from complete dependence on family to independence is at once exhilarating and terrifying, and at other times it is filled with profound despair. Highly charged affect is quite typical of adolescents and is often intense and labile. In the separation process, the adolescent must in one way or another reject his or her parents and at other times seek parental approval and nurturing —even babying. It is not unusual, for example, to see an adolescent girl's bed covered with stuffed animals, and at the same time to overhear her talking about moving out of the house or fantasizing about sexual intrigues.

Although the normal adolescent is capable of understanding the consequences of his or her actions, teenagers typically live "in the moment" and are occasionally blind to risky behavior. Because normal myelination proceeds in a caudal-to-rostral pattern, emotionally charged moments may not enjoy efficient use of the more sophisticated regions of the brain until individuals are in their mid-20s. These neurobiological processes are compounded by cultural norms that create a sense of grandiosity and

invulnerability. However, at other times, the adolescent's omnipotence evaporates as the teenager fears a multitude of potential dangers, from physical illness to kidnapping. Finally, the quest for identity carries with it substantial self-consciousness, characterized by fear of public exposure and shame. These fears may result in reluctance to share feelings or discuss issues that are potentially humiliating.

In addition, all adolescent behavior is influenced by personal traits, such as temperament, and by the environment, including relationships with peers, siblings, and parents and the community in which the teenager lives. Consequently, there is a great deal of heterogeneity in adolescent behavior. To further complicate matters, developmental advances are never linear. Progress is punctuated by periods of regression to earlier stages. It takes considerable clinical acumen or a great deal of familiarity with a particular teenager to understand what is normal for one adolescent or pathological for another.

Some researchers feel that viewing adolescence as a tumultuous period creates a setting in which adolescent tribulations and "out-of-control" behavior are viewed as normative and therefore not deserving of clinical attention. Researchers such as Offer have instead suggested that whereas most adolescents are in fact well adjusted and not psychologically troubled, the tendency to view adolescence as "normally difficult" potentially risks missing the population of youth who are legitimately suffering. In fact, research suggests that approximately 20 percent of the adolescent population in Western settings is seriously in need of clinical intervention. Some of this attention is needed for managing particularly charged crisis points in normal development for youth at risk, whereas other attention is required to treat the onset of many psychiatric disorders that achieve greater prevalence during adolescence. Among these are mood and anxiety disorders, schizophrenia, eating disorders, and substance use disorders. The suicide rate increases tremendously throughout adolescence: Data from the Centers for Disease Control and Prevention (CDC) show that the suicide is the third leading cause of death in youth aged 10 to 24 years. In addition, data from the Youth Risk Behavior Survey (YRBS) published by the CDC show that 17 percent of teens "seriously considered" suicide during the 12-month period prior to September, 2012. There is also a continuing trend toward increasing numbers of Hispanic girls both considering and completing suicide. Thus, the psychiatric treatment of adolescents has significant public health repercussions.

Much of these challenges are driven by the paradoxical lack of careful thought among adolescents, who nevertheless clearly have the capacity to embrace rational and abstract problem-solving approaches. In short, often the overall judgment and risk assessment of adolescent thinking is markedly different from that of adults. For example, whereas adults would rarely accept a challenge to drive recklessly on the freeway, such a scenario

is not unheard of in late adolescent populations in spite of their growing capacity for sound judgment and abstract reasoning. Some researchers have suggested that the affectively charged environment of the adolescent leads to an increased willingness to take risks, sometimes called a “hot cognition” in the cognitive developmental literature. Again, this is best understood in the setting of the incomplete neurodevelopment outlined earlier. Thus, whereas adolescents might appear adult-like in their thinking during calmer moments, affectively volatile environments can pose a particular risk to his or her well-being. These observations serve to accentuate the need to view adolescent behavior as potentially in need of clinical attention.

Another area of controversy involves the interplay of media and popular culture in affecting adolescent behavior. Many researchers have suggested that violent or disturbing media increases the likelihood of psychosocial problems within adolescent populations. These studies have led to policy statements by a number of professional organizations, including the American Academy of Pediatrics and the American Psychological Association, in which suggestions for curtailing media impact are made clear. Other researchers have noted the potential use of the media in both understanding adolescent thinking and in allowing adolescents to explore difficult issues through displacement and projection. From a clinical perspective, therefore, the issue of media effects on adolescence is extremely important, and any discussion of the psychiatric treatment of adolescence must take into account this key controversy.

Among the major concerns surrounding digital media use in adolescence is the ways cell phones and social media are handled: 63 percent of teens exchange text messages daily, 22 percent log onto social media more than ten times daily, and 50 percent more than once a day. Over 90 percent of teens are online and 80 percent use social media such as Facebook, Twitter, Snapchat, or Instagram. There are considerable dangers of social media use in adolescents.

Though most teens think peers are nice, 88 percent witnessed others being cruel or mean. Considerable attention has been given to the serious impact of cyberbullying among adolescents. Cyberbullying increases dramatically from grade school through high school and may have significant psychiatric consequences, including depression, suicidal thoughts and behavior, anxiety and posttraumatic stress disorder, social isolation, low self-esteem, poor academic performance, and alcohol and substance use. Furthermore bullying has been shown to have lifelong psychiatric consequences, particularly anxiety and mood disorders. More than 90 percent of children use cell phones for going online and text messaging by Middle School.

Many teens reference their sexual behavior, substance use, and/or violence through digital media. Sexting is a particularly worrisome problem

in teens. Sexting is the use of text messaging or other means to post sexually explicit photos. Up to 20 percent of teens are found to engage in sexting, and while studies vary widely, probably 10 percent have sent an image, and as many as 20 percent have received images. The reasons for sexting are highly variable, but usually involve the desire to maintain or seek a relationship, peer pressure, boredom, joking around or showing off. While most teens feel the recipient would want to receive the photo and keep it to themselves, forwarding photos is a significant problem. Most teens are unaware that digital media affords recipients making a photo "go viral" and expand to a very wide audience. The consequences of such exposure may be profound and serious. Many teens who engage in sexting are deeply shamed, anxious, and depressed and few are aware that there could be potential legal problems including a risk of being charged with disseminating child pornography and placement on a sex-offenders registry if convicted.

While there may be dangerous consequences of social media, there are potential benefits. For many teens, the immediate contact affords nurturing social relationships, fostering separation from parents and increased autonomy, helping shy or anxious teenagers develop social interactions, and providing forums for discussing moral, social, or religious issues in concert with the adolescent's increased cognitive capacity. However, given the potential serious risks, it is imperative that clinicians know the media use of their patients. Media literacy is a critical and uncharted aspect of the care of adolescents, and one that should not be neglected.

Recently, there has been burgeoning concern regarding addictive Internet behavior among adolescents. Proponents of these concerns note that many people become enamored of Internet activity during adolescence at the expense of normal functioning. These adolescents are said to ignore more normal social behavior and spend increasing time engaged in Internet activity. Clinicians who are more skeptical of the concept of Internet addiction point to the lack of good prospective studies, confounding issues such as differing uses of the Internet among those who display worrisome Internet behavior, and poor understanding of molecular mechanisms such as dopamine reward systems that have been postulated in other addictive problems. In addition, what is considered socially normal increasingly involves the Internet. Thus, fundamental questions, such as how much time on the Internet is too much or whether activities such as Internet gaming and gambling hold the same risks as Internet pornography, need to be better explored.

Finally, in considering media effects on adolescents, one must consider the ongoing debate regarding massively increased use of video game activity among teens. Because many of these games adhere to increasingly realistic and complex narrative themes, there is concern that especially violent video games will desensitize adolescents to aggressive behavior.

However, there are investigations that suggest that video games offer opportunities for virtual displacement of aggression that might otherwise be acted out in the "real world." As with research concerning Internet use, data regarding video games are quite mixed, and much study is needed for a better understanding of the nuances of these activities.

WHEN DOES AN ADOLESCENT REQUIRE TREATMENT?

Psychiatric treatment is indicated for any individual when there is a disturbance in thought, affect, or behavior that affects normal functioning. For adolescents, this would include impairments in a teenager's biological functions, such as eating, sleeping, and other daily personal routines, as well as problems with relationships with family and peers, problems in school, and problems in occupational and recreational activities. Although it is true, as noted earlier, that normal adolescence involves a certain degree of emotional and/or behavioral volatility, most adolescents overall have a rather smooth course, punctuated by crisis points that usually resolve quickly.

The major issue for parents, caretakers, and clinicians is making the discrimination between a normative developmental crisis that will eventually self-correct and a behavioral change that indicates a deeper problem worthy of clinical attention. This distinction may not be easy for a number of reasons. First, there are often discrepancies between the way an adolescent looks and how he or she feels. For example, an adolescent with a depressive disorder who is contemplating suicide may appear no different than before the episode, somewhat quiet and contemplative, or mildly irritable. On the other hand, teenagers with profound insecurity and social phobia may present as arrogant and seemingly fearless. Second, although many adolescents are capable of effective communication, there is often a reluctance to share deeply troubling thoughts or feelings with friends, parents, or other adults. This may be due to developmental lags in expression, but more commonly it is due to the deep self-consciousness and feelings of vulnerability that are typical of normal adolescence. Consequently, many adolescents are misunderstood by parents and peers, which only compounds a sense of aloneness and personal distress. Third, there may be cultural taboos, stigma, or a sense of betraying family expectations that obstruct a teenager from expressing emotional or psychological problems. Despite these obstacles, there are important questions clinicians should keep in mind in determining whether an adolescent is in need of treatment:

- Has there been a developmental lapse, for example, a falloff from the trajectory with respect to functioning in school, home, friends, and recreational activities?
- What is the level of emotional distress?

- What is the meaning of the problem to the adolescent? Although a problem may seem minor for an adult, it may carry tremendous weight for the adolescent.
- How vulnerable is the adolescent to stress? Is there a range of effective coping skills, or is the adolescent limited in managing stress?
- What supports does the adolescent have with other adults at home or in the community?
- Does there appear to be a recent onset of a psychiatric disorder? Is there a past history of psychiatric disorder and/or treatment? Is there a positive family history of such disorders?
- How impulsive is the adolescent?
- Is there family psychopathology, such as mood disorders, substance use disorders, or severe conflict?
- Are there major changes in the family, such as divorce, unemployment, and so on, that might seriously disturb the adolescent?
- Are there concerns about physical or sexual abuse or neglect?
- Have there been recent losses in the family or in peer relationships, such as loss of a friend, romantic relationship, failure to make an athletic team, and so on? How vulnerable is the adolescent to loss?
- How capable is the adolescent in managing affect, such as anger or sadness?
- Are there other acute or chronic medical conditions in the adolescent or in the family?
- Are there cognitive limitations, such as mental retardation?
- Have there been any major media events, such as suicides, school shootings, terrorism, or copycat phenomena, that have a strong effect on the adolescent?
- Has there been a significant trauma to the adolescent outside of the family, for example, date rape, sexual harassment, bullying, scapegoating, and so on?

Answering these questions requires gathering data from multiple sources, including the adolescent, family members, teachers, pediatrician, medical record, or other members of the community. The research literature demonstrates that in children and adolescents there is a poor correlation between self-report and report from others regarding behavioral symptoms. Hence, data collection must include more than one informant.

MODALITY OF PSYCHIATRIC TREATMENT

The psychiatric treatment of adolescents can occur on a number of fronts. Treatment may focus on individual or group settings and may encompass interventions from pharmacological, psychosocial, and environmental

perspectives. Thus, the treatment of adolescents encompasses considerations of individual need, family and community resources, clinical impression, and biological history. Furthermore, treatment often brings together multiple aspects of the adolescent's world. Legal, academic, medical, and psychological issues all play potential roles in any clinical intervention.

Biological Interventions

Biological treatments primarily involve medication interventions. Agents such as antidepressants, conventional mood stabilizers, typical and atypical antipsychotics, anxiolytics, and psychostimulants all are used in adolescent populations. Obviously, choosing a pharmacological regimen is based on clinical and diagnostic impression, as well as the family's and the adolescent's willingness to embrace medication treatment. Furthermore, as for all medication treatments, both the patient and the family must have a sound understanding of the risks, benefits, and potential limitations of this particular modality. Finally, any consideration of medication treatment in the child population should include an understanding that randomized and controlled trials of specific medications are not always available and that many of the most commonly prescribed medications have not been approved by the U.S. Food and Drug Administration (FDA) for use in children. Nonetheless, the last 10 years have seen a dramatic increase in clinical investigation into the efficacy of pediatric psychopharmacology, and clinicians should understand that keeping abreast of these rapid developments is paramount in pursuing medication options for adolescents. In addition, the FDA black box warnings associated with the increased risk for suicide in antidepressant use in children and adolescents need to be taken into consideration when prescribing these agents. Families must be told of the FDA warnings and helped to understand the risks and benefits of these medications despite these concerns. The American Academy of Child and Adolescent Psychiatry website has instructions for how to discuss these complex issues with parents and teens. The risk posed by having these warnings includes the documented decline in the overall willingness to diagnose and treat psychiatric conditions in young people. Finally, as with pharmacological treatment of adults, treatment of adolescents may be most efficacious when medications are combined with psychosocial therapies.

A 15-year-old boy was noted by his parents to have a difficult time in leaving his computer in the morning. Although the problem was initially relatively well managed with gentle reminders that he was going to be late to school, over time, the boy showed a constant tendency to check his standing on an Internet social networking site. This behavior resulted in frequent school tardiness, eventual school refusal, and increasingly poor hygiene. He was referred by his pediatrician for treatment with a psychiatrist, with whom he endorsed concerns that if he did not constantly update his website, his "social score" would decrease, and that this decrease

would translate into worsening social standing at school. He did not enjoy the constant maintenance of his website but felt powerless to stop and was ashamed of his behavior. He was started on fluoxetine and received a course of cognitive-behavioral therapy for obsessive-compulsive disorder. Over the next few months, his behavior gradually improved, with decreased anxiety and greater social engagement.

Psychosocial Interventions

Individual and group interventions include psychoanalytic, psychodynamic, cognitive-behavioral, and interpersonal therapy approaches. All of these modalities, for both individual and group treatments, have been found to be effective for the treatment of a wide range of adolescent problems, including adolescent depression, anxiety disorders, substance abuse, and suicidality.

Psychoanalytic and Psychodynamic Therapies. Psychoanalytic and psychodynamic treatments seek to help the adolescent to come to terms with and thus better understand unconscious conflicts and maladaptive defense mechanisms. In addition, they may be useful in facilitating more-adaptive coping skills. These goals are accomplished by a close relationship between therapist and patient and the use of unconscious processes such as transference, countertransference, and resistance. The best candidates for psychodynamic and psychoanalytic psychotherapy are adolescents who have adequate cognitive and language functions to use insight, self-reflection, and an ability to communicate effectively with the therapist. They also must have sufficient ego strength to tolerate clarification, confrontation, and interpretation. Because psychodynamic therapy may precipitate some regression and will, at times, be stressful in uncovering conflicts, the adolescent should be able to tolerate affect and have sufficient impulse control to thwart acting out. Commonly used psychodynamic therapies include classical psychoanalytic psychotherapy, supportive psychotherapy and psychodynamic psychotherapies based on ego psychology, object relations theory, and self-psychology. As in the treatment of most adolescent disorders, psychodynamic psychotherapy is often conducted in conjunction with other modalities, such as pharmacotherapy, family therapy, and/or other psychosocial therapies. Although these forms of psychotherapeutic treatments are probably the most commonly applied interventions in the treatment of adolescents, the open-ended nature of these modalities has made randomized trials difficult. Many clinicians, however, feel strongly that a skilled psychodynamic treatment is immensely beneficial to suffering adolescents.

A 14-year-old girl gymnast began increasing her daily exercise and restricting her diet. She had never achieved menses, due to her low weight. She became fixated on the size of her thighs and belly and lost so much weight that her coach and pediatrician did not allow her to participate in athletics. She became increasingly terrified of getting fat, and she secretly exercised any chance she could. She was a perfectionist in academics, as well as in gymnastics. Because of excessive

weight loss, she was hospitalized, and the diagnosis of anorexia nervosa nonpurging type was established without comorbid disorders. Following hospitalization with modest weight gain, an outpatient plan was initiated including regular meetings with her pediatrician for monitoring of vital signs, weekly meetings with a nutritionist, weekly family therapy, and individual psychodynamic psychotherapy. In her psychotherapy, over the course of 2 years, she was able to understand that her perfectionism was a defense against low self-esteem, that she was very angry with her mother, who lived vicariously through her gymnastics and did not validate who she really was, and that she was, as an only child, the focus in her family, holding her parents' marriage together. Over time, she was able to understand the function of her anorexia as a way to avoid her maturation as a woman, as a dysfunctional means of developing some kind of identity, and as a passive-aggressive attack on her parents. With time, her weight increased, she was able to resume her athletics, and she developed close friends.

New research has focused on the process of mentalization and the theory of mind in the psychodynamic treatment of adolescents. Roughly speaking, mentalization can be considered the process by which individuals come to understand the thinking and emotional states of others. In a sense, the recursive thinking that becomes possible in adolescence is directly tied to this therapeutic skill. Thus, one attempts to build on the adolescent's capacity for empathy as his or her brain matures to the point at which empathy is more readily accessible.

Cognitive-Behavioral Therapy. Cognitive-behavioral therapy (CBT) has gained increasing popularity in recent years, and there are a number of studies suggesting its efficacy for problems such as adolescent depression, eating disorders, anxiety disorders, suicidality, and obsessive-compulsive disorder. CBT aims to help the adolescent to recognize cognitive distortions that lead to faulty interpretations and maladaptive behavior. Techniques such as relaxation and contingency reinforcement are used to help the teenager to tolerate affective charged environments. Treatment is often manual-driven, with a specific number of sessions and homework assignments between appointments. This form of treatment does not focus on aspects of unconscious motivations and conflicts. Instead it focuses on correcting distorted cognitions by modifying information processing.

A 13-year-old boy was noted by his parents to be extremely worried about contracting germs from his surroundings. He insisted on washing his hands more than 25 times a day and would often spend more than 1 hour a day making sure that his body and hands were clean and "free of contamination." After evaluation, the diagnosis of obsessive-compulsive disorder was made, and the boy and his parents opted for cognitive-behavioral therapy. The therapist started by helping the boy to rank his fears on a scale of 1 to 10, with 10 assigned to objects that he felt were most germ-laden and 1 assigned to those objects that were of little or no concern. The boy was then taught both relaxation techniques and how to engage in thought interruption, and he used these techniques in a stepwise fashion to overcome first the more benign elements on his list, then the most severe. Eventually, he was able to modify his behavior by combining his understanding of what worried him most with his capacity to master and eventually overcome his fears.

Behavioral Therapy. Behavioral therapy has been used for years in the

treatment of adolescent anxiety disorders, such as simple phobias, panic disorder, separation anxiety, and school refusal, as well as disorders of conduct and oppositionality. This modality is based on principles of operant conditioning and reinforcement and often uses important techniques such as systematic desensitization and contingency management. Behavioral therapy differs from CBT by focusing on behavior modification and not on issues such as faulty cognitions and distortions.

Interpersonal Therapy. Interpersonal therapy (IPT) was first developed in 1984 by Gerald Klerman and colleagues in an effort to provide a short, psychodynamically informed treatment of depressed adults. IPT is based on the central premise that depression occurs in the setting of interpersonal conflict and that therapy should focus in a time-limited fashion on specific interpersonal issues as a means of ameliorating depressed mood. Like CBT, IPT is manual driven and often includes homework between sessions. Although it has not been studied in the treatment of adolescents as much as CBT, IPT has been shown to be effective for the treatment of nonsuicidal, nonpsychotic adolescent patients.

Dialectical Behavioral Therapy. Dialectical behavioral therapy (DBT) was developed by Marcia Linehan, given the realization that patients with affective dysregulation, suicidal, and other self-destructive behavior did not respond adequately to CBT or psychodynamic psychotherapy. Her model involves replacing ineffective, maladaptive behavior in response to internal or external stimuli with skillful, adaptive behavior. In DBT, maladaptive responses are viewed as learned responses to painful affect. DBT for adolescents was developed by Drs. Alec Miller, Jill Rathus, and Marcia Linehan. It involves learning emotional regulation skills, developing strategies for stress tolerance, and enhancing interpersonal communication, and combines these modalities with mindfulness techniques. Typically, DBT involves utilization of an individual therapist and participation in a skills-based group, and requires that parents meet in groups, sometimes with their adolescents. One aspect of this form of therapy is the availability of a skills coach to help the adolescent "in the moment" to learn to utilize skills for mastery in real time.

DBT has been shown to be effective for adolescents with Borderline Personality Disorder, suicidal ideation and comorbid depression, bipolar disorder, eating disorders, and adolescent with impulsive and aggressive disorders.

Mindful Meditation. Mindful Meditation has become increasingly used as a treatment for a number of psychiatric and medical conditions in adolescents. While more research is needed, there is evidence for its

efficacy in treating anxiety, depression, pain syndromes, and use in modulating affective dysregulation, impulsivity and aggression. It is often a standard component of DBT for adolescents. Research has demonstrated that meditative techniques are helpful for general wellbeing and stress management.

Grouping psychosocial interventions into categories such as psychodynamic, CBT, IPT, and DBT formats potentially creates the false impression that practitioners practice distinct forms of each modality. Although most clinicians are most comfortable focusing on one or two of these therapeutic formats, almost all who treat children and adolescents use particular aspects of multiple treatments. Psychodynamic therapies can be aided by contingency management, and effective CBT is made possible with a strong and in some cases unconsciously motivated relationship. Indeed, Ablon and colleagues showed that the best indicator of success in all forms of treatment is less related to the therapeutic modality use than to universally important features such as the trust that the patient places in the clinician and the relationship that that the clinician has with the patient.

Environmental and Systemic Treatments

In addition to psychotherapeutic and biological interventions, the treatment of adolescents often includes therapeutic involvement in the adolescent's environment. *Family therapies,* for example, aim to help the adolescent to negotiate his or her changing role in the family dynamic. Because adolescence is characterized by a movement away from the family of origin, many families become potentially distressed and volatile as the adolescent asserts his or her opinions within the family environment. Family interventions help to clarify family responsibilities and responses as these changes take place.

Similarly, *school interventions* seek to help an adolescent's teachers and school to better understand the specific dynamics between an adolescent's behavior and the school's response to that behavior. School consultants can make huge changes in a teenager's ability to continue along the social and academic trajectory by helping teachers to better understand the adolescent's behavior, by recommending and in some cases providing outpatient evaluation, and by suggesting behavioral techniques within the classroom. Sometimes relatively small but psychologically informed interventions such as simply changing classrooms can be immensely beneficial for a troubled teenager.

Finally, there is a growing appreciation of the importance of bringing interventions to the environments in which adolescents are most troubled. *Multisystemic therapy (MST)* has gained increasing popularity for the treatment of conduct disorders among adolescent populations. MST is based on the premise that conduct-disordered teens need interventions in

schools, at home, in the office, and in the neighborhood. These interventions must be integrated and coordinated, and current studies document the efficacy of this multifaceted approach in reducing adolescent transgressions of conduct.

All forms of adolescent treatment are complicated by issues particular to adolescent development. Because much of the adolescent patient population is younger than 18 years old, parents must legally and ethically provide informed consent before allowing their children to be treated either with medications or psychotherapy. However, given the emphasis that healthy adolescents place on separation and autonomy, it is clinically paramount that this population understand the risks and benefits of psychiatric interventions and be willing participants in their treatment. Only in extremely rare circumstances, such as severe psychotic states or in the presence of dangerous suicidality, is it clinically useful to prescribe psychiatric medications and other treatments against an adolescent's will. In addition, adolescents may feel differently about their treatment when they are given the opportunity to discuss their feelings without their parents present. For example, side effects such as sexual dysfunction characteristic of many psychotropic agents need to be discussed with appropriate patients in the absence of their parents. In general, clinicians need to treat adolescents as much as possible like the adults they are striving to become, while at the same time keeping in mind all of the legal and ethical considerations that are paramount in the treatment of any underage patient.

Hospital and Residential Treatment

Most adolescents can be treated in the outpatient setting. However, at times, inpatient, residential, or day hospital care might be required. In recent years, largely due to third-party payers and the advent of managed care, access to these more restrictive environments has been quite limited. The primary indication for inpatient settings is typically danger to self or others. Adolescents might be admitted to short-term residential or day hospital programs, many of which are in unlocked settings, if there is an acute deterioration in behavior with an inability to function in the home and school environments. The most common diagnostic categories requiring hospitalization include severe depression, suicidal behavior, bipolar disorder, obsessive-compulsive disorder, psychosis, eating disorders, substance abuse, posttraumatic stress disorder, and conduct disorder. Inpatient, residential, and day hospitalizations offer the adolescent a therapeutic milieu for observation and multimodal treatment. The adolescent and family may use these services for medication management, behavioral containment, and the acquisition of better problem-solving skills. In many health care systems, a continuum of services has been established, with initial inpatient care leading to step-

down units, such as unlocked residential or day hospital programs that offer a less restrictive environment, and gradual re-entry into the community. A primary function of hospital-based care is to develop a comprehensive after-care program with coordination between outpatient services and utilization of school-based programs and other community services.

EFFECTIVE ADOLESCENT THERAPY

The treatment of adolescent patients is quite different from the treatment of latency-age or adult patients. Adolescents can seem both extremely childish and surprisingly mature. In addition, the same adolescent population may undergo rapid and unnerving developmental changes in a relatively short amount of time. These issues substantially complicate the clinician's task. Forming an effective alliance with an adolescent patient requires special attention to already important clinical considerations such as confidentiality and safety. For example, the patient who admits to using illegal drugs may not want his or her parents to know but may acknowledge a desire for help in understanding this growing drug problem. Although patient safety is always the most important clinical consideration, clinicians must view each situation on a case-by-case basis. Often, adolescents come to their treaters for guidance on issues that they are not able to discuss at length with their parents. A friendly, honest, flexible, and available clinician is absolutely essential for the adolescent to feel comfortable in coming forward with important and pressing issues.

Clinicians also need to be curious about their adolescent patients' worlds. Adolescents will express themselves wonderfully in displacement, finding voices in the music they listen to, in the television they watch, and in the stories and movies they embrace. Although it is neither reasonable nor possible for a clinician to stay abreast of the pantheon of popular culture, inviting adolescent patients to discuss aspects of their culture that excite them is a valuable means of building trust and of gaining insight.

FUTURE DIRECTIONS

The psychiatric treatment of adolescents is by definition a complex endeavor. Because adolescent behavior manifests as a function of biological, developmental, psychological, and cultural phenomena, the clinician must be prepared to take into account all aspects of the adolescent's world to effect change. Furthermore, no clear hierarchy exists among the various determinants of adolescent difficulties. The same behavioral problems among different adolescents can have strikingly different determinants. One irritable teen might have a biological predisposition to depression, whereas another might suffer primarily from a chaotic home life. A tendency to address only one aspect of an adolescent's presentation risks oversimplifying the extent to which

adolescents suffer and misses a valuable opportunity to be maximally effective. Indeed, it is this very complexity that makes the treatment of adolescents so rewarding and fulfilling.

References

Bostic JQ, Bagnell A. Psychiatric school consultation: an organizing framework and empowering techniques. *Child Adolesc Psychiatr Clin N Am*. 2001;10:1–12.

Centers for Disease Control and Prevention. *Youth Risk Behavior Surveillance System*. Available at http://www.cdc.gov/HealthyYouth/yrbs/index.htm.

Cuffe S, Hall W. Suicide and SSRI medications in children and adolescents: an update. Available at http://www.aacap.org/AACAP/Medical_Students_and_Residents/Mentorship_Matters/Devel

Davanzo PA, McCracken JT. Mood stabilizers in the treatment of juvenile bipolar disorder: advances and controversies. *Child Adolesc Psychiatr Clin N Am*. 2000;9:159–182.

DeAngelis T. Is Internet addiction real? *Monitor Psychol*. 2000;31(4). Available at http://www.apa.org/monitor/apr00/addiction.html

Englander E. *Research Findings: Massachusetts Aggression Reduction Center Survey 2011*. Available at http://webhost.bridgew.edu/marc/MARC%20REPORT-Bullying%20In%20Grades%203–12%20in%20MA.pdf

Erikson E. *Identity and the Life Cycle*. New York: W. W. Norton; 1979.

Fair DA, Cohen AL, Dosenbach NU, et al. The maturing architect of the brain's default network. *Proc Natl Acad Sci U S A*. 2008;105:4028–4032.

Freud A. *Normality and Pathology in Childhood: Assessments of Development*. New York: International University Press; 1965.

Goel V, Dolan RJ. Reciprocal neural response within lateral and ventral medial prefrontal cortex during hot and cold reasoning. *Neuroimage*. 2003;20;2314–2321.

Gogtay N, Giedd JN, Lusk L, et al. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci U S A*. 2004;101:8174–8179.

Gorrindo T, Fishel A, Beresin EV. Understanding technology throughout development: what Erik Erikson would say about toddler tweets and Facebook friends. *Focus*. 2012;10(3):282–292.

Groves S, Backer HS, van den Bosch W, Miller A. Dialectical behavior therapy with adolescents. *Child Adolesc Ment Health*. 2012;17(2):65–75.

Henggeler SW, Schoenwald SK, Borduin CM, Rowland MD, Cunningham PB. *Multisystemic Treatment of Antisocial Behavior in Children and Adolescents*. New York: Guilford Press; 1998.

Hofmann SG, Sawyer AT, Witt AA, Oh D. The effect of mindfulness-based therapy on anxiety and depression: a meta-analytic review. *J Consult Clin Psychol*. 2010;78(2):169–183.

Jellinek MS. Impact of media. *J Am Acad Child Adolesc Psychiatry*. 2001;40:1124.

Kelly JA. Children's adjustment in conflicted marriage and divorce: a decade review of research. *J Am Acad Child Adolesc Psychiatry*. 2000;39:963–973.

Klerman G, Weissman MM, Rounsaville BJ, Chevron ES. *Interpersonal Psychotherapy of Depression*. New York: Basic Books; 1984.

Ko CH, Yen JY, Yen CF, Lin HC, Yang MJ. Factors predictive for incidence and remission of

internet addiction in young adolescents: a prospective study. *Cyberpsychol Behav*. 2007;10(4):545–551.

*Kutner L, Olson C. *Grand Theft Childhood: The Surprising Truth about Violent Video Games and What Parents Can Do*. New York: Simon and Schuster; 2008.

Loeber R, Burke JD, Lahey BB, Winters A, Zera M. Oppositional defiant and conduct disorder: a review of the past 10 years, part I. *J Am Acad Child Adolesc Psychiatry*. 2000;39:1468–1484.

Malone CA. Child and adolescent psychiatry and family therapy: an overview. *Child Adolesc Psychiatr Clin N Am*. 2001;10:395–413.

March JS. Cognitive-behavioral psychotherapy for children and adolescents with OCD: a review and recommendations for treatment. *J Am Acad Child Adolesc Psychiatry*. 1995;34:7–18.

Noam G. Clinical-developmental psychology: towards developmentally differentiated interventions. In: Sigel I, Renninger KA, eds. *Handbook of Child Psychology. Vol. 4: Child Psychology in Practice*. 4th ed. New York: Wiley; 1998:585.

O'Malley F. Mentalizing in the psychotherapy of a disturbed adolescent girl. *Bull Menninger Clin*. 2003 Spring;67:150–157.

*O'Malley F. Contemporary issues in the psychiatric residential treatment of disturbed adolescents. *Child Adolesc Psychiatr Clin N Am*. 2004;13:255–266.

Pataki C, Bostic JQ, Schlozman S. The functional assessment of media in child and adolescent psychiatric treatment. *Child Adolesc Psychiatr Clin N Am*. 2005;14(3):555–570, x.

Pelham WE Jr, Fabiano GA. Behavior modification. *Child Adolesc Psychiatr Clin N Am*. 2000;9:671–688, ix.

Reinecke MA, Ryan NE, DuBois DL. Cognitive-behavioral therapy of depression and depressive symptoms during adolescence: a review and meta-analysis. *J Am Acad Child Adolesc Psychiatry*. 1998;37:26–34.

Sargent J. Variations in family composition: implications for family therapy. *Child Adolesc Psychiatr Clin N Am*. 2001;10:577–599.

Stubbe DE, Weiss G. Psychosocial interventions. Individual psychotherapy with the child, and family interventions. *Child Adolesc Psychiatr Clin N Am*. 2000;9:663–670, ix.

Villani S. Impact of media on children and adolescents: a 10-year review of the research. *J Am Acad Child Adolesc Psychiatry*. 2001;40:392–401.

Volkmar FR. Childhood and adolescent psychosis: a review of the past 10 years. *J Am Acad Child Adolesc Psychiatry*. 1996;35:843–851.

*Wiseman CV, Sunday SR, Becker AE. Impact of the media on adolescent body image. *Child Adolesc Psychiatr Clin N Am*. 2005;14(3):453–471, viii–ix.

55 Child Psychiatry: Special Areas of Interest

▲ 55.1 Adoption and Foster Care

SANDRA B. SEXSON, M.D., AND DALE AARON PEEPLES, M.D.

Adoption and foster care are social institutions that provide substitute care for children and adolescents who are unable to be cared for by their biological parents in their own homes. Foster care represents a temporary solution to out-of-home care, whereas adoption is aimed at a permanent assumption of parental responsibilities by parents other than the biological parent(s). Recent data suggest that more than 400,000 children are in formal foster care in the United States in a given year, with more than 640,000 being served by child welfare systems each year and close to 255,000 entering the system and slightly fewer exiting the system annually. All of these numbers represent a decrease from a decade ago. Informal fostering may also occur outside of the public system, adding to the numbers of children and adolescents affected. Statistics for adoption are even harder because adoptions occur in both the private arena and public child welfare system. It is estimated that approximately two and a half to three and a half percent of the US population is adopted. Child welfare data suggest that more than 50,000 children are adopted annually from the child welfare system. Best estimates suggest public adoptions account for 41 percent of all adoptions, with international adoptions accounting for 13 percent. The remaining 46 percent or so are adopted through adoption agencies and private lawyers. Child welfare statistics suggest that, although it has been trending down over the past decade, the number of children in the United States waiting in foster care for adoption at any one time remains well over 100,000.

The staggering statistics regarding foster care and adoption demand that psychiatrists and other mental health professionals address the complex needs of this vulnerable population of children and adolescents in the United States. Children in foster care frequently experience multiple placements, and many of those who are reunified with their families of origin subsequently are placed into care again. Children who are adopted

are less likely to have a disruption in placement. In recent years, parental neglect and parental inability to provide adequate care are the most common reasons that children are removed from their homes. Fifty years ago these reasons typically revolved around incapacitating parental illness or death or unstable financial situations. Today the reasons for parental incapacitation are heavily influenced by parental alcohol and drug abuse and all the iterations of poverty. Other, less common, reasons include physical abuse, abandonment, and sexual abuse. Although many children show resilience in the face of all of these stresses, the risk for mental and physical health problems in children in foster care and, to a lesser and somewhat more controversial extent, children in adoptive settings is well higher than that of the general population. Although it appears that many children enter placement situations with vulnerabilities in these areas, the consensus among professionals is that multiple disruptions in placements over time add to this vulnerability, increasing the risk for serious and prolonged mental health problems. This section reviews the history of adoption and foster care in the United States; identifies trends in the management of out-of-home placement, both foster and adoption, over recent years; and reviews the limited literature on outcomes, in an effort to inform the practicing psychiatrist how best to approach the care of this very vulnerable population.

FOSTER CARE

History of Foster Care

Throughout history there are accounts of children cared for by families or institutions other than their biological parents. Biblical accounts of Moses' rescue by the Pharaoh's daughter and biblical directives to men that they marry their brother's widow offer examples of nonrelated foster care and kinship care, respectively. For centuries, children have been abandoned by parents unwilling or unable to care for them. Early historical accounts suggest that churches and monasteries assumed responsibility for these children, who were often used as child labor, a practice that continued into the 19th century, when children were sent from cities, where they were seen as dangerous, to farm families, who took them in and used them as cheap labor. Older children, particularly, were placed in apprenticeships without supervision. These practices often removed children from their communities and families of origin permanently because of the distances involved in separation. In the 19th century, societies began to develop orphanages where children could be housed and cared for in groups, a practice that was found to be less financially expensive for the state. Pejorative attitudes toward abandoned children pictured them as coming from bad families and being at high risk of bad outcomes. Even during these earlier times, however, family reunification was a possibility. At the

beginning of the 19th century, mothers could reclaim their children by paying a fine and reimbursing the state for the costs incurred during the time the child was in the state's care.

In the United States in the middle of the 19th century, Charles Loring Brace developed a program to remove abandoned children from the streets of New York City and send them to families in which they often became child laborers and were permanently separated from their family of origin and their culture. However, his concept that these children were better served in family homes rather than institutions set the stage for foster care as it is known it today. Foster care, involving temporary out-of-home placement, became institutionalized in the United States around 1909. However, over the next decades children lived out their lives in foster placement with little attention being paid to permanency planning and with their situations frequently complicated by multiple placements, which often resulted in the children being lost in the system. In the early 1960s, the government instituted a plan (Aid to Dependent Children) to foster family preservation. *Permanency planning* to avoid the loss of children within the system became a goal, with reunification (return to the family of origin), adoption, or long-term foster placement as more permanent options. However, the emotional and developmental needs of children were not addressed until the Adoption Assistance and Child Welfare Act of 1980. This act was aimed at preventing unnecessary separation of children from their families by providing more services for the prevention of removal, as well as efforts toward reunification. It also sought to address the problem of the unstable and continuous cycle of children being in multiple foster homes over extended periods of time. It established mandated 6-month reviews of each child in foster care. Permanency planning was emphasized, with reunification the primary goal and adoption as the alternative. Over the years this act was expanded to provide adoption support for special needs children. Under Title IX of the Social Security Act, foster care children generally receive health care through Medicaid.

Many children who were removed from their homes were eventually returned to families of origin (up to 75 percent in earlier studies with around 50 percent as recent as 2013). A significant number of those reunited subsequently return to foster placement. The Adoption and Safe Families Act of 1997 sought to provide safeguards to protect children from neglect and abuse while promoting timely permanency plans and holding the child welfare system more accountable. Concerns that prolonged efforts toward reunification had been unsuccessful led to more stringent guidelines regarding permanency planning. States were required to move toward freeing a child for adoption when a child had been in foster care for 15 of the prior 22 months, with permanency hearings within 12 months after the child going into foster care. Kinship care in the form of permanent placement with a family member became a more popular option in lieu of

reunification. The limitations on the aggregate time in foster care also discouraged the prior practice of moving children from foster care back to the family of origin and then back into foster care again over and over, delaying decisions on other, more permanent options.

More recently child welfare legislation has attempted to address those foster children who "age out" of the system, that is, those who remain in the foster care system until transition into adulthood. Initial steps began in 1986 with the addition of an amendment to Title IV-E of the Social Security Act—the Independent Living Initiative. This provided states with funds for preparation for independent living as youth aged out of foster care. This was replaced in 1999 by the Foster Care Independence Act which increased funding to states to help with transition to independent living, and also provided the option for states to extend Medicaid coverage to the age of 20. In 2014, all youth aging out of foster care are eligible for Medicaid coverage until the age of 26 as a result of the Affordable Care Act of 2010. In addition to targeting access to health care, the federal government has also recognized the need of coordinating other services as children age out of foster care. The Fostering Connections to Success and Increasing Adoptions Act of 2008 requires the development of a transition plan at least 90 days prior to a youth aging out of foster care. Ideally this will help address housing issues, health insurance, education, mentoring, and employment services, as young adults exiting foster care have little or no "family-type" support. Each year more than 20,000 adolescents "age out" of foster care with limited connections or support systems.

Table 55.1–1.
Ethnic Representation of Children in Foster Care

Ethnic Group	Percentage of the Children in Foster Care	Percentage of the Children in General Population	Relative Risk of Foster Care Placement
Black non-Hispanic	26	13	About two times that of the general population
White non-Hispanic	42	63	About two-thirds that of the general population
Hispanic	21	17	Slightly above that in general population
Native American	2	1	About two times that of the general population
Biracial	6	2	About three times that of the general population

Who Is in Foster Care?

In the United States, there are currently more than 400,000 children in foster care. Although child abuse and neglect occur across all racial and ethnic groups, as a percentile, children of color are overrepresented in the

foster care system compared to the general population. Table 55.1–1 shows the comparisons relative to the percentiles in the general population, demonstrating a much greater risk for biracial, African-American, and Native American children than for white, and Hispanic children. Not only are there more African-American, non-Hispanic, and Native American children in foster care proportionately, but in addition, those in foster care are more likely to remain there for longer periods of time and are less likely to experience permanent reunification or adoption (Table 55.1–1).

It should be noted that there are some positive trends seen over the past decade regarding ethnicity and foster care. Although African-American children remain over represented in the foster care system, they have experienced the greatest reduction as a group of children receiving services in foster care, by over 47 percent. Almost all ethnicities of children have seen a drop in their numbers in foster care over the past decade. The one exception to this is biracial children, which had a 65 percent increase during that time. Additionally, Hispanic children experienced a marginal drop in true numbers (only 2.5 percent) but this is due to being members of a rapidly expanding demographic. These changes in the demographics of foster care are largely driven by local policies. Rates in large urban populations are responsible for much of the decline, as just 10 counties in the United States were responsible for half of the national decline. At a state level, 90 percent of the nation's decline is accounted for by changes in just 10 states. While California, New York, and Florida are responsible for the greatest decreases, the most growth in foster care placement was seen in Texas and Arizona.

Recently, a significant amount of media attention has been paid to unaccompanied minors crossing into the United States from the southern border. Children arrive in the United States from all over the world, and they are often fleeing desperate circumstances. At a bare minimum unaccompanied refugee minors experience the stress of adjusting to a new culture and separation from family. In addition, many have been exposed to war, rape, torture, and severe deprivation prompting their immigration to the United States. Subsequently PTSD, depression, and anxiety are common issues among refugee children. In 2002 Congress transferred the care of unaccompanied children from the Immigration and Naturalization service to the Department of Health and Human Services (HSS) in an effort to move away from an adult oriented detention model to a child welfare-based model. Unaccompanied children are defined as individuals under the age of 18 who have no lawful immigration status and who are without a parent or legal guardian. There are several options to house these children while finding more permanent placement-institutional shelters, group homes, independent living, and foster care. Placement in foster care is the most common of these, and the federal government has programs in 17 cities to arrange placement. Following that, placement with a sponsor,

family reunification, or return to country of origin are all possibilities. The benefits for these children as they enter foster care are the same as those for children born in the United States. State child welfare systems monitor their care while in the child welfare system. Additionally foster parents receive training for cultural sensitivity to help with the children's adjustment to life in a new country. All this is not without cost, and, in 2008, prior to the recent influx of children, approximately $127 million dollars a year was spent on the welfare of these children. Since then, the numbers of children entering the United States without an accompanying adult have skyrocketed. Historically, approximately 7,000 children a year were served, but beginning in 2012, the numbers almost doubled and estimates for 2014 number at over 57,000 children. Most of these children are males over the age of 14 who are found at the US border, primarily fleeing Guatemala, El Salvador, and Honduras. Eighty-five percent of youth served by HSS are placed with family members in the United States.

Native Americans have 565 federally recognized tribes. Among these there is a varying capacity of management of child welfare programs. Some Native American children are managed by state protective services. In the state system, they are over-represented relative to their portion of the population. Estimates about the number of children served by tribal organizations are unavailable, and therefore not included in national statistics. In 1978, the Indian Child Welfare Act was passed which reaffirmed and empowered Native American tribes to manage child welfare issues. It was put forth in response to the relatively high rates of Native American children in foster care, and the fact that they were often placed in homes away from their reservation and their culture. Specific plans for services vary from tribe to tribe, but federal funds are available and many tribes use supplemental funds as well. In addition to the challenge of operating a number of different models, services to children on reservations are also challenged by the rural nature of the setting, limited funding, and frequent turnover in local government, which can impact policy. Federal services are available to help Native American communities implement culturally sensitive practice model improvement projects.

Children of all ages are in foster care. Estimates in 2013 suggested that under 7 percent of those in care are less than 1 year of age, 33 percent are between 1 and 5 years of age, 22 percent are between 6 and 10 years of age, 22 percent are between 11 and 15 years of age, 16 percent are between 16 and 18 years of age, and close to 1 percent are older than 19 years of age. Many of the older children entered foster care younger than 1 year of age. Infants stay in care longer and are less likely to be reunited with their parents. Males continue to represent slightly over half of children in foster care at 52 percent.

As a group children entering foster care comprise a high-risk population for both physical and mental health problems. Studies suggest that one-

third of children in foster care have at least one chronic medical problem or disability and over 60 percent will have mental health problems during their lifetime. Estimates suggest that as many as 80 percent of children in foster care were exposed to drugs or alcohol in utero and 40 percent had low birth weight either from prematurity or in utero growth failure; both factors that increase the risk for many medical problems and developmental and mental health problems. It is not surprising that studies suggest that from 50 to 90 percent (compared to 10 to 20 percent in the general population and 35 percent or so in children living below the poverty level) of children who are placed in foster care have significant mental health problems necessitating mental health interventions. In addition, children involved in the child welfare system are at much higher risk for mental health problems than their socioeconomically matched counterparts. Of interest, those entering the foster care system are more likely to get mental health care than those who remain in their homes, regardless of clinical need, leading some to suggest that foster care is a gateway to mental health services for many children. Significant nonclinical predictors of the use of mental health services for children in foster care include older age and being white. In addition, children who have been sexually abused are more likely to access mental health services, as are children living in nonrelative foster care placements versus those in kinship care. Those placed for reasons of physical neglect are less likely to receive mental health services. Some studies also report decreased mental health services for younger children in foster care.

There is a growing body of literature illustrating the physiological changes children in foster care may experience due to early exposure to abuse and neglect. Neurodevelopmental impairment also results from early neglect. In addition to general growth delays, head circumference is decreased in children coming from institutional placements. The Bucharest Early Intervention Project (BEIP) is a longitudinal randomized controlled trial following children exposed to institutional care and comparing their growth and development to those who were able to find placement in foster care and a community control sample. In this study, when children were followed up at age 5, electroencephalography (EEG) patterns consistent with ADHD were present more often in institutionalized children than their peers, reflecting immature cortical development. In other adoption studies, both magnetic resonance imaging (MRI) and diffusion tensor imaging reveal decreased total brain volume in both white and gray matter. For children who found placement in foster care as opposed to an ongoing institutional setting, there was an increase in white matter volume suggesting some neural plasticity and the potential for some catch up once living arrangements improve. With respect to specific brain regions, studies are mixed regarding both the amygdala and the hippocampus, with some finding increased volume and others reporting decreased volume.

Despite undergoing similar experiences in terms of neglect and deprivation, outcomes for children do vary, suggesting that some individuals might inherit resistances to these stressors, while others might be more vulnerable to negative environmental effects. Polymorphisms in genetic sequences may offer some explanation for these different outcomes. Variation in the serotonin re-uptake transporter gene (5HTT) is associated with moderating the influence of early adverse experiences and the development of anxiety and depression. Studies in adopted children comparing groups adopted before and after the age of 6 months found that variation in the 5HTT allele did impact the development of emotional problems in children adopted after the age of 6 months. In the English and Romanian Adoptees (ERA) study, variation in dopamine reuptake receptors (DAT/DRD4) was associated with ongoing ADHD symptoms at the ages of 11 and 15, suggesting that children with the high-risk haplotype did not outgrow ADHD symptoms as rapidly as their peers. In the BEIP study, allele variation in the dopamine reuptake catechol-O-methyltransferase gene resulted in significantly lower rate of depressive symptoms in children with one of the variations.

Reviews of the literature consistently demonstrate that minority children, particularly African-American children, are less likely to receive services than their white counterparts are when both had high behavior dysfunction scores on behavioral screening instruments. This disparity suggests that caregiver, child welfare, and court official's attitudes regarding mental health service needs may be culturally biased, influencing which children in need get mental health services. Another concern in the mental health treatment of children in foster care has been the demonstrated increase in prescriptions of psychotropic medications to the foster care population, with probably less emphasis on psychosocial interventions. Given that the foster care population is very much at risk of mental health problems, it may not be surprising that this population receives psychotropic medications at rates of 37 percent compared to rates of 4 percent in the general population. The concern is whether evidence-based psychosocial interventions, which are often difficult to manage in unstable living environments or in homes with multiple foster children, are being employed when indicated. Certainly foster children are at higher risk for prenatal exposure to substances, malnutrition, and family history of mental illness, all of which could necessitate a pharmacological intervention. However, psychological factors such as grief, adjusting to a new environment, and attachment issues may manifest as internalizing and externalizing symptoms causing a psychiatrist to rush to medicate when a psychosocial intervention is in order. This is a particular risk when the history is frequently obtained from sources that only know a portion of the child's history.

Reasons Youth Enter Foster Care

The reasons that children and adolescents require placement outside their families of origin are complex, involving problems with the parents, child factors that contribute to failure of the nuclear family, and societal or environmental conditions. Biological parents' problems may place their children at risk of neglect or abuse. Substance abuse, mental illness, intellectual disability, domestic violence, and lack of social support systems (low socioeconomic status and being a single parent) all negatively affect the parents' ability to provide adequate care for their children. The child may also bring factors, such as temperament, into the situation that increases the risk of parental failure. Children with difficult temperaments and those with the inability to self-regulate effectively make parenting more difficult. However, even the child with an easy temperament may be at increased risk of neglect in a chaotic and demanding family. Children with cognitive and physical disabilities also may be more difficult to parent. Children who have temper tantrums and oppositional behavior bring more stress to the already vulnerable parent–child interaction. The more needs a child has, the more difficult it is for parents to provide care that is adequate. Finally, adverse environmental conditions may contribute to the failure of capable parenting. Examples that increase stress in families include poverty, unemployment, decreased or absent social supports, and poor nutritional resources. Effective interventions to prevent the need for foster placement, as well as interventions aimed at providing substitute care, must address all of these components if they are to be successful. Too frequently the child is moved to foster care, but many of the issues involved in the ongoing lack of fit for the child in the family are not aggressively addressed.

Outcomes

Outcomes for children and adolescents placed in foster care are hard to assess, but ongoing studies of youth aging out of foster care have provided some information about their transition into young adulthood. At age 26, they are still disadvantaged relative to their peers in many respects. Five percent are incarcerated. Thirty percent have had recent episodes of homelessness or unstable housing ("couch surfing" with friends). Less than half are employed (even excluding those who are incarcerated), and of those working, a third are employed part-time. Relative to peers, they earn approximately $18,000 less per year and 45 percent have experienced an economic hardship in the past year (lacking money for rent/utilities). Over half have health insurance, but two-thirds rely on a government program for this. Foster children face a number of academic challenges. Multiple placements lead to school absences and disruption. Foster children tend to miss twice the amount of school as their peers, and 34 percent had more

than five changes in their schools. At the age of 18 foster care youth are twice as likely to have had out of school suspensions, and three times as likely to have an expulsion. Preparing for graduation, the average youth in foster care reads at a seventh grade level, and only 50 percent graduate high school on time. Of those that graduate 20 percent go on to college, with less than 10 percent of those enrolling in college going on to graduate with a bachelor's degree. Outcomes during foster care look at issues related to the disruption of care, given that improving placement stability is perceived as a major goal in foster care placement. Over a 12-month period, as many as 50 percent of children in placement experience a disruption, resulting in movement to another home or a more restrictive environment. Child characteristics thought to increase the risk of placement disruption include older age, general health problems, a mental health diagnosis, family history of mental health issues, and presence of behavior problems, especially those related to conduct disorders. Studies assessing the impact of gender are mixed, but recent reviews suggest females may now be at greater risk of disruption. Children removed for physical and sexual abuse are also more likely to have trouble with a placement compared to those removed for neglect. In addition, the experience of stress in the foster family in 6 months prior to the child's placement has been associated with increased rates of disruption, while foster parents that are trained to deal with special needs are more likely to continue caring for challenging children. Emerging data that links the number of behavior problems reported by the foster parent per day with increased disruption of placement suggest the need to focus interventions on reducing the child's behavior problems and increasing the foster parents' ability to deal with them. Children in kinship placements seem less likely to be disruptive than those in nonkinship foster care. Reasons for this are unclear. Typically there are fewer children in kinship placements, which may affect a family's ability to manage a given child's behavioral problems. In addition, attachments in kinship placements may be easier to foster. Disruptions appear to be less likely when foster parents are warm and child centered and when foster children are open to fostering. There are well-established, evidence-based interventions focused primarily on teaching parents to intervene appropriately with the children in their care. However, few of these have been implemented with foster parents on a wide-scale basis. Policy and systems also play a role in placement stability for children. There are times when the state may mandate a move for a child to reunite siblings, switch to a long-term placement, or merely try to deal with the reality of funding issues. When children have a high level of turnover in caseworkers, they are at greater risk of changing placements. Subsidies for guardianship and adoption placement, as well as hiring graduate level social work staff, are associated with improved outcomes in placement. Without a doubt, multiple disruptions lead to increased problems for the

child and for the caregivers.

Since the 1980s, child welfare systems have focused more on reunification with the family of origin as a permanent placement option. However, over time some leaders in child welfare have questioned whether this goal is always appropriate because many children who are reunited go back into foster care. Lower rates of reunification occur when the child is very young, an adolescent, or African-American. When the child is removed because of neglect there is a lower likelihood for successful reunification. Family variables that decrease the rates of reunification include poverty and failure to visit when children are in foster care. Longer time in foster care, as well as kinship placement, also lowers the reunification rates. Still, over the years, child welfare has assumed that reunification is the best outcome for children in foster care based on the need for continued contact with the biological family and the fact that extended periods in foster care usually lead to less stability. However, recent data has accumulated suggesting that reunification may not always be the best option. Reunited children in several studies have had more behavioral problems, engaged in more risk behaviors, and had poorer life-course outcomes. Better outcomes for those who were not reunited included higher scores on intelligence testing, increased general well-being, and fewer problems with repeated criminal arrests. Some of the major parental issues that lead to foster placement—specifically substance abuse—are not easily remedied in the parent, limiting the likelihood of successful reunification. Intensive interventions to support families with mental illness, substance abuse, homelessness, and severe poverty are not readily available in most settings. Reunification in the right situation is still a laudable goal, but child welfare agencies should openly confront the pros and cons in each situation in an attempt to make a permanent decision in a timely manner that is best for the individual child.

Over the past three decades kinship care has become more popular as a formalized placement alternative to traditional foster care options. Federal laws such as the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 and the Adoption and Safe Families Act of 1997 have given preference for child placement to kinship arrangements. Most states have options for reimbursing kinship care, although frequently at lower levels than nonkinship foster care. Theoretical advantages to kinship care include the preservation of the child's sense of family and cultural identity; increased access to relatives other than the kinship provider; greater stability in placement; maintenance of religious, ethnic, and racial values; and, potentially, the existence of previously established relationships with the children that may facilitate attachment relationships. Kinship care provides a way for the child's family to be preserved even when reunification with the family of origin is not possible. Children in kinship placements as compared to those in foster care also show greater

mental health stability and fewer behavioral issues. However, there are numerous concerns about kinship care. As a group, kinship caregivers are more likely to be single, older, with more health concerns, less education, higher rate of unemployment, and from lower socioeconomic status. Frequently, kinship caregivers are grandmothers who parented the foster child's parent, which causes some concern among child welfare agencies. Child welfare workers report having more difficulty supervising kinship placements, citing that kinship caregivers frequently do not adhere to parental visitation restrictions. Some of these negative differences may be alleviated with more equitable resources and support services and increased supervision commensurate with that provided to nonkinship foster parents. Continued research on kinship care is necessary to provide child welfare agencies with reliable data on the long-term outcomes of children reared in either kinship or nonkinship foster placements to inform future decisions.

Additionally, it is being recognized that the majority of kinship arrangements are informal, meaning the court and the public child welfare system have not helped broker the arrangement between the birth parent and the caregiver. Approximately 1.76 million children are in this type of living arrangement, six times as many when compared to families that have formalized the agreement. Unfortunately, this creates obstacles for both the caregiver and the child. Without a proof of legal rights, obtaining services ranging from health care to school enrollment may be impossible without assistance. Grandparents find themselves most commonly serving in the role of the informal kinship caregiver. Often they are reluctant to seek out legal action to terminate birthparents' legal rights. They retain hope that the social circumstances (often substance use) that led to the current living arrangement will correct itself, and the child will be able to live with the birth parent once more. Although this arrangement is often prompted out of compassion for the biological parent, it limits a child's access to social services, and places a child at risk in terms of permanency of placement, as they can be removed from the home at the whim of the birth parent. Policies vary state to state, but informal kinship caregivers have options beyond adoption that allow them to seek services for children. Unfortunately, these generally require some out of pocket expense for legal fees, which may be an obstacle to some families. Seeking guardianship will allow a caregiver to access services for a child, but in some states it would also require the termination of parental rights. Legal custody as opposed to guardianship is appointed by the family court instead of probate court, but otherwise quite similar in that it often requires the proposed custodian to provide proof of the birthparents' inadequacy, and in some instances parental rights will be terminated. Some states offer a de facto custody option. This is also managed by the family court and requires that the care provider prove that they provided for the child's needs for a period of time

(often 6 months to a year) without the involvement of the birth parent. Biological parental rights are not terminated, and no concerns about the birthparents' inability to raise the child need be discussed in court. However, once the ruling is made the birth parent is not allowed to remove the child from the de facto custodian without court approval.

Treatment Foster Care

Treatment foster care (TFC) refers to interventions focusing on at-risk children who are unlikely to otherwise be managed in the community and would thus find themselves in an institutional placement such as a mental health facility, juvenile corrections, or a secure residential facility. These are children who have experienced significant trauma, abuse, and neglect which have begun to manifest as major mental illness or budding antisocial behavior. TFC interventions are similar to a number of older services, such as "specialized foster care," "wraparound" foster care, and multidimensional TFC. All try to help at-risk children by developing strong community links, therapeutic training for foster parents, providing treatment directly in the home, limiting the number of children in the home, providing crisis intervention services, and treating the foster parent as a professional as well as providing additional payment to the foster family. Children show an improvement in social skills, a reduction in problem behaviors, and an increase in placement permanency. Such interventions may help reduce antisocial behavior, and subsequent incarcerations for youth in foster care. Additionally, children receiving such services have also shown improvements in school attendance, homework, and employment.

Mental Health Problems

Children in foster care struggle with many psychological issues because the reasons that they enter foster care have almost universally potential negative effects on psychological development and emotional adjustment. The events leading to the need for placement are adverse, often involving abuse, neglect, exposure to violence, and inconsistent parenting. Despite the fact that the foster placement may be a safer environment, it carries with it the loss of familiar surroundings and familiar caretakers, caretakers to whom the child usually has some attachment despite the abusive or neglectful relationship. In addition, foster placements typically are not planned but occur abruptly with little opportunity for preparation. How children handle these major stresses depends on the reasons for placement, the age and/or developmental level of the child, the old and new home environments, and the support systems available for the transitions. Each separation emphasizes the rejection and feelings of abandonment, compromising the ability of the child to develop effective emotional

attachments. Younger children are particularly vulnerable to attachment problems, which may affect their ability to relate effectively in the future. Their inability to trust their caregivers to provide them with consistent predictable care leads them to mistrust future relationships and exhibit behaviors of avoidance, anger, and/or oppositionality. These behaviors in turn make it hard for even the most nurturing of adults to continue to offer the child much-needed nurturance. Attachment to foster parents may also be complicated by complex and ambivalent feelings of being unfaithful to the biological parents when children open up to a relationship with a new caregiver. Typically children removed from their birth parents blame themselves, perceive themselves as unlovable, and feel guilty about the separation. When children cannot trust their caretakers, their psychological development is complicated by the use of primitive defense mechanisms, and they are at risk of developing pathological narcissism, which may lead to an inability to express empathy. Sometimes children have experienced care in their families of origin or in deprived placements that are so pathological that they may develop reactive attachment disorder or disinhibited social engagement disorder, which requires immediate and intensive intervention to avoid catastrophic outcomes.

Children who have experienced pathological care during their younger years also have difficulties regulating their behaviors, leading to problems with modulation of their emotions, aggression, and even attention and activity levels. Self-regulation is learned through a protective process with a predictable caretaker, an experience many foster children lack. In addition to the pathological care, there may be other risk factors for regulatory problems, including prematurity, low birth weight, and exposure to drugs and/or alcohol in utero. The world of the abused and/or neglected child is one of extreme unpredictability, which leads to high levels of anxiety and fear. They are unable to sleep, a problem that contributes even further to a pervasive inability to regulate behavior.

Not only have young children in foster care experienced loss of their primary caretakers, who, in many cases, were unpredictable and pathological in their caregiving, but many have also been traumatized overtly by physical and sexual abuse, as well as exposure to violence. Some may meet full criteria for acute or PTSDs, but many will have subthreshold symptoms that need to be addressed. The stress systems in the brain may be activated in such a way as to keep the child in a chronic state of arousal, an issue that more recent research suggests may impair the child's ability to manage stress in the future. Early and intensive interventions may remediate these problems at a neurobiological level.

Children in foster care also have much higher risks for developmental disabilities than children who are not in foster care. Risk factors that lead to foster care may affect development. These include such negative influences as direct effects of maltreatment (brain injury, spinal cord

injury, burns), as well as prenatal exposure to toxins such as alcohol and other drugs. Parental incapacities may have genetic components that affect the child as well, including intellectual impairments and language delays. Developmental disabilities also increase the risk that children will go into foster care and stay in out-of-home care longer, that foster placement will be less stable, and that other problems either within the child or the families involved will be exacerbated.

What about the Foster Parents?

Foster parents are asked to provide a nurturing and accepting environment for children who have had numerous negative and unpredictable experiences with adults in their environment, who come to a "temporary" relationship with ambivalence and difficulties with trust and attachment, and who then frequently test in numerous ways the foster parents' acceptance of them within their families. Foster families have to deal with frequent, occasionally less than helpful, intrusions into their lives by caseworkers and other supervisors of foster care while sometimes getting little if any support for this challenging and difficult task. Furthermore, they have to do this knowing that their attachment to the child is likely to be disrupted, often under situations that they do not support, such as sending a child back to parents who have a history of being abusive. Foster parents experience grief and loss as children enter and leave their care. They experience rejection and oppositionality from children under major stressful conditions. Although most placement agencies provide some education to prepare parents for the challenges of foster care, ongoing training specific to the individual child's needs and support systems to help the foster parent deal with their sense of loss and grief while continuing to provide nurturing support may improve outcomes for both the foster families and the children placed in their care. Professionals working with foster children and foster parents must address these important issues to facilitate optimum outcomes. The development of rituals for entering and leaving families and careful documentation through scrapbooks and photo albums may provide continuity in the life of the child moving across placements, as well as a historical record for the foster family.

The Psychiatrist's Role

The child and adolescent psychiatrist, along with the entire team of mental health professionals, is in a unique position to advocate for and provide effective interventions for the highly vulnerable population of foster children in the United States. In 2012, the American Academy of Child and Adolescent Psychiatry issued a guide outlining best practices in caring for children in foster care. Recent trends have led to increased use of psychotropic medications for children with Medicaid coverage. Due to a

combination of legitimate concern for safety issues surrounding some of these medications and state budgetary concerns, increasing attention is paid to prescribing practices of physicians treating children in child welfare services. Subsequently before considering a medication, it is imperative to conduct a comprehensive mental health assessment (including, whenever possible, reaching out to biological parents for history), and to identify any medical or developmental issues. Immediate intensive interventions are needed that address the breadth of the problem—not just prescribing medications, but also offering intensive parent training to foster parents specific to the child they must parent and implementing other psychosocial treatments along with psychopharmacological treatments that are evidence based for the disorders identified. Unfortunately evidence based nonpharmacological interventions are frequently unavailable or poorly reimbursed, including intensive parent training. Child and adolescent psychiatrists often struggle with little continuity of information about the child's functioning, therefore limiting appropriate decision making as well as few available resources for consistency in evidence-based psychosocial interventions.

When medication is indicated, educating all parties involved in the child's welfare is essential. Estimates of the number of children in foster care using these medications range from 13 to 37 percent depending on the study, compared to 4 percent of age matched peers. When it is determined that a child will benefit from a pharmacological intervention, consider a conservative approach. Avoid starting multiple medications or making multiple changes at once. Follow up with the patient in a timely manner when changes are made—ideally 2 weeks after an initial visit. If a medication does not appear beneficial, confirm that the dose and duration are appropriate, and then remove the medication by a gradual taper if it truly poses no benefit. When possible, use FDA-approved options and discuss "off label" use with all parties when such an approach is necessary. Whenever possible, become familiar with the preferred medications on the Medicaid formulary and use those. When polypharmacy is necessary, clearly explain the role of all medications to the child and the caregivers. In preschool age children, be very judicious about medication use.

It is necessary to keep the best interests of the individual child always at the forefront in assessment and intervention strategies and advocate for the child for permanency planning that considers the child's attachments, predictability, nurturance, support, and safety. Insisting that foster parents and families are involved in the treatment is also mandatory, avoiding the all-to-frequent custom of having a caseworker (or even a "transport" person with no information) brings the child to the psychiatric appointment without the involvement of the caregiver. Custody assignments should foster the child's sense of belonging within a system and should avoid disruption of strong psychological ties unless physical or emotional well-

being is compromised. Psychiatrists are uniquely qualified to evaluate these attachments and support systems and recommend and/or implement treatment interventions that are evidence based. In addition, as care providers, psychiatrists must advocate for the availability of effective services, which, while never readily available, may become even more limited as an effect of the expanding influence of Medicaid and managed care.

Looking to the Future

Although progress in understanding the problems and needs of children whose families of origin cannot successfully care for them continues, the field of foster care faces unanswered questions. Maltreatment and neglect have established negative effects on the growing child. Multiple placements with failure to provide consistent care over time definitely lead to negative outcomes. Nonetheless, many questions as to how best to intervene are often controversial and, more often, difficult to implement (even when evidence indicates an appropriate intervention) because of limited resources. Children and adolescents in foster care must have physical and mental health assessments and interventions beyond medication management from the very beginning of care. Foster parents need to be involved in these interventions. Agencies need to provide more education and individualized support to foster families who are providing the care for these challenging youth. Further research is needed to assess the viability of family preservation, early determination of the need for out-of-home permanency planning, and the advantages and disadvantages of kinship care over nonkinship care. Implementing more organized supports for adolescents who "age out" of foster care is a continuing challenge.

ADOPTION

History of Adoption

Adoption is a practice through which children whose parents are unable to provide for their care are supplied parents, often providing the adoptive parent(s) much-desired children. The practice is ancient, documented throughout history. Prior editions of this volume have summarized the history from ancient Babylonian times through today's practices. It was not until the 1850s that adoption was first recognized legally in the United States. Previously, informal adoptions often resulted in indentured servitude and child labor. In 1851, Massachusetts passed a law on adoption that integrated concepts that are still important. The law required judicial approval, biological parent or guardian consent, and the determination that the adoptive family was able to care for the child. Under these new laws, children were accorded legal membership in the family, although sometimes without access to inheritance. Children's needs became an

established priority.

Historically, adoption records, when available, were not confidential. Typically the birth certificate was not changed. In the early 20th century, social workers implemented the concept of the *closed adoption*, establishing secrecy among the triad of adoptee, birth family, and adoptive family. Social workers tried to match adoptees with adoptive parents based on race, religion, and physical and intellectual attributions. These efforts at secrecy were felt to protect the privacy of all involved, enhance permanency, and help the child avoid the social stigma of illegitimacy. Sometimes adoptees were not even told that they were adopted. Obviously many children were then, and still are, considered unadoptable. The practice of secrecy is controversial. Historically, concerns for all parties in the adoption process were made to argue in favor of secrecy. First, among these was the idea that the adopted child might experience confusion about their identity development and who their "real parents" were. Second, the adoptive parents might feel that they did not have the right to act as the child's "full parents." Finally, there was concern that the birth mother might not be able to move on from the loss of giving her child up for adoption due to the ongoing contact, thus prolonging her grieving process. Initial studies went on to refute such concerns. Additionally, as the face of adoption shifted in the United States, openness became unavoidable in many cases. Children adopted later in life through foster care are obviously aware of the family of origin. Similarly, children cared for by extended kin are often encouraged to maintain a relationship with the biological parents.

Open adoptions, in which records are open and adoptive children may actually have contact with biological parents, have become more common. The degree of openness varies from family to family, ranging from a one-time meeting between the biological mother and the adoptive parents to ongoing regular contact with the biological family. Although research is limited, recent studies eliciting feedback from adoptive parents have suggested that contact between the family of origin and the parents undertaking the adoption results in a higher comfort level with the adoption process for the adopting family, as well as instilling greater empathy in them toward the birth family. Adoptive parents have indicated that they are able to control boundaries with birth families, and generally note that they also feel they have a good level over control of communication. As this contact improved empathy toward the family of origin, it also may lead to more positive perceptions about the adopted child's behavior. Having an agreed upon and predictable structure for future contact is also seen as a positive to families. As contact continues, adoptive parents note fewer concerns that the biological parents may seek reunification, as trust between the two parties grows. Contact also reduces fears that adopted children have about their biological parents. They report less stress and worry about the safety and well-being of their family of

origin. Similarly, it is felt that some connection and openness can help them to better deal with mental health issues that the biological parent may be experiencing. This preliminary data are limited because the responses are almost totally from the adoptive parents, studies are small and nonstandardized, generally recruitment of subjects is through convenience samples, and most are focused on the adoption of white infants. Still the data offer some newer insights into the management of openness in adoption.

With societal changes, such as increased use of birth control, legalization of abortion, and the decreased stigma of single parenthood, the latter third of the 20th century saw fewer infants available for adoption, and both older children in foster care and children from other countries became potential adoptees for families in the United States. The goal for adoption of older children shifted from historical beginnings, when the benefit was for service to the adoptive family, to the psychological benefit of the child. Children from other countries began to be adopted in the United States after World War II and again after the Korean War, raising the controversy over transracial or transcultural adoptions, one that is not resolved. However, international adoptions have expanded rapidly over the last two decades, with Americans adopting children from many other countries, including China, Romania, Russia, India, Guatemala, and Colombia. In many cases, both with international adoptions and to a lesser extent with adoptions from foster care, the physical appearances of the child and the parent are obviously different, leading to the designation of "visible" adoptions.

Epidemiology of Adoption

Adoption is a well-established institution involving huge numbers of children and adults over the years. Although relevant statistics have not been routinely compiled, estimates suggest that between 2.5 and 3.5 percent of children in the United States are adopted, with more than 2 percent in nonrelative adoptive placements and another 1.5 percent in relative adoptions, including stepparents, who account for a significant number of "relative" adoptions. Foster care adoptions, although increasing somewhat with the push toward permanency planning, only represent about 15 percent of adoptions. Approximately 120,000 to 130,000 children are adopted each year, but this number has varied widely, with only about 50,000 annually in the 1940s and highs of 175,000 in the late 1960s and early 1970s.

Children are adopted in a number of different ways. Infants may be relinquished by their biological parents and adopted through private agencies or through attorneys, physicians, or clergy. These children are placed in adoptive homes either from the hospital or after a few weeks in "private" foster care. More and more of these adoptions are "open," with

contact initially with biological parents, particularly mothers, and in some cases, continued contact between the adoptive and biological parent(s) and in fewer cases ongoing contact. Although statistics on this group are not being collected, it is estimated that approximately 50,000 children per year are adopted in this manner. A second group comprises international adoption—the practice of adopting children from outside the United States. This group continued to increase over much of the last two decades often averaging close to 20,000 adoptions annually, until recent declines due to geopolitics. In 2013, there were just a little over 7,000 international adoptions. These adoptions typically are arranged through agencies, and total costs average a little under $30,000 and take 6 to 18 months to complete. The adopted children, although not newborns, are typically younger than those adopted from foster care, with almost 90 percent younger than 4 years of age and more than half of those under 1 year of age. These children may have had less than optimal care in orphanages prior to adoption. They were removed from their countries of origin and have little or no information about their biological families. Finally, children may be adopted from the foster care system. As many as 50,000 children or adolescents are adopted annually through the child welfare system. However, many more are waiting for adoption. These children typically have been in temporary placements, sometimes for a number of years. These adoptees are more likely to be from a racial or ethnic minority, and more than half are 6 years of age or older and may have experienced abuse and neglect, as well as multiple foster care placements.

Legislation continues to head in a progressive direction, gradually removing some of the barriers for adoption and opening the door for nontraditional adoptions, including single parent, gay, and transracial possibilities. As same sex marriages are still not recognized in all states, and it remains a relatively new phenomena in those states in which it exists, many gay and lesbian couples are unable to adopt children jointly. Fifty percent of gay and lesbian families that adopt go through the public system. Consequently, these families do a great service to many children remaining in foster care, and are generally also more willing to adopt transracial children. As openly gay families having the ability to adopt children is a relatively new phenomenon, longitudinal studies are just beginning to provide some outcome data. In early childhood, there appears to be no difference in internalizing or externalizing behavior by parent and teacher report in children raised in gay, lesbian, or heterosexual families. Further, there are no differences in gender role behavior in these children. Play preferences in young children appear less gender stereotyped in children reared by homosexual parents, particularly by lesbian mothers. Studies have also found that reduced parental stress is associated with better child adjustment. Consequently, state and federal policies that minimize discrimination against gay and lesbian parents can benefit their

children as well.

Approximately 40 percent of all adoptions are transracial, meaning that the race, ethnic or cultural origin of the child is different than that of one or both of the adoptive parents. This is most common in international adoptions, where 84 percent are transracial, but also seen in private adoptions (21 percent) and domestic adoptions (28 percent). Transracial adoptions in the United States initially involved only overseas adoptions, but in the 1960s, some transracial domestic adoptions began to occur. Parents were initially encouraged to de-emphasize the importance of racial differences to facilitate the child's integration into the new family. More recently, parents are encouraged to discuss race, and focus on the positives of cultural diversity, as well as acknowledging to their child that they are living in a racialized world to help validate any experiences with microaggression, or overt racism that may be directed at the child. Recently many studies have begun to focus on two issues- how to promote the development of a racial/ethnic/cultural identity for a child, and how to promote a child's awareness and ability to cope with racism. Contact with members of the child's racial community is helpful but occurs infrequently in transracial adoption. Having such connections may help children cope better with experiences of racism. Also of importance in the United States are the experiences of African-American children adopted into white families compared to African-American children adopted into African-American families. Researchers have described more identity issues in these populations, with African-American children reared in African-American families identifying themselves as African-American much more frequently than those reared in transracial settings. Findings are complicated because biracial children were included in the "African-American" population. Although highly contentious, both politically and socially, review articles conclude that transracial adoption leads to a majority of satisfactory outcomes, similar to other samples of adopted children after infancy.

Obviously adoptees enter the adoptive family with different life experiences. Those who are adopted as newborns experience family life and development much like children reared in biological families. However, the majority of others come from preadoption experiences of severe deprivation, traumatic life experiences, and multiple disruptions in caregivers before adoption.

Developmental Issues in Adoptive Families

Although most adopted children seem to progress along typical developmental trajectories, often even demonstrating developmental catch-up after adoption, there are some developmental challenges encountered in the adoptive family that are peculiar to the realities of the situation. Numerous authors suggest that there is an element of loss, embodied in the

process of relinquishment, in the psychological development of most adoptees, and that loss, particularly loss of the biological child that the adoptive parents were not able to have, may be an important psychological issue for the adoptive parents as well. These losses may affect the attachments of the child to the parents and vice versa, attachments that are necessary for typical development to progress. Benoit, Nickman, and Rosenfeld grouped these losses into three categories: *Overt* losses of environmental stability, including relationships; *covert* losses, which embody how the children think about the reasons for adoption and the effects that being "given up" may have on developing self-esteem; and finally *status* loss, which arises from feeling stigmatized by the fact of the adoption, a loss experienced generally but more commonly in transracial or international adoptees. These elements of loss may contribute to difficulties that the child may have in establishing strong attachments to the adoptive family. The parents may also have ambivalence regarding the adopted child, sometimes having failed to mourn the loss of their own fertility. This ambivalence, along with extended-family attitudes that may be unsupportive of adoption, can impair the adoptive parents' ability to attach wholeheartedly to the adopted child. Delays in finalization of the adoption may also cause adopted parents to have some reservations regarding attachment for fear of further disappointment and loss. Thus, one of the first major developmental tasks for the adoptive family is the facilitation of successful attachment. Attachment difficulties are more complicated when adoptions occur after multiple disruptions in care over extended periods of time or after prolonged periods of institutional care. In these cases immediate psychological interventions are imperative.

Adoptive parents have the additional responsibility of telling their adoptive children about the adoption and what they might know about the family of origin. Although in the past secrecy was frequently attempted, most professionals today recommend early disclosure and ongoing processing of the fact of adoption throughout the developmental years. Many feel that early disclosure, even before children can fully comprehend the information, is recommended. Addressing the reality of the adoption early and in an integrated manner allows children to develop healthy identities as adoptees as they progress developmentally, both cognitively and emotionally, over time and keeps the topic open as something easily talked about within the family. This may become very important when adopted children begin to wonder about their families of origin and begin to ask questions that the adoptive parents must answer as honestly and openly as possible. Anxieties about rejection may spill over into the adoptive family, causing children to worry about being removed from their adoptive homes. These issues are often heightened during adolescence, when children are normally struggling with issues of identity and separation from the nuclear family. Increased curiosity about families of

origin may be difficult for the adoptive families to address. Adolescents' typical desire to know more about their biological families, even to meet them, may feel like rejection to the adoptive families. When there is adolescent turmoil within families, adolescents may idealize the biological families as a way to exert independence within their adoptive families. It is unclear how more "open" adoptions, particularly those that foster open relationships with biological families, may affect these issues common in closed adoptions.

Over the developmental spectrum adoptive parents may find parenting challenges not experienced by biological parents. Frequently the adoptive parents enter the relationship with unresolved grief regarding their infertility, making them potentially less emotionally available to adopted children. Because of genetic differences as well as early experiences, the goodness of fit temperamentally and psychologically may not be optimal, making it more difficult for the adoptive parents to understand and respond empathically to the adopted children. Any problems the children have, any differences they may manifest, often tend to be viewed through the eyes of the adoption rather than as being normal variants within human development. Support for the adoption from extended families tends to decrease the effect of some of these challenges.

Risk for Mental Health Problems

It has long been established that adopted children are over-represented in both outpatient and inpatient mental health settings. Some of these findings predated the recent increase in numbers of children adopted from settings in which they had been neglected or abused prior to adoption. Some of this over-representation may result from evidence that suggests that adoptees are referred for problems of less severity than nonadoptees, leading researchers to hypothesize that adoptive families were more likely to seek interventions than nonadoptive families. Consistently, adoptees demonstrate more risk, however low, for externalizing behaviors, including substance abuse, conduct symptoms, aggression, and possible suicidal behavior, behaviors that are more distressing to parents and the community and more likely to lead to mental health referrals. Nonetheless, analyses controlling for the severity of the mental health problems demonstrated an increased risk in adoptees over that in nonadoptees. Researchers postulate that the reasons for this over-representation in clinical populations are complex and may represent an actual increase in risk for problems in this population along with a referral bias among parents who, for a multitude of reasons, including higher educational and socioeconomic levels, may be more willing to access mental health interventions for their children. Many of the studies suggesting increased risk fail to differentiate clinical from nonclinical samples, age of children when adopted, and the type of comparison group. Some suggest that the

over-representation statistics in clinical populations may be colored by a small group of very serious cases. Large nonclinical studies have yet to be completed, but the ones that have been done continue to show risk factors for the adopted population, albeit at lower levels. Drug use, antisocial behavior, and school adjustment are compromised, although adopted adolescents have fewer problems with withdrawn behavior and have increased prosocial behavior compared to nonadoptees.

Over the years it has been assumed that the older the child at the time of adoption, the higher is the risk for mental health problems. Studies in the 1990s suggested, however, that, although those children adopted as infants statistically demonstrated fewer behavioral and psychiatric problems, about 25 percent of them did have problems. Having a nonadopted sibling seemed to be a risk factor in these cases. In addition, studies have shown that deprivation in early childhood and later adoption resulted in the highest level of maladjustment. Yet, among these children, more than one fourth of them demonstrated no serious problems as adolescents. Other studies have suggested few problems in those children adopted as infants, and all tend to show increased difficulties in adjustment relative to the age of adoption, with infants showing the least problems and children adopted after the age of 10 years showing the most problems when compared to matched controls.

Comparison groups in adoption outcome research have been problematic as well. If the comparison group is that of nonadopted children from intact families, adopted children show more problems with adjustment. However, adopted children seem to have better adjustment than those raised in foster care or institutions and even those returned from temporary care to families where there is a history of and potentially ongoing impaired parenting. The complex nature of evaluating outcome demands addressing these issues, identifying risk factors that children bring to adoption, as well as risks within the adoptive family, and developing early interventions to help adoptive families and adoptees to address these risk factors early and intensively so as to improve outcome across the population.

Adopted children are frequently the source for research evaluating genetic versus environmental contributions to the development of mental illness. Researchers looking at schizophrenia have demonstrated that adoptees at high genetic risk for schizophrenia may be protected in optimal adoptive environments, whereas those who experience disordered parenting situations are more likely to develop schizophrenia, suggesting complex interactions between genetic and environmental contributions to the manifestation of the disorder.

International Adoptions

The practice of international adoptions has fluctuated over the last several

decades, with a gradual decline over the past 10 years. A recent high was in 2004, when 23,000 children were brought to the United States. More recent statistics place numbers closer around 7,000 to 9,000. Geopolitics play a strong role in the viability of international adoption, and if international relations sour, a large source of children available for adoption may suddenly become unavailable, as illustrated in Russia's ban of United States adoptions in 2013. The serious deprivation in the early lives of many of these children has focused attention on the importance of early childhood experiences. Developmental problems, particularly motor delays, as well as a broad spectrum of medical problems are commonly identified in children adopted internationally, particularly in those reared in orphanages. Approximately 12 percent of all adoptees have at least one chronic disability, twice the rate among all children in the same age range. Overall rates are similar among international and domestic adoptees. However, international adoptees who immigrated to the United States later in life, and consequently experienced longer periods of deprivation prior to adoption, have higher rates of impairment. The country of origin also influences disability status, as rates from children of Russian or Eastern European origin are much higher than those children hailing from China and Korea. Likely this is due to a greater reliance on institutional care in Eastern Europe as well as alcohol consumption rates that are twice as high in Russia as they are in China. All international adoptees should be given a full medical and developmental evaluation within the first 2 weeks after arrival in the United States, with interventions instituted immediately for infections, lack of immunizations, and early intervention for growth and developmental delays. International studies suggest that a large minority have increased mental health problems, particularly in the home setting. Preadoption conditions appear to be more predictive than age of adoption. The longer the child is in a deprived early environment, the greater is the risk for attachment difficulties. Interventions that foster attachment in these high-risk children should be implemented. Evidence suggests that children who are able to establish secure attachments with their adoptive mothers attain better developmental recovery, both cognitive and social, in middle childhood. Most studies show good outcomes for internationally adopted children. Factors that may increase the risk for developing mental health problems include male gender, older age at placement, preadoption care in an orphanage, and placement with a single parent or in a white collar family. The story of international adoption over the last several decades highlights the effect of extreme early deprivation and the possible recovery in developmentally supportive families. It has become apparent that adoptions after 24 months of age in these severely emotionally deprived children complicate recovery. Longer deprivation leads to more significant developmental delays and less complete recovery. In addition, recovery, when it occurs, typically does so within the first 2 years after

placement. The final chapter on long-term outcomes is yet to be written, but the evolving literature suggests that children severely deprived for long periods of time may continue to have problems with oppositional behaviors, attachment difficulties, odd or bizarre behaviors including social behaviors, and language delays. It is obvious that adoption improves the outcomes in many of these severely deprived children, many of whom are beginning to tell stories of the confusing move from their countries of origin, being placed frequently in families who do not speak a language they have ever heard, along with the abrupt discontinuation of any exposure to their language of origin. Those who have age-appropriate competence in their language of origin typically have better outcomes. Examination of the typical developmental process through which these young children adjust to English as a second language warns clinicians against diagnosing selective mutism in these children who are adjusting to a world totally dominated by a new language. Nonetheless, retrospectively these international adoptees seem to balance these confusing and benefit greatly from the more empathic and developmentally appropriate care in their adoptive families.

The Role of the Psychiatrist

Many child and adolescent psychiatrists work intensively in outpatient, inpatient, and residential treatment settings. Some are involved early in the adoption of high-risk children through consultation with schools, child welfare agencies, and juvenile court settings. Although there is little research in the area, case reports identify the need for empathy for both the adoptive parent and the adoptee. Psychiatrists working with adoptive families need to explore their countertransference issues related to adoption, including issues related to abandonment and rescue, which may impair their abilities to work with the dyadic issues in a balanced, empathic, and objective manner. It is imperative that psychiatrists educate themselves regarding the complicated research regarding outcomes to avoid presuming that adopted adolescents will have problems, a practice that may lead to self-fulfilling prophecies. Embarking on a psychotherapeutic process with an adoptee must fully include the parents to increase open communication within the family. At times adoptive parents harbor fears regarding the child's heredity, which may lead to the expectation for negative outcomes. Although without empirical evidence, group interventions for both adopted children and for their families may have potential as effective treatment methodologies. Issues that commonly arise in psychotherapy with adoptees involve problems with self-esteem and fears of additional losses or rejection. However, recent studies have demonstrated the development of good self-esteem in most adoptees. Issues related to curiosity about biological families are common concerns, sometimes leading to the opening of birth records without consent of the

birth parent. Another area of increasing focus relates to the common concern about treatment for attachment problems, which may be too frequently presumed in adoptive situations of early deprivation, causing clinicians to miss other treatable diagnoses. Controversial techniques such as holding have been discouraged by some national organizations. Newer interventions to foster attachment involving group educational programs on sensitive parenting coupled with focused feedback on the interactions between the parent and the child led to better infant outcomes, as well as enhanced parent–child attachments. Although the reality of adoption must enter into the psychotherapeutic equation, psychiatrists are cautioned to integrate issues related to adoption with an overall individualized assessment of the child in the family, avoiding the tendency to ascribe all psychopathology to issues related to adoption. Psychiatrists are obviously more likely to work with adoptive situations in which the children and families are not doing well. It is important to remember that, whereas adoptive status may influence the situation and negative experiences prior to the adoption exert great effect on the outcomes, each family needs to be evaluated thoroughly, and targeted interventions must be implemented in an effort to preserve the family because adoption disruptions in older adoptions are higher compared to adoptions in infancy. Data are difficult to come by regarding disruptions, but overall rates are around 10 percent.

Another consideration regarding treatment is that adolescence is the most challenging time for adopted children for both behavioral and emotional problems, regardless of the stage in life at which the child was adopted. This is partially due to the universal challenges of adolescence, but heightened due to the loss and trauma many adopted children experienced early in life. These experiences resurface during adolescence and can catch the adoptive parent off guard. Many adoptive parents envisioned a “happily ever after” scenario and believed that their love would be enough to help a child process issues related to grief, identity, and early childhood stressors. Consequently, adoptive families are four times as likely to seek counseling services and seven times as likely to seek residential treatment when compared to birth families. Problems may become so severe that a youth is at risk of re-entering foster care or the adoption is legally dissolved. From the limited data available, these outcomes are rare with approximately 10 percent of children who are adopted through the child welfare system finding themselves placed back into foster care. Of that 10 percent, approximately one-third may have the adoption terminated. Another 10 percent have extended disruptions to their living situation such as running away, living with another family member, or having an extended placement in a residential treatment facility or a juvenile corrections facility. As federal policy only began to encourage adoption from foster care in the 1980s and 1990s, postadoption services to families with special needs children are relatively new. They

tend to vary significantly from state to state, and it can prove a challenge to locate all services for families in need. Encouraging families to seek out services which improve their knowledge and skills is vital, as is reminding them that this is a sign of family strength, not parental inadequacy, as so many may believe. Education is where intervention begins and a discussion of the adopted child in light of the impact that early abuse and deprivation can cause will provide a greater understanding of the child, increase empathy, and help them look for ways to provide support and build on a child's strengths to promote resilience. It is also helpful to discuss realistic expectations of a child and to encourage parents to take care of their own well-being to promote better interactions with the child.

Psychiatrists may also work with the biological parents and families of children who have been given up for adoption or whose children have been removed because of abuse and neglect. These persons often have pre-existing psychopathology, which might become more complicated because of removal of their children. In addition, for those whose children are in foster care and for whom reunification is a goal, intensive psychiatric interventions and supports may be the most significant influence on ultimate outcome.

Future Directions

It is clear that for many children and families adoption meets a special need and facilitates a positive psychological outcome. For a minority, psychological adjustment is more challenging. Identifying more specifically the risk factors for less-than-optimum outcomes and developing evidence-based interventions to address critical issues like attachment, externalizing behaviors, and increasing empathic responses of adoptive families to high-risk children is the challenge for the future. Few current interventions take into account the complex nature of repeated trauma, attachment, and identity issues faced by adopted children. Policies that facilitate adoptions earlier when there is little hope for family reunification are imperative to interrupt the prolongation of temporary placements, which increase the risk of bad adoption outcomes. Additional work is also needed to provide adoptive families with post-adoption support to address the needs of the child as they are identified. Egregious policies that remain law in some states, which require adoptive parents to formally "abandon" children in order to obtain residential treatment, call for immediate correction. More needs to be learned about birth parents and optimal management of open adoptions or late reunions, which are becoming more frequent but about which very little is known. Finally, continued efforts are needed consistently to keep the best interests of the child at the forefront in all decisions regarding permanency planning, whether it is related to reunification, long-term foster care, or adoption.

REFERENCES

AACAP Committee on Community-Based Systems of Care. A guide for community child serving agencies on psychotropic medication for children and adolescents. 2012;1–24.

*Alavi Z, Calleja NG. Understanding the use of psychotropic medications in the child welfare system: causes, consequences, and proposed solutions. *Child Welfare*. 2012;91(2):77–94.

Carlson BE, Cacciatore J, Klimek B. A risk and resilience perspective on unaccompanied refugee minors. *Social Work*. 2012;57(3):259–269.

Carnochan S, Moore M, Austin MJ. Achieving placement stability. *J Evidence Based Social Work*. 2013;10(3):235–253.

Committee on Early Childhood, Adoption and Dependent Care. Developmental issues for young children in foster care. *Pediatrics*. 2000;106(5):1145.

Council on Foster Care, Adoption, and Kinship Care and Committee on Early Childhood. Health care of youth aging out of foster care. *Pediatrics*. 2012;130(6):1170–1173.

Courtney M, Dworsky A, Brown A, Cary C, Love K, Vorhies V. *Midwest Evaluation of the Adult Functioning of Former Foster Youth: Outcomes at Age 26*. Chicago, IL: Chapin Hall at the University of Chicago; 2011;1–112.

Gibson PA, Singh S. Let's help caregivers and children in informal kinship care: de facto custodian legislation. *Child Welfare*. 2010;89(3):79–97.

*Grotevant HD, McDermott JM. Adoption: biological and social processes linked to adaptation. *Ann Rev Psychology*. 2014;65:235–265.

Holtan A, Ronning JA, Handegard BH, Sourander A. A comparison of mental health problems in kinship and nonkinship foster care. *Eur Child Adolesc Psychiatry*. 2005;14:200.

Kreider RM, Cohen PN. Disability among internationally adopted children in the United States. *Pediatrics*. 2009;124(5):1311–1318.

Leslie LK, Hurlburt MS, James S, Landsverk J, Barth R. Outpatient mental health services for children in foster care: A national perspective. *Child Abuse Neglect*. 2004;28(6):699.

*Macdonald G, Turner W. Treatment foster care for improving outcomes in children and young people. *Cochrane Library*. 2008;1–76.

National Working Group on Foster Care and Education. Fostering success in education: national factsheet on the educational outcomes of children in foster care. 2014;1–20.

*Nickman SL, Rosenfeld AA, Fine P, MacIntyre JC, Pilowsky J. Children in adoptive families: overview and update. *J Am Acad Child Adolesc Psychiatry*. 2005;44(10):987.

Scannapieco M, Iannone MA. Native American Indian child welfare system change: implementation of a culturally appropriate practice model across three tribal child welfare systems. *Child Welfare*. 2012;91(3):157–172.

Sheridan M, Drury S, McLaughlin K, Almas A. Early institutionalization: neurobiological consequences and genetic modifiers. *Neuropsychol Rev*. 2010;20(4):414–429.

Simkiss DE, Stallard N, Thorogood M. A systematic literature review of the risk factors associated with children entering public care. *Child Care Health Dev*. 2013;39(5):628–642.

Smith SL, Maza P, Magruder J, Sciamanna J, Howard J. Keeping the promise: the case for adoption support and preservation. *The Donaldson Adoption Institute*. 2014;1–95.

U.S. Department of Health and Human Services, Administration on Children, Youth and

Families, Children's Bureau. The AFCARS report preliminary FY 2013 estimates as of July 2014. No. 21. 2014;1–6.

U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Children's Bureau. How many children were adopted in 2007 and 2008? Child Welfare Information Gateway. 2011;1–31.

U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Children's Bureau. Trends in foster care and adoption: FFY 2002-FFY 2013. (Based on data submitted by States as of July 21, 2014). 2014;1–3.

U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Office of Data, Analysis, Research, and Evaluation. Recent Demographic Trends in Foster Care. 2013;1–6.

U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Office of Refugee Resettlement. Unaccompanied Alien Children Program Fact Sheet. 2014;1–2.

U.S. Department of State, Bureau of Consular Affairs. FY 2013 Annual Report on Intercountry Adoption. 2014;1–5.

*Winokur M, Holtan A, Batchelder KE. Kinship care for the safety, permanency, and well-being of children removed from the home for maltreatment (review). *Cochrane Library*. 2014;1:1–239.

Wolfgram SM. Openness in adoption: what we know so far—a critical review of the literature. *Social Work*. 2008;53(2):133–142.

▲ 55.2 Child Maltreatment

WILLIAM BERNET, M.D., AND BRADLEY W. FREEMAN, M.D.

Although the federal government and every state have programs to investigate allegations of abuse, prevent abuse, and educate caregivers, child maltreatment remains a terrible misfortune for millions of children and families. Child maltreatment affects young people of all ages—infants, children, and adolescents—and all social classes, ethnic groups, and religions. For centuries, the recognition of child maltreatment has been compromised by pervasive denial. That is, perpetrators typically deny they abused anybody. Maltreated children sometimes deny they were abused. Mandated reporters and otherwise concerned citizens, such as friends and relatives of the victim, may fail to notice and underreport abuse either purposefully or unwittingly. There are sociological phenomena through which entire communities deny the reality of child maltreatment because denial protects individuals from the awareness of a horrible aspect of life.

A diverse group of professionals—including law enforcement, school personnel, pediatricians, dentists, emergency department workers, pathologists, radiologists, mental health professionals, and attorneys—participate in the identification, protection, evaluation, and treatment of

child abuse victims. There are several roles for psychiatrists and other mental health professionals in these cases. As clinicians, for example, they may provide treatment of children who have behavioral or emotional problems caused by abuse. They may evaluate and treat perpetrators of abuse, many of whom were victims of abuse earlier in their lives. Aside from their role as clinicians, mental health professionals may evaluate children for a forensic or legal purpose. The forensic evaluation may assist the court in determining what happened to the child. In addition, child psychiatrists may be asked to evaluate the credibility of children who may have been abused and the competency of children to testify. In civil lawsuits pertaining to child maltreatment, psychiatrists and other mental health professionals may testify about the nature, extent, and cause of the child's psychological injuries. Finally, mental health professionals may address the problem of child maltreatment on the level of public policy by educating legislators and judges about the psychiatric aspects of abuse and the developmental needs of children.

DEFINITIONS

Fifth Edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5)

The somewhat terse classification system provided by the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) lists *child physical abuse*, *child sexual abuse*, *child psychological abuse,* and *child neglect*. The four mental conditions appear in the chapter "Other Conditions That May Be a Focus of Clinical Attention" and in the section "Abuse and Neglect." These categories should be used when the focus of clinical attention is severe mistreatment of one individual by another. The coding system of DSM-5 allows the evaluator to distinguish each type of maltreatment: confirmed abuse from suspected abuse; an initial encounter from a subsequent encounter; parental abuse from nonparental abuse; whether the abuse previously occurred during the person's childhood; and whether the mental health services are for the perpetrator of parental or nonparental abuse.

Federal Law

The Child Abuse Prevention and Treatment Act (CAPTA) was passed in 1974, and has been amended several times, most recently in December 2010. In federal law, *child abuse and neglect* mean, as a minimum, any recent act or failure to act on the part of a parent or caretaker, which results in death, serious physical or emotional harm, sexual abuse or exploitation, or an act or failure to act which presents an imminent risk of serious harm. In federal law, *sexual abuse* means the employment, use, persuasion, inducement, enticement, or coercion of any child to engage in or to assist

any other person to engage in any sexually explicit conduct or simulation of such conduct for the purpose of producing a visual depiction of such conduct. Sexual abuse also includes rape (and in cases of caretaker or familial relationships, statutory rape), molestation, prostitution, or other forms of sexual exploitation of children, or incest with children.

State Law

A large mass of legal definitions and guidelines exists at the state level. The legal definitions of terms related to the maltreatment of children vary from one jurisdiction to another, so clinicians should be aware of the definitions used in their locale. The following generic definitions are used in this section.

Neglect

The most prevalent form of child maltreatment, *neglect* is the failure to provide adequate care and protection for children. Children can be harmed by malicious or ignorant withholding of physical, emotional, and educational necessities. Neglect includes failure to feed children adequately and failure to protect them from danger. Physical neglect includes abandonment, expulsion from home, disruptive custodial care, inadequate supervision, and reckless disregard for a child's safety and welfare. Medical neglect includes refusal, delay, or failure to provide medical care. Educational neglect includes failure to enroll a child in school and allowing chronic truancy.

Physical Abuse

Physical abuse may be defined as any act that results in a nonaccidental physical injury, such as beating, punching, kicking, biting, burning, and poisoning. Some physical abuse is the result of unreasonably severe corporal punishment or unjustifiable punishment. Physical abuse can be organized by the site of injury: damage to skin and surface tissue, damage to the head, damage to internal organs, and skeletal damage.

Psychological Abuse

Psychological abuse occurs when a person conveys to a child that he or she is worthless, flawed, unloved, unwanted, or endangered. The perpetrator may spurn, terrorize, ignore, isolate, or berate the child. Psychological abuse includes verbal assaults (e.g., belittling, screaming, making threats, blaming, or expressing sarcasm), exposing the child to domestic violence, overpressuring through excessively advanced expectations, and encouraging or instructing the child to engage in antisocial activities. The severity of psychological abuse depends on (1) whether the perpetrator intends to inflict harm on the child and (2) whether the abusive behaviors

are likely to cause harm to the child. In contentious divorces, one of the parents may indoctrinate the child to fear or dislike the other parent, thus causing parental alienation between the child and the rejected parent. Depending on the circumstances, parental alienation may be identified by one or more of these DSM-5 terms: child psychological abuse, parent–child relational problem, or child affected by parental relationship distress.

Sexual Abuse

Sexual abuse of children refers to sexual behavior between a child and an adult or between two children when one of them is significantly older or uses coercion. The perpetrator and the victim may be of the same sex or the opposite sex. The sexual behaviors include touching breasts, buttocks, and genitals, whether the victim is dressed or undressed; exhibitionism; fellatio; cunnilingus; and penetration of the vagina or anus with sexual organs or objects. Sexual abuse may involve behavior over an extended time or a single incident. Developmental factors must be considered in assessing whether sexual activities between two children are abusive or normative. In addition to the forms of inappropriate sexual touching, *sexual abuse* also refers to sexual exploitation of children—for instance, conduct or activities related to pornography depicting minors and promoting or trafficking in prostitution of minors.

Ritual Abuse

Cult-based *ritual abuse*, which includes satanic ritual abuse, is physical, sexual, or psychological abuse that involves bizarre or ceremonial activity that is religiously or spiritually motivated. Typically, multiple perpetrators abuse multiple victims over an extended period of time. Ritual abuse is a controversial concept; some professionals believed in the 1990s that ritual abuse was a common phenomenon, whereas others were skeptical about most allegations and descriptions of ritual abuse.

Perpetrators of Abuse

There is some lack of consistency in who may be defined as an *abuse perpetrator*. Usually, a person must be a parent or designated caregiver to be charged with neglect, physical abuse, or psychological abuse. Another adult (e.g., a stranger) who injures a child would be charged with battery, not with child abuse. However, a caretaker or any other person could be charged with child sexual abuse. State laws vary in this regard.

HISTORY OF CHILD MALTREATMENT

Historical and literary sources have documented child maltreatment for thousands of years in cultures from all parts of the world. Although some forms of abuse were considered deviant or prohibited in particular cultures,

most often the abusive practices were accepted or even encouraged by the societies in which they occurred. The long-term trend in European and American history has been to move from blatant, socially acceptable child abuse to strong societal disapproval. According to Lloyd deMause, at least four historical periods can be identified.

Ancient Period

In the ancient period, adults seduced, violated, and murdered their children in an unashamed and socially acceptable manner. Human sacrifice, including children, was practiced by Greeks, Romans, Celts, Phoenicians, and others. The Hebrew Bible criticizes human sacrifice several times, so it apparently was practiced during that era. Greeks, Romans, and Egyptians commonly practiced heterosexual and homosexual pedophilia. The most notorious sadistic child abusers were the Roman emperors Tiberius, Caligula, Claudius, and Nero.

Medieval Period

In the medieval period, Christianity induced guilt, and parents could no longer abuse their children with impunity. When they did, they asserted they were following God's will. The notion of an incest taboo—a universal prohibition against incest that presumably transcends cultures—is an ironic manifestation of societal denial. Some authors think the opposite is true—that incest has been universal for most people in most places at most times.

Early Modern Period

In the 18th through early 20th centuries, the sense of guilt and shame increased to such a degree that incest and child abuse could no longer be considered acceptable. However, abusive practices continued in the underworld, out of sight of polite society. Infanticide through exposure or poisoning was a long-accepted way to dispose of an unwanted or defective child. Foundling hospitals were established to care for thousands of abandoned babies, but few children survived the dismal routine of institutions. Charles Dickens (*Oliver Twist*, 1837 to 1839) and Victor Hugo (*Les Miserables*, 1862) laced their story lines with social criticism and graphically portrayed man's inhumanity. Dickens, for instance, recorded the inhuman aspect of industrial society, including child battery, debasement of children for sensual or recreational purposes, and starvation and deprivation of the laboring child.

Toward the end of this period, societal awareness of child maltreatment increased. In the 1860s, a French forensic pathologist, Ambroise Tardieu, described severe child abuse after performing autopsies on children who had been beaten to death. Child abuse came to public attention in the United States through the case of Mary Ellen, an 8-year-old girl who was

severely maltreated. She was discovered by church workers in New York City in 1874, but they found the only agency available to help was the Society for the Prevention of Cruelty to Animals. Thus, they founded the Society for the Prevention of Cruelty to Children. In 1875, New York became the first state to adopt a child protection law, which became the model for other states.

Late Modern Period

In the latter half of the 20th century, child abuse became the subject of widespread clinical and governmental concern. Abusive practices were actively identified, rather than ignored and denied. Modern recognition of child abuse was signaled in 1946 by a pediatrician and radiologist, John Caffey, who noticed a syndrome of children with multiple skeletal injuries and chronic subdural hematomas. Until the 1960s, physical abuse of children was considered rare, partly because physical discipline of children was generally more acceptable and partly because of societal denial concerning violence toward children. In 1962, Henry Kempe and his colleagues brought child maltreatment to the attention of physicians and legislators in a historic article, "The Battered-Child Syndrome." Within 5 years, child abuse reporting laws were adopted by every state.

It took a separate societal realization during the 1970s to acknowledge the extent of sexual abuse. Incest was known to occur, but most people believed that it was unusual and happened primarily among deviant families. It is now known that incest and other forms of sexual abuse are not rare.

In 1974, the federal government passed the CAPTA, which established uniform national standards for the identification and management of child abuse cases. Child abuse legislation, originally directed toward case discovery and perpetrator punishment, now focuses on prevention, rehabilitation, and family reunification. Individual states independently define maltreatment, formulate investigative procedures, collect data, and organize service delivery systems.

COMPARATIVE NOSOLOGY

The first edition of the DSM (DSM-I) (1952) and the second edition of the DSM (DSM-II) (1968) made no mention of child maltreatment. The Group for the Advancement of Psychiatry (1974) mentioned child maltreatment among the pathogenic factors of childhood mental disorders: "The deleterious effects ... of open rejection or neglect, occasionally involving willful injury or inadequate feeding, are well documented." However, the Group for the Advancement of Psychiatry did not include child maltreatment in their proposed system of classification.

When the third edition of the DSM (DSM-III) (1980) was published, the

American Psychiatric Association Task Force on Nomenclature and Statistics introduced a chapter called "V Codes for Conditions Not Attributable to a Mental Disorder That Are a Focus of Attention or Treatment." The V Codes, which were modeled on the practice in the ninth edition, clinical modification, of the *International Classification of Diseases* (ICD-9-CM), did not include specific categories for child maltreatment. However, it was intended that child abuse would be coded as a *parent–child problem*, the category used "when a focus of attention or treatment is a parent–child problem that is apparently not due to a mental disorder of the individual (parent or child) who is being evaluated." The DSM-III specifically stated that "an example [of parent–child problem] is child abuse not attributable to a mental disorder of the parent." The emphasis of *parent–child problem* was apparently on identifying perpetrators of abuse, not the child victims of abuse.

The revised third edition of the DSM (DSM-III-R) (1987) provided a similar definition for *parent–child problem*: "This category can be used for either a parent or a child when the focus of attention or treatment is a parent–child problem that is apparently not due to a mental disorder of the person who is being evaluated." However, the example involving child abuse was replaced by "conflict between a mentally healthy adolescent and her parents about her choice of friends."

In the fourth edition of the DSM (DSM-IV) (1994), the title of the chapter changed to "Other Conditions That May Be a Focus of Clinical Attention." In addition, a specific section was introduced, "Problems Related to Abuse or Neglect," which included terms related to child maltreatment. The editors noted that this category was "included because of the clinical and public health significance of these conditions." For the first time, it was possible to diagnose clearly *physical abuse of child*, *sexual abuse of child*, and *neglect of child*. Furthermore, the DSM-IV made it possible to distinguish whether the person being evaluated was the perpetrator of abuse or the victim of abuse. The terminology regarding child maltreatment in the DSM-IV-TR (2000) was the same as that in the DSM-IV.

The fifth edition of the DSM (DSM-5) (2013) includes child maltreatment in the same chapter as in the previous edition, "Other Conditions That May Be a Focus of Clinical Attention." A new section was included titled "Abuse and Neglect." This section encompasses abuse of all persons, young and old. Diagnosing abuse with the current DSM nomenclature has become more specific. The child can be diagnosed as being a victim of *child physical abuse, child sexual abuse, child neglect*, and *child psychological abuse*. It remains possible to differentiate whether the clinical service was for the victim or the perpetrator of abuse. For adults, the provider can indicate if the abuse or neglect is from the individual's childhood history.

Additionally, the diagnoses of parent–child relational problem and child affected by parental relationship distress are also included in the same chapter but under the section titled "Relational Problems." That section includes the language, "In the extreme, these close relationships can be associated with maltreatment or neglect, which has significant medical and psychological consequences for the affected individual." The description of parent–child relational problem in DSM-5 indicates the child is affected in either the behavioral, cognitive, or affective domain.

The tenth edition of the ICD (ICD-10, Version 2007) has a slightly different way of organizing child maltreatment diagnoses. The ICD-10 has a chapter titled "Injury, Poisoning and Certain Other Consequences of External Causes" and a section titled "Other and Unspecified Effects of External Causes." Within that section is a heading for "Maltreatment Syndrome," which includes the following diagnoses: neglect or abandonment; physical abuse; sexual abuse; psychological abuse; other maltreatment syndromes; and maltreatment syndrome, unspecified. The ICD-10-CM (Clinical Modification) includes codes that enhance the specificity of the abuse or neglect diagnosis. For example, the abuse can be coded as either *suspected* or *confirmed* and the type of encounter can be denoted as either *initial* or *subsequent*.

EPIDEMIOLOGY

Each year, the Children's Bureau, an agency within the Department of Health and Human Services, collects data on child maltreatment. The results are published in an annual document called *Child Maltreatment*. The agency estimated that, in 2014, approximately 6.6 million alleged victims were reported to child protective services. Of those reports, 702,000 victims were identified. The substantiated cases were distributed as follows: 75 percent neglect, 17 percent physical abuse, 8.3 percent sexual abuse, and 6 percent emotional/psychological maltreatment. The Children's Bureau estimated 1,580 children died as a result of maltreatment in 2014. Approximately 70.7 percent of these deaths were children younger than 3 years of age. Additionally, 12.2 percent of fatalities had been identified as victims and received family preservation services in the prior 5 years, and 1.8 percent had been removed from foster care and reunited with their families in the prior 5 years.

The data were analyzed for patterns of maltreatment by the sex and age of victims. With regard to the age distribution of victims, the age group from 0 to 1 years of age had the highest victimization rate (24.4 per 1,000 children). The majority of identified child victims were either White (44.0%), Hispanic (22.7%), or African-American (21.4%). African-American children had the highest victimization rate, 15.3 per 1,000, and American-Indian or Alaska Native children had the second highest rate at 13.4 per 1,000 children. The rate of victimization declined with increasing

age of victims; adolescents aged 17 years were maltreated at a rate of 3.5 per 1,000 children.

All of these figures are approximations because the actual amount of abuse is unclear. The reporting of abuse and the annual victimization rate increased during the 1980s and reached a high in 1993 (when the victimization rate was 15.2 per 1,000 children in the population). This increase resulted from greater public awareness and willingness to report child abuse, improvement in data collection techniques by individual states, and local economic conditions that placed a larger number of families under stress. The victimization rate fell after 1993 and reached a low of 9.2 per 1,000 children in 2012.

ETIOLOGY

Physical Abuse

Although child abuse occurs at all socioeconomic levels, it is highly associated with poverty and psychosocial stress, especially financial stress. Child maltreatment is strongly correlated with less parental education, underemployment, poor housing, welfare reliance, and single parenting. Child abuse tends to occur in multiproblem families, that is, families characterized by domestic violence, social isolation, parental mental illness, and parental substance abuse, especially alcoholism. The probability of maltreatment may be increased by risk factors such as prematurity, mental retardation, and physical handicap. In addition, the risk of child abuse increases in families with a larger number of children.

Sexual Abuse

Social, cultural, physiological, and psychological factors all contribute to the breakdown of the incest taboo. Incestuous behavior has been associated with alcohol abuse, overcrowding, increased physical proximity, and rural isolation that prevent adequate extrafamilial contacts. Some communities may be more tolerant of incestuous behavior. Major mental disorders and intellectual deficiency have been described in some perpetrators of incest and sexual abuse.

DIAGNOSIS AND CLINICAL FEATURES

Abused children manifest a variety of emotional, behavioral, and somatic reactions. These psychological symptoms are neither specific nor pathognomonic: The same symptoms may occur without a history of abuse. The psychological symptoms manifested by abused children and the behaviors of abusive parents can be organized into clinical patterns. Although it may be helpful to note whether a particular case falls into one of these patterns, that in itself is not diagnostic of child abuse.

Physically Abused Children

In many cases, the physical examination and radiological evaluation show evidence of repeated suspicious injuries. Abused children display behaviors that should arouse the suspicions of the health professional. For example, these children may be unusually fearful, docile, distrustful, and guarded. However, they may be disruptive and aggressive. They may be wary of physical contact and show no expectation of being comforted by adults, they may be on the alert for danger and continually size up the environment, and they may be afraid to go home.

The literature regarding the psychological consequences of physical abuse and neglect indicates a wide range of effects: affect dysregulation, insecure and atypical attachment patterns, impaired peer relationships involving increased aggression or social withdrawal, and academic underachievement. Physically abused children exhibit a range of psychopathology, including depression, conduct disorder, attention-deficit/hyperactivity disorder (ADHD), oppositional defiant disorder, dissociation, and posttraumatic stress disorder (PTSD).

Physically Abusive Parents

Abusive parents typically delay seeking help for the injuries. The history given by the parents is implausible or incompatible with the physical findings. The parents may blame a sibling or claim the children injured themselves. The characteristics of abusive parents include a history of abuse in their early lives, a lack of empathy for the child, unrealistic expectations of the child, and an impaired parent–child attachment, especially to children who are defective in some way.

Carol, 4 years of age, had a change in her behavior at preschool approximately 3 months after the birth of her sister. Her teacher saw Carol push other children and hit a classmate with a wooden block, causing a laceration of the child's lip. When Carol's teacher took her aside to talk about her behavior, she noticed what seemed to be belt marks on Carol's abdomen and forehead. The teacher reported possible child abuse to protective services. In addition, the family was referred for psychiatric evaluation.

Carol's baby sister was colicky and slept only for short periods of time throughout the day and night. She stopped crying only when her mother held her. Her mother therefore had little time for Carol, and Carol's father took over her care on evenings after day care and on weekends. He began to drink more than usual and became increasingly irritable. The parents argued over the mother's attention to the infant and the requirement that the father take care of Carol. Carol, who was a bright, curious, and talkative child, constantly asked questions and often asked to carry the baby. When refused, she would lie on the floor and have a tantrum. She also began to have difficulty falling asleep and awoke repeatedly during the night. Carol's father was unable to cope with her requests for attention and often told her to "shut up" and slapped her when she continued her demands. On many occasions, he responded to her tantrums or repeated questions by hitting her with his belt.

While protective services monitored the situation, Carol and her parents began a family therapy program that included parenting training and behavioral therapy for Carol, which was coordinated with the preschool. Carol's father attended Alcoholics Anonymous (AA) meetings and stopped drinking. He was able to control his anger at his daughter. Six months later,

Carol's aggressive behavior ceased. She was doing well with peers, was sleeping through the night, and stopped having temper tantrums.

Sexually Abused Children

A variety of symptoms, behavioral changes, and diagnoses sometimes occur in sexually abused children: anxiety symptoms, dissociative reactions and hysterical symptoms, depression, disturbances in sexual behaviors, and somatic complaints.

Anxiety Symptoms. Anxiety symptoms include fearfulness, phobias, insomnia, nightmares that directly portray the abuse, somatic complaints, and PTSD.

Dissociative Reactions and Hysterical Symptoms. The child may exhibit periods of amnesia, daydreaming, trance-like states, hysterical seizures, and symptoms of dissociative identity disorder.

Depression. Depression may be manifested by low self-esteem and suicidal and self-mutilative behaviors.

Disturbances in Sexual Behaviors. Some sexual behaviors are particularly suggestive of abuse, such as masturbating with an object, imitating intercourse, and inserting objects into the vagina or anus. Sexually abused children may display sexually aggressive behavior toward others. Other sexual behaviors are less specific, such as showing genitals to other children and touching the genitals of others. A younger child may manifest age-inappropriate sexual knowledge. In contrast to these overly sexualized behaviors, the child may avoid sexual stimuli through phobias and inhibitions.

Somatic Complaints. Somatic complaints include enuresis, encopresis, anal and vaginal itching, anorexia, bulimia, obesity, headache, and stomachache.

These symptoms are not pathognomonic. Nonabused children may also exhibit these symptoms and behaviors. For example, normal, nonabused children commonly exhibit sexual behaviors such as masturbating, displaying their genitals, and trying to look at people who are undressing.

Approximately one-third of sexually abused children have no apparent symptoms. Most adults who were abused as children have no significant abuse-related symptoms. However, the following factors tend to be associated with more severe symptoms in the victims of sexual abuse: greater frequency and duration of abuse, sexual abuse that involved force or penetration, and sexual abuse perpetrated by the child's father or stepfather. Other factors associated with poorer prognosis are the child's

perception of being less believed, family dysfunction, and lack of maternal support. In addition, multiple investigatory interviews appear to increase symptoms.

Long-Term Consequences of Maltreatment

Many adults who were abused as children have no significant abuse-related symptoms. However, epidemiological studies—such as the Adverse Childhood Experiences (ACE) Study—have shown a link between childhood adversity and adult psychopathology in diverse populations. It is likely this vulnerability is mediated by persistent physical consequences of the maltreatment on the child's developing brain.

Intrafamilial Sexual Abuse

Incest may be defined strictly as sexual relationships between close blood relatives, that is, between a child and the father, uncle, or sibling. Because of increased reporting, sibling incest is an area of growing concern. In its broader sense, incest includes sexual intercourse between a child and a stepparent or stepsibling. Although father–daughter incest is the most common form, incest can also involve father and son, mother and daughter, and mother and son.

Intrafamilial sexual abuse and other sexual abuse that occur over a period of time are characterized by a particular pattern or sequence of steps. Victims of sexual abuse recount a gradual progression of boundary violations by the perpetrator, starting with tiny invasions and escalating to serious, overwhelming intrusions. Healthy, self-confident children rebuff the intrusions directly (via temper tantrums and verbal disagreements) or indirectly (through silence and distancing maneuvers) or by adopting a strategy that causes the offender to refrain.

Sexual abuse that occurs over a period of time evolves through five phases: engagement, sexual interaction, secrecy, disclosure, and suppression.

Engagement Phase. The perpetrator induces the child into a special relationship. The daughter in father–daughter incest has frequently had a close relationship with her father throughout her childhood and may be pleased at first when he approaches her sexually.

Sexual Interaction Phase. The sexual behaviors progress from less to more intrusive forms of abuse. As the behavior continues, the abused daughter becomes confused and frightened because she never knows whether her father will be parental or sexual. If the victim tells her mother about the abuse, the mother may not be supportive. The mother often refuses to believe her daughter's reports or refuses to confront her husband

with her suspicions. Because the father provides special attention to a particular daughter, her brothers and sisters may distance themselves from her.

Secrecy Phase. The perpetrator threatens the victim not to tell. The father, fearful of his daughter may expose their relationship and often jealously possessive of her, interferes with the girl's development of normal peer relationships.

Disclosure Phase. The abuse is discovered accidentally (when another person walks in the room and sees it), through the child's reporting it to a responsible adult, or when the child is brought for medical attention and an alert clinician asks the right questions.

Suppression Phase. The child often retracts the statements of the disclosure because of family pressure or because of the child's mental processes. That is, the child may perceive that violent or intrusive attention is synonymous with interest or affection. Many incest survivors rally around their perpetrators, seeking to capture any modicum of tenderness or interest. At times, affection for the perpetrator outweighs the facts of abuse, and children recant their statements about sexual assault, regardless of substantiated evidence of molestation.

A financially comfortable couple lived in a pleasant, clean house in a nice neighborhood, but they had no friends. Their four teenagers never had visitors. One day, the oldest girl, 17 years of age, went to the police and told them she had a baby at home and her father was the baby's father. The girl said her father had been having sexual relations with her for more than 4 years and he was now doing the same with her younger sisters. The mother admitted knowing about the situation for years, but she had not reported it to the authorities for fear of losing her husband.

Extrafamilial Sexual Abuse

Of course, sexual abuse is not limited to incest. Children may be abducted and sexually abused by strangers. A perpetrator may observe a playground and identify a child who is not closely supervised. A pedophile may molest this child and hundreds of other children before he or she is apprehended. For the child, this is usually a single, isolated experience.

However, children may be repeatedly abused by trusted adults, such as teachers, counselors, family friends, and clergy. In this scenario, the pedophilic perpetrator grooms the child over a period of time. He or she gains the friendship of children through enjoyable activities and gifts, introduces sexual activities that may seem innocent and even pleasurable, and progresses to more intrusive activities. The pedophile encourages secrecy.

A solo sex ring is a form of child sexual abuse that involves one adult

perpetrator and multiple child victims, who may know about each other's sexual activities with the perpetrator. A sex ring may also involve multiple perpetrators and multiple victims.

Neurobiological Aspects of Child Maltreatment

Genetics. Recent research has begun to elucidate the interaction between genetics and environmental factors (referred to as G × E interactions). Although G × E interactions have been studied for nearly a century in biology and general medicine, the mapping of the human genome has made it possible to study complex interactions among specific genes and between genes and the environment. For example, a G × E interaction appears to play an important role in the neurobiological response to childhood maltreatment. In several studies, maltreated male children with a genotype conferring high levels of monoamine oxidase A (MAOA) expression were less likely to develop antisocial behavior as adults. In other studies examining the role of genetics in modulating the response to stress, a functional polymorphism in the promoter region of the serotonin transporter gene (known as *5HTT* or *SLC6A4*) was found to affect the likelihood of developing depression and suicidality in response to stressful experiences. Individuals with one or two copies of the "short allele" were more likely to develop depressive symptoms than were individuals with two "long alleles."

Neuroimaging. Childhood maltreatment increases the person's risk of major depression, substance abuse, anxiety disorders, and PTSD, and maltreated individuals with those disorders tend to have an earlier age of onset, greater symptom severity, more comorbidity, a greater risk of suicide, and poorer treatment response than nonmaltreated individuals with the same conditions. It is noteworthy that imaging findings associated with these disorders—such as reduced hippocampal volume and amygdala hyperreactivity—are observed more consistently in individuals with a history of child maltreatment. Trauma victims who develop PTSD have changes in brain activation that are consistent with the individual's experienced symptoms. For example, individuals with PTSD who report reliving their traumatic events in the form of flashbacks have reduced activation in the medial prefrontal cortex and anterior cingulate cortex (observed on functional magnetic resonance imaging [fMRI]), which is consistent with less regulatory control over affective arousal states. However, individuals with PTSD who experience dissociation when exposed to trauma memories have increased activation in the medial prefrontal cortex and anterior cingulate cortex, which is consistent with robust inhibition of affective arousal states.

Chronic physical abuse and repeated sexual abuse cause changes in the child's developing brain that persist into adulthood. Chronic severe abuse

can lead to a syndrome known as complex PTSD or disorder of extreme stress not otherwise specified (DESNOS). This disorder, which was evaluated but rejected by the Posttraumatic and Dissociative Disorders Sub-Work Group for DSM-5, is characterized by alterations in the following: ability to modulate emotions; attention or consciousness; somatization; self-perception; perception of the perpetrator; relations with others; and systems of meaning. The core of this disorder is a breakdown in the ability to regulate internal states. Trauma exposure changes the way in which children perceive environmental stimuli, especially their perception of threats, and how they organize and process information. Specifically, individuals with a history of trauma are more vulnerable to reacting with subcortically initiated responses that are irrelevant or even harmful in the present situation.

EVALUATION PROCESS

The evaluation of a child or an adolescent who may have been physically or sexually abused depends on its circumstances and context. Practitioners must consider whether they are conducting a forensic evaluation, which has legal implications and may ultimately be used in court, or a clinical evaluation, which is done for a therapeutic purpose. A forensic evaluation emphasizes collecting accurate and complete data to determine—as objectively as possible—what happened to the child. Was the injury an accident, was it self-inflicted, or was it a result of parental abuse? Was the child sexually abused, or was he or she indoctrinated to believe that he or she was abused? The data collected in a forensic evaluation must be preserved in an audio and video recorded format. The results of the forensic evaluation are organized into a report that is read by attorneys, a judge, and others. However, the emphasis in a therapeutic evaluation is to assess psychological strengths and weaknesses, to make a clinical diagnosis, to develop a treatment plan, and to lay the foundation for continuing psychotherapy. The clinician may also be interested in determining what happened to the child, but it is not so essential to distinguish facts from fantasies. Compared with the forensic evaluation, the psychotherapist does not need to keep such detailed records and ordinarily does not prepare a report for court.

In addition to distinguishing a forensic examination from a therapy meeting, a number of factors may affect the evaluation of a child who was abused or may have been abused: whether one is a pediatrician in an emergency department or a child psychiatrist in an office, whether a parent or another person is suspected of the abuse, the severity of the abuse and the victim's relationship to the perpetrator, whether physical signs of abuse are obvious or absent, the age and gender of the child, and the degree of anxiety, defensiveness, anger, or mental disorganization the child exhibits. Often, the examiner must be creative and persistent.

From the psychiatric perspective, the interview is usually the primary source of information and the physical examination is secondary. In practice, children who may have been neglected or sexually abused are interviewed first and are later given a physical examination and other tests. A child who has been physically abused is more likely to have a physical examination that may be followed by a psychiatric interview.

Parent Interviews

The evaluator obtains a history from the parents (separately, in most cases) and other pertinent informants, as well as from the child. The emphasis of the interview depends on the circumstances.

Suspected Physical Abuse. The examiner should consider the possibility that the parents are not telling the truth. The physician becomes a detective because the parents who bring the injured child to the emergency department may also be perpetrators of the abuse. To obtain treatment, such caregivers lie about how the injury occurred.

When the child is brought to the emergency room, a detailed and spontaneous account of the injury should be obtained promptly from parents or other caregivers before secondary details and rationalizations cloud the information provided. The interviewer should allow the caregiver to explain, to expound, to derail, or to detour the story line. An abuser or codependent parent may claim to have happened on the injured child in a coma or bleeding from some unknown trauma or to have noticed significant bruising, burns, or a crooked extremity while bathing the child. Comparing the parents' histories can provide valuable insight into how power is wielded in the family unit.

A 30-day-old girl was transferred from a rural hospital to a university medical center because of supposed near sudden infant death syndrome (SIDS). The child was unresponsive and required mechanical ventilation. A brain scan revealed bilateral subdural hematomas, subarachnoid hemorrhage, and hemorrhage in the parenchyma of the brain. An x-ray skeletal survey showed two posterior rib fractures. An ophthalmologist observed extensive retinal hemorrhages. After the child was admitted to the Pediatric Intensive Care Unit, the child abuse consultant interviewed the parents separately. The mother, 28 years of age, said that she had recently started a new job. The baby had been perfectly fine when she had left her in the care of her live-in boyfriend, the child's biological father. The father, 24 years of age, said that when he checked on the baby, he found her to be not breathing, blue, and unresponsive. He ran to report this to a neighbor and then called 911. The child abuse consultant told the father the child must have been injured in some way and asked whether he had any explanation for this injury. The father said, "I shook the baby after I found her not breathing." The consultant concluded that severe child abuse had occurred and notified child protective services and the local police department so they could initiate and coordinate their investigation.

Suspected Sexual Abuse. In cases of suspected sexual abuse, the examiner should consider the possibility that the parents are not telling the truth. However, this situation is more complex than that of suspected

physical abuse. For example, the mother may wish to avoid the discovery of father–daughter incest by blaming the child's genital injury on another child or a stranger. In another scenario, the mother may concoct an allegation of incest when the child had never been abused at all. The first version protects a father who is guilty; the second version implicates a father who is innocent.

The examiner should determine how the allegation originally arose and what subsequent statements were made; determine the emotional tone of the first disclosure (e.g., whether the disclosure arose in the context of a high level of suspicion of abuse); determine the sequence of previous examinations, the techniques used, and what was reported; try to determine whether the previous interviews were likely to have distorted the child's recollections; review, if possible, transcripts, audiotapes, and videotapes of earlier interviews; seek a history of overstimulation, prior abuse, or other traumas; and consider other stressors that could account for the child's symptoms. The examiner should also ask about exposure to other possible male and female perpetrators.

In Either Case. Whether physical or sexual abuse is involved, a pertinent psychosocial history should be collected and organized, including the following:

1. Symptoms and behavioral changes that sometimes occur in abused children.
2. Confounding variables, such as psychiatric disorder or cognitive impairment, which may need to be considered.
3. Family attitude toward discipline, sex, and modesty.
4. Developmental history from birth through periods of possible trauma to the present.
5. Family history, such as earlier abuse of the parents, substance abuse by the parents, spouse abuse, and psychiatric disorder in the parents.
6. Underlying motivation and possible psychopathology of adults involved.

Collateral Information

The evaluator should consider requesting collateral information from the following, after obtaining authorizations: protective services, school personnel, other caregivers (e.g., babysitters), other family members (e.g., siblings), the pediatrician, and police reports.

Child Interview

Several structured and semistructured interview protocols have been introduced that were designed to maximize the amount of accurate information and to minimize mistaken or false information provided by children. These approaches include the Cognitive Interview, which

encourages witnesses to search their memories in various ways, such as recalling events forward and then backward. The Step-Wise Interview is a funnel approach that starts with open-ended questions and, if necessary, moves to more specific questions. The interview protocol developed at the National Institute of Child Health and Human Development (NICHD) includes a series of phases and uses detailed interview scripts. The Cornerhouse Forensic Interview Protocol known as RATAC (Rapport, Anatomy Identification, Touch Inquiry, Abuse Scenario, and Closure) is a semistructured, nondirective interview method for alleged victims of child sexual abuse.

Although these protocols may be particularly important in a forensic context, experienced clinicians endorse flexibility and consistent good-hearted behavior by the interviewer. As when seeing any patient, the evaluator must size up the situation and use techniques that are likely to help the youngster become comfortable and communicative. One victim might need a favorite object (e.g., a teddy bear or a toy truck); another might need to have a particular person included in the interview. Some children are comfortable talking; others prefer to draw pictures. An unrelated joke, a shared cookie, or the picture on the evaluator's wall may lead to a disclosure of abuse. Important comments might be made while chatting during the break time instead of during the structured interviews.

Interview Process. The interviewer of the child who may have been abused should remember the following principles:

1. Audio and video record the interview.
2. Use a minimum number of interviews (perhaps two or three), because multiple interviews may encourage confabulation.
3. Avoid repetitive questions, either–or questions, and multiple questions, and try to avoid leading and suggestive questions.
4. Use restatement, that is, repeat the child's account back to the child (which allows the interviewer to see whether the child is consistent and ensures that the interviewer understands the child's report).
5. Conduct the examination without the parent present (if the child is young, consider having a family member in the room).
6. Use an examination technique that is appropriate to the child's age and developmental level, such as drawings and play re-enactment.
7. Determine the child's terms for body parts and sexual acts; do not educate or provide new terms.

Interview Content. The interview should not take the form of an interrogation. The interviewer should note the child's affect while discussing these topics and be tactful in helping the child manage anxiety. Young children may not be able to report all of the relevant information. The examiner should explore the following:

1. Whether the child was told to report or not to report anything.
2. Who the alleged perpetrator was.
3. What the alleged perpetrator did.
4. Where it happened.
5. When it started and when it ended.
6. Number of times the abuse occurred.
7. How the child was initially engaged, and how the abuse progressed over time.
8. How the alleged perpetrator induced the child to maintain secrecy.
9. Whether the child is aware of specific injuries or physical symptoms associated with the abuse.
10. Whether photography or videotaping took place.

The Investigative Interview Protocol. The usual clinical interview may need some modification for evaluating a child who may have been abused. The Investigative Interview Protocol, which was developed at the NICHD, consists of the following components: introduction, rapport building, training in episodic memory, transition to substantive issues, investigating the incidents, eliciting information that has not been mentioned by the child, collecting information about the disclosure, closing the interview, and concluding with a neutral topic.

INTRODUCTION. The interviewer introduces himself or herself and explains in general terms the purpose of the meeting. The interviewer explains the importance of telling the truth. The child can say, "I don't understand your question" or "I don't know." The child can correct the interviewer if the interviewer makes a mistake.

RAPPORT BUILDING. The interviewer encourages the child to talk about the child's activities and interests but avoids focusing on television, videos, and fantasy.

TRAINING IN EPISODIC MEMORY. Asking the child to describe two specific past events assesses the child's memory and models the form of the interview for the child. Using nonleading, open-ended questions, the interviewer asks the child to tell everything that happened on two specific days. The interviewer trains the child to give a free narrative account, which is the pattern through the rest of the interview.

MAKING THE TRANSITION TO SUBSTANTIVE ISSUES. The interviewer asks the child to focus on the reason for the interview by making a nonleading statement such as, "I understand that something may have happened to you. Tell me everything that happened from the beginning to the end."

INVESTIGATING THE INCIDENTS. The interviewer starts with open-ended

questions, such as "Tell me everything about what happened." If the child describes a specific incident, elicit additional information in a nonleading manner by saying, for example, "You mentioned [a particular activity]. Tell me everything about that." After posing open-ended questions, the interviewer proceeds with focused questions that relate to information already mentioned by the child. For example, "You mentioned a friend was there. What was her name? What was she doing?" This process of investigating the incidents is repeated, which are typically "the last time," "the first time," and "another time that you remember well."

ELICITING INFORMATION THAT WAS NOT MENTIONED BY THE CHILD. The interviewer may proceed with direct questions only if he or she has tried other approaches and forensically important information is missing. For example, "When you told me about the time in the kitchen, you mentioned the man took off his pants. Did something happen to your clothes? Tell me all about that." The interviewer may need to formulate a direct question by making use of outside information. For example, "I heard you told your mother that something happened at school. Tell me everything about that."

OBTAINING INFORMATION ABOUT THE DISCLOSURE. The interviewer wants to know how the child came to disclose the alleged abuse. Using nonleading questions, the interviewer may say, "Tell me what happened after [the last incident]" and "Does anybody else know what happened?"

CONCLUDING THE INTERVIEW. At the end of the interview, the interviewer says, "You have told me lots of things today, and I want to thank you for helping me." The interviewer may also say, "Is there anything else you think I should know?" and "Are there any questions you want to ask me?" The interviewer should not compliment the child with comments such as, "You were very brave to talk about those things that happened." The interviewer may conclude with a discussion of a neutral topic, such as, "What are you going to do today after you leave here?"

Extended Forensic Interview. In typical child abuse investigations, a single-session forensic interview model is widely recommended by Child Advocacy Centers across the United States. Interview protocols, which try to define the ideal single-session interview, assume the child is a willing and able participant who can provide a significant account of the alleged abuse. If the single interview results in an equivocal account of abuse, the interview is flawed, or the details need further clarification, another interview following the same protocol may be required. These single-session interviews may not suffice for children who have difficulty relating to a stranger about emotional, traumatic information. Such children may benefit from a multisession interview process.

An extended forensic interview (EFI) model was investigated by the

National Child Advocacy Center to help determine if such a protocol would be superior to the single-session model for certain children. It was investigated for use in children who either (1) did not disclose but there was other compelling evidence for abuse such as medical findings, (2) the child was unable to provide a full disclosure during the single interview, and (3) the allegations remain unresolved (founded or unfounded) after the single interview. It was determined through field trials at multiple sites that a six-session EFI was the most efficient model. The results from the field testing showed the EFI had utility and had an equal rate of court outcomes as the single interview model. Thus, for single-session forensic interviews that may have initially been unfounded, the EFI helped to resolve the ambiguity surrounding the allegation in almost two-thirds of cases. Like other interviews, the EFI should be electronically recorded.

PHYSICAL EXAMINATION

A forensic medical examination is usually performed by a pediatrician, preferably one with special expertise in child abuse. No part of the human anatomy is unrelated to the question of possible child maltreatment. Fingernails, hair, oral cavity, eardrums, skin surfaces, clothing, and genitals must be assessed, and the results must be recorded. A comprehensive examination must direct attention to all aspects of the victim's health status, including maturity and developmental assessments, dental care, immunizations, and the review of systems.

Damage to the Skin and Surface Tissues

Bruises, burns, bite marks, abrasions, and lacerations should be noted. Harsh punishment of the child may produce bruised lower back and buttocks, slapped cheeks, pinched earlobes, and choke marks about the neck. Injuries resulting from abusive toilet training may include pinch bruises on the penis or scrotum and perineal burns. Feeding-related injuries generate lacerations to the oral cavity and labial frenulum tears. Choke marks may be attributed to attempts at cardiopulmonary resuscitation (CPR). Sharp-bordered, rectilinear bruises may reflect strap marks; these markings can cover an entire body surface and can leave buckle or eyelet impressions.

The human hand leaves various impression patterns. Grab and pinch marks are diagnosed by bruise-colored fingerprints. Negative imprints of fingers are slapping injuries, whereas oval-shaped bruises may result from grabbing during violent shaking. Human bites cause distinctive, paired, crescent-shaped bruises. The age of the biter can be determined by measuring the point-to-point distance between the center of the third tooth on each side (the canines). A distance of more than 3 cm suggests permanent teeth and, hence, an adult or a child older than 8 years of age.

Dating the onset and duration of bruises is problematic. As bruises age and resolve, the skin coloration passes through several hues, including red, blue, purple, green, yellow, and brown. The timing and coloration may be affected by variables such as the depth of the bruise and the amount of bleeding into the bruise. Although characteristic bruises indicate abuse, child abuse may occur without bruising. In particular, most children with fractured bones caused by inflicted injuries do not have an associated bruise.

Immersion burns generate a saved-and-scorched pattern, typically designated as *stocking/glove* or *donut distribution* burns. The skin pressed against a tub filled with scalding water is protected from the heat and remains unburned. Skin in direct contact with the water is scalded.

A 16-month-old boy was brought to the hospital emergency room because of burns from scalding water. The child had patches of erythema on his abdomen, down the sides of his legs, and on the medial aspect of his upper arms. The diapered area had been protected from the heat. While the emergency room staff proceeded with their assessment and treatment, the child abuse consultant interviewed the parents separately. The father reported that he had been trying to heat a baby bottle by placing it in the bathtub and running hot water on it. The father returned in a few minutes and found the toddler, who apparently had climbed into the tub, in approximately 2 in of hot water. The mother confirmed she had been busy and had asked the father to heat the bottle in the bathtub. The consultant thought the pattern of the burns was consistent with the story; the pattern did not look like that seen in forcible submersion. Although the parents' accounts were unusual, they corroborated each other and were consistent with the injuries. The consultant concluded that this was not likely to be child abuse but might constitute negligence. The consultant notified child protective services.

Damage to the Head

Damage to the head includes scalp hematomas, lacerations, brain injuries, and whiplash. Head trauma is a worrisome injury. Even minor insults to the head prompt most caregivers to seek immediate medical attention. More than 95 percent of serious intracranial injuries sustained in the first year of life result from physical abuse. The cause of injury, typically, is violent shaking, to-and-fro whiplash, or slamming. Retinal tearing and hemorrhage may be caused by shaking injuries. Subdural bleeding ranks among the most dangerous inflicted injuries, often resulting in death or serious crippling sequelae. Falls are often blamed for injuries, and a fall of from 1 to 3 feet can result in linear skull fracture and epidural hematoma. Such falls rarely result in subdural hematomas or clavicle or humerus fractures.

Inflicted black eyes are more common than serious eye injuries. Victims smacked about the eyes with an open or closed hand have both eyelids swollen with massive bruising. Generally, black eyes sustained from accidents involve trauma to one eye. Although other scenarios are possible, the onset of bilateral black eyes immediately after facial trauma generally indicates intentional injury.

Traumatic alopecia and subgaleal hematomas are caused by pulling the hair. Alopecia areata (noninflammatory hair loss) is characterized by loose hairs at the periphery of the bald area and is easily distinguished from inflammation or boggy swelling of the scalp caused by violent lifting of the child by the hair. Subgaleal hematomas may result when the aponeurosis connection between the occipital and frontalis muscles is wrenched off the calvaria, permitting rapid filling of the remaining space with blood.

Damage to Internal Organs

Rarely does a child sustain cranial and abdominal traumas from a single innocent fall. Significant abdominal blunt trauma should be suspected in an atypically still child because intra-abdominal injuries are the second-most-common cause of death in battered children. Damage to internal organs includes blunt trauma to the chest; bruised abdominal organs; occult liver lacerations; duodenal hematoma; pancreatic contusion or laceration; superior mesenteric artery laceration; and spleen, kidney, and bladder rupture.

Skeletal Damage

Skeletal damage is usually documented with x-rays taken immediately and also, if necessary for better definition, repeated approximately 2 weeks after the injury. The most distinctive radiological features of infant abuse are metaphyseal lesions, called *corner fractures* or *bucket-handle fractures*, depending on the perspective of the x-ray. In addition to metaphyseal lesions, the following are highly specific for child abuse because they are unlikely to occur accidentally: posterior rib fractures, scapular fractures, spinous process fractures, and sternal fractures. The following are considered moderately specific for child abuse: multiple fractures, especially when they are bilateral; fractures of different ages; epiphyseal separations; vertebral body fractures and subluxations; fractures of the fingers; and complex skull fractures. Radiologists are astute at locating evidence of trauma and can provide insight into causality and patterns of old, healing, or new fractures. A child may not report a fracture if more injuries are likely to result after disclosure.

Intoxication and Poisoning

Psychotropic and analgesic medication, alcohol, and illicit drugs have all been used by adult abusers to render a child or adolescent more susceptible to assault and less capable of offering accurate witness account. There have also been reports of administration of the following: salt; ipecac; laxatives; acetaminophen and aspirin, which are readily available; prescribed medications, including insulin; and water to the point of intoxication.

Physical Findings of Sexual Abuse

Labial and perineal tissues that are susceptible to friction-induced traumas manifest abrasions, contusions, lacerations, petechiae, and edema. Examination may reveal bloodied tissue, genitorectal tears, and infected secretions. The genital examination should be performed by practitioners with training and experience in interpreting and documenting findings.

Scarring after sexual abuse may be minimal or not evident if vaginal penetration has been gentle, consensual, or gradual. Vaginal and rectal tissues rapidly heal after injury, and abrasions can be repaired within several days. Thus, the lack of physical findings may indicate a return to normalcy after rapid healing of minor lesions or trauma perpetrated in a manner that left no material evidence. In other words, a normal genital examination does not mean the child was not sexually abused. Most victims of child sexual abuse have a normal or nonspecific genital examination.

Major insults to the penile head (scaldings, amputation) produce noticeable scars. Penile strangulation with a ligature may be inflicted as a response to enuresis. The more common traumas (e.g., hematomas, abrasions, and lacerations) are superficial injuries and heal rapidly without scarring.

Anal findings include dilation with loss of anal-rectal tone, flattened rectal rugae, lacerations, and wedge-shaped and linear scars. Acute trauma may result in swelling and spasm secondary to submucosal hemorrhage. Most rectal injuries, massive or minor, heal quickly and nearly completely, leaving little evidence of abuse. Minor but persistent irritants, such as pinworms, lichen sclerosis, atopic dermatitis, and rare hematological disorders, can be longer-standing rectal irritants and may confound examination outcomes.

SPECIAL TESTS AND LABORATORY EXAMINATION

An unconscious child or a child with a suspected head injury should have a computed tomography (CT) scan of the head. If the x-ray skeletal survey is negative but there is a high suspicion of abuse, the clinician may consider radionucleotide bone scan or bone scintigraphy. This technique is more sensitive than x-rays in detecting evidence of soft tissue trauma, as well as bone trauma, in child abuse.

When sexual abuse is suspected, the victim requires evaluation for sexually transmitted diseases (STDs). If the examination occurs within 72 hours of the alleged abuse, the clinician should consider cultures of the genital, anal, and oral mucosa. The following STDs are essentially diagnostic of child sexual abuse, unless they were acquired perinatally: gonorrhea, syphilis, and *Chlamydia trachomatis*. Positive serology for the human immunodeficiency virus (HIV) is diagnostic of child sexual abuse, unless it was acquired perinatally or by blood transfusion. The following

STDs raise suspicions of child sexual abuse and should ordinarily be reported to protective services: herpes simplex virus in a genital location, unless there is a clear history of autoinoculation; condylomata acuminata or anogenital warts, except that infants and young children may acquire anal and genital warts perinatally; and *Trichomonas vaginalis*.

Of course, pregnancy is diagnostic of child sexual abuse, as is the presence of semen or sperm acid phosphatase.

DIFFERENTIAL DIAGNOSIS

Physical Findings

Pediatricians who examine children who may have been abused must consider a wide range of congenital, infectious, toxic, and traumatic conditions in their differential diagnoses. The following examples are not meant to be exhaustive.

Conditions Manifested by Skin Lesions. A mongolian spot—a birthmark—may look like a bruise. Lichen sclerosis may look like sexual abuse. Chicken pox and impetigo may look like cigarette burns.

Conditions Manifested by Bone Abnormalities. Osteogenesis imperfecta, congenital syphilis, copper deficiency, and rickets must be considered. CPR sometimes causes rib fractures. However, the physical mechanics of CPR causes lateral rib fractures, whereas child abuse typically causes posterior rib fractures.

Conditions Manifested by Bruises. Vitamin K deficiency from liver disease, salicylate toxicity, bleeding disorders, and erythema multiforme must be considered.

Conditions Manifested by Eye Hemorrhages. Valsalva's maneuver causing increased intracranial pressure and hypertension must be considered.

Self-Inflicted Injuries. Scratches from itchy irritations, factitious illness, head banging, and Lesch–Nyhan syndrome must be considered.

Psychiatric Considerations

The psychiatric evaluation of youngsters who may have been abused may involve assessing the patient's credibility. Although children generally tell the truth when they talk about abuse, sometimes children make false denials (saying that they were not abused, when actually they were) or false allegations (saying they were abused, when actually they were not).

Possible Explanations of Denials of Abuse. A false denial or retraction may occur for several reasons. The child may have been pressured by the perpetrator or family members to recant the allegation. The pressure may consist of bribery, mockery, or threats of injury. The child may be protecting a parent or other family member, even without external coercion (i.e., the child takes on this responsibility through role reversal). The child might be frightened or distressed by the investigation and may decide to withdraw participation. For instance, an interviewer could induce a false denial by asking overly challenging questions. A child may be inhibited by shame or guilt; the child may mistakenly assume responsibility for what happened. Finally, the child may have accommodated to the abuse, consciously or unconsciously, instead of objecting to it.

Possible Explanations of Allegations of Abuse. A false allegation of abuse may occur for several reasons.

ALLEGATIONS ORIGINATING WITH ADULTS. Sometimes a false allegation arises in the mind of a parent or another adult and is imposed on the child.

Parental Misinterpretation and Suggestion. The parent may have misinterpreted an innocent remark, a neutral piece of behavior, or a benign physical condition as evidence of abuse and may have induced the child to endorse this interpretation. This happens in child custody disputes, as well as in other settings.

Parental Delusion. The parent and child may share a *folie à deux*, or the child may simply give in and agree with a delusional parent.

Parental Indoctrination. A parent may have fabricated the story and may have induced the child to collude in presenting it to the authorities.

Interviewer Suggestion. Previous interviewers may have asked leading or suggestive questions. An interviewer who believes abuse occurred may unwittingly shape a child's responses until the child validates the interviewer's assumptions.

Group Contagion. In epidemic hysteria, people modify what they have heard to meet their emotional needs. Thus, rumors may become more convincing as they are retold.

UNCONSCIOUS OR NONPURPOSEFUL MENTAL MECHANISMS IN THE CHILD

Fantasy. A younger child may confuse fantasy with reality.

Delusion. Although rare, older children and adolescents may experience delusions about sexual activities in the context of a psychotic illness.

Misinterpretation. Sometimes children misunderstand what happened and

later report it inaccurately.

Miscommunication. The child may misunderstand an adult's question, and the adult may later misinterpret or take the child's statement out of context.

Confabulation. Children fill gaps in their memory with whatever information makes sense to them and others at the time.

CONSCIOUS OR PURPOSEFUL MENTAL MECHANISMS IN THE CHILD

Pathological Lying. Also called pseudologia phantastica, this refers to children who tell stories with such zeal that they may become convinced of their truth.

Innocent Lying. Children may make false statements because that seems to be the best way to handle the situation. Developmentally, this is more likely to happen with younger children.

Deliberate Lying. Children may choose to avoid or to distort the truth for some personal advantage. This happens more with older children and adolescents.

PERPETRATOR SUBSTITUTION. The child may have been sexually abused and may exhibit symptoms consistent with abuse but may identify the wrong person as the perpetrator, resulting in a false allegation. The child may do this to protect the actual offender, or the child may displace the memories and accompanying affects onto another individual.

Possible Explanations of Allegations of Ritual Abuse. Other phenomena may be mistaken for cult-based ritual abuse, including pseudoritualistic abuse, psychopathological repetitive abuse, sexual abuse by pedophilic sex rings, false memories of abuse, and hoaxes.

The parents of a 9-year-old boy were divorced, and he spent the summers visiting his father. When he returned from visitation, the mother noticed injuries: A healing cut on the boy's right ear and a collection of superficial, vertical scratches from his umbilicus to his pubis. She had the boy disrobe and took digital images of his abdomen. Several days later, the mother took the boy and the images to the boy's regular psychotherapy appointment. She told the therapist that she was concerned the boy had been physically and sexually abused. The therapist interviewed the boy individually, who explained in a matter-of-fact manner his ear was cut accidentally when he had his hair cut and his abdomen was scratched when he went body surfing at a beach. The therapist concluded the boy had not been abused and reassured the mother. The therapist did not notify child protective services.

SPECIAL ISSUES

Fetal Abuse

Fetal abuse can be direct or indirect. A mother or a father may directly

attack their fetus through the abdominal wall or via the vagina. The risk is greatest during late pregnancy, when one of the parents becomes disenchanted with the prospect of a new baby. Indirect fetal abuse by failing to protect the fetus from alcohol, nicotine, or drugs is more common. Fetal alcohol syndrome is an important cause of mental retardation in the United States.

Neonatal abstinence syndrome (NAS) is another example of potential maltreatment to a fetus. Newborns who have been exposed to maternal drug use can undergo withdrawal after their birth. This can be due to illicit substance use by the mother or iatrogenic addiction (e.g., maternal treatment of a pain condition with opioid medication). The newborn's symptoms are reflected by the type of substance exposure. Symptoms of NAS encompass high-pitched cry, jitteriness, tremors, convulsions, sweating, fever, mottling, excessive sucking/rooting, poor feeding, vomiting, and diarrhea. A toxicology workup can be done in addition to thorough history taking, although the mother may be reluctant to disclose substance use. It is estimated that 3 percent of the 4.1 million women of child bearing age who abuse drugs will continue to use drugs during their pregnancy. From 2000 to 2009, the rate of infants born with NAS rose from 1.2 to 3.4 per 1,000 hospital births per year. Maternal substance use has both medical and developmental concerns for the newborn as well as legal, health, and economic impact for the mother.

Factitious Disorder Imposed on Another

In *factitious disorder imposed on another* (previously called factitious disorder by proxy and Munchausen syndrome by proxy), a parent, usually a mother, provides a history that fascinates and beguiles family physicians and expert consultants. The mother gives enough data to suggest a serious but poorly understood condition that requires extensive, painful testing. The child is often drawn into the improbable tale out of fear of contradicting the history given by the mother. Professionals become enmeshed in the apparent validity of the history, but no one can verify or witness the extraordinary symptoms that support the pathology of the child's alleged illness. Parents involved in the deception often have narcissistic, histrionic, or borderline personality disorder and thrive on the center-stage position they assume. In families in which factitious disorder by proxy is found, more than one child may be involved consecutively or simultaneously. In extreme cases, children can be killed, seriously injured, terrorized, and experience severe adverse effects from medications or other treatments.

A 12-year-old boy was referred for psychiatric evaluation because of oppositional behavior. The psychiatrist interviewed the youngster and his mother and learned they had an enmeshed relationship and their affect toward each other ranged from loving to embittered. An older

brother had finished school and worked full time but still lived at home. The patient's medical history was unusual, in that the mother claimed the boy had a long history of fainting spells that defied diagnosis. The mother, but no medical professional, had witnessed these spells. The mother was a secretary in a medical practice consisting of neurologists and neurosurgeons, and she had induced her employers to repeatedly work up the child with physical examinations, laboratory testing, and electroencephalograms (EEGs). In fact, one of the neurosurgeons did burr holes to place EEG electrodes directly on the surface of the child's brain. No neurological diagnosis was made, and factitious disorder by proxy was suspected. The child's neurologist said that no further testing would be done and encouraged psychiatric treatment. The psychiatrist helped the patient and his mother achieve age-appropriate psychological separation, and there were no more fainting spells.

Sex Trafficking

Sex trafficking has been likened to slavery by the Federal Bureau of Investigation. Children become involved in the sex trade for different reasons. Some youth are abandoned by their families and turn to prostitution as a means to survive on the street. Others are recruited into the sex trade through forced abduction, pressure from parents, or through connections the parents have with sex traffickers. These children are often taken far from their homes and become unable to develop relationships with peers or adults other than their abusers. It is estimated that almost 300,000 American youth are at risk of becoming victims of the commercial sex trade. The average age in which children enter the trade is between 12 and 14 years for girls and 11 and 13 years for boys.

The use of violence, drugs, coercion, and other methods of manipulation are commonplace among the traffickers. If a youth in the sex trade becomes addicted to drugs, the trafficker can use the addiction as a means of control. The National Center for Missing and Exploited Children created the Child Sex Trafficking Team in 2001 due to the need for specialized services in this area. Training of police officers, detectives, and emergency room staff can help to identify youth involved in trafficking.

Cultural Variations

Parenting practices in US society vary widely. They are partly determined by the cultural heritage of the family, and the cultural context of the alleged act should be considered in evaluating suspected neglect and abuse, including assessing whether the discipline of a child is abusive. Some unusual bruises result from cultural healing practices. Coining or cupping—techniques used in many European and Asian countries—are used to draw toxins from the skin surface and to remove feverish mucus from the system. In coining, a coin is heated in hot oil and massaged against the skin. In cupping, a heated glass cup is used to create suction against the skin. Skin scraping and moxibustion may also mimic physical abuse.

Female Genital Mutilation

Commonly called *female circumcision*, female genital mutilation is performed on an estimated 2 million girls worldwide each year. It is practiced across diverse socioeconomic classes and different ethnic, cultural, and religious groups. Commonly, girls are circumcised between 4 and 10 years of age, but the procedure may be performed on infants, may be postponed until just before marriage, or may be done after the birth of the first child. In some cultures, it is considered part of a ceremonial induction into adult society. Female circumcision, in the mildest form, consists of clitoridectomy—the anatomical equivalent of penile amputation. In its most severe form, total infibulation involves removal of the clitoris and labia minora plus incision of the labia majora to create raw surfaces that are then stitched together. This practice has been widely criticized, including opposition by the World Health Organization and other major health care groups. In the United States, female circumcision is generally considered child abuse. Efforts have been made to educate immigrant communities on the health risks and legal liabilities of the practice. Compromise may be possible by finding a way to satisfy the cultural requirements without using mutilation, such as adopting a nonmutilative ritual incision that results in only small scars on the labia.

Male genital mutilation (circumcision) is not usually considered child abuse, but it does have its critics. Although this simple procedure is one of the most common operations performed worldwide, complications may result such as hemorrhage and infection. Advocates of male circumcision cite research that it is somewhat protective against HIV infection, other STDs, and penile cancer.

Religious Practices

Certain religious practices may be construed as forms of abuse if they are likely to harm children. For example, adherents of Christian Science believe illness should be treated with prayer instead of medical care. As a result, children may sustain chronic illness, disability, or death. This is an area of flux: Some state laws and some courts have allowed parental discretion to withhold conventional medical treatment for legitimate religious reasons, whereas others have not.

Corporal Punishment

Within families, corporal punishment occurs often enough in the United States that it cannot easily be labeled deviant. Physical punishment by parents is permitted in 49 states by either statute or court ruling. Most child and adolescent psychiatrists discourage physical discipline but realize many capable and apparently nurturing parents spank their children. If physical discipline is used as a last resort, parental anger and frustration may exceed acceptable limits. In its most extreme form, a spanking turns

into a beating. After 10 to 20 minutes of such abuse, a child will bleed into the muscles and other tissues of the back, buttocks, and legs and may lose enough blood to die from hypovolemic shock.

Although professionals may disagree about the propriety of parental spanking of children, there is a consensus against the paddling of children by teachers. Professional organizations (e.g., the American Medical Association, the American Academy of Pediatrics, the Society for Adolescent Medicine, and the American Academy of Child and Adolescent Psychiatry) have gone on record favoring banning of corporal punishment as a disciplinary measure in schools. Although almost every developed country has banned corporal punishment in schools, it is still allowed in more than 20 states, especially in the southeastern United States. New Jersey and Iowa are the only two states that have banned physical punishment in both public and private schools. Within the welfare system, 49 states have prohibited physical punishment in foster care settings.

Teamwork

It is draining to establish rapport, to delve into hidden agendas, to uncover covertly presented data, to address multiple human needs, and to meet the legal obligations associated with child abuse cases. Clinicians are thus encouraged to participate in a multidisciplinary team. Team participation provides consultation and professional support essential for continued health, hope, and positive social interactions within the grim realities of case discovery.

In cases of suspected child abuse and neglect, the physician should make an assessment based on a history, a physical examination, a skeletal survey (when indicated), and photographs; should request appropriate surgical, medical, and psychiatric consultations; should diagnose the suspected maltreatment; and perhaps should secure the child's safety by admitting the child to a hospital. Usually, other members of the multidisciplinary team report the case to the appropriate child protection agency, provide skilled nursing, maintain security, arrange out-of-home placement, arrange a program of care for the child and the parents, and arrange for social service follow-up. These activities, taken together, should be a coordinated team effort. One feature of the team activity is regular meetings of the child abuse committee.

REPORTING LAWS AND PRACTICES

In every state, the District of Columbia, and all areas controlled by the federal government (e.g., military bases), physicians are required to report child abuse to the police or the appropriate child protection agency. The typical statute says a physician who has "reason to believe" or "reason to suspect" child abuse is required to report it. In most states, other

designated individuals, such as nurses, psychologists, school officials, police officers, district attorneys, and providers of child day care and foster care, are also required to report child abuse.

Underreporting

A significant factor in reporting errors is the failure of clinicians and other mandated reporters to consider child maltreatment in the differential diagnosis of child injuries. Frequently, mandated reporters fail to pursue the diagnosis of child maltreatment because of diverse cultural and socioeconomic factors. Lack of experience in dealing with the issues and the judicial quagmire that result when reporting many child maltreatment cases are also an important deterrent. Underreporting of abuse also occurs when well-meaning but misguided therapists attempt to monitor and to treat their abusive clients without involving the proper authorities. Some therapists think they are above the law; others simply have no idea what the law says about reporting child abuse. Perhaps the most important barrier to seeing and reporting abuse is denial. Reporters, both professionals and laypersons, are repulsed by the reality of child maltreatment and cannot accept the thought that abuse exists or is perpetrated by individuals like themselves, persons who in addition might not regret their harmful actions. The most obvious negative consequence of underreporting is the abused child and siblings are left in dangerous circumstances and are vulnerable to further abuse.

Overreporting

Although physicians are required to report abuse when they have reason to believe or to suspect (depending on the jurisdiction) it occurred, they are not required to report other people's beliefs or suspicions. Some clinicians have the mistaken idea that, if a mother expresses a concern that her child may have been abused, the suspected abuse must be reported, even when the clinician has investigated and has no suspicion that abuse occurred (e.g., an overly anxious mother who imagines her child was sexually abused at day care, even though an objective examiner found no indications abuse had occurred). The negative consequences of overreporting include increasing the load on an already overburdened child protective system and an unnecessary invasion into the lives of families, who are subjected to investigation, suspicion, and criticism.

TREATMENT

Child

The first part of the treatment of child abuse and neglect is ensuring the child's safety and well-being. The child may need to be removed from an abusive or neglectful family to ensure protection but emotionally may feel

even more vulnerable in an unfamiliar setting. Because of the high risk of psychiatric symptoms in abused and neglected children, a psychiatric evaluation is in order. Along with providing specific treatments for any mental disorders present, the therapist may have to deal with the immediate situation and the long-term implications of the abuse or neglect. Psychotherapeutic issues to be addressed with an abused child include dealing with the child's fears, anxieties, and self-esteem; building a trusting relationship with the therapist in which the child is not exploited or betrayed; and gaining a perspective, over time, of the factors that contributed to the child's victimization. These issues may be addressed in individual or group psychotherapy or both.

The intensity and duration of treatment depend on the severity of the symptoms. If an abused child has no posttrauma symptoms, it probably is adequate to provide a few counseling sessions that include psychoeducation and anticipatory strategies for recognizing problems in the future. For a child who has been minimally assaulted (e.g., a relatively innocuous single episode of fondling by a stranger), it may be adaptive for the child to forget the incident and to move on.

Abused children who have mild to moderate psychological symptoms have been treated with a form of cognitive-behavioral treatment that has four components: psychoeducation, anxiety management, exposure, and cognitive therapy. *Psychoeducation* includes teaching the child and the caregivers about the nature of maltreatment, offenders, and strategies to avoid future abuse. *Anxiety management* includes teaching children how to use relaxation and other coping strategies to reduce fearful responses to memories and reminders of abuse. *Exposure* refers to talking, drawing, or writing about the abuse experiences. It is usually helpful to have the child share the painful, ugly details of the abusive experiences with the therapist. This may take time and may require therapeutic creativity and flexibility. Although some children may be able to relate the bad things that happened through ordinary conversation, other children may need a variety of play, artistic, or other projective techniques. Some children avoid the painful memories through conscious (suppression) and unconscious (repression) mechanisms. *Cognitive therapy* techniques are used to help the child replace cognitive distortions about the event or negative feelings about himself or herself.

Abused children and adolescents with very severe symptoms (such as complex PTSD or DESNOS) require extensive treatment. These victims of abuse have been treated with a multimodal approach, including specific psychotherapies such as dialectical behavioral treatment (DBT), psychotropic medication, and perhaps hospitalization or residential treatment.

Parents

Developing an intervention plan for an abused or neglected child requires assessing the parents' behavior and functioning in several respects. The psychopathology and the prognosis for achieving adequate parenting skills of the abusive parent and the one who allowed the abuse to occur should be assessed. The evaluator should determine whether the parent's dysfunction is confined to this child or involves other children and whether the dysfunction is short term or long term (reflecting a lifelong pattern). The evaluator should assess the parent's willingness to participate in the intervention plan and the availability of personnel and physical resources to implement the various intervention strategies. One must consider the risk of additional physical or sexual abuse if the child remains in the home.

On the basis of the information obtained, several options can be selected to improve the parent's functioning: (1) eliminating or diminishing the social or environmental stresses; (2) lessening the adverse psychological effects of the social factors on the parents; (3) reducing the demands on the parent to a level that is within their capacity through day care utilization or provision of a housekeeper or babysitter; (4) providing emotional support, encouragement, sympathy, stimulation, instruction in parental care, and training in planning for, assessing, and meeting the needs of the child (supportive casework); and (5) resolving or diminishing the parent's inner psychic conflicts (psychotherapy). Some clinics provide group counseling for nonoffending parents.

Incestuous Behavior

The first step in the treatment of incestuous behavior is its disclosure, which means overcoming the denial and collusion of the family members. Usually, the perpetrator leaves the home, or the child is removed by protective services, so incest is not likely to recur. Often, the family does not reunite because the perpetrator is imprisoned or the parents' divorce. In those circumstances in which it is possible for the family to reunite, family therapy is useful to re-establish the group as a functioning unit and to develop healthier role definitions for each member. When the participants experience severe psychopathology, treatment must be directed toward the underlying illness. While the participants are learning to develop internal restraints and appropriate ways to gratify their needs, the external control provided by therapy helps to prevent further incestuous behavior. At times, legal agencies are involved to help enforce external controls.

PREVENTION

Primary Prevention

Measures that reduce the occurrence of child abuse—*primary prevention*—can be classified on the basis of their audience or intended recipient.

Universal interventions targeted to the general public include such strategies as a poster campaign admonishing parents not to shake their babies, warnings about consuming alcohol and drugs during pregnancy, and raising public awareness through the media, civic associations, church groups, and parent–teacher organizations. To some extent, children can be taught to avoid or to resist being abused. However, ultimately, the adults in a community must create safety and security for their children. *Selective interventions are targeted to populations at high risk of child abuse*. In general, for example, there should be less child abuse in a community with fewer single parents (use of birth control and greater involvement of fathers in families) and greater financial security (local jobs and social support systems). Governmental regulations may be helpful, such as the licensing of day care centers and regulations concerning the operation of residential programs for children. Finally, one can sometimes *intervene with specific persons or families that are at high risk*. Counseling and parenting classes may deter parents from being abusive, although this has not been substantiated. Treatment of parents who abuse substances and services for families with a history of spouse abuse may be helpful. Individuals can help friends and relatives who appear stressed and who are struggling with their parenting responsibilities.

Secondary Prevention

Early detection and prompt intervention when child abuse occurs comprise secondary prevention. Education of the medical profession, members of allied health fields, and all who come in contact with children aids early detection. Emergency department personnel, pediatricians, teachers, and counselors have many opportunities to identify abused children. It is hard to know how best to use the limited resources available to child protection agencies. Some critics think too much effort is put into assessing the millions of referred cases, with relatively little attention given to interventions. For instance, an agency might put many hours into collecting information, might arrive at a decision about whether a case is founded, and might engage in expensive legal measures but might offer little or no aid to help a marginal family function better. In general, child abuse and neglect prevention and treatment programs should try (1) to prevent the separation of parents and children if possible, (2) to prevent the placement of children in institutions, (3) to help the parents attain self-care status, and (4) to help the family become self-sufficient. As a last resort, to prevent further abuse and neglect, children may have to be removed from families who are unwilling or unable to profit from the treatment program.

FUTURE DIRECTIONS

When “The Battered Child Syndrome” was published in 1962, the concept

of child maltreatment had a rather narrow scope, that is, physical child abuse. During the half century since that landmark event, the understanding of child maltreatment has grown to include medical, educational, and other forms of neglect; child sexual abuse and the exploitation of children in child pornography; verbal or psychological abuse; observing family violence; the neglect of children in orphanages; and the abuse of adolescents in institutional settings such as residential treatment centers. In the future, the scope of child maltreatment will continue to enlarge and there will be descriptive research regarding less common forms of child abuse, such as factitious disorder by proxy; parental alienation, in which a child of divorced parents loses his or her relationship with one of the parents; and abusive corporal punishment in schools. The phenomenon and consequences of psychological abuse will be studied more thoroughly.

The sequelae of child maltreatment have been studied through retrospective and limited prospective research. The study of psychological abuse will be of particular interest to mental health professionals, especially since both *child psychological abuse* and *spouse or partner psychological abuse* were added to DSM-5. The definitions of these mental conditions refer to "nonaccidental verbal or symbolic acts . . . that result, or have reasonable potential to result, in significant psychological harm . . . ." The research project that addressed adverse childhood experiences (ACEs) took child psychological abuse into consideration, in that two of the ACEs directly reflected child psychological abuse: verbal abuse by a parent or other adult in the household and feeling threatened with physical injury. The ACE researchers found that an emotionally abusive family environment accentuated negative outcomes later in life. Authors from many countries have noted that causing parental alienation in a child—inducing the child to fear or hate a parent, resulting in psychological parentectomy—is a particularly pernicious contemporary form of child psychological abuse.

In the future, there will be systematic studies from an international perspective of child maltreatment, providing information on, for example, child soldiers in Africa, child suicide bombers in the Middle East, child labor in Asia, the trafficking of children and adolescents for prostitution, and the deaths of large numbers of children due to lack of food and medical care.

A major focus of research in the next few years will pertain to resilience. This research will address the question of why some maltreated children manifest very serious psychological and behavioral symptoms, whereas other abused children are free of symptoms both in the short and the long term. At least part of the answer to this question will involve genetic predispositions, specific life experiences, and the interaction between genetic and environmental factors.

In the more distant future, there will be greater understanding of the chain of events by which trauma damages young victims, starting with specific genetic and other congenital vulnerabilities and predispositions that interact with maltreatment and other forms of adverse childhood experiences. These predispositions and experiences lead to changes in cognitive and affective pathways in the brain, which will be observed through neuroimaging. Alterations in the brain will be shown to cause changes in behavioral and emotional functioning, which result in dysfunctional relationships and additional symptoms and disorders. Although society's goal is the primary prevention of child maltreatment, that objective will not be fully achieved. The research outlined here will inform the secondary prevention of the consequences of child abuse, as well as the biopsychosocial treatment of abused children.

REFERENCES

*American Academy of Child and Adolescent Psychiatry. Practice parameters for the forensic evaluation of children and adolescents who may have been physically or sexually abused. *J Am Acad Child Adolesc Psychiatry*. 1997;36:423.

American Professional Society on the Abuse of Children (APSAC). *Practice Guidelines: Forensic Interviewing in Cases of Suspected Child Abuse*. Chicago, IL: American Professional Society on the Abuse of Children; 2012.

Anderson J, Ellefson J, Lashley J, et al. The CornerHouse forensic interview protocol: RATAC. *Journal of Practical and Clinical Law*. 2010;12:193.

Bernet W. False statements and the differential diagnosis of abuse allegations. *J Am Acad Child Adolesc Psychiatry*. 1993;32:903.

Byrd AL, Manuck SB. MAOA, childhood maltreatment, and antisocial behavior: meta-analysis of a gene-environment interaction. *Biol Psychiatry*. 2014;75(1):9.

Ceci SJ, Bruck M. *Jeopardy in the Courtroom: A Scientific Analysis of Children's Testimony*. Washington, DC: American Psychological Association; 1995.

Corwin DL, Keeshin BR. Estimating present and future damages following child maltreatment. *Child Adolesc Psychiatr Clin N Am*. 2011;20(3):505.

Deblinger E, Mannarino AP, Cohen JA, Steer RA. A follow-up study of a multi-site, randomized, controlled trial for children with sexual abuse–related PTSD symptoms. *J Am Acad Child Adolesc Psychiatry*. 2006;45:1474.

Edwards VJ, Holden GW, Felitti VJ, Anda RF. Relationship between multiple forms of childhood maltreatment and adult mental health in community respondents: results from the adverse childhood experiences study. *Am J. Psychiatry*. 2003;160:1453.

Finkelhor D. Prevention of sexual abuse through educational programs directed toward children. *Pediatrics*. 2007;120:640.

Flaherty EG, Sege RD, Griffith J, Price LL, Wasserman R. From suspicion of physical child abuse to reporting: primary care clinician decision-making. *Pediatrics*. 2008;122:611.

Friedrich WN, Fisher JL, Dittner CA, Acton R, Berliner L. Child sexual behavior inventory: normative, psychiatric, and sexual abuse comparisons. *Child Maltreat*. 2001;6:37.

Gershoff ET. *Report on Physical Punishment in the United States: What Research Tells Us About Its Effects on Children*. Columbus, OH: Center for Effective Discipline; 2008.

*Helfer ME, Kempe RS, Krugman RD, (eds). *The Battered Child*. 5th ed. Rev. Chicago: University of Chicago Press; 1999.

*Jenny C, ed. *Child Abuse and Neglect: Diagnosis, Treatment and Evidence*. St. Louis, MO: Saunders; 2010.

Kellogg N, American Academy of Pediatrics, Committee on Child Abuse and Neglect. The evaluation of sexual abuse in children: subject review. *Pediatrics*. 2005;116:506.

Kellogg N, American Academy of Pediatrics, Committee on Child Abuse and Neglect. Evaluation of suspected child physical abuse. *Pediatrics*. 2007;119:1232.

Kempe CH, Silverman FN, Steele BF, Droegemueller W, Silver HK. The battered-child syndrome. *JAMA*. 1962;181:17.

*Lamb ME, Orbach Y, Hershkowitz I, Esplin PW, Horowitz D. A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: a review of research using the NICHD Investigative Interview Protocol. *Child Abuse Negl*. 2007;31:1201.

*Myers JEB. *The APSAC Handbook on Child Maltreatment*. 3rd ed. Thousand Oaks, CA: Sage; 2010.

Nemeroff CB, Seligman F. The pervasive and persistent neurobiological and clinical aftermath of child abuse and neglect. *J Clin Psychiatry*. 2013;74:10.

Poole DA, Lamb ME. *Investigative Interviews of Children: A Guide for Helping Professionals*. Washington, DC: American Psychological Association; 1998.

Sanders MJ, Bursch B. Forensic assessment of illness falsification, Munchausen by proxy, and factitious disorder. *Child Maltreat*. 2002;7:112.

Spataro J, Mullen PE, Burgess PM, Wells DL, Moss SA. Impact of child sexual abuse on mental health: prospective study in males and females. *Br J Psychiatry*. 2004;184:416.

Teicher MH, Samson JA. Childhood maltreatment and psychopathology: a case for ecophenotypic variants as clinically and neurobiologically distinct subtypes. *Am J Psychiatry*. 2013;170(10):1114.

Trickett PK, Mennen FE, Kim K, Sang J. Emotional abuse in a sample of multiply maltreated, urban young adolescents: issues of definition and identification. *Child Abuse Negl*. 2009;33:27.

U.S. Department of Health & Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. 2016. *Child Maltreatment 2014*. Available at http://www.acf.hhs.gov/programs/cb/research-data-technology/statistics-research/child-maltreatment

van der Kolk BA, Roth S, Pelcovitz D, Sunday S, Spinazzola J. Disorders of extreme stress: the empirical foundation of a complex adaptation to trauma. *J Trauma Stress*. 2005;18:389.

▲ 55.3 Children's Reaction to Illness and Hospitalization

SUSAN BECKWITT TURKEL, M.D., JULIENNE JACOBSON, M.D., AND MARYLAND PAO, M.D.

HISTORY

Hospitalization of young children was a rare event before the mid-19th century, as most children were cared for at home, and hospitals were considered unsafe for them. When young children were admitted to early general hospitals their mothers typically remained with them to assist in providing their care. This was routine and met with little opposition. Some hospitals would only admit children on the condition their mothers were present. From the 1850s through the 1880s, hospitals encouraged regular visits from parents and siblings. These visits were viewed as important for the well-being of the sick children. Outbreaks of infectious diseases might temporarily restrict visits, but they were reinstated as soon as possible. Any restrictions on visits would be lifted for seriously ill or dying children. The goal was to create, at least in part, a "homeliness" where children could feel safe and comfortable.

The transformation of the hospital from poorhouse to headquarters for scientific medicine started in the period before the First World War, and had profound implications on medical care, research, and education. The exclusion of parents from the hospital began in the late 19th century, and hospitals gradually restricted parental access and removed toys, pictures, and pets from the wards. These changes reflected changing ideas about the spread of infectious diseases, the growing professionalism of nursing, the development of the specialty of pediatrics, and funding changes for health care. Visitors, including parents and siblings, were restricted or excluded from the children's wards because of the belief that the emotional upset the visits created would be detrimental to the child's recovery. Fear of infection led to banning personal contact between parents and their children, and glass panels were provided so parents could see their children on the ward. The practice of enabling mothers to remain with their hospitalized children never completely disappeared, but mainstream practitioners during this period were unconvinced of the importance of having the mother in the hospital with her sick child.

Nursing was highly regimented during this period, and children and parents often perceived the staff as uncaring. Although pediatric doctors and nurses cared about their patients' emotional well-being, they were taught to be disciplined, professional and scientific, and not to show affection or emotion. Children were traumatized by their illness, by separation from their parents and families, and also by the lack of warmth in the care they received.

Little changed until after the Second World War. At that time the prevailing practices were in conflict with emerging psychoanalytic and developmental theories of the social and emotional growth of children, and new thinking about the child's need for parental interaction and emotional care. John Bowlby, Deputy Director of the Tavistock Clinic in England, established a separation research unit in 1948, which led to his formulation of attachment theory. Bowlby warned that separation or denial of a warm, intimate and continuous relationship with the mother or mother substitute, or maternal deprivation, could have far-reaching effects on the child's character and mental development. Other psychoanalytic studies used hospitalization as an example of the first early separation from home. Psychoanalysts were aware of the lack of integration between the practice of child guidance, psychoanalysis, and pediatrics. The profound emotional reactions of children to physical illness were noted, as was the likelihood that mothers might be unable to make objective observations when they were emotionally upset by the stresses of their child's condition. The need to prevent the negative psychological impact of hospitalization on children was soon accepted by most psychiatrists and some pediatricians, but this did not yet change hospital practice.

James Robertson, a student working with John Bowlby, translated Bowlby's theories into film, starting with "A 2-year-old Goes to Hospital" in 1953. This film emphasized the detrimental effects of an admission that separated a young child from his family, friends, and mother's care. It demonstrated the phases of protest, despair, and detachment (or denial) through which small children pass when isolated from their mothers for a length of time. Robertson illustrated how children who had been left in the hospital exhibited greater behavioral problems when they returned home. His film was widely shown to both professional and lay groups and was instrumental in accelerating change in hospital practice.

Following the efforts of Robertson and others, public outcry led to the formation of a British parliamentary committee, chaired by Sir Harry Platt, which investigated the hospital care of children. The committee's report, "The Welfare of Sick Children in Hospital," published in Britain in 1959, led to the revolution in pediatric hospital care in Britain and other western countries. The committee found that children were expected to conform to ward routines, lie quietly in their cots, and were not allowed to play; their parents were allowed to visit only during restricted visiting hours. The Platt Report recommended unrestricted visiting and accommodation for parents, especially for children less than 5 years old. The committee sought to minimize potential damage caused by the separation of children and mothers and to encourage appropriate stimulation for the child in the hospital.

The recommendations of the Platt report took time to be implemented. Gradually the views of pediatricians and pediatric nurses changed from

those prevailing before the Second World War and became less discordant with the views of child psychiatrists. Pediatricians and child psychiatrists now agree about the significant impact of separation, illness, and hospitalization on children and their families.

The length of hospitalization and the reasons children and adolescents require hospitalization have changed considerably. Children used to be admitted to the hospital for much longer periods than they are today. Changes in morbidity and treatment modalities, along with alternatives to inpatient care such as day surgery, home intravenous antibiotics, ventilators and advances to maintain technologically dependent children at home, have led to shorter and less frequent hospitalizations. The average length of stay has been reduced by almost half and same day discharges have almost tripled in the past 20 years. The leading causes for hospitalization recently have been asthma, gastrointestinal disorders, pneumonia, and seizures; inpatient surgeries have declined as outpatient procedures have increased. Hospitalizations lasting more than 7 days now are likely to be for psychiatric reasons, chemotherapy and treatment of malignancies, or to address complications of procedures or prior treatment.

As survival rates have increased, at least 90 percent of children with chronic conditions now survive to young adulthood. Inpatient hospitalization for children with more common conditions has shifted to outpatient care, leaving proportionally more complicated, chronically ill children in inpatient settings.

Hospitalized children now are significantly medically complex. This small proportion of children account for most of health care expenditures: 20 percent of children with complex chronic illness and children with special health care need account for 80 percent of costs. The prototypical medically complex child has a multisystem disorder, is probably technologically dependent, has frequent inpatient admissions, requires multiple medications, multiple subspecialists, and optimal care coordination across inpatient and outpatient settings. Care is usually provided at tertiary care centers. These children have three times longer lengths of stay, over 10 times greater charges, significantly higher gastrostomy and tracheostomy placements, intensive resource use, and 15 times inpatient mortality. Seriously and chronically ill children are "growing up in the hospital," and the lives of these children and their families may differ considerably from their peers.

Table 55.3–1.
Estimated Prevalence of Chronic Diseases and Conditions in Children Ages 0–20 Years in the United States

	Prevalence Estimates Per 1,000	
Disorder	**1980**	**2007**

Asthma (ages 0–17 years)	20–53	52–127
Juvenile rheumatoid arthritis (ages 0–15 years)		0.1–0.2
Systemic lupus erythematosus (ages 1–12 years)		0.05–0.5
Cystic fibrosis (in Caucasians)	0.2	0.3125
Ulcerative colitis (estimated based on incidence)		0.4–0.5
Crohn disease (estimated based on incidence)		0.7–1.0
Type 1 diabetes	1.8	1.6–2.5
Congenital heart disease (based on data in newborns)	2.0–7.0	8.0–12.0
Pediatric malignancies		0.011–0.014

CONTEMPORARY PEDIATRIC HOSPITALS

Contemporary pediatric hospitals and wards are staffed and structured differently than adult units, and children's hospitals typically include features not found in most adult settings. The physical layout of a children's hospital or pediatric ward is usually designed to facilitate the child's primary caregiver staying with the child. Sleeping arrangements, showers, laundry, and other facilities are made available to increase individual patient and parent comfort and facilitate extended stays with children. Private rooms or conference rooms exist to allow doctors and parents to confer in privacy.

The last 50 years have seen significant changes in the approach to health care for children. Accommodations are now made to encourage visits even from large or extended families, placing the experiences of children and their families at the heart of health care. To decrease stress in waiting areas, toys and activities are available in brightly colored, child-friendly spaces. In order to provide maximum continuity of care, every child may be assigned a primary nurse who will continue to be the child's primary nurse in subsequent hospitalizations. While most large medical centers have hotels nearby for family members, children's hospitals often provide additional settings such as Ronald MacDonald House, where entire families may stay in small, subsidized apartments. Social workers, parent advocates, groups for parents and for siblings, and educational groups provide support to families. Disease-specific summer camps for patients and siblings and other recreational group activities for sick children are characteristically managed through children's hospitals or disease-specific charitable organizations. Parents, initially focused on potentially life-threatening physical concerns, often express deep gratitude for these services later and are grateful for additional support of their child's psychosocial development.

Child life programs have become standard at children's hospitals. They are available on most pediatric inpatient units and are found in some outpatient pediatric settings. Their purpose is to help minimize disruption in a child's development by hospitalization, surgery, or other potentially frightening or distressing medical experiences. Child life specialists provide

recreational activities, often in designated places where children can engage in therapeutic play. Child life specialists help provide information and preoperative preparation and can rehearse useful coping strategies for medical procedures with children. Preoperative surgical preparation programs have been shown to be effective in reducing anxiety, which may improve children's abilities to cope with the stress of a surgical procedure more effectively. The goal is to help children understand their feelings about their experiences in the hospital and facilitate better adjustment.

Emphasis has been placed on minimizing physically painful experiences. Regular assessment of pain in children using age-appropriate measures has been widely implemented, and the use of nonpharmacological and pharmacological techniques suitable for young patients are more standardized and better understood. These changes have been important in further decreasing the stress of hospitalization for children and adolescents.

Education is a major part of each child's life, and hospital schools and teachers are present in most pediatric settings. Some children's hospitals have arrangements with the local school system to provide teachers, and others employ their own. Either way, collaboration with the child's school is important in helping maintain academic progress as much as possible.

Principles of behavior management are typically incorporated into the culture of pediatric units, where staff can play an important role in modeling and encouraging age-appropriate behavioral management by parents. The absence of limits and consequences for misbehavior can be harmful for children, leading to acting out and difficulties in regulation of emotion and behavior. Although mental health professionals have limited opportunities to monitor behavior directly, and cannot substitute for parents and staff on the unit, frequent consultation with psychologists, social workers, and child psychiatrists can be helpful to the care team, to patients, and to families.

Helping children, adolescents, and their families adapt to hospitalization should involve the child, parents, family, and hospital staff. In current family focused models of care, a child's care is planned around the whole family with parents considered integral to the treatment process. Family-centered care has proven to be valuable in helping families and children cope with the stress, fear, and altered parenting roles that accompany a child's illness and hospitalization. While the creation of a nurturing and developmentally sensitive hospital environment is primarily the responsibility of staff, the assistance of family members is essential in establishing structure and maintaining stability for sick children. "Anticipatory guidance" about parenting struggles in and out of the hospital can be helpful. Parents of children with life-threatening illnesses often feel guilty and need support to maintain usual parenting practices. Parents of chronically ill children should be included in discussions with

hospital staff about realistic rules; this leads to increased compliance with daily schedules and consistent behavioral expectations.

Inpatient pediatric care requires constant review by nursing and medical staff as well as by patients and families to remain relevant and effective. While past initiatives have created the current environment of pediatric inpatient care, hospital staff should remain open to adopting additional practices as further studies identify better alternatives.

A CHILD'S UNDERSTANDING AND RESPONSE TO ILLNESS

Sick children are not a homogeneous group. They come to the hospital having had very different life experiences, irrespective of their medical condition. The presence of an acute, life-threatening, or chronic physical illness by itself is insufficient to produce psychiatric problems, but it may make the child more vulnerable to stresses such as conflict and confusion in the family, and may exacerbate any existing psychiatric illness or behavior problems. How a child will react to illness and hospitalization is a reflection of their age and understanding of their situation, their parents' response, and the nature of the illness and its treatment. There is little evidence that greater severity of the medical condition is itself associated with more negative emotional outcomes, but some observations of greater vulnerability for negative emotional response have been made.

Preschool children are hospitalized more than older, and children younger than 7 years are at risk of more frequent and potentially lasting emotional and behavioral problems after hospitalization than older children. The risk of later aggressive and hyperactive behaviors can usually be predicted by the mother's emotional state and the child's premorbid behavior. Baseline internalizing or externalizing behaviors may be the best predictor of increased behavior problems 3 to 6 months after hospitalization, and medically ill children with greater premorbid emotional and behavioral problems have more psychiatric disorders following their illness. When illness is experienced as an insurmountable obstacle or the cause of rejection by loved ones, the child may react with extreme emotional distress, and suicidal ideation and behavior may occur. Overall, 20 percent of children with chronic conditions have behavioral and emotional symptoms, double the rate of the general population.

Conditions that are less visible, have a remitting–relapsing course, or uncertain prognosis may be associated with greater emotional problems. Males with chronic conditions appear to have more emotional sequelae than females. Children with brain damage have poorer overall adaptation to illness, and adopted children and their parents also seem to have more problems with hospitalization. However, the personality characteristics of sociability, flexibility, and physical attractiveness appear to be protective, and may mitigate against more severe emotional problems. Children from families with clear boundaries, the capacity to balance competing family

needs, the ability to maintain flexibility, and the tendency to ascribe positive meanings to life events, fare best.

Chronic and serious physical illnesses involve recurrent separations and frequently upsetting events, which can be distressing to both parents and children. Surgery, medication side effects, procedures, and diagnostic tests can be frightening, confusing, unscheduled, disruptive to sleep, and painful. The burdens of chronic illness are pervasive and may produce chronically stressful situations, resulting in maladaptive behavior, emotional symptoms, and added financial burdens, all contributing to conflict, guilt, and anxiety in the family.

Medical illness and hospitalization can pose serious threats to the developmental process, and impact the physical and emotional environment, disciplinary context, schooling, family life, and the neurobiology of the chronically ill child, and thus profoundly affect cognitive, emotional, social and sexual developmental trajectories. Invasive procedures, separation from family, friends and school, lost opportunities to develop interests, and a diminished sense of self-mastery further complicate development. Environmental restrictions, such as isolation for infection control, and limited activities while in the hospital lead to fewer opportunities to discover personal strengths and weaknesses. Chronically ill children often have difficulty finding something to be "good at" and have trouble feeling confident and competent. Delayed growth and puberty are common in chronically ill children and adolescents, especially when malnutrition, prolonged corticosteroid treatment, or chronic inflammatory processes are prominent. These growth delays typically may have additional negative psychological implications.

Table 55.3–2.
Impact of Illness and Hospitalization

Age	Child's Response	Possible Effects of Hospitalization	Possible Effects of Chronic Illness
Infancy 0–18 months	Unable to understand Fear abandonment Perceive pain	Anxiety and confusion with separation	Decrease access to environment Parents guilt, anger, and grief may interfere with attachment Difficult to establish trust Greater helplessness
Toddler 18 months–3 years	Fear abandonment Little capacity for reasoning or understanding experience	Regression Negativism Restricted activities Important maintain routines Transitional objects help	Delay motor and language development Parental reluctance to set limits Problems with bladder and bowel function
Preschool	Fear mutilation	Fear abandonment	Parental overprotection

3–5 years	Magical thinking Attributes illness and pain to punishment for bad behavior	Fear mutilation Hospital play beneficial	Regression Limited initiative Compromised social skills
School age 6–12 years	Illness punishment for bad behavior or due to own ideas of cause Understand simple bodily structure and function Understand purpose of treatment	Fear pain Benefit from clear expectations and concrete explanations Important maintain roles	Alienation from peers Fewer social interactions Limited activities Parental overprotection
Adolescent 13–18 years	Abstract thinking Understand body processes and implications of illness Understand cause and effect and hypothetical reasoning Comprehend mind-body issues	Fear alienation from peers Value privacy Appearance important Want participate in treatment decisions	Concerned about appearing different from peers Concerned about medication side effects Risk-taking behavior Treatment noncompliance Immaturity and overly dependent on parents

One week after discharge from the hospital, a child often demonstrates increased anxiety about separation, may experience sleep problems, and may be more aggressive. These behavioral changes should quickly resolve as the child adjusts to the return home, but this may not always happen. Repeated admissions of children coming from dysfunctional or disadvantaged families are very likely to be associated with emotional and behavioral problems following hospitalization, but all children may be at risk. The most difficult period of adjustment to a hospitalization may be near the beginning, when new and painful experiences and separation are most distressing. Prolonged periods of hospitalization bring the risk of developing maladaptive behaviors. In a meta-analysis of studies of children's behavior after hospitalization, children hospitalized for 2 to 3 days showed more behavioral problems than those hospitalized for longer or shorter periods, and overall, briefer hospitalizations of older children were the least likely to have detrimental effects.

Posttraumatic stress disorder (PTSD) has been described following critical illness and surgery. Serious accidents and hospitalizations following acute physical trauma may be associated with PTSD, while the risk of PTSD in children and parents after simple planned surgery is almost nonexistent. Children and parents after newly diagnosed cancer and children and parents following solid organ transplantation may also be at risk of PTSD. Parental distress and characteristics of the family environment are more correlated with the development of PTSD symptoms in sick children than family composition, course of hospitalization, or the extent of a child's injuries. In contrast, commitment, help, support, and the open expression of feelings in the family can facilitate a child's better adjustment to newly diagnosed cancer and associated intensive medical treatment. Sometimes

the treatment of life-threatening medical illness may be experienced as repeated trauma, which might not impact during or immediately after treatment, but rather appear as long-term effects on affect modulation and interpersonal relationships.

Attempts have been made to mitigate the negative consequences of hospitalization, not only in the way hospitals are structured and function but also in provision of preventative educational or behavioral interventions. In a meta-analysis of a variety of interventions, psychoeducation was generally not effective with younger children, and any benefits from the experimental interventions disappeared by 4 weeks after discharge. Other programs, which focus on increasing parents' knowledge and understanding of the range of behaviors and emotions that young children typically display during and after hospitalization, and which support direct parent participation in their child's physical and emotional care are more effective, and help reduce the mother's emotional stress, depression and anxiety, reduce the child's negative behavioral symptoms, and increase the child's adaptability up to 12 months after hospitalization.

When psychiatric and behavioral symptoms are severe or persistent, a thorough psychiatric assessment, based on multiple brief examinations of the child and information gathered from parents, other family members, hospital staff, and teachers, is needed. A child's biological vulnerability to depression and anxiety may be inferred from a history of previous psychiatric symptoms or psychiatric treatment or from a family history of a mood, anxiety disorder, or other psychiatric disorder. Psychiatric syndromes of adjustment disorder, major depression, anxiety, delirium, and psychosis can be diagnosed in medically ill children and require intervention.

Behavioral and cognitive approaches, such as distraction, guided imagery, and relaxation for reducing procedural distress, are well established in the pediatric hospital setting. Supportive individual and family therapy and cognitive behavioral and family therapy may be useful in reducing anxiety and in providing patients and families with adaptive strategies to use in coping with illness and its treatment. Psychopharmacotherapy may be indicated based on severity and duration of psychiatric symptoms.

RESPONSES TO SPECIFIC ILLNESSES

Another way to evaluate the impact of illness and hospitalization on children and adolescents is to review reports of children's responses to specific disorders. As diagnosis and treatment become more effective, children are living longer and forced to cope with a variety of disorders. Several examples are described below.

Asthma

Asthma, a reversible reactive airway disease triggered by immunological and nonimmunological factors, is the most prevalent chronic illness of childhood, affecting more than 1 in 20 school aged children. While asthma can present at any age, 75 percent of patients present in the first 3 years of life. The sensation of chest tightness, respiratory distress, and anxiety may be particularly frightening to young children, while adolescents with new onset asthma seem to adapt more easily than expected. There is a high comorbidity in asthma with anxiety and depressive disorders, although this risk may be less when asthma is well controlled. Anxiety plays a greater role when a child's symptoms are mild, and it is more difficult to quantify symptoms when asthma is severe. Trouble breathing typically triggers anxiety, and proinflammatory cytokine release is found in both asthma and depression. Maximizing the child's ability to perceive symptoms and to comply with treatment can reduce the morbidity and mortality of asthma.

Rheumatological Disorders

Rheumatological disorders in pediatric patients are typically more severe clinically than the same disorders in adults and may have significant psychiatric and medical consequences. These conditions are treated with potent immunosuppressive medications, steroids, and immune modulators that themselves may be associated with psychiatric side effects. Psychiatric symptoms may also be a reflection of an autoimmune process or related underlying vasculitis, or the psychiatric disorder may be unrelated and coincident. In determining the long-term effects of psychiatric complications of pediatric rheumatological disorders, parental functioning may play a role, and parental distress and maternal depression have been associated with greater behavior problems in the child.

Juvenile idiopathic arthritis (JIA) can lead to social isolation, chronic pain, and skeletal deformities, which may lead to depression, anxiety, low self-esteem and lower quality of life. In studies of children and adolescents with JIA, those with less severe physical disabilities were often more affected psychologically, and more emotional and social problems and pain are noted in children whose parents feel stressed, pessimistic and guilty.

Systemic lupus erythematosus (SLE) in children and adolescents is associated anxiety, depression, and cognitive dysfunction. Pain, fatigue, and depression may be related to the effects of a chronic illness, medications, or to SLE itself. Central nervous system involvement, with both neurological and psychiatric symptoms, occurs in more than half of children and adolescents with SLE, usually early in the course of the disease. Neurocognitive impairment is common, and complaints of memory and concentration problems are frequent even when not documented on neuropsychological evaluation.

Cystic Fibrosis

Survival of patients with cystic fibrosis (CF) has been considerably extended by improved therapy of pulmonary infections, improved nutritional management, and an organized system of centralized care with physicians experienced in caring for the complications of CF. The psychological and social functioning of patients with CF is similar to that of healthy peers until their disease becomes severe and their physical and social activities become increasingly limited. Patients with CF are at increased risk of psychiatric problems, such as depression, and their ability to cope with their disease and comply with treatment are important for their quality of life. The psychological and social problems associated with CF may be intensified during adolescence, when more problems with independence, sexuality, relationships, and noncompliance emerge. It is difficult to develop independence when required daily care is complex and may require assistance, and it is difficult to develop self-esteem when chronic productive cough, physical limitations, and peer relationships are disrupted by recurrent hospitalizations. Boys and girls seem to respond differently: girls seem to have a better body image and to be more distressed coughing in public; and for boys, short stature and pubertal delay may have a greater impact on the quality of life than the issue of shortened survival.

CF is a common indication for lung transplantation. Most transplants are living donor lobar transplants, and are usually performed in adolescents. The mean survival is greater than 10 years posttransplant, which allows recipients to develop long-term goals and plan for independence.

Inflammatory Bowel Disease

Inflammatory bowel disease (IBD) is also often associated with depression and anxiety. Children with a good knowledge of their illness, an internal locus of control, a healthy family climate, and an open social network do well, and can maintain good mental health functioning despite severe illness. Disease severity appears to be a significant predictor of depressive symptoms in adolescents. The relationship between social support and adjustment observed with other conditions is also noted in children with IBD; those with high levels of social support from both family and peers show significantly better adjustment than those with limited support from either source. However, parental depression is associated with greater depressive severity in children more than adolescents with IBD. Cognitive behavior therapy and pharmacotherapy, with fluoxetine most studied, have been shown to be helpful in treating depression in children and adolescents with IBD.

Insulin-Dependent Diabetes Mellitus

Type 1 insulin-dependent diabetes mellitus (IDDM) affects many aspects of everyday living and requires a high degree of personal responsibility to manage effectively. Poor diabetic control in childhood is associated with greater risk of long-term complications and may lead to cognitive difficulties, making careful management more critical. Depression and anxiety are common in children and adolescents with IDDM, and risk is higher during the first year after diagnosis, in those with pre-existing anxiety, and in those whose mothers have psychiatric problems. Adolescence and associated physiological changes of puberty can lead to increased insulin resistance, at the same time peer relationships become more complex and parental supervision may be withdrawn, all leading to greater problems with compliance and glycemic control. There appears to be a direct relationship between the diabetic adolescent's emotional and physical well-being. Adolescent juvenile diabetics, who are considered well-controlled medically, tend also to be well-adjusted at home and with peers, are appropriately independent, and have good self-esteem and a sense of internal control.

Diabetes is associated with increased risk of eating disorders, which are associated with poorer glycemic control and increased risk of diabetic comorbidities. Intentional omission of insulin to lose weight or as form of suicidal behavior may occur with type 1. Similarly some patients with type 2 will intentionally omit oral hypoglycemic drugs to lose weight. Binge eating disorders are also common. Patients tend to deny their manipulation of their diet or insulin, but poor glycemic control, missed clinical appointments, and recurrent episodes of diabetic ketoacidosis or hypoglycemia should lead to suspicion of an eating disorder.

Heart Disease

Depression is common in patients with congenital heart disease, and can exacerbate the physical consequences of the disorder. Children with congenital heart disease may have more medical fear and fear of the unknown, more physiological anxiety, more depression, and more delinquent behavior than the general pediatric population. Those with cyanotic malformations may be at higher risk, which may be further exacerbated by maternal anxiety. Pain and separation from parents is more upsetting to younger children, while older children and adolescents are more focused on their physical status and the long-term implications of their condition. Accommodating to complete heart block and the implantation of a cardiac pacemaker often produces emotional distress; those children and adolescents without psychiatric symptoms who undergo the procedure demonstrate intellectualization, identification with medical personnel, and denial during the medical emergency. Problems in adjustment that occur following cardiac transplantation reflect the patient's previous level of emotional adjustment more than a specific response to the

procedure.

Ventricular assist devices (VADs) have become widely used in the management of end-stage heart failure in children particularly for cardiomyopathy or myocarditis patients. Although overall outcomes are favorable, medical complications are common, especially for children with complex congenital heart defects. VADs are most useful as a bridge to recovery or to eventual transplant.

Surgical Procedures

Some surgical procedures may be more disturbing than others, and children usually have more emotional problems after genitourinary surgery than after tonsillectomy. Although early psychoanalytic studies on the effect of tonsillectomy concluded that this operation was a stressful experience for every child, especially those with a history of previous trauma, currently the majority of children do well. Despite mild adverse behavioral changes in some younger children, more than 90 percent undergo the procedure without negative consequences.

Hemophilia

Hemophilia is an example of an inherited X-linked coagulation defect that serves as an illustration of individual and family response to a chronic and potentially lethal disorder. Most boys and their families adapt well and are able to maintain age-appropriate function at home and school despite their illness. The few who show poor adjustment tend to have over-protective, over-anxious, or guilt-ridden mothers, and the boys respond either with social withdrawal and fear, or with inappropriate recklessness.

Cancer

Cancer is the second leading cause of death for children ages 5 to 14 years and the fourth leading cause of death in adolescents ages 15 to 19 years. Progress in the treatment of pediatric malignancies has led to an increase in the number of children, adolescents and young adults surviving free of neoplastic disease, but this increased survival has come at a price. Survivors may suffer endocrinopathies, growth abnormalities, intellectual decline, neurocognitive and psychosocial dysfunction and increased rates of anxiety, depression, dementia, personality and behavioral disturbances. They also have an increased actuarial risk of developing a second malignant neoplasm. The prevalence rate of psychiatric disorders in children and adolescents with cancer is mixed; ranging from rare to as many as half of the patients with anxiety disorders, PTSD, and depression. Psychological interventions have shown promise in decreasing distress associated with diagnosis and treatment, but have not shown an effect on long-term clinical outcomes. Survivors of childhood malignancies may have impaired quality

of life, compromised social functioning and emotional well-being, including suicidal ideation, and diminished academic and occupational success as adults.

A CHILD'S RESPONSE TO ILLNESS AND HOSPITALIZATION AT DIFFERENT AGES

A child's behavioral response and concept of illness and its causes follow a sequence of developmental stages related to their perceptual ability and cognitive development. This section is specifically limited to a description of a child's response to illness and hospitalization at different ages.

Infants (Birth to 18 Months)

Serious illness and hospitalization separate infants from their parents and sets them in a strange, frightening, unpredictable environment. This can provoke anxiety and confusion putting them at risk of disruption in the development of basic trust. Infants younger than 7 months rarely protest separation from their mothers and seem to accept substitute caretakers in the hospital. They may then appear briefly but dramatically upset upon return home. Infants older than 7 months protest initially, appear uncooperative and subdued in the hospital, and then require a readjustment period upon return home. The more infants are provided with contact and comfort from parents when distressed by illness or hospitalization, the better basic trust can be protected. Infants subjected to unexpected medical procedures may have difficulty forming attachments and become anxious. Repeated exposure to pain and the absence of social stimulation in neonates may affect later pain sensitivity, brain volume, neuroendocrine functioning, behavior, and cognition. If parents are unable to hold and soothe their babies because of the nature of an injury or medical care, the babies are at risk of apathy and withdrawal, the attachment process may be disrupted, and separation anxiety may develop. Infants may then refuse to interact with parents, or become clingy, demanding, and difficult to soothe.

It is important to maintain consistent nursing staff and infant routines, preferably with parents or familiar caregivers in the baby's room. If they cannot stay attempts should focus on maintaining regular times for parental visits. Parents should be encouraged not to leave only after infants are asleep as this adds to an infant's anxiety and distrust. Normal stranger and separation anxiety may become exacerbated by frequent staff changes, so attention to providing familiar affectionate staff on a regular schedule may be helpful. Both parents and babies benefit from staff support.

Parents may need staff to model appropriate interactions and interventions with their baby.

Toddlers (18 Months to 3 Years)

Toddlers' primary developmental goal is to establish early autonomy and separation-individuation from their parents. These tasks require a secure base. They are difficult to accomplish without having first established basic trust, which may not have occurred effectively if the toddler was ill or hospitalized as an infant. Toddlers fear abandonment when separated from their parents or primary caregivers. Like infants, they benefit from maintenance of routines and stable relationships, which support early independence and efforts at developing competency. Transitional objects may be particularly helpful in this age group, so bringing toys, blankets or other familiar objects from home to the hospital or doctor's office may help calm an anxious toddler.

The requirements of medical care can be infantilizing and carry high risk of regression, increased dependence, and negativism. It is easy for parents to become over protective or overly permissive during a toddler's illness and behavior problems may result. If these parenting patterns persist after recovery, tantrums and oppositional behavior increase, which leads to parental anger and threat of rejection, followed by guilt and more permissiveness; a vicious cycle of behavior problems can develop. Providing anticipatory guidance to parents during a toddler's hospitalization and follow-up behavioral management and support after discharge can be helpful.

Preschool Age (3 to 5 Years)

Preschool age children seem to be more upset by hospitalization than older children or infants, but they are the most likely to benefit and respond positively when their environment is supportive. When the child is prepared for hospitalization, parental visits are consistent and parents participate in their child's care during hospitalization, the potential emotionally traumatic effects of hospitalization need not occur.

The preschool age is the period of "magical thinking" when children typically attribute illness and pain to punishment and hospitalization and separation to rejection. Limited understanding and concrete thinking contribute to fears and misconceptions, leading to feelings of guilt and distress. Children this age have a very rudimentary concept of the body and its parts. They fear mutilation, and all parts of the body are considered equally essential. This limited understanding, in association with fears of castration, may cause procedures such as phlebotomy and surgery to be especially terrifying. Children are often unable to articulate their fears. They may not understand adult attempts to describe bodily structures and functions, illness effects, or treatment plans. Children of this age often benefit from preparation for procedures through medical play. When material is presented in a manner congruent with cognitive function, comprehension improves and fear is reduced.

School Age Children (6 to 12 Years)

School age children are typically active and focused. Developmentally, they are working to become secure in their own abilities and accomplishments. During this period they acquire academic and social skills and begin to internalize parental and societal values. School age children, like younger children, often believe that their illness is a punishment for misbehavior. Their strong respect for authority and rules often leads school age children to feel illness or injury is the consequence of not following rules properly. Because their understanding of anatomy and physiology is more complete, children this age have less fear of mutilation than younger children. However, because their thinking remains concrete, they cannot comprehend hypothetical possibilities or complicated explanations. Adult attempts to explain illness, injury, or anticipated medical procedures may be confusing, especially when explanations are complex or abstract. Medical explanations can be easily misunderstood or misinterpreted, leading to confusion and increased fear which the child may be reluctant or unable to disclose. Drawings, photographs and clear explanations at an appropriate cognitive level will better help children understand and master their experiences. It is important to minimize physical pain whenever possible. Difficulty coping with pain, and anxiety about pain, may compromise a child's ability to understand and participate in care. Again, potentially painful or frightening experiences may be less stressful with adequate preparation and explanation. School age children should be encouraged to express their fears, ask questions, and talk about their feelings to engage them in the process and help increase their sense of mastery of their illness.

It is important to maintain routines and relationships for school age children during illness or hospitalization. Children should be supported to meet their responsibilities for school, maintain connections with peers, and participate in as many of their usual activities as possible, even when ill or hospitalized. These older children have an increasing awareness of the personal significance and implications of their illness. The ability to adapt to new, challenging experiences increases with age and expanding cognitive function. Coping with illness and hospitalization may increase children's resilience and teach skills useful in other settings.

Adolescents (13 to 18 Years)

Adolescence is a complex process of transition from childhood to adulthood and is a challenging time for most people. Adolescents' ability to understand their illness, their body and its functions, and the causes, nature, and treatment of their illness expands as cognitive skills increase. As formal operations develop in individuals of normal intelligence, adolescents are able to hypothesize cause and effect, understand

hypothetical suppositions and multifactorial explanations, and participate in decision making and assent/consent for treatment. Adolescents highly value their privacy and confidentiality, and are more skeptical of simplistic explanations or superficial reassurance.

Adolescents are usually aware of the implications of their illness. They may be concerned about its effects on their relationships, fertility and sexuality, and vocational possibilities. They fear the loss of independence and control, and of their sense of physical well-being and integrity. Younger adolescents especially do not want to be different from their peers. They are exquisitely sensitive to changes in their appearance and dread permanent disfigurement. They understand the permanence of death and fear it.

Issues of puberty, autonomy, personal identity, sexuality, education, and vocational choice are challenging to adolescents. They are more difficult for an adolescent who is also coping with an acute or chronic illness. Illness can be associated with medical setbacks, impaired physical or mental abilities, forced dependence, and perceived poor prognosis. When an illness is severe or disfiguring, an adolescent's identity, self-image, and ego-development can be negatively affected. Identity formation is especially challenging in the hospital setting, and adolescents may come to rely on their illness or disability as their identity, rather than setting goals or aspirations beyond the hospital environment. Biological, psychological, and social developments are highly interrelated. Delays in one area, such as body image or sexual development, lead to consequences in other domains. Delayed puberty in boys may lead to low self-esteem and apparent immaturity, which in turn can foster dependence on the family and eventual reduced educational qualifications and increased under employment or unemployment.

Not every young person with a chronic condition has significant psychopathology. Adolescents with a chronic illness are generally well adjusted and perceive their illness as less severe than their physicians assume they do. When adolescents perceive their illness to be severe, they are more likely to have symptoms of depression and low self-esteem. Many who live with chronic illness or disabilities are well adjusted socially and emotionally. Those adolescents with the greatest uncertainty of their outcome often have more difficulty accepting and discussing their condition, and may avoid disclosure to peers. In a study of chronically ill adolescents with a variety of disorders compared to matched controls, there were no differences in levels of anxiety or self-esteem between the two groups, and both held a similarly positive outlook on life. This finding may be related to the ability of adolescents to deny the worst aspects and consequences of their illness in order to function effectively, or it may be related to their experience coping with multiple minor stresses resulting in greater resilience and more successful coping strategies their healthy peers.

Compliance with treatment, rules, and authority may be problematic for

some medically ill adolescents, and they may be more likely to demonstrate risky behaviors than their healthy peers. Medically ill adolescents require the same anticipatory guidance and close supportive follow-up as any adolescent at risk.

It is important that adolescents participate in their own care and treatment decisions, although in most cases, diagnosis and treatment information is still directed to the parents. Hospitalized adolescents prefer to have medical information provided directly to them by health care professionals, but they often want their parents present for support. Adolescents would like to have discussions of diagnosis and prognosis as soon as possible after admission, even when their condition may be terminal. Adolescents often recognize the imminence of their death and want to discuss their feelings. They may request information that the parents are reluctant to disclose, and may be ready to discontinue treatment when their parents are not. In these situations, it is important that the health care team is ready to support both the adolescent and parents and to provide ways to bridge their differences. Advance care planning is encouraged by the Institute of Medicine and the American Academy of Pediatrics, and developmentally appropriate tools for discussing end-of-life care are available for adolescents and young adults.

FAMILY'S RESPONSE TO CHILD'S ILLNESS

The family of a seriously or chronically ill child is as affected by the illness and frequent hospitalizations as is the child, and the social impact of the illness is felt throughout the family system. Family systems are dynamic and can be sources of risk or protection for a child with a serious medical condition. Chronic conditions present the entire family with an additional set of tasks, responsibilities, and worries. There is personal impact on a parent, financial impact on the family, and implications for family and social relationships. Parents must deal with the day to day management of the illness, with the increased financial burden associated with medical bills and time lost from work, and with the uncertainty about treatment and their child's prognosis. This can place additional strain on a marriage and can significantly affect the parents' relationships with other children. Relatives, teachers, and friends may be available to provide support, but they also may be negatively impacted by the child's illness.

Nine aspects of a resilient family process have been identified: (1) balancing the chronic condition with other family needs, (2) maintaining clear family boundaries, (3) developing communication competence, (4) attributing positive meanings to the situation, (5) maintaining family flexibility, (6) maintaining a commitment to the family as a unit, (7) engaging in active coping efforts, (8) maintaining social integration, and (9) developing collaborative relationships with professionals. Family cohesion, open communication, and the ability to adapt to conflict are

indicative of healthy family function and psychological adjustment in families with chronically ill children.

Parental temperament and coping skills strongly influence how their child copes. Children take their cues for coping and adaptation to illness primarily from their parents, especially their mothers. Parents who are able to isolate their own fears and anxieties during crises are better able to enhance their child's understanding of the illness and to facilitate self-care and independence in their child. Parental anxiety can have a negative effect, and children with very anxious mothers often have less distress during procedures when their mothers are not present.

Parents tend to remain at the hospital during the initial diagnosis and early therapeutic interventions, and often return when illness worsens or complications arise. Once the first crisis passes, the mother is most likely to take on the day-to-day care of the child. She is often the parent who remains in the hospital with the child, accompanies the child to physician visits, provides support for procedures, and acts as advocate for the child with practitioners.

Studies of mothers of children with different chronic disorders have consistently found the mean level of reported adjustment problems to fall at about one standard deviation above the mean for the general population. Research on fathers' reactions is consistent with studies in mothers, although fathers are less well studied. Fathers traditionally take on additional financial burdens associated with a child's illness, have to endure separation from their spouse and child during hospitalizations, and deal with feelings of helplessness due to the inability to protect their child from harm. Differences in coping style may contribute to emerging family dysfunction. Parents, like children with a chronic disorder, are at risk of maladjustment, but there are considerable individual variations and many do well. Parents may experience marital distress and impaired sexual relationships, especially after their child's diagnosis of cancer. But parents of children with cancer may also experience their marriages as satisfactory and their partners as supportive, and higher divorce rates have not been found. While marital discord is associated with increased psychological problems in chronically ill children, a mutually supportive marriage can benefit the family's ability to cope with the stress of chronic illness. Parents may respond differently to traumatic brain injury in their children: fathers are more likely to use denial and report greater injury-related distress, while mothers are more likely to demonstrate acceptance and emotion-focused coping. Mothers may also be better able to adjust over time, while fathers' adjustment and marital satisfaction may not improve.

Siblings of a chronically ill child are in a very difficult position. Siblings often are asked to take increasingly adult roles to "help out," while receiving less attention and support from their parents. The sibling may love and worry about their sick brother or sister, and at the same time

resent them, causing feelings of guilt. Sibling rivalry and jealousy increase when the focus of both parents and the family's resources are directed toward the ill child. It is further exacerbated when mothers spend more time with their sick children than the healthy siblings. Siblings of chronically ill children are more likely to develop depression, anxiety, a negative self-image, and fewer quality peer activities. Younger siblings may be more resilient to these consequences, and life threatening or severe illnesses may exacerbate the risk.

Sibling adjustment typically parallels the family's adjustment, and finding ways to help the mother feel confident in managing her child's illness is integral to her ability to provide the optimal home environment for both the sick child and the healthy siblings. Parents of children with cancer, CF, and epilepsy rated the siblings' quality of life better than the siblings themselves did.

Siblings may be asked to donate hematopoietic stem cells for transplantation. Donor siblings often feel they "had no choice" about donating. Both donors and recipients may have post-traumatic stress symptoms following the transplant, and nondonor siblings may report more anxiety and lower self-esteem. Teachers have observed donors to have both more adaptive coping and more school problems and attention-seeking behaviors than nondonor siblings. Donors report the psychological aspects of donating outweigh any physical consequences, whether their sibling lives or dies. If the recipient dies, the donor sibling frequently feels guilty. They often use spirituality and religion as primary coping strategies. The donation process and having an ill sibling are associated with a range of positive and negative emotions. Siblings often feel isolated and that their family life interrupted. Siblings want to be included in decisions, be provided information, be given choices, have help sharing their feelings, be provided opportunities to meet with peers, and experience a supportive hospital environment.

Families who actively seek information and services, work to solve problems and express their feelings, and are able to balance personal, family, and illness needs show better adaptation to a child's illness and hospitalization. Maintaining supportive relationships with people in the community can also be an important protective factor. Informal support networks provide emotional and informational support and sometimes tangible aid, but when they are not enough, mental health services and formal support groups are needed.

When families are given a new devastating diagnosis, such as childhood cancer, the family may respond as though the event is a matter of chance beyond their control, and exhibit denial, detachment, intellectualization, or substance abuse. Optimally, the family views the illness as a challenge that is controllable and changeable, making long-term difficulties in psychological adjustment less likely.

Health care providers are very influential and play a powerful role in the lives of families with a sick child. Parents look to them for essential medical information and guidance. Pediatricians need to be aware that anything they say, whether well reasoned or not, can have an enormous impact on parents and their children, with potential effects that may be immediate or long lasting. Medical communication delivered in a thoughtful and sensitive manner can promote competent family functioning and resilience, despite life-threatening illness. However, when health information is not thought through before being given, it may be misconstrued, misinterpreted, or misunderstood, and may lead to a constellation of abnormal behaviors and altered parent-child interactions that are referred to as the vulnerable child syndrome (VCS), first described by Green and Solnit in 1964.

The cardinal features of VCS include (1) pathological separation difficulties, such as poor sleep in infants or recurrent unexplained abdominal pain or school phobia in older children; (2) infantilization with overprotective, overly indulgent, and overly solicitous parent, and an overly dependent, disobedient, irritable, argumentative, and uncooperative child; (3) bodily over concerns, when the child has many physical complaints and the parent (usually the mother) has difficulty differentiating between the child's discomfort and their own distress, and cannot distinguish mild, self-limited symptoms from potentially serious illness; and (4) separation anxiety, delay starting school and school underachievement. Predisposing factors to VCS include a history of infertility, miscarriage, illness during pregnancy; prematurity; a life-threatening illness; or a child with a chronic medical illness, especially if there was hospitalization during the first year. VCS is linked to more frequent ER visits and inpatient hospitalizations. Depending on how their parents perceive the event, even older children seen in an emergency department for minor head trauma or hospitalized for gastroenteritis or asthma may be at risk of VCS.

The prevalence of VCS in a community sample was estimated at 10 percent in 1995, but maybe higher today with improved survival of premature children and children with serious chronic illnesses. The degree of parental perception of vulnerability does not necessarily correlate with the severity of neonatal problems, birth weight, or current health. Rather, higher levels of maternal anxiety at neonatal discharge better predict parental perception of child's vulnerability and are associated with worse developmental outcomes. Affected parents often have difficulty with setting age-appropriate limits and demonstrate permissiveness alternating with excessive attempts to control their child's activities. They are also likely to seek higher levels of medical service and have more dissatisfaction with the medical care provided.

A child's perception of their own vulnerability does not necessarily correlate with their parents' perception. When children rate themselves as

vulnerable, most parents agree; when parents rate their children as vulnerable, only about half the children agreed.

Effective communication is at the heart of the preventing and managing VCS. If the child stays healthy and communication between parents and pediatricians is effective, VCS does not necessarily persist into school age. Persistent parental fear that a child remains vulnerable can be elicited by asking, "has your child ever been seriously ill?" VCS is likely when parents then respond by relating traumatic accounts of their child's perceived near death experiences. The physician's awareness of a parent with excessive anxiety or guilt, and of their own anxiety, can decrease the risk of unintentionally exacerbating parental anxiety and belief in their child's vulnerability when discussing diagnosis, prognosis, and anticipated treatment. The parents' frightening experience should be acknowledged, and parents should be provided a more accurate assessment of their child's health and told when time-limited medical problems are resolved and will not recur. The importance of encouraging appropriate autonomy in the child should be emphasized with discussion of potential adverse effects of excessive parental concern on a child's development and behavior. Reassurance and regular follow-up help to distinguish normal development and behavior from problems needing psychiatric consultation.

TRANSITION

As children with chronic and severe medical problems are living longer into adulthood, the previously unanticipated problem of transitioning a pediatric patient to adult care must be addressed. Adolescence describes the time from childhood to adulthood, when the teen must change from pediatrician to internist and move from shared responsibility to individual responsibility. At some time adolescence ends and adulthood begins, even for patients with special health care needs.

Adolescence can be a difficult time, especially for those with chronic illness who face unique challenges in developmental tasks. Effective transfer to adult care involves coordination and good communication with the patient, family, and adult physicians so that the anxieties of all, including the pediatrician, are adequately addressed.

There are two different approaches for transfer of adolescents with chronic medical problems to adult medical care. Patients may be directly transferred to internists. This is typical for patients with pediatric onset of adult disease. Alternatively, patients may be transferred to a specialized center focusing on pediatric disorders in adults. They will then be seen in pediatric clinics with separate areas for young adults.

Transition is not the same as transfer of care. Transition is a multifaceted process that addresses the psychosocial, educational, and vocational needs of adolescents as they move from child-oriented to adult-oriented care. It is predicated on effective communication between

pediatrician and internist and between the patient, parents, and providers. Transition implies a guided educational and therapeutic process rather than an administrative event.

Pediatric care is typically family centered. The pediatrician looks to the parent for information and to provide instructions. Adults speak to their physicians directly. It is useful for the adolescent preparing for transition to begin talking to the pediatrician directly, without their parents' assistance.

Transitions of care are periods of vulnerability and potential risk. Adult providers are often unfamiliar with the management of rare and often complex chronic childhood conditions. Young adults who have survived with pediatric care are at risk of deterioration after they move on to adult providers. Those with the less severe and visible conditions may have the greater problems moving from pediatric to adult care. They may experience no preparation, with minimal often last minute planning, and poor communication and coordination of transfer of care.

It is important that pediatric relationships are appropriately terminated as part of the transition process. The patient's level of developmental maturity and ownership of their own care is the most prominent determinant of transition outcomes. In order for the process of transition to be effective, the patient needs to have demonstrated adequate self-management skills, long-term social support, self-reliance, autonomy, and independence. Noncompliance with treatment is particularly prevalent in adolescents. Effective management requires attention to psychological and social issues as well as medical factors.

A positive sense of the future is associated with a greater sense of control in adolescents. Both in turn support their emotional well-being and growth, increasing the likelihood of a more successful transition. When accomplished effectively, transition can provide the young adult with a sense of the future and of future responsibilities that continued pediatric care fails to reinforce. A carefully planned transition to adult health care should improve self-reliance, enhance autonomy and independence, and support young people in attaining their maximum potential and having meaningful adult lives.

Table 55.3–3.
Recommended Approach to Transition in Adolescence

Period	Task
Early adolescence Ages 10–13 years	Introduce concept of separate and confidential visits with MD. Straightforward reward-oriented approach. Begin discussion of transfer of care. Need for privacy begins.
Mid adolescence Ages 14–16 years	Begin discussion of health living and risky behaviors in nonjudgmental approach. Expect increased introspection. Begin visits to new setting and meeting new providers.

Late adolescence Ages 17–21 years	Increased self-responsibility. Transition to adult providers. Future-oriented approach. Progression of personal identity and code of ethics. Physician seen alone. Evolving role of social supports.

SOCIAL MEDIA

The Internet and social media sites such as Facebook, Twitter, and YouTube are a few of the new online tools available that provide useful applications to foster information gathering, development of user-generated content, social interaction, and real-time collaboration. These technologies encompass blogs and patient experience journals, social networks, video-sharing and photo-sharing sites, wikis, and myriad other media, and are pervasive around the world. A growing majority of patients, including children and adolescents, particularly those with chronic conditions, are seeking out social media and other online sources to acquire health information, connect with others affected by similar conditions, and play a more active role in their health care decisions. Many patients and families can develop large support networks through sharing and disseminating information about their child's illness on specialized hospital sites such as CaringBridge or CarePages. Internet support groups and interactive peer video games can decrease some of the isolation patients and families feel with rare or unknown conditions or due to necessary hospital medical isolation practices. Applications that monitor and track symptoms (e.g., adherence, mood, nutritional intake, chemotherapy side effects, and blood cell count) may be very useful as well as applications that facilitate delivery of stress management and relaxation techniques (e.g., biofeedback, breathing exercises, mindfulness, and imagery) are available though not yet widely disseminated.

These tools are not a substitute for clear communication, clinical, psychological, and psychiatric assessment and treatment. Practitioners, patients, and families need to consider threats to confidentiality, privacy, and professional and therapeutic boundaries when communicating via social media.

REFERENCES

Armstrong FD. Neurodevelopment and chronic illness: mechanisms of disease and treatment. *Ment Retard Dev Disabil Res Rev*. 2006;12:168–173.

Bedell JR, Giordani B, Amour JL, Tavormina J, Boll T. Life stress and the psychological and medical adjustment of chronically ill children. *J Psychosom Res*. 1977;21:237–242.

*Burns KH, Casey PH, Lyle RE, MacBird T, Fussell JJ, Robbins JM. Increasing prevalence of medically complex children in US hospitals. *Pediatrics*. 2010;126:638–646.

Chambers PL, Mahabee-Gittens EM, Leonard A. Vulnerable child syndrome, parental perception

of child vulnerability, and emergency department usage. *Pediatr Emerg Care*. 2011;27:1009–1013.

Committee on Psychosocial Aspects of Child and Family Health. American Academy of Pediatrics. The new morbidity revisited: a renewed commitment to the psychosocial aspects of pediatric care. committee on psychosocial aspects of child and family health. *Pediatrics*. 2001;108:1227–1230.

*Cooley WC. Adolescent health care transition in transition. *JAMA Pediatr*. 2013;167(10):897–899.

Crowley R, Wolfe I, Lock K, McKee M. Improving transition between pediatric and adult healthcare: a systematic review. *Arch Dis Child*. 2011;96:548–553.

Fitzgerald M, Beggs S. The neurobiology of pain: developmental aspects. *Neuroscientist*. 2001;7:246–257.

Freud A. The role of bodily illness in the mental life of children. *Psychoanalytic Study of the Child*. 1952;7:69–81.

*Fritsch SL, ed. Interface between pediatrics and children's mental health. *Child Adolesc Clin North Am*. 2010;19:183–450.

George DR, Rovniak LS, Kraschnewski JL. Dangers and opportunities for social media in medicine. *Clin Obset Gynecol*. 2013;56:453–472.

Glazer JP, Hilden JM, Poltorak DY, eds. Pediatric Palliative Care. *Child Adolesc Clin North Am*. 2006;15:567–826.

Green M, Solnit AA. Reactions to the threatened loss of a child: a vulnerable child syndrome. *Pediatrics*. 1964;34:58–66.

Kellerman J, Zeltzer L, Ellenberg L, Dash J, Rigler D. Psychological effects of illness in adolescents: I. Anxiety, self-esteem, and perception of control. *J Pediatr*. 1980;97:126–131.

Koopman HM, Baars RM, Chaplin J, Zwinderman KH. Illness though the eyes of the child: the development of children's understanding of the causes of illness. *Patient Educ Couns*. 2004;55:363–370.

McKenna MP, D'Alessandro D. Social networks and the practice of medicine: harnessing powerful opportunities. *J Pediatr*. 2011;158:1–2.

Newacheck PW, Hafon N. Prevalence and impact of disabling chronic conditions in childhood. *Am J Public Health*. 1998;88:610–617.

O'Connor MJ, Rossano JW. Ventricular assist devices in children. *Curr Opin Cardiol*. 2014;29:113–121.

*O'Dougherty M, Brown RT. Chapter 12: The stress of childhood illness. In: L. Eugene Arnold, ed. *Childhood Stress*. New York: John Wiley and Sons, Inc.; 1990:325–350.

Owens PL, Thompson J, Elixhauser A, Ryan K. *Care of Children and Adolescents in U.S. Hospitals*. Rockville, MD: Agency for Healthcare Research and Quality; 2003.

Patterson J, Blum RW. Risk and resilience among children with disabilities. *Arch Pediatr Adolesc Med*. 1996;150:692–698.

Rollins J, Mahan C, Bolig R, Mahan CC, eds. *Meeting Children's Psychosocial Needs across the Health-Care Continuum Austin*. Austin, TX: PRO-ED, Inc.; 2005.

Sharpe D, Rossiter L. Siblings of children with a chronic illness: a meta-analysis. *J Pediatr Psychology*. 2002;27:699–710.

Simon TD, Berry J, Feudtner C, et al. Children with complex chronic conditions in inpatient hospital settings in the United States. *Pediatrics*. 2010;126:647–655.

*Snell C, DeMaso DR. Chapter 2: Adaptation and coping in chronic childhood physical illness. In: Shaw R, DeMaso D, eds. *Textbook of Pediatric Psychosomatic Medicine*. Arlington, VA: American Psychiatric Publishing, Inc.; 2010:21–31.

Stuber ML, Shemesh E, Saxe GN. Posttraumatic stress responses in children with life-threatening illnesses. *Child Adolesc Psychiatr Clin North Am*. 2003;12:195–209.

Thomasgard M, Metz WP. The vulnerable child syndrome revisited. *J Dev Behav Pediatr*. 1995;16:47–53.

*Thompson RH, Vernon DT. Research on children's behavior after hospitalization: a review and synthesis. *J Dev Behav Pediatr*. 1993;14:28–35.

Vermaes IP, van Susante AM, van Bakel HJ. Psychological functioning of siblings in families of children with chronic health conditions: a meta-analysis. *J Pediatr Psychol*. 2012;37:166–184.

Vitulano LA, Lewis M, eds. Psychologic aspects of chronic disease. *Child Adolesc Clin North Am*. 2003;12:389–598.

Wallender JL, Varni JW. Effects of pediatric chronic physical disorders on child and family adjustment. *J Child Psychology Psychiatry*. 1998;39:29–46.

Wamboldt MZ, Wamboldt FS. Role of the family in the onset and outcome of childhood disorders: selected research findings. *J Am Acad Child Adolesc Psychiatry*. 2000;39:1212–1219.

Weiner L, Zadeh S, Battles H, et al. Allowing adolescents and young adults to plan their end-of-life care. *Pediatrics*. 2102;139:897–905.

Zeltzer L, Kellerman J, Ellenberg L. Dash J, Rigler D. Psychological effects of illness in adolescents: II. Impact of illness in adolescents—crucial issues and coping styles. *J Pediatr*. 1980;97:132–138.

▲ 55.4 Psychiatric Sequelae of HIV and AIDS

MARK DEANTONIO, M.D.

In June 1981, the first cases of what was later called acquired immunodeficiency syndrome (AIDS) in the United States were reported in *Morbidity and Mortality Weekly Report*. Since that time, this syndrome has become one of the greatest public health challenges. In 2012, an estimated 35.3 million people were living with human immunodeficiency virus (HIV) worldwide. HIV continues disproportionately to affect certain geographic regions (e.g., sub-Saharan Africa and the Caribbean) and subpopulations (e.g., women in sub-Saharan Africa, men who have sex with men, injection-drug users, and sex workers). In the last decade, there have been significant changes in the understanding, epidemiology, and treatment of this disease in children and adolescents, modifying psychiatric sequelae. In the United States, the prevention of vertical transmission and the surveillance of blood make it rare for a child or adolescent younger than

15 years of age to become infected with HIV. However, adolescent sexual behavior and drug use, especially when combined with psychiatric disturbance, put individuals at increasing risk of HIV infection in adolescence and young adulthood. Psychiatrists in the United States might never encounter an HIV-infected child but have a great opportunity to decrease the risk of infection when psychiatrically treating adolescents who are at risk. Gaining an understanding of psychiatric sequelae of HIV and AIDS and the dynamics of adolescent risk behavior will allow the physician to craft effective interventions to prevent transmission in youth.

EPIDEMIOLOGY

In 2014, the Centers for Disease Control and Prevention (CDC) published the HIV Surveillance Report documenting the incidence and rate of HIV infection in the United States from 2008 to 2012. The incident of HIV infection remained unchanged during this period being between 37,000 and 38,000 per year. Over this 5-year period, the rates for persons aged 13 to 14, 20 to 24 and 25 to 29 increased. The rate for persons aged 35 to 39, 40 to 44, 45 to 49, 50 to 54, 55 to 59, 60 to 64, and 65 and older decreased. The rates remained stable for individuals aged 15 to 19 and 30 to 34. Due to HIV infection surveillance during pregnancy, HIV infection remains low in children (under 13 years old) in the United States with an unchanged rate of 0.5/100,000 during this period. In 2012, the highest rates were for persons aged 20 to 24 years and 25 to 29 years with rates of 36/100,000. Though the rates of adolescents aged 15 to 19 were unchanged (10/100,000), it is notable that behavior leading to HIV infection often begins years before seroconversion to HIV positive occurs. Therefore, the high infection rates noted in the 20- to 29-year-old group can be due to risk behaviors which started in adolescence.

From 2008 through 2012, the rate for Asians increased from 5.8/100,000 to 6.1/100,000. The rates for Black/African Americans decreased from 63.5/100,000 to 58.3/100,000 as wells as for Hispanic/Latino from 20.6/100,000 to 18.5/100,000. The rates for whites remained the same from 6.8/100,000 to 6.7/100,000. It is difficult to interpret rate changes in Native Americans or Native Hawaiians as overall incidence is low (less than 250 per year in each group). Over this 5-year period the rate for females decreased; the rate for males remained stable. In 2012, males accounted for 80 percent of diagnoses of HIV infection among adults and adolescents.

From 2008 through 2012, the annual number of diagnosed HIV infections attributed to male-to-male sexual contact increased among adult and adolescent males. The numbers of infections due to injection drug use, male-to-male sexual contact and injection drug use, or heterosexual contact decreased. Among adult and adolescent females, the numbers of infections due to injection drug use or heterosexual contact decreased. In 2012,

among adult and adolescent males and females, the diagnosed infections attributed to male-to-male sexual contact (67 percent) including male-to-male sexual contact and injection drug use and those due to heterosexual contact (26 percent) accounted for approximately 93 percent of HIV infections in the United States.

Geographically, from 2008 through 2102, there were not significant changes by regions in the United States. The rates in the Northeast, the South, and the West decreased, while the rates in the Midwest slightly increased. In 2012, the rates were 20.7/100,000 in the South, 16.2/100,000 in the Northeast, 11.5/100,000 in the West, and 9.2/100,000 in the Midwest.

The changing US epidemiology shows relative control of HIV infection in children under 13, but a significant rise in rates of individuals in their twenties. Sexual behavior specifically male-to-male sexual contact is the primary cause of infection with injection drug use and heterosexual contact decreasing as a cause. The behaviors that caused infection for individuals in their twenties often begin in adolescence. These trends highlight the importance of interventions to change sexual behavior in adolescence.

There is another way in which HIV infection affects children and adolescents: When their parents or family members are infected. The number of US youth affected by HIV in a family member is growing because the majority of infected women are mothers; 60 percent have children younger than 18 years of age, and 45 percent live with their children. Living with chronically ill parents often with comorbid substance abuse has significant consequences for the affected family members, with potential psychiatric and social sequelae.

ETIOLOGY

Psychiatric sequelae of HIV infection in children and adolescents are due to multiple and interacting etiologies. These include direct and indirect effects of HIV infection on the brain, stresses of living in families often affected by substance abuse, psychiatric disorders, chronic illness, and death. Adolescents who contract HIV infection have a high likelihood of pre-existing psychiatric diagnosis, cognitive deficits, histories of sexual or physical abuse, ongoing substance abuse, homelessness, and chaotic family dynamics. The addition of HIV infection to these adolescents adds a profound stress and challenge to an already overwhelming existence.

HIV is transmitted through blood-to-blood or semen-to-blood contact. The virus can also be transmitted by penile–vaginal intercourse, especially if other sexually transmitted diseases (STDs) coexist. Receptive anal intercourse and the use of contaminated syringes are the most common causes of HIV infection in the United States. HIV-infected women can transmit the virus during pregnancy, delivery, or through breast milk. It also can be transmitted through untested blood products. There are four

identified types of human retroviruses, which belong to two distinct groups: the transforming human T cell lymphotropic virus (HTLV) types I and II, and cytopathic HIV-1, which is more common in the United States, and HIV-2. HIV-1 is further divided into groups, classes or subtypes, and circulating recombinant forms, all of which indicate geographic origin.

Testing to diagnose HIV infection has evolved to provide quicker, more accurate, and more cost-effective results. Current testing approaches directly detect the virus or demonstrate antibody to the virus or a component, which appears 2 to 12 weeks after acute infection. A definitive diagnosis among infants younger than 15 months old cannot be made because the antibody crosses the maternal–fetal blood barrier, so all infants of HIV-infected mothers will have maternal HIV antibodies whether infected or not. Once an individual is diagnosed as infected, the level of plasma HIV RNA (viral load) and the absolute numbers of peripheral CD4+ T cells are used to monitor clinical status.

In 1985, when HIV testing first became available, the main goal of such testing was to protect the blood supply. At that time, professional opinion was divided regarding the value of HIV testing for individuals and whether HIV testing should be encouraged because no consensus existed regarding whether a positive test predicted transmission of the HIV virus to sex partners or from mother to infant. No effective treatment existed. In 1987, the implications of positive HIV serology became evident. HIV testing and counseling was recommended for all persons in high-risk groups and all persons seeking treatment for STDs. In 1995, after perinatal transmission of HIV was demonstrated to be substantially reduced by the administration of zidovudine to HIV-infected pregnant women and their newborns, it was recommended that all pregnant women be counseled and offered testing for HIV. In 2001, the CDC recommended that all pregnant women should have HIV screening as a part of routine prenatal care. Although the woman needed to be aware of what they were being tested for, it was also felt that the previously recommended pretest counseling was a barrier to testing and should be simplified. At this time, the CDC recommended that HIV testing be offered routinely to all patients in high–HIV-prevalence health care settings, despite risk factors. In low-prevalence settings, targeted HIV testing on the basis of risk screening continues to be recommended.

In September 2006, the CDC again revised recommendations for HIV testing of adults, adolescents, and pregnant women. Due to the changes in demographic of the HIV/AIDS epidemic in the United States, increasing proportions of infected persons are younger than 20 years of age, women, members of racial or ethnic minority populations, persons who reside outside metropolitan areas, and heterosexual men and women who frequently are unaware that they are at risk of HIV infection. As a result, the effectiveness of using risk-based testing to identify HIV-infected persons has diminished. In addition, prevention strategies that incorporate

universal HIV screening have been highly effective (blood donors and pregnant women). Perinatal transmission rates have been reduced to less than 2 percent due to prophylactic administration of antiretroviral drugs, scheduled cesarean delivery, and avoidance of breast-feeding. Studies document that the majority of adults who are aware of their HIV infections substantially reduce sexual behaviors that might transmit HIV after they become aware that they are infected. Therefore, to increase the diagnosis of HIV infection, destigmatize the testing process, link clinical care with prevention, and ensure immediate access to clinical care for persons with newly identified HIV infection, the CDC now recommends the adoption of routine HIV testing in all health care settings for individuals from the age of 13 to 64 years. The testing should be undertaken only with the patient's knowledge and understanding that HIV testing is planned. Patients can refuse testing. Persons likely to be at high risk of HIV infection—injection drug users and their sex partners, sex workers, partners of HIV-infected persons, and persons who have had or whose sex partners have had more than one sex partner since their most recent HIV test—all should be tested annually. This significant change in testing recommendation reflects the ever-changing nature and better understanding of HIV/AIDS.

Pre-exposure prophylaxis or PrEP is a new tool in HIV prevention. It is a prevention option for individuals who are at high risk of getting HIV. It is a pill with two antiviral medications (tenofovir and emtricitabine) which are taken daily. While PrEP has been shown to be safe and efficacious for adults, the data on its safety and efficacy for adolescents are not clear. The CDC to date has not come out with a definitive recommendation, stating that the risks and benefits of PrEP for adolescents should be weighed carefully in the context of local laws and regulations about autonomy in health care decision making by minors.

Disclosure of a child's or adolescent's HIV status brings up complex issues requiring careful assessment of an individual's circumstance. Knowledge of one's HIV status is associated with emotional distress and increased rates of depression. The level of distress depends on how the disclosure occurs, under what circumstances, and the reaction of the people involved. Telling a child that he or she has HIV appears to be influenced by several factors, including older child age, higher child cognitive functioning, recent major life events, and higher family expressiveness. Many clinicians and caregivers are ambivalent about disclosing to children because of fears that the children will inappropriately disclose to the community at large and face stigmatization as a consequence. Current pediatric practice guidelines encourage disclosure by adolescence, but neurocognitive deterioration may complicate timing. Clearly it is critical that HIV status be disclosed to adolescents before they become sexually active. Simply telling adolescents that they are HIV positive may not reduce risky sexual behavior or disclosure of that status to their sexual partners.

This potential lack of disclosure underscores the importance of emphasizing safer sex practices, including condom use, for all adolescents regardless of HIV status.

Neurodevelopmental Sequelae of HIV Infection

HIV is found primarily in microglia and macrophages and not directly in neurons. The virus is thought to have neurotoxic components and to induce a variety of inflammatory factors that cause neuronal cell death. The clinical presentation of HIV neurological involvement can vary from mild neurocognitive impairment to static encephalopathy, progressive encephalopathy, and dementia. There has been more progress in understanding the effects of HIV on the brain in adults than in children, owing to the dramatic decrease in childhood infections. Studies of infected patients 13 to 18 years of age are limited to case reports. In the earliest asymptomatic stages, abnormalities may show up in cognitive abilities evaluated during neuropsychiatric testing but not necessarily as impairments in functioning. In children and adolescents, those infected also likely have cognitive deficits due to in utero drug and alcohol exposure, attention-deficit disorder, and learning disabilities, all of which increases maternal HIV infection risk and complicate neuropsychiatric testing. The clinical presentation of HIV central nervous system (CNS) involvement may vary with age of infection and mode of transmission. In children or adolescents infected through blood transfusion, encephalopathy generally presents later and follows a more indolent course. By contrast, in almost one-fourth of vertically HIV-infected children encephalopathy can present and contribute significant morality and mortality as an AIDS-defining illness even when the child is not severely immunocompromised. Early progressive encephalopathy is usually seen in young, untreated infants with high viral loads and rapid clinical deterioration. Vertically HIV-infected children who are labeled “slow progressors”—those with lower viral loads and higher CD4 counts—often exhibit no cognitive deficits throughout childhood and adolescence.

The most common computed tomographic findings associated with encephalopathy are basal ganglia calcifications, brain atrophy with enlarged ventricles, and enlarged cortical sulci. HIV-infected children have a higher incidence of ischemic strokes and cerebral artery aneurysms, which can rupture. Magnetic resonance imaging (MRI) has been recommended for infected children who present with a lack of neurocognitive progression or dysfunction even though they are otherwise asymptomatic.

In adolescents with advanced AIDS, a gradual progression into dementia is seen. It appears as a subcortical dementia process with motor impairment, cognitive disturbance, and a predominance of apathy over depression. Psychotic symptoms may be present. In addition, opportunistic

CNS infections such as toxoplasmosis, progressive multifocal leukoencephalopathy, and CNS lymphoma can complicate the CNS clinical picture. There is no specific treatment for HIV involvement of the CNS. When progressive encephalopathy is present in children and adolescents, increased doses of antiviral agents are used but there have not been studies documenting efficacy in children and adolescents.

CLINICAL FEATURES

Despite more than 30 years of studying HIV/AIDS in children and adolescents, there have been few studies of psychiatric diagnoses in these individuals. Although there is clearly defined neurological pathology due to HIV infection, psychiatric illness specifically due to HIV infection has not been quantified. Instead psychiatric diagnoses are related to inheritable psychiatric disorders and reactions to having a chronic illness in the context of often chaotic and traumatic environments.

Early studies of emotional and behavioral disorders in perinatally HIV-infected children showed high rates of anxiety, depression, adjustment problems, and attention-deficit/hyperactivity disorder. However, results among studies were inconsistent and failed to use appropriate control groups. Studies of behavioral disorder in HIV-infected children that included uninfected control groups of similar background generally reported high rates of emotional and behavioral disorder in both groups. One controlled study found elevated rates of attention-deficit/hyperactivity disorder in a sample of prenatally drug-exposed children, with no difference between HIV-infected and uninfected children. Another controlled study found elevated rates of internalizing and externalizing symptoms in HIV-infected children and demographically matched control. However, rates of symptomatology were actually higher in the control group. The Woman and Infant Transmission Study Group Study compared HIV-infected children with uninfected siblings born to HIV-positive mothers. Findings confirmed earlier data revealing high rates of hyperactivity and impulsivity in both the infected and uninfected children. The Pediatric AIDS Clinic Trials Group reported a higher frequency of psychiatric hospitalization among perinatally infected children than in the general pediatric population and HIV-uninfected members of the cohort. The majority of patients were admitted because of depression and a minority due to behavioral disorders. Studies of vertically infected adolescents and psychiatric illness are limited to case reports. The group whose AIDS progresses significantly in adolescence exhibits a dementia process, which may include psychosis and agitation. The group of vertically infected adolescents who are labeled “slow progressors” due to their mild degree of immune suppression appear to present with depression, anxiety, and chronic adjustment disorders related to chronic illness.

Adolescents who become infected as adolescents due to sexual or

injection substance abuse have not had their psychiatric diagnoses well studied. There is evidence that having psychiatric illness in adolescence, especially when coupled with substance abuse and cognitive and learning difficulties, increases the high-risk behaviors that can lead to HIV infection. This was most recently documented in a study in 2010. It has been documented that a significant number of HIV-infected adolescents have depressive disorders. Conduct disorder and substance abuse are also present to a significant degree. The majority of these disorders predated their HIV infection. There is evidence in the adult and adolescent literature that having depression while being HIV infected can alter the function of killer lymphocytes. This causes an increase in activated CD8 T lymphocytes and can increase viral load. Treatment with selective serotonin reuptake inhibitors and decrease in depressive symptoms reverse these changes. Clearly, more study needs to be made regarding psychiatric sequelae of adolescent-onset HIV infection.

TREATMENT

Psychiatrists play an important role in the treatment of children and adolescents who are HIV infected. HIV infection has changed over the years to a chronic illness in the United States, which requires multiple layers of care in a variety of settings over many years. Although pediatricians have expertise in managing developmental issues in chronic illness in youth, psychiatrists have a system-oriented approach arising from their work with patients who depend on the support systems provided by family, school, and community. Parents of HIV-infected children and adolescents have a high rate of psychiatric illness, substance abuse, and chaotic and dysfunctional family interactions. Psychiatry brings expertise to the treatment team in crafting interventions to promote family stability and family support of the patient's medical treatment. Finally, a significant number of children and adolescents who are HIV infected have pre-existing psychiatric illness, which requires psychiatric expertise to manage. In children and adolescents receiving antiretroviral treatment, lack of almost perfect adherence to daily regimes continues to be the greatest barrier to sustaining a clinically effective response. Lack of adherence can also lead to the development of HIV strains resistant to the available antiretrovirals. Although adherence issues are common with all chronic illnesses in children and adolescents, it is more of an issue in HIV-infected children and adolescents living in situations in which parents may be physically or mentally ill or with children and adolescents in a foster care or residential setting with inadequate supervision. The psychiatrist has a critical role in working with the treatment team, the child or adolescent, and caregivers to problem solve and intervene around issues of adherence.

The HIV epidemic adds to the challenges of psychiatric treatment of children and adolescents in inpatient and residential settings. The

American Academy of Child and Adolescent Psychiatry has the following recommendations concerning HIV and inpatient psychiatric treatment:

- HIV infection or risk of HIV infection is not a reason to deny a child admission to a psychiatric hospital.
- All patients admitted to a psychiatric hospital should be treated as if they were HIV infected. Even if all patients were tested on admission for antibodies to HIV, one is still not sure of infection status due to the potential latency period of antibody response documenting infection.
- Universal precautions should be employed for any secretions, including blood from any psychiatric inpatient. There is no need for a change in the management of an HIV-infected child if universal precautions are used for all inpatients.
- HIV positive or potentially positive patients do not require individual rooms or toilet facilities. HIV infected, or potentially infected, patients do not need to be isolated in regard to the milieu or peer groups. Such patients should participate in all aspects of the inpatient treatment program.
- The Academy agrees with the 2006 CDC recommendation of the adoption of routine HIV testing in all health care settings from the age of 13 to 64 years. The testing should be undertaken only with the patient's knowledge and understanding that HIV testing is planned. Patients can refuse testing. Persons likely to be at high risk of HIV infection, including injection drug users and their partners, sex workers, partners of HIV-infected persons, and persons who themselves or whose sex partners have had more than one sex partner since their most recent HIV test all should be tested annually. The Child and Adolescent Psychiatrist should balance the adolescent's request for privacy with the families wishes. There are specific laws regarding confidentiality of HIV antibody testing results and who authorizes consent in minors. These laws vary from state to state. Prior to testing, ensure the laws are understood.
- Results of patients' HIV antibody tests will be maintained in a confidential manner. A patient's HIV status should be shared only with those staff who need to know the status to appropriately care for the child while an inpatient. Sharing of HIV status should, of course, be in compliance with applicable state and federal law.

There is little published research on the psychopharmacological treatment of HIV-infected children and adolescents. The Pediatric AIDS Clinical Trials Group is beginning a study of treatment of attention-deficit/hyperactivity disorder in HIV-infected children, with no published results as of the time of this writing. All published studies are case reports, which are difficult to generalize. Interactions between psychotropic and antiretroviral medications are common in the HIV-infected child and adolescent. Most medications used in the treatment of psychiatric

disorders are metabolized by the cytochrome P450 (CYP) enzymes, especially the subgroups 3A4 and 2D6. Of the antiretrovirals, the protease inhibitors tend to pose the greatest risk of changes in drug levels that are significant in the HIV patient. Ritonavir is the most potent of the protease inhibitors, with the greatest effect on the inhibition of 3A4 and to a lesser extent 2D6. Enfuvirtide, a fusion inhibitor, is not an inhibitor of CYP enzymes and does not alter the metabolism of CYP 3A4, 2D6, 2C19, or 2EI substrates. Therefore, it has no significant effect on blood levels of psychiatric medications or other antiretrovirals. Efavirenz has been shown in vivo to induce CYP 3A4. Other compounds that are substrates of CYP 3A4 may have decreased plasma concentrations when coadministered with efavirenz. In vitro studies have demonstrated that efavirenz inhibits 2C9, 2C19, and 3A4 isozymes. Coadministration of efavirenz with drugs primarily metabolized by these isozymes may result in altered plasma concentrations of the coadministered drug. Drugs that induce CYP 3A4 activity would be expected to increase the clearance of efavirenz. Some antiretroviral drugs have been reported to cause psychiatric symptoms in HIV-infected youth, including mania with zidovudine and attention-deficit problems with efavirenz. It is important for psychiatrists to remain in contact with the other medical personnel providing treatment for the HIV-infected child or adolescent, maintain updated lists of all medication, and frequently consult updated web-based information sites regarding drug–drug interactions.

In terms of the strategy of psychopharmacological treatment, thoughtfulness is primary. The concern a physician has is that the psychiatric symptoms present are really due to the “organicity” of HIV infection and so should only be treated by maximizing antiretroviral treatment. Although this point of view needs to be considered, one can make psychiatric diagnoses and use psychopharmacology despite a possible underlying organicity caused by HIV. The most common disorders presenting in this population are attention-deficit/hyperactivity disorder, generalized anxiety disorder, major depressive disorder, bipolar affective disorder, psychotic disorder not otherwise specified, and schizophrenia. One should use the same medications with same dose titration as one would with an uninfected child or adolescent. The exception would be to start and maintain lower doses of medication in the patients with clear demented or encephalopathy. Stimulants in general have few drug–drug interactions, and lithium has none because it is excreted renally. In adults it has been reported that the HIV virus can target the destruction of cells in the basal ganglia, which can sensitize an individual to extrapyramidal side effects when he or she uses an antipsychotic medication. This observation has not been reported in children and adolescents, but because children and adolescents overall are more sensitive to extrapyramidal side effects, this should be of concern. Therefore, one should first use atypical

antipsychotics when treating children who have psychosis and adolescents who are HIV infected.

No specific psychotherapeutic interventions had been documented as being efficacious to treat psychiatric disorders in HIV-infected children and adolescents. However, there are studies showing efficacy of insight-oriented, cognitive behavioral, group therapies in HIV-infected adults with depression. Such treatments also document improved adherence to medication regimes. It is likely that adolescents would have a similar response to psychotherapies for depression. However, individual, family, and group interventions are an important adjunct to the successful treatment of HIV-infected youth. For infected adolescents, prevention programs emphasize reducing transmission acts, enhancing quality of life, increasing self-care behaviors, improving medical adherence, coping with medication side effects, dealing with appropriate disclosure of HIV status, and interacting with health care providers. Preventive interventions for youth living with HIV using therapeutic approaches based on cognitive behavioral principles have been studied. One intervention included two primary components: (1) increasing positive health behaviors, particularly around medication adherence and (2) enhancing altruistic motivation to reduce transmission acts. The intervention emphasized coping with being HIV positive, adopting healthy routines, partaking in health care decisions, reducing substance use and risky sexual behavior by identifying risk triggers, and increasing self-efficacy to use condoms and negotiate safe sex. The intervention sought to build skills and improve emotion regulation. These are similar to the strategies used to enhance psychological functioning for youth in psychiatric care. Adolescents who received the intervention reported more lifestyle changes, improved coping, fewer sexual partners, and a lower percentage of unprotected sexual acts than did youth who did not receive the intervention. Additional sessions emphasized increasing life satisfaction and emotional strength. The additional sessions reduced emotional distress and psychiatric symptoms among the infected youth, including somatization and anxiety. Programs have also been developed to help the adult caregivers' of children and adolescents with HIV infection. These programs teach effective parenting skills, parental monitoring of medication adherence, and increasing social support among family members and the community. Some of the more innovative interventions also address the caregivers' mental health and quality of life.

Despite potential long-term intergenerational consequences of HIV illness, few services are available to assist children and adolescents of HIV-infected family members. The few programs with published data have yielded positive effects for affected adolescents. These programs address coping with stigma, dealing with illness-related tasks, dealing with parental disclosure, increasing maternal well-being and social support, developing effective parenting skills, decreasing victimization, custody planning,

maintaining positive family routines, and handling emotional reactions to AIDS. Two- and 4-year follow-up studies of intervention effects have been both promising and disappointing. Two years after the intervention, treated adolescents reported significantly fewer conduct problems, less emotional distress, lower levels of family-related stress, and higher self-esteem than adolescents in the control group. Outcomes were comparable for youths who were bereaved and nonbereaved by parental AIDS and across ethnic and gender subgroups. Unfortunately, the positive effects eroded by the 4-year follow-up, underscoring the need for sustained support and treatment for these families. HIV prevention programs for affected children and adolescents must address intergenerational imitation of HIV risk behaviors and substance abuse and help them to cope with HIV-related psychosocial challenges such as bereavement. Unfortunately, psychosocial interventions for the affected have not been broadly implemented or prioritized.

CONCLUSION

Child and adolescent psychiatry, with its developmental and systems approach, has a unique position from which to manage the mental health sequelae of HIV infection in children and adolescents. Although this epidemic has created challenges, it has also given opportunities. The epidemic has forced a re-evaluation of child and adolescent sexual development and a more comprehensive understanding of the development of sexual orientation. It has prioritized the need for the assessment of sexual and drug-taking behavior as part of an adolescent psychiatric assessment. Finally, the AIDS epidemic demands that psychiatrists constantly monitor high-risk behaviors in adolescent patients and implement strategies to reduce these behaviors. Such interventions will improve the quality of life and prognosis of adolescent psychiatric patients regardless of HIV infection status.

REFERENCES

American Academy of Child and Adolescent Psychiatry. HIV and psychiatric hospitalization of children and adolescents. 2009. Available at http://www.aacap.org/cs/root/policy_statements/hiv_and_psychiatric_hospitalization_of_ch

American Academy of Child and Adolescent Psychiatry. Adolescents, HIV and substance abuse in treatment programs. 2009.

Bekker L, Hosek S. HIV and adolescents: focus on young key populations. *Journal Int AIDS Soc*. 2015;18(2Suppl 1):1–94.

*Benton T. The immunology of adolescent HIV: Current evidence and new directions. The Scientific Proceedings of the American Academy. *J Child Adolesc Psychiatry*. 2006;47.

Branson BM, Handsfield HH, Lampe MA, Janssen RS, Taylor AW. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. *MMWR Morb Mortal Wkly Rep*. 2006;55(14):1.

*Brown L, Hadley W, Stewart A, Lescano C, Whiteley L. Psychiatric disorders and sexual risk among adolescents in mental health treatment. *J Consult Clin Psychol*. 2010;78(4):590.

Brown LK, Kennard BD, Emslie GJ, et al. Effective Treatment of Depressive Disorders in Medical Clinics for Adolescents and Young Adults living with HIV. *J Acquir Immune Defic Syndr*. 2016;71(1):38–46.

Brown LK, Whiteley L, Harper GW, Nichols S, Nieves A. Psychological symptoms among 2032 youth living with HIV: a multisite study. *AIDS Patient Care and STDs*. 2015;29(4):212–219.

Centers for Disease Control and Prevention. The global HIV/AIDS pandemic: 2006. *MMWR Morb Mortal Wkly Rep*. 2006;55(31):841.

Centers for Disease Control and Prevention. HIV Surveillance Report, 2012;24. Available at http://www.cdc.gov/hiv/library/reports/surveillance/.

*Centers for Disease Control and Prevention, Health Resources and Services Administration, National Institutes of Health, American Academy of HIV Medicine, Association of Nurses in AIDS Care, International Association of Providers of AIDS Care, National Minority AIDS Council, and Urban Coalition for HIV/AIDS Prevention Services. Recommendations for HIV Prevention with Adults and Adolescents with HIV in the United States, 2014: Summary for Clinical Providers. 2014. Available at http://stacks.cdc.gov/view/cdc/26063.

Donenberg GR, Emerson E, Bryant F, Wilson H, Weber-Shifrin E. Understanding AIDS-risk behavior among adolescents in psychiatric care: links to psychopathology and peer relationships. *J Am Acad Child Adolesc Psychiatry*. 2001;40(6):642.

Dwight LE, Thomas RTH, Douglas SD, Gettes DR, Morrison M. Association of depression with viral load, CD8 lymphocytes, and natural killer cells in women with HIV infection. *Am J Psychiatry*. 2002;159:1752.

Forstein M, Cournos F, Douaihy A, Goodkin K, Wainber ML. *Guideline Watch: Practice Guideline for the Treatment of Patients with HIV/AIDS*. Washington, DC: American Psychiatric Association; 2006.

Garofalo R, Kuhns LM, Hotton A, Johnson A, Muldoon, A, Rice D. A randomized controlled trial of personalized text message reminders to promote medication adherence among HIV-positive adolescents and young adults. *AIDS and Behavior*. 2015;20(5):1049–1059.

Gaughan DM, Hughes MD, Oleske JM, Malee K, Gore CA. Psychiatric hospitalizations among children and youths with human immunodeficiency virus infection. *Pediatrics*. 2004;113:544.

Idele P, Gillepie A, Porth T, Suzuki C, Maby M. Epidemiology of HIV and AIDS among adolescents: current status, inequities, and data gaps. *J Acquir Immune Defic Syndr*. 2014;66(2):S144–S153.

*Lester P, Chesney M, Cooke M, Weiss R, Whalley P. When the times comes to talk about HIV: factors associated with diagnostic disclosure and emotional distress in HIV-infected children. *J Acquire Immune Defic Syndr*. 2002;31:309.

Lyon ME, Trexler C, Akpan-Townsend C, Pao M, Selden K. A family group approach to increasing adherence to therapy in HIV-infected youths: results of a pilot project. *AIDS Patient Care STDS*. 2003;17:299.

Maartens G, Celum C, Lewin S. HIV infection: epidemiology, pathogenesis, treatment and prevention. *Lancet*. 2014;384:258.

Mirani G, Williams PL, Chernoff M, et al. Changing trends in complications and mortality rates

among US youth and young adults with HIV infection in the era of combination antiretroviral therapy. *Clin Infect Dis*. 2015;61(12):1850–1861.

Pelton J, Forehand R. Orphans of the AIDS epidemic: an examination of clinical level problems of children. *J Am Acad Child Adolesc Psychiatry*. 2005;44(6):585.

Pilowsky DJ, Zybert PA, Hsieh P, Vlahov D, Susser E. Children of HIV-positive drug-using parents. *J Am Acad Child Adolesc Psychiatry*. 2003;42(8):950.

Pilowsky DJ, Zybert PA, Vlahov D. Resilient children of injection drug users. *J Am Acad Child Adolesc Psychiatry*. 2004;43(11):1372.

Rotheram-Borus MJ, Lee MB, Gwadz M, Draimin B. An intervention for parents with AIDS and their adolescent children. *Am J Public Health*. 2001;91:1294.

*Rotheram-Borus MJ, Lee MB, Murphy DA, Futlerman D, Duan N. Teens Linked to Care Consortium: efficacy of a preventive intervention for youths living with HIV. *Am J Public Health*. 2001;91:400.

Rotheram-Borus MJ, Murphy DA, Wight RG, Lee MB, Lightfoot M. Improving the quality of life among young people living with HIV. *Eval Prog Plan*. 2001;24:227.

Rotheram-Borus MJ, Stein JA. Problem behaviors among adolescent whose parents are living with AIDS. *Am J Orthopsychiatry*. 1999;69:228.

Thurstone C, Riggs P, Klein C, Mikulich-Gilbertson S. A one-session human immunodeficiency virus risk-reduction intervention in adolescents with psychiatric and substance use disorders. *J Am Acad Child Adolesc Psychiatry*. 2007;46(9):1179.

Wolbaci-Lowes J, Jed SL, Johnson SC, Fletcher CV, Brown B. *A Pharmacist's Guide to Antiretroviral Medications for HIV-Infected Adults and Adolescents*. Mountain Plains AIDS Education and Training Center, University of Colorado at Denver, and Health Sciences Center. Available at http://www.aidsetc.org/pdf/p02-et/et-03-00/pharm2005.

▲ 55.5 Adolescent Substance Use Disorders

SEAN R. LENOUE, M.D., AND PAULA D. RIGGS, M.D.

INTRODUCTION

Substance use disorders (SUDs) account for about 6 percent of all deaths worldwide and cost an estimated $700 billion in the United States, alone. The vast majority of adults (approximately 80 percent) with chronic addiction started using drugs or alcohol as teenagers. Although the majority of adolescents in the United States report experimenting with drugs or alcohol before graduating from high school, most do not progress to a SUD. The estimated risk of developing an SUD by the age of 18 is 23.8 percent. Risk of progression to an SUD is increased in those who have pre-existing disruptive behavior disorders or other mental health problems, academic problems, and/or have experienced chronic adversities or maltreatment during childhood. Adolescent-onset SUDs are more likely to progress to chronic addiction in adulthood and also increase the risk of

developing mood and anxiety disorders as well as antisocial personality disorder.

At-risk youth can be identified in children as young as 3 to 5 years of age, often presenting with impulsivity, aggression, hyperactivity, inattention, poor persistence, and emotional lability. This constellation of symptoms, characterized as "behavioral dysregulation," is highly heritable and a strong family history of SUD is often present. Such children often continue to have poor self-control, develop disruptive behavioral disorders (oppositional defiant disorder [ODD], attention-deficit/hyperactivity disorder [ADHD], conduct disorder [CD]) in later childhood, and are at high risk of developing SUDs during adolescence. It is important for school personnel, pediatric primary care, and mental health clinicians to screen children for risk factors and refer for evaluation and interventions that may prevent or reduce their risk of early-onset SUD. It is also important to highlight that repeated use of drugs or alcohol can alter brain function in ways that reinforce or promote continued use and development of SUD even in youth without known risk factors. For example, youth with chronic painful medical conditions, acute or prolonged injuries, or multiple medical or dental surgical procedures may be at risk of developing addiction to prescription opioid pain medications even in the absence of pre-existing apparent genetic or environmental risk factors.

Research advances over the past two decades have led to the understanding that addiction as a neurobiologically based medical illness. Initial exposure to substances of abuse increases firing in the dopaminergic pathways of the ventral striatum, producing a subjective "high" that promotes further drug use. However, with prolonged and repeated drug use, dopaminergic neurotransmission gradually diminishes, producing less intense "highs," greater post-use dysphoria, and intense drug craving precipitated by substance-specific withdrawal symptoms. Drug craving can also be triggered by environmental "cues" (i.e., conditioned stimuli) that have frequently been paired with substance use (i.e., conditioned response). At this stage, neural circuitry in the dorsal striatum, rather than the ventral striatum, drives compelling mental preoccupation, decision making, and drug-seeking behaviors that promote habitual use making it much more difficult to consider reducing or discontinuing use.

Greater understanding of the neurobiology of addiction as well as genetic and environmental contributions to adolescent-onset SUD have informed the development of more effective substance prevention and treatment. This chapter focuses on specific aspects of the etiology, epidemiology, prevention, and treatment relevant to adolescent SUDs.

EPIDEMIOLOGY

Prevalence of Use

Substances that can be legally used by adults (e.g., tobacco and alcohol) have historically been those most accessible and commonly used by children and adolescents. For example, inhalants found in many household products (e.g., spray paint and gasoline) and over-the-counter "cold medications" containing dextromethorphan are most easily accessible and commonly abused by children and younger adolescents. Alcohol continues to be the most commonly abused substance by adolescents. However, both alcohol use and cigarette smoking have gradually declined over the past decade and were at the lowest levels in recorded history of the Monitoring the Future (MTF) Survey in 2014. While this is good news, it is important to recognize that despite the decline in alcohol use, almost a quarter of high school seniors reported at least one episode of binge drinking (5 or more drinks on one occasion) in the past 2 weeks in 2014. Binge drinking is especially risky behavior given the potential for serious negative consequences (e.g., motor vehicle accidents, and risky sexual behaviors). Similarly, the decline in cigarette smoking appears to be offset by a significant increase in the use of e-cigarettes among teenagers, the prevalence of which currently exceeds tobacco smoking. Also in 2014, the prevalence of marijuana use surpassed tobacco use among high school students, which is unprecedented in the history of the survey. Increased marijuana use among high school students is likely underpinned by growing national acceptance and the increasing number of states that have legalized medical and/or recreational marijuana use in the United States. Recent national surveys indicate that more than half of Americans (53 percent) favor marijuana legalization. Approximately half of Americans say that they have tried marijuana and, in 2014, more than 2 million reported using marijuana for the first time. Almost two-thirds (64 percent) of high school seniors believe that little harm comes from regularly using marijuana and both regular and daily marijuana use are at 30-year peak levels. This is particularly concerning from a public health standpoint in light of a growing body of research indicating that regular marijuana use during adolescence—a time of rapid brain development—increases the risk of psychosis and is associated with reduction in adult IQ and persistent neurocognitive deficits that may not be fully reversible even with abstinence. Regular use of marijuana before age 17 is also associated with poorer academic achievement, greater risk of high school dropout, as well as unemployment or underemployment in adulthood.

The prevalence of several other illicit drugs including cocaine, methamphetamine, and ecstasy decreased somewhat in 2014 (6 percent of 8th graders, 11 percent of 10th graders, and 16 percent of 12th graders). However, nonmedical use of prescription opiates and heroin use have become a national epidemic and public health crisis in the United States. According to the Centers for Disease Control and Prevention (CDC), deaths related to heroin overdose quadrupled between 2002 and 2013 in the

United States and more than doubled among young adults ages 18 to 25. Approximately 80 percent of young adult heroin users report addiction to prescription opiates before transitioning to injection heroin use. In 2013, almost 2 percent (1.7 percent) of adolescents ages 12 to 17 reported nonmedical use of prescription pain relievers. Between 2012 and 2014 the annual prevalence of heroin use for 8th, 10th, and 12th graders ranged from 0.2 to 0.5 percent. Taken together, these trends indicate the need for more effective substance screening, prevention, and earlier treatment interventions for adolescents.

CLINICAL EVALUATION

Screening

Primary care clinicians can play an important role in identifying youth who are likely to develop or have substance problems as well as motivating and encouraging behavioral changes. Fulfilling this potential requires effective screening and behavioral counseling interventions (BCIs) that are brief and feasible to implement in busy primary care practice settings. Screening, Brief Intervention, Referral and Treatment (SBIRT) initiatives and programs have been developed and adopted in many states to encourage primary care providers to utilize standardized screening procedures and assessment instruments. A number of research-based substance screening and assessment instrument tools have also been developed for use in primary care medical settings to assist clinicians in identifying high-risk youth and/or those who have started using or abusing substances (e.g., CRAFFT and Project ASSERT) A Rapid Teen Substance Use Screening Tool for Clinicians (S2BI). The S2BI tool may be administered by a pediatrician or filled out by the teen on a computer or tablet. Teens are asked to disclose any use of tobacco, alcohol, or marijuana during the past year, and any positive responders are then asked the same multiple choice questions about four additional categories of drugs (non-medical use of prescription drugs, illicit drugs, herbs/synthetic drugs, inhalants) and questions about frequency of use (e.g., “once or twice,” monthly, and weekly).

Most adolescent drug and alcohol screening assessments have been developed for use in primary care medical settings. Additional research is needed to evaluate the effectiveness of existing screening tools in other settings (e.g., mental health; school-based). Moreover, none of the existing drug or alcohol screening assessment instruments for children or adolescents meets sufficiently rigorous evidence-based criteria to be recommended by the United States Prevention Services Task Force (USPSTF). This is important and relevant because the Affordable Care Act (ACA) of 2010, recognizing the independence and methodological rigor of the Task Force, requires insurers to provide first dollar reimbursement only for screening or BCIs receiving a grade of “A” or “B” by the Task Force. Of

the 11 Task Force recommendations (to date) the following are related to substance screening and BCIs for children and adolescents: (1) the Task Force recommends that primary care clinicians provide education and/or brief counseling to prevent tobacco use in school-aged children and adolescents; (2) the Task Force determined that there is currently insufficient evidence to assess the balance of benefits and harms of screening and counseling for alcohol misuse, illicit drug, and nonmedical pharmaceutical use in youth less than 18 years of age in primary care medical settings. These recommendations apply to children and adolescents who have not already been diagnosed with an SUD.

Prevention

The Institute of Medicine (IOM) recognizes three categories of prevention interventions: (i) universal prevention targets the population as a whole; (ii) selective prevention targets high-risk groups of individuals; and (iii) indicated prevention targets individuals who manifest risk behaviors or early signs and symptoms of the illness. According to the most recent Cochrane review only the following three interventions were determined to have sufficient evidence to be deemed efficacious or effective: (i) The Unplugged Program (directed toward 12 to 14 years old), (ii) Life Skills Training Program (directed toward 13 to 17 years old), and (iii) Good Behavior Game (directed toward 6 to 7 years old). All are universal, school-based prevention interventions that are generally adopted school-wide and delivered by teachers trained in the intervention and delivered in the classroom setting. The Cochrane review did not identify any selective or indicated interventions or non–school-based/multi-component community-based prevention interventions meeting rigorous criteria for effectiveness. However, recent studies, not included in the Cochrane review, but meeting similarly rigorous criteria support the efficacy of at least one selective school-based prevention intervention known as the Project Toward No Drug Abuse (directed toward 14 to 17 years old), and one indicated school-based prevention intervention known as the Teen Marijuana Check-Up (TMCU).

Treatment

Most psychosocial substance treatment interventions address drug-taking in general. However, treatment of withdrawal syndromes or other medications used in substance treatment are generally substance-specific or used to treat co-occurring psychiatric disorders (e.g., ADHD and mood disorders). Compared to numerous studies in adults, relatively few randomized controlled trials have evaluated the safety and efficacy of medications for substance abuse or co-occurring psychiatric disorders in adolescents (shown in Tables 55.5–1 and 55.5–2).

Evidence-based psychosocial treatments of adolescents with SUDs supported by at least one randomized controlled trial include Motivational Enhancement Therapy/Cognitive Behavioral Therapy (MET/CBT) and MET/CBT + Contingency Management (CM) (both group and individual format) and family-based interventions (Multi-Systemic Therapy [MST], Functional Family Therapy [FFT], Multidimensional Family Therapy [MDFT], Brief Strategic Family Therapy [BSFT]), and Adolescent Community Reinforcement Approach (ACRA). Overall, these interventions show fairly comparable but modest treatment effect sizes on reductions in substance use but relatively low rates of abstinence ranging from less than 20 percent (most family-based and group MET/CBT interventions) to 30 percent (individual MET/CBT). At least two randomized controlled trials have reported that approximately 50 percent achieve abstinence during treatment when CM/abstinence-based incentives are added to individual MET/CBT compared to 30 percent abstinence with MET/CBT alone. This is consistent with results of numerous studies in adults. Moreover, individual MET/CBT has most consistently been associated with sustained or increasing posttreatment effect size in both adolescents and adults with SUD.

Table 55.5–1.
Medications Used for the Treatment of Substance Use Disorders in Adolescents

Medication	Trial	Targeted SUD	Main Outcomes
N-Acetyl-cysteine (NAC)	Gray et al., 2010	Cannabis use disorder	► Open-label trial, 24 participants (aged 18–21 years) received NAC 1200 mg twice per day. ► Significantly greater reductions in cannabis craving via the Marijuana Craving Questionnaire (MCQ) and cannabis use (via semi-quantitative urine drug screens).
Buprenorphine	Woody et al., 2008	Opiate use disorder	► Randomized 152 opioid dependent adolescents/young adults (aged 15–21 years) in this 12-week study. ► Participants either received a 2-week detoxification with Buprenorphine-Naloxone (Bup-Nal) or 9 weeks of Bup-Nal maintenance followed by a 3-week taper. ► Those maintained on Bup-Nal were significantly more likely to remain in the study at 12 weeks (70% vs. 21%) and had significantly less injection drug use and use of opioids or marijuana. ► However, the 6-, 9-, and 12-month follow-up groups were not significantly different in reported opioid use, cannabis use or injection drug use. ► Thus, 3 months Bup-Nal did not alter long-term course suggesting that even longer term

			maintenance may be indicated—additional research needed. ► Approved in United States for the treatment of opioid dependence in those 16 years and older.
Bupropion Nicotine replacement therapy (NRT)	Killen et al., 2004 Moolchan et al., 2005 Muramoto et al., 2007	Tobacco use disorder	► Bupropion and NRT alone and in combination have been shown to be well tolerated with a relatively good safety profile in adolescents. ► There is evidence of efficacy in adolescents for NRT delivered as patch and for Bupropion at 300 mg/day dose in at least short term.

Table 55.5–2.
Medications Targeting Co-Occurring Psychiatric Disorders in Adolescents with Substance Use Disorders

Medication	RCT	Targeted Disorder	Main Outcomes
Fluoxetine	Riggs et al., 2007	MDD	► Randomized 123 adolescents/young adults (aged 13–19 years) to fluoxetine (20 mg daily dose) or matching placebo. ► Participants in both groups received weekly individual MET/CBT as outpatient substance treatment during the 16-week trial. ► Compared to placebo, subjects treated with fluoxetine had significantly greater reductions in depression severity based on CDRS-R scores (primary depression outcome measure) and higher rates of depression remission (70% vs. 52%). ► Higher than expected depression remission in both groups suggests that the individual MET/CBT (targeting SUD) received by participants in both groups may have contributed to depression treatment response. ► Significant reductions in substance use were seen in both groups but no group difference. ► Regardless of treatment assignment, participants whose depressions remitted by the end of treatment showed significantly greater reductions in substance use compared to nonremitters, whose substance use did not significantly decline from baseline levels of use despite comparable treatment compliance and completion.
Lithium carbonate	Geller et al., 1998	Bipolar disorder	► Randomized 25 adolescents/young adults (aged 12–18 years) with dipolar disorder and co-occurring SUD to 6 weeks of treatment with lithium or matching placebo. ► Lithium treatment was efficacious for both bipolar disorder and comorbid substance use disorder in adolescence. ► Longer-term follow-up study needed.
Pemoline	Riggs et al., 2004	ADHD	► Randomized 69 adolescents/young adults (aged

			13–19 years) with ADHD and SUD recruited from the community to Pemoline titrated to a once daily 37.5-mg dose or matching placebo. ► Reductions in ADHD symptom severity were significantly greater in the Pemoline treatment group compared to placebo on 1 rating scale [Clinician Global Impression-Improvement (CGI-I)] but not both primary outcome measures (ADHD DSM-IV symptom checklist scores; adolescent report). ► Substance use did not decrease significantly in either group in the absence of specific psychosocial treatment for SUD.
Osmotic-release methylphenidate (OROS-MPH)	Riggs et al., 2011	ADHD	► Randomized 303 adolescents/young adults (aged 13–18 years) with ADHD and at least one nontobacco SUD to OROS-MPH (titrated to 72 mg daily dose) or matching placebo. ► Participants in both groups participated in weekly individual MET/CBT targeting SUD throughout the 16-week medication trial. ► Results showed significant reductions in ADHD symptom severity (ADHD-RS scores; primary outcome measure) and past-28-day substance use but no difference between groups. ► OROS-MPH was superior to placebo on some secondary outcome measures: (1) greater improvement in problem solving and coping skills related to their drug use and (2) greater number negative urine drug screens. ► Results suggest that adolescents in substance treatment (even if not yet abstinent) may benefit from concurrent treatment of co-occurring ADHD with OROS-MPH (which showed a relatively good safety profile and low abuse/diversion liability with regular monitoring in this setting).
Atomoxetine	Thurstone et al., 2010	ADHD	► Randomized 70 adolescents/young adults (aged 13–19 years) referred to outpatient substance treatment and meeting diagnostic criteria for ADHD and at least one nontobacco SUD to atomoxetine or matching placebo. ► Participants in both groups concurrently received individual MET/CBT targeting SUD throughout the trial. ► Results showed significant reductions in ADHD symptom severity (ADHD-RS) and substance use but no difference between groups on primary ADHD or substance outcome measures. ► Adverse events were consistent with those reported for atomoxetine in non–substance-involved adolescents with ADHD despite nonabstinence in most participants.

An important limitation of existing substance treatment interventions is that none of them specifically address or integrate or concurrently treat SUDs and co-occurring psychiatric disorders, despite the fact that the vast majority of adolescents referred for substance treatment have psychiatric comorbidity. In addition to long-standing systemic and economic barriers, relatively few studies have focused on development and testing of integrated treatment approaches. Only five randomized controlled trials, enrolling at least 60 subjects (≥30 subjects per group), have evaluated the safety and efficacy of medications for a co-occurring psychiatric disorder in adolescents with SUD (Table 55.5–2) with brief summaries of main outcomes.

DISCUSSION AND CONCLUSION

The past two decades of research have significantly increased the current understanding of risk factors and the neurobiology of addiction that often begins during adolescence. These advances have informed the development of more effective substance prevention and treatment interventions and greater recognition of addiction as a neurobiologically based medical illness. Health care reforms such as the ACA (2010) and the Mental Health and Addiction Parity Act (2008) have paved the way for more rapid integration of mental health and addiction treatment into the mainstream medical health care system in which addiction psychiatry/addiction medicine is likely to function and operate similarly to other medical subspecialties that provide treatment and long-term clinical management of patients with chronic medical conditions (e.g., diabetes, cardiovascular disease, and asthma). As such, medical clinician training programs (physicians, nurses, physician assistants) will need to significantly enhance clinical training in screening, diagnosis, treatment, and clinical management of addiction and chronic pain including medication-assisted treatment (MAT) to address rising number of individuals with opioid use disorders, which has become national public health crisis. Additional research is needed to develop more effective screening, earlier identification of at-risk youth, as well as earlier interventions for substance-involved adolescents to reduce the risk of progression to more severe and chronic addiction in adulthood.

REFERENCES

Botvin G, Baker E, Dusenbury L, Botvin E, Diaz T. Long-term follow-up results of a randomized drug abuse prevention trial in a white middle-class population. *JAMA*. 1995;273(14):1106–1112.

Budney AJ, Higgins ST, Delaney DD, Kent L, Bickel WK. Contingent reinforcement of abstinence with individuals abusing cocaine and marijuana. *J Appl Behav Anal*. 1991;24(4):657–665.

Bukstein OG, Bernet W, Arnold V, et al; American Academy of Child and Adolescent Psychiatry.

Practice parameters for the assessment and treatment of children and adolescents with substance use disorders. *J Am Acad Child Adolesc Psychiatry*. 2005;44:609–621.

Centers for Disease Control and Prevention (CDC). (2015). Today's Heroin Epidemic: More people at risk, multiple drugs abused. Available at http://www.cdc.gov/vitalsigns/heroin/; Accessed October 21, 2015.

Croft B, Parish SL. Care integration in the Patient Protection and Affordable Care Act: implications for behavioral health. *Adm Policy in Men Health*. 2013;40(4):258–263.

Crowley TJ. Adolescents and substance-related disorders: research agenda to guide decisions on Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-V). *Addiction*. 2006;101(1):115–124.

Curry SJ, McNellis RJ. Behavioral counseling in primary care perspectives in enhancing the evidence base. *Am J Prev Med*. 2015;49(3 Suppl 2):S125–S128.

Dalwani MS, Tregellas JR, Andrews-Hanna JR, et al. Default mode network activity in male adolescents with conduct and substance use disorder. *Drug Alcohol Depend*. 2014;134:242–250.

Faggiano F, Vigna-Taglianti F, Burkhart G, et al; EU-Dap Study Group. The effectiveness of a school-based substance abuse prevention program: 18-Month follow-up of the EU-Dap cluster randomized controlled trial. *Drug Alcohol Depend*. 2010;108:56–64.

Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. *Am J Prev Med*. 1998;14:245–258.

Fergusson DM, Horwood LJ, Ridder EM. Conduct and attentional problems in childhood and adolescence and later substance use, abuse and dependence: results of a 25-year longitudinal study. *Drug Alcohol Depend*. 2007;88S:S14–S26.

Foxcroft DR, Tsertsvadze A. Universal school-based prevention programs for alcohol misuse in young people. *Cochrane Database Systematic Reviews*. 2011;9:1–69.

Gates S, McCambridge J, Smith LA, Foxcroft D. Interventions for prevention of drug use by young people delivered in non-school settings. *Cochrane Database Syst Rev*. 2006; (1):CD005030.

Griffin KW, Botvin GJ. Evidence-based interventions for preventing substance use disorders in adolescents. *Child Adolesc Psychiatr Clin North Am*. 2010;19(3):505–526.

Johnston LD, O'Malley PM, Miech RA, Bachman JG, Schulenberg JE. *Monitoring the Future National Survey Results on Drug Use: 1975–2014: Overview, Key Findings on Adolescent Drug Use*. Ann Arbor: Institute for Social Research, The University of Michigan; 2015.

Jonas DE, Garbutt JC, Amick HR, et al. Behavioral Counseling After Screening for Alcohol Misuse in Primary Care: A Systematic Review and Meta-analysis for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2012;157:645–654.

Lanier D, Ko S. *Screening in Primary Care Settings for Illicit Drug Use: Assessment of Screening Instruments — A Supplemental Evidence Update for the U.S. Preventive Services Task Force. Evidence Synthesis No. 58, Part 2. AHRQ Publication No. 08–05108-EF-2*. Rockville, Maryland: Agency for Healthcare Research and Quality; 2008.

Merikangas KR, He JP, Burstein M, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication–Adolescent

Supplement (NCS-A). *J Am Acad Child Adolesc Psychiatry*. 2010;49(10):980–989.

National Institute on Drug Abuse (NIDA) 2015, Substance Abuse - Trends and Statistics. Available at http://www.drugabuse.gov/related-topics/trends-statistics#costs; Accessed May 29, 2016.

National Research Council (US) and Institute of Medicine (US) Committee on Prevention of Mental Disorders and Substance Abuse Among Children, Youth, and Young Adults: Research Advances and Promising Interventions; O'Connell ME, Boat T, Warner KE, eds. *Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities*. Washington, DC: The National Academies Press (US); 2009.

Norman AL, Pulido C, Squeglia LM, Spading AD, Paulus MP, Tapert SF. Neural activation during inhibition predicts initiation of substance use in adolescence. *Drug Alcohol Depend*. 2011;119(3):216–223.

Riggs PD, Hall SK, Mikulich-Gilbertson SK, Lohman M, Kayser A. A randomized controlled trial of pemoline for attention-deficit/hyperactivity disorder in substance-abusing adolescents. *J Am Acad Child Adolesc Psychiatry*. 2004;43(4):420–429.

Riggs PD, Mikulich-Gilbertson SK, Davies RD, Lohman M, Klein C, Stover SK. A randomized controlled trial of fluoxetine and cognitive behavioral therapy in adolescents with major depression, behavior problems, and substance use disorders. *Arch Pedi Adolesc Med*. 2007;161:1026–1034.

Riggs PD, Winhusen T, Davies RD, et al. Randomized controlled trial of osmotic-release methylphenidate with cognitive-behavioral therapy in adolescents with attention-deficit/hyperactivity disorder and substance use disorders. *J Am Acad Child Adolesc Psychiatry*. 2011;50(9):903–914.

Stanger C, Ryan SR, Scherer EA, Norton GE, Budney AJ. Clinic-and home-based contingency management plus parent training for adolescent cannabis use disorders. *J Am Acad Child Adolesc Psychiatry*. 2015;54(6):445–453.e2.

Substance Abuse and Mental Health Services Administration, *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings, NSDUH Series H-48, HHS Publication No. (SMA) 14–4863*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014.

Swendsen J, Burstein M, Case B, et al. Use and abuse of alcohol and illicit drugs in US adolescents: results of the National Comorbidity Survey-Adolescent Supplement. *Arch Gen Psychiatry*. 2012;69(4):390–398.

Thurstone C, Riggs PD, Salomonsen-Sautel S, Mikulich-Gilbertson SK. Randomized, controlled trial of atomoxetine for attention-deficit/hyperactivity disorder in adolescents with substance use disorder. *J Am Acad Child Adolescent Psychiatry*. 2010;49(6):573–582.

United Nations Office on Drugs and Crime, World Drug Report 2015 (United Nations publication, Sales No. E.15.XI.6).

United States Preventive Services Task Force (USPSTF). Drug Use, Illicit: Screening; 2008. Available at http://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/drug-use-illicit-screening; Accessed May 27, 2016.

United States Preventive Services Task Force (USPSTF). Alcohol Misuse: Screening and

Behavioral Counseling Interventions in Primary Care; 2013. Available at http://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFi misuse-screening-and-behavioral-counseling-interventions-in-primary-care; Accessed May 27, 2016.

Volkow ND, Baler RD, Compton WM, Weiss SR. Adverse health effects of marijuana use. *N Engl J Med*. 2014;370(23):2219–2227.

Volkow ND, Koob GF, McLellan AT. Neurobiologic Advances from the Brain Disease Model of Addiction. *N Eng J Med*. 2016;374(4):363–371.

Volkow ND, Swanson JM, Evins AE, et al. Effects of Cannabis Use on Human Behavior, Including Cognition, Motivation, and Psychosis: A Review. *JAMA Psychiatry*. 2016;73(3):292–297.

World Health Organization (WHO). *Global Status Report on Alcohol and Health. Management of Substance Abuse - Alcohol. World Health Organization;* 2012. Available at http://www.who.int/substance_abuse/publications/global_alcohol_report/en/; Accessed August 2, 2015.

▲ 55.6 Forensic Child and Adolescent Psychiatry

PETER ASH, M.D., AND ELISSA P. BENEDEK, M.D.

INTRODUCTION

Overview

Forensic psychiatry is a subspecialty of psychiatry that focuses on the interface between psychiatry and the law. It may involve consulting to courts, evaluation of individuals involved in the legal system, working in corrections, having involvement in legal issues related to practice, teaching, and performing research in forensic psychiatry. Psychiatrists seeking board certification in forensic psychiatry are now required to do a 1-year fellowship in an accredited forensic psychiatry training program. No such specialized training is available for child and adolescent forensic psychiatry, although, increasingly, trainees are getting more exposure to the area in residencies, and most forensic fellowships offer some additional experience in child and adolescent forensic psychiatry. Continued practice in this specialized area may involve certification, recertification, and maintenance of certification.

Forensic evaluations differ from clinical evaluations in that there is no doctor–patient relationship with an implicit duty of care, and confidentiality is limited inasmuch as most reports are being prepared for use in court. Forensic evaluations also differ in that evaluations are focused on answering a specific question. Often the question will have a specific legal definition, which may vary from jurisdiction to jurisdiction. For example, while most jurisdictions use a “best interests of the child” test in

determining custody in divorce proceedings, in many jurisdictions this test has a specific statutory definition that lists a variety of factors to be considered. It is important for evaluators to be clear on the relevant test in their jurisdictions. The forensic examiner approaches each case with a degree of skepticism beyond that used in most clinical settings. Malingering and secondary gain need to be considered in litigants and criminal defendants. To strive for objectivity, the forensic examiner seeks out a wide base of collateral information, known as "discovery" material, which typically includes legal, medical, psychiatric, and academic records, interviews with other family members, depositions, and police reports and arrest records, when relevant. As noted, confidentiality and patient advocacy do not exist in forensic evaluations; thus, guiding ethical principles are not the traditional *primum non nocere*—first do no harm—and beneficence. In fact, were the forensic psychiatrist striving to act in the examinee's best interests, it would skew objectivity. Appelbaum suggests that, in forensic psychiatry, primacy be given to the value of justice based on the ethical principles of truth telling and respect for persons. This primacy applies regardless of whether the forensic psychiatrist has been retained by an attorney for one side or is court appointed.

Dr. A was very concerned and sympathetic as she listened to Becky, a 12-year-old girl, discussing the emotional and physical abuse she claimed she had endured while visiting her father during an unsupervised visitation. Later, while interviewing Becky's mother, Dr. A listened to a verbatim account of the abuse. It was only upon later reflection that Dr. A recognized that Becky had been coached to parrot a word-for-word version of her mother's story, and that in her sympathy for this allegedly traumatized young woman she had lost objectivity.

Guidelines

Traditionally, child and adolescent forensic psychiatry has focused on issues of child abuse and neglect, custody, and delinquency. Beginning in the early 1980s, many child and adolescent psychiatrists were drawn into civil suits as expert witnesses in cases alleging child sexual abuse. Some experienced role confusion, in which they were trying to serve as both therapist and expert witness, seemingly unaware of this inherent conflict. Added difficulties were created by the lack of clear guidelines as to how one actually assessed children for suspected sexual abuse and the paucity of good research data regarding the long-term effects of child sexual abuse.

The American Academy of Child and Adolescent Psychiatry (AACAP) has issued practice parameters for forensic evaluations which include welcome principles of practice. Anyone undertaking forensic examinations of children or adolescents should become familiar with the American Academy of Psychiatry and the Law's *Ethical Guidelines of Forensic Psychiatry*.

Civil litigation involving children may involve other types of trauma,

malpractice, and class action suits against institutions for neglect or abuse. As more youths enter the juvenile justice system, there is a growing need for forensic psychiatrists who are willing to perform evaluations for the court or work within juvenile corrections providing evaluation and treatment for detainees and committed youths. There is a high prevalence of psychiatric disorders in this population, and many youth remain undiagnosed and untreated.

FORENSIC EVALUATIONS

Referrals

Most referrals for forensic evaluations come from attorneys, and some come directly from the court or from child protection agencies or parents. At the outset, it is important to clarify whom the attorney represents, whether the case is civil or criminal, and what the legal issue is. A useful practice is to ask the referring party to put these questions in writing in a cover letter. This also enables the attorneys to think more carefully about exactly what they want to learn from the forensic examination and refreshes the memory of the examiner doing the forensic exam. Conflicts of interest need to be ruled out in taking a case. These might include having treated a relative of the examinee, being on the faculty of a facility being sued, having social ties with the involved parties, or having a personal or work relationship with the physician defendant. Forensic psychiatrists need to consider whether they have sufficient expertise in the area in question. Lack of expertise in the topic being litigated does not necessarily preclude involvement as long as one notifies attorneys of this fact. Some attorneys may be willing to provide background materials on issues not often encountered, such as toxic buildings.

Ms. Smith, an attorney, called Dr. W. and informed him that she was representing a hospital that was being sued by the parents of a youth who committed suicide on an adolescent inpatient psychiatric unit. She had heard of Dr. W.'s fine reputation, both as a therapist of adolescents and as an expert witness, and she hoped that he would take the case. Dr. W. immediately recognized a conflict of interest, in that he did some of his training at this facility and believed that might compromise his objectivity. Furthermore, the defendant psychiatrist was someone with whom he had a collegial relationship. Finally, inasmuch as his practice was entirely outpatient, he informed Ms. Smith that he did not feel qualified to testify on issues involving the standard of care on an inpatient unit. Dr. W. declined to take the case. Dr. J., however, was an adjunct unpaid professor at the same university, had not practiced inpatient psychiatry for years but was not at all reluctant to become involved in the case and did not see a conflict of interest. Needless to say Dr. J. was cross-examined on just these issues.

Other factors to consider when deciding whether to accept a case are the extent of work anticipated; whether the referring attorney is asking just for record review or wants an evaluation as well; whether a written report is expected; the timeline of the case; and whether it is feasible to meet the case's discovery date—the time when attorneys must provide the other side

with their expert's name and opinions. Fees should be discussed and a retainer may be obtained up front. It is always unethical to accept a case on a contingency fee, in that it would compromise objectivity. Plaintiff's lawyers, in contrast, act as advocates for their clients, and so they often work on contingency fees in civil suits.

Dr. G, an inexperienced psychiatrist (not a forensic psychiatrist), agreed to participate in a civil case involving a potential traumatic brain injury that a child might have incurred in an automobile accident on a contingency fee basis and agree that his payment would be a percentage of the amount the child patient recovered. A more experienced forensic psychiatrist would have realized that this violated ethical principles.

Informed Consent

Unless a court order requires the evaluee's participation, prior to continuing the interview, informed consent should be obtained. Since in most cases minors are deemed legally incompetent to provide informed consent, consent is usually obtained from parents or guardians. However, respect for persons suggests that minors should be given information about the evaluation in a manner suitable for their developmental level, which includes explaining in age-appropriate language who has retained the expert, the purpose of the evaluation, that the evaluator is not there as a treating doctor, the lack of confidentiality, and that the expert may be called to testify. It is important to assist the juvenile in understanding the information presented. Assent to the evaluation should be requested. If the minor does not assent, the clinician needs to consider carefully whether it is appropriate to proceed with the evaluation.

Dorothy, an intellectually disabled 12-year-old, said that she understood that the evaluation would not be confidential but then could not explain what confidential meant. The evaluator then made a further effort to explain the concept in simpler language, and asked her again to explain it back.

In some cases in which an examination is court-ordered, informed consent need not be obtained, but the juvenile should still be informed of the nature of the evaluation. If a juvenile refuses to cooperate with a court-ordered examination, he or she should be informed of the possible legal ramifications of failure to comply with a court order. It is prudent to clarify that the forensic examination is not a treatment relationship, and examinees may need to be reminded of that in the course of the evaluation, particularly if they persist in viewing the examiner as being there to help them. If informed consent is required, a written release should be obtained for dissemination of the written report, and the examinee should be made aware of the limited confidentiality of the evaluation.

Written Report

The written report becomes the backbone of any forensic evaluation. Care needs to be taken to ensure that it addresses the relevant legal issues and contains the foundations for the expert's opinions. A thorough, well-written report often leads to an out-of-court settlement, avoiding the cost and trauma of going to trial. If the case proceeds to trial, the report becomes the basis of the expert's testimony. Reports should be free of any inflammatory or prejudicial statements, and parties should be treated with respect. A good check on one's possible bias is to ask, "How might my report have differed if written for the other side?" The expert must be aware that many people besides the judge may read the report, including plaintiffs, defendants, and their parents.

Dr. G. wrote in a report that the youngster "had the brain of a banana." Needless to say this was an inappropriate comment.

Dan, an adolescent whose parents were involved in a contentious child custody dispute, was concerned about the possible information that might be contained in the expert's report of his evaluation. He waited by the rural mailbox daily in an attempt to read the report before his parents and was successful in intercepting it. The examiner had made mention of his latent homosexuality. Dan was distressed about this report with just cause.

CUSTODY ISSUES

The Framework

Custody conflicts may arise between parents, parents and grandparents, parents and the state, gay, and lesbian parents, and cases involving surrogates. Child custody cases, particularly those involving abused or neglected children, often tug at the forensic psychiatrist's heart, and it is easy to become overly involved and engage in rescue fantasies. Empathy is admirable, but the forensic examiner also must be able to pull back and try to assess situations objectively. It is not easy when waging sides give divergent histories and attempt to sway the examiner to a polarized position. Custody battles often bring out the worst in parents and attorneys, leading the examiner to feel caught in the crossfire and repelled by the self-serving actions of some parents. Order often may be restored if forensic examiners set firm limits with involved parties and are clear about their roles. Examiners need to remember that the judge is the ultimate decision maker. Countertransference issues should be monitored and addressed continually. Parties should be treated with parity, and the examiner should keep an open mind and avoid making premature decisions.

Dr. H. regularly consults with colleagues about contested custody cases because of his concern that countertransference may be interfering with his decision making.

Legal Issues

The forensic examiner must be aware of statutory language regarding the basis for child custody decisions; these vary from state to state. When states have guidelines regarding factors to be considered in awarding custody, clinicians need to be aware of these. Since the early 1900s, custody decisions have been guided by the concept of the child's best interests, although critics have attacked this concept as vague and not always realistic and question the concept that there is only one psychological parent, particularly in the era of dual-career families. Custody issues are the only forensic situations in which the expert has an advocacy role in speaking for the child's best interests. In contrast, in most other forensic arenas, advocacy interferes with objectivity. Child and adolescent forensic psychiatrists may be asked to advise the courts on post-divorce issues of visitation, infant visitation, and parental relocation. They also may be called on to perform evaluations for the court or for protective services in cases of termination of parental rights and adoption. These involve complex issues and the thorny question of how long a child should be kept waiting for a permanent home.

Efforts should be made to meet with both parents in custody disputes and, when appropriate, observe the child interacting with each parent. Unilateral parental evaluations in divorce situations are of very limited utility, in that the examiner has only second-hand, and often biased, information about the parent not seen. Furthermore, the examiner cannot make comparisons between the parents if only one parent has been seen. However, custody situations arise when a child is in state custody and the evaluation is requested specifically in regard to one parent's mental health and parenting capacity, and a second parent may not be in the picture. Here, the question is not which parent should the child live with, but whether this parent is capable of caring for and meeting the particular developmental needs of the child. If not, is improvement likely within the near future in keeping with the child's developmental needs? In addition, there may be attachment problems that have interfered with bonding between the child and parent. Observations of supervised visits between parent and child may yield much useful information.

Ms. Smith was seen for an evaluation as to whether her parental rights should be terminated with her 5-year-old child after her 1-year-old son was murdered by her live-in boyfriend. Ms. Smith did not believe that her boyfriend had murdered her child as he had not yet been convicted. Her history was replete with past physical and emotional abuse by her boyfriend and others. She was compliant with all evaluation appointments and court-ordered psychiatric counseling until such time as her boyfriend was convicted. Even then she believed him—she protested his innocence and became less compliant with therapeutic sessions. The evaluator expressed concern about the safety of the 5-year-old child.

Joey, age 8, had been in a therapeutic foster home for 2 years, having been removed from his

home along with his younger sister (who was subsequently returned home) owing to profound neglect as well as abuse. He was receiving appropriate services; remained volatile, with extensive developmental problems; and typically became more aggressive and regressed after weekly supervised visits with his family. Joey's guardian *ad litem* (an attorney appointed by the court to represent the child) requested that Dr. J. perform a forensic evaluation in regard to whether visits should continue. Dr. J. reviewed extensive records, evaluated each parent, obtained history from them and the foster parent, and then observed a visit. Joey's little sister totally dominated the visit, and the parents were at a loss to control her aggressive and hyperactive behavior. Joey was passive and not very engaged with either parent. According to the visitation supervisor, this was a fairly typical visit. When Dr. J. met individually with Joey, Joey expressed concerns that his sister was still being abused at home, and his need to check on her seemed to be the only reason he wanted to continue the visits. Dr. J. told Joey that both parents were getting help and doing better with their parenting, but he remained unconvinced. Joey evidenced a positive relationship with his foster father and, in contrast, had little to say about his parents. Dr. J. recommended a psychiatric evaluation on the sister, but Joey's parents did not follow through with one. She further recommended cutting visits back to monthly, but Joey's anxiety and aggressive behavior persisted around these limited visits. It also became apparent that the parents were not up to the demands of caring for two special-needs children. Dr. J. recommended ceasing efforts at parental reunification but maintaining some contact between Joey and his sister.

Parental Capacity

Assessments of parental capacity must take into consideration the child's developmental needs, the fit between parent and child, and parental strengths and ability to protect, advocate, discipline, and provide emotional and cognitive support. In addition, the psychiatrist needs to assess the parents' cognitive capacities and social supports, as well as their histories and special issues, such as prior sexual abuse, substance abuse, domestic violence, and mental illness.

Making Recommendations

Recommendations to the court should address the legal issues at hand, be feasible, and be based on the child's developmental needs.

Dr. Q., in an effort to be fair to both parents, recommended that a school-age child alternate years between parents who lived three states apart. He apparently failed to consider the effect such an arrangement might have on the child's schooling and social life. The case was relitigated when the child began resisting transitions, and a more traditional shared-custody arrangement, with primary residence with one parent during the academic year, was arranged.

Forensic examiners need to remember that the court is the ultimate decision maker and not be offended if their recommendations are not followed. Sometimes this may occur because the court has become aware of new evidence or because of legal issues. However, a report that fails to demonstrate the basis for the expert's reasoning and opinions may also lead to recommendations not being followed.

CIVIL LITIGATION

Torts

A *tort* is a claim of wrong done to another person for which a legal remedy, most often a monetary award for the plaintiff, is sought. The plaintiff must demonstrate by a preponderance of evidence that there has been a breach of duty owed by the defendant to the plaintiff, as a result of which the plaintiff suffered an injury. Torts may be deliberate, as when an individual sets out to harm someone, or unintentional, as when the act results from negligence (as in malpractice and personal injury cases). The plaintiff may seek compensatory damages for pain and suffering, special damages for medical or psychiatric care, and, in intentional torts, punitive damages. In contrast to family law cases, a jury hears these cases, and a long period often elapses before the case goes to trial. The statute of limitations in each state limits the period during which an action may be brought and usually runs from the time when an injury is discovered. However, with juveniles, in most states, the statute does not begin to run until the child reaches the age of majority.

Susan, a 10-year-old girl, was severely bitten by a neighbor's dog and sustained a cosmetic injury to her lower lip. Her father instituted a suit against the neighbor in the role of her best friend. A forensic evaluation revealed the Susan had little concern about her injury. She was doing well in school, had many friends, and had future plans to become a teacher. Her concerns centered on her mentally ill mother who had been institutionalized and then ran away from the family. Susan wondered about her health and her whereabouts. Susan's father by contrast obsessed about her injury.

Posttraumatic Stress Disorder

A common claim is that posttraumatic stress disorder (PTSD) resulted from the negligent or intentional behavior on the part of a defendant. The forensic psychiatrist may be asked to examine the plaintiff by either side. Sometimes, the forensic psychiatrist's role is limited to record review and critiquing the forensic report of the other side. Evaluating a child or adolescent for alleged PTSD and commenting on issues of causality and future damages is a complex and challenging task. Very young children may have difficulty verbalizing the subjective symptoms of PTSD, and it may manifest itself differently in children than adults. Re-enactment of a trauma may be common, but flashbacks are less likely than with adults. Parents may not be the best observers of their children's behavior and may veer toward minimizing or exaggerating psychopathology. Heightened arousal in PTSD may be confused with attention-deficit/hyperactivity disorder (ADHD) or coexist with it. Children, especially those living in inner cities or war zones or those exposed to domestic violence, may experience cumulative and chronic PTSD that clouds issues of causality. Prior trauma may sensitize them, thereby heightening their vulnerability to subsequent trauma and the development of PTSD. Age at time of trauma

may affect which symptoms develop but not necessarily symptom severity. The course of PTSD is often affected by parental responses to trauma—hence the need to explore family dynamics and other environmental mediating factors. In addition, the forensic psychiatrist should be aware of studies on the longitudinal course of PTSD in childhood.

Key Points

Key points in the forensic assessment of children involve getting thorough developmental and family histories, including whether the child has experienced prior victimization or witnessed violence or sexuality and, if so, how he or she has dealt with it. In addition, it is important to note from other sources the exact onset of symptoms. For example, pediatric records may indicate that a child's sleep disturbance and bedwetting preceded a trauma, contrary to the history given by parents. The forensic psychiatrist should be aware of the salient literature on memory in children and the suggestibility of child witnesses. The child who has been through multiple interrogations may feel that he or she has not been believed and start to embellish his or her responses or try to please the interviewer by giving the desired responses. Leading questions and either–or questions that give forced choices can be hazardous.

Questions should be put to the child in an unbiased way that encourages the use of narrative recall. It is important to elicit the child's recall of traumatic events both verbally and through the use of spontaneous play and drawings, which may reveal traumatic themes. The child may evidence behavioral signs of trauma within the interview, such as heightened startle response or, in the case of the sexually abused child, undue familiarity. Although young children are not likely to malinger, statements made by interviewers or parents can readily influence them. Source misattribution, in which children confuse something they have read or been told about with something that actually happened to them, may also lead to erroneous allegations. In research settings, children who were repeatedly urged to think about nonevents eventually came to believe they really happened.

Eddie, a 4-year-old nursery school child, claimed that she had been repeatedly sexually abused by her nursery school teacher. In addition, she reported the she had accompanied her classmates to a secret location in a nearby park where there were sacrifices of animals that they were forced to observe. After a transcript of the initial evaluator's interview of Eddie was obtained it was possible to conclusively opine that she had been asked leading questions repeatedly and then complimented when she gave the desired answers. In addition she had been told that other classmates had told the interviewer about their horrific experiences with the nursery school teacher.

Pitfalls

Common pitfalls in the forensic evaluation of damages in civil litigation include attributing all of a child's symptoms to the trauma being litigated,

failure to get an adequate prior history, bias that leads to selective interviewing, the failure to pursue disconfirming data, and the need to give attorneys the opinions they seek. Some examiners immediately equate trauma with PTSD and fail to consider other diagnoses, or even whether the trauma in question meets the threshold of trauma necessary to meet the diagnostic criteria of the latest diagnostic manual (DSM-5). Similar problems may be encountered with plaintiffs alleging trauma-related depression or other psychiatric disorders.

Dr. A. was called by a defense attorney to review discovery material in a case that alleged permanent harm and suffering in 6-year-old Tony, who was alleged to have been sexually abused at age 3 years in his day care center. Dr. G., the forensic expert for the plaintiff, had evaluated the child and performed psychological testing on him and concluded that the boy's conduct problems were all related to the alleged abuse, which the child had difficulty recalling. However, his early history on the boy was cursory, and he had little information about the mother, who was a single parent, and did not review medical records. In her thorough review of discovery material, Dr. A. learned that Tony had witnessed extensive domestic violence and his mother's rape, shown signs of hyperactivity since age 2 years, and exhibited much anxiety related to his mother's safety and several separations from her at times when she was unable to care for him owing to depression. Tony also had delayed language development. Dr. G, at the time of his deposition, was asked why he had not asked about these matters. He said he considered the mother's personal life a private matter and did not see its relevance to the litigation. Dr. A, when deposed, pointed out that many other factors beside the alleged abuse might account for Tony's behavioral problems.

JUVENILE JUSTICE CASES

Overview

The US juvenile justice system, which is more than 100 years old, is in a state of rapid change. The founding premise that wayward youths could be rehabilitated through education, special services, treatment, and redirection has been challenged as more hard-core and violent youths enter a juvenile justice system that lacks sufficient dispositions and funding. Until fairly recently, there have been few data on effective treatments and a paucity of adequate dispositions and funding for this population. In response to the upsurge of adolescent violence which peaked in the early 1990s, out of a concern for public safety many states passed mandatory waiver statutes which typically specified that any juvenile in a certain age range who was charged with certain specified crimes would automatically be waived to adult jurisdiction. Other forms of waiver to adult court include judicial waiver, in which a juvenile is waived after a hearing, and discretionary waiver, in which the district attorney by statute is allowed discretion to prosecute certain classes of juvenile defendants in either juvenile or adult court. Juvenile crime has actually decreased dramatically in the last two decades, and in response there is a trend to loosen some of the broad waiver statutes. Studies suggest that adolescent defendants waived to adult jurisdiction are more likely to recidivate once released than

are youths who committed comparable crimes and remained in the juvenile justice system.

The Juvenile Court

The juvenile court remains a civil court in which judges have wide discretion, procedures remain more informal than in adult court, and records may remain sealed. However, in actuality there are many loopholes and exceptions to the concept of the sealed juvenile record. If a juvenile is waived to the adult court, his or her records go along. Juvenile records may also be released to insurers, the military, and employers and for purposes of licensing and national security.

It was not until the 1960s that juveniles were afforded any due process rights. In a seminal 1967 case, *In re Gault,* Justice Fortas charged that "in most juvenile courts the child receives the worst of both worlds: That he gets neither the protections accorded to adults nor the solicitous care and regenerative treatment postulated for children." As a result of this case, juveniles were afforded the rights to notification, counsel, confrontation and cross-examination, and the privilege against self-incrimination. They do not, however, have the right to jury trials or transcripts of proceedings. In 1970, *In re Winship* held that the standard of evidence in juvenile hearings must be beyond a reasonable doubt, the same standard used in the adult criminal court.

Preadjudicatory Examinations

Courts often request psychiatrists to perform preadjudication examinations on juveniles. These are designed to help the judge to make appropriate dispositions. These examinations are similar to diagnostic evaluations but go further in that they also consider mitigating factors, risk factors for continued delinquency and the risk the youth may pose to the community. In addition, an assessment of amenability to rehabilitation and recommendations as to what sort of treatment the youth needs may have bearing on the type of facility to which the youth is committed.

Waiver Evaluations

Waiver, or *bindover,* as it is known in some jurisdictions, refers to the transfer of a juvenile from the juvenile court to the adult court, in which proceedings are criminal, not civil, and public, not closed, and sentences may be harsher and extend into adulthood. Most adolescent defendants who now enter the adult criminal system do so as the result of mandatory waivers or as a result of prosecutorial discretion. With judicial waivers, evaluations are used for determining which court should have jurisdiction of the youth, not for disposition as in a preadjudicatory examination. Thus, the focus is on helping the judge weigh factors that have a bearing on the

decision of whether to waive a youth. In addition to developmental and psychosocial history and diagnostic issues, the examiner considers whether the youth has benefited from any interventions, be they therapeutic or within the juvenile justice system, thus far. Amenability to treatment that can be provided within the juvenile system and ability to contain the youth within the juvenile justice system are also considerations. Special needs that the adult correctional system might not be able to meet need to be identified, and consideration also may be given to age, size, and maturity of the youth. Risk of future violence must be assessed, and short-term prognoses may be made based on known risk factors.

Competency Examinations

Competency is a term that may be used in reference to many legal issues, such as competency to stand trial, waive *Miranda* rights, plead guilty, and waive the right to counsel. The court may also request an evaluation in regard to competency to stand trial if waiver seems likely. An adolescent must be competent to stand trial in adult court. Until the mid-1990s, very few states required that a youth be competent to stand trial in juvenile court, under the theory that many youth were immature and so incompetent, and that the role of the juvenile court was to promote rehabilitation. Since the mid-1990s, an increasing number of states have required that a youth be competent to stand trial in juvenile court, as well. Competency examinations focus on the individual's current ability to understand the charges against him or her and their severity and potential consequences, understand the trial process, be able to work with his or her attorney in developing a defense, and be able to participate in courtroom proceedings. While in many respects, the evaluation of competency to stand trial of a youth is similar to that of an adult, issues of developmental immaturity are much more likely to render a youth incompetent. Juvenile defendants who are found incompetent to stand trial pose a challenge for the courts, in that a major component of restoration is simply to let the minor mature. However, prosecuting a youth several years later who is now in a different developmental stage from how he was when the crime was committed is problematic.

Criminal Responsibility Assessments

In contrast to competency evaluations, which assess present state of mind, assessments regarding criminal responsibility attempt to reconstruct the defendant's state of mind at the time of the crime. Retrospective assessments of state of mind are always more difficult and rely on not just the defendant's version of events but also a great deal of discovery material. *Insanity* is a legal term that is used as a defense to relieve defendants of responsibility for the crime. States have different legal definitions of

insanity, and a few states allow no insanity defense. Legal definitions of insanity typically include a cognitive prong that gets at whether the defendant, at the time of the crime, as a result of mental disease or defect, lacked substantial capacity to appreciate the wrongfulness of his or her conduct. *Substantial* implies serious mental disorders and does not cover Axis II diagnoses or, usually, PTSD. The exact wording of the insanity defense varies from state to state. Some states have an additional volitional prong that refers to the inability of the defendant to restrain himself or herself as a result of an abnormal state of mind. In addition, the law requires that the defendant possessed a guilty state of mind (*mens rea*)—that is, intention at the time of the crime. Mental abnormalities may affect the ability to form intent or lead to diminished responsibility for the act. Thus, a murder charge might be reduced to manslaughter if the defendant was not capable of premeditation. A few states now permit a guilty but mentally ill plea, which results in the same sentence as if found guilty, but supposedly promotes more treatment while incarcerated. While youth waived to adult criminal court can potentially make use of their state's insanity defense, only a minority of states allow the use of an insanity defense in juvenile court.

Sentencing Evaluations

In juvenile court, if the court has already obtained a preadjudicatory evaluation covering the youth's developmental status, treatment needs, and amenability to rehabilitation, that report will often be used in sentencing. Courts or attorneys may utilize similar reports in the sentencing phase if a prior evaluation has not been done or if a supplemental evaluation is deemed useful.

For youth tried as adults, forensic evaluations are increasingly utilized in sentencing decisions. For serious crimes, a series of US Supreme Court cases have changed the landscape with regard to the sentencing of adolescents. In 2005, in *Roper v. Simmons,* the Court held that "evolving standards of decency that mark the progress of a maturing society" had changed to the extent that the death penalty for youth under 18 years of age was "cruel and unusual punishment" and thus unconstitutional. The Court cited three factors regarding adolescents that made them less culpable as a class, and thus not deserving of death: (1) that adolescents were developmentally immature, and thus more impulsive and more prone to respond to peer pressure, (2) that their behavior was often affected by adverse environmental circumstances over which they had no control (such as where they lived and the circumstances of their families), and (3) that their personalities were still in flux and so their actions could not be seen as reflecting enduring character traits. Five years later, in the case of *Graham v. Florida,* using similar reasoning about adolescent culpability, and also noting that neuroscience evidence about the development of the adolescent

brain supported an understanding of why adolescents thought differently from adults, the Court held that a sentence of life without parole (LWOP) was unconstitutional for all nonhomicide offenders. The *Graham* court also said that "A State is not required to guarantee eventual freedom to such an offender, but must impose a sentence that provides some meaningful opportunity for release based on demonstrated maturity and rehabilitation," leaving it somewhat unclear whether this imposed a duty on the state to foster rehabilitation, or just left to the defendant who has received a life sentence an opportunity to demonstrate rehabilitation at a subsequent parole hearing. Two years after *Graham,* in 2012, the Supreme Court, again using the *Roper* analysis of adolescent culpability, ruled that even for murder, a mandatory LWOP sentence for adolescents was unconstitutional, although a state could still give an LWOP sentence following a sentencing hearing. Since adolescent culpability turns on psychological factors, forensic evaluations of these defendants have an important part to play in sentencing hearings where life and LWOP sentences are being considered. Although the Court holdings pertain directly to issues of death and LWOP sentences, the factors of reduced culpability of adolescents have been used in sentencing hearings of adolescents facing lesser sentences and in sentencing hearings of young adults who are slightly older than 18 years of age but who arguably, in many developmental respects, resemble adolescents under age 18.

Blended Sentences

Some courts, such as New Mexico's, have started using criminal sanctions within the juvenile court, and others such as Michigan are experimenting with blended sentences by which the youth, on reaching the age of majority, is automatically transferred from a juvenile correctional facility to an adult one. In Minnesota, judges may impose extended juvenile dispositions with stayed criminal sanctions. Courts also are using mandatory minimum sentences for certain crimes. A very promising development has been that of youth drug courts for first-time offenders charged with drug-related offenses. Close court supervision is combined with outpatient drug treatment and a team approach, which appears to be cutting down on recidivism.

Interventions

The last decade has seen many changes in the juvenile justice system. There has been much more attention to the prevalence of mental illness in this population and the implementation of evidence-based treatments for conduct disorders, substance abuse problems, and comorbid psychiatric disorders. However, many facilities and programs continue to rely on outdated interventions in which juveniles are treated in groups with their

deviant peers. Numerous studies attest to the lack of efficacy of many programs for this population, and research has documented the negative effect that delinquent youth have on one another. Innovative community-based treatment programs, such as intensive probation, are proving effective in helping youth on probation cut down on reoffending. Another effective approach is multisystemic therapy, which was designed to target and influence the known correlates of antisocial behavior in youth. This may occur with youth living either at home or in foster placement. It uses intensive services that focus on problems within the child and his or her family and environment and promotes prosocial behavior. It is a cost-effective treatment that provides an alternative to out-of-home placement and has shown to be effective with delinquents. Collaborative problem solving is another intervention that is being used effectively in many juvenile facilities. Therapeutic foster care also has been shown to cut down on recidivism in some studies. A number of school-based programs now have good evidence of efficacy in reducing violence. In addition, diversion programs and mental health and drug courts are showing promise in keeping youth out of correctional facilities.

EVIDENTIARY STANDARDS

Prior to 1993, the standard for admissibility of expert scientific testimony was whether the opinion was based on theories generally accepted in the scientific community, the so-called *Frye* standard. Judges did not have to decide themselves on the scientific merit of a theory; they only needed to ascertain if it was generally accepted. In 1993, in a case interpreting the federal rules of evidence, *Daubert v. Merrill Dow,* the Supreme Court adopted stricter evidentiary standards requiring that expert testimony be based on scientific knowledge. Federal judges now had the job of deciding whether a theory was actually scientific, not whether it was generally accepted. Under the reliability prong of *Daubert,* the Court listed four nonexclusive factors to be considered when considering the reliability of theory or techniques as cited in Federal Rule 702. These included (1) testability, (2) subjection to peer review, (3) potential rate of error or standards regarding operation, and (4) acceptability in the scientific community, the remnant of the earlier standard set forth in *Frye*. Under the relevance prong, the Court asked the following: (1) Does the testimony logically advance a material aspect of the case? (2) Is there a valid connection to disputed aspects of the case? (3) Is the testimony sufficiently reliable? And (4) Does it rise to the level of intellectual rigor of an expert in the relevant field? *Daubert* is binding in federal courts and is followed by most but not all state courts. It has affected testimony in cases of alleged child sexual abuse and has limited admissibility of hearsay testimony in child custody cases.

FUTURE DIRECTIONS

This is an exciting time to be working in the field of child and adolescent forensic psychiatry because expanding research allows for opinions that are based on research data in addition to specialized knowledge of child and adolescent development, which will aid the courts in their deliberations. This subspecialty is gaining recognition and a body of literature along with clinical guidelines. Child and adolescent psychiatrists also have an important role in conducting research in the areas of prevention, intervention, and understanding the effect of trauma on the developing brain.

The juvenile justice system is in need of further reform. Many programs are inordinately expensive, and their efficacy had not been documented. In fact, in some programs, interventions have led to more reoffending or violent behavior. In many jurisdictions, juvenile offenders receive poor legal representation due to excessively large caseloads assigned to juvenile defenders and inadequate compensation rates. The recognition by the courts of the decreased culpability of adolescents, and the rapidly expanding knowledge of the underlying neurobiology of the developing brain, gives hope that public policy will continue to move in the direction of taking account of developmental information in rehabilitating adolescent offenders. Standards for determining child custody also need to be reviewed in light of changing family structures and the vagueness of "the best interest of the child" standard. Finally, it is hoped that forensic psychiatry fellowships will offer more flexible and diverse programs for those trainees wanting to specialize in child and adolescent forensic psychiatry.

ACKNOWLEDGMENT

The authors would like to acknowledge Dr. Schetky for her contribution to previous versions of this chapter.

REFERENCES

*American Academy of Psychiatry and the Law. *Ethics Guidelines for the Practice of Forensic Psychiatry*. Available at http://www.aapl.org/ethics.htm. Accessed January 6, 2016.

American Medical Association and American Academy of Child and Adolescent Psychiatry. *Brief for the American Medical Association and American Academy of Child and Adolescent Psychiatry as Amici Curiae in Support of Neither Party, in Evan Miller v. State of Alabama,* U.S. Supreme Court Case Nos. 10-9646, 10-9647; January 13, 2012.

Appelbaum PS. A theory of ethics for forensic psychiatry. *J Am Acad Psychiatry Law*. 1997;25:233–247.

Ash P. But he knew it was wrong: evaluating adolescent culpability. *J Am Acad Psychiatry Law*. 2012;40:21–32.

Ash P, O'Leary PJ. An annotated bibliography for the testifying child and adolescent psychiatrist.

Child Adolesc Psychiatr Clin North Am. 2011;20:577–590.

*Benedek EP, Ash P, Scott CL. *Principles and Practice of Child and Adolescent Forensic Mental Health*. Washington, DC: American Psychiatric Publishing; 2010.

Brodsky SL, Gutheil TG. *The expert witness: More maxims and guidelines for testifying in court*. 2nd ed. Washington, DC: American Psychological Association; 2016.

Candilis P, Weinstock R, Martinez R. *Forensic Ethics and the Expert Witness*. New York: Springer; 2007.

Ceci SJ, Kulkofsky S, Klemfuss JZ, Sweeney CD, Bruck M. Unwarranted assumptions about children's testimonial accuracy. *Ann Rev Clin Psychology*. 2007;3:311–328.

Cohen JA, Bukstein O, Walter H, et al; AACAP Work Group On Quality Issues. Practice parameter for the assessment and treatment of children and adolescents with posttraumatic stress disorder. *J Am Acad Child Adolesc Psychiatry*. 2010;49:414–430.

Daubert v. Merrill Dow Pharmaceutical. 509 U.S. 579; 1993.

Dodge KA, Bierman KL, Coie JD, et al; Conduct Problems Prevention Research Group. Impact of early intervention on psychopathology, crime, and well-being at age 25. *Am J Psychiatry*. 2015;172:59–70.

Drozd L, Saini M, Olesen N. *Parenting plan evaluations: applied research for the family court*. 2nd ed. New York: Oxford; 2016.

Elkington KS, Teplin LA, Abram KM, Jakubowski JA, Dulcan MK, Welty LJ: Psychiatric disorders and violence: a study of delinquent youth after detention: *J Am Acad Child Adolesc Psychiatry*. 2015;54:302–312.

Everson MD, Sandoval JM, Berson N, Crowson M, Robinson H. Reliability of professional judgments in forensic child sexual abuse evaluations: unsettled or unsettling science? *J Child Sex Abus*. 2012;21:72–90.

Federal Bureau of Investigation. *The School Shooter: A Threat Assessment Perspective*. 2002. Available at http://www.fbi.gov/publications/school/school2.pdf.

Graham v. Florida G, 130 S. Ct. 2011, 560 U.S.__ (2010).

Grisso T. *Evaluation of Juveniles' Adjudicative Competency*. Sarasota, FL: Professional Resource Press; 2005.

Grisso T. Adolescent offenders with mental disorders. *Future Child*. 2008;18:143–164.

Grisso T, Vincent G, Seagrave D, eds. *Handbook of Mental Health Screening and Assessment for Juvenile Justice*. New York: Guilford Press; 2005.

In re Gault, 387 U. S. 1; 1967.

In re Winship, 397 U.S. 358; 1970.

Kraus LJ, Thomas CR, Bukstein OG, et al. Practice parameter for child and adolescent forensic evaluations. *J Am Acad Child Adolesc Psychiatry*. 2011;50:1299–1312.

**Miller v. Alabama, 567 U.S. __, 132 S. Ct. 2455;* 2012.

Mississippi Youth Justice Project: Mississippi. *An Assessment of Access to Counsel and Quality of Representation in Youth Court Proceedings*. Montgomery, AL: Southern Poverty Law Center; 2007.

Montgomery v. Louisiana, 577 U.S. ___ ; 2016.

Righthand S, Kerr B, Drach K. *Child Maltreatment Risk Assessments: An Evaluation Guide*. New

York: Routledge; 2011.

Roper v. Simmons, 543 U.S. 551; 2005.

Strasburger L, Gutheil T, Brodsky A. On wearing two hats: role conflict in serving as both psychotherapist and expert witness. *Am J Psychiatry*. 1997;154:448–456.

U.S. Department of Justice Office of Justice Programs. *Promising and Proven Programs on Youth Violence Prevention*. Available at http://ojp.gov/programs/yvp_programs.htm. Accessed January 20, 2015.

▲ 55.7 Ethical Issues in Child and Adolescent Psychiatry

ADRIAN NATAN SONDHEIMER, M.D., AND ARDEN D. DINGLE, M.D.

Ongoing scientific developments and discoveries continue to advance and enhance the practice of medicine, often in areas relevant to the care of children and adolescents. The medical profession is no stranger to such continued accrual of new knowledge and to the potential benefits and risks associated with these advances, especially those that at times suggest quantum leaps ahead. In past centuries, the pace of change in the "knowledge base" was slower, but societies still needed medical personnel, designated "healers," whose responsibility was to diagnose ailments, provide treatments, care for the sick, give advice to the afflicted and their families, and develop new understandings of maladies. Since time immemorial, society's healers have encountered and considered conflicting theoretical stances, goals, and approaches. Steady advances in knowledge in the pursuit of providing optimal care have always led to comparisons of consequent risks versus benefits, and practitioners often had to confront the reality that proposed new solutions to problems can induce harmful effects.

These age-old conflicts among and between competing views and goals in the pursuit of "doing good" are ethical ones, and physicians—today's "healers"—exemplify their professional roles when giving these conflicts their due consideration. Situations in which such conflicts, or ethical dilemmas, arise include the degree to which physicians impose their wills on a patient when the patient's wishes oppose the physician's; the potential for physicians' personal aggrandizement or hubris to influence medical decision making; and the recommendation of treatment options that would benefit an individual patient but might conflict with the needs or desires of the family, community, and/or society.

Ethics is the discipline that focuses on moral principles or values that guide attitudes and behaviors while examining notions of right and wrong in relation to these human activities. Acting in an ethical and professional manner has been a fundamental tenet of medical practice throughout its history and remains an essential aspect of medical practice. The ethics of medicine has always intended to concern itself with the pursuit of patients' best interests and the societal good while respecting individual patient's personal rights. The codifications of musings about these considerations serve to advise, admonish, or assist people to behave in desirable ways, and these written compilations reflect the philosophical worldview and religious, spiritual, political, cultural and economic circumstances of their societies.

CODES OF MEDICAL ETHICS'HISTORICAL, NATIONAL, AND INTERNATIONAL

Medical ethics is the application of the ethical discipline to medicine. The individuals designated to treat the somatic and psychological ills of the community's members—today's physicians—are commonly accorded privileges and respect for their specialized knowledge and abilities. To promote this trust and protect the public, the healers and their communities have produced formalized written codes, or guidelines, governing the behaviors and conduct expected of physicians. These codes of medical ethics focus substantially on the treatment obligations of the healers to the individuals residing within these communities.

Several themes and exhortations are commonly found in medical codes of ethics that date back many hundreds of years and emerge in myriad cultures. These include primacy of respect for human life; physician responsibility for the patient; emphasis on virtue and duties; provision of help; avoidance of harm; equality of care for rich and poor; setting of fees; methods used to establish diagnoses, with the expected employment of clinical observation and reasoning; use of appropriate therapies; confidential nature of patient–physician communications; prohibition of physician–patient sexual relationships; and provisions for punishment or sanctions in the event of poor technical performance, code violations, or both.

In the United States, the formation of the American Medical Association (AMA) in 1847 was accompanied by the creation of a code of ethics. A society organized in 1844 by 13 medical administrator superintendents of "American Institutions for the Insane" became the foundation for the creation of the American Psychiatric Association (APA) in 1921. In 1973, the APA produced its first code of ethics, which was annotated and revised in subsequent years.

Considerations of children's needs were largely absent in the formative years of these professional organizations. The dawn of the 20th century saw the beginnings of professional reflections concerning children's health needs as separate from those of adults. The American Academy of Pediatrics was founded in 1930, and in 2007 that body adopted for the first time a document focusing on the mission, values, and professionalism distinctive to its practitioners. APA members, independent psychiatrists, pediatricians, psychoanalysts, and other medical practitioners with a primary interest in children's mental health founded the American Academy of Child (and Adolescent) Psychiatry (AAC[A]P) in 1953. The AACAP introduced its code of ethics in 1980; it was expanded, with accompanying annotations, in 1995, revised in 2007, and substantially rewritten in 2009. In 2014, the code was further revised and enhanced by the inclusion of discussions related to social media, expansion of research matters to include scholarly activities, and a focus on the explicit natures of potential boundary crossings. The AACAP code covers ethical issues

relevant to clinical, administrative, and scholarly practice in child and adolescent psychiatry. Its narrative is organized around principles and major specified areas of concern inclusive of: Developmental Perspective, Promoting the Welfare of Children and Adolescents (Beneficence), Minimizing Harmful Efforts (Nonmaleficence), Assent and Consent (Autonomy), Confidentiality (Autonomy/Fidelity), Third Party Influence (Fidelity), Scholarly and Research Activities, Advocacy and Equity (Justice), Professional Rewards, and Legal Considerations.

This code uniquely introduces a developmental perspective, thus covering an essential aspect of caring for children, adolescents, and their families. It affirms that abiding by these ethical principles is considered an essential aspect of being a competent and capable child and adolescent psychiatrist (CAP). Further, continued learning and maintenance of the ability to identify, analyze, and make decisions about ethical dilemmas are considered a key skill set for all medical students, residents, and physicians. Ethical CAPs should possess the knowledge of ethical principles and their use but, even more fundamentally, understand that this knowledge and skill set are an essential aspect of their professional demeanors and should be reflected in their behaviors. The code advises that a key aspect of the ethical practice of CAP is a developed awareness of one's personal beliefs and the impact of these beliefs on the clinical and professional choices made by the practitioner.

Ethical thought is universalizable, that is, ethical principles and reasoning apply in all situations, nationally and internationally. However, when ethical reasoning is utilized, attention must always be paid to situational context and culture. Thus, while emphases on, for example, competence, honesty, avoidance of exploitation, confidentiality, boundaries, and protection of research participants transcend individual nation-states, their examination and potential resolutions are partially contingent on geographic location. Many ethical and related codes have been produced by international bodies, including those of the World Medical Association (originated 1949), World Psychiatric Association (originated 1996), and the United Nations Convention on the Rights of the Child (1989). This last document established the rights of children to survival, development, protection, and participation in society and, by implication, their health care. In 2006, the International Association for Child and Adolescent Psychiatry and Allied Professions (IACAPAP) produced the document Ethics in Child and Adolescent Mental Health, which created principles addressing the rights of children and behavioral guidelines for the professionals who treat them worldwide.

While children and families in developed countries can struggle to access services, the resources available to developed versus resource-poor countries, as well as the natures of their governance and cultures, serve as the greatest differentiators in mental health approaches to children living

in dissimilar locales. In the former, for example, the practitioner frequently is concerned with the ethics surrounding recommendation of brand versus generic medications in comparison to the practitioner whose pharmacological armamentarium often is empty. Similarly, a locale's culture often determines whether a child's protests are welcomed as a sign of developing autonomy versus being viewed as a violation of elders' decision making for the group, the culturally sanctioned norm. At the worst, displaced refugees, homelessness, immigration, sex-trafficking, and the production of orphans (secondary to war, capture and disease) account for huge numbers of children in predominantly resource-poor countries. A paucity of mental health professionals in these locations further ensures that little attention will be paid to children's psychiatric needs. The ethical imperative to improve children's life circumstances and to provide appropriate care, however, remains unaffected by the surround.

MEDICINE AND MENTAL DISTURBANCE

From a historical perspective, within the context of medically influenced writings, mental illness was first formally addressed by the Greek physician Hippocrates and, several centuries later, by Galen in Rome. The ancient Romans dealt with the definition of mental disturbance, the nature of caregivers, guardianship, approaches to care, and legal competence. Thus, Roman society and numerous others that followed, by recommending approaches that they considered to be correct and beneficial medical and judicial professional behaviors, were instrumental in contributing to the creation of modern codes of behaviors expected of healers. Focusing on medical illness, inclusive of those who suffered from mental disturbance, the codes and laws forged by the ancient Romans and reconsidered and expanded in subsequent eras explicitly addressed the ethical conflicts that daily confront medical practitioners without separating the somatic and psychological spheres.

CHILDREN AND ETHICAL CONSIDERATIONS

When and how did children come into the picture? Some historians have posited that prior to the 17th century, in the Western world, children generally were not viewed as developmentally distinct from adults. Pediatrics was first established as a unique specialty in the late 1800s. Over the last few centuries, children and adolescents have been recognized as distinct groups engaged in ongoing growth, as well as knowledge and skill acquisition, in emotional, physical, intellectual, language, moral, and interpersonal spheres. In the last decades, increasing recognition and support has accrued to the perspective that children have inalienable rights to development and participation in society, as described in the UN Convention on the Rights of the Child. The substages of child and adolescent development imply that children's maturation in developmental

spheres correlate with chronological age and, therefore, that expectations of children's capacities for comprehension, reasoning, and sound making of decisions will vary with the individual child's degree of attainment. From an ethical perspective, insofar as these capacities remain immature during development, children generally are not granted decision-making rights; rather, these responsibilities remain with their guardians. When the goals of the guardians and their children conflict, however, problems may ensue. Furthermore, children are commonly cared for by individuals or agencies concerned with their welfare, such as schools, individual health practitioners and institutional providers, governmental welfare and protective services, and juvenile justice systems. Information exchange among these parties is necessary for the provision of care, but the transmission of data among them can occur without a child's knowledge, leaving affected children vulnerable to the potential misuse of confidential information. Thus, the factors of continuous developmental maturation, potentials for children–guardian conflicts, and numerous avenues of and needs for information transmission frequently lead to ethical dilemmas concerning children.

Ethical Dilemmas and Reasoning

Practitioners might recognize the presence of a child-related ethical dilemma when experiencing internal conflict about potential divergent approaches to the provision of clinical care. For example, CAPs consulting to a pediatric clinic, with 15 minutes available to see a 6-year-old distressed patient, must consider whether to spend this limited time with the patient and guardian or with the guardian alone. If psychiatrists choose the latter, then they can anticipate obtaining a coherent history without needing to "manage" the individual patient, an approach that could rapidly and succinctly provide the needed medical input. However, direct observation and interaction with the child might prove crucial to an adequate database. How to provide the best care—an ethical imperative?

Ethically problematic situations like this one are commonly ambiguous —they involve gray areas. In the face of such a conflict, given the pressure of time and clinical demands, harried clinicians might not register the ethical dilemmas related to the principles of *beneficence* (promotion of well-being) and *nonmaleficence* (avoidance of causing harm). Therefore, it is most advisable for clinicians to resist an immediate response, consider the alternatives, engage in a reasoning process—largely a risk–benefit analysis—and ultimately make their choice based on a resolution that is least damaging and ideally most helpful to the patient.

Ethical Principles and Forensics

Medical ethics is a field of inquiry based on principles that are considered

universal, inclusive of the obligations to pursue *beneficence, nonmaleficence, autonomy* (promotion of individual freedom of action and choice), *fidelity* (faithfulness, loyalty, and respect for the rights of others), and *justice* (treatment of individuals with fairness and equality). Parenthetically, within this last category, consideration may also be given to the concept of *social justice,* that is, the treatment of groups of individuals with fairness and equality. These principles often serve as the foundations for derivations of legal precepts and administrative decisions. For example, the legal concepts of the individual's right to privacy and informed consent for medical procedures largely derive from the principles of autonomy and fidelity. Additional concerns include the possibly conflicting rights of individuals and society, as in the case of proposed quarantine for contagious infectious disease which pits conflicting concepts of individual autonomy against those of communal social justice. Ethical dilemmas in medical practice often involve the operation of several of these principles simultaneously, and the principles may conflict and compete.

The consideration of multiple principles in a reasoning process, however, allows for leeway in decision making. The law, by contrast, although generally based on ethical deliberations, defines acceptable choices and behaviors more narrowly. Not uncommonly, the discussion of ethical considerations segues into an examination of legal ones, conflating the two distinct disciplines. At times, the legal response can be at odds with powerful ethical arguments. For example, in custody disputes that include significantly involved nonbiologically related foster parents, certain jurisdictions adhere to laws that mandate awarding custody to the closest biological relative, even to one not involved in the child's current care. Such compulsory rulings might conflict with evidence of the establishment of close and long-standing emotional ties between children and their nonrelative foster parents. In such instances, implementation of the law could appear to cause harmful, in contrast to beneficial, effects.

Although several distinct ethical concerns may exist simultaneously, over time one approach to the dilemmas usually takes form as the most beneficial, or least harmful. The CAP clinician, administrator, researcher, or teacher, dealing with a problematic case, must strive to arrive at the least damaging resolution, commonly using a risk–benefit analysis. Awareness of relevant laws is necessary to this reasoning process, but the law is not necessarily the deciding factor.

CHILD-RELATED CONCERNS

Areas of ethical concern common to child and adolescent psychiatry practice include confidentiality, parent–child conflicts in desires and rights, assent to or refusal of treatment or research, physician responsibility for continuation or termination of care, child advocacy and protection, agentry, boundaries, and professional responsibilities.

Dilemmas arise in many contexts, including inpatient and outpatient facilities, private offices, and institutional settings; in the course of obtaining histories and providing treatments; abuse and neglect; divorce and custody; and forensic settings (e.g., court). Participation in administrative and research/scholarly activities often raises ethical issues. Conflicts arise owing to increased physician awareness of, and sensitivity to, issues of culture and religion, changes in health care delivery and economics, heightened understanding of the scientific basis underlying disease etiology and expression, the influence of various institutions involved in health care, and the presence of electronic methods of communication, used both professionally and socially. Consequently, new dilemmas, or old dilemmas in new garb, confront CAP practitioners, administrators, researchers, and educators on a daily basis.

Multiple Agentry

Angela, a pregnant 15-year-old, was informed that her 16-week-old fetus was anencephalic. Obstetrics personnel urged termination of the pregnancy, but Angela refused to abort. One day later, she became acutely suicidal. In the emergency department, a psychiatric evaluation found her imminently dangerous to self. Angela refused to agree to psychiatric hospitalization and an involuntary commitment process was begun. Subsequently, the receiving psychiatric hospital refused to admit Angela without a prior termination of pregnancy. Angela's history indicated that the state child protective services agency, which had functioned as Angela's legal guardian for 5 years, had 1 year earlier removed Angela's then 7-month-old child from her care after a finding of neglect.

CAPs are obliged to recommend and pursue optimal care for their patients. In this case, doing so would require a guarantee of the patient's safety, and, therefore, in Angela's case, hospitalization is indicated. The practitioner, however, must also take into account other influences, such as, in this case, obstetric and pediatric personnel who strongly argue for the termination of the pregnancy to spare anticipated future suffering of the patient and fetus. Other positions requiring consideration include those of the receiving psychiatric hospital that refuses to admit Angela without prior termination of the pregnancy; Angela's foster parent, who refuses to countenance abortion as an option for a child living in her house; the child protective services agency, Angela's guardian, which refuses to take a decision-making role, arguing that the patient is competent to make decisions; and the patient, who refuses to assent to either psychiatric hospitalization and termination of the pregnancy. While Angela's right to make the decision about her hospitalization varies depending on the state in which she lives, she retains the right to decide about her pregnancy throughout the entire United States. The CAP has the obligation to be Angela's advocate in order to ensure that her best interests are protected and she does not experience harm. Thus, working with Angela to better understand her views and helping her to make good decisions, while

utilizing appropriate strategies to manage her distress, is essential.

CAPs commonly face issues of multiple agentry—the need to be responsive to a case's interested parties, who often are pursuing conflicting agendas. As above, it is fundamental that the CAP protect and advocate for the individual patient. That is, CAPs must act as agents for the minor patient while not permitting themselves to be swayed by the goals of the other parties, their own personal biases, the sources of income that pay for the physicians' efforts, or their allegiances to fellow professionals. CAPs have the obligation to act as the agents for their minor patients, despite the engagement in the care of these patients by parents or guardians with whom the CAP may spend more time than with the child and who often are paying the physician's fees.

"I want you to treat my daughter. She's been misbehaving at home and at school. This has nothing to do with my recent separation from her mom, and the fact that my new girlfriend is living in our home. I'm not going to say another word about this—concentrate your efforts on my daughter!"

"I am the only normal one in my family. You've heard what happens at home—my parents are crazy. Who behaves like this? I'll tell you, people who have problems. They should be in treatment, not me."

Who is the patient? When dealing with children and adolescents, it is often useful to use a family systems perspective that focuses on the dynamics between and among the various parties in the nuclear household. This stance can be viewed as a beneficial approach for all the involved parties because it permits CAPs to function as agents for both the family as a unit and its individual members. When the guardians refuse such an approach, however, practitioners face at least two ethical dilemmas. CAPs need to provide optimal care to the child while not alienating the parents and must avoid inappropriately labeling the child with psychopathology when the primary disturbance may reside with the parents. Such an approach requires clinical dexterity and due regard for the autonomy of all the interested parties while maintaining as primary the needs of the child.

Effects of Culture and Religion

As the intermingling of differing, multiple cultures continues to be the ever-growing norm in the United States, CAPs are frequently called on to deal with conflicting notions of expectable behaviors and deviations from them.

"I brought my 13-year-old daughter to the emergency department because she's hitting me. She's sinful, and she's crazy. We've always gone to church together, 5 days a week, and now she refuses to go. She has to go, because otherwise the Devil will take hold of her, as He did to me in my teens, when I had her. Her hitting me, her mother, is just the beginning of His hold on her, and through her the Devil will infect my younger children."

> "I don't know why my parents are bothering you with this. They want me to be more observant of our rituals, participate more in our culture, and continue with my education. Both of my parents didn't finish high school and most people my age back home are working at a job. I'm trying to do what they want, but now I found a job and they have no reason to be mad that I'm quitting school."

Varied culturally and religiously defined misbehaviors are found within an array of ethnicities and denominations. CAPs must rely on the profession's scientific knowledge base to evaluate case data in order to arrive at psychiatric diagnoses rather than using the belief systems and mores preferred by a culture or religion. However, CAPs should also be aware of, acknowledge, and have respect for the importance of cultural and religious perspectives to patients and their families. Treatment plans and dispositions should be approached in a similar manner, using resources peculiar to the belief system to the extent possible, without jeopardizing the care of the child. In addition, CAPs should note that separation-individuation phenomena manifest and provoke consternation in many, perhaps all, cultures. In the interest of providing good care, CAPs are ethically bound to pay attention to both culture-specific and developmentally normative behaviors. Not infrequently, education concerning normal child and adolescent development, conveyed in a caring manner, is much appreciated by the parent living in an otherwise self-contained culture.

Countertransference

The five aforementioned scenarios likely engendered, in the evaluating and treating psychiatrists, affective responses toward the protagonists. *Countertransference,* the term colloquially used to describe these phenomena, is closely related to CAPs' affective responses to the raised issues. What are the psychiatrists' attitudes toward abortion, teenagers who engage in sexual activity and become parents, parents reluctant to acknowledge the effect of their behaviors on their children, assaultive youths, individuals for whom religious and cultural beliefs and mores are paramount, and youngsters striving for autonomy? It is important for CAPs to engage in self-inquiry and determine the reasons for their countertransferential responses to these factors so as to ensure that these phenomena do not negatively affect optimal care for the children (and adults) under their care. Overidentification by CAPs with the child or adult is dangerous, as are rescue fantasies that focus on children who seem to live with little in the way of adequate adult protection, advocates, or both. CAPs in earlier stages of their training may tend to ally themselves with their patients, whereas older ones, many of whom have become parents, may tend to express greater empathy with the guardians. Ultimately, CAPs must

focus on the needs of the child and avoid choices that are more reflective of the individual psychiatrist's intrapsychic dynamics.

PROVISION OF CARE

From a clinical perspective, ethical considerations should begin with the first patient contact and extend past the final meetings.

Assent and Consent

> "Hello, doctor. I would like to make an appointment for my 9-year-old boy who probably has ADHD (attention deficit hyperactivity disorder), and I know he has lots of behavior problems."

> "Oh, I'm not telling him that we're coming to see the doctor. He gets so upset and I figured you could explain it better."

Questions that should follow immediately on such a seemingly innocent and sensible inquiry about an appointment ought to address whether the involved child agrees to a medical procedure (i.e., the psychiatric evaluation and interview). Will the child assent to this procedure, is he averse to it, and/or is he being coerced by the parent to attend? To what extent should children's objections to the initiation of an evaluation, or to treatment that might be recommended at its conclusion, be respected—that is, to what extent is children's *autonomy* considered? Have the children been reasonably well apprised by the guardians of the reasons for the proposed evaluation, and, if not, by the practitioners in the course of this first meeting—in other words, has the assent been informed?

Insofar as children ordinarily are legally considered minors until age 18 years, children are generally unable to give legal *consent* to proposed medical procedures, but they may *assent* to them. It is the children's guardians who provide consent. Further complications can arise when disagreements exist between children's legal guardians regarding proposed care, as may occur in the context of divorced parents. Often one of the parties is legally designated the guardian ordinarily responsible for obtaining medical care for their children. In any case, in emergencies, the urgent need for provision of care becomes the physician's paramount consideration, irrespective of parental guardianship. In situations short of emergencies, however, lack of consent or opposition by one of the guardians raises ethical and legal dilemmas. The parental disagreement might be exacerbated by psychiatrists' actions favoring the stance of one of the parents, which could negatively affect or harm the child by aggravating the parental conflict. Standing in contrast, psychiatrists must also consider the children's rights to an evaluation and their need for care.

> "I'm pregnant, my parents don't know, I'm not sure who the father is, and I'm not sure what to

do about it. I'm 15 years old, I can think for myself and make my own decisions, I have the money to pay your fee, and I want to discuss this with you without my parents knowing about it."

"I don't see why you want to talk to my parents. Yes, I live with them but they really don't know anything about me and I'll be leaving soon when I graduate from high school."

These adolescents are not simply assenting to treatment, but are requesting care from psychiatrists. A 15-year-old pregnant girl, in some states' jurisdictions, is automatically considered an adult or "emancipated minor" by dint of her pregnant state, whereas in others she remains legally a minor. "Mature" or "emancipated" minors can legally access care for themselves. Ordinarily, however, care only can be provided, except in emergency situations, with the consent of an adult. At what age does the adolescent develop the cognitive and emotional maturity to be able to give consent? Is that decision based on chronological age (e.g., attainment of age 18 years), or is it situation or case dependent? Given that adolescence is a stage of rapid physical and cognitive maturation and change, this fluidity in growth renders ethical questions regarding autonomous decision-making more complex and less certain than when dealing with similar issues in younger children. As adolescents advance toward cognitive maturity, respect for their desires and opinions should progressively be factored in to professional decisions in the interests of respecting the youths' entitlement to *autonomy*.

Confidentiality

It is the physician's responsibility to protect, and not convey to others, information about a patient, unless the patient has given permission for this transmission. The patient has a right to expect adherence by the physician to this obligation, which serves as a basis for a trusting physician–patient relationship. The principle of confidentiality is often subject to limitations in work with children, in that guardians, patients, and agencies want and require information about child patients, and children often are unable to comprehend the risks and benefits involved in the waiver of their rights to confidentiality. In addition, electronic medical records and other electronic methods of communication can mean that, without adequate built in safeguards, multiple individuals and institutions may potentially access confidential patient information.

"It's simple, doc. You've known my kid for a while. She's seriously disabled, she's 10-years-old, and we need the state benefits to which she's entitled. Just fill out this ten-page form—it asks for all kinds of information about her—nooooo problem!"

"Doc, can you give me a copy of her chart? The school needs the records to start their evaluation process and it'll be faster if I just deliver them the information."

A signed release from guardians permitting transmission of information about children might satisfy legal concerns, but ethical issues remain that warrant exploration. To what extent do children's disabilities affect their capacity to provide assent? Would state/school involvement foster a child's development or hinder it? Is all, some, or none of the requested information pertinent to the child's condition? Is the information requested legally protected, or might it be accessible or transmitted to parties in the future without the patient's or guardian's knowledge? Psychiatrists must evaluate these issues by simultaneously playing the role of child advocates and attempting to protect children from current and possible future misuse of the information that the practitioners might provide. Such questions should be addressed with both guardians and, to the extent possible, patients, because families need to be aware of the ways in which information provided to third parties might subsequently be used. They also should be educated about the potential risks of information disclosure when various forms of electronic communication, such as e-mail, social websites and cell phone texting, are employed.

"About those emails I've sent you about my daughter. You know, asking whether she talks to you and tells you the truth or she lies? Please don't tell her about those messages or about any of the info I've given you about her."

Adolescent patient to psychiatrist: *"I don't really talk on the phone. Can you text? I'll definitely get any messages then. And that way I can let you know what's happening; it would be great to talk between sessions."*

Commonly, parents or guardians are eager to know details concerning information the CAP has obtained in the course of treatment of their minor. The use of a family therapy approach at the onset of treatment, with an explicit message emphasizing free but judicious exchange of information, can circumvent many confidentiality concerns. Informing the family members at that time that information requested or supplied outside of treatment sessions would likely be shared with all obviates otherwise sequestered communications. As well, it is warranted to notify all participants at the outset that potential self-harmful or externally directed aggressive acts do not fall under the confidentiality obligation. In forensic contexts, minors and their guardians need to be informed from the outset of the lack of confidentiality concerning personal information (i.e., that the information included in forensic reports will be made available to all the interested parties and be part of the public record). Ultimately, practitioners should always exercise tact and restraint in the provision of information to all concerned parties, and all the more so when the transmitted information is likely to be reviewed by non-mental health professionals.

Separately, the advent of commonly used digital media and electronic communications, while positively facilitating information exchange of all kinds, does not in any way eliminate the need for alertness to the existence of ethical considerations and legal concerns in these spheres. The considerations that govern face-to-face care apply equally to electronic communications with patients and their families, despite their apparent ease of use. Thus, potentials for boundary violations, confidentiality and privacy transgressions, negligence, and defamatory messages are abundant and sometimes not obvious. Further, the seductive simplicity of seeking and finding information about patients and their families, easily obtainable via the Internet, could result in ethical violations when performed without the patients' and families' knowledge and, similarly, if not necessary for the patients' care. All of these concerns and behaviors involve potential violations of the patients' rights to expect respect by physicians for their *autonomy* and avoidance of breaches of *fidelity*.

Fidelity and Mutual Trust

Fidelity, honesty, and *integrity* are three cardinal qualities expected of physicians that contribute to the measure of a physician's professionalism. Patients are entitled to expect physicians to deal with them openly, stay with and not abandon them through the course of their illnesses or disturbances, and make attention to their needs primary before a physician focuses on possible self-interests. These expectations, however, may come under pressure when a physician feels threatened or hemmed in.

> "You've seen my 16-year-old son for a month. He's angry and unhappy, fighting with me, and failing school. I think he's using drugs. I want you to tell his pediatrician to order a tox[icology] screen and have it performed without his knowledge when he sees the doctor next week. My son would refuse if asked directly. If it turns out negative, he won't know I suspected him of drug use; if positive, I'll keep the results to myself and decide later what I want to do about it."

The patient's guardian requests of the psychiatrist to deal with the patient in a nonstraightforward manner to obtain important information. This request is ethically problematic because the creation of mutual trust between physician and patient is a fundamental tenet for the establishment of a sound treatment relationship. A physician–patient relationship in which neither side could expect honest exchange would be doomed from the outset, even if the knowledge sought in a clandestine manner were significant. The means to ends should be discussed in a candid exchange, even if some of the desired ends are, in the short run, thus rendered more difficult to achieve. Short of emergencies, there is no place for a paternalistic approach that would justify withholding vital information from a patient. It is crucial to the establishment and maintenance of successful therapeutic relationships that all family members understand

this principle.

CAPs and others involved in the care of children are periodically faced with behaviors that adults experience as repugnant, threatening, or both.

An inpatient pediatric attending addresses the consulting CAP: "The nursing staff on my unit have been telling my patient, a 15-year-old boy with hepatic failure, that they wish he would die. He developed severe hepatitis as a child and became a 'darling' of the unit during his stays there, while receiving and then rejecting two liver transplants. He is back here now being evaluated for a third transplant, but 2 weeks ago he was charged with sexual abuse of a child, and the staff has completely turned on him."

CAPs tend to have considerably more experience with victims of abuse than with (alleged) perpetrators. Physicians' efforts, therefore, are mostly devoted to the repair of those individuals who have experienced emotional injury. Children who are accused of doing (or who do) bad things, however, also require care. Nondiscrimination is a fundamental ethical imperative. Provision of the physicians' care is dependent on a patient's needs and not on the whims, preferences, or biases of the practitioners. Thus, it is incumbent on CAPs to act as advocates for the abusing, malfeasant, and nasty child as much as for the victim deserving of sympathy. In other words, despite possibly harboring negative feelings toward patients, physicians must maintain *fidelity* to them and continue loyally to provide the distasteful children with professional service.

Despite such exhortations, situations arise in which CAPs' negative feelings toward children may interfere with the provision of good care, in particular when CAPs feel threatened by the children.

A 15-year-old male teenager in a detention facility, adjudicated for murder, calmly says to the examining CAP: *"I know where you live. I don't like you. You stay in my face and I'll kill you. That's not a threat—that's a promise."*

Psychiatrists, like all human beings, are entitled to self-preservation and considerations of self-interest. To the extent possible, CAPs should avoid allowing these concerns to jeopardize the health of their patients. Termination of care may be appropriate in the event CAPs believe themselves to be in danger, but only if provisions are made beforehand for care to be provided by other practitioners.

Boundaries

Physician behavior toward patients should always be initiated and maintained with the patient's best interests in mind. Medicine's professional ethos emphasizes strict physician adherence to the value of integrity. In the context of the physician–patient relationship, the concept of boundaries refers to limits within the relationship that ought not be breached. Maintenance of such boundaries helps preserve the integrity of

the relationship. Although some physician behaviors clearly violate the physician–patient boundary and are impermissible (e.g., sexual relations between physician and patient; physician use of confidential patient-derived information for the physician's gain), other behaviors reside in gray areas. Physicians who divulge personal information to patients, or overlook timely payment for services from patients reporting financial difficulties, may be engaging in good, compassionate practice or in enticing and exploitative behaviors that service the physicians' needs, or both.

> "I saw you eating dinner at a good restaurant a few weeks ago, so I figured you liked it. You've helped our family so much that I wanted to say thank you with a gift card to that place. Don't worry, it isn't a big deal."

> "Can you come to my graduation? I couldn't have made it without you."

> "Can you lend me fifty cents? I'm really hungry and don't have enough money for the potato chips."

> "The case is a great example of what I am presenting on, and I need to send in my talk tomorrow. It would be too difficult to find the patient and ask her permission to use her information. I'll change the details so no one will know who the patient is."

Work with children and adolescents can lend itself to the potential for boundary infringements. CAPs who passively enjoy the spectacle of a child's oppositional or adolescent's seductive behaviors and therefore neglect to set appropriate and timely limits, agree to treat children despite knowing the family socially, accept significant presents, prolong children's care owing to an attraction to a parent, or discuss with colleagues or write for publication about a patient for the purpose of self-aggrandizement are, in all these instances, potentially violating the patient's boundaries.

Patients justifiably expect not to be exploited by their physicians and to feel secure in the treatment setting. Despite best intentions to engage in self-monitoring, physicians' positive countertransferences toward patients could result in behaviors that end in a special or favored status for particular patients. In those instances, it is highly recommended that physicians recognize these types of behaviors and their potential for leading toward more frank violations. Rapid consultation with respected colleagues is the preferred antidote to strengthen and secure physicians' self-control and provide patients with sound and unimpaired care.

CURRENT AND FUTURE ISSUES

Although the ethical principles of beneficence, nonmaleficence, autonomy, fidelity, and justice remain immutable, their consideration must be applied to constantly changing contexts that include advances in scientific

knowledge, innovations in means and modalities of communication, changes in financial structures, shifts in societal priorities, newly created treatment approaches, and novel research foci with patients. These arenas continually evolve, with further developments undoubtedly to come. The implications for child and adolescent psychiatry practice have been and will be considerable, leaving physicians struggling to keep pace with these rapid changes and the freshly minted ethical dilemmas they pose. The following subsections focus on eight relevant areas.

Genomics

In 2000, the Human Genome Project concluded a successful effort to decode the DNA sequence of an individual's 23 paired cellular chromosomes. The consequences of this achievement, as well as the ongoing knowledge accumulated from continued research in this area that includes DNA-based tests concerning specific genetic disorders, focus on chromosomal changes, and gene products, have profound implications for illness prognostication, treatment, and counseling; the relative importance of environment and genetics in the transmission of psychiatric disorders; and payments for the treatment of these disorders.

For example, current capabilities include the potential use of gene chip technologies to determine the optimal pharmacological treatment of children resistant to specific medications. Future CAPs will no doubt be in a position to consider genetic information, unique to their patients, as a basis for deciding whether to treat an illness preventively that has not emerged or has barely been expressed. The proposed treatments might be applied with currently available psychotropic medications or with yet-to-be-developed gene therapies. Informed *assent* or *consent,* or both, and the weighing of the numerous risks versus benefits of these postulated interventions, need to be mandated as ethical components of the process. Similarly, CAPs' notations in the medical records of genetic information and related considerations could lead, without strong adherence to the principle of *confidentiality,* to inadequate safeguards for the protection of patients' genetic data. This breach, in turn, could lead to the stigmatization of patients and possible discrimination against them by potential employers or agencies responsible for hiring and the provision of health and life insurance coverage.

There appears to be a consensus that children ought not to be genetically tested for precursors to adult-onset psychiatric disorders, such as Huntington and Alzheimer diseases. However, technologies might soon become available that would be capable of screening children for childhood-onset psychiatric disorders, and a strong argument could be made that not to do so would put the child at undue risk. As knowledge expands, the enhanced ability to predict children's development of behavioral, cognitive, and affective traits and disorders is likely to become a

factor that CAPs will consider. It is probable, therefore, that the CAPs of the future will be concerned with the principles of doing good and avoiding harm in the context of working to a considerably greater extent in tandem with medical geneticists and counselors.

Treatment of Psychiatric Prodromes

Clinical researchers have postulated that the appearance of specific behavioral signs and symptoms in childhood or adolescence may be pathognomonic (i.e., currently prodromal) of the later development of full-blown severe psychiatric disorders, such as schizophrenia. Several institutions chose, therefore, to investigate this contention by considering the combination of the appearance of these early-onset behaviors together with a positive family history for the illness as predictive of the likelihood of its later development in the identified youths. Because of the debilitating nature of the full-blown illness, it was the investigator's intentions to provide treatment—for example, neuroleptic medications—to these youths and subsequently to monitor outcomes. Relevant ethical concerns were and are many. Aside from the complicated matters of informed consent and assent, dilemmas include accuracy of the identifications and intrinsic difficulties in monitoring outcomes. Specifically, it is at best difficult to determine whether it was preventive pharmacological treatment that forestalls later full expression of the illness, given that the expression might not have occurred under any circumstances. Similarly, the development of adult-onset schizophrenia in a prodromal individual not exposed to preventive treatment might occur despite earlier exposure to neuroleptic medications. And, complicating the situation, the data on treating prodromal individuals has at best been equivocal, resulting in the discontinuation of these programs at several of the experimenting institutions. Furthermore, recent work suggests that psychosocial, family-based interventions may be more effective for prodromal individuals than antipsychotic drugs. In sum, the theoretical prodrome is not definitively predictive; it simply postulates vulnerabilities and increasing possibilities of the later emergence of full-blown adult-onset illness.

Within this paradigm, later development of the illness in a previously untreated youth would imply that there was an earlier withholding of a potentially beneficial intervention. Further complicating the picture, a proposed treatment—for example, neuroleptic medication—might be known to have the potential for producing severe deleterious effects. Finally, knowledge of genetic factors that might more accurately predict the future development of the illness is a promising anticipated advance in molecular biology, but one that remains as yet to be achieved. Of course, the theoretical basis for the treatment of childhood psychiatric prodromes is not limited to schizophrenia, and similar ethical considerations, involving uncertainty of outcomes, potential dangers and benefits of

interventions, and the effects of withholding interventions, would apply to many of the biologically based disorders.

Information Technologies

Substantial advances in electronic data collection, reporting, storage, retrieval, and communication are continuing to revolutionize medical practice, thus enhancing the possibilities for sustained continuity of care while raising novel ethical issues for child and adolescent psychiatry practice. The advent of the electronic medical record increases the possibility of transmission of confidential information concerning minors, without a guardian's consent, over the time period that the information is electronically stored (theoretically forever). The child, who owing to age lacks control over personal information transmission one day becomes an adult. At that time, potential employers', health insurers', and public agencies' electronically facilitated access to this information could negatively affect the individual, owing both to the stigma that continues to attach to mental disturbance and the effect of pre-existent illness on employer hiring practices and employee acquisition of health benefits. Will an adult have control over personal health data entered into information banks when that individual was a minor? Who owns this information, who is empowered to make these decisions, and what standards are to be used?

The Health Insurance Portability and Accountability Act of 1996 was established to create transaction standards and privacy rules overseeing the electronic transmission of personal medical information. Although the Act, which became effective in 2003, was intended to provide federally mandated privacy protections for this information, critics contend that loopholes exist. In the interest of advocacy for patients, it has become the CAP's professional responsibility to alert children's guardians, as well as the adolescent who may have arrived at a cognitive maturity enabling comprehension of the implications, to these privacy concerns.

In a related phenomenon, individuals of all ages, in huge numbers, are increasingly using the Internet and other electronic platforms to store information and communicate. Simultaneously, electronic systems are increasingly used to provide treatment, as in telemedicine, and many websites exist to complement traditional therapies. In-place security protections of these systems vary. Patients' medical information could be accessed by strangers. Many people fail to appreciate the implications of putting out information electronically on social media, e-mail and via texting. Children and adolescents, in particular, due to their developmental immaturity, often lack the capacity to truly understand the immediate and long term implications of sharing personal data via these mechanisms. Use of these systems by professionals creates potentially significant additional complications, for providers as private citizens as well as in their roles as health care personnel. In addition to the challenges of keeping patient

information secure, maintenance of boundaries in the therapeutic relationship, while providing effective treatment interventions, may frequently be threatened. These communication methods can be effective, personally and professionally, but their use requires caution, vigilance and self-monitoring.

Funding and Provision of Patient Care

In the United States, health care costs have been a major expenditure for the past four decades for individuals, employers and the government. At the same time, many have voiced concerns about the quality of provided care. Various administrative interventions have been implemented in the attempt to improve care as well as to contain costs. The Affordable Care Act, enacted in 2010, the most recent major health reform, aims to improve access to affordable health insurance for more Americans, enhance the quality of health care, regulate the health insurance industry, and decrease health care spending. It is too early to determine the impact of this legislation on the practice of child and adolescent psychiatry though the law, increasing the number of individuals having coverage for mental health care, in turn highlights the shortage of CAPs and other mental health providers in most areas of the country. An earlier reform enacted during the 1980s, colloquially termed managed care, had a clear effect on child and adolescent psychiatry practice. Managed care's primary goal was cost containment, which often complicated and compromised the delivery of effective psychiatric care for children and adolescents. Its innovations included gatekeeper requirements, prior approval for treatment—even in emergencies, and limitations on treatment length and access to providers. Then and now, many of the institutions providing insurance payments for mental health care are for-profit companies with obligations to their shareholders, suggesting possible conflicts of interest negatively affecting their customers and patients.

Sixteen-year-old Julie was psychiatrically hospitalized for expressing suicidal thoughts after having been physically assaulted by her father. Julie's parents had divorced 5 years previously, and Julie since lived with each parent at different times. Six months before the hospitalization Julie became pregnant, was "kicked out" of her mother's home, and obtained an abortion while living with her father. The recent assault was the father's response to Julie's demand to spend time "chilling" with her drug-dealing boyfriend. During the admission Julie exhibited symptoms of a major depressive disorder, although her suicidal ideations had diminished in frequency and intensity. Hospital utilization review informed the attending CAP that the patient's insurance coverage provided a maximum 5-day stay and demanded that the patient be discharged on the 5th day. The parents stipulated "impossible" expectations of Julie in the event that she might come to live with them; Julie demanded both discharge from the hospital and the right to live with her boyfriend; and a state child protective services investigation found no need for its further involvement in the case.

One of the raisons d'être of the current versions of managed care is the

pursuit of the ethical principle of *distributive justice*. In the case of health care provision, that translates into the funding of basic care for the largest number of people by controlling costs expended on the care of individual patients. Furthermore, the fact that many managed care and health insurance companies are for-profit concerns, that is, their primary goal is to benefit their shareholders, is an obvious additional consideration when examining the motivations contributing to administrative stances and decisions. Nevertheless, the CAP, despite awareness of these "macro" concerns, continues to be bound by the professional obligation to provide beneficial care for each individual patient, independent of funding. Although, in the previous case, Julie's hospitalization coverage is about to expire (in theory so that there will be coverage available for other individuals in need), Julie continues to require care—for her depression, protection, and sustenance. The attending CAP remains ethically and legally responsible for attempting to make the best, or least damaging, decisions regarding Julie's needs for care in the immediate future. If the CAP deemed that further hospital stay was indicated, it would be incumbent on the CAP to request additional funding from the insurance company, appeal the company's refusal of funding in writing, or arrange for hospitalization elsewhere. If comfortable with the discharge option but not with the current dispositional choices, the CAP would be obligated to pursue the state child protective services agency to re-examine the case, given that it is this agency's responsibility to provide alternative placement possibilities. Although funding mechanisms may change, the CAP's ethical constraint to provide beneficial care is immutable.

Dr. Green, a CAP, had been providing a combined treatment approach of psychotropic medication and psychotherapy to Ben, an 11-year-old boy who had developed posttraumatic stress disorder (PTSD) 6 months earlier after exposure to a stressful event. Ben's mother informed Dr. Green that her health insurance benefits, obtained via her employment, had changed. The new insurer limited coverage by CAPs to brief medication checks, and benefit payments for psychotherapy were available exclusively for nonmedical mental health personnel. Dr. Green believed that the psychotherapists recommended by the insurer had inadequate training to handle Ben's case.

While frustrating and time consuming, CAPs have the obligation to advocate with insurance companies in their attempts to obtain what they consider optimal care for their patients. It is notable that in some parts of the United States, most commonly in large, urban areas, some CAPs have managed the dilemmas and conflicts arising from dealings with insurance companies by choosing to not participate with them and, instead, engage in fee-for-service practices. While this approach does allow these psychiatrists to provide treatments as they deem appropriate, it raises major ethical issues related to access to care and larger considerations of social injustices. *Split treatment,* defined as the CAP providing psychiatric

evaluations and medication management while nonmedical mental health practitioners provide verbal therapies, has become the norm in numerous health care settings. Consequently, many psychiatrists have far less knowledge of their patients than do their peers who provide combined treatments, while they yet retain an ethical responsibility for their patients' optimal care. In addition, these CAPs often have little contact with the nonmedical mental health practitioner providing verbal treatment. Reduced knowledge of patients, and self- and externally-induced pressures to treat more patients in less time, raise the possibilities of poorer attention to patients' needs despite the ethical constraints to provide beneficial care, not do harm, and make the patient's interests the primary concern.

Advocacy

The impetus to contain health costs, with an emphasis on fiscal restrictions, has significantly highlighted the imperative for child advocacy in both public and private systems caring for youth. Children lack the autonomy and power that cognitive maturity, access to employment, and voting rights provide. Their guardians, insurers, or both, inadvertently or deliberately, may make decisions that are not in the children's best interests. Therefore, proper treatment and placement venues, possible needs for evaluation and treatment, the foci and choices of treatments and the need for their continuation, the value of extended physician–patient relationships, and recommendations for adjunctive investigations and care are all matters that may require advocacy by the CAP for a child.

A 10-year-old boy with a history of assaultive behaviors was brought to a hospital emergency department after having attempted to stab a peer with a knife 1 hour before. The patient was still agitated, expressing rage at this peer, and possessed negligible insight and responsibility for his behavior. Nevertheless, the boy's father insisted on discharge of the patient to his supervision in order that they keep their long-established plans to fly the following day to a distant city. Independent of the family plan, the patient did not want to be hospitalized. The father assured the evaluating CAP that he would immediately obtain psychiatric care for his son in the new location. The patient's insurer expressed ambivalence about hospitalization, especially after learning of the father's wishes, but reluctantly agreed to cover the costs in the event the CAP chose this course of care.

The guardian's concerns are legitimate, and his intentions may be well meaning. The insurer's concerns are legitimate as well, although considerably influenced by financial considerations. Owing to safety considerations, however, the CAP must opt for psychiatric hospitalization of the child. The practitioner is obligated to advocate for the child's immediate well-being, even in the face of opposition from the guardian, child, and insurer. *Advocacy* implies choosing the least damaging or most beneficial course for the child in dealings with parents, guardians, agencies, insurers, and the patients themselves. Individuals, agencies, or insurers may attempt to dissuade CAPs from their chosen course because the

interests of the payment providers may conflict with the immediate needs of the child. As the child's potential advocates of last resort, however, CAPs should provide to the interested parties the reasons underlying their position as well as the maintenance of their medically indicated stance, despite possible resistance or antagonism.

Administration

CAPs serve as administrators in a variety of institutional settings and wear multiple hats. Common to CAP administrators are responsibilities for the provision of clinical mental health services to child and adolescent populations, responses to ever-increasing budgetary pressures, and, in that context, maintenance of staff morale and cohesion. Common administrative maneuvers during times of budgetary constraints include attempts at financial savings via personnel layoffs and cutbacks, reduction in services to patients and their families, and shifting of clinical priorities and foci. For example, the CAP administrator may consider whether funds formerly intended for adolescent inpatient care should be shifted to child day-hospital services or vice versa; whether a child outpatient clinic should be maintained or scuttled if expenses exceed income; and whether a child and adolescent psychiatry employee should be assigned an increase in caseload without a corresponding increase in salary or time availability. These considerations, maneuvers, and decisions cause distress to the administrators, patients, and professional staff.

The CAP administrator is constrained by the ethical injunctions to not do harm and to maintain fidelity to patients. Fiscal realities also necessitate considerations of *distributive justice,* translated into the attempt to provide reasonable care to the greatest number. Considerations of these fiscal concerns can come at the expense of patients currently receiving treatment. The administrator must seek to ensure that care remains available to them —if necessary, via referral.

Administrators are also responsible for the behaviors of those who report to them.

Mrs. Smith complained to the CAP director of a mental health clinic: "My 8-year-old has a diagnosis of oppositional defiant disorder and, with him present, Dr. Jones accused me at length of being the cause of his troubles. Since then, my son doesn't listen to anything that I tell him to do. You better get rid of that Dr. Jones."

The mental health administrator has the reasonable expectation that child and adolescent psychiatry personnel will perform competently and professionally, which includes according patients and their guardians appropriate respect. Administrators must also investigate patients' complaints. When allegations have foundation, the administrator must determine whether a professional's misbehavior or poorly chosen words

reflect a momentary aberration or are one in a series of inappropriate actions. Psychiatric work with children and adolescents has the potential for stimulating poorly restrained countertransferential responses that arise primarily from a professional's unresolved conflicts, possibly stemming from the physician's childhood, in contrast to behaviors reflecting an objective response to the clinical presentation. It is the CAP administrator's responsibility to ensure appropriate professional behavior among the caregivers under his or her supervision. The administrator must also distinguish between a single untoward incident that may be responsive to personal counseling in contrast to the consistent expression of ingrained unprofessional behaviors that would require a significantly stronger response. In both situations, the administrator must endeavor to ensure, as much as possible, that the behaviors do not recur and that professionalism is maintained.

Effect of Pharmaceutical Industry

For a number of years concerns have been voiced about the complicated relationships between medicine and the pharmaceutical industry. Pressure has been brought to make public traditionally private information detailing the financial relationships between individual physicians and pharmaceutical companies. As well, drug companies are now required to make all data derived from their studies publicly available, including negative results. Further, more and better data has been collected on the number of children and adolescents prescribed multiple psychotropic medications, often with unclear indications, while progressively younger children are affected by this phenomenon. These and related concerns are promoting continued discussion of the ethical dimensions of industry–physician and industry–psychiatrist relationships and stimulating ongoing advocacy for transparency and change. Medical schools have implemented policies regarding these relationships, imposing restrictions on interactions and exchanges between industry representatives and physician trainees and practitioners. Professional organizations have re-evaluated their relationships with pharmacological companies and instituted new and more stringent policies that govern their dealings with each other.

The industry, "Pharma," produces the medications prescribed by physicians. It also funds the bulk of CAP research, directly and indirectly funds the work of academic researchers, and underwrites and provides financial support for professional journals, conferences, and lectures. In response to negative publicity concerning Pharma's contributions to individual prescribers, subtly designed to influence prescribing practices, presentations by "opinion leaders" and "freebies" supplied to clinicians in the forms of small gifts, dinners, and pharmaceutical samples continue, though not to the same degree as was true in recent years. The industry engages in these practices, in largest part, because the component

companies are for-profit concerns, beholden to their shareholders, and their primary goal is to maximize purchase of their products. The results of Pharma's largesse to CAP professionals can result in increased discovery and dissemination of new knowledge, expanded evidence-based understandings, and improved patient care. However, given the potential for unconscious biases that are the consequences of funding sources for research as well as donation of "presents" to prescribers, practitioners need to remain intelligent consumers of the professional literature in order to judge the purported value of research results and their interpretations against possible sources of external influence.

Because gifts and samples tend to promote reciprocally positive responses from recipients, clinician CAPs might unwittingly be induced to prescribe more expensive, in contrast to less expensive, medications, or brand in contrast to equally effective generic drugs, thus unnecessarily increasing costs for their patient's families. The current influence of Pharma on clinical practice has the potential for leading to an increased emphasis on drug studies in professional journals, at the possible expense of studies devoted to the nonpharmacological components of the CAP's treatment arsenal. This possibility, coupled with a changed health care indemnification environment, in turn can lead to excessive CAP dependence on drug prescription and, on occasion, to its excessive use. In addition, the expansion of criteria to include youngsters in diagnostic categories from which they had been formerly essentially excluded, as in the case of bipolar disorder, could be perceived by some observers as a consequence of the insidious effect of the influence of Pharma as easily as it would be viewed by others as a real and considerable psychiatric advance.

Thus, the relationships between Pharma and child and adolescent psychiatry as a profession, as well as with individual CAPs, raise numerous potential and actual conflicts of interest. Can these conflicts be managed, or should they be eliminated, by sundering all but indirect contacts and contributions between the industry and CAPs? As illustrated, the ethical dilemmas are many and profound, and the choices made by individual practitioners and professional organizations are likely to significantly affect future patient care, research funding, basic science knowledge, and professional education.

Research, Neuroethics, and Psychopharmacology

Arguably, the major recent advances in CAP knowledge and treatment have resulted from quantum leaps in knowledge of neural function, particularly at the brain receptor and molecular levels. This knowledge has correlated with a considerable expansion in the knowledge base of neuropsychopharmacology, accompanied by the continued development of new products available for the treatment of disabling conditions.

As is true for all treatment techniques, the decision of whether to

intervene pharmacologically requires a risk–benefit analysis. The potential dangers of acute or chronic side effects must be weighed against the long-term effects of unhindered chronic psychopathology. There remains a dearth of knowledge concerning the effects of psychotropic medications, either short or long term, on possibly vulnerable developing brains. Although some negative effects have correlated with medication treatments (e.g., aspirin with Reye syndrome and central nervous system stimulants possibly with growth retardation), studies have not, by and large, followed psychotropic-taking children for years into adulthood. Many psychiatric disorders first manifesting in childhood or adolescence (e.g., autism spectrum, bipolar, Tourette syndrome, and obsessive-compulsive disorders) are likely to remain manifest as chronic conditions in the ensuing decades, whereas the long-term effects of the psychotropic medications prescribed for their treatment are uncertain.

In a related concern, psychotropic medications are commonly prescribed by physicians for "off-label" use. For example, many medications prescribed by CAPs for child psychiatric disorders (e.g., schizophrenia and bipolar disorder) have FDA approval of prescription only for adults. Thus, medications that target severe behavioral and cognitively disruptive behaviors of childhood, which include antipsychotic, mood-stabilizing, antidepressant, and cardiovascular agents, often do not have FDA approval for their prescription to children. In addition, prescriptions of these medications for newly defined psychiatric conditions, such as disruptive mood dysregulation disorder, have as yet to obtain FDA approval for this purpose, and their use in these circumstances would again be termed "off-label." Although considered clinical judgment can result in the conclusion that these medications in fact might be indicated and beneficial, a rigorous evidence base for that opinion could be weak or nonexistent. Because many childhood behavioral disorders contain elements of various disease entities without, however, fully satisfying the established diagnostic criteria, pharmacological management is often attempted based on coupling the overt behaviors with hints or suggestions of other signs or symptoms. Also, practitioners attempt to treat with medications because adequate and appropriate psychotherapeutic and environmental interventions often are not accessible or available. These diagnostic uncertainties and absences of full ranges of care, coupled with inadequate evidence bases, always pose the ethical conflict between the desire to provide good care—for example, by using what are thought to be appropriate and indicated psychopharmacological agents—and the possible iatrogenic precipitation of harm.

The lack of knowledge concerning long-term outcome indicates the continued need for research to expand the child psychiatric scientific database. Political jockeying around the creation of a mandate to study the effects in children of medications prescribed to adults continues to be an

ongoing process. Such a directive has not occurred. Rather, the pharmaceutical industry has been encouraged to expand drug trials to include children via the use of incentives, for example, by increasing the time period allotted for exclusive patents. In any event, until an increase occurs in the number of drug trials focusing on children, prescribing new medications to youngsters without advance knowledge of potential ill effects could put them at risk. Thus, informed consent is crucial in these circumstances. A description of the treatment protocol, risks, and benefits that might accrue consequent to the provision or withholding of the novel treatment, and avoidance of coercion with respect to patient participation, should be expected by the prospective participants in what are essentially experiments whose predictable outcome is indeterminate. Although consent to a child's participation in a research protocol must come from fully informed guardians, the child should be advised of all the pertinent factors to the extent of his or her capacity to understand, and noncoerced assent obtained for participation to occur.

The phenomenal increases in knowledge of molecular biology have widened the spectra of innovative drug treatments, not only for diseases of the brain but also for the enhancement of what until now has been considered cognitive or affective functioning, or both, within that broad range called "normal." Stimulant medications, for example, are known to be effective in increasing attention and concentration not only for children diagnosed with ADHD but also in normal populations. Novel antidepressant medications, effective for the treatment of major depressive disorders, in coming years might also prove useful in forestalling the "downs" in the ups and downs of ordinary life. Might the drugs of the future enhance the sociability of the nondisabled shy younger child or, conversely, reduce the boisterousness of the mildly annoying exuberant child? One needs only to look at controversies regarding possible steroid augmentation of athletic performance to comprehend that similar improvement or magnification of abilities could equally be applied to brain function. It is an open question whether the popular response to steroid-related athletic enhancement might be similar to, or differ from, the responses to medication-induced brain enhancement. In any case, these and related concerns are subject to musings by neuroethicists, those bioethicists who discern such futuristic scenarios as imminent.

A similarly relevant matter concerns the significant increase in the number of psychotropic medications prescribed for children in the United States and elsewhere in the world over the last few decades. Potential explanations for this numerical explosion are many, including a gain in sophistication of professional knowledge of indications for interventions with medications in children, the number of studies indicating medication efficacy (coupled, unfortunately, with earlier withholding from publication of studies productive of negative findings), the range of medication choices

among chemical classes of psychotropic medications, access to medical and psychiatric care by larger numbers of children, dollars devoted to the advertising of pharmaceuticals that target physicians (including CAPs), and CAP professional education that might increasingly emphasize "biological" interventions at the expense of time devoted to training in the verbal therapies. Publicly aired concerns about this increase, in concert with suggestions that the administration of SSRIs to depressed youngsters increased the risk of suicidality in the children, led to FDA health advisories that culminated in "black box" warnings about this danger. The factors of a significant illness (depression) associated with a potentially fatal outcome (suicide), occurring in a high-risk population (adolescents) whose treatment (SSRI antidepressants) might slightly and inconclusively exacerbate the incidence of the self-harmful behavior, all occurring in the context of a pharmaceutical industry that has significant interests in the marketing of its products, has left CAPs on uncertain ground as they attempt to weigh the risks versus the benefits of their therapeutic choices and do what is most beneficial for their patients.

Following the FDA advisories, later studies revealed that the rate of SSRI prescriptions for this population dropped sharply, a development followed in turn by a seemingly significant increase in suicidal behaviors in the target population. This is a fine example of the ethical dilemmas CAPs face as they grapple with an enlarging psychotropic armamentarium, broadening clinical presentations for which medications are prescribed, and increasing economic pressures on and inducements for the practitioners to write prescriptions. In the end, it is incumbent on CAPs to adopt the ethical stance of weighing the objective evidence, while distancing themselves from personal considerations and biases, and making choices that foster the best interests of the child.

TEACHING ETHICS AND REPORTING VIOLATIONS

The American Council of Graduate Medical Education mandates the teaching of ethics in child and adolescent psychiatry training programs in order that future CAPs be aware of the potential emergence of ethical conflicts during their careers. Topics deserving didactic exposure during training commonly include matters generic to all physicians, such as confidentiality, boundaries, professional responsibilities, and relationships with health care providers, companies, and industries, as well as topics focused more specifically on child and adolescent psychiatry practice. The latter would likely include the topics of child advocacy and protection, informed consent and assent, refusal of treatment, child–guardian conflicting rights, agentry, compensation source, autonomy and paternalism, and research-related risks and benefits. In addition, closer attention likely would be paid to the pitfalls potentially accruing to communications via electronic transmissions, and the ethical responsibility

for the practitioner to be aware of federal, state and local laws governing medical and child and adolescent psychiatry practice. Clearly, ethical concerns regarding children arise from the youngsters' needs for adult protection and nurturance, given their physical and cognitive immaturities. Related training in and knowledge of psychodynamics, the origins of motivations, and systems theory would augment, and can easily be incorporated into, the CAP trainee's ethics education.

During this training, CAP trainees should become informed about the AACAP *Code of Ethics*. The document addresses the needs and rights of children and their guardians, as well as the CAP's responsibilities to both. On occasion, trainees and/or experienced practitioners may experience uncertainty regarding an ethical dilemma concerning a child, and referral to the *Code* may help to guide the response.

The AACAP Ethics Committee also serves as a potential resource for advice concerning ethical dilemmas. After deliberations by the committee, inquiries receive suggestions for approaches to resolution. The potential value to be derived from requesting consultation from colleagues, when faced with ethical dilemmas in general and from the committee in particular, cannot be overstressed. In recent years, inquiries presented to the AACAP ethics committee have included concerns about administrative orders to withhold information from families; "double booking" of patient appointments; inadequate research supervision; noninclusive and/or biased custodial evaluations; the use of agencies and credit bureaus to collect fees; potential breaches of patient confidentiality; poor or potentially harmful patient care; organizationally blocked access to professional consultation coupled with inadequate supervision; video monitoring on CAP inpatient units; disclosure of case details in, and obtaining patient consents for, publications; abandonment of children by professionals in the context of threatening parents; potential supply of sexual paraphernalia to patients; practice of child psychiatry without formal training; the impact of television and other media on children younger than the age of 3 years; requests of families for funding of research; treatment by a CAP parent of that psychiatrist's child; reporting alleged abuse/neglect to state child protection agencies; and potential adoption of former patients. In addition, the AACAP Ethics Committee has developed a range of educational materials, disseminated in various forums and maintained on the AACAP ethics website, related to child and adolescent psychiatric practice. Of note, in 2010, the AACAP implemented a mandate requiring all AACAP members to obtain attested to ethics education credit every 2 years.

Periodically, fulfilling a different function, the committee receives complaints concerning practitioners alleged to have engaged in unethical practices in matters concerning children and adolescents. When further investigation is indicated, complainants are referred to appropriate

organizations and agencies, such as the APA and AMA district branches and state medical licensing boards. Investigations by these entities could result in "no findings" or recommendations for sanctions, which may include reprimand, suspension, or loss of licensure. Abrogation of licensure by a state medical licensing board, consequent to findings of an ethics violation, automatically results in loss of membership in the AACAP, the guild representing the profession's ethos.

The role of the healer in society has a long history, and society's designated medical personnel have always been faced with ethical dilemmas. With the rapid increase in relevant scientific knowledge and related technological achievements, the dilemmas have grown more complex and complicated. The basic understandings, however, remain the same and undiluted, that is, the challenge of doing good and not engaging in harmful behaviors, equally for all patients. CAPs have the additional burden of caring for the immature child, as well as maintaining good relations with his or her guardians and caretakers. Whether the CAP is wrestling with the age-old conflicts surrounding privacy and confidentiality or dealing with the newest developments in genetics and transmission technologies, ethical responsibilities represent a never-ending challenge for all practitioners and for the subspecialty as a profession. In the face of unceasing change, therefore, teaching of ethics, discussions of current dilemmas, revisions of ethical codes, and monitoring of practitioners and practices are a must for maintenance of the profession's integrity.

REFERENCES

*American Academy of Child Psychiatry. *Code of Ethics*. Washington, DC: American Academy of Child Psychiatry; adopted in 2009, revised in 2014.

*American Psychiatric Association. *The Principles of Medical Ethics with Annotations Especially Applicable to Psychiatry*. Washington, DC: American Psychiatric Association; 2013.

Anzia DJ, LaPuma J. An annotated bibliography of psychiatric medical ethics. *Acad Psychiatry*. 1991;15:1–17.

Bloch S, Green SA, eds. *Psychiatric Ethics*. London: Oxford Publishing; 2010.

Carandang C, Santor D, Gardner D, Carrey N, Kutcher S. Data safety monitoring boards and other study methodologies that address subject safety in "high risk" therapeutic trials in youths. *J Am Acad Child Adolesc Psychiatry*. 2007;46:489–492.

Carlisle J, Shickle D, Cork M, McDonagh A. Concerns over confidentiality may deter adolescents from consulting their doctors. A qualitative exploration. *J Med Ethics*. 2006;32:133–137.

DeJong SM. *Blogs and Tweets, Texting and Friending: Social Media and Online Professionalism in Health Care*. Philadelphia, PA: Elsevier Publishing, 2013.

Deveaugh-Geiss J, March J, Shapiro M, Andreason PJ, Emslie G. Child and adolescent psychopharmacology in the new millennium: a workshop for academia, industry and government. *J Am Acad Child Adolesc Psychiatry*. 2006;45:261–270.

*Dingle A, Stuber M. Ethics education. *Child Adolesc Psychiatric Clin N Am*. Philadelphia, PA:

Elsevier Saunders; 2008;17(1):187–207.

Dyer A. *Ethics and Psychiatry: Toward Professional Definition*. Washington, DC: American Psychiatric Press; 1988.

*Enzer N. Ethics in child psychiatry—an overview. In: Schetky DP, Benedek EH, eds. *Emerging Issues in Child Psychiatry and the Law*. New York: Brunner/Mazel; 1985:3–21.

Fuentes J, Martin-Arribas M. Bioethical issues in neuropsychiatric genetic disorders. *Child and Adolescent Psychiatric Clinics of North America*. Philadelphia, PA: Elsevier Saunders; 2007;16(3):649–661.

Group for the Advancement of Psychiatry. How Old Is Old Enough? *The Ages of Rights and Responsibilities*. New York: Brunner/Mazel; 1989.

Hepper F, Fellow-Smith E. Off-label prescribing in a community child and adolescent mental health service: implications for information giving and informed consent. *Clin Manage*. 2005;13:29–34.

Leckman J, Leventhal B. A global perspective on child and adolescent mental health. *J Child Psychol Psychiat*. 2008;49:221–225.

McFarlane W, Levin B, Travis L, et al. Clinical and functional outcomes after 2 years in the early detection and intervention for the prevention of psychosis multisite effectiveness trial. *Schizophr Bull*. 2015;41(1):30–43. Available at http://schizophreniabulletin.oxfordjournals.org/content/early/2014/07/25/schbul.sbu108.full

Miller RB, Veatch RM, Smith DH. *Children, Ethics, and Modern Medicine*. Bloomington, IN: Indiana University Press; 2003.

Omigbodun O. Developing child mental health services in resource-poor countries. *Int Rev Psychiat*. 2008;20:225–235.

Parens E, Johnston J. Troubled children: diagnosing, treating and attending to context. A Hastings Center special report. *Hastings Cent Rep*. 2011;41(2):S1–S32.

Ptakowski KK. Advocating for children and adolescents with mental illnesses. *Child Adolesc Psychiatric Clinics North Am*. 2010;19(1):131–138.

Recupero PR. Ethics of medical records and professional communications. *Child Adolesc Psychiatric Clini North Am*. 2008;17:37–51.

Roberts LW, Dyer AR, Belitz G. Caring for children. In: Roberts LW, Dyer AR, eds. *Ethics in Mental Health Care*. Washington, DC: American Psychiatric Publishing; 2004:119–135.

Rosenberg E, DeMaso DR. A doubtful guest: managed care and mental health. *Child Adolesc Psychiatric Clin North Am*. 2008;17:53–66.

Schetky DH. Boundaries in child and adolescent psychiatry. In: Schetky DH, ed. *Child and Adolescent Psychiatric Clinics of North America*. Philadelphia, PA: Saunders; 1995;4(4):769–778.

Schetky DH. Conflicts of interest between physicians and the pharmaceutical industry and special interest groups. In: Dell ML, ed. *Child and Adolescent Psychiatric Clinics of North America*. Philadelphia, PA: Saunders; 2008;17(1):113–125.

Sondheimer A. Ethics and risk management in administrative child and adolescent psychiatry. *Child Adolesc Psychiatric Clin North Am*. 2010;19(1):115–129.

*Sondheimer A, Jensen P. Ethics and child and adolescent psychiatry. In: Bloch S, Green SA, eds. *Psychiatric Ethics*. London: Oxford Publishing; 2010:385–407.

Sondheimer A, Klykylo W. The ethics committees of the American Academy of Child and Adolescent Psychiatry and the American Psychiatric Association: history, process, education, and advocacy. In: Dell ML, ed. *Child and Adolescent Psychiatric Clinics of North America*. Philadelphia, PA: Elsevier Saunders; 2008;17(1):225–236.

Soulier MF. Ethics of child and adolescent forensic psychiatry. In: Benedek E, Ash P, Scott C, eds. *Principles and Practice of Child and Adolescent Mental Health*. Washington, DC: American Psychiatric Publishing Inc.; 2010:13–27.

Spicer CM, Graber GC, Figueroa P, Goldstein D, Kimura R. Codes, oaths, and directives related to bioethics. In: Reich WT, ed. *Encyclopedia of Bioethics*. Vol. 5. New York: Macmillan; 1995.

United Nations. General Assembly. Convention on the Rights of the Child, 20 November 1989. *Annu Rev Popul Law*. 1989;16:95, 485–501. Available at http://www.unicef.org/crc.

Vitiello B. Ethical considerations in psychopharmacological research involving children and adolescents. *J Clin Psychopharmacol*. 2003;171:86–91.

Walter G, Rey JM, Soh N, Bloch S. Publishing ethics in child and adolescent psychiatry: essentials for authors and readers. *Child Adolesc Psychiatric Clin North Am*. 2008;17:149–163.

Walter G, Rey JM, Thomas CR, Martin A. Conflict of interest. In: Martin A, Scahill L, Kratochvil CJ, eds. *Pediatric Psychopharmacology Principles and Practices*. 2nd ed. New York: Oxford University Press; 2011:738–751.

Zito J, American Society of Clinical Psychopharmacology. Pharmacoepidemiology: recent findings and challenges for child and adolescent psychopharmacology. *J Clin Psychiatry*. 2007;68:966–967.

▲ 55.8 School Consultation

JEFF Q. BOSTIC, M.D., ED.D., SHARON HOOVER STEPHAN, PH.D., AND ALEXA BAGNELL, M.D.

INTRODUCTION

Approximately 55 million youth attend the 130,000 public and charter elementary and secondary schools across the United States. This positions schools and psychiatry to be vital partners in both responding to emerging student psychopathology and configuring a positive environment for children to develop lifelong coping skills and resilience. For most children, school is the one natural setting where professionals are consistently available to monitor how they are functioning and learning, and to intervene when necessary. School psychiatry increasingly assists schools in early recognition of youth with difficulties to prevent future worsening of problems. Psychiatrists' roles within schools vary and range from the traditional clinician developing and advocating a treatment plan for a patient, to systemic, mental health promotion, early intervention and prevention programming for the school itself.

HISTORY AND LEGISLATION INFLUENCING MENTAL HEALTH IN SCHOOLS

Mental health clinicians have collaborated with school personnel for over 100 years. Sigmund Freud advocated that psychoanalysis had educational as well as clinical benefits for society, and Anna Freud, trained first as a teacher and later as a child psychoanalyst, consulted with educators to help them better understand student behaviors and classroom dynamics. In the latter half of the 20th century, mental health consultation to schools underwent expansion stimulated by broad sociocultural events. First, Gerald Caplan used school-based interventions to assist teachers working with refugee students displaced after World War II. Second, the civil rights movement in the 1960s led to the passage of federal educational rights legislation. This legislation initiated systematic identification of students with emotional and behavioral disabilities, and created a need for clinical advice about assessing and treating these students in the least restrictive school settings. Third, the increased recognition since the late 1960s of problem behaviors among students, such as substance abuse, unprotected sexual intercourse, and more recently, bullying and school violence, caused schools to turn to mental health clinicians for advice. Fourth, the decrease in the 1990s in psychiatric hospitalizations and residential placements, coupled with the lack of access to mental health care among those most in need, resulted in proliferation of school-based mental health treatment, and delivery of psychiatric services within schools, as well as "transition" programs, to assist students coming from inpatient units back into school life. Technological improvements, including telepsychiatry access in schools, and monitoring "apps" to support students have provided additional mental health access, and have facilitated home-school-community coordination and delivery of care. Fifth, federal legislation, such as the No Child Left Behind Act of 2001, mandated the education system to meet the needs of all children and youth in providing appropriate supports for academic success. The President's New Freedom Commission report of 2003, *Achieving the Promise: Transforming Mental Health Care in America,* recognized the importance of mental health in learning and social/emotional development and outlined recommendations for addressing gaps in the provision of mental health services to youth, including a specific recommendation to expand school mental health programs. In 2004, the Garrett Lee Smith Memorial Act (P.L. 108–355, October 2004) became the first federal legislation in the United States to provide specific funding for youth suicide prevention and early intervention programs, to intervene, treat, and prevent mental and behavioral health problems that can lead to suicide in young people. Sixth, psychiatrists have been enlisted more recently in responding to traumatic events impacting or occurring at schools, such as natural disasters (e.g., surrounding hurricanes or tornados), or catastrophic human events such as murders or war. Finally, increased proactive, resilience-building for students, facing seemingly ever-increasing stress loads, have expanded the range of mental

health efforts at school, including social–emotional learning curricula, mindfulness, and executive skill coaching.

MENTAL HEALTH PROBLEMS AND SCHOOL OUTCOMES

Approximately 20 percent of students aged 6 to 17 years currently suffer from significantly impairing psychopathology. Less than one-fifth of these students receive any treatment for these disorders and only 1 in 10 with substance abuse. Minority groups and the uninsured are less likely to receive mental health services. One-quarter of public schools have no counselors, and over half do not have a social worker, yet schools are being asked to deal with more of the mental health needs of their students. Of those youth receiving services, 70 percent report school as a primary source.

Youth mental health has significant impact on school outcome. First, mental health impacts all students in a school. Most notably, Hattie's exhaustive meta-analysis of variables associated with academic achievement revealed that disruptive behaviors at school and emotional distress are among largest impediments to all students' academic achievement in a school setting. School dropout rates continue to approximate 10 percent and are significantly higher in youth with mental health problems. The cause of not completing school is attributable to an underlying psychiatric disorder in half the cases. Reports of increased bullying, including cyberbullying, and school shootings accent the importance of recognizing and responding to the mental health needs of students.

School psychiatry has expanded beyond individual mental health problems to address school wide and community issues including mental health promotion, school violence, bullying, substance abuse, and discipline. Psychiatrists and other mental health professionals continue to refine child psychiatry's role in schools incorporating business and educational principles in effecting change and improving system functioning. Modern school consultation focuses more on early identification and intervention at the individual and system level to help attain short-term educational and behavioral goals and to prevent later long-term negative outcomes. Mental health literacy has emerged as a focus for children to learn how to optimize and maintain good mental health, to understand mental disorders and treatment, to decrease stigma, and to enhance help-seeking when needed, and is being developed actively in Canada (http://teenmentalhealth.org/curriculum and http://www.myhealthmagazine.net).

Schools today face significant demands to inculcate ever-increasing knowledge, to prepare students with more complex vocational skills, to socialize children, amidst increasingly virtual communities, to interact with others appropriately and increasingly, and to maintain their health and

safety. These varied, often competing agendas impact every intervention proposed by the school consultant. Moreover, public, private, parochial, and charter schools may not only vary widely in their priorities but also in their system hierarchy and their accountability to the community. Familiarity with each individual school, its priorities, and its staff is a prerequisite to any meaningful consultation. Accordingly, evaluating the "school" itself often helps clarify what unique school strengths or weaknesses should be factored in to recommendations, as well as resources that may expand or limit options for interventions. This is summarized in Table 55.8–1.

Roles of School Personnel

As the demands on schools have intensified, staff roles have changed such that consultants must have realistic expectations about who can perform specific interventions with students as recommendations are provided. The consultee may be anyone within the school hierarchy, and the consultant must address needs while being mindful of how the consultation will impact everyone else within the system. The consultant must also consider who is in a position to implement recommendations when providing consultation.

Teachers remain the front-line staff most involved with students. Elementary teachers usually have approximately 25 students for 6 hours each day. By middle school, teachers usually teach in a specific content area, providing instruction to approximately 150 different students each day. *Aides* or Paraprofessionals ("Paras") often do not possess a 4-year college degree or teacher certification, but may work under a classroom teacher, or work with a specific student who has different teachers. *Special education teachers* are credentialed to provide alternative instruction to smaller groups of students with learning disabilities such as dyslexia, nonverbal learning disorders, or emotional disorders interfering with learning. For students to receive instruction from special education teachers, they must have an Individualized Education Program (IEP).

School counselors are often the primary mental health support within a school and sometimes provide psychotherapy and crisis support to students, although often their primary role in secondary school is to assist students in academic or vocational planning and class selection. *School psychologists* are either employed by the school district or have a contract with the school to test and help construct IEPs for students to address their specific learning difficulties. They may also provide individual and group therapy to students, but are often working with many schools. *School adjustment counselors and/or social workers* have been added to school staff in many locations, where their primary role is to provide psychosocial interventions to students, and sometimes their families. *School nurses and nurse practitioners* address acute health care needs of students, provide

health education and administer medications to students; school nurses usually travel between several schools, and approximately one-third of their service in schools is mental health care. *Occupational therapists* work with students individually or in small groups to help students with basic activities of daily living, and alternative strategies for students who have sensory integration issues to help them learn to cope with various stimuli. *Speech-language therapists* meet with students who have communication and social skills difficulties individually and in small groups.

Table 55.8–1.
Framework for Evaluating a School

School Component	What to Observe	Potential Questions for Staff
School Building	Safety/Security: Does the building appear safe? Is the building comfortable (temperature, chairs/desks, lighting, noisy)? Does the building prize students (student art on the walls, recognition of student achievements, evidence of student-parent-teacher alliance)? Is this a place a child would want to be?	How does one enter/exit the building? Where do students go if they are having a hard time? Where are students' classwork or projects kept or displayed? What happens after school is over? Do students stay in the building before or after school?
Classrooms	What do classrooms look like? What size are classrooms? How many students are in classrooms, and how many adults? How many learning areas are there, and are they separated so children can be in a quiet place within classrooms? How stimulating (visual, auditory, tactile) are classrooms, and do the students appear over/understimulated?	What's the average number of students in each classroom? How are teachers encouraged to set up their classroom? If a student is having a hard time in the classroom, where does he/she go? How did you (teacher) decide what to put up in your classroom?
School Atmosphere	How is a stranger greeted? Does the school seem organized for students or for staff? Do most students appear engaged (alert, attentive, answering/asking questions) with instruction? How do students and staff interact (smiling, directives, calm, tense)? Do children thrive here?	Who should I meet when I enter the building? How do students move between classes/to lunch/recess? What do students do when they've finished classroom assignments? What do staff expect from students here? What do staff most worry about here regarding students?
School Staff Administrators	Are administrators present/accessible? What kind of tone/impression does the administrator convey? Do staff appear comfortable around the administrator? Who does this administrator best serve (students, teachers, parents, other	What kind of interactions does the administrator have with students (discipline, earned reward time, common interests discussed in halls/lunch)? What is the administrator's priority in this building?

	administrators)? What led to this administrator being selected for this building (student needs, up/down move for this administrator)?	What do staff seek from the administrator (support, camaraderie, ideas, discipline, avoidance)?
Teachers	Do teachers want to be in this school (eager to be with students or staff)? Do teachers stay in this building? How do teachers engage students (time to work, demonstrate content, model enthusiasm, surprise students)? How often do teachers alter the instructional approach (every 10 minutes, something different [lecture, student reading, class discussion], different modalities [visual, auditory, tactile] employed) during instruction?	How many teachers left this building last year? Average over the last several (5) years? What is the student: teacher ratio? What's the average length of time teachers have been in this district/building? How many teachers here have advanced (masters, doctoral) degrees? What kind of teachers do best in this school (independent, orderly, collaborative, creative)?
Special Staff	Are other school staff present? How do teachers and other school staff interact (take student out, co-teach, friendly/tense)?	What kind of special educational staff work here? How often are they in this building (always, weekly)? How do teachers describe other staff (particularly contributions of special educators)? What kind of other teachers are in the classroom (parent aides, "paraprofessional" aides)? Do they work with particular students, with everyone in the classroom?
Support Staff	Are support staff friendly? Do they appear open about discussing the school and students? Do support staff work well with other staff?	What kind of support staff are in this building? How long have most of them been here? How are support staff paired up with other staff?

School administrators coordinate between and within the school and the community. *Principals* and *vice-principals* manage all services (teaching, discipline, custodial) within their school building and report to the *superintendent*. The superintendent guides educational activities among all the schools within a school district and reports to an elected school board in public schools, or an appointed board in private, parochial, or charter schools.

Role of Psychiatrist in Consultation

The contact for the school will depend on the consultee, the question, and the goals of the consultation. Systemic and school-wide issues require the support and involvement of the school administrators, as well as school staff. Individual student issues may only require the teacher and/or

guidance counselor becoming involved, though may involve administrators and other support staff depending on the nature of the consultation. The consultant should approach communication as bidirectional, with the goals of gaining important expertise from the school in terms of academic, social and behavioral observations, as well as sharing of mental health expertise and evidence-based practices and treatment within the limits of confidentiality and taking into account relevant legal and ethical factors.

SCHOOL CONSULTATION FRAMEWORK

The psychiatric school consultant should consider how *micro problems can lead to macro solutions,* that is, how individual student problems can be addressed to improve circumstances for all subsequent students and staff. Five components should be considered in every consultation:

1. Who is the actual **Consultee,** and what are the *confidentiality* parameters?
2. What is the Consultation **Question,** and what is the consultee seeking and wishing will happen?
3. How is the larger **System** experiencing this problem?
4. What **Legal or Ethical Factors** should be considered?
5. How does the consultant understand this problem **Biopsychosocially**?

The Consultee and Confidentiality Parameters

The consultant must establish procedures to clarify who is appropriate to contact or to evaluate. When requested to meet with students, the consultant must obtain parental permission before interviewing any student. If parents refuse for the consultant to evaluate their child, the consultant may still observe, unobtrusively, the student, and consult to staff about their questions on how best to work with this student. Clarification of confidentiality should occur, sometimes recurrently, with staff, students, and parents. Anyone meeting with the consultant should understand how information is shared and whether written clinical information is placed in the student or personnel records.

Clarification of the Consultation Question and Needs/Wishes of the Consultee

Consultants help others define the problem and envision solutions. While consultees may ask for interventions to assist a student, sometimes they actually wish that the consultant will address administrators or recommend the student be removed from this consultee's classroom. Consultants must "read between the lines" to clarify both the consultee's overt request and underlying desires. The success of every consultation depends on the consultant understanding the consultee's concerns and wishes, such that both agree on the problem and on realistic goals for any intervention.

The System Reaction to the Problem

The problem will affect various school staff members differently, and each may have different goals for this consultation. Consultants should clarify who requested a consultation, who knows about the consultation, who has consented to the consultation, and who should be informed about the findings for this consultation. The more that differing goals can be aligned, the higher the probability that participants will invest in a proposed solution.

Legal and Ethical Factors

Special educational laws may provide the student different opportunities, yet families or schools may not see the applicability of such laws in a particular case. Parents may decline psychiatric consultation offered by the school, or may refuse to give consent for a psychiatrist to consult with school staff about their child. Attuning to parental concerns, wishes, and fears often helps diminish resistance to consultation, although consultants may still address teacher questions without seeing the child if a parent refuses for consultants to meet with their children. If abuse is suspected, the consultant may help staff articulate concerns to the appropriate agency, or in acutely dangerous situations such as suicidality or homicidality, the consultant can assist consultees in facilitating emergency treatment or warning others who may be at risk. In the case of consultation regarding an individual student, and particularly if a student will be interviewed or evaluated by the child psychiatrist, a written consent form is generally required, signed by the parent and/or guardian outlining permission for release of medical information to the school. Parents cannot, however, prevent a clinician from observing a student and consulting directly with school staff. Legal and ethical ramifications should always be considered for every proposed intervention. Accordingly, unusual or paradoxical interventions, such as recommending students avoid particular teachers or encouraging parents to not help their child with assignments, are not appropriate; rather, the consultant should support collaborative, transparent efforts by the student, family, and school to improve difficult situations.

Biopsychosocial Understanding of the Problem

Child psychiatrists integrate people's biological vulnerabilities to psychopathology, their past experiences, and current family, peer, school, or other social stressors which might contribute to dysfunctional behaviors. Social, interpersonal, or family problems are the most commonly reported mental health problems in most schools. Considering these different influences may provide multiple intervention targets for a particular problem and may call for employment of diverse existing services in the

school and community. For example, a disruptive, hyperactive student may need referral for a medication evaluation, a behavioral plan at school and home, and connection to social groups such as sports teams, choir or band, and summer camps. The resources and skills available will differ between districts, schools, and classrooms and plans need to work within existing resource capabilities and commitments in order to have best chance of success.

SCHOOL CONSULTING TECHNIQUES

Ally

1. *Validate consultees' perceptions before proposing any solutions.* Validating perceptions cultivates consultee trust that the consultant understands their predicament. If in doubt, ask questions or pose solutions as questions ("What would likely happen if we . . . ?").
2. *Bind anxiety.* Everything the consultant does either increases or decreases anxiety within the system. The more people who share the problem, the less anxious each individual will be. While consultants may not know the answer to various problems, thinking through how they would find an answer can diminish consultee's anxiety ("Hmmm, while we don't know everything we need at this point, wonder if we attempt _____, that will at least provide us clarity, or "data" about an approach so that we can figure out what's more (and less" likely to work with this student.").
3. *Create respect for everyone involved.* Minimizing splitting of staff, students, and parents decreases the probability that individuals will oppose an intervention. The consultant can attempt to separate personalities from the problem by focusing on terms used, by identifying the common goals between parties, and by identifying the circumstances each party uniquely faces ("So . . . it seems where we all agree, and can start, is _____.") Similarly, giving credit for ideas to various staff can increase their investment and also demonstrate how diverse perspectives can be integrated.

Align

4. *Find the good intent gone awry.* When students or staff act inappropriately, the consultant can back up to the good intention that motivated the person's (mis)behavior so that self-esteem, and thus willingness to attempt different behaviors, is maximized. For example, if a teacher "yells" at students, the consultant can "backtrack" to the diligent efforts by the teacher to get the student's attention and to feel responsible that the student learns the material. Then the consultant is positioned to examine alternatives (visually signal the child, use a very soft voice, etc.), which will seem less critical of the teacher's yelling.

5. *Help others to see the child (or staff member) differently*. Maladaptive behaviors may be used to solve problems when no other solution is available (e.g., talking in class may be the only way a student "knows" to slow down instruction or to obtain positive attention from others since the student cannot garner positive attention by doing the academic task) ("It sounds like [student] remaining at home allows [student] to monitor the parent's sobriety; suspect we'll need to find another way for the student to be reassured the parent is stable, or the student will continue to report somatic symptoms to check up on the parent, since that seems more a priority for the student right now than school participation.")
6. *Connect others*. Opportunities for various staff to be empowered and to collaborate may optimize benefits to the staff and the student, by utilizing existing services in the school and community. Staff who frequently see the student, and who understand or have compassion for the student, often have most impact on the student, as well as the staff and even the culture of the school. ("It sounds like [staff member] can check in with the student on Mondays and Thursdays, while [different staff member] can check in Tuesdays and Wednesdays—when will these staff members have an opportunity to compare their findings?").
7. *Appeal to shared values by giving people options they cannot argue with*. Explore and identify appropriate desires of students and staff so that interventions can be provided to realize those appropriate goals. For example, identifying that a student seeks to attend college, or hold a particular job, provides a "frame" for posed interventions ("You've indicated you want to attend/work at _____, and they require you show up on time, so we have to practice going to sleep at 10 PM so that you can be ready the next morning.").

Mobilize

8. *At every opportunity, expand the consultee's skills*. Whenever possible, the consultant attempts to assist the consultee in selecting and implementing skills helpful for that situation. The consultant aspires to provide the consultee facility with multiple skills that the consultee can use subsequently independently.
9. *Use the consultee's own words to frame interventions*. Framing interventions with terms used by the consultee, by staff, and by students, increases each person feeling heard and improves the probability of each investing in proposed solutions.
10. *Identify one step up from the current situation*. Consultees may need help seeing the steps in a sequence toward appropriate behavior (e.g., cursing may be "a step up" from physically assaulting others). Even small positive changes generate momentum to achieve greater changes.
11. *Move toward anticipating problems rather than reacting to them*.

Efforts to help consultees see how problems arise, and how they might be prevented in the future can empower consultees to find and face problems early. More importantly, initial reactions may not represent optimal solutions, but instead create additional conflicts or problems.

TIERED RESPONSE TO CHILDREN WITH EMOTIONAL DISABILITIES

The categorization of children into "eligible or not" for special education is undergoing significant revision, as the importance of early detection and timely response has become clear. To enhance school planning for all students, and to accelerate interventions for students who show difficulties accessing the curriculum, including because of emotional or behavioral issues, schools have been encouraged to initiate "Response to Intervention" (RTI) plans for such students. RTI has also been called a "multi-tiered" approach, as RTI is not "part of" special education, but instead a mechanism for addressing students, whether identified as special education or not, who struggle at school. This tiered system recognizes the value of both school and community interventions for students at these various tiers, as summarized in Figure 55.8–1.

Screening

RTI relies on screening to clarify targets for intervention for students. This begins with universal screening for all students, both for academic and for emotional issues that impede learning. If students screen "positive" and require more specialized interventions, then increasingly specific mental health assessments are employed, and for students having significant specific difficulties (e.g., clinical depression, anxiety, ADHD), these students may receive specific assessment tools to clarify targeted needs and to monitor progress. Example assessment tools across these tiers are provided in Table 55.8–2.

In the academic arena, the first tier of RTI emphasizes "scientific" instruction for all students, so that approximately 80 percent of all students would receive solid, effective instruction and make progress, and thus not need any further educational interventions. Tier 2 (and Tier 3) includes students more likely to be identified for special education. Tier 2 includes "pull-out" or small group instruction for approximately 15 percent of students usually having mild difficulties. Tier 3 is for those 3 to 6 percent of students who require much more specialized programming to make progress at school, and these students are often identified also as qualifying for special education. The parallel emotional and behavioral RTI model assumes a similar distribution of students, with approximately 80 percent of students demonstrating successful emotional and behavioral skills with universal mental health promotion and prevention programming, and the remaining students benefiting from varying intensity levels of mental

health intervention.

Interventions

While RTI is not a specific component of special education, it allows immediate response without significant evaluations or team meetings to address more quickly emerging student problems. RTI is not designed to delay provision of services for students requiring special education, but rather as an augmentation to regular education to address individual differences rapidly with scientifically supported techniques. RTI plans are required to be "evidence-based," meaning that interventions shown/proven effective can be implemented, and not just anything the school or parent might devise. A number of tiered programs have such empirical support, have been employed in diverse schools, and are shown in Table 55.8–3.

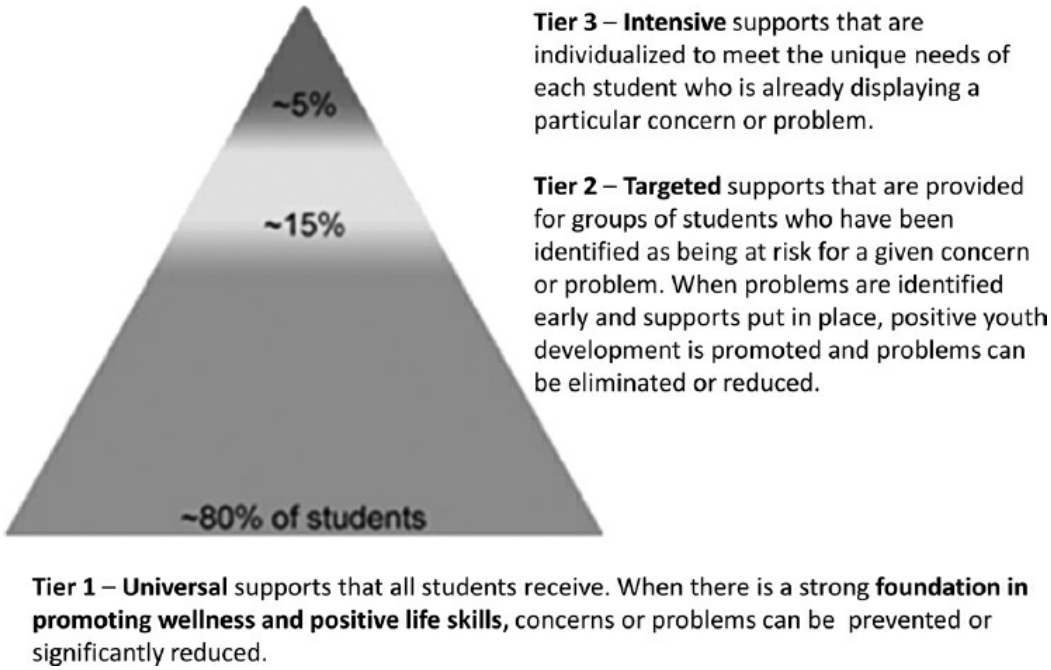

FIGURE 55.8–1. Multitiered system of support. (From Stephan SH, Sugai G, Lever N, Connors E. Strategies for integrating mental health into schools via a multitiered system of support. *Child Adolesc Psychiatric Clin North Am.* 2015;24(2):211–231.)

Accommodations and Modifications for Specific Disabilities

Psychiatrists are most frequently called on to evaluate students in tiers 2 or 3, to determine whether a disability (e.g., emotional, developmental, or medical ["other health impaired"]) is present, and to recommend relevant services, behavior management strategies, or placement needs appropriate for the student. Any child, who is not progressing appropriately in school, has the right to an evaluation to assess whether they have a disability. This can be requested by the teacher, clinician and/or parents, but requires parental consent. Psychiatrists should be aware of the laws protecting students with special education needs, and updates for these in the United

States can usually be found at http://idea.ed.gov.

Students who do not progress through the standard curriculum may qualify for special educational services. Legal protections have evolved to provide every student in the United States with an appropriate education. Although the Fourteenth Amendment to the United States Constitution prohibits discrimination through its equal protection clause, more recent legislation has guaranteed the educational rights of disabled students. *Section 504 of the Rehabilitation Act* (1973) ensures that all disabled children receive a "free and appropriate" public education (**FAPE**) in the "least restrictive environment" (**LRE**). Students qualifying for a 504 Plan may receive special *accommodations* to help them meet educational requirements. Theoretically, *accommodations* refer to classroom changes that help the student meet requirements, without "lowering" academic standards.

Table 55.8–2.
Example Screening Tools for Universal, Selected, and Indicated Student Needs

Tiers	Instrument	Comments
1	Systematic Screening for Behavior Disorders (SSBD) http://store.cambiumlearning.com	Well-validated, efficient, preferred instrument for internalizing issues; inexpensive; normed for grades 1–6 (elementary)
1	Student Risk Screening Scale (SRSS)	Technically adequate, free, quick (15 minutes for entire class), easy to interpret; for grades K-12
1, 2	Strengths and Difficulties Questionnaire (SDQ) www.sdqinfo.org	Well-validated, better for externalizing issues; free, can be scored online to decrease administrative time; for ages 3–17
1, 2	Behavior and Emotion Screening System (BESS) www.pearsonassessments.com	Well-validated, wide range of problem behaviors; can be done online; can be expensive and timely to score; for grades PreK-12
2	Adaptive Behavior Assessment System	Assesses adaptive skills across lifetime; software-based scoring and reporting
2	Behavior Assessment System for Children (BASC-2)	Measures behaviors associated with internalizing and externalizing problem behaviors and academic competence; rules out response bias; web-based screening capacity available
2	Child Behavior Checklist (CBCL)	Well-validated, several subscales (including school subscale) that factor into internalizing and externalizing domains
2, 3	Disruptive Behavior Disorders Scale (DBD)	Assesses the presence and severity of symptoms of inattention, hyperactivity/impulsivity, conduct disorder, and ODD; parent and teacher versions; helpful for diagnostic purposes or evaluating medication effects
3	Revised Child Anxiety and Depression Scale (RCADS)	Subscales for distinct anxiety/depression disorders; can be repeated; for grades 3–12
3	Screen for Child Anxiety Related Disorders (SCARED)	Can be repeated as needed; may need to read to children <11 years

3	UCLA PTSD Index	Part can be repeated to monitor progress; for ages 7–18 years

A more complete list (and copies of the instruments) of free assessment tools is available at http://csmh.umaryland.edu/Resources/ClinicianTools/index.html.

Table 55.8–3. Example Tiered Interventions for School Mental Health

Tier	Program	Goal
1	**Good Behavior Game**	Classroom intervention with a set of evidence-based strategies and a classroom game to increase self-regulation and cooperation and decrease unwanted behaviors
1	**Life Skills Training**	Group or classroom intervention to promote general social skills, self-management, drug resistance, and violence prevention
1	**Project Achieve**	Group or classroom intervention to prevent alcohol, tobacco, and marijuana use and violence
1	**Project Alert**	Group or classroom intervention to prevent alcohol, tobacco, and marijuana use and violence
1	**Second Step**	Classroom curriculum that teaches socioemotional skills to decrease impulsive and aggressive behavior and increase social competence
2	**Check and Connect**	Individualized intervention for students showing warning signs of disengagement and at risk of dropping out. The program monitors school performance and provides mentoring, case management, and other supports
2	**Coping Power**	Group intervention for children at risk of aggressive behaviors, drug use, and delinquency. Uses cognitive-behavioral techniques to teach children how to cope with anger, anxiety, decrease impulsivity, and develop and improve social, academic, and problem-solving skills
2	**Strengthening Families Program (SFP)**	Increases resilience, reduces risk factors to improve social competencies and school performance, and to reduce problem behaviors, delinquency, and alcohol and drug abuse in high-risk children
3	**Aggression Replacement Training**	CBT-based intervention to help children and adolescents improve social skill competence and moral reasoning, better manage anger, and reduce aggressive behavior
3	**Coping Cat**	CBT-based intervention to help children recognize and clarify anxious thoughts and feelings, develop plans for effective coping, and self-evaluate and reinforce positive coping skills
3	**Cognitive Behavioral Intervention for Traumatized Students (CBITS)**	CBT-based group intervention to reduce symptoms of PTSD, depression, and behavioral problems; improve peer and parent support; and enhance coping skills
3	**Interpersonal Psychotherapy for Depressed Adolescents (IPT-A)**	Short-term manualized intervention focusing on developmental and interpersonal needs of adolescents (12–18 years old); builds communication and social problem solving skills
3	**Stark School-Based**	CBT-based intervention for children with depression that teaches self-control techniques, social skills, assertiveness, and relaxation training, imagery,

Additional safeguards have been provided by Public Law 94–142 (1975), revised as the *Individual with Disabilities Education Act* (**IDEA,** PL 101–476, 1990), enacted to provide a free appropriate public education (FAPE) which emphasizes "special education and related services designed to meet their unique needs . . . " The IDEA now provides for educational services to include disabled children under the age of 6, and up to, when necessary, age 21. It ensures that all disabled children receive an Individualized Education Plan with parental input and consent, and includes due-process guarantees if parents disagree with this plan or with the student's placement in an alternative setting or school. The IDEA prevented "troublesome" children from being refused admittance, or from being ejected because the school contended that it did not offer an appropriate program. It goes beyond Section 504 by allowing *modifications,* so the student does not necessarily have to meet the same academic requirements as students in regular education. For example, a wheelchair-bound student would not be required to meet physical education requirements applicable to students without a disability.

The *Americans with Disabilities Act* (**ADA,** 1990) provides additional legal protections for disabled children. The ADA prohibits the denial of educational services, programs, or activities to students with disabilities and prohibits discrimination against all such students once enrolled. Students with suspected disabilities can be identified through observations by their parents, teachers, other professionals, or through school-based screening (e.g., vision/hearing tests and group achievement tests). If parents suspect their child has a disability, they may request an evaluation to determine if a disability is present and interferes with educational progress. This is usually provided in writing to the school's principal. A child with a suspected disability is usually referred to a child study team at the school to provide accommodations. If this proves insufficient, a multidisciplinary team of school-based professionals may conduct a comprehensive individual analysis of all suspected areas of disability. Classroom evaluation/observation, educational and/or psychological testing, medical examinations (including those of psychiatrists), and parental interviews or visits are usually performed to determine eligibility for special educational services. If the parent feels the evaluation is inaccurate, an independent evaluation may be requested, although the school is not "bound" to follow recommendations from an outside evaluator. A framework commonly employed by school districts for determining eligibility is presented in the following section.

Determining Special Education Eligibility

1. Does the child have *one or more of the following types of disability* (documented by the medical evaluation/diagnosis, educational/psychological testing, etc.)?
   Autism
   Developmental Delay
   Intellectual Disability
   Sensory-Hearing, Vision, Deafness, Blindness
   Neurological
   Emotional
   Communication
   Specific Learning
   Other Health
2. If one or more of these disabilities is present, is the child *making effective progress in school*? If the student is being *re-evaluated* to determine if a disability is still impacting the child, would the child continue to make progress in school without the currently provided special education services?
3. Is the lack of progress a *result of the child's disability*?
4. Does the child *require specially designed instruction in order to make effective progress in school* or does the child require related services in order to access the general curriculum?

If the child requires specialized instruction because he or she is *not* making effective progress in school because of a documented disability (includes social and interpersonal progress–not just academic progress), the student is eligible for an Individualized Education Program (IEP) under IDEA. If any of the above conditions are *not* met, the child is *not* eligible for IDEA services, but may be eligible for a 504 Plan or for other services. Section 504 services apply to any person who has a "physical or mental impairment which substantially limits a major life activity" and extends to a wider group of students than those under IDEA. Initially, schools often attempt a 504 plan for students with disabilities minimally impacting educational progress, and then perform IEP evaluations with modifications if the 504 Plan appears insufficient. A child who qualifies for special educational services under IDEA must have a written IEP developed by a team of educators. The IEP specifies in what setting the special services will be provided and this setting must be appropriate to the child's needs and least restrictive of his/her interactions with nondisabled peers. A behavior management plan can be written into their IEP, with the goal of preventing suspensions or expulsions for disruptive behavior. School districts usually attempt to provide special education services on-site or within their district, but when appropriate services are unavailable the school must arrange for an alternative placement out-of-district. Psychiatrists are expected to recommend *services* to address the student's disability, rather

than specific settings or schools. This process is summarized in the example case below. See Table 55.8–4 for specific interventions.

Table 55.8–4.
School Service Plans for Students with Psychiatric Disorders

Type of School Service Plan	Response to Intervention (RTI)	504 Plan	Individualized Educational Plan (IEP)
Purpose	To respond quickly to data that suggests student not making sufficient academic progress	To ensure that all students have equal opportunity to learn, even if have a disability; focus is student's opportunities as compared to other students in that school	To remediate symptoms of a student's disability; student's unique needs are the focus
Criteria to receive this plan	Data (either from universal screening, or from progress monitoring) suggests insufficient academic progress	Student has an impairment that limits a major life activity, but may not require specialized instruction	Student has a disability which interferes with educational progress, and which requires specialized instruction
Who develops this plan	Support staff, teacher and administrator, usually with parental input	Teacher, administrator (often the school's designated "504 Coordinator," school counselor and usually parent, student (if appropriate)	Educational team, including staff certified in special education; may include evaluations by school psychologist, social worker; parent may bring friends, advocates, own evaluators to be part of team
What is usually provided	Changes within classroom (Tier 1) or mental health supports (Tiers 2 and 3) to enable student to perform better based on "evidence-based" supported interventions	Changes within classroom or school building to enable student to complete curriculum expectations	Changes within classroom setting(s) to provide student different instruction, and may substantially alter what is required of student
Example of what is provided	Student is allowed to sit closer to teacher during instruction; student is met arriving at school by familiar staff to decrease anxiety	Student is allowed more time to complete tests; student may be provided device to hear better	Student may leave regular language arts class and receive specialized reading program; student may be exempted from course requirements
Which staff deliver	Usually regular education staff,	Usually regular education staff	Staff with specialized training (special education

services	though support staff may deliver services to students at Tiers 2 and 3		teachers, speech therapists, occupational therapists, etc.)
Where the student receives services	Regular classroom, small group or individual support depending on response to interventions at each Tier	Regular classroom with regular peers "to the maximum extent appropriate"	Wide ranging, from regular education classrooms (inclusion) to pullout for special education classrooms, to off-site day school programs, to 24 hours/day residential schools
Review of the plan	As frequent intervals using ongoing progress monitoring data	Plan reviewed at least every year	At least every year plan is reviewed, and every 3 years the student is retested to see if still qualifies
Disciplinary actions	Usually not applicable	If "manifestation hearing" indicates student's impairment or disability caused misbehavior, then student cannot be suspended/expelled; school is not required to provide free, appropriate education for suspended or expelled students	If "manifestation hearing" indicates student's disability caused misbehavior, then student cannot be suspended/expelled; if student is suspended or expelled, school must still provide free, appropriate education
Appeal recourses	None provided	School may alter 504 Plan immediately should circumstances indicate need; "notice" may be provided verbally; family may appeal to the Office for Civil Rights if perceive school is discriminating against child because of a disability	School must provide "prior written notice" before changes in educational plan or placement are made; family may appeal decisions or plan to local then state departments of education

DT, a 12-year-old 6th grader, struggled at middle school, not keeping up with assignments, losing materials and homework, and making inappropriate comments to other students and teachers. The student performed well on cognitive testing, with verbal skills at the 80th percentile, nonverbal skills 77th percentile, working memory 50th percentile, and processing speed at 37th percentile. There were no traumatic events described, and the student was typically happy, although inconsistent in class, and other peers were distancing from DT because she was talking at the wrong time and not picking up on social cues. The family reported the student had done well in elementary school, but was now overwhelmed and needed a specialized school placement. These symptoms and difficulties in working memory alerted the child psychiatrist to ADHD, which usually evolves by elementary school. In this case, the child had been able to manage effectively enough at elementary school. As the social, academic, and organizational demands increased markedly at middle school, the child was now experiencing greater impairments. The school worked with the family to encourage the family to discuss with their primary care clinician evaluating DT for ADHD (and possible treatments), while the school enacted accommodations for DT's inattention, disorganization, and social skills.

SCHOOL ACCOMMODATIONS FOR MENTAL HEALTH PROBLEMS

Accommodations and Modifications for Students with Selected Psychiatric Disorders

According to the National Comorbidity Survey, an estimated 7.2 million Americans prematurely terminated their education because of early onset psychiatric disorders. The failure to complete educational milestones has negative implications from an individual and societal perspective. It is associated with an increased risk of behavioral and legal problems, peer and social difficulties, psychiatric disorders including substance abuse and depression, and effects on later-life opportunities including occupational achievement, financial security and health and well-being lifestyle behaviors. Conversely, children who do well in school appear to develop factors protective against such negative outcomes.

Psychiatrists may be asked for specific classroom and school environment modifications by the school, the legal system, and/or the parents. The sample interventions in Table 55.8–5, derived from www.schoolpsychiatry.org, may be appropriate for particular students having various psychiatric disorders, depending on the needs of the child and the resources of the school. Such interventions should always be paired with the individual needs of each student rather than categorically applied to every student with that disorder, so interventions addressing several comorbid conditions are often required. Interventions may be adapted to a student's individual service plan, as modifications/accommodations, or more informally implemented in regular classroom settings. Success or failure of these interventions often helps clarify whether a child can remain in the current school environment or requires a different, usually more restrictive, educational setting. While students are entitled to an education in the least restrictive environment, this parameter may necessitate a student receive instruction in a regular classroom, receive some instruction outside the regular classroom, receive all instruction in a self-contained special education classroom, go to a program at another site, or attend an off-site 24 hour/day residential school. The psychiatrist's report will be most helpful if it provides the child's diagnoses ("rule-out" diagnoses are not helpful to schools), appropriate accommodations for symptoms that impair school progress, recommendations for services (e.g., social pragmatics training through speech therapy), and program recommendations (e.g., staff trained to work with students with autism spectrum disorders). From these recommendations, the student's educational team can collaborate with other providers and staff to construct an individualized plan, and then to identify a specific placement to address the student's unique needs. Table 55.8–5 provides examples of these interventions.

Delivery Models of Mental Health in Schools

Direct service models typically reflect a partnership between communities and schools to support an array of mental health services in the school, and/or linked to the community. School-based health centers (SBHCs) and comprehensive school mental health systems (CSMHSs) are two models that reflect such partnerships, providing mental health care to students via a collaboration with community providers.

School-Based Health Centers (SBHCs). School-based health care is a health care service model that is designed for children and adolescents and is responsive to challenges such as inadequate or inconsistent access to primary care, restricted access to mental and oral health services, cultural and linguistic isolation from the mainstream health care delivery system, and lack of insurance or underinsurance. SBHCs provide children and adolescents with access to high-quality, comprehensive clinical and preventive health care that is culturally and developmentally appropriate and delivered in a setting that is trusted, familiar, and immediately available—their school. SBHCs integrate into schools, building an interdisciplinary team of providers who deliver prevention and integrated care that effectively meets the complex medical, behavioral, and social health needs of children and adolescents. Currently, SBHCs serve approximately 1 percent of schools in the nation, though the Department of Health and Human Services has called for a marked increase in their numbers (30 percent increase in the next 3 years). SBHCs are found now in most States, within rural, urban and suburban communities, and across elementary, middle, and high school settings. Public and private resources have been targeted to their establishment in low-income and underserved communities where health disparities are more likely to exist and families have limited access to health and mental health services. Evaluations to date indicate that SBHCs significantly improve children and adolescents' access to and utilization of medical, mental health, and substance abuse services, particularly among disadvantaged populations. While SBHCs provide primary care, approximately 75 percent include mental health as a component of service delivery, and psychiatric consultants often serve students in a treatment capacity, as well as system consultation and care management.

Comprehensive School Mental Health Programs (CSMHPs). Comprehensive school mental health programs (CSMHPs) provide a full array of tiered services, including universal mental health promotion activities, selective prevention, and indicated early intervention services such as clinical assessment and treatment. CSMHPs rely on collaborative partnerships between school systems and community programs so that school-employed mental health professionals are meaningfully supported

by community mental health programs. Increasingly, community mental health staff partner with school-employed mental health staff and educators to provide this more comprehensive integrated model. From a national sample of 1,064 school districts, around 50 percent reported contractual/formal agreements with community-based agencies to augment service provision. Community mental health program partners may include community mental health centers, hospitals, and universities. The shared goals of these partnerships emphasize improving student behavior, addressing barriers to learning, increasing access to mental health care, and improving student attendance and academic achievement.

Table 55.8–5.
School-Based Accommodations/Modifications

Target Symptoms	Interventions To be Implemented by Teacher/School Staff	Examples Specific Tactics for Staff to Employ
ANXIETY		
Worry	Forewarn the student of transitions, and have "tasks" for the student to focus on during transitions	If the student is worried about a school trip, provide tasks that distract from anxiety, such as checking attendance, or holding the door at the site.
	Devise a desensitization approach agreeable to the student	If the student fears speaking in front of the class, allow the student to: have the speech read by a peer; read the speech into a recorder outside class; do the speech with a peer reading some part.
	Help the student examine other perspectives	The student says "I can't go to the school assembly because everyone will notice that I'm nervous." Ask the student: "how would your best friend/someone you admire handle a situation like this? What does your friend think you should do?"
	Provide the student with competing responses to negative thoughts or behaviors	The student says "I'm afraid I'll start crying in class." Ask the student: "if you start to feel sad, what can you do before you start to cry? Can you distract yourself by doodling?"
Emotionality	Provide specific steps the student can take to relax or provide a relaxation ritual	Take three deep breaths; tense fingers or toes for five seconds, then relax.
	Provide alternative foci to distract the student from somatic symptoms	Provide the student a phrase to think of or an activity (doing three problems then standing up, 10 problems then walking to the fountain).
Separation Difficulty	Provide desirable activity/responsibility for the student upon entering school	Before school begins, allow the student to feed fish, clean boards, play with peers, or discuss music/sports with another student who shares interests.
	Provide times during	Have a brief script/message the student can phone to

	the school day for the student to convey brief (30 seconds) messages to his/her family	a parent ("Hi Mom, made it through Science–now we're going to make penguins in art class. I love you.); prepare the parent not to over-react to tearful messages.
	Have a parent send notes to the student to read as a reward for staying at school for increasing intervals	Have a parent place a brief note to the student in his/her lunchbox that the student can read at lunchtime ("Dear [name], I know you're at lunch now–enjoy the cupcake. Have a great afternoon, and when I pick you up today we'll get to play soccer. Love, Dad").
Social Fears	Allow the student to observe several other students before attempting a task	If the student resists speaking in class, allow the student to observe others, focusing on how they start, how long they speak, where they look, and how they stop.
	Have the student rehearse social skills in a smaller or more relaxed setting	In a small group facilitated by a counselor, have students review and role play how to make and keep friends. Give students homework to practice skills in other settings (classroom, playground, home play date, etc.).
Obsessive Thoughts	Establish acceptable teacher comments to "unstick" the student when he/she is obsessing	With the student and parents, identify useful statements to break an obsessive cycle ("Move to your seat on 3 – 1, 2, 3." "Now think about your coach – what would he/she say to your "rut," "stuck moments?").
	Allow the student to dictate or tape record if he/she cannot touch the pencil or paper	Teacher or voice dictation software can "transcribe" the student's ideas to avoid touching/erasing the paper.
Compulsions	Allow the student to alter the work sequence	If the student gets stuck doing problems in a certain way, allow the student to start with even numbers, or start from the end and work backwards to #1.
	Have the student identify and substitute less disruptive compulsive behaviors	Allow the student to touch underneath the desk, flex fingers, or do versions of compulsions that are not disruptive to others.
MOOD SYMPTOMS		
Sad Mood	Check in with the student to quantify his/her mood status each day	Allow the student to complete a mood scale (10 = very happy to 1 = very sad) at first check-in time each morning.
	Help the student identify "all the evidence" surrounding his/her negative perceptions of self or events	Examine specific events that led to the student's conclusion that "I'm no good." What happened that led to this conclusion? Did anything happen that disputes your conclusion? What else happened in your other classes that day?
	Help the student	The student says "I'm no fun. No one wants to be

	identify automatic negative thoughts	around/play with me." Ask the student: "What happened that made you think this?" or "what evidence leads you to reach this conclusion?"
	Acknowledge the student's feelings (rather than dispute/argue with feelings)	"It's sad when . . . .," "discouraging," "frustrating" rather than "it's not so bad," or "aw, come on, things will be better soon."
Irritable Mood	Model appropriate responses to replace irritable responses	Provide an alternative, appropriate comment: "I know you have something important to say, and I want to hear it, but I can't hear your point when you use sarcasm. It sounds like your point (without sarcasm) is _____." Could you please say your point again without relying on sarcasm?"
	Allow the student to take him/herself out of a situation (self-timeout) when irritability is starting to disrupt others	Provide specific steps for the student to remove him/herself from situations ("I [student)] feel myself getting frustrated, so I am going to leave this kickball game and swing on the swings for five minutes. After I watch others, and think I understand what I'm supposed to do, I'll try kickball again.").
	Provide opportunities for the student to "fix" problems or inappropriate classroom behaviors	Allow the student a chance to "redo" with an appropriate comment or behavior; if the student tears up paper, allow him/her to tape it together.
Manic Mood	Allow the student to complete schoolwork or tests in a less stimulating environment	Identify a calm, comfortable area (with limited distractions, noise, and sensory stimuli) where the student can complete work/tests.
	Allow the student alternative modes of expression if he/she cannot verbalize in a useful way or is speaking too rapidly	Provide the student with an opportunity to write, dictate, or draw ideas.
	Allow the student to have homebound instruction during manic periods	Determine the conditions that warrant homebound instruction (e.g., the student is sleeping during the school day while awake at night; the student's fears or delusions are causing conflicts in the classroom). Home instruction may need to be provided by specialized (therapeutically trained) staff.
Appetite Changes	Encourage the student to use snacks and/or physical activity to enhance functioning	Provide snack breaks when the student finishes tasks. Encourage the student to be physically active at appropriate times by taking attendance to the office, standing and moving books, etc.
	Change food reinforcers to nonedibles	Instead of chips or candy, reinforce the student with stickers, computer time, or points to earn a lunch date with a preferred peer (if increased appetite/weight).
Fatigue; Energy Loss	Augment classroom	Tape-record or videotape lectures and provide

	instruction with recordings of instruction	recordings to the student.
	Identify study partners who can support and assist with assignments	Allow a peer to study with the student and to assist in academic assignments.
	Introduce physical activity throughout the day	Provide the student physical responsibilities (taking attendance to the office, emptying trash, putting up class materials) throughout the day to encourage physical activity.
	Grade the student based on work completed or attempted (rather than work assigned)	Grade items completed (e.g., even numbers only), and do not count the items the student does not get to (odd items).
Suicidal or Self-Harm Thoughts	Identify appropriate methods for expressing feelings of hopelessness or self-destruction	Encourage the student to draw pictures, write songs or poems, or use sculpting to depict feelings of sadness, anger or despair.
	Establish a hierarchy of people for the student and staff to contact if the student has suicidal thoughts	Specify multiple staff and treatment providers, and the order or circumstances for contacting them, when the student is feeling or appearing unsafe.
Psychosis Symptoms		
Hallucinations Delusions	Identify and avoid the student's exposure to known distressing stimuli	Clarify with the student or family the items, places, or topics that trigger delusional thinking and find alternatives, particularly when the student is stressed or currently experiencing psychosis symptoms.
	Allow the student alternative schoolwork or activities to avoid provoking delusions	If the student cannot proceed with a task, provide an alternative, "grounding" task which requires little creativity, such as reading, moving items, or doing rote tasks.
	Devise steps to employ when the student is delusional or hallucinating	Employ a series of steps to de-escalate the student when he/she is becoming more delusional (first: change topic, second: change activity, third: change setting, fourth: change staff).
ATTENTION-DEFICIT/HYPERACTIVITY SYMPTOMS		
Inattention	Define classroom expectations in positive terms	"Please sit down, get our your pencils and paper, and let's look at the board to find our first task."
	Give the student duplicate materials	Keep a set of classroom materials in the student's desk and another set at home in case essential items are forgotten.
	Affix materials to the student's desk	Keep regularly needed materials such as pencils attached to the student's desk where they cannot be misplaced.

	Provide organizational devices in the classroom	Use colored folders to distinguish math vs. reading vs. social studies work; use a clipboard to hold papers or current activities.
	Keep a sample model of a correctly formatted paper for the student to refer to	Keep formatted (even laminated) paper showing where name, date, etc., should appear on the student's assignments.
	Have the student repeat directions back to the teacher	After giving directions, have the student repeat back the sequence
	Provide check-in points during lessons	"Raise your hand when you finish the first three problems, and we'll come by to check them."
	Provide untimed or extended time for tests or assignments	Allow the student to work at his/her own pace, even "rest," so he/she can demonstrate all that he/she has learned.
	Diminish external distractions	Minimize noise in the classroom by using headphones, tennis balls on chair legs, or rugs.
	Use a daily progress book or email between school and parents	Send a small book or email between school and home to keep parents and teacher aware of the student's daily progress, and of any events that might influence the student's attention.
Hyperactivity	Clarify volume and movement expectations of the student before unstructured activities	"As we go through the hall, we'll only use our whispering voices if we have to speak, then in the lunchroom we'll remember to use our 'inside' voices."
	Specify acceptable personal space	Clarify space boundaries with visuals such as "each person has the space of four tiles." Put tape on the floor to demonstrate that others stand outside this distance.
Impulsivity	Establish a waiting routine	Develop with the student a multistep plan for waiting, such as "count to five then raise your hand and look the teacher in the eyes."
	When behavior is inappropriate, first remind the student what he/she is expected to do, then reinforce efforts closer to classroom expectations	During misbehavior, clarify what you want the student to do rather than describing the misbehavior ("next time you turn in a paper, make sure your name is at the top" instead of "don't turn in a paper without your name at the top").
AUTISM SPECTRUM SYMPTOMS		
Communication Difficulties	Create situations to motivate language use	Structure situations so that language becomes helpful (choose the flavor of ice cream from a list).
	Allow the student time to process information and respond	Provide sufficient time for the student to understand direction and to respond (wait 10 seconds for student response).
	Ensure that the teacher is positioned strategically to	Align at the student's eye level. Touch the student's desk or chair while quietly saying something like, "Look at the——(checklist or material) and listen. I

	engage the student's attention	need you to know . . . . . ." Pause to give the student time to shift attention.
	Provide choice boards for the student to communicate preferred activities	Show the student different options for tasks (picture of a book [for reading], food [for snack], counting cubes [for math]).
	Develop visual cues to reduce sensory overload	Devise visual cues such as hand signals or use of pictures to diminish reliance on verbal and physical prompts.
	Use scaffolding techniques to promote spontaneous language	Add "parts" (sentence starters, transition statements to connect ideas) to facilitate conversation with the student.
Difficulties with Social Interaction	Pair the student with a "typical" peer/buddy to help carry out social interactions in structured settings	Team the student with a "typical" peer during a structured recess kickball game to show the student how to kick, run, and catch.
	Divide social skills into successive steps and teach the steps incrementally	Break down social encounters and teach multiple ways to accomplish each part (you can introduce yourself by saying your name, by asking what the other person is doing, by showing the other person an interesting object, by just saying "hi," or by having a friend who already knows someone introduce that person).
	Provide explicit teaching about how to start conversations, respond to comments, and end conversations	Describe specific phrases and behaviors to create conversations ("stand this far from a person, look at their eyes, say 'hi', ask if they want to play four square with you," "say 'maybe another time' if they say 'no.')
Restrictive Routines/Interests	Provide alternative tasks, particularly when the student is sensory overloaded	Allow the student several choices within his/her daily tasks. The student can choose the order of tasks, or when/where to take a break.
	Specify the student's routine for asking questions or describing topics when the student seeks or presents information	Explain that the procedure for asking questions is to limit him/herself to two questions, then allow others to ask ("You can ask the two most important questions to begin work.")
Sensory Issues Include extreme over-sensitivity or under-responsiveness to sound, light, or other sources of stimulation	To control sensory inputs, allow the student to be first or last in line or to leave class early	Allow the student to change classes five minutes before the other students, or have him/her hold the door open for other students and then join the end of the line as the "caboose."
	Teach students who use distracting vocalizations or other self-stimulating	Practice humming loudly and softly and have the student role-play appropriate times to hum loudly and appropriate times to hum softly.

	behaviors to employ other intrusive vocalizations or behaviors

Additional specific interventions are available at www.schoolpsychiatry.org.

Table 55.8–6.
School Consultation Websites

http://www.schoolpsychiatry.org: mental health information for school staff, parents, and clinicians; interventions for psychiatric disorders and symptoms; rating scales to assess disorders and monitor treatments.

http://csmh.umaryland.edu: up-to-date information about national school mental health training, practice, research, and policy.

http://www.schoolmentalhealth.org: a repository of useful resources for school clinicians, educators, families, and students on school mental health.

http://smhp.psych.ucla.edu: clearinghouse of important mental health, school, and education materials.

http://ies.ed.gov/ncee/wwc/: information on broad categories of findings of "what works" in schools including academics and mental health.

http://www.ldonline.org: information on classroom changes for students with learning disabilities, including ADHD.

http://www.ideapartnership.org: up-to-date information on changes in the IDEA parameters.

http://www.wrightslaw.com: information about legal aspects of education, including IDEA 504 Plans.

http://www.r14success.org: Response To Intervention school programming across academic and emotional areas.

Many schools have mental health supports less comprehensive than in SBHCs or CSMHPs, yet provide important components of student mental health. Examples include school system consultation, prevention and intervention programs, early identification screening programs; web-based mental health resources, and on-site or video consultation with off-site mental health clinicians. The model adopted by a school often depends on resources, state law, and student and school needs and priorities.

Selection of Evidence-Based Intervention and Prevention Programs

There are many prevention and early intervention models targeting both high-risk and all children in schools. Successful school models integrate the school, family, and community in coordinating services and instituting reform. Common elements in successful intervention and prevention programs include: youth being connected to a trusted adult, access and coordination of appropriate programs and services with ongoing evaluation, program continuity with school support, and emphasis on early identification and intervention. Psychiatrists can help schools identify and implement appropriate programs depending on the needs and resources of the school system. Universal prevention programs address the school system as a whole and aim to better the overall educational climate and

learning environment by coordinating community and outside resources. Specific intervention/prevention models target a wide range of risk behaviors including anxiety, depression, substance abuse, and bullying/violence. Several repositories of evidence-based and promising practices have been developed to assist practitioners and agencies (including school districts and schools) select scientifically supported interventions based on designated criteria (e.g., age, setting, and target outcome). The following sites include several evidence-based school interventions for students with mental health needs.

Websites for Evidence-Based Student Interventions

DELINQUENCY/VIOLENCE/AT-RISK BEHAVIORS:

1. **FindYouthInfo.gov:** This program/site features evidence-based programs that prevent and reduce delinquency or other youth problem behaviors (e.g., drug and alcohol use). Programs focus on one of the following problem behaviors: delinquency; violence; youth gang involvement; alcohol, tobacco, and drug use; family functioning; trauma exposure; or sexual activity/exploitation.
2. Promising Practices Network (PPN) on Children, Families and Communities (http://www.promisingpractices.net/programs.asp): The PPN site features summaries of programs and practices that improve outcomes for children.
3. Blueprints for Violence Prevention Model Programs Selection Criteria (http://www.colorado.edu/cspv/blueprints/model/criteria.html): Blueprints programs have the highest standards by showing evidence of a *deterrent* effect with a strong research design, *sustained* effect, and multiple site *replication*.

SOCIAL-EMOTIONAL LEARNING (SEL):

4. The Collaborative for Academic, Social, and Emotional Learning (CASEL) (http://www.casel.org/programs/selecting.php): "CASEL Select" programs provide outstanding coverage in five essential SEL skill areas, have at least one well-designed evaluation study demonstrating their effectiveness, and offer professional development supports beyond the initial training.

MULTITIERED SYSTEMS OF BEHAVIORAL SUPPORT:

5. Positive Behavioral Intervention System (PBIS): https://www.pbis.org/school/school-mental-health/ interconnected-systems: PBIS provides a school-wide program to enhance positive behavior in the school system.

MENTAL HEALTH AND SUBSTANCE ABUSE:

6. The National Registry of Evidence-based Programs & Practices (NREPP)

(http://www.nrepp.samhsa.gov): Intervention summaries are provided to help determine whether a particular intervention may meet identified needs.

SUICIDALITY:

7. Suicide Prevention Resource Center: Best Practices Registry (BPR) For Suicide Prevention (http://www.sprc.org/featured_resources/bpr/index.asp).

School Violence. The awareness of violence occurring at school (greatest in middle schools) has gained increased attention, and psychiatrists are often contacted to assist in the management of such events. Initial efforts emphasize supporting and empowering the staff and identifying and responding to those in distress, both staff and students. While initially efforts have been to provide universal responses to all students in the school, increasingly, efforts emphasize working with those students and staff most troubled by the event or at-risk. Psychological First Aid for Schools (http://www.nctsn.org/content/psychological-first-aid-schoolspfa), for example, is an evidence-informed intervention designed to assist school staff, students, and their families in the aftermath of a school crisis, disaster, or other violent or traumatic event. The intervention empowers school administrators, educators, and support staff to promote safety and comfort, stabilization, information gathering about current needs and concerns, practical assistance to students and families (including connection with social supports), information about coping and linkages to other community resources.

Psychiatrists are also asked increasingly to evaluate the potential "risk" of students acting violently toward others. A collaborative effort by the U.S. Dept. of Education and the U.S. Secret Service examined 37 instances of targeted school violence and the characteristics of 41 perpetrators. This Safe Schools Initiative (SSI) concluded that the most effective way to prevent an attack at school is to enhance the positive relationship between students and staff. In addition, the SSI clarified the circumstances of school violence: (1) attacks at school are rarely sudden, impulsive acts; (2) other people knew, before the attack, of the attacker's ideas or plans to attack; (3) no "profile" exists to accurately identify which students will attack a school, and any screening approach would both identify an unacceptable and excessive number of potential assailants as well as fail to accurately detect violent perpetrators; (4) most attackers had access to weapons, had felt themselves bullied or mistreated, had struggled with losses or failures and suicidality, and had indicated needs for health; despite prompt response by law enforcement, most events were stopped by other means/entities. For psychiatrists evaluating students at risk of school violence, important topics have been identified to better gauge risk, and are summarized below:

Determining If a Student Poses a Threat

1. What are the student's **motives** surrounding the threat (e.g., retaliate against others, be famous, and have others fear or respect)?
2. Have there been any **communications** suggesting ideas or intent to attack (e.g., disclosures by others, writings, drawings, and social media comments)?
3. Has the student shown **inappropriate interest** in:
   a. *School attacks or attackers*?
   b. *Incidents of mass violence*?
   c. *Weapons* (including recent acquisition of weapons)?
4. Has the student engaged in **attack-related behaviors** (e.g., attempted to make weapons/bombs and harmed animals or others)?
5. Does the student have the **capacity** to carry out an act of targeted violence (e.g., the means to obtain weaponry and confederates to assist)?
6. Is the student experiencing hopelessness, desperation, and/or **despair**?
7. Does the student have a **trusting** relationship with responsible adult(s)?
8. Does the student see violence as **acceptable** or desirable, or the only way to solve problems?
9. Is the student's "story" **consistent** with his/her actions?
10. Are other people **concerned** about the student's potential for violence (e.g., have others alerted school staff, law enforcement, or students about this student?)
11. What **circumstances** might affect the likelihood of an attack (e.g., if some event occurs, or doesn't, holidays/anniversaries of related events [Hitler's birthday], other school assaults)?

References

Barrett P, Lowry-Webster H, Holmes J. *The Friends Program*. Brisbane: Australian Academy Press; 1998.

Benningfield MM, Riggs P, Stephan SH. The role of schools in substance use prevention and intervention. *Child Adolesc Psychiatr Clin North Am*. 2015;24(2):291–303.

Borum R, Cornell DG, Modzeleski W, Jimerson SR. What can be done about school shootings? A review of the evidence. *Educ Research*. 2010;39:27–37.

*Bostic JQ, Bagnell A. School psychiatric consultation: an organizing framework and empowering techniques. School Consultation/Intervention. *Child Adolesc Psychiatric Clin North Am*. 2001;10(1):1–12.

Bostic JQ, Brunt C. Cornered: an approach to bullying and cyberbullying, and forensic implications. *Child Adolesc Psychiatry Clin North Am*. 2011;20(3):447–465.

Bostic JQ, Hart LJ. Building better brains: evidence-based interventions to enhance contemporary schooling. *Child Adolesc Psychiatry Clin North Am*. 2012;21(1):69–80, viii-ix.

Botvin GJ, Griffin K, Diaz T, Scheier LM, Williams C, Epstein JA. Preventing illicit drug use in adolescents: long-term follow-up data from a randomized control trial of a school population.

Addict Behav. 2000;25(5):769–774.

Brener ND, Weist M, Adelman H, Taylor L, Vernon-Smiley M. Mental health and social services: results from the School Health Policies and Programs Study 2006. *J Sch Health*. 2007;77:486–899.

Brock S. *Best Practices in School Crisis Prevention and Intervention*. 2nd ed. Bethesda, MD: National Association of School Psychologists; 2012.

Daly BP, Burke R, Hare I, et al. Enhancing No Child Left Behind-School mental health connections. *J Sch Health*. 2006;76:446–451.

Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol*. 2008;41(3–4):327–350.

Fazel M, Hoagwood K, Stephan S, Ford T. Mental health interventions in schools in high-income countries. *Lancet Psychiatry*. 2014;1(5):377–387.

Fein RA, Vossekuil B, Pollack W, Borum R, Modzeleski W, Reddy M. *Threat Assessment in Schools: A Guide to Managing Threatening Situations and to Creating Safe School Climates*. Washington, DC: U.S. Secret Service; U.S. Department of Education; 2002.

Fonagy P, Twemlow SW, Vernberg E, Sacco FC, Little TD. Creating a peaceful school learning environment: the impact of an antibullying program on educational attainment in elementary schools. *Med Sci Monit*. 2005;11(7):CR317–325.

Garrett Lee Smith Memorial Act, Public Law 108–355, S.2634, 108th US Congress, October 21, 2004. Available at http://frwebgate.access.gpo.gov/cgi-in/getdoc.cgi?dbname=108_cong_public_laws&docid=f:publ355.108

Gillham JE, Reivich KJ, Freres DR, et al. School-based prevention of depressive symptoms: a randomized controlled study of the effectiveness and specificity of the Penn Resiliency Program. *J Consult Clin Psychol*. 2007;75:9–19.

*Hattie J. *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London: Routledge; 2008.

Jaycox LH. *Cognitive Behavioral Intervention for Trauma in Schools (CBITS)*. Longmont, CO: Sopris West; 2003.

Joshi SV, Hartley SN, Kessler M, Barstead M. School-based suicide prevention: process, and role of trusted adults and peers. *Child Adolesc Psychiatric Clin North Am*. 2015;24(2):353–370.

Kam CM, Greenberg MT, Walls CT. Examining the role of implementation quality in school-based prevention using the PATHS curriculum. Promoting Alternative Thinking Skills Curriculum. *Prev Sci*. 2003;4(1):55–63.

Kataoka SH, Nadeem E, Wong M, et al. Improving disaster mental health care in schools: a community-partnered approach. *Am J Prev Med*. 2009;37(6 Suppl 1):S225–S229.

Kessler RC, Angermeyer M, Anthony JC, Graaf R, Demyttenaere K, Gasquet I, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatr*. 2007;(3):168–176.

Kutcher S, Bagnell A, Wei Y. Mental health literacy in secondary schools: a Canadian approach. *Child Adolesc Psychiatric Clin North Am*. 2015;24(2):233–244.

Mattison RE. Consultation interactions between special education teachers and child psychiatrists. *Child Adolesc Psychiatr Clin North Am*. 2001;10(1):67–82, viii.

National Centre for Education Statistics, US Department of Education. Projections of Education Statistics to 2016. December 2007. Available at http://nces.ed.gov/programs/projections/projections2016/ Accessed January 2, 2008.

Rutter M, Maughan B, Mortimore P, Ouston J, Smith A. *Fifteen Thousand Hours: Secondary Schools and Their Effects on Children*. Cambridge, MA: Harvard University Press; 1979.

Stallard P, Simpson N, Anderson S, Carter T, Osborn C, Bush S. An evaluation of the FRIENDS programme: a cognitive behaviour therapy intervention to promote emotional resilience. *Arch Dis Child*. 2005;90(10):1016–1019.

Stein BD, Jaycox LH, Kataoka SH, et al. A mental health intervention for schoolchildren exposed to violence: a randomized controlled trial. *JAMA*. 2003;290(5):603–611.

*Stephan SH, Sugai G, Lever N, Connors E. Strategies for integrating mental health into schools via a multitiered system of support. *Child Adolesc Psychiatric Clin North Am*. 2015;24(2):211–231.

Stephan S, Weist M, Kataoka S, Adelsheim S, Mills C. Transformation of children's mental health services: the role of school mental health. *Psychiatr Serv*. 2007;58:1330–1338.

*Walter HJ, Berkovitz IH. Practice parameters for psychiatric consultation to schools. *J Am Acad Child Adolesc Psychiatry*. 2005;44(10):1068–1083.

Weist MD, Youngstrom EA, Stephan S, et al. Challenges and ideas from a research program on high-quality, evidence-based practice in school mental health. *J Clin Child Adolesc Psychol*. 2014;43(2):244–255.

Weisz JR, Sandler IN, Durlak JA, Anton BS. Promoting and protecting youth mental health through evidence-based prevention and treatment. *Am Psychol*. 2005;60(6):628–648.

▲ 55.9 Prevention of Psychiatric Disorders in Children and Adolescents

WILLIAM R. BEARDSLEE, M.D., DAVID A. MRAZEK, M.D., AND PATRICIA J. MRAZEK, PH.D.

INTRODUCTION

Over the last two decades, there has been significant expansion of empirical research that demonstrates it is possible, in some instances, to prevent mental illness and problem behaviors, particularly in children and youth. In 2009, the Institute of Medicine issued a report focused on children, youth, and families entitled, "Preventing Mental, Emotional, and Behavioral Disorders Among Young People." This report chronicled several major advances since the last IOM report in 1994, including an increased number of efficacy and effectiveness trials, a deeper understanding of risk and protective factors for mental illness, and advances in understanding the neuroscience that underlies preventive intervention. In addition, this report highlighted the fact that the major mental illnesses of adulthood have their onsets in childhood and adolescence, with the onset of symptoms preceding the onset of full-blown disorders, in most cases, by several years.

Such advances in preventive intervention research suggest that it is now possible to set public health goals designed to decrease the incidence of behavioral problems and psychiatric disorders within a given population. Practical strategies for correcting problematic developmental trajectories have been identified, and the salient characteristics of target populations have been defined. Rigorous preventive interventions have been implemented across diverse populations and been shown to be effective. Finally, informative longitudinal studies have followed children into adulthood and demonstrated differential outcomes and the important influences of risk and protective factors at multiple levels. The study of the prevention of psychiatric disorders offers an unusual opportunity to study mechanisms by which disorder is either expressed or prevented and, as such, must be an important part of overall research on development. Preventive intervention research also highlights the influence of broader forces on children's development such as neighborhood characteristics, wealth disparities, and lack of access to resources. Yet, although psychiatrists and other mental health professionals are in key positions to identify those children who are at the highest risk for the development of psychiatric disorders and to institute appropriate interventions, to date, many psychiatrists have been reluctant to embrace developments in preventive intervention. As a reflection of democratic principles and also an emphasis of the Affordable Care Act, it is important that the same quality of

treatment and prevention services be available to all children and families. This requires being able to implement and disseminate evidence-based preventions and treatment widely. A major emphasis of both IOM reports of 2009 was the need to study implementation and dissemination much more fully. This is a focus of a current IOM forum.

Part of the reluctance of psychiatrists and other mental health professionals to engage in advocacy and practice for preventive interventions is that accepting the idea of prevention requires a major paradigm shift from a traditional concentration on the diagnosis and treatment of those seriously ill, in which the focus is on what can be done for the child or adolescent immediately. Prevention requires asking a different question: what does the child, the family, and even perhaps the broader community need 1, 3, and 5 years down the line in order to ensure the best possible trajectories of healthy development for the child and the family? In addition, as the field of prevention has incorporated an increased understanding of genomic variations in the onset of illness, and of the biological, as well as the psychosocial, context of risk and protective factors for illness, preventive interventions are increasingly conceptualized as including both biological and environmental components. The acceptance of preventive interventions has been slow because many of the investigators and clinicians who have been involved with the interdisciplinary development of preventive interventions are not familiar with molecular genetics.

In fact, focusing on prevention requires mental health professionals to recognize that mental health and physical health are inseparable and, hence, preventive interventions in one realm often exert a positive effect in another. Successful preventive interventions are inherently interdisciplinary, drawing on findings and insights from psychology, sociology, medicine, adult and child psychiatry, social work, many medical disciplines, and the field of developmental neuroscience multiple approaches to genetics. Mental, emotional, and behavioral disorders are developmental. That is, they have their onsets at different points across the lifespan, requiring a series of interventions that can be positioned at different points across the lifespan.

The development of a new domain of science can be facilitated by the introduction of new knowledge from related fields that require major reassessment of previous beliefs. This reassessment can subsequently lead to integration of these new insights into the traditional core knowledge base of the science.

Every medical specialty is being challenged to consider how genomic variations influence the onset of illnesses as well as the selection of treatments that can be tailored to be maximally effective for each patient. This new orientation provides an opportunity to begin to practice "individualized personalized medicine." This approach to diagnostic

evaluation focuses on biological vulnerabilities that can be identified in each individual patient. Consequently, preventive interventions, in the same way as therapeutic interventions, will increasingly be conceptualized to include both biological and environmental components. Consideration of molecular genetics is essential in this regard.

This chapter focuses on preventions for children, youth, and families. The current understanding of prevention research is guided by four key principles. First, because improving the health of a population is an essential goal, preventive interventions must eventually be able to be delivered to entire populations. Second, health disparities remain a growing and profound problem, and preventive interventions can help in addressing health disparities. Third, awareness of cultural and linguistic competence and the need for cultural humility must guide the adaptation and delivery of preventive interventions. Finally, given the implementation of the Affordable Care Act with its strong focus on prevention and its emphasis on mental health services as one of the ten essential services, a focus on prevention is timely and necessary.

Definitions

In an attempt to clarify the research evidence and advance the science of prevention, the 1994 IOM Committee on the Prevention of Mental Disorders recommended the adoption of an alternative definitional taxonomy for mental health interventions: prevention, treatment, and maintenance. The use of the term "preventive intervention" in this classification system was restricted to interventions provided prior to a patient receiving a psychiatric diagnosis. The term "treatment" was reserved for interventions targeted at patients with a diagnosis. The term "maintenance" referred to the care provided to patients with chronic illnesses who had not had a remission as a consequence of their primary treatment.

The IOM Committee recommended that preventive interventions, which were previously referred to as primary prevention, be subdivided into three categories, which were designed to be administered to populations prior to the initial onset of a psychiatric diagnosis. These categories were referred to as universal, selective, and indicated preventive interventions, and were differentiated from each other based on their targeted populations. In consideration of the interdisciplinary nature of prevention, this fundamental categorization was continued in the 2009 report.

Universal preventive interventions are targeted to the general public or to a specified group. By definition, universal interventions are provided to individuals who have all levels of risk, including no risk factors at all. A common example of a universal preventive intervention is fluoridation of a water supply to improve the dental health of a community. *Selective*

preventive interventions are targeted to individuals whose risk of developing mental disorders has been demonstrated to be increased. These identified risk factors may be associated with imminent onset or with risk at a later point in development. Biological markers that indicate predisposition for mental disorder but are detected in individuals that are free of symptoms represent an example of medical criteria that can be used to initiate a selective preventive intervention. *Indicated preventive interventions* are targeted to high-risk individuals identified as having detectable signs or symptoms associated with the ultimate development of a mental disorder. Symptoms or signals, which are early warning signs of future development of a psychiatric illness, may arise years before a patient develops a sufficient number of symptoms to meet the full diagnostic criteria that are required to establish the occurrence of the psychiatric disorder. Indicated preventive interventions can be initiated early in the natural course of a mental disorder when the first signs and symptoms appear.

Sufficient evidence has accumulated to include mental health promotion as a fourth major category of preventive intervention (see Fig. 55.9–1). Mental health promotion aims to enhance an individual's ability to achieve age-appropriate developmental tasks or developmental competence and to have a positive sense of mastery and well-being, as well as to strengthen an individual's ability to cope with adversity.

The term *risk factor* is used in both prevention and treatment, but the implications of risk for these two concepts are different. In prevention, a risk factor must predate the onset of the illness and can be the characteristic that is used to identify a target population. The reduction of risk factors that are both causal and malleable is a goal of preventive interventions. Once the diagnosis of an illness has been made, risk factors that are subsequently identified may be used to guide treatment decisions and to provide insight into prognosis. One well-studied example of a risk factor is parental mental illness. Another is undergoing severe loss, such as a family member.

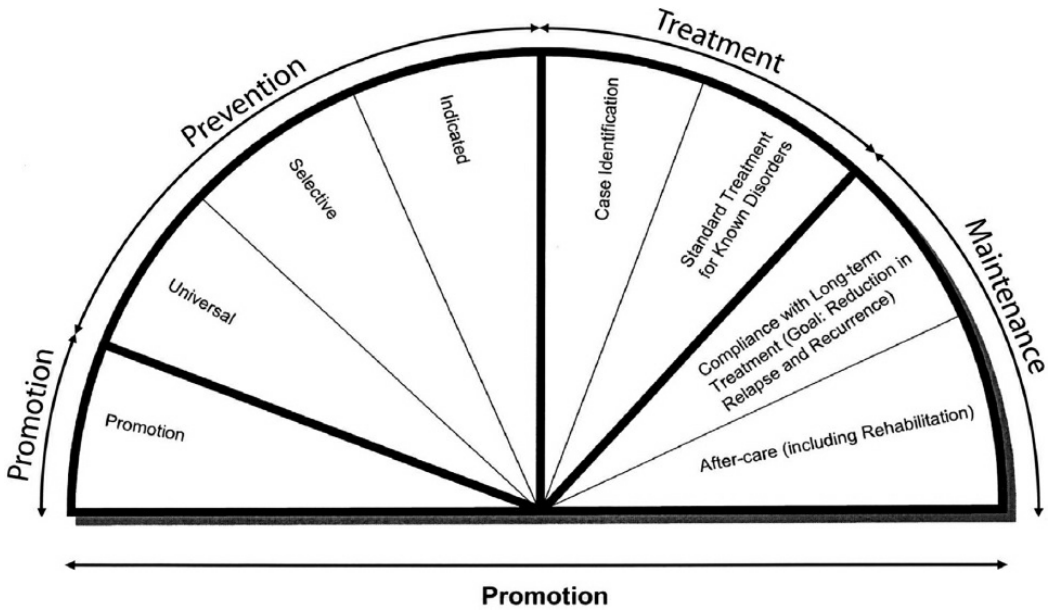

FIGURE 55.9–1. The mental health intervention spectrum for mental disorders. (Reprinted with permission from Mrazek PJ, Haggerty RJ, eds. *Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research*. Washington, DC: National Academy Press; 1994.)

Protective factors decrease the risk of the onset of mental illness. An individual with one or more protective factors is at a lower risk for the development of a specific psychiatric illness than an individual without the protective factor. Most protective factors are nonspecific and have been perceived to be universally beneficial. Common examples of protective factors include a supportive family environment, adequate family financial resources, and access to high-quality education.

Increasingly, research has highlighted the ways in which biological and environmental risk and protective factors interact. The study of epigenetics has emphasized the important ways in which environmental influences can profoundly affect gene expression across generations. Exploration of gene-environment correlations and gene-environment interactions offer important ways to study mechanisms.

Examples of Research on the Prevention of Mental Health Problems in Youth

Figure 55.9–2 presents a schematic representation of the different interventions that are likely to be effective across the span of childhood. Early in life, interventions for high-risk families (e.g., home visitation, center-based care) have been shown to be valuable. As children enter school, both academic enrichment activities and social skills training are valuable, as are various kinds of parent training. As youngsters enter adolescence, there is evidence that some programs can cut the incidence of major depression and substantially reduce substance abuse.

Across the bottom of the diagram, three types of factors that occur across the span of childhood and adolescence are emphasized: the

importance of addressing families in adversity; the importance of interventions that strengthen the fabric of schools and communities; and policy interventions such as those that provide ready access to preventive and treatment services, and that support families. A central theme in the 2009 report is that preventive interventions can reduce or minimize toxic biological and psychological processes, richly reinforce self-regulatory prosocial behavior, teach prosocial skills and values, and foster acceptance. The overall conclusion of the report was, "The scientific foundation has been created for the nation to begin to create a society in which young people arrive at adulthood with the skills, interests, assets, and health habits needed to live healthy, happy, and productive lives in caring relationships with others."

In what follows, rather than trying to comprehensively review all prevention approaches, the focus is on preventive interventions that are particularly relevant to psychiatric practice, and that illustrate more general points about prevention approaches. Specifically, this chapter reviews key interventions targeting family adversity, externalizing behaviors, depression, psychosis, and substance abuse.

Preventions Targeting Family Adversity. Examples of effective parenting programs in risk situations are the effective programs designed and implemented by Wolchik, Sandler, and colleagues at the University of Arizona: The New Beginnings Program (NBP) for families facing divorce, and the Family Bereavement Program (FBP) for families facing bereavement.

Using a randomized, controlled clinical trial design, two well-described preventive approaches were tested in the NBP. The first targeted mothers and consisted of 11 group sessions and two individual sessions; the second provided this same program for the mothers, as well as 11 sessions for the children. Both interventions were compared to a control group where the mother and child received three books on how children adjust after a divorce. In follow-up studies, researchers found very few differences between the mother-only and mother and child interventions, but argue that this was the result of focusing primarily on the outcome of coping. Wolchik and colleagues suggest that the lack of benefit of the child program may be due to little opportunity for the children to practice more adaptive coping strategies and a presentation that may have been too didactic for the young children. The results of a 15-year follow-up demonstrated that the young adults who received the preventive intervention (NBP) were less likely to develop an internalizing disorder than were young adults in the control group during the time period since the 6-year follow-up, and young adults who developed a disorder had a slower rate of onset. Also, between the 6- and 15-year follow-up assessments, males in the NBP group displayed a lower rate of substance-related disorders and reduced

substance use in the past year.

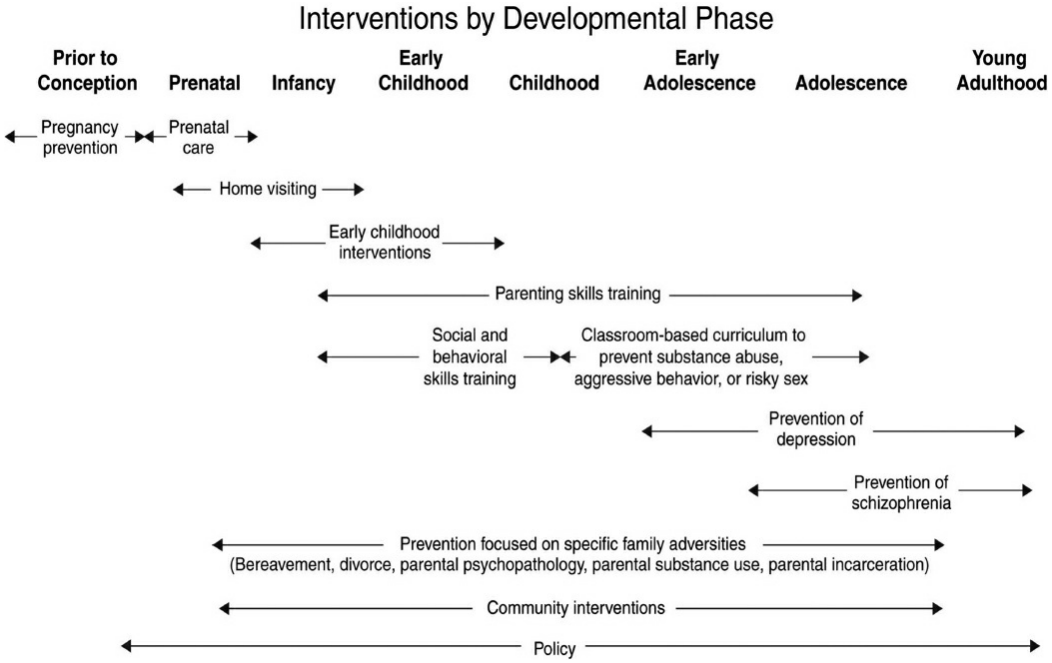

FIGURE 55.9–2. Interventions and their targeted developmental stages.

For families struggling with the loss of a parent, Sandler and colleagues developed the FBP. With a similar structure to the NBP, FBP provides caregivers and children with 12 group sessions, 4 of which include shared activities for both the caregiver and the child or adolescent. Caregivers also participate in two individual sessions. The control group received 3 books, discussing child, adolescent, and adult grief, over the course of 3 months. A 6-year follow-up indicated that effective parenting, operationally defined as caregiver warmth, acceptance, and consistent discipline, had increased in families receiving the FBP. An examination of mediating effects showed that improved parenting at postintervention and at an 11-month follow-up mediated the effect of intervention on the outcome of effective parenting 6 years after the completion of FBP. Although not directly measured in this study, a similar intervention that increased effective parenting had a depression prevention effect on parents, such that parents who participated in the active intervention had fewer cases of depression. These results demonstrate that relatively brief interventions have the potential for long-term sustained effects in families that have experienced a major stressor, such as divorce or the death of a parent. More generally, interventions that are highly specific and increase capacity of parents to parent effectively have a strong knowledge base.

Preventions Targeting Youth Externalizing Behaviors. Webster-Stratton originally developed a treatment intervention for young children with conduct disorder and their families. This intervention, The Incredible Years

Parenting Series, includes therapist-led group discussion that emphasizes parent training. Parents view standardized videotapes that present parents and their children engaging in both desirable and problematic interactions. Findings demonstrate that this program improved parent and child interactions and significantly reduced the children's conduct problems and increased cooperation, compared to a control group. This treatment was later modified to become an indicated preventive intervention for children with early symptoms of conduct disorder but who did not meet diagnostic criteria. More recently, this intervention was modified further to allow it to be used as a selective intervention for children enrolled in Head Start programs who were all considered to be at high risk but were not symptomatic at the time of the intervention. Teachers in the active intervention group used more positive classroom management techniques, and their students had fewer conduct problems and more emotion regulation than control students. Since then, this program has been widely disseminated.

Poduska, Kellam, and colleagues document the long-term effects of a universal classroom–based intervention with children during their first and second grade years. Their intervention, the "Good Behavior Game" (GBG), is directed at improving the classroom management skills of teachers, socializing children to become better students, and reducing aggressive and disruptive behavior. A randomized, controlled trial was carried out in 40 first grade classrooms in 18 elementary schools in Baltimore, Maryland. Significant positive results were found, but these outcomes varied by gender, baseline level of risk, and cohort. Generally, the children with the highest level of risk from the fall of the first grade through the seventh grade benefited more than other children who were assessed to be at lower risk. The most aggressive boys in first grade who received the GBG were the children who demonstrated the greatest reduction in aggressive behavior over the course of first through seventh grades. These boys were less likely to develop an antisocial personality disorder as an adult and had a 50 percent reduction in lifetime illicit drug use. These results point to the benefit of early behavioral interventions.

Preventions Targeting Depressive Disorders. Quite fruitful work has been done in the area of the prevention of depression. There are both specific risk factors for depression, such as having a parent with depression, a depressogenic cognitive style, or a prior depressive episode, and nonspecific risk factors, such as poverty, exposure to violence, and social isolation. Research suggests that children who have either specific or nonspecific risk factors are at increased risk for disorder, and that interventions targeting these risk factors may prevent the onset of depressive symptoms and disorders. Recent meta-analyses support the value of preventive intervention approaches targeting depression and show

that it is possible to actually prevent episodes of major depression in both adults and youth.

An example of a strong depression prevention program is the Coping with Depression Course (CWD), developed by Lewinsohn in the mid-1970s. After finding much success in the reduction of depressive symptoms in adults who received this prevention and treatment course, Clarke and colleagues adapted the CWD for use with an indicated sample of adolescents and demonstrated the prevention of unipolar depressive disorder. Later, Clarke and colleagues used the CWD course with adolescents of depressed parents who themselves had subsyndromal symptoms of depression. Only 9.3 percent of these adolescents went on to develop a major depressive disorder, compared to 28 percent of adolescents who were at similar risk but did not receive the preventive intervention. Garber and colleagues undertook a four site replication of this study, again using adolescents at double risk, because their parents had experienced major depression and they themselves either had prior depression and/or symptoms of depression. This study demonstrated a strong and sustained effect from the prevention program. At a 33-month follow-up, adolescents in the group cognitive behavioral prevention program had significantly fewer onsets of depressive episodes, compared to those in the comparison group; however, for adolescents whose parents were currently depressed at the time of enrollment, the average rate of onset did not differ statistically between the intervention and nonintervention groups.

A second example of a preventive intervention targeting youth depression was developed by Compas and colleagues, who explored preventive interventions for children with depressed parents using a family group cognitive behavioral (FGCB) strategy. Children between the ages of 9 and 15 who had a parent with a history of Major Depressive Disorder participated, along with their parents, in either the 12-session FGCB intervention, or they received 3 sets of written materials about depression in the mail every 8 weeks. Eighteen and 24-month follow-ups demonstrated that the FGCB preventive intervention significantly reduced the onset of the children's depressive episodes for up to 2 years after the intervention ended compared to the control group. Additionally, this preventive intervention reduced the children's internalizing and externalizing symptoms. In another randomized control trial of this intervention, Compas and colleagues found a mediating role of parental behavior, more specifically parental guilt induction. Similar to Sandler and colleagues' FBP, it was found that a reduction in negative parental behaviors improved family functioning; 1 year after the end of the intervention, reductions in parental guilt induction were associated with significantly lower internalizing symptoms and lower levels of negative cognitive style in the youth participating in the FGCB program.

An illustration of a preventive intervention that reduces the risk associated with depression is the Preventive Intervention Project, which targets relatively healthy children, aged between 8 and 15 years, who have a biological parent with a clinical mood disorder. Two manual-based programs were tested. The first intervention consisted of a two-session group lecture format for the depressed parents; the second intervention used a clinician-facilitated format consisting of 6 to 11 sessions, including separate meetings with parents and children, as well as a family meeting. In both formats, psychoeducational information about mood disorders was provided, and parents were helped to better support the initiatives of their children. In the clinician-facilitated format, efforts were made to link information about mood disorders to the depressive illness experiences of the parents in the intervention. At the 4.5-year follow-up, parents who received either the lecture format or the clinician-facilitated intervention reported improvements in their behaviors and attitudes regarding their children, and the internalizing symptom scores of the children who had received either of the interventions improved when they were assessed at follow-up. Moreover, there was a significant relation between the number of positive changes that the parents reported in their behaviors and attitudes regarding their children and the improved understanding of course and treatment of mood disorders that was achieved by their children. This family-based approach is particularly applicable to implementation in a clinical setting, which could be accomplished by adding a relatively small number of additional sessions to the standard treatment of those adult patients with mood disorder who have children in the therapeutic range. Beardslee's family-based depression prevention approach has been adapted successfully for use in different ethnic and cultural groups, both in the United States and around the world.

Prevention of Early-Onset Psychosis. Several investigators have focused on children and youth at high risk for the development of schizophrenia and the prodromal symptoms of psychosis. The goal of much of this work has been to identify risk factors or biological markers that could inform studies designed to prevent the onset of an initial episode of psychosis, or to optimally treat a patient at the very earliest point in the initial onset of the illness. Although these intervention strategies all have been multidimensional, one critical aspect of these interventions is the prudent use of antipsychotic medications. For example, McGorry and colleagues gave young first-episode psychosis patients a low dose of risperidone therapy to treat their early symptoms; results revealed that risperidone was very effective in reducing psychotic symptoms and improved the individual's quality of life, treatment experience, and adherence.

In the United States, with efforts to reduce the duration of untreated psychosis (DUP) from early symptoms to engaging in treatment, Srihari,

McGlashan, and colleagues have developed the Specialized Treatment Early in Psychosis: Early Detection (STEP-ED) program, an adaptation of McGlashan and Johannessen's Early Treatment and Intervention in Psychosis (TIPS) study. The TIPS Early Detection in Psychosis Study, conducted in Norway and Denmark, provided the general public, schools, and Early Detection teams with information about prodromal, first-episode schizophrenia symptoms. This program shortened the DUP by an average of 1.5 years. In the upcoming Specialized Treatment Early in Psychosis project, the researchers seek to adapt the TIPS model for Americans, and hope to improve presentation, engagement, and early outcomes through the reduction of DUP. STEP-ED will employ a public education campaign by utilizing social marketing techniques, including interactive websites, posters on buses, advertisements on social media sites, television, and radio, professional outreach at primary care sites, religious, and educational settings, to promote help-seeking behavior. Srihari, McGlashan, and colleagues will compare the effects of the STEP-ED program to a usual detection program.

Another way to prevent the onset of psychosis, other than reducing the DUP, is to focus directly on the programs designed to detect symptoms early and intervene to prevent psychosis. McFarlane and colleagues created and implemented the Family-aided Assertive Community Treatment (FACT) program and conducted a 2-year follow-up study. Adolescents were assigned to one of three groups, clinically low-risk (CLR), clinically high-risk (CHR), or very early first-episode psychosis (EFEP), based on their score on the Positive Symptom Scale of the Prodromal Symptom Scale. Individuals designated to the CHR and EFEP groups received FACT, which incorporates a variety of interventions, including psychoeducational multifamily group (PMFG) therapy, elements of assertive community treatment, supported education and employment, and psychotropic medication. Those in the CLR group received community care, which consisted of monthly monitoring through phone assessments. Results demonstrated that 2 years after treatment, FACT was effective in reducing symptoms, whether positive, negative, disorganized, or general symptoms. Participants in the FACT condition, who were at risk or experiencing very early psychosis, also had improved Global Assessment of Functioning scores, work and school participation, and overall global outcomes. The findings suggest that preventive interventions can be effective in preventing the onset or progression of psychosis, and that family interventions are extremely important in the treatment and prevention of psychotic symptoms. Given consistent evidence that demonstrates that subsequent episodes of psychosis are associated with a less optimistic prognosis and probable neurodegenerative cerebral changes, the development of preventive interventions for schizophrenia and bipolar disorder is a high public health priority. The identification of children who are at risk as a

result of carrying susceptibility genes for these disorders is a research strategy that has not yet been systematically implemented.

Prevention of Substance Abuse Disorders. Many researchers have developed preventive intervention programs for substance abuse in children and in adolescents. Botvin and Eng developed a universal prevention program, Life Skills Training (LST), which has been shown to be quite successful in reducing substance abuse in middle and high school students. During 30 lessons over the course of 3 years, students participating in LST were taught skills, information, and attitudes to resist temptations to use substances in adolescence. In dozens of efficacy trials, LST has reduced cigarette smoking, alcohol consumption, methamphetamine usage, and marijuana use as much as 80 percent compared to control groups. Similar results have been found with minority group students. Some LST trials have discovered that elementary school programs have positive long-term effects. The LST intervention's effects have been maintained up to 12 years after the completion of the program. Booster program designed for the LST program have demonstrated maintenance of and even bolstered the original intervention's effects.

Likewise, Kumpfer, and colleagues developed the Strengthening Families Program: for Parents of Youth 10 to 14 (SFP). This 14-week, family-focused substance abuse prevention program has demonstrated significantly greater reduction rates of substance use in high school students, compared to school-based alcohol prevention programs. SFP also has generated many significant positive effects in family functioning and parent-child relationships. SFP has been successfully adapted for many cultural groups, such as Hispanic, African American, Native American, and Asian Pacific Islander families, and has been widely adapted internationally.

Now that children and adolescents have greater access to mobile devices and computers, Internet-based prevention programs have become increasingly popular. In a meta-analysis by Champion and colleagues, seven school-based substance abuse prevention programs delivered through either the Internet or CD-ROMs, including Consider This, ASPIRE, CLIMATE, Head On, and several other programs, were assessed for their effectiveness in preventing alcohol, tobacco, and cannabis use in adolescents. Five of the seven programs were found to reduce drug and alcohol consumption in the follow-ups, which were completed between 6 and 34 months after the cessation of the program. Although the reviewers discuss how the effect sizes were modest, the reduction of substance use after the implementation of Internet and computer prevention programs provides evidence for how these new methods for disseminating interventions have the potential to be very effective and can be used by individuals in a variety of urban and rural settings.

Issues in the Prevention of Mental Illness

Barriers to Providing Clinical Preventive Interventions. Despite the very strong scientific evidence about the benefits of preventive interventions, there are relatively few psychiatrists who are involved in the administration of evidence-based preventive interventions. There appear to be three major barriers to the adoption of these interventions. First, there is no clinical infrastructure for delivering preventive and early intervention services. Provision of services to large numbers of children requires the coordination of multiple agencies, disciplines, and service delivery systems. Without a well-developed and accessible clinical infrastructure, most families are unable to receive preventive services even where treatment services exist. A second barrier to the provision of early clinical intervention is that there are insufficient resources to pay for these services. The lack of reimbursement through either public payment plans such as Medicaid or by private insurers prevents most psychiatrists and other mental health clinicians from providing preventive services. A third barrier to initiating early interventions is the absence of an academic curriculum designed to train general and child psychiatrists in the assessment of children who demonstrate very early symptoms of psychiatric disorders. Information and training in available evidence-based, early-intervention, manual-based protocols are not routinely provided to psychiatry residents.

There is a significant and growing public health awareness of the disparity between the recognition of the need for both prevention and early intervention services and the absence of a plan for developing an efficient and effective systematic delivery of such services. An immediate goal should be the development of new clinical initiatives that would begin to provide appropriate services to the highest-risk populations. The Affordable Care Act has a strong emphasis on prevention, and it may well be possible to increase reimbursement for psychiatric services, using mechanisms such as medical homes, collaborative care models, and braided funding.

Economic Issues. Ideally, psychiatric administrators should take responsibility for the management and funding of preventive interventions. However, to achieve this goal, new mechanisms to reimburse the provision of preventive services within the mental health field must be created. One issue will be a revision of diagnostic and billing codes to define specific evidence-based prevention services and to ensure payment for the delivery of these services. The standard medical reference for procedures that will be paid for by insurance companies is the Current Procedural Terminology (CPT) codes of the American Medical Association. CPT codes have been designed to define virtually all clinical procedures that have been accepted as legitimate services. Unfortunately, payment for many of the current CPT

codes that could be used to support services designed to prevent psychiatric illness have not been classified as traditional psychiatric interventions that are covered by medical insurance. Furthermore, even if insurance policies were to change and begin to make payments for preventive interventions that were authorized, many of the available CPT codes that are related to prevention-related services have been assigned very low relative value units (RVUs). When a CPT code is given a low RVU assignment, payment for the service will be quite low. Although a common compendium of CPT codes are established for all medical subspecialties, virtually no mental health specialists are able to use available codes for risk reduction and anticipatory guidance because of the difficulties that exist in receiving adequate payment. This needs to change and appropriate provision for payment needs to be made.

Another administrative problem is that payment for psychiatric services has traditionally required that a psychiatric diagnosis be identified for the patient prior to the clinician receiving payment. By definition, children who are receiving preventive services do not have a psychiatric diagnosis that would be included in the 10th revision of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) at the time of the provision of a preventive intervention. Instead, these children may have problems that could be classified using a "V code." Unfortunately, many insurance companies make the argument that the services provided to deal with problems that are categorized using V codes are not medically necessary. Therefore, these services are not eligible for medical insurance, and these payers deny payment to any psychiatric patients that have not been assigned a traditional Axis I diagnosis.

Training Issues. The next generation of psychiatric clinicians must be better trained in the science of the prevention of mental disorders. Much needs to be done to achieve this goal. For evidence-based prevention programs to be implemented on a national scale, it is absolutely necessary that clinicians be appropriately trained to provide these specified interventions. Supervised training experiences to provide trainees with appropriate skills will have to be incorporated into the traditional residency curriculum as one of the "core competencies" that must be developed during residency training. The ability to identify children at risk and deliver evidence-based preventive interventions must become an academic expectation of training. However, for this goal to be achieved, the importance of preventive interventions will need to become more explicitly stated in the residency training requirements that are available to provide guidance to the design of clinical training programs. Ultimately, board certification and recertification must require the demonstration of clinically relevant applications of preventive interventions as part of the core knowledge base that is examined.

Research Issues. Continuing a strong research program in prevention is essential. As noted, it provides one of the best ways of understanding the mechanisms by which either disorders are expressed or the balance of risk and protective factors is altered and disorders are prevented. Research investigators should be encouraged to assess multiple risk factors and multiple outcomes crossing the domains of mental health, substance abuse, physical health, and learning disorders. Priority should be given to sequential preventive interventions that can strategically target high-risk children prenatally, in infancy, in early childhood, in elementary school, during the teenage years, and finally in young adulthood. In addition, research efforts must consider and track high-risk youth and the different interventions they need at different stages (e.g., high-quality child care, daycare, and early education followed by access to high-quality schools). Finally, preventive intervention studies must aim to access the mechanisms by which such interventions work. Greater attention to genetics, epigenesis, and gene environment interactions is clearly warranted as this work goes forward.

The following case example illustrates how newer pharmacogenomic approaches offer great promise. George was 16 years old when he started experiencing depressive symptoms after his girlfriend decided to end their relationship. George had a family history of bipolar disorder and, by using pharmacogenomic testing, a clinician familiar with preventive approaches to mental illness was able to treat him safely.

When George presented to the clinician, he reported sleeping for more than 10 hours a night, rarely leaving his home, and having little interaction with his friends. His maternal uncle had struggled with bipolar disorder for many years. Although his parents had no history of mania or hypomania, both his mother and his father had been successfully treated with selective serotonin reuptake inhibitors for episodes of major depressive disorder.

A series of eight outpatient cognitive-behavior therapy sessions targeting his depressive symptoms proved to be minimally effective. Subsequently, George was placed on a dose of 20 mg of fluoxetine per day while being followed closely for the possible emergence of suicidal ideation. Within 2 weeks, he reported that he needed to sleep for only 1 or 2 hours each night and was experiencing marked surges of energy.

His psychiatrist became concerned that George was becoming hypomanic and was at risk for developing bipolar disorder. Prior to selecting another medication to manage these new symptoms, he ordered pharmacogenomic testing of a panel of genes with the goal of minimizing additional adverse reactions. George was determined to have a normal genotype of his cytochrome P450 2C19 gene, indicating that he could effectively metabolize 2C19 substrate medications. However, his cytochrome P450 2D6 and 2C9 genotypes documented that he was a poor metabolizer of both 2D6 and 2C9 substrate medications. In addition, his serotonin transporter genotype was associated with an increased risk for the onset of depression, and his GRIK3 genotype indicated that he was at an increased risk for the development of bipolar disorder.

George was started on a low dose of lithium, and the fluoxetine that he had been receiving was discontinued. After 2 months of outpatient treatment on lithium, he complained of having less energy and had become very discouraged about his ability to complete his planned applications to attend college. Given the results of his pharmacogenomic genotype testing, he was placed on a dose of 10 mg of escitalopram per day in addition to the lithium. After 4 weeks, he reported feeling much better and had not experienced a recurrence of hypomanic symptoms. He was subsequently able to complete his college applications without further difficulty.

George was maintained on lithium and escitalopram throughout his first year at college.

During the summer before the beginning of his second year away at school, both medications were discontinued and he remained euthymic.

Although fluoxetine is an effective and economical treatment for major depressive disorder in adolescents, fluoxetine is primarily a 2D6 substrate medication. Of note, the cytochrome P450 2C9 enzyme plays a secondary role in the metabolism of fluoxetine. Patients with both of these pathways compromised are at very high risk of developing side effects while taking standard doses of fluoxetine.

The clinical observation that manic activation may occur in depressed adolescents when treated with selective serotonin reuptake inhibitors has been frequently reported. However, clinicians have concluded that the benefits of treating many depressed adolescent patients with antidepressant medications have outweighed the risks of activating bipolar symptoms in the relatively few adolescent patients who develop manic symptoms when treated with these medications. Fortunately, it is now possible to identify genetic risk factors for activation and offer alternative medications. Managing the treatment of depression in a manner that minimizes activation can be conceptualized as an example of an indicated preventive intervention designed to prevent the onset of a first full episode of bipolar disorder. More generally, this case illustrates a way of thinking that is likely to be more useful in both treatment and prevention going forward.

Looking Forward to the Future. There is a debate in the prevention field regarding the cost–benefit ratio for universal interventions. Some universal prevention studies have been shown to provide the most benefits to the children who are at highest risk. In contrast, many children who are at minimal risk do not benefit at all. This pattern of response supports the view that early screening and interventions targeted to help populations with specific high risks, including early prodromal symptoms, could be considered for some disorders. Integrating the new, rapidly evolving personalized medicine approaches will become increasingly important in the future as well understanding how to effectively disseminate widely those prevention approaches that are strongly supported empirically.

Inasmuch as there are potent risk factors for multiple disorders (poverty or exposure to violence), more attention also needs to be devoted to addressing these public health drivers. The IOM recognized this in emphasizing the role of poverty: “One factor lurks in the background of every discussion of the risks for mental, emotional, and behavioral disorders, and antisocial behavior: poverty [. . .] Although not the focus of this report, there is evidence that changes in social policy that reduce exposure to these risks are at least as important for preventing mental, emotional, and behavioral disorders in young people as other preventive

interventions. We are persuaded that the future mental health of the nation depends crucially on how, collectively, the costly legacy of poverty is dealt with."

It is now possible to implement evidence-based selective and indicated preventive interventions to minimize risk factors associated with the onset of mental disorders. Broader implementation of these interventions for children and adolescents at risk for depression, conduct disorder, and psychosis will require continued political commitment. Given that these evidence-based interventions are predicted to have an ultimate cost savings effect on public health expenditures, there is real reason to be optimistic about their expansion and to recognize the need to ensure that such approaches are widely implemented.

ACKNOWLEDGMENT

This chapter is dedicated to the extraordinary work and life of Dr. David Mrazek who passed on May 6, 2013.

REFERENCES

Allen D, Coombes L, Foxcroft DR. Cultural accommodation of the Strengthening Families Programme 10–14: UK Phase I study. *Health Educ Res*. 2007;22(4):547–560.

Beardslee WR, Brent DA, Weersing VR, et al. Prevention of depression in at-risk adolescents: longer-term effects. *JAMA psychiatry*. 2013;70(11):1161–1170.

Beardslee WR, Solantaus TS, Morgan BS, Gladstone TR, Kowalenko NM. Preventive interventions for children of parents with depression: international perspectives. *Med J Aust*. 2013;199(3 Suppl):S23–25.

Beardslee WR, Wright EJ, Gladstone TR, Forbes P. Long-term effects from a randomized trial of two public health preventive interventions for parental depression. *J Fam Psychol*. 2007;21(4):703–713.

Botvin GJ, Griffin KW. Life skills training: preventing substance misuse by enhancing individual and social competence. *New Dir Youth Dev*. 2014;(141):57–65.

Champion KE, Newton NC, Barrett EL, Teesson M. A systematic review of school-based alcohol and other drug prevention programs facilitated by computers or the internet. *Drug Alcohol Rev*. 2013;32(2):115–123.

Clarke GN, Hornbrook M, Lynch F, et al. A randomized trial of a group cognitive intervention for preventing depression in adolescent offspring of depressed parents. *Arch Gen Psychiatry*. 2001;58(12):1127–1134.

Compas BE, Forehand R, Thigpen JC, et al. Family group cognitive–behavioral preventive intervention for families of depressed parents: 18- and 24-month outcomes. *J Consult Clin Psychol*. 2011;79(4):488–499.

Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. *Am J Psychiatry*. 2008;165(10):1272–1280.

D'Angelo EJ, Llerena-Quinn R, Shapiro R, et al. Adaptation of the preventive intervention

program for depression for use with predominantly low-income Latino families. *Fam Process*. 2009;48(2):269–291.

Garber J, Clarke GN, Weersing VR, et al. Prevention of depression in at-risk adolescents: a randomized controlled trial. *JAMA*. 2009;301(21):2215–2224.

Hagan MJ, Tein JY, Sandler IN, Wolchik SA, Ayers TS, Luecken LJ. Strengthening effective parenting practices over the long term: effects of a preventive intervention for parentally bereaved families. *J Clin Child Adolesc Psychol*. 2012;41(2):177–188.

Institute of Medicine (IOM). *Workshop 1: Individual Processes and Adolescent Risk Behavior. Institute of Medicine, Board of Children, Youth, and Families*. Washington, DC: 2008.

Institute of Medicine (IOM). *Harvesting the Scientific Investment in Prevention Science to Promote Children's Cognitive, Affective, and Behavioral Health: Workshop Summary*. Washington DC: The National Academies Press; 2014.

McFarlane WR, Levin B, Travis L, et al. Clinical and functional outcomes after 2 years in the early detection and intervention for the prevention of psychosis multisite effectiveness trial. *Schizophr Bull*. 2015;41(1):30–43.

McGlashan TH, Johannessen JO. Early detection and intervention with schizophrenia:rationale. *Schizophr Bull*. 1996;22(2):201–222.

McGorry PD, Cocks J, Power P, Burnett P, Harrigan S, Lambert T. Very low-dose risperidone in first-episode psychosis: a safe and effective way to initiate treatment. *Schizophr Res Treatment*. 2011;2011:631690.

Meaney MJ. Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annu Rev Neurosci*. 2001;24:1161–1192.

Merry S, McDowell H, Hetrick S, Bir J, Muller N. Psychological and/or educational interventions for the prevention of depression in children and adolescents. *Cochrane Database Syst Rev*. 2004;(1):CD003380.

*Mrazek PJ, Haggerty RJ, eds. *Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research*. Washington, DC: National Academies Press; 1994.

National Research Council, Institute of Medicine. The etiology of depression. In: *Depression in Parents, Parenting, and Children: Opportunities to Improve Identification, Treatment, and Prevention*. Washington, DC: National Academies Press; 2009.

O'Connell ME, Boat T, Warner KE, eds. *Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities*. Washington, DC: National Academies Press; 2009.

Poduska JM, Kellam SG, Wang W, Brown CH, Ialongo NS, Toyinbo P. Impact of the Good Behavior Game, a universal classroom-based behavior intervention, on young adult service use for problems with emotions, behavior, or drugs or alcohol. *Drug Alcohol Depend*. 2008;95 Suppl 1:S29–44.

Srihari VH, Tek C, Pollard J, et al. Reducing the duration of untreated psychosis and its impact in the US: the STEP-ED study. *BMC Psychiatry*. 2014;14:335.

Webster-Stratton C. Mothers' and fathers' perceptions of child deviance: roles of parent and child behaviors and parent adjustment. *J Consult Clin Psychol*. 1988;56(6):909–915.

Whitbeck L, Smith J. *Outcomes of a culturally-modified family program for Ojibway Indian families*. Washington, DC: Society for Prevention Research; 2001.

Wolchik SA, Sandler IN, Tein JY, et al. Fifteen-year follow-up of a randomized trial of a preventive intervention for divorced families: effects on mental health and substance use outcomes in young adulthood. *J Consult Clin Psychol*. 2013;81(4):660–673.

▲ 55.10 Child Mental Health Services Research

BONNIE T. ZIMA, M.D., AND REGINA BUSSING, M.D.

Child mental health services research integrates advances in child psychiatric assessment and treatment in examining unmet need for services, service use across care sectors (e.g., specialty mental health, primary care, schools, child welfare, juvenile justice, informal care), and the quality of care along its continuum (e.g., prevention, early detection, assessment, treatment, follow-up). It is a specialty area within the larger context of health services research in which the overarching goal is to improve the alignment between clinical need, delivery of recommended care, and health policies to sustain quality care.

Conceptually, it can be viewed as the second arm of translational research that translates scientific advances from "bedside to community." Westfall and colleagues expanded this model by inserting a third translational step, "practice-based research," which translates research findings to patients (e.g., treatment guideline development) and real world practice settings (e.g., dissemination and implementation research). This area of research is in contrast to the first translational arm which translates scientific advances in basic science to clinical research, shifting the generation of knowledge from "laboratory to bedside." To meet this challenge, health services research uses a compendium of quantitative and qualitative research methodologies that brings together expertise from multiple medical specialties and disciplines (i.e., social sciences, biostatistics, economics, health policy, administration), community partners, and consumers. For child mental health services research, the breadth of expertise also includes the perspectives of primary caregivers, teachers, social workers, case managers, probation officers, and child advocates.

Early stages of child mental health services research typically describe access to care and its predictors, extent of unmet need for services among children identified as clinically meriting treatment, and the delivery of care (e.g., service use, filled medication prescriptions) in usual care settings. Findings at this stage are typically a wake-up call alerting others of the potential magnitude of a problem in child mental health care, such as rising prescription rates of atypical antipsychotic medication among very young children. Some studies may also explore how mutable factors at several levels (i.e., system, community, family, child) independently predict who

receives care while adjusting for clinical need, to identify target areas for reducing disparities in access to or quality of care. Examples of questions driving such research are: Among children with a psychiatric disorder and impairment, who is more likely to access care? What are the potential mechanisms underlying disparities in child mental health care? How can technology (e.g., telepsychiatry) improve access and continuity of care?

With the development of treatment guidelines and quality measures, the scope of child mental health services research has expanded to describe the gap between recommended and usual care, and the predictors of greater adherence to recommended care processes. Improving the quality of child mental health care is of public health significance because child psychiatric disorders are prevalent, disabling, and costly. Use of treatments with established efficacy in community settings is typically low, and underuse of recommended treatment for mental health problems during childhood may negatively impact a child's developmental trajectory. Examples of questions common to quality of care research are: What proportion of children receives guideline-concordant care? What organizational, provider, family, or child factors are predictive of acceptable care? How should evidence-based practices (EBPs) be modified for use in community-based settings while preserving their effectiveness? Does provision of nationally recommended care processes relate to improved outcomes for children and their families? And if so, for whom and at what cost? How should care for child mental health problems be integrated into primary care settings in such a way that provision of care is seamless across care sectors and tailored to the changing needs of the child and family? What is the role of emerging technologies and how can they be applied to improve the quality of care? How does the extent of community partnership in developing new models of care improve engagement, uptake, and sustainability? Such information is necessary to inform the development and dissemination of quality improvement interventions delivered to children with mental health problems and their families.

As an introduction to child mental health services research, this section briefly describes the historical context and summarizes findings from select landmark studies and health policies. Access to care is described as a separate component of quality, following Avedis Donabedian's conceptualization of quality that argues that accessibility is a separate yet influential component of quality. This is followed by an overview of a well-established model of quality of care, examples of quality of care research and national quality measures. Because the study of the quality of care for children with psychiatric disorders and their families spans the care continuum, the reader is encouraged to read cross-referenced sections and chapters to stimulate further inquiry about how to close the gap between recommended and actual care.

HISTORICAL CONTEXT

As outlined in Table 55.10–1 advances in clinical, epidemiology and health services research as well as health policy initiatives shape the historical context from which child mental health services research has emerged. The first half of the twentieth century was characterized by increasing federal initiatives that raised public awareness of the need to improve the quality of life for children, such as a child's right to preventive health care, safe working conditions, and a free public education, particularly for those identified as seriously emotionally disturbed or handicapped. After the mid-century mark, child psychiatry was identified as a relatively new medical subspecialty, with its first national professional organization established in 1953 and board certification in 1959. During the next three decades, landmark epidemiological studies provided some of the first prevalence estimates of child mental health problems in community-based populations and their correlates. These studies were significant because the findings raised awareness of the substantial proportion of children who met criteria for a psychiatric disorder and the need to improve early detection and access to child mental health services.

During the 1980s the pediatric community identified psychosocial conditions, such as poverty, family dysfunction, and child exposure to traumatic events, as part of the new morbidity of child health that included learning and behavioral disorders. The conceptualization of child health was further broadened the following decade to include attention-deficit/hyperactivity disorder (ADHD) within the definition of chronic conditions among children with special health care needs. Within this same time period, a philosophy for a system of care for children was operationalized. The two core values were that care should be child centered and community based. These values were later echoed a decade later in the guiding principles underlying the pediatric medical home. Remarkably, these principles share similar dimensions of "quality" as identified in the "articles of faith" underlying good medical care by R. I. Lee and L. W. Jones more than 50 years earlier. Nevertheless, sustained articulation of these values and principles is important because it provides a relatively consistent framework from which programs serving children with mental health problems can plan and implement new models of care for children with mental health problems. Table 55.10–2 summarizes the guiding principles of these two conceptual models.

During the 1990s, federal initiatives increasingly focused on improving child mental health care and promoting basic science, clinical, and health services research for childhood psychiatric disorders. In preparation for a national epidemiological study to estimate the prevalence of childhood mental disorders and service use, the Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) Study was conducted. The MECA Study was a multisite, community-based study that further refined

standardized assessment methods of psychiatric diagnosis, impairment, mental health service use, and risk factors and also provided prevalence estimates of childhood psychiatric disorders and mental health service use among four diverse geographic areas. Despite these advances and others, to date, only one study, the National Health and Nutrition Examination Study, used standardized diagnostic interviews among a nationally representative sample of children ages 8 to 15 years to estimate the prevalence of ADHD among US children.

Table 55.10–1.
Brief Historical Context from Which Child Mental Health Services Research Is Developing

	Event(s)	Significance
1909	White House Conference on Dependent Children	Raised public awareness of poor living conditions and high cost of homes for emotionally disturbed children. National recommendations for improved child programs stimulated the development of child welfare and foster care system.
1912	Federal Children's Bureau created	Establishment of the oldest federal agency for children. Located in the U.S. Department of Health and Human Services' Administration for Children and Families, the agency is responsible for assisting states in the delivery of child welfare services; providing grants to states, Native American tribes, and communities to operate a range of child welfare services (e.g., child protective services, family preservation and support, foster care, adoption, and independent living); and making major investments in staff training, technology, and innovative programs for children.
1919	The Child Welfare League of America chartered	Establishment of the oldest and largest national nonprofit organization dedicated to developing and promoting policies and programs to protect US children and strengthen US families.
1920	The Dawson Report published in England	Considered to be one of the founding documents in the history of epidemiology by introducing the concept of primary care. Lord Dawson believed health care in a given district should be based on "primary health centres," which he termed "institutions equipped for services of curative and preventive medicine to be conducted by the general practitioners of that district." Difficult cases requiring special treatment would be referred to "secondary health centers." In addition, "secondary services" would be equipped for the care of infectious diseases and chronic ailments such as epilepsy or mental illness. "Teaching hospitals," with medical schools, would treat the most difficult illnesses and provide physician training.
1922	Pediatric Section of AMA approves the Sheppard-Towner Act while the AMA House of Delegates condemns it.	The Sheppard-Towner Act supported federal funding to states for health services for mothers and children. The AMA opposed it, calling it "an imported socialist scheme." Following this disagreement, AMA leadership promptly ruled that no section could independently adopt a resolution. This ruling was described as, "sterilizing the Pediatric Section [that] made the

		American Academy of Pediatrics inevitable."
1930	White House Conference on Child Health and Protection	Children's Charter established, outlining a set of 19 principles in the areas of education, child labor, vocational training, recreation, family welfare, health, growth and development, disabled children, and related areas, including the Bill of Rights for the Handicapped Child.
1930	American Academy of Pediatrics established	On June 23–24, 1930 at Detroit's Harper Hospital, 35 pediatricians drew up the constitution and by-laws for AAP. The purposes of the AAP were: 1. To create reciprocal and friendly relations with all professional and lay organizations that are interested in the health and protection of children. 2. To foster and encourage pediatric investigation, both clinically and in the laboratory, by individuals and groups.
1933	Roger I. Lee, Lewis Webster Jones, and Barbara Jones, *The Fundamentals of Good Medical Care,* a publication of the Committee on the Costs of Medical Care	Defined quality of care as follows: "Good medical care is the kind of medicine practiced and taught by the recognized leaders of the medical profession at a given time or period of social, cultural, and professional development in a community or population group." Identified eight articles of faith underlining the concept of good care: 1. Good medical care is limited to the practice of rational medicine based on the medical sciences. 2. Good medical care emphasizes prevention. 3. Good medical care requires intelligent cooperation between the lay public and the practitioners of scientific medicine. 4. Good medical care treats the individual as a whole. 5. Good medical care maintains a close and continuing personal relationship between physician and patient. 6. Good medical care is coordinated with social welfare work. 7. Good medical care coordinates all types of medical services. 8. Good medical care implies the application of all the necessary services of modern scientific medicine to the needs of all the people.
1935	L. Kanner, *Child Psychiatry*	First English-language textbook on child psychiatry.
1938	Fair Labor Standards Act signed into law by President Franklin Delano Roosevelt	Prohibited interstate shipment of goods made in firms that employ children younger than age 16 years or children younger than age 18 years in hazardous occupations.
1950	Mid-Century Conference on Children and Youth held	Prepared the Pledge to Children, which summarizes commitments by adults to children to promote education, strengthen family life, safeguard civil equality, improve workplace conditions, enhance democratic citizenship and institutions, and advocate for the development of the child's individual talents and potential.
1953	American Academy of Child and Adolescent Psychiatry established	The leading national professional medical association, currently composed of more than 6,500 child and adolescent psychiatrists and other interested physicians dedicated to treating and improving the quality of life for children, adolescents, and families affected by these disorders.
1958	Community-based	First wide-scale epidemiological study of problem behaviors

	epidemiological survey of mental health problems among those living in a northeastern US community	among a US community-based sample. Using systematic assessment methods for problem behaviors, investigators identified high rates of fear and worries among children aged 6–12 years in a general population sample, suggesting that symptoms are in part developmental. Established the feasibility of child self-report of emotional and behavioral symptoms. Used multiple-informant data and reported discrepancies in mother and child reports for some symptoms.
1959	Board certification in child and adolescent psychiatry offered for the first time by the American Board of Psychiatry and Neurology	National application of standards for clinical knowledge and skills in child and adolescent psychiatry.
1961	Kerr White and colleagues, "The Ecology of Medical Care," published in *New England Journal of Medicine*	Seminal article that promoted the development of health services research. The authors "observed that traditional indexes of public health—including mortality and morbidity—do not adequately describe the actions of individual patients and doctors." They argued that doctors should combine research of pathological processes with exploration of the use of medical care resources, their relative availability in different contexts, and the patient's decision-making processes. In their view, "the patient may be a more relevant primary unit of observation than the disease, the visit, or the admission." Findings from this study examining the relationship between US and UK populations and the experience of illness and different levels of medical treatment suggested that health care delivery and the training of physicians did not bear any logical relationship to the actual experience of illness in a given population. These data suggested that greater attention should be devoted to primary, continuing medical care as opposed to episodes of hospitalization or consultation of specialists. Furthermore, medical education and research should take greater account of the environment of health care and consider more systematically how health care resources should be deployed.
1964–1975	Isle of Wight studies	Extensive longitudinal cohort study over 5 years of intellectual and educational retardation, psychiatric disorder, and physical handicap of the entire population (estimated *N* = 3,500) of children aged 9–11 years. Screening for neuroepileptic disorders among entire population of 5- to 15-year-olds. Survey of 10-year-old children to directly compare to children of the same age range living in an inner borough of London. Additional studies examining the mental health profile of young children age 5–7 years. Innovative research designs that included the following: 1. Multiple-informant data (child, parent, teacher, school records, psychiatric records, and pediatric records).

		2. Two-stage screening strategy. 3. Comprehensive battery of assessments. 4. Identification of psychiatric diagnoses using social impairment as a criterion for severity and symptom pattern as a criterion of type. 5. Longitudinal design to study age of onset, identification of risk factors, and description of course of illness over time. 6. Causal hypotheses of risk factors (e.g., organic brain dysfunction, family discord, reading difficulties). 7. Standardized neurodevelopmental assessments used to relate neurological findings to psychiatric disorders. 8. Examination of differences in participants and nonparticipants.
1966	Avedis Donabedian, "Evaluating the Quality of Medical Care," published in *Milbank Memorial Quarterly*	Proposed conceptual model for the evaluation of quality of care that remains widely used in health services research. Identified three main domains for examining quality: (1) structure, (2) process (technical, interpersonal), and (3) outcomes.
1968–1982	RAND Health Insurance Experiment	The most important health insurance study ever conducted. Addressed two key questions in health care financing: (1) How much more medical care will people use if it is provided free of charge and (2) what are the consequences for their health? Using a large-scale experimental design, different types of health insurance plans were offered to people in five sites across the country. The study examined the relationship between level of insurance coverage and service use rates and health outcomes. As the rate of cost sharing for health care rose, the consumption of health care fell. Of note, cost sharing had particularly adverse effects on the health of vulnerable groups, such as low-income children.
1969	Joint Commission on Mental Health for Children	Raised public awareness of high unmet need for child mental health services and overuse of mental health care delivered in restrictive treatment settings. Recommended establishment of a child advocacy system to coordinate federal, state, and local actions, development of community services focused on prevention and remediation, and delivery of child services in settings resembling normal living conditions. Identified need for research examining the effectiveness of child mental health services.
1974	Large epidemiological study of 2,034 youth age 6–18 years in New York	Included a sample of 1,000 youth in the welfare sector and 1,034 randomly selected youth. Prevalence estimates of psychiatric disorders were higher among children in the welfare sector than in the general population (21.0% vs. 17.5%). Documented high unmet need for services, such that <50% of children with serious impairment were referred for care.
1975	Education for All Handicapped Children Act passed (Public Law 94–142)	Federal law whose purpose is to ensure that all children with disabilities have access to a free appropriate public education that emphasizes special education and related services designed to meet their unique needs; to ensure that the rights of children with disabilities and their parents or guardians are protected; to assist states and localities to provide for the education of all

		children with disabilities; and to assess and ensure the effectiveness of efforts to educate children with disabilities.
1977	President's Commission on Mental Illness and Health	Identified need for studies of the incidence and prevalence of mental disorders in the United States. Recommended an increase in the provision of child mental health services and the capacity to provide a continuum of care to children.
1979	National Institute of Mental Health (NIMH) contracts to develop standardized assessment methods of children and adolescents in general population samples	Stimulated a wave of epidemiological studies examining the reliability and validity of checklists, rating scales, and diagnostic interviews in geographically diverse populations in the 1980–1990s. Promoted the development of assessment tools for use in community-based populations, such as the Child Behavior Checklist, Clinical Global Assessment Scale, Child and Adolescent Psychiatric Assessment, Diagnostic Interview for Children, and Diagnostic Interview Schedule for Children.
1980	Cross-national replication of epidemiological study of 100 3-year-old children	Raised awareness of early onset of mental health problems among preschool-age children in the United States. An estimated 11% of children age 3 years screened positive for a behavior problem.
1980	NIMH Epidemiologic Catchment Area Study	First large-scale, multisite epidemiological study of psychiatric disorders among 20,291 adults in US households and institutions. Stimulated further development of the Diagnostic Interview Schedule and provided prevalence estimates of treated and untreated psychiatric disorders in the community.
1981	Select Panel for the Promotion of Child Health	Acknowledges significant improvement in services to children with disabilities (including serious emotional disturbances) but identifies wide variation in the availability of child-related services. Developed national recommendations to improve the delivery of services to children, including greater access to care, expansion of services to include prevention and early detection and treatment of developmental problems, improved quality of care in residential treatment settings, enhanced community support for child programs, and greater involvement of families in the delivery of mental health services.
1982	Children's Defense Fund report, *Unclaimed Children*	Raised public awareness of high unmet need for child mental health services in the United States. Identified children with severe emotional and behavioral problems as "unclaimed" by the public agencies responsible for serving them.
1984	Child and Adolescent Service System Program (CASSP) established	NIMH-funded program to support states in the development of interagency efforts to improve the system of care for children with mental health problems. Provided technical assistance to community-based programs in the adaptation of systems-of-care values and principles to the delivery of services in children with severe emotional and behavioral problems.
1985	Institute of Medicine	Documented the high societal burden of mental disorders and

	report, *Research on Mental Illness and Addictive Disorders: Progress and Prospects*	substance abuse, progress in the assessment and treatment of mental illness, and need for additional federal funding for research.
1986	Beth A. Stroul and Robert M. Friedman, *A System of Care for Severely Emotionally Disturbed Children and Youth*	Pivotal document from the CASSP Technical Assistance Center, Georgetown University Child Development Center, in Washington, DC, that articulated the core values and guiding principles for the system-of-care philosophy for children.
1986–1992	Medical Outcomes Study	A large, 2-year longitudinal study of variations in practice style and outcomes of care for patients with specific "tracer conditions," including adult major depression, in three US metropolitan areas. Poor functioning among patients with depressive disorder or depressive symptoms was comparable to or worse than that associated with major medical conditions (hypertension, diabetes, advanced coronary heart disease, angina, arthritis, back problems, gastrointestinal disorders). Detection and treatment of depression were significantly lower among depressed patients who visited a medical clinician (45.9% and 51.2%, respectively) compared with those who visited a mental health specialist (78.2% and 86.9%, respectively).
1989	Effectiveness of "typical" psychotherapy in community-based clinics examined	Compared 6-month and 1-year symptom levels among a nine-clinic sample of children completing a course of therapy (*N* = 93) and those who dropped out after intake (*N* = 60). No significant main effects of therapy were found. Raised awareness of the need to reconsider the generalizability of outcomes studies of controlled laboratory interventions when assessing the effectiveness of therapy under usual care conditions.
1989	Agency for Health Care Policy and Research established (later renamed Agency for Healthcare Research and Quality)	The health services research arm of the US Department of Health and Human Services (HHS). Supports research in the following areas: 1. Quality improvement and patient safety. 2. Outcomes and effectiveness of care. 3. Clinical practice and technology assessment. 4. Health care organization and delivery systems. 5. Primary care. 6. Health care costs and sources of payment. Major source of funding and technical assistance for health services research and research training at leading US institutions. A science partner, working with the public and private sectors to build the knowledge base for what works—and does not work—in health and health care and to translate this knowledge into everyday practice and policy making.
1989	System of Care for Severely Emotionally Disturbed Children and Youth	Introduced a philosophy underlying a system of care for children with severe emotional and behavioral problems. Identified core values for system of care for children: Care should be child centered and community based.

		Outlined 10 guiding principles for system of care.
1989–1995	Fort Bragg Demonstration Project	Large-scale effectiveness study of providing a continuum of services to children and their families with four components: 1. Implementation study 2. Quality study 3. Mental health outcome study 4. Cost/use study A longitudinal cohort study comparing outcomes and costs of providing a model continuum of care (Rumbaugh Clinic in Fort Bragg catchment area) vs. traditional mental health services (Fort Campbell and Fort Stewart catchment areas). Compared with traditional mental health services, provision of a continuum of care was related to the following: 1. Improved access to care. 2. Higher adolescent and parent satisfaction with services. 3. Greater costs. 4. Reduced rates of hospitalization or residential treatment. Continuum of care did *not* produce better clinical outcomes than traditional mental health services.
1989	Institute of Medicine report on Research on Children and Adolescents with Mental, Behavioral, and Developmental Disorders	Comprehensive review by national experts of the dimensions, causes and determinants, and interventions for childhood psychiatric disorders, as well as child mental health research infrastructure. Raised public awareness of an estimated 12% of US children having a diagnosable mental illness, advances in the assessment and treatment of childhood disorders, promising research directions, and resource needs for research.
1989	Agency for Health Care Policy and Research created.	Supported by the Omnibus Budget Reconciliation Act of 1989. Mission included "to support research designed to improve the quality, safety, efficiency and effectiveness of health care for all Americans." Activities included large multidisciplinary, multiinstitutional projects that focused on patient results in certain medical conditions, development of practice guidelines, and participating in the development of a health care reform plan.
1990	National Plan for Research on Child and Adolescent Mental Disorders	Comprehensive 5-year plan of how NIMH intended to stimulate new basic and clinical research, promote research training, and provide leadership and coordination to stimulate advocacy and research dissemination in child mental health.
1990	The 1990s declared by the US Congress the "Decade of the Brain"	Provided national public endorsement for federal, state, and community initiatives to support a wide range of mental health research, as well as initiatives to improve the quality of care for the mentally ill.
1990	The National Committee for Quality Assurance is created.	A large, nonprofit organization dedicated to improving health care quality through measurement, transparency and accountability. Products include the Healthcare Effectiveness Data and Information Set (HEDIS), a set of quality measures used to report on quality of care at the health plan level.
1991–1996	Methods for the Epidemiology of Child and Adolescent Mental Disorders	Pilot studies followed by multisite study of 1,285 child–caregiver pairs among four geographic areas using household interviews (Hamden, East Haven, West Haven, Connecticut; DeKalb, Rockdale, Henri Counties, Georgia; Westchester County, New

	(MECA) Study	York; San Juan, Puerto Rico). Stimulated the development of standardized assessment methods of psychiatric diagnosis, impairment, mental health service use, risk factors, and administrative infrastructure for large-scale epidemiological study. Produced prevalence estimates of psychiatric disorders with and without impairment criteria and mental health service use.
1992	Passage of the Comprehensive Community Mental Health Services Program for Children and Adolescents with Severe Emotional Disturbances	Authorized federal funding of more than $100 million/year for implementation of systems of care in 38 different communities. Supported the development of community-based continuum-of-care programs for children, including assessment, case management, outpatient treatment, day treatment, in-home services, respite care, therapeutic foster care, and group homes for children with severe emotional disturbances.
1992	American Academy of Pediatrics, Ad Hoc Task Force on Definition of the Medical Home	The guiding principles for the medical care of infants, children and adolescents should be accessible, continuous, comprehensive, family centered, coordinated, and compassionate. Comprehensive care encompass the following services: 1. Provision of preventive care. 2. Ambulatory and inpatient care for acute illness 24 hours a day, 7 days a week. 3. Provision of care over an extended period of time to enhance continuity. 4. Identification of the need for subspecialty consultation and referral and knowing from whom and where these can be obtained. 5. Interaction with school and community agencies to be certain that special health needs of the individual child are addressed. 6. Maintenance of a central record and data base containing all pertinent medical information about hospitalizations. This record should be accessible, but confidentiality must be assured.
1997–1999	Development of practice parameters by the American Academy of Child and Adolescent Psychiatry	Twenty practice parameters were developed for the assessment and treatment of specific psychiatric disorders, as well as child mental health problems (e.g., suicide, sexual abuse).
1998	National Advisory Mental Health Council's Clinical Treatment and Services Research Workgroup, *Bridging Science and Service*	Work group of researchers, policy makers, and mental health advocates charged by NIMH to advise the council on strategies for increasing the relevance, speeding the development, and facilitating the use of research-based treatment and service interventions for mental illnesses into both routine clinical practice and policies guiding local and national mental health service systems. Developed an action plan to meet the following goals: 1. Increase the usefulness of NIMH research for individuals with mental illnesses, clinicians, purchasers, and policy makers through informed priority setting. 2. Selectively expand the NIMH portfolio in the domains of

		efficacy, effectiveness, practice, and service systems research to foster integration across these fields and to expedite the implementation of research-based practices and policies. 3. Identify research innovations in design, methods, and measurement to facilitate the translation of new information from bench to trial to practice. 4. Strengthen NIMH's leadership and administrative activities to provide the infrastructure to achieve the aforementioned goals stated in a timely manner.
1998	National Institutes of Health Consensus Development Conference on Diagnosis and Treatment of Attention Deficit Hyperactivity Disorder	Review the scientific evidence, public open discussion, and closed-panel deliberations by a 13-member multidisciplinary national expert panel. Findings from this conference support the following: 1. The validity of the diagnosis of attention-deficit/hyperactivity disorder (ADHD). 2. The substantial effect of ADHD on children, their families, and society. 3. Effectiveness of psychosocial and medication treatments. 4. Need for additional research on the long-term effects of stimulant medication treatment. 5. Need to improve early detection, access to care, and quality of evaluation and treatment. 6. Need for further basic science and clinical and intervention research.
1999	White House Conference on Mental Health	First major public introduction to mental illness in the United States and scientific advances in the assessment and treatment of psychiatric disorders. Called for a national anti-stigma campaign for persons with mental illness.
1999	Surgeon General's Call to Action to Prevent Suicide	Developed a national plan to increase awareness and prevent suicide in the United States.
1999	Surgeon General's Report on Mental Health	First Surgeon General's report dedicated to mental health and mental illness. Communicated three main messages: (1) Mental health is fundamental to health; (2) mental disorders are real health conditions; and (3) the efficacy of mental health treatments is well documented, and a range of treatments exists for most mental disorders. Included a chapter on children and mental health.
1999	Multimodal Treatment Study of Children with ADHD (MTA study)	First large-scale, multisite, randomized clinical trial of treatment of a childhood mental disorder. Compared the effectiveness of intensive behavioral treatment, tightly managed medication treatment, combined behavioral and medication treatment, and standard community care for ADHD. A randomly selected sample of 579 children age 7.0–9.9 years recruited from six university-based clinic referral sources, including mental health settings, pediatricians, advertisements, and school notices. Oversampled children from families that are more likely to fully participate in assessment and treatments and less likely to require additional treatments that were incompatible with standardized ADHD treatment protocols.

		Children were randomly assigned to one of four treatment arms. The following were found when using core ADHD symptom reduction as the outcome: 1. Medication management was superior to behavioral treatment and routine community care. 2. Combined treatment did not yield significantly more benefit than medication management alone.
1999	The Agency for Health Care Policy and Research is renamed the Agency for Healthcare Research and Quality	The agency was reauthorized under the Healthcare Research and Quality Act of 1999. Its mission is to produce evidence to make health care safer, higher quality, more accessible, equitable, and affordable. Development of practice guidelines was removed from the new mission and it was made explicit that the agency is not a regulatory body.
1999	National Quality Forum created.	This nonprofit, membership-based organization was created in response to recommendations from the President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry. The main activities are: (1) convene working groups to foster quality improvement in both public and private sectors; (2) endorse consensus standards for performance measurement; (3) ensure that consistent, high-quality performance information is publicly available; and (4) seek real time feedback to ensure measures are meaningful and accurate.
2000	White House Meeting on Child Mental Health	Major public meeting specifically addressing the need to improve the diagnosis and treatment of children with emotional and behavioral disorders.
2000	US Surgeon General's Conference on Children's Mental Health, Developing a National Action Agenda	Developed recommendations for a National Action Agenda on Children's Mental Health based on feedback from a broad cross section of more than 300 child mental health stakeholders
2000	National Institute of Mental Health and the Food and Drug Administration (FDA) Workshop on Psychopharmacology for Young Children, Clinical Needs and Research Opportunities	A nationally sponsored meeting of researchers, practitioners, ethicists, industry staff, and family and patient representatives, whose mission was to examine the current state of knowledge regarding psychopharmacology for young children and discuss a variety of topics relevant to research in this age group, including safety, efficacy, investigational methods, and ethical aspects. The meeting was in response to the increasing use of psychotropic medication among preschool-age children, for which the efficacy and safety of psychotropic medication have not been systematically evaluated. Recommendations included additional research in developmental psychopathology to clarify diagnosis and provide sensitive and specific methods for clinical trials.
2000	National Research Council and Institute of Medicine, *From Neurons to Neighborhoods: The Science of Early Childhood Development*	A 17-member committee of national experts convened to update scientific knowledge about the nature of early child development; disentangle knowledge from erroneous popular beliefs; and discuss the implications of this knowledge base for early childhood policy, practice, professional development, and research. With 11 national recommendations, this monograph advocated for improving the national capacity to meet high unmet need for mental health services among preschool

		children and conducting research examining the effectiveness of early intervention programs for children.
2000	Clinical practice guidelines for ADHD developed by the American Academy of Pediatrics	National recommendations for the evaluation and diagnosis of ADHD, as well as treatment of the school-age child with ADHD in the primary care sector.
2000	Agency for Health Care Policy and Research evidence report, *Treatment of Attention-Deficit/Hyperactivity Disorder*	As part of the Evidence-Based Practice Program, a multidisciplinary research team conducted a systematic review of the scientific evidence for the treatment of ADHD to develop a base from which clinical guidelines, performance measures, and other quality improvement tools could be developed.
2000–2002	Development of quality indicators for childhood psychiatric disorders by American Psychiatric Association Task Force on Quality Indicators for Children	Conceptualized the quality of care for children in five main domains: Prevention, access, quality/process/appropriateness, satisfaction/perceptions of care, and outcomes/effectiveness. Developed multiple recommendations for quality care and corresponding sample indicators.
2001	Surgeon General's Conference on Children's Mental Health, *A National Action Agenda*	Made the following recommendations: 1. Promote public awareness of children's mental health issues and reduce the stigma associated with mental illness. 2. Continue to develop, disseminate, and implement scientifically proven prevention and treatment services in the field of children's mental health. 3. Improve the assessment and recognition of mental health needs in children. 4. Eliminate racial/ethnic and socioeconomic disparities in access to mental health care services. 5. Improve the infrastructure for children's mental health services, including support for scientifically proven interventions across professions. 6. Increase access to and coordination of quality mental health care services. 7. Train front-line providers to recognize and manage mental health issues and educate mental health care providers about scientifically proven prevention and treatment services. 8. Monitor the access to and coordination of quality mental health care services.
2001	National Advisory Mental Health Council's Workgroup on Child and Adolescent Mental Health Intervention Development and Deployment, *Blueprint for Change: Research on Child*	A 21-member work group of national experts that reviewed major research findings on child and adolescent psychiatric disorders over the prior decade and identified three major research priority areas: 1. Basic science and the development of new interventions. 2. Intervention development, moving from efficacy to effectiveness. 3. Intervention deployment, moving from effectiveness to dissemination.

	and Adolescent Mental Health	
2001	Institute of Medicine's Committee on Quality of Health Care in America, *Crossing the Quality Chasm*	Proposed six aims for improving the quality of care for the 21st century: Health care should be the following: 1. Safe: Avoiding injuries to patients from the care that is intended to help them. 2. Effective: Providing services based on scientific knowledge to all who could benefit and refraining from providing service to those not likely to benefit (avoiding underuse and overuse, respectively). 3. Patient-centered: Providing care that is respectful of and responsive to individual patient preferences, needs, and values and ensuring that patient values guide all clinical decisions. 4. Timely: Reducing waits and sometimes harmful delays for both those who receive and those who give care. 5. Efficient: Avoiding waste, including waste of equipment, supplies, ideas, and energy. 6. Equitable: Providing care that does not vary in quality because of personal characteristics such as gender, ethnicity, geographic location, and socioeconomic status.
2002	President's New Freedom Commission on Mental Health	National commission whose mission was to conduct a comprehensive study of the US mental health service delivery system, including both the private and public sector providers and advise the president on methods of improving the system. The main goals were the following: 1. Review the current quality and effectiveness of public and private providers and federal, state, and local government involvement in the delivery of services to individuals with serious mental illnesses and children with serious emotional disturbances and identify unmet needs and barriers to services. 2. Identify innovative mental health treatments, services, and technologies that are demonstratively effective and can be widely replicated in different settings. 3. Formulate policy options that could be implemented by public and private providers and federal, state, and local governments to integrate the use of effective treatments and services, improve coordination among service providers, and improve community integration for adults with serious mental illnesses and children with serious emotional disturbances.
2002	Agency for Healthcare Research and Quality, *Systems to Rate the Strength of Scientific Evidence*	Identified 19 generic systems that fully address key quality domains for a particular type of study and 7 systems that address all three quality domains grading the strength of a body of evidence. These systems may be useful to the development of clinical practice guidelines or other health-related policy advice.
2002	International Consensus Statement on ADHD	A consensus statement developed by a consortium of international scientists that supports the validity of ADHD and its adverse effect on the lives of children and their families using scientific evidence. The statement was developed as a reference on the status of the scientific findings about ADHD to dispel periodic public perception that ADHD is not a real medical condition.
2003	President's New	Recommendations for Transformed Mental Health System:

	Freedom Commission on Mental Health: Transforming Mental Health Care in America	1. Americans understand that mental health is essential to overall health. 2. Mental health care is consumer and family driven. 3. Disparities in mental health services are eliminated. 4. Early mental health screening, assessment, and referral to services are common practice. 5. Excellent mental health care is delivered and research is accelerated. 6. Technology is used to access mental health care and information.
2004	New York/Montreal 2-year study of ADHD treatment with long-term methylphenidate and multimodal psychosocial treatment	Among two clinically based samples of stimulant-responsive children ages 7–10 years who meet diagnostic and impairment criteria for ADHD only (i.e., no comorbidity), combination treatment was not found to be superior over medication treatment alone. Outcomes were symptom reduction, impairment, classroom behavior, academic achievement, and social functioning.
2004	U.S. Food and Drug Administration (FDA) black box warning on pediatric antidepressants	The FDA-directed manufacturers of all antidepressant medication to add a boxed warning that reported (1) antidepressants increase the risk of suicidal thinking and behavior in children and adolescents with major depression and other psychiatric disorders, (2) anyone considering the use of an antidepressant in a child or adolescent for any clinical use must balance the risk of increased suicidality with the clinical need, (3) patients who are started on therapy should be observed closely for clinical worsening, suicidality, or unusual changes in behavior, and (4) disclose whether the particular drug is approved for a pediatric indication(s) and, if so, which one(s).
2005	Caring for California Initiative	First statewide study of the quality of care for children with ADHD, major depression, and conduct disorder in publicly funded outpatient specialty mental health programs.
2006	Institute of Medicine's (IOM) Committee on Quality of Health Care in America, *Improving the Quality of Health Care for Mental and Substance Use Conditions*	Recommended that the prior IOM health care system redesign strategy was applicable to mental and substance use health care but required some adaptation to address greater obstacles to patient-centered care, a weaker quality measurement and improvement infrastructure, more separated care-delivery arrangements, less use of information technology, a more educationally diverse workforce with a greater use of solo practice, and a differently structured marketplace.
2006	National Advisory Mental Health Council's Workgroup on Services and Clinical Epidemiology Research, *The Road Ahead: Research Partnerships to Transform Services*	Six issues driving recommendations shaping the future research agenda to transform mental health services are (1) partnerships in research, (2) quality of care, (3) fairness, (4) recovery, (5) communication, and (6) ongoing evaluation.
2007	Treatment for Adolescents with	Among 327 patients from 13 academic and community sites randomized to combination therapy, fluoxetine therapy, and

	Depression Study (TADS) 36-wk outcomes	cognitive behavior, combined treatment appears to be superior to monotherapy using symptom reduction and overall clinical improvement rating as outcomes.
2007	NIMH MTA study, 3-year follow-up	At 36 months—2 years after treatment protocols (medication only, combined, behavior only, usual community care)—the advantage of medication only or combined treatment was not sustained. Treatment groups did not vary in parent- and teacher-rated ADHD and oppositional defiant disorder symptoms, reading achievement scores, social skills, and functional impairment.
2007	National Stigma Study-Children	First large-scale, nationally representative survey of public knowledge, attitudes, and beliefs on children's mental health reported substantial rates of stigma (35–57%) and negative perceptions of medication treatment (52–85%).
2009	The Children's Health Insurance Program Reauthorization Act	Mandated the identification, refinement and development of child health care measures for voluntary use in Medicaid and Children's Health Insurance Programs. Called for submission of initial core set of 24 quality measures to Department of Health and Human Services by 2010. Funded the 5-year Pediatric Quality Measures Program.
2010	Patient Protection and Affordable Care Act	Landmark legislation for health care reform. Mandated development of a National Quality Strategy and earmarked support for the Patient-Centered Outcomes Research Institute.

In response to increasing adaptation of system of care principles to the delivery of community-based services for children and their families, a large-scale effectiveness study was conducted, namely, the Fort Bragg Demonstration Project. Provision of a continuum of care was found to provide more services and be more costly but did not produce better clinical outcomes. These thought-provoking findings encouraged the field to conceptually expand outcomes for children beyond clinical status (e.g., to system-, community-, family-, and parent-level outcomes) and to integrate the influence of varying service use (e.g., intensity, persistence, type) in the determination of effectiveness.

Table 55.10–2.
Core Values of Children's System of Care and Pediatric Medical Home

Children's System of Care	Pediatric Medical Home
Mental Health Care	Medical Care
Access to a comprehensive array of services	Accessible
Smooth transitions to the adult service system	Continuous
Early identification and intervention	Comprehensive
Full participation of families and surrogate families	Family-centered
Integrated and coordinated care across child-serving agencies	Coordinated
Inclusion of case management	
Individualized services	Compassionate
Least restrictive, most normative environment	
Protection of child rights and effective advocacy efforts	
Equitable and culturally sensitive service delivery	

Scientific advances in the assessment and treatment of ADHD also laid the foundation for conducting the first large-scale, randomized, clinical trial of two treatment modalities. The Multimodal Treatment Study of Children with ADHD, or MTA study, examined the effectiveness of highly standardized psychosocial, medication, and combined treatment protocols for ADHD. Based on the use of reduction of ADHD symptoms as the main outcome at 14 months, medication management was found to be superior to intensive psychosocial treatment and usual community care, and combined psychosocial treatment was found to be of no significantly greater benefit than medication management alone. Findings from the MTA study, however, should not be interpreted as conclusive that psychosocial interventions are inferior to medication treatment. Instead, alternative analyses from this rich data set have illuminated key methodological issues, such as selection of outcome measures, use of composite versus individual outcomes as indicators of effectiveness, and the effect of moderators (e.g., sociodemographic factors, clinical severity, parental education, and income) and mediators (e.g., treatment adherence) in the assessment of an intervention's effectiveness. Nevertheless, treatment effects were found not to be sustained. At 36 months, the treatment groups did not significantly differ on any outcome measure. In addition, based on saliva assays for methylphenidate as an objective indicator of medication adherence, only 53.5 percent of children were consistently adherent to stimulant medication treatment during initial 14-month treatment period. Further, the findings are not generalizable to community-based populations.

In the early 21st century, a report from the Institute of Medicine's (IOM) Committee on Quality of Health Care in America raised national awareness that the "American health care delivery system was in need of fundamental change." The wide difference between the care that should be received and the care that was actually being received was described as the quality chasm. A follow-up report found that the IOM's recommendations to transform the health care system were generally applicable to mental and substance use disorders but also identified distinctive characteristics of health care for mental and substance use conditions. Harold Pincus and colleagues noted that "the [Committee's] fundamental conclusion [is] that coordination and collaboration with the rest of medicine are essential (and vice versa), but it is also critical to ensure access to the depth and expertise provided by psychiatrist and other mental health and substance use specialists."

Within this same time period, national professional organizations developed practice parameters and treatment guidelines that enabled child mental health services researchers to develop quality indicators based on evidence of effectiveness and clinical consensus. Using documented

adherence to quality indicators in the medical record to estimate quality of care, in 1998 to 1999, the Caring for California Initiative (CCI) was the first statewide quality-of-care study for children with common psychiatric disorders in publicly funded outpatient mental health programs. The CCI found that on average less than one half (48.5 percent) of medical records passed all eligibility quality domains at the probable acceptable care level, a rate similar to general health care among US adult and child populations. Medication safety monitoring was particularly weak, such that roughly one fourth (28.5 percent) of the records of children receiving psychotropic medication treatment had documented monitoring of at least one clinically indicated vital sign or laboratory study. National concerns regarding the safety of psychotropic medication were also raised during this time period. In 2004, the U.S. Food and Drug Administration (FDA) issued a black box warning on all antidepressant medication to alert health care providers of an increased risk of suicidality in children and adolescents. Further, high national rates of stigma related to medication treatment raised awareness of the need for further examination of the mechanisms underlying engagement, uptake, and adherence to recommended care for children with psychiatric disorders and their families.

Although improving the quality of child mental health care has been a long-standing national priority, two federal health policies in the early 21st century legislatively mandated the development and use of quality measures for child health and mental health. In 2009, the Children's Health Insurance Program Reauthorization Act (CHIPRA) was signed. This law mandated the identification, refinement, and development of child health care measures for voluntary use in Medicaid and Children's Health Insurance Programs. It called for submission of initial core set of 24 quality measures to Department of Health and Human Services by 2010, and funded a 4-year Pediatric Quality Measures Program (2010 to 2014). Within this program, seven centers of excellence were funded to develop new measures and refine existing ones for "potential core enhancements" as well as 10 demonstration programs to develop, test, and evaluate and/or report on adherence to quality measures.

During the following year, the landmark Patient Protection and Affordable Care Act was signed in 2010. Within this complex legislation, the development of the National Quality Strategy was mandated. This strategy to transform health care had three aims: (1) better care; (2) healthy people and healthy communities; and (3) affordable care. The six priority areas were safety, person or family engaged as partners in care, effective communication and coordination of care, effective prevention and treatment, work with communities to promote wide use of best practices, and make quality of care affordable. This legislation also funded the Patient Centered Outcomes Research Institute (PCORI), to support research that informs better choices among alternative treatment strategies, supports a

patient-centered orientation and directs attention to individual and system differences that may influence strategies and outcomes. Within this funding source, research focusing on child health and mental health is included.

ACCESS TO CARE

Significant gaps continue to exist between clinical need and mental health service use for children with psychiatric disorders in the community. A number of conceptual models were developed to explain differential access to care. Of note, these models attempted to identify factors that can be changed through quality improvement interventions (referred to as *mutable factors*) to improve the likelihood of contact with services among children meriting assessment and treatment. A brief review of historically relevant access models is provided, because their frameworks continue to inform current quality improvement efforts.

Conceptual Models of Access to Care

Donabedian's early model of the medical care process acknowledges the complex interactions between client and provider behaviors, as well as the influence of contextual factors, such as the characteristics of the settings (e.g., formal and informal treatment settings and social, economic, and cultural factors) that take place in the process of seeking and receiving care (Fig. 55.10–1). From the client's perspective, service use and, ultimately, health outcomes are initiated by the recognition of need. This step is followed by the decision to look for care and the processes of seeking care, assuming the sick role, and maintaining and terminating care. On seeking care, provider behaviors also become influential and include diagnosis, therapeutic decision making, and potential collaboration with other providers during these processes.

Subsequent conceptual models further specify factors that are influential to the access to care. Ronald Andersen's Behavioral Model of Health Services Use, for example, proposes that predisposing (e.g., sociodemographic factors such as gender), enabling (e.g., availability of insurance or treatment resources), and need (e.g., illness severity) factors form the basis of the decision-making process. Alternatively, the Health Belief Model suggests that individual psychological attributes, such as attitudes, beliefs, and perceived benefits of treatment, represent the major determinants of seeking care. Other models of access, such as Bernice Pescosolido's Family Network-Based Model of Access to Child Mental Health Services, emphasize the role of social networks in this decision-making process.

Overall, help-seeking pathway models propose a series of links between the initial recognition of a mental health problem and the eventual use of

mental health services, framing these steps within the context of an "illness career" and network experiences thought to influence career progression as well as adherence to professional treatment suggestions. Few studies, however, have implemented study designs complex enough to analyze network influences, temporal sequencing, and combined multisector care patterns for youth. Studies addressing access to care have been most commonly based on the Andersen model, using cross-sectional designs to identify relevant predictors.

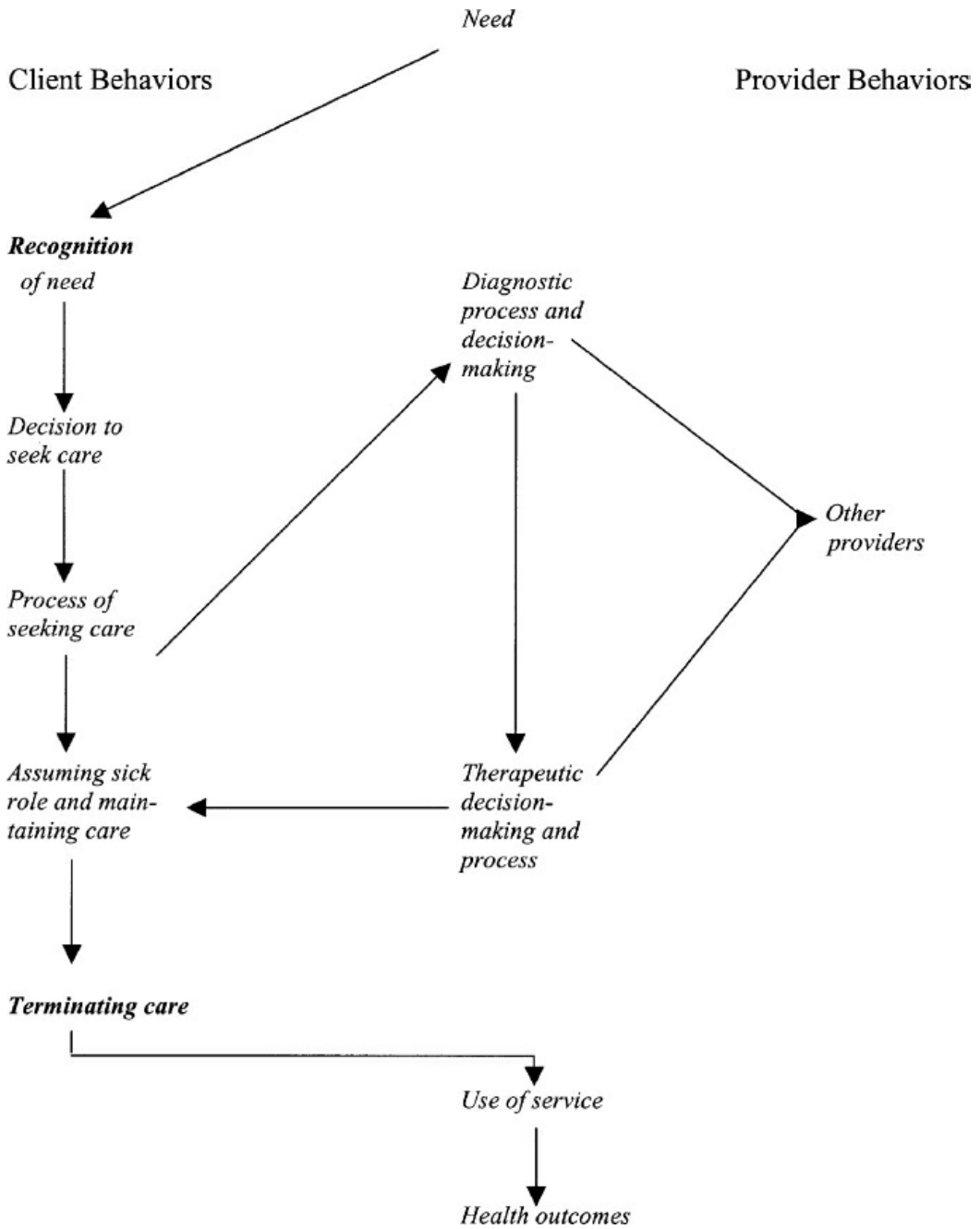

FIGURE 55.10–1. Model of the medical care process. The *arrows* represent only a small portion of significant interactions. (Reprinted with permission from Donabedian A. Promoting quality through evaluating the process of patient care. *Med Care*. 1968;7:183.)

Stigma as Barrier to Access

The 1999 Surgeon General's Report on Mental Health identified stigma as an important influence on access and use of mental health services. Youth with mental disorders face significant stigma in its various forms, including public (expressed as prejudice and discrimination), courtesy (stigma suffered by family members or associates of individuals with mental disorders), and self-stigma (stigmatized individual's acceptance of negative views by others). Successful recovery from mental illness requires awareness of stigma and of its potential adverse consequences and effective means to combat it. To examine factors that shape prejudice and discrimination toward children with mental disorders, the recent National Stigma Study-Children used vignettes depicting several stigmatizing scenarios, including ADHD, depression, asthma, or "normal troubles." Results indicate that scenarios of ADHD and depression and those depicting adolescents elicited the highest rejection responses. Respondents who thought a child suffered from a mental illness rather than from asthma or "normal" troubles indicated that they would seek social distance from the affected child, apparently resulting from perceptions that a child with mental illness is dangerous. In a comparison of child and adult depression vignettes, more respondents evaluated childhood depression as serious (83 percent) than adult depression (51 percent) and endorsed treatment of all types, including coerced care, for children with depression. However, high stigma surrounding childhood disorders poses unique access barriers for youth and their families and merits further study to identify ways to combat these harmful perceptions. General societal anti-stigma programs have been developed and implemented, consisting of three paradigms—protest, education, and contact—but insufficient guidance exists for treatment providers on how to combat stigma experienced by individual patients and their families.

QUALITY OF CARE

The Institute of Medicine defined quality as "The degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge." According to Elizabeth McGlynn and colleagues, quality-of-care research is "the conceptualization, measurement, and assessment of the outcomes, process, and structure of both the health delivery system and the context in which it operates." The study of the quality of care thus forges critical feedback loops among those receiving, providing, administering, and planning the delivery of care. For children, most of the early quality-of-care research focuses on descriptive studies of psychotropic medication prescription patterns and the effectiveness of model treatment programs (e.g., continuum care, multisystemic therapy, wraparound services,

therapeutic foster care, and medication treatment algorithms). Findings from these studies have raised questions about the appropriateness and safety of psychotropic medication use among children in the community. They have also underscored the complexities of developing rigorous study methods that can adequately test whether children and families in the community benefit from receiving care that is valued as "good" based on clinical consensus and scientific evidence. Together, these early studies stimulate further research that will ideally link child clinical need with more detailed information about the structure, processes, and outcomes of care over time from child, family, payer, provider, family, and child perspectives.

In response to national recommendations to integrate care, a body of research examining collaborative and integrated care models across primary and specialty mental health care is emerging. With the recommended use of quality measures, new research is examining the feasibility and clinical validity of applying national quality measures to child-level data in large community-based samples is also underway. To introduce this dimension of child mental health services research, a brief review of a well-established conceptual model for evaluating quality of care is provided, followed by summaries of select early quality-of-care research findings and national quality measures as exemplars.

Models of Quality of Care

Components of quality of care are often conceptualized as structure, process, and outcomes, following Donabedian's model of quality of care. Structural factors include system and clinic organizational factors, as well as provider, child, and family characteristics. Examples are health plan benefit structure, care sector, clinic volume, provider caseload, provider training and discipline, and parent education, as well as child insurance coverage, clinical profile, and sociodemographic characteristics. Processes of care are dichotomized into technical and interpersonal aspects of care. A common technical care process is medication treatment, which may, for example, be assessed using quality indicators or criteria pertaining to medication evaluation, use, and monitoring. For children, variables within the domain of interpersonal style may include the assessment of aspects of parent-provider communication and shared decision making. For example, how well a child and/or parent believes his or her concerns are listened to or how much involvement the primary caregiver had in the decision process of making the diagnosis, developing a treatment plan, and participating in the care of his or her child over time. Outcomes may also include child-, parent-, family-, community-, and system-level characteristics. Outcomes for children, parents, and families include symptom reduction, improved functioning, child and parent satisfaction, reduced parent strain, and improved family cohesion. Examples of

community- and system-level outcomes are improved school attendance and reduced costs, respectively. Using this well-established paradigm, most of the early quality-of-care research examines discrete dimensions of quality, such as effectiveness or concordance with a profession's normative standards.

Building on the Donabedian quality of care framework, McGlynn and colleagues proposed an integrative approach for examining quality of care for mental health. As depicted in Figure 55.10–2, the ordering of the components of quality of care are reversed, such that outcomes influence both care processes and structural factors and processes of care directly affect change in structural characteristics. They also suggested that in examining quality of care for mental health problems, an analytical strategy be used that "simultaneously examines the outcomes, the processes associated with producing those outcomes, and the structural context in which the process was delivered." Further, interpretation of findings related to clinical outcomes should be done carefully. Statistical techniques are often applied to adjust for the clinical severity and complexity of a child when examining quality of care. Missing data is also a common problem. For example, clinic follow-up visit rates are often poor, and without follow-up data, it is difficult to determine if the child dropped out of care because no further care is needed or not.

Quality of Care for Child Mental Health

Psychotropic Medication Prescription Trends. National trends in psychotropic medication use, based on secondary databases (i.e., pharmacy claims, encounter data, physician office visits, household surveys), have raised awareness of the need for additional research examining the safety and efficacy of psychopharmacological treatment regimens in larger, community-based child populations, particularly for younger children and those in foster care. Rates of psychotropic medication prescription also may vary by sociodemographic characteristics, psychiatric disorder, insurance status and Medicaid eligibility category. Among a statewide sample, a retrospective analysis of data on mental health and pharmacy services was conducted for 1,491 children age six and younger who were in foster care in 2010 and had at least 365 days in foster care during 2009 to 2011. Approximately 12 percent of children age 6 and younger in foster care for at least 1 year received at least one psychotropic medication over the 3-year study period. With each year of age, the prevalence of ADHD medication and antipsychotic medication and duration increased. In addition, based on data for 2- to 5-year-old children from the 1994 to 2009 National Ambulatory and National Hospital Ambulatory Medical Care Surveys, psychotropic prescription rates were 0.98 percent from 1994 to 1997, 0.83 percent from 1998 to 2001, 1.45 percent from 2002 to 2005, and 1.00 percent from 2006 to 2009. Boys, white children, older children (i.e.,

4 to 5 vs. 2 to 3 year olds), and those lacking private insurance were more likely than children from other groups to receive psychotropic prescriptions. Further, when examining 10-year trends among a statewide sample of children enrolled in Medicaid, the prevalence of antipsychotic use increased from 1.2 percent in 1997 to 3.2 percent in 2006. The increase in odds of antipsychotic use varied by Medicaid eligibility and was greatest for children enrolled in State Children's Health Insurance Program, followed by children in foster care, Temporary Assistance for Needy Families, and Supplemental Security Income. Increases in rates of psychotropic medication were greatest for youths with externalizing behavior disorders and bipolar disorder. The overall proportion youths using antipsychotics between 1997 and 2006 also increased significantly more among African Americans and Hispanics than among whites.

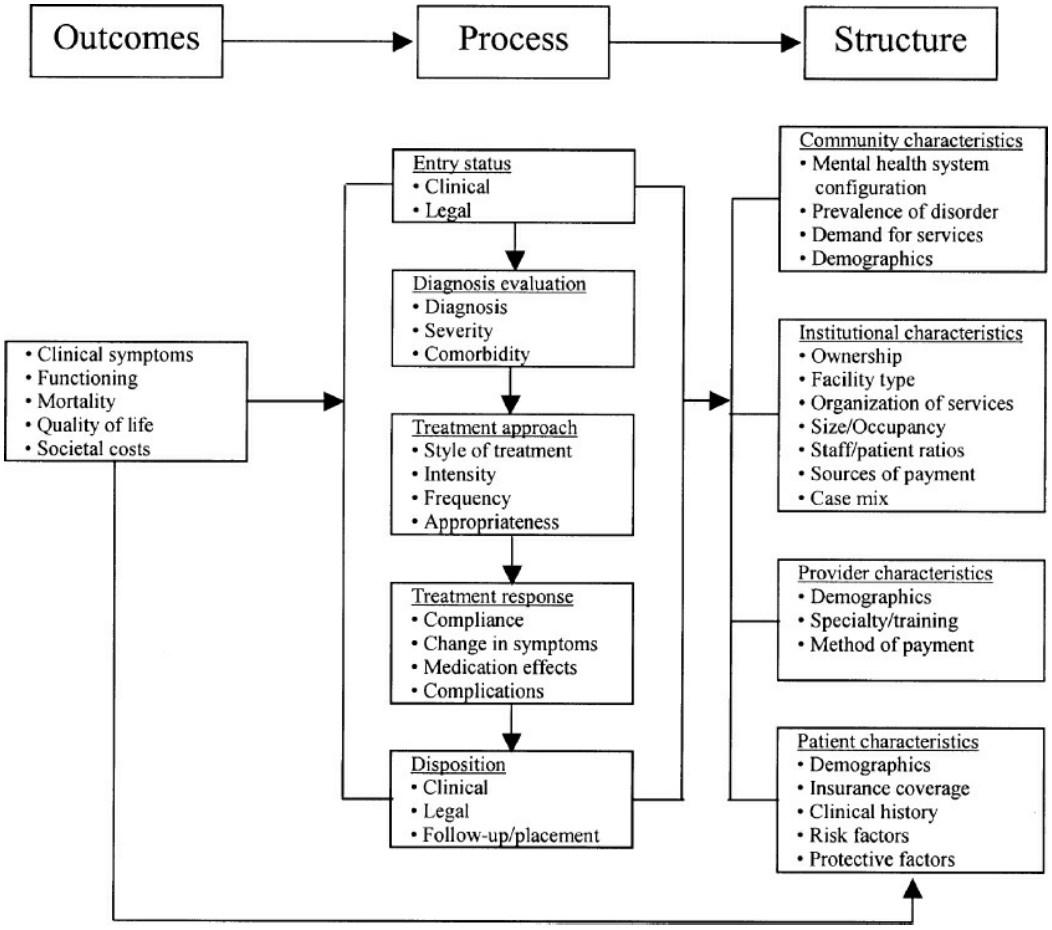

FIGURE 55.10–2. An analytical strategy for research on the quality of mental health services. (Reprinted with permission from McGlynn EA, Norquist GS, Wells KB, Sullivan G, Liberman RP. Quality-of-care research in mental health: responding to the challenge. *Inquiry*. 1988;25:157–170.)

Overall, trends in psychotropic medication use among children should not be interpreted as suggesting inappropriate use of such treatment for children. To begin to address this question, research that examines the

relationship between a child's clinical need for medication treatment and use of appropriate medication treatment over time is warranted. A longitudinal perspective is particularly key for child psychiatric disorders because they are often chronic, and target symptoms and functioning may change with child maturation, contextual changes (e.g., psychosocial stressors), and treatment response.

Gaps between Recommended and Usual Care. Findings from studies examining care processes for child psychiatric disorders suggest that gaps exist between recommended and usual care. Medical record data from a statewide sample of children receiving care for common psychiatric disorders in publicly funded outpatient mental health programs documented that adherence to quality of care indicators for clinical assessment was high (78 to 95 percent) but adherence to quality indicators related to service linkage, parental involvement, use of evidence-based psychosocial treatment, and patient protection were moderate to poor (8.0 to 74.1 percent). Medication safety monitoring was particularly weak, such that roughly one-fourth (28.5 percent) of the records of children receiving psychotropic medication treatment documented the monitoring of at least one clinically indicated vital sign or laboratory study. Using child-level data from multiple data sources, quality of care for ADHD in primary care and specialty mental health clinics in a large managed care Medicaid program was examined using a longitudinal cohort study design. Continuity of care was poor, and in primary care children received on average only one to two follow-up visits a year. In the specialty mental health clinics, less than one-third of children received any stimulant medication, and in both sectors, stimulant medication refill prescription persistence was poor (31 to 49 percent). With few exceptions, ADHD diagnosis, impairment, academic achievement, parent distress, and parent-reported treatment satisfaction, perceived benefit, and improved family functioning did not differ between children who remained in care and those that received no care.

Table 55.10–3.
Operational Definition of ADHD Medication Follow-up Quality Measure Developed by the National Committee on Quality Assurance

Initial Phase: Percentage of children age 6–12 years of age as of the Index Prescription Episode Start Date with ambulatory prescription dispensed for an ADHD medication and who had one follow-up visit with a practitioner with prescribing authority during the 30-day initiation phase.
Continuous and Maintenance Phase: Percentage of children who remained on ADHD medication for at least 210 days and who in addition to the visit in the Initiation Phase, had at least two additional follow-up visits with a practitioner within 270 days (9 months) after the Initiation Phase ended.

National Quality Measures. In response to national recommendations to improve the quality of health care, there has been a proliferation of

quality measures for behavioral health care, mostly dedicated to the care of adults and care in inpatient treatment settings. For children, within the CHIPRA initial core set of 24 quality measures for child health, three were related to mental health care. Between 2009 and 2011, the National Quality Forum reviewed 101 candidate measures reviewed for the Patient Outcomes (Phase III): Child Health and Child Health Measures Projects, and 15 pertained to child mental health care.

Across both initiatives, the well-established quality measure for ADHD medication treatment and follow-up were selected or endorsed. The operational definition of the measure is described in Table 55.10–3. This measure was developed by the National Committee on Quality Assurance (NCQA) and released in 2006. The data source is health plan-level data, but research examining application to real world data is emerging. One recent study found that even within integrated care delivery systems less than one-half of children received ADHD care consistent with NCQA guidelines for initiation and continuation/maintenance of ADHD medications, with higher quality rates for mental health than nonmental health providers. The 2012 NCQA Accreditation report sets the benchmark for the 90th percentile of the national initiation rate at 47 percent for commercial insurers and 51 percent for Medicaid, and of the continuation/maintenance rate at 55 and 63 percent, respectively. This is in notable contrast to benchmarks for asthma care, where the benchmark is set at 96 percent for commercial insurers and 93 percent for Medicaid for appropriate medication use. The study also highlights an important limitation of this widely used quality measure, namely, that of those youth started on ADHD medication, nearly two-thirds were ineligible for quality analyses because they did not receive continuous treatment (210 out of 300 days) required for inclusion in the denominator.

In addition, research to guide further refinement of quality measures and development of new ones is needed. The 2014 National Quality Strategy annual report found however that not all measures are useful, some are redundant, reporting adherence to quality measures can be burdensome, and there were concerns that quality measurement diverted resources away from other quality improvement activities. The development of new child mental health quality measures also poses methodological challenges. The constraints of existing data infrastructure, at the state and provider levels, must be addressed to enhance the capacity to capture data that links measure adherence to improved care and meaningful outcomes.

References

American Academy of PediatricsAd Hoc Task Force on Definition of the Medical Home: The medical home. *Pediatrics*. 1992;90:774.

Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? *J*

Health Soc Behav. 1995;36:1–10.

Bickman L. Implications of a children's mental health managed care demonstration evaluation. *J Ment Health Adm*. 1996;23:107–117.

Bussing R, Narwaney KJ, Winterstein AG, et al. Pharmacotherapy for incident attention-deficit/hyperactivity disorder: practice patterns and quality metrics. *Curr Med Res Opin*. 2014;30(8):1687–1699.

Chirdkiatgumchai V, Xiao H, Fredstrom BK, et al. National trends in psychotropic medication use in young children: 1994–2009. *Pediatrics*. 2013;132(4):615–623.

*Donabedian A. Promoting quality through evaluating the process of patient care. *Med Care*. 1968;6:181–202.

dosReis S, Tai MH, Goffman D, Lynch SE, Reeves G, Shaw T. Age-related trends in psychotropic medication use among very young children in foster care. *Psychiatr Serv*. 2014;65(12):1452–1457.

Froehlich TE, Lanphear BP, Epstein JN, Barbaresi WJ, Katusic SK, Kahn RS. Prevalence, recognition, and treatment of attention-deficit/hyperactivity disorder in a national sample of US children. *Arch Pediatr Adolesc Med*. 2007;161(9):857–864.

*Institute of Medicine Committee on Quality Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academy Press; 2001.

Jensen PS, Arnold LE, Swanson JM, et al. 3-year follow-up of the NIMH MTA Study. *J Am Acad Child Adolesc Psychiatry*. 2007;46(8):989–1002.

Martin A, Van Hoof T, Stubbe D, Sherwin T, Scahill L. Multiple psychotropic pharmacotherapy among children and adolescent enrollees in Connecticut Medicaid managed care. *Psychiatr Serv*. 2003;54:72–77.

McGlynn EA, Norquist GS, Wells KB, Sullivan F, Liberman R. Quality of care research in mental health: responding to the challenge. *Inquiry*. 1988;25:157–170.

Pappadopulos E, Jensen PS, Chait AR, et al. Medication adherence in the MTA: saliva methylphenidate samples versus parent report and medication effect of concomitant behavioral treatment. *J Am Acad Child Adolesc Psychiatry*. 2009;48(5):501–510.

*Pescosolido BA, Perry BL, Martin JK, McLeod JD, Jensen PS. Stigmatizing attitudes and beliefs about treatment and psychiatric medication for children with mental illness. *Psychiatr Serv*. 2007;58:613–618.

Pincus HA, Page AE, Druss B, Appelbaum PS, Gottlieb G, England MJ. Can psychiatry cross the quality chasm? Improving the quality of health care for mental and substance use conditions. *Am J Psychiatry*. 2007;164:712–719.

Stroul BA, Friedman RM. *A System of Care for Severely Emotionally Disturbed Children and Youth*. Washington, DC: CASSP Technical Assistance Center, Georgetown University Child Development Center; 1986.

*U.S. DHHS. *National Quality Strategy. Better Care. Affordable Care. Healthy people/healthy communities. Working for quality: achieving better health and health care for all Americans*. 2014.

Westfall JM, Mold J, Fagnan L. Practice-based research—"blue highways" on the NIH roadmap. *JAMA*. 2007;297(4):403–406.

Zima BT, Bussing R, Tang L, et al. Quality of care for childhood attention-deficit/hyperactivity disorder in a large managed care medicaid program. *J Am Acad Child Adolesc Psychiatry*. 2010;49(12):1225–1237.

Zima BT, Hurlburt MS, Knapp P, Ladd H, Tang L. Quality of publicly-funded outpatient specialty mental health care for common childhood psychiatric disorders in California. *J Am Acad Child Adolesc Psychiatry*. 2005;44:130–144.

*Zima BT, Murphy JM, Scholle SH, et al. National quality measures for child mental health care: background, progress, and next steps. *Pediatrics*. 2013;131(3):S38–S49.

Zito JM, Burcu M, Ibe A, Safer DJ, Magder LS. Antipsychotic use by Medicaid-insured youths: impact of eligibility and psychiatric diagnosis across a decade. *Psychiatr Serv*. 2013;64(3):223–229.

▲ 55.11 Impact of Terrorism on Children

WANDA P. FREMONT, M.D., AND JAMES S. DEMER, M.D.

Over the last few decades, terrorist attacks, threats, and their aftermath have been increasingly frequent in countries throughout the world. The effects of terrorist activities have a significant clinical impact on children and their families. A growing amount of research documents the effects of this exposure, especially data that help to differentiate reactions to terror from those associated with other types of violence. In addition, the indirect effects of terrorism, including exposure to media and television coverage of violent images, have been noted to be of increasing concern. To develop a proactive and strategic response to these reactions, clinicians, educators, and policy makers must understand the psychological effects of terrorism. Many of the effects of terrorism-induced trauma on children are similar to the effects of other human-caused and natural disasters. Therefore, it is essential for the clinician to have an understanding of children's reactions to trauma. The important contributing influences on children's reactions to trauma include risk and protective factors, developmental factors, neurobiological consequences, and family and community strengths and vulnerabilities. Recently, with the increase in terrorist bombings and access to media coverage worldwide, more studies have examined the unique effects of terrorism-induced trauma on children. This research covers a range of violent acts that are classified as terrorism. These include isolated events in countries not at war (the Oklahoma City bombing and September 11 terrorist attacks) and repeated terrorist attacks in areas of political conflict (terrorist activities in Northern Ireland, state terrorism in Guatemala, the Scud missile attacks during the Persian Gulf War, and bombings and kidnappings in Israel). Earlier studies completed in the 1970s and 1980s were primarily descriptive and anecdotal in nature. More recent studies have been more rigorous and focus on protective and risk

factors that affect children's reactions. They include longitudinal studies that assess the long-term consequences of terrorist activities. Diagnostic implications, clinical interventions, and treatment strategies have developed as a result of the increasing knowledge obtained about the effect of terrorism on children.

DEFINITION OF TERRORISM

The U.S. Defense Department defines terrorism as "the calculated use of violence or threat of violence to inculcate fear: Intended to coerce or to intimidate governments and societies in the pursuit of goals that are generally political, religious, or ideological." The goal of terrorism is not only to cause visible disaster, but also to inflict psychological fear and intimidation at any time—during periods of peace or conflict. Unlike family or community violence or the trauma resulting from war, terrorist activities are not confined to a specific geographical area or time. Terrorist events occur suddenly without any forewarning and frequently result in severe trauma. The threat persists indefinitely. New terrorist threats and actions may exacerbate underlying anxieties. Terrorism capitalizes on media coverage. Powerful visual images create strong emotional responses, including fear, panic, despair, and rage, even in individuals who live far from the traumatic events. The events are given well-recognizable labels (e.g., "the blood bus," "the explosive refrigerator in Israel," "the Chowchilla bus kidnappings," "the Scud missile attacks," "the 9/11 terrorist attacks").

CHILDHOOD REACTIONS TO TERRORISM-INDUCED TRAUMA

Although many of the reactions of children to terrorism-induced trauma are similar to children's reactions to other forms of human-caused violence and natural disasters, terrorism-induced trauma results in unique stressors and reactions in children. The clinician must have a good understanding of the variables that affect children's responses to trauma and the variety of responses that children manifest. It is therefore important to review the literature on childhood reactions to trauma and their clinical manifestations. Children vary in their reactions to traumatic events. Some suffer from anxiety and bad memories that dissipate with time and emotional support. Other children may be more severely traumatized and experience long-term problems that significantly interfere with their functioning. The reader is encouraged to refer to the section in this text on posttraumatic stress disorder in children and adolescents (Section 52.2).

Several factors contribute to a child's differential response to trauma. These include the child's age and level of psychological maturity, the severity and level of exposure to trauma, the duration of exposure to violence, predisposing risk factors, protective factors and resiliency, the influence of parents' responses to disaster, family protective factors, the degree of social and community support (including educational, political,

and religious influences), and the extent and context of displacement and relocation.

Children's emotional reactions may develop immediately after the trauma or may occur later. *Acute stress disorder* (ASD) is the most common psychiatric disorder following a traumatic event. *Posttraumatic stress disorder* (PTSD) is a reaction that may occur after a period of time. Children or adolescents suffering from acute stress disorder or PTSD may re-experience the trauma by having nightmares or recurring flashbacks of the trauma. They may become fearful, distractible, anxious, or depressed. Their play and study may be affected in addition to their pattern of sleep. In young children, repetitive play may occur in which themes or aspects of the trauma are expressed. Their dreams may be frightening but without any recognizable content. Trauma-specific re-enactment may occur. Adolescents and children may feel numb and withdraw from the external world or avoid situations that arouse recollection of the trauma. Symptoms of increased arousal may occur that include hypervigilance, sleep difficulties, irritability, difficulty concentrating, and outbursts of anger.

In addition to acute stress disorder and PTSD, children may develop comorbid disorders, including anxiety disorders (e.g., generalized anxiety disorder, social phobia, specific phobias, panic disorder, agoraphobia, obsessive-compulsive disorder), depression, and substance abuse–related disorders. Children may present with partial or variable symptoms of these disorders.

Children who are indirectly exposed to a traumatic event may also manifest symptoms of anxiety, including trauma-specific fear, fear of being alone, the habit of clinging to others, and event-specific fears.

Children's responses to terrorism-induced trauma include many of the symptoms and reactions to nonterrorist traumatic events. However, terrorism-induced trauma results in unique stressors and reactions in children. Terrorism, by its nature, is likely to cause psychological stress. Because terrorists can strike at any time and at any location, the threat of terrorism is omnipresent and enduring. The unpredictable, indefinite threat of terrorist events, the profound effect on adults and communities, and the effect of extensive terrorist-related media coverage exacerbate underlying anxieties and contribute to a continuous state of stress and anxiety. The resulting anticipatory fear and anxiety may be severe.

RISK AND PROTECTIVE FACTORS FOR CHILDREN EXPOSED TO TERRORISM

Several risk and protective factors influence children's reactions to terrorist events. These include factors relating to the history of the terrorist event and individual, family, and community strengths and vulnerabilities.

History of the Terrorist Event

The level of exposure to terrorist actions, either directly or indirectly, affects a child's response. The degree of exposure to terrorist actions is related to the prevalence of PTSD. The more severe the traumatic event, the greater is the risk of the child developing posttraumatic symptoms. Children who experience loss directly are more symptomatic. Suffering physical injury or witnessing the death or physical injury of others is associated with higher rates of PTSD and comorbid depression and anxiety. The degree of personal loss (i.e., the relationship of the child to the victim) is also correlated with the severity of PTSD symptoms. Knowing an injured or deceased person increases the risk of symptom development. Children who lose a parent manifest the highest rates of PTSD. In addition to the level of trauma, the duration of exposure predicts the risk for development of psychiatric problems in children.

Individual Factors

Individual factors and vulnerabilities that affect a child's differential response to terrorist events are related to developmental factors. These include a child's age and level of psychological maturity. Additional individual factors influencing responses include neurobiological responses, predisposing risk factors, protective factors, and effects of resiliency.

Childhood Development Influences on Terrorism Reactions. How children understand and conceptualize traumatic events resulting from terrorist activities is greatly influenced by their developmental level with respect to cognition, emotional, and social factors.

PRESCHOOL-AGE CHILDREN. Preschool-age children, between the ages 3 and 5 years, express their concerns for safety in terms of separation from parents and other primary caregivers. At this age, they are not able to cognitively understand the finality of death and do not fully understand the actual events that are occurring. Preschoolers rely on cues about danger from caretakers. They are more comforted than older children by reassuring and supportive parental behaviors. They may manifest their anxiety by having difficulties in separating from parents. Regressive behaviors, including thumb sucking, enuresis, and fear of the dark, are not uncommon in young children. In addition, children age 5 years and younger may manifest symptoms of fear, anxiety, sleep problems, and aggressive behaviors.

LATENCY-AGE CHILDREN. Latency-age children, 6 to 11 years, are prone to derive a sense of safety from clear-cut rules. At younger ages, they have difficulty abstracting and thus tend to polarize complex situations into "right" and "wrong" or "good" and "bad." As they mature, they develop an increasing sense of empathy and altruism. Children in this age group recognize the finality of death and are capable of fearing for others.

Latency-age children are extremely influenced by indirect effects of violence and trauma. Their reactions to terrorist events may include depression, PTSD, and anxiety. Symptoms of anxiety include irrational fears, sleep problems, nightmares, somatic complaints (headaches, stomachaches), irritability, angry outbursts, and school avoidance. Signs of depression may include sadness, withdrawal, feelings of hopelessness, and problems with attention. These problems may result in academic difficulties.

ADOLESCENTS. Adolescents, ages 12 to 18 years, are likely to be vulnerable to idealistic notions and an unrealistic sense of immunity to harm and omnipotence. This may influence their judgment regarding the degree of danger in their environment. When exposed to an extreme act of violence, such as a terrorist attack, young adolescents may feel overly responsible for protecting their parents from danger. Adolescent responses to terrorist events are more similar to adult responses. They may exhibit symptoms of acute or posttraumatic stress disorder, including intrusive thoughts, hypervigilance, and emotional numbing. Additional symptoms may include nightmares, sleep disturbances, avoidance, and withdrawal. They are at increased risk to develop depression, substance abuse problems, and peer problems. Shame, humiliation, self-blame, and guilt may result from a sense of powerlessness.

Neurobiological Consequences of Trauma in Children. Research on chronic stress such as child abuse and neglect and acute stress such as disaster, war, or medical trauma supports the hypothesis that trauma produces long-lasting effects on brain functioning. The human brain undergoes significant development during childhood and adolescence. Trauma may have a particularly devastating effect on the developing brain and may cause permanent changes in neurotransmitter systems that are stimulated by stress. Early childhood exposure to traumatic events influences the neurobiological response to stress in the developing brain in such a way that it predisposes affected individuals to mood and anxiety disorders later in life. An important mediator of the stress response is corticotropin-releasing factor (CRF). Charles Nemeroff and colleagues hypothesized that stress causes the hypersecretion of CRF. Dysregulation of the CRF system in the pituitary gland and brain may explain some of the symptomatology of traumatic syndromes, such as an increased startle response, and symptoms of anxiety and mood disorders. More research is needed to understand whether the neurobiological changes are reversible or permanent. An understanding of neurobiological consequences of trauma is important to determine treatment choices.

Predisposing Risk Factors. All children who are exposed to a traumatic

event experience a threatened sense of safety and potential for personal danger. However, children with predisposing risk factors are at greater risk to develop symptoms of anxiety and depression. Risk factors include exposure to past traumatic events during childhood, childhood anxiety, childhood conduct problems and antisocial behaviors, learning disabilities, and a history of psychiatric disorders. Children who have survived the death of a parent or another significant person are more vulnerable to the effects of trauma. In addition, each child reacts differently, depending on temperament and sensitivity and whether he or she tends to internalize or externalize experiences and emotions.

Protective Factors and Resilience. Many children who are exposed to trauma do not develop pervasive chronic posttraumatic symptoms. Factors associated with resilience among children after exposure to traumatic events are related to both psychological and biological characteristics. Resilience is related to the variability in psychological interpretation of physiological arousal. Initial biological responses to a fearful event include an increase in cortisol level in the brain and an increase in heart rate. Recovery from a traumatic event includes the ability to respond to the event in a supportive environment without extreme avoidance and in the absence of persistent physiological arousal. The development of a perpetual state of fear, with both emotional reaction and abnormal physiological response, leads to posttraumatic symptoms. Factors related to resilience include a child's characteristics and resources (persistence, motivation, adaptability, optimism, willingness to approach novel events, self-esteem, intelligence, social skills), developmental stage, family characteristics (support, open communication, consistency), and community support. Understanding risk and protective factors that contribute to a child's response to trauma is important to develop strategic interventions. Early intervention facilitates the process of recovery in vulnerable individuals.

Family Factors and Vulnerabilities

Children's response to trauma and their ability to cope with these disasters are strongly influenced by their parents' response to the violent events. There is a positive correlation between children's and parents' symptomatology. Parental involvement in mediating stress reactions in children exposed to terrorist attacks is a protective factor and helps to reduce a child's vulnerability to the trauma. Parents who underestimate their children's reactions may unintentionally hinder their children's process of mastering the trauma. Terrorist incidents affect adults profoundly. Therefore parents, adult caretakers, and teachers may not be able to provide the support and reassurance needed to help avoid potential long-term emotional harm to their children. Increased levels of distress and less adequate coping skills have been noted in children whose parents

responded to traumatic events with negative emotions. Positive coping responses in children are associated with parents who responded with positive emotional reactions. Family factors that have been shown to increase resiliency and buffer stress responses in children include a stable, secure emotional relationship with at least one parent, a parental model of constructive coping mechanisms, and physical proximity of children to parents.

Community Factors

Children's responses to terrorism are significantly influenced by community members' responses. An increased degree of social support and positive community influences increases resilience. Community educational, political, and religious support fosters moral development in children. The presence of a caring adult who is able to provide emotional support, encourage self-esteem, and promote competence has a moderating influence on acute and long-term mental health outcomes. Community value systems, beliefs, and ideology contribute to resiliency by giving meaning to dangerous events, allowing children to identify with cultural values, and enabling children and adults to function under extreme conditions.

CHILDREN'S RESPONSES TO MEDIA EXPOSURE TO TERRORIST EVENTS

Media coverage of terrorist activities has a significant effect on children's emotional responses. The media plays an important role in providing information to the public, especially during times of disaster. The images portrayed are powerful and intensely emotional. They are frequently repeated over and over again. Terrorist organizations depend on the media dissemination of their activities to increase responses of fear, thus creating an environment of persistent threat. Studies have shown that one-third or more children watch the news. Extensive research has examined the influence of violence and aggression on children and their families. Recently, with the increase in terrorist bombings and access to media throughout the world, more studies have examined the effects of children's reactions to media coverage of terrorism. The results research, covering a range of violent acts, has shown that remote exposure, such as viewing terrorism on television, has a significant clinical effect on children's responses.

Because media exposure is so extensive and the effect on children's reactions is significant, it is important to provide a summary of relevant factors pertaining to children's exposure, interpretation, and constructive and destructive effects of television and the media.

Several studies have examined the extent of children's exposure to the media, which has dramatically increased over the last decades. Forty-one

percent of US households have three or more televisions. Approximately 33 percent of children ages 2 to 7 years and 56 percent of children ages 8 to 16 years have a television in their room. US children ages 2 to 17 years spend an average of 19 hours, 40 minutes per week watching television. During 81 percent of the time that children ages 2 to 7 years watch television they are alone and unsupervised. Day care centers use television during a typical day 70 percent of the time. The average US child spends 900 hours in school and 1,023 hours watching television. In a survey of children and media conducted by Fairbanks and colleagues, 65 percent of children reported watching a television news program, and 44 percent of children read a newspaper. More than 50 percent of the children reported feeling angry, sad, or depressed after watching the news.

Childhood Developmental Influences on Media Responses

Children's understanding and interpretation of television and the media are related to their cognitive and psychological developmental levels. Children are less able than adults to regulate their emotions after watching traumatizing images.

Preschool-Age Children. Preschool-age children are most affected by visual images and emotional sounds. They are more upset by images of dead bodies or bloodied survivors and crying victims or witnesses than falling buildings or announcer commentaries. They are not able to understand the extent of the tragedy. Preschoolers cannot distinguish between live pictures and replays. If events are shown over and over again, they believe that the catastrophe is occurring repeatedly in real life. They are likely to believe that the terrorism is happening at the same time they are watching the coverage, and they believe that it is occurring in close proximity to their home.

Latency-Age Children. Latency-age children (ages 6 to 11 years) have a better understanding of the event but will interpret it concretely. They have difficulty in understanding the effect of the terrorist threats and adult responses to these threats. They are more sensitive to injured children or those whose parents were killed. School-age children are more likely to focus on their safety and their family's safety.

Adolescents. Adolescents are more capable of abstraction than are younger children. They are equally affected by visual images; however, they are also concerned about the effect of threats on the future and community responses. They are able to understand more complex, philosophical, and existential issues such as social justice, politics, and theology. Their personal struggles with identity, meaning, and place in society may be compromised seriously or distorted by the perception of the world and its

human interactions as dangerous, untrustworthy, and immoral. This worldview may have a negative effect on their sense of the future.

Although there has been extensive research about children's exposure to the entertainment media, there is far less scientific research about children's exposure to the news, especially exposure to terrorist media coverage. Media coverage of terrorist events has both positive and negative consequences.

Constructive Effects

Several constructive effects of the media have been emphasized. The media plays a vital role in disseminating information quickly to local, national, and international regions. Media coverage facilitates family and community cohesion. It facilitates the dissemination of educational information, and it can serve as a vehicle for discussion in older children and adolescents. The media, the Internet, and the telephone may serve as an extensive support system. Television stations are able to broadcast advice on how to help children to cope with stress and anxiety related to disastrous events. Data obtained after the September 11, 2001, attacks noted that sustained advertising through multiple media outlets was effective in informing individuals about counseling resources and encouraging them to seek mental health services.

Destructive Effects

The increased incidence of negative effects and stress-related symptoms in children exposed to media coverage of terrorist events has been examined recently in studies of the aftermath of the *Challenger* disaster, the Oklahoma City bombing, the Scud missile attacks and the Gulf War, the World Trade Center attacks, and the conflicts between Palestinians and Israelis. Repeated media exposure is especially detrimental for young children because of their cognitive immaturity. The effect of repeated coverage may promote the trauma and increase the intensity and duration of the effects. The media may also have a negative effect on parents and caregivers, which directly affects the coping abilities of their children. More research is needed to measure the effects of media coverage of terrorist events and to determine whether they have detrimental or positive effects on mental health outcomes.

LITERATURE REVIEW OF TERRORIST EVENTS

Early research on the effects of terrorism indicated that children exposed to violence are greatly affected by the experience. These studies examined children's responses to terrorist activities in Northern Ireland, state terrorism in Guatemala, and terrorist attacks in Israel. Children manifested many psychological symptoms, including anxiety, depression, anger,

feelings of hopelessness with regard to the future, difficulties with trust, anticipatory fears, and an altered sense of security. More recent studies have used statistically valid structured assessments to document children's reactions and symptoms. Research has been conducted in the aftermath of the 1993 World Trade Center bombing, the 1995 bombing in Oklahoma City, the 1991 Gulf War and its effects on Palestinian and Kuwaiti children, Scud missile attacks and terrorist activities in Israel in the last two decades, the 1998 American Embassy bombing in Kenya, and the September 11, 2001, terrorist attacks. These studies emphasized the importance of thorough assessment of children's responses and the influence of the media on children's reactions to a terrorist event. The rates of PTSD in children exposed to terrorist activities range from 28 to 50 percent. Other psychiatric problems have been identified, including anxiety, separation problems, depression, mood changes, sleep difficulties, regressive symptoms, and behavioral problems.

The importance of the influence of risk and protective factors on responses to terrorism has been identified. Research has increasingly focused on factors (personal, family, and community) that contribute to the persistence of psychological problems and factors that foster resilience. The results of these studies have contributed to the design of prevention and intervention strategies to help children and their families to cope and recover from the effects of terrorist attacks.

The 1991 Persian Gulf War and Scud Missile Attacks

Repeated terrorist attacks, including the Iraqi Scud missile attacks and terrorist activities on civilian sites in Israel, occurred during the war in the Persian Gulf in 1991. Further data measuring children's affective reactions and coping skills to terrorist events were gathered during studies conducted by several researchers during and after the attacks. The results of the research supported earlier data that correlated the degree of exposure to trauma to the prevalence of PTSD. Children who directly experienced loss due to terrorist attacks were more likely to develop symptoms of PTSD. The degree of personal loss was also correlated with the number of trauma symptoms. Knowing an injured or deceased person increased the risk of symptom development. The studies in Israel also focused on risk and protective factors and long-term consequences of terrorist activities. Several studies addressed the importance of parental involvement in mediating stress reactions in children. Earlier research identifying the protective effects of family relationships was conducted on Cambodian adolescent refugees who survived 4 years under the Pol Pot regime from 1976 to 1979, during which 1 to 2 million people were exterminated. Adolescents who were able to re-establish contact with a family member had fewer psychiatric symptoms than adolescents who lived alone or in a foster home. The natural families were successful in

maintaining traditional child-rearing practices. Clear adult–child boundaries and high expectations for conformity and responsibility were found in these families. Naomi Bat-Zion and Rachel Levy-Shiff noted that parents' responses were a central mediating factor for Israeli children's reactions to Scud missile attacks. Negative emotional expressions by parents were associated with increased levels of distress in children, whereas positive emotional manifestations were associated with increased coping efforts.

Laor and colleagues studied the long-term consequences on Israeli children of the Scud missile attacks. Children of mothers with poor psychological functioning showed more stress symptomatology in families who were displaced. The association between children's and mother's symptoms was stronger among younger children (ages 3 to 4 years). Unhealthy family interactions—either disengaged or enmeshed styles—exacerbated children's stress. These children manifested more symptoms, and their capacity to recover from the trauma was jeopardized. Conversely, children with adequate family cohesion manifested less stress in reaction to trauma and were better able to recover from the initial effect of the trauma.

Correlation between media exposure of terrorist activities and posttraumatic stress symptoms was verified in several studies. The effect of the media on children living in a war zone was also studied in Palestinian and Kuwaiti children during the Gulf War. Palestinian children who were exposed to traumatic events indirectly by the media had a higher incidence of anticipatory anxiety and cognitive expressions of distress than children who were exposed directly, who in turn manifested more symptoms of PTSD.

1993 World Trade Center Bombing

The first major terrorist attack in the United States occurred on February 26, 1993, when international terrorists bombed the World Trade Center in New York City. Six people were killed, and thousands of persons were trapped inside the building. One study compared the reactions of children who were present in the building at the time of the explosion to a community sample of children who were indirectly exposed (children whose exposure was limited to interpersonal relationships with persons directly affected and media coverage of the event). Children who were directly exposed to the attack were more symptomatic. Children in the comparison group had mild symptoms, which did not interfere significantly with their functioning. A marked discrepancy was noted between the parents' report of PTSD symptoms and fear in their children and the children's direct report of their fears. In a follow-up study 9 months later, parents of both groups noted a decrease in PTSD symptoms and fear in their children. However, the children did not report any significant decrease. The results of this study emphasize the importance of evaluating

children directly in addition to evaluating their parents.

1995 Oklahoma City Bombing

On April 19, 1995, the Alfred P. Murrah Federal Building, a US government office complex in downtown Oklahoma City, Oklahoma, was bombed. Until the September 11, 2001, attacks, it was the deadliest act of terrorism on US soil. A total of 168 people were killed, including 19 children, and more than 800 people were injured. Until then, few studies had examined children's reactions to terrorist events. Several studies were conducted in the aftermath of the attack, most notably those of Betty Pfefferbaum and colleagues. The significance of the effect of media exposure to terrorist attacks was identified in these studies. In one study, preschool-age children who attended day care in a facility near the Federal Building were examined. Although many of these children had minor injuries, none of them died or experienced serious injuries. Parents tended to underestimate symptoms in their children. All children met the exposure criterion for PTSD, although none of their parents reported diagnostic criteria for the disorder. Staff reports of children indicated that they met the exposure and re-experiencing criteria (posttraumatic play, discussion with peers). Only three children met the criteria based on parent report. The importance of direct observation of young children after exposure to trauma is essential for an accurate assessment of their reactions. The results of these studies also supported the literature documenting developmental differences in responses to trauma. Children were more likely to exhibit regressive behaviors and functional problems than the numbing and hyperarousal criteria. None of the preschool children reported symptoms of avoidance, restricted range of affect, or sense of foreshortened future that are typical of older children and adolescents.

Research on the Oklahoma City bombing also addressed the effects of media exposure to terrorist events and of indirect exposure on children's responses. Lenore Terr and colleagues first studied these effects of distant trauma after the Space Shuttle *Challenger* explosion in 1986. She examined the responses of children who had watched the explosion on television rather than directly witnessed the disaster. Terr and her colleagues defined distant trauma as "the reaction (memory, thinking, symptoms) to a disastrous event, observed at the time of the event, but from a remote, and safe distance." Children's responses included symptom patterns similar to PTSD. Children's reactions to distant trauma also included trauma-specific fear, fear of being alone, and clinging behaviors. Children who lived on the East Coast were more symptomatic than children who lived on the West Coast. East Coast children had a significant increase in dreaming, drawing, and behavioral enactment, trauma-specific fears (e.g., death and dying, explosions, fires, airplanes), and clinging behaviors. Children throughout the United States had diminished expectations of the future. Latency-age

children were more affected than adolescents. Most symptoms diminished within 1 year; however, for children with partial PTSD, symptoms (i.e., posttraumatic play for young children and hopeless attitudes about the future for adolescents) often persisted. Terr and colleagues defined a spectrum of trauma-related conditions that resulted from remote exposure to violence, including viewing media coverage of disaster events. Because one of the goals of terrorists is to inflict fear in the broader community, it is important to examine the effects of less direct and distant forms of exposure.

Pfefferbaum and colleagues assessed more than 3,000 middle and high school students 7 weeks after the Oklahoma City bombing. Most of the children heard or felt the blast, more than 40 percent knew someone who was killed or injured, and approximately two-thirds reported extensive television viewing of the bombing. Seven weeks after the event, one-third of the children reported continuing worry about themselves or their families. Female gender, acute reactions of fear and arousal, interpersonal exposure, and bomb-related television viewing were associated with an increase in posttraumatic stress symptomatology. Children who experienced personal losses were at greater risk of watching significantly more terrorist-related television coverage than were children without direct losses. This further exacerbated the traumatic experience. Children who witnessed the bombings on television reported trauma-related stress for more than 2 years. In addition to broadcast media, print media (i.e., newspaper and magazine) exposure was also examined. Print media exposure correlated more strongly with enduring posttraumatic stress than did exposure to broadcast media. Most of the children reported mild posttraumatic stress reactions and minimal effect on functioning. The clinical significance of the results is unclear, and further research is needed to determine whether they are at risk for future stress reactions.

1998 American Embassy Bombing in Nairobi, Kenya

More than 200 individuals were killed and thousands were injured on August 8, 1998, when a terrorist bomb destroyed the American Embassy in Nairobi, Kenya. Pfefferbaum and colleagues extended their studies of children who were indirectly exposed to terrorism, exploring issues that were not examined in the Oklahoma City studies, including exposure to other traumatic events. More than 500 indirectly exposed children in a convenience sample of Nairobi schoolchildren were assessed 8 to 14 months after the embassy bombing. Most of the children in the study lived in a community in which trauma was common and had experienced other crimes of violence and the death, illness, or injury of family or friends. Children who were at greater risk of developing symptoms of posttraumatic stress to the bombing incident had experienced a higher number of previous traumatic events, more severe posttraumatic stress reactions to

past events, and stronger acute reactions. Little or no impairment of children's functioning was noted. Similar to the Oklahoma City school samples, the children in this study experienced little direct exposure to the terrorist actions, and posttraumatic stress reactions and impaired functioning were relatively low. Future studies are needed to explore the significance of risk and protective factors, as well as a more comprehensive set of outcomes, including long-term longitudinal data.

September 11, 2001, Attacks

The deadliest terrorist events were the attacks on September 11, 2001. There were 2,973 fatalities, not including the 19 highjackers—246 on the four airplanes, 2,602 in New York City at the World Trade Center, and 125 at the Pentagon. These events significantly affected US and world history, governmental policies, and the lives of individuals in the United States and other countries. Indeed, "pre-9/11" and "post-9/11" have become commonly used terms worldwide. The terrorist attacks and their aftermath, including repeated media coverage of the events, have had a powerful effect on children and their families.

April 2013 Boston Marathon Bombings

On April 15, 2013 two pressure cooker bombs exploded during the Boston Marathon killing 3 people and injuring an estimated 264 others. A manhunt for the suspects followed. The subsequent investigation determined that the bombings were an act of terrorism. Comer and colleagues studied the psychological impact of exposure to this event on children. He examined the rates of PTSD in children who witnessed the event and the impact of media exposure. Children who witnessed the bombings were six times more likely to develop symptoms of PTSD. Two to six months after the attacks surveys from parents, living within 25 miles of the bombing and manhunt, showed that 11 percent of the parents identified PTSD symptoms in their children. This rate is similar to what was found among 9/11 New York City children 6 months after the attacks. Watching the manhunt on television was even more strongly associated with problems such as conduct issues or anxiety than exposure to the bombings themselves. Most children watched 1.5 hours of coverage on the day of the bombings, with 20% watching greater than 3 hours. About one-third of parents surveyed tried to restrict children's access to the media. Recognizing the potential psychological impact of media exposure on children, researchers advised that, "if there is a child who shouldn't watch an R-rated movie they should not be watching news with the same amount of violence and aggressive behavior."

Children's Reactions. Research on children's reactions to these attacks

corroborates the results of previous research studies. The degree of exposure, proximity to the disaster area, association with parental reactions, experience of distant trauma (including media coverage), and history of preexisting trauma were associated with increased stress symptoms. Children's reactions included symptoms of posttraumatic stress, depression, panic attacks, anxiety, separation anxiety, and agoraphobia. The largest study of children's responses was commissioned by the New York City Board of Education and was supported by the U.S. Department of Education through Project SERV (School Emergency Response to Violence). The study, conducted in collaboration with the New York City Department of Health and Mental Hygiene, surveyed 8,266 randomly selected children in grades four to twelve in the New York City metropolitan area 6 months after the September 11 attacks. Christine Hoven and colleagues surveyed children who attended schools in different areas of proximity to Ground Zero. Those children closest to the World Trade Center were most likely to have seen or heard the planes hitting the towers, seen the flames, been exposed to the dust and smoke, and struggled with fleeing to safety. Children farther away from the towers were more indirectly exposed to the attacks, including interpersonal exposure through direct victimization of a family member or friend or through extensive media exposure.

Seventy percent of all children had either direct or indirect exposure. The incidence of PTSD in this group increased from 2 to 10.5 percent following September 11. Children who lived outside of New York City showed an equal prevalence of symptoms of PTSD and other psychiatric problems (depression, panic attacks, anxiety, conduct problems) as children living within the city. The diagnosis of separation anxiety doubled and that of agoraphobia tripled. Children who were exposed to trauma before the attacks were more likely to experience PTSD and psychiatric problems. More than one half of the total sample of children who exhibited symptoms of PTSD had experienced previous traumatic events. This high frequency of previous trauma exposure emphasizes the need to address and identify prior trauma experiences when conducting needs assessments and intervention strategies following a terrorist event.

Parental stress reactions were also noted to have a significant effect on children's responses. The results of several studies noted a correlation between parents' responses and the frequency and severity of posttraumatic stress symptoms in children. Gerry Fairbrother and colleagues assessed predictors of posttraumatic stress reactions in children using a parent report administered by a New York City telephone survey. Parents reported that almost 20 percent of their children experienced severe or very severe posttraumatic stress reaction, and approximately two-thirds of children had moderate posttraumatic symptoms. Parental reactions were associated with the severity of posttraumatic reactions in

children. Mark Schuster and colleagues surveyed 560 adults throughout the United States 3 to 5 days after September 11 using random-digit dialing. Parents with substantial stress reactions were more likely to report that their children had symptoms. Thirty-five percent of adults surveyed reported that their children had one or more stress symptoms. Forty-seven percent of adults noted that their children were worried about their own safety or that of a loved one. This data continue to support the importance of evaluating parents' reactions to a disaster when assessing children's response and needs for treatment.

Further studies after the September 11 attacks assessed children's psychological and behavioral responses by surveying their parents, either by telephone or web-based interviews and assessments. William Schlenger and colleagues assessed psychological responses 1 to 2 months after the September 11 attacks using a web-based epidemiological survey based on the Posttraumatic Stress Disorder Checklist and the Brief Symptom Inventory. Adult reports indicated that children who were distressed by the events manifested symptoms of sleep disturbances, irritability, and fear in separating from their parents. Another study examined behavioral problems in New York City children by comparing three cross-sectional telephone surveys of adults. Two surveys were conducted after the September 11 attacks. The first survey was conducted 4 and 5 months and the second between 6 and 9 months after the attacks. The results of these surveys were compared to a representative survey conducted before September 11. The results of the study indicated that behavioral problems decreased in the first few months after the attacks but then returned to the predisaster levels after 6 to 9 months. These results support findings in an earlier study of children's behaviors before and after Hurricane Andrew.

Children's Reactions to Media Coverage. The effect of extensive terrorist-related media coverage on children's stress reactions were noted in several different studies after the September 11 attacks. Media coverage was extensive, and repeated coverage of the graphic images of planes striking the World Trade Center caused considerable concern about the added emotional effects of viewing the traumatic events. Results of a national study conducted a few days after September 11 concluded that children ages 5 to 19 years watched 3 hours of television news on the day of the attacks. Only 8 percent of children did not watch any coverage of the event. Parents who reported that their children were upset restricted their children's television viewing. In children whose viewing of television was not restricted, the number of hours of television viewing correlated with the number of stress symptoms. A study commissioned by the New York City Board of Education 6 months after the attacks noted that 62 percent of children spent much of their time and 33 percent of children spent some of their time learning about attacks on television. Approximately 75 percent of

children increased the amount of time they spent reading newspapers and magazines, and more than 30 percent obtained information from the Internet. A total of 10.5 percent (75,000) of students manifested symptoms of PTSD at 6 months after September 11, including children who were not affected directly by the event. Among these children, the prevalence of stress was higher among children who spent more time watching television coverage of the events than children who spent less time. The author noted, however, that these results did not determine a direct causality between television viewing and PTSD rates. Exposure to the images on television may have exacerbated children's stress responses; however, anxious children may have been more likely to seek information from the media than were nondistressed children.

The effects of television viewing of the September 11 attacks also were examined in adults. Specific disaster-related television images were associated with PTSD and depression among persons who were exposed directly to the disaster. The number of hours of television coverage viewed and an index of the content of that coverage were associated with PTSD symptoms.

Studies in several other geographical areas in the United States support the effects of media exposure. The indirect exposure of September 11 via television, the Internet, and printed media was examined in a study of elementary schoolchildren in the southeastern United States by Saylor and colleagues. PTSD symptoms were reported in children who saw reports on the Internet (vs. television and printed material), saw images of death or injury, or feared that a loved one might have died in the attacks. There was no measurable benefit to seeing "positive" or "heroic" images. Older children and boys were noted to have greater media exposure and more trauma-specific PTSD symptoms.

The correlation between television viewing of the September 11 attacks and an increased sense of insecurity and stress symptoms was also demonstrated in a study of elementary school–age children in Washington, DC. A few studies examined reactions of children who lived on the West Coast. One month after the attacks on the World Trade Center, adolescents in California experienced heightened perceptions of vulnerability to death, which extended beyond the terrorist attacks and generalized to unrelated risks. After exposure to the terrorist attacks, even from across the country, adolescents internalized a perception of the world as more risky and dangerous. Children in an ongoing longitudinal study who were assessed after the 9/11 attacks demonstrated posttraumatic stress symptoms and fears at levels comparable to children who directly experienced disasters.

Counseling Interventions. Few studies examined the frequency of counseling interventions following the September 11 terrorist attacks. In a telephone survey 1 to 2 months after the September 11 attacks, more than

100 parents were asked to describe the experiences and reactions of their children, most of whom were in school or day care at the time of the disaster. More than 20 percent of the parents studied reported that they received counseling. Parental posttraumatic stress, male gender, and having more than one sibling living in the household were factors that were correlated with receiving counseling. Two additional phone surveys of 423 and 112 parents, respectively, reported that 10 and 22 percent of symptomatic children received mental health services. These services were more likely to have been accessed by children whose parents reported having current symptoms of PTSD. Gerry Fairbrother noted that only 27 percent of children who had been identified in his first study to have severe or very severe stress reactions received any mental health treatment 4 to 5 months after the September 11 terrorist attacks.

In another survey, military health system beneficiaries in the Washington, DC, area after the September 11 attack on the Pentagon were studied using the Electronic Surveillance System of the Early Notification of Community-Based Epidemics (ESSENCE) database. The number of visits to behavioral health clinics for anxiety disorders and acute stress reactions in children increased after the attack.

Project Liberty. In the aftermath of the September 11 attacks, the largest federal health program in US history, Project Liberty, was created to provide mental health services to approximately 1.5 million people. It was funded through grants totaling $155 million from the Federal Emergency Management Agency (FEMA). The New York State Office of Mental Health (NYSOMH) administered and designed the program and collaborated with almost 200 local agencies to provide service delivery to New York City and ten surrounding New York counties. The overall goal of Project Liberty was to provide free and anonymous community-based disaster services to help individuals to recover from the psychological distress that they experienced as a result of the attacks. Services provided included disaster-related outreach, individual and group counseling, psychoeducation, and referral services for both adults and children. The services were developed based on the assumption that although most persons' distress reactions were disturbing to them, they constituted normal reactions to a traumatic event. Most services were planned to be of short duration, with the expectation that reactions would dissipate over a short period of time. Interventions helped people to identify their trauma responses and understand them as normal reactions to a traumatic event and to reconnect people with their preexisting social supports. Populations identified to need greatest support included victims' families, survivors and their families, displaced individuals, emergency and recovery workers, children, the elderly, persons with limited financial and social support resources, certain cultural and ethnic groups, and persons with a history of psychiatric problems. Project

Liberty provided face-to-face counseling and educational and outreach services to an estimated 1.2 million individuals in the New York City metropolitan area between the time services began a few weeks after September 11 through December 31, 2003, when services were terminated. An important component of Project Liberty was to provide assurance/quality improvement activities to assess the quality of the delivery and effect of services. Data on demographic variables, symptom clusters, variables predictive of further need versus recovery, and long-term outcome were gathered. In addition, data were collected to monitor how well the services provided adhered to treatment manual and guidelines.

As part of the quality assurance program, counselors completed encounter logs that included checklists for reactions to the September 11 attacks. These results were used to help guide program planning and responses to service recipients. Nancy Covell and colleagues collected data on service utilization and reaction patterns among children who received Project Liberty counseling services. Nine percent of Project Liberty community-based service recipients who received counseling between September 2001 and December 2003 were children. Census data for the 15 boroughs and counties serviced by the project indicated that children constituted 25 percent of the population. During a later phase of the project, funded by Project Liberty and organized through the New York City Department of Education, an additional 36,337 children received counseling services between January 2003 and December 2004. With these additional services, approximately 20 percent of Project Liberty's total services for individual (including family) counseling were provided to children. The most commonly reported reactions for children's responses paralleled responses recorded for adults. Reactions included sadness, fear, anxiety, concentration difficulties, anger, irritability, intrusive thoughts and images, and difficulty sleeping. Children's responses differed according to their age. Elementary school–age children were more likely to display signs of withdrawal and isolation, anxiety, and difficulties with concentration and memory. Older children (ages 12 to 17 years) were more likely to exhibit symptoms similar to adults, including avoidance and "numbing" reactions, in addition to substance abuse.

CLINICAL INTERVENTIONS AND TREATMENT

The effectiveness and comparative advantages of specific treatment modalities and interventions for children exposed to terrorism have not been extensively studied. Only recently have researchers developed rigorously designed studies to examine the effectiveness of treatments in children exposed to non–terrorism-induced trauma (e.g., sexual abuse, single-incident traumas, natural disasters, community violence, and war-induced trauma). The results of this research are promising and have significant implications for intervention strategies and treatment of

children exposed to terrorism. These include community, school-based, individual, and family interventions and treatments. Essential components used in all treatment approaches include addressing protective and risk factors. Emphasis is placed on providing assistance to parents and caregivers and including them in treatment. Psychoeducation, cognitive restructuring, exposure, and coping skills management are frequently used treatment strategies. Guidelines and suggestions specifically related to media coverage of terrorist events are also crucial interventions to prevent exacerbation of terrorist-induced trauma. Resources for information on children and terrorism are provided in Table 55.11–1. Guidelines for parents, health care providers, and educators are listed in Table 55.11–2, and guidelines related to media coverage are outlined in Table 55.11–3.

Community-Based Interventions

Based on their research into the effects of the 1995 Oklahoma City bombings, Robin Gurwitch and colleagues identified essential responses to community-wide acts of terrorism that included early community-based intervention; clinical needs assessment to identify children at risk; multimodal, trauma-loss–focused treatment programs; and program evaluation of treatment efficacy. Early community-based interventions focus on safety and protection. Restoration of rest and sleep, emotional support by parents and caregivers in the community, and stress-related symptom reduction are emphasized. Community-wide screening is critical to identifying children at risk. Screening may take place in schools, primary care settings, or neighborhood centers. Although there is no consensus on the most effective method of assessment, numerous structured interviews and self-reports are available. Screening identifies children at risk and children in need of acute, trauma-related service. An essential component of the screening process is distinguishing normal, developmentally appropriate reactions to trauma from abnormal reactions. Screening information may be used to develop specialized treatment programs based on the unique needs of the child, family, or community. Screening results also serve to provide pretreatment data for evaluating outcome studies of treatment effectiveness. Triage and referral, based on knowledge of psychopathology and risk factors, are essential components of intervention. Active involvement of parents, caregivers, health care providers, educators, and community leaders is emphasized.

Table 55.11–1.
Mental Health Resources on Children and Terrorism

American Academy of Child and Adolescent Psychiatry: http://www.aacap.org
American Academy of Pediatrics: http://www.aap.org/terorrism/
American Psychiatric Association, *Children and Terrorism:* http://www.apa.org/terrorism

American Psychiatric Association, *Disaster Psychiatry Handbook:* http://www.psych.org/disasterpsych/pdfs/apadisaster

American Psychiatric Publishing (APPI): http://www.appi.org/disaster_relief.cfx

American Psychological Association: http://www.apa.org

American Red Cross, *Facing Fear: Helping Young People Deal with Terrorism and Tragic Events:* http://www.redcross.org/disaster/masters/facing/fear

Center for Disease Control National Advisory Committee on Children and Terrorism (NACCT): http://www.bt.cdc.gov/children/

Center for the Study of Traumatic Stress: http://www.centerforthestudyoftraumaticstress.org

Federal Emergency Management Agency (FEMA), *FEMA for Kids: Resources for Parents and Teachers:* http://www.gov/kids/teachers.htm

Citizen Corps Newsletter: http://www.citizencorps.gov/news/enews/Home Box Office, "Through a Child's Eyes; September 11, 2001": http://www.hbo.com/schedule

Institute of Medicine, *Preparing for the Psychological Consequences of Terrorism—A Public Health Strategy:* http://www.iom.edu/report.asp?id=11573

International Society for Traumatic Stress Studies (ISTSS): http://www.istss.org

National Care Information Center, *Terrorism and Children:* http://www.nccic.org/poptopics/cope.pdf

National Child Traumatic Stress Network, *Psychological First Aid Field Operations Guide:* http://www.nctsnet.org/nctsn_assets/pdfs/pfa/PFA9–5Final.pdf

National Institute of Mental Health, U.S. Department of Defense, Justice, and Veterans Affairs, and the American Red Cross, *Mental Health and Mass Violence: Evidence-Based Early Psychological Intervention for Victims/Survivor of Mass Violence:* http://www.nimh.gov/healthinformation/massviolence_intervention.cfm

National Mental Health Association: http://www.nmha.org

National Mental Health Information Center, *Psychosocial Issues for Children and Families In Disasters: A Guide for the Primary Care Physician:* http://www.mentalhealth.samhsa.gov/publications/allpubs/SMA95–3022/defaulty.asp

New York University Child Study Center: http://www.aboutourkids.org

Substance Abuse and Mental Health Services Administration (SAMHSA) National Mental Health Information Center: http://www.mentalhealth.samhsa.gov/cmhs/EmergencyServices/

U.S. Department of Education, Emergency Planning, Office of Safe and Drug-Free Schools: http://www.ed.gov/emergencyplan/

U.S. Department of Health and Human Services, *Terrorism and Other Public Health Emergencies: A Reference Guide for the Media:* http://www.hhs.gov/emergency/mediaguide/PDF/

Community and school interventions, family therapy, adult therapy, and group therapy have been used to treat traumatized children and families, help them with coping skills, and address issues related to trauma and loss, but no systematized studies have shown their effectiveness. Psychoeducational group meetings (i.e., parents or community meetings) reduce symptoms of individual, family, and community fear and arousal levels.

School-Based Interventions

School-based interventions to treat PTSD symptoms have received empirical support in the literature. Annette La Greca and colleagues developed a school-based intervention manual to help children cope after Hurricane Andrew. Their manual has been distributed to schools

internationally and has been used in other disasters, including the Oklahoma City bombing. Robin Gurwitch and Anne Messenbaugh expanded the work of La Greca and colleagues after Hurricane Andrew and the work of Storm and colleagues after the Australian bushfires. They developed a manualized treatment program for use following community traumatic events, including acts of terrorism. The program, Healing After Trauma Skills (HATS), emphasizes the active involvement of parents and caregivers and includes educational materials on basic safety skills and psychological stress responses. It uses treatment exercises to address symptoms and behavioral difficulties associated with trauma and loss. Similar manuals and materials have been developed in the aftermath of the September 11 terrorist attacks. Among the materials that have been widely distributed internationally are *Helping America Cope: A Guide to Help Parents and Children Cope with the September 11th Terrorist Attacks* by La Greca and colleagues, and the American Red Cross publication *Facing Fear: Helping Young People Deal with Terrorism and Tragic Events*. The preliminary results of the interventions and manuals developed to aid children after terrorist attacks are promising. However, outcome data have not been examined, and further research is needed to evaluate their effectiveness.

Table 55.11–2.
Guidelines for Parents, Health Care Providers, and Educators: Terrorism and Trauma

Helping children and adolescents to overcome psychological problems in the aftermath of terrorist events is an important challenge for families and communities. Proactive and preventative methods to help children to cope with loss and anxiety may help to facilitate positive posttraumatic adjustment and allow for healthier outcomes in children's growth and development. Several important general guidelines and suggestions are recommended for parents, health care providers, and educators to help children in the aftermath of terrorist attacks.

Adults should focus on positive ways of understanding and coping with their responses to terrorist events. Since September 11, 2001, many resources have been made available to help parents to cope with the terrorism-related fears and anxieties. The National Mental Health Association, Centers for Disease Control and Prevention, and American Academy of Child and Adolescent Psychiatry are among some of the agencies that have published extensive information.

Any parent who is overwhelmed by the effects of the disaster should seek help from primary care providers, educators, mental health agencies, and clergy. If parents are not able to manage their responses to the trauma, they will not be effective in reassuring their children and helping them to discuss their concerns.

Children cannot be sheltered entirely from knowing about or reacting to the disturbing event. Their fears of not knowing or understanding what has happened may be more disturbing than the truth. It is therefore important for parents to present information to children according to their developmental level. Children should be encouraged to express their concerns, and their feelings should be validated. Listening is more important than talking. All questions should be treated with respect and not ignored or dismissed. Responses should be made with calm reassurance and empathy. If a child's question causes anxiety or discomfort, the adult should not remain silent, because the child might interpret the silence as a sign of danger. It is permissible to tell the child that

the question cannot be answered immediately but will be responded to after the adult has had time to think about it.

Younger children may act out their feelings rather than verbalize them. When possible, children should be encouraged to express their feelings in words. Answers to their questions should be honest and brief. More information than is requested should not be provided.

Older children may be hesitant to talk about their feelings. Discussing terrorist activities with them, without overwhelming them, and encouraging questions or correcting misperceptions are helpful.

Some children's and adolescents' responses may be confusing. Their defenses may be covering up their fears and anxieties. For example, they may make comments such as "We should not be upset, nothing has happened to us." Just as adults should be aware of their own defenses and coping styles, they should be sensitive to children's defenses and ways of coping.

If a child chooses not to talk about the events, the decision should be respected. Children may not be ready to share their experiences immediately; they may wait several days or weeks. Adults should make sure that children know that they are available to listen and discuss their feelings whenever the children need them.

Adults need to reassure children that they are safe and will be protected. Maintaining familiar routines helps children to know what to expect and is comforting to them. Reassuring children that they are loved and that they will be cared for will give them a sense of security. However, they should not be offered false assurance. Children should be taught to manage their fears and anger constructively by showing them how to handle their emotions in healthy ways. They can be made to feel more in control by allowing them to make decisions (meals, what to wear, what movie to see, etc.)

Children may be encouraged to draw pictures or write about their feelings. An important way for children to work through their worries is play. At times, it may be helpful for an adult to redirect the play into caring or helpful themes (such as pretending to make meals for emergency helpers or building hospitals). Focus and attention should be given to those who are helping in the community and nation at that time. Children may feel empowered by participating in events to help those who have suffered. Offering children opportunities to channel their feelings into positive actions allows for healthier outcomes.

Children should be given a reasonable sense of hope for the future. An example may be giving children historical references to how a country overcame previous threats to become a strong nation. Other examples include sharing past problems that children have had with classmates or family members and how they resolved their difficulties, or by giving examples of conflict and resolution within the community.

Children should be educated about racism and stereotyping. They should be told that discrimination and prejudice are unacceptable. Children learn from adults, who must be careful to monitor their own morals and values in front of their children.

Most children will develop healthy coping skills if given the support and reassurance from adults. Some children need more help, especially if they are at greater risk or have experienced the loss of a loved one. Children who may need more intensive help should be referred to a mental health professional. Children's prolonged responses to anxiety and stress may interfere with their ability to function at home, at school, or with peers. The signs of stress may be similar to those seen in adults: Avoidance (refusing to go places that remind them of a traumatic event), emotional numbing (diminished response or lack of feeling), re-experiencing trauma (through nightmares or daytime recollections), and hyperarousal (tendency to be easily startled, sleep disturbances). Children's responses also may differ from those of adults and may include disorganized or agitated behavior rather than fear and helplessness, constant repetitive play of themes of trauma that preoccupy and distract them from their daily activities, and frightening dreams that are not related to terrorist activities. In these cases, children should receive support and help from a mental health professional to prevent these symptoms from continuing to cause problems in the future as children grow.

Individual and Family Interventions

Children and families who manifest significant psychiatric impairment and dysfunction need more intensive health treatment. Cognitive-behavioral therapy (CBT) has been the most rigorously studied treatment for traumatized children. Randomized, controlled studies have shown the effectiveness of CBT in targeting symptoms of PTSD and anxiety in sexually abused children. Additional studies have extended CBT methods to children exposed to other types of trauma. Cognitive restructuring (reprocessing the traumatic event and identifying traumatic triggers), relaxation training, anger management skills, and grief management are important methods of treating PTSD after disasters. However, the effectiveness of CBT for treating children exposed to chronic trauma, comorbid disasters, past history of trauma exposure, and serious family dysfunction has not been studied.

Table 55.11–3.
Guidelines Related to Media Coverage

Because the media plays a major role in informing the public about terrorism and may be a causative agent in symptom formation, it is critical for parents and families to have guidance about the use and misuse of media in traumatic times. Several media-related guidelines have been identified to help prevent traumatizing children specifically from media coverage of terrorist events. Parents, educators, health care providers, politicians, and journalists must increase their awareness about the potentially harmful effects of media exposure on children. Parents should closely monitor disaster-related television viewing by children.

Children's exposure should be limited. Adults would also benefit from limiting their own exposure to reduce their stress reactions so that they and their young children are better able to cope.

Vulnerable persons who have risk factors, such as young children and parents with previous traumatic experiences, should be especially cautious. The use of media should take into consideration a child's developmental level. Young children should not watch adult-oriented shows. Older elementary school–aged children may tolerate being exposed to news "briefs" in which simple concise statements of the facts are made (e.g., reports that occur on major networks between television shows) but may not be ready to watch news shows that spend 30 minutes with interviews, photos, and more detailed discussion of devastation.

Parents and caregivers are advised to watch the coverage with children when they are permitted to watch television. This supervision provides adults with the opportunity to observe children's signs of distress and be available to answer questions and discuss relevant topics.

School staff also must be aware of the importance of monitoring media exposure and should consider policies related to viewing coverage of traumatic events.

Constructive uses of the media should also be taken into consideration. The media may be helpful for parents by providing them access to valuable information to help them when dealing with children in the aftermath of terrorist attacks. Media reports can be used to help reassure children that many adults are working to maintain a safe and protected environment for them Adults may choose to offer taped news reports or specific types of media discussion that will help a child feel secure in the knowledge that adults are taking charge and responding in ways that are protective. Films about resilience, survival, family and community consolidation, and triumph over adversity may play a constructive role by giving children a reasonable sense of hope for the future. Long after an event, children may watch shows with their parents to help them understand terrorism, for example, "Through a Child's Eyes; September 11, 2001," which is an Emmy Award–winning HBO children's special that deals with young school–age children's responses to the September 11 attacks.

Minimal data are available on the effectiveness of pharmacotherapy for treating symptoms of PTSD in children. Sertraline is the only medication that has been approved by the U.S. Food and Drug Administration (FDA) to treat PTSD in children ages 12 to 18 years. Although selective serotonin reuptake inhibitors (SSRIs) have not been approved by the FDA, they have been shown to have therapeutic effects on symptoms of depression in children. Because symptoms of anxiety and depression are common in children following trauma, treatment with SSRIs has been considered a treatment option. However, caution must be used because of concern about the potential increase in suicidal ideation when these medications are used in children. Close monitoring for adverse side effects of suicidal thoughts are necessary when prescribing SSRIs in young patients.

Certain limitations exist when applying trauma-specific interventions to children who have experienced ongoing violence and the continual threat of terrorism. Crisis intervention alone does not prevent the long-lasting effects on children exposed to terrorism. Compounded effects of multiple traumas have not been addressed. The unpredictability of terrorist attacks, the indefinite nature of violent threats, the lack of specific geographical boundaries, and the effects of media coverage result in unique stressors and pose specific challenges for the treatment of terrorism-induced trauma in children.

FAMILY TRAUMATIC GRIEF

Ben was a 7-year-old boy whose maternal uncle, a firefighter, died in the September 11, 2001, World Trade Center attacks. His uncle did not return with his company on that day. Ben was first seen in treatment 1 month after September 11. He was at school when the attacks occurred. His father picked him up. After they returned home, Ben, his parents, and his 5-year-old sister watched the news and the repeated images of the falling towers. Ben knew that his uncle was a firefighter, but he did not know whether he had responded to the attacks. Ben's mother was very distraught as she tried desperately to find out what happened to her brother. Ben's father stayed home from work during the first week after September 11 and assumed most of the household chores and child care, which his mother had been primarily responsible for prior to the attacks. Ben was very worried about his mother. He was aware that the family was concerned about his uncle, and he asked frequent questions about him. His parents had been ambivalent about what to tell their children. Initially, when they did not know for certain whether Ben's uncle had died, they were hesitant to accept the possibility of his death and maintained hope that he was alive. They told their children that they were hopeful that he would be found. Both parents made a conscious effort to make sure that their children were not exposed to television coverage. However, after the children were in bed, they spent several hours watching the news coverage. They talked to their family and friends by phone and tried to protect their children by talking in another room while the children were playing. After Ben's parents were notified of his uncle's death, they struggled with their feelings of shock and grief. They did not tell their children about their uncle's death for several days. By continuing to "not know," Ben's parents may have reacted to their terror and loss by using protective defenses of shock and disbelief. Ben and his sister continued to ask questions: "Was Uncle Gary one of the firefighters who died trying to rescue the people in the tower?" The children's direct questioning challenged the adults' defenses of denial. Ben's mother was irritated by his frequent questions and then felt guilty for being angry with him. Ben in turn became upset, whiny, and oppositional. He also became very clingy and insisted on being with his parents, following them to each room of the house. He did not want to go to the bathroom by himself. At night, Ben would join his parents in their bedroom, which was very uncharacteristic of him.

Ben's parents eventually told their children that their uncle had died. Ben continued to ask questions about his uncle's death. "How did he die? Did he burn to death?" His mother, in a state of traumatic bereavement, did not know how to answer his questions. Ben's father tried to reassure his son that his uncle was now in heaven.

When school resumed the following week, Ben refused to go. He complained of "belly aches" in the morning. When his parents tried to help him to get ready for school and reassure him, he started crying. He refused to get dressed for school and had temper tantrums. Ben's mother took him to school. He went willingly with her, but when it was time for him to leave her and enter the classroom, he started crying and refused to let go of her. She struggled seeing him in so much anguish, and after speaking with the school counselor, decided to take Ben home for the day. The school counselor noted that several of the children had not yet returned to school. Ben had no prior history of school problems and had always enjoyed attending school.

Ben continued to have difficulties separating from his mother and continued to refuse to go to school. He no longer asked questions about his uncle, and his parents avoided talking about him in front of their children. Ben continued to complain of stomachaches. His parents became increasingly concerned and made an appointment for him to see his pediatrician. They noticed that Ben's physical complaints were more frequent on the days he was expected to go to school, but they wanted to make sure that he did not have an underlying medical problem. Ben's pediatrician completed a full medical workup and did not find any underlying organic problems that could be contributing to his complaints of abdominal pain. She reassured Ben's parents that no further medical tests were needed. She explained to them that she believed Ben's current difficulties were related to the recent trauma that the family had experienced and referred them to a child therapist.

The child therapist met with Ben's family initially. Ben's parents focused on their worries about their son. They felt that his problems were related to the recent traumatic events and the death of his uncle. They acknowledged that they felt overwhelmed and were grieving the loss of his uncle. Ben said that he was worried about his parents, especially his mother. "She pretends she is OK, but I don't believe her." The therapist then met with Ben alone. Ben said he was worried about his mother. He was also afraid that another attack would occur. The therapist commented that this was a scary time, and a lot of children were frightened, sad, and mad. Ben was quiet and did not answer. While they were talking, Ben drew a picture of his uncle's fire-truck. The therapist asked Ben about his uncle. He said that he liked to spend time with his uncle who had "awesome baseball cards." Ben and his uncle's family would get together every Sunday. Ben commented that the Sunday before his uncle died, he went to a friend's birthday party while his family visited his uncle's family. He was very sad that his uncle died, but he also felt "bad" that he chose to go to his friend's party instead of spending time with his uncle. The therapist assured Ben that his uncle knew that Ben loved him. She explained to Ben that he did not cause his uncle's death. Ben responded, "I know, but I felt like I jinxed him."

When the therapist met with Ben's parents alone, she asked them how they were coping. Ben's mother noted that in addition to mourning the death of her brother, she was also very frightened that another attack would occur. She had frequent nightmares and was afraid to leave the house. She was afraid for her husband's safety. He worked long hours and commuted to his job in Manhattan by subway. Her main supports were her sister, who lived out of state. She had a few close friends, but their families were also affected by the September 11 attacks, and she did not want to be a burden to them by talking about her fears and loss. Ben's father was also worried about his wife. He noted that he coped with stress "by being a workaholic." Both parents continued to watch the television coverage of the events, usually at night after their children had gone to bed. They stated that they tried to avoid talking about the terrorist attacks in front of their children because they "did not want to upset their children further."

The therapist continued to meet with Ben and his family. Family sessions were held, in addition to individual sessions with Ben and sessions with his parents. Family sessions focused on encouraging the family to communicate more directly with each other and to express their feelings of grief and anxiety. The therapist helped to normalize Ben's parents' reactions by explaining to them that their feelings of fear were normal reactions to an extreme situation. She encouraged them to use their supports, including their friends and family. She helped them to develop coping strategies. Ben's mother had stopped exercising after the attacks, an activity she had previously enjoyed. Ben's father was able to reduce his work hours and spend

more time with his family during the crisis. Both parents were advised to decrease the amount of time they spent watching television. The therapist explained to them how their reactions and defenses affected their children. Ben's parents were given guidance regarding how to talk to their children. They were encouraged to listen to their children, and they were taught how to understand their children's feelings and behaviors. In particular, the therapist helped Ben's parents to understand how his difficulties with separation and school refusal were related to their responses and anxieties, especially his mother's worries. The therapist helped Ben's parents to understand his experience and how he was feeling guilty that he may have caused his uncle's death by not visiting him. The importance of returning to a regular routine and ordinary daily life was emphasized. This included making sure that Ben's parents were consistent in helping him to attend school every day. To facilitate his return to school, the therapist spoke with his school counselor to collaborate on treatment. It was agreed that Ben's mother would drop him off at school and the school counselor would meet him at the school door. Ben would meet with the counselor every day during the first week. He would also be allowed to call his mother during the first week. The therapist and school counselor agreed to collaborate and communicate regularly to make appropriate adjustments to the treatment plan as needed. In addition, once Ben returned to school, he would participate in a peer group that the school counselors had developed to help children with similar experiences and losses.

The therapist also met with Ben individually. She helped to reassure him that the adults in his life, including his parents and teacher, would do their best to make sure that he and the other children were safe. Ben's therapist helped to normalize his feelings, explaining to him that other children also experienced feelings of sadness, anger, fear, and confusion. She facilitated his symbolic expression of his feelings by supporting and joining him in play and in art projects. She listened to him when he expressed his feelings verbally and clarified his feelings by helping him to find words to name his emotions. She avoided probing questions, and followed Ben's lead. Ben's attachment and identification with his uncle was actively facilitated by helping him remember and talk about his uncle.

FUTURE DIRECTIONS

The continuing frequency of terrorist events around the world and their effect on children and families underscore the urgency for an effective public health approach for providing for the needs of children and their families. Mental health organizations and providers must be educated and prepared to help children in the aftermath of terrorist activities. Terrorism is a community mental health issue because terrorism is a form of psychological warfare whose target is to intimidate and cause fear. Several mental health activists have made recommendations to help communities organize and plan appropriate interventions and treatments. Collaboration among community, educational, and mental health care organizations is essential. In addition, collaborative efforts between mental health organizations nationally and internationally have been initiated and continue to grow and develop. In the United States, the mental health approach to disaster has developed over several decades, primarily based on experiences with natural disasters. Following the 1995 Oklahoma City bombings, many modifications to the federal approach were made, including more effective crisis intervention methods and better triage and referral services, to address the unique effects of violent acts of terrorism. The effect of the events of September 11, 2001, further emphasized the difficulties in identifying the extent of services needed, the scope and depth of personnel and funding necessary to provide services, and the anticipated

duration of treatment. The lack of sufficient data and comprehensive information prevented adequate funding by federal mental health agencies to address mental health needs of children and their families. Obtaining money and resources will continue to be a significant challenge. Advocacy for program development and funding to increase community services for children affected by terrorism is needed.

Trained mental health providers who have experience working with children are needed to identify children at risk and provide community-wide services and more intensive treatment interventions. Standards for mental health providers who work with children and families exposed to terrorist events should be established.

The correlation between media exposure of terrorist events and traumatic stress symptoms in children is well documented. Collaboration between community agencies and the media is necessary. Public policy limiting media coverage is a controversial issue but needs to be addressed.

The research on childhood reactions to terrorism-induced trauma highlights the need for a systematic public health approach that provides community-based interventions, screening of children at risk, triage and referral services, community-based trauma and loss programs, and multimodal outcome studies. During the last decade, policy and procedures for the conduct of organized interdisciplinary research have received more attention. These advances resulted in more cooperation between researchers to address the issue of how ethically to obtain research data during times of crisis. Planning to incorporate research infrastructures into disaster preparedness plans is necessary. Further research on terrorism and its unique effects on children and families should be encouraged. More research on the effects of early interventions and treatment is needed to provide for reliable mental health response planning. Further research identifying children who are at risk for developing PTSD and related symptoms will allow for more efficient triage services. Studies on longitudinal effects and the long-term effect on development in children exposed to terrorist violence will provide information useful for initial and later treatment interventions. More reliable and objective measures are needed to quantify parents' reactions and the effects of social supports. Methods for evaluating a child's intrinsic resiliency will help to differentiate coping models learned from parents. Assessment of the capacity of the public health system to respond to terrorist events will help to determine optimal response recommendations for the future. Efforts should be focused on increasing community services for children affected by terrorism and encouraging further research on terrorism and its unique effects on children and families. Responses from mental health providers and organizations, educators, spiritual leaders, and community organizations will be better able to help children to cope and will minimize long-term psychological problems to enhance healthier outcomes and

growth.

REFERENCES

Ahern J, Galea S, Resnick H, Kilpatrick D, Bucuvalis M. Television images and psychological symptoms after the September 11 terrorist attacks. *Psychiatry*. 2002;65:289–300.

Allwood MA, Bell-Dolan D, Husain SA. Children's trauma and adjustment reactions to violent and nonviolent war experiences. *J Am Acad Child Adolesc Psychiatry*. 2002;41:450–457.

Bat-Zion N, Levy-Shif R. Children in war: stress and coping reactions under the threat of Scud missile attacks and the effect of proximity. In: Leavitt L, Fox N, eds. *The Psychological Effects of War and Violence on Children*. Hillsdale, NJ: Lawrence Erlbaum; 1993.

*Bonger B, Brown LM, Beutler LE, Breckenridge JN, Zimbardo PG, eds. *Psychology of Terrorism*. New York: Oxford University Press; 2006.

Coates SW, Rosenthal JL, Schechter DS. *September 11: Trauma and Human Bonds*. Hillsdale, NJ: Analytic Press; 2003.

Comer JS, Dantowitz A, Chou T, et al. Adjustment among Area Youth after the Boston Marathon Bombing and Subsequent Manhunt. *Pediatrics*. 2014;134(1):7–14.

Covell NH, Allen G, Essock S, Pease EA, Felton CJ. Service utilization and event reaction patterns among children who received project liberty counseling services. *Psychiatric Serv*. 2006;57:1277–1282.

*Danieli Y, Dingman RL. *On the Ground after September 11: Mental Health Responses and Practical Knowledge Gained*. New York: Haworth Press; 2005.

Fairbrother G, Stuber S, Galea S, Pfefferbaum B, Fleishman AR. Unmet need for counseling services in children in New York City after the September 11th attacks on the World Trade Center: Implications for pediatricians. *Pediatrics*. 2004;113:1367–1374.

Fremont WP, Pataki C, Beresin E. The impact of terrorism on children and adolescents: terror in the skies, terror on television. *Child Adolesc Psychiatric Clin N Am*. 2005;14:429–451, viii.

Gurwitch RH, Messenbaugh A. *Healing after Trauma Skills: A Manual for Professionals, Teachers, and Families Working with Children after Trauma/Disaster*. Oklahoma City, OK: Children Medical Research Foundation; 2001.

Hoven CW, Duarte CS, Lucas CP, Wu P, Mandall DJ. Psychopathology among New York City public school children 6 months after September 11. *Arch Gen Psychiatry*. 2005;62:545–552.

*La Greca AM, Sevin SW, Selvin EL. *Helping America Cope: A Guide to Help Parents and Children Cope with the September 11th Terrorist Attacks*. Coral Gables, FL: 7-Dippity; 2001.

La Greca AM, Silverman WK, Vernberg EM, Roberts MC, eds. *Helping Children Cope with Disasters and Terrorism*. Washington, DC: American Psychological Association; 2002.

La Greca AM, Vernberg EM, Silverman WK, Vogel AL, Prinstein MJ. *Helping Children Cope with Natural Disasters: A Manual for School Personnel*. Coral Gables, FL: University of Miami; 1994.

Laor N, Wolmer L, Cohen D. Mother's functioning and children's symptoms 5 years after a Scud missile attack. *Am J Psychiatry*. 2001;158:1020–1026.

Laor N, Wolmer L, Mayes L, Gershon A, Weizman R, Cohen DJ. Israeli preschool children under Scuds: a 30-month follow-up. *J Am Acad Child Adolesc Psychiatry*. 1997;36:349–356.

Leavitt L, Fox N, eds. *The Psychological Effects of War and Violence on Children*. Hillsdale, NJ:

Lawrence Erlbaum; 1993.

Nemeroff CB. Neurobiological consequences of childhood trauma. *J Clin Psychiatry*. 2004;65(Suppl 1):18–28.

*Neria Y, Gross R, Marshall R, eds. *9/11: Mental Health in the Wake of Terrorist Attacks*. Cambridge: Cambridge University Press; 2006.

Pfefferbaum B, Nixon SJ, Tucker RM, Tivis R, Moore VL. Post traumatic stress responses in bereaved children following the Oklahoma City bombing. *J Am Acad Child Adolesc Psychiatry*. 1999;39:1372–1379.

Pfefferbaum B, Seale T, McDonald N, Brandt E, Rainwater S. Posttraumatic stress two years after the Oklahoma City bombing in youths geographically distant from the explosion. *Psychiatry*. 2000;63:358–370.

Pynoos RS, Nader K. Prevention of psychiatric morbidity in children after disaster. In: Shaffer D, Philips I, Enzer NB, eds. *OSAP Prevention Monograph-2. Prevention of Mental Disorders, Alcohol and Other Drug Use in Children and Adolescents (DHHS Publication ADM 89–1646)*. Washington, DC: U.S. Government Printing Office; 1989:535.

*Rosenfeld L, Caye J, Ayalon O, Moodi L, eds. *When Their Worlds Fall Apart*. Washington, DC: NASW Press; 2005.

Schlenger WE, Caddell, JM, Ebert L, Jordan BK, Rourke KM. Psychological reactions to terrorists attacks: findings from the national study of Americans' reactions to September 11. *JAMA*. 2002;288:581–588.

Schuster MA, Stein BD, Jaycox LH, Coliins RL, Marshall GN. A national survey of stress reactions after the September 11, 2001 terrorist attacks. *N Engl J Med*. 2001;345:1507–1512.

Thabet AM, Abed Y, Vostanis P. Emotional problems in Palestinian children living in a war zone: a cross-sectional study. *Lancet*. 2002;359:1801–1804.

Wilson JP, Friedman M, Lindy J, eds. *Treating Psychological Trauma and PTSD*. New York: Guilford Press; 2001.

Yehuda R. Risk and resilience in posttraumatic stress disorder. *J Clin Psychiatry*. 2004;65(Suppl 1):29–36.

▲ 55.12 Impact on Parents of Raising a Child with Psychiatric Illness and/or Developmental Disability

ALICE R. MAO, M.D., DIANE E. TREADWELL-DEERING, M.D., MATTHEW NEIL BRAMS, M.D., AND PIETER JOOST VAN WATTUM, M.D.

INTRODUCTION

The birth of a child is generally anticipated with joy and a myriad of feelings. Parents have conscious and unconscious wishes about the child's appearance, temperament, personality, and other characteristics based on past experiences, hopes, and dreams. Expectations develop about the child's developmental course and future, including education, career, and generational continuity. Parents also have expectations for themselves in their new roles which are shaped by past experiences and other important role models. One of the important tasks of parenthood is the adaptation and modification of these fantasies and hopes to match the reality of the actual child and his or her abilities.

The task of raising a child requires fostering autonomy while providing appropriate limit setting at each developmental stage. The child has an increasing need for separation and individuation as he or she moves into adolescence and then into adulthood, which must be combined with nurturing parental structure and support. Ultimately, the goal is for the child to become an adult who is able to successfully form relationships, support himself or herself, and take on a socioculturally defined mature family and societal role. Unfortunately, if a child is not developing normally or is suffering from a psychiatric disorder that is affecting multiple domains of functioning, his or her ability to achieve these goals is compromised and caregiver burden is increased.

When a child is diagnosed with a psychiatric or a developmental disorder, parents typically undergo a complicated adjustment and grieving process. This may include anger or sadness that their child is not developing normally, coping with the narcissistic injury that the "imperfect" child is a negative reflection of their parental competency. In addition, parents may irrationally blame themselves for causing a disorder to manifest in their child. Disorders with heritable causative factors may lead to parental self-blame and/or blame from nonaffected family members. The contrast between the imagined parenting experience of raising a normally developing child and the reality of caring for a child with a psychiatric and/or developmental disorder may lead to intense disappointment and increased caregiver burden.

This section reviews the effect on parents and caregivers of raising a child with serious psychiatric disorders, specifically bipolar disorder (BPD),

attention-deficit/hyperactivity disorder (ADHD), schizophrenia, and autism/developmental disorders. Unique stressors faced by parents raising children with these disorders compared to normally developing children are summarized. Psychological responses and coping in response to these stressors and research on specific interventions developed to help parents and family members cope effectively are reviewed. Individual psychotherapeutic work with parents of psychiatrically ill children may be enhanced by understanding the life-changing consequences of dealing with a chronic mental illness or developmental disability. Importantly, the parent's ability to implement complex and multimodal treatment interventions for their child is affected by their responses to their situation. Early and effective intervention for parents after their child receives a psychiatric diagnosis may improve the long-term outcome for the child.

IMPACT ON PARENTS OF RAISING A CHILD WITH AN AUTISTIC SPECTRUM DISORDER

Introduction

In the mid-1940s, Leo Kanner and Hans Asperger independently reported on disorders in childhood that were characterized by difficulties in social interaction, presence of repetitive or stereotyped behaviors, and language abnormalities. Initial reports included observations of similarities between parents and their affected children. Asperger suggested possible genetic factors; however, prevailing theories attributed the children's symptoms to their parents' behaviors. Currently, the autistic spectrum disorders (ASDs) are clearly recognized as biologically based neurodevelopmental syndromes.

Unique Stressors

Psychiatrists caring for parents of people with ASDs will be better prepared if they are aware of the myriad of issues that may affect families. This subsection points out the unique stressors associated with raising a child diagnosed with ASD; highlights issues related to the life cycle of the family; summarizes findings from studies of parents' coping with raising a child with ASD; discusses the psychiatric disorders common in these parents—including the "broad autism phenotype" (BAP)—and suggests some approaches to therapeutic work with this population.

The parenting of any child with special needs, especially if accompanied by intellectual disability (ID), is stressful. However, parents of children with ASD cope with unique and more intense sources of stress. These include the uncertainty of diagnosis, the prolonged duration of the diagnostic process, the extreme variability in the child's developmental course, the effect of the child's atypical social communication, the child's typically normal physical appearance, and the burden on parents of

creating and organizing the child's treatment program.

Diagnosis of ASD is often delayed when primary care providers do not screen for or recognize signs of developmental disorders. Referrals to subspecialists may be time consuming and lack definitive results. Siklos and Kerns reported that parents saw an average of 4.5 professionals and waited almost 3 years to obtain a diagnosis. More than 50 percent of parents were dissatisfied with the diagnostic process. Mandell and Palmer found varying ASD prevalence rates when comparing states based on their practicing number of pediatricians and education-related spending. This is further evidence of continued difficulty in obtaining a timely diagnosis and accessing care. Suboptimal disciplining techniques used prior to the correct diagnosis result in caregiver loss of confidence and self-efficacy.

The spectrum of autistic disorders is very broad; therefore, predictions about a specific child's progress are imprecise. Irregular patterns of skill development are common. Developing appropriate expectations for a child and predicting future development is challenging for most clinicians because of wide variations in developmental progression.

Abnormalities in social communication appear to be a particular source of distress for families. Safronoff and Farbotko in 2002, found autistic children's difficulties with verbal and nonverbal communication result in parents feeling unable to meet their child's wants or needs. The inability to understand their child and his or her behavior is significantly more frequently reported by mothers of autistic children than control mothers.

Lack of consistent behavior across time, people, and environment is frustrating and confusing for parents. Children with more obvious impairments, such as those affected with Down syndrome, are excused and tolerated by others, but the typically normal physical appearance of a child with autism leads others to expect appropriate behavior. Fear of their child's unpredictable, disruptive, and unusual behavior, as well as others' responses to it, leads parents to avoid public situations, which in turn leads to restriction of family activities, social isolation, and increased burden in completing everyday household chores, such as grocery shopping.

Typical recommended interventions for children with ASD are obtained from several different systems—public education, ancillary health care therapists, and hospital-based, center-based, and community private practitioners. The family is typically left in charge of coordinating a complicated network of professional providers, who have differing opinions, recommendations, physical locations, and financial/insurance requirements. The lack of curative interventions from traditional care providers, the lack of empirically supported therapies, the nature of ASD symptoms, and the chronicity of these disorders have led to the emergence of unconventional treatments by alternative practitioners with poorly supported claims of cure or rapid improvement. The family is confronted with the challenge of obtaining intensive treatment for their child during a

critical time frame with little professional guidance.

Lifespan Issues

Each developmental stage presents the family with unique and specific issues. Early childhood is characterized by the family's realization of an abnormal developmental trajectory, the diagnostic process, and issues of definitive diagnosis. Parental guilt, mourning the "lost," normally developing child, the responses of extended family members, and the possible additional diagnosis of ID are all issues that color this period. The child's irregular sleep patterns, hyperactive behavior, lack of effective communication, and peculiar eating and elimination habits, as well as safety issues and the lack of ability to obtain appropriate day care or child care, are practical concerns facing most parents. Accepting the diagnosis, finding community services, and planning for the future are tasks for the family. Families report that education about the disorder and information about resources were the most helpful interventions they received. Other useful resources were learning and behavior assistance, speech therapy, and family support.

With a good school program, middle childhood may be a relatively easier life phase. Behavior problems often improve; families have adjusted their expectations of the child and have accepted the delays in development, abnormal social interaction, and communication. Developing an educational program that includes the necessary special education services, offered in the setting that is most successful for the child's learning, is the goal of collaboration with the school system. The focus may be on the development of adaptive behaviors in the areas of self-care, domestic chores, and social behaviors. Parents may grieve anew as they recognize the child's inability to participate in team sports, engage in birthday parties, and other social activities that they see normally developing children doing. Sibling issues, which may occur at any time, are common. The affected child behaves oddly, is not a typical playmate, and requires much of the parents' time and energy. Parents often overestimate siblings' abilities to understand the implications of autism and underestimate siblings' worries.

By adolescence, almost all families report worries about their child's future. Public education services end, and adult services may be uncertain. Aggression, increased sexuality, and increased sense of independence without the skills for self-management are among behavioral concerns. Higher-functioning children may become aware of their differences from peers. Nonaffected siblings may be leaving home, creating a higher burden of care for parents who have relied on their help. The aging of grandparents may also lead to less support and greater family strain.

The most pressing issues for adulthood are appropriate living arrangements and vocational opportunities. The move toward

deinstitutionalization has eliminated the option of residential placement for most families, and many adults with autism continue to live with their families. There are limited opportunities for appropriate social experiences and employment. Young adults with comorbid ID may find services through this diagnosis. Individuals with higher cognitive skills often continue to lack job-related skills and remain un- or underemployed. Seltzer et al. compared children with a developmental disability versus typically developing children and found 57 percent of parents aged 50 or older live with their child with developmental disorder versus 16 percent of normally developed children. Guardianship and long-term estate planning are worrisome issues for parents. The divorce rate has been found to be higher in parents of autistic children when compared with parents raising normally developed children (25 percent vs. 13 percent). Distinct from typically developing children, where the divorce rate is high until approximately age of 8, the likelihood of divorce remains high through early adulthood for parents of children with ASD. The financial strain for therapeutic interventions cannot be overlooked; Lord and Bishop found the economic impact is 3 to 5 million dollars more than for a typical developing child.

Psychological Responses to Raising a Child with an ASD

Stress. Many published studies have documented that parents of children with ASD report more stress than parents of children with other developmental disorders or of typical children. A variety of factors have been found to cause and/or mediate this, the most important of which is gender. The preponderance of studies examining gender differences in parental response soundly endorse the observation that mothers report significantly more stress than fathers in general and within given family units. Explanations for this finding include that mothers shoulder the greater burden of child care, even if they work outside the home; perceived parenting competence contributes more to mothers' self-esteem than to that of fathers; and fathers use more effective coping strategies than mothers.

Most studies support a positive correlation between the amount of stress mothers report and the behavioral problems of the child. These behaviors include oppositionality, noncompliance, aggression, irritability, and unpredictability. None of these are core features of ASD, but they often accompany these disorders. Neither the severity of the ASD per se nor the level of adaptive functioning was found to correlate with reported stress or poor maternal outcome. However, these were not universal findings, and other authors have found that autism severity or the amount of services the child received, which was interpreted as a measure of autism severity, did correlate with levels of maternal stress.

Several studies have aimed at understanding factors other than the

child's autism that affected the well-being of individual family members. Factors such as the partner's depression and anxiety were positive predictors of stress for both mothers and fathers. Attributes of the affected child were predictors for maternal but not paternal stress.

Factors that have been found to decrease stress for parents include positive coping styles, belief in the efficacy of treatment, and social support. Parent education and the acquisition of behavioral management skills have also been found to be beneficial. Greater marital quality is predictive of lower levels of stress.

Coping. Coping is defined as "a process by which individuals respond to threats of stress." These strategies are generally divided into two categories: (1) emotion-focused strategies, which aim to reduce or mask feelings of distress, including engaging in denial, hoping for miracles, having fantasies, using food or drugs, and distancing and avoiding; and (2) problem-focused strategies, which aim to solve or change the problem, including confrontive coping, fighting for what is wanted, expressing anger and other feelings, taking chances, and engaging in positive reappraisal. Several studies have examined the consequences of particular coping styles for parents of ASD children. Emotion-focused styles have been found to correlate with increased depression, isolation, and marital problems. Pottie found that escape, blaming, withdrawal, and helplessness lead to decreased positive mood, whereas problem focused, social support, positive reframing, emotional regulation, and compromise coping resulted in a positive mood. Conversely, problem-focused coping was associated with lower levels of depression. Moreover, coping style more robustly predicted levels of depression than locus of control and social support, although these factors were also contributory.

Research on coping over time suggests that, in general, young adults are more likely to use problem-focused strategies, and middle-aged adults use more emotion-focused methods. Gray found evidence for this in his longitudinal study of coping with children with ASD. Religion and faith became much greater in relative importance as a coping strategy; attachment to the child and greater appreciation for his or her good qualities also emerged as more common strategies. In general, the use of an array of coping strategies is beneficial, although Gray found that parents used fewer strategies over time. Studies show mixed findings regarding increased effectiveness of coping as parents age. Fitzgerald reported that advanced age of affected ASD children was associated with better maternal mental health, implying adaptation on the part of the mothers. However, Smith et al. found high levels of depression in parents of both toddlers and adolescents.

A rare focus on fathers revealed that fathers of children with ASD or Down syndrome reported more use of wish-fulfilling fantasy and

information seeking than control fathers. More effective paternal coping has been hypothesized as the explanation for decreased levels of stress reported by fathers compared to mothers.

Studies show that fathers use avoidance and engagement in outside activities more frequently as coping strategies than mothers.

Tunali and Powers proposed a model of cognitive strategies that would facilitate coping for mothers of children with autism. These strategies included greater emphasis on being a "good mother" rather than on career development; greater frequency of recreation with family members rather than individual recreation; less emphasis on individual leisure; less emphasis on others' opinions of child-rearing practices; greater emphasis on parenting and partner support than on intimacy and shared leisure in marital relationships; and higher tolerance of ambiguity.

Quality of Life. Several authors have used the conceptual framework of "quality of life" (QOL) to understand the impact of rearing a child with an ASD. Defined by the World Health Organization as "individuals' perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns," QOL is a multidimensional concept that includes physical and mental health, relationships, and personal beliefs and values in relationship to the environment. On the World Health Organization Quality of Life-BREF, parents of children with pervasive development disorder (PDD) were compared to parents of children with mental retardation (MR), cerebral palsy (CP), and typical development. Parents of children with PDD showed a significant impairment in QOL as compared to the other groups, whereas few differences were noted between the MR, CP, and typically developing groups. Within the PDD group, mothers had lower QOL than fathers. These differences were posited to be caused by the stress of raising children with severe developmental disorders, the presence of the BAP in the parents or other family members, and parental psychiatric disorders. Analysis of data from the 2003 National Survey of Children's Health (NSCH) revealed significantly poorer QOL in families of children with ASD when compared to families of children with ADHD or typical controls, most notably in the areas of burden of care and isolation. In a population-based study of psychological functioning and coping, mothers of children with ASD rated their overall health as comparable to mothers of nonaffected children, but they reported poorer mental health. In a study of health-related QOL in parents of children with high-functioning autism (HFA) or Asperger syndrome, mothers had lower physical health scores than fathers of affected children and control mothers.

Psychiatric Disorders and the Broad Autism Phenotype

As data were gathered supporting the role of genetics in autism, the idea

gained momentum that the genetic liability for ASD may also be expressed in nonautistic relatives of autistic probands with qualitatively similar but milder behavioral and cognitive features of autism. The concept of BAP is supported by co-twin studies and family studies of siblings and parents. Subsequent work has refined the description of this phenotype, which includes social impairments, communication impairments, and rigidity. Social difficulties include lack of interest in social interaction and lack of close, confiding relationships. In one study, 37 percent of parents of children with ASD reported no friendships, compared with only 4 percent of parents with Down syndrome controls. Communication problems are pragmatic, referring to deficits in the social aspects of language. Clear-cut obsessional and repetitive behaviors occur infrequently; however, rigidity, characterized by resistance to and difficulty in accepting change, is more common. These differences in language, anxiety, rigidity, and communication may cause parents to feel inadequate to help in the behavioral treatment of their children who experience similar impairments. Bolton reported that the BAP was seen in 20 percent of siblings of those with autism versus 3 percent of siblings in the control group. Similar rates for parents have been reported. Bishop et al. reported that 10 of 65 (15.4 percent) of mothers and 17 of 46 (40 percent) of fathers were found to have BAP. Ingersoll et al. (2011) found parental BAP was associated with less social support, increased stress and higher rates of depression. Ingrams et al. found expressing BAP used maladaptive coping strategies.

Studies that focused on the stress of raising a child with autism have found increased rates of several psychiatric disorders in parents of affected children, most robustly major depressive disorder (MDD) and social phobia. Piven and Palmer reported major depression in 33 percent and social phobia in 14.6 percent of proband parents in comparison to 11.7 and 3.3 percent, respectively, in Down syndrome parent controls. Smalley et al. found a lifetime prevalence of major mood disorders in 40.3 percent of parents and 32.4 percent of siblings of children with autism, compared to 19.2 percent of parents with children with tuberous sclerosis or seizures. With regard to major depression, the increased risk does not appear to be due to raising a child with ASD, given that the majority of depressed parents had their first episode prior to the birth of their affected children. Moreover, these psychiatric disorders have not been found to be associated with the presence of BAP. Potential explanatory models for these findings have involved overlap of the risk alleles that are common to both autism and recurrent mood or anxiety disorders or a result of assortative mating of parents with BAP, with spouses with major depression, or social phobia. Other anxiety disorders, mood disorders, and substance use disorders have also been reported to occur more frequently in parents of children with ASD in comparison to control groups of general population parents or

parents of children with other developmental disorders.

Summary

The task of parenting a child with an ASD is a stressful, lifelong process that significantly affects all of those involved. Those caring for such parents need to be aware of the nature and magnitude of this undertaking. Encouragement to become active in the child's therapies, join available support services, and use positive coping strategies is warranted. Monitoring for the development of depression and anxiety disorders and conducting an evaluation for the broader autism phenotype are also indicated. Although specific data are not available for ASD, cultural factors color parents' beliefs about their children's mental health conditions and need to be considered by mental health practitioners who treat the parents.

IMPACT ON PARENTS OF RAISING A CHILD WITH BIPOLAR DISORDER

Introduction

Clinicians and researchers have become increasingly aware of the tremendous burden that is placed on parents and caregivers who are raising a child with BPD. Previous theories implying that the dysfunctional family system is causing the illness in the child, have been replaced by the realization that the child's illness is negatively impacting the family. There is growing evidence to support interventions for family members and caregivers of patients with BPD because decreasing stress may improve long-term outcomes for the family and the child.

Unique Stressors

Research has shown that children, adolescents, and adults who have BPD have severe global impairments in multiple areas of functioning. They are at higher risk for academic failure, poor interpersonal relationships, substance abuse, psychosocial and occupational difficulties, legal problems, suicide attempts, and neuropsychologic deficits. In many cases, the parents of children with BPD may also be struggling with their own mood disorders. The child's unpredictable course may create additional stressors that exacerbate parental illness, leading to impaired parenting. Emotional dysregulation in parent and child may increase conflict and compound family stress. Not only is an episode of illness stressful for each of the family members, but also the fear of relapse is present even when the illness is stabilized.

The diagnosis of pediatric-onset BPD is imprecise. However, there is agreement that the presentation of BPD in children and adolescents differs substantially from the presentation in adults. In contrast to the discrete episodes of manic or depressive symptoms observed in adults, children usually exhibit ultradian cycling, irritability, explosive temper outbursts,

and decreased frequency of interepisode recovery. Due to the unpredictable nature of the illness presentation, accompanying affective instability, and often violent or aggressive behaviors, the caregiver role can be demanding, distressing, and harmful to health and QOL. Frequently, the diagnosis of BPD is delayed because the child's behavior is initially labeled as oppositional or defiant; therefore, appropriate pharmacotherapy and psychotherapy interventions are also delayed.

Lifespan Issues

Parents of children who begin exhibiting symptoms of BPD are generally bewildered by the child's mood swings, disruptive behaviors, and general inability to conform to their same-aged peers. Most often the children appear developmentally normal and have normal intellectual functioning, and so they do not come to the attention of school staff or teachers until they exhibit disruptive behaviors or receive failing grades. Practitioners are often reluctant to diagnose a child with BPD. These children have been described as "diagnostically homeless" because they often exhibit symptoms that can fit the criteria for multiple disorders, such as ADHD, mood disorder not otherwise specified, conduct disorder, and intermittent explosive disorder. When the diagnosis of BPD is finally made, parents have described experiencing an array of emotions ranging from relief to depression.

Adolescents with BPD are even more challenging to parents because their desire for autonomy, coupled with impaired judgment, can lead to harmful consequences. Parents struggle to set limits with adolescents exhibiting promiscuous behavior, dangerous driving habits, substance abuse, and oppositional behaviors. Frequently, the adolescents are not compliant with pharmacotherapy because the medications have side effects such as weight gain or sedation. Some adolescents perceive that the medications alter their personality, making them more subdued and socially inhibited. They frequently take little responsibility for the consequences of their actions and tend to perceive their parents as being overly controlling rather than appropriately protective. Parents struggle with guilt, self-blame, and feelings of helplessness because they can no longer physically restrain the adolescents if they choose to leave the home or refuse to take their medications.

Parents of adults with BPD face the greatest obstacles because they have limited legal recourse to protect or help their children if treatment is refused. Because adults have euthymic episodes with longer periods of interepisode recovery, they may cycle into depression or mania without parents being aware of the polarity switch. Consequently, reinitiating treatment may be delayed. During acute episodes, adults with BPD frequently alienate parents, spouses, or other caregivers with their aggressive, abusive, or impulsive behavior. Parents may be overly taxed by

the responsibility of emotionally and often financially supporting their adult child.

Caregiver Burden

The effects on caregivers of living with an ill relative have been defined as family burden. The burden of caring for a psychiatrically ill family member has been divided into two categories:

1. *Objective burden* is any disruption to family or household life that is verifiable and observable. Examples of objective burden include high rates of separation and divorce in marriages with a psychiatrically ill spouse; restricted social and leisure activities, often related to fears of stigma; and financial difficulties.
2. *Subjective burden* involves the personal feelings of carrying such a burden, such as feeling distressed, unhappy, or upset. The types of subjective burden experienced include guilt if the relative feels responsible for contributing to the illness; rejection toward the patient; anger that the illness has spoilt the relative's life; and grief and a sense of loss contrasting the present with memories of the person before the illness occurred and the diagnosis was made. Subjective burden experienced by relatives has been linked to the development of affective symptoms.

Review of Research on Caregiver Burden for Bipolar Disorder. There are limited data on degree of stress experienced by caregivers of children with pediatric BPD. Some information about the consequences of caring for a family member with BPD can be obtained from studies of caregivers of adult bipolar patients. Factors affecting family burden experienced by caregivers of adult patients with BPD include prolonged illness and high levels of dysfunction among patients, as well as caregiver illness beliefs and coping style.

The Family Experience Study, a longitudinal study of the primary caregivers for patients with BPD, examined the cost and consequence of caring for a relative or friend with BPD. Primary caregivers were evaluated within 1 month after patients entered the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD), using measures of burden, coping, physical health, mental health, resource utilization, and cost. At study conclusion, more than half of the caregivers experienced moderate or high-burden levels, which were related to patient problem behaviors, interference with caregivers' ability to fulfill other role obligations, or disruption of household routine. Caregivers experiencing higher-burden levels reported more physical health problems, depressive symptoms, behavior associated with greater health risk such as smoking and drinking, health service use, and more limited social support than less

burdened caregivers. They also reported having to provide more financial support to their bipolar relative.

Most adult patients with BPD spend more time in periods of remission or with subthreshold symptoms than in an acute episode; however, most studies of caregiver stress are completed when patients are acutely ill, which may overestimate stress. In one study, investigators focused on stress in caregivers of euthymic adult bipolar patients aged 18 to 65, during periods of subthreshold symptomatology. Caregivers showed a moderate level of subjective burden on the Social Behavior Assessment Schedule. The variable most highly related to caregiver distress was found to be the patient's behavior (hyperactivity, irritability, sadness, and withdrawal). Parents were most worried about their children's abilities to perform at work, in school and in social relationships.

Subjective and Objective Burden. Factors that increase caregiver's subjective burden are poorer patient social and occupational functioning, a relapse during the previous 2 years, and history of rapid cycling. Rapid cycling is defined as having four or more episodes of mania or depression per year and is found in a significant fraction of patients with BPD. It has been associated with greater disability and worse prognosis as evidenced by higher use of intensive psychiatric interventions.

Risk factors that have been associated with increased objective burden levels include caregiver responsibility for medication administration, violence from the patient, caregiver social isolation, rejection by the patient during periods of illness, and patient irritability leading to more frequent arguments and conflicts. Some studies have suggested that the effect of clinical variables can be mediated by caregiver coping styles and their level of understanding of the patient's illness and symptoms.

Raising a child with BPD can be complicated if the parent is also suffering from a mood disorder. Research has shown that the offspring of parents with BPD have increased risk for inheriting mood disorders. A meta-analysis of 17 studies of offspring of parents with BPD found that 52 percent of offspring had challenges with the interpersonal, cognitive, and behavioral manifestations of BPD are even less capable of coping with the burden of raising a child with BPD.

Interventions

Family interventions for BPD are generally psychoeducational, with the goals of teaching family members the signs and symptoms of BPD, developing strategies for intervening early with new episodes, and providing consistent medication regimens. The psychoeducational approach recognizes that lack of information about the disorder, compounded by uncertainties about the future, can increase caregivers' stress. Thus, in addition to providing information, clinicians address the

patients' and family members' affective reactions to the illness, its prognosis, and expected treatments and assist them in developing coping strategies relevant to their individual situations.

Although there are some randomized, controlled trials that highlight the positive benefit for the patient of family-focused treatment in addition to pharmacotherapy in BPD, few trials using contemporary methodologies have analyzed the specific effects of caregiver-focused interventions. A randomized, control trial published in 2008, studied the impact that interventions for caregivers had on relapse rates in the euthymic bipolar patients they cared for. Caregivers in the experimental group were provided twelve 90-minute sessions that focused on knowledge of BPD and training in coping skills. The patients did not attend the groups, but were assessed monthly during both the intervention and at the 12-month follow-up. The patients whose caregivers attended the intervention group experienced fewer recurrences than those in the control group. In addition, when type of episode was analyzed separately, significant differences were found in the number of subjects who experienced a hypomanic/manic relapse and the time to such an episode. The differences in depressive and mixed episodes did not reach statistical significance.

Relapse prevention may improve by facilitating early identification of the first signs of an episode, leading to earlier intervention. Additional benefits include dispelling caregivers' misconceptions about the illness and issues such as caregiver criticism or overinvolvement. Some studies have shown that family psychoeducation also increases a mutual understanding with the patient. It also increases positive interactional behaviors and reduced expressed emotion. High expressed emotion with negative affective style and family burden have been related to poor illness outcome.

A 2002 pilot study of manual-driven, multifamily psychoeducation group (MFPG) treatment for families of children with mood disorders represents a rare study focused on pediatric subjects. It addressed two questions: Do families of children with BPD and families of children with MDD or dysthymic disorder (DD) differ at treatment entry, and do they benefit equally from intervention? Conclusions were that BPD families entered treatment with more impaired children and more extensive treatment histories than families of children with MDD or DD; however, all families benefited from intervention. A qualitative analysis revealed that families in the study gained emotional support, improved coping skills, and decision making as it related to their child's illness, shifted their attitudes with regard to the illness, and experienced improved interaction with their affected children. Clinical impressions suggested that interventions combining families of children with BPD and depressive spectrum illnesses were clinically feasible and potentially beneficial. The authors suggested that intervention modalities geared toward families of children with BPD may also include families of children with other mood disorders, given that

both groups may benefit from these types of interventions.

Conclusions

Although most treatment research has focused on pharmacotherapy for pediatric BPDs, a growing literature suggests that psychosocial interventions are also important to provide families with an understanding of symptoms, course, and treatment of BPDs. Interventions to teach youth and parents methods for coping with symptoms (e.g., problem solving, communication, emotion regulation, cognitive-behavioral skills) and prevent relapse have been shown to be helpful to parents. Identifying and addressing stressful aspects of raising an individual afflicted with BPD could reduce the emotional burden and potential negative impact on the caregiver's mental health and QOL. Gaining support from other parents who are raising children with BPD may reduce feelings of guilt, self-blame, and inappropriate feelings of responsibility for their child's condition. Moreover, providing support for the parents may ultimately improve the child's outcome. With support and education, parents are better able to implement multimodal treatment interventions needed to improve adaptive and functional outcomes for their child. In cases of multigenerational BPD, treatment of the parent's mood disorder can also have a positive effect on the child's symptoms, which can bolster parental perceptions of self-competence. In conclusion, psychosocial interventions that involve families, psychoeducation, and skill building may offer added benefit to pharmacotherapy interventions.

IMPACT ON PARENTS OF RAISING A CHILD WITH ATTENTION-DEFICIT/HYPERACTIVITY DISORDER

Introduction

Parenting a child ADHD has challenges similar to those faced by parents of children with other developmental and/or behavioral disorders, including increased parental stress, marital discord, and maternal depression. Advocacy issues and long-term care concerns, although present, are generally less urgent. Nonetheless, rearing children with ADHD involves separate and unique aspects of care that are influenced by both child-derived and parent-derived characteristics. Complicating this scenario is the likely interaction among these characteristics.

Effect of Child Characteristics on Parenting Stress

Several authors report increased parenting stress in mothers with children of ADHD. Child characteristics of ADHD that specifically correlate with parental stress are the severity of ADHD symptoms, aggression, oppositional behavior, and, to a lesser degree, the presence or absence of medical illness in the child. This highlights the effect of comorbid and

associated symptoms on the level of stress within the ADHD family.

Unlike most children with more severe disabilities, children with ADHD, especially of a more moderate level of severity, are expected to master the developmental challenges of typical childhood, adolescence, and adulthood. From kindergarten through middle school, they are expected to relate to peers, join extracurricular peer group activities, participate in sports, and succeed in mainstream education. In adolescence, they drive cars, date, work, and attend school. They are expected to enter early adulthood with the typical goals of learning a trade, going to college, obtaining employment, and establishing long-term relationships and friendships. However, due to the core symptoms of inattention, impulsivity, and/or hyperactivity in these children, these goals are not always readily met. Studies show that children with ADHD, especially those with a hyperactive/impulsive component, are at much higher risk for accidental injuries and trauma, emergency room visits, motor vehicle accidents, grade retention and high school dropout, drug use, unplanned pregnancies, and sexually transmitted diseases. In addition, more than half of these children were fired from one or more jobs by age 21 years. These consequences of ADHD create stress within the family and can lead to a higher incidence of parental divorce or marital discord and child abuse.

Effect of Parent Characteristics on Parenting Stress

Parenting stress is also correlated with specific attributes of the parents who are rearing children with developmental disabilities. For example, maternal depression and the perception of one's share of caretaking burden have been linked to increased stress in parents raising children with autism. Similarly, in Anastopoulos' study of mothers of children with ADHD, parenting stress, as measured by the Parenting Stress Index (PSI), correlated highly with maternal psychopathology (e.g., depression, BPD) and lack of employment outside of the home. Seipp et al. reported that during observed mother–son interactions with boys with ADHD, mothers with greater overt hostility had less positive parenting behavior, which was associated with greater oppositionality in the son. Moreover, mothers with sons with ADHD and comorbid oppositional defiant disorder (ODD) were less responsive to their child's needs and more overreactive than control mothers. This finding may be more directly related to specific maternal or child characteristics or the interplay between the two.

Rearing a child with ADHD is greatly complicated if one or both parents share this diagnosis with their child and struggle with similar cognitive and behavioral manifestations of the illness. It is estimated that 40 to 60 percent of adults with ADHD are rearing an affected child. Because the prevalence of ADHD in adults is estimated at around 4.5 percent, there are millions of parents with ADHD rearing children with ADHD. This situation creates a multitude of additional management issues, increasing the

family's stress level. The parent without ADHD may be unfairly burdened with responsibility for managing the child with ADHD. Parental ADHD, especially if untreated or poorly managed, may contribute to parental neglect, inadequate supervision, inconsistent and inappropriate discipline, and increased parent–child conflict. Furthermore, maternal depression and early maternal negative parenting techniques are predictors of future conduct problems in children with ADHD. Conversely, early positive parenting techniques, including positive feedback, praise, and optimistic affect, are associated with fewer conduct issues.

Mrs. M, a 34-year-old homemaker and mother of two, had been married to a pharmaceutical sales representative for 10 years. Both her 8-year-old son and her husband were diagnosed with ADHD, although only her son was medicated. He continued to have significant impulsive aggression directed toward his younger sibling, serious academic struggles, and poor peer relationships secondary to "immature behavior." Her husband had poor job performance, with deficient organizational skills and failure to attend important meetings due to "forgetfulness." He was reluctant to take medications because of fears of side effects and "drug dependence." Mrs. M's family of high achievers compared her unfavorably to her more accomplished sister and her family. Moreover, her family had a history of depression, and Mrs. M presented with several neurovegetative signs, but worries of stigma interfered with her pursuit of treatment. She and her husband had discussed divorce. She had no sources of emotional support and felt overwhelmed.

Clinical Implications

The assessment of patients who are parents needs to include consideration of the patient's ability to meet the demands of parenthood, especially in the presence of a childhood psychiatric disorder. The consequences of rearing a child with ADHD should not be underestimated, especially given the frequency with which the disorder also affects one of the child's parents. Proper attention to child-based and parent-based characteristics associated with ADHD, as well as to the interplay of these factors, can lead to decreased stress for the parents and improved functioning for the child and for the family as a whole.

IMPACT ON FAMILIES OF RAISING A CHILD DIAGNOSED WITH SCHIZOPHRENIA

Introduction

The National Institute of Mental Health estimated that in 2002, 1.1 percent of the population suffered from schizophrenia. The estimated yearly cost of schizophrenia in the United States was 62.7 billion dollars in 2002. Globally, the cost is estimated to be between 1.5 percent and 3 percent of national healthcare budgets. This subsection discusses the burden borne by parents and siblings of a child with schizophrenia. Whereas the other disorders discussed in this section are usually observed earlier in the child's life, schizophrenia typically develops between the mid-teens to early 30s, with rare cases of earlier onset after a significant period during which the

child appeared to be on track to become a healthy, independent adult.

Early psychodynamic theories of schizophrenia focused on the role of mothers of patients with schizophrenia. They were often called "schizogenic" or "schizophrenogenic," and their lack of maternal warmth was, without any sound scientific evidence, blamed for their child's development of schizophrenia. Only in the mid-1960s was this concept countered, and since then schizophrenia research has been approached from a biopsychosocial perspective based on the notion that expression of the disorder is due to a genetic predisposition combined with exposure to environmental and individual stressors. This puts the role of the parent and family in a different light.

Caregiver Burden

To better assess the issues with which families of schizophrenic patients grapple, Hoenig and Hamilton introduced the concept of "family burden," which was defined earlier in this section. The Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) study found four factors that caregivers identified as most burdensome in their experience with their ill family member: (1) patient problem behavior, (2) the patient's impairment in activities of daily living, (3) lack of patient helpfulness, and (4) resource demands and disruptions in the caregivers' personal routine. Other factors related to level of burden were social and professional support, age, and education level. Family members who were able to share their experience with others and who received financial support experienced a lower level of burden; younger and less educated caregivers experienced more burden than older and more educated caregivers.

Parents and siblings of patients with schizophrenia often grieve the loss of the child or sibling they knew before the onset of illness. They may also experience feelings of embarrassment and stigma related to their family member's illness, which, in turn, may lead to isolation. Aging parents frequently worry about what will happen to their children when they can no longer care for them. In a study of 157 mothers of children with serious mental illness, only 11 percent reported definite plans for their offspring's future residence; 66 percent had completed financial plans; and 75 percent hoped a family member would assume care, but only 25 percent thought that this would actually occur. Older parents need professional help with planning for the future, in particular estate planning, to enable case management and ongoing services after the parent's death.

Siblings play a unique role in the life of the patient with schizophrenia and may have specific needs. They have additional levels of burden, such as guilt about being healthy, fear of developing schizophrenia, and fear of passing on "bad genes" to their offspring. Several coping mechanisms have been identified in siblings, including normalization, caregiving, and grief. Smith et al. identified several factors that predict whether siblings will be

involved in the future care of the ill sibling. Pull factors, which predict involvement, are early experiences in socialization, strong sibling relationship, and personal caregiver gains, such as developing personal strengths while caring for the sibling. Push factors, which predict disengagement, are geographic distance and a belief that the ill sibling has control of his or her symptoms.

Cultural Aspects of Burden

Across cultures, families of patients with schizophrenia often suffer from stigma that accompanies the diagnosis of a psychotic disorder. Several studies have indicated that black caregivers, on average, experience less distress than Hispanic and white caregivers: a seemingly surprising finding in light of the greater likelihood of additional stressful conditions such as prejudice, discrimination, and lower income that African Americans face. In addition, African American family members tend to be more involved in caregiving activities, which are also associated with higher levels of distress. Potential protective factors, such as religious involvement and social support, do not appear to differ according to race and ethnicity. Studies indicate that QOL for caregivers who live with a chronically ill family member is lower than those living alone or with healthy family members and the financial burden is significant. QOL for siblings is better than for parents and children of schizophrenic patients but, across cultures, women tend to experience more distress than men because they tend to provide more care for ill children or siblings.

Treatment and Family Burden

A multidimensional approach is required to assess and understand the burden experienced by families of schizophrenic patients. The Schizophrenia Patient Outcomes Research Team (PORT) recommended that, in addition to pharmacotherapy, patients with ongoing family contact should be offered family psychosocial intervention lasting at least 9 months, emphasizing education about the illness, family support, crisis intervention, and problem-solving skills training. This may relieve the burden experienced by the family, who in turn is better able to take care of the ill relative. Several support models exist, including single- and multiple-family group models. Although models vary, effective programs share the following characteristics and assumptions:

1. Schizophrenia is an illness.
2. Program leadership is provided by mental health professionals.
3. Participating family members are not seen as patients.
4. Family therapy is not included, to avoid the suggestion that the family is to blame for the relative's illness.
5. The focus remains on patient outcome, including prevention of relapse

and improvement of functioning and well-being of the patient.

Families participating in such programs can reduce relapse rates by up to 50 percent, which is comparable to pharmacotherapeutic effect sizes. However, according to PORT, participation in such programs is limited. Although 75 percent of patients had contact with family members, only 31 percent of families had received information, advice, or treatment. Barriers to access included patient factors (concerns about family burden and confidentiality), family factors (lack of awareness regarding benefits, time constraints, reluctance to become more involved), and provider factors (staff resources, time and training, reimbursement issues, lack of awareness of benefits). Community-based programs, such as the National Alliance for the Mentally Ill (NAMI), Family-to-Family, and Journey of Hope programs, are run by volunteers from patients' families and focus on family needs, advocacy, and system reform. Participation in these programs leads to improved overall well-being for family members, greater facility in navigating the mental health care system, and increased understanding of schizophrenia. Shorter duration and lower costs may make them more accessible to family members.

Conclusions

Schizophrenia is a chronic illness that often causes an enormous burden for the patient's family. However, during the last decade, research has increasingly shown that support for patients and their families through psychoeducation and support programs can have a significant effect on the perceived burden by both families and the patient. It can lead to improved outcome for the patient, with fewer episodes of relapse and better QOL, whereas families are more able to understand the diagnosis and its consequences for their child or sibling and to take on the role of caregiver. In addition, these programs reduce the financial burden of schizophrenia on the health care system. However, many factors hamper implementation of these programs; it is the role of the treatment provider to offer the best possible and evidence-based care to patients and their families.

SUMMARY

This section has described the consequences of having a psychiatrically ill child on his or her parent by considering four diverse disorders that usually first appear in childhood or young adulthood. Although not all inclusive, these disorders were chosen to illustrate themes generally experienced by any family so affected. Common to all of these disorders are the grief and disappointment that parents experience when they realize that their child is not developing typically. Self-blame is frequent, particularly in heritable disorders. Factors that shape parental response include age of onset of the disorder, chronicity of the disorder, the extent to which the disorder affects

acquisition of basic developmental milestones and functioning, the presence of aggression, and the stigma associated with the specific disorder. Presence of the disorder or an associated condition in a parent, as well as in the child, greatly affects the outcome for all members of the family. The lack of appropriate services, especially when the youngster reaches adulthood, is a prevalent dilemma for parents. Psychiatric practitioners need to consider the profound effect that raising a child with a serious psychiatric disorder has on a parent who is their patient.

REFERENCES

Allik H, Larsson JO, Smedje H. Health-related quality of life in parents of school-age children with Asperger syndrome or high-functioning autism. *Health Qual Life Outcomes*. 2006;4:1.

Chronis AM, Lahey BB, Pelham WE Jr, et al. Maternal depression and early positive parenting predict future conduct problems in young children with attention-deficit/hyperactivity disorder. *Dev Psychol*. 2007;43:70–82.

Cohen IL, Tsiouris JA. Maternal recurrent mood disorders and high-functioning autism. *J Autism Dev Disord*. 2006;36:1077–1088.

Fristad MA, MacPherson HA. Evidence-based psychosocial treatments for child and adolescent bipolar spectrum disorders. *J Clin Child Adolesc Psychol*. 2014;43(3):339–355.

Gerdes AC, Hoza B, Arnold LE, et al. Maternal depressive symptomatology and parenting behavior: exploration of possible mediators. *J Abnorm Child Psychol*. 2007;35:705–714.

Giallo R, Gavidia-Payne S. Child, parent and family factors as predictors of adjustment for siblings of children with a disability. *J Intellect Disabil Res*. 2006;50:937–948.

*Gray DE. Coping over time: the parents of children with autism. *J Intellect Disabil Res*. 2006;50:970–976.

Hastings RP, Kovshoff H, Ward NJ, degli Espinosa F, Brown T, Remington B. Systems analysis of stress and positive perceptions in mothers and fathers of pre-school children with autism. *J Autism Dev Disord*. 2005;35:635–644.

Herring S, Gray K, Taffe J, Tonge B, Sweeney D, Einfeld S. Behaviour and emotional problems in toddlers with pervasive developmental disorders and developmental delay: associations with parental mental health and family functioning. *J Intellect Disabil Res*. 2006;50:874–882.

Kersh J, Hedvat TT, Hauser-Cram P, Warfield ME. The contribution of marital quality to the well-being of parents of children with developmental disabilities. *J Intellect Disabil Res*. 2006;50:883–893.

Kowatch RA, DelBello MP. Pediatric bipolar disorder: emerging diagnostic and treatment approaches. *Child Adolesc Psychiatr Clin N Am*. 2006;15:73–108.

*Lecavalier L, Leone S, Wiltz J. The impact of behaviour problems on caregiver stress in young people with autism spectrum disorders. *J Intellect Disabil Res*. 2006;50:172–183.

Mandell DS, Novak M. The role of culture in families' treatment decisions for children with autism spectrum disorders. *Ment Retard Dev Disabil Res Rev*. 2005;11:110–115.

Mandell DS, Palmer R. Differences among states in the identification of autistic spectrum disorders. *Arch Pediatr Adolesc Med*. 2005;159:266–269.

*Marcus L, Kunce LJ, Schopler E. Working with families. In: Volkmar FR, Paul R, Klin A, Cohen

D, eds: *Handbook of Autism and Pervasive Developmental Disorders*. 3rd ed. Vol. 2. Hoboken, NJ: Wiley; 2005:1055–1086.

Miklowitz DJ. The role of the family in the course and treatment of bipolar disorder. *Curr Dir Psychol Sci*. 2007;16:192–196.

Montes G, Halterman JS. Psychological functioning and coping among mothers of children with autism: a population-based study. *Pediatrics*. 2007;119:e1040–e1046.

Mugno D, Ruta L, D'Arrigo VG, Mazzone L. Impairment of quality of life in parents of children and adolescents with pervasive developmental disorder. *Health Qual Life Outcomes*. 2007;5:22.

*Murray C, Johnston C. Parenting in mothers with and without attention-deficit/hyperactivity disorder. *J Abnorm Psychol*. 2006;115:52–61.

*Perlick DA, Rosenheck RA, Kaczynski R, Swartz MS, Canive JM, Lieberman JA. Components and correlates of family burden in schizophrenia. *Psychiatr Serv*. 2006;57:1117–1125.

*Perlick DA, Rosenheck RA, Miklowitz DJ, et al; STEP-BD Family Experience Collaborative Study Group. Prevalence and correlates of burden among caregivers of patients with bipolar disorder enrolled in the Systematic Treatment Enhancement Program for Bipolar Disorder. *Bipolar Disord*. 2007;9:262–273.

Reinares M, Colom F, Sanchez-Moreno J, et al. Impact of caregiver group psychoeducation on the course and outcome of bipolar patients in remission: a randomized controlled trial. *Bipolar Disord*. 2008;10:511–519.

Reinares M, Vieta E, Colom F, et al. What really matters to bipolar patients' caregivers: sources of family burden. *J Affect Disord*. 2006;94:157–163.

Schieve LA, Blumberg SJ, Rice C, Visser SN, Boyle C. The relationship between autism and parenting stress. *Pediatrics*. 2007;119(Suppl 1):S114–S121.

Schwichtenberg A, Poehlmann J. Applied behaviour analysis: does intervention intensity relate to family stressors and maternal well-being? *J Intellect Disabil Res*. 2007;51:598–605.

Seipp CM, Johnston C. Mother–son interactions in families of boys with attention-deficit/hyperactivity disorder with and without oppositional behavior. *J Abnorm Child Psychol*. 2005;33:87–98.

Shu BC, Lung FW. The effect of support group on the mental health and quality of life for mothers with autistic children. *J Intellect Disabil Res*. 2005;49:47–53.

Siklos S, Kerns KA. Assessing the diagnostic experiences of a small sample of parents of children with autism spectrum disorders. *Res Dev Disabil*. 2007;28:9–22.

Singh MK, DelBello MP, Stanford KE, et al. Psychopathology in children of bipolar parents. *J Affect Disord*. 2007;102:131–136.

Smith LE, Seltzer MM, Tager-Flusberg H, Greenberg JS, Carter AS. A comparative analysis of well-being and coping among mothers of toddlers and mothers of adolescents with ASD. *J Autism Dev Disord*. 2008;38(5):876–889.

Smith MJ, Greenberg JS, Mailick Seltzer M. Siblings of adults with schizophrenia: expectations about future caregiving roles. *Am J Orthopsychiatry*. 2007;77:29–37.

Tonge B, Brereton A, Kiomall M, Mackinnon A, King N, Rinehart N. Effects on parental mental health of an education and skills training program for parents of young children with autism: a randomized controlled trial. *J Am Acad Child Adolesc Psychiatry*. 2006;45:561–569.

Wu EQ, Birnbaum HG, Shi L, et al. The economic burden of schizophrenia in the United States in 2002. *J Clin Psychiatry*. 2005;66:1122–1129.

▲ 55.13 Pediatric Sleep Disorders

JESS P. SHATKIN, M.D., AND ANNA IVANENKO, M.D., PH.D.

INTRODUCTION

Although precise estimates of the frequency of sleep disorders in children and adolescents are somewhat variable among studies, the understanding of the epidemiology of sleep pathology and its relationship to comorbid psychiatric diagnoses has grown substantially in recent years. Research in cognition and brain development over the last 15 years also has significantly advanced the understanding of the genetic and neurobiological bases of sleep disorders in children and adolescents. With an increasing understanding of sleep disorders, treatment interventions have also become more refined and useful.

This section reviews the research and clinical methods used to study sleep, new findings on the neurodevelopmental and cognitive effects of sleep loss and fragmentation, and the clinical description and treatment of sleep disorders in children and adolescents and their interface with psychiatric disorders.

SLEEP EVALUATION METHODS

One of the primary tools for studying sleep is *polysomnography* (PSG), which incorporates a variety of measures of physiological changes experienced during sleep. The American Thoracic Society developed a set of recommendations for the interpretation of pediatric PSG that has become the standard for sleep labs nationally. With the aid of portable technology, PSG is increasingly being performed at home, although the methods and results are not quite as sophisticated as when performed in the lab. While PSG remains the gold standard for sleep studies, other, less complicated and less expensive techniques are available for studying a variety of sleep disorders in children and adolescents.

Actigraphy was developed in the early 1970s and has come into increasing use in both research studies and clinical practice. Actigraphy allows for the study of sleep–wake patterns and circadian rhythms via the assessment of body movements. The device is typically worn on the wrist and can easily be adapted for home use.

A number of survey instruments exist for detecting problematic sleep in children and adolescents, including self-report questionnaires (such as the Sleep Disturbance Scale for Children, the Child Sleep Questionnaire, the Child and Family Sleep History Questionnaire, and the Sleep Disorders

Inventory for Children), sleep diaries, and parent report forms. Sleep logs or diaries provide information on bedtime, sleep latency, wake time, and the number of nocturnal awakenings, in addition to exercise, and the use and timing of caffeine, alcohol, and drug use. Sleep logs are typically completed for a 2-week period by parents and caregivers or adolescents themselves. Although subjective, sleep logs provide a wealth of useful information for treating insomnia and sleep-phase disorders and are available readily in many forms on the Internet. Finally, the Epworth Sleepiness Scale, commonly used with adults as a subjective measure of the propensity to fall asleep in certain situations and to monitor treatment efforts, has recently been modified for use in children and adolescents.

NEUROCOGNITIVE EFFECTS OF SLEEP DISRUPTION

Most of the research examining neurocognitive sequelae of sleep loss and fragmentation in children has focused on the effects of sleep restriction and sleep-disordered breathing on daytime performance in children and adolescents. As little as one night of acute sleep restriction in children can result in significant attention deficit. A study conducted in a sample of 7- and 8-year-old healthy children subjected to 1 hour of sleep restriction for 1 week demonstrated impaired attention and concentration, along with significant reorganization in brain cortical electroencephalographic (EEG) response as measured by multiarray P300 (cognitive) evoked potentials. By contrast, one additional hour of sleep in children aged 8 to 12 years over 1 week resulted in parent and child reports of improved emotional functioning and objective improvements in memory, attention, and math fluency. Similarly, in children 7 to 11 years old, one additional hour in bed, averaging 27 minutes of additional sleep, demonstrated improvements in teachers' blinded ratings of emotional lability and impulsivity.

Obstructive sleep apnea syndrome (OSAS), which is characterized by sleep fragmentation due to multiple arousals and disruption in ventilation following episodes of complete or partial airway obstruction during sleep, has been studied with respect to its effect on neurocognitive functions in children. Numerous studies report academic deficits and learning problems in children with OSAS that appear to be reversible after treatment of the condition. Hyperactive, inattentive, and disruptive behaviors resembling those found in attention-deficit/hyperactivity disorder (ADHD) have been increasingly reported in children with snoring and OSAS. Similar behavioral changes are seen in some children with periodic limb movement disorder (PLMD) and restless legs syndrome (RLS).

In adults, sleep deprivation adversely affects functions mediated by the prefrontal cortex, particularly attention and working memory, upon which most cognition depends. Deduction and critical thinking (logic-based tasks) show little change in accuracy with sleep deprivation, but do reflect cognitive slowing. Divergent skills involved in decision-making, however,

are greatly affected by sleep loss and involve trouble assimilating changing information, updating strategies based upon new information, innovation, risk assessment, maintaining interest, insight, communication, and memory. Unfortunately, a lack of sleep, in addition to increasing levels of responsibility, lead to somewhat of a perfect storm, contributing to the high-risk behavior so commonly observed among adolescents.

A number of studies have indicated reduced intelligence scores among children with sleep-disordered breathing compared to normal controls, including significantly impaired general conceptual ability. Lower academic achievement has also been observed in children with OSAS compared to children without reported histories of sleep-disordered breathing when matched for age, race, gender, socioeconomic status, and school attended.

SLEEP DISORDERS IN CHILDREN

It is estimated that approximately 25 percent of children will suffer from a sleep problem at some point during childhood. The types of problems affecting children vary greatly in their frequency and severity, ranging from the more common bedtime resistance and anxiety to the occasional primary sleep disorders, such as narcolepsy and OSAS. Regardless of the primary complaint, parents frequently report sleep difficulties among their children, and these reports are remarkably consistent with most researchers identifying up to 50 percent of preschool children, roughly 30 percent of school-aged children, and approximately 40 percent of adolescents as having sleep difficulties. Much higher rates of sleep disorders are generally reported among those with neurodevelopmental and psychiatric disorders. Among children with intellectual disability, for example, 30 to 80 percent have sleep problems, while 50 to 90 percent of children with autism spectrum disorders are affected. Despite a high overall prevalence, however, sleep problems are not often discussed during visits with primary care practitioners and are therefore left undiagnosed and untreated.

Classification of Sleep Disorders

The *International Classification of Sleep Disorders*, Third Edition (ICSD-3), and the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5), are the primary classification systems for sleep disorders. While the ICSD-3 and the DSM-5 cover much of the same ground, the ICSD-3 classification is more comprehensive and precise, while the DSM-5 specifies 12 main categories of sleep disorders (see Table 55.13–1).

Sleep-Disordered Breathing

The term *sleep-disordered breathing* (SDB) encapsulates a variety of difficulties, including primary snoring upper airway resistance syndrome

(UARS) and OSAS, and occurs relatively frequently in otherwise healthy children. Although accurate figures are difficult to obtain, it is generally estimated that 2 to 5 percent of all children may be affected by OSAS, and 7 to 12 percent of children may suffer from primary snoring.

Whereas adults with OSAS most often present with *excessive daytime somnolence* (EDS), children rarely exhibit this symptom and instead tend to present with externalizing behaviors, such as hyperactivity, aggression, social withdrawal, and learning difficulties. In one study, fewer than 15 percent of children with OSAS exhibited EDS at all. Diagnosis of OSAS further requires a consistent clinical history and physical examination. To augment these two diagnostic measures and reveal the extent of the effect on sleep, children should also undergo overnight PSG.

Table 55.13–1.
DSM-5 Classification of Sleep Disorders

- Insomnia disorder
  - Hypersomnolence disorder
  - Narcolepsy
  - Breathing-related sleep disorders
    - Obstructive sleep apnea
    - Central sleep apnea
    - Sleep-related hypoventilation
  - Circadian rhythm sleep–wake disorders
    - Delayed sleep-phase type
    - Advanced sleep-phase type
    - Irregular sleep–wake type
    - Non–24-hour sleep–wake type
    - Shift-work type
  - Parasomnias
    - Nonrapid eye movement sleep-arousal disorder
    - Nightmare disorder
    - Rapid eye movement sleep-behavior disorder
    - Restless legs syndrome
- Substance/medication-induced sleep disorder
- Other specified insomnia disorder
- Unspecified insomnia disorder
- Other specified hypersomnolence disorder
- Unspecified hypersomnolence disorder
- Other specified sleep–wake disorder
- Unspecified sleep–wake disorder

Data from American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

Children with OSAS present with loud snoring, sweating, agitated sleep, and a variety of related sleep problems (e.g., persistent enuresis, night terrors, somnambulism, bruxism, etc.). EDS, somnambulism, bruxism, and

anxiety are also frequently noted in adolescents with OSAS. Physical examination often demonstrates craniofacial abnormalities and adenotonsillar hypertrophy, and PSG indicates frequent apneas/hypopneas and oxygen desaturations. Outcomes of OSAS can include significant health problems, such as failure to thrive, hypertension, pulmonary hypertension with or without cor pulmonale, and mental health problems.

The most common treatment for pediatric SDB is adenotonsillectomy, which cures the problem in approximately 85 percent of affected children. A continuous positive airway pressure (CPAP) mask is often used for children who fail surgical treatment or for whom surgery is not indicated.

Insomnia

Pediatric insomnia has been much less studied and is more poorly understood than adult insomnia. A number of efforts have been made to define pediatric insomnia, estimate its prevalence among different pediatric populations, and develop clinical guidelines for behavioral and pharmacological treatment.

The prevalence of pediatric insomnia is estimated at about 1 to 6 percent among general pediatric populations, with a much higher prevalence in children with neurodevelopmental and chronic medical and psychiatric conditions. When bedtime resistance and disruptive nighttime awakenings are included, the prevalence of sleep-disrupted behavior approaches 25 to 50 percent in preschool-aged children. Psychiatric symptoms have been reported in almost 50 percent of children with persistent insomnia. It has been proposed that persistent insomnia may also represent an early sign of emotional distress in susceptible children with poorly preserved sleep homeostasis.

The ICSD-3 describes two forms of behavioral insomnia of childhood as unique diagnostic entities to emphasize sleep difficulties that result from inappropriate sleep associations or inadequate parental limit setting. *Limit-setting type* is due to parents having difficulties establishing and enforcing bedtimes, while *sleep-onset association type* is characterized by reliance on maladaptive and inappropriate sleep associations such as rocking, watching television, feeding, and falling asleep in the parents' bed. The child is usually unable to fall asleep in the absence of these conditions at both bedtime and following nocturnal arousals. Behavioral insomnia generally leads to sleep fragmentation and increased nighttime awakenings, insufficient sleep duration, later bedtimes, and daytime sleepiness.

Good sleep hygiene and behavioral interventions, such as extinction or bedtime fading, are the first recommended treatments for pediatric insomnia. Children who do not respond to behavioral interventions may be considered for pharmacological management of their insomnia. There are no U.S. Food and Drug Administration (FDA)–approved medications for

the treatment of insomnia in children. Nonetheless, a large study of over 1,000 child and adolescent psychiatrists found that 96 percent recommended at least one prescription medication and 88 percent advised an over-the-counter medication for sleep in a typical month. Despite the lack of FDA approval, however, limited empirical evidence supports the use of melatonin and clonidine for childhood insomnia. Meanwhile, benzodiazepines, guanfacine, pyrimidine derivatives (e.g., zaleplon and zolpidem), sedating antidepressants (e.g., trazodone and mirtazapine), sedating antipsychotics (e.g., quetiapine), and sedating antihistamines (e.g., diphenhydramine, doxylamine, and hydroxyzine) are often employed to improve symptoms of insomnia in children and adolescents but have virtually no solid evidence to support their use.

Coincident with the increasing use of herbal medications and dietary supplements in the United States, a number of studies have been conducted over the last decade that demonstrate that melatonin is an effective, safe, and well-tolerated agent, especially in cases of sleep-initiation insomnia caused by circadian factors. Several placebo-controlled studies of melatonin in normal, healthy children with idiopathic sleep-initiation insomnia have shown that melatonin administered at bedtime reduces sleep-onset latency time and increases total sleep time.

Little information is available on the use of the non-benzodiazepine hypnotic agents zaleplon and ramelteon, but studies have found both zolpidem and eszopiclone ineffective for ADHD-related insomnia. Additionally, hypnogogic hallucinations and paradoxical agitation have been demonstrated when adult doses of these agents have been prescribed to children and adolescents. At this time there is no evidence supporting the use of these drugs in the treatment of pediatric insomnia.

The most effective treatment for both decreasing sleep latency and improving sleep maintenance in adults is cognitive behavior therapy (CBT). Core components of CBT for insomnia (CBT-I) include sleep education, stimulus control (e.g., removing clocks and cell phones from the bedroom, maintaining a quiet sleep environment, etc.), sleep restriction (e.g., limiting the amount of time in bed until the patient is sleeping 90 percent of the time they are in bed), circadian rhythm maintenance (e.g., going to bed and awakening at the same time each day), sleep hygiene, addressing cognitive distortions (e.g., catastrophizing such as “if I don’t fall asleep soon, I’ll be a complete wreck tomorrow”), and arousal reduction (e.g., relaxation techniques such as belly breathing and biofeedback). Numerous randomized studies of adults demonstrate that CBT-I is even more effective for insomnia than sleep medications. New studies have now demonstrated the benefits of CBT-I for adolescents as well.

Narcolepsy

Major strides in the understanding of the etiology of narcolepsy have been

taken in recent years as remarkable discoveries have identified a new neuropeptide, hypocretin or orexin, which is linked to animal models of narcolepsy. The exact prevalence of childhood narcolepsy is unknown. The youngest patient on record was diagnosed at age 12 months. In retrospective studies, about one-third of adults with narcolepsy indicate onset prior to 15 years of age and roughly 15 percent prior to 10 years of age.

Children with narcolepsy are commonly misdiagnosed with other neurological and/or psychiatric disorders, such as epilepsy, ADHD, mood disorders, and psychotic disorders. Excessive daytime sleepiness is the most common presentation of narcolepsy in school-aged children, although according to some reports 50 to 70 percent of childhood narcolepsy patients present with symptoms of cataplexy. Nighttime sleep is frequently disturbed in patients with narcolepsy, and periodic limb movements during sleep are also common. The diagnosis of narcolepsy requires sleep laboratory assessment, including nocturnal PSG and a multiple sleep latency test.

Effective treatment of narcolepsy in children and adolescents demands establishing a regular sleep schedule, providing extensive education on sleep hygiene, and implementing scheduled daytime naps if possible. Excessive daytime sleepiness in children and adolescents with narcolepsy is commonly treated with psychostimulants such as methylphenidate and dextroamphetamine. Modafinil, a nonstimulant alerting agent, has been successfully prescribed to children and adolescents with narcolepsy in a dose range from 100 to 600 mg/day in divided doses.

Tricyclic antidepressants such as clomipramine, protriptyline, and imipramine have long been known to be effective for symptoms of cataplexy. For the last decade, serotonin-specific reuptake inhibitors and mixed-action antidepressants have also been increasingly prescribed to patients with cataplexy. Sodium oxybate (γ-hydroxybutyrate [GHB]) has been approved by the FDA for the treatment of cataplexy, and there has been at least one report of its use in children with narcolepsy. Due to the complexity of narcolepsy symptoms and their effect on child development, a variety of treatment modalities, including individual and family therapy, school advocacy, and academic interventions, should be used to optimize symptom control.

Parasomnias

Parasomnias are disruptive physical acts that occur during specific sleep stages or during the transition between sleep stages and include non-REM types (somnambulism and sleep terrors), REM sleep behavior disorder, somniloquy, bruxism, nightmares, and RLS. Although the incidence of general psychopathology in children with non-REM sleep parasomnias is quite low, the overall prevalence of having at least one parasomnia by age

13 years has been reported to be as high as 78 percent. The events, although rarely requiring clinical intervention, are often very troubling for parents.

Somnambulism, or sleepwalking, is relatively common in childhood, with annual prevalence rates approaching 17 percent. Although somnambulism rarely results in harm to the subject, it is possible for an episode of somnambulism to become a confusional arousal, in which the likelihood of accidental violence (to oneself or others) is much greater. The prevalence of confusional arousal is less well established, although a 4 percent incidence was noted in a longitudinal study of 212 children of ages 6 to 16 years selected at random in Stockholm.

Sleep terrors are somewhat less common than somnambulism and are difficult to quantify because of the self-report nature of the data and the fact that parents often confuse them with nightmares. Annual incidence rates have been reported as high as 6 percent and prevalence rates as high as 17 percent among children aged 13 years and younger. Most non-REM parasomnias are first noticed by parents of affected children. Although a large percentage of children have at least one incidence of somnambulism, far fewer present with recurrent and disruptive somnambulism. Sleep terrors follow a similar trend, although they are less common than sleepwalking. Numerous studies have demonstrated a connection between anxiety levels and parasomnias in children, showing that increased anxiety correlates with increased prevalence of sleepwalking and sleep terrors.

Although the treatment of non-REM parasomnias has received minimal formal study, there are numerous measures that have been shown to be effective. In addition to safety measures, the use of scheduled awakenings and naps, and in severe cases clonazepam or tricyclic antidepressants, may be effective. In addition, a formal sleep study may be helpful in revealing SDB as a comorbid disorder for which treatment may reduce the apparent severity of the parasomnia.

Sleeptalking, or *somniloquy*, is considered the most common parasomnia, with a reported prevalence of greater than 50 percent in children between the ages of 3 and 13 years. Somniloquy is often comorbid with sleepwalking and sleep terrors, suggesting a common pathophysiology among the three disorders. Nightmares are common among children and adolescents and are only considered abnormal when they are recurrent and cause excessive distress. Although sometimes associated with posttraumatic stress disorder (PTSD), up to 70 percent of children report that the content of their nightmares is influenced by media seen on television or at the movies. Effective treatment for nightmares generally includes CBT and rescripting, but can also include hypnotherapy and medication. REM sleep behavior disorder, in which subjects become aroused during REM sleep and “act out” their dreams due to loss of body paralysis, is not generally observed in children and is more closely associated with brain damage and dementia.

RLS is characterized by an irresistible urge to move the legs commonly associated with discomfort and pain, which is partially relieved by movement. In the adult population, RLS has an estimated prevalence of 10 to 15 percent. PLMD, no longer classified by DSM-5, is characterized by episodic limb movements during sleep, generally involving the leg along with the extension of the ankle and toe.

Only recently has the presence of RLS in children and adolescents been elucidated with studies generally demonstrating a prevalence rate of RLS among the general pediatric population to average around 2 percent. There is considerable interest in the relationship between RLS, PLMD, and ADHD. The prevalence of RLS in patients with ADHD is estimated to be between 10.5 and 44 percent, emphasizing the need to consider both diagnoses and the possible interrelationships between them in children and adolescents.

Prior to the diagnosis of RLS, a child often presents with sleep disturbances and daytime impairments, including extreme fatigue, irritability, and hyperactivity, and the discomfort in the extremities that accompanies the disorder is often confused with growing pains. Other qualitative signs suggestive of RLS include an urge to move the legs with a worsening sensation when sitting or lying down, the alleviation of discomfort through movement, and worsening symptoms in the evening. Several studies have demonstrated an autosomal-dominant inheritance for early-onset RLS. Family history, therefore, is an important indicating factor in diagnosing RLS in children and adolescents.

Most treatments for RLS/PLMD include a pharmacological component, although proper sleep hygiene is also very important. Children and adolescents with RLS/PLMD should reduce caffeine intake, eliminate alcohol and tobacco use, avoid stimulating activities prior to bedtime, and minimize sleep deprivation. Correlation among low ferritin levels, iron deficiency, and RLS symptoms has been identified in both children and adults. Providing ferrous sulfate supplementation for those children who are deficient has been demonstrated significantly to reduce PLMD symptoms. Other medications have also demonstrated efficacy in reducing symptoms. One randomized double-blind, placebo-controlled study of children with RLS aged 7 to 12 years found a combination of carbidopa/L-DOPA to be safe and effective. In adults with RLS, dopaminergic medications, including levodopa, pramipexole, and ropinirole, are most often used. Gabapentin has also been shown to improve RLS symptoms in children. Although these medications are prescribed to children and can be effective in reducing symptoms, none has been approved by the FDA for the treatment of RLS/PLMD in children. As a result, the medications should be administered with great caution and intensive monitoring of potential side effects.

Circadian Rhythm Sleep–Wake Disorders

Circadian rhythm sleep–wake disorders occur due to a misalignment between the internal circadian rhythm and the sleep–wake schedule. Over 10 percent of children are thought to be affected by circadian rhythm sleep problems. Delayed sleep-phase type occurs in as many as 5 to 10 percent of adolescents, while advanced sleep-phase type is more common in the elderly and occasionally in preschool children.

A delay in the normal sleep phase is typical in the peripubertal period, leading to later bedtimes. Extensive homework and extracurricular obligations coupled with early school start times frequently result in excessive daytime sleepiness, academic difficulties, irritability, mood instability, attention and memory impairment, and family conflict. Recognizing the detrimental effects of delayed sleep phase on adolescents, the American Academy of Pediatrics has published a policy statement urging a delay in the school start time so that high school students can achieve adequate sleep.

Treatment of delayed sleep phase typically involves sleep education and hygiene, along with gradual advancement of the sleep phase. Bright light morning therapy (5,000 to 10,000 lux), along with low-dose melatonin (approximately 500 mcg) about 4 to 5 hours before the desired bedtime, have both been found helpful in advancing the sleep phase after several weeks. Chronotherapy, or the gradual delay in bedtime by 3 hours each night until the desired bedtime is achieved, can also be effective, although the schedule is very difficult to maintain.

Nocturnal Enuresis

Although no longer classified by DSM-5 as a sleep–wake disorder, *nocturnal enuresis* is a frequent occurrence, affecting approximately 30 percent of 4 year olds, 10 percent of 6 year olds, 5 percent of 10 year olds, 3 percent of 12 year olds, and 1 percent of individuals aged 15 years or older. *Primary nocturnal enuresis* (never consistently dry at night) is likely to have a multifactorial etiology and may be due to difficulties with bladder musculature stability, central nervous system arousability, pontine reflex function, internal sphincter tone, functional bladder capacity, nocturnal urine production, and a maturational delay in antidiuretic hormone (ADH) secretion. *Secondary nocturnal enuresis* (previously dry for at least 6 months), by contrast, is typically caused by urinary tract infections, diabetes mellitus, and psychological factors.

The first step in treatment is to educate the family and patient regarding proper sleep hygiene. All caffeine and alcohol intake should be discontinued because both decrease ADH secretion, and late-night fluid intake should be strictly limited. On occasion, the introduction of an afternoon nap will be beneficial by virtue of decreasing the quantity of

nighttime deep sleep, thereby making it easier for the patient to awaken to the internal cues of bladder fullness. Finally, parents sometimes find that briefly awakening the child after 2 or 3 hours of sleep and sitting him or her on the toilet will decrease the frequency of enuresis without a significant impairment in sleep efficacy. Should these educational strategies be ineffective, both behavioral and pharmacological treatments exist for nocturnal enuresis.

Behavioral treatment most commonly involves the use of an enuretic alarm, which has the highest cure rate and lowest relapse rate. Other behavioral interventions include bladder training to increase the number of micturitions during the day or to enlarge bladder capacity. The use of reward systems, such as star charts, has also been shown to be useful. Finally, cognitive therapy (e.g., explaining the enuretic process and maintaining a daily diary), motivational therapy, pelvic floor muscle training, and biofeedback also have some reported utility.

The two medications most commonly accepted as appropriate pharmacological treatment for pediatric nocturnal enuresis are desmopressin acetate and imipramine. The use of amitriptyline and anticholinergic agents, such as oxybutynin and tolterodine, however, has also been described.

Sleep Disorders in Children with Psychiatric Disorders

Sleep problems often occur in the context of mental illness, either as primary concerns (e.g., primary insomnia) or secondary concerns (e.g., as a consequence of depression or anxiety). All sleep problems identified early in childhood have been later correlated with anxiety, depression, attentional difficulties, family or parental distress, continued sleep problems, aggression, delinquency, and social problems. Some of the more common psychiatric disorders among children and adolescents and their effect on sleep (and vice versa) will be reviewed here.

Attention-Deficit/Hyperactivity Disorder

Numerous studies based on parent reports have demonstrated the prevalence of sleep disorders among children with ADHD to be higher than that for healthy controls, controls with other psychiatric illness, and healthy sibling controls by up to fivefold. Actigraphy has shown children with ADHD to have greater variation in sleep onset time, wake time, and single-night sleep duration than controls, and in at least one study there were significantly more bedtime struggles and longer total sleep duration in the ADHD group. Although children with ADHD often have difficulty settling down to sleep and may sleep more total hours than controls, restricting the amount of sleep a child receives in an experimental setting can also result in symptoms of ADHD, including inattention, hyperactivity,

impulsivity, and poor cognitive performance.

ADHD is also associated with snoring during sleep. Habitual snoring has been reported to be three times more common in children with ADHD than control children drawn from child psychiatry and pediatric clinics. In addition, children with ADHD have been found to have a greater frequency of PLMD while maintaining normal sleep architecture. Children with ADHD also suffer from SDB at a greater rate than expected. SDB has been associated with poor memory performance on standardized psychometric tests and lower academic achievement in children, whereas in adults, deficits in attention, concentration, memory, and verbal and nonverbal intelligence have been noted.

Children with ADHD may also suffer circadian rhythm abnormalities. Among children with ADHD, there is greater inconsistency or instability in time to sleep onset, wake time, and sleep duration than in controls. Furthermore, a high prevalence of ADHD has been shown among children with chronic sleep-onset insomnia, and sleep-onset insomnia in children with ADHD is associated with later dim-light melatonin release.

Fragmented or disrupted sleep, rather than lessened total sleep time, may be the true culprit behind the daytime behavioral changes and disruptions seen in some children. To that end, recommendations to families with a disruptive or hyperactive child should include an emphasis on basic sleep hygiene. Many children with ADHD take medication regularly, which can further complicate an understanding of their sleep cycle. Much investigation, however, has suggested that children with ADHD still demonstrate more sleep disturbances than controls regardless of medication use, particularly with regard to sleep-onset latency, sleep efficiency, and total sleep time.

Opinions vary greatly on how to manage insomnia due to either ADHD or the stimulant medications that are used to treat it. Typical recommendations include lowering the dose of the stimulant, changing to a shorter-acting preparation, adding a low dose of stimulant (if the insomnia is believed to be secondary to hyperactive “rebound”), or changing to an entirely different stimulant or treatment, such as atomoxetine. If adjunctive pharmacotherapy is used, typical medications include melatonin, clonidine, antihistamines (e.g., diphenhydramine, cyproheptadine), and antidepressants (e.g., trazodone, mirtazapine).

Autism Spectrum Disorders

Children with *autism spectrum disorders* (ASDs) are also susceptible to sleep difficulties, with prevalence estimates ranging as high as 80 percent. Younger children and those with more severe cognitive or intellectual disabilities tend to demonstrate increased sleep problems, which often persist over time. The most frequently reported sleep problems among those with ASD include difficulty falling asleep, restless sleep, frequent

awakenings with difficulty falling back to sleep, early-morning awakenings, irregular sleep–wake patterns, shortened duration of sleep, dyssomnias, and parasomnias.

Numerous hypotheses have been put forth to explain the reported sleep difficulties among children with ASD, including failure to recognize environmental and social cues, poorly developed circadian rhythms secondary to social deficits, altered melatonin production, and abnormalities in the hypothalamic–pituitary–adrenal axis and serotonin systems. Many treatments have been proposed for children with ASD with comorbid sleep difficulties.

Various behavioral treatments have been used with varying success in the ASD population. Most favorably, the use of proper sleep-hygiene techniques, stimulus control, graduated extinction, scheduled awakenings, and daytime nap restriction have all shown utility among children with autism and intellectual disability. Light therapy, in combination with chronotherapy, may also be useful in children with autism who have circadian rhythm abnormalities. Pharmacological treatments also have proven to be useful for many, but not all, children with ASD.

Various prescribed medications have reported utility in the treatment of sleep difficulties in the ASD population, including antihistamines, sedating antidepressants, α-2 agonists, benzodiazepines, and antipsychotics. Melatonin has also demonstrated utility in treating children with neurodevelopmental disabilities and autism.

Mood Disorders

Sleep complaints are among the most prevalent symptoms of major depression in children and perhaps the most prevalent symptom among adolescents. Insomnia has been reported to affect nearly 90 percent of depressed adolescents, whereas *hypersomnia* affects a considerably smaller proportion, perhaps around 25 percent. Some adolescents—around 10 percent—continue to experience insomnia even after remission of depression. The use of actigraphy in diagnosing sleep problems among depressed children and adolescents highlights the compromised sleep quality and abnormal circadian rhythms these individuals experience while depressed. PSG studies in children with depression often demonstrate decreased REM latencies, increased total REM, and increased sleep-onset time.

Many children who are or will be later diagnosed with bipolar disorder present with sleep-related symptoms. When depressed, these children may experience exhaustion, distractibility, irritability, and impulsivity at levels significantly greater than at baseline. In contrast, children experiencing mania may present with unusual amounts of energy and a decreased need for sleep. Such extremes of sleep behavior warrant clinical assessment.

Seasonal affective disorder (SAD) has a reported prevalence of 3 to 4

percent in children and adolescents, for whom more than 20 percent of their parents report that their children are sleepier and generally more tired during winter months. Studies of children with a diagnosis of SAD have found that they did not differ from controls in the level of nocturnal activity but did experience significantly lower diurnal activity levels and had less robust circadian rhythms, with a reduction in amplitude.

There have been a small number of studies examining the effects of pharmacological and behavioral interventions in depressed children and adolescents suffering from sleep problems. A combination of improved sleep hygiene and cognitive behavioral therapy has been shown to be effective in managing the insomnia and awakenings associated with depression. In those cases that require further treatment, the most effective drug therapy appears to be a combination of an antidepressant with the short-term use of a sedative or hypnotic. As noted previously, however, some children and adolescents experience a paradoxical reaction to these medications, which severely limits their utility.

Anxiety Disorders

Although it is difficult to establish causality between early sleep problems and later onset of anxiety and depression, it is clear that anxiety and sleep problems are intimately connected in childhood and that the presence of a sleep problem at age 4 years is significantly correlated with the development of depression and anxiety by age 15 years. Fears at nighttime are common among children, with up to 75 percent of those aged 4 to 12 years reporting such fears, which generally take the form of animals, fictitious characters (e.g., witches, monsters), being kidnapped, or being teased by peers. Although no specific numbers are available, for the majority of children with DSM-5 anxiety disorders, sleep problems typically follow. In addition to the aforementioned concerns, anxiety is also believed to predispose children to parasomnias and nightmares.

The effects of anxiety on sleep are profound, and the clinical presentation depends largely on the specific disorder. Individuals with *generalized anxiety disorder* (GAD) have a wide range of sleep disturbances, which can include insomnia and hypersomnia. These GAD sleep symptoms may be due to central nervous system hypervigilance and hyperarousal such that the individual is unable to relax sufficiently to fall asleep and maintain good sleep quality. Patients with panic disorder tend to present with insomnia, restless and fragmented sleep, and nocturnal panic attacks. These symptoms often result in trouble initiating and maintaining sleep, increased sleep latency, and reduced sleep efficiency.

Children and adolescents with PTSD may experience significant sleep difficulties. Individuals with PTSD often have trouble initiating sleep and experience significant nighttime arousals and disruptive nightmares. Because children with PTSD may have flashbacks to their traumatic

experiences, bedtime can become a stressful time in which they experience increased anxiety and are unable to relax sufficiently for restful sleep. The accompanying nightmares, which are a characteristic symptom of PTSD, and increased limb movements, which are associated with nightmares, lead to disrupted REM sleep.

Cognitive and behavioral techniques, systematic desensitization, and imagery rehearsal therapy have demonstrated efficacy in decreasing nightmares and are integral parts of treating the sleep disturbances accompanying obsessive compulsive disorder (OCD), PTSD, and GAD. Cognitive behavior therapy (CBT), along with systematic desensitization, has been demonstrated to improve the sleep disturbances associated with OCD. In cases that are particularly severe, serotonin-specific reuptake inhibitors such as fluoxetine, sertraline, and paroxetine further aid in the reduction of symptoms. Children and adolescents with PTSD generally have a complex constellation of symptoms, and the primary psychiatric diagnosis must be treated first before the sleep disturbances will resolve. CBT and play or expressive therapies, in which the children process the traumatic events triggering the PTSD, have been shown to be very effective, and antidepressants and anxiolytics may hasten the recovery process in combination with other therapies.

References

*Adolescent Sleep Working Group; Committee on Adolescence; Council on School Health. School start times for adolescents. *Pediatrics*. 2014;134(3):642–649.

Alfano CA, Reynolds K, Scott N, Dahl RE, Mellman TA. Polysomnographic sleep pattern of non-depressed, non-medicated children with generalized anxiety disorder. *J Affect Disord*. 2013;147(1–3):379–384.

Aran A, Einen M, Lin L, Plazzi G, Nishino S, Mignot E. Clinical and therapeutic aspects of childhood narcolepsy-cataplexy: a retrospective study of 51 children. *Sleep*. 2010;33(11):1457–1464.

Badin E, Haddad C, Shatkin JP. Insomnia: the sleeping giant of pediatric public health. *Curr Psychiatry Rep*. 2016;18(5):47.

*Chervin RD, Archbold KH, Dillon JE, et al. Inattention, hyperactivity, and symptoms of sleep-disordered breathing. *Pediatrics*. 2002;109(3):449–456.

Cortese S, Brown TE, Corkum P, et al. Assessment and management of sleep problems in youths with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2013;52(8):784–796.

Cortesi F, Giannotti F, Ivanenko A, Johnson K. Sleep in children with autistic spectrum disorder. *Sleep Med*. 2010;11(7):659–664.

England SJ, Picchietti DL, Couvadelli BV, et al. L-dopa improves restless legs syndrome and periodic limb movements in sleep but not attention-deficit/hyperactivity disorder in a double-blind trial in children. *Sleep Med*. 2011;12(5):471–477.

Fallone G, Acebo C, Arnedt TA, Seifer R, Carskadon MA. Effects of acute sleep restriction on

behavior, sustained attention, and response inhibition in children. *Percept Mot Skills*. 2001;93(1):213–229.

*Gozal D. Sleep-disordered breathing and school performance in children. *Pediatrics*. 1998;102(3 Pt 1):616–620.

Gruber R, Cassoff J, Frenette S, Wiebe S, Carrier J. Impact of sleep extension and restriction on children's emotional lability and impulsivity. *Pediatrics*. 2012;130(5):1155–1161.

Gruber R, Sadeh A, Raviv A. Instability of sleep patterns in children with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2000;39(4):495–501.

Guilleminault C, Palombini L, Pelayo R, Chervin RD. Sleepwalking and sleep terrors in prepubertal children: what triggers them? *Pediatrics*. 2003;111(1):e17–e25.

Hendricks MC, Ward CM, Grodin LK, Slifer KJ. Multicomponent cognitive-behavioural intervention to improve sleep in adolescents: a multiple baseline design. *Behav Cogn Psychother*. 2014;42(3):368–373.

Ivanenko A, Crabtree VM, Gozal D. Sleep and depression in children and adolescents. *Sleep Med Rev*. 2005;9(2):115–129.

Jacobs GD, Pace-Schott EF, Stickgold R, Otto MW. Cognitive behavior therapy and pharmacotherapy for insomnia: a randomized controlled trial and direct comparison. *Arch Intern Med*. 2004;164(17):1888–1896.

Johnson EO, Chilcoat HD, Breslau N. Trouble sleeping and anxiety/depression in childhood. *Psychiatry Res*. 2000;94(2):93–102.

Malow B, Adkins KW, McGrew SG, et al. Melatonin for sleep in children with autism: a controlled trial examining dose, tolerability, and outcomes. *J Autism Dev Disord*. 2012;42(8):1729–1737.

Malow BA, Byars K, Johnson K, et al. A practice pathway for the identification, evaluation, and management of insomnia in children and adolescents with autism spectrum disorders. *Pediatrics*. 2012;130(Suppl 2):S106–S124.

*Mindell JA, Emslie G, Blumer J, et al. Pharmacologic management of insomnia in children and adolescents: consensus statement. *Pediatrics*. 2006;117(6):e1223–e1232.

Murali H, Kotagal S. Off-label treatment of severe childhood narcolepsy-cataplexy with sodium oxybate. *Sleep*. 2006;29(8):1025–1029.

*O'Brien LM. Cognitive and behavioral consequences of obstructive sleep apnea. In: Sheldon SH, Kryger MH, Ferber R, Gozal D, eds. *Principles and Practice of Pediatric Sleep Medicine*. 2nd ed. Philadelphia, PA: Elsevier; 2014:231–238.

Owens JA, Rosen CL, Mindell JA, Kirchner HL. Use of pharmacotherapy for insomnia in child psychiatry practice: a national survey. *Sleep Med*. 2010;11(7):692–700.

Picchietti DL, Bruni O, de Weerd A, et al; International Restless Legs Syndrome Study Group (IRLSSG). Pediatric restless legs syndrome diagnostic criteria: an update by the International Restless Legs Syndrome Study Group. *Sleep Med*. 2013;14(12):1253–1259.

Ramtekkar U, Ivanenko A. Sleep in children with psychiatric disorders. *Semin Pediatr Neurol*. 2015;22(2):148–155.

Rosen G. Circadian rhythm disorders in children. In: Barkoukis T, Matheson J, Ferber R, Doghramji K, eds. *Therapy in Sleep Medicine*. Elsevier; 2012:465–474.

Shatkin JP, Pando M. Diagnosis and treatment of common sleep disorders in adolescence.

Adolesc Psychiatr. 2015;4:1–13.

Smedje H, Broman JE, Hetta J. Associations between disturbed sleep and behavioural difficulties in 635 children aged six to eight years: a study based on parents' perceptions. *Eur Child Adolesc Psychiatry*. 2001;10(1):1–9.

Taylor DJ, Zimmerman MR, Gardner CE, et al. A pilot randomized controlled trial of the effects of cognitive-behavioral therapy for insomnia on sleep and daytime functioning in college students. *Behav Ther*. 2014;45(3):376–389.

Vriend JL, Davidson FD, Corkum PV, Rusak B, Chambers CT, McLaughlin EN. Manipulating sleep duration alters emotional functioning and cognitive performance in children. *J Pediatr Psychol*. 2013;38(10):1058–1069.

56 Adulthood

CALVIN A. COLARUSSO, M.D.

As science cracks the genetic code and medicine successfully treats and eliminates diseases that in past centuries made survival to middle or old age rarities, the prospect of vast numbers of humans living beyond one hundred years is increasingly becoming a reality. Consequently, the task for future generations of mental health scientists, and authors of this chapter, will be to provide a greater understanding of what are the normal developmental experiences and expectations for those who live to one hundred and beyond.

Anticipating that future and relying on what is currently known about the mental and emotional life of those in "old age," this chapter includes material on healthy development and the rich quality of life and exquisite understanding of the human condition which is increasingly becoming the norm in the decades of the seventies, eighties, and nineties.

MATURITY

Success and happiness in adulthood are made possible by achieving a modicum of maturity—a mental state, not an age. However, the capacity for maturity is a direct outgrowth of the engagement and mastery of the developmental tasks of young and middle adulthood. From a developmental perspective, maturity may be defined as a mental state found in healthy adults that is characterized by a detailed knowledge of the parameters of human existence, a sophisticated level of self-awareness based on an honest appraisal of one's own experience within those basic parameters, and the ability to use this intellectual and emotional knowledge and insight in a caring way in relation to one's self and others.

The achievement of maturity in midlife leads to the emergence of the capacity for wisdom. Those who possess wisdom have learned from the past and are fully engaged in life in the present. Just as important, they anticipate the future and make the necessary decisions to enhance prospects for health and happiness. In other words, a philosophy of life has been developed that includes an understanding and acceptance of the person's place in the order of human existence. That worldview, which varies widely from individual to individual, produces fulfillment when the following aspects of the human condition are accepted and integrated.

The body must be cared for—in health, through regular checkups, exercise, and a proper diet, and in sickness, through prompt treatment and proper care. Caring for the body is not an end in itself, but it is critically important because sentience, the essence of human experience, springs from a healthy brain and body.

Human beings are individuals who are alone with themselves, separated and individuated from all others.

Paradoxically, human beings cannot survive or develop on their own. Helpless at birth and relatively dependent throughout childhood, even the most self-sufficient individuals require the sustaining presence of others. Humans exist in a framework of interdependence, a basic characteristic of all human relationships, be it the parental view of the child as a confirmation of his or her sexuality, the need of the child for the parents' loving care and protection, or the reversal of these roles between child and aging progenitor. The mature adult (unlike the child who takes, uses, controls, and dominates) mutes the grandiose expectations of childhood and propels the self toward interactions characterized by caring and mutuality, thus striking a balance between personal needs and those of others.

Change is a constant in life. A basic aspect of this change is the shifting nature of significant relationships. Adult involvement with loved ones, such as children, parents, colleagues, and friends, is in constant realignment. Healthy marriages deepen in significance, whereas others break up on the shoals of midlife development. Parents die. Children grow, leave, and return with new family members. As opposed to old age and, in some respects, childhood as well, the task for the mature person is to sort out, to categorize, and to set priorities among relationships, while in the process balancing emotional needs and realistic demands and responsibilities. The shifting nature of relationships stimulates the achievement of greater maturity by forcing a constant redefinition of who one is in relationship to others. Mature individuals mourn lost relationships but can remain focused on current and future ones.

All human beings—regardless of wealth, position, power, achievement, appearance, or cultural background—are on the same developmental course. All are born and will die. All have a body with the same functions. All have the same emotional needs for closeness and love and the same vulnerability to loss and deprivation.

Few individuals have an exaggerated importance. The wish for such grandiose prominence is universal, the result of the untempered narcissism of infancy and childhood. In reality, most individuals are important to themselves and a relatively limited number of others who know and love them. The mature individual accepts this fact without despair and uses the knowledge to set realistic goals and priorities that will increase the chances for personal happiness and fulfillment.

Personal time is limited. Everyone will die. Young children do not have the cognitive capacity to understand the notion of personal death. Spurred on by the thrust of physical growth and a seemingly endless future, adolescents and young adults think and act as though they are immortal. The true acceptance of time limitation and personal death occurs in midlife. Then the mature individual, religious or not, stimulated by an awareness of the aging process in the body, the maturation of children, the death of parents and friends, and the arrival of grandchildren, accepts the inevitability of a personal end. As with the realization of the limited importance of each individual, this painful recognition, which precipitates midlife crises in some, stimulates the mature individual to seek fulfillment in each moment, to define what is truly important, and to plan the future to actualize those priorities.

Money and possessions have limited intrinsic value. They are a means to an end, tools for enriching life and improving the human condition of loved ones and the broader community. Ownership of tangible objects is temporary; sooner or later they will be lost, left behind, or given to others.

Work occupies a central position in adult life. Considered a drudgery by some, the wise person recognizes its extraordinary value. In addition to the obvious function of earning a living, work is organizing, an activity that provides purpose and direction, a meaningful way to manage time, and an environment in which to form sustaining relationships. Satisfied midlife workers, at the peak of power and position in the workplace, facilitate the development of the skills and capabilities of younger colleagues while fully realizing that these individuals will, sooner or later, replace them and assume control of the levers of power. Wishes to hold down and to attack younger workers are sublimated into generativity, Erikson's term for enhancing the development of the next generation.

Midlife is the phase of life in which the experience of being human can be realized and enjoyed most fully. Then, the combination of physical health and vigor, power and prestige in the workplace, accumulation of wealth and possessions, and meaningful relationships from within the midst of three or four generations provide the potential for a life overflowing with richness and complexity. The mature life is one in which the triumvirate of human experience—love, work, and play—are successfully balanced to bring true fulfillment.

Unfortunately, the joys of midlife do not last forever. Old age lies ahead. Although the hope and statistical expectation is for many years of mental competence and independence, physical and mental decline, increased dependence, and eventually, death must be anticipated. Late adulthood has its own great pleasure, when there is a focus on continued mental and physical activity, a dominant preoccupation with the present and the future, and involvement with and facilitation of the young. Then, death can be met with feelings of satisfaction and acceptance, the natural end point of

human existence that follows a life lived and well loved.

WISDOM AND FULFILLMENT IN OLD AGE

When individuals have lived into their 70s, 80s and beyond, they have an ability to understand the human condition with a clarity and depth that is reserved for those who are living at the outer limits of the human life span as it exists today. Having experienced many of the joys and tribulations that are an inevitable part of human existence, their wisdom allows them to define the essential components of happiness and fulfillment as they approach the end of life (Fig. 56–1).

Physical and Mental Health

Human happiness and fulfillment are built on a solid foundation of physical and mental health. Physical health increases the ability to function in the world in the relative absence of pain and impairment and most importantly, provides an internal environment in which the brain can function optimally. Human thought and emotion originate in the brain and are totally dependent on brain function for existence. The mind resides within a physical body that profoundly affects mental life and development throughout life as it matures in childhood and ages in adulthood. Thus, caring for the body is absolutely essential to the maintenance and enhancement of sentience, the essence of human existence. Humans live in their bodies. Humans' mental life is totally linked to their physical being, but humans ARE their thoughts and emotions.

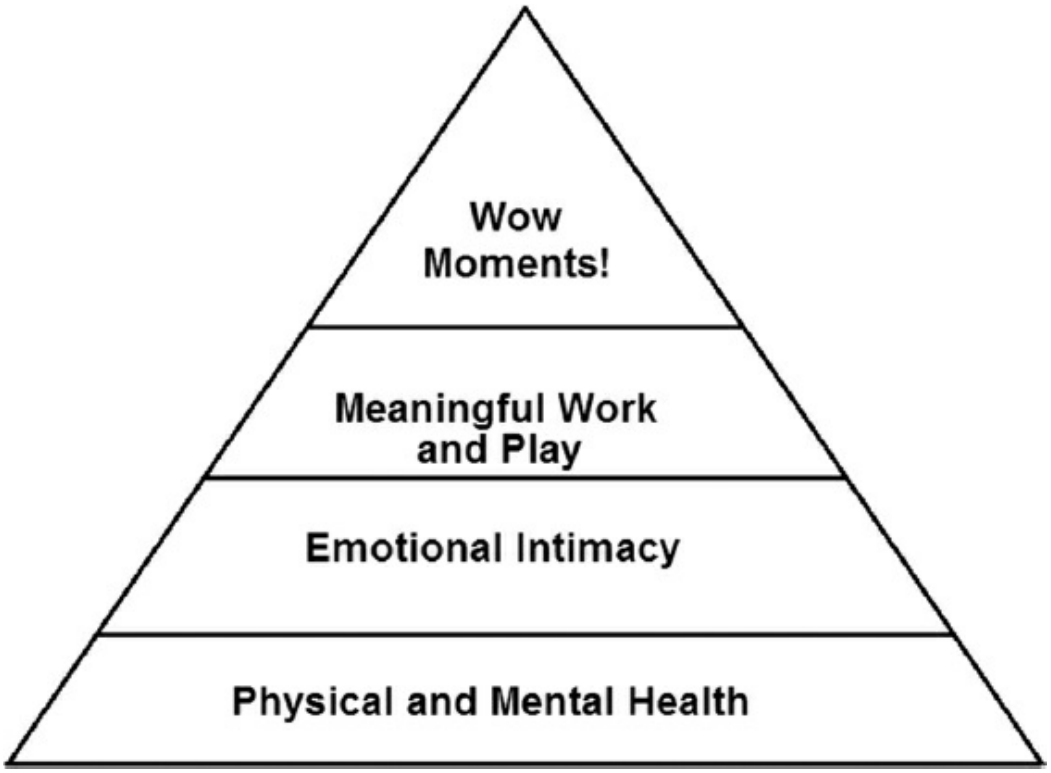

FIGURE 56–1. The four essential components of happiness and fulfillment.

Mental health, which may be roughly defined as the absence of significant mental illness—we are all neurotic to one degree or another—allows the individual to experience the inner world as a palace rather than a prison and to experience genuine intimacy in relationships with others. For those who are victims of biologically based mental illnesses, or who suffer from crippling and pleasure-robbing neuroses or worse forms of psychologically determined pathology, the quest for happiness and fulfillment may be difficult or impossible to achieve at any point in life.

Emotional Intimacy

We come into the world alone and leave it alone. But in the decades in between birth and death people live in a state of interdependence with others. When those relationships promote mental and emotional development, the result is emotional intimacy. Because people remain vulnerable to the attitudes and actions of those individuals with whom they are interdependent, these very same relationships can, and do, cause people to experience pain, rejection, and abandonment throughout their lives. There is nothing more exquisitely human than to be both emotionally enriched by those that one love and need, and at the same time continually vulnerable to the pain which their insensitive words and actions, whether intended or not, may cause. Living on this razor edge of pain and pleasure is lifelong and experienced differently as people mature and age. Consider the following:

The rapturous fusion of the infant with a loving mother or the total disintegrating despair caused by maternal rejection or abandonment.

The toddler's elation as experienced within a loving, mother–child, symbiotic dyad, or the terrifying experience of angry voices or physical abandonment during a true temper tantrum.

The empowering grandiosity, when accepted benignly, of the oedipal-age child's challenge and the verbal or physical retaliation of the insecure parent.

The enormously satisfying feeling of acceptance by the latency-age peer group and the cruel rejection of the child who is thought to be different.

The dream-like perfection of the first adolescent infatuation versus the total despair of the inevitable first rejection.

The joy of sharing the miracle of parenthood with a spouse and the crushing pain of losing a child due to illness, divorce, or death.

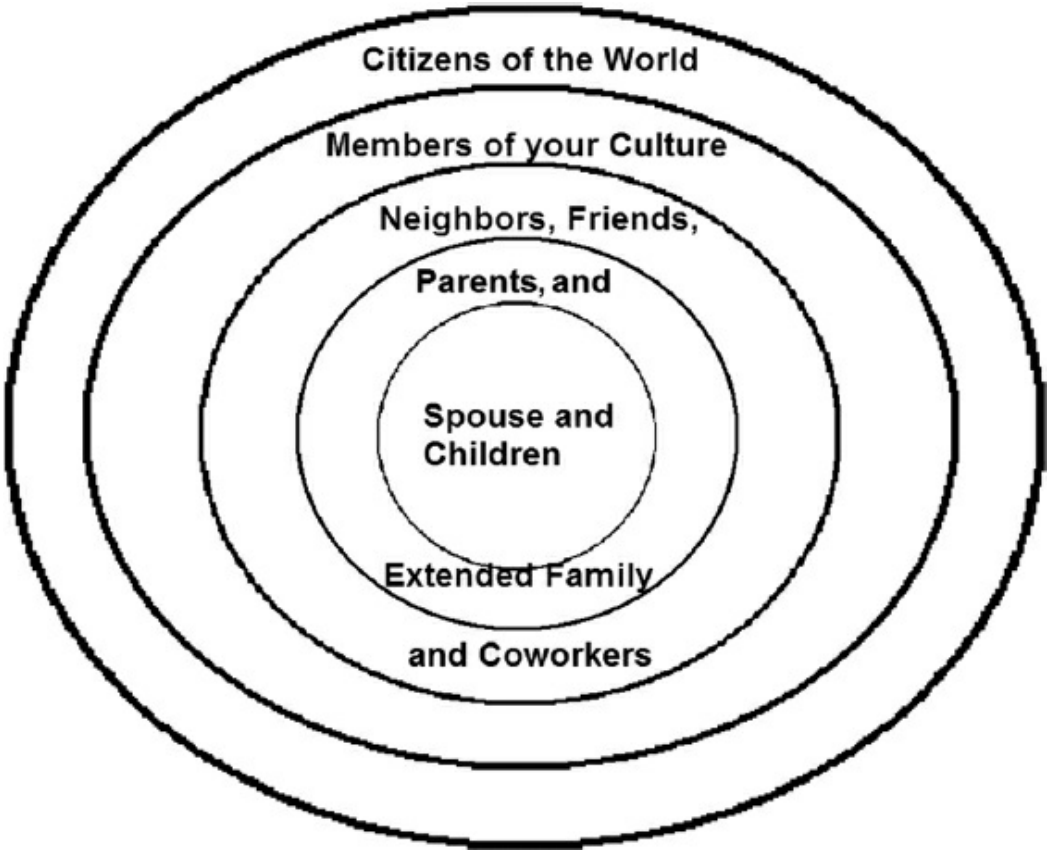

FIGURE 56–2. Relationships.

The tender intimacy of caring for aging parents and the sadness experienced when they die.

The joyous generativity experienced when giving and relating to younger generations and the jarring pain associated with the realization that all will be lost through death.

In addition to organizing people's lives, relationships that lead to intimacy, stability, and fulfillment occur in varying degrees of intensity. The nature of these relationships may be conceptualized in the form of concentric circles. At the center of the bull's-eye is involvement with spouse and children. In the next concentric ring are relationships with parents and extended family, followed in the third concentric ring by interactions with friends, coworkers, and neighbors. As people get further away from the warmth and sustenance provided by those who love and care about them, they are defined, and to a degree sustained, by the realization that they share all aspects of human experience with members of their particular nation and culture and the seven billion others who currently inhabit this planet (Fig. 56–2).

Work

Meaningful work is presented here as an essential ingredient which is necessary for happiness and fulfillment throughout adulthood and

increasingly in late adulthood. For many adults, including those in their 70s, 80s and beyond, work is a good thing, not something to be minimized or avoided.

The capacity for work begins to develop in the second and third years of life as parents introduce toilet training and other limit-setting experiences. As parents repeatedly say "No!" they frustrate the toddler's wish for total freedom, replace chaos with order, and impose routines and restrictions; they help the child learn to begin to tolerate frustration, delay gratification, and establish the capacity to control mind and body—essential components of a mature capacity to work and of happiness.

As they interact with their child around his or her "jobbies" (stools) and gently insist on adherence to societal demands for bowel and bladder control and cleanliness, parents facilitate the emergence of a core of strength which forms the basis on which the ability to work is built. Character emerges out of toilet training and other limit-setting processes. As Anna Freud described it, "In short, what takes place in this period is the far-reaching modification and transformation of the {personality} which—if kept within normal limits—supply the individual personality with a background of highly desirable, valuable qualities." (1965, p. 77).

During the oedipal years ambition and initiative, then directed primarily toward parental competitors, are integrated into the future worker's armamentarium. To quote Erikson, "while autonomy concentrates on keeping potential rivals out, and therefore, can lead to jealous rage most often directed against encroachments by younger sibling, initiative brings with it anticipatory rivalry with those who have been there first and may, therefore, occupy with their superior equipment, the field toward which one's initiative is directed." (1963, p. 255). Success in the workplace is predicated on the ability to face potential rivals and challenges, to work with and learn from those with "superior equipment," and to direct one's initiative beyond the limits imposed by current-day procedures and knowledge. The child, who successfully traverses the oedipal stage and its developmental tasks, is at the same time internalizing the capacities and characteristics required for success and fulfillment in the workplace.

Latency is the first phase of development in which children in the Western society are expected to work. Their form of work is contained in the "work," in the word homework. In order to be a successful worker-learner, many capabilities must be present to meet the demands of the school-workplace. These capacities also form the nucleus of the skills required by the adult worker, including the ability to comfortably control the body and utilize its potential; to modulate gratification and frustration and remain focused on a task for increasingly long intervals; to possess intellectual skills, and to relate to peers and superiors in a manner which facilitates accomplishment and productivity. All of these characteristics emerge in the elementary school classroom. Erikson recognized the

importance placed on work as a future expectation when he described the developmental dilemma of latency as industry versus inferiority (1963, p. 255).

When the latency-age child succeeds in the classroom, the newly internalized superego becomes "loving and beloved" (Schafer, 1960) in response to work and achievement. Such an individual is self-motivated, focused, and goal oriented, continually responding to a powerful inner compass that directs him or her toward higher levels of achievement. These are the qualities which, 20 years later, produce outstanding academic achievement, new discoveries, the guts to begin a new business, and satisfaction in the workplace.

Adolescence is the first phase of development in which work is performed in actual jobs in the economy for compensation. The desire to partake of the pleasures of adolescent/adult life, necessity, and parental insistence push adolescents into the workforce where, for the most part, they earn money and learn about the demands and expectations connected with work. When necessity requires that adolescents work, they quickly learn about the demands and limitations of employment and are usually forced into low-paying jobs that provide valuable experience but also detract from time which could be more profitably (in every sense of that word) spent on studies, preparing for the future when educational credentials literally translate into dollars, a high standard of living, and meaningful, fulfilling work. Whatever the reason may be that causes adolescents to enter the work force, by midadolescence, or earlier, they have developed the physical, emotional, and cognitive capacities necessary to perform the same work functions as adults.

Work becomes a central, often dominant, activity in young adulthood. When the childhood developmental line leading to adult work is relatively conflict free, there is a relatively smooth transition from high school to on-the-job training or college and beyond. The transition from learning and play to adult work may be gradual or abrupt, but at some point in the late teens or twenties the realization gradually dawns that the pleasure of play and the luxurious life of the student, dripping with freedom and free time, must be subordinated to the temporal and emotional demands of work.

Work becomes a fulfilling activity when the desire to learn and to succeed is driven by healthy curiosity and ambition. Abundant self-esteem is the reward for advancement in the work place and the gradual movement toward self-sufficiency. In addition, friendships are formed with others who are engaged in the same or similar work. Children play together. Young adults work together and in the process become friends.

As they meet, fall in love, and intertwine their lives, young adults begin to work for each other. The gratification derived from work is enormously intensified when children appear. With their presence, mothers and fathers double their efforts to provide a bountiful present and secure future for

their offspring.

While all of this is occurring, slowly but surely, work begins to permeate the deepest recesses of the psyche, becoming an integral part of one's identity, often a source of pleasure, even of security because of the nurturing aspects of its familiar routines. For some lucky ones there are moments, even days, when work becomes a *preferred* activity, replacing play and even relationships as a significant source of pleasure. Love, work, and play—as young adulthood progresses, work becomes elevated to equal status with the other two, completing the triumvirate of activities and relationships which produce lasting adult happiness and fulfillment.

Midlife workers are at the top of the heap. Capable, accomplished, and powerful, they are reaping the rewards of years of effort during which skills were mastered and seniority earned. For these individuals the narcissistic gratifications emanating from work may be considerable. A stable work situation may compensate for the painful realities of life related to aging, parental death, the turmoil surrounding having adolescent and young adult children, and the crush of responsibility related to being in the midst of three or four generations. On occasion, work may become the main source of emotional sustenance and stability, a refuge from doubts, debt, and difficult relationships.

Relationships with younger coworkers stir up considerable ambivalence because of the recognition that educating the next generation is the means of one's replacement. For those with self-confidence and sources of fulfillment outside the workplace, the anger at and envy of promising subordinates is not acted on to any significant degree. Instead, the fantasies and feelings are recognized and processed at a mental level, often sublimated into *generativity,* Erik Erikson's term for caring for the next generation.

Thoughts of retirement are frequently blotted out by those at the peak of their capabilities, but prospects for happiness in later life are greatly enhanced by the ability to take that searing look into the future, to anticipate, to plan. Retirement is clearly not for everyone. For some it does bring true happiness and fulfillment. For others, including many in the mental health field, work remains a sustaining activity well into late adulthood.

Play

Playing peekaboo, hopscotch or football, listening to rock and roll, Sondheim or Sebelius; reading romance novels or watching a Shakespearean play—what do they all have in common? All are expressions of one of the most ubiquitous and intriguing human activities—play. Psychoanalytic theory postulates that play is not a random, carefree action, free from the psychic restraints of more mundane human pursuits. Like all other thought and behavior, it is molded by the forces of the mind and the

environment into nearly endless forms which fascinate people from shortly after birth until the end of life and bring much pleasure, happiness, and relief of tension; thus its inclusion, along with work, as one of the four essentials for happiness and fulfillment.

In a brilliant study of children's play, psychoanalyst, Robert Waelder (1932), determined that children play so incessantly because they have to, not because they want to; because their minds are not fully developed and thus they are not easily able to understand or integrate the multitude of new ideas, situations, and various forms of stimulation which they encounter every day. Thus, children live in a state of "traumatic stimulation." In their play they attempt to master those experiences that are too overwhelming to be integrated in one fell swoop. Two decades later psychoanalyst Lilli Peller (1954) expanded Waelder's ideas, relating the particular forms of children's play to phase-specific developmental challenges. In other words, play will be different at different ages because of gradually increasing levels of mental sophistication and the need to address new challenges that arise as the developmental stages of childhood are traversed.

But what of adults? Better able to meet the challenges of everyday living, they have less need to play but play for the same reasons that children do; namely to deal with the stresses, conflicts, and developmental challenges of adult life. Psychoanalyst Mortimer Ostow (1987) suggested that play provides, not for the unrestrained pursuit of pleasure, but rather for an exposure to realistic or realistic-like challenges, the overcoming of which relaxes tension and replaces it with pleasure. By providing a mechanism for the reduction of stress and the mastery of traumatic experience, play becomes a means of pushing the individual toward new levels of mental and physical competence. Play in adulthood, like that in childhood, consists of both thought and action, and provides both pleasure and fulfillment because it is free of significant negative consequences and can be terminated at will.

In a 1993 paper on the nature of play in adulthood, Colarusso described four levels of play, covering the life cycle from infancy through adulthood. Only levels III and IV are relevant to the consideration of adulthood. Level III play, which could be subtitled *Learning the Games of Your Culture,* occurs for the first time during the latency phase of development (ages 6 to 11 years) because the latency-age child is increasingly capable of a variety of complicated motor skills, has the ability to read and write, has independent peer relationships, and has the capacity to adhere consistently to rules and tolerate frustration. Once the capacity for level III play emerges in latency, it continues to be used throughout life. Although the games of any culture continue to be played throughout life, the underlying developmental themes and conflicts that energize the play change dramatically from developmental phase to developmental phase. When adults around the

world watch football, as billions did during the recent World Cup, they are reacting to a form of play first learned in latency.

But for many adults, particularly older ones, happiness and fulfillment are frequently experienced in level IV play, mental play, the form used continually throughout adult lives to bring the relief of tension, momentary pleasure, and the false illusion that people have control of this life and their place in it.

Watching and Listening: Spectator Play

When Grandpa takes his Grandson to a soccer match they both cheer every goal, both hard at work-resolving developmental conflicts. Through identification with the players on the field, freckle-faced grandson is dealing with latency-age phase-specific issues such as competitiveness and sexual identity. At the same time, 65-year-old Grandpa, also through identification with the players on the field, is mourning for his lost body of youth and coming to terms with success or failure in his chosen fields of endeavor, now that he no longer has the abundant future still available to the younger man playing in the match. Nurtured by their play and each other, when the game is over, the score soon forgotten, both Grandson and Grandpa return to the "real" world.

Listening to music is another extremely common form of play. Highly evocative and capable of stimulating a wide variety of sexual and aggressive fantasies, its ethereal, nontangible nature is consistent with the abstract nature of level IV play. From hard rock to Haydn, the endless variety of musical expression enhances its developmental usefulness to individuals of all ages. Great composers, such as Richard Strauss in his "Four Last Songs," have used music to confront the inevitability of death.

Writing a novel or a play is another form of play that brings great pleasure to both author and reader, as together they create an imaginary world of fantasy. Stimulated by the author's creation, the reader can revel, without fear of discovery or recrimination, in whatever sexual or aggressive fantasies he or she wishes. Unlike the child who is not primarily interested in communication, the author must take into consideration the effect of his or her creation on the audience, filtering fantasies through the organizing and critical influences of the ego and superego. By so doing, the author blends play with work, a blurring of boundaries that occurs frequently in adulthood. On the other hand, the playgoers or readers, who have no responsibility for the creative work, are closer to pure play as they sit in their seats, keep their mental and emotional responses to themselves, and experience a brief respite from the relentless pressures of work and reality.

Thus, it is clear that work and play are essential activities that help grown-ups deal with the huge responsibilities of adulthood and experience pleasure in the process. Work provides the means of caring for one's self and loved ones. Play provides a safety valve from the pressure cooker of

responsibility and furnishes an outlet for the powerful emotions generated therein. In the world of play and make-believe, time stands still, fantasy dominates reality, and the player is in control of life and death, if only for a moment or two.

Wow Moments!

And now to the last component, and certainly the most purely enjoyable one of all the ingredients for happiness and fulfillment, the term Wow Moments! is an attempt to describe in words the feeling of elation that accompanies the awareness, which hopefully occurs many times each day, that life is often truly wonderful, endlessly interesting, and full of beauty. Such moments of elation are obviously highly individual and vary enormously from person to person.

Paradoxically as it may seem, the ability to appreciate how stunningly wonderful life can be is based on the realization that nothing lasts forever, that change is a constant in life, and that the sense of elation experienced during Wow Moments! is transitory, like life itself. Although such moments can occur in childhood as well as adulthood, the awareness of the transient nature of all things, and therefore of Wow Moments! is experienced most fully in middle and late adulthood.

Freud's comments on the subject are particularly relevant to this discussion. In his paper "On Transience" (1916 [1915]) written during the First World War, Freud's attention was drawn to thoughts about loss and impermanence. Speaking about a friend who was a poet, he noted: "The poet admired the beauty of the scene around us but felt no joy in it. He was disturbed by the thought that all this beauty was fated to extinction, that it would vanish when winter came, like all human beauty and all the beauty and splendour that men have created or may create. All that he would otherwise have loved and admired seemed to him to be shorn of its worth by the transience which was its doom." (p. 305).

It is precisely this gradually growing, adult awareness that time is limited and that all things and all beings will eventually be destroyed, or die, that makes Wow Moments! so wonderful, so full of elation—and so transient.

Consider the following examples of Wow Moments!

The sense of exhilaration while taking the first breath of pure air on a sunlit morning beneath a cloudless sky.

The sense of physical, mental, and psychological fusion experienced by an elderly couple who have shared their lives for five or more decades.

The sense of deep satisfaction experienced by a therapist when after several years of therapy, a young adult patient who was terrified of commitment because of a childhood filled with abuse and abandonment was able to overcome his fear of intimacy and ask his girlfriend to marry him.

The feeling of physical intactness and competence experienced by those in their seventies and eighties after a hard workout in the gym or after watching an all too infrequent perfect iron shot land on the middle of the green on the golf course.

The sense of awe created while observing two loving parents care for their young children; the almost palpable experience of watching the next generation being created within the crucible of the nuclear family.

The transcendental joy of listening to the last movement of Mahler's second symphony or Mozart's clarinet concerto or Gershwin's rhapsody in blue.

The thrill of attending a World Cup and sharing the experience with more than a billion other human beings, becoming in that moment a true citizen of the world.

The tingling up and down of one's spine when listening to the Beatles or Soprano, or Marilyn Horne sing Donizetti.

The sense of wonder and connection with past generations and cultures while visiting Machu Picchu, Easter Island, Egypt, or Angkor Wat.

The recognition of individual insignificance when gazing at the heavens on a clear night while realizing at the same time, that in some incomprehensible way, every human being is an integral part of it all.

FUTURE DIRECTIONS

As more and more men and women are living into their 80s, 90s, and even to 100 and beyond, the frontier for the study of adulthood lies in understanding the nature of the developmental processes in late (60 to 80 years) and late late (80 to 100 plus years) adulthood, and the effect of the longer life span on their children and grandchildren. For example, caring for a very elderly parent when the "child" is in his or her 70s or 80s can stretch out the time spent engaging the development task of caring for parents to 30 or 40 years. What effect does that interaction have on the ability to the elderly offspring to focus primarily on their own engagement of the developmental tasks of late adulthood and relationships with his or her spouse, children, and grandchildren?

For example, one patient in her late 60s was struggling to deal with the care of her 103-year-old mother. Although still mentally sharp, if somewhat forgetful, mother had numerous physical problems such as pneumonia and hip surgery after a fall. The husband of the care-taking daughter was becoming deeply resentful of the time, energy, and money that were being expended to care for his mother-in-law. He was not in good health himself and wanted his wife to be more caring and available for him and for them as a couple. In addition, the couple's desire to spend time with their children and grandchildren was compromised. Being caught in the middle of generations (i.e., the "sandwich" generation) for 10 or 20 years, as was the case for most families in the past, has expanded in this family into a

fourth decade.

As the life span lengthens, researchers and clinicians are confronted with new challenges to provide patients and their families with guidelines for this new era in human development.

REFERENCES

Aronson P. The markers and meanings of growing up: contemporary young women's transition from adolescence to adulthood. *Gend Soc*. 2008;22:56–82.

Block J. *Lives Through Time*. Berkeley, CA: Bancroft; 1971.

*Colarusso CA. *Child and Adult Development: A Psychoanalytic Introduction for Clinicians*. New York: Plenum Press; 1992.

Colarusso CA. Play in adulthood: A Developmental consideration. *Psychoanal Study Child*. 1993;48:225–245.

Colarusso CA. Separation-individuation processes in middle adulthood: the fourth individuation. In: Akhtar S, Kramer S, eds. *The Seasons of Life: Separation-Individuation Perspectives*. Northvale, NJ: Jason Aronson; 1997:73–94.

Colarusso CA. A developmental line of time sense. In late adulthood and throughout the life cycle. *Psychoanal Study Child*. 1998;53:113–140.

Colarusso CA. A child-analytic case report: a 17-year follow-up. In: Cohen J, Cohler BJ, eds. *The Psychoanalytic Study of Lives over Time*. San Diego, CA: Academic Press; 2000:17.

*Colarusso CA. Separation-individuation phenomena in adulthood: general concepts and the fifth individuation. *J Am Psychoanal Assoc*. 2000;48:1467–1489.

Colarusso CA. The absence of a future: the effect of past experience and current developmental conflicts on a midlife analysis. *J Am Psychoanal Assoc*. 2006;54:919–943.

Dederick JG, Miller HL. Transitions into adulthood: are they the same for women and for men? In: Wainrib BR, ed. *Gender Issues Across the Life Cycle*. New York: Springer; 1992:44–55.

Diamond MJ. Masculinity unraveled: the roots of male gender identity and the shifting of male ego ideals throughout life. *J Am Psychoanal Assoc*. 2006;54:1099–1130.

Dworsky A, Courtney M. Addressing the mental health service needs of foster youth during the transition to adulthood: how big is the problem and what can states do? *J Adolesc Health*. 2009;44(1):1–2.

Erikson EH. *Childhood and Society*. 2nd ed. New York: Norton; 1963.

Freud A. *Normality and Pathology in Childhood: Assessments of Development*. New York: International Universities Press; 1965, 77.

Freud S. *On Transience*. Standard Edition, vol. 14. London: Hogarth Press; (1916 [1915]):73–102.

Gilmore K. Psychoanalytic development theory: a contemporary reconsideration. *J Am Psychanal Assoc*. 2008;56(3):885–907.

*Guttmann DL. The country of old men: cross-cultural studies in the psychology of later life. In: *Occasional Papers in Gerontology*. 5th ed. Ann Arbor, MI: Institute of Gerontology, University of Michigan Press; 1969.

Kimmel DC. Adult development and aging: a gay perspective. In: Garnets LD, Kimmel DC, eds. *Psychological Perspectives on Lesbian and Gay Male Experiences*. New York: Columbia University Press; 1993:517–534.

*Levinson DJ, Darrow CN, Klein EB, Levinson MH, McKee B. *The Seasons of a Man's Life*. New York: Knopf; 1978.

Lyketsos CG, Chen LS, Anthony JC. Cognitive decline in adulthood: an 11.5-year follow-up of the Baltimore Epidemiologic Catchment Area study. *Am J Psychiatry*. 1999;156:58–65.

Mashour GA. Toward a general theory of unconscious processes in psychoanalysis and anesthesiology. *J Am Psychoanal Assoc*. 2008;56(1):203–222.

Nemiroff RA, Colarusso CA. *The Race against Time: Psychotherapy and Psychoanalysis in the Second Half of Life*. New York: Plenum Press; 1985.

Neugarten BL, Berkowitz H, Crotty WJ, et al. *Personality in Middle and Late Life*. New York: Atherton; 1964.

Notman MT. When a husband dies. In: *The Race Against Time: Psychotherapy and Psychoanalysis in the Second Half of Life*. New York: Plenum Press; 1985.

Ostow M. Play and reality. *Psychoanal Study Child*. 1987;42:193–203.

Peller LE. Libidinal phases, ego development, and play. *Psychoanal Study Child*. 1954;9:178–198.

Person ES. Masculinities, plural. *J Am Psychoanal Assoc*. 2006;54:1165–1186.

Plotkin F. Treatment of the older adult: the impact on the psychoanalyst. *J Am Psychoanal Assoc*. 2000;48:1591–1616.

Roberts BW, Mroczek D. Personality trait change in adulthood. *Curr Dir Psychol Sci*. 2008;17(1):31–35.

Rusbridger R. The internal world of Don Giovanni. *Int J Psychoanal*. 2008;89:181–194.

Schafer R. The loving and beloved superego in Freud's structural theory. *Psychoanal Study Child*. 1960;15:163–188.

Scott WR. Approaching adulthood: the maturing of institutional theory. *Theory Soc*. 2008;37(5):427–442.

Stevens-Long J. Adult development: theories past and present. In: Nemiroff RA, Colarusso CA, eds. *New Dimensions in Adult Development*. New York: Basic Books; 1990:125.

Vaillant GE. *Adaptation to Life*. Boston, MA: Little Brown; 1977.

Vaillant GE, Vaillant CO. Natural history of male psychological health, XII: a 45-year study of predictors of successful aging at age 65. *Am J Psychiatry*. 1990;147:31–37.

Valenstein AF. The older patient in psychoanalysis. *J Am Psychoanal Assoc*. 2000;48:1563–1589.

van Gennep A. *The Rites of Passage*. Chicago: The University of Chicago Press; 1960.

van Haren NE, Hulshoff Pol HE, Schnack HG, et al. Progressive brain volume loss in schizophrenia over the course of the illness: evidence of maturational abnormalities in early adulthood. *Biol Psychiatry*. 2008;63(1):106–113.

Waelder R. (1932). The psychoanalytic theory of play. In: Guttman SA, ed. *Psychoanalysis*. New York: International Universities Press; 1976:84–100.

Wallerstein JS. The psychological tasks of marriage: part 2. *Am J Orthopsychiatry*. 1996;66:217–227.

57 Geriatric Psychiatry

▲ 57.1 Overview

57.1a Introduction to Geriatric Psychiatry

DILIP V. JESTE, M.D.

AGING OF THE POPULATION

The population of the United States is aging rapidly, with both an increased number and increased proportion of older adults. The number of Americans aged 65 and over has increased from 3 million (4 percent of the total population) in 1900 to 43.1 million (13 percent of the population) in 2012. Yet, this historical rise will be dwarfed by that which has started in the year 2011 when the first of the baby boomers (those born between the years 1946 and 1964) reach the age of 65—the traditional definition of "old age." The number of people over 65 will rise to 72 million in 2030 and 84 million (20 percent of the population) by 2050 (Fig. 57.1a–1).

The average life span in the United States has increased from 47 years in 1900 to 79 years today. The more optimistic estimates project an increase in life expectancy at birth from the current figures of 76 years for men and 81 years for women to 86 years for men and 92 years for women by 2050. Men and women who are 65 years old today are expected to live an average of 16 and 19 years, respectively. By 2050, the average life spans at age 65 would have increased to 25 and 30 years, respectively. Thus, reaching age 65 would not indicate imminent death, but rather many more years of life. One of the unwelcome scenarios that could torpedo such an age wave would be a continuation of the current epidemic of obesity in younger and middle-aged individuals, leading to rising incidence of diabetes and other metabolic and vascular diseases that may shorten life spans.

The fastest growing segment of the population in the United States is that of individuals over the age of 85. The 85-and-over age group is expected to increase from about 6 million today to more than 18 million in 2050. Also striking is the growth of the centenarian population. While

there were about 5,000 centenarians in the United States in 1900, today there are around 53,000 people in that age group in the United States, or 1.73 centenarians per 10,000 people in the total population. By 2050, that number will have gone up to about 835,000. A particularly notable fact is that whereas most of the centenarians in the United States today are white, nearly half of those in 2050 will belong to ethnic minority groups.

The aging of the population is not restricted to the United States, but is a global phenomenon. The world is getting older at a pace that has taken demographers by surprise and presented governments with economic and social challenges. Whereas there were about 605 million people over age 60 in the year 2000, that number will reach 2 billion in the year 2050. Two-thirds of all the people in the entire history of the world who have reached the age of 65 are alive today. The number of centenarians in the world will increase from approximately 316,000 at the present time to an incredible 2.2 million by 2050. To put the latter number in perspective, presently the entire population of countries such as Kuwait, Latvia, Macedonia, and the United Arab Emirates is roughly 2 million people each.

The rates of aging of the population vary widely by region. In some developed countries today, the proportion of older persons is close to 20 percent, that is, 1 in 5. During the next several decades, that proportion will reach 1 in 4 and in a few countries as high as 1 in 2 (i.e., 50 percent). More highly developed countries have more stable populations and will experience proportionately less growth of their older populations than developing countries that are undergoing a rapid increase in the number of aged individuals. Figure 57.1a–2 illustrates the changing age structure of the population in more developed versus developing countries from 2005 to 2050.

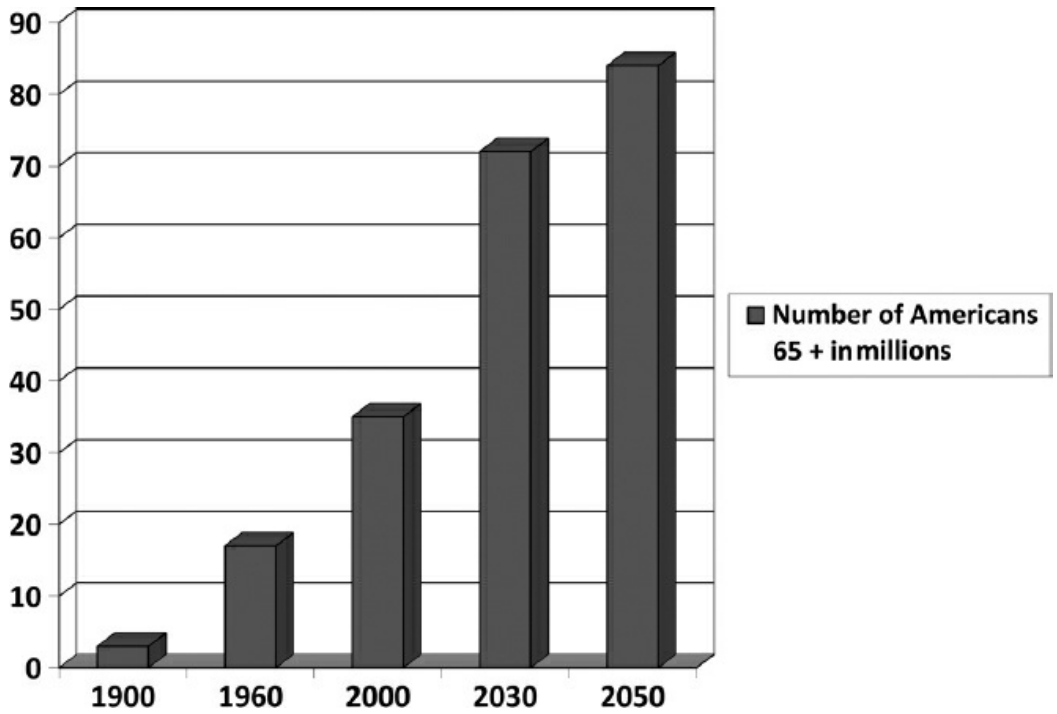

FIGURE 57.1a–1. Number of Americans 65+ (in millions). (U.S. Census Bureau; 2014.)

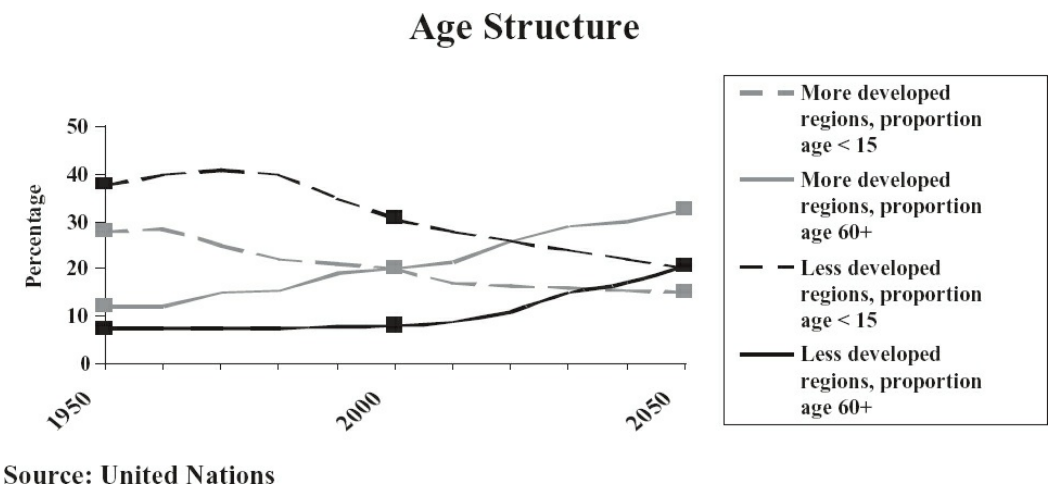

FIGURE 57.1a–2. Source: United Nations, Department of Economic and Social Affairs, Population Division.

The definitions of middle age and old age, as well as some of the traditional thinking about chronological aging, may need to be revised. It has been suggested that middle age be defined as age between 51 and 70 years, "young old" between 71 and 90 years, and "exceptionally old" over age 90. With increasing longevity, the timetable for brain aging may have changed as well.

The growth of the older population during the 20th century is attributable to several factors. During the first half of that century, the main factor responsible for the demographic change was a progressive and dramatic decrease in perinatal and infant mortality due to better treatments of infections, improved nutrition, and better prenatal and

postnatal care. In the latter half of the last century, the greater longevity resulted primarily from an increase in average life span, that is, adults lived longer as a result of declines in mortality from pneumonia, heart disease, and other chronic diseases.

Aging is associated with increased heterogeneity. Tables 57.1a–1 to 57.1a–6 show number of people over age 55 in the United States by age and gender, by marital status, by educational attainment, by earnings of full-time workers, by types of household, and by monetary income. The wide range of diversity in socioeconomic status in older adults is remarkable. Factors such as education, marital status, and income have a potential bidirectional relationship with a number of mental illnesses. For example, adults with schizophrenia are likely to have lower levels of education, be unmarried, and have lower income than the general population. These variables could be considered potential risk factors for the illness; it is more likely, however, that these factors are secondary to the development of a serious mental illness. In either case, they are of considerable importance in management of persons with schizophrenia.

WHAT IS NORMAL AGING?

"Normal" aging must be differentiated from disease. Some bodily functions decline with age, but health problems are not inevitable. Most diseases of old age such as heart disease, cancer, and dementia result from pathological processes rather than from normal aging. Disabilities resulting from diseases can usually be treated and sometimes prevented. For example, most hearing problems can be at least partially corrected with hearing aids while the risk of heart disease can be reduced by healthy habits with respect to diet and exercise. The same applies to mental health and many of the mental illnesses in older people.

One point to stress is that aging is associated with more, and not less, diversity. It is a common misconception that older adults look, think, and behave alike. Indeed, there are more differences among older people than among younger adults. Even in the same person, different organs "age" at different rates. It has been suggested that the notion of the chronological age ("how old are you?") should be abandoned, and instead, the stages of aging that a given individual exhibits in different areas of functioning should be considered. Furthermore, just as a child may be at different stages of development mentally and physically, an older adult may have divergent aging of his or her brain, liver, kidneys, and heart. As a matter of fact, not even all the parts of the brain age at the same rate in a person. Thus, an aged person may have relatively intact memory while having deficits in motor function.

It must be recognized that although age cutoffs are needed for defining old age for societal purposes (e.g., qualifying age for starting Medicare or for Social Security), such cutoffs are artificial and arbitrary. No magical

change occurs overnight on the 65th or any other birthday. Also, with increasing longevity, the age cutoffs used currently will need to change.

There is often a tendency to confuse the effects of old age with those of its common, but not inevitable, associations. Thus, medical morbidity increases with aging, but many physical diseases can be managed better, and some (e.g., nutritional deficiencies) can be prevented. The diverse treatment of various physical illnesses by different specialists resulting in polypharmacy has an adverse influence on the treatment of mental disorders in the older adults. Similarly, psychosocial stressors such as financial hardship and lack of appropriate community services impact management of older people with psychiatric symptoms.

Table 57.1a–1.
Population by Age and Gender in the US: 2012 (Numbers in Thousands. Civilian Noninstitutionalized Population[a])

	Both Genders		Male		Female	
Age	**Number**	**Percent**	**Number**	**Percent**	**Number**	**Percent**
All ages	308,827	100.0	151,175	100.0	157,653	100.0
Under 55 years	229,349	74.3	114,685	75.9	114,664	72.7
55–59 years	20,470	6.6	9,879	6.5	10,592	6.7
60–64 years	17,501	5.7	8,278	5.5	9,223	5.8
65–69 years	13,599	4.4	6,461	4.3	7,139	4.5
70–74 years	9,784	3.2	4,519	3.0	5,265	3.3
75–79 years	7,331	2.4	3,122	2.1	4,209	2.7
80–84 years	5,786	1.9	2,421	1.6	3,365	2.1
85 years and over	5,006	1.6	1,809	1.2	3,196	2.0
Under 55 years	229,349	74.3	114,685	75.9	114,664	72.7
55 years and over	79,478	25.7	36,489	24.1	42,989	27.3
Under 60 years	249,820	80.9	124,564	82.4	125,256	79.5
60 years and over	59,007	19.1	26,611	17.6	32,397	20.5
Under 62 years	257,164	83.3	128,100	84.7	129,064	81.9
62 years and over	51,664	16.7	23,074	15.3	28,589	18.1
Under 65 years	267,320	86.6	132,842	87.9	134,478	85.3
65 years and over	41,507	13.4	18,332	12.1	23,174	14.7
Under 75 years	290,704	94.1	143,822	95.1	146,882	93.2
75 years and over	18,123	5.9	7,353	4.9	10,771	6.8

[a]Plus armed forces living off post or with their families on post.

From U.S. Census Bureau, Current Population Survey. *Annual Social and Economic Supplement.* Washington, DC; 2012. Internet release date: December 2013.

Table 57.1a–2.

Marital Status of the Population 55 Years and Over by Gender and Age in the United States: 2012 (Numbers in Thousands. Civilian Noninstitutionalized Population[a])

Gender and Marital Status	Total 55 Years and Over		55–59 Years		60–64 Years		65–74 Y
	Number	Percent	Number	Percent	Number	Percent	Number
Both genders	79,478	100.0	20,470	100.0	17,501	100.0	23,383
Married	49,255	62.0	13,649	66.7	11,542	66.0	15,317
Spouse present	48,170	60.6	13,348	65.2	11,271	64.4	15,073
Spouse absent	1,085	1.4	301	1.5	270	1.5	244
Widowed	12,729	16.0	864	4.2	1,204	6.9	3,410
Divorced	10,935	13.8	3,507	17.1	2,911	16.6	3,213
Separated	1,298	1.6	511	2.5	343	2.0	313
Never married	5,261	6.6	1,939	9.5	1,502	8.6	1,130
Male	36,489	100.0	9,879	100.0	8,278	100.0	10,980
Married	26,240	71.9	6,828	69.1	5,939	71.7	8,325
Spouse present	25,711	70.5	6,684	67.7	5,811	70.2	8,199
Spouse absent	529	1.5	144	1.5	127	1.5	126
Widowed	2,561	7.0	204	2.1	242	2.9	690
Divorced	4,552	12.5	1,601	16.2	1,229	14.8	1,278
Separated	602	1.6	244	2.5	152	1.8	146
Never married	2,535	6.9	1,002	10.1	717	8.7	541
Female	42,989	100.0	10,592	100.0	9,223	100.0	12,404
Married	23,015	53.5	6,821	64.4	5,603	60.8	6,992
Spouse present	22,459	52.2	6,664	62.9	5,460	59.2	6,875
Spouse absent	556	1.3	157	1.5	143	1.5	117
Widowed	10,168	23.7	660	6.2	962	10.4	2,720
Divorced	6,383	14.8	1,906	18.0	1,681	18.2	1,936
Separated	696	1.6	267	2.5	191	2.1	168
Never married	2,726	6.3	937	8.9	785	8.5	588

[a]Plus armed forces living off post or with their families on post.

From U.S. Census Bureau, Current Population Survey. *Annual Social and Economic Supplement.* Washington, DC; 2012.

Internet release date: December 2013.

DEMOGRAPHICS OF OLDER PEOPLE WITH MENTAL ILLNESSES

Aging and aging-associated disorders (e.g., metabolic and vascular diseases) may influence expression and management of mental illnesses. Yet, psychiatric epidemiological studies have tended to exclude older adults. The common notion that psychiatric disorders are rare among older individuals is just plain wrong. This misconception, prevalent even among nongeriatric mental health experts, is based on studies such as the Epidemiologic Catchment Area (ECA) study conducted by the National Institute of Mental Health in the early 1980s. The ECA study was one of the largest epidemiological investigations of psychiatric disorders using, what was at that time, state-of-the-art methodology including the then-new DSM-III criteria. The ECA study reported that the prevalence of any psychiatric disorder (excluding dementia), either over the previous year or over the entire lifetime of the individuals, was considerably lower among older compared to that in younger adults (Table 57.1a–7). This decrease in the prevalence of mental illness with aging was thought to be due, in part, to higher mortality associated with serious mental illnesses. Yet, an important explanation for this ECA finding could be that psychiatric disorders were underdiagnosed among older adults for a variety of reasons.

Table 57.1a–3.
Educational Attainment of the Population 55 Years and Over by Gender and Age in the United States: 2012 (Numbers in Thousands. Civilian Noninstitutionalized Population[a])

Gender and Educational Attainment	Total 55 Years and Over		55–59 Years		60–64 Years		65–7
	Number	Percent	Number	Percent	Number	Percent	Number
Both genders	79,478	100.0	20,470	100.0	17,501	100.0	23,383
Less than 9th grade	5,322	6.7	796	3.9	799	4.6	1,567
9th–12th grade (no diploma)	6,344	8.0	1,296	6.3	935	5.3	2,115
High school graduate	26,531	33.4	6,546	32.0	5,227	29.9	8,067
Some college or associate's degree	19,343	24.3	5,505	26.9	5,005	28.6	5,300
Bachelor's degree	12,803	16.1	3,838	18.8	3,261	18.6	3,530
Advanced degree	9,134	11.5	2,489	12.2	2,274	13.0	2,805
Less than high school	11,666	14.7	2,092	10.2	1,734	9.9	3,682

graduate							
High school graduate or more	67,811	85.3	18,378	89.8	15,767	90.1	19,702
Less than bachelor's degree	57,541	72.4	14,143	69.1	11,966	68.4	17,048
Bachelor's degree or more	21,937	27.6	6,327	30.9	5,535	31.6	6,335
Male	36,489	100.0	9,879	100.0	8,278	100.0	10,980
Less than 9th grade	2,389	6.5	361	3.7	380	4.6	748
9th–12th grade (no diploma)	2,754	7.5	677	6.9	426	5.1	951
High school graduate	11,220	30.7	3,206	32.5	2,402	29.0	3,318
Some college or associate degree	8,574	23.5	2,500	25.3	2,194	26.5	2,454
Bachelor's degree	6,476	17.7	1,891	19.1	1,652	20.0	1,886
Advanced degree	5,076	13.9	1,243	12.6	1,224	14.8	1,623
Less than high school graduate	5,143	14.1	1,038	10.5	806	9.7	1,699
High school graduate or more	31,347	85.9	8,840	89.5	7,472	90.3	9,281
Less than bachelor's degree	24,937	68.3	6,744	68.3	5,402	65.3	7,471
Bachelor's degree or more	11,552	31.7	3,134	31.7	2,876	34.7	3,509
Female	42,989	100.0	10,592	100.0	9,223	100.0	12,404
Less than 9th grade	2,933	6.8	435	4.1	419	4.5	819
9th–12th grade (no diploma)	3,590	8.4	619	5.8	509	5.5	1,164
High school graduate	15,311	35.6	3,340	31.5	2,825	30.6	4,750
Some college or associate's degree	10,769	25.1	3,005	28.4	2,811	30.5	2,845

Bachelor's degree	6,327	14.7	1,947	18.4	1,609	17.4	1,644
Advanced degree	4,058	9.4	1,246	11.8	1,050	11.4	1,182
Less than high school graduate	6,524	15.2	1,054	10.0	928	10.1	1,983
High school graduate or more	36,465	84.8	9,537	90.0	8,295	89.9	10,421
Less than bachelor's degree	32,604	75.8	7,399	69.9	6,563	71.2	9,578
Bachelor's degree or more	10,384	24.2	3,193	30.1	2,659	28.8	2,826

[a]Plus armed forces living off post or with their families on post.

From U.S. Census Bureau, Current Population Survey. *Annual Social and Economic Supplement.* Washington, DC; 2012.

Internet release date: December 2013.

Table 57.1a–4. Earnings of Full-Time, Year-Round Workers 55 Years and Over by Gender and Age in the United States: 2012 (Numbers in Thousands. Civilian Noninstitutionalized Population[a])

Gender and Earnings[b]	Total 55 Years and Over		55–59 Years		60–64 Years		65 Years
	Number	Percent	Number	Percent	Number	Percent	Number
Both genders	21,985	100.0	10,945	100.0	6,897	100.0	4,143
Under $5,000	168	0.8	67	0.6	55	0.8	45
$5,000–$9,999	264	1.2	113	1.0	68	1.0	83
$10,000–$14,999	646	2.9	284	2.6	170	2.5	191
$15,000–$19,999	997	4.5	413	3.8	333	4.8	250
$20,000–$24,999	1,528	7.0	736	6.7	459	6.7	333
$25,000–$34,999	3,358	15.3	1,621	14.8	1,037	15.0	699
$35,000–$49,999	4,350	19.8	2,144	19.6	1,443	20.9	763
$50,000–$74,999	5,113	23.3	2,639	24.1	1,684	24.4	791
$75,000–$99,999	2,247	10.2	1,259	11.5	631	9.2	357

$100,000 and over	3,315	15.1	1,668	15.2	1,016	14.7	632
Male	12,437	100.0	6,098	100.0	3,780	100.0	2,558
Under $5,000	98	0.8	41	0.7	31	0.8	26
$5,000–$9,999	93	0.7	42	0.7	13	0.4	37
$10,000–$14,999	307	2.5	121	2.0	75	2.0	111
$15,000–$19,999	388	3.1	160	2.6	124	3.3	104
$20,000–$24,999	611	4.9	286	4.7	168	4.4	158
$25,000–$34,999	1,642	13.2	770	12.6	456	12.1	415
$35,000–$49,999	2,174	17.5	1,035	17.0	733	19.4	406
$50,000–$74,999	3,071	24.7	1,576	25.8	981	25.9	514
$75,000–$99,999	1,486	12.0	821	13.5	414	11.0	251
$100,000 and over	2,568	20.7	1,247	20.4	785	20.8	536
Female	9,548	100.0	4,847	100.0	3,117	100.0	1,585
Under $5,000	70	0.7	27	0.6	24	0.8	19
$5,000–$9,999	171	1.8	71	1.5	55	1.8	46
$10,000–$14,999	339	3.6	163	3.4	96	3.1	80
$15,000–$19,999	609	6.4	254	5.2	209	6.7	146
$20,000–$24,999	917	9.6	451	9.3	291	9.3	175
$25,000–$34,999	1,716	18.0	851	17.6	581	18.6	283
$35,000–$49,999	2,177	22.8	1,110	22.9	710	22.8	357
$50,000–$74,999	2,043	21.4	1,063	21.9	703	22.6	277
$75,000–$99,999	760	8.0	438	9.0	217	7.0	106
$100,000 and over	747	7.8	421	8.7	230	7.4	95

[a]Plus armed forces living off post or with their families on post.

[b]Earnings is the sum of wage and salary income and self-employment income.

From U.S. Census Bureau, Current Population Survey. *Annual Social and Economic Supplement.* Washington, DC; 2012. Internet release date: December 2013.

According to the ECA study, 13 percent of people over age 65 would meet the criteria for psychiatric disorders other than dementia. This is, however, likely to be a serious underestimate because of factors such as misattribution of psychiatric symptoms (by patients, caregivers, and raters) to other causes such as cognitive impairment, physical disorders, or "normal aging"; a lack of age-appropriate diagnostic criteria for major psychiatric disorders including substance abuse or dependence; and underreporting of psychological symptoms due to forgetfulness and social stigma. The "real" prevalence of psychiatric disorders other than dementia in older adults is probably at least 25 percent higher, that is, over 16 percent (Jeste et al., 1999). Additionally, over 10 percent of older adults have dementia, usually of the Alzheimer type; neuropsychiatric behavioral disturbances are common in patients with dementia, with at least a third of them (i.e., about 3 percent of the total older adult population) exhibiting psychotic and/or depressive symptoms. Thus, a total of nearly 20 percent of persons over age 65 would have diagnosable psychopathological symptoms.

Furthermore, psychiatric symptoms that do not meet the traditional criteria for psychiatric disorders (that had been developed for use in younger adults) but are nonetheless quite disabling are much more common in the older adults than the ECA study suggests. For example, a diagnosis of major depression requires the presence of at least five symptoms from a list of nine symptoms included in the DSM-5. If a person has four of those symptoms s/he would not meet the criteria for major depression, but may still be quite disabled. Such subsyndromal depression is much more common than major depression in the older adults; yet the patient may not receive the necessary therapeutic attention because the patient was not diagnosed with major depression. Table 57.1a–4 shows the prevalence of mental illnesses among younger (age 30 to 44 years) versus older (>65 years) adults.

Table 57.1a–5.
Households by Type and Age of Householder 55 Years and Over in the United States: 2012 (Numbers in Thousands. Civilian Noninstitutionalized Population[a])

	Total 55 Years and Over		55–59 Years		60–64 Years		6!
Household Type	Number	Percent	Number	Percent	Number	Percent	Numb
Total households	49,622	100.0	12,037	100.0	10,742	100.0	14,51
Family households[b]	29,333	59.1	8,044	66.8	6,857	63.8	8,924
Married couple	23,979	48.3	6,408	53.2	5,685	52.9	7,596
Male	1,369	2.8	489	4.1	307	2.9	317

householder[c]							
Female householder[c]	3,985	8.0	1,147	9.5	866	8.1	1,010
Nonfamily households	20,289	40.9	3,993	33.2	3,884	36.2	5,593
Male householder	7,514	15.1	2,032	16.9	1,705	15.9	2,027
Female householder	12,775	25.7	1,961	16.3	2,180	20.3	3,566

[a]Plus armed forces living off post or with their families on post.

[b]Households in which at least one member is related to the person who owns or rents the occupied housing unit (householder).

[c]No spouse present.

From U.S. Census Bureau, Current Population Survey. *Annual Social and Economic Supplement.* Washington, DC; 2012.

Internet release date: December 2013.

Table 57.1a–6. Monetary Income and Age of Householder 65 Years and Over in the United States: 2012

	65 Years and Over					
		65–74 Years				
Characteristic	Total 65 Years and Over	Total	65–69 Years	70–74 Years	75 Years and Over	Mean Age of Householder
Total	28,729	15,911	9,229	6,682	12,818	51.0
Under $5,000	625	326	200	126	299	47.0
$5,000–$9,999	1,439	590	326	264	849	53.0
$10,000–$14,999	2,899	1,232	680	552	1,667	58.0
$15,000–$19,999	2,894	1,169	616	553	1,725	56.0
$20,000–$24,999	2,295	1,006	469	537	1,289	54.0
$25,000–$29,999	2,117	958	538	420	1,159	53.0
$30,000–$34,999	1,877	983	492	491	894	52.0
$35,000–$39,999	1,600	977	519	459	622	51.0
$40,000–$44,999	1,453	798	471	327	655	51.0
$45,000–$49,999	1,208	646	378	268	562	49.0
$50,000–$54,999	1,181	744	376	368	438	50.0
$55,000–$59,999	952	636	374	262	316	49.0
$60,000–$64,999	935	638	421	217	297	49.0
$65,000–$69,999	802	537	316	220	265	49.0
$70,000–$74,999	722	517	340	177	205	49.0
$75,000–$79,999	656	472	281	192	183	49.0
$80,000–$84,999	513	364	226	138	149	48.0

$85,000–$89,999	404	310	222	88	94	48.0
$90,000–$94,999	409	249	166	84	160	49.0
$95,000–$99,999	310	227	149	78	83	48.0
$100,000–$104,999	334	224	126	99	110	49.0
$105,000–$109,999	204	144	99	45	60	47.0
$110,000–$114,999	243	151	94	57	92	48.0
$115,000–$119,999	180	126	72	54	54	48.0
$120,000–$124,999	206	127	105	22	79	48.0
$125,000–$129,999	182	121	82	39	61	49.0
$130,000–$134,999	161	133	101	32	28	48.0
$135,000–$139,999	187	144	100	45	43	51.0
$140,000–$144,999	148	110	72	38	38	48.0
$145,000–$149,999	130	102	60	43	28	49.0
$150,000–$154,999	144	107	73	34	36	48.0
$155,000–$159,999	110	91	45	46	19	50.0
$160,000–$164,999	71	53	49	4	18	48.0
$165,000–$169,999	59	45	33	12	15	48.0
$170,000–$174,999	65	49	36	13	16	50.0
$175,000–$179,999	66	48	35	13	18	49.0
$180,000–$184,999	54	36	24	12	18	49.0
$185,000–$189,999	43	33	23	10	10	49.0
$190,000–$194,999	51	40	18	23	11	48.0
$195,000–$199,999	35	33	27	6	2	49.0
$200,000 and over	761	613	395	217	149	51.0
Median income (dollars)	35,611	44,426	48,956	39,194	27,322	(X)
Standard error (dollars)	439	921	1,345	1,127	425	(X)
Mean income (dollars)	53,513	63,873	68,963	56,842	40,654	(X)

Standard error (dollars)	799	1,309	1,869	1,507	752	(X)
Gini ratio	0.482	0.475	0.478	0.464	0.458	(X)
Standard error	0.0051	0.0073	0.0102	0.0095	0.0073	(X)
Income per household member	29,836	33,301	35,068	30,709	24,804	(X)
Standard error (dollars)	456	690	979	815	470	(X)
Total persons in households	51,527	30,517	18,149	12,368	21,009	(X)

From U.S. Census Bureau, Current Population Survey. *Annual Social and Economic Supplement.* Washington, DC; 2012.

Internet release date: December 2013.

There is expected to be a disproportionately greater increase in the number of mentally ill older adults than that of the overall older population. There are several reasons for this demographic change. Excess mortality (from suicide or physical comorbidity) in early life among seriously mentally ill patients such as those with schizophrenia, depression, and substance dependence is likely an important reason for a lower prevalence of nondementia psychiatric disorders in older compared with younger adults. With an improved overall standard of living, better treatment for physical disorders, and more effective treatments for psychiatric illnesses, the high mortality in mentally ill young adults should begin to decrease, resulting in a greater number of those who will reach old age. Also, with the general increase in the number of people older than 65 years, there will be more older adults in the community at risk of developing late-onset psychiatric disorders. Furthermore, aging baby boomers are expected to have a higher risk of depression, anxiety disorders, and substance use disorders than those born before the World War II. For all these reasons, it is estimated that the number of older mentally ill adults in the population will increase by at least 10 percent from 2000 to 2030, that is, it will be about 22 percent. Thus, the number of psychiatrically ill older adults will rise by 275 percent, from 4 million in 1970 to 15 million in 2030, while comparable number for those aged 30 to 44 years will increase by only 67 percent, from over 9 million to 15 million (Fig. 57.1a–3). These estimates are educated crude guesses; nonetheless, they provide some understanding of the magnitude of the challenge facing psychiatry.

Table 57.1a–7.
Prevalence of Psychiatric Disorders among Younger versus Older Adults in the United States

DSM-III-Diagnostic Categoryα	Younger Adults (30–44 Years)—ECAα		Older Adults (ê65 Years)—ECAα		Prevalence of Clinically Significant Symptoms
	1-Year Prevalence	Lifetime Prevalence	1-Year Prevalence	Lifetime Prevalence	
Affective Disorders					
Any	2.7% men	6.6% men	0.6% men	1.6% men	15–25%
Major depression	7.9% women	15.3% women	1.5% women	3.3% women	
Dysthymia	3.9%	7.5%	0.9%	1.4%	
Bipolar I	Not available	3.8%	Not available	1.7%	
Bipolar II	1.2%	1.4%	0.1%	0.1%	
	0.3%	0.6%	0.1%	0.1%	
Anxiety Disorders					
Panic disorder	0.7% men	1.8% men	0.04% men	0.1% men	17–21%
Phobic disorder	1.9% women	3.1% women	0.4% women	0.7% women	
Generalized anxiety disorder	6.1% men	10.5% men	4.9% men	7.8% men	
Obsessive-compulsive disorder	16.1% women	22.6% women	8.8% women	13.7% women	
	3.6%	4.9–6.8%	2.2%	2.6–4.3%	
	2.1%	3.3%	0.9%	1.2%	
Alcohol Abuse/Dependence	14.1% men	27.9% men	3.1% men	13.5% men	7–8% community-dwelling older adults consume 12 to 21 drinks/week; 10–15% of older primary care patients may have alcohol-related problems
	2.1% women	5.5% women	0.5% women	1.5% women	
Drug Abuse/Dependence	Not available	6.7%	Not available	0.1%	
Schizophrenia	1.5%	2.3%	0.2%	0.3%	
Antisocial Personality	1.5%	3.7%	0.0%	0.3%	
Cognitive Impairment					
Severe	0.3% (35–54 years old)	Not available	1.0% (55–74 years old)	Not available	
			5.0% (75 years or	Not available	

			older)	
Mild	3.1% (35–54 years old)	Not available	7.5% (55–74 years) 19.1% (75 years or older)	Not available Not available
Any psychiatric disorder (excluding cognitive impairment)	23%	39%	13%	21%

Note: Although not a psychiatric disorder per se, suicide is generally secondary to a major mental illness such as depression or schizophrenia. Suicide rates increase with advancing age (especially in Caucasian men), and seem to be increasing among more recent cohorts of older persons. Among older patients in primary care, the prevalence of suicidal ideation is 0.7 to 1.2%.

[a]The ECA (Epidemiologic Catchment Area) Study used DSM-III criteria.

From Jeste DV, Alexopoulos GS, Bartels SJ, et al. Consensus statement on the upcoming crisis in geriatric mental health: Research agenda for the next two decades. *Arch Gen Psychiatry.* 1999;56:848–853.

Table 57.1a–8. Overview of the Education and Training of Professionals in Geriatrics in the United States

	Nurses	Physicians	Social Workers
Total jobs held (2006)	RNs: 2.5 million LPNs: 749,000	633,000	595,000
Geriatric specialization or certification	Less than 1% of RNs and about 2.6% of APRNs certified	7,128 certified in geriatric medicine; 1,596 certified in geriatric psychiatry	About 4% of social workers specialize
Academic leadership	76% of baccalaureate programs have at least one full-time “expert,” 29% have a certified faculty member	Less than 1% of faculty specialize; all programs have an identifiable leader in geriatrics	40% of schools have no faculty knowledgeable in aging
Exposure to geriatrics in schools	One-third of baccalaureate programs require exposure; 94% of fundamental courses integrate geriatric content	98% of schools require some form of exposure	80% of BSW students have no coursework in aging
Advanced geriatric training programs	Less than 100 master’s and post-master’s programs; five programs in geropsychiatric nursing	Medicine: 139 fellowship programs (468 first-year positions)	29% of MSW programs offer aging certificate, specialization, or concentration
Number of advanced geriatric trainees	Approximately 300 geriatric APRNs produced annually	Psychiatry: 58 fellowship programs (142 first-year positions)	Unknown
Explicit testing	Yes	Medicine: 253 in first	No general

on nongeriatric board certification exams	year; 34 in second year	certification; national licensure exam organized by approaches own

RN, Registered Nurse; LPN, Licensed Practical Nurse; APRN, Advanced Practice Registered Nurse; BSW, Bachelor of Social Work; MSW, Master of Social Work; DSW, Doctor of Social Work.

From Institute of Medicine; Committee on the Future Health Care Workforce for Older Americans, Board on Health Care Services. *Retooling for an Aging America: Building the Health Care Workforce.* Washington, DC: The National Academies Press; 2008.

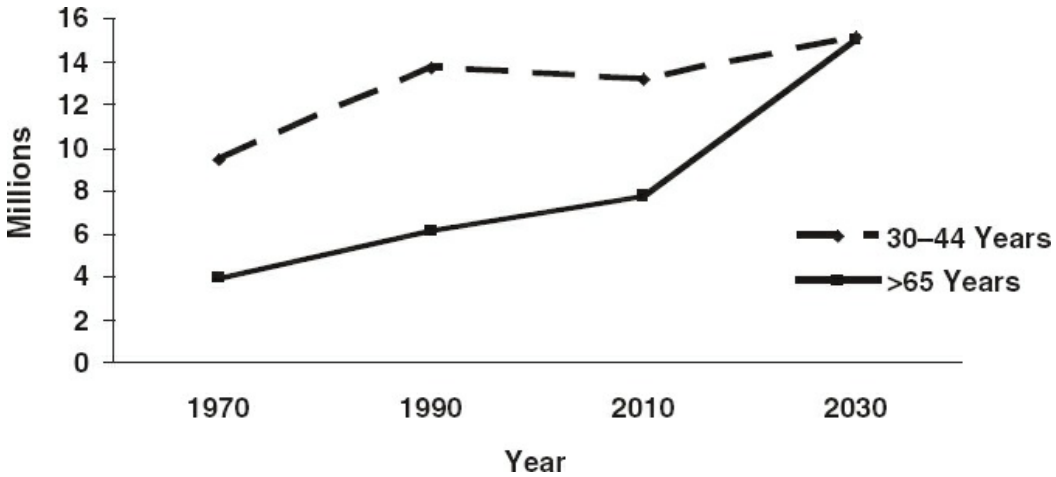

FIGURE 57.1a–3. Estimated number of people with psychiatric disorders in the United States. (From Jeste DV, Alexopoulos GS, Bartels SJ, et al. Consensus statement on the upcoming crisis in geriatric mental health: research agenda for the next two decades. *Arch Gen Psychiatry.* 1999;56:848–853.)

It is also to be noted that not only will the number of older adults with mental illness increase dramatically over the next few decades, but also the types of the patients one will see will change to some extent. For several disorders, there will be higher proportions of patients at the two ends of the severity spectrum. For example, expected advances in the ability to detect Alzheimer disease (AD) in earlier stages and to decrease the rate of decline via use of effective neuroprotective agents may result in greater number of patients who are more mildly affected and who will present different management problems from patients with advanced dementia. Issues regarding employment, financial planning, driving, home care, and day care will become more prominent. At the same time, reduced mortality from severe forms of physical illness (e.g., strokes, heart disease) will lead to increased number of severely demented patients with comorbid medical conditions living longer, raising different clinical, research, and ethical issues.

GERIATRIC PSYCHIATRY AS A SUBSPECIALTY

There are unique diagnostic and therapeutic considerations in the management of older adults with mental illnesses. Just as children are not small adults, the older adults are not merely chronologically more aged people. There are important psycho-bio-social differences that have an impact on the diagnosis, course, and treatment of mental (and physical) illnesses in different age groups. Because of age-related changes in the functioning of liver, kidneys, and brain, older people can only tolerate lower doses of most medications and are at a greater risk of developing a number of side effects. They also have special needs for psychosocial treatments (e.g., they are less likely to be living with their families and more likely to have financial difficulties). Additionally, they tend to have major physical illnesses and disabilities, have visual and hearing impairments, and have even less access to health care than younger adults. As a result, fewer than half of older Americans with mental illness receive the needed mental health services. They face the dual stigma of mental illness and aging. These problems are further compounded by therapeutic pessimism (and even nihilism) fostered by rampant societal as well as professional ageism.

Old-age psychiatry has had a long and distinguished tradition in the United Kingdom and other countries in Europe. In recent years, it has started to come of age in the United States too. Geriatric psychiatry became the second subspecialty of psychiatry (after child psychiatry) recognized by the U.S. Accreditation Council for Graduate Medical Education and by the American Board of Psychiatry and Neurology. There are several geriatric psychiatry organizations such as the American Association for Geriatric Psychiatry, the International Psychogeriatric Association, and the International College of Geriatric Psychoneuropharmacology. There are also several peer-reviewed journals devoted to geriatric psychiatry including the *American Journal of Geriatric Psychiatry*, the *International Psychogeriatrics*, the International Journal of Geriatric Psychiatry, and the *Journal of Geriatric Psychiatry and Neurology*. In basic sciences, there is a growing interest in the molecular biology of aging and neuroplasticity of aging.

Unfortunately, the increasing need for a geriatric mental health workforce is being met with a diminishing availability of experts in the field. There is a severe shortage of workforce trained to provide the necessary mental and physical health care for the growing number of older adults (Table 57.1a–8). Despite an urgent need for a trained workforce with expertise in geriatric mental disorders, the number of American Board-certified geriatric psychiatrists, who have maintained their subspecialty certification, has dropped from over 2,200 a decade ago to fewer than 1,800 today. This number is projected to decrease further to 1,650 by the year 2030, amounting to less than one geriatric psychiatrist for every 6,000 older adults with mental health and substance use disorders. Out of more

than 40,000 psychiatrists in the United States, only about 5,000 list "geriatric or gerontological psychiatry" as one of their three primary areas of interest.

The Institute of Medicine (IOM) has reported that the "tipping point" of an inevitable shortfall in providers has been passed, and it highlights the urgency of expanding and strengthening the geriatric health care workforce. Yet, little is being done to address this challenge. Already geriatric psychiatric services are in high demand and specialists are in short supply. Trainee pool is a critical issue. Currently, there are over 100 geriatric psychiatry fellowship training slots available nationwide, but the number of slots filled has declined from more than 80 percent in 2000 to under 50 percent today. Psychiatry residents, often struggling with large student loans, are drawn away from geriatric psychiatric fellowship, at least partly because of financial concerns. Rates of Medicare reimbursement are decreasing, and the reimbursement system as a whole makes psychiatric care difficult to access for older patients. The dual stigma against aging and mental illness also makes geriatric psychiatry less attractive to many trainees.

COMMON MENTAL ILLNESSES IN OLDER PEOPLE

A brief description of a few of the mental disorders of late life is provided below. A detailed description of each illness, its causes, and treatments is provided in subsequent chapters of the book.

Alzheimer Disease and Other Dementias

AD is the most common form of dementia. It is a progressive brain disorder that results in cognitive deficits such as memory loss, impaired thinking, and abnormal behavior. The number of older persons with AD will increase substantially over the next three decades for several reasons. The prevalence of dementia doubles every 5 years after the age of 60. Thus, 1 percent of the people aged 60 to 64 years are affected; 2 percent of those aged 65 to 69, 4 percent of those between 70 and 74, 8 percent of those 75 to 79 years old, 16 percent of those aged 80 to 84, and 30 to 45 percent of those aged 85 years and older. These figures provide a reasonable guide to the number of patients with AD and other dementias anticipated in over the next three decades. From a different perspective, one of the "hottest" areas of research is developing drugs that delay the progression of brain degeneration in patients with dementia. AD and other dementias are seriously underrecognized at present. In various studies, investigators have found that primary care practitioners recognize only a minority of the AD patients identified by researchers using a structured diagnostic interview. A greater awareness of the need for early detection of AD among professionals as well as among lay public can be expected. Currently, many

researchers are working on developing biological markers to detect early AD. Reliable and valid biological markers will diminish the present dependence on clinical (largely subjective) judgment in diagnosing early AD and enhance the rate of detection of AD even before obvious dementia has set in. Furthermore, as new therapies for the treatment of very early AD (even before dementia is obvious) become available in the coming years, improved recognition of patients before they manifest the disease clinically will increase the number of identified cases in the population.

Interestingly, there are also factors that could serve to reduce the incidence of AD in the future cohorts of the aged individuals. One such factor is education. The prevalence of AD is lower among those with more years of formal education. There will be greater number of older baby boomers with a high school diploma compared to the current cohort of the older adults. Education also has a variety of secondary effects that may reduce the frequency of dementia. For example, head injury has been identified in several studies as a risk factor for AD, and head trauma is less common among those with more education who are more likely to wear safety belts, use protective head gear in contact sports, and avoid driving while intoxicated. Such factors associated with a reduced risk of AD point to the way in which may be able to lower the incidence of AD may be able to be lowered by proactive societal actions.

Neuropsychiatric behavioral disturbances such as agitation, psychosis, and depression are common in patients with AD, and are a frequent cause of institutionalization. About a third of the patients with AD have depressive symptoms, whereas the prevalence of psychosis is about 50 percent. Agitation is even more frequent, especially in later stages. A number of intervention studies of pharmacological as well as psychosocial treatments have been performed for these neuropsychiatric symptoms, with variable results.

Dementias other than AD, such as Lewy body dementia, frontotemporal dementia, and vascular dementia are receiving increasing attention too, mainly in the areas of brain imaging and neuropathology. Again, behavioral disturbances are a common phenomenon in these patients.

Anxiety Disorders

Anxiety disorders are the most common disorders among the older adults with the exception of dementia. There are a number of differences between younger versus older adults in terms of psychopathology of these disorders. For example, older phobic patients tend to have fear of situations or inanimate stimuli, such as heights or lightning, whereas younger individuals are more likely to have a fear of animals. Because of physical and psychosocial changes associated with aging, it might be difficult to distinguish between phobias and nonpathological avoidance behaviors in older patients, such as avoiding leaving home due to arthritis or a fear of

crime. Older adults might be more likely than younger patients to manifest a mixture of anxiety and depression. Such comorbidity could be a result of cooccurrence of two common conditions, or could reflect one being an epiphenomenon or a complication of the other. A syndrome of "mixed anxiety–depressive disorder" might have both heuristic and clinical utility for older patients.

Depression

Numerous studies have shown that depression goes undiagnosed in sizable proportions of older adults. This is due, in part, to frequent denial of depression with its associated social stigma and attributing the symptoms of depression to "normal" aging or to physical illnesses. In the past three decades, some of the SSRIs have become nearly household names, leading to more recognition and acceptance of depression in the public at large. This trend will continue, resulting in reduced social stigma associated with depression. Another important reason for the underdiagnosis of depression in the primary care setting, where most older persons are seen, is the inadequacy of geriatric mental health training at the medical school level. With improvements in medical education, this situation will change in the years to come, resulting in greater detection of depression in the older adults.

Major depression manifests somewhat differently in the older compared to younger adults. Physical symptoms such as change in appetite and body weight, constipation, and sexual dysfunction are more common in the older adults; on the other hand, the physical symptoms in older adults could be a result of an undiagnosed physical illness (e.g., pancreatic cancer). While the point prevalence of subsyndromal depression in the younger adults is comparable to or lower than that of major depression, the prevalence of subsyndromal depression (18 percent) is three times that of major depression (6 percent) in people aged 65 years or older. Thus the predominant form of depression in the older adults appears to be subsyndromal depression. Notably, subsyndromal depression is also associated with significant increases in psychosocial impairment in major domains of everyday functioning.

A number of large epidemiological investigations in several different countries have shown that the risk of depression, as well as depression-related suicides, alcoholism, and substance abuse is significantly higher in the baby boomers than in those born before the World War II. The reasons for this type of "cohort effect" (i.e., different rates of an illness in people from different generations) are not known, but could be related to several different factors: genetic factors such as new mutations among recent generations. A number of environmental factors may also contribute to greater stress and increased risk of depression among the baby boomers; these include changing family structure with nuclear families, higher

divorce rates, changed roles for working women, greater urbanization, and a decrease in "social connectedness" due to geographical mobility and job changes. Finally, depression is known to be much more common among those with physical illnesses than in the physically healthy people. This observation leads to the development of preventive interventions in individuals with physical disorders such as heart disease and cancer.

Schizophrenia and Other Psychotic Disorders

The century-old concept of schizophrenia, based on Kraepelin's definition of dementia praecox, is that this is a disorder with onset in early life, that is, adolescence and early adulthood ("praecox") and progressive deterioration ("dementia"). Yet studies in older patients show that schizophrenia can have onset in middle age (late-onset schizophrenia). Late-onset schizophrenia is more common in women, and is usually of paranoid type. In terms of its course in late life, schizophrenia is typically not associated with a dementing course. Indeed, a small minority of patients with schizophrenia have a sustained remission of their illness in old age.

Psychotic symptoms may manifest for the first time after age 65. This is generally not schizophrenia, but is best labeled very late-onset schizophrenia-like psychosis.

Substance Use Disorders

As better treatments become available for substance use disorders (including alcoholism), the focus in the war on drugs during the coming decades will shift from punishment to treatment. Studies using currently available diagnostic criteria for substance abuse have underestimated the prevalence of drug abuse in the older adults because these criteria were developed and validated only in young and middle-aged samples. For example, DSM-5 criteria for substance use disorders include development of tolerance to the effects of that substance (e.g., a given amount of a stimulant produces less stimulation than it used to), resulting in a need for increased consumption over time. Yet, liver and kidney damage common in old age would produce higher concentrations of that substance in the blood, while age-associated brain changes would serve to enhance the effects of the substance on the brain. Hence, there is likely to be decreased tolerance to that substance in an older individual (i.e., a given amount produces more effect than it used to) leading to a need for decreased consumption. This older person, although still dependent on the substance, may not meet the "official" criteria for diagnosis. One DSM-5 criterion for substance use disorders is a report of being repeatedly unable to carry out major obligations at work, school, or home, such absences or poor work performance, suspensions or expulsions from school, and neglect of children. Such adverse consequences are much less likely to occur in older

adults who often live alone and are unemployed, resulting in underestimation of substance use disorders.

It is generally believed that aging is associated with reductions in and changing patterns of drug use (i.e., reductions in illicit drug use). One theory posits that factors associated with "maturation" associated with aging processes, and increased number of individuals who have been dependent on illicit drugs would become abstinent as they age. This is, however, not entirely supported by empirical data which suggest that there may be increasing number of older addicts and that persons who have been dependent for more than 5 years do not become abstinent as they age. In addition to an increased number of aging "lifelong" drug abusers, a greater number of individuals who, after decades of abstinence, return to drug or alcohol problems are also likely to be seen. Psychosocial stressors including isolation, various types of losses, loneliness, and onset of illness have been related to late-onset drug use.

There is a greater use and acceptance of psychotropic medications among the baby boomers than among those born before the World War II. In contrast to younger substance abusers who most often abuse illicit drugs, substance abuse problems in the older adults may result from misuse or abuse of over-the-counter (OTC) and/or prescription drugs. Older individuals use prescription drugs three times as frequently as the general population, and the use of OTC medications by this group is even more extensive. The estimated expenditure on prescription drugs by the older adults in the United States is $15 billion annually, a fourfold greater per capita expenditure on medications compared to younger individuals. Psychoactive drug use (most often sedative-hypnotic, anti-anxiety, and analgesic) is particularly problematic, including long-acting drugs not recommended for the older adults, and inappropriately high doses of benzodiazepines.

Comorbid physical as well as psychiatric disorders appear to be important predictors of prescription drug use. As greater number of people with chronic physical and psychiatric disorders survive into older age, there will be increased prevalence of substance use disorders in the older population.

The jury is still out whether the legalization of marijuana for medical use or adult recreational use in a growing number of states may negatively affect the health of older adults. Marijuana is the most commonly abused drug in the United States and around the world, but it is likely that with legalizing marijuana, the number of marijuana users would increase.

POSITIVE SIDE OF AGING

Contrary to the concept that aging means progressive and inevitable deterioration of the brain and mind, a number of recent studies across the world shows that psychological well-being and happiness increase with age,

especially after middle age. Some studies have found that subjective feeling of well-being (a sense of contentedness) is high at the beginning of adult life, then declines progressively until it hits the rock bottom in middle age, producing the so-called mid-life crisis, but then starts to rise again. Other investigations report a progressive rise in well-being from age 20 to 80+. Interestingly, the level of happiness in the early 80s is similar to that in the early 20s. The increase in happiness after middle age occurs even as physical, cognitive, and fertility functions decline. This finding has been reported by researchers from several different countries and cultures. Notably, perfect physical health is neither necessary nor sufficient to feel happy as people grow older. Research shows a paradox of aging: physical health and some aspects of cognitive function (for instance, memory for names and faces) decline with age, while happiness, mental health, and management of interpersonal relationships improve.

What are the possible explanations for this counter-intuitive result? Suggested factors include changes in attitude such as greater acceptance of one's physical limitations, contentedness with past accomplishments in life, reduced preoccupation with peer pressure, a more realistic appraisal of one's own strengths and limitations, and increased emotional stability. Intriguingly, a recent study, based on the reports by the individual caretakers of chimpanzees and orangutans in zoos and other protected areas, found an increase in these great apes' well-being after middle age. Thus, better well-being in older age might be rooted, in part, on biological changes in the brain.

Several aspects of aging can be aided significantly by improved health care, greater societal value placed on the older adults, and a more positive conceptualization of old age. Most psychiatrists today would not agree with Freud's assertion in 1924 that near or above the age of 50, the elasticity of mental processes is lacking and, therefore, older individuals are "no longer educable." The so-called problems of aging should become opportunities for a better understanding and improvement of the entire life process.

The aging of the population is generally viewed with concern and alarm about its impact on social security, health care (including Medicare), housing, jobs for younger generations, etc. Yet, this is also a grand opportunity to make positive use of the growing older adults' population so that the new era, dubbed by some as "Silver Tsunami," can be labeled the "Golden Wave." The most important advance in neuroscience in the past two decades has been the demonstration of neuroplasticity of aging—the fact that brain growth and development continue into old age. Furthermore, such a positive outcome is related, not so much to the genes people inherit from their parents, but to their behavior, attitude, and environment. With appropriate physical activities, targeted nutrition, and psychosocial stimulation along with necessary health care and social support, successful aging need not be a dream, but can become a

widespread reality. As the quality of life of older adults improves, they become more productive, contribute more to the rest of the society, and their health care costs go down. The seniors who learn and teach can be wonderful and wise mentors for younger generations. Older citizens can lead efforts at social entrepreneurship in a collaborative manner.

SUCCESSFUL COGNITIVE AND EMOTIONAL AGING

Research on aging has typically focused on disease and disability. However, during the past two decades, this focus has slowly begun to expand to incorporate studies of "successful or healthy aging." Understanding why some people seem to adapt to challenges they encounter in late life when others cannot, would facilitate the development of interventions and policies geared toward older populations. Yet, there are many times as many studies of aging-related disorders as there are on successful aging. One problem in research on successful aging is that it involves multiple components (e.g., physical health, psychological well-being) that extend beyond longevity, and as a result, there are numerous criteria for defining it but no consensus. In 1987, Rowe and Kahn proposed that successful aging involved three components: (1) absence/low probability of disease and disease-related disability, (2) high cognitive and physical functioning, and (3) active engagement with life. According to this model, however, many cognitively intact and socially active older adults would not be categorized as aging successfully because of the presence of chronic physical diseases and disabilities. Also, such criteria fail to take into account the perspectives of older adults themselves about what defines and determines successful aging. Whereas researchers tend to focus on physical health as a critical factor to successful aging, older adults tend to stress the role of psychological factors such as adaptation and acceptance, along with emotional well-being. One focus group study indicated that older adults viewed successful aging as involving a foundation of health and security, along with attitudes and adaptive behaviors that can compensate for difficulties with that foundation, and enable engagement in meaningful activity and stimulation. It is particularly noteworthy that a number of aspects of successful aging identified by older adults in this study (e.g., attitude, health care, financial security) are potentially modifiable.

Successful aging is a complex and multidetermined trait and state, with numerous genes and environmental factors making small contributions to the overall presentation of the phenotype. No single factor is either necessary or sufficient to determine the outcome in a specific individual. Cognitive preservation may, however, be a prerequisite of independent living, positive adaptation, mastery, and growth, all of which seem to be the major determinants of "success."

IMPLICATIONS FOR FUTURE PROGRESS

The impending crisis in geriatric mental health with the projections of dramatic increase in the number of mentally ill older people as the baby boomers reach old age should be a cause of great concern. However, the good news is that there is something that can be done about this issue. First, begin by recognizing the fact that the expected increases in the ranks of older mentally ill adults are not necessarily a bad thing. It is better to diagnose illnesses that are treatable rather than keeping them from being detected in the older adults. Advances in the pharmacological and psychosocial management of psychiatric disorders over the past two decades have already improved the treatment of several serious mental illnesses in older people. Previously unimaginable advances are taking place in the understanding of brain function in normal aging as well as in various neuropsychiatric disorders. New lines of investigation are pointing to possible ways of not just managing but even possibly curing and preventing several diseases for which there was little hope in the past. Brain stem cell implants, gene therapy, and vaccines for AD (despite the initial disappointing results) are among a host of therapies that are moving from the realm of science fiction to reality in a matter of years instead of decades. Early diagnosis of disorders such as depression or AD is important in order to start treatments that will prevent or at least slow their progress. Similarly, it is better that young adults with schizophrenia who currently have a shorter life span than the general population live into old age. Hence, an increase in the number of older adults with mental illnesses will be, in part, a welcome development.

At the same time, some of the increased prevalence of mental illness in older people may be preventable. For example, a proportion of cases of late-onset depression or substance use disorders could be prevented by appropriately treating individuals at high risk of developing these disorders. Thus, support groups for bereaved persons could lower the risk of developing major depression. Similarly, persons with stroke or myocardial infarction are at a high risk of depression, and the incidence of poststroke and postmyocardial infarction depression could be reduced with use of suitable antidepressant treatments and/or psychosocial interventions following strokes or heart attacks. There are new emerging neurobiological data suggesting that physical exercise may be as beneficial for the brain as for the body. Research also needs to be conducted to determine if engaging in specific types of mental activities may lead to delayed onset of dementia.

The less developed regions of the world will undergo proportionately larger demographic shifts in the decades ahead than the more developed ones that have had stable population base for many years. Yet, the less developed nations also have fragile economies and will have limited opportunity to increase health services and treatment for aged individuals. This global situation places a premium on developing diagnostic

technologies and treatments that are sufficiently inexpensive and can be widely applied.

Reducing social stigma associated with aging is as critical as reducing social stigma associated with mental illness. Mental illnesses are brain disorders and as such, should receive the same consideration in terms of insurance coverage.

The age-old stereotype that late life is a gray and dark period of life can be changed to try to make it a "golden" one for senior citizens. There are several aspects of aging that can be aided significantly by improved health care. If the societal value placed on the older population becomes more positive and the scientific openness to conceptualizing old age increases, it will become clear that the so-called problems of aging are actually opportunities for a better understanding and improvement of the whole life process. To quote Alex Comfort, "No pill or regime known, or likely, could transform the latter years of life as fully as could a change in our vision of age and a militancy in attaining that change." The same applies to the attitude toward mental health and mental illness in old age.

References

Alexopoulos GS, Streim J, Carpenter D, Docherty JP; Expert Consensus Panel for Using Antipsychotic Drugs in Older Patients. Using antipsychotic agents in older patients. *J Clin Psychiatry*. 2004;65:5–99.

Bryant C, Mohlman J, Gum A, et al. Anxiety disorders in older adults: looking to DSM5 and beyond . . . . *Am J Geriatr Psychiatry*. 2013;21(9):872–876.

Charney DS, Reynolds CF 3rd, Lewis L, et al. Depression and bipolar support alliance consensus statement on the unmet needs in diagnosis and treatment of mood disorders in late life. *Arch Gen Psychiatry*. 2003;60:664–672.

DeMichele-Sweet MA, Sweet RA. Genetics of psychosis in Alzheimer's disease: a review. *J Alzheimers Dis*. 2010;19:761–780.

Depp C, Vahia IV, Jeste D. Successful aging: focus on cognitive and emotional health. *Annu Rev Clin Psychol*. 2010;6:527–550.

Fingerhood M. Substance abuse in older people. *J Am Geriatr Soc*. 2000;48(8):985–995.

Howard R, Rabins PV, Seeman MV, Jeste DV; The International Late-Onset Schizophrenia Group. Late-onset schizophrenia and very-late-onset schizophrenia-like psychosis: an international consensus. *Am J Psychiatry*. 2000;157:172–178.

Iglewicz A, Meeks TW, Jeste DV. New wine in old bottle: late-life psychosis. *Psychiatr Clin North Am*. 2011;34(2):295–318.

Institute of Medicine; Committee on the Future Health Care Workforce for Older Americans, Board on Health Care Services. *Retooling for an Aging America: Building the Health Care Workforce*. Washington, DC: The National Academies Press; 2008.

Institute of Medicine. *The Mental Health and Substance Use Workforce for Older Adults: In Whose Hands?* Washington, DC: The National Academies Press; 2012.

Jeste DV, Alexopoulos GS, Bartels SJ, et al. Consensus statement on the upcoming crisis in

geriatric mental health: research agenda for the next two decades. *Arch Gen Psychiatry*. 1999;56:848–853.

Jeste DV, Blazer DG, First M. Aging-related diagnostic variations: need for diagnostic criteria appropriate for elderly psychiatric patients. *Biol Psychiatry*. 2005;58:265–271.

Jeste DV, Savla GN, Thompson WK, et al. Association between older age and more successful aging: critical role of resilience and depression. *Am J Psychiatry*. 2013;170:188–196.

Jeste DV, Twamley EW, Eyler Zorrilla LT, Golshan S, Patterson TL, Palmer BW. Aging and outcome in schizophrenia. *Acta Psychiatr Scand*. 2003;107:336–343.

Jorm AF, Korten AE, Henderson AS. The prevalence of dementia: a quantitative integration of the literature. *Acta Psychiatr Scand*. 1987;76:465–479.

Kuerbis A, Sacco P, Blazer DG, Moore AA. Substance abuse among older adults. *Clin Geriatr Med*. 2014;30(3):629–654.

Lohr JB, Palmer BW, Eidt CA, et al. Is post-traumatic stress disorder associated with premature senescence? A review of the literature. *Am J Geriatr Psychiatry*. 2015;23(7):709–725.

McKeith IG, Dickson DW, Lowe J, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. *Neurology*. 2005;65(12):1863–1872.

Ortman JM, Velkoff VA, Hogan H. *An Aging Nation: The Older Population in the United States, Current Population Reports*. Washington, DC: U.S. Census Bureau; 2014.

Rowe JW, Kahn RL. Human aging: usual and successful. *Science*. 1987;237:143–149.

Schuitevoerder S, Rosen JW, Twamley EW, et al. A meta-analysis of cognitive functioning in older adults with PTSD. *J Anxiety Disord*. 2013;27(6):550–558.

Simoni-Wastila L, Yang HK. Psychoactive drug abuse in older adults. *Am J Geriatr Pharmacother*. 2006;4(4):380–394.

Thorp SR, Ayers CR, Nuevo R, Stoddard JA, Sorrell JT, Wetherell JL. Meta-analysis comparing different behavioral treatments for late-life anxiety. *Am J Geriatr Psychiatry*. 2009;17(2):105–115.

U.S. Census Bureau, Current Population Survey. *Annual Social and Economic Supplement*. Washington, DC; 2012.

Weintraub D, Hurtig HI. Presentation and management of psychosis in Parkinson's disease and dementia with Lewy bodies. *Am J Psychiatry*. 2007;164:1491–1498.

57.1b Epidemiology of Psychiatric Disorders

CELIA F. HYBELS, PH.D., AND DAN GERMAN BLAZER, II, M.D., PH.D.

Epidemiology is the basic science of public health. The data provided from epidemiological studies are critical for recognizing the relative burden of varieties of psychiatric disorders. Not only does epidemiology present the situation of psychiatric disorders in the population, it serves as a means for identifying causal factors which can be followed up with clinical and basic science investigations. Data from epidemiological studies are also critical in service planning nationally and at the level of local communities. Perhaps at no time in recent years have epidemiological studies of older adults been of more importance given the increasing numbers of older adults and the frequency of disorders encountered, some of which are apparently increasing in frequency, such as substance use disorders.

Epidemiology as a science is concerned with the patterns of disease in populations. In geriatric psychiatry, this includes identifying the proportion of older adults with a psychiatric disorder (diagnosed using operational criteria such as DSM-5) during a specified time period (*prevalence*) and the frequency of disease occurrence or the proportion of older adults who develop a psychiatric disorder within a specified time period (*incidence*). Incidence and duration of a disorder affect its prevalence. For example, disorders that have a slow progressive course, such as Alzheimer disease, are likely to increase in prevalence if the incidence increases, if survival with the illness increases (duration), and if no treatment is found. But it is the search for determinants of disease patterns (e.g., what characteristics are shared among persons in whom these disorders are most likely to be found and how these identified characteristics affect the course or duration of the disorder) and the identification of risk factors (e.g., which characteristics increase the probability of disease occurrence) that is the goal of most epidemiological research.

Case definition is critical to epidemiology, yet presents special challenges for psychiatric disorders in late life. Not only can these disorders occur for the first time in late life, but their presentation in older adults may also be different from that seen in younger adults. Many older adults may fail to meet specified criteria for psychiatric disorders yet endorse an array of symptoms that are clinically meaningful, suggesting that case definition must include psychiatric symptoms as well as disorders. In addition, through their comorbidity with physical illness and other mental disorders, psychiatric symptoms in older adults present diagnostic, prevention, and treatment challenges to health care professionals. This chapter provides an update to epidemiological data presented in the previous edition of this text while at the same time retaining descriptions and data which remain relevant and have not changed since the previous

edition.

DEMOGRAPHICS OF AGING

The last century has seen a dramatic increase in the number of adults age 65 or older, what some have described as the silver tsunami. In 1900, there were 3.1 million in this age group in the United States, while by the year 2010 the number had increased to 40.3 million. Similarly, those 85 or older, 0.1 million in 1900, numbered 5.5 million in 2010. By the year 2050, the population of persons 65 years of age or older in the United States is expected to be approximately 88.5 million, and those 85 or older are expected to number 19 million. The proportion of adults age 65 years or older is projected to increase from 13.0 percent in 2010 to 20.2 percent in 2050. The proportion age 85 or older is projected to increase from 1.8 percent in 2010 to 4.3 percent in 2050. Similar increases in the proportion of older adults are also expected outside of the United States as better treatments for chronic disease and other physical illness become available.

These projections suggest that even if the proportion of older adults with a current psychiatric disorder remains unchanged the number of older persons with a disorder will increase over the next several decades. In addition, the proportion of younger adults with a psychiatric disorder (excluding the dementias) is much higher than the proportion among older adults, suggesting that as current younger cohorts age the proportion of older adults with a psychiatric disorder may increase as well as the burden due to a higher number of older adults. As better treatments for mental illness are developed, the proportion of adults with a history of mental disorder surviving to old age may increase. Some forecasters have estimated that the number of older adults with a psychiatric disorder in the United States will be 15 million by the year 2030. These numbers emphasize the importance of understanding the epidemiology of psychiatric disorders in older adults.

PREVALENCE OF PSYCHIATRIC DISORDERS IN OLDER ADULTS

From 2001 to 2003, the World Health Organization (WHO) conducted a series of household surveys across 14 countries to estimate the prevalence, severity, and unmet need for treatment of mental disorders, the World Mental Health surveys. In the United States, the survey was called the National Comorbidity Survey Replication (NCS-R), and 9,282 respondents age 18 or older were interviewed. The WHO Composite International Diagnostic Interview (CIDI) was the instrument used to generate fourth revised edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR) diagnoses. As shown in Figure 57.1b–1, among adults age 60 or older in the NCS-R, the most prevalent lifetime disorders were major depression (10.6 percent), specific phobia (7.5 percent), social phobia (6.6 percent), and alcohol abuse (6.2 percent). A total of 26.1

percent of the respondents age 60 or older had one or more lifetime psychiatric disorders, 11.6 percent had two or more lifetime disorders, and 5.3 percent had three or more. These lifetime prevalence estimates were lower than those in other age groups. Among those 18 to 29, 30 to 44, and 45 to 59, a total of 22.3, 22.5, and 15.9 percent, respectively, had three or more lifetime disorders.

The Collaborative Psychiatric Epidemiology Surveys (CPES) were sponsored by NIMH to combine three nationally representative surveys: the NCS-R described above, the National Survey of American Life (NSAL) which oversampled adults of African-American and Afro-Caribbean ethnicity, and the National Latino and Asian American Study (NLAAS). By using the same interview structure and the WHO version of the CIDI across the three studies, the CPES allowed for more intensive study of the epidemiology of psychiatric disorders within minority populations. For example, among those age 55 or older in the CPES, non-Hispanic Whites and Latinos consistently had a higher prevalence of psychiatric disorders. African Americans had a lower prevalence of major depression and dysthymia, and Asian Americans were less likely to report any affective disorder. The lifetime prevalence of major depression among adults age 55 or older in the NSAL was 6.7 percent among African Americans and 11.5 percent among Blacks of Caribbean ethnicity, compared to 14.3 percent among Whites. The 12-month prevalence of major depression showed a different pattern: 2.0 percent among African Americans, 4.7 percent among Whites, and 10.7 percent among Caribbean Blacks. Among older Asians in the NLAAS, the 12-month prevalence of major depression was 1.72 percent among those ages 60 or older compared to 1.55 percent among those ages 45 to 59 years, 5.47 percent among those ages 30 to 44 years, and 7.76 percent among those ages 18 to 29 years.

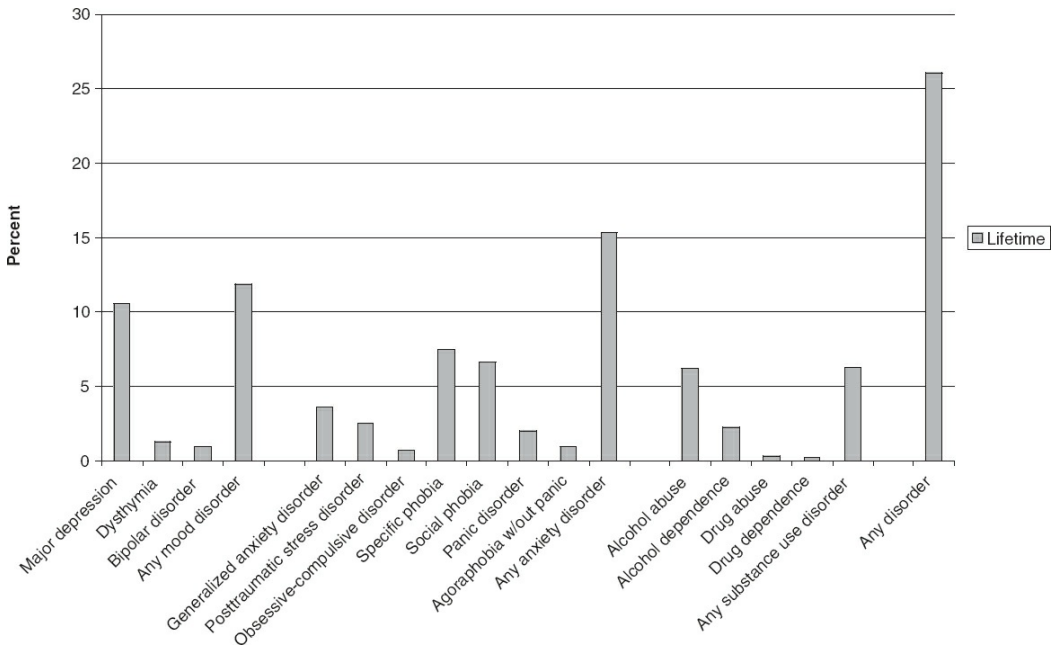

FIGURE 57.1b–1. Lifetime prevalence of selected psychiatric disorders among adults 60 or older in the National Comorbidity Survey Replication (NCS-R).

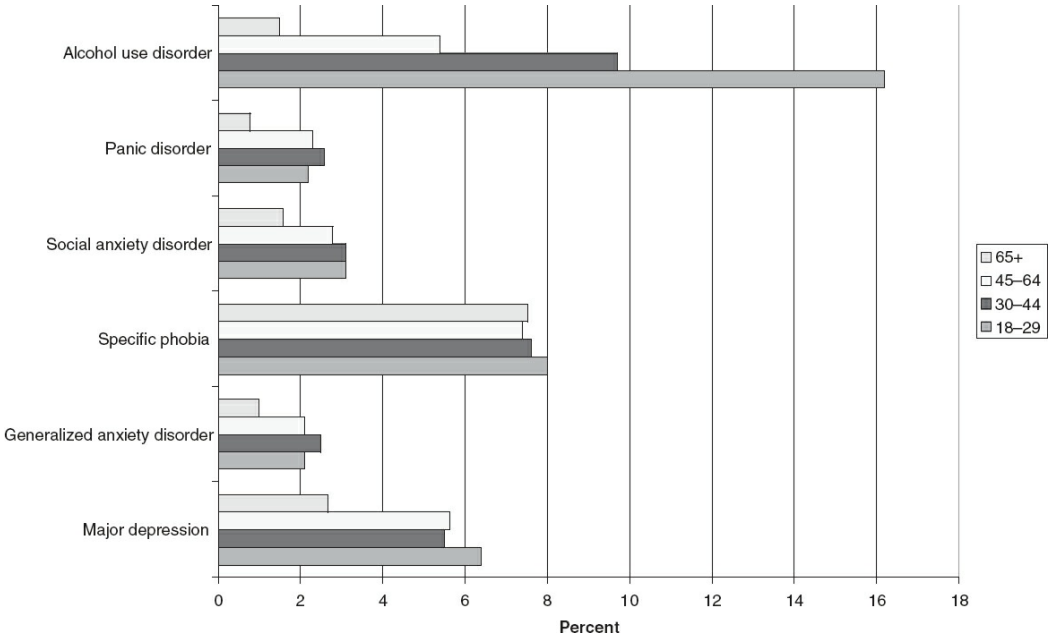

FIGURE 57.1b–2. Twelve-month prevalence estimates of selected psychiatric disorders in the National Epidemiologic Survey of Alcoholism and Related Conditions (NESARC) by age.

Another large-scale study conducted in the United States from 2001 to 2002 also provided national estimates of DSM-IV-TR disorders. In the National Epidemiologic Survey of Alcoholism and Related Conditions (NESARC), a total of 43,093 persons age 18 or older were interviewed. A structured interview, the Alcohol Use Disorder and Associated Disabilities Interview Schedule-DSM-IV Version (AUDADIS-IV), was used to generate diagnoses. The 12-month prevalence estimates for selected disorders across age groups are shown in Figure 57.1b–2. Among those age 65 or older, the 12-month prevalence of major depression was 2.7 percent, social anxiety disorder was 1.6 percent, generalized anxiety disorder was 1.0 percent, specific phobia was 7.5 percent, panic disorder was 0.8 percent, and alcohol abuse was 1.2 percent. Across all disorders, the prevalence is lowest among older adults in comparison to younger adults. In the NCS-R, the 12-month prevalence of any mood, anxiety, or substance use disorder was 8.5 percent among those age 65 or older compared to 27.6 percent among those ages 18 to 44 years and 22.4 percent among those ages 45 to 64 years.

The estimates from the CPES and the NESARC add to a growing body of literature reporting lifetime, 12-month, and current prevalence estimates across multiple countries. A description of the epidemiology of psychiatric disorders would not be complete without recognizing the contributions of one of the earliest studies of the prevalence of psychiatric disorders, the landmark Epidemiologic Catchment Area (ECA) survey conducted in the

late 1970s and early 1980s in five US communities—New Haven, Connecticut; Baltimore, Maryland; St. Louis, Missouri; Durham, North Carolina; and Los Angeles, California. A total of 18,571 persons participated, including 5,702 persons age 65 or older. Psychiatric disorders were measured using the Diagnostic Interview Schedule (DIS), one of the first structured interviews from which DSM diagnoses could be generated. These structured interviews could be administered by trained lay interviewers, allowing population estimates to be reported.

Across all five ECA sites, the 1-month prevalence of any DIS disorder in adults age 65 or older was 12.3 percent, compared to 16.9 percent in adults 18 to 24 years of age, 17.3 percent in those 25 to 44 years of age, and 13.3 percent in those 45 to 64 years of age. For all disorders, with the exception of substance use, alcohol abuse and dependence, and severe cognitive impairment, the prevalence was higher in older women than that in older men. The prevalences of any DIS disorder within the previous month, excluding the dementias, were 10.5 percent in men 65 years of age or older and 13.6 percent in older women.

Geriatric psychiatrists have suggested that these studies may underestimate the burden of psychiatric illness in older adults and that the frequency of clinically significant psychiatric symptoms in older adults may be much higher. Even so, data from studies conducted in the United States as well as other countries over the last several decades among persons age 60 or older have added much information concerning the prevalence of and risk factors for various psychiatric disorders and their associated subclinical forms in late life. Although they have published varying prevalence estimates of the disorders, these studies have concluded that mental illness in older adults is a significant public health problem.

The reported prevalence and incidence of psychiatric disorders in late life can be influenced by a variety of factors, including case definition, sampling issues, potential bias, and cohort effects.

Case Definition

Research studies generally define cases in one of two ways: (1) persons with psychiatric symptoms and (2) persons who meet criteria for specific psychiatric disorders as defined by the nomenclature. If strict diagnostic criteria are applied, then the prevalence of a psychiatric disorder is generally lower than that if the criteria for a case are more relaxed, such as not reaching the threshold of a DSM diagnosis. For example, the DSM-5 as well as previous editions since 1980 defines cases by the presentation of specific symptoms, often within a specific time period.

There are several instruments that have been used in epidemiological studies to identify persons with specific DSM disorders. Most of these instruments were originally developed for DSM-III through DSM-IV-TR. Large scale epidemiological studies are only beginning to apply DSM-5

criteria. Yet the criteria in DSM-5 have not varied such that the previous estimates are at significant variance from DSM-5 for older adults. The one exception is the addition of the meta-categories of mild (roughly equivalent to mild cognitive impairment [MCI]) (see Alzheimer Disease and Other Dementias) and major neurocognitive disorders (which basically is equivalent to the combination of all dementias). The DIS, previously mentioned, was developed for the ECA program and was based on the nomenclature in place at the time at which the ECA studies were conducted (DSM-III). Other instruments have been developed since that time. The version of the CIDI used in the CPES corresponds to DSM-IV-TR criteria and, like the DIS, can be administered by trained lay interviewers. The AUDADIS-IV used in the NESARC is able to generate DSM-IV diagnoses. The Structured Clinical Interview for DSM-IV-TR (SCID) is another structured instrument to make clinical diagnoses but must be administered by a health professional. It can be expected that modifications to these instruments will be made in order to generate DSM-5 diagnoses. Instruments have also been developed that correspond to the tenth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) criteria. For example, the Geriatric Mental State Automated Geriatric Examination for Computer-Assisted Taxonomy (GMS-AGECAT) has been used in many studies in Europe to identify clinically significant symptoms of depression, dementia, and other disorders that allow clinicians to identify cases which can be mapped to ICD or DSM criteria.

Other measures used widely in clinical and community epidemiological research identify specific symptoms rather than diagnostic criteria for selected disorders. For example, the Center for Epidemiologic Studies—Depression (CES-D) Scale and the Patient Health Questionnaire (PHQ-9) screen for depressive symptomatology, whereas the Mini-Mental State Examination (MMSE) screens for cognitive impairment. It is important to note that many epidemiological studies of mental illness in older adults include individuals as cases who do not necessarily meet the diagnostic criteria for a particular psychiatric disorder (such as clinically significant depressive symptoms that do not meet criteria for major depression). Prevalence and incidence estimates can vary, therefore, depending on the instrument used to define the cases and whether symptoms or disorders are being reported.

Many older adults have clinically significant symptoms documented by such instruments that, in number or duration, do not meet the DSM-5 or ICD-10 criteria, yet are associated with decreased physical and social functioning and overall quality of life, particularly subsyndromal depression and MCI. As described later in this section, the prevalence of these subclinical forms of the disorders is high and, for some disorders, may be more representative of the disorder in older adults.

Other factors affect case identification. Prevalence estimates may be based on occurrence of symptoms within the past month, 6 months, 1 year, or lifetime. For example, the 1-month prevalence is usually much lower than the lifetime prevalence. Age of onset may affect the lifetime prevalence of a disorder. Older adults may not recall the age at which symptoms were first experienced or even if they have experienced a prior episode. The reported prevalence can be affected if the disorder had to appear before a certain age to satisfy the diagnostic criteria. For example, previous restrictions on the diagnosis of schizophrenia to onset before 45 years of age reduced earlier prevalence estimates of schizophrenia among older adults. Case detection in late life is also complicated by the fact that many psychiatric disorders appear for the first time at a younger age and then reappear in late life. Finally, it is difficult to distinguish psychiatric symptoms in older adults that have a psychogenic origin from those that result from medication use or a physical illness. Older adults may experience symptoms of depression or cognitive impairment that are the result of medication use and may be misclassified as having a psychiatric disorder. Similarly, older adults may experience virtually all of the symptoms of major depression, yet these symptoms may be attributed to a physical illness.

Sampling Issues and Sources of Bias

Estimates of prevalence and incidence are determined by the target population about whom inferences are made. For example, estimates of the prevalence of psychiatric disorders in clinical populations are generally higher than those obtained when the target population is community-dwelling older adults. The prevalence of dementia in older adults is higher in samples from long-term–care facilities in comparison to samples of community-dwelling older adults. Psychiatric disorders for which adults may be less likely to seek treatment may be quite prevalent in community samples yet rarely observed in clinical samples. For example, one of the most surprising findings from the ECA surveys was the high prevalence of phobic disorder in the community population of elders. Because individuals may be unlikely to seek treatment, the population prevalence was not known until a large-scale community sample was undertaken.

Older adults may be more likely than younger adults to experience a physical illness or to be in a long-term–care facility, which can complicate or preclude their participation in a study that uses a representative sample, such as the CPES and the NESARC, resulting in an underestimate of the true prevalence of psychiatric disorders among older adults. Some older adults may not be cognitively able to give consent and to participate in research studies. If individuals with dementia are more likely to have depressive symptoms than those without dementia and demented subjects are excluded, then the prevalence of depressive symptomatology is

underestimated.

Finally, older adults may refuse to consent to participate in research studies, or their family may refuse to allow them to participate. If these persons are more likely to have clinically significant symptoms of mental illness, then differential patterns of nonresponse can result. In particular, studies in which elderly volunteers are recruited may be more likely to enroll healthier subjects than one would get in a random sample.

There are several types of biases that can affect the prevalence of psychiatric disorders in older adults reported from research studies. Misclassification bias is an important issue. Specifically, if the current nomenclature is less applicable to older adults, then elderly adults may be misclassified as not meeting criteria for the disorder. For example, older adults with an array of depressive symptoms but no report of sadness would not meet criteria for DSM-5 major or minor depression. Recall bias is important in studies of older adults who may selectively recall symptom information. Selection bias should also be considered, in that prevalence estimates can be affected by selective survival (i.e., only the healthiest adults survive to old age). As is discussed later in this section, suicide rates are high in late life, which can affect the prevalence of disorders such as major depression.

Cohort Effects

The prevalence of psychiatric disorders can be affected by external circumstances affecting a particular cohort. Specifically, it is important to determine if the observed prevalence of a psychiatric disorder in a particular age group is likely to remain stable over time (i.e., age is a major determinant of the disorder) or whether the prevalence is likely to be associated with a particular cohort. For example, individuals born 80 years ago may carry less burden of psychiatric disorders such as drug or alcohol abuse in comparison to adults born in more recent years. In addition, the oldest-old today may have fewer psychiatric disorders, because they have always been healthier, having survived such occurrences as the Great Depression and World War II. When the prevalence of psychiatric disorders is compared across studies, it is important to note in which time period the particular research was conducted and whether a particular characteristic of a cohort could affect the reported prevalence.

EPIDEMIOLOGY OF SELECTED PSYCHIATRIC DISORDERS IN OLDER ADULTS

Alzheimer Disease and Other Dementias

There have been a number of epidemiological studies conducted to determine the prevalence and incidence of dementia in older persons in community and institutional populations. The focus has shifted from severe cognitive impairment in the ECA to estimates of types of dementia,

such as Alzheimer disease and vascular dementia, to predementia syndromes such as MCI. Studies have also examined subtypes of MCI to identify those that may be more likely to progress to more severe forms of dementia. In DSM-5, a new diagnostic category of mild neurocognitive disorder was created which subsumed MCI. There are not yet studies available to provide the prevalence of this disorder.

The prevalence estimates of cognitive impairment and dementia in persons age 65 or older range from approximately 1 to 20 percent, depending on how cognitive impairment or dementia was defined in the research study as well as the population sampled. In the ECA surveys, the prevalence of severe cognitive impairment in noninstitutionalized adults age 65 or older was 4.9 percent. In the Canadian Study of Health and Aging, a large representative sample of adults age 65 or older residing in the community (N = 9,008) and in institutions (N = 1,255), the prevalence of dementia was estimated to be 8 percent. In the Aging, Demographics, and Memory Study (ADAMS), the prevalence of dementia was 13.9 percent. The prevalence of severe dementia reported from epidemiological studies is somewhat lower than that when mild and severe forms of dementia are combined, and ranges from less than 1 to 7 percent. It is important to note that older adults in the early stages may meet criteria for Alzheimer disease but not meet criteria for severe cognitive impairment or severe dementia.

The prevalence of dementia in institutional samples is generally higher than that reported in community samples. A study in Edmonton, Canada reported that more than two-thirds of the institutional population (69 percent) had mild or severe cognitive impairment, whereas the prevalence of mild impairment in the household sample was 3.3 percent, and severe impairment was not found outside the institutional sample. The prevalence of cognitive impairment among older adults seen in primary care is over 15 percent.

The prevalence of dementia and cognitive impairment is higher in older age groups. In the ECA study, the prevalence of severe cognitive impairment were 2.9 percent in adults 65 to 74 years of age, 6.8 percent in those 75 to 84 years of age, and 15.8 percent in those adults 85 years of age or older. The prevalence of dementia increases with age, with the prevalence estimated to double every 5.1 years in older adults.

Alzheimer disease is the most common type of dementia among older adults and accounts for over 50 percent and perhaps as much as 75 percent of the dementia cases. The prevalence of Alzheimer disease in older adults ranges from 3 to 10 percent and also increases with age. Among those 60 to 65 years of age, the prevalence is less than 1 percent but increases exponentially to be 24 to 33 percent among those 85 or older. Difficulties with sampling prevent investigators from accurately determining if the prevalence of dementia continues to increase or levels off after 90 years of age.

Vascular dementia is found among older adults with cerebrovascular disease. It is not clear whether the cerebrovascular disease is a risk factor for this type of dementia or whether the two disorders share common risk factors. Vascular dementia is also often comorbid with Alzheimer disease. Although less prevalent, other subtypes, including dementia with Lewy bodies, frontotemporal dementia, mixed dementia, Parkinson dementia, and unspecified dementia, have all been noted in community samples of older adults.

In the ECA studies, the prevalence of severe cognitive impairment in those 65 to 74 years of age was higher in men (4.2 percent) than that in women (1.9 percent), but the prevalence was higher in females in the older age groups. Among those age 85 or older in the ECA, the prevalence of severe cognitive impairment was 8.2 percent in men and 19.5 percent in women. Alzheimer disease has been found more often in older women in some studies, yet a true difference in prevalence by gender remains to be determined.

The incidence of dementia and Alzheimer disease also rises exponentially with age at least to as old as 85 or 90 years of age. The estimated incidence of dementia among persons 60 to 69 years of age is 0.7 to 3.5 per 1,000 person-years at risk and then doubles approximately every 5 years of age. It is not clear if the incidence of vascular dementia follows a similar exponential increase with age. Women may have a higher incidence of Alzheimer disease in very old age.

A number of epidemiological studies have focused on syndromes such as MCI and cognitive impairment no dementia (CIND). MCI is generally defined as some evidence of cognitive difficulties on neuropsychological testing or cognitive decline greater than expected for a person's age or education but failing to meet criteria for dementia. These individuals generally fail to meet dementia criteria because the cognitive difficulties do not interfere with daily activities. The prevalence of MCI and CIND ranges from 3 to 39 percent among community-dwelling older adults and 6 to 85 percent in clinical and institutional settings. The incidence of MCI in older adults is 8 to 58 persons per 1,000 per year. MCI is associated with a twofold increase in mortality as well as an increased risk for dementia.

More than half of adults with MCI progress to dementia within 5 years, with the rate of progression being 5 to 35 percent per year. Other studies have shown, however, that a significant proportion of older adults with MCI have normal function 1 year later. Research has focused on identifying subtypes of MCI that may be linked to degenerative, vascular, or psychiatric factors and examining which subtypes are more likely to progress to more severe forms of dementia. For example, some research suggests that older adults with amnestic MCI have a high rate of progression to dementia, particularly Alzheimer disease.

Several risk factors for Alzheimer disease, dementia, and cognitive

decline have been identified. A family history of dementia, Down syndrome, or Parkinson disease is a risk factor for dementia. In one family study, the relative risk of developing Alzheimer disease in those with one first-degree relative with dementia was 2.6 in comparison to those with no relative. The relative risk was 7.5 in those with two or more first-degree relatives with dementia. The relative risk associated with one or more relatives with dementia decreased with increasing age of onset, but even in those with onset after 80 years of age, the familial risk was still significant. Studies have found the epsilon 4 allele of apolipoprotein E (ApoE) on chromosome 19 to be a susceptibility gene for the development of Alzheimer disease. Head trauma such as that experienced by those in military service or involved in professional sports such as football has been suggested as another risk factor for Alzheimer disease or dementia. Other possible risk factors include decreased brain reserves, low education and occupational attainment, and reduced mental and physical activity in late life. The risk of dementia may be higher in older adults who use alcohol. Risk factors for vascular dementia include hypertension, diabetes mellitus, and cardiovascular disease.

Several possible protective factors have been identified. Higher education is associated with a decreased risk of Alzheimer disease, perhaps owing to increased cognitive reserves. Regular exercise or physical activity may decrease the risk of cognitive decline. Taking nonsteroidal anti-inflammatory or cholesterol-lowering (statins) drugs has also been shown to be a protective factor against the onset of Alzheimer disease. Some factors, such as estrogen, have been studied as protective factors for the onset of Alzheimer disease, but results have been inconclusive. There is also some evidence that dietary intake of nutrients such as vitamin B_{12} and folate, antioxidants, omega 3 fatty acids, and unsaturated fatty acids may decrease the risk of Alzheimer disease.

A history of moderate to severe depression has been identified as a risk factor for dementia, but it is difficult to establish causality. Depression may be the result of an awareness of early cognitive decline. Depression and dementia may also share common risk factors. Dementia may also be a risk factor for functional decline. Older adults with dementia are at an increased risk of mortality, and survival may be worse in patients with vascular dementia than in patients with Alzheimer disease. The differential diagnosis between depression and dementia in older adults can be challenging, as patients with depression can experience symptoms of cognitive impairment.

Mood Disorders

Depressive symptoms are common among older adults and are associated with functional impairment and decline, mortality, increased service

utilization, and decreased quality of life. Although the mood disorders are a significant health concern for older adults, the prevalence of DSM-5 major depression is low in comparison to that of younger adults.

The lifetime prevalence of any DSM-IV-TR mood disorder was 11.9 percent in the NCS-R among those ages 60 or older, approximately 50 percent of the lifetime risk among younger adults. The lifetime risk for major depression is higher than that for dysthymia and bipolar disorder. Overall, major depressive disorder carries a lifetime risk of 8.2 to 10.6 percent in older adults. Current prevalence estimates are much lower. In the NESARC, the 12-month prevalence of major depression was 2.69 percent among adults 65 or older, lower than that observed in other age groups (6.39 percent among those ages 18 to 29, 5.52 percent among those 30 to 44, and 5.62 percent among those 45 to 64).

The 1-month prevalence of any mood disorder in community-dwelling adults 65 years of age or older in the ECA surveys was 2.5 percent. The prevalence of major depression was 0.7 percent (0.9 percent in women and 0.4 percent in men). The prevalence of dysthymia was 1.8 percent (2.3 percent in women and 1.0 percent in men). Overall, the prevalence of major depression in community studies of older adults ranges from 0.9 to 9.4 percent depending on the time period assessed.

The prevalence of dysthymia is approximately 2 percent among older adults. Dysthymic disorder was replaced in DSM-5 by the diagnosis "persistent depressive disorder" which also includes chronic depression. Overall, older adults with dysthymia generally report late age of onset, no major depression, and a low family risk of mood disorders. Patients often have experienced stressors such as bereavement. Older age is significantly associated with the prevalence of DSM-5 chronic depressive disorder.

Most elderly adults with depressive symptoms in the community do not meet criteria for major depression or dysthymia. In the Longitudinal Aging Studies Amsterdam (LASA), a longitudinal study of 3,056 community-dwelling older adults 55 to 85 years of age, the prevalence of *minor depression*, defined as positive for depression by the CES-D but not meeting DSM-III-R criteria for major depression, was 12.9 percent. Across various epidemiological studies, the estimated prevalence of minor depression or clinically significant depressive symptomatology among community-dwelling older adults ranges from 1 to 49 percent.

The prevalence of major depression is generally higher in clinical samples than that reported from community samples. The prevalence of major depression in older adults in medical care facilities is approximately 5 to 10 percent. In institutional settings, the prevalence of major depression is estimated in some studies to be more than 42 percent. Older persons with depressive symptoms generally do not seek medical help, so the prevalence of depressive symptoms in primary care may underestimate the true prevalence of clinically significant depressive symptoms not meeting

criteria among older adults.

In the ECA, the incidence of major depression among adults age 65 or older was estimated to be 1.25 per 100 person-years of risk, with a higher incidence among older women (1.48 per 100 person-years) in comparison to men (0.90 per 100 person-years). The incidence of first onset of major depression declines with age, especially for men. In the NESARC, the 3-year incidence of major depression was 3.3 percent among those age 60 or older.

Epidemiological studies have confirmed that, although major depression is less common among older adults than among younger adults, depressive symptoms associated with impairments in functioning may be common in late life and increase with advancing age. This high prevalence of symptoms suggests that the current nomenclature may be less applicable for older adults. Clinically significant depressive symptoms in older adults, although prevalent, also can be serious and chronic. In the LASA study, depressed elderly were followed for 6 years. A total of 32 percent of the sample had a severe, chronic course of depression, whereas 44 percent had an unfavorable but fluctuating course. Those with subthreshold or minor depression had the best outcome, followed by those with major depressive disorder, dysthymia, and, last, double depression, or depression and dysthymia together. Those with subthreshold depression were at a high risk of developing DSM-IV-TR mood disorders.

Major depression and dysthymia as well as depressive symptoms are more common in older women in comparison to men. Racial and ethnic differences in the prevalence of depressive symptoms and major depression have not been consistently observed in older adults residing in the community. Risk factors and correlates of depression in late life include factors such as being widowed or otherwise unmarried, impaired functional status, impaired social support, perceived loneliness, perceived poorer health, and low internal locus of control. Older adults with major depression are more likely to have a family or personal history of major depression. In addition, loss of a partner in the previous year may be associated with risk. Risk factors for depressive symptoms may also include, in addition, fewer years of education, living in an urban area, lower socioeconomic status, having one or more chronic physical illnesses, reporting uncontrolled pain, having MCI, experiencing other recent stressful life events, and personality characteristics, such as neuroticism. In addition to these social and psychological risk factors, depression in late life may be linked to biological risk factors such as underactive serotoninergic neurotransmission and hypersecretion of cortisol.

A “vascular depression” subtype has been observed in older adults. Depressive symptoms in this subtype are associated with white matter hyperintensities in neuroimaging. Vascular depression may have a later age of onset or onset after age 60. Late-onset depression may also be associated

with stressful life events while earlier onset may be associated with a family history of depression.

Bipolar disorder is rare in older adults, with a 12-month prevalence of 0.1 percent and a 1-month prevalence of less than 0.1 percent. The lifetime risk is lower among older adults, 1.0 percent among those age 60 or older in comparison to 3.5 percent among those 45 to 59, 4.5 percent among those 30 to 44, and 5.9 percent among those 18 to 29 years of age. Late-onset bipolar disorder is rare but may be associated with cerebrovascular disease, present with psychotic symptoms, and be less likely to be associated with external stressors.

Major depression and depressive symptomatology are frequently comorbid with anxiety disorders and anxiety symptoms in older adults. Some studies have also found alcohol problems and depression to be frequently comorbid, especially in men. Therefore, major depression and depressive symptoms in late life may be a risk for other psychiatric disorders, or these disorders may place the older adults at a greater risk for depression.

Although disability is a risk factor for late-life depression, depressive symptoms are also a risk factor for functional decline and disability. Depression can also predict cognitive decline. Depression is often present in older patients with neuropsychiatric disorders such as Alzheimer disease or Parkinson disease. Depressive symptoms are associated with an increased risk of mortality in older adults, although this risk is confounded by factors such as physical health and health behaviors, especially alcohol and tobacco use. Finally, major depression is a risk factor for suicide in older adults.

Anxiety Disorders

Epidemiological studies consistently show that anxiety disorders are less prevalent in older adults in comparison to younger adults. Yet, the prevalence of anxiety disorders in this older age group is high when compared to that of other disorders. In the NCS-R, the lifetime prevalence of any DSM-IV-TR anxiety disorder was 15.3 percent among those age 60 or older. Of the disorders assessed, specific phobia (7.5 percent) and social phobia (6.6 percent) had the highest lifetime prevalences. The lifetime prevalence of other disorders in this older age group included generalized anxiety disorder (GAD) (3.6 percent), posttraumatic stress disorder (PTSD) (2.5 percent), panic disorder (2.0 percent), agoraphobia without panic (1.0 percent), and obsessive-compulsive disorder (OCD) (0.7 percent). The lifetime prevalence for each individual anxiety disorder assessed was the lowest in this older age group in comparison to that in other age groups, suggesting that as the number of older adults increases the number with a lifetime prevalence of any anxiety disorder will also increase. Current prevalence estimates are lower than lifetime prevalence estimates. The

current (1-month) prevalence of any DSM-IV-TR anxiety disorder estimated from the Australian National Health and Well-Being Survey was 1.3 percent. In the ECA surveys, anxiety disorders were the most prevalent psychiatric disorders (excluding the dementias) in persons age 65 or older. The 1-month prevalence of any anxiety disorder among noninstitutionalized older adults in the ECA was 5.5 percent, more than two times the prevalence of any mood disorder in this age group. The prevalence was lower in older men (3.6 percent) in comparison to that in older women (6.8 percent). The ECA data did not include estimates of generalized anxiety disorder, which suggests that the overall prevalence may have been underestimated. In addition, the ECA estimated the prevalence of anxiety disorders based on DSM-III criteria, whereas the later studies used DSM-IV-TR criteria. It is important to note that the diagnostic criteria for generalized anxiety disorder changed between DSM-III and DSM-IV-TR, which may account for another difference in estimates across studies. In general, among community studies of older adults, the overall prevalence of any DSM-IV-TR anxiety disorder ranges from 0.7 to 18.6 percent, depending on how *caseness* is defined. Similar to what has been observed in depression, the prevalence of anxiety symptoms in community-dwelling older adults is higher than the prevalence of anxiety disorders and may be as high as 24 percent.

The prevalence of anxiety disorders in institutionalized adults is higher than that observed among community-dwelling elders. In a survey in Edmonton, Canada, the prevalence of any anxiety disorder in older institutionalized women was 7.1 percent, compared to 1.4 percent in women in the community, and the prevalence was 4.1 percent in institutionalized men, compared to 2.7 percent in older men in the community. The prevalence of anxiety symptoms among older adults in clinical samples may exceed 40 percent.

Phobic Disorders

The ECA provided data on the prevalence of any phobic disorder as well as phobic subtypes based on DSM-III criteria. Later work has focused more on the prevalence of specific subtypes of phobia. In the ECA, the 1-month prevalence of any phobic disorder was 4.8 percent. The prevalence was twice as high in older women (6.1 percent) in comparison to that in men (2.9 percent). The 1-year prevalence of any phobic disorder in the ECA was twice as high in older African-American men as that in white and Hispanic men. The prevalence was also higher among African-American women in comparison to that in older white and Hispanic women. In community studies of older adults, the prevalence of any phobic disorder ranges from 0 to 12 percent, with the prevalence consistently higher in women. This wide range is again due primarily to case definition as well as to which phobic subtypes are assessed (usually whether specific phobia is included).

The incidence of any phobic disorder in older adults is estimated to be 4.29 cases per 100 person-years of risk, which is similar to the incidence in younger age groups. In the ECA, incidence was more common in older women (5.52 per 100 person-years) than that in men (2.66 per 100 person-years). Some studies have suggested that, although specific or simple phobia usually has an onset in younger children or adults, agoraphobia can often have its onset in late life. Phobic disorder overall appears to have an earlier age of onset in women. Most cases of phobic disorder among older adults are new cases and are not in individuals with a lifetime diagnosis of phobic disorder. In the NESARC, the 3-year incidence among those ages 60 or older was 1.6 percent for generalized anxiety disorder, 1.4 percent for specific phobia, 0.8 percent for panic disorder and 0.6 percent for social phobia.

Specific Phobia. The most common phobic subtype is specific phobia. In the NESARC, the 12-month prevalence of specific phobia among adults 65 years of age or older was 7.5 percent. Unlike other disorders, this prevalence did not really differ from those among younger adults (8.0 percent among those 18 to 29, 7.6 percent among those 30 to 44, and 7.4 percent among those 45 to 64 years). The lifetime estimates were not significantly higher than the current prevalence estimates. The lifetime estimates of specific phobia range from 6 to 7.5 percent across studies and are lower than that observed in younger age groups. Specific phobia is more common in females than males.

Social Phobia. Older adults also have a decreased lifetime risk of social phobia or social anxiety disorder in comparison to that of younger age groups. In the NESARC, the lifetime prevalence of social anxiety disorder among those 65 years of age or older was 3.0 percent, compared to 5.6 percent among those 45 to 64 years of age. The same pattern was observed in the NCS-R where the lifetime prevalence of social phobia (6.6 percent) was approximately half the risk of that among adults younger than 60. In contrast to specific phobia, current prevalence estimates of social phobia across all age groups are generally lower than lifetime estimates. The 12-month prevalence of social phobia or social anxiety disorder among those 65 years of age or older ranges from 0.2 to 1.6 percent across studies. Whether there are significant differences by sex is not clear.

Panic Disorder and Agoraphobia. Comparisons of prevalence estimates across studies are made somewhat difficult because of the changes in the nomenclature in the way panic disorder and agoraphobia are defined, whether the emphasis is on agoraphobia with and without panic or panic disorder with and without agoraphobia. Overall, however, both panic disorder and agoraphobia are less common in older adults than either

social phobia or specific phobia. Regardless of how these disorders are grouped, the lifetime prevalence of panic disorder and agoraphobia without panic is lower among those age 60 or older compared to that of younger adults. Current estimates of these disorders are also lower than lifetime estimates across all age groups. In the NESARC, the 12-month prevalence of panic disorder among those ages 65 or older was 0.8 percent (0.7 percent for panic disorder without agoraphobia and 0.1 percent for panic with agoraphobia). Regardless of whether studies look at panic and agoraphobia and panic separately or together, the current prevalence of both disorders is less than 1 percent. The prevalence of both agoraphobia and panic disorder is higher in females.

Epidemiological studies of community-dwelling older adults have found that panic disorder is rare, and most of the few cases identified are in older women. The prevalence of panic disorder is slightly higher among those in institutions (1 percent) than that among community-dwelling elders. The 1-year prevalence of panic disorder in the ECA was higher among older white (0.3 percent) and Hispanic (3.4 percent) women than that in African-American women (0 percent). There were no racial or ethnic differences among men. The incidence of panic disorder in older adults has been estimated to be 0.04 percent per 100 person-years at risk, suggesting that the few cases of panic disorder present in late life had an earlier onset.

Generalized Anxiety Disorder

The lifetime prevalence of generalized anxiety disorder among adults age 65 or older is 2.6 to 3.6 percent, half the lifetime risk seen in younger adults. In epidemiological studies of older adults, the 12-month prevalence of generalized anxiety disorder is lower than the lifetime risk and ranges from less than 1 to 2.4 percent. Older females are over twice as likely as older males to meet criteria for generalized anxiety disorder within the past year. In some studies, generalized anxiety disorder accounts for the majority of the cases of anxiety disorder in older adults, whereas, in other studies, specific phobia is much more common. In the Durham, North Carolina, ECA, the 1-year prevalence of generalized anxiety disorder among adults age 65 or older was estimated to be 2.2 percent, with a higher prevalence in women (2.6 percent) in comparison to that in men (1.8 percent). The prevalence was higher in older white men (1.5 percent) than that in African-American men (0.9 percent), whereas there were no racial or ethnic differences for women. In the Durham ECA, approximately 3 percent of those with generalized anxiety disorder had an onset of the disorder at age 65 or older. Later studies have confirmed that generalized anxiety disorder is more than twice as prevalent in older women in comparison to older men.

Obsessive-Compulsive Disorder

The overall estimated lifetime prevalence of obsessive-compulsive disorder from epidemiological studies is estimated to be less than 0.1 to 0.8 percent. The 1-month prevalence of obsessive-compulsive disorder in older adults in the ECA was 0.8 percent, with a prevalence of 0.7 percent in older men in comparison to 0.9 percent in older women. Among institutionalized older adults, the prevalence of obsessive-compulsive disorder is estimated to be as high as 4.7 percent in women and 1.4 percent in men. The annual incidence in those ages 65 or older is less than 1 case per 100 person-years but is estimated to be lower in late life in men than at earlier ages, whereas, in women, the incidence is slightly higher in those age 65 or older.

Posttraumatic Stress Disorder

In the CPES, the lifetime prevalence of PTSD among adults age 55 or older was 3.7 percent. Of those, 1.8 percent had PTSD that persisted into late life. Older adults with persistent or pre-late life PTSD were more likely to have a disability than adults without a history of PTSD.

Risk Factors for Anxiety Disorders

Across all age groups, anxiety disorders are more prevalent in women than men, but these gender differences are less apparent in the older age groups. Risk factors for the onset of anxiety disorders or anxiety symptoms in late life include having another anxiety disorder or other psychiatric disorder, neuroticism, being female, experiencing a traumatic life event, and having an impaired social network. Having any anxiety disorder is also associated with having one or more physical disorders in older adults. In the NCS-R, the presence of a 12-month anxiety disorder among those age 65 or older was associated with female sex, being unmarried, having fewer years of education, and having three or more chronic conditions. Anxiety symptoms, generalized anxiety disorder, and agoraphobia can be associated with decreased quality of life in older adults. Social phobia and specific (simple) phobia in older adults are unlikely to be associated with severe anxiety symptoms, and these phobic subtypes are therefore less likely to lead to decreased functioning.

Data from epidemiological studies suggest, however, that comorbidity between depressive and anxiety disorders is common in late life. Depression has been observed in patients with phobic disorder and generalized anxiety disorder and among residents of long-term–care facilities with anxiety symptoms. Anxiety disorders are also frequent among patients with diagnosed depression and depressive symptoms. Anxiety symptoms are also prevalent in persons with dementia, suggesting that one reason for the decline in prevalence of anxiety disorders with age is the inability of older cognitively impaired subjects to participate in studies. There is less comorbidity of anxiety and medical illness and anxiety

and alcohol abuse in older adults than is seen in those who are younger.

Schizophrenia

Although the prevalence of schizophrenia in older adults is low compared to that of other psychiatric disorders, this small number of adults has complex clinical and service needs. With the number of adults over age 65 projected to increase over the next 50 years, older adults with schizophrenia will increase in number and present particular challenges as they will have spent less time in mental institutions than earlier generations of older adults.

The current prevalence of schizophrenia in community samples in those age 65 or older can range from 0.1 to 0.5 percent. In the ECA surveys, the 1-month prevalence of schizophrenia in noninstitutionalized adults age 65 or older was 0.1 percent, compared to 0.7 percent in those 18 to 24 years of age, 0.9 percent in those 25 to 44 years of age, and 0.4 percent in those 45 to 64 years of age. The prevalence was the same in women and men among older adults. The prevalence was slightly higher in older African Americans (0.3 percent) in comparison to that of whites (0.1 percent). The overall lower prevalence reported by those ages 65 or older may be due in part to fewer persons with the disorder surviving to an older age. Also, the ECA prevalence estimates were based on DSM-III criteria, which specified that the illness had to begin before 45 years of age. In the 1980s, there was evidence that some individuals with schizophrenia become ill in middle age or late life. The DSM-III-R included a separate category for persons who met the symptom criteria but the age of onset was 45 years or later. The DSM-IV-TR and DSM-5 do not differentiate persons by age of onset but note differences among those individuals with a later onset compared to earlier onset.

More recent research has provided updated prevalence estimates. For example, in a study conducted in Finland, the lifetime prevalence of all nonaffective psychotic disorders among those age 65 or older was 2.32 percent (1.71 percent for men and 2.67 percent for women). The lifetime prevalence of schizophrenia in this age group was 0.92 percent (0.78 percent for men and 1.00 percent for women). The 12-month prevalence among mental health patients age 60 or older in an Amsterdam Catchment Area was 0.71 percent, subdivided to 0.55 percent for schizophrenia, 0.10 percent for schizoaffective disorder and 0.03 percent for delusional disorder. In this group, the prevalence was 0.35 percent for early onset, 0.14 percent for late onset, and 0.05 percent for very late onset schizophrenia like psychosis.

As noted for other disorders, the prevalence of symptoms of suspicion or paranoid symptoms in older adults is much higher, between 4 and 10 percent. This prevalence is even higher when older adults in institutions or selected residential settings are included and may be as high as 25 percent.

Incidence rates differ by gender. The incidence is greater in men up to age 39 years and then greater in women over the 50 to 70 years age group. The incidence of late-onset schizophrenia has been estimated at 12.6 per 100,000 persons per year.

Multiple studies have found a later age of onset of schizophrenia among women and an overrepresentation of women in the late-onset group, with an even higher frequency of late-life paranoid symptoms in comparison to men (reflecting the old classification of paraphrenia). In addition, late-onset patients are more likely than early onset patients to have shown good premorbid educational, occupational, and social functioning. Those with a later onset tend to have a greater frequency of paranoid symptoms but fewer negative symptoms, less thought disorder, and better neuropsychological performance in comparison to those with an earlier onset of schizophrenia. Risk factors for late-onset schizophrenia may include psychosocial factors, such as retirement, bereavement, and decline in physical functioning. Relatives of late-onset schizophrenia patients have been reported to have a lower risk of schizophrenia than relatives of patients with an earlier onset of disease, whereas other studies have not found differences in family history between the two groups. Some studies have reported that patients with very late life onset of the disease may have a high prevalence of sensory deficits.

Schizophrenia in late life may be comorbid with other psychiatric illness, such as depression. The prevalence of psychotic symptoms in older patients with depression may be as high as 45 percent. Paranoid symptoms are also not uncommon in persons with dementia. Studies have reported that the prevalence of paranoid symptoms may range from 12 to 19 percent among older adults with dementia in comparison to 1 to 3 percent among those without dementia.

Studies looking at the outcome of schizophrenia in older adults suggest that men may be more severely ill in the first decade of the illness but then improve, whereas women may initially have mild symptoms but become more ill as they age. Older schizophrenia patients experience relatively greater age differences in cognitive functioning than healthier individuals. Older adults with schizophrenia and severe cognitive impairment may have poorer outcomes than those without cognitive problems. Older adults with schizophrenia are also at risk for physical illness. They may underreport symptoms or be less compliant with treatment regimens for the medical illness, and cognitive deficits may affect symptom reporting. The prevalence of successful aging in older adults is significantly lower among adults with schizophrenia compared to healthy older adults, due in part to poorer physical health in the schizophrenia group.

Adverse outcomes of schizophrenia symptoms have been noted. For example, in a longitudinal study of 347 85-year-old community-dwelling Swedish elders, the prevalence of any psychotic symptom was 10 percent,

the prevalence of hallucinations was 7 percent, and the prevalence of delusions was 6 percent. Hallucinations were associated with major depression, disability, and visual deficits. Delusions were associated with disability, and paranoid ideation was associated with visual deficits. Hallucinations, delusions, and paranoid ideation were each associated with incident dementia. In addition, hallucinations and paranoid ideation were associated with increased mortality in older women but not in older men.

Alcohol Use Disorders

In the NESARC, the 12-month prevalence of DSM-IV alcohol abuse among adults age 65 or older was 1.2 percent, whereas the prevalence of alcohol dependence was 0.2 percent. These estimates were lower than those for younger adults. These 12-month estimates are also significantly lower than the lifetime estimates for this older age group—12.7 percent for alcohol abuse and 3.4 percent for alcohol dependence. Both of these lifetime estimates reported from the NESARC are higher than those observed in the NCS-R for adults age 60 or older (6.2 percent for alcohol abuse and 2.2 percent for alcohol dependence). Among adults age 50 or older in the United States, the estimated 12-month prevalence of alcohol abuse and dependence in men was 2.9 percent and 2.1 percent, respectively, in comparison to 0.7 percent for both abuse and dependence in older women. The prevalence of subthreshold alcohol use disorders was higher—8.4 percent in men and 4.4 percent in women. In the NESARC, the 3-year incidence among those age 60 or older was 1.0 percent for alcohol abuse and 0.4 percent for alcohol dependence.

Many persons with a lifetime history of alcohol use and dependence may not survive to an older age. In addition, there may be a cohort effect. That is, social norms have changed over the years. Today's oldest old adults may have consumed less alcohol throughout their lives, whereas the use of alcohol has become more acceptable to those adults growing up in later decades. With the ageing of the population, the number of older adults with alcohol use disorders is expected to increase.

As with some of the other disorders, the diagnostic criteria may be less applicable for older adults, resulting in an underestimate of the burden of alcohol-related problems in this population. For example, middle-aged and older adults show an elevated prevalence of binge drinking that does not meet criteria for an alcohol use disorder. Older adults have an increased sensitivity to the effects of alcohol due to decreases in lean body mass, which results in higher blood alcohol concentration per level of alcohol consumed. Some research has suggested that the recommended limits of alcohol use for older adults should be set lower than those for younger adults. Therefore, the quantity of alcohol reported consumed may not be a good indicator of alcohol problems in late life.

Screening criteria for younger adults may be less relevant to older

adults. For example, those who no longer drive will not be stopped for driving while impaired, those who are retired will not experience work problems, and those who are widowed will not experience marital conflicts related to alcohol. It can be difficult to measure alcohol-related changes in social functioning and social decline in older adults because many older adults have cut back on activities for reasons unrelated to alcohol use. In older adults, therefore, attention must be focused on alcohol-related problems and drinking in excess of standardized limits for older adults. Even changing the case definition is still somewhat problematic for older drinkers, because the diagnosis relies on self-report. Many older adults may not reliably report their use of alcohol. Although the prevalence of alcohol problems in older adults is generally higher in hospital emergency department samples and among hospital admissions, relying on case identification in clinical settings is also problematic. Alcohol problems may be masked by symptoms of physical illness. Physicians are less likely to diagnose alcoholism in an older patient than a younger patient, and even when cases are diagnosed, elderly patients are less likely than younger patients to be referred for treatment. Health care workers may perceive excessive alcohol consumption as understandable given age and circumstances.

Recognizing these limitations, research studies have shown that alcohol consumption decreases with age. However, alcohol use is not uncommon in late life, and over half of older adults report alcohol use within the past year. The prevalence of alcohol problems or excessive consumption is higher than the prevalence of DSM-IV-TR alcohol use disorder and estimated to be between 1 and 20 percent in community samples and 14 and 44 percent in hospital samples. The prevalence of heavy drinking varies according to how stringent the criteria are. When *heavy drinking* is defined as consuming five to seven drinks per week, the prevalence among older adults is 11 to 25 percent, whereas the prevalence is 7 to 8 percent when heavy drinking is defined as 12 to 21 drinks per week. The prevalence in the United States also varies by geographic region, with more heavy drinking in the Northeast than in the South.

Researchers have identified two groups of older adults who drink in excess: An *early-onset* alcoholism, which includes long-term alcoholics who survive into old age, and a *late-onset* alcoholism, which includes those older adults with a recent onset of problem drinking. Those with an early onset are more likely to have a comorbid psychiatric disorder. Those who begin to drink heavily in old age appear to have different social and medical characteristics. Persons with a late onset of alcohol problems may increase their drinking in response to stressful life events, such as retirement and bereavement. Other identified risk factors for late-onset alcoholism include a history of alcohol use, being single, separated, or divorced, more opportunities to drink, loneliness and social isolation, boredom,

psychological problems such as depression or anxiety, lack of social support, and physical illness causing pain or insomnia. Approximately one-third of older individuals with drinking problems had a later onset. The incidence of late-onset problem drinking may be as high as 4 percent in the general population. The prognosis for late-onset problem drinkers, however, is generally more favorable than that for early onset problem drinkers. Over the 1-year follow-up period in the ECA across all age groups, for example, the incidence of alcohol abuse declined with age until 60 years of age and then increased, particularly for men 75 years of age or older.

Alcohol-related problems are more prevalent in older men in comparison to women, as are the quantity and frequency of alcohol use. Racial or ethnic differences in alcohol use are less known, although white older adults may tend to use more alcohol and older African Americans and Hispanics may tend to binge drink more frequently. Older adults with lower income consume less alcohol than those with higher income. Physical health problems are associated with past or current alcohol use in older adults. Psychiatric comorbidity is more likely in persons who consume alcohol than in general populations of older adults. Alcoholism is correlated with anxiety and affective disorders and tobacco or drug dependence, and a history of alcohol use has been associated with cognitive impairment. In the Canadian Study of Health and Aging, the occurrence of all types of dementias, except Alzheimer disease, was higher in those with definite or questionable alcohol abuse.

Other adverse outcomes of alcohol use include increased hospitalization, increased incidence of psychiatric disorders, and increased medical disorders. Alcohol use may interfere with the effectiveness of medications. Overall, the prevalence of physical illness is higher in older alcoholics than that among nondrinking older adults. For example, cirrhosis of the liver is one of the leading causes of death in late life. Although some studies have reported increased mortality due to alcoholism in older adults, it is important to note that there can be a protective effect of moderate alcohol use in late life, resulting in a decreased likelihood of mortality.

Suicide

The last half of the 20th century generally saw a decline in suicide rates among older adults but recent decades have shown an increase. In 1950, the death rate from suicide for persons age 65 or older in the United States was 30.0 per 100,000, while in 2010 the rate was 14.9. In 2010, the death rate for suicide was 17.6 per 100,000 persons for those 85 years of age or older, 15.7 for those 75 to 84 years of age, and 13.7 for those 65 to 74 years of age. For those 25 to 64 years of age, rates of suicide ranged from 15.0 among those 25 to 44 years of age to 18.6 among those 45 to 64 years of age. The rates of suicide in 1950 for adults age 45 or older were higher than

those in more recent years for the corresponding age group. As shown in Figure 57.1b–3, there was a decline in suicides in older adults from 1950 to 1980. The suicide rates were higher in 1980s and then lower in 2000 and 2010. These changes, however, were almost totally explained by declines and increases in suicide deaths by older men. The highest rates of suicide in the United States are observed among older men, particularly white men. In 2010, the suicide rates per 100,000 for white males were 26.3 among those 65 to 74 years of age, 34.9 among those 75 to 84 years of age, and 50.8 among those 85 years of age or older. Because many older adults may commit suicide using less noticeable methods, such as medication overdose or cessation of eating, which can be difficult to detect, especially in older adults who live alone, it is likely that suicide rates in this age group are underestimated.

Older adults tend to give fewer warnings of suicidal intent. The prevalence of suicidal ideation among older adults is lower than that observed among younger adults, but the odds of depression among adults with suicidal thoughts is higher among those in the older age group. Older adults invest more planning and preparing for the suicidal act than younger adults and may also use more violent methods. Attempted suicides are less frequent than in younger age groups, and older adults are more successful in their suicide attempts. Among younger adults, suicide attempts occur 200 times as frequently as completed suicides, whereas, in older adults, one out of every four attempts is successful. Overall, the rates of completed suicide are higher in older adults.

Among older adults, suicide is associated with several psychiatric disorders, particularly major depression, but other disorders such as substance use disorder may be associated with suicide. Studies of risk factors for suicide are difficult to conduct, because the suicide rate, although higher in older adults than that in younger adults, is still low. In an analysis of the multisite Established Populations for Epidemiologic Studies of the Elderly (EPESE) data, there were 12 suicides among the sample of 14,456 adults 65 years of age or older followed for 10 years. Suicide was predicted by depression, perceived heath status, poor sleep quality, and absence of a confidant. Other risk factors identified through research studies include personality traits such as neuroticism and stressful life events such as disruption of social ties. Physical illness and functional decline may be risk factors, but their effect may be mediated by depression. The presence of a firearm is also associated with risk of suicide, even after controlling for psychiatric illness.

USE OF HEALTH SERVICES

Older adults with mental disorders are less likely to receive any treatment and, among adults who receive health care treatment, less likely than younger adults to receive mental health specialty treatment. Being in an

older cohort is associated with failure to make initial contact after first onset and delay in making contact among those who eventually seek treatment. Many older adults with psychiatric disorders or clinically significant symptoms remain untreated. If help from a health professional is sought, then it is generally from the individual's primary care provider. Older adults may be less likely to receive referrals from primary care to mental health specialty care than younger adults. Older adults may not bring problems to the attention of providers, and providers may not recognize the significance of clinical symptoms that do not meet criteria for a psychiatric disorder. Disorders or clinically significant symptoms may be more difficult to recognize in older adults with comorbid physical illness. Additionally, older adults may attribute their symptoms to a physical illness. Reasons for nonuse of mental health services in this age group are many, and can include being in a racial or ethnic minority group, being uncomfortable discussing personal problems, being married or cohabitating, being in a middle versus higher income group, having a mild disorder, and having low versus high perceived cognitive impairment. In the CPES, racial differences in mental health service use were noted. Twenty percent of whites compared to 13 percent of blacks reported use. Differences by region were only significant in the South where whites were twice as likely as blacks to use services. Among adults age 65 or older in the CPES, 6.5 percent received some mental health care in the previous year, although 65.9 percent of those with major depression and 72.5 percent with anxiety did not receive care. Seventeen percent of those with a perceived need for mental health care did not receive it. Older age was positively associated with mental health service satisfaction and perceived amount of benefit.

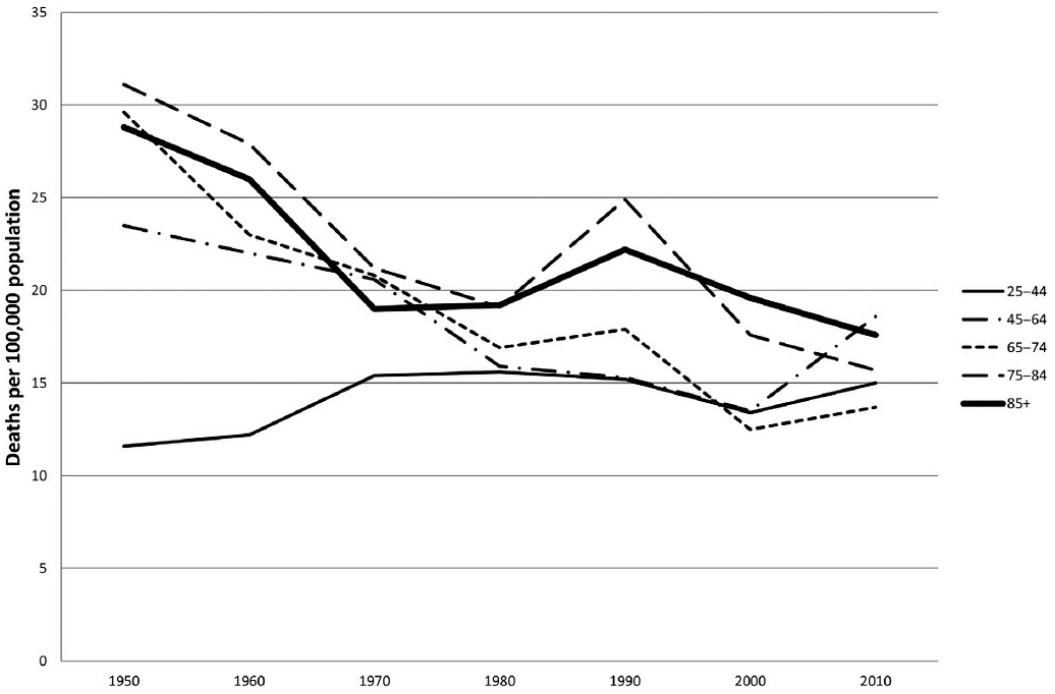

FIGURE 57.1b–3. Death rates per 100,000 for suicide in the United States from 1950 to 2010.

With the projected increase in the number of adults with a psychiatric disorder in this age group, planning for increased use of health services will be important. Training primary care providers to recognize and treat or refer adults with disorders will be an important part of medical education. The public heath focus should include interventions to treat these symptoms and disorders in older adults to reduce the likelihood of negative outcomes.

FUTURE DIRECTIONS

The future of psychiatric epidemiology for older adults will be shaped by at least three factors. First, work on the DSM-5 has been completed, and this iteration of the nomenclature is more evidence based. Changes have been included in the Neurocognitive Disorders section which now includes Mild Neurocognitive Disorder to be distinguished from Major Neurocognitive Disorder. Therefore, the use of existing epidemiological data for secondary data analysis could provide critical assessment of the criteria for disorders based on the natural clustering of symptoms. Statistical methods such as item response theory could be quite valuable and relatively inexpensive as a means for informing this nomenclature. In addition, existing assessment methods in a number of both cross-sectional and longitudinal data bases provide opportunities for analyses of the usefulness of these criteria both cross-sectional and longitudinal. For example, many surveys of older adults, both community-based and from Alzheimer's Disease Centers, provide a wealth of data on the initial cognitive deficits and progression of the dementias.

Second, psychiatric epidemiology in the future must incorporate biological, especially genetic, markers. For example, though DSM-5 does not include specific use of markers as yet, the widespread and relatively inexpensive approaches to exploring the genome and other markers, such as tau in neurocognitive disorders associated with Alzheimer disease provide the opportunity to obtain extensive individual-specific information about the neuropathology of the disease over time. Yet the complexity of psychiatric disorders almost invariably results from an interaction of the genome and the social environment. This interaction can no better be explored than in clinical and community-based epidemiological studies, often genetic add-on studies to existing epidemiological surveys. This interaction may be especially beneficial, for example, in exploring the frequency, distribution, and complex causes of vascular depression later in life.

Following upon the second factor, psychiatric epidemiologists working with older adults must be better able to define and explore the social

environment of older adults. Recent studies have distinguished individual social risks (such as socioeconomic disadvantage) from community-based factors such as poverty and crime rates as risks for late-life depression. Yet better measures of the individual as well as neighborhoods will advance such studies. For example, many psychiatric disorders are undoubtedly aggravated by the stress experienced by older adults who are isolated in unsafe neighborhoods.

Psychiatric epidemiologists have many opportunities to advance the extensive work reviewed in this section, and such opportunities will be critical given the ageing of society.

References

Aranda MP, Chae DH, Lincoln K, Taylor RJ, Woodward AT, Chatters LM. Demographic correlates of DSM-IV major depressive disorder among older African Americans, Black Caribbeans, and non-Hispanic Whites: results from the National Survey of American Life. *Int J Geriatr Psychiatry*. 2012;27(9):940–947.

Beaudreau SA, O'Hara R. Late-life anxiety and cognitive impairment:a review. *Am J Geriatr Psychiatry*. 2008;16(10):790–803.

Blazer DG 2nd, Hybels CF. Origins of depression in later life. *Psychol Med*. 2005;35(9):1241–1252.

*Blazer DG, Wu LT. The epidemiology of alcohol use disorders and subthreshold dependence in a middle-aged and elderly community sample. *Am J Geriatr Psychiatry*. 2011;19(8):685–694.

Byers AL, Arean PA, Yaffe K. Low use of mental health services among older Americans with mood and anxiety disorders. *Psychiatr Serv*. 2012;63(1):66–72.

Byers AL, Covinsky KE, Neylan TC, Yaffe K. Chronicity of posttraumatic stress disorder and risk of disability in older persons. *JAMA Psychiatry*. 2014;71(5):540–546.

*Byers AL, Yaffe K, Covinsky KE, Friedman MB, Bruce ML. High occurrence of mood and anxiety disorders among older adults: The National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2010;67(5):489–496.

*Chou KL, Mackenzie CS, Liang K, Sareen J. Three-year incidence and predictors of first-onset of DSM-IV mood, anxiety, and substance use disorders in older adults: Results from Wave 2 of the National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry*. 2011;72(2):144–155.

Cohen CI, Cohen GD, Blank K, et al. Schizophrenia and older adults. An overview: directions for research and policy. *Am J Geriatr Psychiatry*. 2000;8(1):19–28.

Conwell Y, Duberstein PR, Caine ED. Risk factors for suicide in later life. *Biol Psychiatry*. 2002;52(3):193–204.

Cooper C, Sommerlad A, Lyketsos CG, Livingston G: Modifiable predictors of dementia in mild cognitive impairment: a systematic review and meta-analysis. *Am J Psychiatry*. 2015;172:323–334.

Dennis M, Baillon S, Brugha T, Lindesay J, Stewart R, Meltzer H. The spectrum of suicidal ideation in Great Britain: comparisons across a 16–74 years age range. *Psychol Med*. 2007;37(6):795–805.

Devanand DP. Comorbid psychiatric disorders in late life depression. *Biol Psychiatry*. 2002;52(3):236–242.

Devanand DP. Dysthymic disorder in the elderly population. *Int Psychogeriatr*. 2014;26(1):39–48.

Eaton WW, Kramer M, Anthony JC, Dryman A, Shapiro S, Locke BZ. The incidence of specific DIS/DSM-III mental disorders: data from the NIMH Epidemiologic Catchment Area Program. *Acta Psychiatr Scand*. 1989;79(2):163–178.

*Eden J, Le M, Maslow K, Blazer D, eds. *The Mental Health and Substance Use Workforce for Older Adults: In Whose Hands*? Washington, DC: National Academies Press; 2012.

Ford KL, Bryant AN, Kim G. Age differences in satisfaction with and perceived benefit from mental health services: results from the collaborative psychiatric epidemiology surveys. *Int J Geriatr Psychiatry*. 2013;28(8):831–840.

Forrester SN, Gallo JJ, Smith GS, Leoutsakos JS. Patterns of neuropsychiatric symptoms in mild cognitive impairment and risk of dementia. *Am J Geriatr Psychiatry*. 2016;24:117–125.

Garrido MM, Kane RL, Kaas M, Kane RA. Use of mental health care by community-dwelling older adults. *J Am Geriatr Soc*. 2011;59(1):50–56.

Gonzalez H, Tarraf W, Whitfield K, Vega WA. The epidemiology of major depression and ethnicity in the United States. *J Psychiatr Res*. 2010;44(15):1043–1051.

Grant BF, Goldstein RB, Saha TD, et al. Epidemiology of DSM-5 alcohol use disorder—results from the National Epidemiologic Survey on Alcohol and Related Conditions III. *JAMA Psychiatry*. 2015;72:757–766.

Grant BF, Hasin DS, Blanco C, et al. The epidemiology of social anxiety disorder in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry*. 2005;66(11):1351–1361.

Grant BF, Hasin DS, Stinson FS, et al. Prevalence, correlates, co-morbidity, and comparative disability of DSM-IV generalized anxiety disorder in the USA: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Psychol Med*. 2005;35(12):1747–1759.

Grant BF, Hasin DS, Stinson FS, et al. The epidemiology of DSM-IV panic disorder and agoraphobia in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry*. 2006;67(3):363–374.

*Gum AM, King-Kallimanis B, Kohn R. Prevalence of mood, anxiety, and substance-abuse disorders for older Americans in the National Comorbidity Survey-Replication. *Am J Geriatr Psychiatry*. 2009;17(9):769–781.

Hasin DS, Goodwin RD, Stinson FS, Grant BF. Epidemiology of major depressive disorder: Results from the National Epidemiologic Survey on Alcoholism and Related Conditions. *Arch Gen Psychiatry*. 2005;62(10):1097–1106.

Hasin DS, Stinson FS, Ogburn E, Grant BF. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Arch Gen Psychiatry*. 2007;64(7):830–842.

Howard R, Rabins PV, Seeman MV, Jeste DV. The International Late-Onset Schizophrenia Group. Late-onset schizophrenia and very-late-onset schizophrenia-like psychosis: An

international consensus. *Am J Psychiatry*. 2000;157(2):172–178.

Ibrahim F, Cohen C, Ramirez PM. Successful aging in older adults with schizophrenia: prevalence and associated factors. *Am J Geriatr Psychiatry*. 2010;18(10):879–886.

Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):593–602.

Kim J, Choi NG. Twelve-month prevalence of DSM-IV mental disorders among older Asian Americans: comparison with younger groups. *Aging Ment Health*. 2010;14(1):90–99.

Kim G, Parton J, DeCoster J, Bryant A, Ford KL, Parmelee P. Regional variation of racial disparities in mental health service use among older adults. *Gerontologist*. 2012;53(4):618–626.

Langa KM, Larson EB, Karlawish JH, et al. Trends in the prevalence and mortality of cognitive impairment in the United States: is there evidence of a compression of cognitive morbidity? *Alzheimers Dement*. 2008;4(2):134–144.

Loewenstein DA, Czaja SJ, Bowie CR, Harvey PD. Age-associated differences in cognitive performance in older patients with schizophrenia: a comparison with healthy older adults. *Am J Geriatr Psychiatry*. 2012;20(1):29–40.

Meesters P, de Haan L, Comijs HC, et al. Schizophrenia spectrum disorders in later life: Prevalence and distribution of age at onset and sex in a Dutch catchment area. *Am J Geriatr Psychiatry*. 2012;20(1):18–28.

Miloyan B, Byrne GJ, Pachana NA. Threshold and subthreshold generalized anxiety disorder in later life. *Am J Geriatr Psychiatry*. 2015;23:633–641.

Murphy JA, Byrne GJ. Prevalence and correlates of the proposed DSM-5 diagnosis of chronic depressive disorder. *J Affect Disord*. 2012;139(2):172–180.

Oslin DW. Late-life alcoholism: Issues relevant to the geriatric psychiatrist. *Am J Geriatr Psychiatry*. 2004;12(6):571–583.

Palmer BW, McClure FS, Jeste DV. Schizophrenia in late life: findings challenge traditional concepts. *Harvard Rev Psychiatry*. 2001;9(2):51–58.

Panza F, D'Introno A, Colacicco AM, et al. Current epidemiology of mild cognitive impairment and other predementia syndromes. *Am J Geriatr Psychiatry*. 2005;13(8):633–644.

Plassman BL, Langa KM, Fisher GG, et al. Prevalence of dementia in the United States: the aging, demographics, and memory study. *Neuroepidemiology*. 2007;29(1–2):125–132.

Plassman BL, Langa KM, Fisher GG, et al. Prevalence of cognitive impairment without dementia in the United States. *Ann Intern Med*. 2008;148(6):427–434.

Regier DA, Boyd JH, Burke JD Jr, et al. One-month prevalence of mental disorders in the United States. Based on five epidemiologic catchment area sites. *Arch Gen Psychiatry*. 1988;45(11):977–986.

Reynolds K, Pietrzak RH, Mackenzie CS, Chou KL, Sareen J. Post-traumatic stress disorder across the adult lifespan: findings from a nationally representative survey. *Am J Geriatr Psychiatry*. 2016;24:81–93.

Robins LN, Regier DA, eds. *Psychiatric Disorders in America*. New York: The Free Press; 1991.

Sachdev PS, Lipnicki DM, Crawford J, et al. Risk profiles for mild cognitive impairment vary by age and sex: the Sydney memory and ageing study. *Am J Geriatr Psychiatry*.

2012;20(10):854–865.

Steffens DC, Potter GG. Geriatric depression and cognitive impairment. *Psychol Med*. 2008;38(2):163–175.

Stinson FS, Dawson DA, Chou SP, et al. The epidemiology of DSM-IV specific phobia in the USA: results from the National Epidemiologic Survey on alcohol and related conditions. *Psychol Med*. 2007;37(7):1047–1059.

Trollor JN, Anderson TM, Sachdev PS, Brodaty H, Andrews G. Prevalence of mental disorders in the elderly: The Australian National Mental Health and Well-Being Survey. *Am J Geriatr Psychiatry*. 2007;15(6):455–466.

van der Werf M, Hanssen M, Kohler S, et al; RISE Investigators. Systematic review and collaborative recalculation of 133693 incident cases of schizophrenia. *Psychol Med*. 2014;44(1):9–16.

Wang PS, Berglund P, Olfson M, Pincus HA, Wells KB, Kessler RC. Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):603–613.

Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):629–640.

Woodward AT, Taylor RJ, Bullard KM, Aranda MP, Lincoln KD, Chatters LM. Prevalence of lifetime DSM-IV affective disorders among older African Americans, Black Caribbeans, Latinos, Asians and non-Hispanic White people. *Int J Geriatr Psychiatry*. 2012;27(8):816–827.

▲ 57.2 Assessment

57.2a Psychiatric Assessment of the Older Patient

DAVANGERE P. DEVANAND, M.D., AND GREGORY H. PELTON, M.D.

In the elderly, there is interplay between biological, psychological, and social issues in the presentation of mental illness. Therefore, a thorough evaluation, including assessment of the patient's internal and external environment, is necessary for accurate diagnosis and treatment. In neuropsychiatric disorders, localization of signs and symptoms to specific brain regions should also be attempted to improve understanding of the patient's condition and to facilitate clinical management.

The advent of a variety of brain imaging techniques, genetic testing, and other developments have helped enhance diagnostic assessment, but these approaches are not adequate substitutes for comprehensive history taking and clinical assessment of the older patient. Structured and semistructured interviews and rating scales can be used to supplement the diagnostic interview process.

In elderly patients, the fear of aging and its effects are common. They commonly state that since they are getting old, their symptoms, for

example, depression or cognitive impairment, are normal and are to be expected. In such a patient, it is useful to point out that among people in their own age group, only a small minority has their specific condition and that most people function well.

The evaluation should aim at obtaining a comprehensive assessment of the patient's mental and physical state, and input is often required from the family and significant others. The historical background and values of the elderly patient within a social class and culture influence the perception and presentation of illness and disease.

HISTORY TAKING

The principles in gathering history are similar to those in adults generally: establish rapport, make the patient comfortable, and maintain privacy and confidentiality. Older people have lived a long time, have undergone many stresses, are resilient, and expect to be treated with respect.

The elderly patient expects direct inquiry and examination. Hand contact and reaching out and touching the patient are often useful in establishing rapport. A medical approach is often expected, and respected, by the older patient.

The psychiatrist needs to take an active role and involve the family and significant others whenever feasible. The psychiatrist should explain clearly the nature and purpose of the evaluation, and inquire about the patient's complaints and the patient's and family's goals for the evaluation. The patient's limitations may require modification in the interview technique. In the patient with hearing impairment, the psychiatrist may need to move close to the patient's good ear and speak slowly, loudly, and clearly. Cognitively impaired patients may find it very stressful to perform cognitive tasks in the mental status examination, and may refuse to complete the tests. It is important to explain the nature and purpose of the cognitive tests before administering them, and frequent breaks between tests can help the patient complete the cognitive assessment.

Details about medical illnesses and medications taken are important in elderly patients. Information from other physicians, hospitals, and therapists should be obtained whenever needed to supplement the evaluation. In severely impaired patients, particularly patients with dementia, information from caregivers and professionals involved in social support services is useful. Most elderly patients first see their primary care physician for their psychiatric symptoms, and this is the setting where the majority of them are treated. Therefore, primary care physicians commonly refer their patients to a psychiatrist if diagnosis and management become difficult.

The history should include the following components: identifying data (name, age, sex, education, marital status), chief complaint, history of present illness including onset, duration and symptom course over time,

past psychiatric history, medical history and medications taken, alcohol and substance use history, work history and living situation, and family history of psychiatric, neurological, and other medical disorders.

The initial evaluation session may be long in an older patient, especially if extensive information from collateral sources needs to be obtained. Further, a second visit may be necessary to provide enough time to gather all the additional information from tests and procedures conducted in the interim in order to make an informed diagnosis. In a patient with cognitive decline, laboratory tests to identify potential causes (e.g., syphilis, vitamin B_{12} deficiency, hypothyroidism) of cognitive impairment, and neuropsychological testing and structural/functional brain imaging procedures (CT or MRI, SPECT or PET) may need to be ordered at the first visit. In these cases, the second visit becomes an informing visit at which the patient and family receive the opinion about the diagnosis and likely prognosis. Some diagnostically puzzling cases may need more visits and follow-up to help clarify the patient's condition. In particular, the psychiatrist needs to keep an open mind about the possibility that the initial diagnosis may need to be revised based on the patient's clinical course and response to treatment, for example, change in cognitive status after initiation or withdrawal of medications that have an adverse impact on cognition, for example, benzodiazepines or medications with anticholinergic properties.

HISTORY OF PRESENT ILLNESS

The patient should be encouraged to describe history of the current illness and symptoms and history of the current illness in an open-ended interview, with more structured questions following this initial inquiry. The onset and course of symptoms should be evaluated in relation to other changes including loss of loved ones, relation to daily activities, medication usage, response to treatment if applicable, and intercurrent medical illnesses. Many elderly patients consider insomnia and fatigue to be a normal part of the aging process, and these areas may require explicit inquiry. The time course of development of the present illness is very important in elderly patients. In a patient with cognitive decline, a history of gradual deterioration suggests dementia, episodic cognitive deficits with confusion suggest delirium, focal neurological symptoms or signs suggest cerebrovascular disease, and a short, downhill course may indicate a rapidly progressive dementia like Creutzfeldt–Jakob disease.

Ascertaining age and age at onset of symptoms is critical in the elderly patient. There is considerable evidence that patients with late-onset (onset after 50 to 60 years of age) major depression are less likely to have a positive family history of affective disorder and more likely to have evidence of cerebrovascular disease and cerebral atrophy than elderly patients with early-onset depressive illness. Late-onset schizophrenia

differs in some respects from early-onset schizophrenia, with a greater preponderance of females, paranoid subtype, and some differences in neuropsychological test profiles. Patients with early-onset Alzheimer disease (onset in the 40s or 50s) typically have presenilin I or presenilin II gene mutations, whereas late-onset Alzheimer patients usually do not have these gene mutations though some still have a positive family history of Alzheimer disease.

Based on the clinical presentation, inquiry should focus on pertinent symptoms to make a diagnosis of major depression or psychosis, or anxiety disorder or other axis I or axis II disorder. Potentially serious symptoms that should be reviewed with the patient include depressed mood, suicidal thoughts, lack of interest, agitation/aggression and impulsivity, delusions, and hallucinations.

Insomnia is common in the elderly. During the aging process, there is a small, gradual decline in the number of hours slept per night. Therefore, any significant change in sleep pattern should alert the psychiatrist to the presence of an underlying psychiatric disorder, most commonly a depressive illness. Some patients may downplay insomnia, anorexia, and lack of interest as normal features of the aging process, but this constellation of symptoms should trigger a thorough evaluation for the presence of a depressive disorder. Insomnia, apathy, and anhedonia also occur in early dementia, and the presence of these symptoms should lead to a systematic assessment of cognitive abilities. A gradual decrease in libido and sexual functioning occurs in some older adults, but a rapid deterioration in sexual interest should trigger investigation into the underlying cause. Diabetes, hypertension, and the iatrogenic effects of medication are common causes of sexual dysfunction, but so is depression. Somatization is typical of elderly depressed patients, but there are cultural differences in the degree to which this occurs.

Older adults face a variety of stresses: medical illnesses, physical disability, sensory impairments in vision and hearing, loss of family and friends, loss of social support, changes in relationships with children and significant others, alterations in living situation, and loss of income. One or more of these factors may be related to the onset of the clinical syndrome. The death of a spouse is extremely stressful, as is the death of an adult child. Patients with major depression perceive significantly greater negative emotional impact of the same life events than patients with dysthymic disorder and healthy controls, particularly for interpersonal conflicts.

In the elderly, there is a high rate of delirium after admission to a hospital. This may be related to the lower “cognitive reserve” of older compared to young adults. If there is a suspicion of delirium, potential risk factors should be evaluated, including fever, urinary tract infection, urinary retention, or fecal impaction.

The following symptoms need to be asked of patients with cognitive

decline: difficulties in memory or thinking or concentration, forgetting of recent events, forgetting where things are located, problems in finding the right word to say, getting lost in familiar places, and difficulty in managing finances like balancing the checkbook or paying for items in a store. These symptoms may be the first indicators of dementia. However, the isolated symptom of forgetting names is common during the normal aging process and by itself is not indicative of dementia.

Specific inquiry should also be made about pain, visual and hearing impairment, and possible drug toxicity. For patients in institutional settings, the interpersonal interactions with other residents are important.

PSYCHIATRIC HISTORY

Evaluation of psychiatric history should include the course of illness and types of treatment, that is, psychotherapy, medications, and ECT, and history of response or lack of response to these treatments. Treatments that worked well before in the patient, or worked well in family members, should be given priority.

History of depression, anxiety disorders, and psychosis/schizophrenia, and treatments received in prior and current episodes, should be obtained. Personality disorders are less common in elderly depressed patients compared to young adults with depression. Obsessive-compulsive and avoidant subtypes are more common than borderline, histrionic, narcissistic, and antisocial personality disorders in older adults. Leading questions should be asked about alcohol or substance abuse, history of head trauma, infections like meningitis or encephalitis, stroke, seizure disorder, and other neurological disorders.

FAMILY INTERVIEW AND HISTORY

In the cognitively impaired patient, it is advisable to see the patient alone, then the informant (usually a family member) and/or the patient together with the informant. The informant interview provides an independent assessment of the patient's environment, for example, the cognitively impaired patient may have allowed the home to become disorganized or left the stove on with adverse consequences. In severe dementia, interviewing the informant (or informants) both separately and in a joint interview with the patient is often necessary. In these patients, impairment in memory and other cognitive functions may be so severe that nearly all the history needs to be obtained from the informant. Even at the stage of mild cognitive impairment, the informant's report of the patient's function has been shown to be a much stronger predictor of dementia on follow-up than the patient's self-report of functional ability.

Patients may minimize symptoms in the presence of other people, but may be more revealing about symptoms like depressed mood and suicidal ideation when interviewed alone. Paranoid ideation may not be revealed

when family members are also present in the room. A joint interview with the patient and family (or other informant) may become necessary if there are conflicts about the patient's habits and behavior that need to be resolved, for example, the patient wants to continue driving after developing dementia. However, the family member's report may not always be accurate. For example, in patients with dementia, there is consistent evidence that the informant's report that the patient is depressed often reflects the informant's own depression rather than the patient's mood state.

A genogram is helpful in delineating the familial relationships and the important psychiatric, neurological, and other medical illnesses in different generations. Inquiry should be made about a family history of psychiatric hospitalization, unipolar depression or bipolar disorder, schizophrenia, anxiety disorders, alcohol or substance abuse, suicide, Alzheimer disease, stroke, epilepsy, Creutzfeldt–Jakob, Huntington disease, Parkinson disease, and other neurological disorders. Interviewing more than one family member, and preferably more than one generation, improves accuracy regarding family history.

In elderly patients with functional impairment, family and other caregiver support need to be assessed. The time spent by family members, services provided by family members, perception of family support, and tolerance by family members of changes in behavior are elements that need to be assessed and considered in the development of a treatment plan. Services that may be needed include transportation, nursing services, physical therapy, continuous supervision, homemaker and household services, meal preparation, administrative, legal and protective services, financial assistance, and coordination of delivery of services. These issues may be of greater concern to the patient and family than the clinical symptoms, and usually require coordination with nurses, case managers and social workers in the health support system for the older patient.

The patient and family usually find the transition to assisted living or a nursing home to be very difficult. Guilt, frustration, anger, helplessness, isolation, depression, and demoralization on the part of the caregiver can reduce caregiver effectiveness and impact adversely on the patient. This can lead to an accident like falling and fracturing a hip, inappropriate nursing home placement, or, in some instances, elder abuse. It is important to inquire how the caregiver is coping with the stress of having to take care of the patient, and appropriate referral for the caregiver to seek evaluation and treatment may become necessary.

PERSONAL HISTORY

Details should be obtained about the patient's marital status, living situation, educational background, work history, and the nature of relationships important to the patient. Premorbid personality does carry

into old age, but its impact on illness in the elderly varies considerably. Personality itself may change with aging. This change is believed to be environmentally determined, but it could also be based on changes in neurobiology, for example, catecholaminergic neurons in the locus ceruleus decay markedly with aging and this change may be related to the relatively low rate of late-onset anxiety disorders in the elderly population. Religion is often important to the elderly patient and needs to be included in the history.

The interviewer may be embarrassed to ask about sexual history in the elderly patient, but many older adults do have symptoms of sexual dysfunction and want to talk about it. Inquiry about sexual orientation and sexual dysfunction should cover any history of sexually transmitted diseases, erectile dysfunction (including impotence) in men, and anorgasmia in men and women. Sexual dysfunction can be a side effect of medication, particularly SSRIs and related antidepressants. Diabetes and hypertension are also common culprits.

MEDICAL HISTORY

The elderly patient typically suffers from several medical illnesses and takes a large number of medications. Patients may not keep track of all the medicines they are taking and their list often excludes medications obtained over-the-counter or alternative medications sold in health food stores. Therefore, it is important to inquire about all these categories of medications, and it is a good idea for the psychiatrist to ask the patient to bring all the current medications in a bag. Benzodiazepines, steroids, and anticholinergics can lead to cognitive impairment and even delirium; L-dopa and dopamine agonists prescribed for Parkinson disease can lead to psychosis, and the use of some antihypertensives and steroids have been associated with depression.

Medical history should cover all major illnesses, surgery, hospitalizations, trauma, and treatments. A variety of medical disorders can first present with disturbances in mood, cognition, or psychotic features including infections, electrolyte and metabolic imbalances, and cardiovascular and cerebrovascular disease. Depression can lead to weight loss and malnutrition. Dietary history is important in this regard; the older adult living alone may not obtain adequate nutrition and develop vitamin and other deficiencies.

ALCOHOL AND SUBSTANCE ABUSE

Alcohol abuse is not uncommon in the elderly, and explicit inquiry is usually necessary to elicit this history. Even when alcohol consumption has been constant in amount for many decades, it can lead to cognitive decline in susceptible patients as they grow older because of decreased cognitive reserve due to neuronal loss in the brain. Neuropsychiatric complications

of alcohol abuse and dependence, now subsumed under alcohol use disorder in DSM-5, include varying degrees of amnesia, Korsakoff syndrome, and intention tremor and ataxia (cerebellar damage). Other substance use disorders are less common in elderly patients, but some do use marijuana regularly. Even with the advent of new hypnotics, benzodiazepine use remains fairly widespread. This is of particular concern in the elderly because benzodiazepines can contribute to excessive sedation, unsteady gait and falls, and cognitive deficits. Patients with chronic pain, particularly back pain, may be dependent on narcotics.

MEDICAL EXAMINATION

If a medical examination has not been conducted recently by the patient's primary care physician, the psychiatrist needs to complete this component or arrange to have it done by another physician. The neurological examination is particularly important, and needs to include evaluation of sensorium, cognition, motor strength and function, sensation, reflexes, extrapyramidal signs, vision, and hearing. Neurological deficits, endocrine disorders, cardiovascular disease, occult malignancy, and infections are potential causes of cognitive impairment, agitation, depression, and occasionally psychosis. Urinary tract infection, which can occur without fever, is a common underlying cause of cognitive decline and delirium, particularly in the nursing home.

MENTAL STATUS EXAMINATION

The mental status examination is essential. Serial mental status examinations may be necessary, especially if mental functioning fluctuates over time.

OBSERVATION OF BEHAVIOR

The first part of the mental status examination is general observation of the patient's appearance, personal hygiene, motor activity, any unusual behaviors, speech, attitude toward the examiner, and the presence of hearing and visual aids. Poor hygiene suggests dementia, but conversely the demented patient may look well groomed because of the caregiver's efforts.

In older patients, paranoid delusional thoughts occur not only in schizophrenia and delusional disorder but also in dementia, particularly Alzheimer disease, and during the course of delirium. Apathetic, withdrawn behavior is common in both depression and dementia. Disturbances of motor activity are frequent in elderly patients, and bradykinesia with slow gait is common in the 8th to 10th decade of life. A slow shuffling gait with mask-like facies, resting tremor, and rigidity indicate Parkinson disease or extrapyramidal side effects of antipsychotic medication. Elderly patients

are particularly prone to the neurological side effects of antipsychotics. Lip-smacking and involuntary buccofacial movements suggest tardive dyskinesia in patients with prior antipsychotic exposure, though this has become less common since the advent of second-generation antipsychotics. Impulsive, disinhibited behavior with minimal to mild cognitive impairment may be the first signs of frontotemporal dementia, though increasing apathy is, in fact, more common than disinhibited behavior in frontotemporal dementia.

LANGUAGE AND THOUGHT DISORDER

Speech and language are important to assess in older adults. The quality, quantity, and fluency of speech need to be assessed. Speech can be normally responsive, rapid, pressured, loud or emotional, slow, monotonous, lacking spontaneity, or slurred. Aphasia or slurred speech occurs in stroke and in a variety of neurological disorders. Mutism occurs in late stages of primary progressive aphasia, dementia from a variety of causes, and amyotrophic lateral sclerosis. Repetitive speech with perseveration on the same topic is common in patients with dementia, and tends to be secondary to the loss of recent memory that is a hallmark of the illness.

The patient should be able to understand and answer questions appropriately, and think in abstract terms. Paucity of speech with long latency is characteristic of major depression, particularly the melancholic subtype. Tangentiality and looseness of associations are characteristic of schizophrenia. The patient with Alzheimer disease may display repetition, perseveration, and fragmented ideas. In contrast, patients with subcortical dementias like Huntington or Parkinson disease usually do not show these patterns of thought process.

Thought content abnormalities include preoccupations or obsessions, including preoccupation with health problems in the elderly depressed patient. Delusions can be persecutory, grandiose, somatic, erotic, guilty, or nihilistic. Hypochondriacal delusions are common in delusional depression, for example, bowels are not working or have been destroyed. Delusional depression in the elderly is associated with a high suicide risk. Bizarre, fixed delusions are common in schizophrenia, but they are rare in dementia where transient persecutory ideas are more common. Distinguishing fact from fiction can be difficult in the elderly patient, who may have had personal effects stolen and may not have a persecutory delusion as the evaluation may suggest.

PERCEPTION

Perceptual disorders include hallucinations and illusions. Hallucinations can be simple or complex and organized. Transitory hallucinations can occur in dementia, and in patients with hearing or visual deficits. In these

patients, it is important to distinguish an illusion, that is, perceptual distortion of a real sensory stimulus, from a hallucination that may indicate more severe psychopathology. Memory loss and disorientation help to identify the problem as being related to dementia. A number of other perceptual abnormalities occur in a variety of neurological disorders involving damage to specific brain regions. These syndromes include micropsia, which occurs when objects look smaller than their real size, macropsia, which occurs when objects look larger than their real size, visual agnosia, which is the inability to recognize visible objects, prosopagnosia, which is the inability to recognize familiar faces, astereognosis, which is the inability to recognize objects by touch, and autotopagnosia, which is the inability to recognize parts of one's own body.

MOOD AND AFFECT

Mood is the subjective and affect is the objective emotional state. Many elderly patients describe their mood accurately as depressed or anxious, but some tend to deny and suppress it. For example, an elderly woman who is crying may say that she is not depressed. Somatization is common in older patients, particularly gastrointestinal symptoms like constipation. Some patients are not sure if they are depressed because they have begun to lose interest in activities but do not experience a clear-cut depressed mood. Many of these patients have dysthymia or minor depression. Depression is common in a variety of other neuropsychiatric syndromes, including dementia, stroke, and Parkinson disease. Mania is not common in the elderly, but can occur as a primary late-onset mania or secondary to dementia or stroke, or other neurological insult. Labile affect is believed to be typical of vascular dementia, but it remains unclear if labile affect is more common in vascular compared to other forms of dementia. Pseudobulbar affect is uncontrollable laughing or crying in the absence of the corresponding subjective mood experience, and occurs in neurological disorders like amyotrophic lateral sclerosis and frontotemporal dementia, and less commonly in Alzheimer disease. Medication side effects or toxicity need to be considered as possible etiologies for depression. Of note, late-onset anxiety disorders, including panic disorder, are relatively rare.

SUICIDAL AND HOMICIDAL IDEATION

Inquiry should be made about suicidal and homicidal ideation. Suicidal ideation should be probed by first asking whether the patient thinks that life is not worth living. A positive response should lead to follow-up questions about whether there is a suicide plan or history of a suicide attempt. If this line of inquiry indicates suicidal risk, further questions should focus on whether a gun or other means to commit suicide are easily available, and the appropriate clinical intervention with the patient and family members should then be made. Studies consistently show that

elderly males, particularly white males, are the demographic group with the highest risk of committing suicide, and that they often use guns as the method of suicide. Of note, the majority of patients who commit suicide have visited their primary care physician a few days to weeks before they act.

SLEEP

Approximately half of older adults complain of sleep difficulties. Careful assessment of sleep habits, including factors such as alcohol and caffeine intake that can disrupt sleep, is important. Other factors contributing to sleep impairment include age-related changes in sleep architecture and circadian rhythms, sleep disorders including sleep apnea, dementia, depression, other medical illness, and polypharmacy. The impact of treatment, or withdrawal of treatment, for example, benzodiazepines, needs to be assessed in the context of these factors that can underlie sleep disorders.

SENSORIUM AND COGNITION

Evaluation of sensorium and cognition is critical and essential in elderly patients. Whenever there is suspicion of cognitive impairment by history, it is important to administer a screening test like the Mini Mental State Examination (MMSE) or Montreal Cognitive Assessment (MOCA).

At the start of this portion of the examination, the patient should be told that some of the questions are easy and some a little more difficult, and that people score across a wide range in these tests. This instruction is important because the patient with significant cognitive deficits may not agree to complete the test if they feel that they are taking an examination and that their deficits are being judged critically. Skills that can be tested in this component of the evaluation, and which can be assessed more thoroughly with detailed neuropsychological testing, include attention and concentration, learning, recall, recognition after a delay (memory), comprehension, verbal fluency, repetition, and naming (language), visuospatial abilities, calculation, praxis, and the executive functions of initiation, perseveration, response control, synthesis and integration, and changing mental sets (frontal lobe functions).

A patient needs to be able to focus and be attentive in order to complete the tests. If a patient is distractible and unable to concentrate, the subsequent components of the test may not be valid. Educational achievement and the patient's primary language are important variables to consider in the interpretation of cognitive tests. A patient with low education or a patient for whom English is not the primary language is likely to score poorly on some components of the cognitive assessment, but memory is affected less by these factors. Therefore, a patient with low education who scores 25/30 on the MMSE but recalls 3/3 objects at 5

minutes is less likely to be developing dementia than a patient with a postgraduate degree who scores 28/30 on the MMSE but recalls only 1/3 objects at 5 minutes.

COGNITIVE TESTING

A brief evaluation of cognitive ability can be performed with the MMSE. A score of less than 24 out of 30 is considered indicative of a cognitive deficit. A variety of conditions, including major depression, can be associated with low MMSE scores. In patients with mild cognitive impairment, deficits in recall of three objects after a 5-minute delay, and impairment in the orientation items, suggest an increased likelihood of developing dementia. However, the MMSE by itself is not sufficient to make a diagnosis of dementia. In the early stages of dementia, attention is preserved but short-term memory is impaired. In contrast, attention is impaired in depression but short-term memory is relatively intact, that is, what is registered in the brain is generally retained. Delirium is characterized by disorientation, which occurs only in the mid to late stages of most dementing disorders. If the results of the cognitive screening leave the diagnosis in doubt, which is often the case, referral for detailed neuropsychological testing is usually indicated. Neuropsychological testing is useful in making the diagnosis of dementia in its early stages, but its utility in differentiating subtypes of dementia, for example, Alzheimer versus vascular, or in diagnosing depression in patients with dementia, is not as well established.

INSIGHT

Insight refers to the ability of patients to understand their own symptoms and their potential causes. Insight is impaired in most forms of moderate to severe mental illness, and can disappear in severe dementia. Judgment can be assessed during the course of the interview by inquiring about the patient's opinion and proposed course of action in incidents that they describe in their own history, or in a hypothetical incident. Judgment is compromised in psychotic disorders and dementia.

DAILY FUNCTIONING

Assessment of daily functioning is essential in the evaluation of the older patient, unlike the situation in most young adults. Activities of daily life (ADL) can be classified into basic ADL (e.g., eating, dressing, toileting), instrumental ADL (cooking, shopping, laundry), and more complex social and cognitive functions requiring greater mental effort. Basic and instrumental ADL need to be assessed in the elderly patient. The Pfeffer functional activities scale is widely used in large-scale studies of patients with mild cognitive impairment and Alzheimer disease.

NEUROLOGICAL EVALUATION

Psychiatric, cognitive, and neurological examinations overlap considerably in older adults. While patients with manifest neurological disorders need consultation by a neurologist, the geriatric psychiatrist needs to be able to perform the neurological examination because many patients present with syndromes that overlap between the two specialties. In patients who have cognitive disorders or focal neurological symptoms or signs, a neurological examination is essential. All components of the neurological examination are important, but in the elderly patient presenting to a psychiatrist, the areas of frontal lobe function, basal ganglia function (examination for extrapyramidal signs), motor and sensory function, reflexes, coordination, and evaluation of vision and hearing are essential.

RATING SCALES

Although rating scales are used primarily in research studies, they can serve several possible functions in clinical practice: address specific clinical problems and document change, teach, communicate with colleagues, and strengthen the chart or electronic medical record for medicolegal and insurance purposes. The SCID can be administered to make comprehensive DSM diagnoses, but it requires special training and takes too long to be recommended for routine clinical practice. Some of the widely used scales to assess specific domains of psychopathology and cognitive and daily functioning will be summarized here.

Depression

The Hamilton Rating Scale for Depression (HRSD), Montgomery–Åsberg Depression Rating Scale (MADRS), Geriatric Depression Scale, and Center for Epidemiological Studies-Depression Scale (CES-D) have been used in a variety of research studies. The HRSD and the MADRS are particularly useful in patients with melancholic depression. Another widely used instrument is the Beck Depression Rating Scale, a self-report instrument. Epidemiological studies have typically used the self-report CES-D. The Geriatric Depression Scale, which can be self-rated or observer-rated, is also useful. To assess depression in patients with dementia, the Cornell Scale for Depression in Dementia has been developed.

To assess cognitive impairment, a plethora of rating instruments is available. The most widely used instrument is the MMSE, and the MOCA is comparable in its utility. Education and cultural background can affect performance on these tests, which is one reason why the score obtained should be evaluated in the context of the entire clinical picture.

To assess psychosis, the Brief Psychiatric Rating Scale (BPRS) is commonly used in young adults and can also be used in elderly patients. To address the unique features and methods of assessment of psychosis in dementia, a variety of scales have been developed. These include the

Neuropsychiatric Inventory (NPI), the Columbia University Scale for Psychopathology in Alzheimer's Disease (CUSPAD), CERAD Behavior rating Scale, and the Behavioral Pathology in Alzheimer's Disease (BEHAVE-AD) rating scale. The Cohen-Mansfield agitation inventory is available for use in nursing home patients. Currently, the NPI, which uses a time-efficient, decision-tree approach, is the most widely used scale to evaluate the effects of pharmacological treatment of psychosis and agitation in patients with dementia.

To assess the neurological side effects of antipsychotic medications, the Unified Parkinson's Disease Evaluation Scale and the AIMS scale can be used.

COMPETENCE

The psychiatrist may be asked about the patient's level of competence to give consent for medical treatment, give power of attorney, sign or change a will, or manage personal finances. Competency is a legal term, and the criteria for judging it have never been precisely defined. Competency needs to be judged in a specific framework and context, that is, a patient may be competent to execute one task but not another. There are some specific measures to evaluate financial abilities, but testamentary capacity remains a judgment call. Both cognitive abilities and general functioning need to be evaluated before making this judgment.

SYNDROMES SPECIFIC TO GERIATRIC PATIENTS

There are several syndromes common in geriatric patients who do not fit easily into the current diagnostic nomenclature. These syndromes cut across medical and psychiatric diagnoses, are widely recognized, and need to be incorporated into the evaluation. Apathy is common, and it can be difficult to separate from depression. Frailty is a syndrome that is associated with increased morbidity and mortality, and gait and grip strength are important metrics in the evaluation of frailty. Technology-based assessments that can monitor activity levels and sleep to inform diagnosis are now used mainly in research studies, but over time they are likely to be included in the standard clinical assessment of patients. Similarly, telemedicine can be extended to telepsychogeriatrics, particularly given the aging of the population and the need to provide psychiatric assessment and treatment to the homebound elderly.

SUMMARY

The evaluation concludes with a summary of the key findings and differential diagnoses, which leads to a diagnosis. An etiological formulation using the biopsychosocial model is the next step. Finally, these conclusions are built into the treatment plan and management strategies

for the patient.

References

Alexopoulos GS, Abrams RC, Young RC, Shamoian CA. Cornell scale for depression in dementia. *Biol Psychiatry*. 1988;23:271–284.

Brown PJ, Devanand DP, Liu X, Caccappolo E; Alzheimer's Disease Neuroimaging Initiative. Functional impairment in elderly patients with mild cognitive impairment and mild Alzheimer disease. *Arch Gen Psychiatry*. 2011;68(6):617–626.

Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology*. 1994;44:2308–2314.

DeMers S, Dinsio K, Carlson W. Psychiatric care of the older adult: an overview for primary care. *Med Clin North Am*. 2014;98(5):1145–1168.

Devanand DP, Miller L, Richards M, et al. The Columbia University Scale for Psychopathology in Alzheimer's disease. *Arch Neurol*. 1992;49:371–376.

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189–198.

Glover J, Srinivasan S. Assessment of the person with late-life depression. *Psychiatr Clin North Am*. 2013;36(4):545–560.

Karimi S, Dharia SP, Flora DS, Slattum PW. Anticholinergic burden: clinical implications for seniors and strategies for clinicians. *Consult Pharm*. 2012;27(8):564–582.

König A, Aalten P, Verhey F, et al. A review of current information and communication technologies: can they be used to assess apathy? *Int J Geriatr Psychiatry*. 2014;29(4):345–358.

Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist*. 1969;9:179–186.

Lesser RP, Fahn S, Snider SR, Cote LJ, Isgreen WP, Barrett RE. Analysis of the clinical problems in parkinsonism and the complications of long-term levodopa therapy. *Neurology*. 1979;29:1253–1260.

Lin JS, O'Connor E, Rossom RC, Perdue LA, Eckstrom E. Screening for cognitive impairment in older adults: a systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2013;159(9):601–612.

Mackenzie TB, Robiner WN, Knopman DS. Differences between patient and family assessments of depression in Alzheimer's disease. *Am J Psychiatry*. 1989;146:1174–1178.

McKeith IG, Galasko D, Kosaka K, et al. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. *Neurology*. 1996;47(5):1113–1124.

Morris JC, Blennow K, Froelich L, et al. Harmonized diagnostic criteria for Alzheimer's disease: recommendations. *J Intern Med*. 2014;275(3):204–213.

Neikrug AB, Ancoli-Israel S. Sleep disturbances in nursing homes. *J Nutr Health Aging*. 2010;14(3):207–211.

Ramos-Ríos R, Mateos R, Lojo D, Conn DK, Patterson T. Telepsychogeriatrics: a new horizon in the care of mental health problems in the elderly. *Int Psychogeriatr*. 2012;24(11):1708–1724.

57.2b Complementary, Alternative, and Integrative Medicine in Geriatric Psychiatry

HARRIS A. EYRE, M.B.B.S. (HONS.), BERNHARD T. BAUNE, M.D., PH.D., AND HELEN LAVRETSKY, M.D.

INTRODUCTION

Complementary and alternative medicine is "a group of diverse medical and health care systems, practices, and products that are not presently considered part of conventional medicine." Some of these therapies provide promise as novel and effective therapies for treatment and prevention of late-life psychiatric disorders. "Complementary" generally refers to using a nonmainstream approach together with conventional Western medicine. "Alternative" refers to using a nonmainstream approach in place of conventional Western medicine. "Integrative medicine" is another term often used in this CAM field, it aims to combine alternative and complementary medicine with evidence-based Western medicine. From here on, such therapies are referred to as Complementary, Alternative and Integrative Medicine (CAIM). The shift in the societal attitudes to be more inclusive of these practices is reflected in the recent change in the titles of the National Center for Complementary and Alternative Medicine (NCCAM) to the National Center of Complementary and Integrative Health (NCCIH).

CAIM interventions have varying levels of efficacy and evidence, they can include mind–body practices, conventional physical activity, dietary interventions, natural products, as well as body-based practices and other medical system practices. Mind–body practices refer to practices where both physical and mental activity may be combined during training. These can be more exercise-focused (i.e., yoga, tai chi, qigong) or more practice and stress-reduction focused (i.e., mindfulness and relaxation). Physical activity refers to any bodily movement produced by skeletal muscles that requires energy expenditure and may include aerobic, resistance, stretching and toning, or combination activities. Dietary interventions explore the effect of dietary patterns on health and illness and may include Western, traditional, and Mediterranean dietary patterns, as well as caloric restriction, among others. Natural products may include herbs, vitamins, minerals, natural supplements, and probiotics. Body-based practices may include spinal manipulation (e.g., chiropractic, osteopathic medicine), aromatherapy, and massage therapy. Other medical systems include traditional Chinese medicine (including acupuncture), Ayurvedic medicine, or homeopathy. These CAIM interventions are largely based on therapies, which have been utilized in many civilizations and religions for hundreds and thousands of years. It is not since the past 50 to 100 years that empirical science has begun to explore the evidence behind these therapies. For a graphical representation of the CAIM subtypes, refer to Figure 57.2b–1.

There are a number of major reasons why geriatric mental health providers need to recognize the importance of these interventions. (1) *CAIM therapy use is high and rising*. Research suggests 12-month prevalence of any CAIM usage (excluding prayer) in the United States is around 35 to 50 percent; Baby Boomers (adults born from 1946 to 1964) report significantly higher rates of use than the Silent Generation (born from 1925 to 1945) for chronic conditions. (2) *The global population is aging*. The number of older persons (aged ≥60 years) globally is expected to more than double, from 841 million people in 2013 to more than 2 billion in 2050. (3) *CAIM therapies are increasingly cited in clinical guidelines*. It is now commonplace to note the use of therapies like conventional exercise in the management of psychiatric conditions. (4) *CAIM therapies are a source for innovative interventions and have a growing, quality empirical research basis*. Empirical researches into these therapies are exploring their clinical efficacy in trials and the neurobiological mechanism underlying their effects. (5) *CAIM therapies can help to lower utilization of conventional medicines*. CAIM therapies may be used in replacement of conventional medicines for the management of mild illnesses (e.g., relaxation techniques for mild anxiety instead of benzodiazepines which are known to have adverse effects with elderly populations). (6) *CAIM therapies can have significant side effects and CAIM–drug interactions*. With the high use in geriatric populations, this is therefore an important reason why clinicians need to better understand CAIM therapies.

This chapter aims to provide an overview of the importance of understanding CAIM therapies in geriatric psychiatry.

THE EPIDEMIOLOGY OF COMPLEMENTARY AND INTEGRATIVE MEDICINE USE

To appreciate the importance of CAIM therapies in geriatric psychiatry, it is important to explore the epidemiology of, and drivers behind usage. This section will explore these topics in general adult, old age, and psychiatric populations.

CAIM Use in Geriatric Populations

CAIM usage in general adult populations appears to be high, however, there is variability within studies. A recent systematic review of prevalence studies from 1998 onward explored the use of these therapies internationally and included 51 reports from 49 surveys in 15 countries. Estimates of 12-month prevalence of any usage ranged from 9.8 to 76 percent; and from 1.8 to 48.7 percent for visits to health practitioners. Estimates of 12-month prevalence of any usage (excluding prayer) from surveys using consistent measurement methods showed remarkable stability in Australia (49 percent, 52 percent; 1993, 2004) and United

States (36 percent, 38 percent; 2002, 2007). National estimates of CAIM use were highest in East Asian countries such as Japan, South Korea, and Malaysia. Studies consistently show that "average" CAIM users are more likely to be female, better educated, middle-aged, and report poorer health status than nonusers.

Specific data on geriatric populations suggest high use of CAIMs, however, there is variability depending on study populations. A recent study compared CAIM usage between Baby Boomers (adults born from 1946 to 1964; n = 7,734) and the Silent Generation (born from 1925 to 1945; n = 4,682) through analysis of the National Health Interview Survey in the USA. Baby Boomers were more likely to use CAIM within the past year (43.1 percent vs. 35.4 percent) with chronic pain independently associated with greater CAIM use. This data therefore suggest that given Baby Boomers are already using 30 percent more CAIMs than their older counterparts, with the aging Baby Boomer population across the world and associated increases in chronic and painful conditions, CAIM therapies will only continue to increase. By 2030, more than 6 in 10 Baby Boomers will be managing at least 1 chronic disease.

FIGURE 57.2b–1. Types of complementary, alternative, and integrative therapies.

There may be a number of drivers for CAIM use in geriatric populations. The use of CAIMs appears to be influenced by a complex mix of economic and sociocultural factors. First, it appears CAIM use is tied to the movement for self-care. Therefore, CAIM use may provide people with a better sense of control and responsibility over their illness. CAIM therapies such as tai chi and yoga appear to improve patients' quality of life. CAIMs are used in addition to, or in replacement of pre-existing pain medications. CAIM therapies can offer a less harmful management of chronic pain by avoiding unnecessary complexity, polypharmacy, and side effects. In low resource environments where access to biomedical services is poor, there is evidence of a pervasive reliance on traditional healers, even for severe diseases. Delayed medical care—more likely in low resource environments—is associated with CAIM use in the United States. In high resource settings where biomedical services are more readily available, a

significant amount of CAIM therapies are used for health promotion and illness prevention. CAIM therapies are often used as adjuncts to conventional therapies for severe diseases like cancer; and to self-manage chronic health complaints like lower back pain. It appears CAIM usage is driven more so by values and beliefs rather than dissatisfaction with biomedicine. The influence of increasing marketing of CAIM therapies is anecdotal but must be considered.

Persons with Mental Illnesses

Persons with mental illnesses appear to use higher numbers of CAIMs than the general population. In national samples of general adult populations, persons with depression and anxiety symptoms use CAIM therapies more than persons without such disorders. One study indicated that over half of such persons reported CAIM use within the prior year. The presence of psychiatric illness also increases the odds of visiting a CAIM practitioner. Also of significance, the use of CAIM therapies for anxiety and depression was often in conjunction with traditional treatments, raising concerns for how these treatments interact with conventional psychiatric treatments. Relatively less information exists specifically about CAIM use among older adults with mental illnesses. Data from the 2002 National Health Interview Survey indicate that nearly 82 percent of older adults with self-reported depression or anxiety had used CAIM within the prior year (as opposed to 65 percent of older adults without these symptoms). Most of the CAIM use reported by older adults with depression, anxiety, or both was described as being used for nonpsychiatric symptoms. A smaller regional study indicated that older adults in geriatric psychiatry clinics who used CAIM had higher scores for depression, pain, and sleep disturbance, with mind–body practices being particularly associated with depression in older adults.

CAIM MEDICINE IN AGE-RELATED COGNITIVE DECLINE AND DEMENTIA

As of 2013, there were an estimated 44.4 million people worldwide with dementia. This number will increase to an estimated 75.6 million in 2030, and 135.5 million in 2050. Unfortunately for those with Alzheimer disease, pharmacological agents temporarily treat symptoms without having an effect on the underlying pathophysiology of the disease. Conventional treatments for dementia are limited, therefore innovative management strategies are required. There is promising evidence for the use of various CAIM therapies in dementia; alternatively, there is mixed, contradictory, or no evidence for other therapies. This section provides an overview of the salient findings for therapies which show promise for evidence-based use in clinical populations.

Conventional Physical Activity

Conventional physical activity is one of the most promising therapy for the treatment and prevention of age-related cognitive decline and dementia. Importantly, there are a variety of subtypes of physical activity, from aerobic activity, to resistance, to stretching and toning. Currently, the American College of Sports Medicine recommends that exercise programs for older adults include both aerobic and nonaerobic physical activities, such as resistance training, balance training, and stretching, for optimal general health. See Table 57.2b–1 for the latest clinical guidelines on physical activity for older adults.

A recent randomized controlled trial (RCT) examined the efficacy of resistance and aerobic training in the improvement of cognitive functions in subjects with subjective Mild Cognitive Impairment (MCI). MCI—a well-recognized risk factor for dementia—represents a critical window of opportunity for intervening and altering the trajectory of cognitive decline. The study occurred over 6 months and involved 86 community-dwelling women aged 70 to 80 years. Physical activity protocols included twice-weekly resistance training, twice-weekly aerobic training, or twice-weekly balance and toning (BAT) training (i.e., control). Resistance training improved selective attention/conflict resolution and associated memory compared with BAT. In contrast, aerobic physical activity improved general balance and mobility and cardiovascular capacity. The study also found aerobic physical activity improved verbal memory and both resistance and aerobic physical activity improved spatial memory. A recent stringent meta-analysis has examined the efficacy of exercise on cognition in older adults with MCI. MCI was diagnosed on documented criteria or via the Mini-Mental State Exam (MMSE). Fourteen RCTs with 1,695 participants aged 65 to 95 years were utilized. These studies had a duration of 6 to 52 weeks. Overall, 42 percent of effect sizes (ESs) were potentially clinically relevant (ES >0.20) with only 8 percent of cognitive outcomes statistically significant. The meta-analysis revealed negligible but significant effects of exercise on verbal fluency (ES: 0.17 [0.04, 0.30]). No significant benefit was found for additional executive measures, memory, or information processing. The authors critically appraised RCT methods and concluded on a moderate quality with the majority of trial samples being too small for sufficient power. There is clearly some effect for exercise at moderate to high levels of exertion. In addition, nowadays there is an emerging evidence for the effect of low levels of physical activity or “sedentary behaviors” like sitting. Physical inactivity may increase the risk of Alzheimer disease by 82 percent. Approximately 13 percent of Alzheimer disease cases worldwide may be attributable to sedentary behavior. A 25 percent reduction in sedentary behavior could potentially prevent more than 1 million Alzheimer disease cases globally. Sedentary behaviors may contribute to risk of Alzheimer disease and dementia by increased risk of cardiometabolic risk factors—diabetes, hypertension, obesity—diseases

associated with increased risk of dementia.

Table 57.2b–1.
Clinical Suggestions for Geriatric Populations

If 65 years or older, generally fit, and have no limiting health conditions

- ► At least 150 minutes of moderate-intensity aerobic activity (i.e., brisk walking) every week and muscle-strengthening activities on 2 or more days a week that work all major muscle groups (i.e., legs, hips, back, abdomen, chest, shoulders, and arms).

OR

- ► 75 minutes of vigorous-intensity aerobic activity (i.e., jogging or running) every week and muscle-strengthening activities on 2 or more days a week that work all major muscle groups.

OR

- ► An equivalent mix of moderate- and vigorous-intensity aerobic activity and muscle-strengthening activities on 2 or more days a week that work all major muscle groups.

Other comments

- ► If there is a limiting condition, it is recommended the health professional and patient consult a qualified exercise physiologist.
- ► 10 minutes at a time is fine for physical activity.
- ► The following activities count as muscle-strengthening activities: yoga, heavy gardening.

Adapted from the Centers for Disease Control and Prevention Report. Physical activity is essential to healthy aging; 2015.

Recommendations and Summary. Our clinical recommendation is physical activity should be encouraged for all members of society and particularly older adults. More work is required to understand if physical activity recommendations or dementia prevention are the same as generic "healthy aging" recommendations (as in Table 57.2b–1). While shorter-term interventions are useful in the research setting, care must be taken to ensure a sustained engagement in physical activity. Care must also be taken to create an age- and physical-function appropriate prescription given risks of adverse health effects (e.g., falls, cardiovascular events).

Mind–Body Practices

Mind–body practices are varied and can be best divided into exercise and nonexercise type practices. Mind–body physical activity practice includes yoga, tai chi, qigong. Mind–body nonphysical activity practice includes relaxation, mindfulness, meditation, and acceptance therapies. Mind–body physical activity practices are often defined as "physical activity executed with a profound inwardly directed contemplative focus." Mindful physical activity contains the following key elements: (1) a noncompetitive, nonjudgmental meditative component; (2) mental focus on muscular movement and proprioceptive awareness combined with a low to moderate level of muscular activity; (3) centered breathing; (4) a focus on anatomic alignment (i.e. spine, trunk, and pelvis) and proper physical form; and (5)

energy centric awareness of individual flow of intrinsic energy, vital life force, qi, etc. Mind–body nonphysical activity practice doesn't emphasize the physical activity component of the above. There is an emerging body of empirical research exploring the effects of mind–body therapies in geriatric psychiatry disorders.

Mind–Body Physical Exercise. A recent meta-analysis critically evaluated the effects of tai chi on individuals aged 60 and old with and without cognitive impairment. In this study, 20 eligible studies with a total of 2,553 participants were identified that met inclusion criteria for the systematic review; 11 of the 20 eligible studies were RCTs, 1 was a prospective nonrandomized controlled study, 4 were prospective noncontrolled observational studies, and 4 were cross-sectional studies. Overall quality of RCTs was modest, with 3 of 11 trials categorized as high risk of bias. Meta-analyses of outcomes related to executive function in RCTs of cognitively healthy adults indicated a large effect size when tai chi participants were compared with nonintervention controls ($g = 0.90$; $p = 0.04$) and a moderate effect size when compared with exercise controls ($g = 0.51$; $p = 0.003$). Meta-analyses of outcomes related to global cognitive function in RCTs of cognitively impaired adults, ranging from MCI to dementia, showed smaller but statistically significant effects when tai chi was compared with nonintervention controls ($g = 0.35$; $p = 0.004$) and other active interventions ($g = 0.30$; $p = 0.002$).

There is a relative dearth of studies examining yoga and its effects on memory enhancement. One main study of which the authors are aware explores the effects of yoga and exercise in 135 healthy men and women aged 65 to 85 years. In this study subjects were exposed to 6 months of Hatha yoga classes, walking classes or wait-list control; subjects were screened with a variety of mood and cognition tests (e.g., Stroop test). Interestingly, after the intervention there were no effects from either of the active interventions on any of the cognitive and alertness outcome measures. The yoga intervention produced improvements in physical measures (e.g., timed one-legged standing, forward flexibility) as well as a number of quality-of-life measures related to sense of well-being and energy and fatigue compared to controls. This type of study needs to be replicated in populations with MCI to better understand potential effects.

Mind–Body Practices and Stress-Reduction Techniques. Relaxation techniques, such as breathing exercises, guided imagery, and progressive muscle relaxation, are designed to produce the body's natural relaxation response. These relaxation therapies are shown to lower stress levels in elderly subjects, and so are recommended for clinical use.

A recent pilot randomized trial with mindfulness-based stress reduction (MBSR) aimed to test the safety and feasibility of MBSR in older adults

with MCI. The study found adults with MCI can safely participate and adhere to an MBSR program. The qualitative interviews revealed that most enjoyed the program and described improved mindfulness skills, well-being, interpersonal skills, acceptance/awareness of MCI, and decreased stress reactivity. Most data suggest trend toward improvement for measures of cognition and well-being, however, a more formal, large scale RCT is needed.

Recommendations and Summary. Firm clinical recommendations for these activities are not possible given the paucity of sufficient evidence at this stage. However, the authors recommend mind–body therapies as useful stand-alone or adjunctive therapies for the management of mild neuropsychiatric symptoms (e.g., cognitive impairment and psychological distress) and physical health (e.g., balance, strength, coordination). Care must be taken to create an age-, physical- and mental-function appropriate prescription. There recommendations are based from clinical experience, emerging evidence, and the safe nature of these activities.

Natural Products

Natural products may include herbs, minerals, natural supplements, vitamins, and probiotics. A range of these products have been explored with regard to the treatment and prevention of cognitive impairment. There are too many of these products to thoroughly review in this chapter, so only the main products, with a larger body of research and/or epidemiological usage, are reviewed here with further reading recommended.

Omega-3 Polyunsaturated Fatty Acid Supplements. One of the most widely utilized therapies is Ω3 polyunsaturated fatty acids (PUFAs). Over 37 percent of individuals who reported using a nonvitamin, nonmineral natural product in the 2007 NHIS reported using Ω3 PUFAs. Oily fish, such as salmon, mackerel, herring, and sardines are a rich source of Ω3 PUFA which are essential for brain development. It is thought that Ω3 PUFAs may benefit neurodegenerative disorders due to antioxidant and anti-inflammatory effects. Previous research from observational studies has suggested that increased consumption of fish oils rich in Ω3 PUFA may reduce the chance of developing dementia. A recent Cochrane Database systematic review explored the effects of Ω3 PUFAs on the prevention of dementia and cognitive decline in cognitively healthy older people. Information was available from 3 RCTs including 3,536 participants in total. This meta-analysis found no evidence to support a preventative effect following 24 or 40 months of intervention. RCTs suggest that selected patients with an MMSE score >27 were more likely to identify a positive effect of Ω3 PUFA supplementation. It appears that studies of longer

duration are required to further explore this area. The Cochrane review authors' clinical recommendation emphasizes clinician–patient discussions of evidence. Care must be taken with potential side effects of mild nausea, mild increased bleeding, and a fishy aftertaste.

Ginkgo Biloba Supplements . Ginkgo biloba leaf extract is among the most widely sold herbal dietary supplements in the United States. Its purported biological effects include scavenging free radicals, lowering oxidative stress, reducing neural damage, reducing platelet aggregation, and anti-inflammatory, antitumor, and antiaging activities. No standards or guidelines regulate the constituent components of Ginkgo biloba leaf extract, nor are exposure limits imposed. Safety evaluation of Ginkgo biloba leaf extract is being conducted by the U.S. National Toxicology Program. Many RCTs of *Ginkgo biloba* in patients with various types of dementia have yielded contradictory results. Of the studies that revealed cognitive improvement, the effect was minor and did not last more than 6 months. A recent RCT demonstrated the efficacy of once-daily dosed Gingko biloba (240-mg extract EGb 761) in outpatients aged 50 years or older with mild to moderate AD or vascular dementia (VaD) over 24 weeks. Both types of dementia showed improvement in cognition and neuropsychiatric symptoms compared to placebo. However, evidence does not suggest Gingko biloba to prevent dementia. In a trial of AD prevention in older adults, ages 72 to 96 years old, 120 mg of *Gingko biloba* twice a day was not effective in reducing the overall incidence of dementia in the elderly with normal cognition or MCI. Therefore, *Ginkgo biloba* will not prevent the development of AD but may have a minor impact on cognitive decline in those already afflicted with a dementia. However, it should be noted that the data supporting these positive findings come mostly from smaller studies which lack a placebo control group. As with other natural products, Ginkgo biloba comes with risks. It is recommended to use caution when using *Ginkgo biloba* in those who are at increased risk of hemorrhage as it may potentiate bleeding. Nevertheless, a short RCT of males using EGb 761 did not show any effect on coagulation or bleeding time. The short-term use of Ginkgo biloba is acceptable under some conditions, dependent on clinician–patient discussions, with the potential risks being considered.

B-Vitamin Supplements. B-vitamin supplementation is a potential therapy to prevent and treat age-related cognitive decline. It is considered standard procedure to check a vitamin B_{12} level when considering a diagnosis of dementia because it is seen as a reversible cause of cognitive decline. Moreover, a cross-sectional study showed various measures of vitamin B_{12}, including methylmalonate, homocysteine, and cystathionine to be associated with episodic memory deficits, slowed perceptual speed and

decreased total brain volume in a sample of adults aged 65 years and older. Interestingly, data have related elevated homocysteine levels to amyloid and glutamate toxicity and cognitive decline in animal studies. Additionally, vitamin B_3 (niacin) deficiency also referred to as pellagra leads to the classic triad of dementia, dermatitis, and diarrhea. Therefore, it is logical that B vitamin supplementation should benefit cognition. However, the evidence does not support a robust response in cognition with B vitamins. A multicenter RCT of adults over 50 years old with mild to moderate AD and normal vitamin B_{12} and B_9 levels showed no benefit in cognition with supplementation. Another RCT in China failed to show an improvement with supplementation in cognition with older participants, age 60 years or older, with either AD or VaD and normal vitamin levels. However, the evidence appears to be more promising in individuals without cognitive impairment. A recent RCT of older adults, ages 60 to 74 years old, without dementia and normal vitamin B_{12} and B_9 levels showed a benefit in immediate and delayed recall with supplementation. The evidence does not support vitamin B supplementation in treating dementia when levels are normal, however, some data support supplementation improving cognition prior to development of dementia. Additionally, such use of supplemental vitamins may come at a risk, the Iowa Women's Health Study demonstrated an increase in mortality with the use of various dietary supplements including vitamins B_6 and B_9. This study supports the fact that is important for patients and clinicians to acknowledge that being natural is not always safer than "conventional" medicine.

For further details on other utilized CAIM therapies for the treatment and prevention of age-related cognitive decline, refer to Table 57.2b–2.

Potential Side Effects and CAIM–Drug Interactions. As eluded to above, when encountering natural products in a clinical environment, it is important to consider potential side effects and CAIM–drug interactions. Table 57.2b–3 outlines an array of clinical findings and potential pitfalls of natural products. It is important to note that CAIM use may exacerbate polypharmacy in the elderly, which is a risk factor for drug interactions, medication errors, and hospitalization. A survey of 271 British seniors found a mean of 5.91 (range 4 to 7) herbal and nutritional supplements and 2.26 prescription drugs.

Table 57.2b–2.
Other Natural Products and Supplements Suggested for the Treatment of Cognitive Impairment

Herb/Supplement	Suggested Dose	Possible Mechanism of Action	Evidence Base	Side Effects

Bacopa monniera	300 mg/day	An Ayurvedic medicinal plant that acts as a free radical scavenger and in animal models reduces Aβ; may also boost cholinergic function	No convincing evidence: 9 RCTs; 518 subjects in meta-analysis. Limited evidence given poor study design.	GI upset
Curcumin (yellow curry)	2,000–8,000 mg/day	Antioxidant, anti-inflammatory, induces heat shock proteins; antiamyloid effects in animals	Mixed results; no convincing evidence: 2 studies find no effect; 1 study finds beneficial effect.	GI upset
Dehydroepiandro-sterone (DHEA)	100 mg/day	Adrenal steroid that declines with aging, is lower in Alzheimer disease patients, and is neurotrophic in animals	No convincing evidence of benefit. 3 RCTs	Acne, balding, insulin resistance, dyslipidemia, mood changes, hepatic dysfunction, possible effects on hormone-sensitive cancers
Hydergine	3 mg/day	Combination of 4 ergot derivatives with vasodilatory effects and possible effect on mono-amine and cholinergic transmission	2000 Cochrane meta-analysis of 19 trials found some evidence of modest efficacy but many trials conducted before standardized criteria for diagnosing dementia	GI upset, psychosis, flushing, blood pressure changes
Lecithin	3,600 mg/day	An acetylcholine precursor	Review of 11 randomized trials showed no consistent benefit	GI upset, rash, headache, dizziness
Resveratrol	Unknown (phase 1 trials 2,500–5,000 mg/day)	Polyphenol in the skin of red grapes and red wine with antioxidant and antiamyloid properties in animals	Clinical trial linking to increased cerebral blood flow. Needs further study on cognitive effects	Possible estrogen-like effects, as its chemical structure is similar to phytoestrogens
Vitamin E	800–2,000 IU/day	Free radical scavenging	No convincing evidence. Cochrane review—3 studies.	High dose increases all-cause mortality.

GI, gastrointestinal; RCT, randomized controlled trial.

Table 57.2b–3.
Potential CAIM–Drug Interactions between Natural Products and Prescription Medications

Herb/Supplement	Prescription Drug(s)	Possible Interaction
Chamomile	Sedatives (or alcohol)	Oversedation
Garlic extract	Protease inhibitors	Decreased efficacy against HIV
Ginkgo biloba	Anticoagulants	Excessive bleeding
	Omeprazole (CYP2C19 substrates)	Decreased drug levels (2C19 induction)
	Antipsychotic	Priapism
Ginseng	Anticoagulants	Excessive bleeding
	Monoamine Oxidase Inhibitors	Mania, Hypertension, Headache
	Sulfonylureas, Insulin	Hypoglycemia
	Stimulants	Insomnia, Anxiety, Mania
Kava	Sedatives (or alcohol)	Oversedation
	Levodopa	Increased Parkinsonism
Omega-3 fatty acids	Warfarin	Excessive bleeding
St. John wort	Protease inhibitors	Decreased efficacy against HIV
	Cyclosporine	Organ rejection
	Serotonergic antidepressants	Serotonin syndrome
	Oral contraceptives	Ineffective contraception
	Tetracycline	Photosensitivity
Valerian	Sedatives (or alcohol)	Oversedation
	CYP3A4 substrates	Increased drug levels (3A4 inhibition)
Vitamin E	Anticoagulants	Excessive bleeding

Body-Based Practices

Body-based practices may include massage therapy and spinal manipulation. The authors are unaware of any strong evidence supporting these therapies in the prevention and treatment of cognitive disorders, however, they may have merits in stress reduction, pain management, and treatment of anxiety disorders. Massage therapy includes many different techniques in which practitioners manually manipulate the soft tissues of the body. It has found some use in dementia care populations.

Aromatherapy. Aromatherapy, a therapy related to massage, has found some interesting results in small-scale studies. Aromatherapy is the use of essential oils extracted from plants for health purposes. Trials of aromatherapy have at times combined the application of the essential oils with massage while at other times the patients simply inhaled the aromas of the oils. One study reported the combination of aromatherapy (with lavender oil, commonly proposed to have sedating effects) and massage tended to yield less agitation than either condition alone. Another small study in China was positive for simple inhalation of lavender oil versus control oil in reducing agitated behaviors. The most methodologically

sound study of aromatherapy used lemon balm oil (*Melissa officinalis*) in a double-blind, placebo-controlled trial for managing agitation in severe dementia and found significant benefit for aromatherapy. Interestingly, the oil was applied to the skin (though not officially through "massage"), and this plant compound when administered orally. A case example of the potential benefits of CAIM therapies is below.

Mrs. M was an 80-year-old nursing home resident with advanced dementia. She was a "wanderer"; was frequently agitated and rarely verbal. She was a former dancer and responded to music of the 1940s and 1950s by calming down, dancing and singing along. She also responded to aromatherapy and brief massage using oils—especially lavender oil—by calming down and attending more to her grooming. She also liked participating in a group prayer and singing along. A rare smile would fill her face instead of fear, and she looked peaceful for a period of time. Overall, periods of better behavior and quality of life have improved for this patient.

Medical Systems

Chinese or Oriental Traditional Medicine. Acupuncture is a part of Oriental Traditional Medicine (OTM)—it is a technique in which practitioners stimulate specific points on the body—most often by inserting thin needles through the skin. Several reports of positive effects of acupuncture for cognitive and psychiatric symptoms of VaD have emerged from China, but a 2007 Cochrane review found the methodologies of most of these studies to be highly problematic, leaving no convincing evidence of positive benefit.

Ayurvedic Medicine. Ayurveda, which translates as "the science of life" and has been described as "knowledge of how to live," is a complete natural health care system that began in India more than 5,000 years ago and has been used within the context of mental health and aging for memory enhancement and anxiety reduction. Promoting immune system function through anti-inflammatory remedies, Ayurveda is widely used contemporarily in India as a system of primary health care. With the emphasis on developing an individualized regimen for patients that includes dietary, meditation, herbal, and other components, interest in Ayurveda is growing worldwide to treat a number of maladies including depression. Through yoga and/or transcendental meditation techniques, Ayurveda promotes lifestyle change and education about how to release stress and tension. Furthermore, Ayurveda is geared toward treating chronic disorders associated with the aging process, and promising preliminary results related to use of Ayurveda have been found in studies of depression, anxiety, sleep disorders, hypertension, diabetes mellitus, Parkinson disease, and Alzheimer disease. Of concern regarding the quality of certain Ayurvedic products, in 2008, an NCCAM-funded study examined the content of 193 Ayurvedic products purchased over the Internet and

manufactured in either the United States or India. The researchers found that 21 percent of the products contained levels of lead, mercury, and/or arsenic that exceeded the standards for acceptable daily intake.

CAIM MEDICINE IN LATE-LIFE MOOD DISORDERS

Late-life depression is a significant mental illness in older aged populations, with a rising burden expected this century. The most recent global data on late-life depression comes from a contemporary analysis of the World Health Organization's (WHO) Global Burden of Disease Study 2010. This study identified depressive disorders as a leading cause of burden internationally, and suggests major depressive disorder (MDD) was also a contributor of burden allocated to suicide and ischemic heart disease. Depressive disorders were the second leading cause of years lived with disability (YLD) in 2010. The burden of depressive disorders was highest in adults of working age (15 to 64 years) with 60.4 YLDs in 2010, versus adults ≥65 years with 6.1 million YLDs. In addition to the high and rising burden of late-life depression, conventional antidepressant therapies are limited by modest efficacy. In late-life depression, a recent meta-analysis of clinical trials suggests a response rate of 48 percent and a remission rate of 33.7 percent were found, both very similar to response and remission rates found in adult patients. Clearly alternative therapies need to be considered to enhance treatment efficacy.

Conventional Physical Activity

Conventional physical exercise has received considerable attention for its potential role in the treatment of depression, based initially on results of observational studies. A large (N = 6,828) cross-sectional study of men and women aged 20 to 88 years old found that aerobic fitness and physical activity levels as measured by a maximal treadmill test and self-report, respectively, significantly related to degree of depressive symptoms and emotional well-being. Interestingly, these relationships occurred in a graded, dose-dependent fashion, with progressively higher levels of fitness associated with increasingly lower levels of depressive symptoms and higher degrees of self-reported well-being. For physical activity levels, a similar dose–response relationship was observed; however, the greatest benefit was seen at between 11 and 19 miles per week of walking, jogging, or running.

The effects of aerobic exercise in depressed adults age 50 and older was examined and compared to either monotherapy with sertraline as a standard antidepressant or a combination of both approaches. The study found that after 16 weeks, all three groups demonstrated equivalent significant reductions in depressive symptoms as measured by scores on the *Hamilton Depression Scale* and the *Beck Depression Inventory*.

Furthermore, a 10-month follow-up study showed that participants in the exercise group who had improved to the point of remission during the initial study showed lower rates of relapse than those taking antidepressants, and independent continuation of exercising related to a reduced probability of being depressed at the end of the trial. Another study of exercise for older adults found that 10 weeks of weight-bearing exercises classes provided as an adjunctive treatment significantly lessened depressive symptoms in a group of older adults who had not responded to antidepressant therapy alone when compared to a health education course. A systematic review examining the role of physical exercise in the treatment of depression in older adults was conducted in 2009. This study noted 11 RCTs with a total of 641 participants aged >60 years. The study noted short-term (0 to 3 months) positive outcomes were found in nine studies, although there was variation in the type, intensity, and duration of exercise. The efficacy of exercise in the medium- and long-term (3 to 12 months and >12 months, respectively) was less clear. Issues were noted, with only five studies showing appropriate allocation of concealment, five studies showing intention to treat analysis and blinding occurred in seven studies. The effectiveness of a number of lifestyle interventions promoting increases in conventional physical exercise has been recently reviewed. Overall, interventions were relatively successful in initiating increased levels of physical activity in the short run; however, maintenance of these gains were diminished over time, and there was no convincing evidence that behavioral reinforcement strategies were beneficial. Another review of exercise interventions found that while many older adults increase physical activity levels in response to a wide variety of interventions, the amount of increase is seldom enough to convincingly impact health outcomes in a positive manner.

Recommendations and Summary. In summary, preliminary evidence points to exercise as having a significant potential to lessen symptoms in depressed older adults, and furthermore, to prevent relapse of depression and/or development of depressive symptoms in older adults experiencing acute stressors associated with aging. Therefore, the authors recommend physical exercise as a stand-alone therapy in the treatment of mild depression, and an adjunct in moderate depression. Recommendations for engagement are given in Table 57.2b–1.

Mind–Body Practices

Mindful exercise interventions have shown promise in addressing depressive symptoms in older adults. For example, a study of 82 older adult participants with depression randomized to either 16 weeks of qigong practice or newspaper reading groups found that qigong participants showed significantly greater improvements in mood, self-efficacy, and

personal well-being. Practice of yoga typically benefits from instruction by expert instructors and requires the dedication by participants to multiple weekly sessions and continual use for maximal benefit. Prior review of published RCTs of yoga for depression in adults revealed that while all trials found benefit trial methodologies have generally been weak with lack of blinding, short duration of the intervention, variable outcome measures, limited information about subjects, randomization procedures, compliance, and dropout rates. Comparative studies of yoga have likewise been limited, with one trial demonstrating that yoga was as effective as TCAs and another trial showing that yoga may provide benefit as an augmentation strategy for antidepressant treatment. Yoga is commonly used in combination with other treatments for depression, anxiety, and stress-related disorders. Data on use of yoga for anxiety and depression in older adults are more limited; however, 1 significant study of 69 older adults in India did compare the impact of yoga to Ayurveda or a wait-list control condition on sleep and depressive symptoms. Participants in the yoga group practiced physical postures, relaxation techniques, regulated breathing, devotional songs, and attended lectures for more than 7 hours a week during the course of the 6-month trial. Practice of yoga significantly impacted quality of sleep and level of depressive symptoms when compared to the two control conditions, neither of which demonstrated significant effects. In particular, depressive symptoms, as measured by the short form of the *Geriatric Depression Scale*, decreased in the yoga group from a baseline average of 10.6 to 8.1 by 3 months and 6.7 by 6 months. The average time to fall asleep decreased in the yoga group by 10 minutes, while total number of hours slept increased by 60 minutes, and resulted in a greater feeling of being rested after 6 months. A recent study compared the effectiveness of laughter yoga, group exercise therapy, and control in decreasing depression in older adult women (60 to 80 years). In this study, 70 depressed women were chosen if their Geriatric Depression Score was >10. This study went for 10 sessions and found a significant improvement in depression score with both yoga and group exercise therapy as compared with control. While the evidence supporting the use of these therapies in the treatment of late-life mental illnesses is not strong, there are some interesting early results which require further high-quality research.

Recommendation and Summary. We recommend the use of mind–body therapies as a stand-alone or adjunctive antidepressant treatment in the management of geriatric depression, based on individual preference and the severity of depression. Further high-quality data is needed to further inform these recommendations.

Natural Products

Ω3 PUFA Supplementation. Researchers support an antidepressant

effect of Ω3 PUFAs. Low levels of Ω3 PUFAs are seen in individuals suffering from depression as well as in suicidal patients. In general adult populations, a comprehensive review found that actual diagnosis of depression predicted a response to Ω3 PUFA supplementation compared to simply having depressive symptoms, supporting Ω3 PUFA as a beneficial treatment for individuals suffering from a clinical depression as compared to minor depression. Another meta-analysis showed that Ω3 PUFA supplementations with ≥60 percent of EPA were beneficial to depression. Therefore for the general adult population, evidence supports a role for Ω3 PUFA supplementation in treating depression. The data supporting Ω3 PUFA supplementation in the geriatric depression are less extensive. There are no meta-analyses available for the geriatric population, however, some RCTs looking at the antidepressant effect of Ω3 PUFA exist. A small RCT of clinically depressed females, age 66 to 95 years old, who were living in a nursing home that found Ω3 PUFA supplementation improved depression and quality of life. However, another RCT of older adults, age 60 to 80 years old who had recently suffered a myocardial infarction failed to show an antidepressant effect of Ω3 PUFA supplementation. However, when a subsample of individuals from this study who had been diagnosed with clinical depression were evaluated, Ω3 PUFA supplementation did provide an antidepressant effect. These findings are consistent with previous studies in general adult populations in which Ω3 PUFA provide benefit for clinical depression.

St. John Wort. St. John wort (SJW) (Hypericum perforatum) is a wildflower that has been used for medicinal purposes for thousands of years. Various standardized extracts of this natural dietary supplement have been studied. In older adults, the available evidence base for use of SJW as an antidepressant is limited. A small RCT comparing SJW (800 mg extract LoHyp-57) and fluoxetine (20 mg) in moderately depressed older adults (60 to 80 years old) found equivalent efficacy of these agents in reducing depression over 6 weeks. Additionally, a more recent study showed improvement in depressive symptoms with SJW (600 mg extract LI160) in a sample of mixed-aged adults (18 to 70 years) with atypical depression. Therefore, the few studies available which include older adults support the use of SJW for depression. A recent meta-analysis of 23 randomized trials, including outpatients with mild-to-moderate depression found improvement in depressive symptoms compared to placebo. Most studies included in the meta-analysis used dosages of SJW ranging from 300 to 1,000 mg daily. Another meta-analysis showed similar efficacy and better tolerability of SJW when compared to standard antidepressants such as tricyclics and Selective Serotonin Reuptake Inhibitors (SSRIs). Clinical trials have used various divided doses of SJW (600 to 1,800 mg daily), with the most common dose being 300 mg three times daily. Therefore, there is

a strong evidence basis for SJW in treating the general adult population suffering from depression. SJW is relatively well tolerated and safe when used alone, however, when taken in combination with other medications, there can be numerous potentially dangerous drug–drug interactions. SJW acts as a monoamine oxidase inhibitor (MAOI) making its combination with standard antidepressants concerning for precipitating a serotonin syndrome. SJW is also an inducer of cytochrome P-450 3A4 liver enzyme system, which leads to reductions in multiple medications such as warfarin, methadone, alprazolam, clopidogrel, cyclosporine, and verapamil. SJW has also been seen to precipitate a manic episode. For an older adult who would rather not take conventional western medications, SJW can be considered as an alternative natural remedy for mild to moderate depression. However, the ideal patient is free of comorbid medical conditions and all patients using SJW should be informed about potential side effects and drug–drug interactions.

S-Adenosyl-L-Methionine. S-adenosyl-L-methionine (SAMe) is a naturally occurring compound normally synthesized in the body formed from the amino acid L-methionine as part of a multi-step metabolic pathway involving folic acid and vitamin B_{12}. Initial oral formulations of SAMe were relatively unstable, necessitating parenteral administrations. However, the more recent development of stable oral forms of SAMe has allowed for more widespread testing and use of this compound. SAMe is thought to play a role in depression since it is a necessary cofactor for the synthesis of serotonin, norepinephrine, and dopamine. Additionally, cerebrospinal fluid levels of SAMe have been shown to be decreased in depressed patients and levels rise in response to treatment. Clinical studies in the general adult population support an antidepressant effect of SAMe. Meta-analyses found that SAMe is superior to placebo and equivalent to TCAs. This analysis included mostly parenteral formulations and when they were compared to oral formulations, parenteral SAMe showed a slight superiority over oral formulations. Therefore there is strong evidence to support the antidepressant effect of parenteral SAMe in the general adult population for an antidepressant effect. The data supporting oral formulations of SAMe are growing. A more recent open trial looking exclusively at oral SAMe in a mixed age population (18 to 80 years old) supports its clinical benefit as an augmentation strategy in patients showing an incomplete response to conventional antidepressants. Additionally, a small, 6-week open trial of oral SAMe in the general adult population found that response and remission rates were significantly higher than placebo. Among the trials, the oral doses of SAMe ranged from 400 to 1,600 mg/day, with most trials showing benefit with dosing between 800 and 1,600 mg/day. However, dosages up to 3,600 mg/day have been used with success and safety in older populations, supporting higher

dosages may be necessary for clinical benefit. Therefore, there is strong evidence to support the use of SAMe as an antidepressant with parenteral formulation and the support is growing for oral administration. SAMe is generally well tolerated, however, important safety concerns such as increased anxiety, mania, and hypomania have been reported in patients with bipolar disorder. There was one case study of a 61-year-old female suffering from depression, with a family history of bipolar disorder, who took SAMe and attempted suicide by self-burning. Patients with bipolar disorder should not take SAMe without concomitant administration of mood stabilizers coupled with clinical monitoring. Therefore, SAMe can be beneficial for depression, however, the data are lacking for the geriatric population. An ideal candidate for a trial of SAMe could include a mildly symptomatic patient who would not have serious effects if there was a delay in treatment or a person who has failed multiple trials of conventional antidepressants and/or is unable to tolerate side effects.

For other natural products and supplements which may be used for the treatment of late-life depression, see Table 57.2b–4.

Other Therapies

Acupuncture. Prior reviews of RCTs using acupuncture to treat depression have found some evidence for benefit; however, the majority of included studies have been of relatively low quality with placebo acupuncture treatment at times yielding similar results to treatment acupuncture. The Cochrane database review of studies examining acupuncture for depression found insufficient evidence to determine the efficacy of acupuncture compared to standard antidepressants, sham acupuncture, or wait-list controls. Most recently, an RCT comparing electro-acupuncture to control acupuncture in 53 subjects diagnosed with mild to moderate major depression found significant but similar improvements in depressive symptoms for both techniques. For depression in older adults, the literature is limited, with a pilot study of acupuncture for depression in older adults finding positive effects on measures of mood and well-being.

Table 57.2b–4.
Other Natural Products and Supplements Suggested for the Treatment of Late-life Depression

Herb/Supplement	Suggested Dose	Possible Mechanism of Action	Evidence Base	Side Effects
Inositol	6–19 mg/day	Altering serotonin, dopamine, glutamate levels.	Mixed findings: 7 RCTs in depression explored (MDD, PMDD). One positive, some negative.	GI upset

Saffron	30 mg/day	Serotonergic effects, antioxidant, anti-inflammatory, neuroendocrine effects, neuroprotective effects.	5 trials found. Promising findings which need further, high-quality assessment.	Headache, nausea, anxiety, decreased appetite.
Vitamin D	600–300k IU/daily or weekly	Neuroprotection.	Promising findings requiring further research. Meta-analysis of 7 RCTs.	Fatigue, headache.

MDD, major depressive disorder; PMDD, premenstrual depressive disorder; GI, gastrointestinal; RCT, randomized controlled trial.

FUTURE DIRECTIONS

Gaining a better understanding of the clinical effect of CAIMs, as well as their safety is clearly critical for the management of geriatric psychiatry conditions. At present, there is insufficient data on the prevalence, effectiveness, efficacy, safety, and health economic benefits of most CAIM treatments. This means patients, health providers, and other stakeholders are unable to access reliable and rigorous evidence on CAIMs. When considering future directions for this field, it is important to consider the most recent NCCIH strategic plan "Exploring the Science of CAM: Third Plan 2011–2015." This plan has five main objectives which are (1) to advance research on mind and body interventions, practices, and disciplines, (2) advance research on CAIM natural products, (3) increase understanding of real-world patterns and outcomes of CAIM use and its integration into health care and health promotion, (4) to improve the capacity of the field to carry out rigorous research, and (5) to develop and disseminate objective, evidence-based information on CAIM interventions.

Specific to psychiatric clinical practice, the NCCIH authors' have further recommendations for future CAIM trials and clinical utilization. *Researchers must disclose conflicts of interest* matching current trends in pharmaceutical trials. *CAIM trials should aim to replicate the highest trial standards*. This means the adequate blinding, allocation concealment, randomization, complete outcome data presentation, and thorough reporting. *CAIM trials should utilize adequate control groups*. Control groups for holistic treatments should have similar elements to active groups; control groups for biologically based treatments can use placebo. *Randomized clinical trials must include treatment and prevention of all mental disorders in older adults* in order to build a comprehensive evidence for such practices. *Care must be taken to ensure the reliability, safety, and consistency of CAIM supplements*. It is important that doctors and patients are utilizing the same preparations which are free from toxins or other potentially harmful ingredients. *Research must be conducted in other geriatric psychiatry disorders,*for example, anxiety disorders and oncology-related psychological issues. Clinical trial research must occur

more intensively in nutritional psychiatry. For example, Prevención con Dieta Mediterránea (PREDIMED) and the Supporting the Modification of lifestyle In Lowered Emotional States (SMILES) trial.

CONCLUSION

Neuropsychiatric disorders in older adults such as mood, anxiety, sleep, and cognitive disorders are among the most common reasons for older adults to use CAIM therapies. The use of CAIM therapies in geriatric psychiatry is promising with a growing body of empirical research. There is a wide range of research quality within this field, making firm conclusions difficult. Therefore, the authors recommend high-quality research continue into this area to establish efficacy and safety. When promoting safety of various natural products, clinicians can offer important information for those patients who are less comfortable with "conventional" medicine and prefer alternative "natural" therapies, such as advising of drug–CAIM interactions and directing them toward the evidence-based indications. It is clear that CAIM use is significant in older adult populations, will likely increase, so empirical research is essential.

REFERENCES

Abbott R, Lavretsky H. Tai Chi and Qigong for the treatment and prevention of mental disorders. *Psychiatr Clin North Am*. 2013;36(1):109–119.

Barnes DE, Yaffe K. The projected effect of risk factor reduction on Alzheimer's disease prevalence. *Lancet Neurol*. 2011;10(9):819–828.

Burns A, Perry E, Holmes C, et al. A double-blind placebo-controlled randomized trial of Melissa officinalis oil and donepezil for the treatment of agitation in Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2011;31(2):158–164.

Chodzko-Zajko WJ, Proctor DN, Fiatarone Singh MA, et al; American College of Sports Medicine. American College of Sports Medicine position stand. Exercise and physical activity for older adults. *Med Sci Sports Exerc*. 2009;41(7):1510–1530.

de Rezende LF, Rey-Lopez JP, Matsudo VK, do Carmo Luiz O. Sedentary behavior and health outcomes among older adults: a systematic review. *BMC Public Health*. 2014;14:333.

Farina N, Isaac MG, Clark AR, Rusted J, Tabet N. Vitamin E for Alzheimer's dementia and mild cognitive impairment. *Cochrane Database Syst Rev*. 2012;11:CD002854.

Ferrari AJ, Charlson FJ, Norman RE, et al., Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med*. 2013;10(11):e1001547.

Gates N, Fiatarone Singh MA, Sachdev PS, Valenzuela M. The effect of exercise training on cognitive function in older adults with mild cognitive impairment: a meta-analysis of randomized controlled trials. *Am J Geriatr Psychiatry*. 2013;21(11):1086–1097.

Ho TF, Rowland-Seymour A, Frankel ES, Li SQ, Mao JJ. Generational differences in complementary and alternative medicine (CAM) use in the context of chronic diseases and pain: baby boomers versus the silent generation. *J Am Board Fam Med*. 2014;27(4):465–473.

Kennedy DO, Wightman EL, Reay JL, et al. Effects of resveratrol on cerebral blood flow variables and cognitive performance in humans: a double-blind, placebo-controlled, crossover investigation. *Am J Clin Nutr*. 2010;91(6):1590–1597.

Lavretsky H. Complementary and alternative medicine use for treatment and prevention of late-life mood and cognitive disorders. *Aging Health*. 2009;5(1):61–78.

Linde K, Berner MM, Kriston L. St John's wort for major depression. *Cochrane Database Syst Rev*. 2008;(4):CD000448.

Martinez-Lapiscina EH, Clavero P, Toledo E, et al. Mediterranean diet improves cognition: the PREDIMED-NAVARRA randomised trial. *J Neurol Neurosurg Psychiatry*. 2013;84(12):1318–1325.

Mather AS, Rodriguez C, Guthrie MF, McHarg AM, Reid IC, McMurdo ME. Effects of exercise on depressive symptoms in older adults with poorly responsive depressive disorder: randomised controlled trial. *Br J Psychiatry*. 2002;180:411–415.

NCCAM. National Center for Complementary and Alternative Medicine. 2014. http://nccam.nih.gov/.

Papakostas GI. Evidence for S-adenosyl-L-methionine (SAM-e) for the treatment of major depressive disorder. *J Clin Psychiatry*. 2009;70(5):18–22.

Paul RT, McDonnell AP, Kelly CB. Folic acid: neurochemistry, metabolism and relationship to depression. *Hum Psychopharmacol*. 2004;19(7):477–488.

Pilkington K, Kirkwood G, Rampes H, Richardson J. Yoga for depression: the research evidence. *J Affect Disord*. 2005;89(1-3):13–24.

Saper RB, Phillips RS, Sehgal A, et al. Lead, mercury, and arsenic in US- and Indian-manufactured Ayurvedic medicines sold via the Internet. *JAMA*. 2008;300(8):915–923.

Selkoe DJ. Preventing Alzheimer's disease. *Science*. 2012;337(6101):1488–1492.

Sublette ME, Ellis SP, Geant AL, Mann JJ. Meta-analysis of the effects of eicosapentaenoic acid (EPA) in clinical trials in depression. *J Clin Psychiatry*. 2011;72(12):1577–1584.

Sydenham E, Dangour AD, Lim WS. Omega 3 fatty acid for the prevention of cognitive decline and dementia. *Cochrane Database Syst Rev*. 2012;6:CD005379.

Varteresian T, Lavretsky H. Natural products and supplements for geriatric depression and cognitive disorders: an evaluation of the research. *Curr Psychiatry Rep*. 2014;16(8):456.

Walker JG, Batterham PJ, Mackinnon AJ, et al. Oral folic acid and vitamin B-12 supplementation to prevent cognitive decline in community-dwelling older adults with depressive symptoms—the Beyond Ageing Project: a randomized controlled trial. *Am J Clin Nutr*. 2012;95(1):194–203.

Wayne PM, Walsh JN, Taylor-Piliae RE, et al. Effect of tai chi on cognitive performance in older adults: systematic review and meta-analysis. *J Am Geriatr Soc*. 2014;62(1):25–39.

57.2c The Aging Brain

DOUGLAS R. GALASKO, M.D.

INTRODUCTION

Aging is an important factor in health and disease. Over the adult life span,

the body undergoes changes that generally result in loss of functional capacity. These may be driven by fundamental processes that underlie aging, and can be exacerbated by the presence of disease and influenced by lifestyle factors. The brain is not unique in experiencing changes in association with aging that also tend to decrease its capabilities. These changes can be described at different levels, including clinical evaluation, risk and protective factors, biomarkers such as brain imaging, and underlying mechanisms and pathophysiology. Although the current understanding of the mechanisms that underlie brain aging is incomplete, there has been much progress in identifying the contributions of aging and disease processes to changes in cognition, mood, and behavior in the elderly. Many myths about aging have arisen, and old age has often been viewed as a time of physical and mental decline and infirmity. This is challenged by recent research in successful aging—much progress has been made in defining the cognitive strengths and abilities that are preserved in healthy brain aging and in teasing out contributions of diseases to loss of ability. By delineating the clinical features, biomarkers and pathophysiology of conditions that are not part of healthy brain aging and result in impairment of brain function, such as Alzheimer disease (AD) and vascular pathology, specific treatments may be developed to target these age-associated disease processes. Factors that may limit age-related impairment in cognition and other aspects of brain function and result in greater resilience in the face of aging are also receiving attention; these include cognitive stimulation, social and physical activity, and nutrition. This chapter provides an overview of concepts and challenges in studying and measuring brain aging and the effects of aging on brain anatomy. Tools and approaches to measure cognitive aging and neuroimaging studies that have revolutionized how brain aging is studied will be reviewed. The nature and impact of common age-associated brain pathologies and diseases is described. Next, the burgeoning area of identifying factors that may modulate the effects of brain aging and understanding how their effects may be mediated will be reviewed. Finally, approaches that may help to promote healthier brain function during aging are summarized.

CONCEPTS AND CHALLENGES IN STUDYING BRAIN AGING

How are metrics for brain aging defined and established? Some important questions and issues that are critical to this measurement question are listed in Table 57.2c–1. First, is it useful to try to define approximate stages in brain aging based on transitions at specific ages, even though the underlying processes are likely to be continuous and cumulative? Although it is easy to identify age by the passage of time, biological markers of aging processes have not been developed. By default, “milestones” used to define aging are thus measured by chronological age, and may impose artificial divisions and definitions. For example, many aging studies set a cutoff of

65 or older as elderly. Even disease definitions, such as dementia, have considered onset before the age of 65 as presenile and onset after 65 as senile. This arbitrary cutoff at 65 was selected based on an age for retirement benefits that was established in 1889. In developed countries, average life expectancy now exceeds 75, and many people live into their ninth and tenth decades. The number of centenarians in the United States and Europe has steadily increased in the recent past and will continue to do so. So, the concepts of young-old, aged 65 to 75, and oldest-old, aged 85 and older, have arisen. However, there are no agreed-upon measures or markers other than the passage of time that distinguish groups of people in their eighth, ninth, and tenth decades. Studies have not yet shown clearly whether age-associated decline of brain abilities accelerates at particular ages. Indirect support for acceleration comes from epidemiological studies of age-associated syndromes such as dementia, which show an exponential increase in incidence after the age of 65. Therefore, the inadvertent inclusion of some subjects with preclinical dementia influences many studies of "normal aging" with a likely outcome that cognitive decline will appear to accelerate in very old age.

Table 57.2c–1.
Important Questions Regarding Brain Aging

1. Should benchmarks or reference ages be used to calibrate the extent of age-related changes? Are there biomarkers of chronological age?
2. Is brain aging a steadily progressive process or does it accelerate?
3. How can "normal" aging be distinguished from disease-related changes?
4. Are some brain pathological changes inevitable?
5. What factors are associated with more or less severe age-associated brain changes?
6. Are some brain regions more susceptible to age-related changes than others?
7. Do plasticity and cognitive reserve attenuate the loss of function associated with brain aging?

Another question about staging is how to develop benchmarks for "normal" brain structure and performance against which to measure effects related to aging and disease. One approach is to collect normative data for groups of apparently healthy individuals of varying age, that is, to age adjust. This may have clinical utility, for example, to define laboratory test cutoffs or identify suboptimal performance as a likely indicator of disease. Clinical, pathological, and imaging research studies across the life span suggest that the human brain matures, establishes connectivity, and achieves optimal function by the time people reach their mid-to-late 20s. Some decline can be measured beginning as early as the 30s. Should brain structure and performance at age 30 be used as a benchmark to calibrate the magnitude of changes associated with many years of aging? This may be helpful for estimating relative changes in the number of neurons in particular brain regions or changes in connections between neurons, and

may also shed light on how human cognitive performance changes. Many studies have followed people from age 60 or older to derive data on brain structure or performance and compare subjects in later decades to people in their 60s. However, this may underestimate the magnitude of changes across the adult age span. This may particularly affect pathology studies, where studies across the age span have indicated that a disorder such as AD expresses itself clinically only after decades of slowly emerging pathology.

Cross-sectional comparisons across the age span may be inflated by secular changes that may exaggerate differences between younger and older individuals. For example, levels of education, nutritional habits, and the widespread use of medications such as statins and antihypertensives have changed markedly over the past 50 years. These changes may result in differences between cohorts born decades apart, when assessed through cognitive testing during life and even in studies of postmortem brain tissue. Longitudinal data relating measures of clinical performance of the brain to underlying structure are becoming available from imaging studies, but the period covered by serial imaging is usually measured during the framework of a few years rather than decades. Secular changes do not inevitably result in better brain health; inactivity, obesity and increasing rates of type 2 diabetes among adults in Westernized societies may counteract improvements in education and cardiovascular management.

Almost any physiological index shows an increase in its variability in association with aging. This holds true for many aspects of cognitive and motor performance in the elderly. The reasons for increased variability are not well understood, but both genetic and environmental factors may contribute. Identifying factors that contribute to decline in brain performance associated with aging, particularly those that may be modifiable or preventable, may lead to interventions aimed at forestalling or attenuating the degree of age-associated decline that is considered to fall within the spectrum of normal.

Are there boundaries between brain aging and disease? Can disorders like AD be clearly demarcated from normal or optimal brain aging? Aging is associated with an increased prevalence of lesions such as senile plaques (SP) and neurofibrillary tangles, as well as vascular brain changes. Even though the clinical syndrome of progressive dementia associated with AD is different from the lesser degrees of cognitive change associated with aging, the key pathological features of AD are present to some extent in most individuals aged 90 or older. If Alzheimer pathology is an inevitable feature of aging, it is possible that extension of life span will lead to the inevitable development of clinical AD. Even if AD pathology is inextricably linked to the aging process, the pathology builds up slowly, over a decade or longer, and the task of understanding factors that may lead to earlier or later age of onset of clinical dementia and the quest to delay the expression of clinical dementia for as long as possible remains extremely important. Because of

the high frequency of AD lesions in the elderly, it is difficult to carry out studies of brain aging in which the definition of normal requires the exclusion of any degree of AD pathology. It may be possible to dissect the impact of aging from that of AD. In this quest, biomarkers for AD pathology have been used in recent studies to illustrate that there are differences in longitudinal cognitive performance and in magnetic resonance imaging (MRI) indices of brain functional connectivity in people who test positive for amyloid biomarkers compared to those who test negative.

Studying Brain Tissue: Neuroanatomy, Genomics, and Imaging

Neurons vary greatly in structure and connectivity throughout the brain. An important question is whether aging selectively affects certain types of neurons or specific brain regions, or whether changes are relatively uniform and widespread. To provide a systematic answer to this question would require an enormous amount of work obtaining quantitative measures or neuronal numbers and axonal and dendritic structure throughout the brain from a series of tissue samples across a wide age span. Studies of neurons during aging have typically targeted brain regions that whose connections or neurotransmitter circuitry relate to specific brain processes, for example, hippocampal neurons in relation to memory.

An issue that troubled the field for years is the extent of neuron loss with aging. Interest was sparked by an initial series of studies by Harold Brody, supported by other investigators, which suggested that up to 50 percent of neurons were lost between the ages of 18 and 95 in the superior-frontal and superior-temporal cortices and 20 to 30 percent losses occurred in the precentral gyrus and visual cortex. It was later shown that processing of brain tissue caused shrinkage, leading to overestimation of the density of neurons in brains from younger individuals. After appropriate adjustment and when automated counting methods were used, markedly lower estimates of neuron loss were obtained. But the number of neurons is only one of many conceivable measures of neuronal well-being. Recent studies have evaluated the dendritic arbor or synaptic density as indices of neuronal connectivity, and have assessed the strength of neurotransmitter circuits by measuring levels of neurotransmitters or their receptors in the brain. These approaches, assessing structure, connectivity, and biochemistry, provide a more nuanced picture of age-associated changes.

Advances in technology are enabling comprehensive studies of protein and gene expression in discrete brain regions and even in single neurons. To interrogate mechanisms of brain aging, large-scale omic analysis is now feasible. In particular, transcriptional analysis using arrays to measure relative mRNA expression, followed by bioinformatic approaches, allows signaling pathways to be identified and analyzed. For example, the BRAINEAC project (www.braineac.org) analyzed mRNA expression in 10

brain regions, in brains from over 100 people covering the adult age span.

In recent years, neuroimaging studies have brought another dimension to the study of brain aging. It is possible to obtain morphological, functional, and metabolic information about the brain in living people with excellent precision and reliability. The resolution of MRI is sufficient to obtain overall and regional quantitative measures related to the brain. Multimodal brain imaging studies, combining measures of brain structure, connectivity, blood flow, and chemistry in the same individual, are being brought into play to help to understand structure–function relationships.

Thresholds of Pathology, Clinical Disorders, and Cognitive Reserve

To clarify the relationship between the pathology and the clinical expression of brain disorders, many attempts have been made to define thresholds, for example, a burden of pathology or of cell loss that is typically present by the time that a clinical phenotype manifests. For example, studies of postmortem brain tissue of patients with Parkinson disease have suggested that motor symptoms do not appear until 40 percent or more of dopaminergic neurons in the substantia nigra are lost. This exceeds the extent of neuron loss found in this brain region in studies of aging. Impaired reserve and plasticity may play a role in whether age-associated changes in the brain may manifest as disease or not. Subthreshold changes related to aging may increase vulnerability to pathology. For example, an individual with mild Alzheimer pathology may maintain intact cognitive function with aging due to sufficient brain reserve. However, the addition of one or more small strokes may be enough to impair cognition beyond a tipping point, resulting in clinically recognizable dementia. Similarly, age-related decreases in the strength of norepinephrine or dopamine circuits may increase the vulnerability for the development of depression, which may be precipitated by a life event or stressor that would not have done so at a younger age. Finally, brain aging may increase someone's vulnerability to develop adverse cognitive side effects as a result of medications, and also to develop delirium in the setting of systemic illness.

Cognitive reserve may be thought of as resilience that allows performance to be preserved in the face of damage, and it may protect against the onset of a clinical disease phenotype. Two types of reserve have been proposed by Dr. Yaakov Stern: brain reserve refers to differences in brain structure that may result in tolerance of pathology, and cognitive reserve describes differences between individuals in how tasks are performed. Brain reserve may depend on the intrinsic richness of connections between brain structures or possibly on the total number of mature neurons that remain at the end of brain development. Factors such as education, carrying out cognitively demanding activities or physical exercise may assist in the maintenance of these connections during aging.

Neuroplasticity is an important component of reserve. This classically refers to the brain's ability to adapt and strengthen connections or potentially to form new connections between neurons and circuits. Studies of recovery after stroke or ablative brain surgery support this possibility. However, it is generally noted that aging leads to less effective recovery after injury, and the brain is no exception. Neurons have been classically considered to be fully differentiated postmitotic cells, which makes them particularly vulnerable to cumulative damage sustained over the years. Although neurons do not divide, they have the capacity to form new synapses and carry out remodeling. This limited view of the regenerative potential of neurons has changed; exciting research findings over the last two decades have challenged the doctrine that neurons are not actively replaced in mammals. Neural progenitor cells had been identified in birds and rodents and shown to play roles in brain plasticity. A prominent example is neurogenesis that accompanies the acquisition of new songs by birds. In the 1990s, evidence emerged that similar progenitor cells exist in the brain in primates, including humans. In brain regions such as the hippocampus, new neurons can be generated throughout adult life as progenitor cells in the subventricular and granular zones differentiate. Although the roles of these progenitor cells in health and disease are still being defined, their potential to proliferate, differentiate, and migrate expands the repertoire of compensatory changes within the brain. In rodent and primate models, brain injury, for example, due to stroke, leads to the stimulation of neurogenesis. Recent studies have suggested that depression may be associated with deficient neurogenesis.

MECHANISMS IMPLICATED IN BRAIN AGING

The mechanisms and processes that lead to age-related changes throughout the body and ultimately limit life span are not well understood. Aging has been characterized by one of its leading scientists, Dr. Leonard Hayflick, as the effect of an increase in entropy on molecules such as proteins, lipids, and deoxyribonucleic acid (DNA), resulting in the accumulation of structural changes that eventually lead to impaired physiological function at the level of molecules, cells, and ultimately the organism. Many mechanisms may contribute to age-related damage that may impair brain function or reduce brain reserve, as summarized in Table 57.2c–2.

Oxidative stress is particularly relevant to the brain, whose long-lived neurons use oxygen more extensively than any other organ. Although there are many antioxidant defenses and repair mechanisms, over time oxidative damage affects the structure of DNA, ribonucleic acid (RNA), lipid membranes, and proteins. These changes may become irreversible, resulting in impaired function of the affected molecules and of organelles. Mitochondrial damage associated with aging may lead to impairment of energy metabolism; this is well documented for the substantia nigra.

Proteins that undergo oxidative changes such as nitrosylation or carbonylation may be prone to cross-linking and to aggregation. In addition, they impose a load on the degradative machinery of neurons, which may be unable to remove these abnormal proteins effectively or break them down completely. Diabetes, which is increasingly prevalent with aging, may promote oxidative stress through mechanisms such as glycation of proteins and generation of advanced glycation end-products. Caloric restriction has been widely studied in animals as a method that does appear to extend life span, and decreasing oxidative damage and improving insulin signaling are two potential mechanisms that may explain these effects. Whether this is effective for life span extension in humans is unknown, and caloric restriction has not yet been proven to benefit brain health.

Misfolding of proteins, with the adoption of stable insoluble amyloid configurations, is an important theme in many age-associated diseases. For example, aggregates of amyloid-β protein (Aβ) and of the microtubule-associated protein tau form SP and neurofibrillary tangles, the hallmark lesions of AD. In families with early-onset AD, mutations in genes related to Aβ (amyloid protein precursor [APP] and presenilin 1 and 2) are associated with a relative or absolute increase in the production of longer forms of Aβ or produce mutant forms of Aβ that are more prone to aggregate. Evidence to date does not support the overproduction of Aβ in sporadic AD or with aging. It is possible that only a small change in the expression levels of certain proteins is needed to initiate pathology, or aging may predispose to amyloid deposition in the brain through other mechanisms. Impaired degradation or clearance of Aβ is implicated in age-associated plaque formation. Factors that modulate the misfolding, aggregation, and clearance of proteins (proteostasis) during aging continue to receive much research attention.

Table 57.2c–2.
Mechanisms Implicated in Brain Aging

1. Oxidative stress—damage to DNA, proteins, lipids
2. Mitochondrial damage
3. Protein misfolding and aggregation
4. Decreased endocrine support (e.g., estrogen and testosterone)
5. Decreased neurotrophic factor support
6. Impaired axonal transport
7. Impaired neurogenesis

Another age-related factor may be lack of trophic support of certain types of cells or their connections. Estrogen deficiency after menopause or androgen decline with aging may impair the function of neurons that respond to these hormones. Delivery of neurotrophic factors may be

important for synaptic plasticity and for certain populations of neurons. For example, in aged rodents and primates, the administration of nerve growth factor (NGF) to the brain can lead to enhancement of cholinergic neurotransmitter circuits and improvement of performance on cognitive tasks related to memory. Axonal transport is required for neurons to convey proteins and other components needed to maintain the structure and function of their lengthy processes. Impaired axonal transport is one factor that has been implicated in the decrease of growth factors in the aging brain.

Inflammation is a mechanism under exquisite regulatory control throughout the life span. Peripheral inflammation has been invoked as one mechanism related to aging in general. For example, a number of cytokines and inflammatory biomarkers, such as C-reactive protein, are increased in serum with aging, and various types of auto-antibodies can be detected with increasing frequency during aging. In the brain, an increase in microglial activation and in other mechanisms of innate immunity may contribute to damage or to lack of appropriate neuronal or synaptic support.

Although there is no obvious selection or survival advantage for genetic factors that influence brain aging, several genes have been identified that may have an impact on brain aging. As an extreme case, rare genetic causes of premature aging syndromes, such as progeria, result in hallmarks of premature aging, such as cataract development, skin changes, increased risk of cancer, and immune alterations during childhood. However, these syndromes are not clearly associated with accelerated brain aging or decline. There are rare genetic mutations associated with early-onset familial AD, inherited in an autosomal dominant pattern. However, these genes have not been shown to have sequence changes that alter the risk or age at onset of late-life AD. The most common variant associated with the risk of late-onset AD is the gene for apolipoprotein E (APOE), which codes for a lipoprotein involved in cholesterol transport within the brain. There are three APOE alleles, e2, e3, and e4. People who carry one copy of the e4 allele have a two- to threefold increased risk of developing AD, whereas those with two copies have a four- to eightfold increased risk. The e2 allele has a small effect on decreasing the risk of developing AD and has also been shown to be overrepresented in people who live into their 90s or 100s.

The ability to generate new neurons in certain regions of the adult brain has also been shown to decrease with aging in animal studies. Although direct data are lacking from studies of human brain tissue, a decrease in the production of neural progenitor cells may play a role in brain aging. A recent study developed an innovative magnetic resonance spectroscopic index of neural progenitor cells and demonstrated that a specific signal could be detected in cultured cells and in animals. A small group of human volunteers spanning the age range from 20 to 70 showed attenuation of this

signal with age in the hippocampus, an area known to harbor progenitor cells. At present, this represents the best available evidence for age-related decline of neural progenitor cells in humans.

STRUCTURAL CHANGES ASSOCIATED WITH BRAIN AGING

Brain atrophy accompanies aging, but the extent to which this represents loss of neurons is controversial. The belief that thousands of neurons are lost every day derives from studies with methodological flaws. More recent studies indicate that neuronal cell loss is less conspicuous and may be selective for certain brain regions.

Methodological issues are important in brain aging studies. By design, aging studies of human brains are cross-sectional. Secular effects, for example, generational changes in nutrition, general health, and comorbidity can complicate comparisons in brains of individuals in their 30s and 40s with those in their 70s and 80s. Brain weight and size are influenced by height and gender. Counts of neurons should attempt to adjust for these factors. Also, the size of brain regions and the number of neurons vary from person to person, by as much as two- to threefold. To detect significant changes in the face of high variability requires a large number of samples or a large effect size. Ideally, subjects whose brains are used in normal aging studies should have been formally evaluated close to their time of death to establish that they did not suffer from cognitive impairment. As mentioned above, the inclusion of brains with AD pathology is inevitable if an aging study includes individuals in their 80s and 90s. However, staging systems of AD pathology have been developed, allowing people with a substantial lesion burden to be excluded.

Postmortem tissue preparation is highly important—early studies that claimed that extensive neuron loss accompanied aging were flawed because of shrinkage of brain tissue during its preparation. Technical approaches to imaging and counting neurons have improved in recent years, including automated counting of neurons based on profiles, and the availability of cell-specific stains. Stereological approaches to sample neurons in a relatively unbiased manner have become routine. Comparative studies of brain tissue from aging primates and rodents have the advantage of minimizing postmortem changes and optimizing the preparation of brain tissue. These provide valuable data that complement direct studies on human tissue.

Total brain weight reaches a maximum in young adults around the age of 20, followed by a progressive decline thought to begin around the ages of 30 to 40 and reaches its nadir after the age of 85 (Table 57.2c–3). By that age, on average, brain weight has decreased by 11 percent compared to the maximum. Similar findings of an age-associated decline have been found in studies of chimpanzee brains. Neuroimaging studies have provided both cross-sectional and longitudinal data showing that brain volume decreases

with age, accompanied by an increase in the volume of the ventricles. Quantitative MRI studies have shown that brain volume decreases across the adult age span. Some studies have suggested that brain volume loss may accelerate with aging, with about 0.1 to 0.2 percent of brain volume lost per year between the ages of 30 and 50, increasing to about 0.3 to 0.5 percent per year after the age of 70. Preclinical disease related to aging, including vascular changes, amyloid and tau deposition may contribute to this apparent acceleration. Although the whole brain, including gray and white matter, shows loss of volume, regions such as the frontal and parietal cortices and hippocampus are affected to a greater degree than the occipital cortex.

Table 57.2c–3.
Structural Changes Associated with Brain Aging

1. Decline of brain weight
2. Neuron loss
3. Neuronal atrophy
4. Synaptic loss
5. Pruning of dendritic trees
6. Nucleolar atrophy
7. White matter changes
8. Gliosis

What explains the decline in brain weight or volume? Neuron loss, as discussed, has been overestimated in the past. After the introduction of stereological methods (including the dissector and its modification, the optical dissector) by Hege Gundersen and colleagues, studies by Bente Pakkenberg and Gundersen reported less substantial loss of neurons across the adult life span, representing about 9.5 percent. Other studies have questioned whether there was substantial loss of gray matter, attributing age-associated declines to loss of white matter.

Neuron atrophy rather than loss could contribute to the age-related decline in brain weight. For example, a classic study by Dr. Robert Terry and colleagues used a computer-assisted method to estimate neuron numbers and size in the superior-temporal, medial-frontal, and inferior-parietal cortical regions from ages 24 to 100. They found that the number of large neurons decreased in all of these regions, accompanied by a corresponding increase in the number of small neurons. Atrophy of pyramidal cells has been well documented. In particular, pyramidal neurons in the hippocampus and cerebral neocortex show atrophy. This has been associated with shrinkage in the nucleolus or loss of cell body RNA.

Regional and Microstructural Changes

The question of whether certain regions are particularly vulnerable to neuron loss with aging has been revisited. Because of their important roles in memory and learning, the hippocampus and its subregions, such as the entorhinal cortex, CA1, and CA2, have been singled out for study. The use of high-resolution MRI, for example, at 4 or 7 Tesla strength allows regional anatomical information to be acquired in great detail. For example, in a study of healthy individuals spanning the age range from 20 to 85, the CA1 subregion of the hippocampus showed the greatest degree of volume loss with age, whereas the entorhinal cortex, CA2, CA3/4, and dentate gyrus showed much less change. How does this compare with pathology studies? Teresa Gomez-Isla and colleagues applied stereological techniques to the entorhinal cortex and found that there was no appreciable loss of neurons in cognitively healthy individuals in their 70s and 80s; by contrast, there was a substantial degree of neuronal loss in AD, even when clinically very mild. When they similarly studied the superior-temporal cortex, an area representative of the association cortex, once again there was no significant loss of neurons with aging. Stereological studies of the CA1 subfields have been inconsistent. Some investigators have found decreases in neurons in CA1 and the subiculum with age, but others failed to find these effects. Reasons for these discrepancies are unclear. Studies of monkeys, using similar stereological methods, have shown no appreciable loss of neurons with age in the entorhinal cortex, CA1, CA2, or subiculum. However, despite the intact number of neurons, decreases in *N*-methyl D-aspartate (NMDA) receptors were evident in aged monkey brains. This raises the possibility that changes short of neuronal death may occur during aging and may provide a substrate that could explain cognitive decline.

Changes in Synapses and Circuits

Neurons function by exchanging information across circuits and networks. The shape and branching of dendrites varies markedly between different types of neurons and within individual neurons of a particular class. Dendrites have protrusions called spines, which contain excitatory synapses. These vary in density across the length of a dendrite, and increase in number with successive branches of dendrites. Even in the absence of overt neuronal loss, changes in dendritic branches or spines could impair neuronal circuit function. A consistent finding in anatomical studies is age related, prominent in pyramidal neurons, in areas such as the prefrontal, precentral, and superior-temporal cortices. Studies using fluorescent dyes have revealed a marked loss of dendritic spines at all levels of dendrites in aged monkeys. Dendritic changes may be linked to levels of neurotransmitters and strength of certain types of neurochemical circuits. For example, the glutamate R2 and NMDAR1 receptors show decreased

expression in monkeys in association with aging, in areas such as the prefrontal cortex for glutamate R2 and along projections from the superior-temporal cortex for NMDAR1. Detailed studies of postsynaptic non-NMDA glutamate receptors in monkeys have suggested that these remain functionally unchanged in primates in areas such as the prefrontal cortex. However, there appears to be an increase in the level of inhibitory synaptic input to pyramidal cells in this area. This could be attributed to increased release of γ-aminobutyric acid (GABA) from presynaptic neurons. These changes in morphology and biochemistry, in the absence of overt neuron loss, could form a substrate for declining performance on cognitive tasks that assess memory, learning, and cognitive processing speed. Of interest, circuits that are highly vulnerable to undergoing degeneration in AD may also display synaptic changes similar to those described above in association with aging.

Neuroimaging and electrophysiological methods (e.g., EEG and magnetoencephalography [MEG]) have been used to develop functional brain maps, and to identify overall and regional brain connectivity. Functional MRI (fMRI) evaluates cerebral responsiveness to activation in human subjects. This takes advantage of the BOLD signal—a signal related to a decrease in deoxyhemoglobin concentration and an increase in cerebral blood flow that accompany local neuronal activation. Such studies can explore the activation of brain circuitry in response to cognitive tasks. Many fMRI studies have been performed to evaluate neural activation of regions and circuits in the brain in relation to aging. For example, the extent of activation of the hippocampus in response to episodic memory tasks decreases in cognitively intact subjects aged 70 or older compared to younger adults. Resting state fMRI takes advantage of correlated regional blood flow changes at short time scales to investigate and map the degree of large-scale coordination between brain systems. A default network of functionally interacting anterior and posterior brain regions is associated with internally directed mental states and is important in forming a substrate for memory acquisition and planning. Early studies exploring aging, for example, those by Buckner and colleagues, reported a large age-related decrease in functional connectivity between these systems that correlated with worse cognitive performance, particularly prominent after the age of 70. The most likely structural basis for this functional decline was disruption of white matter integrity. The ability to image preclinical amyloid pathology by amyloid PET scanning has led to a reappraisal that connectivity changes in cognitively intact elderly individuals are greater in those with positive amyloid scans than without.

In summary, the level of understanding of structural changes in the brain that accompany aging remains descriptive. It will be a great challenge to define the mechanisms that underlie these changes. Brain imaging provides tools to study effects of risk factors that may explain some of the

variability of brain changes in later life. They also may help to interpret the effects of interventions that may decrease the degree of age-associated decline or enhance the function of brain networks or connectivity, as discussed below.

LESIONS IN THE AGING BRAIN

Various types of pathological markers are visible on microscopic examination of brain tissue from elderly individuals (Table 57.2c–4). Markers such as corpora amylacea, lipofuscin, or neuromelanin, may represent cellular responses to aging that are of uncertain clinical consequence because they do not correlate with clinical disorders or with decline in function. Others such as amyloid deposits and neurofibrillary tangles are markers of AD when they exceed a threshold. In elderly people who lack defined disorders such as dementia, it is likely that the overall burden of these lesions may contribute to cognitive decline attributed to aging. For example, circumscribed distribution of tau aggregates has been recognized as a common part of aging, in the absence of significant amyloid deposits, and the acronym primary age-related tauopathy (PART) was recently proposed. PART may be asymptomatic or associated with mild cognitive changes.

Lipofuscin is the most common age-related marker seen in tissues throughout the body. Postmitotic cells such as neurons accumulate lipofuscin more extensively than mitotic cells. Lipofuscin is particularly evident in neurons in the inferior olive, dorsal root ganglia, anterior horn cells, the dentate nucleus of the cerebellum, and large neurons in the cerebral cortex. Lipofuscin consists of protein and lipid, mostly oxidized and contained within lysosomal membranes. Metals, particularly iron, are also present. Lipofuscin is evident as intracellular autofluorescence or can be seen using lipid stains or via electron microscopy. It is thought that lipofuscin represents the remnants of incomplete mitochondrial breakdown by lysosomes. The accumulation of small amounts of lipofuscin may allow cells to sequester aged mitochondria or cross-linked proteins. When large areas within cells are occupied by lipofuscin, it is less clear that this is adaptive, and it is possible that this could adversely affect cellular processes.

Corpora amylacea are round, cytoplasmic glycoprotein inclusions most commonly seen in astrocytes and are inevitably found in the brain after the age of 40. They may arise due to effects of oxidative stress and iron imbalance within astrocytes. It is uncertain whether they represent a cellular defense that sequesters iron or oxidized components of mitochondria or whether their presence may impair the ability of astrocytes to function optimally. The number of corpora amylacea is increased markedly in the brain in AD.

Neuromelanin may be found within neurons as early as 3 years of age

and increases throughout the life span. It is particularly abundant in the substantia nigra and the locus ceruleus. The function of neuromelanin is not known. It contains melanins, derivatives of dopamine, and lipofuscin. Neuromelanin likely forms after the uptake of extracellular dopamine and its oxidation and may represent a means to sequester potentially damaging dopamine byproducts. Neuromelanin avidly binds iron, which may be protective, but could act as a source of oxidative stress in association with aging.

Table 57.2c–4.
Lesions That Accumulate in the Brain with Age

1. Lipofuscin
2. Corpora amylacea
3. Neuromelanin
4. Amyloid deposits, including senile plaques
5. Vascular amyloid
6. Neurofibrillary tangles and argyrophilic grains
7. Lewy bodies
8. Aggregates of TDP43 and hippocampal sclerosis
9. Vascular (ischemic) lesions

Protein misfolding is a prominent feature in many organs of the body accompanying aging and associated with disease. In the brain, aggregates of Aβ form extracellular SP and also can deposit in the walls of meningeal and intraparenchymal arterioles. Tau forms aggregates with the ultrastructure of paired helical filaments. Tau aggregates may be found in dystrophic neurites, argyrophilic grains, and neurofibrillary tangles. Both plaques and tangles are found in association with aging and in greater abundance in patients with AD. Because there is no unique morphology of plaques or tangles that characterize AD but is absent from cognitively intact elderly individuals, pathological criteria used to diagnose AD are quantitative. The presence of both plaques and tangles is necessary for a diagnosis of AD, and the distribution and burden of tangles has a much stronger correlation with cognitive abilities and the presence of dementia than the presence or severity of amyloid. Several sets of pathological criteria for AD have been developed. These have been updated in the light of increasing knowledge of the anatomical distribution of tau lesions. National Institute on Aging and the Reagan Institute (NIA-Reagan) criteria, combine the presence of abundant neuritic plaques with the anatomical extent of tau pathology (limited to entorhinal cortex, spread to hippocampus, or spread to neocortex) to assign a low, intermediate, or high probability that AD is present. These were recently modified to also score the extent and nature of Aβ pathology as a “phase” of brain amyloidosis.

Studies of individuals with intact cognitive performance close to the time of death have reported an age-associated increase in the number of

plaques and tangles. In fact, about one-third of cognitively normal subjects in their 80s meet NIA-Reagan criteria for intermediate probability of AD. In people without dementia, tangles are usually limited to the hippocampal formation. Spread of these lesions to involve association areas in the temporal neocortex, and more widely, is more clearly linked to the development of clinical evidence of cognitive decline and dementia. In very elderly individuals with AD, there is also a lower tangle burden compared to younger individuals with Alzheimer disease. This suggests that brain reserve is diminished in association with aging, so that only a small increase in tangles is necessary to lead to clinical expression of dementia. The question of whether AD pathology contributes to age-associated changes in general cognitive performance is less clear. Autopsy studies have suggested that amyloid deposition and tangle formation may be found to an extent in the brains of people in their 50s. However, tau pathology does not occur through the brain but instead is usually restricted to hippocampal subregions and the entorhinal cortex, and may represent PART. Age-related declines in cognitive speed and executive function correlate weakly with plaques and tangles.

Two other lesions occur variably in association with brain aging. Hippocampal sclerosis, with loss of neurons in the CA1 and sometimes other hippocampal regions accompanied by gliosis, is not considered part of "normal" aging. It increases in prevalence with age, and can be found in about 15 to 20 percent of brains of people over the age of 90. Hippocampal sclerosis contributes to cognitive loss and dementia when present, and is usually associated with aggregates of a protein called TDP43. Another protein, α-synuclein, forms aggregates within neurons, including structures called Lewy bodies, which are perinuclear intraneuronal inclusions. These are the identifying pathology of Parkinson disease. They occur in subcortical areas such as the locus ceruleus and substantia nigra, and in cortical and limbic regions in small numbers and have been reported in up to 20 percent of elders whose brains were studied at autopsy. Like hippocampal sclerosis, α-synuclein aggregates are more likely to be present in the brains of people with cognitive impairment, and these may be better thought of as markers of preclinical disease processes rather than adaptive changes of brain aging.

Age-associated vascular changes in the brain have received increasing focus. The many faces of vascular pathology include cerebral infarcts, lacunar infarcts, microangiopathy, and amyloid angiopathy. Some of these lesions, in particular infarcts, are not inevitable accompaniments of aging but are related to vascular risk factors such as hypertension and diabetes. Cerebral amyloid angiopathy (CAA), with deposition of Aβ in the walls of leptomeningeal and medium-sized and small arterioles, is present in at least 30 percent of elderly individuals and can occur in the absence of the diagnosis of AD. Vessels laden with amyloid may rupture, resulting in lobar

hemorrhage, which may cause neurological deficits. MRI studies indicate that microhemorrhages, visible on gradient echo MRI sequences, may occur in association with CAA, and if extensive, are associated with impaired cognition. The widespread use of computed tomography (CT) imaging and more recently MRI in studies of aging and cognition has led to an appreciation of white matter changes. On CT, these appear as a periventricular halo of hypodensity. On MRI, they are most clearly seen on T_2-weighted images, where they have various appearances, including periventricular hyperintensity, capping of the ventricular horns, and hyperintense spots of varying size, which may coalesce to form larger cotton–wool-like confluent areas in the fiber tract projections of the centrum semiovale. Initially referred to as leukoaraiosis, these changes were attributed to incomplete ischemias affecting the white matter tracts. Capping of the ventricles is not related to ischemia but to age-related disruption of the ependymal layer, increased accumulation of interstitial fluid, and subependymal gliosis. Neuropathological findings vary, and often include pallor and loss of myelin. Cognitive testing in groups of elderly individuals with white matter changes has revealed decreased performance on timed tests and on tasks dependent on executive function. White matter changes are more extensive on MRI in people with vascular risk factors such as diabetes and hypertension. These vascular factors may worsen the severity of age-dependent changes in the structure of long projecting tracts that extend through the white matter.

Detailed microscopic studies of blood vessel morphology have raised further questions about vascular factors in age-associated cognitive decline. For example, a decrease in the diameter of capillaries in hippocampal subregions, independent of AD-related lesions, raises the possibility that microvascular changes could contribute to neuron loss or damage. It is possible that these changes could alter the available cerebral blood flow to neurons in the hippocampus, resulting in oxidative stress, damage, and eventually death of neurons. Intervention to treat hypertension and diabetes in middle and later life could help to forestall some of the cognitive decline attributed to aging, although clinical trials of antihypertensive treatment have had mixed results on cognitive outcomes to date.

Another vascular measure is cerebral blood flow, which can be measured using positron emission tomography (PET). More recently, the magnetic resonance technique of arterial spin labeling (ASL) has been used to measure cerebral blood flow. This method magnetically tags water protons in the carotid and vertebral arteries at the base of the brain, the major conduits of blood supply to the brain. As these tagged protons enter the brain, they alter the local magnetization state in proportion to blood flow. There have been relatively few large-scale studies of cerebral blood flow with aging. In general, total cerebral blood flow decreases with age, with studies reporting changes of about 0.5 to 0.6 percent/year. Recent

studies that have examined regional cerebral blood flow have identified heterogeneous changes: decreases with age seem to occur mainly in limbic and tertiary association cortices, with little change in regions such as the primary visual cortex. Vascular risk factors may adversely affect global and regional cerebral blood flow.

NEUROCHEMICAL CHANGES IN AGING

Many neurotransmitters and neuropeptides in the brain show age-related alterations. Two of these, dopamine and acetylcholine, will be highlighted because of their importance in cognitive and affective function, and their potential importance in neuropsychiatric disorders. Much of the decline in neurotransmitter concentrations is due to presynaptic changes, often secondary to loss or atrophy of subcortical projecting neurons, for example, the pigmented cells of the locus ceruleus and substantia nigra (which produce norepinephrine and dopamine, respectively) or the nucleus basalis (which produces acetylcholine). Aging may also affect postsynaptic processes that affect transmitter function, including decreased density of postsynaptic receptors and possible decreased efficiency of signal transduction.

Populations of dopaminergic neurons show marked changes during aging. In monkeys, the number of neurons in the substantia nigra that stain positive for tyrosine hydroxylase decreases by as much as 40 to 50 percent. This may be associated with increased α-synuclein content in individual neurons, a potential factor related to decreased strength of dopaminergic transmission in the nigrostriatal system. These changes may account for slowing of movement and possibly, slowing of cognitive processes in association with aging. Recently, detailed anatomical studies in primates have shown that dopamine innervation from the midbrain projects to cover extensive areas throughout the cortex. Two broad groups have been identified, ventral tier neurons that project predominantly to the striatum, and dorsal tier, whose cells express calbindin and project to limbic, cortical, and selected striatal areas. More detailed anatomical studies may be required to evaluate the specific contribution of these different types of cells in relation to aging. A decline in the strength of dopamine transmission in the mesolimbic and mesocortical pathways may contribute to increased vulnerability to disorders such as depression in aging.

PET imaging has been used to obtain quantitative data for indices of dopaminergic neurotransmission during life. Studies have found that there are age-associated declines in levels of the dopamine transporter, dopamine D_2 receptors, $5HT_2$ receptors, and $5HT_{1a}$ receptors. Differential patterns of decline with aging have been reported. For example, dopamine transporters and D_2 receptors show a linear decline with age, but serotonin $5HT_2$ receptors decline rapidly during the fifth decade of life, and then

decline more slowly.

Cholinergic neurons in the nucleus basalis of Meynert and the septal nuclei project widely to the hippocampus and neocortex. They function to activate regions of the brain in relation to sustained attention, long-term potentiation, and memory formation. In animal studies, lesions to cholinergic nuclei result in deficits in memory abilities. Aged rodents, monkeys, and humans show a decrease in the levels of markers of the cholinergic system with age. In AD, there is a more dramatic decrease which parallels the severity of dementia. Although the mechanisms of these age-related declines are unknown, one possible factor is diminished neurotrophic support. Cholinergic neurons respond to NGF, and it is possible that decreased retrograde transport of NGF from targets to the cell body may result in atrophy of the cholinergic cell bodies. Treatment of aged animals, with grafted cells genetically engineered to secrete NGF, resulted in improvement of cognitive test performance and increased levels of cholinergic markers. However, human studies of gene therapy with NGF in AD have been negative.

Table 57.2c–5.
Brain Performance Measures That Worsen with Aging

1. Cognition
   - Mental speed
   - Fluid reasoning
   - Executive function
   - Retrieval
   - Episodic memory
   - Free recall worse than recognition
2. Movement
   - Gait slowing
   - Simple reaction time slowing
   - Upward gaze
   - Balance changes (vestibular, sensory, motor, and brain)

CHANGES IN COGNITIVE AND MOTOR ABILITIES ASSOCIATED WITH AGING

Much research has documented changes in cognitive abilities, in particular memory, accompanying aging (Table 57.2c–5). On tests of episodic memory, both the acquisition of new verbal or visual information and recall after a delay show decline with age. These are evident even when comparing the performance of 40 year olds to that of 20 year olds. On average, subjects at age 80 perform approximately 1.0 to 1.5 standard deviations below the mean for young adults. Whereas performance appears to decrease throughout the adult age span, decline may be a little steeper after the age of 60. Other aspects of cognition, including working memory, perceptual speed, spatial visualization, and inductive reasoning show similar age-related decline. Age-related effects on different types of

cognitive tests are correlated, suggesting that widespread brain changes underlie them. One such factor is cognitive slowing; if time limits are removed from certain types of cognitive tests and the criterion for scoring is accuracy, then elderly individuals show improved performance. Semantic memory and cumulative knowledge, as indexed by measures such as vocabulary, tend to increase with age and remain preserved in old age. However, retrieval of semantic knowledge, including words and names, is impaired in aging, as demonstrated by performance on timed tests.

Many aspects of movement show evidence of slowing with aging. One of the more obvious features is the slowing of gait and tendency to stoop that are so common in the oldest-old. Elderly individuals also show a decline in balance: difficulty being able to stand on one leg or perform a tandem walk usually begins after the age of 70. This is most likely multifactorial, with contributions from declines in nerve and muscle function, changes within the spinal cord, sensory systems (vision and vestibular function), and brain changes involving structures such as the basal ganglia. Simple reaction time and movement time are both slowed with aging, resulting from changes to white matter tracts as well as neuronal networks that initiate and coordinate rapid movements.

AGE-ASSOCIATED BRAIN DISEASES

A number of disorders show a predilection for the aged brain (Table 57.2c–6). Neurodegenerative diseases rank high on the list; among these, AD is the brain disorder most intimately related to aging. Both the incidence and the prevalence of AD show an exponential increase after the age of 60. Parkinson disease has a prevalence of approximately 1 to 2 percent in the general population. The prevalence of Parkinson disease increases with aging and may be as high as 4 to 5 percent in individuals aged 85 or older. However, the frequency of Lewy bodies—the signature pathology of Parkinson disease—in the substantia nigra in series of autopsies from people aged 60 to 90 is higher, typically around 5 to 8 percent. This is consistent with the concept that reserve may protect against the clinical expression of Parkinson disease. Amyotrophic lateral sclerosis (ALS) is much less common than AD or Parkinson disease. Its onset is more common among people aged 60 or older, than at a younger age in the absence of a causative gene mutation.

Table 57.2c–6.
Age-Associated Brain Diseases

1. Alzheimer disease
2. Vascular dementia
3. Normal pressure hydrocephalus
4. Parkinson disease
5. Amyotrophic lateral sclerosis

6. Depression
7. Seizures
8. Sleep disorders
9. Essential tremor

The incidence of stroke increases with age, including both clinically apparent strokes and evidence of stroke on neuroimaging in the absence of a history of stroke (silent infarcts). In population-based studies, silent infarcts are much more common than symptomatic strokes in the elderly, and the overall prevalence of infarcts on MRI may be as high as 20 percent in the elderly. The prevalence of MRI-defined infarcts is associated with aging and with cardiovascular risk factors such as hypertension.

BRAIN AGING IN THE OLDEST-OLD AND IN CENTENARIANS

According to the US 2000 census, there were over 50,000 Americans older than 100 years. Living to 100 or older is a milestone that is especially worthy of celebration if it is accompanied by reasonable physical and cognitive function. A number of studies have tried to determine the prevalence of dementia in centenarians. Difficulty with performing daily activities is common among centenarians, often due to physical impairment or changes in vision and hearing. Because of marked differences in methods of cognitive assessment and incomplete sample ascertainment, estimates vary widely. Longitudinal data such as those that have emerged from the 90+ study of nonagenarians living in Leisure World in Southern California, have identified a marked acceleration in the risk of dementia after the age of 90 with a prevalence of at least 50 percent by age 100.

In the oldest-old, correlations between pathological lesions of AD and cognitive status are weaker than those in younger individuals. For example, the Oregon Brain Aging Study followed people aged 85 or older without cardiovascular or major age-associated diseases. Subjects who died without having shown cognitive decline or dementia had scanty SP or NFT, although there was a correlation between the presence of increasing burden of these lesions and impaired performance on cognitive testing. The Leisure World 90+ study includes cognitive testing, clinical evaluation for the presence or absence of dementia, and a high rate of brain examination at autopsy. Whereas counts of SP weakly distinguished people with dementia from those without, the strongest discriminating features for dementia were reduction of presynaptic biochemical markers, overall tangle burden and the presence of hippocampal sclerosis. Synaptic density may represent a dynamic and adaptive marker that is initially related to the establishment of connectivity during development, remodeled according to use and plasticity, and undermined by age- and disease-related processes.

PROSPECTS FOR HEALTHY BRAIN AGING

In the absence of effective treatment that slows the progression of AD, it is important to consider how to prevent AD and preserve cognitive abilities during aging. Population-based studies of aging and dementia have suggested that there may be modifiable factors associated with the risk of dementia and or age-associated cognitive decline. These include aspects related to lifestyle and diet as well as medical interventions (Table 57.2c–7). These types of interventions are being evaluated in randomized clinical trials beginning in mid or later life. Early results from one of these studies conducted in Finland are encouraging, suggesting that cognitive performance may improve as a result of multicomponent intervention in people who have an unhealthy lifestyle.

Converging evidence suggests that physical activity may modify brain aging. In animal studies, maintained physical activity or exercise is linked to decreased markers of oxidative damage in the brain with aging, increased neurogenesis, increased brain perfusion, and lower expression of amyloid pathology in transgenic mice used as models of plaque deposition and AD. Epidemiological studies have shown that regular physical activity is associated with a decreased risk of developing AD. Vigorous aerobic exercise may not be necessary, as evidence suggests that walking may be associated with risk reduction. Self-reported higher levels of exercise have been associated with better cognitive test scores in elderly individuals. There have been many controlled trials of exercise intervention in elderly subjects; a recent meta-analysis of 18 randomized trials found that aerobic exercise training had beneficial effects for healthy older adults who did not previously exercise regularly. Effects were stronger for tests of executive function and reaction time than for other cognitive processes and were statistically significant only for interventions that used exercise sessions that were at least 30 minutes long. Before making a firm recommendation that exercise will benefit cognitive performance or delay dementia, further research is needed to confirm the findings of intervention studies, establish what types of exercise are beneficial, and determine how long and frequent each exercise session should be. In addition to defining a minimal amount of exercise that may show a positive cognitive effect, studies need to identify exercise regimens that are feasible for most elderly individuals. Along the way, it will be important to clarify how exercise benefits the brain.

Cognitive stimulation, using methods such as environmental enrichment, is associated with reduced deposition of amyloid in transgenic mouse models related to AD. A number of human studies have found that engagement in cognitively demanding pursuits throughout adult life is associated with better cognitive performance and with a decreased risk of incident AD. For example, the Rush Memory and Aging Project enrolled over 1,200 people in Chicago, with an average age of 80, who were

determined to be free of dementia and followed and reevaluated this cohort annually. Subjects were questioned about types of cognitively active pursuits (such as reading books, attending music or plays, playing card games) and their frequencies. Those subjects who had a moderate or high level of cognitive activity had a lower rate of age-associated cognitive decline, which remained significant even after controlling for age, race, education, depressive symptoms, and chronic medical conditions. Further analyses have shown that subjects with higher levels of cognitively demanding pursuits had a 2.6-fold lower risk of developing dementia than subjects with a low level. Other studies have reported similar types of findings, which have given rise to the "use it or lose it" hypothesis. Many factors have been related to higher levels of cognitive activity in later life, including general health, physical function and mobility, and differences in diet and lifestyle, that may be part of a package that contributes to brain function in aging. Many intervention studies have been carried out in this area, and in recent years, these have included computerized training of cognition. The Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) study was a large-scale intervention that enrolled over 2,800 subjects aged 65 or older. Participants were randomly assigned to receive specific modules of training for memory, reasoning, speed of processing, or no intervention. A subset of subjects received booster training at 11 or 35 months after enrollment. Cognitive training resulted in improvement in the area of cognitive performance that was trained, but did not carry over into other areas. The effect persisted for the 5 years of the study. Reasoning training resulted in slightly less decline on self-reported instrumental activities of daily living; otherwise there was no functional benefit associated with training. This rigorous study highlights a key feature of cognitive intervention, training to task. Computer-assisted training appears to influence the specific area that is trained, and it has been difficult to show broader improvement in other cognitive abilities, or an impact of training on functional abilities. Unless such evidence emerges, there is no clear reason to recommend computer games and interventions to improve brain health and function in aging.

Table 57.2c–7.
Interventions Proposed to Help Maintain Brain Abilities with Aging

1. Control hypertension
2. Treat diabetes and vascular risk factors
3. Regular physical activity
4. Mental activity, e.g., cognitively demanding pastimes
5. Social activity, e.g., friend, family, and community interactions
6. Diet: similar components to a heart-healthy diet, e.g., a diet rich in antioxidants and low in saturated fat, or a Mediterranean-type diet

Nutrition and brain aging is another area of controversy. It is difficult to randomize people to different types of diet for long-term studies, so research efforts in this area have used dietary questionnaires, short-term interventions, and supplements with vitamins or other nutrients. Many components of diet have been linked to cognitive performance with aging. Some examples of factors that may be protective include eating a Mediterranean or heart-healthy diet or consuming foods with a relatively high content of naturally occurring vitamins with antioxidant activities such as vitamin E and vitamin C. Studies of diet need to consider other factors, such as obesity, cardiovascular risk factors, and lifestyle that may also be associated with dietary choices. Intervention studies regarding supplemental antioxidant vitamins, such as vitamin E and vitamin C, have not shown cognitive benefits or reduction of the risk of developing dementia. Homocysteine levels have been associated with the risk of AD and stroke in several studies. In the United States, fortification of grain products with B group vitamins has lowered the population's homocysteine levels. Intervention studies of lowering homocysteine in people with AD had no benefits on cognition. Omega-3 fatty acids are important nutrients for brain function and growth during development. Clinical trials that have used omega-3 supplementation have had no benefits in AD, and have not shown clear effects in altering cognitive decline associated with aging.

Cardiovascular risk factors, in particular hypertension and diabetes, are associated with worse cognitive function with aging. For example, in the ACTIVE study discussed above, participants with hypertension had a faster decline on logical reasoning tasks, and diabetes was associated with faster decline on a test of cognitive processing speed. Population-based studies have found that midlife hypertension is a consistent risk factor for accelerated age-associated cognitive decline, stroke, and incident AD. However, although clinical trials of treatment of hypertension in elderly individuals have shown clear benefits for reducing heart attack and stroke, they have had less consistent benefits on slowing cognitive decline. These studies have been limited by problems such as dropout during long periods of follow-up and the difficulty of maintaining an untreated comparison group. As obesity has increased in adult populations in Western societies, attention has focused on whether metabolic syndrome is associated with cognitive performance in the elderly. Studies have found that components of metabolic syndrome, such as diabetes and abdominal girth, appear to be related to poorer performance on a variety of cognitive measures in cross-sectional studies in the elderly. Much of this risk appears to be due to hyperglycemia, although in some studies biomarkers of inflammation measured in the blood also contribute. Clinical trials are currently in progress to examine whether interventions to maintain stringent control of type 2 diabetes are associated with benefits in cognitive performance.

Even in the absence of evidence from rigorous intervention studies, it

seems reasonable to recommend maintaining physical and cognitive activity, controlling cardiovascular risk factors, and eating a heart-healthy diet to try to preserve general health with aging, with the bonus that there may be positive spin-offs regarding cognitive abilities. These aspects of lifestyle were included in the recommendations of a recent report prepared by the Institute of Medicine. That report also mentioned avoiding delirium and being cautious in use of prescription medications with potential cognitive side effects. It also highlighted the important planning roles that families, communities, and government need to take in developing the environment to support healthy cognitive aging.

Our ability to describe and quantify changes in the brain associated with aging has burgeoned in the past few decades, with parallel advances in studying aging in animal models. Advances in brain imaging continue to have a huge impact on human studies of aging and cognition. Although great progress has been made at a descriptive level, researchers are still handicapped by not understanding which mechanisms are most important in driving the aging process and how they interact to result in biochemical and structural changes in the brain. Because aging is associated with a decline in brain abilities as well as an increased vulnerability to disorders such as dementia and cerebrovascular-associated cognitive impairment, continued research into mechanisms that lead to damage as well as those that support resilience and plasticity within the brain could pay an enormous dividend by improving health and vitality as society continues to benefit from gains in longevity.

References

*Andrews-Hanna JR, Snyder AZ, Vincent JL, et al. Disruption of large-scale brain systems in advanced aging. *Neuron*. 2007;56:924–935.

Biagi L, Abbruzzese A, Bianchi MC, Alsop DC, Del Guerra A, Tosetti M. Age dependence of cerebral perfusion assessed by magnetic resonance continuous arterial spin labeling. *J Magn Reson Imaging*. 2007;25:696–702.

Brier MR, Thomas JB, Snyder AZ, et al. Unrecognized preclinical Alzheimer disease confounds rs-fcMRI studies of normal aging. *Neurology*. 2014;83:1613–1619.

Calvert JF Jr, Hollander-Rodriguez J, Kaye J, Leahy M. Dementia-free survival among centenarians: an evidence-based review. *J Gerontol A Biol Sci Med Sci*. 2006;61:951–956.

Colcombe S, Kramer AF. Fitness effects on the cognitive function of older adults: a meta-analytic study. *Psychol Sci*. 2003;14:125–130.

Corrada MM, Brookmeyer R, Paganini-Hill A, Berlau D, Kawas CH. Dementia incidence continues to increase with age in the oldest old: the 90+ study. *Ann Neurol*. 2010;67:114–121.

Crary JF, Trojanowski JQ, Schneider JA, et al. Primary age-related tauopathy (PART): a common pathology associated with human aging. *Acta Neuropathol*. 2014;128:755–766.

*Dickstein DL, Kabaso D, Rocher AB, Luebke JI, Wearne SL, Hof PR. Changes in the structural complexity of the aged brain. *Aging Cell*. 2007;6:275–284.

Drachman DA. Aging of the brain, entropy, and Alzheimer disease. *Neurology*. 2006;67:1340–1352.

Gorelick PB, Scuteri A, Black SE, et al. Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*. 2011;42:2672–2713.

Green MS, Kaye JA, Ball MJ. The Oregon brain aging study: neuropathology accompanying healthy aging in the oldest old. *Neurology*. 2000;54:105–113.

Harmell AL, Jeste D, Depp C. Strategies for successful aging: a research update. *Curr Psychiatry Rep*. 2014;16:476.

Hayflick L. Biological aging is no longer an unsolved problem. *Ann N Y Acad Sci*. 2007;1100:1–13.

Iadecola C. The pathobiology of vascular dementia. *Neuron*. 2013;80:844–866.

*Imhof A, Kövari E, von Gunten A, et al. Morphological substrates of cognitive decline in nonagenarians and centenarians: a new paradigm? *J Neurol Sci*. 2007;257:72–79.

Jack CR Jr, Wiste HJ, Weigand SD, et al. Age, sex, and APOE ε4 effects on memory, brain structure, and β-amyloid across the adult life span. *JAMA Neurol*. 2015;72:511–519.

Karr JE, Areshenkoff CN, Rast P, Garcia-Barrera MA. An empirical comparison of the therapeutic benefits of physical exercise and cognitive training on the executive functions of older adults: a meta-analysis of controlled trials. *Neuropsychology*. 2014;28:829–845.

Klempin F, Kempermann G. Adult hippocampal neurogenesis and aging. *Eur Arch Psychiatry Clin Neurosci*. 2007;257:271–280.

Martin M, Clare L, Altgassen AM, Cameron MH, Zehnder F. Cognition-based interventions for healthy older people and people with mild cognitive impairment. *Cochrane Database Syst Rev*. 2011;(1):CD006220.

Masliah E, Mallory M, Hansen L, DeTeresa R, Terry RD. Quantitative synaptic alterations in the human neocortex during normal aging. *Neurology*. 1993;43:192–197.

Morrison JH, Baxter MG. The ageing cortical synapse: hallmarks and implications for cognitive decline. *Nat Rev Neurosci*. 2012;13:240–250.

Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *Lancet*. 2015;385:2255–2263.

*Pakkenberg B, Gundersen HJ. Neocortical neuron number in humans: effect of sex and age. *J Comp Neurol*. 1997;384:312–320.

Rahimi J, Kovacs GG. Prevalence of mixed pathologies in the aging brain. *Alzheimers Res Ther*. 2014;6:82.

Robinson JL, Molina-Porcel L, Corrada MM, et al. Perforant path synaptic loss correlates with cognitive impairment and Alzheimer's disease in the oldest-old. *Brain*. 2014;137(Pt 9):2578–2587.

Salthouse TA. When does age-related cognitive decline begin? *Neurobiol Aging*. 2009;30:507–514.

Stephan BC, Matthews FE, McKeith IG, Bond J, Brayne C; Medical Research Council Cognitive Function and Aging Study. Early cognitive change in the general population: how do different definitions work? *J Am Geriatr Soc*. 2007;55:1534–1540.

*Stern Y. Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurol*. 2012;11:1006–1012.

Wilson RS, Bennett DA, Bienias JL, Mendes de Leon CF, Morris MC, Evans DA. Cognitive activity and cognitive decline in a biracial community population. *Neurology*. 2003;61:812–816.

57.2d Psychological Changes with Normal Aging

JENNIFER J. DUNKIN, PH.D., AND KIMBERLY ANN CORP, M.S., M.F.T.

INTRODUCTION

Advances in medicine and improvements in living conditions over the past century have resulted in greater and greater number of persons reaching old age. This "graying" of the population is expected to reach its peak over the next 60 years, as the Baby Boom cohort reaches old age. The oldest-old group of elderly individuals—those 85 years of age and older—is expected to grow at the fastest rate. Although on the surface these demographic trends reflect favorably on society's attempts to eradicate disease, reduce infant mortality, and improve standards of living, recent debates have focused on the "hidden costs" of these changes. For example, aging is associated with increased risks of, among other factors, physical illness, personal losses, and cognitive impairment. If these risks translate into significant impairment for the majority of older people, then there will be increased costs to society as greater and greater number of people reach old age.

From both public health and quality-of-life perspectives, it is important to determine how to maximize the health and well-being of the aged population. It is with this in mind that a number of large-scale longitudinal studies have focused on the concept of "successful aging" to determine what characteristics are present in individuals who appear to reach old age with favorable outcomes. The results of such studies have been important in shaping public policy as well as changing societal views on the nature of the aging process itself. For example, although there are clearly unique life challenges that face the elderly, it may be surprising that the majority of research studies suggest that most elderly people report high levels of life satisfaction, cope well with the challenges of late life, and demonstrate considerable resilience in the face of functional and personal losses. Nevertheless, recent studies suggest that even successfully aging individuals are prone to their own negative aging self-stereotypes and that these self-stereotypes cause distress and appear to be related to poorer functional health. Thus, the challenges as a society are to increase the number of elderly persons reaching old age successfully and to decrease negative views of the aging process.

DESIGN AND SAMPLING ISSUES

Design Issues

Most studies in the field of aging are cross-sectional in design, meaning that groups of individuals of different ages are sampled at one point in time. Although meaningful conclusions about differences between age

groups can be drawn from cross-sectional studies, it is inappropriate to draw inferences about change over time from such studies, because individuals are not studied prospectively and there is considerable potential for mistaking cohort differences for longitudinal change. For example, there are vast differences between the average 30-year-old and the average 70-year-old besides age, including average number of years of education, exposure to historical events such as the Great Depression and World War II, and comfort with revealing personal information. Thus, what may resemble change over time may instead be baseline cohort differences. Because of these issues, it is likely that different trajectories of change will be necessary for each individual age cohort. In addition, cross-sectional designs tend to overestimate the relationships between variables that change with time.

Sampling Issues

Until recently, most studies of psychological issues in aging have used predominantly Caucasian, well-educated, healthy, and financially comfortable samples of older adults. Studies have primarily used samples of subjects in the so-called young-old age group (65 to 75 years of age), or, less frequently, the old-old age group (75 to 85 years of age), but rarely the oldest-old group (85 years of age and older). Furthermore, early longitudinal studies of personality development and adjustment used only male samples. The result is that comparatively little is known about minority elderly, elderly women, inner-city and lower-income elderly, and the oldest-old. Similarly, many psychometric and neuropsychological tests that measure cognitive functioning do not have adequate normative data for the oldest age groups. Attrition rates are also important phenomena: Studies rarely compare individuals who drop out of studies with those who remain in the study. This may overestimate the amount of stability and underestimate the amount of decline in psychological functions, as those who decline in functioning tend to have higher attrition rates than those who do not decline. In terms of cognitive aging, studies have varied considerably in terms of whether individuals with medical conditions known to affect brain function were included in the sample. Although such individuals may be “typical” of their age cohort, their inclusion in studies makes it difficult to determine which cognitive changes can be expected as a function of the aging process itself and which changes may be a function of epiphenomena associated with aging, such as medical conditions.

MENTAL HEALTH AND PERSONALITY

Mental Health over the Life Span

With the exception of dementing disorders and delirium, mental disorders are no more common in the elderly as compared to younger age groups.

Epidemiological studies have found that prevalence estimates of common disorders such as major depression, anxiety disorders, substance abuse, and psychotic disorders are generally lower for those over the age of 65 versus younger age cohorts, a finding recently replicated by newly revised prevalence estimates of mental disorders using the NIMH Epidemiological Catchment Area Program data, as well as the National Comorbidity Survey. Furthermore, recent longitudinal studies that have followed individuals for up to 25 years have found no evidence that those rates of emotional distress increase with age.

Age differences in the experience and expression of both positive and negative emotions have also been studied to determine if older adults report more negative emotional experiences than younger adults. In a recent study using a series of large ethnically diverse samples, older adults reported experiencing fewer negative emotions such as sadness, anger, and fear but more positive emotions such as happiness than those in the younger age cohorts. In another study, even among older persons with physical illness, more positive than negative emotions were observed, lending further support to the notion that even advanced age and illness do not necessarily lead to negativity in outlook. In addition, older adults reported a greater sense of emotional control over negative internal emotional states than younger subjects. To explain their data, the authors proposed an *emotional control* model, suggesting that aging is associated with an increased ability to inhibit negative emotional states and maintain positive emotional states.

A recent study found interesting racial differences between elderly blacks and whites on measures of psychological distress. The study included over 1,000 high-functioning older adults between the ages of 70 and 79 and found that blacks were less distressed than whites, particularly at lower educational levels. The finding remained even after controlling for social support, socioeconomic status, and functional status. The authors suggest that elderly blacks may be afforded higher status and stay more engaged within their communities than elderly whites. If so, then the place of older blacks within their communities could serve as a model for other ethnic or racial groups.

Stage Theories of Personality Development

Early personality theorists proposed that development was completed by the end of childhood or adolescence. One of the first developmental theorists to propose that personality continues to develop and grow over the course of the life span was Erik Erikson. Erikson believed that development proceeded through a series of *psychosocial stages*, each with its own conflict that is resolved by the individual with a greater or lesser degree of success. Erikson termed the crisis of the last epoch of life *integrity versus despair* and believed that successful resolution of this

crisis involved a process of life review and achieving a sense of peace and wisdom through coming to terms with how one's life was lived. For example, Erikson proposed that successful resolution of this crisis would be characterized by a sense of having lived one's life well, whereas a less successful resolution would be characterized by feeling that life was too short, that one did not choose wisely, and a sense of bitterness that one will not have another chance to live life over.

Erikson's theory is descriptive and few studies have attempted to empirically validate it. However, a recent longitudinal study of approximately 500 subjects from two age cohorts found not only that the earlier age cohort scored significantly higher on a measure of integrity than the later age cohort but also that scores for both age cohorts on this measure had declined significantly by the final time of testing. These data suggest that the conflict of integrity versus despair may have a more favorable outcome in earlier age cohorts versus later age cohorts, raising the possibility that changing societal values have negatively impacted the struggle for integrity. Although more empirical data are needed, Erikson's theory remains one of the most eloquent descriptions of the potential of late life.

Personality over the Life Span: Stability or Change?

Although Erikson and other stage theorists focused on unique developmental tasks and stages central to each phase of life, other studies have focused on defining core personality traits within the individual and determining their course over the life span. Studies have consistently revealed five basic personality traits: extraversion, neuroticism, agreeableness, openness to experience, and conscientiousness. Further, these personality traits generally remain remarkably stable over the life span. A number of large, well-designed longitudinal studies that have followed individuals over periods ranging from 10 to 50 years have found strong evidence for stability, although with a few notable exceptions. For example, some studies have found that neuroticism decreases as individuals move from young and middle adulthood to old age, after which it remains stable. Other studies have found slight decreases in extraversion as individuals move into the oldest-old age category, possibly linked to deterioration in health and sensory loss, particularly hearing.

Personality factors may even contribute to mortality in old age. For example, recent studies have found that high levels of extraversion and low levels of neuroticism are associated with a reduced risk of mortality in individuals 65 years of age and older. One explanation for this finding may be that personality traits mediate physical, social, and cognitive activity levels that lead to less healthy and supportive lifestyles.

ADJUSTMENT AND COPING IN THE ELDERLY

Individuals can expect to experience periods of change, disruption, and even crisis throughout the life span. There is no doubt that older adults experience significant challenges as they enter and negotiate the last phase of life, including bereavement, loss of socially valued roles such as that of worker, and chronic and accumulating health problems. How age affects the process of managing change and discontinuity has been an important research focus, with some authors proposing that people become less effective as they age, due to depletion in resources, whereas others have hypothesized that people become better at managing stressful life events over time, secondary to a process of learning more effective strategies. These questions have been investigated from a wide range of theoretical perspectives.

Age Differences in Coping Strategies and Coping Effectiveness

"Coping" is a general term that refers to strategies that individuals use to manage stressful life situations that involve perceived or actual threats. It includes strategies designed to reduce emotional distress, such as distraction, passive avoidance, and positive reappraisal, and strategies designed to solve the problem, such as direct action, confrontation, and information-seeking. Cross-sectional studies of age differences in coping strategies have had somewhat mixed results, with some studies finding no differences between age groups in their use of coping strategies, and other studies finding that older adults preferentially use strategies designed to reduce emotional distress, whereas younger adults use more problem-focused coping, although the effect sizes are relatively small. This finding has often led to the conclusion that older adults are more passive in their coping style, yet when the type of stressor is controlled, these differences are often eliminated or reduced. In addition, older adults are often attempting to cope with situations that are less changeable than those faced by younger individuals, and therefore, the use of coping techniques that attempt to reduce emotional distress rather than solve the problem may be a more effective approach. In this vein, a recent study found that older adults who modify their goals and accommodate and adapt to losses and stressors rather attempting to exert control or deny losses have a better outcome.

Coping effectiveness across the life span has been another area of interest for gerontologists. The effectiveness of a coping strategy is defined by its outcome, usually measured by level of distress, depression, or health symptomatology. Clearly, if individuals become less effective at coping with the demands of life as they age, this would have considerable implications for clinical practice and the kinds of interventions that might be used with older adults. However, studies of coping effectiveness have generally not revealed age differences; that is, the old appear to cope as well as the young, and their outcome, in terms of depression and other stress-related

symptomatology, is just as favorable. For example, in a study investigating age differences in coping across a broad range of stressors using a sample of over 2,000 men ranging in age from the late 40s to over 90, there were no significant age differences in the reporting of negative emotional states in response to stressors or perceived efficacy in coping.

Role Transitions

Role transitions play an important part in the aging process, and there is no domain where it is more important to understand issues of coping and adaptation than during transitional phases of life. There are many transitions experienced by the elderly, but older adults are especially likely to experience two major role transitions: Retirement, or the transition from worker to nonworker, and bereavement, or the transition from spouse to widow. These major transitions present the individual with the challenge of having to redefine identity and to create new roles to replace those that have been lost. Indeed, there is evidence that the elderly who are successful in creating new roles for themselves experience a greater sense of well-being than those who are less successful at this process.

Retirement. Retirement is an age-graded event, in that individuals in society are expected to retire at a particular age, regardless of health status or personal preference. It is almost entirely a phenomenon of Western, industrialized, and postindustrialized societies and prior to the late 19th century was largely a personal decision. The transition and adjustment to retirement has been characterized by social gerontologists as a *psychosocial crisis*, and the image of the successful executive deteriorating or dying soon after retirement has been popularized in the press. However, recent studies have suggested that most recent retirees rank retirement quite low in terms of its impact on a list of psychosocial stressors, with 17 percent of men and 11 percent of women reporting dissatisfaction with retirement. Nevertheless, there is some heterogeneity among retirees, and the factors that predict a poor adjustment to retirement include an involuntary or unexpected retirement, low socioeconomic status, health problems, a competitive and highly achievement oriented personality (the so-called Type A behavior pattern), inability to set future goals, and general difficulty adjusting to change. Further, despite early data suggesting that there is an increased risk of death in the first 2 years after retirement, recent studies controlling for health status at retirement have failed to reveal any associations between retirement and mortality. In addition, when preretirement mental and physical health are controlled for, there are no clear associations between retirement and subsequent physical or mental health consequences. Thus, it is important to consider retirement in the context of an individual's life and other changes that they may be experiencing rather than in isolation.

Few studies of retirement have specifically investigated ethnic differences in adjustment. For many ethnic minority groups, retirement represents a significant loss of income, because many are economically disadvantaged and earn lower wages relative to their Caucasian counterparts. Available data suggest that many continue to work past retirement age even in the face of disability and health problems. A recent study suggests that middle-aged African-American males can expect to spend, on average, 2.5 fewer years retired than Caucasian males and therefore can expect to spend more time in the workforce with fewer years of retirement and leisure in their old age.

Spousal Bereavement. Demographic data suggest that 51 percent of women and 14 percent of men over the age of 65 will be widowed at least once. Bereaved spouses consistently rank spousal loss as among the most stressful of all life experiences. Early work on bereavement and the elderly hypothesized that older adults might be less able to cope with the loss of a spouse than younger adults, due to declining physical capabilities and subsequent decrements in the ability to manage stress. Subsequent research has revealed age differences in coping with bereavement but in the opposite direction as predicted: Younger bereaved spouses report greater symptomatology on measures of physical and mental health following bereavement than older spouses. Several factors might account for this difference, including cohort effects, with younger cohorts having a greater comfort level in acknowledging symptoms, or whether the death was expected or unexpected, with unexpected deaths causing greater symptomatology and occurring with greater frequency in younger age groups.

Other recent data are consistent with the idea that, as a group, older adults appear to have a more favorable outcome than expected following the death of a spouse. A recent longitudinal study comparing over 120 older bereaved spouses with age-matched nonbereaved controls found that depressive symptoms peak within the first few months after a death but decline significantly within a year, and after 2 or 3 years of follow-up, although bereaved groups still report significantly higher levels of depressive symptoms than age-matched controls, these differences are very small. Even during the first few months following bereavement, few older adult widows (between 5 and 17 percent) score in the severely depressed range on depression rating scales. Interestingly, although elderly bereaved spouses report that loneliness and performing activities of daily living are factors that complicate adjustment, many older spouses also report positive feelings following a death, such as relief, in contrast to younger samples. There are nevertheless important exceptions; studies suggest that elderly survivors of spouses who committed suicide, who have low social support and who report greater psychopathology at the time of death, are all at a

higher risk of poorer long-term adjustment. In addition, elderly women who are highly dependent on their spouses for assistance appear to be at a higher risk for adverse outcomes and may require more professional intervention.

Recent data reveal an interesting relationship between bereavement and an excess risk of mortality in older adults. Most studies have revealed that actually younger bereaved spouses face a greater risk of mortality than older recent widows. In addition, elderly white widows appear to be at a greater increased risk of mortality than elderly black widows, who may not face an increased risk at all.

Resilience of the Aging Self

The above review of coping and adaptation in aging suggests that older adults appear to cope well with adversities and challenges, despite in many cases the more chronic and accumulating nature of the stressors that they face. The concept of *resilience* has been offered as an explanatory factor for this phenomenon and is defined as the process of both maintaining adaptive behavior and positive emotions when presented with stress as well as recovery from adversity. Resilience may involve a process of *selective optimization*, or an accommodative process of rescaling goals to match one's limitations and the demands of the environment, and investing the fewer resources available in personally salient activities that maintain self-esteem. For example, a recent study found that the psychological impact of physical disability was less for those elders over 80 than that in those younger than 80, suggesting that a process of scaling down expectations with increasing age may occur and may be adaptive.

Maintaining a strong sense of self-worth in the face of accumulating losses may be an important part of resilience in aging. For example, a recent study found that self-esteem actually declined very little over time in a sample of elders with low socioeconomic status and low levels of education, despite experiences of personally meaningful losses. Rather, these older adults appeared to engage in cognitive strategies to protect their sense of self and well-being. In another study, even individuals over the age of 100 who experienced multiple losses in many domains report similar levels of happiness to middle-aged and young-old individuals, and attitudes such as optimism and self-efficacy are more predictive of happiness than objective resources available.

SOCIAL FUNCTIONING IN LATE LIFE

Research over several decades has consistently validated the importance of social support in "buffering" the effects of stress: Given equivalent stressors, those individuals with strong social ties appear to have better outcomes than those who are isolated or perceive that support is

unavailable or unhelpful from those in their network. Yet, the elderly appear to have reduced social networks, the frequency of their social contacts is reduced as compared to those of younger adults, and this may be especially true for the oldest-old, whose social networks have been shown to be one-half the size of those of the young-old. This finding was previously thought to be evidence for the *disengagement* hypothesis of aging, that is, that aging involves a process of withdrawal from society and others and the development of a more inward focus. However, for many older adults, the reduction in social contact is not voluntary and may be more related to disability and reduced mobility than preference. The disengagement hypothesis has been recently reframed as *socioemotional selectivity*, based on the finding that although overall social contacts do decrease in late life, it is only contact with acquaintances and other peripheral individuals that decreases; social contacts with very close friends and family remain stable, and older adults' satisfaction with these close relationships remains high. It may be that as individuals age and resources become fewer, those resources are adaptively invested in maintaining close relationships as opposed to maintaining acquaintanceships. For older adults with a *limited future time perspective* or a greater awareness of mortality, there is even more motivation to maintain emotionally meaningful relationships. However, this phenomenon may be at least partially dependent on the external circumstances of older people's lives. For example, in a recent study of over 500 elderly individuals over the age of 70, the size of social networks and the feelings of connectedness to that social milieu were related to the presence and availability of the nuclear family. Older adults with nuclear family members tended to have social networks that were both larger as well as more emotionally close than older adults without a nuclear family. However, those without a nuclear family reported greater feelings of closeness to friends than those with a nuclear family, suggesting that older adults can adapt to the absence of family by creating satisfying relationships with others.

The social structure of oldest-old has recently been studied, and data suggest that this group may be particularly vulnerable to social isolation due to multiple losses. In a recent study, most individuals over the age of 85 had at least one surviving relative, but approximately one-quarter had little or no contact with that relative. Most social contact was derived from children, and for those individuals without children, social contact and support were not provided from more distant relatives, except within the African-American community, where other relatives and friends were available to provide help and support. For the oldest-old, then, the family, especially adult children, appears to play a crucial role in providing social support.

COGNITIVE AGING

The study of cognition and aging has been an extraordinarily productive area of research over the past several decades. Nevertheless, the existing literature is fraught with difficulties in interpretation, with confusion or disagreement about the inclusion or exclusion of subjects with medical problems and multiple medications and the inadvertent inclusion of subjects with early dementia, which tends to overestimate the effects of age by lowering the overall mean of elderly groups. Furthermore, until magnetic resonance imaging (MRI) scanning was widely used, studies unknowingly included subjects with subcortical ischemic disease, confounding the effects of age with disease. In addition, design issues are important; cross-sectional studies, in which individuals of different age groups are compared at one point in time, almost always show broader and sharper age effects than longitudinal studies due to possible confounding of age effects with cohort effects.

Summary of Age Effects on the Major Cognitive Domains

The broad domain of cognition has been divided into subdomains that include attentional processes, language functions, visuoperceptual abilities, learning and memory, processing speed, and the executive functions.

Attentional Processes. Attentional processes are a complex group of behaviors by which individuals select information for cognitive processing while ignoring competing information, concentrate on information over a sustained period of time, divide attentional focus between two or more aspects of a task at a time, and shift attentional focus from one set of stimuli to another and back again. Basic attention span is the amount of information that can be apprehended at one time and demonstrates little or no relationship to age. Similarly, sustained concentration or vigilance is not affected by the aging process. However, complex attentional tasks such as divided attention that place more demand on cognitive capabilities show significant declines with advancing age.

Language Functions and Verbal Abilities. Language functions remain remarkably stable with age, with few exceptions. Vocabulary and expressive word knowledge may even increase over the life span. Verbal expression and verbal reasoning skills may demonstrate subtle changes with age, as some studies have shown that spontaneous language may become less precise and more repetitious over time, although most of these conclusions were drawn from cross-sectional studies and may be confounded with cohort effects. Finally, the ability to spontaneously retrieve names of objects from the lexicon appears to show reliable but small decreases with age, although these effects are heavily influenced by educational

achievement.

Visuoperceptual and Visuoconstructive Abilities. Visuoperceptual abilities are those functions that involve the recognition and appraisal of nonverbal relationships, specifically those involving shapes, patterns, and objects with spatial relationships to one another. Visuoconstructive abilities are those involving the reproduction of nonverbal material, such as the direct copy of simple and complex designs. Visuoperceptual abilities demonstrate relatively consistent age effects across studies, with older adults generally performing as well as younger adults on tests of simple object and pattern recognition but performing more poorly than younger adults on measures of visual closure and spatial orientation. Similarly, in terms of visuoconstructive functions, older adults perform about as well as younger adults when copying simple two-dimensional designs but more poorly than younger adults when copying three-dimensional complex designs and when planning and strategy are necessary for good performance. It should be noted that age effects are the largest on visuoconstructive tasks that are timed and removing the speed component of the task often considerably reduces but does not remove the age difference.

Learning and Memory. Effects of age on learning and memory are among the most well-documented and well-studied effects in the neuropsychological literature. However, "memory" is an extremely complex set of functions and systems, and the aging process affects some components of memory to a greater degree than others. Sensory memory, or the extremely brief sensory record of a stimulus, primary memory, exemplified by the ability to recall a string of digits long enough to repeat them, and tertiary memory, or memories from childhood, show no significant age effects or mild and subtle effects that can be detected in the laboratory. Research studies have consistently shown that secondary memory, exemplified by the ability to learn information and recall it after a time delay is most the strongly affected by age. Nevertheless, the age-related effects on secondary memory are relatively small: Studies suggest that performance declines at a rate of 10 percent per decade of life beginning in midlife. However, once older adults learn information, it is retained well; studies suggest that older adults retain information at rates only slightly below that of younger adults. In addition, available data suggest that spontaneous recall demonstrates considerably stronger age effects than recognition memory, which is the ability to recognize information that has been previously presented.

Processing Speed. Speed of information processing and pure motor speed demonstrate the most reliable age differences and decrements with

age, whether measured cross-sectionally or longitudinally. Indeed, reduced processing speed in aging may be a crucial component of many age-related effects on cognition, such that in a recent study age effects on a number of neuropsychological tests were considerably diminished after processing speed was partialled out, even on tests where speed of performance was not measured. In addition, recent studies suggest that speed of processing is an important mediator of memory performance, such that some authors have proposed a speed of processing model to account for age-related decrements in memory and other cognitive processes.

Executive Functions. The term "executive functions" refers to superordinate cognitive skills involved in abstract reasoning, problem-solving, planning, response inhibition, the ability to shift cognitive set, and a working memory construct that is defined as the ability to simultaneously hold and process information. Early studies suggested that the executive functions decline with age. However, recent studies that use carefully selected samples and exclude subjects with histories of chronic medical conditions known to affect central nervous system (CNS) functioning or documented evidence of subcortical white or gray matter lesions on MRI scanning reveal very few age differences on executive tasks. Because of the relationship of performance on tests of executive functioning to the presence of medical conditions, especially diseases of the cardiovascular system, these studies suggest that some of these age effects may be preventable.

Models of Cognitive Aging

Taken together, studies suggest that there is a profile of cognitive changes with age that includes significantly reduced information processing speed and pure motor speed and mild decrements in spontaneous recall, executive skills, complex attentional processes, and complex visuoperceptual and visuoconstructive abilities. It should be emphasized that although age-related changes in cognition are statistically significant, most elderly persons still perform within normal limits on cognitive testing. The extent to which these observed changes are related to the process of aging itself, the effects of particular medical conditions on the brain and CNS functioning, and cohort differences in education and experience are not well understood at present, although a recent longitudinal study with a 10-year follow-up interval revealed that significant decline in cognitive functions (excluding psychomotor speed) was apparent only in individuals suffering from poor health. A number of models have been proposed to account for these changes, and these models have included both psychosocial and biological interpretations. First, it is unlikely that all age-related effects on cognition are secondary to cohort differences, because cohort-sequential studies, in which individuals from different age cohorts

are followed prospectively, reveal clear longitudinal decline across age cohorts. Unfortunately, studies specifically addressing this important question are few and have only investigated intellectual abilities and not functions such as memory or other neuropsychological variables. It is likely that cohort effects exert the strongest influence on those abilities most dependent on formal education, but this hypothesis has not been formally tested. It is also unlikely that simple disuse accounts for significant aspects of age-related cognitive decline.

The Cattell–Horn model of intellectual changes over the life span proposes that age affects different cognitive processes differently. The model proposes that general intelligence can be subdivided into *crystallized intelligence*, referring to abilities and information learned through exposure to education and life experience, and *fluid intelligence*, referring to abilities involved in novel problem-solving. The model proposes that crystallized intelligence remains stable or increases over the course of the life span and may even increase into old age, due to continuing exposure to information. In contrast, fluid intelligence, thought to be dependent on the integrity of brain function, was hypothesized to decline with advanced age. Available data have generally supported this model, with studies showing that there is little evidence of decline on measures of such crystallized abilities such as word knowledge or general fund of information, whereas measures of fluid abilities such as novel problem-solving, as discussed above, generally reveal some decline. However, the model, as originally proposed, is limited in that it only discusses changes in intellectual functions, whereas other cognitive functions, such as learning and memory, are not included. Furthermore, the same limitations of studies of executive functions, apply to the fluid intelligence concept.

The processing speed hypothesis hypothesizes that many of the observed age-related changes in cognition are in fact secondary to reductions in speed of processing. The model proposes that the speed at which information can be processed limits both the quantity and the quality of cognitive operations. Slowing affects not only how much information is processed within a given period but may also dilute and weaken the associations between neural networks, leading to more impoverished cognitive processes. The model has received considerable empirical support, with most studies finding that a general speed factor is an important mediator of other cognitive processes, such as memory.

Neuropsychological models of age-related decline have attempted to integrate findings from cognitive testing with findings from brain imaging and neuropathological studies to suggest possible etiologies of cognitive changes. The most promising of these is a model variously termed the frontal deficit hypothesis or the executive function hypothesis, based on the finding that the pattern of cognitive performance in normal aging, although

less pronounced, is similar to that seen in many dementing processes affecting primarily subcortical–frontal circuitry. In addition, neuropathological data suggest that frontal–subcortical structures are the most vulnerable to aging, with many studies revealing that neuronal loss and degeneration of dendritic arborization are most prominent in the frontal cortex and in some subcortical areas. Further, there are definite changes in the vasculature of the brain with age that result in ischemic white and subcortical gray matter disease that predominantly affects cognitive processes mediated by frontal–subcortical circuits. Finally, a recent study found that age-related changes in the dopamine system, particularly declines in D_2 receptor availability in some basal ganglia structures, were related to age-related declines in performance on both motor and executive systems tasks. However, most studies have found that the executive functions appear to decline at the same rate as other cognitive functions and are not among the first age-related cognitive changes to be observed. This casts some doubt on their central role as mediators of age-related cognitive decline.

Finally, some authors have proposed a "common cause" model, suggesting that decrements in sensory functions, including visual and auditory acuity, may explain age-related changes in cognitive processes. As with the processing speed hypothesis, empirical support for this idea is generally strong, with most studies finding that sensory function is an important predictor of age differences in a number of cognitive processes. Nevertheless, it is clear that an as-yet-unnamed variable or group of variables remains unexplained in accounting for variance in cognitive aging, because most studies find that age-related variance remains even after both speed and sensory abilities are partialled out.

It is important to note that the line between a "normal" process of cognitive aging and a pathological process is not well drawn at present. For example, there is considerable disagreement in the field about the inclusion or exclusion of elderly subjects with medical conditions known to affect brain function or individuals with white matter lesions in studies of normal aging and cognition. Because these conditions affect such a high proportion of functional elderly persons, their exclusion may lead to erroneously high estimates of normal expectations of performance and the subsequent classification of many individuals as impaired. In addition, many older individuals without known cerebrovascular risk factors or other medical conditions have periventricular white matter lesions on MRI scanning, confusing further the line between disease processes and normal aging. At present, until this demarcation is better defined, the study of "typical" aging, which attempts to characterize the majority of elderly people, as well as the study of "optimal" aging, which attempts to characterize the ideal outcome of the aging process as well as characterize changes that are a function of aging itself and not epiphenomena of other conditions, are both

important pursuits, as long as the research questions are appropriate to the sample being used.

Finally, it may be that older adults' performance on cognitive testing may be affected by their own ageist stereotypes. A recent study found that when, prior to cognitive testing, older adults were "primed" with positive images of aging, their performance improved on measures of memory, as well as measures of memory self-efficacy, in contrast to older adults who were primed with negative images of aging.

SUCCESSFUL AGING

In recent years, there has been a paradigm shift in gerontology, from a view of aging as inevitable deterioration to one that includes the potential for development and growth. In this section, a number of studies were reviewed that, taken together, support this notion that the process of aging can be managed well by the majority of older adults. In this vein, the construct of *successful aging* was created. In the process, however, there has been considerable debate about how to define successful aging. It has become clear that if one defines successful aging as the absence of deficits in physical, cognitive, and social functioning, then few older adults would be so classified. A recent study examined both objective and subjective indicators of successful aging and found that only 10 percent of their sample could be classified as "successfully aged" based on objective criteria that included quantitative assessments of physical, cognitive, and social functioning and measures of well-being. However, when successful aging was defined as a subjective sense of having adapted well to the demands of life or was self-rated, many more individuals were characterized as having aged successfully. In a recent review of data-based studies, the proportion of successfully aged individuals ranged from 0.4 to 95 percent, thus underscoring this ongoing debate. In general, older adults rate social engagement and positive thinking as more important than physical functioning in their conceptualizations of successful aging. It is likely that any definition should include both objective and subjective definitions of success.

Studies have found that high education, absence of alcohol abuse and cigarette smoking, absence of depression, maintaining appropriate weight, regular exercise, and social support are all significant predictors of successful aging. Social class, however, does not appear to be an important predictor. Intrinsic psychological factors such as resilience, spirituality, and optimism have also been associated with successful aging. As more is learned about successful aging, both in terms of its psychosocial correlates and biological correlates, it may be that this knowledge can be translated into useful interventions that can improve the overall number of elderly reaching old age successfully.

IMPLICATIONS FOR CLINICAL PRACTICE

One of the major points of this section is that, with the exception of cognitive processes, age in and of itself is rarely an important predictor of outcome. Studies of mental health over the life span, coping and adjustment, and personality often fail to reveal any significant age effects, when other factors are controlled. In addition, research has revealed an often surprising degree of resilience in older adults. Therefore, the clinician working with older adults should not expect that age itself is a risk factor for poor outcomes. Unfortunately, this notion has probably been partially responsible for the underidentification of depression in many elderly cohorts; many clinicians used to believe that depression was a "normal" response to the aging process. Clearly, the life themes that older adults struggle with are different from those of younger adults, and the clinician should be prepared to deal with issues of loss, increasing dependency, and the struggle to come to terms with the past. In addition, the concept of resilience may be useful in psychotherapeutic work with older adults; clinicians can encourage patients to selectively invest resources in activities that they find especially gratifying.

The important exception to some of these conclusions may be individuals in the oldest-old category, or those over the age of 85. Research has shown that the oldest-old may be at a higher risk of social isolation than younger age groups, especially those without living children. Clinicians should be alert to the unique life situations faced by individuals in this age group, who may be especially vulnerable not only to isolation but also to increasing frailty and the cumulative deleterious effects of chronic medical conditions. Interventions that both provide social support as well as encourage alternative sources of social support such as senior centers may be especially helpful for individuals in this age group without living family.

In terms of cognitive processes, there are clear decrements in some but not all cognitive processes over the life span. However, with the exception of processing speed, these changes are generally of small magnitude. The result is that the typical changes in cognition should not significantly impact the clinical situation, either the psychotherapeutic encounter or the psychiatric consultation. Some older adults may require more repetition and explanation of material, especially material that is novel and highly abstract, but this can easily be managed with modifications of psychotherapeutic approaches. The pattern of cognitive decline seen in normal aging discussed in this section should alert the clinician to changes that deviate from this pattern that might be cause for concern, especially such things as difficulties with recalling information even after cueing, problems with basic attentional processes, problems with constructing basic two-dimensional designs, or language impairment. At the same time, clinicians need to be alert to the effects of negative age-related stereotypes

and its effects on both performance and self-concept.

Finally, it is worth noting that because research findings are based on the central or average tendencies of groups of individuals, the conclusions drawn often mask underlying heterogeneity. This may be especially true with the elderly, as research has shown that the variability of almost every construct studied, such as personality, cognition, health status, and income, increases over the life span. It is likely that this increased heterogeneity is a complex interactive process of increasingly individualized expressions of genetic predispositions in response to accumulating demands from the environment over the life span. Thus, the clinician working with older adults may have less precise guidelines or standards for normative expectations for elderly patients, because the elderly can present with a much wider range of behavior, personality, and life experience than is found in younger adults.

REFERENCES

Aldwin CM, Sutton KJ, Chiara G, Spiro A 3rd. Age differences in stress, coping, and appraisal: Findings from the Normative Aging Study. *J Gerontol B Psychol Sci Soc Sci*. 1996;50(4):P179–188.

Baltes PB, Lindenberger U. Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychol Aging*. 1997;12(1):12–21.

Berg AI, Johansson B. Personality change in the oldest-old: Is it a matter of compromised health and functioning? *J Pers*. 2014;82(1):25–31. doi: 10.1111/jopy.12030.

Carr D, House JS, Kessler RC, Nesse RM, Sonnega J, Wortman C. Marital quality and psychological adjustment to widowhood among older adults: a longitudinal analysis. *J Gerontol B Psychol Sci Soc Sci*. 2000;55(4):S197–207.

Collins AL, Smyer MA. The resilience of self-esteem in late adulthood. *J Aging Health*. 2005;17(4):471–489.

Depp CA, Jeste DV. Definitions and predictors of successful aging: a comprehensive review of larger quantitative studies. *Am J Geriatr Psychiatry*. 2006;14(1):6–20.

Erikson E. *Childhood and Society*. New York: Norton; 1950.

Ferraro K, Wilmoth J. *Gerontology: Perspectives and Issues*. 4th ed. New York: Springer Publishing Company; 2013.

*Folkman S. Coping across the life span: Theoretical issues. In: Cummings EM, Greene AL, Karraker KH, eds. *Life Span Developmental Psychology: Perspectives on Stress and Coping*. Hillsdale, NJ: Lawrence Erlbaum & Associates; 1991:3.

Gross JJ, Carstensen LL, Pasupathi M, Tsai J, Skorpen CG, Hsu AY. Emotion and aging: Experience, expression, and control. *Psychol Aging*. 1997;12(4):590–599.

Hofer SM, Sliwinski MJ. Understanding ageing: An evaluation of research designs for assessing the interdependence of ageing-related changes. *Gerontology*. 2001;47(6):341–352.

Irving P. *Upside of Aging: How Long Life is Changing the World of Health, Work, Innovation, Policy, and Purpose*. Hoboken, NJ: Wiley Publishing; 2014.

Jopp D, Rott C. Adaptation in very old age: exploring the role of resources, beliefs, and attitudes for centenarians' happiness. *Psychol Aging*. 2006;21(2):266–280.

Kubzansky LD, Berkman LF, Seeman TE. Social conditions and distress in elderly persons: findings from the MacArthur Studies of Successful Aging. *J Gerontol B Psychol Sci Soc Sci*. 2000;55(4):P238–246.

Kim J, Lee HY, Christensen MC, Merighi JR. Technology access and use, and their associations with social engagement among older adults: do women and men differ [published online ahead of print October 8, 2016]? *J Gerontol B Psychol Soc Sci*.

Kok AA, Aartsen MJ, Deej DJ, Huisman M. Socioeconomic inequalities in a 16-year longitudinal measurement of successful ageing. *J Epidemiol Community Health*. 2016;70(11):1106–1113.

Kulik L, Shilo-Levin S, Liberman G. The contribution of role satisfaction, meaning in life and work-home conflict to explaining well-being among working grandparents. *Anxiety Stress Coping*. 2016:1–14.

Lang FR. Regulation of social relationships in later adulthood. *J Gerontol B Psychol Sci Soc Sci*. 2001;56(6):P321–326.

Levy BR. Mind matters: cognitive and physical effects of aging self-stereotypes. *J Gerontol B Psychol Sci Soc Sci*. 2003;58(4):P203–211.

*Lund DA. Conclusions about bereavement in later life and implications for interventions and future research. In: Lund DA, ed. *Older Bereaved Spouses: Research with Practical Applications*. New York: Hemisphere Publishing Corporation; 1989:3.

Martin P, Long MV, Poon LW. Age changes and differences in personality traits and states of the old and very old. *J Gerontol B Psychol Sci Soc Sci*. 2002;57(2):P144–152.

MacLeod MD, Saunders J. Episodic memory and age-related deficits in inhibitory effectiveness. *Exp Aging Res*. 2017;43(1):34–54.

Mehta M, Whyte E, Lenze E, et al. Depressive symptoms in late life: Associations with apathy, resilience and disability vary between young-old and old-old. *Int J Geriatr Psychiatry*. 2008;23(3):238–243.

Mejia ST, Ryan LH, Gonzalez R, Smith J. Successful aging as the intersection of individual resources, age, environment, and experiences of well-being in daily activities. *J Gerontol B Psychol Sci Soc Sci*. 2017;72(2).

Montross LP, Depp C, Daly J, et al. Correlates of self-rated successful aging among community-dwelling older adults. *Am J Geriatr Psychiatry*. 2006;14(1):43–51.

Mroczek DK, Spiro A 3rd. Modeling intraindividual change in personality traits: findings from the normative aging study. *J Gerontol B Psychol Sci Soc Sci*. 2003;58(3):P153–165.

Narrow WE, Rae DS, Robins LN, Regier DA. Revised prevalence estimates of mental disorders in the United States: using a clinical significance criterion to reconcile 2 surveys' estimates. *Arch Gen Psychiatry*. 2002;59(2):115–123.

Ong AD, Bergeman CS, Bisconti TL, Wallace KA. Psychological resilience, positive emotions, and successful adaptation to stress in later life. *J Pers Soc Psychol*. 2006;91(4):730–749.

Schaie KW. A lifespan developmental perspective of psychological aging. In: Laidlaw K, Knight B, eds. *Handbook of Emotional Disorders in Later Life: Assessment and Treatment*. New York: Oxford University Press; 2008:1.

*Staudinger UM, Marsiske M, Baltes PB. Resilience and reserve capacity in later adulthood:

Potential and limits of development across the life span. In: Cicchetti D, Cohen DJ, eds. *Developmental Psychopathology II: Risk, Disorder, and Adaptation*. New York: John Wiley & Sons; 1995:801.

Thumala Dockendorff DC. Healthy ways of coping with losses related to the aging process. *Educ Gerontol*. 2014;40(5):363–384.

Vliet EC, Manly J, Tang MX, Marder K, Bell K, Stern Y. The neuropsychological profiles of mild Alzheimer's disease and questionable dementia as compared to age-related cognitive decline. *J Int Neuropsychol Soc*. 2003;9(5):720–732.

von Faber M, Bootsma-van der Wiel A, van Exel E, et al. Successful aging in the oldest old: Who can be characterized as successfully aged? *Arch Intern Med*. 2001;161(22):2694–2700.

Wilson RS, Krueger KR, Gu L, Bienias JL, Mendes de Leon CF, Evans DA. Neuroticism, extraversion, and mortality in a defined population of older persons. *Psychosom Med*. 2005;67(6):841–845.

Wilson RS, Mendes de Leon CF, Bienias JL, Evans DA, Bennett DA. Personality and mortality in old age. *J Gerontol B Psychol Sci Soc Sci*. 2004;59(3):P110–116.

57.2e Neuropsychological Evaluation

BARTON W. PALMER, PH.D., AND GAURI N. SAVLA, PH.D.

INTRODUCTION

As medical technology advances rapidly, people are living longer, and are active well into "old age," typically considered to be 65 years of age and above. Nevertheless, this particular demographic is also at greatest risk for cognitive deficits related to cerebrovascular, systemic, psychiatric, and neurodegenerative illness. Comprehensive clinical care of older adults, therefore, should almost invariably include some consideration of the presence and influence of neuropsychological deficits. Cognitive decline is a core component of neurodegenerative conditions such as Alzheimer disease, Dementia with Lewy Bodies (DLB), Parkinson disease, as well as cardiovascular conditions such as strokes. Chronic metabolic illnesses (e.g., diabetes) and long-standing autoimmune diseases (e.g., multiple sclerosis) may also be associated with cognitive impairments.

Psychiatric disorders with typically earlier onset that often extend into later life, such as primary psychotic disorders or severe mood disorders, are also often (but not always) associated with mild to moderate deficits in neuropsychological functioning; these deficits tend to remain more or less stable over the course of the illness (i.e., there are rarely sharp declines in later life without the co-occurrence of a new-onset neurodegenerative illness). There is some emerging evidence that cognition in bipolar disorder may continue to deteriorate over time, with repeated affective episodes. This finding, if confirmed, could have particular relevance to older adults with bipolar disorder who have lived with the condition for decades. An important point to consider, however, is that cognitive deficits in

schizophrenia and bipolar disorder have a much stronger predictive relationship to level of adaptive functioning than symptoms of the primary psychopathology.

The bottom line is that an evaluation of neuropsychological functioning can be a critical component of effective diagnosis, and long-term treatment and care planning in geriatric psychiatry settings.

WHAT IS A NEUROPSYCHOLOGIST?

Neuropsychology as a field of scientific inquiry is not a profession or degree unto itself but as a clinical discipline it falls within the broader field of clinical psychology. The term "clinical neuropsychologist" refers to a licensed clinical psychologist, who, in addition to the broader training required for a doctorate and licensure in clinical psychology, has also completed additional training in structural and functional neuroanatomy, neuropathology, and neurobehavioral syndromes. A neuropsychologist, therefore, is an expert in scientifically grounded brain–behavior relationships.

WHAT IS A NEUROPSYCHOLOGICAL EVALUATION?

Standardized neuropsychological tests that measure cognitive abilities and deficits form the core of neuropsychological evaluation. However, neuropsychological evaluation involves much more than administrating and scoring tests. Valid and clinically useful interpretation of test performance requires the neuropsychologist to consider the examinee's performance within the broader context of his or her presenting complaint as well as prior medical history, psychosocial history, and functioning. Thus, in addition to the standardized assessment of neurocognitive abilities, comprehensive clinical neuropsychological evaluations generally include a thorough review of available medical records as well as a clinical interview of the patient and family or other collateral information sources. Personality and emotional functioning are also generally assessed via clinical interview, behavioral observations, clinical rating scales, and/or standardized psychological tests. Neuroimaging is also increasingly used to complement the results of these evaluations.

HISTORICAL DEVELOPMENT

The emergence of neuropsychology as a clinical discipline can be traced back to the early to mid-20th century, particularly the 1930s and 1940s when several pioneers of the field such as Ward Halstead and Arthur Benton began their studies of brain–behavior relationships using psychometric methods. Many of the tests developed or adapted by Halstead (and his battery as subsequently expanded by Ralph Reitan), as well as those by Benton remain core components of contemporary

neuropsychological test batteries. There was also early work on psychometric testing of intelligence and other cognitive functions, as exemplified by the work of Alfred Binet, as well as the Army aptitude testing programs that emerged during World War I. (These are, by necessity of space, only selected examples of a richer and more complex history of psychometric testing.) Another key contributor to the development of clinical neuropsychology was David Wechsler. In 1939, Wechsler published the first version of his Bellevue Intelligence Examination, followed in 1945 by its companion, the Wechsler Memory Scale. Many of the subtests in the Wechsler batteries were adapted from earlier sources. As collective batteries and with appropriate revisions and updates in the ensuing decades, the Wechsler scales have remained among the most widely used measures of cognitive functioning in patients as well as healthy individuals.

The development of modern clinical neuropsychology was fostered by the profound advances in regard to theory and empirical knowledge of brain systems and processes, brain–behavior relationships, and the effects of various neuropathological conditions on everyday functioning. The nature and quality of standardized normative data for valid interpretation of neuropsychological test results have also improved significantly. These advances have occurred in close collaboration with the growth in related fields, such as imaging, behavioral neurology, and cognitive neuroscience as well as in psychometric theory and practice.

COMMON USES OF NEUROPSYCHOLOGICAL EVALUATIONS

Neuropsychological assessment can be invaluable in geriatric mental health care settings; it can aide in the distinction between normal versus pathological cognitive and functional changes, differential diagnosis among neuropsychiatric and neuropathological conditions, and determination of long-term functional needs for treatment and care planning.

Despite the remarkable advances in clinical brain neuroimaging over the past few decades, their clinical relevance (particularly in regard to diagnostic specificity) remains equivocal. For instance, one of the most common concerns with older patients is distinguishing normal age-related changes from early Alzheimer disease or other neurodegenerative conditions. A second frequent presenting concern is determining whether an individual's reports of memory or other cognitive deficits are due to depression or primary neurocognitive dysfunction. With present technology, such distinctions are difficult, if not impossible, to make based on brain imaging results alone but can be readily made by considering the pattern of neuropsychological test performance together with the individual's current and prior medical and psychosocial functioning.

As an example, consider a patient presenting with a report of "poor memory." The term "memory" has a variety of meanings in colloquial use,

such as remembering to perform a task ("prospective memory"), finding misplaced objects (which may reflect attention or organizational skills rather than memory per se), and remembering recent events or newly acquired information ("episodic memory"). Even when the concern is narrowed down to episodic memory, it is important to distinguish between a range of underlying memory-related tasks or processes, including difficulties with initial learning or encoding information or difficulties efficiently retrieving information, and actual loss of the memory trace. Results from standardized memory tests are invaluable in making such distinctions. For instance, if a patient shows poor performance on initial learning or recall trials of a test, but after a delay, he or she is still able to recall most of whatever amount of information he or she learned, then his or her deficit is likely one of learning or encoding rather than forgetting. Comparison of cued recall, performance on a recognition format, or both to free recall permits identification of patients with inefficient retrieval rather than actual forgetting. These types of distinctions are key in the differential diagnostic process. Rapid forgetting, that is, the inability to retain recently learned information, is relatively uncommon with normal aging and even in patients with depression but is quite common among patients with cortical dementias. One would not diagnose a patient based on his or her performance on an isolated test score (most healthy adults will manifest one or more impaired range scores on some measure when given a large test battery), but the presence of rapid forgetting would be an important piece of evidence more consistent with the presence of cortical dementia than with normal aging or depression.

Even when a patient's diagnosis is known, neuropsychological assessment can be valuable in the documentation of the trajectory and rate of cognitive change. For instance, there is considerable variability between patients with neurodegenerative conditions in terms of the rate of cognitive decline. Repeated evaluation with a targeted battery can be useful in determining how fast a particular person is likely to reach a point requiring a more structured or intensive level of care. Conversely, for a patient who has experienced a stroke or other acute brain trauma, repeated neuropsychological testing can be useful in documenting the rate, trajectory, and degree of recovery of cognitive functioning.

Neuropsychological evaluation in geriatric mental health care can also be useful in treatment plans involving the remediation of deficits in everyday functioning. For example, it may be known that a particular patient's "forgetfulness" is causing difficulties in everyday functioning (such as missed appointments, unpaid bills, etc.). In such instances, a key goal of neuropsychological evaluation is to identify areas of spared cognitive functioning, in addition to documenting the type and degree of any cognitive deficits. With such information, the specific cognitive processes underlying any apparent deficits in independent functioning can

be identified, and information about spared abilities can be used to arrange the environment in a way that maximally draws upon the individual's strengths and reduces the potential deleterious functional effects of impaired abilities.

ASSESSMENT DOMAINS AND COMMONLY USED MEASURES

The length and extent of cognitive testing performed by a neuropsychologist frequently varies depending upon the nature of the presenting problem and the general context or purpose of the evaluation. Some cognitive screening procedures require 30 minutes or less. However, a comprehensive neuropsychological evaluation generally requires several hours to complete. The latter type of testing is particularly valuable in identifying cognitive strengths and weaknesses across a range of more specific neurocognitive abilities. Some of the common constructs evaluated as well as examples of tests frequently used to assess these constructs are described below.

Cognitive Screening

The majority of brief cognitive screening measures were designed specifically to stage severity of Alzheimer disease or other dementias so may be insensitive to other functionally relevant forms of cognitive impairment and tend to be less helpful in identifying areas of clear cognitive strengths.

One of the most widely used dementia screening measure is the Folstein Mini-Mental State Examination (MMSE). The MMSE has a potential range of 0 to 30, with scores below 24 generally considered "impaired." Advantages of the MMSE include ease of administration and scoring; it typically takes less than 10 minutes to complete, can be reliably administered and scored by trained clinic staff, and can be tolerated even by very low functioning patients. Also, it is so widely used that most clinicians have a tangible sense of how scores at a given range are likely to correspond with actual functioning. MMSE scores can thus serve as a useful means of quickly communicating the severity of a patient's overall deficits. Disadvantages of the MMSE include its sensitivity to education (highly educated patients with mild Alzheimer disease may perform very well on the MMSE, whereas healthy individuals with low education may earn scores in the "impaired" range) and its relative insensitivity to mild cognitive deficits and deficits associated with noncortical dementias. One caveat, although not always appreciated by users, is that the MMSE is a copyrighted test that requires purchase for legal use.

Another brief screening instrument for cognitive performance is the Montreal Cognitive Assessment (MoCA). Similar to the MMSE, the MoCA is brief (less than 10 minutes), easy to administer and score, and yields

total scores ranging from 0 to 30, with lower scores representing worse cognitive impairment. It was originally designed to screen for Mild Cognitive Impairment (MCI), rather than cortical dementia per se, and thus may be more sensitive to subtle impairments in the nondemented range of functioning. Another advantage of the MoCA includes availability in multiple languages in addition to English. The copyright holder permits use free of charge and without written permission for noncommercial clinical and educational use, and with written permission in many other applications (see MoCA website for current details: http://www.mocatest.org/permission.asp).

The Mattis Dementia Rating Scale-Second Edition (DRS-2) is another widely used scale of dementia screening; it generates a total score (maximum possible points = 144), indicating overall cognitive functioning, as well as five subscale scores, that is, Attention, Initiation/Perseveration, Construction, Conceptualization, and Memory, and takes from 15 to 20 minutes (for a cognitively intact individual) to approximately 25 to 30 minutes (for an impaired individual) to administer. Like the MMSE, it is prone to age and education effects, but available age- and education-corrected norms facilitate interpretation of DRS-2 scores. The DRS-2 is sensitive to mild dementia, although the ideal "cut-off" score seems to vary slightly among studies; one cross-validated study applied a cut-off of 129 of 144 and yielded a 98 percent sensitivity and 97 percent specificity while distinguishing patients with Alzheimer disease from healthy comparison subjects. Overall DRS-2 scores as well as subscale scores have been shown to be good predictors of functional outcome (basic and instrumental activities of daily living) among patients with Alzheimer disease. The cognitive subscale of the Alzheimer's Disease Assessment Scale (ADAS-Cog) is another widely used screen for cognition among those with Alzheimer disease.

An intermediate option between a brief screening instrument and a lengthy comprehensive neuropsychological battery is to use an instrument such as the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). The RBANS takes approximately 30 minutes to administer and assesses a number of domains of interest. Computerized cognitive screens, such as the MicroCog, were developed to reduce time and administration burden and are also available for use.

A screening measure is useful for a quick and cursory evaluation of cognitive functioning, but is limited in the extent to which it can assess the breadth and depth of neuropsychological abilities. A comprehensive neuropsychological test battery is the best way to measure more nuanced aspects of cognition across multiple domains of functioning.

Crystallized Knowledge

Crystallized knowledge refers to information and facts typically learned in

formal school settings. Tests that measure crystallized knowledge include the Wechsler Adult Intelligence Scale-4th Edition (WAIS-IV) Vocabulary subtest, in which the examinee is asked to define a common word, the WAIS-IV Information subtest, in which the examinee is asked to answer "common knowledge" questions, and the Wide Range Achievement Test-4th Edition Reading subtest that measures one's ability to recognize or read isolated words. Crystallized knowledge tends to be resistant to most forms of neurocognitive injury and tends to be relatively stable (or even improve) over the life span; word reading and vocabulary scores are therefore, often used, together with educational history, to estimate levels of premorbid cognitive functioning.

Receptive and Expressive Language

Deficits in language functions may reflect specific or localized lesions (such as a stroke in the left frontal lobe) or diffuse neuropathology such as that occurs with Alzheimer disease. Common receptive language measures include such tasks as reading comprehension or comprehension of increasingly complex aurally presented instructions, such as measured by the Token Test or the Complex Ideational subtest from the Boston Diagnostic Aphasia Examination. The Boston Naming Test may be the most widely used measure of expressive language functions in assessments of adults and is particularly useful with patients reporting word finding difficulties. The task involves presenting the examinee with line drawings of objects one at a time and asking him or her to identify each one in succession. If the examinee cannot identify the item within 20 seconds, then he or she is given a cue to see if that aids responding. Confrontational naming is often one of the first abilities to deteriorate in patients with cortical dementia, such as in early Alzheimer disease. Patients with subcortical dementias may have difficulty retrieving the names of the objects, but they tend to do better with cueing.

Another widely used set of tasks that are sometimes categorized under expressive language are the letter and category fluency (word generation) tasks. For instance, in the "FAS" test, the examinee is given three 1-minute trials to name as many words as he or she can think of that begin with the letters "F," "A," and then "S," respectively. Category fluency involves having the examinee rapidly name words from a given category, such as types of animals. Because word knowledge tends to be organized by meaning (semantics) rather than by the initial letters, healthy individuals generally find the category fluency task considerably easier than the letter fluency task. However, with some neurodegenerative conditions, such as Alzheimer disease, the semantic links begin to deteriorate, so the typical advantage of category fluency over letter fluency performance no longer exists. This is another example of how considering the pattern of performance across tasks can be useful in the differential diagnostic process.

Attention and Vigilance

The construct of attention comprises a range of specific functions that include consciousness (attentional, arousal, and temporal orientation), simple attention (the ability to repeat recently presented stimuli, most commonly, strings of digits), selective attention (distinguishing target stimuli or information from irrelevant stimuli), sustained attention (ability to stay mentally "on track"), and vigilance (the ability to wait/look out for a particular stimulus or event over time). Some aspects of working memory, best described as a "mental workbench," also fall under the larger domain of attention (with the more complex aspects constituting executive functions). This function of working memory allows for a limited amount of information to be actively held for short periods of time for further processing and requires active rehearsal or allocation of attentional resources for it to be retained; a classic example is the holding of a telephone number in one's working memory so long as is needed to dial the telephone, and the number may be retained via active rehearsal over the delay period.

Attentional functions are among the first to be affected by a variety of neurological and psychiatric conditions. Attentional decline may therefore be readily observed in older clinical populations. Alterations in consciousness are common in elderly patients with delirium, for example, which in turn may be caused by multiple etiologies (e.g., substance use, prescription medication, urinary tract infections). Other aspects of attention may be moderately to severely impaired in patients with dementia, especially subcortical dementias such as those associated with Parkinson disease. DLB is also associated with impairments in both consciousness and arousal and simple and sustained attention. Mild to moderate impairments in attention are present among patients with Alzheimer disease, but they tend to be able to perform basic attention tasks fairly well even in more advanced stages of the disease. Fluctuating attention, especially within a single attention-oriented task, is often associated with geriatric depression or anxiety.

Orientation and simple attentional tasks are included in the MMSE, MoCA, and DRS-2. The MMSE and MoCA include the serial 7s subtraction task, while the DRS-2 includes some cancellation tasks. A common measure of simple attention and working memory is the WAIS-IV Digit Span subtest, in which the examinee is asked to repeat strings of digits of increasing length and is then asked to repeat strings of digits in reverse order. Cancellation tasks are common measures of sustained attention or vigilance; for example, in the Digit Vigilance Test from the expanded Halstead–Reitan Battery (HRB), the examinee is asked to rapidly identify and cross out every time a target number appears in a much larger array of numbers. The various versions of the Continuous Performance Test (CPT),

such as the "identical pairs" version, in which the examinee is shown a series of numbers rapidly flashed on a computer screen and is required to make a response after any number is shown twice in a row, is also an example of such a task.

Learning and Memory

Memory complaints are prevalent among older individuals, and there is consistent evidence that some aspects of memory, specifically, episodic declarative memory, deteriorate with advanced age. The function of an objective memory evaluation is to determine whether memory decline is above and beyond age-appropriate memory decline. Therefore, normative data on standardized memory tests become even more important to accurately interpret the target person's test results. Also, given that memory components of episodic memory tests are typically preceded by learning new material (e.g., word lists, stories, geometric designs), it is also important to rule out the possibility of a learning disorder (which may have remained previously undiagnosed or undetected among older patients). The presence of a learning disorder is determined by performances on achievement tests (e.g., reading, spelling, simple arithmetic) and the individual's educational or occupational history (repetition of a grade, difficulties in school, special education programs, extra tutoring, and behavioral problems in school being common red flags). If a learning disorder exists, then memory test results must be interpreted with caution, and education-corrected norms will give skewed results.

Referral questions for neuropsychological evaluation among geriatric patients typically include a request for "memory assessment," often because it is the primary cognitive complaint of patients, their caregivers, or both. There are several reasons why older individuals may experience impairment in episodic learning and memory, such as dementia, depression, anxiety, and even physical conditions, such as pain. Procedural memory (e.g., driving a car, making dinner) tends to fade at a slower rate and may only be seen in moderate to severe stages of dementia. Finally, autobiographical memory tends to remain intact for the most part, except in patients with very advanced stages of dementia.

Neuropsychological evaluation typically includes measures of simple registration (e.g., asking the patient to repeat three or five words, as in the MMSE and MoCA, respectively) and both auditory and visual episodic learning and memory. Tests of episodic learning and memory typically begin with a series of learning trials followed by short and long delay recall trials. The time between the learning and delay trials is filled with unrelated tasks (typically, another neuropsychological test or "distraction stimuli") so that the patient cannot rehearse the test stimuli. Delayed recall trials may include free recall, cued recall, recognition paradigms, or a combination of these. Cued recall and recognition trials help identify retrieval versus

encoding deficits; in the former, the information is encoded but can only be retrieved with help (e.g., cues or recognition), and in the latter, the information is never encoded in the first place (so neither cues nor recognition will help).

The most commonly used tests of auditory learning and memory include the Hopkins Verbal Learning Test-Revised, Rey Auditory Verbal Learning Test, and the California Verbal Learning Test-2nd Edition (CVLT-II), all of which are list-learning tasks, and Logical Memory (from the Wechsler Memory Scale-4th Edition [WMS-IV]) or the Story Memory Test in which the stimuli are stories or short passages. Visual learning is often measured with the Brief Visuospatial Memory Test-Revised, Visual Reproductions (from the WMS-IV), and the Rey–Osterrieth Complex Figure Test. These tests require the examinee to learn and remember geometric figures by copying them or via immediate recall, recalling them by drawing as many figures as possible from memory. These tests therefore also have a visuoconstructional element in addition to a visual memory component. The Visual Reproduction test also has recognition components.

Two classic distinctions made between memory test score profiles among older individuals are those of cortical versus subcortical dementia and dementia versus late-life depression. Cortical profiles of list-learning tasks, such as the CVLT-II, seen most often among patients with Alzheimer disease, include a flat learning slope across the learning trials, rapid forgetting (i.e., little to no words in short and long delayed recall), high recency effect (i.e., most words recalled from the last third of the list), high number of intrusions (i.e., words not on the list, similar to confabulations seen in informal assessments of patients with Alzheimer disease), high number of repetitions, and no benefit from cueing or recognition formats. Patients with subcortical dementias, such as Parkinson disease, also have flat learning slopes and difficulty recalling any words during free recall trials, but are unlikely to have intrusions or repetitions and do better with cues or on the recognition trial. The performance of patients with depression is marked by overall slowness of responding and a lower output of words during free recall. They will almost never have intrusions or repetitions, and they tend to recall words from all over the list. They tend to do significantly better with cueing and recognition, often achieving scores within the normal range for their age.

Visuospatial Cognition

The assessment of visuospatial cognition or perceptual organizational skills is an important component of both screening and comprehensive neuropsychological evaluations. Broadly, these functions consist of perceptual ability (the ability to discern forms or patterns), visuospatial judgment (the location of objects in space), and constructional ability (the

manipulation of objects in three-dimensional space). Intact visual acuity is an obvious precondition to all subdomains of visuospatial cognition, and adequate use of hands is necessary for the constructional tasks. Visual field defects and hemiparesis can also cause impairments in visuospatial cognition. Some of the most common neurological conditions affecting visuospatial cognition among older adults are vascular incidents (lesions causing contralateral hemispatial neglect), Balint syndrome (associated with bilateral lesions in the posterior parietal cortex), and dementias caused by Lewy body, Parkinson, and Huntington disease. Patients with Alzheimer disease also have impairments in visual cognition, but they are less severe compared with those among individuals with other dementias. Incipient Alzheimer disease is associated with difficulties with constructional ability, but other visuospatial abilities may only get severely impaired in advanced stages of the disease.

Visuospatial abilities are commonly measured with subtests of the WAIS-IV such as Picture Completion, in which the examinee must identify the important missing element in a series of pictured scenes, and Matrix Reasoning, a visual reasoning task wherein the examinee must identify the missing element in a matrix by discerning the rule underlying the relationship between each of the several subparts. The Hooper Visual Organization Test is another widely used measure, wherein the examinee is shown pictures of objects that have been cut and randomly arranged in pieces; he or she must mentally rotate the various pieces and put them together, and then name the object. Constructional skills are measured by graphomotor tasks, such as the Rey–Osterrieth Complex Figure Test, in which the examinee is presented with a complex geometric figure, which he or she must copy (the order in which the figure is copied is noted by the examiner), or, most commonly, the Block Design subtest (from the WAIS-IV), in which the examinee is required to rapidly match a two-dimensional design by combining three-dimensional red-and-white blocks.

Executive Functions and Working Memory

There is wide consensus among neuropsychologists that intact executive functions are central to optimal daily functioning; yet, the precise definition of the construct remains elusive. The term, “executive functions” is used to describe an array of subdomains, including abstraction and concept formation, planning, response inhibition, cognitive flexibility, and some aspects of working memory. The products of all these executive functions typically involve novel problem solving, initiating and maintaining goal-directed behavior, or disengaging from such behaviors when appropriate to the goals or in response to environmental cues. Executive functions were also previously called “frontal cortical functions” because they were thought to be mediated by the frontal cortex. However, there is now overwhelming evidence that these functions are not limited to

the frontal cortex but that they involve several frontal–subcortical circuits as well. Therefore, lesions in frontal as well as subcortical brain structures can disrupt any of the executive functions.

People with Alzheimer disease often show signs of executive dysfunction even in early stages of the disease. People with frontotemporal dementia and those with any of the several forms of subcortical dementias (such as those due to Parkinson disease) have significant impairments in cognitive flexibility, abstract thinking and conceptualization, and fluency. Executive functions may also be more sensitive to low or inconsistent effort; therefore, patients with depression may do poorly on some or all executive functioning measures.

The Wisconsin Card Sorting Test (WCST) is the most commonly used measure of executive functioning, specifically, concept formation and cognitive flexibility. The task involves asking the examinee to discern an unspoken rule that defines correct versus incorrect matches and also requires the examinee to change strategies when appropriate. The Booklet Category Test, another popular test of conceptualization and abstract thinking, similarly requires the examinee to discern the pattern or rule underlying the sequence in a sequence of presented stimuli, but that rule may change over various subsections.

The Delis–Kaplan Executive Function System (D-KEFS) is a battery of nine executive functioning tests and has normative data for adults up to 89 years. The D-KEFS includes the Trail Making Test, Verbal Fluency Test, Design Fluency Test, Sorting, Tower Test, Color–Word Interference, the Word–Context Test, the Proverb Test, and Twenty Questions. These tests, some of them modified from popular, existing tests (e.g., the Trail Making Test is modified from the Trail Making Test, Parts A and B, a.k.a., the Partington Pathways Test), measure a broad range of executive functions and are useful to include in a comprehensive battery for use with older adults.

Mental Processing Speed/Psychomotor Speed

The term “mental processing speed” (or information processing speed) refers to the speed with which one thinks and executes a task. As the tasks used to assess this construct generally include a motor response component, it is also commonly labeled “psychomotor speed,” reflecting the fact that test performance is determined not only by the speed with which one thinks, but also by the integration of those thoughts with the motor response. Whether construed as psychomotor speed or pure mental processing speed, tests in this domain tend to be among the most sensitive to virtually any form of neurocognitive injury or dysfunction, including that caused by psychiatric conditions such as depression. Indeed, the steady, if gradual, stowing of processing and psychomotor speed is the most ubiquitous and consistent change seen with normal aging, beginning at

least by the early part of the fourth decade of life. When examining an elderly patient for possible presence of pathological motor disorders, it is important to consider his or her psychomotor speed relative to that expected from an age-comparable normative sample.

The Digit Symbol (a.k.a., "Coding") test is a prototypical example of processing/psychomotor speed tests, and is among the most widely used tests to assess this domain. The Digit Symbol test consists of a stimulus key with the numbers 1 through 9, each paired with a unique geometric figure. Beneath this stimulus key is a large matrix of numbers with spaces below each number. The patient's task is to rapidly fill in the symbol paired with each number. The stimuli used for this task have appeared in largely unaltered form since the early 1900s; the stimuli used in the Army Mental Test battery during World War I are almost identical to those used in the adult versions of Wechsler intelligence test batteries, up through the 1997 edition of the Wechsler Adult Intelligence Scale (WAIS-III). The Digit Symbol subtest stimuli were substantially redesigned for the WAIS-IV, published in 2008, but there is little reason to expect that the modifications will alter the sensitivity of the task to age-related or other psychomotor slowing.

Another widely used measure of psychomotor speed is the Trail Making Test Part A (originally published as Partington Pathways Test in the 1930s), which requires the patient to connect a series of circled numbers in order as fast as possible, and measures scanning, visuospatial sequencing, and motor speed. The D-KEFS includes a number of interesting modifications of the Trail Making Test, including a Motor Speed subtest that may prove helpful in distinguishing pure motor slowing from slow mental processing.

OTHER ASPECTS OF COMPREHENSIVE NEUROPSYCHOLOGICAL EVALUATION

Assessment of Mood and Anxiety

Emotional functioning is extremely important to monitor as part of any assessment with older adults. Depression is often undetected among older adults, although increasing awareness of the frequency of psychiatric problems in this population has led to more routine assessments of depression, anxiety, and suicidality in psychiatric, neurological, and general medical treatment settings. Two useful and quick measures of depression are the Geriatric Depression Scale and the Beck Depression Inventory-2nd Edition (BDI-II). The former is particularly helpful for use with older patients and is validated for use with patients with Alzheimer disease. Occasionally, if the patient has been able to tolerate a lengthy neuropsychological battery and may have potential emotional problems, he or she may be given the Minnesota Multiphasic Personality Inventory-2nd Edition (MMPI-II). The MMPI can give a detailed profile of the patient's emotional functioning as well further highlight targets for intervention.

Activities of Daily Living/Everyday Functioning

It has been common practice in neurology and psychiatry to informally assess older patients' basic and instrumental activities of daily living (ADLs and instrumental ADLs [IADLs], respectively). Basic ADLs include personal functions such as feeding oneself, showering, dressing, and using the toilet, among other things. Intake interviews as part of neuropsychological assessments tend to focus more on evaluating IADLs, such as cooking, driving, managing finances, and so forth (i.e., skills that are more complex than basic ADLs, yet necessary for independent everyday functioning).

ISSUES IN THE ASSESSMENT OF OLDER ADULTS

Sensory Impairment

Hearing impairments tend to be fairly common among older adults in general and can considerably impede verbally administered neuropsychological tests (e.g., the CVLT-II). The potential examinee's hearing should be at least informally assessed to determine whether it is adequate for the verbal neuropsychological tests. If hearing is deemed too impaired, he or she can still be assessed with largely visual tests (assuming they have adequate visual acuity). Visual acuity may be informally assessed, for example, using the Rosenbaum Pocket Screener. Visually impaired patients can be assessed with a variety of tests, including all of the verbal memory tests, some attention tests, all of the verbal D-KEFS tests, and the Tactual Performance Test, which is a measure of visuospatial functioning.

Interpretation of Test Results in the Context of Normal Aging

Normal aging is itself associated with some changes in cognitive functioning, including improvements (at least during the first seven decades of life) in crystallized verbal knowledge, and slowing of mental processing speed, psychomotor speed, and subtle declines in some aspects of memory and executive functions. A key advantage of the standardized testing emphasized in neuropsychological assessment over other neurobehavioral examinations is that it ascertains that the interpretation of an examinee's responses is not based solely on the absolute level or content of responses (or pathognomonic signs) but also in reference to the level of performance expected from a neurologically healthy individual. The latter is determined by comparing the examinee's test scores to demographically appropriate norms. The most widely used neuropsychological tests have normative data that are developed and collected using standardized administration and scoring procedures.

REFERENCES

Boake C. From the Binet-Simon to the Wechsler-Bellevue: tracing the history of intelligence testing. *J Clin Exp Neuropsychol*. 2002;24(3):383–405.

Brewster PW, Melrose RJ, Marquine MJ, et al. Life experience and demographic influences on cognitive function in older adults. *Neuropsychology*. 2014;28(6):846–858.

*Bush SS, Martin T, eds. *Geriatric Neuropsychology: Practice Essentials*. New York: Taylor & Francis; 2014.

Caselli RJ, Locke DE, Dueck AC, et al. The neuropsychology of normal aging and preclinical Alzheimer's disease. *Alzheimers Dement*. 2014;10(1):84–92.

*Grant I, Adams KM, eds. *Neuropsychological Assessment of Neuropsychiatric and Neuromedical Disorders*, 3rd ed. New York: Oxford University Press; 2009.

Jamora CW, Ruff RM, Connor BB. Geriatric neuropsychology: implications for front line clinicians. *NeuroRehabilitation*. 2008;23(5):381–394.

Kane RL, Kane RA, eds. *Assessing Older Persons: Measures, Meaning, and Practical Applications*. New York: Oxford University Press; 2000.

Kolb B, Whishaw IQ. *Fundamentals of Human Neuropsychology*, 6th ed. New York: Worth Publishers; 2008.

*Lezak MD, Howieson DB, Bigler ED, Tranel D. *Neuropsychological Assessment*, 5th ed. New York: Oxford University Press; 2012.

Morris RG, Brookes RL. Neuropsychological assessment of older adults. In: Goldstein LH, McNeil JE, eds. *Clinical Neuropsychology: A Practical Guide to Assessment and Management for Clinicians*, 2nd ed. Hoboken, NJ: Wiley-Blackwell; 2013:347–373.

Noggle CA, Dean RS, Bush SS, Anderson SW, eds. *The Neuropsychology of Cortical Dementias*. New York: Springer; 2015.

*Ravdin LD, Katzen HL, eds. *Handbook on the Neuropsychology of Aging and Dementia*. New York: Springer; 2013.

Salmon DP, Bondi MW. Neuropsychological assessment of dementia. *Annu Rev Psychol*. 2009;60:257–282.

Savla GN, Moore DJ, Palmer BW. Neuropsychological assessment. In: Birren JE, ed. *Encyclopedia of Gerontology*, 2nd ed. Oxford,UK: Elsevier Academic Press; 2007.

Snyder PJ, Nussbaum PD, eds. *Clinical Neuropsychology*. Washington, DC: American Psychological Association; 1998.

*Strauss E, Sherman EMS, Spreen O. *A Compendium of Neuropsychological tests: Administration, Norms, and Commentary*, 3rd ed. New York: Oxford University Press; 2006.

Winocur G, Craik FI, Levine B, et al. Cognitive rehabilitation in the elderly: overview and future directions. *J Int Neuropsychol Soc*. 2007;13(1):166–171.

Young J, Meagher D, Maclullich A. Cognitive assessment of older people. *BMJ*. 2011;343:d5042.

57.2f Neuroimaging

LISA T. EYLER, PH.D., AND ZVINKA Z. ZLATAR, PH.D.

INTRODUCTION

The brain, long recognized as the site of pathology in mental illness, has

only recently become accessible for study in the living human. Modern neuroimaging techniques now permit detailed measurement of the brain's size, shape, and function. These methods provide a window into the disordered organ in mental illness and therefore provide important clues about the pathological processes that cause these disorders. In addition, neuroimaging techniques have the potential to improve the ability to diagnose and treat some of the most complex and disabling brain disorders.

In geriatric mental illness, the advantages of neuroimaging are even greater because they allow one to observe directly how the brain's aging processes might combine with disease processes. Neuroimaging methods have been used widely to examine changes in brain structure and function associated with normal aging and diseases of aging such as dementia. These findings provide a basis for examining the effect of aging on mental illness and for understanding new onset psychiatric disorders in late life. Furthermore, neuroimaging studies have provided some of the growing evidence suggesting that mental disorders may involve a process of accelerated aging.

Theoretically, there are several ways that neural processes related to aging and mental illness could combine. Out of many theoretical possibilities, there is empirical support for several scenarios. One possibility is that aging processes and mental illness produce additive effects, even when they involve similar brain systems. In this case, an older patient would have different brain findings compared to a younger patient, but the differences would parallel those seen in normal aging. For instance, brain volume in a certain region that is abnormally small in young patients with the disorder might further shrink with age, but at the same rate as seen in healthy individuals. An alternative scenario is that the effects of age and mental illness on the brain may involve similar neural systems, and these influences may interact with each other. The synergistic nature of this interaction could take several forms. One possibility is that aging processes could trigger new onset of a disorder in late life, perhaps among individuals with a genetic or environmental liability. A second possibility is that aging might alter the brain profile observed in younger individuals beyond that predicted by the effects of aging alone. For example, if functional brain response during a cognitive challenge is influenced by age and the mental disorder in an interactive fashion, a steeper decline in levels of activation with age among patients than among healthy individuals may be observed.

Most studies of the brain in geriatric psychiatry have so far sampled from two main populations: (1) elderly patients with a typical age of onset earlier in life (which, in this chapter, is referred to as early-onset elderly, or EOE) or (2) elderly patients with new onset of a disorder late in life (late-onset elderly, or LOE). In general, studies of EOE compare these individuals to healthy individuals with similar ages, or, more rarely, include

a comparison to younger patients with a comparable age of onset. Studies of LOE frequently compare findings to a group of age-comparable healthy participants and may include a group of EOE of comparable age. A fully factorial design crossing age of onset and current age, and including young and old healthy individuals is rarely used, even though it is ideally suited to disentangle the effects of age from the age of onset and to examine the nature of how age and illness interact. Longitudinal studies, perhaps the gold standard for examining age-related changes, are very rare in neuroimaging studies of geriatric mental illness, but can provide important insights into the trajectory of mental illness across the adult life span.

OVERVIEW OF NEUROIMAGING TECHNIQUES

Modern neuroimaging techniques can be divided into two primary categories: (1) those that measure the size and shape of tissues and structures within the brain, known as structural neuroimaging, and (2) those that measure the activity of the brain, known as functional neuroimaging.

Structural Neuroimaging

Structural neuroimaging uses techniques such as computed tomography (CT) and magnetic resonance imaging (MRI) to create images of different tissue types within the brain, namely gray matter, white matter, and cerebral spinal fluid. MRI is the preferred structural neuroimaging technique for research and clinical uses in mental illness because it lacks ionizing radiation, can easily be repeated, and generally provides good spatial detail and a sharp division between tissue types (see Fig. 57.2f–1). The digital images collected using MRI can be manipulated, either through use of mathematical algorithms or human experts, in order to measure quantities such as volume of substructures or tissue types, number and size of certain abnormal occurrences such as white matter lesions, and cortical surface area and thickness. In addition, detailed information about the shape of three-dimensional structures and cortical folding patterns can be extracted from these images and mathematically compared between individuals using the latest computerized techniques. Measurement of the organization, integrity, and connectivity of the brain's myelinated fiber tracts can also be achieved using diffusion tensor imaging (DTI), which is a form of MRI that is sensitive to the constrained and directional diffusion of water molecules in white matter tracts. Finally, the degree of local neuronal integrity can be inferred through the measurement of chemical compounds such as *N*-acetyl aspartate using a variant of MRI called magnetic resonance spectroscopy (MRS).

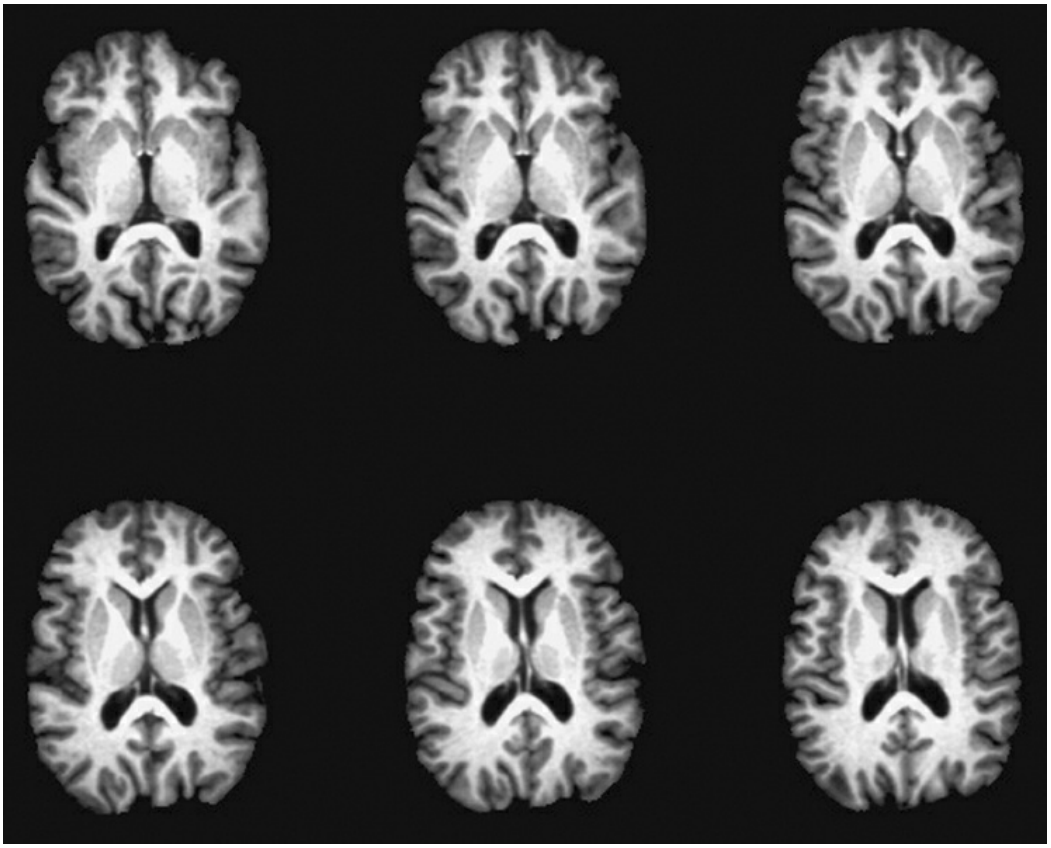

FIGURE 57.2f–1. High-resolution T_1-weighted magnetic resonance image of brain structure showing six axial slices at the level of the lateral ventricles.

Although MRI-based structural imaging provides a wealth of quantitative data about the static properties of the living brain, there are several limitations that should be kept in mind when evaluating its use in the study of geriatric mental illness. In addition to practical constraints due to the expense and lack of portability of the instrument, safety issues associated with MRI exclude a proportion of patients, particularly older ones, who may have pacemakers or metal in soft tissue. Furthermore, scanning is difficult for those who cannot comfortably lie still for long periods of time (a more frequent concern in the elderly), those who are very obese (which may preferentially exclude some mentally ill patients for whom weight gain is a side effect of medication), and those who have claustrophobia (an anxiety that may be more common among mentally ill patients). These difficulties may reduce the generalizability of research findings due to a biased sample and may also limit the usefulness of MRI as a clinical tool for some patient groups. Also, image quality and reliability is severely reduced in MRI by participant movement during the scan. To the extent that certain groups may have greater trouble staying still (due, e.g., to motoric side effects of medication), this too can lead to a biased sample at the time of data analysis. A final caveat regarding MRI structural measures is that, like all assessments, they are subject to potential problems with measurement error, reliability, and validity. Variability

between and within individuals across multiple measurements can occur due to mechanical fluctuations in the hardware and due to variability in processing of the data including that introduced by the use of human experts. In the case of validity, measurements based on MRI images, which are, after all, only pictures of the brain, may or may not have systematic biases relative to measurements of the brain itself. Studies comparing features of postmortem brain dissections with measurements based on MRI images of the same specimens are underway to help address this issue.

Functional Neuroimaging

In contrast to the static measures provided by structural neuroimaging, several techniques are now available that attempt to measure the dynamic function of the brain. In some cases, activity is measured during a cognitively unstructured resting state; in other cases, the brain's response to a sensory, cognitive or emotional challenge is assessed. Positron emission tomography (PET), functional magnetic resonance imaging (fMRI; see Fig. 57.2f–2), electroencephalography (EEG), and magnetoencephalography (MEG) are the most frequently used techniques to examine the brain in action; each has its strengths and weaknesses. PET scanning uses radiolabeled compounds to measure blood flow, cerebral metabolism, or receptor occupancy. PET measures of blood flow have the advantage of permitting quantification of absolute levels of regional perfusion, and measures of cerebral glucose metabolism and receptor occupancy are only available using PET technology. More recently, PET techniques have been used to measure in vivo brain beta-amyloid deposition, the hallmark neuropathology of Alzheimer disease, which previously could only be achieved postmortem. Disadvantages of PET relative to other functional neuroimaging measures include poorer spatial resolution compared to MRI-based measures, coarser temporal resolution, limits on the ability to repeat scans due to concerns about exposure to radioactive compounds, and less availability for research and clinical use because of the need to synthesize the necessary agents on site.

Functional MRI uses endogenous sources of contrast to create images of blood oxygenation and blood flow. The most common signal measured by fMRI is blood oxygen level dependent (BOLD), which takes advantage of the fact that more oxygen is delivered to the vessels surrounding active neurons than is consumed by them, creating a surplus of oxygenated hemoglobin in the local area. Oxygenated hemoglobin is less disruptive to the MRI signal than deoxygenated hemoglobin, so the BOLD signal is relatively greater in active regions. The BOLD signal can be used to determine changes in blood oxygenation at the whole brain or regional levels, and it can be measured at rest or during a specific task (cognitive, affective, motor, etc.). Data obtained by measuring the BOLD signal can be analyzed in different ways to answer distinct questions. For example, group

differences (age, disease state, etc.) and within person change in the location and shape of the BOLD signal in response to different tasks can be measured, which can help to elucidate whether age or disease states manifest in differential brain activity patterns. Brain activity changes can also be measured over time in the same individuals, or before and after a particular intervention to assess whether or not the intervention changed brain activation patterns, and purportedly, neural activity. Since the brain is a network of interconnected and spatially distributed regions, it is also important to study how connectivity between different regions and within brain networks may be affected by age or disease states. Functional connectivity analyses apply the BOLD signal to study temporal connectivity patterns between different brain regions and networks to investigate how these connections may be affected by age or disease states. Correlations of the BOLD signal between regions can be measured either during a cognitive challenge task or during rest. Data from resting state fMRI scans have recently been used to identify reliable large-scale networks of correlated BOLD signal and to examine patterns and properties of these networks. Another MRI-based functional technique, called arterial spin labeling (ASL), can be used to quantify the amount of blood flowing into an area by measuring the inflow of an applied magnetic "tag." Both of these methods have millimeter spatial resolution, temporal resolution on the order of seconds, and have the advantage of being able to measure brain structure in the same scanning session as the functional images. A disadvantage of BOLD fMRI relative to ASL is that it does not measure an absolute quantity, so activation must always be compared between an experimental condition and some baseline condition. The most appropriate choice for a baseline comparator has been widely debated. As MRI-based techniques, the same caveats mentioned above for structural scans also apply to BOLD and ASL, with an even greater concern about the effect of participant movement, especially movement that is coordinated with the cognitive activity of interest. In addition, fMRI images are susceptible to unique artifacts that make visualization of particular areas of the brain, such as the orbitofrontal cortex, more challenging. Finally, both PET and fMRI are blood-based measures, so differences between groups might reflect vascular rather than neuronal deficits, although measuring baseline cerebral blood flow using ASL can help to disambiguate group brain activation differences by better characterizing the vascular contributions to the BOLD signal.

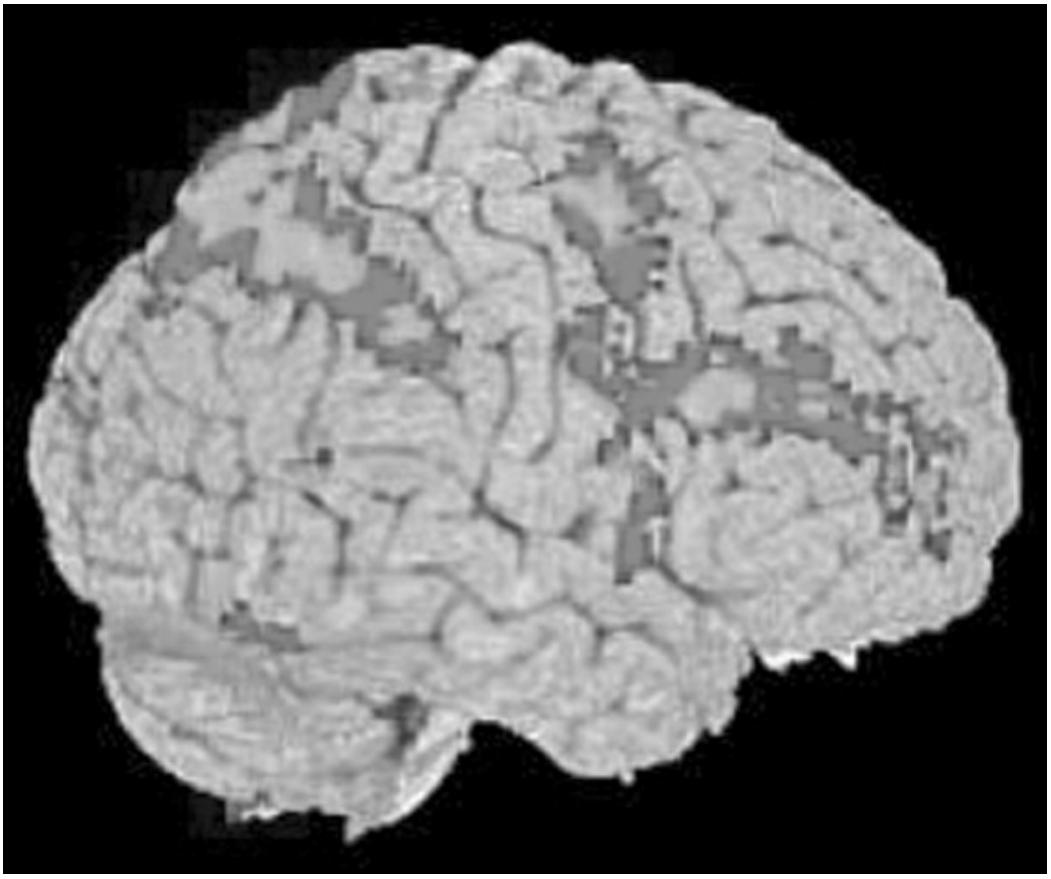

FIGURE 57.2f–2. Processed image of areas of significant blood oxygen level–dependent (BOLD) response as measured by functional magnetic resonance imaging. Brightest contrast within dark areas indicate those areas where change in the BOLD signal over time was most highly correlated with the behavioral paradigm presented, and these activation clusters are superimposed onto a three-dimensional rendering of the cortical surface based on a high-resolution magnetic resonance scan.

In contrast, EEG and MEG both measure the electrical activity of the brain using detectors at the surface of the scalp. The temporal resolution is, therefore, on the order of milliseconds, similar to the timing of neuronal processes. The spatial resolution, however, is generally coarser than MRI-based measures and localization of deep brain sources is especially problematic, although mathematical source localization techniques have recently improved the ability to pinpoint the origin of the electrical and magnetic signals within the brain.

LATE-LIFE DEPRESSION

Course and Features in Late Life

Unipolar depression often has a recurring course; so many individuals with onset early in life continue to suffer from the disorder into old age. In addition, a proportion of individuals will manifest depression for the first time in late life. The combined prevalence of EOE and LOE patients with depression varies tremendously and it ranges from ~1 to 35 percent in

community-dwelling samples, with major depression in ~1.8 percent and minor depression in ~9.8 percent of the community-dwelling population. The prevalence of major depression is higher in nursing home settings (12 to 14 percent). In general, the core symptoms of late-life depression (LLD) are similar to those seen in younger patients and include emotional distress, somatic complaints, and cognitive impairment. Cognitive deficits are more prominent among elderly patients with depression compared to younger depressed individuals, particularly in the realm of executive function. LOE with depression may have more pronounced cognitive deficits than EOE patients and there appears to be a stronger association between depression and dementia among LOE patients.

Features of LLD that contrast with depression earlier in life include greater variability in treatment response, high medical comorbidity, increased nonsuicidal mortality, and less reported family history of depression. Some of these features may be particularly linked to late-onset depression. LOE patients with depression suffer from less psychiatric comorbidity and have fewer relatives with depression, but have even greater rates of medical comorbidity compared to EOE patients. The distinction between late- and early-onset depression is complicated, however, by the fact that chronic, subclinical depression symptoms may have existed, or mild mood episodes may have occurred earlier in life, without proper diagnosis or treatment among some LOE patients. Thus, some LOE with depression may be more properly classified as EOE with a milder initial course, although there are also clear cases of first onset of mood problems late in life.

Structural Neuroimaging

Compared to age-matched healthy individuals, patients with LLD have smaller brain volumes of several cortical regions, including prefrontal and orbitofrontal cortex and anterior cingulate, and of subcortical regions including the basal ganglia, hippocampus, and amygdala. In addition, LLD is characterized by cerebrovascular disease with white and gray matter hyperintensities, particularly in subcortical structures and frontal projections. The integrity of frontal and temporal white matter connections as measured by DTI also appears to be reduced in geriatric depressed patients, implicating a possible frontostriatal disconnection syndrome. Finally, at least one study using MRS has revealed disruptions of the neurochemical properties of the frontal white matter in geriatric depressed patients.

Hippocampal volume reductions may be particularly pronounced among LOE with depression, and a recent study that compared surface characteristics of hippocampal shape between age-matched EOE and LOE patients showed more pronounced deformations in several hippocampal subregions among LOE compared to EOE patients. Furthermore, a recent

study that followed LOE- and EOE-depressed patients over 4 years found that EOE had greater severity of depressive symptoms, while LOE was associated with more severe cognitive and neurological changes, and with accelerated rate of hippocampal volume loss compared to EOE. EOE-depressive symptoms have also been associated with greater disturbance in cortical thickness compared to LOE. White matter hyperintensities are also strongly linked to late onset of depression: a recent meta-analysis found an odds ratio greater than 4 for these abnormalities in LOE versus EOE with depression as well as greater severity among LOE than EOE. White matter abnormalities have also been linked to worse functional outcomes and mobility measures in depressed older adults compared to those never depressed.

The direction of causality of the associations between LLD and these structural abnormalities is not yet clear. Smaller cortical and subcortical volumes may be predisposing factors for depression (since they are also seen in younger adult samples), but there is also a suggestion that stress hormones associated with depressive episodes may cause volume loss. The association between smaller hippocampal volume and longer duration of illness is consistent with the latter interpretation, while the recent finding that hippocampal atrophy was associated with higher subsequent depressive symptoms and with shorter-term symptom exacerbation in women followed for over a 10-year period is suggestive of the former. In terms of white matter damage, the established link between stroke and late-onset depression has led to the conclusion that at least some LLD is vascular in origin. However, it also may be the case that history of depression predisposes to cardiovascular problems and that the observed white matter lesions are a result of the disorder rather than the cause. Several ongoing longitudinal studies should help to disentangle these possibilities. In at least one of these, initial results suggest that presence of white matter hyperintensities independently predicts depressive symptoms 1 year later.

Given the observed cerebrovascular pathology and increased risk of dementia in those with LLD, especially in those with LOE depression, recent studies have begun to investigate the association between LLD, cognitive decline, and brain structure as a predictor of conversion to dementia. Studies comparing four groups of individuals (cognitively normal, LLD, amnestic mild cognitive impairment [aMCI], and comorbid LLD and aMCI) have found increased white matter vulnerability in the comorbid group, suggesting that white matter disruptions may be a marker of disease severity. A similar study found that gray matter volumes were associated with mood-regulating circuit structures in LLD, while associations with regions subserving multicognitive domain functions were found in those with aMCI only. The group with comorbid LLD and aMCI showed gray matter atrophy in widespread regions implicated in early

Alzheimer disease, such as the dorsal anterior cingulate cortex, the medial-temporal lobe, and prefrontal cortices.

The relationship of structural brain abnormalities to treatment response in LLD also has been examined. The results, unfortunately, have been mixed. Some studies find that increased white matter hyperintensities and volume deficits are associated with poor short-term response to antidepressants, but others have found no such association. Importantly, since studies have found gender differences in the underlying neural substrates of LLD, such as higher white matter disease as a consequence of depressed mood only in men, future studies should take into account gender differences when attempting to parcel out the contribution of structural and functional neural deficits to LLD diagnostic features and disease course.

Functional Neuroimaging

In LLD, disruptions in metabolism, functional activity, and network connectivity have been identified. Hypometabolism and hypoperfusion have been observed compared to age-matched healthy individuals. A recent PET study of individuals from age 20 to 79 found a significant negative relationship between age and glucose metabolism in the rostral and dorsal anterior cingulate cortex that was not observed in a similarly aged healthy comparison group. This negative relationship with age was most pronounced in a subgroup of patients with a higher degree of illness severity; there was no such correlation among first-episode patients. Studies during cognitive tasks have revealed underactivation of the dorsal anterior cingulate and hippocampus during a verbal fluency task, decreased prefrontal and increased caudate activation during sequence learning, and decreased temporolimbic response during episodic memory among individuals with LLD compared to age-matched healthy volunteers. A role for illness duration and severity is also suggested in at least one cognitive challenge study that found anterior cingulate hypoactivation only in the LLD subgroup that had suffered multiple episodes. Potential functional brain differences between depressed LOE and EOE have not yet been examined.

Several studies have examined the neurophysiology of LLD. When measuring brain stem auditory–evoked responses using EEG, the wave V latency (linked to functioning of the pons) was found to be longer among patients than comparison participants, particularly those with comorbid vascular disease. The latency of P300 response has also been found to be increased in elderly depressed patients. More recent studies have examined evoked potentials during a go/no-go paradigm and found greater responses during inhibitory trials on waveforms thought to reflect anterior cingulate cortex conflict monitoring. As of late, many studies have attempted to combine imaging modalities in an effort to better understand the neural

changes associated with health and disease. One such study combined EEG and MRI modalities and found changes in electrophysiological markers of cerebral arousal and information processing in LLD, some of which were related to white matter hyperintensities. Another multimodal imaging study found elevated glucose metabolism with corresponding atrophy in frontal and posterior regions using PET and MRI in a small sample of older adults with depression, supporting a possible compensatory metabolic response resulting from structural atrophy in these regions.

Disruptions in several major brain neurocircuits important for cognitive and emotional functioning have been identified in LLD using several imaging modalities (resting state fMRI, structural MRI, DTI, EEG, MEG). A recent review of the literature indicated that individuals with LLD exhibit dysfunction in the default mode network (medial prefrontal cortex, posterior cingulate, precuneus, and medial temporal lobe), the cognitive control network (dorsolateral prefrontal cortex, dorsal anterior cingulate, and posterior parietal cortex), the affective/frontolimbic network (amygdala, subgenual anterior cingulate, hypothalamus, orbitofrontal cortex, and nucleus accumbens), and corticostriatal networks (connections between frontal cortical regions and the basal ganglia and thalami). Above and beyond alterations in discrete network systems, abnormal function in whole-brain network topographical properties have been identified in LLD using newer methods such as graph theory. Graph theory–based techniques can estimate network efficiency (integration) and resilience (ability to maintain its structural and functional integrity when challenged) and recent studies using this technique have found reduced global efficiency, network strength, and network resilience in LLD. Furthermore, white matter disease burden may mediate these network alterations in LLD.

The role of cognitive impairment and risk for dementia in LLD has been further explored in recent studies. One such study followed people with depression for 2 years and found that those with persistent cognitive impairment at follow-up had decreased baseline activity in dorsal anterior cingulate, hippocampus, inferior-frontal cortex, and insula compared with noncognitively impaired depressed patients. Similarly, resting state functional connectivity analyses were used to measure hippocampal frontal connectivity similarities and differences in LLD and aMCI. While aMCI was associated with overall diminished connectivity, LLD was associated to increases and decreases in connectivity in cortical and subcortical structures involved in mood regulation. Interactive effects of LLD and aMCI were found in frontal regions associated with emotion processing and cognitive functioning, where LLD and aMCI showed diminished hippocampal connections whereas the comorbid group exhibited enhanced connectivity.

Given the variability in treatment response in geriatric depression and

concerns about poor prognosis in this age group, there has been considerable interest in using functional neuroimaging to understand and improve treatment response. For example, a large body of work has used PET methodology to examine neurotransmitter binding. Among patients with midlife depression, however, only modest differences in binding of transporters or receptors for serotonin and dopamine compared to healthy individuals have been observed. Other studies have used metabolic or electroencephalographic imaging in the context of treatment studies to examine change with treatment. Geriatric patients treated with sleep deprivation and paroxetine showed a decrease in metabolism in the anterior cingulate cortex that was correlated with symptom improvement. A more recent study administered citalopram for a period of 8 weeks to older adults with LLD compared to a control group while evaluating changes in glucose metabolism over a 2-year follow-up. Compared to the control group, patients in the treatment group had higher glucose metabolism in the anterior cingulate and insula, with most patients being remitted at follow-up. Furthermore, both groups showed decreased metabolism in posterior association cortices implicated in dementia, which were more extensive than the decreases in brain volume. Even though the latter study included a very small sample, investigations such as this one can help to elucidate the underlying mechanisms of treatment response in LLD.

Summary and Clinical Implications

Structural MRI and DTI have consistently revealed white matter abnormalities in patients with LLD, although this is not the case for all studies. The current evidence associating cerebrovascular and cardiovascular disease with LLD is sufficiently strong that a cerebrovascular workup should be considered, especially when late-onset, elderly, depressed patients present in the clinic. Recent studies have identified that comorbid MCI and depression is associated with more frontal, parietal, and temporal white matter atrophy compared to those without depression over a 2-year period. Similarly, in individuals with MCI and depression, higher beta-amyloid pathology was found in those with LOE compared to EOE, while higher beta-amyloid load was found in comorbid MCI and depression compared to depressed individuals without cognitive impairment. Given the increased conversion rates from MCI to dementia in older adults with depressive symptoms, abnormal hippocampal volume or shape, increased cerebrovascular disease burden, and metabolic changes in areas associated with dementia presenting in concert with memory impairment among elderly patients with late-onset depression may be prognostic signs of cognitive decline. The care of such patients should include routine assessments of cognitive functioning and behavioral interventions, such as increasing engagement in exercise and

other cognitive and social activities, to preserve brain health and boost cognitive function. As with depression in young adulthood, depression in the elderly is associated with impaired glucose metabolism and abnormal brain activation and network connectivity patterns. The prognostic value of abnormal metabolic findings and brain activation patterns as predictors of treatment outcome remains to be determined, although some progress has been made through short-term longitudinal studies. The current literature, however, indicates that abnormal findings in subdivisions of the anterior cingulate might be as useful in predicting outcome and monitoring treatment response in LLD as they are in predicting treatment response in younger depressed patients.

LATE-LIFE BIPOLAR DISORDER

Course and Features in Late Life

In contrast to the growing body of evidence regarding late-life unipolar depression, less is known about the course of bipolar disorder in late life. Although some authors have suggested that symptoms of bipolar disorder may eventually "burn out," longitudinal studies show a recurrent and chronic course into old age. Epidemiological studies find a low community prevalence of bipolar illness in older individuals (between 0.1 and 0.5 percent), but among geriatric psychiatry visits, 8 to 10 percent of cases have bipolar disorder diagnoses and nearly 25 percent of bipolar disorder patients are over the age of 60. Approximately 6 to 8 percent of patients over age 60 have first experienced symptoms of bipolar disorder in late life. Clinically, these late-onset cases appear broadly similar to those with more typical onset in their early 20s. In general, older patients with bipolar disorder are more likely to be women, have lower substance abuse comorbidity, high physical illness comorbidity, and low perceived social support. Although manic symptoms may be less severe in older patients, hospital stays are longer suggesting slower resolution of symptoms. Older patients are more likely than younger ones to take valproate rather than lithium, although treatment guidelines specific to older patients are lacking.

A small number of recent studies have examined cognitive performance among elderly, nonsymptomatic individuals with bipolar disorder. Older bipolar patients performed worse compared to age-comparable healthy individuals on broad measures of cognition, and, when tested with comprehensive neuropsychological batteries, had deficits in multiple domains, including those linked to the function of prefrontal cortex such as working memory, attention, and problem solving. Whereas in younger samples, cognitive deficits of bipolar patients tend to be intermediate between those of patients with schizophrenia and those of healthy individuals, one study found that older patients with bipolar disorder were

as impaired as schizophrenia patients in most domains, including abilities like attention/working memory and reasoning/problem solving. This raises the question of whether cognitive performance may worsen with age among patients with bipolar disorder and whether the degree of worsening is greater than expected due to the effects of age alone. Two studies have addressed this question in a cross-sectional design. When groups of younger and older bipolar patients were compared, the older group had greater cognitive impairment, and the degree of deficit compared to unipolar patients and spouse controls was greater among the older group than among the younger group. Longitudinal studies of cognitive changes in bipolar disorder are difficult to conduct and interpret due to changes in mood state that increase within-subject variability. Manic patients over the age of 60 showed steeper cognitive decline than a comparison group, but several other studies in which mood state was required to be the same at baseline and follow-up showed no evidence of more rapid cognitive changes. Future work is needed to understand which patients might be at risk for cognitive declines and how variability of mood state contributes.

Structural Neuroimaging

Despite a growing interest in cognitive performance in geriatric bipolar disorder, structural brain abnormalities associated with bipolar disorder in late life have not been extensively studied. The few studies comparing elderly bipolar patients to age-equivalent healthy individuals have shown structural abnormalities, including more frontal white matter hyperintensities, smaller right caudate volume, greater hippocampal volume, and reduced fractional anisotropy (FA) in the corpus callosum. Some studies of predominantly young to middle-aged adult patients have examined the relationship of age to structural abnormalities and generally found more abnormality with older age, although other studies found no association. Neuronal integrity of the right lenticular nucleus as assessed by MRS was also negatively related to age in one study. Two cross-sectional studies directly compared the age correlation between patients and comparison subjects. One found that bipolar patients had a much larger negative correlation between age and total gray matter volume than healthy individuals. The other found that patients and healthy controls had a similar slope of age-related change in dopamine D_2-receptor occupancy. Evidence from longitudinal studies is mixed: one found that gray matter in hippocampus, fusiform gyrus, and cerebellum declined more over 4 years among bipolar patients than among healthy individuals, but another showed no change over 2 years. No changes in white matter volume were observed over 4 years in patients whose age at first assessment averaged 40 years, and DTI measures of FA showed a similar pattern of change over 2 years in bipolar patients compared to a matched group. The mixed

neuroimaging findings emphasize the heterogeneity among bipolar patients in terms of age-related change. Much more work is needed to understand the role that other age-related factors such as duration of illness, medication history, substance abuse history, and medication comorbidity may play in observed age-related decrements. It will be particularly important to conceptualize bipolar disorder in late life as a multisystem disorder and account for the complex interplay of immunity, inflammation, blood sugar regulation, and cardiovascular function. Each of these is likely to have independent and interacting effects on the brain and therefore influence the course of neurocognitive aging in late-life bipolar disorder.

Onset of bipolar disorder in late life also has been associated with structural brain abnormalities. LOE (onset >45 years) patients were found to have decreased caudate volume relative to age-matched healthy controls and decreased brain volume relative to an EOE group, and another study found that LOE patients had bilateral-temporal sulcal and cortical sulcal atrophy compared to normal, age-matched individuals. Other studies examined white matter hyperintensities and silent cerebral infarctions and found them to be more prevalent among LOE compared to EOE bipolar patients. Consistent with this, among a group of elderly patients with mania and a wide range of ages of onset, later age of onset was related to greater frontal white matter hyperintensities. A negative relationship of age of onset to neuronal integrity (as assessed by the *N*-acetyl aspartate to creatine ratio measured by MRS) of the right lenticular nucleus has also been observed.

Functional Neuroimaging

A growing body of evidence suggests that young to middle-aged patients with bipolar disorder have abnormalities of resting blood flow and metabolism, and altered electrical and blood-based brain response to both cognitive and emotional challenges. These dysfunctions are apparent during episodes of mania and depression but also during periods of euthymic mood. Unfortunately, the authors are not aware of any functional neuroimaging studies that have specifically studied older patients with bipolar disorder or examined a late-onset group of patients. Clearly, this is an important area for future exploration, as functional imaging studies in geriatric psychiatry could help to improve diagnostic specificity, explore mechanisms of treatment response, and understand the interaction of age and age of onset with disease-related pathophysiology.

Summary and Clinical Implications

Fewer studies have examined brain structure and function among elderly patients with bipolar disorder compared with studies of late-life unipolar

depression. The studies that have been completed suggest greater involvement of the basal ganglia in late-onset bipolar disorder compared with late-onset unipolar depression. White matter abnormalities can be observed in both late-life bipolar disorder and unipolar depression, reducing their usefulness in the differential diagnosis of the two conditions. Even fewer studies have examined abnormal brain activity or metabolism in elderly bipolar disorder. The research that has been completed in young and middle-aged individuals shows that abnormal brain metabolism can be observed even in euthymic states.

LATE-LIFE SCHIZOPHRENIA

Course and Features in Late Life

Schizophrenia symptoms typically occur for the first time in late adolescence and the course of the disorder is then generally fluctuating but unremitting into old age, although some proportion of individuals can experience true remission. Thus, there are many elderly individuals who have been suffering from schizophrenia most of their life and are still symptomatic. In addition, new onset of schizophrenic symptoms can occur in middle age, and these individuals have a very similar clinical picture to early-onset cases, albeit slightly less severe. For instance, LOE with schizophrenia onset in middle age have similar levels of positive symptoms, functional impairment, and family history of psychosis, but fewer negative symptoms, lower average doses of antipsychotic medication compared to EOE patients, and somewhat better adjustment in young adulthood. A final group of elderly patients with psychosis have onset after age 60 and have been termed very–late-onset schizophrenia-like psychosis. These individuals are more predominantly female, have lower rates of thought disorder and negative symptoms, have a lower genetic load for schizophrenia, and are more susceptible to antipsychotic side effects, suggesting that they may suffer from a qualitatively different disorder.

Cognitively, LOE patients with onset in middle age have deficient neuropsychological function relative to healthy individuals and display a pattern that resembles that of EOE patients, except that the deficits of LOE patients in learning and abstraction are less severe than those seen in EOE patients. Widespread cognitive deficits also are observed in LOE patients with onset after age 60. There is some debate about whether cognitive performance declines with age among EOE patients with schizophrenia more rapidly than would be expected based on age alone. Studies of EOE elderly who have been institutionalized suggest that a subgroup of patients with very poor outcome shows a steep decline in cognitive function late in life that mimics a dementia. In contrast, EOE outpatients living in the community were found to have declines in neuropsychological performance that did not differ from those of age-comparable healthy individuals.

Structural Neuroimaging

Structural brain abnormalities such as ventricular enlargement, reduced hippocampal size, and reduced whole-brain volume are present in first-episode patients with schizophrenia, and some findings (especially ventricular enlargement) appear to worsen over time, both near the time of onset and in more chronic patients. In one study, some patients followed for 10 years after their first episode had continual ventricular enlargement at a significantly faster rate than that observed in healthy individuals. Progression may be particularly common within the subgroup of EOE patients with very poor outcome and steep cognitive declines in old age. These patients have larger ventricles than both healthy individuals and better outcome patients, show progressive increase in ventricle size, have a more posterior distribution of gray and white matter volume deficits, and evidence more severe and widespread deficits in white matter organization as measured by FA based on DTI. Unfortunately, few of the longitudinal studies in chronic patients have included a healthy comparison group, so it is unknown whether the rate of change later in life is more precipitous than that seen in healthy aging. LOE with schizophrenia also show structural deficits relative to age-matched healthy individuals, although at least one study has suggested that the volume of the thalamus is greater in LOE than EOE. Other studies show that LOE schizophrenia patients have more white matter hyperintensities than those with typical-onset schizophrenia. Some studies have also reported more signs of white matter abnormality and focal cerebrovascular findings among patients with very–late-onset schizophrenia-like psychosis, but other investigations that carefully ruled out organic pathologies have not. One study found greater cerebellar atrophy among very–late-onset schizophrenia-like psychosis patients compared to EOE with schizophrenia.

Functional Neuroimaging

Several functional neuroimaging studies using PET and fMRI have compared the brain function of EOE patients with schizophrenia to age-matched comparison groups and found abnormalities similar to those seen in younger patients, including hypofunction of frontal cortex and abnormal hippocampal activation. Additional PET studies have directly examined age-related change in brain functioning within schizophrenia; although most of the studies were cross-sectional, the samples were generally small and mostly middle aged, and brain function was measured predominantly at rest (see Table 57.2f–1). In general, the reviewed cross-sectional studies have found evidence for negative correlations of hemispheric and regional brain blood flow and metabolism with age, and those that included a control group showed steeper negative slopes among patients in some regions including frontal and temporal cortex and the caudate. In contrast,

one study found regions with more positive age-related slopes among patients, and the only longitudinal study observed no change in degree of hypofrontality across 18 years. Also, it is important to note that the largest study to date, while finding a negative correlation between age and mean hemispheric blood flow, also observed a strong relationship of duration of illness to hypofrontality. In a follow-up study, the same authors found that the relationship of duration of illness to anterior–posterior blood flow ratio persisted despite controlling for and statistically covarying age in the analysis. In a recent investigation, correlations of age with functional brain response to a verbal learning task as measured by fMRI were examined in both patients with schizophrenia and healthy individuals. The authors did not find significant negative correlations with age among the schizophrenia patients; patients specifically failed to show normal positive age correlations in regions associated with the "default mode" (a network of brain regions that is more active during passive baseline conditions than during active experimental tasks).

LOE patients with schizophrenia have been found to have deficits in brain function relative to healthy individuals that are similar to those seen in separate studies of younger patients with schizophrenia and EOE patients. Specifically, deficient blood flow in frontal and temporal regions has been consistently observed in LOE with schizophrenia, and elevated D_2-dopamine receptors in LOE has also been found. The only direct comparison of brain function in LOE compared to EOE patients with schizophrenia found reduced blood flow in temporal cortex among the LOE individuals.

The electrophysiology of late-life schizophrenia also has been examined. EEG abnormalities, such as decreased amplitude of the P300-orienting response, have been observed in young individuals with schizophrenia and in EOE with schizophrenia compared to healthy individuals. The latency of the P300 response has also been found to relate positively to age among patients with schizophrenia. Since P300 latency and amplitude become more abnormal in healthy aging, it is not clear whether the age-related changes in schizophrenia are greater than expected. The mismatch negativity component, in contrast, does not seem to be impaired in first-episode younger patients, whereas deficits are commonly found in more chronic and older patients with schizophrenia. Longitudinal studies will be needed to confirm that these differences reflect steeper age-related decreases among the schizophrenia group rather than cohort effects. Interestingly, general EEG and P300 deficits have not been found among LOE with schizophrenia, although an increased latency of the N400 semantic incongruity response was observed in LOE patients but not EOE patients compared to controls in one small study.

Summary and Clinical Implications

Like young patients with schizophrenia, ventricular enlargement and reduced hippocampal volume have been observed when brain images of older patients with schizophrenia are compared with age-matched, healthy individuals. Cortical thinning of the posterior third of the superior temporal gyrus on gray matter parcellated images, another consistent imaging finding in younger patients with schizophrenia, has not been systematically investigated in older patients. Similar patterns of abnormal resting cerebral metabolism and blood flow have been observed in younger and older patients with schizophrenia. Whether the pattern of regional brain abnormality for resting blood flow and metabolism differs in some subtle manner between younger and older patients with schizophrenia awaits an appropriately powered study comparing these two groups. Despite the interesting fMRI data showing a complex pattern of increased and decreased brain activation among patients with schizophrenia, little research using cognitive challenge tasks to study brain activation in older patients has been reported. Findings from current P300 and N400 studies of late-onset schizophrenia patients suggest that differences in functional brain activation of early- and late-onset patients might exist. In contrast to growing interest in the use of functional and structural brain imaging to predict and measure treatment response in geriatric mood disorders, these techniques have not been widely applied in the study of psychosis in the elderly.

Table 57.2f–1.
Studies of the Relationship of Age to Brain Function among Patients with Schizophrenia

Citation	N	Age	Sex	Chronicity/Medication Status	Imaging Technique	Task
Ingvar & Franzen, 1974	9 older SZ; 11 younger SZ	Old SZ: 61.3 (7.0); range = 49–71; young SZ: 24.5 (5.2); range = 17–33	Old SZ: 0M:9F; young SZ: 7M:4F	Old SZ: chronically institutionalized; young SZ: duration of 8 months–15 years; all medicated (almost all typical antipsychotics)	Xenon injection	Rest (ol ob na yo Ra pr ma
Mathew, 1988	108 SZ; 108 HC	SZ: M = 31.5 (8.9) F = 38.3 (12.4);	SZ: 60M:48F; HC: 60M:48F	Chronic (mean duration 11 years); 46 were medication withdrawn for 2 weeks, rest on	Xenon inhalation	Rest

		HC: M = 32.0 (13.1) F = 38.9 (13.1)		mostly typicals		
Mathew & Wilson, 1990	Long duration SZ: 27; short duration SZ: 27; HC: 2 matched groups of 27 each	Long duration SZ: 41.8 (10.4); short duration SZ: 31.4 (12.8); HC matched to long duration: 41.5 (11.4); HC matched to short duration: 32.4 (13.9)	Long duration SZ: 11M:16F; short duration SZ: 17M:10F; HC matched to long duration: 11M:16F; HC matched to short duration: 17M:10F	Medicated inpatients, long duration: 21.6 (8.5) years; short duration: 2.4 (1.7) years	Xenon inhalation	Rest
Goldstein, 1990	SZ: 15; major affective: 15; HC: 15	SZ: 34.6 (11.9); major affective: 36.9 (9.7); HC: 36.2 (12.2)	SZ: 8M:7F; major affective: 4M:11F; HC: 9M:6F	SZ: duration 2 to 28 years, 12/15 medicated; major affective: duration 5–26 years, 14/15 medicated	Xenon inhalation	Rest
Cantor-Graae, 1991	SZ: 7 (of the 11 young SZ reported in Ingvar & Franzen, 1974)	Time 1: 23.0 (5.3), range = 17–32; Time 2: 41.4 (5.0), range = 34–49	5M:2F	Chronic: 22.9 (2.9) years duration of illness; typical anti-psychotics; 5/7 had ECT	Xenon inhalation	Rest, flu W ass co tas
Siegel, 1994	18 SZ; 22 HC	SZ: 29.6 (7.2); range =	SZ: 18M:0F; HC: 22M:0F	Never medicated; most first episode	FDG PET	Degra sti CP

		21–45; HC: 28.2 (7.2); range = 19–42				
Shihabuddin, 1998 (age results reported in Buchsbaum & Hazlett, 1997)	SZ: 18; HC: 24	SZ: 38.5 (14.8), range = 18–65); HC: 37.0 (13.1), range = 21–52	SZ: 12M;6F; HC: 15M:9F	Never medicated (7/18) or medication withdrawn for 12 days–2 years (11/18)	FDG PET	Verba lea an me tas
Schulz, 2002	SZ: 49	SZ: 32.6 (8.1), range = 20–51	49M:0F	Chronic (mean duration 10.25 [7.5] years); medication withdrawn for 3 weeks	[15O] water PET	Rest

LIMITATIONS OF EXISTING LITERATURE

There are several limitations of the existing literature that make it difficult to draw firm conclusions about the neural systems involved in late-life mental illness. First, as mentioned in the beginning of the chapter, in order to begin to disentangle the effects of aging and illness, a cross-sectional design should compare three groups of patients (early-onset young, early-onset old, and late-onset old) to two groups of healthy individuals (young and old). Although few studies have compared early- and late-onset older adults to each other, none of the reviewed studies contained all of these groups, so, for many of the neuroimaging findings, it is difficult to know if the magnitude of age-related change is greater than expected due to normal aging and also whether the effects are different depending on age of onset. This is especially important for studies investigating late-life depression and cognitive impairment, which have not studied the effects of age of onset with regards to longitudinal conversion to dementia. Ideally, longitudinal studies would follow early- and late-onset patients over many years to identify treatment response and monitor the associations between

depressive and cognitive symptoms. Some such studies are underway in the realm of unipolar depression, but are also needed for other disorders. Second, as can be seen from the above review, many of the neuroimaging abnormalities (e.g., reduced frontal volume, white matter hyperintensities, hypofunction of the prefrontal cortex) are common to multiple psychiatric disorders. Direct comparisons between groups of elderly patients with different disorders are rare, although at least one study did find differences between the structural abnormalities of late-life mood disorders compared to late-life psychotic disorders. Further work is needed to determine the specificity of late-life neuroimaging findings to particular disorders. A third limitation of the existing literature is that very little is known about neural systems in late-life anxiety disorders or geriatric patients with personality disorders. These gaps will need to be filled in order to fully understand the pathophysiology of late-life mental illness.

FUTURE DIRECTIONS

In addition to addressing the limitations above, future neuroimaging studies in geriatric mental illness will benefit from combining multiple imaging technologies. For example, to better elucidate the functional systems that are involved in late-life mental illness, results from metabolic and blood flow imaging techniques can be combined with mapping of white matter tracts using DTI. Similarly, combining the spatial resolution of fMRI with the temporal resolution of EEG or MEG is likely to yield sensitive measures of pathology. More work is also needed to investigate the functional and structural connectivity of brain networks associated with geriatric psychiatric disorders and how disruptions in these networks predict treatment outcomes and comorbidities or functional impairments. The future is also likely to bring an increased emphasis in the use of functional neuroimaging to examine predictors and mechanisms of treatment response. It is likely that metabolic, blood flow, blood oxygenation, and neurochemical imaging all will be brought to bear in these types of studies, which ultimately promise to help address difficulties in treating geriatric patients. One day, it may be possible to use neuroimaging measures to aid in diagnosis and guide treatment decisions, although much work is needed to establish the validity and reliability of such techniques before they can be put to clinical use. Studies that combine imaging information with genetic testing are becoming more common, and could also be useful to help bring individualized medicine to geriatric psychiatry. Neuroimaging techniques are powerful tools that have already provided important clues in the search for knowledge about etiology, course, and treatment of late-life mental disorders. In the coming years, advances in these techniques and their use in combination with each other and with genetic technologies will lead to even greater advances to the benefit of the growing numbers of elderly mentally ill.

REFERENCES

Aizenstein HJ, Khalaf A, Walker SE, Andreescu C. Magnetic resonance imaging predictors of treatment response in late-life depression. *J Geriatr Psychiatry Neurol*. 2014;27:24–32.

*Alexopoulos GS. Frontostriatal and limbic dysfunction in late-life depression. *Am J Geriatr Psychiatry*. 2002;10:687–695.

Alexopoulos GS, Murphy CF, Gunning-Dixon FM, et al. Event-related potentials in an emotional go/no-go task and remission of geriatric depression. *Neuroreport*. 2007;18:217–221.

Andreescu C, Tudorascu DL, Butters MA, et al. Resting state functional connectivity and treatment response in late-life depression. *Psychiatry Res*. 2013;214:313–321.

Ballmaier M, Narr KL, Toga AW, et al. Hippocampal morphology and distinguishing late-onset from early-onset elderly depression. *Am J Psychiatry*. 2008;165(2):229–237.

Barak Y, Aizenberg D, Mirecki I, Mazeh D, Achiron A. Very late-onset schizophrenia-like psychosis: clinical and imaging characteristics in comparison with elderly patients with schizophrenia. *J Nerv Ment Dis*. 2002;190:733–736.

Brambilla P, Harenski K, Nicoletti M, et al. Differential effects of age on brain gray matter in bipolar patients and healthy individuals. *Neuropsychobiology*. 2001;43:242–247.

Buxton RB. *Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques*. New York: Cambridge University Press; 2002.

Delaloye C, de Bilbao F, Moy G, et al. Neuroanatomical and neuropsychological features of euthymic patients with bipolar disorder. *Am J Geriatr Psychiatry*. 2009;17:1012–1021.

DeLisi LE. Regional brain volume change over the life-time course of schizophrenia. *J Psychiatr Res*. 1999;33:535–541.

Elbejjani M, Fuhrer R, Abrahamowicz M, et al. Hippocampal atrophy and subsequent depressive symptoms in older men and women: results from a 10-year prospective cohort. *Am J Epidemiol*. 2014;180:385–393.

Eyler LT, Kaup AR, Mirzakhanian HM, Jeste DV. Schizophrenia patients lack normal positive correlation between age and brain response during verbal learning. *Am J Geriatr Psychiatry*. 2009;17:43–55.

*Hahn C, Lim HK, Lee CU. Neuroimaging findings in late-onset schizophrenia and bipolar disorder. *J Geriatr Psychiatry Neurol*. 2014;27:56–62.

*Herrmann LL, Lemasurier M, Ebmeier KP. White matter hyperintensities in late life depression: a systematic review. *J Neurol Neurosurg Psychiatry*. 2008;79(6):619–624.

Kalayam B, Alexopoulos GS, Kindermann SS, Kakuma T, Brown GG, Young RC. P300 latency in geriatric depression. *Am J Psychiatry*. 1998;155:425–427.

Konarski JZ, Kennedy SH, McIntyre RS, Rafi-Tari S, Soczynska JK, Mayberg HS. Relationship between regional brain metabolism, illness severity and age in depressed subjects. *Psychiatry Res*. 2007;155:203–210.

Lloyd AJ, Ferrier IN, Barber R, Gholkar A, Young AH, O'Brien JT. Hippocampal volume change in depression: late- and early-onset illness compared. *Br J Psychiatry*. 2004;184:488–495.

Moorhead TW, McKirdy J, Sussmann JE, et al. Progressive gray matter loss in patients with bipolar disorder. *Biol Psychiatry*. 2007;62:894–900.

Nobuhara K, Okugawa G, Sugimoto T, et al. Frontal white matter anisotropy and symptom

severity of late-life depression: a magnetic resonance diffusion tensor imaging study. *J Neurol Neurosurg Psychiatry*. 2006;77:120–122.

Raz N, Lindenberger U, Rodrigue KM, et al. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cereb Cortex*. 2005;15:1676–1689.

Sachs-Ericsson N, Corsentino E, Moxley J, et al. A longitudinal study of differences in late- and early-onset geriatric depression: depressive symptoms and psychosocial, cognitive, and neurological functioning. *Aging Ment Health*. 2013;17:1–11.

*Sajatovic M, Forester BP, Gildengers A, Mulsant BH. Aging changes and medical complexity in late-life bipolar disorder: emerging research findings that may help advance care. *Neuropsychiatry (London)*. 2013;3:621–633.

Sheline YI, Sanghavi M, Mintun MA, Gado MH. Depression duration but not age predicts hippocampal volume loss in medically healthy women with recurrent major depression. *J Neurosci*. 1999;19:5034–5043.

Smith GS, Gunning-Dixon FM, Lotrich FE, Taylor WD, Evans JD. Translational research in late-life mood disorders: implications for future intervention and prevention research. *Neuropsychopharmacology*. 2007;32:1857–1875.

Smith GS, Reynolds CF 3rd, Pollock B, et al. Cerebral glucose metabolic response to combined total sleep deprivation and antidepressant treatment in geriatric depression. *Am J Psychiatry*. 1999;156:683–689.

Sullivan EV, Pfefferbaum A. Diffusion tensor imaging and aging. *Neurosci Biobehav Rev*. 2006;30:749–761.

Tadayonnejad R, Ajilore O. Brain network dysfunction in late-life depression: a literature review. *J Geriatr Psychiatry Neurol*. 2014;27:5–12.

Taylor WD, McQuoid DR, Payne ME, Zannas AS, MacFall JR, Steffens DC. Hippocampus atrophy and the longitudinal course of late-life depression. *Am J Geriatr Psychiatry*. 2014;22:1504–1512.

*Teodorczuk A, O'Brien JT, Firbank MJ, et al. White matter changes and late-life depressive symptoms: longitudinal study. *Br J Psychiatry*. 2007;191:212–217.

Truong W, Minuzzi L, Soares CN, et al. Changes in cortical thickness across the lifespan in major depressive disorder. *Psychiatry Res*. 2013;214:204–211.

57.2g Genetics of Late-Life Neurodegenerative Disorders

SEETHALAKSHMI RAMANATHAN, M.B.B.S., D.P.M., AND STEPHEN J. GLATT, PH.D.

Genetic studies of any complex condition, including late-life neurodegenerative disorders, attempt to identify the elemental biologic underpinnings of these disorders to facilitate their early identification, early intervention, and ultimately, prevention. In addition, the identification of genes that influence the development and/or progression of these illnesses may foster medicinal chemistry applications that target faulty DNA sequences or the altered RNA transcripts or proteins produced by such genes. Such discoveries could usher in the long-awaited era of "personalized or precision medicine" in which the treatment of any given

late-life neurodegenerative disorder can be tailored to individual patients based on their genetic and biologic background.

The search for genes influencing late-life neurodegenerative disorders falls under the broader umbrella of genetic epidemiology, which has been defined as a science that deals with etiology, distribution, and control of disease in groups of relatives and with inherited causes of disease in populations. Genetic epidemiologists examine the distribution of illness within families with the goal of finding genetic and environmental causes of illness. Genetic epidemiologic research on a particular disorder such as Alzheimer disease (AD) tends to follow a series of questions in a logical progression. This sequence proceeds as follows: First, we ask, "Is the phenotype familial?" or "Does it run in families?" which can be established by family studies. Second, "What is the relative magnitude of genetic and environmental contributions to the phenotype?" which can be established by twin studies. Third, "How is the phenotype transmitted from generation to generation?" which can be addressed by segregation analyses. Fourth, "If genes mediate this transmission, where are they located?" a question that may be answered by linkage studies. Fifth, "What are the specific genes and variants in those genes that influence risk for the phenotype?" which can be addressed by association studies, including genome-wide association studies (GWASs). Finally, "How do the specific genes functionally cause the illness?" which is addressed by mechanistic studies looking into gene and protein expression and further functional assays. These are difficult questions to answer for any trait, but particularly so for phenotypes as complex as late-life neurodegenerative disorders, which can often be confused or misconstrued with other medical, psychiatric, or substance use disorders. Fortunately, the wide variety, specificity, and power of available methods have allowed genetic researchers to begin to find meaningful answers to some of these questions for each late-life neurodegenerative disorder.

GENETIC METHODS AND PRINCIPLES

Family Studies

The first question that must be asked and answered when attempting to delineate the genetic and environmental components of a complex late-life neurodegenerative disorder is, "Does the phenotype run in families?" or, "Is this phenotype familial?" This question can be answered through the use of family studies. The basic design of the family study begins with the ascertainment of a group of subjects who are affected by the disorder (cases) and a comparable group of comparison subjects who do not have the disorder. Next, the biologic relatives of these index subjects, or "probands" are ascertained and evaluated for the presence of the illness, or in some cases, subthreshold forms of the illness. The rate of the disorder or

associated "spectrum conditions" among family members of affected probands is then compared to the rate of the disorder among family members of comparison probands to determine the familial risk or relative risk. If a late-life neurodegenerative disorder has a genetic etiology, then biologic relatives of cases should have a higher likelihood than relatives of comparison subjects of carrying the gene or genes that influenced illness in their relative, and thus they should be at greater risk for the illness themselves. In addition, the risk to relatives of cases should be correlated with their degree of relationship to the proband, or the amount of genes they share in common. First-degree relatives such as parents, siblings, and children, share 50 percent of their genes, on average, with the proband. Thus, first-degree relatives of cases should be at greater risk for the disorder than second-degree relatives (grandparents, uncles, aunts, nephews, nieces, and half-siblings), who share only 25 percent of their genes with the proband.

Twin Studies

Once a disorder has been established as familial, it becomes necessary to determine if that pattern is attributable to the inheritance of genes or to shared familial and other environmental factors. It is also important to quantify the contribution that genes make relative to that made by environmental factors, as this may encourage or discourage future molecular genetic studies; it may also influence the decisions made by individuals seeking genetic counseling and the usefulness of family information in identifying high-risk individuals to target for early intervention and prevention protocols. These questions can be answered by twin studies. In twin study designs, the ratio of concordant:discordant MZ twin pairs is compared to the ratio of concordant:discordant DZ twin pairs. If the concordance for a late-life neurodegenerative disorder is higher among MZ twin pairs than DZ twin pairs, this is a good indication that there is a genetic contribution to the disorder; if MZ and DZ twin pairs have approximately equal concordance rates, environmental factors are more strongly implicated. Concordance rates among twin pairs are then used to estimate the "heritability" of a disorder. Heritability refers to the ratio of genetic to phenotypic variances, or the proportion of variance in risk for a late-life neurodegenerative disorder that is accounted for by variability in genetic factors. A heritability of 1.0 (or 100 percent) indicates that all variability in the phenotype is due to genetic factors alone. In contrast, a heritability of zero attributes all phenotypic variation to environmental factors.

Segregation Analyses

Once twin and/or adoption studies demonstrate that the familial

transmission of a disorder is due, at least in part, to genetic factors, it becomes necessary to identify the manner in which those genetic factors exert their influence. The mode of transmission of a late-life neurodegenerative disorder through multiple affected families can be modeled statistically through the use of segregation analysis. Segregation analysis is a flexible procedure that can test various known and hypothesized modes of genetic transmission; the familiar mendelian model of single-gene transmission of a trait is only one of the many transmission mechanisms that can be tested. Mendelian diseases have recognizable inheritance patterns (autosomal dominant, autosomal recessive, and X-linked) and generally are caused by mutations in a single gene within a particular family. In contrast to Mendelian diseases, complex diseases display no clear inheritance pattern. Complex diseases are multifactorial in their etiology and are caused by an interacting web of genetic and environmental risk factors. The genetic risk factors for such complex diseases are generally referred to as *risk* or *susceptibility genes*, in which genetic variants that are common in the population act to increase risk of developing the disease. Unlike mutations in mendelian genes, susceptibility genes are neither necessary nor sufficient to cause disease.

Linkage Studies

Knowing that genetic factors are involved in the etiology of a late-life neurodegenerative disorder—and to what degree and in what manner—is essential for designing optimal molecular genetic studies to reveal the chromosomal location of the responsible risk genes. To identify regions of chromosomes that have a high likelihood of harboring risk genes for (especially mendelian forms of) late-life neurodegenerative disorders, linkage analysis is a highly appropriate strategy. Families are ascertained for linkage analysis through a proband affected with the disorder of interest. Each individual in the family is then genotyped as a series of DNA markers (not necessarily in genes) spaced evenly throughout the genome, and the cosegregation of these DNA markers with the illness is tracked in each pedigree. Evidence for cosegregation at each marker locus is summed across pedigrees to derive an index of the likelihood of the obtained patterns of marker-phenotype cosegregation given the sampled pedigree structures.

Association Studies

Once regions of certain chromosomes have been implicated from linkage analysis as harboring a risk gene for a disorder, the next step is to identify what specific gene is segregating through families to give rise to that linkage signal. A gene can be selected for such an analysis subsequent to linkage analysis as a means to follow up on evidence for increased genetic

similarity at a locus among affected individuals in a family (i.e., a "positional candidate gene" approach). Alternatively, specific genes can be examined in the absence of linkage information if there is some compelling physiologic reason to suspect that the gene influences risk for a given disorder (i.e., a "functional candidate gene" approach). For example, dopamine-system genes, such as receptors and transporters, are commonly examined as functional candidates for Parkinson disease (PD), based on the neural structures involved in its pathology and treatment. In contrast to linkage analysis, which uses random DNA markers as proxies for nearby risk genes, genetic association analysis is an appropriate method for determining if a particular gene variant has a direct effect on risk for an illness, or is in very tight linkage disequilibrium (LD) with such a gene.

Genome-Wide Association Studies

In the last few years, the completion of the human genome project and advancements in the genotyping of single nucleotide polymorphisms (SNPs) laid the foundation for a more broad examination of the human genome: GWASs. GWASs, as the name suggests, examine the entire genome for causal associations with specific traits. This broad approach accords GWAS more flexibility in exploring complex diseases, including late-onset Alzheimer disease (LOAD). GWAS techniques examine numerous markers distributed throughout the genome in order to understand the contribution to disease of various types of SNPs, including LD-based markers that are highly correlated with the causal allele, markers for missense mutations associated with mendelian diseases, and convenience-based markers that are chosen based on logistical conveniences such as cost.

Clearly such a broad examination has advantages as well as disadvantages. One advantage that GWAS has over other methods is its unbiased nature; however, the examination of the entire genome requires large samples to establish relationships with adequate power. Another unfortunate constraint of most current genome-wide association methods is that, unlike candidate genes which are often typed on functional polymorphisms, the DNA markers included in GWASs are often intergenic or otherwise do not constitute any recognized functional motifs. Researchers have been following and attempting to replicate the extensive list of SNPs identified by GWASs, and are now concluding that while these early "hits" are useful in understanding the pathology of various diseases, the added benefit of continued GWAS analysis is diminishing. This has led to forays into newer areas such as pooling summary statistics from multiple GWASs by meta-analysis, pooling individual data from independent GWASs by mega-analysis, and examining functional outcomes associated with implicated loci, including their identification as expression quantitative trait loci (eQTL), or their impact on protein interaction

networks.

Transcriptomics

Individual candidate-gene studies and GWASs have opened up a new catalog of risk genes, triggering new research into functional aspects of genes, including their potential as biomarkers. Transcriptomics, or gene expression analysis, measures the mRNA levels of all known human genes in a given sample simultaneously. In addition to mRNAs, transcriptomics also aims to understand noncoding RNAs and small RNAs, transcriptional structure of genes, splicing patterns, and other posttranscriptional modifications. These transcripts, whose expression is regulated by both genetic (i.e., DNA sequence) and environmental (i.e., chemical or biologic inputs transduced onto the genome *via* epigenetics) factors have both a trait-like quality and a state-like quality, and thus may better fulfill the definition of a biomarker than the DNA sequence variation alone (which, of course, is static and cannot reflect the state of disease or health as it changes through development). This search for biomarkers is not without its limitations, including bias (due to potential sample contamination), overfitting of statistical models due to overparameterization, and lack of statistical power. Nonetheless, transcriptomic studies have been particularly useful in identifying biomarkers that can predict various disorders.

MAJOR OR MILD NEUROCOGNITIVE DISORDER DUE TO ALZHEIMER DISEASE

AD is a progressive neurodegenerative disorder characterized by an insidious onset and progressive loss of memory and other cognitive functions. AD is the most common form of dementia in older adults. Two broad subtypes of ADs have been identified: an early-onset subtype, defined as occurring before age 65 years, accounts for no more than 5 percent of cases, and a late-onset subtype, defined as occurring on or after age 65 years, accounts for 95 percent or more of cases.

Early-onset AD has been characterized as a familial, autosomal-dominant genetic disease with causative mutations in several genes. Early family studies of early-onset AD revealed linkage to a locus on chromosome 21q, which was later identified as the gene for amyloid precursor protein (*APP*) which encodes the beta amyloid protein found in amyloid plaques. However, since linkage to chromosome 21 was not found in most families with early-onset AD, other potential loci were investigated. This led to the discovery of other gene mutations in presenilin 1 (*PSEN1*) on chromosome 14 and in presenilin 2 (*PSEN2*), based on homology to *PSEN1*, on chromosome 1. The net effect of the mutations in these genes is overproduction of the more pathogenic form of beta amyloid; that is, more beta amyloid 42 relative to beta amyloid 40.

The less complicated genetic nature of early-onset AD paved the way for a greater understanding of the genetically complex LOAD, which is sporadic or familial with no clear segregation patterns in families. AD has been estimated to be 60 to 80 percent heritable with influence from nonshared environmental factors. Thus far, *APOE* is the only gene that has been reproducibly and consistently identified as a risk for the disease, with just a single copy of *APOE ε4* allele increasing the risk of AD by approximately 300 percent. Recent studies have also identified that *APOE* affects age of developing dementia in a gene-dosage dependent manner. On the other hand, the *APOE ε2* allele has been shown to reduce risk for AD relative to the *ε3* allele. *APOE ε4*, however, does not account fully for the genetic variance in AD. Effects have been noted for no less than 40 candidate loci in LOAD; however, the effects of each variant are minimal. Some of these genes include the ones implicated in early-onset AD (*APP*, *PSEN1,* and *PSEN2*) and other tauopathies such as frontotemporal dementia, progranulin (*GRN*), and microtubule-associated protein tau (*MAPT*). A rare mutation in the triggering receptor expressed on myeloid cells (*TREM2*) was recently identified as a risk factor that can triple one's risk of developing AD, providing support to the role of neuroinflammation in AD. Among the variants, an arginine-to-histidine substitution at amino acid 47 (R47H) had the strongest association with AD, having an effect size similar to *APOE ε4*. Some of the others include bridging integrator 1 (*BIN1*), clusterin (*CLU*), complement component [3b/4b] receptor 1 (*CR1*), phosphatidylinositol-binding clathrin assembly protein (*PICALM*), siglec-3 (*CD33*), GRB2-associated binding protein 2 (*GAB2*). A meta-analysis of GWASs of clinicopathologic AD using data from 4,914 brain autopsies confirmed the association of *ABCA7*, *BIN1*, *CASS4*, *MEF2C*, and *PICALM* with neurofibrillary tangles and neuritic plaques, both of which are involved in the pathology of AD. Another recent meta-analysis of GWASs identified 20 genes that were statistically significantly involved. When these loci were explored in a pathway study, four functional modules became evident including ubiquination, endocytosis, immunity, and cholesterol metabolism. In addition to the role of the immune system in AD, genes involved in the insulin-like growth factor (ILG) and genes regulating the cholinergic and glutamate receptors have also been implicated in AD. Studies examining the role of genes in the behavioral and psychological symptoms associated with AD have implicated a number of genes including serotonergic genes (serotonin receptor [*HTTLPR*] and transporter [*SCL6*]), dopaminergic receptor and metabolic genes (dopamine receptor [*DRD1* and *DRD3*] and catechol-*O*-methyl transferase [*COMT*] genes), and nicotinic genes in psychotic symptoms.

Most transcriptomic studies in AD have focused on areas that demonstrate increased amyloid accumulation or the medial temporal region. Broadly, studies have noted dysfunction of genes that mediate two

broad biologic processes: calcium channel signaling and neuroinflammation. These have also been identified in peripheral blood-based markers, including blood mononuclear cells and leukocytes. Gene expression studies of the entorhinal cortex and hippocampus have identified mitochondrial dysfunction in AD.

MAJOR OR MILD NEUROCOGNITIVE DISORDER DUE TO PARKINSON DISEASE

PD is a neurodegenerative disease characterized clinically by a progressively worsening movement disorder and dementia in 20 to 30 percent of cases late in the course of disease. Similar to AD, PD largely follows a complex pattern of inheritance with only 5 to 10 percent aggregating in families and showing an autosomal-dominant or recessive inheritance. It is therefore not surprising that early family studies failed to find much support for a familial aggregation of the disorder, causing many to suspect that the etiology of PD was largely (if not exclusively) driven by environmental factors. However, later studies ultimately supported a genetic contribution to the etiology of the disorder. The largest family study has been the Mayo Clinic study, which established that biologic relatives of a patient with PD are at an approximately 70 percent elevated risk of themselves developing the disorder as compared to relatives of unaffected individuals. It appears that genes may only (or more strongly) contribute to the risk for early-onset PD (less than 66 years of age). Relatives of such early-onset PD patients are at a 2.6-fold increased risk of developing PD as compared to relatives of patients with onset of PD later than 66 years, who were not at an increased risk for the disorder. Segregation analyses further suggest that early-onset PD may be driven by an autosomal-recessive mechanism.

Late-onset PD (like AD), on the other hand, has a more complex etiology involving environmental factors and not excluding some major genes of additive or recessive effects. Linkage studies of PD have identified certain distinct loci as possibly harboring risk genes for the disorder in chromosomes 1, 2, 4, 5, and X. Numerous functional candidate genes have also been reliably implicated in the etiology of PD by association analysis. GWASs have further added to this list by identifying unsuspected candidate genes for PD, as well as other late-life neurodegenerative disorders. Genes encoding alpha-synuclein (*SNCA*, which accumulates in Lewy bodies), *MAPT* (which is a constituent of tangles in AD), and leucine-rich repeat kinase 2 (*LRRK2*) have also been repeatedly associated with PD. A recent meta-analysis of GWASs found that together *MAPT* and *SNCA* contribute to up to 25.6 percent of the population attributable risk for PD. Mutations in parkin (*Park2*), *PTEN-induced kinase 1* (PINK1) and *oncogene DJ1* have been associated with autosomal recessive parkinsonism, which is early onset and responsive to DOPA. However, considering the significant clinical role of the dopaminergic system in PD, other dopaminergic system

candidate genes, such as *DRD1* and *DRD2* which code for dopamine receptors, are not similarly implicated. Similarly, genes involved in other neurodegenerative disorders such as APOE have not been implicated. Another meta-analysis of genome-wide association data of 13,708 cases with PD and 9,522 comparison subjects identified 26 loci, six of which were novel candidate genes (*SIPA 1L2, INPP5F, MIR4697, GCH1, VPS13C*, and *DDRGK1*). The genes at these 26 risk loci meditate diverse biologic processes, but collectively they are enriched with genes that bind nucleotides, and genes that are involved in dopamine biosynthesis. Similar to GWAS in AD, these new findings are triggering new studies aimed at understanding the various functional aspects of the implicated genes. Transcriptomic studies involving substantia nigra pars compacta (SNpc) tissues have noted altered expression of genes involved in dopamine neurotransmission, mitochondrial function, protein degradation, and synapse dynamics.

MAJOR OR MILD NEUROCOGNITIVE DISORDER WITH LEWY BODIES (DEMENTIA WITH LEWY BODIES)

Dementia with Lewy bodies (DLB) is characterized by cognitive decline, parkinsonian signs and symptoms, fluctuations in cognition and attention, and visual hallucinations. Individuals with DLB can also have REM sleep disturbance, repeated falls, autonomic dysfunction, and a heightened sensitivity to neuroleptics. Based on the distribution and density of Lewy bodies, DLB can be categorized into three types: the brainstem-predominant type, the limbic (transitional) type, and the neocortical type (also known as diffuse Lewy body disease [DLBD]).

DLB has a great deal of clinical overlap with both AD and PD. DLB differs from AD in that the onset of hallucinations is much earlier in DLB. Similar to PD, individuals with DLB demonstrate parkinsonian signs; however, these motor manifestations are markedly more in PD as compared to DLB. DLB has been found to be largely sporadic and has been noted to be complex in nature, resulting from genetic and environmental interactions. Considering this sporadic nature, most genetic studies of DLB have been limited to family studies. Family studies have noted that DLB can be autosomal dominant or recessive. Reflecting the overlap in clinical signs, DLB shares a considerable genetic overlap with AD and PD. For example, the *APOE4*, synuclein genes (*SNCA* and *SNCB*) and *PSEN1* genes have been found to be risk factors in DLB.

MAJOR OR MILD FRONTOTEMPORAL NEUROCOGNITIVE DISORDER (PICK DISEASE AND OTHER FRONTOTEMPORAL DEMENTIAS)

Pick disease, or frontotemporal lobar degeneration, is a progressive dementia characterized clinically by prominent behavioral and language deficits and neuropathologic signs including atrophy in the frontal and

temporal lobes, *tau*-containing intraneuronal inclusion bodies (Pick bodies), and ballooned neurons. There are several clinical syndromes of frontotemporal lobar degeneration including frontotemporal dementia, semantic dementia, and progressive nonfluent aphasia. Other related neurodegenerative diseases with similar clinical features and *tau* neuropathology include corticobasal ganglionic degeneration, progressive supranuclear palsy, and motor neuron disease. Collectively, these neurodegenerative diseases are commonly referred to as "tauopathies."

Frontotemporal dementia has rarely if at all been evaluated within the context of family, twin, or segregation studies; thus, it is difficult to estimate the familial risk, heritability, and mode of inheritance of this group of disorders. However, within the context of several linkage studies, it has been shown that frontotemporal dementia may be inherited in an autosomal-dominant manner with age-dependent penetrance (i.e., risk increases with age). One of the most routinely identified loci linked to frontotemporal dementia has been chromosome 17q21. Approximately10 to 20 percent of familial cases are accounted for by mutations in the *MAPT* and *GRN* genes, which are located in this region. In addition to these, *C9ORF72*, genetically linked to chromosome 9p, has been recently identified as having a mutation that may singly account for the most number of cases (11.7 percent) of familial FTD. The association with 9p has been found by at least two GWASs, thus giving credence to the importance of 9p in FTD. While *MAPT* is associated with an increased deposition of phosphorylated tau, *GRN* and *C9ORF72* have been associated with deposition of the transactive response DNA binding protein of 43 kDa (*TDP-43*). GWAS has also identified associations with immune system processes and possibly lysosomal and autophagy pathways in FTD through 6p21.3, encompassing the *HLA* locus (immune system), and 11q14, encompassing *RAB38/CTSC* (transcripts of which are related to lysosomal biology).

MAJOR OR MILD VASCULAR NEUROCOGNITIVE DISORDER (VASCULAR DEMENTIA)

Vascular dementia (VaD) is the second most common form of dementia in older adults. VaD is caused by the effects of cerebrovascular disease and events (i.e., "stroke") that initiate a cascade leading to CNS damage and, ultimately, a dementia syndrome in some individuals. Evidence for the role of genetic factors in VaD (other than familial VaD) has largely been limited.

While familial forms of VaD are uncommon (accounting for less than 10 percent of cases of VaD), cerebral autosomal-dominant arteriopathy with subcortical infarcts and leucoencephalopathy (CADASIL) and hereditary cerebral hemorrhage with amyloidosis-Dutch type (HCHWA-D) are among the better studied genetic syndromes of cerebrovascular disease and dementia. CADASIL is a syndrome of subcortical small vessel disease with

lacunar strokes, migraine, and dementia. Discovery of linkage to chromosome 19q12 in two unrelated families and, subsequently, mutations in the *NOTCH3* gene explain the cause of most cases of CADASIL. *NOTCH3* is expressed in vascular smooth muscle cells and is important in cell proliferation and differentiation. FCAA, or familial cerebral amyloid angiopathy, is another heritable form of VaD, which is characterized by the deposition of amyloid on the vessel walls. A number of distinct forms of FCAA have been identified. Hereditary cerebral hemorrhage with amyloidosis-Dutch type, HCHWA-D, is an autosomal-dominant syndrome primarily of hemorrhagic strokes and dementia. Since extensive deposits of beta amyloid protein in brain vasculature had been found in HCHWA-D, linkage to the *APP* gene was tested for and was found in two unrelated families with HCHWA-D. A point mutation in the *APP* gene was found in two patients, providing evidence that *APP* gene mutations can cause HCHWA-D. Italian and Flemish types of HCHWA also are caused by point mutations in the *APP* gene and exhibit autosomal-dominant transmission. The Icelandic type of HCHWA, also with autosomal-dominant transmission, is caused by point mutation in the cysteine protease inhibitor cystatin C gene, while the British type of FCAA has been associated with mutations in the *ITM2B* gene (also called *BRI2*).

The genetic underpinnings of most cases of VaD however remain largely undetermined. This deficiency stems largely from the heterogeneity of the illness. VaD and AD frequently occur together as "mixed dementia." *APOE* gene variants are known to affect risk for cerebrovascular disease, and presenilins have been shown to interact with *NOTCH3*—suggesting mechanistic overlap if not the possibility of additional genetic overlap between AD and VaD, which may be driving the comorbidity. Studies examining candidate genes have noted inconsistent associations with genes coding for proteins involved in lipid metabolism, inflammation, and the renin–angiotensin system. A complex disease such as VaD warrants exploration into epistatic interactions. Studies have identified that epistatic interactions may act as a stronger risk factor for VaD. For example, the combined genotype of tumor necrosis factor (*TNF*α T-1031C) and heat shock 70-kDa protein 1A (*HSP70-1* A-110C) showed a much stronger odds ratio as compared to the individual genotypes. Similar synergistic relationships have been noted between transforming growth factor beta 1 (*TGF-β1*) and angiotensinogen (*AGT*) genes, and between family with sequence similarity 134 member B (*FAM134B* [rs10041159]) and tumor necrosis factor receptor superfamily member 19 (*TNFRSF19* [rs9317882]). A recent GWAS also noted an association with a variant near the androgen receptor (*AR*) gene on the X chromosome. While the mechanistic role of this association is unclear, this finding replicates past reports of associations with CAG repeat lengths on AR and VaD.

MAJOR OR MILD NEUROCOGNITIVE DISORDER DUE TO PRION DISEASE (CREUTZFELDT–JAKOB DISEASE AND OTHER TRANSMISSIBLE SPONGIFORM ENCEPHALOPATHIES)

Creutzfeldt–Jakob disease (CJD) is one of several neurodegenerative diseases, characterized by progressive dementia and ataxia, called *transmissible spongiform encephalopathies* (TSEs). TSEs are rare diseases that include CJD, kuru, Gerstmann–Sträussler syndrome, and fatal familial insomnia. The sporadic form of CJD is the most common TSE, with one to two incident cases per million people worldwide. The TSEs are unique in that they are inherited, sporadic, or acquired by infection. An abnormal isoform of the prion protein has been implicated as the causal agent in the TSEs. The prion protein exists in one of two states: PrP^{C}, for the normal cellular form of the protein, and PrP^{SC}, for the abnormal isoform isolated from the brains of individuals with TSE. The genetic epidemiologic study of this class of late-life neurodegenerative disorders has largely bypassed behavioral genetic approaches such as family and twin studies; thus, it has never been firmly established that these disorders have a familial or genetic basis. Data from Finland showed that 6 of 30 cases observed nationwide from 1975 to 1984 were familial cases, and all belonged to the same extended kindred. Further evaluation of this pedigree identified autosomal-dominant transmission of an early-onset long-duration form of the illness.

Despite the dearth of behavioral genetic data on CJD and related disorder in general, molecular genetic methods have been used often in the study of these disorders, largely due to their ease of application. Much of the focus of earlier studies had been on the *PRNP* gene, which encodes the prion precursor protein. In fact, most of the affected individuals in the Finnish family described above also carried a nonsynonymous mutation in codon 178 of *PRNP*, whereas none of 10 healthy controls carried the same variant. Since then *PRNP* (specifically an amino-acid polymorphism at codon 129 coding for methionine or valine) has been confirmed in a number of studies as a reliable risk factor for both variant or vCJD and sporadic or sCJD. Over the last 5 years, a number of other risk genes have been identified, although *PRNP* still remains the strongest universal factor. Variations in associations have been noted across populations for sCJD; SNPs at the *ZBTB38–RASA2* locus (a zinc finger transcriptional activator that binds methylated DNA) have been associated with CJD in UK populations, while an SNP in the *CHN2* gene has been found to be associated with vCJD in German population. Some of the other associations include myotubularin-related protein 7 gene (*MTMR7)* and neuronal PAS (per-ARNT-sim) domain-containing protein 2 gene (*NPAS2*), which have been identified to increase risk independently, and β-site APP cleaving enzyme 1 (*BACE1)* influence the risk of sCJD by interacting with PRNP gene.

MILD NEUROCOGNITIVE DISORDER

The DSM-5 introduced a new category of neurocognitive disorder—mild neurocognitive disorder (or mNCD). The introduction of mNCD in DSM-5 will likely help formalize a uniform definition which, in turn can help further our understanding of factors that drive progression to dementia. mNCD essentially refers to cognitive impairment, the severity of which does not amount to dementia. mNCD can be characterized based on the pathology into different subtypes. Two of the most popular subtypes include mNCD secondary to AD and mNCD secondary to vascular processes. As the titles indicate, the underlying pathogenic and genetic mechanisms are similar to the major neurocognitive disorders. Understanding the genetics of mNCD becomes critical for purposes of early intervention and prevention and also furthering personalized medicine. Thus far, most of this research has been limited to individuals with the amnestic type of mNCD and progression to AD. Thus, it is natural that most studies have focused on *APOE*. While individual studies produced varying results (likely driven by ancestral differences in genetic background and allele frequencies), a recent meta-analysis revealed that *APOE4* was associated with an increased relative risk (2.09) of conversion to AD. Similarly, studies have noted that the *ACT* and *CHRNA7* genes may modify risk of progression to dementia, with the former increasing risk and the latter being protective.

FUTURE DIRECTIONS

Over the last decade, significant strides have been made in understanding the genetics of late-life neurodegenerative disorders. Despite the initial successes, association studies remain plagued by limitations, including their propensity to produce false-positive results and their limited breadth in terms of identifying causal variants rather than markers. Regarding the former, genes identified as associated with a late-life neurodegenerative disorder in an initial study often overestimate the true effect size and subsequently fall victim to the winner's curse, wherein the same magnitude of an effect cannot be replicated. Regarding the latter, association methods have traditionally focused on one or, at most, a handful of genes at once, whereas linkage analysis constitutes a genome-wide survey of (relatively) unselected markers. Unfortunately, the prior probability of selecting the right candidate gene (out of ~25,000 human genes) and the right polymorphism (out of more than 11,000,000 in the human genome) for analysis is understandably remote. Most candidate genes for late-life neurodegenerative disorders have been targeted based on their expression within systems widely implicated in the disorder (e.g., amyloid plaques in AD and functional candidate genes in the dopamine neurotransmitter systems in PD). This approach has been critical in clarifying the nature of

dysfunction within these recognized candidate pathways; however, it has not been optimal for identifying additional novel risk factors outside of these systems. GWASs have attempted to fill some of these deficits. The result has been the elucidation of a more complex profile of gene involvement. The discovery of the complexity led to the identification of newer and advanced methods including imputation, which relies on the knowledge that certain gene sequences are usually coinherited. GWAS often identifies a large number of loci, some of which are noncoding. This has led to the identification of expression quantitative trait loci (eQTLs). eQTLs are regions in the genome that regulate nearby genes. In their attempts to integrate functionality with observations, researchers are also beginning to look into one of the limitations of GWASs—the tendency to filter out "near-misses." Preliminary work in this area has led to the identification of an "inflammatory subtype" of PD, which may have a greater involvement of genes associated with the immune system.

The multifactorial polygenic etiologies and heterogeneity of late-life neurodegenerative disorders have proved to be a huge challenge in the discovery of their underlying genetic bases. The multifactorial polygenic nature of these diseases dictates that the "signals" obtained in linkage and association studies will be numerous and of very low intensity, while the heterogeneity of these disorders increases the "noise" against which these already faint signals must be detected. Some of the most effective ways of combating these complexities are through the maximization of power, both by increasing sample size and by studying homogeneous groups of affected individuals. To identify the numerous genes with small effects on risk for late-life neurodegenerative disorders, various data-pooling strategies, such as combination of study-level data by meta-analysis or direct combination of primary data by mega-analysis, can be particularly effective. However, to overcome the obstacles introduced by heterogeneity, delineating and studying smaller more homogeneous subgroups of affected individuals may have the greatest beneficial effects on power. When used in tandem, the practices of pooling and splitting will yield maximal power for genetic studies, and hasten the discovery of the full compendium of risk genes for each heritable late-life neurodegenerative disorder.

Along the same lines, researchers are beginning to explore other approaches such as epigenomics, which has led to the study of DNA methylation and microRNA expression. As an example, in DLB, studies have noted an epigenetic regulation of the *SNCA* gene. The p.A53T mutation in *SNCA* can cause an overall imbalance in allelic expression by epigenetically silencing the mutant gene. Other promising avenues for elucidating risk genes include the analysis of gene–gene interactions as well as the aggregation of risk genes into polygenic scores which in aggregate account for a larger proportion of variance in susceptibility than any single variant in isolation. As an example, a genetic risk score, based on

ABCA7, *BIN1*, *CD2AP*, *CLU*, *CR1*, *MS4A4E*, *MS4A6A*, and *PICALM*, was noted to predict the rate of progression (a twofold increase in rate) to AD from mNCD. But, perhaps, one of the most interesting shifts in paradigm is the recommendation to examine neurodegenerative disorders from the bottom-up, that is, identifying common biologic pathways in neurodegenerative disorders that become individualized, leading to the unique clinical and pathologic characteristics. Recent GWASs have emphasized the biologic pathways that have subserved shared mechanisms between AD and PD such as apoptosis, autophagy, mitochondrial dysfunction, ubiquitin–proteosome system, oxidative DNA damage, inflammation, lipid metabolism, and vascular changes.

CLINICAL IMPLICATIONS

Since the first draft of the human genome was produced in 2001, scientists have touted the promise of genetic analysis for identifying risk factors for common diseases and the development of personalized (or, more recently, "precision") medicine. Yet, the vast potential of genetic studies to change the clinical practice of psychiatry is just beginning to be tapped. Recently, commercial genetic tests for (especially mendelian forms of) late-life neurodegenerative disorders have become available to individuals with a strong family history of illness (e.g., *APP*, *PSEN1*, and *PSEN2* for AD, *MAPT* for frontotemporal dementia and Pick disease, *PRNP* for CJD, and *PARK*s *1,2,6,7*, and *8* for PD). However, individuals in the general population cannot be tested for the possession of particular genotypes to determine their cumulative risk of developing a particular disorder. Similarly, over the past decade, progress has been made in understanding the factors that contribute to conversion to AD. However, despite these advances, the leads that have been generated from molecular genetic studies are simply not yet understood well enough.

Once genetic risk factors for these conditions become established and widely recognized, they can serve many purposes for early intervention and prevention efforts. For example, understanding risk factors involved in conversion to dementia could speed the initiation of appropriate treatments, which consequently may promote better prognoses through early intervention efforts. Needless to say, the most ideal situation would be preventing the development of cognitive impairment and perhaps, identification of individuals at high risk of developing cognitive impairment through the development of biomarkers will be useful. The operationalization of the definition of early cognitive impairment will clearly advance the investigation into early identification and prevention. Advances in the area of transcriptomics have also been encouraging researchers in their attempts to develop personalized interventions in terms of biomarker tests, including both diagnostic and prognostic tests, as well as understanding the functional biologic consequences of possessing

certain risk-conferring genetic variants. All of this information may someday be harnessed to identify the earliest biologic changes associated with a trajectory toward a late-life neurodegenerative disorder, and this in turn may help identify specific molecular targets around which the next wave of pharmacotherapies can be designed. Furthermore, the delineation of numerous, distinct etiologies for what are now considered unitary disorders (e.g., PD) may allow, in the future, for these disorders to become recognized as distinct and treated as such (e.g., *PARK1*-positive PD, *PARK2*-positive PD, etc.). Recent successes in identifying specific risk genes for these disorders have been encouraging. Replicating and refining these results, and translating them into meaningful clinical applications, signals exciting times ahead for research on the genetics of late-life neurodegenerative disorders.

ACKNOWLEDGMENTS

The authors would like to thank Matthew Weinerth and Totadri Dhimal for assistance in summarizing recent GWASs.

REFERENCES

Adams HH, de Bruijn RF, Hofman A, et al. Genetic risk of neurodegenerative diseases is associated with mild cognitive impairment and conversion to dementia. *Alzheimers Dement*. 2015;11(11):1277–1285.

Baizabal-Carvallo JF, Jankovic J. Parkinsonism, movement disorders and genetics in frontotemporal dementia. *Nat Rev Neurol*. 2016;12(3):175–185.

Beecham GW, Hamilton K, Naj AC, et al. Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. *PLoS Genet*. 2014;10(9):e1004606.

Bennett DA, Yu L, Yang J, Srivastava GP, Aubin C, De Jager PL. Epigenomics of Alzheimer's disease. *Transl Res*. 2015;165(1):200–220.

Bertram L. Alzheimer's genetics in the GWAS era: a continuing story of 'replications and refutations.' *Curr Neurol Neurosci Rep*. 2011;11(3):246–253.

Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet*. 2007;39(1):17–23.

Blacker D, Haines JL, Rodes L, et al. ApoE-4 and age at onset of Alzheimer's disease: the NIMH genetics initiative. *Neurology*. 1997;48(1):139–147.

Corder EH, Saunders AM, Risch NJ, et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat Genet*. 1994;7(2):180–184.

Croes EA, Alizadeh BZ, Bertoli-Avella AM, et al. Polymorphisms in the prion protein gene and in the doppel gene increase susceptibility for Creutzfeldt-Jakob disease. *Eur J Hum Genet*. 2004;12(5):389–394.

Ferrari R, Hernandez DG, Nalls MA, et al. Frontotemporal dementia and its subtypes: a genome-wide association study. *Lancet Neurol*. 2014;13(7):686–699.

International Parkinson Disease Genomics Consortium; Nalls MA, Plagnol V, Hernandez DG, et

al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet*. 2011;377(9766):641–649.

Joutel A, Corpechot C, Ducros A, et al. Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature*. 1996;383(6602):707–710.

Leblanc GG, Meschia JF, Stuss DT, Hachinski V. Genetics of vascular cognitive impairment: the opportunity and the challenges. *Stroke*. 2006;37(1):248–255.

Lee C, Kim Y. Complex genetic susceptibility to vascular dementia and an evidence for its underlying genetic factors associated with memory and associative learning. *Gene*. 2013;516(1):152–157.

Levy E, Carman MD, Fernandez-Madrid IJ, et al. Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. *Science*. 1990;248(4959):1124–1126.

Levy-Lahad E, Wasco W, Poorkaj P, et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science*. 1995;269(5226):973–977.

Lloyd SE, Mead S, Collinge J. Genetics of prion diseases. *Curr Opin Genet Dev*. 2013;23(3):345–351.

Meeus B, Theuns J, Van Broeckhoven C. The genetics of dementia with Lewy bodies: What are we missing? *Arch Neurol*. 2012;69(9):1113–1118.

Morton NE. *Outline of Genetic Epidemiology*. Basel: Karger; 1982.

Nelson PT, Wang WX, Partch AB, et al. Reassessment of risk genotypes (GRN, TMEM106B, and ABCC9 variants) associated with hippocampal sclerosis of aging pathology. *J Neuropathol Exp Neurol*. 2015;74(1):75–84.

Rogaev EI, Sherrington R, Rogaeva EA, et al. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature*. 1995;376(6543):775–778.

Saunders AM, Strittmatter WJ, Schmechel D, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology*. 1993;43(8):1467–1472.

Schrijvers EM, Schurmann B, Koudstaal PJ, et al. Genome-wide association study of vascular dementia. *Stroke*. 2012;43(2):315–319.

St. George-Hyslop PH, Tanzi RE, Polinsky RJ, et al. The genetic defect causing familial Alzheimer's disease maps on chromosome 21. *Science*. 1987;235(4791):885–890.

Tan EK. The role of common genetic risk variants in Parkinson disease. *Clin Genet*. 2007;72(5):387–393.

Tan EK, Khajavi M, Thornby JI, Nagamitsu S, Jankovic J, Ashizawa T. Variability and validity of polymorphism association studies in Parkinson's disease. *Neurology*. 2000;55(4):533–558.

Tanzi RE, Gusella JF, Watkins PC, et al. Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. *Science*. 1987;235(4791):880–884.

Tournier-Lasserve E, Joutel A, Melki J, et al. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy maps to chromosome 19q12. *Nat Genet*. 1993;3(3):256–259.

Van Broeckhoven C, Backhovens H, Cruts M, et al. Mapping of a gene predisposing to early-onset Alzheimer's disease to chromosome 14q24.3. *Nat Genet*. 1992;2(4):335–339.

Van Cauwenberghe C, Van Broeckhoven C, Sleegers K. The genetic landscape of Alzheimer disease: clinical implications and perspectives. *Genet Med*. 2015;18(5):421–430.

▲ 57.3 Psychiatric Disorders of Late Life

57.3a Assessment of Functioning

DAVID J. MOORE, PH.D., AND PARIYA L. FAZELI, PH.D.

INTRODUCTION

Older adults are the fastest growing segment of the population in the United States. By 2030, the number of Americans aged 65 and older will more than double to 72.1 million; roughly 1 in 5 Americans will be over 65. Given that older persons represent such a large part of the American population, it is increasingly important that clinicians and researchers in the field of psychiatry be prepared to work with and care for older adults. Older adults, like any other segment of the population, will require care for physical and mental health problems, and they may be at increased risk for impairment in both of these areas. Decline in everyday functioning may be associated with mental and physical health problems. Clinicians are increasingly asked to make decisions regarding patients' capacity to live independently, as well as to determine capacity to conduct other daily tasks such as managing health care and finances. This chapter defines and reviews domains of everyday functioning and discusses factors (e.g., neurocognitive impairment) that are likely to interfere with daily functioning in older adults. Additionally, this chapter discusses various assessment methods and specific clinical populations, and we make recommendations for future instrument development. The authors hope that this chapter will serve as a resource for clinicians and spur additional needed research on everyday functioning difficulties among older adults.

WHAT IS EVERYDAY FUNCTIONING?

Before discussing the various assessment modalities and other factors related to everyday functioning, a few terms must be defined. Quite simply, *everyday functioning* refers to all of the activities in which people engage in their daily lives. By this definition, *everyday functioning* is different for every person. The nature of one's daily activities and behaviors depends on numerous factors including culture, socioeconomic status, and the presence or absence of medical or psychiatric conditions. Despite the wide range of possible daily activities, certain basic behaviors are nearly universal. They range from basic *Activities of Daily Living (ADLs)* such as eating and grooming to more complex behaviors, often termed *Instrumental Activities of Daily Living (IADLs)*, such as managing finances and medications. Some scientists would confine assessment of functioning to physical behaviors, but we take a broader view. For instance,

the authors believe that social skills and quality of life (QoL), neither of which is based on physical ability, fit nicely under the umbrella of everyday functioning.

Even though many useful everyday functioning instruments exist, so far researchers have failed to reach a consensus on the instrument or instruments that would represent the "gold standard." Several factors may have contributed to this. First, the assessment of everyday functioning in older adults is important to a wide variety of disciplines, such as geriatrics, gerontology, occupational therapy, psychiatry, psychology, nursing, general practice, neurology, and psychiatry, each of which has its own particular focus. This diversity of focus is apparent in the literature as well. Because of the multitude of disciplines conducting research on this topic, different terms are often used when referring to very similar constructs. As a result, no unified search term for "everyday functioning" is available in library reference databases (e.g., MeSH) in contrast to such terms as "neuropsychology" or "cognition." The most widely used terms for everyday functioning are *ADL*, which refers to very basic skills needed to survive, and *IADL*, which covers higher-order tasks that represent a level of independence that may or may not be present among individuals with intact ADLs. These terms are discussed and defined in more detail below.

Many individuals experience declines in their ability to perform everyday tasks as they age. Some of these declines may be considered normal, whereas other impairments may result from a disease process. Declines in daily functioning can result from numerous factors including physical, medical, cognitive, and psychiatric factors. Daily functioning ability is so important that it is included as a diagnostic criterion in the determination of many psychiatric and cognitive disorders. For instance, a diagnosis of major neurocognitive disorder due to *Alzheimer Disease* (AD) according to the fifth edition of *Diagnostic and Statistical Manual for Mental Disorders* (DSM-5) requires cognitive deficits that "interfere with independence of everyday activities." Due to its typically late onset, many individuals who are diagnosed with AD are no longer working. The specific means used to determine "independence of everyday activities" are typically left to the discretion of the clinician making the diagnosis. Along these same lines, a diagnosis of mild neurocognitive disorder due to AD—thought to be a transition between normal cognitive functioning and dementia—requires that "cognitive deficits do not interfere with the capacity of independence in everyday activities." Again, the methods for determining whether daily living skills are intact are left to the clinician, although recent standardized methods for assessing everyday functioning in mild neurocognitive disorder have been employed.

DOMAINS OF EVERYDAY FUNCTIONING ASSESSMENT

Daily functioning tasks lie on a continuum that ranges from tasks that are

necessary to function independently (e.g., ability to get oneself to the toilet and use it) to tasks that are optional but common (e.g., driving a vehicle in the United States). Tasks falling in between these extremes, such as managing one's finances, are not critical to living independently, but inability to perform them requires strong environmental or personal supports such as a structured living facility or a committed family member or caregiver. Although we prefer not to classify daily functioning skills by category, several terms are used in the literature. These terms should be familiar to clinicians so that they are able to effectively communicate with other members of a multidisciplinary clinical team.

Activities of Daily Living

As discussed previously, some basic daily functions seem to cut across cultures. Termed ADLs, these are personal care behaviors and basic physical skills necessary in daily life. They include walking, grooming, bathing, and dressing.

Some ADLs are more physically based than others (Table 57.3a–1). They are sometimes simply referred to as physical functions. The assessment of physical functioning is often left to physiatrists, physical therapists, rehabilitation specialists, and general practice physicians. Often, these practitioners will use the *Functional Independence Measure* (FIM) (later described in the Clinician Rating of Everyday Functioning section) or a similar measure to describe the level of physical dependence. Perhaps the most basic form of everyday functioning assessment is determining whether an older individual has the physical capabilities to perform basic motor tasks. Motor-based assessments such as timed gait and grip strength may provide additional information in this area. For example, the Timed Up and Go Test (TUG) tests the time it takes to rise from an arm chair, walk 3 m, turn, walk back, and sit down, and has been shown to be well correlated with scores on scales of balance, gait speeds, and ADL, and is predictive of falls. Another measure of physical functions is the Short Physical Performance Battery (SPPB), which evaluates balance, gait, strength, and endurance by examining ability to stand with the feet together in side-by-side, semi-tandem, and tandem positions; time to walk 8 ft; and time to rise from a chair and sit back down 5 times. This test has been reported to predict mortality, disability, and nursing home admissions.

Table 57.3a–1.
Typical Everyday Functioning Domains by Activities of Daily Living (ADL), Instrumental Activities of Daily Living (IADL), and Others

ADLs	IADLs	Beyond IADLs
Physically based ADLs	Financial management	Social Functioning

Transferring
- ► Chair, bed, wheelchair
- ► Toilet
- ► Tub/shower

Walking
- ► Given distance
- ► Climbing stairs
- ► Uneven surfaces

General ADLs

Grooming/hygiene
Bathing
Toileting
Continence
- ► Bladder
- ► Bowel

Dressing
Eating/feeding

Medication management
Cooking
Laundry
Cleaning
Communication
- ► Written
- ► Spoken
- ► Reading
- ► Telephone

Shopping
Housekeeping
Gardening
Driving
Hobbies
Safety

Driving
Quality of Life
Multi-tasking

Other ADLs are required for autonomous functioning at home (e.g., grooming, bathing, dressing). Persons incapable of conducting these basic activities require significant assistance and are likely to live in a structured setting (e.g., nursing home) or receiving assistance in completing these tasks from a full-time caregiver. A list of common ADLs can be found in Table 57.3a–1.

Instrumental Activities of Daily Living

Assessments of IADL were developed to assess certain higher-level abilities that are required for autonomous functioning within society. Examples of IADLs include managing medications, balancing a checkbook, and preparing a meal (Table 57.3a–1). Assessments of IADLs are targeted at older persons who live independently and have little problem with basic ADLs but may need some assistance with higher-order tasks. Most assessments of everyday functioning either focus on ADLs or IADLs, and assessment approaches are discussed below.

Beyond IADLs

There are several abilities that go beyond IADLs but can nevertheless be important for higher levels of daily functioning. These include social functioning, QoL, and driving. Furthermore, everyday functioning abilities do not exist in isolation. Thus, a discussion of multi-tasking and what it might mean for everyday functioning in specific cases is also pertinent here.

Social Functioning. Ability to function socially can be important for many aspects of an older adult's life. For instance, social skills are essential to maintaining friendships and remaining actively engaged with other

persons. Social skills are also critical in many everyday situations such as articulating an error on a recent bill or discussing health care choices with a physician. Assessments of social skills may be particularly important in conditions known to impact these abilities, including dementia and schizophrenia.

Driving. Many older adults continue to drive in later life. Although they typically drive fewer miles per year than do younger drivers, they are at greater risk for accidents than all other age groups with the exception of teenagers. Risk for accidents seems to increase most noticeably after age 75. After adjusting for miles driven per year, those over age 75 have the greatest number of fatal crashes of any group.

Given that older adults are at significant risk for driving-related accidents and even death, it is important for clinicians to understand what factors put these individuals at risk so that they can appropriately intervene. Equally important is to avoid stigmatizing older adults who are still capable of safe driving. Decisions regarding the advisability of continued driving can be difficult, since driving often gives older persons a sense of independence and allows them to maintain meaningful social relationships that may be otherwise hampered. The final decision regarding continued driving typically rests with each state's Department of Motor Vehicles (DMV). While many states have adopted special provisions for older drivers, such as no mail renewal after a certain age, very few states require individuals to take an exam or a road test once a certain age is reached. Some states require clinicians to report patients diagnosed with dementia to the DMV (e.g., California). The DMV then evaluates the driver to determine if he or she may retain driving privileges.

Neuropsychological abilities play a large role in determining whether persons of any age should be allowed to continue driving; however, given that dementia is more prevalent in older adults, it is especially important in this population. Some cognitive domains appear to be more strongly associated with impaired driving performance. They include memory, executive functioning, and processing speed. Some evidence exists that cognitive training may help improve processing speed among older adults.

Interactions may exist between driving ability and medication use. Studies have shown that persons over 65 exposed to psychoactive medications, particularly cyclic antidepressants, have a significantly increased rate of involvement in motor vehicle collisions. Furthermore, depression, which is common among older adults can contribute to decreased responsiveness and poor concentration and lead to an increased risk for driving accidents.

Quality of Life. QoL, while not in itself a functional ability, is perhaps the most important factor contributing to successful daily functioning.

Certainly, improved QoL is an assumed treatment goal with any medical or psychological intervention. QoL can be viewed from two perspectives. The global perspective considers all aspects of a person's life; these assessments are typically both subjective and multidimensional. Alternatively, health-related QoL (HRQoL) focuses exclusively on the effects of health, illness, and treatment. HRQoL can be assessed generically or with reference to a specific disease. The most commonly used generic HRQoL instrument is the 36-Item Short-Form Health Survey. Although several HRQoL instruments are used with older adults, the authors are unaware of any specifically designed for this population. Physical functioning is often a part of HRQoL questionnaires and may unknowingly bias these instruments against older adults. A sampling of QoL and HRQoL instruments is shown in Table 57.3a–2. By nature, these instruments are primarily self-report, although some collateral-report measures are available. In 2011, an expert panel of the National Institute on Aging recommended that older adults, particularly those with multiple chronic conditions, complete a brief but comprehensive (i.e., assesses physical symptoms and mental health symptoms) composite screening measure such as the SF-8/SF-36 or the Patient Reported Outcomes Measurement Information System 29-item Health Profile (PROMIS-29). This recommendation was made to identify domains where additional follow-up measures may be warranted (e.g., depression, daily functioning).

Multi-Tasking. One problem for virtually all assessments of everyday functioning is that they are conducted in isolation; that is, each domain of functioning is assessed separately. It is known, however, that "multi-tasking" is the norm in most people's lives, and that parts of many everyday tasks are embedded within the performance of other tasks. For example, few people preparing to cook pasta will drop all other activities simply to sit and wait for the water to boil before placing the pasta in the water. Rather, most higher-functioning individuals will engage in some other behavior such as preparing some other part of the meal, calling a friend, or watering the plants. Attention typically shifts away from one daily task to another and must be reoriented in order for the original task to be completed. Few instruments, if any, have succeeded in capturing the multi-tasking nature of daily functioning. The few that have been developed are mentioned in the performance-based assessment section.

Table 57.3a–2.
Commonly Used Quality of Life (QoL) and Health-Related Quality-of-Life (HR-QoL) Measures for Older Adults

Quality of Life
Comprehensive Assessment and Referral Evaluation (CARE)
EASY-Care

Functional Assessment Inventory (FAI)
IOWA Multidimensional Functional Assessment Questionnaire (OMFAQ)
Perceived Well-being Scale (PWB)
Philadelphia Geriatric Centre Multilevel Assessment Instrument (PGCMAI)
Quality-of-Life Profile: Seniors Version (QoLPSV)
Health-Related Quality of Life (HR-QoL)
36-Item Short-Form Health Survey (SF-36)
12-Item Short-Form Health Survey (SF-12)
8-Item Short-Form Health Survey (SF-8)
EuroQoL Five-Dimension Questionnaire (EQ-5D)
Feeling Thermometer Medical Outcomes Study (MOS)-derived questions
Quality of Well-Being (QWB) Scale
Sickness Impact Profile (SIP)
Patient Reported Outcomes Measurement Information System 29-item Health Profile (PROMIS-29)

FACTORS INFLUENCING EVERYDAY FUNCTIONING

Given the complexity of most everyday functioning tasks, many factors may significantly affect performance. These include, but are not limited to, cognition, mood, physical health, culture and ethnicity, and the physical environment. Research on the relative impact of each of these factors is fairly limited, and each factor has shown varying levels of influence. A discussion of each of these factors and the evidence supporting its role in everyday functioning difficulties is provided below.

Cognition and Everyday Functioning

Of all the factors pertinent to daily functioning, the strongest evidence exists for cognitive impairment as a significant contributor to decline in many everyday functioning abilities. However, this literature discussing the relationship between cognition and everyday functioning is not as well developed and complete as one might expect. A recent comprehensive review of the cognitive correlates of functional status found that (1) the proportion of variance accounted for by cognition was modest, (2) some cognition domains had a stronger relationship to functional status than others, (3) cognitive impairments in executive control may be particularly strong correlates of impaired medical and financial decision-making, and (4) global cognitive screening instruments may be relatively strong predictors of functional status. As a rule, neuropsychological tests have undergone a more rigorous test development process, than have everyday functioning assessments, and consequently their reliability and validity are better established.

Many of the early researchers on functional abilities (e.g., Lawton, Brody) hypothesized that daily functioning skills would be lost in the opposite order from that in which they were acquired. Indeed, more complex daily tasks seem to be more susceptible to cognitive impairments than are basic tasks.

The role of neurocognitive impairment for daily functioning in older adults is not the same for all disorders. For instance, some individuals may have cognitive difficulties that have long affected their daily functioning (e.g., an older person with schizophrenia), whereas others may have experienced recent and sudden changes (e.g., an older person who recently suffered a stroke). Moreover, the profiles of neuropsychological impairment vary greatly among the disorders found in older adults, and these distinct cognitive profiles may lead to different declines in daily functioning abilities. For instance, memory impairments offer a unique set of difficulties for everyday functioning as compared to motor impairments. A person with declarative memory impairment may have difficulty remembering to take his medications, whereas a person with motor difficulties may remember to take his medications but have difficulty opening the pill bottle.

Mood and Everyday Functioning

Depressed mood and other psychiatric difficulties appear to contribute less to everyday functioning impairment than do cognitive difficulties. Obviously, persons who are severely depressed or experiencing florid psychotic symptoms may experience significant impairment of daily functioning; however, persons who are only mildly to moderately impacted by psychiatric symptoms appear to suffer less resultant functional disability than do persons with comparable impairments in cognition. This issue is discussed further in the "Clinical Populations and Everyday Functioning" section.

Physical Disability and Everyday Functioning

Persons who experience physical disability typically have relatively intact daily functioning with the appropriate assistive devices and compensatory strategies. For example, persons recently confined to a wheelchair may need special accommodations installed in their dwellings. On the whole, however, compensating for physical disability is usually easier than compensating for cognitive difficulties. Several methods for assessing physical functioning are found in the "Assessment of Everyday Functioning" section.

Culture and Ethnicity in Everyday Functioning

Ethnicity and culture can play a substantial role in the types of everyday activities in which people engage as well as in the support that they receive in accomplishing those activities. Certainly individualistic cultures differ from collectivist cultures in the amount of support that each offers. For example, Latinos and many Asian cultures have, comparatively speaking, a collectivist approach to functional impairment in family members.

Households are frequently multigenerational, encompassing parents, grandparents, and other older relatives as well as the householders and their children. Hence, families in these cultures often provide greater support to older family members in carrying out daily tasks than do persons from families with individualistic cultural background (e.g., Caucasian Americans). Latino and Asian families, by virtue of their family-oriented support system, may be less likely to seek treatment for an older family member who is experiencing decline and may prefer to handle the matter within the family. Another consideration is the shame associated in some cultures (e.g., Japanese Americans) with acknowledging mental health difficulties in oneself or a family member. In such cases, persons may try to convey mental or cognitive difficulties by presenting with physical complaints (e.g., referring to a decline in cognitive abilities as a headache).

Given such cultural considerations, using everyday functioning assessment instruments designed primarily for English-speaking, Caucasian persons with individuals of other cultural backgrounds or who do not speak English as a first language can be problematic. Persons from different cultures may perform poorly on assessments of a daily functioning task for cultural reasons unrelated to functional capacity. For example, in one study, Black and Latino subjects over 55 years of age who had recently been prescribed an antidepressant attributed intentional nonadherence to concern about side effects and stigma; they also attributed less importance to antidepressants than other medications. In some cases, everyday functioning instruments have been translated and culturally modified for use with other, non–English-speaking populations, for example, Chinese elders as shown in Table 57.3a–3; however, there is clearly a need for additional translations, adaptations, and culturally appropriate modifications.

Environment and Everyday Functioning

An older person's physical environment also can negatively affect everyday functioning abilities. Persons with limited financial resources are often living in physical environments that exacerbate their everyday functioning difficulties. For instance, persons with limited ability to climb stairs may live in apartments on upper floors of buildings without elevators, or they may not be able to afford to have bars installed over the toilet to assist with transferring. Likewise, a person of lower socioeconomic status (SES) may not to be able to afford to hire a health care assistant (e.g., for help with managing medications).

Table 57.3a–3.
Multidomain Performance-Based Batteries of Everyday Functioning

Battery Name	Brief Description	Site	Scoring	Admin. Time (Min.)	Notes
Dementia (Primarily Alzheimer Disease–AD, some dementia NOS)					
Activities of Daily Living Situational Test (AST) (Skurla, 1988)	4 subscales; 38 individual items. Item examples include making instant coffee, and finding and calling a phone number.	Clinic	Items scored from 0–4. Total score ranges from 0–152.	15–20 min.	
Cognitive Performance Test (CPT) (Burns, 1994)	6 tasks; ratings of specific behaviors are used to place individuals into 1 of 6 levels of disability. Task examples include getting dressed, and preparing a piece of toast.	Clinic	Tasks scored from 1–6. Total score ranges from 6–33.	45 min.	Requires t and spe equipm Availabl specific subpop (Elderly Israelis and wit Dement
Direct Assessment of Functional Abilities (DAFA) (Karagiozis, 1998)	10 domains; Designed to measure Pfeffer Functional Activities Questionnaire (PFAQ) IADL items. Examples include writing a check and making a sandwich.	Clinic	Items are scored from 0–3.	45–90 min.	The DAFA been translat adapted use witl Kong Ch Elders.
Direct Assessment of Functional Status (DAFS) (Loewenstein, 1989)	7 subscales; 53 individual items. Item examples include balancing checkbook and writing out a check.	Clinic	Items scored from 0–2. Total score ranges from 0–106.	25–40 min.	DAFS has a been us with otl geriatri samples older schizop patients

Functional Performance Measure (FPM) (Carswell, 1995)	3 subscales; 8 subsections; total of 40 items. Participants are videotaped doing tasks and then the videotapes are rated. Item examples include washing and drying hands, and putting on and taking off sweater.	Clinic	Each sub-section scored from 0–50. Total score ranges from 0–400.	<20 min.	
Refined ADL Assessment Scale (RADL) (Tappen, 1994)	5 domains; 14 individual items. Item examples include washing hands and brushing teeth.	Clinic	Tasks are broken into component parts and scored from 1–6.	5–60 min. for eating subtask Total >90 min.	
Structured Assessment of Independent Living Skills (SAILS) (Mahurin, 1991)	10 domains; 50 individual items. Item examples include zipping a jacket, and using a key.	Clinic	Items scored from 0–3. Total score ranges from 0–150. Both time and accuracy contribute to the score.	60 min.	An abbrev version SAILS is availabl Eating a money subscal validate older M America
Texas Functional Living Scale (TFLS) (Cullum, 2001)	5 subscales; 21 individual items. Item examples include making change, and remembering to take medications.	Clinic	Items scored from 0–5. Total score ranges from 0–52.	15–20 min.	
Bay Area Functional Performance Evaluation (BaFPE) (Houston,	2 general domains; 5 tasks in the first domain, 7 items in the second. Item	Clinic	Rated from 1–4.	>45 min.	This inforr is basec the revi BaFPE.

1989)	examples include filling out a bank deposit slip and drawing a person.				
Older Psychiatric Patients (Primarily schizophrenia, some mixed psychiatric groups)					
Kohlman Evaluation of Living Skills (KELS) (Thomson, 1992)	5 subscales; 17 areas assessed. Item examples include budgeting money, and recognition of dangerous situations.	Clinic	Item scores range from 0–1; ½ points are possible.	20–30 min.	Test sold b publish KELS is a combin of self-r NP, and perform based t
Performance Assessment of Self-Care Skills (PASS) (Rogers, 1984; Holm, 1999)	In the home version, participants are observed conducting various tasks primarily with their own tools and props. Example items include sweeping and mailing bills.	Home	Most recent version is scored on a 4-point ordinal scale ranging from 0–3. Higher scores indicate better performance.	>90 min.	Contains b physica and IAD section
UCSD Performance-based Skills Assessment (UPSA) (Patterson, 2001)	5 subscales; item examples include shopping for items in mock grocery store and planning a recreational activity.	Clinic	Each subscale transformed to 0–20 score. Total score ranges from 0–100.	30 min.	
Mixed (Psychiatric and Dementia)					
Independent Living Scales (ILS) (Loeb, 1996)	7 subscales; 68 individual items. Example items include how to use a telephone book, and writing a check.	Clinic	Items scored from 0–2.	45 min.	Test sold b publish Combin of NP a perform based t ILS has used wi schizop patient

					older ac
Occupational Therapy Evaluation of Performance and Support (OTEPS) (Nadler, 1993)	4 domains; 42 individual items. Item examples include making a sandwich and identifying medication dosage from a bottle.	Clinic	Items scored from 0–2. Total score ranges from 0–84.	Not provided	Also know OTAPS.
Performance Test of Activities of Daily Living (PADL) (Kuriansky, 1976)	16 individual tasks; item examples include lifting food onto spoon and to mouth, combing hair, and drinking from cup.	Clinic	Scored 0 (completes task with minimal assistance) to 1 (unable to complete on own).	Not provided	
St. George Hospital Memory Disorders Clinic Occupational Therapy Assessment Scale (OTAS) (Fairbrother, 1997)	8 domains; 30 individual items. Example items include using a telephone book, and offering correct money in a simulated shopping activity.	Home	Items scored from 0–3. Total score ranges from 0–90. Score is based on time and the number of prompts.	30–40 min.	
Older Healthy Individuals Activities of Daily Living Test (ADL-T) (Warren, 1989)	4 tasks; item examples include boiling an egg and writing a check.	Home and Clinic	Scores ranged from 0–16.	>90 min.	
Day out Task (DOT; Schmitter-Edgecombe et al., 2012)	8 subtasks; laboratory apartment; participants told to imagine they are planning for a day out (beginning with meeting a friend at	Laboratory	Accuracy of subtasks (i.e., complete/efficient; complete/inefficient; incomplete/inaccurate; never attempted), overall task accuracy, and task sequencing.	No specific time, but average completion times between 10–14 min.	Used in M healthy adult cc

	museum at 10 AM and later traveling to friends house for dinner), the 8 tasks must be completed to prepare for the day; participants are told to multitask.				
Everyday Problems Test (EPT) (Willis, 1993)	7 domains; 84 individual items. Participants are shown printed material of everyday items. Examples include ordering merchandise from a catalog and determining amount to pay from phone bill.	Clinic	Scoring available in published test.	<60 min.	Test sold b publish There is al Everyda Problen for Cog Challen Elderly (EPCCE)
Observed Tasks of Daily Living (OTDL) (Diehl, 1995)	3 domains; 31 individual items (21 useable items). Example items include baking a cake in a microwave oven and loading a pill reminder.	Home	Scoring available from author.	90 min.	

Note: Tests preceded by a < or a > indicate approximation of administration time based on the text provided. Readers are encouraged to examine the literature for additional studies pertinent to their clinical or research question that may apply their candidate measures.

From Moore DJ, Palmer BW, Patterson TL, Jeste DV. A review of performance-based measures of functional living skills. *J Psychiatr Res*. 2007;41(1-2):97–118.

ASSESSMENT OF EVERYDAY FUNCTIONING

Standardized assessment instruments and procedures have been developed to evaluate everyday functioning. Functional assessments typically aim to determine the level of ability or disability in a given domain of daily functioning. Depending on the individual being evaluated, an assessment may focus on basic ADLs (e.g., bathing or grooming), physically-based ADLs like mobility (e.g., ability to walk without assistance), or IADLs (e.g., cooking, financial management). Table 57.3a–1 shows some of the domains typically assessed.

Some assessments focus on capacity (what the person can do), while others focus on performance (what the person actually does). Glass describes three "tenses" of functioning: (1) a *hypothetical* tense representing the capacity of the person, (2) an *experimental* tense defined as what the person is capable of doing, and (3) *enacted* tense which represents what the patient actually does. Furthermore, Glass describes common errors in classifying older adults' functional status. The first occurs when persons provide a low self-report of functioning but actually function at a higher level; these persons are "overachievers," and their reports as compared to behavior represent a discrepancy between the hypothetical and the enacted tense. The opposite situation also occurs: some persons report high functioning but engage in few activities that would qualify for that designation. These persons can be termed "underachievers" and also show a discrepancy between the hypothetical and the enacted tense.

When interpreting functional assessment data, the clinician or researcher must take into account whether the person being assessed has ever done a given task before. For example, performing a financial assessment with a person who has never managed her own finances has little to no meaning.

There are many reasons for conducting a functional assessment. In many cases, functional assessments are used to get a "snapshot" of an individual's current level of functioning. Functional assessments are rarely done as part of routine care; rather, they are often conducted following some acute event (e.g., stroke) or referral complaint (e.g., patient is having difficulty managing medications). An evaluation in this context serves as a baseline for future evaluations and, ideally, would be sensitive to improvement or worsening of abilities over time. Functional strengths and weaknesses can be identified, and interventions can aim to improve areas of disability and utilize areas of intact functioning to compensate for deficits. In the optimal situation, improvements would be observed following a treatment intervention to improve the targeted area of functional impairment. It is important to place particular emphasis on those abilities that are critical for functioning within a given environment. For instance, medication management may not be important for an individual living in a structured living facility that administers medications,

but bathing, grooming, and ability to transfer remain important. In this way, a functional assessment can be used as a tool for discharge planning and estimating the level of functional assistance that a particular individual may need.

Often a person receiving a functional assessment has significant pre-existing cognitive or physical disability, but persons with more subtle difficulties may be even more appropriate candidates for a functional assessment. For example, a person with AD is likely to have many functional impairments, and decisions regarding housing placement and the level of care may have been determined prior to the administration of a functional assessment. On the other hand, a person who is having mild cognitive difficulties, and who evidenced impairments on tests of arithmetic and memory, may be a perfect candidate for a financial capacity assessment if he or she is still managing personal finances. The authors argue that functional assessments should be used at earlier stages of clinical care to facilitate gradual life changes (e.g., enlisting a family member to help with finances), rather than wait for a radical upheaval of all aspects of daily life at once.

Everyday functioning assessments are not just for individual-level clinical purposes. These assessments can be used with representative samples of clinical populations to draw general demographic conclusions. For instance, research using the Social Skills Performance Assessment (SSPA) has shown that older schizophrenia patients have difficulties with social skills. As a result, Cognitive-Behavioral Social Skills Training was developed and has targeted these problems among persons with schizophrenia and improved the functioning of this group.

There are several different methods for assessing everyday functioning. The most common are clinical interview, self- or collateral report, clinician ratings of behavior, laboratory performance-based measures, and direct observation. Several reviews in the literature provide additional information about the everyday functioning assessment instruments designed for older adults.

Each assessment method has inherent strengths and weaknesses (Table 57.3a–4). For example, some types of instruments are easily administered, but may not be reliable (e.g., self-report), whereas others may provide more reliable data but are not feasible to administer routinely (e.g., direct observation). Choosing the type of functional assessment to use is largely driven by the referral question, available resources, and the person's particular needs. In the sections below, common functional assessment methods are reviewed, and some of the instruments frequently used with each type are discussed.

Clinical Interview of Everyday Functioning

Time is often at a premium when psychiatrists and other physicians

conduct clinical interviews. Despite this, a detailed clinical review of a patient's daily functioning can be extremely helpful in establishing current levels of functioning as well as for treatment planning. Understandably, a thorough assessment of everyday functioning cannot be conducted at every clinical visit; however, psychiatrists can make referrals to other disciplines (e.g., physical and occupational therapy, clinical psychology, neuropsychology) that may be able to spend more time both conducting a clinical interview as well as other assessment modalities described below.

Certain elements are essential to a thorough clinical interview designed to assess a person's daily functioning. First, it is important to get a sense of the persons' daily activities. Open-ended questions such as, "Can you describe a typical day for me?" are often helpful. This puts the onus on the patient to elaborate his or her daily activities. Queries should be made about difficulties with ADLs and IADLs. All assessments should include questions about other important domains such as social functioning, driving, and safety. Table 57.3a–5 lists some of the critical elements of an everyday functioning interview. Similarly, follow-up questions pertinent to any daily functioning domain of concern should be phrased in ways that give the patient the opportunity to elaborate any possible concerns or difficulties. It is sometime helpful to disarm patients by saying things like, "All persons have difficulties managing their medications. Can you tell me about the last time that you had difficulties managing yours?"

It is also important to elicit patients' histories with particular daily tasks. For instance, older men may have little experience cooking and preparing meals, while older women may have little experience managing finances. If a patient has recently lost a significant other or a support system that used to perform a particular task, poor performance may not represent a decline in daily functioning ability, but rather a lack of experience. Some assessment performances, however, may represent declines, and it is useful to pose questions such as the following: Was the onset of these difficulties gradual or sudden? Has there been continued decline? Was the decline associated with any other recent change (e.g., medical problem, change in medications)?

Table 57.3a–4.
Strengths and Weaknesses of Various Functional Assessment Modalities

	Cost	Ease of Use	Reliability	Validity	Predictive
Self-report	Inexpensive	Easy	High	Weak	Weak
Collateral report	Inexpensive	Easy	Questionable	Questionable	Questionable
Clinician-rating	Moderate	Moderate	Moderate	Moderate	Moderate
Performance-based	Moderate	Moderate	High	Strong	High
Direct	Expensive	Difficult	High	Strong	High

Table 57.3a–5.
Topics to Cover in Clinical Interview of Everyday Functioning

For All Older Adults

- Current daily activities?
- Difficulties with:
  - Physical demands (e.g., climbing stairs, walking)
  - ADLs (e.g., bathing, dressing, feeding)?
  - IADLs (e.g., cooking, medication/financial management, driving)?
  - Social skills?
  - Safety (e.g., two-story house in frail person)?
  - Decision making?

Additional questions for persons with difficulties

- Experience with daily activities of difficulty?
- Onset (e.g., gradual/discrete)?
- Progression (e.g., continue to decline)?
- Decline associated with critical event (medical, change in medication, grief)?
- Availability, reliability, commitment of support declines in daily functioning?
- Concrete plan for dealing with daily functioning difficulties?

Among patients exhibiting daily functioning difficulties, it is important to find out whether they have support for daily tasks. Clinicians should assess the availability, commitment, and feasibility of the proposed support system. Where possible, clinicians should get consent from the patient to follow up with the identified support person or network to determine if the person or people identified are truly capable of providing support in performing daily tasks.

Self-Report of Everyday Functioning

Self-report is the most widely used modality for assessing problems with daily functioning. Self-reports are limited by the knowledge and objectivity of the respondent. Although self-report is subject to inaccuracies, particularly overestimation of functional abilities, it has several advantages including cost effectiveness, ease of administration, and portability. Furthermore, in persons without significant psychopathology (e.g., depression, personality disorders) or health problems (e.g., chronic pain, fatigue), self-endorsement of functional difficulties is generally held to equate to actual difficulties; however, the opposite does not always prove to be true. That is, persons often fail to report, or minimize, difficulties that they may be having. Among older persons with mild neuropsychological impairment or dementia, self-reports of functional status may have little to no utility. Even among older individuals without cognitive problems, overestimation of abilities is common.

Another disadvantage of most self-report measures is that they are not

particularly sensitive to change over time. Individuals tend to hold fairly constant views of themselves regardless of recent personal events and tend not to report changes in daily functioning abilities until they become drastic. Self-report measures appear to elicit more accurate data and correlate more robustly with performance-based measures when queries are specific and pertinent to the individual's difficulties (e.g., about stair climbing) rather than abstract (e.g., about mobility in general).

Self-report data concerning functional abilities can be collected in several different ways, including structured, semi-structured, or unstructured interviews. Interviews allow clinicians to hear, in the patient's own words, the difficulties that he or she is having. Clinicians can then ask follow-up questions that help specify the particular difficulty or difficulties. This is the approach most commonly used by general practitioners. Where indicated, problem areas can be addressed with more formal assessments.

Another way of assessing functional status is through the use of standardized self-report questionnaires. These instruments vary in their levels of reported reliability and validity. The outcome is generally a score that represents the level of functional impairment. Some of the most widely used instruments of self-reported daily functioning are the Katz ADL scale, the Lawton–Brody IADL scale, and the Patient's Assessment of Own Functioning Inventory (PAOFI). Additional self-report assessment instruments are listed in Table 57.3a–6.

Katz ADL Index. The Katz ADL Index uses a 3-point Likert-type scale to assess ability to bathe, dress, toilet, transfer, feed oneself, and maintain bowel and bladder continence. It was designed for use primarily with older hospitalized patients who function at lower levels, and hence it is unlikely to identify functioning difficulties among most community-dwelling outpatients. Another limitation is the scale's hierarchical structure. Functional difficulties are assumed to accrue in this order: bathing, dressing, toileting, transferring, continence, and feeding. Research has shown that alternative hierarchies of functions may apply equally as well.

Lawton–Brody Instrumental Activities of Daily Living. The Lawton–Brody IADL scale is another structured questionnaire that focuses on higher-level domains needed for independent functioning at home and in the community. The scale covers eight domains: telephone use, shopping, food preparation, housekeeping, laundry, transportation, medication management, and financial management. This scale is often used in conjunction with the Katz scale to provide indicators of both basic ADLs and IADLs. Although the questionnaire is often used in self-report mode, the clinician or caregiver can also complete it.

Table 57.3a–6.

Self-Report Measures of Everyday Functioning

Self-Report measures of Functional Status
Lawton–Brody IADL
Katz ADL
Functional Status Questionnaire
Health Assessment Questionnaire
World Health Organization Questionnaire
Short-Form Health Survey (SF-36)
Sickness Impact Profile (SIP)
Functional Disability Index Score (FDIS)
Older American Resources and Services (OARS)
Patients Assessment of Own Functioning Inventory (PAOFI)
Rapid Disability Rating Scale–2 (RDRS-2)
Questionnaire on Clinical and Preclinical Disability (QCPD)
Unmet Needs Indices (UNI)
Identification of Seniors at Risk (ISR)
Instrumental Activities of Daily Living Compensation (IADL-C) Scale
Everyday Competence Questionnaire (ECQ)

Patient's Assessment of Own Functioning Inventory. Another self-report instrument of everyday functioning is the PAOFI. The PAOFI is a 41-item questionnaire that is widely used in conjunction with neuropsychological testing. Respondents report the frequency with which they have had difficulties with memory, language and communication, manual dexterity, sense-perception, higher-level cognitive and intellectual functions, and work and recreation. The main focus is on cognitive symptoms, and the scale is often used in conjunction with ADL and IADL questionnaires such as those described above.

Collateral Report of Everyday Functioning

Reports from caregivers, family members, or other persons who are close to the individual with everyday functioning difficulties can be extremely valuable. In situations where the collateral reporter knows the patient well, is capable of providing an accurate report, and is able to maintain objectivity, collateral reports can be one of the best ways of gathering information regarding a patient's functional status. Accurate collateral reporters provide the same types of information as a self-report assessment, but they do so from the perspective of a direct observer. Unlike clinical direct observers, collateral reporters are typically part of the individual's daily routine, which obviates some of the problems common to traditional direct observation (see below). Unfortunately, collateral reports also have limitations. Close family members or friends may have reasons for distorting their reports. For example, they may minimize the level of functional disability in an effort to protect the patient from unpleasant but

inevitable transitions, or they may overemphasize the level of impairment in order to obtain medical or psychosocial assistance for the patient more quickly. The collateral reporters may also have cognitive or psychiatric impairments of their own that limit their ability to render accurate judgments. Finally, in many cases, older patients are socially isolated such that a knowledgeable informant is unavailable. Where collateral reports are available, the clinician must decide how likely they are to be accurate. Many of the self-report measures listed in Table 57.3a–6 also have collateral-report versions. Additionally, the Physical Self-Maintenance Scale (PSMS) was specifically designed for use by a person with significant knowledge of the skills of the patient, and the Specific Levels of Functioning (SLOF) was designed for caretakers (detailed below).

Specific Levels of Functioning. Another measure is the SLOF, which is a 43-item caretaker report of a patient's behavior and functioning in the following domains: physical functioning (e.g., vision, hearing), personal care (e.g., toileting, eating, grooming), interpersonal relationships (e.g., initiating, accepting, and maintaining social contacts; effectively communicating), social acceptability (e.g., verbal and physical abuse, repetitive behaviors), activities (e.g., shopping, using telephone, paying bills, use of leisure time, use of public transportation), and work skills (e.g., employable skills, level of supervision, punctuality). The SLOF rates the amount of assistance that the patient requires to perform real-world tasks (personal care, activities), and it rates levels of physical functioning and the frequency of certain behaviors (interpersonal relationships, social acceptability, work skills). An important aspect of collateral reports such as the SLOF is how well the caretaker or informant knows the patient. Performance-based measures of everyday functioning discussed below have been shown to predict real-world outcomes as determined by the SLOF.

Clinician Rating of Everyday Functioning

Another common functional assessment approach is the clinician rating. Such an assessment is typically conducted in a clinic or hospital and can be administered by a physician, nurse, occupational or physical therapist, or a specifically trained rater. In contrast to self- or collateral reports, the clinician rating has the advantage that the clinician is generally familiar with specific daily functioning difficulties and the particular behaviors that may put individuals at risk for poor daily functioning. Many clinician-rated instruments have reported good interrater reliability, although it is likely that experienced clinicians will produce more accurate ratings and classifications. These instruments can measure both actual behavior and capacity. For instance, some clinician instruments measure what they see patients actually doing and others ask patients to conduct a given behavior

(capacity). A disadvantage of clinician ratings is that they are typically conducted outside of the patient's habitual environment (i.e., clinic or hospital) rather than in the patient's home or some other everyday setting. Therefore, they may not generalize to the patient's everyday setting (particularly if the patient still lives independently).

Barthel Index. One of the most widely used clinician-rated scales is the Barthel Index. To complete this and other, similar instruments (Table 57.3a–5), the clinician rates several different ADLs on a Likert-type scale that ranges from dependent to independent. The Barthel Index assesses more basic functions such as feeding, bathing, grooming, dressing, bowel movements, bladder control, toilet use, transfers from bed to chair and back, mobility, and stair use. Items on the index mostly range from 0 to 10 (some from 0 to 5) and have behavioral anchors for midpoints (e.g., for the transfers item, a rating of 5 indicates major help from one or two people). The Barthel Index and the Katz ADL (which assesses bathing, dressing, toileting, transfers, continence, and feeding) are most appropriate for individuals who have significant physical or cognitive impairments.

Functional Independence Measure. The FIM is a clinician-rated assessment that is widely used within rehabilitation settings for the determination of functional dependence/independence. It is important to note that the FIM rates observed performance, not capacity; this limits the measure's usefulness to inpatient or other settings where the clinician has frequent observation of the patient. Each of the FIM's 18 items is rated on a 7-point Likert-like scale that ranges from total dependence to functional independence. The scale requires trained raters (with at least 80 percent accuracy), and certification is provided for a fee by the Uniform Data System for Medical Rehabilitation (UDS). The FIM features both a motor and a cognitive component. The items covered are listed in Table 57.3a–7.

The Functional Assessment Measure (FAM) was created to supplement the FIM and includes 12 items related to cognitive, behavioral, communication, and community functioning. These items are also listed in Table 57.3a–7. The FIM + FAM comprises a total of 30 items and is often used to evaluate change over the course of an inpatient stay. Both the FIM and the FAM were designed primarily for persons with brain injury, but they are also widely used in other settings that provide treatment for older adults.

Performance-Based Measures of Everyday Functioning

Performance-based measures of everyday functioning use role play and mock settings to simulate traditional everyday functioning tasks. These tasks are administered under consistent conditions and use standardized instructions. For instance, rather than ask a patient or a collateral

informant about the patient's cooking skills and safety (which is limited by the insight, candor, and objectivity of the patient or the informant) or observe the patient at home and watch him or her prepare a meal (which may be too time- or labor-intensive), a performance-based measure will have the patient prepare a meal using a mock kitchen based in a laboratory or hospital. Performance-based measures assess capacity or what a person "can do," and represent the enacted tense of the Glass (1998) model, under standardized, optimal conditions rather than in the patient's habitual environment. This distinction is important because in some disorders, what impairs performance is not a lack of capacity, but rather a failure to translate functional capacity into self-initiated functional performance.

Table 57.3a–7.
Everyday Functioning Skills Assessed by the Functional Independence Measure and the Functional Assessment Measure

FIM	FAM
Eating	Swallowing
Grooming	Car transfer
Bathing	Community access
Dressing upper body	Reading
Dressing lower body	Writing
Toileting	Speech intelligibility
Bladder management	Emotional status
Bowel management	Adjustment to limitations
Bed, chair, wheelchair transfer	Employability
Toilet transfer	Orientation
Tub and shower transfer	Attention
Walking/wheelchair locomotion Stairs	Safety judgment
Comprehension	
Expression	
Social interaction	
Problem solving	
Memory	

It should be noted that the distinction between clinician ratings and performance-based measures of functioning is fundamental. Performance-based measures typically use a mock setting and are scored similarly to a neuropsychological test, with little ambiguity in criteria. By contrast, clinician ratings occur in less-standardized settings (hospital or clinic), are generally geared toward more basic ADLs, and typically require a more experienced or highly trained rater to score appropriately.

A variety of performance-based measures have appeared over the past 30 years in discipline-specific journals (e.g., occupational therapy, clinical and educational psychology, psychiatry, neurology). The authors conducted a review of these measures. Table 57.3a–8, adapted from the authors' review, shows the functional domains assessed by commonly used

performance-based measures. The authors present measures that assess multiple domains as well as some that focus on a specific domain. Tables 57.3a–3 and 57.3a–9 briefly describe these measures and provide information on scoring and approximate assessment times (more details in Moore DJ et al., 2007). In addition to these tables, which primarily show assessments of traditional ADLs and IADLs, Table 57.3a–10 shows some common performance-based measures of physical functioning and simple motor skills. As there is no specific gold standard measure, a broad representation is presented in the tables.

Performance-based measures of everyday functioning, like all assessment tools, have strengths and weaknesses. The strengths are many. Foremost are the standardized methods of administration, which leave little ambiguity in the scoring. As a result, inter-examiner reliability tends to be high. Additionally, these assessments focus on real-life skills in a safe setting (e.g., cooking tasks can be controlled) and are face valid. Furthermore, they do not rely on the patient's insight, or ability to answer questions on a self-report questionnaire which might be influenced by education or reading level, and they do not depend on the knowledge or objectivity of a collateral or proxy reporter. These instruments are good targets for intervention studies designed to improve everyday functioning and have been implemented in this role (e.g., the UCSD Performance-based Skills Assessment, UPSA). The UPSA is a real-world outcome measure for the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) trial, which is designed to assess the effectiveness of antipsychotic medications in improving cognition. The consensus appears to be that performance-based measures have more strengths than weaknesses and offer benefits over other assessment approaches that make them useful for clinical treatment studies.

However, performance-based measures do have weaknesses. One is that the test environment is contrived (e.g., it is not the patient's own kitchen, but a mock one). This decreases the patient's ability to follow behavioral scripts or schemas that allow him or her to successfully complete a task at home but not elsewhere. Performance-based assessments may achieve high reliability at the cost of validity. Perhaps the greatest weakness of performance-based measures of functional capacity, however, is that they measure capacity, not actual performance. For instance, a person may very well have the capacity to complete the task, but may not have the initiative or supportive environment in order to complete the task on his or her own. Finally, most of the current performance-based measures are designed to only assess one given functional domain at a time, whereas in daily life a person is often completing many functional tasks simultaneously.

While there are far more single-domain performance-based tests and/or multiple-domain tests that test each separately, there have been some multi-tasking tests developed. For example, a novel performance-

based multi-tasking assessment, which involves balancing the demands of four everyday tasks (i.e., financial management, cooking, medication management, and telephone communication) has been used in the HIV population and shown convergence with both neuropsychological performance as well as self-reported everyday functioning. For the aging population there is also the Day Out Task (DOT), which is a laboratory-based multi-tasking qualitative assessment using an automated smart home apartment. In this task participants must complete an interweaved set of activities needed to prepare for a day out. In addition to assessing everyday functioning and multi-tasking in elderly populations, this test was also designed to help researchers develop smart home adaptive technology interventions to help older adults age in place and live independently for longer.

Virtual Reality Performance-Based Measures of Everyday Functioning

While virtual reality (VR)-based assessment has been used to assess driving ability, there are several VR, performance-based everyday functioning assessments that have been developed recently that show promise for geriatric everyday functioning assessment. Moreover, many of these take into account multi-tasking. These VR tasks that assess multi-tasking aspects of everyday functioning may be particularly sensitive to executive dysfunction accompanying age-related neurodegenerative disorders involving frontal pathology (e.g., planning, task switching).

The VR multiple errands test (VMET) was adapted from a non–VR version of the MET, due to the need for a more feasible approach to assess multi-tasking and executive abilities in the context of everyday functioning. Several versions have been established, but in general participants complete a set of diverse subtasks within a simulated shopping mall and/or office environment, and include a planning component, tapping into the role of metacognition and metacognitive strategies in everyday task performance.

Table 57.3a–8.
Domains Assessed by Several Performance-Based Batteries of Everyday Functioning

Instrument	Communication	Hygiene	Dress	Eating	Time	Safety	Finances	Cooking	S
PERFORMANCE-BASED BATTERIES									
Dementia (primarily Alzheimer disease, but									

some dementia NOS)								
Activities of Daily Living Situational Test (AST) (Skurla, 1988)	X		X					X
Direct Assessment of Functional Status (DAFS) (Loewenstein, 1989)	X	X		X	X		X	
Structured Assessment of Independent Living Skills (SAILS) (Mahurin, 1991)	X		X	X	X		X	X
Functional Performance Measure (FPM) (Carswell, 1993)		X	X					X
Cognitive Performance Test (CPT) (Burns, 1994)	X	X	X					X
Refined ADL Assessment Scale (RADL) (Tappen, 1994)		X	X	X				
Direct Assessment of Functional Abilities (DAFA) (Karagiozis, 1998)							X	X
Texas Functional Living Scale (TFLS) (Cullum, 2001)	X		X		X		X	
Older Psychiatric Patients (Primarily schizophrenia, some mixed psychiatric groups)								
Kohlman Evaluation of Living Skills (KELS) (Thomson, 1992)	X	X				X	X	
Performance	X	X	X	X		X	X	X

Assessment of Self-Care Skills (PASS) (Rogers, 1984; Holm, 1999)								
UCSD Performance-based Skills Assessment (UPSA) (Patterson, 2001)	X						X	X
Mixed (Psychiatric and Dementia)								
Performance Test of Activities of Daily Living (PADL) (Kuriansky, 1976; Guralnik, 1989)	X	X	X	X		X		
Occupational Therapy Evaluation of Performance and Support (OTEPS). Also known as OTAPS (Nadler, 1993)	X	X				X	X	X
Independent Living Scales (ILS) (Loeb, 1996)	X				X	X	X	
St. George Hospital Memory Disorders Clinic Occupational Therapy Assessment Scale (OTAS) (Fairbrother, 1997)	X	X	X				X	X
Older Healthy Controls								
Activities of Daily Living Test (ADL-T) (Warren, 1989)	X						X	X
Everyday Problems Test (EPT) (Willis,	X						X	X

1993)				
Everyday Problems Test for Cognitively Challenged Elderly (EPCCE) is derived from the EPT (Willis, 1998)				
Observed Tasks of Daily Living (OTDL) (Diehl, 1995)	X			X

Table 57.3a–9. Single Domain Performance-Based Batteries of Everyday Functioning

Test	Test Description	Site	Scoring	Admin Time (Min.)	Other
Dementia (Primarily Alzheimer Disease—AD, some dementia NOS)					
Beck Dressing Performance Scale (BDPS) (Beck, 1988)	In most recent version, participant is rated on 42 discrete component steps of dressing for males and 45 steps for females (e.g., button shirt).	Home	Participants are scored on a scale from 0–7 with 7 indicating maximal assistance and 0 representing no assistance.	<30 min.	
Financial Capacity Index (FCI) (Marson, 2000)	6 domains; 14 individual tasks of financial ability. Domains range from basic monetary skills to financial judgment.	Clinic	Cutoff scores place individuals into levels of capacity.	30–50 min.	Modified version (FCI-8) assesses 8 domains and has 19 tasks.
Kitchen Task Assessment (KTA) (Baum, 1993)	Rate the participants' ability to cook pudding from provided directions. 6 components are assessed. Components scored include performance of all steps, and judgment and safety.	Clinic	The level of support from the tester is scored from 0–3.	<30 min.	
Psychiatric Patients (Primarily schizophrenia, some mixed psychiatric groups)					
Medication	Asks participant to	Clinic	Comprehensive		

Management Ability Assessment (MMAA) (Patterson, 2002)	walk through 4 fictitious medication regimens.		manual available upon request.		
Individual medication points vary. Total score ranges from 0–25.	15 min.	Older Schizophrenia Patients			
Mixed (Psychiatric and Dementia)					
Time and Change Test (T&C) (Froehlich, 1998)	Tests participant's ability to tell time and make change.	Clinic	Screen for dementia or not based on cut scores.	2 min.	
Older Healthy Individuals					
Drug Regimen Unassisted Grading Scale (DRUGS) (Edelberg, 1999)	Patient is asked to perform each of four medication-related tasks: identification, access, dosage, and timing. Participants used their own medications.	Clinic	Participant given a score of able (1) or unable (0) for each task for each medication.	>20 min.	
Functional Ability to Take Medications (FATM) (Meyer, 1989)[a]	Patient is presented with 6 specific medication-related tasks (e.g., opening a pill bottle).	Clinic	Items are scored as either yes or no.	<10 min.	Instrument was later slightly modified.

Note: Tests preceded by a < or a > indicate approximation of administration time based on the text provided. Readers are encouraged to examine the literature for additional studies pertinent to their clinical or research question that may apply their candidate measures.

[a]Abbreviation assigned by the author of the current chapter.

From Moore DJ, Palmer BW, Patterson TL, Jeste DV. A review of performance-based measures of functional living skills. *J Psychiatr Res*. 2007;41(1–2):97–118.

Another measure is the Multi-tasking in the City Test (MCT), an errand-running (e.g., shopping, attending appointments, etc.) task that takes place in a virtual city. Similar to the VMET, the MCT measures multi-tasking and executive abilities in an everyday setting and requires participants to create a plan before they begin. Unlike the VMET test, tasks must be performed with little explicit rule constraints because the goal of this test was to explore whether and how often subjects participate in behavior that is

clearly not goal-directed when there is no explicit instruction or "rules" directing against such behaviors

Table 57.3a–10.
Performance-Based Tests of Physical Functioning

Physical Performance Test (PPT)
Established Populations for Epidemiologic Studies of the Elderly (EPESE)
Timed Manual Performance Test (TMP)

Table 57.3a–11provides a list and description of several recently developed VR assessments. Although many of these tests were not necessarily designed for older adult populations, they may prove useful in healthy and pathological aging. The advantages of VR-based assessments are: (1) increased safety, (2) improved standardization and objectivity in terms of stimulus delivery, measurement/scoring, and error reduction, (3) ability to detect subtle deficits, and (4) portability which makes them ideal for clinical settings and has potential for remote delivery. Many of these tasks are new and are therefore still in the development and validation stages; however, they do show promise for geriatric assessment in future research and practice (Table 57.3a–11lists the populations that have been used in to-date). It is important for these VR assessments that participants and patients are given a change to familiarize themselves with the computerized interface, especially when they have no prior computer experience.

Table 57.3a–11.
Virtual Reality (VR) Performance-Based Everyday Functioning Assessments

Test	Test Description	Scoring	Admin Time (Min.)	Populations Used In To Date	O
Nonimmersive Virtual Coffee Task (NI-VCT; Allain et al., 2014)	14 steps; virtual environment simulates kitchen; patients must make a cup of coffee; requires multiple cognitive processes: serial ordering of several steps (open the drawer, etc.), object selection (coffee filter, etc.), and technical manipulations (coffee machine) to complete goal.	5 outcomes: Total completion time; percentage of task steps completed with or without error (range 0–100%); total number of errors (e.g., omissions, commissions).	15 min. for instructions and training. No time limit, but in Allain et al., 2014 it took AD patients ~11 min. and healthy elderly controls ~6 min.	AD patients and healthy older adult controls	A

Virtual Errands Test (VET; McGeorge et al., 2001); also called Virtual Multiple Errands Test (V-MET)	12 errands (in McGeorge version); diverse subtasks (e.g., collecting a book, meeting a colleague) within simulated environment; includes a planning component, which taps into the role of metacognitive strategies in everyday tasks. McGeorge version more vocationally oriented and contains work-related errands, while other versions of the task include shopping-related errands.	Planning; total number of errands completed.	Unlimited planning time; 20 min. for test.	Healthy adult controls; brain damage patients; Parkinson disease.	A
Multi-tasking in the City Test (MCT; Jovanovski et al., 2012)	8 tasks; virtual environment has several stores (e.g., drug store, post office); specific task requirements/errands listed on screen, along with errands completed, amount of money in wallet, and clock displaying time; designed to be open-ended with few explicit rules (i.e., complete as many errands as possible in 15 min.).	Planning; completion time; number of tasks completed; qualitative errors (i.e., task failures, task repetitions, inefficiencies).	Unlimited time for planning; 3 min. for instructions and 15 min. for test.	Stroke; TBI; Healthy young adult controls; not used in aging yet.	
Virtual Reality Functional Capacity Assessment (VRFCAT; Ruse et al., 2014)	4 scenarios (i.e., navigating a kitchen, getting on a bus to go to a grocery store, finding/purchasing food in a grocery store, and returning home on a bus); 12 objectives; three factors identified in development: Reasoning and Problem Solving, Speed of Processing, and Working Memory.	Completion time for each objective; accuracy/errors for each objective.	Not specified; time limits for each task; for healthy controls range across 12 objectives was ~25–120 sec., for schizophrenia was ~25–225 sec.)	Healthy adult controls and schizophrenia (developed for schizophrenia and other CNS disorders; studies are underway)	

VRMMA (Virtual Reality Apartment Medication Management Assessment Kurtz et al., 2002)	3 types of pills must be "taken"; exploration phase then actual "test" phase. In ''testing'' phase, participant told several seconds into task a post-it note would appear telling them which medications to take and when to take them (e.g., ''Remember to take: 2 yellow, 4 red, and 1 blue pill in 15 min.'' and to use all items in apartment that were explored during the practice phase; virtual apartment consists of a living room, bedroom, bathroom, and kitchen.)	Dependent measures include: quantitative errors; qualitative errors; time discrepancy; total distance traveled; clock checks; reminder note checks.	If an attempt to take medications not initiated by 18 min., notice appears saying ''time has expired, please go take your medication now.'' If 23 min. elapses and participant still has not taken medication, the program terminates.	Schizophrenia and healthy adult controls.	D

Direct Observation of Everyday Functioning

Some measures employ direct observation, which may be more valid than subjective reports. Direct observation has typically been used in inpatient settings where clinicians can easily record behaviors. Although many consider direct observation to be the "gold standard" of functional assessments, it is in fact more of a theoretical technique than a practical tool. That is, the most accurate representation of functional ability would consist of observing someone in her natural environment as she conducts her daily tasks without knowledge that she is being observed. In most cases, direct observation of this type is unethical and unrealistic. Even where the patient's consent has been obtained for direct observation, it is labor intensive and provides data only about the behaviors engaged in during the observation period. Furthermore, a person who knows he is being observed will frequently alter his behavior, sometimes radically.

CLINICAL POPULATIONS AND EVERYDAY FUNCTIONING

The following text discusses selected domains of everyday functioning that are of concern among specific clinical populations of older adults or that have received substantial research (e.g., functional impairments in older persons with schizophrenia). This list of clinical populations is in no way exhaustive, but it gives a sense of the ways in which functional impairment can differ from group to group.

Normal Aging

Cognitive declines associated with normal aging seldom reach a level sufficient to impact daily functioning. Among older persons without dementia or some other condition known to affect cognition, the neuropsychological domains that show the greatest decline with aging are processing speed and psychomotor speed. Declines in these domains are unlikely to significantly influence basic ADLs such as grooming, bathing, or using a telephone. Normal declines in these domains may, however, have some effect on speeded IADLs such as driving. Driving relies heavily on intact processing and psychomotor speed, and as discussed above, driving is a domain that warrants particular attention from clinicians when assessing older adults.

Dementias and Neurological Conditions

The cognitive profiles of individuals with dementia can be quite varied. AD, the most common of the dementias, often produces difficulties with memory (particularly with delayed recall) as well as language difficulties (e.g., anomia). Difficulties with memory can lead to problems with many IADLs such as financial and medication management. Tasks such as cooking and meal preparation are highly dependent on the order in which they are conducted, and a person with memory impairment may forget a critical step or unknowingly repeat a task. See the following case study (Mrs. Smith) for an example of a thorough assessment of neuropsychological, mood, and daily functioning in a patient suspected of having AD.

DEMOGRAPHICS

Mrs. Smith is a 71-year-old, right-handed, Caucasian woman.

EDUCATIONAL AND VOCATIONAL HISTORY

The patient completed 14 years of education. After completing 2 years of college she accepted a position as a secretary in a law office and remained in that position until she retired at the age of 55.

BEHAVIORAL OBSERVATIONS

The patient was alert and oriented to day of the week, year, and city, but was incorrect by 2 days on the exact date. She knew the month and she could name the president, governor, mayor, and the name of the building she was in. The patient was pleasant to work with. Her mood was positive and her affect was congruent with her mood. She mentioned that her children were worried about her memory lately, particularly her daughter. She interacted with the examiner in an appropriate manner. She took a few breaks to stretch her legs. Her frustration tolerance was good. She showed very good attention and low distractibility. Gross visual acuity with glasses was intact and her hearing was adequate. She was able to understand test instructions with minimal repetition. Overall, the patient seemed to put forth good effort on all of the tests. The results are considered to be a valid estimate of her current cognitive and everyday functioning.

GENERAL COGNITIVE FUNCTIONING

Mattis DRS-2 out of a total possible score of 144:		
Total Raw Score:	127	Scaled score: 6
Attention Raw Score:	36	Scaled score: 11
Initiation/Perseveration Raw Score:	37	Scaled score: 11
Construction Raw Score:	6	Scaled score: 10
Conceptualization Raw Score:	32	Scaled score: 7
Memory Raw Score:	16	Scaled score: 2

VISUOSPATIAL FUNCTIONING

Performance on the Clock Drawing Test was normal, at the 50th percentile, consistent with her performance on the DRS construction score.

PSYCHOMOTOR SPEED

Trail Making Test Part A and Trail making Test Part B were severely impaired, at the <1st percentile.

EVERYDAY FUNCTIONING ASSESSMENT

The patient was administered several measures of everyday functioning. Specifically, the patient was administered portions of the performance-based Independent Living Scale (ILS): Managing Money; the self-report Minimum Data Set (MDS) questionnaire that queries on the frequency that one needs help on and difficulty experienced when performing 20 IADLs; the SPPB; and an on-road driving assessment. On the ILS: Managing Money her performance fell within the mildly impaired range when compared to matched peers. For example, the patient inaccurately calculated how much change she should get back if she purchased something for $1.19 and gave the clerk $5.00 (her answer was $4.79). When presented with mock check materials and specific directions for completing the check, she correctly filed out all of the check, except writing the wrong date. She was able to calculate the balance in her checking account. On the self-report MDS, she reported that she only needs help "some of the time" with cooking, and experienced no difficulty with any of the activities. On the SPPB the patient scored 8 points total of 12 possible points, indicating relatively intact physical functioning: 3 out of 4 points on the balance tests (missed 1 point on holding the tandem stand for less than 10 seconds); 2 out of 4 points on the gait speed tests; and 3 out of 4 points on the chair stands. Finally, on the on-road drive, the patient's discrete maneuvers were rated unsatisfactory within the majority of 8 domains evaluated but fully intact in the remaining domains.

CORROBORATING REPORT FROM MRS. SMITH'S DAUGHTER

The patient's daughter stated that her mother developed memory problems over the last 12 months. She stated that Mrs. Smith is able to remember to take her daily medications with no issues, but has trouble remembering appointments and birthdays. Her daughter also reports that Mrs. Smith does require some assistance with managing finances and does not drive alone. Her daughter mentions that since Mrs. Smith has started being more diligent about using a pocket calendar to keep track of events and appointments, her memory functioning is better.

MOOD ASSESSMENT

On the Center for Epidemiologic Studies Depression Scale (CES-D) (a self-report measure of depressive symptomatology), the patient endorsed two items, indicating that she was not experiencing significant clinical depressive symptoms.

SUMMARY AND CONCLUSIONS

Results of the current evaluation indicate that the patient's overall level of cognitive functioning was overall in the moderately impaired range for her age (DRS-2 Total score = 127); however, her memory and executive functioning were severely impaired, fitting the profile of amnestic Mild Cognitive Impairment (MCI) due to AD. The patient is overall independent in most IADLs, however her memory impairments are having at least a mild effect on some domains, such as managing finances and keeping track of appointments, which is consistent with her daughter's reporting. The etiology of her cognitive deficits is unknown, however vascular disease does not appear to be a primary contributor, as she has no medical

comorbidities, vascular or otherwise (e.g., no depressive symptomology). It was recommended that she receive assistance with her complex financial and medical decisions. She could continue to manage her own medication administration with some supervision and aid (e.g., weekly pill organizer). Her driving, while not impaired at this stage, should be closely monitored. She was advised to not drive alone and only drive in familiar areas that are proximally closer to her home. Finally, her family was cautioned that the observed difficulties may continue progress and plans should be made now for this possibility.

Dementias and Neurological Conditions (Continued)

The use of memory aids such as calendars, electronic reminders, and checklists may help patients to circumvent their problems for a brief period; however, as dementia progresses, cognitive difficulties tend to worsen and compensatory strategies may fail. Many persons in later stages of AD become completely dependent in all their ADLs. Persons in early stages of AD can appear to have few daily functioning difficulties. They may be particularly capable of maintaining appropriate over-learned social interactions (e.g., introducing themselves to a new person); however, if pressed for details about newly obtained information, they may have difficulty recalling.

Other neurodegenerative conditions can present interesting challenges for everyday functioning assessments. For example, persons with Lewy Body Dementia can have profound impairments of visuospatial skills. They are likely to have great difficulty with higher-order tasks dependent on intact visuospatial functioning such as driving, whereas they may have less difficulty with social skills. Persons with Parkinson disease may have a slightly different set of daily functioning difficulties. Parkinson disease causes particular difficulties with initiation of motor movements and more generally with motor skills. Initiation problems may hinder a person from starting activities that they are cognitively capable of planning and executing. Difficulties with fine motor skills may prevent the accomplishment of basic tasks like opening medication bottles.

Psychiatric Conditions

Older persons suffer from the same types of psychiatric conditions as younger persons, although in some cases (e.g., depression), incidence may increase in older adulthood. Accordingly, assessment instruments need to be designed to capture daily functioning difficulties in persons with a serious mental illness. This group manifests particular deficits that may be important to evaluate. The following text reviews the difficulties with everyday functioning among these individuals as well as some of the instruments used to assess those difficulties.

Depression. Some of the most consistent evidence linking depression to neuropsychological dysfunction comes from studies of older people, in whom disturbances in executive function can be prominent and these

impairments may relate to everyday functioning difficulties. Given that depression is common in older age, and that cognitive problems can increase among older adults, this is a particularly important concern. Some investigators have suggested that impaired cognition may be limited to patients experiencing somatic symptoms of depression. Other studies have suggested that cognitive problems are not necessarily more common with increasing age bands, but rather that "premature" aging may occur in certain cognitive domains (e.g., nonverbal memory, word generation, and certain frontal lobe skills). Given that difficulties with daily functioning can be more common among older adults, the presence of depression can exacerbate these difficulties.

Particularly among older individuals, the presence of depressive symptoms is associated with a decline in performance of ADLs. Thirty to 60 percent of individuals who have been hospitalized for depression experience a drop in their ability to perform ADLs, including bathing, dressing, toileting, transferring, and eating. One study of 572 recently hospitalized older individuals found that persons with more depressive symptoms were significantly more likely to be dependent on others to help them perform ADLs as compared to individuals with no or few depressive symptoms. This finding was also true for IADLs, which include taking medicine, handling finances, transportation, and using the telephone. Elderly patients increased their risk of being hospitalized by failing to comply with prescribed medication regimens. Other studies have confirmed that depressive symptoms are strongly associated with poor functional performance, and depression scores for older individuals were significantly higher for those who reported experiencing ADL decline. Among community-based samples of older adults, neuropsychological impairment and depression appear to be two of the strongest predictors of daily functioning problems, even when controlling for baseline cognitive function, alcohol consumption, and chronic health conditions. Unintentional nonadherence to psychotropic medications has been shown to be associated with greater cognitive impairment.

Among adults with a primary major depressive episode and severe depressive symptoms (mean Beck Depression Inventory = 34.3 [SD = 11.0]) cognitive impairment was strongly associated with impairment in IADLs (e.g., medication taking and finance handling), whereas more severe depressive symptomatology and older age were strongly associated with impairments in basic ADLs. However, patients' satisfaction in role functioning and relationships improved with age.

Schizophrenia. Daily functioning difficulties are common among persons with schizophrenia. As a result, there are several performance-based instruments to capture everyday functioning difficulties common among older persons with schizophrenia that have been developed. These

instruments include the UPSA, including newer computerized versions; the Medication Management Adherence Assessment (MMAA); and the SSPA. Performance on these everyday functioning measures is strongly correlated with global neuropsychological functioning. A recent study showed that processing speed, episodic memory, and executive functions were associated with general everyday functioning (UPSA scores), whereas working and episodic memory and verbal fluency were associated most strongly with social skills (SSPA scores). These results suggest that neuropsychological abilities and social skills have differential effects on general everyday functioning skills. Consistently, neuropsychological impairment has been shown to be a stronger correlate of functional status than psychiatric symptoms among persons with schizophrenia. Negative symptoms of schizophrenia (e.g., flat affect) tend to be correlated with neuropsychological impairment; however, positive symptoms are rarely associated with neuropsychological or daily functioning impairments.

Recent research indicates that scores on the UPSA are associated with "real-world" outcomes, thereby making the UPSA a potentially useful tool to aid in clinical decision making. For example, a recent study found that UPSA scores of 75 or above were predictive of patients' ability to reside independently (e.g., alone in a home or apartment). In this same study, the UPSA was significantly better at predicting residential independence than the Mattis Dementia Rating Scale (DRS), a commonly used instrument that assesses global cognitive functioning. In another study, higher scores on the UPSA were associated with greater involvement in community responsibilities, such as working for pay, engaging in volunteer work, and attending school, among Hispanic/Latino patients with psychosis. In contrast, the DRS was not significantly associated with greater engagement in community responsibilities. These two studies may suggest that performance-based assessments of daily functioning are more proximal to community functioning than is global cognitive functioning.

Recently Hospitalized Older Individuals

Unfortunately, many older persons never receive an assessment of their everyday functioning abilities until they are hospitalized. Because information about how the person was functioning prior to hospitalization may be sparse, it is often difficult to determine how much change in daily functioning resulted from the hospital stay and related events. The hospitalization may result directly from a recent decline in functional abilities or from some unrelated cause (e.g., a need for hip replacement or coronary bypass surgery). Regardless of the reason, ADLs among elderly patients are more likely to deteriorate and less likely to improve after hospitalization. Specifically, it has been shown that 30 to 60 percent of older individuals who have had a recent hospital stay acquire at least one new dependency in daily functioning. A recent study showed that the

variables most often associated with functional decline among older patients who have been hospitalized are age, lower current ADL or IADL abilities, worse cognitive functioning, lower preadmission IADL, depression, and length of hospital stay. Few screening instruments have been developed to measure the risk of functional decline among elderly individuals who have been recently hospitalized; however, a recent review of the literature identified three useful screening instruments: Hospital Admission Risk Profile (HARP), Identification of Seniors at Risk (ISAR), and Care Complexity Prediction Instrument (COMPRI).

Persons Living with HIV

Neurocognitive impairment commonly occurs in HIV infection, particularly among older individuals living with HIV. Neurocognitive problems may stem from the disease itself, HIV medication side effects, comorbidities associated with HIV (e.g., stroke, diabetes) or a combination of factors. The severity of the HIV-associated neurocognitive disorders (HAND) varies from asymptomatic forms that do not interfere with daily functioning to more moderate and severe forms that interfere moderately or significantly with everyday functioning. Specifically, neurocognitive impairment in HIV has been associated with driving, medication management abilities, and other IADLs. Many of the daily functioning assessments used with older HIV-infected persons, such as UPSA and Lawton & Brody ADL Scale, were originally developed in non–HIV-infected populations. While translating these assessments to the older HIV-positive population may be appropriate in terms of measuring functioning, the etiology of the everyday functioning impairment should be considered in the context of the unique risk factors in this heterogeneous population.

DEMOGRAPHICS

Mr. Anderson is a 62-year-old, right-handed, African-American male.

EDUCATIONAL AND VOCATIONAL HISTORY

The patient completed 13 years of education and received his high school diploma. He denied having any problems learning to read or write and never skipped or repeated a grade. He worked as a cook in a restaurant for approximately 30 years.

BEHAVIORAL OBSERVATIONS

The patient was alert and oriented to day of the week, date, year, and city. The patient was diagnosed with major depressive disorder in his early 20s and was diagnosed with HIV infection at age 50. He has been worried for some time that he may have some neurocognitive difficulties. He complains of slowed ability to process information, difficulty learning new strategies, and difficulty planning and initiating new activities. During the assessment the patient was apathetic and appeared somewhat lethargic, but was pleasant and cooperative. Gross visual acuity and hearing were intact. Overall, the patient seemed to put forth very good effort on all of the tests. The results are considered to be a valid estimate of his current cognitive and everyday functioning.

GENERAL COGNITIVE FUNCTIONING

The patient was administered a comprehensive neuropsychological test battery. Using T-scores

adjusted for age, education, sex, and ethnicity, Mr. Anderson was found to be impaired in the domains of learning and executive functioning. His performance was within normal limits on verbal functioning, speed of information processing, recall, motor, and attention/working memory. His global neuropsychological performance is mildly impaired (impairment ≥2 domains).

EVERYDAY FUNCTIONING ASSESSMENT

The patient was administered the UCSD Performance-based Skills Assessment (UPSA), which assesses overall ability to manage everyday functional tasks including ability to plan and organize events, manage finances, communicate with others, arrange for transportation needs, and conduct household chores. Overall, his performance fell within the moderately impaired range (overall score = 63). His UPSA score suggests some reduced capacity in maintaining his independent living situation.

CORROBORATING REPORT FROM MR. ANDERSON'S THERAPIST

The patient's therapist stated that the patient has experienced some mild difficulty performing day-to-day tasks including managing his own medications and finances. Specifically, the therapist stated that on occasion the patient has reported that he sometimes has difficulty in organizing and managing his medications. The therapist has increasingly become concerned with the patients neurocognitive and daily functioning difficulties. The therapist indicated that these neurocognitive and functional impairments have been stable over the past several years and do not appear to be worsening.

MOOD ASSESSMENT

The patient scored 11 on the Hamilton Depression scale, which is indicative of mild symptoms of depression. The patient denied suicidal ideation.

SUMMARY AND CONCLUSIONS

Results of the current evaluation indicate that the patient's overall level of cognitive functioning was mildly impaired as compared to other individuals of his demographic background. The patient also demonstrated moderate impairments on a measure of everyday functional abilities (UPSA). The patient's therapist corroborated many of the patient's daily functioning difficulties. These test results suggest mild to moderate impairments in cognitive and everyday functioning. The longitudinal course of these impairments has been stable. That is, cognitive and everyday functional abilities appear to have remained stable over the past several years, which is again consistent with published literature on this topic. Based on his current level of cognitive functioning and deficits in functional skills, the patient was recommended to seek additional assistance in instrumental activities in daily living activities. It was also recommended that the patient receive functional skills training classes, which may help improve his ability to perform essential everyday skills (e.g., mange finances, communicate with others). Particularly important is his ability to self-administer his medications. Providers should consider additional training in managing medications as well as methods to reduce errors in administering medications (e.g., blister packs, electronic reminders, pill organizers).

RECOMMENDATIONS FOR FUTURE INSTRUMENT DEVELOPMENT

Assessments of daily functioning continue to evolve. Below, the authors offer recommendations for future development.

Everyday functioning is multi-faceted; however, most assessment instruments are designed to test a given domain in isolation. For instance, the MMAA assesses medication management abilities independent of other abilities. Medication management in "real life" occurs over a much longer time span and is intertwined with other daily tasks. The ideal assessment of daily functions would test multiple domains of functioning concurrently.

Some multi-tasking tests of this nature have been developed, but additional assessments of this type are clearly warranted.

Another desirable component of future assessment instruments is that they be adaptive. That is, instruments should have the flexibility to adjust to the level of mastery of a particular individual. If an item was too difficult, the individual could be tested with easier items, or vice-versa. A great deal of research has been done on adaptive testing, particularly with computer-administered, standardized achievement tests. Hopefully this technology can be applied to the assessment of functional capacity.

VR technology may offer a way of creating more realistic mock environments in which multiple aspects of functioning could be tested. Assessments could even be designed to recreate a patient's unique habitual environment, and many assessments have already been developed and are currently being tested. Some argue that older adults are too resistant to digital technologies for such technologies to be useful with this population, but the degree of resistance may be exaggerated. Furthermore, performing VR tasks may not require high levels of technical knowledge. These techniques may be useful both for assessments and for interventions. These technologies have already been implemented for the treatment of certain disorders (e.g., PTSD).

Very few functional-assessment instruments provide demographically adjusted normative data, with the most important variable being age. This constitutes a significant limitation as compared to other types of assessment, such as neuropsychological tests, for which the assessor typically has an indication of where a given score falls on a continuum. Ideally, all functional assessments would have normative information. In the absence of normative data, an evaluator who is more experienced is likely to have a better sense of what a given score means in the context of other objectively gathered data. The authors understand that normative studies are not easily funded; however, the authors believe that future studies of this type should be a research priority.

In addition to demographically corrected normative data, there are significant advantages to the co-norming of different assessment modalities together. If evidence could be obtained of how certain instruments work with one another, then better insight into a given patient's true difficulties would be available. For example, if a self-report, collateral report, and a performance-based measure were concurrently administered to a large sample of older adults with varying diagnoses and ages, the results might help us to understand the distinct strengths and weaknesses of the various assessment modalities. It might be found that persons who score below a given level of self- and collateral-reported functioning would derive no benefit from receiving a performance-based assessment. On the other hand, in persons who report a relatively intact level of functioning there might still be difficulties on performance-based assessments due to

inaccuracies in self- or collateral report. The combinations are many, but ideally it would be helpful to know how the most widely used instruments worked together.

Finally, the discussion of norming and co-norming is moot if the instruments in question have no validity for predicting actual functioning. Predictive validity is difficult to obtain without a widely recognized gold standard of assessment for everyday functioning abilities; however, the authors believe that using a multi-modal assessment strategy allows for an accurate representation of daily functioning. Until reliable and valid instruments with strong predictive validity are established, there will be difficulty in establishing the true usefulness of everyday functioning assessments.

SUMMARY

Older adults are at significant risk for difficulties with everyday functioning. These difficulties can range from impairments on basic behaviors critical for independent functioning, such as toileting to optional daily tasks, such as driving. Several factors may influence daily functioning, and cognitive impairment appears to be a strong predictor of daily functioning difficulties. Numerous everyday functioning assessment methodologies are available and include clinical interview, self- and collateral report, clinician ratings, performance-based measures, and direct observation. Performance-based measures may offer significant advantages such as improved reliability and a better estimate of functional capacity as compared to other assessment approaches such as self-report. More research is clearly needed to develop quick, reliable and valid assessments of everyday functioning for older adults; the field is currently burgeoning, and improved assessments will continue to be developed in the near future.

ACKNOWLEDGMENT

The authors would like to thank Virginia Wadley, Ph.D. for contributing to one of the case presentations in this chapter, and Thomas L. Patterson, Ph.D. for contributing to a prior version of this chapter.

REFERENCES

Cardenas V, Mausbach BT, Barrio C, Bucardo J, Jeste D, Patterson T. The relationship between functional capacity and community responsibilities in middle-aged and older Latinos of Mexican origin with chronic psychosis. *Schizophr Res*. 2008;98(1–3):209–216.

Carlsson G, Haak M, Nygren C, Iwarsson S. Self-reported versus professionally assessed functional limitations in community-dwelling very old individuals. *Int J Rehabil Res*. 2012;35(4):299–304.

Carswell A, Spiegel R. Functional assessment. In: Wilcock GK, Bucks RS, Rockwood K, eds. *Diagnosis and Management of Dementia*. New York: Oxford University Press; 1999.

Coman L, Richardson J. Relationship between self-report and performance measures of function: a systematic review. *Can J Aging*. 2006;25(3):253–270.

Farias ST, Mungas D, Reed BR, Harvey D, Cahn-Weiner D, Decarli C. MCI is associated with deficits in everyday functioning. *Alzheimer Dis Assoc Disord*. 2006;20(4):217–223.

Glass TA. Conjugating the "tenses" of function: discordance among hypothetical, experimental, and enacted function in older adults. *Gerontologist*. 1998;38(1):101–112.

Grant I, Franklin DR Jr, Deutsch R, et al. Asymptomatic HIV-associated neurocognitive impairment increases risk for symptomatic decline. *Neurology*. 2014;82(23):2055–2062.

Green MF. Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. *J Clin Psychiatry*. 2006;67(10):e12.

Harvey PD, Velligan DI, Bellack AS. Performance-based measures of functional skills: usefulness in clinical treatment studies. *Schizophr Bull*. 2007;33(5):1138–1148.

Hoogerduijn JG, Schuurmans MJ, Duijnstee MS, de Rooij SE, Grypdonck MF. A systematic review of predictors and screening instruments to identify older hospitalized patients at risk for functional decline. *J Clin Nurs*. 2007;16(1):46–57.

Jovanovski D, Zakzanis K, Campbell Z, Erb S, Nussbaum D. Development of a novel, ecologically oriented virtual reality measure of executive function: the Multitasking in the City Test. *Appl Neuropsychol Adult*. 2012;19(3):171–182.

Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The Index of Adl: a standardized measure of biological and psychosocial function. *JAMA*. 1963;185:914–919.

Mackin RS, Nelson JC, Delucchi KL, et al. Association of age at depression onset with cognitive functioning in individuals with late-life depression and executive dysfunction. *Am J Geriatr Psychiatry*. 2014;22(12):1633–1641.

Mausbach BT, Bowie CR, Harvey PD, et al. Usefulness of the UCSD performance-based skills assessment (UPSA) for predicting residential independence in patients with chronic schizophrenia. *J Psychiatr Res*. 2008;42(4):320–327.

McCall WV, Dunn AG. Cognitive deficits are associated with functional impairment in severely depressed patients. *Psychiatry Res*. 2003;121(2):179–184.

Moore RC, Fazeli PL, Patterson TL, et al. UPSA-M: Feasibility and initial validity of a mobile application of the UCSD Performance-Based Skills Assessment. *Schizophrenia Research*. 2015;164(1-3):187–192.

Moore DJ, Palmer BW, Patterson TL, Jeste DV. A review of performance-based measures of functional living skills. *J Psychiatr Res*. 2007;41(1–2):97–118.

Morris JN, Morris SA. ADL assessment measures for use with frail elders. *J Ment Health Aging*. 1997;3(1):19–45.

Palgi Y, Shrira A, Zaslavsky O. Quality of life attenuates age-related decline in functional status of older adults. *Qual Life Res*. 2015;24(8):1835–1843.

Patterson TL, Lacro J, McKibbin CL, Moscona S, Hughs T, Jeste DV. Medication management ability assessment: results from a performance-based measure in older outpatients with schizophrenia. *J Clin Psychopharmacol*. 2002;22(1):11–19.

Patterson TL, Moscona S, McKibbin CL, Davidson K, Jeste DV. Social skills performance assessment among older patients with schizophrenia. *Schizophr Res*. 2001;48(2–3):351–360.

Pearson VI. Assessment of function in older adults. In: Kane RL, Kane RA, eds. *Assessing Older Persons: Measures, Meaning, and Practical Implication*. New York: Oxford University Press; 2000.

Royall DR, Lauterbach EC, Kaufer D, Malloy P, Coburn KL, Black KJ; Committee on Research of the American Neuropsychiatric Association. The cognitive correlates of functional status: a review from the Committee on Research of the American Neuropsychiatric Association. *J Neuropsychiatry Clin Neurosci*. 2007;19(3):249–265.

Schmitter-Edgecombe M, McAlister C, Weakley A. Naturalistic assessment of everyday functioning in individuals with mild cognitive impairment: the day-out task. *Neuropsychology*. 2012;26(5):631–641.

Schneider LC, Struening EL. SLOF: a behavioral rating scale for assessing the mentally ill. *Soc Work Res Abstr*. 1983;19(3):9–21.

Scott JC, Woods SP, Vigil O, et al. A neuropsychological investigation of multitasking in HIV infection: implications for everyday functioning. *Neuropsychology*. 2011;25(4):511–519.

Working Group on Health Outcomes for Older Persons with Multiple Chronic Conditions. Universal health outcome measures for older persons with multiple chronic conditions. *J Am Geriatr Soc*. 2012;60(12):2333–2341.

Zimmerman SI, Magaziner J. Methodological issues in measuring the functional status of cognitively impaired nursing home residents: the use of proxies and performance-based measures. *Alzheimer Dis Assoc Disord*. 1994;8(Suppl 1):S281–S290.

57.3b Psychiatric Problems in the Medically Ill Geriatric Patient

ANNA BORISOVSKAYA, M.D., BETH ANN LABARDI, M.D., MATTHEW ALAN SCHREIBER, M.D., PH.D., AND MARCELLA PASCUALY, M.D.

Since the beginning of the 20th century, a dramatic increase in life expectancy has produced an aged population of unprecedented size and longevity. According to demographic forecasts, one out of five people will be over the age of 65 in the year 2060, up from one in seven in 2012. In the United States, a 65 year old will typically remain independent until he or she reaches the age of 75, and will live an additional 6.5 years once he or she becomes dependent on others.

People over the age of 65 account for twice the average number of prescription drugs, one and a half times the number of physician contacts, and three times the number of hospital-bed days. Overall, the aging population accounts for more than one-third of the total personal health expenditures in the United States. Among the top 5 percent of highest health care spenders who collectively make up one-half of total spending, adults over age 65 account for 40 percent.

Over half of the population that is 65 and older report having at least one chronic disability, and a third report having a chronic condition that is severe enough to limit their activities. For the psychiatrist, medical comorbidity and disabilities, which aggregate with age, are the most

important factors distinguishing older patients from their younger counterparts.

Some of the most common disorders in the aged population include degenerative arthritis and cardiovascular disorders, mainly hypertension and heart disease. In addition, geriatric syndromes such as falls, vision and hearing loss, urinary incontinence, and impaired cognition disproportionately impact health care costs and quality of life. One recent study found that about 65 percent of adults over age 65 have two or more chronic diseases, and this increases to 82 percent in individuals over the age of 85.

Estimates of the prevalence of psychiatric disorders among the medically ill and disabled elderly vary widely across clinical settings, and reliable data on comorbidities are still sparse. For example, 30 percent of older adults in primary care have at least one active psychiatric diagnosis: 10 to 15 percent have clinically significant depressive syndromes such as major depression, dysthymia, or minor depression; 5 to 10 percent have dementia; and 2 to 5 percent have alcohol abuse or dependence or psychoses. The lowest prevalence rates for psychiatric symptoms and disorders are found among medically healthy persons living independently in the community, and prevalence increases with severity of medical illnesses and extent of functional dependency (Fig. 57.3b–1). The relationship between functional health and psychiatric comorbidity has been most clearly established for depression, but emerging data suggest similar trends for anxiety disorders, sleep disorders, and psychoses. In addition to impairing day-to-day functioning and quality of life, depression in medically ill older adults has been associated with decreased self-care and adherence to treatment regimens for chronic medical conditions, worse medical outcomes, increased health care costs, and mortality from suicide and medical causes.

The profound changes in the demographics of chronic disease have resulted in an increased need for psychiatric care of patients with comorbid medical and mental disorders. New models of practice, based on effective collaboration between psychiatrists and primary care providers or medical specialists, are required to address the needs of older patients. Clinical care of psychiatrically disturbed medical patients poses several special challenges for the psychiatrist. Clinical syndromes may vary in presentation from the classical disorders for which treatment guidelines are well established to situations where common disorders such as depression and anxiety are masked by medical comorbidities and functional disability, in place of overt expression of emotional symptoms. Thus, one needs a working familiarity with the common medical and functional problems of the elderly, and a full range of skills in psychotherapeutic and somatic treatment modalities. In spite of these inherent challenges, most psychiatrically disturbed older persons can be substantially helped when

approached within a systematic framework.

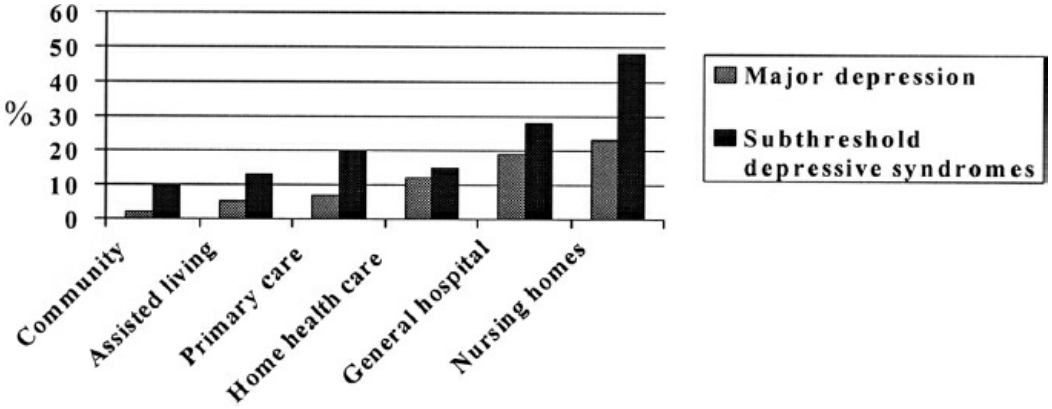

FIGURE 57.3b–1. Prevalence of depression in older adults.

APPROACH TO THE PATIENT

Mastery of a few central concepts helps integrate the range of clinical data essential in the successful evaluation and management of psychiatrically disturbed, medically ill geriatric patients (Table 57.3b–1). The adaptive reserve of all vital organs, including the brain, tends to decline with aging, even in the absence of manifesting disease. This loss of physiological resilience is generally detectable in clinical settings only after a sudden or severe deterioration in the face of acute illness or psychosocial stress. Age-related loss of adaptive reserve underlies the downward spiral and multiple organ failures often seen in sick elderly. It is responsible for the narrowed margin of safety observed for many drugs, changes in the presentation of disease, and increase in morbidity of medical and psychiatric disorders. For example, acute myocardial infarction in an octogenarian may present as painless delirium, while a mild chronic anxiety disorder may become fulminant panic in a patient with rectal bleeding who, ultimately, may not have a serious medical problem. Such complex patterns of causation are typical of psychiatric symptoms in the very old and in persons with *multimorbidity*—the most common chronic medical condition—used to define the presence of several comorbid medical conditions. Sometimes causal factors emerge clearly only after a successful intervention, or not at all.

A 78-year-old man with bipolar disorder since age 40 became manic while residing in a nursing home. His situation had been complicated by chronic alcoholism and mild dementia. He had no history of renal disease, and his serum creatinine was normal. After an initial period of manic grandiosity, when he was often seen near a local tavern, he soon deteriorated to crawling on his hands and knees, growling and barking like a dog, and striking out at all who approached him. His lithium level was 1.2 mEq/L with a maintenance dose of 750 mg/day. After reduction in lithium dose to achieve a level of 0.6 mEq/L, the temporary addition of low-dose olanzapine, and close attention to adequate fluid intake and preventing access to alcohol, both the manic syndrome and the superimposed delirium cleared within 3 weeks without the need for psychiatric hospitalization.

Medical causes should be considered whenever an elderly patient develops a new psychiatric symptom or disorder. When psychotropic medications are used, clinicians are advised to carefully watch for the emergence of side effects, allow sufficient time for medications to reach effective plasma and receptor site concentrations, and seek optimal therapeutic outcomes. Particularly with newer-generation psychotropic medications, the differences between "geriatric" and "adult" doses are minimal; careful titration of medications from a low dose to a usual adult dose may be necessary to achieve the desired therapeutic benefit. Close collaboration with other clinicians and caregivers is essential for maximal safety and effectiveness at every stage of treatment. The greatest overall therapeutic gain, at the smallest cost in treatment-related adverse events, is usually achieved by combining psychotherapeutic, social, environmental, and somatic interventions. A rehabilitation model that specifies progress toward attainment of specific goals in stepwise increments is a useful way to frame the ongoing process of care that is often most appropriate for chronically ill older adults.

Table 57.3b–1.
Key Concepts in Psychiatric Care of the Medically Ill Elderly

Decreased organ reserve alters response to illness, treatment, and social stressors.
Comorbid medical illness affects presentation and treatment of psychopathology.
Nondrug treatments are preferable when effective.
Combining treatment modalities optimizes gains.
Comfort, function, and safety are the major goals of treatment.
Coordinated care is better care.
Family and caregivers are essential to patients' recovery and must be involved in the treatment.

An 83-year-old woman with long-standing diabetes and ischemic and hypertensive cardiovascular disease had become bedfast over several years in a nursing home. Two months before psychiatric consultation was requested, she developed a psychotic illness characterized by delusions of being beaten during the night, aggressive outbursts, and pervasive despondency, which led to disruptive nocturnal phone calls to her daughter and to the local fire station. Initial clinical assessment disclosed a moderately severe dementia in a personable but angry and frightened woman, eager for relief. Collateral history confirmed progressive cognitive deterioration over the course of her chronic medical illness, on a background of long-standing dysthymia. Trials of several antipsychotic and antidepressant medications had exacerbated her symptoms. A review of her medications showed that 4 of the 12 drugs she was taking on a regular basis were not essential for her medical management. A head computed tomography (CT) scan showed mild cerebral cortical atrophy, moderate periventricular white matter hyperintensities, and numerous lacunar infarcts. The onset of her psychosis was eventually associated with the initiation of aggressive treatment of arterial leg ulcers, including round-the-clock debridement dressings, oral antibiotics, and narcotic analgesics. Control of her diabetes had deteriorated, and her insulin dose had been increased. Nursing staff members had provided excellent physical care, but were resentful of her racist attitudes and physical aggressiveness, which exacerbated her loneliness and paranoia. Her treatment plan was as follows: discontinuation of night-time debridements (allowing elimination of narcotics for pain), simplification of her drug regimen, preventing nocturnal hypoglycemia, installation of a night light in her room, problem-solving discussions with staff, weekly psychotherapy, and

wheelchair walks, which represented her first opportunity to go outdoors for months. Her distressed phone calls ended soon after. Her mood, sleep, and interactions with staff improved gradually, and her delusions and racist remarks disappeared. Foci of psychotherapy were discovery of traumas she experienced in Cossack Russia, psychoeducation about her delirium, and discussions of her desire to die rather than live in a deteriorated state. She developed a firm therapeutic alliance and regained the dignified and compassionate manner that characterized her response to adversity in the past.

PSYCHIATRIC SYNDROMES IN MEDICALLY ILL PATIENTS

General Principles

Systemic medical illness can mimic or cause almost any psychiatric syndrome, often through more than one pathophysiological mechanism. They can also coexist with psychiatric illnesses that were present prior to the onset of such medical problems. The *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition (DSM-5) further simplified the classification of psychiatric disorders due to medical diseases, dividing them into cognitive disorders, psychotic disorders, depressive and bipolar disorders, anxiety disorders, obsessive-compulsive disorders, catatonic disorders, and personality changes. Changes from DSM-IV-TR include addition of obsessive compulsive and related disorders due to another medical condition, removal of amnestic disorder and sleep disorder due to a medical condition, and specifying depressive or bipolar illness due to another medical condition. Distinguishing a "medical" psychiatric syndrome from a primary psychiatric disorder is often difficult. Clues to the secondary nature of a psychiatric syndrome include partial expression, first onset in old age or a change in clinical features with recurrence of an earlier-onset disorder, lack of family history, a recent change in health or medical treatment, and concurrent cognitive processing deficits that point to impaired brain function.

Psychiatric disorders not caused directly by the effects of a medical illness interact with chronic medical conditions in complex ways. This is best understood for depression in the setting of chronic disease: Depression can amplify somatic symptoms, leading to excess functional impairment and overestimation of disease severity; it decreases motivation and brings about despair about the efficacy of treatment, leading to poor treatment adherence, poor self-care, and finally, poor outcomes. In such cases, consideration should be given to the newly introduced diagnosis in DSM-5, *Psychological Factors Affecting Other Medical Conditions*, denoting psychological or behavioral factors that affect a medical condition adversely. Comorbid medical illness, in turn, can delay or prevent recognition and treatment of psychiatric disorders that might be readily treated in primary care.

Medications as a Cause of Psychiatric Symptoms

Most drugs, in sufficient doses, can affect central nervous system function in a vulnerable host. Drugs affecting the brain, either as their primary site of therapeutic action (e.g., all psychotropic drugs, narcotic analgesics, anti-Parkinson agents, and many antihypertensive agents) or as a secondary site not essential to their therapeutic effect (e.g., digitalis [Lanoxin], cimetidine [Tagamet], glucocorticoids, and nonsteroidal anti-inflammatory agents, such as aspirin and indomethacin), can cause psychiatric symptoms in medically ill elderly patients. As an example, corticosteroids used for the treatment of inflammatory states such as rheumatoid arthritis may lead to depression, while thyroid replacement and even antidepressants can cause anxiety.

Clues to a possible drug-induced psychiatric syndrome include presence of associated systemic or neurological signs and symptoms, such as recent loss of appetite, reduced sense of well-being, blurred vision, slurred speech, dizziness, gait instability, or somnolence. A patient's complaint of being drugged, while occasionally delusional, should always raise suspicion of drug toxicity.

Identifying a drug as the primary cause of a change in mental state can be accomplished in several ways. Some drugs produce classical symptoms in overdose, such as nausea and distorted color vision in digitalis overdose, or tinnitus and dizziness with aspirin. In other instances, a neuropsychiatric syndrome results from pathophysiological mechanisms common to a number of drugs with very different primary indications but similar pharmacological properties. An example is the central anticholinergic syndrome, characterized in its most extreme form by tachycardia, pupillary dilation, visual hallucinations, and delirium in patients receiving medications with known affinity for muscarinic receptors, such as tricyclic antidepressants (TCAs), but also caused by pharmacodynamic interactions of a vulnerable brain with multiple drugs, each having weak individual antimuscarinic actions.

Blood levels often don't help in determination of whether a medication is causing a neuropsychiatric problem, except for the relatively few drugs readily measured in serum or plasma and for which therapeutic ranges are well established. However, medication toxicity may occur in older adults without elevated plasma levels, and improvement when a suspect drug is decreased or discontinued is the most crucial test of its role in pathogenesis.

An 87-year-old woman came to a psychiatrist with a mild depression associated with a recent cross-country move and hepatic cirrhosis of unknown etiology with mild renal insufficiency. Treatment was initiated with fluoxetine 20 mg/day and supportive psychotherapy. Over the ensuing 9 months, she was hospitalized several times with intractable vomiting for which no cause was found. When she again presented with the same picture and, now, ascites and severe demoralization, a fluoxetine level was obtained. The combined concentration of fluoxetine + norfluoxetine was within the expected range at 138 ng/mL (normal range 100 to 930 ng/mL). Discontinuation of fluoxetine eliminated the nausea, vomiting, and demoralization within a

week.

Drug and Disease Interactions

Psychiatric symptoms arising from the interactive effects of medications and underlying disease may be severe enough to lead to psychiatric hospitalization. In a series of nearly 16,000 consecutive adult psychiatric admissions, about 1 percent was attributable to adverse drug reactions. Elderly patients were most often affected, particularly those with neurological disease or those receiving multiple medications. Nearly 30 percent of elderly persons develop delirium after hospitalization for acute medical illness. Particularly at risk are those with cognitive impairment at the time of admission, the very old (>80 years), patients with fractures or symptomatic infection, and those receiving neuroleptics or narcotics. Early detection of mental disorders due to drug–disease interactions depends on familiarity with the common medical illnesses of older people, their potential impact on brain function, and the pharmacological properties of drugs used in medical management and their special potential for interactive effects. Table 57.3b–2 shows examples of some drug–drug and drug–disease interactions that may be encountered in clinical practice.

Differential Diagnosis

Differential diagnosis of psychiatric syndromes in the medically ill older patient begins with a mental status examination that incorporates a careful assessment of several cognitive dimensions, including attention and orientation (both impaired in delirium), memory and learning, and executive control function, the ability to understand, organize, and carry out goal-directed cognitive and motor sequences. Thorough collateral information and an evaluation of the patient's past and current medical conditions, emphasizing disease of vital organs and painful conditions are essential. A complete medication history and a focused physical examination are important, including at a minimum a review of vital signs and a screening neurological examination to check for abnormal gait or movements, focal areas of weakness, or abnormal deep tendon reflexes. The judicious use of appropriate laboratory tests may clarify the possibilities in complex situations (Table 57.3b–3). Finally, the gathered data can be synthesized into a list of possible differential diagnoses and probable causes. For example, in a patient with blood pressure in the low-normal range, pitting ankle edema, apathy, weakness, and depressed mood, who has been taking a high-potency diuretic for many years, an electrocardiogram (ECG) and echocardiogram may be needed to determine whether CHF or the diuretic is the most salient contributor to the depressive syndrome, whether it is safe to try discontinuing the diuretic, and what special precautions are required in prescribing an antidepressant.

In doubtful cases, where the differential diagnosis includes a primary psychiatric disorder not caused by physical effects of illness, the electroencephalogram (EEG) may be helpful. The EEG in patients with psychiatric syndromes due to metabolic, toxic, or circulatory causes typically shows diffuse slowing, although it cannot implicate specific diseases or drugs, and the interpretation may be confounded when an underlying dementing illness is present. An increase in dominant frequency accompanying recovery lends weight to the formulation.

Table 57.3b–2.
Examples of Drug–Drug/Drug–Disease Interactions in the Elderly

Drug	Indication	Adverse Effect
Cimetidine	Gastroesophageal reflux	Delirium Elevated drug levels (e.g., TCAs, benzodiazepines, neuroleptics, propranolol, theophylline)
Metoclopramide	Gastroparesis	Parkinsonism, inappropriate prescription of levodopa
TCAs	Depression, anxiety, insomnia, chronic pain	Confusion Constipation Tachycardia, cardiac conduction defect, hypotension Altered effect of antihypertensive agents (↑ effect of diuretics, ↓ effect of α_2 agonists, e.g., clonidine)
SSRIs	Depression, anxiety, obsessive-compulsive symptoms	Apathy Insomnia Restlessness, agitation, extrapyramidal symptoms, akathisia Central bradycardia ↑ Potential for diverse drug–drug interactions Destabilization of Coumadin dosing Gastrointestinal distress, diarrhea
NSAIDs	Arthritis	Cognitive impairment ↑ Serum lithium levels ↑ Blood pressure, ↓ hypertension control
Neuroleptics/atypical antipsychotics	Dementia with psychosis	Death ↑ Cognitive impairment Extrapyramidal symptoms Falls Aspiration pneumonia
Ginkgo biloba	Impaired memory, arterial insufficiency	↓ Clotting time Restlessness (high doses)
Glucocorticoids	Inflammation, COPD	Depression, mania, psychosis, cognitive impairment (especially at high doses), insomnia

COPD, chronic obstructive pulmonary disease; NSAIDs, nonsteroidal anti-inflammatory drugs; SSRIs, selective serotonin reuptake inhibitors; TCAs, tricyclic antidepressants.

When a drug is believed to cause or contribute to a psychiatric disorder, all drugs should be sorted into those with a high potential for causing psychiatric symptoms (such as anticholinergic and sedating drugs), those clearly required for the management of serious medical disease, and those that may be modified or stopped on a trial basis. Familiarity with drug side effects, with interactions, and with the pathways for their elimination may assist in the detection of otherwise occult medication toxicity syndromes. Close adherence to the "rule of maximal simplification" of medication regimens can produce clinical improvement in up to half of all psychiatrically disturbed medically ill older persons.

A 73-year-old man with severe chronic obstructive lung disease (COPD) came for the treatment of anxiety with insomnia and severe daytime fatigue. He had been treated with fluoxetine for several years after initial psychiatric evaluation, to which clonazepam had been added by his pulmonologist in doses up to 2 mg/day, without relief. Shortness of breath and activity-induced panic attacks were his most pervasive complaints. Physical examination and imaging disclosed hypertension (190/120), persistent tachycardia, and signs of emphysema but no indication of CHF. A walk with oximetry led to oxygen desaturation from 93 to 83 percent after 3 minutes. His medication list included two antihypertensive agents, only one of which he took, since the other was a diuretic that prevented him from leaving home unless assured of constant proximity to a bathroom. He was also on three scheduled inhaled medications and albuterol when needed (prn). It turned out that he was using his prn albuterol at least 12 times daily, but his technique was ineffective in achieving drug deposition in smaller airways. Pulmonary function tests showed no improvement with bronchodilator therapy. Diagnoses of activity-dependent oxygen debt and albuterol toxicity were made, manifested by treatment-resistant anxiety and hypertension, tachycardia likely to endanger his heart if uncorrected, and severe exercise intolerance. Reduction in albuterol use to four times daily, addition of a second antihypertensive drug without diuretic activity, and prescription of continuous oxygen therapy with judicious oral morphine significantly improved his shortness of breath and activity tolerance, and reduced his blood pressure and tachycardia to safer levels.

Elderly patients who present with previously undiagnosed cognitive impairment warrant aggressive assessment. The most prevalent dementia syndrome among geriatric patients is Alzheimer disease, dementia with Lewy bodies a close second. Other common causes of cognitive impairment in older patients include vascular dementia, frontotemporal dementia, dementia due to Parkinson disease, Korsakoff syndrome, and potentially reversible cognitive impairment (if caused, e.g., by depression, drugs, B_{12} deficiency, and others). (See chapter on Dementias for additional discussion.) A meta-analysis of rates of reversible dementia found 9 percent of cases had a partially or completely reversible etiology, but that a total of only 0.6 percent of cases actually reversed with appropriate clinical management. These figures likely reflect general improvements in geriatric medical care.

Table 57.3b–3.
Practical Evaluation of Psychiatrically Disturbed Medically Ill Older Patients

Establish rapport quickly.

Define the reason for the consultation or visit.

Learn and use a few rapid screening instruments:

- ► Cognition—e.g., Mini-Cog
- ► Depression—e.g., Patient Health Questionnaire-9 (PHQ-9)
- ► Neuropsychiatric symptoms—e.g., Neuropsychiatric Inventory (short version)
- ► Functional ability—e.g., Functional Activities Assessment

Expect more than one psychiatric diagnosis (e.g., delirium plus chronic depression).

Make a complete problem list:

- ► Medical
- ► Psychiatric
- ► Social
- ► Functional

Identify all drugs in current use:

- ► Prescription drugs
- ► Over-the-counter medications
- ► Vitamins, minerals, botanicals, and "natural" remedies
- ► Alcohol and tobacco
- ► Include those used infrequently
- ► Ask patients or caregivers to bring all medications from home in a paper bag for your inspection

Interview capable collateral sources:

- ► Family or other primary caregivers
- ► Primary physician
- ► Staff of residential facilities

For acute exacerbations of chronic problems, ask "why now?"

If problem unexplained, search for:

- ► Undiagnosed medical illness
- ► Drug effect or interaction
- ► Occult injury
- ► Undetected social problem (abandonment, abuse)

Conduct a screening physical examination: vital signs, gait, neurological signs.

Obtain appropriate corroborating laboratory studies; consider broad-spectrum tests to screen for suspected disease, e.g., complete blood count, blood glucose, serum electrolytes, blood urea nitrogen and creatinine, liver enzymes, thyroid-stimulating hormone and free thyroxine, vitamin B_{12} and folate, syphilis serology if indicated, urinalysis, electrocardiogram, chest radiograph, electroencephalography for delirium, erythrocyte sedimentation rate for occult systemic disorders, hip radiographs for a delirious patient who will not walk, neuroimaging (head computed tomography for screening, or magnetic resonance imaging when microvascular disease is a possibility; single photon emission computed tomography for more precise diagnosis).

Completing an evaluation of a patient with dementia often requires two or more visits, particularly for frail elderly patients, as well as a thorough collateral history gathering. Medical illness can interfere with the adequate diagnosis of dementias, or unnecessarily delay antidementia pharmacotherapy. Also often overlooked in the management of demented, medically ill patients is the opportunity to improve cognition by eliminating medications with mild but cumulative central nervous system effects that worsen both cognition and function.

Approach to Treatment: Overall Recommendations

When treating an older adult with medical comorbidities, some overarching principles are worth calling to mind. Where appropriate, involve family and other supports. Knowing the extent, strength, and reliability of a patient's social support system affects treatment planning, and can aid in predicting what additional resources may be needed and how likely a given treatment is to be successful. Likewise, economic status can shape treatment and its delivery, especially in older adults managing multiple medical conditions, medications, Medicare copays, and hospital bills. The living environment should be safe and adequately accessible to needed medical resources, and should be addressed directly in a multidisciplinary treatment plan if it is not. Do not neglect the patient's preferences in treatment, but continually assess for medical decision-making capacity.

Always consider medical illnesses and treatments as potential causes of or exacerbating factors of the symptoms one is addressing, and work collaboratively with other health care professionals involved in the care of an elderly patient with psychiatric and medical comorbidity. See the end of this chapter for more discussion on the collaborative models of care.

It is wise to consider available nonpharmacological options for treatment of a given condition first, and weigh their potential appropriateness prior to turning to pharmacological strategies. These options could involve a variety of interventions from individual therapy modalities to structured activity and physical exercise, and improved caregiver communication.

Prescribing medications in the elderly poses a particular challenge, for reasons such as polypharmacy, pharmacokinetic (e.g., increased drug volume, decreased renal and hepatic clearance) and pharmacodynamic (e.g., increased sensitivity to certain classes of drugs) changes that occur with aging, overall reduced functional reserve that limits compensatory homeostatic mechanisms, and underlying medical and neurological diseases that impact drug metabolism and tolerance, variables that should be viewed as a constantly evolving process requiring close continued attention.

Polypharmacy is a significant concern in this population. A survey of community dwelling older adults (between the ages of 57 and 85 years) reported that 5 or more prescriptions were used by 29 percent of the overall group surveyed and this number increased to 36 percent in the group aged 75 to 85 years. In addition, almost half of the prescription users also took an over-the-counter medication on a regular basis. Herbals and dietary supplements use are also on the rise in the elderly. The Beers Criteria is a useful guide to assess inappropriate drug prescribing in the elderly due to ineffectiveness or high risk of adverse events.

These are the general guidelines that the authors recommend when prescribing medications for an elderly patient:

- Keep goals of care and patient's current health and cognitive status in mind.
- Before adding a medication, carefully review the current regimen, simplify the dosing schedule, and discontinue any unnecessary medications. Consider the cost of medications.
- Change one medication at a time, as much as possible.
- Be patient and use a slow pace of titration.
- Focus on potential drug–drug interactions and keep anticholinergic load to a minimum.
- Be equally concerned about overprescribing and underprescribing.
- Educate the patient and caregivers about side effects and expected improvements.
- Periodically review effectiveness, need for dose adjustment, or continued use. Standardized scales are invaluable feedback devices.
- Be vigilant of treatment adherence.
- Do not neglect cognition, as it may be a target of treatment and a factor complicating treatment of medical and psychiatric conditions.

PSYCHIATRIC SYMPTOMS ASSOCIATED WITH SPECIFIC MEDICAL DISORDERS

Vascular Disease

Vascular disease is being increasingly recognized as a cause of late-onset psychopathology as well as a factor exacerbating pre-existing disorders such as recurrent major depression, bipolar disorder, and Alzheimer disease. Ischemic and hemorrhagic strokes cause a variety of disturbances of behavior and emotionality related both to effects of the lesions on brain function and to psychological problems in adaptation to disability. Depression is often comorbid with vascular disease, occurring in roughly one-third to a half of all patients following an acute cerebrovascular accident. Though it has long been thought that lesions in the anterior left hemisphere are particularly predisposed to depression, this has more recently been disputed. The most recent review of available evidence indicates that right hemisphere lesions correlate with depression in the acute stroke recovery period, but that lesion location is otherwise not predictive in this regard. Persistence of milder depressive syndromes after the acute phase has been associated with factors related to functional impairment, including cognitive deficits, weak social supports, and inadequate living environment. The link between mood and vascular health is notable for its converse relationship: depression is an independent risk factor for stroke.

The importance of cerebrovascular disease in psychiatry extends well

beyond poststroke depression. The vascular depression hypothesis supposes that cerebrovascular disease may be etiologically related to geriatric depressive syndromes. The clinical presentation of this entity is characterized by cognitive deficits, psychomotor retardation, lack of insight, and disability disproportional to depression severity, as well as a suboptimal response to antidepressants. The most common neuropsychiatric outcomes of cerebrovascular disease are depression, anxiety, fatigue, and apathy. Less common but equally distressing are emotional lability, personality changes, psychosis, and mania. There is often no clear relation to location of the brain lesion, and management is challenging given lack of evidence base.

Cerebrovascular disease can cause psychomotor slowing, impaired executive function, and significant functional dependency even in the absence of depression. Associated clinical features include encoding-based memory deficits and impaired gait and urinary control. Neuroimaging studies, especially those using magnetic resonance imaging, typically demonstrate ischemic changes in subcortical and deep white matter and/or lacunar infarcts in the thalamus, the basal ganglia, and sometimes the upper brainstem. Interference with the function of frontal-subcortical circuits appears to be a common underlying anatomical feature of lesions that cause depression and executive dysfunction, and can be seen as a partial disconnection syndrome. Many patients with subcortical vascular disease have minimal or no discernible mood change, showing instead only cognitive, motivational, and neurological signs (often quite subtle). A diagnosis of vascular neurocognitive disorder may be appropriate in such cases, but minor depression may also be diagnosed when apathy and disengagement are functionally impairing. Vascular brain diseases can be caused by intrinsic disease of small and/or large vessels supplying the brain, due to atherosclerotic narrowing, occlusive plaque emboli, medial (muscular) hypertrophy, vascular amyloidosis, and cardiogenic causes. Among the latter, several pathogenic mechanisms are recognized: recurrent or severe hypotensive episodes and arrhythmias, CHF, and complications in patients after coronary artery bypass grafting. A third of older adults who undergo such a procedure develop permanent cognitive impairment, the causes of which are still debated. All physicians caring for such patients, including psychiatrists, need to be alert to the potential to improve long-term outcomes with aggressive management of vascular risk factors.

In fact, prevention of further vascular insults should be a key component in the psychiatric treatment plan of patients suspected of having vascular psychopathology. Relevant measures include fasting lipids, homocysteine, vitamin B_{12}, and folate levels; a careful history and ECG to identify cardiac arrhythmias or, if already recognized, coordination of care and education of other physicians about the possible cognitive consequences, monitoring serial blood pressure measurements to identify

hypertension and hypotensive episodes. Appropriate therapeutic steps can be taken after identification of patient's particular risks, such as prescription of folate/vitamin B_6/vitamin B_{12} combinations, antihypertensives, antithrombotic agents, or lipid-lowering drugs. Depression related either to an acute cerebrovascular accident or to cerebrovascular disease is responsive to pharmacological treatment with antidepressants, but a ceiling effect short of full recovery is likely and may call for adjunctive therapies, particularly stimulants such as methylphenidate or amphetamines, and psychosocial measures to provide activity structure and supportive or problem-solving psychotherapy. Antidepressants are also worth considering for their preventative role. Treatment of nondepressed stroke survivors with selective serotonin reuptake inhibitors (SSRIs) has been shown to significantly lower the incidence of poststroke depression, and has now been incorporated into standard practices. When the executive dysfunction or gait impairment is significant, extra caution is needed in choosing psychotropic medications. A broad treatment strategy is needed for optimal results: securing structure and support to compensate for defective executive function; supervised exercise to improve and maintain gait and balance; and management of frequently associated urinary incontinence.

Heart disease is a well-recognized cause of depression and anxiety. Medical and psychological factors both play a part, and a bidirectional causal relationship exists. Depression is an independent risk factor for coronary heart disease and myocardial infarction, and untreated or inadequately treated major depression is associated with adverse outcomes in individuals with known coronary disease. Effective psychiatric treatment often requires both psychotherapeutic and pharmacological modalities. SSRIs are usually the drugs of choice. They may cause bradycardia, by a central medullary mechanism, but lack primary cardiotoxicity, and are considered safe and effective in most patients with heart disease. Patients who fail to respond to SSRIs and don't have cardiac conduction defects or CHF can generally be safely treated with other antidepressants such as bupropion, mirtazapine, duloxetine, or venlafaxine. While electroconvulsive therapy is an undisputed myocardial stressor and should be avoided in the acute post–myocardial infarction period, treatment can be given without major complications if anesthesia and treatment are carefully individualized.

A 72-year-old man had a history of recurrent major depression since early adulthood as well as coronary heart disease. Depressive episodes were seasonal and consisted of low mood, anergia, amotivation, and increased appetite. With treatment with an SSRI and seasonal phototherapy, he suffered no major episodes in a 3-year period of relative stability. In January, he suffered a cardiac arrest, and had a prolonged course of cardiopulmonary resuscitation complicated by a difficult intubation in the field. During the hospitalization that followed, he was diagnosed with a diffuse anoxic brain injury, primarily affecting the subcortical watershed areas of the right

hemisphere greater than the left. After his acute rehabilitation, he was left with residual left-sided weakness and new psychiatric complaints. He described his mood as "blah," couldn't find anything that interested him, and had profound psychomotor retardation, exceeding his physical limitations. He expressed worry about becoming like his father who had had Alzheimer disease. His wife took over management of his medication and finances, and he wasn't allowed to drive. He reported that his previously effective antidepressant medication was no longer helping. In consultation with his cardiologist, low-dose methylphenidate was added to his medication regimen. Six weeks later, his affect was brighter, his speech and movements were less labored, and he had been able to dabble in his old hobby of collecting military memorabilia.

Parkinson's, Alzheimer's, and Other Neurological Diseases

Depression occurs with disproportionate frequency in Parkinson disease, affecting up to half of all patients. This depression is believed to be caused, in large part, by the neurodegenerative process itself; its presence is associated with evidence of damage to noradrenergic and serotonergic systems, and is thought to identify a group of Parkinson patients with a more rapid course of disease. Nevertheless, antidepressant drugs or other treatments can be effective. Venlafaxine (Effexor) and pramipexole (Mirapex) have been found to be clinically useful; the latter especially so for targeting motor symptoms concomitantly. Though evidence is lacking for use of specific SSRIs, they are commonly used and relatively safe. TCAs may be effective, but are less commonly used due to concerns about adverse effects in older patients. Selegiline (Deprenyl), an anti-Parkinson agent that selectively inhibits monoamine oxidase B, may be a useful addition to treatment of depression in Parkinson's. Cognitive-behavioral therapy (CBT) has also proven effective. Hallucinations are both a feature of later-stage Parkinson disease and consequence of anti-Parkinson drugs, and may respond to lowering the doses of these medications. If that is not feasible, atypical antipsychotic drugs with less impact on nigrostriatal dopaminergic pathways can be helpful. Clozapine (Clozaril) and quetiapine (Seroquel) are the most extensively studied, and though clozapine is consistently efficacious, quetiapine is usually prescribed first because of the risk of agranulocytosis and the requirement for frequent blood monitoring with clozapine. Use of ziprasidone (Geodon) is supported by limited but promising evidence. Motor deterioration outweighing impact on psychosis has been observed with olanzapine (Zyprexa), risperidone (Risperdal), and aripiprazole (Abilify). The majority of depressed Parkinson patients also have cognitive deficits, which may occur as an independent neurocognitive disorder unrelated to depression. Major neurocognitive disorder due to Parkinson disease commonly affects learning and memory, initiative, planning, and capacity for independent functioning, in excess of the constraints imposed by the motor disability alone. These cognitive deficits tend to persist after successful treatment of any mood disorder, and are not substantially improved by anti-Parkinson agents. The currently available evidence supports the use of cholinesterase inhibitors in these patients,

with a positive impact on global assessment, cognitive function, hallucinations, behavioral disturbance, and activities of daily living rating scales. The rivastigmine (Exelon) transdermal patch is better tolerated than oral formulations, though exacerbation of tremor remains a concern. Response to memantine (Namenda) in this disorder has been disappointing.

In Alzheimer disease, depression affects up to one-third of patients, especially in milder stages, and may be the major presenting problem. In recent years, the cumulative evidence is found to be insufficient to support the use of antidepressants in this population; they also lead to increased rates of adverse events, including falls. Trials of antidepressants may be best reserved for those with moderate or severe depression, and risks and benefits should always be carefully weighed. Apathy and other neuropsychiatric manifestations of major neurocognitive disorder due to Alzheimer disease may be modestly improved by cholinesterase inhibitors. With the progression of dementia, classical presentations of depression become less frequent, while agitation, psychosis, and aggressive behavior with or without dysphoria become more common. Mixed behavioral syndromes are the rule in middle- and later-stage dementia. Assaultive or threatening behavior demands thorough medical and socio-environmental assessment to identify precipitating and promoting factors, but pharmacological interventions are often required. Second-generation antipsychotics are commonly used off-label in these cases, a controversial practice due to the black box labeling by the Food and Drug Administration (FDA). Recent reviews conclude that use of these agents confers modest benefits for short-term treatment of aggression and psychosis as well as a small but significant benefit on caregiver burden. This is offset, however, by the substantial risk of serious adverse events, including a 1.5 to 1.8 times increased mortality risk. Trials of valproate for these dementia-related behaviors have been disappointing. Existing evidence for use of SSRIs in this condition is promising but limited.

A 90-year-old male Holocaust survivor with major neurocognitive disorder due to Alzheimer disease was brought to a psychiatrist with increased tearfulness and speaking about the loss of his family in concentration camp settings. His family reported that this was a change for him, and that he had not previously focused so much on these troubling memories. They also reported that the television in the home was often left on to a channel covering World War II and other historical topics, at the caregiver's preference. The family was instructed to ensure different programming was on television, and as a result, the caregiver spent more time interacting with the patient. His mood improved. A year later, the man returned, now with aggressive behaviors of grabbing and yelling out, especially around toileting. The patient was assessed for any physical cause of discomfort, but none could be found. The caregiver received additional information on communicating with individuals with Alzheimer disease, and providing distractions and rewards around these difficult times. The patient's behavior improved only minimally with these interventions, and a low dose of quetiapine was started. The behaviors improved, and the psychiatrist was able to gradually taper off the medication after 3 months.

Posttraumatic and postanoxic encephalopathies, hydrocephalus, central nervous system tumors and infections, multisystem atrophy, Shy–Drager syndrome, Huntington disease, demyelinating diseases such as multiple sclerosis, idiopathic inflammatory diseases such as systemic lupus erythematosus and temporal arteritis, and occupational neurodegeneration such as chronic organic solvent encephalopathy may all be associated with depressive, manic, delirious, and psychotic syndromes. For depressions in patients with these disorders, antidepressant trials are generally warranted, though drug intolerance may be problematic and require the use of innovative and combination treatment strategies. The management of other psychiatric complications is less well established.

Chronic Respiratory, Metabolic, and Endocrine Disorders

COPD is the most prevalent of the chronic respiratory diseases of the second half of life. Cigarette smoking and, in the developing world, exposure to biomass fuels, are significant risk factors. COPD is strongly associated with depression, anxiety, and, in advanced cases, a special form of panic disorder precipitated by exertion. Over 40 percent of patients with COPD suffer from depressive symptoms. Smoking is a shared risk factor in depression and COPD and the rates of depression are even higher in COPD patients who are actively smoking. Untreated depression in COPD worsens overall outcomes with more COPD exacerbations per year, contributes to persistent smoking, and has been associated with increased risk of death, poorer social functioning, increased physical disability, poor adherence to treatment, and health care overutilization.

Psychological complications are common, and include feelings of hopelessness and helplessness that contribute to impaired treatment and rehabilitation adherence, anxiety-related dyspnea and secondary panic symptoms fostering further inactivity, fear of sudden death by asphyxiation, intense self-blame in patients whose illness is related to smoking and social isolation due to shame, low mood, and lack of motivation. For milder dysphoric states not meeting criteria for a depressive diagnosis, a multidisciplinary rehabilitation and counseling approach improves mood in many patients. In such cases, the efficacy of antidepressants in COPD has not been demonstrated either for alleviation of depression or improvement of dyspnea. However, available evidence suggests that, in patients with syndromal depression and absence of left ventricular disease or hypercapnia, antidepressant pharmacotherapy is beneficial for depression, anxiety, and somatic symptoms. A risk and benefit analysis regarding the use of antidepressants needs to be considered in each patient. Comprehensive nonpharmacological interventions for anxious depressed patients with COPD provide individualized education about the disease, its management, and adherence to recommended treatment, teach patients to modify their symptoms with

appropriate techniques such as breathing retraining, and offer supervised physical training. The use of behavioral activation, CBT and collaborative care approaches should also be considered. Patients with pulmonary disability and those who require home oxygen therapy may benefit from self-help groups. Patients with lung cancer may develop psychiatric complications similar to those occurring in COPD; in addition, paraneoplastic psychiatric syndromes with psychotic, cognitive, and/or affective symptoms have been reported.

Deterioration in pulmonary or cardiopulmonary status in patients with COPD can lead to acute and chronic encephalopathies, to sustained hypoxemia and hypercapnia, or to the central nervous system effects of drugs used in management (oral or inhaled bronchodilators, high-dose prednisone, cough or dyspnea suppressants, or benzodiazepines). Psychotic states may occur, usually related to acute impairment of brain function by severe exacerbation of illness. Prior to the establishment of guidelines for oxygen therapy in COPD, neurocognitive impairment was thought to be quite widespread, but recent studies questioned the prevalence of cognitive disorders due to COPD prior to severe and end-stage disease.

Chronic renal and hepatic disease can lead to well-known delirium states due to their detrimental effects on nutrition, metabolism, clearance of toxic metabolites, and drug disposition. In addition to these transient alterations of mental function, the prevalence of other psychiatric disorders in these conditions approximates those for chronic disease in general. Pharmacological treatment considerations include evaluating the effects of disease on effective half-life and plasma drug levels and adjusting prescribing accordingly. More specifically in renal disease, chronic hemodialysis has been associated with neurocognitive decline, but improvements in this area have occurred since aluminum was eliminated from dialysate. There is also substantial comorbidity with depression. For patients with end-stage renal disease, suicide risk appears to be less elevated than previously thought. However, complex end-of-life issues associated with dialysis termination remain a clinically challenging issue where nephrologists, palliative care specialists, and psychiatrists can work closely together with patients. With hepatic disease, an area of notable concern is the effect of hepatitis C and its treatment which is associated with depressive states.

Diabetes is implicated in both depression and cognitive impairment. Depression rates are at least twice those of the general population and can be associated with poorly controlled hyperglycemia, lessened adherence to medications and reduced self-care, all of which can contribute to increased diabetic complications and elevated mortality rates. In short-term studies, treatment for depression improves both glucose control and quality of life. In epidemiological studies, diabetes (in its more common type 2 form in particular) has been found to be associated with dementia, which is

exacerbated by the associated increased rates of comorbid cardiac and vascular disease. However, the pathogenesis of this neurocognitive risk is not yet understood. It is believed that the adverse effects of diabetes on cognition can be modified by optimal glucose control, with longitudinal studies in this area continuing.

Hormonal changes are useful to consider in clinical assessment in the medically ill geriatric population. For example, even subtle thyroid abnormalities can play a role in mimicking or worsening psychiatric syndromes. Another area of interest and continuing investigation relates to changes in sex hormone levels which can be deranged with illness and evolve substantially with age. There is suggestive evidence that sex hormone dynamics can have effects on both general well-being and the presentation of depression and related symptoms.

Chronic Pain and Musculoskeletal Disorders

In older patients, the prevalence of chronic pain ranges between 36 and 80 percent, depending on the setting. Patients suffering from pain often restrict activity to prevent further deterioration and to avoid exacerbation of pain, which leads to deconditioning, isolation, and worsening disability. Over half of older adults with depressive disorders have chronic pain, with depression more likely to be present in patients with multisite pain and higher pain intensity and frequency. Older patients who have both chronic pain and depression are more likely to experience insomnia and suicidal ideation; they have a higher level of impairment and poorer quality of life.

The causes of chronic pain in older adults include musculoskeletal disorders, neuropathies, and pain related to cancer, with musculoskeletal disorders being the predominant cause. These include fractures, osteoarthritis, rheumatoid arthritis, microcrystal disorders, myositis, vasculitis, and other arthropathies. Aging musculoskeletal tissues are less resilient due to the loss of ligament elasticity, redistribution of fat, osteoporosis, and subsequent decrease of bone strength. *Sarcopenia,* the loss of skeletal muscle mass, is a nearly universal aging phenomenon, contributing to disability and functional loss in older patients. The elderly are also at higher risk for infections that may affect the joints due to the comorbid predisposing disorders (peripheral vascular disease, diabetes mellitus), immunosuppression, and surgical interventions that are prevalent in this population. Falls are both a risk factor for incurring injuries that may result in chronic pain and a consequence of chronic pain.

Sixty to 70 percent of all joint disease in the elderly is osteoarthritis, often comorbid with regional pain syndromes. A common complaint in ambulatory patients over the age of 75 is chronic low back pain, while knee osteoarthritis remains the most common cause of disability in older persons. Rheumatoid arthritis is to blame for 10 percent of all joint pain complaints, with patients who are taking biological disease-modifying

drugs being at highest risk for depression, suicidal ideation, and poor quality of life. The number of older adults with arthritis or joint symptoms is projected to be around 41.1 million by the year 2030.

Chronic pain is undertreated in older adults for several reasons, among them poor access to care, patients' concern about the meaning of pain and cost of its diagnosis and treatment, and the patients' and providers' belief that pain is concomitant with aging, although there's evidence that a significant portion of older patients with chronic pain hope that physicians will pay more attention to their pain complaints. Physicians may fail to assess the patients' pain, or believe in it when it is reported. There's also a persistent fear that prescribing opioid medications, often necessary for control of pain when it doesn't respond to other agents, will result in addiction. A recent survey of hospitalized geriatric patients found prevalence of pain around 67.3 percent but only 49 percent of these cases were treated; 75 percent of patients felt their pain medication was of low or no efficacy. Patients with cognitive impairment, particularly those with severe dementia and inability to communicate, are at risk for undertreatment of pain, due to failure to identify pain and use appropriate tools for measuring pain.

Treatment of chronic pain in older adults is a complex undertaking, requiring multidisciplinary approach. A comprehensive assessment should include a thorough history of pain, past medical history, history of substance abuse, physical and mental status examination, thorough review of medications, and examination of how pain affects the patient's quality of life and function. Self-report of pain is the gold standard for evaluation of pain in any patient, although this may be challenging in cognitively impaired, noncommunicative patients.

Pharmacological treatment should be individualized for each patient; options for treatment of pain exist but each one has its pros and cons. While opioids are an effective way to treat noncancer and possibly neuropathic pain, their duration of action in older adults is prolonged, and long-acting medications may be of particular concern. The consensus statement by the expert panel from the International Forum on Pain Medicine recommends that doses of opioids should be lower, and a longer time interval should be used between doses. Further, consideration of any concomitant medications affecting respiratory function and pulmonary conditions should occur, so as to avoid problems with hypoventilation and respiratory depression. The use of opioids in patients with history of substance abuse should be carefully monitored. Further, it should be remembered that opioids may cause depression, the risk of which increases with the duration of exposure to opioids. TCAs, gabapentin, and pregabalin can manage neuropathic pain, but they are sedating. TCAs are also anticholinergic with risks of cognitive impairment, urinary retention, constipation, visual changes, and cardiac conduction disorders, among

others. Some NSAIDs may cause gastrointestinal bleeding and discomfort.

Nonpharmacological treatments should be utilized whenever possible: these include biofeedback, physical therapy, chiropractic, acupuncture, aquatic therapy, weight control, corticosteroid joint injections, and surgery if indicated. Psychoeducation, psychotherapy (particularly, CBT and/or mindfulness) and psychiatric medications such as antidepressants should be recommended if depression and/or anxiety symptoms are comorbid with pain. Due to increasing interest regarding assessment and treatment of pain in older patients, recent years have seen publications of excellent guidelines. The authors recommend "*Guidance on the Management of Pain in Older People*" published in *Age and Ageing* 2013;42:i1–i57, by the British Geriatrics Society.

A high index of suspicion regarding the presence of untreated pain should be applied when working with patients with dementia. A change in their behavior, including aggression and agitation, could be related to pain. Other pain manifestations may include changes in facial expression, verbalizations, body movements, as well as onset of crying, irritability, and increased confusion. Using scales developed specifically for evaluation of pain in patients with cognitive impairment will improve outcomes. An example of such a scale is the Pain Assessment in Advanced Dementia (PAINAD), applicable in long-term care and acute geriatric care settings. Investigation of possible causes of pain should be undertaken with assistance of collateral information, full medical history, and physical examination. Finally, a trial of an analgesic resulting in improvement in behavior may confirm the diagnosis of pain.

An 81-year-old man with previously diagnosed dementia related to Alzheimer disease sustained a fall resulting in a right elbow fracture. After the orthopedic surgery, he had a difficult postoperative course complicated by severe confusion and disrupted sleep–wake cycle. A geriatrician who examined the patient found a confused, thin, elderly man with a marked level of rigidity and cogwheeling in the left arm (right arm could not be examined due to pain). He was disoriented and his level of consciousness fluctuated markedly during the interview. He had difficulty focusing on the content of questions, and was unable to state the reason why he was in the hospital. The patient's son reported that his father's memory and executive functioning have been declining for over 4 years, and the course of this disease was marked by visual hallucinations, usually of animals like cats, which were distressing for the patient and for his family. Rigidity and falls became a problem 2 years after the onset of cognitive impairment. After the surgery, the patient was described as markedly below his recent cognitive baseline. The geriatrician diagnosed the patient with delirium, related to the recent surgery and the number of sedating medications on board, including morphine, fentanyl, zolpidem, and lorazepam. Discontinuation of the above medications and initiation of scheduled acetaminophen and tramadol for pain led to significant improvement in the patient's mental status and eventual discharge from the hospital to a rehabilitation facility. The geriatrician subsequently reexamined the patient and diagnosed him with dementia with Lewy bodies rather than Alzheimer disease, recommending avoidance of long-acting opioids and sedating medications, as well as initiation of rivastigmine patch to address cognitive impairment. The patient also worked with a physical therapist to improve balance and coordination, and prevent falls.

SPECIAL ISSUES IN THE CARE OF ELDERLY PATIENTS

Impaired Vision and Hearing

The importance of sensory impairments in patients who develop psychiatric disorders in later life justifies special attention to these issues during evaluation and treatment. Degenerative diseases of the eye and ear impair responsiveness to the interpersonal and social environment, increase feelings of vulnerability among older persons, and cause hallucinations in the affected sensory modality in some individuals. For example, visual hallucinations in the absence of psychiatric disease (Charles Bonnet syndrome) occur in 10 to 50 percent of patients with low vision, a potentially distressing experience. In addition, deafness originating in childhood disease but exacerbated by presbycusis appears to be a risk factor for late-onset paranoid illness, and hearing and visual loss both contribute to the onset and persistence of depression. Patients with hearing loss are at higher risk for accelerated cognitive decline. Cognitive impairment is also frequently present in patients affected by glaucoma, which may jeopardize their adherence to treatment and cause otherwise preventable vision loss. Treatment of eye disease, improvement in corrective lenses, appropriate prescription of hearing aids, and attention to environmental lighting and sound may enhance the efficacy of specific interventions for psychiatric disorder in sensory-impaired patients.

Falling

Falling and fall-related injuries occur in 30 percent of older adults living in the community and in more than 50 percent of those living in nursing homes each year. According to the recent data from Centers for Disease Control and Prevention, the number of deaths related to falls among older adults increased by 7,000 from 2005 to 2013. Complications of falling include hip fractures (25 per 1,000 women over age 85), other fractures, lingering painful soft tissue injuries, and head injuries with behavioral complications. Such injuries are widely feared by the old as harbingers of death. Risk factors for falling include conditions that cause postural instability, general muscular deconditioning, painful or deforming food conditions, visual, auditory, vestibular, proprioceptive, and central neurological impairments, decrease in cognition and speed of processing information, depressive symptoms, environmental hazards, inadequate ambulatory supports, and all classes of psychotropic medications. In hospitalized patients, the presence of mental status deficits is one of the strongest risk factors for falls. Overall, the use of medications with sedative effects is the single most powerful predictor of falling in the elderly, all other things being equal. These include drugs used for the management of insomnia, anxiety, depression, and agitation (prescribed and over the counter), alcohol, and drugs used in the treatment of medical diseases that

sedate by themselves or enhance the sedative effects of other drugs. The presence of multiple risk factors geometrically increases hazard.

Psychiatric complications of falls include acute cognitive disturbance due to head injury, and a common but rarely recognized phobic anxiety state (fear of further falls, which leads to limiting participation in activities of daily living, further deconditioning and curtailment of opportunities for social exchange, and sometimes permanent retirement to a wheelchair). Persistent fall phobias may become less common with aggressive rehabilitation after hip fracture and dissemination of exercise programs for frail elders.

Urinary Incontinence and Urinary Tract Infections

Urinary incontinence affects up to one-third of older adults living in the community and can cause significant suffering and functional impairment, usually worsening as the problem becomes more severe. Notably, urgency urinary incontinence causes a higher degree of disability than stress urinary incontinence. Patients may not mention urinary incontinence to their physician unless specifically asked and may suffer silently, limiting their trips away from home and their opportunities for social activities. There are a number of causes for urinary incontinence, including urinary tract infections, altered mental status, restricted mobility, medications such as diuretics, or more chronic anatomical changes. Urinary tract infections are one of the most common infectious diseases in this age group, and are significantly associated with multiinfarct dementia in very old women. Patients with acute or untreated chronic urinary symptoms should be referred to their primary care physician or to a gynecologist or urologist for a thorough work-up.

Frailty

Frailty is a clinical syndrome comprised of several physical symptoms, including weight loss, exhaustion, poor hand-grip strength, slow walking speed, and physical inactivity, according to the definition most commonly accepted in the literature, by Fried et al. Other commonly encountered symptoms of frailty include weakness, inactivity, balance and gait abnormalities, and decreased food intake. Frailty is generally thought of as an age-related state of vulnerability related to reduced reserve and capacity of the organism to withstand stress and maintain homeostasis, a syndrome characterized by the decline of multiple physiological systems. A number of medical problems can lead to frailty, including chronic infections and inflammatory conditions, cancer, congestive heart failure (CHF), and thyroid abnormalities. Some common psychiatric problems may also manifest as frailty, among them depression, psychosis, grief, and dementia. Frailty is associated with, and predictive of, physical disability, falls,

hospitalizations, and increased mortality. The prevalence of this syndrome in the elderly is between 14 and 24 percent, depending on the definition of frailty, and is generally higher among women and African Americans.

Recent data indicate that cognitive function is significantly poorer in patients who are frail; moreover, a higher level of physical frailty is associated with a faster rate of global cognitive decline. Impaired cognition and frailty may be independent syndromes, but they often overlap and share common risk factors. Those include hormonal changes, nutrition, cardiovascular risks, and mental health and, most importantly, increasing age. The systemic changes related to chronic inflammatory conditions, poor nutrition, and declining function of endocrine and immune systems affect neurons in the hippocampi and microglia, placing the brain at risk for cognitive impairment, delirium, and dementia.

An association of significance also exists between depression and frailty. A recent study of men older than 75 years of age found that those with current symptoms of major depressive disorder had a fourfold increase in mortality, mostly due to comorbidity with frailty. Another cross-sectional study of elderly men found that those with frailty were much more likely to also suffer from anxiety and depression. Among hospitalized patients, frailty is often comorbid with delirium, conferring a poorer prognosis and decreased survival than either condition alone.

The best approach to treatment of frailty and associated mental health conditions starts with a comprehensive geriatric assessment, whether conducted during an acute admission or on outpatient basis. Encouraging studies of using exercise to improve endurance and increase strength and balance in frail patients have been conducted, but there is little known at this time on the potential value of nutritional changes or pharmacological agents. Interventions of most benefit are likely to be multidisciplinary and multidimensional, addressing the whole person rather than a single disease or organ system at a time, and paying attention to the mental health problems of an aging, frail patient.

Overreliance on Drugs for Managing Mental Disorders

Psychiatrists and other physicians tend to underutilize psychotherapeutic techniques in the care of sick or frail older patients. Judiciously prescribed medications are likely to remain the mainstay of treatment with these patients. However, care of psychiatric illness can be greatly enhanced by a psychotherapeutic orientation, and there's evidence to suggest that older adults prefer counseling and psychotherapy to pharmacotherapy. While access to formal office-based individual psychotherapy is often a limited resource, there is accumulating evidence that brief therapeutic interventions, delivered in the primary care setting, can provide considerable benefit and are easier to access for these patients. As mental health care becomes more available in an integrated care setting,

psychotherapeutic interventions may prove highly efficacious as a complement to, and at times an alternative to, long-term medication use. For example, a recent review examining prevention of major depressive disorder among older adults with subsyndromal depression found psychotherapy to be a safe and cost-effective method to reduce the burden of the disease on the patients and to the society. Psychotherapy could potentially be of benefit in reducing the risks of adverse effects from medication overuse and polypharmacy as well. Brief structured psychotherapies can be delivered individually, with key family members, or in groups in both specialty mental health and primary care settings. More broadly, the psychiatrist's use of psychotherapeutic techniques from diverse orientations can include the capacity to develop emotional connections with patients and empathic integration of the needs of the patients, their families, and professional caregivers into a comprehensive and workable treatment plan. In this way, elements of psychotherapy can make substantial contributions both to medical management and to patients' comfort and capacity to persevere in the face of demoralization from serious and ongoing illness.

A HEALTH-SYSTEMS APPROACH TO THE CARE OF THE MEDICALLY ILL GERIATRIC PATIENT

Older adults are less likely than younger patients to seek and receive specialty mental health services and, when referred by their primary care providers, are less likely than younger patients to follow through on such referrals. Older adults, their family members, and their physicians often attribute enduring but treatable psychiatric symptoms to "normal aging," stressful life events, grief, social isolation, losses, financial difficulties, care-taking responsibilities, dementia, or medical illness, resulting in hopelessness and therapeutic nihilism in patients and providers. Older adults may also fear stigmatization associated with mental disorders or psychiatric treatments. As a result, the majority of this subset of population receives care for mental and emotional problems exclusively in the general medical sector, and many who could benefit from treatment of psychiatric disorders do not receive it. It is worth noting that Caucasian men over the age of 85 are at highest risk for suicide in the United States, primarily due to their use of firearms during suicide attempts, and that many older adults visit primary care providers in the weeks leading up to the suicide attempt. Primary care visits should be an opportunity to intervene; however, there is evidence that depression in older adults is often unrecognized in this setting, which is particularly problematic for men and members of ethnic minority groups. Recent data indicate that accuracy of diagnosis of depression in primary care usually improves with repeated visits, and it is recommended that those patients suspected to have depression should be reassessed rather than have one-off visits; use of rating scales can also aid

the diagnosis.

Even if depression is diagnosed, only a minority of patients receive successful treatment despite robust evidence for the efficacy of treatments such as antidepressant medications and psychotherapy in old age. In recent years, the rate of antidepressant prescribing in primary care has increased substantially, but many depressed older adults are started on an antidepressant with little follow-up. As a result, many stop taking medication prematurely, stay on inadequate doses for too long, or do not change medication or receive psychotherapy if an initial trial of an antidepressant fails.

Despite most practice guidelines recommending that primary care providers should play a central role in the detection of cognitive impairment and dementia, there is a concerning lack of diagnosis of dementia in this setting, with rates of underdiagnosis ranging between 50 and 80 percent. Thus, patients with declining cognitive ability, usually in great need of treatment and advice on planning for the future, are not able to take advantage of interventions available to them. The lack of diagnosis is partly due to lack of recommendations for systematic screening of all at-risk patients, and partly due to lack of comfort on the part of primary care physicians when it comes to mental health in geriatric patients, since their training does not emphasize geriatrics or mental health. Further, when dementia is diagnosed, the diagnosis is often undisclosed to patients and their caregivers, and treatment is not offered—largely due to providers' lack of confidence and comfort with the pharmacological options available.

Older adults may have difficulty accessing specialty mental health care, especially when it is provided only in off-site specialty clinics. This problem is only compounded by the dearth of geriatric psychiatrists: by the year 2030, there will be only one such specialist per 20,195 older Americans. Some improvements in access have been made in the recent years, such as the Medicare Improvement for Patients and Providers Act of 2008 reducing the copayment for outpatient mental health services to 20 percent, down from 50 percent, but barriers in private insurance plans regarding psychiatric medications and mental health services such as psychotherapy still exist.

Efforts to improve care of common mental disorders in primary care have included training of primary care providers, development of treatment guidelines for depression or dementia, and the incorporation of routine screening into primary care practice as recommended for depression by the National Preventive Services Task Force. Attention must be paid to medical education—encouraging physicians to specialize in geriatric medicine and psychiatry, making an effort to promote geriatric education in medical schools and residencies, even changing the structure of medical textbooks that are often organized in chapters covering a single disease entity, rather than teaching management of patients with

multimorbidity. Clinical guidelines should be redesigned to address multimorbidity. All these measures are necessary, but they should be accompanied by promoting effective models of care.

The present overarching organization of primary care is designed for short visits meant to take care of one problem at a time—and geriatric patients with mental health problems such as depression or dementia, especially those with several comorbid medical conditions or frailty, fall so far outside of such a simplistic paradigm that there is no schema in which to tackle their complex presentations. Collaborative models of care have been demonstrated to substantially improve clinical outcomes. In these approaches, mental health specialists collaborate with primary care providers to manage individuals with comorbid medical and mental disorders within the primary care setting. Such interventions have included systematic case finding using simple screening instruments such as the Patient Health Questionnaire-9 (PHQ-9), patient and family education and activation, clinician education and support, systematic assessment and treatment matching, proactive outcomes tracking, and integration of primary care and specialty mental health services. These models may use a geriatric psychiatry care manager with special expertise in the recognition and treatment of common mental disorders or a mental health specialist such as a psychologist or psychiatrist working with a primary care clinic to provide consultation and treatment on site. On-site services can reduce barriers to access and communication between primary and specialty providers, maintain the comfort of patients accustomed to receiving their care in a single setting, and result in better care for both medical and psychiatric problems. In a similar fashion, geriatric psychiatrists have also developed successful models for integrated care provided on site in nursing homes. Psychiatrists and other mental health specialists working in these general medical settings can also help refer patients to further specialty mental health care as needed or link them to helpful community services.

This kind of collaborative care may be especially important for older adults with multimorbidity, management of which requires close attention to both medical and mental disorders. This most common chronic medical condition requires a comprehensive treatment approach that addresses all problems and syndromes present in one elder patient seamlessly, accounting for interactions between treatments and conditions and patient's own preferences. The National Quality Forum, in its publication of *Multiple Chronic Conditions Measurement Framework*, proposed focusing on patient-centered quality measures such as optimizing function, ascertaining patient-important outcomes, and avoiding inappropriate care. Centers for Medicare & Medicaid Services and private health insurance plans should redesign payment from focusing on specific diseases and procedures to value-based quality services promising relief for individuals with multiple medical and mental health problems. In this regard,

accountable care organizations (ACOs) and patient-centered medical homes (PCMHs) hold particular promise. Through their emphasis on high-quality primary care, ACOs and PCMHs may help transform the failing health care system.

The PCMH is a model or health care delivery of a "medical home" where a primary care provider will design and coordinate comprehensive, evidence-based, cost-effective care for the patient. All aspects of care, including end-of-life care and mental health care, are coordinated and provided through the PCMH. An essential component of success of the PCMH is a trusting relationship between the primary care provider and the patient. While this is a new and emerging model of health care, emerging data encourages optimism—we can expect that the PCMH will increase patient satisfaction and decrease health care costs. Following the paradigm of providing quality coordinated care, the last few years have seen emergence of the ACO, defined by the Centers for Medicare and Medicaid Services as "an organization of health care providers that agrees to be accountable for the quality, cost, and overall care of Medicare beneficiaries enrolled in the traditional fee-for-service program who are assigned to it." Similar to the PCMH, the ACO has a strong basis in primary care. The ACO is also driven by intertwining the quality and cost of care, and by magnifying outcome improvements. As new PCMHs and ACOs continue to emerge, and already existing organizations continue to improve, it should be remembered that providing care for a whole patient, geriatric or not, means attending to their psychosocial concerns including mental health, substance abuse, living and family situation, and caregivers if applicable (for patients with physical and mental disabilities, dementia, etc.). Financial incentives that reward addressing behavioral health care on par with medical problems, and reward for improved outcomes rather than fee-for-service approaches, must be implemented.

References

American Geriatrics Society 2012 Beers Criteria Update Expert Panel. American Geriatrics Society updated Beers Criteria for potentially inappropriate medication use in older adults. *J Am Geriatr Soc*. 2012;60(4):616–631.

*American Geriatrics Society Expert Panel on the Care of Older Adults with Multimorbidity. Guiding principles for the care of older adults with multimorbidity: an approach for clinicians. *J Am Geriatr Soc*. 2012;60:E1–E25.

Connolly BS, Lang AE. Pharmacological treatment of Parkinson disease: a review. *JAMA*. 2014;311(16):1670–1683.

DeGruy FV, Etz RS. Attending to the whole person in the patient-centered medical home: the case for incorporating mental healthcare, substance abuse care, and health behavior change. *Fam Syst Health*. 2010;28(4):298–307.

Elderon L, Whooley MA. Depression and cardiovascular disease. *Prog Cardiovasc Dis*.

2013;55(6):511–523.

Fine PG. Chronic pain management in older adults: special considerations. *J Pain Symptom Manage*. 2009;38(2 Suppl):S4–S14.

Freedman GM. Chronic pain. Clinical management of common causes of geriatric pain. *Geriatrics*. 2002;57:36–41.

*Fried LP, Tangen CM, Walston J, et al; Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci*. 2001;56:M146–M156.

Gorevic PD. Osteoarthritis: a review of musculoskeletal aging and treatment issues in geriatric patients. *Geriatrics*. 2004;59:28–32.

Hackett ML, Köhler S, O'Brien JT, Mead GE. Neuropsychiatric outcomes of stroke. *Lancet Neurol*. 2014;13(5):525–534.

*Helfgott AW. The patient-centered medical home and accountable care organizations: an overview. *Curr Opin Obstet Gynecol*. 2012;24:458–464.

*Herr K. Pain assessment strategies in older patients. *J Pain*. 2011;12(3 Suppl 1):S3–S13.

Huang AX, Delucchi K, Dunn LB, Nelson JC. A systematic review and meta-analysis of psychotherapy for late-life depression. *Am J Geriatr Psychiatry*. 2015;23(3):261–273.

Inelmen EM, Mosele M, Sergi G, Toffanello ED, Coin A, Manzato E. Chronic pain in the elderly with advanced dementia. Are we doing our best for their suffering? *Aging Clin Exp Res*. 2012;24(3):207–212.

Kaiser RM. Physiological and clinical considerations of geriatric patient care. In: Blazer DG, Steffens DC, eds. *The American Psychiatric Publishing Textbook of Geriatric Psychiatry*. 4th ed. Arlington, VA: American Psychiatric Publishing, Inc.; 2009.

Lee SY, Franchetti MK, Imanbayev A, Gallo JJ, Spira AP, Lee HB. Non-pharmacological prevention of major depression among community-dwelling older adults: a systematic review of the efficacy of psychotherapy interventions. *Arch Gerontol Geriatr*. 2012;55:522–529.

Li M, Zhang XW, Hou WS, Tang ZY. Impact of depression on incident stroke: a meta-analysis. *Int J Cardiol*. 2015;180:103–110.

Mayeda ER, Whitmer RA, Yaffe K. Diabetes and cognition. *Clin Geriatr Med*. 2015;31(1):101–115.

Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet*. 2009;374:609–619.

Miu DKY, Chan KC. Under-detection of pain in elderly nursing home residents with moderate to severe dementia. *J Clin Gerontol Geriatrics*. 2014;5:23–27.

*Moak GS. Treatment of late-life mental disorders in primary care: we can do a better job. *J Aging Soc Policy*. 2011;23(3):274–285.

Ní Mhaoláin AM, Fan CW, Romero-Ortuno R, et al. Frailty, depression, and anxiety in later life. *Int Psychogeriatr*. 2012;24(8):1265–1274.

NIHCM Foundation. *The Concentration of Health Care Spending*. Washington, DC: NIHCM Foundation Data Brief; 2012.

*Pergolizzi J, Böger RH, Budd K, et al. Opioids and the management of chronic severe pain in the elderly: consensus statement of an International Expert Panel with focus on the six clinically most often used World Health Organization Step III opioids (buprenorphine, fentanyl, hydromorphone, methadone, morphine, oxycodone). *Pain Pract*. 2008;8(4):287–313.

Robertson DA, Savva GM, Kenny RA. Frailty and cognitive impairment—a review of the evidence and causal mechanisms. *Ageing Res Rev*. 2013;12:840–851.

Robinson RG, Jorge RE, Moser DJ, et al. Escitalopram and problem-solving therapy for prevention of poststroke depression: a randomized controlled trial. *JAMA*. 2008;299(20):2391–2400.

Rolinski M, Fox C, Maidment I, McShane R. Cholinesterase inhibitors for dementia with Lewy bodies, Parkinson's disease dementia and cognitive impairment in Parkinson's disease. *Cochrane Database Syst Rev*. 2012;(3):CD006504.

*Schwarz S, Froelich L, Burns A. Pharmacological treatment of dementia. *Curr Opin Psychiatry*. 2012;25:542–550.

Taylor WD, Aizenstein HJ, Alexopoulos GS. The vascular depression hypothesis: mechanisms linking vascular disease with depression. *Mol Psychiatry*. 2013;18(9):963–974.

Tinetti ME, Fried TR, Boyd CM. Designing health care for the most common chronic condition—multimorbidity. *JAMA*. 2012;307(23):2493–2494.

Unützer J, Park M. Older adults with severe, treatment-resistant depression. *JAMA*. 2012;308(9):909–918.

Wei N, Yong W, Li X, et al. Post-stroke depression and lesion location: a systematic review. *J Neurol*. 2015;262(1):81–90.

57.3c Sleep–Wake Disorders in Older Adults

JEANNE E. MAGLIONE, M.D., PH.D., JANA R. COOKE, M.D., AND SONIA ANCOLI-ISRAEL, PH.D.

INTRODUCTION

Up to half of older adults report chronic difficulties with their sleep. A number of changes occur in sleep architecture and sleep patterns with normal aging. However, problems with sleep are not an inevitable part of aging and are associated with significant morbidity. Sleep disturbances in older adults are generally related to specific sleep disorders, comorbid conditions, or both.

SLEEP AND AGING

Older adults report a number of subjective sleep complaints including difficulty falling asleep, waking up often during the night, waking up too early in the morning, and needing to nap during the daytime or evening. Studies using polysomnography (PSG) to objectively measure sleep have confirmed that older people spend more time in bed, have more awakenings at night (decreased sleep efficiency), and take longer to fall asleep compared with their younger counterparts. The proportion of time spent in the lighter stages of sleep (stages 1 and 2) is greater and the amount of slow-wave (deep) and rapid eye movement (REM) sleep are decreased in older adults compared to young adults. Almost all of the above changes in sleep architecture occur between early and middle adulthood. A meta-analysis that included 65 studies and 3,577 subjects ranging from 5 to 102 years in age suggested that sleep efficiency (the percentage of time in bed spent sleeping) alone continues to decrease after the age of 60 years, whereas the other sleep parameters remain relatively stable. Older adults are, on average, sleepier than their younger counterparts during the daytime when tested objectively, suggesting that they are less likely to obtain an adequate amount of restorative sleep at night.

CONSEQUENCES OF POOR SLEEP IN OLDER ADULTS

Insufficient sleep is associated with significant morbidity and increased mortality. Patients with difficulty sleeping report decreased physical health status, mental well-being, and quality of life compared to those without sleep difficulties. In older adults, subjectively and objectively assessed sleep problems are associated with an increased risk of falls, along with ambulation balance, and visual difficulties—even after controlling for medication use. Chronic sleep difficulties at any age can lead to deficits in attention, response times, short-term memory, and performance level. In older adults subjective reports of too little sleep (less than 4.5 hours per night), objectively measured difficulty falling asleep, and fragmented sleep

have all been associated with poor cognition. Accordingly, poor sleep is also associated with increased risk for long-term care placement. Subjective sleep difficulties and objectively measured sleep disturbances are also associated with increased mortality. In a longitudinal study of 185 healthy older adults, lower sleep efficiency (less than 80 percent) increased the risk of all-cause mortality almost twofold. In another large prospective study, women over the age of 69 who reported sleeping 9 to 10 hours in a 24-hour period (including napping) were almost 60 percent more likely to die from cardiovascular or other noncancer causes compared to women who slept between 7 and 8 hours.

SLEEP–WAKE DISORDERS IN OLDER ADULTS

Contributors to sleep disturbances in older adults can include comorbid medical or psychiatric illnesses, medications, circadian rhythm disturbances, or certain specific sleep–wake disorders.

Comorbid Medical and Psychiatric Illnesses

Complaints of sleep difficulties are strongly correlated with complaints about health and depression. In the 2003 National Sleep Foundation survey of adults aged 65 years and older, those with more medical conditions reported significantly more sleep complaints. As the number of medical conditions increased, the likelihood of having sleep difficulties increased. Studies of the prevalence of sleep disturbances in patients with chronic medical diseases have reported that patients with arthritis and chronic pain have difficulty falling asleep and staying asleep. Shortness of breath (due to chronic obstructive pulmonary disease or congestive heart failure), nocturia (due to benign prostatic hypertrophy), and neurological deficits (related to residual stroke deficits or Parkinson disease) are all associated with complaints of insomnia.

Psychiatric disorders including mood disorders and anxiety are also commonly comorbid with insomnia. In a large cross-sectional survey, insomnia was present in 65 percent of respondents with major depression, 61 percent of those with panic disorder, and 44 percent of those with generalized anxiety disorder. Older adults are more likely to experience some major life stressors such as the death of a spouse, medical illness and injury, disability and functional decline, and lack of social contact, all of which can increase their risk for developing depression and anxiety.

Medications

Medications used to treat chronic medical and psychiatric conditions can contribute to or even cause sleep disruptions. It is particularly important to consider the effects of medications in older adults who are often taking multiple prescription and over-the-counter medications. A number of

medications are known to disturb sleep, including β-blockers, bronchodilators, corticosteroids, decongestants, and diuretics, as well as other cardiovascular, neurological, psychiatric, and gastrointestinal medications. Clinicians should advise patients to take sedating medications prior to bedtime and stimulating medications and diuretics earlier in the day to avoid further disruptions in sleep.

Insomnia Disorder

The DSM-5 defines insomnia as a disorder with "a predominant complaint of dissatisfaction with sleep quantity or quality," associated with at least one of the following: difficulty initiating sleep, maintaining sleep, and/or early-morning wakening. The symptoms must cause significant distress or impairment in functioning and must occur at least 3 nights per week for at least 3 months. All possible etiologies (e.g., primary sleep disorders, comorbid medical or psychiatric conditions) must be considered before making this diagnosis.

Treatment of Insomnia. Research has found that behavioral therapy may, in many cases, be more effective than medications for insomnia. In older adults who are at greater risk for serious adverse events related to pharmacotherapy for insomnia and who are often at greater risk for polypharmacy, nonpharmacological approaches should be the first treatment options considered.

BEHAVIORAL TREATMENTS. Cognitive-behavioral therapy (CBT) for insomnia has been shown to have excellent short- and long-term efficacy in both younger and older adults. This approach combines stimulus control and/or sleep restriction with cognitive restructuring, good sleep hygiene (a set of guidelines for healthy sleep–wake habits), and relaxation. A program of CBT typically consists of 6 to 10 weekly sessions administered by a specially trained clinician. Stimulus control is based on the belief that insomnia may be the result of maladaptive classical conditioning. Patients are instructed to eliminate all in-bed activities (i.e., reading, watching television) other than sleep. If they are unable to fall asleep within about 20 minutes, they are told to get out of bed and to do something relaxing outside of the bedroom such as sit quietly reading or listening to music. Once they feel sleepy, they are told to return to bed and attempt to fall asleep. If they are again unable to fall asleep within about 20 minutes, the pattern of getting out of bed until feeling sleepy repeats itself. The goal is to break the connection between the bed and wakefulness. Sleep restriction limits the time spent in bed beyond the amount spent asleep at night. Patients are allowed to be in bed for the amount of time they believe they typically sleep plus 15 minutes. As sleep efficiency increases, the amount of time allowed in bed increases in 15-minute intervals. Cognitive restructuring involves

changing unrealistic beliefs or irrational fears about sleep and/or the loss of sleep. Finally, patients are also educated about ways to modify sleep hygiene practices that may be interfering with their sleep (Table 57.3c–1).

In some settings, the use of CBT for insomnia has been limited by a shortage of clinicians trained to administer the therapy. This obstacle, combined with the burden to patients associated with the required 6 to 10 session therapy course has inspired the development of brief behavioral interventions for insomnia. For example, brief behavioral therapy for insomnia (BBTI) was designed to be administered by nurses within the primary care setting. The treatment is administered over four sessions, three of which last 30 minutes or less and two of which occur by phone. BBTI has an exclusively behavioral focus; the core components being stimulus control and sleep restriction. Several randomized trials have shown BBTI to be effective in older adults.

PHARMACOLOGICAL TREATMENTS. In some cases it may be appropriate to use medications to treat insomnia (Table 57.3c–2). In general, the most effective way to include medications in a treatment plan is in combination with a nonpharmacological intervention so that the medication provides short-term relief while patients undergo CBT or brief behavioral interventions, which are associated with lasting benefit.

Although benzodiazepines are commonly prescribed for insomnia in older adults, there are significant risks associated with this class of medications, particularly in older adults, including an increased risk of falls, cognitive impairment, and driving accidents. In 2012, after reviewing all available data, the American Geriatrics Society updated the Beers Criteria for potentially inappropriate medication use in older adults to include a strong recommendation based on high-quality evidence to avoid benzodiazepine use in older adults. As such, this class of medications should no longer be considered a first-line treatment for insomnia in older adults.

Table 57.3c–1.
Sleep Hygiene Rules

1. Do not spend too much time in bed
2. Maintain a consistent sleep–wake time
3. Get out of bed if unable to fall asleep
4. Restrict naps to 30 minutes in the late morning or early afternoon
5. Exercise regularly
6. Spend time outside each day
7. Eat a light snack (i.e., milk, bread) before bed
8. Avoid caffeine, tobacco, and alcohol after lunch
9. Limit liquids in the evening

Table 57.3c–2.
U.S. Food and Drug Administration–Approved Nonbenzodiazepines for Insomnia

Name	Indication	Time in Bed Required (hours)	[a]Medication Appears in Beers Criteria Recommendations
Zaleplon	Sleep initiation	4	Yes
Zolpidem	Sleep initiation	7–8	Yes
Zolpidem MR	Sleep initiation/sleep maintenance	7–8	Yes
Eszopiclone	Sleep initiation/sleep maintenance	8+	Yes
Ramelteon	Sleep initiation	—	
Low-dose doxepin	Sleep maintenance	—	Yes
Suvorexant	Initiates sleep/maintains sleep	7–8	n/a

[a]See text for specific Beers Criteria cautions regarding potentially inappropriate use of sleep medications in older adults.

The more selective short-acting nonbenzodiazepines (type 1 γ-aminobutyric acid [GABA] benzodiazepine receptor agonists), including eszopiclone, zaleplon, and zolpidem, have been shown to be effective in the short-term management of insomnia in the older adults. Despite an improved side-effect profile, safety concerns remain including increased risk for falls, cognitive symptoms, and complex sleep-related behaviors. The 2012 Beers Criteria strongly recommends avoiding use of these medications in patients who are at risk for falls, delirium, or for patients with neurocognitive disorders and recommends against chronic use in older adults.

A handful of industry-sponsored studies suggest that low doses of the tricyclic antidepressant doxepin (≤6 mg/day) may be effective and safe for promotion of sleep maintenance in older adults. The 2012 Beers Criteria strongly recommends avoiding use of doses >6 mg/day in older adults due to concerns about its significant anti-cholinergic properties, sedating properties, and tendency to induce orthostatic hypotention. A strong recommendation to avoid use of doxepin at any dose in older adults who are at risk for delirium, or who have neurocognitive disorders was also included. Tertiary TCAs can also worsen urinary retention and constipation, common problems in older adults.

Ramelteon, a melatonin receptor agonist is another consideration. A few studies suggest Ramelteon may be safe and effective in older adults, but more research is required in this area. Industry-sponsored phase 2 and 3 clinical trials thus far suggest that the newest FDA-approved sleep medication, suvorexant, an orexin receptor antagonist, may also be safe

and effective in older adults with insomnia though experience with this medication is, at this point, limited.

Although it is still not uncommon to see other classes of medications (e.g., antihistamines, antidepressants, antipsychotics, anticonvulsants) prescribed for insomnia, there is no systematic evidence to support this practice and, in fact, the risks likely outweigh the benefits in older adults.

Circadian Rhythm Sleep–Wake Disorders

In humans, the sleep–wake cycle is controlled by an endogenous clock operating over an approximate 24-hour period called the circadian rhythm. This rhythm is synchronized by external time cues ("zeitgebers") such as the light–dark cycle, which serves as the major external time cue in humans. Circadian rhythm sleep–wake disorders (CRSDs) may develop if dyssynchrony between the endogenous clock and external environmental zeitgebers occurs.

A number of factors associated with aging can contribute to the development of CRSDs. The circadian pacemaker, located in the suprachiasmatic nucleus, is known to degenerate with age, which may result in weaker, more disordered rhythms. The nocturnal secretion of endogenous melatonin, known to play an important role in the sleep–wake cycle, gradually decreases with age, which may result in disrupted circadian rhythms. In addition, the environmental cues needed to synchronize the circadian rhythm of the sleep–wake cycle may be absent or weak in older patients. For example, one study showed that daily bright light exposure averages about 60 minutes for healthy older adults, 30 minutes for patients with Alzheimer disease living at home, and 0 minutes for nursing home patients. Studies of sleep in institutionalized patients have found that this low level of light exposure is associated with sleep fragmentation and CRSDs. Gradual decreases in the amplitude of the circadian rhythm is also associated with aging and may lead to an increase in the frequency of nighttime awakenings and the severity of daytime sleepiness.

The most common CRSD experienced by older adults is advanced sleep phase syndrome (ASPS), which is characterized by habitual (and involuntary) sleep and wake times that are several hours earlier than societal norms. Individuals with ASPS will typically get sleepy in the early evening, timed with the drop in body temperature, and wake up in the early morning (Fig. 57.3c–1). Although sleep itself is normal, delays in sleep onset due to obligations or societal customs may result in insufficient sleep time (as these individuals experience an early-morning awakening regardless of what time they went to bed) and excessive sleepiness during the day. The prevalence of ASPS is about 1 percent in middle-aged and older adults, increasing with age. Men and women are affected equally. A careful and detailed sleep history, sleep diaries, and activity monitoring with wrist actigraphy can all be useful in making the diagnosis.

Treatment of Advanced Sleep Phase Syndrome. For older adults with ASPS, it can be sufficient treatment for them to adopt an earlier bedtime which would then allow for a sufficient amount of total sleep. For many patients, however, this may not be desirable as it may interfere with lifestyle or quality of life. Treatment of CRSDs focuses on strengthening and synchronizing the sleep–wake cycle. Bright light is the strongest cue for circadian entrainment and is therefore the most common and the most effective treatment for circadian rhythm shifts. For older patients with ASPS, including those who reside in nursing homes, evening light exposure has been shown to strengthen the sleep–wake cycle and to delay circadian rhythms. Patients with ASPS should be advised to avoid bright light during the morning and educated on sleep hygiene practices, including adherence to a regular sleep schedule. Early-evening artificial bright light via a bright light box has also been shown to improve sleep continuity in both healthy and institutionalized older patients.

Melatonin supplements given in the evening, could, in theory, result in a phase advance in older adults causing or worsening ASPS, although this has not been studied. Given its popularity as a "natural" treatment for insomnia, clinicians should review the medication list carefully and consider discontinuation of evening melatonin in patients being treated for ASPS. There is currently no evidence base to support prescribing morning melatonin therapy to delay sleep phase in ASPS. Given its sedating properties, clinicians should use care when considering a trial of melatonin therapy in this population.

Obstructive Sleep Apnea Hypopnea

Obstructive sleep apnea-hypopnea syndrome (OSA) describes a spectrum of intermittent respiratory events occurring during sleep, ranging from primary snoring, to increased respiratory effort without an increase in airflow which results in arousal from sleep (flow-limitation events, typically called upper airway resistance), to substantial decreases in airflow (hypopneas), to complete cessation of airflow (apneas), or some combination of these events.

The prevalence of OSA increases or stabilizes with age, as has been shown in most longitudinal and cross-sectional studies. In a large series of randomly selected community-dwelling elderly 65 to 95 years of age, researchers reported that 81 percent of the study subjects had an apnea-hypopnea index (AHI) of 5 or greater, with prevalence rates of 62 percent for an AHI of 10 or greater, 44 percent for an AHI of 20 or greater, and 24 percent for an AHI of 40 or greater. In a longitudinal study, the prevalence only changed with associated changes in body mass index. The Sleep Heart Health Study reported prevalence rates of OSA by 10-year age groups, finding increasing rates with increasing age for subjects with an AHI of 15 or greater.

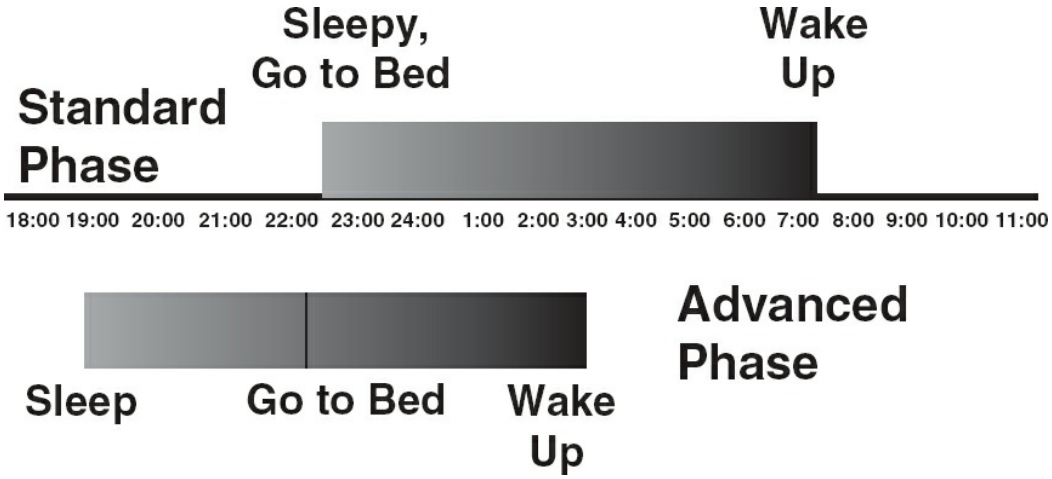

FIGURE 57.3c–1. Standard phase of sleep versus advanced phase of sleep. (From Ancoli-Israel S. *All I Want is a Good Night's Sleep*. St. Louis, MO: Mosby; 1996.)

Institutionalized elderly patients appear to have higher rates of OSA than those elderly patients who live independently. Patients with dementias involving degeneration in brainstem areas, responsible for regulating respiration and other autonomic functions relevant to sleep maintenance, may have a higher risk of developing OSA. Institutionalized elderly patients with more severe dementia have been reported to have more severe OSA than those older patients with milder or no dementia. This study also reported that patients with more severe OSA had lower cognition scores during testing, suggesting that more severe OSA was associated with more severe dementia.

Risk factors for OSAs in the older adult are similar to those found in the younger population, including increasing age, gender, and obesity. Recent data suggest that the pathological mechanism of OSAs may differ from that of younger adults with a more collapse upper airway playing a great role in older adults. Other conditions that increase the risk of developing OSAs include the use of sedating medications, alcohol consumption, family history/genetic factors, menopausal status, race, smoking, and craniofacial morphology.

Snoring and excessive daytime sleepiness (EDS) are the predominant symptoms of OSAs in the elderly. Less common presentations include insomnia; particularly sleep maintenance complaints, nocturnal confusion, nocturia, and daytime cognitive impairment such as difficulty with concentration and short-term memory loss. Unintentional dozing and falling asleep at inappropriate times during the day are common manifestations of the EDS in older patients with untreated OSAs. As with younger adults, excessive sleepiness can lead to both social and occupational difficulties as well as cognitive dysfunction, which is of particular relevance in those patients with underlying cognitive impairment at baseline.

The presence of OSAs is associated with significant morbidity, including cardiovascular disease, depression, and cognitive dysfunction. The presence of untreated OSAs has recently been linked to an increased risk of developing dementia. Older patients with severe OSAs (defined as an AHI of 30 or greater) have consistently been found to have impairments in attention, concentration, memory consolidation, immediate and delayed recall, executive tasks, planning and sequential thinking, visuomotor skills, and manual dexterity. Older adults with milder OSAs (AHI 10 to 20/hr) may experience similar deficits.

OSA is an independent risk factor for the development of systemic hypertension; it is one of the only known causes of hypertension. The exact relationship between OSAs and other cardiovascular disease (including myocardial infarction, pulmonary hypertension, stroke, arrhythmias, and congestive heart failure) in the older adult remains unclear, although results from a few recent studies have provided some insight. The Sleep Heart Health Study found a positive association between the severity of OSAs (based on overnight PSG) and the risk of developing cardiovascular disease, including coronary artery disease, congestive heart failure, and stroke. A recent study in acute ischemic stroke patients reported that OSAs was common, particularly in older male patients with diabetes mellitus whose stroke occurred at night.

As with younger adults, clinicians should use a stepwise approach in evaluating elderly patients with suspected OSAs. A complete sleep history focused on symptoms of OSAs, specifically EDS, unintentional napping, snoring, trouble with maintaining sleep, and symptoms of other sleep disorders (e.g., restless legs syndrome [RLS], insomnia), should be obtained, preferably in the presence of a bed partner or caregiver. The patient's medical history should be reviewed, with particular attention paid to OSAs-associated medical conditions, along with medications, the use of alcohol and/or tobacco products, and evidence of cognitive impairment. If these are suggestive of OSAs, an overnight sleep recording should be obtained for diagnostic purposes.

Treatment of Obstructive Sleep Apnea-Hypopnea Syndrome. Continuous positive airway pressure (CPAP) is the gold standard and most effective treatment of OSAs and has been shown to improve cognition functioning cardiovascular disease outcomes and EDS in older adults. CPAP compliance can be an issue for any patient with OSAs, regardless of age, and clinicians should not assume that elderly patients will be noncompliant simply due to age. Studies support the idea that patients with mild and moderate Alzheimer disease can also be compliant with CPAP treatment, the only factor associated with poor CPAP compliance in these studies was the presence of depression—not age, severity of dementia, or severity of OSAs.

Alternative treatments for OSAs include use of an oral appliance and upper airway surgery. An upper airway stimulator ("pacemaker"), that must be surgically implanted, has recently been approved by the FDA for treatment of OSAs. Clinicians should counsel all patients on weight loss, smoking cessation, and abstinence from the consumption of alcohol if indicated. Sedating medications, especially those known to result in respiratory depression, including benzodiazepines, muscle relaxers, narcotics, should generally be avoided in older adults with OSAs because these agents can increase the number and duration of apneas.

Most experts agree that the treatment of OSAs in the elderly should be guided by the patient's symptoms and the severity of the OSAs. If an older adult has comorbid conditions such as hypertension, depression, cognitive dysfunction, or cardiovascular disease or if the OSAs is severe, treatment should be considered. Age alone should not be a determining factor in treatment decisions.

Periodic Limb Movements in Sleep/Restless Legs Syndrome

Periodic limb movements in sleep (PLMS), a disorder characterized by clusters of repetitive leg movements during sleep, is common in older adults. The leg jerks or kicks occur typically every 20 to 40 seconds and may recur several hundred times over the course of the night. Each jerk may cause an arousal or brief awakening, resulting in sleep fragmentation, and therefore patients with PLMS may complain of daytime sleepiness. The etiology of PLMS is unknown, although it can be seen in patients with fibromyalgia and in combination with other sleep disorders, including sleep disordered breathing (SDB) and narcolepsy. The diagnosis of PLMS requires an overnight polysomnogram showing a calculated periodic limb movement index (PLMI; the number of limb movements per hour of sleep) greater than or equal to 5.

The prevalence of PLMS increases dramatically with age, with reported prevalence of up to 45 percent in community-dwelling adults older than the age of 65 years. It is equally prevalent in older men and women. Despite its high prevalence in older adults, the severity of PLMS does not appear to progressively worsen with increasing age.

RLS, a disorder strongly linked to PLMS, is characterized by leg dysesthesia ("a creepy crawling" or "restless" sensation) that occurs while the person is in a relaxed awake state and can only be relieved by movement. Epidemiological studies reveal the prevalence of RLS in older adults ranges from 9 to 20 percent and, like that of PLMS, increases with age. Studies have also found that RLS occurs twice as often in older women than in older men. RLS can be secondary to anemia or end-stage renal disease, and there is a growing body of evidence linking other chronic illnesses with RLS, including rheumatoid arthritis, chronic obstructive pulmonary disease, asthma, fibromyalgia, diabetes, Parkinson disease, and

multiple sclerosis.

Table 57.3c–3.
Four Essential Questions for Restless Legs Syndrome Diagnosis

1. Do you experience the urge to move your legs (or other parts of your body), typically accompanied or caused by an uncomfortable and/or unpleasant sensation in the affected body part?
2. Does the urge to move or the uncomfortable sensation begin and/or worsen during periods of rest or inactivity?
3. Is the urge to move or uncomfortable sensation partially or totally relieved by movement?
4. Is the urge to move or the uncomfortable sensation worse in the evening or at night compared to the daytime, or does it only occur in the evening or at night?

The diagnosis of RLS is a clinical one, based on history alone—PSG has no role in the diagnosis. Table 57.3c–3 lists the four essential questions that clinicians should ask to make the diagnosis of RLS. Questionnaires such as the Cambridge-Hopkins diagnostic questionnaire for RLS (CH-RLSq), and the RLS Diagnostic Index (RLS-DI) can also be helpful. Exclusion of other conditions that may be mistaken for RLS, such as neuropathy, myalgia, lower extremity edema, leg cramps, positional discomfort, arthritis, pruritus, and vascular insufficiency, is critical. In patients with significant cognitive impairment, signs of discomfort in the legs such as rubbing or kneading the legs and increased motor activity, which are present or worse during inactivity and which are lessened by activity, may be observed.

Treatment of Periodic Limb Movements in Sleep and Restless Legs Syndrome. Medications that suppress the leg movements and/or the related arousals are the therapies currently recommended for both RLS and PLMS. Several dopamine agonists (ropinirole, pramipexole, and rotigotine transdermal) and more recently gabapentin enacarbil have been approved by the U.S. Food and Drug Administration (FDA) for the treatment of RLS. Off-label use of carbidopa/levodopa has also been shown to be effective in some studies.

Rapid Eye Movement Sleep Behavior Disorder

REM sleep behavior disorder (RBD) is more common in older adults, with a predilection to affect men (about 90 percent). The prevalence, which increases with age, is reported to be 0.5 percent in the general population and may be as high as 2 percent in those over 60 years of age and 6 percent in those over 70 years of age. Nearly 80 percent of patients with RBD eventually develop neurodegenerative conditions; specifically synucleinopathies, including Parkinson disease, multisystem atrophy, or neurocognitive disorders with Lewy bodies.

RBD is characterized by elaborate and complex motor activities

associated with dreams ("acting-out dreams") due to the intermittent absence of normal skeletal muscle atonia during REM sleep. Patients may punch, kick, yell, or jump/run out of bed, and these actions usually correlate with dream imagery. Vivid dreams consistent with the patient's aggressive and/or violent behavior may be recalled on waking. Sleep behaviors can place the patient and/or bed partner at risk for physical injuries, which occur in 32 to 64 percent of cases.

The diagnosis of RBD requires a thorough sleep history, in the presence of the bed partner if possible, and an overnight polysomnogram with video recording to confirm the relationship between REM sleep and the complex motor behaviors.

Treatment of Rapid Eye Movement Sleep Behavior Disorder. The most important intervention for all patients with RBD is to make appropriate modifications to ensure environmental safety. Patients should remove heavy or potentially injurious objects from the bed's vicinity and consider locking doors and windows at night. Measures can be taken to avoid being injured by falling out of bed, such as the use of bedrails, a low-profile bed, or mattress on the floor. Historically cautious use of clonazepam has been the mainstay of pharmacological treatment for RBD. However, this practice is based on low-level evidence including exclusively retrospective case series reports, and use of clonazepam can be especially problematic in older adults due to significant associated risks (e.g., cognitive impairment, delirium, falls, fracture, and motor vehicle accidents). Studies also suggest that melatonin may be effective and better tolerated in older adults. High-quality prospective clinical trials comparing these two medications for the treatment of RBD in older patients are needed. Maintaining good sleep hygiene and educating the patient and their bed partner are also important aspects of RBD treatment.

SLEEP IN NEUROCOGNITIVE DISORDERS

Recent studies support a potential connection between sleep disturbances, especially obstructive sleep apnea, fragmented sleep, abnormal duration of sleep, and development of cognitive impairment suggesting that the sleep–wake cycle may play an important role in cognitive aging. Patients with neurocognitive disorders, both those living independently and those institutionalized, experience disrupted sleep which may be due, in part, to neurodegeneration in areas of the brain responsible for sleep regulation. It has been reported that 19 to 44 percent of community-dwelling patients with dementia complain of sleep disturbances, and the prevalence is estimated to be even higher in patients with dementia who are institutionalized. Disrupted sleep commonly results in napping, both intentional and unintentional, and can be manifested as abnormal nocturnal behaviors, including wandering and confusion, which are

significant contributors to caregiver stress and the need for patient institutionalization.

Patients with neurocognitive disorders can have primary sleep disorders such as SDB, PLMS, RLS, or RBD or may have disturbed sleep related to pain, medications, circadian rhythm changes, or depression. Therefore, disturbed sleep related to a neurocognitive disorder should be a diagnosis of exclusion.

As with all older adults, treatment of specific sleep disorders in older patients with neurocognitive disorders should be targeted for that specific sleep disorder. Patients with SBD should be treated with CPAP if appropriate, whereas patients with circadian rhythm disturbances should be treated with bright-light therapy. Given the higher risk-to-benefit ratio of many of the pharmacological treatments in this patient population, nonpharmacological interventions such as regular physical activity and restricting time in bed are preferred. Medications, including sedative–hypnotics, antihistamines, and antipsychotics, all have significant side effects such as residual daytime sleepiness (the so-called "hangover effect") and impairment motor and/or cognitive function. Further, the efficacy and safety of sleep medications have not been adequately studied in patients with neurocognitive disorders and so there is little evidence base to support their use.

SLEEP IN NURSING HOMES

Institutionalized older adults experience more fragmented and disturbed sleep than those living in the community. Nursing home patients have poorer sleep quality, more fragmentation of sleep, more advanced sleep–wake cycles, more daytime sleepiness, and are given sedative–hypnotics more frequently than older adults living in the community or in assisted living facilities. During any 24-hour period of time, institutionalized patients with major neurocognitive disorders may not spend a single hour completely awake or completely asleep.

Excessive nighttime noise and inadequate ambient light exposure both significantly contribute to sleep disruption in nursing home patients. In a study of institutionalized patients, most were found to be exposed to less than 10 minutes of bright light per day and patients with more light exposure had fewer sleep disruptions. Finally, nursing home patients may spend large amounts of time in bed, and inactivity is known to disrupt circadian rhythms.

Patients should have a careful sleep evaluation including assessment for the presence of primary sleep and circadian rhythm disorders. Assessment for contributing medications and for medical and psyciatric illnesses should also be performed. Where appropriate, reduction in nighttime light and noise, increases in daytime light exposure, and improvements in sleep hygiene should be implemented. The authors recommend avoiding the use

of sleep medications wherever possible. If medications are employed they should be used sparingly, cautiously, and only short term.

FUTURE DIRECTIONS

For many older adults, aging is associated with significant changes in sleep. There are a variety of potential contributors including medications, comorbid conditions, circadian rhythm disruption, and specific sleep–wake disorders. As with younger adults, the diagnosis requires a thorough sleep history and often an overnight sleep recording. Furthermore, also as with younger patients, treatment should address the primary sleep problem and can result in significant improvement in both quality of life and daytime functioning.

REFERENCES

Allen RP, Picchietti DL, Garcia-Borreguero D, et al. Restless legs syndrome/Willis-Ekbom disease diagnostic criteria: updated International Restless Legs Syndrome Study Group (IRLSSG) consensus criteria–history, rationale, description, and significance. *Sleep Med*. 2014;15(8):860–873.

American Geriatrics Society 2012 Beers Criteria Update Expert Panel. American Geriatrics Society updated Beers Criteria for potentially inappropriate medication use in older adults. *J Am Geriatr Soc*. 2012;60(4):616–631.

Aurora RN, Zak RS, Maganti RK, et al. Best practice guide for the treatment of REM sleep behavior disorder (RBD). *J Clin Sleep Med*. 2010;6(1):85–95.

Blackwell T, Yaffe K, Ancoli-Israel S, et al; Study of Osteoporotic Fractures Group. Poor sleep is associated with impaired cognitive function in older women: the study of osteoporotic fractures. *J Gerontol A Biol Sci Med Sci*. 2006;61(4):405–410.

Boot BP, Boeve BF, Roberts RO, et al. Probable rapid eye movement sleep behavior disorder increases risk for mild cognitive impairment and Parkinson disease: a population-based study. *Ann Neurol*. 2012;71(1):49–56.

Buysse DJ. Insomnia. *JAMA*. 2013;309(7):706–716.

Buysse DJ, Germain A, Moul DE, et al. Efficacy of brief behavioral treatment for chronic insomnia in older adults. *Arch Intern Med*. 2011;171(10):887–895.

Dam TT, Ewing S, Ancoli-Israel S, Ensrud K, Redline S, Stone K; Osteoporotic Fractures in Men Research Group. Association between sleep and physical function in older men: the osteoporotic fractures in men sleep study. *J Am Geriatr Soc*. 2008;56(9):1665–1673.

Foley DJ, Monjan AA, Brown SL, Simonsick EM, Wallace RB, Blazer DG. Sleep complaints among elderly persons: an epidemiologic study of three communities. *Sleep*. 1995;18(6):425–432.

Hornyak M, Trenkwalder C. Restless legs syndrome and periodic limb movement disorder in the elderly. *J Psychosom Res*. 2004;56(5):543–548.

Kripke DF, Garfinkel L, Wingard DL, Klauber MR, Marler MR. Mortality associated with sleep duration and insomnia. *Arch Gen Psychiatry*. 2002;59(2):131–136.

Liu J, Wang LN. Ramelteon in the treatment of chronic insomnia: systematic review and meta-

analysis. *Int J Clin Pract*. 2012;66(9):867–873.

McCleery J, Cohen DA, Sharpley AL. Pharmacotherapies for sleep disturbances in Alzheimer's disease. *Cochrane Database Syst Rev*. 2014;3:CD009178.

McGrane IR, Leung JG, St. Louis EK, Boeve BF. Melatonin therapy for REM sleep behavior disorder: a critical review of evidence. *Sleep Med*. 2015;16(1):19–26.

Michelson D, Snyder E, Paradis E, et al. Safety and efficacy of suvorexant during 1-year treatment of insomnia with subsequent abrupt treatment discontinuation: a phase 3 randomised, double-blind, placebo-controlled trial. *Lancet Neurol*. 2014;13(5):461–471.

Morin CM, Bootzin RR, Buysse DJ, Edinger JD, Espie CA, Lichstein KL. Psychological and behavioral treatment of insomnia: update of the recent evidence (1998–2004). *Sleep*. 2006;29(11):1398–1414.

Neikrug AB, Ancoli-Israel S. Sleep disorders in the older adult—a mini-review. *Gerontology*. 2010;56(2):181–189.

Neikrug AB, Liu L, Avanzino JA, et al. Continuous positive airway pressure improves sleep and daytime sleepiness in patients with Parkinson disease and sleep apnea. *Sleep*. 2014;37(1):177–185.

Nishihata Y, Takata Y, Usui Y, et al. Continuous positive airway pressure treatment improves cardiovascular outcomes in elderly patients with cardiovascular disease and obstructive sleep apnea. *Heart Vessels*. 2015;30(1):61–69.

Ohayon MM, Carskadon MA, Guilleminault C, Vitiello MV. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. *Sleep*. 2004;27(7):1255–1273.

Peever J, Luppi PH, Montplaisir J. Breakdown in REM sleep circuitry underlies REM sleep behavior disorder. *Trends Neurosci*. 2014;37(5):279–288.

Pollak CP, Perlick D, Linsner JP, Wenston J, Hsieh F. Sleep problems in the community elderly as predictors of death and nursing home placement. *J Community Health*. 1990;15(2):123–135.

Russell T, Duntley S. Sleep disordered breathing in the elderly. *Am J Med*. 2011;124(12):1123–1126.

Stone KL, Ewing SK, Ancoli-Israel S, et al. Self-reported sleep and nap habits and risk of mortality in a large cohort of older women. *J Am Geriatr Soc*. 2009;57(4):604–611.

Stone KL, Ewing SK, Lui LY, et al. Self-reported sleep and nap habits and risk of falls and fractures in older women: the study of osteoporotic fractures. *J Am Geriatr Soc*. 2006;54(8):1177–1183.

Troxel WM, Conrad TS, Germain A, Buysse DJ. Predictors of treatment response to brief behavioral treatment of insomnia (BBTI) in older adults. *J Clin Sleep Med*. 2013;9(12):1281–1289.

Wennberg AM, Canham SL, Smith MT, Spira AP. Optimizing sleep in older adults: treating insomnia. *Maturitas*. 2013;76(3):247–252.

Yaffe K, Falvey CM, Hoang T. Connections between sleep and cognition in older adults. *Lancet Neurol*. 2014;13(10):1017–1028.

Yaffe K, Laffan AM, Harrison SL, et al. Sleep-disordered breathing, hypoxia, and risk of mild cognitive impairment and dementia in older women. *JAMA*. 2011;306(6):613–619.

Zee PC, Attarian H, Videnovic A. Circadian rhythm abnormalities. *Continuum (Minneap Minn)*. 2013;19(1 Sleep Disorders):132–147.

57.3d Anxiety Disorders

CAROLYN I. RODRIGUEZ, M.D., PH.D., AND MURRAY B. STEIN, M.D.

INTRODUCTION

In DSM-5, the chapter on anxiety disorders has been revised to include separation anxiety disorder and selective mutism. No longer included are obsessive-compulsive disorder (OCD) (which is included with the obsessive-compulsive and related disorders) or posttraumatic stress disorder (PTSD) and acute stress disorder (which are included with the trauma- and stressor-related disorders). The term "anxiety disorders" used in this chapter nonetheless is used to refer to the all the DSM-5 diagnoses within anxiety-, trauma- and stressor-related, and obsessive-compulsive and related disorders. When we refer to these disorders in aggregate (i.e., when there are no specific distinctions between them we are highlighting), we will use the term "anxiety disorders" otherwise we will specify the distinctions. The majority of research described in this chapter employed a DSM-IV framework. If applicable, we included recent research findings using the revised DSM-5 nosology.

Recently, there has been an increasing evidence base that anxiety in late life has greater clinical consequences and is more prevalent than previously thought. Although much less is still known about late-life anxiety than about other common late-life psychiatric conditions such as depression or dementia, current studies show that in community samples, anxiety is more common than depression in older adults. In addition, anxiety disorders in later life can greatly impact quality of life, increase disability levels, and impact recovery from illness. When compared to the body of research of anxiety in younger adults or children, limited data suggest that older adults present with the same anxiety disorders as do younger adults, although several additional anxiety syndromes may be particularly relevant to the elderly. Virtually all epidemiological research is consistent in finding that the prevalence of anxiety, trauma- and stress-related, and obsessive-compulsive and related disorders is lower among older adults than among younger adults, particularly with respect to panic disorder and social anxiety disorder (social phobia). Rates vary considerably by study, however, leading to contradictory conclusions: either anxiety, trauma- and stress-related, and obsessive-compulsive and related disorders in the elderly are extremely rare or they constitute the most common class of geriatric psychiatric conditions, with substantially higher prevalence rates than mood disorders.

Diagnostic issues are complicated by medical illness and medications,

which can cause or contribute to anxiety-like symptoms. Clinicians may also attribute symptoms to realistic, age-appropriate health problems, worries, or fears.

Although these factors may lead investigators to believe that anxiety disorders are uncommon in older adults, it is also possible that anxiety is underdiagnosed in the presence of medical conditions or depression. Some clinicians may conclude that anxiety or avoidance is an inevitable or adaptive response to aging and therefore not pathological; in many cases, this conclusion may be inaccurate. Differences in the presentation of certain anxiety disorders in older adults may also muddy the diagnostic waters. Finally, there is insufficient research on differences in presentation or etiology between late-onset anxiety disorders and disorders originating earlier in life that have persisted until or recurred in old age. Additional research on these topics is necessary to clarify the many unresolved issues surrounding anxiety in the elderly.

DEFINITION, HISTORY, AND COMPARATIVE NOSOLOGY

The definition, history, and comparative nosology of anxiety, trauma- and stressor-related, and obsessive-compulsive and related disorder diagnoses, as defined by the various versions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) and the *International Statistical Classification of Diseases and Related Health Problems* (ICD). To date, no geriatric-specific anxiety disorders have been included in the diagnostic nomenclature.

EPIDEMIOLOGY

Prevalence

The prevalence of anxiety disorders in the elderly is controversial. Investigations have found rates ranging from 0.7 to 19 percent. A recent review of epidemiological studies in Europe found a range of 4 to 14 percent. Because only one of those studies surveyed PTSD, it is likely that the true prevalence of anxiety disorders may be underestimated. By comparison, another recent review found prevalence rates of major depression ranging from 0.9 to 9 percent among community-dwelling older adults. Thus, it appears likely that, as in younger adults, anxiety disorders as a class are more common than major depression in the elderly.

The prevalence of anxiety disorders most likely decreases among older adults, and the gender difference, with women reporting higher rates than men, appears to attenuate among the oldest old. There appears to be a lower comorbidity of anxiety disorders and alcohol abuse among older adults than among younger adults, although this may be influenced by low rates of substance use disorders among the elderly. Most studies conclude that generalized anxiety disorder (GAD) and phobias are the most common

anxiety disorders among older adults, with panic disorder and social anxiety disorder (social phobia) much less common than in younger age groups. Estimated prevalence rates for specific anxiety disorders are listed in Table 57.3d–1.

Table 57.3d–1.
Prevalence Estimates of Anxiety Disorders in Older Adults

Generalized anxiety disorder	0.7–9%
Specific phobia	3–5%
Social anxiety disorder (social phobia)	0.2–2%
Panic disorder	0.1–1%
Agoraphobia	0.4%
Obsessive-compulsive disorder	0.1–2%
Hoarding disorder	14%
Posttraumatic stress disorder	8.7%

Incidence

Incidence studies of late-life anxiety are relatively rare, so most data on the age of onset come from retrospective reporting. The incidence of agoraphobia and possibly OCD in women may increase in late life, whereas specific phobia, OCD in men, and panic disorder are typically continuations from earlier life or the result of another psychiatric or medical illness. Approximately 30 to 40 percent of GAD cases arise for the first time in later life.

Prevalence within Special Groups

Anxiety disorders are common in diseases of the elderly. Rates as high as 30 percent have been reported in patients with cardiac conditions such as coronary heart disease and atrial fibrillation. In chronic obstructive pulmonary disease patients, prevalence rates of panic disorder range from 8 to 32 percent, with GAD rates ranging from 10 to 16 percent. GAD may occur in 15 to 20 percent of patients with neurological conditions such as stroke and Alzheimer disease.

Although rates of anxiety disorders in special populations are typically higher, this is not the case for nursing home patients, among whom the rate is approximately 5 percent. A review of prevalence data on anxiety disorders among caregivers for individuals with dementia found rates ranging from 4 to 16 percent. In a review and meta-analysis of studies involving widows, prevalence rates were 12 percent for PTSD, 10 percent for panic disorder, and 31 percent for GAD. By comparison, the prevalence of major depression was 22 percent.

Some data suggest that anxiety may be more common in older individuals from minority groups, although a recent study of older African

Americans found 6.6 percent reporting an anxiety disorder (vs. 2.4 percent for major depression), consistent with other research. The distribution of disorders, however, was different from what has been reported in other samples. For example, the most common anxiety disorders in this sample were PTSD (2.9 percent) and social anxiety disorder (social phobia) (2.1 percent). The oldest old and those in the South were less likely to report anxiety disorders. Almost one-fourth (24 percent) of older Puerto Rican primary care patients met criteria for an anxiety disorder, primarily specific phobia (12 percent), GAD (11 percent), or panic attacks (11 percent). Across all ethnic groups, phobias appear to be associated with lower income.

Although older veterans report rates of PTSD that are substantially lower than the rates found in younger veterans (e.g., 6 vs. 19 percent in one study), rates of PTSD are high among aging former prisoners of war (POWs) (e.g., 20 percent). The prevalence of PTSD appears to be highest among Holocaust survivors (e.g., 45 percent).

Comorbidity with Mood Disorders

As is the case with younger adults, comorbidity of anxiety and mood disorders is very common. Data from the Epidemiologic Catchment Area survey and two recent Dutch studies found that 21 to 37 percent of community-dwelling patients with an anxiety disorder (as defined by DSM-IV) experienced comorbid depression. Among nursing home patients, 42 percent of those with a DSM-IV anxiety disorder experienced comorbid major depression. Although these rates suggest that the disorders frequently overlap, it appears that most DSM-IV anxiety disorders in older adults are unaccompanied by a mood disorder. Studies using DSM-5 criteria are not available.

ETIOLOGY AND RISK FACTORS

Although much less research has been conducted on anxiety disorders than on depression in older adults, it appears that risk factors are similar for both conditions. Female gender, cognitive impairment, chronic health conditions, poor self-rated health, and functional limitations are among the most important factors increasing vulnerability. Psychological risk factors for both anxiety disorders and depression include personality traits (i.e., neuroticism), dysfunctional coping, and psychopathology. Marital status (unmarried) appears to be associated with depression but not anxiety, and being childless, having lower income, and experiencing traumatic events are more typically associated with anxiety than with depression. Older age appears to be protective for anxiety but a risk factor for depression. Of interest, one study that examined depression and anxiety disorders simultaneously found that although risk factors were similar for both classes of conditions, odds ratios associated with virtually all risks were

higher for anxiety disorders than for depression.

In nursing home patients, anxiety disorders appear to be related to younger age, lower levels of social support, better cognitive performance, depression, and history of stroke. Comorbid anxiety and depressive disorders are associated with the presence of pain, loneliness, and perceived lack of adequate care.

With respect to specific anxiety disorders, female gender, marital status (divorced or widowed), and comorbid psychiatric illness (specifically major depression and social anxiety disorder [social phobia]) are risk factors for agoraphobia in older adults. Late-onset agoraphobia is often associated with medical illness or traumatic events such as victimization or falls. Panic disorder is more strongly associated with disability than with health problems. Phobic disorders are associated with lower income and early parental loss. Risk factors for PTSD include female gender, neuroticism, low self-efficacy, and childhood trauma, and there are some reports of new-onset or worsening PTSD symptoms with cognitive decline.

DIAGNOSIS AND CLINICAL FEATURES

There is no evidence that anxiety disorders in older adults present substantially differently in older than in younger adults. Similar to what has been reported in younger samples, frequency and difficulty controlling worry, sleep disturbance, fatigue, and muscle tension appear to be the cardinal features of late-life GAD.

Older adults with panic disorder report fewer and less severe panic symptoms, lower levels of arousal, and less depression than younger individuals. Late-onset panic disorder appears to be associated with lower levels of cognitive and somatic distress during panic attacks. OCD may differ somewhat between older and younger patients, with older persons reporting more contamination and religious obsessions but fewer rituals involving symmetry or counting.

Older adults appear to develop PTSD and other anxiety syndromes after exposure to traumatic events less frequently than do younger adults. Older PTSD patients report more social dysfunction than younger patients, and in older adults, external locus of control appears to be associated with more avoidance, social dysfunction, and depression. Older PTSD patients may present more somatic than psychological symptoms and may report less severe PTSD symptomatology.

Anxiety Syndromes of Particular Relevance to the Elderly

Agoraphobia, fear of falling, hoarding, and complicated grief merit special mention because they may be anxiety syndromes with particular relevance to the elderly.

Agoraphobia. As shown in Table 57.3d–1, prevalence rates of

agoraphobia vary widely in older adults, suggesting a lack of diagnostic clarity. Distinguishing between adaptive and excessive avoidance is challenging, and patients typically believe that their behavior is reasonable, given age-associated medical conditions and physical limitations. Because panic disorder is rare in the elderly, Agoraphobia Without History of Panic Disorder (DSM-IV) is the most common symptom presentation. In DSM-5, where Panic Disorder and Agoraphobia have been diagnostically uncoupled, the diagnosis of Agoraphobia would apply.

In one study of older adults that used conservative criteria that defined agoraphobia as avoidance due to a panic attack rather than panic-like symptoms, the prevalence of comorbid panic disorder was still much lower (8 percent) than the prevalence of major depression (25 percent) or social phobia (20 percent). This suggests that older adults may limit their activity (i.e., develop agoraphobia) more often than do younger adults in response to panic attacks that do not meet criteria for panic disorder. Only 11 percent of late-onset patients with agoraphobia had histories of panic attacks in another investigation with looser criteria for agoraphobia; instead, they developed their fear in response to medical illness or traumatic events.

Across the lifespan, clinical features of agoraphobia are fairly consistent; however, age may impact the type of situation feared. The most common situations feared in older adults include being in line, being in shops, and being in open spaces. Older individuals may fear or avoid these situations due to thoughts that help may not be available in the event of embarrassing symptoms (e.g., fear of falling, fear of incontinence).

Fear of Falling. Although under DSM-5 diagnostic criteria, fear of falling can only be classified as a Specific Phobia, Other Type if the patient perceives it to be excessive or unreasonable, it is a common and disabling anxiety syndrome in older adults. As many as 25 percent of community-dwelling older adults report fear of falling. Although individuals with a past fall are more likely to develop fear of falling than those without such history, most individuals reporting fear do not have a history of falls.

Fear of falling has potentially serious consequences, including, paradoxically, increased risk of falls. Approximately half of those with fear of falling restrict their activities due to fear, and the risk of falling is highest among those who do, even among those who have never experienced a fall. Fear of falling may lead to maladaptive changes in gait parameters, further increasing the risk of future falls, and may interfere with rehabilitation following a fall. For example, in one investigation of patients recovering from hip fracture surgery, fear of falling predicted poor functional recovery, although pain and depression did not.

As is true with other anxiety syndromes, fear of falling is frequently, although not always, comorbid with other anxiety disorders and depression. In one sample of individuals with high fear of falling and a

recent fall, 36 percent had depression and 19 percent had an anxiety disorder, most commonly agoraphobia without panic, and anxiety was a strong predictor of fear of falling even after controlling for factors such as dizziness, visual impairment, number of medications, and depressive disorders and symptoms.

Hoarding. Once considered a subtype of OCD, hoarding disorder is now considered a separable syndrome and is a new diagnostic category in DSM-5. Nationally representative epidemiological studies of hoarding disorder with DSM-5 criteria are not available, but recent smaller epidemiologic surveys estimate a 1.5 to 6 percent prevalence of hoarding behaviors (not necessarily at the level of a hoarding disorder diagnosis) in the general population. Hoarding symptoms may be nearly three times more prevalent in older adults compared with younger adults. Older adults may also show more severe hoarding symptoms than do middle-aged adults.

Elderly hoarders are typically female, never married, and living alone. Hoarding is typically a chronic condition with an early onset and progressive course. For those with a late onset of symptoms, hoarding likely begins secondary to a disease process occurring in later life, such as dementia. Pest infestations, fire hazards, and falling are risk factors for older adults with hoarding behaviors.

Persistent Complex Bereavement Disorder. Older adults experience bereavement at much higher rates than do younger adults, and as many as 20 percent of bereaved individuals suffer prolonged distress and disability. However, approximately 10 to 12 percent of individuals will continue to experience persistent symptoms that do not resolve naturally and interfere with functioning. Persistent Complex Bereavement Disorder is a proposed disorder included in the section of "Conditions for Further Study" in DSM-5. Symptoms may include persistent yearning, sorrow, and preoccupation with the deceased or the circumstances of the death, reactive distress to the death or social/identify disruption. Risk factors for persistent complex bereavement disorder include dependency on the deceased person prior to death, death of a child, and being female. Further research is needed to establish the validity and utility of this diagnostic entity.

PATHOLOGY AND LABORATORY EXAMINATION

Most research on the biology of late-life anxiety has focused on PTSD, and much of it has been performed by a single research team in samples of Holocaust survivors and older veterans. Studies in PTSD and other late-life anxiety disorders have examined cognitive performance, hypothalamic–pituitary–adrenal (HPA) axis function and other markers of autonomic arousal, neuroimaging, and genetics.

Posttraumatic Stress Disorder

With respect to neuropsychological functioning, older persons with PTSD display poorer explicit memory, visuospatial functioning, attention, and working memory than those with trauma exposure but no PTSD and those without trauma exposure. Among those with PTSD, improvement in recall may be associated with improvement in PTSD symptoms.

Other research shows evidence of HPA axis dysregulation associated with late-life PTSD. Individuals with PTSD show flatter circadian salivary cortisol patterns and lower urinary cortisol excretion than do those without PTSD, and better explicit memory performance is associated with lower salivary cortisol levels. Urinary cortisol excretion and glucocorticoid receptor–mediated responsiveness appear to be associated with age of first trauma exposure, and cortisol level may be associated with severity of childhood trauma. Moreover, cortisol administration may enhance working memory performance in older veterans with PTSD. Glucocorticoid receptor, plasma adrenocorticotropic hormone (ACTH), and cortisol suppression in response to dexamethasone administration are greater in older PTSD patients. Older PTSD patients may have higher levels of dehydroepiandrosterone (DHEA) and DHEA sulfate ester (DHEAS) and a lower cortisol/DHEA ratio. The cortisol/DHEA ratio appears to be predicted by the severity of childhood trauma and current PTSD symptoms, and both variables are predicted by symptom improvement and coping.

In terms of brain structure, although one investigation found no differences in hippocampal volumes in Holocaust survivors relative to controls, among those with PTSD, smaller left hippocampal volumes were found in those who had no prior history of trauma before the index trauma. These findings suggest that smaller hippocampal volume may be a risk factor for the development of PTSD.

Other Anxiety Disorders

Elderly female social anxiety disorder (social phobia) patients in one sample showed greater hemodynamic responses, cardiac vagal attenuation, and sympathetic vasomotor activation than those without social anxiety disorder (social phobia) in response to a public speaking challenge. Older adults with specific phobias and GAD show greater and longer-lasting cortisol response to stress than nonanxious older adults. Older GAD patients may have memory and executive function deficits comparable to those of patients with major depression relative to normal older adults, and older adults with depression and comorbid anxiety disorders show greater memory declines over time than those with depression alone. In GAD, working memory deficits have been associated with higher levels of anxiety symptoms, and errors in temporal orientation are associated with poorer response to cognitive-behavioral therapy (CBT).

There is limited information to date on how—and indeed, if—separation anxiety disorder and selective mutism manifest in older adults. Both of these are disorders that are typically diagnosed in childhood, and the extent to which they persist or have onset in later adulthood is unknown.

Genetic Factors

A few studies have explored genetic risk factors for late-life GAD. As in younger adults, genetic influences appear to be shared with depression and partially, although not entirely, with neuroticism. A large, population-based study of middle-aged and older adults (mean age 60 years) found no associations between a brain-derived neurotrophic factor coding variant (Val66Met, rs6265) and current or lifetime GAD.

DIFFERENTIAL DIAGNOSIS

The most common differential diagnoses are medical conditions and depression. In primary care or specialty medical settings, somatic symptoms of anxiety may be difficult to distinguish from those of cardiovascular, endocrine, gastrointestinal, metabolic, neurological, or pulmonary disorders or from the effects of medications, including antihypertensives, β-agonists, corticosteroids, digitalis, nonsteroidal anti-inflammatory drugs (NSAIDs), and thyroid medication. A workup to identify medical conditions that may be contributing to anxiety in older adults should comprise a complete medical history and physical examination and laboratory tests including complete blood count, thyroid function, and levels of electrolytes, vitamin B_{12}, and folate. It should be noted, however, that anxiety can coexist with medical conditions and may lead to excess disability and, potentially, exacerbation of the medical condition itself. Treatment of anxiety as a co-occurring disorder may be appropriate, particularly when the medical condition is stable.

Because few older adults with anxiety disorders that are not accompanied by depression present to mental health settings, it is common for comorbid anxiety disorders to be overlooked in depressed older patients. Several symptoms of major depression and GAD overlap, and rates of comorbid anxiety disorders are as high as 40 to 50 percent in samples of depressed older adults presenting to mental health clinics.

Specific training about anxiety disorders for geriatric mental health providers can boost recognition rates, especially in depressed older adults. Table 57.3d–2 summarizes differences in the typical presentation of depression and anxiety in older adults. Because anxiety typically develops before depression and follows a chronic rather than episodic course, questions about symptoms such as worry, feeling tense or nervous, or finding it difficult to relax at times when the individual was not depressed may clarify whether a depressive episode is superimposed on a pre-existing

anxiety disorder.

COURSE AND PROGNOSIS

The course of anxiety disorders in older adults is typically chronic. Prognosis is often poor; recent investigations have found higher mortality risk associated with generalized anxiety and PTSD symptoms and panic attacks, even after controlling for medical factors and depression.

With respect to specific anxiety disorders, GAD may be characterized by periods of remission and relapse occurring over decades but with longer symptomatic periods in older than in younger adults. One influential view holds that late-life GAD and depression exist on a continuum, with GAD representing the mildest form of the disorder and comorbid GAD and depression representing the most severe syndrome. Providing more support for the hypothesis that GAD and depression exist on a continuum, one epidemiological study found that only 12 percent of individuals with GAD at baseline continued to have GAD without major depression at 3-year follow-up. Almost half (48 percent) had remitted and 40 percent went on to develop major depression, either with or without comorbid GAD. Ten-year follow-up data from the same study indicated that excess mortality was associated with depression in men but not in women; GAD either with or without comorbid depression was not associated with mortality. A 40-year follow-up of individuals retrospectively diagnosed with GAD found evidence that the disorder may convert to a somatoform disorder over the lifespan.

A 50-year follow-up study of OCD patients found remission rates of approximately 60 percent. By contrast, a 50-year follow-up of panic disorder patients found only 6 percent of patients not meeting criteria for any psychiatric disorder at the final assessment point. In this study, prevalence of panic disorder with and without agoraphobia began to decline after approximately age 40 years, whereas incidence of somatoform disorders and agoraphobia without panic rose steadily from age 20 through 60 years and then leveled off. Although PTSD symptoms are most prominent immediately after the trauma, they can persist for decades, and some evidence suggests that they may emerge for the first time or increase in later life, especially in the context of cognitive decline.

Consequences

Consequences of anxiety disorders include activity limitations, higher numbers of general practitioner and home health visits, loneliness, poor life satisfaction, poor perceived health, and perceived lack of help. Older GAD patients report poorer role functioning, social functioning, and energy levels than individuals with other chronic conditions such as heart disease and diabetes. Older adults with PTSD, even in the absence of depression,

experience more days spent in bed and days with limited activity and lower levels of life satisfaction than those without PTSD after controlling for factors including age, chronic diseases, functional limitations, social network size, and cognitive functioning. They also make more visits to mental health providers and have higher rates of benzodiazepine use.

Table 57.3d–2.
Characteristics of Depression and Anxiety in Older Adults

Typical Presentation	Depression	Anxiety
Somatic symptoms	Vegetative (e.g., poor appetite, fatigue)	Autonomic (e.g., palpitations, dizziness, gastrointestinal distress)
Activity level	Typically psychomotor retardation	Agitation or subjective feeling of restlessness
Enjoyment of life	Diminished	Not diminished, but patient may claim that he or she has no time to enjoy life
Sleep disturbance	Early-morning awakening	Difficulty falling or staying asleep
Thought content	Rumination about past losses or failures, suicidal ideation	Worry about current or future threats or dangers
Behavior	Decreased activity	May exhibit excessive activity (e.g., cleaning, taking on too many projects or responsibilities)
Family relationships	May be isolated or disconnected from family	May be overinvolved or intrusive with family
Social networks	Diminished	Often not diminished, may hide symptoms from network
Onset	May have early or late onset	More typically early onset
Course	Episodic	Chronic
Attribution	Patient may believe that symptoms are part of normal aging or a natural response to life events	Patient may believe that symptoms are part of his or her personality or a natural response to life events

In nursing home patients, anxiety disorders and depression are associated with lower levels of well-being, although not necessarily with higher levels of disability. Patients with depression, either alone or combined with anxiety, report lower levels of well-being than patients with pure anxiety. Individuals with anxiety and/or depression receive less assistance with activities of daily living (ADLs), make more visits to medical specialists, and take more medications, especially antidepressants. Use of antidepressants is typically more common in patients with combined anxiety and depression relative to others. Patients with anxiety make more visits to medical specialists than those without anxiety or depression.

Anxiety as a Risk Factor for Other Medical Conditions

Anxiety may also serve as a risk factor for common medical conditions of later life. Studies have reported increased risk of vascular problems, respiratory disease, chronic pain, gastrointestinal conditions, and cancer associated with PTSD.

Substantial evidence links anxiety to heart disease. Symptoms of phobic anxiety may predict the onset of ventricular arrhythmias in patients with coronary artery disease. Generalized anxiety disorder, panic attacks, and PTSD symptoms lead to elevated risk of coronary heart disease, even in the absence of major depression. PTSD symptoms are also associated prospectively with nonfatal myocardial infarction. Reported lifetime prevalence of anxiety disorders is considerably higher than current prevalence in coronary heart disease, suggesting that anxiety might represent a risk factor rather than a consequence of heart disease.

TREATMENT

Anxiety disorders in the elderly are frequently undetected and untreated. Older adults are significantly less likely than younger adults to receive treatment for anxiety. When treatment takes place, it usually involves benzodiazepines or other anxiolytic medications rather than antidepressants or psychotherapy.

Pharmacotherapy

Surprisingly little research has been conducted on treatment of anxiety in older adults. Recent evidence suggests that selective serotonin reuptake inhibitor (SSRI) pharmacotherapy is superior to psychotherapy as an acute treatment for geriatric anxiety. Data from controlled trials support the use of escitalopram (Lexapro), duloxetine (Cymbalta), pregabalin (Lyrica), venlafaxine XR (Effexor), citalopram (Celexa), and sertraline (Zoloft) for late-life GAD.

Data are even more sparse with respect to treatment of other anxiety disorders, but evidence from controlled trials and open-label studies suggests the efficacy of sertraline, citalopram, and escitalopram, as well as of imipramine (Tofranil) and alprazolam (Xanax), for panic disorder in older adults. The SSRIs citalopram, escitalopram, and sertraline may be particularly effective treatments for anxiety in the elderly based on their low propensity for drug interactions and short half-lives. Fluvoxamine (Luvox) has demonstrated efficacy in an open-label trial including older patients with GAD, panic disorder, or OCD.

For PTSD, only one small open-label pharmacotherapy study has been conducted in adults with a mean age greater than 60 years, and studies with mean ages of 50 years or older have been veteran samples comprised almost exclusively of men. Randomized, controlled trials in this age group support the use of sertraline and mirtazapine, and an open-label trial

supports the efficacy of citalopram. Prazosin may improve sleep quality and PTSD-related nightmares; its impact on other PTSD symptoms is less clear.

Benzodiazepines are often the only mental health treatment received by older adults with anxiety disorders, and they are prescribed more frequently to older adults than to younger adults. Although benzodiazepines are effective in reducing anxiety and may be useful as an adjunctive, as-needed, or short-term treatment strategy, long-term daily use is associated with falls, fractures, and cognitive decline in the elderly and is not recommended. Among benzodiazepines, lorazepam has a relatively short half-life and no active metabolites and is therefore preferable to alprazolam or clonazepam, which are more likely to accumulate and achieve toxic levels in older adults. Benzodiazepine taper should be very gradual in the elderly, even after a short duration of use.

Comorbid Anxiety and Depressive Disorders. Older adults with depression and comorbid anxiety disorders or significant anxiety symptoms typically demonstrate delayed acute treatment response and increased risk of relapse of depression. They may also report more suicidal ideation. For these reasons, it is especially important to treat older adults with depression and coexisting anxiety to complete remission. Open-label trials support the potential utility of escitalopram for comorbid depression and GAD in older adults.

Augmentation and Maintenance Treatment. No randomized trials have tested augmentation strategies with anxious older adults who fail to respond to a first-line treatment. A group of older adults with anxiety disorders (mostly GAD) who did not respond to an initial trial of an antidepressant or benzodiazepine responded to adjunctive risperidone (average dose 1.2 mg per day).

Although data suggest the continued efficacy of citalopram for GAD after 24 weeks of treatment, with an additional 20 percent of patients responding in the second 12 weeks, no research has examined maintenance therapies for late-life anxiety. Given the chronic course of anxiety disorders, studies on this topic should be a high priority.

Clinical Management. Because SSRIs may initially exacerbate symptoms of anxiety, older adults should be started on a low dose (Table 57.3d–3). In clinical practice, however, older adults with anxiety or depression are frequently maintained on minimum or even subtherapeutic doses, with corresponding limited benefit. If the patient is tolerating the medication, the dose should be increased at least to the minimum target dose within 4 weeks. If the patient has not made a full response (e.g., achieved at least a 50 percent decrease in symptoms) within another 4 to 6 weeks, the dose should be gradually increased, with continued increases up

to the maximum until response is achieved or the patient is no longer able to tolerate the medication.

Older adults with anxiety are frequently more resistant to the idea of pharmacotherapy and more sensitive to perceived side effects than younger adults with anxiety or older adults with depression, attributing to the drug symptoms that preceded the initiation of a medication trial. Therefore, the specific agent may be less important than good clinical management, including information about the safety and nonaddictive potential of the medication, reassurance about the transient and noncatastrophic nature of side effects, availability of the physician at any time in the event of bothersome symptoms, and frequent visits initially (ideally weekly for the first month and biweekly thereafter until response is achieved). Patients may need continued reassurance at every dose change and, in the event of emergent side effects, should be encouraged to stay the course for at least 1 or 2 weeks. Clinicians may need to think creatively to facilitate this process, for example, by suggesting that patients split the dose between morning and evening administration or take over-the-counter medications temporarily to counteract side effects (e.g., Imodium for diarrhea).

Table 57.3d–3.
Usual Starting, Minimum Target, and Maximum Dose of Commonly Used Antidepressants for Older Adults with Anxiety Disorders

Medication	Starting Dose (mg/day)	Minimum Target Dose (mg/day)	Maximum Dose[a] (mg/day)
Citalopram	5	20	20[b]
Escitalopram	10	10	20
Sertraline	12.5	50	200
Fluvoxamine	25	100	300
Venlafaxine extended-release	37.5	75	225
Mirtazapine	7.5	15	45

[a]Higher doses are sometimes used on a case-by-case basis; lower doses are often used in the oldest old.

[b]The U.S. Food and Drug Administration revised recommendations for citalopram are as follows: the maximum recommended dose of citalopram is 20 mg per day for patients older than 60 years of age.

Psychotherapy

Some evidence suggests that SSRI medications may be more effective than psychotherapy as an acute treatment for late-life anxiety disorders. Research, mostly with older GAD patients, provides support for the effectiveness of CBT relative to waiting list and active control conditions such as supportive therapy and attention placebo. One study has demonstrated the superiority of CBT to supportive psychotherapy in a well-characterized sample of older adults with anxiety disorders (52 percent

with panic disorder), and all participants in this investigation were taking medications but had failed to respond prior to receiving psychotherapy. Response to CBT for GAD in older adults may be compromised by memory impairment, executive dysfunction, or failure to complete at-home practice. A recent study has demonstrated that the sequence of antidepressant medication augmented with CBT in older adults with GAD leads to worry reduction in the short term. This study also found that both maintenance medication and CBT could prevent relapse compared with placebo, providing two strategies for improving outcomes in GAD in older adults.

For older individuals who refuse to take medications or for whom they are contraindicated, relaxation training may be the most effective component of a behavioral treatment for GAD or distressing anxiety symptoms that do not meet criteria for a specific anxiety disorder. Recommending a good relaxation tape with breathing or meditation exercises, progressive muscle relaxation, or imagery may be helpful; if after a few weeks of practice the patient does not report sufficient gains, it may be worthwhile to reopen a discussion about medications.

For anxiety disorders other than GAD, some evidence supports the use of CBT for panic disorder in older adults, although results from a study that included a substantial proportion of older panic disorder patients suggest the superiority of sertraline over CBT. There are no published data beyond case reports documenting the behavioral treatment of specific phobia or OCD, all of which use exposure-based protocols, and no data on the treatment of older adults with agoraphobia without panic disorder or social anxiety disorder (social phobia). Case reports of PTSD treatment in older adults have described gains associated with psychodynamic and supportive approaches, CBT, eye movement desensitization and reprocessing, life review, and exposure therapy. A recent pilot study of prolonged exposure therapy in older veterans with PTSD showed significant decrease in their symptoms of PTSD.

Recommendations on CBT with older adults and on behavioral treatment of anxiety are included in Sections 57.4i and 14.7 of this volume, respectively. Recommendations specific to late-life anxiety include (1) ensuring that sufficient time is devoted to learning skills such as relaxation training; (2) practicing skills multiple times in session to ensure that the rationale and procedures are well understood prior to assigning at-home practice; (3) ensuring that anxiety levels become sufficiently high during exposure exercises, which may be facilitated by, for example, having patients hyperventilate briefly before engaging in imaginal exposure; (4) not avoiding difficult or painful issues or worries if appropriate (e.g., imaginal exposure to issues surrounding an ill spouse's death); and (5) deemphasizing more abstract cognitive restructuring techniques in favor of simpler strategies (e.g., coping statements).

Mrs. H was a 69-year-old married white woman with a college degree. She reported that she was "always a worrier" but believed that she was acting in the best interests of her family by looking out for their well-being. She indicated that her worrying had caused problems within her family; for example, her daughter was annoyed with Mrs. H for sending her classified ads and articles about employment resources to encourage her to find a new job. She further stated that her concern about her children led her to telephone them daily, and she had once contacted the police when her call was not returned by the next day.

In addition to worrying about her children, she indicated that she was concerned about her husband's health, their finances, and her health, particularly her recent diagnosis of gastroesophageal reflux disease. Although she believed that her worry was normal under the circumstances and she was resistant to the idea of taking medication, she indicated that sleep disturbance, fatigue, and difficulty concentrating were making it difficult for her to accomplish daily tasks.

Mrs. H was started on 50 mg per day of sertraline but called the psychiatrist's office after 3 days to report that she had discontinued the drug after experiencing an upset stomach and a burning sensation in her throat. She agreed with the doctor's recommendation for her to restart the medication at a reduced dose of 25 mg per day and to take chewable calcium tablets as often as needed. He called her back 2 days later to inquire about the success of this strategy. Although Mrs. H reported that the symptoms continued, she agreed to continue the medication for another week, until her next scheduled appointment. At that point, she was unwilling to consider raising the dose but agreed to stay the course for another 2 weeks, at which point she reported that the side effects were "unpleasant but manageable" and reluctantly agreed to try 50 mg per day. Over the next 3 months, her dose was gradually increased to 200 mg per day, with diminishing side effects. Mrs. H reported that although she continued to worry about her family, the thoughts were less distressing, and she was sleeping better and had more energy.

Mrs. P was an 82-year-old African-American woman whose daughter brought her to the clinic out of concern that she was depressed. Although Mrs. P lived independently and had been sleeping and eating well, she had refused to leave her home unaccompanied since experiencing a noninjurious fall 3 months before the appointment. She denied feeling sad and reported that she enjoyed watching television, listening to music on the radio, and having family and friends from her church visit her in her home. She stated that although she used to take walks for exercise and do her grocery shopping at the corner market, she now experienced considerable anxiety about not having help available should she fall outside of her home. Her family was trying to decide whether she should come to live with one of her children. Because of her strong preference to continue living on her own, she agreed to come to the clinic for an evaluation.

After consulting with her physician, who reported that she did not have orthostasis or other medical conditions increasing her risk of falling, the psychiatrist decided first to have Mrs. P evaluated by a physical therapist, who tested her gait and taught her to use a cane. Although Mrs. P reported that she was using the cane to feel more secure within her home, she continued to avoid unaccompanied trips outside of her home. She was then referred to a psychologist for behavioral treatment.

Mrs. P and the psychologist collaboratively developed a hierarchy of situations involving exposure to trips outside of the home that varied in duration, distance, and availability of family members for support. Her family was enlisted to help with the treatment by initially being available by telephone while she was engaged in exposure. As treatment progressed, she was "weaned" from making calls to family members immediately before, during, and after her walks. At this point, which took six sessions over an 8-week period, Mrs. P reported that she was able to leave her home alone, taking her cell phone but not using it routinely during every excursion. Treatment was deemed successful and terminated.

FUTURE DIRECTIONS

Anxiety disorders are more common than usually recognized in older adults. In the mental health clinic setting, they are typically seen as a co-

occurring diagnosis with depression, although this may change as public awareness about anxiety and its treatment grows. Coexisting anxiety and depression typically represents a more severe disorder than depression alone. Because onset is often early in life, older adults with anxiety disorders have usually experienced a chronic course. Distress and impairment associated with these disorders can be substantial. Data on biological factors involved in late-life anxiety implicate HPA axis dysregulation, and genetic, cognitive, and neuroanatomical influences are under investigation. Although data are still sparse, SSRI medications appear to represent the first-line acute treatment for geriatric anxiety. Very few maintenance studies have been performed, and more research is needed. Psychotherapy has shown some efficacy in older adults for GAD and panic disorder with promising pilot studies in PTSD and has not been adequately investigated for phobias, including agoraphobia. A focus on anxiety disorders across the life span is needed; one practical step toward this is including older participants in clinical trials within multiple settings, including community and residential facilities. Alternatives to CBT and combinations of pharmaco- and psychotherapy have not been investigated.

References

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

Ayers CR, Saxena S, Espejo E, Twamley EW, Granholm E, Wetherell JL. Novel treatment for geriatric hoarding disorder: an open trial of cognitive rehabilitation paired with behavior therapy. *Am J Geriatr Psychiatry*. 2014;22(3):248–252.

Bower ES, Wetherell JL, Mon T, Lenze EJ. Treating anxiety disorders in older adults: current treatments and future directions. *Harv Rev Psychiatry*. 2015;23(5):329–42.

Bryant C, Mohlman J, Gum A, et al. Anxiety disorders in older adults: looking to DSM5 and beyond . . . . *Am J Geriatr Psychiatry*. 2013;2(9):872–876.

Lenze EJ, Rollman BL, Shear MK, et al. Escitalopram for older adults with generalized anxiety disorder: a randomized controlled trial. *JAMA*. 2009;301:295–303.

Lenze EJ, Wetherell JL. Bringing the bedside to the bench, and then to the community: a prospectus for intervention research in late-life anxiety disorders. *Int J Geriatr Psychiatry*. 2009;24:1–14.

Reid JM, Arnold E, Rosen S, et al. Hoarding behaviors among nonclinical elderly adults: correlations with hoarding cognitions, obsessive–compulsive symptoms, and measures of general psychopathology. *J Anxiety Disord*. 2011;25(8):1116–1122.

Samuels J, Grados MA, Riddle MA, et al. Hoarding in children and adolescents with obsessive–compulsive disorder. *J Obsessive Compuls Relat Disord*. 2014;3(4):325–331.

Thorp S, Stein MB, Jeste D, Patterson T, Wetherell J. Prolonged exposure therapy for older veterans with posttraumatic stress disorder: a pilot study. *Am J Geriatr Psychiatry*. 2012;20(3):276–280.

Timpano KR, Exner C, Glaesmer H, et al. The epidemiology of the proposed DSM-5 hoarding

disorder: exploration of the acquisition specifier, associated features, and distress. *J Clin Psychiatry*. 2011;72(6):780–786.

Wetherell JL, Petkus AJ, White KS, et al. Antidepressant medication augmented with cognitive-behavioral therapy for generalized anxiety disorder in older adults. *Am J Psychiatry*. 2013;170(7):782–789.

57.3e Geriatric Mood Disorders

JENNIFER ROSE GATCHEL, M.D., PH.D., JAMES M. WILKINS, M.D., BRENT P. FORESTER, M.D., ROBERT EMMETT KELLY, JR., M.D., AND GEORGE S. ALEXOPOULOS, M.D.

Geriatric mood disorders cause suffering, increase medical burden, worsen the outcomes of many medical conditions, and contribute to disability. Investigators at the World Health Organization conclude that depression produces a greater decrement in health than angina, arthritis, asthma, and diabetes. This finding holds across countries and demographic characteristics. The complexity of geriatric mood disorders and comorbid medical illnesses requires coordination of care among health professionals and continuity of care, often difficult to achieve within the current health care delivery system. Moreover, the behavioral disability of persistent geriatric mood disorders often requires expensive support systems.

Despite the availability of relatively efficacious treatments, geriatric depression continues to be underdiagnosed. The failure to recognize geriatric depression stems from the clinical complexity of the syndrome, symptom overlap with medical comorbidities, social bias, and barriers of the settings in which most depressed elderly patients are treated. Depression does not have pathognomonic symptoms or signs. Many of the symptoms of depression (e.g., loss of interest, lack of energy, physical complaints) can be caused by medical illnesses. Moreover, depressed patients often have several chronic medical illnesses, which make it difficult to determine whether a symptom is caused by depression, a medical illness, or both. Depressed elderly patients often experience anhedonia (depression without sadness) rather than depressed mood, a symptom difficult to ascertain during a brief examination. Ageism among patients and health professionals may lead to diagnostic complacency based on the view that older individuals who lose their physical prowess and social status have valid reasons to be depressed. Finally, two-thirds of depressed elderly patients are treated in busy primary care offices. Such settings are geared to acute care of medical problems. Their fast pace and the emphasis on medical rather than mental health problems are significant barriers to effective diagnosis. Although late-life depression is underdiagnosed, many nondepressed older patients receive antidepressants. Incorrectly attributing physical symptoms to depression may lead to premature discontinuation of medical workup and failure to identify a medical cause for physical complaints. Moreover, inappropriate use of antidepressants may increase the cost of care and cause drug interactions.

The past decade has witnessed an acceleration of the understanding of the neurobiology of late-life mood disorders. Future studies designed using the NIMH Research Domain Criteria (RDoC) guidelines will hopefully

identify critical neurobiological targets paving the way for novel pharmacotherapy and psychotherapy treatment approaches of late-life mood disorders.

EPIDEMIOLOGY

The National Comorbidity Survey Replication discovered that the overall prevalence of major depression among persons aged 65 years or older was estimated to be 1.5 percent in women and 0.2 percent in men, with an overall prevalence of 1 percent, approximately one-fourth of that in younger adults. However, when combining symptoms of major or minor depression with reported treatment for depression, the cumulative depression prevalence in later life is noted to be as high as 11.19 percent with similar rates for men (10.19 percent) and women (11.44 percent). Further, in this population study of late-life depression, Caucasians and Hispanics had nearly three times the prevalence of depression found in African Americans.

Subthreshold depression has a median community point prevalence of 9.8 percent, approximately two to three times more prevalent than major depression. Subthreshold depression is lower in community settings versus primary care settings and highest in long-term care settings. Approximately 8 to 10 percent of older persons with subthreshold depression developed major depression per year. Subsyndromic depression is more frequent than syndromic depression and similarly does not increase with age.

The prevalence of major depression has been increasing in recent birth cohorts compared to cohorts born in the early part of the 20th century. Depression appears to have an increasingly earlier onset in cohorts born in recent years, and there is a narrowing in the differential risk for depression between genders.

Depression frequently occurs in medical populations. For patients treated in primary care settings, depression was identified in 17 to 37 percent of patients. Approximately 30 percent of these patients had major depression, and the remainder had a variety of depressive syndromes that could benefit from medical attention. Major depression occurs in roughly 25 percent of older home care patients, 11 percent of medically hospitalized patients, and up to 48 percent for individuals in long-term care settings.

As the world population ages, prevalence rates for older adults with bipolar disorder (BD) are rising, now estimated to account for 25 percent of all bipolar patients. While the estimated lifetime prevalence of BD in the U.S. population for individuals aged 60 years and older is 1 to 2 percent, rates in specific clinical populations vary from 3 percent of nursing home residents, up to 7 percent of psychiatric outpatients, 10 percent of psychiatric inpatients, and 17 percent of individuals presenting for psychiatric treatment in ER settings. Further, approximately 70 to 95 percent represent cases with onset prior to age 50 years that has persisted

into later life. Approximately 69 percent of late-life bipolar patients are women.

In summary, mood disorders occur in the older geriatric population at a rate lower than that of younger adults. However, significantly higher rates of depression occur in medically ill, disabled, or institutionalized elderly persons. With the aging of birth cohorts who appear to be particularly prone to depression, an unprecedented number of older adults are expected to require psychiatric attention.

SUBTYPES OF LATE-LIFE DEPRESSION

Depression in Medically Ill Patients

Geriatric depression typically occurs in the context of medical disorders. Sometimes these conditions predispose or trigger depression. For example, subclinical hypothyroidism increases the risk for depression more than fourfold while nearly a third of patients develop depression within a year after a myocardial infarction. Overall medical severity and disability appear to be important risk factors for depression. For these reasons, geriatric depression is viewed as a heterogeneous entity.

Older adults with chronic disease have a 50 percent greater risk for depression than older, medically healthy adults. Depression increases the risk of coronary disease by over 60 percent while also increasing the risk of cancer and a shortened survival time, regardless of cancer etiology. Individuals with malnutrition, metabolic syndrome, and post hip fracture surgery are all at elevated risk of depression. Data from a large health maintenance organization population suggest that the health care cost of depressed medical patients is twice that of nondepressed patients with similar levels of medical morbidity, though only a small proportion of the increased costs are spent on specialty mental health care.

Depression impairs social function more than heart and lung disease, arthritis, hypertension, and diabetes and leads to greater impairment in physical function than arthritis, hypertension, and diabetes. This effect seems to be mediated through decreasing adherence to treatment, as well as through biological mechanisms such as changes in sympathetic–parasympathetic balance, changes in serotonin-mediated thrombogenesis, and changes in the immune system.

Depression is underdiagnosed and undertreated in primary care settings. Approximately 11 percent of depressed high utilizers of primary care services receive adequate antidepressant treatment, whereas 34 percent receive inadequate treatment and 55 percent receive no treatment. The pattern of inadequate treatment of depression also extends to long-term care settings where 23 percent of nursing home residents diagnosed with depression receive no treatment.

Depression in Patients with Dementia

In DSM-V, cognitive disorders are described on a spectrum: mild and major neurocognitive disorders (MaNCD) and their etiological subtypes. Despite a recent change in nomenclature, the term dementia continues to be used frequently in clinical settings, and is retained here for consistency. Depressive manifestations of varying intensity occur in approximately 50 percent of patients with dementia. Examination of the interval between the occurrence of depression and development of dementia suggests that depression is both a risk factor for AD and a prodrome. A meta-analysis concluded that a history of depression doubles the risk for developing dementia in late life. Depression occurs in approximately 25 percent of patients with cerebrovascular disease and in up to 50 percent of those with Parkinson disease. The similarity of depressive manifestations with symptoms and signs of dementing disorders, for example, loss of interest, decreased energy, agitation or psychomotor retardation, poor concentration, and apathy often poses diagnostic problems. Sad mood and psychic rather than vegetative features have been found useful in distinguishing depressed patients with dementia from patients with dementia alone.

Depression with Reversible Dementia

Some elderly depressed patients develop a dementia syndrome that improves or completely subsides after remission of depression. This syndrome has been termed *pseudodementia*, *dementia syndrome of depression*, or *depression with reversible dementia*, and it has attracted considerable clinical interest. Abnormal executive function, processing speed, and working memory may persist after remission of mood symptoms in many patients with geriatric depression.

Depressed elderly patients who remain with some cognitive impairment even after improvement of depression usually have an early-stage dementia whose cognitive manifestations are exaggerated when the depressive syndrome is superimposed. Although some disagreement exists, it appears that even those individuals with nearly complete cognitive recovery develop high rates of irreversible dementia (about 20 percent per year) on follow-up. Many of these patients may remain nondemented for 1 to 2 years before the development of irreversible dementia. The heterogeneity of the course of depression with reversible dementia suggests that patients suffering from this syndrome can be ordered along a continuum. At the one end of the continuum lie cases in which the cognitive disturbance results predominantly from the depressive syndrome itself. On the other end are cases in which the intellectual impairment originates from a progressive subclinical neurodegenerative disorder with some contribution by the depressive syndrome. Therefore, identification of a reversible dementia

syndrome in elderly depressive patients constitutes an indication for a thorough diagnostic workup and frequent follow-up aimed at the identification of treatable neurological disorders.

Psychotic Depression

Psychotic depression occurs in 20 to 45 percent of hospitalized elderly depressives and in 3.6 percent of elderly depressives living in the community. Patients with psychotic depression have delusions, whereas hallucinations are less frequent. The usual themes of depressive delusions are guilt, hypochondriasis, nihilism, persecution, and sometimes jealousy. Depressive delusions can be distinguished from the delusions of patients with dementia, in that the latter are less systematized and less congruent to the affective disturbance.

Late-Life Bipolar Disorder

Although community prevalence rates of BD are lower in older adult populations than in younger adult ones, recurrent episodes of mania and hypomania persist into later life and affect 5 to 10 percent of psychiatric patients. Late-onset mania often occurs in patients who have had recurrent depression episodes since early life. Unipolar mania, though relatively uncommon, is an early-onset disease. Similar to unipolar depressive disorder, late-life BD occurs in the context of cognitive impairment, high medical burden, disability, and limited psychosocial support. These factors influence the course of late-life BD. Importantly, decrements in physical health-related quality of life in this population are not an inevitable consequence of aging, and adequate attention to medical comorbidity at any age may have a tangible impact.

Cognitive impairment occurs in 40 to 60 percent of symptomatic elderly bipolar patients and may endure after improvement. Deficits in attention, executive functions, and memory have been noted. Elderly bipolar manic patients may have more severe executive function deficits than unipolar depressed elders. These impairments may reflect, in part, processes involved in the pathophysiology and pathogenesis of mood disorders, including dysfunction of corticostriatal circuits and other limbic pathways. Cognitive impairment may be a prodrome of dementia and may also be associated with poor outcomes in late-life mania.

Older adults with BD have higher stroke risk factors than both aged controls and unipolar elders. Late-onset mania is especially associated with vascular burden and high mortality due to cerebrovascular disease. Neuroimaging studies have revealed a high frequency of signal hyperintensities and silent strokes in older bipolar patients. Structural abnormalities are most prominent in BD with late age at onset of first episode. Cardiac pathology, renal impairment, and thyroid dysfunction are

also common in aged bipolar patients.

BD leads to behavioral disability that worsens the symptom course in BD. Disability in the elderly can lead to institutionalization and influence morbidity and mortality. Geriatric bipolar patients report low perceived social support compared to same-age controls. Bipolar patients with a history of frequent or lengthy episodes of illness lack social support. The effect of social support on treatment outcomes of BD in old age requires further investigation.

Depression and Disability

According to the World Health Organization, unipolar depression alone is responsible for more than 1 in 10 years lived with disability worldwide. Depression is among the 10 diseases that increase disability-adjusted life years (DALYs) the most. DALY is a measure of years of life lost to premature death and years lived with a disability. In a 12-year follow-up study, individuals with major depression at baseline were four times more likely to report incident disability after controlling for medical illness and sociodemographic factors. Whereas depression increases the risk of disability, the opposite is also true. Increases in disability over time predict the emergence of depressive symptoms. Impairment in instrumental activities of daily living is related to depressive symptomatology and particularly to anxiety, helplessness, hopelessness, and worthlessness. Disability is associated with executive dysfunction and with late-onset depression.

Bereavement

Depressive symptoms or syndromes often follow a significant loss. In DSM-IV, bereavement within 2 months of the death of someone close was an exclusion criterion for the diagnosis of depression. Older persons appear to be at lower risk for developing depressive symptoms or syndromes than younger adults during the first months after loss of a spouse. However, by the end of the second year, older and younger individuals have the same rates of major depression. In fact, the prevalence of major depression continues to increase during the second year of bereavement. At the end of the second year after loss, 14 percent of bereaved elderly individuals have major depression.

SUICIDE RISK

Older adults are at significant risk of death by suicide relative to younger adults. Caucasian men older than 80 years are the group with the highest suicide risk. Severity of depression, psychotic depression, alcoholism, recent bereavement, abuse of sedatives, and recent development of disability are the most important risk factors in depressed elderly.

Accumulation of chronic medical conditions and isolation that comes with divorce or widowhood as well as cognitive difficulties may also contribute to the increased risk in this population. Aging reduces suicide attempts but increases their lethality. Suicide victims are demographically different from individuals who attempt suicide. About 60 percent of those who commit suicide are men while 75 percent of those who attempt suicide are women. The majority of suicide victims use guns or hang themselves while 70 percent of suicide attempters take a drug overdose, and 20 percent cut or slash themselves. In late life, suicide attempts are often more carefully planned and are preceded by fewer warnings of suicidal intent.

Indeed, unlike the high frequency of suicide attempts in late life, reports of suicidal ideation decrease in older adults. Despite its lower frequency, suicidal ideation is a risk factor for suicide in elderly patients. Severity of depression is the strongest predictor of the course of suicidal ideation. In elderly patients with major depression, suicidal ideation is most common in those with previous serious suicide attempts, high severity of depression, and poor social support.

Depression is the most common psychiatric diagnosis in elderly suicide victims, unlike younger adults in whom substance abuse alone, or with comorbid mood disorders and psychosis are the most frequent diagnoses. Major depression is identified in 80 percent of suicide victims older than 74 years of age. Many elderly suicide victims are widowed and fewer are single, separated, or divorced, compared with younger adults. Depressive symptoms, medical status, perceived health status, cognitive difficulties, and perceived burdensomeness and hopelessness were identified as risk factors for suicidal ideation and death in older adults. Insomnia, one of the most common symptoms of depression at any age and particularly in late life, is an independent risk factor for suicidal ideation in adults and suicidal death in the elderly. It is unclear whether treatment of insomnia can mitigate this risk.

Suicide is also a leading cause of death in patients with bipolar illness. By one estimate, 25 to 50 percent of all patients with BD will attempt suicide at least once in their life, and 8 to 19 percent will die of suicide. Risk factors for suicide in BD include younger age of illness onset, history of past suicide behavior, family history of suicide acts, comorbid borderline personality disorder, substance use disorders, and hopelessness.

The clinical characteristics of depressed elderly suicide victims suggest that identification of depression and appropriate care can reduce the incidence of suicide. The majority of elderly suicide victims see their physicians within a few months of their death and over a third, within a week of their suicide. Among depressed elderly seen in primary care in the 12 months prior to a suicide attempt, less than 5% received appropriate depression treatment prior to the attempt. Therefore, reliable assessment of suicide risk is critical, since protective measures may avert suicide. There

is no evidence that questions about suicide can increase the likelihood of suicidal behavior. Aggressive treatment of depression may prevent suicide, as suicidal ideation is closely associated with depressive symptomatology in older adults. Worsening depression and suicidal ideation may indicate the need for psychiatric hospitalization and/or ECT.

BIOLOGICAL DYSFUNCTION

The Late-Onset Hypothesis

An early hypothesis of late-life depression postulated that major depression with the onset of first episode in late life includes a large subgroup of patients in whom neurological brain abnormalities contribute to the development of the depressive syndrome. Clinical findings supported the "late-onset hypothesis." Relative to older adults with early-onset major depression, individuals with late-onset major depression have less frequent family history of mood disorders, higher prevalence of dementing disorders, and higher rate of dementia development on follow-up, although, some disagreement exists. Most elderly patients with major depression and an initially reversible dementia have late-onset depression. Despite improvement of cognitive functions with remission of the depressive syndrome, many of these patients develop irreversible dementia during long-term follow-up. The "late-onset depression" hypothesis pointed out the importance of aging and disease-related brain changes in late-life depression.

The Depression-Executive Dysfunction Hypothesis

This hypothesis postulates that a "depression-executive dysfunction (DED)" syndrome exists with a clinical presentation resembling medial-frontal lobe syndrome. Clinical studies documented that the DED syndrome is common and characterized by psychomotor retardation, reduced interest in activities, apathy, impaired insight, and severe behavioral disability, but less pronounced depressive ideation and vegetative signs. Patients with depression and executive dysfunction have poor, slow and unstable response to antidepressants. Similarly, depressed older adults who reported executive functioning complaints at baseline had a slower response to escitalopram treatment than those without executive functioning complaints.

The DED hypothesis assumes that executive dysfunction is the clinical expression of frontolimbic dysfunctions predisposing to late-life depression. The limbic system evaluates reward-related stimuli and organizes stress responses, and higher neocortical structures exert control over these processes. A recent study provided evidence for insula metabolism as a treatment-specific biomarker in major depression. Specifically, insula hypometabolism was associated with remission in

response to cognitive-behavioral therapy and poor response to escitalopram, while insula hypermetabolism was associated with remission to escitalopram and poor response to cognitive-behavioral therapy in a mixed-age sample. Reduced bilateral activation of the dorsal-anterior cingulate and the hippocampus was demonstrated in severely depressed, nondemented elderly patients performing a word-activation task. Low resting functional connectivity within the cognitive control network and high-resting functional connectivity within the default mode network distinguished depressed from normal elderly subjects. Low resting functional connectivity within the cognitive control network predicted low remission rate and persistence of depressive symptoms and signs, apathy, and dysexecutive behavior after treatment with escitalopram. In contrast, resting functional connectivity within the default mode network was correlated with pessimism but did not predict treatment response. Patients with the DED syndrome often present with apathy. Resting functional connectivity studies of apathy in late–late-life depression showed that apathy is characterized by abnormalities in the cognitive control network as well as in reward and salience networks.

White matter abnormalities compromising connections between limbic and dorsal cortical structures may interfere with limbic–dorsal balance and lead to chronic depressive syndromes. Subcortical white matter hyperintensities (WMHs) have been associated with both executive dysfunction and nonremission of late-life depression. Complementing these findings, diffusion tensor imaging studies documented that reduced white matter integrity in distributed cerebral networks (dorsal and rostral anterior cingulate cortex [ACC], dorsolateral-prefrontal cortex [DLPFC], hippocampus, posterior cingulate, insula, neostriatum, and the midbrain as well as select temporal and parietal regions) is associated with poor response of late-life major depression to a serotonin reuptake inhibitor. Depressed, older, short serotonin transporter 5-HTTLPR allele carriers had both a lower remission rate than long allele homozygotes, and more microstructural white matter abnormalities in frontolimbic areas. These abnormalities may be caused by inflammatory responses, vascular disease, prominent aging-related changes, degenerative brain disease, or an accumulation of those and other factors.

The Inflammation Hypothesis

The inflammation hypothesis postulates that aging related immune processes promote changes in the emotional and cognitive neural networks predisposing to geriatric depression and/or triggering metabolic brain changes mediating the depressive syndrome in late life. The hypothesis is supported by emerging findings suggesting that: (1) aging leads to a proinflammatory state; (2) immune responses of the CNS can influence the function of some of the emotional and cognitive networks pertinent to

geriatric depression; (3) elevation of peripheral cytokines is associated with depressive symptoms in older adults; (4) some antidepressants reduce the expression of several inflammation markers in the periphery; and (5) pretreatment cytokine plasma levels may influence the response to antidepressants.

Aging promotes a chronic state of neuroinflammation, and animal and human studies suggest that this proinflammatory state may compromise neural systems related to depression. In aged rodents, the cerebral cortex produces more inflammatory cytokines than the cortex of adult animals. In older animals, the inflammatory cytokine IL-6 has higher concentrations in the hippocampus, cerebral cortex, and cerebellum than in the same regions of younger animals. In normal human adults, enhanced activity of the subgenual anterior cingulate cortex and reduced functional connectivity (FC) of the subgenual anterior cingulate with the amygdala, medial-prefrontal cortex and nucleus accumbens during emotional face processing was modulated by a peripheral inflammatory response (IL-6). An SNP encoding IL-1b has been associated with both reduced activity of the anterior cingulate and the amygdala in response to emotional probes and with poor response of major depression to antidepressant. Patients treated with the cytokine interferon-α have greater activation of the dorsal-anterior cingulate than controls. In brain autopsies of suicide victims, increased cytokine concentrations have been found in the orbitofrontal cortex and the DLPFC.

Epidemiologic studies have shown an association between elevation of peripheral cytokines and depressive symptoms in older adults. Increased IL-6 was the most consistent finding. Elevation of IL-1Ra was a risk factor for developing depressive symptoms during a 6-year follow-up in older adults. Clinical studies of depressed patients, documented a plasma elevation of several inflammatory markers (TNF-α, IL-6 and IL-1b) and a correlation with the severity of depressed mood and cognitive symptoms of depression.

Some tricyclic antidepressants (TCAs) may reduce some proinflammatory cytokines, increase anti-inflammatory cytokines, and reduce polysaccharide-induced “sickness behavior” in animals. Pretreatment cytokine plasma levels may influence the response to antidepressants in humans. Selective serotonin reuptake inhibitor (SSRI)-resistant patients had higher TNF-α plasma levels while euthymic patients with a history of depression had TNF-α levels similar to those of normal controls. A higher level of TNF-α was shown to predict failure of depression to respond to escitalopram. The largest study to date reported that patients with baseline low C-reactive protein (CRP <1 mg/L) levels had greater improvement of depression when treated with escitalopram than with nortriptyline. In contrast, patients with higher CRP levels had greater improvement when treated with nortriptyline than with escitalopram. In

treatment-resistant major depression, baseline concentrations of high--sensitivity CRP (>5 mg/L), TNF and TNF-soluble receptors were associated with favorable response to the TNF antagonist infliximab.

The Vascular Depression Hypothesis

This hypothesis postulates that cerebrovascular disease predisposes, precipitates, or perpetuates some late-life depressive syndromes. The hypothesis was based on the high comorbidity of depression and vascular risk factors; the high incidence of depression in stroke; the high prevalence of WMHs in late-onset depression; the high frequency of cognitive impairment in depressed patients with vascular risk factors; and the similarity of cognitive abnormalities of depression to those associated with WMH. Recent studies show greater WMH severity in specific white matter fiber tracts in late-life depression, including the cingulum bundle, uncinate fasciculus, and superior longitudinal fasciculus. Fiber tract structural connectivity is positively correlated with resting state functional connectivity of connected regions.

A syndrome of "subcortical ischemic depression" has been proposed based on WMH literature. This syndrome is defined by the combination of depression and MRI evidence of subcortical ischemic changes. Age, lassitude, and a history of hypertension were positively associated with the diagnosis of subcortical ischemic vascular depression, whereas a family history of mental illness and loss of libido were negatively associated with this diagnosis.

WMH and microstructural brain abnormalities predict poor response of late-life depression to antidepressants. Increases in WMH volume over 2- and 4-year intervals are associated with nonremission or relapse. Microstructural white matter abnormalities in multiple frontolimbic brain areas, including the rostral- and dorsal-anterior cingulate, DLPFC, genu, hippocampus, posterior cingulate, left superior corona radiata, and right inferior longitudinal fasciculum are associated with nonremission.

Disconnection of structures regulating mood and cognition may be one of the mechanisms by which cerebrovascular disease promotes late-life depression. Another mechanism may be vascular dysregulation leading to reduced cerebral blood flow (CBF). Mild CBF reduction may impair cognitive and affective processes, while greater CBF reduction in the context of autoregulatory deficits may cause ischemic injury.

Vascular diseases impair vascular regulation. Vascular changes, including increased intima-media thickness, increased arterial stiffness, and endothelial dysfunction are pronounced in late-life depression and endothelial function may be most impaired in nonresponders to antidepressants. In late-life depression, perfusion deficits are more severe, particularly in the medial- and lateral-prefrontal cortices, subcortical, and temporal structures.

Arguably, at least three interconnected processes may lead to "vascular depression," that is, disconnection of networks responsible for regulation of mood and cognition, reduced regional CBF, and inflammatory responses related to vascular disease. As an example, inflammation leads to vascular remodeling, accelerates vascular damage, and may contribute to perfusion deficits. In turn, regional perfusion deficits adversely affect gray matter function and contribute to the development of WMHs that disrupt regional structural and functional connectivity.

Genetic Studies

Polygenic inheritance has been hypothesized in depression. Multiple genes are associated with common psychiatric disorders, which immeasurably complicates the analysis of associations between psychiatric illness and genetic makeup. Despite convincing evidence for genetic factors contributing and predisposing to depression, there have been few robust molecular genetic findings. A genome-wide associated study (GWAS) of more than 9,000 cases of major depressive disorder (MDD) did not reveal loci that exceeded genome-wide significance. This is consistent with a model of a heterogeneous disorder and/or one in which the majority of genetic variance is due to joint effects of multiple loci of small effects.

Candidate gene studies focusing on biological pathways implicated in MDD have provided conflicting results. Meta-analysis of 26 candidate genes yielded 7 genes with significant findings: *5-HTTP/SLC6A4, APOE, DRD4, GNB3, HTR1A, MTHFR, SLC6A3*, though these genes have not been detected in GWAS.

Based on the observation that SSRIs have antidepressant actions, genetic polymorphisms in the serotonin transporter gene have been the subject of intense research activity. Some promising initial findings included (1) an association between a serotonin 2A-receptor gene variant and citalopram antidepressant treatment response (individuals homozygous for the A allele had an increased rate of treatment response); and (2) glutamate receptor SNPs associated with the development of suicidal ideation during citalopram treatment. Only the serotonin 2A-receptor finding has been replicated, though showing an inverse allelic association as compared to the initial finding.

In many cases, it is possible to identify associations between particular genes and characteristics (endophenotypes) that may predispose an individual toward the development of depression. One such endophenotype is hippocampal volume loss. In a cohort of 138 older adults aged 60 and above, genetic variants at three *AGTR1* SNPs (a gene encoding the angiotensin-II type 1 receptor of the renin–angiotensin system [RAS]) were independently associated with accelerated hippocampal volume loss in the right hemisphere over a 4-year follow-up period. In addition, the *AGTR1* risk alleles predicted worse episodic memory performance and interacted

with the *APOE4* allele to accelerate right hippocampal volume loss. These initial findings suggest that targeting molecules in the RAS may have the benefit in late-life depression. Future work will involve investigating the effects of genetic variation on other relevant endophenotypes of MDD and BD in older adults.

Genetic studies using SNP heritability approaches have provided evidence suggesting that some loci may contribute to both conditions, particularly involving calcium channel signaling. GWAS in BD has identified several risk-conferring variants including markers within the *CACNA1C* gene (encoding the α subunit of the L-type voltage-dependent calcium [Ca2+] channel CaV1.2) and the *ANK3* gene (encoding ankyrin G, a cytoskeletal protein involved in scaffolding in the axon initial segment and in nodes of Ranvier), which may be mechanistically linked through providing signaling integrity in GABAergic neurons. Initial findings have shown that variants in these genes may functionally impact neuronal circuitry involved in affective processing in individuals with BD carrying the risk variants. An additional study analyzing four independent genome-wide data sets along with data contributed to the Psychiatric Genomics Consortium Bipolar Group identified the *CACNA1C* gene as well as pathways involved in hormonal regulation, second messenger systems, glutamate systems, and neuronal development, as contributing to the genetic predisposition and underlying neurobiology of BD.

A Model of Late-Life Depression

Accumulation, and perhaps a synergy, of noxious factors is consistent with the clinical reality of geriatric populations in whom a variety of processes contribute to brain dysfunction, leading to depression and to cognitive impairment. Etiological factors, directly or by causing brain abnormalities increasing vulnerability to depression, promote metabolic changes mediating the expression of depressive syndrome. Etiological factors include cerebrovascular lesions, perfusion-related abnormalities, and inflammatory and repair responses. Abnormalities in frontolimbic and frontostriatal functional architecture resulting from these processes may serve as vulnerability factors for depression. Heredity, epigenetic factors, and psychological vulnerability represent additional predisposing factors. Biological responses to stress may promote processes serving as etiological or vulnerability factors of late-life depression.

Brain Structure Studies

Neuroradiological studies suggest frontostriatal and limbic impairment in depression. WMHs prevalent in geriatric depression are mainly located in subcortical structures and the medial-orbitofrontal cortex. Increase in volume of WMHs over a 2-year period predicted lower rates of sustained

late-life MDD remission during the same period; and worsening of white matter grade in the elderly over a 5-year period has been associated with cognitive decline.

Histopathological and neuroimaging studies of individuals with mood disorders have provided evidence of microscopic and macroscopic changes to several brain regions. Low volumes of the subgenual anterior cingulate, the head of the caudate nucleus, the putamen, and the hippocampus have been noted in mixed-age depressed patients. Reduced hippocampal volume has also been found in depressed elderly patients. Small left hippocampal volume in geriatric depression was associated with later development of dementia. Reduction in orbitofrontal cortex volume has been reported in depressed elderly patients compared to normal controls. Abnormalities in neurons of the DLPFC have been observed in unipolar and BD. Moreover, reduction in glia of the subgenual prelimbic anterior cingulate gyrus has been reported in unipolar-depressed patients. Finally, amygdala volume findings in depressed patients have been mixed.

Brain Function Studies

Increased metabolism has been noted in limbic regions, including the amygdala, the rostral-anterior cingulate, and the posterior orbital cortex. In contrast, dorsal neocortical areas, such as the lateral and DLPFC, and the dorsal-anterior cingulate, as well as the caudate have reduced blood flow during depression. Reduced bilateral activation of the dorsal-anterior cingulate and the hippocampus was demonstrated in severely depressed, nondemented elderly patients performing a word-activation task.

The limbic system evaluates reward-related stimuli and organizes stress responses. Reciprocal pathways link limbic structures with striatal and neocortical sites. It has been proposed that abnormalities in the reciprocal relationship between ventral limbic and dorsal cortical neural systems are a critical physiological disturbance in depression. White matter abnormalities compromising connections between limbic and dorsal cortical structures may interfere with limbic–dorsal balance and lead to chronic depressive syndromes.

A recent study provided evidence for insula metabolism as a treatment-specific biomarker in MDD. In this positron emission-based study, insula hypometabolism (relative to whole-brain levels) was associated with remission in response to cognitive-behavioral therapy and poor response to escitalopram, while insula hypermetabolism was associated with remission to escitalopram and poor response to cognitive-behavioral therapy in a mixed-age sample. This finding will require replication in an independent prospective sample and also in an older-aged sample, but offers promise for the use of brain-based biomarkers in guiding treatment selection in MDD.

COURSE AND OUTCOME

Geriatric Depression

Knowledge of risk factors and periods of high risk for particular adverse outcomes enable the clinician to use appropriate diagnostic methods and offer preventive treatments with favorable risk–benefit ratios. Despite advances in antidepressant therapies, longitudinal studies of 1 to 6 years duration suggest that 7 to 30 percent of geriatric patients have a chronic major depression. If partially remitted subjects are considered chronic, the rate of chronicity reaches 40 percent. Chronicity of depression may be predicted by history of a long current episode or long previous episodes, coexisting medical illness, high severity of depression, nonmelancholic presentation, and delusions.

Naturalistic treatment studies of geriatric populations observed a 13 to 19 percent rate of relapse/recurrence at 1 year, while a 15 percent relapse rate is observed in patients receiving controlled antidepressant treatment. This percentage is lower than the relapse/recurrence rate (34 percent) reported in younger adults. However, if the follow-up of elderly patients with depression is extended from 3 to 6 years, the recurrence rate increases to 38 percent. A 20-year follow-up study of mixed-age patients with depression showed that 95 percent of patients had a recurrence. Patients with only minor or subsyndromal depression have an increased risk (approximately fivefold) of major depression compared with unaffected individuals after 1 year. Depression outcomes are also predicted by initial medical burden, self-rated health, and subjective social support.

Relapse/recurrence appears to have different predictors than recovery of depression. The likelihood of relapse/recurrence may be high in patients with history of frequent episodes, late age of illness onset, history of dysthymia, intercurrent medical illness, and possibly high severity and chronicity of the index depressive episode.

Geriatric Bipolar Disorder

In geriatric BD, relapse and hospital readmission after manic episodes are common as episodes become more frequent and the interval between episodes becomes shorter with age. In comparison to younger patients with BD, there is evidence to suggest that geriatric BD patients have less severe manic episodes, are more likely to relapse into depression after a manic episode, are less likely to experience psychotic symptoms, are more likely to experience rapid cycling between episodes, and are less likely to attempt suicide.

EVALUATION AND TREATMENT OF GERIATRIC DEPRESSION AND BIPOLAR DISORDER

Overview of Geriatric Depression Evaluation and Treatment

The objectives of geriatric depression treatment include the following: (1) remission of depression, (2) reduction in the risk of relapse and recurrence, (3) improvement of cognitive and functional status, and (4) development of skills or provision of supports needed for coping with handicaps and psychosocial adversity. Clinicians should expect an improvement in both disability and excess medical morbidity with adequate treatment.

Geriatric depression can be effectively treated with psychotherapy, pharmacotherapy, or combination treatment. Appropriate behavioral rehabilitation techniques should be combined with antidepressant treatment in order to enable elderly patients to regain function as the depressive syndrome subsides. Long-term outcomes of elderly depressed patients can be improved by (1) identifying and treating comorbid psychiatric and medical conditions, (2) minimizing medication side effects, (3) striving for full remission of acute symptoms, and (4) providing psychoeducation for patients, families, and health care professionals about the illness and the importance of adhering to treatment recommendations, both acutely and after symptom remission.

The evaluation of geriatric depression should include assessment of psychopathology, medical and neurological status, functional impairment, and psychosocial factors as well as suicide risk. For instance, dementing disorders, use of drugs that may cause depression (e.g., steroids), and benzodiazepine and alcohol abuse often contribute to late-life depression and worsen its course. Therefore, attention to these problems should be part of the assessment of depressed elders. A further complexity of assessment includes the risk of missing a bipolar depression diagnosis, as such a diagnosis would alter clinical management and contraindicate the use of antidepressant monotherapy. Rating scales, including the Hamilton Depression Rating Scale, the Cornell Scale for Depression in Dementia, the Geriatric Depression Scale, and the Mini-Mental State Examination or Montreal Cognitive Assessment (MOCA) can be useful in assessment.

Adults with comorbid medical illness including both chronic/deteriorating processes and acute/debilitating events are preferentially affected by late-life depression. Further, a number of other medical conditions such as viral infections, endocrinopathies (e.g., thyroid and parathyroid abnormalities), vitamin B_{12} deficiency, and malignancies (e.g., lymphoma and pancreatic cancer) are often associated with depression. Medical conditions likely to respond promptly to treatment (e.g., amphetamine withdrawal) should be treated first and antidepressants used only if depression persists. For many other medical conditions, however, depression may not remit until an antidepressant agent is used. For example, depression in a hypothyroid elderly patient rarely responds to thyroid supplementation alone. Similarly, an antidepressant trial often is

ineffective before hypothyroidism is corrected. Thyroid supplementation and an antidepressant drug, however, offer the highest probability of antidepressant response.

Some drugs can contribute to or cause late-life depression, such as steroids, antiparkinsonian drugs, benzodiazepines, and β-blockers. If the depression occurs in a patient receiving methyldopa (Aldomet), cimetidine (Tagamet), clonidine (Catapres), hydralazine, propranolol (Inderal), or reserpine, switching to another medication is recommended. If the depression occurs in a patient treated with digitalis, estrogens, tamoxifen (Soltamox), vinblastine, or vincristine, the medication should be continued and depression should be treated with antidepressants. There is no agreement on the best strategy for depression that occurs in patients treated with benzodiazepines, corticosteroids, progesterone, propoxyphene (Darvon), or psychostimulants. As interferon-α treatment frequently is associated with depression, increased frequency of psychiatric follow-up and starting or increasing antidepressant treatment prophylactically should be considered before initiating such treatment in patients with a history of depression.

The evaluation of geriatric depression should also include assessments of the psychosocial situation and functional capacity of a patient, as functional difficulties can lead to physical limitations and social isolation, which can further worsen the course of depression. As a more extreme example, poststroke patients experience a "psychosocial storm" characterized by individual strain (e.g., difficulties with ADLs, demanding rehabilitation for the patient) as well as family/caregiver strain (e.g., alteration of usual family roles in accommodating the increased need of the poststroke family member) that increase feelings of helplessness, hopelessness, and depression in both the patient and the family/caregiver. Indeed, depressive symptoms seem to be increased in caregivers for patients with chronic illness, which is particularly pronounced in caregivers for patients with dementia. Other psychosocial factors to consider include race and ethnicity, as there is evidence for disparities in antidepressant medication prescriptions and symptom severity between different racial and ethnic groups. Further, evaluation of adverse childhood experiences (e.g., physical, sexual, and emotional abuse) should be a part of the assessment, as the negative effects of repeated adverse childhood experiences do not diminish in later life and are associated with depression in older adults.

Psychotherapy for Geriatric Depression

For the treatment of major depression in later life, current practice guidelines recommend a combination of pharmacotherapy and psychotherapy, with psychotherapy alone often sufficing as a first-line treatment for minor depression in later life. Psychotherapy may play an

essential role in the care of bereaved depressed elderly patients or in patients who have significant life crises, lack social support, develop disabilities, or lack coping skills to deal with their life situations. Psychotherapy may be helpful in supporting elderly patients to cope with their symptoms and adhere to treatment before antidepressant drugs take effect, and should be considered in patients who cannot tolerate biological treatments. Telepsychiatry may offer an efficacious option for psychotherapy delivery for homebound elders.

A number of psychotherapeutic modalities including interpersonal therapy, cognitive-behavioral therapy, and problem-solving therapy have been found effective in depressed elderly patients without significant cognitive impairment or medical burden. For patients with executive dysfunction, a number of psychotherapies have been found effective including a modification of problem-solving therapy (PST) where the therapist is more directive, the focus is on less complex problems to facilitate learning, and the program is more structured. Additionally, problem-adaptation therapy (PATH) was developed for older individuals with major depression and intermediate cognitive impairment and disability; this modality combines problem-solving therapy, environmental adaptation tools, and caregiver involvement to target the behavioral limitations inherent in depression and disability.

Extending the focus on the environment, ecosystem-focused therapy (EFT), has been developed to support patients through the "psychosocial storm" of an acute medical illness where depression is often a major contributor to disability and a hindrance in recovery. EFT has been developed as a treatment for poststroke depression. A meta-analysis showed that antidepressants are only weakly efficacious in poststroke depression and the efficacy of conventional psychotherapies is questionable. The aim of EFT is to help the patient develop new perspectives and skills to confront the significant changes to their ecosystem brought on by these acute changes. A preliminary study showed that EFT is more efficacious than education in stroke and depression.

Although psychotherapy has proven to be an effective treatment in this population, it appears to be an underutilized resource. One idea behind the underutilization is that the complexity of available psychotherapies makes their implementation difficult in a general community setting. To streamline both the delivery and language of psychotherapy interventions, it has been proposed that neurobiological constructs be used to identify biologically meaningful behavioral targets that can serve as the basis for simplified and effective treatments. This view of treatment is supported by the RDoC project and has been successfully applied to treatment of late-life depression with the development of "Engage," a psychotherapy using simple-to-learn strategies that target accepted neurobiological domains of late-life depression.

Pharmacotherapy for Geriatric Depression

Overview. Elderly patients benefit from the same psychopharmacological agents as younger patients. Benzodiazepines should generally be avoided. Residual insomnia in depressed elderly patients should be treated with trazodone (Desyrel) and, if it fails, with zolpidem (Ambien), or zaleplon (Sonata). Residual anxiety should be treated with an increase of the dosage of the antidepressant to maximum levels rather than prescribing a benzodiazepine.

Clinical practice guidelines recommend that SSRIs and other second-generation antidepressants (e.g., serotonin-norepinephrine reuptake inhibitors [SNRIs], bupropion, mirtazapine) are the first-line treatment for late-life depression. Late-life depression responds more slowly to antidepressants than depression of younger adults. Executive dysfunction, slow cognitive processing speed, WMH, microstructural white matter abnormalities, and small hippocampal volumes each predict a slower and less complete response to antidepressant treatment. Elderly patients who respond to antidepressant treatment usually require 8 to 12 weeks to achieve full remission. The minimal length of an antidepressant trial should be 3 to 4 weeks before a switch to another antidepressant or an augmentation agent is used. Patients who show no signs of improvement by the 4th week, however, will not likely respond to further treatment with the same agent.

Selective Serotonin Reuptake Inhibitors. Studies of elderly outpatients have shown that SSRIs are equally effective as tertiary amine tricyclics in the acute treatment of late-life depression. The dosages of SSRIs should be increased gradually. The starting daily dosages are 10 to 20 mg for fluoxetine, paroxetine, and citalopram, 5 to 10 mg for escitalopram, and 25 to 50 mg for sertraline. For some patients, daily dosages of 20 mg of fluoxetine, paroxetine, and citalopram, 10 mg of escitalopram, and 100 mg of sertraline are sufficient, although higher dosages may be required. The most frequent side effects of SSRIs are insomnia, akathisia, nausea, anorexia, and pseudoparkinsonism. Inappropriate secretion of antidiuretic hormone (SIADH) leading to hyponatremia is much more common in the elderly than in younger adults. Further, SSRIs may lead to an increased risk of bleeding, fragility fractures, poor mobility, and stroke.

Drug interactions should be considered in elderly patients receiving SSRIs. Fluoxetine and paroxetine inhibit the cytochrome P450 2D6 liver cytochrome isoenzyme. This pathway is essential for the hydroxylation of nortriptyline and desipramine, and the metabolism of antipsychotics and type 1A antiarrhythmic drugs (encainide, flecainide), β-blockers, and verapamil, whose plasma levels may be raised when fluoxetine or paroxetine is administered. The metabolite of fluoxetine, norfluoxetine, is

an inhibitor of the 3A4 isoenzyme. The 3A4 isoenzyme is responsible for the metabolism of alprazolam (Xanax), carbamazepine (Tegretol), quinidine, erythromycin, and astemizole and may lead to an increase in the plasma level of these agents. Sertraline, citalopram, and escitalopram have a weak inhibitory action of the 2D6 isoenzyme and do not inhibit the 3A4. Fluvoxamine (Luvox) inhibits the P450 3A4 and the 1A2 cytochrome isoenzymes, but does not significantly inhibit the 2D6 isoenzyme. Drugs metabolized by the 3A4 should be avoided in patients receiving fluvoxamine. Similarly, theophylline should be cautiously prescribed in fluvoxamine-treated patients since fluvoxamine may produce a threefold decrease in theophylline clearance by inhibiting the 1A2 isoenzyme.

Serotonin-Norepinephrine Reuptake Inhibitors. Venlafaxine XR (extended release) leads to high remission rates (asymptomatic states) and has been found effective in hospitalized depressed patients as well as in drug-resistant depression and in depressed patients with chronic pain. For this reason, venlafaxine should be considered in severe depression and in depression unresponsive to other agents. Elderly patients appear to require dosages comparable to those of younger adults. Daily dosages of 112.5 to 225.0 mg are adequate for the majority of elderly patients. Nausea, a rather frequent side effect of venlafaxine, can be minimized by a slow increase of the daily dosage. Blood pressure should be monitored, especially in patients receiving dosages above 225 mg daily, since venlafaxine can increase blood pressure. Venlafaxine has limited drug–drug interactions and may be selected in elderly patients who require treatment with drugs interacting with other antidepressants. Venlafaxine has a short elimination half-life but the extended-release preparation can be administered in a single dose. Nonetheless, abrupt discontinuation of venlafaxine XR may cause a withdrawal syndrome.

Duloxetine, an SNRI with relatively balanced serotonergic and noradrenergic action, has demonstrated efficacy in late-life depression as well as treatment of comorbid pain. Following a week of 30 mg/day, the dose can usually be increased to 60 mg/day after a week of treatment, if tolerated. For patients who remit on a dose of 60 mg, remission will typically occur by week 12 of treatment with additional time on that dose not associated with an increased chance of remission. For patients who remit on a dose of 60 mg, remission will typically occur by week 12 of treatment; additional time o that dose is not associated with an increased chance of remission.

Levomilnacipran is an SNRI with higher selectivity for norepinephrine reuptake inhibition than other SNRIs and has proven effective for major depression, although studies specific to late-life depression are lacking. However, there is evidence to suggest that individuals older than 60 years of age may have a greater antidepressant response to this medication than

younger individuals. Levomilnacipran is largely metabolized by cytochrome P450 3A4. The medication is intended for once-daily dosing with a starting dose of 20 mg daily. The target dose range is 40 to 120 mg daily.

Tricyclic Antidepressants. The secondary amines, nortriptyline and desipramine, are the most frequently used TCAs in geriatric depression. They have lower anticholinergic and sedative effects than the tertiary amines amitriptyline, doxepin, and imipramine. Nortriptyline appears to have a lower potential for orthostatic hypotension than do other TCAs. The plasma levels of antidepressants required for the treatment of depressed elderly patients are similar to those needed by young adults (nortriptyline plasma levels of 60 to 150 ng/mL and desipramine above 115 ng/mL). Elderly patients often develop therapeutic blood levels of nortriptyline or desipramine while on low daily dosages (nortriptyline 1.0 to 1.2 mg/kg of body weight and desipramine 1.5 to 2.0 mg/kg). Older patients are more likely than younger patients to develop delirium, constipation, urinary retention, dry mouth, and orthostatic hypotension. For this reason, the dosages of antidepressants should be increased at slower pace than in younger adults. Elderly patients on nortriptyline often develop higher plasma levels of the nortriptyline metabolite 10-hydroxy-nortriptyline than do young adults, even when the plasma levels of the parent compound are similar. High 10-hydroxy-nortriptyline levels may contribute to cardiac conduction defects in elderly patients.

Orthostatic blood pressure, the electrocardiogram (ECG), and urination should be monitored frequently in elderly patients receiving nortriptyline or desipramine. Pretreatment systolic orthostatic hypotension has been found to correlate with antidepressant response to nortriptyline in some elderly patients. Therefore, nortriptyline, rather than desipramine, should be considered in such patients and orthostatic blood pressure and subjective symptoms of orthostasis should be monitored carefully. TCAs have anticholinergic properties. For this reason, these drugs should be avoided in patients with prostatic hypertrophy or patients with narrow angle glaucoma. Nortriptyline and desipramine have properties similar to those of type 1A antiarrhythmic drugs (quinidine-like drugs). When administered to patients with right or left bundle branch block, tricyclics may cause second-degree block in approximately 10 percent of cases. For this reason, an ECG should always precede the use of tricyclics in the elderly. A multicenter study (CAST-II) demonstrated that type 1A antiarrhythmics increase cardiac mortality in postmyocardial infarction patients. Although TCAs were not used in this study, the type 1A antiarrhythmic properties of TCAs suggest cautious use in depressed patients with ischemic heart disease.

MAO Inhibitors. Limited information exists on the use of these agents

in the old-old. The MAO inhibitors, however, have been studied in depressed patients who are in their late 60s to early 70s. MAO inhibitors are effective in patients with major depression as well as depressed patients with panic attacks. Low dosages of MAO inhibitors (e.g., phenelzine [Nardil] 30 to 45 mg daily or tranylcypromine 20 to 30 mg daily) should be used in the elderly. Orthostatic hypotension is the most frequent important side effect of MAO inhibitors. This side effect is of concern in the elderly because it may lead to falls and fractures especially of the hip or the humerus. Other side effects include weight gain, insomnia, lack of energy, and daytime somnolence in phenelzine-treated patients and nervousness, insomnia, and excessive perspiration in tranylcypromine-treated patients. Peripheral neuropathy occurs in a small percentage of patients on MAO inhibitors and often responds to pyridoxine. Sympathomimetic amines, monoamine precursors, TCAs, SSRIs, venlafaxine, concomitant administration of two MAO inhibitors, and tyramine-rich food may cause a hypertensive crisis and should be avoided in patients on MAO inhibitors. The selegiline patch (Emsam) does not require diet restrictions at the lowest 6-mg dose and may be safer during diet violation than other MAOIs at higher doses. The advantage of transdermal administration is that intestinal monoamine oxidase is not inactivated to the same degree as for oral administration, reducing the risk of hypertensive crisis triggered by ingested monoamines.

Other Antidepressants. Bupropion, a noradrenaline and dopamine reuptake inhibitor, has proven efficacy in younger adults as well as older adults with major depression based on a number of limited trials of the IR, SR, and XL formulations. Bupropion does not cause cognitive impairment or sedation, is safe in overdose, and lacks significant cardiotoxicity; it should be considered in elderly cardiac patients. Bupropion has few drug interactions, but it should not be prescribed in patients receiving MAO inhibitors. Bupropion may exacerbate pre-existing hypertension and regular monitoring of blood pressure is required. Seizures have been reported in 0.4 percent of patients treated with bupropion, but the risk of seizures can be minimized by the use of sustained release preparations, slow introduction of bupropion (100 mg daily), and restriction of the total daily dosage to 450 mg. Most elderly patients require a total daily dose of 300 mg.

Mirtazapine has been increasingly used in frail depressed nursing home patients with anorexia. Mirtazapine antagonizes the α-2 presynaptic inhibitory receptor as well as the 5-HT2 and the 5-HT3 serotonin receptors. Mirtazapine is an antagonist of histaminergic receptors and has a mild antagonist effect on the α-1 adrenergic and muscarinic receptors. These effects account for the sedative, hypotensive, and anticholinergic side effects of mirtazapine. Mirtazapine is not an inhibitor of the P450

cytochrome. Side effects include somnolence, dizziness, increased appetite, weight gain, and hyponatremia. Mild increase of cholesterol and triglycerides has been observed in 15 and 6 percent of patients, respectively. The starting dosage of mirtazapine is 15 mg and should be administered at bedtime to reduce sedation. Daily dosages from 30 to 45 mg are sufficient as a rule for the treatment of late-life depression.

Vortioxetine inhibits the serotonin reuptake protein, is an agonist of the serotonin 1A and an antagonist of the serotonin 3 receptor. It is also an antagonist of the serotonin 1D and 7 receptors and a partial agonist of the serotonin 1B receptors, although with somewhat lower affinity. Vortioxetine has been found effective in late-life depression and it may improve memory and executive dysfunction, although the clinical significance of the cognitive findings is unclear at this point. The starting dose is 10 mg daily, and a range of 5 to 20 mg dosages is efficacious. Vortioxetine neither inhibits nor augments liver enzymes. But it is metabolized by the 2D6 isoenzyme and its dosage should be adjusted when coadministered with 2D6 inhibitors or inducers.

Psychostimulants may be ineffective in elderly patients with major depression. Dextroamphetamine and methylphenidate (Ritalin), however, appear to improve apathy and anergy in medical patients or cognitively impaired patients. Dextroamphetamine or methylphenidate at dosages 5 to 10 mg in the morning and at noon is sufficient. Psychostimulants have rapid onset of action, minimal side effects, limited potential for tolerance, and low risk for addiction.

Complementary and Alternative Medicines. There are a number of natural supplements with some evidence of efficacy in geriatric populations. For instance, omega-3 fatty acids, St. John wort, and *S*-adenosyl-*L*-methionine (SAMe) have shown some benefit for depression in a limited number of small studies in older adults. St. John wort must be used with caution in geriatric populations due to drug–drug interactions through induction of cytochrome P450 3A4.

Somatic Treatments

ECT. ECT should be considered in both geriatric depression and mania because of its efficacy, rapid onset of action, and safety. ECT may be chosen in patients with severe mood syndromes unable to tolerate the long waiting imposed by the gradual introduction of antidepressants or mood stabilizers and the slow onset of drug action in the elderly. In addition, ECT can be effective in treating neuropsychiatric comorbidities such as catatonia and parkinsonism often present in the depressed elderly.

ECT treatment characteristics must be chosen with an eye to maximizing treatment efficacy while minimizing cognitive side effects. An example of one such consideration is that ECT administration as often as

three times weekly appears beneficial in producing a more rapid recovery than ECT given twice or once a week, yet less frequent ECT is associated with less cognitive impairment in the time period shortly after ECT. No differences in cognition have been found, however, 1 month after ECT completion between groups that had ECT thrice or twice weekly. ECT treatment variables to consider are electrode placement (right unilateral vs. bilateral), treatment frequency, and stimulus intensity (energy), duration, pulse width, and pulse frequency. The effects of these choices are described elsewhere in this book. High-dose, ultra-brief, right unilateral ECT treatment has become a common choice due to reports of equal efficacy and reduced cognitive side effects compared with bilateral ECT.

The mortality rate for all ages following a single ECT treatment is approximately 1 per 80,000 treatments, corresponding to 1 per 10,000 ECT-treated patients. This rate may be higher in the elderly with severe medical conditions, but it appears to be lower than that for some medication treatments (TCAs), and there is some evidence that it may be compensated for by decreases in mortality following hospitalization (compared with depressed patients receiving other forms of treatment or no treatment at all). The majority of deaths (67 percent) related to ECT are due to cardiac complications and occur immediately after ECT or within a few hours from treatment. Compared to younger adults, elderly patients and patients with pre-existing cardiac disease are more prone to cardiovascular events occurring in the context of ECT, including ischemic syndromes, arrhythmias, decompensation of heart failure, and transient severe rise in blood pressure. With adequate medical evaluation, however, monitoring during and after ECT, and appropriate intervention, most cardiovascular events related to ECT have a benign outcome.

Geriatric patients require a longer time to recover their memory than younger adults, especially after bilateral ECT. Moreover, geriatric patients sometimes develop prolonged confusion after ECT. The increase in falls with ECT has been reported as up to 15 percent of elderly patients. Controlled studies, however, are needed in order to examine whether and to what extent ECT-related confusion is responsible for these falls. The practice of favoring ECT over antidepressants in medically ill depressed patients may have led to inflated reports of ECT-related morbidity. Anecdotal literature suggests that ECT was used uneventfully in patients over 100 years old and in elderly patients with aortic aneurysm, cardiac pacemaker, history of myocardial infarction or stroke, transient ischemic attacks, severe hypertension, arrhythmias, or need for anticoagulants.

Other Therapies

Vagal nerve stimulation (VNS) is an FDA-approved treatment for chronic and recurrent depression. Through an implanted generator, electrical signals are delivered to the left cervical vagus nerve where it is postulated

that stimulation of the afferent fibers leads to modulation of mood-related networks (e.g., amygdala and locus coeruleus). Although studies of efficacy in geriatric depression are lacking, VNS does not carry an increased risk of surgical complications and does not necessarily increase the risk for cardiac abnormalities in this population. Transcranial magnetic stimulation (TMS) is also FDA approved for the treatment of depression. In this noninvasive procedure, an electromagnetic field is generated that depolarizes neurons leading to modulation of cortical activity. Although this procedure is well tolerated with minimal cognitive side effects, there is some evidence that TMS is less effective in older patients. Another preliminary study with TMS, however, showed significant reduction of depressive symptoms in poststroke depression. Bright light treatment is another noninvasive treatment with some evidence of efficacy in geriatric depression.

Treatment of Bereavement

Most bereaved elderly individuals do not develop clinically significant depressive symptoms. For this reason, prophylactic treatment is not recommended for the entire bereaved elderly population. Self-help groups, counseling, group therapy, and individual psychotherapy should be considered for individuals with an intense reaction to loss or history of depression. Bereaved persons with depressive symptomatology that does not meet criteria for major depression may benefit from brief focused psychotherapy. Interpersonal therapy lends itself well for this purpose since it addresses issues related to loss, depression, and change of roles. Although antidepressant drugs can be effective in bereaved elderly with depressive symptoms, clear guidelines are not available. If major depression develops during the bereavement period, antidepressant treatment should be considered. Some clinicians suggest that psychotherapy and/or antidepressant drugs should be administered if bereavement lasts longer than 1 year.

Complicated grief, a persistent complex bereavement disorder characterized by prolonged acute grief, feeling that life has no meaning, avoidance behaviors, and dysregulated emotions, has been identified in DMS-V as a condition in need of further study. Complicated grief treatment (CGT) has been developed as a 16-week treatment with components of prolonged exposure, IPT, and motivational interviewing and has been found to be more efficacious in symptom improvement than IPT in older adults. It appears that antidepressants may enhance the response to CGT but that antidepressants alone may be ineffective.

Treatment of Psychotic Depression

For treatment of psychotic depression, the recommended treatment includes combinations of antidepressants (SSRIs or venlafaxine XR) and

atypical neuroleptics (except clozapine), although there are limited clinical trials for older adults. In one multicenter 12-week double-blind randomized controlled trial (STOP-PD), combination pharmacotherapy with olanzapine and sertraline was associated with higher remission rates during the trial period compared to olanzapine plus placebo in patients. ECT treatment is another option, especially if combination drug treatment fails or is poorly tolerated.

Treatment of Depression in Dementia

In elders with major depression and mild-to-moderate dementia, the historical treatment of choice was combination of an antidepressant with psychosocial intervention, although there are limited data to support the use of antidepressant medication. Indeed in a large, well-powered study not only were antidepressants (i.e., sertraline and mirtazapine) found ineffective in depressed demented patients, but they also were associated with more adverse effects than placebo. Psychosocial interventions such as PATH and caregiver-focused treatment may be a better option for first-line treatment in this population.

Treatment-Resistant Depression

Change in antidepressant strategy should take into consideration that depressed older patients take longer time to respond to antidepressants than younger adults. If a patient has had little or no response to an adequate antidepressant trial, switching to a different antidepressant rather than using an augmenting agent is an appropriate first step. If a patient has had little or no response to an initial SSRI, the patient may be switched to an SNRI, or bupropion SR or XL. If a patient has had little or no response to venlafaxine, the patient may be switched to an SSRI. If the patient has had little or no response to a TCA or bupropion, switching to an SSRI or SNRI is appropriate. Although studies in geriatric subjects are lacking, research in younger patients suggests that switching from one SSRI to another is an effective strategy.

Augmentation therapies are recommended in depressed elders with partial response to the initial antidepressant, although there is limited agreement on the use of augmentation therapies. Bupropion, lithium, or nortriptyline may be added if the initial antidepressant is an SSRI. If the initial antidepressant is a TCA, lithium or an SSRI may be added. If the patient was on bupropion, an SSRI or lithium may be added. An additional augmentation strategy in older adults is the combination of an SSRI or SNRI plus mirtazapine. In a study of middle-aged adults (mean age in mid-40s), mirtazapine (30 mg) plus fluoxetine (20 mg), venlafaxine (225 mg) or bupropion (150 mg) achieved greater remission rates (46 to 58 percent) after 6 weeks than fluoxetine monotherapy (25 percent). Among those who

responded, double-blind discontinuation of one agent led to relapse in 40 percent of the cases.

Addition of methylphenidate may be helpful to accelerate treatment response to an SSRI antidepressant. Treatment with an atypical neuroleptic, alone or in combination with other medications, is sometimes effective for treatment-resistant depression. Augmentation of antidepressants with aripiprazole or quetiapine XR has been found effective in younger adults, but controlled studies are lacking in older adults. It should be noted, however, that the use of antipsychotic agents (both atypical and typical) in older individuals with dementia has been associated with a greater risk of mortality and increased risk of cerebrovascular events. When medication treatment fails, ECT may offer the best chance for response and remission.

Continuation and Maintenance Treatment

The intent of continuation treatment (6 months) is to prevent relapse of the acute episode, while maintenance treatment is offered after the first 6 months of remission in order to prevent a new depressive episode. Continuation and maintenance treatment should use the same dosages of antidepressants as those of acute treatment. Continuation treatment should be offered to all patients with late-life depression. Patients at high risk for recurrence should have maintenance treatment. The strongest predictor of recurrence is a history of three or more lifetime episodes of depression. For such patients, experts recommend treatment for 3 years or longer. Disagreement exists concerning the optimal maintenance treatment duration in elders with fewer episodes. Other predictors of depression recurrence are high severity of the initial episode and persisting anxiety, and in mixed-age populations, late age of illness onset is a strong predictor of recurrence. In patients with psychotic depression who achieve remission following treatment with an antidepressant and an antipsychotic, continuation of the antipsychotic drug is recommended, typically for 6 months, but some experts recommend shorter or longer treatment. For depressed patients who failed antidepressants but responded to ECT, antidepressants (other than those to which the patient previously failed to respond) plus a mood stabilizer should be used for continuation/maintenance treatment. Continuation/maintenance ECT is another option.

Adherence. Approximately 70 percent of patients fail to take 25 to 50 percent of their medications. Older patients and their families often have difficulty accepting the diagnosis of depression and understanding the need for long-term treatment. Concurrent medical illnesses may interfere with antidepressant response or attainment of adequate dosages. Alcohol and other substance use disorders may undercut pharmacotherapy. Difficulties

accessing health care may hinder the ability of elderly, especially functionally dependent elderly, to obtain adequate treatment.

Care Delivery

Several interventions have been developed to improve the treatment of geriatric depression in the primary care setting. The Improving Mood-Promoting Access to Collaborative Treatment (IMPACT) study focused on older adults and offered access for up to 12 months to a depression care manager who was supervised by a psychiatrist and an expert primary care physician. The depression care manager offered education, care management, and support of antidepressant management by the primary care physician or problem-solving therapy (a brief structured psychotherapy). Depressed elderly primary care patients receiving the IMPACT intervention were more likely to receive antidepressants or psychotherapy than usual care patients. They had lower depressive symptomatology and higher overall function, response and remission rates, quality of life, and satisfaction with treatment. These benefits were sustained 12 months after the end of the intervention.

Positive outcomes have also been reported by the Prevention of Suicide in Primary Care Elderly: Collaborative Trial (PROSPECT). The lynchpin of the PROSPECT intervention was the use of care managers, who were trained social workers, nurses, and psychologists who worked alongside with primary care physicians. They followed the patients and monitored their psychiatric symptoms and signs, treatment adherence, and side effects, and provided appropriately timed and targeted recommendations to physicians according to an algorithm. Patients treated in practices offering the PROSPECT intervention had a higher likelihood to receive antidepressant treatment, a greater decline in the rate of suicidal ideation, and a more favorable course of depression over a period of 24 months. The beneficial effects of the PROSPECT intervention were most pronounced in primary care patients with major depression rather than minor depression.

The concept of an integrated care model has been extended to depressed patients with chronic medical illnesses. The personalized adherence intervention for depression and COPD (PID-C) is an integrated approach implemented by care managers to improve adherence to treatment for both COPD and depression with a focus on helping patients to overcome barriers to treatment and offering interventions that support patients in using treatment prescribed by their own physicians. PID-C reduced depressive symptoms and dyspnea-related disability more than usual care.

Overview of Geriatric Bipolar Disorder Evaluation and Treatment

The goals of geriatric BD include symptom amelioration (i.e., resolution of

mania, remission of depression, and maintenance of euthymia) as well as improvements in functional status and prevention of recurrences. The evaluation of geriatric BD should include a comprehensive assessment exploring comorbidity (psychiatric, medical, and neurological), functional status, and psychosocial factors. Further, substance use disorders are common in geriatric BD patients and should be routinely assessed and treatment offered in the evaluation period. Medications should be evaluated as well, given that medications such as corticosteroids and dopaminergic agents such as stimulants and antiparkinsonian drugs can precipitate mania.

Bipolar elderly patients often have limited, slow, or fragile treatment response. Guidelines for dosage and duration of treatment with antimanic agents cannot be extrapolated directly from younger adult studies because aging, comorbid neurological diseases, high medical burden, and poor social support may directly or indirectly, through poor adherence, interfere with treatment response of late-life BD and increase vulnerability to side effects.

Pharmacotherapy for Geriatric Bipolar Disorder

Lithium is the best investigated drug in geriatric mania. Lithium may be less effective in mania with comorbid neurological and medical disorders, than in uncomplicated mania. Age-associated decline in renal clearance in elderly patients leads to high lithium plasma levels even when low dosages are used. Lithium should be introduced slowly, starting with daily dosages of 150 mg. About one-half of the dosage required for young adults is sufficient for the elderly. The half-life of lithium is about 24 to 36 hours in patients who are in their 70s. Therefore, steady-state pharmacokinetics is anticipated 5 or more days after the stabilization of the daily dosage. In some elderly patients, lithium plasma levels of 0.3 to 0.6 mEq/L were found to be effective. Controlled studies, however, are needed to establish the efficacy of low lithium plasma levels in the elderly. The onset of action of lithium is slow and may require several days or weeks.

The side effects of lithium are pronounced in geriatric patients. Elderly patients often develop a fine tremor or myoclonus at lithium plasma levels considered nontoxic in young adults. Lithium can produce parkinsonian symptoms or worsen pre-existing Parkinson disease. Delirium can develop at plasma levels below the therapeutic range. Neuroleptics increase the risk of delirium in elderly patients on lithium. Lithium may worsen cognitive function, especially in demented patients. Sinoatrial block may be caused by lithium in predisposed patients; digitalis and β-blockers increase the risk for sinoatrial block. Salt depletion, thiazide diuretics, and nonsteroidal anti-inflammatory drugs may raise the lithium plasma levels and result in toxicity. Chronic kidney disease appears to be common in older adults treated with lithium.

The antiepileptics sodium divalproex, carbamazepine, and lamotrigine are effective for preventative treatment in BD. Anticonvulsants are effective in both lithium responders and lithium nonresponders. Anticonvulsants may be more effective than lithium in patients with rapidly cycling mania or dysphoric mania. Although there is limited research in elderly populations, they should be considered especially in patients at high risk for lithium toxicity or history of poor response to lithium. Drowsiness and sedation are frequent dose-dependent side effects of valproate and carbamazepine. Many patients are habituated to sedation. Therefore, slow introduction of valproate or carbamazepine may reduce this problem. In the elderly, the principal side effects of valproate are gastrointestinal distress and ataxia. Frequent side effects of carbamazepine are ataxia and confusion. Rarely, carbamazepine can cause hyponatremia. Approximately 0.4 percent of patients receiving valproate and 2.0 percent of patients receiving carbamazepine develop leukopenia with a white count below 4,000/mm^3. Valproate may cause thrombocytopenia in older patients much more frequently than in younger adults. Carbamazepine-treated patients have a reduction in their white blood cell count during the first 2 weeks. For this reason, periodic monitoring of the white count is indicated especially during the early phases of treatment.

Lamotrigine is generally well tolerated and does not cause weight gain, sedation, sexual side effects, or cognitive impairment. It helps mainly for depression, which for most bipolar patients is a greater concern than mania or hypomania. Mild rash is common and the risk of serious, but not life-threatening rash is approximately 1 in 250. The risk of Stevens–Johnson syndrome is approximately 1 in 1,000, or perhaps lower, with cautious uptitration as recommended to reduce the rash risk. Lamotrigine should be stopped immediately if a rash develops. Simultaneous treatment with valproate can more than double lamotrigine's half-life, so such patients should start at no more than 25 mg every other day with dose increases no more than once every 2 weeks in the first weeks. Effective doses are usually 100 to 200 mg/day. Thus, a drawback with lamotrigine is that it can take weeks to reach a therapeutic dose.

Agitated manic elderly patients may be treated with low dosages of lorazepam. Lorazepam is favored over other benzodiazepines because it undergoes only phase II metabolism (glucuronidation), a process minimally affected by aging. Effective alternatives to lorazepam are atypical and high-potency neuroleptics in oral (e.g., risperidone 1 to 2 mg, olanzapine 5 to 10 mg daily) or injectable (e.g., haloperidol 0.5 to 5 mg daily) form. These drugs should be used early in the treatment of acute mania and should be withdrawn when the agitation is controlled and therapeutic plasma levels of a mood stabilizer are established.

In addition to agitated mania, antipsychotic agents may also be useful in the treatment of mania as well as bipolar depression. For instance, there

is evidence that atypical antipsychotics including olanzapine, risperidone, and quetiapine may be effective in the treatment of geriatric mania while quetiapine and lurasidone may have some efficacy in bipolar depression. The increased risk of weight gain, hyperlipidemia, and diabetes associated with at least some atypical neuroleptics raises the risk of cardiovascular disease.

Antidepressant use in BD is controversial due to concern about antidepressant-induced mania or mixed states, although studies of such complications are limited in older populations. Recent work has shown that antidepressant use is as high as 72 percent in older adults with BD.

Other Treatments for Geriatric Bipolar Disorder

ECT is efficacious in late-life bipolar depression. Limited data support the efficacy of ECT in the treatment of mania in older adults. TMS may be a treatment option for geriatric bipolar depression, but further study is needed to further clarify the efficacy and safety in this population. Cognitive-behavioral therapy, family-focused therapy, and interpersonal social rhythm therapy have been shown to be effective as adjuncts to medications in shortening the time to remission from a depressive episode. Psychosocial interventions targeting adherence to treatment such as the customized adherence enhancement (CAE) therapy have been found effective in geriatric BD in improving treatment adherence, reducing mood-related symptoms and signs and global psychopathology, and improving functioning.

FUTURE DIRECTIONS

The medical and neurological comorbidity of geriatric mood disorders, although it complicates care, has served as a window to mechanisms contributing to geriatric mood disorders. The clinical association of depressive symptoms with impairment in specific cognitive functions has been a central stimulus to research in geriatric depression. Studies of the integrity of white matter in distinct brain regions have begun to identify abnormalities leading to disconnection states contributing to late-life depression and executive dysfunction. Advances in neuroimaging and electrophysiology make possible the identification of structural, functional, and processing abnormalities predisposing depressed elderly patients to poor response to antidepressants or to mood stabilizers. Genetics research identified some polymorphisms that increase the risk of resistance to some antidepressants. Neuroimaging studies are underway to identify the mechanisms by which genetic factors increase resistance to antidepressants and also to investigate endophenotypes of geriatric mood disorders and biomarkers that may predict response to specific treatment interventions.

With the recently advanced hypotheses of geriatric depression, there is

the promise to guide novel treatment development. For example, studies of the "DED syndrome" demonstrated that executive dysfunction and structural and functional abnormalities of the cognitive control system are associated with poor response to antidepressant drugs. However, appropriately modified problem-solving therapy reduced both depression and disability more than supportive therapy in depressed, nondemented patients with major depression and executive dysfunction. PATH was developed in order to address the clinical needs of depressed older adults with cognitive impairment up to moderate dementia. A neuroplasticity-based computerized cognitive remediation-geriatric depression treatment (nCCR-GD) was developed to target executive dysfunction in late-life depression. The underlying assumption was that remediation of these deficits may improve the functional brain networks shared by ED and depression. In a preliminary study, nCCR-treated depressed older adults showed greater improvement in executive functions than historical controls treated with escitalopram. Future work may investigate whether nCCR alone or in combination with TMS or pharmacotherapy may ameliorate specific cognitive dysfunction and improve outcomes in geriatric depression and BD.

While the inflammation hypothesis of late-life depression has generated interest in the role of anti-inflammatory agents in the treatment of late-life depression, many questions remain. The position of inflammation in the chain of biological events leading to late-life depression needs to be clarified. Assuming that the brain's inflammatory responses have regional specificity, it is unclear whether the localization of these responses influences the clinical profile of late-life depression and its course. In addition, the role of peripheral inflammatory indices in determining the choice of antidepressants needs to be further understood.

Minocycline, a microglia inhibitor, has been shown to reduce depression-like behavior in animal models of depression. Anecdotal evidence suggests that minocycline can augment the action of antidepressants in major depression. Etanercept, an agent that reduces TNF-α–mediated cellular response, reduced symptoms of depression in a placebo-controlled study of patients with psoriasis. The cyclooxygenase-2 inhibitor, celecoxib, has been found to augment the efficacy of reboxetine and fluoxetine in patients with MDD. In treatment-resistant major depression, baseline high concentrations of high-sensitivity CRP predicted favorable response to the TNF antagonist, infliximab. Adequately powered randomized controlled trials need to clarify which anti-inflammatory agents have antidepressant action and for whom.

Studies need to ascertain whether antihypertensive, antihypercholesterolemic, and antiplatelet agents and diabetes control improve the outcome of vascular depression. Longitudinal assessment of depressive symptomatology during antiplatelet secondary prevention trials

can evaluate the efficacy of these agents in improving the course of vascular depression. Modulators of the RAS improve cerebral perfusion and may improve cognition, although they have not been studied in vascular depression. Finally, anti-inflammatory agents may have a role in vascular depression but studies are lacking.

Advances in biology and psychotherapeutic interventions, along with clinical and psychosocial studies of the comorbidity and social complexity of geriatric mood disorders, are generating knowledge on which to build comprehensive models of care. For programs such as PID-C to become part of the mainstay of clinical treatment of late-life depression, their efficacy, implementation potential, and fiscal feasibility need to be demonstrated. The IMPACT and PROSPECT interventions are concrete examples of treatment models that address the clinical complexity of geriatric depression and the barriers imposed by the primary care setting. Reimbursement constraints remain a barrier to the implementation of such programs. Studies are underway that build on the principles of these models and focus on existing services supported by the available reimbursement system. Such studies offer focused training to clinicians and interventions at the organizational level to increase functional collaborative relationships between primary care practices and already funded independent mental health clinics. Further research is needed to find efficient interventions and to weigh costs versus benefits in various elderly populations with distinct needs.

References

Alexopoulos GS. The vascular depression hypothesis: 10 years later. *Biol Psychiatry*. 2006;60(12):1304–1305.

Alexopoulos GS. Clinical and neurobiological findings and novel treatment developments in late-life depression. *Psychiatric Ann*. 2014;44:126–129.

Alexopoulos GS, Arean P. A model for streamlining psychotherapy in the RDoC era: the example of 'Engage.' *Mol Psychiatry*. 2014;19(1):14–19.

Alexopoulos GS, Hoptman MJ, Kanellopoulos D, Murphy CF, Lim KO, Gunning FM. Functional connectivity in the cognitive control network and the default mode network in late-life depression. *J Affect Disord*. 2012;139(1):56–65.

Alexopoulos GS, Kiosses DN, Heo M, Murphy CF, Shanmugham B, Gunning-Dixon F. Executive dysfunction and the course of geriatric depression. *Biol Psychiatry*. 2005;58(3):204–210.

Alexopoulos GS, Morimoto SS. The inflammation hypothesis in geriatric depression. *Int J Geriatr Psychiatry*. 2011;26(11):1109–1118.

Alexopoulos GS, Murphy CF, Gunning-Dixon FM, et al. Serotonin transporter polymorphisms, microstructural white matter abnormalities and remission of geriatric depression. *J Affect Disord*. 2009;119(1–3):132–141.

Alexopoulos GS, Reynolds CF 3rd, Bruce ML, et al; PROSPECT Group. Reducing suicidal ideation and depression in older primary care patients: 24-month outcomes of the PROSPECT

study. *Am J Psychiatry*. 2009;166:882–890.

Avari JN, Alexopoulos GS. Models of care for late-life depression of the medically ill: examples from chronic obstructive pulmonary disease and stroke. *Am J Geriatr Psychiatry*. 2015;23(5):477–487.

Banerjee S, Hellier J, Dewey M, et al. Sertraline or mirtazapine for depression in dementia (HTA-SADD): a randomised, multicentre, double-blind, placebo-controlled trial. *Lancet*. 2011;378:403–411.

Dombrovski AY, Lenze EJ, Dew MA, et al. Maintenance treatment for old-age depression preserves health-related quality of life: a randomized, controlled trial of paroxetine and interpersonal psychotherapy. *J Am Geriatr Soc*. 2007;55(9):1325–1332.

Drysdale AT, Grosenick L, Downar J, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med*. 2016 Jan;23(1):28–38.

Eller T, Vasar V, Shlik J, Maron E. Pro-inflammatory cytokines and treatment response to escitalopram in major depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry*. 2008;32(2):445–450.

Flint J, Kendler KS. The genetics of major depression. *Neuron*. 2014;81(3):484–503.

Forester BP, Ajilore O, Spino C, Lehmann SW. Clinical characteristics of patients with late life bipolar disorder in the community: data from the NNDC Registry. *Am J Geriatr Psychiatry*. 2015;23(9):977–984.

Forester BP, Gatchel JR. Medical co-morbidity, brain disease, and the future of geriatric psychiatry. *Am J Geriatr Psychiatry*. 2014;22(11):1061–1065.

Gildengers AG, Mulsant BH, Al Jurdi RK, et al; GERI-BD Study Group. The relationship of bipolar disorder lifetime duration and vascular burden to cognition in older adults. *Bipolar Disord*. 2010;12:851–858.

Goes FS. Genetics of bipolar disorder: recent update and future directions. *Psychiatr Clin North Am*. 2016;39(1):139–155.

Herrmann LL, Le Masurier M, Ebmeier KP. White matter hyperintensities in late life depression: a systematic review. *J Neurol Neurosurg Psychiatry*. 2008;79(6):619–624.

Kiosses DN, Ravdin LD, Gross JJ, Raue P, Kotbi N, Alexopoulos GS. Problem adaptation therapy for older adults with major depression and cognitive impairment: a randomized clinical trial. *JAMA Psychiatry*. 2015;72(1):22–30.

Lewandowski KE, Sperry SH, Malloy MC, Forester BP. Age as a predictor of cognitive decline in bipolar disorder. *Am J Geriatr Psychiatry*. 2014;22:1462–1468.

McDonald WM. Neuromodulation treatments for geriatric mood and cognitive disorders. *Am J Geriatr Psychiatry*. 2016;24(12):1130–1141.

Meyers BS, Flint AJ, Rothschild AJ, et al; STOP-PD Group. A double-blind randomized controlled trial of olanzapine plus sertraline vs olanzapine plus placebo for psychotic depression: the study of pharmacotherapy of psychotic depression (STOP-PD). *Arch Gen Psychiatry*. 2009;66:838–847.

Milaneschi Y, Corsi AM, Penninx BW, Bandinelli S, Guralnik JM, Ferrucci L. Interleukin-1 receptor antagonist and incident depressive symptoms over 6 years in older persons: the InCHIANTI study. *Biol Psychiatry*. 2009;65(11):973–978.

Morimoto SS, Gunning FM, Wexler BE, et al. Executive dysfunction predicts treatment response

to neuroplasticity-based computerized cognitive remediation (nCCR-GD) in elderly patients with major depression. *Am J Geriatr Psychiatry*. 2016;24(10):816–820.

Müller N, Schwarz MJ, Dehning S, et al. The cyclooxygenase-2 inhibitor celecoxib has therapeutic effects in major depression: results of a double-blind, randomized, placebo controlled, add-on pilot study to reboxetine. *Mol Psychiatry*. 2006;11(7):680–684.

Raison CL, Rutherford RE, Woolwine BJ, et al. A randomized controlled trial of the tumor necrosis factor antagonist infliximab for treatment-resistant depression: the role of baseline inflammatory biomarkers. *JAMA Psychiatry*. 2013;70(1):31–41.

Smoller JW, Allison M, Cochrane BB, et al. Antidepressant use and risk of incident cardiovascular morbidity and mortality among postmenopausal women in the Women's Health Initiative study. *Arch Intern Med*. 2009;169(22):2128–2139.

Taylor WD, Aizenstein HJ, Alexopoulos GS. The vascular depression hypothesis: mechanisms linking vascular disease with depression. *Mol Psychiatry*. 2013;18(9):963–974.

Towfighi A, Ovbiagele B, El Husseini N, et al. Stroke 2016 epub ahead of print.

57.3f Alzheimer Disease and Other Neurocognitive Disorders

GARY W. SMALL, M.D., AND STEPHEN T. CHEN, M.D.

INTRODUCTION

In part because of medical technologic advances, people in industrialized societies are living longer, and the number of older persons is growing rapidly. Because age is a major risk factor for developing Alzheimer disease and other neurocognitive disorders (NCDs), the number of persons suffering from these illnesses has increased as well. In most cases, the cause of Alzheimer disease is not known. Rarely, one of several genetic mutations having an autosomal dominant inheritance pattern causes the disease with symptoms emerging in mid-life. Scientists are actively searching for more effective treatments and possible cures, and improved knowledge of potential risk and protective factors may contribute to preventive strategies. Along with such approaches, accurate early diagnosis, symptomatic treatments, and family interventions are essential to minimize the suffering of patients and family members and to maximize the quality of life for patients.

HISTORY

In 1906, the German psychiatrist and neurologist Alois Alzheimer first described a middle-aged patient who had suffered from a progressive dementia that affected language, memory, and behavior. After the patient's death at age 55 years, Alzheimer applied new staining techniques to the patient's brain tissue and demonstrated the presence of what are now termed neurofibrillary tangles and neuritic plaques in the neocortex and other brain regions. For years, Alzheimer disease was considered a presenile dementia, partially because some plaques and tangles occurred in

elderly persons without dementia and some elderly persons with dementia had few plaques and tangles. Such conflicts were resolved, however, in the late 1960s, when the degree of dementia was shown to correlate with the number of neuritic plaques in neocortical association areas. Moreover, causes of senile dementia other than Alzheimer disease were also recognized.

Over the last few decades, the recognition that cognitive decline may be pathologic rather than normal in aging has helped to drive scientific advances but also served to stigmatize the condition. Many older people have been reluctant to seek help for memory symptoms because of fear that they may be suffering from Alzheimer disease. Thus, symptomatic treatments are often begun after patients are already suffering from moderate degrees of dementia. However, if treatments that delay the onset of Alzheimer disease can be developed, there may be less emphasis on treatment of the disease state and greater attention to prevention and deceleration strategies, reducing the stigma and increasing access to care.

COMPARATIVE NOSOLOGY

A definitive diagnosis of Alzheimer disease can be made only by histopathologic examination of brain tissue, generally postmortem. Therefore, most cases are diagnosed based on clinical information and, in some settings, additional diagnostic studies.

Diagnostic criteria for dementia and Alzheimer disease were recently updated in two important references. In the fifth edition of the *Diagnostic and Statistical Manual* (DSM-5) published in 2013, the syndrome of dementia from the revised fourth edition (DSM-IV-TR) is subsumed under the new category "Major Neurocognitive Disorders." A comparison of their respective diagnostic criteria is shown in Table 57.3f–1. The essential features are similar, with some key differences. While dementia in DSM-IV-TR requires memory impairment plus one other type of cognitive disturbance that are gradual in onset and progressive, a major NCD requires a significant decline in at least one cognitive domain based on subjective concern and clinical assessment. For both diagnoses, specific etiologies, the most common being Alzheimer disease, may be added as specifiers; DSM-5 provides additional diagnostic criteria for these specific NCD subtypes.

In 2011, a work group formed by the National Institute on Aging and the Alzheimer Association (NIA-AA) published diagnostic guidelines for "probable" and "possible" Alzheimer disease. Criteria for both are similar to those for a major NCD, but require cognitive impairment based on history and objective cognitive assessment and a minimum of two cognitive or behavioral domains affected. Probable Alzheimer disease further specifies that cognitive impairment is insidious in onset, progressive, and prominent in memory, language, visuospatial, or executive functioning. Possible

Alzheimer disease should be considered when the course is atypical or cannot be reliably documented and there is evidence of other etiologies, for example, cerebrovascular disease or Lewy body dementia.

Both DSM-5 and the NIA-AA group recognize milder levels of cognitive impairment or a predementia stage of Alzheimer disease. In DSM-5, "Mild Neurocognitive Disorders," which feature "modest" cognitive decline and no functional impairment, was added as a new diagnostic category to encourage earlier diagnosis and intervention, which may more effectively modify the course of illness than that in the later stages. The NIA-AA guidelines refer to the predementia stage as "mild cognitive impairment (MCI)." Core clinical criteria for MCI include concern for cognitive changes, impairment in one or more cognitive domains, preferably measured by serial objective testing, preserved independence in daily activities, and no evidence of dementia and other etiologies of cognitive impairment. While either minor NCD or MCI can be diagnosed with only clinical data, the DSM-5 and the NIA-AA guidelines both refer to research biomarkers for Alzheimer disease that may aid diagnosis, particularly in the earlier stages of disease. Such biomarkers include structural neuroimaging with magnetic resonance imaging (MRI), molecular neuroimaging with positron emission tomography (PET), cerebrospinal fluid analysis, and genetic testing.

EPIDEMIOLOGY

NCDs are the most common psychiatric disorders in older persons and a growing public health problem in the United States. Alzheimer disease accounts for about 60 percent of late-onset NCDs. It is estimated to afflict from 5 to 10 percent of people of age 65 years and accounts for the most striking rise in dementia incidence in the very old. Alzheimer disease and other NCDs cost society each year over $400 billion in direct (i.e., actual dollar expenditures) and indirect costs (i.e., resource losses not involving dollar expenditures). The greatest risk factor for developing Alzheimer disease is age: its incidence and prevalence double every 5 years after age 60 years. Over 5 million Americans suffer from Alzheimer disease. As the baby boom generation (born between 1946 and 1964) ages, this number is expected to triple by 2050, to as many as 16 million. The prevalence of MCI may be as high as 19 percent in people older than age 65 years and 29 percent in those older than age 85 years.

Table 57.3f–1.
Comparison of Diagnostic Criteria for Major Neurocognitive Disorders in Fourth-Text Revision and Fifth Editions of the Diagnostic and Statistical Manual (DSM-IV-TR vs. DSM-5)

	DSM-5 Major Neurocognitive

	DSM-IV-TR Dementia	Disorders
Core Criteria	Multiple cognitive deficits: Memory plus one or more of the following: - Aphasia (language) - Apraxia (motor tasks) - Agnosia (identifying objects) Gradual onset and continuing cognitive decline	Decline in one or more of the following cognitive domains: - complex attention memory/learning - language - perceptual-motor - executive functioning - social cognition
Functional Criteria	Cognitive deficits cause significant impairment in social or occupational functioning and represent significant decline	Cognitive deficits interfere with independence in everyday activities
Exclusion Criteria	Core criteria do not occur exclusively during delirium Not better accounted for by another axis I disorder	Core criteria do not occur exclusively during delirium Not better accounted for by another mental disorder

ETIOLOGY

The search for possible causes for Alzheimer disease remains inconclusive. Scientists have made significant advances, approaching this search from various angles, which include genetics, neurochemistry, neuropathology, and neuroimaging studies. Hopefully, this search will eventually lead to treatments that may delay or even halt the disease process.

Alzheimer disease was first characterized and remains pathogenetically defined by its histopathologic features, neuritic plaques (extracellular deposits of amyloidogenic proteins), and neurofibrillary tangles (abnormal intracellular cytoskeletal filaments). Reduction in neuron number and synapses and an accumulation of plaques and tangles occur in the frontal, temporal, and parietal lobes; that finding is consistent with the involvement of association and limbic brain structures.

Extensive study of the formation of plaques and tangles has led to two prevailing hypotheses to explain the mechanisms that lead to Alzheimer disease: the cholinergic and the amyloid cascade hypotheses. The cholinergic hypothesis says that a dysfunctional cholinergic system causes the symptoms of Alzheimer disease. Clinical evidence suggests that anticholinergic medications impair cognitive function in older adults, and biochemical and neuropathologic evidence confirm a presynaptic cholinergic deficit in brains with Alzheimer disease. The amyloid hypothesis states that Alzheimer disease is caused by the excess accumulation and deposition of β-amyloid to form neuritic plaques. Additionally, the accumulation of β-amyloid is believed to result in neuronal dystrophy and neurofibrillary tangles and induce cholinergic dysfunction, oxidative stress, inflammation, and vascular damage.

Discoveries in the area of genetics and molecular biology have been critical to our understanding of the etiology of Alzheimer disease and

shown relationships between genes and pathogenesis. The increased frequency of Down syndrome (trisomy 21) patients who invariably develop Alzheimer disease neuropathology and dementia by their 40s led to finding significant linkage to chromosome 21 in some early-onset families. Excitement escalated when the gene coding for the amyloid precursor protein (APP) found in senile plaques was localized to the same chromosome 21 region. APP mutations, however, are relatively rare. Most early-onset pedigrees not segregating APP mutations show mutations of a chromosome 14 gene (presenilin 1). Another form of early-onset Alzheimer disease is represented by families of Volga German origin, in whom chromosome 1 mutations (presenilin 2) have been identified.

Studies of the common late-onset disease (symptoms beginning after age 60 years) showed evidence for linkage or association or both for a chromosome 19 region. Apolipoprotein E (APOE) was a candidate Alzheimer disease susceptibility gene partly because APOE and cerebrospinal fluid bound to immobilized β-amyloid peptide and antisera to APOE stain senile plaques and neurofibrillary tangles. In addition, APOE localized to the same region of chromosome 19 identified in linkage studies.

APOE has three allelic variants (2, 3, and 4). Everyone inherits one allele from each parent, so that five common genotypes are possible (2/3, 3/3, 2/4, 3/4, and 4/4). The sixth genotype (2/2) is rare. In the general population, approximately 3 percent of persons have the 4/4 genotype, 20 percent have the 3/4 genotype, and most people have the 3/3 genotype. The APOE-4 allele increases risk and decreases age at dementia onset in a dose-related fashion, that is, Alzheimer disease risk is lowest for the 3/3 genotype, higher for the 3/4 genotype, and highest for the 4/4 genotype. By contrast, the APOE-2 allele has a protective effect, so that Alzheimer disease risk is lower for people with the 2/3 genotype than for those with the 3/3 genotype. Susceptibility to Alzheimer disease from APOE affects many races and has been confirmed worldwide.

More recent research has focused on inflammatory pathways and mediators that contribute to the neurodegenerative process. Pathologic β-amyloid can activate phagocytic microglia and astrocytes to secrete cytokines, chemokines, and other proinflammatory molecules, such as prostaglandins and free radicals, which can have deleterious effects on the brain. A number of inflammatory processes are overrepresented in brains with Alzheimer disease. Microglia are highly activated and surround β-amyloid plaques. Interleukin-1 overexpression promotes the synthesis and processing of APP and thus of amyloid production and plaque deposition, and also induces the activity and mRNA expression of acetylcholinesterase, which reduces cholinergic activity. Chemokines and their receptors are upregulated and may contribute to plaque-associated inflammation and neurodegeneration. Complement activation fragments, reactive astrocytes, and activated microglia are all highly colocalized with plaques containing

aggregated β-amyloid. Chronic inflammation is a well-known contributor to vascular disease, and may be the common pathway to the prevalent coexistence of vascular and Alzheimer disease pathology. Another common link may be APOE-4, a risk factor for developing both atherosclerosis and late-onset Alzheimer disease. These relationships suggest that neurovascular damage is a primary occurrence and that subsequent injuries, including Aβ deposition, amplify and/or exacerbate vascular damage that then leads to neurodegenerative processes and ultimately cognitive decline.

Our current database suggests that Alzheimer disease may be a heterogeneous condition, and therefore have multiple causative factors. Investigations comparing familial (one or more affected first-degree relatives of proband or large autosomal dominant kindreds) and sporadic Alzheimer disease patients, as well patients with early-onset (65 years or younger) and late-onset dementia, have found clinical and biologic differences among them. Other studies suggest differences in disease expression between men and women, cerebral glucose metabolism, and neuroendocrine function, and a possible interaction among age at onset, sex, and family history of dementia. Case-control studies have identified prior head trauma and depression as risk factors for Alzheimer disease, and higher educational achievement, larger brain size, and cardiovascular conditioning as protective factors.

In a sense, finding the etiology of Alzheimer disease is like solving a huge jigsaw puzzle: Although researchers are beginning to piece together a few corners, the puzzle is far from solved. In some cases, the entire cause may lie in a single genetic defect; in others, a mixture of genetic and environmental factors may play a role. The likelihood of multiple causative factors suggests that there are multiple approaches for developing effective prevention and treatment strategies.

DIAGNOSIS AND CLINICAL FEATURES

NCDs in late life usually begin with milder forms of cognitive decline and progress gradually. Early on, pathologic cognitive decline may be difficult to distinguish from normal aging, in which there is commonly an awareness of mild memory changes, but with minimal objective cognitive deficits. However, as neurodegeneration progresses, MCI, often an intermediate state between normal aging and NCD, develops and place patients at increased risk for developing an NCD. Patients with primarily memory deficits may have the amnestic subtype of MCI, which is more likely than other MCI subtypes to progress to dementia due to Alzheimer disease.

The clinical hallmark of Alzheimer disease is memory impairment. Short-term memory is affected initially. Patients have particular difficulty in learning new information, as shown by their inability to recall three

objects after 5 minutes or even after just two distractions. Long-term memory may be affected, and patients may not recall important past personal information. Other cognitive areas may be impaired, including abstract thinking, judgment, personality, and language.

In a typical case of Alzheimer disease, the onset is so insidious that family members have difficulty recognizing when the impairment began. Diagnosis is often missed or delayed because of the insidious onset and gradual progression of symptoms, but also due to lack of attention to cognitive issues during medical screenings. Prevalence studies of community samples have documented high proportions of undiagnosed cases. A complete history of the patient's cognitive and everyday functioning from someone who knows the patient well, physical and neurologic examinations, and a mental status examination are essential in the clinical evaluation. The history should include screening questions for behavioral changes, including apathy, paranoia, agitation, insomnia, anxiety, and depression. On examination, the patient may show *apraxia* (inability to carry out motor commands, even though comprehension and motor function are intact), *agnosia* (inability to recognize objects, despite intact sensory function), or difficulties in visuospatial skills (inability to copy two-dimensional and three-dimensional figures or to assemble blocks). Brief standardized mental status tests (e.g., the Mini-Mental State Examination, the Mini-Cog) are useful screening tools, and more extensive neuropsychologic batteries provide details on the nature and degree of the cognitive deficits.

PATHOLOGY AND LABORATORY EXAMINATION

Because of neuronal loss, the brains of Alzheimer disease patients demonstrate atrophy greater than expected for their ages. In clinical practice, radiologists interpret structural scanning results based on visual readings, which are less sensitive in detecting subtle atrophy. In research settings, image analysis programs are used to quantify regional MRI and CT volumes. These analyses have found that medial temporal or hippocampal atrophy measures can distinguish patients with a clinical diagnosis of Alzheimer disease from controls. Atrophy in the entorhinal cortex and hippocampus is associated with increased risk for progressive memory decline and development of Alzheimer disease.

PET scans measuring regional glucose metabolism after injection of fluorodeoxyglucose (FDG) can differentiate patients with Alzheimer disease from cognitively intact individuals and those with MCI and other NCDs. Patients with Alzheimer disease show hypometabolism in posterior cingulate, parietal, temporal, and frontal regions, and these patterns can be observed in patients with MCI as well (Fig. 57.3f–1). Combining genetic risk assessment with PET is a potential strategy for the early detection of Alzheimer disease. Imaging studies using FDG-PET have shown brain

metabolic deficits in nondemented middle-aged and older persons with the APOE-4 allele. In clinical practice, FDG-PET scans have been approved for reimbursement by the U.S. Centers for Medicare and Medicaid Services to assist with the differential diagnosis of Alzheimer disease and frontotemporal dementia, making this technique more accessible to patients and their physicians.

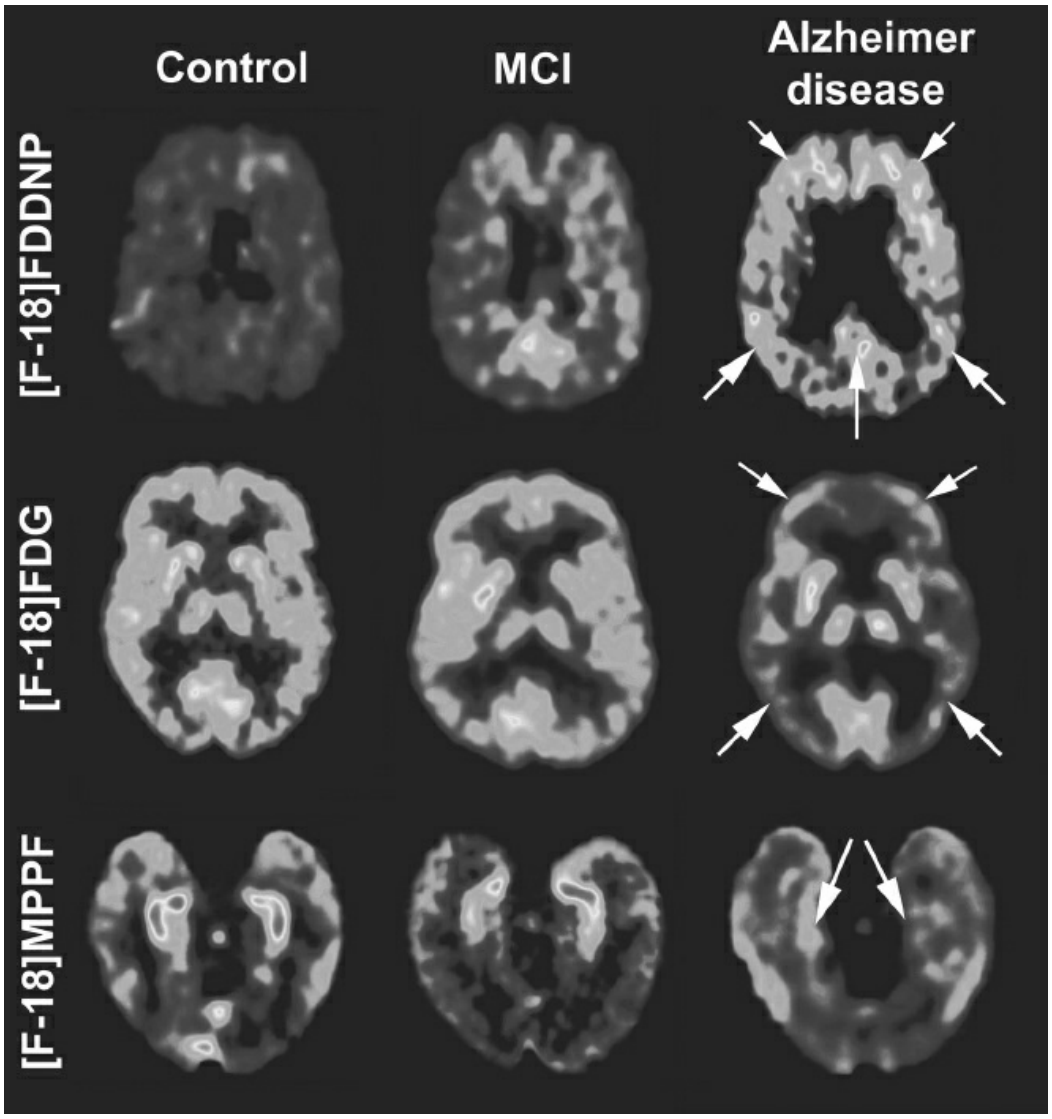

FIGURE 57.3f–1. Representative examples of brain positron emission tomographic (PET) images from a control subject (**left**), a patient with mild cognitive impairment (**center**), and a patient with Alzheimer disease (**right**), including the use of the agents FDDNP (**top**), FDG (**middle**), and MPPF (**bottom**). For each imaging method, the patient with mild cognitive impairment shows intermediate values between those of the control and the patient with Alzheimer disease. Absolute scales are comparable within subjects with a specific probe, and warmer colors indicate higher values. In the patient with Alzheimer disease, the *arrows* point to frontal, parietal and posterior cingulate **(upper)**, frontal and temporal-parietal (**middle**), and medial temporal (**lower**) regions. (From Kepe V, Barrio JR, Huang SC, et al. Serotonin 1A receptors in the living brain of Alzheimer disease patients. *Proc Natl Acad Sci USA*. 2006;103(3):702–707.)

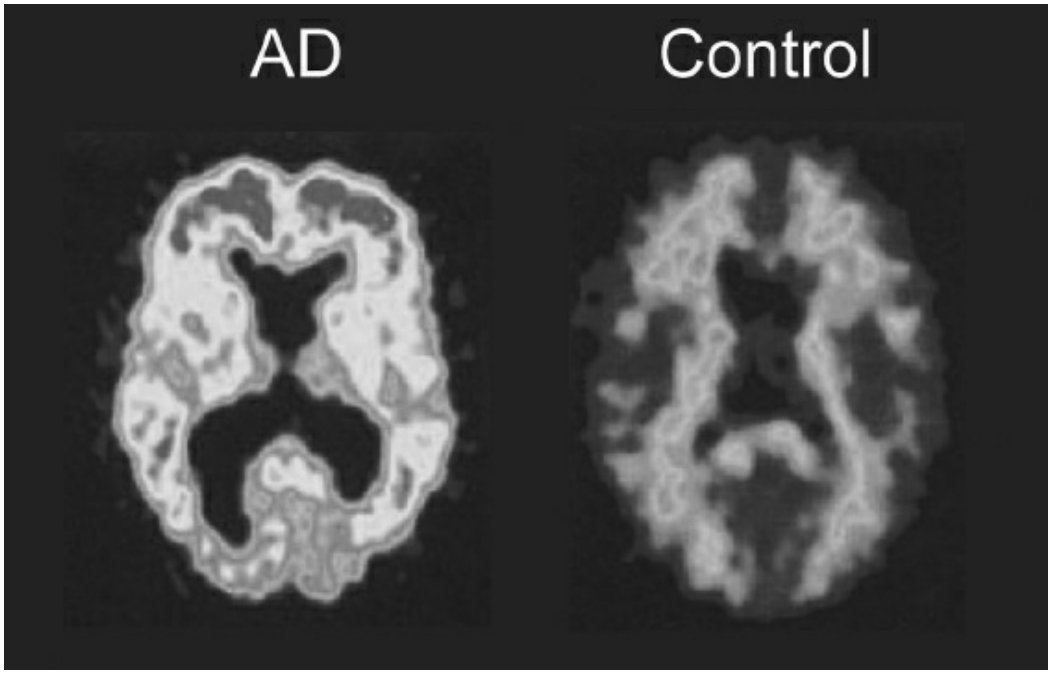

FIGURE 57.3f–2. PIB-standardized uptake images showing higher Pittsburgh Compound-B (PIB) retention in a 79-year-old patient with Alzheimer disease than in a 67-year-old healthy control. (Reproduced with permission from Klunk WE, Engler H, Nordberg A. Imaging brain amyloid in Alzheimer disease with Pittsburgh Compound-B. *Annals Neurol.* 2004; 55(3):306–319.)

Several small-molecule probes for use with PET imaging have been found to provide in vivo measures of amyloid plaques, and one compound (FDDNP) can measure both amyloid plaques and tau tangles. PET studies using Pittsburgh Compound-B (PIB), a derivative of thioflavin-T amyloid dye, show significantly greater cortical retention in patients with Alzheimer disease compared with controls (Fig. 57.3f–2). PET scans after FDDNP injection have been found to differentiate patients with Alzheimer disease from those with MCI and cognitively intact controls and can predict future cognitive decline (Fig. 57.3f–1). FDDNP-PET scans also can be displayed as three-dimensional cortical surface projections, which show signal patterns remarkably similar to those expected from autopsy studies demonstrating regional brain accumulation patterns of plaques and tangles (Fig. 57.3f–3). Another PET ligand, MPPF, provides visualizations of hippocampal pyramidal neurons, which also demonstrates different signal levels according to the patient's degree of cognitive deficit (Fig. 57.3f–1). In recent years, new PET ligands are being developed to image amyloid plaques or tau tangles. One goal of this approach is to use these imaging tools to accelerate research in disease-modifying drug development for treating Alzheimer.

Single-photon emission computed tomography (SPECT) scanning provides cerebral blood flow patterns of parietal and temporal hypoperfusion, which can be helpful in diagnosing Alzheimer disease. Comparison studies suggest that SPECT has lower diagnostic accuracy than

PET, but advances in voxel-based statistical analysis have improved the diagnostic accuracy of SPECT. Although SPECT has been used more widely because of its availability and low cost, the greater use of PET for oncology studies and reimbursement from Medicare has made it more widely available and accessible.

DIFFERENTIAL DIAGNOSIS

The differential diagnosis of Alzheimer disease can be a challenge. Except for measures of regional cerebral glucose metabolism using PET scanning, most potential biologic markers are still in a development phase, and the clinician is left with conventional methods of evaluation: the patient's history, a mental status and physical examination, structural imaging, and laboratory tests. The history provides important information, and several informants should be sought for corroboration. Even with the most thorough evaluation, diagnostic errors occur. When patients are followed to autopsy for neuropathologic confirmation, accuracy of prior clinical diagnosis can be as low as 60 percent.

Major vascular NCD is estimated to account for 10 to 20 percent of major NCDs in late life. Clinically, the disorder is characterized by a sudden onset of illness and stepwise decline in cognitive function rather than the insidious onset and gradually progressive course of Alzheimer disease. The frequent coexistence of vascular disease and Alzheimer pathology (estimated to occur in about one-third of all major NCDs) impairs diagnostic accuracy.

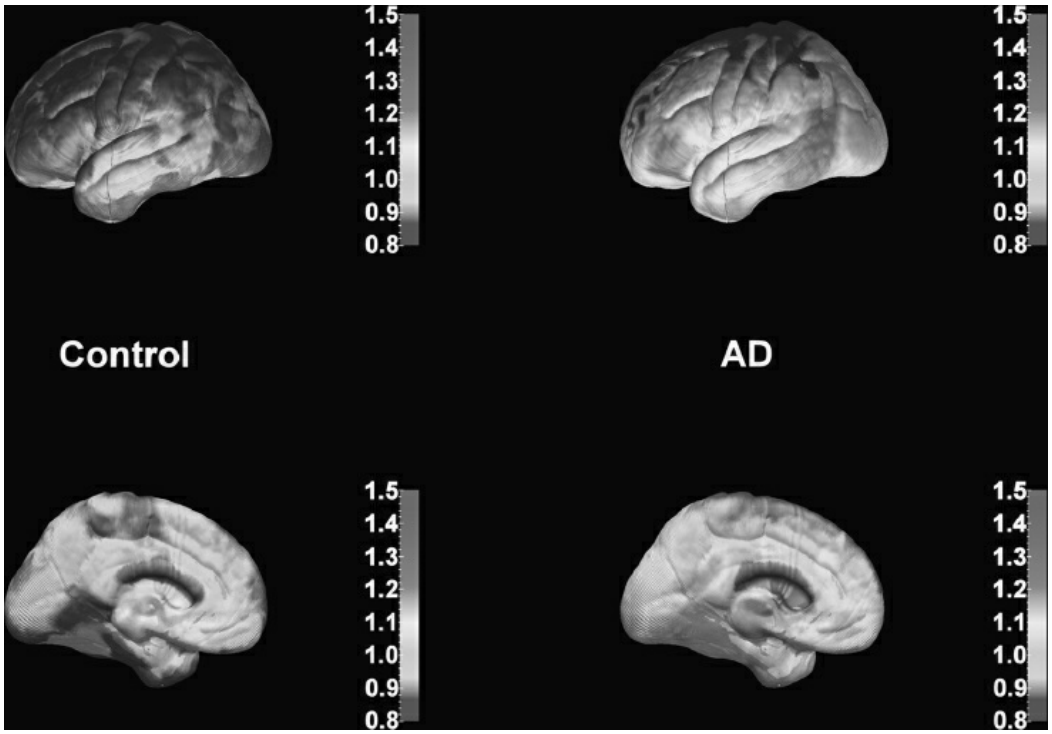

FIGURE 57.3f–3. Three-dimensional cortical surface projection images of FDDNP-positron emission tomography (PET) scans from a control subject and a patient with

Alzheimer disease showing lateral (**upper**) and medial (**lower**) brain surfaces. Warmer colors indicate higher levels of plaques and tangles. (Reproduced with permission from Small GW, Bookheimer SY, Thompson PM, Cole GM, Huang S-C, Kepe V, et al. Current and future uses of neuroimaging for cognitively impaired patients. *Lancet Neurol.* 2008;7(2):161–172.)

Even though Alzheimer disease and vascular disease account for the majority of cases of major NCDs, other conditions must be considered and sometimes can be differentiated from Alzheimer disease by their clinical presentation. For example, frontotemporal lobar degeneration is characterized by marked personality and behavioral changes, executive dysfunction, and relative preservation of other cognitive domains. By contrast, patients with Lewy body disease may evidence visual hallucinations, delusions, fluctuating mental status, and sensitivity to antipsychotic drugs.

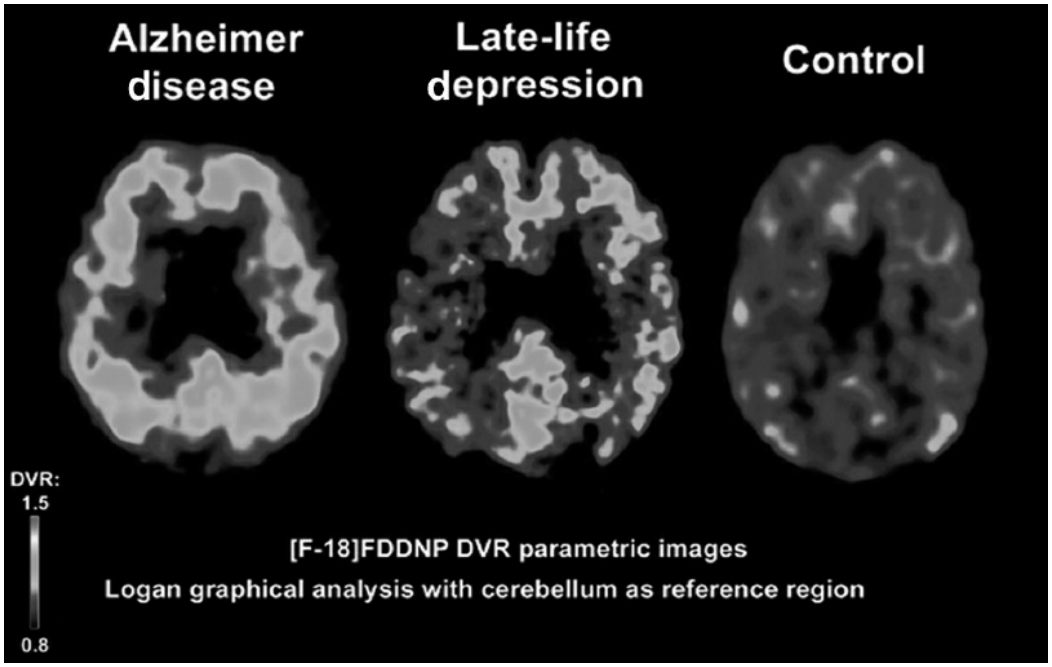

FIGURE 57.3f–4. 2-(1-{6-[(2-[18F]fluoroethyl) (methyl)-amino]-2-naphthyl}ethylidene) malononitrile ([18F]FDDNP) binding in a patient with Alzheimer disease, late-life depression, and a healthy control individual. [18F]FDDNP relative distribution volume (DVR) parametric images with Logan graphic analysis with the cerebellum as the reference region are shown.

Alcohol and drugs may cause memory loss and other cognitive dysfunction. For older patients, drugs deserve special emphasis as a cause of memory impairment because of widespread polypharmacy and age-related changes in body function and pharmacokinetics. An older person's brain can be extremely sensitive to change in its environment, and almost any physical illness can effect such a change and produce a dementia-like syndrome. Numerous potentially treatable physical and mental conditions must be considered whenever the diagnosis of dementia is entertained

(Table 57.3f–1).

The clinical presentation of late-life depression may mimic that of an NCD. The earlier term "depressive pseudodementia" is no longer considered useful because the cognitive impairment associated with depression is real. The reported frequency of mistaking depressed patients for patients with an NCD ranges from 5 to 15 percent. Psychomotor retardation, impaired concentration, and apathy characteristic of the depressed patient may be mistaken for symptoms of an NCD, especially when these symptoms are accompanied by memory loss. In the dementia syndrome of depression, however, the complaints of memory tend to exceed the actual memory impairment observed on neuropsychologic testing. The fact that NCDs and depression often coexist further clouds the diagnostic picture. About 30 percent of patients with a progressive NCD also suffer from some form of depression.

Depression is a risk factor and a prodrome of dementia, and common neurobiologic mechanisms may influence both conditions. Although depression occurring in the course of dementia may reflect a psychologic reaction to cognitive losses, PET scanning of patients with late-life depression show increased signals indicating higher levels of amyloid plaques and tau tangles in patterns that overlap with those observed in Alzheimer dementia (Fig. 57.3f–4).

Many people in their 50s and older complain of forgetfulness and worry about its meaning. *Age-associated memory impairment,* as defined by a workgroup convened by the National Institute of Mental Health (NIMH), applies to healthy persons 50 years and older who have subjective complaints of memory loss and mild deficiencies in memory performance compared to young adults. About 40 percent of persons in their 50s and up to 85 percent of those age 80 years and older suffer from age-associated memory impairment. Preliminary data suggest that the condition is stable over intervals of up to 4 years and thus is presumed to reflect normal aging and does not usually progress to dementia. Although it may take them longer to remember, most elderly people continue to recall information they have learned well and important events from the past. Because age-associated memory impairment is so common, it is also referred to as normal aging.

COURSE AND PROGNOSIS

Alzheimer disease is a progressive disease that eventually leads to death, not directly from the dementing process, but usually from intercurrent illnesses. The rate of progression varies, and it is difficult to estimate how rapid a decline a patient may have. Some patients become extremely demented within 1 year; others experience plateaus for several years. The majority experience behavioral disturbances such as agitation and psychosis at some point during the course of illness. Other causes of NCDs,

including frontotemporal lobar degeneration, vascular disease, and Lewy body disease, are gradually progressive as well. Occasionally, treatable causes of dementia (e.g., drug toxicity, hypothyroidism) can be cured when the underlying cause is treated.

TREATMENT

The first step in treating NCDs is attempting to identify any underlying illnesses that may be causing memory impairment and other symptoms of the syndrome. If none is found or if the symptoms persist after appropriate specific treatment, then symptomatic treatments may be tried.

Table 57.3f–2.
Some Potentially Reversible Conditions That Might Resemble Alzheimer Disease and Other Neurocognitive Disorders

Substances
Anticholinergic agents
Antihypertensives
Antipsychotics
Corticosteroids
Digitalis
Opiates
Nonsteroidal anti-inflammatory agents
Phenytoin
Polypharmacy
Benzodiazepines
Psychiatric disorders
Anxiety
Depression
Mania
Delusional (paranoid) disorders
Metabolic and endocrine disorders
Addison disease
Cushing syndrome
Hepatic failure
Hypercarbia (chronic obstructive pulmonary disease)
Hypernatremia
Hyperparathyroidism
Hyperthyroidism
Hypoglycemia
Hyponatremia
Hypothyroidism
Renal failure
Volume depletion
Miscellaneous conditions
Fecal impaction
Hospitalization

Psychopharmacotherapy

A variety of pharmacologic treatments are available for Alzheimer disease (Tables 57.3f–2 and 57.3f–3). Many of these approaches diminish symptoms and increase quality of life, but none of them can halt the dementing process. Most pharmacologic agents now available or in development target a specific symptom (e.g., agitation or memory loss) and are derived from the known neurobiology of the disease (e.g., a specific neurotransmitter deficit) or hypothesized antidementia approaches (e.g., antiamyloid, antitau, anti-inflammation, or antioxidation).

One approach to improving memory function has been to enhance cholinergic activity in the brain. Autopsies of patients with Alzheimer disease have found reduced levels of choline acetyltransferase in the brain, a finding consistent with a central cholinergic deficit. Trials of cholinergic agonists have temporarily improved memory in young and old normal volunteers. Trials of several agents, including physostigmine (Antilirium), choline (Mega-B), and lecithin (PhosChol), have yielded contradictory results (some effects vs. no measurable effects on memory in patients with dementia) or else have not yet been replicated.

Table 57.3f–3.
Medications Used or Under Investigation for the Treatment of Alzheimer Disease

Cholinesterase inhibitors
Tacrine[a]
Donepezil[a]
Rivastigmine[a]
Galantamine[a]
Sustained-release physostigmine
Eptastigmine
Cholinergic agonists
Xanomeline
Milameline
Antioxidants
Vitamin E (α-tocopherol)
Hormones
Conjugated estrogens
***N*-Methyl** D-**aspartate antagonist**
Memantine[a]
Other potential cognitive enhancers
Acetyl-L-carnitine
Curcumin
Ginkgo biloba
Nonsteroidal anti-inflammatory drugs

Omega-3 fatty acids

[a]Approved by the U.S. Food and Drug Administration.

Table 57.3f–4.
Some Drugs Used or Under Investigation for the Treatment of Behaviors Associated with Alzheimer Disease and Other Neurocognitive Disorders

Antipsychotics
Aripiprazole
Haloperidol
Olanzapine
Quetiapine
Risperidone
Anticonvulsants
Carbamazepine
Divalproex sodium
Antidepressants
Citalopram
Sertraline
Paroxetine
Sertraline
Trazodone
Anxiolytics
Buspirone
Lorazepam
Clonazepam
Antiparkinsonian Drugs
Selegiline
β-Adrenergic receptor antagonists
Propranolol
Pindolol

The U.S. Food and Drug Administration (FDA) has approved four cholinesterase inhibitors (tacrine [Cognex], donepezil [Aricept], rivastigmine [Exelon], and galantamine [Razadyne]) and one *N*-methyl D-aspartate (NMDA)–receptor antagonist (memantine [Namenda]) for the treatment of Alzheimer disease. These drugs have proven benefits for memory, other aspects of cognition, behavioral disturbances, overall functioning, and even caregiver burden. Tacrine has the disadvantage of causing reversible elevations of serum transaminase levels and is therefore rarely used.

These symptomatic treatments were originally indicated only for mildly to moderately ill patients with Alzheimer disease. More recently, the efficacy of cholinesterase inhibitors has been demonstrated in severe Alzheimer disease, Lewy body disease, and dementia associated with Parkinson disease. Memantine is indicated for moderately to severely ill

patients with Alzheimer disease. Clinical practice often combines a cholinesterase inhibitor with memantine, supported by data that this combination provides beneficial response compared to cholinesterase inhibitors alone.

Clinical trials point to the importance of early detection and treatment with cholinesterase inhibitor drugs. In randomized, placebo-controlled inpatients with Alzheimer disease, these drugs result in less cognitive decline compared with placebo after 6 months of treatment. During the following 6-month open-label treatment periods, the patients who were originally treated with placebo were given active drug, and at 1 year, better cognitive performance was observed in patients who began drug treatment from the beginning of the trial compared with those who had been placebo-delayed for 6 months.

Studies using cholinesterase inhibitor drugs in patients with MCI have been inconclusive. For example, a randomized, placebo-controlled trial in patients with MCI found that donepezil delayed progression to Alzheimer disease compared with placebo after 1 year of treatment; however, no difference between drug and placebo was found after 3 years of treatment.

Studies have examined a variety of other agents that may improve cognitive functioning, including estrogen, nonsteroidal anti-inflammatory agents, and antioxidant supplements. Unfortunately, any clinical benefit for these various treatments has been inconclusive. The widespread use of over-the-counter preparations should be evaluated routinely, as some do carry risk. For example, while vitamin E (α-tocopherol) at 2,000 IU daily lowered rates of functional decline in patients with moderately severe Alzheimer disease in one study, meta-analyses have found an increased risk of myocardial infarction in patients taking daily vitamin E supplementation, particularly doses greater than 400 IU.

As scientists uncover the basic pathogenetic mechanisms of Alzheimer disease, several therapeutic strategies for preventing or diminishing insoluble amyloid accumulation are under investigation, including secretase inhibitors or modulators, active or passive vaccines, aggregate (oligomer) and tau kinase inhibitors, cholesterol-lowering statin drugs, anti-inflammatory drugs, omega-3 fatty acids, and the yellow curry pigment, curcumin.

Psychiatric medications may be helpful for treating behavioral disturbances associated with dementia, such as depression, agitation, psychosis, and anxiety. Patients who have dementia with concurrent depression may improve following treatment with antidepressant medications, though data are limited. Antidepressants with minimal anticholinergic effects (e.g., serotonin reuptake inhibitors) are preferred over tricyclics in terms of minimizing risk for worsening cognition. One meta-analysis found only five randomized controlled trials that studied the efficacy of treating patients with dementia and depression with

antidepressants. A total of 82 patients were treated with active drug (sertraline, imipramine, clomipramine, and fluoxetine) versus 83 with placebo. Response and remission rates for depression were significantly greater for the pooled antidepressant group than for that receiving placebo. The study found no significant differences in overall rates of withdrawal or those due to adverse events. Antidepressants may also have a role in treating agitation. In controlled trials, sertraline and citalopram have demonstrated efficacy over placebo in patients with dementia and agitation.

Antipsychotic drugs may be effective in treating psychotic symptoms and agitation. Multiple controlled trials and meta-analyses of controlled trials of antipsychotic treatment in dementia indicate that antipsychotics have a significantly greater effect than do placebo, but the effect sizes are mostly less than 20 percent. As there is no evidence that one agent is superior to others, side-effect profile is a primary consideration when choosing a specific agent.

First-generation, or "conventional" antipsychotics may be generally categorized as "high-potency" or "low-potency" antipsychotics, according to how potently they occupy dopamine-2 receptors. High-potency agents such as haloperidol are more likely to cause parkinsonian symptoms, whereas such low-potency drugs as chlorpromazine cause sedation, postural hypotension, and anticholinergic effects. Less frequently, tardive dyskinesia and neuroleptic malignant syndrome may develop with either. Clozapine, while considered an "atypical" antipsychotic, can have similar side effects as low-potency antipsychotics, and also requires routine blood-count monitoring because of the risk of agranulocytosis. Because of the lower associated risk for neuromotor side effects, the newer atypical antipsychotic drugs, such as quetiapine, risperidone, olanzapine, and aripiprazole, have received greater attention in recent years. However, these atypical agents are associated with other adverse effects, including metabolic syndrome, vascular events, and cardiac arrhythmias, which warrant vigilant monitoring. Overall, conventional and atypical antipsychotic medications have been associated with increased mortality rates in elderly patients with dementia, which prompted black box warnings by the FDA for all antipsychotics currently available. Investigators have more recently challenged the basis for these warnings with data showing no significant differences in mortality in this population between antipsychotics and placebo.

Benzodiazepines and anticonvulsants have been used and studied for the treatment of agitation. However, there is little evidence for their use in this setting, and they have undesirable adverse effects, including some particularly noxious to the elderly patient with dementia, such as sedation, confusion, memory impairment, dysarthria, and ataxia. Moreover, benzodiazepines have recently been associated with increased risk for

Alzheimer disease.

Other Interventions

Attention to the Environment. Attending to patients' environments is essential to their well-being, and can mitigate the need for pharmacotherapy for behavioral disturbances. Patients with cognitive losses are sensitive to their surroundings and do best with optimal stimulation. Understimulation may cause withdrawal; overstimulation may cause confusion and agitation. Familiar and constant surroundings and caregivers maximize the patient's cognitive functions and minimize behavioral disturbances. If moves cannot be avoided, placing familiar objects (e.g., photographs and furniture) in the new environment help create a more comfortable and homelike atmosphere. Daily routines—regularly scheduled awakening, bedtime, meals, and physical exercise—increase a patient's sense of security. Memory and orientation can be facilitated by prominent displays of clocks and calendars, daylight exposure, a night-light, checklists, and diaries. Medication schedules should be simplified as much as possible. Newspapers, radio, and television, appropriate to the patient's cognitive level, can be useful in maintaining a patient's contact with and awareness of the outside world. Daily simple yet meaningful tasks that the patient can complete provide a valuable sense of usefulness and accomplishment for the patient. Specific therapies such as music and activity therapies have also been shown to improve behavior.

Family Intervention. Addressing the needs of the patient with Alzheimer disease would not be complete without addressing the needs of the patient's care system. While the identified patient is the one suffering from the disease and in need of treatment, often the adult child initiates contact with the physician and has his or her own reasons for evaluation and treatment. Considerable skill and sensitivity are necessary to maintain an alliance with the adult child while respecting the patient's autonomy, dignity, and privacy. Psychotherapeutic intervention with family members is a critical aspect of treatment. Caregiver and family stress can result in suboptimal care for the patient, as they themselves might experience frustration, anxiety, fatigue, anger, resentment, guilt, and even depression. Interpersonal conflicts among the family members may arise and require resolution to help the patient. Children and siblings might become anxious about their own risk for developing Alzheimer disease. Education and counseling about the nature of the patient's illness help relatives to cope with their feelings. Relatives may need reassurance that their emotional reactions are common and that talking about them can bring relief. Many relatives also need help in grieving the loss of the person they once knew. The Alzheimer's Association, a national organization of family members

with local chapters throughout the United States, has been at the forefront of providing educational and emotional support for family members. Sometimes, a family member's needs may exceed this level of support and require individual attention, in the form of treatment for anxiety or depression, or a full physical and neuropsychologic evaluation, which may either allay anxiety about having the disease or occasionally uncover early signs of cognitive impairment, which may lead to more anxiety, but also the opportunity for early intervention.

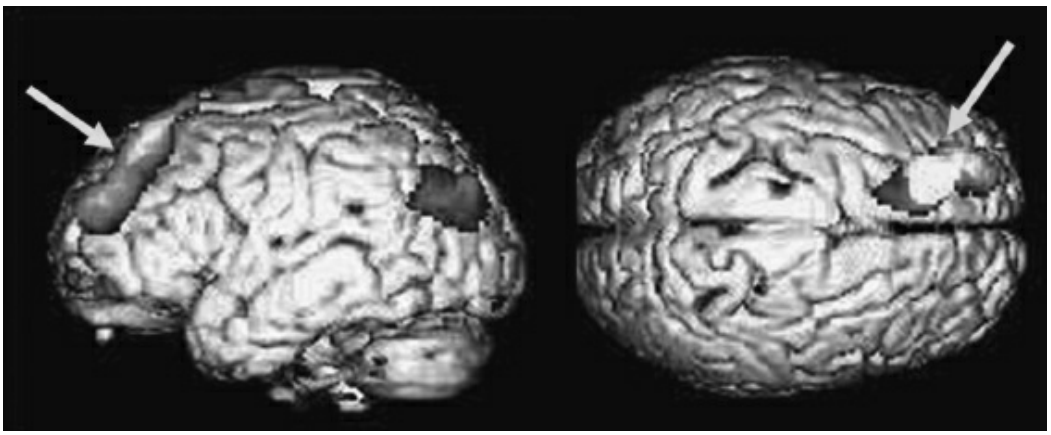

FIGURE 57.3f–5. Statistical parametric mapping of results of fluorodeoxyglucose (FDG)-positron emission tomography (PET) scanning shows a 5 percent decline in metabolic activity in the left dorsolateral prefrontal cortex after a 2-week healthy lifestyle program, including memory training, physical conditioning, and healthy diet, whereas no such decline was observed in the control group. The color scale highlights the location of all cortical voxels showing significantly greater decline ($p < 0.01$) in the intervention group compared with the control group. The left image is a left lateral viewpoint; the right image viewpoint is from the top of the brain. The *arrows* point to the voxels of peak significance ($Z = 3.82$, $p < 0.001$). (Reproduced with permission from Small GW, Silverman DHS, Siddarth P, Ercoli LM, Miller KJ. Effects of a 14-day healthy longevity lifestyle program on cognition and brain function. *Am J Geriatr Psychiatry.* 2006;14(6):538–545.)

Prevention Strategies. The MacArthur Foundation studies on aging have shown that early lifestyle choices, more than genes, determine our health and vitality as we age. Epidemiologic, survey, and animal data suggest that these lifestyle choices can specifically impact brain health. Reducing modifiable cardiovascular risk factors, depression, low education, and physical inactivity by just 10 to 25 percent could potentially prevent millions of cases of Alzheimer disease worldwide. Some of these same risk factors have been shown to be associated with subjective memory impairment, a possible precursor to MCI and Alzheimer disease, in cognitively intact persons of all ages. These same individuals reported better memory if they engaged in healthy lifestyle habits such as regular exercise, nutritious diet, and not smoking. One controlled study demonstrated that participating in a healthy lifestyle program combining mental and physical exercise, stress reduction, and healthy diet for just 2 weeks was associated with significant positive effects on cognitive

performance and brain metabolism (Fig. 57.3f–5). Whether lifestyle choices can prevent or delay the onset of Alzheimer disease and other NCDs remains to be determined. However, modifying one's lifestyle to incorporate some of these healthy behaviors come with minimal risk and offer other proven health benefits. Until more effective pharmacotherapies are available, prevention strategies through healthy lifestyle choices may be the safest and most beneficial way to reduce one's risk for developing Alzheimer disease and other NCDs.

A recent telephone survey conducted with Gallup Poll of a representative community sample of 18,614 US respondents, ages 18 to 99 years, addressed whether potentially modifiable risk factors for AD were associated with self-awareness of memory problems. As predicted, depression, less education, less exercise, and hypertension were significantly associated with self-awareness of memory problems, which increased with age and number of risk factors.

References

Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. *JAMA*. 2014;311(10):1052–1060.

Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011:270–279.

Barnes DE, Yaffe K. The projected effect of risk factor reduction on Alzheimer's disease prevalence. *Lancet Neurol*. 2011;10:819–828.

Bjelakovic G, Nikolova D, Gluud LL, Simonetti RG, Gluud C. Antioxidant supplements for prevention of mortality in healthy participants and patients with various diseases. *Cochrane Database Syst Rev*. 2012;3:CD007176.

Burns A, Iliffe S. Dementia. *Br Med J*. 2009;338:b75.

Chen ST, Siddarth P, Ercoli LM, Merrill DA, Torres-Gil F, Small GW. Modifiable risk factors for Alzheimer disease and subjective memory impairment across age groups. *PLoS ONE*. 2014;9(6):e98630.

Donix M, Small GW, Bookheimer SY. Family history and APOE-4 genetic risk in Alzheimer's disease. *Neuropsychology Review*. 2012;22(3):298–309.

Dubois B, Feldman HH, Jacova C, et al. Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurol*. 2010;9:1118–1127.

Kales HC, Kim HM, Zivin K, et al. Risk of mortality among individual antipsychotics in patients with dementia. *Am J Psychiatry*. 2012;169:71–79.

Kepe V, Bordelon Y, Boxer A, et al. PET imaging of neuropathology in tauopathies: progressive supranuclear palsy. *J Alzheim Dis*. 2013;36:145–153.

Kumar A, Kepe V, Barrio JR, et al. Protein binding in patients with late-life depression. *Arch Gen Psychiatry*. 2011;68:1143–1150.

McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's

disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7:263–269.

Merrill DA, Siddarth P, Kepe V, et al. Vascular risk and FDDNP-PET influence cognitive performance. *J Alzheimers Dis*. 2013;35:147–157.

Miller KJ, Siddarth P, Gaines JM, et al. The Memory Fitness Program: cognitive effects of a healthy aging intervention. *Am J Geriatr Psychiatry*. 2012;20:514–523.

Petersen RC, Caracciolo B, Brayne C, Gauthier S, Jelic V, Fratiglioni L. Mild cognitive impairment: a concept in evolution. *J Intern Med*. 2014;275:214–228.

Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimer's Dement*. 2013;9:63–75.e2.

Raji CA, Eyre H, Wei SH, et al. Hot topics in research: preventive neuroradiology in brain aging and cognitive decline. *AJNR Am J Neuroradiol*. 2015;36:1803–1809.

Small GW. Treating dementia and agitation. *JAMA*. 2014;311:677–678.

Small G, Bullock R. Defining optimal treatment with cholinesterase inhibitors in Alzheimer's disease. *Alzheimers Dement*. 2011;7:177–184.

Small GW, Siddarth P, Ercoli LM, Chen ST, Merrill DA, Torres-Gil F. Healthy behavior and memory self-reports in young, middle-aged, and older adults. *Int Psychogeriatr*. 2013;25:981–989.

Small GW, Siddarth P, Kepe V, et al. Prediction of cognitive decline by positron emission tomography of brain amyloid and tau. *Arch Neurol*. 2012;69:215–222.

Small G, Vorgan G. *The Alzheimer's Prevention Program*. New York: Workman; 2012.

Vandenberghe R, Adamczuk K, Dupont P, Laere KV, Chételat G. Amyloid PET in clinical practice: its place in the multidimensional space of Alzheimer's disease. *Neuroimage Clin*. 2013;2:497–511.

Yokomizo JE, Simon SS, Bottino CM. Cognitive screening for dementia in primary care: a systematic review. *Int Psychogeriatr*. 2014;26:1783–804.

57.3g Delirium

SUSAN W. LEHMANN, M.D., AND BENJAMIN LIPTZIN, M.D.

INTRODUCTION

Delirium is an acute cognitive disorder of attention and awareness which disproportionately affects geriatric patients. Reports indicate that delirium is present in one-third of hospitalized older patients and occurs in 15 percent of older patients presenting to emergency departments. The incidence of delirium is high among individuals with dementia, especially Alzheimer disease. Delirium is usually an indicator of severe underlying medical illness or metabolic derangement. However, it may also occur in vulnerable individuals with relatively minor medical provocations, such as constipation or mild dehydration.

Delirium is frequently unrecognized, yet the costs of not detecting

delirium are significant. Delirium is associated with poor outcomes, both in the hospital and after discharge. Patients who are delirious are more likely to have longer lengths of hospitalization. Older patients with delirium are at increased risk for nursing home placement and are at increased risk for prolonged cognitive impairment and dementia. Delirium is associated with high rates of morbidity as well as increased risk of mortality. Indeed, it has been estimated that the US healthcare costs which are attributable to delirium range from 40 to 150 billion dollars annually. Studies indicate that delirium may be preventable in up to 30 to 40 percent of cases, and so assessment for delirium has become a significant focus for patient safety management in acute care hospitals.

DEFINITION

Delirium is a multifactorial syndrome whose hallmark features are that of disturbance in attention, awareness, and cognition. The symptoms develop acutely, usually over hours or days, in the setting of serious medical illness or in the postoperative period. Symptoms include impaired ability to focus or sustain attention as well disorientation to the environment. Delusions or perceptual disturbances such as hallucinations (usually visual) are common, as are incoherent and disorganized speech. Patients usually have disrupted sleep–wake cycles and may be lethargic and drowsy or agitated and hyperalert. A key feature of delirium is that these disturbances tend to fluctuate in severity during the day.

The *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition (DSM-5) classifies delirium as a neurocognitive disorder. Other terms which are sometimes used by physicians to describe delirium, but are not recognized by DSM-5, are "acute confusional state," "acute brain failure," "toxic encephalopathy," or "metabolic encephalopathy."

HISTORY

The syndrome of delirium has been described in medical writing for over 2,000 years. It was recognized as early as the fifth century BC, when Hippocrates recorded a detailed description of the clinical features of delirium, noted that the syndrome was linked to physical disease, and suggested that it carried a poor prognosis. In the first century ad, the Roman, Celsus, distinguished delirium from major mood disorders and hysteria. In the second century, Galen also noted that the brain could be disturbed *secondarily* in diseases such as pneumonia. In antiquity and in medieval times, delirium was recognized, usually in the context of febrile illness, and was called by a variety of different names.

In his classic textbook, *General Psychopathology,* first published in 1913, the German psychiatrist, Karl Jaspers reported his observations of delirious patients as follows: "Delirium is characterized by the patient

turning away from external reality . . . He is dominated by an anxiety which is often extreme and a drive for aimless activity . . . When left alone the patient is always on the verge of sleep which however is never fully achieved."

In 1935, Harold Wolff and Desmond Curran published a study of 106 patients with delirium hospitalized in the United States and United Kingdom, with a careful cataloging of psychobiologic functions and case reports for 20 of the patients. These authors made several observations of note—for example, that patients recede into vagueness and hallucinations when external stimulation ceases, and that summoning relatives to the bedside of the delirious patient does much to restore function. Most important, however, was that they established for this cohort of patients that symptomatology was independent of etiology in delirium.

In the 1940s, John Romano and George Engel performed a series of EEG studies in delirious patients, and demonstrated that the degree of EEG slowing was correlated with the severity of the disturbance in consciousness. These authors hypothesized that delirium represented a state of "cerebral insufficiency."

More recently Zbigniew J. Lipowski wrote extensively about delirium and emphasized both the prevalence and seriousness of delirium as a public health problem. His work ushered in the contemporary era of research about delirium and his book, *Delirium: Acute Confusional States,* published in 1990, is considered a classic in the field. Sharon Inouye and her colleagues in geriatric medicine have carried out a series of investigations regarding the prevention of delirium which led to the development of a successful preventive multistep approach to delirium prevention and management. This work has highlighted the importance of recognizing delirium as an indicator of quality of care.

COMPARATIVE NOSOLOGY

The *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition (DSM-5) defines delirium as a neurocognitive disorder whose key features are disturbances of attention and awareness which develop acutely and which fluctuate over time. Typically, these changes in attention and awareness ebb and flow throughout the day such that the patient may appear lucid and talkative at one point in time, and then a few hours later might appear lethargic and confused. In addition, there are impairments in other aspects of cognitive functioning, such as orientation to time and place, memory, expressive and/or receptive language, and visuospatial abilities. Abnormal perceptions, especially visual and tactile hallucinations, are common. There must be evidence from the patient's history and physical examination or laboratory investigations to indicate that these disturbances are due to an underlying medical illness or illnesses or due to substance intoxication or withdrawal.

The DSM-5 criteria include specification about whether the delirium is acute (i.e., lasting a few hours or a few days) or if it is persistent (i.e., lasting weeks or months). In addition, the DSM-5 recognizes three psychomotor subtypes of delirium: hyperactive, hypoactive, and mixed level of activity.

For delirium, due to another medical condition, the DSM-5 requires that there be evidence from the history, physical examination, or laboratory findings that the delirium is due to a specific underlying medical condition. In clinical settings, it is unusual for there to be one single, identifiable cause that is clearly the etiology of the delirious state, and, more commonly, delirium is attributable to multiple causes. DSM-5 also specifies substance intoxication delirium and substance withdrawal delirium, which should be used when the symptoms of disturbed attention, awareness, and impaired cognition are most significant from a clinical standpoint and require clinical attention.

In the *International Classification of Diseases*, 10th revision (ICD-10), delirium due to medications or substances is specified separately by the causative medication or substance.

EPIDEMIOLOGY

Incidence and prevalence rates for delirium vary widely, depending on the clinical setting. Prevalence in the community is approximately 0.4 percent of adults and 1.1 percent of people aged 55 and older. Among patients over age 70 admitted to a hospital, about one-third have delirium. Of these individuals, about half are delirious at the time of admission, and the rest develop delirium during their hospitalization. Onset of the presenting symptoms of delirium is frequently a reason that a patient is brought to an emergency department for evaluation. Overall, the incidence of delirium during hospitalization ranges from 6 to 56 percent, with higher rates being more commonly seen among patients with dementia. Incidence rates of delirium also vary across hospital services. Incidence rates of delirium are highest in intensive care units, postoperative clinical units, and in palliative care settings. Among residents in nursing homes who are aged 75 and older, the prevalence of delirium is as high as 60 percent. Among patients with terminal illnesses, up to 80 percent of individuals become delirious around the time of death.

Table 57.3g–1.
Risk Factors for Delirium

Pre-existing cognitive impairment (dementia, executive dysfunction)
Previous history of delirium
Multiple medical problems
Multiple medications (especially opioids, anticholinergics, benzodiazepines)
Advanced age
Structural brain disease (e.g., stroke or traumatic brain injury)

Alcohol or sedative dependence
Chronic liver disease
Severe medical illness
Terminal illness
Chronic psychiatric illness
Undernutrition (low albumin and prealbumin)
Compromised cerebral circulation
Deafness or hearing impairment
Blindness or visual impairment
Sleep deprivation
Immobilization
Stress, including psychosocial stress

Risk factors for delirium are listed in Table 57.3g–1. These include pre-existing characteristics which predispose an individual to delirium as well as factors which precipitate an episode of delirium. Predisposing risk factors include pre-existing cognitive impairment, advanced age, previous history of delirium, severe medical illness, alcohol or sedative dependence, hearing or vision impairment, and poor nutritional status. Precipitating factors include bed rest or immobilization, medications (especially opioids, anticholinergic medications, and benzodiazepines), physical restraints, sleep deprivation, and low hematocrit. Awareness of risk factors for delirium, especially those which are modifiable, is critical for delirium prevention efforts.

ETIOLOGY

While delirium can be caused by a single etiologic factor, most commonly the etiology is multifactorial. Predisposing risk factors, as described above, confer vulnerability to delirium. Development of delirium depends upon the complex interplay of predisposing and precipitating factors. In older individuals with dementia, a relatively minor new condition, such as a urinary tract infection, can trigger an episode of delirium. At the same time, delirium often develops as a consequence of serious medical and/or metabolic derangement, and therefore, is considered a medical emergency. Recognition of the multifactorial etiology for delirium is essential to clinical management because a multidimensional approach to both prevention and treatment is needed.

Table 57.3g–2presents a list of etiologies for delirium frequently seen in older patient populations. Among hospitalized elders, the most important causes are infections and medications. Infections can include urinary tract infections, pneumonia, and sepsis. Many medications have been implicated as being deliriogenic in vulnerable populations and their use should be avoided, if possible. In particular, opioids, anticholinergic medications, and benzodiazepines are frequent causes of delirium in older patients. Anticholinergic medications include atropine, diphenhydramine,

benztropine, and oxybutynin (Ditropan) as well as medications with strong anticholinergic side effects such as imipramine, ranitidine (Zantac), dipyridamole (Persantine), warfarin (Coumadin), isosorbide, theophylline (Theo-Dur), nifedipine (Adalat, Procardia), and prednisolone (Prelone). Among opioid medications, meperidine appears to be most likely to induce delirium, and hydromorphone is the least likely to cause delirium. Among the benzodiazepines, the most lipophilic drugs with the longest elimination half-lives are most likely to cause delirium, including chlordiazepoxide (Librium), flurazepam (Dalmane), and diazepam (Valium).

Table 57.3g–2.
Etiologies of Delirium in Elderly Patients

Systemic Illnesses
Infections: Pneumonia, urinary tract infection, sepsis, influenza
Metabolic derangements: Dehydration, hypoxia, hypoglycemia, hyperammonemia, uremia, hyponatremia, thiamine deficiency, hyperthyroidism
Cardiovascular conditions: Arrhythmia, congestive heart failure, myocardial infarction, severe hypertension
Medications
Anticholinergics
Benzodiazepines, other sedative hypnotics (e.g., barbiturates)
Opioids
Antiarrhythmics
Certain antibiotics (e.g., fluoroquinolones, clarithromycin)
Digoxin
Interferons
Withdrawal States
Alcohol
Benzodiazepines, other sedative hypnotics
Primary Brain Diseases
Stroke or transient ischemic attack
Trauma: Brain injury, subdural hematoma
Infection/inflammation: Abscess, meningitis, encephalitis, vasculitis
Aneurysm
Epilepsy: Nonconvulsive status epilepticus, postictal states
Surgery or Trauma
Hip fracture repair
Open heart surgery (e.g., coronary artery bypass grafting)

A 79-year-old woman with metastatic ovarian cancer was admitted to the hospital with acute onset of vomiting, inability to stand, and delirium. She had been treated with warfarin for a hypercoagulable state. On admission workup she was found to have a cerebellar hemorrhage with evidence of increased intracranial pressure. Anticoagulation was reversed, the hemorrhaging stopped, and intracranial pressure was normalized. Digoxin was started for newly diagnosed atrial fibrillation. She returned to her baseline level of cognitive function, which was completely normal. She was transferred to a rehabilitation facility to improve her ability to ambulate, which was still affected by the cerebellar hemorrhage. Soon after her admission to the rehabilitation unit, she was reported to be disoriented. She began to accuse the staff of waking her up at 3 AM. Labs were ordered to determine the cause of her mental status change. Her digoxin level was elevated at 2.4 ng/mL (normal range 0.5 to 2.0 ng/mL). Her dose of digoxin was reduced and her symptoms of delirium resolved over several days

without other intervention.

The definitive pathophysiology of delirium remains unclear but evidence indicates that several sets of biologic factors appear to be involved. It is thought that delirium represents a reversible syndrome of neuronal dysregulation caused by disruption of oxidative metabolism. Certain brain structures, such as the cerebral cortex and hippocampus, are relatively more sensitive to this disruption than are other parts of the brain. At the cellular level, disruption of oxidative metabolism results in an inability to maintain ionic gradients. In addition, there is abnormal synthesis, release and metabolism of neurotransmitters, and inability to dispose of neurotoxins.

The two neurotransmitter systems which are believed to have the greatest causal relationship to delirium are the cholinergic and dopaminergic systems. Cholinergic deficiency and dopamine excess have both been implicated in delirium. Decreased cholinergic neurotransmission is believed to underlie the cognitive symptoms seen in patients with delirium. Cholinergic deficiency also correlates with adverse cognitive effects of anticholinergic medications. Acetylcholine is the primary neurotransmitter of the ascending reticular activating system (RAS), which regulates alertness. There are neocortical and limbic inputs to this system which control attention. Cholinergic, dopaminergic, and serotonergic neurons are especially sensitive to oxidative stress. In hypoxia, not only are derangements of these systems found, but there is also an increase in release of glutamate. It has been hypothesized that the presence of dopamine is required for at least some glutamate-mediated neuronal injury in hypoxia. Dopamine blockade, with agents such as haloperidol, may be protective by reducing neuronal damage in hypoxia in vulnerable brain areas such as the hippocampus. Increased dopaminergic neurotransmission is also believed to underlie the perceptual disturbances often seen in delirium and antidopaminergic antipsychotic drugs effectively treat these disturbances. Other neurotransmitter disturbances of GABA, serotonin, norepinephrine, and histamine may also be involved in delirium pathogenesis. GABA hyperactivity, for example, is implicated in hepatic encephalopathy. Delirium tremens is associated with GABA hypoactivity, and hyperactivity of glutamate and norepinephrine.

Neuroinflammation is another likely mechanism of brain injury in delirium and may help explain how systemic infection can disrupt cerebral oxidative metabolism. It has been hypothesized that peripheral inflammation triggers circulating proinflammatory cytokines which, in turn, promote endothelial activation, disruption of the blood–brain barrier, and microglial activation. Neuroinflammation of the brain can lead to microglial overactivation, with further neuronal injury. Elevations in proinflammatory cytokines have been found with diverse etiologies of

delirium, including the postoperative state (e.g., total hip replacement), influenza, heatstroke, and acute hepatic encephalopathy. It is believed that proinflammatory cytokines may have a direct neurotoxic impact on the brain.

DIAGNOSIS AND CLINICAL FEATURES

Delirium is a clinical diagnosis based on the history, physical examination, and mental status examination. However, the first step is suspicion for the disorder based on clinical observation. Because the features of delirium are quite variable and, by definition, tend to fluctuate over time, it is often overlooked or unrecognized by clinicians. Patients with delirium may seem so drowsy that they fall asleep while eating or speaking. At other times, the patient may seem inattentive to questions, with poor eye contact, easy distractibility by irrelevant stimuli and requiring questions to be repeated because of difficulty sustaining focus on the conversation. These fluctuations tend to occur throughout a 24-hour period so that a patient might be relatively alert and lucid with appropriate orientation to the environment for a few hours during different parts of the day, alternating with periods of confusion, incoherence, or somnolence. This waxing and waning of orientation and level of awareness might not be evident to an observer who interacts with the patient during only one of these periods during the day. In the acute hospital or nursing home setting, where clinical staff change with each shift, it is important for physicians to be alert to changing descriptions of the patient's mental state and behavior across different shifts to specifically look for this pattern of symptom fluctuation. It is not uncommon to find that hospital staff who work on evening shifts and day shifts have different views of the patient's level of cognitive and behavioral disturbances.

The speech of the patient with delirium is often incoherent, rambling, irrelevant to the topic and disorganized. Specific language impairments include dysarthria (impaired speech articulation), dysnomia (impaired naming), and dysgraphia (impaired writing ability). The patient's ability to reason abstractly, retain information, and direct his or her thinking is impaired. As a result, in the medical setting, patients with delirium may temporarily lose the capacity for informed consent. Frequently, patients who are delirious have sleep–wake cycle reversals and may waken during the night believing it is time to get washed and dressed. Disorientation to the environment is common, with patients mistaking the hospital for home and beginning to wander out of the hospital bed looking for the kitchen or a family member.

Other features commonly associated with delirium include transient delusional thoughts and emotional disturbances such as paranoia, fear, and disruptive behaviors. Delusional thinking is usually paranoid in nature with patients becoming fearful that routine or specialized medical care is being

done to them with a nefarious purpose. Perceptual disturbances due to delirium can be among the most frightening symptoms to patients and to their families. These problems vary in severity and frequently include misinterpretations of items and people in the patient's environment (e.g., belief that the physical therapist is a family member or mistaking the bang of a closing door for a gunshot). Illusions, which are misperceptions of actual stimuli, are also common in delirium. An example of an illusion might be a delirious patient misperceiving an intravenous line to be a snake. Hallucinations are perceptions in the absence of actual stimuli and are usually visual but they may also be tactile or auditory in nature. The hallucinations which occur in delirium range in severity from simple colored spots or geometric figures to moving objects, such as white handkerchiefs falling, to more complex phenomena, such as animals, people, monsters, panoramic views, and scenes acted out. Typically, these hallucinations are felt as threatening to patients. It is easy to see how they can engender feelings of fear and anxiety.

Disruptive behaviors can be present more or less continuously or they can occur unpredictably. These behaviors range from disorganized, purposeless actions such as picking at bandages or repetitive vocalizations to pulling out Foley catheters or attempting to get out of bed and ambulate without assistance. Distorted perceptual experiences can lead a patient to incorrectly interpret a medical situation as dangerous, leading him or her to attempt to pull out intravenous lines or endotracheal tubes or to respond aggressively to a nurse trying to assist the patient. While disruptive behaviors can occur at any time of the day, they are more common at night.

When delirium is suspected, the process of making the diagnosis includes a careful history, focusing on the timeline of symptom development. A key feature which distinguishes the impairments of delirium from those of dementia is acute development of symptoms. Delirium can present almost immediately after head trauma or it may develop over a period of hours or days in a hospitalized patient with pneumonia or other infection. An elderly outpatient with metabolic disturbances or medication toxicity might develop symptoms more gradually over a period of days to weeks. Unlike other cases of medical diagnosis, important aspects of the patient's history usually come from individuals other than the patient him/herself. Speaking with family members of the patient, or other knowledgeable informants who can provide a clear description of the patient's baseline level of cognitive functioning, is crucial in establishing that an acute change has occurred. When delirium occurs in a hospitalized patient, it is important to review nursing notes to help establish the timeline of abrupt onset of symptoms and the course of fluctuating symptoms. Other important aspects of the history include careful review of all medications prescribed to the patient and the timing of when new medications were begun or when dosages were

changed. A complete medication history review must include over-the-counter medications as well as prescription medications. In older patients, medications are the most common cause for delirium, with sedative-hypnotic, anticholinergic, opioid analgesic, and antidepressant medications being most potentially problematic. A thorough history must also include assessment of alcohol use, history of pain and its treatment, and any new physical symptom.

The physical examination of the patient should include careful evaluation for possible underlying causes, with assessment of vital signs including oxygen saturation and thorough cardiac and pulmonary examination. The neurologic examination should look for evidence of focal findings. Motor system abnormalities that can be seen in delirium include tremors, myoclonus, asterixis, increased deep tendon reflexes and increased tone. Evidence of adrenergic system overactivity can be seen in delirium due to alcohol or sedative withdrawal and also with anticholinergic toxicity states. Signs of hyperadrenergia include tachycardia, hypertension, sweating, pallor, flushing, and pupil dilation.

The mental status examination of the patient begins with assessment of the patient's level of awareness and arousal. It should be noted whether the patient is somnolent or lethargic, and if so, how much stimulation is required to arouse the patient. The patient's thinking and speech should be assessed, including whether speech is coherent or rambling and whether answers to questions are logical or disorganized. A number of simple tests can be used to evaluate the patient's level of attention. Often it is evident through interviewing the patient that he/she is having difficulty sustaining focus on the conversation and/or is being easily distracted by other stimuli in the room. A common test known as digit span requires patients to repeat 5 digits forwards and 3 digits backward. An alternative approach would be to ask patients to recite the days of the week or months of the year backward. In addition to impairment of awareness and attention, patients with delirium will have a change in at least one other area of cognitive functioning. This might include memory, learning, or disorientation. Perceptual disturbances, such as misinterpretation, illusions, and hallucinations should be assessed. A bedside cognitive test can be used to screen patients who are alert enough to cooperate with formal testing. Commonly used bedside cognitive tests are the Mini Mental State Examination (MMSE), Mini-Cog, and Montreal Cognitive Assessment (MoCA).

Several psychomotor phenotypes of delirium are described in DSM-5. Patients who have the hyperactive subtype of delirium exhibit hypervigilance, restlessness, distractibility, and agitation. Individuals with hyperactive delirium are easier to identify but may be more difficult to manage because of problematic behaviors which can include refusing to cooperate with medical care, combativeness, and wandering. Patients with

hypoactive delirium present primarily with withdrawal, lethargy, sparse speech, and decreased motor activity. Patients with mixed level of activity may appear to have a normal level of psychomotor activity even though attention and awareness are impaired. This subtype also includes individuals who have a rapidly fluctuating level of psychomotor activity between hypoactive and hyperactive states.

A number of assessment instruments have been developed to diagnose delirium and evaluate its severity. The most useful and commonly used bedside tool is the Confusion Assessment Method, or CAM. The CAM identifies key features of delirium and has been validated in many studies. The CAM-ICU is a version of the CAM that has been validated for use in the intensive care unit and does not require verbal responses from the patient. The Delirium Rating Scale-Revised-98 (DRS-R-98) has also been validated in both general medical populations and in geriatric populations. While an assessment tool does not take the place of a mental status examination, it can be very useful as a screening measure to help detect patients who are likely to be delirious. Daily use of a standardized screening instrument, such as the CAM, can be important for prompt identification of acute changes and fluctuations in cognition and attention in hospitalized older patients who are at risk of developing delirium. This can also be used to train nurses or other caregivers to recognize delirium.

PATHOLOGY AND LABORATORY EXAMINATION

The syndrome of delirium is diagnosed on the basis of the clinical history and the examination of the patient at the bedside. The diagnosis of delirium requires a prompt and careful laboratory evaluation to determine the etiology or etiologies. Delirium frequently heralds serious medical illness and therefore should be responded to as a medical emergency. A standard laboratory evaluation of delirium is shown in Table 57.3g–3. As indicated in Table 57.3g–3, if the basic evaluation fails to reveal a likely cause, then other selected laboratory investigations may be initiated. Depending on the particular case, further workup may include a brain imaging study or a lumbar puncture for cerebrospinal fluid studies. For instance, a lumbar puncture should be done if there is concern about possible meningitis or subarachnoid hemorrhage. A brain computed tomography (CT) or magnetic resonance imaging (MRI) should be done if there is a high possibility of stroke or head trauma causing brain hemorrhage. Arterial blood gas evaluation should be done if there is concern for hypercarbia in a patient with chronic obstructive pulmonary disease. Detailed neuropsychologic testing is not usually obtained because the attentional impairment of delirium compromises the validity of higher-order testing.

Table 57.3g–3.
Laboratory Workup for Delirium in the Elderly Patient

- Basic evaluation
  - Blood tests
    - Complete blood count (CBC) with differential and platelets
    - Chemistries:
      - Albumin
      - Blood urea nitrogen (BUN)
      - Creatinine (Cr)
      - Creatine phosphokinase (CPK)
      - Electrolytes: Sodium (Na^+), potassium (K^+), chloride (Cl^-), bicarbonate (HCO_3^-), calcium (Ca^{2+}), phosphate (phos), magnesium (Mg^{2+})
      - Glucose
      - Liver function tests
    - Erythrocyte sedimentation rate or C-reactive protein
    - Syphilis serology
    - Medication levels
    - Toxicology screen (serum and urine)
- Urinalysis (with culture and sensitivity testing, if indicated)
- Electrocardiogram
- Chest radiograph

Other laboratory studies to consider, depending on clinical indications: brain imaging (computed tomography or magnetic resonance imaging), electroencephalogram, arterial blood gas, cerebrospinal fluid studies, vitamin B_{12} level, folate level, thyroid-stimulating hormone, heavy metal screening, human immunodeficiency virus testing, and serum anticholinergic activity (available through research laboratories).

Electroencephalography (EEG) has a characteristic pattern in delirium and typically shows generalized slow wave activity in the delta and theta range. In hyperadrenergic states, such as alcohol or sedative withdrawal, the EEG will show increased low-voltage fast (beta) activity. However, the EEG is rarely useful in determining the underlying cause of delirium and is not routinely performed.

DIFFERENTIAL DIAGNOSIS

Although the symptoms of delirium can mimic a number of psychiatric conditions, in the elderly patient, the most important differential diagnosis is between delirium and dementia. To distinguish between delirium and dementia requires obtaining a thorough history from family members and caregivers, or from nursing staff who have recently been caring for the patient, and who can provide a description of the patient's baseline level of function and timeline of recent acute changes. Since dementia is a recognized risk factor for delirium, the two diagnoses often coexist in the same patient. Information from family members about the patient's premorbid functioning—such as when the patient last drove a car, balanced a checkbook, or dressed independently—can help identify a more slowly

developing dementia. Description of a more rapid decline in general and cognitive functioning over days to weeks, and/or a rapidly fluctuating course, suggest delirium evolving in a patient with pre-existing dementia. For patients hospitalized on acute medical units, it is imperative to review reports from evening and night shift nursing staff in addition to day shift staff to look for fluctuations in attention, cognitive abilities, and behavior over the course of a 24-hour period. Table 57.3g–4 lists features that can help distinguish delirium from dementia.

Patients who are delirious can appear depressed or manic and frequently are anxious and psychotic. Patients with the hypoactive subtype of delirium tend to be withdrawn, minimally reactive, with minimal speech, poor oral intake, and flat affect and can be mistaken as being depressed. While older patients who are depressed frequently exhibit impairment in short-term memory and in concentration they do not experience fluctuations in awareness and rapid fluctuations in cognitive functioning that are characteristic of delirious patients. Conversely, patients with the hyperactive subtype of delirium may be agitated, incoherent, and psychotic and may be mistaken as being manic. A patient who presents with an acute change in attention and cognition with fluctuating symptoms should be evaluated for delirium, whether or not the patient has a history of depression or bipolar illness, to prevent the possibility of not detecting a serious underlying medical disorder.

COURSE AND PROGNOSIS

An episode of delirium can be observed to progress through prodromal, overt, and resolving phases. During the prodromal period there are typically symptoms such as anxiety, restlessness, irritability, mild confusion, difficulty concentrating, poor sleep, nightmares, daytime drowsiness, social withdrawal, increased sensitivity to light and noise, illusions, and hypnopompic and hypnagogic hallucinations. There is growing recognition of a subsyndromal delirium, in which patients present with some, but not all, of the characteristic features of delirium. The prodrome may persist as subsyndromal delirium or it may evolve into overt delirium over a period of days or weeks. Patients with subsyndromal delirium have outcomes intermediate between those with no symptoms of delirium and those who meet the full diagnostic criteria.

Table 57.3g–4.
Differential Diagnosis: Delirium versus Dementia

	Delirium	Dementia
Onset	Acute	Often insidious
Duration	Variable; usually brief	Chronic, unless reversible
Consciousness	Impaired, fluctuating	Normal

Orientation	Impaired	Variable
Memory	Impaired	Impaired
Attention	Impaired	May be normal
Thought	Disorganized; transient delusions; content rich	Often stereotyped; content poor
Perception	Florid hallucinations (especially visual)	No hallucinations or simple hallucinations
Insight and judgment	Poor, but may return partially during lucid intervals	Variable, depending on etiology and stage
Mood and affect	Often anxious; may appear depressed	May be depressed or anxious
Sleep	Disturbed	May be normal
Electro-encephalogram	Abnormal: generalized slowing and slow posterior dominant frequency	Normal or mild slowing of posterior dominant frequency

Delirium that develops in the context of severe medical illness, such as sepsis, may progress to stupor, coma, and death if appropriate and prompt treatment of the illness is not administered. With appropriate treatment of the underlying medical illness and metabolic abnormalities, delirium will generally resolve. During the resolution phase, psychotic symptoms lessen, sleep cycles normalize, and clarity of awareness returns but cognitive deficits may be more persistent. The stage of resolving delirium in elderly patients may last from days to months. Persistent delirium is associated with poor functional recovery from surgery and from medical illnesses and is also associated with increased length of hospital stays and with higher medical costs. Patients who are recovering from delirium will usually require additional assistance for activities of daily living and close monitoring of their medical condition and cognitive functioning.

In the absence of neurodegenerative disease or dementia, the signs and symptoms of delirium usually resolve completely. In patients with Alzheimer disease or Parkinson disease, cognitive deficits can remain such that patients function below their previous baseline levels. Recent research has indicated that patients with Alzheimer disease may experience more rapid cognitive decline following an episode of delirium. In addition, there is also growing evidence that in vulnerable older adults, delirium can lead to permanent cognitive impairment and increases incident dementia. The pathophysiologic mechanism underlying delirium-induced dementia is not clear, but may involve processes other than those associated with Alzheimer disease. A serious medical illness that leads to delirium can also be thought of as a “stress test” for the brain that can unmask underlying cognitive deficits that were not obvious previously.

Most patients who have had a serious episode of delirium, retain little or no recall for events they experienced while delirious. Patients who had a mild-to-moderate episode of delirium are more likely to have memories of their experience. Some patients develop residual anxiety or even a posttraumatic stress disorder (PTSD) as a consequence of the fearful emotions and frightening perceptual disturbances they experienced when

they were delirious.

The presence of delirium is associated with a significant elevation in mortality risk for elderly patients, both during the hospitalization and in the 6 to 12 months after discharge from hospital. This association remains even after controlling for severity of underlying illness and other confounding factors. Delirium symptoms which persist at discharge from hospital are associated with a high rate of nursing home placement. Patients who have experienced an episode of delirium and have recovered remain at increased risk for a recurrent future episode of delirium.

PREVENTION

In recent years, there has been considerable interest in efforts to prevent delirium. The most widely used strategy is a multicomponent nonpharmacologic approach devised by Sharon Inouye and colleagues known as the Hospital Elder Life Program (HELP). This multidimensional strategy has been used for hospitalized patients with diverse medical problems. The HELP interventions include cognitive reorientation, therapeutic activities and early mobilization (such as daily exercise and walking), reduced use and reduced doses of psychoactive drugs, nonpharmacologic promotion of sleep, maintenance of oral hydration and adequate nutrition, and provision of appropriate hearing and vision aids. The HELP program has been instituted in over 200 hospitals and has been found to decrease both the incidence and duration of delirium.

Delirium prevention efforts have also focused on patients undergoing surgery since delirium is a common postoperative complication, especially in older adults. Studies have looked at patient baseline characteristics as well as intraoperative and postoperative management and their association with delirium. Current evidence indicates that the route of intraoperative anesthesia (e.g., general, spinal, or epidural) does not affect development of delirium. However, the total dose of anesthetic drugs may have a role in causing delirium. A recent double-blind randomized study found that using lighter sedation for elderly patients undergoing hip fracture repair surgery led to a 50 percent drop in the risk of postoperative delirium. In addition, minimizing use of benzodiazepines and certain opioid pain medications, especially meperidine, in the postoperative period can help decrease the incidence of delirium.

TREATMENT

The most important aspect of treatment is to identify and treat underlying medical illnesses and derangements causing delirium. In many cases, timely and appropriate medical treatment will be lifesaving to the patient. When delirium is detected in elderly patients, it should be considered a medical emergency, resulting in immediate workup to determine the underlying cause or causes and to initiate treatment. In older patients, a

single cause for delirium usually cannot be identified, as delirium is more commonly multifactorial in etiology. Most of the time, an older patient with delirium will be hospitalized in order to adequately evaluate the patient and address treatable causes.

Table 57.3g–5lists critical medical interventions which should be used in managing elderly patients with delirium. Besides diagnosing underlying medical causes, attention should be given to all possible contributors to the delirious state, especially modifiable predisposing factors. These include the elimination of all unnecessary medications, treating hypoxemia and anemia, correcting vision and hearing impairments which lead to sensory deprivations, maintaining good hydration and nutrition, preventing urinary retention and fecal impaction, and minimizing pain. Reducing the use of indwelling urinary catheters can prevent catheter-associated urinary tract infections and associated delirium. Close monitoring is necessary for the patient's safety and to prevent aspiration, falls, or inadvertent self-harm. Physical restraints should be avoided as much as possible. Use of physical restraints has been associated with worsening of psychotic symptoms and with medical complications such as deep vein thrombosis, pulmonary embolism, and cardiac dysrhythmias in delirious patients.

Table 57.3g–5.
Medical Interventions for Delirium in Elderly Patients

Ensure safety	Close observation by family member or staff may be indicated
Ensure adequate hydration and nourishment	Patient cannot monitor intake or feeding; this usually means intravenous hydration is indicated Eating and drinking should be observed to ensure adequate intake, and to avoid choking/aspiration
Ensure adequate oxygenation	Consider supplemental oxygen via nasal cannula to keep O_2 saturation >95%
Attention to bowel/bladder function	Watch intake/output Ensure regular bowel movements via appropriate bowel regimen
Investigate common etiologies of delirium	Perform the following: Chest radiograph to rule out pneumonia Urinalysis to rule out urinary tract infection Physical/lab exam to rule out dehydration/electrolyte disturbance Consider medication effects (not just toxicity, but ANY medication linked temporally to delirium onset) Hold nonessential medications
Treat pain adequately	Recommended drugs include scheduled acetaminophen for mild pain and low-dose hydromorphone for severe pain Avoid meperidine
Increase activity level as appropriate and tolerated	Get patient out of bed to chair as soon as possible Include written plan in chart to maintain mobility

Table 57.3g–6lists environmental, nonpharmacologic interventions which have been shown to be beneficial in the management of patients with delirium. These include taking steps to help restore a normal sleep–wake

cycle by minimizing light, noise, and disruptions during the night while increasing exposure to light and activity during the day. Many patients benefit from being moved closer to a window, where natural daylight and night cues can be appreciated. It is important for staff and family to be educated about appropriate ways to interact with the patient to promote cognitive functioning. Delirious patients should not be left alone for long periods. Family members and close friends of the patient should be encouraged to visit and should be allowed to stay with the patient overnight, if they so desire. Hospital staff and family members should reorient the patient to time, place, and person frequently during the day. Both overstimulation and understimulation should be avoided, as either extreme can exacerbate symptoms of delirium. The patient should be given only one task to do at a time, rather than multistep instructions. For patients who are anxious or fearful, frequent calm reassurance by family and staff can minimize agitation.

Table 57.3g–6.
Environmental Interventions for Prevention and Management of Delirium in Elderly Patients

Appropriate lighting	Bright during day Dark at night
Promote sleep at night	Low nighttime noise level Pagers on silent mode Telephone ring turned down Television turned off Voices kept low Doors closed as patient care permits Minimize necessary awakenings
Frequent reorientation	Staff to reorient patient to time, place, and person verbally each shift Clock and calendar in room Familiar possessions in room
Education	Family to be educated about the cause, course, and prognosis of delirium Staff to be educated about behavioral dysregulation in delirium and its management
Promote sensory function	Hearing aids on and function checked Glasses on and cleaned
Comfort/reassurance	Personal effects from home at bedside Family and close friends encouraged to visit and stay with patient, if possible Reassurance provided to patient

There are no FDA-approved medications to treat the behavioral disturbances in delirium. The first-line approach to managing agitation and psychotic symptoms in delirium is to try the nonpharmacologic interventions previously described, including reorientation, support, and reassurance. However, pharmacologic treatment may be necessary for agitated delirious patients for whom verbal reassurance is inadequate and

for patients who exhibit behaviors which are dangerous to themselves or others. Pharmacologic treatment is also indicated for patients with psychotic symptoms causing significant fear and anxiety.

In deciding which medication to use in treating behavioral symptoms of delirium, it is important to determine whether alcohol or sedative withdrawal is a cause of delirium. For patients experiencing alcohol or sedative withdrawal delirium, benzodiazepines are the medications of choice. Lorazepam or oxazepam may be preferable to chlordiazepoxide, especially in patients with decreased liver metabolism, because they are not metabolized by the liver and have shorter half-lives. It should be noted that benzodiazepines are not without adverse effects, especially in elderly individuals. These drugs can affect respiration and hemodynamics and can cause dysrhythmias. They can increase confusion and the risk of falls, and therefore, hip fractures. Paradoxical reactions, in which increased agitation and disinhibition are seen rather than sedation, are observed with greater frequency among patients with liver disease or low serum albumin levels. In the treatment of alcohol or sedative withdrawal, a common dosing guideline for lorazepam is as follows. Severity of withdrawal is monitored using a structured instrument such as the Clinical Institute Withdrawal Assessment for Alcoholism, revised (CIWA-Ar).

Lorazepam 1 to 2 mg is given orally, intravenously or intramuscularly every 4 hours on a scheduled basis on day 1. This dose is decreased by 50 percent on day 2, and it is decreased again by 50 percent on day 3. Score the CIWA-Ar every 4 hours. If the score increases to greater than 8 to 10, then the frequency of lorazepam doses is increased to every 1 to 2 hours until symptoms are controlled, then the original protocol is restarted. Patients with full blown delirium tremens will likely require higher and more frequent doses.

A 78-year-old male was admitted to the emergency department after a 5-week alcohol binge. He was known to have a long history of alcohol dependence, with multiple prior episodes of delirium tremens and probably alcohol-related dementia. On admission, clinical examination and laboratory testing revealed malnutrition, urinary tract infection, and pneumonia. The patient was started on oral lorazepam, but within 24 hours was transferred to the intensive care unit with tachycardia and visual and auditory hallucinations. A lorazepam infusion was started, along with a beta blocker for control of heart rate. Over the next 24 hours, the patient's condition stabilized. He was switched from the lorazepam infusion to intravenous bolus lorazepam with good control of symptoms. As his medical problems were treated, his mental state improved. The beta blocker was discontinued, and the lorazepam was tapered slowly over several weeks. He was discharged off medications 25 days after admission.

For delirious patients who are not experiencing alcohol or sedative withdrawal, benzodiazepines are avoided because they may exacerbate delirium and increase risk of falls. Antipsychotic medications are the agents of choice for treatment of psychotic and behavioral symptoms. Haloperidol, a typical antipsychotic, is highly effective for this indication, and, when

dosed appropriately, has minimal adverse effects. It does not cause significant hypotension, is not anticholinergic, does not depress respiration and does not cause or impair glucose regulation. Haloperidol can be administered orally or intravenously, and the latter route may be preferred in patients who are medically ill and unable to take oral medications. Intravenous dosing is also usually preferred for patients in intensive care units. The risk of QTc interval prolongation is minimized by staying within dosing guidelines and adhering to recommendations regarding QTc interval checks. Before haloperidol is initiated, a baseline QTc interval is obtained. If the interval is 450 ms or greater in a male, or 470 ms or greater in a female, intravenous haloperidol should not be used. Any patient whose QTc interval is prolonged beyond 25 percent of baseline during treatment should have haloperidol discontinued. There is evidence that intravenous haloperidol is less likely to cause extrapyramidal symptoms than oral doses of haloperidol.

A conservative but usually effective dosing schedule for an elderly delirious patient is as follows: haloperidol 0.25 mg to 0.5 mg intravenously every 8 hours on a scheduled basis, and 0.25 mg to 0.5 mg intravenously every 6 hours as needed to keep agitation and psychotic symptoms under control. If oral dosing of haloperidol is used, higher doses are given (e.g., 0.5 mg to 1 mg by mouth every 8 hours, and 0.5 mg to 1 mg by mouth every 6 hours as needed). Younger patients may require and tolerate higher doses of medication. How long the acute dose of haloperidol is maintained depends on how the patient responds clinically. If the patient becomes calm, and psychotic symptoms lessen within the first 12 to 24 hours (as usually happens), this dose of haloperidol is continued for 2 to 3 days and then is tapered over the next 2 to 3 days. If the patient does not become calm within this time frame, consideration should be given to a missed or inadequately treated medical condition prolonging delirium. In this case, haloperidol would likely be continued for a longer period.

Atypical antipsychotic medications have been shown to be just as effective as haloperidol in controlling symptoms of agitation and psychosis in delirious patients. These medications are generally well tolerated when used for acute episodes of delirium. One limitation is that they are given orally rather than intravenously. The two atypical antipsychotic agents which are most commonly used for delirium in elderly patients are risperidone and quetiapine, with dosing schedules as follows:

Risperidone 0.25 to 0.5 mg twice daily on a scheduled basis, and 0.25 to 0.5 mg every 6 hours on an as-needed basis.

Quetiapine 25 mg at bedtime to 25 mg twice daily on a scheduled basis, and 25 mg every 6 hours on an as-needed basis.

While delirium is recognized as a condition of cholinergic deficiency, a number of studies which looked at using cholinesterase inhibitors to treat

or prevent delirium did not find these agents to be beneficial. As a result, cholinesterase inhibitors should not be newly prescribed to either prevent delirium or to treat delirium in elderly patients.

An 84-year-old man was admitted from home with acute coronary syndrome. Medical history was significant for coronary artery disease, myocardial infarction, left carotid endarterectomy, chronic renal insufficiency, and renal cell carcinoma. He underwent three-vessel coronary artery bypass grafting. On the first postoperative day he was noted to be irritable, combative, and confused. He told his family that he was brought to the hospital for "experimentation," and had persistent tactile hallucinations of leaves and other objects in his mouth. Nighttime restlessness progressed to nocturnal insomnia and agitation. The patient had marked fluctuations in alertness during the day, with worsening agitation in the evening hours. During his agitated episodes, he pulled out his intravenous line several times. During a period of relative lucidity, an MMSE was performed and he scored 10/30. A psychiatric consultant recommended 0.5 mg of haloperidol orally every 8 hours on a scheduled basis for symptomatic treatment of his hallucinations and agitation. Medical complications (which included dehydration, atrial fibrillation, hypotension, and pneumonia) were managed as they arose. Although agitation and psychotic symptoms were improved, the patient required intermittent one-to-one observation by staff. After 7 days, his mental state began to improve as his medical condition stabilized. He was discharged on postoperative day 17, with some residual memory problems and insomnia, but without evidence of psychotic symptoms or agitation.

FUTURE DIRECTIONS

Animal models could be helpful in understanding the underlying pathophysiology of delirium. Studies have looked at the role of interleukins, GABA, dopamine, and acetylcholine—but more studies are needed. An animal model could also be helpful in testing preventive or treatment strategies. More studies of possible prophylactic or treatment effects of different antipsychotics would be useful. A possible role for melatonin or ramelteon, a melatonin receptor agonist, would be helpful based on preliminary studies. Dexmedetomidine has also been looked at in the prevention or treatment of delirium in postoperative patients and should be explored further. Different approaches to sedation for surgery also are promising areas for investigation. In particular, using light sedation, where possible, seems to be less likely to cause delirium.

REFERENCES

Bourne RS, Tahir TA, Borthwick M, Sampson EL. Drug treatment of delirium: past, present and future. *J Psychosom Res*. 2008;65:273–282.

Breitbart W, Alici Y. Agitation and delirium at the end of life: "we couldn't manage him." *JAMA*. 2008;300:2898–2910.

Cole MG, McCusker J, Ciampi A, Belzile E. The 6- and 12-month outcomes of older medical inpatients who recover from subsyndromal delirium. *J Am Geriatr Soc*. 2008;56:2093–2099.

Davis DH, Muniz Terrera G, Keage H, et al. Delirium is a strong risk factor for dementia in the oldest-old: a population-based cohort study. *Brain*. 2012;135:2809–2816.

de Rooij SE, van Munster BC, de Jonghe A. Melatonin prophylaxis in delirium: panacea or paradigm shift? *JAMA Psychiatry*. 2014;71(4):364–365.

Ely EW, Shintani A, Truman B, et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. *JAMA*. 2004;291:1753–1762.

Grover S, Kumar V, Chakrabarti S. Comparative efficacy study of haloperidol, olanzapine and risperidone in delirium. *J Psychosom Res*. 2011;71:277–281.

Hshieh TT, Fong TG, Marcantonio ER, Inouye SK. Cholinergic deficiency hypothesis in delirium: a synthesis of current evidence. *J Gerontol A Biol Sci Med Sci*. 2008;63:764–772.

*Inouye SK, Bogardus ST Jr, Charpentier PA, et al. A multicomponent intervention to prevent delirium in hospitalized older patients. *N Engl J Med*. 1999;340:669–676.

*Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Ann Intern Med*. 1990;113:941–948.

Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet*. 2014;383:911–922.

Laurila JV, Laakkonen ML, Tilvis RS, Pitkala KH. Prediposing and precipitating factors for delirium in a frail geriatric population. *J Psychosom Res*. 2008;65:249–254.

Lipowski ZJ. *Delirium: Acute Confusional States*. New York: Oxford; 1990.

Liptzin B, Laki A, Garb JL, Fingeroth R, Krushell R. Donepezil in the prevention and treatment of post-surgical delirium. *Am J Geriatr Psychiatry*. 2005;13:1100–1106.

MacLullich AM, Anand A, Davis DH, et al. New horizons in the pathogenesis, assessment and management of delirium. *Age Ageing*. 2013;42:667–674.

Marcantonio ER. Postoperative delirium: a 76-year-old woman with delirium following surgery. *JAMA*. 2012;308:73–81.

*Mayo-Smith MF. Pharmacological management of alcohol withdrawal. A meta-analysis and evidence-based practice guideline. American Society of Addiction Medicine Working Group on Pharmacological Management of Alcohol Withdrawal. *JAMA*. 1997;278:144–151.

McAvay GJ, Van Ness PH, Bogardus ST Jr, et al. Older adults discharged from the hospital with delirium: 1-year outcomes. *J Am Geriatr Soc*. 2006;54:1245–1250.

Meagher DJ, Moran M, Raju B, et al. Motor symptoms in 100 patients with delirium versus control subjects: comparison of subtyping methods. *Psychosomatics*. 2008;49:300–308.

Nadelson MR, Sanders RD, Avidan MS. Perioperative cognitive trajectory in adults. *Br J Anaesth*. 2014;112:440–451.

Saczynski JS, Kosar CM, Xu G, et al. A tale of two methods: chart and interview methods for identifying delirium. *J Am Geriatr Soc*. 2014;62:518–524.

*Schuckit MA. Recognition and management of withdrawal delirium (delirium tremens). *N Engl J Med*. 2014;371:2109–2113.

Shehabi Y, Bellomo R, Reade MC, et al. Early goal-directed sedation versus standard sedation in mechanically ventilated critically ill patients: a pilot study. *Crit Care Med*. 2013;41(8):1983–1991.

Siddiqi N, Young J, Cheater FM, Harding RA. Educating staff working in long-term care about delirium: The Trojan horse for improving quality of care? *J Psychosom Res*. 2008;65:261–266.

Sieber FE, Zakriya KJ, Gottschalk A, et al. Sedation depth during spinal anesthesia and the development of postoperative delirium in elderly patients undergoing hip fracture repair. *Mayo Clin Proc*. 2010;85(1):18–26.

Strijbos MJ, Steunenberg B, van der Mast RC, Inouye SK, Schuurmans MJ. Design and method

of the hospital elder life program (HELP), a multicomponent targeted intervention to prevent delirium in hospitalized older patients: efficacy and cost effectiveness in Dutch health care. *BMC Geriatr*. 2013;13:78.

Teslyar P, Stock VM, Wilk CM, Camsari U, Ehrenreich MJ, Himelhoch S. Prophylaxis with antipsychotic medication reduces the risk of post-operative delirium in elderly patients: a meta-analysis. *Psychosomatics*. 2013;54:124–131.

Trzepacz PT, Mittal D, Torres R, Kanary K, Norton J, Jimerson N. Validation of the Delirium Rating Scale-revised-98: comparison with the delirium rating scale and the cognitive test for delirium. *J Neuropsychiatry Clin Neurosci*. 2001;13:229–242.

Xia ZQ, Chen SQ, Yao X, Xie CB, Wen SH, Liu KX. Clinical benefits of dexmedetomidine versus propofol in adult intensive care unit patients: a meta-analysis of randomized clinical trials. *J Surg Res*. 2013;185:833–843.

Young J, Leentjens AF, George J, Olofsson B, Gustafson Y. Systematic approaches to the prevention and management of patients with delirium. *J Psychosom Res*. 2008;65:267–272.

57.3h Schizophrenia and Delusional Disorders

IPSIT V. VAHIA, M.D., AND CARL I. COHEN, M.D.

Over the past century, the vast majority of schizophrenia studies have primarily involved young adult patients. Indeed, only 1 percent of the literature on schizophrenia is dedicated to issues of aging. There has been a paucity of research on outcome of early-onset schizophrenia into later life as well as schizophrenia of late onset. Although research regarding late-onset schizophrenia is a relatively recent endeavor in the United States, late-onset psychosis has interesting historical ties. Accounts suggest that in 1693, Sir Isaac Newton experienced an episode of psychosis at the age of 51. His presentation included paranoid delusions, social withdrawal, and references to conversations that never occurred. Some individuals have attributed Newton's symptoms to metal poisoning or life events, but others have refuted such claims. This apparent case of late-onset schizophrenia-like psychosis that remitted after approximately 18 months, although debatable, is just one example of cases and ongoing research that seem to challenge the original Kraepelinian notion of dementia praecox in terms of age of onset and long-term course of schizophrenia. Indeed, a case of a new onset case of schizophrenia at age 100 has been reported by Cervantes et al. In this chapter, we examine older persons with early- and late-onset schizophrenia. We also briefly describe delusional disorder in older patients, which is sometimes viewed as part of a continuum with schizophrenia.

HISTORY AND COMPARATIVE NOSOLOGY

Emil Kraepelin's description of dementia praecox as a disorder of progressive decline with onset during adolescence or early adulthood has greatly influenced subsequent conceptualization of schizophrenia. Nonetheless, Kraepelin himself questioned the term *dementia praecox* after observing that the onset of this disorder was not limited to younger patients, nor was the course necessarily progressive. He used the term *paraphrenia* to describe a subset of patients with an onset of delusions and hallucinations that were primarily paranoid in nature and with a better prognosis. This group of patients had a number of symptoms similar to those of dementia praecox but with less alteration of emotion and volition. Kraepelin was said to refer to paraphrenia as the "darkest area of psychiatry."

Subsequently, several other European investigators examined the concept of late-onset schizophrenia in greater detail. The initial studies of late-onset schizophrenia were conducted by Manfred Bleuler, who described late-onset schizophrenia as being similar to early-onset schizophrenia, except for an age of onset after 40 years. Bleuler found that

late-onset cases comprised 15 percent of his schizophrenia patients and that roughly one-half of the patients with late-onset schizophrenia had symptoms very similar to those seen in early-onset individuals. Martin Roth et al. introduced the idea that late-onset schizophrenia could be a distinct entity after applying the term *late paraphrenia* to a group of patients with an onset of the disorder after the age of 60. These patients were generally women with sensory deficits, who had no family history of schizophrenia and presented with delusions and hallucinations. Years later, the term *late paraphrenia* was included in the ninth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-9). Kurt Leonhand and other German contemporaries used the term *paraphrenia* to refer to paranoid schizophrenic patients without specifying age of onset. Felix Post categorized nonaffective paranoid disorders beginning after the age of 50 into three groups: simple paranoid psychoses, schizophrenia–like illness, and paranoid schizophrenia. The former resembled delusional disorder in our current nomenclature, and the latter two categories were distinguished by the absence and presence of Schneiderian symptoms, respectively. The use of the terms *paraphrenia* and *late paraphrenia* led to some confusion because Kraepelin did not consider late age at onset a feature of paraphrenia, and paraphrenia is a more inclusive term and not synonymous with schizophrenia.

Criteria such as age of onset of schizophrenia have changed over the years. The first and second editions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-I and DSM-II, respectively) did not specify age of onset criteria for schizophrenia. DSM-I included the category "involutional psychotic reaction" for psychotic symptoms occurring during the involutional period, and DSM-II used the term *involutional state* or *involutional paraphrenia* to describe a delusional state with onset during the involutional period. DSM-II did not specify an age cutoff for the diagnosis of schizophrenia. Differing from the previous DSM editions, DSM-III criteria restricted the diagnosis of schizophrenia for those with symptom onset before age 45. DSM-III-R included a late-onset category for patients with initial presentations at the age of 45 or later. The revised third edition also replaced the term *paranoid disorder* with *delusional disorder*, modified the diagnostic criteria to include a minimum duration of 1 month, and expanded the scope of delusions to include grandiose, erotomanic, somatic, jealous, and unspecified types. The criteria for delusional disorder in DSM-IV and DSM-5 have not changed. The ICD-10, DSM-IV, and DSM-IV-TR criteria did not include codable diagnoses for late-onset schizophrenia. However, DSM-IV and DSM-IV-TR mentioned differences between cases of schizophrenia with onset after 45 years compared to those with earlier onset. DSM-5 does not distinguish types of schizophrenia based on age, and in fact, has eliminated the syndrome-based subtypes (paranoid, disorganized, undifferentiated, residual, and catatonic).

Research indicates that late-onset forms of schizophrenia are more similar than different from earlier-onset forms, but there may be some subtle distinctions. This remains an area of debate, with some studies questioning the existence of a late-onset form.

To further examine issues surrounding late-onset schizophrenia, such as diagnostic and nomenclature inconsistency and the need for further study, an international late-onset schizophrenia group met to review the literature and reach consensus regarding the previously discussed topics. The group concluded that there was proof to support two late-onset illness classifications and proposed the following: late-onset (onset after the age of 40 years) schizophrenia and very-late–onset (onset after the age of 60 years) schizophrenia-like psychosis. However, subsequent research has frequently combined these categories, and as a result, the literature on distinctions between these subtypes remains limited.

EARLY-ONSET SCHIZOPHRENIA

Course of Illness

Schizophrenia typically develops in the second or third decade of life but greater numbers of schizophrenia patients are surviving into old age. Between 80 and 85 percent of persons aged 55 and over developed schizophrenia before age 45. In the United States, the prevalence estimate for schizophrenia is 0.6 to 1.0 percent in people aged between 45 and 64 years and 0.1 to 0.5 percent in people aged 65 years and older. Over the first quarter of the 21st century there will be a doubling of the number of persons aged 55 and over with schizophrenia, to about 1.1 million in 2025. By that time, about one-fourth of persons with schizophrenia will be in this older age bracket. Likewise, worldwide, the number of people aged 60 years and older with schizophrenia will more than double between 2014 and 2050, and an estimated 10 million people will be in this age category by 2050.

There is substantial heterogeneity in outcome of schizophrenia in later life. Long-term studies observing the symptoms and functioning in early-onset schizophrenia suggest that the course of schizophrenia may not be as pessimistic as previously thought, and findings tend to challenge the notion implied in the term *dementia praecox*. Based on several long-term studies carried out in Europe ranging from 22 to 37 years, Ciompi found that approximately 20 to 27 percent of patients attained complete symptomatic remission, 22 to 33 percent attained mild end states, 24 to 29 percent attained intermediate end states, and 14 to 18 percent attained severe end-states. Within the latter group, there is a very poor outcome group of about 5 to 10 percent, which corresponds most closely to Kraepelin's description of dementia praecox. Studies conducted in seven countries have reported similar results to those reported by Ciompi. However, cross-sectional data

have obscured heterogeneity in course and outcome in later life. Recent studies conducted in New York City by Cohen et al. have challenged the notion of schizophrenia in later life being a quiescent or stable end stage: about two-fifths of persons fluctuated between remission and nonremission. Moreover, outcome in later life is now conceptualized multidimensionally (e.g., positive and negative symptoms, cognition, mood, social functioning, and quality of life [QoL]), and that these dimensions are only weakly intercorrelated and fluctuate over time. Thus, persons may be doing well in some dimensions and not others, and that improvements in one dimension may not induce improvements in other dimensions. Furthermore, looking at a more comprehensive definition of recovery, Auslander and Jeste found only 8 percent of older outpatients attained "sustained remission," based on symptom remission, social functioning, and medication dosage. These findings were similar to Ciompi's findings regarding "global mental health" in which he found only 12 percent was completely free of disturbances, and 19 percent suffered from mild disturbances of some kind. Jeste et al. found that although the course of schizophrenia appears stable in later life, older outpatients remained impaired in various measures of psychopathology, functioning, and well-being as compared to an age-matched comparison group. In other words, schizophrenia continues to be symptomatic in older age in most patients and the possibility of complete recovery exists in only a small proportion of persons.

Overall, life course patterns reflect the wide variations in individual histories, and are consistent with a biologic-personality-environmental model of mental illness. Heterogeneity of outcome is the rule as schizophrenia progresses into later life. Despite scarce data, some findings suggest that the diversity seen in younger individuals might be even greater in older age. Although the extremes of this continuum attract most attention, a majority of older patients with schizophrenia can be identified at some point in between these extremes, moving with time in one direction or the other.

While a prediction of the long-term course of the illness in an individual cannot be accurately made, the general factors associated with favorable outcome include female gender, later onset of symptoms, paranoid subtype, less severe negative symptoms, and better premorbid functioning. However, with age, the strength of association between these variables and outcome weakens. Consequently, predicting a probable long-term course for schizophrenia patients is difficult. Apart from the relatively small group of poor outcome patients who worsen over time and who are largely found among chronically institutionalized patients, community-dwelling persons with schizophrenia—who represent about 85 percent of the older population—have been generally found to have a nonprogressive course, but with likely fluctuations over time in individual dimensions.

Dimensions of Outcome

Recent investigations have focused on dimensions of psychopathology, comorbidity, and functioning in schizophrenia in the elderly.

Positive Symptoms

Long-term studies of outcome have demonstrated that there is an improvement in positive symptoms with age. In cross-sectional prevalence studies, roughly half of patients living in the community demonstrated significant clinical improvement in positive symptoms. However, a longitudinal study (mean 4.5 years follow-up) from a subset of a New York City community sample aged 55 and over found only one-fourth had persistent symptomatic remission.

Negative Symptoms

Negative symptoms are thought to predominate the clinical picture of schizophrenia in later life. However, estimating negative symptoms in older schizophrenia patients can be challenging, as the clinical picture is often confounded by depression, demoralization, cognitive decline, medication effects, and poverty. The few studies of negative symptoms in community dwelling older adults found prevalence rates of between one-quarter to one-half of the sample. Notably, longitudinal data from the New York City sample described above found that 44 percent fluctuated between the presence and nonpresence of negative symptoms. Thus, contrary to earlier notions of schizophrenia progressing to "burnt-out" in later life, negative symptoms do not typically dominate the clinical picture, appear to be no higher than in younger schizophrenia populations, and may fluctuate over time.

Cognition

While cognitive decline in schizophrenia has been well established, there is ongoing debate on whether the trajectory of cognitive function shows a static course over time, in line with the supposed neurodevelopmental nature of schizophrenia, or exceeds that of normal age-related decline, suggesting neurodegeneration. Shah et al.'s review of longitudinal studies of cognitive decline found slightly more studies indicating a faster cognitive decline in later life than expected with normal aging, although this difference disappeared if only studies with control groups were included. Furthermore, any differences in cognitive decline between early- and late-onset groups were inconclusive. Another review by Rajji et al. found heterogeneity of outcomes on the basis of residential status. Community-dwelling patients exhibited stability into their late 60s except for cognitive decline that is expected from normal aging, whereas a greater than age-normal cognitive decline in patients with long periods of

institutionalization, especially after age 65 years. Thompson et al. recently postulated a more granulated picture for community dwelling middle-aged and older persons with schizophrenia (ages 40 to 100). On 3.5-year follow-up, three trajectories were identified: "Stable" 50 percent; "Modest Decline" 40 percent; and "Rapid Decline" 10 percent.

There have been few longitudinal studies examining the incidence of dementia among older persons with schizophrenia. Work in this area has been impeded by challenges in distinguishing cognitive deficits dating back to earlier in the disease course from normal age-associated cognitive decline. Consequently, nearly half of persons with schizophrenia may meet cognitive cut-offs for mild dementia. A recent review noted only three longitudinal studies comparing persons with schizophrenia to controls, and two demonstrated an increased prevalence of dementia and one did not; none of the studies controlled for risk factors or identified the types of dementia that were found. An unresolved issue is whether persons with schizophrenia may be more prone to dementia because of their higher rates of metabolic syndrome, which is a risk factor for both Alzheimer disease and vascular dementia.

Depression

Among the studies on depression in older adults with schizophrenia, prevalence rates of depression ranged from 11 to 75 percent, depending on the level of severity that was examined. At all levels of depression, the prevalence was significantly greater than their age peers in the general population, although rates were similar to those of younger persons with schizophrenia. Because there is a substantial minority of persons who do not experience depression, it may be inappropriate to consider a depression to be a core feature of the disorder, although it is a common feature with prevalence rates that remain high throughout the lifespan.

A recent prospective study found that persistent depression (defined as the presence of syndromal or subsyndromal depression) occurred in 44 percent of persons, 30 percent remained persistently nondepressed, and 26 percent fluctuated between depression and nondepression.

The potential for suicide remains high in later life. Among depressed persons aged 40 and over with schizophrenia, 23 percent had current suicidal ideation and 49 percent had made a previous attempt. In another study that included depressed and nondepressed persons with schizophrenia aged 55 and over, 13 percent had current suicidal ideation, and 30 percent had made a suicidal attempt during their lifetime.

Medical Comorbidity

Life expectancy is dramatically shortened in schizophrenia (by up to 20 years). The mortality risk in schizophrenia is 2 to 3 times higher than in the

general population, and this differential mortality gap may be increasing. There is increasing recognition of high prevalence of cardiovascular disease, diabetes, and metabolic syndrome in patients with schizophrenia, that at least partially can be attributed to lifestyle factors (smoking, poor dietary habits, limited physical activity) and the effects of antipsychotic medication, especially second-generation agents. Studies have also noted significantly higher rates of congestive heart failure, chronic obstructive pulmonary disease, and hypothyroidism. Interestingly, studies have suggested the possibility of lower cancer rates. Kredentser et al. found deaths from cancer to be higher among middle-aged persons with schizophrenia but not in persons over age 59. Cardiac disease is the chief cause of mortality in older schizophrenia patients, who have higher rates of medical illness, greater severity of medical illness, and more untreated medical conditions than their age peers.

In addition, part of the excess morbidity and mortality in schizophrenia may be inherent to the disorder itself. The search for the biologic underpinning of this supposedly intrinsic propensity for accelerated physiologic aging now focuses on inflammation and oxidative stress. Older schizophrenia patients were also found to be less compliant with nonpsychotropic medication. More importantly, although they have access to care similar to their age peers in the general population; older schizophrenia patients tend to receive less adequate care. Higher medical comorbidity has been associated with depression and positive symptoms.

Community Integration and Social Functioning

Social functioning has been defined by Weissman et al. as "the ability to establish and maintain relationships with friends and family as well as to undertake work and leisure activities and cope with day-to-day activities." Daily adaptive functioning refers to the ability to function independently. Long-term studies indicate that social functioning is often restored in older schizophrenia patients and can even develop further after onset of schizophrenia, although as with other domains, there may be movement in positive and negative directions over time. According to one study, up to 30 to 40 percent of schizophrenia patients may be able to regain capacity to work in late life. Higher levels of daily adaptive functioning have been associated with better cognitive function, fewer negative symptoms, better physical health, and independent living. Cognitive capacity is an important predictor of both social and adaptive functioning.

A more recent body of work has focused on community integration in schizophrenia. Community integration can be conceptualized as "normalization" of functioning in society and comprises assessments of functioning and perceived QoL within several dimensions, viz., independence, physical, psychological, and social integration. In the cross-sectional study conducted in New York City, community-living older adults

had roughly twice the level of high community integration than those with schizophrenia (41 percent vs. 23 percent). In a longitudinal study using the same sample, 34 percent of persons fluctuated between high and low community integration, suggesting that there is movement toward improved social integration in later life but also the potential for increased social isolation and diminished functioning.

Quality of Life

QoL indicates an individual's own perception of their position in life in relation to their goals, expectations, and concerns. It is an indicator of individuals' own appraisal of their condition. Lower QoL has been found to be associated with depression, positive and negative symptoms, cognitive deficits, financial strain, poor social support, and impaired social skills. Lower QoL may also predict suicidality. Although QoL indices have been found to be lower in older community-dwelling schizophrenia patients than their age peers matched for race and income, the absolute difference between the groups is small, and both scored in the moderate range of QoL, and substantially higher than persons with chronic pain.

In summary, the research data on the outcome dimensions of schizophrenia in older patients provide some support for optimism in later life. However, based on recent longitudinal studies, we now have a more nuanced understanding of the dynamics of various domains. Thus, favorable outcome in one category does not necessarily result in favorable outcome in another category. This suggests that it may be more profitable to develop treatment strategies that are specific to each outcome category, although treating one category may have some impact on other categories. Because each of the outcome categories has been found to be associated with a variety of clinical and social variables, it is likely that effective treatments will require biologic and psychosocial methods. Jeste et al. have described an exaggerated "paradox of aging" among older adults with schizophrenia. These patients may demonstrate accelerated physical aging (increased and earlier medical comorbidity and mortality). However, their cognitive aging rate appears to stabilize following an initial, persistent, occurrence of mild neurocognitive disorder. Conversely, as these patients age, there is potential for improvement in psychosocial function, decrease in psychosis, relapse and hospitalization, and improvement in self-management and life quality.

Treatment

As we noted above, optimal treatment for older patients with schizophrenia will most likely involve a combination of pharmacology, nonpharmacologic therapy, and community-based approaches. Unfortunately, there are only a few studies assessing treatment approaches in older schizophrenia

patients. However, the limited available data point in the direction of a multimodal approach.

Pharmacotherapy

Evidence-based pharmacologic treatment of the older population with schizophrenia remains problematic given the absence of placebo-controlled, randomized, double-blinded clinical trials that specifically compare efficacy of various antipsychotic medications in this population. A Cochrane review of antipsychotic drugs for persons aged 65 and over with schizophrenia identified only three randomized control trials (RCTs) of 252 people. One RCT assessed drugs that are no longer available. The other two studies noted no differences between risperidone and olanzapine, and olanzapine and haloperidol. The use of antipsychotic agents in older adults has been complicated by recent federal warnings regarding increased incidence of cerebrovascular events and death with their use in older populations with dementia. Most clinicians have relied on evidence based on open-label trials and retrospective reviews that point toward the effectiveness of risperidone, clozapine, olanzapine, quetiapine, ziprasidone, and aripiprazole for use in older schizophrenia patients. No RCTs of the newer antipsychotic agents (paliperidone, iloperidone, asenapine, and lurasidone) have been done in elderly people with schizophrenia, and only asenapine was studied for short-term safety. One controlled study suggested that citalopram can be effective for middle-aged and older adults with schizophrenia and subsyndromal depression. An expert review panel written in 2004 recommended risperidone (1.25 to 3.5 mg/day) as the first line of treatment. Quetiapine (100 to 300 mg/day), olanzapine (7.5 to 15 mg/day), and aripiprazole (15 to 30 mg/day) were viewed as good second-line treatments.

The increased risk of extrapyramidal side effects and tardive dyskinesia (TD) in the elderly population on antipsychotics poses an additional challenge for clinicians. Despite low dosing, the cumulative 1-year incidence of TD is 29 percent among older patients (mean age 65 years) on first-generation antipsychotic agents, whereas in young adults, the annual cumulative incidence of TD is 4 to 5 percent. Atypical antipsychotics have substantially lower risk of TD than conventional neuroleptics, even in very high-risk patients such as middle-aged and older adults with borderline TD. However, the atypical antipsychotics have a less favorable metabolic side-effect profile than conventional antipsychotics. Expert reviews have recommended starting antipsychotic medication at 25 percent of the recommended adult dose and maintenance doses at 25 to 50 percent of the adult dose.

Finally, there is some evidence that two groups of older persons with schizophrenia may be candidates for discontinuing their medications: those who have been in sustained remission and those in whom

medications have been ineffective in modifying symptoms.

Suzuki and Uchida found nonrelapse rates of between 35 and 81 percent in 9 studies of middle-aged and elderly patients with schizophrenia whose medication had been stopped for 2 weeks to 18 months.

Nonpharmacologic Approaches

Use of nonpharmacologic treatment approaches to improve outcomes in older schizophrenia patients has received attention over the last decade. Several RCTs have shown favorable results for the use of cognitive remediation therapy for cognitive deficits, social skills therapy for difficulties with social interaction, and cognitive behavioral therapy for problems with social relatedness, psychosis, and mood. Older schizophrenia patients seem to respond well to behavioral learning techniques and can be taught a variety of instrumental roles and social skills, and these skills seem to provide long-term benefits and improved outcome. For example, a combined approach—cognitive-behavioral social skills training (CBSST)—was found to improve coping skills, social functioning, and facilitate a more objective understanding of subjective psychotic phenomena. More importantly, improvements in functioning as a result of CBSST were found to persist 1 year after conclusion of treatment.

Placing older schizophrenia patients into jobs and providing on-site training and support was found to improve rates of both paid and volunteer work. This approach was significantly more successful than the traditional "train-then-place" approach.

Community-Based Approaches

Assertive Community Treatment (ACT) and case management were both found to improve outcomes in older schizophrenia patients. There is also increasing use of self-help groups and voluntarism. Several model programs have combined collaborative care and case management strategies to confront the medical and psychiatric issues facing many patients as well as the need to facilitate communication between the mental health team and the primary care providers. Case management strategies include augmenting the patient's social and cognitive skills in conjunction with physical health–enhancing strategies.

LATE-ONSET SCHIZOPHRENIA

Epidemiology

The prevalence of late-onset schizophrenia and very-late–onset schizophrenia-like psychosis remains unclear. A review of studies of late-onset schizophrenia found that approximately 20 to 25 percent of patients with schizophrenia had an onset of the disorder after age 40. A Dutch case-register study reported a 1-year prevalence rate of 0.55 percent for

schizophrenia in persons over age 59 years, of which 64 percent had early-onset schizophrenia. These figures suggest that the loss of early-onset individuals (to death or recovery) may at least partially be compensated for by the influx of cases with a later onset. Epidemiologic study of late-onset schizophrenia is complicated by a number of factors, including the use of various definitions of *late onset*, uncertainty in the determination of age of onset, and questions regarding a correct diagnosis. A review of studies of late-onset schizophrenia found that approximately 23 percent of patients with schizophrenia were reported to have experienced the onset of the disorder after age 40, with 13 percent in the fifth decade of life, 7 percent in the sixth decade, and 3 percent in later decades.

Among persons aged 45 to 64 years there is an incident rate of 12.6 per 100,000 per year for new-onset schizophrenia.

Rates of psychotic symptoms or paranoid ideation have been estimated as high as 14 percent in older nondemented community dwelling adults, with rates among urban African Americans being twice that of Caucasians. A study of "very-old" nondemented individuals (aged 85 and 95 years) living in the community or institutions in Sweden found the prevalence of any psychotic symptoms in 85 year olds was 10.1 percent (hallucinations, 6.9 percent; delusions, 5.5 percent) and prevalence in 95 year olds was 7.1 percent (hallucinations, 6.7 percent; delusions, 10.4 percent).

Late-onset schizophrenia affects women between 2 and 10 times more often than men. The robustness of this finding, coupled with the higher incidence rates of early-onset schizophrenia in men, have led to the hypothesis that estrogen-mediated dopaminergic inhibition may protect younger women from schizophrenia, whereas the relative estrogen deficiency associated with menopause may be related to the increased incidence of late-onset schizophrenia in women. However, the incidence of schizophrenia is still considerably higher in younger versus older women, and treatment interventions based on the estrogen hypothesis have been largely unsuccessful. An alternative hypothesis is that older men may be protected in some way.

Etiology

The precise etiology of late-onset schizophrenia (onset age 45 to 60) and very-late–onset schizophrenia-like psychosis (onset over 60) is largely unknown. The International Late-Onset Schizophrenia Group concluded that schizophrenia with onset in middle age is a neurodevelopmental disorder, and that differences with early-onset schizophrenia are more of degree than of kind. The International Late-Onset Group also postulated that "Very-Late–Onset Schizophrenia-Like Psychosis" had some features that suggested a neurodegenerative component including more brain abnormalities and neuropsychologic deficits. A potential problem with this tripartite classification system is its reliance on age rather than clinical

presentation, and some persons over 60 can resemble earlier-onset cases.

Although some studies found that late-onset schizophrenia was less frequently associated with a family history of schizophrenia than the early-onset form, a more recent study reported that the family history of late-onset schizophrenia appeared to be similar to that of early-onset schizophrenia. Familial association of schizophrenia appears to be more common in early-onset and late-onset schizophrenia than in very-late–onset schizophrenia-like psychosis. Late-onset patients function somewhat better than do early-onset schizophrenia individuals during adolescence and adulthood. Late-onset schizophrenia patients seem to have a similar pattern, but lower severity, of cognitive impairments when compared to early-onset schizophrenia patients. Late-onset schizophrenia patients perform significantly worse than do healthy comparison subjects on measures of executive functioning, learning, motor skills, and verbal memory. Making a diagnosis of late-onset schizophrenia can be challenging. The onset of hallucinations and delusions in later life can arise de novo or be associated with mood disorders, sensory deficits, polypharmacy, substance abuse, medical disorders, or dementia.

The role of social and environmental stressors may also have a role in development of schizophrenia and schizophrenia-like psychosis after age 40. A relationship between trauma and onset of schizophrenia in childhood and early adulthood has been established previously in the literature. A relationship between history of trauma and onset of schizophrenia may also exist. A retrospective study that reviewed records of child survivors of the holocaust found that development of schizophrenia after age 40 was associated with higher degree of persecution in concentration camps before age 15. While this study has limitations and the hypothesis needs to be investigated further, the findings were remarkable for the strength of association between late-onset schizophrenia and degree of persecution. Studies in the United Kingdom found that the risk of developing very late-onset schizophrenia like psychosis is greater among older migrant populations than British-born elderly residing in the United Kingdom. In these studies, differences in the background age–gender structure of these populations were insufficient to explain this distribution. Like the study of holocaust survivors, the authors of the British studies have speculated that adverse psychosocial events associated with migration may precipitate schizophrenia in later years in vulnerable individuals.

The role of cognitive deficits in late-onset disorder is not fully understood. A sizeable proportion of patients with very-late–onset-type schizophrenia may go on to develop a type of dementia. Whether this proclivity to develop dementia is a characteristic of this subtype or suggests initial misdiagnosis is unclear. Brichant-Petitjean et al. observed that people with late-onset disorder occupy an intermediate position between those with early-onset type and healthy controls with respect to cognitive

deficits.

No specific biologic markers of schizophrenia with a later onset have been identified.

Hahn et al.'s review of neuroimaging studies that compared early-onset and late-onset schizophrenia reported differences in eight of ten studies, but across studies these were not consistent. Similarly, measures of event-related brain potentials (N400) and neurofibrillary tangle densities found no differences between early-onset and late-onset groups. However, a study reported in 2014, found raised C-reactive protein concentrations in both people with late-onset and very-late–onset schizophrenia. These findings need to be replicated. There also may be some differences with normal controls. For example, MRI studies found that patients with late-onset schizophrenia had larger right temporal horns and larger third ventricles compared to healthy controls, similar to the changes seen in younger patients with schizophrenia. In a study comparing the electroencephalograms of late-onset individuals to age and gender-matched controls, the only significant difference was a generalized slowing in the electroencephalograms of the late-onset schizophrenia patients, which might have been partially accounted for by the effects of antipsychotic medications. The investigators concluded that there did not appear to be an underlying dementia in late-onset schizophrenia.

Few data exist specifically examining functioning in late-onset schizophrenia, although examinations of middle-aged and older schizophrenia patients, of whom some were late-onset patients, have been conducted. In a group of middle-aged and elderly patients with schizophrenia or schizoaffective disorder, patients demonstrated significantly greater disability in all areas of social functioning compared to healthy controls. Although marriage rates have been shown to be higher in persons with late-onset schizophrenia compared to early-onset schizophrenia patients, social isolation was greater when comparing late-onset patients to the general population. As compared with healthy subjects, middle-aged and older schizophrenia outpatients were less likely to be married and have children, and more likely to live alone, and they had fewer friendships. However, older schizophrenia patients did not appear to be estranged from family members.

Although there were no differences between early and late-onset groups in positive symptoms, family history, brain abnormalities, memory retention, or minor physical abnormalities, the late-onset group was more likely to be female versus no gender differences in early-onset, have the paranoid subtype of schizophrenia, have lower levels of negative symptoms, have less impairment in learning, abstraction, and flexibility, better premorbid functioning with respect to work and marriage, and require lower daily dosages of antipsychotic medications. These differences suggest that late-onset schizophrenia may be a neurobiologically distinct form of

schizophrenia.

"Very-Late–Onset Schizophrenia-Like Psychosis" is distinguished from the other two types by more brain abnormalities and neuropsychologic deficits; more females; greater prevalence of persecutory and partition delusions; higher rates of visual, tactile, and olfactory hallucinations; lower genetic load; more sensory abnormalities; and absence of negative symptoms or formal thought disorder.

Diagnosis and Clinical Features

In terms of symptoms, early- and late-onset patients are more similar than different. The essential features of schizophrenia are a mixture of characteristic positive and negative symptoms associated with cognitive impairment and marked social or occupational dysfunction. The most common features of late-onset schizophrenia are persecutory delusions that may be bizarre and auditory hallucinations. The severity of positive symptoms appears to be similar to that in early-onset schizophrenia. A comparative study of early-onset and late-onset schizophrenic patients found no difference between the two groups in the prevalence of delusions of reference, bizarre delusions, or lack of insight. Patients with the late-onset form, however, tended to have more persecutory delusions with and without hallucinations, organized delusions, and abusive auditory hallucinations or hallucinations with a running commentary. The most common symptoms in the late-onset group were delusions, and approximately two-thirds had bizarre delusions. In general, late-onset illness is different from early-onset schizophrenia in that the late-onset type is associated with less severe negative symptoms, lower daily doses of antipsychotics needed for treatment, and less frequently observed loose associations and inappropriate affect. Consistent with these findings, late-onset schizophrenia patients typically meet criteria for the paranoid or undifferentiated subtype of schizophrenia.

In clinical samples, very old age at onset is related to a low prevalence of thought disorder and affective blunting and a higher prevalence of visual hallucinations. Variations in symptom profiles with onset at different ages do not necessarily imply differences in pathophysiology or etiology. Although similar symptoms could arise from different etiologic or pathologic processes, symptom differences could be related to cohort differences or age-associated central nervous system (CNS) differences that are independent of the illness.

Differential Diagnosis

The principal disorders in the differential diagnosis of late-onset schizophrenia include earlier-onset schizophrenia, mood disorders with psychotic features and schizoaffective disorder, delusional disorder,

psychosis associated with cognitive disorders, psychotic disorder due to a general medical condition, and substance-induced psychotic disorder.

Early-Onset Schizophrenia

Determining the onset of schizophrenia can be challenging, as many schizophrenic patients have an insidious onset of symptoms or premorbid personality traits that are difficult to differentiate from prodromal symptoms. Late-onset patients tend to be similar to earlier-onset schizophrenia of paranoid type in symptoms. Earlier onset and thus longer duration of schizophrenia are generally associated with more severe negative and disorganized symptoms and greater cognitive deficits (especially in the areas of learning and abstraction). In addition, late-onset schizophrenia is characterized by less severe symptoms with somewhat better preservation of affect and social functioning.

Mood Disorder

Differentiating schizophrenia from mood disorder with psychotic features and schizoaffective disorder is complicated in that mood disturbances are common during the prodromal, active, and residual phases of schizophrenia. As described earlier, prevalence rates of co-occurring clinical depression have been found in up to three quarters of older persons with schizophrenia, although mood symptoms may not meet criteria for major depression. Furthermore, mood-congruent or mood-incongruent psychotic features may occur in major depression or bipolar disorder, although in both disorders psychotic features tend to be more mood-congruent, for example, grandiosity erotomania, special powers among bipolar individuals and delusions of poverty, guilt, sin, or physical illness among depressed persons. Auditory hallucinations occur less commonly in depression and may be vague or condemnatory. Although uncommon, catatonic features are more commonly seen in severe depressive disorders, a history of mood symptoms may help to discriminate between conditions based on DSM-IV-TR criteria. Mood disorder with psychotic features beginning after 45 years of age is differentiated from late-onset schizophrenia by its affective symptoms that precede psychotic symptoms.

In schizoaffective disorder, there must be a mood episode that occurs simultaneously with the active-phase symptoms of schizophrenia and they must be present for a substantial proportion of the active and residual periods of the illness. Additionally, delusions or hallucinations must be present for at least 2 weeks in the absence of prominent mood symptoms. If the mood symptoms have a duration that is brief in relation to the total duration of the disturbance, occurs only during the prodromal or residual phases, or do not meet full criteria for a major depressive disorder, manic disorder, or mixed episode, then a diagnosis of schizoaffective disorder

cannot be made. Patients with remissions between schizophrenic episodes may be diagnosed as schizoaffective on one occasion and schizophrenia with mood disturbances on another occasion depending on levels of symptom severity and duration.

Delusional Disorder

Delusional disorder is differentiated from schizophrenia by a lack of prominent auditory or visual hallucinations and an absence of deterioration in areas of functioning outside the delusional scope. The delusions are necessarily nonbizarre and involve situations that may occur in real life, such as theft, suffering from a disease, or spousal infidelity. Delusional disorder may be difficult to differentiate from the paranoid type of schizophrenia because the latter is not characterized by prominent disorganized speech, disorganized behavior, or flat or inappropriate affect, and is often associated with less decline in functioning than in some of the other subtypes of schizophrenia. Nonetheless, the presence of auditory hallucinations or bizarre delusions seen in schizophrenia should differentiate the disorders. Furthermore, poor psychosocial functioning in delusional disorder is directly related to the delusional beliefs themselves.

Psychosis Associated with Cognitive Disorders or due to a General Medical Condition. A patient who presents with psychotic symptoms after the age of 45, particularly if the symptoms develop after the age of 60, requires a thorough evaluation for underlying medical disorders. Dementia is a common disorder in elderly persons, and approximately half of patients with a diagnosis of dementia of the Alzheimer type may present with psychotic symptoms at some point during the course of their dementia. Delusions in Alzheimer disease tend to be nonbizarre, episodic, and preceded by cognitive decline. Typically, delusions in Alzheimer disease involve someone stealing from them or entering their house. Visual hallucinations are more common in Alzheimer disease than in schizophrenia, and auditory hallucinations may be slightly less common. Other types of dementia, such as alcohol-induced persisting dementia and vascular dementia, also need to be ruled out. Psychosis due to a general medical condition is diagnosed from the history, physical examination, or laboratory tests that show the delusions or hallucinations are the direct physiologic result of a general medical condition. Metabolic disturbances, neurologic and endocrinologic disorders are the most common causes of psychotic symptoms. Misuse or abuse of prescription and over-the-counter medications must also be considered.

Substance-Induced Psychotic Disorder. The prevalence rates of substance use disorders are estimated at 2 to 3 percent in older woman and 10 percent in older men. These figures are expected to increase as the baby

boomers reach old age. A number of different substances, such as amphetamines, cocaine, and phencyclidine, can produce symptoms similar to those of schizophrenia, such as delusions, hallucinations, or a mixture of positive and negative symptoms. Many substances can induce psychotic symptoms during withdrawal states, for example benzodiazepines, barbiturates, and alcohol. Ideally, substance-induced psychotic disorders would be differentiated from schizophrenia after a period of abstinence; however, other evidence, such as whether the psychotic symptoms appear to be exacerbated by the substance and decrease when it has been discontinued or whether the psychotic symptoms have been instigated and maintained by the substance use, need to be considered.

Course and Prognosis

The course of late-onset schizophrenia is usually chronic but may be interrupted by partial remissions and exacerbations. Over the past decade, there has been substantial change in our understanding of prognosis and outcomes in schizophrenia in late life. The prognosis is somewhat better than that in early-onset schizophrenia. Patients may be quite responsive to antipsychotics used in lower doses. However, mortality (especially from suicide) exceeds that in the general population and is probably comparable to that in early-onset schizophrenia.

Studies of cognitive outcome in late-onset schizophrenia have yielded mixed findings. A 2-year study compared early- and late-onset schizophrenia patients, normal controls, Alzheimer disease patients with psychosis, and Alzheimer disease patients with a Mini-Mental State Examination (MMSE) of at least 24 at baseline. This study suggested that late-onset schizophrenia patients have a rate of cognitive decline that is similar to early-onset patients and normal controls. However, another study which followed late-onset schizophrenia patients and normal controls over 5 years, demonstrated significant cognitive decline in the late-onset schizophrenia group. The course of late-onset schizophrenia therefore remains unclear.

Treatment

A Cochrane review of antipsychotic treatment for elderly people with late-onset schizophrenia reported only one small RCT comparing two drugs and the reviewers concluded that this study provided little usable data. Based on very limited data, late-onset schizophrenia patients appear to have good symptomatic improvement with antipsychotic treatment compared to early-onset illness individuals. Although appropriate therapy requires individualization, it appears that substantially reduced doses of antipsychotics are necessary in late-onset schizophrenia patients. Maintenance therapy is frequently required, because many of these

patients relapse if they discontinue the medication. Elderly patients are more susceptible to certain antipsychotic side effects, such as sedation, anticholinergic toxicity, and extrapyramidal symptoms. As mentioned in the previous section, only a small number of open-label and controlled prospective trials have been conducted to examine the efficacy and safety of antipsychotics in schizophrenia, with the majority of literature pertaining to conventional antipsychotics. Atypical antipsychotics have become the agents of choice for older adults with psychosis, owing to their improved side-effect profile compared to conventional agents. However, there is a lack of published data specifically examining the use of atypical antipsychotics in patients with late-onset schizophrenia, and existing investigations have frequently involved mixed samples of early- and late-onset patients. Thus, results must be inferred from trials of atypical antipsychotics conducted in young adults with schizophrenia.

Psychosocial treatment is a vital aspect of the management of late-onset schizophrenia, and treatments specifically designed for older adults with schizophrenia, such as cognitive behavioral therapy and social skills training also have shown promise.

DELUSIONAL DISORDER

Epidemiology

Studies suggest that between 2 and 8 percent of the elderly population experiences some type of paranoid symptoms. One study indicated that although approximately 2 percent of elderly psychiatric patients met DSM-III criteria for paranoid disorder, another 13 percent had paranoid ideation. DSM-IV-TR estimates the population prevalence of delusional disorder to be approximately 0.03 percent. Delusional disorder usually first appears in middle to late adulthood and the sociodemographic profile is similar across cultures. The overall prevalence of delusional disorder is slightly higher among women than among men, and the average age of onset is earlier for men (40 to 49 years) than for women (60 to 69 years). Owing to its later age at onset, the lifetime morbidity risk is estimated by DSM-IV-TR to range between 0.05 and 0.1 percent.

Etiology

Several factors have been postulated to contribute to the development of delusional disorder: a family history of schizophrenia; avoidant, paranoid, or schizoid personality disorder; hearing loss; or being an immigrant or of low socioeconomic status. The available evidence implicating each of these factors needs confirmation. From the viewpoint of psychological defense mechanisms, delusions are generally attributed to an excess and inappropriate use of projection. The patient tends to hold others responsible for any biopsychosocial losses that may accompany normal

aging. Studies aimed at establishing a genetic component to delusional disorder have been unable to either confirm or establish such a component. This is because all studies to date have been limited by unreliable sampling (owing to the subjective nature of the diagnosis of delusional disorder). Studies that have had relatively sound inclusion criteria have been limited by small sample sizes.

There are no clear neuroanatomical changes associated with delusional disorder. There is mixed evidence that hearing or visual abnormalities might play a role in the development of delusional disorder, with Maher observing that a subset of patients develops delusions in the context of sensory impairment (Reinhardt paper 116).

Family, twin, and adoption studies have suggested that a close association between delusional disorder and schizophrenia is unlikely. Until robust molecular data are available, a genetic role in etiology of delusional disorder or a common link between delusional disorder and schizophrenia cannot be fully appreciated.

Diagnosis and Clinical Features

Persecutory, somatic, or erotomanic delusions in the elderly patient are usually secondary to another neuropsychiatric disorder (e.g., schizophrenia, mood disorder, or dementia). Occasionally, however, a primary delusional disorder is implicated. A clearer understanding and study of this disorder has been held up by inconsistencies in nomenclature and diagnostic criteria.

The essential feature of delusional disorder is one or more persistent nonbizarre delusions that have been present for at least 1 month. Diagnostic criteria for schizophrenia or schizoaffective disorder must not be met. Auditory or visual hallucinations, if present, are not prominent. Psychosocial functioning, although variable, generally is not considerably impaired outside of the delusional realm, and behavior is not noticeably odd or bizarre. The DSM-5 describes the following subtypes of delusional disorder based on the predominant delusional theme: erotomanic, grandiose, jealous, persecutory, somatic, mixed, and unspecified.

Common types of delusions seen in elderly patients with delusional disorder are skin infestation by insects; organs, such as a liver, are cancerous (despite a lack of evidence); the person is being spied on, followed, or poisoned; and the spouse is unfaithful (without factual support). Social, marital, or work problems can result from the delusional beliefs of delusional disorder. Ideas of reference are common, and understanding of such ideas is usually consistent with the content of the patient's delusional beliefs. Anger, irritability, and dysphoric mood can occur, especially in the persecutory and jealous types of delusional disorder. It is thought that major depressive episodes likely occur more frequently in persons with delusional disorder than in the general

population. In addition, individuals with delusional disorder may have high lifetime comorbidity with affective disorders, such as obsessive-compulsive disorder (OCD) and body dysmorphic disorder.

Differential Diagnosis

The considerations in the differential diagnosis of delusional disorder are similar to those for schizophrenia with onset in late life. Delusional disorder can be differentiated from schizophrenia by the absence of other characteristic symptoms of the active phase of schizophrenia, such as prominent auditory or visual hallucinations, bizarre delusions, and negative symptoms. Although mood disorders with psychotic features commonly involve nonbizarre delusions without prominent hallucinations, the temporal relationship between the mood disturbance and the delusions is an important consideration when differentiating the disorders. Somatic delusions and hypochondriasis are sometimes difficult to distinguish; however, in hypochondriasis, the fears of having a serious disease are held with less intensity than in delusions. It is also difficult to distinguish between delusional disorder and OCD. Obsessions may develop into sustained delusional beliefs that may make an additional diagnosis of delusional disorder appropriate. Organic etiology such as Wilson disease and stroke must be ruled out prior to diagnosing delusional disorder.

Course and Prognosis

The course of delusional disorder may vary but tends to be chronic, especially in those with the persecutory type of delusional disorder. A waxing and waning of the preoccupation with the delusional beliefs often occurs, and in other cases, full periods of remission may be followed by subsequent relapses. In some cases, the disorder remits within a few months, often without subsequent relapse. Some evidence suggests that the jealous type may have a better prognosis than the persecutory type. Generally, intellect and occupational activities are maintained, but marital and social functioning is compromised. The data on prognosis are limited, although the likelihood of complete recovery seems low.

Treatment

There are no available studies on the treatment of late-life delusional disorder. Antipsychotic agents often are effective, especially in agitated delusional patients, but they must be administered cautiously in elderly persons. Expert consensus guidelines recommend the use of atypical antipsychotic medications over conventional agents as follows: First line: risperidone 0.75 to 2.5 mg/day; second-line: olanzapine 5 to 10 mg/day or quetiapine 50 to 200 mg/day. Commonly, persons deny their illness. Consequently, nonadherence to antipsychotics is a common problem in the

treatment of patients with delusional disorder. For refractory cases, modified electroconvulsive therapy has been reported as successful. In addition to pharmacologic therapies, psychosocial treatments represent an important therapeutic option. Cognitive-behavior therapy has been demonstrated as an effective treatment approach.

REFERENCES

Abdallah C, Cohen CI, Sanchez-Almira M, Reyes P, Ramirez P. Community integration and associated factors among older adults with schizophrenia. *Psychiatr Serv*. 2009;60(12):1642–1648.

Almeida OP, Hankey GJ, Yeap BB, Golledge J, Norman PE, Flicker L. Mortality among people with severe mental disorders who reach old age: a longitudinal study of a community-representative sample of 37892 men. *PLoS ONE*. 2014;9(10):e111882.

Arnold SE. Contributions of neuropathology to understanding schizophrenia in late life. *Har Rev Psychiatry*. 2001;9(2):69–76.

*Auslander LA, Jeste DV. Sustained remission of schizophrenia among community-dwelling older outpatients. *Am J Psychiatry*. 2004;161(8):1490–1493.

Bleuler M. Late schizophrenic clinical pictures. *Fortschr Neurol Psychiatr*. 1943;15:259–290.

Brichant-Petitjean C, Legauffre C, Ramoz N, Ades J, Gorwood P, Dubertret C. Memory deficits in late-onset schizophrenia. *Schizophr Res*. 2013;151(1–3):85–90.

Brodaty H, Sachdev P, Koschera A, Monk D, Cullen B. Long-term outcome of late-onset schizophrenia: 5-year follow-up study. *Br J Psychiatry*. 2003;183(3):213–219.

Castle DJ, Wessely S, Howard R, Murray RM. Schizophrenia with onset at the extremes of adult life. *Int J Geriat Psychiatry*. 1997;12:712–717.

Cervantes AN, Rabins PV, Slavney PR. Onset of schizophrenia at age 100. *Psychosomatics*. 2006;47(4):356–359.

Ciompi L. Catamnestic long-term study on the course of life and aging of schizophrenics. *Schizophr Bull*. 1980;6(4):606–618.

Cohen CI, Cohen GD, Blank K, et al. Schizophrenia and older adults: an overview: directions for research and policy. *Am J Geriatr Psychiatry*. 2000;8(1):19–28.

*Cohen CI, Iqbal M. Longitudinal study of remission among older adults with schizophrenia spectrum disorder. *Am J Geriatr Psychiatry*. 2014;22(5):450–458.

Cohen CI, Natarajan N, Araujo M, Solanki D. Prevalence of negative symptoms and associated factors in older adults with schizophrenia spectrum disorder. *Am J Geriatr Psychiatry*. 2013;21(2):100–107.

Copeland JR, Dewey ME, Scott A, et al. Schizophrenia and delusional disorder in older age: community prevalence, incidence, comorbidity and outcome. *Schizophr Bull*. 1998;24(1):153–161.

Corey-Bloom J, Jernigan T, Archibald S, Harris MJ, Jeste DV. Quantitative magnetic resonance imaging of the brain in late-life schizophrenia. *Am J Psychiatry*. 1995;152:447–449.

Desai AK, Grossberg GT. Differential diagnosis of psychotic disorders in the elderly. *Schizophrenia into later life: Treatment, research, and policy*. 2003:19–41.

Granholm E, Holden J, Link PC, McQuaid JR, Jeste DV. Randomized controlled trial of cognitive

behavioral social skills training for older consumers with schizophrenia: defeatist performance attitudes and functional outcome. *Am J Geriatr Psychiatry*. 2013;21(3):251–262.

*Howard R, Rabins PV, Seeman MV, Jeste DV; The International Late-Onset Schizophrenia Group. Late-onset schizophrenia and very-late-onset schizophrenia-like psychosis: an international consensus. *Am J Psychiatry*. 2000;157(2):172–178.

Ibrahim F, Cohen CI, Ramirez PM. Successful aging in older adults with schizophrenia: prevalence and associated factors. *Am J Geriatr Psychiatry*. 2010;18(10):879–886.

Jeste DV, Harris MJ, Krull A, Kuck J, McAdams LA, Heaton R. Clinical and neuropsychological characteristics of patients with late-onset schizophrenia. *Am J Psychiatry*. 1995;152(5):722–730.

Jeste DV, McAdams LA, Palmer BW, et al. Relationship of neuropsychological and MRI measures with age of onset of schizophrenia. *Acta Psychiatr Scand*. 1998;98(2):156–164.

Kredentser MS, Martens PJ, Chochinov HM, Prior HJ. Cause and rate of death in people with schizophrenia across the lifespan: a population-based study in Manitoba, Canada. *J Clin Psychiatry*. 2014;75(2):154–161.

Loewenstein DA, Czaja SJ, Bowie CR, Harvey PD. Age-associated differences in cognitive performance in older patients with schizophrenia: a comparison with healthy older adults. *Am J Geriatr Psychiatry*. 2012;20(1):29–40.

Meesters PD, Comijs HC, de Haan L, et al. Symptomatic remission and associated factors in a catchment area based population of older patients with schizophrenia. *Schizophr Res*. 2011;126(1–3):237–244.

Morimoto K, Miyatake R, Nakamura M, Watanbe T, Hirao T, Suwaki H. Delusional disorder: molecular genetic evidence for dopamine psychosis. *Neuropsychopharmacology*. 2002;26(6):794–801.

Ostling S, Skoog I. Psychotic symptoms and paranoid ideation in a nondemented population-based sample of the very old. *Arch Gen Psychiatry*. 2002;59(1):53–59.

Rajji TK, Voineskos AN, Butters MA, et al. Cognitive performance of individuals with schizophrenia across seven decades: a study using the MATRICS consensus cognitive battery. *Am J Geriatr Psychiatry*. 2013;21(2):108–118.

*Reinhardt MM, Cohen CI. Late-life psychosis: diagnosis and treatment. *Curr Psychiatry Rep*. 2015;17(2):1.

Salzman C, Tune L. Neuroleptic treatment of late-life schizophrenia. *Harv Rev Psychiatry*. 2001;9:77–83.

Thompson WK, Savla GN, Vahia IV, et al. Characterizing trajectories of cognitive functioning in older adults with schizophrenia: does method matter? *Schizophr Res*. 2013;143(1):90–96.

*Vahia IV, Palmer BW, Depp CA, et al. Is late-onset schizophrenia a subtype of schizophrenia? *Acta Psychiatr Scand*. 2010;122(5):414–426.

57.3i Personality Disorders

MARC E. AGRONIN, M.D.

Multiple normative and pathologic changes during the aging process affect personality, and these same forces apply to the consideration of personality

disorders in late life. These disorders represent a heterogeneous array of chronic affective and behavioral disturbances that cause enormous personal and interpersonal distress and pose significant challenges to clinical management, especially in the setting of comorbid medical and psychiatric conditions. Changing and inherently limited diagnostic schemes combined with a general lack of previous study make it difficult to understand and plot the trajectory of a given personality disorder in late life. Clinical perspectives on personality stability versus change will vary depending on the methods of assessment. There is limited empirical evidence and a range of unsubstantiated opinions as to whether personality disorders persist with minimal or no change, convert into more age-appropriate forms with similar levels of pathology, or remit over time. Understanding all of these factors is critical to improving diagnosis and optimizing clinical management.

DEFINITIONS

Defining the concept of *personality* itself is key to a fuller understanding of personality disorders. According to the *Bantam Medical Dictionary,* personality is "an enduring disposition to act and feel in particular ways that differentiate one individual from another." George Vaillant qualifies this definition, suggesting that the enduring aspects of personality, which he terms *temperament,* consist largely of various inherited elements such as personality traits and IQ. He bases this belief on longitudinal studies of core personality traits (e.g., extroversion, introversion) that have generally found stability with aging. On the flip side, Vaillant holds that elements of personality or *character* defined by adaptive styles can change significantly with aging, affected by both "environment and maturation." Putting together both elements of stability and change, Henry Murray's statement that "the history of the personality is the personality" is most helpful when considering aged individuals. Lacking accurate longitudinal data from an older individual's younger days, the average clinician is forced to make decisions about personality functioning based on more recent history.

The distinction between normal and abnormal personalities is based largely on the expression of certain personality *traits* (as opposed to more acute psychopathologic symptoms) that are repeatedly associated with unhappiness, interpersonal conflict, and/or occupational dysfunction. Although everyone has a unique personality with relative strengths and weaknesses, the current diagnostic model requires the categorization of certain personalities as pathologic along the lines of various subtypes. This categorical model has been criticized for the ways in which it provides pejorative labels and obscures the richness of personality. In late life, it is particularly problematic given the large number of physical and psychosocial stresses that can distort an individual's pattern of adaptation.

An alternative dimensional approach was proposed but not adopted in

the *Diagnostic and Statistical Manual of Mental Disorders,* fifth edition (DSM-5). Personality disorders are defined in DSM-5 as chronic psychiatric disorders characterized by enduring and pervasive patterns of inflexible and maladaptive inner experiences and behaviors. These dysfunctional features are rooted in early development, deviate from cultural expectations, and lead to significant social and occupational impairment across several spheres of function: perception and interpretation of experiences, emotional expression, interpersonal relations, and impulse control.

There are ten subtypes of personality disorders listed in the DSM-5 classification, grouped into three broad phenomenologic clusters: cluster A, the *odd* cluster, includes paranoid, schizoid, and schizotypal subtypes; cluster B, the *dramatic* cluster, includes antisocial, borderline, histrionic, and narcissistic subtypes; and cluster C, the *anxious* cluster, includes avoidant, dependent, and obsessive-compulsive subtypes. Depressive and passive-aggressive (or negativistic) personality disorders were two subtypes that were listed in the appendix of the *Diagnostic and Statistical Manual of Mental Disorders,* fourth edition text revision or DSM-IV-TR but are no longer included in DSM-5. Each personality subtype is defined by 8 to 10 diagnostic criteria, and an individual must meet a specified number of them in any combination in order to qualify for a diagnosis. Individuals may demonstrate a single disorder or a combination of maladaptive traits from several disorders. Individuals who have symptoms characteristic of a personality disorder in general, but do not meet criteria for any specific disorder, may be diagnosed as having an *unspecified personality disorder*.

Although there are several designated pediatric categories for personality disorders, there are no equivalent, age-adjusted geriatric subtypes, despite growing evidence that changes in both personality and personality disorders in late life can lead to enduring but *different* patterns of personality dysfunction. However, when personality change results from physical damage to the brain from substance abuse or a medical condition, the term *personality change due to another medical condition* can be used (with the medical condition specified and then coded separately), with several designated subtypes: labile, disinhibited, aggressive, apathetic, and paranoid. Common causes of personality change include traumatic head injury, stroke, endocrine dysfunction, encephalitis, and Alzheimer disease.

EPIDEMIOLOGY

Community-based epidemiologic studies have found rates of personality disorders in adults ranging from 10 to 20 percent, with one large sample of over 43,000 individuals in the United States yielding a prevalence rate of 14.8 percent. Epidemiologic studies focusing only on individuals 65 years or older have found a slightly lower prevalence range between 5 and 13 percent. These rates increase two- to threefold in individuals being treated

in both outpatient and inpatient psychiatric settings for comorbid psychiatric diagnoses, particularly major depression. For example, a study of 546 elderly psychiatric inpatient admissions found personality disorders in 13 percent of all patients, which increased to 24 percent of those diagnosed with major depression. In all studies, the method of diagnosis has produced variations in prevalence rates. Epidemiologic surveys and more impressionistic clinical diagnoses have generally found higher rates of personality disorders and yet a more limited spectrum of subtypes compared to criterion-based structured clinical interviews.

The most common personality disorders seen across late-life studies have been paranoid, avoidant, dependent, and obsessive-compulsive subtypes. In late life, nearly two-thirds of those diagnosed with a personality disorder also meet criteria for two or more subtypes, or for the unspecified category. It is important to bear in mind that the lifelong course of specific subtypes will influence prevalence. For example, a decrease in the prevalence of antisocial personality disorder in late life no doubt results from increased rates of remission, death, institutionalization, and avoidance of psychiatric settings with age. Decreasing rates of other personality disorders in late life may be due to the fact that they are difficult to differentiate from other specific psychiatric diagnoses, or may change in ways that obscure their recognition.

There are limited data on the prevalence of personality disorders in long-term care settings, due to the great difficulty obtaining longitudinal data in addition to the challenge of interviewing individuals who may be suffering from varying degrees of cognitive impairment. One survey that based diagnoses on clinical interviews found that nearly 11 percent of a modest sample of nursing home residents had a primary or secondary diagnosis of personality disorder. An identical diagnostic rate was found in a sample of intellectually disabled adults older than 40 years who were admitted to a long-term care institution—although in that study, the congenital diagnosis might call into question the personality disorder diagnoses. A review of referrals to a nursing home found that 15 percent were for unspecified personality problems, although it is not clear how many of these individuals had actual personality disorders. Rates for personality disorder subtypes in long-term care facilities have not been established in any study.

DIAGNOSIS

In spite of the very specific diagnostic criteria provided by DSM-5, clinicians immediately encounter a plethora of diagnostic challenges when assessing older individuals. Sometimes the older personality, viewed from an ageist stance, is erroneously believed to be more dependent, stubborn, and hypochondriacal and less adaptive in the face of stress, compared to younger personalities. Such an attitude may lead clinicians to view many

disordered behaviors as normal in late life, preventing accurate diagnosis. At the same time, many clinicians might not consider the possibility of a personality disorder in an older patient, or might be daunted by the large number of diagnostic criteria. As a result, these diagnoses are often deferred or merely lumped under an unspecified category. Lack of knowledge either of specific criteria or how to apply them in late life has been shown to create discrepancies between largely impressionistic clinical diagnoses and those determined through more rigorous structured interviews.

The criteria themselves foster misdiagnosis since they were developed for adults, and do not take into consideration age-associated changes in lifestyle and behaviors that render many of them obsolete in later life. For example, the diagnosis of antisocial personality disorder includes criteria such as "repeatedly performing acts that are grounds for arrest," "repeated physical fights or assaults," and "consistent irresponsibility" with work or financial obligations. For an older person who is retired from work, residing in congregate living or a long-term care facility with limited mobility, or physically frail, these criteria will simply not apply. Diagnostic criteria for borderline personality disorder such as self-injurious behaviors and self-damaging impulsivity with respect to spending, sex, driving, or eating are also less applicable to older adults, who in general tend to be less impulsive. As of yet, there are no geriatric criteria that can be substituted, and attempts by clinicians to make diagnoses based on fewer criteria, or to "age-adjust" other criteria, threaten the very validity of the diagnosis.

Not surprisingly, a recent analysis of data from the National Epidemiologic Survey on Alcohol and Related Conditions, which studied over 43,000 individuals, found that young and old adults endorsed nearly 30 percent of personality disorder criteria differently, despite having equal degrees of personality dysfunction. The end result was that older adults were more likely to have schizoid and obsessive-compulsive personality disorder diagnoses, while younger adults were more likely to have avoidant and dependent personality disorder diagnoses. The same diagnostic criteria, then, yielded different disorders based on age.

Another diagnostic dilemma is the need for accurate longitudinal descriptions of a personality on which to base the diagnosis. Unlike symptoms of most other psychiatric disorders, which are readily reported or observed, personality disorder symptoms can be more difficult to elicit through self-report given their pejorative nature, and when observed require repeated encounters across different settings to be accurate and reliable. Instead, clinicians often have to rely on clinical records, which typically lack longitudinal data, or on existing informants who can rarely provide critical information dating back to an older patient's young adulthood.

An important caveat for any clinician attempting to make a late-life

diagnosis is that personality dysfunction does not necessarily indicate a personality disorder. Significant, repeated, and sometimes devastating losses in late life, such as the death of a spouse or of a series of close friends, a chronic debilitating illness, chronic pain, physical disability, the loss of finances or livelihood, or the overall loss of personal attributes that once defined an individual (e.g., strength, beauty, athletic, or sexual prowess), can lead to severe adjustment reactions that mimic personality disorders. Without clear longitudinal information, it is difficult to determine whether a patient with pervasive dysfunctional interpersonal relations, for example, is demonstrating a lifelong pattern or a more recent reaction to life events. Many symptoms of major depression and generalized anxiety disorder overlap with depressive, negativistic, avoidant, and depressive symptoms of several personality disorders. Similarly, symptoms of persistent bipolar mania can mimic antisocial, narcissistic, and borderline traits. For this reason, personality disorder diagnoses are best deferred during episodes of acute psychiatric syndromes. Perhaps one of the most difficult decisions to make is whether changes in demeanor brought on by neurocognitive disorders and associated neuropsychiatric syndromes such as apathy are merely expressions of the underlying disorder or represent a true personality change.

CLINICAL FEATURES AND COURSE

Since its inception in 1952, iterative versions of the DSM have provided a heterogeneous and fluid array of personality disorder subtypes up until the more enduring categorization that appeared in the third edition of the DSM (DSM-III) in 1980. Despite the relative stability in the nomenclature since then, the DSM has never been clear on the course of personality disorders into late life, only suggesting that several remain relatively unchanged while others attenuate. The possibility of a late-onset personality disorder is not addressed in the DSM, aside from the diagnosis of a personality change disorder. Unfortunately, there are limited data on most subtypes, and current thinking rests largely on case studies.

For the clinician working with elderly individuals with known or suspected personality pathology, the main question is: What do these disorders look like in late life? Are they fundamentally the same as in previous years, or does aging change them in ways that either attenuate symptoms and cause remission or lead to similar degrees of pathology but with novel, age-appropriate manifestations? Consider the following clinical vignette.

Mr. L was an 82-year-old man who suffered from brittle diabetes, morbid obesity, and chronic obstructive lung disease. After his wife died, Mr. L was unable to manage his own care needs and had to be admitted to an assisted living facility (ALF). He was oxygen dependent and used an electric scooter to get around. Mr. L quickly established a reputation as a rude and aggressive man. He would badger other residents and staff in the elevator when they did not

get out of his way quickly. He would make unrealistic demands of the waiters in the dining room, and insult them when he did not get his way. He accused the nursing staff of deliberately withholding treatment after one of his medications went missing from his tray. Over time, staff grew increasingly tired of his verbally abusive and disruptive behaviors, and was afraid to help him since he often threatened to sue them for shoddy care. They described him as a difficult, paranoid, and hateful person.

These behaviors built to a crisis one day when Mr. L became enraged at a deaf resident standing in front of him in the elevator who could not hear him berating her to move out of his way. He pushed her down and ran over her foot with his scooter, causing a serious skin laceration and bruise. He completely denied his actions, however, accusing the injured woman of deliberately feigning injury after blocking his scooter and getting bumped. He screamed at all staff who tried to intervene, and threatened to kill anyone who got in his way. As a result, Mr. L was hospitalized involuntarily and not allowed to return to the ALF.

Additional information on Mr. L indicated that he had been an extremely aggressive but quite successful real estate broker as a young and middle-aged man. His daughter estranged herself from him due to his abusive behaviors, and his son limited his contact with his father whom he described as arrogant, paranoid, and verbally abusive to those who disagreed with him. Mr. L's behaviors were chronic but seemed to have improved in the several years prior to his becoming ill, perhaps because his business success enabled him to get involved in several gratifying philanthropic endeavors. However, with the onset of illness and then the death of his wife, he could not cope with being dependent on others for help. He was completely unable to trust staff or forge stable relationships in the ALF, and became increasingly angry, depressed, and paranoid. Having to adjust to institutional living proved impossible, and his brittle coping style was eventually overwhelmed, resulting in a crisis characterized by aggression and paranoid accusations. Putting together known history and current observations, Mr. L's clinical team diagnosed him with a paranoid and narcissistic personality disorder.

In this clinical vignette, several themes appear that characterize the course and prognosis of many personality disorders throughout life. First, dysfunctional behaviors do not arise de novo but have a history, and have pervaded the individual's life in one form or another since young adulthood. In many cases, there may be long periods of increased stability in middle age that represent symptomatic remission, typically bolstered by stable relationships, successful occupational pursuits, and intensive therapeutic assistance. Not all personality disorders, however, will benefit from these stabilizing factors—particularly those with avoidant, schizoid, and schizotypal features. As seen in the illustrative case, major late-life stresses can exacerbate existing pathologic traits or lead to reemergence of previous traits in ways that are different in kind but equally problematic as earlier in life. Lacking a longitudinal history and based strictly on observed behaviors, many personality-disordered individuals are typically diagnosed with major depression, bipolar disorder, or even late-onset schizophrenia given the intensity of their symptoms. Although these diagnoses may present comorbidly, they are driven in part by the underlying personality disorder.

Cluster A Disorders

In late life, the diagnostic criteria for the three odd cluster disorders coalesce around a lack of desire for and openness toward interpersonal relationships. The tendency of the paranoid personality disorder to read

malevolent intent into the actions of others leads to a lifetime of angry conflicts with others, while the aloof and strange demeanor of the schizoid and schizotypal individual precludes the formation of close relationships. As a result, odd cluster personality disorders tend to enter later life without the benefit of past interpersonal experiences that can, for less dysfunctional individuals, teach social skills and provide emotional support. It is believed that the pathologic cognitive and behavioral patterns of the odd cluster disorders tend to remain the same or even worsen with aging.

Paranoid, schizoid, and schizotypal individuals face age-associated stresses with variable responses. On the one hand, the lack of close relationships means that these individuals may feel less emotional impact from the inevitable loss of family members or longstanding social contacts. However, when medical illness and disability occur, these same individuals often have great difficulty seeking help and coping with forced dependency on others, especially when placed in long-term care institutions. The result is often the increased expression of longstanding dysfunctional behaviors and defense mechanisms, including paranoid ideation, social isolation, and odd or idiosyncratic beliefs or rituals. There may also be an increased risk of psychosis, to the extent that these disorders are considered potential precursors for late-onset schizophrenia.

Cluster B Disorders

Although lumped together under the "dramatic" cluster domain, these four personality disorders are heterogeneous in both expression and course over the life cycle. A central theme, however, is that while their psychopathology results in severe and, in some cases, deadly consequences, there is potential for symptomatic quiescence, improvement, or even remission in middle age when there are significant personal or occupational achievements or the presence of consistent interpersonal relationships. In addition, individuals with more personal, social, and/or financial resources and with less psychiatric comorbidity tend to fare better. By the same token, age-associated losses in relationships, social and occupational roles, physical attributes such as strength and beauty, and financial status can reverse many of these gains and prove devastating to the personality. For those individuals who had improved over their lifetimes, the stress of major losses can lead to the reemergence of previous dysfunctional cognitive and behavioral patterns, but in ways not always recognizable by the current diagnostic nomenclature. These relapses are inevitably complicated by comorbid psychiatric disorders such as major depression, neurocognitive disorders, anxiety disorders, and substance-related and addictive disorders.

There are several factors that limit the number of older individuals with antisocial personality disorder who are seen in clinical settings. First, their history of reckless and criminal behaviors and high rates of substance

abuse often result in either death or incarceration long before they reach middle age. Second, these individuals have poor insight into their problems, few supportive relationships, and lack of motivation to get help unless there is an acute issue that prevents their ongoing activities. As a result, both young and old antisocial individuals tend to avoid psychiatric settings.

Despite this rather grim assessment, clinical data reveal, somewhat surprisingly, that up to a third of antisocial individuals demonstrate significant clinical improvement if they survive into middle age, and about 12 to 33 percent no longer meet diagnostic criteria after 10 to 15 years. Individuals with poor social skills and higher degrees of aggression tend to do worse. The most salient changes include a shift away from aggressiveness, promiscuity, and criminal behaviors but replaced, in part, by substance abuse, depression, and somatization. All of these changes may reflect a general trend of decreased impulsivity and aggression with aging.

In long-term care settings, antisocial behaviors may include resistance to care, surreptitious substance abuse, abuse of other residents, and disregard for institutional rules regarding personal safety, hygiene, and schedules. Based strictly on observable behaviors, these older individuals might not meet criteria for antisocial personality disorder, even though a longstanding lack of conscience—the hallmark of the disorder—is driving more age-appropriate dysfunctional behaviors.

In a similar fashion, individuals with borderline personality disorder demonstrate the potential for significant improvement and even remission in middle age, with one longitudinal study of 100 individuals finding that 75 percent no longer met criteria after 15 years. This improvement was bolstered by personal and occupational successes and associated with increased physical attractiveness, higher intelligence, and self-discipline. Poorer outcomes were associated with a history of abuse, antisocial tendencies, high levels of aggression, and substance abuse. Even with improvement, however, there remained a significant risk for depression, substance abuse, and suicide.

Clinical case studies have described borderline individuals who demonstrate reemergent symptoms later in life, but with a different appearance. Earlier defining symptoms of self-mutilation, impulsive and risky behaviors, and poor self-identity are attenuated and replaced by mood lability and depression, nonadherence to treatment, overuse of prescription medications, and alcohol abuse. In long-term care settings, however, the erratic and rageful interpersonal style seen in younger individuals with borderline personality disorder is often still present, creating havoc on the milieu.

Over their lifetimes, individuals with narcissistic personality disorder often leave a wake of hurt and disgruntled personal and business relationships in their path, and though aware of their own anger or

estrangement from others, they rarely attribute it to their own behaviors. Their sense of entitlement and demands on others may bring success in middle age, but with time they are particularly vulnerable to age-associated losses of personal attributes, gratifying occupational and social roles, and supportive if not somewhat subservient relationships. It becomes increasingly difficult for them to replace lost or jettisoned relationships over time, and they may find themselves in great need due to illness or disability but without adequate social supports. The very process of aging proves to be a significant narcissistic injury, only worsened when they have to rely on the help of others they deem below them. In long-term care, the narcissistic personality disorder often presents as simultaneously depressed, rageful, and imperious toward staff, leading to frequent conflicts.

Histrionic personality disorder is probably the least studied in late life, and has not been reported much in epidemiologic surveys or in clinical case studies. Like the narcissistic individual, however, histrionic individuals often lack deep, supportive social relationships, and are equally buffeted by age-associated losses in physical appearance that allowed them to maintain some state of psychological equilibrium in the past. It has been suggested that somatization may be one dysfunctional avenue for older histrionic individuals to obtain attention and social gratification.

Cluster C Disorders

The anxious cluster disorders are thought to involve a lifetime of chronic personality dysfunction without significant remission, perhaps explaining why they tend to be some of the most commonly seen personality disorders in late life. Comorbid major depression is common and is often the presenting diagnosis when the underlying personality disorder is not appreciated. Avoidant and dependent individuals, in particular, often enter later life without having been able to cultivate many supportive relationships. As a result, the loss of a spouse or someone who served as their main support can be devastating and precipitate severe anxiety, depression, and incapacity. Given their poor social skills and inability to activate the social resources to meet their daily needs, these individuals may end up in dire circumstances, bereft of support and at high risk in the community of abuse and neglect. The same psychopathology may be somewhat adaptive in long-term care settings, however, where forced dependency is a welcome respite for them and alleviates the difficult responsibility of having to assume self-care.

Individuals with obsessive-compulsive personality disorder tend to be less isolated in later life than avoidant or dependent individuals, and have often cultivated sufficient social supports due to their rigid insistence on routines and schedules. These individuals are more self-sufficient and may do relatively well with aging as long as they are able to find activities, part-

time jobs, or other tasks to keep them busy and organized. Cognitive and physical disability can threaten this stability, however, especially when obsessive-compulsive personalities are unable to control their daily life according to preset rules and routines and have to depend on less reliable caregivers or long-term care staff. The latter circumstances can lead to significant conflicts with caregivers, as well as angry behavioral outbursts, anxiety and panic attacks, and depression.

TREATMENT

Given both the chronicity and pervasiveness of personality disorders, the initial goal of treatment is not to provide a cure, but to reduce the frequency and intensity of symptoms. True symptomatic improvement is challenging and takes months, if not years to accomplish, and will always be slowed down when comorbid medical and other psychiatric disorders are present. Similarly, treatments for major depression have been found to take longer and to be less effective in the presence of underlying personality disorders. Thus, realistic target symptoms should be established at the outset of treatment so as not to fall into the trap of building up excessive expectations that end up disappointing the patient and his or her social supports as well as burning out the therapist.

There are several necessary treatment approaches for any older individual with a personality disorder. First, a general survey of active medical, neurologic, and medication-related factors that may be causing acute changes in personality expression must be ascertained to ensure that adequate treatment is being provided. Many of these factors, including cerebral dysfunction, metabolic abnormalities, pain, and unmet physical or emotional needs, may be triggering and/or fueling the interpersonal conflicts or disruptive behaviors that represent the personality disorder. Normal psychotherapeutic or pharmacologic strategies will fail unless these factors are addressed. This first step will necessitate significant involvement of multidisciplinary treatment team members, other treating physicians, and caregivers in order to implement the treatment plan as well as to serve as informants on response—particularly for individuals with cognitive impairment and/or physical disabilities.

Second, comorbid psychiatric disorders such as major depression, panic disorder, and substance-related and addictive disorder, among others, must be under treatment in order to reduce the psychological stress on the personality, as well as to clarify the personality disorder diagnosis. Diagnostic false positives are more likely during episodes of other psychiatric disorders, prompting many clinicians to defer personality disorder diagnoses during this time. A particular challenge at this juncture involves dealing with personality disorders within the setting of Alzheimer disease or other types of neurocognitive disorders, which can change personality itself or distort previously dysfunctional expressions of

personality in ways both positive and negative. Some individuals may be more labile and impulsive, while others may be more apathetic and less likely to act out, especially when there is frontal or temporal lobe dysfunction. Unfortunately, it is difficult to predict the course of any given individual personality or personality disorder within the setting of neurocognitive disorders, even when the specific regions of cerebral or subcortical damage are known. Although there are a number of existing acetylcholinesterase agents and one glutamate-receptor antagonist that can treat the cognitive, behavioral, and functional symptoms of Alzheimer disease, there is no evidence that any of these agents will have a discernible effect on personality disorders.

Third, clinicians must be willing and able themselves or have available to them a multidisciplinary team that can actively engage the older personality-disordered individual over time and through a variety of treatment modalities in order to foster a stable alliance. For older borderline, histrionic, narcissistic, dependent, and obsessive-compulsive subtypes, episodic clinical visits with an unpredictable schedule and therapy with too narrow a focus on medication management are liable to impair the therapeutic relationship. It is best to have several team members working together to provide an adequate density of time and attention in order to maintain stability and manage crises. On the flip side, older schizoid, schizotypal, paranoid, and avoidant subtypes might do best with shorter and more intermittent clinical visits, with a specific focus geared toward their stated needs. In all circumstances, there are inevitable medical and social issues—and often crises—that arise and require the involvement of a clinician who has made a firm commitment to the patient. Although such an approach seems obvious, it is all too often overlooked in a time-limited and overstretched mental health system.

Fourth, older individuals with personality disorders are often seen in long-term care settings, providing a structured setting in which to engineer environmental or behavioral plans. These plans require the input and creativity of mental health, social work, and nursing staff to be developed, implemented, and revised if necessary. The following clinical vignette illustrates how these plans may work.

Mr. P was a 78-year-old man with severe spinal stenosis who suffered from chronic back pain and was confined to a wheelchair. He had previously been diagnosed with both narcissistic and obsessive-compulsive personality disorders, as well as major depression. Shortly after being admitted to a nursing home, he frequently became enraged and verbally threatening toward nurses whom he accused of not adequately maintaining the battery on his electric wheelchair. On one occasion he struck a nurse who was trying to reconnect a lead to the battery, accusing her of failing to follow his detailed instructions. A multidisciplinary clinical team met to discuss the situation, and decided to condition Mr. P's access to the electric wheelchair on his cooperative behavior. A single staff member was assigned to work with Mr. P to maintain the battery, which limited the targets of his wrath. He earned hours in the electric wheelchair for each day that he went without cursing or abusing staff. As a result, the treatment team noticed a significant reduction in his abusive behaviors.

The key to success with any plan lies in the concerted and well-coordinated efforts of staff to implement its elements, monitor for improvement, and maintain communication. Pitfalls include disputes among staff members that often reflect the personality disorder's inner dynamics, lack of consistency, and lack of patience. When plans fail, staff members may inappropriately resort to punitive actions to deal with dysfunctional behaviors or attitudes, such as withholding attention or care, neglecting legitimate needs, or using chemical restraints unnecessarily. In long-term care settings, there are federal guidelines stemming from the 1987 Omnibus Budget Reconciliation Act or OBRA that ideally should prompt attention to behavioral strategies as well as protect residents from unnecessary and excessive use of psychotropic medications.

Psychotherapy

In older individuals with several subtypes of personality disorders and minimal to no cognitive impairment, both individual and group psychotherapeutic modalities may be helpful. Schizoid, schizotypal, paranoid, antisocial, and avoidant individuals tend to be more reluctant to engage in insight-oriented therapy, and might do best through the indirect fostering of therapeutic relationships. In long-term care settings, this can be achieved through nonpsychiatric therapies (e.g., physical or occupational therapy) or activities (small jobs, arts and crafts, and music) that provide vital involvement, socialization, and meaning. Older individuals with borderline, narcissistic, histrionic, obsessive-compulsive, and dependent personality disorders will do better with direct forms of psychotherapy, especially cognitive-behavioral therapy (CBT). Dialectical behavior therapy (DBT) is a particularly useful form of CBT that was originally developed for individuals with borderline personality disorder but has been adapted for individuals with major depression and personality disorders. DBT focuses on helping individuals become mindful of both the triggers and the consequences of their behaviors and then develop strategies to change them. Studies in younger individuals with DBT have found less overall depression and anxiety, fewer suicidal and self-injurious behaviors, and a decreased number and length of psychiatric hospitalizations.

Group psychotherapy can take on many forms for elderly individuals, ranging from psychoeducational forums to single-focus groups for socialization, bereavement, or caregiving to more formal therapy groups such as are found in a structured outpatient program. Group selection for the older personality-disordered individual will depend on his or her tolerance for a group setting, ability to participate without either monopolizing the group or acting out, and level of motivation. Cognitively

intact individuals with reasonably good social skills are the best candidates.

Pharmacotherapy

Psychotropic medications can be used across several levels in the older patient with a personality disorder: (1) crisis management; (2) treatment of comorbid psychiatric disorders, such as major depression; (3) treatment of the personality disorder itself; and (4) treatment of the most salient target symptoms associated with the personality disorder. The first two levels are covered extensively in other parts of this text, and are inevitably faced with every personality disorder. Treatment of the personality disorder itself is quite difficult given the heterogeneity of presentations, and the literature has almost exclusively focused on borderline, schizotypal, and antisocial subtypes. Even in the majority of those studies, the main focus of treatment was on selected target symptoms that were most problematic, such as impulsivity, aggression, and psychosis. All three of these common symptoms in both dramatic cluster and odd cluster personality disorders have been found to be partially responsive to antipsychotic agents, antidepressants, and mood stabilizers such as divalproex (Depakote) and carbamazepine (Tegretol). Many of the early studies looked at the use of tricyclic antidepressant agents and monoamine oxidase inhibitors to treat symptoms of personality disorders. Although they demonstrated some degree of efficacy, these latter agents are less practical in elderly patients due to their anticholinergic properties and propensity to cause orthostasis. In general, there is no evidence to date that clearly supports one particular agent or medication class as more efficacious than another for any particular symptom or personality disorder. As it stands, given the lack of data for older individuals, the selection of specific psychotropic medications for both young and old individuals with personality disorders will follow the same guidelines.

In older individuals, clinicians must hearken to the familial caveat to "start low, go slow, but go" when prescribing medications. However, another equally important caveat is to know when the use of medications might actually be counterproductive. It is risky to prescribe to patients with poor adherence to medications or hazardous use of them due to cognitive impairment, denial of illness, or impulsivity. In all of those cases, more structure is needed to administer the medications, either through a trusted family member or a visiting nurse, or by transferring the patient to an assisted living or long-term care facility. Patients with obsessive-compulsive and dependent traits sometimes become fixated on medication doses and dosing schedules and spend excessive time discussing this with the prescribing doctor instead of dealing with more relevant issues. Taking a medication becomes a way for them to pathologically structure their relationship with the doctor, such as through increased dependency. In all of these situations, the clinician must weigh the potential benefits of the

medication against the potential harm to the patient through misuse or the resultant exacerbation of dysfunctional behaviors.

References

Abrams RC, Horowitz SV. Personality disorders after age 50: a meta-analysis. *J Personal Disord*. 1996;10(3):271–281.

Abrams RC, Spielman LA, Alexopoulos GS, Klausner E. Personality disorder symptoms and functioning in elderly depressed patients. *Am J Geriatr Psychiatry*. 1998;6:24–30.

Agronin M. Personality disorders in the elderly: an overview. *J Geriatr Psychiatry*. 1994;27(2):151–191.

*Agronin ME, Maletta G. Personality disorders in late life: understanding and overcoming the gap in research. *Am J Geriatr Psychiatry*. 2000;8(1):4–18.

*Balsis S, Eaton NR, Cooper LD, Oltmanns TF. The presentation of narcissistic personality disorder in an octogenarian: converging evidence from multiple sources. *Clin Gerontol*. 2011;34(1):71–87.

*Balsis S, Segal DL, Donahue C. Revising the personality disorder diagnostic criteria for the Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (DSM-5): consider the later life context. *Am J Orthopsychiatry*. 2009;79:452–460.

Balsis S, Woods CM, Gleason ME, Oltmanns TF. Overdiagnosis and underdiagnosis of personality disorders in older adults. *Am J Geriatr Psychiatry*. 2007;15(9):742–753.

Black DW, Baumgard CH, Bell SE. A 16- to 45-year follow-up of 71 men with antisocial personality disorder. *Compr Psychiatry*. 1995;36(2):130–140.

Grant BF, Hasin DS, Stinson FS, et al. Prevalence, correlates, and disability of personality disorders in the United States: results from the national epidemiologic survey on alcohol and related conditions. *J Clin Psychiatry*. 2004;65(7):948–958.

Ingenhoven T, Lafay P, Rinne T, Passchier J, Duivenvoorden H. Effectiveness of pharmacotherapy for severe personality disorders: meta-analyses of randomized controlled trials. *J Clin Psychiatry*. 2010;71(1):14–25.

Kunik ME, Mulsant BH, Rifai AH, et al. Personality disorders in elderly inpatients with major depression. *Am J Geriatr Psychiatry*. 1993;1(1):38–45.

Lynch TR. Treatment of elderly depression with personality disorder comorbidity using dialectical behavior therapy. *Cogn Behav Pract*. 2000;7:468–477.

McGlashen TH. Schizotypal personality disorder. Chestnut Lodge follow-up study: IV. Long-term follow-up perspectives. *Arch Gen Psychiatry*. 1986;43:329–334.

McGlashen TH. The Chestnut Lodge follow-up study. III. Long-term outcome of borderline personalities. *Arch Gen Psychiatry*. 1986;43:20–30.

*Oltmanns TF, Balsis S. Personality disorders in later life: questions about the measurement, course, and impact of disorders. *Annu Rev Clin Psychol*. 2011;7:321–349.

Paris J, Brown R, Nowlis D. Long-term follow-up of borderline patients in a general hospital. *Compr Psychiatry*. 1987;28:530–535.

*Schuster JP, Hoertel N, Le Strat Y, Manetti A, Limosin F. Personality disorders in older adults: findings from the National Epidemiologic Survey on Alcohol and Related Conditions. *Am J Ger Psychiatry*. 2013;21(8):757–768.

Segal DL, Coolidge FL, Rosowsky E. *Personality Disorders and Older Adults: Diagnosis, Assessment, and Treatment*. Hoboken, NJ: John Wiley & Sons, Inc.; 2006.

Stevenson J, Datyner A, Boyce P, Brodaty H. The effect of age on prevalence, type and diagnosis of personality disorder in psychiatric inpatients. *Int J Geriatr Psychiatry*. 2011;26(9):981–987.

Thompson LW, Gallagher D, Czirr R. Personality disorder and outcome in the treatment of late-life depression. *J Geriatr Psychiatry*. 1988;21(2):133–153.

Zweig RA, Agronin ME. Personality disorders. In: Agronin ME, Maletta GJ, eds. *Principles and Practice of Geriatric Psychiatry*. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2011:523–544.

57.3j Alcohol and Substance Abuse in Older Adults

STEVE KOH, M.D., MICHAEL C. HANN, M.D., AND DAVID W. OSLIN, M.D.

The demographics of the United States are shifting as the "baby-boom" generation ages and advances in medical science and public health lead to longer life spans. Projections based on U.S. Census data indicate that by the year 2030, approximately 18 percent of the U.S. population will be over the age of 65 years. This aging trend is not unique to the United States. The United Nations Department of Economic and Social Affairs projects that those above age 60 will account for about 21 percent of the total global population by 2050. Even within the older adult group, the oldest old (above age 80) is increasing. Furthermore, the ratio of males to females is increasing due to improvements in male life expectancy. In 2013, this ratio in those above age of 60 was 76:100. By 2050, this will increase to 87:100. Population trends indicate that the aging cohort of adults who are currently considered middle age will reflect the substance use and misuse trends of younger cohorts.

As the U.S. population ages, classically held views on alcohol and substance misuse are becoming increasingly inaccurate. Indeed, although alcohol and drug misuse are among the foremost causes of disability worldwide and contribute to negative health outcomes, the misuse of and addiction to alcohol, prescription medications, marijuana, and illicit substances are often not examined or addressed in the geriatric population and have grown to the levels of an "invisible epidemic." Data from the National Household Survey on Drug Abuse suggests that by the year 2020, 2.7 million Americans aged 65 years or older will misuse prescription drugs and 4.4 million older adults will require some form of substance abuse treatment. Other beliefs include the view that older women do not misuse substances and that older adults do not benefit from substance misuse treatment. Form of ageism is currently common in clinical practice and these issues are not well explored by clinicians.

Table 57.3j–1.

Barriers to Detection of Substance Misuse in the Geriatric Population

Barriers to the Clinician	Barriers to the Older Adult
Assumptions of lower prevalence of substance misuse in the geriatric population	Poor insight into substance misuse and reluctance to voluntarily seek help
Failure to screen appropriately for substance misuse in older population	Lack of knowledge about the mental health system and treatment options
Masking of symptoms of substance misuse by other comorbid physical and mental/cognitive ailments	Feelings of shame, denial, desire to continue using, pessimism about treatment and recovery
Discomfort with addressing substance misuse with older adults	Cognitive problems–substance-induced amnesia, underlying neurocognitive disorder
Absence of collateral information from family members and caretakers	Family members and/or caretakers may not adequately report concerns of substance misuse
Assumption that older adults may not benefit from substance misuse treatment and prevention	Belief that other medical issues may be more important to address than substance misuse

Adapted from Crome I, Dar K, Janikiewicz S, Rao T, Tarbuck A. Our Invisible Addicts: First Report of the Older Persons' Substance Misuse Working Group of the Royal College of Psychiatrists. College Report CR 165, June 2011.

Several common barriers impede the identification of substance misuse in the elderly, some of which are outlined in Table 57.3j–1. Many stereotypes of the typical older adult with substance abuse are prevalent in clinical care. For example, an older patient needing addiction treatment is often deemed to have been a life-long heavy drinker with failing health related to alcohol use and a limited to nonexistent social network. Also, the relatively healthy older adult who is drinking "moderately" is not considered to be in need of help and is often not assessed fully for the potential effect of alcohol on overall health and well-being. As the baby-boom cohort enters the geriatric population, unique challenges will face providers; along with the increased size of this cohort with changes in age distribution, there will be higher rates of excessive drinking, binge drinking, and illicit drug use.

An additional potential barrier to identification is the relative rarity of severe use disorders in the elderly, where the prevalence of moderate to severe alcohol use disorder is less than 4 percent and even lower in women, and illicit drug use disorders are currently less than 1 percent. On the other hand, 10 to 20 percent of older adults are drinking above recommended limits and are therefore at risk of having or developing problems even if they do not meet criteria for a use disorder.

For older adults, lack of understanding regarding the definition of at-risk use and who may benefit from treatment interferes with assessment or treatment-seeking behavior, which may be further compounded by denial. Similar to providers, patients may perceive physical symptoms (e.g.,

fatigue, sleep problems, anxiety, confusion) as normal or attribute them to a comorbid condition. Furthermore, age- and substance-related declines in memory may result in underreporting of past and current substance use. These complications in dealing with older patients in identifying alcohol and/or substance misuse is illustrated in Table 57.3j–2.

DSM-5 TERMINOLOGY

The *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edition (DSM-5) criteria for substance use disorder are considered identical for younger and older adults. Whereas distinctions were made between substance abuse and dependence in previous editions of the DSM, DSM-5 combines these diagnoses into a single disorder measured on a continuum from mild to severe. The essential feature of a substance use disorder is a cluster of cognitive, behavioral, and physiological symptoms indicating that the individual continues using the substance despite significant substance-related problems. Each of the 10 separate classes of substances is addressed as a separate use disorder (e.g., alcohol use disorder, stimulant use disorder, etc.) and is assessed clinically along a severity spectrum scale.

Table 57.3j–2.
Case Study on the Barriers to Identification of Alcohol and/or Substance Misuse in the Geriatric Population

Mr. Doe is a 74-year-old man who was evaluated in the Emergency Department (ED) after a fall in which the patient sustained a hip fracture. In addition to hip pain, the patient endorses some ill-described upper abdominal discomfort. The patient's blood alcohol level (BAL) was 0.04 in the ED. The patient lives alone but takes care of himself, and has a small social network of a few friends. He has family but they do not live nearby, and he is financially independent. When asked about general health habits, he reports that he does not smoke, but drinks every day at dinner and before bed. The ED physician asked the CAGE questions, which upset the patient who responded negatively to all items. The patient is taking hydrochlorothiazide and Lisinopril for hypertension and was recently prescribed Ambien as a sleep aid and a histamine antagonist for gastritis by his primary care physician (PCP). The ED physician documented that the patient's fall was likely secondary to use of Ambien, and declined to investigate the patient's drinking habits further due to the patient's irritation with the topic.

Review of the patient's medical record indicated that the patient received a mental health evaluation 6 months earlier at the request of his family. The family was worried about him living alone and wanted him to move into an assisted living facility. He was determined to be able to perform daily activities, although he was somewhat slow,

The case study illustrates several diagnostic dilemmas faced by clinicians who treat older drinkers:

- Mr. Doe did not seem to meet any dependence criteria.
- The patient, the ED physician, and the psychiatrist failed to make the connection that alcohol misuse was probably contributing to Mr. Doe's gastritis and to his presentation in the ED.
- Mr. Doe was, however, drinking well above the recommended limits and was therefore an at-risk drinker, even if he did not present as the typical "alcoholic."

Take home points:

- Diagnostic criteria and symptoms for alcohol and prescription drug misuse are not as applicable for older adults and may be confounded with comorbid medical illnesses.
- Lack of recognition by health care providers is often compounded by lack of understanding of the benefits of reduced substance use when the

and he rarely left home. There was no sign of dementia, but there was minor impairment in reaction time and problem solving. When asked about alcohol use, he reported that he drank one drink at dinnertime and two drinks of sherry before bed. He described this pattern of drinking since he was widowed 15 years ago.

patient is not physiologically dependent or there is a general belief that there are few, if any, effective, accessible treatments to target substance misuse.

► Improved dissemination of current treatment methods and their efficacy as well as continued development of other effective treatments is essential.

DSM-5 defines substance use disorder as a maladaptive pattern of substance use occurring within the previous 12 months leading to clinically significant impairment or distress. Diagnostic criteria include recurrent substance use resulting in failure to fulfill major roles or obligations, recurrent substance use in physically hazardous situations, and continued substance use despite persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of the substance. Other criteria include the presence of tolerance and withdrawals, intense cravings, consumption of the substance in larger amounts or time periods than intended, persistent desire or unsuccessful attempts to cut down, and excessive time spent obtaining, using, or recovering from the substance. Finally, an individual with a substance use disorder may reduce or sacrifice important social, occupational, or recreational activities and continue to use the substance despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by the substance. A patient's substance use disorder is categorized as "moderate" if they meet two to three criteria and "severe" if they meet four or more criteria.

For the purposes of terminological clarity and to reflect the transition to DSM-5 criteria, the term "abuse" will only be used generally in this chapter and will not be used to refer to the clinical entity described in previous versions of the DSM. However, "physiological dependence," which is defined as evidence of tolerance or withdrawal (Severity Spectrum Scale Items 4 and/or 5), continues to be clinically relevant and will be used in this chapter in accordance with the DSM-5 conceptualization of this term.

Although the diagnostic criteria for substance use disorders are the same for younger and older adults, some of the criteria may be less relevant for the latter. For example, problems related to work or the family may not be relevant for a retired or widowed older adult. Older women tend to drink at home alone, thus would be unlikely to have legal or work-related difficulties. Similarly, an older adult with physiological dependence on a benzodiazepine with chronic insomnia may have no other symptoms of a sedative, hypnotic, and anxiolytic use disorder and may be more prone to falls and cognitive impairment from the benzodiazepine use itself. On the other hand, health problems are often more prominent in the geriatric

population and may be a motivating reason for changing alcohol and substance use.

HISTORY

Substance misuse among older adults as a distinct entity was first described in the research literature in 1964. In the 1960s and 1970s, researchers believed that substance misuse was rare in the elderly and thus devoted little attention to this population as compared to younger adults. Health providers assumed that individuals with life-long addictions either died early or spontaneously improved in later life, and older adults with alcohol misuse were considered to likely have had a long history of drinking. Even the DSM-III-R, published in 1987, reads: "In males, symptoms of alcohol dependence or abuse rarely occur for the first time after age 45." Thus, until the past two decades, alcohol and drug misuse in the elderly was rarely discussed in the substance abuse or in the gerontological literature. Even today, as the frequency of substance problems climbs in older adults, it remains largely an invisible epidemic currently underdiagnosed and undertreated. Screening and preventative measures for older populations are often missing in routine primary care visits. More recently, studies have been conducted examining alcohol and prescription misuse in older adults and investigating treatment approaches better suited to this population. Illicit drug use in the elderly is believed to be uncommon in the general population and is not well researched.

ALCOHOL

Spectrum-of-Use Categories

As the DSM-5 criteria may be less ideally applied for older adults, alternate categories of substance use may be more appropriate, particularly for alcohol use. Individuals who drank less than one or two alcoholic drinks in the previous year are categorized as *abstainers*. Approximately 50 to 70 percent of older adults report abstinence, making this the most common late-life drinking pattern. The reasons for abstaining, however, may have implications regarding treatment and counseling about drinking. Although some individuals do report life-long abstinence, many individuals stop drinking alcohol due to the onset or presence of an acute or chronic illness, while others abstain due to a previous history of alcohol problems. Identification of the latter group allows preventive monitoring and screening to help avoid future relapse and also alerts the clinician to increased vulnerability for mental health problems and cognitive declines. *Low-risk*, *social*, or *moderate drinkers* include individuals who drink within the recommended guidelines and are not using any contraindicated medications or having any problems related to use.

An *at-risk* or *excessive drinker* is defined as someone who does not

meet alcohol use disorder criteria but is drinking at a level that increases the likelihood of problems or negative consequences. Drinking above recommended limits would be included in this category (more than one drink per day and, in particular, more than three drinks in a day). *Problem drinking* is the term often used in place of abuse for older adults drinking at a level that is causing detrimental medical, psychological, or social effects.

The Spectrum-of-Use Categories and the DSM-5 Severity Spectrum Criteria are both based on knowledge from young and middle-aged adults, and the applicability and validity in older persons is questionable. Physiological tolerance may occur in older adults without increased consumption over time due to the physiological effects of aging. DSM-5 diagnostic criteria are also susceptible to self-report biases, including memory impairment, lack of insight, and unwillingness to acknowledge symptoms. Other signs and symptoms may alert the physician to potential alcohol and substance misuse in older adults (Table 57.3j–3). In addition, collateral information, particularly when cognitive impairment is present, from family, friends, employers, or others is an important source of information to supplement or confirm the patient's report.

Epidemiology

Alcohol misuse is often underreported and underestimated in older adults. While typically drinking tends to decrease with age, some studies have shown that aging cohorts' rate of drinking is rising. Furthermore, the aging baby-boom generation, with its higher rates of alcohol and substance use, is more likely than previous generations of older Americans to misuse alcohol. In the past few decades, prevalence estimates of at-risk or problem drinking in the elderly community range from 1 to 15 percent, with the variability attributed to definitional issues and sample variation. In the most recent National Survey on Drug Use and Health (NSDUH), 48 percent of those 60 to 64 years old were reported to have consumed alcohol in the past month, with about 10.1 percent reporting binge use (defined as five or more drinks on the same occasion in the past 30 days) and 2.7 percent reporting heavy use (five or more drinks on the same occasion on each of 5 or more days in the past 30 days). For those over the age of 65, 38.4 percent reported alcohol use in the past month with 6.0 and 1.6 percent reporting binge and heavy use, respectively. It should be noted that these are definitions developed for the general population and are not those generally used for older adults.

Table 57.3j–3.
Common Signs and Symptoms of Potential Substance Misuse and Abuse in Older Adults

Anxiety
Blackouts, dizziness
Depression, mood swings
Disorientation
Family problems
Financial problems
Headaches
Incontinence
Increased tolerance to alcohol or medications
Legal difficulties
Memory loss
New difficulties in decision making
Poor hygiene
Falls, bruises, and burns
Idiopathic seizures
Sleep problems
Social isolation
Poor nutrition

Adapted from Barry KL, Blow FC, Oslin DW. Substance abuse in older adults: review and recommendations for education and practice in medical settings. *Subst Abus*. 2002;23:105–131. Available at http://www.informaworld.com

Gender-related changes in alcohol consumption are also quite relevant diagnostically in screening for alcohol misuse. Women have historically tended to curb their drinking as they age more than men; however, it is unclear if that pattern will continue as women from the baby-boomer generation enter older age. Older women are more susceptible than their male counterparts to the adverse effects of alcohol misuse, including medical morbidity, and are more apt to have comorbid medical conditions, depression, anxiety, and abuse of psychoactive medications. Furthermore, as men live longer, it will be important to continue the screening for alcohol misuse and consider preventative measures.

Rates of alcohol use disorders are greater in health care settings relative to community-based samples. In a recent study of older adults in primary care, 4.1 percent were defined as at-risk drinkers (8 to 14 drinks per week), and 4.5 percent were heavy (more than 14 drinks per week) or binge drinkers. Mental health clinics have even higher rates of alcohol use disorders.

Etiology

The etiology of problem drinking in older adults has not been well studied. It is a common myth that older problem drinkers have a life-long history of alcohol-related problems. This myth is more accurate for illicit drugs and nicotine, whereas as many as one-third of older adults seeking treatment for alcohol misuse develop their problems in later life. Thus, although

drinking patterns are often stable over time, increased drinking in later life is not unusual due perhaps in part to more free or unstructured time, a rise in disposable income, and losses of loved ones. Older adults also potentially face several stressful life transitions (i.e., retirement, menopause, the "empty nest" phenomenon, loss of friends, reduced physical function) and the adoption of new roles (i.e., widow[er], caretaker) as they approach an older age, thereby increasing the vulnerability to alcohol misuse. Treatment approaches for late-onset and life-long problem drinkers are currently the same; however, the distinction is important for clinicians as the former are more likely to go undiagnosed and therefore untreated.

Recommended Limits

Certain age-related factors, including physiological factors that increase sensitivity to alcohol and drugs in older adults, have led the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the Center for Substance Abuse Treatment (CSAT) to set lower recommended limits for alcohol use for older adults. The maximum drinking limit for women and men over the age of 65 is no more than an average of one standard drink per day, versus two standard drinks per day for younger men. Table 57.3j–4 provides a description of how to measure a standard drink. In addition, older adults should not consume more than three standard drinks in 1 day (i.e., binge drinking).

Table 57.3j–4.
Alcohol Conversion Chart

"1 Standard Drink"	
Beverage Type	**Quantity (oz)**
Beer	12
Wine	5
Fortified wine	3
Hard liquor (80-proof distilled spirits)	1½ (i.e., "a shot")
Malt liquor	8
Liqueur or aperitif	4
Additional Conversions	
Beverage Type/Quantity	**# Standard Drinks**
Beer:	
► 1 (6) pack of 16 oz cans/bottles	8
► 1 quart	3
► 1 L (33.8 oz)	2.8
Wine (e.g., red white, chianti):	
► 1 bottle (25 oz)	5
► 1 bottle (40 oz)	8
► Fifth	6

► 1 magnum	12
► ½ gallon	16
Fortified wine (e.g., sherry, port):	
► Pint	7
► Fifth	12
Hard liquor (e.g., rum, gin, tequila, vodka):	
► Fifth	18
► Pint	11
► Quart	20
► ½ gallon	40
► 1 bottle (25 oz)	17
► 1 bottle (40 oz)	27
Malt liquor (e.g., Olde English):	
► 12 oz	1½
► Forty (40 oz)	5

Standardized Screening Instruments

Brief and low-cost assessments can be used to screen for alcohol misuse to help in the prevention or early intervention of older adults at risk and can also be used to facilitate a diagnosis of alcohol use disorder. These instruments may be administered by verbal interview or by paper-and-pencil or computerized questionnaire with equal reliability and validity. Including follow-up questions about consequences, health risks, and interpersonal issues related to substance use can assist in the determination of a DSM-5 diagnosis of substance use disorder.

The most appropriate screening assessments used in the geriatric population are the Short Michigan Alcoholism Screening Test–Geriatric version (SMAST-G) and the Alcohol Use Disorders Identification Test (AUDIT). The SMAST-G was developed specifically for screening older adults in a variety of settings, and its psychometric properties exceed those of other screens for identification of alcohol use disorders in this population (Table 57.3j–5). The AUDIT and the abbreviated version, the AUDIT-C (Table 57.3j–6), assess frequency of drinking and binging in the past year. The AUDIT-C score ranges from 0 to 12, with a higher score indicating increased likelihood of alcohol use impacting health and safety. For older adults, a score of 2 or greater suggests the need for further assessment or recommendation of a brief alcohol intervention.

Other screening instruments that are commonly used in younger patients are not ideal for older adults for various reasons. The CAGE questionnaire is the most widely used alcohol-screening test. However, while the CAGE questionnaire has shown high specificity for detecting alcohol use disorders, there is a relatively low sensitivity for problem

drinking. Poor insight into drinking is often a barrier for older adults' self-reporting of amount of drinking and for isolated older adults there may not be enough social interaction to report if others are criticizing the drinking pattern. If a CAGE questionnaire is used, a modified version that only inquires about recent problems is most appropriate for older adults, and the threshold for indicating possible problems is one positive response.

Table 57.3j–5.
Short Michigan Alcohol Screening Test—Geriatric Version (SMAST-G)

In the past year:	Yes	No
1. When talking with others, do you ever underestimate how much you actually drink?	(1)	(0)
2. After a few drinks, have you sometimes not eaten or been able to skip a meal because you didn't feel hungry?	(1)	(0)
3. Does having a few drinks help decrease your shakiness or tremors?	(1)	(0)
4. Does alcohol sometimes make it hard for you to remember parts of the day or night?	(1)	(0)
In the past year:		
5. Do you usually take a drink to relax or calm your nerves?	(1)	(0)
6. Do you drink to take your mind off your problems?	(1)	(0)
7. Have you ever increased your drinking after experiencing a loss in your life?	(1)	(0)
8. Has a doctor or nurse ever said they were worried or concerned about your drinking?	(1)	(0)
9. Have you ever made rules to manage your drinking?	(1)	(0)
10. When you feel lonely, does having a drink help?	(1)	(0)
TOTAL SMAST-G SCORE (0–10)		
Two or more positive responses are indicative of an alcohol use problem.		

Reprinted with permission of the University of Michigan Alcohol Research Center. © 1991 The Regents of the University of Michigan.

Table 57.3j–6.
AUDIT-C Alcohol Screening

1. How often did you have a drink containing alcohol in the past year?

—	Never	(0 points)
—	Monthly or less	(1 point)
—	Two to four times a month	(2 points)
—	Two to three times per week	(3 points)
—	Four or more times a week	(4 points)

If you answered "never," score questions 2 and 3 as zero.

2. How many drinks did you have on a typical day when you were drinking in the past year?

—	1 or 2	(0 points)
—	3 or 4	(1 point)
—	5 or 6	(2 points)
—	7 to 9	(3 points)
—	10 or more	(4 points)

3. How often did you have six or more drinks on one occasion in the past year?

—	Never	(0 points)
—	Less than monthly	(1 point)

—	Monthly	(2 points)
—	Weekly	(3 points)
—	Daily or almost daily	(4 points)

Possible range = 0–12. For older adults, a score of 3 or more is considered positive.

Differential Diagnosis

Clinicians must be careful when diagnosing psychiatric conditions that occur only in the presence of intoxication or withdrawal as both of these states may mimic psychiatric symptoms, including depressive, anxiety, and psychotic symptoms. Diagnosing based on historical self-report of symptoms leads to questionable reliability and validity, particularly with older patients and alcoholics who may have cognitive deficits as part of their illness.

Alcohol problems and depression are common comorbid conditions in late life. Depression symptoms can precede (primary or independent depression), occur simultaneously (indeterminate), or develop as a consequence (reactive or secondary depression) of significant alcohol use. Repeated use of depressants can induce mood disturbance and may affect the mood enough to meet criteria for a major depressive episode. A diagnosis of primary depression may lead to a treatment plan targeting both disorders, whereas a secondary depression diagnosis would suggest treating the addiction and monitoring the depression. In the latter case, treatment with antidepressants is not likely to be effective, whereas termination of the substance use can resolve the mood disturbance. In one notable study, of 42 percent of patients who began detoxification with significant symptoms of depression, only 6 percent reported continued significant depressive symptoms at 4 weeks.

The high prevalence of chronic medical conditions in older adults may affect the ability of the clinician to determine decreased functioning due to substance misuse. Alcohol problems contribute to cognitive deficits in later life. Wernicke–Korsakoff syndrome is well defined and often caused by alcohol dependence, but alcohol-related dementia may be difficult to differentiate from Alzheimer disease due to a lack of well-specified criteria. A study of older adults seeking treatment found 23 percent had dementia associated with alcohol dependence. In older adults, sedative intoxication can also be confused with progressive dementia. Withdrawal of the substance tends to stabilize cognitive and functional status for substance-related cognitive impairment and not for dementias.

Sleep disorders also frequently co-occur with excessive alcohol use. Alcohol use and aging are both associated with changes in sleep patterns. Alcohol has been shown to decrease sleep latency, decrease stage IV sleep, and precipitate or aggravate sleep apnea. Age-related changes, which can be exacerbated by alcohol use, depression, comorbid medical conditions, and medication use, include increased rapid eye movement (REM)

episodes, decrease in REM length and stage III and IV sleep, and increased awakenings. On the other hand, abstinent alcoholics with poor sleep due to insomnia, frequent awakenings, and REM fragmentation demonstrate normalization of sleep periodicity and temporary suppression of REM sleep upon drinking, suggesting that alcohol may also be used to self-medicate sleep disturbances. Differential diagnosis, again, involves a historical account of the initiation of sleep problems and determining if symptoms subside with substance abstinence.

Course and Prognosis

As mentioned, drinking patterns often are stable over time but increased drinking in later life is not unusual due to factors idiosyncratic to an older population. In addition, aging patients can experience increased problems due to alcohol without increased consumption due to physiological changes that increase sensitivity to alcohol and drugs, increased medication use, and development of chronic health conditions.

Alcohol misuse is related to adverse effects in several physical health and psychiatric domains. Intoxication can lead to increased risk of falls and motor vehicle accidents, especially in conjunction with medication interactions or other substance misuse. Chronic alcohol misuse is associated with multiple disease outcomes in several organ systems. More than 90 percent of older patients with a diagnosis of alcohol misuse have a history of depression and depressed older adults are three to four times more likely to have an alcohol use disorder compared with nondepressed older adults. Like depression, late-life anxiety disorders can both be a consequence of problematic alcohol use as well as a primary psychiatric illness. Generalized anxiety disorder is the most common type of anxiety disorder in the older population, and those with the disorder are more than twice as likely to have a substance use disorder as those without an anxiety disorder.

Following depression, substance use disorders, particularly alcohol use disorder, are the second most common category of psychiatric disorders associated with completed suicide in the older population. Drinking among older adults elevates suicide risk through interactions with many factors that are more prevalent in this age group than in younger adults, factors such as depressive symptoms, medical illness, negatively perceived health status, and low socioeconomic support. Together, alcohol misuse and comorbid psychiatric illness are a potentially lethal combination.

Moderate Use

Before discussing the negative prognoses for alcohol and substance abusers, it is important to acknowledge that moderate alcohol consumption may have positive effects on physical health and mental well-being. Low-

risk alcohol consumption has been linked with a reduction in cardiovascular disease risk for men and women. For older adults without cardiovascular disease, moderate alcohol use has been associated with lipoprotein subclass distribution, including fewer small low-density lipoprotein particles and increased large- and medium-sized high-density lipoprotein particles. In addition, moderate alcohol use has been correlated with a lower incidence rate of type 2 diabetes mellitus in older adults and may reduce the risk of cognitive impairment and dementing illnesses.

Despite these potential health benefits, however, abstinence is recommended for individuals taking particular medications and those with a history of alcohol or drug abuse. In addition, although moderate drinking may protect against certain health problems, any alcohol consumption may be detrimental after a diagnosis of a chronic health condition (such as cardiovascular disease or diabetes). Furthermore, there is no evidence that initiation of drinking by nondrinkers will translate into reduced health risks, and moderate use may contribute to several negative health outcomes. Moderate alcohol consumption may reduce stroke risk due to blocked blood vessels but may increase stroke risk due to hemorrhaging or bleeding. The risk of breast cancer in women is increased about 50 percent in women who drink three to nine drinks per week compared to fewer than three drinks per week. In addition, there is a linear relationship between alcohol use and bone mineral density at the hip. Even limited alcohol consumption has been shown to impair one's driving ability and may increase the chance of accidents and fatal injuries due to falls, motor vehicle accidents, and suicides. In addition, depression, memory problems and other cognitive changes, liver disease, cardiovascular disease, and sleep problems have been associated with moderate alcohol use. Mental health providers also need to be cognizant that alcohol can negatively impact other mental health disorders. Even in the moderate use range, greater alcohol consumption adversely affects depression treatment. The converse is also true; reduction of moderate alcohol use during depression treatment enhances treatment outcomes. Limited alcohol use may also negatively impact the course and prognosis of dementing illnesses and may precipitate or exacerbate onset of personality changes and behavioral disturbances in patients with dementia.

Severe Use

Severe alcohol use disorder is associated with increased probability of morbidity and mortality from disease-specific disorders including acute pancreatitis, alcohol-induced cirrhosis, and alcohol-related cardiomyopathy. Comorbidity of substance misuse with other medical and psychiatric conditions may impact the course, treatment, and prognosis of this condition. Substance misuse in earlier life can also impact later life, including increasing the likelihood of developing late-life mental health

problems. Epidemiological surveys have shown that comorbid alcohol use and psychiatric conditions are common in younger age groups. Less is known about later life, although a few studies have shown high frequency of physiological alcohol dependence in older adults with psychiatric problems. In a study of Veterans Administration patients from 60 to 69 years of age, an affective disorder was present in 21 percent of the patients presenting for alcohol treatment, 43 percent of whom had major depression. In another study of community-dwelling older adults, almost half of the 4.5 percent of patients with a history of alcohol misuse had a comorbid depression or dysthymia diagnosis. Comorbid depression and alcohol misuse are not only common in later life but may have a reciprocal effect on each other. Depressed alcoholics are more likely than nondepressed alcoholics to have a complicated course of depression, manifest by increased suicide risk and social dysfunction. In addition, a history of alcohol abuse as a younger adult is predictive of more severe and chronic depression course in late life.

PRESCRIPTION AND OVER-THE-COUNTER MEDICATION

Epidemiology

Due to the elevated use of pharmaceutical drugs in older adults, the chance of misuse of prescriptions and over-the-counter medications, including sedatives or hypnotics, narcotic and nonnarcotic analgesics, diet aids, decongestants, and other medications, is also greater. Although the literature is sparse, it is estimated that up to 11 percent of older women misuse prescription drugs, and by the year 2020, it is projected that 2.7 million adults aged 50 years and older will be using prescription drugs for nonmedical purposes.

Older adults are also far more likely to take multiple medications (polypharmacy); in a study of community-dwelling older adults, 71 percent were taking at least one prescription medication while 10 percent reported taking five or more medications. Community surveys show that 60 percent of elderly adults are taking an analgesic, 22 percent are taking a central nervous system (CNS) medication, and 11 percent are taking a benzodiazepine. Psychoactive medications with the potential for misuse are used by at least one in four older adults, and recent studies estimate the misuse of psychoactive prescription medications by older adults will increase from 911,000 in 2001 to 2.7 million in 2020. Benzodiazepine sedative hypnotics and opioid analgesics are the two main classes of prescription medications with abuse potential used by older adults. It is also important to note that the rate of adverse drug reactions increase with increasing numbers of medications taken and many medications interact negatively with alcohol.

Benzodiazepines

Benzodiazepines, in particular, are one of the most common classes of prescription medications in the United States, and their use increases with age and tend to be chronic in late life. While it is well established that benzodiazepines increase the risk of cognitive impairment, delirium, fall injuries, fractures, and motor vehicle accidents in older adults, there remains insufficient research on the impact of these medications in the older population. Studies on this topic are made more difficult because the prescription medication misuse is often unintentional. According to the CSAT, 23 percent of all prescriptions written for older adults are for benzodiazepines. Despite the wide use of benzodiazepines, the Beers criteria recommends benzodiazepine use be limited in the older population to seizure disorders, REM sleep disorders, benzodiazepine withdrawal, alcohol withdrawal, severe generalized anxiety disorder, periprocedural anesthesia, and end-of-life care. Commonly, sleep disorder is the indication cited for benzodiazepine use in the older adult population, accounting for approximately 59 percent of total prescriptions; this is despite the fact that long-term benzodiazepine use does not help with sleep restoration. Sleep disorders occurring later in life are thought to be related to decreased activity of the suprachiasmatic nucleus, contributing to a disruption of circadian rhythms in older adults including the nocturnal secretion of endogenous melatonin. However, the iatrogenic contribution of medications, including stimulants, antihypertensives, decongestants, chemotherapy, and respiratory medications, cannot be overlooked.

Age-based biases are also commonplace among physicians and caregivers and contribute significantly to misdiagnosis of comorbid conditions and the misuse of benzodiazepines. Physicians and family members often overlook clinical signs of sleep disorders in older adults, including somnolence and fatigue, as signs of normal aging. Common sleep disorders, including sleep-disordered breathing and periodic limb movement syndromes, have been shown to benefit from nonbenzodiazepine evidence-based treatments. Overnight recording of the apnea–hypopnea index is indicated in patients with symptoms of sleep disorder breathing, including high snoring severity, unintentional napping, and/or excessive daytime sleepiness. Primary insomnia in the older adult may be initially treated using numerous nonpharmaceutical tools to improve sleep hygiene, such as limiting naps, adjusting medications, avoiding stimulants like coffee, environmental improvements, examining sleep problems, and other specific treatments. Nonbenzodiazepines sedatives like zolpidem is often considered to be safe and effective in older adults, but recent research shows that they may lead to as many fractures, falls, and delirium as benzodiazepines.

Despite clear indications that benzodiazepines are not the preferred

agent for long-term treatment of anxiety, this is the second most common reason for prescription of a benzodiazepine to an older adult, often in the absence of first-line antidepressant medications. Risk factors for anxiety disorder in the older adults include: female gender, low socioeconomic status, family history, external locus of control, current level of perceived stress, poor primary support, and poor social support. While SSRI and SNRI medications are associated with falls and an increased risk of hyponatremia in older adults, their adverse-effect profile is favorable in the chronic treatment of anxiety compared to benzodiazepines. Evidence for nonpharmacological treatment of anxiety in the older adult population is conflicting regarding cognitive behavioral therapy, physical therapy, and other behavioral approaches.

Opioid Analgesics

Data from multiple agencies, including the Centers for Disease Control (CDC), have documented an alarming rise in prescription opioid analgesic abuse with a concordant rise in attributed mortality. The opiate's impact on the older population is second only to alcohol, accounting for up to 22 percent of substance-related inpatient admissions. While a large percentage of the older population suffers from some form of chronic pain often resulting in pain management with opioid analgesics, there is currently no evidence for the long-term effectiveness of opioid analgesics in the treatment of persistent noncancerous pain conditions in older patients. Despite growing concerns for opioid analgesic misuse, it is important to note that pain continues to be undertreated in adults.

Studies have examined the use of methadone in older adults for the treatment of chronic opioid misuse and physiological dependence and found that they were no different than their younger counterparts in terms of medical and psychiatric problems or employment, but did significantly better in treatment. Despite this more positive prognosis, older adults are thought to be more susceptible to the stigma of drug addiction, old age, psychotropic medication use, depression, poverty, race, and HIV status associated with methadone treatment.

The American Geriatric Society recommends that acetaminophen be used as a first-line agent in the treatment of pain in the older adult population. An appropriate second-line for pain management includes SNRIs, gabapentin, and pregabalin alone or in combination with acetaminophen. NSAIDs should be used sparingly as older adults are considered more susceptible to the potential gastrointestinal adverse effects of these agents. Tricyclics have been shown to be somewhat efficacious in the treatment of pain in the general population; however, their use in the older population is not recommended given their significant adverse effect profile.

Etiology

Declining physical health and physiological changes associated with aging affect the metabolism of medications and thus the potential for medication misuse. Previous and comorbid psychiatric diagnoses, poor health status, social isolation, and female gender also increase the risk for prescription drug misuse. Many cases of psychoactive medication misuse have been shown to be unintentional on the part of the patient or physician. The drugs are usually obtained with an appropriate prescription for a specific health-related purpose, either a physical condition or emotional distress. An older patient may misunderstand the directions, a confusion that may be compounded by multiple medications, or misuse may occur due to a prescription that is not appropriate for an older adult. Unintentional misuse can progress into abuse to obtain the desirable effects. Physiological dependence on psychoactive medications without meeting additional DSM-5 substance use disorder criteria can also result from even therapeutic dosages for brief periods.

Aging also involves increased sensitivity to over-the-counter and prescription medications, with women demonstrating even greater sensitivity. With advancing age, lean body mass and total water volume decrease relative to total fat volume, thereby decreasing total body volume and subsequently increasing the serum concentration, absorption, and distribution of alcohol and drugs in the body. Furthermore, reduction in liver mass and decreased hepatic blood flow results in slower drug metabolism and changes in renal function can lead to build up of some substances. At the same time, the response to CNS-active drugs is frequently exaggerated due to age-related decline in CNS function and increased pharmacodynamic sensitivity for some psychoactive drugs, thus amplifying the response to sedative–hypnotic and mind-altering drugs. The presence of dementia also increases the likelihood of intoxication and impaired functioning. Age-related changes in drug sensitivity vary by drug, however, and pharmacodynamic changes associated with aging continue to be an area of investigation.

TOBACCO

Large-scale epidemiological studies have shown that smoking is harmful for all age groups. Although often overlooked in the elderly, smoking contributes significantly to poor health and disability in older adults. Impressions that late-life smoking cessation does not provide benefit and do not lead to over changes in health outcome are patently false. Even in older age, smoking cessation can decrease overall morbidity and improve quality of life. While smoking rates tend to decrease with age, one of five smokers is 50 years or older. Of community-dwelling individuals, 15.2 percent of adults aged 65 to 74 and 8.4 percent of those older than 75 are

smokers. This rate is even higher in older African American men, where 22.1 percent of individuals 65 to 74 and 12 percent of 75-years-old and older smoke. The importance of clinicians continuing to have low threshold for smoking screening and to work with patients toward smoking cessation do not change while taking care of older patients.

MARIJUANA

The United States is in the midst of a cultural shift regarding its perception of and the potential medicinal applications considered appropriate for marijuana. As of 2014, 23 states have approved some type of medicinal marijuana legislation. The increasingly legal use of and easing access to marijuana is also effecting the substance use patterns of older adults. Data from the NSDUH indicated that 2.8 percent of older adults were past-year marijuana users in 2008. As the baby-boom cohort ages, the rate of increase of past-year use of marijuana is expected to rise given that this cohort is more likely to have used illicit substances than prior cohorts. Marijuana abuse in older adults appears to be concentrated in the "young-old" group, with 2008 NSDUH data indicating that almost all older adult past-year users were aged 50 to 64. Additionally, more than two-thirds of older adult past-year marijuana users initiated use before or at age 18 and an additional 27 percent between ages 19 and 29. The large proportion of older adults with past-year use of marijuana that reported first-time use at these earlier ages suggests that a large portion of this population are long-term users rather than mid- to late-life users who initiated use for medical reasons.

Various studies have been conducted to evaluate the efficacy and safety of medical cannabinoids in a variety of conditions in adult patients, and in a substantial number of studies, older subjects were included. However, there is a general lack of evidence concerning the use of cannabinoids specifically in older patients, resulting in scarcity of data to guide treatment decisions. A review of studies performed that included older patients in the study population indicated no efficacy of medical marijuana on dyskinesia, breathlessness (versus placebo), chemotherapy-induced nausea and vomiting (versus prochlorperazine), or oral THC in the symptomatic treatment of behavioral problems in dementia. However, sample size and methodological quality may have biased these conclusions. Further research on cannabinoids specifically in the older population is needed.

ILLICIT DRUGS

Epidemiological data for illicit drug use in the elderly are limited. Data from the early 1990s in the Epidemiologic Catchment Area Study estimated lifetime prevalence rates for substance misuse to be 0.12 percent for older men and 0.06 percent for older women, and a lifetime history of illicit drug use to be 2.88 and 0.66 percent, respectively. More recent work, however,

suggests that illicit substance abuse is increasingly common in adults seeking treatment. For example, 38 percent of veterans in a Veterans Affairs older adult treatment program reported recent illicit substance use. The recent NSDUH showed an increase from 2.7 to 4.3 percent for illicit drug use in the past month for adults aged 50 to 59 years from 2002 to 2006; other age groups demonstrated no change or declines in use. This trend likely reflects the aging of the baby-boom cohort and is predictive of further increases in illicit substance use and abuse in older adults in the next decades.

SCREENING

Given the high utilization of medical services by older adults, and because a large percentage do not self-identify as substance misusers, physicians and health care staff can be critical for identifying individuals with alcohol and drug problems. Screening for alcohol and other drugs should be included as part of the routine mental and physical health care for adults over the age of 60 and should be updated annually or whenever problems arise that may be related to alcohol or drug use. To ensure successful incorporation, screening should be simple and consistent with clinical screening practices already in place. Routine screening including toxicology screening (urine, salvia) enables identification of those with problematic substance use along with those at risk of misusing drugs and alcohol and helps determine if additional assessment is necessary.

ASSESSING FREQUENCY AND QUANTITY OF USE

Despite these barriers, successful screening techniques have been developed and implemented. Assessing the quantity and frequency of alcohol and drug use generally involves prospective recording and monitoring of use, retrospective accounts of daily use over a defined period of time (i.e., time-line follow-back method [TLFB]), or questions regarding average use. Although the prospective diary method is the gold standard, it is often impractical for screening assessments, as it requires more time and multiple visits. The TLFB method is the most commonly used technique in addiction studies, and 7-day TLFB assessments are highly correlated with the prospective method. Two difficulties arise with the TLFB. First, the week of measurement may not represent the typical week. Second, the TLFB method tends to underestimate use for less frequent users (compared to nonusers or frequent users). The TLFB also takes longer to administer than general questions about use and is more accurate when administered by an interviewer versus self-administration due to variation in individuals' definitions of a standard drink (see Table 57.3j–3 for description of standard drink).

Although general questions about average quantity and frequency of use over a certain time period is the least correlated with the prospective diary

and likely underestimates moderate drinking, it is the easiest assessment to administer and therefore the choice for many clinicians. Of note, this is the method used in the AUDIT-C and it recommended that this standard questionnaire be used over general questions. These questions may be used independently to screen and identify at-risk patients and can be easily embedded in the context of other health behaviors such as exercise, weight, and smoking. Additional measures should be used, however, when attempting to recognize problem drinking or diagnose substance use disorder to identify how the use is affecting the patient dynamically, socially, and physiologically.

PATHOLOGY AND LABORATORY EXAMINATION

Several laboratory values indicate recent use or misuse, including blood alcohol or blood acetate, a metabolite of alcohol. Long-term markers of alcohol use include γ-glutamyl transferase (GGT), mean corpuscular volume, high-density lipoprotein level, and carbohydrate-deficient transferrin (CDT). High blood levels of alcohol without signs of intoxication indicate high tolerance to alcohol. Illicit drug use can often be determined by a simple urine test. The majority of drugs are detectable in a urine drug screen for 4 days or greater, with some detectable after several weeks. Other physical signs of intoxication for different substances can be determined by physical examination. Intoxication with sedative, hypnotic, or anxiolytic substances or opioid medications results in decreased blood pressure, respiratory rate, pulse, and body temperature, while amphetamine or cocaine intoxication generally induces the reverse.

TREATMENT

Previous addiction studies have often excluded patients over the age of 65, resulting in a significant knowledge gap regarding treatment outcomes and the neurobiology of addiction in later life. Nevertheless, naturalistic studies have shown that older adults respond as well as or better to treatment than middle-aged adults, and they are more likely to complete the treatment. Furthermore, brief interventions have been demonstrated to be effective at reducing alcohol consumption for older, at-risk drinkers. Therefore, older adults can be receptive and responsive to treatment, and providers need to recognize that they need not be a substance abuse professional to play a significant role in addressing addiction issues and improving the lives of patients.

Despite the limited research, a variety of treatment strategies are available for and have been used with older adults, including individual, family, and group-based approaches, cognitive-behavioral and medical or psychiatric therapies, community-based services and outreach, and formal substance abuse treatment. Ideally, treatment approaches are individualized given the specific medical, psychiatric, social, and cultural

needs of the patient. More practically, treatment modalities may be limited by managed health care environments and community availability. Currently there is no best treatment option for older adults; however, less intense options are recommended first in this population. Future studies are needed to determine when patients need to be referred to specialty addiction care versus receiving treatment in primary care or mental health clinics. It is likely, however, that older patients with less severe disease can be treated most appropriately in primary care and mental health care settings, where it has been shown they would be more apt to engage in care, and patients with more severe disease should be treated in conjunction with specialty care.

Brief Interventions/Therapies

Low intensity, brief interventions or brief therapies are cost-effective, practical techniques that may be useful as an initial treatment approach with at-risk and problem drinkers across a range of clinical settings. Brief interventions are time limited and nonconfrontational and range from unstructured counseling and feedback to more formal therapy. As the interventions were developed using concepts and techniques from the behavioral self-control literature, one trademark of brief interventions is to encourage individuals to change their behavior through motivational interviewing. As older adults with substance misuse often experience shame about their substance problem, nonconfrontational and supportive approaches are well suited and labels (addict, alcoholic) are best avoided.

Randomized clinical trials of brief interventions for alcohol problems among older populations have demonstrated efficacy of the treatment, that is, reduced alcohol consumption in at-risk drinkers compared to controls, while also demonstrating that older adults can be engaged in brief intervention protocols and find the protocols reasonable. Most trials to date have been conducted in primary care settings using only one to four sessions, including some telephone contacts, and follow-up over 1 year. Brief interventions for older adults are likely to be effective in mental health settings as well; thus, geriatric mental health care providers are encouraged to gain familiarity with brief intervention therapy both as a treatment tool and as a way to increase adherence to specialty treatment.

Psychosocial Treatments

Little research has been conducted on the efficacy of psychological treatments for substance misuse in older adults. One study of older veterans found that individuals who completed a 16-week relapse-prevention group intervention, including cognitive-behavioral and self-management approaches, were more likely to be abstinent at a 6-month follow-up than noncompleters. Another study that included older adults (7

percent were aged 60 or over) demonstrated effectiveness for three different manual driven individual psychological treatments (cognitive-behavioral therapy, motivational enhancement therapy, and 12-step facilitation) that spanned 12 weeks. In this study, alcohol consumption was reduced 1 year posttreatment in adults with alcohol dependence.

Twelve-Step Programs

Community-based programs for many years have centered on the traditional 12-step peer support model of recovery and rehabilitation, which includes group support and encouragement to help group members achieve and maintain sobriety. Studies have shown that older adults engage in and attend 12-step programs with comparable frequency to other age groups. On the other hand, the degree of involvement in the programs has been shown to be less in older adults, and the likelihood of engagement in subsequent formal aftercare is reduced. The 12-step model has clearly been successful for many individuals; however, more treatment options are available now than there was when 12-step programs were developed. In addition, certain factors need to be considered including perceived stigma, level of comfort disclosing personal information in a group setting, degree to which age-appropriate issues are discussed, and transportation and health problems that may interfere with the likelihood of attendance.

Pharmacotherapy

Pharmacological agents have been demonstrated to be effective in treatment of physiological dependence in conjunction with psychosocial treatments. Historically, long-term treatment of older physiologically alcohol-dependent adults did not involve pharmacological approaches. Disulfiram, inhibitor of aldehyde dehydrogenase, was initially the only medication approved for alcohol treatment, and it was used infrequently in older adults due to the likelihood of adverse effects. The U.S. Food and Drug Administration (FDA) approved the opioid receptor antagonist naltrexone in 1995 based on clinical trials demonstrating its safety and effectiveness in preventing relapse and reducing alcohol cravings, making it the first pharmacological agent in 50 years to be approved for treatment of alcohol misuse. The original clinical trials were conducted in middle-aged adults; however, naltrexone has also been shown to be effective in older adults. Results from a double-blind, placebo-controlled, randomized trial demonstrated that in veterans aged 50 to 70 years, twice as many placebo-treated subjects relapsed to significant drinking as compared to those receiving naltrexone. Most recently a depot formulation of naltrexone was approved and potentially offers an advantage in improving adherence to treatment. Acamprosate was approved by the FDA in 2004 and has demonstrated safety and efficacy similar to naltrexone, although the

response in geriatric populations has not yet been well studied. The mechanism of action is unclear but is believed to involve normalization of the glutamate and γ-aminobutyric acid neurotransmitter systems which then modifies the reward pathway of alcohol use. Disulfiram is not recommended because of the potential complications of a disulfiram reaction in a medically compromised patient. Topiramate, which is not FDA approved for alcohol use disorders is also gaining popularity but should be considered with caution in patients with cognitive impairment.

Pharmacological treatments for illicit and prescription drug misuse have not been well studied in older adults. In general, when prescribing for older adults, health care providers need to consider lower initial doses and time-limited dosing for psychoactive medications as well as avoidance of substances with long half-lives. In addition, preference should be given to less complex drugs with simpler pathways and fewer side effects, and close monitoring of patient's reactions is advised. Buprenorphine, an FDA-approved treatment for opioid addiction, has not been studied in older adults, but there is no particular reason the treatment should not be effective and well tolerated in older adults. Depot naltrexone is considered a safe first-line treatment for opioid addiction even in older adults.

Detoxification and Withdrawal

When patients discontinue or drastically reduce substance use after a period of heavy use, withdrawal symptoms are likely. A risk of alcohol and benzodiazepine withdrawal exists during hospitalizations when the treatment team is unaware of the patient's use of these substances. Due to potentially life-threatening complications, clinicians treating substance abusers need a fundamental understanding of withdrawal symptoms and potential difficulties. Familiarity with the most common withdrawal symptoms and anticipated symptom time course is crucial for all clinicians. The first steps should consist of performing a history and physical examination, and considering the patient's medications and medical comorbidities. The goal of the clinician treating a patient potentially in withdrawal from alcohol is patient safety, and preventing potentially serious complications such as seizure, delirium, and death. Consideration of treating the older patient in highly monitored setting should be made. Competence in detoxification management is particularly important for clinicians practicing in settings where withdrawal management is available. Understanding the management of alcohol withdrawal with benzodiazepines is necessary; however, detoxification of older adults should not include stabilization with high doses of benzodiazepines or barbiturates with intermediate to long half-lives. These drugs have the potential to accumulate, and the resulting toxicity can cause cognitive impairment and may affect overall functioning and continuation of treatment. Off-label use of gabapentin in alcohol withdrawal is another

alternative and may have less side effects for older adults than benzodiazepines. The Clinical Institute Withdrawal Assessment of Alcohol Scale, Revised (CIWA-Ar) is a guideline for administering pharmacological agents for the treatment of alcohol withdrawal based on a review of symptoms and physical examination findings that has been validated on a younger population. While useful, it should be carefully applied to older patients, especially in context of potential cognitive impairment.

Use of medications to treat alcohol withdrawal must be balanced between adequate treatment of withdrawal symptoms and the avoidance of over treatment, which can lead to over sedation and other adverse events. As older adults tend to be on a number of medications, medication interactions should be considered prior to initiating treatment. Commonly used benzodiazepines used to treat alcohol withdrawal are oxazepam, diazepam, chlordiazepoxide, and lorazepam. Diazepam and chlordiazepoxide require less frequent dosing and lead to fewer interdose symptoms. However the pharmacokinetics of these medications is altered not only as people age, but also by hepatic dysfunction. Due to its short half-life and metabolism through glucuronidation only, lorazepam is best for treatment of alcohol withdrawal in older adults. In inpatient hospital settings, lorazepam is also preferred because it can be given intravenously.

Alcohol withdrawal symptoms range from mild and almost unnoticeable to severe and life threatening. Classic symptoms include autonomic hyperactivity (increased pulse rate, blood pressure, and body temperature), restlessness, disturbed sleep, anxiety, nausea, and tremor. Severe withdrawal can involve auditory, visual, or tactile hallucinations, delirium, seizures, and coma. Benzodiazepine, opioids, and cocaine are also associated with potentially life-threatening withdrawal symptoms. Physiological benzodiazepine dependence in the elderly requires very gradual tapering to avoid withdrawal symptoms, which may be different from younger age groups. More specifically, benzodiazepine withdrawal may include confusion and disorientation in older adults as opposed to anxiety, insomnia, and changes in perception. Prescribers should also be aware that withdrawal symptoms have been demonstrated to have a longer duration in older adults and can potentially complicate other medical and psychiatric conditions. Nevertheless, there is no evidence currently suggesting increased susceptibility to withdrawal or longer treatment for withdrawal symptoms in the elderly population.

Moderators and Correlates of Treatment

Certain factors may impact treatment response and adherence in older adults receiving treatment. Age-appropriate treatment is considered important in improving treatment completion and attendance rates at group meetings as compared to treatments that do not consider the unique aspects of aging. Older adults in a study of male veterans with alcohol

problems randomly assigned after detoxification to age-specific or standard mixed-age treatment were 2.9 times more likely at 6 months and 2.1 times more likely at 1 year to report abstinence when in the age-appropriate treatment group.

Matching the treatment to the patient also affects treatment outcome, although this has not been well evaluated in older adults. In one study of older primary care patients who were at-risk drinkers, 60.4 percent of patients referred to an integrated care model using a brief intervention modality attended at least one session versus 33 percent of patients referred to specialty care. This result occurred despite elimination of copayments, availability of insurance claims at most sites, and appointment availability within 2 weeks of the initial assessment.

Certain patient-level characteristics may also impact treatment outcomes. At a 6-month follow-up assessment of patients treated at a private outpatient chemical dependency program, older women with severe alcohol use disorder were significantly more likely to report abstinence from drugs and alcohol for the previous 30 days than men (79.3 vs. 54.0 percent). Among patients who were not abstinent, men reported a mean of 4 heavy drinking days, whereas women denied any heavy drinking. Favorable outcomes for older adults have also been associated with longer treatment retention and the absence of social network members who encourage substance use.

FUTURE DIRECTIONS

There remains a significant knowledge gap regarding addiction in older adults, including the etiology and neurobiology of addiction for this population. Future studies including individuals over the age of 65 are necessary to improve the understanding of treatment outcomes in older adults, including the impact of nutrition, vitamin supplementation, and comorbid medical illnesses. Further research is needed to validate DSM-5 diagnostic criteria for substance use disorders in older adults and to define the point at which patients with substance misuse should be treated in specialty care as opposed to primary care or mental health clinics. With the aging cohort and substance misuse pattern, more attention needs to be paid to prescription medication overuse and other illicit substances use and their impact in older population. Expansion of knowledge base on the potential risks and benefits of marijuana use and misuse in the geriatric population, given the current trend of legalization and expanding use, will be important.

REFERENCES

Barry K. Center for substance abuse treatment SAMHSA. consensus panel chair, brief alcohol interventions and therapies in substance abuse treatment. In: *Treatment Improvement Protocol*. Washington, DC: U.S. Government Printing Office; 2001.

Barry KL, Blow FC, Oslin DW. Substance abuse in older adults: review and recommendations for education and practice in medical settings. *Subst Abus*. 2002;23:105–131.

Barry KL, Oslin DW, Blow FC. *Alcohol Problems in Older Adults: Prevention and Management*. New York: Springer Publishing; 2001.

Blow FC. *Short Michigan Alcoholism Screening Test—Geriatric Version (SMAST-G)*. Ann Arbor, MI: University of Michigan Alcohol Research Center; 1991.

Blow FC. *Substance Abuse Among Older Adults. Treatment Improvement Series Protocol (TIP) Series No. 26. Center for Substance Abuse Treatment*. Rockville, MD: U.S. Department of Health and Human Services; 1998.

Blow FC, Oslin DW. *Older Adults and Addiction: Clinician's Manual*. Center City, MN: Hazelden; 2007.

Choi NG, DiNitto DM, Marti CN. Treatment use, perceived need, and barriers to seeking treatment for substance abuse and mental health problems among older adults compared to younger adults. *Drug Alcohol Depend*. 2014;1(145):113–120.

Coogle CL, Owens MG. Screening and brief intervention for alcohol misuse in older adults: training outcomes among physicians and other healthcare practitioners in community-based settings. *Community Ment Health J*. 2015;51(5):546–553.

DiNitto DM, Choi NG. Marijuana use among older adults in the U.S.A.: user characteristics, patterns of use, and implications for intervention. *Int Psychogeriatr*. 2011;23(5):732–741.

Draper B, Ridley N, Johnco C, et al. Screening for alcohol and substance use for older people in geriatric hospital and community health settings. *Int Psychogeriatr*. 2015;27(1):157–166.

Han B, Gfroerer JC, Colliver JD, Penne MA. Substance use disorder among older adults in the United States in 2020. *Addiction*. 2009;104(1):88–96.

Katz IR, Sands LP, Bilker W, DiFilippo S, Boyce A. Identification of medications that cause cognitive impairment in older people: the role of oxybutynin chloride. *J Am Geriat Soc*. 1998;46(1):8–13.

Lemke S, Moos RH. Treatment and outcomes of older patients with alcohol use disorders in community residential programs. *J Stud Alcohol*. 2003;64(2):219–226.

Moore AA, Whiteman EJ, Ward KT. Risks of combined alcohol/medication use in older adults. *Am J Geriatr Pharmacother*. 2007;5(1):64–74.

Morse SA, Watson C, MacMaster SA, Bride BE. Differences between older and younger adults in residential treatment for co-occurring disorders. *J Dual Diagn*. 2015;11(1):75–82.

Oslin DW. Alcohol use in late life: disability and comorbidity. *J Geriatr Psychiatry Neurol*. 2000;13(3):134–140.

Oslin DW. Late-life alcoholism: issues relevant to the geriatric psychiatrist. *Am J Geriatr Psychiatry*. 2004;12(6):571–583.

Sanchez HA, Durai UN, Miles KM, Llorente MD, Costantino G. Alcohol consumption among older adults in primary care. *J Gen Intern Med*. 2007;22(1):92–97.

Satre DD, Blow FC, Chi FW, Weisner C. Gender differences in seven-year alcohol and drug treatment outcomes among older adults. *Am J Addict*. 2007;16(3):216–221.

Satre DD, Mertens JR, Arean PA, Weisner C. Five-year alcohol and drug treatment outcomes of older adults versus middle-aged and younger adults in a managed care program. *Addiction*. 2004;99(10):1286–1297.

Satre DD, Mertens JR, Weisner C. Gender differences in treatment outcomes for alcohol dependence among older adults. *J Stud Alcohol*. 2004;65(5):638–642.

Simoni-Wastila L, Yank HK. Psychoactive drug abuse in older adults. *Am J Geriatr Pharmacother*. 2006;4(4):380–394.

Substance Abuse and Mental Health Services Administration. *Results from the 2006 National Survey on Drug Use and Health: National Findings (Office of Applied Studies, NSDUH Series H-32, DHHS Publication No. SMA 07–4293)*. Rockville, MD: Author; 2007.

United Nations, Department of Economic and Social Affairs, Population Division. (2013). *World Population Ageing 2013*. ST/ESA/SER.A/348.

Van den Elsen GA. Efficacy and safety of medical cannabinoids in older subjects: a systematic review. *Aging Research Reviews*. 2014;14:56–64.

Wadd S, Papadopolos C. Drinking behaviour and alcohol-related harm amongst older adults: analysis of existing UK datasets. *BMC Res Notes*. 2014;7:741.

Widlitz M, Marin D. Substance abuse in older adults. an overview. *Geriatrics*. 2002;57:29–34.

▲ 57.4 Treatment of Psychiatric Disorders

57.4a Treatment of Psychiatric Disorders: General Principles

BRUCE G. POLLOCK, M.D., PH.D.

Age-related health problems lead to a disproportionate use of prescription medications in older patients. Nevertheless inappropriate prescriptions, excessive dosage, preventable drug interactions, and unrecognized adverse drug reactions contribute substantially to morbidity and mortality. It is not uncommon to find side effects from one medicine treated with another medication and for a prescriber not to be aware of all the medications a patient is taking. The average older American uses four prescription drugs and an additional four over-the-counter medications or herbal preparations putting them at particular risk for drug–drug interactions. Moreover, pharmacokinetic and pharmacodynamic changes associated with aging and illness render older patients more sensitive than younger adults to adverse effects of medications.

For the above reasons, it is perhaps not surprising that the most common treatable illness in older patients is iatrogenic. Two-thirds of admissions to an acute geriatric medical unit were found to be preceded by at least one inappropriate medication. Importantly, psychotropics are among the most common medications associated with preventable adverse drug events among the aged.

Despite, or perhaps because of, their extensive medication use, the old are often excluded from clinical and regulatory trials. Without randomized, prospective drug trials, it is not truly possible for clinicians to decide whether a treatment is efficacious or even harmful. The absence of controlled studies makes it difficult to determine the optimal drug class or

drug within a class for treating elderly psychiatric patients. Encouragingly, there has been progress in pharmacotherapy trials for the cognitive and psychiatric symptoms accompanying Alzheimer disease as well as for late-life depression and generalized anxiety disorder.

Although elderly patients frequently require lower doses of medications to mitigate side effects, age, alone, does not entirely guide appropriate dosage. Often, there is greater physiologic variation among individual elders than there are group differences between young and old. It is important to consider individual patient characteristics, such as cardiovascular or renal function and genetic and acquired differences in body composition or drug metabolism, together with a drug's physicochemical properties.

In addition to using rational pharmacotherapy, the physician must form an alliance with patients and their families to improve adherence to treatment. Nonadherence with therapy is a major problem for psychiatric patients, and this dilemma is exacerbated with age. Moreover, the cost and complexity of medication regimens, inherent to the multiple illnesses of the aged, are further complicated by communication difficulties arising from impaired hearing, cognitive impairment, and language and cultural differences.

PHARMACOKINETIC CONSIDERATIONS

Pharmacokinetics provides a way of describing and predicting drug concentrations in plasma and various tissues over time. It has been estimated that 70 to 80 percent of adverse drug reactions are dose, and therefore concentration, related. Concentration relationships for adverse effects were clearly evident with tricyclic antidepressants (TCAs) and neuroleptics. Cholinesterase inhibitors have been associated with dose-related, cholinergic (nausea, vomiting, diarrhea) side effects, which can be particularly severe if doses are escalated too rapidly, or if cytochrome P450 (CYP) 3A4 metabolism is inhibited for those treated with donepezil (Aricept) or galantamine (Reminyl). For newer antidepressants and antipsychotics, explorations of drug concentration variability, let alone response relationships, were deemed unnecessary. Nonetheless, therapeutic indices for most drugs narrow with age and regulatory trials conducted in younger, healthier subjects are often conducted at or near their maximum-tolerated dosages to increase the probability of demonstrating efficacy. Failure to address this issue resulted in notable difficulties with the introduction of bupropion (Wellbutrin) and clozapine (Clozaril) (seizures at high doses), as well as fluoxetine (Prozac) and risperidone (Risperdal) (agitation, akathisia, extrapyramidal syndromes with excessive doses) in older patients. Olanzapine (Zyprexa) has shown a curvilinear dose–response relationship for treating elderly patients and quetiapine (Seroquel) is currently used over a wide dose range in elders. In

fact, more than 25 percent of neuropharmacologic drugs introduced in the past two decades have undergone recommended postmarketing dosage reductions for safety reasons.

The selective serotonin reuptake inhibitors (SSRIs) are first-line treatments for depression in the elderly. Although they have relatively flat dose–response curves, unnecessarily high SSRI drug concentrations may increase side effects and cost. For example, SSRI plasma concentrations have been linearly associated with adverse cognitive and psychomotor effects in older subjects. Dose-dependent effects of SSRIs have also been reported for fragility fractures, falls, and decrements in bone density.

It is important to appreciate the critical lack of pharmacokinetic information relevant to the majority of elders encountered in clinical practice. The limited information that does exist for older subjects is largely derived from "classic" or "two-stage" pharmacokinetic modeling. Although single- and multiple-dose pharmacokinetic studies are a standard part of phase I drug development, they are almost exclusively carried out on young, physically healthy, unmedicated men. For example, the only published geropharmacokinetic data on the disposition of fluoxetine (which has a geriatric indication) is limited to a study of single doses in 11 healthy subjects (ranging in age from 65 to 77 years). Traditional pharmacokinetic methodology requires an intensive sampling regimen (i.e., a large number of precisely timed samples) usually obtained from a small number (e.g., 6 to 12) of volunteers. The pharmacokinetic parameters are determined for each individual separately and average pharmacokinetic parameters are calculated as the mean of the individually determined parameters. Single-dose studies often fail to provide accurate estimates of pharmacokinetic parameters for older patients. Multiple-dose kinetic studies are especially important to detect nonlinear phenomena wherein metabolic clearance ceases to be proportional to dose.

An alternative method is that of population pharmacokinetics using mixed-effects modeling. This statistical approach permits identification of individual and overall population pharmacokinetic parameters based on only a few samples per subject and using each data point to inform the entire analysis. Because a greater number of typical patients are evaluated, the results are much more applicable to clinical practice than those of traditional pharmacokinetics. Incorporating population pharmacokinetic analyses into clinical trials with older subjects may prove to be a significant addition to the ability to determine sources of variability in psychotropic plasma concentrations and response in elders, particularly those with medical comorbidity.

Age-associated pharmacokinetic differences may be due to changes in absorption, distribution, metabolism, or elimination of a drug.

Absorption

With age, there is a decrease in the rate of gastric emptying and intestinal blood flow. Most studies evaluating drug absorption, however, have shown no difference in the rate or extent of absorption in geriatric versus younger patients. Nevertheless, antacids, high-fiber supplements, laxatives, and the cholesterol-binding resin cholestyramine (Questran) may significantly diminish the absorption of medications, such as TCAs and anticonvulsants. Drugs with anticholinergic properties may decrease absorption of L-dopa (Larodopa), owing to delayed gastric emptying in conjunction with the gastric metabolism of L-dopa. The potential gain in compliance achieved by administering medications at the same time should be weighed against the possibility of drug interactions in the absorption phase.

There is increasing interest in the genetic differences and drug interactions affecting functioning of the important drug efflux transporter, P-glycoprotein (MDR1), which can transport a large number of clinically used drugs out of the intestine, thereby playing a significant role in absorption. For example, variants in the *ABCB1* gene, which encodes P-glycoprotein, have been found to influence the bioavailability of digoxin (Lanoxin).

Distribution

Clinicians frequently overlook the fact that a drug's half-life is directly proportional to its apparent volume of distribution. Adipose tissue increases from 33 percent of body mass in younger women to 48 percent in the old, and from 18 to 36 percent in men. For drugs that distribute in body fat (the vast majority of psychotropics), the volume of distribution increases with age, leading to a longer half-life and drug accumulation. For water-soluble drugs, such as lithium or digoxin, volumes of distribution will be diminished in older patients, reducing the margin of safety (i.e., transient increases of these drugs' plasma concentrations will not be buffered to the same degree in elders). Lipid-soluble benzodiazepines (e.g., diazepam [Valium] and clonazepam [Klonopin]) with long half-lives have been associated with an increased risk of falls and hip fractures in contrast to more water-soluble benzodiazepines with intermediate half-lives, such as lorazepam (Ativan).

The volume of distribution is also affected by the extent of drug bound to plasma proteins. Age-associated illness may cause reductions in serum albumin and possible increases in α-1-acid glycoprotein. The majority of psychotropics are transported in the blood extensively bound to these plasma proteins. Drug displacement interactions from plasma proteins were previously the subject of considerable concern, because the protein-bound drug acts as an enormous, pharmacologically active reservoir. It is now recognized that not only is free drug pharmacologically active, but it is also more available for metabolism and tissue distribution. Therefore, an

increase in the absolute concentration of free drug, caused by reduction in plasma proteins or a displacement interaction, would be buffered and equilibrium rapidly reestablished if there is no concurrent metabolic inhibition caused by the interacting drug. When plasma drug monitoring is used to adjust TCA or anticonvulsant doses, it should be appreciated that the total drug concentrations (free + protein bound) are usually reported. In the face of hypoalbuminemia or a displacement interaction (e.g., phenytoin and aspirin), the total drug level may appear low and doses may, in consequence, be inappropriately increased. If the displacing drug also causes inhibition of drug metabolism (e.g., valproic acid [Depakote] and aspirin), total drug levels may appear to be acceptable, while free drug concentrations are in fact significantly elevated. The use of unbound (free) drug levels in these situations has been found to be essential for lidocaine (Xylocaine), theophylline (Theo-Dur), valproic acid, phenytoin (Dilantin), and digitoxin (Crystodigin).

The blood–brain barrier forms an interface between the circulating blood and the brain and possesses various carrier-mediated transport systems for small molecules to support and protect central nervous system (CNS) function. For example, the blood-to-brain influx transport systems supply nutrients, such as glucose and amino acids. Age and illness, such as diabetes, may affect the integrity of the blood–brain barrier, increasing the drug exposure of the brain. P-glycoprotein, located at the luminal membrane of endothelial cells, functions as a clearance system for drugs entering the brain as well as for metabolites and neurotoxic compounds produced in the brain. Drugs recognized by these transporters are expected to show lower blood–brain barrier permeability. Examples of drugs now known to be substrates for P-glycoprotein include digoxin, erythromycin (Ery-Tab), lidocaine, morphine (Astramorph), nifedipine (Adalat), nortriptyline (Pamelor, Aventyl), propranolol (Inderal), risperidone, and verapamil (Calan). Several P-glycoprotein substrates may also inhibit its function, such as erythromycin, risperidone, and verapamil, permitting greater CNS penetration.

Metabolism

The purpose of drug metabolism is to transform lipid-soluble compounds into more water-soluble derivatives that can be more effectively eliminated and may be less active pharmacodynamically. Biotransformation of drugs is categorized as belonging to either phase I (oxidation or reduction) or phase II (conjugation). A large group of heme-containing enzymes, which vary in substrate preferences, collectively called CYP enzymes, is responsible for oxidative (phase I) drug metabolism. As more is learned about both the diversity and specificity of the CYP enzymes, a structure is emerging for categorizing potential drug–drug interactions and possible age-associated changes. Available evidence suggests that there is not a uniform age-

associated decline in the quality of activity of these enzymes. Nonetheless, reductions in hepatic mass and blood flow with aging place greater emphasis on interindividual differences in drug metabolic capacity. These metabolic differences may be genetic or the result of interactions from multiple medications. Isoenzyme specificity suggests that inhibition or induction of a given CYP will affect all drugs metabolized by that specific enzyme. A drug may act as a substrate for one CYP at a time or simultaneously for more than one. A drug can act as both a substrate and a competitive inhibitor for the same isozyme, or it can act as a noncompetitive inhibitor for a different enzyme.

CYP 2D6 is the isozyme specifically responsible for hydroxylation of many antidepressants (e.g., nortriptyline, desipramine [Norpramin, Pertofrane], duloxetine [Cymbalta], paroxetine, venlafaxine [Effexor], and vortioxetine [Brintellix]) and several antipsychotics (perphenazine [Trilafon], thioridazine [Mellaril], iloperidone [Fanapt], and risperidone). Approximately 7 percent of the Caucasian population are genetically poor 2D6 metabolizers, which has been shown to impact on nortriptyline doses and adverse effects of perphenazine in older patients. Concern has also been raised about the potential cardiotoxicity of venlafaxine in poor 2D6 metabolizers or those taking concomitant diphenhydramine (a CYP 2D6 inhibitor). Studies conducted in older, unmedicated volunteers have not found an age-associated decline in CYP 2D6 activity. Nevertheless, many drugs commonly used by older patients competitively (e.g., perphenazine) or noncompetitively (e.g., paroxetine, fluoxetine, bupropion) inhibit this enzyme, leading to potentially significant interactions.

CYP 2C19 is responsible for the metabolism of diazepam, mephenytoin (Mesantoin), escitalopram (Lexapro), and omeprazole (Prilosec), as well as several barbiturates. In addition, the demethylation of tertiary TCAs is also partly dependent on 2C19. The incidence of the poor metabolizer phenotype in a young Caucasian population is approximately 3 percent but was 18 percent in a Japanese sample. There is some evidence that 2Cl9 functioning may decline with age. CYP 2C19 is responsible for metabolizing the narrow-therapeutic-index drugs phenytoin, tolbutamide (Orinase), and warfarin (Coumadin), and is inhibited by fluvoxamine (Luvox) and fluoxetine.

CYP 3A4 is the most prevalent drug-metabolizing enzyme and is responsible for metabolizing the most medications. Because it is normally a high-capacity enzyme, serious and surprising toxicity has occurred when the 3A4-mediated clearance of terfenadine (Seldane), astemizole (Hismanal), cisapride (Propulsid), midazolam (Versed), triazolam (Halcion), mibefradil (Posicor), and cerivastatin (Baycol) was diminished. Also, because hepatic metabolism of CYP 3A4 drugs is typically so rapid, many of these drugs are perfusion limited (i.e., dependent on hepatic blood flow, which may decline by 40 percent between ages 25 and 65 years).

Therefore, it often appears that drugs specifically metabolized by CYP 3A4 decline with age—for example, alprazolam (Xanax) and triazolam clearances in the elderly are only 50 to 80 percent of those in young adults. In particular, gender- and age-related pharmacokinetic differences have been demonstrated for several 3A4-typed drugs, such as sertraline (Zoloft), mirtazapine (Remeron), and nefazodone (Serzone). CYP 3A4 activity may be inhibited by grapefruit juice, protease inhibitors, "mycin" antibiotics, and "azole" antifungals. Among antidepressants, nefazodone and fluvoxamine are the most potent inhibitors, followed by fluoxetine, through its demethylated metabolite. The very long half-life of norfluoxetine may result in interactions occurring many weeks after initiation of fluoxetine. Similarly, the inhibition of the CYP 3A4-mediated clearance of "statins," resulting in muscle weakness, may be misattributed in depressed patients. The 3A4 enzyme also may be induced by other drugs, such as carbamazepine (Tegretol), phenytoin, topiramate (Topamax), modafinil (Provigil), barbiturates, steroids, and St. John's wort, increasing the risk of therapeutic failure for CYP 3A4 drugs. Many CYP 3A4 inhibitors (e.g., verapamil) and inducers (St. John's wort) also have been found to interact at the P-glycoprotein drug transporter, amplifying their CYP 3A4-mediated effects.

CYP 1A2 metabolizes clozapine, olanzapine, fluvoxamine, asenapine (Saphris), and theophylline and contributes to the demethylation of some tertiary TCAs. This isozyme is induced by cigarette smoking and eating cruciferous vegetables and charcoaled meats, as well as by certain medications, such as omeprazole and phenobarbital (Luminal). Estrogen replacement therapy in postmenopausal women has been found to inhibit CYP 1A2 metabolism, whereas treatment with growth hormone causes induction. Inhibition and induction magnify an already marked, 60-fold variation in 1A2 activity. This degree of metabolic variation mandates that theophylline and clozapine be closely monitored.

There is more limited information about CYP 2B6, which substantially metabolizes bupropion. In vitro evidence has suggested that fluoxetine, paroxetine, and sertraline may cause CYP 2B6 inhibition. Given bupropion's risk of concentration-related seizures and the frequency of its use with SSRIs, this may be of potential concern.

Beyond oxidative metabolism, phase II metabolism in older patients has been a neglected area for systematic study. In general, drugs such as lorazepam and oxazepam (Serax) accumulate less than more lipid-soluble drugs and undergo single-step conjugative elimination. Glucuronidation, however, may not be an immutable process, being significantly affected by malnutrition. Despite the common assertion that phase II metabolism does not change with age, significant reductions in oxazepam's elimination have been found in those older than 80 years.

Excretion

The average glomerular filtration rate decreases by 50 percent between the ages of 20 and 70 years. Age alone, however, cannot be used to predict renal function, because at least one-third of older patients may maintain relatively stable renal function. Serum creatinine may not appear elevated in the face of declining renal function because there is often a decline in muscle mass, the major source of creatinine, with age and debilitation. Minimally, the use of estimates from standard nomograms or equations should be used to estimate creatinine clearance, given that it is often impractical to collect a timed urine sample. Decline in renal clearance may directly affect excretion of such medications as lithium, gabapentin (Neurontin), pregabalin (Lyrica), topiramate, amantadine (Symmetrel), paliperidone (Invega), memantine (Namenda), rivastigmine (Exelon), and desvenlafaxine (Pristiq). The clearance of hydrophilic (polar), psychotropic drug metabolites may also be reduced in older patients. Paliperidone is the active 9-OH metabolite of risperidone and accumulations of active hydroxylated metabolites of antidepressants, such as nortriptyline, bupropion, and venlafaxine, have been observed in older patients with uncertain clinical consequences.

The nonsteroidal anti-inflammatory drugs (NSAIDs) and angiotensin-converting enzyme (ACE) inhibitors reduce lithium's clearance (possibly through effects on renal blood flow). Given the widespread use of these medications in elders and their increased risk of lithium toxicity, this potential interaction deserves emphasis. Low-salt diets and the possibility that SSRIs may cause inappropriate antidiuretic hormone secretion, with concomitant hyponatremia, are additional sources of interactive complications with lithium.

PHARMACODYNAMICS

As patients age, there is a general reduction in homeostatic mechanisms (e.g., postural control, orthostatic circulatory responses, thermoregulation, visceral muscle function, laryngeal reflexes, hypoxic responses, cognitive function). This reduction may interfere with adaptation to environmental changes and may be manifest as adverse drug experiences. All psychotropics may increase risk of falls and hip fractures and may increase cognitive impairment. Physiologic (postural or action) tremors may be exacerbated by numerous drugs, including β and dopamine agonists, xanthines, valproic acid (valproate), steroids, and antidepressants. Reduced β-receptor sensitivity with aging contributes to diminished baroreceptor reflex responses, whereas coronary artery disease, reductions in the heart's intrinsic sinus rate, and damage to conduction pathways predispose to arrhythmias. The fact that SSRIs may exacerbate bradycardia should also be borne in mind. In older patients chronically treated with

antipsychotics, the prevalence of obesity, poor nutritional status, and increased triglycerides may contribute additional cardiovascular risk factors.

Specific receptor and neurotransmitter changes associated with senescence, such as reductions in dopamine or acetylcholine content, result in increased sensitivity to D2 antagonists and antimuscarinic agents. Side effects from conventional antipsychotics and tricyclics were common and frequently troubling in older patients. In particular, the association of these medications with significant cognitive impairment, declines in activities of daily living, and significant falls (with concomitant risk of hip fracture) was of grave concern in the frail elderly. Prospective studies demonstrated that there was a very rapid and high incidence of onset of tardive dyskinesia in older patients taking conventional antipsychotics for relatively brief periods (cumulative incidence reported as 26 to 45 percent over 1 year). Although atypical antipsychotics are now extensively used in older patients, meta-analyses of 17 controlled trials found that these medications increase mortality (4.5 percent) compared to placebo (2.6 percent) in elders suffering from dementia within the relatively short duration of the trials conducted (typically 10 weeks). Moreover, the side effects of these medications may outweigh their potential benefits. Some of these side effects are concentration related, such as extrapyramidal symptoms with risperidone and anticholinergicity with clozapine or olanzapine. Concentrations of atypical antipsychotics may be significantly affected by concomitantly administered medications that inhibit or induce CYPs lA2 (asenapine, clozapine, olanzapine), 3A4 (aripiprazole, iloperidone, quetiapine, ziprasidone), or 2D6 (aripiprazole, risperidone).

Two groups of drugs cause anticholinergic effects: those used therapeutically for their anticholinergic properties (e.g., oxybutynin [Ditropan], benztropine [Cogentin]) and those that have anticholinergic effects that are incidental to their main actions (e.g., amitriptyline [Elavil], cimetidine [Tagamet], digoxin, paroxetine). Peripheral manifestations of anticholinergicity include dry mouth, tachycardia, blurring of vision, urinary retention, and constipation. Common difficulties in older patients, such as constipation, xerostomia, glaucoma, and urinary retention, may be markedly exacerbated by even modestly anticholinergic medications. Anticholinergic effects on the CNS may range from subtle changes, such as mild confusion and impairment of recent memory, to physical decline. Patients with a dementing illness are most at risk from these central anticholinergic effects. Nonetheless, studies have found that even medications with modest anticholinergicity are additive and affect cognitive performance in a wide cross section of elders. In fact, understanding of the nature of cognitive decrements associated with “normal” aging may be distorted by modest anticholinergic burden. It should also be noted that medications with prominent noradrenergic

actions can also cause dry mouth, urinary retention in older men, and tachycardia, in addition to exacerbating hypertension and tremor.

Older patients have been found to require smaller doses of anticoagulants to achieve the same degree of anticoagulation as obtained in younger individuals. Although confounded by effects of other medications and hepatic dysfunction, increased sensitivity of elders to anticoagulants is probable. Anticoagulants figure prominently in drug-related morbidity and mortality. Genetic tests may, in the near future, help to detect the estimated one-third of patients with variants that make them slow metabolizers of warfarin (CYP 2C9) or more sensitive to its effects (vitamin K epoxide reductase activity [VKORC1]). Anticoagulant–antidepressant interactions may be pharmacokinetic as well as pharmacodynamic. Fluvoxamine poses a risk for a pharmacokinetic interaction through inhibition of CYP 2C9, reducing the clearance of warfarin's active S-enantiomer. In addition, inhibition of CYPs 1A2 (fluvoxamine), 2C19 (fluvoxamine, fluoxetine), and 3A4 (fluvoxamine, fluoxetine) will cause R-warfarin to accumulate, which also contributes to reduced S-warfarin clearance. However, increased bleeding times with SSRIs alone or in combination with anticoagulants also may be possible owing to SSRIs' effects on platelet serotonin and thereby attenuating aggregation. Concurrent therapy need not be avoided, but close monitoring is advisable whenever an SSRI is combined with warfarin. Both concerns (increased risk of gastrointestinal bleeding in elders) and potential benefits (protective effects against myocardial infarction in smokers) have been raised by epidemiological studies of SSRIs.

In addition to problems that may be anticipated, because most of the data on a new drug's effects have been obtained in younger patients, unexpected side effects are not infrequent in elders. In older patients, SSRIs and venlafaxine have been associated with the syndrome of inappropriate antidiuretic hormone secretion (SIADH). Although SIADH is a rare adverse effect of many psychotropic medications, the majority of reports of antidepressant-induced SIADH have involved patients older than 65 years. In an analysis of 760 reports of hyponatremia associated with SSRIs, the median time to onset of the hyponatremia was 13 days (range, 3 to 120 days). SSRIs in older patients also have been associated with parkinsonism or worsening of motor disability in patients who have idiopathic Parkinson disease. It is possible that this is a result of serotonergic-influenced reduction in dopaminergic transmission in a population with diminished dopamine reserves. Concern has also been raised recently from large population studies. One found chronic use of SSRIs associated with a twofold increased risk of fragility fractures and lower bone mineral density. Another associated SSRI therapy with a nearly fivefold increased risk of completed suicide in older patients during the first month of treatment compared with other antidepressants. Although the absolute risk of suicide may be low, there is a clear need for close

monitoring of older patients, especially males with comorbid medical illness.

The greatest threats to successful pharmacotherapy in older patients are patient nonadherence with medication regimens and drug–drug interactions. High-risk medications, including anticoagulants, antihypertensives, and antiarrhythmics, have been consistently emphasized in long-term, comprehensive surveys of adverse drug reactions in older patients. Physicians need to be particularly vigilant and knowledgeable when psychotropics are coprescribed with these high-risk medications. As patients age, pharmacodynamic changes and concomitant illness will significantly alter their response to medication. Psychiatrists treating older patients should use a set of routine procedures as follows: (1) identify patient risk factors, such as medical status, use of high-risk medications, diet, and alcohol use; (2) take a thorough drug history involving caregivers and relatives, being sure to include over-the-counter medications and update regularly in a prominent place in the patient's record; (3) educate patients and caregivers about their medications and dosing regimens and monitor adherence; (4) be knowledgeable about drug actions, interactions, and clearance; (5) use the lowest effective doses and carefully graduated increases, and avoid unnecessary polypharmacy, but do not undertreat; (6) frequently reassess psychomotor function, including mental status and balance, as well as orthostatic blood pressures and specific laboratory tests when appropriate; and (7) if there is a deterioration in patients' physical or cognitive states, always suspect an adverse drug reaction.

FUTURE DIRECTIONS

Developments in clinical trial design and the completion of large effectiveness trials serve further to heighten concerns about the efficacy and side effects of currently available pharmacologic therapy for the symptomatic relief of psychiatric symptoms in elders. Controlling for heterogeneous drug exposure in conjunction with improved understanding of the genomic basis of cognitive, mood, and psychotic dysfunction and its complex interactions with age-related ischemic and other environmental CNS changes may lead to more targeted molecular therapeutics. Regulatory policy and public health initiatives must be amplified to increase the representation of older subjects in clinical trials and improve the relevance of the therapeutic evidence base. For medications approved for release, it is to be hoped that the cross-linking of computerized medical records will lead to improved postmarketing surveillance of medications in elders.

REFERENCES

*American Geriatrics Society 2015 Beers Criteria Update Expert Panel. American Geriatrics Society updated Beers criteria for potentially inappropriate medication use in older adults. *J Am Geriatr Soc*. 2015;63:2227–2246.

Andrade C, Sandarsh S, Chethan KB, Nagesh KS. Serotonin reuptake inhibitor antidepressants and abnormal bleeding: a review for clinicians and a reconsideration of mechanisms. *J Clin Psychiatry*. 2010;71:1565–1575.

Bellelli G, Morandi A, Davis DH, et al. Validation of the 4AT, a new instrument for rapid delirium screening: a study in 234 hospitalised older people. *Age Ageing*. 2014;43:496–502.

Budnitz DS, Lovegrove MC, Shehab N, Richards CL. Emergency hospitalizations for adverse drug events in older Americans. *N Engl J Med*. 2011;365:2002–2012.

Chew ML, Mulsant BH, Pollock BG, et al. Anticholinergic activity of 107 medications commonly used by older adults. *J Am Geriatr Soc*. 2008;56:1333–1341.

Danielsson B, Collin J, Jonasdottir Bergman G, Borg N, Salmi P, Fastbom J. Antidepressants and antipsychotics classified with torsades de pointes arrhythmia risk and mortality in older adults —a Swedish nationwide study. *Br J Clin Pharmacol*. 2016;81:773–783.

*Flockhart D. *Cytochrome P450 Drug Interaction Table*. Indiana University School of Medicine, Division of Clinical Pharmacology: http://medicine.iupui.edu/clinpharm/ddis/main-table. Accessed January 9, 2017.

*Fox C, Smith T, Maidment I, et al. Effect of medications with anticholinergic properties on cognitive function, delirium, physical function and mortality: a systematic review. *Age Ageing*. 2014;43:604–615.

Frankenthal D, Lerman Y, Kalendaryev E, Lerman Y. Intervention with the screening tool of older persons potentially inappropriate prescriptions/screening tool to alert doctors to right treatment criteria in elderly residents of a chronic geriatric facility: a randomized clinical trial. *J Am Geriatr Soc*. 2014;62:1658–1665.

Ho T, Pollock BG, Mulsant BH, et al. R- and S-citalopram concentrations have differential effects on neuropsychiatric scores in elders with dementia and agitation. *Br J Clin Pharmacol*. 2016;82:784–792.

Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet*. 2014;383:911–922.

Jin Y, Pollock BG, Frank E, et al. Effect of age, weight and CYP2C19 genotype on escitalopram exposure. *J Clin Pharmacol*. 2010;50:62–72.

*Looper KJ. Potential medical and surgical complications of serotonergic antidepressant medications. *Psychosomatics*. 2007;48:1–9.

Mark TL, Joish VN, Hay JW, Sheehan DV, Johnston SS, Cao Z. Antidepressant use in geriatric populations: the burden of side effects and interactions and their impact on adherence and costs. *Am J Geriatr Psychiatry*. 2011;19:211–221.

Meyers BS, Jeste DV. Geriatric psychopharmacology: evolution of a discipline. *J Clin Psychiatry*. 2010;71:1416–1424.

Nebes RD, Pollock BG, Perera S, Halligan EM, Saxton JA. The greater sensitivity of elderly APOE ε4 carriers to anticholinergic medications is independent of cerebrovascular disease risk. *Am J Geriatr Pharmacother*. 2012;10:185–192.

Pollock B, Forsyth C, Bies R. The critical role of clinical pharmacology in geriatric psychopharmacology. *Clin Pharmacol Ther*. 2009;85:89–93.

Pollock BG, Ferrell RE, Mulsant BH, et al. Allelic variation in the serotonin transporter promoter affects onset of paroxetine treatment response in late-life depression. *Neuropsychopharmacology*. 2000;23:587–590.

Pollock BG, Laghrissi-Thode F, Wagner WR. Evaluation of platelet activation in depressed patients with ischemic heart disease after paroxetine or nortriptyline treatment. *J Clin Psychopharmacol*. 2000;20:137–140.

Pollock BG, Mulsant BH. Between Scylla and Charybdis: antipsychotic and other psychotropic medications in older nursing home residents. *CMAJ*. 2011;183:778–779.

Porsteinsson AP, Drye LT, Pollock BG, et al; CitAD Research Group. Effect of citalopram on agitation in Alzheimer disease: the CitAD randomized controlled trial. *JAMA*. 2014;311:682–691.

Qato DM, Alexander GC, Conti RM, Johnson M, Schumm P, Lindau ST. Use of prescription and over-the-counter medications and dietary supplements among older adults in the United States. *JAMA*. 2008;300:2867–2878.

*Reuben DB, Herr KA, Pacala JT, Pollock BG, Potter JF, Semla TP. *Geriatrics at Your Fingertips*. 18th ed. New York: The American Geriatrics Society; 2016.

Richards JB, Papaioannou A, Adachi JD, et al; Canadian Multicentre Osteoporosis Study Research Group. Effect of selective serotonin reuptake inhibitors on the risk of fracture. *Arch Intern Med*. 2007;167:188–194.

Sandson NB, Marcucci C, Bourke DL, Smith-Lamacchia R. An interaction between aspirin and valproate: the relevance of plasma protein displacement drug-drug interactions. *Am J Psychiatry*. 2006;163:1891–1896.

Sjöberg C, Bladh L, Klintberg L, Mellström D, Ohlsson C, Wallerstedt SM. Treatment with fall-risk-increasing and fracture-preventing drugs before and after a hip fracture: an observational study. *Drugs Aging*. 2010;27:653–661.

Snowdon J, Day S, Baker W. Current use of psychotropic medication in nursing homes. *Int Psychogeriatr*. 2006;18:241–250.

Steinmetz KL, Coley KC, Pollock BG. Assessment of geriatric information on the drug label for commonly prescribed drugs in older people. *J Am Geriatr Soc*. 2005;53:891–894.

Tjia J, Briesacher BA, Peterson D, Liu Q, Andrade SE, Mitchell SL. Use of medications of questionable benefit in advanced dementia. *JAMA Intern Med*. 2014;174:1763–1771.

Wallerstedt SM, Gleerup H, Sundström A, Stigendal L, Ny L. Risk of clinically relevant bleeding in warfarin-treated patients–influence of SSRI treatment. *Pharmacoepidemiol Drug Saf*. 2009;18:412–416.

57.4b Antidepressants and Mood Stabilizers

CARL SALZMAN, M.D.

GOALS AND RATIONALE FOR THE PHARMACOLOGICAL TREATMENT OF DEPRESSION IN LATER LIFE

Clinically significant depression in late life is a **disorder that encompasses a spectrum of severity** ranging from debilitating major depression that has serious consequences for the life of the patient to minor depression that can impair quality of life. As in younger people, the course of depression in the elderly is characterized by exacerbations and remissions as well as by chronicity. Multiple pathways contribute to late-life depression, including neurobiological and psychological factors, physical illnesses and their treatments, cognitive and sensory decline, as well as common social and economic problems. Among community-dwelling residents over the age of 65, approximately 13 to 27 percent may suffer from subsyndromal depression, 15 percent exhibit depressive symptoms but do not meet diagnostic criteria for any disorder, 1 to 2 percent suffer from major depression, and 2 percent are living with dysthymia. The prevalence of depression of all forms is much higher in nursing homes and in hospitals.

There appears to be considerable heterogeneity in the onset and course of late-life depression. Diagnostic patterns are also variable among older individuals. Clues to depression include frequent medical visits with persistent complaints of pain, headache, fatigue, and insomnia; and high utilization of services. Additional clues to depression in elderly hospitalized patients include delayed recovery, treatment refusal, and problems with discharge. In the nursing home patient, depression should be suspected in the presence of apathy, failure to thrive, agitation, and delayed rehabilitation. Elderly individuals who are depressed may also have structural and functional changes in the brain including ventriculomegaly, white matter hyperintensities, and vascular changes. Depression in later life should be treated because it usually persists in the absence of treatment, causing distress to patients' caregivers, as well as leading to excess health care use and increased substance abuse. Mood disorders also play a greater role in suicides of elderly patients than in those of younger patients; the group at highest risk for suicide is white males aged 80 to 84 years. A majority of suicide victims visited a primary care physician in the last month of life.

Treatment of depression in elderly adults may improve day-to-day functioning and self-care, and can improve physical health and cognitive functioning. Although treatment benefits usually outweigh the risks, late-life depression is often untreated, especially in the frail elderly either

because the depression is not recognized or from fear that treatment will further damage an older individual. The risks of no treatment or of inadequate treatment, however, include chronic depression, cognitive impairment, worsening of physical illness, poor compliance with medical treatment, social impairment, reduced quality of life (satisfaction, ability to cope), and suicide. **Anxiety is a common component of late-life depression and high levels of anxiety may delay the response to antidepressant treatment.**

The goals of pharmacological treatment of both major and minor depressive disorders in later life include remission, if not substantial reduction, of depressive symptoms, as well as the prevention of relapse and recurrence. Hence, long-term use of **antidepressants** is probably the rule, not the exception. Treatment goals also include restoring psychosocial functioning, relieving excess disability, improving patient compliance with other medical therapies that **might be refused** because of hopelessness or pessimism caused by depression. **Antidepressants may also** ease the adaptation to irreversible losses. Expert **consensus guidelines now recommend that antidepressant treatment should be combined with psychotherapy. Since elderly individuals with mood disorders are often taking multiple medications, it is also necessary for the prescribing clinician to remain alert for possible adverse drug interactions. These will be discussed in each of the following sections.**

PHARMACOLOGICAL TREATMENT OF DEPRESSION

Antidepressants used to treat elderly depressed adults include **the same drugs that are given to younger adults. In general,** aging results in altered metabolism and excretion of some antidepressants, **prolonging their effect. For example,** the clearance of some antidepressants such as citalopram **and nefazodone** is delayed, causing drug accumulation and high blood levels. The sensitivity of the central nervous system to **antidepressant drugs is also increased by age: m**ost (but not all) elderly individuals are more sensitive to the therapeutic effects of antidepressants as well as to their side effects. **Since elderly individuals are often taking multiple medications, there is an increased risk of adverse drug interactions with antidepressants.**

Accumulating data suggest that elderly patients respond **as** well to antidepressant treatments **as younger adults.** Data **also** suggest that **the** time to response does not differ from younger adults, **but clinicians often notice that it may** require longer time periods to achieve remission or full therapeutic response.

SELECTIVE SEROTONIN REUPTAKE INHIBITORS

SSRI antidepressants are now considered **to be** the first choice for the

treatment of elderly depression because of their efficacy and favorable side-effect profile. Studies of SSRIs do not demonstrate superior antidepressant efficacy of any one member of this class over another. Late-life depressions associated with dementia as well as with medical illness respond to SSRIs. Most SSRIs can be taken once daily which increases compliance especially for elderly patients who are taking multiple medications. **Although** SSRIs do not cause orthostatic hypotension, they have been implicated in an increased prevalence of falls in elderly nursing home residents. **At therapeutic doses, they are** less likely to cause cognitive impairment, or anticholinergic and cardiovascular adverse effects. The limitations of these agents include nausea, headache, restlessness, sexual dysfunction, decreased heart rate, occasional insomnia, and occasionally inappropriate antidiuretic hormone secretion. Some SSRIs, especially fluoxetine and paroxetine (and sertraline, to a lesser extent), interfere with cytochrome-metabolizing enzymes. Since hepatic metabolism is the primary route for drug detoxification and elimination, inhibition of these enzymes (especially CYP450 2D6 and 3A4) may raise the blood levels of other medications, causing unwanted side effects and toxicity. Medications whose metabolism may be inhibited include cardiac antiarrhythmic drugs, antidepressants and neuroleptics, **some** benzodiazepine anxiolytics, steroids, some antihistamines and antibiotics, and antifungal agents. **Fluoxetine and paroxetine, in particular, are strong inhibitors of the metabolizing hepatic enzyme CYP450 2D6. This enzyme is necessary to convert tamoxifen to its active metabolite; inhibition by either of these SSRI antidepressants will prevent this conversion and render the tamoxifen useless. Citalopram and escitalopram do not affect hepatic metabolizing enzymes and may be less likely to cause drug interactions.**

Examination of the role of polymorphisms of the serotonin transporter protein gene indicates that in elderly individuals (as in young adults) the "s" variant is associated with a decreased reduction in serotonin reuptake and a less robust response to SSRI antidepressants.

Starting doses of SSRIs should be low (paroxetine 10 mg, sertraline 25 mg, citalopram 10 mg, **escitalopram** or fluoxetine, 10 mg) **followed by** gradual dosage increases, particularly in those above age 75. **Therapeutic** dosages will **vary,** ranging from 20 to 60 mg a day of paroxetine, 50 to 200 mg a day of sertraline, and 10 to 60 mg a day of fluoxetine or of citalopram **or escitalopram.** Therapeutic blood concentrations of SSRIs are not available for routine clinical use. Studies of augmentation of an SSRI antidepressant with other medications such as methylphenidate, bupropion, thyroid, or estrogen have shown **equivocal efficacy in the elderly.**

Newer antidepressants with selective serotonin reuptake inhibiting properties are now available. All are effective antidepressants; there are no

data to suggest that any are superior to existing SSRIs. These drugs include vilazodone (given in the usual daily dose of 40 mg taken with food) which combines serotonin reuptake blockade with direct effects on synaptic serotonin receptors.

SSRIs, but not tricyclic antidepressants, are associated with an increased rate of bone loss in older women, **especially in** the hip. This bone loss is associated with a twofold increase **of** clinical fragility fracture since there is an increased risk of falls in nursing home residents who are taking SSRI (or tricyclic) antidepressants.

DUAL-ACTING ANTIDEPRESSANTS

There are nine currently available antidepressants that enhance neuronal transmission in central nervous system mood-regulatory circuits by inhibiting the presynaptic reuptake of both norepinephrine and serotonin: venlafaxine, desvenlafaxine, duloxetine, milnacipran, levomilnacipran, mirtazapine, trazodone, nefazodone, and the recently released vortioxetine.

Venlafaxine, an inhibitor of noradrenaline and serotonin reuptake, is an effective and safe antidepressant for the elderly. Compared with other antidepressants, it may induce slightly higher rates of remission of depressive symptoms in nonelderly adults; data are **scant** for the elderly. Patients respond to a relatively wide range of daily doses, and no dosage adjustment is necessary based on age. The use of venlafaxine entails several precautions. Headache and nausea may be minimized by using with lower dosages of 25 mg daily and using the extended-release formulation, gradually titrating up to a therapeutic dosage. Venlafaxine may cause modest increases in diastolic blood pressure which should be monitored. Venlafaxine use is not associated with drug interactions or enzyme inhibition. **Desvenlafaxine is also available for the treatment of depression. This metabolite of venlafaxine is an effective antidepressant at 50 mg/day. Duloxetine is an effective antidepressant at doses similar to those used for younger adults (up to 60 mg/day). Its side-effect profile is mild, and there have been clinical reports that this medication (like other dual-acting antidepressants) may be helpful in controlling pain from arthritis of rheumatoid disorders. Milnacipran is an effective antidepressant for the elderly at doses of 75 to 100 mg/day. At these doses, it does not adversely affect cognitive function or psychomotor performance. Levomilnacipran is now available and is effective in geriatric patients at doses of 20 to 40 mg/day. This antidepressant is primarily cleared through the kidneys, so elderly patients with renal impairment probably should not be treated with this particular antidepressant. Mirtazapine, through its unique receptor activity, is also considered a dual-acting drug. Mirtazapine is an effective antidepressant in the**

elderly at doses of 15 to 30 mg/day. However, it is quite sedating, so starting doses should be low; mirtazapine may be helpful as a sleep-inducing medication even in the absence of depression. A rapidly dissolving tablet is very useful for nursing home patients who have difficulty swallowing a pill.

Although not officially classified as dual-acting drugs, trazodone and nefazodone also affect both norepinephrine and serotonin neurotransmission. These drugs are weak serotonin reuptake inhibitors and are less commonly used to treat depression. Trazodone, however, is an effective drug for inducing sleep onset. Nefazodone strongly inhibits the CYP450 3A4 metabolizing enzyme; it should not be prescribed together with benzodiazepine anxiolytics, steroids, antibiotics, antifungal medications, and certain antibiotics. When prescribed with mood-stabilizing medications (antiseizure drugs such as carbamazepine, lamotrigine, and divalproex sodium), blood levels of the mood stabilizer may rise, producing unwanted side effects.

Vilazodone and vortioxetine and new antidepressants that are available for clinical use. Research has been scant on the clinical efficacy of these two drugs in the depressed geriatric population.

In general, these dual-acting antidepressants are not likely to produce serious drug interactions with other medications. It is good practice, however, to minimize the total number of medications that are coprescribed with these antidepressants and avoid using more than one antidepressant at the same time. Drugs that increase serotonin function may interact with these dual-acting drugs to produce a serotonin syndrome and should be avoided.

BUPROPION

Bupropion is an effective and safe antidepressant that is often prescribed to elderly patients because of its activating properties and relatively benign side-effect profile (no weight gain or significant anticholinergic properties). Advantages of bupropion include efficacy especially in psychomotorically retarded elderly patients, and no adverse effects on cognition. Bupropion does not cause sedation, and may even be stimulating for some elderly patients which may interfere with sleep. Bupropion may be helpful in depressed elderly patients with Parkinson disease. However, because of its activating properties, bupropion may increase anxiety in some elderly individuals, especially those who are already anxious. Bupropion is a strong inhibitor of the metabolizing hepatic enzyme CYP450 2D6. This enzyme is necessary to convert tamoxifen to its active metabolite; inhibition by bupropion will prevent this conversion and render the tamoxifen useless. As with other antidepressants, starting doses should be lower than those recommended for younger adults, and dosage increments should be gradual.

TRICYCLIC ANTIDEPRESSANTS

Tricyclic antidepressants, nortriptyline and desipramine in particular, are effective medications for major depressive disorder in the elderly. Blood levels of these two drugs necessary to achieve therapeutic effect in the elderly are the same as in younger adults, thus providing the clinician with a laboratory value to guide dosing. Hydroxy metabolites of the tricyclics are potentially cardiotoxic in elderly individuals, and may be present at higher plasma levels than in younger adults because **their** renal excretion is decreased by advancing age. An electrocardiogram can be used to monitor patients in order to avoid approaching cardiotoxicity which is signaled by a widening of the QTc interval.

Annoying **peripheral anticholinergic** side effects of tricyclics, particularly constipation and dry mouth, can become severe problems in the elderly leading to bowel obstruction and poor dental hygiene. Central nervous system anticholinergic side effects, which usually consist of decreased memory and attention, can progress to a delirious state with severe confusion, disorientation, and paranoid ideation. Orthostatic hypotension, another common adverse effect of tricyclic treatment of late-life depression may increase the risk of falls; oversedation may interfere with participation in daytime activities.

To prevent these common tricyclic antidepressant side effects, relatively low starting doses with small dosage increments are usually employed. The prescribing maxim "start low (initial dose) and go slow (subsequent dosage increases)" applies to elderly individuals as a class of patients, although not necessarily to any individual patient.

Plasma levels of imipramine, desipramine, and nortriptyline can be routinely obtained. These levels correlate with therapeutic effects as well as with side effects. Prescribing clinicians who wish to use a tricyclic should follow plasma levels of the tricyclic (preferably the secondary amines, desipramine or nortriptyline) as doses are being adjusted to the therapeutic range. Using these levels along with an EKG to monitor the appearance of a potentially dangerous cardiac arrhythmia provides the prescribing clinician with two sources of data to guide the use of these medications.

MONOAMINE OXIDASE INHIBITORS

Monoamine oxidase inhibitors **may be** effective in depressed elderly patients. Their advantages include few anticholinergic effects and no effects on cardiac conduction. **MAO inhibitors are potentially dangerous for elderly individuals and should not be used except by those prescribing clinicians who have experience with these medications. The potential for toxic interactions between these drugs and other medications (e.g., serotonin-enhancing drugs; meperidine or other opiates, tricyclic antidepressants) or**

certain commonly eaten foods (e.g., aged cheese, pickled meats and fish) is high in the elderly population. Consequently, if these medications are to be used, they must be prescribed only to those elderly patients who are able to follow the dietary and drug interaction limitations.

For **appropriate and compliant** elderly patients, problem interactions with food and drugs can be minimized through patient education and collaboration among treating physicians. Limitations in their clinical utility include both hypotension and hypertension, insomnia, and somnolence. The starting dosages of tranylcypromine and phenelzine are 10 mg a day and 15 mg a day in two doses, respectively. In general, maximum dosages do not typically exceed 40 to 50 mg a day for tranylcypromine and 75 to 90 mg a day for phenelzine.

Transdermal selegiline (Emsam) is now available as a skin patch, although there is scant literature regarding its use in the elderly. This MAO inhibitor decreases the problems of drug and food interactions, that may occur when high doses of the patch or the interacting medication or food are ingested.

STIMULANTS

Stimulants, such as methylphenidate (Ritalin) are not classified as antidepressants and should not be used as the solo medication to treat an elderly depressed individual. **However,** low doses of methylphenidate (2.5 to 5.0 mg/day in divided doses) **may be** helpful in activating a withdrawn apathetic older patient, as well as increasing motivation and social participation. Methylphenidate may also enhance the therapeutic effect of antidepressants which have only produced a partial response. Low doses are not associated with tachycardia, hypertension, or cardiac toxicity; doses administered after 3 or 4 PM may interfere with sleep. **Other activating drugs (modafinil, armodafinil) may also be used for mild activating effects. Stimulants and other activating drugs should not be given after 2 to 3 PM in order to avoid interfering with sleep onset at bedtime.**

TREATMENT OF PSYCHOTIC DEPRESSION IN LATER LIFE

Psychotic depression is a severe illness with profound depressive symptomatology accompanied by delusions, often of somatic dysfunction, hopelessness, or guilt. Psychotic depression poses a health risk from progressive weight loss and exhaustion as well as from suicide, usually by violent means. For this reason, recognition and treatment of psychotic depression can be lifesaving. Both young and elderly patients with psychotic depression respond poorly to short-term treatment with antidepressants given alone. Combinations of antidepressants and

antipsychotic medications are effective in psychotic depression, in both controlled studies of younger patients and naturalistic treatment studies of elderly patients. Although tricyclics were the most frequently prescribed antidepressants in conjunction with a typical antipsychotic drug such as perphenazine (Trilafon), clinicians are now using a combination of SSRIs or **dual-acting** antidepressants **together with** atypical neuroleptics to reduce side effects **of older medications.** Electroconvulsive therapy, **often** rapidly effective and safe, remains the treatment of choice **for** severe geriatric psychotic depression.

TREATMENT OF DEPRESSION ASSOCIATED WITH MEDICAL ILLNESS AND WITH COGNITIVE IMPAIRMENT

An NIH Consensus Development Conference on the Diagnosis and Treatment of Depression in Late Life stated that "the hallmark of depression in older people is its comorbidity with medical illness." Compelling evidence from placebo-controlled clinical trials indicates that antidepressant medications are efficacious in patients with common comorbid medical disorders as diverse as stroke, Parkinson disease, Alzheimer disease, cancer, ischemic heart disease, chronic obstructive pulmonary disease, and arthritis. It appears that all antidepressants are therapeutic, at least to some degree, even in rehabilitation units where treatment of depressed patients can facilitate recovery from acute illness or injury and in a residential care setting, where patients were characterized by extreme old age and the interacting effects of many of the common disabling medical disorders of later life. **Until recently,** tricyclic antidepressants were the most widely studied drugs for the treatment of depression associated with another illness. They **have demonstrated** safety and effectiveness for poststroke patients, as well as for elderly depressed patients with Parkinson disease, cardiovascular illness, and cancer. However, interactions between **tricyclic** antidepressants and other **nonpsychiatric** medications used to treat **medical or neurological** illness can make older patients highly vulnerable to the development of side effects. Thus, careful monitoring of both potential adverse effects and plasma concentrations is necessary to optimize safety of these agents in frail elderly patients. SSRI and other nontricyclics are now more widely used in order to avoid this potential for adverse drug events; SSRIs have been shown effective for depressed patients with cardiovascular illness, stroke, Parkinson disease, and cancer.

Depression and cognition intersect in two ways: elderly individuals with failing cognition or who are suffering from a dementia commonly are depressed; depressed elderly individuals frequently complain of impaired memory and concentration. This interaction between depression and diminished cognitive function has been well studied. Regardless of whether the primary illness is dementia or is depression, the

depression should be treated, following the basic principles that apply **to** older patients. Medication use should be supervised, and caregivers should assess the benefits and monitor for adverse effects of treatment. Patients with dementia of the Alzheimer type are highly sensitive to increased cognitive impairment from anticholinergic medications. Therefore, antidepressants with the lowest anticholinergic effects are preferred, which include the SSRIs, **low doses of** bupropion, venlafaxine, mirtazapine **and vilazodone.** Low doses of nortriptyline are sometimes helpful as well. Memory and concentration often improve as a depression responds to medication; in mild to moderate degrees of dementia, successful antidepressant treatment can restore or even transiently improve memory and concentration. However, as the dementia worsens, antidepressants may cause increasing agitation and must be discontinued.

ESSENTIALS OF PATIENT MANAGEMENT

Several principles of clinical practice should guide the clinician who wishes to treat a depressed elderly individual. Careful psychiatric and medical diagnoses go hand in hand; older depressed individuals must be medically evaluated, especially for thyroid dysfunction and levels of B_{12} and folate. For some older people, this evaluation should include a neurological examination to rule out Parkinson disease or dementia. **It is well established that inflammation and inflammatory disorders are commonly associated with mood disturbance, especially with depression. Chronic illnesses such as diabetes are also commonly associated with depression. Poststroke and postmyocardial infarction depressions are similarly well known and failure to treat the depression may increase morbidity or subsequent mortality.** A comprehensive list of all medications (prescribed or self-purchased) including herbs, holistic remedies, and homeopathic preparations must be obtained since medication-induced (or worsened) depression is common in the elderly. The clinician must also determine whether or not alcohol (or other drugs) is being used excessively or inappropriately before embarking on a treatment program. It may be necessary for the clinician to obtain this information from a spouse, child, or other source as the depressed older person may not offer an accurate history.

Managing the depressed geriatric patient also includes educating the patient and family about the nature of depression as a true illness, emphasizing that it is treatable. Patients and their families need to understand that response to medication may take longer than for younger adults, that treatment adherence is essential, and that adverse effects can be expected but are amenable to supportive countermeasures. Weekly contact during the early phases of short-term treatment is recommended for individualizing treatment, providing encouragement and support,

reviewing questions and concerns, and monitoring safety (including suicidal potential). Common errors in practice include underdosing, discontinuing treatment too soon, failure to consider possible drug interactions, and polypharmacy. If a patient has not begun to respond by 6 to 8 weeks, compliance should be reevaluated, dosage or plasma concentration should be optimized, and referral to a geriatric psychiatrist experienced in treatment resistance should be considered.

Late-life depression is often a chronic illness and tends to recur, especially among those whose first depression develops after the age of 60. One must bring a longitudinal model of depression to treatment. Short-term treatment is intended to induce remission but incomplete remissions are not unusual and commonly lead to relapse. Recurrences after a longer period of recovery are also common. Studies continue to demonstrate the therapeutic benefits of long-term treatment of depression, and it is recommended, therefore, that treatment continue at least 6 months beyond recovery and for longer periods after a second or third episode. Many older patients have chronic depression that requires indefinite maintenance therapy. Long-term prevention of new episodes is best achieved by maintaining patients on the same dosage of antidepressant medication that was used successfully to treat the acute episode.

TREATMENT OF BIPOLAR DISORDERS IN LATER LIFE

Bipolar disorders constitute 5 to 10 percent of the diagnoses of elderly psychiatric inpatients. The diagnosis of a manic episode is usually not difficult in older individuals, although advanced age is associated with greater degrees of irritability and less euphoria. Most manic elderly patients have a history of prior manic episodes; onset of a first manic episode in late life is unusual, and often is associated with another illness such as thyroid dysfunction, or alcohol or substance abuse. Elderly individuals who have suffered repeated episodes of mania and depression over the course of their lives are likely to be quite dysfunctional when manic, with rapid shifts of mood. **Episodes of euphoria decrease with advanced age and are replaced by "dysphoric" or "mixed" mood states characterized by anger, irritability, paranoid ideation, and depression.** Diagnostic principles follow the suggestions for evaluation of an elderly depressed patient, therefore, including a careful and comprehensive medical evaluation, laboratory screening, and history of past or current drug and alcohol use.

In principle, the management of mania in elderly patients is similar to that in younger patients, although there may be a relatively poor response to standard pharmacological treatment in patients with long-standing bipolar illness. Treatment of the acute manic episode is focused on two separate categories of dysfunction: (1) management of manic behavior, impaired sleep, and (2) regulation of excessive or inappropriate mood

states. Agitated hyperactive elderly manic patients or those who are psychotic, as well as patients with impaired sleep may be helped by treatment with neuroleptics. All **antipsychotic medications** are effective for these symptoms, although current practice favors the use of **second-generation drugs** in order to avoid excessive extrapyramidal symptoms caused by conventional neuroleptics. Risperidone, olanzapine, and quetiapine are commonly used for this purpose; clozapine is also therapeutic but its use in older patients is associated with numerous potentially serious side effects; **other medications (e.g., ziprasidone, aripiprazole, lurasidone, olanzapine) may also be helpful, but research studies in the elderly of any of these medications are scant and do not provide adequate prescribing information. In general, therefore,** starting doses of an atypical neuroleptic should be low, with gradual dose increases; clinicians should resist the pressure to rapidly increase doses to control behavioral disruption. When an atypical neuroleptic is given at bedtime, sleep is often promoted which, in turn, may help diminish the manic symptoms.

Regulation of mood is accomplished by the use of mood-stabilizing medications; principles of their use follow guidelines for younger adults, although elderly individuals are significantly more sensitive to side effects. Lithium, for example, commonly produces side effects at doses and blood levels that are not toxic in younger people. For example, delirium can develop at lithium concentrations at or even below the **young adult** therapeutic lithium plasma concentration range. Delirium and cerebellar dysfunction may last for weeks after discontinuation of lithium use. Lithium-induced tremor is more frequent in older patients than in younger patients. Lithium can worsen Parkinson disease or produce parkinsonism in neurologically intact elderly patients, **and can also** worsen cognitive function, especially in demented patients. Sinoatrial block can be caused by lithium, a risk that is increased by the concurrent use of digitalis and β-adrenergic receptor antagonists (β-blockers). Elderly patients often receive drugs that increase lithium plasma concentrations, including thiazide diuretics, nonsteroidal anti-inflammatory drugs, or angiotensin-converting enzyme inhibitors. Age-induced decrements in glomerular filtration rates and renal blood flow reduce lithium excretion. Starting dosages of 75 to 150 mg a day may reduce adverse effects and improve compliance. These very low doses may successfully treat the disturbed mood with lower, nontoxic blood levels than in younger adults (e.g., 0.2 to 0.4 mM/L). **Because of the potential for toxicity and impairment of renal function, lithium use should be restricted to selected elderly patients whose physical health and ability to comply with prescribing guidelines and laboratory studies is evident.**

It is now well established that some anticonvulsant medications also are excellent mood stabilizers and may be effective and safe for elderly

individuals. Like lithium, however, anticonvulsants are used primarily to prevent recurrence of dysregulated mood states. For example, the anticonvulsants carbamazepine (Tegretol) and divalproex sodium (Depakote) have demonstrated prophylactic effectiveness in elderly individuals. Current prescribing practice favors the use of lamotrigine (Lamictal) as a stabilizing anticonvulsant for the elderly. Although a few individuals may develop a skin rash with this drug (less likely in the elderly), the side-effect profile is benign and with low therapeutic doses (12.5 to 100.0 mg/day) there may be few or no side effects. Oxcarbazepine (Trileptal) is also sometimes used to stabilize mood in younger adults, but there are no studies of its effectiveness in the elderly. Some evidence suggests that anticonvulsants are more effective than lithium in rapidly cycling patients and in patients with dysphoric mania, although studies have not been conducted in elderly patients.

Most anticonvulsants are metabolized by hepatic metabolizing enzymes, primarily CYP 3A4. Hematological and liver function tests are needed prior to treatment with anticonvulsants to ensure their safe use. Carbamazepine induces its own metabolism so blood levels may need to be adjusted during treatment. This drug can also cause leukopenia (white blood cell counts below 4,000/mm) in approximately 2 percent of cases; it should not be used for routine geriatric patient treatment. Valproate inhibits the CYP450 3A4 metabolizing enzyme and can cause elevations of other medications that use this metabolic pathway (e.g., lamotrigine, some benzodiazepines, and antidepressants). Valproate can also cause increased ammonia levels and toxic neurological states, although it is not clear that elderly individuals are more susceptible to this side effect. Lamotrigine, currently the preferred mood-stabilizing anticonvulsant for the elderly, is metabolized by CYP450 3A4 and its blood levels can be affected by other medications (e.g., valproate, fluoxetine) that inhibit this enzyme. Oxcarbazepine can cause a sudden decrease in serum sodium, followed by a change in mental status that may be mistaken for a worsening psychiatric condition. Levetiracetam has been used as a mood stabilizer in a small number of elderly patients. Other new anticonvulsants have not been tried in the elderly and cannot be recommended.

Prescribing guidelines for lithium and anticonvulsant treatment of late-life mood dysregulation follows the general geriatric prescribing principle: “start low and go slow.”

Continuation and maintenance treatment is necessary in elderly bipolar patients who achieve remission of a manic or depressive episode. Although studies of geriatric populations are lacking, studies of younger patients suggest that lithium, carbamazepine, and valproate are effective in preventing both depressive and manic episodes. For this reason, continuation and maintenance treatment with these drugs should be considered for elderly patients. Elderly patients, in particular, should be

carefully and frequently monitored for alterations in medical condition, especially hepatic function, as well as cognitive function. As the older bipolar patient continues to age, dosage modifications are likely to be necessary.

References

Dols A, Kupka RW, van Lammeren A, Beekman AT, Sajatovic M, Stek ML. The prevalence of late-life mania: a review. *Bipolar Disord*. 2014;16:113–118.

Ellison JM, Sivrioglu EY, Salzman C. Pharmacotherapy of late-life depression: evidence-based recommendations. In: Ellison JM, Kyomen HH, Verma S, eds. *Mood Disorders in Later Life*. 2nd ed. New York: Informa Healthcare USA, Inc.; 2009:239–290.

Forester BP, Jordan B. Bipolar disorder in later life. In: Ellison JM, Kyomen HH, Verma S, eds. *Mood Disorders in Later Life*. 2nd ed. New York: Informa Healthcare USA, Inc.; 2009:65–90.

Garnock-Jones KP. Vortioxetine: a review of its use in major depressive disorder. *CNS Drugs*. 2014;28:855–874.

Jacobson SA. *Clinical Manual of Geriatric Psychopharmacology*. 2nd ed. Arlington, VA: American Psychiatric Publishing, Inc.; 2014:536 pp.

Maust DT, Oslin DW, Thase ME. Going beyond antidepressant monotherapy for incomplete response in nonpsychotic late-life depression: a critical review. *Am J Geriatr Psychiatry*. 2013;21:973–986.

Robinson DS, Kajdasz DK, Gallipoli S, Whalen H, Wamil A, Reed CR. A 1-year, open-label study assessing the safety and tolerability of vilazodone in patients with major depressive disorder. *J Clin Psychopharmacol*. 2011;31:643–646.

Sajatovic M, Gildengers A, Al Jurdi RK, et al. Multisite, open-label, prospective trial of lamotrigine for geriatric bipolar depression: a preliminary report. *Bipolar Disord*. 2011;13:294–302.

Sheline YI, Disabato BM, Hranilovich J, et al. Treatment course with antidepressant therapy in late-life depression. *Am J Psychiatry*. 2012;169:1185–1193.

Tedeschini E, Levkovitz Y, Iovieno N, Ameral VE, Nelson JC, Papakostas GI. Efficacy of antidepressants for late-life depression: a meta-analysis and meta-regression of placebo-controlled randomized trials. *J Clin Psychiatry*. 2011;72:1660–1668.

57.4c Antianxiety Drugs

ALASTAIR J. FLINT, M.B.

Anxiety disorders are common in late life, with 5 to 10 percent of community-dwelling older adults meeting criteria for a current anxiety disorder. Despite this prevalence, there has been relatively little research on the pharmacologic treatment of anxiety in the elderly and most treatment decisions in this group have to rely on research conducted in younger adults. Extrapolating findings from younger to older adults is however potentially problematic, given the pharmacokinetic and pharmacodynamic changes that occur with aging. In particular, the

efficacy, tolerability, and safety of drugs in the elderly can be affected by age-related physiologic changes (such as the increased volume of distribution of lipophilic drugs, decreased function of liver and kidneys, and changes in organ receptor sensitivity), the increased frequency of comorbid medical and neurologic conditions, and the fact that older adults take more medications than their younger counterparts, thereby increasing the risk of drug–drug and drug–body interactions. This chapter will highlight issues that are of particular relevance to the elderly in the pharmacologic treatment of anxiety.

Double-blind randomized controlled trials (RCTs) are the gold standard for evaluating treatment efficacy. Unblinded ("open-label") and uncontrolled studies are of limited use in guiding treatment because of the significant biases and confounds that can affect their findings. There have been 10 double-blind RCTs of the pharmacologic treatment of anxiety disorders in persons aged 60 years or older (Table 57.4c–1), but only eight studies pertain to drugs that are available in North America. Eight of the 10 RCTs focused on generalized anxiety disorder (GAD), while two studies included participants with nonspecific anxiety. The category of anxiety disorders has historically included phobic disorders, panic disorder, GAD, posttraumatic stress disorder (PTSD), and obsessive compulsive disorder (OCD). Why then have treatment studies in the elderly focused almost exclusively on GAD? Knowledge of the epidemiology of anxiety disorders in later life helps answer this question. First, phobias and GAD are the most prevalent anxiety disorders in the elderly. However, the most common presentations of phobia in later life, that is, specific phobia and agoraphobia without panic, tend not to come to psychiatric attention. In addition, the treatment of choice for these two phobic disorders is behavioral therapy, not pharmacotherapy. Therefore, GAD is the most frequently occurring late-life anxiety disorder in psychiatric practice that is medication responsive. Second, panic disorder, OCD, and social anxiety disorder (social phobia) have a low incidence and prevalence in later life, and therefore clinicians and researchers may view them as less compelling compared with disorders with a higher incidence and prevalence. In contrast, approximately half of the cases of syndromal generalized anxiety in the elderly are late onset, often in association with major depression. Finally, syndromal and subsyndromal generalized anxiety is the commonest manifestation of anxiety associated with medical illness and cognitive impairment in older persons, again frequently in association with depression.

Table 57.4c–1.
Published Double-Blind Randomized Controlled Trials of the Pharmacologic Treatment of Anxiety Disorders in Persons Aged 60 Years

or Older

Study	Anxiety Disorder	Drug (Daily Dosage)[a]	Duration of Treatment (Weeks)	Efficacy	Comments
Lenze et al., *JAMA.* 2009;301:295.	GAD (*n* = 177)	Escitalopram (10–20 mg)	12	Drug>placebo	NIH-funded
Lenze et al., *Am J Psychiatry.* 2005;162:146.	GAD (*n* = 31) Panic (*n* = 2) PTSD (*n* = 1)	Citalopram (20–30 mg)	8	Drug>placebo	NIH- and industry-funded
Katz et al., *J Am Geriatr Soc.* 2002;50:18.	GAD (*n* = 184)	Venlafaxine ER (37.5–225 mg)	8	Drug>placebo	Industry-funded retrospective analysis of pooled data
Davidson et al., *Hum Psychopharmacol.* 2008;23:519.	GAD (*n* = 73)	Duloxetine (60–120 mg)	9–10	Drug>placebo	Industry-funded retrospective analysis of pooled data
Koepke et al., *Psychosomatics* 1982;23:641.	Anxiety neurosis (*n* = 220)	Oxazepam (30 mg)	4	Drug>placebo	
Bresolin et al., *Clin Ther,* 1988;10:536.	GAD (*n* = 63)	Ketazolam (15 mg)	2	Drug>placebo	Not approved for sale in North America
Frattola et al., *Clin Neuropharmacol.* 1992;15:477.	GAD (*n* = 33) Adjustment disorder with anxious mood (*n* = 7)	Alpidem (75–150 mg)	3	Drug>placebo	Industry-funded. Withdrawn from the global market because of liver toxicity
Bohm et al., *J Clin Psychopharmacol.* 1990;10:47S.	Anxiety state (*n* = 40)	Buspirone (5–30 mg)	4	Drug>placebo	Industry-funded
Montgomery et al., *Br J Psychiatry.* 2008;193:389.	GAD (*n* = 273)	Pregabalin (150–600 mg)	8	Drug>placebo	Industry-funded. Uncertain clinical significance of drug–placebo difference
Mezhebovsky et al., *Int J Geriatr Psychiatry.* 2013;28:615.	GAD (*n* = 450)	Quetiapine (50–300 mg)	9	Drug>placebo	Industry-funded

[a]Placebo was the comparator agent in each trial.

The 10 RCTs in Table 57.4c–1 have evaluated the efficacy and tolerability of antidepressants, benzodiazepines, nonbenzodiazepine anxiolytics, an anticonvulsant, and an atypical antipsychotic in the treatment of late-life GAD. The remainder of this chapter will discuss these classes of drugs, as well as commenting on miscellaneous drugs that have been examined in younger persons with anxiety.

ANTIDEPRESSANTS

In recent years, there has been growing awareness that antidepressant medications have a broad spectrum of efficacy in the treatment of anxiety disorders in younger adults. RCTs in mixed-age patients have found that antidepressants of various classes are beneficial in the treatment of GAD, panic disorder, social anxiety disorder, PTSD, and OCD (Table 57.4c–2).

RCTs of antidepressant pharmacotherapy in older people with an anxiety disorder found that escitalopram (Lexapro) and citalopram (Celexa) are more efficacious than placebo in treating GAD (Table 57.4c–1). In addition, retrospective analyses of data pooled from RCTs found that venlafaxine (Effexor) and duloxetine (Cymbalta) had greater efficacy than placebo in older persons with GAD. In the escitalopram, citalopram, and venlafaxine studies, approximately two-thirds of older participants were classified as responders after 8 to 12 weeks of treatment. The venlafaxine analysis also compared response rates between younger and older participants and found almost identical rates of response to drug and placebo between the two age groups. Escitalopram, citalopram, and venlafaxine were well tolerated in these trials. On the other hand duloxetine, at dosages of 60 to 120 mg/day, was associated with significantly more nausea and weight loss than placebo and the duloxetine group had a significantly higher rate of discontinuation than placebo (22 percent vs. 0 percent).

Table 57.4c–2.
Drugs Found Efficacious in Randomized Controlled Trials Treating Anxiety Disorders in Mixed-Age Samples

Disorder	Drug
Generalized anxiety disorder	SSRI: escitalopram[a], fluoxetine, paroxetine[a]
	SNRI: duloxetine[a], venlafaxine[a]
	TCA: imipramine
	Other antidepressants: trazodone
	Benzodiazepine: various
	Azapirone: buspirone[a]
	Other: hydroxyzine, pregabalin, quetiapine
Panic disorder with or without agoraphobia	SSRI: citalopram, fluoxetine[a], fluvoxamine, paroxetine[a], sertraline[a]

	SNRI: venlafaxine[a]
	TCA: clomipramine, desipramine, imipramine
	MAOI: brofaromine, phenelzine
	Benzodiazepine: alprazolam[a], clonazepam[a], diazepam, lorazepam
Social anxiety disorder	SSRI: fluvoxamine, paroxetine[a], sertraline[a]
	SNRI: venlafaxine[a]
	MAOI: brofaromine, moclobemide, phenelzine
	Benzodiazepine: alprazolam, clonazepam
Posttraumatic stress disorder	SSRI: fluoxetine, paroxetine[a], sertraline[a]
	SNRI: venlafaxine
	TCA: amitriptyline, imipramine
	MAOI: phenelzine
	Other: prazosin (for nightmares)
Obsessive compulsive disorder	SSRI: fluoxetine[a], fluvoxamine[a], paroxetine[a], sertraline[a]
	TCA: clomipramine[a]

[a]FDA approved.

SSRI, selective serotonin reuptake inhibitor; SNRI, serotonin and norepinephrine reuptake inhibitor; TCA, tricyclic antidepressant; MAOI, monoamine oxidase inhibitor.

Selective Serotonin Reuptake Inhibitors

Of the selective serotonin reuptake inhibitors (SSRIs), escitalopram and citalopram are well suited to older persons. They have linear pharmacokinetics across the dose range (i.e., an increase in dose results in a predictable increase in plasma concentration), no clinically active metabolites, elimination half-lives of 33 hours or less, and little propensity for drug–drug interactions. In addition, sertraline (Zoloft) is also suitable in older persons because of a favorable pharmacokinetic profile. Although data from RCTs in mixed-age patients have demonstrated the efficacy of sertraline in several anxiety disorders (Table 57.4c–2), there are no controlled data pertaining to sertraline in late-life anxiety disorders, although it has been found to be efficacious, safe, and well tolerated in late-life depression. On the other hand, paroxetine and fluoxetine have less favorable pharmacokinetic profiles in older adults and paroxetine has the potential to cause anticholinergic adverse effects such as worsening cognitive function, blurred vision, dry mouth, constipation, and urinary hesitancy; therefore, these drugs are not recommended as first-line SSRIs in the elderly.

If started at too high a dose, SSRIs may initially exacerbate symptoms of anxiety and, therefore, a lower starting dose (e.g., 2.5 mg/day of escitalopram, 5 mg/day of citalopram, or 25 to 50 mg of sertraline) is often indicated in elderly anxious individuals. Once it has been established that the patient is tolerating the starting dose, the dose should then be gradually increased to the usual therapeutic range (Table 57.4c–3). Both citalopram and escitalopram have been associated with dose-dependent prolongation

of the QT interval on EKG and, as a result, it is recommended not to exceed 20 mg/day of citalopram in persons aged 60 years or older and 10 mg/day of escitalopram in persons 65 years or older. Anxious patients frequently misattribute somatic symptoms of anxiety to the adverse effects of medication. Thus, in addition to starting with a low dose of antidepressant, adherence with medication can be enhanced by discussing potential side effects in advance, frequent monitoring of the patient during the first few weeks of treatment, and availability of the clinician by telephone between appointments to address any questions or concerns that the patient may have about treatment.

Table 57.4c–3.
Usual Starting Dose, Initial Target Dose, and Recommended Maximum Dose of Selected SSRIs and SNRIs in Older Adults with Anxiety Disorders

Drug	Usual Starting Dose (mg/day)	Initial Target Dose (mg/day)	Recommended Maximum Dose (mg/day)
Escitalopram	2.5	5	10
Citalopram	5	10	20
Sertraline	25	50–100	200
Venlafaxine ER	37.5	150	300
Duloxetine	30	60	120

SSRI, selective serotonin inhibitor; SNRI, serotonin and norepinephrine reuptake inhibitor.

Based on the above considerations, escitalopram is a reasonable first-line choice for the treatment of GAD in later life, given that it has FDA approval for this condition and there are efficacy and tolerability data specific to the elderly. In the absence of pharmacotherapy, data pertaining to anxiety disorders other than GAD in the elderly, sertraline would be a reasonable first-line SSRI for panic, social anxiety, PTSD, and OCD, for the following reasons: it has FDA approval for the treatment of these disorders based on studies of mixed-age samples, it has a favorable pharmacokinetic and tolerability profile in older patients, and there is considerable flexibility in dosing without the concern of QT prolongation at higher doses.

Serotonin and Norepinephrine Reuptake Inhibitors

Venlafaxine (Effexor) is a serotonin and norepinephrine reuptake inhibitor (SNRI), although it has been postulated that its noradrenergic effect occurs only at higher dosage (possibly 150 mg/day or higher). The doses used to treat GAD in older persons in RCTs were relatively low (75 to 150 mg/day in the majority of older subjects), raising the possibility that the serotonergic action of the drug may be sufficient to treat GAD in many older patients. Like escitalopram, citalopram, and sertraline, venlafaxine

has linear pharmacokinetics and lacks propensity for significant drug–drug interactions. The usual starting dose of venlafaxine ER in older persons is 37.5 mg/day, with dose increments of 37.5 mg every 4 to 5 days to reach a minimum target dose of 150 mg/day. The dose range of venlafaxine ER is wide (75 to 375 mg/day), and the usual therapeutic dose for GAD has not been established in older patients. However, if there has not been satisfactory improvement of GAD after 4 weeks of treatment at a dose of 150 mg/day, the dose should be gradually increased, guided by response and tolerability (Table 57.4c–3). Venlafaxine is a suitable antidepressant for older persons with GAD, and on the basis of RCTs conducted in mixed-age samples, it can also be considered as a first-line option for panic, social anxiety, and PTSD. Given that venlafaxine and SSRIs are different drug classes, switching from one class to the other (i.e., from an SSRI to venlafaxine or vice versa) is a reasonable approach in the case of inadequate response to a therapeutic trial of the first drug.

Other Antidepressants

In addition to SSRIs and SNRIs, selected tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs) have been found to be efficacious in a range of anxiety disorders in younger adults (Table 57.4c–2), although there have been no comparable studies in the elderly. With the exception of desipramine (Norpramin), the TCAs that have been examined in younger persons have been tertiary amine compounds, which are potentially problematic in the elderly. In particular, older adults may be vulnerable to the anticholinergic, sedative, hypotensive, and adverse cardiac effects of tertiary amine TCAs, especially if they have concomitant physical or cognitive problems that could be aggravated by these effects. A retrospective analysis of data pooled from three RCTs of the treatment of poststroke depression, found that nortriptyline (Aventyl) (a secondary amine TCA that is preferred over tertiary amine TCAs in the elderly) was more effective than placebo in treating generalized anxiety associated with the depression. Thus, if a TCA is considered for an older anxious person, the recommendations would be despiramine in the case of panic disorder and either nortriptyline or desipramine (which is a metabolite of imipramine [Tofranil]) in the case of GAD. Although OCD rarely starts as a primary disorder in late life, older persons with chronic OCD may present for treatment. Clomipramine (Anafranil), a serotonergic tertiary amine TCA, can be a useful treatment for OCD, especially in cases that have failed to adequately improve with SSRIs. While the use of clomipramine has potential limitations in the elderly because of its side-effect profile, it is worth considering, in combination with cognitive behavior therapy, in selected older persons who have OCD that has failed to respond to other medications.

The first-generation MAOI phenelzine (Nardil) should be kept in mind

as a potential treatment for refractory panic disorder, social phobia, or PTSD, or for refractory anxious depression. However, the need to avoid tyramine-containing foods and certain medications to prevent development of a hypertensive crisis with phenelzine, and the drug's potential to cause significant supine and orthostatic hypotension and weight gain, will limit its use to carefully selected individuals.

BENZODIAZEPINES

Despite their widespread use, benzodiazepines have a limited role in the treatment of late-life anxiety. First, they are not a useful treatment for two of the three most common anxiety disorders in later life, that is, specific phobias and agoraphobia without panic. Second, the frequent comorbidity of GAD with depression in later life suggests the need for a medication with both antidepressant and anxiolytic effects. Benzodiazepines are not an effective treatment for the core symptoms of depression. On the other hand, antidepressant medication is an effective treatment for GAD even when depression is absent. Third, there are a number of risks associated with benzodiazepines in the elderly, including cognitive impairment, psychomotor impairment, excessive daytime sedation, instability of gait, falls, and hip fractures. Finally, anxiety disorders tend to be chronic conditions, with most patients requiring long-term treatment. Long-term use of benzodiazepines can be associated with psychological and physical dependence and tolerance, so that it can be difficult to discontinue the drug if its benefits diminish and/or adverse effects develop. Persons with a history of alcoholism or drug abuse may be at particular risk of becoming dependent on a benzodiazepine.

In situations where benzodiazepines are indicated (e.g., acute situational anxiety or as a short-term adjunct to an antidepressant until the antidepressant becomes effective), lorazepam (Ativan) or oxazepam (Serax) are preferred in the elderly because they have relatively short elimination half-lives (10 to 15 hours), no active metabolites, and they do not undergo oxidative metabolism in the liver (thus their clearance is not affected by age). As a result, these drugs do not have the potential for cumulative toxicity in the elderly which is a risk associated with longer-acting compounds such as diazepam (Valium) or clonazepam (Klonopin). Oxazepam is slowly absorbed, so is not very effective as a hypnotic unless it is taken at least an hour before bedtime. The usual dosage range in the elderly is lorazepam 0.5 to 1.5 mg/day and oxazepam 10 to 30 mg/day. Although alprazolam (Xanax) is FDA approved for the treatment of GAD and panic disorder, there are no data pertaining to its risk–benefit profile in older persons with these conditions. Alprazolam has high potential for dependence and abuse. Its relatively short elimination half-life can result in rebound anxiety, which can lead to patients taking more frequent doses and higher daily doses than prescribed. For these reasons, alprazolam is

not recommended in the elderly.

As noted, most elderly patients who are prescribed a benzodiazepine do not require the drug on an ongoing basis. However, there are occasions when long-term benzodiazepine treatment may be necessary. In particular, some patients who have been maintained on a benzodiazepine for years can experience a significant worsening of anxiety when an attempt is made to withdraw the medication. In such cases, if the benzodiazepine is not causing significant side effects, it is usually least disruptive to leave the individual on the medication and monitor him or her on a regular basis. If side effects do become an issue as the person grows older, dose reduction rather than complete discontinuation of the drug may be the best compromise. In addition, some patients do not experience adequate relief of anxiety symptoms with a nonbenzodiazepine medication alone, and in these cases longer-term treatment with an adjunctive benzodiazepine, given at the lowest possible dose, may be necessary.

BUSPIRONE

Buspirone (Buspar), an azapirone derivative and a serotonin-1A receptor partial agonist, is a nonbenzodiazepine anxiolytic that is more efficacious than placebo in the treatment of GAD. It is not indicated in the treatment of other anxiety disorders. Although buspirone and benzodiazepines have comparable efficacy in patients with GAD who persist with treatment, several studies have found that patients are more likely to discontinue buspirone than a benzodiazepine, for reasons other than side effects. It has been suggested that the higher rate of discontinuation of buspirone compared with a benzodiazepine may be due to buspirone's delayed onset of action.

Although buspirone is an efficacious treatment for GAD, it has a limited role in the treatment of generalized anxiety in late life. First, it is not a substitute for antidepressant medication in the treatment of generalized anxiety that is comorbid with, or secondary to, major depression. Second, because buspirone has a delayed onset of action, it is not suitable as an acute adjunctive treatment of generalized anxiety symptoms associated with major depression. Third, there is evidence that buspirone may have diminished efficacy in persons who have previously been treated with a benzodiazepine. Therefore, its usefulness in the management of elderly patients with chronic GAD, many of whom have been treated with benzodiazepines, is unclear. Finally, the twice-a-day to three-times-a-day dosing schedule of buspirone increases the risk of poor adherence with this medication.

Unlike benzodiazepines, buspirone does not cause sedation, psychomotor impairment, cognitive impairment, or depressed respiratory drive. This side-effect profile suggests that it would be preferable to a benzodiazepine in the ongoing management of generalized anxiety in

nondepressed patients who have chronic obstructive airways disease, sleep apnea, or a neurologic disorder. However, select antidepressants may also be effective for anxiety and well tolerated in these patients.

The usual starting dose of buspirone in older patients is 5 mg twice-a-day, followed by dose increments of 5 mg every 4 to 5 days to reach a target dose of 20 to 30 mg/day, given in divided doses. If a person fails to show significant improvement after 4 to 6 weeks at this target dose, the dose can be further gradually increased to a maximum of 45 mg/day, as tolerated. The main side effects of buspirone are dizziness, headache, and nausea. Gradual dose titration is essential in order to minimize the risk of these effects.

Sometimes a physician elects to switch a patient from long-term benzodiazepine therapy to buspirone. Because buspirone has a delayed onset of action and it does not suppress benzodiazepine withdrawal symptoms, a patient should not be abruptly switched from a benzodiazepine to buspirone. Instead, the patient should take a therapeutic dose of buspirone for 2 to 4 weeks before the benzodiazepine is gradually tapered. Despite this approach, some patients are less satisfied with the anxiolytic effect of buspirone compared with that of a benzodiazepine.

OTHER DRUGS

Pregabalin

Pregabalin (Lyrica) is an anticonvulsant drug that binds to the alpha-2-delta subunit of the voltage-dependent calcium channel in the central nervous system and decreases the release of neurotransmitters including glutamate, norepinephrine, and substance P. Unlike benzodiazepines, pregabalin does not bind to GABA receptors and does not alter GABA metabolism. Several industry-funded RCTs in mixed-age patients have found pregabalin to be more efficacious than placebo in the treatment of GAD, although it does not have FDA approval for this indication. Although a statistically significant difference between 150 and 600 mg/day of pregabalin and placebo was found in patients aged 65 years or older with GAD, the clinical significance of the difference is uncertain: There was only a two-point greater reduction in HAM-A scores with drug versus placebo (12.9 vs. 10.7). Moreover, pregabalin and placebo groups did not differ in response or remission rates at the end of the 8-week study. Pregabalin was well tolerated in this group of older patients.

Quetiapine XR

First-generation antipsychotic medications have long been used in the acute treatment of agitation, but they are not recommended as a treatment for primary anxiety disorders because of the risk of tardive dyskinesia associated with their longer-term use. Quetiapine XR (Seroquel XR) is a

second-generation atypical antipsychotic with a low risk of neurologic side effects. A recent industry-funded study found that 50 to 300 mg/day of quetiapine XR was more efficacious than placebo in treating GAD in older adults, with significant differences in response and remission rates by the end of 8 weeks of treatment. While quetiapine has a low risk of parkinsonism and tardive dyskinesia, it can be associated with significant weight gain and metabolic dysfunction, and as a result cannot be recommended as a routine first-line treatment of GAD in the elderly.

Beta Blockers and Antihistamines

Beta blockers (e.g., 10 to 20 mg of propranolol [Inderal]) can be effective for performance anxiety when given about 1 hour prior to the feared performance. Therefore, they can be useful in patients with specific types of social anxiety disorder such as public speaking or eating in public. However, beta blockers should be avoided in anxious patients with sinus bradycardia, hypotension, or asthma.

RCTs conducted in Europe have found that the antihistamine drug hydroxyzine (Atarax) is more efficacious than placebo and comparable in efficacy to a benzodiazepine and buspirone in the treatment of GAD in mixed-age adults. Hydroxyzine has not been studied as a treatment of GAD in late life, but its sedative and anticholinergic effects (increasing the risk of daytime somnolence, cognitive impairment, and falls and fractures) limit its usefulness in older patients.

References

Bryant C, Mohlman J, Gum A, et al. Anxiety disorders in older adults: looking to DSM5 and beyond. *Am J Geriatr Psychiatry*. 2013;21:872–876.

*Flint AJ. Epidemiology and comorbidity of anxiety disorders in the elderly. *Am J Psychiatry*. 1994;151:640–649.

Iaboni A, Bronskill SE, Reynolds KB, et al. Changing pattern of sedative use in older adults: A population-based cohort study. *Drugs Aging*. 2016;33:523.

Kimura M, Tateno A, Robinson RG. Treatment of poststroke generalized anxiety disorder comorbid with poststroke depression: merged analysis of nortriptyline trials. *Am J Geriatr Psychiatry*. 2003;11:320–327.

*Lenze EJ, Rollman BL, Shear MK, et al. Escitalopram for older adults with generalized anxiety disorder: a randomized controlled trial. *JAMA*. 2009;301:295–303.

*Lenze EJ, Wetherell JL. A lifespan view of anxiety disorders. *Dialogues Clin Neurosci*. 2011;13:381–399.

Maust DT, Chen SH, Benson A, et al. Older adults recently started on psychotropic medication: where are the symptoms? *Int J Geriatr Psychiatry*. 2015;30:580–586.

Mezhebovsky I, Mägi K, She F, Datto C, Eriksson H. Double-blind, randomized study of extended release quetiapine fumarate (quetiapine XR) monotherapy in older patients with generalized anxiety disorder. *Int J Geriatr Psychiatry*. 2013;28:615–625.

Mohlman J, Bryant C, Lenze EJ, et al. Improving recognition of late-life anxiety disorders in Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition: observations and recommendations of the advisory committee to the lifespan disorders work group. *Int J Geriatr Psychiatry*. 2012;27:549–556.

Olfson M, King M, Schoenbaum M. Benzodiazepine use in the United States. *JAMA Psychiatry*. 2015;72(2):136–142.

Paquin AM, Zimmerman K, Rudolph JL. Risk versus risk: a review of benzodiazepine reduction in older adults. *Expert Opin Drug Saf*. 2014;13:919–934.

*Pinquart M, Duberstein PR. Treatment of anxiety disorders in older adults: a meta-analytic comparison of behavioral and pharmacological interventions. *Am J Geriatr Psychiatry*. 2007;15:639–651.

Rosnick CB, Wetherell JL, White KS, et al. Cognitive-behavioral therapy augmentation of SSRI reduces cortisol levels in older adults with generalized anxiety disorder: A randomized clinical trial. *J Consult Clin Psychol*. 2016;84:345.

*Tannenbaum C, Martin P, Tamblyn R, Benedetti A, Ahmed S. Reduction of inappropriate benzodiazepine prescriptions among older adults through direct patient education: the EMPOWER cluster randomized trial. *JAMA Intern Med*. 2014;174:890–898.

van't Veer-Tazelaar PJ, van Marwijk HW, van Oppen P, et al. Stepped-care prevention of anxiety and depression in late life: a randomized controlled trial. *Arch Gen Psychiatry*. 2009;66:297–304.

Wetherell JL, Petkus AJ, White KS, et al. Antidepressant medication augmented with cognitive-behavioral therapy for generalized anxiety disorder in older adults. *Am J Psychiatry*. 2013;170:782.

57.4d Antipsychotic Drugs

KELLY C. LEE, PHARM.D., M.A.S., B.C.P.P., AND VIVIAN HOOK, PH.D.

Antipsychotic medications constitute the treatment of choice for psychoses associated with schizophrenia or schizoaffective disorder. Antipsychotics are frequently prescribed in the elderly for the treatment of psychosis and severe behavioral disturbances, such as agitation and aggression. Although nonpharmacological therapies may be helpful in milder states of behavioral disturbances associated with cognitive disorders such as dementia, antipsychotics are often necessary to treat severe symptoms. Age-related physiological changes in the elderly can increase sensitivity to adverse events from these medications. Treatment is also often complicated by the frequent presence of comorbid medical conditions and medication interactions in the elderly. Despite the complexities, antipsychotic medications are commonly prescribed to older patients in inpatient, outpatient, residential, and nursing home settings.

Antipsychotics have been broadly classified into two groups. The older agents (i.e., introduced in the United States before 1990) are referred to as typical or conventional antipsychotics and primarily block dopamine

receptors (primarily D2) to exert antipsychotic effects. Typical antipsychotics, while effective for positive symptoms of schizophrenia, are associated with undesirable extrapyramidal symptoms and endocrine changes.

Atypical antipsychotics have absence or decreased incidence of extrapyramidal effects and tardive dyskinesia, as well as an expanded spectrum of activity against negative symptoms associated with schizophrenia. Because of their improved neurological profile, atypical antipsychotics have generally replaced typical antipsychotics. Unfortunately, due to studies that ultimately led to black box warning labeling for antipsychotics when used in elderly patients with dementia, the benefits of using these agents must be carefully weighed against the potential risks.

PHARMACOLOGICAL ACTIONS

Pharmacokinetics and Pharmacodynamics

Pharmacokinetics can generally be described as the body's effect on the absorption, distribution, metabolism, and elimination of drugs. Numerous age-related changes in physiology may affect pharmacokinetics. Changes in pharmacodynamics (the drug's effect on the body) are also likely to occur; however, these mechanisms are not as well understood relative to the knowledge of age-related pharmacokinetic changes. Neurotransmitter and receptor function changes in the central nervous system may increase sensitivity to drugs and toxicity may occur at lower plasma levels. In general, these pharmacokinetic and pharmacodynamic changes may lead to higher blood levels of medications, prolonged pharmacological effects, and a greater risk for side effects.

Receptor Activity

The specific mechanism of action of antipsychotics has not been elucidated, but most of the research attention has focused on receptor blockade of dopamine and serotonin. While the typical or conventional antipsychotics have been shown to work through blockade of the central dopamine receptors (specifically D2), atypical antipsychotics have been referred to as serotonin–dopamine antagonists as they block both 5-HT_{2A} and D2 receptors. Antipsychotic drugs also affect other receptors, usually resulting in adverse effects rather than an enhanced therapeutic response. For example, histamine-1 (H_1) blockade is responsible for sedation and possibly weight gain. Blockade of the 5-HT_{2C} receptor may also play a role in weight gain associated with some atypical agents. Blockade of the α_1 receptors causes orthostatic hypotension, whereas α_2 blockade plays a role in sexual dysfunction. Muscarinic receptor blockade causes the classical

anticholinergic effects such as dry mouth, blurred vision, constipation, and urinary retention. An understanding of the receptor-binding properties of the typical and atypical agents can predict their adverse effect profiles (Table 57.4d–1).

THERAPEUTIC INDICATIONS FOR ANTIPSYCHOTICS

Early-Onset Schizophrenia

Historically, it has been thought that all schizophrenia patients uniformly follow a poor prognostic course with no return to premorbid level of functioning. Fortunately, long-term studies in chronic schizophrenia have been more optimistic. Approximately one-third of the patients either undergo remission or are left with only mild symptoms over the long term, while 15 to 20 percent of patients get worse. The remaining patients appear to have a stable course. Patients who have benefited from maintenance antipsychotic treatment with prolonged absence of psychotic symptoms may warrant a trial off medication. Older patients typically have fewer and less severe positive symptoms than younger patients, while negative symptoms tend to persist into late life. Relapse and exacerbation may occur in older patients with chronic schizophrenia just as they do in younger patients, and investigators have struggled to identify factors that may predict outcome. Hence, if a drug withdrawal is attempted, careful monitoring is required.

Table 57.4d–1.
Relative Adverse Effects for Selected Antipsychotics

Medication	Brand Name	Sedation	Hypotensive Effects	Anticholinergic Effects	Extrapyramidal Effects
Haloperidol	Haldol	+	+	+	+++
Clozapine	Clozaril	+++	+++	+++	–
Risperidone	Risperdal	+	++	+/–	++
Olanzapine	Zyprexa	++	++	+	+/–
Quetiapine	Seroquel	++	++	+/–	–
Ziprasidone	Geodon	+	++	+/–	+/–
Aripiprazole	Abilify	+	+	+/–	++[a]
Paliperidone	Invega	+	+	+/–	++
Asenapine[b]	Saphris	++	++	+	++[a]
Iloperidone[b]	Fanapt	+	++	+	+
Lurasidone[b]	Latuda	++	+	+	++[a]
Brexpiprazole[b]	Rexulti	+	+	+/–	++[a]

Medication effects: –, none; +/–, rare; +, mild; ++, mild-to-moderate; +++, marked.

[a]Includes akathisia.

[b]Risk for adverse effects listed is based on limited experience.

Late-Onset Schizophrenia

Approximately 20 to 30 percent of patients have onset of symptoms after 40 years of age and are considered to have late-onset schizophrenia (LOS). In general, regardless of the age of onset of schizophrenia (LOS vs. early-onset schizophrenia [EOS]), patients with schizophrenia are similar in terms of positive symptoms, such as delusions and hallucinations, chronicity of the course of the disorder, and family history. However, LOS differs from EOS in several aspects. LOS seems to be more common in women, have better premorbid functioning with successful occupations and marital histories, fewer negative symptoms, and less severe neurocognitive impairment. Additionally, patients with LOS typically require lower antipsychotic doses and have less positive symptoms compared with EOS patients. Onset after age 65 usually signifies very LOS-like psychosis, which is typically secondary to general medical conditions such as dementia or neurodegenerative disorder. Additional studies are needed to identify genetic variants, biomarkers, and changes on neuroimaging studies that may differentiate LOS from EOS.

Delusional Disorder

Delusional disorder refers to the presence of one or more persistent, nonbizarre delusions and may be misdiagnosed as schizophrenia. Delusions commonly involve situations that may occur in real life, such as being poisoned, having a disease, or experiencing spousal infidelity. Delusional disorder generally presents in middle or late life and can be differentiated from schizophrenia by the absence of prominent auditory or visual hallucinations and by the lack of functional deterioration outside the area of delusions. Psychiatric comorbidities, especially depressive disorders, are common among those with delusional disorder. Mild symptoms may be treated without medications, but antipsychotics may be helpful for the agitated patient whose delusions have a negative impact on social relationships or functional behaviors. Antipsychotic noncompliance is a common problem in those with delusional disorder, although some suggest that women may be more adherent to pharmacological treatments than men.

Dementia with Psychosis or Severe Behavioral Disturbances

The most common types of dementia are Alzheimer disease and vascular dementia. Patients who develop dementias have cognitive impairment resulting in memory and executive functioning decline but also noncognitive dysfunction that can be more distressful. A majority of elderly patients with dementia develop behavioral disturbances such as aggression, agitation, delusions, and other neuropsychiatric symptoms during their course of illness. These changes in mood, affect, and behavior can lead to

early institutionalization, increased caregiver burden, and increased health care costs. When behavioral problems arise, early recognition and prompt treatment can improve quality of life of patients, their families, and caregivers and reduce health care costs. In addition to medications used for cognitive dysfunction, nonpharmacological interventions and antipsychotics have been widely used for treating behavioral disturbances associated with dementias. Treatment guidelines have generally recommended careful assessment, diagnosis, and nonpharmacological treatment approaches, such as psychosocial and environmental interventions, prior to pharmacological approaches. When nonpharmacological approaches fail, the provider, together with the patient and his or her family, should examine pharmacological options carefully. The family should be informed about the relative risks and benefits of antipsychotics and other treatments, as well as the risks of not using medication, so that they can participate effectively in clinical decision making. Documentation of informed consent should be obtained.

The use of antipsychotics in those with dementia has become controversial due to the FDA black box warning regarding the increased incidence of cerebrovascular events and death. In the CATIE-AD study of atypical antipsychotics in patients with Alzheimer disease, olanzapine [Zyprexa], quetiapine [Seroquel], and risperidone [Risperdal] were also associated with significant decline in cognitive functioning over 36 weeks compared to placebo. While the effectiveness of antipsychotics for noncognitive symptoms in dementia may be marginal, a short-term trial is generally warranted to control symptoms of mania, psychosis, hallucinations, and aggressive behaviors that could endanger others. While there is no significant difference in effectiveness between atypical and typical antipsychotics, atypical antipsychotics are generally preferred due to better side effect profile, particularly in terms of extrapyramidal side effects and tardive dyskinesia. Antipsychotics should be used in the lowest effective doses and careful monitoring for adverse events is recommended.

The decision to discontinue antipsychotics in patients who have responded to the drug must be carefully weighed against the possibility of worsening symptoms. In a study of patients with Alzheimer disease who initially had a response to risperidone for psychosis or agitation/aggression, the rate of relapse was higher among those who were then randomized to placebo versus those who continued risperidone (60 vs. 33 percent, respectively). A Cochrane Database review concluded that chronic antipsychotics may be withdrawn from patients, but the risk of relapse may be greater for patients with severe neuropsychiatric symptoms. Other pharmacological alternatives include benzodiazepines that may be effective for behavioral problems; however, they have been linked to adverse outcomes such as increased confusion, falls, hip fractures, and excessive sedation. Information on the efficacy and safety of other

pharmacological options such as selective serotonin reuptake inhibitors (SSRIs) and other antidepressants, carbamazepine (Tegretol), divalproex sodium (Depakote), and cognitive enhancers is limited and inconclusive. These psychotropics should be used in short therapeutic trials and ongoing use of pharmacological agents for the treatment of behavioral disturbances should be continuously evaluated and adjusted.

Parkinson Disease with Psychosis

Psychotic symptoms can occur in up to 40 percent of patients with Parkinson disease. The symptoms may be present continuously or intermittently over the course of a month and predominantly consist of hallucinations. Most hallucinations are visual and of these, half of the patients also have auditory hallucinations. Hallucinations also tend to be different from schizophrenia and generally have less "emotional content." Delusions can occur in 5 to 10 percent of patients treated with antiparkinsonian drugs and patients are typically paranoid in nature, such as in spousal infidelity. Psychosis is stressful for patients and can potentially cause more distress to caregivers. Psychosis is a poor prognostic indicator for institutionalization, dementia, and mortality. Managing psychosis is one of the most difficult challenges in the care of patients with Parkinson disease. The dopaminergic drugs that can improve the motor disability caused by Parkinson disease are associated with the emergence of psychotic symptoms and appear to be dose-related. Conversely, the use of antipsychotic drugs to reduce psychosis can exacerbate motor symptoms of parkinsonism. Additionally, anticholinergic medications can cause delirium with agitated behavior or psychosis. Management of psychosis in Parkinson disease begins with withdrawal or lowering of the dose of the dopaminergic agent without disrupting motor signs. Due to the complexity of balancing the dopaminergic requirement to control both symptoms, psychosis presents a serious threat to the patient's ability to maintain independence. If psychosis persists, antipsychotic medications with the lowest liability for exacerbating parkinsonism may be helpful (clozapine [Clozaril] and quetiapine). Aripiprazole (Abilify) is considered third line by the American Academy of Neurology task force recommendation, although the drug has been reported to worsen motor symptoms in Parkinson disease patients.

Delirium

Delirium is a cognitive disorder that is characterized by fluctuating levels of consciousness, inattention, memory impairment, and perceptual disturbances. Delirium may be accompanied by illusions, visual hallucinations, and delusions. Delirium can arise from acute onset of a new medical illness, an exacerbation of a pre-existing chronic illness, a drug reaction or interaction, or a fever or trauma. A careful medical,

neurological, and laboratory evaluation to identify the conditions that are contributing to the mental status change is required. Drugs with anticholinergic properties are common causes of delirium in the elderly. Diagnosis in the older patient may be difficult because peripheral anticholinergic signs are not always present. Recognition and treatment of delirium are important to reduce the risks of morbidity and mortality in these patients. Nonpharmacological treatment such as environmental changes, behavioral support, and redirection are key strategies for treating delirium. Antipsychotic medication should be prescribed only when necessary, to maintain safety and basic care, and these medications can contribute to the mental status changes. The lowest effective dosage should be used, and treatment should continue no longer than required to manage marked distress or behavioral disturbance.

Depression

Patients with delirium and dementia are at high risk of depression and depression may be a predictor for cognitive decline. Suicide risk is also important to assess since suicidal ideation can be seen in up to 30 percent of older adults with major depression. Psychotic symptoms may also occur in patients suffering from depression. In contrast to nonpsychotic depression, elderly patients with psychotic depression are at increased risk for relapse, suicide, and hospitalization. Major depressive disorder with psychotic features responds better to treatment with a combination of an antipsychotic and an antidepressant than to treatment with either agent alone. Electroconvulsive therapy is highly effective in major depressive disorder with psychotic features and at times may be considered a first-line treatment, even in the elderly. Aripiprazole is also indicated for adjunctive treatment for nonpsychotic depression; however, the benefit of using multiple medications that may present compliance challenges in elderly patients must be clearly warranted.

Bipolar Disorder

Bipolar disorder is not uncommon in the elderly and often leads to functional impairment and substantial utilization of health care resources. While conventional mood stabilizers such as lithium and valproate are still widely used for bipolar disorder, atypical antipsychotics have also emerged as first-line options for control of manic symptoms. Currently, aripiprazole, asenapine [Saphris], olanzapine, quetiapine, risperidone, and ziprasidone [Geodon] are approved for mania. Aripiprazole, quetiapine, and ziprasidone are indicated for maintenance treatment while quetiapine, olanzapine (in combination with fluoxetine [Prozac]), and lurasidone [Latuda] are also indicated for bipolar depression. Atypical antipsychotics may offer advantages over drugs like lithium due to their wider therapeutic

range and less drug–drug interactions. However, metabolic and cardiovascular considerations should still be considered in elderly patients. Unfortunately, comparative effectiveness or safety data for antipsychotics in elderly with bipolar disorder are lacking.

TREATMENT PRINCIPLES

Pretreatment Evaluation

If an elderly person presents with new-onset psychosis, an exacerbation of psychosis, or severe behavioral disturbance, he or she should have a thorough medical examination (including neurological assessment) prior to initiating a new, or modifying a pre-existing antipsychotic regimen. At minimum, the evaluation should include a comprehensive history, physical examination, and laboratory screening including a toxicology screen for substances of abuse. This may identify medical or substance-induced problems that may be contributing to the behavioral disturbances (e.g., infection or pain), or that may represent a special concern during treatment (e.g., recent myocardial infarction).

Antipsychotic Selection

The choice of an antipsychotic is usually based on a number of factors and selection should be highly individualized. Factors include specific symptoms, experience with previous treatment (efficacy and side effects), simplicity of regimen (available dosage forms, avoidance of drug interactions, ease of obtaining a therapeutic dose), and cost considerations. With few exceptions, there is limited evidence for superior efficacy of one antipsychotic over another and clinicians should base their decision on the above patient-specific factors.

Dosage Forms

Although all of the antipsychotics are available in oral formulations, only a few are available as injectables (e.g., olanzapine, aripiprazole, ziprasidone, risperidone, and paliperidone [Invega]). Short-acting intramuscular preparations of an antipsychotic may be preferred over oral medications if a patient is agitated and not likely to be cooperative. Intramuscular preparations bypass the gastrointestinal tract and the first-pass metabolism, and have a faster onset of action. For example, most short-acting intramuscular antipsychotics reach a maximum plasma level within 30 to 60 minutes and clinical effects are seen within 15 minutes. While short-acting intramuscular preparations offer administration advantages, care should be taken to avoid repeating doses without first assessing for a response. Alternatively, rapidly dissolving tablet formulations (e.g., risperidone, olanzapine, aripiprazole, and asenapine) offer the advantage of improved adherence in some patients resistant to taking medications or

prone to behaviors such as "cheeking" medications. Except for people with schizophrenia or schizoaffective disorder with a history of poor medication-taking behaviors, long-acting depot medications are not usually prescribed for older people with psychosis.

General Dosing Recommendations

There is considerable heterogeneity in age-related pharmacokinetic and pharmacodynamic alterations. Thus, the efficacy and adverse effects of a particular antipsychotic dosage may vary considerably. In addition, the time required for antipsychotic response is typically longer than in younger patients. Initial dosage should be relatively low, and dose increases should be titrated cautiously, but up to an effective level. Antipsychotic medication dosage in an older patient with a primary psychotic disorder (e.g., schizophrenia) is usually one-half the usual dosage in young patients, although some older patients will respond best to the higher dosages typically prescribed for younger patients. Optimal medication dosage for older patients with primary cognitive or neurological disorder (e.g., dementia) and psychosis is usually 25 to 50 percent of doses typically prescribed to younger patients (Table 57.4d–2).

Adherence Issues

Older patients commonly take several medications simultaneously and therefore have an increased risk of adverse drug reactions. A complete drug history should be obtained and concomitant medications simplified, if possible, prior to implementing antipsychotic treatment. Medication adherence is a problem in all age groups but specific problems may be observed in the elderly. Physical problems, such as impaired vision, may increase the likelihood that instructions will be misread or that one medicine will be mistaken for another. Patients with cognitive impairment may forget to take their medication and drug regimens with complex administration schedules should be avoided. Drug administration issues also present challenges for elderly patients, such as special handling requirement for asenapine (no food or liquid for 10 minutes after dose) and dietary requirements for ziprasidone and lurasidone (350 to 500 kcal required with dose). Psychological factors, such as fear of side effects or becoming dependent on medications, may also affect adherence.

Table 57.4d–2.
Recommended Dosing Guidelines for Antipsychotics

	Dementia		Schizophrenia	
	Initial Dosage in Elderly Patients	Usual Maximum Dosage in Elderly Patients	Initial Dosage in Elderly Patients	Usual Maximum Dosage in Elderly Patients

Medication	Brand Name	(mg/day)[a]	(mg/day)[a]	(mg/day)[a]	(mg/day)[a]
Haloperidol	Haldol	0.25–0.50	0.25–2.0	0.5–1.0	5.0
Clozapine	Clozaril	12.5	75–100	12.5–25.0	100–300
Risperidone	Risperdal	0.25–1.0	1.0–2.0	0.5–2.0	2.0–4.0
Olanzapine	Zyprexa	1.25–5.0	10	2.5–5.0	5.0–15.0
Quetiapine	Seroquel	12.5–50	200–300	12.5–50.0	50–300
Ziprasidone	Geodon	20	Unknown	20	Unknown
Aripiprazole	Abilify	5	15	5–10	10–30
Paliperidone	Invega	3	Unknown	3	3–12
Asenapine[b]	Saphris	10	Unknown	10	20
Iloperidone[b]	Fanapt	2	Unknown	2	24
Lurasidone[b]	Latuda	40	Unknown	40	160
Brexpiprazole[b]	Latuda	1	Unknown	1	4

[a]Routine orally administered initial and maximum dosage ranges are listed. Higher or lower doses may be indicated in some cases.

[b]Optimal dosage ranges listed are based on limited experience.

Drug Interactions

Both pharmacokinetic and pharmacodynamic drug interactions can occur with antipsychotic agents. Cytochrome P450 (CYP) enzymes, especially 1A2, 2D6, and 3A4, are responsible for the metabolism of many antipsychotics. In general, antipsychotics are primarily metabolized hepatically by CYP enzymes but are not major inducers or inhibitors of CYP-mediated metabolism. Thus, antipsychotics do not commonly affect the pharmacokinetics of other medications but are themselves affected by other medications. Induction or inhibition of these enzymes by other drugs may occasionally provide clinically important drug interactions. Drugs that have potential to have serious interactions with antipsychotics include fluvoxamine (1A2 inhibitor); erythromycin, ketoconazole, and ritonavir (3A4 inhibitors); quinidine, risperidone, fluoxetine, and paroxetine (2D6 inhibitors); and cimetidine (multiple systems). Conversely, enzyme inducers such as carbamazepine (3A4) or cigarette smoking (1A2) can reduce plasma concentrations of antipsychotic drugs by increasing the metabolism of the antipsychotic agent and lead to less-than-effective plasma concentration. For example, carbamazepine has been shown to reduce plasma concentrations of haloperidol by 50 percent. Cigarette smoking (1A2) has been shown to increase the clearance rate of clozapine and olanzapine. A pharmacodynamic interaction occurs when medications with similar unwanted properties are concurrently prescribed, such as low-potency typical agents or clozapine with other drugs possessing strong anticholinergic effects (e.g., diphenhydramine or hydroxyzine). Pharmacodynamic drug interactions of greatest concern involve additive orthostasis, anticholinergic effects, QTc prolongation, and sedation (Table

57.4d–3).

PRECAUTIONS AND ADVERSE REACTIONS

Mortality

The black box labeling of antipsychotics by the FDA was prompted by 15 placebo-controlled trials that showed a 1.6- to 1.7-fold increased risk of death among those with dementia taking atypical antipsychotics. Specific causes of these deaths were attributed to heart-related events (e.g., heart failure, sudden death) or infections (mostly pneumonia). In the comparative effectiveness review in 2011 that included all currently available antipsychotics, with the exception of lurasidone, there was a small but statistically significant increased risk of death (effect size between 0.12 and 0.20) in elderly patients with dementia associated with atypical antipsychotics. One prospective cohort study showed conflicting results in that the all-cause crude mortality rate was lower than previous reports and age was the only significant predictor for mortality. Another cohort study also showed that antipsychotics were not associated with early time to nursing home admission or time to death and psychosis was the only significant predictor.

Cerebrovascular Events

Cerebrovascular adverse events, such as strokes, transient ischemic attacks, and fatalities, have been observed in elderly patients with dementia-related psychosis prescribed with aripiprazole, olanzapine, quetiapine, or risperidone. In 17 placebo-controlled trials, higher number of cerebrovascular adverse events was observed in patients treated with atypical antipsychotics compared to patients treated with placebo. Data analyses have suggested that increased age (older than 80 years), concurrent use of benzodiazepines, sedation, pulmonary infections (pneumonia), and medical risks (prior strokes, hypertension, coronary artery disease, and hyperlipidemia) may increase the risk of experiencing a cerebrovascular adverse event.

Whether the mortality and cerebrovascular adverse event risks have a similar underlying mechanism is yet unclear. In a study of over 30,000 Medicare and Patient Assistance Program enrollees, typical antipsychotics were associated with higher risks of stroke and death compared with atypical antipsychotics over 180 days. Interestingly, stroke did not explain the difference in mortality risk between the two drug groups. Orthostatic hypotension seems a likely mediating variable in some patients as wells as falls and oversedation.

Weight Gain

Weight gain has been associated with typical and atypical antipsychotics in

young and older adults. Weight gain is most prominent within first 4 to 12 weeks of treatment with antipsychotics, especially atypical antipsychotics. The cause of antipsychotic-induced weight gain is multifactorial and may include a change in food preferences, increased food or fluid intake, carbohydrate craving, or a sedentary lifestyle. Blockade of the histamine and 5-HT_{2C} receptors are also contributory. Histamine H_1 blockade causes sedation that may produce inactivity. H_1 blockade may also increase weight by interfering with normal satiety signals from the gut, resulting in overeating. In the CATIE-AD study, elderly patients who were treated with olanzapine, quetiapine, or risperidone had an average weight gain of 1.0, 0.7, and 0.4 lb, respectively compared with a monthly weight loss of 0.9 lb for those who received placebo. Weight gain was more prominent in those with a BMI of less than 25 at baseline and this is consistently seen in other age groups. Clozapine has also been associated with significant weight gain. Weight gain associated with antipsychotic treatment should be taken seriously because it may contribute to or exacerbate medical conditions such as diabetes, coronary artery disease, and hypertension. If significant weight gain is observed, a switch to another antipsychotic with lower weight gain liability may be considered as long as treatment effectiveness is not compromised.

Table 57.4d–3.
Dosage Formulations and Drug Interaction Potential for Selected Antipsychotics

			Drug Interaction Potential	
Medication	**Brand Name**	**Available Dosage Forms**	**Pharmacokinetic: Major Cytochrome P450 Pathway**	**Pharmacodynamic: Medications with the Following Effects May Exacerbate Adverse Effect Profile of Atypical Antipsychotic**
Haloperidol	Haldol	T, S, IM, LA	2D6	Anticholinergic, CNS effects
Clozapine	Clozaril	T, ODT	1A2, 2D6, 3A4	Anticholinergic, bone marrow suppressors, CNS effects, hypotension
Risperidone	Risperdal	T, S, ODT, LA	2D6	CNS effects, hypotension
Olanzapine	Zyprexa	T, ODT, IM, LA	1A2	Anticholinergic, CNS effects, hypotension
Quetiapine	Seroquel	T, XR	3A4	CNS effects, hypotension
Ziprasidone	Geodon	C, IM	3A4 (minor)	CNS effects, hypotension, QTc prolongation
Aripiprazole	Abilify	T, IM, S, ODT, LA	2D6, 3A4	CNS effects
Paliperidone	Invega	T, LA	2D6 (limited), 3A4 (limited)	CNS effects

Asenapine	Saphris	ODT	1A2, 2D6	CNS effects, hypotension
Iloperidone	Fanapt	T	2D6, 3A4	CNS effects, hypotension, QTc prolongation
Lurasidone	Latuda	T	3A4	CNS effects
Brexpiprazole	Rexulti	T	3A4, 2D6	CNS effects

Dosage forms: C, capsule; CNS, central nervous system; IM, short-acting intramuscular injection; LA, long-acting intramuscular injection; ODT, orally disintegrating tablet/granule; S, solution; T, tablet; XR, extended-release oral tablet.

Metabolic Effects

Diabetes mellitus and other metabolic disorders have been shown to be more common in adult and elderly patients with schizophrenia than in the general population. Typical and atypical antipsychotics that are used to treat schizophrenia can also worsen glucose intolerance, diabetes, and hypertriglyceridemia in those with or without a history of diabetes mellitus. In rare incidences, these events may lead to diabetic ketoacidosis and death. These complications may or may not be associated with weight gain. Clozapine and olanzapine have been associated with 1.3- to 1.4-fold higher prevalence of diabetes, followed by quetiapine and risperidone. Potential mechanisms for the glucose dysfunction include H_1-receptor antagonism, 5-HT_{2C}-receptor antagonism, reduced functioning of pancreatic cells or pancreatic cell damage, and altered regulation of insulin resistance. In older patients with various psychiatric disorders (e.g., schizophrenia, mood disorders, PTSD, or dementia), treatment with antipsychotics (aripiprazole, olanzapine, quetiapine, and risperidone) for at least 6 months was associated with high incidence of metabolic syndrome. Body mass index, blood pressure, fasting blood glucose, LDL and HDL cholesterol, and triglycerides were the primary metabolic markers studied. Patients also experienced serious and nonserious adverse events at the 2-year follow-up (23.7 and 50.8 percent, respectively). In addition, there was no significant change in psychopathology associated with any of the atypical antipsychotics in the study. Hypertriglyceridemia seems to be due to increased adiposity resulting in higher free fatty acid concentration, insulin resistance, and increased VLDL production. Clozapine and olanzapine, again, have been associated with higher risks of hypertriglyceridemia compared to other antipsychotics in nonelderly patients. In the CATIE-AD trial and other systematic reviews, the association of weight gain and BMI was demonstrated with atypical antipsychotics but there was no association with glucose, total cholesterol, and triglyceride levels. The impact of metabolic syndrome and other adverse events on overall health of elderly patients must be considered in the decision to use these drugs and careful and proactive monitoring of weight, glucose levels, and lipid panel is warranted.

Tardive Dyskinesia

Tardive dyskinesia is a syndrome characterized by abnormal involuntary movements, typically involving the orofacial region, limbs, and trunk. Essential features of tardive dyskinesia include involuntary movements of the tongue, face, and jaw. Most tardive dyskinesia cases are relatively mild. In some older patients, a form of tardive dyskinesia may develop that is severe enough to impair functioning. Severe oral dyskinesia may result in dental and denture problems that can progress to ulceration and infection of the mouth as well as muffled or unintelligible speech. Severe orofacial tardive dyskinesia can impair eating and swallowing, which in turn can produce significant health problems. Gait disturbances due to limb dyskinesia may leave patients vulnerable to falls and injuries. Increasing age is the most consistently observed risk factor for the development of tardive dyskinesia. The annual incidence rates of developing tardive dyskinesia are three- to five-fold greater in older patients compared to younger adults. In addition, elderly patients seem to have earlier tardive dyskinesia symptoms and have more persistent symptoms than younger patients. Other risk factors include higher mean daily and cumulative antipsychotic doses, and presence of extrapyramidal signs early in treatment. The risk of antipsychotic-induced tardive dyskinesia is suggested to be lower with atypical agents compared to typical agents. This lower risk was confirmed in a 2-year study of risperidone and olanzapine in antipsychotic-naïve elderly patients. There is some indication that the risk of tardive dyskinesia may be higher in women, nonwhites, patients with affective psychiatric disorders, and patients taking concomitant anticholinergic agents; however, the heterogeneity of available studies makes it challenging to determine this association conclusively.

Anticholinergic Side Effects

Elderly patients are much more sensitive to the anticholinergic side effects of medication than are younger adults. Anticholinergic medications may aggravate previously existing comorbidities or concomitant medications with anticholinergic properties in elderly patients. Peripheral anticholinergic symptoms may include dry mouth, constipation, urinary retention, tachycardia, blurred vision, and aggravation of narrow-angle glaucoma. Dry mouth may cause problems with dentures; constipation may lead to ileus and perforation; urinary retention may be further aggravated by prostatic hypertrophy; and blurred vision may further aggravate pre-existing visual problems.

The central muscarinic blockade in the elderly also increases their risk for delirium and impaired cognition, attention, memory, and self-care capacity. The higher risk of falls and hip fractures in the elderly associated with antipsychotics is due in part to the anticholinergic, sedative, and orthostatic properties of these medications. In more severe cases, anticholinergic toxicity may be mistaken for an exacerbation of psychosis

and the antipsychotic dose may be increased. Severe anticholinergic symptoms due to high doses of antipsychotics may manifest as increased irritability, aggressiveness, and visual hallucinations.

Sedation

Sedation is a common side effect of antipsychotics that can increase the risk for confusion and falls in the elderly. Antipsychotics with sedative properties are often prescribed to calm agitated patients or to treat sleep disturbances. In a comparative effectiveness review, sedation was more common in elderly patients taking olanzapine or quetiapine than risperidone but this was not statistically significant. Sedation is likely to occur in the elderly at much lower doses than those prescribed to younger patients, and elderly patients are at risk for developing daytime sedation from residual effects of nighttime dosing of the antipsychotic. Elderly patients may appear confused or have other cognitive impairments due to oversedation and should be evaluated appropriately.

Hypotension

Because of decreased adrenergic tone and cardiac output, elderly patients may be more vulnerable to orthostatic effects and may have severe orthostasis at relatively low doses. Orthostatic episodes are most likely to appear early in treatment, and the patient should be monitored closely during this period. Patients (or their caregivers) should be instructed to rise slowly after being seated or supine and take extreme care if they awaken during the night. Lowest doses of antipsychotics should be used and patients should be adequately hydrated throughout the day.

Cardiovascular Effects

In addition to orthostatic hypotension, antipsychotics may also cause tachycardia and electrocardiogram (ECG) changes. Tachycardia may occur as a result of the anticholinergic effects of antipsychotic medications on vagal inhibition, or secondary to orthostatic hypotension. ECG changes such as prolongation of the QT and PR intervals, ST-segment depression, and T-wave flattening have been observed with antipsychotics. The most clinically important ECG change is prolongation of the QTc interval (which is the QT interval adjusted for heart rate). QTc prolongation may lead to the development of ventricular tachyarrhythmias such as torsades de pointes and ventricular fibrillation, which can cause syncope, cardiac arrest, or sudden cardiac death. All antipsychotics have the potential to prolong the QTc interval but to varying degrees. Thioridazine, pimozide, ziprasidone, and iloperidone (Fanapt) appear to have the largest effect on QTc interval. Effects are usually dose-related, although that may not be the case for ziprasidone or iloperidone. Other risk factors for QTc prolongation

and torsades de pointes include advanced age, female gender, and electrolyte imbalance (hypokalemia, hypomagnesemia). Patients with a history of syncope, cardiac disease, or electrolyte imbalance or a family history of early sudden cardiac death should have an ECG before antipsychotic treatment. While thioridazine and pimozide are no longer clinically used, ziprasidone and iloperidone are still therapeutic options for patients with schizophrenia. These drugs should be avoided in patients with a known history of QTc prolongation, recent acute myocardial infarction, or uncompensated heart failure. Clinical and ECG monitoring of cardiac function is recommended during antipsychotic therapy, particularly in patients treated with antipsychotics with known higher risks of QTc prolongation or those who have symptoms that could indicate torsades de pointes (e.g., dizziness, palpitations, syncope). These agents should also not be prescribed with other drugs that prolong the QT interval.

Other Adverse Effects

Clozapine use is associated with a dose-related risk of seizure, and is the only antipsychotic agent associated with the rare development of myocarditis. In rare cases, the orthostatic hypotension can be accompanied by cardiac collapse and/or cardiac arrest. Due to clozapine's significant risk of agranulocytosis, mandatory monitoring of white blood cell counts and absolute neutrophil counts are required. Age and certain ethnicities appear to be risk factors for development of agranulocytosis. Clozapine, thus, is used specifically in patients for whom previous therapies were inadequate or not tolerated, or who have severe tardive dyskinesia.

References

*Alexopoulos GS, Streim J, Carpenter D, Docherty JP; Expert Consensus Panel for Using Antipsychotic Drugs in Older Patients. Using antipsychotic agents in older patients. *J Clin Psychiatry*. 2004;65(Suppl 2):5–99.

Atti AR, Ferrari Gozzi B, Zuliani G, et al. A systematic review of metabolic side effects related to the use of antipsychotic drugs in dementia. *Int Psychogeriatr*. 2014;26(1):19–37.

Ballard C, Orrell M, YongZhong S, Moniz-Cook E, Stafford J, Whittaker R, Woods B, Corbett A, Garrod L, Khan Z, Woodward-Carlton B, Wenborn J, Fossey J. Impact of Antipsychotic Review and Nonpharmacological Intervention on Antipsychotic Use, Neuropsychiatric Symptoms, and Mortality in People With Dementia Living in Nursing Homes: A Factorial Cluster-Randomized Controlled Trial by the Well-Being and Health for People With Dementia (WHELD) Program. *Am J Psychiatry*. 2016;173(3):252–262.

Bingham KS, Whyte EM, Meyers BS, Mulsant BH, Rothschild AJ, Banerjee S, Flint AJ; STOP-PD Study Group. Relationship Between Cerebrovascular Risk, Cognition, and Treatment Outcome in Late-Life Psychotic Depression. *Am J Geriatr Psychiatry*. 2015;23(12):1270–1275.

Declercq T, Petrovic M, Azermai M, et al. Withdrawal versus continuation of chronic antipsychotic drugs for behavioural and psychological symptoms in older people with

dementia. *Cochrane Database Syst Rev*. 2013;(3):CD007726.

*Devanand DP, Mintzer J, Schultz SK, et al. Relapse risk after discontinuation of risperidone in Alzheimer's disease. *N Engl J Med*. 2012;367(16):1497–1507.

Downing LJ, Caprio TV, Lyness JM. Geriatric psychiatry review: differential diagnosis and treatment of the 3 D's—delirium, dementia, and depression. *Curr Psychiatry Rep*. 2013;15:365.

Dubovsky SL, Frobose C, Phiri P, de Greef R, Panagides J. Short-term safety and pharmacokinetic profile of asenapine in older patients with psychosis. *Int J Geriatr Psychiatry*. 2012;27:472–482.

Friedman JH. Atypical antipsychotic drugs in the treatment of Parkinson's disease. *J Pharm Pract*. 2011;24(6):534–540.

Gareri P, De Fazio P, Manfredi VG, De Sarro G. Use and safety of antipsychotics in behavioral disorders in elderly people with dementia. *J Clin Psychopharmacol*. 2014;34:109–123.

Graff-Guerrero A, Rajji TK, Mulsant BH, Nakajima S, Caravaggio F, Suzuki T, Uchida H, Gerretsen P, Mar W, Pollock BG, Mamo DC. Evaluation of Antipsychotic Dose Reduction in Late-Life Schizophrenia: A Prospective Dopamine D2/3 Receptor Occupancy Study. *JAMA Psychiatry*. 2015;72(9):927–934.

Jackson JW, Schneeweiss S, VanderWeele TJ, Blacker D. Quantifying the role of adverse events in the mortality difference between first and second-generation antipsychotics in older adults: systematic review and meta-synthesis. *PLoS One*. 2014;9(8):e105376.

Jackson JW, VanderWeele TJ, Viswanathan A, Blacker D, Schneeweiss S. The explanatory role of stroke as a mediator of the mortality risk difference between older adults who initiate first- versus second-generation antipsychotic drugs. *Am J Epidemiol*. 2014;180(8):847–852.

Jalbert JJ, Daiello LA, Eaton CB, Miller SC, Lapane KL. Antipsychotic use and the risk of diabetes in nursing home residents with dementia. *Am J Geriatr Pharmacother*. 2011;9:153–163.

*Jin H, Shih PA, Golshan S, et al. Comparison of longer-term safety and effectiveness of 4 atypical antipsychotics in patients over age 40: a trial using equipoise-stratified randomization. *J Clin Psychiatry*. 2013;74(1):10–18.

Lenze EJ, Mulsant BH, Blumberger DM, Karp JF, Newcomer JW, Anderson SJ, Dew MA, Butters MA, Stack JA, Begley AE, Reynolds CF 3rd. Efficacy, safety, and tolerability of augmentation pharmacotherapy with aripiprazole for treatment-resistant depression in late life: a randomised, double-blind, placebo-controlled trial. *Lancet*. 2015;386(10011):2404–2412.

*Leucht S, Cipriani A, Spineli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*. 2013;382:951–962.

Lin ST, Chen CC, Tsang HY, et al. Association between antipsychotic use and risk of acute myocardial infarction: a nationwide case-crossover study. *Circulation*. 2014;130(3):235–243.

Lopez OL, Becker JT, Chang YF, et al. The long-term effects of conventional and atypical antipsychotics in patients with probable Alzheimer's disease. *Am J Psychiatry*. 2013;170(9):1051–1058.

*Maglione M, Maher AR, Hu J, et al. *Off-Label Use of Atypical Antipsychotics: An Update. Comparative Effectiveness Review No. 43*. (Prepared by the Southern California Evidence-Based Practice Center under Contract No. HHSA290-2007-10062-1.) Rockville, MD: Agency

for Healthcare Research and Quality; September 2011. Available at https://www.effectivehealthcare.ahrq.gov/ehc/products/150/778/CER43_Off-LabelAntipsychotics_20110928.pdf

Maglione JE, Thomas SE, Jeste DV. Late-onset schizophrenia: do recent studies support categorizing LOS as a subtype of schizophrenia? *Curr Opin Psychiatry*. 2014;27:173–178.

Mehta S, Chen H, Johnson M, Aparasu RR. Risk of serious cardiac events in older adults using antipsychotic agents. *Am J Geriatr Pharmacother*. 2011;9:120–132.

Rado J, Janicak PG. Pharmacological and clinical profile of recently approved second-generation antipsychotics: implications for treatment of schizophrenia in older patients. *Drugs Aging*. 2012;29:783–791.

Rafaniello C, Lombardo F, Ferrajolo C, et al. Predictors of mortality in atypical antipsychotic-treated community-dwelling elderly patients with behavioural and psychological symptoms of dementia: a prospective population-based cohort study from Italy. *Eur J Clin Pharmacol*. 2014;70:187–195.

Rosenheck RA, Leslie DL, Sindelar JL, et al; Clinical Antipsychotic Trial of Intervention Effectiveness-Alzheimer's Disease (CATIE-AD) investigators. Cost-benefit analysis of second-generation antipsychotics and placebo in a randomized trial of the treatment of psychosis and aggression in Alzheimer disease. *Arch Gen Psychiatry*. 2007;64:1259–1268.

Sahlberg M, Holm E, Gislason GH, Køber L, Torp-Pedersen C, Andersson C. Association of Selected Antipsychotic Agents With Major Adverse Cardiovascular Events and Noncardiovascular Mortality in Elderly Persons. *J Am Heart Assoc*. 2015;4(9):e001666.

Sajatovic M, Forester BP, Gildengers A, Mulsant BH. Aging changes and medical complexity in late-life bipolar disorder: emerging research findings that may help advance care. *Neuropsychiatry (London)*. 2013;3(6):621–633.

Seitz DP, Gill SS, Herrmann N, et al. Pharmacological treatments for neuropsychiatric symptoms of dementia in long-term care: a systematic review. *Int Psychogeriatr*. 2013;25(2):185–203.

Vigen CL, Mack WJ, Keefe RS, et al. Cognitive effects of atypical antipsychotic medications in patients with Alzheimer's disease: outcomes from CATIE-AD. *Am J Psychiatry*. 2011;168(8):831–839.

*Woerner MG, Correll CU, Alvir JM, Greenwald B, Delman H, Kane JM. Incidence of tardive dyskinesia with risperidone or olanzapine in the elderly: results from a 2-year, prospective study in antipsychotic-naïve patients. *Neuropsychopharmacology*. 2011;36:1738–1746.

Wustmann T, Pillmann F, Friedemann J, Piro J, Schmeil A, Marneros A. The clinical and sociodemographic profile of persistent delusional disorder. *Psychopathology*. 2012;45:200–202.

57.4e Antidementia Drugs

LON S. SCHNEIDER, M.D.

The medications available for prescription in the United States for treating Alzheimer disease include the cholinesterase inhibitors, donepezil (Aricept), rivastigmine (Exelon), and galantamine (Razadyne, formerly Reminyl); and the noncompetitive *N*-methyl-D-aspartate (NMDA)

antagonist memantine. The cholinesterase inhibitors are specifically indicated in their U.S. Food and Drug Administration (FDA)–approved labeling for "the treatment of mild to moderate dementia of the Alzheimer type." Donepezil is further indicated for severe Alzheimer disease as well. Memantine is indicated for the treatment of moderate to severe Alzheimer disease; the FDA specifically did not approve an indication for its use in mild Alzheimer disease. A capsule containing both donepezil and memantine is also marketed for moderate to severe Alzheimer disease. These drugs, thus, are considered as "antidementia" drugs.

Antidementia drugs are, by definition, medications intended to improve cognitive function in patients with dementia. The term has regulatory significance with the FDA, because it differentiates these medications from "cognitive enhancers," medications that might be expected to improve cognition without reference to a particular illness, and from medications that might improve such behavioral syndromes as depression, aggression, or psychosis in people with dementia.

Worldwide, a range of putative cognitive enhancing agents is offered with and without prescription, including piracetam derivatives; cholinesterase inhibitors; an NMDA antagonist, memantine (Namenda); selegiline (L-deprenyl); antioxidants; vitamins; food supplements, so-called nutraceuticals; and various plant derivatives, including *Ginkgo biloba* standardized extract. In the United States, food supplements and nutraceuticals are regulated by the FDA only with respect to apparent safety and inappropriate advertising claims. No specific health claim or disease-treatment claim should be made by the marketers of nutraceuticals and food supplements, and efficacy evidence for the conditions being treated does not have to be addressed. Rather, marketers of nutraceuticals and food supplements may make a "structure/ function claim," characterizing the supplement as intended to affect normal structure or function or the mechanism by which it acts to maintain such structure or function (21 CFR 101.93).

Although it has been suggested that other medications available in the United States might have efficacy in patients with Alzheimer disease, clinical trial results indicate that selegiline, vitamin E, vitamin E, *Ginkgo biloba* extract, conjugated estrogens, and anti-inflammatories do not have salutary effects—in some cases they have cognitively impairing effects—and should not be used to treat dementia.

Alzheimer disease is the most common dementia associated with aging and hence is the most common target for antidementia therapeutics. Vascular dementia, dementia with Lewy bodies, Parkinson disease dementia, frontotemporal dementia, and dementias associated with hippocampal sclerosis are less common, although there is a large degree of overlap of the pathology of Alzheimer disease with these other pathologies.

Although the causes of Alzheimer disease are unknown, except for rare

early-onset familial and genetic cases, a substantial part of the pathophysiology is sufficiently understood to allow for the rational development of symptomatic or potentially disease-modifying treatments.

The framework for clinical trials in Alzheimer disease developed slowly during the last two decades, after the advent and general acceptance of research-based diagnostic criteria for the dementia of Alzheimer disease (i.e., the essentially identical National Institute of Neurological Communicative Disorder [NINCDS] Alzheimer's Disease and Related Disorders Association [ADRDA] criteria for probable Alzheimer disease, and the *Diagnostic and Statistical Manual of Mental Disorders,* fourth edition text revision [DSM-IV-TR] criteria for dementia of the Alzheimer type), a better understanding of its underlying pathology, and the development of mechanism-based pharmacological therapeutics. For example, the pathophysiology of Alzheimer disease is characterized, in part, by early loss of synapses, death of neurons, the deposition of insoluble β-amyloid peptide fragments, atypical inflammatory responses, and the excess production of toxic, oxidizing free radicals, thus providing targets for drug development.

Until the 1980s, medications were proposed for dementia treatment based on clinical experience or prevailing theories of dementia and aging. Examples of such earlier claimed treatments (and their rationale) include nonspecific psychostimulants (to increase arousal), vasodilators (to increase blood flow), and hyperbaric oxygen (to enhance metabolism). One medication, ergoloid mesylates, was approved in the United States for the ill-defined condition "senile mental decline," a condition based on vague diagnostic constructs, such as "cerebral insufficiency" or "cerebral deterioration." The mechanism of action of ergoloid mesylates is unknown but may involve effects on complex neurotransmitter and vascular effects. However, it is no longer available or marketed in the United States.

REGULATORY AND METHODOLOGICAL CONSIDERATIONS

FDA guidelines for establishing that a drug has antidementia efficacy include requirements that patients meet accepted diagnostic criteria for a primary well-accepted dementia, such as Alzheimer disease; that trials use double-blind, placebo-controlled parallel-group designs; that at least one trial be of 6 months' duration; and that improvement be obtained on an accepted cognitive instrument and a clinician's global assessment, to ensure a clinically meaningful effect as compared to a potentially measurable but trivial cognitive effect. Improvements in activities of daily living also may be considered as an index of clinical meaning in lieu of clinical global assessments.

From a regulatory perspective, behavioral changes as a result of treatment are of secondary importance despite the clinical importance of such improvement. Specifically, improvement in behavior or functional

activities in the absence of improved cognition is not a sufficient outcome to allow a drug to be labeled "antidementia." A concern is that antidepressants or antipsychotics might be misleadingly advertised as being effective for dementia when their intended use is to treat depressive or psychotic symptoms or aggression.

For clinical trials in Alzheimer disease, the Alzheimer's Disease Assessment Scale—cognitive subscale (ADAS-cog) has become the standard cognitive change measure, and a clinician's interview-based impression of change instrument or an activities of daily living inventory have become the measures of "clinical meaningfulness." The ADAS-cog provides a summary score reflecting measures of word recall, naming, following commands, constructional and ideational praxis, orientation, word recognition, spoken language, comprehension of spoken language, word finding, and recall of test instructions. Although it has proven itself sufficiently sensitive to change with cholinesterase inhibitor treatments in mild to moderate Alzheimer disease, it is not as sensitive to change in mild dementia and mild cognitive impairment due to Alzheimer disease (also known as prodromal Alzheimer disease), and people with severe dementia have difficulty performing the test. More recently, another cognitive battery, the Severe Impairment Battery, has been used as an index of cognitive change in Alzheimer disease trials of patients who are moderately to more severely cognitively-impaired. It is group of brief tests assessing social interaction, memory, orientation, language, attention, praxis, and construction.

For mild Alzheimer disease and mild cognitive impairment due to Alzheimer disease a Neuropsychological Test Battery (an NTB) has been offered that—in its most common form—consists of three tests of immediate and delayed memory, verbal and category fluency, and digit span; and the Repeatable Battery For The Assessment of Neuropsychological Status (RBANS) which similarly assesses memory, language, attention, and visuospatial/constructional ability. Regulatory guidelines for marketing approval do not exist for the treatment of other dementias (e.g., vascular dementia, dementia with Lewy bodies, frontotemporal dementia), mild cognitive impairment (other than MCI due to Alzheimer disease), other cognitive disorders, or age-related cognitive decline (DSM-IV-TR), or for cognitively unimpaired persons, mainly because of a lack of general expert consensus on the validity of the proposed diagnostic criteria. One exception is Parkinson disease dementia, a dementia with onset after the Parkinson disease, and for which rivastigmine is approved by the FDA for this indication.

Recently, however, FDA has proposed draft guidelines for early stage Alzheimer disease, including the entities MCI due to Alzheimer disease and preclinical Alzheimer disease. The guidelines allow for only a composite outcome including assessments of cognition and function combined in order to approve a drug for MCI, and a single cognitive outcome for

provisional approval for a drug for preclinical Alzheimer disease. No drugs have been approved under these guidelines.

Although there are efforts to gain recognition of functionally significant but mild cognitive loss, such as age-associated memory impairment, in which cognition is lower than a young-adult comparison, or age-related cognitive decline, in which memory is impaired compared to an age-appropriate comparison, the vagueness of the definitions and underlying pathology limits their general acceptance, and hence the acceptance of "treatments" for such unrecognized conditions. People with age-associated memory impairment generally do not worsen cognitively at a rate greater than the population, whereas those with mild cognitive impairment have an increased rate of worsening and developing dementia. The results of several clinical trials for improving cognition in mild cognitive impairment using the cholinesterase inhibitors donepezil, galantamine, and rivastigmine and vitamin E, anti-inflammatories, and piracetam (Nootropil) have not shown significant cognitive effects for the drugs.

With respect to unimpaired older persons, nearly all of whom can claim that memory, learning, and other aspects of cognition are not as good as they used to be, there are societal questions of the appropriateness of treatment with cognitive enhancers even if they are demonstrated to be effective. Similarly, generally accepted guidelines do not exist for the treatment of cognitive deficits associated with schizophrenia, attention-deficit disorder, or depression, despite general recognition that these cognitive deficits are important areas of clinical attention. There have been attempts recently to provide clinical criteria for "cognitive impairment associated with schizophrenia" and cognitive outcomes criteria for clinical trials.

PHARMACOLOGICAL TREATMENT OF SYMPTOMATIC BEHAVIORS OR BEHAVIORAL AND PSYCHOLOGICAL SYMPTOMS OF DEMENTIA (BPSD)

The majority of patients with Alzheimer disease experience such behavioral symptoms as depression, anxiety, psychosis, agitation, aggression, apathy, withdrawal, and vegetative disturbances (e.g., sleep or appetite change). These signs and symptoms frequently lead to the use of psychotropic medication treatment (both appropriate and inappropriate), physical restraint, and institutionalization. Behavioral symptoms burden caregivers and are considered among the most stressful aspects of caregiving. A variety of psychotropic medications, however, may be effective for pharmacological intervention and improve many of these symptoms. Evidence suggests that these symptoms occurring in the context of Alzheimer disease can represent manifestations of two syndromes: Psychosis of Alzheimer disease and depression of Alzheimer disease.

The treatment of symptomatic behaviors begins with a careful characterization of the target symptoms. Such symptoms as agitation,

aggression, or delusions can result from chronic or recurrent psychiatric disorders, represent the first manifestation of a new psychiatric disorder, or indicate a change in medical status. (The term *agitation* often is used to describe a range of nonspecific behaviors in the elderly beyond excess psychomotor activity, including wandering, pacing, uncooperativeness, and verbal outbursts. Similarly, the commonly used phrase *behavioral and psychological disorders of dementia* is vague and unspecific, and probably should not be used.) A comprehensive medical evaluation might uncover potentially treatable underlying physical conditions that can exacerbate behaviors. Moreover, patients with dementia are at increased risk for developing superimposed deliria and consequent disruptive behaviors.

Patients with dementia are sensitive to a variety of factors that may influence behavior, including sensory impairment, environmental disruption, and altered daily routines. Consequently, environmental manipulations may be helpful. Many dementia patients experience anger, sadness, and denial, which may improve with psychosocial and psychotherapeutic interventions. Various behavior management techniques, including cognitive-behavioral management, may be helpful for more moderately to severely impaired individuals. Supportive psychotherapy and counseling may be helpful for patients with more mild cognitive impairments. Caregivers often need greater awareness that their own behaviors influence the patient. Counseling about the nature of the patient's illness not only helps relatives cope better, but also is associated with better outcomes for patients, such as delayed institutionalization.

Although the nonpharmacological and pharmacological treatment of symptomatic behaviors in dementia patients is beyond the scope of this section, the suggested interventions are necessary preparatory approaches to the pharmacological treatment of cognitive impairment as well.

DRUGS TO TREAT DEMENTIA

From a conceptual consideration, antidementia drugs can be grouped into two broad categories: Those intended to improve cognitive function symptomatically and those intended to modify the course of illness over the longer-term, as manifested by the potential attenuation of cognitive and functional decline. In practice, these two categories overlap. Medication development for dementia typically focuses on one or the other, depending on the mechanism of action of the new drug. The only drugs currently approved, however, are for symptomatic treatment, and their efficacy and FDA marketing approval was based on 6-month-long, placebo-controlled clinical trials. Table 57.4e–1 lists available drugs to treat the cognitive symptoms of dementia.

Table 57.4e–1.

U.S. Food and Drug Administration–Approved and Marketed Medications for Alzheimer Disease

Medication	Preparations	Initial Dosage	Maintenance Dosage	Comment
Donepezil (Aricept)	5- and 10-mg tablets, and orally disintegrating tablets 23-mg tablet	5 mg q.d.	5 or 10 mg q.d.	10 mg may be somewhat more efficacious, but with more adverse effects. 23-mg tablet is indicated after at least 3 months use of 10 mg; it is not overall more effective than 10 mg.
Rivastigmine (Exelon)	1.5-, 3.0-, 4.5-, and 6.0-mg capsules 4.6-, 9.5-mg, and 13.3-mg transdermal patch	1.5 mg b.i.d. 4.6 mg/24 hrs	Capsules: 3.0, 4.5, or 6.0 mg b.i.d. Patch: 4.6 mg/24 hrs to 13.3 mg/24 hrs	Oral doses of 4.5 mg b.i.d. may be most optimal. May be taken with food. Higher doses of the patch are well-tolerated.
Galantamine (Razadyne)	4-, 8-, and 12-mg capsules; solution, 4 mg/mL 8-, 16-, and 24-mg extended-release capsules	4 mg b.i.d. 8 mg q.d.	8 or 12 mg b.i.d. 16 or 24 mg q.d.	Effective dosage range is 16–24 mg per day; 16 mg per day is the modal optimal dose and has fewer adverse events.
Memantine (Namenda)	5- and 10-mg tablets 7-, 14-, 28-mg tablet	5 mg q.d. 7-mg q.d	10 mg b.i.d. 28-mg q.d.	10 mg q.d. was effective in a trial in nursing home patients. Clinical effects of 28-mg q.d. similar to 10 mg. b.i.d., but has not been compared directly.
Combined memantine and donepezil (Namzaric)	14-mg memantine and 10-mg donepezil capsule; 28-mg memantine capsule; and 10-mg donepezil in one capsule	14-mg memantine/10-mg donepezil	28-mg memantine/10-mg donepezil	Bioequivalence of this combination with co-administration of memantine and donepezil was the basis for marketing approval. Indicated for patients maintained on memantine 10–28 mg per day and donepezil, 10 mg per day.

Note: For the cholinesterase inhibitors: Initial dosages should be maintained for at least 2 and preferably 4–6 weeks before increasing. Adverse events may occur with dosage titration. The U.S. Food and Drug

Administration (FDA)–approved indication for these drugs is "indicated for the treatment of mild to moderate dementia of the Alzheimer type." Donepezil is indicated for severe Alzheimer disease and rivastigmine for Parkinson disease dementia as well. Websites for prescribing information on cholinesterase inhibitors are donepezil, http://www.aricept.com/productinfo.htm; rivastigmine, http://www.pharma.us.novartis.com/product/pi/pdf/exelon.pdf; galantamine, http://razadyneer.com/razadyneer/pages/pdf/razadyne_er.pdf

Galantamine formerly had the brand name Reminyl. For memantine: Starting dose is 5 mg q.d., titrated by 5 mg per day each week to a target dose of 10 mg b.i.d. The FDA-approved indication for memantine is for "moderate to severe dementia of the Alzheimer type." For prescribing information, see http://www.namenda.com. For memantine—extended-release starting dose is 7 mg q.d., increased to 28 mg q.d.

http://www.allergan.com/assets/pdf/namendaxr_pi

Cholinergic Hypothesis

The cholinergic hypothesis of dementia is based on observations of massive decline of the cholinergic basalocortical projections; reductions in the activity of choline acetyltransferase, the key enzyme in acetylcholine synthesis; a marked loss of cholinergic cell bodies in the nucleus basalis in Alzheimer disease; correlations between choline acetyltransferase reduction or nucleus basalis cell reduction and neuritic, β-amyloid–containing plaque densities in cortical areas; and correlation of the cholinergic deficits with the decline in cognitive test performance.

Many other neuronal systems are affected besides the cholinergic system, including pronounced declines in the noradrenergic projections from locus ceruleus to cortex, and general changes in glutamate systems. Notably, there is greater research emphasis on more fundamental pathogenetic processes affecting neurodegeneration, including β-amyloid peptide production and deposition, τ protein phosphorylation resulting in neurofibrillary tangles, inflammatory responses, and the use of medications that might be neuroprotective.

Nevertheless, the observation that cholinergic deficits represent the most pronounced transmitter deficit throughout the dementia process has led to considerable focus on cholinergic restitutive strategies. The cholinergic hypothesis of dementia treatment implies that if the cholinergic deficit is responsible for cognitive or behavioral changes, then potentiation of central cholinergic function should improve cognitive and behavioral impairment associated with Alzheimer disease.

Thus far, there have been three main cholinergic approaches to Alzheimer disease treatment: Use of precursor loading; use of direct, cholinergic agonists; and cholinesterase inhibition.

Precursor Loading and Direct Cholinergic Stimulation

Early investigations focused on the use of acetylcholine precursors,

including choline and phosphatidyl choline (lecithin), analogous to the use of dopamine precursors in Parkinson disease. Numerous trials with precursors, however, have not shown significant results, failing to show a reliable improvement in cognitive function. Thus, even though these products are sold in health food stores, their use cannot be recommended.

Direct cholinergic agonists enhance responsiveness to released acetylcholine and minimize the presynaptic inhibition of further release by intrasynaptic acetylcholine concentrations. Thus far, however, direct muscarinic or nicotinic receptor agonists have not proven useful in enhancing cognition in Alzheimer disease. The older and newer, relatively more muscarinic receptor subtype-specific investigational medications have not proven useful either. Marketed cholinergic agonists, such as arecoline and bethanechol (brand name, Urecholine), cannot be recommended because of their inadequate central nervous system (CNS) penetration, high frequency of cholinergic side effects, and minimal, if any, efficacy.

Although nicotinic agents currently under development, including transdermal nicotine, nicotinic α_7 receptor modulators, and newer nicotinic α_4, β_2 receptor modulators, may improve attention, arousal, and concentration in adults, their usefulness in such cognitive impairment syndromes as Alzheimer disease has not been adequately investigated. They also may have a role in attention-deficit disorder and cognitive impairment associated with schizophrenia. Two α_7 nicotinic receptor modulators are in later-stage clinical development, encenicline (Forum) and MK7622 (Merck) in AD.

Cholinesterase Inhibitors

The most successful approach to the treatment of Alzheimer disease thus far has been cholinesterase inhibition. In early studies, physostigmine demonstrated modest cognitive improvements but required individualized and very frequent dosing because of its very short duration of action (1 to 2 hours), thus limiting any potential clinical usefulness.

Cholinesterase inhibitors inhibit CNS acetylcholinesterase in the synaptic cleft, thus preventing the degradation of endogenously released acetylcholine (Table 57.4e–2). However, progressive presynaptic cholinergic neuron loss continues, so there is a progressively diminishing production of acetylcholine over time.

The three cholinesterase inhibitors that are currently marketed for treatment of the cognitive symptoms of Alzheimer disease, and tacrine which is no longer marketed, can be grouped into two classes according to their structure and mode of inhibition: (1) reversible inhibitor tertiary and quaternary amines (i.e., tacrine, donepezil, or galantamine) and (2) pseudoirreversible inhibitor carbamates (i.e., rivastigmine). The

irreversible organophosphate cholinesterase inhibitor metrifonate was clearly efficacious at higher doses but was not approved for marketing because of myasthenia and respiratory depression, possibly owing to its effects at the myoneural junction in muscles of respiration.

INDIVIDUAL CHOLINESTERASE INHIBITORS: CLINICAL STUDIES, DOSING, AND ADVERSE EFFECTS

For the most part, the several published multicenter trials of cholinesterase inhibitors were designed with the objective of obtaining marketing approval from the FDA. The protocols were fairly similar to each other, generally selecting outpatients with mild to moderate Alzheimer disease, usually with Mini-Mental State Examination (MMSE) scores between 10 and 26, inclusively. Patients in these trials were generally physically healthy, usually treated for 6 months or less, and more likely to be female, and they had a mean age of approximately 72 to 74 years, a decade lower than the median age of Alzheimer disease patients in the United States.

Tacrine

Tacrine is a noncompetitive reversible inhibitor of cholinesterase. It binds near the catalytically active site of the acetylcholinesterase molecule to inhibit enzyme activity. It has other actions as well, including blocking sodium and potassium channels, and direct activity at muscarinic receptors. Statistically significant treatment effects for the 30 mg q.i.d. and 40 mg q.i.d. dosage groups were found on the ADAS-cog, a clinician's interview-based impression of change rating, and activities of daily living in 3- and 6-month clinical trials. Tacrine will not be discussed further because it is no longer marketed. Additional information can be found in the previous editions of this book.

Table 57.4e–2.
Pharmacodynamics and Pharmacokinetics of Marketed Cholinesterase Inhibitors

Drug	Pharmacodynamics	Absorption	Bioavailability (%)	Peak Plasma (hours)	Half-Life (hours)	Protein Binding (%)	Me
Donepezil (Aricept)	Noncompetitive, reversible acetylcholinesterase inhibitor	Not affected by food	100	3–4	70	96	Via
Rivastigmine (Exelon)	Noncompetitive, pseudoirreversible cholinesterase inhibitor, both butyryl and acetyl cholinesterase	Delayed by food	40	1.4–2.6	<5	40	Hy

	inhibitor, may differentially affect different acetylcholinesterase subtypes						
Rivastigmine transdermal patch (Exelon)	(as above)	Not affected by food	100	8–16	3	40	Tra
Galantamine (Razadyne)	Competitive, reversible cholinesterase inhibitor, modulates nicotine receptors	Delayed by food	90	1	7	18	Via

Note: Pharmacodynamic effects of some cholinesterase inhibitors are longer than their elimination half-lives. Drugs that inhibit or induce the cytochrome enzymes above might be expected to increase or decrease blood levels.

Donepezil

Donepezil is a long-acting, piperidine-based, highly selective and reversible acetylcholinesterase inhibitor. Two phase III clinical trials provided pivotal evidence for the drug's FDA approval in late 1996. Additional randomized clinical trials include a trial of 6 months' duration, a Scandinavian study of 12 months, a study of nursing home patients, a 6-month trial in which a large proportion of patients were more severely impaired than in previous trials, and an assessment of activities of daily living.

One non–industry-sponsored, randomized, placebo-controlled trial followed patients over several years, reporting very modest cognitive effects over 2 years but no significant effects on loss of function, nursing home placement, or health economic measures.

Pharmacokinetics. Oral bioavailability approaches 100 percent, with peak concentrations in 4 hours and linear pharmacokinetics. Donepezil is extensively bound to plasma proteins, including α_1-acid glycoproteins. The drug accumulates at a constant rate, reaching steady state in approximately 2 weeks. It has a relatively slow clearance, with a long elimination half-life of 70 hours. It is both excreted unchanged in the urine and extensively metabolized by cytochrome P450 (CYP) 2D6 and 3A4 hepatic enzymes to active and inactive metabolites.

Drug Interactions. Inhibitors of CYP 3A4 and CYP 2D6 (such as ketoconazole [Nizoral] and quinidine [Quinaglute]) and inducers of CYP 2D6 and CYP 3A4 (such as phenytoin, carbamazepine, dexamethasone, and phenobarbital) could inhibit donepezil metabolism or increase its rate of elimination but may not be clinically significant in view of its high protein binding and very low therapeutic concentrations of 2 to 12 nG/mL.

Dosing. Donepezil is initiated at 5 mg per day and then increased to 10 mg per day after 4 to 6 weeks. Raising the dose earlier increases the risk for cholinergic adverse events. Five or 10 mg per day is an effective dose; 10 mg tends to be somewhat more effective than 5 mg when the various trials as a group are evaluated. The efficacy or appropriateness of higher doses has not been established.

A 23-mg formulation was marketed in 2010 intended for patients who are being maintained on 10 mg and for whom a higher dose is thought desirable. In a randomized, controlled trial, efficacy compared to patients who continued on 10 mg per day was not demonstrated.

Rivastigmine

Rivastigmine is a pseudoirreversible, selective acetylcholinesterase inhibitor, and butyrylcholinesterase inhibitor. Four phase III randomized, placebo-controlled, 26-week-long clinical trials were completed of similar design but differing mainly in dosing methods. Two have been published, and the results of the other two have been included in secondary reports.

In the two published trials, doses were titrated weekly during the first 7 weeks to one of two dosage ranges, 1 to 4 mg per day or 6 to 12 mg per day, and dose decreases were not permitted, possibly contributing to lesser tolerability and seemingly more side effects during these stages of treatment.

A transdermal patch formulation has been marketed recently based on a placebo-controlled trial in mild to moderate Alzheimer disease comparing a 17.4-mg patch, 9.5-mg patch, and 6 mg of orally administered rivastigmine twice per day, and another trial using 13.3-mg patch in severe Alzheimer disease.

Pharmacokinetics. Rivastigmine is well absorbed; very little binds to protein. Although the elimination half-life is less than 2 hours, enzyme inhibition lasts approximately 9 hours. Rivastigmine is not metabolized by the hepatic microsomal system. Rather, after binding to acetylcholinesterase, the carbamate portion of rivastigmine is slowly hydrolyzed (hence the "pseudoirreversibility"), conjugated to a sulfate, and excreted renally. Thus, its metabolism is essentially extrahepatic and unlikely to have significant pharmacokinetic interactions.

The pharmacokinetics of the transdermal patch shows a time to maximum concentration of 8 to 16 hours and a 3-hour elimination half-life after the patch is removed.

Drug Interactions. Rivastigmine's metabolism does not depend on liver CYP enzymes; therefore, no drug interactions related to hepatic cytochromes have been observed.

Dosing. The starting dose of rivastigmine is 1.5 mg twice a day, taken with meals. If this dose is well tolerated after a minimum of 2 weeks of treatment, it may be increased to 3 mg twice a day. Subsequent increases to 4.5 mg and then 6 mg twice a day should be based on good tolerability of the current dose and may be considered after a minimum of 2 weeks of treatment. Higher daily doses are associated with better efficacy than are lower doses. The initial transdermal patch is applied at 4.6 mg (5 cm^2) per day and then increased to one 9.5-mg (10 cm^2) patch per day after 4 weeks. The dosage level may be increased to a 13.3 (15 cm^2) mg patch per day after a subsequent 4 weeks; and the usual maintenance doses may be 9.5-mg or 13.3-mg patch per day.

Galantamine

Galantamine, an alkaloid originally extracted from *Amaryllidaceae* (*Galanthus woronowii,* the Caucasian snowdrop) but now synthesized, is a reversible, competitive inhibitor of acetylcholinesterase with relatively less butyrylcholinesterase activity.

Because competitive inhibitors compete with acetylcholine at acetylcholinesterase binding sites, their inhibition is, theoretically, dependent on intrasynaptic acetylcholine concentration in that they will be less likely to bind to sites in brain areas that have high acetylcholine levels. Another characteristic is allosteric modulation of nicotinic receptor sites, possibly enhancing cholinergic transmission by presynaptic nicotinic stimulation.

At least four large clinical trials involving more than 2,400 subjects have been published, generally showing efficacious doses between 8 mg twice a day and 16 mg twice a day, with few adverse effects at the low dose. Four milligrams twice a day was not an effective dose.

Pharmacokinetics. Galantamine is well absorbed and has approximately 90 percent bioavailability, little protein binding, and linear pharmacokinetics, with peak concentrations in approximately 1 hour and an elimination half-life of approximately 7 hours. It is metabolized by CYP 2D6 and 3A4, glucuronidated, and excreted in the urine. There is an approximately 25 percent decrease in clearance of galantamine in CYP 2D6–poor metabolizers.

Drug Interactions. Drugs that are potent inhibitors for CYP 2D6 or 3A4 may increase the availability and plasma levels of galantamine. Coadministration of paroxetine (Paxil) and ketoconazole increases bioavailability of galantamine by 40 and 30 percent, respectively. Galantamine clearance is decreased approximately 25 to 33 percent by concurrent administration of CYP 2D6 inhibitors, such as amitriptyline (Elavil), fluoxetine (Prozac), and quinidine; cimetidine (Tagamet)

increased bioavailability by approximately 16 percent. Galantamine itself does not inhibit CYP 3A4 or 2D6 and has little effect on the metabolism of other drugs.

Dosing. Initial dosing is 4 mg twice a day and should be raised to 8 mg twice a day after 2 to 4 weeks. For patients who are tolerating medication but not responding, the dose can be raised to 12 mg twice a day after another 4 weeks. The initial dose of the extended-release formulation is 8 mg per day and raised to 16 mg per day after 2 to 4 weeks with the option to raise to 24 mg per day.

Adverse Effects of Cholinesterase Inhibitors

The most common adverse events with cholinesterase inhibitors are cholinergically mediated and involve those related to cholinergic excess, including nausea, vomiting, diarrhea, dyspepsia, anorexia, fatigue, and weight loss. In addition, myalgia or muscle cramps occur in approximately 7 to 9 percent of patients, indicating direct myoneural effects (but do not seem to be associated with rivastigmine) which specifically inhibits only certain CNS acetylcholinesterase subtypes. The frequency of the cholinergic side effects generally is less than 15 percent and often no greater than placebo in clinical trials. In general, they tend to be mild and short lived, lasting less than 2 or 3 days, and to occur around the time that dosage is increased. Medication dosage can usually be maintained or temporarily lowered and then raised. Fatigue also tends to occur, and occasionally there is a decrease in heart rate of 4 to 7 beats per minute, possibly due to enhanced vagal effects.

Direct reversible hepatotoxicity was associated only with tacrine, more commonly in women than men, and is the most common abnormal finding with this drug which is no longer marketed.

Contraindications to Cholinesterase Inhibitors

There are few absolute contraindications to cholinesterase inhibitors, and they relate to the unique chemical properties of each drug. Because of the actions of cholinesterase inhibitors, these drugs require caution when used in patients with significant asthma, significant chronic obstructive pulmonary disease, gastrointestinal ulcer disease, cardiac conduction defects, or clinically significant bradycardia. Appropriate considerations are involved in general anesthesia as well, because they may prolong the effects of succinylcholine-type drugs.

Donepezil. The most common gastrointestinal adverse effects of donepezil include nausea, vomiting, diarrhea, and anorexia. Additionally, some patients develop muscle cramps, headache, dizziness, syncope,

flushing, insomnia, weakness, drowsiness, fatigue, agitation, and sleeplessness. Weight loss occurred at twice the rate of placebo in the nursing home study, but not in outpatient trials. Adverse effects occurred at a higher rate when the titration from 5 to 10 mg was made in 1 week compared to 6 weeks.

Rivastigmine. Adverse effects with orally administered rivastigmine were primarily gastrointestinal and occurred in the high-dose (6 to 12 mg per day) group. They occurred mainly during dose escalation and led to withdrawal in one study in 23 percent of the high-dose group, 7 percent of the low-dose group, and 7 percent of the placebo group. Adverse effects that occurred in the higher-dose group were sweating, fatigue, asthenia, weight loss, malaise, dizziness, somnolence, nausea, vomiting, anorexia, and flatulence. In the maintenance phase, dizziness, nausea, vomiting, dyspepsia, and sinusitis occurred more in the 6- to 12-mg-per-day group than in the placebo group.

Adverse effects with the transdermal patch at the 17.4-mg dose were similar to 6 mg twice per day of oral rivastigmine, with the 13.3-mg dose somewhat less, and with the 9.5-mg dose were not significantly greater than placebo.

Galantamine. Galantamine's principal adverse effects were nausea, vomiting, diarrhea, anorexia, weight loss, abdominal pain, dizziness, and tremor. They were more frequent earlier in the course of treatment and during the dosage titration from 16 mg a day to 24 mg a day and higher. In one trial, the effective dosage of 16 mg per day was not associated with overall greater adverse events than experienced by the placebo-treated group.

Clinical Application and Treatment Implementation for Cholinesterase Inhibitors

The differential pharmacology of the cholinesterase inhibitors requires treatment considerations, such as once-per-day versus twice-per-day dosing, oral versus transdermal administration, and acetylcholinesterase inhibition versus butyrylcholinesterase inhibition. Answers are not known.

Donepezil and rivastimine transdermal patch have been demonstrated to be effective in more severe patients as well as in outpatients with mild to moderate Alzheimer disease (MMSE scores between 10 and 26). Some experts believe that cognitive efficacy also is apparent in the more mildly and more severely cognitively impaired, but very few studies have been done in mild Alzheimer disease or mild cognitive impairment and did not show advantageous effects.

Duration of Treatment and Discontinuation

Placebo-controlled clinical trials with cholinesterase inhibitors have lasted only 6 months, with only a few exceptions. In some trials, when medication is discontinued, patients return within 3 to 6 weeks to the cognitive level of the placebo patients, indicating an overall symptomatic effect and that treatment efficacy is longer than 6 months. Three placebo-controlled trials suggest that in some cases efficacy could be longer than 1 year on cognitive and functional measures. A few analyses of databases or open-label extensions of clinical trials suggest that patients who tolerate and continue to take cholinesterase inhibitors over at least 1 year might experience a delay in nursing home placement compared to those who cannot tolerate or do not take them. However, these analyses are biased and probably subject to the "healthy patient effect," whereby patients who are destined to have a milder course of illness or who have active caregivers are the ones who *can* take these medications, so that the apparent long-term effect compared to historical controls is illusory. These observational studies have not been supported by a recent, long-term, placebo-controlled trial.

Discontinuation of cholinesterase inhibitors occasionally has been associated with worsening of cognition and occasional episodes of confusion in some patients. It is generally recommended that medication be tapered before discontinuing. In one trial in which donepezil was tapered and discontinued, or continued, in moderate to severe outpatients who had been treated for at least 3 months, continuing donepezil was associated with slightly better cognitive and functional scores over the course of a year. Adding memantine to donepezil did not show an advantage over continuing donepezil alone; but giving memantine to the patients who discontinued donepezil mitigated any differences in cognitive and function compared to continuing donepezil.

MEMANTINE: CLINICAL STUDIES, DOSING, AND ADVERSE EFFECTS

Overstimulation of NMDA receptors has been implicated as a causative mechanism in Alzheimer disease dementia. In preclinical studies, memantine, a moderate-affinity, uncompetitive NMDA receptor antagonist, protects against the effects of glutamate and calcium-mediated neurotoxicity. It has been available in Germany for the treatment of spasticity, Parkinson disease, and other neurological disorders since 1982, and was approved by the European Union in 2002 and the United States in 2003 for treating moderately severe to severe Alzheimer disease.

Mechanism of Action

Memantine's action is hypothesized as being an open channel blocker without appreciable activity until increased glutamate levels trigger the receptor and cause the ion channel to open. It then enters the channel, occupying a site within it, preventing calcium influx and the depolarization

and hyperactivation of the neuron. Because it is cleared from its site of action quickly, it might provide partial blocking of the channel, depending on intrasynaptic levels of glutamate. As glutamate increases, causing channels to open, memantine may block more, and as glutamate decreases, it may block less, allowing greater neuronal activity. In some respects, memantine can be viewed as a modulator of glutamatergic activity.

Clinical Studies in Moderate to Severe Alzheimer Disease

The effect of memantine on cognition generally has been mixed in clinical trials in moderate to severe Alzheimer disease (i.e., MMSE scores <15). Memantine (10 mg b.i.d.) was investigated for the treatment of moderate to severe Alzheimer disease in a 28-week, placebo-controlled trial of 252 patients. The decline of cognitive function was lessened in memantine patients compared to the deterioration observed in placebo.

Another 24-week, randomized, placebo-controlled trial of memantine in moderate to severe Alzheimer disease showed greater improvement in cognition, global assessment, and function compared to placebo. In this trial, memantine or placebo was administered to patients who had been maintained on stable doses of donepezil for at least 6 months. In fact, average time on donepezil was approximately 2.5 years, with nearly 90 percent on it for more than 1 year before being randomized. Notably, the placebo-treated patients (i.e., patients who continued on donepezil throughout the 6 months of the trial) continued to decline at a rather rapid rate, suggesting that memantine substantially stabilized symptoms over this time, whereas any effect of the cholinesterase inhibitor may have been lost earlier over the disease course. Based on these two trials memantine received marketing approval in the United States.

A third moderate to severe Alzheimer disease trial, however, which was similar in design to the other 6-month trials, failed to show significant effects for memantine compared to placebo.

A fourth moderate to severe Alzheimer disease trial failed to show that memantine was more effective than placebo over the course of a year when given to patients being maintained on donepezil. A fifth clinical trial, also in moderate to severe Alzheimer disease did not find an overall clinical effect for memantine compared to placebo over the longer course of 2 years on average, again in patients maintained on cholinesterase inhibitors, while by contrast it showed positive effects on function for vitamin E.

In a 12-week, randomized, placebo-controlled trial of memantine 10 mg per day in nursing home patients with Alzheimer disease or vascular dementia, memantine showed significant improvement over placebo in global assessment and activities of daily living.

Using the higher dose formulation, memantine, 28 mg and intended for once per day administration showed efficacy over 24 weeks. The 28-mg per day dose has not been directly compared to the 10 mg b.i.d. dose in a single

trial. The placebo-controlled trial of 28 mg per day did not show greater effects than the previous 10 mg b.i.d. placebo-controlled trials.

Clinical Studies in Mild to Moderate Alzheimer Disease

Two of three mild to moderate Alzheimer disease trials (i.e., MMSE scores from 10 to 23), one with patients taking cholinesterase inhibitors, did not show significant results, and a third showed significance only with one analysis. Based on these results, the FDA did not approve memantine for mild Alzheimer disease.

Pharmacokinetics

Memantine absorption is not affected by food intake; bioavailability is 100 percent. Plasma protein binding is 45 percent, and peak plasma concentrations occur 3 to 7 hours after oral intake. It is widely distributed and rapidly diffuses across the blood-brain barrier. The majority of the parent compound is excreted unchanged in the urine, in part by tubular secretion. There is minimum hepatic metabolism; the elimination half-life is 60 to 80 hours.

The 28-mg formulation, an extended-release capsule, shows a C_{max} and 24-hour AUC that is 48 and 33 percent higher than the original memantine formulation with an elimination half-life from 60 to 80 hours.

Drug Interactions

Interactions with memantine are few. There are potential interactions with hydrochlorothiazide, resulting in a lower plasma level of the diuretic or a potentially reduced effect. There are no known interactions with cholinesterase inhibitors. Drugs that alkalinize urine would reduce renal clearance of memantine.

Dosing

Dosing is fairly straightforward, with a titration to 10 mg b.i.d. over the course of 3 weeks, starting at 5 mg q.d. and increasing the daily dose by 5 mg each week; or with 7-mg capsules q.d., and increasing by 7 mg q.d. each week to a dose of 29 mg q.d.

Adverse Events

Interestingly, agitation occurred less often in memantine treatment groups than in placebo groups. In the memantine add-on trial, confusion, generally rated as mild, occurred four times more frequently in the memantine group, and headache twice as frequently. Of note, gastrointestinal symptoms were less frequent, with diarrhea or fecal incontinence occurring half as frequently in the memantine group,

suggesting that the actions of memantine may mitigate some of the cholinergic effects of donepezil.

Clinical Application and Treatment

Memantine is indicated for moderate to severe dementia of the Alzheimer disease type, which generally corresponds to patients with MMSE scores of 14 or less. It generally can be considered as initial treatment for this group if patients have not been treated previously. If they have, then memantine can be added to a cholinesterase inhibitor as well. Because the clinical trials that showed efficacy were no longer than 6 months, treatment duration is unclear, and issues are similar to those of the cholinesterase inhibitors. Whether discontinuation is associated with worsening symptoms is not known.

Although memantine is commonly used with or without a cholinesterase inhibitor in mild Alzheimer disease, clinical trials assessing efficacy in mild to moderate Alzheimer disease have not demonstrated cognitive efficacy for memantine. This is due to a lack of efficacy observed in the mild subgroup of patients in these trials and memantine is not FDA approved for mild Alzheimer disease. Nevertheless, it is commonly used in mild Alzheimer disease with nearly half of mild Alzheimer disease patients in observational research studies receiving it.

OTHER MEDICATIONS AND APPROACHES TO ALZHEIMER DISEASE AND COGNITIVE ENHANCEMENT

Although cholinesterase inhibitors have been the most studied class of medications for Alzheimer disease, a broad range of other drugs and supplements are available for prescribing in some countries or are undergoing clinical trials. Some of these may be approved for marketing in the United States, but for other indications. These approaches might improve symptoms or slow neurodegenerative processes, or both.

Atypical inflammatory responses, production of oxidizing free radicals, and the production from amyloid precursor protein of insoluble and possibly toxic β-amyloid peptide fragments are part of the neurodegenerative process. Therefore, anti-inflammatories and antioxidants might have a role in the treatment of progressive dementias. The neurotrophic and cholinergic properties of estrogens were hypothesized to be useful in women.

Huperzine A

Huperzine A is a plant extract that has selective acetylcholinesterase inhibition. It is sold as a food supplement and is a traditional Chinese treatment. A phase II trial failed to demonstrate efficacy at a lower dose of 200 mcg b.i.d., but with a suggestion of a cognitive effect at 400 mcg b.i.d.

Monoamine Oxidase Type B Inhibitors

Some positive effects have been reported with selegiline (L-deprenyl, Eldepryl®), a monoamine oxidase (MAO) inhibitor that relatively selectively inhibits MAO type B at a 5- to 10-mg daily dose. At this dose, its overall effect may be to increase CNS levels of dopamine and trace neurotransmitters, such as phenylethylamine, without affecting norepinephrine levels. Selegiline is marketed for the treatment of Parkinson disease, for which it has demonstrated effects in maintaining motor function, and in some Eastern European countries as a cognitive enhancer.

From a different perspective, selegiline and other MAO type B inhibitors may have neuroprotective potential. Elevated levels of MAO type B in the brains of Alzheimer disease patients might lead to elevated levels of neurotoxins, such as 6-hydroxydopamine, quinones, and free radicals. It has been postulated also that exposure to substances in the environment could lead to neurotoxic products (via oxidative biotransformation). In either case, inhibition of MAO on a chronic basis could theoretically retard progression of the illness based on antitoxic mechanisms. Also, selegiline may have neuronal "rescue" effects, based on neurotrophic-like properties not involving MAO inhibition. It is notable that rodents and dogs treated with selegiline have increased survival. Selegiline is also marketed in the United States as an FDA-approved veterinary medicine for "canine cognitive dysfunction."

A 12-month placebo-controlled trial of transdermal selegiline in patients with milder Alzheimer disease revealed no significant effects overall. Oral selegiline, 5 mg twice a day, along with and separately from vitamin E, 1,000 IU twice a day, has been associated with prolonged maintenance of activities of daily living and avoidance of nursing home placement in moderately to severely impaired Alzheimer disease patients, but with no improvement in cognitive function. An adequate efficacy trial has yet to be completed in Alzheimer disease.

Another, recent, 12-month, phase II, placebo-controlled trial of the MAO-I, RO4602522 (EVT 302) did not show efficacy in 495 patients with moderate Alzheimer disease.

Ladostigil is a derivative of rasagiline and rivastigmine. It acts as a cholinesterase inhibitor and MAO-I; it has antioxidant properties and modulates the processing of amyloid precursor protein. It is currently being assessed in a mild to moderate Alzheimer disease trial, and in a phase II trial in mild cognitive impairment. In sum, there is little evidence for the efficacy of MAO type B inhibitors.

Ergoloid Mesylates

Ergoloid mesylate was a commonly used putative cognitive-enhancing drug

until the advent of the cholinesterase inhibitors and is no longer marketed in the United States. It was used primarily for patients with dementia, age-associated cognitive symptoms, or memory complaints. The agent has mixed and complex actions on central α-adrenergic, dopaminergic, and serotonergic receptors.

Nootropics

The term *nootropic* is derived from the Greek *noos* (mind) and *troepin* (forward) and refers to a drug class with diverse actions, structurally related to piracetam. They are pyrrolidine derivatives of γ-aminobutyric acid (GABA), although they do not appear to have GABA-like effects, and are postulated to have a neuroprotective effect on the CNS against hypoxia, electroconvulsive treatment, and drug intoxication. They also may enhance the CNS microcirculation by reducing platelet activity and adherence of red blood cells to vessel walls, and may stimulate central cholinergic activity. They include piracetam, oxiracetam (Neuromet), pramiracetam (Neupramir), and aniracetam (Draganon). They may improve memory and learning, but an antidementia mechanism has not been established.

Clinical trials of nootropics in Alzheimer disease have shown mixed results with piracetam in the treatment of dementia in the elderly. Pramiracetam and oxiracetam have been evaluated in large-scale multicenter studies with no sufficiently significant clinical effects in dementia or in mild cognitive impairment. Nootropics, however, are marketed for the treatment of dementia and cognitive impairment in countries other than the United States. There is some evidence for efficacy in attention-deficit disorder.

Ginkgo Biloba

The ability of an extract from the leaves of *Ginkgo biloba* to improve cognitive function in Alzheimer disease patients is the subject of ongoing work. Preclinical evidence suggests that certain flavonoid components in *Ginkgo* might have neuroprotective effects, including scavenging of oxygen free radicals; protection of membranes against lipid peroxidation; inhibition of β-amyloid formation, aggregation, and toxicity; and antiapoptotic effects, and may increase hippocampal muscarinic receptors and enhance choline uptake.

A preparation (EGb 761) is approved in Germany for "disturbances of cerebral function related to dementia syndromes," supported by various clinical studies. However, a clear role for *Gingko,* as a symptomatic agent or in the prevention of Alzheimer disease, has not been determined. One 52-week placebo-controlled, double-blind trial of 309 patients with Alzheimer disease or vascular dementia demonstrated that 120 mg per day of standardized *Ginkgo* extract (EGb 761) conferred a statistically

significant improvement on the ADAS-cog scores compared to placebo. A 6-month placebo-controlled trial of 120 and 240 mg per day of EGb 761 for mild to moderate Alzheimer disease, however, did not show efficacy. A recent Cochrane meta-analysis does not support advantageous effects of *Ginkgo biloba* extract for Alzheimer disease or as a cognitive enhancer. A randomized, placebo-controlled trial (the Gingko Evaluation of Memory Study [GEMS]) funded by the National Center for Complementary and Alternative Medication evaluated *Ginkgo biloba* EGb 761 extract, 240 mg per day, in delaying the onset of Alzheimer disease or mild cognitive impairment in 3,072 subjects aged 75 years and older who were followed over 5 years and found no effect. Another prevention study in 200 subjects 85 years or older, also sponsored by the National Center for Complementary and Alternative Medication, did not find an effect for 240 mg per day of EGb 761 in delaying progression to mild cognitive impairment. Finally, a multicenter prevention trial in France involving 2,854 patients followed for 5 years did not show a positive effect for EGb 761 for delaying dementia or improving cognition.

Based on these studies, there is little or no evidence to recommend *Ginkgo biloba* for the treatment of Alzheimer disease or vascular dementia. Studies in normal elderly subjects have been inconsistent, presenting a confusing picture for its use, perhaps partly owing to variations in dose, duration, and methodological issues such as lack of blinding.

Vitamin E

The antioxidant effects of vitamin E (DL-α-tocopherol) might have an impact on reducing clinical progression. In the prototypical antioxidant trial, vitamin E (1,000 IU twice a day) and selegiline (5 mg twice a day), each given alone to moderately impaired patients with Alzheimer disease, delayed the time until patients required nursing home placement, died, or lost activities of daily living. Overall, the delayed time to these end points was approximately 7 months compared to placebo. There were no effects on cognition on this group of patients, who were more severely impaired than in other studies. Adverse effects included falls and syncope in selegiline patients (9 and 10 percent, respectively), α-tocopherol patients (14 and 7 percent, respectively), and the combination (22 and 16 percent, respectively).

A multicenter trial in amnestic mild cognitive impairment patients reported no significant delay in conversion to Alzheimer disease or in cognitive effects with α-tocopherol over 3 years of treatment. The outcomes of the previous moderate to severe Alzheimer disease trial was supported by a trial comparing vitamin E with memantine and placebo over 2 years on average and showing improved function with vitamin E and not with memantine.

Taken together, considering the absence of cognitive efficacy and the potential risks for treatment, vitamin E cannot be recommended as treatment.

B Vitamins and Homocysteine Reduction

Elevated levels of homocysteine are a risk factor for cardiovascular disease, vascular dementia, and Alzheimer disease. Plasma concentrations rise when homocysteine's metabolism to methionine or cysteine is impaired, which may occur under a number of circumstances, including natural aging. Increased blood levels of homocysteine are associated with damage to vascular endothelial cells through lipid peroxidation and release of endothelium-derived relaxing factor, and may increase the risk of vascular thrombosis. More relevant to Alzheimer disease, homocysteine potentiates the neurotoxic effects of β-amyloid peptides in vitro. Homocysteine plasma levels can be reduced by as much as 30 percent using regimens of B vitamins (e.g., B_{12} 1 mg per day, B_6 25 mg per day, and folic acid 5 mg per day). A placebo-controlled clinical trial of these vitamins over 18 months in patients with mild to moderate Alzheimer disease confirmed that substantial reduction in homocysteine could be achieved but did not show cognitive or functional effects compared to placebo.

HMG-CoA Reductase Inhibitors

Epidemiological evidence suggests that cholesterol lowering and the use of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors ("statins") may impact the development or progression of Alzheimer disease possibly because cholesterol levels modulate the enzymatic processing of amyloid precursor protein and consequently the production of β-amyloid, a main component of neuritic plaques in the Alzheimer disease brain. Furthermore, cholesterol-fed rabbits demonstrate increased β-amyloid in brain, and hypercholesterolemia induced by a high-cholesterol diet results in increased levels of β-amyloid in the CNS. Statin administration in vitro reduces intra- and extracellular β-amyloid levels in neuronal cultures.

Observational studies of patient records demonstrated a 60 to 73 percent lower prevalence of Alzheimer disease in lovastatin (Mevacor)- or pravastatin (Pravachol)-treated patients, and a markedly reduced risk for Alzheimer disease as assessed in an analysis of a clinical database.

Clinical trials, however, generally have not supported the use of statins for Alzheimer disease. A randomized, placebo-controlled trial of pravastatin (the PROSPER study) in 5,804 subjects aged 70 to 82 years with risk factors for vascular disease reported favorable cardiovascular outcomes but no significant effect on cognitive function or disability during an approximately 3-year follow-up period. In a randomized, placebo-

controlled trial of the effect of simvastatin (Zocor) on cardiovascular outcomes in a subgroup of 5,806 patients aged 70 or more years, there was no significant difference in the proportion of patients with cognitive impairment or in the incidence of dementia over the average 5-year follow-up period.

A small-scale trial of atorvastatin (Lipitor) indicated some positive cognitive effects. Two larger placebo-controlled trials, one with simvastatin and another with atorvastatin, including 400 and 641 patients, respectively, treated over 18 months did not show positive effects for the drugs in symptoms or in slowing the clinical decline of Alzheimer disease patients. Thus, these drugs should not be considered as useful for preventing or treating Alzheimer disease.

Docosahexaenoic Acid

A reduced risk of Alzheimer disease associated with fish consumption in epidemiological studies is also associated with omega-3 fatty acids. Docosahexaenoic acid (DHA) is an omega-3 fatty acid found in fish oil, a component of neural membrane phospholipids. Several animal studies show DHA to have antiamyloid, antioxidant, and neuroprotective mechanisms. A carefully conducted multicenter, randomized, placebo-controlled trial of DHA in patients with mild to moderate Alzheimer disease did not show efficacy for DHA.

Souvenaid

Souvenaid (Nutricia N.V.) is marketed outside the United States as a medical food. It contains uridine, choline, phospholipids, eicosapentaenoic acid, DHA, selenium, and vitamins E, C, B12, and B6; folic acid. The rationale is that these are precursors and cofactors for the development of neuronal membranes and intended to support the formation of synapses. This represents both a structure/function food supplement claim and a medical food claim. Clinical trials in Alzheimer disease, however, have not been supportive however of a clinical or functional effect.

Axona

A particular liquid formulation of medium chain fatty triglycerides composed of glycerin and caprylic acid (AC-1204 or Axona®) is intended to create chronic ketosis in an effort to enhance mitochondrial metabolism. A rationale is that reduction in cerebral glucose uptake and utilization is observed in Alzheimer disease. Therefore, the AC-1204 formulation is intended to improve cellular metabolism by providing an alternative energy substrate to glucose. Caprylic acid is metabolized into the ketone bodies acetoacetic acid and β-hydroxybutyric acid which in turn are converted to acetyl-CoA and energy by the tricarboxylic acid (TCA or Krebs) cycle.

The formulation was initially marketed as a medical food for Alzheimer disease and then withdrawn by the FDA because of absent clinical trials efficacy evidence and claims. There is an ongoing phase III trial to assess potential efficacy for its use as a drug rather than as a medical food.

UNPROVEN AND UNSAFE MEDICATIONS TO SLOW CLINICAL PROGRESSION OR DELAY ONSET OF ALZHEIMER DISEASE

Anti-inflammatories

Markers of inflammation in the brains of Alzheimer disease patients suggest a potential therapeutic benefit of anti-inflammatories. Epidemiological studies with nonsteroidal anti-inflammatory drugs (NSAIDs) indicate that these agents may be potentially effective for reducing the incidence or delaying the onset of Alzheimer disease. As examples, (1) patients with rheumatoid arthritis who receive chronic anti-inflammatory treatment have reduced risk for Alzheimer disease compared to controls without rheumatoid arthritis, (2) twins who used anti-inflammatory medications had lesser risk for Alzheimer disease than did their co-twins, and (3) the risk of Alzheimer disease was reduced in cohorts of people using anti-inflammatories in middle-age years before the typical onset of Alzheimer disease.

A placebo-controlled trial of indomethacin (Indocin) initially supported a potential role for at least NSAIDs. Subsequent randomized, placebo-controlled trials of NSAIDs in Alzheimer disease patients have failed to show a clinical benefit. Moreover, 1-year-long, placebo-controlled trials of low-dose prednisone (10 mg) or of NSAIDs by the National Institute of Aging Alzheimer's Disease Cooperative Study and long-term trials of celecoxib (Celebrex) and rofecoxib (Vioxx) did not show advantageous effects.

To investigate the effect of NSAIDs in delaying the onset of Alzheimer disease, an ongoing multisite trial (the Alzheimer's Disease Anti-Inflammatory Prevention Trial) randomized approximately 2,528 patients aged 70 years or older with a family history of dementia to naproxen (Aleve), celecoxib, or placebo, and followed them for 5 to 7 years before the trial was ended early due to cardiovascular toxicity concerns. Results, however, indicated nontherapeutic and cognitively toxic effects with increased risk for Alzheimer disease with both treatments.

In view of the negative clinical trials with NSAIDs and considering their frequent adverse effects, including gastrointestinal bleeding, especially with long-term use, NSAIDs cannot be recommended for the prophylaxis or treatment of Alzheimer disease.

Estrogens

Preclinical evidence amply demonstrates that estrogens improve cognitive

function through cholinergic neuroprotective and neurotrophic effects. For example, estradiol (Estrace) replacement increases choline acetyltransferase and acetylcholine after ovariectomy and enhances learning in ovariectomized rats. Clinical evidence of a possible effect for estrogen replacement therapy (ERT) has included observations of an inverse relationship between ERT and death-certificate diagnoses of dementia, several case-control and cohort studies in which 40 to 50 percent reductions in the risk for developing Alzheimer disease were reported, and several smaller clinical reports suggesting that estradiol, estrone (Estrol), or conjugated equine estrogens (CEE, Premarin) enhance cognitive function in Alzheimer disease.

Acute adverse events with ERT are frequent and generally dose dependent and include endometrial hyperplasia, uterine bleeding (when taken with a progestin), breast tenderness, nausea, and venous thrombosis, among others. Because of the enhanced risk for endometrial cancer when estrogens are given to women with intact uteri, administration of progestins with estrogens has been recommended for women with intact uteri.

During the last few years, a number of clinical trials have been carried out, and except for one very small trial using transdermal estradiol, results with CEE (without a progestin) did not substantiate their efficacy. A 20-week-long and a 1-year-long trial of CEE in hysterectomized women were negative. Indeed, there was a trend in women receiving hormone replacement to worsen cognitively, and—although a relatively small sample size—have greater adverse events, including a few instances of deep vein thrombosis.

Without evidence for efficacy, then, safety is of great concern. The Women's Health Initiative trial followed a large number of women randomly assigned to hormone replacement therapy (CEE with or without the progestin medroxyprogesterone [Provera or PremPro when combined with CEE]) or placebo to assess potential cardioprotective effects of the hormone replacement. In 2002, the arm of the trial that included combination treatment was halted because of a substantial increase of cardiovascular adverse effects. The results of the cognitive outcomes in 4,532 women in the subgroup older than 65 years were reported in 2003, finding that the CEE/medroxyprogesterone-treated women had an increased risk for developing dementia. In 2004, the arm of the trial that included CEE treatment alone also was halted for lack of effect and increased strokes with estrogens. CEE alone was associated with an increased incidence of dementia and worse cognition over the approximately 5-year follow-up period.

It is possible but not at all proven that estrogen treatment in midlife around menopause could be neurotrophic and preventative of the onset of Alzheimer disease, but beyond epidemiological studies, evidence is lacking.

The lack of any salutary cognitive effects, combined with the substantial cardiovascular, stroke, and breast cancer risk over the long term, strongly suggests that there is no role for the use of combined estrogen/progestin or CEE-alone intervention in women over 65 years of age.

ADJUNCTIVE AND COMBINATION THERAPY

It is not known whether combining medications such as cholinesterase inhibitors with NMDA antagonists or vitamins confers added efficacy. There have only been pilot studies comparing combinations of cholinesterase inhibitors with such drugs as selegiline, suggesting only their feasibility and not efficacy. The combination of selegiline and vitamin E did not confer any added benefits over either alone in one study and was associated with greater adverse effects.

The combination of a cholinesterase inhibitor with memantine is another example of the potential use of combined therapy. Although such issues as what drug to use first or whether to start both simultaneously have not been worked out, nor is there evidence that each drug separately has an effect that the other does not, it appears that the medications can be used together safely.

As mentioned, a clinical trial with memantine in patients who were being maintained on donepezil showed positive effects in moderate to severe dementia with an actual reduction in treatment-emergent adverse events in patients who had memantine *added on* to long-term donepezil treatment. However, as above, similar trials in mildly to moderately impaired patients did not show salutary effects for memantine, although the addition of memantine did not increase adverse effects. Nevertheless, it makes theoretical sense that cognitive-enhancing drugs with complementary mechanisms of action may be more effective than monotherapy. Consequently, some experienced clinicians use both drugs in combination even as evidence for efficacy is lacking.

TREATMENT OF MILD COGNITIVE IMPAIRMENT

Operational criteria for mild cognitive impairment are generally taken to include a subjective cognitive (memory) complaint; MMSE within normal limits; neuropsychological tests for memory greater than 1.5 standard deviation below age-appropriate norms; activities of daily living within normal limits; and a Clinical Dementia Rating scale score of 0.5, indicating no dementia. Criteria for mild cognitive impairment due to Alzheimer disease and used for clinical trials, however, are slightly different requiring as well progressive worsening and a delayed memory deficit on the Logical Memory II delayed recall below 9 in most cases.

In research clinics a diagnosis of “amnestic mild cognitive impairment” or “mild cognitive impairment of the amnestic type,” marked by predominant memory impairment, implies a conversion rate to Alzheimer

disease of about 12 to 15 percent per year. However, mild cognitive impairment is not necessarily a predementia stage of Alzheimer disease because many people who fulfill criteria do not progress to dementia and some improve to not being cognitively impaired. The meaning of mild cognitive impairment is further confounded as it is viewed also as a transitional state on the road to the development of Alzheimer disease, as an "at-risk" state with an increased likelihood for dementia over a specified time, and as the very early diagnosis of Alzheimer disease before the full dementia syndrome is recognized.

Although criteria for mild cognitive impairment can be defined by consensus, the diagnosis as variously used suffers from vagueness in conceptualization, with the condition overlapping with Alzheimer disease. The application of mild cognitive impairment criteria seems to vary, however, and the diagnosis does not necessarily predict onset of Alzheimer disease. For example, although all mild cognitive impairment trials nominally used the Mayo Clinic criteria for mild cognitive impairment, when the criteria from three clinical trials were applied to the same clinic cohort, differing prevalences were found from 17 to 51 percent.

Mild cognitive impairment has been a therapeutic target for several clinical trials of cholinesterase inhibitor medications marketed for Alzheimer disease. These trials have ranged from 2 to 4 years in duration. None showed positive effects for the cholinesterase inhibitor in delaying time to dementia onset or at improving cognitive scores beyond 6 to 12 months. Given the vagaries of the definition of mild cognitive impairment, that many people with mild cognitive impairment will not progress, the adverse effects of the medications, and the lack of efficacy in the trials, the use of cholinesterase inhibitors for people with mild cognitive impairment cannot be recommended.

A development for clinical research and trials is to define mild cognitive impairment "due to Alzheimer disease" or the equivalent "prodromal Alzheimer disease" by the above clinical and cognitive criteria but to also require a positive amyloidβ PET scan or low amyloidβ CSF levels to enhance the likelihood that the research participant will progress to Alzheimer disease.

TREATMENT OF VASCULAR DEMENTIA

Although a range of treatments has been proposed for the cognitive deficits of ischemic vascular dementia, more recently conducted trials using particular research criteria for vascular dementia have yielded the best efficacy evidence. Six-month trials of the cholinesterase inhibitors donepezil and galantamine have generally demonstrated very modest cognitive efficacy similar to that shown in Alzheimer disease but often without evidence of global improvement or improvement in activities of daily living. Similarly, two 6-month multicenter trials of memantine also

tended to show improvements in cognition. These trials relied on a cognitive battery designed for patients with Alzheimer disease.

Overall, however, there is insufficient overall evidence for efficacy in vascular dementia. One limitation is that vascular dementia is a heterogeneous condition that is difficult to define clinically and often overlaps with Alzheimer disease. These drugs are not approved for use in vascular dementia. The results of the trials demonstrate that cognitively enhancing medications might be useful in other conditions besides Alzheimer disease.

TREATMENT OF PARKINSON DISEASE DEMENTIA

Patients with Parkinson disease have an increased risk for dementia and about 25 percent may develop dementia after the onset of the illness. The dementia, occurring after the onset of parkinsonian motor symptoms, can be characterized by a particular pattern of impairments in executive function, memory, and attention that is different from other dementias such as Alzheimer disease.

From an FDA perspective, however, a diagnosis of the Parkinson disease dementia is made in patients in whom "a progressive dementia syndrome occurs at least 2 years after a diagnosis of Parkinson disease" in the absence of other causes of the dementia. This definition came about as a result of the FDA approving the cholinesterase inhibitor rivastigmine for "mild to moderate dementia associated with Parkinson disease." Thus, rivastigmine used in the same dosing range as with Alzheimer disease is the only drug indicated for treating the dementia.

β-AMYLOID–BASED RESEARCH INTERVENTIONS IN CLINICAL TRIALS

Other drugs with a broad range of mechanisms are in development. They include β- and γ-secretase inhibitors that prevent the cleavage of amyloid precursor protein into insoluble β-amyloid; vaccines such as synthetic β-amyloid in order to generate β-amyloid antibodies; the use of manufactured antibodies against β-amyloid; blocking the receptor for advanced glycation endproducts (RAGE) from binding amyloid; preventing amyloid fibril formation; and a range of other "antiamyloid" approaches. "Antiamyloid" approaches have been most actively researched and are summarized in Table 57.4e–3.

Table 57.4e–3.
Amyloid-a Peptide-Based Research Interventions in Clinical Trials

Active Immunotherapy

A synthetic Aβ antigen that triggers an immune response. Atrial of AN 1792 (Wyeth) was terminated early due to T-cell, immune-mediated, aseptic meningoencephalitis. No significant differences were found between antibody responder and placebo groups on most outcomes. Other vaccines are under development. A phase II trial of a vaccine consisting of short peptides mimicking native Aβ (AD02 and

AD03, Affiris) was negative. ACC-001(Elan) failed to show efficacy in two phase 2 trials in 520 subjects with mild–moderate and prodromal AD over 24 months. CAD106 (Novartis), an antigen designed to induce N-terminal Aβ-specific antibodies against $A\beta_{1-6}$ without an Aβ-specific T-cell response, is in phase II trials; and ACI-24 (AC Immune) is a liposome vaccine intended to generate an antibody response against aggregated Aβ peptides that is under development.

Passive Immunotherapy

Manufactured monoclonal antibodies may be delivered by intravenous infusion or subcutaneous injection. Bapineuzumab (AAB-001, Janssen), targets fibrillar Aβ, and failed to show efficacy in four, phase III clinical trials. Solanezumab targets soluble, monomeric Aβ; two phase III trials were negative, a third trial, including only mild AD patients is underway. Gantenerumab mainly targets Aβ plaques; a phase 2 trial in prodromal AD was negative; a mild to moderate AD trial is ongoing. Crenezumab targets monomeric or oligomeric Aβ; a phase II trial was negative; a phase III trial is underway, as are trials in autosomal dominant AD. Aducanumab (BIIB037, Biogen) targets Aβ aggregates showed a strong, dose-dependent effect at reducing brain fibrillary amyloid and is in phase III trials.

Human intravenous immunoglobulin G contains polyclonal antibodies against Aβ but did not show evidence for efficacy in a phase II trial.

Secretase Inhibitors

Both β-secretase and γ-secretase abnormally cleave APP creating presumably toxic Aβ protein fragments. Several γ-secretase inhibitors, including semagacestat (Lilly), avagacestat (BMS), and have failed in phase II and III trials because of clinical toxicity.

β-Secretase inhibitors (β-site amyloid precursor protein cleaving enzyme 1, BACE-1 inhibitors) are also under development and in phase II and III trials including the BACE-1 competitive inhibitor MK-8931

Selective $A\beta_{42}$-lowering Agents

Some nonsteroidal anti-inflammatory drugs (NSAIDs) may have direct $A\beta_{42}$ lowering properties by allosterically modulating γ-secretase to produce shorter, less toxic Aβ fragments. The R-enantiomer of flurbiprofen (R-flurbiprofen, tarenflurbil) has no significant cyclo-oxygenase inhibition and is being tested in two phase III, 18-month, placebo-controlled trials with the first to be completed in 2008.

Anti-fibrillogenesis Drugs

Inhibitors of fibrillogenesis or aggregation of Aβ may reduce toxicity and plaque burden. The curry spice curcumin, tramiprosate (homotaurine or Alzhemed), and scyllo-inositol (ELND005) were not shown to be effective. Other drugs are under development.

Moreover, there is a new focus on the development of neuroprotective medications or drugs that might cause neurogenesis or enhance plasticity; and drugs and antibodies directed against tau protein targets. Research and development efforts, of course, continue.

FUTURE DIRECTIONS

Current pharmacological treatment options for improving cognition in Alzheimer disease include the cholinesterase inhibitors and memantine. Other available drugs, food supplements, and vitamins, including anti-inflammatories, available estrogens, B-complex vitamins, vitamin E, medium chain fatty acids, and *Ginkgo biloba* extract have not been proven efficacious, some show considerable toxicity in clinical trials, and cannot be recommended for use.

Ideally, it would be best to treat Alzheimer disease early before the onset of the dementia; however, there are no current drugs for this and

cholinesterase inhibitors have been demonstrated to be ineffective in this regard.

Although no currently available antidementia treatment has shown dramatic effects in improving the cognitive impairments of Alzheimer disease, the improvements nonetheless may be clinically meaningful for some people. Therapeutic results are usually modest, affecting a minority of patients. Cholinesterase inhibitors are the best proven efficacious symptomatic treatments for Alzheimer disease. Patients assessed and treated in clinical trials, however, were usually outpatients with mild to moderately severe dementia and few concomitant medical illnesses. Duration of effect beyond 6 months to 1 year (2 years in one trial) and long-term safety are not known.

The efficacy of these drugs for other cognitive disorders remains to be established, although the FDA has approved rivastigmine for Parkinson disease dementia. The prospects for future treatments with greater efficacy in Alzheimer disease and for cognitive enhancement in elderly patients are encouraging.

References

ADAPT Research Group, Lyketsos CG, Breitner JC, et al. Naproxen and celecoxib do not prevent AD in early results from a randomized controlled trial. *Neurology*. 2007;68:1800–1808.

Aisen PS, Cummings J, Schneider LS. Symptomatic and nonamyloid/tau based pharmacologic treatment for Alzheimer disease. *Cold Spring Harb Perspect Med*. 2012;2:a006395.

*Birks J. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database Syst Rev*. 2006;1:CD005593.

Birks J, Grimley Evans J. Ginkgo biloba for cognitive impairment and dementia. *Cochrane Database Syst Rev*. 2007;(2):CD003120.

*Birks J, Grimley Evans J, Iakovidou V, Tsolaki M. Rivastigmine for Alzheimer's disease. *Cochrane Database Syst Rev*. 2000;(4):CD001191.

*Birks JS, Harvey RJ. Donepezil for dementia due to Alzheimer's disease. *Cochrane Database Syst Rev*. 2006;(1):CD001190.

*Courtney C, Farrell D, Gray A, Hills R, Lynch L. Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double-blind trial. *Lancet*. 2004;363(9427):2105–2115.

DeKosky ST, Williamson JD, Fitzpatrick AL, et al. Ginkgo biloba for prevention of dementia: a randomized controlled trial. *JAMA*. 2008;300:2253–2262.

Dysken MW, Sano M, Asthana S, et al. Effect of vitamin E and memantine on functional decline in alzheimer disease: the Team-AD VA cooperative randomized trial. *JAMA*. 2014;311:33–44.

Emre M, Aarsland D, Albanese A, Byrne EJ, Deuschl G. Rivastigmine for dementia associated with Parkinson's disease. *N Engl J Med*. 2004;351(24):2509–2518.

Emre M, Aarsland D, Brown R, et al. Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov Disord*. 2007;22(12):1689–1707; quiz 1837.

Farlow MR, Salloway S, Tariot PN, et al. Effectiveness and tolerability of high-dose (23 mg/d)

versus standard-dose (10 mg/d) donepezil in moderate to severe Alzheimer's disease: a 24-week, randomized, double-blind study. *Clin Ther*. 2010;32:1234–1251.

Feldman HH, Ferris S, Winblad B, Sfikas N, Mancione L. Effect of rivastigmine on delay to diagnosis of Alzheimer's disease from mild cognitive impairment: the InDDEx study. *Lancet Neurol*. 2007;6(6):501–512.

Gilman S, Koller M, Black RS, Jenkins L, Griffith SG; AN1792(QS-21)-201 Study Team. Clinical effects of Abeta immunization (AN1792) in patients with AD in an interrupted trial. *Neurology*. 2005;64(9):1553–1562.

Grossberg G, Manes F, Allegri RF, et al. The safety, tolerability, and efficacy of once-daily memantine (28 mg): a multinational, randomized, double-blind, placebo-controlled trial in patients with moderate-to-severe Alzheimer's disease taking cholinesterase inhibitors. *CNS Drugs*. 2013;27:469–478.

Henderson ST, Vogel JL, Barr LJ, Garvin F, Jones JJ, Costantini LC. Study of the ketogenic agent AC-1202 in mild to moderate Alzheimer's disease: a randomized, double-blind, placebo-controlled, multicenter trial. *Nutr Metab (Lond)*. 2009;6:31.

Howard R, McShane R, Lindesay J, et al. Donepezil and memantine for moderate-to-severe Alzheimer's disease. *New Engl J Med*. 2012;366:893–903.

*Kavirajan H, Schneider LS. The efficacy and adverse effects of cholinesterase inhibitors and memantine in vascular dementia: a systematic review and meta-analysis of controlled trials. *Lancet Neurol*. 2007;6:782–792.

Kozauer N, Katz R. Regulatory Innovation and Drug Development for Early-Stage Alzheimer's Disease. *New Engl J Med*. 2013;368:1169–1171.

*Leber P. Criteria used by drug regulatory authorities. In: Qizilbash N, Schneider L, Chui H, eds. *Evidence-Based Dementia Practice*. Oxford: Blackwell Science; 2002.

Loy C, Schneider L. Galantamine for Alzheimer's disease and mild cognitive impairment. *Cochrane Database Syst Rev*. 2006;1:CD001747.

McKeith I, Del Ser T, Spano P, Emre M, Wesnes K. Efficacy of rivastigmine in dementia with Lewy bodies: a randomised, double-blind, placebo-controlled international study. *Lancet*. 2000;356:2031–2036.

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease. Report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*. 1984;34:939–944.

*McShane R, Areosa Sastre A, Minakaran N. Memantine for dementia. *Cochrane Database Syst Rev*. 2006;(2):CD003154.

Olin J, Schneider L, Novit A, Luczak S. Hydergine for dementia. *Cochrane Database Syst Rev*. 2002;(2):CD000359.

Petersen RC, Thomas RG, Grundman M, Bennett D, Doody R; Alzheimer's Disease Cooperative Study Group. Vitamin E and donepezil for the treatment of mild cognitive impairment. *N Engl J Med*. 2005;352(23):2379–2388.

*Qizilbash N, Schneider LS, Chui H, Tariot P, Brodaty H, eds. *Evidence-Based Dementia Practice*. Oxford: Blackwell Science; 2002.

Quinn JF, Raman R, Thomas RG, et al. Docosahexaenoic acid supplementation and cognitive

decline in Alzheimer disease: a randomized trial. *JAMA*. 2010;304:1903–1911.

*Rabins PV, Blacker D, Rovner BW, Rummans T, Schneider LS. American Psychiatric Association Practice Guidelines Work Group on Alzheimer's Disease and Other Dementias. Practice guidelines for the treatment of patients with Alzheimer's disease and other dementias, second edition. *Am J Psychiatry*. 2007;164(12):3.

Rafii MS, Walsh S, Little JT, et al. A phase II trial of huperzine A in mild to moderate Alzheimer disease. *Neurology*. 2011;76:1389–1394.

Reisberg B, Doody R, Stoffler A, Schmitt F, Ferris S, Memantine Study Group. Memantine in moderate-to-severe Alzheimer's disease. *N Engl J Med*. 2003;348(14):1333–1341.

Sano M, Ernesto C, Klauber MR, Schafer K, Woodbury P. A two-year, double-blind randomized multicenter trial of selegiline and α-tocopherol in the treatment of Alzheimer's disease. *N Engl J Med*. 1997;336:1216.

Schneider LS. Treatment of Alzheimer's disease with cholinesterase inhibitors. In: Galasko D, ed. *Clinics in Geriatric Medicine: Alzheimer's Disease and Dementia*. Oxford: WB Saunders; 2001:337.

*Schneider LS. Issues in design and conductance of clinical trials for cognitive-enhancing drugs. In: McArthur RA, Borsini F, eds. *Animal and Translational Models of Behavioural Disorders, Volume 2: Neurological Disorders*. New York: Elsevier; 2008.

Schneider LS, DeKosky ST, Farlow MR, Tariot PN, Hoerr R, Kieser M. A randomized, double-blind, placebo-controlled trial of two doses of Ginkgo biloba extract in dementia of the Alzheimer's type. *Curr Alzheimer Res*. 2005;2:541–551.

Schneider LS, Insel PS, Weiner MW; Alzheimer's Disease Neuroimaging Initiative. Treatment with cholinesterase inhibitors and memantine of patients in the Alzheimer's Disease Neuroimaging Initiative. *Arch Neurol*. 2011;68:58–66.

Schneider LS, Mangialasche F, Andreasen N, et al. Clinical trials and late-stage drug development for Alzheimer's disease: an appraisal from 1984 to 2014. *J Inter Med*. 2014;275:251–283.

Shah RC, Kamphuis PJ, Leurgans S, et al. The S-Connect study: results from a randomized, controlled trial of Souvenaid in mild-to-moderate Alzheimer's disease. *Alzheimers Res Ther*. 2013;5:59.

Shumaker SA, Legault C, Kuller L, Rapp SR, Thal L, et al; Women's Health Initiative Memory Study. Conjugated equine estrogens and incidence for probable dementia and mild cognitive impairment in postmenopausal women: Women's Health Initiative Memory Study. *JAMA*. 2004;291:2947–2958.

Tariot PN, Farlow MR, Grossberg GT, Graham SM, McDonald S, Gergel I; Memantine Study Group: Memantine treatment in patients with moderate to severe Alzheimer disease already receiving donepezil: a randomized controlled trial. *JAMA*. 2004;291:317–324.

Vellas B, Coley N, Ousset PJ, et al; GuidAge Study Group. Long-term use of standardised ginkgo biloba extract for the prevention of Alzheimer's disease (GuidAge): a randomised placebo-controlled trial. *Lancet Neurol*. 2012;11:851–859.

57.4f Electroconvulsive Therapy and Other Neurostimulation Treatments

MUSTAFA M. HUSAIN, M.D., SHAWN M. MCCLINTOCK, PH.D., AND PAUL E. CROARKIN, D.O.

Chapters 34.35 and 34.36 comprehensively detail the application of electroconvulsive therapy (ECT) and other neurostimulation therapies, respectively. This particular section applies to the utility of ECT and other neurostimulation therapies in a geriatric population.

ECT is by far the single most effective neurobiological treatment for major depression. In the elderly, it is often the treatment of choice for severe depression. Depressed geriatric patients are especially at high risk for sequela of depression, including psychotic symptoms, malnutrition, weight loss, and inactivity. As a result, ECT is commonly employed in geriatric depression with the added benefit that it is relatively safe and well tolerated compared to medications. ECT often become the first-line treatment recommendation for patients who are acutely suicidal, psychotic, or present with severe melancholia. Recent research studies suggest that among patients receiving ECT, older age is associated with rapid remission. In some instances, mean time to remission is as low as 3 weeks.

After one ECT stimulation, there is increased hippocampal expression in the dentate gyrus of neurotropic and transcription factors, including the neurotropic factors: brain-derived neurotropic factor (BDNF), vascular endothelial growth factor (VEGF), and fibroblast growth factor (FGF) and transcription factors, including: c-fos (which plays an important role in signal transduction, cell proliferation, and differentiation) and the arachidonic acid pathway (COX 2). There are cell proliferation and neurotropic effects of ECT also in the prefrontal cortex, amygdala, and hypothalamus. After several stimulations, there is release of even more neuropeptide factors, including neuropeptide Y, thyrotropin-releasing hormone, and an increase in nerve growth factors. Chronic ECT works by changing gene expression. Thus, relapse after ECT is thought to be due to disrupted re-alteration of genes. Chronic ECT also helps strengthen second messenger systems. With the emergence of treatment-resistant depression, defined as depression that does not improve after adequate and substantial antidepressant treatment, ECT has found a new role in antidepressant algorithms.

Mr. B, a 67-year-old Caucasian man with a long-standing history of depression dating back to his college years, was referred by his outpatient psychiatrist for ECT. In his younger days he self-medicated with cocaine, marijuana, and LSD. He reported worsening depression for the past 10 years, despite medication trials of Prozac, Wellbutrin, Cymbalta, Abilify, Adderall, Pristiq, and 25 years of outpatient psychotherapy. His severely depressed mood made him feel "stuck"; he also complained of anhedonia, poor concentration, anergia, amotivation, poor appetite, and hypersomnia. He was "on the edge" of considering death. Home medications included: Lithium 600 mg daily, Effexor 75 mg thrice daily, Ativan 0.5 mg twice daily as needed, and Trazdoone 50 mg.

He was cleared for ECT and underwent 10 index treatments and his MADRS dropped from 35 to 3. He started playing golf again, gardening, cooking, and per wife was back close to his "normal" self. He transitioned to maintenance ECT, during which time he underwent repair of an ascending aortic aneurysm. After clearance, he continued with ECT because he felt he was on one of the most effective treatment regimens. He continued maintenance treatments for approximately 7 months, but then relapsed, as indicated by a MADRS of 44. He underwent 12 index treatments over 6 weeks duration, was taken off both Lithium (due to interaction with Celebrex) and Effexor, and started on Brintellix, titrated from 10 to 15 mg. Over the next three months he received maintenance treatments, with MADRS back down to a 3 during his last five most recent maintenance treatments and remarkably without any cognitive complaints. Brintellix was discontinued during this time due to GI side effects, and he was switched back to Celexa, titrated to 40 mg.

BRIEF HISTORY AND BACKGROUND

Convulsive therapy was introduced in the early part of the 20th century by Ladislas von Meduna as a treatment for severe catatonia in dementia praecox by use of camphor oil injections. In the late 1930s, Ugo Cerletti and Luigi Bini refined convulsive therapy by using electricity to induce therapeutic convulsions, and this formed the precursor of modern-day ECT. Due to its broad applications, and at times misapplications to a variety of psychiatric diseases in the 1940s, as well as its associated cognitive adverse effects, ECT received substantial criticism from both inside and outside the psychiatric community. These critiques of ECT resulted in its limited use, but they also procured its scientific and systematic development as a clinically beneficial and safe treatment modality.

The use of ECT in elderly populations varies by country. In some countries, particularly those with low to middle economies, its use is limited due to financial limitations and legal restrictions. In the United States, its use is higher in elderly patients relative to younger patients based on its effectiveness and elderly patients' intolerance to pharmacological therapy. The National Institute of Mental Health (NIMH) reported in 1980 that approximately 33,000 patients received ECT across the nation, but there was no age demographic reported. A report of elderly patients who received Medicare between the years of 1987 and 1992 found that the rate of ECT use increased each successive year. For example, from 1987 to 1992, the number of patients treated with ECT increased from 12,000 to 15,560, and the use per 10,000 elderly Medicare beneficiaries minimally increased from 3.9 to 4.2 percent. In 2002, a survey of psychiatric hospitals in the United Kingdom showed that of all the ECT treatments administered, 46 percent of the patients were age 65 or older. Exact estimates of patients who receive ECT in the United States are difficult to ascertain as many states do not require reporting laws. However, in the state of Texas, which does have mandated reporting laws, annual reports of ECT use showed that in 2007 a total of 1,772 patients received ECT and of those, 22 percent were over the age of 65, with the

eldest patient treated age 94. Recent population projections estimate that by 2040 the overall population in general will increase by 41 percent, and those over the ages of 65 and 75 will grow by 160 and 268 percent, respectively. Thus, there could be a substantial increase in ECT use given the higher numbers of elderly patients.

Temporal trends in the use of ECT with elderly patients showed that it has increased over time. Trend analysis that examined 12 million Ontario (Canada) residents between 1992 and 2004 compared ECT use between younger (less than age 65) and older (65 years or older) patients. Results found that ECT use remained stable over the 12-year time span with an overall net increase of only 1.1 percent. Interestingly, relative to younger patients, elderly patients were three times more likely to be treated with ECT.

Surveys conducted outside the United States found conflicting results between practitioners who do and do not favor the use of ECT in elderly patients. Of the 306 geriatric psychiatrists surveyed in the United Kingdom in 2002, the majority felt that ECT is a well-tolerated and beneficial treatment for geriatric patients with severe psychiatric disorders, even those patients who are 80 years of age and older. Importantly, the survey pointed out that the role of the ECT consultant needs to be clarified as often the responsibilities of that position may be ambiguous. On the other hand, a 2004 survey of practitioners in the Netherlands suggested that there is hesitation to utilize ECT in elderly patients as a first, second, or third option. While 84 percent of the 152 psychiatrists noted that the use of ECT should be broadened in elderly patients, approximately 54 percent agreed that it should be a first or second treatment option. Equally important, only 15.1 and 13.9 percent indicated ECT as a first-line treatment for major depression with psychotic features or severe suicidal risk, respectively. Collectively, these surveys suggest that the practice of ECT is influenced by practitioner opinion and experience, and that usage estimates vary per country studied.

CLINICAL INDICATIONS

Elderly patients present with multiple psychiatric and medical illnesses. The frequency of reoccurrences and severity of depressive symptoms often increases with age. Severe depression in the elderly confers high levels of functional impairment and morbidity. The elderly are vulnerable to dehydration, weight loss, immobility, and complications from an array of comorbid medical conditions. ECT is indicated for older patients with severe psychiatric illnesses including affective and psychotic disorders, catatonia of any etiopathology, and, in certain cases, neurological disorders (Table 57.4f–1). Within the affective disorders spectrum, ECT can be used to treat both unipolar and bipolar depression, with and without psychotic features. For unipolar depression, ECT may be especially beneficial for

elderly patients with psychotic or melancholic symptoms.

Mrs. C is a 70-year-old Caucasian woman with 2-year history of severe recurrent major depression and three-and-a-half year history of Parkinson disease, initially treated with Sinemet. She was referred by her outpatient psychiatrist for ECT consultation. She was taking Remeron 45 mg nightly and Wellbutrin XL 150 mg daily. She had been switched from Sinemet to Ropinirole, which due to lack of effect was discontinued in November 2014.

Prominent symptoms included: decreased self-care, decreased appetite with 20 pound weight loss within a year, decreased sleep, marked social withdrawal, spending 22 hours a day in bed, and decreased mobility. She was per family "alive, but not living" and "her smile was not there anymore." She stopped watching TV and preferred keeping the shades down. Mrs. C's Parkinson disease was apparently masking her comorbid depression. She was initiated with ECT treatments. She received 11 index course treatments over 6 weeks and her MADRS improved from 47 to 9 with only mild memory dysfunction, namely with concentration and brief postictal confusion. After the 11th treatment, she began maintenance ECT. She reported feeling less tension and sadness. Her family reported increased interest, initiating outdoor activities with her husband, and overall a brighter affect. She also had improvement with Parkinson disease symptoms after the fifth treatment—both her tremors and handwriting improved. Her outpatient psychiatrist decided once she was in the maintenance phase to switch her from Remeron and Wellbutrin XL to Brintellix, which was titrated from 5 to 10 mg. At last ECT treatment, husband reported continued improvement, especially with motivation.

Table 57.4f–1.
Indications for Electroconvulsive Therapy (ECT)

- ► Major depressive disorder with melancholic features
- ► Major depression with acute suicidality
- ► Major depressive disorder with psychotic features
- ► Major depressive disorder intolerant or resistant to pharmacotherapy
- ► Bipolar disorder with acute mania
- ► Bipolar disorder with depression
- ► Bipolar disorder resistant or intolerant to pharmacotherapy
- ► Catatonia (of any etiopathology)
- ► Schizophrenia intolerant or resistant to pharmacotherapy
- ► Prior positive response to ECT

Severe depression is a common psychiatric illness in geriatric patients with estimates in the range of 2 to 5 million patients with severe and subsyndromal depression, respectively. Depression in all forms late in life results in significant mental and physical complications, impaired neurocognitive function, decreased activities and instrumental activities of daily living, and increased morbidity. The World Health Organization reported that among all medical and psychiatric diseases, severe depression was the second leading cause of world-wide disability. Of significant health interests, patients with depression later in life have an increased risk for developing psychosis, suicidal ideation, suicidal intent, and suicidal completion. Thus, it is critical to provide a therapy that has rapid onset of action in order to abate the acute crises.

Geriatric patients with schizophrenia spectrum illnesses may also benefit from ECT, particularly if the psychotic disorder is a new onset,

acutely exacerbated, or treatment refractory. Schizophrenia affects approximately 1.1 percent of the US population, and most geriatric patients with schizophrenia tend to manifest severe positive and negative symptoms that complicate their clinical prognosis and limit their adaptive functioning. Originally introduced for the treatment of schizophrenia, ECT is helpful in ameliorating both the positive and negative symptoms in patients who are nonresponsive to typical or atypical neuroleptic treatments.

Elderly patients often present with affective or psychotic disorders which are secondary to a medical diagnosis. A thorough examination regarding the etiology of the psychiatric pathology should be undertaken to determine the association. In those instances where the underlying cause is related to the primary medical illness, medical care should be implemented. If no improvement is seen in the psychiatric illness, treatment should be initiated beginning with less invasive options (i.e., pharmacotherapy, psychotherapy, social interventions) and incrementally building up to the administration of ECT. However, in cases where there is increased lethality (i.e., catatonia), ECT may be considered as a first-line treatment. Delirium requires an exhaustive evaluation to identify and address underlying etiologies. However, ECT may improve the neuropsychiatric symptoms of this serious syndrome.

Neurological disorders such as Parkinson disease and refractory epilepsy, which afflict elderly patients and result in severe affective or psychotic disorders, are indications for employing ECT. Elderly patients with Parkinson disease may be at increased risk for severe depression, which may improve with ECT. These patients also often have side effects from drug–drug interactions; hence, ECT may be a favorable alternative to pharmacotherapy. Caution needs to be taken though when administering anesthetic agents during the ECT treatment course. Of note, some of the antiparkinsonian medications can result in depressive and psychotic symptoms. A collection of case studies reporting on a total of 22 patients with Parkinson disease found ECT to be beneficial for depressive and parkinsonian motoric symptoms, though considerable relapse in both motor and affective symptoms occurred. Neuroleptic malignant syndrome, a life-threatening condition related to the use of neuroleptic medications, historically in some cases has been treated with ECT. However, ECT is not recommended as a first-line treatment, and prudent use requires careful monitoring of cardiac adverse effects due to pre-existing unstable blood pressure and muscular rigidity, which could be exacerbated during ECT sessions. Because ECT has anticonvulsant properties as seen with subsequent increases in the seizure threshold after each treatment, ECT can be used to treat elderly patients with treatment-resistant epilepsy or those patients with status epilepticus. The risks and benefits must be weighed when applying ECT as other treatments may be more optimal such

as vagus nerve stimulation, which is approved by the U.S. Food and Drug Administration (FDA) for treatment-resistant epilepsy.

Although not routinely performed, ECT can be administered to elderly patients with dementia. Patients with dementia, such as Lewy body or Alzheimer dementia, frequently develop severe affective disorders, which can be treated with ECT if other therapeutic options fail. Case reports and clinical practice have found ECT to be safe and tolerable and in most cases efficacious, with a 40 percent response rate in elderly patients with dementia. The primary dementia does not appear to be exacerbated by ECT. Although results are discordant, recent work has purported that beta amyloid and apolipoprotein E ε4 allele carriage may serve as biomarkers for ECT outcome in general. Nonetheless, in patients who are diagnosed with dementia, the utmost attention must be paid to the progression of the dementia and other neurocognitive effects in order to minimize cognitive impairments. As such, it is prudent to have a neuropsychological evaluation before commencing with ECT. While ECT is rarely utilized to treat combative or agitated behavior associated with dementia, even when other treatments are ineffective, there is limited research that has documented efficacy for those behavioral symptoms. The use of ECT as a therapeutic modality with patients who are diagnosed with dementia may be legally limited in certain states, and the treatment team would benefit from legal consultation regarding consenting procedures and inclusion of the family in acute and longitudinal care plans.

PRE-ECT EVALUATION

A comprehensive evaluation for elderly patients undergoing ECT should include a complete psychiatric history, neurocognitive evaluation, and medical/surgical history (Table 57.4f–2). The evaluation should examine potential proscriptive factors that could limit the effectiveness of ECT in geriatric patients.

Table 57.4f–2.
Medical Evaluation for Elderly Patients Undergoing Electroconvulsive Therapy

- Physical examination
- Neurological examination
- Complete blood count
- Electrolytes
- Thyroid function
- Electrocardiogram
- Anesthesia workup
- Oral/dental examination
- Brain scan (such as magnetic resonance imaging or computed tomography)
- Chest x-ray
- Spinal series (if clinically indicated)

A thorough and comprehensive psychiatric interview should be conducted that ascertains historical and current psychiatric diagnostic and treatment information. Given that elderly patients may have an extensive psychiatric history, it will be advantageous to obtain prior psychiatric records and to include a family member in the interview process to provide corroborating information. The psychiatric diagnosis needs to be elucidated in order to confirm the utility of ECT. Although elderly patients rarely present with primary personality pathology, this psychiatric domain needs to be assessed as research suggests that personality disorders can limit ECT efficacy. For instance, one small study found that patients with borderline personality disorder had limited benefit from ECT and showed greater relapse after the acute ECT course. However, patients in that study were under the age of 65 and limited research has examined the effects of ECT in elderly patients with personality disorders.

Documentation of the psychiatric treatment history is of vital importance as it will have implications for the ECT treatment course. The historical use of ECT in patients has been found to be a predictor of future safety and efficacy. For instance, patients who benefited from ECT are more likely to benefit from future use of ECT, and those patients who showed no benefit may not benefit from ECT. Nonetheless, the prior ECT procedures need to be examined to ensure that ECT was properly conducted and followed American Psychiatric Association–recommended procedures.

Psychotropic medications that were and were not beneficial need to be recorded in order to determine the level of treatment resistance and to plan the prospective acute, continuation, and maintenance treatment courses. Those psychotropic medications that were found to be beneficial may possibly be added to the acute ECT course to enhance the speed of response and therapeutic advantages, and can be used in the continuation and maintenance phases to prolong remission. Collection of the antidepressant treatment history can be facilitated by standardized forms such as the Antidepressant Treatment History Form (ATHF). For example, the ATHF allows for the quantification of antidepressant medications in order to determine the adequacy of the treatment and the level of therapeutic resistance. Collaboration between the ECT practitioner and the geriatric psychiatrist at the initial psychiatric interview can lay the foundation for successful psychiatric treatment and provide an accurate understanding of the elderly patient's psychiatric pathoses.

A clinician privileged to administer ECT and an anesthesiologist should collect comprehensive medical and surgical histories. Ideally, this evaluation will occur as close as possible to the initial ECT session. Prior surgeries, ECT treatments, and exposure to anesthesia should be

ascertained. Specifically, personal or family histories of malignant hyperthermia or other difficulties with anesthesia must be queried. Elderly patients may have multiple gastrointestinal, dental, cardiac, respiratory, musculoskeletal, or neurological disorders that could complicate anesthesia or ECT procedures. Cardiovascular diseases are a major area of concern for increasing morbidity and mortality in the treatment of elderly patients. These include recent myocardial infarctions (generally occurring within 6 months), congestive heart failure, coronary artery disease, aneurysms, and valvular disease. Brain tumors or conditions that increase intracranial pressure also present unique safety concerns. These conditions require careful clinical judgment, and in some cases (i.e., large, rapidly growing, or metastatic tumors) will present a clear contraindication to ECT. Recent hemorrhagic and ischemic strokes or conditions that predispose patients to cerebral vascular bleeding also pose significant safety risks. Poor dentition, loose teeth, and dentures should be evaluated as special precautions (i.e., customized bite block) may be required during the ECT treatment.

Elderly patients also require a careful physical examination with emphasis on the neurological examination (including funduscopic examinations to assess for papilledema) and dental, cardiac, respiratory, gastrointestinal, and musculoskeletal systems. Medical or specialty consultation for elderly patients can assist the ECT practitioner with the optimization or understanding of an illness rather than providing only "clearance" for ECT. The laboratory evaluation should be individualized and congruent with clinical indications. However, ECT centers should mandate minimum standards for this patient population, which should include a complete blood count (CBC), electrolytes, chest radiographs, and an electrocardiogram. Spinal radiographs should be considered in any geriatric patient with back or neck pain. Brain computed tomography (CT), magnetic resonance imaging (MRI), and electroencephalogram (EEG) are not routine or mandated, but may be more beneficial and prudent in this population as compared to young adults.

The elderly patient's current and prior pharmacotherapy should be reviewed due to interactions between different medications and ECT. Mood stabilizers, such as lithium, taken during the ECT treatment can increase the incidence of delirium after the ECT treatment, and can impact the seizure expression. The unique classes of antidepressant medications differentially impact ECT with certain classes enhancing and others interfering with ECT. Monoamine oxidase inhibitors (MAOIs) can cause drug–drug interactions with anesthetic agents and should be discontinued before beginning ECT treatment. Tricyclic antidepressant medications (e.g., nortriptyline) are usually safe to use with ECT and may increase the therapeutic benefits. Use of selective serotonin reuptake inhibitors (SSRIs) in combination with ECT has received limited attention, but use of selective serotonin-norepinephrine reuptake inhibitors (SNRIs), such as

venlafaxine, has been shown to be beneficial when combined during the acute and maintenance ECT course. Most antipsychotic medications have a synergistic effect with ECT; that is, the combined use of psychotropic medications and ECT can magnify the therapeutic effects. However, augmenting ECT with reserpine should be avoided as it could increase the incidence of mortality. For patients with comorbid anxiety who use benzodiazepines, usage of those medications can reduce the seizure duration and should be tapered or discontinued before treatment initiation. Also, anticonvulsant medications should be tapered or discontinued 24 hours before each ECT treatment as they inhibit seizure induction and may result in increased stimulus dosage relative to seizure threshold in order to elicit beneficial seizure activity. More research is warranted to decipher pharmacotherapy and ECT augmentation strategies, particularly with newer antidepressant and antipsychotic medications.

NEUROCOGNITIVE ASSESSMENT

Neurocognitive assessment of an elderly patient who receives ECT should be conducted before, during, and after the treatment course by a qualified clinical neuropsychologist. The assessment before and after the treatment course should comprehensively assess neurocognitive functions including premorbid intellectual ability, processing speed, attention, verbal and visuospatial memory, working memory, and executive function. Such a comprehensive evaluation allows for the establishment of a baseline to quantify the neurocognitive impact after the treatment course. Also, the baseline testing will aid in ruling out a diagnosis of primary dementia or delirium. Measuring neurocognitive functioning during the treatment course provides treatment-specific information to gauge the effects of ECT.

A thorough neurocognitive examination of patients considered for ECT may not be necessary unless there are indications of severe neurocognitive impairment. For example, research has found that certain cognitive domains such as executive function, attention, and processing speed are minimally impacted by ECT. The cognitive domain that has received the majority of attention is memory. An acute course of ECT has an immediate impact on anterograde learning and memory, but these effects tend to be transient and resolve within a week after the end of the acute phase. Thus, the practitioner designing and administering the neurocognitive battery needs to balance the information to be acquired with patient burden. Especially for elderly patients, neurocognitive assessments can be taxing and counterproductive as the information can be compromised due to fatigue and decreased effort.

Table 57.4f–3.
Neurocognitive Assessment Before, During, and After Electroconvulsive

Therapy (ECT)

Before and after ECT	**Global Cognitive Function** ► Attention ► Executive function ► Learning and memory ► Confrontational naming ► Verbal fluency ► Visuospatial construction **Retrograde Memory** ► Personal information ► Impersonal information
During ECT	**Time to Orientation** ► Eyes open ► Name ► Age ► Day of week ► Place ► Date of birth
After ECT	**Retrograde Memory** ► Personal information ► Impersonal information **Executive Function** ► Cognitive flexibility ► Social interaction

At minimum, before and after the ECT course, the neurocognitive domains to be assessed include global cognitive functioning and retrograde memory (Table 57.4f–3).

In most clinical practices, global cognitive functioning is typically assessed and usually with a standardized measure such as the Mini-Mental State Examination (MMSE) or the Montreal Cognitive Assessment (MoCA). A short measure such as the MMSE or MoCA gathers information to determine whether further tests need to be administered to clarify if any other neurocognitive deficits are present. Measuring global cognitive function succinctly taps into many cognitive domains in a short period of time (e.g., within 5 to 10 minutes), and is usually well tolerated by most elderly patients. Unfortunately, the MMSE, MoCA, and other common measures of global cognitive function are insensitive to the effects of ECT. Consequently, the baseline information will provide limited data to guide clinical treatment. Should those measures be used, it would be beneficial to use age- and education-adjusted normative data, such as found in the Mayo's Older Americans Normative Studies, to account for the age-specific effects on those measures and to fully appreciate the ECT-associated effects.

The data generated from global cognitive function measures can be strengthened by also assessing retrograde memory. To accomplish this,

standard scales can be employed that measure autobiographical information for personal and impersonal details. Personal information regards unique information such as a prior birthday or employment history, whereas impersonal information regards information more of a "pop-culture" nature such as a famous person, famous television program, or a well-known event. If a standard scale is not used, the patient can be quizzed using information gathered from the clinical interview and corroborating information from a significant other. An investigation of 50 patients who were randomized to receive either right unilateral or bitemporal ECT found that patients showed a propensity to recall more personal information as opposed to impersonal information, regardless of the ECT stimulus dosage or efficacy outcome. However, relative to patients randomized to receive right unilateral ECT, those who received bitemporal ECT showed greater retrograde amnesia for personal and impersonal information. An important aspect of this study was the measurement and quantification of retrograde amnesia with a novel scale, which had the advantage of being directly translated into clinical practice. To increase the validity of the retrograde memory assessment, corroborating information should be gathered from a significant other.

During the treatment course, a main domain of cognitive function to assess is orientation, especially the time required to achieve full reorientation after each ECT session. Research has consistently shown that ECT has an impact on orientation and causes a period of disorientation immediately after each ECT session. A historical study compared bitemporal and right unilateral ECT, and found that bitemporal ECT resulted in lengthier times to recover full orientation compared to right unilateral ECT. Although the study involved a small number of patients, it generated invaluable information and set a precedent for quantifying disorientation with an easy-to-use structured method. This method involves a clinician to administer an orientation questionnaire. With a stop-watch or other method to measure time, timing starts as soon as the patient is able to breathe without assistance and ends when the patient accomplishes five out of six orientation criteria. These criteria include opening eyes and correct verbalization of name, place, day of week, age, or date of birth. In one study, time to orientation was associated with the magnitude of retrograde amnesia (i.e., the lengthier time required to achieve orientation results in greater retrograde amnesia); however, this finding has yet to be replicated. Factors such as type of lead placement, dosage level, and pulse waveform, have been found to affect time to orientation, with bitemporal placement, higher dosages relative to seizure threshold, and sine wave resulting in longer times to achieve orientation. Ultimately, neurocognitive functioning is important to assess in order to quantify adverse effects of ECT and to assist in clinical decision making.

CONSENT

Informed consent is not a one-time event; rather, it is a continual process. As such, the initial consenting meeting is a vital component of the ECT treatment as it sets the stage for the consent process. Consenting procedures before ECT is initiated are important, particularly for geriatric patients, given that there may be neurocognitive compromises related to the underlying illness. Many patients with psychiatric illnesses may lack the complete ability to provide informed consent, particularly those with a psychotic disorder, and this can conflict with the provision of treatment. Table 57.4f–4 provides suggested recommendations when obtaining informed consent.

Prior to signing the consent form, prospective patients need a complete and thorough explanation of what ECT consists of, the procedures that will take place, the associated effects (both positive and negative), the length of the treatment, alternatives to treatment, and the continued care after the acute ECT course is completed. All of this should be written in the consent form that the patient will sign. Having an open forum of Socratic questions and answers makes this process informative and respects the patient's autonomy. This initial meeting can be further enhanced by watching a video of the ECT treatment with the patient and clarifying any related questions. The inclusion of the video ensures a comprehensive explanation of the ECT treatment and utilizes multimodal means of learning to help ensure that the patient is fully informed. Further, the initial consenting session can facilitate the establishment of trust between the patient and the doctor, clarify misconceptions, and help demystify ECT as many elderly patients may view it as a negative treatment.

Table 57.4f–4.
Suggested Informed Consent Procedures for Geriatric Patients Receiving Electroconvulsive Therapy (ECT)

- ► Follow all local, state, and federal guidelines for the proper administration of ECT.
- ► Qualify the patient's capacity to provide informed consent.
- ► Engage in an open conversation with the patient and a family member if available regarding the ECT treatment and the rationale for administration.
- ► Utilize multimedia aids such as videos documenting an ECT treatment. Elicit feedback from the patient and answer any questions that arise.
- ► The consent form must describe all aspects of the ECT course including indications and contraindications in individual patients for ECT, ECT treatment procedures, anesthesia, adverse effects and respective benefits, alternatives to treatment, and rights to decline ECT treatment.
- ► Use a standardized questionnaire to assess if the patient understands the consent form.
- ► Have a witness cosign the informed consent form.
- ► Document the consent process in the medical record.
- ► Provide the patient with a copy of the signed consent form and place a copy in the chart.
- ► Informed consent is a process, not a one-time activity, that must be monitored throughout the ECT course.

Building a supportive network for the patient is advantageous during and after the ECT treatment, and the consenting session is an ideal setting to lay that framework. The physician obtaining consent should invite a family member or caregiver to attend the informed consent meeting. This will allow for a thorough discussion of the ECT treatment process, allow the patient to feel that he or she is not alone, and provide the patient with someone to talk to after the discussion. Although it may not be required, having a family member cosign the consent form provides further assurances that consenting procedures are being followed. For example, a retrospective chart review of 62 geriatric patients who received ECT in the New York state area found that 75 percent of those charts contained a family member's cosignature, even in cases where the patient was documented to have capacity to provide informed and voluntary consent. For elderly patients who are cared for at home, the family member is a vital component to the consent process and the overall treatment, as he or she will be primarily responsible for ensuring that the patient receives the appropriate aftercare. Also, the family member may have unanswered questions and negative stigmas associated with ECT that can be clarified in the consenting session.

A common procedure to perform before treating a patient is to assess the patient's capacity to provide informed consent, especially for those patients who may have neurocognitive impairments. To provide informed consent, a patient must have the ability to provide voluntarily and fully informed consent. To assess for consenting, the clinician (i.e., nurse, physician) should review the medical chart, acquire the patient's educational history and information, and administer a global cognitive examination such as the MMSE. These procedures will help guide the clinician in determining the level of capacity by which prospective plans can be arranged. For instance, should the patient have limited understanding of his or her condition and treatment options, additional education regarding ECT, including a family member, and having the patient answer a questionnaire may help to increase the decision-making ability. An empirical study conducted at the Mayo Clinic comparing geriatric (mean age: 75.6, standard deviation [SD] = 5.87) and nongeriatric patients (mean age: 41.9, SD = 8.01) randomized to an additional half-hour session of education regarding ECT provided in a one-on-one format with a psychiatrist found that the provision of extra education increased the patient's knowledge base to inform his or her treatment decision. An elaborate battery of neuropsychological tests can be administered to comprehensively assess for any neurocognitive impairment; however, those procedures may not be necessary and could overburden the patient. Nonetheless, they are available to the physician and could be a source of additional information.

Although it may be rare, some patients may be unable to voluntarily

consent. This could be due to myriad factors including the psychiatric illness (i.e., catatonia), limited ability to make decisions, or blatant refusal. For those patients unable to voluntarily provide consent, the consenting physician should ask for an ECT consultation. The outside ECT consultation, usually performed by the chief of ECT services for the hospital, provides for a secondary opinion of the patient's decisional ability and the immediacy for treatment. Particularly for those patients who refuse ECT, the secondary opinion provides further justification for the dire need of treatment based on the safety of the patient and others, and can be used in legal proceedings to attain court-ordered treatment. As in voluntary consent, it is beneficial to include a family member in this process. The inclusion of the family member can help to mitigate any coercion or perception of coercion of the patient.

The consent process should be completely documented in the patient's medical record including the number of meetings, who was in attendance, educational aids used, neuropsychological tests administered (if any), and the signed consent form. A copy of the signed consent form should be given to the patient. Since the consenting requirements and regulations can vary by state, it is critical to understand the respective laws for the consenting of the patient to receive ECT. For example, in the state of Texas, the patient is reconsented before each ECT treatment is provided. The reconsenting of the patient at each treatment further substantiates that consent is not a one-time decision; rather, it is a continual process and the patient or family member could have new questions arise through each phase of the treatment.

ECT PROCEDURES

Anesthesia

All ECT treatments are provided while the patient is under general anesthesia and for elderly patients an anesthesiology consultation is of significant importance in order to determine the optimal anesthetic doses and minimize side effects. The anesthesia consultation should occur before the first treatment. Anesthetic doses should be re-evaluated before each treatment with subsequent dosage adjustments based on the outcome of the prior treatments. Different classes of anesthetic medications are delivered including neuromuscular blocking agents and hypnotic agents. The interactions between anesthesia and ECT can affect clinical outcome, making it essential to have input and collaboration between the anesthesiologist and the ECT practitioner. For example, certain hypnotic agents can raise the seizure threshold, thereby inhibiting the production of seizure activity and resulting in a postponement of clinical efficacy. Hence, the provision of ECT in elderly patients is dependent upon multidisciplinary team work.

Age-related physiological changes must be taken into account when providing anesthesia during ECT. Many elderly patients age 65 years and older will have slow conduction in peripheral nerves, increased systolic blood pressure, reduced cardiac function, decreased hepatic metabolism, and decreased circulation time. These physiological changes can result in potential delayed onset of anesthetic action as well as prolonged elimination. This can affect the ECT treatment in various ways. For example, as the seizure threshold is already high in elderly patients, proper dosage and timing of administering hypnotic agents must take into account the eliciting of a therapeutic seizure; otherwise, the ECT treatment may be ineffectual. Also, the administration of succinylcholine, a muscle relaxant, must be timed accordingly so that the body is relaxed upon ECT initiation in order to minimize muscle contraction and to allow the patient to regain muscle control in parallel with regaining consciousness.

Different hypnotic agents have been used during ECT treatments, each with their respective physiological properties that impact the ECT expression. A small study that compared the effects of different doses of methohexital, propofol, and etomidate in 10 patients with an average age of 54 years (range: 27 to 81 years) who received bilateral ECT found that propofol and methohexital at high doses significantly suppressed EEG seizure activity. Also, the use of etomidate, although it had minimal effects on seizure expression, resulted in prolonged recovery of cognitive function. A separate small study in seven patients with an average age of 42 years found that methohexital, with part of the dose substituted with remifentanil, produced prolonged seizures relative to methohexital alone. With the different anesthetic agents available and various administration techniques, expertise is paramount in synergistically pairing ECT and anesthesia.

With the use of anesthetic agents, there is the potential risk of increased confusion or disorientation after the ECT treatment. This side effect becomes critical when providing ECT to elderly patients as they are at an increased risk of ECT-associated adverse cognitive side effects in addition to normal age-related cognitive decline. One small study compared the cognitive adverse effects of anesthesia with etomidate or ketamine in ten patients and found that short-term memory, as assessed by asking the patients to recall the names of four objects, was less affected by ketamine. The significance of this study is that it showed that different types of anesthetic medications can differentially affect cognitive function after the ECT treatment; however, the conclusions were limited by a small sample size, underdosage of the ECT treatment (i.e., provided right unilateral ECT at 200 percent above seizure threshold), and the fact that short-term memory was only measured with one nonstandardized cognitive test. On the contrary, in small documented cases, ketamine has been found to result in visual hallucinations, and thus is rarely used in current clinical practices.

Treatment Parameters

ECT treatment parameters can significantly affect its efficacy and tolerability. The three parameters that interact are pulse waveform, electrode placement, and stimulus dosage. Regarding pulse waveform, the first waveform introduced, and the one that generates the greatest neurocognitive adverse effects, is the sine waveform. Although this waveform is rarely used due to technological enhancements in ECT devices and newer pulse waveforms (e.g., ultra-brief), a large-scale research study found that it is employed in some psychiatric practices. However, the majority of academic and public institutions that provide ECT do so with brief or ultra-brief pulse waveform. These two waveforms result in equivalent efficacy of sine wave, but produce minimal neurocognitive side effects. Ultra-brief is the latest waveform developed and has been associated with faster rates of time to orientation after treatment, less neurocognitive adverse impact, and equivalent efficacy.

There has been and there still remains debate as to what is the appropriate site placement for the ECT electrodes. There have been various electrode placement sites studied over the years of ECT research, and three sites have received considerable attention, those being bitemporal, right unilateral, and bifrontal (Fig. 57.4f–1). For all electrode placements, the patient's scalp will need to be prepared by applying a topical prep solution to clean the scalp, which will reduce impedance and allow maximum passage of the electrical stimulus.

For bitemporal electrode placement, each electrode is placed on each temple 1 inch above the midpoint of the line extending from the tragus of the ear to the external canthus of the eye. For right unilateral, one electrode is placed on the right temple 1 in above the midpoint of a line extending from the tragus of the ear to the external canthus of the eye and the other electrode is placed 1 in to the right of the vertex. For bifrontal, each electrode is placed 2.5 in above the outer external canthus of the eye. Bitemporal electrode placement was the first electrode placement used and has been shown to be significantly advantageous relative to the other two placement sites in terms of speed of response and remission. However, right unilateral placement, six times above the seizure threshold, has been shown to produce equivalent response and remission rates despite requiring more time while also having minimal adverse neurocognitive effects. Bifrontal placement has received nominal attention, but has shown high efficacy rates coupled with low adverse neurocognitive effects.

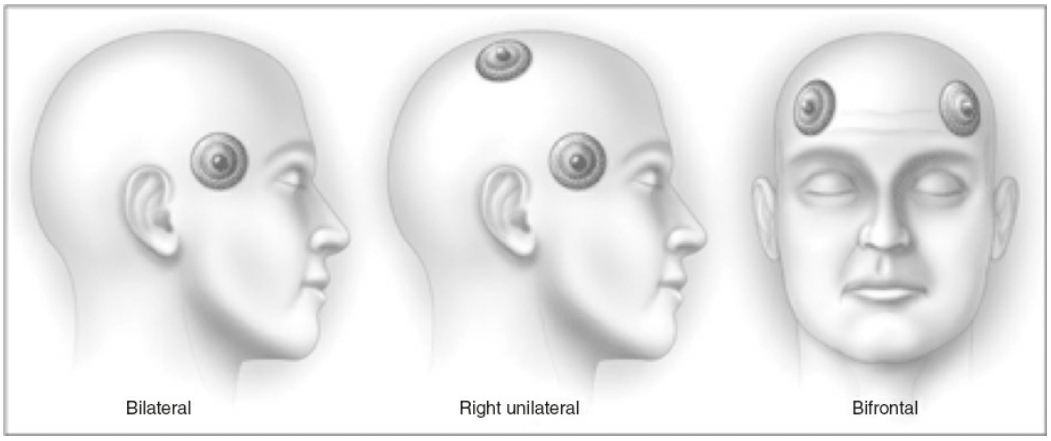

FIGURE 57.4f–1. Electrode placement sites. (From Lisanby SH. Electroconvulsive therapy for depression. *N Engl J Med. 2007;357(19):1939–1945.)*

Stimulus dosing can be performed with one of two methods, either the age-based method or the empirical titration procedure. Both procedures have been found to produce successful inductions of therapeutic seizures. In general, with the age-based method, stimulus dosage is adjusted by age, with younger age groups treated at low doses and older age groups treated with higher doses. Some clinicians prefer age-based titration methods because they may limit the number of subconvulsive stimuli. Regarding this method, either full-age or half-age dosing formulas can be applied. The full-age formula (dose = patient age multiplied by five) is more appropriate to use with right unilateral ECT as it typically equates to doses 250 percent above the seizure threshold. On the other hand, the half-age formula (dose = patient age multiplied by 2.5) is more fitting with bilateral placement and results in doses approximately 250 percent above the seizure threshold. The age-based methods, though, do not account for intraindividual variability, which could equate to the provision of markedly elevated stimulation with significant cognitive adverse effects.

The empirical titration procedure has been shown to sustain efficacy while limiting adverse effects through containment of unnecessary energy. This procedure involves the delivery of an initial stimulus dose, followed by restimulation with increased intensity until a 20- to 25-second interval of motoric or EEG manifestation of a seizure is observed. As found in clinical and research practice, seizure threshold increases with age; thus, elderly patients require higher doses to achieve grand malseizures. Even titration procedures are altered for elderly patients. Whereas with younger patients, those below the age of 50, the first stimulus may be administered at 5 percent, for elderly the starting stimulus is administered at 10 percent. Stimulus dosing is also influenced by the choice of electrode placement. For subsequent treatments after the ECT titration session, bilateral and bifrontal electrode placement dosing is administered at moderately suprathreshold levels, defined as 150 percent above the seizure threshold. On the other hand, right unilateral ECT placement dosing is administered

at moderately to markedly suprathreshold stimulation, which is 250 to 600 percent above the seizure threshold (2.5 to 6 times threshold). These dosing strategies have been extensively researched and tested in order to ascertain the therapeutic effects.

Specific ECT Recommendations

During the administration of ECT, the induced seizure should be monitored via EEG and motoric response recordings. The EEG monitoring shows characteristic patterns of the grand malseizure as well as the postictal suppression. The sites for recording the EEG are the frontal-mastoid and frontal-frontal montage, which will maximize the EEG seizure expression. During the tonic phase, hypersynchronous polyspike complexes are followed by polyspike and slow-wave complexes, and during the clonus phase are followed by sudden or gradual suppression. Most ECT-induced seizures last less than 90 seconds and consequently are deemed inadequate if they do not show bilateral generalization on EEG manifestations. For monitoring motoric expressions, the cuff technique can be used. This procedure involves applying a blood pressure cuff to a limb (i.e., arm or leg) before the administration of a muscle relaxant. Commonly, the cuff is applied just above the ankle of the left leg in order to observe ECT motor manifestations.

Continuous monitoring of hemodynamic functioning continues throughout the ECT treatment session. This is essential for anesthesia and the rapid shifts in parasympathetic and sympathetic outflow that occur after the electrical stimulus. As the patient receives the electrical stimulus, increased parasympathetic tone can produce asystole or significant bradycardia. These hemodynamic effects can be aggravated by subconvulsive stimuli or medications. The ECT-associated bradycardia can be mitigated by pretreatment with anticholinergics such as atropine or glycopyrrolate. During the seizure, heart rate and blood pressure generally elevate and are followed (sometimes after many minutes) by a period of bradycardia. As a result, geriatric patients with any cardiac condition often require additional monitoring, management, and medications to dampen these effects.

Patients with Comorbid Medical Diagnoses. Many elderly patients have pre-existing cardiovascular diseases such as hypertension, angina, arrhythmias, aneurysms, conduction abnormalities, and previous myocardial infractions. Consultation with a cardiologist prior to ECT for optimum management is warranted in these situations. Possible concerns involve the increased vagal (parasympathetic) tone during and shortly after each electrical stimulus, which can evoke brief bradycardia or asystole. The onset of the therapeutic seizure leads to elevations in heart rate and blood pressure. Postictally, there is a second period of heightened vagal tone

evidenced by bradycardia and ectopic beats in some patients. With careful management, cardiovascular risk can safely be mitigated in these situations. Anticholinergics such as glycopyrrolate and atropine given prior to ECT can subvert cardiac arrhythmias. Short-acting β-blockers such as esmolol and labetalol are often helpful in controlling hypertension and tachycardia during treatments. Other cardiac agents (such as nifedipine, hydralazine, trimethaphan, and nitroglycerin) can be used if indicated.

Intracranial mass lesions or increased intracranial pressure presents a serious risk factor for ECT. Optimally, subdural hematomas will be removed prior to treatment. Elderly patients with depression and who present with previous cerebral infarctions can benefit from ECT. However, the amount of time that should pass after a stroke to maximize safety is unknown. Otherwise, the incidence of cerebrovascular complications is minimal.

Osteoporosis, spinal disc disease, and spondylolysis may necessitate a higher level of muscle relaxation during ECT. Neurosurgical consultation for ideal management prior to treatment is advisable in these situations. Pulmonary status must also be maximized prior to treatment. Unique ventilation strategies may be employed for patients with lung disease.

ECT leads to transient increases in intraocular pressures. Patients with glaucoma should receive eye drops prior to treatment. Collaboration with ophthalmologists for patients with retinal detachments and acute close-angle glaucoma is essential. Patients with loose, uneven, or decayed teeth are at risk for dental injury with ECT. Dental consultation and unique block strategies may need to be employed in this setting.

ACUTE CLINICAL SAFETY AND EFFICACY

ECT is regarded as one of the safest and most effective somatic treatments for severe major affective disorders, particularly those for which alternate treatments have been ineffective. Its safety compares favorably to other procedures that require general anesthesia. The American Psychiatric Association Task Force on ECT, National Network of Depression Centers Task Force on ECT, and the National Institute for Health and Clinical Excellence in the United Kingdom suggest that ECT is an efficacious acute treatment. In many systematic controlled trials, ECT has consistently been found to be superior to sham ECT (i.e., patients received only anesthesia) and pharmacotherapy.

The remission rates following an acute course of ECT have been found to vary between 20 and 80 percent with an effect size as large as 0.9. This large range is dependent on an interaction of both ECT treatment parameters and patient clinical factors. For instance, in order for ECT to be effective, it must be appropriately applied with the respective therapeutic dosage, pulse waveform, and electrode placement. A controlled trial that compared fixed high-dose right unilateral and bitemporal ECT in elderly

patients with major depression showed both electrode placements to be similar in terms of clinical outcome with comparable remission rates. In the Consortium for Research in ECT (CORE) study, which investigated the safety and efficacy of continuation ECT, approximately 80 percent of the total sample achieved remission during the acute phase with bitemporal, dose titrated, brief pulse waveform ECT. Patients were treated three times a week and standardized depressive rating inventories were collected within 24 to 48 hours after each treatment to systematically monitor mood change. Thus, a major conclusion of this study was that ECT can result in substantial acute efficacy when treatment is judiciously applied following a standardized therapeutic protocol with continual monitoring of patient symptom severity.

The significant benefit of ECT rests not only with its therapeutic benefits, but also with the speed at which response and remission are achieved. Analyses from the CORE trial of 253 patients found that approximately 54 and 83 percent of the sample achieved response (decrease in depression severity as measured by a standardized rating scale by a minimum of 50 percent) by the third and sixth ECT treatments, respectively. In terms of remission (level of depression severity found in the normal population), 65 percent of the sample met criteria for remission by the tenth ECT treatment. Elucidating the speed with which ECT works, between the fourth and ninth ECT treatment, remission rates increased from 10 to 60 percent. Given that pharmacotherapy may take many weeks to achieve steady state and produce results, and psychotherapy likewise may take time for results to show, ECT is a prime therapeutic choice for those patients who are in critical condition due to acute and severe psychiatric illness.

Early retrospective studies that utilized chart reviews found conflicting results, with some indicating ECT to be ineffective for elderly populations and others finding it to be both safe and efficacious. With technological improvements and advances in ECT practice, recent controlled clinical trials have found ECT to be highly effective in elderly patients. Comparisons with tricyclic antidepressants, monoamine oxidase inhibitors, and even selective serotonin reuptake inhibitors found ECT to be superior. In fact, age has been found to be an influencing factor for the therapeutic effects of ECT. A report from the CORE trial showed that age was positively associated with treatment response and subsequent remission. In that report, 90 percent of the elderly cohort relative to 64 percent of the younger cohort responded with ECT. Further examination of the data discovered that the greater response was only valid for those patients with psychotic features and suggested that the presence of psychosis and older age interacted and resulted in higher response rates. That is, elderly patients without psychotic features and younger patients showed comparable response rates, but elderly patients with psychotic features

showed higher rates of response relative to younger patients.

Electroconvulsive therapy is a highly effective treatment for the majority of elderly patients, even those who are 75 years and older. A prospective study of 268 patients examined three categories of elderly patients: those younger than 60 years of age, those between 60 and 74 years, and those aged 75 or older. Although the younger cohort had relative lower seizure thresholds, no differences were found between groups concerning the number of treatments or electrode placement. The middle cohort was found to have higher response rates compared to the younger cohort, but no efficacy differences were found between the middle and older groups. This study found that ECT is beneficial for the oldest-old and confirmed earlier clinical reports that older age is positively associated with higher seizure threshold levels.

The effectiveness of ECT is paralleled by its timeliness in onset of response, which may result in shorter and less costly hospital stays. A small controlled trial of 13 patients over the age of 65 found a 21 percent reduction in depression severity (as rated on a standardized scale) after the first ECT treatment, and a 36 percent reduction after the second. The depressive symptom impacted most after the first treatment with a decrease of 36 percent was "reported sadness." On average, four treatments resulted in a decrease of depression severity by 50 percent, though the investigators cautioned that response varied per individual case and that further research was required to assess the time to response and remission in elderly patients. To a degree, these data are synonymous with the findings of the CORE trial, which found ECT to have a quick onset of action.

Electroconvulsive therapy may be judiciously applied to a spectrum of geriatric patients with comorbid illnesses including cerebrovascular disease or dementia; however, there is a dearth of information from systematic randomized controlled trials. Retrospectively, chart reviews and clinical reports support ECT to be efficacious in the treatment of severe mood disorders in elderly patients with concomitant dementia, neurocognitive impairments, or stroke. The presence of cerebrovascular abnormalities did not minimize the therapeutic response; nonetheless, when applying ECT to geriatric patients in the presence of concomitant cerebral pathologies and dementias, prudent observation and monitoring of symptomatology and adverse neurocognitive effects become paramount.

Optimal benefit from ECT in elderly patients is dependent upon eliciting therapeutic seizures. Both clinical and research practice have shown that elderly patients have higher seizure thresholds and thus require higher doses of electrical stimulation. One controlled trial reported that in 20 patients, two required stimulus doses that exceeded the levels available on commercial ECT devices in the United States. The FDA has limited the ECT devices to 576 millicoulombs; as such, a small percentage of elderly

patients with remarkably high seizure thresholds may not benefit due to the inability to elicit seizure activity.

CONTINUATION AND MAINTENANCE ECT SAFETY AND EFFICACY

A cogent constraint of ECT is the high relapse rate after completion of the acute treatment course with proneness for relapse being greater during the first week and month after remission. Consequently, though elderly patients may show better response and remission with ECT relative to younger patients, they may also have higher rates of relapse. One study that followed 51 elderly patients for 1 year during and after treatment with either ECT or pharmacotherapy treatment for major depression found that 41 percent of the total sample was rehospitalized at least once within the first year, and 46 percent (6 of 13) of the ECT group was rehospitalized during the first month after acute treatment.

Many factors may contribute to the high relapse found following a successful acute course of ECT including residual depressive symptoms, concomitant physical illness, and psychosocial factors. Residual depressive symptoms are those symptoms that did not abate with antidepressant treatment that could substantially increase the risk of reverting back to baseline status, or could be a precursor to the development of a new depressive episode. Many elderly patients who suffer from comorbid physical illness may be burdened with additional stressors that complicate continued recovery from the acute phase. Poor physical health, specifically cardiovascular disease and diabetes, has been linked to both affective and anxiety disorders. Thus, elderly patients with comorbid physical medical illnesses may be more at risk for re-experiencing a return of psychiatric pathology after a period of remission. Finally, elderly patients with increased psychosocial stressors such as lack of autonomy, living alone, or decreased activities and instrumental activities of daily living may be more susceptible to relapse occurrence. Nonetheless, the goal of continuation and maintenance therapy is to sustain the period of remission while simultaneously limiting the associated adverse effects.

The question that arises is what is the best method to achieve the above goal, a question that has received much attention. The CORE trial was designed to determine whether continuation pharmacotherapy (lithium plus nortriptyline) or ECT (bilateral electrode placement 150 percent above the seizure threshold) resulted in a better outcome, and the results indicated that both were equivalent in prolonging remission. Of the patients randomized to the continuation ECT condition, 37.1 percent relapsed and 46.1 percent retained remission. For those in the continuation pharmacotherapy group, 31.6 percent relapsed and 46.3 percent retained remission. The time to relapse was similar in both groups with patients in the ECT arm relapsing after an average of 9 weeks and those in the pharmacotherapy arm after an average of 6 weeks. Although both of these

therapeutic modalities were deemed better than no treatment after attaining remission in the acute ECT phase, they were moderately effective in retaining the antidepressant benefits produced during the acute ECT phase.

A probable benefit of maintenance therapy with ECT is the lack of drug–drug interaction. Elderly patients usually receive multiple medications for management of physical health-related issues. The resulting polypharmacy from mixing a multitude of medications such as antidepressants, mood stabilizers, analgesics, and cardiac medications could result in iatrogenic effects. Moreover, elderly patients may be sensitive to pharmacological properties of medications and proper administration must be monitored to ensure attainment of the full therapeutic effects while limiting harmful reactions. In such cases where elderly patients are extremely sensitive to the effects of medications, or they have a coexisting medical condition that limits the use of psychiatric medications, ECT can be a viable alternative. However, given the cardiovascular adverse effects of ECT in elderly patients, especially those with pre-existing hypertension or unstable cardiac activity, continued ECT use needs to be combined with careful and comprehensive monitoring of cardiovascular functioning. For instance, one study of patients over the age of 75 found the development of transient hypertension and sinus tachycardia during and after the administration of ECT. Although the transient phenomena were considered "normal," they still merit monitoring and evaluation.

ECT ADVERSE EFFECTS

ECT is one of the safest and well-tolerated somatic treatments with a documented mortality rate of 0.002 percent. The majority of side effects are limited to the time during and immediately after an ECT treatment, with the majority of side effects persisting for a short period of time after a complete acute course, and others lasting for a longer period of time (i.e., 6 months and longer) (Table 57.4f–5). Three broad side-effect categories for ECT are systemic, neurophysiological, and neurocognitive.

Table 57.4f–5.
Side Effects Related to Electroconvulsive Therapy

- Headache
- Hypertension
- Cardiac arrhythmia (bradycardia, tachycardia, asystole)
- Nausea
- Tongue bite
- Temporomandibular joint pain
- Disorientation
- Delirium
- Anterograde amnesia

- Retrograde amnesia
- Prolonged seizure
- Precipitation of mania
- Drug–drug interactions
- Anesthetic complications
- Aspiration
- Breathing complications
- Myalgias
- Dental complications
- Postseizure myocardial infarction
- Postseizure stroke
- Tardive seizure
- Scalp burn (very rare)
- Bone fractures (very rare)
- Spinal disc prolapse (very rare)
- Compression fracture (very rare)

Systemic Effects

Mortality associated with ECT is most commonly attributed to cardiovascular or respiratory complications. Death can result from myocardial infractions or ruptured arrhythmias, ruptured aneurysms, and brain herniations. Pulmonary aspiration is rare, but particularly important to prevent in patients with chronic lung disease. Short-lived increases in blood pressure, pulse, cerebral blood flow, and intracranial pressure present risks for elderly patients, but in the majority of cases are well tolerated. Ideally, the modern practice of using muscle relaxants prevents fractures that often occurred in the early days of ECT practice. Patients can still suffer from dental complications (i.e., fractured teeth, temporomandibular joint pain) or complain of myalgias, but these risks are extremely rare with the advent of anesthesia and placement of bite blocks. Muscle pain can be related to the use of depolarizing agents such as succinylcholine, and headaches can be successfully treated with over-the-counter analgesics and often improve with time. Pretreatment for musculoskeletal pain can also be beneficial in certain cases. Some elderly patients complain of nausea and vomiting following ECT sessions. Antiemetics such as ondansetron, metoclopramide, and prochlorperazine can be administered to treat these symptoms or prophylactically prior to ECT.

Neurophysiological Effects

Research involving humans and animals indicate that neither ECT nor brief seizures impart structural brain damage. This is further substantiated by carefully controlled imaging studies. However, seizures lasting more than 180 seconds should be aborted with an anticonvulsant anesthetic agent or a benzodiazepine. Lengthy seizures and status epilepticus could theoretically lead to neuronal injury if the patient becomes hypoxic. This is unlikely,

though, as ECT patients are well oxygenated and most prolonged seizures are easily suppressed. Spontaneous seizures after treatment are rare, but in their event require careful evaluation and management. Postictal delirium can involve disorientation, agitation, and confusion and is best managed with supportive interventions, minimization of environmental stimuli, and judicious use of medications such as benzodiazepines.

Neurocognitive Effects

The neurocognitive effects of ECT are likely the most limiting factor for the provision of this therapeutically beneficial treatment. Numerous studies have systematically examined the neurocognitive sequelae associated with ECT, taking into account various ECT parameters and neurocognitive domains, as well as the effects of the underlying psychiatric disease. Although research and clinical findings have varied per patient sample and method of ECT administration, the overarching conclusion is that ECT does result in transient negative cognitive effects, particularly in the memory domain.

The ECT parameters studied have included sites of electrode placement, stimulus intensity, and pulse waveform. There are three prominent electrode placement sites used clinically: bitemporal, right unilateral, and bifrontal. Anecdotally, bitemporal is thought to produce the fastest speed of response, but also carries the burden of having the most neurocognitive adverse impact. Right unilateral, on the other hand, is thought to have similar properties in terms of efficacy, yet a more benign neurocognitive profile, and bifrontal too has been suggested to have high efficacy with benign neurocognitive effects. Numerous clinical trials studying the impact of bitemporal and right unilateral ECT have consistently found right unilateral to be more cognitively sparing relative to bitemporal, notably with regard to memory functions.

Examinations of bifrontal ECT have been limited due to small sample sizes and inconsistent procedures in terms of dosing strategies. However, the CORE group in its large-scale, controlled trial of three electrode placements compared the cognitive effects of bitemporal, right unilateral, and bifrontal placements administered with brief pulse waveform during an acute ECT course. The study found that all three electrode placements were equivalent both in efficacy and neurocognitive adverse effects. However, that study was limited by missing data and patients who exited the study early. A study conducted in Belgium compared bifrontal and right unilateral electrode placements administered with ultra-brief pulse-width during the acute course. That study found that both bifrontal and right unilateral were equivalent in both efficacy and neurocognitive outcome.

Studying electrode placement must take into account the interactions of stimulus dosing and pulse waveform. Research has found a consistent direct correlation between stimulus dosage and neurocognitive

impairment. With the introduction of stimulus dose titration in the 1980s, dosage of ECT is limited to the amount needed to provide a therapeutic seizure. This has helped to minimize the negative neurocognitive side effects that were found in earlier use of ECT when higher doses were applied. An empirical example of the effects of stimulus dosage was found in a controlled, double-masked, clinical trial that compared bitemporal and right unilateral ECT. The treatment conditions were bitemporal ECT at 150 percent above the seizure threshold and right unilateral ECT at 50, 150, or 500 percent above the seizure threshold. Efficacy and neurocognitive assessments were performed before the treatment, after the sixth and last treatments, and 1 week after the last treatment. The investigators found that the neurocognitive effects of right unilateral ECT were dependent on the stimulus dosage utilized to provide treatment. When administered at higher doses, statistically and clinically significant differences were found between the right unilateral groups, with the higher dose (i.e., 500 percent above seizure threshold) resulting in greater time to reorientate and more anterograde memory impairment (i.e., memory for new information). However, even at the highest right unilateral dose, bitemporal electrode placement resulted in more cognitive impairment in the domain of retrograde amnesia.

Another controlled investigation examined the acute neurocognitive effects of only right unilateral ECT delivered at either a fixed high dose (403 millicoulombs) or moderate titrated dose (2.25 times the initial seizure threshold). The findings mimicked those of other studies and showed that right unilateral ECT at a fixed high dose produced greater neurocognitive effects than the moderate titrated dose in the domains of time to orientation, global cognitive function, and retrograde memory. Thus, when deciphering the neurocognitive effects, electrode placement and stimulus dose must both be examined in order to sufficiently understand the interaction. With the recent advent in studying the interaction effects of pulse waveform, there has been inconsistent information. For example, randomized, controlled clinical trial found that ultra-brief, relative to brief pulse, was associated with less neurocognitive effects, regardless of electrode placement. However, another randomized controlled clinical trial that compared brief and ultra-brief pulse waveform in patients treated only with right-unilateral electrode placement, empirical dose titrated ECT found that both pulse waveforms produced similar clinical and neurocognitive outcomes. That is, the study suggested that ultra-brief pulse waveform has no neurocognitive advantages relative to brief pulse waveform. Further research is warranted to clarify both the clinical and neurocognitive effects of ultra-brief pulse waveform.

Time to reorientation (i.e., the time needed after the end of each ECT session to regain orientation) is greatly impacted immediately after each treatment with resolution following in time after treatment. Like other

cognitive domains, orientation time is affected by electrode placement, stimulus dosage, and pulse waveform with bitemporal electrode placement, higher stimulus dosages, and sine waveform resulting in lengthier times to reorientation. In a study that compared bitemporal and right unilateral ECT, both electrode placements showed relatively similar patterns of orientation recovery over the course of treatment, indicating that subsequent ECT treatments do not incrementally prolong orientation recovery time. For instance, in that study the first and last two treatments were compared and no within-group differences were found regarding reorientation time, even though between-group differences were observed. Of interest, with right unilateral ECT, age was positively related to recovery time with older patients regaining orientation at a faster rate than younger patients. The positive association between age and reorientation time was mitigated by the lengthy reorientation time in the bilateral group. Research of the pulse waveforms has consistently found that sine waveform results in the longest time to regain orientation, followed by brief pulse, and finally ultra-brief pulse.

When assessing neurocognitive effects, acute and long-term effects must be examined. Acutely, ECT results in modest impairment of many domains including visuospatial and verbal learning and memory, and global cognitive function. However, those effects typically resolve quickly within 1 week to 1 month after the end of the acute ECT course. A study of 45 depressed patients who received ECT (administered with right unilateral or bitemporal electrode placement) compared those above and below the age of 60 on a battery of neurocognitive tests at set time points. The study found that immediately after, and at long-term follow-up, patients showed improvement on neurocognitive measures of attention, verbal fluency, and memory for visual shapes. The investigators concluded that the improvement in neurocognitive function may have been related to decreased depression severity or the ECT-associated activation of dopaminergic and serotonergic neurotransmitters.

Consequently, in terms of short- and long-term neurocognitive effects, the domain of neurocognitive function greatly impacted is retrograde memory. Initially, retrograde memory was thought to resolve over time after the end of the acute phase, with substantial recovery of personal memories. Earlier research found that retrograde amnesia could be predicted by the global level of cognitive functioning before the initiation of ECT.

A recent study of community ECT practices in the New York area, though, showed evidence that retrograde amnesia is a persistent effect of ECT and that some autobiographical information may remain lost. Seven New York hospitals (three community hospitals, two private hospitals, and two university hospitals) were included in the study and a total of 347 patients who were longitudinally followed for 6 months provided outcome

data. Focusing on the neurocognitive outcome, the seven treatment facilities differed due to ECT practice differences at the respective sites. For instance, one center administered ECT with a sine waveform that was associated with more effects on psychomotor speed and memory for autobiographical information. The significance of the site differences indicated that there is no one uniform practice of ECT administration and that outdated methods that result in greater neurocognitive effects are still employed.

More importantly, the study found that retrograde memory deficits persisted for at least 6 months and were related to ECT technique, female gender, lower premorbid intelligence, and advancing age. Of equal importance, the number of ECT treatments with bitemporal electrode placement was associated with the amount of retrograde amnesia, even at the 6-month time point. However, that study was recently critiqued as it used an autobiographical memory interview that has no reported psychometric properties. Thus, those findings have now been called into question, and further research is needed with psychometrically sound measures of autobiographical information to confirm or refute those findings. Nonetheless, elderly patients may be more prone to the ECT-associated adverse cognitive effects and treatment with right unilateral electrode placement, empirical dose titration, and ultra-brief pulse waveform may help to mitigate the cognitive effects.

Autobiographical memory, the area mainly affected by ECT, can be categorized into two types: One that focuses on personal memories and one that focuses on impersonal memories (i.e., public events). A clinical trial of patients treated with either right unilateral or bitemporal ECT, who were compared to a healthy control group, found that memory for impersonal events was more affected than memory for self, and greater deficits were found in those patients treated with bitemporal ECT. The pattern of memory loss showed that more recent events, those closest in time to the delivery of ECT, were the most vulnerable to the ECT effects. Also, those events that were patient determined to have low salience were most vulnerable to the effects of ECT.

A neurocognitive domain that is often overlooked in both clinical and research arenas is the patient's view of his or her cognitive experience, which is commonly referred to as subjective cognitive functioning. Many elderly patients with affective disorders will complain of memory difficulties, short attention spans, and slow psychomotor speed. Although the neurocognitive functions may be intact based on formal neurocognitive testing, the patient's subjective experience needs to be considered, especially with regard to ECT. Due to the negative stigma associated with ECT, which some elderly patients refer to as the "shock," their subjective view point is a factor in the efficacy of the treatment. Prior research found that patient self-rating of neurocognitive function was related to a

depression severity self-rating scale, with higher scores on the depression scale correlating to greater subjective cognitive impairment scores. During and following an acute ECT course, patients showed significant improvement in subjective cognitive function, indicating that their memory had improved regardless of electrode placement type (bitemporal or right unilateral). The change in self-evaluation of neurocognitive functions was correlated with improvement in mood, and was also related to the degree of retrograde amnesia.

OTHER NEUROSTIMULATION THERAPIES

Although ECT is a highly effective treatment for major affective disorders and other severe treatment-resistant psychiatric conditions, its use is limited due to high relapse rates after the acute phase and associated adverse cognitive effects. In order to remedy the negative effects of ECT and to increase the delivery of efficacious treatments, other neurostimulation modalities are being developed as antidepressant therapies. These neurostimulation modalities include transcranial direct current stimulation, vagus nerve stimulation, rapid transcranial magnetic stimulation, magnetic seizure therapy, and deep brain stimulation.

Transcranial Direct Current Stimulation

Transcranial direct current stimulation (tDCS) applies a weak direct current (1 to 2 mA) with scalp electrodes. Studies suggest that anodal stimulation increases cortical excitability while cathodal stimulation increases cortical inhibition. It is thought that tDCS has variable effects on long term potentiation (LTP) or long term depression (LTD) depending on how it is delivered. As a result, this tool has been applied for a wide array of studies investigating neuroplasticity and cognitive neuroscience. Generally, tDCS is well tolerated and has the potential to alter cortical activity beyond they stimulation session. Therapeutic research has focused on depression, stroke rehabilitation, cognitive remediation, and more recently studies of dementia. Contemporary work suggests that clinical outcomes with tDCS may be enhanced by coupling it with cognitive remediation, antidepressants, or other concurrent treatments.

Vagus Nerve Stimulation

Vagus nerve stimulation (VNS) was approved by the FDA in 2005 as an adjunctive antidepressant treatment for patients with treatment-resistant depression. VNS involves implanting a device similar to a pacemaker that produces brief, intermittent electrical stimulation to the vagus nerve. This neurostimulation modality was investigated in two large-scale, randomized, controlled clinical trials. The first, an open-label trial, found VNS to be safe and tolerable, producing limited side effects grouped into

two categories, surgery related and stimulation related. These side effects were transient and resolved over time; even those related to stimulation lessened over time and many patients reported substantial adaptation to side effects such as increased cough and hoarseness during active stimulation.

In terms of efficacy, the second trial, double masked and sham controlled, found no difference between active and sham VNS during the acute phase. However, during the long-term open-label phase in which all patients received active treatment, depression severity decreased as measured by both clinician and patient reports, and modest response (29.8 percent) and remission (17.1 percent) rates were found. Although VNS has shown some benefit, it remains unclear how to adequately provide dosing. In the initial study, the majority of the patients received high-dose VNS, whereas in the second study, VNS doses were lower.

Of significance, there were no negative cognitive side effects found and, conversely, patients showed minimal improvement in the cognitive domains of psychomotor speed and executive functions. However, this improvement was hypothesized to be related to the improvement in depression as opposed to directly related to VNS. Nonetheless, VNS could be an option for geriatric patients with treatment-resistant depression who have not benefited from other available antidepressant therapies and who for any medical reasons are unable to undergo ECT.

Rapid Transcranial Magnetic Stimulation

Rapid (or repetitive) transcranial magnetic stimulation (rTMS) is a noninvasive neurostimulation technique that applies alternating magnetic currents, which then produce electrical current in cortical areas. rTMS has been under significant investigation as a treatment for major affective disorders and other neuropsychiatric illnesses. A large randomized, controlled clinical trial comparing active and sham rTMS to treat unipolar major depressive disorder during the acute phase (6 weeks) found no statistically significant difference between groups; however, active rTMS did produce clinically beneficial response and remission rates. The patients in the trial were actively taking concomitant pharmacotherapy and were treated for 6 weeks. Those who did not respond to the acute phase were offered open-label active treatment for an additional 6 weeks, and those who responded were naturalistically followed to monitor the long-term rTMS effects. While this study did not specifically address safety and efficacy in a geriatric population, elderly patients were included in the study and those data are expected to be analyzed.

In a separate study, investigators assessed the use of open-label rTMS in 18 geriatric patients (mean age: 61.2, range: 55 to 75). These patients were diagnosed with moderate to severe unipolar major depressive disorder, and many of them were resistant to polypharmacy. All patients

tolerated rTMS treatment, with side effects limited to mild headaches and scalp discomfort. Overall, there was significant improvement in global functioning and symptoms of depression and anxiety. However, the study found that the presence of general medical comorbidities and acute onset of depression decreased the therapeutic response of rTMS.

The use of rTMS is attractive to many clinicians as it has been shown to be safe and well tolerated. The main side effects reported include temporary headaches and scalp pain at the site of stimulation. As rTMS has been investigated for more than 20 years, safety guidelines have been generated by expert scientists to guide clinical decision making regarding rTMS treatment parameters. A limitation of using rTMS in a geriatric sample is the presence of cortical atrophy in the prefrontal cortex (the usual site of stimulation). Cortical atrophy is inversely related to treatment outcome with rTMS and has been found to be a predictor of nonresponse. In 2008, the FDA-approved rTMS for treating adult patients with unipolar major depressive disorder for whom one antidepressant treatment has failed to work.

Magnetic Seizure Therapy

Magnetic seizure therapy (MST) is a neurostimulation therapy currently being studied that utilizes rTMS at high frequencies (i.e., 50 to 100 Hz) to induce therapeutic seizures. The goal of MST is to focally induce therapeutic seizures, containing the cortical spread of the seizure, in order to maximize efficacy while limiting adverse effects. Initial research in primate and human models found MST to be safe and well tolerated with evidence of limited negative effects. An initial trial conducted in ten humans who each received two MST treatments within an acute course of ECT found that MST, relative to the ECT treatments, resulted in faster time to achieve orientation and generation of more words on a test of verbal fluency. To date, there have been two randomized, controlled clinical trials of MST in the treatment of severe mood disorders. One trial has been completed and one trial is ongoing. In the first trial, 20 patients with unipolar or bipolar major depressive disorder were randomized to receive focal MST, with the magnetic stimulating coil placed over the frontal cortex, or nonfocal MST, with the coil placed over the vertex. The results found the treatment to be modestly efficacious with response and remission rates of 50 and 15 percent, respectively. Of benefit, MST showed limited cognitive effects in the domain of retrograde amnesia for autobiographical memories. Also, there were no adverse effects in the domains of global cognitive functioning, verbal and visual anterograde learning and memory, or time to orientation. Patients also reported subjective improvement in cognitive functioning.

The second randomized trial will provide further insight into the efficacy of MST by evaluating high-dose MST (delivered at 100 Hz).

Preliminary studies in nonhuman primates and humans found high-dose MST to be safe and tolerable and the current trial will comprehensively evaluate the efficacy and cognitive effects. Although not initially targeted for geriatric patients, MST could have a role in future algorithms given its limited adverse cognitive effects. Thus, geriatric patients who may be more susceptible to the negative cognitive effects of ECT may be able to receive MST and benefit from therapeutic seizures coupled with limited cognitive sequelae.

Deep Brain Stimulation

Deep brain stimulation involves the surgical implantation of a pacemaker-like device that sends continuous electrical stimulation to a specifically targeted region in the brain. Deep brain stimulation has been safely used with beneficial results in patients with severe movement disorders such as Parkinson disease. The procedure is considered invasive, but safe, as the surgery takes place while the patient is awake and the electrode leads are placed directly in cortical tissue.

There has been recent interest in using deep brain stimulation to treat severe affective disorders based on initial anecdotal observations that patients with movement disorders appeared brighter during the course of treatment. A small open-label pilot study of six patients (age range: 37 to 59 years) treated with deep brain stimulation found that four patients responded (i.e., at least 50 percent decrease in depression severity) and two showed no response. Neurocognitive testing found that deep brain stimulation had no adverse effects and in turn many patients showed improvement in visuomotor functioning and verbal fluency. Aside from the development of local infections from the postsurgical implantation of the connector leads, which resolved with antibiotic treatment, no other adverse events were reported. To remedy the risk of infection, the surgical procedures were modified to implant the stimulating generator and the electrodes in one surgery session. A recent 6-month, naturalistic study of 54 patients treated with deep brain stimulation for Parkinson disease demonstrated improvement in motor and mood symptoms. Further, baseline, psychiatric comorbidities did not have a differential effect on outcome in this sample. The use of deep brain stimulation in treating psychiatric illnesses warrants further investigation based on the preliminary optimistic response, positive neurocognitive effects, and limited adverse effects.

FUTURE DIRECTIONS

Algorithm development is a crucial future step to optimize the application of ECT and neuromodulatory interventions in the elderly. This work will aid clinicians in sequential treatments or complex regiments with

multimodal approaches. Further work to identify risk factors for neurocognitive side effects and minimizing this burden for patients will improve outcomes. Optimal continuation and maintenance treatment is also still poorly understood.

With the advent of newer somatic therapies such as tDCS, rTMS and MST, additional research is warranted to understand their effects in geriatric populations. Research thus far has found these neurostimulation therapies to be cognitively benign, and possibly cognitively enhancing with promising clinical efficacy in relatively younger cohorts, and this will have significant therapeutic implications when treating elderly patients, especially those with pre-existing cognitive deficits. Magnetic seizure therapy in particular, though presently in developmental stages, could be a new frontier in convulsive therapies and psychiatric practice through its proposed mechanism of action that facilitates therapeutic seizure induction coupled with little neurocognitive effects.

With newer neurostimulation and convulsive therapies added to the psychiatric therapeutic armamentarium, future directions will need to begin to address the issue of developing high-caliber treatment algorithms tailored to the individual geriatric patient. This could involve utilizing different neurostimulation modalities at different treatment courses. For example, patients who benefited from acute ECT may receive continuation TMS to prolong the period of remission. Also, the combination of neurostimulation modalities with pharmacotherapy requires the understanding of how to refine that administration in order to enhance the clinical efficacy, particularly in elderly patients who are prescribed multiple medications. The goal of these treatment algorithms should focus on defining the optimum delivery of somatic therapies that maximize therapeutic benefits and minimize unnecessary medications and adverse reactions. In addition, these evidence-based algorithms for geriatric patients should underscore the importance of neurocognitive assessments and effects when determining clinical management with somatic and multimodal therapies. These scientific advances in geriatric psychiatry will contribute to the continued evolution of novel and advanced somatic therapies.

References

Bjolseth TM, Engedal K, Benth JS, Dybedal GS, Gaarden TL, Tanum L. Clinical efficacy of formula-based bifrontal versus right unilateral electroconvulsive therapy (ECT) in the treatment of major depression among elderly patients: a pragmatic, randomized, assessor-blinded, controlled trial. *J Affect Disord*. 2015;175:8–17.

Blumberger DM, Hsu JH, Daskalakis ZJ. A review of brain stimulation treatments for late-life depression. *Curr Treat Options Psychiatry*. 2015;2:413–421.

Bosboom PR, Deijen JB. Age-related cognitive effects of ECT and ECT-induced mood improvement in depressive patients. *Depress Anxiety*. 2006;23:93–101.

Brunoni AR, Valiengo L, Baccaro A, et al. The sertraline vs. electrical current therapy for treating depression clinical study: results from a factorial, randomized, controlled trial. *JAMA Psychiatry*. 2013;70(4):383–391.

Chopra A, Abulseoud OA, Sampson S, et al. Mood stability in Parkinson disease following deep brain stimulation: a 6-month prospective follow-up study. *Psychosomatics*. 2014;55(5):478–484.

Feliu M, Edwards CL, Sudhakar S, McDougald C, Raynor R. Neuropsychological effects and attitudes in patients following electroconvulsive therapy. *Neuropsychiatric Dis Treat*. 2008;4:613–617.

*Gardner BK, O'Connor DW. A review of the cognitive effects of electroconvulsive therapy in older adults. *J ECT*. 2008;24:68–80.

*Husain MM, Rush AJ, Fink M, Knapp R, Petrides G. Speed of response and remission in major depressive disorder with acute electroconvulsive therapy (ECT): a Consortium for Research in ECT (CORE) report. *J Clin Psychiatry*. 2004;65:485–491.

Kellner CH, Husain MM, Knapp RG, et al. Right unilateral ultrabrief pulse ECT in geriatric depression: phase 1 of the PRIDE study. *Am J Psychiatry*. 2016;173:1101–1109.

Kellner CH, Husain MM, Knapp RG, et al. A novel strategy for continuation ECT in geriatric depression: phase 2 of the PRIDE study. *Am J Psychiatry*. 2016;173:1110–1118.

Kellner CH, Knapp R, Husain MM, et al. Bifrontal, bitemporal and right unilateral electrode placement in ECT: randomized trial. *Br J Psychiatry*. 2010;196(3):226–234.

Kimball JN, Rosenquist PB, Dunn A, McCall V. Prediction of antidepressant response in both 2.25 x threshold RUL and fixed high dose RUL ECT. *J Affect Disord*. 2009;112:85–91.

Lee WH, Lisanby SH, Laine AF, Peterchev AV. Minimum electric field exposure for seizure induction with electroconvulsive therapy and magnetic seizure therapy [published online ahead of print December 9, 2016]. *Neuropsychopharmacology*.

*Lisanby SH. Electroconvulsive therapy for depression. *N Engl J Med*. 2007;357:1939–1945.

Lisanby SH, Husain MM, Rosenquist PB, Maxiner D, Gutierrez R. Daily left prefrontal repetitive transcranial magnetic stimulation in the acute treatment of major depression: Clinical predictors of outcome in a multisite, randomized controlled clinical trial. *Neuropsychopharmacology*. 2009;34:522–534.

Lozano AM, Mayberg HS, Giacobbe P, Hamani C, Craddock RC, Kennedy SH. Subcallosal cingulate gyrus deep brain stimulation for treatment-resistant depression. *Biol Psychiatry*. 2008;64:461–467.

Mantovani A, Lisanby SH. Transcranial magnetic stimulation in major depression. In: George MS, Belmaker RH, eds. *Transcranial Magnetic Stimulation in Clinical Psychiatry*. Washington, DC: American Psychiatric Publishing, Inc.; 2007.

McClintock SM, Choi J, Deng ZD, Appelbaum LG, Krystal AD, Lisanby SH. Multifactorial determinants of the neurocognitive effects of electroconvulsive therapy. *J ECT*. 2014;30(2):165–176.

McClintock SM, Trevino K, Husain MM. Vagus nerve stimulation: Indications, efficacy and methods. In: Swartz C, ed. *Electroconvulsive and Neuromodulation Therapies*. New York: Cambridge University Press; 2009.

Navarro V, Gasto C, Torres X, Masana G, Penades R. Continuation/maintenance treatment with

nortriptyline versus combined nortriptyline and ECT in late-life psychotic depression: a two-year randomized study. *Am J Geriatric Psychiatry*. 2008;16:498–505.

Nordenskjold A, von Knorring L, Engstrom I. Predictors of the short-term responder rate of electroconvulsive therapy in depressive disorders—a population based study. *BMC Psychiatry*. 2012;12:115.

Oudega ML, van Exel E, Stek ML, et al. The structure of geriatric depressed brain and response to electroconvulsive therapy. *Psychiatry Res: Neuroimaging*. 2014;222:1–9.

Rhebergen D, Huisman A, Bouckaert F, et al. Older age is associated with rapid remission of depression after electroconvulsive therapy: a latent class growth analysis. *Am J Geriatr Psychiatry*. 2015;23:274–282.

Segrave RA, Arnold S, Hoy K, Fitzgerald PB. Concurrent cognitive control training augments the antidepressant efficacy of tDCS: a pilot study. *Brain Stimul*. 2014;7:325–331.

Semkovska M, McLoughlin DM. Objective cognitive performance associated with electroconvulsive therapy for depression: a systematic review and meta-analysis. *Biol Psychiatry*. 2010;68:568–577.

Sienaert P, Vansteelandt K, Demyttenaere K, Peuskens J. Randomized comparison of ultra-brief bifrontal and unilateral electroconvulsive therapy for major depression: cognitive side-effects. *J Affect Disord*. 2010;122:60–67.

Spaans HP, Sienaert P, Bouckaert F, et al. Speed of remission in elderly patients with depression: electroconvulsive therapy v. medication. *Br J Psychiatry*. 2015;206:67–71.

Spaans HP, Verwijk E, Comijs HC, et al. Efficacy and cognitive side effects after brief pulse and ultrabrief pulse right unilateral electroconvulsive therapy for major depression: a randomized, double-blind, controlled study. *J Clin Psychiatry*. 2013;74:1029–1036.

Sutton TA, Sohrabi HR, Rainey-Simth SR, Bird SM, Weinborn M, Martins RN. The role of APOE-ε4 and beta amyloid in the differential rate of recovery from ECT: a review. *Transl Psychiatry*. 2015;5:e539.

Tielkes CE, Comijs HC, Verwijk E, Stek ML. The effects of ECT on cognitive functioning in the elderly: a review. *Int J Geriatr Psychiatry*. 2008;23:789–795.

Vaughn McCall W, Lisanby SH, Rosenquist PB, et al. Effects of a right unilateral ultrabrief pulse electroconvulsive therapy course on health related quality of life in elderly depressed patients. *J Affect Disord*. 2017;209:39–45.

Wilkins KM, Ostroff R, Tampi R. Efficacy of electroconvulsive therapy in the treatment of nondepressed psychiatric illness in elderly patients: A review of the literature. *J Geriatr Psychiatry Neurol*. 2008;21:3–11.

57.4g Psychosocial Factors in Psychotherapy of the Elderly

JOEL I. SADAVOY, M.D., AND VIRGINIA WESSON, M.D.

DEFINITION

The term *psychosocial,* as used here, refers to social and environmental elements in an individual's life that interact with and affect the individual's mental health and intrapsychic processes. Such elements may be

conceptualized at two levels: The intermediate level of factors arising within the immediate environs of the patient (e.g., friendship or family relations) and the more macro level of society and cultural norms (i.e., societal attitudes, culture, and ethnicity). Historical knowledge can be an important factor in creating an accurate clinical context for the patient's symptoms. While history obviously is not changeable, it can inform therapy by providing context for decision making. For example, for the oldest-old, key life decisions such as whether to marry or have children may have been heavily influenced by the experience of major wars, local conflicts, or worldwide economic depression. In turn, such decisions can strongly affect the degree to which an individual will age successfully. The aging demographic continues to evolve as elders live longer. In the last census in 2010 there were 53,000 centenarians in the United States with similar increases in number of individuals living to extreme old age in other industrialized countries. It is still too early to fully understand the psychosocial implications of living to extreme old age, such as the impact of outliving most of one's family and age cohort.

RELATIONSHIPS

The passage of time leads to major change in a patient's social network. In youth, time is often viewed as an ally, an opportunity for growth, and a time when new interpersonal roles and novel social partners are preferred. In old age, time is no longer an ally, eroding many life achievements and increasing the preference for emotionally close social relationships. Perhaps most profound are the changes associated with loss of relationships through death or other factors.

There is a convincing body of research in support of the concept that better social relations is related to better health and longer life in the elderly and its absence is a risk factor for psychiatric disorders, particularly depression, and stress-related physical illnesses and disability. What is less clear in the literature on social relations is which elements are most critical, for example, social support, social network, social integration, or social participation. The majority of data support the concept that elders who are part of what have been termed more socially endowed, diverse or resourceful networks have better mental health. Such networks are characterized by a greater number and variety of social ties, with whom the individual is intensely involved. It is also unclear whether all relationships confer health benefits. It may be that social relationships with friends and confidantes (the so-called discretionary or non-kin relationships) have a greater effect on mental health than relationships with kin, particularly children. There remains some debate as to whether it is the nature of the network itself or whether it is the individual's perception of and satisfaction with support that is the most important protective factor. Both seem to be important factors and should be taken into account when evaluating

individual patients. It seems fairly clear that beyond the nature of the network, the presence of negative interactions is a source of stress while perceived positive relations buffer stress. Hence, it is perhaps not surprising that strong social support helps to buffer depression, and married adults are at lower risk of depression than unmarried or widowed adults. There is an ongoing debate over the elimination of the bereavement exclusion criterion from a diagnosis of depression in the Diagnostic and Statistical Manual V. A recent study, for example, has found a high rate of significant depressive symptoms in bereaved elders who lost their spouse (although major depressive disorder was not associated with bereavement). It is estimated that one in six men and almost half of women over the age of 65 are widowed, and of these about a third will experience complicated bereavement, increasing their risk of depression and associated morbidity and mortality. Moreover, even uncomplicated bereavement can be associated with depressive symptoms which, while they may not reach threshold criteria for a diagnosis of major depression, nevertheless also may impair function and increases their risk of medical comorbidity and mortality. However, psychosocial resources including religious and spiritual belief, and practice, social and family support and Erikson's concept of vital involvement expressed in commitment to life tasks are moderating or protective factors which might be modifiable and enhanced by therapy.

Bereavement, while almost always difficult, may also be a release from an oppressive relationship, liberating the remaining spouse to resume old pursuits or take on new roles and responsibilities. For example, one spouse spoke eagerly about her ability to return to the academic research and teaching she had relinquished due to the demands of providing care to her ailing spouse.

During care of the elderly, as with younger adult patients, it is important to make no assumptions about the significance of loss but rather to explore the specific meaning of loss to each patient. It is important, too, to be aware that the bereaved spouse may need time and energy to develop new roles and relationships and they may struggle with how to identify themselves beyond being a spouse and partner. Other social changes may emerge that are supportive following bereavements. For example, widowed older women receive more help from friends than married women while widowed older men are the most likely to receive support adjusting to practical and emotional changes following the loss of their spouse.

Beyond bereavement, marital relationships in old age also are changed by physical or cognitive impairment, and shifting patterns of dominance and responsibility. For example, the basic nurturing role previously carried out by a loving wife may now fall into the unaccustomed hands of her husband in the face of her decline with Alzheimer disease. He experiences not only loss, but also the need to offer comfort and care, behavior and

feelings with which he has always been awkward. Guilt, resentment, and some anger sometimes emerge as a consequence. Conversely, the physical or mental decline of one spouse may elicit kindness and nurturing late in life that was absent in earlier years. For example, a haughty, narcissistic wife became "chastened" by her awareness of her own decline. While she agonized over her own losses, she also realized that the husband she previously demeaned and devalued was now her greatest ally. He welcomed the new role and the respect he began to feel from her, revealing a nurturing, loving side she had never seen before.

Sociocultural trends in relationships will impact the aging process. For example, in North America there has been a trend, especially among well-educated, high socioeconomic groups, to delay child-bearing in favor of focusing on career and financial advancement, only to find that fertility has been significantly reduced. This in turn often leads to inadvertent and unwanted childlessness and deep late-life regrets. In addition, greater participation of women in the workforce may preclude them from taking on the more traditional role of caregiver and further contribute to the relative lack of children able to support aging parents. Cultural patterns of marriage have also changed. Japan's almost universal marriage tradition stood in contrast to patterns in Europe where marrying at a later age or not marrying has been much more common. Interestingly, however, Japanese women are increasingly delaying marriage and childbirth, with a high percentage of the unmarried expressing the intention to never marry. This inevitably leads to a dramatically lower birthrate. The result will be a growing cohort of childless elders in many societies. Childlessness would seem to disadvantage individuals as they age but research suggests that childless individuals compensate for not having the support of children by strengthening relationships with friends and collateral kin such as siblings, nieces, and nephews. Nevertheless, in times of increased need for support in old age, childless elders are left without adequate levels of care. Women are more likely to be childless than men and childless elders of both sexes are much more likely to live alone or in institutions and have fewer face-to-face contacts than their peers who have children. However, interestingly there appears to be little difference in psychological well-being between childless seniors and those who are parents.

A social safety net can mitigate social pressures associated with old age. For example, in the United States social insurance provides the financial resources to allow many single elders to live alone. Similarly, housing alternatives for seniors including congregate living, supportive housing, and community-based social welfare services reduce isolation and dependency on others. However, it is not only practical support that is crucial to the mental well-being of elders. Rather, for many, it is the presence of true confidants that is key to warding off demoralization and depression. In this regard, informal care networks may be equally or more

relevant.

Despite enhanced social support, elders, especially women, without supportive families often face an emotionally impoverished old age. This is particularly true if governments erroneously assume that families will care for elders. Childless elders or those emotionally or geographically distant from their children will obviously be disproportionately disadvantaged by such policies and more likely to be alone. Clearly these issues have a central effect on the focus of psychotherapy. Problem-solving or interpersonal psychotherapy approaches for isolated elders will construct and address a different array of problems and issues than for elders who have strong support structures, such as the management of issues of abandonment or deficits in intimacy.

The nature of intimacy similarly shifts, often in the direction of distance and isolation. The equal partnership between married spouses may shift to a caregiver–care recipient relationship with consequent loss of mutuality and intimacy. Friendships, easily maintained during younger, more active years, often become much more difficult to sustain, especially in the face of physical impairments. Driving is more of a challenge, evening and night travel more fearsome, and gathering places less welcoming and accommodating. Healthy and active elders are more able to maintain equitable friendships where both friends can contribute equally. More frail elders receive more than they give, becoming the so-called "overbenefited" in their friendships. Equitable friendships or relationships in which elders (women) give more than they receive lead to higher morale than in overbenefited women. Interestingly, older men seem to establish more equitable relationships than women.

Affiliation and intimacy seem to be more important to women than to men, although such data should be approached with caution as there have been few studies on the needs of men for intimacy in old age. Some work has shown that, when compared to married men and married and widowed women, widowed men less often have relatives living close by and are less likely to be members of community organizations but may be no more likely to describe themselves as lonely. The data on the importance of intimacy have not been well verified for all cultural groups. For example, there's some evidence that black women place more importance on functional aid (as opposed to intimate affiliation) than do white women. Studies of Japanese elders have demonstrated an association between social support and self-rated health and mortality, results that were also noted in studies in Sweden and Finland.

Friendships are more than just emotionally supportive. Data suggest that the existence of a close friend and to a lesser extent kinship contacts increase the likelihood of maintaining functional independence.

The pattern and meaning of change in friendship networks in old age are different for men and women. For example, married men may derive

much of their self-esteem from their career and family and value friendships significantly less than married women. In old age, men may establish increased casual friendships but there is little indication of increased close relationships and overall data suggest that men (in contrast to women) decline in many measures of friendship. After retirement, men may establish casual friendships but turn to their spouses for personal meaning and self-worth previously derived from their work. With the death of a spouse, married men may be particularly vulnerable to isolation because of their relative lack of friendships and other community connections developed earlier in life. In addition, their spouse may have been their primary or only confidante and may also have been their link to the community. Married women may be relatively protected in bereavement because they are more likely to have other confidantes, both relatives and friends, and belong to community organizations. As a result, they have the opportunity to find other meaningful roles and relationships that may mitigate the loss of their spouse. However, for both men and women, social support from family or friends can only go so far. It is more difficult to lessen the internal loneliness that stems from the loss of the person who is often the spouse's most significant attachment figure and companion. Indeed, other relationships may not be able to alleviate that loneliness and the loss of the life partner must be mourned and therapy may provide an important venue for working through. While loneliness is a natural human experience it may not be easy to reliably measure using standard instruments. Why is this? While lack of social contact and support are key antecedents that can lead to loneliness, it is not equivalent to social isolation. Social isolation is an objective state and more easily identified and quantified while loneliness is an internal dysphoric and unpleasant experience which is related to the experience of deficiencies in social relationships; for example the feeling, right or wrong, that while there may be many people around to talk to nobody truly understands. This is a common expression of loneliness among elders such as those who are caring for a spouse with dementia. There are a number of loneliness measures both direct (self-report of loneliness) and indirect (measuring factors associated with loneliness without using the word itself to avoid the problem of stigma of self-identification) which yield different results. For example, a direct question revealed greater prevalence of loneliness among elderly women while an indirect questionnaire showed greater incidence among elderly men.

Illness and other age-related events like bereavement take their toll on friendship networks. These networks become depleted although the number of contacts may be less important than their quality. The maintenance of fewer emotionally meaningful relationships may provide sufficient social connectedness and protection against loneliness. Nonetheless, even for naturally outgoing personalities, the avenues for

establishing and maintaining new friendships are often blocked. One outcome is the erosion of relationships in which the old person's history, accomplishments, desires, and ambitions are known by his or her friends. Increasingly, aging fragments and isolates. Data on adaptation in the oldest-old are not consistent. The oldest-old, in contrast to the young-old, often become overburdened by illness and frailty, and even previously adaptive personality types with strong positive self-assessments often cannot successfully adjust to their disabilities. Some data indicate that subjective well-being can be sustained through positive self-referential attitudes. Results of these changes include physical distance from longstanding friendships and often, although not invariably, reduced frequency of contacts, which increase feelings of aloneness and create reliance on more casual opportunistic relationships such as with neighbors. While certainly not universal, this process of change contributes to creating a special context imposed by old age on the therapeutic process. Old patients, often through no fault of their own and despite their efforts, may present with a burden of aloneness. This is especially true for those who are unmarried or bereaved. This can create in therapists the anxious sense that they must be more than a treater, responding to an unspoken request for friendship and becoming for the patient the only place and person to offer a safe, intimate emotional haven. The therapist may find these challenges to be somewhat daunting as they may stir up the therapist's own fears about illness, incapacity, and death. The therapist may feel overwhelmed and worry about being engulfed by the patient's helplessness and dependency. In response, the therapist may distance themselves from or reject the patient, labelling them as chronically ill, beyond help and incapable of change. The reverse reaction may also occur in the therapist who wishes to halt or turn back the aging process. In this situation, the therapist may see therapy as life sustaining or even lifesaving. As a result, they may avoid termination rather than face the inevitability of the patient's and, perhaps, their own death.

Blanket rules about friendship and its impact in old age are inadvisable since the nature of friendships differs for each person depending on his or her psychological makeup. People differ in their need for, as well as their capacity to have, friends; their ability to maintain them; the nature and intensity of their friendships; their ability to do without them; and their standards for with whom they allow intimate confiding relationships. For example, Donald Winnicott suggests that those who never fully invested in friendships earlier in life may be better able to "make do" in old age with more superficial acquaintances without experiencing aloneness. Such individuals may be more invested in activities and personal expressions of creativity that are not dependent on contact with others (i.e., a well-developed capacity for self-soothing and self-sustenance). The implication for psychotherapy is that therapists are most effective when they explore in

depth with patients the nature of their friendship networks, their satisfaction with the level of intimacy and support, and strategies either for mourning lost relationships that cannot be replaced or for ways of enhancing their friendship networks to fulfill their needs. Many patients do not volunteer such information in therapy or may not be aware of their needs; hence, it must be elicited and explored actively. No specific psychotherapeutic techniques have been developed for dealing with these issues, but by extrapolation from other work (for example, on depression), interpersonal psychotherapy, cognitive-behavioral approaches, problem-solving therapy, or perhaps reminiscence therapy may be effective tools.

Assessing social isolation includes evaluating ease of relating to others; feeling of isolation from others; extent of sharing feelings with a significant other; connecting with others when needed; feelings of engagement with others; and feelings of aloneness and friendlessness.

Beyond friendship, sexual and other physical intimacy may diminish or disappear from the life of the old person—no one to touch or be touched by. The literature suggests that sexuality among the elderly can remain vibrant and pleasurable although the determinants of sexual activity are varied and may change over time. For example in one large study positive determinants of sexual activity included a positive attitude toward sexuality, sexual debut before age 20, having a very happy relationship and a physically and mentally healthy partner, good global health and mental health, being married/cohabiting, and drinking alcohol more than three times a week. While true as an isolated fact, the data also must be seen in the full psychosocial context, in which the elderly over a certain age are often widowed and without partners and with limited opportunity to form new close bonds. Lack of a suitable partner is the key predictor of sexual activity among women. One study found that 82 percent of residents of subsidized independent-living facilities had not had intimate sexual contact in the past year, although many had at least one experience of hugging or other touching. Hence, for example, intimate touching may become the domain of the medical encounter during physical examination or the nonsexual hugging of children, grandchildren, or friends. Physical isolation becomes especially evident in retirement or nursing homes or other institutions where predominantly old women live in an unusually gender-homogenous environment. Whatever sexual or intimate interest remains is put in cold storage, often forever. Having acknowledged this reality, it is also true that sexual activity and pleasure vary dramatically, often directly associated with the availability of sexual opportunities in relationships. Psychotherapists and other caregivers need to be sharply aware of the patient's needs and reactions under these circumstances, elicit the impact on the patient, and incorporate the realities into the treatment, either working toward change or to acceptance. The therapist must also be aware of countertransference reactions in this area. Unresolved conflicts over

parental sexuality may lead the therapist to avoid any discussion of sexuality or to regard the older client as asexual. As in all psychotherapies, the therapist must monitor their reactions to the patient through active self-examination. By bringing these reactions into awareness, the therapist can minimize their potential to have detrimental effects on the therapy and, instead, use them to better understand the patient.

The psychosocial phenomenon of loss of intimate relationships in old age may have special impact on the form of the psychotherapeutic relationship. Therapists are often and easily experienced by the isolated patient as a wished-for substitute for the lost relationship. Therapists may feel alone with the patient in their depleted world as well as uncomfortable pressures to adopt the roles or behaviors of friend, confidant, protector, advisor, or advocate. The therapist who identifies too closely with the patient's problems may be immobilized by pity or sadness that impairs their ability to empathize with or promote change in the patient. On the other hand, the therapist may shrink away from the patient's sadness, thereby increasing rather than easing the client's isolation and grief.

While old age is a time of substantial change and requires great adaptability, satisfactions may still abound. However, they tend to be encountered within a narrower frame, focused more on family relationships and very close friendships. For many, the loss or unavailability of friendships/relationships promotes a kind of centripetal emotional force that draws some elderly persons into their families and somewhat away from other relationships particularly when faced with the challenges of failing health. The loss of confiding relationships may be especially clinically relevant, since some research suggest that these relationships are protective against depressive disorder. Effective psychotherapy and environmental interventions that support the patient's need for new relationships can counter some of the effects of loss of confidants. While attention to relationships during therapy is important regardless of the patient's age, it is especially necessary with elders. Assessment and intervention must elicit not only the presence of instrumentally supportive relationships, but also which of them are sustaining of self-esteem and maintenance of identity.

The increasing importance of family relationships with advancing age adds an element to the psychotherapy of older adults not found as prominently among younger adult patients. Therapists often must take into account age-induced dependency issues in families and the conflicts these issues engender, such as forced intimacy and caregiving situations with adult children. For example, conflicts arise over the different perceptions of need that the elderly parent has. Children may be overly anxious and protective, aging parents inappropriately resistive of needed help or exaggerating of their helplessness. While family assistance and caring are crucial supports for the aging person, the therapist must be wary of

assuming that family can substitute for friends. Indeed, friendship interaction is positively correlated with self-esteem, while kinship interaction is not. Whatever the pattern, the psychotherapist of the elderly must take the family system into account, sometimes involving the other family members directly in the therapy or helping the elderly person develop appropriate interpersonal patterns of interaction. Flexibility of psychotherapeutic technique and interventions is essential. The usual adult–therapist stance broadens when treating the elderly as psychosocial issues associated with frailty, dependency, and relational shifts impinge often rendering individual assessment and focal intervention inadequate. Impelled by these realities, the therapist often must combine overt roles of advocate, intermediary, and educator with the more traditional psychotherapeutic roles and interventions. Sometimes this means reaching out actively to families, social agencies, or medical caregivers, roles that are often unaccustomed for psychotherapists. While therapists need to be flexibly supportive, they also need to examine their motivations when intervening on their patients' behalf. Transference- and countertransference-induced forces can lead therapists to inappropriately cross therapeutic boundaries, unwittingly adopting the idealized roles of adult–child or spouse seemingly in response to real psychosocial needs, but actually impelled by dimly understood or unconscious dynamic forces in the therapeutic relationship. Therapists dealing with relationship changes in aged patients also must be prepared to deal with their own feelings of loss, grief, fear, and loneliness and wishes to idealize, dominate, or avoid painful psychological issues that may be generated during treatment of the psychosocial stresses of old age.

ROLE CHANGES

Occupational role changes impose relatively unique challenges on the elderly. This group is almost universally forced to give up important sources of self-esteem, creative activity, social networks, and the organizing effects of job expectations. The importance of work is exemplified by the finding that nonemployed older women have lower self-esteem. Reactions to retirement differ. Retirement may be eagerly anticipated and handled well, especially among those who used to be called blue-collar workers and those with the resources, resilience, and creativity to use newfound leisure time in a satisfying way. Others, such as some professionals, often avoid retirement, and may find the loss of their occupational roles very difficult. Either way, the psychotherapist must be keenly aware of the centrality that work-related roles play in structuring life and the changes to this structure imposed by aging. These factors are especially relevant in the so-called young-old, 65 to 75 years, who are more likely to have freshly encountered the issues of occupational role change. For example, an emotionally distant marital relationship may have been quite stable and acceptable when work

substituted for unavailable marital closeness. After retirement the emptiness of this form of relationship becomes much more evident and can lead to depressive re-evaluation of life and choices. While difficult, this is also a time of therapeutic opportunity. Since retreat into work-related defenses is no longer possible, patients are forced to confront themselves and re-examine their long-held beliefs about others, especially their spouses.

Here, examining with patients the sources of their feelings both longstanding, deriving from early, perhaps unconscious developmental factors and more immediate, for example, associated with distorted modes of perception and thoughts about a partner who seems to want to care but is constantly rebuffed, is important. In the more advanced years, these issues generally recede, either adapted to and worked through or overwhelmed by other more pressing issues. However, when fresh, the psychosocial pressures of occupational role change, especially when combined with other events, such as spousal bereavement, often emerge as states of vulnerability, fearfulness, and impaired capacity to seek new, alternative sources of gratification.

Beyond its function as a source of personal gratification and self-esteem, the workplace serves an important social function. For example, one of the important considerations in retiring for women is the finding that they tend to give up work-related relationships after retirement. Working may have disrupted other familial and social roles and may not have allowed time for employed women to nurture or even maintain other relationships. In retirement, women may feel disconnected and that they have no role or purpose in life, particularly as other relationships may no longer be active. Hence, retirement creates a more impoverished social life for women, increasing their vulnerability to emotional reactions that accompany lost relationships.

As implied already, parenting roles may change dramatically in old age. For example, well-established patterns of parental dominance and reliance by children on a strong authoritative parent often suddenly reverse as parents become more frail and dependent. Conflict, guilt, or fear may invade a previously stable system. Particularly vulnerable are family systems whose members adapted to earlier conflicts by establishing distance, for example, children moving away from the pressures created by demanding, unreasonable, critical parents. In such circumstances, the increasing needs of aging parents may impose pressures on children to care for parents for whom they have had longstanding conflicted feelings. The psychotherapist must be careful to examine and understand the parent–child system and beware of taking stereotyped approaches to the aging parents, since the patterns of parental role changes are quite varied. For example, frailty and fear of being alone and incapacitated often leads elderly patients to complain about the inattentiveness of adult children.

Taken at face value, the psychosocial needs of the patient often are indisputable. However, the pattern of the family members' reactions to the patients' needs are rooted in the long history of their relationships with each other. One daughter found herself trapped providing care to her aging parents and subjected to a pattern of mistreatment well established earlier in her life. She found visits to her parents' home very difficult but, if she didn't go, her mother would attempt repeatedly to contact her and berate her when a visit eventually occurred. The daughter chose to go regularly to forestall intrusions and having to cope with her mother's disapproval and agitation. Her sense of obligation toward her parents kept her involved in a situation where she felt that there was little or nothing she could do to bring about positive change for herself or her parents. Another child spoke about her sadness at what she saw as the "changing of the guard" meaning she was now caring for the people who previously cared for her. To be effective, the therapist must be sensitive to the emotional forces that are preventing the family from adapting to the new, age-induced needs of the parent.

Many elders, mostly women, become caregivers to a spouse who develops dementia. They are subject to an array of stressors that lead to what has been termed caregiver burden. Burden is associated with disproportionate rates of both psychiatric and physical morbidity. Psychosocial factors as well as personal characteristics impact the resilience of caregivers and their capacity to manage this new and generally unwelcome new role. As with other stressors of old age social support and engagement are strong modifiers of burden as are the availability of instrumental support such as homecare, respite care, and easy access to professional support. These factors are often modifiable elements in treatment of caregiver burden when incorporated into management of caregivers. However, beyond these instrumental psychosocial factors, management of caregivers must be integrated with therapeutic interventions to strengthen adaptive, coping skills, and self-efficacy together with enhancing skills for management of dementia behaviors. It has been shown that fully integrated comprehensive approaches to caregivers which incorporate skills training, problem solving training, instrumental psychosocial support, and psychotherapeutic principles can be effective in enhancing key caregiver capacities.

SOCIETAL ATTITUDES AND RESPONSES

Biases about aging and the aged abound, and the older individual is confronted with them. Some of the North American perspectives on aging are quite negative—for example, the elderly are perceived as no longer productive and therefore marginal to mainstream life or, because of frailty, imposing undue economic burdens on overstretched pension and public support systems. As a result, there may be a tendency to dismiss the

problems of the elderly or a reluctance to provide help. The effective psychotherapist of the elderly incorporates these societally engendered issues into the therapy, recognizing with the patient the nature and impact of the messages, and working on appropriate responses. In this role, the therapist can help to buffer the sometimes harsh messages the old person hears daily about those in his or her age group. A particular issue is the difference between the patient's self-perception—more youthful, involved, and competent—and the perception of younger adults who can only see and react to the apparently universal aging appearance—lost skin turgor, decreased muscle mass, gray hair, and postural change. In other words, the older person's external and age-related life issues obscure the underlying depth of feeling and identity and can delay a more complete and deeper understanding by the young of the truer self-perceptions held by the older individual. Older patients in psychotherapy often feel both misunderstood by society at large and ashamed of themselves for their diminished capability and resultant helplessness, loss of purpose or role in life, or the burden they may feel they impose on others. In the presence of shame, individuals may fear new social situations and avoid them. The therapist will be most effective if these issues are articulated and explored with aging patients, helping them counter societally induced loss of self-esteem and feelings of alienation by confronting distortions and developing techniques to defend against inappropriate stereotyping by those around them.

ECONOMIC AND ENVIRONMENTAL FACTORS

For many, old age is a time of economic and environmental change. Retirement and the many difficulties sometimes imposed by living on fixed incomes, social security, or supplements provided by children can induce real or fantasized anxiety and preoccupation about financial security. These factors can influence psychotherapy in a number of ways, for example, inducing envious transferential feelings in patients or guilty countertransference in therapists.

Financial restrictions, often real, but sometimes imagined, lead to restricted social activity, feelings of isolation, humiliation, and shame, as well as uncertainty and fear of the future. Like advancing frailty or bereavement, economic difficulties must be actively dealt with in psychotherapy. This involves determining with patients their actual financial situation, separating distortions from reality, uncovering latent meanings about financial security, and then introducing both practical approaches to coping and interpretive understanding of the underlying meaning of the situation to patients.

Living accommodation is another important psychosocial variable that often changes in old age. The key element that induces psychological stress associated with moving to a new setting in old age is that it is generally a forced move, precipitated by physical or cognitive decline, economic

factors, or isolation. Even more than death, the specter of abandonment to an impersonal chronic care facility is most feared by the elderly. The prospect of such a move provokes various forms of abandonment anxiety, including the knowledge that this probably will be their last abode before death. This emotional reaction becomes more acute for married couples when the move necessitates separation of the partners and placement in different settings appropriate to each one's care needs. Moving to a long-term care facility also may generate internal conflicts associated with forced interaction with other elderly residents. Emotional responses may occur for a variety of reasons. For example, aged patients routinely express their anxiety at being confronted with the pain and disability of their age-peers. In them, they see mirrored their own future and become afraid. These feelings can be powerful motivators, preventing patients from accepting realistic moves to settings where they can receive needed care. Pressing a reluctant, frail patient to accept needed help in an institution can pose ethical dilemmas for the therapist. Not the least of these is the reality that many or most of these institutions are barely adequate to the task of satisfying the needs of their residents. Inevitably, no matter what the setting, the new resident entering long-term care will encounter severe psychosocial stressors, including some degree of loss of individual identity, regimentation to the institutional routines, and loss of privacy. Confronting these realities can be difficult not only for the patient, but also for the sensitive therapist who is only too aware of the limitations of the resources that they must rely upon.

The psychotherapist of an aged patient facing environmental dislocation will be most effective if they can uncover the patient's fears and anxieties—abandonment and loneliness, shame, infirmity, pain, or death—and work toward both realistic appraisal of the future and development of skills to combat the specific feared outcome of the environmental transition.

ETHNOCULTURAL ISSUES

Culture, ethnicity, and race have a profound impact on aging, which must be understood and dealt with by psychotherapists, especially when practicing in geographical areas made up of diverse communities. For example, elders among racial/ethnic minorities are less socially engaged with reduced access to social groups and non–church-related activities. This appears to be associated with greater vulnerability to less robust recovery from illness and mortality. Recent data suggest that racial/ethnic minority residents of long-term care facilities are less socially engaged than white residents although nursing home residents in general have low levels of social engagement.

Depending on the specific community, the presence of mental illness may carry with it a profound stigma that stops both patients and their

families from seeking care. These traditional cultural values are likely to be more evident in less acculturated members of ethnocultural communities. The therapist must be careful to evaluate the level of acculturation of a given patient and avoid stereotypical presumptions about cultural values, for example, assumptions about family support and filial piety.

When patients from minority ethnocultural communities do seek care, they often come with socioculturally sanctioned feelings of shame at their inability to cope with their emotional problems. Compounding the more universal problematic aspects of aging, aging members of ethnocultural communities also must deal with barriers to mental health care imposed by language- and immigration-related features, such as the feelings associated with leaving their homeland, unanticipated family conflicts, loss of status, or inaccessible spiritual support. Often, newly arrived or even well-established, but poorly integrated, aged immigrants are carrying a heavy burden of isolation, for example, unfamiliar with how to travel beyond their neighborhood, unduly reliant on often overburdened family members, feeling like aliens when they may have expected to feel like respected elders, encountering racist attitudes, or struggling with inadequate finances. While perhaps self-evident, it is crucial to the establishment of an effective therapeutic alliance for psychotherapists to convey to the patient that they are sensitive to these issues by enquiring about them in an empathic, caring, and specific manner. Therapists must be aware of the influences of their own culture on their choices, assumptions and prejudices, guard against cultural superiority and be willing to adopt an attitude of cultural humility and flexibility in adjusting their practice to accommodate cultural concerns. Most important is the overt demonstration to patients that the therapist recognizes them as individuals who will not be stereotyped by virtue of their culture, ethnicity, or race.

Therapeutic approaches and technique are influenced by cultural factors. Traditional cultural perspectives on aging and mental health forge the individual's view of the determinants of mental health and in turn the acceptability of and trust in therapeutic interventions. To be effective, the therapist must suspend preformed assumptions about the elder's background and culture and explore in an open, inquiring manner how cultural values influence explanations for symptoms and illnesses, diagnosis, and treatment. Therapists need to evaluate the degree to which culturally determined communication patterns affect therapy and adapt accordingly. These factors include four forms of communication patterns: (1) proxemics, the study of the perception and use of personal space (e.g., Anglos are more distant, while Latinos, Black Americans, and the French want much closer stances); (2) kinesics, the study of bodily movements—for example, facial expression or eye contact (e.g., among some Asians smiling is often an expression of embarrassment rather than mirth, and smiling may be seen by them as a sign of weakness rather than

friendliness); (3) paralanguage, which refers to vocal cues such as loudness of voice or inflections, and is easily misinterpreted (e.g., loud, rambunctious behavior is common in some groups such as Americans but is easily misunderstood as aggressive in low-volume cultures such as Japanese); and (4) high-low context communication, which refers to overt, intellectualized communication (low context) versus subtle or hidden communication rooted more in situations or internalized in the person. High-context communication is exemplified by how a mildly spoken "yes" can actually mean "no" when spoken by a Filipina challenged by her superior. In this context her "yes" is the expression of someone who is afraid to contradict her superior.

Once the patient's cultural perspectives are understood, the therapist's task becomes one of synthesizing this understanding and helping the patient integrate his or her traditional views with Western medical concepts (e.g., integrating the patient's view that he or she is suffering because of failure to honor ancestors with Western concepts of guilt). The successful culturally competent psychotherapist will be able to translate the patient's metaphors and belief systems, understanding both their concrete and symbolic meaning. For example, understanding that because physical symptoms are acceptable while mental symptoms are not, the patient may only feel able to convey distress in physical terms such as, "I have no energy," "My heart is uncomfortable or heavy," and so on.

Decision making in many cultures is group not individually based, as is more common in Western culture. It is necessary to therapeutic success that the therapist recognizes the importance of family hierarchy in decision making in some cultures and maintain an authentic stance of nonjudgmental collaborator with both patient and family.

An example of cultural knowledge that can be essential to the psychotherapy of a traditional, elder Native American, for example, is the centrality to the Native American of sacred tribal land. Tribal land usually refers to that which is occupied by an individual's tribe prior to their resettlement. Being torn away from traditional lands was a universal experience for Native Americans, disrupting many aspects of basic culturally based practices. For example, most religious ceremonies, culture, and custom revolve around tribal lands. Clearly, this form of dislocation, even though it occurred long ago, has had a profound effect on the personal identity of individuals, their sense of belonging in the world, their feeling of trust of those who are in power, and their sense of belonging within the wider society.

Of course, no one can hope to know all about every culture, but one basic ingredient is recognition that the elderly patient struggles with many culturally determined issues and conflicts that they may feel they have no right to speak about openly. For example, among groups who experience racism, elders may feel uncertain about how safe it is to reveal their anger

about white dominance and may view the treating doctor as a symbol of the very oppression and loss that they suffered in the past. Early in the therapy, therefore, it is useful to clearly indicate acceptance of the patient's uncertainty and, where appropriate, actively open such issues for discussion. This can be reinforced in the first assessment when the therapist can ask about the patient's feelings about traditional beliefs, experiences with the mainstream including racism, the difference between their private thoughts and their public beliefs, and their sense of belonging in society and in the therapy.

Psychiatrists and other therapists will be most effective if they are able to take a truly inquiring attitude with the patient and to be open to the patient's explicit attempts to educate them. In so doing, the therapist must also take care that the patient is not taking refuge in cultural differences to explain away all emotional reactions. It is important to carefully distinguish between patients' reluctance to reflect on and talk about themselves and true difficulties with expression. Having said this, it is also true that patients who speak a language fluently do not always use that language effectively for emotional expression. Patients may not be able to fully express their emotional state in English if it is their second language and can only convey the true meaning when they revert to their language of origin. This sometimes is relevant especially in the early stages of dementia if language skills begin to decline slightly.

Another important cultural competence of the therapist is the ability to accurately gauge the degree of closeness or distance required by the patient. The willingness or ability of individual patients to delve into important areas such as sexuality or interpersonal conflict with family members is often colored by feelings of shame and humiliation and prohibition against discussing such matters beyond the family, some of which are culturally determined. It may be helpful to preface inquiry into these areas by first acknowledging with patients their potential sensitivity and asking how they feel about exploring such issues. In other words, it is important to address the underlying culturally determined resistance that the patient may bring to the therapy before attempting to get to the substance itself.

The effective therapist treating patients from another culture will have the capacity to match the nature of the therapy to the level of belief and sophistication about Western methods that the patient holds. For example, the idea that there are social and interpersonal origins for feeling is not universally held around the world. Moreover, the implication that early life experience affects later life development may pose culturally determined problems in the therapy.

Factors associated with immigration may have special relevance to psychotherapy. For example, often hidden, but of great psychological importance, especially in refugee groups, are histories of severe deprivation

and trauma associated with war, political oppression, or famine. Because psychological trauma is a function of life experience, the prevalence of psychological trauma in the elderly will vary based on the life experiences of particular aged cohorts, for example, those who lived in times of political oppression, war or deprivation versus those who lived in relative peace and tranquility. One recent Canadian study for example found that 11.1 percent of the older adult patients presented with posttraumatic stress symptoms but less than 2 percent met DSM-IV criteria for full Posttraumatic Stress Disorder. The data are clear that late-life effects of earlier adult trauma endure, but their presence must be actively sought out by the therapist, as patients frequently keep silent about their experiences and their persisting effects.

SPIRITUAL ISSUES

Frequently omitted from psychotherapy inquiry, spiritual issues may play a profound role in the lives of many elderly patients. Religious or philosophical beliefs often are a source of great comfort in enabling appropriate acceptance of change, facing aloneness, or confronting death. Aging patients may return to religion in later years, both as a source of comfort and as a social resource. Some suggest that isolated older adults may benefit from becoming more involved in religious activity. While attending to spirituality and religious needs may be important it is often not addressed in psychotherapy and use of pastoral counseling is limited. Psychotherapy will be more acceptable to the patient and often more effective when it is carried out with a carefully elicited understanding of the patient's beliefs and attempts to use the strengths inherent in the patient's faith as an ally in treatment. However, the therapist must take care to deal with religion from the patient's perspective, not their own.

FUTURE DIRECTIONS

Regardless of the type of psychotherapy employed, the therapist of the aged patient must understand, acknowledge, and incorporate into the treatment the array of psychosocial factors that form the context of the patient's problems. The complex set of factors taken together is the matrix of the patient's life within which he or she will be supported or failed. In psychodynamics terms, this matrix may be viewed as the holding environment, the adult version of that which is provided by an infant's parents. As with children, when the holding environment fails in its supportive and enhancing functions, the individual comes under psychological stress. It is the person in his or her environment that is reflected in the environment within the person. Psychotherapists will be most helpful and effective when they work simultaneously with the external and internal worlds of the elderly patient.

REFERENCES

Anderson L. Loneliness research and intervention: a review of the literature. *Aging Ment Health*. 1998;2:264–274.

Aykan H. Effect of childlessness on nursing home and home health care use. *J Aging Soc Policy*. 2003;15:33–53.

Beckman N, Waern M, Ostling S, Sundh V, Skoog I. Determinants of sexual activity in four birth cohorts of Swedish 70-year-olds examined 1971–2001. *J Sex Med*. 2014;11:401–410.

Carstensen LL, Fung HH, Charles ST. Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motivation and Emotion*. 2003;27:103–123.

Chiu M, Wesson V, Sadavoy J. Improving caregiving competence, stress coping, and mental well-being in informal dementia carers. *World J Psychiatry*. 2013;3(3):65–73.

Fernandez ME, Mutran EJ, Reitzes DC, Sudha S. Ethnicity, gender and depressive symptoms in older workers. *Gerontologist*. 1998;38:71–79.

Field D. Continuity and change in friendships in advanced old age: findings from the berkeley older generation study. *Int J Aging Hum Dev*. 1999;48(4):325–346.

Giles LC, Glonek FV, Luszcz MA, Andrews GR. Effect of social networks on 10 year survival in very old Australians: the Australian longitudinal study of aging. *J Epidemiol Community Health*. 2005;59:574–579.

Ginsberg TB, Pomerantz SC, Kramer-Feeley V. Sexuality in older adults: behaviours and preferences. *Age Ageing*. 2005;34(5):475–480.

Goh A, Eagleton T, Kelleher R, et al. Pastoral care in old age psychiatry: addressing the spiritual needs of inpatients in an acute aged mental health unit. *Asia Pac Psychiatry*. 2014;6:127–134.

Hawthorne G. Measuring social isolation in older adults: development an initial validation of the friendship scale. *Social Indicators Research*. 2006;77:521–548.

Hurd Clarke L. Older women and sexuality: experiences in marital relationships across the life course. *Can J Aging*. 2006 Summer;25:129–140.

Jopp D, Rott C. Adaptation in very old age: Exploring the role of resources, beliefs, and attitudes for centenarians' happiness. *Psychol Aging*. 2006;21:266–280.

Koenig HG. Religious attitudes and practices of hospitalized medically ill older adults. *Int J Geriatric Psychiatry*. 1998;13:213–224.

Koenig HG. The role of religion and spirituality in the mental health of older adults. In: Sadavoy J, Jarvik L, Grossberg G, Meyers B, eds. *Comprehensive Textbook of Geriatric Psychiatry*. New York: W.W. Norton; 2004.

Koropeckyj-Cox T, Call VR. Characteristics of older childless persons and parents— cross-national comparisons. *J Fam Issues*. 2007;28(10):1362–1414.

Kreager P. Where are the children? In: Kreager P, Schrüder-Butterfill E, eds. *Ageing without Children: European and Asian Perspectives*. Oxford, UK, and New York: Berghahn Books; 2004:1.

Larsson K, Silverstein M. The effects of marital and parental status on informal support and service utilization: A study of older Swedes living alone. *J Aging Stud*. 2004;18:231–244.

Leong F, Lee SH. A cultural accommodation model for cross-cultural psychotherapy: illustrated with the case of Asian Americans. *Psychotherapy (Chic)*. 2006 Winter;43(4):410–423.

Li Y, Cai X. Racial and ethnic disparities in social engagement among US nursing home residents. *Med Care*. 2014;52:314–321.

Litwin H. The association between social network relationships and depressive symptoms among older Americans: what matters most? *Int Psychogeriatr*. 2011;23(6):930–940.

Lonnee-Hoffmann RA, Dennerstein L, Lehert P, Szoeke C. Sexual function in the late postmenopause: a decade of follow-up in a population-based cohort of Australian women. *J. Sex Med*. 2014;11:2029–2038.

Luo Y, Hawkley LC, Waite LJ, Cacioppo JT. Loneliness, health, and mortality in old age: a national longitudinal study. *Soc Sci Med*. 2012;74(6):907–914.

McFarland MR. Use of culture care theory with Anglo- and African American Elders in a long-term care setting. *Nurs Sci Q*. 1997 Winter;10:186–192.

Mendes de Leon CF, Gold DT, Glass TA, Kaplan L, George LK. Disability as a function of social networks and support in elderly African Americans and whites: the Duke EPESE 1986–1992. *J Gerontol B Psychol Sci Soc Sci*. 2001;56(3):S179–S190.

Nicolaisen M, Thorsen K. Who are lonely? Loneliness in different age groups (18–81 years old), using two measures of loneliness. *Int J Aging Hum Dev*. 2014;78(3):229–257.

Pin S, Guilley E, Spini D, Lalive d'Epinay C. The impact of social relationships on the maintenance of independence in advanced old age: Findings of a Swiss longitudinal study. *Z Gerontol Geriatr*. 2005;38(3):203–209.

Préville M, Lamoureux-Lamarche C, Vasiliadis HM, et al. The 6-month prevalence of posttraumatic stress syndrome (PTSS) among older adults: validity and reliability of the PTSS scale. *Can J Psychiatry*. 2014;59:548–555.

Sadavoy J. Survivors. A review of the late life effects of prior psychological trauma. *Am J Ger Psychiatry*. 1997 Fall;5(4):287–301.

Schwarzbach M, Luppa M, Forstmeier S, König HH, Riedel-Heller SG. Social relations and depression in late life-a systematic review. *Int J Geriatr Psychiatr*. 2014;29:1–21.

Sikorski C, Luppa M, Heser K, et al; AgeCoDe Study Group. The role of spousal loss in the development of depressive symptoms in the elderly—implications for diagnostic systems. *J Affect Disord*. 2014;161:97–103.

Smith J, Borchelt M, Maier H, Jopp D. Health and wellbeing in the young old and oldest old. *J Social Issues*. 2002;58:715–732.

Smith TF, Hirdes JP. Predicting social isolation among geriatric psychiatry patients. *Int Psychogeriatr*. 2009;21:50–59.

Tang F, Heo JG, Weissman M. Racial differences in social engagement and health status among older women. *Soc Work Public Health*. 2011;26:110–122.

Vikström J, Bladh M, Hammar M, Marcusson J, Wressle E, Sydsjö G. The influences of childlessness on the psychological well-being and social network of the oldest old. *BMC Geriatr*. 2011;11:78.

Walker RB, Luszcz MA. The health and relationship dynamics of late-life couples: a systematic review of the literature. *Ageing and Society*. 2009;29:455–480.

57.4h Individual Psychotherapy

JOEL I. SADAVOY, M.D., AND LAWRENCE W. LAZARUS, M.D.

INTRODUCTION

Psychotherapy for the elderly has evolved in step with the evolution of general adult psychotherapy. Any debate over whether psychotherapy is as effective for the elderly as it is for younger patients is now largely resolved. The data are clear from an array of outcome studies and meta-analyses that most forms of psychotherapy are demonstrably effective in elders treated for depression and anxiety. However, as with psychotherapy in general adult populations, there are few evidence-based guidelines that help the clinician determine which form of psychotherapy will work for a given patient. Some suggest that elders prefer shorter-term therapies although clinical experience indicates that there remains wide variation in what elders want and expect from therapy. Hence, therapists will be most effective if they have an array of techniques to choose from and determine on a case-by-case basis the specific form of therapy to be used.

Overall meta-analytic reviews of randomized controlled trials (RCTs) of various psychotherapies for elders have confirmed that the psychotherapies are generally effective. Overall effect sizes derived from meta-analytic studies of psychotherapy in older adults range from 0.63 to 0.78 with higher values for specific interventions such as 0.85 for 7 studies of cognitive therapy and 1.05 for 8 studies of reminiscence therapy compared to no intervention or placebo. Life review (LR) is a more systematic technique of reminiscence (R) and data indicate it is more effective. In life review therapy for depression in older persons, the effect size is more than twice that of reminiscence (LR-ES 1.28 vs. R-ES 0.52) and four times greater when well-being is measured (LR-ES 1.02 vs. R-ES 0.24).

Because of the complexity of the challenges of late life that impact on the adaptation of elders and their vulnerability to emotional disturbance and disorders, individual psychotherapy often needs to integrate other interventions into the therapeutic process. In addition to manualized models of psychotherapy in elders, models that incorporate combinations of therapy may be necessary to address the complex factors affecting elders' adaptation. For example, social support sometimes mediated through the use of simple technology like the telephone and psychoeducational models have shown positive synergistic effects with psychotherapy on depression outcome and systematic telephone-based psychotherapy has been effective in reducing distress in aging dementia caregivers.

Psychotherapies are effective treatments for depression in late life and may sometimes be considered as a first-line intervention. Studies comparing the efficacy of various individual therapeutic approaches for

depression in geriatric outpatients (e.g., cognitive-behavioral, interpersonal, problem solving, reminiscence and life review, and psychodynamic psychotherapy) have not yet been able to demonstrate the superior efficacy of one approach over another. At the Palo-Alto V.A., for example, a comparison of cognitive, behavioral, and brief psychodynamic therapy for depressed elderly outpatients found comparable remission ratios (53 percent had remitted after about 18 sessions). At 2-year follow-up, 77 percent of those whose depression had initially remitted were still depression free. The majority of patients who did not respond to the initial treatment remained depressed at the 2-year follow-up despite continued treatment. An important variable limiting reliable generalizability of the effectiveness in depression of the most well-researched psychotherapies in the elderly, cognitive-behavioral therapy (CBT) and problem-solving therapy (PST), is the fact that these studies have generally been conducted on young-old cohorts which are well white and well educated.

A recent meta-analysis of evidence-based psychosocial therapies for anxiety revealed support for CBT, cognitive therapy, relaxation training, and supportive therapy both for self-reported as well as diagnosed anxiety in older adults at posttreatment. One comprehensive analysis reported an overall mean effect size for treatment versus no treatment or placebo control of 0.55 just above the median (0.5), providing some although not robust support for psychological treatment of late-life anxiety.

Strongest support is for CBT for generalized anxiety disorder (GAD) but is weaker for subjective anxiety, where relaxation training has demonstrated more efficacy especially when practiced long term.

For more severely depressed outpatients, a combination of psychotherapy with an antidepressant may be more effective than either alone although many elders prefer psychotherapy over medication, Interpersonal psychotherapy (IPT), for example, has shown effectiveness in enhancing social adjustment when used in combination with an antidepressant. Even here, however, initially positive research outcomes have been called into question by more recent data. One problem in interpreting the data is that the variables of psychotherapy such as frequency of intervention, is not always standardized from one study to another limiting generalizability of specific studies.

Age-related brain changes probably affect the response of elders to psychotherapy although empirical data are modest. For example, reduced bilateral cortical thickness may be a marker of individuals at higher risk for poor response to psychotherapy.

In this section, the authors review the history of, and theoretical basis for, psychotherapy with the elderly, give brief descriptions of various therapies with illustrative clinical vignettes and suggestions for future research.

HISTORY

Modern psychotherapy began with Sigmund Freud's attempt to construct a universal model of understanding human behavior. In the prebiological era of psychiatric treatment, psychoanalytic theory formed the basis for most psychotherapeutic interventions. Unfortunately, Freud held a particularly gloomy, but highly influential, view of the potential for psychoanalytic treatment of patients over the age of 50. He believed that there was too much material to be dealt with and that the elderly had inelastic mental processes and hence were not educable. These ideas produced inordinate pessimism about the potential for psychological change in elderly adults. As the psychotherapies matured and began to be applied to a broader range of patients, therapists came to recognize that age per se does not preclude therapeutic change. Indeed, in the 1950s and 1960s several analysts wrote enthusiastically of their positive experiences in treating elderly patients. In retrospect, this enthusiasm resulted from treatment of generally vigorous elderly patients who were good candidates for analytically oriented psychotherapy.

THEORETICAL ISSUES

Effective psychotherapy, at a minimum, appears to require a strong, early, working alliance regardless of which specific psychotherapy technique is used. The working alliance is commonly understood to be agreement on the goals and tasks of therapy in the context of a positive affective bond between patient and therapist. Elders, not unlike younger patients, vary in their expectations and understanding of psychotherapy. They often come to psychotherapy with general wishes to feel better but uncertain what to expect. As with younger patients, the alliance in psychotherapy is enhanced when: the therapist steps back with the patient to inquire about the nature and form of the problems; helps the patient define agreed upon goals for therapy and articulates them clearly such as reduction of anxiety symptoms or improved relationship with a significant other; there is a personal bond between therapist and patient, in which the patient views their therapist as caring, understanding, and knowledgeable; both patient and therapist agree on the means to achieve goals.

Alliance is further strengthened when the therapist implicitly or explicitly conveys to the patient their understanding and knowledge of the age-related tasks and challenges facing the patient. There are many models for framing this understanding, depending on the type of therapy being used, but at the root, each model makes use of common themes now understood to underpin the adaptional challenges of those negotiating the geriatric phase of life.

A major developmental task of aging is finding restitution for the inevitable biopsychosocial losses associated with this stage of the life cycle.

Therefore, loss is one of the most prominent issues discussed by elderly patients. For many, it is the cumulative effect of multiple repeated losses, before sufficient time has passed for mourning and resolution, that is so devastating. The older person struggles to maintain ego integrity and a sense of purpose in life at a phase of the life cycle often characterized by diminishing coping resources, increasing self-doubt, and the onslaught of assaults and challenges to self-esteem and a positive self-image. The magnitude of an elderly person's reaction to a loss depends, to some degree, on the amount of pride and emotional investment in the lost function. Beyond lowered self-esteem, elderly adults experience life changes that can damage their sense of security and safety. For some, the world becomes uninviting and dangerous, particularly if they have encountered real events such as falls, personal attack, or inaccessible help at a time of need.

Perhaps counterintuitively, fear of death is not as prominent in the elderly as in younger adults and hence is not a major theme in geriatric psychotherapy. Clinical experience suggests that other fears especially focused on health concerns, incapacitation, memory loss, pain, abandonment, and helplessness are more evident. For example, in one psychotherapy-study cohort of seniors over age 60, reasons for seeking treatment were depression related to interpersonal relationships, health conditions, grief/loss, finances, housing and challenges due to executive dysfunction.

Self-worth, security, and independence depend in part on the individual's ability to maintain physical capacities. Beyond obvious challenges posed by physical illness or decline, impairment of important psychological functions becomes pronounced when frailty or illness forces restricted function and imposes a new, different, and often unacceptable self-image and personal identity.

Psychological adaptiveness in old age varies widely. Vulnerability to states of depleted self-esteem, insecurity, fear, or anxiety is heightened when an individual cannot adapt flexibly and is too wedded to maintaining or retrieving that which must be relinquished. Problems of psychological adaptation and diminished self-esteem occur in late life because of such factors as overdependence on social or work roles, overemphasis on the maintenance of control and mastery in all situations, and unresolved problems with self-esteem arising from earlier developmental stages. Difficulties with self-esteem may be rooted in maladaptive, unmodified beliefs; unconscious issues such as dependency conflicts; or interpersonally generated psychological reactions like unresolved grief.

The majority of elderly persons adapt quite well to the challenges imposed by negative societal stereotypes about aging and the losses noted above that threaten to erode self-esteem. Several psychological self-protective mechanisms help to maintain a positive self-image, such as

focusing on past successes, discounting messages that do not fit with lifelong self-concepts, refusing to apply general myths and misconceptions about aging to the self, and choosing to interact with those who provide pleasurable self-esteem enhancing experiences. Some selectively screen out unwanted perceptions such as those that challenge impaired cognitive abilities. While often adaptive and self-protective, such selectivity may also distort reality and promote maladaptive denial.

GENERAL TECHNIQUES

In the initial interview, the psychiatrist explores the major reason the patient has come for the interview and is active and empathic in exploring patient and family resistance to therapy. The patient and family are asked to bring with them all currently used prescription and over the counter medications. Psychotherapy with the elderly can only be successful if it is integrated with effective management of relevant somatic factors affecting adaptation and emotional responses. Hence, careful attention is paid to previously overlooked medical problems and medication adverse effects that may contribute to current psychiatric signs and symptoms.

The elderly patient is usually interviewed individually before important information and concerns are obtained from the family with the consent of competent patients. Elderly patients respond positively to the therapist's listening, empathic, and encouraging attitude. Some patients like to sit close to the therapist and may welcome such expressions of spontaneous warmth as a natural touch on the patient's arm. However, the therapist evaluates each patient's sensitivity about intimacy and does not assume that familiarity is welcome just because the patient is old. Inquiring about how the patient has successfully coped with adversity in the past helps to recall effective coping strategies and offers hope that current stresses and loses can be similarly overcome.

Psychologically less sophisticated patients, who are accustomed to talking about their physical complaints when visiting a physician, initially may feel more at ease if the therapist begins with a medically oriented focus. The therapist may then shift the focus to psychological and social problems.

In general, the therapist of the elderly patient is more active, engaging, supportive, and flexible than with younger patients. The therapist is knowledgeable about the normative, age-related changes that occur in the psychological, cognitive-intellectual, and social life of older people; geriatric medicine; psychopharmacotherapy; family therapy; and community resources. In both office and hospital settings, it is especially useful to work with a mental health colleague, such as a social worker, who is knowledgeable about family therapy and community resources. Patients must feel that they have obtained something from the initial interview (e.g., empathic understanding; hopefulness) to increase the likelihood that they

will return for subsequent sessions. Telephone reminders to selected patients or family members in advance of the next appointment impart the therapist's interest and improve patient compliance. Elderly patients, particularly the frail and dependent, may hesitate to express any negative reactions or doubts about therapy. The therapist might wonder out loud whether the patient may find it difficult, for a variety of reasons, to return for a second visit. Such exploration increases the likelihood that therapy will continue.

Realistic, explicit treatment goals are negotiated with the patient and, as appropriate, family. If treatment goals are overly ambitious, both patient and therapist may experience frustration and disappointment. Many elderly patients, whether consciously or unconsciously, have a specific treatment goal in mind, which once accomplished, leads them to terminate treatment.

SUPPORTIVE PSYCHOTHERAPY

More frail patients, such as those who are cognitively impaired or chronically ill, may require such psychotherapeutic techniques as support of healthier defense mechanisms, ventilation, advice, help with acceptance of diminished capacities, and increased dependency needs. Rather than trying to achieve intrapsychic change, the therapist strives to improve and refine the environment to the particular needs of the patient. For example, the therapist may help the family to arrange a more caring, nurturing environment and assist them to become better caregivers.

The patient's life history and psychodynamics (obtained from the patient or family members) enables the therapist to institute specific supportive therapeutic strategies.

An elderly woman with progressive cognitive impairment and lifelong personality traits of dependency and passivity is anxious and depressed about entering a nursing home. She may adapt well to the home if assigned to a roommate who is nurturing and protective.

Alvin Goldfarb has described a brief, supportive therapy technique for institutionalized, cognitively impaired patients. The therapist improves patients' foundering self-esteem, sense of control, and safety by permitting them to develop an apparent special relationship with the therapist, who is perceived as benevolent and powerful. The patients believe they have some control over the benevolent physician. This is accomplished in small, subtle ways. For example, the physician may elicit the patients' preferences for frequency of sessions, daily timetables or diet, then in a realistic or symbolic way the therapist acquiesces to the patient's wishes while maintaining a quiet caution about being perceived as being unduly manipulative.

INTERPERSONAL PSYCHOTHERAPY

There are not a lot of data supporting the efficacy of IPT in elders. However, what there suggests it is useful as either a primary intervention for depression and possibly as an adjunct to antidepressants although this latter data are inconsistent.

Nevertheless, older adults often have to cope with loss of friends and family, to adapt existing social skills to changing life circumstances and to resolve conflict with significant others, especially adult children. Hence, a therapy that helps the older patient cope with and resolve interpersonal problems would seem well suited to elderly patients. Recent cross-sectional study data in elders confirmed that loneliness is a significant risk factor for depression in both men and women although especially so for women. Interestingly and perhaps counterintuitively, social network size was not associated with risk of depression, although this is consistent with other findings, that is, it is the level of perceived intimacy and social support that is most critical rather than the overall size of the social network.

IPT is a time-limited (usually 16 weeks), manualized psychotherapy that focuses on one or two interpersonal problems, such as the death of a family member. It has been adapted and studied in the acute, continuation, and maintenance treatment of late-life depression.

In a series of controlled studies at the University of Pittsburgh, 79 percent of elderly outpatients with major depression who were treated in the acute and continuation phases with a combination of IPT and nortriptyline achieved full remission. Over the subsequent 3 years, 80 percent of the patients with recurrent major depression who had remitted and who continued on combined treatment, compared with 10 percent who received placebo, were still in remission.

LIFE REVIEW OR REMINISCENCE THERAPY

There is considerable empirical support from many RCTs of the efficacy of individual life review and reminiscence therapy for the treatment of depression, compared to no treatment control groups although support for group intervention is not strong. Along similar lines there is some evidence that autobiographical memory practice can reduce symptoms of depression when compared with "usual care."

Reminiscence is characterized by the progressive return of memories of past experiences, especially those that were meaningful and conflictual. It is theorized that reviewing the past helps the patient find greater meaning in the present. To varying degrees, elderly patients in therapy reminisce about the past, search for meaning in their life, and strive for resolution of past interpersonal and intrapsychic conflicts. Life review therapy is the therapeutic promotion of reminiscence by systematically enhancing this reminiscing process making it more conscious and deliberate. The therapist

may guide the process by encouraging the patient to write or tape an autobiography. Techniques include reunions with family and friends and looking through memorabilia such as scrapbooks or picture albums. This technique has been reported to resolve old problems, increase tolerance of conflict, relieve guilt and fears, and enhance self-esteem, creativity, generosity, and acceptance of the present.

Some authors have cautioned about possible negative consequences of encouraging life review, particularly for a depressed patient with obsessive tendencies or life histories filled with multiple misdeeds not amenable to a positive review. Evoking memories in patients who have suffered massive trauma or whose lives have been seriously affected by deprivation, overwhelming abandonments or humiliation will be emotionally challenging for patients and therapeutic technique is directed to providing an environment of safety and trust so that the evoked emotion can be contained modulated and perhaps reframed to enhance self-acceptance.

COGNITIVE-BEHAVIORAL THERAPY

The elderly are a heterogeneous population. There is no consensus on how much adaptation is necessary in CBT technique for elders. However, it is probable that when frailty, sensory impairment, or cognitive decline is present, therapy will be facilitated if some adaptations are employed. For example, with some elders the therapist is often more active with a focus on here-and-now issues. The pace of therapy is sometimes slower than with younger adults, and when appropriate, the therapeutic alliance may be enhanced by permitting the patient to adopt a senior, teaching role with the younger therapist. To compensate for specific deficits such as hearing loss or cognitive slowing, information may be presented in different modalities, such as written material or taped sessions for the patient to take home. Homework has been found to be feasible with many elders and helpful to the process including use of formal thought records. With older patients, material may have to be reviewed more often and termination is often managed gradually rather than abruptly.

There is some evidence to suggest that CBT is most effective for older outpatients with mild–moderate depression who are relatively healthy with high motivation for treatment and absence of a personality disorder. These characteristics are not unlike those of candidates for psychodynamic therapy and IPT.

PROBLEM-SOLVING THERAPY

Four RCTs investigated the effectiveness of problem-solving therapy for older adults. Three found that problem-solving therapy was better than other therapies and better than no treatment although not all trials are positive. There is some indication that PST is especially useful for depression associated with executive dysfunction where problem solving

and decision making may be particularly impaired. A large comparative trial of PST and supportive therapy found that the more structured and focused intervention of PST versus a less structured supportive method was more effective for treating late-life depression associated with executive dysfunction.

PST has been adopted for use in a variety of disorders and elderly populations with demonstrated effectiveness. It has been shown to be effective in reducing symptoms of poststroke apathy compared to placebo, late-life major depression compared to supportive therapy and reminiscence therapy, minor depression and dysthymia compared with paroxetine and placebo, vision-related quality of life in patients with macular degeneration.

PST is a brief solution-focused intervention that draws on methods commonly employed in other modes of structured therapy such as CBT or IPT, that is, direction and coaching, hierarchical lists created, homework assigned, relating problems encountered to evoked emotions, situational analysis of antecedents and precipitants for the problem and patient-generated solutions guided by the therapist are core to the method. Problems can range from major adverse life events, such as bereavement or illness, to everyday coping such as transportation, banking, or shopping. Once a problem is identified and its determinants are specified, a brief list of the pros and cons of potential but realistic solutions is generated by the patient and therapist taking into account the focal nature of the therapy and the time constraints of the brief format. Patients with mild cognitive impairment can utilize PST if they retain sufficient cognitive capacity to absorb the method and retain key elements from one session to the next. They may require additional supportive strategies, such as written notes, memory devices, environmental adaptations, and caregiver involvement.

PST may be most effective when delivered face to face either individually or in group formats but other acceptable and feasible methods of PST delivery have been shown to be effective for harder to reach patients, such as Internet-based therapy for depressed home-care elders. Improvement in quality of life and a reduction in anxiety symptoms have been demonstrated through use of videophone in hospice caregivers.

PST has been widely adopted for use with informal caregivers of those with dementia improving coping skills, self-efficacy, and depression. Frontline workers dealing with patients with dementia and their families in the community are often overwhelmed by the complexity of problems they must address during relatively short interventions. When professional workers in the community are trained in PST it has been shown that, compared to a control group of caregivers in standard care, PST training helps them organize their approach to complex problems. This in turn appears to enhance their ability to help caregivers, thereby significantly improving the capacity of caregivers on key measures of function and

emotion such as coping, mastery, and competence while significantly reducing emotion-oriented coping, burden, and stress.

BRIEF (TIME-LIMITED) PSYCHODYNAMIC PSYCHOTHERAPY

Brief, or time-limited, psychodynamic therapy can be considered for elderly patients with clearly defined, age-related problems that can be expected to resolve within a brief period of psychotherapy, such as an adjustment disorder, unresolved grief reaction, or new-onset anxiety disorder. Setting a time limit at the beginning reinforces the patients' confidence in their ability to resolve problems within a circumscribed period of time, focuses and accelerates the therapeutic process, diminishes fear of dependency, and accommodates limited financial resources. Brief therapy uses the therapist's psychodynamic understanding of the patient and the patient's transference to clarify and help resolve the patient's emotional reactions to a current life stress.

In a pilot study of the process and outcome of brief psychodynamic psychotherapy of patients with an adjustment disorder with depressed mood, the women showed greater improvement and were more active and assertive than the men who tended to become more dependent on the therapist. The patients used the therapist to validate competency and normalcy, which helped to support and restore feelings of mastery and self-esteem. Symptomatic improvement often exceeded the achievement of insight or self-understanding.

INSIGHT-ORIENTED PSYCHOTHERAPY

While studies of dynamic psychodynamic therapies are relatively sparse in number, several RCTs of psychodynamic therapy (including brief dynamic therapy) for older adults have shown that it is an effective treatment for depression comparable to other forms of therapy such as CBT in one head-to-head trial. Overall, data suggest that about 70 percent of people who had received dynamic psychotherapies were in remission at 2-year follow-up, broadly consistent with meta-analytic data on the overall efficacy of psychotherapy in elders. A rich array of both earlier and recent literature describing successful clinical cases of psychotherapy with the elderly, convincingly support empirical data and the position that, for appropriately selected cases, age per se does not interfere with the effectiveness of psychodynamic therapy including psychoanalysis.

The selection of patients and the basic approach to insight-oriented therapy for older patients who are physically and cognitively intact is largely the same as that for younger patients. Appropriate patients have the ability to form a therapeutic alliance, the capacity for introspection and to mourn loses and the motivation for change. Older patients may be particularly suited for this therapy given their appreciation of the finitude of life, increased tendency toward introspection, and a more philosophical

view of themselves and life.

During longer-term psychotherapy, the condition and circumstances of the older patient evolve and change. The development of physical illness may require medical referral, collaboration with family members, or advocacy on behalf of the patient for community and other support services. Such active intrusions into the patient's real-life problems are not commonly dealt with by psychodynamic therapists, but when treating elderly patients, techniques often need to be adapted to their changing needs. Therapists are most effective when they learn to integrate different therapeutic modalities while maintaining a psychodynamic focus on the therapeutic relationship, the transference, and the goal of inner conflict resolution. Indeed, the necessity to effectively integrate different modalities of treatment is one of the hallmarks of the therapist's flexibility necessary to treat and manage the evolving therapeutic issues of elderly patients.

CLINICAL ISSUES

Indications for Individual Psychotherapy

Elderly adults are a heterogeneous population. The continuum ranges from the normative or healthy older patient whose overall functioning is retained to the frail elderly patient. Five categories are normative elderly; normative elderly in crises; independent frail elderly; dependent, physically frail elderly; and dependent, cognitively impaired, frail elderly. Indications for various modalities of individual therapy are strongly influenced by the position of the patient on the normative-frail continuum. Moreover, because capacities change over time, therapeutic approaches for a given patient may also need to change.

With advancing age, crisis is frequently the stimulus for seeking out psychiatric care, often initiated by a family member. A small qualitative study found that common reasons for elders seeking psychotherapy included depression related to interpersonal relationships, health conditions, grief/loss, finances, housing and challenges due to executive dysfunction, that is mild cognitive dysfunction.

Individual therapy during crisis focuses on mourning for lost important relationships or personal capacities and working through conflicts over separation, redirecting energy to more realistic goals, redefining relationships especially with regard to accepting appropriate dependency and reestablishing or developing appropriate strategies to enhance mastery, sense of control, and self-esteem. Family education about ways of helping the impaired elder is of critical importance. Psychotherapy during crisis is often integrated with medication and environmental therapies. Problem-solving techniques, brief IPT, and CBT are useful approaches during the crises, although longer-term supportive or psychodynamically oriented therapy may evolve.

Behavioral techniques using operant behavioral conditioning have been applied to a wide range of target symptoms in chronically ill elderly patients, including incontinence, pain, insomnia, wandering, socially inappropriate behaviors, reduced level of activity and participation, toileting, bathing, and feeding. Environmental cues are manipulated to increase, decrease, or shape behaviors. These techniques are useful for institutionalized and frail elderly patients to reduce the excess disability that often accompanies dementia. An ABC approach to behavioral management of disturbed behavior may be useful. This technique includes accurate neuropsychological evaluation; defining specific measurable and observable behavior for modification; determining the antecedents (A) to the onset of the disturbed behavior and systematically recording them; recording the type, frequency, and duration of the baseline (B) conditions; tailoring a treatment program (A&B) to alter the behavior, which includes caregiver counseling, introduction of reinforcing, and extinguishing techniques, that is, consequences (C); instituting environmental changes and psychotherapy to assist caregivers; and program evaluation. A major deficiency of behavioral interventions is their labor intensiveness and the need to maintain the program. Once the interventions are withdrawn, behaviors tend to return to pretreatment levels in these institutionalized patients. Cognitive-behavioral interventions in long-term care settings, while often effective when introduced by a consultant, generally fail if the consultant leaves. Front-line staff in these settings usually lack the training to sustain sophisticated cognitive-behavioral programs on their own.

Various forms of individual psychotherapy have shown therapeutic value for states of bereavement. Bereavement can lead to chronic states sometimes called complicated or prolonged grief, which is characterized by persistence of acute grief symptoms complicated by cognitive emotional and behavioral symptoms. Complicated grief is common, found in 9 percent of bereaved older women. Psychotherapeutic interventions are particularly effective for bereavement complications such as chronicity, depression, or anxiety disorders. Most interventions include some education, support, attention to the relationship with the deceased before death and as it has changed since, and supportive assurances of "normality" for bereaved individuals who need to be reminded that their feelings and experiences are not deviant or abnormal. Complicated grief therapy (CGT) has structured this process and includes grief monitoring, imaginal and situational revisiting (promoting memory and using aids such as pictures or situations), imaginal conversations, building personal emotional and instrumental supports, defining and planning personal aspirations and opportunities for self-care and promoting positive emotions. RCT data suggest that, when compared to IPT, CGT may be highly effective for the elderly.

Psychotherapy of various types has been recommended for the

treatment of anxiety disorders. Preliminary data suggest that CBT can be quite effective in older patients with panic disorder, and may thus be considered as an alternative to pharmacological treatment for selected patients or as an adjunct to pharmacological treatments when patients have behavioral symptoms that can be targeted.

Psychotherapy has shown some value in the management of elderly patients struggling with alcohol abuse. Inclusion of cognitive-behavioral training that addresses themes of self-advocacy, self-esteem, and relapse prevention may aid recovery.

Individual psychotherapy is often helpful in the treatment of sexual dysfunction. For example, premature ejaculation usually responds readily to sex education, reassurance, and sex therapy techniques practiced at home alone or with a cooperative partner.

Limitations

Elderly adults are underrepresented in psychiatric outpatient clinics (2 to 4 percent of patients seen) and even more so in private psychiatric practice. Older patients are often not referred for therapy possibly because primary care physicians are not fully aware of its effectiveness in the age group. Overall a minority (25 percent in one study) of depressed elders receive psychotherapy but, even of these, only one-third got ongoing consistent therapy. Young-old, more highly educated elders are more likely to receive psychotherapy.

The application of psychotherapy, therefore, appears to be limited by the willingness of practitioners, both psychiatric and primary care, to use it for elderly patients. Completion of a series of regular therapy sessions, whether short or long term, is often limited by practical considerations such as difficulties with insurance and personal finances, transportation, or weather.

Raised in an era when shame was associated with psychiatric problems and treatment, some elderly persons feel embarrassed and indignant when their physician or family suggests consultation with a psychiatrist. Many elderly persons, as well as health professionals, believe that depression and anxiety are expected concomitants of aging or routinely attribute abnormal psychological states to physical illness. The symptoms themselves, such as depression, apathy, and hopelessness, prevent older persons from seeking treatment. Patients whose ethnoracial backgrounds differ markedly from that of their therapist may have quite different notions and expectations about a visit to a psychiatrist. Many cultures view mental disorder as shameful. Individuals may fear damage to their families' reputations. For example, elderly patients from some ethnic backgrounds have expressed fear that their grandchildren may be less marriageable if there is mental illness in the family. Beyond cultural barriers, other access barriers limit care for these groups including language incompatibilities, lack of

knowledge of where to obtain psychiatric services, economic issues, a tendency to seek care only after a crisis, and fragmented, difficult-to-access, mental health care services.

Adult children of an aging parent may minimize the parent's psychiatric symptoms, be ambivalent about assuming a parenting role, and may share their parent's negative attitudes toward psychiatry. All these barriers result in impaired access to care, poor compliance with medication and psychotherapy and a high dropout rate.

Strategies to overcome these formidable barriers in the very first session include active exploration of patient and family resistances to treatment, discussion of how ethnic differences between therapist and patient can interfere with treatment, ambivalent and negative attitudes about psychotropic medication, and by encouraging the patient to share reactions to the first session. If the therapist intuitively suspects the patient is resistant to treatment, active inquiry about problems that may interfere with subsequent sessions such as shame or practical issues like transportation, are explored. It is important that the patient feels that something has been achieved in the very first session. In this way the seeds of the therapeutic alliance are planted with both the patient and family in the initial visits. Even for the very impaired depressed patient bereft with hopelessness, conveying hope and a willingness to help can facilitate the beginning of a therapeutic relationship.

Complications

Whatever the form of psychotherapy, the course may be complicated by issues of transference and countertransference. The patient's feelings about therapy will be influenced by education, socioeconomic status, culture, and ethnicity that can restrict or enhance the impact of treatment.

Transference is age independent (of both patient and therapist), the determinants stemming from the timeless unconscious. However, developmental theory expands the traditional perspectives by demonstrating that the transference derives not only from significant infantile experience, but also from internalized objects encountered in adult life, for example, a spouse or a child.

While there are no age-specific transferences, the psychodynamic impact of life experiences during old age may increase the emergence of certain forms. For example, frailty leads to dependency conflicts and fears of abandonment, emotions that can easily evoke parental transference reactions toward the health professional derived from reactivation of the patients' need for a protector (parental figure). This form of transference emerges readily in dependent, institutionalized patients, even when psychotherapeutic contact is relatively brief and intermittent. Bereavement, especially of a spouse, or other loss (e.g., sexual capacity) may lead to an eroticized spousal transference to the therapist in which the

patient unconsciously relates to the therapist as an erotic love object who becomes a fantasy substitute for the lost object.

Reverse transference describes a phenomenon of geriatric psychotherapy in which the patient experiences and relates to the therapist as a child and takes a parental role toward the therapist in the treatment dyad. Therapists may experience this transference interaction as demeaning or patronizing and feel uncomfortable with the seeming reversal of dominance. Young therapists may find this especially challenging. It is generally a transient phase of therapy that defensively protects the patient from deeper levels of transference-based affect associated with conflicts over dependency.

Somewhat different is filial transference, in which patients deal with the therapist as one of their own children, enacting such unconscious conflicts as rage for not fulfilling wishes for caring, or guilt derived from earlier failures in the parental role.

Countertransference responses of therapists are often intense in the treatment of elderly patients, sometimes blocking the therapist's willingness to engage the patient in individual psychotherapy. The complexity or chronicity of problems of late life may lead therapists to be unduly concerned that older patients will become dependent on them and engender in some therapists reluctance to embark on treatment for fear of the patient's dependency. Similarly, it is not uncommon for younger therapists to avoid treating elders because of their anxiety over having to deal with frailty or decline, in older patients. Uncertainty about their ability to help the older person may induce feelings of guilt or anxiety in the therapist and lead to withdrawal from or avoidance of provision of psychotherapy.

Less experienced therapists may initially doubt the applicability of their skill to elderly patients, thus reinforcing pre-existing uncertainties about the mastery of their craft. The patient's frailty sometimes induces especially strong withdrawal from using psychotherapy because therapists are unwilling to encounter and confront their own hopelessness and fears of abandonment or mortality. Some therapists withdraw because of fear of being engulfed by the patient's demands and dependency. These factors may be compounded by the added therapeutic burdens of working at a slower pace, dealing with sensory deficits and cognitive decline.

Countertransference parallels transference reactions in that therapists may develop strong urges to rescue the patient and do too much, reactions that may derive from unresolved guilt toward their own parents.

Therapists must deal within themselves with the repetitive need for therapy to focus on mourning and acceptance of loss. This is in contrast to the more hopeful perspective of ego development and mastery, which is common in psychotherapy of younger adults and easier for therapists to deal with. Treating elderly patients holds the danger that helpless anxiety

or despair may emerge in therapists who have not successfully dealt with their own issues of loss and/or advancing age or frailty.

Therapists of elderly patients are often surprised by the emergence of erotic countertransference during more intensive therapy and may recoil from or deny its presence or feel uncomfortable about aspects of the patient's sexual feelings, thereby depriving the therapy of crucial areas of work.

The sometimes overwhelming effects of cumulative life trauma are painful for therapists, who may retreat into mechanistic or formulistic approaches, inhibiting their own empathetic relatedness.

A young therapist was surprised when a difficult elderly woman complained angrily about feeling manipulated and exposed when sent home to write her life story, a technique the therapist had read was suitable for the elderly. In discussion of the case with his supervisor, it became evident that he felt overwhelmed by the patient's painful life and uncertain of what would be helpful. Rather than engaging her in an attempt to understand her feelings, he prematurely employed a cognitive intervention that made her feel infantilized and rejected.

The process of termination may mobilize the very conflicts and issues that led to therapy, such as the loss of a loved one. Nevertheless, the termination process affords another opportunity for the patient to work through conflicts about the finitude of life and relationships, feelings about abandonment and loneliness, and dependency.

In approaching termination, elderly patients need to be assured of the therapist's continued future availability should the stresses associated with advancing age precipitate regression, conflict, and decompensation. While some elderly patients terminate successfully, for many therapy is an intermittent process to be resumed as needed. The therapist can ease the patient's return to therapy by anticipating, with the patient, reasons that may inhibit reinstating therapy such as fear of disappointing the therapist and feeling ashamed because symptoms have recurred. Termination is sometimes imposed by unanticipated events, such as illness. Both patient and therapist must deal with the emotional impact. The focus of the relationship shifts from coping with life to dealing with illness and in the extreme, death.

Contraindications

There are few contraindications to psychotherapy, although the form of treatment is carefully tailored to the needs of the patient. Psychoanalytic psychotherapy is approached with caution. The therapist first ensures that the patient retains sufficient ego capacity to deal with regression during therapy, to work through adaptive mourning, and has the capacity to replace lost sources of self-esteem.

As in younger populations, CBT is indicated for the treatment of acute or chronic depressive disorders, concurrent with antidepressant

medications when indicated, and for symptoms associated with personality disorder. Mild cognitive impairment secondary to dementia is not a contraindication to CBT, although suicidal depression or severe cognitive decline requires alternative therapies.

Goals of Treatment

Psychotherapy is most effective when it is used for specific indications with defined goals. Goals are modified to fit the function, cognition, and psychological capacity of each patient. Generally, the greater the frailty, the more modest are the goals. Goals for patients able to undertake dynamic psychotherapy include psychological mastery over the past and its unresolved grief; resolution of adult and childhood sources of shame and guilt; resolution of current interpersonal conflicts, especially with spouse or adult children; working through conflicts associated with loss of self-esteem and establishing substitute sources of gratification for those that had to be relinquished; coming to terms with failure to achieve ideal aspirations; and, at the most advances level, processing and integrating unconscious conflicts experienced in the therapeutic transference relationship. Pollock has termed the creativity and energy that may be released by successful resolution of losses *mourning-liberation*.

In crisis therapy, the goal is to reestablish the patient's best premorbid level of function. Examples include acceptance of appropriate levels of dependence or controlling excess, psychologically induced morbidity.

For the more debilitated, frail, elderly patients, psychotherapy goals are often symptom-focused behavioral changes. These include focal interventions; for example, dealing with dementia-related behaviors such as aggression, excess withdrawal, agitation, or delusions. With these patients, therapy goals include strengthening the roles of family caregivers and fostering a more appropriate, supportive environment for the patient.

ETHICAL ISSUES

Frail, institutionalized elderly patients have limited capacity to use traditional psychotherapies. They are easily viewed by therapists as beyond help. In allocating scarce resources, therapists struggle with whether to expend effort on this population or reserve it for their more vigorous elderly patients. This is a difficult ethical dilemma since frail elderly patients are the most dependent, helplessly reliant on caregivers, who in turn are burdened with the role of deciding who shall or shall not receive certain forms of care.

While most forms of psychotherapy require the active participation of patients, and therefore their inherent consent, behavioral interventions may be applied to the patient externally. In using behavioral techniques with frail elderly patients, therapists are advised to exercise caution in

distinguishing between externally imposed goals of therapy, that is, a focal symptom that is troublesome to staff and the overall welfare of an often helpless patient whose ability to consent to, and participate in, setting treatment goals is limited.

Clinicians are often confronted with ethical concerns over confidentiality and the patient's decision-making capacity. Both family members and therapists may inadvertently infantilize more frail patients and override their right to participate fully in the therapeutic decision-making process. The psychotherapist must carefully evaluate a patient's ability to make decisions. Similarly, the family's right to know about the patient and participate in therapy is carefully balanced by the preeminent right of the competent elderly patient to confidentiality in the therapeutic relationship. In practice, this is often a difficult issue for therapists, especially when families are intrusive, demanding, or litigious.

FUTURE DIRECTIONS

Conclusions about the efficacy of psychotherapy have remained restricted to elderly patients motivated to participate in research who are typically healthy, community-residing Caucasians in their 60s and 70s. Studies of the efficacy of psychotherapy in other populations (e.g., minority elders, the frail medically compromised and the very old) and in other treatment settings (e.g., long-term care, hospital) may help clinicians to adapt psychotherapy to these relatively unstudied elders. Future research should study predictors of treatment outcome, including subtypes of depression, the role of medical comorbidity, and how therapist and patient characteristics influence outcome. Given the current understanding that motivation and compliance with psychotherapy are major factors governing success, it would be helpful to understand elderly persons' resistance to therapy and strategies to prevent premature termination.

REFERENCES

Areán PA, Raue P, Mackin RS, Kanellopoulos D, McCulloch C, Alexopoulos GS. Problem-solving therapy and supportive therapy in older adults with major depression and executive dysfunction. *Am J Psychiatry*. 2010;167(11):1391–1398.

Ayers CR, Sorrell JT, Thorp SR, Wetherell JL. Evidence-based psychological treatments for late-life anxiety. *Psychol Aging*. 2007;22(1):8–17.

Chiu M, Pauley T, Wesson V, Pushpakumar D, Sadavoy J. Evaluation of a problem-solving (PS) techniques-based intervention for informal carers of patients with dementia receiving in-home care. *Int Psychogeriatr*. 2015;27(6):937–948.

Choi NG, Marti CN, Bruce ML, Hegel MT, Wilson NL, Kunik ME. Six-month postintervention depression and disability outcomes of in-home telehealth problem-solving therapy for depressed, low-income homebound older adults. *Depress Anxiety*. 2014;31(8):653–661.

Choi N, Marti CN, Conwell Y. Effect of problem-solving therapy on depressed low-income

homebound older adults' death/suicidal ideation and hopelessness. *Suicide and Life-Threatening Behavior*. 2016;46:323–336.

Francis JL, Kumar A. Psychological treatment of late-life depression. *Psychiatr Clin North Am*. 2013;36(4):561–575.

Gonçalves DC, Byrne GJ. Interventions for generalized anxiety disorder in older adults: systematic review and meta-analysis. *J Anxiety Disord*. 2012;26:1–11.

Gould RL, Coulson MC, Howard RJ. Cognitive behavioral therapy for depression in older people: a meta-analysis and meta-regression of randomized controlled trials. *J Am Geriatr Soc*. 2012;60(10):1817–1830.

Gustavson KA, Alexopoulos GS, Niu GC, McCulloch C, Meade T, Areán P. Problem-solving therapy reduces suicidal ideation in depressed older adults with executive dysfunction. *Amer J Geriatr Psychiatry*. 2016;24(1):11–17.

Hsieh HF, Wang JJ. Effect of reminiscence therapy on depression in older adults: a systematic review. *Int J Nurs Stud*. 2003;40:335–345.

Karlsson B, Johnell K, Sigstrom R, Sjoberg L, Fratiglioni L. Depression and depression treatment in a population-based study of individuals over 60 years old without dementia. *Am J Geriatr Psychiatry*. 2016;24(8):615–623.

Kiosses DN, Alexopoulos GS. Problem-solving therapy in the elderly. *Curr Treat Options Psychiatry*. 2014;1(1):15–26.

Klug G, Lacruz ME, Emeny RT, Hafner S, Ladwig KH, Huber D. Aging without depression: a cross-sectional study. *Psychodyn Psychiatry*. 2014;42(1):5–22.

Knavelsrud C, Bottche M, Pietrzak RH, Freyberger HJ, Renneberg B, Kuwert P. Integrative testimonial therapy: an Internet-based, therapist-assisted therapy for German elderly survivors of the World War II with posttraumatic stress symptoms. *J Nerv Men Dis*. 2014;202(9):651–658.

Laidlaw K. A deficit in psychotherapeutic care for older people with depression and anxiety. *Gerontology*. 2013;59:549–556.

Lavretsky H. Resilience, stress, and mood disorders in old age. *Annu Rev Geronto Geriatr*. 2012;32(3):49–72.

Lewis MI, Butler RN. Life-review therapy: putting memories to work in individual and group psychotherapy. *Geriatrics*. 1974;29:165–173.

Linde K, Sigterman K, Kriston L, et al. Effectiveness of psychological treatments for depressive disorders in primary care: systematic review and meta-analysis. *Ann Fam Med*. 2015;13:56–68.

Mackin RS, Tosun D, Mueller SG, et al. Patterns of reduced cortical thickness in late-life depression and relationship to psychotherapeutic response. *Am J Geriatr Psychiatry*. 2013;21:794–802.

Newton NA, Jacobowitz J. Transferential and countertransferential processes in therapy with older adults. In: Duffy M, ed. *Handbook of Counseling and Psychotherapy with Older Adults*. New York: John Wiley and Sons; 1999.

Nordhus I, Pallesen S. Psychological treatment of late-life anxiety: an empirical review. *J Consult Clin Psychol*. 2003;71(4):643–651.

Payman V. Psychotherapeutic treatments in late life. *Curr Opin Psychiatry*. 2011;24:484–488.

Pinquart M, Forstmeier S. Effects of reminiscence interventions on psychosocial outcomes: a meta-analysis. *Aging Ment Health*. 2011;16:541–558.

Plotkin, DA. Older adults and psychoanalytic treatment: it's about time. *Psychodyn Psychiatry*. 2014;42(1):23–50.

Qualls SH, Knight BG. Psychotherapy for depression in older adults. In: *Wiley Series in Clinical Geropsychology*. Hoboken, NJ: John Wiley & Sons Inc.; 2006.

Sabir M, Henderson CR, Kang S, Pillemer K. Attachment-focused integrative reminiscence with older African Americans: a randomized controlled intervention study. *Aging Ment Health*. 2016;2(5):517–528.

Sadavoy J. Personality disorders in late life. In: Laidlaw K, Pachana N, eds. *Oxford Handbook of Clinical Geropsychology*. Great Clarendon: Oxford University Press; 2014.

Schachter J, Kächele H, Schachter JS. Psychotherapeutic/psychoanalytic treatment of the elderly. *Psychodyn Psychiatry*. 2014;42(1):51–63.

Shear MK, Wang Y, Skritskaya N, Duan N, Mauro C, Ghesquiere A. Treatment of complicated grief in elderly persons: a randomized clinical trial. *JAMA Psychiatry*. 2014;71(11):1287–1295.

Sirey AA, Bruce ML, Alexopoulos GS. The treatment initiation program: an intervention to improve depression outcomes in older adults. *Am J Psychiatry*. 2005;162:184–196.

Spector A, Charlesworth G, King M, et al. Cognitive-behavioural therapy for anxiety in dementia: pilot randomised controlled trial. *Br J Psychiatry*. 2015;206:509–516.

Tremont G, Davis J, Papandonatos GD, et al. A telephone intervention for dementia caregivers: background, design, and baseline characteristics. *Contemp Clin Trials*. 2013;36(2):338–347.

Westerhof GJ, Bohlmeijer ET. Celebrating fifty years of research and applications in reminiscence and life review: state of the art and new directions. *J Aging Stud*. 2014;29:107–114.

Woods RT, Bruce E, Edwards RT, et al. REMCARE: reminiscence groups for people with dementia and their family caregivers – effectiveness and cost-effectiveness pragmatic multicentre randomised trial. *Health Technol Assess*. 2012;16(48):v–xv, 1–116.

57.4i Cognitive-Behavioral Therapy

ERIC GRANHOLM, PH.D., JOHN R. MCQUAID, PH.D., AND PATRICIA A. AREáN, PH.D.

DEFINITION

Considerable progress has been made in developing cognitive-behavioral therapy (CBT) interventions for use in elderly people with anxiety and depression, and researchers have recently begun to modify and evaluate newer CBT interventions for older people with psychotic disorders.

CBT is a form of psychotherapy that attempts to modify the interpretations of experiences that determine feelings and behaviors. The premise of CBT is that cognition (the process of acquiring knowledge and forming beliefs) can influence feelings and behaviors. CBT interventions help people to identify and correct thoughts and misinterpretations of experiences that are at the root of problematic behavior. Behavioral interventions (increasing rewarding activities, practicing new skills) both

improve mood and provide opportunities to test the accuracy and benefit of beliefs that may contribute to psychological symptoms and impaired functioning.

ADAPTING CBT FOR OLDER CLIENTS

When provided optimally, CBT is individualized for the specific client, regardless of age or aging-related issues. Problems and goals are identified in a collaborative manner, activities and skills are designed for the specific needs of the client, and the therapist and client continually review progress to determine whether treatment is providing the desired benefit. Clinicians should keep in mind that treatment modifications should not be made solely on the basis of a client's age, but should be made when issues related to cognitive impairment, disability and illness are present that can influence the effectiveness of treatment. These issues can occur in any age, but are particularly prevalent in older adults, especially common among those who are 75 years old or older. Therefore, before making any major accommodations, a good clinician assesses the need for making them first rather than solely making assumptions on the client's age.

Ageist Beliefs and Age-Related Challenges

One of the few age-specific modifications clinicians need to make has to do with cohort beliefs about seeking help and ageism. Older adults are less likely to seek help for mental health problems than younger adults, and have been noted to hold erroneous beliefs about aging. Cohort-related beliefs can devalue therapy (e.g., "Don't air your dirty laundry") and ageist beliefs held by both therapists and clients (e.g., "I'm too old to learn or change") can interfere with participation in therapy. Just as they can be applied to beliefs that increase psychiatric symptoms, thought-challenging interventions can also be used to modify ageist beliefs that interfere with treatment participation. Similarly, using CBT techniques to address interpersonal and behavioral challenges related to aging is also necessary, such as role playing age-relevant interpersonal situations (e.g., talking to a doctor about health concerns or eyeglasses) and practicing solving around age-specific concerns (e.g., finding transportation, coping with hearing or vision problems).

Transportation Barriers

Transportation and ambulation problems, which are more common among older clients, can further exacerbate problems with isolation and withdrawal from society. It may be necessary to provide transportation to therapy sessions or conduct therapy at convenient locations in the community. Providing treatment in the community can also reduce the stigma associated with going to a mental health clinic, which can especially

deter older clients.

Cognitive Impairment

Addressing neurocognitive impairment is another important factor to consider when designing CBT interventions for the elderly. Due to normal aging and a lifetime of comorbid factors that can impact cognition in older clients, especially those with severe mental illnesses (e.g., poor education, lack of work, physical illness, poor nutrition, poverty), cognitive functioning may be impaired to a greater extent in older clients. Simplifying and slowing down discussions, repeated practice of skills, and neurocognitive compensatory aids (e.g., reminder notes, treatment workbooks, writing down information in session, and acronyms to help with skill recall) can be used to help alleviate cognitive impairments. Given fine motor and vision difficulties that are also common in older clients, cognitive compensatory aids and educational materials should use large fonts and writing spaces. Clinicians who have worked with younger populations may need to modify their expectations about the speed at which change may take place, without assuming that older clients cannot make change.

Medical Problems

Many medical impairments increase in frequency as people age. With longer life spans, there is increased exposure to a variety of medical ailments, including cancer, heart disease, musculoskeletal disorders, and other chronic health problems. These changes in health have a number of direct and indirect effects on psychiatric symptoms and treatment. Physical suffering can directly influence mood, anxiety, and psychotic symptoms. In addition, reductions in physical health can interfere with engaging in rewarding activities, demands of health care can reduce financial resources, and sensory impairments secondary to health problems can interfere with both daily activities and interpersonal interactions. All of these problems can also interfere with participation in CBT, which emphasizes the active practice of new skills. Therefore, CBT interventions need to focus on increasing client activities within the constraints of their medical condition. In addition, cognitive interventions can be beneficial if they help clients to shift perspective from the negative aspects of illness and to focus on thinking that can improve their experience in the moment, even in the context of medical problems.

Role Changes

Aging is associated with transitions through different roles across stages of life. Some of these changes can be disruptive for people, particularly if they are unexpected or involve loss of important sources of reinforcement. For

example, retirement may be a positive experience for many, but for some it can be very disconcerting, particularly individuals who defined their worth primarily through their productivity. Other role transitions involve profound loss, such as changing from a spouse to a widow(er) or divorcee, or from a homeowner to a member of an assisted-living community. In addition, for some individuals role changes include unexpected demands of caregiving at an older age, either of a spouse, child, or grandchild, or alternatively losing independence and needing the aid of a caregiver. These role shifts can involve both changes in self-image as well as physical and emotional demands. These changes can contribute to depression by either reducing access to rewarding activities (e.g., work) or by activating beliefs the client has that can be depressogenic (e.g., "I'm unproductive and useless"). They can exacerbate anxiety symptoms as individuals face uncertainty and increased perceived risk (e.g., "my wife handled the finances, how can I manage without her?").

When clients are facing such role changes, CBT can be particularly beneficial. On the pragmatic side, behavioral exercises such as scheduling, information gathering, and contacting support resources can help patients more effectively navigate the role change. Cognitive interventions can address depressogenic thoughts, challenge beliefs that are inconsistent with the new role, and generate thoughts that are likely to improve functioning. These efforts can work in service of modifying schema that are less beneficial in the new role (e.g., seeing one's value as connected to productivity; feeling dependent on others) to a schema that is more functional given the realities of the situation (e.g., seeing one's self as innately valuable; remembering that new skills can be learned).

Interpersonal Events

Older adults are vulnerable to changes in their relationships that can have important implications for their mood and functioning. Deaths of spouses, other family members, friends, and acquaintances can both provoke grief and represent loss of important social supports. Interpersonal relationships can end through other means as well, including adult children moving away from home or the patient leaving the work environment. In addition, changes in relationships associated with aging, such as having a child become a caregiver of the older person or needing to become a caregiver of an older person, can strain those relationships.

Issues of loss, isolation, and improving social support, leisure activities, and interpersonal communication skills, therefore, are even more important for older clients. Again, the combination of behavioral activities, such as communicating with supports or developing new activities that can increase positive contacts, as well as thought challenging to reduce cognitions that exacerbate the pain associated with interpersonal events can be important themes in CBT for the elderly.

Culture and Age Cohorts

Broadly defined, culture provides the context that guides the development of beliefs and expectations. That context changes over time. Among geriatric populations, different age groups have different expectations. The oldest members of society, having lived through the Great Depression and World War II, grew up with different cultural expectations than their children, who were born during and after World War II and came of age during the 1950s and 1960s. It is important for therapists to assess each person's developmental history with a focus on understanding the beliefs and values that arose from that history. This assessment can then help in developing interventions that address unhelpful thoughts and behaviors while supporting the core values of the client. In addition, the assessment is critical to prevent clinicians from relying on general assumptions about a particular age cohort, but rather to have an understanding of how the experiences shaped the person with whom they are working.

In reviewing the domains above, two major themes arise: *Loss* (of health, role, loved ones, finances) and *danger* of further losses in these domains. Loss tends to be associated with risk for depression, and perceived danger with risk for anxiety. Treatment for older individuals in particular needs to address coping with such losses in a constructive manner, helping clients differentiate between likely and unlikely threats, and identifying coping strategies for upcoming risks.

Traditional CBT emphasizes change orientation, attempting to change how one acts and thinks to improve mood. However, emphasizing change alone can be limited in effectiveness and potentially invalidating of the client, if the therapist does not also acknowledge the unchangeable nature of some challenges in the older person's life. Some CBT experts refer to this as "the healthy management of reality." From this perspective, the role of CBT is to help clients accurately understand the challenges they face, and then when appropriate modify their actions and thoughts so they can live a meaningful and rewarding life within the context of those challenges. Recent models that have derived from CBT and incorporate acceptance and mindfulness approaches (including dialectical behavior therapy, acceptance and commitment therapy, and mindfulness-based cognitive therapy) all emphasize both skills in change as well as acceptance of unchangeable realities.

TECHNIQUES

Stages of CBT

In general, CBT can be framed in three stages. The first stage is assessment and initial treatment and includes definition of clients' current problems and goals, education about the model of treatment, establishment of a collaborative relationship, and agreement on a treatment plan. In addition,

the therapist begins to develop a case formulation. The formulation is a model that attempts to explain the current difficulties in terms of a schema (an organizing set of deeply held beliefs that define the client's perception of self and determine the information that the client attends to, remembers, and recalls, and how that information is interpreted). One goal of this phase is to instill hope that improvement can occur, if possible through the client experiencing some symptom relief secondary to an initial exercise or homework assignment. The middle stage of therapy focuses on skills acquisition, with the primary goal of the client learning to implement CBT interventions independently. The final stage, relapse prevention and termination, emphasizes consolidation of skills, problem-solving obstacles to skill use, identifying potential risks for symptom exacerbation, and developing a plan to apply learned skills in the case of symptom recurrence.

Assessment and Initial Treatment

CBT emphasizes objective measurement of constructs whenever possible, and this begins in the assessment phase. When working with older clients, this will include assessing not only symptom severity, but also the resources available and the challenges faced, and then trying to define all these factors in objective terms. Number of supports, income relative to expenses, health limitations, and caregiver responsibilities are all-important aspects to assess. In addition, standardized measurement of symptoms provides a baseline to assess progress in treatment. In particular, the use of measures with normative data for older populations (such as the Geriatric Depression Scale) is important because several symptoms of depression and anxiety (e.g., sleep changes, reduced energy, memory, and concentration difficulties) can occur in older populations independent of depressive symptoms, and therefore measures need to account for these potential confounds.

As noted above, a key goal of the initial assessment is the development of a case formulation. A case formulation provides a framework for choosing interventions that match the specific needs of the client. Case formulation is particularly important when dealing with special populations such as older adults because it guides both the assessment and the treatment. A well-defined case formulation increases the likelihood that areas such as those reviewed above are addressed and incorporated into treatment.

Juanita was a 69-year-old divorced, unemployed Mexican American woman, who supported herself on her Social Security benefits. She lived in a one-bedroom apartment. Due to her daughter having been arrested for drug possession, she was serving as guardian of her 5-year-old grandson. She presented with the following concerns:

1. Depressed mood: Juanita reported daily crying, anhedonia, increased appetite and weight gain, feeling tired "all the time," feelings of guilt and shame, and thoughts of death.

2. Anxiety symptoms: Juanita met criteria for generalized anxiety disorder. She reported constant excessive daily worry about both major stressors (her daughter's legal troubles; financial stress) as well as other factors (her children, her friends, safety).
3. Daughter's arrest and custody of grandson: Her daughter had been arrested approximately 1 year prior to treatment, leading to Juanita receiving custody of her grandson. Her grandson frequently cried at night for his mother, and Juanita felt overwhelmed by trying to provide for his care. Her daughter was incarcerated several hours from where Juanita lived, preventing visitation.
4. Health problems: Juanita had type II diabetes. She had difficulty managing her symptoms in part due to anxiety about checking her glucose levels. Juanita was obese and worried that her diabetes would progress and lead to amputation of her foot.
5. Social isolation: Although Juanita had family and friends in the area, she rarely contacted them or answered her phone.

In the intake interview, Juanita stated, "I tried to be a good mother, but I failed. And now I'm going to fail my grandson." When discussing goals, Juanita stated that her greatest concern was her daughter's and grandson's futures. In discussing this goal in terms of what Juanita could influence, she agreed that she would like to provide a safe and loving environment for her grandson, and to be a support for her daughter. These goals required her taking several actions around caring for herself, including:

1. Reconnect with family and friends to increase pleasant activities and get help with caring for her grandson
2. Improve her diabetes care and reduce weight so that she could be physically better able to care for her grandson
3. Generate additional income by re-establishing her home-based business sewing traditional dance dresses for a local dance studio
4. Make arrangements to visit her daughter in prison with her grandson on a regular schedule

In this case, several of the described factors that are associated with depression and anxiety in older people were apparent. Juanita had clear medical problems that limited her activities, she had limited resources, and she was isolated from supports (albeit by her own actions). She also experienced a clear loss (her daughter being incarcerated) and a demanding role change (becoming caregiver to her grandson). In the context of the meaning of these events ("I failed"), it is not surprising for her to be experiencing depressive symptoms. Of particular note, her anxiety and depressive symptoms were exacerbating both her health problems (interfering with diabetes management) and interpersonal functioning (avoiding supports).

In reviewing the goals, the initial theme expressed by Juanita was a desire to care for others. However, the therapist was concerned that without taking care of Juanita's own issues, she would be unable to achieve her goals. Therefore, the therapist helped focus on self-care in terms of necessary preparation for her goal of caring for others. In this case, the proposed formulation identified a potential core belief held by Juanita of "I am undeserving; I need to care for others to get love." Rather than directly challenging this initially, the therapist first worked with strategies to achieve the goals Juanita had identified. Over the course of therapy, an additional goal of schema change (moving from "I'm undeserving and need to care for others to get love" to "I am a good and deserving person who loves to care for others") was instituted, but initially therapy focused on engaging Juanita in treatment so that she could gain some symptom improvement before taking on the more difficult tasks of acceptance and schema change.

In working with the therapist, these general goals of providing for her grandson were translated into more specific goals that could meet her goals and be beneficial to her in terms of improving her mood and functioning.

Middle Stage Interventions

Although the interventions in this stage of therapy are generally divided into behavioral and cognitive interventions, they frequently have cognitive and behavioral aspects. Behavioral interventions can lead to changes in cognition (e.g., doing an activity and finding out it is more enjoyable than predicted), while cognitive interventions can be useful in helping patients to do behavioral interventions (e.g., challenging the belief that one has to "feel" like doing something before doing it). Research has shown that cognitive or behavioral exercises lead to change in both cognition and behavior. In general, interventions focus initially on assessing baseline thoughts and activities and their relationship to mood and other outcomes. Clients then learn change strategies and coping skills in this stage of therapy, and assess whether these skills are effective in improving outcomes.

Psychoeducation. This initial assessment of thoughts and activities and their relationship to mood and other outcomes provides an opportunity to introduce the cognitive model and the idea that thoughts influence feelings and behaviors. The groundwork for future cognitive work is laid by using the client's own experience to illustrate how thoughts might be interfering with their goals by producing unwanted moods and behaviors. With all clients, but particularly with older clients with cognitive impairments, it is helpful to draw on a pad or marker board or use a handout to illustrate associations between thoughts, feelings, and behaviors, as shown in Figure 57.4i–1.

Depending on individual client needs, it may also be important to provide education about symptoms associated with specific diagnoses, the action of medications, medication adherence strategies (especially in the context of memory impairments), and complementary interactions between medication and psychotherapy treatments. It is often useful in this context to provide education about the vulnerability-stress-coping model of mental illness, focusing on the interaction of biological vulnerabilities, stress triggers, and coping strengths in symptom exacerbation or relapse. Finally, depending on the individual client, it may be essential to provide education about triggers and signs of crisis and develop a safety plan for prevention of suicide or relapse/hospitalization.

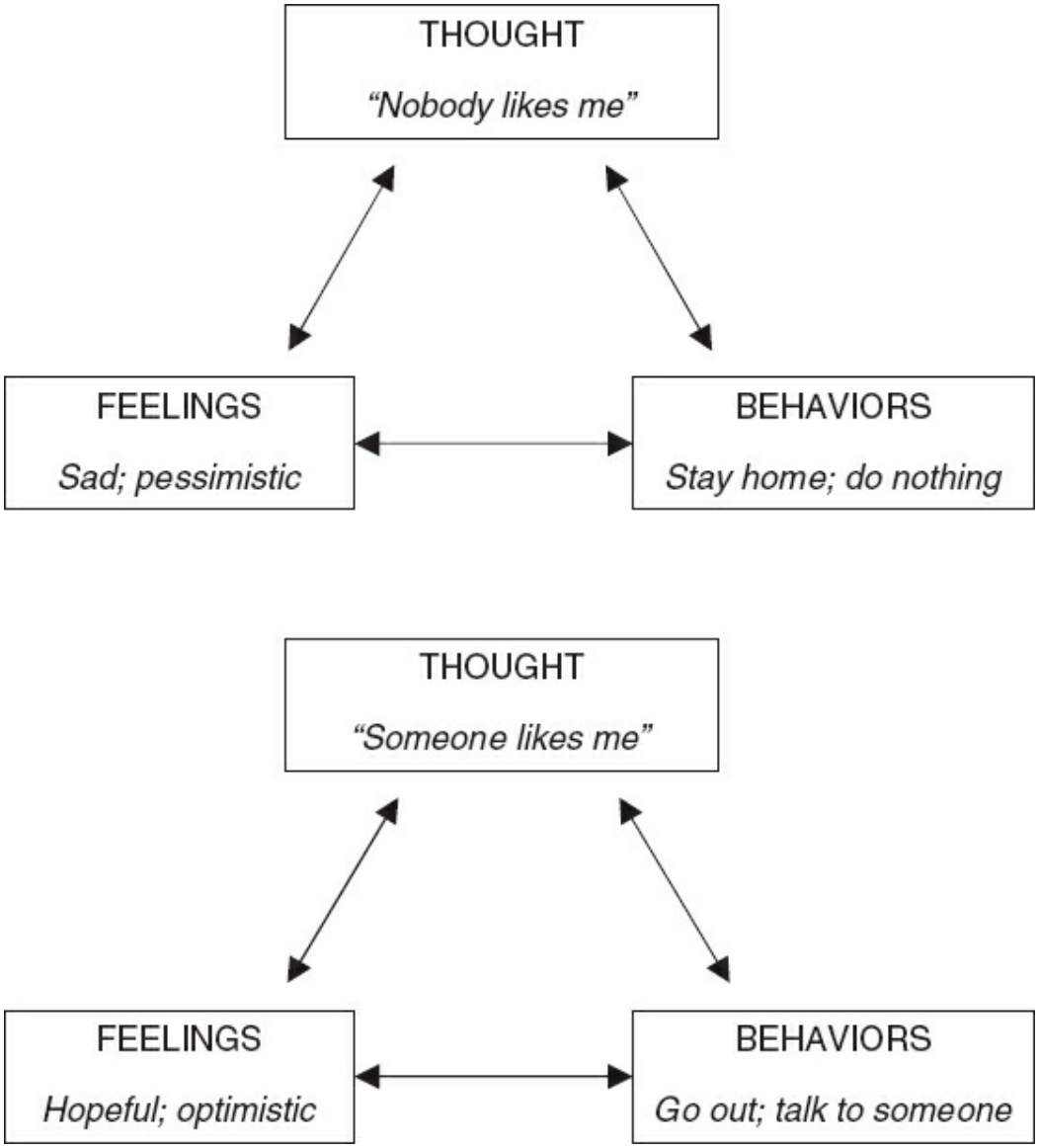

FIGURE 57.4i–1. Generic cognitive model illustrating relationships between unhelpful (*top*) and helpful (*bottom*) thoughts, feelings, and behaviors.

Behavioral Interventions. Behavioral interventions in CBT for mood disorders tend to emphasize activities that increase positive affect, gather evidence to challenge unhelpful beliefs, or both. Initial interventions can solely consist of activity tracking (e.g., recording a baseline of all activities that a person is doing throughout the day). Activity scheduling is simply identifying one or more activities that clients may find rewarding and planning it into their day. Keys to successful activity planning are generating multiple options, working with the clients to identify choices they believe they can complete, and problem solving potential obstacles to the activity. In addition, tracking mood levels before and after the activity can help clients notice whether there is any benefit. In the elderly, medical problems can significantly limit motility and activity and older clients can have beliefs (e.g., "But, I've been doing it this way for years") that make

changing routines can be particularly challenging. However, activities do not have to involve strenuous physical activity to enhance mood or alter cognitions, and most clients can manage some change in routine if care is taken to address on interfering cognitions and problem solve around limitations.

A variant on activity scheduling is pleasure predicting. This is a useful technique when clients doubt that they will enjoy an activity. In addition to rating mood before and after the exercise, clients are asked before they do the activity to predict how much they will enjoy it. Comparing the predicted outcome with the actual outcome helps challenge the belief "I won't enjoy it" and helps to highlight the tendency to underestimate the benefit of negative events.

For anxiety disorders, behavioral interventions include gradated exposure, which involves engaging in feared activities in an organized fashion. Using principles of systematic desensitization, clients break down a feared activity (e.g., shopping in a store) into specific steps that are arranged from least anxiety provoking to most (e.g., shopping in a large grocery store early when no one else is there versus shopping in a small crowded shop at the busiest time of the day). Clients practice doing less anxiety-provoking tasks first until their anxiety level is reduced, and then move on to the next task. Clients both become less physiologically reactive to the feared situation, as well as develop more helpful cognitions (e.g., "I can handle this.").

Clients in anxiety treatment can also be taught a variety of relaxation techniques, such as progressive muscle relaxation or breathing retraining, to facilitate management of their anxiety symptoms. The limited research available suggests that relaxation alone can be as effective as CBT for reducing anxiety symptoms in older clients. Recommendations specific to late-life anxiety include (1) ensuring that sufficient time is devoted to learning relaxation training skills, (2) emphasizing in vivo rather than imaginal exposure exercises, and (3) de-emphasizing more abstract cognitive restructuring techniques (e.g., directly disputing dysfunctional thoughts).

A related exercise that has been incorporated into both anxiety and depression treatment, and to some extent acceptance interventions for psychosis, is mindfulness meditation. Mindfulness is defined as paying attention in a purposeful way, in the moment, nonjudgmentally. Clients are taught to allocate their attention to a particular stimulus (e.g., their breath), and to just notice this. As distractions arise (e.g., internal thoughts, external sounds), clients learn to just notice these, as well as their own reaction to them. They then return attention to the task at hand. Mindfulness is a particularly useful skill in depression and anxiety for helping clients learn to observe their thoughts. When people can notice their thoughts, they can then question the accuracy of them and do not

need to act on those thoughts.

Juanita had stopped communicating with other family members, even though this had been a rewarding activity in the past. Therefore, therapy focused on this as an opportunity for behavioral activation.

Therapist: One thing you told me is that you used to enjoy time with your son and his family. Why don't we work on arranging some time with them?

J: Okay.

T: Okay, so what day and time do you think makes sense?

J: Well, my son works, so it would be best on the weekend, maybe in the afternoon.

T: How do you usually arrange to get together?

J: I just call him up and tell him I'm coming over.

T: Is there anything stopping you now?

J: I just don't want to be around people. I just feel like I'll bring everyone down. That would just make things worse.

T: Remember, part of the goal is to test beliefs like that. Would you be willing to go visit and then write down your evidence as to whether you brought everyone down and things got worse?

J: Okay.

This example highlights several important skills: Activity planning (spending time with the family), identifying obstacles (thoughts about her response), and challenging thoughts (gathering evidence from the experience).

Cognitive Interventions. One of the key skills that clients learn is thought challenging. The key steps to thought challenging are:

1. Noticing clients' thoughts and identifying their effect on mood and behavior
2. Learning to test the helpfulness and/or accuracy of those thoughts
3. Generating a new, more accurate or useful thought
4. Examining the effect of the new thought on mood and behavior to determine if it is an improvement

For some clients, the initial task of noticing thoughts is challenging. For clients who have difficulty even noticing their thoughts, a form such as the automatic thoughts checklist can be helpful. This is a list of thoughts associated with depression. Clients read through the list and check those that they experience. Once clients can recognize their thoughts, there are a variety of forms for teaching clients to examine and change them. With older clients a straightforward model, such as the "3C's" form, can be useful, because it is a simple model (Table 57.4i–1). Clients are initially taught to write down thoughts in the "catch it" column and the relationship to their mood in the "check it" column, as well as any errors or problems with the thought. Once clients can recognize their cognitions, they learn strategies to question and challenge thoughts that are unhelpful. Strategies include identifying common unhelpful thoughts (e.g., all-or-nothing thinking, predicting the future) and examining evidence both in support of

the unhelpful thought and in contradiction to it. Throughout these exercises, clients are taught to determine the effectiveness of these efforts by rating changes in their mood (Table 57.4i–2).

Table 57.4i–1.
3C's Homework Form

Catch-It, Check-It, Change-It (3C's) Worksheet	
Situation: What was happening? Where and when?	**Feelings:** How was I feeling at the time? **Actions:** What was I doing?
CATCH IT: What was I thinking in this situation? What went through my mind? Which thought best explains how I was feeling and what I was doing?	
CHECK IT: Check to see if the thought is accurate by listing the evidence for and against the 'Catch It' thought. Also, check to see if the thought is a mistake in thinking.	

FOR	AGAINST	MISTAKES IN THINKING
		(Check any that apply) ► All or nothing ► Mind reading ► Fortune telling ► Jumping to conclusions ► Catastrophizing ► Emotional reasoning

CHANGE IT: What would be a more accurate and helpful thought? If there was mistake in thinking or if the thought was inaccurate, develop a more helpful thought based on the evidence you listed.

Many people, including therapists and older people themselves, hold unhelpful beliefs about the older adults. These can include thoughts such as, "Older people can't learn/change," "Others do not want to be around older people," "Older people are a burden," or "Older people don't have much to give." Beliefs about managing suffering ("Don't air your dirty laundry") or about therapy ("Therapy is for crazy people") can also interfere with the acceptability of treatment. Therapists need to attend to both of these sets of beliefs, both in how therapy is initially presented and in the course of the actual treatment.

Table 57.4i–2.
Common Types of Unhelpful Thinking

All or nothing	Viewing situations in terms of absolutes: "I am a complete failure"
Selective attention	Attending to negative information and/or ignoring positive information (e.g., noticing health problems and ignoring supportive family members)
Shoulds	Holding rigid rules and expectations; refusing to accept realities outside one's control: "I should be able to do all the things I used to do."
Emotional reasoning	Judging a situation based on one's emotional response, rather than objective evidence: "I was feeling anxious; therefore, I must have looked foolish."
Mind reading	Making assumptions about others' perceptions: "They think I'm too old."

Predicting the future/probability estimation errors	Overestimating the likelihood of a negative event: "I know that I will fall and break my hip if I go to the store."
Catastrophizing	Underestimating one's ability to cope with a negative event: "I can't stand it if I get sick."
Overresponsibility	Feeling guilt for situations outside of one's control: "It is my fault my grandchildren are unhappy."

Mike was a 75-year-old retired, gay identified white male whose long-time partner and husband had died approximately 2 years before initiating treatment. He was referred for care by a close friend who was concerned that he was still grieving the death of his husband and had become isolated.

Mike had grown up in a traditional Christian family on a farm in a small Midwest town, the middle of three siblings (an older brother and younger sister). He reports having a loving mother and distant father, but feeling separated from his parents and siblings when he developed awareness that he was homosexual. He joined the Navy in part to "escape" small town life. Mike had served as a Navy corpsman including deployment on an aircraft carrier during the Vietnam War. He had no direct combat exposure, but was involved in caring for wounded aboard the ship.

Mike found his role rewarding and served in the military approximately 15 years. However, he ended up leaving the military prior to retirement because rumors of his sexual orientation were beginning to spread among his unit. Upon leaving the military he attended a state university in a large city and obtained his bachelor's degree and teaching credential, and had a successful career as a middle school teacher.

Mike had very limited dating or sexual experience prior to or during his military service, fearing consequences for his career. However, while in school he was active in the gay community of his city, dating frequently. He also reports drinking heavily at this time. He met his partner, Bill, and reports that within a year they had developed a committed relationship. Bill was approximately 10 years older than Mike, and Mike describes him as a partner as well as a mentor helping him navigate the "gay scene." Over the next 10 years with Bill's support he was able to disclose his sexual orientation to his parents and siblings, and his brother and sister in particular were supportive and developed a close bond. His parents were initially "disappointed" but eventually "came around."

Mike reported that his relationship with Bill was marred by several periods of heavy drinking during which Mike had anonymous sexual encounters with other men. When Bill found out about these incidents, they had a series of arguments and confrontations that led to Mike attending Alcoholics Anonymous (AA). While they maintained their relationship and married when gay marriage was legalized, and Mike remained sober for the following 30 years, Bill would frequently remind Mike of his infidelities and criticize him.

Approximately 10 years prior to the initiation of CBT, Bill had developed a neurodegenerative disease that quickly limited his physical functioning, and left him bed-ridden for the last 4 years of his life. Mike served as his primary caretaker throughout his decline, retiring from his job as a teacher to take care of Bill. During this time, as Bill developed symptoms of dementia, he would continue to recall Mike's earlier infidelities and remind him of how he had hurt Bill in the past.

When initially presenting for treatment, Mike shared that since Bill had died he continued to feel both extreme sadness at the loss and guilt about his previous behavior. However, he also reported anger at Bill's inability to forgive him, which in turn led to additional guilt. In addition, he reported having "no purpose" now that Bill had passed. Although he and Bill had a large network of friends, he was not interacting with any. While he was still attending AA, he no longer took leadership roles in the organization, sponsored other members or participated in social events. He only spoke to his brother and sister when they called, and had taken to screening their calls to avoid them.

At the beginning of therapy Mike was meeting criteria for major depression, including daily sad mood, anhedonia, decreased energy, pervasive guilt, poor appetite, insomnia, and thoughts of death (he denied suicidal ideation, plan or intent, but thought that the world would be better off without him). He stated he was irritable at AA meetings, and took benign comments as

criticisms directed at him. He also reported he had not cleared out any of Bill's belongings, and that he criticized himself for this as well.

Therapy focused on three areas: Acknowledging the loss of Bill, developing goals for Mike that were consistent with his values, and increasing his engagement with people who cared about him. Based on his history, the therapist developed a case formulation that focused on the core belief "I'm defective." This manifest in secondary assumptions of "I deserve to be punished" and "I can never be forgiven for the harm I've done." This in turn led him to be both very self-critical and to interpret those around him as being critical as well.

In the course of therapy, Mike learned to examine his thoughts, and the relationship between thoughts, actions, and feelings. As an example, he stated that one of his goals was to start cleaning up his home, including donating most of Bill's clothes to charity. However, in session Mike reported that he had been unable to donate the clothes and got upset when he thought about it. The therapist helped Mike to examine his thoughts using the 3C's form: In reviewing the thoughts, Mike was guided in identifying thinking errors, in this case, "should" thoughts. He was then guided through questioning the assumptions underlying the thoughts, to generate alternative, more helpful thoughts:

Therapist: What do you tell yourself when you looked at Bill's clothes?
Mike: If I get rid of them, I'm being unfaithful. I'm betraying him.
T: When you say that, what emotions come up?
M: Shame, guilt . . . and some anger
T: I can see where the shame and guilt come from; how does that thought lead to anger?
M: I get angry at myself for even wanting to give away his clothes. I shouldn't want to.
T: I see. You're telling yourself there's something wrong with you if you want to give his clothes away; "I shouldn't want to do this." So what would be a more helpful thing to say to yourself?
M: I don't know.
T: Well, what is the reason you want to give away the clothes? Is it because you didn't love Bill?
M: No! It's just that it is so hard to see those reminders. I think it would help me move forward.
T: Do you think that Bill would not want you to move on?
M: No, he told me that he wanted me to be happy after he was gone, or at least he did when he wasn't upset with me.
T: So it sounds like the "should" thought is not true; there are good reasons for wanting to give the clothes away, and it makes sense to want to do so. What would be a more helpful thing to say?
M: Maybe that it's ok to give the clothes away; it is doing something for others, and Bill would like that.
T: And what's your evidence that this is true?
M: Well, he told me to move on when he was gone, and he did like donating stuff to charity.
T: Anything else? If the roles were reversed, what would you want Bill to do, if you were gone and your stuff caused him pain?
M: Oh, I'd want him to donate it. I wouldn't want him to hold onto it.
T: So what might be a more useful thing to tell yourself?
M: He'd be ok with me giving his stuff away, because it would be helping others
T: On a scale of 0 to 100, how much do you believe that's true?
M: I would say about an 80; we had talked about this.
T: When you say that to yourself, what happens to your emotions?
Mike: They get better.

In this example, the therapist asked questions to guide Mike in questioning his assumptions, and to notice evidence that alternative thoughts may be true. The therapist then helped Mike formulate a more helpful thought, and tested its usefulness by asking how it affected his feelings, using the 3C's form.

Situation	Catch it	Check it	Change it
Deciding on what to do with Bill's things	"If I get rid of his clothes I'm betraying him" "I shouldn't want to give his things away"	Emotions: Shame (90) Guilt (90) Anger (60)	"Giving away Bill's clothes would be helping others, and he'd like that." Emotions: Shame (20), Guilt (20) Anger (30)

Socratic Questioning and Techniques for Getting at Thoughts. Also called "guided discovery," the overall purpose of Socratic questioning is to challenge accuracy and completeness of thinking. These are open-ended questions phrased so people rarely respond with a "yes" or "no." It is important that the tone of the discussion remain collaborative, and the therapist not betray judgment of the client's interpretations. Socratic questioning may be used at almost any stage of therapy, and may be particularly useful when resistance is detected. There are several different categories of Socratic questioning, defined below with examples:

CONCEPTUAL CLARIFICATION QUESTIONS. These are basic "tell me more" questions that encourage deeper thinking:

- *What exactly does this mean?*
- *How does this relate to what we have been talking about?*
- *What is the nature of . . . ?*
- *What do we already know about this?*
- *Can you give me an example?*
- *Are you saying . . . or . . . ?*

PROBING ASSUMPTIONS. These are directed at uncovering assumptions underlying beliefs. This can be particularly helpful for identifying underlying core schema.

- *You seem to be assuming . . . ?*
- *How did you choose those assumptions?*
- *Please explain why/how . . . ?*
- *How can you verify or disprove that assumption?*
- *What would happen if . . . ?*
- *What would it mean about you if . . . ?*
- *Do you agree or disagree with . . . ?*

PROBING RATIONALE, REASONS, AND EVIDENCE. These uncover reasoning in arguments.

- *Why is that happening?*
- *How do you know this?*
- *Why do you think X causes Y . . . ?*
- *Are these reasons good enough?*
- *Would it stand up in court?*
- *How might it be refuted?*
- *What evidence is there to support what you are saying (or an alternative)?*
- *When did you first decide this was happening?*

QUESTIONING VIEWPOINTS AND PERSPECTIVES. Most arguments are given from a particular position. Show that there are other, equally valid viewpoints.

- *Another way of looking at this is . . . ; does this seem reasonable?*
- *Is there an alternative way of looking at this?*
- *What is the difference between . . . and . . . ?*
- *Why is it better than . . . ?*
- *What are the strengths and weaknesses of . . . ?*
- *How are . . . and . . . similar/different?*
- *What if you compared . . . and . . . ?*
- *How could you look at this another way?*

PROBING IMPLICATIONS AND CONSEQUENCES. The argument given may have logical implications that can be forecast. Do these make sense? Are they desirable?

- *Is . . . a helpful thought? Is . . . helping you reach your goal?*
- *How does this thought affect how you feel or what you do?*
- *Then what would happen?*
- *What are the consequences of that assumption?*
- *What are the implications of . . . ?*
- *How does . . . affect . . . ?*
- *Why is . . . important?*

Mindfulness and Acceptance-Based Interventions. These interventions can be particularly helpful in helping clients to observe their thoughts and emotions, and to also tolerate distressing situations. Key exercises include body scans (where the client learns to notice each part of the body in order, in a nonjudgmental way) and breath-focused meditation (paying attention to breathing without trying to change it). Clients then learn to use the skills of nonjudgmental observation practiced in these exercises in attending to experiences in their everyday life.

Schema Change. CBT also addresses schema, defined as core beliefs and associated cognitive biases that are held by people. Common core beliefs in mood disorders focus on achievement ("I am worthless unless I am productive") or affiliation ("I am unlovable"), while in anxiety disorders they focus on vulnerability ("I can't cope; I am insufficient"). Due to the stigma and multiple failure experiences associated with severe mental illness, people with schizophrenia may have both types of core beliefs. In paranoia, core beliefs may also focus on persecution as an explanation for failure ("poor me") or as a deserved punishment ("bad me"). Schema change involves the other treatment techniques previously defined. In addition, the hypothesized schema is reviewed with the client. The costs

and benefits of the schema are discussed in the context of the current problems and goals, and an alternative schema is derived that is more functional for clients while still being consistent with their values.

Mike frequently expressed the thought "I'm a screw up; I'm not good enough." His therapist noted that Mike seemed to believe that there was something inherently flawed or defective about himself and he deserved to be punished. Mike agreed with this. He noted that intellectually, he knew he was trying to be a good person, but emotionally he always felt there was something innately "broken" about him. Via guided discovery, Mike tied this sense of himself as flawed to struggling with his sexual orientation as a young man, as well as his infidelities with Bill and Bill's lasting anger about them. The therapist suggested to Mike that he consider another perspective; what would the people in his life say they valued most about him—Would they focus on his flaws or his gifts? Mike was able to share that his family and friends had all said how much they respected the love and dedication with which he cared for Bill. He also still received cards from former students sharing how important he had been to them as a teacher. In addition, although Bill frequently criticized him, Bill also was very loving and appreciative. In fact, most of the time, they were very loving toward each other, and Mike realized he tended to focus on the negative times. Finally, Mike and his therapist reflected on the strength and commitment it took to have stayed sober for 30 years, and the support he had provided to others in AA, particularly younger gay men struggling with their own sexual orientation.

Mike was willing to consider a more helpful view of himself, and came up with an alternative self-statement; "I am a caring guy." He developed some new secondary assumptions; "I've done things that I'm ashamed of, but I have worked hard to acknowledge my mistakes and make amends" and "I've been a good influence in a lot of lives."

While Mike agreed that these statements were true, he stated "they don't feel true." Therefore, assignments were designed where Mike would act as if the new formulation was true, and see how it worked for him. This included increasing his contacts with family members and friends, and behaving like he deserved to be happy (e.g., doing enjoyable things). He found contacting people hard initially, but he was greeted with enthusiasm. As he started joining friends for activities, his mood started to improve and he found it easier to act on tasks around the house. He still would miss Bill, but now had more support to cope and friends who expressed understanding about his loss.

Termination and Relapse Prevention. As clients achieve symptom improvement and functional gains, the emphasis of CBT shifts to preparing for termination. Treatment emphasizes clients developing their own interventions and experiencing mastery of skills. In addition, the therapist helps the client identify both persistent problems that need to be managed as well as potential future problems, and then identify the skills that can be used to address those problems. Particularly for older clients, termination can be stressful if they have lost other relationships recently. Therapists need to emphasize the goal of the client developing skills, and the importance of independent practice to adequately learn the skills. However, it is also helpful to schedule follow-up "booster" sessions. This both provides the client with some sense of support and allows a chance to problem solve difficulties in implementation of relapse prevention plans.

Modifications for Psychosis

Psychosis is the most expensive mental disorder to treat, with

hospitalization costs accounting for the majority of expenditures. The cost of care for psychosis patients over 65 years old is even higher than for patients aged 30 to 64. Best practice guidelines and nationalized healthcare systems recommend or mandate that CBT be delivered to consumers with schizophrenia as part of a comprehensive treatment program. For example, CBT is among the recommendations made by the American Psychiatric Association, Department of Veterans Affairs, and the Schizophrenia Patient Outcomes Research Team (PORT), a project sponsored by the National Institute of Mental Health to develop evidence-based treatment guidelines for schizophrenia in the United States. PORT client survey findings, however, indicate poor conformance with this recommendation, especially for older people with psychotic disorders. Providers with expertise in delivering these interventions for older people with psychosis are very much needed.

In principle, all of the stages and techniques of CBT discussed above can be applied to the treatment of older people with psychosis, with some modifications but not fundamental differences. All of the techniques discussed above are appropriate for the treatment of the multiple problems faced by consumers with psychotic disorders. In this section, a brief discussion of the vulnerability-stress model and the general cognitive model of schizophrenia is provided, followed by descriptions of specific CBT interventions applied to psychosis, negative symptoms and poor functioning in older people with schizophrenia.

Cognitive Model of Schizophrenia. In the context of the vulnerability-stress-coping model of schizophrenia, personal enduring vulnerability traits, like cognitive and neurobiological dysfunction, interact with environmental stressors and coping responses to contribute to symptom formation. Neurocognitive deficits exacerbated by stress hyperreactivity can lead to decompensation and the emergence of delusional beliefs and hallucinations, as well as diminished capacity to accurately evaluate anomalous beliefs. People with psychosis, like healthy individuals, have great difficulty distancing themselves from their highly charged beliefs, but also have the added problem of insufficient cognitive resources available to evaluate and correct them. Rigid thinking, perseveration, cognitive biases, impaired abstract reasoning and memory, and cognitive distortions can play an important role in the development of delusional beliefs by leading to exaggerated appraisals of negative life events that become highly charged delusional belief systems.

In addition, multiple failure experiences due to cognitive deficits, odd behavior, and social withdrawal can lead to dysfunctional performance beliefs and social aversion attitudes (e.g., “I am inferior,” “Why bother; I will fail,” “If you can’t do something well, why do it at all?” and “People are against me”), which can exacerbate maladaptive behaviors, such as social

withdrawal and suspicious avoidance. In the context of generic cognitive model applied to schizophrenia, dysfunctional attitudes like these may contribute to negative symptoms and poor functioning. Negative appraisals about one's self and one's ability to perform goal-directed tasks contribute to passivity and avoidance of productive activities, as a defense against anticipated failure and negative evaluations by others. This avoidance and lack of effort manifests as diminished motivation (avolition-apathy), loss of pleasure (anhedonia-asociality), impoverished speech (alogia), and blunted affect. CBT can be used to address and challenge beliefs about psychotic symptoms, social aversion attitudes, and defeatist performance beliefs, including expectancies ("It won't be fun"), self-efficacy beliefs ("I always fail"), and anomalous beliefs ("Spirits will harm me") that interfere with functioning. In addition, by including behavioral interventions, such as social skills training, skills deficits that can also interfere with social functioning can be addressed. By challenging self-defeating beliefs that interfere with skill performance while addressing skill capacity, people are more likely to engage in functional behaviors and use the skills they have.

Psychoeducation for Psychosis. Education is a key component in helping clients with psychotic symptoms make sense of their experiences. Psychoeducation is best accomplished as a two-way process, whereby the clients themselves come up with psychological explanations of their symptoms using guided discovery with inclusion of information from the therapist only where necessary. It is typically helpful to review the history and symptoms at the onset of psychotic illness in the context of the initial stressors that may have triggered it, using reflections like, "So, with all these stressors going on, including your father's death, you suddenly began hearing this voice, and tried to make sense of it." Further, "Given that it seemed that nothing else could explain where this voice was coming from, I can see why you decided it was your father talking." Comments like this emphasize both that data were gathered that led to a conclusion about the source of the voice (an empirical approach) and that there were life events that contributed to the experience. This can set the stage for gathering evidence that might dispute these long-held conclusions, as well as identifying potential relapse triggers for relapse prevention plans. It can be challenging to review the first episode in older clients, when the first episode was often more than 30 years ago. Collateral informants and record reviews can be helpful here, but the most important information is the evidence and context recalled by the client. In older clients, reviewing the impact of recent losses and bereavement (and sometimes anticipated losses or illness of loved ones) on symptom exacerbation can be useful to illustrate the impact of stress on symptoms. In the context of the vulnerability-stress model, it is often useful to focus on "normalization," that involves presenting the idea that any individual is vulnerable to

experiencing psychotic symptoms if exposed to certain circumstances and sufficient stresses (e.g., sleep deprivation, unusual severe stressors, or substance abuse). This "normalization" discussion can serve to reduce stigma associated with severe mental illness. The focus on normalization and open questions and reflections directed at an honest understanding of how the client arrived at delusional beliefs often facilitates rapport and builds a trusting relationship. Time and patience are needed in this process, as trust can be slow to develop with some clients (e.g., with paranoia).

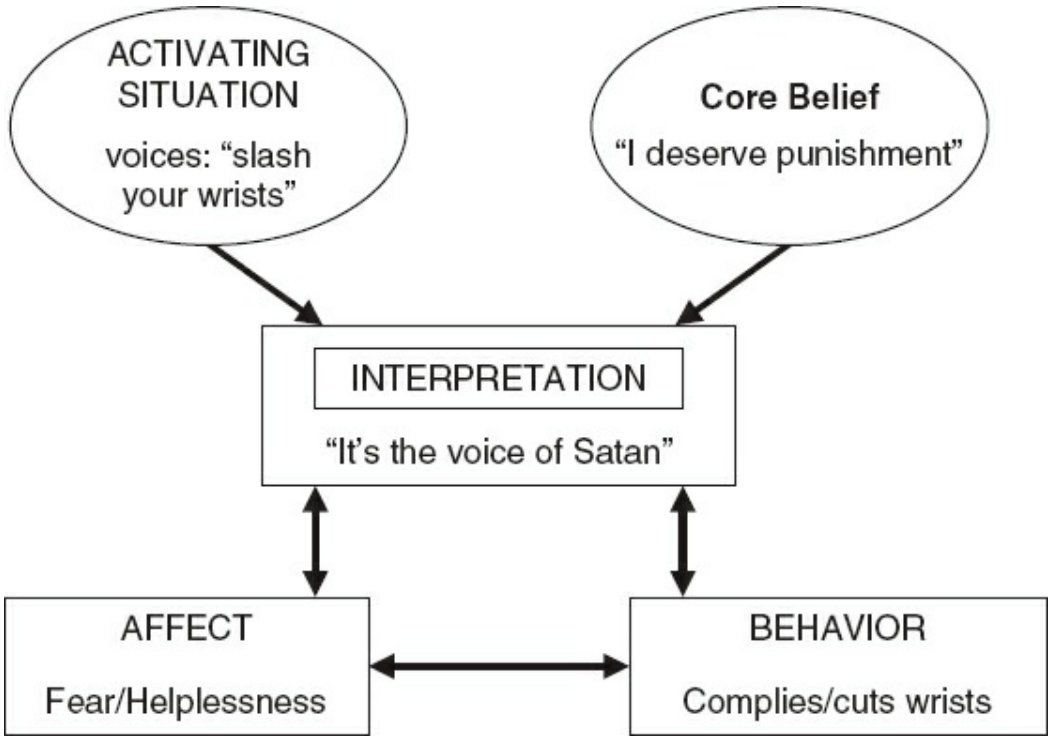

FIGURE 57.4i–2. Generic cognitive model applied to psychosis.

Clients are introduced to the general concepts of CBT, including the relationship between thoughts, behaviors, and feelings (cognitive model). Multimedia presentation of information can help compensate for neurocognitive impairment associated with both psychotic illness and aging. For example, drawing figures, like the one below, on a pad (or handout) for clients to take home can help compensate for cognitive impairment (Fig. 57.4i–2).

This example helps to illustrate the importance of interpretations or thoughts about experiences, such as auditory hallucinations. It is important to stress that it is not the voice itself that leads to unwanted feelings or behaviors, but the thought about the voice. It is more distressing to think Satan is making threatening demands (Satan is a powerful, controlling, external agent) than to think voices are due to brain changes caused by illness. If the interpretation is that voices are due to illness (no power to

harm, controllable, internal origin), then fear, helplessness, and risk of self-harm can be reduced. The core belief ("I deserve punishment") is provided in Figure 57.4i–2 to illustrate the importance of developing a formulation about highly charged beliefs that drive interpretations. It is typically not useful to complicate the model by including this aspect in the psychoeducation process early in therapy, although it may be added later.

Behavioral Interventions. Behavioral activation, pleasure predicting, and relaxation training techniques described above are all relevant to the older client with psychosis, and these techniques are used in essentially the same way. In addition, an important treatment target for older people with schizophrenia or other psychotic disorders is to improve communication skills and psychosocial interactions (e.g., asking someone to share a leisure activity; making positive requests for support from friends, family or professionals; expressing anger and assertiveness). The predominant therapeutic technique used to improve communication skills is the behavioral role play. Role plays are also useful for generating thoughts that can interfere with adaptive social interaction (e.g., "He thinks I'm dumb"; "He won't want to talk to me"; "He wants something from me"). These thoughts can then be targeted for change using thought challenging interventions (e.g., the 3C's). Low self-efficacy and defeatist performance beliefs ("I will forget what to say"; "I will mess it up") can also be identified by using scaling techniques, where the therapist asks clients to rate their expectations for success in the role play on a scale of 1 to 10 before practicing and again rate performance after practice. The therapist can then compare these ratings to point out how people perform better than they think they will (pre-post ratings almost always increase), so clients can raise expectations for success to increase the likelihood that they will use the skills in the real world. In this way, the role play itself can serve as a behavioral experiment to test out defeatist beliefs. By teaching new social skills, while changing the defeatist thoughts that interfere with skill performance, negative symptoms and functioning can improve.

Role plays can focus on a number of topics. Clients are more motivated to learn skills when role play topics are linked to the client's personal goals (e.g., role playing how to ask someone to coffee, if the client's goal is to make a friend). Expressing positive and negative feelings and making a positive request can be emphasized to improve assertive, nonaggressive, clear, and comfortable sharing of feelings in social interactions. Improving everyday leisure activities and psychosocial functioning are common role play topics for these skills (e.g., asking a roommate to change his or her behavior; asking someone to go to the movies). Another common focus of role plays is on identifying warning signs of psychotic symptom exacerbation and practicing reporting these to the doctor. It is sometimes useful to make a video recording of role plays and play back the recording

in session. With the help of the therapist, clients critique the strengths and weaknesses of their own performance, according to a list of key skills displayed on a handout or poster board (e.g., eye contact, speech volume). Clients sometimes have little insight into the impact of their negative symptoms (e.g., low voice volume, poor eye contact, blunted affect) on their social skills until watching themselves on video. The group therapy setting is ideal for role plays. Members with better skills model for others, and this leads both to a sense of mastery for the skilled client and an increased perception that the skill is learnable for the unskilled client. Clients can sometimes accept feedback better from peers than from therapists, and situations played between group members are more similar to in vivo experiences than role plays between therapist and patient.

Given the loss and isolation experienced by older clients, one important focus is to help clients to recognize and identify support persons in their lives. Older clients with psychosis often have a difficult time recognizing that some people in their life truly are support persons, such as health care providers, board-and-care workers, and peers. Often, clients believe that only family or relatives can act as support persons. Because many older clients with schizophrenia no longer have connections with their family or relatives, it becomes increasingly important to identify support persons and help clients to recognize that many others in their life outside of family can act as supports.

Cognitive Interventions. Thought-challenging interventions (e.g., using the 3C's) for clients with psychosis are delivered in essentially the same way as for clients with depression or anxiety. Clients are introduced to the general concepts of CBT, including the relationship between thoughts, actions, and emotion (generic cognitive model); identifying automatic thoughts; and challenging unhelpful thoughts by examining evidence for beliefs and mistakes in thinking (e.g., jumping to conclusions, mind reading, all-or-none thinking). Through in-session discussion and homework assignments, clients are taught to identify thoughts; identify relationships between thoughts, feelings, and behaviors; and identify mistakes in thinking. Thought records are used to elucidate the relationship between situational triggers (e.g., people speaking outside the client's room) and thoughts (e.g., "They've come to hurt me") and subsequent changes in feelings (e.g., fear) or behaviors (e.g., isolation, aggression) in response to thoughts. Thoughts that lead to distressing feelings or maladaptive behaviors are labeled as "unhelpful," and are examined for accuracy. Socratic questioning is used to help clients examine the logic of beliefs and generate more adaptive alternatives to mistakes in thinking or thoughts without sufficient evidence. The primary thoughts targeted are typically beliefs about voices (e.g., "My voices are from an external being," "My voices have power to harm me," "My voices cannot be

controlled"), beliefs about daily events that clients use as evidence to confirm delusions ("That person on the bus was part of the conspiracy"), and dysfunctional performance beliefs (e.g., "Why bother, I will fail," "I am too sick") that interfere with adaptive functioning.

Using the 3C's technique for thought challenging is particularly helpful to compensate for neurocognitive impairment that is common in older clients with severe mental illness. The 3C's technique simplifies learning and serves as a mnemonic device to help clients remember and use cognitive techniques in everyday life. Table 57.4i–3 shows examples of a 3C's thought record for beliefs about psychosis. Thought records like this are completed both in session and as homework assignments. Situational antecedents are typically first linked to the feelings and behaviors that arise in them. This is done to emphasize that clients need to use these unwanted feelings and behaviors as an alert or "red flag" reminder to examine their thinking ("Catch it"). Once the thought linked to the feeling or behavior is identified and written down, possible mistakes in thinking are identified and the evidence for and against the thought is reviewed ("Check it"). If the thought does not "check out" (e.g., there is evidence the thought is unhelpful), it is replaced in favor of a more balanced, more helpful thought ("Change it"). The group setting can facilitate thought challenging, because other members are often better able to notice delusional thinking and articulate logical flaws, contrary evidence, and alternative thoughts than the client, who has an emotional conviction attached to the inaccurate thought.

To gather evidence to evaluate thoughts, behavioral experiments are conducted inside and outside of sessions (homework). For example, one common belief is that a voice is from an external agent that is omnipotent (e.g., the voice of God or the Devil). A behavioral experiment to test this belief might involve training clients in distraction techniques to stop or quiet hallucinations, such as humming, listening to music, shadowing what the voices say, watching TV, or counting. If clients can control the voices to any extent using these techniques (in session if hallucinations are triggered or at home for homework), this is discussed as evidence against the omnipotence of the voices. How can simply humming or turning on the TV control God or Satan? The following therapist–client interaction illustrates a behavioral experiment.

Pat had been experiencing some paranoia that was causing her to isolate at home.

P: I'm really confused because I had so many coincidences happen that I was writing them down. Then my friends will say, "Are you taking your medication?" when I tell them about what's been happening. But then they can't deny that there are so many of them and it is really weird. And then it becomes so weird that mathematically it would be like getting hit by lightning two or three times. It's just that the odds are such long shots that all this would be happening. Then I start thinking that I must be under some kind of mind control and it's preplanned because the odds of this happening couldn't possibly be this strong. So

this has been difficult. This is confusing.

T: What kinds of things have been coincidences in the past few weeks?

P: Just like a matter of 2 days. What I'm thinking about or what I'm doing, the next thing I'll see it in print and or it will be said on the TV. Whatever I'm thinking about, one thing or another after another. Just like, I was thinking about hermaphrodites and people who are male and female and I open my book and it's right there. Talking about androgynous people, male and female people as I was just thinking it, and then several conversations like that came up. What I was talking about or what I was thinking happened right then and there, at that moment, and I keep trying to ignore it 'cause people have said something about free associating or something like that, too much. It just got to be too much of it happening. Either I'm at some really strong path with God or I'm under some kind of mind control. I don't know which.

T: How does that impact your life? How does it get in the way of you obtaining your goals?

P: When I feel my mind controlled, for me, I end up deciding not to move at all. Like the only way is to contain it. Like it's a disease, infection. I just feel diseased, like I have an infection. Like nobody should be near me.

T: Because we might catch it?

P: Yeah or something like that.

T: So we could catch the coincidence disease?

P: Yeah.

T: What would happen to us?

P: You'd be in the same situation.

T: Which is?

P: Being controlled.

T: So we would feel the same ways you're feeling. Well, are you feeling that way right now?

P: Yeah, pretty much.

T: What if I came and sat next to you? Can we try this out and see if I catch what you've got? How about it?

P: I don't think you would know right away.

T: How long would it take me to know?

P: Just until you start noticing those types of coincidences all happening to you.

T: How long would it take me to know?

P: I don't really know.

T: Approximately?

P: I don't know. I'm just guessing how all this works.

T: Well, would you mind? I mean because you're going to be in therapy for several months. So do you think it could happen during that time?

P: Yeah.

The therapist sat next to Pat for the next 2 months and revisited the idea weekly that the "coincidence disease" was contagious. Over time, Pat continued to notice coincidences, but the inferred controlling nature of these occurrences decreased over time, in part, because her conviction in the belief that it was a disease diminished. This is an example of thought chaining—if others don't catch it, maybe it is not a disease, and if I don't have a disease, then the cause of coincidences is not a controlling disease. As conviction in this delusional belief diminished, Pat was willing to try more activities outside the home, which provided additional opportunities for behavioral experiments to test out possible causes of coincidences, with a focus on generating alternative hypotheses.

Table 57.4i–3.
Example of 3C's Thought Record for Beliefs about Psychosis

	Catch It (What's the Thought Going	Change It (What's a More

Situation	Feelings/Behaviors (Consequences of Thoughts)	Through Your Mind?)	Check It (Mistakes In Thinking and Evidence Against the Thought)	Helpful and Accurate Thought?)
On the bus people looking at me	Afraid; angry; Left-got off	"They want to hurt me."	Mind reading; jumping to conclusions; just looking at me doesn't mean they want to hurt me; maybe they think I look like a friend or someone they know; maybe they think I'm attractive	"There are lots of reasons people would look at me on the bus."
While going to fill out a job application, voices say: "You can't do anything right"; "Nobody would ever hire you"	Sad; hopeless; went home; did not apply for the job	"The voices are right; I could never get a job."	All-or-none thinking; predicting the future; the voices have been wrong before; I used to have a job	"Maybe the voices are wrong; I can check it out if I try and see if I get hired."

The following case example illustrates how cognitive-behavioral interventions can be used to address command hallucinations, paranoia, and medication adherence.

Mark is a 61-year-old unmarried Caucasian male who was unemployed and taking an art class at a local community college. He served in the US Army for 6 years from ages 18 to 24, but had not worked since his discharge from active duty. He was not in combat and denied trauma (was trained in a logistics/support role). He was medically discharged from the military for Psychosis NOS and received some Veteran's benefits income. He received care from the local VA facility, including antipsychotic medication, and was diagnosed with schizophrenia, paranoid type.

He lived in a group home in a financially depressed area of his city. He had several psychiatric hospitalizations, often stemming from physical altercations, including assaulting an individual he thought was going to hurt the pastor of his church. He was arrested but not charged in this assault. He reported frequently hearing voices, including command hallucinations, especially when he forgot to take his medications, and ascribed these assaults to the voices ("The voices tell me to stop bad people from hurting other people"). He also experienced difficulty with paranoid ideation, reporting he often thought people were watching and talking about him on public transportation and at his group home.

Mark had a strong identity as an artist, and pursued art consistently for many years since leaving the military. He sculpts using clay in his bedroom, and showed his therapist pictures of his sculptures. However, he reported that the limited space in his room prevented him from working on his art consistently. He also recognized that when his symptoms increased, it interfered with his working on his art. Therefore, the consumer and therapist developed goals increase his opportunity to practice art and better manage his hallucinations and paranoia by learning cognitive-behavioral strategies and improving his medication adherence.

Treatment involved using the 3Cs and behavioral experiments to examine thoughts about the omniscience of his command hallucinations (e.g., "They know who will hurt my friends") and medications (e.g., "It's no big deal if I forget my meds"). For example, he would Catch thoughts about the omniscience of his voices ("They can tell who has evil in them"), Check the evidence against the thought ("They were wrong about the guy they said would hurt my pastor") and the consequences ("I get arrested when I listen to them"), and Change the thought to be more realistic and helpful ("Maybe the voices are wrong"). After his conviction about the

omniscience of the voices was reduced, he was more willing to ignore them and not comply with their commands to protect others in order to test whether they their threats would come true. By not complying, he was conducting a behavioral experiment that produced further evidence that he did not have to do what they say to protect himself or others. He also reported that his voices were from "good spirits; maybe angels" who would warn him about "people with evil in them." The fact that he thought the voices were benevolent and might protect him from harm contributed to poor medication adherence, because he did not want to take medications that quieted or eliminated the voices that he perceived as helpful. The therapist used Socratic questioning to challenge these beliefs about the voices as a benevolent and omniscient external agent:

T: You mentioned that you have noticed that the medications can make the voices go away, right?
M : Yes, I have noticed that.
T: That makes me a little confused, though. How can something you put in your body—medications that change your brain chemistry—make angels or something outside of your body stop talking?
M: Well, I'm not sure but it does.
T: But does it make sense that a chemical inside you could change an angel or some kind of powerful supernatural spirit? How does that work?
M: I don't know but I get nervous when it's been a while since the voice has given me any warnings or told me who to watch out for.
T: I can see how you might want to stop taking the medication for a couple days when you get nervous about that, so the voices might come back, but does it make sense that medications would make angels or spirits come and go. What if the voice isn't from an angel or spirit? What if it is from inside you and not an outside spirit? That would help explain why changing your brain chemistry can stop it. What do you think about that possibility?
M: Well, that part adds up, I guess, but I'm just not sure.
T: And we already discovered that the voices are not always correct about who is dangerous. That also makes me think the voices are from you rather than an angel, because people are more likely to be wrong about things than angels. We all make mistakes, but angels probably don't.

CLINICAL ISSUES

CBT and Medications

The use of CBT forms one part of a package of care that often includes the use of medications, where relevant and appropriate, as well as other services (e.g., case management, vocational rehabilitation, couples/family therapy). This is particularly true for psychosis. For depression, meta-analyses of CBT compared to antidepressants indicate that effects are similar (d = 0.35), and that response to CBT among older adults is similar to effects seen in younger adults. A recent meta-analysis including head-to-head comparisons of antidepressants and CBT in older adults found no differences between treatments, but the authors conclude the number and quality of studies was too low to make definitive conclusions. For late life GAD, results are further mixed. Some studies indicate that CBT is as effective as antidepressant medications for late life GAD, but others contend that CBT is not an effective treatment at all. Although older adults report a preference for talk therapies over medications, the two are often prescribed together. This generally has meant that cognitive therapy is used

alongside medication in a synergistic way, such that CBT improves medication adherence and medication controls acute symptoms that might interfere with CBT. It is recommended that, if the CBT therapist is not the psychiatrist, a two-way release of information be obtained from each client in order to facilitate continuity of care and communication.

Individual versus Group Format

There are several advantages to using these interventions in a group format. Groups are widely used in mental health clinics (increasing the likelihood that these interventions will be acceptable to such sites); treatment can be provided for a greater number of clients in groups than could be accomplished in individual sessions; groups can impact social support systems and allow practice of communication skills; groups provide a shared identity with others (i.e., healing power of connections with others, "me too" and "not alone" influences); confronting or observing others in groups may provide better insight into one's own symptoms and skills deficits; and groups can provide social reinforcement for change. One important difference between group and individual CBT formats is that groups tend to emphasize skills training over detailed case formulation. There is less time in groups to fully explore the unique content and history of each client's presenting concerns in developing a formulation. There is also greater flexibility to tailor therapy to the individual needs of the client in individual therapy.

Goal-Setting Challenges

Setting goals for therapy with older clients who have been ill for decades can present unique challenges encountered less frequently with more recent onset clients. Particularly for older clients who have had a chronic or recurrent course of a disorder, there may be low expectations for treatment success. Information-processing biases also often lead to a focus on failures and ignoring past achievements. By acknowledging a goal, older clients must accept the risk of yet another failure. This risk assessment is often exaggerated into risk of symptom exacerbation and rehospitalization. For patients with psychosis in particular, by this stage of illness, older people have also been informed that they have a severe, chronic brain disease with little hope of recovery. Older clients have had much less exposure to modern recovery and psychosocial rehabilitation models that have only recently begun to transform health care systems to focus more on strengths and hope for functional achievement. Hopelessness in providers can be passed on to older chronic clients. While not as potentially profound as with psychosis, other depressive and anxiety disorders that have followed a chronic course can also facilitate hopelessness for change.

These factors can create challenges for eliciting a goal for therapy.

Setting a treatment goal may require patience and many sessions. Cognitive therapy tools can be used to address dysfunctional performance beliefs, negative expectancies, and ignoring the positive and other attributional styles that can lead to hopelessness and reluctance to set a meaningful goal. Older clients may also initially offer a goal that they have learned providers want to hear (e.g., "I want to take my medications more regularly"), rather than a more personally meaningful goal. Goals may need to be revised as more meaningful goals emerge. It may be useful to ask clients to recall a time when they were functioning better in the past or imagine a time years in the future when they are much better and ask about daily behaviors. Ask, "What would you be doing if you weren't feeling this way?" This can stimulate a helpful discussion about desired, meaningful functional behaviors (relationships, volunteer work, leisure activities). The key is to get a personally meaningful goal.

We have found it most helpful to focus on specific, functional behavioral goals (increasing socialization, relationships, leisure activities, living skills to promote independent housing, school, work/volunteering) chosen by clients. Emotional or symptom goals (e.g., "I want to stop being sad" or "I don't want to hear voices") are less helpful, unless framed in terms of their impact on functional behaviors. For example, one can ask, "How would you know you weren't sad anymore? What would you be doing?" The response is often something like, "I would be getting up and doing more things." The therapist can then ask what things, and frame the functional behavior as the goal (e.g., doing things with friends, taking a class). Goals should also be specific and stated as attainable behaviors. For example, the therapist can ask, "How will you know the goal has been achieved? What will you be doing differently each day if you achieve the goal?" The therapist's goal is to help older clients set meaningful, realistic, attainable, specific goals stated in terms of functional behaviors.

Importance of Homework

One goal of CBT is to provide skills that generalize and are used by the client outside therapy in everyday life. To that end, treatment should include both skill practice in sessions as well as homework assignments. In-session practice is a crucial means by which clients understand the sometimes difficult concepts in this intervention. Homework assignments are crucial for generalizing these skills to the "real world," but homework adherence can sometimes be challenging. The term "at-home practice" is often used, rather than "homework," to emphasize the importance of practice outside of sessions and to reduce the negative connotation associated with "homework" assignments back in school days. Adherence to at-home assignments can vary considerably between clients. Older clients in particular may have expectations that the active treatment factor in therapy is disclosure rather than practice, and this can interfere with

willingness to complete outside assignments. To facilitate homework completion, the therapist guides, instructs, and explains the rationale for completing home activities, and activities are practiced in session. The therapist goes over several examples, so that clients can more easily do the remainder of the assignment at home. An analogy with learning to play a musical instrument can be used to stress the importance of practicing skills (e.g., "Just like learning to play the guitar: If you only practice when you are at your music lesson, you won't learn songs as fast as if you practice in between lessons"). Principles of operant conditioning are also used to facilitate homework completion (e.g., positive attention and praise to those who complete homework). The therapist makes clear to clients that homework completion is expected and failure to complete assignments is the topic of thorough problem solving to help the client do a better job next time. The authors have had success in obtaining 80 percent homework completion in 80 percent of older clients with schizophrenia in their CBT for psychosis groups. Previous research with other populations has shown that CBT incorporating homework assignments helps clients improve about 60 percent more than those in treatment without homework.

Indications

CBT treatment targets may change as people age. There is evidence that CBT is efficacious for a range of disorders in older adults, including depressive disorders, and psychosis. In addition, while there are a number of pharmacologic treatment options available, clients often desire psychotherapy options. Aging in people with schizophrenia is typically associated with improvement in positive symptoms and reduced hospitalization, whereas approximately 60 percent of older people with schizophrenia still reside in assisted care settings (e.g., board-and-care homes). Although positive symptoms in these older clients with severe mental illness may often be fairly well controlled through pharmacological treatment, these clients still have negative symptoms and severe psychosocial functioning deficits, which are important treatment targets of CBT in older people with psychosis.

Contraindications

A number of factors may interfere with psychotherapy overall. For all older adult clients, it is important to assess the possible causes of many psychological symptoms, such as fatigue, appetite, and sleep disturbance, because medical comorbidity and normal age-related changes may contribute. Comorbidities, such as substance abuse and medical illness, may interfere with attendance in CBT. This barrier can be overcome by conducting therapy in the client's home or senior center in the neighborhood or in the presence of a caregiver, or by finding a resource for

transportation. In addition, CBT interventions can include comorbidities as treatment targets (e.g., reducing substance use through harm reduction techniques, improving health status through addressing obstacles to treatment adherence). Poor insight (awareness of having a mental illness that requires treatment) is prevalent in some psychotic disorders and predicts poor course of illness and functioning. There is also some evidence that insight moderates treatment outcome in CBT for psychosis. Severe disorganization in psychotic clients presents a number of challenges to the logical, hypothesis-testing process of cognitive therapy, but some of these clients succeed despite thought disorder, and there are no clear data on the impact of this potential barrier. The authors have found that disorganization frequently fluctuates, as well, so a more behavioral intervention focus can be used until clients are more able to communicate more linearly. Neurocognitive impairment is often assumed to interfere with CBT, but this assumption may not always be accurate. Progressive dementia will likely interfere with CBT and neurocognitive impairment does predict poorer outcome in serious mental illness, in general, as well as treatment outcome in psychosocial interventions. However, there is no evidence that neurocognitive impairment reduces the benefit of CBT in serious mental illness relative to standard pharmacological care alone. That is, even participants with moderately severe neurocognitive impairment show added benefit from receiving CBT plus pharmacotherapy, relative to receiving only pharmacotherapy. With little clear evidence about which clients will and won't benefit from CBT, the authors recommend that willing clients be given a chance to succeed. The biggest barrier to success is provider hopelessness; clients who are never referred for CBT because they are thought to be too old or too impaired in some way can never succeed.

Cognitive Therapy versus Cognitive Remediation

Two broad categories of cognitive interventions have been developed: Cognitive remediation, which focuses on rehabilitation of specific cognitive functions (e.g., memory, attention), and CBT, which trains participants to use metacognitive strategies to identify and challenge unhelpful or inaccurate thoughts. These areas of intervention have several factors in common and may lie on a continuum from more "molecular" to more "molar" cognitive rehabilitation approaches, but they are often incorrectly equated. At one end, cognitive remediation targets specific problems of neurocognitive processing, and at the other end, CBT targets more global cognitive constructs, such as thinking and self-concept. Clients with severe mental illness typically have neurocognitive impairment and may benefit from cognitive remediation interventions and/or medications that attempt to improve neurocognition. Combining these interventions may enhance outcome in CBT, but cognitive remediation is not cognitive therapy.

RESEARCH AND EVALUATION

Depression

Recent studies indicate that CBT is as effective as pharmacotherapy for treatment of major depression. A meta-analysis of treatments for late-life depressive disorders found that pharmacotherapy and psychotherapy (most often, CBT) were similarly effective in terms of symptom reduction. There is currently little evidence of differential effectiveness of pharmacotherapy versus psychotherapy in the treatment of major depression and related mood disorders, and therefore recommendations for treatment should be based on factors such as patient preference and possible medical implications. In particular, older clients may have more difficulty with side effects of medications or have concerns about interactions with other medications, making CBT a particularly appropriate option. Finally, treatment preference studies overwhelming suggest that older adults prefer talk therapies over medications, so these treatments should be accessible in settings that serve older adults.

Anxiety

There are fewer well-designed studies available of CBT for late-life anxiety disorders, especially trials comparing CBT with pharmacotherapy for late-life anxiety disorders, and the combination of medication and psychotherapy has not yet been formally tested in an adequate clinical trial. Most randomized, controlled trials have compared CBT to other psychotherapies or control conditions, and most have examined generalized anxiety disorder or else have enrolled clients with a range of anxiety diagnoses without differentiating between them. The minimal amount of available evidence does suggest, however, that medications, particularly selective serotonin reuptake inhibitors, appear to be more effective than psychotherapy as an acute treatment for late-life anxiety disorders. In general, results of psychotherapy trials also suggest that CBT has some benefit, although clients do not on average reach a symptom-free status. CBT is superior to wait list and usual care, and not significantly different from other active treatments such as relaxation alone or supportive therapy. Cognitive-behavioral interventions that include relaxation training appear to be more effective than those without. For older individuals who refuse to take medications or for whom they are contraindicated, relaxation training may be the most effective component of a behavioral treatment, but direct comparisons have not been conducted. At this point, there is need for better controlled trials of CBT for specific anxiety disorders, and in particular trials that compare CBT with pharmacotherapy as well as trials examining combined pharmacotherapy and psychotherapy, especially in older clients. Nonetheless, the current literature indicates that CBT is a potentially useful approach for older

clients with anxiety disorders.

Schizophrenia

Only a handful of studies have examined any form of psychotherapy interventions for older people with schizophrenia. Over 50 clinical trials of CBT for psychosis have been conducted with middle-aged and younger people with schizophrenia. These clinical trials found significant improvements in hallucinations and delusions (typically the primary treatment target), community functioning, negative symptoms, social anxiety, and mood. Available clinical trials, however, have concentrated on the adult population up to 65, so their generalizability to older groups needs to be considered cautiously. The authors developed, manualized, and are conducting randomized, clinical trials of a novel psychosocial rehabilitation intervention called cognitive-behavioral social skills training (CBSST) in older people with schizophrenia (mean age 58). CBSST combines techniques from both CBT and social skills training to target the multidimensional deficits that contribute to disability in aging people with schizophrenia. In two separate clinical trials, the authors have found significant improvements in social skills and community functioning in CBSST relative to standard care and relative to a goal-focused supportive contact control group, and these gains in functioning were maintained at 1-year post-treatment follow-up. Based on the evidence for the efficacy of CBSST for older people with schizophrenia, CBSST is registered in the National Registry of Evidence-based Programs and Practices of the Substance Abuse and Mental Health Services Administration (nrepp.samhsa.gov).

References

*Ayers CR, Sorrell JT, Thorp SR, Wetherell JL. Evidence-based psychological treatments for late-life anxiety. *Psychol Aging*. 2007;22:8–17.

Beck AT. *Cognitive Therapy and the Emotional Disorders*. New York: International Universities Press; 1976.

Burns DD. *The Feeling Good Handbook: Using the New Mood Therapy in Everyday Life*. New York: Morrow; 1989.

*Cuijpers P, Karyotaki E, Pot A, Park M, Reynolds C 3rd. Managing depression in older age: psychological interventions. *Maturitas*. 2014;79:160–169.

Cuijpers P, van Straten A, Smit F, Andersson G. Is psychotherapy for depression equally effective in younger and older adults? A meta-regression analysis. *Int Psychogeriatr*. 2009;21(1):16–24.

Cuijpers P, van Straten A, Warmerdam L, Andersson G. Psychotherapy versus the combination of psychotherapy and pharmacotherapy in the treatment of depression: a meta-analysis. *Depress Anxiety*. 2009;26(3):279–288.

Dwight Johnson M, Apesoa-Varano C, Hay J, Unutzer J, Hinton L. Depression treatment preferences of older white and Mexican origin men. *Gen Hosp Psychiatry*. 2013;35:59–65.

Gatz M, Fiske A, Fox LS, Kaskie B, Kasl-Godley JE. Empirically validated psychological treatments for older adults. *J Ment Health Ageing*. 1998;4:9.

Gould RL, Coulson MC, Howard RJ. Cognitive behavioral therapy for depression in older people: a meta-analysis and meta-regression of randomized controlled trials. *J Am Geriatr Soc*. 2012;60:1817–1830.

Gould RL, Coulson MC, Howard RJ. Efficacy of cognitive behavioral therapy for anxiety disorders in older people: a meta-analysis and meta-regression of randomized controlled trials. *J Am Geriatr Soc*. 2012;60:218–229.

Granholm E, Holden J, Link PC, McQuaid JR, Jeste DV. Randomized controlled trial of cognitive behavioral social skills training for older consumers with schizophrenia: Defeatist performance attitudes and functional outcome. *Amer J Geriatr Psychiatry*. 2013;2:251–262.

*Granholm E, Holden J, McQuaid JR. *Cognitive Behavioral Social Skills Training (CBSST) for Schizophrenia: A Practical Guide and Consumer Workbook*. New York: The Guilford Press; 2015.

Gum AM, Arean PA, Hunkeler E, et al. Depression treatment preferences in older primary care patients. *Gerontologist*. 2006;46(1):14–22.

Jayasekara R, Procter N, Harrison J, et al. Cognitive behavioural therapy for older adults with depression: a review. *J Ment Health*. 2014;30:1.

Kingdon DG, Turkington D. *Cognitive Therapy of Schizophrenia*. New York: The Guilford Press; 2005.

*Laidlaw K, Thompson LW, Dick-Siskin L, Gallagher-Thompson D. *Cognitive Behaviour Therapy with Older People*. UK: Wiley and Sons; 2003.

Secker DL, Kazantis N, Pachana N. Cognitive behavioral therapy for older adults: Practical guidelines for adapting therapy structure. *J Rat Emot Cogn Behav Ther*. 2004;22:93–109.

Stanley MA, Bush AL, Camp ME, et al. Older adults' preferences for religion/spirituality in treatment for anxiety and depression. *Aging Ment Health*. 2011;15:334.

Thase ME, Kingdon D, Turkington D. The promise of cognitive behavior therapy for treatment of severe mental disorders: a review of recent developments. *World Psychiatry*. 2014;13:244–250.

*Thompson LW, ed. *Handbook of Behavioral and Cognitive Therapies with Older Adults*. New York: The Guilford Press; 2008.

Thorp SR, Ayers CR, Nuevo R, Stoddard JA, Sorrell JT, Wetherell JL. Meta-analysis comparing different behavioral treatments for late-life anxiety. *Amer J Geriatr Psychiatry*. 2009;17:105–115.

Wetherell JL, Gatz M, Craske MG. Treatment of generalized anxiety disorder in older adults. *J Consult Clin Psychol*. 2003;71:31–40.

Wolitzky-Taylor KB, Castriotta N, Lenze EJ, Stanley MA, Craske MG. Anxiety disorders in older adults: a comprehensive review. *Depress Anxiety*. 2010;27:190–211.

57.4j Family Intervention and Therapy with Older Adults

CAROL ANN PODGORSKI, PH.D., LAUREN N. DECAPORALE-RYAN, PH.D., ANN ELIZABETH CORNELL, PSY.D., AND DEBORAH A. KING, PH.D.

INTRODUCTION AND DEFINITION

Over the past decade demographic trends have altered the composition of the older adult population in substantive ways. The older adult profile began to change in 2011, when the Baby Boom generation, currently one-quarter of the US population, began to turn 65. Demographers report that older adults are now more diverse in some meaningful ways. Older adults are now comprised of two generational cohorts—the "traditionalists" born before 1946, and the "Baby Boomers" born between 1946 and 1964. These groups differ in terms of experiences, values, attitudes, and preferences. The Baby Boom generation is likely to have families that are more geographically dispersed; experienced more divorces; had fewer biological children; amassed more blended families; and are increasingly living in multi-generational households. Baby Boomers are more likely to be healthier, but with the increased longevity of "the traditionalists," 70 percent of those over 65 have functional limitations. Similar to previous generations, their needs are met predominantly by informal supports consisting of families and friends.

The infiltration of Baby Boomers into the over-65 population has also triggered shifts in health care preferences and in relationships with medical professionals. This information-seeking, authority-challenging generation was instrumental in forming and shaping the movement toward consumer-directed health care. In addition, the newest older adults are more accepting of the notion of "mental health" and less resistant to seeking mental health care. Thus, the demand for mental health services among older adults is increasing, and the nature of the older adult patient–mental health clinician relationship will continue to evolve to reflect multiple generational influences.

Other changes in health care delivery approaches have increased the need for clinicians to hone their skills for working with families. The Patient Protection and Affordable Care Act, passed in 2010, has placed a greater emphasis on integrated care through establishment of federally qualified patient-centered medical homes. According to the Agency for Healthcare Research and Quality, "integrated behavioral health care" is the care a patient experiences as a result of a team of primary care and behavioral health clinicians, working together with patients and families, using a systematic and cost-effective approach to provide patient-centered care for a defined population. Another widely adopted health care practice affecting health care delivery is the emergence of patient- and family--centered care. In 2012, influential organizations, including the Institute of Medicine and The Joint Commission, adopted patient-centered care as a fundamental aim toward enhancing quality of care. An outgrowth of patient-centered care has been the emergence of family-centered care. The primary goal of patient- and family-centered care is to promote patients and families as active participants in care and decision making by treating

them with respect and dignity; sharing complete and unbiased information in ways that are useful and affirming; and encouraging their participation and collaboration. Inherent in these principles is the need for clinicians to be skilled in building and navigating relationships within families, as well as among families and health and social service providers. The adoption of integrated health care and patient- and family-centered care have created a shift in health care delivery from *serving* patients to *partnering* with patients and families.

Older adults today often have modified extended families as byproducts of single-parenthood, multiple marriages, and blended families. While families may span great geographical distances, they nevertheless remain able to mobilize support between and among elders, adult offspring, and other family members as needed. When elders do not have relatives nearby, they often form intimate social bonds with friends, neighbors, or others who are not related by blood or legal ties. Thus, an appropriate definition of the aging "family" must be flexibly inclusive.

Family intervention and family therapy are not one and the same. Family intervention is the process of bringing together two or more family members to address problems causing distress or impairing function. In late-life families, family intervention involves educating about the unique medical and psychosocial challenges of aging and their effects on the older adult and family. Family intervention also may involve facilitating communication among family members regarding age-related concerns, but it does not involve intentionally changing the way a family functions. In contrast, family therapy is an intervention that brings family members together to alter the family structure or functioning, thus facilitating adaptation to the challenges of aging. Family therapy is considered when assessment reveals problems in the family's pattern of organizing itself or interacting in response to age-related concerns, and when family members explicitly agree to engage in a deliberate process of change. The decision to engage in family therapy typically reflects an acknowledgment that the family's natural adaptive capabilities are overwhelmed by the nature or intensity of challenges or perhaps that the family does not have the knowledge or skills necessary to deal with the situation at hand.

HISTORY

Family therapy with older adults has its roots in two distinct but overlapping theoretical approaches to the understanding of human behavior. First, family systems theory maintains that the family functions as a network of interconnected elements so that change in any one individual's status or behavior will affect that of others in the family. Accordingly, there is a natural tendency for the family system to attempt to maintain balance or homeostasis. Second, family life cycle theory views the family system as moving as a unit through time, encountering both

predictable, normative transitions and unique, family-specific challenges. Life events such as a change in health or functional status, loss of a loved one, or remarriage of a parent, challenges the family's natural homeostatic mechanisms and requires a dynamic process of adaptation in order to reestablish equilibrium. The clinician working with aging families interprets the behavioral or emotional problems of an older adult as indications that the family system is struggling with transitions or crises associated with the life cycle stage. The therapist's task is to help family members organize or reorganize their resources and family roles in a manner consistent with the life cycle challenge they are confronting.

Although much of family developmental and systems theories has focused on the first half of the family life cycle, more recent depictions consider the structural and cultural heterogeneity of aging families. Life cycle theorists now consider late-life families as being comprised of blended families, families with single parents or grandparents raising children, and families of gay, lesbian, and transgendered individuals. Moreover, life cycle theories have started to include description and differentiation of developmental tasks during the second half of life, including adjustment to retirement, attainment of grandparenthood, coping with loss and illness, and passing on family legacy. Recent theoretical accounts also address stages of family reorganization when its elders decline due to progressive illnesses. One such account contends that adult children undergo predictable stages of denial, gradual acceptance, and eventual assumption of power over their parents' affairs. Such new models help therapists frame interventions that are appropriate to the tasks associated with a family's needs for adaptation.

THEORETICAL ISSUES

A variety of specific approaches to family intervention and therapy with older adults have their foundations in systemic and life cycle themes. The contextual approach emphasizes family ethics and seeks to help families balance the intergenerational "ledger" so that patterns of caregiving are congruent with the family's history and resources. The developmental-interactional model categorizes late-life family problems as situational, transitional, or chronic, integrating life cycle and structural-strategic approaches to help the family frame solutions to the problems in a manner consistent with their unique history and style. Recent approaches to grief therapy combine elements of family systemic, symbolic interactional, and narrative approaches to facilitate family storytelling, or "meaning making," thus strengthening family ties among the living as well as helping families redefine their relationships with those who have died. Cognitive behavioral therapy approaches, such as problem-solving therapy, are increasingly being used with older adult couples, caregivers, and families alike.

Strength-Vulnerability Model of Health and Illness

The strength-vulnerability model postulates that the mental health of individual family members is determined over time by the interaction of individual and family life cycle processes. Although not empirically validated, this clinical approach builds on systemic and life cycle theories, as well as research linking health outcomes to a variety of personality and psychosocial factors. Specifically, the model focuses on strengths and vulnerabilities at the level of the individual older adult (e.g., biological health, personal themes, and individual life history) as these interact with health-enhancing and risk factors at the larger levels of family and other social systems (e.g., family events, themes, relational processes).

According to this theory, attachment and caregiving form the bond of affection between family members that must be present as a foundation for the development of other interpersonal processes in the family. Communication is the development of shared meaning between family members. This process includes sharing important themes and content, as well as constructive communication of affect. Shared problem solving is the ability to develop solutions to problems and carry them out cooperatively. Implicitly or explicitly, this process involves the delineation of family roles as they pertain to decision making and authority. Mutuality involves a long-term commitment to the relationship and the dynamic ability to reinvent the relationship over time in the face of life cycle transitions. These processes are thought to develop sequentially in families, so that some element of attachment and caregiving is necessary for effective communication. Likewise, some effective communication is needed to successfully and collaboratively solve problems. Finally, effective problem solving is a prerequisite to the ability to reorganize and reinvent family relationships with mutual understanding through periods of crisis and life change. The epigenetic pattern is actually circular and dynamic, involving periodic reworking of various relational abilities, rather than being a linear progression through a series of discrete stages. Thus, if late-life transitions and crises are met with successful mutuality, attachment and caregiving will take a new and strengthened form. With unsatisfactory resolution of change, the bedrock of attachment and caregiving will be undermined.

TECHNIQUES

Family therapists are tasked with developing interventions that are congruent with the family's stages of relational and life cycle development. Thus, a detailed family assessment is the foundation upon which family interventions are developed.

Assessment

During the initial sessions, the therapist evaluates the family's unique

strengths and vulnerabilities, including at a minimum the following key points which can be easily remembered using the acronym LIFE: **Life cycle transitions** faced by the family and by individual family members—for example, raising children, job demands, financial stressors, retirement, grandparenthood, loss of spouse, or onset of chronic or degenerative illness; **Individual roles** assumed by family members—for example, caregiver, breadwinner, matriarch or patriarch, and scapegoat; **Family relational processes**—that is, the quality of attachment or caregiving bonds, communication, and problem solving; and **Ethnic or cultural factors**, including spiritual beliefs and practices impacting the family's functioning.

Levels of Intervention

The family assessment determines whether the intervention will most appropriately be psychoeducational, with a focus on providing information pertinent to a problem or illness, or therapeutic, with a goal of changing the family's structure and quality of relational functioning. Several levels of therapy are considered, each reflecting work at successively deeper strata of family relational functioning. These levels serve as general guidelines for considering the type and intensity of treatment appropriate for a particular family.

Level One: Psychoeducational Intervention and Community Referral. Families who exhibit solid attachment bonds, open communication, and good problem-solving skills do not require ongoing therapy but, rather, benefit from brief education regarding a specific late-life issue. For example, when elders develop new medical disorders, such as Alzheimer disease, families benefit from information specific to the disorder and its management as well as referral to community agencies and support groups. Families can usually organize themselves for effective problem solving regarding the care and support of the older adult. Despite their potential benefit, only a small percentage of families avail themselves of these resources. In the case of chronic, deteriorating illnesses, however, the family may benefit from brief, intermittent contact with the therapist at those transition points when the disease worsens or they acknowledge readiness to face the next challenge. Such families typically do not require regular, ongoing professional support.

Level Two: Restoring Effective Problem Solving. Families often have solid, positive bonds of attachment with good communication skills and histories of effective problem solving, yet they temporarily are overwhelmed by the intensity and unfamiliarity of a late-life problem, such as severe illness, functional decline, or the approaching death of a previously capable family member. In these cases, effective problem solving

may be temporarily stymied because of paralyzing grief reactions or sudden disequilibrium in the previously stable family structure. The family may need to reorganize lines of authority and decision making to bring them into accordance with the demands of a later-life transition. For example, families typically go through a natural transition during which the younger generation gradually assumes more of a role in an elder's affairs. This reciprocal process requires not only commitment and involvement on the part of the younger family members, but also willingness on the part of the elders to allow this transition to occur without becoming overly demanding, resentful, or needy. Therapists recognize that the history of parental authority and caregiving continues to influence intergenerational dynamics well beyond the actual parenting years.

One common example of family difficulty at this level is when an elderly matriarch or patriarch experiences a decline in health or cognitive capacity, throwing the surviving family members into a state of disequilibrium. The spouse and adult offspring in this situation are unaccustomed to assuming decision-making authority and typically need assistance in reorganizing family roles. After receiving clear and direct information about the status of the elder, reflecting on the meaning that they as well as the elder assign to this situation, and having an opportunity to grieve these changes, adaptive adult children at this level of relational development can often mobilize their resources to solve family problems.

Level Three: Improving Patterns of Communication. Some families who care deeply about each other and have solid bonds of attachment communicate little about serious problems. Some may have a history of arguments and resentment stemming from past crises. Some family rules prohibit the exchange of certain emotions or information. Others have been able to function with little overt communication in earlier stages of the life cycle, because they were not confronted with major loss or illness. When faced with challenges of aging, they have difficulty developing a common understanding of the problem and thus cannot work collaboratively to find and carry out solutions. Families at this level benefit from active modeling of open, positive communication. Basic communication techniques, such as active or reflective listening, can greatly benefit these families, as can clear and consistent modeling of how to talk openly but respectfully about the challenges before them.

QUESTIONS FOR THE PATIENT

Communication and Problem Solving. How is information about important events shared in your family? Who typically participates in making decisions? Is one person responsible for making a final decision? Is anyone excluded from important family decisions? How do family members resolve conflict when it occurs?

Family Roles and Expectations. What role does the older adult play in the family? Who are the individuals responsible for caring for the older adult? Who if anyone is primarily responsible for providing care? Does anyone live with the older adult? How much time in person, on the phone or otherwise is each person able to contribute to the older adult's care? Who if anyone is responsible for helping the older adult with his/her activities of daily living (ADLs)/instrumental activities of daily living (IADSs)? Are you aware of local resources that may be able to help the older adult? What have your experiences with the medical system been like in the past? What are your expectations for the older adults care? What do you know about his/her end-of-life wishes? (if the older adult is unable to express these himself/herself).

Level Four: Improving Bonds of Attachment and Caregiving. Families with negative, weak, or ambivalent bonds of attachment typically have a history of conflict, marginal functioning, or insufficient nurturance between family members. In some cases, adult children who were abused or neglected as children struggle to take care of the one they perceive as a former abuser. Such families are at risk for elder abuse, and some may require referral to professional caregivers or long-term care facilities. In many instances, however, techniques including family life review or contextual family therapy interventions can be used to induce empathy for the older adult's position, thereby decreasing negativity, increasing feelings of safety, and making it possible to forge more positive bonds of affection between the generations.

Family life review is the process of bringing multiple generations together to weave a coherent intergenerational family story. Typically, this process is facilitated by constructing a family genogram, which is a graphic, multigenerational picture (family tree) that encapsulates family history, structure, and relationships over time. The therapist encourages the family to begin with the story of the elders' families of origin, gradually working down through the themes and experiences of subsequent generations. Learning about the elders' upbringing, cultural context, and early family circumstances often gives younger family members a different, more empathic perspective. Some offspring or grandchildren then can establish more mature "adult-to-adult" relationships with their elders, rather than continuing to relate solely from the vantage point of children whose needs for nurturance and care were unmet.

CLINICAL ISSUES

Clinical Applications across Health Care Settings

The strength-vulnerability model of health and illness is an appropriate framework to use for conceptualizing and delivering patient- and family-

centered care to families with older adults in multiple therapeutic contexts. Applications of this model across health care settings in which mental health clinicians practice are described here.

Outpatient Psychiatry

The majority of older adults with behavioral health symptoms receive mental health services through some type of outpatient setting. Outpatient services are provided in a variety of settings, including private practices, community mental health centers, hospital-based clinics, or programs embedded within nonprofit organizations. Geriatric specialty clinics offer a range of services including psychotherapy, social work, and case management which can help to support the work of mental health clinicians. Many older adults gain access to psychiatry through consultation models that exist in some integrated care settings. In this model, primary care physicians and psychiatrists are colocated, and physicians are able to consult with a psychiatrist onsite, or they can refer patients within the practice to see the psychiatrist for an evaluation and treatment recommendations. Occasionally, independent living communities may help organize or provide access to psychiatric services for residents. In any type of an outpatient setting, psychiatry can play an important role in assessing and identifying family, social, contextual, environmental, and physiological factors that can support or exacerbate the patient's mental health issues.

In any outpatient setting, clinicians should focus on: conducting a comprehensive evaluation of mood, cognition and functional capacity, paying particular attention to issues that may impair the patient's ability to perform ADLs or IADLs; gather information about the patient's social support network including family, friends, and community resources; provide psychoeducation on the identified mental health issue and how such issues may impact the older adult's ability to care for themselves; identify potential barriers to treatment including financial resources, transportation, housing, and access to health care; and gather information about the older adult's physical health and collaborate with the patient's physician or other members of the health care team. As with any population, there are challenges involved when treating older adults in an outpatient setting. The most salient issue is the constraint of time. Completing a comprehensive evaluation that includes an assessment of family history and function within a 30-minute intake poses a formidable challenge. The medical comorbidities with which many older adults present often add to the complexity of the assessment and treatment planning processes.

Mrs. D was an 89-year-old, white, widowed female of Italian descent, referred to an older adult outpatient specialty clinic for symptoms of depression and anxiety. She reported a long history of depression and anxiety and estimated that her current symptoms began 6 months ago when

her family decided she could no longer live alone after she gave herself food poisoning. This incident made the family realize how difficult it was for the patient to complete her ADLs due to arthritic pain, decreased mobility and stamina, and an unstable gait.

The patient's family consisted of her four daughters, three of whom lived locally. Each of the daughters contacted the psychiatrist separately to provide information about the patient's mental and physical health. During the intake, which two of the daughters attended, the patient was irritable and expressed her concerns about her eldest daughter's ability to care for her properly. The daughters argued openly about the patient's needs until one daughter left. As the psychiatrist gathered more information it was clear that this family had deep emotional wounds that negatively impacted all of their relationships. The daughters felt "burnt out" and stated that they cared for their mother "out of obligation, not love." Two of the daughters reported having been physically and emotionally abused by the patient.

Based on the strength-vulnerability model, this family is best described as having negative, weak and ambivalent bonds of attachment (Level 4). Communication and, therefore, problem solving were ineffective. They were unable to make decisions as a group and each wanted to be in a position of authority. Working on the hypothesis that the family needed to work on a family life review to encourage more positive family communication, the psychiatrist referred the patient and her family for family therapy. Meanwhile, the psychiatrist took active steps to create opportunities for positive communication and problem solving. The patient and her daughters were asked to create a list of activities with which the patient would require assistance, and the degree to which she would need assistance. The psychiatrist provided guidance regarding the ADLs and IADLs to consider. Family members were encouraged to review the list and to decide which type of help each was able to provide. An information system including the patient's needs, medical status, and psychological status was devised and shared with all family members on a biweekly basis. The combination of these interventions helped the family learn how to communicate and problem solve more effectively while working with a counselor to improve the overall quality of the attachment bonds.

Inpatient Medicine Service

Over 40 percent of older patients seen in the emergency department are admitted to the hospital. These patients not only face physiological problems but also suffer psychosocially, which is even more likely if diagnosed with chronic medical conditions. Estimates suggest that nearly 75 percent of older adults who enter the hospital functionally independent will leave requiring support with ADLs. The need for adequate discharge planning and improved long-term disposition is dependent upon assessment of their holistic needs, including functional status, cognitive capacity, mood, and social network/support. Though hospitalists are often relied upon to conduct such assessments, their time is limited and many have not been trained extensively in the unique needs of geriatric patients or in psychiatric interview techniques. Psychiatry can play a supportive role in acute medical settings by offering consultation to the medical team and to patients and their families, thereby ensuring a safe transition from hospital to home or an alternate placement. In addition to this consultative role, psychiatrists offer valuable help to patients coping with biopsychosocial changes and transitions.

To do this work most effectively, it is important to utilize a patient- and family-centered approach that recognizes the older adult as an activated patient who wishes to participate in the development of treatment and discharge plans. Psychiatry supports the patient and family by recognizing

the patient's preferences, helping them to make informed decisions, listening to their concerns and questions, providing them with education about the mind–body connection in the context of acute medical illness, and informing them of available resources. In an acute medical setting, psychiatrists can anticipate conducting a comprehensive assessment of mood, cognition, and functional capacity, paying attention to the illness story and changes that have occurred as result of hospitalization; evaluating patient's social supports and availability of caregiving; providing consultation on medication management of symptoms, including labile mood, poor appetite, and insomnia and identifying resources to address other patient and family needs.

Conducting a comprehensive assessment with an older adult in the medical setting is similar to the health assessments psychiatrists conduct with adult patients, but requires additional detailed attention to the functional capacity of the patient as outlined in Recommended Questions for the Patient.

RECOMMENDED QUESTIONS FOR THE PATIENT

Personal/Social Functioning

Where do you currently live? Where will you go after discharge? Where do you hope to go? Is there anything you (patient/family) worry about regarding discharge? Is there anything about the physical layout of your home that concerns you? Prior to your hospitalization, did you require help with ADLs/IADLs (specify each)? How does your illness limit you in your ability to engage in routine activities? Who has helped you in the past when you've become ill? Who will help when you go home? Are you the caregiver to a loved one (and, if yes, who will help you care for them now)? Have you experienced any other transitions or changes recently? Do you have transportation needs?

Experience of Past and Present Illness

Have you recently been hospitalized? What caused you to become sick enough to return to the hospital? Why were you readmitted? What could have been done differently before you left the hospital last time? What has gotten in the way of you taking care of yourself? What needs to happen while you are here/before you are discharged so that you remain healthy? How well do you follow medical recommendations? What gets in the way of taking care of yourself?

Psychiatry is also often tasked with providing psychoeducation to the treatment team, patient, and family. It is well established that a patient's physical functioning, number of comorbid medical conditions, and history of depressive episodes can result in comorbid mood and anxiety disorders in late life. When patients and families receive adequate information it

serves to validate the experienced change in mood, normalizes the experience of lability, acknowledges the mind–body connection, explains the anticipated recovery from changes in personality and mood, and provides patients and families with hope for improved quality of life following illness.

Clinicians can also provide supportive listening, paying attention to the patient and family experience of illness, the subsequent loss of function and independence, and fears of death. Through dialogue, the consultant can identify whether the development of improved coping skills and implementation of relaxation techniques at the bedside would be beneficial. The family draws support from acknowledgment that they are also impacted by the hospitalization. Caregivers are described as "invisible patients," as their needs are often not assessed or addressed in health care. They report that hospitalization causes disruptions to their everyday routines, distress, and difficulty meeting their own personal needs. Caregivers are often elderly themselves and require assistance with their own care in order to best serve the patient. Consideration for caregiver burden, burnout, risk of elder abuse, and determination of the family's capacity to provide care are essential components of a discharge plan.

Mr. B was a 76-year-old man of Jamaican descent. He was hospitalized following a stroke, suffering from dysarthric speech and changes in mobility. His wife of 47 years reported that their youngest daughter, who lived with them, had been admitted to the same hospital that week for management of congestive heart failure. Their six other children expressed concern about their mother's capacity to care for both her husband and daughter post discharge.

The patient was very tearful throughout his hospitalization. Mrs. B was increasingly distressed by the patient's continued dysarthric speech, which was exacerbated by his tearfulness. She wanted to know what was "wrong" with him, articulating that he had never suffered from such a labile mood and expressing uncertainty of how best to support him. Though the Medicine service recognized the patient's symptoms as common following a stroke, they did not have adequate time to address Mr. or Mrs. B's distress or understand it fully. The Medicine service felt strongly that a supportive consult from Psychiatry would be best suited to meet the patient's and family's needs and made a referral.

A psychiatry resident met Mr. B, asking questions about his medical experience. Mr. B expressed regret that he had not been a more spiritual man, noting that his survival would not have been possible without religious intervention. His wife frequently nodded her head, saying little except to occasionally repeat what Mr. B had said when concerned the resident had not understood him. Amidst the couple's sobbing the resident asked, "How much do you understand about why you feel this way?" Again, Mr. B noted religious reasoning. The resident was able to validate the family's tie to spirituality while simultaneously providing psychoeducation about the mind–body connection and impact of stroke on mood.

Mr. B heard this information and asked if he would always be a burden on his wife in the aftermath of this acute event. By then one of their sons had arrived and inquired about what types of care his father would likely require. Though occupational therapy had evaluated the patient, the family had not yet heard the results. The resident reviewed the results, noting areas in which the patient would require full assistance. He was then able to facilitate a conversation focused on problem solving, helping Mr. and Mrs. B identify their personal wishes for his recovery and the ways in which this would be facilitated best. At the close of the visit, the resident ensured that Mr. and Mrs. B were given resources regarding therapeutic services, including support groups and respite services for Mrs. B. The referring provider was given recommendations for medication management of symptoms, which had already been shared

with the family.

In this case, Psychiatry identified what the family understood about the patient's health and mood, and provided education on both topics. The family learned new coping strategies and received referrals to appropriate services both in the hospital and out. Through these interventions, Psychiatry supported Medicine, completed tasks that would have otherwise been the responsibility of the hospitalist and may not have been addressed as thoroughly. Psychiatry's efforts enabled the patient to become active in his own recovery and his family to become participants in discharge planning. Psychiatry helped Medicine transition the patient safely home, reduced the likelihood of readmission, and increased patient–family satisfaction.

Long-Term Care

While the number and variety of housing and residential care options for older adults have evolved to include retirement communities and assisted living facilities, this proliferation has not obviated the need for traditional long-term care facilities or nursing homes. Among the 1.5 million long-term care residents in the United States, 80 percent have one or more psychiatric diagnoses. Since the 1990s, nursing homes have absorbed increasing numbers of older adults with severe mental illness and serious behavioral problems and, hence, the prevalence of mental illnesses in long-term care populations is likely to rise. Despite the high prevalence of mental disorders in this population, less than one-fifth of nursing home residents with psychiatric diagnoses receive mental health treatment. In facilities that do provide mental health services, psychiatrists typically work in one of two models—as consultants or as members of multidisciplinary teams. At present, the consultant model is used more widely, and psychiatrists are most often consulted to address behavioral manifestations associated with adjustment issues, disease and infirmity, and dementia. A survey of geriatric psychiatrists in the United States found that the most common treatment recommendations in long-term care included psychiatric medications followed by changes in the general medical regimen, staff support interventions, medical diagnostic testing, behavioral interventions, staff training, individual or group psychotherapy, and family psychotherapy. These data suggest that geriatric psychiatrists tend not to rely solely on pharmacotherapy, but instead recommend a more diverse range of interventions in accordance with evidence-based practices. In some facilities psychiatry is the discipline responsible for intervening with family issues, especially conflicts between staff and families; in others, disciplines such as social workers, marriage and family therapists, and mental health counselors are available to serve in this role. Since 40 percent of long-term care residents generally reside with family prior to placement, family involvement after admission is commonplace. Additionally, long-term care placement is a significant life cycle event for older adult families that often triggers the manifestation of a spectrum of family dynamics.

Maria was a 47-year-old Caucasian woman of Italian-American descent and daughter of her 83-year-old father and nursing home resident. Her father, Mr. F., lived in the facility for 6 months following a discharge from the hospital following a stroke that left him with limited function on his left side. Prior to the stroke Mr. F. had lived in the in-law apartment attached to the home where Maria lives with her husband and son. Mr. F. moved into the apartment reluctantly 2 years ago following his wife's death. He did not want to leave his home, his church, or the diner in his neighborhood where he met "the guys" for coffee a couple of times a week. Mr. F. had been well known in the community "in the day" because he owned and operated a popular barbershop for many years. Maria ultimately persuaded him to move so that she could help him more and he would be near family. He continued to grieve the loss of his wife, and spent most of his time watching baseball games and old movies and helping his daughter with her garden. His health was stable for a while despite hypertension and high cholesterol, but Maria began to become concerned about when he started to lose his appetite, appeared noticeably thinner, and started to become a bit withdrawn. Shortly thereafter Mr. F. had a stroke, was hospitalized and then went to a rehab facility for 2 months. He recovered some function while in rehab but not enough to be able to return home safely.

Mr. F. did not resist moving into a nursing home as he believed he would never be "home" again anyway. After a month or so of adjusting to his recent physical limitations, he began to eat dinner in the dining room. Being a "ya gotta do what ya gotta do kind of man" the nursing staff and therapists from rehab liked him. He was quiet but friendly and very compliant in therapy. He met a couple of the other residents on the unit, began to watch baseball with another resident, and occasionally played cards or checkers when invited. The nursing home staff was pleased with his adjustment to the facility.

Maria was not as pleased. She didn't understand why he was not getting stronger. She repeatedly advocated for rehab services for her father at least 4 days per week. The social worker overheard Maria commenting to another visitor that "they were giving up on him." The social worker did her best to validate Maria's feelings while explaining that her father was receiving as much rehab as he could tolerate. Maria felt better after speaking with the social worker and continued to advocate for more rehab now that "the social worker said that she agreed with her."

Maria was also displeased with the food her father was served. She was concerned about his cholesterol and feared that he could have another stroke. She questioned his dietician about the menu choices he recommended. She advocated for fresh fruits and vegetables and talked about how important gardening had always been to him. She questioned the physician about alternative medicine and use of herbal supplements to reduce his risk of stroke. She decided to back down on that fight and took matters into her own hand by supplementing his diet with lots of green, leafy vegetables from her garden.

At the time of the 6-month family meeting the nurse manager opened by reporting that Mr. F. was doing really well, that he had developed a few friendships on the unit, that his mood had improved, and then declared, "We are all fond of him and I think he is happy here." Maria became agitated and started to voice other complaints. She asked if his bed could be moved because the sun glare obstructed his view of the TV. The nurse responded politely that she had discussed this with Mr. F. and he said that he preferred it the way it is because he did not wish to give up his view of the outdoors. This upset Maria because she felt that her father no longer needed her to advocate for him.

OPPORTUNITIES FOR INTERVENTION WITH MARIA

Provide Education

- To assure her that the meals meet nutritional standards
- To address her questions about herbal supplements
- To explain about the potential effects of green, leafy vegetables on clotting when blood thinners are involved

Improve Communication

- Develop and deliver "team" messages as a way to reinforce that all team members understand and can explain the plan of care

- Respect the delicate balance between allowing Maria to be "the expert" of her father and assuring her that the staff have relationships with him and are getting to know him also

Create Collaborative Team Meeting Formats

- Avoid "Team shares information–Family shares concerns" hierarchy
- Allow Maria an opportunity to add items to the agenda
- Come prepared with answers to questions Maria typically asks

Provide Maria With Opportunities To Be A Partner In Care

- Assure Maria that she maintains a special role in her father's life
- Encourage her to bring in sauce or other foods her father used to enjoy
- Show Maria how her contributions have shaped her father's plan of care or the way in which care is delivered

Palliative Care Service

Current research demonstrates that patients with advanced, life-threatening illness benefit from care that addresses their psychosocial needs, including loss, grief, quality of life, and suicidal ideation. Though rates of depression are as high as 58 percent in this population, psychiatric needs often go unrecognized or untreated. Lack of treatment is partially the result of limited availability of psychiatry consultation and clinicians' perceptions of their insufficient competence to meet the needs of this population.

Patients and families facing terminal illness are often seeking improved quality of life and maximum comfort. Psychiatry can support a patient–family-centered experience that enhances shared decision making about end-of-life decisions, honors patient preferences regarding place of care, and recognizes caregiver fatigue. Families require continued education not only about the disease process, but also about palliative care treatment options. Though palliative care teams are equipped to have these conversations, the dialogue can sometimes come as a surprise to families, and clinicians in psychiatry and medicine can continuously assess the family's readiness for palliative interventions. Additionally, psychiatry is uniquely positioned to help address family conflicts that may arise with respect to different perceptions around end-of-life options, and to help families transition from their in-hospital experience to post-loss bereavement experience and provide grief counseling.

Ms. P was a 67-year-old, Asian American when she presented to the hospital. Widowed for over a decade before, and suffering from amotivation and anhedonia, she had paid little attention to her own health and not modified her diet despite medical recommendations. She lived with her daughter and grandson and followed her grandson's dietary habits, eating largely sweet snacks and carbohydrates that were easy for her daughter to cook. The patient had suffered from uncontrolled diabetes for several years, episodes of gastroparesis, and chronic kidney disease. Upon hospital admission she was diagnosed with end-stage kidney disease and there was discussion of starting dialysis; alternatively, they noted that without dialysis, Ms. P would likely

die. In the interim, palliative care was monitoring the patient's medication management of pain, and had started a PCA. Ms. P was quiet whenever members of the medical team entered her room, as was her daughter; however, her daughter shared her preferences quite openly outside of the hospital room, adamant that her mother should receive dialysis.

Psychiatry was consulted to enhance communication between mother and daughter, and to help the team better assess Ms. P's interest in dialysis. Knowing that Ms. P and her daughter remained silent in front of the team, the psychiatrist recommended that a family meeting take place to include psychiatry, medicine, and palliative care. In doing so, the team was able to model (and subsequently encourage) effective communication. The psychiatrist demonstrated use of open-ended questions, facilitating a conversation that prompted participation from all team members, Ms. P, and her daughter. Ms. P eventually articulated that she was not interested in living a life that required intensive medical care and wanted to be with her husband. Her daughter was tearful but able to verbally accept her mother's choice. Palliative care was then able to provide insight about expectations of the disease process. With this information, the psychiatrist then invited Ms. P to consider other personal preferences and wishes. The psychiatrist supported further dialogue between mother and daughter. At conclusion of the conversation, Ms. P expressed gratitude that she might die with dignity, and was happy to know that she could live her remaining days at home with supportive services and her daughter and grandson nearby.

When Ms. P died a few weeks later, her daughter mailed a letter of thanks to the palliative care team and psychiatrist. She explained the value she perceived in preparing her for the future, avoiding treatment her mother viewed as futile, and improving their communication. She lamented the fact that a family meeting had not occurred sooner in her mother's care, wondering if the last decade of their relationship might have been different. She shared that she previously found her mother's quiet, stoic nature to be one of avoidance rather than resilience, and that the family meeting helped her view her mother more positively. She concluded that in giving her mother dignity at the end of her life, the psychiatrist had reduced the family's experience of guilt and grief.

Indications

Family therapy is most commonly used in older adults when the presenting problem is one in which the patient's distress or capacity for problem solving is rooted in a relational context. Family therapy is also indicated when the clinical situation at hand could likely be resolved with brief, solution-focused interventions that are designed to address specific therapeutic goals. Most serious mental health problems of older adults are more successfully addressed when family members are involved in the treatment. Although caution and clinical judgment must be exercised as to the type and timing of family interventions, substantial evidence indicates that psychiatric problems worsen or resolve less completely when conflictual or chaotic family environments are not addressed in the treatment. A major developmental task that families face throughout the life course is finding a balance between supporting individual members to be competent and independent while at the same time remaining connected to and part of the family as a whole. Families are often confronted with this challenge as they encounter life cycle transitions such as the birth of a child, marriage, retirement, an aging parent, or the loss of a loved one.

Limitations

Older adults are most likely to share psychiatric complaints with their primary care physician, and even then their report of symptoms is often in the form of somatic concerns. It is therefore quite difficult to reach aging families and implement therapeutic interventions without flexibility, integration, and collaboration. Though family therapy is most commonly provided by marriage and family therapists, social workers, and clinical psychologists with specialty training in family psychology, making referrals to family therapists is not always straightforward as requirements for licensure differ by state and availability of clinicians is often limited. Reimbursement for family therapy services varies by payer, type of mental health provider, setting, and state. Psychiatrists face similar barriers in the implementation of this work. In addition to complexities regarding access, a major limitation is the pervasive attitude among clinicians that potential for change in late-life families is negligible.

Contraindications

Family therapy with older adult clients is contraindicated if a patient's insight or capacity to change is impeded by cognitive impairment, paranoia, or active substance abuse. Conjoint family therapy is typically contraindicated in situations that involve extreme anger on the part of any involved family member or situations in which an assessment has indicated that therapy may increase the likelihood of abuse or other harm. Additionally, it is important to be mindful of the setting in which therapeutic interventions are being applied. For example, in an inpatient medical setting where an episode of care is brief, longstanding transgressions and disagreements are unlikely to be resolved. It is contraindicated to initiate exploration of such work without a plan for follow-up after a hospital discharge, or if a referral for outpatient care is not yet established.

Goals of Treatment

Family therapy interventions are used to reduce clinical symptoms that impact the individual and the family system. As noted, such interventions can be utilized to improve function by effectively changing the family's structure, patterns of interaction, problem-solving abilities, capacity for empathy and understanding, communication skills, and sometimes the meaning that a family ascribes to a certain event or situation. In the context of current health care reform, addressing a family's response to medical illness, death and dying, and the aging process is critical. The cases presented in this chapter provide examples of the multiple layers of family dynamics that accompany the transitions older adult families often confront.

ETHICAL ISSUES

Providing treatment to older adult families at critical times in an elder's life is embedded within layers of complexities. The settings in which the care is provided often introduce additional ethical considerations that may interfere with the provision of optimal patient- and family-centered care. Some ethical issues that clinicians are likely to encounter are discussed here.

End-of-Life Care

Patients, families, and care teams encounter difficult life-and-death decisions, which address medical decision-making, comfort management, and emotional and spiritual well-being. Though formal documents and legal statutes in the United States emphasize the importance of individual autonomy and independent decision making, research suggests that older adults want family members involved in end-of-life decisions. Those suffering from chronic and terminal illness may experience cognitive changes that impede their ability to make decisions independently and benefit from familial support. Despite this, research indicates that older adults and their children often disagree on key factors about end of life. Clinicians can support patients and families by facilitating open dialogue about end-of-life preferences, especially when patients still have capacity to share their wishes.

Clinicians further assist families to make decisions collectively by addressing the family's relational dynamics and life cycle considerations. Such conversations may occur rapidly in acute medical situations that require immediate decision making, or more slowly at the outset of a chronic illness. In some instances, family members express concern that they do not support their loved one's preferences regarding the use of life-sustaining measures. Such scenarios may require development or enhancement of families' natural problem-solving abilities, modeling of adaptive communication, or grief work. In such situations, clinicians are challenged to balance the patient's needs while acknowledging the larger family experience.

Capacity and Decision Making

Decisional capacity refers to a person's ability to make his or her own decisions in a variety of contexts including those related to health care. The issue of consent or refusal of treatment arises regularly when working with older adults, particularly in the inpatient setting. Some have argued that the medical conditions older adults face can contribute to feelings of helplessness, thus rendering these patients more open to the undue influence of others. Though clinical methods are available to assist with assessing capacity, this determination is made amidst multiple contextual

factors. Psychiatrists are thoroughly trained to assess capacity, but it is important to not only determine the patient's ability to understand and reason, but also to ascertain a sense of the patient's values prior to the onset of illness. Such evaluations and subsequent information sharing with the patient attempt to honor patient autonomy and appropriate family inclusion. It is important to understand shared versus conflicting values held by the patient and family, and the risk of conflict and potential coercion when values and wishes do not align. Further, if a patient is deemed "incompetent," clinicians also consider whether the surrogate decision-maker is acting with the patient's values and best interest in mind.

Cultural Considerations

Each generation of older adults gives rise to a unique set of values, beliefs, and family dynamics. These factors impact how older adults and families interact with the health care system and can play an important role in treatment decisions and outcomes. As increasing numbers of Baby Boomers reach the age of 65, it is important for clinicians to recognize the different social and demographic characteristics that define this unique cohort. Individuals in this group tend to be more educated than their predecessors, have more access to health care information, hold higher and more specific expectations of the health care system, and have a tendency to challenge authority. Baby Boomers are more open to receiving mental health services than previous generations. Despite these overarching themes, Baby Boomers comprise a heterogeneous group. Thus, careful evaluation and treatment planning are fundamental to providing care that honors and respects these elders and their families.

In addition to generational considerations, US Census data indicate that the fastest growing segment within the over-65 population is Latinos. There is a need for practice guidelines addressing the feasibility and acceptability of family therapy approaches with this population.

Elder Abuse, Neglect, and Exploitation

Elder mistreatment, including abuse and neglect, is a public health problem of significant concern. Estimates indicate that over 500,000 older adults in the United States alone are the victims of abuse or neglect. Community-based studies indicate that between 7 and 10 individuals over the age of 65 experienced abuse in the prior year, with the majority of abusers being family members. Older adults who are cognitively or functionally impaired are at higher risk for abuse, and some who experience severe disability are unable to report maltreatment. Elder abuse is one of the strongest predictors of long-term care placement. Advanced age, female gender, being in a spousal relationship, African-American race, living with others, physical and economical dependency on caregivers,

isolation, intergenerational conflict, alcoholism, and mental illness have all been associated with increased risk of abuse. Additionally, published reports suggest that family members who perceive a high level of caregiver burden, have a history of their own mental health needs, or have a history of substance abuse, may be more likely to engage in abusive behaviors, particularly when seeking compliance from the patient.

Clinicians must be alert to the needs of patients and their caregivers. It is helpful to look for signs of abuse or neglect, including repeated, unexplained bruises or injuries, malnutrition, fearfulness around particular family members, or, in the case of elders who are being financially exploited, a discrepancy between their ability to pay for personal commodities and their reported level of income. Clinicians can reduce the risk of abuse by conducting thorough assessments of caregiver burnout and by attending to the caregivers' experience of an aging relative. Family therapy is recommended to mediate conflicts that might give rise to mistreatment, but more research is needed to understand the treatment techniques that are most effective.

The majority of states now have mandatory reporting laws for elder abuse, as well as state or county-funded protective service programs for older adults who have been abused or are at risk. Practitioners are responsible for knowing state reporting laws. Although estimates indicate that one-third of actual cases of child abuse are reported, some studies estimate that only 1 in 14 cases of elder abuse is reported. Older adults are often reluctant to disclose the problem for fear of retaliation from the abusive family member. Further limitations to the frequency of reporting are thought to be related to the professional's fear of offending the patient or family.

RESEARCH AND EVALUATION

Despite the growing knowledge base on family systems and life cycle approaches to the treatment of older adults, evidence of the utility of these methods remains sparse. Controlled outcome studies are needed to confirm the general effectiveness of systemic family therapy with older adults and to clarify which approaches are applicable to the range of problems confronting families in later life. In contrast to the paucity of outcome studies of family systemic approaches to therapy, there are numerous studies of the effectiveness of psychoeducational interventions and cognitive-behavioral psychotherapies that are focused on family caregivers of older adults. Many of the most promising advances in the field are rooted in ecosystemic interventions that are delivered outside of the traditional health care practitioner's office. These interventions have been demonstrated to have positive effects on such outcomes as caregiver burden, depression, subjective well-being, knowledge and skills, and physical health, as well as on the level of symptoms expressed by the older

adult receiving care. At present, research on caregiving, elder abuse, and end-of-life care remain the most fertile areas for investigation of family therapy approaches.

References

American Association of Geriatric Psychiatry. Long-term care forum: the role of the geriatric psychiatrist in the nursing home. *Am Assoc Geriatr Psychiatry*. 2002;1:1–16.

Arean P, Huh T. Problem-solving therapy with older adults. In: Qualls SH, Knight BD, eds. *Psychotherapy for Depression in Older Adults*. Hoboken, NJ: John Wiley & Sons, Inc.; 2006:133.

Bartels SJ, Moak GS, Dums AR. Mental health services in nursing homes: models of mental health services in nursing homes: a review of the literature. *Ment Health*. 2002;53(11):1390–1396.

Binstock RH, George LK. *Handbook of Aging and the Social Sciences*. San Diego, CA: Academic Press; 2011.

Clissett P, Porock D, Harwood RH, Gladman JR. Experiences of family carers of older people with mental health problems in the acute general hospital: a qualitative study. *J Adv Nurs*. 2013;69:2707–2716.

Hargrave TD, Hanna SM. *The Aging Family: New Visions in Theory, Practice, and Reality*. New York: Brunner/Mazel; 1992.

Irwin SA, Ferris FD. The opportunity for psychiatry in palliative care. *Can J Psychiatry*. 2008;53:713–724.

King DA, Quill T. Working with families in palliative care: one size does not fit all. *J Palliat Med*. 2006;9:704–715.

King DA, Wynne LC. The emergence of "family integrity" in later life. *Fam Process*. 2004;43:7–21.

*Kissane DW, Bloch S. *Family Focused Grief Therapy: A Model of Family-Centered Care during Palliative Care and Bereavement*. Philadelphia, PA: Open University Press; 2002.

Lachs MS, Pillemer K. Elder abuse. *Lancet*. 2004;364:1263–1272.

*Lin M, Giles H. The dark side of family communication: a communication model of elder abuse and neglect. *Int Psychogeriatr*. 2013;25:1275–1290.

Marriott A, Donaldson C, Tarrier N, Burns A. Effectiveness of cognitive-behavioural family intervention in reducing the burden of care in carers of patients with Alzheimer's disease. *Br J Psychiatry*. 2000;176:557–562.

Nadeau JW. Family construction of meaning. In: Neimeyer RA, ed. *Meaning Reconstruction and the Experience of Loss*. Washington, DC: American Psychological Association; 2001.

Pruchno R. Special issue baby boomers: not your mother's old age: baby boomers at age 65. *Gerontologist*. 2012;52:149–152.

*Qualls SH. Family therapy with older adult clients. *J Clin Psychol*. 1999;55(8):977–990.

Qualls SH. Realizing power in intergenerational family hierarchies: Family reorganization when older adults decline. In: Duffy M, ed. *Handbook of Counseling and Psychotherapy with Older Adults*. New York: Wiley and Sons; 1999.

Rolland JS. Helping families with anticipatory loss. In: Walsh F, McGoldrick M, eds. *Living*

Beyond Loss: Death in the Family, 2nd ed. New York: WW Norton; 2004.

*Shields CG, King DA, Wynne LC. Interventions with later-life families. In: Mikesell RH, Lusterman DD, McDaniel SH, eds. *Integrating Family Therapy: Handbook of Family Psychology and Systems Theory*. Washington, DC: American Psychological Association; 1995.

Sorensen S, King DA, Pinquart M. Care of the caregiver: individual and family interventions. In: Loboprabhu SM, Molinari VA, Lomax JW, eds. *Supporting the Caregiver in Dementia: A Guide for Health Care Professionals*. Baltimore, MD: Johns Hopkins University Press; 2006:168.

Teri L. Training families to provide care: effects on people with dementia. *Int J Geriatr Psychiatry*. 1999;14:110–116; discussion 116–119.

The Joint Commission. *Advancing Effective Communication, Cultural Competence, and Patient- and Family-Centered Care. A Roadmap for Hospitals*. Oakbrook Terrace, IL: The Joint Commission; 2010.

*Walsh F. Families in later life: challenges and opportunities. In: Carter B, McGoldrick M, eds. *The Expanded Family Life Cycle: Individual, Family, and Social Perspectives*. 4th ed. Needham Heights, MA: Allyn and Bacon; 2010.

57.4k Group Therapy

KENNETH M. SCHWARTZ, M.D., AND MOLYN LESZCZ, M.D.

DEFINITION

Contemporary group therapy practice of the psychogeriatric population has expanded significantly as noted in earlier editions of this chapter. Relevant models of new applications are emerging steadily to meet the increasing needs of elderly patients with complex medical and mental health issues, although evaluation and research have yet to keep pace with these clinical developments. Foci of intervention include the aging-related losses of interpersonal, emotional, social, physical and cognitive capacities. The group therapies provide demonstrably effective treatment achieved through improving social support and engagement while enhancing individual patient and family member coping capacities and skills. Group therapy is provided in ambulatory, hospital, and institutional settings and ideally should also be available in community and primary care settings. In all instances the group leader's belief in the continued potential for growth and change in later life is of overarching importance: it is at the heart of clinical effectiveness.

HISTORY

The roots of group therapy can be traced back to the work of social workers in a variety of settings for seniors including homes for the aged, community centers, and mental institutions in the earlier part of the 20th century, providing group interventions combining activity and social treatments, emphasizing instruction, inspiration, and social support. In the following decades, interest in psychodynamic group approaches with elderly gained prominence as Maurice Linden, an American psychiatrist and pioneering group therapist, wrote a series of transformative articles on the topic. His early conclusions that group therapy provided psychotic and "senile" institutionalized patients opportunities for resocialization, happiness and some self-sufficiency have stood the test of time. The approach was well tolerated by patients, satisfying to staff, and economical. Group therapy has since broadened its reach with the provision of treatment in acute care settings, day hospitals, ambulatory, and community settings. Accompanying this has been a shift in the emphasis from the powerful and caring leader to interpersonal models emphasizing the importance of relationships between group members in healing and change. As well, not only are psychotherapeutic interventions routinely integrated with pharmacotherapy, but the recognition of the limits of singular models has encouraged integration within group psychotherapy itself, synthesizing the benefits of cognitive-behavioral, dynamic, and developmental models. As is

common in the field of group therapy, practitioners continue to respond to emerging clinical need with diverse populations with renewed emphasis on creating evidence-based models of intervention that are applied and then evaluated.

Current models include (1) verbal groups for the cognitively intact which include groups for burdened caregivers; (2) verbal groups for the cognitively impaired, including cognitive stimulation, resocialization, remotivation, and reality orientation groups; (3) activity- and creativity-centered groups, including dance movement, drama, music, art, and poetry groups; (4) a variety of time-limited homogeneous groups addressing issues such as retirement, bereavement, and medical illness; and (5) self-help and advocacy groups.

THEORETICAL ISSUES

Elderly individuals cover a large range, from baby boomers to centenarians, with an increasing diverse range of diagnoses and abilities. The need for more effective psychological treatments will become more, not less important, as the aging population becomes more psychologically sophisticated, having utilized psychotherapy at earlier stages of life.

The themes encountered in geriatric group therapy include reactions to loss, both developmental and traumatic, and to decline, with the resultant discontinuity with a prior, more valued sense of self. Consequently, themes revolve around issues of maintaining dignity, increased need for help and dependency, changing roles and/or conflictual relationships with adult children and mourning of losses of spouse and/or health.

Depression is a major public health concern because it is associated with high prevalence, morbidity, chronicity and health care cost, and affects both the cognitively intact and cognitively impaired elderly. Loneliness and the lack of social support are profound risk factors in depression. Group therapy provides opportunities to grieve losses and to re-engage with interests and relationships. It provides real relatedness alongside the acquisition of skills needed to generate continuing social integration beyond treatment. Because the experience of relatedness generated in the group is the result of the members' own actions, feelings of dependence may be reduced and feelings of personal competence may be increased. Enhanced social support produces a primary effect of improved well-being and a secondary, stress-buffering effect, which is particularly important for the elderly in light of the likelihood of further losses in the future. Higher levels of social support predict a significantly superior clinical course in the treatment of depression.

Often group members experience shame at having lost their prior sense of self, viewing the current incompetent mental and/or physical replacement to which they must adjust as bad or defective. In group, they can be assisted to learn that their presumed defect is merely human and

that other members also struggle with similar thoughts and feelings. The sharing of painful and secreted concerns can reduce the shame and stigmatization that isolate the individual. Through the therapeutic mechanism of altruism, members help one another, simultaneously raising self-esteem for both the receiver and provider of support. Realistic hope can similarly be nurtured countering feelings of demoralization and despair. Hope is malleable and can evolve with support to focus on wishes and aspirations even in the face of decline, loss, and limited time. The social milieu of the group, with members at different points along a developmental continuum can restore hope and also provide members with an opportunity to relate to, share with, and care about others dealing with similar challenges.

In a day hospital program for depression, each week typically saw the arrival of one to two new members and the impending departure of senior members. The group therapists determined that the task of introducing the integrated group therapy and orienting new members into the group would enhance the independence of senior members. Mr. Y, an 80-year-old widower, enthusiastically seized the opportunity to orient the new members to the group. He began by saying that he recalled his own start in the program 16 weeks before and how scared he was at starting something new. Another member commented that it was quite normal to feel apprehensive and discouraged starting out, adding how initially Mr. Y was so withdrawn and lacking in energy that he could barely get out of bed, never mind begin and contribute to the group as he was now doing. A third member recounted how he also benefited and encouraged the newcomers to give the group time to help them also get better. Although Mr. Y was sad at the prospect of being discharged from this very supportive and accepting environment, he would continue in follow-up sessions that he was confident would help him stay free from depression as he was leaving with many new skills and ideas that he had learned through therapy. In response to a question from a new member asking what he was leaving with, Mr. Y described a restored sense of confidence and usefulness, the capacity to challenge his tendency to withdraw, and a new set of skills that helped him better understand himself and others leading to improved relationships with his children.

A metatheoretical synthesis of dynamic and developmental approaches along with cognitive-behavioral approaches into a seamless whole emerges from the recognition that effective psychotherapy is facilitated by understanding of the individual's private, subjective experience residing beneath more manifest interpersonal behavior, bridging emotional insight of oneself and others with skill acquisition. In all instances the overarching goal is to increase each group member's sense of self-efficacy. Restoring a sense of personal effectiveness and the recognition of cause and effect in one's life is essential. This integrative model emphasizes the individual's sense of agency, as opposed to viewing oneself as a victim, of his or her own relational world. It combines interpersonal feedback within the here-and-now of the group and the processing of extra-group events and assigned homework, adding an intensity and immediacy that enlivens cognitive and behavioral approaches. Behavioral activation and skills training build atop this. Allied with these more traditional and recent approaches are mindfulness-based cognitive therapy groups and dialectical behavior

therapy aimed at increasing reflective capacity, reducing reactivity and enhancing skills in regulating negative emotions. Adapted for older adults, these approaches also focus on teaching skills to enhance openness to new experience and in creating meaning over the course of life. Problem-solving therapy is a structured approach that is derived from cognitive-behavioral principles and teaches older individuals to address problems by systematically examining and finding solutions to problems, without depth exploration. Life review approaches encourage reflection on historical material or reminiscences, either spontaneously produced or actively elicited by the therapists that is in turn explored and linked to the group member's current life and relationship. Past accomplishments and areas of competence and pride are highlighted to help restore feelings of worth, status, and competence through the articulation of past successes as the declaration of one's credentials. Members are encouraged to respond to the reminiscences of others and to explore their own emotional response to the material presented at the level of the here-and-now. Incorporating the patient's reminiscing about an earlier, more robust sense of self not only creates opportunities for psychological restoration by recollection of the past, but also sets the stage for contemporary employment of these valued elements of the self. The examination of actual models of group therapy described in the literature underscores this integrative process regardless of the specific noted designation of the intervention.

The elderly's awareness of the limits imposed by aging, loss, and decline focuses attention on existential issues regarding death, meaning in life, and personal responsibility. Group therapy should support exploration of these concerns with a view to facilitating members' meaningful engagement with life. The event of death obviously ends life, but the idea of death can vitalize life by focusing attention on both reducing regrets of things undone or unsaid and living life before time runs out.

Mrs. S, an 84-year-old woman in a group spoke of "how at memorials people often joked or were blocked from expressing their emotions." She contrasted this with the group which allowed members to share their vulnerability. Another member commented he used to think vulnerability was bad or was now a sign of weakness. Mrs. S agreed and went on to say she was once afraid of aging but was now helped by believing she was still growing because "I no longer get so angry and I can now approach my children to talk. I'm having the best time of my life, but I'm still sad. Why? I don't want my life to end as I've never had this in my life before."

Combined pharmacotherapy and psychotherapy are progressively employed in the acute, maintenance, and continuation phases of treatment. Combining drug and psychosocial therapies demands that attention be paid to the meaning each treatment has to the individual while being cognizant of the risks of devaluing one modality or the other. For example, a medication group where participants are encouraged to discuss their shame and fears associated with taking antidepressant medications can lead to

sustaining durable treatment responses, reducing subsequent affective morbidity, chronic psychosocial invalidism, and poor adherence to medical treatment. The universality of the group, that is, the realization that other group members taking medications may suffer from similar fears or shame coupled with the increased awareness that others benefit from these medications often leads to hope and improved acceptance of the use of medication.

Activity- and creativity-centered groups (e.g., dance, movement, pottery, music, or drama) encourage social engagement through shared participation in a pleasurable activity that may rekindle a connection with the previous sense of self through the use of the creative process. Group activities elicit feelings of pleasure and mastery; challenge self-denigrating and constrictive distortions that individuals may have about themselves; and counter the trend toward homogenization. Although the focus in the group is on a physical activity, integrating a psychodynamic appreciation to the meaning of the activity groups amplifies the use of the physical activity to promote social interchange. This has been demonstrated to assist in role transition, acceptance of loss in body function and integration of new goals, interests, and meaning in life.

Groups for caregivers are used widely in response to the increasing number of frail or cognitively impaired individuals being cared for in the community by their families. Increasing longevity is a challenge for both society and family members generally unprepared with little knowledge and few skills to deal with their loved one's physical and especially cognitive decline associated with dementia. These caregivers are often highly stressed, isolated, overwhelmed, and have an increased risk of psychological, physical, and financial problems, yet they remain the key resource to the implementation of any interventions for the identified patient. Groups may provide an opportunity for working through feelings of loss and abandonment, as well as anger at the burdens imposed on the caregiver. Attention to the process of the giving and receiving of mutual care within the actual groups provides a direct opportunity to address self-management and self-care issues. Education, social support, and skills training all may be integrated in these groups. Using simulation of challenging caregiver scenarios is also being applied to good effect in facilitating direct application of specific and relevant coping skills. Growing attention to individuals with mild cognitive impairment is evident in groups that provide psychoeducation, support, and coping strategies for couples with the objective of reducing anxiety, aligning expectations, and jointly coping while reducing marital stress.

Admittedly, cognitive impairment makes it more difficult to process new information, relate to what others are experiencing and participate in a group environment. Nonetheless, cognitive stimulation, resocialization, remotivation, and reality orientation groups for the more cognitively

impaired elderly play a significant role in institutional and community settings. Wellness group models teach skills, such as relaxation, guided imagery, and deep breathing, with the objective of fostering in members a sense of personal integrity and dignity to reduce the emotional distress that may underlie agitation and behavioral disruption. Members' sense of who they are within their environment is consistently validated. Interpersonal engagement with the present may be achieved through the structure and content of the group. Group interventions employ maximal multisensory stimulation in the form of sound, sight, and even pleasant aromas so that cognitively impaired individuals maximize opportunities for engagement with their environment and the retention of their cognitive and social abilities through sensory stimulation and reinforcement. Multisensory stimulation focuses on realigning sensory stimulus imbalance with gentle and humane exposure to language, sound, visual smell, texture, and touch stimuli: complex cognitive capacity is not required.

Group models are also developing outside of traditional venues. A group music intervention using percussion instruments with familiar music may lead to reduction of anxiety and agitated behaviors of those with dementia and to alleviate the burden of nursing caregivers in institutional settings. Telephone group therapy increases access for those unable to attend therapy sessions due to infirmity or distance. Technological advances and growing individual competence with technology will expand the reach of group therapies, as reflected in the growth of Internet support groups. Internet-based psychotherapeutic support group intervention for family caregivers of individuals with neurodegenerative illnesses already demonstrate the feasibility of such groups to benefit members with the reduction of caregiver stresses.

TECHNIQUES

Technical expertise in the application of the group therapies is surpassed in importance only by the therapist's posture of respect, hopefulness, genuine warmth, and empathy. Study after study has reiterated the findings that effective group therapy is predicated upon a strong sense of group cohesion and the leader's capacity for both receptive (understanding) and expressive (communicative) empathy. Beyond the relationship between member and leader, successful treatment hinges on the establishment of group cohesion. Modifications of therapy approaches are sometimes necessary for group work with the elderly in order to achieve and maintain cohesion. These include the use of cognitive reframing, supportive techniques, and practical arrangements for problems in hearing, vision, mobility, and other disabilities. Nonetheless, cohesion may be difficult to achieve with demoralized, withdrawn, and self- or other-devaluing patients. The therapist needs to be mindful that cohesion emerges both from the emotional bond shared by group members and the group's effectiveness

with its designated tasks. Therapist interventions follow from the understanding of the tasks and goals of the therapy. The rationale for the group and the therapist's intervention should be transparent and clear. Co-therapy may be useful in providing an appropriate level of focus and activation, continuity during therapist absences, and the added opportunity to address countertransferential reactions through co-therapy discussion and the co-constructed space for thinking. Within institutional settings, co-therapy also facilitates transmission of information between the group and other institutional care providers to ensure maximum treatment integration and reduce institutional resistances to the group interventions.

Attention must be paid to the maintenance of group norms and group integrity, which may be challenged by patient or institutional resistances. Patient resistance may be difficult to discern or be interwoven with illness or realistic reliance on others for transportation services. In institutional settings, an alliance between the group therapist and administration is essential. The group therapy inevitably reflects the culture and norms of the larger milieu in which it is embedded. The principle of isomorphy notes that important concerns, values, and fears are transmitted throughout the institutional system and gain expression at every level from patient to leader to group to ward to institution. Hence, an opportunity for dialogue and interchange between all the stakeholders in the treatment environment is invaluable, while striving to ensure appropriate privacy and confidentiality within the circle of care.

A key aspect of this is to ensure that group sessions in institutional settings are scheduled in an accessible location without conflict with other scheduled activities. Active outreach including the arrangement of portering may be required to promote regular attendance. Groups for the elderly may often be held at midday in ambulatory settings and may be contiguous with lunch, to encourage opportunities for social interaction around mealtime. In institutional settings and day hospital programs, the treatment frame is also affected by the fact that group members will often have involvement with one another outside of the group. Opportunities for engagement and interaction outside of the group can be employed as part of the prescription of behavioral interventions and to facilitate homework. In all instances, a group norm should be established that protects the confidentiality of the group members and encompasses extra-group contact within the treatment frame.

Therapy groups generally consist of 6 to 10 patients, but activity groups or groups for depressed or frail populations may be larger and benefit from the increased energy in the room. Duration of sessions is shaped by members' capacities and may vary from 30 minutes daily for cognitively impaired patients to 75 minutes weekly for more cognitively intact patients. Many group formats are ongoing, but the majority of cognitive-behavioral groups and groups for burdened caregivers are closed and time limited,

lasting from 6 to 20 sessions. The limits of briefer treatments are being recognized, and treatments are often extended to include follow-up sessions that meet at a reduced frequency, but over a longer period of time, to extend treatment into maintenance and continuation phases. It is not uncommon in some settings to see members attend a monthly maintenance phase meeting for many years or for group participants to continue to meet informally after the group sessions end.

The ideal treatment setting will offer an array of group therapy options, phasing and tailoring interventions to patient requirements with graded levels of intensity and depth. The therapy should be experienced as a self-esteem–enhancing and sustaining experience and a nonfailure experience for each member. Therapist activity must focus on promoting group interaction in ways that maintains the dignity of participants. The therapist should achieve a balance between support and expression and may need to intervene with active prompting, reinforcement, reframing, and modulation of emotional intensity, emphasizing the invitation for what is desired rather than the rebuke for what is not. An important therapeutic task is the empathic confrontation of projective defenses and resistances to engagement that isolate individuals and undermine group cohesion. The therapist may need to be quite active in reinforcing progroup or procohesion exchanges, bolstering patient feelings of worth and belonging. It is essential not to permit depressive affect to overwhelm the group as the depressed elderly are particularly prone to the effect of demoralization. Patients and therapists must understand that demoralization is a state—it is a specific emotional reaction and not an obligatory trait of aging.

Mrs. X, an 86-year-old widow, was a member of a group designed to help people adjust to living in a nursing home. Initially, frightened by her belief that a nursing home was a place where one went to die, she was pleased that the group members, some of whom had been in the home for years, reassured her that one could still have a good quality of life. She suffered from chronic depression and a history of unsatisfactory interpersonal relationships due to a perceived lack of attention, including from her son. Mrs. X entered the group unaware of how her responses when upset would offend other residents and staff leading them to stay away. With the group's help, she became aware of the consequences of her responses both inside and outside the group. Following the group's encouragement, she was able to appropriately invite her son to visit. That visit and subsequent visits went well and her mood and her adjustment improved.

An outbreak of viral illness then necessitated the cancellation of two group meetings. Upon resumption, Mrs. X was weepy as she complained about the lack of attention offered by her son and staff during the outbreak. Other group members shared their demoralization during this period. One group member, however, wondered if Mrs. X and other group members were more demoralized because of the break and enforced isolation of some of the ill members. Mrs. X then wondered if she was particularly affected because of her sensitivity to interpersonal rejection, as previously pointed out to her. The group acknowledged Mrs. X's increased insight and her affect brightened in conjunction with the other group members.

Self-psychology makes a valuable contribution in psychodynamic group approaches by virtue of the understanding it provides about the inner

world of the individual and the consequences of loss of self-object and self-sustaining relationships, capacities, and functions. Empathic awareness of the impact of loss and decline as experienced subjectively may diminish the risk of unwitting collusion with patient demoralization or engagement in negative interpersonal cycles of criticism and rejection. Without this perspective, the patient's narcissistic defences of devaluation, externalization, or self-inflation may recruit hostile and damaging feedback. The group may be used constructively in a self-sustaining mode by providing mirroring, idealized, adversarial, or twinning self-object transferences, thereby creating an opportunity for each member of the group to feel a valued member within a group of peers. This in turn may promote a restoration of a prior, more cohesive sense of self.

Life review and reminiscing techniques employ the naturally occurring process of patients' recollections and review of their lives. It can serve as a vehicle for making meaning of one's life viewing it as valuable and of consequence; and aid in resolving old conflicts and accepting losses. When used in a group format it also creates a sense of social identification among group members. Reminiscing can be used as a main treatment focus or to stimulate cohesion in beginning groups as an introductory exercise. Reminiscing can be spontaneous or prescribed in a series of sessions, following a chronological life course sequence. Either milestoning, which emphasizes only positive events, can be utilized, or the review process can be comprehensive and autobiographical, providing an opportunity for mourning, conflict resolution, and acceptance. Encouraging a focus on the past is not without risk. Facilitating the group members' return to dialogue within the here-and-now of the group, eliciting feedback and interaction, ensures that morbid self-absorption, preoccupation with painful memories or regrets, and guilt do not exacerbate the individual's withdrawal from the group or adherence to the past as a way of avoiding engagement with the present.

Aging often involves a shift from primary control and autonomy in which one effects change in the environment to deal with obstructions to one's needs and potentials, to secondary control in which one's beliefs and attitudes must be realigned to address these obstructions, for example, decreased mobility and access to social engagement and opportunities. Failing to do so may generate profound difficulties with self-criticism and beliefs about one's incompetence. Effective modification of thoughts and behaviors plays a major role in addressing the lack of perceived control. Therapists can challenge the "evidence" that patients have to support their fears. While recognizing age-related limitations of time and energy, the belief that their remaining life will be one of unhappiness and despair can be challenged, often through the countervailing experience in the here-and-now of the group. In turn, the behavioral responses employed to cope with change to life situations can also be evaluated for their effectiveness, and by

these means positive adjustment can be promoted and successful aging and an improved mood can be resumed.

Mr. Z, an 82-year-old widower in group in a retirement home, described his apparent decline. He first thought using a cane was a sign of getting old. Still not feeling old, he soon was told that for safety reasons he needed a walker. He imagined everyone he encountered was feeling sorry for him. Months later, and now in a wheelchair, he was struggling with not feeling sorry for himself. A co-member also in a wheelchair commented on similar struggles, particularly in light of having earlier led an active life. A third member encouraged both to look at what they could still do to help others who could not see or remember as well as them. Mr. Z gratefully received this feedback as one of his worries was whether he could still be useful and of value at his stage of life in an institutional setting.

Cognitive-behavioral therapy emphasizes conscious cognition and learning and adapts behavioral strategies and the cognitive therapy model to the group setting, affording opportunities for the acquisition of cognitive and behavioral skills. Cognitive therapy addresses the patient's problems with dysfunctional thinking patterns that reinforce negative cognitions and beliefs about the self. Behavioral approaches focus on remediating the patient's reduction of environmental or social reinforcement that results from his or her withdrawal, diminished interpersonal skills, or increase in adverse life events.

This approach emphasizes the teaching of skills with a consistent focus on patients assuming responsibility, crediting themselves for improvements, and stimulating a sense of mastery. Larger problems are deconstructed into smaller, more manageable components. Therapeutic strategies include challenging the patient's processes of generalizing, catastrophizing, magnifying, and black-and-white thinking. Patients are helped to examine evidence for and against their beliefs and by finding alternative explanations for their negative thoughts; reattribute meaning to events; schedule activities with the assertion of pleasurable activities and affects; utilize time meaningfully; and challenge both the discarding of positive options and the failure to take credit for partial positive gains. Patients may be taught to log the relationship between mood and either pleasurable or aversive experiences and to increase the frequency of pleasure-focused activities. Thought record logs, modified for use with the elderly, have been developed by Larry Thompson and colleagues. The Daily Record of Unhelpful Thoughts (DRUTS) facilitates this recording and review process. Active commitment to the assigned homework impacts directly on the outcome of therapy and may require input and support from family members. Behavioral prescription may be quite explicit and includes modeling, role play, didactic sessions, psychoeducation, behavioral exposure to anxiety stimuli, and the review of homework with the provision of feedback about each group member's efforts in using prescribed skills to re-engage with pleasurable experiences.

In a day hospital for depression, Mrs. A, a 78-year-old widow, commented upon entry to the group that she was depressed and that "the group will not help me." Other members of the cognitive-behavior group suggested a thought record could improve Mrs. M's mood. Group members then assisted Mrs. A in finding evidence that did not support her thoughts. In a later session focused on creating personal action plans, Mrs. A observed she had withdrawn from her friends, was not returning phone calls and was declining offers to socialize. Encouraged by the other group members, she decided on the action plan of returning a phone call to a friend. In the next session, she reported that she was able to complete the action plan of phoning the friend who was happy to hear from her that her mood improved as a result. Having achieved this success and seeing the effect on her mood, Mrs. A decided that she would call a friend several times a week.

Cognitive-behavioral therapy groups are quite structured and follow a predictable format of a check-in, homework review, summary, assignment of new homework, and discussion of anticipated obstructions to completion of the homework. It does require a baseline of cognitive capacity and in most instances respecting the impact of slower cognitive processing with greater respect paid to pacing, appropriate repetition, and reinforcement allows them to benefit. The group setting is ideal for this model, linking together mutual support, modeling, and the collaborative challenging of common cognitive and behavioral contributions to distress and depression. Early studies that showed that individual cognitive-behavioral therapy was superior to group cognitive-behavioral therapy likely reflect the inadequate attention to the impact of working with group process, an important consideration, even in highly structured groups.

Groups for cognitively impaired patients focus on patient engagement, with the structure of the group activity employing multiple modes of cognitive stimulation. Interpersonal and group processes are generally not addressed. These groups employ a broad range of sensory and cognitive stimuli to activate and engage patients, with consistent reinforcement and validation provided by the leader(s). The therapist should be guided by the principle of conferring meaning to every patient communication rather than dismissing these expressions as irrelevant. Cognitive exercises; activity and word games; life review; discussions; participation in arts and crafts; pleasant aromas; music; baking; guided imagery; relaxation; and physical exercise are components of these groups. What links all such groups is the moment-to-moment focus on esteem-enhancing patient engagement with the world around them. Groups of 6 to 10 participants work well. The group should be small enough that each member actively participates. Future growth in these approaches is likely, reflecting the growing role and positive effects of nonpharmacological interventions for dementia recommended in contemporary practice guidelines.

Group therapy interventions for caregivers work best when they integrate a variety of techniques which are structured and intense; require active involvement of caregivers, include education, support and respite; promote knowledge transfers, skill building and competency; and are

combined with individual support.

CLINICAL ISSUES

Indications

The choice of a group modality for the elderly may take into consideration both practical issues, such as the availability of various group types, along with affordable and convenient transportation, and patient issues, such as the type and extent of psychiatric illness and/or cognitive impairment.

Group therapy may be used as an adjunct to pharmacotherapy, individual treatment or as an independent treatment. With respect to selection criteria, structured groups, especially those that are more homogeneous in terms of membership, ego functioning, and cognitive capacities and not age alone are more effective than unstructured groups. There is a potential for poorer treatment outcomes in groups that lack either focus or appreciation for the potential to amplify depression by intensifying affects related to illness and death beyond some members' capacity. As with all treatments, therapeutic success is enhanced by the linking of patient goals and abilities to the tasks and techniques of the treatment. Mixing cognitively impaired and intact patients together is likely to jeopardize cohesion and group effectiveness by generating a wish among members to disidentify and withdraw from the more impaired participants.

Although pregroup preparation has not been extensively researched with the elderly, the evidence for its usefulness in other patient populations is quite substantial. Pretherapy preparation outlining the group treatment's goals and methods demystifies an otherwise anxiety-provoking situation, normalizes and destigmatizes treatment, helps to establish a therapeutic alliance, and sets group norms regarding regular attendance, confidentiality, and extra-group socialization. This process may increase group therapy tenure and task adherence and foster patient hopefulness. Role play or simulations may be utilized to induct the patient into the therapy role.

A written summary of the group therapy logistics, norms, tasks, and expectations may be useful to supplement the preparation interview. Pregroup preparation should also explore treatment resistances. Identification of ego-syntonic character difficulties, in particular narcissistic pathology or interpersonally alienating behaviors, may be useful in containing behaviors that can result in early peer rejection from the group.

Dropout rates in group therapy in general range between 10 and 50 percent in younger adults and are likely the same with the elderly. The likelihood of premature termination from treatment increases with increasing severity of depression, comorbid physical illness, and comorbid characterological difficulties of devaluation and externalization.

Burdened caregivers are a notably vulnerable group. Limits in health care resources coupled with the an aging population has resulted in the common situation of caregivers routinely providing at home the kind of care that in a previous generation was provided by professionals. Caregiver stress emerges from strain around safety, managing patient problematic behavior, self-care, social support, and personal emotional well-being. Caring for the cognitively impaired individual generates greater psychological strain and burden than caring for the medically ill. Family caregivers are often disproportionately vulnerable to both psychological and medical illness.

Contraindications

In general, individuals who require acute and active management because they are in crisis, suicidal, actively abusing substances, or acutely psychotic do not do well in general groups. Severe cognitive or sensory impairment or medical illness that obstructs the individual's engagement with the group enterprise may also require exclusion from group therapy. Some contraindications are relative and can be remedied by employing groups that are homogeneous for that condition. A program that has very few dropouts or therapy failures likely has set the bar for entry too high and may be missing opportunities to treat some patients who could benefit.

Goals

Notwithstanding the range of group formats described, consensus exists regarding realistic goals of treatment. Although obviously shaped by particular patient needs, in general, the goals of treatment include: reduction of symptomatic disturbance; reduction of isolation; restoration of interpersonal skills and engagement within a social network; correction of psychological distortions and cognitive beliefs; restoration of self-esteem; improved cognitive function; acquisition of problem-solving and coping skills; grieving and adaptation to loss; and realistic acceptance of one's needs and limitations with effective use of available resources. In institutional settings, group therapy may provide a forum to deal with the imposed inhibition of personal autonomy associated with routine, regulation, and restricted opportunity, thereby enhancing quality of life by addressing the dissonance between the individual's wishes and the imposed environment. At other times, addressing interpersonal tensions between residents may be the focus of treatment. A men's group may have particular value in institutional settings in which men are in a minority and feel isolated and particularly distant from former pursuits and interests.

Objectives of groups for burdened caregivers include education about the illness process and specific ways of providing care with the acquisition of strategies for dealing with disruptive behavior and problem-solving and

stress reduction skills. Working through loss and grief and facilitating appropriate disengagement that may lead to more realistic decisions about placement are also important objectives. Additional areas of focus include legitimizing the needs of the caregiver and promoting self-care, working through anger and guilt, and addressing relationships with health care professionals. These objectives may be integrated or addressed in a series of modules emphasizing important aspects sequentially.

Groups for the cognitively impaired aim to reduce excess disability and foster more appropriate social behavior and social functioning, thereby optimizing quality of life. Groups that emphasize creative expression aim to restore a sense of competence and mastery through the activity, be it art, drama, sculpting, or dance. Wellness and exercise groups aim to maintain patient well-being through physical activity and stress reduction provided in the context of a social milieu.

ETHICAL ISSUES

Proper patient selection and preparation and competent therapist functioning contribute to an environment that is generally well tolerated and experienced by participants as respectful and helpful. There is growing awareness of the differential impact group therapists can achieve, even when applying the same model of intervention. The interplay of therapist attributes and models of care may result in wide differences in impact.

Ethical care demands the therapist commit to ongoing training, evaluation of one's clinical impact and supervision or consultation. The ending of group may be resisted because of the important social support and self-restorative experience offered by the group and because it provides a sense of stability and consistency in the face of the participant's many life changes. This necessitates an active focus on the transfer of skills gained in treatment into the external world, recognizing that treatment is a bridge and not a destination. Therapeutic groups must be well integrated collaboratively into the overall milieu in which the individual resides. Alterations in the frame of group therapy required within institutional settings, along with other factors that contribute to a higher likelihood of extra-group contact both occurring and even being endorsed, warrants special attention to issues of confidentiality. Extra-group contact and the risk of subsequent subgrouping may be managed by diminishing secretiveness around the extra-group relationships and actively bringing them to the attention of the group.

Societal stereotypes of aging often lead to barriers hindering the elderly from being referred and appropriately treated. For too many, the idea of aging stirs up feelings of loss and deterioration, chronic medical illness and challenges the wish the public has for beauty and remaining young. As a result, attention to therapist countertransference and reactions to the clinical work is particularly important in the practice of group

psychotherapy. Realistic expectations guard against such negative countertransference as denial, hopelessness, and therapeutic despair. A reflective capacity for self-examination is particularly important in working with the elderly because of potential issues of identification with patients, filial countertransference, and avoidance of the issue of sexuality. Ageist beliefs in group members, such as aging inevitably being associated with senility, depression, illness, or being too old to change, when present, must be identified and worked through.

Group therapy is compatible with clinically indicated pharmacotherapy or individual therapy. Its concurrent use is complementary and not antagonistic. However, collaboration between treatment providers is the critical, at times neglected component of effective conjoint therapy.

An additional important ethical concern relates to the removal of barriers to access of care that are rooted in ethnic and cultural patient factors. The concept of cultural competence is gaining prominence in the field of mental health as an important variable influencing treatment effectiveness. Cultural competence relates to the therapist's appreciation for and understanding of the patient's specific cultural context regarding illness, aging, mental health, family relationships, and, of course, treatment. The ideal group therapy will be provided in the primary language of the participants and in accessible settings within the specific communities. Recognition of cross-cultural factors may require attention to the culturally resonant translation of therapy models.

RESEARCH AND EVALUATION

Substantial evidence has emerged documenting the effectiveness of the range of group therapies with the elderly. Meta-analysis of group psychotherapy studies for those over age 55 suffering from a variety of disorders such as depression, anxiety, and dementia document the average effect size for pre–post-studies was quite similar to those of younger adults and adolescents. Older adults who received group treatment were more improved compared to those who received no treatment or a control treatment, and the results seemed to be enduring. The young-old were more likely to benefit, but those with an average age over 76 still reported significant improvements over the course of therapy. Although group interventions for both functional and organic problems demonstrated a significant improvement at posttreatment, the treatments for functional and psychosocial disorders were more effective than treatments for organic problems, such as dementia. However, more group therapy research outcome studies are required as the number of groups providing adequate posttreatment follow-up data is limited.

Reviews underscore that conceptually sound and clinically relevant group approaches appear to be equally effective, equivalent to individual therapy, and significantly superior to no treatment in the area of geriatric

depression. Meta-analysis in this area documents the most effective and most robust evidence for cognitive-behavioral therapy groups compared to waiting-list controls. Dynamic group therapy approaches appear equivalent; however, further research studies evaluating the benefits of psychodynamic and interpersonal therapies in geriatric depression are required. The current evidence is most clear for mild or moderate depression, but even for severe depression group therapy may play an important role, allied with pharmacotherapy, in the acute, maintenance, and continuation phases of treatment. Combined pharmacotherapy and psychotherapy confer the best protection against depression and relapse and improve social adjustment in the longer term while improving medication compliance in the near term as well. Comparative trials of maintenance-phase group with individual therapy or pharmacotherapy or guided self-help groups are required.

A review of relevant treatment studies for depression in long-term care suggest that both group and individual psychotherapies hold promise in being efficacious with various kinds of residents, from those closer to a full mental capacity and relatively robust physical health, to residents with significant cognitive impairment or physical frailty. The evidence suggests that behavioral approaches that decrease pleasant events and cognitive approaches that challenge distorted cognitions both through increased self-esteem and self-mastery are effective, as are interpersonal therapies that address relationship issues or reminiscence or life-review approaches that seek to re-establish meaning in life.

Meta-analysis of psychosocial interventions in geriatric depression note that psychosocial interventions, where group psychotherapy is particularly prominent, demonstrate a significant effect size relative to placebo or no treatment comparisons. Positive outcome correlates with the degree of interpersonal engagement in the sense of interpersonal competence and effectiveness experienced by members of the group.

A meta-analysis of group cognitive-behavioral psychotherapy for subclinical depression in the elderly demonstrates statistically significant but modest reductions in depressive symptoms posttreatment compared with waiting-list controls. However this is not maintained at follow-up. Therefore, there is a need for more research to investigate the different outcome measures on the effect sizes, as the baseline scores in this population are relatively small, and the expected level of change for any intervention is at best modest.

Group interventions for individuals with anxiety disorders are effective, as significant improvements with worry, anxiety, and quality of life emerge posttreatment and endure 1-year follow-up. Mindfulness-based cognitive therapy groups show significant improvements in reported anxiety, ruminative thoughts, sleep problems, and decreased depressive symptoms. The use of dialectical behavior therapy as an augmentation strategy of

antidepressant medication is successfully used to treat chronic depression and personality disorders. Adapted for older adults, the approach focuses on teaching skills to enhance openness to new experience and in creating meaning over the course of life. Chronicity of symptoms may reduce treatment benefit and demand a longer treatment. Mild cognitive impairment may warrant modification of therapy technique, but does not preclude a positive response.

Group cognitive-behavioral therapy alone is efficacious in reducing comorbid anxiety and depression in the elderly and gains are maintained for at least 3 months. An integrative group therapy model combining a more structured cognitive-behavioral therapy with more process-oriented interpersonal therapy for depression and anxiety in later-life shows statistically significant changes in pre- and postgroup scores on several outcome measures for anxiety and depression and quality of life.

A review of two randomized clinical trials using Standard Dialectical Behavior Therapy (DBT) skills training groups as an augmentation strategy of antidepressant medication to treat chronic depression and personality disorders in the elderly demonstrates significant advantages for decreasing remission rates for medication and DBT compared to medication alone in the treatment of chronic depression. As well, DBT plus medication demonstrates superiority compared to medication alone in decreasing interpersonal sensitivity and interpersonal aggression at posttreatment and 6-month follow-up in a sample of elderly adults with comorbid depression and personality disorder. Research is underway to explore the effectiveness of the DBT model that has been modified to better meet the needs of the elderly.

A systematic review of the literature of group and individual cognitive interventions programs for individuals with mild cognitive impairment of the amnestic type demonstrate equal efficacy of both training formats. Statistically significant improvements after either treatment are obtained on objective memory and on subjective measures of memory, quality of life, or mood. It is suggested for further research to combine group and individual approaches as both approaches have unique benefits.

A group program for patients with mild cognitive impairment and their significant others based on cognitive-behavior principles combined with psychoeducational and cognitive rehabilitation elements incorporates slight adjustments in the CBT method because of cognitive impairments of patients. Pre- and posttreatment scales demonstrate that compared to a waiting-list control group, the group treatment helps patients to be better able to accept their illness, with the female patients especially showing a decrease in helplessness cognitions. No changes on subjective measures of general well-being or distress were experienced. It is suggested that efficacy of this developing group intervention may be enhanced by tailoring the content to patients sharing particular pertinent characteristics and by

incorporating more sex-specific themes and training.

Groups for burdened caregivers generally demonstrate that participants are very satisfied with the group experience. Complexities in the measurement of subjective and objective burden, coupled with the likely progressive deterioration of the care recipient may confound other outcome measures. The acquisition of knowledge is easy to measure and demonstrate successfully. Reductions in burden are harder to demonstrate. Both professional- and peer-led groups show significant effects with the former tending to emphasize the provision of information and problem solving and the latter emphasizing peer support and networking. The best results emerge from integrative interventions that employ a broad focus, recognizing that there are multiple domains of concern for caregivers. This is consistent with a summary of the reviews of the caregiver intervention literature that demonstrates that more recent intervention studies are more effective when they employ combined interventions that target both the caregiver and patient.

A pilot study of evidence-based interventions for family caregivers of individuals with behavioral and psychological symptoms of dementia demonstrates improvements in carers' self-perception of competence, significant reductions in emotion-focused coping, enhanced mastery and reduced burden. Carers with more compromised baselines benefited most from the intervention comprised of three integrated components: group psychotherapy, problem-solving techniques, and skill acquisition for specific current challenging interactions in caregiving. Systematic use of standardized patients to role play the spouse or parent with dementia is employed. Limitations of the study include lack of a control group of carers and no outcome measures of care recipients.

FUTURE DIRECTIONS

An ever-increasing need for effective and meaningful group therapy interventions for the elderly and their families can be anticipated. Challenges and opportunities that lie ahead include creating models of intervention that are culturally accessible to a heterogeneous population of seniors. The field will also need to develop and test interventions for mood disorders that address common comorbid conditions including anxiety, substance abuse, personality disorders, and, of course, cognitive impairment. In allied fashion, integrative group interventions for the treatment and management of patients with dementia and their families, both together and separately, will need to be developed. Groups for caregivers, both family and professional, of older adults with late-life psychiatric and medical illness and behavioral and psychological symptoms of dementia also need to be developed. In all domains, greater efforts will need to be made in the training of group therapists as theory is linked to practice to empirical evaluation of the therapy.

REFERENCES

Baker MN, Baker HS. Self psychological contributions to the theory and practice of group psychotherapy. In: Alonso A, Swiller HI, eds. *Group Therapy in Clinical Practice*. Washington, DC: American Psychiatric Press; 1993.

Bernard H, Burlingame G, Flores P, Greene L, Joyce A; Science to Service Task Force. *Practice Guidelines for Group Psychotherapy*. New York: American Group Psychotherapy Association; 2007.

Cheston R, Ivanecka A. Individual and group psychotherapy with people diagnosed with dementia: a systematic review of the literature. *Int J Geriatr Psychiatry*. 2017;32(1):3–31.

*Chiu M, Wesson V, Sadavoy J. Improving caregiving competence, stress coping, and mental well-being in informal dementia carers. *World J Psychiatry*. 2013;3:65–73.

Elias SMS, Neville C, Scott T. The effectiveness of group reminiscence therapy for loneliness, anxiety and depression in older adults in long-term care: a systematic review. *Geriatr Nurs*. 2015;36(5):372–380.

Fain DS, Sharon A, Moscovici L, Schreiber S. Psychotropic medication from an object relations theory perspective: an analysis of vignettes from group psychotherapy. *Int J Group Psychother*. 2008;58:303–326.

Helmes E, Ward BG. Mindfulness-based cognitive therapy for anxiety symptoms in older adults in residential care. *Aging Ment Health*. 2015;1–7.

*Jean L, Bergeron ME, Thivierge S, Simard M. Cognitive intervention programs for individuals with mild cognitive impairment: systematic review of the literature. *Am J Geriatr Psychiatry*. 2010;18:281–296.

Joosten-Weyn Banningh LW, Prins JB, Vernooij-Dassen MJ, Wijnen HH, Olde Rikkert MG, Kessels RP. Group therapy for patients with mild cognitive impairment and their significant others: results of a waiting-list controlled trial. *Gerontology*. 2011;57:444–454.

Krishna M, Lepping P, Jones S, et al. Systematic review and meta-analysis of group cognitive behavioural psychotherapy treatment for sub-clinical depression. *Asian J Psychiatry*. 2015;16:7–16.

Leszcz M. Integrated group psychotherapy for the treatment of depression in the elderly. *Group*. 1997;21:89–113.

*Leszcz M. Group therapy. In: Sadavoy S, Lazarus LW, Jarvik LF, Grosberg GT, eds. *Comprehensive Review of Geriatrics Psychiatry II*. 3rd ed. Washington, DC: American Psychiatric Press Inc.; 2004.

Linden M. Group psychotherapy with institutionalized senile women: study in gerontologic human relations. *Int J Group Psychother*. 1953;3:150–170.

Lynch TR, Cheavens JS, Cukrowicz KC, Thorp SR, Bonner L, Beyer J. Treatment of older adults with co-morbid personality disorder and depression: a dialectical behavior therapy approach. *Int J Geriatr Psychiatry*. 2007;22:131–143.

Madan R. Cognitive behavior therapy (Individual and group). In: Schwartz K, Madan R, eds. *A community outpatient model for the treatment of depression. The Baycrest Psychiatric Day Hospital Intervention Training Manual*. Toronto, Ontario: Baycrest; 2013.

Marziali E, Damianakis T, Donahue P. Internet-based clinical services: virtual support groups for family caregivers. *J Technol Hum Serv*. 2006;24:39.

Meeten F, Whiting S, Williams CM. An exploratory study of group mindfulness-based cognitive therapy for older people with depression. *Mindfulness*. 2015;6(3):467–474.

Otsuka T, Tanemura R, Noda K, Nagao T, Sakai H, Luo ZW. Development of computer-aided cognitive training program for elderly and its effectiveness through a 6 months group intervention study. *Curr Alzheimer Res*. 2015;12(6):553–562.

*Payne KT, Marcus DK. The efficacy of group psychotherapy for older adult clients: a meta-analysis. *Group Dynamics: Theory, Research and Practice*. 2008;12:268.

Saiger G. Group psychotherapy with older adults. *Psychiatry*. 2001 Summer;64:132–145.

*Schwartz K. Remembering the forgotten: psychotherapy groups for the nursing home resident. *Int J Group Psychother*. 2007;57:497–514.

Schwartz K, Schwartzberg SL. Psychodynamically informed groups for elders: a comparison of verbal and activity groups. *Group*. 2011;35:17–31.

Sitzer JI, Twamley EW, Jeste DV. Cognitive training in Alzheimer's disease: a meta-analysis of the literature. *Acta Psychiatr Scand*. 2006;114:75–90.

Sochting I, O'Neal E, Third B, Rogers J, Ogrodniczuk JS. An integrative group therapy model for depression and anxiety in later life. *Int J Group Psychother*. 2013;63:502–523.

Sole C, Mercadal-Brotons M, Galati A, De Castro M. Effects of group music therapy on quality of life, affect, and participation in people with varying levels of dementia. *J Music Ther*. 2014;51(1):103–125.

Thompson LW, Powers DP, Coon DW, Takagi K, McKibbin C. Older adults. In: White JR, Freeman AS, eds. *Cognitive-Behavioral Group Therapy for Specific Problems and Populations*. Washington, DC: American Psychological Association; 2000:235.

Wang CC, Tzeng DS, Chung WC. The effect of early group psychotherapy on depressive symptoms and quality of life among residents of an apartment building for seniors. *Psychogeriatrics*. 2014;14(1):38–46.

Wetherell JL, Petkus AJ, Alonso-Fernandez M, Bower ES, Steiner AR, Afari N. Age moderates response to acceptance and commitment therapy vs. cognitive behavioural therapy for chronic pain. *Int J Geriatr Psychiatry*. 2016;31(3):302–308.

Wuthrich V, Rapee R, Kangas M, Perini S. Randomized controlled trial of group cognitive behavioural therapy compared to a discussion group for co-morbid anxiety and depression in older adults. *Psychol Med*. 2016;46(4):785–795.

Yalom ID, Leszcz M. *The Theory and Practice of Group Psychotherapy*. 5th ed. New York: Basic Books; 2005.

57.4l Counseling and Support Needs of Dementia Caregivers

LIAT AYALON, PH.D., AND PATRICIA A. AREáN, PH.D.

INTRODUCTION AND DEFINITION OF CAREGIVING

What does caregiving mean? According to Webster's online addition, a *caregiver* is someone, either family member or provider, who has responsibility for meeting the physical and psychological needs of an infant, child, or dependent adult. While anyone may assume the function of

caregiver, primary caregivers are individuals who help the patient with at least one instrumental activity of daily living, such as bathing, eating, dressing, or getting to appointments, or a person who supervises a disabled patient. Two types of primary caregivers exist: *Formal caregivers,* individuals who are paid to meet the physical and psychological needs of patients, and *informal caregivers,* individuals who provide those needs voluntarily. This usually involves family members and friends. Although the reasons for assuming these responsibilities vary between the two types of caregivers, the burden of caring for a demented person is still high for both types of caregivers. Currently, informal caregivers are the majority of the caregiving population. Less than 10 percent of all older adults received their care solely from paid carers. In 2012, more than 15 million informal caregivers provided care to individuals with AD and other dementias. This care has a tremendous economic value mounting to $216 billion and 17.5 billion hours of care. Given the increase in life span, the decrease in childbirth, the entrance of women into the workforce, the high divorce rate, and the popularity of the nuclear family system, it is expected that formal long-term care alternatives will become more popular in future years.

Who is most likely to be a caregiver? According to the Behavioral Risk Factor Surveillance Survey, 62 percent of all informal caregivers are women. According to Kaycee M. Sink and colleagues, caregivers are on average 64 years old and largely female (72 percent) and nearly half of caregivers are spouses. About two-thirds of caregivers also work outside the home in addition to their caregiving duties. According to the National Alliance for Caregiving (NAC/AARP), 30 percent of all caregivers also cared for children younger than 18 and thus, are torn between caring for an older relative and younger children.

Researchers found no demographic differences between caregivers of individuals with dementia and those with milder cognitive impairments. Most caregivers of dementia patients care for people with Alzheimer disease, although a fair number also care for people with Lewy body disease, frontotemporal dementia, and vascular dementia. Each of these illnesses poses a unique set of problems for caregivers. Alzheimer disease is typically characterized initially by memory impairments, but as the illness progresses, can result in personality and behavioral changes that may be upsetting to caregivers. Lewy body disease shares this same course but is coupled with physical disability related to tremors and unusual hallucinations that many caregivers find hard to understand and manage. Frontotemporal dementia is particularly vexing, as patients who develop this illness tend to be younger and physically healthy (men in their 50s), and while their memory remains intact for a long time in the course of the illness, behaviorally these patients become impulsive, inappropriate, and highly agitated. These disorders are not mutually exclusive. A person with Alzheimer disease may also have vascular disease; a person with Lewy body

disease may also develop frontotemporal dementia. The more complicated the dementing illness in the patient, the more caregivers have to contend with. Because Alzheimer disease is the most common dementia afflicting older people, caregivers of Alzheimer disease have received the most attention, with the majority of research focused on informal caregiving.

The Mental Health Consequences of Caregiving

Caregivers are generally at greater risk for mental health and health problems than noncaregivers. Those caring for older adults with dementia tend to present with more negative emotional consequences compared with those who care for older adults who do not suffer from dementia. The most common mental health consequences of caring for a dementia patient are depression, anxiety, and a condition called *burnout*. These three conditions are often lumped together under the term *caregiver distress*. Studies find mixed results as to which condition is most common. It is largely believed that depression is the most common mental health consequence of caregiving. The Aging, Demographics and Memory Study (ADAMS) found that as many as 44 percent of the caregivers of people with dementia were depressed, but only 27 percent of the caregivers of older adults with cognitive impairment but no dementia indicated depressive symptoms. Others have shown that as many as 40 percent of caregivers of older adults with Alzheimer disease are depressed, but only 5 percent of noncaregivers report depressive symptoms and meet the clinically significant cutoff of depression.

It is important to note that burnout is not necessarily the same condition as depression or anxiety, but these conditions do co-occur. According to Schulz, caregivers of dementia patients are particularly vulnerable to burnout, a psychological and physical state of exhaustion resulting from excessive and prolonged stress. Caregivers who are family members are particularly vulnerable to burnout, and many speculate that this emotional consequence to caregiving explains why 90 percent of elder abuse is committed by family members. Stress and burnout are not one and the same. People under stress can generally feel hopeful about the future, and when no longer in the stressful situation they begin to feel like their normal selves. Burnout is far more severe and is characterized by frustration; feeling powerless and hopeless; feeling drained of emotional energy; becoming detached from others, which results in withdrawal and isolation; and feeling trapped, irritable, and sad.

Caregiver distress, be it depression, anxiety, or burnout, does not necessarily remit after caregiving responsibilities end. According to Carol S. Aneshensel and colleagues, caregivers who are less distressed before the death of the care recipient tend to have better mental health outcomes, whereas caregivers with significant emotional distress while caregiving tend to become even more depressed after the loss of the care recipient.

Caregiver burden before death tends to increase the odds of becoming depressed; self-esteem and emotional support tend to have protective roles.

The Effects of Caregiving on Physical Health

In addition to mental health consequences, there is a notable increase for poor health outcomes in caregivers. Compared with caregivers to older adults without dementia, caregivers to older adults with dementia are more likely to say that caregiving worsens their health. Some researchers speculate that poor health outcomes are the result of poor health care behaviors; caregivers generally do not have sufficient time to attend to their own medical appointments, and therefore potentially preventable health problems arise, such as hypertension and heart disease. Recent research also suggests that the stress related to caregiving may also influence health outcomes. Brent T. Mausbach and colleagues found that an exaggerated plasma norepinephrine response in caregivers is linked to an increased allostatic load related to cardiovascular disease. These researchers have also associated caregiver stress to reduced immune cell ss2-receptor sensitivity and increased D-dimer marker of hypercoagulability. Roland von Känel and colleagues have also found that caregiver distress is associated with increased frailty, affecting interleukin-6, C-reactive protein, and D-dimer. Janice Kiecolt-Glaser and colleagues found that caregiver distress has a direct effect on immune function through increases in the pro-inflammatory cytokine interleukin-6. Taken together, these studies suggest that morbidity and mortality related to caregiving have a physical basis, and that beyond mental health consequences, caregiving distress has significant and direct impacts on health.

HISTORY AND THEORETICAL ISSUES

Attention to the needs of informal caregivers has been increasing over the past three decades. This is primarily due to the changing demographics and the realization that caregiving to older adults with dementia carries substantial economic, emotional, physical, and relational tolls. Several theoretical models have been offered to account for the negative effects associated with caregiving. Most theoretical models explain caregivers' burden, depression, and ill health by proposing a relationship between primary stressors such as the older adult's functional or cognitive impairment and the caregiver's outcomes. This relationship is thought to be mediated through the amount of informal support provided by the caregiver and the use of formal support.

TECHNIQUES FOR PREVENTING AND MANAGING CAREGIVER DISTRESS

Intervention models are often based on the above-detailed conceptualization of caregiver distress, and many are multicomponent

models that address several risk factors, including education about the dementing illness, emotional support, information about resources, and counseling in the form of skills training, anger management, and mood management. The most widely studied interventions are skills-training classes, educational interventions, and support groups. Interventions that are administered over a long period of time and target the entire family are considered particularly effective.

Skills-Training Classes

There is considerable evidence in support of the efficacy of structured classes aimed at educating caregivers in stress management, anger management, and self-care. These interventions typically are multicomponent, in that they teach caregivers about the causes of dementia, explain the reasons for behaviors they will see in the care recipient, and provide them with strategies for managing their anger, mood, and burden. Most of these interventions are behaviorally focused; they teach caregivers how to refocus their understanding of the care recipient's behavior, make use of stress-reduction strategies like meditation and relaxation training, increase pleasant activities, and find opportunities to take a break from their caregiving responsibilities. The success of these interventions is very well documented. According to Mary Mittelman and colleagues, interventions aimed at mediating the factors related to caregiver distress (coping, appraisals, support) have significant effects on distress even 4 years after treatment. Interventions aimed at anger management and burnout are also effective. According to David Coon and colleagues, anger management and skills-building interventions have positive impacts on caregiver appraisal and self-efficacy.

Skills-training classes appear to be effective across ethnic groups. According to one study, a multicomponent intervention that targets stress management and self-care strategies was better than education and support alone in Caucasian, Hispanic, and African-American caregivers. Caregivers with low-pretreatment self-efficacy seemed to respond better to these interventions than those with high self-efficacy, although across groups, everyone experienced benefit from these skills-building groups. One of the largest studies of caregiver treatment is the REACH trial, a study of 167 caregiver–recipient dyads, which found that coping and stress management classes were far more effective in managing caregiver depression and burden, than interventions focused solely on how to manage behavior related to dementia. These findings were robust for 2 years after caregivers participated in the coping skills interventions. The REACH trial has also shown to improve self-reported health status.

Education and Support Groups

Research shows that even minimal education classes that focus on caregiver-coping strategies and information about dementia have benefit for caregivers in mitigating distress. In one study, a minimal information control condition was compared to a skills-building program for African-American and Caucasian caregivers. The Caucasian caregivers benefited from the information control condition better than the skills-building condition, indicating that even the provision of information about dementia and caregiving can have substantial benefit to caregivers. Robert Woods and colleagues compared a mental health nursing program to support caregivers to community educational programs and found that community educational programs were as effective in reducing caregiver distress as the more intensive nursing program.

CLINICAL ISSUES

Indications

Not everyone who cares for a demented person feels the experience is a negative one. Some people, while experiencing some stress, are able to cope well with the role of caregiver. Spouse caregivers are more likely to report depressive symptoms, greater financial and physical burden, and lower psychological well-being. This is only partially accounted for by the fact that spouse caregivers also provide more hours of care. Compared with men, women caregivers are more likely to report depressive symptoms. Those caring for older adults with frontotemporal dementia are more likely to report distress and strain compared with caregivers of older adults with Alzheimer disease. Other studies have shown that certain ethnic groups cope better with caregiving than others; for instance, Hispanic and Black caregivers are more likely to report high levels of burden compared with White and Asian caregivers. A number of caregiver characteristics seems to be related to whether or not the caregiver will be distressed by his or her responsibilities. The most notable and well-studied characteristics related to the development or prevention of caregiver distress are the caregiver's culture, coping style, social support, and formal support; relationship between the caregiver and care recipient; and, finally, characteristics of the care recipient. While each of these factors contributes to distress, most researchers agree that these factors interact with one another, and no one factor alone contributes to caregiver distress.

Limitations

One of the most complicated aspects of treatment for distress is access to these effective interventions. Several studies show that caregivers rarely feel they have time to devote to their own well-being. A number of studies have been done to see if telemedicine-type of support could benefit caregivers. The studies have largely been mixed, indicating that simple

telephonic support does not result in as good outcomes as in-person group support. However, a study of 299 caregivers receiving a multimedia caregiver intervention over the Internet found that Internet interventions of this nature provided a low-cost, convenient, and effective means of overcoming caregiver distress, as well as being flexible enough to tailor the intervention to the needs of the caregivers. Further, web-based support combined with in-person caregiving groups appeared to have a substantial impact on depression in 225 caregivers. Finally, Sara Anne Smith and Paul Bell studied a CD-ROM plus printed material intervention for rural caregivers, and found this intervention to result in greater use of available support services, use of behavioral interventions to manage care-recipient aberrant behavior, and reductions in distress. Research of this nature is beginning to show that access to effective psychosocial interventions is possible for caregivers. However, a recent systematic review found only low levels of support for the efficacy of Internet-based supportive interventions.

Mrs. Y was a 56-year-old Caucasian woman whose husband of 30 years was diagnosed with Alzheimer disease 3 years ago. Over the years, Mr. Y had become increasingly forgetful and impulsive, often leaving food on the stove to burn, walking to the corner store and becoming lost, and buying expensive items that the family did not need. In the last year, Mr. Y bought two new cars and three high-definition televisions, despite the fact that he already owned a 2-year-old car in good condition and the Y did not watch television. Since Mr. Y's dementia became more apparent, he had to stop working, and Mrs. Y, as a consequence, also had to cut back on her working hours to accommodate Mr. Y's needs. In addition to Mr. Y's forgetfulness and impulsivity, his sleep patterns also became disturbed, and he often woke at night and wandered around the house, which disturbed Mrs. Y's sleep.

Mr. and Mrs. Y had two children, both of whom lived in different states; thus, Mrs. Y had to cope with caregiving demands on her own. As a result of a part-time job and an increasing worry that leaving Mr. Y at home alone would result in either a fire in the home or more unnecessary excursions to car dealers and electronics stores, Mrs. Y became quite distressed and experienced several colds and flus that she was not able to care for. One such illness resulted in a case of pneumonia, and it was at this point in time that Mrs. Y came to the attention of the clinical team.

Although Mr. Y had ongoing neurology appointments in a prestigious university hospital, it was clear that Mrs. Y had not made use of the resources available to her through this service. Her reason for not attending the support groups, working with the social worker to link her to respite care, and making use of community resources was her lack of time between taking care of Mr. Y and her work to help support them financially. While in the hospital, the team social worker visited Mrs. Y to talk to her about the importance of making use of available services and the need to care for herself first, so that she could then better care for Mr. Y. From this discussion, it was clear that Mrs. Y had become depressed and burnt out. Not only was she suffering physically from her caregiving responsibilities, but she was also feeling less interested in usual activities, had lost her appetite, had begun feeling hopeless and helpless, had trouble concentrating at work, and wished she or her husband would die. This last symptom left her feeling incredibly guilty; she loved her husband, and although forgetful and impulsive, he remained a loving companion to her. She was also clearly resentful toward her children; when she entered the hospital, her daughter had reluctantly flown in from her home to care for Mr. Y, and had made it clear she would only stay until Mrs. Y returned from the hospital.

A clinical team began to work on reducing Mrs. Y's caregiver distress in the following way. First, the social worker began to make arrangements for Mr. Y to attend a day treatment program that would provide services within the Y's financial means. Mr. Y's schedule would be mapped to Mrs. Y's work schedule, so that she would not have to worry about Mr. Y being

home alone. Temporary evening respite care was also arranged while Mrs. Y recovered from her pneumonia. This respite care eventually was reduced to 1 day a week so that Mrs. Y could have time for herself and attend a skills-building class for dementia caregivers. Mrs. Y was reluctant to join the class—in fact, she only attended intermittently; however, she did make use of the neurology department's online caregiver support program, in conjunction with the occasional classes, and reported that she found the website helpful.

Six months after her hospitalization for pneumonia, Mrs. Y was feeling both physically and emotionally better. She still struggled with her husband's increasing dementia, but had become more active, reengaged with friends, and was more productive at work. One year after her hospitalization, Mr. Y was moved to an assisted-living facility because of his rapidly deteriorating condition. Mrs. Y then joined a caregiver support group, as she missed her husband and was dealing with guilt related to assisted-living facility placement. This group was particularly useful to Mrs. Y in helping her see this placement as a necessary step in her and Mr. Y's health, gave her strategies for making her visits to Mr. Y pleasant, and reinforced her need to remain connected to her friends.

Goals of Interventions

Given the diverse nature of the various caregiving interventions, potential goals for treatment are also diverse. Goals may include increasing knowledge about dementia, improving the caregiver's coping skills with problem behaviors, reducing caregiver burden, depression or anxiety, increasing well-being, reducing the care recipient's problem behaviors or increasing social support. The exact intervention goals should determine the intervention offered.

ETHICAL ISSUES

Caregivers of older adults with dementia face multiple ethical challenges. As the disease progresses, dementia caregivers have to make decisions about various treatment options, end-of-life care, and participation in drug trials. Some of these issues are likely to surface, during treatment, as a source of great distress and agony to some dementia caregivers. A different ethical issue that should be addressed by researchers is the amount of burden put upon dementia caregivers, who often accompany their loved ones during drug trials and as a result become secondary participants in dementia research.

RESEARCH AND EVALUATION

Successful interventions exist, with the majority of caregiving interventions including both a supportive and an educational component. Longer interventions tend to be more effective than shorter ones. More attention should be placed on clinically significant interventions, rather than merely statistically significant ones. There is also increasing interest in establishing consensus on outcome measures for caregivers' interventions. In the past, access to interventions was complicated by the fact that caregivers, who rarely feel they have time to attend classes, were not likely to take the time to attend these sessions. Recent research suggests that interventions can be successfully delivered over the telephone and Internet. However, these

interventions may not meet the need for social support so often present among caregivers of older adults with dementia and their efficacy should be further examined.

REFERENCES

*Alzheimer's Association. 2013 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2013;9(2):208–245.

Aneshensel CS, Botticello AL, Yamamoto-Mitani N. When caregiving ends: the course of depressive symptoms after bereavement. *J Health Soc Behav*. 2004;45(4):422–440.

Ayalon L, Green O. Live-in versus live-out home care in Israel: satisfaction with services and caregivers' outcomes. *Gerontologist*. 2015;55:628–642.

Balducci C, Mnich E, McKee KJ, et al. Negative impact and positive value in caregiving: validation of the COPE index in a six-country sample of carers. *Gerontologist*. 2008;48:276–286.

*Boots LM, de Vugt ME, van Knippenberg RJ, Kempen GI, Verhey FR. A systematic review of Internet-based supportive interventions for caregivers of patients with dementia. *Int J Geriatr Psychiatry*. 2014;29(4):331–344.

*Elliott AF, Burgio LD, Decoster J. Enhancing caregiver health: findings from the resources for enhancing Alzheimer's caregiver health II intervention. *J Am Geriatr Soc*. 2010;58(1):30–37.

Haley WE, Bergman EJ, Roth DL, McVie T, Gaugler JE, Mittelman MS. Long-term effects of bereavement and caregiver intervention on dementia caregiver depressive symptoms. *Gerontologist*. 2008;48:732–740.

Haley WE, Gitlin LN, Wisniewski SR, et al. Well-being, appraisal, and coping in African-American and Caucasian dementia caregivers: findings from the REACH study. *Aging Ment Health*. 2004;8(4):316–329.

Kiecolt-Glaser JK, Dura JR, Speicher CE, Trask OJ, Glaser R. Spousal caregivers of dementia victims: longitudinal changes in immunity and health. *Psychosom Med*. 1991;53:345–362.

Mahoney R, Regan C, Katona C, Livingston G. Anxiety and depression in family caregivers of people with Alzheimer disease: the LASER-AD study. *Am J Geriatr Psychiatry*. 2005;13(9):795–801.

Mausbach BT, Ancoli-Israel S, von Känel R, et al. Sleep disturbance, norepinephrine, and D-dimer are all related in elderly caregivers of people with Alzheimer disease. *Sleep*. 2006;29(10):1347–1352.

*Mausbach BT, Chattillion EA, Roepke SK, Patterson TL, Grant I. A comparison of psychosocial outcomes in elderly Alzheimer caregivers and noncaregivers. *Am J Geriatr Psychiatry*. 2013;21(1):5–13.

Mausbach BT, Dimsdale JE, Ziegler MG, et al. Depressive symptoms predict norepinephrine response to a psychological stressor task in Alzheimer's caregivers. *Psychosom Med*. 2005;67(4):638–642.

Mausbach BT, Mills PJ, Patterson TL, et al. Stress-related reduction in personal mastery is associated with reduced immune cell beta2-adrenergic receptor sensitivity. *Int Psychogeriatr*. 2007;19(5):935–946.

Mittelman MS, Roth DL, Haley WE, Zarit SH. Effects of a caregiver intervention on negative

caregiver appraisals of behavior problems in patients with Alzheimer's disease: results of a randomized trial. *J Gerontol B Psychol Sci Soc Sci*. 2004;59:P27–P34.

Overton E, Appelbaum PS, Fisher SR, Dohan D, Roberts LW, Dunn LB. Alternative decision-makers' perspectives on assent and dissent for dementia research. *Am J Geriatr Psychiatry*. 2013;21(4):346–354.

*Pinquart M, Sörensen S. Spouses, adult children, and children-in-law as caregivers of older adults: a meta-analytic comparison. *Psychol Aging*. 2011;26(1):1–14.

Schulz R, Newsom J, Mittelmark M, Burton L, Hirsch C, Jackson S. Health effects of caregiving: the caregiver health effects study: an ancillary study of the Cardiovascular Health Study. *Ann Behav Med*. 1997;19:110–116.

Smith SA, Bell PA. Examining the effectiveness of the Savvy Caregiver Program among rural Colorado residents. *Rural Remote Health*. 2005;5:466.

*Sörensen S, Conwell Y. Issues in dementia caregiving: effects on mental and physical health, intervention strategies, and research needs. *Am J Geriatr Psychiatry*. 2011;19(6):491–496.

von Känel R, Dimsdale JE, Mills PJ, et al. Effect of Alzheimer caregiving stress and age on frailty markers interleukin-6, C-reactive protein, and D-dimer. *J Gerontol A Biol Sci Med Sci*. 2006;61(9):963–969.

Wong C, Merrilees J, Ketelle R, Barton C, Wallhagen M, Miller B. The experience of caregiving: differences between behavioral variant of frontotemporal dementia and Alzheimer disease. *Am J Geriatr Psychiatry*. 2012;20(8):724–728.

Woods RT, Wills W, Higginson IJ, Hobbins J, Whitby M. Support in the community for people with dementia and their carers: a comparative outcome study of specialist mental health service interventions. *Int J Geriatr Psychiatry*. 2003;18:298–307.

▲ 57.5 Financial Issues in the Delivery of Geriatric Psychiatric Care

STEVE KOH, M.D., AND HELEN H. KYOMEN, M.D.

INTRODUCTION

Financing medical care in the United States is often complex and involves multiple factors. This is especially true for the aging population where complications for providing care are confounded by multiple economic concerns. Despite the passage and implementation of the Patient Protection and Affordable Care Act (PPACA) in 2010, universal health care is not yet a reality and health care cost continues to be a significant percentage of total personal and governmental expenditures. Mental health care continues to suffer from a complicated interplay between federal, state, county, and private payers and consumer demands. Even with the passage of the Mental Health Parity Act (MHPA) and subsequent Mental Health Parity and Addiction Equity Act (MHPAEA), mental health care is often carved out from general health care coverage and patients suffer from inadequate financial support. Older adults with fixed incomes are reliant

on social security and government health care coverage. Many individuals become care givers to loved ones with chronic illnesses, placing additional financial burden on the household. Economic factors such as payment mechanisms and funding sources are key determinants of access to geriatric psychiatric care and health care delivery services, and the actions of providers.

Since the 1990s, both the government and industry have sharpened their focus on health care expenditures and it has become more and more clear that cost effective health care cannot ignore comorbid mental illness. Now widely recognized as a cost multiplier when present as a comorbid condition, the need for improved mental health treatment to achieve better outcomes have been reported by multiple sources. While not perfect, recent changes to the health care system and increased attention to cost savings, outcome measures and population-based medicine allows for innovative models of care. These developments highlight the relationship between patients, providers, and larger systems crucial in providing the best care for older adults.

This chapter will give a brief overview of current demographic trends, the historical background and updates on the Medicare and Medicaid system, out-of-pocket health care costs to older adults, and changes brought on by MHPAEA and PPACA. It will conclude with a discussion of potential future directions for the mental health care of older adults.

DEMOGRAPHIC TRENDS AND THE COST OF HEALTH CARE FOR OLDER ADULTS

Rational planning for mental health care services for the future requires an understanding of the magnitude and diversity of the projected growth patterns in the elderly population in the United States. Furthermore, the differences between sexes and ethnic groups are important to note as there continues to be significant financial resource differences between them. In the mid-1950s, only 1 of every 11 Americans was older than the age of 65. In each of the last several decades, the proportion of elderly has grown substantially. By 2000, there were more than 35 million people older than the age 65, constituting approximately 12 percent of the American population. Since then, the growth of the baby boomers has contributed to more than 76 million adults over the age of 55 in 2010. By 2030, this number will increase to 112 million, accounting for about 30 percent of the total population. By 2050, it is projected that there will be 88.5 million individuals over the age of 65. This is not unique to the United States. The United Nations estimates that by 2050, 21 percent of the total global population will be above the age of 60. Additionally, the number of oldest-old adults (85 years of age and above) continues to increase. It is expected that by 2040, 14.6 million Americans will be among the oldest-old.

The differences in the number of older men versus women will have important economic consequences. The economic effects of lower

educational levels, traditional work roles, and income discrimination experienced by women in young and middle age will be exaggerated as retired women depend on pensions and social security derived from preretirement income. Currently, older women outnumber men by 25.1 million to 19.6 million. While in 2013, the ratio of older males to females was 76:100, by 2050, it is projected to increase to 87:100. This is due to continued increases in life expectancy for both sexes and improved health care. In 2013, the median income of older men was larger compared to women by $29,327 to $16,301. Older men are more likely to be married than women (72 to 46 percent) and 35 percent of women were widows. While 28 percent (12.5 million) of older adults live alone, 8.8 million are women and almost half of older women above the age of 75 live alone. Notably, many older adults continue to be the primary financial support for their household and often are responsible for the care of grandchildren who live with them.

Similarly, ethnic minorities make up a growing proportion of the older-than-65 population. In 1990, approximately 11 percent of people older than 65 in the United States were of nonwhite ethnicity (non-European descent). In 2013, 21.2 percent of older Americans were members of racial or ethnic minority, which are approximately 9.5 million individuals. This is expected to increase to 21.1 million by 2030. Currently, older African Americans make up 8.6 percent of the population while Hispanics represent 7.5 percent. It is expected that older Hispanics will become the largest ethnic minority group by 2030. Life expectancy at birth for Caucasians exceeds that of African Americans, although this gap is decreasing. In 1993, there was a 7.1-year gap. Currently it is 5.4 years for males and 3.8 years for females. This change has been due to decreases in the rates of human immunodeficiency virus, heart disease (in women), homicide, and unintentional injuries. Still, many older African-American individuals have considerably higher rates of poverty, illness, and violent deaths than Caucasian individual in the same age group. The economic and social discrimination of the chronic underprivileged associated with American minority status is exaggerated by the socioeconomic consequences of older age. In this population, accrued social and financial resources are limited; access to preventive and acute health care is poor; and the likelihood of exposure to tobacco, alcohol, other toxins, and violence is higher. In addition, barriers to governmental services such as housing, transportation, meals, and income maintenance are more likely. All of this results in a cycle of medical, economic, and social poverty that materially affects the quality of life and has a substantial impact on the provision and financing of geriatric psychiatric care for elderly minorities. Further compounding this situation is the fact that there are significant state-to-state discrepancies in ethnic minority life expectancies.

Retirement from employment has important effects on economic status,

access to health care, and insurance. Retirement reduces personal income and increases reliance on government assistance and retirement plans. Reduction in income after retirement threatens stability of social class for older people and increases the likelihood that some people will experience economic poverty for the first time, during late life. These poverty prevalence rates are noticeably higher for women and minorities. In 2012, the major sources of income for older Americans were distributed between the following: social security (86 percent), income from assets (51 percent), private pension (27 percent), government employee pensions (14 percent) and earnings (28 percent). Social security constituted more than 90 percent of beneficiaries' income. In 2013, over 4.2 million older adults were below the poverty level based on income alone. If regional variants in cost of living, noncash benefits received, and nondiscretionary expenses are taken into account (the so-called Supplemental Poverty Measure), the poverty level of older persons reaches 14.6 percent. A major contributor to this increase is high medical out-of-pocket expenses. This cost is magnified in geriatric mental health care and in long-term care, where out-of-pocket copayments, deductibles, and direct purchase of services account for a substantial component of total costs.

As long-term financial stability is questioned, many individuals are electing to delay retirement or continue to work part-time. Since the Great Recession (spanning from 2007 to 2009), the average retirement age increased. Thirteen percent of Americans continue to work past 65 years of age and 8 percent continue to work part-time. Even though median income peaks between the ages of 45 and 54, between 2012 and 2013, those above age 65 saw percentile increases in their median income. This indicates that more Americans are working into their 60s. Yet, many indicate that there is added pressure to retire despite their worry about long-term security. The single most important factor for concern about retirement is health care cost. Retirement directly affects health insurance premiums, the availability of certain insurance plans, and access to some health care delivery systems.

Overall fiscal impact on health care by the older population is undeniable. While their earnings are decreasing, health care needs of the growing older population continue to rise. The cost of health care in the United States is far greater than other industrialized nations. In 2011, the United States' health care expenditure per capita of $8,508 was about $3,000 more than the next largest health care expenditure per capita country. Older adults make up significant portion of this high cost. Almost half of all health care spending is consumed by just 5 percent of the population, and these are focused in age groups below 5 and above 65. In fact, older adults account for more than 36 percent of total health care cost. Those between 65 and 70 years of age represent 29 percent, and those above the age of 80 account for 14 percent of the top 5 percent of spenders.

Among those with Medicare–Medicaid coverage, the elderly and disabled account for more than 70 percent of Medicaid spending despite representing just 25 percent of the Medicaid population. The top five most expensive health conditions to treat are heart disease, cancer, trauma, mental disorder, and pulmonary conditions. Twenty-five percent of the US population reportedly has one of the top five chronic conditions (mood disorder, diabetes, heart disease, asthma, and hypertension). The rising Medicare and Medicaid cost can partially be explained by the aging population and mental illness accounting for a significant portion of the cost. All of this has led to attempts to restructure Medicare and social security and to even raise the retirement age in the United States.

HISTORICAL BACKGROUND

Historically, government policy has been important in shaping the status of psychiatric care for older people. During the 19th century, financial responsibility for the care of the chronically mentally ill and people with major neurocognitive disorder was transferred from local to state authorities, resulting in the replacement of asylums and almshouses by the massive state hospital systems of the first part of the 20th century. Ultimately, deinstitutionalization in areas of limited community resources resulted in the transfer of a substantial number of older people with severe and persistent mental illnesses from state hospitals to community nursing homes with sparse psychiatric resources. Beginning in approximately 1965, reimbursement for nursing home care under Medicaid effectively shifted some of the financial responsibility for this population from states alone to the states and federal government.

The introduction of Medicare and Medicaid in the mid-1960s created a private market for older adult health care. Private practitioners and institutions were reimbursed fee-for-service at market prices by these insurance programs. Adequate financial returns for hospital-based services, reinforced by the growth of commercial indemnity insurance, drove the development of for-profit hospital care, particularly in freestanding psychiatry. Simultaneously, severe limitations in reimbursement for ambulatory psychiatric services under Medicare impeded the growth of private practice interest in geriatric psychiatric care while creating imbalanced incentives for the use of hospital resources. Federal-level sensitivity to these new expenditures and social concerns about inappropriate care led to the nursing home reform provisions of the latter part of the 1980s and to the formulation and implementation of a set of guidelines for the use of psychiatric interventions in that environment. The health care cost reform of the last generation has similarly reshaped geriatric psychiatric service delivery. In the recent past, Medicare was more attractive than some commercial payers to hospitals and individual providers. Nonpsychiatrist mental health specialists, newly eligible for

participation in Medicare and other payment programs, created the availability of a greater diversity of providers, with psychotherapeutic services much more broadly available to the elderly.

However, as the population ages and health care cost continues to rise, models to decrease expenditures, increase cost-sharing and improve access to care have dominated health care discussions. For geriatric mental health care, the focus has been on achieving mental health care parity, developing cost-effective, innovative programs, and providing access to care to growing population with limited providers. As will be noted later, inequality in mental health care coverage has made the treatment of mental illness in older adults difficult and the impact of MHPAEA can dramatically improve the care received by older adults. Furthermore, the impact of changes to Medicare and Medicaid by PPACA is still being examined and is likely to have a lasting effect on psychiatric care for older adults.

MEDICARE SYSTEM

The Medicare program was established in July of 1966 to provide medical care benefits for older people participating in the Social Security program. Since then, a number of modifications were made to the Medicare program, in both the Hospital Insurance program, Part A, and the Supplemental Medical Insurance program, Part B. The Medicare Prescription Drug, Improvement, and Modernization Act of 2003 provided the most extensive changes to Medicare since its initiation.

In 1972, Medicare was expanded to cover younger people with disabilities and older adults outside of social security who were willing to pay a premium. In 1973, the program was further expanded to provide medical coverage for people with end-stage renal disease. In 1997, the Balanced Budget Amendment of 1997 (BBA 1997) established Medicare Part C, or Medicare+Choice. Medicare Part C included Medicare managed care plans (such as HMOs, provider-sponsored organizations [PSOs], and preferred provider organizations [PPOs]); Medicare private, unrestricted fee-for-service plans; and medical savings accounts. In the 2003 revision, Medicare was extensively updated to include: (1) expansion and renaming of the Medicare+Choice program as Medicare Advantage; (2) prescription drug programs and benefits, including a new program called “Medicare Part D”; (3) medication therapy management programs for those with multiple chronic conditions, multiple prescriptions, and likely high drug expenses; (4) coverage of certain screening and disease management programs such as initial physical examinations and cardiovascular disease and diabetes screening; (5) changes in reimbursements to hospitals and physicians, including ongoing refinement of the prospective payment system for inpatient psychiatric care; (6) improved rural health care provisions; (7) a demonstration project for recovery audits to identify underpayments or overpayments and collect overpayments; and (8)

changes in the acquisition systems for durable medical equipment and other products to provide for a competitive bidding process. In 2013, Medicare covered more than 52 million people, or almost 17 percent of the American population. Over 83 percent of this Medicare-covered population in that year was 65 or older.

Growth in the older adult population, advances in medical technology, more extensive use of pharmaceutical treatments, and other increases in the cost of health care have contributed to the extraordinary growth in health care expenditures, both in the Medicare program and otherwise, that has occurred since the mid-1990s. Although health care costs have continued to grow in recent years, especially among the elderly, various efforts to contain Medicare costs have been successful in slowing health care inflation rates. Each of these changes has profoundly affected access to appropriate psychiatric care for older people, and each has impacted the scope and configuration of geriatric mental health specialty care. Aggregate Medicare cost reductions have resulted from decreases in reimbursement levels of hospital and physician providers, increased requirements for out-of-pocket payments by Medicare beneficiaries, and substantial shifts from indemnity insurance to managed care. In addition, as part of the Deficit Reduction Act of 2005, the Centers for Medicare and Medicaid Services (CMS) (replacement agency for the Health Care Finance Administration) has been allowed, starting in 2008, to begin adjusting payments for "never events." "Never events" are "serious and costly errors in the provision of health care services that should never happen" and include hospital-acquired infections, death or serious disability related to a medication error, and death associated with falling while being cared for in a health care facility. "Never events" lead to increased health care costs due to the need to treat the consequences of the error. Additional Medicare cost reduction has been attributed to the Affordable Care Act (ACA) of 2010 and sequestration, which started in 2013. Still, Medicare spending represented 20 percent of national health spending in 2013 and grew 3.4 percent to about $586 billion, but was a slowdown from a growth of 4.0 percent in 2012. This slowdown was thought to be due largely to slower enrollment growth and impacts of the ACA and sequestration.

Medicare was designed to provide coverage for the treatment of acute conditions, and this has been the primary orientation of the program over the years. Medicare had not covered most preventive and pharmaceutical services since the program's inception. The 2003 law provided for some change in this focus with coverage of some preventive services and the addition of pharmaceutical benefits. Long-term care and dental services continue to be excluded from traditional indemnity Medicare. Other services, including mental health care, had been subject to limitations in payment and substantial copayments. However, the Wellstone-Domenici Mental Health Parity and Addiction Equity Act of 2008, enacted into law

on October 3, 2008 and effective for most health insurance plans on January 1, 2010, was designed to change this. This law required that group health plans (with over 50 employees that offer both medical/surgical and mental health/substance-use benefits) provide coverage for treatments of mental health/substance-use disorders that is no more restrictive than the coverage for treatments of medical/surgical conditions. Equity coverage is applied to financial requirements as well as treatment limitations. Format guidance and rule-making as to how this law is implemented continues to evolve.

The Medicare program, in summary, has several components: Hospital Insurance, or Part A, covers inpatient care, limited home care, skilled nursing facility care, and hospice services. Part B, the Supplemental Medical Insurance program, covers services provided by physicians and other professional providers, durable medical equipment, ambulatory medical laboratory tests, some outpatient hospital care, and group therapy services of psychiatric hospitals and comprehensive outpatient rehabilitation facilities. Part C includes Medicare-managed care plans (such as HMOs, patient safety organizations [PSOs], and PPOs); Medicare private, unrestricted fee-for-service plans; and medical savings accounts. Under Medicare Part D, effective starting January 1, 2006, Medicare beneficiaries have been eligible to obtain prescription drug coverage by enrolling in a private prescription drug plan or in an HMO or other plan that accepts Medicare capitation payment and offers a drug benefit.

Hospital Insurance (Part A)

Americans older than age 65 who participate in the Social Security program are entitled to Part A Medicare coverage. The Part A Hospital Insurance Trust Fund is supported by social security payroll taxes paid by employers, employees, and the self-employed. In 2013, 52.3 million people had Medicare coverage, and the Hospital Insurance Trust Fund had assets of $280.5 billion.

Medicare coverage for inpatient care is limited to relatively brief inpatient stays, presumably for the stabilization of acute symptoms. Part A provides coverage for episodes, which are time-limited “spells” of illness. A spell is an episode of inpatient care that starts with hospital admission and ends after the first period of 60 consecutive days outside the hospital after hospital discharge. If a person is discharged and readmitted numerous times during a given episode of illness and has not been out of hospital for 60 or more consecutive days, he or she is considered to have remained within the same “spell.” For treatment of any condition in a general hospital, there is no limit on the number of spells or days covered within an individual’s lifetime. However, no more than 150 inpatient days are covered during each episode.

The first 60 days of each spell are fully paid after a deductible equal to

the average cost of 1 day of hospitalization has been met. The next 30 inpatient days are paid at the same rate less a 25 percent copayment. The last 60 days of covered hospitalization requires a 50 percent copayment. Days 91 to 150 are identified as "lifetime reserve days." Medicare pays for those days during any single spell of illness but only once in a beneficiary's lifetime. For example, if a patient requires five hospitalizations totaling 110 days without a period of discharge of at least 60 days, he or she would have 40 days of remaining coverage in his or her lifetime for future episodes that lasted more than 90 days. Individuals may save these days for future hospitalizations and use other resources to pay for any part of the costs of days 91 to 150.

At the time that Medicare was first signed into law, there were large populations of older adult patients in state psychiatric hospitals. Federal legislators feared at that time that the enactment of Medicare would shift the burden of care for that population from state-supported resources to acute care facilities reimbursed with federal monies. As a consequence, Medicare coverage for inpatient treatment in freestanding psychiatric hospitals was limited to a total of 190 days over the course of a person's lifetime. The law provides that if an individual becomes eligible for Medicare during the course of a psychiatric hospitalization, the Medicare intermediary may deem that less than the full 150 days allocated for that spell of illness will be covered. Given improvements in treatment and increased life expectancy, this limitation can easily create undue hardship and compromised continuity of care. For example, a person with a recurrent major affective disorder who experiences several depressive episodes requiring inpatient care after age 65 could easily exhaust the lifetime freestanding psychiatric hospital benefit.

Part A provides coverage for nonacute services, including skilled nursing facility care if these services are used briefly and for "remediable" conditions. For a beneficiary to obtain benefits for skilled nursing facility care, he or she must have been hospitalized for at least 3 consecutive days for an acute condition, and skilled nursing facility admission must follow within 30 days of hospital discharge. Medicare may fully cover up to 20 days and partially cover days 21 to 100 of skilled nursing facility care for each episode of illness. Similarly, coverage for home health services is limited to care for acute and remediable conditions.

From 1966 until 1983, Medicare Part A paid for hospital care using a retrospective cost-based reimbursement system. This system compensated providers on a fee-for-service basis for all of those costs deemed allowable under Medicare without an incentive for utilization management. In 1982, after years of rapid and sustained growth in Medicare expenditures for inpatient care, the Reagan administration signed into law the Tax Equity and Fiscal Responsibility Act (TEFRA). TEFRA created incentives to reduce hospital payments by setting targets for operating costs, and it provided

supplemental payments for compliance with specified objectives. Reimbursement was transitioned from a retrospective cost-based approach to a payment system designed to reflect patient mix, geographic location, and training costs. This TEFRA mandated system is known as the prospective payment system.

The prospective payment system established by TEFRA uses diagnosis-related groups to classify patients who require similar care into 23 major diagnostic categories and assigns each Medicare patient discharge to one of 468 diagnostic groups. These groups are derived from principal and secondary diagnoses; procedures rendered; and, to a lesser degree, age, sex, complications, and discharge status. The diagnosis-related group system was revolutionary in creating a "case-based" approach to reimbursement. Under the prospective payment system, hospitals are paid a fixed amount for each case according to the corresponding diagnosis-related group regardless of the resources that are consumed during the hospitalization. If patients require extraordinary resources or extremely prolonged hospital stays, Medicare classifies the patient as an "outlier" and provides additional payment to the hospital, although at a per diem rate considerably below actual costs. This set of conditions creates an incentive for hospitals to provide efficient care and create an infrastructure of ambulatory services that can shorten or substitute for hospitalization.

Fourteen of the original diagnosis-related groups applied to the treatment of psychiatric disorders. Several studies had indicated that, although diagnosis-related groups are modestly effective in predicting use for medical and surgical diagnoses, their ability to predict use for psychiatric diagnoses is poor. Accordingly, the Health Care Finance Administration, the federal agency responsible at that time for Medicare/Medicaid fiscal management, exempted freestanding psychiatric hospitals and distinct psychiatric units in general hospitals from the diagnosis-related group system until a fairer system could be ascertained. On November 15, 2004, a newly designed Medicare prospective payment system for inpatient psychiatric services with updated psychiatric diagnosis-related groups was established by the CMS. Under this new system, a federal per diem rate was established, providing for adjustments for changes in the hospital market basket index, age, comorbidities, and psychiatric diagnosis. In this system, 15 psychiatry-related diagnosis-related groups were established, with weights ranging from 22 percent above average for diagnosis-related group 424 ("Operating Room Procedure with Principal Diagnosis of Mental Illness") to 12 percent below average for diagnosis-related group 523 ("Alcohol/Drug Abuse or Dependence without Rehabilitation Therapy and without Complications and Comorbidities").

Until the new prospective payment system for inpatient psychiatric services was initiated, freestanding psychiatric hospitals and psychiatric

units in general hospitals were reimbursed retrospectively by Medicare. Ever since TEFRA was implemented in 1983, payment for treatment in psychiatric units had been capped at a rate established for each facility based on use during its first fiscal year of operation after October 1, 1983. Until 1998, if the actual cost of care per case was greater than the capped rate, the hospital absorbed the loss. However, if the cost was less than the capped rate, a small incentive payment was retained by the hospital. The Balanced Budget Act (BBA) 1997 modified and limited these reimbursements by establishing differential caps, with or without wage adjustment, and also provided less generous bonus and relief payments for prospective payment system–exempt facilities. The Balanced Budget Refinement Act of 1999 (BBRA 1999) addressed some of the flawed policy and excessive payment reductions resulting from BBA 1997. The BBRA 1999 allowed for wage adjustments at the 75th percentile limit established by the BBA 1997 for prospective payment system–exempt facilities that received their first Medicare payment before October 1, 1997, and also increased bonus payments for psychiatric hospitals. The BBRA 1999 also stipulated that a new Medicare prospective payment system for inpatient psychiatric services was to be initiated by April 1, 2004. The final rule establishing the prospective payment system for inpatient psychiatric facilities was published on November 15, 2004, in the Federal Register. The main features are the following:

1. A base per diem prospective payment system rate of $575.95 was determined in 2004, effective for the first payment period from April 2004 through June 2005. Subsequent payment periods, starting on July 1 of each year, had adjustments to the base rate determined by estimated changes in the hospital market basket index. In addition, the base rate was further adjusted at the patient level for age, comorbidities, and psychiatric diagnoses. For patients 65 years of age and older, payments increased by 10 to 17 percent. In 2015, the base per diem rate was $728.31.
2. There were adjustments to increase payments for services provided to patients with one or more of 17 designated comorbid conditions, as well as adjustments for psychiatric diagnoses (one of 15 psychiatry-related diagnosis-related groups mentioned previously). Rates were also adjusted to reflect area wage differences, rural location, and the presence of teaching programs.
3. Rates were 19 to 31 percent higher for the initial day of admission, depending on whether the inpatient facility had or was a psychiatric unit in an acute care hospital with a qualifying emergency department. There were smaller adjustments for subsequent days until the 10th day, when these auxiliary payments ended and adjustments to decrease payment below the base per diem rate from 1 to 5 percent would occur through the

21st day. After the 21st day, the rate would be 8 percent less than the base per diem rate. The payment system also included "outlier payments" for hospitalizations resulting in significantly higher costs.

4. Once the cost of an admission exceeded a designated threshold (initially $4,200) above the diagnosis-related group payment, Medicare paid part of the additional costs from an outlier pool of funds equal to 2 percent of the estimated total payments for inpatient psychiatric care. Of note, the payment system did not offer adjustments for disproportionate share status, patient gender, and distinct units, although these elements were all considered by CMS.

Supplemental Medical Insurance (Part B)

Medicare Part B is a medical insurance program that is voluntary and requires a monthly premium payment ($104.90 in 2015) for participation. Approximately 97 percent of Medicare Part A beneficiaries also participate in the Part B program. Part B is supported by the Supplementary Medical Insurance Trust Fund, which includes general revenues from the Treasury, as well as premiums. Premiums cover approximately 25 percent of the total cost of the program. Part B benefits are subjected to an annual deductible ($147 in 2015) as well as a 20 percent copayment. Since enactment of the MHPAEA, group health plans and health insurance issuers are required to ensure that financial requirements (such as copays, deductibles) and treatment limitations (such as visit limits) applicable to mental health or substance use disorder benefits are no more restrictive than the predominant requirements or limitations applied to substantially all medical/surgical benefits. MHPAEA supplements prior provisions under the MHPA, which required parity with respect to aggregate lifetime and annual dollar limits for mental health benefits.

From 1984, when TEFRA and prospective payment were initiated, until 1992, the rate of growth of Medicare physician costs was more than double the rate of growth for inpatient hospital costs. Policy analysts attributed this growth to uncontrolled price escalation and substantial increases in the number and intensity of services consumed. In 1989, inflation in physician-related costs and mounting evidence that physicians' behaviors in treating complex medical illnesses were influenced by the high prices of the procedures they performed led to a sweeping reform of Medicare Part B through the Omnibus Budget Reconciliation Act of 1989 (OBRA-89).

OBRA-89 created standards for expenditures and procedure volumes in an attempt to bring the total cost of physicians' services under control. The law required the secretary of the U.S. Department of Health and Human Services (HHS) to identify specific volume performance standard growth rates for surgical, primary care, and nonsurgical nonprimary care services at the beginning of each fiscal year. These standards must be related to medical inflation and changes in the size of the Medicare population as well

as to changes in service volume and intensity. At the end of each year, the secretary must compare actual growth and expenditures with this volume performance standard and determine any changes in fee schedules for the following year.

OBRA-89 also created a national uniform Medicare fee schedule to replace local methods that were traditionally used by intermediaries to set fees for each procedure and for each provider. In 1992, a Resource Based Relative Value Scale (RBRVS) was developed to determine relative values of approximately 7,000 physician procedures and services. In this scale, the relative value of each of these services is derived from three components: Work, practice, and malpractice expenses, and the opportunity costs of postgraduate specialty training required. The proportion of each component in determining the relative value is based on a weighted average of specialty-specific practice expense and malpractice data. In an effort to change provider behavior and reduce incentives to perform highly reimbursed technical procedures, the work component is derived from estimates of the time intensity and stress associated with delivery of the procedure. Approximately 60 percent of the value of each fee is adjusted for local costs. The product of each relative value and a monetary conversion factor determine the fee for each service.

OBRA-89 was a landmark for behavioral health in its provision for direct reimbursement of services provided by psychologists and social workers. It also set criteria for nonphysician providers regarding consultation with a physician. Nonphysician providers must document that a patient has been informed that consultation with the patient's primary care physician (PCP) is desirable for the PCP to evaluate medical conditions that could contribute to a patient's well-being. The provider must also document any written or verbal communications with the PCP about treatment unless the patient specifically prohibits such contact. The law does not require consultation with a psychiatrist. The law stipulates that reimbursement of services provided by psychiatric nurses and other therapists requires direct physician supervision.

In general, direct patient contact is required for any reimbursable service. Consultation or other services rendered by telephone and renewals of prescriptions that do not include direct evaluation of a clinical condition are not reimbursable. Medicare Part B allows billing for services required to obtain clinical information from relatives or close associates of unreliable or uncommunicative patients. Family counseling services are reimbursable only when their purpose is to facilitate patient care.

There are two ways in which providers may participate in Part B of the Medicare program:

1. A provider may accept assignment and agree to accept the applicable Medicare fee as payment in full. Medicare's fiscal intermediary pays the

provider the approved amount less copayment and deductible directly. The provider is legally required to bill the patient directly for the copayment and the deductible.

2. A provider may choose not to accept assignment. Providers who do not accept assignment may charge patients up to 115 percent of the approved Medicare fee. Medicare reimbursement for providers who do not participate in Medicare Part B is limited to 95 percent of the approved fee less deductibles and copayment. The provider may then bill the patient directly for the balance of charges up to the 115 percent limit.

All providers must accept assignment from Medicare beneficiaries who also participate in the Medicaid program. Assignment must be accepted for all Medicare recipients who are at or below the federal poverty level regardless of whether they have Medicaid as a secondary payer. In an increasing number of states, medical licensure authorities and insurance commissions require that all physicians who participate in the Medicare program accept assignment for patients seen within that state.

Medicare Risk in Managed Care

Since the early 1970s, Medicare has encouraged older Americans to enroll in capitated health plans. Since that time, HMOs have been able to contract directly with Medicare. Between 1985 and 1997, Medicare payments to risk-contracting HMOs for aged and disabled beneficiaries were based on adjusted actuarial estimates of the per capita cost of such services incurred in a beneficiary's county of residence, using adjusted average per capita cost (AAPCC) methods. The AAPCC had been based on the individual's county of residence, age, and sex and whether a patient is covered by Medicaid, lives at home or in an institution, and works or is retired or disabled. These county per capita payment rates, adjusted for the average beneficiary, have been published annually as the County Rate Book.

HMOs, although allowed to earn "normal profit margins," must use savings above a predetermined rate to provide extra services for elderly members. For many years, to penetrate a local market, HMOs frequently waived premiums and offered to include usually uncovered services in the benefit package to induce enrollment. Unlike traditional Medicare, HMOs allow access to primary care and other preventive services in the absence of a primary diagnosis. In addition, managed care plans usually have provided coverage for some portion of outpatient pharmaceutical costs.

The nature of HMO risk-contracting with Medicare provides an incentive for HMOs to market to healthy older adult populations who are likely to consume services valued at considerably less than the 95 percent of AAPCC criterion used to set capitation rates. Indeed, healthy populations, even if older than 65 years, yield substantially greater profit margins for HMOs, leaving sicker, more disabled people to be cared for via

direct payments from CMS in the traditional Medicare program. Several studies have indicated that capitated plans generally enroll healthier and less costly beneficiary populations than those who remain in fee-for-service Medicare. In one study, in the 6 months preceding enrollment, new members of capitated plans had health care costs that were 35 percent lower than the average costs of patients in traditional fee-for-service Medicare. Some plans profit through marketing to "younger older adults" (people aged 65 to 70) in settings in which there is a greater likelihood of health and independence. Studies indicate that Medicare has overpaid some capitated plans by 5 to 20 percent because selection factors resulted in enrollees with better health, on average.

To reduce financial incentives rewarding the enrollment of healthy persons and to pay more fairly for the care of frailer populations, the BBA 1997 required Medicare to link payments to HMOs to the health status of their beneficiaries by the beginning of the year 2000. In this system, frail enrollees are eligible for higher payments than healthier cohorts are. This should encourage plans to discontinue avoiding the enrollment of more chronically ill populations and to bring a broader base of the general older adult population into managed care. In enacting Medicare Part C in 1997 to create Medicare+Choice, a new methodology was introduced to pay organizations that enter into Medicare Part C contracts. The equivalent of the County Rate Book is based on the greatest of three amounts: (1) a new blended payment rate based on a calculation method combining local and national rates in setting county rates, (2) a new minimum specified rate amount (e.g., $367 per month per enrollee in 1998), or (3) an amount that is based on a 2 percent increase over the prior year's rates, with the 1997 County Rate Book serving as the baseline.

The percentage of beneficiaries enrolled in Medicare managed care increased steadily into the 1990s. Whereas only 3.6 percent of older Americans were enrolled in HMOs in 1990, this proportion increased to 10 percent in 1996 and to more than 15 percent in 1998. Since 2001, enrollment has been declining due to the departure of many health insurance programs from the Medicare managed care market due to reimbursement rates that are increasingly inadequate to cover medical utilization costs. The Medicare+Choice program, a major component of the BBA 1997, renamed in 2004 to Medicare Advantage, is the Medicare managed care component. It is designed to make it easier for various kinds of organizations, including PSOs and PPOs, to become participants in Medicare managed care. Medicare Advantage has been organized to encourage greater competition among health plans enrolling Medicare beneficiaries.

Managed care plans caring for Medicare recipients have applied the same paradigms to behavioral health care that they have used for their younger commercial populations. Many of these plans, perhaps a majority,

"carve out" the behavioral health benefit and capitate the risk for these services to managed care organization (MCO) vendors. Most of the MCOs are large for-profit entities that function as psychiatric and substance abuse insurers in numerous markets regionally and nationally. MCOs generally establish local networks of ambulatory and inpatient providers; establish fee-for-service pricing with those providers; provide medical management, care management, utilization review, and credentialing functions; and also pay claims. These MCO systems are generally very effective in reducing inpatient use. This effectiveness allows them to negotiate significantly discounted prices for services and to limit the size of available networks, thereby reducing the management effort required to control service use.

A review of the credentialing and procedure manuals for a group of the largest MCOs nationally found virtually no mention of requirements for specialty training or expertise in the care of older adults or for special consideration of medical comorbidity or frailty. There were no requirements for services to be sited at locations in which older adults reside or where access to care may be facilitated for frail or disabled people. This review suggests that the current model for delivery of managed behavioral health services to older adult Medicare beneficiaries is unlikely to reflect recent advances in behavioral neuroscience. Similarly, many carve-out programs exclude or limit comprehensive assessment, neuropsychological testing, and other routine components of major neurocognitive disorder evaluation.

A number of not-for-profit entities have developed the capacity to be included in full-risk agreements with their medical and surgical colleagues. This has allowed an emergence of a new model of behavioral health care, carving in services while locating providers on site or adjacent to primary care delivery. These programs are serving considerably smaller populations than most MCOs, and differences in actual outcomes have yet to be demonstrated.

Managed care continues to be a priority in Medicare policy. However, reductions in payments to managed care entities by Medicare coupled with complex paperwork requirements have prompted many managed care plans to leave the Medicare program in recent years. The tension between the financial effects of the BBA 1997 and BBRA 1999 laws and government efforts through the 2003 Medicare Advantage provisions to attract beneficiaries to managed care plans is likely to persist over the course of the coming years.

Insurance Coverage Supplementing Medicare

More than 90 percent of older Medicare beneficiaries have supplemental Medicare coverage for service costs not covered by Medicare directly. These supplemental coverages include employer or union-sponsored retiree health plans, Medigap policies, Medicare+Choice/Medicare Advantage,

and dual Medicaid–Medicare coverage.

In July 1991, Congress enacted mandatory federal Medigap standardization requirements as part of the OBRA of 1990 (OBRA-90). It provided requirements for 10 basic policies. All of these policies must include payment of the Medicare Part A hospital and Part B provider coinsurances. In addition, all policies must pay any applicable 50 percent copayment for outpatient psychiatric services. More comprehensive policies may elect to cover deductibles, skilled nursing facility coinsurance, outpatient prescription pharmaceuticals, and preventive in-home services. Medigap regulations provide incentives for the use of PPOs by Medicare recipients. This allows insurers to offer lower premiums to members who agree to use a PPO for all of their care. An increasing number of states have authorized insurers to offer this option.

The Medicare Prescription Drug, Improvement, and Modernization Act of 2003 impacts Medigap policies that have a drug benefit. People currently enrolled in a Medigap plan without a drug benefit are not affected. However, those who have a Medigap policy with drug benefits have two choices: (1) They may keep their current Medigap plan with drug benefits. However, if they later decide to enroll in a Medicare Part D plan, a late enrollment penalty may be charged if the Medigap policy is not equivalent to Medicare Part D. (2) They may enroll in a Medicare Part D plan. Those who choose to do this can either enroll in another Medigap plan without drug coverage to cover Medicare's cost sharing for other services or keep their current Medigap plan but drop its drug benefits and pay a lower premium. People who reach age 65 after January 1, 2006, are able to enroll in Medigap policies without drug benefits. After January 1, 2006, people were not able to purchase a supplemental policy to help offset their deductibles and copayment for drug policies. Regarding employer-sponsored retiree plans, the government pays 28 percent of the costs that companies incur for each retiree, to encourage companies to maintain their drug coverage for retirees; this amounted to between $250 and $5,000 per retiree in 2006.

On April 16, 2015, the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) was authorized. Among various actions, this replaced the Sustainable Growth Rate formula (never actually put into action due to multiple short-term Congressional "patches" since 2002). Starting in 2020, it will affect insurance coverage supplementing Medicare by decreasing funding to supplemental Medicare plans, reducing "first dollar" coverage of supplemental Medigap plans.

MEDICAID SYSTEM

Enacted in 1965 as a social insurance program to pay for the medical care of Americans with limited financial means, Medicaid is funded through a combination of federal and state matching funds. To receive these

matching funds, states must provide a basic benefit package, including hospital, physician, and nursing facility services, as well as a wide range of supportive services. The program allows states to impose restrictions on the services that they fund and on the level of reimbursement. Medicaid is a means-tested and needs-based program thus, an individual must meet certain criteria to be eligible. The eligibility varies widely from state to state, but the most important criteria are poverty and financial resources. With PPACA, the eligibility was increased to 133 percent of poverty line. The federal government is required to pay for the cost of Medicaid expansion at 100 percent rate till 2016, 95 percent in 2017, 94 percent in 2018, 93 percent in 2019, and 90 percent starting in 2020 and thereafter. Medicaid expansion does not apply nationwide, as states are able to reject the PPACA Medicaid expansion. Therefore, as of this writing, Medicaid coverage continues to depend on the state.

For older Americans with mental illness, Medicaid is important in two ways. First, it is the most important payer for long-term care services. The nursing home is an important site of care for older adults with severe and persistent mental illness, including major neurocognitive disorder, schizophrenia, and depression associated with medical illness and disability. Deinstitutionalization from state facilities shifted the financial burden of care for many people with these diagnoses from the states (in state hospitals) to cost sharing with the federal government via its partial support of Medicaid. The nursing home reform provisions of the OBRA of 1987 (OBRA-87) require preadmission screening for nursing home applicants to determine if active psychiatric illness is present. If active treatment for an identified psychiatric disorder cannot be provided appropriately in the nursing home, alternative referral for this care is mandated. Nursing homes must demonstrate efforts to provide appropriate care for behaviorally disturbed people, and appropriate training in management of difficult behaviors is required of all staff members of long-term care facilities participating in the Medicaid program. Currently, Medicaid continues to be the major coverage option for long-term services and supports. Its scope is broader than Medicare and encompasses services delivered in hospitals, institutions, home, and communities. As Americans continue to age, it is estimated that over 70 percent of those above the age of 65 will need long-term services and those above the age of 85 are four times more likely to need long-term services. In 2013, 51 percent of the estimated cost of $310 billion for long-term services was covered by Medicaid (19 percent was out-of-pocket). The second area of importance is the dual Medicare–Medicaid eligible patients. These are individuals who not only meet eligibility criteria for Medicare but also Medicaid (often due to financial hardship or disability). Currently, 8.3 million Americans meet this dual eligibility and 62 percent of Medicaid expenditure or about $91.8 billion were for long-term care. 4.6 million are low-income seniors. With

the expansion of Medicaid, it is expected that this number will increase. The 2015 dual eligible standards based on percentage of federal poverty level amounts to $1,001 for individuals and $1,348 for couples/families with slightly higher rates for Alaska and Hawaii. Asset limits are $7,280 and $10,930, respectively. Even with Medicaid expansion, the dual eligible individuals do not constitute the majority of the Medicaid population. However, they account for over 70 percent of Medicaid spending.

OUT-OF-POCKET COST

As discussed above, maintaining one's health is the single most expensive cost during old age. Despite governmental support in the form of Medicare and Medicaid, many older Americans bear out-of-pocket costs in health care. Part of the cost is due to the simple fact that increases in federal budget and growth in Medicare spending has prompted political and policy interest in raising cost-sharing obligations of older adults. Out-of-pocket costs can include the following: monthly premiums for physician services under Medicare, prescription drug coverage under Medicare, custodial long-term care services and supports, routine dental care (including dentures), routine vision care or eyeglasses, and hearing exams and aids. To cover these expenses, many opt to get supplemental insurance to cover the cost not covered or only partially covered by Medicare (Medigap, employer or union-sponsored retiree coverage or employer-sponsored insurance for those who are still working). Those who cannot obtain supplemental insurance due to the high cost and are below the federal poverty level can receive Medicaid supplements. However, even with Medicaid expansion under PPACA, some older adults with low income do not qualify for Medicaid due to their saved assets. Although there have been measurable reductions in recent years in per capita out-of-pocket Medicare beneficiary spending due to the growth of managed care and cost containment measures, the effects of out-of-pocket payments on older adults, particularly those with limited resources, continues to be problematic.

Recent studies give some estimates of out-of-pocket expenses for health care. In 2010, Medicare beneficiaries spent $4,734. Premiums and supplemental insurance accounts for about 42 percent, and of the remaining, long-term facility costs are the highest accounting for about 18 percent. This is followed by medical providers/supplies (14 percent), prescription drugs (11 percent) and dental care (6 percent). The cost rises with age and there are notable differences between men and women as they approach and pass age 85. For women, who pay more out-of-pocket expenses then men, the difference can be attributed to rising long-term care cost. Chronic illness such as Alzheimer disease and Parkinson disease that are associated with aging have profound impacts on rising health care cost ($8,305 and $5,841 respectively). Additionally, the out-of-pocket cost

can increase with those with mental illness requiring more Medicare-covered services, as well. As noted in the Mental Health Parity section, coverage of mental health treatment is not equitable, hence older adults with chronic mental health issues requiring frequent or long psychiatric hospitalizations can face catastrophic costs when obtaining adequate treatment.

Since Medicare and Medicaid do not cover different levels of long-term care, privately paid long-term insurance plans are available and add to additional personal cost. Prescription drug continues to be the fastest-growing segment of health care cost for older Americans. To help with this, by 2020, PPACA will close Medicare Part D's "donut hole" and will reduce the cost by an average of $2,000 per year. Furthermore, PPACA restructures Medicare Part B and D premiums to reduce cost for those in low-income brackets and increase contributions from the wealthy. These measures, as well as coverage of preventative care, may work to decrease overall out-of-pocket costs.

MENTAL HEALTH PARITY

A significant financial barrier to mental health care has been inequality of coverage of mental treatment by insurers. There has been a long history of benefit inequality between mental health and general health. Historically, the inequality was shown by higher mental health insurance premiums, limitations on mental health care coverage, and carving out mental health care from general health insurance. Importantly for older adults, lack of mental health care parity in Medicare meant that outpatient mental health treatment services required up to a 50 percent copayment while general care required just 20 percent. This posed significant barriers for older Americans to access mental health care.

While President Kennedy's administration first required Federal Employees Health Benefits Program (FEHBP) to cover psychiatric illness at the same level as general medical care, by 1975, even FEHBP scaled back on mental health coverage. Efforts in the 1980s to improve private insurance benefits were largely driven by individual states and no consistent parity level care or financial coverage existed. The MHPA originated during President Clinton's administration in context of his ultimately failed health care reform. By the time it was passed in 1996 as an amendment to the Veterans Affairs, Housing and Urban Development legislation, it just provided mental health parity for annual and lifetime dollar limits for group plans with 50 or more employees. It did not apply to annual hospitalization day or visit limits nor alcohol/substance treatments. Most health plans found loop holes to circumvent the law and restrictions were placed on psychiatric hospitalization days and outpatient visits. They also put larger emphasis on cost sharing. Despite its limitations, MHPA pushed the agenda toward mental health parity.

After significant work, in 2008, MHPAEA was passed to address the loopholes present in the MHPA. It underwent the implementation stage and became effective on April 5, 2010. It became applicable beginning or after July 1, 2010. MHPAEA requires that if an employer offers behavioral health coverage, it must be equal to those for general medical/surgical benefits. It continues to prohibit use of annual and lifetime dollar limits for mental health benefits and now includes addiction benefits. While the law is applicable to all US employers with more than 50 employees, it does not mandate employers to provide mental health and substance abuse benefits and does not extend to individual insurance market. MHPAEA does allow individual states to impose stricter parity requirements. Importantly, it also mandates that health plans that provide out-of-network coverage for medical/surgical benefits also provide equal mental health benefits. After the implementation of MHPAEA, Medicare payment liability limitations were gradually reduced and by 2014, 100 percent limitation was in effect. Thus, just like other medical treatments, mental health treatment had cost sharing of 20 percent. As of this writing, the implementation of mental health parity continues to be played out on the ground level. The law did not create any funding or mechanism to monitor, enforce or implement parity at the local level. This forces many psychiatrists and patients to continue advocating and working toward mental health parity. Therefore, it is important for mental health providers and organizations to closely monitor the continued carving-out mental health benefits to third parties so that the coverage is at parity. Also, the process of carving-in mild to moderate mental illness to primary care must be evaluated to ensure that adequate, high-quality mental health care is being delivered in medical homes.

PATIENT PROTECTION AND AFFORDABLE CARE ACT

Since the passage of the Medicare and Medicaid in 1965, the most significant health care reform legislation in America is the PPACA. Its overarching goals are to increase the quality and affordability of health insurance, lower uninsured rates and reduce overall costs of health care. After its passage, it successfully underwent challenges in court on individual mandate and federal subsidies mechanisms and is currently the law of the land. PPACA's numerous provisions started to take effect in 2010 and has a rollout period to last till 2020. Since its passage, by 2015, Centers for Disease Control and Prevention reported 11.4 million fewer Americans without insurance and polls indicated a decrease in the percentage of uninsured from 18 to 11.4 percent. While it is outside the scope of this chapter to give an overview of all provisions, several, as it pertains to care for older Americans, are worth noting.

For older Americans with mental illness, keeping out-of-pocket health care cost low, ensuring mental health parity, helping low-income

individuals, and ensuring adequate coverage of mental health treatments are important. PPACA works on several levels to help achieve these goals. First, the law prohibits insurers from denying coverage due to pre-existing conditions and sets minimum standards for health insurance policy. As most states have set certain mental health coverage as a minimum standard at parity level, PPACA works to ensure that mental illness treatment is adequately covered. The law also makes significant changes to the health insurance market by creating exchanges which allow individuals and employers to compare policies and prices of insurance. This along with sliding scale mechanisms for low-income individuals and families (between 100 and 400 percent of federal poverty level) allows for a more affordable and robust insurance market. Additionally, Medicaid eligibility has been dramatically increased. Now individuals and families with incomes up to 133 percent of federal poverty level qualify. This expansion of Medicaid is arguably the single most important aspect of the law, which has worked to decrease the number of uninsured in America. Furthermore, PPACA made several important reforms to Medicare. The reimbursement model has been restructured from fee-for-service to bundled payments. This encouraged cost savings at point of care. It also created Independent Payment Advisory Board to help reduce cost and created several measures to monitor for and control Medicare fraud and abuse. PPACA also improves preventative care coverage, which will be very important to improve the lives of older adults by potentially screening for and possibly preventing chronic diseases earlier. As mentioned earlier, it also closes out the "donut hole" in Medicare Part D. By 2020, this will be closed out and reduce prescription cost by an average of $2,000 per year. Mechanisms of PPACA can help serve older Americans with mental illness by improving outreach and education, improving care transitions, creating Medicare–Medicaid Coordination Office, and encouraging innovation.

It is generally hoped that PPACA will help offset the cost of health care in several ways. Its provisions will help reduce the number of bankruptcies due to medical care costs. Not only is medical cost the leading reason for personal bankruptcies, it also forces individuals to be locked into a job solely for health insurance coverage. PPACA is expected to reduce Medicare and Medicaid spending and thus leading to overall cost savings. Another facet of the law is to encourage new models of care delivery, one in which long-term, health care quality and health information transparency is encouraged. Innovative models like medical homes and accountable care organizations (ACOs) are just some examples being tested in the market now. These are largely designed to integrate health care providers to provide comprehensive care to the patient and to measure, report, and improve on quality of care. It also attempts to increase access for medically underserved populations. Access challenges is a complex problem and involves not only financial access but also provider access. Addressing the

critical shortage of primary care services and providers, PPACA expands community health centers and invests in public health programs. Finally, it creates new Medicaid options to promote community-based care and care giver support for those requiring long-term care.

Mental health treatment is not provided in isolation. Research and statistics argue that mental illness' impact on overall health cannot be overstated. It is integral to health and without it, true health care reform is not possible. As Americans continue to age and the impact of chronic mental illness grows, it is hoped that PPACA will have lasting positive changes on health care outcome and cost.

CURRENT AND FUTURE DIRECTION

As of this writing, the financial implications and impact on geriatric psychiatric care is actively evolving. MHPAEA and PPACA have made significant changes to the insurance market (private and public) and their impact will be felt for years to come. The growth of the older adult population will continue to be a major challenge faced by health care policy makers and providers. Without continued innovation and change, older adults will continue to consume health care services at an unsustainable rate. The recent emphasis on cost reduction in service delivery, while technology advances and the size of the frail older population grows at an unprecedented rate, has created conflict, threatening stability and coherence in the humanistic, scientific, and economic spheres. Advances in behavioral neuroscience and in particular, geriatric psychiatry, can improve quality of life, while extending life spans and improving the economic productivity of the fastest-growing segment of the population. Despite some concerns, there are ample opportunities for geriatric psychiatrists to get involved and to take on leadership roles.

Some immediate impact from MHPAEA and PPACA has to do with improving the inequality in mental health care coverage and treatment. Before this legislation, among employer-sponsored health insurance, 2 percent excluded mental health benefits and 7 percent excluded substance use benefits; rates were much higher in individual insurance markets. By the inclusion of mental health and substance abuse services as one of ten required components of Essential Health Benefit, PPACA supports the parity regulation from MHPAEA. While this does not mean that all Americans will benefit from increased parity level access to mental health care, it significantly increases the number of patients who are able to seek care. This poses challenges to geriatric psychiatrists as there simply are not enough providers to take on the increase in volume of work. Additionally, the growing emphasis on quality outcome measures begs the question of how to adequately represent quality geriatric psychiatric care when chronic mental illness follows a different time course and expected outcomes. The push to integrate mental health with general health also

challenges the traditional geriatric psychiatry practice model. Geriatric psychiatrists must be more transparent and collaborative with their primary care colleagues and expanding mental health providers (such as mid-level providers like physician assistants and nurse practitioners, therapists and peer specialists).

Strategic goals that underlie the continuing health care reform initiatives that will impact how geriatric psychiatry will be practiced are the triple aim of better care, affordable care, and better health for populations. By leveraging technological and medical advances, geriatric mental health care is well suited to achieve these goals. Some emerging models that will impact geriatric psychiatry in the future are ACOs, integrated care, patient-centered medical homes, and population health. All of these models attempt to provide cost-effective, best quality care not only to the individual, but to large populations. The Center for Medicare and Medicaid Services defines ACOs as groups of doctors, hospitals, and other health care providers who come together voluntarily to give coordinated high quality care to Medicare patients. One of the goals is to provide coordinated care to the chronically ill. It benefits within the Medicare program by sharing the cost savings. Integrated care is one response to the often disjointed and uncoordinated health care being delivered. World Health Organization defines it as a concept that brings together inputs, delivery, management and organization of services related to diagnosis, treatment, care, rehabilitation, and health promotion. The integration model then is designed to improve the overall access, quality, and efficiency of care delivery. One of its outcome goals is to achieve improved patient satisfaction as well. The patient-centered medical home is a model developed by PCPs where a health care provider leads a team centered on comprehensive and continuous care of an individual with the goal of maximizing health outcomes. It embraces the whole person orientation, care coordination by a personal physician, and emphasizes quality and safety. Finally, population health focus is an approach that tries to improve large patient populations by reducing health disparities, influencing social determinants of health, and developing strategies to have specialized but hard to find providers focused on certain diagnosis or patient subgroups. Population health management endeavors to focus interventions to specific patients and to avoid worse outcomes in the future. By focusing specialists, like geriatric psychiatrists, to complicated patients as identified within a patient registry and/or within a large organization, it seeks to maximize the impact of the specialist. The overall result would be identification and prevention of potentially high-cost patients.

Mental health care has been and will continue to be intimately related to financial aspects of health care delivery models. MHPAEA and PPACA have fundamentally changed the way in which Medicare, Medicaid, private insurance, and out-of-pocket cost impact mental health care of older

Americans. As of this writing, the full consequences of these landmark legislations are unknown, but it is hoped that they will improve the well-being of patients. For geriatric psychiatrists, participation and acquiring leadership roles in new models of care will be crucial to effectively provide treatment for the increasing number of older adults.

References

Administration on Aging, Aging Statistics 2014. Available at http://www.aoa.acl.gov/Aging_Statistics/Profile/2014/2.aspx. Accessed September 2015.

Altman D, Frist WH. Medicare and medicaid at 50 years: perspectives of beneficiaries, health care professionals and institutions, and policy makers. *JAMA*. 2015;314(4):384–395.

Barry CL, Huskamp HA, Goldman HH. A political history of federal mental health and addiction insurance parity. *Milbank Q*. 2010;88(3):404–433.

Berwick DM. Making good on ACOs' promise—the final rule for the medicare shared savings program. *N Engl J Med*. 2011;365:1753–1756.

Bitton A, Martin C, Landon B. A nationwide survey of patient centered medical home demonstration projects. *J Gen Intern Med*. 2010;25(6):584–592.

Centers for Medicare and Medicaid Services: 2015 Medicare Costs. Available at https://www.medicare.gov/Pubs/pdf/11579.pdf. Accessed September 15, 2015.

Centers for Medicare and Medicaid Services: CMS.gov. https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nhe-fact-sheet.html. Accessed January 15, 2017.

Chen C, Ackerly D. Beyond ACOs and bundled payments: medicare's shift toward accountability in fee-for-service. *JAMA*. 2014;311(7):673–674.

Cubanski J, Swoope C, Damico A, Neuman T. *How Much is Enough? Out-of-Pocket Spending Among Medicare Beneficiaries: A Chartbook*. The Henry J. Kaiser Family Foundation. 2014.

Davis K, Stremikis K, Schoen C, Squires D. *Mirror, Mirror on the Wall, 2014*. Update: How the U.S. Health Care System Compares Internationally, The Commonwealth Fund, 2014.

DeNavas-Walt C, Proctor B, US Census Bureau. *Current Population Reports, P60–249, Income and Poverty in the United States: 2013*. Washington, DC: US Government Printing Office; 2014.

Institute for Healthcare Improvement. Triple Aim Initiative: Better Care for Individuals, Better Health for Populations, and Lower Per Capital Costs. Available at http://www.ihi.org/engage/initiatives/tripleaim/Pages/default.aspx. Accessed August 2015.

Katon W. Health reform, research pave way for collaborative care for mental illness. Interview by Bridge M. Kuehn. *JAMA*. 2013;309(23):2425–2426.

Keckley PH, Hoffmann M. Accountable care organizations: a new model for sustainable innovation. *Deloitte Center for Health Solutions*. 2011:1–19.

Knickman JR, Hunt KA, Snell EK, Alecxih LM, Kennell DL. Wealth patterns among elderly Americans: implications for health care affordability. *Health Aff (Millwood)*. 2003;22(3):168–174.

Kyomen HH, Gottlieb GL. Financial issues in the delivery of geriatric psychiatric care. In: Sadock BJ, Sadock VA, Ruiz P, eds. *Comprehensive textbook of psychiatry/IX*. Philadelphia, PA:

Lippincott Williams & Wilkins; 2009:4185–4193.

McClellan M, McKethan AN, Lewis JL, Roski J, Fisher ES. A national strategy to put accountable care into practice. *Health Aff (Millwood)*. 2010;29(5):982–990.

McDonough JE, Adashi EY. Realizing the promise of the Affordable Care Act—January 1, 2014. *JAMA*. 2014;311(6):569–570.

Medicaid Behavioral Health Services. Available at http://www.medicaid.gov/medicaid-chip-program-information/by-topics/benefits/mental-health-services.html. Accessed August 2015.

Merlls M. Health Policy Brief: Accountable Care Organizations. *Health Affairs*. 2010:1–6.

Patient-Centered Primary Care Collaborative. Defining the Medical Home, A Patient-centered Philosophy that Drives Primary Care Excellence. Available at https://www.pcpcc.org/about/medical-home. Accessed July 2015.

Potetz L, Cubanski J, Neuman T. *Medicare Spending and Financing, A Primer*. The Henry J. Kaiser Family Foundation; 2011.

Rosenbaum S. The Patient Protection and Affordable Care Act: Implications for Public Health Policy and Practice. *Public Health Rep*. 2011;126(1):130–135.

Smith J, Medalia C. *U.S. Census Bureau, Current Population Reports, P60-250, Health Insurance Coverage in the United States:* 2013, Washington, DC: U.S. Government Printing Office; 2014.

United Nations, Department of Economic and Social Affairs, Population Division (2013). World Population Ageing 2013. ST/ESA/SER.A/348.

United States Department of Labor. *The Mental Health Parity and Addiction Equity Act of 2008 Fact Sheet (MHPAEA),* January 29, 2010.

US Census Bureau: Population Projections 2014. Available at http://www.census.gov/population/projections/data/national/2014.html. Accessed June 2015.

US Government Publishing Office. 75 FR 37187 – Patient Protection and Affordable Care Act: Preexisting Condition Exclusions, Lifetime and Annual Limits, Rescissions, and Patient Protections. Available at http://www.gpo.gov/fdsys/granule/FR-2010–06–28/2010–15278/content-detail.html. Accessed September 2015.

US Government Publishing Office. Public Law 111 – 148 – Patient Protection and Affordable Care Act. Available at http://www.gpo.gov/fdsys/granule/PLAW-111publ148/PLAW-111publ148/content-detail.html. Accessed September 2015.

What's Medicare Supplement Insurance (Medigap)? Available at https://www.medicare.gov/supplement-other-insurance/medigap/whats-medigap.html. Accessed September 15, 2015.

▲ 57.6 Special Areas of Interest

57.6a Psychiatric Aspects of Long-term Care

ABHILASH K. DESAI, M.D., AND GEORGE T. GROSSBERG, M.D.

The prevention and treatment of psychiatric disorders is a critical

component of comprehensive, state-of-the-science care for all individuals who live in long-term care facilities. Psychiatric disorders in long-term care residents are prevalent and complex; have heterogeneous etiologies, clinical presentations, and treatment outcomes; and are the subject of extensive research. Early and comprehensive management of psychiatric disorders has the potential not only to promote the dignity, well-being, and safety of residents in long-term care facilities but also to improve the quality of life of their caregivers and reduce the costs to society by preventing visits to emergency departments and hospitalizations.

At least 20 percent of older adults will spend their last years in a long-term care facility, and another 40 percent will have a short stay in a skilled nursing facility for rehabilitation (subacute care) after hospitalization for acute care. Long-term care facilities primarily include nursing homes and assisted living homes but can also include the many certified family homes and group homes that provide care to older adults who are unable to live independently due to disabling physical and/or psychiatric problems. More than 3 million individuals in the United States live in long-term care facilities, with 1.5 million of them in approximately 16,000 nursing homes. This number of 3 million is expected to double by 2030. The proportion of residents older than 85 has increased dramatically in the last decade and currently comprises the largest age group (37.1 percent). Residents in nursing homes and assisted living homes have grown progressively more debilitated, requiring more assistance with basic activities of daily living, such as bathing, grooming, and toileting. A growing number of residents are younger, having experienced stroke, traumatic brain injury, or other disabling neuropsychiatric disorders. Adults who have developmental disability (especially intellectual disability, Down syndrome, cerebral palsy, autism spectrum disorder) are also living longer and often live their last years in a long-term care facility.

More than 90 percent of the long-term care population at some point will experience clinically significant psychiatric symptoms. (Table 57.6a–1shows the prevalence of common psychiatric disorders in the long-term care population.) Thus, psychiatric care is a core component of the day-to-day care of most of the residents in long-term care facilities. The long-term care population is not only primarily older adults who have substantial disability but also those who have considerable cognitive impairment and/or medical burden, making the availability of mental health professionals (psychiatrists, psychiatric nurse practitioners and physician assistants, psychologists, and licensed counselors) who have additional training in geriatric medicine and psychiatry, key to high-quality care for this vulnerable population.

THE ROLE OF THE PSYCHIATRIST IN LONG-TERM CARE

A psychiatrist typically has the role of a clinical consultant in long-term

care, although some may serve as medical director. Psychiatrists can also play a vital role in administrative aspects of care, especially in moving the culture of the facility toward a person-centered approach and improving collaboration with primary care clinicians and pharmacists to reduce inappropriate use of medication. Psychiatrists can play a key role in educating and training the staff to understand behavioral symptoms of major neurocognitive disorders, in handling psychiatric emergencies, and in helping the facility develop policies and procedures regarding when to transfer a resident to a local emergency department for treatment in an acute inpatient psychiatric unit. The authors recommend adhering to the American Medical Directors Association's clinical practice guidelines on delirium, acute problematic behavior, dementia, depression, pain management, sleep disorders, and transitions to other care. Psychiatrists trained in geriatrics (e.g., geriatric psychiatrists, psychiatrists who have completed training for nursing home medical directors provided by the American Medical Directors Association) can also play the role of the facility medical director or codirector. Many residents and their family members are reluctant to accept psychiatric help due to perceived stigma. Psychiatrists can play a crucial role in alleviating such concerns by being available for informal meetings with the resident and their family to clarify the unique expertise they provide and the great potential for psychiatric help in reducing suffering and improving quality of life of the resident. Psychiatrist's presence in monthly support groups at the facility for the residents and/or their family members may also help reduce stigma for psychiatric care.

Another role which psychiatrists can fill in long-term care is via telepsychiatry. Telepsychiatry involves providing psychiatric assessment and management remotely (typically using a two-way computer system with audio–video conferencing). Telepsychiatry can help overcome a key barrier to improving psychiatric care in long-term care facilities: lack of availability of mental health professionals with geriatric expertise, especially in rural areas. Although the benefits and risks of telepsychiatry have not been subject to rigorous research, the authors support the use of telepsychiatry for facilities that do not have a mental health consultant with geriatric expertise primarily for the goals of reducing the use of psychopharmacologic agents (especially antipsychotics, benzodiazepines, and multiple antidepressants), differentiating delirium from other psychiatric disorders, educating and supporting families, evaluating residents who have suicidal ideation, helping with decisions about palliative and end-of-life care, and educating and training the staff to implement individualized psychosocial environmental interventions.

APPROACH TO EVALUATING AND MANAGING PSYCHIATRIC SYMPTOMS AND DISORDERS

A systematic approach to evaluating and managing psychiatric disorders using evidence-based practices and interventions has the potential to reduce suffering dramatically and to prevent unnecessary visits to the emergency department and hospitalization. The authors recommend four steps to evaluating and managing psychiatric disorders in the long-term care population.

Assessment of Safety

The first step is to assess safety. The psychiatrist may need to address behaviors that pose an imminent danger to the resident or others (behavioral emergencies), such as violent or suicidal behaviors, before attempting to assess underlying psychiatric problems. Suicidal behaviors can be direct (e.g., suicidal ideas, such as voicing thoughts to cut oneself with a knife or scissors, or suicide attempt, such as attempting to strangle oneself with a cord) or indirect (e.g., refusing dialysis, refusing food or life-preserving medications such as warfarin). Completed suicides, although rare in long-term care, may be more common than in the community. If a behavioral emergency cannot be resolved, the resident may need to go to the nearest emergency department for urgent evaluation and stabilization.

Team-Based Evaluation to Identify Causes of Psychiatric Symptoms and Disorders

The second step is for a multidisciplinary team (resident, family members, primary care clinician, psychiatrist, nurse, nursing assistant, and social worker) to identify all factors that may contribute to psychiatric symptoms. Psychiatric symptoms are typically due to several specific reversible or modifiable causes. (Table 57.6a–2lists common reversible causes of psychiatric symptoms in the long-term care population.) Obtaining the resident's perspective is essential, even if the resident has significant cognitive impairment. Diagnoses listed in the long-term care resident record need to be validated as they are often incorrect. In complicated cases (e.g., presence of chronic mental illness, refractory and severe behavioral and psychological symptoms of neurocognitive disorder, concerns of abuse and neglect, and complex polypharmacy), the psychiatrist may need to consult a professional who has geriatric expertise (e.g., geriatric psychiatrist, geriatrician).

Action Plan

In the third step, the team collaboratively devises and institutes an action plan (treatment plan) that is individualized and directed at the specific target psychiatric symptoms, diagnoses and differential diagnoses identified in the second step. An action plan generally consists of a multimodal approach combining a variety of psychosocial environmental

interventions targeted at caregivers and residents to address psychiatric symptoms that are most distressing or problematic. An action plan should address all the potentially modifiable causes of psychiatric symptoms identified during evaluation. Professionals from allied fields (e.g., pharmacist, occupational therapist, physical therapist, speech therapist, recreational therapist, music therapist, activity therapist, and art therapist) and professionals who have knowledge of complementary and alternative medicine (e.g., aromatherapy, massage therapy, and horticultural therapy) can play valuable roles. Psychopharmacologic interventions may be the first line in behavioral emergencies related to violent behaviors, relapse of pre-existing psychiatric disorder (e.g., bipolar disorder), new onset of major depression, persistent and distressing psychotic symptoms, and persistent aggressive behavior. (Table 57.6a–3lists key principles of psychopharmacology with the long-term care population.)

Table 57.6a–1.
Prevalence of Common Psychiatric Disorders in Long-Term Care

Psychiatric Syndrome (Point Prevalence)	Key Example(s)	Comments
Major neurocognitive disorders (MNCD) (60–90%)	Alzheimer disease (AD) Vascular cognitive impairment (VCI) Parkinson disease dementia Dementia with Lewy bodies Frontotemporal dementia Traumatic brain injury and chronic traumatic encephalopathy Hypoxic/anoxic encephalopathy Alcohol-related dementia	Many residents have both AD and VCI. Most residents are in advanced stages of major neurocognitive disorder. Behavioral and psychological symptoms of major neurocognitive disorder are the most common reasons for psychiatric consultation. Many residents, especially in subacute care, have minor neurocognitive disorder and, when there is superimposed depression or delirium, may be misdiagnosed as having major neurocognitive disorder.
Depressive disorders (20–40%)	Depression due to MNCD, stroke, cerebrovascular disease, or another general medical condition (most common) Major depression (single episode and recurrent) Persistent depressive disorder (dysthymic	Adequate control of pain is needed to improve depression, and optimal treatment of depression is needed to improve pain for residents who have pain and depression. Drug–drug interactions between antidepressants (e.g., SSRIs[a]) and drugs used to treat medical comorbidity (e.g., tramadol, diuretics, warfarin) are common and need to be monitored. A schedule of pleasant events and physical exercise have been found to be effective in the treatment of depression.

	disorder) Adjustment disorder with depressed mood or with mixed depression and anxiety Depression due to medications	
Delirium (5–15%)	Hypoactive delirium (most common) Hyperactive delirium Mixed delirium	Delirium is especially prevalent among residents newly transferred from an acute care hospital (subacute settings in skilled nursing facilities). Subsyndromal delirium is even more common and commonly manifests with new onset of agitation and/or new onset of falls.
Psychotic disorders (5–10%)	Psychotic symptoms due to MNCD (most common) Psychotic symptoms due to Parkinson disease and dopaminergic agents used to treat Parkinson disease Schizophrenia and schizoaffective disorder	Psychotic symptoms are also common in the context of delirium. Psychotic symptoms are a common cause of persistent physical aggression. As adults who have schizophrenia and schizoaffective disorder live longer, their prevalence in long-term care is expected to rise further.
Hypomania and mania (2–5%)	Bipolar disorder (types I and II, most common) Medication-induced hypomania or mania Hypomania or mania due to stroke or another general medical condition	As adults who have bipolar disorder live longer, their prevalence in long-term care is expected to rise further. Corticosteroids are the most common cause of medication-induced hypomania or mania.
Anxiety disorders (lack of rigorous data on prevalence; clinical experience indicates high prevalence)	Anxiety symptoms and bathing phobia in the context of major neurocognitive disorder (most common) Generalized anxiety disorder Panic disorder Fear of falling	Anxiety symptoms are often comorbid with depressive symptoms. Premorbid subsyndromal generalized anxiety disorder is often exacerbated by development of a major neurocognitive disorder.
Trauma-related disorders (lack of rigorous data on prevalence; clinical experience indicates growing prevalence)	Posttraumatic stress disorder (PTSD)	War veterans, victims of abuse (physical, sexual, emotional), and victims of a motor vehicular accident should be identified at the time of admission to a long-term care facility so that investigation for presence of PTSD can be done soon after admission.

Impulse-control disorders (5–15%)	Sexually inappropriate behaviors in the context of major neurocognitive disorder (especially VCI)	Dopamine agonists (pramipexole, ropinirole [Requip], rotigotine [Neupro]) may cause impulse-control problems, including sexually inappropriate behaviors and excessive gambling or spending sprees.
Substance-use disorders (lack of rigorous data on prevalence; clinical experience indicates growing and significant prevalence)	Alcohol (most common) Tobacco Prescription opioids Benzodiazepines	Hazardous use of alcohol, tobacco, prescription opioids, and benzodiazepines that do not meet the criteria for substance-use disorder are prevalent in long-term care and should be treated vigorously. As more baby boomers enter long-term care, the prevalence of substance-use disorders (especially alcohol and marijuana) in the long-term care population is expected to grow.
Other psychiatric disorders with growing prevalence	Disorders of the sleep–wake cycle Eating disorders Intellectual disability Autism spectrum disorder Personality disorder	These disorders are often comorbid with other psychiatric disorders. They need to be identified and the treatment plan modified accordingly.

[a]SSRIs, selective serotonin reuptake inhibitors.

Table 57.6a–2. Common Reversible Causes of Psychiatric Symptoms in Long-Term Care

Cause	Key Example(s)	Potential Psychiatric Symptoms
Medications	Benzodiazepines Anticholinergics Corticosteroids Dopamine agonists	Cognitive impairment Agitation Hypomania, mania Psychotic symptoms
Physical health conditions	Pain, neuropathy, constipation, urinary tract infection, dehydration, delirium, anemia, vitamin B_{12} deficiency, vitamin D deficiency, low blood pressure due to overtreatment of hypertension	Agitation Depression Cognitive impairment
Sensory deficits	Hearing deficit Vision deficit	Depression Agitation Paranoia Visual hallucinations Cognitive impairment
Disorders of the sleep–wake cycle	Insomnia Obstructive sleep apnea Excessive daytime sleepiness REM sleep behavior disorder Restless leg syndrome	Increased or decreased sleep Nighttime agitation Cognitive impairment

Psychiatric conditions	Major neurocognitive disorders Major depression (single episode or recurrent) Generalized anxiety disorder Posttraumatic stress disorder (PTSD) Bipolar disorder Schizophrenia	Agitation/aggression Social withdrawal Insomnia Cognitive impairment Nightmares Poor food and fluid intake Delusions Hallucinations Hypomania/mania
Long-term care staff approach	Impatience Lack of empathy Overexpectation	Agitation/aggression Depression
Long-term care staff's understanding of psychiatric symptoms/disorders	Staff's opinion that the behaviors are intentional ("he can stay in bed all day, as he is 90 years old and is tired")	Agitation Depression Insomnia
Long-term care staff distress	Caregiver anxiety and depression Caregiver burnout Excessive caregiver burden (having to care for too many residents)	Agitation Depression
Negative environmental factors	Nighttime noise Low lighting Boredom Overstimulation Improper or inadequate incontinence care routines Therapeutic diet Lack of places to wander Physical restraint	Agitation Depression Anxiety Poor food and fluid intake Insomnia
Unmet needs of residents	Toileting, hunger Physical activity Spend time with nature To be productive, to be needed Spiritual needs To express and vent Freedom from restraint	Agitation Depression

Response to Treatment

The fourth step involves appropriately monitoring adverse effects of interventions as well as response to treatment so that the psychiatrist can evaluate whether to modify or continue the interventions. If the problem is not improving, the psychiatrist should discontinue the intervention. If the intervention included an antipsychotic, the psychiatrist should make systematic attempts to taper the dosage and discontinue the medication after 6 to 12 months.

DELIRIUM AND PSYCHIATRIC ASPECTS OF COMORBID MEDICAL CONDITIONS

Age above 85, presence of a neurocognitive disorder, and polypharmacy are the three most important risk factors for delirium in the long-term care

population. More than 70 percent of residents have major neurocognitive disorder, and the majority of the remaining 30 percent have minor neurocognitive disorder. Residents who have advanced stages of major neurocognitive disorder have low cognitive reserve and thus are prone to develop delirium even with mild conditions such as urinary tract infection or a new medication that has anticholinergic activity. The psychiatrist should consider acute change in mental status, that is, the sixth vital sign (pain being the fifth vital sign and heart rate, respiratory rate, temperature, and blood pressure being the first four vital signs). New onset of falls and agitation are the two most common early signs of developing delirium in long-term care populations. New onset of agitation is one of the most common reasons for psychiatric consultation, and the psychiatrist may be the first health care professional to diagnose delirium.

Table 57.6a–3.
Key Principles of Psychopharmacology in Long-term Care

1. Prescription of even a single psychopharmacologic agent is a complex decision. It should be undertaken judiciously and, in most situations, after psychosocial environmental interventions have been tried and failed.
2. Start low, go slow, and keep going. Initial doses of psychotropic drugs need to be low or, for frail residents, very low. The dose increase needs to be slower than for younger adults. Final therapeutic doses for some residents may need to be as high as those for younger adults.
3. The benefit–risk ratio for all pharmacologic agents (including psychopharmacologic drugs) is lower than that for younger adults, necessitating close monitoring for adverse effects and prompt discontinuation of the drug if adverse effects develop or if there is no therapeutic response.
4. The psychiatrist should evaluate for periodic dose reduction and discontinuation of psychotropic drugs, and documentation should include the justification for continued use of psychopharmacologic agent(s).
5. The psychiatrist should promptly reassess the benefit–risk ratio of all psychotropic drugs whenever there is significant decline in the resident's overall health and functioning or decline in renal and/or hepatic function.
6. Psychopharmacologic therapies should comply with regulatory guidelines for the use of medications in nursing homes.
7. The psychiatrist should provide residents and their surrogate decision makers optimal information regarding the risks and benefits of psychopharmacologic therapies and alternatives before initiating therapy and periodically thereafter.

There are three types of delirium: hypoactive (most common), hyperactive, and mixed. Psychotic symptoms are common during delirium and may cause the resident to become severely agitated or physically aggressive, further complicating management. The Confusion Assessment Method (CAM) is a useful tool for improving accurate diagnosis of delirium in the long-term care population. (Table 57.6a–4lists common causes of delirium in the long-term care population.) Mania can also mimic delirium. Management of delirium involves ensuring safety and identifying and treating its various causes. Judicious short-term use of atypical antipsychotics may be appropriate and necessary for the management of

severe physical aggression and psychotic symptoms in the context of delirium.

Ms. A is a 65-year-old woman who was recently transferred from a local hospital to a skilled nursing facility. She was referred for psychiatric evaluation because she was not eating, appeared depressed, and was not sleeping well. The nursing staff suspected depression. During assessment, the psychiatrist discovered that Ms. A had undergone a colonic resection for ulcerative colitis and had excessive liquid drainage from an ileostomy. During examination, Ms. A was inattentive and somewhat drowsy and had psychomotor retardation. The psychiatrist diagnosed hypoactive delirium, possibly due to disorders of fluids and electrolytes. A subsequent basic metabolic panel showed hyponatremia, elevated blood urea nitrogen and serum creatinine (higher than baseline), confirming the diagnosis. Treatment with intravenous fluids at the facility quickly brought resolution of the presenting symptoms.

Table 57.6a–4.
Common Causes of Delirium in Long-term Care

Cause	Key Example(s)	Comments
D: Drugs	Anticholinergic drugs Benzodiazepines Opioids Steroids	Drugs listed in Beers list of drugs that are inappropriate for older adults often cause cognitive impairment and delirium. Many drugs have low anticholinergic activity and, if used with other drugs that have anticholinergic activity, may increase anticholinergic load sufficient to cause delirium. Benzodiazepine withdrawal can cause delirium.
E: Embolic–thrombotic	Pulmonary emboli Deep vein thrombosis (DVT) Cerebrovascular accident (CVA)	Residents who spend a large amount of time in bed (e.g., those who have depression, those who are in pain) are at high risk of DVT and its complications. Neuroimaging may be necessary to diagnose CVA accurately.
L: Last phase of life	Residents in end-of-life state	Delirium is common in terminal stages of major neurocognitive disorder.
I: Infections	Urinary tract infection Bronchitis Pneumonia Cellulitis	Asymptomatic bacteriuria and asymptomatic pyuria are prevalent, especially among women, and often misdiagnosed as urinary tract infection. Delirium may be a more common manifestation of pneumonia than fever and cough.
R: Respiratory conditions	Acute exacerbation of chronic obstructive pulmonary disease (COPD)	Acute bronchitis is one of the common causes of exacerbation of COPD.
R: Renal conditions	Acute renal failure (ARF)	Dehydration is a common cause of ARF.
I: Intestinal conditions	Acute intestinal obstruction, including fecal impaction	Inadequate treatment of constipation often leads to intestinal obstruction.
I: Injury	Hip fracture	
U: Urologic	Urinary retention	Drugs that have anticholinergic activity are the most common cause of acute urinary retention.
M: Myocardial	Acute myocardial infarction (AMI)	Delirium may be a more common manifestation of AMI than chest discomfort.

dysfunction	Acute congestive heart failure	Laboratory tests may be the only way to diagnose dehydration accurately.
M: Metabolic	Dehydration	
	Hypothyroidism or hyperthyroidism	
	Hyponatremia	

Clinical pearl: Hypoactive delirium often presents as depression.

Malnutrition

Malnutrition is prevalent in the long-term care population. It commonly manifests as unintentional weight loss or rarely as excessive weight gain and is a common reason for psychiatric consultation. Weight loss is usually accompanied by loss of appetite (including general loss of appetite as well as loss of appetite for food offered at the facility) or refusal of food. Depression is one of the most common causes of a general loss of appetite. Medications such as cholinesterase inhibitors also can cause anorexia and weight loss. Not having someone available to assist with meals, especially in residents with neurocognitive disorders is also a common reason for weight loss in long-term care facility population. Liberalizing diets, optimizing dental health, cutting food into smaller and manageable sizes (finger foods), giving smaller portions at each meal, and increasing the availability of snacks throughout the day can help address the problem of food refusal. The psychiatrist should evaluate all residents who refuse food for depression as well as for psychotic symptoms (especially delusions of food being poisoned). For residents who have depression accompanied by loss of appetite and weight loss, the psychiatrist may consider prescribing mirtazapine (Remeron) as the antidepressant of first choice. Nutritional supplements should be given between meals rather than with meals.

The prevalence of obesity and excessive weight gain is increasing in the long-term care population. This is typically accompanied by excessive consumption of calorie-dense food. Along with sedentary lifestyle, psychopharmacologic medications (e.g., valproate [Depakote], olanzapine [Zyprexa], quetiapine [Seroquel], mirtazapine) and other pharmacologic medications (e.g., gabapentin [Neurontin]) are common causes of excessive weight gain. It may be necessary to discontinue psychopharmacologic agents that are thought to cause or contribute to weight gain and as necessary replace them with alternative psychopharmacologic agents that have lower propensity for weight gain.

Pain

Psychiatric consultation can provide valuable help in improving outcomes and quality of life of residents who are experiencing chronic pain. Depression is common, underrecognized, and undertreated in residents

who have chronic pain. Pain is one of the most common causes of agitation in residents who have major neurocognitive disorder, especially verbal agitation (e.g., yelling, screaming, verbally abusive behavior) and resisting personal care. The psychiatrist should consider pain, the fifth vital sign. Antidepressants, especially serotonin–norepinephrine reuptake inhibitors (SNRIs, such as duloxetine [Cymbalta]), can help reduce the perception of pain and improve function as well as quality of life, even in the absence of significant depression, for residents who have chronic pain. Serious adverse drug–drug interactions between psychotropic drugs and analgesics (e.g., selective serotonin reuptake inhibitors [SSRIs] and tramadol [Ultram] [increased risk of seizure and serotonin syndrome], benzodiazepines and opioids [increased risk of delirium, falls, and injury]) can be prevalent and underrecognized in the long-term care population. The psychiatrist should consider recommending a trial of scheduled acetaminophen (Tylenol) to treat agitation (especially verbal agitation) before prescribing psychopharmacologic agents. Psychiatric care should be a key component of an integrated approach to pain management for all residents who have chronic pain.

MAJOR NEUROCOGNITIVE DISORDERS (DEMENTIAS)

Evaluation and treatment of major neurocognitive disorders is one of the key aspects of psychiatric care of long-term care residents. Residents who have major neurocognitive disorder and have not had a thorough evaluation to clarify its cause(s) may need workup to rule out potentially treatable or reversible factors of cognitive decline, including thyroid dysfunction, vitamin deficiency (e.g., B_{12} deficiency), subdural hematoma, normal pressure hydrocephalus, central nervous system infection, and brain tumor. The authors recommend using the Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), or Saint Louis University Mental Status (SLUMS) to assess residents' cognitive function. The Brief Interview for Mental Status is not as thorough as the previous three scales, although it is part of the Minimum Data Set used to assess all residents in nursing homes in the United States.

Behavioral and psychological symptoms of major neurocognitive disorders (also called behavioral and psychological symptoms of dementia [BPSD]) are a heterogeneous group of symptoms that tend to cluster into five distinct groups—depression, apathy, psychotic symptoms (delusions and hallucinations), sleep disturbance, and agitation—and are distinct from the cognitive deficits that are the hallmarks of all major neurocognitive disorders. Agitation, defined as "inappropriate verbal, vocal, or motor activity that is not explained by needs or confusion per se," can be characterized as either aggressive or nonaggressive. Aggressive behaviors can occur with or without agitation and include fighting, throwing, grabbing, destroying furniture, cursing, and screaming, whereas

nonaggressive agitated behaviors include restlessness, pacing, wandering, hoarding, repetitive questioning, inappropriate disrobing, and verbal outbursts. Agitation is best understood as an expression of a resident's efforts to communicate unmet needs. The majority of individuals who have major neurocognitive disorder (including those who have advanced major neurocognitive disorder) can speak for themselves to a much higher degree than has been recognized to date. The psychiatrist should make all efforts to understand the resident's perspective and have the treatment plan be guided by that perspective. The psychiatrist may use the Agitation Subscale of the Neuropsychiatric Inventory nursing home version (NPI-NH), Cohen Mansfield Agitation Inventory (CMAI), or Pittsburgh Agitation Scale (PAS) to evaluate the severity of agitation and to document response to treatment.

Management of major neurocognitive disorders should begin with discussion with the resident and/or the surrogate decision maker regarding potential benefits of use of cholinesterase inhibitors and memantine (Namenda). Some more capable residents (e.g., residents who have minor neurocognitive disorder due to probable Alzheimer type and highly motivated residents who have mild major neurocognitive disorder) may benefit from and often prefer interventions to slow cognitive decline. For these residents, creating and implementing an individualized brain fitness program that includes regular exercise, Mediterranean diet, cognitive and social strategies, cholinesterase inhibitors, and use of memantine in various combinations is appropriate and recommended.

Table 57.6a–5.
Psychosocial/Environmental Interventions for Neurocognitive Disorders

Psychosocial Interventions Directed Toward Caregivers		
Type of Psychosocial Environmental Intervention	**Examples**	**Practical Suggestions**
Staff education	Education about dementia Education about Parkinson disease Education about common triggers for agitation Education about adverse psychiatric effects of commonly prescribed medications	Case-based education preferred. Recording the education and making it available on website for caregivers to review on their own time. Providing work credit for caregivers who complete the education.
Staff training	Training in approach to the management of behavioral and psychological problems in long-term care residents. Training in effective communication and appropriate documentation.	Hands-on training, especially of new staff and staff with minimal experience. Video recordings of training for staff to review on their own time.
Staff support and incentives	Informal support to staff members Formal support groups for staff	During visits, psychiatrist can provide direct and situation-specific verbal support to

	members Opportunity to take on leadership role in psychiatric aspects of long-term care	caregivers. A break room for staff members where they can "chill" (massage chairs, relaxing music, yoga videos and mats).

Psychosocial Environmental Interventions Directed Toward Residents

Type of Psychosocial Environmental Intervention	Examples	Comments/Resources
Environmental	Design of physical environment that accommodates cognitive impairment and promotes emotional well-being Robust activity programming Structured routines Therapeutic gardens	www.healthdesign.org Consultation, staff education, and periodic supervision of the programs by recreational therapist Activity programs run by activity therapist Cognitive stimulation therapy (www.cstdementia.com) www.timeslips.org www.thegreenhouseproject.org
Music-based strategies	Individualized music Group music	Consultation, staff education, and periodic supervision of the programs by music therapists Availability of piano for the residents to play, music CDs, iPod, mp3 players, radio
Physical activity and exercise	Walking programs Chair yoga Tai Chi Dance exercise programs	Guidance, staff education, and periodic supervision of the programs by physical therapist Staff encouraged to make physical activity a normal part of resident's routine
Aromatherapy combined with massage therapy	Hand or foot massage with a fragrant lotion	Consultation, staff education, and periodic supervision of aromatherapy programs by certified aromatherapist Availability of common therapeutic lotions (e.g., lavender and citrus lotions) for staff to use
Animal-assisted therapy	Dogs and cats visiting	Eden Alternative programs (www.edenalt.org)
Emotional support	Support groups	Stroke support group Early-stage dementia support group
Counseling	Problem-solving therapy Reminiscence therapy Interpersonal therapy Cognitive-behavioral therapy	Facility administration encouraging licensed psychotherapist/counselor to provide psychological therapies to cognitively capable residents Internet-based self-psychotherapy for cognitively capable residents
Spiritual	Maintaining religious rituals	"Postcards to God" program

Psychosocial environmental interventions (also called nonpharmacologic interventions) are recommended as the first line of treatment of all behavioral and psychological symptoms related to major neurocognitive disorders because they are effective and the risks are minor. (Table 57.6a–5lists psychosocial environmental interventions for

neurocognitive disorder.)

The psychiatrist should institute psychopharmacologic interventions to treat agitation related to neurocognitive disorder only after thoroughly evaluating and treating all potentially treatable medical conditions and other modifiable factors and psychosocial environmental interventions have not been adequate in addressing them. Discussions should make it clear that any psychopharmacologic intervention for agitation in the context of neurocognitive disorder is off label (i.e., is not approved by the U.S. Food and Drug Administration for its treatment).

Antipsychotics are the most commonly prescribed psychopharmacologic agents for the treatment of agitation and aggression among residents who have major neurocognitive disorder, and the prevalence of use in the long-term care population is high (20 to 25 percent). There is a strong initiative from the Center for Medicare and Medicaid Services (CMS) in the United States to reduce the use of antipsychotics in the long-term care population. In addition, the United States Congress has passed legislation mandating reduction of antipsychotic use in nursing homes. This is in response to strong evidence that antipsychotics are often used inappropriately in the long-term care population (with the majority of its use being by residents who have major neurocognitive disorder) and the use of antipsychotics by individuals who have major neurocognitive disorder carries high risks of morbidity (e.g., cerebrovascular event, pneumonia, deep vein thrombosis, hip/femur fracture) and increased mortality. (Table 57.6a–6shows a stepped care approach to appropriate use of antipsychotic medications for major neurocognitive disorder.) The psychiatrist may consider prescribing short-term use of low-dose risperidone (Risperdal), olanzapine (Zyprexa), quetiapine (Seroquel), and aripiprazole (Abilify) to treat severe agitation, aggression, and psychosis with dementia in certain clinical situations (e.g., clear risk of harm to the resident or others and psychosocial environmental interventions have not been effective). Atypical antipsychotics are preferred over conventional antipsychotics (e.g., haloperidol [Haldol]), which are associated with higher risks. It is important that the psychiatrists explain to the surrogate decision maker(s) that severe psychotic symptoms and/or persistent aggression are associated with considerable emotional distress, increased mortality, falls, and hospitalization; hence, in certain situations, the use of antipsychotics may be justified. To minimize risks, the authors recommend prescribing the lowest possible dose of antipsychotics for the shortest possible time and regular monitoring of metabolic parameters when appropriate and for extrapyramidal symptoms.

Table 57.6a–6.
Approach to Appropriate Use of Antipsychotic Medications to Manage

Severe Behavioral and Psychological Symptoms of Major Neurocognitive Disorder

Step 1: Presence of at least one of the following:

- ► Persistent severe agitation and or physically aggressive behavior

or

- ► Psychotic symptom causing severe emotional distress

Step 2: Assess safety

- ► The resident poses an imminent danger to self and/or others (address safety first using following options individually or in combination: one-to-one staff monitoring; use of oral or intramuscular antipsychotic medication; transfer to Emergency Department. When safety is ensured, proceed to step 3)
- ► The resident does not pose an imminent danger to self and/or others (proceed to step 3)

Step 3: Treatment team completes a thorough evaluation to look for treatable or reversible causes of target problems

- ► Primary care clinician or psychiatric consultant completes history and examination (including pertinent neurologic and detailed mental status examination)
- ► Complete team discussion to identify treatable or reversible causes
- ► Order laboratory and other tests as necessary and results reviewed or testing not required

Step 4: Treat agitation and/or emotional distress with an individualized treatment plan and without psychopharmacologic intervention

- ► Correct identified causes of target problems
- ► Team discussion identifies and institutes individualized psychosocial environmental interventions

Step 5: Assess response to treatment

- ► Adequate improvement and/or resolution of target problems
- ► Lack of adequate response to treatment (proceed to step 6)

Step 6: Attempt trial of pharmacologic interventions other than antipsychotics

- ► Attempt trial of acetaminophen if there are no contraindications
- ► Attempt trial of citalopram (up to 20 mg/day) or escitalopram if there are no contraindications
- ► Attempt trial of low-dose trazodone if there are no contraindications
- ► Attempt trial of combination of above medications if there are no contraindications

Step 7: Assess response to treatment

- ► Adequate improvement and/or resolution of target problem(s)
- ► Lack of adequate response to treatment (proceed to step 8)

Step 8: Make decision to prescribe antipsychotic medications

- ► Obtain consent from surrogate decision maker for use of antipsychotics after adequately explaining risks and potential benefits of their use

Step 9: Attempt trial of antipsychotic medications

- ► Attempt trial with one of three atypical antipsychotics: risperidone, olanzapine, aripiprazole

Step 10: Assess response to treatment

- ► Adequate improvement and/or resolution of target problems without significant adverse effects
- ► Adequate improvement and/or resolution of target problems but associated with significant adverse effects (proceed to step 11)
- ► Lack of adequate response to treatment (proceed to step 11)

Step 11: Attempt trial of another of the three antipsychotic medications mentioned above

- ► Adequate improvement and/or resolution of target problems (proceed to step 13)
- ► Lack of adequate response to treatment (proceed to step 12)

Step 12: Team discussion (including with surrogate decision maker) regarding further care

► Decide to transfer resident to local Emergency Department for acute inpatient psychiatric treatment or to transfer resident directly to an acute inpatient psychiatric unit in a hospital (preferably a geropsychiatric unit)

► If in keeping with the residents' values and wishes, decide to switch to palliative care (with more liberal use of antipsychotics and, if necessary, short-acting benzodiazepines such as lorazepam to relieve the burden of intractable suffering, even at the cost of excessive sedation) as the resident is in severe to terminal stages of dementia and burdens of transfer to Emergency Department and hospitalization are high and potential benefits low.

Step 13: Attempt to reduce and discontinue antipsychotics after 3–6 months of stabilization

► Attempt to reduce and discontinue antipsychotic medication

Step 14: Assess response to dose reduction and discontinuation

► Lack of relapse of target symptoms

► Relapse of target symptoms (proceed to step 15)

Step 15: Reinstitute antipsychotics and attempt step 13 after 3–6 months of stabilization

Note: The psychiatrist may need to skip steps 5 and 6 if the resident who has major neurocognitive disorder has behavioral and psychological symptoms that are posing immediate danger to self and or others. In behavioral emergencies (e.g., violent behavior), immediate use of one or two doses of antipsychotic medication may be necessary to stabilize the emergency situation before stepped care approach is initiated. If, during step 3, the team concludes that the resident qualifies for end-of-life care, proceed to palliative care directly. Step 12 can be instituted at any point if there is radical change in the resident's clinical status. Collaboration with a pharmacist is encouraged at every step and in creating a facility-wide initiative to reduce inappropriate use of antipsychotics.

Among mood stabilizers, there is some support for the use of carbamazepine (Tegretol) to treat physical aggression in major neurocognitive disorder, but its use carries considerable risks, including falls, drug–drug interaction, and delirium. Valproate has been the most studied of all anticonvulsants for the treatment of behavioral and psychological symptoms of major neurocognitive disorder, but the results to date have been inconsistent or negative and its use carries considerable risks (especially falls and sedation). Hence, the authors do not recommend the use of valproate to treat behavioral and psychological symptoms of neurocognitive disorder.

SSRIs (especially citalopram [Celexa] and escitalopram [Lexapro]) and trazodone (Desyrel) are an option for the treatment of persistent agitation and aggression not responding to psychosocial environmental interventions. The psychiatrist may also consider prescribing antidepressants to treat severe emotional lability. Antidepressants, although safer than antipsychotics, carry considerable risks, including the risk of increased agitation, nausea, insomnia, and falls. The psychiatrist may consider prescribing a trial of a combination of dextromethorphan and quinidine (Nuedexta) to treat pseudobulbar affect as well as for moderate to severe agitation (off label) in the context of major neurocognitive disorder or stroke that is not responding to psychosocial environmental interventions. The psychiatrist may consider a judicious trial of

methylphenidate (Ritalin) for residents who have major neurocognitive disorder or traumatic brain injury and are experiencing severe apathy when psychosocial environmental interventions have not been effective and there are no cardiac contraindications. The use of benzodiazepines carries a high risk of serious adverse effects (especially delirium, falls, and injury) and thus should be avoided. Short-term use of low-dose short-acting benzodiazepines such as (lorazepam [Ativan] or oxazepam [Serax]) is an option for the treatment of prominent anxiety or infrequent severe agitation.

Mr. C is an 80-year-old male, married for 55 years, a retired school teacher and father of three children. He has had major neurocognitive disorder, probable Alzheimer type, for the last 8 years. Mr. C had been experiencing severe agitation and aggressive behaviors over the last several months. His wife reported that Mr. C would be "Dr. Jekyll and Mr. Hyde," at times pleasant and nice but then suddenly would "have a look of hatred" toward her and accuse her of infidelity. Mr. C had been physically aggressive toward his wife, which prompted admission to an assisted living home. At the facility, Mr. C attempted to choke a female resident, twisted the arm of a female staff member, and made several threats that he would "destroy" the place. Mr. C had also voiced paranoid thoughts that "they" (meaning the staff and residents) were trying to steal his money and belongings. Mr. C's primary care physician had done a thorough workup but found no reversible cause, so had prescribed escitalopram (10 mg daily), donepezil (Aricept), and as-needed alprazolam (Xanax). The staff had attempted various individualized psychosocial environmental interventions (e.g., take him for walks, play his favorite music in his room) but the aggressive behaviors persisted. At this point, psychiatric consultation was requested to facilitate transfer to a local acute inpatient geropsychiatric unit. After completing a thorough assessment and review of records, the psychiatrist encouraged the facility to give a trial of antipsychotics before transferring Mr. C to an acute inpatient psychiatric unit. The facility agreed. The psychiatrist recommended initiation of olanzapine 5 mg at bedtime, discontinued alprazolam, started scheduled acetaminophen as there were some reports from the staff that Mr. C occasionally complained of lower back pain, and added as needed orally dissolvable 5 mg olanzapine up to twice a day. Over the next 2 weeks, Mr. C's aggressive episodes improved modestly and the staff were able to engage him in more individualized psychosocial environmental interventions. After 4 weeks, his wife and the staff felt that Mr. C was much better, although he still had episodes of anger.

Clinical pearl: Judicious use of atypical antipsychotics may help reduce aggressive behaviors by residents who have major neurocognitive disorder and prevent transfer to the emergency department.

DEPRESSIVE DISORDERS

The prevalence of depressive disorders is high in long-term care residents. Hence, the authors recommend assessing all residents for depression within 1 month of admission to a long-term care facility and every 6 months thereafter. Residents who have a history of depression, psychiatric hospitalization, and/or antidepressant treatment have shown the strongest association with new onset of depression after admission to a long-term care facility, and thus would benefit from closer monitoring. The period of adjustment to living in long-term care can range from a few weeks to 12 months as residents cope with multiple losses, including independence,

home, privacy, possessions, and control over personal schedule. Additional challenges to living in long-term care include interaction with peers who have severe major neurocognitive disorder. High staff turnover and lack of robust activity programming may also contribute to depression.

Standardized scales recommended for assessing depression in long-term care residents include the Geriatric Depression Scale, used for residents who have mild cognitive impairment or mild dementia, and the Cornell Scale for Depression in Dementia, used for residents who have moderate to severe dementia. The Patient Health Questionnaire-9 or the Saint Louis University "AM SAD" (appetite, mood, sleep, activity, thoughts of death) may be used in place of the Geriatric Depression Scale.

Psychosocial and environmental strategies are appropriate and necessary and may be sufficient for the treatment of mild to moderate depressive disorder. Such interventions include psychotherapy, physical exercise, individualized pleasant activity schedule, bright light therapy, increasing social support, and complementary and alternative treatments such as yoga, Tai Chi, meditation, massage therapy, and aromatherapy. Psychotherapy with the long-term care population primarily involves supportive psychotherapy (including reminiscence therapy) with problem-solving strategies (especially for residents who have executive dysfunction) and/or interpersonal psychotherapy components (especially for residents coping with grief or family conflict). With more capable residents, cognitive-behavior therapy may also be appropriate. Malnutrition, dehydration, constipation, and pain are common comorbid conditions among residents who have depression and need to be optimally treated for successful outcomes.

Although rigorous studies have not consistently found antidepressants to be effective for the treatment of depressive disorders among long-term care residents, most experts and expert consensus guidelines support the use of antidepressants to treat moderate to severe depressive symptoms. The starting doses of antidepressants need to be low, but final effective doses may need to be comparable to those used by younger adults. One of the SSRIs, SNRIs, or mirtazapine is the preferred first-line treatment for depression if the psychiatrist opts for pharmacotherapy. Bupropion (Wellbutrin) and nortriptyline (Pamelor) are usually the second line of treatment. Nortriptyline may be preferred over other antidepressants to treat depression for residents who have Parkinson disease (PD). The psychiatrist should also consider prescribing pramipexole (Mirapex) to treat depression for residents who have PD; it may even be preferred over nortriptyline and other antidepressants, as it also improves motor symptoms of PD. The psychiatrist may consider prescribing short-term use of low-dose methylphenidate to treat depression with severe apathy to speed up the response to antidepressant, especially in the context of rehabilitation if the resident is not participating in physical therapy. The

psychiatrist should avoid prescribing monoamine oxidase inhibitors as well as antidepressants with high anticholinergic activity, such as amitriptyline (Elavil), imipramine (Tofranil), and doxepin (Sinequan), because of the high risk–benefit ratio. For residents who have psychotic depression, adding an atypical antipsychotic medication is recommended. For residents who have inadequate response to an antidepressant, the psychiatrist should consider adding aripiprazole or quetiapine or brexpiprazole as an adjuvant to treat moderate to severe major depressive disorder.

Mr. L is an 88-year-old widowed male, a retired construction worker living in a nursing home. He has a long history of alcohol addiction and alcoholic liver disease. Psychiatric consultation was made, as the staff felt Mr. L was depressed. The staff reported that for the last several weeks, Mr. L had been refusing to get out of bed, was not eating well, lost 15 lb, did not enjoy visits from his grandchildren, and had voiced multiple somatic complaints, including dizziness. During evaluation, the psychiatrist found out that Mr. L had had multiple recent losses: two of his children had died of heart disease in the last year, his pet dog had died 6 months ago (just before he moved into the nursing home), and 3 months ago his house was sold to pay for the nursing home. Mr. L became tearful during the interview. Workup for reversible conditions detected mild vitamin B_{12} deficiency and mild dehydration. There was no any evidence of major neurocognitive disorder, and Mr. L was open to recommendations from the psychiatrist, including antidepressants. The psychiatrist recommended vitamin B_{12} supplementation, improved intake of daily fluids, and weekly individual psychotherapy with a psychologist who visits the facility; the psychiatrist did not recommend antidepressants. Over the next 6 weeks, Mr. L showed significant improvement. He started to come out of his room, participated in activities, improved his food intake, gained 2 lb, and would look forward to visits from his grandchildren.

Clinical pearl: The psychiatrist should routinely consider recommending psychotherapy to treat depression for cognitively capable residents, and long-term care facilities should arrange with local psychotherapists to visit the facility and provide psychotherapeutic services.

PSYCHOTIC DISORDERS

Psychotic symptoms in the context of major neurocognitive disorder are the most common psychotic disorders in the long-term care population, followed by psychotic symptoms due to medication and due to general medical conditions (e.g., stroke, traumatic brain injury). Dopaminergic agents used to treat PD and corticosteroids are the two most common categories of medication that cause psychotic symptoms. Psychotic symptoms in the context of affective disorder (major depression as well as bipolar disorders), schizophrenia, schizoaffective disorder, and delusional disorder are becoming increasingly prevalent due to the aging of adults who have chronic mental illness. Psychotic symptoms may also be the initial or most disturbing manifestation of delirium. Psychotic symptoms are among the most serious complications of major neurocognitive disorder and are associated with a high degree of emotional distress and

accelerated cognitive decline. The psychiatrist should consider new onset of psychotic symptoms among long-term care residents as equivalent to delirium, with a thorough workup, including, as appropriate, neuroimaging.

Psychotic symptoms associated with major neurocognitive disorder are typically mild to moderate in severity and are best managed with psychosocial/environmental interventions. Antipsychotics have not been shown to reduce psychotic symptoms related to neurocognitive disorder, and the authors do not recommend their use unless psychotic symptoms are accompanied by severe aggression and or agitation. Management of psychotic symptoms in the context of PD is challenging, as the majority of residents who have PD have PD-related major neurocognitive disorder. Furthermore, only clozapine has been found to be effective for the treatment of psychotic symptoms in the context of PD, and the use of clozapine carries even higher morbidity and mortality risk for long-term care residents who have PD than for community-dwelling individuals who have PD due to frailty and the presence of significant cognitive impairment. The first strategy is to reduce the dosage of antiparkinsonian agents (such as Levodopa) and use psychosocial environmental interventions to manage psychotic symptoms. If psychotic symptoms are causing severe distress and agitation, the psychiatrist may consider prescribing low-dose quetiapine or aripiprazole. A new drug, pimavanserin is the only drug approved by the Food and Drug Administration for the treatment of psychotic symptoms for individuals who have PD. For residents who have neurocognitive disorder due to Lewy body disease and are experiencing psychotic symptoms, a trial of rivastigmine (Exelon) may help both cognitive and psychotic symptoms to a modest extent.

Ms. T is a 71-year-old divorced female, a homemaker and mother of one daughter and one son, who has a long history of schizophrenia. She was admitted to a nursing home after a stroke. Psychiatric consultation was requested to adjust Ms. T's antipsychotic medication, as she was refusing to participate in physical therapy and making paranoid statements. After completing a thorough evaluation, including review of records from previous psychiatric hospitalization, the psychiatrist confirmed the diagnosis of schizophrenia. Ms. T also had minor neurocognitive disorder due to cerebrovascular disease. The psychiatrist noticed that many staff members felt that Ms. T should be transferred to another facility, as they were not equipped to manage chronic mental illness. Ms. T's primary care clinician was also not comfortable managing chronic mental illness by adjusting antipsychotics. Ms. T was taking quetiapine 200 mg daily at bedtime and risperidone 0.5 mg once daily. The psychiatrist increased risperidone to 0.5 mg twice daily for the first week and then 1 mg twice daily and decreased and discontinued quetiapine over 2 weeks. Ms. T showed significant improvement in psychotic symptoms over the next 3 weeks and began to participate in physical therapy.

Clinical pearl: Individuals who have chronic mental illness such as schizophrenia or bipolar disorder are living longer, and the prevalence of these illnesses in long-term care facilities is expected to increase. The majority of facilities are not prepared to meet their needs. Psychiatric

consultants can play a key role in improving their acceptance by facility staff and help improve the staff's understanding of these illnesses and knowledge about evidence-based strategies to treat them.

DISORDERS OF THE SLEEP–WAKE CYCLE

Insomnia, excessive daytime sleepiness, and reversal of the night–day cycle are common among long-term care residents. Medications are the most common cause of disorders of the sleep–wake cycle. Restless leg syndrome (RLS) and REM sleep behavior disorder (RBD) are prevalent in the long-term care population, are often underdiagnosed, and are amenable to treatment. The primary treatment for insomnia should be psychosocial environmental interventions (e.g., reducing daytime napping, increasing physical activity and exposure to sunlight or bright light, following a regular sleep time, reducing exposure to noise and light at night). The psychiatrist may consider prescribing low-dose trazodone to treat persistent insomnia not responding to psychosocial environmental interventions. The use of hypnotics such as zolpidem (Ambien), zaleplon (Sonata), and eszopiclone (Lunesta) should be restricted to low doses for short periods of time (days to 2 weeks) (e.g., for a new resident adjusting to living in a long-term care facility and experiencing significant insomnia) due to a high risk–benefit ratio and regulatory concerns. Ramelteon (Rozerem) and surovexant (Belsomra) may also be considered for treatment of insomnia. The psychiatrist may consider prescribing low-dose clonazepam (Klonopin) and melatonin to treat RBD and may consider prescribing dopamine agonists and gabapentin enacarbil (Horizant) to treat RLS.

POSTTRAUMATIC STRESS DISORDER

Posttraumatic stress disorder (PTSD) is increasingly seen in the long-term care population, not only among veterans but also among civilians who have a past history of abuse or motor vehicular accident. Long-term care staff needs education to understand psychological and behavioral manifestations of past trauma and training in care practices to accommodate trauma-related challenges residents typically face. For example, male staff members should not be involved in caring for a female resident who has a history of sexual or physical abuse by a male. Physical restraints are not recommended in general but may have additional negative psychological effects on residents who have a history of abuse. The treatment of PTSD typically involves antidepressants (especially SSRIs) and, when feasible, individual psychotherapy. Short-term or as-needed use of short-acting benzodiazepines may be necessary for residents who have PTSD and are experiencing acute anxiety attacks. Low-dose prazosin (Minipress) may be appropriate for the treatment of nightmares in the context of PTSD, although long-term care residents are at higher risk than

the general population of adverse effects of prazosin, such as hypotensive effects, orthostasis, and falls.

ANXIETY DISORDERS

Generalized anxiety disorder (GAD), panic disorder, and phobias, especially bathing phobia, are prevalent in the long-term care population. Bathing phobia is also common among residents who have major neurocognitive disorder and is best managed by staff training in creative and alternative ways to maintain hygiene using strategies such as towel bathing instead of using a shower. The authors recommend the DVD "Bathing without a Battle," created by the University of North Carolina, for training all staff members in person-centered strategies to bathe residents who have advanced major neurocognitive disorder. Treatment of GAD and panic disorder typically involves antidepressants (especially SSRIs) and, when feasible, individual psychotherapy.

PSYCHIATRIC ASPECTS OF PALLIATIVE CARE

More than 25 percent of older adults currently die in a long-term care facility, and the number is expected to rise to more than 40 percent by 2030. The majority of residents have at least one life-threatening illness, and at least 50 percent of these have a life expectancy of 2 years or less. Thus, it is imperative that all these residents be offered palliative care as soon as possible after admission. Psychiatric consultants can help the staff appreciate that most residents and their family members find the goal of increasing comfort much more satisfying and reasonable than aggressive life-prolonging care. The psychiatrist can play a key role in clarifying a resident's advance care directives, especially in the context of advanced major neurocognitive disorder, where it may be difficult to predict life expectancy but the burdens of traditional medical care may still be high.

Delirium and depression are two psychiatric syndromes residents frequently experience at the end of life. Psychiatric consultants may be the first health care professional to identify the resident as in the process of dying and help initiate palliative and hospice care. During end-of-life care, it is appropriate and often necessary to prescribe benzodiazepines to manage anxiety and insomnia, morphine to manage pain and dyspnea, and atypical antipsychotics such as risperidone and olanzapine to manage psychotic symptoms associated with agitation or aggressive behavior. The use of methylphenidate to treat depression during end-of-life care has been recommended for its rapid effect but among long-term care residents, who often have advanced major neurocognitive disorder, methylphenidate may cause delirium or agitation and thus the psychiatrist should closely monitor its use. Palliative sedation may be necessary to relieve the intractable suffering some residents experience during the end stages of life.

Another common reason for referral to a psychiatrist is for evaluating

decision-making capacity. Some common clinical situations for such a referral are: a resident who has end-stage renal disease and wants to discontinue dialysis; a resident who has depression and is requesting a Do Not Resuscitate status.

References

American Geriatrics Society. Guide to the management of psychotic disorders and neuropsychiatric symptoms associated with dementia in older adults. Available at http://dementia.americangeriatrics.org/GeriPsych_index.php. Accessed December 29, 2014.

American Geriatrics Society. Postoperative delirium in older adults. Best Practice Statement from the American Geriatrics Society. Available at http://www.journalacs.org/article/S1072-7515%2814%2901793-1/pdf. Accessed December 29, 2014.

American Geriatrics Society. 2015 updated Beers criteria for potentially inappropriate medication use in older adults. American Geriatrics Society Beers Criteria Update Expert Panel. *Journal of the American Geriatrics Society*. 2015;63:2227–2246.

American Psychiatric Association. The American Psychiatric Association Practice Guideline on the use of antipsychotics to treat agitation or psychosis in patients with dementia. 2016. Availablr at http://psychiatryonline.org/doi/pdf/10.1176/appi.books.9780890426807. Accessed January 17, 2017.

Canadian Coalition for Seniors' Mental Health. *National Guidelines for Seniors' Mental Health: The Assessment and Treatment of Mental Health Issues in Long-Term Care Homes (Focus on Mood and Behavior Symptoms)*. Available at www.ccsmh.ca. Accessed October 1, 2016.

Desai AK, Lo D, Grossberg GT. Depression in older adults receiving hospice care. In: Lavretsky H, Sajatovic M, Reynolds CF III, eds. *Late-Life Mood Disorders*. Oxford: Oxford University Press; 2013.

Desai AK, Grossberg GT. *Psychiatric Consultation in Long-Term Care: A Guide for Health Care Professionals*. 2nd ed. London: Cambridge University Press. (In Press)

Gitlin LN, Kales HC, Lyketsos CG. Nonpharmacologic management of behavioral symptoms in dementia. *JAMA*. 2012;308:2020–2029.

Hui C, Sulzer D. Depression in long-term care. In: Lavretsky H, Sajatovic M, Reynolds CF III, eds. *Late-Life Mood Disorders*. Oxford: Oxford University Press; 2013.

Janssen DJ. Palliative sedation in nursing homes: a good death? *J Am Med Dir Assoc*. 2014;15:541–543.

Kales HC, Gitlin LN, Lyketsos CG; Detroit Expert Panel on Assessment and Management of Neuropsychiatric Symptoms of Dementia. Management of neuropsychiatric symptoms of dementia in clinical settings: recommendations from a multidisciplinary expert panel. *J Am Geriatr Soc*. 2014;62:762–769.

Kitwood T. *Dementia Reconsidered: The Person Comes First*. Philadelphia, PA: Open University Press; 1997.

Kotlyar M, Gray SL, Maher RL Jr, Hanlon JT. Psychiatric manifestations of medications in the elderly. In: Agronin ME, Maletta GJ, eds. *Principles and Practice of Geriatric Psychiatry*. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2011.

Mace NL, Rabins PV. *The 36-Hour Day: A Family Guide to Caring for People Who Have*

Alzheimer Disease, Related Dementias, and Memory Loss. 5th ed. Baltimore, MD: Johns Hopkins University Press; 2011.

McFadden SH, McFadden JT. *Aging Together: Dementia, Friendship, and Flourishing Communities*. Baltimore, MD: Johns Hopkins University Press; 2011.

Moore A, Patterson C, Lee L, Vedel I, Bergman H; Canadian Consensus Conference on the Diagnosis and Treatment of Dementia. Fourth Canadian Consensus Conference on the diagnosis and treatment of dementia. *Can Fam Physician*. 2014;60:433–438.

Morley JE. Certified nursing assistants: a key to resident quality of life. *J Am Med Dir Assoc*. 2014;15:610–612.

Morley JE, Cao Li, Shum CK. Improving the quality of End-of-life care. *J Am Med Dir Assoc*. 2016;17:93–95.

Pitfield C, Shahriyarmolki K, Livingston G. Systematic reviews of stress in staff caring for people with dementia living in 24-hour care settings. *Int Psychogeriatr*. 2011;23:4–9.

Powers BA, Watson NM. Spiritual nurturance and support for nursing home residents with dementia. *Dementia*. 2011;10:59.

Rosenblatt A, Samus QM. The aging patient and long-term care. In: Agronin ME, Maletta GJ, eds. *Principles and Practice of Geriatric Psychiatry*. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2011.

Sabat R. *The Experience of Alzheimer's Disease: Life through a Tangled Veil*. Oxford: Blackwell Publications Inc.; 2001.

Simmons SF, Rahman AN. Next steps for achieving person-centered care in nursing homes. *J Am Med Dir Assoc*. 2014;15:615–619.

Steinberg M, Lyketsos CG. Atypical antipsychotic use in patients with dementia: managing safety concerns. *Am J Psychiatry*. 2012;169:900–906.

Volicer L. Nursing home telepsychiatry. *J Am Med Dir Assoc*. 2015;16:7–8.

57.6b Forensic Aspects

DAVID NAIMARK, M.D., KEITH A. DAVIDSON, J.D., AND ANSAR M. HAROUN, M.D.

The American Board of Psychiatry and Neurology has recognized both geriatric psychiatry and forensic psychiatry as subspecialties, each meriting their own fellowship training programs and their own subspecialty board examination. These two fields have an overlap, and this field of "forensic geropsychiatry" is growing in its knowledge base and literature. Unlike clinical geropsychiatry, which evolves slowly based on the natural progression of knowledge in the basic psychiatric sciences, forensic geropsychiatry may grow rapidly with abrupt changes, based on any new pronouncement by a court of law. Whereas geriatric psychiatry is generally defined by a demographic characteristic (that of age), forensic psychiatry is defined by its ethical duty to tell the truth. This is in contrast to the duty of clinical psychiatry, which is to be helpful to the patient. Much of forensic work within the field of geriatric psychiatry requires attention to this tension between these two competing duties, which may or may not conflict.

CRIMINAL ASPECTS OF FORENSIC GEROPSYCHIATRY

Epidemiologically, there is a relationship between criminal behavior and age, with an overrepresentation of youthful offenders and an underrepresentation of elderly offenders. In absolute terms, criminal behavior remains a primary problem of the young, but the relative rate of increase in criminality is more pronounced in the elderly. This trend may partially be accounted for by the population increase in the total number of elderly, resulting in a corresponding increase in bad behavior. There are, however, suggestions that there may be an independent increase in criminal behaviors in the elderly for other reasons. Possible explanations include the improved efficiency of law enforcement to detect crimes or new patterns of crimes (such as Internet-based crimes) that may be easier for physically disabled elders to commit. It is also useful to look at the relationship of age to subtypes of criminal behavior. One typology of criminal behavior contrasts crimes that are affective (based on emotion, impulsive) with those that are predatory (coldly calculated, premeditated). While youthful offenders are often hotheaded and at high risk for the former, the elderly are more mature, are less hotheaded, and may be at greater risk for the latter.

The expertise of a geriatric psychiatrist may be needed at several stages of a criminal proceeding. At the time of the arrest, when elder defendants are read legal rights, do they understand those rights? If not, any statements made in response to police questioning may not be admissible in a court proceeding. Older individuals (or any other persons who may be

confused, have mild cognitive impairment, or be demented) are at risk for not understanding the warnings that the arresting officers are required to state and the understanding (or lack of understanding) can change the outcome of a legal case by allowing (or disallowing) the statements or confessions. In real life, it may be difficult for a psychiatrist to separate such cases into the two boxes (did or did not understand) as many elders are in a gray zone of partial understanding. As a result, what may be relevant was whether the elder spoke to the police voluntarily or through coercion. While it may be rare for the police to coerce using physical pain, they may be tempted to coerce using psychological pressures to which the elderly may be especially susceptible (see the following discussion of undue influence). A practical problem that may arise from such possible coercion is that the elder may "confess." If the confession is voluntary, it is of course admissible, but if coerced, it may be challenged. Analysis of the voluntariness of such a confession requires knowledge not only of the behavior of those who elicited the confession (usually the police), but also of the unique strengths and frailties of the confessor, both psychological and physical.

Competency to Stand Trial

When a criminal trial starts, the court must be satisfied that the defendant is not only physically present, but also mentally present (i.e., mentally competent to stand trial). The psychiatrist may be asked to give an opinion around the cognitive status of the elder (whether there is any deficit in *understanding* the nature of the criminal proceedings), the social capacities of the elder (ability to *assist* an attorney), and the rationality of the elder (able to understand and assist in a *rational* manner). Further, the psychiatrist must opine whether any deficit is from a mental disease (making the defendant mentally incompetent) or from some other cause. An elder who cannot assist because of a hearing loss is "audiologically" incompetent, but not mentally incompetent. Problems may arise from elderly defendants who are forgetful. If this represents a mental disorder, such as a dementia of the Alzheimer type, then the defendant may postpone or even avoid trial (often a desired outcome), but if the forgetfulness represents age-related cognitive decline, then it may not qualify for a finding of mental incompetence.

JO had a clear psychotic disorder at the time of the trial. He had numerous delusions and hallucinations. Most of these concerned the fidelity of his spouse, though not all. He seemed to have a basic understanding of the trial process, with the exception of two unusual beliefs: While he could tell the evaluator that the judge oversees the trial and tells the jury the law, and the jury is the body that decides, after hearing the evidence, whether he is guilty or innocent, he also had the belief that the judge was an "interloper" and the jury members were "four-footed animals."

Is this defendant incompetent to stand trial? JO's competency to stand trial is likely to be controversial. Is his understanding of the functions of judge and jury adequate? Does it matter

if his conception of the judge and jury is irrational if he knows their function? On the other hand, perhaps he thinks their status as interloper and animals will make them perform their jobs in a severely substandard way—a way that will deprive him of due process.

Involuntary Medication to Restore Competence

If the court finds that a defendant is incompetent to stand trial, the court may seek further opinion from the psychiatrist as to whether the court may or may not order restoration of competency through involuntary medications. Such an analysis in a geriatric case may differ from a nongeriatric case, since the balancing of the defendant's interests (to be free of side effects of psychotropic medication) and the State's interests (to restore the defendant to competency for trial) may be altered. Geriatric patients may have different responses (good and bad) to medications, which must be separately analyzed in each case. In the case of the elder who may have many medical conditions complicating the side-effect profile of the intended psychotropic medication, a different conclusion may be reached than in the case of a nonelder, making it more likely that the court may reject the application for involuntary medication, thus creating a class of "permanently incompetent" persons. Up to the present time, such cases of the permanently incompetent have been rare, since most cases of trial incompetence arise from psychotic-like disorders, which mostly do respond to antipsychotic medications.

At some point in the future, as more elderly enter the criminal justice system and as more elderly are found to be incompetent to stand trial, it is likely that there will be more cases of trial incompetence from cognitive disorders (such as the dementias) rather than from the psychoses. Since some cognitive disorders are reversible and some are not, geriatric psychiatrists will increasingly be asked to separate these two subpopulations. Ultimately, there will be a large group of elderly persons who are permanently incompetent to stand trial. Permanent incompetence is potentially a problem because the Supreme Court has said that one cannot hold a person as incompetent to stand trial for more than the reasonable period of time necessary to determine whether there is a substantial probability that he or she will attain that capacity in the foreseeable future—and these individuals will not. On the other hand, such persons can be held under ordinary civil commitment laws, which would allow most of them to be committed: they have a mental illness that makes them dangerous. Still, the policy problem remains: such persons cannot be tried (since they are incompetent to stand trial), but they cannot be released (since they may repeat criminal behavior), and there is no point in confining them for treatment (since they are incurable).

Mental Health Defenses

Most criminal defendants are immediately or eventually found competent

to stand trial and during the trial may assert any one of several mental health defenses, including psychiatric excuses and/or justifications. In the movies (and hence, in the popular imagination), a common mental excuse is that of legal insanity. This requires a confession of behavioral guilt ("I did it") but an absence of moral guilt, based on some mental deficit (cognitive or volitional). Elderly defendants may specifically be at greater risk for cognitive deficits. The geriatric psychiatrist should clarify the legal standard in the jurisdiction of practice, especially noting what types of cognitive deficits represent an excusing condition. Cognitive deficits around memory ("I don't remember what I did"), while common in the elderly, are unlikely to excuse, whereas cognitive deficits in the areas of understanding are more likely to excuse. Further analysis may be required as to whether the court wants an opinion on the elder's capacity to know, to understand, or to appreciate. An elder who is having some cognitive decline is likely to first lose the capacity to appreciate, then to understand, while the ability to merely know is likely to be the last to be lost.

Restoration of Sanity

If an elder is indeed found to be legally insane, the state must make efforts to restore sanity. In the case of youthful offenders, most of the not guilty by reason of insanity acquittals are likely to be based on a severe mental illness that is treatable, usually of the psychotic type. But in the case of the elder, the acquittal may be based on a mental disorder not of the psychotic type (with a likely good response to medication), but rather of a cognitive type (which may or may not respond to medication). If the cognitive decline is medically reversible, the court should be so informed, and will order the appropriate treatment. If there is an irreversible dementia, however, this will create a problem for the court, since no punishment may be imposed (the defendant is "not guilty"). Although the defendant may be confined for treatment, he or she may not be confined in the absence of treatment. In many court cases, however, the right to treatment is extremely minimal—for example, enough to help keep people out of restraints. The fact that certain people are effectively untreatable does not mean that they cannot be confined if they are dangerous or gravely disabled. Notwithstanding this, a skillful defense attorney may argue that the elder must be treated or released, thus creating a dilemma for the court. The case of an elder being acquitted after a finding of not guilty by reason of insanity is uncommon; more common is a verdict of guilty (or not guilty). In such cases, the court must decide on a sentence.

Sentencing

In sentencing, judges are guided by many factors, including some age-related factors like the epidemiological risk of future harm (lower in the

elderly), the degree of culpability based on gullibility (the elderly may be more gullible, and hence less blameworthy), the likelihood of physical violence (lower in the elderly), and the likelihood of good compliance with medication to achieve remission of mental illness (which may be lower in medically impaired subjects, whose medical status may make tolerating psychotropic medications more difficult). Such objective factors should be noted in a presentencing report for a judge. Subjective factors (whether the elder did or did not inspire pity in the examining doctor) should ideally not be given weight, but if the doctor is temperamentally compassionate and is aware of feelings of pity that may introduce a bias, the doctor should either avoid giving an opinion (which could be invalid) or should give the opinion but clarify why the feelings may not have produced a valid opinion.

The geropsychiatrist should be aware of the possible purposes of sentencing, which can be punitive, nonpunitive, or a combination of the two. If the purpose of the sentence is punitive (to punish), it is uncommon that a court seek a psychiatric opinion, although a skillful defense attorney may argue that the very punishment that may be appropriate for a youthful offender may be cruel or unusual for an elder, who may have a reduced capacity to tolerate discomfort. If the court then moves to a less punitive sentence (perhaps some combination of restitution, incapacitation, or rehabilitation), the geropsychiatrist can offer an opinion as to the merits of any of these. Philosophical problems may arise from sentences that are "rehabilitative," since it may be argued that these are doomed to fail (since the elderly criminal may be too old to reform), or that even if rehabilitation is possible, it fails the test of cost–benefit analysis (the costs of rehabilitation perhaps should not be incurred in late life, when there will be only a short time to reap the benefits).

CIVIL ASPECTS OF FORENSIC GEROPSYCHIATRY

Civil matters, involving a dispute between parties, constitute the majority of the cases for which a forensic geropsychiatrist may be retained. There are two broad legal claims that require the assessment of a subject's mental state: (1) legal capacity, also known as decisional capacity to do a certain act (such as make a contract, create a will, give informed medical consent, or appoint an agent), and (2) undue influence, also known as coercion. Civil commitment issues, substituted decision making (i.e., living wills, advanced directives, and durable powers of attorney), elder abuse, and nursing home malpractice are additional areas for possible geropsychiatric involvement.

Decisional competence is a legal concept that concerns the ability of an individual to comprehend and manipulate the relevant information for a decision. While the doctor analyzes the decisional capacity (a dimensional construct) and forms a penultimate medical opinion, the trier of fact forms a legal judgment about competence (a binary construct), which is the

ultimate issue.

Undue influence is a legal concept where a person's free will is overcome by the excessive persuasion of another, also commonly understood as coercion. While there are a number of factors to establish a legal claim for undue influence, the geropsychiatric involvement primarily concerns determining whether a person is susceptible to undue influence.

Informed Consent

In settings of clinical work and research, informed consent is one of the cornerstones of ethical practice. The informed consent process ideally epitomizes professional ideals and protects individual rights. Often, unfortunately, the informed consent discussion morphs into a procedure done to the patient or subject (i.e., "consenting" the individual). This implies that the elder is a passive recipient of information who then acquiesces to the decision. Optimally, however, the informed consent process can provide a meaningful chance for a physician and elder to participate in a personalized discussion of risks/benefits and alternatives associated with a decision in question. In this process, the patient becomes a partner with the physician in the management of the situation, and hopefully thereby engages in authentic decision making. Common disputes leading to forensic evaluations in this area would be analysis of capacity to (1) consent to medical treatment, (2) consent to be a research subject, and (3) decide to die. In life outside of medicine, issues of decision making present themselves constantly. Unlike in the medical world, however, the rules around the need for the transactor to obtain informed consent from the transactee become markedly murkier (i.e., does a Wal-Mart employee need to obtain informed consent before he allows an elder to purchase socks?). Common disputes leading to geriatric forensic evaluations in this area would be analysis of capacity to (1) make financial decisions/contracts, (2) make a will (testamentary capacity), and (3) drive.

Over the last half-century, multiple state and federal court decisions have greatly influenced the analysis of the concept of informed consent. What the doctor must disclose is measured by two standards. The older rule, and one contained in some statutes, is that the doctor must disclose what a reasonable doctor would disclose. The newer rule is the "reasonable patient" standard: The doctor must disclose all material information that he or she could reasonably expect a patient would want to consider in determining whether to undergo the procedure. Informed consent requires not only that the patient is informed, but also that he or she have the capacity to decide and that his or her decision should be voluntary. The latter was called into question in one case: An involuntarily confined prisoner was deemed unable to consent to experimental psychosurgery because his choice would be coerced as a matter of law.

The first element of informed consent refers to "informed" and asks

whether the information from the database (regarding the specific decision in question) was transmitted to the decision maker. It does not assess the state of mind of the transactee and, as such, is a responsibility of the transactor. With inadequate or absent information, the consent would be uninformed, not informed.

The second element of informed consent addresses the "sent" (from the same root as *science*, as in knowledge) and requires the elder to have the mental capacity to understand and appreciate the information provided. The best information in the universe is useless to the elder whose cognitive plumbing is clogged to the degree that it simply can't pass through the pipes of mental functioning. It is here that the forensic evaluator will perform the analysis of decision-making capacity through a mental functioning examination. Mental status deficits, in and of themselves, do not lead to mental incapacity to make a decision. It is only when the observed defect prevents the information from penetrating into the awareness of the subject that the defect becomes relevant. Put another way, mental state defects are a "threshold test"—some decisions are complicated and require a maximally functional mental state, whereas others are simple and probably require no more than a marginally intact mental state.

The third, and perhaps most misunderstood, element of informed consent concerns the "con" part of consent (distinguishing it from assent), and requires the assessment of whether the elder was able to give a voluntary agreement to the transaction. Lack of voluntariness can be as simple as the overt threatening of an elder unless he or she agrees to a decision or as subtle as the insinuation into a susceptible elder's life to the degree that the elder becomes unduly influenced, and thereby unable to give a voluntary consent. It is important to understand that "influence" is generally construed to be acceptable, as all human beings influence each other repeatedly every day. It is only when the influence progresses to the point that the will of the elder has been substituted by another's will (generally for the purpose of extracting money or other property) that the court may declare that the influence is undue. Although there may be considerable disagreement about when undue influence may exist, there are some circumstances where undue influence likely cannot occur. These circumstances include if the elder (1) knows his or her own mind, (2) can distinguish between his or her interests and the interests of another, and (3) can distinguish a neutral disinterested assertion from an active, persuasive one. These conditions emphasize cognitive constructs (i.e., knowing and distinguishing) but seem to leave out issues of emotional knowing. If there is no emotional vulnerability leading to emotional dependence, undue influence probably cannot occur.

The three elements of informed consent clearly overlap (i.e., mental capacity is related to the ability to give voluntary consent) and the distinction is, therefore, somewhat artificial. It is helpful for clarity, though,

for the forensic evaluator to separate the elements as much as possible.

General Procedure to Assess Undue Influence

While every jurisdiction has different elements for undue influence, the legal concept can best be understood to consist of four elements: (1) vulnerability of the victim, (2) the influencer's apparent authority, (3) the actions or tactics used by the influencer, and (4) the equity of the resulting decision or action. Undue influence can be used to set aside any type of decision or action taken, including financial decisions, contracting, medical decisions and consent, and will or trust creation.

Vulnerability of the Victim. This element is the most relevant to geropsychiatric involvement because it requires a finding of at least one of the following conditions: incapacity, illness, disability, impaired cognitive function, emotional distress, isolation, or dependency.

Influencer's Apparent Authority. An influencer's apparent authority focuses primarily on the relationship between the influencer and the victim. Relationships that create an increased risk of undue influence include fiduciaries, family members, care providers, health care professionals, legal professionals, spiritual advisers, financial advisers, and experts.

Actions or Tactics Used by the Influencer. Actions and tactics focus on the bad acts undertaken by the influencer to suppress the free will of the victim. Actions and tactics include controlling necessities of life, medicine, and access to family and friends, use of affection, intimidation or coercion, initiating changes in personal or property rights, use of secrecy in making financial changes, or making changes at inappropriate times and places.

Equity of the Decision or Action. Equity is a legal concept that refers to the fairness of the result. If a person has a long history of benefitting family members under his or her will, and then suddenly changes that plan to benefit a caretaker, this could be viewed as an inequitable result under the circumstances. While equity, by itself, is not enough to establish undue influence, an apparent inequitable result can be persuasively used to help support the other elements of undue influence.

Presumption of Undue Influence. Most jurisdictions have a general presumption that undue influence exists where: (1) a confidential relationship exists between the influencer and the victim (much like the apparent authority element discussed above), (2) the influencer actively participated in the procurement of the decision or action, and (3) the influencer unduly benefitted from the decision or action taken. Where a

presumption exists, it is then the burden of the influencer to prove the absence of undue influence, which can be difficult to accomplish in most cases.

Competency to Make Medical Decisions or Consent to Be a Research Subject

While the "meat" of the medical/research capacity evaluation is the forensic interview, there are certain standardized assessment instruments that may be useful. These tools (the MacArthur Competence Assessment Tools) have been shown to have good validity and reliability in multiple populations. Included are semistructured interviews for evaluating the four aspects of decision-making capacity as described in the previous section. Recently, a new instrument for assessing decisional capacity for clinical research has been validated. The University of California, San Diego, Brief Assessment of Capacity to Consent (UBACC), a 10-item scale focusing on understanding and appreciation of information contained in a research protocol, was found to have good reliability and validity, taking less than 5 minutes to administer.

Recent research suggests that certain procedures for "enhancing" informed consent for research and treatment may be helpful. Effective interventions include corrected feedback, multiple learning trials, and well-organized simplified consent forms. Even in elderly individuals with psychotic disorders, enhanced consent procedures involving a computerized slideshow have recently been shown to improve understanding of research consent. When elderly people are incapable of deciding whether to participate in research, the question arises whether their family—or some other proxy—may consent on their behalf. The federal regulations allow a legally authorized representative (LAR) to do so, leaving the question of who may serve as an LAR to the states. Study of the state regulations shows fairly wide variability in the state statutes, with most states not weighing in on whether families may consent. The important questions that remain are whether proxies are/should be permitted to consent to research on behalf of an incompetent subject; if so, who may/should serve as available proxies and in what order; and whether there are/should be limits on when proxies may decide in the research context. Further study of these issues is warranted, as is gathering data on what is actually being done and how well it is working.

Competency to Make Decisions about Death and Dying

Over the last generation, there has been a change in societal attitudes to death and dying and this is reflected in the way in which geriatric patients may view this topic. Throughout history, life has been mostly valued and cherished, while death has been feared. The respect given to the healing professions stemmed partly from the ability of the healer to prolong life

and postpone death. But recently, scholars from both medicine and medical ethics have looked more critically at what should be valued in life, its quantity or quality? There is now a school of thought that argues that it is the quality of life that is more valuable, and that if a good enough quality of life is not possible, death should not be feared and postponed at all costs, but rather considered and possibly welcomed.

Elderly patients may raise questions about: (1) treatment refusal, (2) suicide, (3) physician-assisted suicide, and (4) euthanasia. Geriatric psychiatrists are likely to face such questions in elderly patients who approach their end of life, and may wonder how best to proceed. Is the role of the psychiatrist to ascertain the wishes (conscious or unconscious) of the patient? And promote those wishes, even if the wishes are unconventional (not to prolong life, but to shorten it)? But what if those wishes conflict with: (1) the values of the psychiatrist (who may wish to prolong life at all cost), (2) the law, which may criminalize medical behaviors such as physician-assisted suicide (or any action that even resembles that), and (3) societal pressures which may make it difficult for the psychiatrist to support the patient's decision to refuse treatment, resulting in an earlier death, and the risk of a medical malpractice law suit.

If an elderly patient indicates that she needs to make a decision to promote death or dying, the psychiatrist should take a number of clinical and forensic steps, beginning with a determination of the quality of decision making (i.e., is the decision contaminated by mental illness). If the decision to seek death is causally related to a mental disorder, then the ethical task is simple—treat the mental disorder, and reassess. The ethical task is more complex if the decision (not to promote life, but to seek death) is not related to a mental disorder (so there is no disorder to treat), and the decision is rational, suggesting the patient is competent. Under the law, competent patients are generally free to make decisions about their health status, including decisions (e.g., to refuse treatment) that may hasten death. How should the geriatric psychiatrist then proceed?

Strictly following the law (allowing the competent senior to make medical decisions, including those that may promote death) may not seem ethically right. But not following the law, by denying a competent senior the right to make medical decisions is also fraught with dangers! In such cases, where there is a clash (real or apparent) between law and ethics, the clinical geropsychiatrist should seek consultation from a forensic colleague. The forensic specialist may then advise the clinician that it is the doctrine of Informed Consent that best balances the competing demands from law, ethics, and psychiatry. The clinical psychiatrist should ensure that all three elements of informed consent are met. Since such cases have the potential for careful scrutiny by legal agencies, ethics committees, and the media, the documentation of how the elements of informed consent were met should be documented in some detail.

The elements of informed consent may take on added complexity with respect to decision making in death and dying. Much interest is now focused on the potential for abuse in the process of informing, rather than on the content of information. Previously, it was assumed that competent patients were rational and therefore blind to the framing of the information. In fact, one of the defining criteria of rationality is that a rational person is blind to such framing (the way the information is presented). It is now believed that most individuals are not blind to framing effects, but improperly biased by them. This may create a situation where a patient is properly fully informed (in the legal sense), but disinformed in the psychological sense.

A doctor is required to inform the patient of the risks of a medical treatment. The doctor has the choice of framing the information in positive terms ("75 percent probability of success") or in the alternative ("25 percent probability of failure"). The two pieces of information are mathematically equivalent, and a rational agent should be indifferent to how the information is presented. But most patients process these two pieces of information differently, and are likely to reach differing conclusions depending on the presenting frame.

An ethical doctor should inform the patient of the choices open to the patient, and to the risks associated with the choices. If the doctor presents the patient with a forced choice (simplifying the treatment decision to one of two), and if one option leads to certainty, while the other option has an uncertain outcome (most medical treatments do not offer any guarantee of success), the patient may make a medical decision based not on the merits of the case, but on the risk preference of the patient. Most humans are risk-averse (are more comfortable with outcomes that are certain), rather than risk-seeking (taking a gamble on a decision that has an uncertain outcome). The doctor may inform the patient of possible intervention in one of two ways: either emphasizing the advantages (benefits) of participating in the intervention, or emphasizing the disadvantages (risks and costs) of not participating. The two informing processes are logically equivalent, but produce different outcomes. Most patients are likely to be more swayed by the doctor who emphasizes the disadvantages of not participating in the intervention, and therefore are more likely to consent.

Another new area of interest is the distinction between a threat and an offer. Informed "consent" based on a threat is clearly not a valid consent, because it was not given on a voluntary basis. But what if the doctor denies making a threat (which invalidates the consent), but reframes it as an offer (which does not invalidate the consent). For example, if a doctor says to a senior patient: "I wish to inform you of your medical choices; if you seek treatment, you will exhaust your finances, soon become poor, a burden on your children, and they will hate you . . . but if you refuse treatment, you will die an earlier death, with dignity, with your financial assets preserved,

which you can pass on to your children, who will cherish you and your memory . . ."

If the senior makes a medical decision (to refuse treatment), was that decision voluntary, or was it subtly coerced, by a threat pretending to be an offer? Some philosophers argue the distinction is based on whether the proposed outcome (in this case death with dignity) makes the patient better off, or worse off than they are now. But in such cases, this is unclear, a value judgment, and not a medical judgment.

Competency to Make Financial Decisions or Contract

From a legal perspective, some forms of competency have a higher threshold to meet than other forms of competency. For example, competency to make financial decisions or contract is one of the highest levels of legal competency because it requires more mental cognition and reasoning abilities than capacity to make a will (also called testamentary capacity). In fact, in most jurisdictions across the United States a person can lack the capacity to make financial decisions or contract (such as individuals under a guardianship or conservatorship proceeding), but still have the necessary legal capacity to make a will.

Competency to make a financial decision or contract is usually questioned while the patient is alive (in contrast to testamentary capacity, which is often retrospectively questioned when the elder is deceased) and often occurs in the context of a dispute involving an elder's right to give money to a person of his or her choosing, particularly when a suspicion exists that the elder is being financially exploited.

Generally, to be found legally competent to make financial decisions or contract, the subject must be able to understand and appreciate, to the extent applicable (1) the rights, duties, and responsibilities created by or affected by the decision, (2) the probable consequences of the decision, and (3) the significant risks, benefits, and reasonable alternatives involved in the decision. In contrast to testamentary capacity (discussed in more detail below), this standard is vastly different and requires more mental functioning.

A financial decision or contract can also be overturned based on the legal concept of undue influence. In this context, undue influence focuses more on the actions of the wrongdoer and looks to mental functioning only to determine if the subject is susceptible to undue influence. A person can be susceptible to undue influence, but still have the requisite legal mental capacity to make a financial decision or contract. At times, undue influence can be an easier claim to prove in court as the mental functioning aspect is not as stringent and makes up only one of four elements to prove undue influence (the four elements being: (1) vulnerability of the victim, (2) apparent authority of the wrongdoer, (3) actions and tactics of the wrongdoer, and (4) equity of the result).

Diane is an 85-year-old female, who has always been independent, and refuses to talk about her finances with her only daughter, Tina Smith. Tina lives in San Francisco, while Diane lives in Los Angeles. Diane owns the home she lives in as well as three other homes on the same block, which she rents out to various tenants. One of Diane's neighbors, Ned Riley, is a handyman who has helped her repair her properties and various other odd jobs throughout the past 10 years. Diane's mental functioning has slowly declined and she has started having troubling walking. Ned and his wife, Brenda, starting assisting Diane with grocery shopping, taking her to medical appointments, and obtaining her prescription medicine. After several months, Brenda takes over Diane's checkbook and starts paying her bills. Brenda writes all the checks by hand, and then has Diane signs the checks. After 10 months of assisting Diane, Ned and Brenda create a corporation and enter into a contract between Diane and the new corporation to "buy" Diane's three rental properties by paying a $1,000 down payment and agreeing to make payments of $1,000 per month. They set a total purchase price for all three properties at $300,000. Each rental property is worth over $250,000, making this sale arrangement a below-market bargain for Ned and Brenda. After the initial down payment, no other payments are made and all rental income goes directly to Ned and Brenda. After a year of this arrangement, Diane runs low on funds and cannot pay for her prescriptions, which is the first time Diane mentions her finances to her daughter. Tina hires a lawyer and discovers the "sale" arrangement as well as the neighbors' actions with Diane's checkbook.

In this scenario there are two possible claims to set aside the "sale" of the real properties to the sham corporation created by the wrongdoers, Ned and Brenda. The first claim is the lack of capacity to contract, which can be evaluated using the elements pertaining to competency to make a decision. While the above facts do not provide enough details to fully evaluate capacity, the mere fact that the properties were sold at a substantial discount suggests there were problems with Diane understanding the consequences of that decision.

In addition, these facts are typical of the pattern of behavior found in many undue influence claims. The neighbors started by being helpful and slowly took over more and more jobs for Diane. Such actions built trust and goodwill with Diane. Over time, the neighbors used that trusting relationship to convince Diane to enter into a bargain sale of the properties. The first step in evaluating an undue influence claim is to determine Diane's susceptibility to undue influence. Add to that the action and tactics of the neighbors, their position of trust with Diane and the inequity of the result, and the undue influence claim is further established. Capacity and undue influence are separate claims that can be evaluated at the same time, but one claim can succeed where the other fails, or they both can be established where the facts are sufficient to prove both.

Competency to Make a Will

Testamentary capacity has two components. First, the testator must understand what he or she is doing in making a will; the nature and extent of his or her property; and the "natural objects of his or her bounty"—that is, his or her relation to living descendants, spouse, and parents, and others whose interests are affected by the will. Most jurisdictions require evidence of a mental defect that has a direct correlation to the individual's ability to understand one of these three elements. A series of eccentric or odd behaviors will not, by itself, establish a lack of capacity under this standard. A mental defect, however, such as impaired cognitive functioning that directly affects an individual's ability to understand the natural objects of his bounty, would qualify as a lack of capacity under this standard.

Second, the person must not suffer from a mental disorder that involves delusions or hallucinations that have an effect on the particular bequest. In both of these, the mental health evaluator is typically called upon to make

judgments after the fact (some lawyers have evaluations at the time of the will to try to avoid later disputes about testamentary capacity).

The "insane delusion" standard is hard to meet: The delusion must be extreme and the individual must cling to the delusion in the face of uncontroverted contrary facts. For example, a belief that one's family no longer loves one would not be an insane delusion even if it is untrue and not well grounded in evidence: It *could* be true. By contrast, a belief that one's family is trying to kill one with ray guns would be an insane delusion. Or, a belief that one's child is the product of adultery, and not really one's child, will typically not be an insane delusion. A belief that the child has been taken over by aliens, and is not really one's child, would. The delusion must also affect the bequest. If the individual has another independent (sane) reason for the bequest, then the delusion does not vitiate his or her testamentary capacity.

Mr. Jones was an unmarried 75-year-old male with three children. One year prior to his death, he began exhibiting strange behavior. He would repeatedly visit his neighbors and offer to buy all of the furniture in their home even though the furniture was never for sale. Mr. Jones would also make crank calls to local pizza parlors. Over the course of 12 months he refused to let any family members or friends into his home, and would send gifts through the mail that typically consisted of a whole fish or raw meat. Mr. Jones was never diagnosed with any form of dementia or mental defect causing cognitive impairment. Upon his death, Mr. Jones' will was discovered. The will was created 6 months before Mr. Jones' death and he left his entire estate to his church. Did Mr. Jones have legal capacity to create his Will?

Mr. Jones likely had the requisite legal capacity to make his will. The law in this area presumes that decedents have capacity until proven otherwise. Here, there is no evidence of a mental defect. Even though Mr. Jones exhibited eccentric and strange behavior prior to his death, he lacked any mental defect that would directly affect his ability to understand the objects of his bounty, the fact that he was creating a will, and the general nature of his property.

Undue influence is commonly used in attempts to overturn a decedent's will. The elements (1) vulnerability of the victim, (2) apparent authority of the wrongdoer, (3) actions and tactics of the wrongdoer, and (4) equity of the result are the same as using undue influence to overturn a decision or contract.

It is usually the case that lessened capacity makes a subject more susceptible to undue influence. While the lessened capacity may not be enough to prove a lack of capacity, it could provide support of a successful undue influence claim provided the other elements of undue influence are present. In the example above, there are no facts to support the other elements of undue influence, but if a wrongdoer was present and exerting pressure on Bob Jones' decision to leave all his assets to his church, then there could be a claim for undue influence.

Roger is a wealthy business executive, age 83. Roger is married to Jenny, age 81 and they have two daughters, Wendy and Sally. For the past 20 years, Sally has worked in Roger's business where she oversees all of his finances. Roger was advised to start gifting some of his assets to

his children and grandchildren so he would not be subject to estate tax on his death. He instructed Sally to prepare checks of $10,000 each to Wendy, Wendy's two children, Sally, Sally's one child, and Sally's husband. Wendy became enraged when she learned that Sally's husband received a $10,000 check, but her husband did not. A few weeks after the checks were written, Roger died. Wendy then went to Jenny and told her that Sally is stealing money, cannot be trusted with the funds, and that Sally is trying to take Jenny's assets. Jenny becomes alarmed and immediately fires Sally from the business. She cuts off all communications with Sally. Wendy then moves Jenny into her home to care for her. Wendy is given a power of attorney to manage the finances, and Wendy begins overseeing the family business. Jenny then creates a new will, leaving all assets to Wendy and disinheriting Sally (all previous wills gave everything equally to the two daughters, Wendy and Sally). Sally attempts to talk to her mother, but Wendy refuses to let them speak and refuses Sally entry to her home to see Jenny. During this time Wendy continually tells Jenny that Sally is trying to be nice to her so that she can steal more money. Jenny dies and Wendy believes she is entitled to everything.

Was Jenny unduly influenced to create her will leaving everything to Wendy? The first question is whether Jenny was susceptible to undue influence. Here, there are signs that she may indeed be susceptible due to the death of her husband, her age, and her actions in so readily believing Wendy's wild claims. The first step in evaluating this claim would be to investigate Jenny's medical records to determine her level of mental functioning. But Jenny need only have a weakened mental state, not a complete lack of capacity, because the remaining elements focus on the actions of Wendy. Wendy had apparent authority over Jenny both as a trusted daughter and as an agent under Jenny's durable power of attorney. Wendy used suspicious tactics in keeping Jenny secluded and refusing to allow Sally to see Jenny. The resulting will was unlike any past will, where the estate was divided equally. This example also highlights the "poisoning of the mind" that occurs when an elder is secluded and manipulated by a single child. In this example, undue influence is present and the will can be set aside based on this claim.

Competency to Drive

Generally, geriatric psychiatry training does not provide specific instruction in how to complete a driving assessment. Most states require that a physician make a referral to the Department of Motor Vehicles (DMV) if dementia is diagnosed, but do not require the physician to state whether the patient can or can't drive. In fact, it may be preferable for the physician not to be seen by the patient as the decision maker regarding a loss of driving privileges. Loss of driving privileges is, after nursing home placement, the most feared of events for an elderly person. Age alone does not cause impairment behind the wheel; it is the mental and physical changes associated with aging that cause the problem. All drivers who suffer from a cognitive deficit should be screened for a high-risk classification by the DMV. Screening test administrators are looking for four mental requirements in the driver: Proper integration of sensory input, focus of attention, proper association of thought, and appropriate judgment. The presence of mild dementia is not necessarily a cause for a license revocation, but a clinical diagnosis of Alzheimer disease is usually grounds for revocation of driving privileges.

Living Will, Advanced Directive, and Durable Power of Attorney

The device of a living will has emerged as a way for people to control their

medical care in the event they become incapacitated. Typically, they authorize withdrawal of life-saving treatment (or failure to initiate it) in the event the person is terminally ill or in a persistent vegetative state. There are certain issues that are controversial, such as whether they can authorize termination of artificial hydration and nutrition. Another issue is whether living wills can allow termination of treatment in cases other than terminal illness or a persistent vegetative state (e.g., profound dementia). While living wills have been hailed in allowing people to effectuate the kind of care they desire should they become incapacitated, there is considerable evidence that they are not widely enacted or used.

Related to a living will is an advance directive that specifies kinds of treatment one wishes and doesn't wish for in the event of incapacity. For instance, in the mental health context, one might specify which medications one prefers or whether one prefers not to have medication at all. Also related is the durable power of attorney for health care. This is a person appointed by the elder to make decisions for him or her in the event of incapacity. Because it is the elder who designates this person, it is more likely that the person can effectuate the elder's needs and desires.

ADMINISTRATIVE ASPECTS OF FORENSIC GEROPSYCHIATRY

The Elderly Impaired Physician

Life expectancy is generally rising in Western societies, and upper income, highly educated persons with upper middle class lifestyles (like physicians) are likely to live and work even longer than the national average. Impediments to working in old age may include physical infirmities (a parkinsonian tremor may cut short a surgical career) or mental infirmities (mild cognitive impairment). If the affected individual is a physician, this is most likely to come to the attention of a medical board through a complaint made by an unhappy patient, usually after a bad clinical outcome.

The geriatric psychiatrist may be asked to evaluate the elderly physician in order to help the medical board reach a conclusion. The evaluating geropsychiatrist should avoid giving a conclusion ("the doctor should keep or not keep his or her medical license"), but should clarify the specific question that is to be addressed. The report of the doctor who can evaluate the various strengths and weaknesses in the elder's capacity to plan, organize, sequence, and abstract (the executive functions) and identify any impairments (in attention, memory, language, and social skills) that affect those executive functions will be more useful than the report that reaches a final conclusion (which should usually be a function of the board).

Elderly Applicants for Disability Benefits

Some provision, public or private, protects almost all working individuals against the risk of becoming disabled, physically or mentally. In the private

sector, there are many policies, and generally the benefits are related to the premiums. In the public sector, most citizens are covered by the Social Security Administration's policies, which are not designed to generate a profit, but have humane motivations (that no disabled persons be unable to provide for themselves if they lack the capacity to work). The Social Security Administration therefore processes many claims for alleged disability and the geriatric psychiatrist may be asked for a forensic opinion regarding the validity of such a claim.

The geriatric psychiatrist should resist the temptation to give an opinion as to the ultimate issue (i.e., the claimant can or cannot work), which is an administrative, not medical, judgment based on social security law. Rather, the doctor should give an opinion regarding the penultimate medical issues, which are forming an opinion as to the quality and quantity of impairments in four areas: (1) activities of daily living; (2) social functioning; (3) concentration, persistence, and pace; and (4) decompensations. In two of these areas (activities of daily living and concentration, persistence, and pace), there is a normal decline in performance with age, which makes the evaluation difficult. The Social Security Administration (SSA) will ask the psychiatrist to rate any impairments in terms that are poorly defined (mild, moderate, severe), and the difference between a claim that succeeds from one that fails may depend on whether the psychiatrist assesses the impairment as "moderate" or as "severe" in the absence of good definitional guidance around these terms. The SSA guidelines are sensitive to the demands of age; the same level of impairment may result in a failed claim if the applicant is young, but succeed if the applicant is elderly. If impairment is indeed present, the doctor should relate it to a causal factor, which could be a mental disorder (dementia) or a condition that is not a mental disorder (lack of English language skills). Based on this (medical) opinion, the SSA will then determine whether the applicant does or does not qualify for disability benefits.

CORRECTIONAL ASPECTS OF FORENSIC GEROPSYCHIATRY

Opportunities for the geriatric psychiatrist to provide clinical services within the correctional system abound and are likely to increase exponentially over the ensuing years. It is important to recognize that the correctional psychiatrist *treats* the incarcerated elder and hence plays a different role than the forensic psychiatrist, who *evaluates* the incarcerated elder to truthfully answer a medicolegal question. The last 10 years have been significant for a rising awareness of the impending problem of elderly prisoners in the criminal justice system. Virtually every major newspaper in the United States has run a story on how the "graying" of the incarcerated population is likely to present a serious ethical/moral challenge and/or economic crisis for the penal system. There are presently more than

125,000 prisoners over the age of 50 and 35,000 prisoners over the age of 65 in U.S. correctional facilities. Due to sentencing laws and demographic shifts in the population, the U.S. Census Bureau estimates that in 2030, one-third of the jail population will be over 55 years of age.

From a forensic perspective, the geriatric correctional psychiatrist may be asked whether an aging prisoner should go free. There are a number of considerations: 97 percent of elderly inmates are nonviolent offenders and less than 1 percent of those on probation or parole who are returned to jail or prison are older than 60. Furthermore, the poor health conditions of prisoners combined with harsh prison life make the prison environment particularly conducive to fostering disease. There is, however, no legal statute that is likely to provide elders with any specific rights beyond the general right to medical care and the right to be free from cruel and unusual punishment. A possible related question is whether age should be a basis for execution exemption. There are presently over 100 people age 60 or over on death row. The Supreme Court has ruled that it is unconstitutional to execute certain special populations (juveniles and the mentally retarded) due to violation of the Eighth Amendment's prohibition against cruel and unusual punishment. In the case of the elderly prisoner, however, most are presumably fully functioning at the time of the crime. The strongest argument might be if the elder suffers from dementia or some other age-related mental deficiency that impairs capacity to understand why he or she was being executed at the time the sentence was carried out.

NOSOLOGICAL/THEORETICAL ISSUES

In routine clinical practice, most geropsychiatrists use the DSM-V and follow the "spirit" of the text. Good clinical practice (they would argue) has to do with using the diagnostic criteria in a manner that promotes communication and increases patient well-being. Consider the case of an elderly patient recently retired, now bored and unhappy. Since he may respond to antidepressant medication, many psychiatrists would label such a case as one of depression (a mental disorder) rather than "occupational problem" or "phase-of-life problem," both of which are not mental disorders but rather mental conditions (and hence coded with a "V" code). The fact that each one of the specific DSM-5 criteria may not be exactly met is less relevant than the fact that following the spirit of the text, rather than the letter, creates more good than harm.

In clinical practice, the doctor has a financial incentive to diagnose in a manner that promotes reimbursement; usually, the diagnosis and treatment of a mental disorder does result in monetary reimbursement, while the diagnosis and management of a mental condition may not. In forensic cases, the doctor is more likely to be rewarded for telling the truth (or at least whichever "truth" the payer is seeking); in such cases, the

doctor is more likely to consider the mental conditions (as well as the mental disorders) in the differential diagnosis.

In court testimony, the doctor swears to tell the whole truth, and if the diagnoses do not follow the letter of the text, the doctor should truthfully acknowledge that. But the truth may be difficult to determine—forensic geropsychiatrists are often asked to clarify boundaries that may be fuzzy, including the boundaries between (1) cognitive disorder/age-related cognitive decline, (2) antisocial personality disorder/adult antisocial behavior, and (3) depressive disorder/bereavement. The former are all mental disorders, while the latter are not. The boundary between the two is often vague and ill-defined, and there may be temptations to place the boundary in a manner that promotes values that are ethical (compassion) or policy based (access to care), rather than scientific (nosology). While it is admirable for the clinical geropsychiatrist to be compassionate or to advocate for greater access to health care for untreated patients, the forensic geropsychiatrist subscribes to a code of ethics that emphasizes truth telling, rather than the traditional ethical values of promoting health. In an ideal world, there will be no clash between such competing values. But since there is often an adversarial quality to forensic practice, the forensic geropsychiatrist may have to make difficult choices, between acting on values that promote health (the traditional values of clinical ethics) or on those that promote the truth (the values of the forensic expert). If there is a choice between promoting health and promoting patient autonomy, the ethical doctor may be troubled and unsure about the resolution of such a dilemma. With the growth in both the population of the elderly and the number of regulations that govern the care of the elderly, the forensic geropsychiatrist is likely to see more such dilemmas.

Future research in forensic geriatric psychiatry may therefore be of two types: (1) quantitative (data-based) research—especially in the sciences pertaining to nosology (to clarify the proper boundaries of the various mental disorders that may be a basis for forensic opinions) and decision science (to better delineate the pathways to good and bad judgment [cognitive, social, moral, logical] that often form the basis of findings of mental incompetence in the elderly) and (2) qualitative (conceptual) research—integrating findings from the biomedical aspects of geropsychiatry (i.e., whether a brain disease is or is not curable), with the implications of such findings (legal protections for the rights of the elderly mentally ill and ethical concerns regarding allocation of limited health care resources to a population of the elderly that may have unlimited growth).

References

Aday RH. *Aging Prisoners: Crisis in American Corrections*. Westport, CT: Praeger; 2003.

Anfang SA, Appelbaum PS. Forensic evaluation of the older patient. In: Agronin ME, Maletta GJ, eds. *Principles and Practice of Geriatric Psychiatry*. Philadelphia, PA: Lippincott Williams &

Wilkins; 2006.

Bryan AG, ed. *Blacks Law Dictionary*. 9th ed. Eagan, MN: West Publishing Company; 2009.

California Welf. & Inst. Code section 15610.70 (effective January 1, 2014).

Clark RV. 28 Cal.4th. 2002;89:96.

Dukoff R, Sunderland T. Durable power of attorney and informed consent with Alzheimer's disease patients. *Am J Psychiatry*. 1997;154:1070–1075.

Dunn LB, Jeste DV. Enhancing informed consent for research and treatment. *Neuropsychopharmacology*. 2001;24:595–607.

Dunn LB, Lindamer LA, Palmer BW, Golshan S, Schneiderman LJ. Improving understanding of research consent in middle-aged and elderly patients with psychotic disorders. *Am J Geriatr Psychiatry*. 2002;10:142–150.

Grisso T. *Evaluating Competencies: Forensic Assessments and Instruments*. 2nd ed. New York: Plenum Publishing; 2003.

*Grisso T, Appelbaum PS. Comparison of standards for assessing patients' capacities to make treatment decisions. *Am J Psychiatry*. 1995;152:1033–1037.

Grossberg G. Advanced directives, competency evaluation and surrogate management in elderly patients. *Am J Geriatr Psychiatry*. 1998;6:S79–S84.

Haroun A. Psychiatric aspects of wickedness: Introduction. *Psychiatr Ann*. 1997;27:615–616.

Horwitz HL. California's new statutory definition of undue influence—modernization or game changer? *California Trusts Estates Quarterly*. 2013;19(3):1–8.

*Jeste DV, Palmer BW, Appelbaum PS, Golshan S, Glorioso D. A new brief instrument for assessing decisional capacity for clinical research. *Arch Gen Psychiatry*. 2007;64:966–974.

Miller MK, Bornstein BH. Elders on death row. *APA Online Monitor Psychol*. 2007;38(4):60.

Moye J. Guardianship and conservatorship. In: Grisso T, ed. *Evaluating Competencies*. New York: Plenum Publishing; 2003.

*Naimark D. Financial exploitation of the elderly: The evaluation of mental capacity and undue influence. *Am J Forens Psychiatry*. 2001;22(3):5–19.

Naimark D, Dunn L, Haroun A, Morris G. Informed consent and competency-legal and ethical issues. In: Jeste DV, Friedman JH, eds. *Current Clinical Neurology: Psychiatry for Neurologists*. Totowa, NJ: Human Press Inc.; 2006.

Perr IN. Wills, testamentary capacity, and undue influence. *Bull Am Acad Psychiatry Law*. 1981;9(1):15–22.

Perry SJ. Legal implications for failure to comply with advanced directives. *Behav Sci Law*. 2002;20:253–269.

Rinkert J, Naimark D. The elderly driver. *Psychiatr Times*. 2005;12:7.

Saks ER, Dunn LB, Wimer J, Gonzales M, Kim S. Proxy consent to research: The legal landscape. *Yale J Health Care Policy Law Ethics*. 2008;8:37–92.

Singer MT. Undue influence and written documents: Psychological aspects. *J Questioned Document Exam*. 1992;1:4.

Social Security Administration. *Office of Disability Programs. Disability Evaluation Under Social Security. SSA Publication No. 64-039, ICN 468600*. Washington, DC: Social Security Administration; 2005.

*Spar JE, Hankin MB, Stodden AB. Assessing mental capacity and susceptibility to undue influence. *Behav Sci Law*. 1995;13:391–406.

Superior Court of California. County of San Francisco, Undue Influence: Definitions And Applications, A Project Supported by the Borchard Foundation Center on Law and Aging, Final Report (March 2010), Mary Joy Quinn, Project Director. Available at http://www.courts.ca.gov/documents/UndueInfluence.pdf. Accessed February 13, 2014.

Thoreen VL, David KG. What every trusts and estates practitioner needs to know about elder financial abuse. *California Trusts Estates Quarterly*. 2014;20:1.

The Trusts & Estates Section of the State Bar of California. Preserving Senior Independence, An Outline for Speaker Presentations (June 6, 2013). p. 2.

U.S. Gov't Accountability Office, Elder Justice. Stronger Federal Leadership Could Enhance National Response to Elder Abuse, GAO-11-208 (March 2011) p. 48 (Citing MetLife Mature Market Institute et al., Broken Trust: Elders, Family, and Finances: A Study of Elder Financial Abuse Prevention (2009) notes 1 & 4).

U.S. Gov't Accountability Office, Elder Justice. National Strategy Needed to Effectively Combat Elder Financial Exploitation, GAO-13–110 (November 2012).

57.6c Ethical Issues

NATHAN FAIRMAN, M.D., SCOTT A. IRWIN, M.D., PH.D., AND BARTON W. PALMER, PH.D.

Geriatric mental health care is often fraught with complex issues of bioethics. Most of these challenges have parallels in the ethical care of nonelderly mental health patients, as well as in the care of medical patients of all ages. The issues of general psychiatric ethics are considered elsewhere in this volume. The intent of this section is to review the core bioethical principles and consider their application in the context of geriatric mental health care. Although mere identification of core principles does not automatically provide an unambiguous solution for a particular ethical dilemma, such clarification does foster wider recognition of subtle, if common, ethical issues and helps to highlight the areas that need consideration when developing a plan for resolution or management of such dilemmas.

Following a review of the historical roots of bioethics and a presentation of the core principles, several general applications of these principles to patient care and research in geriatric mental health will be reviewed. Finally, a more detailed consideration of one specific component of bioethics—particularly important in geriatric mental health—will be provided, that is, informed consent and decision-making capacity. The rationale for this focus is that many of the commonly encountered ethical challenges in geriatric mental health involve questions of decisional capacity, particularly in cases marked by deficits in cognition or affective states which may be felt to undermine capacity. Moreover, older adults and others with chronic, life-limiting illness will increasingly encounter

decisions that may influence the timing of death—a possibility which, to some, may seem to "raise the stakes" surrounding questions of decisional capacity. Since deficits in cognition and affective conditions such as depression, grief, and demoralization are common among older adults, careful attention to decisional capacity and informed consent can help to ensure sound clinical and ethical decision making at the bedside. In addition, informed consent and decisional capacity constitute an area of bioethics that has received particularly rich empirical attention over the past decade and thereby serve as a model for how empirical data may complement moral, philosophic, legal, and public policy perspectives on bioethical issues in patient care.

HISTORY

Consideration of moral and ethical principles in health care has ancient beginnings, including the Hippocratic Oath from the fifth century BCE. In terms of development of contemporary ethical codes, a key historical marker was reached in 1803, with the publication of Thomas Percival's *Medical Ethics*. Although his was not the first ethics code developed, it is widely credited as providing the foundation for many subsequent formal ethics codes in Western medicine, including that adopted with the founding of the American Medical Association in 1847.

The 1960s and 1970s witnessed the emergence of bioethics as a term and formal discipline of inquiry. During this period, several key bioethics centers were founded, including the Hastings Center in 1970 and the Kennedy Institute of Ethics in 1971. The rising interest in clinical ethics during this period appears to have been partially due to advances in medical care that brought new challenges, such as the introduction of increasingly technical (and costly) medical procedures (and therefore questions about how to fairly allocate finite medical resources), evolving public attitudes about care near the very end of life, and advances in genetics and various aspects of reproductive medicine. In reference to geriatrics, it may also be noted that the 1960s to 1970s was precisely the period when recognition and interest in Alzheimer disease and other neuropsychiatric conditions associated with aging began to gain wider public, professional, governmental, and empirical attention. The increased attention to such conditions, as well as advances in medical care for many chronic but life-limiting conditions, also brought into focus a modern dilemma: balancing quality of life with mere extension of life.

The growing attention to the ethics of clinical care in the later half of the 20th century was also catalyzed by parallel growth in the related field of medical research ethics. In particular, the Nuremberg Code, published in 1948 in response to the Nuremberg trial in the aftermath of World War II, and the Declaration of Helsinki, first published in 1964, represented major multinational collaborative efforts to specify the limits and requirements

for ethical medical research, such as the strong emphasis on the need for voluntary informed consent. Public, professional, and government attention to health care ethics was also fostered by Henry Beecher's now classic 1966 account of unethical research practices, exemplified in 22 contemporaneous studies that had been published in leading peer-reviewed medical journals, and by the public revelations in the early 1970s of the now-infamous Tuskegee and Willowbrook experiments.

Two key bioethics documents were developed in the late 1970s: (1) the *Principles of Biomedical Ethics*, by Tom L. Beauchamp and James F. Childress and (2) *The Belmont Report* from the U.S. National Commission for the Protection of Human Subjects. Beauchamp and Childress described bioethics in terms of four core principles: respect for autonomy, nonmaleficence, beneficence, and justice. Nonmaleficence and beneficence were fused together in *The Belmont Report*, which identified the fundamental principles of ethical medical research as respect for persons, beneficence, and justice.

CORE PRINCIPLES

Whether in geriatric psychiatry, general psychiatry, or general medicine, the core principles of bioethics—described in further detail below—do not, in themselves, explicitly specify the full range of ethical responsibilities clinicians or researchers have to their patients and research participants. Obligations to patients may be rooted in other principles, or may entail other ethically relevant considerations. Nonetheless, most of the cardinal bioethical concerns can be understood as derivatives of these core principles.

Respect for Persons/Autonomy

As defined in *The Belmont Report*, respect for persons includes two potentially conflicting components: (1) respect for the rights of individuals to autonomous decision making, and (2) recognition of the need to protect those with impaired capacity for autonomous decision making. The principle of respect for persons is widely represented in the contemporary ethics codes of various health and mental health professions. For instance, the second decree in the World Medical Association's *International Code of Medical Ethics* reads "A physician shall respect a competent patient's right to accept or refuse treatment." Another example is Principle I, Respect for the Dignity of Persons, from the *Canadian Code of Ethics for Psychologists,* which includes explicit recognition of the need to protect vulnerable individuals as well as respect their wishes and human dignity. Indeed, as shall be seen, many of the prototypic ethical conflicts in geriatric medicine and mental health care can be understood as arising, at least in part, from tensions between these two related (but often competing)

obligations: the duty to respect autonomy and dignity and the duty to protect those with impaired capacity for autonomous decision making.

Beneficence and Nonmaleficence

In *The Belmont Report,* beneficence is conceptualized as encompassing two complementary ethical mandates: "(1) do not harm and (2) maximize possible benefits and minimize possible harms." Some authors, including Beauchamp and Childress, separate these two mandates, identifying the former as the principle of nonmaleficence (construed as an obligation to minimize harm), and reserving the term beneficence to connote an obligation to the promotion of good. There is little disagreement that both components are germane to patient care and clinical research. For instance, in describing the concept of beneficence, Beauchamp and Childress note: "Morality requires not only that we treat persons autonomously and not harm them but also that we contribute to their health and welfare." Notions of beneficence and nonmaleficence are also widely recognized in contemporary codes of bioethics. The World Medical Association's *International Code of Medical Ethics,* for example, dictates that "A PHYSICIAN SHALL act in the patient's best interest when providing medical care." The current ethics code from the American Medical Association includes the statement that "A physician shall, while caring for a patient, regard responsibility to the patient as paramount." Similarly, the Canadian Medical Association's code of ethics lists the first fundamental responsibility of the physician to "Consider first the well-being of the patient." A final example is seen in the current ethics code from the American Psychological Association, which includes beneficence and nonmaleficence together as Principle A, noting: "Psychologists strive to benefit those with whom they work and take care to do no harm."

Justice

From an ethical perspective, the term justice is generally interpreted in reference to "fairness" or "fairness in distribution." Issues of justice are often most readily identified at the systems level, for example, in reference to concerns about resource allocation, access to care, or disparities in health. Nonetheless, justice also pertains to other important domains of medical care: *The Belmont Report,* for example, emphasizes the fairness of benefits and burdens from clinical research. So too, issues of justice often arise in clinical care in geriatric medical and mental health settings. Many of the professional codes of bioethics include at least aspirational goals that clinicians will work toward the betterment of society. For instance, two of the nine principles in the current ethics code from the American Medical Association clearly reflect justice considerations:

VII A physician shall recognize a responsibility to participate in activities contributing to the

improvement of the community and the betterment of public health . . .

IX A physician shall support access to medical care for all people.

Another of many possible examples is Principle D from the ethics code of the American Psychological Association, which reads:

Psychologists recognize that fairness and justice entitle all persons to access to and benefit from the contributions of psychology and to equal quality in the processes, procedures, and services being conducted by psychologists. Psychologists exercise reasonable judgment and take precautions to ensure that their potential biases, the boundaries of their competence, and the limitations of their expertise do not lead to or condone unjust practices.

APPLICATION OF CORE PRINCIPLES OF BIOETHICS TO GERIATRIC MENTAL HEALTH CARE

Several common general applications of the principles of bioethics to geriatric mental health care warrant mention "respect for dignity" and "respect for persons." For example, respect for dignity, an important component of the broader principle respect for persons, is often a paramount consideration in geriatric mental health. As a result of the stigma associated with neuropsychiatric disorders, as well as the loss of function and status sometimes associated with neurodegenerative conditions, patients and family often view health-related decisions in this context through the prism of concerns about dignity. Similarly, the right to confidentiality (closely related to issues of dignity) may also be viewed as an extension of respect for persons. In the context of work with decisionally impaired individuals, confidentiality issues can be particularly complex, as there is often a need to work closely with adult children or other family members, sometimes disclosing information at a level that would be uncommon (and perhaps inappropriate) with a patient whose decisional capacity was intact.

Another key general application of bioethics to geriatric medicine and mental health care involves the issue of professional competency, particularly the skills and knowledge necessary for delivering appropriate treatment to older patients. The need to practice within the scope of one's training and competence is rooted in the principle of beneficence/nonmaleficence, and it is explicitly emphasized in most professional bioethics codes. For instance, in describing the duties of physicians to patients, the World Medical Association's *International Code of Medical Ethics* includes the requirement that "Whenever an examination or treatment is beyond the physician's capacity, he/she should consult with or refer to another physician who has the necessary ability." The ethics code from the Canadian Medical Association enjoins physicians to "Recognize your limitations and, when indicated, recommend or seek additional opinions and services." The issue is stated even more forcefully in the American Psychiatric Association's annotations to the American Medical Association's *Principles of Medical Ethics*: "A psychiatrist who regularly practices outside his or her area of professional competence should be considered unethical."

Two possible responses to the prescription to practice within the limits of one's training and competence include (1) limiting one's practice, or (2) expanding one's training and competence. Given the major growth in the numbers and proportion of older adults with neuropsychiatric conditions, considerations of justice (as well as beneficence and nonmaleficence) should place increasing pressure on general psychiatrists and physicians to pursue the second option, that is, obtain sufficient training in geriatric mental health to at least enable the individual clinician to appropriately discern when consultation or referral to a geriatric mental health expert is warranted. This may be most clearly illustrated in reference to late-life depression. Although rates of clinically significant depressive symptoms are relatively high among older adults, depression in this population tends to be underrecognized (and therefore undertreated). What is more, clinical competence with respect to managing late-life depression includes the knowledge that, even near the very end of life, depression can often be effectively treated. A variety of medications and psychotherapies have been shown to be beneficial in treating depression in palliative care populations, and for this reason, if depression is suspected, the "competent" physician should at least recommend a time-limited therapeutic trial, simply from the standpoint of the principle of beneficence.

Justice considerations also have ethical relevance to the development and testing of psychotropic medications. Although the situation has been improving in the past decade, elderly patients have been traditionally underrepresented in clinical trials of antidepressant and other psychotropic medications. Given potential age-related changes in pharmacokinetics and pharmacodynamics, as well as increased likelihood of drug interactions from polypharmacy, the need for clinicians to extrapolate from data collected from samples of younger patients represents an ethical challenge in terms of justice (as well as nonmaleficence). The latter example also reflects that ethical dilemmas are not always best conceptualized or addressed at the level of individual clinician and patient interactions, but may also occur and require change at the level of public policy—in this case, with respect to setting national (or international) research priorities.

DECISION-MAKING CAPACITY AND INFORMED CONSENT

As the above discussion is intended to illustrate, geriatric mental health care and research activities frequently give rise to a variety of bioethical concerns. In fact, as noted, many of the ethical challenges encountered in the context of geriatric mental health can be understood as arising from conflicts among the core principles of bioethics. Among these, the potential conflict between the two components of the principle of respect for persons (respecting individuals as autonomous agents and protecting those with diminished capacity for autonomous decision making) is one of the most salient and challenging in geriatric mental health care.

Broadly speaking, such conflicts can be particularly difficult to comfortably resolve when the weighing of considerations from competing principles is highly subjective, that is, in those instances wherein "reasonable persons" may disagree and come to divergent conclusions regarding the most appropriate course of action. Consider the controversies around the "right to die" near the end-of-life, often exemplified in discussions about withdrawing or withholding treatment, palliative sedation, or physician aid-in-dying. One can readily identify tensions among the core bioethical principles in such issues, wherein reasonable people might disagree about how to best respect autonomy, balance obligations to beneficence and nonmaleficence, and even address concerns about justice. Some of these tensions are brought to light in the following case:

Sylvia is a 68-year-old woman with advanced chronic obstructive pulmonary disease, a history of breast cancer, and major depression. Though she lives alone, over the past year she has found it increasingly difficult to maintain her independence: she has been in and out of the hospital several times, and after the most recent hospitalization she was discharged to a nursing facility for rehabilitation. She found the brief stay in the nursing facility to be demoralizing—not so much because of the quality of care she received, but rather because the idea of being dependent on others for her most basic needs ran contrary to her own sense of self-worth. She counts her ex-husband, her daughter, and a former colleague among her closest supports, and she has vowed to them that she would "rather die" than to experience "the indignity of wasting away" in a nursing home.

In a subsequent hospitalization—due to yet another bout of pneumonia—she appears more depressed than usual, and there are some new concerns about delirium. She makes an oblique reference to her doctor suggesting that she wishes her life were over: "I'm tired of going through this; when will it end? There must be something you can do to help me . . ."

From the standpoint of bioethics, how should Sylvia's doctor respond? What course of action will strike the right balance among the different—and seemingly opposed—obligations? For instance, how should her doctor negotiate the two components of respect for persons (honoring Sylvia's right to autonomous decisions, while protecting her in the event that her capacity for authentic autonomous decision making is diminished, due to what may be depression or cognitive impairment)? Further, apart from concerns about autonomy, how should the clinician understand "the good" that should be promoted, or "the harm" that should be avoided (so as to respect the dictates of beneficence/nonmaleficence) in a case where a suffering patient, near the end of life, desires that their death come more quickly?

Not surprisingly, in medical ward settings, psychiatrists are frequently asked to provide consultation in cases like Sylvia's—to evaluate the degree to which a patient's refusal of potentially life-extending treatment (or more generally, their desire for hastened death) may reflect compromised capacity for autonomous decision making due to depression, cognitive impairment, or even psychosis. From the perspective of the geriatric mental health clinician, the question of decisional capacity—a component of informed consent, relating fundamentally to the principle of respect for persons—is likely to be instrumental in helping to sort out some of these issues.

Three Components of Informed Consent

Concerns about decision-making capacity relate fundamentally to the doctrine of informed consent, which (in most formulations) involves three interrelated components: first, valid informed consent requires that the

individual's decision be made in the absence of coercion (i.e., the decision must be truly voluntary). Second, informed consent also requires disclosure of all the information relevant to the decision at hand, such as the purpose and nature of the proposed treatment or procedure, viable alternatives, and the foreseeable risks and benefits, as well as potential lack of benefits, from each possible course of action (including doing nothing). Such disclosure is generally documented with a printed consent form that the patient or his or her surrogate signs; however, as cogently noted in a report from the Institute of Medicine:

Conceptually, informed consent should be construed as an evolving decision-making process, rather than as simple "permission." This model of "ongoing conversation and decision making" allows for reinforcement of previously disclosed information, introduction of new information, and respect for autonomous participant decision making. The act of obtaining *written* consent is tangential to the *process* of informed consent and merely provides a mechanism to document and record that ethically relevant communication has taken place.

The third key component of valid informed consent is that the individual has the ability to use the information to make an authentic, informed decision; that is, the individual has "decision-making capacity."

Decision-Making Capacity

Decision-making capacity derives from a set of mental abilities, which together enable a patient to engage in the process of informed consent. Having decisional capacity is a necessary—but not sufficient—element of informed consent, a process which, as noted above, requires the presence of several other important components (absence of coercion, disclosure of relevant information, etc.). Most often, decision-making capacity is distilled into four main elements:

1. Understanding—the ability to comprehend information relevant to the decision at hand;
2. Appreciation—the ability to apply that information to one's own situation and condition;
3. Reasoning—the ability to manipulate and consider pertinent information, so as to rationally assess the different options available; and
4. Communication—the ability to express a stable choice over time.

Note that many experts draw a distinction between "capacity" (a determination made by a physician) and "competency" (a determination made by a judge), but the two terms are often, confusingly interchanged. In general, questions of capacity should relate to specific interventions, at a specific point in time (e.g., a trial of dialysis, or surgical repair of an aneurysm, etc.), rather than a more generalized ability, since the capacity required to engage in a simple, straightforward decision may be

substantially different from the capacity required to engage in a highly complex or risky decision.

In most states, decisional capacity can be determined by any licensed physician (and sometimes also by a small set of nonphysician mental health specialists, such as clinical psychologists). Hence, most consultation requests for "capacity evaluation" can be framed as providing a second opinion, when the primary clinician's assessment is indeterminate or uncertain. In practice, psychiatrists are frequently consulted when serious affective conditions, problems with cognition, or deficits in reality testing are felt to interfere with decision-making capacity.

Capacity determinations are ultimately based on clinical assessment, which is often complemented by an objective measurement of several domains of cognition. Unfortunately, empirical studies suggest poor reliability, even among experienced clinicians, in terms of categorical determinations of an individual patient's capacity to consent to treatment based on standard clinical interviews. However, interrater reliability may improve when the clinicians providing such ratings are given explicit operational guidelines. Several tools have been developed to assist with the assessment of decisional capacity, such as the MacArthur Competence Assessment Tool for Treatment (MacCAT-T), the Capacity to Consent to Treatment Instrument (CCTI), and the Aid to Capacity Evaluation (ACE). The advantages and disadvantages of each instrument (and others) should be considered in reference to the specific purpose of the capacity assessment for an individual patient or research subject. Finally, while published reviews and educational resources often leave the impression that capacity determinations are merely a matter of applying certain criteria in the context of a clinical examination, such assessments in complicated cases are rarely straightforward, and they often involve an uncomfortable degree of subjectivity and judgment.

Impact of Demographics and Disease on Decision-Making Capacity: Is Mental Illness a Proxy for Incapacity?

In 1998, the U.S. National Bioethics Advisory Commission (NBAC) issued a report titled *Research Involving Persons with Mental Disorders That May Affect Decision Making Capacity*. The purpose of the report was to consider whether persons with mental illness require additional protections when enrolled in clinical research, due to the possibility that mental illness may adversely affect their capacity to provide informed consent. Although the intent to protect potentially vulnerable individuals was widely lauded, critics of the report cited concerns that it may have inappropriately singled out persons with mental illness as being uniquely vulnerable to impaired decisional capacity, and thereby, inadvertently risked amplifying the social stigma associated with mental illness. There were further concerns that some of the specific recommendations could

lead to situations in which mental illness was used as a proxy for impaired decisional capacity, thereby making research with certain patient populations pragmatically difficult and hindering the pursuit of research that would be critically important for those very populations.

One clear benefit from the NBAC report was that it catalyzed empirical studies of capacity to consent to treatment or research among persons with schizophrenia, Alzheimer disease, and other neuropsychiatric or medical conditions. The results of that research confirm that some neuropsychiatric conditions are associated with increased risk of impaired decisional capacity, but also consistently reveal substantial heterogeneity within each of the diagnostic groups (as well as within general medical patient groups). This research has also consistently shown, across and within diagnostic groups, that the strongest predictor of impaired decisional capacity is the severity of cognitive deficits, rather than the severity of psychopathology. Indeed, at least among outpatients, positive symptoms in schizophrenia (such as hallucinations or delusions) have little discernable relationship to decision-making capacity. The correlation of age and impaired decisional capacity tends to be modest at best and is virtually zero among persons with schizophrenia. As capacity relates to particular qualities of reasoning in the context of particular interventions, the presence of a psychiatric condition (major depression, Alzheimer dementia, schizophrenia, etc.) cannot be used as a simple substitute in assessing a person's decisional capacity. In short, whether dealing with medical or psychiatric patients and whether dealing with older or younger patients, decisional capacity must be assessed on a case-by-case basis, taking into account both patient factors (cognition, affect, etc.) and factors unique to the decision at hand (complexity, risks, uncertainty, etc.).

Respect for Persons in Persons with Impaired Decision Making

Where does all this leave the geriatric mental health specialist, when faced with a person like Sylvia, who, in the advanced stage of a progressive and terminal illness, experiencing some form of depression as well as possible cognitive deficits, begins to give voice to a desire to die? If Sylvia has decisional capacity, she might, for example, choose to decline antibiotic therapy for her pneumonia and to be discharged home with hospice care, with an expectation that she would soon die. If she lacks capacity, however, should antibiotics be given without her consent? Should she be "forced," against her objections, to return to a nursing facility? From the standpoint of bioethics, uncertainty about these questions reflects tensions arising from the dual obligation to respect the autonomy of persons while protecting people with impaired capacity for autonomous decision making. Hence, in very practical ways, what is done for (and to) Sylvia will hinge on the assessment of decisional capacity.

Often, a careful and systematic process—involving thoughtful

clarification of Sylvia's decision-making capacity, but also comprehensive assessment of other sources of suffering, close coordination with family, realistic assessment of prognosis and the effectiveness of treatment, and therapeutic interventions that are structured in time-limited trials with clear outcomes—can help to achieve an acceptable solution, even in some of the most complicated cases. In Sylvia's case, such a process might include the following:

1. Empathic clarification of the patient's request: what, precisely, is Sylvia referring to when she says "I'm tired of going through this"? As Susan Block and Andrew Billings have observed: "Most patients, in saying that they want to die, are asking for assistance in living—for help in dealing with depression, anxiety about the future, grief, lack of control, dependence, physical suffering, and spiritual despair."
2. Attention to potential sources of coercion (e.g., from financial pressures, or burdens of caregiving).
3. Careful assessment of decisional capacity. When severe depression influences capacity, it often does so through distorted judgments, leading patients to undervalue the likelihood of a meaningful future, to underestimate the possibility that depression could be relieved, or even to be unaware of the presence of depression itself.
4. A thorough exploration of Sylvia's long-held values; how she had made decisions in the past; confirmation from a surrogate that her current choices remain "authentic" in the sense that they are in alignment with her lifelong preferences.
5. Identification and treatment (if consistent with the goals of care) of any potentially reversible causes of delirium.
6. Thorough, expert characterization and appropriate treatment of "depression." Is this another major depressive episode? Grief? Demoralization? An adjustment disorder? etc.
7. Comprehensive review of other forms of distress: physical, emotional, interpersonal, spiritual, practical, etc. Reliance on an interdisciplinary team to respond to the patient's suffering and deliver the plan of care.
8. Interventions structured in time-limited trials with clear outcomes.
9. Ongoing, open communication with the patient, her surrogates, and other relevant stakeholders.

Although treatment "against one's will" is sometimes supported by the obligation to protect vulnerable individuals with impaired decisional capacity, clinicians should also be open to the possibility that in many cases, one can simultaneously respect the person's autonomy and dignity, while also protecting him or her as a person vulnerable to impaired decisional capacity. In particular, efforts to explain and obtain the person's assent, wherever possible, will foster dignity and respect, even when consent must be obtained from the patient's surrogate. As Kim and

colleagues have noted in the context of research consent, some patients may lack the capacity to make the decision on their own, but retain sufficient capacity to identify a trusted proxy decision maker.

References

*American Bar Association, American Psychological Association, Assessment of Capacity in Older Adults Project Working Group. Assessment of older adults with diminished capacity: a handbook for psychologists. 2008. Available at http://www.apa.org/pi/aging/. Accessed March 3, 2009.

American Psychological Association. Ethical principles of psychologists and code of conduct. *Am Psychol*. 2002;57:1060–1073.

American Psychiatric Association. The principles of medical ethics with annotations especially applicable to psychiatry. 2009. Available at https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/Ethics/principles-medical-ethics.pdf

*Beauchamp T, Childress J. *Principles of Biomedical Ethics*. New York: Oxford University Press; 1979.

Beecher HK. Ethics and clinical research. *N Engl J Med*. 1966;274(24):1354–1360.

Block SD. Psychological issues in end-of-life care. *J Palliat Med*. 2006;9(3):751–772.

Block SD, Billings JA. Patient requests for euthanasia and assisted suicide in terminal illness. The role of the psychiatrist. *Psychosomatics*. 1995;36:445–457.

Canadian Psychological Association. *Canadian Code of Ethics for Psychologists*. 3rd ed. Ottawa, Ontario: Author; 2000.

Canadian Medical Association. *CMA Code of Ethics*. 2004. Available at http://policybase.cma.ca/dbtw-wpd/PolicyPDF/PD04-06.pdf

Community tools: Aid to Capacity Evaluation (ACE). University of Toronto Joint Centre for Bioethics. Available at http://www.jcb.utoronto.ca/tools/documents/ace.pdf

*DuBois JM. *Ethics in Mental Health Research: Principles, Guidance, and Cases*. New York: Oxford University Press; 2008.

*Dunn LB, Jeste DV. Enhancing informed consent for research and treatment. *Neuropsychopharmacology*. 2001;24:595–607.

Federman DD, Hanna KE, Rodriguez LL. *Responsible Research: A Systems Approach to Protecting Research Participants*. Washington, DC: Institute of Medicine, National Academies Press; 2003.

Giron MS, Fastbom J, Winblad B. Clinical trials of potential antidepressants: to what extent are the elderly represented: a review. *Int J Geriatr Psychiatry*. 2005;20:201–217.

*Grisso T, Appelbaum PS. *Assessing Competence to Consent to Treatment: A Guide for Physicians and Other Health Professionals*. New York: Oxford University Press; 1998.

Grisso T, Appelbaum PS. *MacArthur Competence Assessment Tool for Treatment (MacCAT-T)*. Sarasota, FL: Professional Resource Press; 1998.

Kim SY, Karlawish JH, Kim HM, Wall IF, Bozoki AC, Appelbaum PS. Preservation of the capacity to appoint a proxy decision maker: implications for dementia research. *Arch Gen Psychiatry*. 2011;68(2):214–220.

Marson DC, Earnst KS, Jamil F, Bartolucci A, Harrell LE. Consistency of physicians' legal standard and personal judgments of competency in patients with Alzheimer's disease. *J Am Geriatr Soc*. 2000;48:911–918.

Marson DC, Ingram KK, Cody HA, Harrell LE. Assessing the competency of patients with Alzheimer's disease under different legal standards. A prototype instrument. *Arch Neurol*. 1995;52:949–954.

National Bioethics Advisory Commission. *Research Involving Persons With Mental Disorders That May Affect Decisionmaking Capacity*. Rockville, MD: National Bioethics Advisory Commission; 1998.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. 1979. Available at https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html

Palmer BW, Savla GN. The association of specific neuropsychological deficits with capacity to consent to research or treatment. *J Int Neuropsychol Soc*. 2007;13:1047–1059.

Rothman DJ. Were Tuskegee and Willowbrook 'studies in nature'? *Hastings Cent Rep*. 1982;12:5–7.

Sessums LL, Zembrzuska H, Jackson JL. Does this patient have medical decision-making capacity? *JAMA*. 2011;306(4):420–427.

Shore D, Hyman SE. An NIMH commentary on the NBAC report. *Biol Psychiatry*. 1999;46:1013–1016.

Sullivan MD, Youngner SJ. Depression, competence, and the right to refuse lifesaving medical treatment. *Am J Psychiatry*. 1994;151:971–978.

*World Medical Association. *International Code of Medical Ethics*. 2006. Available at http://www.wma.net/en/30publications/10policies/c8/

57.6d Gender Issues

HELEN H. KYOMEN, M.D., AND MA'AYAN E. LIEBERMAN, M.D.

INTRODUCTION

Gender and age are important factors in the daily practice and science of medicine and are key components in establishing diagnosis, evaluating risks, and determining treatment including the effective and safe dosages of medications. Due to differences in life expectancy, there are more older women than men. In 2013, there were approximately 128 older women for every 100 older men; in those 85 years and older, there were about 196 women for every 100 men. This has significant clinical and scientific consequences. For example, in current psychiatric practice, psychopharmacological interventions are widely used. In general, prescription drugs are used more frequently by elderly women than men. Overall, women reportedly have a 1.5- to 1.7-fold greater risk of adverse drug reactions. Yet women, specifically those of older age, have been underrepresented in medical investigations. Studies designed to establish

the efficacy, safety, and tolerability of medications have been historically biased for incorporating younger men as study subjects. As a result, established therapeutic guidelines may be inadequate or inappropriate for older women.

Great efforts have been made toward filling the knowledge gap that resulted from the historical underrepresentation of women in research studies. In 1985, the U.S. Public Health Service Task Force issued a report on women's health issues that documented that women were disadvantaged in health care and research. Based on this report, the National Institutes of Health (NIH) released a policy encouraging the inclusion of women in clinical research. In 1990, following an investigative study by the Government Accounting Office (GAO) and urging from the Congressional Caucus on Women's Issues, the NIH established the Office of Research on Women's Health to take responsibility for establishing policy and promoting research on women's health. In March 1993, the U.S. Food and Drug Administration (FDA) stated that it would "end its ban on women's participation in most drug safety tests and require companies to carry out analysis by gender in virtually all applications for new drugs." This declaration followed the GAO October 1992 report "FDA Needs to Ensure More Study of Gender Differences in Prescription Drug Testing." Further guidance was issued in 1993, followed by regulations in 1998 and 2000. Federally funded clinical intervention studies would be required to include women, with data analysis by gender. These legal pronouncements, with continued oversight, help to ensure that women are represented in health care research, which informs current evidence-based medical practice and patient-centered care.

Clearly, patients are better served when gender and age are taken into account. This is especially important for older women as after menopause, morbidity and mortality from conditions such as cardiovascular disease, major neurocognitive disorder, or depression reach or exceed rates seen in men. These differences will be highlighted and illustrated in the sections that follow.

IMPACT OF GENDER ON PSYCHIATRIC CONDITIONS

Major Neurocognitive Disorders

Major neurocognitive disorders are an important health concern in aging populations. Major neurocognitive disorder due to Alzheimer disease and major vascular neurocognitive disorder together account for 80 to 90 percent of all major neurocognitive disorders. Of those diagnosed with major neurocognitive disorder due to Alzheimer disease, about two-thirds are women. Being a woman has been associated with an increased risk for developing major neurocognitive disorder due to Alzheimer disease, although the evidence is not as clear for major vascular neurocognitive

disorder. Research using autopsy data suggests that about half of patients with major neurocognitive disorder due to Alzheimer disease have vascular lesions. There is a growing body of scientific literature on how cardiovascular risk factors may play a role in the development of major neurocognitive disorder due to Alzheimer disease. Obesity in mid-life may influence, to a greater degree in women than in men, the development of major neurocognitive disorder. Hypertension and hypercholesterolemia in both women and men predict a higher risk of developing major neurocognitive disorder due to Alzheimer disease with aging. In women older than 75 years of age, rates of hypertension and hyperlipidemia are higher than in men of similar age. Thus, with greater age, hypertension and hyperlipidemia may play a more prominent role in the development of major neurocognitive disorder due to Alzheimer disease in women. The apolipoprotein E ε4 (ApoEε4) genotype seems to be a risk factor for major neurocognitive disorder due to Alzheimer disease primarily in women and seems to have a more deleterious effect on hippocampal gross pathology and memory performance in women than in men. Diabetes, which increases in frequency with age in women to a greater extent than in men, is associated with an increased risk for cognitive impairment and may influence this trend as well. Whether prevention or reversal of these associated factors can cause a decreased risk of major neurocognitive disorder due to Alzheimer disease in women has yet to be clearly established.

Women with major neurocognitive disorder due to Alzheimer disease are at greater risk of having comorbid conditions undetected and inadequately treated, potentially affecting mortality risk. In a cross-sectional study of 340 community-dwelling elderly with major neurocognitive disorder due to Alzheimer disease, women were at greater risk of having a malnourished state that was undetected by clinicians, despite having a low body mass index (BMI) (less than 21). In a retrospective cohort study of 5,831 men and 17,918 women nursing home residents with major neurocognitive disorder due to Alzheimer disease, women were underweight, less likely to be hospitalized and received fewer medications for each disease compared to men. The women in this cohort were also older and more severely cognitively impaired than the men. In an observational study of 28,367 nursing home residents with major neurocognitive disorder due to Alzheimer disease, it was found that although psychotic and depressive symptoms were equally prevalent in both women and men, women were more likely to receive antidepressants and men were more likely to receive antipsychotics. This may be due, in part, to the fact that men were more likely than women to have behavior problems such as wandering, abusiveness, and socially inappropriate behaviors in this cohort. These findings suggest that greater attention by gender to the nutritional, health, and psychiatric status of elderly patients

with major neurocognitive disorder due to Alzheimer disease may be needed so that appropriate, timely interventions may be proactively implemented.

As major neurocognitive disorder due to Alzheimer disease progresses, a decline in cognitive, functional, and overall health status occurs, with the gradual loss of the ability to live independently. Women and men with major neurocognitive disorder due to Alzheimer disease, if admitted to a nursing home, seem to have different reasons for nursing home admission. A 2-year, four-wave follow-up study with 573 patients with major neurocognitive disorder due to Alzheimer disease and their family caregivers showed that men entered an institution primarily for medical and caregiving needs, but women, who showed less physical and cognitive impairment than men on nursing home admission, entered due to family caregiving patterns. This suggests that providing additional education and support to affect family caregiving patterns might delay nursing home placement for women with major neurocognitive disorder due to Alzheimer disease.

Mortality rates in elderly with major neurocognitive disorder due to Alzheimer disease are higher in women than in men. In a 15-year prospective epidemiological study of 1,670 community-dwelling elderly individuals, it was found that having a major neurocognitive disorder due to Alzheimer disease increased mortality risk overall in women but not in men. Cardiovascular risk factors, the lack of detection and adequate treatment of comorbid conditions, and patterns of family care provided to elderly women with major neurocognitive disorder due to Alzheimer disease may possibly influence mortality. If so, interventions to influence these factors might reduce mortality rates in elderly women with major neurocognitive disorder due to Alzheimer disease.

Depression

Depression in the elderly is prevalent and a significant public health problem. It is two to three times more common in elderly women than in elderly men overall. The comorbidity associated with depression in elderly women is significant. Depressive symptoms statistically significantly increased the risk for major neurocognitive disorder in women 85 years and older. The prevalence of generalized anxiety disorder is higher in elderly with depression. In studies with fairly large cohorts, the association of depression with decreased health status, disability, and decreased quality of life in older women is prominent. In a prospective study of 4,657 elderly women, depression, along with underweight status, obesity, and smoking, was associated with the development of frailty. Greater prevalence of depression was found to contribute substantially to greater disability and decreased quality of life in elderly women compared to men in a study of 1,348 elderly patients. In addition, depression and frailty in

older adults, especially in older women, has been associated with an increased risk of death.

In population-based cohort studies, depression was found to be associated with increased rates of hip bone loss and lower bone mineral density (BMD) in older women compared to younger women. Higher rates of hip fracture in elderly women have been attributed to declining bone strength with age; depression may exacerbate this trend by potentially influencing rates of bone loss and BMD. Of note, some antidepressants and antipsychotics that may be used to treat depression may exacerbate bone loss through causing hyperprolactinemia and changes in bone metabolism, leading to bone demineralization.

It has yet to be clearly determined whether preventing or reducing depression might improve the health status, disability, comorbid conditions, and quality of life in older women, or vice versa. In addition, there may be complex fluctuating, interactive relationships among these factors so that affecting one may or may not clearly lead to an association with decreased depression. In such cases, multipronged interventions that influence all implicated factors together may be most effective clinically.

Depression has been associated with psychosocial and lifestyle features. In a cross-sectional study of 580 geriatric hospital patients, depression was associated with fewer social contacts and supports. This association persisted even after adjusting for physical and cognitive impairment and was stronger for women than for men. In a study of 860 community-dwelling elderly from 60 to 96 years of age, decreased depression was associated with both light and strenuous exercise status in both women and men, with light exercise having a somewhat stronger effect on depression in women than in men. These findings, albeit associations without clearly established causality, raise questions as to whether programs that promote social support and light exercise might affect depression even more in women than in men.

Although depression and some associated conditions are more prevalent in elderly women than in elderly men, there are significant adverse comorbid conditions associated with elderly men. Suicidality is associated with depression, and because depression is more prevalent in elderly women than in elderly men, it may be expected that elderly depressed women have higher suicide rates. However, the highest rates of successful suicide in the United States occur in elderly white men, possibly due to men using more-lethal means in attempting suicide. This trend has continued despite the decrease in suicide rates in all age groups over the last 15 to 20 years. In a community-based sample of 4,051 elderly individuals, major depression with psychotic features was found to increase mortality risk in both women and men, but milder “neurotic depression” was found to increase mortality risk in men only. Determining why women seem to be more “protected” with respect to these factors might help lead to

more effective interventions for both men and women who are elderly and depressed.

Psychosis

In women, the onset of schizophrenia peaks at about age 25 to 30 years, with a second, smaller peak after age 45 years. In men, the age of onset is earlier, at about 20 to 25 years of age. Psychosis is often considered to be uncommon in elderly individuals without major neurocognitive disorder. However, a 3-year longitudinal prospective cohort study of 347 elderly individuals living in community or institutional settings found that the prevalence of any psychotic symptom was 10.1 percent, higher than previously reported. Moreover, hallucinations, delusions, and paranoia were each related to an increased incidence of major neurocognitive disorder in elderly aged 85 to 88 years. Hallucinations and paranoia were associated with increased 3-year mortality in women but not in men.

A descriptive, cross-sectional study of 90 outpatients with early-onset schizophrenia and 34 with late-onset schizophrenia found that a greater proportion of women had late-onset schizophrenia compared to men (41 percent vs. 20 percent). Women overall had more severe positive psychotic symptoms than men, and women with late-onset schizophrenia seemed to have less severe negative psychotic symptoms than men or women with earlier-onset schizophrenia.

In general, these data suggest that women may be more prone to developing psychosis later in life compared to men and that positive psychotic symptoms are more severe and associated with greater mortality in women than in men. However, antipsychotics are frequently used to treat psychotic signs and symptoms, and use of these medications may affect mortality rates. Both typical and atypical antipsychotics are associated with greater risk of death, so that cautious use of these medications becomes very important.

Other Psychiatric Conditions

Generalized anxiety disorder is much more prevalent among the elderly population than previously believed. In general, anxiety has been reported to be more frequent in women than in men. Among the elderly population, this trend seems to continue. In one study of 103 elderly individuals with generalized anxiety disorder, women comprised 54 percent (25/46) of patients with generalized anxiety disorder and 70 percent (40/57) of patients with both generalized anxiety disorder and major depression.

Regarding bipolar disorder, one epidemiological study found that elderly patients with bipolar disorder reported having higher lifetime and 12-month rates of alcohol use, dysthymia, generalized anxiety, and panic disorder than those without bipolar disorder. Elderly women with bipolar

disorder reported more panic disorder, whereas older men with bipolar disorder reported more alcohol dependence.

Much more needs to be understood about anxiety and mania in the elderly population, especially with respect to gender issues, before specific recommendations for treatment can be made confidently. However, if the trend toward more anxiety disorders in elderly women persists, it suggests that the judicious use of antianxiety agents might be especially pertinent for women compared to men.

Deterioration of the quality of sleep occurs in both women and men with aging. However, when women enter menopause, they are at higher risk for even greater disruptions in sleep patterns. Gender-appropriate treatment may include time-limited hormone replacement therapy used in the early stages of menopause.

TREATMENT ISSUES

The management of psychiatric conditions includes psychosocial, behavioral, milieu, and other psychotherapeutic interventions, as well as biologically based treatments such as medication or brain stimulations therapies like electroconvulsive therapy and transcranial magnetic stimulation.

Psychotherapeutic Interventions

When treating geriatric patients, there are unique aspects to transference and countertransference that must be addressed. Older patients may view younger therapists as being similar to their children and less experienced or knowledgeable about life issues. The elderly may feel more stigmatized at having to seek psychiatric help. Although the therapist's understanding of the older patient's relationships with both parents continues to be important, it may be even more essential to determine what the patient's relationships with children and other mental health professionals have been to effectively treat the older patient. Therapists must pay special attention to their attitudes and biases regarding aging, sexuality, disabilities, losses, and death so that these vital concerns in aging are not avoided and optimal adaptations can be facilitated with the elderly patient. Ageist and stereotyped views of elderly adults as "shipwrecks at the end of their voyage—hard of hearing, with low vision, physically disabled, asexual, unattractive, slow, unproductive, and close to death"—deny many elderly the professional psychiatric help they need, in a respectful, gender-sensitive, reality-oriented, compassionate, and knowledgeable manner that promotes ongoing growth, development, adaptation, friendship, and enjoyment of life.

Gender-specific developmental theory for later life has not yet been established. However, clinically, eclectic psychotherapeutic techniques

incorporating aspects of psychodynamic, cognitive-behavioral, interpersonal, and problem-solving therapy have been found to be useful for older women and men. Individual, couples, family, and group therapies may all be helpful in appropriate situations, and peer support and education are vital.

A psychotherapeutic issue of focus in elderly patients includes the need to adapt to being more dependent than in their younger adult years. Older women and men who, during their younger adult lives, prided themselves on being self-sufficient and able to nurture others may find dependency on others especially difficult. When assistance with personal care is necessary, older patients may have more difficulty accepting direct care from caregivers of the opposite gender, especially if the caregiver is a son caring for a mother or a daughter caring for a father. Excessive resistance to dependency can result in self-injurious or self-neglectful behaviors. Not wanting to be a burden to others (especially one's children) is a common concern and can often be successfully addressed in psychotherapy by pointing out the vital contributions that the older patient has made to the family over a lifetime. Unrealistic and excessive feelings of entitlement and dependency on others beyond actual need and capacity, often with an excessive focus on losses and somatic concerns, may also require therapeutic attention. The individual etiologies and meaning of such behaviors need to be explored and clarified to patients in a gender-sensitive manner.

Family participation is a vital component of psychotherapy with dependent elderly persons. More women in the advanced years live alone, have more chronic illness, and may need assistance from adult children. The readiness for emotional, physical, and, at times, financial commitment to parental caretaking can vary widely among adult children in a family. The skilled and sensitive psychotherapist can assess and effectively distribute tasks and can help families to function as a quasi-multidisciplinary team while facilitating access to appropriate community resources. By not assuming that the women in the family will take on all housekeeping and caretaking responsibilities, gender barriers to participating in caregiving can, at times, be overcome. Family involvement in medication management and treatment adherence, as well as many other activities of daily living for the elderly patient, remains a professional standard in the therapeutic interventions for the dependent, frail elderly. Even if professional caregiving services are paid for by insurance, family involvement often is essential for the appropriate access and deployment of such services.

Another psychotherapeutic issue of focus for elderly patients relates to the developmental phase of becoming a grandparent. This subject needs attention in theory, research, teaching, clinical practice, and policymaking. The stereotyped image of the grandmother as an unemployed or retired

baby-sitting, meal preparing, cookie baking, and ever-available individual has become less and less of a reality. Usually, no such expectations are made of the grandfather, whose interest and assistance are often appreciated at whatever level he chooses. Grandfathers, especially when retired, may enjoy the new roles of helper and caregiver in which they may not have participated earlier in their lives. A therapist's role in a grandparent's life can include resolution of intergenerational conflicts and facilitation of the development of this new and challenging identity. Grandmothers who live alone and at a distance are less likely to have the financial resources to visit other family members than grandfathers or grandparents who live together. Clinical observations have revealed that depriving older women of the opportunity of regular visits with adult children and grandchildren can contribute to depression and considerable distress. Other clinical observations reveal that older parents who survive alcohol use disorders from their younger years tend to get depressed when rejected by adult children who cannot forgive their abusive behaviors from earlier years. This rejection inadvertently leads to deprivation of the grandparent role.

A third psychotherapeutic issue of focus involves addressing cultural, societal, and economic predicaments that the elderly patient may be facing. The older population includes persons born around the turn of the previous century, as well as those born after World War II. Many members of the aging population came to the United States during various immigration waves. Thus, cultural heterogeneity is vast in both elderly women and elderly men. Intergenerational and gender-based variations in cultural practice and language may lead to conflict, calling for cross-cultural psychotherapeutic intervention.

The elderly have experienced many societal changes in the United States. Gender roles for women have changed from being primarily defined by men in power and authority to being less stereotyped and more self-determined. Yet, it is common knowledge that women who succeed in this male-dominated world, be it in political, corporate, academic, or other venues, remain more often than not in the minority. Most of these women learn the "male rules of the game," and many modify their lifestyles by remaining unattached, delaying childbearing, enduring the trials and tribulations of fertility specialization technologies, adopting children, or having no children. Communication styles, behaviors, and responsibilities within and between genders, families, and generations have changed markedly. Elderly patients may benefit from psychotherapy in dealing with the causes and consequences of these tremendous fluctuations in societal roles and circumstances.

Coping with financial issues, including payment for health care and extended care facility placement, is often a focus of psychotherapeutic intervention for the elderly. Money management is becoming more

complex. Greater numbers of women rise to this task earlier in life, and more women without experience in financial matters have to rise to the task in later life. The economic effects of lower educational levels, traditional work roles, and income discrimination experienced by women in young and middle age become exaggerated as retired women depend on pensions and Social Security derived from preretirement income. Women are about one-half as likely as men to have pensions, and the amount of their pension is one-half as much as that of men. When men die, they are three times as likely to leave an impoverished spouse. Single, widowed, and divorced women often have to purchase the in-home care that married men are provided through their spouses. Women live longer, have more disabilities, and live alone more often than men. Thus, they often require the services of nursing homes more than men do. Seventy-five percent of nursing home residents are women aged 65 years or older. Women are four times more likely than men to use Medicaid for nursing home care because of their relative indigent status. Due to relatively low Medicaid reimbursement to nursing homes, these women have little choice in the selection of a nursing home. They must share a room, their beds can only be held for a limited period of time in case of hospitalization, and medications must be generic. Older men with spouses to look after them at home use more acute care hospital Medicare dollars than women. Medicare does not pay for skilled nursing home care for more than 100 days (with days 21 to 100 requiring a coinsurance payment), only covers rehabilitative nursing home care after a 3-day hospitalization, and provides limited coverage for prescriptions. In the past, Medicare provided only 50 percent coverage for outpatient psychotherapy visits; with the movement for parity in mental health, Medicare covers 80 percent. Many managed care companies with Medicare contracts have discontinued care provision for the elderly population, leaving many without continuation of much-needed medical care from their primary care physicians. Education and problem solving regarding these issues, and psychotherapeutic support that promotes the patient's psychological strengths and resources, may empower patients to cope with these difficult situations with a greater sense of hope and mastery.

Over the last few decades, online (Internet) computer use has grown in the population in general and may be a source of psychoeducational and peer support for the elderly, especially for those who are more isolated. One study of 174 subjects found no significant main effects for gender and online communication and relationship preferences for friends, family, coworkers, or unknown individuals. However, high Internet users had higher scores on online communication and relationship building. Another study found that although access to computers was equal for both women and men, in the year 2000 women were 0.7366 times and in the year 2002 they were 0.5675 times less likely than men to use computers to obtain

web-based health information.

Psychotherapeutic treatment for family caregivers of frail or disabled elderly is important. Family caregivers often continue to work full-time and also may have children for whom they are caring. These caregivers are at times referred to as "the sandwich generation," caught between the tasks of caring for children and elderly. Many sites of employment continue to be unfriendly to family caregiving. Extended-family members, including grandmothers, who traditionally might have been able to provide more care might be employed and less available for helping in child and dependent elder care, or need care themselves. Reconstituted families on their second and third marriages are spread thin between children of spouses. Availability and quality of community resources for caregiving have remained inconsistent because of inadequate funding and staffing. Consequences of many of these changes are manifested in the crisis of informal caregiving of the elderly who are challenged by mental illness, physical conditions, cognitive disorders, and/or functional deficits.

Pharmacological Interventions

The complexities of incorporating gender issues are often not addressed in the scientific literature specifically, especially among the elderly. However, there has been some research evaluating gender differences in the pharmacokinetics and pharmacodynamics of drugs in animals and younger humans.

Pharmacokinetics. Pharmacokinetic investigations in animals and younger patients have revealed gender differences in various factors that affect drug absorption and bioavailability, drug distribution, metabolism, and elimination. Gender differences have been documented for gastrointestinal motility and metabolizing enzyme effects, but studies have not consistently shown a difference in drug bioavailability between women and men. There are few studies examining differences in bioavailability, and these are generally confounded by variability in other pharmacokinetic factors, such as distribution, metabolism, and excretion. The available evidence suggests that gender-based differences in bioavailability are not clinically significant.

Gender-related differences in drug pharmacokinetics may be due primarily to differences in drug distribution, given that many of the factors influencing volume of distribution differ by gender. In general, women, compared to men, have a higher percentage of body fat, lower mean body weight, decreased plasma volume, and less organ blood flow, leading to differences in the rate and extent of drug distribution. In women, compared to men, due to differences in body fat, lipophilic drugs may have greater volumes of distribution and water-soluble compounds may have less. A larger volume of distribution for diazepam has been observed in women

compared to men, thought to be due to differences in proportions of body fat and gender-related protein-binding changes.

Gender-related differences in plasma protein binding may lead to pharmacokinetic differences for certain agents. Although albumin levels usually are not affected by gender, endogenous and exogenous estrogens decrease levels of α-1 acid glycoprotein (AAG) in plasma and induce hepatic AAG-glycosylation. Exogenous estrogens also increase serum-binding globulins such as sex-hormone–binding globulins, corticosteroid-binding globulin, and thyroxine-binding globulin. Variations in levels of such plasma-binding proteins may change the amount of active agent that is available. This may be important to consider when interpreting serum drug levels.

Pharmacokinetic variability is influenced greatly by gender-related differences in drug metabolism. Hepatic drug metabolism is influenced especially by hepatic enzyme activity, as well as by liver blood flow, which is less in women than in men. The cytochrome P450 3A (CYP 3A) subfamily, especially CYP 3A4 in humans, is active in the initial hepatic metabolism of a large number of medications. Other prominent cytochrome P450 enzymes include CYP 1A2, CYP 2B6, CYP 2C9, CYP 2C19, CYP 2D6, and CYP 2E1. Drugs also can be metabolized through intestinal CYP enzymes. The frequency of hepatic metabolizing enzyme variant alleles has been reported to differ among ethnic groups and by gender in younger humans and animals. Gender-based differences in drug levels have been attributed to gender differences in hepatic enzyme activity, which also are influenced by endogenous sex hormone production and hormonal changes associated with exogenous sex hormone use, pregnancy, and menopause. Enzyme CYP 2D6 has been found to be important in metabolizing antipsychotic agents. It has been reported that tardive dyskinesia develops more frequently in Chinese women with schizophrenia than in men due to the increased frequency of defective CYP 2D6 alleles in Chinese women. The activity of enzyme CYP 1A2 has been found to be prominent in the metabolism of antipsychotics and is higher in females than in males.

Drug excretion usually is through the kidney or the liver. Renal clearance of drugs depends on the glomerular filtration rate, which is directly proportional to weight and thus higher in men than women, in general. Gender-related differences in drug renal excretion rates are most likely due to weight differences.

Pharmacodynamics. Pharmacodynamic changes in drug response based on gender have not been as well studied as pharmacokinetic differences. Gender-related pharmacodynamic changes are differentiated from pharmacokinetic effects by showing that a given plasma level of active drug in both women and men does not result in the same pharmacological outcome.

Many psychotropic medications have been reported to have gender-related modifications in pharmacodynamics. Women have greater impairment of psychomotor abilities from diazepam compared to men. This is thought to be due to heightened diazepam pharmacodynamic sensitivity in women. Compared to men, women show more improvement in psychotic signs and symptoms and greater severity of adverse side effects from typical antipsychotic agents. Further studies need to be done to determine whether these differences are caused primarily through pharmacodynamic or pharmacokinetic processes. Some but not all research suggests that women and men appear to show different effects to various antidepressants, although more work is needed to clarify these findings.

FUTURE DIRECTIONS

The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) includes sections on Gender–Related Diagnostic Issues for many diagnoses. Yet, it could greatly benefit from more evidence-based gender- and age-specific data, especially regarding the neurocognitive disorders. Incorporating gender into the evaluation and management of psychiatric conditions in an evidence-based fashion is important but extremely complex, given that the differences or similarities in responses and outcomes that are observed between elderly women and men may be due to multiple factors. These factors may be affected by contextual, sociocultural roles and attitudes, genetically based, affected by prenatal hormone organizational effects, related to natural or drug-induced hormone fluctuating activational effects, associated with variable circadian/sleep patterns, or influenced by pharmacokinetic/pharmacodynamic characteristics. In elderly individuals, the sum of these, as well as the effects of aging, may play a role. To evaluate these factors carefully, studies ideally would include standardized psychiatric and functional rating scales that are gender sensitive, incorporate sociocultural contextual measures (such as those evaluating immigration status, ethnocultural and religious values, family and caregiving roles, employment status, and socioeconomic level), assess genetic factors, have greater numbers of observations to account for potential behavioral fluctuations that may be related to changes in hormone levels and circadian/sleep patterns of behavior, evaluate pharmacokinetic and pharmacodynamic factors using gender-appropriate methods, be prospective and longitudinal in nature, and enroll both women and men in adequate numbers to be able to carry out meaningful comparative analyses by gender. Routinely incorporating functional neuroimaging investigations into descriptive and psychotherapeutic/drug interventional studies also may enhance knowledge regarding gender-based differences and similarities in brain function. Data on outcomes with absolute and proportional indices, sociocultural contextual measures,

genetics, pharmacokinetics, pharmacodynamics, adverse effects, dropouts, onset and rates of psychotherapeutic or drug intervention effects, and placebo response should be reported by gender. Such all-inclusive studies are challenging to design, carry out, and complete. Due to constraints in resources, time, subject participation, and other practical considerations that can affect scientific rigor, most studies are planned using partial models, focusing thoroughly on a few aspects of any broadly comprehensive study. Approaches such as study replication/semi replication with convergence of outcomes and statistical techniques using meta-analysis or physiology-based whole-body population models for assessing the influence of individual variability on the pharmacokinetics of drugs are used in an attempt to help shed light on complex issues. Still, information garnered from the appropriate analysis of primary data of well-designed, highly powered studies would be most convincing.

The Society for Women's Health Research (http://swhr.org) highlights many specific alarming gender gaps in knowledge that hinder effective prevention, treatment, and cure for conditions that are prevalent in women. Research, clinical, and educational attention and resources need to be expanded to include the older population in ways that are equitable to their numbers, especially older women.

Gender-based life-cycle studies, of which Erik Erikson, George Vaillant, and others have been the pioneers, should be carried out simultaneously for women and men, rather than sequentially for each gender, to provide the necessary information regarding which lifestyles and conditions promote or hinder healthy aging for women and men. Appropriate and useful engagement of aging men and women regarding the benefits of rapidly changing information technologies needs ongoing attention and development. As psychosocial processes affect illness outcome, so does illness affect psychosocial processes. Thus, affective, cognitive, and functional status need routine assessment in primary care, and all medical records from primary care physicians, including gynecologists, need to be available to the psychiatrist before the patient's first visit. The managed care interfaces with necessary state-of-the-art medical and mental health care need serious attention and overhaul. Interdisciplinary work, rather than mental health carve-outs and fragmentation of medical mental health care, has been shown to benefit patient care and to be cost-effective.

The caregiving crisis needs major policy attention.

Education for all primary care and specialties in age- and gender-sensitive diagnosis and interventions needs to start early in medical school and residencies.

Much progress has been made in promoting gender-sensitive research and the prevention, evaluation, and treatment of illnesses. However, tremendous amounts of work still need to be done to promote greater understanding of the similarities and differences between genders in the

elderly patient relative to the description, etiology, treatment, and intervention of psychiatric illnesses and the enhancement of healthy behavior.

REFERENCES

Altmann A, Tian L, Henderson VW, Greicius MD; Alzheimer's Disease Neuroimaging Initiative Investigators. Sex modifies the APOE-related risk of developing Alzheimer disease. *Ann Neurol*. 2014;75:563–573.

Barry LC, Allore HG, Guo Z, Bruce ML, Gill TM. Higher burden of depression among older women: the effect of onset, persistence, and mortality over time. *Arch Gen Psychiatry*. 2008;65:172–178.

Bashir M, Holroyd S. Caring for the elderly female psychiatric patient. *Psychiatr Clin North Am*. 2010;33:475–485.

*Bigos KL, Pollock BG, Stankevich BA, Bies RR. Sex differences in the pharmacokinetics and pharmacodynamics of antidepressants: an updated review. *Gend Med*. 2009;6:522–543.

Brown PJ, Roose SP, Fieo R, et al. Frailty and depression in older adults: a high-risk clinical population. *Am J Geriatr Psychiatry*. 2014;22:1083–1095.

de la Monte SM, Tong M. Brain metabolic dysfunction at the core of Alzheimer's disease. *Biochem Pharmacol*. 2014;88:548–559.

*Franconi F, Campesi I. Pharmacogenomics, pharmacokinetics and pharmacodynamics: interaction with biological differences between men and women. *Br J Pharmacol*. 2014;171:580–594.

Gudmundsson P, Lindwall M, Gustafson DR, et al. Longitudinal associations between physical activity and depression scores in Swedish women followed 32 years. *Acta Psychiatr Scand*. 2015;132:451–458.

Haring B, Leng X, Robinson J, et al. Cardiovascular disease and cognitive decline in postmenopausal women: results from the Women's Health Initiative Memory Study. *J Am Heart Assoc*. 2013;2(6):e000369.

Kerwin DR, Gaussoin SA, Chlebowski RT, et al. Interaction between body mass index and central adiposity and risk of incident cognitive impairment and dementia: results from the Women's Health Initiative Memory Study. *J Am Geriatr Soc*. 2011;59:107.

Kinon BJ, Liu-Seifert H, Stauffer VL, Jacob J. Bone loss associated with hyperprolactinemia in patients with schizophrenia. *Clin Schizophr Relat Psychoses*. 2013;7:115–123.

*Kyomen HH, Goldstein MZ. Gender issues. In: Sadock BJ, Sadock VA, Ruiz P, eds. *Comprehensive Textbook of Psychiatry*. 9th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2009.

LaTorre D, Falorni A. Pharmacological causes of hyperprolactinemia. *Ther Clin Risk Manag*. 2007;3:929–951.

Lorence D, Park H. Gender and online health information: a partitioned technology assessment. *Health Info Libr J*. 2007;24:204–209.

Luppa M, Luck T, Weyerer S, Konig H, Riedel-Heller SG. Gender differences in predictors of nursing home placement in the elderly: a systematic review. *Int Psychogeriatr*. 2009;21:1015–1025.

McDougall FA, Kvaal K, Matthews FE, Paykel E, Jones PB. Prevalence of depression in older people in England and Wales: The MRC CFA Study. *Psychol Med*. 2007;37:1787–1795.

Mirza SS, Wolters FJ, Swanson SA, et al. 10-year trajectories of depressive symptoms and risk of dementia: a population-based study. *Lancet Psychiatry*. 2016;3:628–635.

National Institute on Drug Abuse. Prescription Drug Abuse. Available at http://www.drugabuse.gov/publications/research-reports/prescription-drugs/what-prescription-drug-abuse. Accessed August 15, 2015.

Rathee R, Rallabandi VP, Roy PK. Age-related differences in white matter integrity in healthy human brain: evidence from structural MRI and diffusion tensor imaging. *Magn Reson Insights*. 2016;9:9–20.

Rocca WA. Sex and gender differences. *Arch Gen Psychiatry*. 2008;65:172–178.

Rocca WA, Mielke MM, Vemuri P, Miller VM. Sex and gender differences in the causes of dementia: a narrative review. *Maturitas*. 2014;79:196–201.

Schneider JA, Arvanitakis Z, Bang W, Bennett DA. Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology*. 2007;69:2197–2204.

Schulze J, Glaeske G, van den Bussche H, et al. Prescribing of antipsychotic drugs in patients with dementia: a comparison with age-matched and sex-matched non-demented controls. *Pharmacoepidemiol Drug Saf*. 2013;22:1308–1316.

*Smith S. Gender differences in antipsychotic prescribing. *Int Rev Psychiatry*. 2010;22:472–484.

Sundermann EE, Katz MJ, Lipton RB. Sex differences in the relationship between depressive symptoms and risk of amnestic mild cognitive impairment. *Am J Geriatr Psychiatry*. 2017;25:13–22.

*United States General Accounting Office. Women's health. women sufficiently represented in new drug testing, but FDA oversight needs improvement, July 2001. Available at http://www.gao.gov/new.items/d01754.pdf. Accessed August 15, 2015.

Williams LJ, Pasco JA, Jackson H, et al. Depression as a risk factor for fracture in women: a 10 year longitudinal study. *J Affect Disord*. 2016;192:34–40.

Woods NF, LaCroix AZ, Gray SL, Aragaki A, Cochrane BB; Women's Health Initiative. Frailty: Emergence and consequences in women aged 65 and older in the Women's Health Initiative Observational Study. *J Am Geriatr Soc*. 2005;53:1321–1330.

57.6e The Legal and Ethical Analysis of Elder Abuse

PAUL RICHARD GREENWOOD, L.L.B. (HONS.), ANSAR M. HAROUN, M.D., AND DAVID NAIMARK, M.D.

INTRODUCTION

Americans who are 65 years or older are the fastest growing population in the United States and in much of the developing world. As an example, the U.S. Census Bureau has projected that California's elderly residents will nearly double within the next 20 years—from 3.7 million to more than 6.4 million. The mounting statistics suggest that humans are on the verge of essentially being a "graying world." The term "elder abuse" is now as familiar as spousal abuse and child abuse.

In the past, law enforcement has routinely minimized the importance of cases involving elderly victims. Police officers have tended to stereotype older victims, prosecutors have been reluctant to allow seniors to testify, and judges have underestimated the impact of abuse upon the victim. Within the medical profession there has been a reluctance to accept that an obligation is owed not only to the elderly patient, but also to the community, to report suspected incidences of elder abuse.

Most states have specific statutes in their criminal codes addressing "elder abuse" crimes, yet few prosecutorial agencies take advantage of such laws, in part because of outdated perceptions about seniors. Some of the classic stereotyping of older citizens portrays them as being senile, fragile, long-winded, grumpy, hard of hearing, and unable to see well.

It is not unusual for police officers to write incomplete crime reports when the alleged victim is elderly because of an unfounded belief that the victim is incapable of providing sufficient information that could lead to a successful prosecution, that he may not have a reliable memory, that his senses are not adequate, or that he may not be strong enough physically to pursue prosecution. A number of communities have recognized the failings of such misconceptions and the use of multidisciplinary teams (MDTs) has created an effective response to combat this escalating crime. Psychiatrists are as an essential component of MDTs that can provide much needed information on cognitive issues for law enforcement, adult protective workers, and prosecutors.

DEFINITIONS

For the sake of clarity, a few definitions are proposed. Elder abuse can take three basic forms—physical, financial, and emotional/psychological. Within these broad categories are several other subcategories.

Physical elder abuse includes assault and battery, neglect, and sexual assault. Assault and battery is primarily inflicted by either: (1) opportunists who target the elderly in the streets with muggings and robberies, (2) family members (typically sons, daughters, or grandchildren) who systematically wield aggression against the elderly relative, or (3) care providers who, out of a sense of frustration, lash out in anger against the victim.

Neglect may be present when a care provider willfully or deliberately fails to provide basic services of hygiene, medical support, or nutrition to the victim. The term "willful" encompasses circumstances where a caregiver intentionally "turns a blind eye" to an obvious condition that requires a positive act in order to alleviate a situation that is endangering the health of the elderly patient. Examples of neglect would include a failure to turn a bedridden elder (causing Stage 4 pressure sores) or a deliberate withholding of crucial medications (resulting in deterioration in the elder's health).

Sexual assault typically occurs through force, threats, or deception and culminates in unwanted sexual activity. Such conduct can include rape, nonconsensual kissing or fondling, and the taking of sexually explicit photos or videos of the elderly victim.

Financial elder abuse manifests in several forms:

Theft—Most commonly, thefts of personal items including jewelry and antiques (taken by a care provider) that inevitably wind up in the local pawn shop.

Theft of individual checks—Again, most commonly committed by a trusted care provider, checks are usually ripped from the back of the victim's checkbook and when cashed, are typically out of sequence.

Misuse of an ATM card—Is likely to be perpetrated by somebody who has gained the trust of the victim, or by the use of threats and intimidation.

Credit card fraud—Going beyond just the misuse of personal credit, the perpetrator successfully forges the victim's details into a new credit application that has been obtained without the victim's consent or knowledge.

Theft of assets—Involves large-scale pilfering of assets such as savings accounts, stocks, or real property. Trusted individuals may employ a power of attorney (POA) or quitclaim deed, either of which will be obtained through stealth or intimidation. A POA is a legal document that permits another person to make legal decisions on behalf of the elder. Depending on how the document is drafted it may permit the other party to sign checks, open new accounts, or transfer assets belonging to the elder. A quitclaim deed is a legal document that transfers an interest in real property to another person.

Scams—Today's uncertain financial climate is fertile ground for crooks and conmen to target the elderly with a wide variety of schemes. A few are listed here:

Telemarketing fraud such as bogus charity schemes, illegal sweepstakes, and false investments take advantage of susceptible elders.

Door-to-door salespeople who convince the elderly victim to pay up-front for labor services such as a new roof, driveway, or other home improvement. These con artists may make off with the up-front money or stay around to bilk the senior out of his or her savings.

Pretending to be representatives of utilities or roofing companies or some other trustworthy agency, thieves gain access to a senior's home with the intention of either stealing what they can, or casing the residence for a return visit and a more comprehensive theft

opportunity.

Exchanging a senior's assets in return for a false promise of "lifetime care" relies on a fear that often haunts a senior's contemplation of the future, when he or she may be less self-reliant.

Telephone scams involve a caller pretending to be a relative, such as a grandchild, who is in some kind of trouble and needs money.

Emotional/Psychological elder abuse—Normally presents as verbal threats by a family member. Psychological abuse may include threats, intimidation, humiliation, or other such degradation. Threats may encompass hints of personal violence against the victim, a loved one, or a pet if the victim fails to deliver on financial "commitments."

NOSOLOGY

DSM-IV-TR contained a section on "Problems Related to Abuse or Neglect," which included a subsection on "Abuse of Adult," both physical and sexual, but not psychological! The more recent DSM-5 renames this area as "Adult Maltreatment and Neglect Problems" and continues with Adult Physical Abuse and Adult Sexual Abuse, but also adds Adult Psychological Abuse. This change reflects the growing realization that maltreatment may take many forms, including:

- Physical and sexual abuse, recognized in both DSM-IV-TR and DSM-5.
- Psychological abuse, now recognized in DSM-5.
- Neglect, which has a DSM-5 code for both child neglect and for spouse or partner neglect, but not for adult neglect (where neglect of elders may be coded).
- There remains no DSM-5 code for
- Self-neglect, where there is no perpetrator, but the elder fails to provide for self in a proper manner, thus threatening personal health or safety.
- Abandonment, when a caretaker or legal guardian leaves the elder to fend for herself.
- Exploitation, where a vulnerable elder is the victim of some scam, usually material or financial.

FACTORS ASSOCIATED WITH ELDER VULNERABILITY

Old age can create a unique profile, sometimes enabling exploitation. The elderly are likely to have:

- desirable assets (money in various forms)
- legal rights to control those assets through their decision-making
- some impairment in decisional capacities
- psychological desire to exercise autonomy

Because elders often recognize some of their limitations (in their knowledge or in the process required for wise decision-making), they may develop relationships to enhance such decision-making, and thus be guided by advisors whose selection is based on:

- trust (family members who are trust worthy, but not necessarily wise)
- expertise (lawyers and financial planners who may be clever, but not trustworthy)
- legitimate (blessed by some authority, like a legal guardian or a priest)
- knowledgeable (such as their treating doctor)
- able to reward or punish (their caregiver or whomever gives them spending money)
- charismatic (able to charm the elder)

These characteristics have no relationship to the moral character of the advisor, who may be a saintly figure (determined to promote the interests of the elder), or a "scoundrel" (interested only in self-interest). The combination of an advisor who is motivated only by self-interest and an elder who is vulnerable, accordingly raises the risk of exploitation. Such exploitation may violate the law, or ethics, or both. The law is written in language that is precise, so there is usually a crisp boundary between what is legal and illegal, and such (illegal) exploitation is thus easy to recognize. It may be more difficult to recognize some behaviors as exploitative if they do not violate the law, but are ethically questionable.

MULTIDISCIPLINARY TEAMS

In recent years, in response to the escalation in reported elder abuse incidents, many communities have developed an MDT. Some states have provided for the creation of such teams through legislation. Examples are the establishing of a Financial Abuse Specialist Team (FAST) or an Elder Death Review Team (EDRT). Key members of an MDT are representatives from Adult Protective Services (APS), law enforcement, prosecutors, financial institutions, medical examiners, hospital personnel, mental health case workers, the Public Guardian, and psychiatrists.

APS is the agency in the majority of states that provides social services to elders and adults with disabilities. However, unlike their counterpart Child Protective Services, APS relies heavily upon the cooperation of the "client." If the victim refuses services offered by APS, then typically the investigator has no option other than to walk away and close out the case.

There is no uniform definition for an elder who falls under the jurisdiction of APS. Some states use age as definitive. For example, California defines an elder as anyone 65 years of age or older, while Illinois regards an elder as anyone 60 years or older. Other states use a variety of definitions including "endangered adult" or "vulnerable adult."

Psychiatrists should become familiar with their local APS agency and with their State's definition of "elder." A useful starting point is to contact the National Adult Protective Services Association (NAPSA). Additionally, inquiry should be made of the local APS as to whether an elder abuse prevention MDT exists. Psychiatrists can provide valuable insight at such meetings, particularly when there are discussions about undue influence and consent. Most MDTs have a Memorandum of Understanding (MOU) that clarifies issues of confidentiality and potential conflicts of interest.

An important role of the MDT is to provide ongoing education to professionals about current elder abuse issues. Special attention should be aimed at improving interviewing techniques, the importance of photographs and videotaped victim interviews, and in spotting red flags and indicators of abuse.

Mandated reporters need to understand their duty to report suspected elder abuse. There should be ongoing discussions about enhancing or establishing a public awareness campaign that informs the community of the correct channels for reporting suspected cases of abuse. MDT members should orchestrate community efforts to create billboards, posters, and public service announcements on local television and radio. Slogans such as "abuse is getting old" have been found to be an effective means of getting the message out to the public.

The MDT sessions should use every opportunity to educate personnel who deal with elders in need, such as APS case workers, emergency room nurses and doctors, paramedics, and fire officers in awareness, recognition, and response:

APS workers: after an initial referral is made via a "1-800" telephone number, a caseworker is assigned to pay an unannounced call upon the senior. These workers need to be tuned into looking for possible signs of abuse in order that they will know when to bring in law enforcement.

Emergency room and other medical personnel: sometimes an elderly patient will appear at a hospital suffering from injuries sustained from an assault or through neglect. Medical practitioners need to be able to differentiate between accidental and intentionally inflicted injuries. The victim may be unwilling to explain the cause of such injuries truthfully, and it will be up to the triage nurse to obtain as many details as possible regarding the patient's living conditions and the events leading up to the patient's arrival at the hospital.

Paramedics: at times, these professionals are the first responders to a domestic call for assistance when an elderly person is in need. If the paramedic is not trained to observe certain indicators of abuse or neglect, there is a danger that the police will not be called to the scene. Fire officers, like paramedics, are often the first responders to the home of an elderly victim. They also require training to spot potential abuse or neglect.

Financial workers: the venue for many crimes of financial elder abuse is

a bank, credit union, or stockbroker firm. The perpetrator will often take the elderly victim into a financial institution in order to prepare a POA or arrange for a transfer of monies. The staff at these institutions may or may not be trained to spot various techniques used to exploit seniors. MDT members dealing with personnel from financial institutions can help alert them to red flags and suggest possible responses. For example, financial staff should make it a policy to discuss any alterations in financial arrangements or substantial withdrawals with elderly customers in private, away from the elder's companion. Some states have already established protocols—notably Massachusetts and Oregon—and these protocols provide excellent resources for other states. Several states have enacted laws that provide immunity to banks when reporting a suspected case of financial elder abuse to APS or to law enforcement.

Attorneys who provide legal assistance to elderly clients: need to be educated about the signs of dementia and the impact that cognitive impairment can have on the elderly client's decision making. Psychiatrists can give valuable insight to these professionals about the need for sensitivity in gauging an elder's capacity to make sound legal decisions.

THE RED FLAGS AND INDICATORS OF ABUSE OF AN ELDER

At times, signs of elder abuse may be mistaken for the initial symptoms of dementia or merely the result of increasing frailty. Caregivers may offer logical-sounding explanations for everything from unusual bruising to abnormally frequent check writing. While many caregivers are warm and caring people who are trustworthy, professionals who interact with elders need to examine situations with a keener eye than the general population.

A victim of elder abuse may become increasingly withdrawn from the community and much more dependent upon the perpetrator. If an unannounced visit by a caseworker from APDS is made, the perpetrator may insist that: "this is not a good time." Excuses might include that the victim is sleeping or is unwell. Even if access is given, the caregiver may try to prevent a private conversation between victim and the social worker or friend by leaving the door open or by hovering over the victim. Family and friends who attempt to make telephone calls are often prevented from speaking directly to the victim or the perpetrator will monitor the phone call in some fashion.

Notable physical signs include:

1. Unexplained signs of injury such as bruises, welts, or scars. It is especially concerning if they appear to be symmetrical or shaped like a specific object.
2. Broken bones, sprains, or dislocations.
3. Overdose of prescription drugs or indications of medication not being taken regularly (a prescription has more remaining than it should, or

remains unfilled).
4. Broken eyeglasses or frames.
5. Signs of physical restraints (marks around the wrists or ankles).
6. Unaccounted weight loss, signs of malnutrition, and dehydration.
7. Untreated physical problems (bed sores, urinary tract infections, etc.).
8. Unsanitary living conditions (dirt, bugs, soiled bedding, and clothes).
9. Elder being unbathed and unkempt.
10. Unsuitable clothes for the weather.
11. Living without essentials (no heat or running water, faulty electrical wiring, and other fire hazards).
12. Elder left alone in a public place.

In financial transactions at a bank, the suspect caregiver will often attempt to control the conversation with the bank official and will endeavor to speak on behalf of the victim. The bank clerk may be told that the elderly customer is sick or hard of hearing. The suspect will justify the need for a POA, or will insist that any withdrawals are made in cash.

Other financial red flags include: larger than usual withdrawals, sudden changes in the elder's financial condition, items or cash reported missing from the senior's home, suspicious changes in wills, POA, titles, and policies. Sometimes, names are added to the senior's signature card and bills are uncharacteristically left unpaid. There may be evidence of financial activity that the senior could not have initiated because of physical impairments, contracts for unnecessary services, goods, or subscriptions, or evidence of over- or under-medication.

SENIORS AND THE COURT SYSTEM

For many seniors, their own abuse situation will be their first exposure to the criminal justice system. There are often fears, misconceptions, and feelings of reluctance. It may be important for a treating psychiatrist to discuss such feelings with their patient prior to a hearing where the elder will be required to testify. Moreover, with the elder's permission, it could prove helpful for the psychiatrist to reach out to the prosecutor or victim advocate prior to the hearing to discuss any major concerns.

Competency of a witness to testify—the general rule is that every person is competent to be a witness in a court proceeding. The threshold is low; and typically the witness need only show that he or she understands the significance of taking the oath and be able to appreciate the difference between telling the truth and a lie.

ESTABLISHING A "THEFT" IN ELDER FINANCIAL ABUSE

Sometimes it is difficult to ascertain whether sufficient evidence exists to prove that a suspect has taken property from the victim without consent and with the intent to permanently retain said property.

Three different scenarios often surface:

Scenario One—The victim is able to testify competently that the defendant took money or property without consent. The victim may be asked to look at a check and confirm that the signature is a forgery or that a series of ATM transactions were not authorized.

Scenario Two—The victim is incompetent to testify because of advanced dementia. The transaction took place at a time when the victim was already diagnosed with that advanced condition. A treating physician for the victim will need to lay the appropriate medical testimony foundation to establish that the victim would have been unable to provide the necessary consent at the time of the transaction.

Scenario Three—The victim is marginally competent to testify, although there may be some signs of short-term memory loss. The victim appears to have voluntarily transferred assets to the suspect and appears to understand the nature and extent of the transaction. The suspect may describe the transfer as either a gift or a loan. Law enforcement may have refused to investigate such a scenario on the basis that no apparent crime has been committed. The scenario may be further complicated by the fact that the victim, prior to the transaction, was taken to meet with an attorney (unsurprisingly, chosen by the suspect) to discuss the impending transfer. That attorney will likely testify that the client understood the significance of the transaction and wished to proceed. This last scenario is the most difficult one to prove as a theft. The suspect has clearly manipulated the elderly victim to bring about the intended result. Lawyers refer to that conduct as "undue influence."

Historically, the concept of undue influence has been used widely in civil actions such as will contests or contract disputes. But only a few states recognize undue influence as a criminal concept. As a result, creativity is required in order to introduce evidence of manipulation that will demonstrate that the transaction was completed without the victim's consent. Often a psychiatrist can provide valuable testimony—either as an expert witness to explain undue influence, or as a percipient witness to discuss the victim's susceptibility to manipulation.

Consent is the cornerstone of many financial elder abuse cases. In California, the trial judge in a criminal financial elder abuse case can instruct the jury with the following definition of consent:

To consent to a transaction a person must:

- Act freely and voluntarily and not under the influence of threats, force, or duress
- Have knowledge of the true nature of the act or transaction involved

- Possess the mental capacity to make an intelligent choice whether or not to do something proposed by another person

The judge will then conclude: "Consent requires a free will and positive cooperation in act or attitude."

In grappling with the question of whether an elder was susceptible to undue influence, certain factors should be weighed:

1. The length of the relationship between the victim and the suspect. The longer the friendship, the less likely it is that fraud can be established.
2. The nature of the relationship between the victim and the suspect. If the suspect performed certain services for the victim without any other evident remuneration, an assessment must be made to determine whether the amount of the transfer was a reasonable exchange for services rendered.
3. The spending habits of the victim prior to the transfer. If it can be established that the victim had a history of frugal spending, it might be possible to show that this particular transaction was completely out of character.
4. The amount of the transfer compared to the residual left for the victim. Sometimes, the victim is left with insufficient funds to provide for unexpected emergencies after the transaction has been completed.
5. The documentation that accompanied the transaction. If the suspect maintains that the monies were given in the form of a loan, then it might be expected that supporting paperwork evidencing a loan were drafted at the same time.
6. The question of whether the lawyer who prepared the documentation was introduced by the suspect to the victim.
7. The issue of whether the victim demonstrated any changes in behavior prior to the transaction that might point to signs of short-term memory loss, confusion, or a decrease in the ability to make wise decisions.
8. The methods used by the suspect to develop the friendship with the victim. Did the suspect create a false persona to be more interesting to the elder? Did the suspect gradually isolate the victim from other friendships and family, causing the victim to become increasingly dependent upon the suspect? Did the suspect ply the victim with alcohol to wear down the victim's defenses?

EXCUSES AND JUSTIFICATIONS IN ELDER ABUSE

The psychiatrist should be familiar with common reactions to accusations of elder abuse. The alleged perpetrator may first deny that there was abuse. Since elder abuse requires some combination of actual harm, intended harm, and bad motive, the alleged perpetrator may argue that any of these factors is absent. But if confronted with evidence of proven abuse, the perpetrator may employ some excuse or justification.

In an excuse, the perpetrator admits that abuse happened and further admits that such abuse should not have happened, but nevertheless has an excuse like:

- "it was an accident" (with no intent to perform the action), or
- "it was a mistake" (the action was indeed intended, but executed in some faulty manner), or
- "the victim provoked me," or
- "I was impaired" (tired, intoxicated, overwhelmed), or
- "Most nursing homes treat their elders just as I did, so I was following the norm."

In contrast, if a perpetrator invokes a justification, he alleges that while the abuse was regrettable; given the circumstances, it was the right thing to do:

- "I did provide substandard care (abuse), but if I had demanded more resources for better care, I would have been fired. If that were to happen, then the nursing home would not have any doctor, so the residents would be without any medical coverage."
- "Since some medical coverage (even if it is substandard) is better than no medical coverage; given the circumstances, I was justified in doing what I did."

ETHICAL ANALYSIS OF SUSPECTED ELDER ABUSE

1. The difference between threats and offers–

most people think threats are ethically bad, while offers are ethically acceptable, but the distinction may be difficult to see:

Vignette 1

facts:

a child says to parents:

"either – leave all your assets to me in the will,

or ———I will put you in the cheapest nursing home, since you cannot take care of yourself."

analysis:

this is a threat, and it seems the child is exploiting the parents' lack of capacity for self-care, and this may qualify as elder psychological abuse

Vignette 2

facts:

a child says to parents:

"if you leave all your assets to me in the will,

then—I will not put you in a nursing home, but buy a larger home, and you can live with me in your old age."

analysis:

this is an offer, thereby not exploitative, and likely does not qualify as

elder psychological abuse

In both cases, the parents are invited to consider a deal, and both deals are substantively very similar. But in, one case it is manipulative, while in the other it is not.

2. The difference between shoves and nudges–

An advisor or caregiver should presumably respect the autonomy of the elder; violating such autonomy is unethical. "Shoving" an elder into a decision (forcing a decision onto an unwilling elder) is clearly unethical, but the ethics of "nudging" (where the elder retains autonomy) remains controversial.

Competent elders should be masters of their own agency, but examples where such agency can be changed, altered, or influenced includes:

- Reframing Information

An elder's medical interest would be promoted by taking a first-generation antipsychotic, but the elder refuses such treatment. The patient will only consent to meds that help him "relax." The doctor mentions "Mellaril" and when the patient asks if that will cause him to feel "mellow," the doctor reports affirmatively, and the patient consents to Mellaril (to feel mellow). Since the patient did exercise autonomy, there was no shove. The doctor nudged the patient, by the reframing of information, which was not a lie (Mellaril does have a sedating action, which is "mellowing"). This nudge benefits the recipient of the nudge.

- Overcoming Queasiness

An adult child knows his parents' views on postmortem organ donation and therefore encourages them to do the paper work, but they keep avoiding it, out of queasiness. Finally, the child employs a nudge to get this done. Such a nudge benefits neither the nudger, nor the nudgee, but rather, society as a whole.

- Overcoming Procrastination

An adult child wants his parents to write their will so that he will get his inheritance easily, rather than be forced to go through probate. The elderly parents keep postponing this task, out of procrastination. Finally, the child employs a nudge to get this done. All the benefit of such an action (writing the will) goes to the child, not to the elderly parent (who will be dead).

CONCLUSION

Over the last few years, elder abuse has become one of the most common crimes in the United States. It is important to prepare now to meet the challenge by developing an expertise so that the abuse can be identified, and the abusers held accountable. The responsibility to watch out for the senior population is a collective one. Psychiatrists will play an increasingly

important role in the fight against elder abuse over the next few years. Becoming an active member of a local MDT that develops protocols to combat elder abuse is an effective first step.

CASES

Profile of a Physical Abuser of the Elderly

Very often the abuser of an elder is the son of a widowed woman. He is either divorced and has come back to live at home with his mother because of financial pressures borne by the divorce or he is single, and has never left home. He tends to be in his late-30s to mid-40s, while his mother is usually in her mid-to-late 70s. In the majority of cases, the son is lazy and unemployed. He will have used a variety of excuses as to why he is unable to work, and his mother has accepted such excuses. He will be addicted to alcohol (normally beer), narcotics, or gambling. In order to feed his habit, the son extracts money from his mother. There may come a time when the mother refuses to provide any more money, and this tends to be the flash point for violence. Violence may begin with him breaking objects in the home that he knows the mother values. An escalation of violence may materialize as grabbing his mother's arms and shaking her or pushing her into furniture or throwing an object at her. Characteristically, the mother will not immediately call the police, perhaps out of a sense of failure or shame. Often, neighbors hear the altercations between the son and his mother, but they may also be reluctant to involve the police.

Profile of a Financial Abuser of the Elderly

The profile of the financial abuser is more difficult to describe. The exploiter can be a family member who has been distant for many years; then after the death of a spouse, the elusive relative surfaces to befriend the widow or widower in the hope of acquiring some or all of the anticipated inheritance. Alternatively, the abuser can be a trusted individual such as the victim's attorney, realtor, accountant, pastor, or physician. Ultimately many financial abusers of an elder will rely upon portraying the transfer of assets as either a "gift" or a "loan." For that reason many alleged thefts from elders are mistakenly judged to be "civil" violations, rather than criminal.

REFERENCES

California Office of the Attorney General. Bureau of Medi-Cal Fraud and Elder Abuse. Available at http://ag.ca.gov/bmfea/eldr.php. Accessed January 22, 2009.

Dubble C. A policy perspective on elder justice through APS and law enforcement collaboration. *J Gerontol Soc Work*. 2006;46:35–55.

Dyer CB, Pavlik VN, Murphy KP, Hyman DJ. The high prevalence of depression and dementia in elder abuse or neglect. *J Am Geriatr Soc*. 2000;48(2):205–208.

HelpGuide.Org, Elder Abuse. Types, signs, symptoms, risk factors, and prevention, Available at http://www.helpguide.org/mentalleldr_abuse-JhysicaLemotionaLsexuaLneglect.htm. Accessed February 10, 2009.

Kahan FS, Paris BEC. Why elder abuse continues to elude the health care system. *Mt Sinai J Med*. 2003;70(1):62–68.

Koenig RJ 3rd, DeGuerre CR. The legal and governmental response to domestic elder abuse. *Clin Geriatr Med*. 2005;21(2):383–398.

National Committee for the Prevention of Elder Abuse: Best practice models. Available at http://www.preventelderabuse.org/communities/best.html. Accessed February 10, 2009.

National Research Council. *Elder Mistreatment: Abuse, Neglect, and Exploitation in an Aging America*. Washington, DC: National Academies Press; 2003.

Redding.com. "Police warn of new telephone scam." aan. 28, 2009), Available at http://www.redding.com/news/2009/jan/28Ipolice-warn-new-telephone-scam/. Accessed February l0, 2009.

Sacramento Bee. "Murder charges filed against 5 in elderly swindle." Available at http://www.sacbee.com/state_wire/story/1721848.html. Accessed March 23, 2009.

Survey of State Adult Protective Services. Abuse of adults 60 Years of age and older, The National Committee for the Prevention of Elder Abuse and the National Adult Protective Services Association, prepared for The National Center on Elder Abuse (2006). Available at http://www.ncea.aoa.gov/NCEArootiMain_Sitelpdfl2-14-06%20FINAL%2060+REPORT.pdf. Accessed January 22, 2009.

57.6f Sexuality and Aging

ABESH KUMAR BHATTACHARJEE, M.D., PH.D., AND DANIEL D. SEWELL, M.D.

An estimated 10,000 Americans turn 65 years old everyday and are joining the pool of about 40 million individuals who are presently over 65. As a result, by 2030, the number of individuals in the United States who are 65 years old or older is expected to grow to 72 million. The fastest growing segment of the population is people over 80. Issues affecting older people, therefore, are becoming increasingly more important. Traditionally, aging has been viewed as a period of progressive decline in physical, psychosocial, and cognitive domains, which consequently increased the health care burden. This negative view of old age contrasts starkly with the recently ascendant concept of *healthy aging* or *successful aging*, which is supported by existing research on older adults who continue to function well and to be active, productive, and engaged in life. Self-rated *successful aging* indicates that it is a subjective perception of an individual's quality of life, which comprises physical, social, and spiritual dimensions. Sexual activity and enjoyment are considered components of these quality of life dimensions, particularly the physical and psychosocial dimensions.

Regardless of how long a person lives, experiencing human closeness and intimacy remains important. Individuals and societies react in a variety of ways to sexual intimacy among older people. Despite the rapid growth of the older population and the aggressive marketing of products aimed at sexual health in later life, sexuality in older adults has received relatively limited research attention. A review of cross-cultural data from 106 traditional societies determined that sexual behavior in older men and women was very common though not invariably observed. For example, 84 percent of the studies, which included information about sexual behavior in older females, documented continued sexual activity and interest among older females. The authors of this review concluded that the continued sexual activity among older members of many traditional societies and the cessation of sexual behavior among older members of other traditional societies suggested that cultural, as well as biological factors, might be important determinants of sexual expression in later life in these societies.

Cultural and biological factors also appear to be important determinants of sexual behavior in the industrialized world where many still believe that older people stop having sex or believe that it is dangerous for older people to be sexually intimate. For some, certain cultural taboos from the Victorian era may, even today, inhibit sexual behavior, and provoke both intrapsychic and interpersonal conflict. Older people may internalize the myth of sex in older age being "wrong" or "inappropriate" and some may choose to repress their sexual urges rather than become objects of derision or disapproval (e.g., being considered "a dirty old man"

or, in the case of women, “a cougar”).

Fortunately, over the past 50 years, attitudes about later-life sexuality have been evolving. Contrary to persistent stereotypes, recent studies have indicated that many older men and women remain sexually active, experience sexual desire, and regard sexual activity as an important aspect of their lives. Moreover, multiple aspects of sexual health in older adults are positively associated with self-rated health and indicators of physical function. The Women’s Health Initiative Study found that activity and satisfaction with sexuality were associated with higher self-perceived successful aging and quality of life. The acceptance of hormone replacement therapy for menopausal and postmenopausal women, the increasing popularity of the oral erectogenic agents, and the incorporation of “ED” into the common vernacular, all indicate changing attitudes about sexuality in later life.

It is possible to live a happy and fulfilling life without sex. A clinician who is knowledgeable about sexuality in later life and willing to discuss sexuality with older patients has the challenge of creating an open, safe environment for discussion without implying that any problem necessarily exists. Ironically, a therapist who addresses these issues must be aware that the introduction of the topic may be perceived as very helpful or may unintentionally trigger feelings of inadequacy or incompleteness in patients for whom sex never was or is no longer a central part of life. Sexual disinterest is a matter of concern only if a person finds it troubling or if it causes problems in relating to others. All things considered, it is better to ask with sensitivity about sexual behavior, listen carefully to determine if sexual problems exist, and to remember that whatever reaction ensues will serve as an opportunity for a deeper understanding to develop between the patient and the clinician.

In order to help older individuals who are experiencing a discrepancy between their desire for sexual intimacy and their actual experience of it, knowledge about the following may be useful: (1) the neurobiology of sexual function; (2) the effects of aging on sexual function; (3) the recommended guidelines for how to take a sexual history; (4) the types of sexual dysfunction, their causes, and their recommended treatments; and (5) the impact of institutional living on sexual expression.

NEUROBIOLOGY OF SEXUAL FUNCTION

The neuroendocrine perspective of sexual behavior can be summarized in a five-component model focused on neurological and hormonal mechanisms that lead to behavioral and motivational changes during human development. These components of sexuality are sensorimotor (prenatal and neonatal spinal reflexes, such as penile or clitoral erections), valence or salience (sensory stimuli developed in late juvenile period), attractiveness (a wide range of behaviors intended to increase sexual attractiveness),

receptivity (the willingness to engage in sexual activity), and solicitation or proceptivity (the urge to seek out partners). With reproductive maturity, all become tightly coordinated until later in life when these components may become uncoupled again because of different rates of change associated with aging.

The neuroendocrine model acquires a new level of nuance and complexity when considered through the lens of sociological and psychological conceptualization of sexual motivation. Historically, the scientific literature tends to consider sexual interest and motivation in terms of desire and arousal and places sexual desire in a nested framework of contexts, from dyadic relationships, to communities, to cultures. Such a perspective on sexual motivation shares much with the interactive biophysical model of health, which posits that health is a function of biophysical and psychocognitive dynamics between individuals over time.

Various models of sexual arousal and response have been proposed over the past century. Masters and Johnson described four sequential phases to the "sexual response cycle": excitement, plateau, orgasm, and resolution. Kaplan proposed a triphasic model of sexual response, which highlighted the interdependence of the three components (desire, excitement, and orgasm) and proved to be clinically helpful because the phases correlate well with various common sexual problems. For example, low libido or hypersexuality are problems with phase I or the desire phase, erectile dysfunction (ED) or dyspareunia are problems with phase II or the arousal phase, and anorgasmia or premature ejaculation are problems with phase III or the orgasm phase. The *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5) uses these phases as a basis for classifying the sexual dysfunctions. One of the more recent refinements of the sexual response cycle has been the addition of the concepts of "spontaneous desire," which is believed to be more typical of men, and "stimulus-driven desire" which is believed to be more significant for women and refers to desire aroused by stimuli such as the behavior or appearance of the partner. Hypoactive sexual desire disorder is significant among women. Additional information about the neurobiology of the sexual response cycle can be found elsewhere in this text.

AGE-RELATED CHANGES IN SEXUAL FUNCTION AND BEHAVIOR

Aging may have both positive and negative impacts on sexual expression. Normal age–associated changes in sexual function are summarized in Table 57.6f–1. For women during and after menopause, lowered estrogen levels cause atrophy of the vaginal epithelium, labia majora, and Bartholin glands, as well as shortening and loss of elasticity of the vaginal canal. Reductions in vaginal blood flow, genital engorgement, and vaginal lubrication lead to delay in achieving sexual arousal. Loss of lubrication may also contribute to dyspareunia. While orgasmic capacity is retained

with age, decreased muscle tone may decrease sensitivity, orgasm may require a longer period of direct stimulation, there may be reductions in the number and intensity of vaginal contractions and some women may experience painful uterine contractions during orgasm.

Testosterone in the central nervous system influences female sexual behavior and women experience declines in androgen levels with age, especially between the third and fifth decades of life. There is not yet consistent evidence demonstrating a direct relationship between reduced testosterone levels and change in sexual functioning in women but, in general, lowered levels are thought to contribute to decreases in libido, arousal, responsiveness, genital sensation, and orgasm. Some recent research suggests that both estrogens and androgens may work together to enhance sexual arousal and response in women. For example, the administration of estrogen plus small amounts of testosterone to menopausal women may provide greater improvement of psychological symptoms such as sense of well-being, depression and fatigue as well as of sexual problems related to the effects of estrogen deprivation on the genitalia, decreased libido and anorgasmia. Perhaps even more than in men, however, variability of sexual response in women is likely to be contextual, partner based and relatively independent of biological control systems. In general, for women the response to sexual stimuli is largely influenced by the quality of emotional intimacy with the partner. Frequently, the emotional and relationship well-being a woman experiences contribute more to sexual enjoyment than the physiological response.

Table 57.6f–1.
Normal Age–Associated Changes in Sexual Function

Function	Male Changes	Female Changes
Sex hormones	Decline in testosterone levels	Cessation of estrogen production Decline in testosterone levels
Genitals	Fewer and less functional sperm Decreased penile blood flow Decreased penile smooth muscle relaxation	Cessation of ovulation Decreased pelvic blood supply Shortening and narrowing of vagina Atrophy of vaginal mucosa, labia majora, Bartholin glands Decreased vaginal elasticity Decreased muscle tone
Fertility	Reduced rates of conception	Ceases after menopause
Libido	For some, may decline due to lower testosterone	For some, may decline due to lower testosterone
Arousal	Requires more tactile stimulation Erections take longer to achieve Erections are less firm Erections more difficult to sustain	Requires more tactile stimulation Decreased vaginal swelling/lubrication Decreased clitoral sensitivity

Orgasm	Takes longer to achieve Ejaculation is less forceful Decreased ejaculate volume	Takes longer to achieve Fewer, less forceful vaginal contractions
Refractory period	Lengthened by hours to days	Lengthened

In men, testosterone is necessary for normal libido and optimal erectile function. In men, as in women, testosterone levels also progressively decline with age, particularly beginning around the fifth decade of life, while levels of sex hormone–binding globulin increase. These changes result in more bound testosterone and less free or physiologically available testosterone and are believed to be responsible for the gradual decline in sexual responsiveness in men as they age. Specifically, psychic and tactile stimuli are less effective and full penile erections occur more slowly. Maintenance of erections also requires more direct genital stimulation. The force and volume of ejaculation may diminish, ejaculation may not even occur with every sexual encounter but difficulties with rapid ejaculation tend to lessen with increasing age. Other age-related changes include possibly less intense orgasms, retrograde ejaculation, more rapid penile detumesence, and longer refractory periods.

Despite the clear role that testosterone plays in male sexual behavior in later life, the relationship between them still remains unclear. For example, the observation that men with a deficiency of testosterone may be capable of an erectile response demonstrates some independence of sexual arousal from testosterone levels and underscores that a lack of sexual interest in men does not result only from a lack of testosterone. Many psychological and relationship factors may contribute to libido in men such as the perceived level of attractiveness of the partner, various feelings toward the partner such as resentment or hostility, issues related to maintaining control or feeling controlled, and difficulties dealing with sexual dysfunction in himself or his partner.

Until relatively recently, surveys of the frequency of sexual function in later life have, in general, been limited by methodological shortcomings such as relatively small, nonrandom subject samples of convenience and questions which narrowly define sex as coitus. In addition, the studies of the effects of aging mainly involve white, middle class, well-educated predominantly heterosexual populations, so there is currently much less information about the role that race, culture, sexual orientation, gender identity, education, and socioeconomic conditions play in determining sexuality. According to several major surveys, 50 to 90 percent of men and women over 60 continue to be sexually active, usually defined as having sexual intercourse at least once a month.

Perhaps the two best recent explorations of this issue have come from the National Social Life, Health, and Aging Project (NSHAP) which provided researchers with data from a nationally representative probability

sample of 3,005 community dwelling older US adults in wave 1 of this study (1,550 women and 1,455 men) who were 57 to 85 years old, defined sex or sexual activity as "any mutually voluntary activity with another person that involves sexual contact, whether or not intercourse or orgasm occurs," and included in-home interviews conducted in English and Spanish by professional interviewers. Analysis of this data revealed that sexual activity over the past 12 months declined with age. Overall, men were more likely than women to be sexually active, report a good quality of life, and be interested in sex. These gender differences increased with age and were greater among the 75- to 85-year-old group. For men, 84.4 percent for those 57 to 64, 67.4 percent for those 65 to 74 and 38.9 percent for those who were 75 to 85 years old were sexually active whereas for women 61.9 percent for those 57 to 64, 39.8 percent for those 65 to 74 and 16.8 percent for those who were 75 to 85 years old were sexually active. The wave 2 of this study focused on measuring sexual interest, behavior, sexual experience, frequency, and appeal of physical contact in 3,377 community dwellers who were 62 to 90 years old, and of whom 73.7 percent (n = 2,487) had a spouse or cohabiting partner. It is important to note that 54.8 percent (n = 1,365) engaged in touching genital organs of their partners. Fifty-nine individuals who were 80 to 90 years old had sexual intercourse in the last 12 months, compared with 383 individuals who were 62 to 69 years old. In general, the measures for the salience/valence and receptivity (terms defined in an earlier paragraph under neurobiology of sexual function: neuroendocrine model) components decrease with age. Other noteworthy findings from this study included: the inclination to make one's self sexually attractive to others is stronger in women, two-fifths of the men and women reported dissatisfaction with their partner's frequency of caring behavior that makes later sexual behavior pleasurable, a fifth of women and quarter of men had vaginal sex in the past year and physical contact is frequent among those living with partners, whereas it is rare among those living alone.

Perhaps as the result of biological factors, social conditioning, or some combination of both, some older people were not especially interested in sex when they were younger and, even with the increased acceptance of later-life sexuality, growing older has not increased their interest. For other older individuals, sexual intimacy has become associated with painful emotional conflict resulting from or causing difficulties with their partners. One member of a couple may have a significantly different level of interest or functional ability than the other. For these individuals, the stereotype of "sexless old age" may provide a comfortable shelter from what are experienced as vexing and/or insurmountable obstacles to the achievement of sexual satisfaction. Still other older individuals may have simply grown tired of sex and have replaced sexual intimacy with other satisfying nonsexual activities. Some older individuals become celibate as the result

of a deliberate decision to share sex only with a particular partner who may have become too ill or disabled to have sex or who may have died. The sexual behavior of some older people may be influenced by a religious belief that sex is solely for procreation and, therefore, ends with the completion of menopause. Although it is important to be aware of older people's sexuality, clinicians must be careful neither to oversexualize the aging process, nor to overmedicalize declining sexual function and interest.

Sexually active older adults need and benefit from sex education. Table 57.6f–2 includes strategies for enhancing sexual function in later life. Older adults also need education about HIV and other sexually transmitted diseases (STDs). HIV infection and acquired immunodeficiency syndrome (AIDS) are no longer problems associated only with those who are members of high-risk groups such as gay men, sex workers, and intravenous drug users. Older individuals have also become part of the epidemic although, fortunately, from 2008 through 2012 the estimated annual numbers and rates of HIV/AIDS diagnoses in 65 years old and older aged groups remained stable in the United States. In 2008, according to the Center for Disease Control (CDC), among persons 65 years old and older in the United States, the estimated number of diagnoses of HIV infection was 896 (rate = 2.4 per 100,000 persons) and persons living with HIV infection ever classified as stage 3 (AIDS) was 748 (rate = 2 per 100,000 persons). In 2012, among persons 65 years old and older in the United States, the estimated number diagnosed with HIV infection was 792 (rate 2.1 per 100,000 persons) and stage 3 (AIDS) was 673 (rate = 1.8 per 100,000 persons). In terms of racial and ethnic minority differences, individuals who self-identify as Black/African American or White continue to rank highest among racial and ethnic minorities in terms of the number diagnosed with HIV infection and with AIDS.

Table 57.6f–2.
Strategies for Enhancing Sexual Function in Later Life

Optimize your physical health
Maintain a heart healthy diet
Get sufficient rest
Have your hearing evaluated and corrected if necessary
Learn and practice good dental hygiene which includes flossing and seeing a dentist on a regular basis
Take good care of your skin
Avoid tobacco products
Minimize alcohol intake
Engage in regular physical exercise
Learn and perform specific exercises that can improve your appearance and function such as crunches and knee raises for a protruding abdomen and Kegel exercises for weakened pelvic muscles
Obtain optimal treatment of any medical illnesses
Inform the clinicians who prescribe medications for you that you wish to avoid medications which are likely to cause problems with sexual function

Pretreat symptoms of pain or dyspnea that interfere with sex
Use lubrication to counteract vaginal dryness and sensitivity

Optimize your psychological health

Foster a positive attitude toward sexuality in later life
Put aside unrealistic expectations of sexual performance

Optimize the health of your relationship with your sexual partner

Maintain open, honest communication
Identify and discuss changes in sexual function
Discuss ways to adapt sexual practices to these changes
Focus extra time and energy on sensual foreplay
Select sexual practices and positions that account for declines in strength and endurance

HOMOSEXUALITY AND OTHER SEXUAL MINORITIES IN LATER LIFE

While it was once thought that gay people represented 10 percent of the adult US population, more reliable recent data estimate the prevalence of homosexuality as between 1 and 5 percent. The 2013 National Health Interview Survey (NHIS) was the first national survey to include a sexual orientation question. In this survey, 2.3 percent of survey participants characterized themselves as LGBT, a relatively new term designed to include bisexual and transgender persons under one umbrella with those who identify as gay or lesbian because all of these individuals are members of sexual minorities and, as a result, remain at risk for misunderstanding, rejection, discrimination, and violence due to their sexuality. The use of this term represented an effort by the surveyors to use the most up-to-date vernacular in order to obtain the most reliable and trustworthy data regarding the number of individuals who belong to a sexual minority.

The NHIS included interviews of a nationally representative sample of 34,557 adults aged 18 and older. The NIHS found that 1.6 percent of adults identified as gay and lesbian and 0.7 percent identified as bisexual. A higher percentage of women than men identified as bisexual (0.9 percent compared with 0.4 percent). Among people 65 years old and older, 0.7 percent identified as gay and lesbian and 0.2 percent identified as bisexual. It is important to be aware of the possible role that methodology has on research findings in this area. For example, the sexual orientation/identity question consistently yield a lower percentage of people saying that they were LGBT, on average about 3.5 percent, then the percentage obtained when people were asked if they engage in same-sex behavior, which was around 8 percent. In addition, the question about same-sex attraction yielded the highest response rate of around 11 percent.

In an ongoing effort to include all sexual minorities under one umbrella due to their shared experiences, the term LGBT continues to evolve and a more recent version of this term is LGBTQ where the “Q” stands for questioning or queer. When the “Q” is used to denote questioning, it refers to individuals who remain uncertain of their sexual orientation. In some instances the “Q” is used as a shorthand for “queer” which historically has

been a pejorative term for gay, but which is now being reclaimed by some gay men, lesbians, bisexuals, and transgendered persons as a self-affirming umbrella term. Keeping the geriatric theme of this chapter in mind, the authors recommend caution when using the term "queer" because this term may still be experienced as extremely offensive when used as an epithet, especially among older members of a sexual minority group.

The current older generations of the LGBTQ community grew up during a time of pervasive homophobia, heterosexism, and associated discrimination and hate crimes. Before the turn of the 20th century, few LGBTQ individuals had any sense of connection to others like themselves or acceptance by society at large. Most married opposite sex partners had either lived double lives or suppressed all homosexual sexual inclinations. Not until the Stonewall riots in 1969 gave birth to the modern gay liberation movement, did the LGBTQ community begin to develop a sense of solidarity and systematically begin to pursue greater social acceptance. As a result, many older LGBTQ individuals have endured major traumas personally, politically, and socially. Many older LGBTQ couples are still living completely or partially in the closet. The complicated challenges of the AIDS crisis have in some ways both helped and hindered efforts to increase acceptance and have diverted both attention and resources away from other important issues which impact the LGBTQ community such as cancer research.

Contrary to the long held societal belief that older members of the LGBTQ community are destined to be single, lonely, and unhappy, many have long-term relationships, are emotionally stable, and may be just as likely (or as unlikely) to be successful and happy in their later years as heterosexual individuals. The majority of the potential challenges faced by older LGBTQ couples are identical to those faced by older heterosexual couples but a few of the challenges may be different. Older LGBTQ individuals may find themselves socially isolated with few or no role models for growing older. An older member of the LGBTQ community may experience a lack of support from friends, relatives, neighbors, coworkers, as well as various forms of discrimination from hospitals, various levels of government and other institutions such as insurance companies. This isolation and unequal treatment can become particularly difficult when a partner is ill or dies. For example, some hospitals do not recognize LGBTQ relationships in terms of visitation rights and consultation with medical professionals. At the time of this writing, 37 states, plus the District of Columbia, now allow same-sex marriage, and 13 states prohibit same-sex marriage or civil unions. Seven out of ten US residents now reside where same-sex marriages are legal. The state-by-state trend toward legal recognition of same sex marriage, as well as changes in certain institutional and federal policies, have helped eliminate forms of discrimination and unequal treatment in a variety of situations such as hospital visitation,

income tax benefits, insurance, retirement benefits, and inheritance but the legal rights of LGBTQ couples in states where same sex marriage is not legal, generally remain unclear and unprotected. In January 2015, the US Supreme Court announced its intention to issue a ruling on this issue.

A number of organizations now exist which may provide support and guidance to more senior members of the LGBTQ community. The existence of these organizations may be especially important for the older individual who begins the process of *coming out* when older. *Coming out of the closet*, or simply *coming out*, is a figure of speech for self-disclosure of belonging to a sexual orientation or gender identity minority group. Framed and debated as a privacy issue, coming out of the closet is a multifaceted experience in which each person may have shared unique experiences. These organizations include: the Senior Action in the Gay Environment (SAGE) which offers a variety of social services to older gays and lesbians and has independent affiliates in various cities in the United States and Canada; the Lesbian and Gay Aging Issues Network (LGAIN), which is sponsored by the American Society on Aging, and the National Association of Lesbian and Gay Gerontologists which has created a resource guide for older gays and lesbians. In addition, LGBTQ community centered around the country often sponsor or support social organizations for older members of the LGBTQ community. For example, in San Diego, California the LGBTQ community center supports a social organization for older members of the LGBTQ community called Fellowship of Older Gays (FOG).

HOW TO OBTAIN A SEXUAL HISTORY

Knowing how to take a sexual history is the first step in helping an older person with their sexual quality of life. Many patients think that all psychiatrists are knowledgeable, skillful, and interested in addressing sexual concerns. Unfortunately, education in human sexuality and on how to take a sexual history is not always emphasized in medical school curriculums and residency training programs. Though the ability to take a sexual history is not especially difficult to master, doing so may provoke certain fears such as: (1) the clinician or patient experiencing arousal while talking about sex; (2) not knowing what questions to ask or how to help with a patient's sexual problems; (3) unexpectedly becoming aware of one's own sexual concerns; or (4) having the patient see one's moral repugnance or personal distaste about certain sexual practices. These apprehensions will be reduced by: (1) identifying them in yourself; (2) thinking about them rationally; (3) learning about the broad range of sexual expression; (4) understanding professional boundaries; and (5) acquiring more knowledge and experience regarding the diagnosis and treatment of sexual dysfunctions.

Recent research has found that psychiatrists are less likely to obtain a sexual history from older patients presenting with depressive symptoms

compared with younger patients and they are also less likely to refer to appropriate services if sexual dysfunction is identified. On the other hand, many older people are reluctant to seek help for asexual problems even if they have a significant effect on quality of life. The evaluation of the sexual functioning of an older individual is similar to the evaluation of a younger adult, although for the younger clinician it may be more difficult due to ageist beliefs or fears of appearing disrespectful. General considerations include: (1) establishing rapport which may require that this topic not be explored until at least the second or third contact; (2) introducing the topic by putting it in context (e.g., sexual behavior is both an indicator of health and a contributor to health); (3) accepting that the first few times you introduce the topic, you may feel uncomfortable and may even blush but not letting this stop you; (4) being aware of your own potential prejudices and subconscious feelings; (5) not waiting for the patient or patient's partner to introduce the topic because they may assume that you are only interested in "medical problems" or may view sexual problems as trifling (partner's caring for a partner with dementia may be especially prone to make this assumption because of fears that asking about may portray them as selfish or as inadequate caregivers); (6) beginning with a neutral, nonjudgmental statement like "have you experienced any changes in your sexual life?" or, when the older individual is being assessed for depression, "sometimes when people feel very low and depressed, they lose all interest in sex. Do you think this is an issue for you?"; (7) using open-ended questions and active-listening techniques; (8) demonstrating sensitivity to the patient's values and beliefs; (9) establishing a private and comfortable environment; (10) allowing sufficient time for discussion which means bringing up the issue no later than the midpoint of the time period set aside for the meeting; (11) not making assumptions like the person is heterosexual (less than 10 percent of older persons are questioned about their sexual orientation) or that the person has a sexual problem; (12) being clear that sexual activity includes both coital and noncoital behaviors; (13) providing information about HIV and other STDs including the fact that condom use not only prevents conception but also prevents transmission of STDs; (14) educating patients regarding age-related changes that may affect sexuality, like "often women around the time of the menopause can suffer not only with the hot flushes you have described, but also with sexual problems such as vaginal dryness. Is that something you have experienced?" or "people at your age may experience ED. Is that something you have experienced?" and, if clinically indicated, discussing noncoital methods of achieving sexual intimacy; (15) possibly using humor. (Note: A phenomenological study with experienced nurses in long-term care found that the use of humor was an effective method of addressing the sexual needs of older adults.)

HOW TO ASSESS SEXUAL DYSFUNCTIONS IN OLDER PATIENTS

Because sexual dysfunction in later life can result from psychological, psychiatric, medical or social causes, or from some combination of these factors, an optimal evaluation should cover of all of these areas. For example, psychological factors that impact sexual function in later life include decreased self-esteem or a decreased sense of sexual attractiveness due to physical changes which may be natural changes associated with aging or changes brought about by illness or the treatment of illness such as a mastectomy for breast cancer. Sexual dysfunction may also develop when a member of a couple experiences an illness such as a depression, dementia, or stroke, which precludes emotional intimacy. Social issues may also trigger sexual dysfunction. The lifespan of men in the United States is shorter than that of women (life expectancy in 2010 for US males was 76.2 years and for US females was 81.0 years) which often leaves older women without sexual partners.

In the large recent survey of community dwelling older adults referred to above, among respondents who were sexually active, about half of both men and women reported at least one troubling sexual problem. The most prevalent sexual problems among women were low desire (43 percent), difficulty with vaginal lubrication (39 percent), and inability to climax (34 percent). Among men, the most prevalent problems were erectile difficulties (37 percent), performance anxiety (27 percent), and climaxing too quickly (26 percent). Rating one's health as poor was associated with decreased likelihood of being sexually active and, among those sexually active, being more likely to report sexual problems. The prevalence of discussing sex with a physician since age 50 for men and women was 38 and 22 percent, respectively.

The following are some guidelines on how to optimally assess sexual dysfunction in an older patient. When a sexual problem involves both the person with the dysfunction and his or her partner, then inclusion of both individuals in the assessment is essential. Often each person will have a somewhat different version of the problem. To initiate the interview, the clinician should ask the identified patient to describe the problem and then ask the partner "What do you think about this report?" Such open-ended questions can be followed up by more specific questions designed to pinpoint which phases of the sexual response cycle (desire, arousal, or orgasm) are involved for the patient and the patient's sexual partner. Illnesses which cause pain, malaise, or actual physiological or anatomical limitations may impact any of the phases of the sexual response cycle.

After the joint interview, each person should be interviewed individually. For widowed patients who are in second relationships, it may be difficult to acknowledge in front of the new partner that guilt regarding the deceased partner is contributing to the sexual problem. In fact, the loss of a partner may be the biggest barrier to sexuality for older individuals,

especially for older women. Certain societal taboos, though waning, may still prevent a surviving partner from seeking a new mate. Even when a new mate is found, an older person may experience sexual difficulties as a result of guilt arising from a perception of being unfaithful to a former partner and anxiety about overcoming old patterns of physical intimacy and establishing new ones. Masters and Johnson called this "the widower's syndrome."

It is essential to learn about the role of sexuality in each person's life and about past sexual experiences. Potentially useful questions include: what would be the optimal frequency? have there been any problems before? what has changed in your relationship? and what is the meaning or value of intimacy? Some individuals may view emotional intimacy as an essential precursor to quality physical or sexual intimacy while for others the primary value of sex is to assuage libido.

Elucidating the time course of the development of the sexual problem may help uncover underlying etiologic factors. For example, if the problem's appearance coincided with the onset of an illness of the initiation of a particular medication, then a causal connection should be considered. Asking a question like "sometimes a person taking these medications notices sexual problems; is this something that has affected you at all?" may be helpful. Many medications, including medications with psychiatric indications like selective-serotonin reuptake inhibitors (SSRIs), have the potential to impact one or more aspects of sexual function. An abrupt onset of a change in sexual function temporally related to the introduction of a medication strongly suggests that this medication may be the cause.

In general, a gradual onset suggests an etiology related to chronic medical illness. For example, does osteoarthritic pain limit movement or physical position? does cardiovascular disease reduce physical stamina? are physical changes associated with aging or illness provoking embarrassment in the person or impacting sexual interest in the partner? is a previously unrecognized or suboptimally treated psychiatric or substance use disorder present? and what is the patient's current sexual ability? It may be helpful to elicit a detailed history of the patient's most recent previous sexual experience. The ability to achieve orgasm during masturbation suggests that the sexual problem may have a psychological origin.

THE TREATMENT OF SEXUAL DYSFUNCTION IN OLDER ADULTS

Just as in a younger adult, sexual dysfunction in an older patient may cause significant distress. The clinician must integrate all of the pertinent information into a compassionate and reasonable treatment plan. If there is a contributing psychological problem, individual and/or couples psychotherapy should be provided. If the sexual problem has a relevant physical or biological component, then education and medications may be helpful.

The treatment of sexual dysfunctions may be viewed through the lens of the sexual response cycle. The first phase of the sexual response cycle is "desire." Desire may be influenced by psychological, psychiatric, or medical problems. For example, illnesses whose optimal treatment requires surgery may result in physical disfigurement (e.g., mastectomy), which, in turn, triggers low self-esteem and feelings of being no longer sexually attractive or desirable. Psychiatric illnesses, which may decrease desire include major depression and alcoholism. Mania is often associated with hypersexuality and the various forms of dementia may lead to increases or decreases in libido.

A variety of medical illnesses have the potential to reduce libido. One common pathway for certain medical illnesses to trigger decreased libido is through reduction in androgen levels. Illnesses associated with lowered androgen levels include chronic renal failure, chronic pulmonary disease, hepatic cirrhosis, hyperthyroidism, and HIV. A number of medications such as anticonvulsants and tamoxifen may also lead to lowered testosterone levels. In 2013, the Endocrine Society convened a task force which updated and published online a clinical practice guideline for treatment with testosterone in adult men. These guidelines recommend making a diagnosis of androgen deficiency only in men with consistent symptoms and signs and unequivocally low serum testosterone levels. The measurement of morning total testosterone level should be repeated and in some patients, measurement of free or bioavailable testosterone level is necessary. Initiation of testosterone therapy is recommended for men with symptomatic androgen deficiency to induce and maintain secondary sex characteristics and to improve their sexual function, sense of well-being, muscle mass and strength, and bone mineral density. (Table 57.6f–3summarizes the symptoms and signs of testosterone deficiency for both men and women). Testosterone therapy should not be initiated in patients with breast or prostate cancer, a palpable prostate nodule or induration or prostate-specific antigen greater than 4 ng/mL or greater than 3 ng/mL in men at high risk for prostate cancer such as African American men or men with first-degree relatives with a history of prostate cancer without further urological evaluation, a hematocrit greater than 50 percent, untreated severe obstructive sleep apnea, severe lower urinary tract symptoms with International Prostate Symptom Score above 19, or uncontrolled or poorly controlled heart failure. Men receiving testosterone therapy should be monitored regularly. The task force members who created this guideline favored a cutoff for men of levels less than 300 ng/mL while others favored 200 ng/mL. Finally, another unresolved controversy is the method by which testosterone levels should be measured. The gold standard is equilibrium dialysis but this test is expensive, labor intensive, and not widely available. Most laboratories measure free testosterone by the analog method, which often yields less reliable results.

Table 57.6f–3.
Signs and Symptoms of Low Testosterone

Men
Erectile dysfunction
Infertility
Women
Impaired arousal
Decreased genital sensitivity
Both
Breast discomfort
Decreased libido
Decreased muscle strength
Depression
Hot flushes, sweats
Inability to concentrate
Lethargy
Mild anemia
Sleep discomfort

Based on currently available data, it is clear that sex steroids influence female sexual function but the measurement of androgen levels in women presenting with low sexual function is not helpful and should not be used for the purpose of diagnosing androgen insufficiency in women. Although a recent report suggests that many women following oophorectomy have no decline in sexual function, a subset of women with lowered androgen may experience lowered libido and there is some evidence that testosterone therapy in this setting may be helpful. Whether testosterone treatment is appropriate for women who have normal testosterone levels and lowered libido but no other signs or symptoms of androgen deficiency has not been established. Currently, there are no approved medications for female hyposexual desire disorder.

The Endocrine Society Clinical Practice Guideline reappraised androgen therapy in women after the Task Force established by the Endocrine Society commissioned two systematic reviews of published data and reviewed several other existing meta-analyses. The Task Force recommended against: (1) making a diagnosis of androgen-deficiency syndrome in healthy women because there is a lack of a well-defined syndrome and data-correlating androgen levels with specific signs or symptoms were unavailable; (2) the general use of testosterone for the following indications: infertility; sexual dysfunction other than hypoactive sexual desire disorder; cognitive, cardiovascular, metabolic, or bone health; or general well-being; (3) the routine use of dehydroepiandrosterone due to limited data concerning its effectiveness and safety in normal women or those with adrenal insufficiency; and (4) the routine prescription of testosterone or dehydroepiandrosterone for the treatment of women with

low androgen levels due to hypopituitarism, adrenal insufficiency, surgical menopause, pharmacological glucocorticoid administration, or other conditions associated with low androgen levels because there are limited data supporting improvement in signs and symptoms with therapy and no long-term studies of risk. Evidence supports the short-term efficacy and safety of high physiological doses of testosterone for postmenopausal women with sexual dysfunction due to hypoactive sexual desire disorder. Importantly, endogenous testosterone levels did not predict response to therapy. At present, physiological testosterone preparations for use in women are not available in many countries including the United States and long-term safety data are lacking. Any woman receiving testosterone therapy should be monitored for signs and symptoms of androgen excess.

Sexual problems of the arousal phase of the sexual response cycle include ED, vaginismus, and dyspareunia. ED is defined as the inability to achieve and/or maintain an erection satisfactory for sexual function. Studies have found that the prevalence of ED is significantly higher among older men as compared to those in younger age groups. The Massachusetts Male Aging Study found a tripling of the prevalence of complete ED among men aged 70 years as compared to those aged 40 years. Compared with men in their 20s, those in the oldest age group (70 to 75 years) had a 14-fold higher relative risk for ED. In this study, age was found to be the strongest predictor of ED. Other age-adjusted risk factors significantly associated with ED included hypertension, diabetes, heart disease, obesity, smoking, hypercholesterolemia, lack of physical exercise, poor health, depressed mood, and prostate disease.

There are now a handful of medications that may be helpful to men with sexual dysfunctions involving the arousal or orgasm phases of the sexual response cycle. Phosphodiesterase V inhibitors act peripherally on the penile vascular system to improve tumescence. These medications increase cyclic GMP in the corpus cavernosum, which relaxes smooth muscle in the penile vasculature and fosters blood flow. Three medications in this category are now available: sildenafil, vardenafil, and tadalafil which essentially only differ in their half-lives. Other agents for ED include transurethral and injectable prostaglandin E1, which are less convenient to administer but may appeal to some men who value the immediate and predictable response and who show poor response or intolerance to phosphodiesterase V inhibitors. Alprostadil, a synthetic form of prostaglandin E1, is currently used and it has a unique role in men with ED secondary to diabetes and secondary to radical pelvic surgery (e.g., radical prostatectomy). Based on a variety of factors including effectiveness and patient preference, vacuum pump devices are an alternative treatment for ED. Both manual and battery-operated devices are available and, if determined to be medically necessary, will be paid for by Medicare. Surgical implantation of a penile prosthesis is a third option for treatment

of ED and is generally considered only after the less invasive interventions have proven unsuccessful.

Premature ejaculation and anorgasmia are sexual dysfunctions of the orgasm phase of the sexual response cycle. Increasing age often diminishes, or sometimes even resolves, problems with premature ejaculation. Finally, there are data that bupropion may have some value for improving sexual function. In one study, both normal subjects and patients who took bupropion reported improvements in desire, arousal, and orgasm and these effects appeared to be independent of antidepressant efficacy.

The definitions of female sexual dysfunction have recently been revised in the DSM-5, with merging of hypoactive sexual desire disorder and female sexual arousal disorder into female sexual interest/arousal disorder. There are no FDA-approved medications for female arousal disorders or orgasmic disorders. The value of phosphodiesterase inhibitors for women with sexual dysfunction remains unclear. An open-label nonrandomized study found that sildenafil 50 mg was well tolerated in postmenopausal women with sexual dysfunction and was associated statistically nonsignificant improvement in vaginal lubrication, clitoral sensitivity, and orgasm frequency and overall sexual desire or satisfaction with sexual intercourse did not improve. A controlled study of sildenafil 25 or 50 mg as needed found that both doses resulted in greater arousal, more frequent orgasms, and greater sexual satisfaction compared with placebo. Bupropion may also be useful for sexual dysfunction in women. One small study (n = 20), found that women had significant improvement relative to placebo in all aspects of sexual function including intensity of orgasm on either bupropion-SR 150 mg or 300 mg daily. It is important to note that several small-scale studies have assessed the effects of the above-mentioned drugs. The results are still conflicting and the drugs are not devoid of adverse effects. Currently ongoing phase I and II clinical trials for drugs to treat for female sexual dysfunction include topical alprostadil, bremelanotide (BMT), intranasal testosterone (TBS-2), intravaginal dehydroepiandrosterone (DHEA), sublingual testosterone with sildenafil, apomorphine, bupropion and trazodone.

The treatment of sexual dysfunction can also be viewed through the lens of specific illnesses. Almost any illness may cause sexual dysfunction. Some of these illnesses may be gender specific, for example, “senile” vaginitis and prostate cancer, and others may impact both genders. Those most often associated with sexual dysfunction are illnesses which negatively impact physical stamina or cause pain, those that require medications that adversely impact sexual function, or those that impact the endocrine system, the reproductive system, or the nervous system. Table 57.6f–4 lists some of the illnesses, which may cause sexual dysfunction in later life. Given that the prevalence of serious medical illness increases with age, it is not surprising that sexual dysfunctions are more common in older patients.

In a patient with chronic renal disease, sexual function may be diminished due to fatigue and malaise. Chronic renal disease may also lead to sexual dysfunction due to associated vascular disease, neuropathies, and endocrine abnormalities. For example, both free and total testosterone levels are lowered in chronic renal disease but may return to normal following transplantation. In patients with chronic renal disease, sexual dysfunction may also be related to abnormalities between the pituitary and the Leydig axis. Both vitamin E and erythropoietin have been reported to improve sexual function in men with renal failure.

Table 57.6f–4.
Some of the Medical and Psychiatric Conditions Associated with Sexual Dysfunction in Later Life

Diseases that negatively impact physical stamina or cause pain
AIDS
Chronic fatigue
Chronic obstructive pulmonary disease
Congestive heart failure
Coronary artery disease
Degenerative joint disease
Fibromyalgia
Irritable bowel syndrome
Peripheral vascular disease
Renal failure
Diseases that require medications that adversely impact sexual function
Cancer
Hypertension
Disease of the endocrine system
Diabetes Mellitus
Hypercortisolism (Cushing syndrome)
Hyperpituitarism (acromegaly)
Hypopituitarism
Hyperthyroidism
Hypothyroidism
Reproductive system diseases
Breast cancer
Cervical cancer
Ovarian cancer
Prostate cancer
Testicular cancer
Uterine cancer
Diseases of the nervous system
Alzheimer dementia
Epilepsy
Generalized anxiety disorder
Major depressive disorder
Multiple sclerosis

Obsessive compulsive disorder
Panic disorder
Parkinson disease
Schizophrenia
Substance use disorders

Studies have found that cardiovascular disease including myocardial infarction (MI) may physically and psychologically interfere with sexual functioning and negatively impacts all three phases of the sexual response cycle in older women as well as older men. Following an MI, many patients have concerns about intercourse that include anxiety about inability to perform or fear that one may die during coitus. Angina may contribute to fears of having an MI during sex and reduce libido. In addition, the anxiety and depressive disorders that commonly occur after an MI may themselves cause sexual dysfunction.

Sexual intercourse places roughly the same demands on the heart as a vigorous walk around a city block. There is great intersubject variability in blood pressure and heart rate increases during sexual intercourse. Generally, an individual who can walk vigorously or pedal 600 kpm/min on a stationary bicycle without symptoms, increased blood pressure, or electrocardiogram changes can safely engage in sexual intimacy including coitus. Sexual abstinence is recommended for 8 to 12 weeks after MI, although there is minimal evidence to justify this in uncomplicated patients. After this period of abstinence, gradual resumption of sexual activity is recommended, often beginning with sexual experiences other than sexual intercourse. Because superior positions may heighten cardiac demand, side-by-side or patient inferior positions may be recommended when coitus is initially resumed. Prior to intercourse, in selected patients the use of beta blockers and sublingual nitrates may reduce angina. Patients who benefit from the use of nitrates must be instructed not to use phosphodiesterase inhibitors.

Chronic pulmonary disease impacts sexual function in at least two ways: chronic arterial hypoxia is believed to cause testosterone reductions and associated diminished desire in both men and women and shortness of breath may limit physical ability to engage in sexual activity.

Both insulin-dependent (type I) and noninsulin-dependent (type II) diabetes mellitus (DM) may contribute to sexual dysfunction in men and in women. ED has been reported in up to 50 percent of men with type I DM and in up to 33 percent of men with type II DM. As men living with DM age, the rate of ED may approach 70 percent and comorbid smoking and alcohol abuse predict an earlier onset of ED. Men living with DM are believed to develop ED secondary to neurologic or microvascular lesions and, as a result, the presence of any of the “opathies” will be associated with ED. Sexual dysfunction in women with DM is often overlooked but may

occur at a rate as high as 50 percent. Dyspareunia secondary to decreased vaginal lubrication and difficulty achieving orgasm are among the problems reported by women living with DM.

A patient living with cancer may experience sexual dysfunction due to psychological factors, physical problems directly related to the cancer including pain, and side effects of the treatments. Women diagnosed with breast cancer may experience grieving and fears of death and disfigurement that decrease libido even before any treatment has been provided. Women whose treatment includes mastectomy and who have a history of marital difficulties and who do not discuss with their spouse the possibility of mastectomy are at increased risk for sexual problems. A recent study compared women who had a lumpectomy to women who had a mastectomy and found that having a lumpectomy was associated with a greater sense of sexual desirability but no difference in frequency of sexual experiences or of orgasms.

The treatment of sexual dysfunction following mastectomy includes minimizing medical problems that are inhibiting sexuality. Medical issues that contribute to sexual dysfunction in women undergoing treatment for breast cancer include dyspareunia and other forms of pain, fatigue, ovarian dysfunction from chemotherapeutic agents such as tamoxifen, and nausea and vomiting associated with chemotherapy. If possible, open communication between the patient and her sexual partner is helpful during all stages of treatment. It is helpful for the partner to observe the mastectomy wound soon after surgery. Rapid resumption of sexual activities is also believed to be helpful and alternative sexual positions and vaginal lubrication may be helpful. Finally, breast reconstruction may help improve self-image and restore self-esteem both of which may contribute to improved sexual functioning.

Although testicular cancer is relatively uncommon in older men (incidence is highest at about 35 years of age, then declines to nearly zero by the age of 80), staging surgery may result in ejaculatory dysfunction and ED. Even in the absence of these problems, surgical removal of one or both testicles may cause problems with self-esteem. Prostate cancer increases in frequency with age. The relative success of nonsurgical interventions and the introduction of surgical techniques which are nerve sparing have helped to decrease ED and other sexual problems in men being treated for prostate cancer. Patients may need to wait 2 to 3 years following surgery for maximum sexual function to return. One study reported that radiation therapy caused ED in over 50 percent of the men who were studied.

As is true for male genital malignancies and for breast cancer in women, female genitourinary cancers are often associated with serious emotional problems which impact sexual function. Sexual difficulties with cervical carcinoma vary with treatment approach. In general, surgery has fewer side effects than radiation therapy. Radical hysterectomy compromises sexual

function because of vaginal shortening but this may improve when there is ongoing sexual activity. The most important variable predicting sexual dysfunction following hysterectomy is whether the ovaries are removed. Radiation therapy promotes vaginal atrophy and stenosis due to impaired ovarian function. The extent of surgical treatment for vulvar carcinoma often determines sexual side effects. From a medical perspective, local excision often does not interfere with sexuality. More extensive surgeries may cause numbness.

Sexual problems in patients diagnosed with colon cancer are usually related to chemotherapy-associated side effects or to surgeries resulting in ostomies. Fear of fecal spillage and odor may inhibit sexual enjoyment. Women with ostomies can develop dyspareunia due to fistula formation. Men who require posterior abdominal perineal resection may become impotent due to ablation of autonomic nerves to the penis.

The desire or need to express one's sexuality does not expire with age and for many older people including those living in aged care facility, sexuality continues to be important. A variety of outcomes have been observed when a spouse or significant other has dementia. Research indicates that 73 to 80 percent of the time, the presence of dementia has a negative impact on a couple's sex life. Conversely, about one-quarter of the couples that have been studied continue a warm and satisfying sexual relationship. A significant number of sexually inactive caregiving partners (39 percent in one study) are dissatisfied with the inactivity. Clearly, there is no one right answer for everyone. Ballard and Poer have written that in the absence of a cure or treatment, the most significant thing that can be done for a patient with dementia is to adapt one's approaches with each change in symptoms to make life as meaningful and satisfying as possible. Though written with the patient in mind, this quote and underlying sentiment seems equally relevant to caregiving spouses or partners. For patients with dementia and/or their caregiving partners, the opportunity to discuss their sexual concerns with a health care professional can be tremendously helpful. According to research, however, health care professionals only rarely inquire about the impact that the dementia is having on a couple's sex life.

Among the factors that may influence whether a couple continues to have satisfying sexual experiences are: (1) the type of dementia; (2) the stage of the dementia; (3) the couple's previously established patterns of sexual intimacy; (4) the impact of dementia on the patient's ability to maintain two-sided intimacy; (5) the impact of the dementia on the patient's knowledge of how to engage in and to follow through with varies sexual activities and on the sexual activities may be preferred or even demanded; (6) the influence of the dementia on the patient's personality, libido, grooming, and personal hygiene; (7) the impact of the dementia on the caregiving partner's view of the patient; (8) the effect of caregiving on

the spouse or sexual partner of the patient; and (9) whether the patient with dementia is able to continue living at home or requires institutionalization.

In regard to the partner with dementia, possible changes include remaining interested but developing problems with follow through or with apraxia for one or more sexual acts, becoming hypersexual, or making new or unusual requests regarding sexual activities. At some point, the erosion in intellectual function may make it difficult or impossible to accurately assess the sexual health of the relationship and to provide emotional intimacy.

For a variety of reasons, a caregiving spouse or partner may lose interest in sex. This loss of interest may be the result of overwhelming grief or the physical exhaustion from the demands of caregiving. Sometimes caregiving partners deny their own needs including their sexual needs to assuage survivor's guilt or as a way of showing their respect for and loyalty to the person with dementia. It is also possible that dementia may cause such drastic changes in the patient's values and personality that the caregiving partner sees the patient as a child or in other ways so different that sexual intimacy begins to feel wrong or uncomfortable. Caregiving partners may sometimes need to be reminded that they do not have to participate in an activity that is frightening, upsetting, or no longer satisfying. A caregiving partner has the right not to be abused. Due to the level of the patient's cognitive impairment, a caregiving spouse may wonder if having sex with the partner is a form of rape. Ballard and Poer suggest that if the dementia patient initiates the experience and if the experience is mutually satisfying to both, then it should not be viewed as rape. It is quite possible that the same patient who might be incapable of giving informed consent for medication or research, because of the complexity and novelty of information given, might be capable of giving informed consent to participate in a sexual relationship, a decision that is based more on old learning.

Patients with Alzheimer dementia will usually experience dramatic changes in short-term memory relatively early in the course of illness but personality and social graces do not erode until relatively late in the illness. The preservation of personality and social skills may allow a couple living with Alzheimer dementia to maintain satisfying sexual experiences through most of the illness as long as ED is not problematic. One of the earliest studies of sexual function in patients with Alzheimer dementia found that loss of erection was reported in 53 percent of male patients around the time that they learned of their Alzheimer dementia diagnosis. In addition to ED, another potential situation that may negatively impact the quality of sexual life of a couple living with dementia is when short-term memory deficits trigger delusions of infidelity because the patient with dementia has trouble remembering where the caregiving spouse has gone or how long the

partner has been absent and then begins to suspect that the partner is having an affair.

Patients with frontotemporal dementia (FTD) very often experience personality changes or profound language problems fairly early in the course of their dementia and either of these can negatively impact sexual quality of life. For example, personality changes may cause the patient with dementia to behave so differently that their partner begins to feel as if they are with a stranger. Sometimes the expressive or receptive aphasia seen in patients with FTD is so severe that it is no longer possible to maintain two-side intimacy that for some partners is an essential ingredient for safe and satisfying sexual experiences.

The later the stage of the dementia illness, the more difficult it becomes to maintain sexual intimacy. One study of 20 Caucasian male Alzheimer patients and their nondemented caregiving wives found that the patients with dementia reported no reduction in quality of sexual function regardless of the stage of illness. The wives of these patients, however, reported that the more advanced the stage of the dementia, the worse their quality of sexual life became. The life expectancy of patients with many forms of dementia may exceed 12 years or more. For some healthy, active spouses, this may become an unacceptably long time to live without satisfying emotional and sexual experiences. Certainly, to give and receive affection, physical intimacy, are needs that do not die with the overwhelming responsibilities of becoming a caregiver. As a result, some caregiving spouses may begin a new intimate relationship while continuing to provide for the needs of the ill partner. When this decision is made, the adult relatives of the patient with dementia may have a variety of reactions. Some family members may be very sympathetic. Others may fail to understand the well partner's need for an intimate relationship. Other family members may feel the need to protect the ill partner's interests or may have fears that the well partner's new alliance will threaten financial assets of the ill partner or of their own anticipated inheritance. Sometimes the most difficult aspect of the well partner's decision to begin a new intimate relationship is that this decision makes it impossible for other family members to deny the reality of the tragic illness that has befallen the patient with dementia.

Dementia may also include the development of inappropriate sexual behaviors and the development of behaviors that appear sexual but are actually not. Caregivers of patients with dementia, whether a partner, some other family member, or a paid professional, may experience distress, discomfort, and uncertainty about how to manage inappropriate sexual behaviors. Professionals should assess these behaviors with the assumption that there is a valid need to be met keeping in mind that not all behaviors that appear sexual may actually be sexual. For example, touching the genital area repeatedly may not represent masturbatory efforts but may be

a signal that help with toileting is needed. Climbing in bed with the patient in the room next door may not be an act of sexual aggression but the expression of a desire for comfort, companionship, or protection or simply the expression of a lifelong habit of having a bed partner.

A variety of classes of medications and specific individual medications may cause sexual dysfunction in adults of all ages including older adults. Table 57.6f–5 provides a list of some of the medications, which have the potential to cause sexual side effects. A number of classes of psychotropic medications can negatively impact sexual function. Men may experience ED from serotonin reuptake inhibitors (SSRIs). Both men and women may experience decreased sensitivity of the genitalia, delayed orgasm, or anorgasmia from SSRIs. Any of the psychotropic medications with significant dopamine receptor blocking ability may also cause a variety of sexual problems which are believed to be mediated, at least in part, through medication-associated hyperprolactinemia and associated decreased androgen levels. Drugs, which increase gamma amino butyric acid (GABA) activity, such as benzodiazepine, have been reported to cause less satisfying orgasms for women.

Antihypertensive medications may also impair sexual function although recent research suggests that the beta-adrenergic blockers may not be as problematic as originally believed. Digoxin may decrease sexual desire. Agents which block the histamine receptor subtype 2 (H2 blockers) frequently interfere with penile erections although not all medications in this class may be do this equivalently. For example, cimetidine is more likely than ranitidine to cause ED. Lastly, alcohol, though not a medication, may seriously impair sexual function.

Table 57.6f–5.
Some of the Medications Associated with Sexual Dysfunction in Later Life

Analgesics
Methadone
Morphine sulfate
Anti-infective medications
Ketoconazole
Antineoplastic medications
Tamoxifen
Cardiovascular medications
Amiodarone (Cordarone)
Angiotensin-converting enzyme inhibitors
Calcium channel blockers
Clonidine (Catapress)
Digoxin (Lanoxin)
Disopyramide (Norpace)
Guanethidine (Ismelin)
Prazosin (Minipress)

Reserpine
Spironolactone
Thiazide diuretics

Endocrine medications

Adrenal corticosteroids
Leuprolide (Eligard, Lupron)

Gastrointestinal medications

Cimetidine (Tagamet)

Neurologic and psychiatric medications

Antipsychotic medications
Benzodiazepines
Carbamazepine (Carbatrol, Equetro, Tegretol)
Carbidopa/levodopa (Sinemet)
Divalproex sodium (Depakote)
Lithium carbonate (Eskalith, Lithobid)
Monoamine oxidase inhibitors
Phenytoin (Dilantin)
Primidone (Mysoline)
Selective-serotonin reuptake inhibitors
Trazodone
Tricyclic antidepressants
Venlafaxine (Effexor)

Other

Diphenhydramine (Benadryl)
Cocaine
Hydroxyzine (Atarax, Vistaril)
Nicotine
Alcohol

SEXUALITY IN LONG-TERM CARE

The challenges and difficulties for older people with respect to the expression of their sexuality in the residential care environment have been identified and described in the literature and include negative and judgmental staff attitudes and inadequate staff member knowledge and training. Older members of sexual minorities are especially at risk for challenges and difficulties with comfortably meeting their sexual needs in a residential care environment. A survey of 250 nursing home residents in Texas determined that only 8 percent of them had been sexually active (masturbation or intercourse) during the previous month but 17 percent of the inactive residents said they would have liked to be active but lacked the opportunity.

Although addressing the continuing sexual needs of older residents of long-term care facilities can be a difficult task to accomplish, these should be addressed with the same priority as nutrition, hydration, personal hygiene, and other well-accepted needs. Sexual behavior in residential care facilities is no longer a private matter and is often open to public scrutiny.

In one way or another, sexual behavior in residential care facilities has an impact on staff members, other residents, and the families of residents and is, therefore, much more often inhibited by other residents, staff, and family. Residents of long-term care facilities may experience numerous barriers to achieving sexual satisfaction. Perhaps the biggest of these barriers is the discomfort nurses, physicians, and other health care providers may have in discussing sexuality with older adults. Eighteen states in the United States address the issue of married couples sharing rooms or having conjugal visits unless there is any medical contradiction (Federal regulation 483.10). Only four states allow room for sexual relationship between persons of same gender, as well as persons of the opposite gender who are not married. Such recognition seems to be important to assure that nursing home residents have the same rights as other citizens in this regard. Although there is an increasing emphasis on the application of person-centered approaches to care delivery, many providers of residential care services for seniors still neglect to recognize and address sexuality as a component of well-being. Many are unaware of the importance of the multiple facets of the care environment that impact on resident's ability to express their sexuality and are unaware of what sort of support may be required.

Another barrier to satisfying sexual experiences for long-term care residents is the lack of knowledge and understanding among the residents themselves about later-life sexuality. Studies have found that this lack of knowledge in older adults correlates with less positive attitudes toward sexuality such as embarrassment and apprehension. Finally, family members may also be a barrier to sexual satisfaction for older adults living in a long-term care environment because they may have difficulty understanding and accepting that their older relative is having a sexual relationship with someone new.

For residents of long-term care facilities, emotional intimacy may supplant sexual intimacy and provide similar levels of satisfaction. In one survey of nursing home residents, coital intimacy was rated as the lowest in importance when compared to social, noncoital physical, intellectual, and emotional intimacy. Unmet intimacy needs of older adults may be fulfilled by a simple human touch on the hand, face, or shoulder. It is very important, however, before offering touch to a patient to ask permission and to confirm that the patient accurately understands the intentions of the clinician. Some older patients, especially those with dementia, may misinterpret the touch as a sexual overture.

If a patient living in the long-term care environment expresses a desire for sex, the long-term care facility staff should help arrange a safe and private time and place for the sexual activity and the resident should be advised to keep his or her call light within reach. If necessary, supportive items such as personal lubricant or adaptive equipment such as hospital

beds, side rails, or trapezes may be needed.

Determining whether two patients, both of whom have dementia, should be allowed to share sexual intimacy may provoke feelings of anxiety or discomfort but this has and will continue to be required of health care professionals. Lichtenberg and Strzepek have published guidelines that are based on the principles of informed consent. They recommend assigning one member of the interprofessional treatment team to assess each patient according to three criteria and then to report their findings to the entire team, which then votes on whether the criteria have been met for the patients to have sex. Lichtenberg and Strzepek recommend assessing each patient's awareness of: (1) the relationship (e.g., is the patient aware of who is initiating the sexual contact or is the person being misidentified as someone else such as a deceased spouse?); (2) the potential risks (can the patient describe how he or she will react if the relationship ends?); and (3) the ability to avoid exploitation (does the patient have the ability to state what level of sexual intimacy is desired and to say no to any uninvited or unwanted sexual act?). Although not a part of these guidelines, it may be wise to inform a family member of each patient about the approach that has or will be used to address the situation.

Inappropriate sexual behavior among long-term care residents include using sexually suggestive language, exposing oneself, masturbating, or openly participating in sexual activity with another resident. Whether in the long-term care environment or at home, the approach to sexually inappropriate behaviors is the same. If the problem sexual behaviors are new, a careful evaluation should take place. In order to avoid precipitating a "catastrophic reaction," it is important not to scold, argue, or shame the person. Praising and encouraging acceptable behaviors and ignoring inappropriate behaviors may be helpful. Remaining calm helps to avoid rewarding unwanted behaviors. Sometimes providing attention and affection proactively may prevent or eliminate problem sexual behaviors.

Depending on the specific behavior, different solutions should be explored. For example, for the patient who repeatedly masturbates in public places, providing privacy as well as other things like a vibrator, personal lubricant or condom may be helpful. If a patient disrobes in public, dressing the patient in clothing that makes disrobing difficult may help. For patients who expose themselves unnecessarily during personal care or become sexually aroused, it is important to set limits verbally, involve more than one caregiver, and depending on the patient's sexual orientation, use caregivers whose gender is not the gender which stimulates the patient, and make the interactions as clinical as possible by wearing medical attire like a white coat and gloves, and performing these activities in an examination room rather than in the resident's room.

If behavioral interventions are unsuccessful, a medication trial may be needed. Table 57.6f–6 lists medications that may be helpful including

various forms of estrogen and SSRIs. Mood stabilizers may be useful if hypersexuality is occurring in the context of a manic-like syndrome. Antipsychotic medication may be helpful if the sexual behavior is driven by a delusion of mistaken identity in which the patient is attracted to a person misidentified as the patient's sexual partner.

Below the authors present a case example of a patient who received medication treatment for inappropriate hypersexual behaviors. This case highlights a number of key issues associated with sexually inappropriate behaviors in patients living with dementia: (1) although these behaviors can arise in patients living with a variety of types of dementia, they are most often seen in patients whose dementia has preferentially injured the frontal lobes resulting in problems with judgment and impulse modulation; (2) these behaviors are more likely to occur in later stages of dementia, resulting in sexually inappropriate behaviors occurring in the midst of other types of behavioral challenges; (3) wide variation in both sex hormone levels and their impact on physical characteristics and behavior often limits their value in establishing the etiology of sexually inappropriate behaviors (this patient's measured serum testosterone level was actually low); (4) medoxyprogesterone acetate (MPA) may help reduce but may not eliminate sexually inappropriate behaviors completely; and (5) the use of MPA requires careful monitoring due to potential side effects that limit its usefulness and safety.

Table 57.6f–6.
Medications for Hypersexuality in Patients with Dementia

Medication	Preparation	Suggested Dosage	Potential Side Effects
Hormonal agents			
Antiandrogens			
Medroxyprogesterone acetate (Depo-Provera)	IM injection 150 mg/mL	150–300 mg once a week	Sedation, weight gain, fatigue, hot and cold, flashes, depression
Estrogens (Premarin)			
	0.3-, 0.45-, 0.625-, 0.9-, 1.25-mg tablets	0.625–2.5 mg/day	Increased risk of cardiac death in men with prostate cancer, fluid retention, nausea, vomiting, gynecomastia
(Enjuvia)	0.3-, 0.45-, 0.625-, 1.25-mg tablets		
(Climara)	0.025-, 0.0375-, 0.05-0.06-, 0.075-, 0.10-mg/day transdermal patch	0.05–0.1 mg/day	

Gonadotrophin-releasing hormone analogs			
Leuprolide acetate	Injectable suspensions:	7.5 mg/month	Hot flashes, erectile dysfunction, irritation at injection site
(Eligard)	7.5, 22.5, 30, 45 mg/month		
(Lupron Depot)	3.75, 7.5 mg/month 11.25 mg every 3 months		
Serotinergic agents			
Serotonin reuptake inhibitors			
Citalopram	10 mg/5 mL oral solution, 10-, 20-, 40-mg tablets	20 mg/day	Headache, nausea, diarrhea, loss of appetite, insomnia
Paroxetine	10 mg/5 mL oral suspension 10-, 20-, 30-, 40-mg tablets 12.5-, 25-, 37.5-mg controlled-release tablets	20 mg daily	
Tricyclic antidepressants			
Clomipramine	25-, 50-, 75-mg tablets	150–200 mg/day	Orthostatic hypotension
Trazodone hydrochloride	50-, 100-, 150-, 300-mg tablets	100–500 mg/day	Sedation, hypotension, priapism
Antipsychyotics			
Atypical antipsychotics			
Quetiapine	25-, 100-, 200-, 300-mg tablets	25 mg/day	Orthostasis, sedation, increased death due to cardiovascular events
Mood stabilizers			
Carbamazepine	100-mg chewable tablet 200-mg tablet 100-, 200-, 400-mg extended-release tablets 100-, 200-, 300-mg capsule 100 mg/5 mL and 200 mg/10 mL suspensions	50–300 mg/day	Ataxia, dizziness, rash, sedation, nausea, vomiting, hyponatremia, agranulocytosis, aplastic anemia, falls

Valproate	250-mg capsules 125-mg sprinkle capsule 250-, 500-mg extended-release tablets 250 mg/5 mL syrup 100-mg/mL injection	500–1,500 mg/day	Nausea, sedation, falls, tremor
Other agents			
Pindolol	5-, 10-mg tablets	40 mg/day	Fatigue, hypotension
Cimetidine	200-, 300-, 400-, 800-mg tablets 300-mg/L oral solution	600–1,600 mg/day	Headaches, nausea, arthralgias, delirium

The primary care physician (PCP) of a 91-year-old widowed male, retired engineer with vascular dementia residing in a dementia-specific residential unit, requested a geriatric psychiatry consultation to evaluate depression and inappropriate sexual behaviors. Because of the severity of the patient's cognitive deficits, most of the history was obtained from collateral sources.

Prior to the emergence of the patient's dementia symptoms, the patient had no psychiatric problems or psychiatric treatment of any kind. His cognitive deficits were first recognized several years prior to the time of the consultation request following a stroke which left the patient with residual ataxia and pseudobulbar affect. About 5 months after his stroke, behavioral symptoms including wandering and elopement attempts required the patient to move from an independent living apartment to a dementia-specific assisted living unit. Several months later, a geriatric psychiatry consultation was obtained after the patient slapped the hand of a male peer because this peer had touched the patient's food. Following this incident, the patient told a nurse "if it happens again, I'll break his hand off." The consulting geriatric psychiatrist (CGP) found no undiagnosed medical factors that might have caused the problem behaviors and observed that the patient was struggling with frustration related to both loss of control and problems with memory. The patient stated "If I see a telephone number on the TV that I want to remember, I don't remember that number at all." The CGP recommended discontinuation of quetiapine, given its questionable efficacy and unfavorable risk–benefit profile in patients living with dementia, and initiation of treatment with dextromethorphane/quinidine to target the patient's emotional lability and impulsive behaviors. In addition, the CGP suggested that, if needed, a trial of divalproex might also be reasonable.

Several months later, when sexually inappropriate behaviors emerged, the patient's PCP independently changed the patient's psychiatric medications to escitalopram 10 mg po daily and divalproex delayed release 250 mg po twice daily (which yielded a trough serum level of 30 mcg/dL) and encouraged the staff to use a variety of behavior-modifying techniques including avoidance of the use of ambiguous terms of endearment such as "honey" or "sweetie" and responding to inappropriate sexual behaviors with simple, firm verbal limit setting in order to avoid unintentionally reinforcing these behaviors with excessive positive (laughter) or negative (anger) affect. The patient, however, continued to have frequent and disruptive inappropriate sexual behaviors involving staff members including verbalizing requests for sex, kissing and groping staff members and masturbating in public.

The patient had no known past psychiatric history of hypomanic or manic episodes. The

patient's medical problems included congestive heart failure, DM type 2, hypothyroidism, benign prostatic hypertrophy, a history of sinus surgery, and a stroke 22 months previously. His medications were acetaminophen 500 mg po TID, bumetanide 0.5 mg po BID, clopidogrel 75 mg po daily, divalproex 250 mg po BID, doxazosin mesylate 8 mg po QPM, escitalopram 10 mg po QD, lansoprazole 30 mg po QD, levothyroxine 100 mcg po QD, metoprolol 12.5 mg po BID, probiotic containing saccharomyces boulardii lyo 250 mg po BID, sitagliptin 100 mg po QD, and vitamin B_{12} 1,000 mcg po QD.

On examination, the patient's vital signs were within normal limits. The CGP examined the patient in the common area and observed that he was a casually dressed and groomed Caucasian man who was seated in a wheelchair and had no observed abnormal movements. He was distant, detached and disinterested. He reported feeling "pretty normal," denied any current symptoms of depression or anxiety but was unable to describe his current medications other than saying that he was taking "a bunch" and that he was not aware of any side effects. His affect revealed frustration and irritability, his speech was normal, his process of thought was linear and the content of his speech was noteworthy for concreteness, sarcastic responses to most questions and the absence of clear answers to questions about suicidal ideation, homicidal ideation, delusions, or hallucinations. He was oriented to self and attentive but refused to cooperative with a standardized bedside cognitive screening test.

None of the patient's medications were felt likely to be triggering the patient's sexually inappropriate behaviors and the results of a wide range of laboratory tests obtained by his PCP 1 week prior did not reveal a medical problem that would explain the patient's inappropriate sexual behaviors. These lab results included a random glucose of 123 mg/dL (high), creatinine 1.64 mg/dL (high), albumin 2.9 g/dL (low), total protein 5.3 g/dL (low), hemoglobin 8.6 g/dL (low), hematocrit 25.7 percent (low), TSH 3.09 μIU/dL, prolactin 14.3 μg/L, FSH 15.5 IU/L (high), LH 7.5 IU/L, progesterone 0.5 ng/mL, testosterone 187 ng/dL (low). After obtaining informed consent from the patient's family, a trial of medroxyprogesterone acetate (MPA) 150 mg IM Q monthly was initiated. No other changes in his medications were made in order to clearly assess the impact of the addition of MPA.

One month later, the CGP learned from staff members that the patient was no longer propositioning or attempting to kiss or grope members of the staff although he was masturbating in his room occasionally. When interviewed, the patient reported feeling "in and out" and of being "pushed around" by "people like you (the psychiatrist)." The CGP concluded that the patient was demonstrating a favorable response to the treatment with MPA and recommended continuation of MPA along with ongoing use of nonpharmacologic measures. Two months after initiation of treatment with MPA, the patient had not had any problematic sexual behaviors but had been spending most of the day sleeping, had become apathetic about meals, would eat if spoon-fed and had lost 17 pounds over the preceding 3 months. When interviewed, the patient was oppositional and antagonistic, denied depression and anxiety but reported feeling angry due to "people like you." When the CGP asked the patient to elaborate, he replied "You had better get your ass out of here and never come back." Due to concerns about potential divalproex-associated nausea and weight loss as well as absence of any clear benefit, divalproex was discontinued and a trial of mirtazapine 7.5 mg po QHS was initiated to help with appetite stimulation. Three weeks after this visit, the patient was transferred from the dementia unit to a skilled nursing unit due to poor safety awareness associated with a number of falls. During his first week in the new location he had two episodes of inappropriate sexual behavior: he grabbed a nursing assistant's buttocks and asked her if she wanted to see his "banana" and he was observed patting his groin in a hallway. The CGP examined the patient in a public area of the skilled nursing unit and observed that he was hostile, intermittently scratched himself and patted his exposed abdomen. When asked about the recently concluded music-based activity the patient replied "I was there because they dragged me in there" and he reported feelings of depression because "I have to put up with people like you." No changes in the patient's treatment were recommended. When the patient was assessed approximately 4 months after the initiation of treatment with MPA, the CGP learned that no episodes of sexually appropriate behavior had occurred during the past month but several episodes of verbal and physically aggressive behaviors had been observed including smearing feces, resisting attempts to change his soiled clothing, and attempting to spit and urinate on staff members. The patient informed the CGP that he was seeing a "spirit." His blood glucose at the time of this interaction was determined to be 587 mg/dL. After

consultation with the patient's PCP, a decision was made to discontinue MPA due to concerns that the patient had developed delirium secondary to MPA-associated metabolic/endocrinologic dysregulation.

References

Araujo AB, Mohr BA, McKinlay JB. Changes in sexual function in middle-aged and older men: longitudinal data from the Massachusetts male aging study. *J Am Geriatr Soc*. 2004;52:1502–1509.

Aziz A, Brannstrom M, Bergquist C, Silfverstolpe G. Perimenopausal androgen decline after oophorectomy does not influence sexuality or psychological well-being. *Fertil Steril*. 2005;83:1021–1028.

Bacon CG, Mittleman MA, Kawachi I, Giovannucci E, Glasser DB, Rimm EB. Sexual function in men older than 50 years of age: results from the health professionals follow-up study. *Ann Intern Med*. 2003;139:161–168.

Ballard EL, Poer CM. *Sexuality and the Alzheimer's Patient*. Durham, NC: The Duke Family Support Program; 1993.

Bhasin S, Cunningham GR, Hayes FJ, et al; Task Force, Endocrine Society. Testosterone therapy in men with androgen deficiency syndromes: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2010;6:2536–2559.

Butler RN, Lewis MI. *The New Love And Sex After 60*. Revised edition. New York: Ballantine Books; 2002.

Centers for Disease Control and Prevention. *HIV Surveillance Report, 2012; Vol. 24*. http://www.cdc.gov/hiv/library/reports/surveillance/. Published November 2014. Accessed February 19, 2015.

Dahlhamer JM, Galinsky AM, Joestl SS, Ward BW. Sexual orientation in the 2013 national health interview survey: A quality assessment. *Vital Health Stat 2*. 2014;169:1–32.

Davis SR, Davison SL, Donath S, Bell RJ. Circulating androgen levels and self-reported sexual function in women. *JAMA*. 2005;294:91–96.

Galinsky AM, McClintock MK, Waite LJ. Sexuality and physical contact in National Social Life, Health, and Aging Project Wave 2. *J Gerontol B Psychol Sci Soc Sci*. 2014;69:S83–S98.

Irwig MS. Do hypogonadal older men benefit from testosterone therapy? *Clinical Geriatrics*. 2007;15:32–35.

Kinsey AC, Pomeroy WB, Martin CE. *Sexual Behavior in the Human Male*. Philadelphia, PA: W.B. Saunders Company; 1948.

Kinsey A, Pomeroy W, Martin C, Gebhard P. *Sexual Behavior in the Human Female*. Philadelphia, PA: W.B. Saunders Company; 1953.

Lichtenberg PA, Strzepek DM. Assessments of institutionalized dementia patients' competencies to participate in intimate relationships. *Gerontologist*. 1990;30:117–120.

Lindau ST, Gavrilova N. Sex, health, and years of sexually active life gained due to good health: evidence from two US population based cross sectional surveys of ageing. *BMJ*. 2010;340:c810.

Lindau ST, Schumm LP, Laumann EO, Levinson W, O'Muircheartaigh CA, Waite LJ. A study of sexuality and health among adults in the United States. *N Engl J Med*. 2007;357:762–774.

Mace NL, Rabins PV. *The 36-hour Day*. Revised Edition. New York: Warner Books, Inc.; 1992.

Masters WH, Johnson VE. *Human Sexual Response*. Boston, MA: Little, Brown and Company; 1966.

Nusbaum MR, Lenahan P, Sadovsky R. Sexual health in aging men and women: addressing the physiologic and psychological sexual changes that occur with age. *Geriatrics*. 2005;60:18–23.

Rheaume C, Mitty E. Sexuality and intimacy in older adults. *Geriatr Nurs*. 2008;29(5):342–349.

Rosen RC, Fisher WA, Eardley I, Niederberger C, Nadel A, Sand M; Men's Attitudes to Life Events and Sexuality (MALES) Study. The multinational men's attitudes to life events and sexuality (MALES) study: I. prevalence of erectile dysfunction and related health concerns in the general population. *Curr Med Res Opin*. 2004;20(5):607–617.

Rowland DL. Neurobiology of sexual response in men and women. *CNS Sprectr*. 2006;11(8 Suppl 9):6–12.

Segraves RT. Female sexual dysfunction. *Primary Psychiatry*. 2007;14:37–41.

Shields KM, Hrometz SL. Use of sildenafil for female sexual dysfunction. *Ann Pharmacother*. 2006;40(5):931–934.

Srinivasan S, Weinberg AD. Pharmacologic treatment of sexual inappropriateness in long-term care residents with dementia. 2006;14:20–28.

Stimmel GL, Gutierrez MA. Sexual dysfunction and psychotropic medications. *CNS Spectr*. 2006;11(Suppl 9):24–30.

Ward BW, Dahlhamer JM, Galinsky AM, Joestl SS. Sexual orientation and health among U.S. adults: National health interview survey, 2013. *Natl Health Stat Report*. 2014;77:1–14.

Wierman ME, Arlt W, Basson R, et al. Androgen therapy in women: a reappraisal: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2014;10:3489–3510.

Winn RI, Newton N. Sexuality in aging: A study of 106 cultures. *Arch Sex Behav*. 1982;11:283–298.

Zeiss AM, Davies HD, Wood M, Tinklenberg JR. The incidence and correlates of erectile problems in patients with Alzheimer's disease. *Arch Sex Behav*. 1990;19:325–331.

57.6g HIV and Aging

RAEANNE C. MOORE, PH.D., AND MARíA J. MARQUINE, PH.D.

INTRODUCTION

Once considered a death sentence, HIV has been transformed into a chronic disease as a result of highly active antiretroviral therapy (ART). The life expectancy of HIV-infected (HIV+) adults receiving ART is now approaching that of the general population, and today more than half of HIV+ people in the United States are over age 50. This remarkable success has brought with it a new set of challenges. As people with HIV are becoming older, HIV has become a chronic disease with multiple interacting causes of morbidity and mortality, generating significant public health interest for understanding the multiple factors that can impact aging with HIV.

Recognizing the issues and needs of the increasing older HIV+ population, a White House Conference on HIV and Aging was held in 2010. In addition, the Office of AIDS Research of the National Institutes of Health commissioned a working group to develop an outline of the current state of knowledge and areas of critical need for research in HIV and Aging. The findings and recommendations of the working group are summarized in a published report. While much research is needed to better understand health outcomes in older HIV+ persons, the American Academy of HIV Medicine in collaboration with the AIDS Community Research Initiative of America developed a series of guidelines for clinicians managing older patients with HIV.

EPIDEMIOLOGY

In 2010, the estimated rate of diagnoses of HIV infection among older adults in the United States was 8.5 per 100,000 population. Similar to younger age groups, members of racial/ethnic minority groups are at increased risk of HIV infection among older adults, with rates of infection in 2010 being 41.6 for blacks/African Americans, 15.4 for older Hispanics/Latinos and 3.9 for older non-Hispanic whites. Older males are also at increased risk, with the rate of infection among older males (13.6) being 3.2 times the rate of older females (4.2). As shown in Figure 57.6g–1, the most common forms of transmission vary by gender.

Older persons in the United States are more likely to be diagnosed with HIV infection later in the course of the disease compared to younger persons, which can result in worse disease burden and prognoses. For example, the rates of progression to AIDS in 12 months, among people diagnosed with HIV in 2010, were notably lower among those aged 25 to 29 (24 percent) than older persons (44 percent of people aged 50 to 59, 49 percent of people aged 60 to 64, and 53 percent of people aged 65 and older). Deaths of persons with diagnosed HIV infection may be due to any cause (i.e., may or may not be related to HIV infection). Of the estimated 19,343 deaths among people living with HIV in the United States in 2010, 10,244 (53 percent) were among people aged 50 and older. In 2010, HIV was the 10th leading cause of death among men and women aged 50 to 54.

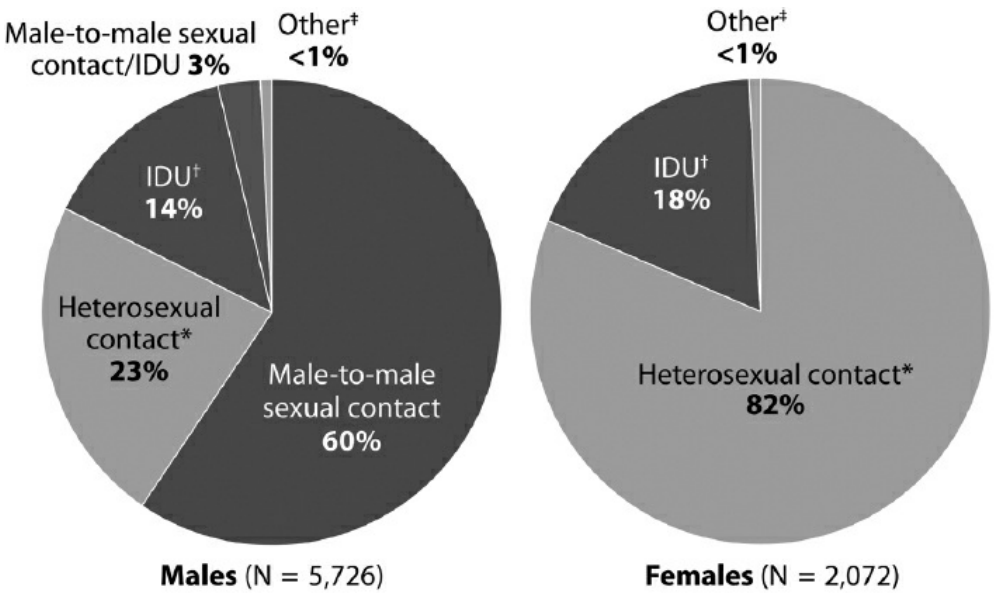

†Injection drug use.
‡Other transmission risk factors include hemophilia, blood transfusion, and risk factors not reported or identified.
*Heterosexual contact with a person known to have, or be at high risk for, HIV infection.
Due to rounding, percentages may not add up to 100%.

FIGURE 57.6g–1. Estimated diagnoses of HIV infection among adults aged 50 and older by gender and transmission category 2010. (From Centers for Disease Control and Prevention. Diagnoses of HIV infection among adults aged 50 years and older in the United States and dependent areas, 2007–2010. *HIV Surveillance Supplemental Report* 2013. February 2013.)

PREVENTION

Persons over the age of 50 appear to have several of the same risk factors for HIV infection than younger persons, which can be exacerbated by issues unique to older adults. Many older adults are sexually active and a number of them are widowed or divorced and dating again. Yet, they may lack knowledge about HIV, and thus be less likely to protect themselves. Older women who are not concerned about becoming pregnant may be less likely to practice safe sex. In addition, age-related thinning and dryness of the vaginal tissue may place older women at increased risk of HIV infection. For older men, erectile dysfunction medication may facilitate sex for some who would otherwise not have been able to engage in vaginal or anal intercourse.

While there is much left to understand regarding the reasons driving later HIV diagnoses among older persons, some of the key reasons may be that older persons are less likely than younger persons to discuss their sexual habits or drug use with their doctors, and that they might mistake HIV symptoms for those of normal aging. In turn, health care providers may also be less likely to ask about these issues and to test older persons for HIV infection. It also appears that HIV stigma might be of particular concern for older HIV+ persons and may prevent them from seeking care and disclosing their HIV status.

Given these prevention challenges the Centers for Disease Control and

Prevention, and its partners are implementing a high-impact prevention approach, which includes a number of initiatives and programs aimed at providing support and technical assistance to health departments and community-based organizations for implementing prevention programs, raising awareness about HIV in older age groups, supporting HIV testing, and optimizing linkage to care.

NEUROCOGNITIVE OUTCOMES OF HIV AND AGING

HIV and aging are independently associated with central nervous system (CNS) dysfunction. The cooccurrence of HIV and older age appears to have additive or synergistic negative effects on the CNS, with several studies observing elevated rates of neurocognitive impairment (NCI) in older HIV+ cohorts. However, the issue of whether the combined effects of HIV and aging on NCI are synergistic or additive remains controversial, with some studies finding a significant HIV and aging interaction and thus demonstrating a synergistic effect, and others showing an additive effect. For example, initial reports from the Hawaii Aging Cohort indicated twice the rate of HIV-associated dementia in older compared to younger HIV+ cohorts. In general, these reports have found particular problems in executive functioning, psychomotor speed, and episodic memory. In contrast, other studies have failed to detect an interaction between HIV and older age. For instance, the Hawaii Aging with HIV Cohort later did not detect a significant interaction between HIV and older age on NCI, but trends were observed on measures of learning, memory, attention, and timed gait.

The discrepant findings across studies have been attributed to a number of factors, including differences in demographics and other HIV disease characteristics across groups of younger and older HIV+ and HIV– adults, and differences in outcomes (global NCI vs. NCI in specific domains, and NCI vs. specific HIV-associated neurocognitive disorder or HAND, among others). Another factor to bear in mind for future studies in this area is that the *Diagnostic and Statistical Manual of Mental Disorders*, 5th edition (DSM-5) uses different criteria for neurocognitive disorders than the established Frascati criteria for HAND. For example, only problems on one neurocognitive domain are required by DSM-5 rather than two (per the Frascati criteria) to define impairment in "global" neurocognitive function implied by the term "cognitive disorder" (whether major or minor). Moreover, the exclusion criteria required for diagnosis of neurocognitive disorders are limited to other psychiatric disorders per the DSM-5 but explicitly include CNS illnesses specific to the setting of immunodeficiency, as noted by both the original AAN criteria and the current Frascati criteria. These disparities in neurocognitive disorder criteria might lead to difficulty with merging research across studies and obscure the findings of future research.

While much is yet to be known regarding key predictors of NCI among older HIV+ persons, a recent study investigated predictors of longitudinal neurocognitive decline in 436 HIV+ adults (age in years: mean = 43.9, standard deviation = 8.4; M follow-up = 35 months). Results from multivariable models showed that key predictors of earlier time to neurocognitive decline included Hispanic ethnicity, notable non-HIV comorbidity, being off ART, having low albumin and low hematocrit, having a lifetime methamphetamine use diagnosis and increased depressive symptoms. These, along with other findings, suggest that a number of different factors (HIV and non-HIV specific) are probably important predictors of NCI in older HIV+ persons.

In an effort to capture both HIV disease and non-HIV disease factors important for health outcomes in older HIV+ persons, multivariate indices indicative of NCI have been developed. One such index is the Veterans Aging Cohort Study (VACS) Index, which is predictive of mortality. The VACS Index combines age, current plasma HIV RNA, and CD4 cell count with non-HIV disease markers. In one recent study, a higher VACS Index score was associated with increased risk for global cognitive impairment, but the effect size was relatively small. Older age, low hemoglobin level, and low CD4 cell count were the components of the VACS Index more strongly linked to NCI.

MENTAL HEALTH OUTCOMES OF HIV AND AGING

Older persons living with HIV face unique emotional issues, including stigma, marginalization, and discrimination. In addition to emotional factors, these adults also often face unique social factors (e.g., lack of appropriate treatment facilities, poor access to health care, financial problems, disability) and unique family factors (e.g., HIV/AIDS in spouse, poor family support). These issues make HIV+ adults more vulnerable to psychiatric illness, decreased well-being, and poorer health-related quality of life (HRQoL). Prevalence rates of mood or anxiety disorders have been reported to be as high as 50 percent among older HIV+ adults, with a similar proportion reportedly having alcohol or other substance use disorders. In addition to being at heightened risk for clinical depression, those aging with HIV are at increased risk for negative mood compared to those without HIV infection. This cooccurrence of mental illness and HIV result in poorer health outcomes, including decreased CD4 T-lymphocytes and increased viral loads, and higher health care HIV costs. However, studies show that among older adults, 40 to 90 percent of mental health concerns are not detected in primary care settings, resulting in a majority of mental health problems going untreated among older HIV+ adults. Efforts to integrate primary health care at Infectious Disease or General Medicine clinics with mental health treatment and to combat the societal stigma that is associated with mental illness and HIV/AIDS are likely to

improve the health and well-being of persons living and aging with this disease.

The role of mental health professionals in treating older persons living with HIV is threefold: (1) identify behavioral factors most saliently impacting the course of HIV and well-being outcomes to determine how and when to intervene; (2) work with the patient's geriatrician and other primary care team members to determine the most appropriate course of treatment; and (3) implement treatment plan, closely monitor progress, and modify as needed.

Emotional Functioning

Depression management is a priority to achieve optimal health outcomes among those aging with HIV. Older HIV+ adults do not show the same age-related declines in depression that is observed among older adults without HIV, making depression all the more relevant in this population. Additionally, research has indicated moderately high levels of overall distress among older HIV+ adults; distress that is compounded by the addition of other chronic health problems. Both antidepressant medications and psychotherapy have showed effectiveness in treating depression among those living and aging with HIV. Effectively treating depressive symptoms can lead to better everyday outcomes, including better medication adherence, sustained virological suppression, reduced likelihood of HIV transmission, improved HRQoL, and longer survival. However, despite the existence of effective treatments for depression, underdiagnosed and untreated depression is highly prevalent. This is likely due to numerous factors complicating the clinical picture of aging with HIV, including individual patient characteristics, somatic symptom overlap between depression and chronic medical illnesses, and competing demands of primary care visits. Physiological changes associated with medication metabolism that may cause symptoms that look like depression also complicate the clinical picture and care planning. Studies have found patient race/ethnicity to be a contributor of depression treatment, with depressed African-American HIV+ patients less likely to receive treatment than their non–African-American counterparts. Additionally, investigators have found HIV+ patients diagnosed as severely depressed in a primary care setting more likely to receive mental health treatment than patients meeting clinical criteria for mild or moderate depression. These investigators also found that if depression was not treated on the day of a patient's primary care assessment, it was unlikely to be treated within the following 90-day period, regardless of depression severity.

It is also important to consider cohort effects when diagnosing and treating older adults living with HIV. The mental health of older adults who have been newly infected with HIV will look different from the mental health of adults who have successfully survived into older age with HIV.

See Figure 57.6g–2 for the authors' model of the inter-relationships between HIV, aging, and cognitive, emotional, and physical functioning on successful well-being outcomes. In addition to duration of infection, individual patients may be experiencing mental health problems for the first time in their lives, or they may have a long-term history of mental health symptoms. These cohort effects impact how older adults perceive mental health problems and treatment, self-identify, experience, and how they describe their emotions and distress. For example, while not recognized by the DSM-5 at this time as a clinical syndrome, many older HIV+ long-term survivors have what some of these patients refer to as AIDS Survival Syndrome. *AIDS Survival Syndrome* has been defined as a latent trauma response that emerges among those who survived the AIDS epidemic. These individuals represent a unique group of older HIV+ adults who lived through a period of history in which they suffered the loss of a majority of their friends, partners, and community, and were faced with the real threat of death to their own personhood. *AIDS Survival Syndrome* is thought to manifest as a cluster of symptoms including one or more of the following: depression, personality changes, flashes of anger, survivor guilt, anxiety, emotional numbness, insomnia, social withdrawal and isolation, hopelessness, substance abuse, sexual risk-taking, and a lack of future orientation. Symptoms can range from mild-to-severe. While empirical evidence is currently lacking on AIDS Survival Syndrome, the authors believe this nomenclature can help clinicians understand and assess the emotional issues faced by this particular cohort of older HIV+ adults, and provide education and develop appropriate treatment plans for these individuals.

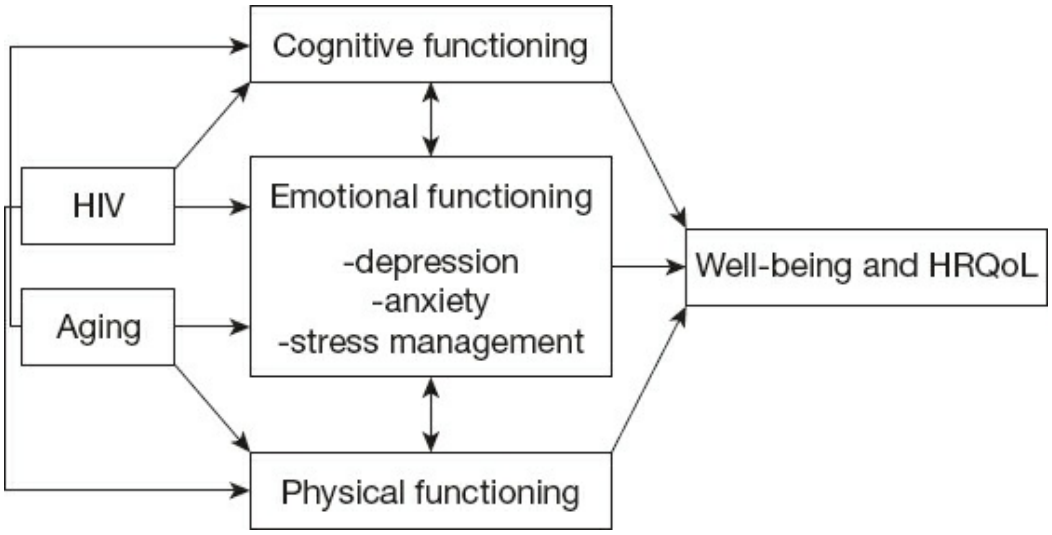

FIGURE 57.6g–2. Trajectory of psychosocial outcomes in aging with HIV. HRQoL, Health-Related Quality of Life.

Substance Use

Prevalence rates of substance use are higher in older HIV+ adults compared to their HIV– peers, and range from approximately 20 to 50 percent. Substance use is a known primary risk factor for HIV mediation nonadherence, particularly in those using hard drugs (i.e., cocaine or amphetamines) or injection drug users, or those with multiple-substance use such as combined cocaine and methamphetamine or combined cocaine and heroin. Excessive alcohol use, with and without other substance use, is also a predictor of poor adherence. A large study of substance use patterns among HIV+ adults aged 50 and older found the poorest adherence and greatest self-perceived impairment from substance use among those using alcohol and cocaine/crack or those using multiple substances. Another large study examining excessive alcohol use among older HIV+ adults with poor antiretroviral adherence found the following clinical characteristics to be related to the most severe unhealthy drinking patterns: African-American race, less than full-time employment, alcohol treatment within the past year, and a positive depression screen. The pronounced role of substance use among older HIV+ adults, together with the known negative impact of substance use on medication adherence and other functional outcomes, underscores the need for ongoing substance use screening and treatment in addition to increased prevention efforts on the part of both primary care clinicians and mental health professionals.

HIV and Severe Mental Illness

The literature on the cooccurrence of severe mental illness (SMI) and HIV/AIDS is limited, yet is known to be associated with greater risk for morbidity or mortality than with either illness alone. The association between SMI and HIV differs at an individual level, and may be due to their simple coexistence or SMI preceding HIV infection. Studies in the United States estimate prevalence rates of HIV infection, among patients with SMI, ranging from 3 to 23 percent. Alternatively, among those with psychosis, HIV infection may have a direct impact on the development of psychosis due to effects of the virus or antiretroviral medications on the CNS.

Patients with schizophrenia, in particular, may have difficulty describing their symptoms and may be more susceptible to HIV-related distress. Clinical evaluation of HIV+ older adults with schizophrenia requires a comprehensive assessment of the onset and course of symptoms. Additionally, caution should be taken when prescribing antipsychotic medications and take into account medications' side effects and interactions with antiretroviral medications.

Some evidence exists for the beneficial effects of antidepressant medication on improving antiretroviral adherence among patients with cooccurring HIV, bipolar disorder, and current mental health symptoms. For example, research examining antiretroviral adherence among patients

with HIV and bipolar disorder found antiretroviral adherence was significantly better than adherence to psychiatric medications. These investigators also found a correlational relationship between antiretroviral adherence and psychiatric medication adherence, indicating increasing adherence in one medication may have an impact on adherence in the other. In other research, a disconnect between memory performance on standard neuropsychological tests, subjective memory complaints, and depressive symptoms in this population has been observed, such that memory complaints were not associated with memory performance but were associated with depressive symptoms. In turn, memory complaints were predictive of problems with everyday functioning (e.g., unemployment, dependence on activities of daily living). Compared to HIV+ adults without bipolar disorder, those with cooccurring bipolar disorder had less awareness of their actual memory abilities. Investigators concluded that both subjective memory complaints and objective memory ability should be considered during treatment planning in this subset of patients with HIV, as both appear to impact everyday functioning behaviors.

Successful Aging with HIV

To date, most studies have focused on factors associated with aging with HIV that are predictive of poor functioning, including physical illness burden, immune dysfunction, cognitive impairment, and psychiatric problems. Nonetheless, there is a growing body of literature that suggests the possibility of successful aging in individuals with HIV, and understanding HIV in this light may help identify individual factors related to better quality of life and well-being. Studies have shown that there is a potential for successful cognitive and emotional aging among older adults with HIV, despite significant medical comorbidities and disease burden. Additionally, the impact of negative mood on overall functioning appears to be buffered by protective psychological factors such as resilience and optimism.

In the healthy aging literature, research has suggested that subjective well-being and positive affect may be stronger predictors of morbidity and mortality in seronegative older adults than negative mood and somatic complaints. There is also a rapidly growing body of healthy aging literature demonstrating the benefits of physical activity on health and well-being outcomes. In HIV+ adults, studies have also demonstrated a positive impact of engagement in active lifestyle factors (namely physical, social, and mental activities) on better cognitive outcomes. Among adults aging with HIV, investigators have found an increased number of active lifestyle factors to be associated with better neurocognitive performance and lower prevalence of HAND. Interventions to enhance positive affect, positive psychological traits, and increase engagement in active lifestyle factors

among older HIV+ adults may have beneficial effects on mental health, cognitive, and daily functioning outcomes, as well as HRQoL.

COMORBIDITY AND MULTIMORBIDITY

As HIV-infected individuals are living longer, they are increasingly experiencing fewer AIDS-defining conditions and a wider range of non-AIDS diseases. Some of the more common comorbidities in older HIV+ persons include hypertension, diabetes mellitus and insulin resistance, cardiovascular disease, pulmonary hypertension, liver disease, kidney failure, non-AIDS cancers, osteoporosis, peripheral neuropathy, and "frailty," which are characterized by weight loss, weakness, and increased risk of disability and death.

Although comorbidities in HIV+ persons on ART are often defined as "non-AIDS," they are associated with HIV infection (HIV-associated non-AIDS or HANA), HIV treatment, and/or behaviors, conditions, and demographics more common among those with HIV infection. Thus, while they are also common in HIV– older persons, these conditions may act differently among older persons with HIV infection. Several mechanisms have been suggested, including microbial translocation, chronic inflammation, oxidative stress, and immune senescence. Yet, much work remains to be done to understand the mechanisms underlying comorbidity in older HIV+ persons.

Guidelines are available for the management of common comorbid conditions in older HIV+ persons. It is important to bear in mind, however, that multimorbidity is common among older HIV+ patients and is more than simple comorbidity. Multimorbidity is conceptualized as the presence of several health problems that have the propensity to synergize to worsen a person's functional capacity and increase mortality more than might otherwise seem apparent. When considering diagnosis and treatment in multimorbidity, the sum is greater than the parts. This multimorbidity can result in overlapping injury to multiple organ systems, transforming HIV into a complex chronic disease. This has prompted the need for better methods to track the impact of comorbidities and multiple organ system dysfunctions on the health of people with HIV.

The VACS Index was developed in this context, as a tool to monitor the impact of dysfunction involving multiple organs commonly affected by HIV, using objective measures that are collected routinely in HIV clinics. It includes age, traditional HIV biomarkers (i.e., HIV-1 plasma RNA and current CD4 count), and non-HIV biomarkers (i.e., indicators of renal and liver function, anemia, and Hepatitis C coinfection). The VACS Index has been consistently validated for estimating risk for mortality among HIV-infected persons, including that in diverse cohorts. It has also been associated with NCI, but the strength of this association is relatively weak.

IMPACT ON EVERYDAY FUNCTIONING

Real-world functioning impairment is a known sequelae of HIV infection, and functional disability expenditures have made HIV among the most expensive chronic illnesses. As such, the NIH Office of AIDS Research, and the U.S. Senate Special Committee on Aging have recently emphasized the significant public health priority for research, improving functioning in those aging with HIV. Real-world functioning in aging and HIV is complex and multidetermined. Age and HIV are associated with increased risk for NCI, which, along with numerous risk factors (i.e., certain demographics, depression, anxiety, memory complaints, social isolation), result in older HIV+ adults having a disproportionate disruption in real-world daily functioning. *Real-world functioning* in defined as what a person actually does in their daily lives. In particular, decreases in job-related activities, self-reported problems with basic activities of daily living (BADLs) and instrumental activities of daily living (IADLs), clinician-reported functional outcomes, ART nonadherence, and driving accidents have been found in older HIV+ adults relative to their younger counterparts. These age-related impairments have been found irrespective of neurocognitive status. In terms of *functional capacity*, defined as a person's abilities that are essential to function independent in the community (i.e., what they are capable of doing, not what they actually do), a synergistic effect of HIV and neurocognitive status has been observed, such that older, cognitive-impaired HIV+ adults perform worse on tasks of medication management and finance management than younger and unimpaired groups.

While older age has been found to exacerbate HIV-associated disability in real-world functioning, not all older HIV+ adults evidence functional impairment, which raises important questions about individual clinical characteristics impacting poor functional outcomes. The most well-studied predictor of disruptions in real-world functioning is NCI. The association between neurocognitive and real-world functional impairment is highly complex, and even subtle neurocognitive deficits, such as shallow encoding and forgetting, have been found to be associated with IADL dependence in this population. Factors believed to influence this relationship include the match between the individual's profile of neurocognitive strengths and weaknesses and the specific demands of the real-world activity (e.g., its complexity and cognitive and functional requirements). Motivation, awareness of cognitive and functional deficits, cognitive reserve (i.e., effects of enrichment of cognitive resources through educational and vocational experiences together with a person's innate intellectual ability), and the use of compensatory strategies also play a role in determining the nature of this relationship. Finally, there are a multitude of additional static and dynamic (i.e., changes in one area may impact other areas of functioning) biological, psychological, and social factors influencing the impact of neurocognitive ability on real-world daily functioning. For example, having sufficient home

and community resources, along with appropriately adjusting daily routines to account for physical and cognitive symptoms, have been associated with successful real-world daily functioning among HIV+ adults.

Cognitive reserve may be especially important to consider within the context of aging with HIV. Research has found lower cognitive reserve to be associated with risk for HIV-associated NCI and greater declines in IADLs. Conversely, higher cognitive reserve has been found to be the sole significant clinical cofactor of successful cognitive aging among the older HIV+ adults, whereas among younger HIV+ adults several cofactors affecting successful cognitive aging have emerged (i.e., higher cognitive reserve, Caucasian race/ethnicity, shorter duration of HIV, and lower prevalence of lifetime alcohol dependence). Thus, the evidence suggests that among older HIV+ adults, lower cognitive reserve may interfere with the adaptive ability to initiate the use of compensatory strategies when engaging in behaviors in their daily lives.

Apathy and depression are also two known reliable predictors of adverse functional outcomes among older HIV+ adults. A link between more severe depressive symptoms and over-reporting neurocognitive and functional impairments has been identified. Furthermore, a link has also been found between less severe depressive symptoms and underreporting impairments. Associations between higher self-reported apathy with greater severity of IADL declines and more numerous cognitive complaints have been observed, suggesting that apathy may make a unique contribution to important everyday functioning outcomes. Fortunately, these risk factors can easily be screened for and are amenable to change. One rigorously designed study found no relationship between antidepressant use better daily functioning among older HIV+ adults with a current or lifetime major depressive disorder diagnosis. While to the authors' knowledge the relationship between antidepressant use and daily functioning among older HIV+ adults has only been explored in one study, it may be important for clinicians to consider treatment incorporating a combination of pharmacotherapy and psychotherapy.

When assessing and treating functional impairments in this population, the patient's medical history, medical comorbidity burden, HIV disease burden, culture, and cohort effects should also be carefully considered. In the current era of effective antiretroviral therapy, older adults recently diagnosed with HIV may not have the same downstream effect of HIV disease-specific biomarkers on real-world daily functioning that can be observed in older survivors; thus, disability risks may vary based on duration of infection. In terms of the cultural impact of age and HIV on real-world functioning, assessing the sociocultural context in which HIV infection occurs is important to determine factors impacting functional outcomes. These sociocultural factors may include (but are not limited to) access to health care and antiretroviral treatment, value of employment

versus disability, HIV-associated stigma, and behavioral expressions distress.

THE SPECIAL CASE OF THE VETERAN POPULATION

The Veterans Health Administration (VHA) is the largest single provider of HIV care in the United States, with 1 out of every 250 veterans in care at VHA living with HIV/AIDS. The VA faces challenges to meeting the needs of the growing population of older HIV+ veterans. Up to 63 percent of HIV-infected veterans have at least one mental health diagnosis, and many veterans in need of mental health treatment are not receiving treatment. Primary barriers to treatment include identified targets for intervention, scheduling and availability of a mental health provider, time and travel to and from a VA facility, and privacy concerns.

Older veterans with HIV are at increased risk for depression as well as loneliness and social isolation. Depression affects over half of HIV+ veterans, making it the most common comorbidity in this population. The VACS found social isolation to be associated with greater risk of hospitalization and death in HIV+ and HIV– older veterans, and reported that this effect may be greater in HIV+ veterans. In a retrospective cohort study using data from the national VHA HIV Clinical Case Registry, researchers examined the impact of mental illness among HIV+ veterans on progression to AIDS-defining illness and death, and found substance use to be predictive of AIDS-defining illness, whereas schizophrenia, bipolar disorder, and substance were predictive of all-cause mortality. In summary, data from these studies indicate the need for increased knowledge of the mental health needs of older HIV+ veterans and increased access to care, specifically care targeting the aforementioned risk factors for poor outcomes.

TREATMENT AND MANAGEMENT

With longer life expectancy and improved health outcomes, it is no longer assumed that adults aging with HIV will need to collect disability payments as income, but instead, many may seek to continue working until retirement age. Given the public health burden of functional impairments associated with HIV, it is imperative to understand the risk and protective factors impacting real-world daily functioning. Additionally, it is equally important to understand these risks and protective factors in the context of sociodemographic characteristics, including sex, age, cohort effect, racial/ethnic minority status, and socioeconomic status. A targeted understanding of the most potent factors impacting real-world daily functioning at an individual level will aid clinicians in determining how and when to intervene in order to improve functioning and HRQoL, as well as to identify functional rehabilitation interventions that can be incorporated into clinical care.

Despite the greater availability of knowledge regarding the neurocognitive consequences of HIV disease compared to the psychiatric effects, there remains a paucity of literature on treating HAND and improving overall health outcomes. Evidence has consistently shown that patients with histories of significant immunocompromise (e.g., low nadir CD4) have increased risk of poor neurocognitive outcomes. Therefore, current practice standards recommend prescribing combined antiretroviral therapy (cART) at the time of HIV diagnosis as a prevention strategy against the neuroinflammatory cascades associated with poor immune functioning. However, a meta-analysis suggests the benefits of cART on HAND effects may be modest among older HIV+ adults. In terms of nonantiretroviral pharmacotherapy for HAND, there is some evidence that certain psychiatric medications used to treat mood problems (e.g., lithium, serotonin reuptake inhibitors) may have neuroprotective benefits in HIV. Thus, treating depression and other mood problems may have the added benefit of strengthening cognitive ability. Clinical trials using neuroenhancing drugs for the treatment of HAND have been largely unsuccessful. Overall, there is currently no first-line clinical pharmacotherapy for HAND among adults aging with HIV. While cognitive and behavioral rehabilitation strategies are increasingly being explored, published studies on cognitive neurorehabilitation approaches for HAND are scarce. Solely using pharmacotherapy for HAND treatment is contraindicated, as there is a strong correlation between neurocognitive deficits and medication nonadherence among persons living with HIV. Psychotherapeutic interventions incorporating cognitive and behavioral components with adherence (i.e., cognitive-behavioral therapy for adherence and depression [CBT-AD]) are potentially efficacious supplements or alternatives to pharmacotherapy. Additionally, there is some promising early evidence of the potential efficacy of restorative cognitive rehabilitation approaches.

Lifestyle modification is another potential functional rehabilitation treatment approach. Evidence-based behavioral treatments exist for a vast array of lifestyle factors, including physical activity, stress reduction, mental activity, social engagement, sleep hygiene, and substance use reduction. For example, randomized controlled trials have demonstrated preliminary evidence for the positive effects of mindfulness-based stress reduction on reducing behavioral avoidance, increasing positive affect, reducing depressive symptoms, and reducing cART side effects (e.g., 30). Although many of the aforementioned lifestyle factors have yet to be tested (via controlled trials), specifically among older adults with HIV, many of these factors protect and promote cognitive reserve, increase brain-derived neurotrophic growth factors, and reduce neuroinflammation and oxidative stress, which would have a direct benefit on cognitive and real-world everyday functioning.

FUTURE DIRECTIONS

As long-term survival among persons living with HIV continues to improve, the need for identifying targets for improving outcomes in this population is clear. Comorbid health problems are the norm among persons living with HIV who are 50 years of age and older. Continued efforts to integrate primary health care, geriatricians, HIV providers, and mental health treatment are needed to address the needs of this population. Primary health care providers and HIV providers are unlikely to assess for HAND or mood symptoms; this has a direct negative impact on treatment. Comprehensive assessment and integrated care planning in this population, regardless of age, is necessary. It is increasingly important for the patients, their care providers, and the health care system to expand the focus of clinical research and practice beyond alleviation of symptoms, to consider issues of real-world functioning and HRQoL.

REFERENCES

Al-Khindi T, Zakzanis KK, van Gorp WG. Does antiretroviral therapy improve HIV-associated cognitive impairment? A quantitative review of the literature. *J Int Neuropsychol Soc*. 2011;17(6):956–969.

American Academy of HIV Medicine. *Updated—The HIV and Aging Consensus Project: Recommended Treatment Strategies for Clinicians Managing Older Patients with HIV*. 2014.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

*Antinori A, Arendt G, Becker JT, et al. Updated research nosology for HIV-associated neurocognitive disorders. *Neurology*. 2007;69:1789–1799.

Balderson BH, Grothaus L, Harrison RG, McCoy K, Mahoney C, Catz S. Chronic illness burden and quality of life in an aging HIV population. *AIDS Care*. 2013;25(4):451–458.

Blackstone K, Tobin A, Posada C, et al; The HIV Neurobehavioral Research Program (HNRP) HIV-infected persons with bipolar disorder are less aware of memory deficits than HIV-infected persons without bipolar disorder. *J Clin Exp Neuropsychol*. 2012;34(7):773–781.

Blazer DG, Hybels CF. What symptoms of depression predict mortality in community-dwelling elders? *J Am Geriatr Soc*. 2004;52(12):2052–2056.

Centers for Disease Control and Prevention: Diagnoses of HIV infection among adults aged 50 years and older in the United States and dependent areas, 2007–2010. *HIV Surveillance Supplemental Report*. 2013. February 2013.

Ciesla JA, Roberts JE. Meta-analysis of the relationship between HIV infection and risk for depressive disorders. *Am J Psychiatry*. 2001;158:725–730.

Fazeli PL, Doyle KL, Scott JC, et al; HIV Neurobehavioral Research Program (HNRP) Group. Shallow encoding and forgetting are associated with dependence in instrumental activities of daily living among older adults living with HIV infection. *Arch Clin Neuropsychol*. 2014;29(3):278–288.

Fazeli PL, Woods SP, Heaton RK, et al. An active lifestyle is associated with better neurocognitive functioning in adults living with HIV infection. *J Neurovirol*. 2014;20(3):233–242.

Gayner B, Esplen MJ, DeRoche P, et al. A randomized controlled trial of mindfulness-based stress reduction to manage affective symptoms and improve quality of life in gay men living with HIV. *J Behav Med*. 2012;35(3):272–285.

Greysen R, Horwitz LI, Covinsky KE, Gordon K, Ohl ME, Justice AC. Does social isolation predict hospitalization and mortality among HIV+ and uninfected older veterans? *J Am Geriatr Soc*. 2013;61(9):1456–1463.

Heaton RK, Franklin DR Jr, Deutsch R, et al; CHARTER Group. Neurocognitive change in the era of HIV combination antiretroviral therapy: the longitudinal CHARTER study. *Clin Infect Dis*. 2015;60(3):473–480.

*High KP, Brennan-Ing M, Clifford DB, et al; OAR Working Group on HIV and Aging. HIV and aging: state of knowledge and areas of critical need for research. A report to the NIH Office of AIDS Research by the HIV and Aging Working Group. *J Acquir Immune Defic Syndr*. 2012;60 Suppl 1:S1–S18.

*Hooshyar D, Goulet J, Chwastiak L, et al; VACS Project Team. Time to depression treatment in primary care among HIV-infected and uninfected veterans. *J Gen Intern Med*. 2010;25(7):656–662.

Justice AC, Modur SP, Tate JP, et al; NA-ACCORD and VACS Project Teams. Predictive accuracy of the Veterans Aging Cohort Study index for mortality with HIV infection: a North American cross cohort analysis. *J Acquir Immune Defic Syndr*. 2013;62:149–163.

Marquine MJ, Umlauf A, Rooney AS, et al; HIV Neurobehavioral Research Program (HNRP) Group. The Veterans Aging Cohort Study Index is associated with concurrent risk for neurocognitive impairment. *J Acquir Immune Defic Syndr*. 2014;65(2):190–197.

Moore DJ, Posada C, Parikh M, et al; HIV Neurobahavioral Research Program (HNRP). HIV-infected individuals with co-occurring bipolar disorder evidence poor antiretroviral and psychiatric medication adherence. *AIDS Behav*. 2012;16(8):2257–2266.

Moore RC, Fazeli PL, Jeste DV, Moore DJ, Grant I, Woods SP; HIV Neurobehavioral Research Program (HNRP) Group. Successful cognitive aging and health-related quality of life in younger and older adults infected with HIV. *AIDS Behav*. 2014;18(6):1186–1197.

Morgan EE, Iudicello JE, Weber E, et al; HIV Neurobehavioral Research Program (HNRP) Group. Synergistic effects of HIV infection and older age on daily functioning. *J Acquir Immune Defic Syndr*. 2012;61(3):341–348.

Nurutdinova D, Chrusciel T, Zeringue A, et al. Mental health disorders and the risk of AIDS-defining illness and death in HIV-infected veterans. *AIDS*. 2012;26(2):229–234.

*Parsons JT, Starks TJ, Millar BM, Boonrai K, Marcotte D. Patterns of substance use among HIV-positive adults over 50: implications for treatment and medication adherence. *Drug Alcohol Depend*. 2014;139:33–40.

Sueoka K, Goulet JL, Fiellin DA, et al. Depression symptoms and treatment among HIV infected and uninfected veterans. *AIDS Behav*. 2010;14(2):272–279.

*Thames AD, Becker BW, Marcotte TD, et al. Depression, cognition, and self-appraisal of functional abilities in HIV: an examination of subjective appraisal versus objective performance. *Clin Neuropsychol*. 2011;25(2):224–243.

Thames AD, Kim MS, Becker BW, et al. Medication and finance management among HIV-infected adults: the impact of age and cognition. *J Clin Exp Neuropsychol*. 2011;33(2):200–

209.
United States Senate Special Committee on Aging. *Hearing: Older Americans: The changing Face of HIV/AIDS in America*. In: *One Hundred Thirteenth Congress, First session, September 18, 2013*. Washington, DC; 2013.
Valcour V, Paul R, Neuhaus J, Shikuma C. The effects of age and HIV on neuropsychological performance. *J Int Neuropsychol Soc*. 2011;17(1):190–195.
Valcour V, Shikuma C, Shiramizu B, et al. Higher frequency of dementia in older HIV-1 individuals: the Hawaii Aging with HIV-1 Cohort. *Neurology*. 2004;63(5):822–827.
Williams EC, Bradley KA, Balderson BH, et al. Alcohol and associated characteristics among older persons living with human immunodeficiency virus on antiretroviral therapy. *Subst Abus*. 2014;35(3):245–253.

57.6h Technology for Seniors

RUJVI KAMAT, PH.D., COLIN A. DEPP, PH.D., AND IPSIT V. VAHIA, M.D.

INTRODUCTION

With the population of older adults expected to reach two billion worldwide by 2050, there will be a serious shortage of skilled care providers to manage the needs of this population. To cope with the increasing demands for health and social care, there is a great need to reorganize care processes and leverage different technologic solutions. Simultaneously, with the rapid maturing of the "aging in place" movement, there is an unprecedented incentive to optimize an array of technologies to help adults maintain independence in their own homes, and avoid or forestall transition to institutionalization. Over the past decade, there has been a revolution in how technology has both evolved and impacted the global population at large. Canadian social scientists, Wellman and Quan-Haase used the term "hyperconnectivity" to describe multimodal human-to-human and human-to-machine interaction using a variety of platforms including face to face, email, using mobile phone technology, and via the Internet. The implication of this term is that most persons will communicate via a multitude of channels, facilitated by technology, and that human interaction with multiple forms of technology will grow substantially.

In response to this societal shift, there has been an explosion in technologies over the past decade, facilitated not just by seamless communication, but also advances in computing and sophisticated sensors that can increasingly provide continuous objective data on a number of parameters. Although older adults have traditionally been considered late adopters of emerging technology, there is growing consensus that this demographic group, with its needs for increased care and monitoring, especially among the oldest-old with multiple medical needs and reduction in independent functioning, represents a significant opportunity for technology-based innovation.

COMPARATIVE NOSOLOGY

Defining Technology in Aging

The term *technology* is derived from the Greek term *"tekhne,"* which means art or skill. It is defined by the Collins English Dictionary as the "total knowledge and skills available to any human society for art, skills, etc." This is a very broad term, subject to multiple interpretations, and inclusive of technology in all its forms and considerations, from use of flint to make fire to the latest developments in robotics. More recently, within the context of aging, researchers have favored the use of the term "gerontechnology" which has been described as "an interdisciplinary field of scientific application of technologic applications that are directed toward the aspiration and opportunities of the older person." These are broad terms, and as may be expected, they cover a very broad range of technologies. Table 57.6h–1 presents an overview and broad classification of types of technology, based on application. This chapter will focus primarily on Information and Communication Technology and Assistive Technology. Other technologies including Medical Technology and Smart Home Technology are also discussed.

There has been a growing cross-disciplinary effort to develop and apply information and communication technologies (ICTs) for acquiring, processing, presenting, and managing health-related data in older adults. The goal of these technologies is to improve health, quality of life, and social connectivity for less impaired older adults, and assist in clinical care and monitoring of more impaired older adults. Although older adults do not use ICTs as commonly as younger individuals, they do use devices such as smartphones in comparable ways, that is, to communicate, access information (health-related and otherwise), entertainment, and financial management. Emerging data suggest that certain factors, such as perceived advantage of technology and design issues contribute to the lower rates of technology use in older adults. These issues may serve to guide the modification and development of ICTs tailored to engage and benefit older adults.

Table 57.6h–1.
Classification of Technology for Aging

Type	Example	Target Population	Role
Medical technology	Pacemakers, implanted medication infusion systems, vagal nerve stimulators	Focused: Older adults with specific medical issues	Targeted medical role
Assistive Technology	Telemedicine systems, sensors	Narrow: Older adults with medical conditions that	Maintenance of physical health and well-being

	(e.g., accelerometers), monitors (e.g., glucometers)	require monitoring	
Information and Communication Technology	Wi-Fi, smartphones, tablet devices (e.g., iPads)	Broad: Older adults with need and ability to connect to the internet or use cellular technology for communication	Staying in touch with friends or family, more integrated social functioning, access to information
Smart Home technology	Personal Emergency Response Systems, Environmental Sensors	Broad: Older adults with cognitive or functional impairments, invested in aging in place	Facilitating remote monitoring of older adults living independently, minimizing risk of in-home adverse events

Assistive technologies have a more narrow scope and focused application, since they are usually developed to target specific medical conditions or functional deficits. It is important to note that a specific technology may be classified in more than one category. For example, a smartphone is usually considered a platform for an ICT. However, it may be used as a form of assistive technology, if a specific app has a targeted role in monitoring. In addition, there is a new generation of cellphone accessories that aim to leverage the computing capabilities of smartphones to transform them for use as medical devices. For example, a California-based company called CellScope has recently launched a device called the Oto, which fits onto the camera of smartphones, and allows the phone to function as an otoscope. Such devices have the potential to facilitate remote consultations and rapid diagnoses.

Technology and Aging in Place

An important goal of technology for older adults is to compensate for age-associated losses in physical and cognitive functioning. There is growing evidence to support the use of assistive technologies as a way to allow older adults to "age in place." Aging in place is defined by the Centers for Disease Control and Prevention (CDC) as "the ability to live in one's own home and community safely, independently, and comfortably, regardless of age, income, or ability level." Social networking, remote monitoring, and assistive devices have shown promise in enhancing opportunities for aging in place. Since a significant proportion of increased healthcare costs associated with aging is related to the need for intensive monitoring and institutionalized care, promoting aging in place is an important approach to preserving independence and also reducing costs. Both ICTs and assistive technologies have a key role in promoting aging in place. The American Association for Retired Persons (AARP), which has taken on a prominent role in advocacy for aging in place, describes case scenarios on their website of how technology, including pillbox sensors, refrigerator

sensors, and personal emergency response systems (PERS) can help promote aging in place.

SPECIFIC APPLICATIONS OF TECHNOLOGY

Telehealth

Medical providers are increasingly using telecommunication devices in the diagnosis and care of patients in remote locations. These technologic modalities include video, voice, robotic, and remote-access technology to facilitate interactions between providers and patients. Telehealth approaches to observe patients and provide intervention services have direct relevance to psychiatry and geropsychiatry. The use of telemedicine has the potential to be cost effective and improve the ease with which patients may be observed and assessed.

Literature regarding the use of telehealth systems with patients with schizophrenia supports the reliability and feasibility of using videoconferencing as an assessment modality for patients with psychiatric illness. Furthermore, this approach appears to be a useful method of delivering interventions in this population and is associated with good functional and psychiatric outcomes.

Building on the feasibility of using telemedicine in cohorts with psychiatric illness, there has been a growing effort to develop concepts, systems, and devices in a geropsychiatry context. There is a pressing need to develop telemedical services for older adults. Telehealth approaches may mitigate unique challenges faced by this population. For example, distance from providers may restrict regular healthcare visits for older patients with limited mobility. Services such as video consultation and remote monitoring devices facilitate the transfer of information between the patient and doctor and monitor the appropriateness of clinical interventions, thus improving the quality and access to health care. A majority of interventional telemedical studies in the geriatric population have focused on monitoring older adults. This may take the form of routine personal monitoring accompanied by supportive contact between the healthcare provider and the patient and/or measurements of key biomarkers using technologic modalities. Overall, such interventions appear to have a positive impact on important health outcomes such as adherence and behavioral change.

Older patients tend to have physical and cognitive limitations and thus have distinctive needs that must be addressed by telehealth protocols. It is particularly promising that telehealth approaches appear to be feasible and useful in a wide variety of geriatric subgroups. For example, Loh and colleagues developed and examined a videoconferencing protocol for diagnosing Alzheimer disease and found that there was good diagnostic agreement between the telemedical modality and the face-to-face

assessment. Telepsychiatry has also been investigated in nursing home populations. Not only do telepsychiatry visits have a positive impact on the quality of patient care, they also appear to be associated with significant financial savings (approximately $30,000 saved over the course of 278 telepsychiatry visits). Taken together, these data support the potential benefit of telemedical applications in aging cohorts and highlight the need for greater empirical attention to the development and use of these novel approaches.

The application of telehealth concepts to geropsychiatry requires attention to the clinical benefits of these approaches, as well as their feasibility, acceptability, and cost–benefit ratio. As these systems spread in use, it is worthwhile to consider whether demographic factors (e.g., age, gender, race, level of education, and employment status) impact the ease with which older adults use these telehealth modalities. Furthermore, the influence of factors such as dementia, psychiatric illness, and caregiver availability also need to be better characterized. These variables can guide the efforts to maximize the usability and acceptability of telehealth interventions in older adults, and potentially, subsequent public policy.

Telemedicine-based approaches also continue to find application in providing clinical services in underserved areas. A recently published report describes a study conducted in Imperial County, California, which has the nation's highest rate of unemployment, and where over 80 percent of the population is Latino, and approximately 25 percent of the population meets Federal criteria for poverty. In this study, the authors compared neuropsychological testing in Spanish conducted in-person, with testing conducted via telemedicine, for a sample of adults meeting clinical indications for testing. The authors found that test performance was comparable across the telemedicine and in-person modalities. This study used a within-subjects design and randomized participants to determine which modality they received first; the findings are a strong indicator that telemedicine may serve as an acceptable and valid proxy for providing services such as neuropsychological testing.

Thus, as telehealth modalities continue to improve in design and implementation, they will have the potential to provide rapid, high-quality health care for aging cohorts.

Mobile Health

The term *mobile health* (or mHealth) loosely refers to the use of mobile phone technology for health care. In the past 5 years, there has been an explosion of interest in the uses of mobile devices in collecting health-related data, delivering health care and facilitating interaction among patient, clinician, and family stakeholders in the service of improving health. There are several reasons why mobile phones are specifically attractive as a potential platform for healthcare delivery. For one, recent

estimates indicate that the penetration of mobile phones is approximately 90 percent in the United States. This penetration has increased at rapid pace over the past 10 years and is nearing complete saturation. Of note, 64 percent of people own a smartphone, which means that the majority of people (at least in the United States) carry with them an increasingly potent device capable of real-time interactive communication and capture of continuous data on a variety of physiological and behavioral factors. Older adults are, as with other technologies, less likely to own a mobile device than their younger counterparts, but as of 2014, 74 percent owned a mobile device. The use and ownership of cell phones cuts across other sociodemographic strata, with poor and underserved subpopulations often more likely to rely on mobile devices. This ubiquity of mobile devices provides a remarkable new conduit for innovation in healthcare delivery, and both academic and consumer developers have responded with the creation of a wide variety of mHealth applications.

The "value proposition" of mHealth is that it could enable delivering care outside of the typical clinic setting, either by augmenting interventions by providing additional interaction to support everyday health sustaining behaviors (e.g., medication adherence), or by extending them beyond the time frame typically associated with the intervention. Remote mobile technologies can automate tasks that are typically performed by clinicians, such as appointment reminders, symptom and risk assessments, and scheduling, which could increase access to health care by reducing the need for face-to-face interaction. Additional benefits of mHealth are the ability to help compensate for prospective memory problems and facilitate self-management of chronic illnesses. Mobile phones can provide timely reminders for beneficial health behaviors and deliver "just-in-time" interventions that could diminish the time gap between symptom emergence and intervention. Given that large volumes of data are available both through patient-reports and sensors linked with the devices, it may also be possible to identify actionable patterns of symptoms over time in chronic recurrent illnesses, and then intervene based on accumulating knowledge of individual patients so as to personalize care to individuals. Finally, provider-facing applications may assist in managing patient needs and diagnoses, reducing the need for brick and mortar clinic settings, which may be particularly attractive in the developing world. To temper enthusiasm somewhat, while collectively these advantages offer great potential, the mHealth field is nascent and so few mHealth interventions are yet embedded in routine care.

mHealth interventions typically involve a number of elements, which include the following: (1) a patient facing platform. This is an interface that provides interaction, and typical platforms include the feature phone that is not connected to the Internet yet capable of voice or SMS interaction, a smartphone, or a dedicated device that is developed for the mHealth

intervention, (2) one or more active communication channels to the patient, which may include SMS, live or asynchronous voice, email or in-app interaction. Alternatively communication can involve passive sensing, or the collection of data without the need for patient response. Sensors are highly varied, but could include activity measured by GPS antenna or accelerometer embedded in the device or ancillary sensors that couple with the device, such as blood pressure or glucose, (3) a set of algorithms that govern the interaction between data captured on the device and a clinical response. These may be automated and do not involve direct clinical oversight or might flag or prompt interaction from clinicians tasked with monitoring the incoming data, (4) a server that stores and allocates interactive elements to individual users in the intervention network and, finally (5) a clinician facing dashboard or utility that enables a supervising clinician the capacity to manage, monitor, and interact with patients in the mHealth intervention. Many iterations of the above architecture are possible, with some apps that are completely consumer facing without clinician oversight and others offering complex layered interaction with a variety of clinicians and other technologic assets such as the electronic medical record.

A number of reviews have addressed the question as to whether mHealth interventions are effective. To date, the evidence is mixed. While recent systematic reviews have indicated that mHealth interventions seem effective for certain commonly researched areas, in particular smoking cessation and HIV medication adherence, other areas were limited in terms of evidence. A key limitation of the literature is the lack of high-quality randomized clinical trials. Another limitation pointed out by de la Vega and Miro is that there are over 100K health-related apps that are currently available for download in the iOS and Android markets, yet few of these have been subjected to any empirical study. Conversely, very few of the applications that have actually been subjected to rigorous randomized trials are actually available to consumers. As such, the mHealth field is somewhat bifurcated into unproven apps directly available to consumers, and academic products that are not translated to products available to consumers.

mHealth solutions for mental health and aging are far fewer in number than are ones for chronic medical illness. Nonetheless, there are some examples of apps designed for chronic mental illness, including schizophrenia, bipolar disorder, and substance abuse. Review studies have indicated that mHealth applications could provide important resources to support access, initiation, engagement, and sustainment of psychiatric care. A recent study published by Depp and colleagues indicates that even among older adults with schizophrenia, the rate of past exposure to mobile phones is high (78 percent had ever used one) and the great majority felt comfortable with use of mobile devices. Despite high rates of prior use, few

patients currently owned mobile devices, which do create a barrier to future implementation without subsidization of mobile devices. Moreover, data gathered by mobile devices on symptoms appear to be valid—one study indicated that daily mood ratings obtained in a sample of patients with bipolar disorder appear to be more consistent with clinician ratings than were identical mood ratings recorded on paper-and-pencil forms. As such, evidence, albeit limited, suggests that patients with chronic and severe psychiatric impairments are capable of using mobile devices and can report accurately on symptoms that would be the likely focus of intervention.

There have been few randomized trials of mHealth interventions in psychiatric illnesses. Depp and colleagues conducted a 24-week randomized controlled trial in 81 patients trial of a four-session psychoeducational intervention for bipolar disorder augmented by automated personalized reminders contrasted to a four-session intervention without mobile device augmentation. Effect sizes for this trial indicated that the mobile augmented condition was associated with greater reduction in depression, by a medium effect size; the magnitude of this effect is notable given the challenges met by many interventions in impacting bipolar depression. However, these positive effects seemed to dissipate at follow up suggesting that gains made by mobile augmentation did not sustain once mobile intervention ceased. Separate studies by Granholm and colleagues as well as Ben Zeev and colleagues have each reported positive changes in medication adherence, psychiatric symptoms, and socialization in open trials of SMS and smartphone application interventions in schizophrenia. Gustafson and colleagues evaluated a smartphone application designed to extend the impact of rehabilitation for alcohol use disorder and found significant reductions in alcohol use in a randomized trial of more than 300 patients. As such, these emerging clinical trials offer an encouraging indication of the feasibility, acceptability, and effectiveness of mobile interventions that augment psychiatric care. Given the large number of ongoing clinical trials of mHealth interventions in mental health conditions (see ClinicalTrials.gov), it is likely that data on the effectiveness of mental health targeting mobile applications will proliferate in the near future.

There are a variety of exciting and emerging directions for mHealth mental health applications. For one, there is increasing evidence that passive sensors (described in detail in the next section) can identify patterns consistent with change in symptoms. For example, vocal prosody appears to covary with mood symptoms in bipolar disorder. If passive sensors can supplant patient reports, it may be quite possible to efficiently monitor and allocate preventative care to patients without the need for patient-initiated interaction. Second, and relatedly, there is increasing interest in the use of predictive analytics in leveraging the intensive longitudinal data generated by mHealth applications to develop dynamic

personalize risk profiles for serious adverse events like suicide. Thompson and colleagues reported preliminary data on a statistical technique called functional data analysis that is optimized for the noisy repeated data that are generated in mHealth and found that a simple algorithm modeling negative affect highly accurately predicted subsequent increases in suicidal ideation in a small sample of bipolar patients. Other approaches have used natural language processing to identify patterns predictive of suicide. Other frontiers include the use of more sophisticated algorithms to intervene with patients depending upon social contexts or enlistment of social networks and peers in managing mental illnesses. As with mHealth in general, while exciting, each of these novel approaches would need to be adequately evaluated in well-designed clinical trials to move beyond proof-of-concept stages to translate to impact on psychiatric care.

In summary, the field of mHealth is a rapidly evolving one that, while yet to fully deliver on its promise, holds great potential for impacting the landscape of mental health care in older adults with psychiatric illnesses. Globally, older adults are increasingly adopting mobile technologies, and the development and tailoring of mHealth interventions to meet the needs of older adults will be a subject of research for years to come. Next, sensors and their role in quantifying health will be discussed.

Sensors/Self-Monitoring Devices

With accelerometers available readily, monitoring of movement, behavior, balance, as well as physiologic variables such as heart rate and oxygen saturation is becoming simpler. Researchers are developing sensors that can perform blood tests by connecting through ports to the computing chips in cellphones. Qualcomm—a San Diego-based communications company—has initiated a competition encouraging developers to produce a "tricorder," a noninvasive handheld diagnostic device, which can monitor multiple physiologic functions and perform diagnostics without availability of sophisticated labs.

Table 57.6h–2outlines various types of sensors, and the types of roles they can play in promoting aging care.

The use of sensor technology in health care and especially for older adults is a rapidly evolving field. Wrist-worn and body-worn accelerometers have found multiple uses among athletes, and to measure levels of physical activity. Most of these sensors rely on accelerometry to detect motion. However, the new generation of sensors incorporate additional elements such as using highly sophisticated computing chips (such as those found in smartphones), machine learning (i.e., the automated optimization of systems by use of statistical inference to make predictions), and wireless transmission of data and visualization (using software that also relies on the computing power of handheld devices).

Chemical sensors and body-worn sensors are noninvasive alternatives

that have shown promise in in vitro testing. These sensors can be easily produced on a mass scale using 3D printer technology and can be mounted on a variety of platforms such as mouthguards for measuring electrolytes and metabolites in saliva), textile (for measuring sweat or urine), bandaids (for measuring chemicals in interstitial fluids). These sensors can also be integrated with microneedles that are minimally invasive and painless. Currently, sensors in development can measure a wide range of electrolytes and metabolites including glucose, lactate, sodium, potassium, pH, and urea.

Wearable sensors can already measure vital signs, and commercially available devices such as the recently launched Apple Watch can measure and transmit heart rate in real time. These sensors have led to the creation of novel "personal health data." Several health care systems are working on developing infrastructure that can incorporate such data with electronic health records to have a more complete set of data on patient health and well-being.

Table 57.6h–2.
Classification of Sensor Technology

Type of Sensor	Examples	Role	Stage of Development (at time of writing)
Implantables	Nanosensors	Monitoring of parameters in very constrained environments, e.g., small blood vessels, GI	Conceptual development and prototyping Clinical validity not yet established
Neurosensors	Bitemporal EEG monitors	Portable noninvasive monitoring of neurologic (e.g., seizures), psychiatric (e.g., anxiety), and physiologic (e.g., pain) parameters	Commercially available. Clinical validity under assessment
Wearable physiologic monitoring devices	Microneedle-based sensors, tattoo sensors	Noninvasive monitoring of electrolyte and metabolite levels	Prototypes developed. In-vitro device validity and reliability under assessment
Wearable external monitors	Accelerometers, gyroscopes	Monitoring of psychical activity patterns and sleep	Commercially available and clinically validated
Environmental sensors	Motion sensors, gas sensors	External continuous monitoring of activity and functioning	Motion sensors are currently commercially available. Gas sensors are in prototype stage

Wireless electroencephalography (EEG) devices are also gaining popularity. These devices consist of a semicircular band with electrodes on either end, which can be placed bitemporally to collect bitemporal EEG

readings. These devices are also being utilized to gather EEG data to predict emotional states, pain, and anxiety.

As the use of wearable devices and sensors grows, there is a need to develop algorithms that can generate relevant clinical alerts based on changes in physiologic and environmental data. Further, medical providers, caregivers, and the wearers themselves will need tools to visualize these data and use that information to predict or respond to abnormal events.

Smart Homes and Senior Design

Advances in ICT have led to a growing interest in the development of "smart homes," that is, living environments equipped with multiple sensors that can monitor and detect a person's activity. The goal of these technologies is to improve the quality of life and promote physical independence of older adults. Within this context, multiple devices and technologies may be embedded within the infrastructure of the residence and may thus function without the need to train residents on their use. For example, a heat sensor mounted on a stove can detect if the stove has accidentally been left on for a prolonged period. Motion sensors can keep track of a person's mobility in the house and may provide information on clinical issues such as triggers of insomnia and markers of agitation such as pacing. Existing prototype smart homes (e.g., the Awarehome at the Georgia Institute of Technology) serve as research environments that guide development of technologies that may aid development of senior-friendly living.

Smart homes may be beneficial for individuals living alone who are unable to seek help in emergencies (e.g., falls, strokes, unconsciousness), older adults with cognitive (e.g., memory loss, visuospatial deficit) and/or physical (e.g., hearing, visual, or mobility) impairments, and individuals needing assistance in performing basic and instrumental activities of daily life. Additionally, smart homes may provide integrated telehealth tools that can connect people living in rural or underserved areas with medical providers. With regard to the use of smart homes, important questions such as health outcomes, clinical algorithms for predicting potential health problems, user perception and acceptance and privacy issues are yet to be fully addressed. Ideally a smart space, depicted in Figure 57.6h–1 for health monitoring should be able to integrate implanted, body-worn, and environmental sensors to measure changes with the individual's body, their behavior and changes in the environment. Such a space would also incorporate mHealth approaches (described above) and machine learning and big data analytics (described below) to generate relevant health information with the capability to be predictive of health events.

PERS are another set of important applications of ICTs. These are relatively inexpensive devices designed to facilitate emergency responses

for older adults. They may reduce mortality rates and hospital utilization and may be especially useful in rapid detection and early response to falls, and use of this technology is a subject of great interest and study.

Overall, it remains to be seen how sophisticated home-based solutions and the growing number of advanced ICT applications are tolerated and utilized by older adults in general and specific groups such as those with disabilities or dementia.

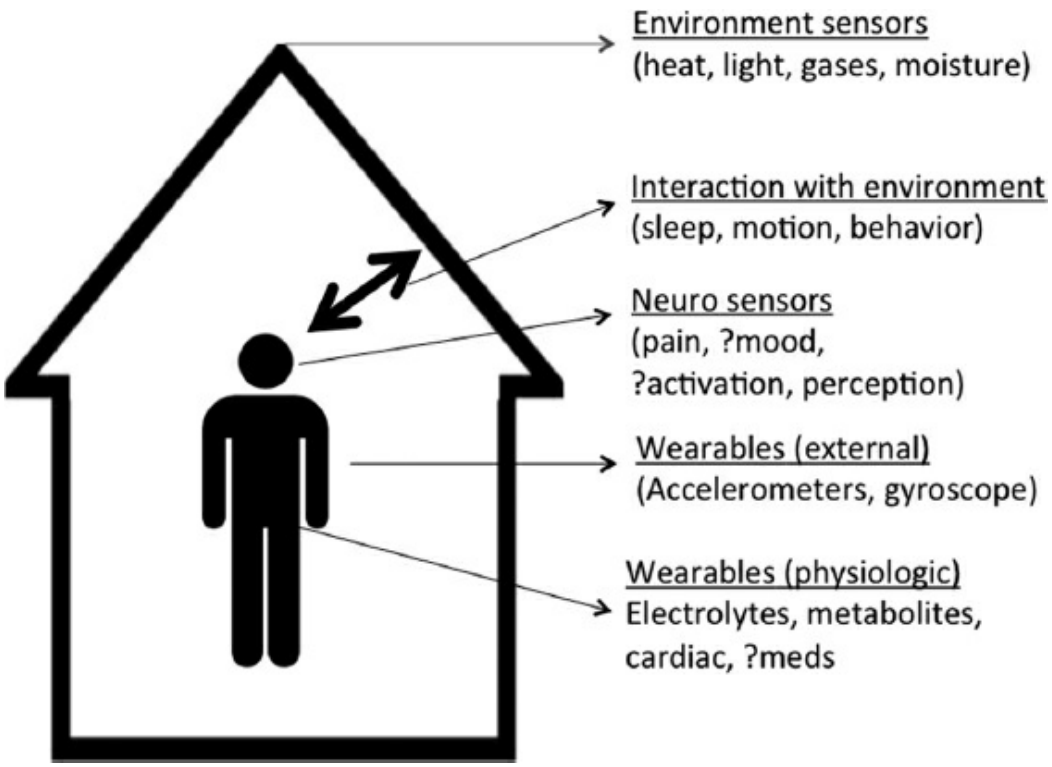

FIGURE 57.6h–1. Smart room. This figure indicates the multiple parameters of interaction between persons and their environment that can be quantified by sensors. This figure includes both existing and in-development sensor technology.

Games

Interest in clinical and cognitive applications of video games and game-based interventions continues to grow. Broadly, computerized games can be classified into three types—(a) games developed for, or played on commercially developed gaming systems such as the Sony Playstation, Microsoft Xbox, or Nintendo Wii, (b) commercially available cognitive games such as Lumosity or Posit Science modules, and (c) games developed for research purposes.

In one study conducted by Rosenberg and colleagues in 2010, researchers utilized the Nintendo Wii system, which relies on the persons using the system to perform physical tasks, to administer an intervention of physical exercise, conducted indoors, to assess improvements in subsyndromal depression in older adults. In a small sample of 19 patients, they noted that exercise using this game, performed for 35 minutes three

times a week resulted in statistically and clinically significant improvements in depressive symptoms. The researchers proposed that use of indoor gaming systems may provide incentives for older adults to engage in physical activity. Additionally, the use of the Nintendo Wii to play games that stimulate sports (e.g., tennis, golf) was found to burn approximately 50 percent of the calories that may otherwise be burned by playing the actual sport.

Several cognitive games have been marketed in the past 5 years, aimed at improving cognitive function in older adults who have experienced cognitive decline, and also at younger adults who are cognitively intact but invested in maintaining a high level of cognitive function. However, evidence of the efficacy of such games remains mixed. Despite early promise, large meta analyses of data from studies of such games indicate that at best, they may result in some domain-specific improvement in the cognitive abilities specifically targeted by each game. However, there is no evidence so far to suggest that such gains may be more generalizable to other domains.

There is also existing evidence of efficacy of a game that was custom developed to require multitasking. This game—a 3D racing game called Neuro Racer—was administered to older adults, and the authors noted that by performing the multitasking required by the game, older adults demonstrated benefits in multiple cognitive domains, including those such as sustained attention and working memory, which were not targeted by the game. Moreover, the resultant cognitive enhancement persisted for up to 6 months. This study provides a counterpoint to previous studies on cognitive games, and suggests that rather than targeting specific domains, games that tap multiple skills may be efficacious in improving cognitive outcomes.

Overall, use of a computer gaming interface has been found to be an acceptable and enjoyable method for a wide range of tasks and tests for older adults. Given the ease with which games can now be accessed on ICTs, this application is likely to grow rapidly over the coming years.

Tablets

According to the most recent available data, 18 percent of older adults use tablet devices. Unlike smartphones, which are most popular in younger adults aged 18 to 25, tablets may be more popular among middle aged and older adults. Older adults (65+) are the only demographic that is more likely to own a tablet or e-reader than a smartphone.

Given the larger screen and the consequent versatility of the touchscreen, tablets are increasingly being used for a variety of applications in older adults.

Several existing cognitive screening tools and other performance-based standardized tests are being adapted to iPad versions, including the UCSD

Performance Skills Assessment. Clinical applications, especially for comprehensive neuropsychological testing using tablets as the platform for administration are beginning to gain acceptance. A study by Mielke and colleagues has demonstrated that use of iPad tablets may be especially acceptable to older adults and that testing scores are comparable to those from regular pen-and-paper testing.

In addition to applications developed for delivery using tablets, there is evidence that simply learning to use a tablet device may be beneficial. Older adults are believed to be more receptive to learning these devices over computers, because of larger screens, intuitive interfaces, and touch screens. Learning to use an iPad may result in improvement in cognitive domains including episodic memory and processing speed, though there is no evidence it can improve visuospatial function. Programs created for teaching older adults to use technology, including tablets may also promote well-being by enhancing intergenerational activity and new learning. In a small pilot study, Kamat and colleagues have demonstrated that iPads may be effectively used to control behavior symptoms in older adults with dementia. In this study, the authors noted that because the iPad is a versatile device that can be used as a platform for a broad range of tasks, it may be an especially effective mechanism for administering nonpharmacologic interventions. Its primary advantage would be the ability to administer multiple interventions (e.g., reminiscence through pictures, music therapy, alternative engagement through videos, cognitive stimulation) using a single device.

An 87-year-old male of Russian origin, with a history of dementia and impulsive agitated behavior, typified by explosive outbursts was hospitalized on a geriatric psychiatry inpatient unit for management of behavioral symptoms. The patient was primarily Russian-speaking and struggled to communicate with staff. It was noted that the patient tended to be withdrawn, and kept largely to himself with minimal involvement in the treatment milieu. He was often irritable, though at other times he could be pleasant and friendly. A comprehensive medical assessment and workup was noncontributory. The patient frequently required two to three doses of antipsychotic medication (risperidone 0.5 mg) to control his episodes of agitation.

On one occasion, an iPad was utilized in one of the group activities, that involved showing patients how to access newspapers from their place of origin. Staff observed that the patient demonstrated a clear change in affect, when he was able to access digital content in Russian through this iPad, which was connected to the hospital's secure Wi-Fi network. Based on this observation, the physicians and clinical staff identified multiple apps that could help facilitate opportunities for the patient to get in touch with Russian language material online. This included Russian newspapers, blogs, videos, and other reading material. Staff also used a translation app and taught the patient how to use this to initiate communication. This gave the patient more opportunities to express himself. After this approach was initiated, staff noted a decrease in the patient's apathy and irritability, greater engagement in the milieu, and a reduction in the number of medication doses he required. It was hypothesized that at least part of his symptoms may have been triggered by a combination of boredom, lack of stimulation, and inability to communicate. The iPad provided a platform to address all three issues, and subsequent clinical impact was clearly observed by the treating staff.

Social Media

Older adults are the fastest growing segment of users of social media such as Facebook; this population is increasingly leveraging this technology to remain connected with family and friends. A 2013 survey by the Pew Research Center found that 45 percent of adults over 65 use Facebook. There is growing interest in creating Facebook interfaces that are simplified to facilitate use by older adults. Improving social connectedness among older adults can have a broad range of implication for enhancing well-being. The behavioral and psychological correlates of social media use in older adults warrant close empirical attention.

Caregiver Applications

ICT applications have the potential to be useful not only to older adults, but also their caregivers. These tools include peer networking, professional support, and resource identification. One innovative intervention for caregivers utilized a computer-mediated, automated, interactive system to reduce caregiver stress and provide assistance with managing disruptive behaviors exhibited by patients with Alzheimer disease. While these applications are promising given the rising number of people serving in caregiver roles, the outcomes associated with these technologies are yet to be fully characterized.

FUTURE DIRECTIONS

Robotics and Nanotechnology

With the rapid growth of technology, there has been a greater emphasis on developing robotic devices that can take over many physically demanding tasks of daily life. Promising ICT applications such as robotic-assistive devices for activities of daily living and other voice-activated technologies have been receiving greater attention for their potential use to reduce the physical demands of daily activities such as cooking, but further research is required to examine how they are suited for the needs of older adults. Robotic devices may even be implantable or flexible and wearable, and may play an assistive role in limb or organ function. Similarly, nanotechnology, which is growing rapidly, is considered a modality for intervention at the cellular level for aging-related disorders, especially those related to telomere dysfunction, reversal of oxidative cellular damage, and fabrication of stem cells and stem-cell–based therapies.

Big Data

Big data is a broad term used to loosely denote large volumes of data, whether collected in a structured or unstructured manner that can be mined and analyzed for information. The rise of big data has paralleled the development of sensor technology and rapid real time- analytics. As sensors allow one to collect extremely large volumes of behavioral and

other data in real time, big data analytics will be essential to allow for rapid, effective analysis of this data to generate information, and decision making. Big data analytics are also likely to play a key role in developing applications aimed at prediction of health events and preventative medicine.

BARRIERS AND LIMITATIONS

Although there have been significant technologic advances, older adults continue to experience barriers to using technology effectively. These include inadequate training to use the devices, lack of knowledge about the technologies available, inadequate design and equipment suitability, and cost/payment issues. Some of these are reviewed below.

Integration with Healthcare System

Despite rapid growth and interest in technology and aging, significant issues persist. For technologies that involve gathering personal health data, there remains incomplete integration with electronic health records. As a result, physicians often view such data (e.g., accelerometer data, sleep data) as burdensome as it involves the additional step of accessing, processing, and interpreting data. However, several companies that develop electronic medical records (e.g., Epic) are in the process of developing health apps that integrate with patients' electronic medical records. Manufacturers of commercial healthcare monitoring devices are also in the process of establishing ways for smoother integration of device data health with records.

Ultimately, it is likely that personal health data will be automatically integrated with electronic medical records as a routine procedure, allowing physicians to review both sets of data simultaneously.

Billing

The ICT applications for older adults hold great promise for the improvement in their quality of life and physical wellness. However, these technologies are often prohibitively expensive. In the United States, a large number of assistive technologies are not reimbursed by public or private insurance systems. The prospect of restricted payment hinders the development and widespread use of technology by those who can benefit from it the most. Congress and the Federal Government may promote the use of health technologies by incentivizing their use by providers and changing the policies regarding reimbursement through the Medicare program.

Literacy and Usability

Although ICTs have considerable potential for benefit in aging contexts,

literacy and physical limitations may pose a barrier to their utility in this population. Unsurprisingly, older adults' perceptions regarding the usability of ICTs appear to be related to their acceptance of technology. Factors such as perceived usefulness, perceived ease of use, and subjective norm have a strong impact on older adults' acceptance and use of ICTs.

Privacy and Ethics

Privacy concerns regarding the data generated by wearable and environmental monitoring systems have significant bearing on the motivation of older adults and their caregivers in adopting these technologies. The security of these data is often trusted to a third party and this may be associated with real or perceived disadvantages, which at present are not well characterized.

Information Overload

Although the large volume of data generated from wearable sensors and smart home devices is central to the appeal of these technologies, there are important issues that warrant consideration. The devices available tap different modalities (e.g., gait, skin conductance, diet, sleep) and provide heterogeneous longitudinal data. There is no consensus regarding the best methods to appropriately aggregate, analyze, and interpret the data collected. Episode Discovery, an unsupervised algorithm, has shown early promise in monitoring the health of elderly persons and identifying unusual or dangerous changes in behavior. Algorithms such as this are yet to be fully characterized with clinical populations.

Age-Related Impairments

Declines in working memory, attention, and processing speed commonly accompany normal aging. These age-related declines in cognition have important implications for technology access and use in older adults. For example, training older adults to use new platforms or technologic skills may be impacted by declines in working memory and attention, while their ability to discriminate and select targets from complex displays may be impacted by declines in information processing skills. These are important aspects that warrant close consideration during the development of technology for older adults.

Lack of Regulation and Evidence

Sensors and other healthcare devices remain both understudied and unregulated. Although the FDA is stepping up its effort at monitoring health sensor technology, significant work will be required to establish the safety, validity, and reliability of these devices, before they are widely deployed. A further issue is that despite a recent wave of technology

targeting older adults, the majority of these technologies are developed by technology firms and marketed without sufficient evidence to establish safety, efficacy, reliability, and validity of these devices to perform the tasks they are designed for. There is a great need for data on how these technologies are used, and how effective they may be. Geriatric psychiatrists have a great opportunity to be leaders in the area of establishing parameters for technology use and development in medicine.

References

AARP. Technology, tablets, social media and more - AARP [Internet]. 2015 [cited 4 June, 2015]. Available at http://www.aarp.org/home-family/personal-technology/

Armontrout J, Torous J, Fisher M, Drogin E, Gutheil T. Mobile mental health: navigating new rules and regulations for digital tools. *Curr Psychiatry Rep*. 2016;18(10):91.

Bandodkar AJ, Wang J. Non-invasive wearable electrochemical sensors: a review. *Trends Biotechnol*. 2014;32(7):363–371.

Ben-Zeev D, Schueller SM, Begale M, Duffecy J, Kane JM, Mohr DC. Strategies for mHealth research: lessons from 3 mobile intervention studies. *Adm Policy Ment Health*. 2015;42(2):157–167.

Chaudhuri S, Thompson H, Demiris G. Fall detection devices and their use with older adults: a systematic review. *J Geriatr Phys Ther*. 2014;37(4):178–196.

Czaja SJ, Lee CC. The impact of aging on access to technology. *Universal Access in the Information Society*. 2007;5(4):341–349.

de la Vega R, Miró J. mHealth: a strategic field without a solid scientific soul. A systematic review of pain-related apps. *PLoS ONE*. 2014;9(7):e101312.

Demiris G, Hensel B. Technologies for an aging society: a systematic review of. *IMIA Yearbook*. 2008;3:33–40.

Depp CA, Harmell AL, Vahia IV, Mausbach BT. Neurocognitive and functional correlates of mobile phone use in middle-aged and older patients with schizophrenia. *Aging Ment Health*. 2016;20:29–35.

Goldwater J, Harris Y. Using technology to enhance the aging experience: a market analysis of existing technologies. *Ageing International*. 2011;36(1):5–28.

Gustafson DH, McTavish FM, Chih M-Y, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. *JAMA Psychiatry*. 2014;71(5):566–572.

Kasckow J, Felmet K, Appelt C, Thompson R, Rotondi A, Haas G. Telepsychiatry in the assessment and treatment of schizophrenia. *Clin Schizophr Relat Psychoses*. 2014;8(1):21–27A.

Loh P-K, Donaldson M, Flicker L, Maher S, Goldswain P. Development of a telemedicine protocol for the diagnosis of Alzheimer's disease. *J Telemed Telecare*. 2007;13(2):90–94.

Mielke M, Machulda M, Hagen C, et al. Performance of the CogState computerized battery in the Mayo Clinic Study on Aging. *Alzheimers Dement*. 2015;11:1367–1376.

Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. *Gen Hosp Psychiatry*. 2013;35(4):332–338.

Or CK, Karsh B-T, Severtson DJ, Burke LJ, Brown RL, Brennan PF. Factors affecting home care patients' acceptance of a web-based interactive self-management technology. *J Am Med Inform Assoc*. 2011;18(1):51–59.

Poulin C, Shiner B, Thompson P, et al. Predicting the risk of suicide by analyzing the text of clinical notes. *PLoS ONE*. 2014;9(1):e85733.

Rabinowitz T, Murphy KM, Amour JL, Ricci MA, Caputo MP, Newhouse PA. Benefits of a telepsychiatry consultation service for rural nursing home residents. *Telemed JE Health*. 2010;16(1):34–40.

Redmond S, Lovell N, Yang G, et al. What does big data mean for wearable sensor systems?: Contribution of the IMIA wearable sensors in healthcare WG. *Yearb Med Inform*. 2014;9(1):135.

Rosenberg D, Depp CA, Vahia IV, et al. Exergames for subsyndromal depression in older adults: a pilot study of a novel intervention. *Am J Geriatr Psychiatry*. 2010;18(3):221–226.

Satariano WA, Scharlach AE, Lindeman D. Aging, place, and technology toward improving access and wellness in older populations. *J Aging Health*. 2014;26(8):1373–1389.

Soulas J, Lenca P, Thépaut A, (eds). Monitoring the habits of elderly people through data mining from home automation devices data. In: *Progress in Artificial Intelligence. Proceedings of the 16th Portuguese Conference on Artificial Intelligence, EPIA 2013, Angra do Heroísmo, Azores, Portugal, 9–12 September, 2013*. Berlin, Heidelberg: Springer; 2013:343–354.

Thompson WK, Gershon A, O'Hara R, Bernert RA, Depp CA. The prediction of study-emergent suicidal ideation in bipolar disorder: A pilot study using ecological momentary assessment data. *Bipolar Disord*. 2014;16(7):669–677.

Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? *PLoS Medicine*. 2013;10(2):e1001382.

Topo P. Technology studies to meet the needs of people with dementia and their caregivers a literature review. *J Appl Gerontol*. 2009;28(1):5–37.

Torous J, Summergrad P, Nassir Ghaemi S. Bipolar disorder in the digital age: new tools for the same illness. *Int J Bipolar Disord*. 2016;4(1):25.

Vahia IV, Kamat R, Vang C, et al. Use of tablet devices in the management of agitation among inpatients with dementia: an open-label study [published online ahead of print August 4, 2016]. *Am J Geriatr Psychiatry*.

Vahia IV, Sewell DD. Late-life depression: a role for accelerometer technology in diagnosis and management. *Am J Psychiatry*. 2016;173(8):763–768.

Vahia IV, Ng B, Camacho A, et al. Telepsychiatry for neurocognitive testing in older rural latino adults. *Am J Geriatr Psychiatry*. 2015;23:666–670.

van den Berg N, Schumann M, Kraft K, Hoffmann W. Telemedicine and telecare for older patients—a systematic review. *Maturitas*. 2012;73(2):94–114.

57.6i Positive Psychiatry of Aging

DILIP V. JESTE, M.D., AND BARTON W. PALMER, PH.D.

There is a growing recognition that medicine should encompass not only treatment of symptoms, but also promotion of well-being. For geriatric

psychiatry this means greater attention to promotion of "successful aging" as well as to other aspects of well-being among elderly people. The notion that mental health is more than the absence of symptoms or disorder is not new. In 1946, the World Health Organization (WHO) published a definition of *health,* which was not merely the absence of disease or disability, but rather "a state of complete physical, mental, and social well-being." While seemingly self-obvious, this definition is potentially revolutionary in its implications for the practice of geriatric as well as general psychiatry. Yet, approximately 70 years later, the implications of that revolution are only now beginning to materialize in psychiatry. Contemporary geriatric and general clinical psychiatric practice and training remain primarily restricted to diagnosis and treatment of mental illnesses, while psychiatric research mainly focuses on elucidating the underlying pathology and ameliorative interventions to treat and prevent relapses of mental illnesses. Yet, largely outside the field of psychiatry, and particular within the field of Positive Psychology founded in the late 1990s by Martin Seligman and colleagues, a growing body of research strongly suggests that positive psychosocial factors (PPSFs) such as resilience, optimism, and social engagement are associated with higher levels of subjective well-being as well as lower morbidity and greater longevity. Although many key questions still need to be addressed through empirical research, the last few years have finally brought together the beginnings of Positive Psychiatry as a potentially key aspect of 21st century psychiatry.

DEFINING POSITIVE PSYCHIATRY

Positive Psychiatry is the science and practice of psychiatry that seeks to understand and promote well-being through assessment and interventions aimed at enhancing behavioral and mental wellness. As a part of psychiatry, itself rooted in the field of medicine, Positive Psychiatry is grounded in biology, and ultimately seeks to determine the neurobiologic underpinnings of the PPSFs associated with positive outcomes, and to use that information to develop more effective biologic and psychosocial preventive, therapeutic, and rehabilitative interventions. As a branch of medicine, an important facet of Positive Psychiatry that at least partially distinguishes it from the complementary field of Positive Psychiatry, is that it is primarily focused on promotion of PPSFs and positive outcomes among those with or at high risk for psychiatric or physical illnesses.

HISTORY

In 1906, William James, a psychologist and physician by training, spoke of the need for a new approach to psychology that would study and apply principles underlying the success of the "mind cure." The concept of *mind cure* referred to the purported healing powers of positive emotions and beliefs. Although in retrospect the mind cure movement itself appears an

anachronism of questionable validity, James' emphasis on positive emotions and beliefs illustrates an early effort to include PPSFs during the emergence of psychology as an independent field. Yet these inclinations were largely neglected until the mid-20th century which brought the Humanistic Psychology movement of Abraham Maslow and colleagues. Their focus was on study and promotion of health and creativity among individuals. But that movement too failed to fully shape psychiatry, as well as mainstream clinical psychology. Instead, Positive Psychiatry traces its most immediate lineage and debt to the Positive Psychology movement initiated by Martin Seligman and colleagues in the mid to late 1990s. In his 1998 Presidential Address, Seligman spoke of the need for "a reoriented science that emphasizes the understanding and building of the most positive qualities of an individual: optimism, courage, work ethic, future-mindedness, interpersonal skills, capacity for pleasure and insight, and social responsibility." In the ensuring decades, Positive Psychology has grown and flourished into an international movement, but much of the research has focused on promoting well-being among the psychiatrically and physically healthy, and it has had less impact on the day-to-day practice of clinical psychiatry. This situation has begun to change. In his May 2013 address as President of the American Psychiatric Association Dilip Jeste commented that "I expect that the future role of psychiatry will be much broader than treating psychiatric symptoms. It will seek to enhance well-being of people with mental or physical illnesses. That is positive psychiatry." Positive Psychiatry is allied with and partially overlaps with Positive Psychology in concepts and goals, but also complements in that psychiatrists are first trained as physicians, an applied biologically rooted profession focused on treating diseases and disabilities.

POSITIVE MENTAL HEALTH OUTCOMES

There has been a growing recognition that medicine should encompass not only treatment of symptoms, but also promotion of well-being. Similarly, the goal of psychiatric interventions should not be just a reduction in psychopathology but also promotion of positive mental states. Some of the better known outcomes are summarized below.

Successful Psychosocial Aging

There is no single consensus definition of successful aging. A 2006 review of the empirical literature revealed 29 definitions in 28 published empirical reports. But there were some common themes. Traditional notions of successful aging have emphasized how much one has not aged, that is, the relative absence of physical, cognitive, and social disability. However, more recent studies focused on the perspectives of older adults themselves showing that self-perceived successful aging is not dependent on physical

health. Large proportions of older adults with chronic physical illnesses, who would not meet objective disability-based criteria of successful aging, perceive themselves as aging successfully. Qualitative studies indicate that older adults consider the ability to adapt to circumstances and positive attitude toward the future as more important than an absence of physical disease and disability. This distinction between physical disability and self-perceived successful aging may help explain a now well-documented paradox of aging. Specifically, although physical health and functioning, as well as some aspects of cognitive performance, tend to decline with increasing age, subjective well-being begins a steady improvement from youth or at least mid-life that continues into old age.

Other Positive Outcomes

Well-Being. Echoing the WHO definition of health, Positive Psychiatry construes *well-being* as representing the presence of pleasant positive experiences and states such as life-satisfaction and happiness (hedonic well-being), as well as eudemonic well-being, including a sense of purpose or meaning in life and self-acceptance. The concept of well-being has risen to prominence in recent years in economics, with self-reported well-being considered as a measure of national development, alongside traditional constructs such as the Gross Domestic Product. Well-being is associated with better health outcomes, and is correlated with reduced mortality/greater longevity.

Low Perceived Stress. The concept of stress has a complex history; the term may be used to refer to objective pressure or tension or the body's objective response to demands or change. In the context of Positive Psychiatry, however, the emphasis is on the individual's perception of stress which appears to be a better predictor of deleterious changes in some biomarkers such as telomere length than is objective stress. Perceived stress specifically represents the degree to which individuals perceive demands, problems, or challenges as exceeding their ability to cope with them. Recurrent unhealthy stress responses lead to cumulative wear and tear.

Posttraumatic Growth. The link between severe stress and posttraumatic stress disorder (PTSD) is self-evident, but not inevitable. Some people who experience significant traumatic events experience a form of personality growth that is referred to as *Posttraumatic growth*. It is characterized by an increased appreciation of life, altered priorities, closer interpersonal relationships, greater sense of personal strength, and/or recognition of new possibilities for one's life. Posttraumatic growth is fostered by the individual's resilience in the face of severe stress—it is not the opposite of PTSD, as they may be copresent within the same individual

in response to the same stress or trauma.

Recovery in Serious Mental Illnesses. Contemporary empirical data paint a brighter picture of the course of psychopathology and disability among older people with serious mental illnesses than traditional conceptualizations. For example, relative to their younger counterparts, older adults with schizophrenia tend to have better psychosocial functioning, including better adherence to medications and self-rated mental health, and lower prevalence of substance abuse and psychotic relapse. There is no consensus definition of recovery (e.g., there are differences of opinion regarding the need for full symptom remission), but there is general consensus that recovery involves decreased functional dependence and reduction of the impact of the illness on one's life. Elyn Saks has described this capacity as finding "wellness within illness."

Prevention. Prevention is one of the key positive outcomes in Positive Psychiatry. The focus on prevention has long roots in community psychiatry where there have been efforts directed at primary, secondary, and tertiary prevention. In recent decades, medicine and psychiatry are experiencing a paradigm shift from an almost exclusive focus on treating pathology to promoting health, as well as from caring for individual patients to enhancing health of communities through preventive measures. Prevention is no less meaningful in the context of geriatric psychiatry. A wealth of empirical data indicate that prevention in later life can be practical and cost-effective. Some examples including findings that treatment or prevention of depression following stroke, myocardial infarction, and/or bypass surgery results in significant reductions of morbidity and mortality. Likewise, interventions aimed at altering partially malleable risk factors such as social isolation, physical disability, and sleep problems can prevent or at least delay the onset of depression. Similarly, the risk of developing PTSD can be decreased with counseling older victims of natural disasters, wars, or tragic events.

POSITIVE PSYCHOSOCIAL FACTORS

A key tenet of Positive Psychiatry is that positive outcomes such as those described above are at least partially mediated by and can be fostered by enhancement of intrapersonal positive psychosocial traits as well as positive environmental/social factors, together subsumed under the rubric of PPSFs. Most of these factors are valued highly by the society, yet until recently they have been generally neglected in the biomedical sciences as imprecise, amorphous concepts that lack relevance to traditional medicine or to neuroscience research. There have been numerous studies in recent years suggesting that the PPSFs are tightly linked to human biology and are associated with positive health outcomes.

POSITIVE PSYCHOSOCIAL CHARACTERISTICS

PPSFs may have as strong (if not stronger) an impact on patient well-being as common health risk factors such as obesity, hypertension, sedentary lifestyle, and smoking, and even overall objective health status. Provision of effective interventions to strengthen PPSFs in older psychiatric patients could substantially increase the prevalence of recovery in seriously mentally ill older adults. Some of the key intrapersonal constructs (Positive Psychosocial Traits) and then Positive Environmental Factors are considered below.

Positive Psychosocial Traits

Resilience. *Resilience* refers to ready recovery from or positive adaptation to illness, depression, or other adversities. Resilience is associated with greater well-being and reduced mortality risk among older adults. People in their 90s who endorse higher levels of resilience have a 43 percent higher likelihood of living up to 100 years compared to their peers with lower resilience. Among people with medical illnesses, resilience is also associated with better treatment adherence, better self-care behaviors, and improved quality of life. Among physically ill patients, resilience is associated with improved physical health, health-related quality of life, self-care, and treatment and exercise adherence. Resilience is also associated with better emotional health and well-being.

Optimism. The term *optimism* refers to a characteristic of people who tend to expect good outcomes. Optimism is associated with lower risk of depression and higher overall subjective well-being, lower mortality, as well as less pain, better cardiovascular, oncologic, and other physical health outcomes, and better immune function. The precise mechanisms through which optimism has these effects is not fully clear but it may partially operate through its effects on health behaviors. Older adults with higher optimism are more likely than less optimistic age-comparable peers to consistently engage in healthy behaviors.

Personal Mastery and Coping Self-Efficacy. *Personal mastery* refers to an individual's expectations of his/her ability to achieve self-desired outcomes. A closely related concept is that of *coping self-efficacy,* that is, the individual's self-perceived ability to perform the requisite coping behaviors needed to overcome life challenges. Individuals high in personal mastery or coping self-efficacy have a belief in control over their life circumstances and future—that is, they have a general sense of control over things that happen to them and have greater confidence they can solve problems in their lives. In contrast, *optimism* more broadly denotes a tendency to expect positive outcomes regardless of whether or not one is

the agent of those outcomes. In the context of chronic life stress, personal mastery, and coping self-efficacy may promote specific, adaptive, goal-oriented behaviors, as well as maintenance of those behaviors, despite ongoing stresses. High levels of personal mastery tend to protect people from the potentially negative effects of stress on general health.

Wisdom. *Wisdom* is a construct with heuristic value in that it appears across most cultures and contrasts with factual knowledge or even intelligence. Wisdom may be defined as a complex trait with several domains including pro-social attitudes and behaviors (compassion, empathy, and altruism), social decision making or problem solving, capacity for self-reflection and understanding/insight, decisiveness, acknowledgment and acceptance of uncertainty, emotional regulation, tolerance of divergent value systems, openness to new experience, spirituality, and a sense of humor. What unites these components is their utility for the self (e.g., greater well-being) and for others (better relationship with others, and serving common good). Cumulative experience of life with age may promote some of the dimensions associated with wisdom. Some, although not all, studies suggest an increase in several domains of wisdom with aging that may be important for an older adult's ability to survive and even thrive in spite of worsening physical health. The so-called Grandma Hypothesis suggests that wisdom transmitted to younger generations may provide fitness advantage and help neutralize the evolutionary pressure associated with loss of fertility in old age.

Spirituality and Religiosity. Religious and spiritual practices are reportedly associated with greater well-being as well as better health outcomes throughout the lifespan. It may be argued that *religiosity* is better conceptualized under the umbrella of positive environmental factors (see next section) as the term *religiosity* is usually reserved in reference to involvement in organized religion, and thus has an inherent social connotation whereas spirituality can be personalized. However, religiosity too ultimately rests on the beliefs and behaviors/choices of an individual. The term *spirituality* refers to an internal experience, that is, the degree to which one's personal beliefs, thoughts, and behaviors focus on transcendent topics such as the meaning of life, belief in a higher being, power or level of consciousness, awareness, or existence. Religious and spiritual practices are associated with positive outcomes such as greater well-being as well as better health outcomes. These positive effects of religious practices may partially rest on the emotional support derived from the religious community. Religious and spiritual beliefs also tend to confer a sense of optimism. Spirituality has been reported to increase with age.

Positive Environmental Factors

In addition to intrapersonal Positive Psychological Traits, the interaction of the person with the environment and the nature of that environment also impact likelihood of positive outcomes. These environmental factors are described below.

Social Engagement and Support. *Social Engagement* refers to integration into a social network or social participation—that is, how well integrated one is into a social network. Within this construct is the number and quality of close relationships, frequency of socialization, and the degree to which one finds meaning or pleasure from social integration/interactions. A related construct of *social support* refers to the degree to which other people are available for emotional and physical support. Empirical research has shown that the quantity and quality of one's social relationships are associated with better mental and physical health outcomes, including lower mortality, as well as more specific health outcomes such as lower prevalence of hypertension, cardiovascular disease, dementia, depression, anxiety, and substance use disorders in young and old adults.

Family Dynamics. A particularly influential form of social engagement and support is the family. There is a large empirical literature showing that family dynamics and degree of a supportive family environment affect mental and physical health including symptom severity, treatment response, and recovery.

Other Environmental Factors. Access to quality physical and mental health care as well as environmental factors such as availability of safe parks and walking paths, community centers, access to grocery stores for healthy eating alternatives, can promote or discourage physical activity and exercise, nutrition, and stress reduction, especially in older adults.

BIOLOGY OF POSITIVE PSYCHIATRY

As noted above, psychiatry, as a branch of medicine, is rooted in biology, and so is Positive Psychiatry. Biology of relevant constructs involves neuroscience as well as genetics and epigenetics involving interactions among genes, environment, and structure and function of the brain. This type of research has implications for developing biologic interventions to enhance positive traits and impact mental and physical health. To date the amount of such research in relation to positive outcomes and PPSFs has been limited to specific aspects only, and there is an obvious need for additional investigations in this arena.

Potential Biomarkers of Positive Mental Health

Allostatic load and telomere length are among putative biomarkers of stress

and aging, but can also be conceptualized as potential biomarkers of well-being in reverse.

Allostatic Load. Allostatic load reflects the body's cumulative physiological response to frequent and sustained exposure to environmental stressors; it represents dysregulation of one or more biologic systems, multiple pathways to disease, and cumulative cost of antecedent challenges which together can serve as an indicator of morbidity, as well as affecting cognitive functioning and psychological factors. There is no single measure of allostatic load, but a commonly used composite includes systolic and diastolic blood pressure, body-mass index, serum cholesterol (with an emphasis on HDL), and plasma norepinephrine and epinephrine.

Telomere Length. Telomere length is an indicator of biologic age, and may represent a cumulative index of the number of cell divisions and a cumulative record of exposures to genotoxic and cytotoxic processes such as oxidation. Telomere length is maintained by the enzyme telomerase, and there is some preliminary evidence that healthy lifestyle changes as well as psychosocial stress reduction interventions can increase telomerase activity.

Oxidative Stress. Oxidative stress, caused by an overproduction of free radicals or an underproduction of antioxidant molecules, damages DNA, protein, lipids, and other macromolecules in many tissues, with the brain being particularly sensitive. Oxidative damage to cellular components by free radicals may be an important component of the aging process. Thus, prolonged or repeated exposure to oxidative stress may accelerate aspects of biologic aging, leading to greater comorbidity and mortality.

Inflammation and Immune Function. Neuroinflammation and immune function also relate to several PPSFs. For instance, low level of optimism has been reportedly associated with increased inflammation markers. Reduced personal mastery resulting from chronic stress may provoke excessive activation of the autonomic nervous system, which may in turn affect immune functioning. Several studies report direct associations of engagement in pleasant events with better health, including reduced plasma concentrations of inflammatory biomarkers and lower catecholamine response to stress.

Genetics and Epigenetics. Genes clearly play a role in determining the lifespan and health status in older age. Many PPSFs are also influenced by genetic factors. For example, resilience and optimism are both partially heritable. Beginning in development, an individual's genes and their

interactions with environmental factors may shape the neural circuitry and neurochemical function expressed in PPSFs. Various genetic polymorphisms affect a person's limbic reactivity and prefrontal-limbic connectivity. Adaptive functioning of fear, reward, emotional regulation, or social-behavior circuits underlies a resilient person's capacity to face stress. Stochastic epigenetic changes occurring during brain development, that drive resilience and hardiness, may affect response to periods of high stress.

About 20 percent to 30 percent of variability in human lifespan appears to be attributable primarily to genetic factors. Overall, longevity seems to be promoted by an absence of genes that predispose people to diseases, rather than by presence of "life-extending" genes. Risk for cognitive impairment in older age is clearly raised by presence of the ApoE4 allele, but this may be attenuated by higher educational attainment or greater "cognitive reserve." In short, genetic vulnerabilities may be amenable to modification through intentional or incidental environmental changes.

Neuroplasticity across the Lifespan

Empirical findings in recent decades have challenged previously held concepts about brain development in later life. There are now convincing empirical demonstrations that neuroplasticity, with associated brain growth and development, can continue into old age. Mechanisms underlying neuroplasticity include compensation for aging-associated degenerative changes (e.g., involvement of larger numbers of neuronal networks in the performance of an activity) as well as synaptogenesis, increase in the number of dendritic spines, and even neurogenesis in selected areas of the brain such as the hippocampal dentate gyrus and periventricular region. Such neuroplasticity is related less to the genes inherited than to the behavior, attitude, and environment which affect expression of those genes. It is this neuroplasticity that may enable PPSFs to affect an individual's well-being throughout the lifespan, although the brain is admittedly less plastic in later life than it is in childhood and early adulthood.

INTERVENTIONS

Research on interventions aimed at promoting PPSFs is in early stages, but seems highly promising. Treatment trials have shown beneficial effects of behavioral strategies buttressing resilience, changing attitudes toward aging, or incorporating meditation, on outcome measures such as well-being, as well as alcohol consumption, biologic response to stress, and mortality rates in older adults. There is considerable diversity in the types of psychosocial or behavioral interventions reported, which include writing gratitude letters, practicing optimistic thinking, acts of kindness, replaying

positive experiences, socializing, life review, mindfulness, cognitive behavior therapy, life coaching, goal training, and forgiveness therapy. Cognitive behavior therapy has been shown to be effective in enhancing optimism and facilitating reappraisal of traumatic events in a more positive light. Increasing optimism has been found to increase adherence to pharmacotherapy in formerly pessimistic individuals. Interventions aiming to improve social engagement, while not specifically targeting depression, appear to have antidepressant effects in older people. Studies of longer-term effects of meditation have reported reduced mortality rates in older adults at risk for negative cardiovascular outcomes.

Interventions to Enhance Healthy Aging

A key component of Positive Psychiatry in the context of working with older adults is the concept of successful or healthy aging. The quest for "antiaging" therapies and interventions has gone on for thousands of years, with as yet little in the way of "cure" for aging being found. But there are interventions that can delay the onset of disability and disease, and prolong "healthspan." In contrast to earlier assumptions, many of these interventions such as physical exercise appear to produce benefits even among older adults.

Lifestyle Interventions. Lifestyle behaviors are a logical place to start when thinking about enhancing positive outcomes in older people. Physical exercise, proper nutrition, activity participation, and regular health screenings and vaccinations are among the most powerful and modifiable predictors of the healthspan. In addition to the positive effect of healthy behaviors on the cardiovascular system, recent research has provided compelling data that improvements in brain structure and functioning can be attained from improving lifestyle habits.

Physical Activity. The axiom "what's good for the heart is good for the brain" has gained research support in recent years. There is strong evidence from randomized controlled clinical trials conducted among sedentary older people that increasing aerobic physical exercise can enhance performance on cognitive tests. Together, the empirical literature suggests that initiating exercise participation in later life may have significant effects that extend beyond physical outcomes, and include cognitive and emotional improvements.

Using animal models and examining the biologic effects of exercise, it is likely that physical activity has multiple pathways leading to improvement in cognitive and emotional health. By reducing hypertension, the risk for cerebrovascular diseases decreases. By reducing obesity, stress on the skeletal system is diminished. Physical activity may also have a secondary benefit of increasing social engagement. Fall risk may be decreased by

increasing muscle mass. Research from animal models suggests that at the cellular level, exercise may provide a mild stressor that may act as a kind of "inoculation" against other kinds of stress, and may promote neuronal growth.

The Centers for Disease Control and Prevention suggests that adults engage in 3 to 5 days per week of moderate intensity of exercise (equivalent to brisk walking) for 30 minutes or more; stretching every day, and strength-building activities 2 to 3 days per week. Only about one-third of Americans meet the minimum exercise recommendations, and the prevalence of meeting these criteria decreases with age. At particular risk for not engaging in recommended levels of physical exercise are older people from ethnic minority groups, who have lower educational attainment, or who have sub- or supernormal body mass indices.

A central challenge to increasing exercise participation is identifying the barriers to exercise in routine life. The "walkability" of the built environment is an important consideration for enhancing engagement in physical activity, especially for older adults because they appear to prefer walking over other exercises.

Diet and Nutrition. Diet and nutrition are clearly related to health and other positive outcomes, and no other strategy for healthy aging has received as much research support as caloric restriction. Caloric restriction has been examined in worms (*Caenorhabditis elegans*), rodents, primates, and humans. Reducing caloric intake has been associated with increasing the lifespan of certain strains of mice and primates. In addition to extending the lifespan, deceleration in other aging phenotypes accompanies caloric restriction. The exact mechanism of caloric restriction is unclear, but it may involve reducing the rate of metabolism, perhaps limiting cellular wear and tear. Whether and how calorically restricted diets can be safe and palatable in humans remains to be seen; however, there are ongoing studies examining the impact of caloric restriction in older overweight or obese adults.

The content of dietary intake is also of interest to promoting health, results from an epidemiologic study suggest a recommended "low-risk" diet that includes a higher ratio of polyunsaturated to saturated fats, high in cereal fiber and marine fatty acids, and multivitamins, low in glycemic load, minimal trans fats. Additionally, moderate alcohol drinking (one to two drinks per day), especially of red wine, appeared to be better in the Nurses Health Study. As with exercise participation, only a minority of older adults conform to such a diet, and a growing proportion of people, including older people, are obese. Indeed, the rising incidence of obesity in the United States may significantly slow the rate at which life expectancy increases over the next 30 years—that is, obesity may decelerate the dramatic gains that have been made in life expectancy. Practitioners

working with older adults need to provide dietary education, explaining the risks associated with unhealthy diets and encouraging healthier alternatives.

Biologic Treatments. The potential for biologic treatments in Positive Psychiatry is largely awaiting further research, but there does seem to be potential for advances in this arena. Recent research suggests that it might be possible to use neurobiologic approaches to enhance resilience in the context of anxiety and mood disorders. With increased understanding of the neural circuitry underlying resilience, new interventions may also be developed using findings indicating that individuals can be trained to modulate their own brain activity with the help of real-time fMRI-based neurofeedback. Biologic interventions may be developed on the basis of improved understanding of neuroplasticity as well as gene variations related to positive traits. In addition, findings from animal and human studies might eventually yield new pharmacologic agents that can maximize the adaptive function of the HPA axis as well as monoamine, neuropeptide, and other neurochemical stress response systems.

REFERENCES

Bangen KJ, Meeks TW, Jeste DV. Defining and assessing wisdom: a review of the literature. *Am J Geriatr Psychiatry*. 2013;21:1254–1266.

Carver CS, Scheier MF, Segerstrom SC. Optimism. *Clin Psychol Rev*. 2010;30:879–889.

Diener E, Chan MY. Happy people live longer: subjective well-being contributes to health and Longevity. Applied Psychology: *Health and Well-Being*. 2011;3:1–43.

Gage FH. Neurogenesis in the adult brain. *J Neurosci*. 2002;22:612–613.

Herrman H, Stewart DE, Diaz-Granados N, Berger EL, Jackson B, Yuen T. What is resilience? *Can J Psychiatry*. 2011;56:258–265.

Jacobs TL, Epel ES, Lin J, et al. Intensive meditation training, immune cell telomerase activity, and psychological mediators. *Psychoneuroendocrinology*. 2011;36:664–681.

Jeste DV. A fulfilling year of APA Presidency: from DSM-5 to positive psychiatry. *Am J Psychiatry*. 2013;170:1102–1105.

Jeste DV, Palmer BW. A call for a new positive psychiatry of ageing [Editorial]. *Br J Psychiatry*. 2013;202:81–83.

Jeste DV, Palmer BW, eds. *Positive Psychiatry: A Clinical Handbook*. Arlington, VA: American Psychiatric Publishing, Inc; 2015.

Jeste DV, Palmer BW, Rettew DC, Boardman S. Positive psychiatry: its time has come. *J Clin Psychiatry*. 2015;76(6):675–683.

Jeste DV, Savla GN, Thompson WK, et al. Association between older age and more successful aging: critical role of resilience and depression. *Am J Psychiatry*. 2013;170:188–196.

Knight T, Ricciardelli LA. Successful aging: perceptions of adults aged between 70 and 101 years. *Int J Aging Hum Dev*. 2003;56:223–245.

Mazzucchelli TG, Kane RT, Rees CS. Behavioral activation interventions for well-being: a meta-

analysis. *J Posit Psychol*. 2010;5:105–121.

Rasmussen HN, Scheier MF, Greenhouse JB. Optimism and physical health: a meta-analytic review. *Ann Behav Med*. 2009;37:239–256.

Rice VH. Theories of stress and its relationship to health. In: Virginia H Rice, ed. *Handbook of Stress, Coping, and Health: Implications for Nursing Research, Theory, and Practice*. Thousand Oaks, CA: Sage Publications, Inc.; 2000.

Sin NL, Lyubomirsky S. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: a practice-friendly meta-analysis. *J Clin Psychol*. 2009;65:467–487.

Southwick SM, Vythilingam M, Charney DS. The psychobiology of depression and resilience to stress: implications for prevention and treatment. *Annu Rev Clin Psychol*. 2005;1:255–291.

Stewart DE, Yuen T. A systematic review of resilience in the physically ill. *Psychosomatics*. 2011;52:199–209.

Uchino BN. Social support and health: a review of physiological processes potentially underlying links to disease outcomes. *J Behav Med*. 2006;29:377–387.

Vance DE, Brennan M, Enah C, Smith GL, Kaur J. Religion, spirituality, and older adults with HIV: critical personal and social resources for an aging epidemic. *Clin Interv Aging*. 2011;6:101–109.

58 Public Psychiatry

▲ 58.1 Public and Community Psychiatry

LEIGHTON Y. HUEY, M.D., ROBERT F. COLE, PH.D., AND JULIAN D. FORD, PH.D.

INTRODUCTION

The subject area of public and community psychiatry embodies a fundamental core of experience and tradition of this medical specialty. In the context of the upheaval and re-examination of American health care before and since the passage of the 2010 Patient Protection and Affordable Care Act (referred herein as the ACA), the experience of public and community psychiatry is poised to serve as a foundation for a transformation of behavioral health care.

Each term in the phrase public and community psychiatry is burdened by a complex past and needs interpretation to clarify its current meaning. The term public can refer to psychiatric programs, treatment, or institutions paid for by public funds or as objects of public policy, whether paid for or not. In recent decades, the scope has been narrowed to encompass chiefly populations of low-income adults disabled by persistent mental illnesses or substance use disorders, and for children, those with serious emotional disorders. These categories of public funding, represented in Table 58.1–1, identify target or priority populations that usually are defined by the presence of at least one qualifying major mental illness or substance use disorder, an assessment in which the illness or disorder is considered severe and persistent and affects the individual's ability to function in multiple areas. An evaluation of the individual's resources and ability to obtain private mental health services is also part of the assessment.

Public psychiatry stands in stark contrast to personal or private psychiatry, the latter being the entrepreneurial practice of delivering services to individuals having some means of purchasing them.

The care and treatment offered under public psychiatry are delivered in a variegated mosaic of inpatient and community-based services that are

more or less integrated into a coherent network sponsored by public agencies. Funding for public and community psychiatric services tends to be provided by federally legislated appropriations that are passed through state, county, and municipal government agencies (such as departments of mental health; substance use treatment services; children, youth, and family services; and public health, social services, education, adult corrections, and juvenile justice agencies). Ultimately, most public psychiatry and community services are provided by not-for-profit community mental health, substance use treatment, child guidance, social services, or health care organizations that either are funded by or subcontracted to government agencies. Thus, the very existence of public and community psychiatric services and the policies and resources determining how they are delivered are extremely dependent on legislative mandates and fiscal appropriations from all levels of public government. Psychiatry practitioners and administrators working in the public sector thus must be knowledgeable and savvy concerning the laws, regulations, and funding mechanisms relevant to their services. Increasingly there is recognition of the prevalence of substance use disorders, both as primary disorders (often in combination with mental disorders, i.e., "co-occurring" substance use and psychiatric disorders) and as secondary disorders (when mental disorder[s] are the condition of the first onset of primary concern, i.e., "comorbid" psychiatric and substance use disorders). While there has been some tortured historical debate on whether the more accurate phrases that should be used are "mental disorders" versus "behavioral disorders" or even "mental and behavioral disorders" as an inclusive term, this chapter will use "mental" and "behavioral" interchangeably. Finally, mention should be made of the significant prevalence of co-morbid physical disorders with mental and behavioral disorders which is true regardless of whether one is discussing "public" or "private" or "community" psychiatry. Therefore, this section specifically includes in "public psychiatry" not only the concept of "mental illnesses," but also that of co-occurring or comorbid serious substance use disorders.

The term community, as historically applied to psychiatry, connotes the mental health status of a specific, geographically defined population. It is synonymous in all respects with "not the state hospital" and is contrasted to institutional care, especially in the move to community mental health in the second half of the 20th century. Community psychiatry in this context refers to psychiatric practice in the community as part of the community effort to care for individuals with major mental illnesses and substance use disorders in a specific geographical catchment area. Community psychiatry initially was an attempt to break with traditional public psychiatry, which was seen as inextricably linked to hospital-based services. At the end of the 20th century, it obtained still broader connotation in terms of the health of the community: community psychiatry activities involve the effect of

psychiatric disorders in individuals on the health of the community and formulate treatment strategies to improve community health. Now, in the 21st century, with the decline of long-term institutional psychiatric care and the ascendance of in-home approaches to delivering integrated psychiatric and psychosocial services, community psychiatry is the dominant force in a de facto merger with public psychiatry.

The term psychiatry originally described the medical specialty focused on the care of persons with mental illness in the state hospital setting. Its domain was strengthened by the dominant role of the medical superintendents of the state hospitals (who would eventually be called "psychiatrists") and their medical colleagues in the latter half of the 19th century and first half of the 20th century. However, in the era of community mental health in the latter half of the 20th century, the term exploded to cover a host of disciplines and new personnel were recruited to staff the community-based supportive treatment systems that were supposed to take the place of the institutions. Nursing, social work, clinical psychology, various therapeutic specialties, and non-degreed direct care providers all became labeled as part of psychiatry. One of the most dramatic changes emerging since the late 1990s is the role of consumers of psychiatric services as providers in their own right through networks of peer-support services and advocacy initiatives. Family members of people with these disorders are also increasingly taking a more assertive role in treatment and support. In fact, a recent national report argued that the concept of "workforce" correctly includes both patients and families as members of the workforce, whether paid or volunteer.

Table 58.1–1.
Eligibility for Mental Health Service Grants

The following definitions are currently used to determine eligibility for services provided by the federal Center for Mental Health Services and by state public Mental Health agencies.

Adults with serious mental illness are persons

- 18 years of age and older
- Who currently or at any time during the past year
- Have had a diagnosable mental, behavioral, or emotional disorder of sufficient duration to meet diagnostic criteria specified within DSM-5
- Which has resulted in functional impairment that subsequently interferes with or limits one or more major life activities and
- Lasts less than 6 months of the past year

Adults with serious and persistent mental illness are persons who are

- Recently severely impaired, and
- The duration of the impairment totals 6 months or longer of the past year.

Children with a serious emotional disturbance are persons

- From birth to as old as 18 years of age
- Who currently or at any time during the past year
- Have had a diagnosable mental, behavioral, or emotional disorder of sufficient duration to meet

diagnostic criteria specified within DSM-5

- That resulted in functional impairment that substantially interfered with or limited the child's role or functioning in family, school, or community activities.

From the Federal Register.

This has broadened the traditional concept of public and community psychiatry, which previously had a narrow focus on the psychiatrist in his or her role as the medical specialist, trained in the discipline of medical science and exercising medical and legal authority to take defined actions on behalf of individual patients. In this paradigm, the public psychiatrist practices psychiatry in publicly funded services and carries out programs that are mandated by public policy. As a public and community psychiatrist, he or she is engaged with a specific geographically defined population using community programs for the care of its individuals with mental illness in order to improve the health of the community as broadly defined.

The experience of public and community psychiatry as it has evolved over the last two centuries in the United States, then, is a particularly valuable asset for the challenges facing US psychiatry as part of a health care system that has been characterized as chaotic and out of control. Three contemporary trends illustrate the importance of this basic (and rather narrow) conceptualization of public and community psychiatry to the daunting challenges facing US health care. The first is the realization that the effect and pervasiveness of psychiatric disorders imposes a global burden on society. The second trend has been the shift in funding for public psychiatric services from state expenditures, tied to the historic funding for state hospitals, to Medicaid, a health insurance model driven by the "medical necessity" for intervention. Finally, the power of reform efforts in general medical care and psychiatry have led to increasing demand for the use of evidence-based interventions defining quality, for safety standards, and for efficiency in the delivery of health care.

The first trend—the recognition of the burden of these disorders—can be most directly linked to a landmark study called the Global Burden of Disease. This study, funded by the World Health Organization (WHO), the World Bank, and several foundations, considered the leading causes of disability and premature death worldwide as a result of various health and situational problems. Of the ten leading causes of disability worldwide, five are neuropsychiatric, with the leading cause of disability being depression. When considering the leading causes of healthy years of life lost using the concept of disability-adjusted life years (DALYs), which reflects premature death and disability, depression was found to be second in the United States and was expected to rank second in the world by the year 2020. Mental health disorders were the fourth most costly health problem in the United States in 2000, 2004, and the most recent year for which data are

available 2008, according to the federal Agency for Healthcare Research and Quality. When the cost to society of trauma-related disorders (which include post-traumatic stress disorder [PTSD]) are added in—and exposure to traumatic stressors has been found to play a role in both causing and moreover severely worsening most psychiatric disorders—the cost to the United States as a society in 2008 was estimated at $138 billion, 50% more than the annual cost of heart disease and almost double the annual cost of cancer. This was true despite the fact that the cost per person was three to four times higher for heart disease and cancer than for mental health or trauma-related disorders, due to expensive high-technology and pharmacological treatments, and high rates of death or total disability. Many more persons, an estimated 50 million or more in the United States in 2008, suffer from mental health or trauma-related disorders than from either heart disease or cancer.

The broadly conceived "healthy city" and the Global Burden of Disease both have particular relevance to public and community psychiatry. The implications and the importance of properly addressing primary and comorbid psychiatric disorders across a spectrum as part of enlightened health policy are compelling. The prevalence of psychiatric disorders warrants a systems-based approach, which is the conceptual strength of public and community psychiatry. A systems-based model of psychiatry involves understanding and drawing on the wide range of biopsychosocial factors that underlie mental illness to develop psychiatric treatment and rehabilitation interventions to reduce the disease burden associated with severe and persistent mental illness. Recent estimates are that mental and behavioral health disorders are more common than previously thought, with about a quarter of adults experiencing one or more of these disorders, with a third receiving treatment, and only 12% of these, having treatment that would be considered adequate. Of the total number of Americans with mental and behavioral disorders, about 6% struggle with the most seriously disabling mental illnesses. Care for a significant number of these individuals is poorly coordinated and there is a recurring whorl of decompensation, emergency room visits, hospitalizations, various outpatient and sometimes residential programs, variable compliance and funding struggles with treatment, diagnostic confusion (e.g., schizophrenia vs. bipolar disorder, borderline personality vs. post-traumatic stress disorder vs. depression, etc.), polypharmacy with lack of attention to guidelines or to metabolic consequences leading to the iatrogenic production of physical health problems for these individuals who then enter the cycle multiple times. The early death of such individuals, primarily because of physical health problems, has been well documented. Many of these individuals end up in the justice system. The historical separation of psychiatry from the rest of medicine, partly self-imposed and partly extruded by mainstream medicine will hopefully be corrected as

delivery systems change and as greater knowledge comes to influence a more focused and effective interdisciplinary effort for the most afflicted in particular.

The second trend relates to financing. The role of Medicaid in shaping public services for people with mental illnesses is difficult to overstate. For state governments, the appeal was almost immediate. When the states assumed the onus of care for people with mental illnesses, largely through the mechanism of state institutions, 100 percent of those expenditures came from state budgets—competing with all of the other priorities facing state legislatures. Medicaid, which is a state–federal partnership, allows states to shift between 50 to 80 percent of their cost for covered services to the federal budget. Because of a provision in the original design of Medicaid, referred to as the "IMD exclusion," states were prohibited from using Medicaid to pay for services delivered in "Institutes for Mental Disease," namely state hospitals. So a policy movement toward community-based care was reinforced by a federal financing shift. For a large portion of the newly financed services, however, the role of the physician was highlighted because services had to meet the "medical necessity" threshold as certified by a medical doctor—typically primary care physicians, not psychiatrists, with the effect of limiting the scope of services for persons with more severe disabilities. More recent federal laws and regulations mandating parity in eligibility and reimbursement for mental health (compared to medical/physical health) have attempted to direct more services to more individuals with mental illnesses. As will be seen, recent analysis of Medicaid state expenditures has documented this trend, which has contributed to a dramatic increase of funding for mental health services in the last 50 years.

The third trend is a period of critical evaluation of US medical practice sparked by the Institute of Medicine (IOM) in its 2001 report entitled Crossing the Quality Chasm: A New Health System for the 21st Century. This is a call to American medicine to reform inadequate service delivery systems and to apply effective practices systematically. It documents the need for a fundamental overhaul of the entire US health care system, addresses the problems caused by insufficient quality in the health care system, and offers a road map for creating a new health system. A key principle in enhancing the quality of health care services is that treatments and practices must have a clinical and scientific evidence base in order to justify their utilization and costs. The Quality Chasm report was a primary influence among a series of reports from the IOM and other authoritative agencies in the two decades that laid the foundations for the health reform that has been codified in the Affordable Care Act (2010). Table 58.1–2 lists a selection from the hundreds of IOM reports and workshops, seminars, etc., that represent the ferment of these years that provided the content of the legislation and is setting the course of its implementation.

While these landmark reports were focused on inadequacies throughout the American health care system, public and community psychiatry qualifies as a primary target for the reform measures.

The connotation associated with public and community psychiatry has been synonymous with care for the poor and underprivileged, care for special populations normally not attended to by the private sector, overwhelmed systems, low reimbursement and inadequate resources, inferior and unresponsive care, and a final common pathway for individuals no longer able to afford private care. This is not for want of their providers' dedication: it is the paucity of the providers' numbers and the general public neglect of this sector. The patients of public and community psychiatry as a group lack the most basic resources to sustain themselves (food, clothing, housing, financial support, social support, etc.), and they are arguably the group least likely to be able to advocate for themselves and to have the capacity to recover from this default position. The needs of the patients, the inadequacies of the delivery systems, the lack of an adequate workforce, and the lack of coverage make the sector of public and community psychiatry a poster child for the attentions of the ACA reforms.

Table 58.1–2.
Precursor Authoritative Reports for 21st Century American Health Care Reform

Managing Managed Care: Quality Improvement in Behavioral Health. Washington, DC: National Academy Press; 1997.
Mental Health: A Report of the Surgeon General by David Satcher, United States. Public Health Service. Office of the Surgeon General. Center for Mental Health Services. Washington, DC: National Institute of Mental Health; 1999.
America's Health Care Safety Net: Intact but Endangered. Washington, DC: National Academy Press; 2000.
Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.
The Future of the Public's Health in the 21st Century. Washington, DC: The National Academies Press; 2002.
Research Training in Psychiatry Residency: Strategies for Reform. Washington, DC: The National Academies Press; 2003.
Key Capabilities of an Electronic Health Record System. Washington, DC: The National Academies Press; 2003.
Academic Health Centers: Leading Change in the 21st Century. Washington, DC: The National Academies Press; 2003.
New Freedom Commission on Mental Health: Achieving the Promise: Transforming Mental Health Care in America, Final Report (DHHS Pub. No. SMA-03-3832). Rockville, MD: U.S. Department of Health and Human Services; 2003.
Improving the Quality of Health Care for Mental and Substance-Use Conditions: Quality Chasm Series. Washington, DC: The National Academies Press; 2006.
Graduate Medical Education That Meets the Nation's Health Needs. Washington, DC: The National Academies Press; 2014.

The reform is opening up a new community paradigm for American health care. It is a shift away from solo or corporately based practice models; emphasis on specialty care to the neglect of primary care; awkward payment methodologies that encourage volume over value and mechanical reliance on defined procedures; little accountability for population health; and a strange separation between main stream private health care and the care of the poor and disabled. It is a shift toward the public health emphasis on education, prevention, early identification and the health of populations; emphasis on primary care and the patient-centered medical home; physicians working with multidisciplinary teams; effective and evidence-based care; organization of accountable clinical entities meeting the individualized needs of each patient. Public and community psychiatry will benefit particularly from the reform. There will be a shift from the psychiatrist as solo therapist to be integrated into the multidisciplinary primary care team where it will be possible to register the earliest identification of developing mental health issues and to provide selective care for emerging psychiatric symptoms, while avoiding the stigma that has been assigned to mental health care. Primary care will also provide an organizational base for chronic illness care for patients burdened with disability and special treatment systems that can sustain a constant and continuing clinical process that will bring these patients to a degree of recovery.

Reform will lead to profound changes for public and community psychiatry. To take full advantage of this period of reform, it will be important to understand its past and the factors that shaped it development and its culture. American public and community psychiatry began in colonial times over two hundred years ago and was one of the first public health programs. It began in the village poorhouses, moved to asylums and stayed there until the middle of the 20th century when the state hospitals were replaced with uneven efforts at community care. Mainstream American medicine, as it was fortified by its use of the natural sciences, leaped ahead of psychiatry in the 20th century, but, in recent decades, psychiatry rapidly catching up as a science of the brain has become more accessible. The concerns of public and community psychiatry, as well as private sector psychiatry, have been prominent as the groundwork of the reform took shape. Before turning to the particulars of the reform agenda, a quick review of its history is appropriate.

HISTORICAL PERSPECTIVE OVERVIEW

Mental health policy and treatment has been a public issue in the United States from the beginning. A broad brush overview of this complex history can distinguish four stages. Colonial times and the early years of the Republic were the era of the poorhouse, during which the great majority of persons with mental illness found refuge or were confined for their own

and others' protection as part of local community management. Care and custody were the operative functions, and treatment was nonexistent. In the 18th century, the ideals of the Enlightenment formed the basis for a paradigm shift toward fostering a better life for persons with mental illness rather than condemning and isolating them, ushering in an era of "moral treatment" (aimed at restoring the capacity for ethical conduct) through a return to the pastoral rural environment of the pre-industrial age ("the Retreat"). In the mid-19th century, the protection and moral rehabilitation of the mentally ill was scaled up and institutionalized through the pioneering work of Dorothea Dix, resulting in an age of the asylum, during which public inpatient facilities, generically called state hospitals in the United States, were constructed in every state. A century later, advances in medicine, psychodynamic psychiatry, biological psychiatry, and cognitive behavioral psychology along with a remarkable policy consensus moved federal legislation to shift from institutionally based care to an era of community mental health. Finally, as health care costs began to soar at the end of the 20th century, and as employer purchasers and government moved to create market-driven health care, mental health policy moved conceptually to tightly organized systems of care that focused on prevention, early intervention, and intensive community care for children with severe emotional disturbances and their families and community support and recovery for adults with serious, acute, and persistent mental illnesses. The policy and concept, although ideal, remain partially operational with the smooth provision of care constantly jostled by bureaucratic, regulatory, and economic potholes.

The movement of history is best understood at its transition points. To get a sense of this history, the following quick review of public and community psychiatry in the United States focuses on three vignettes at the transition points that separate these stages.

From the Poorhouse to the Asylum

> Now for near five years I have been indulged by the gracious Heaven in my long holiday in this goodly house of mine entertaining and entertained by so many worthy and gifted friends and all this time poor Nancy Barron the madwoman has been screaming herself hoarse at the poorhouse across the brook and I still hear her whenever I open my window. — Ralph Waldo Emerson, Journals, 1840.

In the American colonies and in the early days of the Republic, the care of the mentally ill was modeled on the practices of 18th-century England, where the Enclosure Laws disrupted the rural social order and set thousands on the road to seek shelter and employment in the cities. People with mental disabilities were at a particular disadvantage. They were saved from vagrancy by confinement in poorhouses, prisons, and jails, along with criminals, debtors, and other people with disabilities. In the American colonies, most towns would establish a poorhouse, and persons with

mental illness, especially those who were found to be disruptive, would be committed by the actions of local government to extended stays in them. It is not that there were no hospitals. The construction of a public hospital was first proposed in Virginia in 1766. The money was appropriated in 1770, and the Public Hospital at Williamsburg opened in 1773, which was the first public mental hospital in the colonies. In Massachusetts, McLean Asylum opened in Charlestown in 1818. The first large state hospital opened in Worcester, Massachusetts, in 1833.

The physician superintendents of these early hospitals were the reformers of the era, following the ideas of moral management put forth by the French reformer Philippe Pinel. Pinel believed that if people were removed from their urban environments, relocated to an asylum in a beautiful pastoral setting, engaged in agriculture and the trades, lived in close proximity with staff, and were treated with kindness, then soon they would fully recover from the behavioral changes that they manifested as a result of having lived in a stressful, demanding environment. Benjamin Rush, a signer of the Declaration of Independence and considered the founder of American psychiatry, applied many of these theories and methods of his own in the treatment of the insane in the years after 1783 in the Philadelphia Hospital.

The great majority of those who experienced mental illness, however, were kept in poorhouses or jails under deplorable conditions and became a cause for reform. The Unitarian reformer, Dorothea Dix, had begun a 2-year survey of the conditions in which these people lived. In 1843, she compiled her findings in a Memorial to the Legislature of Massachusetts: "I tell what I have seen, to present the state of the insane persons confined within the Commonwealth in cages, closets, cellars, stalls, pens! Chained, naked, beaten with rods, and lashed into obedience!"

After the Massachusetts Legislature appropriated funds for a substantial expansion of the Worcester Hospital between 1844 and 1854, Dix traveled more than 30,000 miles, visiting hundreds of prisons, almshouses, jails, and hospitals, collecting data and arousing public opinion to her cause. In 1848, with the example of federal support given to education, she sought federal support for national hospitals. President Franklin Pierce, who stated, "I cannot find any authority in the Constitution that makes the Federal Government the great almoner of public charity throughout the United States," and vetoed the bill for federal support. Public mental health was, for the next century, exclusively in the hands of the states and the state hospitals that continued to be built only to be quickly filled and become hopelessly overcrowded.

Community Mental Health

I am proposing a new approach to mental illness and to mental retardation. This approach is designed, in large measure, to use Federal resources to stimulate state, local, and private action.

When carried out, reliance on the cold mercy of custodial isolation will be supplanted by the open warmth of community concern and capability. Emphasis on prevention, treatment, and rehabilitation will be substituted for desultory interest in confining patients in an institution to wither away. — John F. Kennedy, Message from the President of the United States Relative to Mental Illness and Mental Retardation, February 5, 1963.

In a little more than a century after President Pierce blocked the federal government from involvement in the problem of mental illness, another president reversed this direction and committed the federal government to the cause of persons with mental disabilities and their families. President John F. Kennedy signed the Mental Retardation Facilities and Community Mental Health Centers Construction Act (PL 88–164) in October 1963. He did so perched on a wave of reform that defined a new era of community mental health. The reform was founded on a century of steady progress in which asylums were built and their weaknesses discovered. The advances of medicine offered new hope to psychiatry, and that optimism gathered the momentum of public opinion. Consumers and advocates for the mentally ill found their voice in the mental hygiene movement. A series of research-oriented institutions was established. Significant therapeutic improvements were discovered. Finally, a coalition of skillful bureaucrats, advocates, and politicians set the scene for the public policy shift that President Kennedy declared would supplant the "cold mercy of custodial isolation" with the "open warmth of community concern and capability."

It has been said that US medicine reached a critical juncture toward the beginning of the 20th century. Science and medicine had established a strong alliance in the major university medical schools, and the profession had established a powerful public credibility as breakthrough after breakthrough transformed medicine. Psychiatrists were not in the forefront of this transformation, but they were profoundly affected by it. One of the most prominent psychiatrists of the period, from 1890, until his retirement at the eve of World War II, was Adolf Meyer, and his life and career may be thought of as representative of what was happening to psychiatry. Meyer was born in Switzerland in 1866, and was trained as a neurologist. He came to the United States in 1892, and worked for 8 years in state hospitals in Illinois and Massachusetts. In this sense, he came from the ranks of the asylum psychiatrists who were the bulwark of the state hospital system that had, by that point, spread throughout the country.

The first public psychiatrists were state hospital superintendents, responsible for all institutional clinical and administrative functions. Within a few short years, the populations of state hospitals grew so that their treatment function was eclipsed by care and custody concerns. The superintendent became more of a manager and administrator than a treatment provider. The ability to organize and to bring order to the institution became a necessary skill.

The New York State Hospital system had established, in 1895, the

Pathological Institute, which had two mandates: the study of mental diseases from the standpoint of cellular biology and the offer of instruction in brain pathology to state hospital physicians. The Institute sought to become the cutting edge of reform—not without opposition from the asylum psychiatrists, who were more oriented toward managerial and administrative concerns and clear demarcation of the diseased state of institutional inmates from the normality of functioning people in the community. Meyer became its director in 1901. When Henry Phipps endowed a psychiatric clinic at Johns Hopkins, Meyer was chosen as its first head. From Johns Hopkins, he played a leading role in American psychiatry for more than three decades.

In the beginning of the 20th century, the concept of dynamic psychiatry opened the possibility of psychiatry practiced in the community. Dynamic psychiatry blurred the distinction between health and disease and saw mental illness on a continuum. This allowed the center of attention for psychiatric practice to move beyond the institutions and to include troubled individuals in the early stages of illness. This approach elevated the significance of the life history and prior experience of each person. It was a short leap of faith to suggest that early community treatment might prevent the onset of severe mental illnesses that, up to then, had required institutionalization. The once strong bond between the asylum and psychiatry was unraveling. Dynamic psychiatry expanded the jurisdictional boundaries to include psychologically troubled individuals, as well as allegedly dysfunctional social structures and relationships. From 1880 to 1940, psychiatrists identified new careers outside of institutions. Psychiatrists articulated novel and sometimes bizarre theories and therapies; expanded jurisdictional boundaries to include not only mental disorders, but also the problems of everyday life; proposed solutions for society's ills; and defined a preventive role. In retrospect, it was an over-reach which subsequently led to questions about the credibility of the field and its foundations. What had been hypotheses initially were interpreted by some as factual, and often damaging (e.g., the so-called "schizophrenogenic mother," asthma as "a cry for help," etc.). It was also during this "pre-chlorpromazine era" that psychiatry alienated itself from the rest of medicine. Despite its good intentions, it was, in retrospect, gross overstatement based more on conjecture and opinion than a scientific basis.

Meyer was a leading figure in advocating "a fresh and courageous pragmatism" and held "a biological conception" of human beings: "We must ... accept the statement that all mental activity must have its physiological side and its anatomical substratum." He developed a genetic-dynamic or developmental approach to mental illness, which he named psychobiology. He stressed the interaction of the organism with its environment and defined mental disorders in behavioral terms, tracing

their origins to defective habits. He recognized the importance of consumer advocacy and supported Clifford Beers, a former mental patient, who wrote a celebrated work, A Mind That Found Itself, and who founded the Mental Hygiene movement.

Meyer was an empiricist, urging his colleagues to assemble and to catalogue facts. He was not a theoretician or a systematizer. Perhaps he realized that biology did not yet have the scientific tools to construct adequate theories about human mental processes. In this regard, he did not favor Sigmund Freud and what he viewed as the tendencies to dogmatic pronouncement. The eminent social historian of American mental health policy, Gerald Grob, observed that "psychoanalytic concepts ... had far greater influence on culture and thought than on medicine in general or psychiatry in particular." This seems to overstate the case, given the great enthusiasm for psychoanalysis that came after World War II and that dominated popular perceptions of psychiatry.

On the eve of World War II, as Meyer retired, public psychiatry had advanced on three fronts. On the first, research institutes and psychopathic hospitals were developed alongside the state asylums in which research and training would benefit the care of persons with mental illness. These hospitals and institutes had limited success, in that their research efforts to establish the efficacy of psychodynamic treatment progressed slowly, and they suffered a degree of mission drift away from the seriously and persistently mentally ill in the asylums and toward less severe mental health problems in the community. Second, an active consumer advocacy movement began in the founding of the National Committee for Mental Hygiene in 1909, which developed into a national movement focusing first on the plight of individuals in the asylums and then on mental health issues in the community. Finally, the ideas of dynamic psychiatry made it possible to think of the various stages of mental illness and the possibility of the prevention, amelioration, and potentially even cure of mental disease.

Before World War II, there were fewer than 3,000 psychiatrists in the United States, with more than one half of them working in the public sector. At war's end, the American Psychiatric Association (APA) still represented hospital psychiatry. However, a bitter and divisive power struggle ensued. By 1950, the APA had been transformed into a visible and active organization that represented psychodynamic and psychoanalytic viewpoints. Adherents to these perspectives controlled virtually all university departments, ensuring that psychiatrists trained in the 1950s and 1960s shared these concepts. By 1957, only 1,700 of the 10,000 APA members worked in state hospitals. Within less than 20 years, the overwhelming majority of psychiatrists in the United States went from working in state hospitals to working in other settings that had little or no interface with the populations having serious mental illness or serious and persistent mental illness or with children with serious emotional disorders.

As World War II ended, two factors entered the US public psychiatry arena that would have a catalytic effect. The first was the corps of 2,400 physicians assigned to psychiatry along with prominent leaders in US psychiatry who had screened recruits and successfully treated neuropsychiatric symptoms in noninstitutional settings. Many psychiatrists left military life convinced of the need for change in the ways in which the needs of persons with severe mental illness and individuals with mental health problems were addressed. They entered the nonmilitary world confirmed in the ideas that dynamic psychiatry had articulated. The second factor was the appearance of the so-called "French drug" chlorpromazine (Thorazine). Chlorpromazine was tested in a Parisian mental hospital after encouraging findings of its effect on trauma patients in the war. The positive findings of these trials led to its marketing under the name Thorazine in the United States in the early 1950s. The new drug was cautiously and ambivalently received by psychiatrists committed to psychodynamic concepts as there were comments that the drug would alter the transference of the therapeutic relationship, but it opened a new world of possibilities for the academics and researchers who came to represent biological psychiatry.

The 1950s were a time of great therapeutic optimism. Biological psychiatrists pointed to the efficacy of drug therapy; their psychodynamic colleagues could point to psychological, psychosocial, and environmental therapies that would reach the most severely disabled institutionalized patients. Milieu therapy (the therapeutic community) arose in this period and had much in common with moral therapy. It appeared that medications were going to allow patients to engage in the therapeutic community, and then other innovations, such as unlocked wards and partial or day hospitals, were going to help some chronic patients transfer to the community.

As all of this momentum for change was building, a remarkable partnership for political action was developing among Robert Felix, the head of the newly formed National Institute for Mental Health, Senator Lister Hill, and Representative J. Percy Priest. Hearings were held aimed at focusing public attention on mental health problems and thus strengthening the case for an enhanced federal role. The Mental Health Study Act was passed in 1955, creating the Joint Commission on Mental Illness and Mental Health (JCMIMH). That Commission's final report, Action for Mental Health, did not make a clear-cut distinction between individuals with serious and persistent mental illness and psychologically troubled individuals. Although the JCMIMH report was a compromise among the 26 participating constituency organizations, nonetheless there was strong support for fundamental change. Faith in the effectiveness of treatment in community programs was based on belief in the efficacy of early intervention, psychodynamic theories, epidemiological efforts to

elucidate the prevalence of mental illness in the community, and the putative identification of environmental variables.

At the end of the decade, with a new administration with an aggressive domestic agenda, the National Institute of Mental Health (NIMH) convened the Interagency Task Force on Mental Health, which developed the Blueprint for Action and, more importantly, activated the political coalition that made President Kennedy's initiative possible. The final policy sidestepped the provisions recommended by the JCMIMH for the refurbishment of the state hospitals and the provision of acute institutional care for the most severely disabled and devoted its attention to community mental health clinics, which it recommended to be enhanced to comprehensive community mental health treatment centers—the community mental health centers (CMHCs).

The national program for comprehensive CMHCs was in place. For the next 15 years, grants flowed to the localities and then through the states to the localities to the new centers, and states proceeded to build two community systems while beginning the process of phasing out the old state hospitals. The immediate discovery was that the CMHCs were able to expand mental health services for the communities but were not at all prepared to care for persons with serious and persistent mental illness who had spent most of their lives in state institutions. The discharge of 90 percent of the state hospital population (the deinstitutionalization process) without adequate community resources available was chaotic and often precipitous. It led to a class of homeless and generally untreated or transinstitutionalized individuals (put into nursing homes, prisons, squalid rooming houses and hotels, etc.) who deteriorated outside the state hospitals. There also arose a peripatetic subpopulation with serious and persistent mental illness who often had comorbid substance abuse and generally avoided any beneficial involvement in treatment for a significant portion of the early phase of their illness. However, the states, to varying degrees, were set on a course of dismantling state hospitals that were becoming increasingly expensive. As federal staffing grants to the CMHCs were phased out and state contract dollars replaced them, the states gradually narrowed the focus of the community programs to its priority clients—the former residents of the state hospitals. Special programming models were developed to provide supportive programming for them that focused on community support and rehabilitative approaches rather than clinical ones.

As community mental health programs developed, it became clear that psychiatrists were, for the most part, playing limited roles: reimbursement could not support them full time, and they gravitated to private practice models that were appropriate for individuals with less severe mental health issues. The psychiatrists who were involved in the early CMHCs were quickly outnumbered by practitioners from the other disciplines and often

sidelined themselves in private practice models in which they could treat educated and functional people with relatively mild disorders. One tongue-in-cheek publication during this era described the "YAVIS" individual as the "ideal" psychotherapy candidate … Young, Attractive, Verbal, Insightful, Successful. Many psychiatrists and others in behavioral health gravitated toward this profile for their ideal caseloads. These practice models limited their ability to treat persons with severe mental health disabilities without appropriate outreach and supportive care.

It was a chaotic period in US mental health services. The different states advanced unevenly in the development of community programs while, uniformly, political pressure built to reduce the census in the expensive state hospitals. In some cases, appropriate community-based programming was ready to receive the newly discharged individuals, but in many cases it was not. There are those who have come to look back in dismay at the massive changes that the federal community mental health programs put in motion. The durable and revered concept of the asylum was to be abandoned in this process. With the picture of many deinstitutionalized former state hospital residents, homeless or living marginally, especially in the nation's cities, the optimism and faith in the effectiveness of the new treatment modalities that drove these policy decisions might be judged (in hindsight) to be unwarranted. The process had been driven by rhetoric, single-mindedness, and unbridled enthusiasm, and, strangely, the NIMH had ignored the more balanced recommendations that supported improvement measures for the hospital facilities and the community programs.

Many mistakes were made, but many valuable lessons were learned. In 1977, President Jimmy Carter signed an executive order creating the President's Commission on Mental Health to review the progress and needs of the national programs and to make recommendations. The reports and monographs prepared by the Commission documented and catalogued what had been gleaned from the development of community programs and the process of deinstitutionalization that had ensued. Overall, they represented a more balanced and informed conceptualization of the agenda and the problems that needed to be addressed. The Commission's findings led to the passage of the Mental Health Systems Act (PL 96–398), which President Carter signed in October 1980. It enabled a National Plan for the Chronically Mentally Ill, which was released a month after Carter's defeat for re-election. Under his successor, President Ronald Regan, the Omnibus Budget Reconciliation Act of 1981, repealed most of the Mental Health Systems Act in favor of block grants to states. The role of federal leadership and national policy was substantially diminished, and the states were largely on their own.

Organized Systems of Care

The psychiatric hospital is a special target of health care corporations. Its profit margins are large, investment costs are low, inpatient psychiatric care is now widely insured, and the complexity (and perhaps ambiguity) of psychiatric diagnosis and treatment makes cost control efforts by insurers and the government difficult. In 1968, for-profit hospital chains were nonexistent in psychiatric care. In 1980, investor chains owned 25 percent of private psychiatric hospitals and, just two years later, these corporations owned a full 43 percent market share. Between 1982 and 1983, the number of investor-owned psychiatric hospitals rose 42.5 percent, from 106 to 151. The increase was partly attributable to investor groups' acquisition of existing facilities, but new construction of psychiatric hospitals, which grew 53.3 percent, accounted for most of the growth. "Ad-Psych" (Adolescent Psychiatry) units are so lucrative that corporations have closed units providing other types of services in order to establish the more profitable adolescent units. Admission data reflect the current domination of the private sector in the provision of adolescent inpatient psychiatric services. In 1980, for example, private facilities admitted 61 percent of all adolescents who underwent psychiatric hospitalization; since that time, juvenile admissions to these facilities have increased substantially. — Lois A. Weithorn, Stanford Law Review, 1988.

US health care costs began to increase dramatically in the 1970s. By the end of the 1980s, they were spiraling. The article just cited by Lois Weithorn in the Stanford Law Review addressed the possible violation of children's rights posed by their involuntary hospitalization in commercially operated psychiatric hospitals. It documented a trend of which US employers—and the insurance companies that they hired to administer their health benefits—were all too aware. While employee health care costs overall were climbing, behavioral health expenditures were increasing by 50 percent each year. Venture capital had discovered a chink in the armor of health care finance. Even though there were lifetime limits for mental health benefits, a substantial amount of money lay in those coffers. Psychiatric hospitals were relatively cheap to build and could be designed to be converted to elderly housing facilities once the gap was closed. The criteria for mental illness were sufficiently permeable and standards of care were sufficiently undefined that decisions for hospitalizations were made on behalf of many patients when more considered judgments would have found other alternatives. In fact, many hospitals launched deliberate marketing campaigns to fill their beds, and many practitioners participated. The trend was an abuse. Its most cynical aspect was that benefit dollars that could have been devoted to the care of persons with severe mental illness were quickly extracted—a $50,000 lifetime limit could be spent in 60 days of private hospital care—so that sick patients were discharged to the care of public sector agencies.

Although several examples of health maintenance organizations (HMOs) had demonstrated their effectiveness in improving the health status of enrolled populations for several decades, US employers were reluctant to require their employees to sign up for a plan instead of the traditional indemnity (or fee-for-service) health insurance. However, the growth of behavioral health expenditures was so dramatic that they overcame their reluctance for this sector. Within a few years, an industry of managed behavioral health organizations was created by the contracting

activities of human resources executives of the major US corporations. The new companies were staffed by a corps of seasoned mental health professionals, most of whom had been prepared in the ranks of the community mental health movement. They introduced, in a business-like fashion, utilization review techniques reinforced by elaborate information and communication technology. They selected panels of providers who could be trusted to care for enrolled members without inappropriate or extended use of hospital or other expensive care. Their impact was dramatic: Within a few years, behavioral health expenditures stabilized, at least temporarily.

The success of the carve out of behavioral health care inspired employers to apply their contracting techniques to the rest of health care. Even though in 1987, only 27 percent were in managed care, by 1995, 73 percent of insured employees were enrolled in some form of managed care. Simultaneously, in the early 1990s, President Bill Clinton undertook a health care reform that attempted to provide universal health care as a highly regulated public entitlement. The Clinton health plan was labeled "managed competition" as a nod to the rapid move to competitive health care markets. However, the originators of that term—an assembly of business and health care experts known as the Jackson Hole Group—were much clearer about an approach in which government and government regulation would play a minimal role.

Nonetheless, a health care marketplace was created in the course of the failed Clinton health care reform. A wide variety of health care entrepreneurs would organize providers to bring their products to the markets, and consumers were encouraged to make choices that would ensure the value of the care that they would receive. The employers, through their contracting practices, sought to apply the principles of industrial organization to health care delivery to advance the ideal of organized systems of health care that were accountable and attractive to value-driven consumers. Although, certainly, some excellent health care delivery systems have developed, most have fallen far short of these ideals in a full array of organizational structures. All of these structures and models fell under the term "managed care."

In a systematic review of the dramatic changes in US health care in the last 15 years, Ginsberg described a "backlash" in the rapid growth of managed care. In the early 1990s, managed care's utilization review techniques and benefits had restrained providers; HMOs and other plan types restructured health care delivery; the insurance industry transformed itself, acquiring HMOs and creating a new array of "products" for employers to purchase; risk was cascaded to hospitals, in many cases making them insurers and encouraging them to consolidate their positions in physician hospital organizations (PHOs); and physicians reorganized their practices into independent practice associations (IPAs) and multi-

and single-specialty groups. By the mid-1990s, the rise of health care costs leveled off (temporarily), but there was widespread dissatisfaction among consumers and the rest of the system. By 2000 many of these measures were relaxed. Instead of benefits that restrained professionals, copays for consumers restrained their utilization. Employers encouraged broader networks, preferred provider organizations, and eliminated gatekeeper access to specialists. The insurance industry consolidated itself through cross-market mergers and large national contracts.

Quality and accountability began to dominate the health care market by the end of the 1990s, and consumers were encouraged to look to quality ratings by major accreditation organizations in making their choices. Managed care has had an indelible impact on the US health care system, introducing issues of quality and accountability, reinforced by the Institute of Medicine's report Crossing the Quality Chasm.

At the End of the Century

The last decade of the 20th century was a period of change, reflection, and revision. While the rise of health care costs were slowed temporarily by managed care, 50 million people were projected to be uninsured by 2020, and health care costs would hit 20% of the Gross Domestic Product (GDP). Health economists considered this unsustainable, seriously eroding the ability of the country to attend to infrastructure, education, and other critical societal factors (see Chapter 58.2 on Reforming Health Care). What has happened to public and community psychiatry was documented in three important retrospectives. Historian Gerald Grob and psychiatrist Howard Goldman have chronicled the complex policy formulation processes that lead to federal mental health policy initiatives and maneuvers in the 1990s and early 21st century. Two health economists, Richard Frank and Sherry Glied, have provided a sweeping and data-rich review of the mental health policies in this period and the performance of the resulting programs. Finally, President George W. Bush's New Freedom Commission on Mental Health in 2004 revisited the task of the last presidential commission under the Carter Administration to provide a thorough review of current mental health policy and formulate a comprehensive policy to address specific problems and plans for the future. All three efforts are reflections on the very eventful changes in the last years of the 20th century as public psychiatry moved from institutions into the community. They are also among the public documents that laid the groundwork for the Affordable Care Act in 2010.

Grob and Goldman tell the story as an alternation between policy idealists seeking radical reform and policy realists intent on incremental reforms. In 1980, the Mental Health Systems Act (MHSA) had authorized a radical reform of 15 years of community mental health programming. As noted earlier, this was repealed in the first months of the Reagan

Administration. The National Plan for the Chronically Mentally Ill, which had been developed by an interagency task force in the course of the congressional review of the MHSA, became a blueprint for incremental reform and established a degree of success in implementing its objectives over the decade of the 1980s. This provided the context for NIMH's Community Support Program (CSP) for adults, the Child and Adolescent Service System Program (CASSP) for children with serious emotional disorders and their families, and the administration of the mental health Block Grant program to the states.

With a meticulous assembly of state/federal Medicaid data, Frank and Glied documented a general improvement in every aspect of mental health care. Remarkably, public financing for mental health care from all sources increased 70-fold from 1956 to 1997. The increase is 100-fold if reported expenditures of $100 billion in 2003 are considered. The picture they present, however, is complex. As the process of deinstitutionalization proceeded and patients were moved from state hospitals to care in community programs, the financing also moved from institutional budgets to the categorical budgets of the state mental health agencies to fund the community programs. As employers moved dramatically to managed care in the 1990s, however, state Medicaid agencies followed suit and contracted with the Behavioral Health Managed Care Organizations (BHMCOs) to manage their mental health benefit. States quickly realized that by moving mental health categorical dollars into their Medicaid program they could leverage Federal Financial Participation (FFP) into their state's general fund. Almost all of the states shifted large portions of their mental health budgets for direct services into the Medicaid program—hence the substantial increases in mental health funding. However, these services ended up in the stewardship of Managed Behavioral Health Organizations (MBHOs), whose products tended to be modeled on private sector mental health care for general populations.

The shift of mental health service dollars to Medicaid and then to the control of the MBHOs was considered a critical mistake both by Grob and Goldman and by Frank and Glied. For Grob and Goldman it derailed the possibility of genuine reform that would fulfill the vision of coherent specialized delivery systems fully integrated in the community. For Frank and Glied it meant a loss of focus when patients were mainstreamed into communities, where they discovered the fundamental fragmentation of US social welfare programs. Frank and Glied argued for "exceptionalism" as opposed to "mainstreaming" for mental health services, with mental health exceptionalism evolving historically into mental health mainstreaming, the latter being what is typically found in state systems today. They proposed, as the next step, a federal regulatory agency (Fig. 58.1–1).

The more formal and official retrospective is that of the 2003 President's New Freedom Commission on Mental Health. Its landmark

finding was the recognition that a measure of recovery is possible for persons with even the most severe and persistent mental illness. The Commission had in hand the recently prepared Mental Health: A Report of the Surgeon General by Satcher, which reviewed scientific advances in the understanding of mental illness. These were, of course, dramatic. Despite this very hopeful news, the Commission's report was confronted with "unmet needs … many barriers to care for people with mental illnesses" and a severely "fragmented mental health delivery system." The states, on a fast track to wind down the state hospitals and, in most cases, to build community support programs, had uneven success. In many large cities, persons discharged from the hospitals with serious mental illness joined the ranks of the homeless or lived in substandard conditions in welfare hotels or board-and-care facilities. The Commission's report proceeded to lay out the goals and priorities through which the mental health system could be rebuilt at the family and community level so as effectively to transform it to enhance the resilience and recovery of individuals who suffer from mental illness.

Ascendance of Biological Psychiatry

In the year 2000, the Nobel Prize in Medicine was awarded to Arvid Carlsson, Paul Greengard, and Eric Kandel "for their discoveries concerning signal transduction in the nervous system"—work that explained at the level of molecular biology the effect of chlorpromazine and the psychotropic medications that followed it. In June of that same year, the Human Genome Project of the U.S. Department of Energy announced the completion of a working draft of the entire human genome sequence. Using sophisticated microarray techniques, it might one day be possible from a simple blood sample to determine what "type" of depression or psychosis an individual has and, in turn, target specific treatment to that specific mental disorder.

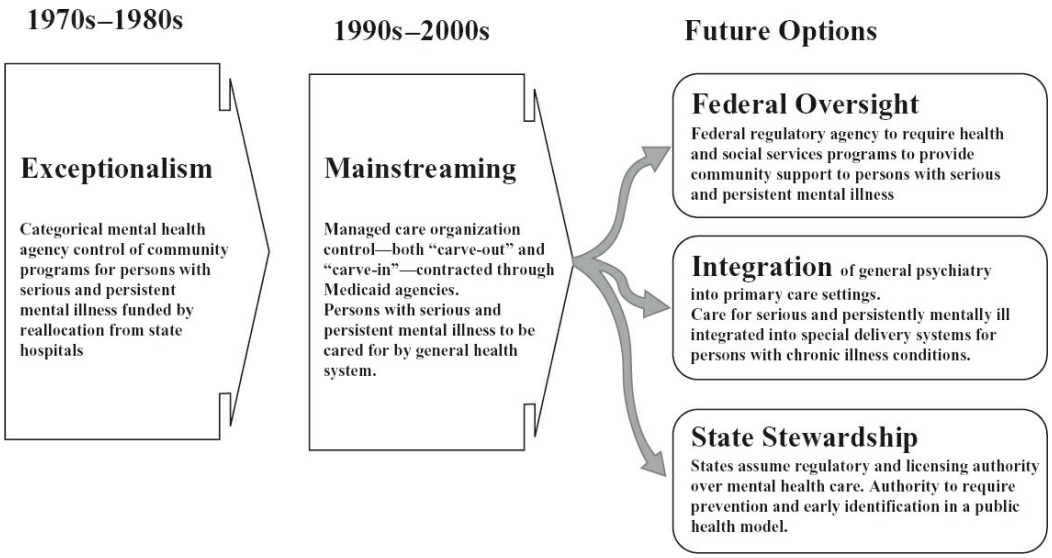

FIGURE 58.1–1. Possible futures for the evolution of mental health exceptionalism versus mental health mainstreaming. (Adapted from Frank FG, Glied SA. *Better But Not Well, Mental Health Policy in the United States since 1950*. Baltimore, MD: Johns Hopkins University Press; 2006.)

Reprise

It may appear that, in two centuries, US mental health has come full circle, that individuals severely disabled by mental illness are back living marginal or tragic lives in US communities and wander through cities without receiving appropriate care. There is certainly much truth in this. However, on the very spot where Nancy Barron was restrained in the poor house and screamed through the day in Emerson's Concord, there now stands a community residential program for adolescent girls with serious emotional disturbances that is part of a community system of care. They are treated by clinicians who understand the brain chemistry underlying their illnesses and who have pharmaceuticals and psychotherapy interventions that can alleviate them. With the increased knowledge-base and treatment evidence base that now informs public and community psychiatry, children and adults with serious and persistent emotional disturbance and mental illness can receive exceptionally high quality evaluations and treatments—if the health system and its psychiatric resources are accessible to them and able to follow through with the care that they need without letting them fall between the cracks. But this is the business of health care reform.

CONTEMPORARY PUBLIC AND COMMUNITY PSYCHIATRY'THE AFFORDABLE CARE ACT'S CALL FOR TRANSFORMATION

The emerging road map of contemporary psychiatry in the 21st century is largely conditioned by the impetus of the ACA for the whole of American Medicine. This has been characterized by the implementers of the new law as the "Triple Aim"—Better Care, Better Health, Lower Costs. Better Care refers to the emphasis on quality and safety which generated the "Quality Chasm" reports of the Institute of Medicine (IOM) including that focused on Behavioral Health and Substance Use Conditions as well as the development of the Patient-Centered Outcomes Research Institute (PCORI) and the work of the Agency for Healthcare Research & Quality (AHRQ). Better Health refers to the emphasis upon population health that has long been the focus of Public Health, its methods of surveillance, epidemiology, health promotion and prevention. And Lower Costs refers to improvements in quality and safety, the application of treatments that are demonstrated as effective, the organization of health service delivery and the integration of health care delivery systems leading to improved efficiency and a better process of care for the individual. The authors will use these three aims as a framework in which to discuss the challenges ahead for Public and Community Psychiatry. With a majority anti-ACA

government from the 2016 election, the outcome remains unclear whether the ACA will be totally replaced or improved upon, and what the results will be in either scenario.

Better Care: Evidence-Based Psychiatry, Integrated with Primary Care, in the Context of Coherent Health Care Delivery Systems

The developments in the last decade leading up to the ACA, as well as the statute itself, have laid the groundwork for the agenda of improving health care for public and community psychiatry. The Surgeon General's Report (2000) celebrated the impressive scientific advances in neuropsychiatry in anticipation of the New Freedom Commission's review of mental health policy and programs (2003). As several prominent medical societies came together to formulate a common definition of a "Patient-Centered Medical Home," SAMHSA and HRSA, crafted a joint commitment to integrate psychiatry into Primary Care in their Center for Integrated Health Solutions. The National Committee for Quality and Accountability (NCQA) provided detailed standards by which a Patient-Centered Medical Home could be certified. HRSA invited CMHCs to apply for status as Community Health Centers or Federally Qualifies Health Centers (FQHCs), as well as strengthening the behavioral health services in the existing FQHCs. In states and localities new programmatic models developed based upon evidence-based findings. Many of these models were "manualized" (i.e., formally documented in written protocols and decision rules that guide clinicians step-by-step in the application of specific therapeutic procedures) to support their fidelity to research-validated practices in implementing treatment and replication as the therapeutic model is disseminated throughout the field. The Agency for Health Research and Quality (AHRQ) supported an ambitious program of research designed to strengthen the evidence base of clinical interventions across the board. These efforts gained more focus by the ACA's creation of the Patient-Centered Outcomes Research Institute (PCORI). Finally, in order to meet the glaring need for an increased Primary Care workforce, the ACA included the concept of a "Primary Care Residency" program to be based in community-based practices as Teaching Health Centers including the FQHCs and CMHCs. Psychiatry was included as one of the nine medical subspecialties that would be active in this new program.

In order to be specific about the dimensions of Public and Community Psychiatry in improving patient care, it will be important to identify the obstacles that Public and Community Psychiatry will confront. The first is the relative absence of psychiatrists working in Public and Community Psychiatry. Surely there are wonderfully dedicated psychiatrists caring for adults and children with behavioral health disabilities in public programs—but their numbers are very small. The second obstacle is the fact that the bulk of psychiatrist time that the heavily stressed and under-resourced

public mental health systems does employ is largely limited to writing medication scripts and supporting medication management to some degree. Finally, there is a strange and persistent historical separation between mainstream American physical medicine and psychiatry, apparently due to the lagging scientific advance of psychiatric practice in the last century and the lack of acknowledgement or knowledge of the dramatic progress in the science of the brain in recent years. But this is changing as it becomes clear that a significant percentage of individuals presenting with physical health disorders have co-morbid psychiatric disorders that must be addressed as well.

Adults with Severe and Persistent Mental Illness. Although psychiatrists that do work in public and community settings encounter patients with mild impairment or acute disorders, the preponderance of their work involves people with severe psychiatric disorders and chronic psychosocial impairment. Adults who have been faced with socioeconomic deprivation for most or all of their lives (often over several generations) are at risk for chronic psychiatric disorders that often are complicated by medical illness, violence, social stigma, and cultural or political marginalization. These adversities, as well as the cognitive, emotional, biological, and interpersonal impairments often experienced by adults with severe and persistent mental illness, may obscure their personal strengths and resilience and the resources potentially available to support them in their families and communities. A major challenge to psychiatrists serving adults who have the dual burden of mental illness and socioeconomic disadvantage is the identification and enhancement of the psychological and sociocultural resources that can promote illness management and reintegration in a support system.

The traditional foci for private sector psychiatric practice—evaluation and diagnosis of psychopathology, psychotherapy and pharmacotherapy—should be essential components in a fully operational program of assisting adults with severe and persistent mental illness to make the best possible adjustment in the least restrictive public and community settings. This would be consistent with the Affordable Care Act's integration of effort across public agencies, professional collaboration, and health systems as learning communities. There is, in fact, a full array of treatment models that include multidisciplinary psychiatric rehabilitation teams that are demonstrated as effective recovery-oriented interventions.

Evidence-Based Manualized Interventions for Adult Psychiatric Rehabilitation. Public and community approaches to psychiatric rehabilitation involve an array of complementary services that must be coordinated to assist people with severe mental illness or behavioral disorders and their families and support systems to manage symptoms, to

access and use resources effectively, and to gain the greatest degree of autonomy in the least restrictive setting possible. Compendia have been developed on a national level in the United States to catalogue evidence-based mental health treatment (e.g., the SAMHSA National Registry of Evidence-based Programs and Practices [NREPP] and National Child Traumatic Stress Network [NCTSN] Empirically Supported Treatments and Promising Practices; the Office of Juvenile Justice and Delinquency Model Program Guide) and prevention (e.g., Blueprints for Violence Prevention) interventions. Several states in the United States also have developed registries listing evidence-based mental health treatments (e.g., California, North Carolina, Oregon, Vermont, Virginia, Washington).

The development of this array of evidence-based program models marks the transition in public community based supportive programs to interventions that are proven to be therapeutic. They have demonstrated that a substantial degree of recovery is possible. They all assume and incorporate a coherent clinical process that is carried on by a multidisciplinary team guided by rehabilitation and mental health specialists—including the diagnostic and clinical expertise and medical/legal authority of the psychiatrist. Not all of these programs explicitly specify psychiatric participation, largely because few psychiatrists have been available to serve these populations. Nonetheless, the engagement of psychiatrists with the treatment of patients with persistent disability will strengthen these proven intervention models. An illustrative but not exhaustive sample of these interventions is detailed in Table 58.1–3.

The need for psychiatry's engagement in the care of adult patients whose mental illness has led to persistent disability cannot be overstated. In the past, little could be offered to these patients. They were cared for and confined in institutions or supported in community programs. Now they can be treated holistically in primary care settings (avoiding the stigma of mental illness) and work with their multidisciplinary team specialized in chronic illness care with a strong psychiatric component.

Children with Severe Emotional Disturbance. Children from all strata of society experience severe psychiatric illnesses (such as developmental, affective, and anxiety disorders) and behavioral problems (such as attention-deficit, oppositional defiant, or conduct disorders). Access to Primary Care can be a barrier for many families, especially those in poverty. Public and community psychiatry treatment is provided by five categorical agencies, which are all responsible for their care and often hopelessly fragmented in their response. In this context of social and governmental disorganization, children with severe emotional disturbance and their families require the cross-agency organization of a clinical process that will direct a coordinated combination psychiatric and

psychosocial services that wrap around the child and family.

It has generally become accepted that children with severe emotional disturbance and adolescents with severe behavioral disorders do better in families that are appropriately supported than having them removed from their families and placed for extended periods in alternative programs. These placements are often more restrictive or protective—such as psychiatric inpatient wards, residential group homes, and juvenile detention centers. Placements are often extended and usually break the continuity of whatever clinical process has begun in their homes. Placement separates children from their natural family, school, peer group, and community environments. While short-term placement in a skilled program working closely with the family and the clinical team can provide a healing respite, if extended and detached from community and family, it disrupts and fragments the clinical process. Home and community-based care for troubled children has been made possible by the development of evidence-based program models for children with severe disorders and their families. These have been developed, tested, and disseminated in replicable manuals and training programs, and prominent models are presented in Table 58.1–4.

Table 58.1–3.
Evidence-Based Manualized Interventions for Adult Psychiatric Rehabilitation

Modality	Goals and Targeted Areas	Results of Evaluation Research
Social and independent living skills	Teaches skills for managing symptoms and preventing relapse (mental illness and addictions), improving social communication with providers, peers, and family, and choosing and engaging in activities that match the person's abilities, limitations, and interests	Social and independent living skills evaluated in randomized, clinical trials show positive outcome, with increased successful community living and reduced rehospitalization
Assertive Community Treatment (ACT)	Provides ongoing care coordination and case management to increase ability to live independently in the community and prevent/manage psychiatric crises	Randomized effectiveness trials with adults and adolescents with serious and persistent mental illness and co-occurring addictive disorders demonstrate increased psychiatric, residential, and vocational stability
Cognitive-Behavior Therapy (CBT)	Teaches skills for recognizing and modifying thoughts and beliefs that increase or prolong distress, and strengthening thoughts and beliefs that are self-affirming	Meta-analyses support CBT's efficacy in reducing the severity of anxiety disorders, somatoform disorders, bulimia, anger control problems, and generalized stress reactions, and offer modest support for CBT as an adjunctive treatment with addictive disorders and serious mental

	and interpersonally assertive.	illness.
Dialectical behavior therapy	Teaches skills for regulating intense emotions, tolerating extreme emotional distress, achieving interpersonal effectiveness, and developing mindfulness	Evaluated in randomized clinical trials with parasuicidal adult patients with borderline personality disorder, resulting in reduced behavioral and psychiatric crises greater than that achieved by expert psychiatrists' treatment
	Originally developed for adults with borderline personality disorder, but has been adapted for use with a wide range of Axis I disorders in which extreme dysregulation and poor impulse control are problematic	Randomized clinical trials and observational studies show positive psychosocial outcomes with adults with severe and persistent mental illness and chronic addictions and with adolescents who are suicidal or self-harming, including those who have substance abuse or posttraumatic stress disorders
Acceptance and commitment therapy	Teaches skills for identifying and managing (rather than avoiding and feeling controlled by) the symptoms of psychosis, affective, anxiety, eating, traumatic stress, and addictive disorders. Acceptance of symptoms facilitated by explanation that psychotic symptoms are extreme versions of normal perception and cognitive phenomena (private experiences)	Randomized clinical trials show evidence of increased awareness but reduced distress and impairment due to psychiatric symptoms across a variety of psychiatric disorders in outpatient and inpatient treatment settings
Relapse prevention therapy	Teaches skills for identifying the warning signs of potential or actual resumption of substance use or psychiatric symptoms	Randomized clinical trials show efficacy in delaying, reducing the severity of, and preventing substance use relapse with adults with chronic addictions and co-occurring or comorbid psychiatric disorders
	Uses cognitive–behavioral skills for proactively modifying patterns of thought and behavior that are likely to escalate into a full relapse	
Interpersonal psychotherapy	Teaches skills for interacting and communicating effectively in close relationships, with a particular focus on overcoming four core interpersonal problems: unresolved grief, role conflict (e.g., dissatisfaction with parenting), role transition (e.g., marital separation), and interpersonal deficits (e.g., isolation and anger	Randomized clinical trials show evidence of reduced severity and risk of relapse of moderate to severe depression; PTSD, anxiety disorders, and eating disorders

	management)	
Behavioral Therapy	Provides systematic reinforcement contingencies to increase adaptive behavior and to extinguish maladaptive behavior	Randomized effectiveness trials show evidence of reduced alcohol, drug, and other addictive behaviors
Motivational Enhancement Therapy	Teaches skills for recognizing and aligning choices and behavior with personal motivations	Randomized effectiveness trials show evidence of reduced alcohol, drug, and other addictive behaviors
Trauma-focused cognitive–behavior therapy for posttraumatic stress disorder (PTSD)	Teaches skills for safely recalling rather than avoiding troubling memories of traumatic events and for recognizing and managing PTSD symptoms, such as intrusive memories or flashbacks, emotional numbing, dissociation, and distress (e.g., guilt, grief), and problems with sleep or anger, or concentration problems due to hyper arousal or hypervigilance	Randomized clinical trials of seven approaches with adults (prolonged exposure therapy, cognitive processing therapy, eye movement desensitization and reprocessing, narrative exposure therapy, seeking safety, skills training for affect/interpersonal regulation with modified prolonged exposure, and emotion-focused trauma therapy) show evidence of safety and efficacy for adults with severe PTSD, co-occurring addictive disorders, problems with anger, dissociation, grief, and guilt, including one study with adults with chronic mental illness
Self-regulation focused cognitive behavior therapy for PTSD	Teaches emotion regulation skills for recognizing and modulating traumatic stress symptoms without requiring re-working of traumatic memories	Randomized effectiveness trials of three approaches with adults (Trauma Affect Regulation: guide for education and therapy; patient-centered therapy; skills training in affect/interpersonal regulation) show evidence of safety and efficacy for adults with severe PTSD, co-occurring addictive disorders, problems with anger, dissociation, grief, and guilt, including in individuals with comorbid serious mental illness, and incarcerated women.

Violence in the Lives of Children. A major impetus for reform of public and community psychiatry, and the mental health services systems more generally, has come from calls for reform from the public and policymakers in response to a more unavoidable crisis than the largely invisible suffering of most individuals with serious mental illness or severe emotional disturbance. This is the epidemic of violence against children in all walks of society. Most evident in the tragic and highly publicized incidents are the rash of mass shootings in schools and colleges, as epitomized by the Columbine high school and Virginia Tech university shootings. The larger issue of endemic violence against children that occurs on a daily basis was the catalyst for a commission in 2012, by the U.S. Attorney General of a Task Force on Defending Childhood Against Violence, which issued a comprehensive report detailing the epidemic scope of the problem and the need for comprehensive initiatives by law enforcement, the courts, child

protective services, family services, and mental health to protect children from violence and hasten recovery from the debilitating psychosocial effects of being victimized by violence. A preponderance of the recommendations directly or indirectly involve enhancing the funding, staffing, delivery systems, and use of evidence-based practices by public and community mental health systems and providers.

The Task Force report was released on December 12, 2012. Two days later, on December 14, 2012, a mass murder was perpetrated in which 26 children and teachers were killed in the Sandy Hook elementary school in Newtown, Connecticut.

The Sandy Hook Tragedy. The shooter was a young man who had a lengthy documented history of autism spectrum and related (e.g., obsessive compulsive disorder) psychiatric illness. A report by the Connecticut Child Advocate's Office was released on November 21, 2014 (http://www.ct.gov/oca/lib/oca/sandyhook11212014.pdf) describing in detail the decidedly mixed role played by psychiatry and mental health providers in identifying and attempting to respond to the shooter's serious emotional disturbance.

The Child Advocate's report describes how early childhood intervention services were determined to be necessary for problems with sensory hypersensitivity and expressive communication, and were provided to him as a preschooler and extended to address subsequent learning problems through Individualized Education Plans developed while he was in elementary, middle, and high school. However, neither a psychiatric admission to a hospital emergency department at age 12, nor a remarkably thorough and prescient biopsychosocial evaluation at age 13 by an academic child psychiatrist, with recommended psychiatric and psychosocial services, prompted the required level of monitoring. Instead, the boy was seen for only a handful of sessions over the next three years by an adult psychiatrist. No medications were prescribed. The boy was kept entirely out of school, despite the academic child psychiatrist's warning that what was needed was an increase in supported socialization (along with appropriate psychotropic medication and social-emotional rehabilitation services). After a year's hiatus on homebound status, the boy returned to middle school. A teacher noted that the child seemed preoccupied with violent fantasies to the point of not being able to complete writing assignments without what the teacher considered to be a disturbing intrusion of violent material. This observation apparently had never been noted when he resumed attendance at public high school the following Fall, with IEP accommodations that allowed him to attend a few classes at the school, take classes at a public university, and ultimately to graduate ahead of schedule. From the point of high school graduation there is no record of further contact by the schools or with any mental health providers, as he became increasingly isolated and embroiled in Internet communities that promoted mass violence and pedophilia.

The implication of this tragic case study for psychiatry is definitely not that psychiatrists can, or ever should be asked to, detect future mass murderers. In fact, there are studies suggesting that the ability to predict future violence, whether self-directed or toward others, is not strong. While there were numerous red flags that might seem, with hindsight, to presage the development of a violent killer, in fact there are thousands of youth who are similarly troubled, impaired, and isolated who never commit acts of violence. Instead, what the case study illustrates is the tragic deterioration that occurs when severely disturbed youth either never come to the attention of mental health providers or, perhaps even more disturbingly, are definitely identified by competent psychiatric and mental health

providers but then are lost to follow-up and never receive adequate evidence-based services.

The Child Advocate's report illuminated gaps and limitations in several interconnected public systems, including those in which public and community psychiatry should be involved. The report made a number of recommendations for enhancing public and community systems serving children and families. These include requiring that universal mental health screening be conducted and followed by timely provision of coordinated prevention, and support services by pediatric providers, schools, child welfare, and social/family services as well as mental health providers. Finally, the report concludes with three recommendations to increase the "Expertise and Services to Support Children with Developmental and Mental Health Challenges": namely, (1) to address the major issue for families—the paucity of services and supports for children and youth, particularly older youth and adults, who have Autism Spectrum Disorders with or without co-occurring mental health challenges; (2) to address the dramatic workforce development needs, increased technical support, and expertise to help mental health, pediatric, and educational providers meaningfully meet the needs of children with complex developmental or mental health disorders, and their families; and (3) to enable schools to provide, or even to import, comprehensive behavioral health or developmental supports to children, and will need significant support to ensure adequate expertise and related services for children with highly specialized needs. This latter recommendation will need to accommodate the schools' federal Special Education mandate to allow collaboration (and cost sharing) with other public agencies.

The challenges presented by the gathering currents of transformation of the health care system are especially acute for child and family mental health needs because of the complexity of responsibility among the child serving agencies. (1) The first need for public and community psychiatry is to work for an integration of effort among all the responsible parties for each individual child and family, starting with the Patient-Centered Medical Home in the primary care organization. Primary Care should be the center of collaboration since it will avoid much of the stigma of mental illness and inappropriate use of labeling; it will enable early intervention for the child and preparation of the family to meet the special needs of the child; and it will encourage the other categorical agencies to commit themselves to a "common plan of care" and draw the needed resources from across the federal funding silos supporting each agency. (2) The second need for public and community psychiatry is to work for funding mechanisms that allow compensation for seasoned clinicians to provide collateral support and consultation for front-line direct care providers who provide day-to-day care for children and their families. This will also make it possible for seasoned clinicians to participate in treatment team meetings

and to guide diagnostic discussions so that the team has a deeper understanding of clinical goals that are to be set and expectations of patient outcomes sought. Under the ACA there is an array of financing mechanisms involving bundled rates, risk-adjusted capitations, and pooled funding streams from the categorical silos that can be used to fund this kind of collateral support and consultation. (3) Finally, there is a critical need for child and adolescent psychiatrists nationally and, again, very few of them work in public and community psychiatry. Somehow the children who are most distressed, whose families are devastated by the severity of poverty, who have suffered profound trauma, and who present the most complex psychiatric symptoms are assigned to the most minimally prepared caregivers—young and untested child welfare workers, overworked probation officers, unqualified special education teachers—often without adequate supervision and with no access to the most highly trained specialists. It is the kind of neglect that leads to agony in the lives of children and their families and squanders community resources.

Table 58.1–4.
Evidence-Based Manualized Intervention for Child Psychiatric Rehabilitation

Modality	Goals and Targeted Areas	Results of Evaluation
Problem-solving skills training (PST)	Teaches children skills and supports parents in helping children with self-monitoring, prosocial goal setting, developing peer positive environments, setting limits with friends, and developing problem-solving and communication skills	Randomized, clinical trials show evidence of improved problem solving and interpersonal skills and modest reductions in oppositional and inattentive behavior
Cognitive behavior therapy	Teaches skills for recognizing and modifying thoughts and beliefs that increase or prolong distress, and strengthening thoughts and beliefs that are self-affirming and interpersonally prosocial and assertive.	Meta-analyses demonstrate efficacy in reducing children's anxiety and depressive symptoms
Parent behavior/Teacher classroom management training	Parent management training changes how parents use tangible incentives and consequences and model effective communication and to help their child change disruptive or defiant behavior and to model behavior and interact with their children	Randomized effectiveness studies with two approaches (Incredible Years; Oregon PMT) demonstrate reduced antisocial, oppositional, impulsive behavior, extending to adolescence for children with severe early antisocial problems
Brief strategic family therapy (BSFT)	A problem-focused family therapy that focuses on improving family relationships by addressing family leadership, alliances, behavioral control, and parental responsibilities; it was developed and used extensively with inner-city children and families of color	In a randomized, clinical trial with inner-city families whose children had a variety of behavior problems and varying degrees of juvenile justice involvement, BSFT was superior to usual community care services

Functional family therapy (FFT)	FFT teaches skills for problem solving and behavior management and delivers to parents and children together	Randomized and quasi-experimental studies show reduced juvenile recidivism for youths receiving FFT
Multidimensional family therapy (MDFT)	MDFT identifies family interaction patterns and restructures them to increase trust, cooperation, and emotional support among all members of the family; MDFT has evolved from a primarily clinic-based intervention to delivering services in the home and community settings to enhance youth engagement and the immediacy of services; MDFT is developing a variant for traumatized youth incorporating trauma affect regulation: Guide for education and therapy (TARGET).	Randomized clinical trials of MDFT across a variety of ethnocultural and socioeconomic subpopulations have shown that it is associated with reduced youth substance abuse and legal problems and enhanced family functioning and school involvement
Multisystemic therapy (MST)	MST involves up to 6 months of individualized, community-based contact by an M.A. or Ph.D. level therapist, supported by a 24-hour back-up team; MST addresses all relevant environments by empowering parents with the skills and resources needed to independently prevent and manage crises and maintain family rules; variants of MST are under development for sexual offending youth, youth in psychiatric crisis, and youth who have experienced childhood maltreatment	In randomized, clinical trials, MST has been shown to reduce juvenile delinquent recidivism and crime severity, psychiatric symptoms, out-of-home placements, and drug use while improving family functioning and school attendance; large-scale dissemination studies in the United States, Canada, and Scandinavia show that MST is superior to usual services on some, but not all, psychosocial outcomes
Intensive in-home child and adolescent psychiatric services (IICAPS)	IICAPS provides low-income families with children with serious emotional disorders a team led by a child psychiatrist and including psychological, social work, and educational consultation and crisis intervention in the family's home	One randomized effectiveness trial and observational studies with a system of community agencies treating psychiatrically impaired, socioeconomically disadvantaged children demonstrate reduced psychiatric morbidity
Trauma-informed affect regulation therapies for posttraumatic stress disorder (PTSD)	Dyadic therapies for young children and their primary caregivers, and group and individual therapies for adolescents, teach children and caregivers skills for affect regulation to reduce PTSD-related emotional and behavioral problems	Randomized controlled trials with young traumatized children (child–parent psychotherapy, parent–child interaction therapy) and youth (trauma affect regulation: guide for education and therapy; trauma and grief components therapy for adolescents) demonstrate improvements in parent-child relationships and child/youth emotional, behavioral, and interpersonal functioning
Multidimensional treatment foster care (MTFC)	MTFC provides training and ongoing support, parent education, family therapy, and youth mentoring to enable	Three randomized controlled trials showed that MTFC reduced adolescents' psychiatric

	foster parents to maintain a structured therapeutic environment for teaching children skills, setting limits, and modeling communication and problem-solving strategies	rehospitalization, arrests and incarceration, running away, and drug use while improving vocational and educational outcome
Cognitive–behavioral intervention for trauma in the schools (CBITS)	CBITS provides 10-session groups in school mental health or counseling settings for preadolescents or adolescents with PTSD due to family or community violence; CBITS includes education about traumatic stress, skills for anxiety management, and desensitization of trauma memories	Two randomized, controlled trials with inner-city children (one with Spanish-speaking youth) showed CBITS to be associated with reductions in PTSD and depression but not in teacher-reported behavioral or emotional problems
Trauma-focused cognitive–behavioral therapy (TF-CBT)	TF-CBT guides traumatized children and their parents in recognizing and managing PTSD symptoms and in desensitizing trauma memories	Randomized, controlled trials with sexually and physically abused children and children who experienced traumatic losses, disasters, or violence showed reductions in PTSD and depression and improvements in social functioning, but no improvement in disruptive or oppositional behavior

Roles of Psychiatrists in Public and Community Psychiatry Multidisciplinary Teams

Psychiatrists in the public sector should work within a team that includes professionals from several disciplines (e.g., psychology, social work, nursing, occupational therapy, rehabilitation or addiction counseling, social services, housing, and employment specialists) and non-degreed direct care workers (e.g., bachelor's degree–level counselors or case managers, high school graduate indigenous outreach workers, peer-support specialists, and family advocates), each of whom brings unique skills and experience to address the varied needs of people with severe and persistent mental illnesses. The team, rather than any single provider, assumes responsibility for the ongoing care of each patient across the many levels of services and often for many years. Its success is based on effective communication and coordination. Communication and coordination should be explicitly focused on the individual's (i.e., patient's and/or family's) stated goals, as well as on the community resources available and the team's technical and logistical capacities. Maximizing the individual's motivation to participate actively and productively in the comprehensive treatment plan that is patient-centered and collaborative is essential.

Psychiatrists can play four primary roles in multidisciplinary psychiatric rehabilitation teams: (1) conducting psychiatric diagnostic and treatment evaluations with significant involvement from members of the team; (2) providing thoughtful pharmacotherapy; (3) coordinating care with physical medicine in order to promote both physical and mental health; and (4)

serving as the team's medical director, mentor, sometimes clinical supervisor, peer, and colleague.

Psychiatric Evaluation and Diagnosis. The goal of evaluation and diagnosis is to develop the most clinically effective individualized approach to treatment and rehabilitation for each patient by culturally-literate clinicians. Psychiatric evaluations in public and community settings must provide for the full participation of the patient (and family where appropriate). The discussion should carefully review the individual's psychosocial strengths and resources rather than weaknesses or failures. A positive approach should be offered to help overcome systemic factors (e.g., eligibility regulations for governmentally funded services or benefits), as well as an emphasis on helping the person to overcome disability and the limits to the individual's or family's access to services or benefits (e.g., welfare-to-work regulations that place time or other eligibility restrictions on types of temporary aid, such as food stamps or vouchers for household supplies or housing). Following a comprehensive assessment, a full multiaxial psychiatric evaluation is in order.

Pharmacotherapy. The psychiatrist's most visible role usually is providing pharmacotherapy. This role will be enhanced as Personalized Medicine in psychiatry becomes operational. The most difficult challenge in public sector mental health settings often is not the technical formulation of an effective medication regimen, but, instead, the arranging of a plan of care so that the patient reliably follows the prescribed regimen (i.e., adherence or compliance). A recent review of interventions designed to enhance adherence to psychotropic medication regimens by patients with schizophrenia found that education alone often was the only strategy used, despite it being less effective than approaches that focus on "concrete problem solving," "motivational techniques," and "reminders, self-monitoring tools, cues, and reinforcements" or that provide "an array of supportive and rehabilitative community-based services." Thus, although the psychiatrist must address technical medical issues accurately to formulate diagnoses and to establish an effective psychotropic regimen, effective pharmacotherapy depends heavily on providing—directly, in encounters with patients, or indirectly, by working closely with nonpsychiatric mental health professionals and non-degreed care workers —practical assistance to patients to enable them to anticipate and to manage the psychosocial stressors or problems that can render even the most technically sound medication regimen completely ineffective as a result of nonadherence by the patient.

Pharmacotherapy must be essentially "therapy" that the prescriber and his/her clinical team administer in a coordinated way. The new funding mechanisms being developed under the ACA will make this possible and

more recognized as the reform continues.

Team Participant-Leader and Mentor. The psychiatrist is being called upon to play a special role as the project or program medical director—with the attendant responsibilities of monitoring the medical safety and well-being of all patients and establishing or supporting management and clinical procedures that support quality of care and a cohesive treatment team. As the team's clinical Mentor, or simply as a team member, the psychiatrist serves as a role model for compassionate and professional behavior in relation not only to patients, but also to all other team members. Formal or informal leadership is especially important when psychiatry residents and medical students work on a team as a training experience. The team psychiatrist serves not only as a mentor and a role model for the core aspects of psychiatric practice, but also should demonstrate the values and skills necessary to integrate psychiatry within the framework of psychosocial interventions that are multimodal, community-based, and longitudinal.

For example, psychiatric rehabilitation adapts and extends the traditional psychiatric services and roles by embedding them within the context of a set of psychosocial interventions designed not only to alleviate psychiatric symptoms, but also to promote optimal readjustment in the community: A recent review of the "requirements for multidisciplinary teamwork in psychiatric rehabilitation" summarized the challenge facing public and community psychiatrists. Psychiatrists who work with persons with severe mental illness and who wish to contribute as key members or leaders of their teams must go beyond their prescription pads to acquire knowledge, attitudes, and skills congruent with contemporary practice guidelines for psychiatric rehabilitation; give respect and support to the family and to multidisciplinary team members; gain their respect and support in return; cope with large caseloads; collaborate with agencies and programs to ensure continuity, consistency, and coordination of care; and value the roles and expertise of staff who provide case management, supported employment, skills training, and family and housing supports. If psychiatrists can incorporate these attitudes and skills, they can expect reciprocity and cohesion from other team members, and clients will benefit.

The shift from a narrow focus on treatment of a solo therapist to a broader psychiatric rehabilitation model, and now to models that stress the recovery of the individual, requires solid training (and continuing education) and hard work, as well as a combination of personal qualities, competencies, and expertise that have been summarized as "persistence, realistic optimism, a collaborative style, teaching ability, and networking and system coordination skills." Surveys of psychiatrists in the public sector indicate that those who embrace the rehabilitation perspective by

expanding their roles to be consultants and teachers for patients, families, trainees, and colleagues from other mental health disciplines are more satisfied professionally than those who focus on diagnosis and treatment alone.

New Paradigms. The paradigm shift active in the current reform calls for virtually all health care to be integrated and coordinated to deliver effective, rational, and cost-efficient care. If the focus is on the role of the individual physician and public and community psychiatrist, the new paradigm calls for practice in an organized system of care accountable for the health status of a designated population, operating through a clinical team, providing comprehensive primary care coupled with chronic illness care management, rehabilitation, or recovery models in an integrated system (Fig. 58.1–2).

This continuum requires four other components: (1) the application of new research to improve care and the application of bench-to-bedside models to translational research; (2) meaningful involvement of patients and families in a shared decision-making and patient-centered paradigm; (3) integration with primary and specialty care systems; and (4) the development of the clinical system as a system of advocacy. Training pre-professionals in such a model is one approach to introduce meaningful and sustained change to health care and its delivery. In short, it is developing the workforce of the future for psychiatry and behavioral health, especially as accountable care models populate the health care landscape.

Better Health: The Tools of Public Health Leading the Complex of Public Agencies to Coherent Health Care Delivery Supporting Population Health

Public health is not simply publicly funded health care but is rather a specific discipline and tradition. It is a complicated field that historically has been defined negatively by the dominance of personal health, that is, the health care delivery centered on individual doctor-entrepreneurs focused on the care of individual patients. Until the advent of managed care in the 1990s, US health care was an industry organized largely in terms of these individual doctors who dominated the corporate entities and associations in hospitals and insurance organizations. Each jurisdiction uniquely defined and organized its public health programs, and the particular pattern of personal health practices had substantial influence—hence the wide variation in public health programs throughout the country.

The passage of the ACA in 2010 has changed the significance of American Public Health. Despite the stunning successes of American Medicine, the series of critical reports by the Institute of Medicine and the process of crafting and passing the Act itself, which aims at universal coverage, opened up all the issues of the Public Health movement. As a discipline and tradition, public health's mission is to assure the conditions

in which people can be healthy. Public health consists of organized community efforts aimed at the prevention of disease and the promotion of health. It involves many disciplines, but rests on the scientific core of epidemiology. It provides as well an essential template for contemporary public and community psychiatry.

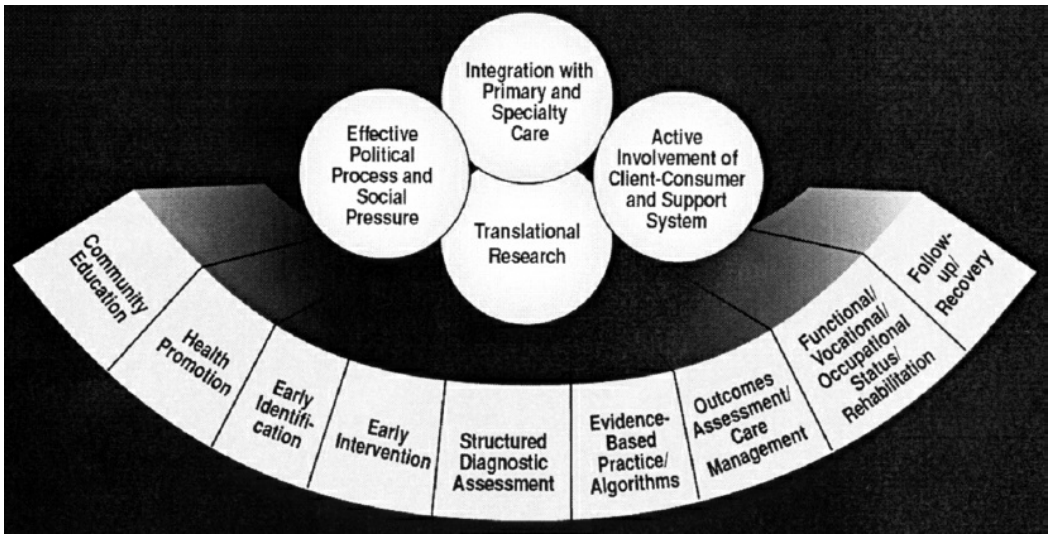

FIGURE 58.1–2. A model clinical and training system. An academic community health system for quality practice integrating public health, acute care, and rehabilitation and recovery models with science and advocacy.

Mental illness and substance use problems certainly are public health concerns. Mental illness affects the lives of as many as one in four individuals across all walks of life, with as many as 6 percent of all adults (i.e., more than 10 million adults in the United States alone) experiencing mental illness that is so severe and persistent as to constitute a major impairment to their ability to live and to work safely, independently, happily, and productively. Mental illness and substance abuse constitute an enormous financial burden on communities and states, as well as on the federal government. As community mental health efforts developed in the last 50 years, mental health professionals increasingly assumed the responsibilities called for by the Institute of Medicine's Committee for the Study of the Future of Public Health. This included fostering the development of competency in people, promoting effective leadership, ensuring a scientifically sound knowledge base, providing tools to monitor health problems and measure progress, operating within a productive organizational structure, having adequate financial resources, and having a solid legal foundation that supports effective action, all motivated by a vision of the public's health that is understood and supported by that public.

The 2000 Surgeon General's report on mental health underscored the necessity of a public health approach to care and rehabilitation for people

experiencing mental illness that is "broader in focus than medical models that concentrate on diagnosis and treatment." Although diagnosis and treatment are core areas of expertise for all psychiatrists, the Surgeon General recommended that—even when psychiatric diagnosis or treatment is the primary focus for practitioners, researchers, or educators—they should ground their professional activities in a vision and knowledge base that is "population-based ... encompass[ing] a focus on epidemiologic surveillance, health promotion, disease prevention, and access to services." These fundamental public health functions define the public health perspective in public and community psychiatry. In the discussion that follows, these four public health components are framed in current terms of health care reform to define a coherent "public health" strategy for public and community psychiatry.

Epidemiology. Based on epidemiological studies of nationally or regionally representative samples of adults and children in countries around the world, more than one half of the population experiences a mental illness involving at least one distinct psychiatric disorder in their lifetime, often as many as three or four comorbid disorders. It is important for psychiatrists to be prepared to identify not just the most obvious psychiatric disorder, but also other potential comorbid disorders, and to recognize the potential effects of culture, community structure and environment, and socioeconomic status on the ways in which people describe and understand mental illness and their expectations for care.

An immediate implication of population-based epidemiological studies is the need for timely identification of people with or at high risk for psychiatric disorders. When systematic diagnostic research studies are done in the most accessible arena for risk factor surveillance—primary and specialty medical care settings, schools, and child and family services programs—the findings consistently demonstrate that most psychiatric disorders are underdiagnosed (by 25 to 75 percent). Subthreshold psychiatric disorders (i.e., significant symptoms and impairment but not sufficient to meet full diagnostic criteria) are even less well identified, although they are potentially more responsive to treatment and often associated with substantial psychosocial impairment. Despite the existence of strong evidence from prospective risk factor studies, as noted earlier, surveillance of the principal risk factors for psychiatric disorders (e.g., family history, life stressors, and prodromal signs and symptoms) is not systematically done by most psychiatrists or by the public mental health treatment system in most states and local communities.

A potentially related problem that has been revealed by psychiatric epidemiology is that at most half of all people with psychiatric or substance use disorders receive treatment for these problems. Results of studies of representative community populations from the United States, Canada,

and the Netherlands consistently reveal that about 7 percent (or 1 in 14 persons) believe that they need mental health treatment but have not received such treatment. For low-income families served by public and community psychiatrists, inability to pay for the costs of treatment was a more significant barrier to getting help in the United States than in countries like Canada and the Netherlands where public mental health services are more universally available (see Reforming Health Care, Chapter 58.2).

Further analyses of data from an epidemiological sample from the United States (surveyed in the early 1990s) and a replication sample surveyed 10 years later in the early 21st century determined that persons with demographic and economic advantages that arguably increased their ability to access health care (e.g., older age, mainstream ethnicity, employment, insurance) and who had a perceived need for help were most likely to get treatment for psychiatric problems. However, the providers of services in the more recent survey were more likely to be nonpsychiatric medical practitioners and less likely to be psychiatrists or other mental health or human services professionals than 10 years earlier. Those persons fortunate enough to receive services for serious mental health problems tend to receive a lesser amount of poorer quality of services than research suggests is necessary to achieve meaningful improvement in the generally severe impairment and disability they are experiencing. Thus, the economically disadvantaged and often socially stigmatized segment of the population that is served by public and community psychiatry consistently is the least likely subpopulation to get this help. These sub-populations increasingly have had to turn to nonpsychiatric providers for the limited mental health treatment that is available, and only a fraction of these sub-populations get any treatment at all. These epidemiology research findings highlight a crucial dilemma facing psychiatric practitioners and systems of care: getting adequate help to people with severe psychiatric problems, particularly to the most disadvantaged families and communities where such help is especially scarce and inaccessible. In some communities, some established psychiatrists will not see individuals having low-paying insurance or will only see people who can afford to pay cash, refusing those with insurance coverage altogether. The implications and ramifications on society and the credibility of the profession of such practices are hopefully apparent in this time of health reform.

A strategic implication for a public health approach to psychiatry is that primary care medicine and, schools and child and family services agencies are the frontline for epidemiological surveillance not only of the incidence and prevalence of psychiatric morbidity per se, but also of the mental health professionals' ability to identify and get psychiatric services to the persons and communities most in need.

Health Promotion

Generally, psychiatry in the last half of the 20th century has been quietly on the sidelines, absorbed in private practice, caring for a segment of society that could appreciate its highly refined approaches to dealing with life's problems. Academic psychiatry, in the meantime, made strides laying the foundations of a science of the brain; and a small cohort of psychiatrists led and mentored the structural changes of the public mental health systems, closing the state hospitals and building community care systems. American health care is now in the midst of a reform driven largely by the visions of public health. The public promotion of community health is an integral part of this; and it is intriguing that the role of public and community psychiatry is to promote mental health care by participating in a series of models of treatment and community engagement.

Psychiatric professionals can contribute substantially to the promotion of public health by working with primary health care professionals and educators to identify and provide front-line treatment and referrals for adults and children who have undetected psychiatric symptoms, subthreshold syndromes, and psychiatric disorders. Collaborative care models bring psychiatric professionals into primary care and school settings as consultants for medical, nursing, and education professionals who work in those settings, as well as for social service, educators and direct treatment providers.

Furthermore, the organization of various forms of intensive outpatient care can benefit people with chronic health conditions, and behavioral health disorders. Achieving recovery and regaining physical or mental health often depends on a person's or family's access to social and psychological resources and integration into supportive social networks. Illness Management (also known as disease management or chronic illness care management) is a framework that has been adapted from medicine to guide mental health professionals in delivering services that go beyond traditional diagnostic treatment to promote the health and recovery of people with mental illness or addiction. A number of approaches to illness management have been scientifically and clinically evaluated and have been found to enhance standard psychiatric treatment.

Patients and families understandably want to feel that treatment is consistent with their personal and cultural beliefs and practices, and that it will improve their lives in ways that matter to them. In a word, the key to actualizing the potential benefits of psychoeducation and psychosocial skills is motivation. If the patient and family view the treatment as enabling them to achieve their specific personal goals, they are most likely to actively participate in, utilize, and adhere to the treatment. There are several evidence-based approaches to communicating with patients and families that enable psychiatrists and other mental and behavioral health providers

to maximize the patient's and family's motivation and adherence in treatment. These include motivational enhancement therapy (also known as motivational interviewing), acceptance and commitment therapy (also known as mindfulness-based cognitive therapy), client-centered therapy (also known as Rogerian therapy), working alliance interventions (also known as the therapeutic relationship), solution-focused therapy, and relapse prevention.

Providing patients and families with education about mental illness and the process of treatment and recovery has been shown to increase their knowledge and potentially to enhance adherence to treatment. An effective approach to promoting adherence to medication treatment is called behavioral tailoring—the careful design of practical reminders, routines, and positive incentives that adapt medication regimens to the preferences, lifestyle, and limitations of each individual patient.

Teaching patients skills for recognizing and coping with stressors, symptoms, or potential relapses of remitted symptoms in a timely and effective manner has been shown to reduce the severity of psychotic, anxiety, mood, or addictive symptoms and to prevent re-hospitalization or the resumption of potentially harmful (e.g., suicide or substance use) behaviors.

For lower-functioning patients, encouraging active, pleasurable social activities or, directly setting up structured opportunities (e.g., social activities) or milieus (e.g., social clubs or partial hospital programs) that foster safe and rewarding social interaction and membership in an ongoing social support system has been found to reduce symptoms of affective, anxiety, addictive, and psychotic disorders, as well as to enhance participants' overall social adjustment, self-esteem, and quality of life.

Illness management thus offers an added set of tools to the psychiatrist's toolkit that promote the health and well-being of patients and do not replace, but rather complement, the more traditional approaches to psychiatric diagnosis and treatment.

The illness management model of health promotion provides psychiatry with a critical role in the integrated delivery system. An illness management strategy promotes mental health while it engages patients as active collaborators in their care and enables them to enhance their health and well-being, as well as to recover from mental illness.

Another approach to health promotion for persons with mental illness or addictive disorders is supported by the Substance Abuse and Mental Health Services Administration's (SAMHSA) Center for Mental Health Services uses services led by persons in recovery (and in the case of children, by their families). Peer- or parent-led and -designed psychiatric services involve consumers ("persons with lived experience") and family members in collaborative efforts with psychiatric professionals to design, deliver, and evaluate educational, support, care coordination, and

counseling services.

Prevention. Prevention is a natural impulse for the Public Health model. A disease or debilitating condition that can be avoided through early intervention and timely treatment can save a neighborhood or community and secure the overall health of the population. The reform that is moving physicians forward reinforces the power of prevention. Health plans who will participate are asked to assume risk for carefully designed benefit packages for their enrollees and they manage this by contracting with networks of providers and health care systems to whom they contractually cascade that risk. The major public payers, Medicare and Medicaid, are creating funding mechanisms that create the same kind of accountability for the most medically vulnerable patients. Smart public and community psychiatry will position itself in accountable clinical entities that are actively engaged in early intervention and intensive, individualized care management.

Application of the traditional public health concepts of primary, secondary, and tertiary prevention has been confusing in psychiatry. Primary prevention involves addressing the root causes of illness with healthy individuals, with a goal of preventing illness before it occurs. Secondary prevention involves the early identification and early treatment of individuals with acute or subclinical disorders or high-risk persons to reduce morbidity. Tertiary prevention attempts to reduce the effects of a disorder on an individual through rehabilitation and chronic illness care management.

The Institute of Medicine, in an effort to clarify different aspects of prevention, developed a classification system with three different categories. Universal interventions are those intended for the general public, such as immunizations or media campaigns providing information about illnesses, early warning signs, and resources for health promotion and timely treatment. Selective interventions focus on individuals at higher-than-average risk (e.g., persons with prodromal symptoms or a family history of psychiatric disorders) to reduce morbidity by enhancing resilience and preventing the onset of illness. Indicated interventions target individuals who are experiencing impairment as a result of illness as early as possible in the course of the illness, to reduce the burden of the illness on the individual, family, community, and treatment system.

Psychiatric disorders most often follow a course over time that begin with an often lengthy period in which prodromal or sub-threshold symptoms or functional problems precede the full onset of an impairing disorder. Psychiatry integrated into primary care is in a position to identify these early symptoms and intervene effectively. In primary care, surveillance for very early universal intervention is possible, and selective interventions can be easily made as symptoms are recognized. Intensive

psychiatric services can be organized in the form of indicated interventions when disabling symptoms appear to establish a path to recovery.

A public health–oriented strategy for disease and disability prevention will involve early identification and intervention in cooperation with consumers, families, health care professionals, and educators. Psychiatric professionals can play a key role by bringing their expertise to the development of appropriate screening, assessment, education, and treatment strategies that, if implemented on a wide scale, may reduce the burden of mental health disability and facilitate recovery. Psychiatrists also can become involved in the social marketing of prevention services and programs to add expertise and credibility to support legislation, regulations, and commercial investment that will provide the economic base necessary for broad-based prevention programs.

Prevention interventions have been found to be effective with adults with a variety of risk factors or preclinical problems. For example, women who have been raped are less likely to develop PTSD if they receive a five-session cognitive–behavioral treatment than if their recovery is left to chance. Men and women identified with sub-threshold symptoms of depression in primary care settings are more likely to remain free from the full syndrome of depression or, if they do become clinically depressed, they will be able to recover rapidly with treatment if their standard medical treatment is enhanced by education about depression and learning skills for coping actively with depressive symptoms or stressors. This includes women who are at risk for postpartum depression, who have been shown to benefit from in-home cognitive behavior therapy in the first year after giving birth. Children who show early signs of severe behavioral problems (e.g., aggression, impulsivity, stealing) have been found to be less likely to become severely and chronically behaviorally impaired and antisocial when they and their caregiver(s) participate in therapeutic programs designed to enable the child and caregiver to deploy executive functions (i.e., anticipating consequences, inhibiting reactive behaviors, problem solving) while simultaneously strengthening their attachment bond with each other.

The Medicaid program includes a mandated program for Early Periodic Screening Diagnosis and Treatment (EPSDT) that broadly covers very early intervention and extends services through the twenty-first year. States have been slow to meaningfully implement the periodic screenings and full use of the program because of the potential financial liability of the program to the states. Consent decrees have in some cases been used to expand these services for children with severe emotional disturbances. EPSDT could provide a complete framework for the early identification of young children with psychiatric risk factors, and would support the effective targeting of comprehensive mental health and parenting services in the primary care setting of the family's "medical home."

Access to Effective Mental Health Care. As noted in the Preface to the Surgeon General's 2000 report, "even more than other areas of health and medicine, the mental health field is plagued by disparities in the availability of and access to its services." First, it is important to acknowledge that access to behavioral health services confronts a full array of what might be called demographic barriers—race (i.e., people of color), age (i.e., children and older adults), gender (i.e., female), sexual preference (i.e., homosexual), language (i.e., monolingual, non-English speaking or English as a second language), national background (i.e., self or family from nations in Central or South America, Africa, or Asia), education (i.e., less than college), and financial status (i.e., low income). These barriers apply not only to mental health care, but also to health care and health insurance of any kind, even to health insurance that specifically covers psychiatric disorders. People whose primary access to mental health or addiction recovery services come through publicly funded programs or agencies tend to have some, if not many, of these characteristics. Public sector mental health providers tend to provide services to people of varied backgrounds but primarily to people who have been marginalized due to these characteristics.

There apparently is not a consensus on how to measure disparities and, there are different definitions employed. The Institute of Medicine (IOM) defines a disparity as a difference in health care quality not due to differences in health care needs or preferences of the patient. It is known that ethnic and racial minorities, when compared to whites, have less access to mental health services, are less likely to receive the care needed, and more often receive care that is of poorer quality. But the patterns of disparities are complex. Even though ethnic minorities in the United States have higher rates of poverty and poor health than Caucasians, all of the minority subgroups (with the exception of Puerto Rican Americans) have lower rates of lifetime behavioral health disorders than white Americans. While overall rates of treatment for psychiatric disorders increased between 1990 and 2003, African Americans were only 50 percent as likely as whites to receive treatment. Nonetheless, there is evidence that African Americans are over-diagnosed with schizophrenia and underdiagnosed with mood disorders which strongly calls into question the field's diagnostic acumen and practice as noted earlier in this chapter. It is hard to ignore the likelihood that Caucasians have easier access to mental health services.

There are also barriers to access that are more structural, such as a shortage of trained mental health professionals. For most people with severe mental illness or addictions access is a serious problem. In The United States, the National Comorbidity Study (NCS) and the National Comorbidity Study-Replication (NCS-R) found that fewer than 40 percent of people with severe psychiatric disorders had received any mental health

treatment in the last year, and fewer than one in six (15 percent) had received minimally adequate mental health services. Young adults, African Americans, people residing in certain geographical areas, people with psychotic disorders, and patients treated by medical but not mental health providers were at highest risk for inadequate psychiatric treatment. Although income was not a predictor of inadequate treatment, it is likely that many of the people who did not receive adequate mental health services were uninsured or had insufficient insurance coverage for mental health conditions. They had no viable source of mental health care other than through a medical provider or clinic or in the public mental health system. A recent study with the NCS and NCS-R epidemiological databases found that persons of minority ethnic-racial backgrounds and those who were not employed or did not have health insurance were least likely to gain access to mental health services from medical, psychological, social work, and counseling professionals. However, race and socioeconomic resources were unrelated to the amount of mental health treatment received by persons who were able to gain access to these services.

Numerous studies have determined that a large number of individuals having mental and behavioral health problems are in jails or prisons. The Treatment Advocacy Center estimates that approximately 350,000 of the combined 2.1 million individuals in prisons and both county or city jails in the 2011–2012 time period were seriously mentally ill. With state hospital beds dropping from a high of about 560,000 in 1955 to 35,000 today, this means that there are ten times as many individuals with serious mental health problems incarcerated in the criminal justice system than in state mental hospitals.

Even when mental illness is identified, people with socioeconomic adversities often do not, or cannot, get adequate mental health services in their communities. For example, although it is estimated that more than a quarter of a million incarcerated adults in the United States have psychiatric disorders, few were detected or received treatment until they reached jail or prison. Federal and state correctional systems have instituted mental health screening and treatment programs to address psychiatric disorders as a health problem for incarcerated adults and to manage the problematic behavior that can occur in controlled settings as a result of mental illness. On returning to the community, the vast majority of prisoners with psychiatric disorders cease to receive more than minimal mental health services: A recent study found that fewer than one in six (16 percent—strikingly similar to the National Comorbidity Study finding) received steady mental health services, and only 1 in 20 with addictions received steady substance abuse recovery services. Thus, access to mental health and addiction treatment services is far better in prison than in the community! This has caused great concern because of the possibility that correctional facilities may be a de facto system of care for low-income

people (often of minority backgrounds) with serious mental illness.

Evidence of serious and pervasive barriers to accessing mental health care can be found in several kinds of crisis situation. Children and adults in poverty increasingly are deferring physical and mental health care until illnesses become chronic and severe and then find it difficult to gain entry into any setting for services except public hospital emergency departments. Children with severe psychiatric or behavioral disorders are often held in emergency department facilities for days and even weeks because the staff cannot locate any treatment facilities with an appropriate level of care that have an opening, or are willing to take the child as a patient (or to accept the child or family's insurance coverage or lack thereof). People who must rely on the public sector services thus face daunting obstacles when seeking appropriate mental health care because of the serious underfunding of practitioners and programs. The economic forces and public policy dilemmas that drive the ever-increasing cost of health care bear directly on public sector psychiatry, as well as on the lives of tens or hundreds of thousands of people who do not receive adequate care.

Access to appropriate behavioral health services is a bramble bush of factors from the demographic barriers that defined diversity to the structural barriers that keep physicians from responding fully to patient needs. The challenge for public and community psychiatry is to take advantage of the opportunities offered by the ACA to level the barriers to access for everyone in need of care. The primary obstacles to access are social disparity, socioeconomic adversity, and the fragmentation of the health and social service systems. Social disparity and socioeconomic adversity will diminish as the reform approaches close-to-universal health care coverage. It is estimated that by 2016 some 24 million Americans will have been enrolled in health plans that provide an appropriate benefit package. The fragmentation of the health and social services systems will be addressed by developments in three areas. First, the consensus model of the Patient-Centered Medical Home will ensure an appropriate expansion of primary care capacity organized so as to integrate ready access to specialty care, including psychiatry under the Medical Home's umbrella. Until recently Primary Care has not been functionally available to patients with severe and persistent mental illness; the increasing integration of psychiatry into primary care will allow for early intervention without stigma and appropriate treatment interventions to avoid more severe problems. Second, the development of new payment mechanisms led by the CMS's Center for Innovation will provide "accountable clinical entities" the flexibility to provide individualized care through risk adjusted rates appropriate for the needs of defined populations. This will support the development of chronic illness care programing including primary care-based programs for persons with serious and persistent mental illness. Third, value-based standards of care based on the efforts of AHRQ and

PCORI to support evidence-based research and to catalogue the most effective clinical practices and programs, will lead to "learning health care systems" for continuous improvement of care as well as lay the foundations for pay-for-performance mechanisms.

These measures, which are being actively pursued in the reform effort, will lead to a transformation of the health and social service systems to create a continuum of care capacity to meet the needs of every stage of mental illness, as well as realize the possibilities of public and community psychiatry.

Psychiatry and Public Agencies

Psychiatry's relationship to public sector agencies has, with some specific exceptions, been one of detachment because of the dominance of the private practice models before the advent of managed care and also largely because of the nature and structure of US social welfare programs. This is a divide that often began during residencies, in that historically many psychiatry training programs eschewed meaningful rotations in public sector settings. In contrast to most industrialized nations in which comprehensive social welfare reforms were integrated, US social welfare programs grew incrementally and categorically, that is, one category of service at a time. Large-scale initiatives, such as President Johnson's War on Poverty, have been implemented in piecemeal fashion by fragmented federal, state, and local government bureaucracies. Various social service agencies have been created through a process and set of alliances that has been called the iron triangle. Advocates form an organization to champion a particular cause, such as blindness, developmental disabilities, primary care, or mental health, among others. They find key legislative sponsors who advance the cause through legislation and appropriation. This creates a bureaucracy and then bureaucrats who join the alliance. Through successive legislative sessions, the alliance of advocates, lawmakers, and bureaucrats builds a stronger categorical agency; for example, adults and families with dependent children or people with biologically based, severe mental illness. From the 1930s, through the end of the 20th century, US social welfare agencies were created on this pattern: The system has resulted in generous funding and hopeless fragmentation of effort at local service delivery levels in which each agency depends on separate funding silos that also deliver conflicting rules and regulations.

It has been said that "mental health is not a place." Psychiatrists have a role at the receiving end of every categorical silo because the clients of each categorical agency experience mental illness or various addictions. Services for children with severe emotional, mental, or behavioral disturbances are a dramatic case of the importance of psychiatrists taking a leadership role in promoting comprehensive rather than fragmented services. Five categorical agencies—child welfare, education, primary health, substance

abuse, and juvenile justice—all have a responsibility to care for these children. Especially for the children with the most severe disabilities, protective services, special education, the juvenile probation officer, and the child's pediatrician all need the consultation and support of a skilled psychiatrist. The situation is similar for adults with severe mental illness, for whom case managers, vocational rehabilitation counselors, basic needs benefits specialists, social workers, psychotherapists, substance abuse counselors, parole or probation officers, legal conservators, visiting nurses, peer specialists, and physicians all may be mandated to deliver services. The need for psychiatric leadership is equally important for those who care for persons with mental retardation or developmental disabilities, for the elderly, for persons in vocational rehabilitation programs, and for persons in correctional programs. Virtually all public resources come through this complex of categorical agencies.

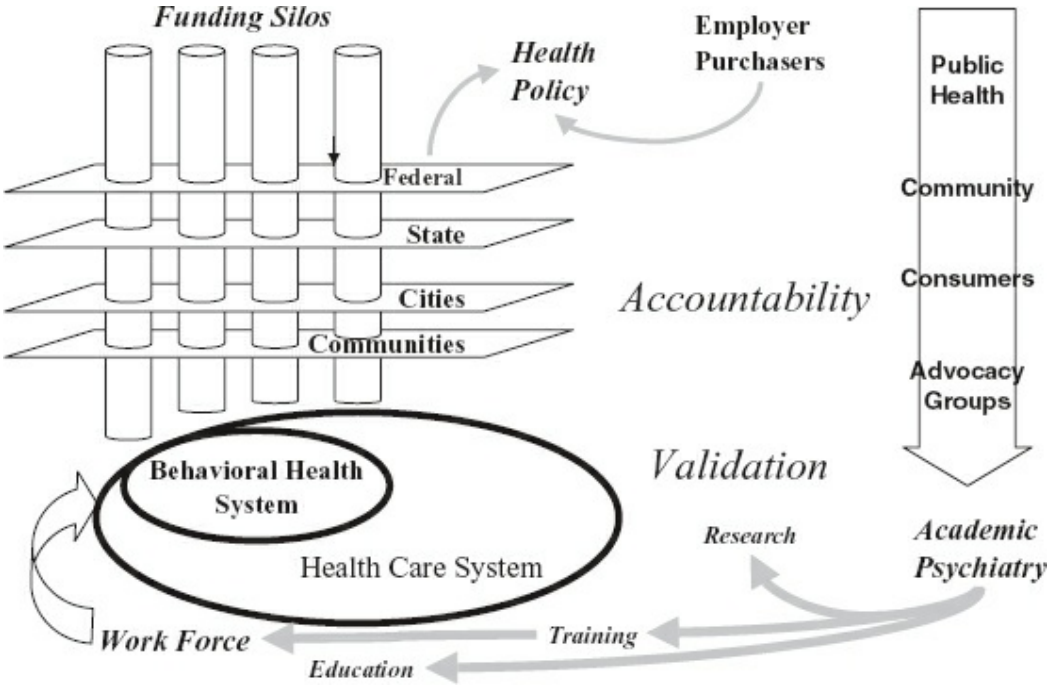

FIGURE 58.1–3. Conceptual map for initiatives and concepts for public and community psychiatry.

Figure 58.1–3 provides a conceptual map of this very complex governmental and social structure in which categorical government programs cascade through the levels of US federal, state and community systems. Health policy is formulated between government and employer purchasers and expressed unevenly through public health functions in communities, consumers, and advocates. Academic psychiatry bears responsibility for research, training, education and workforce development for a behavioral health system that is only loosely integrated into the health care system. The ACA has laid the groundwork for clarifying the accountability of all parties and establishing the validity and efficacy of

health care activities.

It will be impossible to achieve the ideals of the public health perspective on public and community psychiatry unless ways are deliberately devised to engage committed and proficient psychiatrists in meaningful roles in the operations of these agencies. Again, the factors at play in health care reform provide an opportunity to integrate the efforts of the complex and fragmented array of health and social service agencies by emphasis on the primacy of the clinical process (maintained within the PCMH), the revision of funding mechanisms to encourage clinical sponsorship by accountable clinical entities, and emphasis upon the implementation of evidence-based practice standards.

Workforce Challenges. The Health Care reform that is underway will make significant demands on the public and community psychiatry workforce. It has been noted that the National Comorbidity Survey reports that less than 40% of those with severe psychiatric disorders have had an encounter with a psychiatrist in the last year; that one in six (15%) have received minimally adequate care; that one in 20 persons with substance use disorders have received care. Meanwhile, the Bureau of Labor Statistics reports (2013) that, of the 36,000 psychiatrists, about 28% are self-employed in solo practices; 53% are working in private sector hospitals, group practices, or outpatient clinics; about 6% are in academic settings; and around 10% work in public and community psychiatry. The geographical distribution of psychiatrists is about 50% along the Atlantic seaboard, with concentrations in Texas and California as well as Illinois and Missouri. Access for persons with severe and persistent mental illness is already at crisis proportions and, as universal coverage is approached, it will become more critical. There is a recognized imbalance of the specialty care workforce over primary care and new medical schools are opening to meet the primary care need. If ACA-driven coverage for 24 million new patients are enrolled by 2016 as projected, many of them will come from the ranks of the most neglected and most severely disabled mentally ill. This imbalance will be even more critical.

The reform impulse that has been codified in the ACA contains many implications for the preparation of new generations of public and community psychiatrists and the adaption and retraining of the current workforce. The regime of quality, effectiveness, safety, and integration of effort represented by models like the Patient-Centered Medical Home will align psychiatry closely with primary care and its multidisciplinary teams. Constant reference to the evidence base of clinical interventions will enhance the value of accurate diagnostic assessment and the application of clinical treatment and the use of effective programmatic models. And the needs of newly covered patients with serious and persistent mental illness will demand the attention and skills of psychiatrists who can guide them to

recovery.

There are promising developments that will support the changes ahead. The Decade of the Brain sponsored by the NIMH in the 1990s, contributed to this change. Medical students are increasingly stimulated to enter the field by recognizing the importance of psychiatry to health care in general. They realize that this is still a field in which the practitioner (and by extension the interdisciplinary team) is an instrument of change through integration of contemporary knowledge, skillful learned interactions with patients, and a willingness to commit to providing value in quality care.

The Annapolis Coalition on the Behavioral Health Workforce has devoted years to the study of the workforce and to advocacy for needed change. Responding to the New Freedom Commission's call for a national blueprint for workforce development, SAMHSA commissioned the Annapolis Coalition to produce a national action plan for change. That report, published in 2007, identifies seven strategies for change, grouped in three broad domains. Achieving these seven goals would substantively affect the future quality of care: (1) Broadening the concept of the workforce by engaging individuals in recovery and their families when appropriate and (2) by expanding the capacities of communities to identify their needs and to promote behavioral health and wellness; (3) Strengthening the workforce by implementing systematic recruitment and retention strategies, (4) by making training and education relevant, effective, and accessible, and (5) by actively fostering leadership; (6) Implementing structures to support the workforce by enhancing its capacity and workforce development efforts and (7) by implementing a national research and evaluation agenda for behavioral workforce development.

Finally, the Institute of Medicine has issued a report calling for the Reform of Financing and Governance of Graduate Medical Education (2014) which has been for decades funded by Medicare in the residency programs of Academic Medical Centers and in Teaching Hospitals. The Report calls for a ten-year transition period in which the residency programs will be phased out while transformative and innovative residency formats are created in primary care settings as well as interdisciplinary organizations providing intensive care programs for persons with severe disabilities and multiple chronic illness conditions.

Lower Costs: Accountability for Quality, Safety, Efficiency, Effectiveness, and Evidence-Based Care

The goal of achieving lower costs for American health care is recognized as the critical priority of the nation's economic and civil wellbeing. It also presents very difficult challenges for this thriving sector of the American economy. The introduction of market forces by the ACA's health exchanges has yet to be fully demonstrated. But the ACA legislation itself introduces

strong tools to encourage cost control by exercising accountability on the part of individual providers and their health care organizations, for meeting standards for the quality of care (its safety, efficiency, effectiveness), and its recognition of care as being evidence-based. The enhancement of primary care is of critical importance and NCQA has crafted standards for a Patient-Centered Medical Home supported by all the related major medical disciplines. Effective treatment models are being recognized and catalogued by the Administration for Healthcare Research and Quality (AHRQ) and the Patient-Centered Outcomes Research Institute (PCORI). A host of innovative payment mechanisms (e.g., bundled rates) have been launched by the Center for Medicare & Medicaid Innovation (CMMI). Integrated Service Delivery Systems in the form of Integrated Care Organizations (ICOs) are being demonstrated in several states providing risk-adjusted global rates for dual eligible (Medicare/Medicaid) patients with profound disabilities—forty percent of whom are persons with serious and persistent mental illness. Specialized delivery systems models such as Accountable Care Organizations (ACOs) are being demonstrated for targeted populations through the CMS Innovation Center. All of these movements for the exercise of accountability will have an impact and will reshape public and community psychiatry.

The Health Care Marketplace. The most explicit cost control feature of the ACA is clearly the introduction of market forces through the development of the Health Care Exchanges on the part of states or on the part of the federal government on behalf of the states. While the health plans have brought competitive rates to the Exchanges in many instances, it cannot yet be determined the degree to which this will be effective in lowering costs. Many have observed that traditional health insurance organizations may evolve in the coming years into new business models. Health Care insurance, as it is currently known, may not survive. But the creation of a health care marketplace is not the only means for cost control that is presented in the ACA and the health policy ferment that accompanied its passage. First, there is a clear emphasis on the development of accountable clinical entities able to direct the care of patients; second, there is a renewed emphasis on limiting coverage to effective evidence-based care; and third, there is recognition that redesigned financing methods will enhance the quality of care and lower its cost.

Accountable Clinical Entities. As the implementation of the ACA is unfolding, it is clear that there is a growing appreciation of the role of "accountable clinical entities" in transitioning American health care delivery systems from their chronic fragmentation of functions and resources to coordination and integration of effort. By accountable clinical

entities, the authors mean health care systems that support interdisciplinary teams that sustain an individualized, ongoing clinical process for each patient (and family) and control the allocation of resources for the benefit of the patient's health. The statute itself identifies the need to improve and expand the provision of coherent Primary Care services as well as the need to provide more intensive chronic illness care delivery systems; employer purchasers have supported the concept of organized systems of care for over two decades; and the Institute of Medicine's reports on health care quality have repeatedly called for reform of the organizational structures of the health care delivery system.

In 2007, four prominent medical organizations endorsed Joint Principles of the Patient-Centered Medical Home (PCMH): the American Academy of Family Physicians, the American College of Physicians, the American Academy of Pediatrics, and the American Osteopathic Association. The PCMH elaborated the basic concept of a "Medical Home" for children with complex illness or disability proposed by the American Academy of Pediatrics in 1967, to apply to the basic need for expansion of Primary Care. The National Committee for Quality Assurance (NCQA), with the support of employer purchasers, as well as these medical organizations, developed standards for certification of clinical entities as Patient-Centered Medical Homes. Public and private payers have provided monetary incentives for organizations that gained certification. Between 2008 and 2014, over 8,300 practices were certified as PCMHs.

The Center for Medicare/Medicaid Services (CMS) has moved deliberately to support the enhancement of accountable clinical entities. Its Innovation Center, influenced by demonstration projects under the Medicare Modernization Act (2003), has developed a series of "Accountable Care Organization" (ACO) grants that encourage health care delivery systems to assume responsibility for specific high need populations, providing some form of bundled rates risk-adjusted to meet their patients' specific needs. These projects are budget neutral, but any savings are shared between the providing organization and the federal program. From 2010 to 2014, over 400 ACOs have been developed. CMS's Medicare/Medicaid Coordination Office has developed Integrated Care Organizations (ICOs) in a dozen states aimed at populations that are dual eligible, that is, eligible for both Medicare and Medicaid. These programs combine the Medicare and Medicaid dollars (and expenditures from other service silos) to provide a global capitation that is risk adjusted to the needs of the different populations. The dual eligible populations account for the highest health care expenditures. They are either elderly or disabled and they are poor. They suffer the most from the fragmentation of the health care delivery system. As noted, about 40% of them are persons with serious and persistent mental illness.

Health care delivery organizations that function as accountable clinical

entities and assume responsibility for managing, over the long term, the comprehensive care of specific populations will save money and lower costs. Fragmentation of health care services is expensive and these organizations will coordinate care whether it manages a properly enhanced Primary Care operation or an intensive program serving persons with multiple chronic conditions in home and community settings. Accountable clinical entities will be intent on early preventive intervention; they will control and advance effective care; they will be focused on the quality of care and its safety. They will maximize the value of each health care intervention and control costs.

Evidence-Based Care. It is fairly obvious that there can be lower costs if the health care system does "things that work" and avoids practices that have uncertain effect. The dramatic progress of medicine in the last century has been based on a methodical clinical process that constantly measures the effect of every intervention and constantly examines the accuracy of the diagnosis of the health problem. But one of the great stimuluses' to the movement for health care reform was the analysis of utilization in the national Medicare billing data by John Wennberg and his associates. They found surprising variation in practice in communities throughout the country and were able to distinguish interventions that were "preference-sensitive" (for patients and their practitioners) and "supply-sensitive" in that given equipment or training were readily available—and that these far outweighed interventions that were demonstrated as "effective."

Two agencies have emerged in the quest for evidence-based care. Within the U.S. Department of Health and Human Services, the Agency for Healthcare Research and Quality's (AHRQ) has a mission "to produce evidence to make health care safer, higher quality, more accessible, equitable, and affordable," as well as "to work and with other partners to make sure that the evidence is understood and used."

The second agency, the Patient-Centered Outcomes Research Institute (PCORI), a nonprofit nongovernmental organization, was established by the ACA to lead a national effort to develop patient-centered outcomes research (PCOR). With more than $100 million in annual federal funding, PCORI conducts peer review and funds more than 150 clinical research investigations developing and testing innovative approaches to patient-centered health care and its outcomes. As described on the PCORI website (www.PCORI.org), this involves funding clinical effectiveness research with the particular mission to help "people make informed healthcare decisions, and improves healthcare delivery and outcomes, by producing and promoting high-integrity, evidence-based information that comes from research guided by patients, caregivers, and the broader healthcare community."

Psychiatry is prominently represented in the portfolio of more than 350

investigations funded (with a total cost of more than $600 million) by PCORI since its inception in 2010. Studies of patient-centered and community-guided approaches to public sector treatment of depression with low-income women, rural Native American community members, and inner city Latino and African-American adults are three of the 20 research in action projects highlighted by PCORI. Other PCORI-funded mental health treatment studies address serious mental illnesses in adults and children (e.g., autism spectrum disorders), behavioral disorders (e.g., ADHD), substance use disorders, and PTSD.

The work of both of these agencies will create an ongoing framework to determine "what works" in health care. The best evidence for effective care will be systematically catalogued; standards of care will be refined; and ways will be found to improve care while lowering its cost.

Financing Methods. The primary method of payment that has dominated American health care has been the Fee for Service (FFS) system. Activities and procedures that are performed based on a patient's diagnosis are billed and reimbursed by the insurance-payer based on the terms of the insurance contract. This system encourages volume over quality; it is easily gamed; payment rates are constantly in dispute; there is no way to assure expected performance; it reduces medical care into a kind of piece-work.

As noted, CMS has addressed the question of financing or payment methods in the context of the demonstrations launched through its Innovation Center. The models of ACOs include different payment models including bundled capitation rates, risk adjusted to meet the needs of proposed populations. These have demonstrated savings that are shared between the government and the provider organization. CMS's program for the dual eligible populations developed as ICOs by its Medicare/Medicaid Coordination Office provides a kind of capitation determined as a "financial alignment" of the two programs. It is not clear that the service expenditures contracted by other categorical agencies such as a state Mental Health Department, and included in the state's Title XIX claim, are part of this financial alignment.

Other similar initiatives for payment reform include the Medicare shared savings models developed under the Prescription Drug, Improvement, and Modernization Act (MMA) of 2003; the Innovation Center's Multipayer advanced Primary Care Program; the Primary Care Initiative and many other experiments. Each of these program experiments are designed to limit unnecessary costs, to encourage integration of effort across various sectors of health care delivery, to enhance efficiency, effectiveness, and quality of care, and to maximize value and reward performance.

Public and Community Psychiatry in the Mainstream of American

Medicine. Finally, it is important to survey where and how Public and Community Psychiatry will participate in the changes that are afoot for American health care. The state hospitals were the first publicly funded health care, but they were separated in a strange way from the rest of health care as American medicine found its scientific legs in the reforms sparked by the Flexner Report in 1910. The strides that were made by medicine and science in the early 20th century were not possible for psychiatry because the tools were not yet available for a science of the brain. Psychiatry remained in the state hospitals with patients supported with public funds and the separation between psychiatry and the rest of medicine widened. By the end of the century this had changed dramatically and psychiatry is building on a flourishing base of science. The change was advanced in the middle of the century when biological psychiatry and the new medications began to be used and understood. But the hospitals were still operating and the strange separation of psychiatry from the rest of medicine persisted. The former Commissioner of Mental Health in Massachusetts and Director of the NIMH's Joint Commission on Mental Health and Illness (1961), Jack Ewalt, MD who had played a key role in the development of community mental health wrote a small book that methodically detailed the operational decline of the state hospitals at the same time that the Hill-Burton legislation was funding sparkling new general hospitals in communities throughout the country. The meaning of the word "hospital" was changing dramatically and that strange separation remained.

The term Public and Community Psychiatry may change just as dramatically in the process of achieving better care, better health, and lower cost. Community care, organized to be accountable for effectiveness, quality and safety will certainly survive. But, as a policy of universal access to health care is approached, a "public psychiatry" may make little sense. Patients who are impoverished by their disability will be indistinguishable. Public psychiatry itself may disappear as well as that strange separation between psychiatry and the rest of medicine.

References

Belfer M, Kieling C, Baker-Henningham H, et al. Child and adolescent mental health worldwide: evidence for action. *Lancet*. 2011;378(9801):1515–1525.

Bishop T, Press M, Keyhani S, Pincus HA. Acceptance of insurance by psychiatrists and the implications for access to mental health care. *JAMA Psychiatry*. 2014;71(2):176–181.

Bodenheimer T, Grumbach K. *Understanding Health Policy: A Clinical Approach*. New York: Lange; 2016.

Bradford D, Kim M, Braxton L, Marx C, Butterfield M, Elbogen EB. Access to medical care among persons with psychotic and major affective disorders. *Psychiatr Serv*. 2008;59:847–852.

*Cooke M, Irby D, O'Brien R. *Educating Physicians: A Call for Reform of Medical School and Residency*. Josey Bass; 2010.

Costa e Silva JA. Personalized medicine in psychiatry: new technologies and approaches. *Metabolism*. 2013;62 Suppl 1:S40–S44.

Druss B, Marcus S, Campbell J, Cuffell B, Harnett J. Medical services for clients in community mental health centers: Results from a national survey. *Psychiatr Serv*. 2008;59:917–920.

Elhai JD, Ford JD. Correlates of mental health service utilization intensity in the National Comorbidity Survey and National Comorbidity Survey Replication. *Psychiatr Serv*. 2007;57:1108–1115.

*Emauel EJ. *Reinventing American Health Care*. New York: Public Affairs; 2014.

England MJ, Cole R. Preparing for communities of care for child and family mental health for the twenty-first century. *Child Adolesc Psychiatr Clin North Am*. 1998;7:469–481, vii.

Enthoven AC. The history and principles of managed competition. *Health Aff (Millwood)*. 1993;12:24–48.

Ford JD, Chapman J, Mack M, Pearson G. Pathways from traumatic child maltreatment to delinquency: Implications for juvenile and permanency court proceedings and decisions. *Juv Fam Court J*. 2006;57(1):13–26.

Frank FG, Glied SA. *Better But Not Well, Mental Health Policy in the United States since 1950*. Baltimore, MD: Johns Hopkins University Press; 2006.

Frank FG, Glied SA. Better but not best: recent trends in the well-being of the mentally ill. *Health Aff (Millwood)*. 2009;28(3):637–648.

Ginsberg PB. Competition in health care: Its evolution over the past decade. *Health Aff (Millwood)*. 2005;24(6):1512–1522.

Grob GN, Goldman HH. *The Dilemma of Federal Mental Health Policy: Radical Reform or Incremental Change*? New Brunswick, NJ: Rutgers University Press; 2006.

Hoge MA, Morris JA, Daniels AS, Stuart GW, Huey LY, Adams N. *An Action Plan on Behavioral Health Workforce Development*. Cincinnati, OH: The Annapolis Coalition on the Behavioral Health; 2007. Available at http://www.annapoliscoalition.org and http://www.samhsa.gov/matrix2/matrix_workforce.aspx.

Huey LY, Lefley HP, Shern DL, Wainscott CA. Schizophrenia: a complex disease necessitating complex care. In: Buckley PF, Messias EL, eds. *Psychiatr Clin North Am*. 2007;30(3):323–592.

Hyde PH, Falls K, Morris JA, Schoenwald SK. *Turning Knowledge into Practice: A Manual for Behavioral Health Administrators and Practitioners about Understanding and Implementing Evidence Based Practices*. Boston, MA: Technical Assistance Collaborative; 2003.

*Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st-Century*. Washington, DC: National Academy Press; 2001.

Institute of Medicine. *The Future of Public Health in the 21st Century*. Washington, DC. National Academies Press; 2002.

*Institute of Medicine. *Improving the Quality of Health Care for Mental and Substance-Use Conditions*. Washington, DC: National Academy Press; 2006.

Lehman AF, Buchanan RW, Dickerson FB, et al. Evidence-based treatment for schizophrenia. *Psychiatr Clin North Am*. 2003;26:939–954.

Liberman RP. *Recovery from Disability: Manual of Psychiatric Rehabilitation*. Washington, DC: American Psychiatric Publishing; 2008.

Lincoln TM, Wilhelm K, Nestoriuc Y. Effectiveness of psychoeducation for relapse, symptoms, knowledge, adherence and functioning in psychotic disorders: a meta-analysis. *Schizophr Res*. 2007;96(1–3):232–245.

McGuire T, Miranda J. New evidence regarding racial and ethnic disparities in mental health: policy implications. *Health Aff (Millwood)*. 2008;27(2):393–403.

Mohatt DF, Bradley MM. Rural issues in the behavioral health workforce. In: Annapolis Coalition on the Behavioral Health Workforce. An Action Plan for Behavioral Health Workforce Reform. Substance Abuse and Mental Health Services Administration, U.S. Department of Health and Hum an Services; 2007:216. Available at http://www.samhsa.gov/workforce/annapolis/workforceactionplan.pdf.

*New Freedom Commission on Mental Health. *Achieving the Promise: Transforming Mental Health Care in America, Final Report (DHHS Pub. No. SMA-03-3832)*. Rockville, MD: U.S. Department of Health and Human Services; 2003.

Rosen A. The community psychiatrist of the future. *Curr Opin Psychiatry*. 2006;19(4):380–388.

Rosenbloom M. Chlorpromazine and the psychopharmacologic revolution. *JAMA*. 2002;287(14):1860–1861.

*Satcher D. Mental Health. *A Report of the Surgeon General by avid, United States. Public Health Service. Office of the Surgeon General. Center for Mental Health Services*. Washington, DC: National Institute of Mental Health; 1999.

Sowers W, Pollack D, Everett A, Thompson KS, Ranz J, Primm A. Progress in workforce development since 2000: advanced training opportunities in public and community psychiatry. *Am J Psych*. 2011;62(7):782–788.

Tansella M, Thornicroft G. Implementation science: understanding the transition of evidence into practice. *Brit J Psych*. 2009;195:283–285.

U.S. Attorney General. Task force on defending childhood against violence, 2012. Available at http://www.justice.gov/defendingchildhood/cev-rpt-full.pdf.

▲ 58.2 Reforming Health Care

LEIGHTON Y. HUEY, M.D., JULIAN D. FORD, PH.D., AND ROBERT F. COLE, PH.D.

INTRODUCTION

Numerous studies and reports have documented for years the problems in the US health care system and the overall unfavorable comparisons to other industrialized countries when considering quality, access, efficiency, equity, public health, and costs (Table 58.2–1). A 2014 report by the Commonwealth Fund shows the US health care system ranked last among eleven industrialized countries overall in access, efficiency, equity, and achieving healthy lives.

However, with the passage of the Patient Protection and Affordable

Care Act (ACA) in 2010, a watershed event, health care reform in the United States has moved into the forefront of the national discourse despite rear-guard objections to the ACA's enactment. The ACA is the law.

The following chapter traces the movement toward health care reform, first in the 20th century, and then in the 21st century culminating in the passage of the ACA. The context of the larger health care system is discussed, the problems driving the need for reform, along with the role of and implications for behavioral health and psychiatry.

Based upon his considerable influence, virtually all of US health care has rallied around Don Berwick's 2008 "Triple Aim" concept that, simply put, involves:

- Improving the individual patient's experience of care
- Improving the health of populations
- Reducing per capita costs of care

Meeting these three inter-related concepts is required to achieve high-value care as reform shifts from a *volume-based* to a *value-based* system. The concept of the Triple Aim will guide the discussion in this chapter. The larger health care systems in which psychiatrists and other behavioral health specialists work have a profound impact on the delivery and outcomes of both psychiatric and physical health care services. Internationally, and over the last decade in the United States, the dominant theme in reform of health care systems shifted from an emphasis on managed care to a tripartite focus on patient-centered care (as a complement to provider-driven disease management), improvement of the health of entire populations with interventions tested in comparative effectiveness research (as opposed to individual level observational or clinical trial research), and the development of incentives to maximize health care cost-effectiveness (as opposed to containing costs by restricting access to or reimbursement for health care). The Triple Aim serves as a foundation for health care reform (see Fig. 58.2–1).

As shown in Figure 58.2–1, the guiding principle of the Triple Aim is that health care quality should be defined based upon the vantage of the recipient, the patient, and not externally by technical or clinical considerations. Beginning with a patient-centered (as opposed to disease-based) perspective, health care must be teamed up with public health (i.e., prevention and health promotion) and social services in order to fully meet the needs of individuals and families and create a healthy overall population. This biopsychosocial approach integrates otherwise fragmented or missing services, and makes primary (rather than specialty) health care the driver for the delivery of services—containing and reducing costs as well as promoting health. These are not new or radical concepts, but have been recurrent themes in efforts to reform health care for more than a century in the United States as well as playing a central role in the

more successful health care systems in other affluent and industrialized countries internationally.

Table 58.2–1.
Comparison of the US Health Care System to Other Industrialized Countries

COUNTRY RANKINGS
Top 2*
Middle
Bottom 2*

	AUS	CAN	FRA	GER	NETH	NZ	NOR	SWE	SWIZ
OVERALL RANKING (2013)	4	10	9	5	5	7	7	3	2
Quality Care	2	9	8	7	5	4	11	10	3
Effective Care	4	7	9	6	5	2	11	10	8
Safe Care	3	10	2	6	7	9	11	5	4
Coordinated Care	4	8	9	10	5	2	7	11	3
Patient-Centered Care	5	8	10	7	3	6	11	9	2
Access	8	9	11	2	4	7	6	4	2
Cost-Related Problem	9	5	10	4	8	6	3	1	7
Timeliness of Care	6	11	10	4	2	7	8	9	1
Efficiency	4	10	8	9	7	3	4	2	6
Equity	5	9	7	4	8	10	6	1	2
Healthy Lives	4	8	1	7	5	9	6	2	3
Health Expenditures/Capita, 2011**	$3,800	$4,522	$4,118	$4,495	$5,099	$3,182	$5,669	$3,925	$5,643

Notes:*Includes ties. **Expenditures shown in $US PPP (purchasing power parity); Australian $ data are from 2010.

Source: Calculated by The Commonwealth Fund based on 2011 International Health Policy Survey of Sicker Adults; 2012 International Health Policy Survey of Primary Care Physicians; 2013 International Health Policy Survey; Commonwealth Fund *National Scorecard 2011*; World Health Organization; and Organization for Economic Cooperation and Development, *OECD Health Data, 2013* (Paris: OECD, November 2013).

The benefits of designing health care delivery practices and systems based on the Triple Aim are described by the Institute for Healthcare Improvement as:

Organizations and communities that attain the Triple Aim will have healthier populations, in part because of new designs that better identify problems and solutions further upstream and outside of acute health care. Patients can expect less complex and much more coordinated care and the burden of illness will decrease. Importantly, stabilizing or reducing the per capita cost of care for populations will give businesses the opportunity to be more competitive, lessen the pressure on publicly funded health care budgets, and provide communities with more flexibility

to invest in activities, such as schools and the lived environment, that increase the vitality and economic wellbeing of their inhabitants.

The Triple Aim's relevance for psychiatry, and a key role that psychiatry and behavioral health can play in health care reform, was articulated by Wayne Katon, an academic psychiatrist who has been a pioneer and leading proponent and researcher in the area of collaborative mental and physical health care. He stated in a 2013 interview:

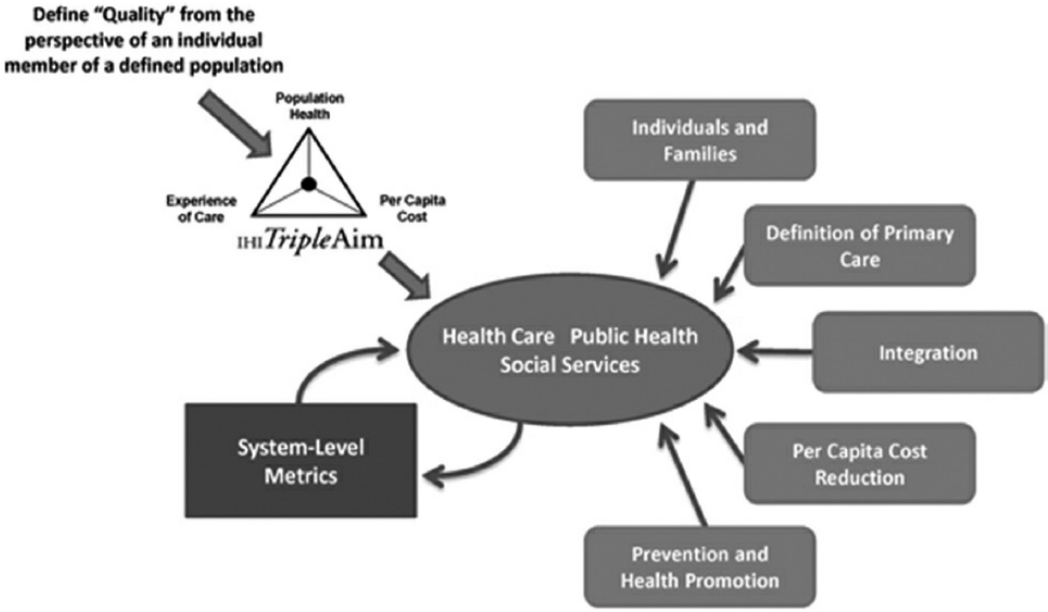

FIGURE 58.2–1. Design of a Triple Aim Enterprise. (Reprinted with permission by the Institute for Healthcare Improvement Triple Aim; 2012.)

Health reform initiatives will bring primary care, hospitals, and other centers into accountable care models in order to improve continuity and quality of care and reduce costs. We don't want psychiatrists to be left behind in the changes that are rapidly going on. Many psychiatrists are involved in efforts to integrate mental health services into primary care and medical specialty care as well as collaborative care and dissemination trials. Health reform has provided a tremendous stimulus for positive change in terms of integrating mental health care into primary care . . . There is also an overrepresentation of people with mental illness among the currently uninsured individuals who may gain access to insurance through health reform. That's another incentive to adopt collaborative care models to help provide mental health care more efficiently. Otherwise, this influx of patients will really stress the system.

Integration of psychiatry into a comprehensive "patient care medical home" that provides a full range of coordinated cost-effective physical and mental health services. The integration continuously monitors the patient's response in order to enable timely adjustments to the services involving a transformation that will affect all psychiatric and behavioral health providers and sub-specialties. This chapter will describe the challenges and opportunities facing psychiatry and behavioral health as a field and psychiatrists and others as practitioners as the Triple Aim guides health care reform in the 21st century.

HISTORICAL OVERVIEW

Health Care Reform in the 20th Century

The reform of the US health care system has been periodically and repeatedly proposed since and before the social reforms of President Franklin Roosevelt. Presidents Harry Truman, Jimmy Carter, Gerald Ford, George H. W. Bush, and Bill Clinton all considered and proposed plans for reform. Earlier proposals for national health insurance set the framework for these proposals. Table 58.2–2 provides a summary of the major periods of legislative activity since the early 20th century and the complex elements that were involved in the proposals. Earlier proposals coupled health with other social welfare programs; later proposals narrowed in on health care as health care costs increased dramatically in the later part of the 20th century. Proposals were designed on either the model of social insurance or that of public assistance or welfare. As the importance of universal coverage became clear, these models merged. The locus of administration or control of the proposed health insurance programs—by employers, the federal government, or the states—became important. In particular, the most critical issue was the financing of the system: Through government funds, the requirement that employers pay for their employees ("employer mandate"), or the requirement that individuals buy their own insurance ("individual mandate").

In the last decade of the 20th century, as health care costs continued to spiral upward, the proposed Clinton Health Security Act generated a tremendous amount of energy but eventually failed to be enacted. Large employers changed their methods of purchasing care, hiring companies to introduce a variety of cost control techniques, generally described as "managed care." Meanwhile "incremental" measures were taken in the 1996 Health Insurance Portability and Accountability Act; the 1997 Balanced Budget Act, which limited payments to hospitals; and the 2003 Medicare Prescription Drug Improvement and Modernization Act, which, in addition to providing partial coverage to Medicare enrollees for medications, also provided the Medicare program with the flexibility to introduce more aggressive health care purchasing techniques to control cost. These measures also supported individual incentives to control cost through new insurance models such as medical savings accounts, which were designed to combine tax-exempt personal savings for health care with catastrophic insurance coverage. The movement toward "managed care" purchasing did slow the rising costs for a few years, but its requirements had to be relaxed due to the strong reaction against control by both patients and providers.

Table 58.2–2.
Five Periods of American Legislative Activity for National Health

Insurance

Dates: Proposed Reform	Benefit	Administration/Financing	Model/Outcome
1912–1919 American Association for Labor Legislation (AALL): Opposed by American Medical Association (AMA), American Federation of Labor, businesses	Medical care, sick pay, funeral expenses for lower-paid workers (<$1,200/year) and dependents	Administered by states; financed by payroll tax collected in regional funds	Model: Social insurance Failed to pass
1946–1949 Wagner-Murray-Dingle Bill: Supported by President Truman; opposed by AMA	Retirement, disability, unemployment, and health insurance: Physician and hospital care	Expansion of Social Security; employer/employee contributions to "social insurance trust fund"	Model: Social insurance Failed to pass
1963–1965 Medicare and Medicaid	Medicare: Hospital care (Part A) physician care (Part B); medications (Part D) Medicaid: Comprehensive health care, long-term care	Medicare administered federally, Medicaid by the states; Medicare financed through Social Security, federal income taxes, and individual premiums; Medicaid financed by general state and federal tax revenues	Medicare: Social insurance model Medicaid: Public assistance (welfare) mode Both programs were enacted
1970–1974 Kennedy-Griffiths Health Security Act: President Nixon countered with an employer mandate plan; Physicians for National Health (PHN) proposed "single payer"	National health insurance for the entire population	Administered federally, eliminating all public and private insurance; Nixon plan featured employer mandate, privately administered, plus Medicaid and Medicare; PNH plan financed by federal tax with single fund in each state	Model: Combined social insurance and public assistance; failed to pass Nixon plan: Maintained public assistance; added social insurance; failed to pass PNH plan: Merged social insurance and welfare; failed to pass
1991–1994 Clinton Health Security Act: Opposed by insurance and pharmaceutical industries and small	Comprehensive health care	Administered federally; financed through governmental funds, employer mandate and individual mandate (with subsidies); voluntary: Tax	Model: Integrated social insurance and public

employers, single-payer advocates; proposed voluntary approaches	incentives for employers; medical savings accounts (MSAs) for individuals	assistance Failed to pass

Managed Care: A Phenomenon on the Path to Health Care Reform

The freewheeling spirit resulting from having generous public and private financial support for health care during roughly the period from 1950 to 1980 led to a heady time for US medicine. Society supported the clinical advances built on research that promised breakthroughs just over the horizon. Practitioners were, by association, part of this new era, and, as authorities in their professions, their determination of what patients needed was funded almost without question by public and private sources. As resources became limited owing to many factors, such as inflationary pressures and legislation, payers began to question what it was that health care was delivering for the dollar; that is, the issue of *value* was raised for the first time in the health economic equation by those paying for health care. A corollary of the issue of value was the question of the *accountability* of practitioners. Subsequent landmark studies demonstrated geographically differing practice-patterns and consequently different costs but similar outcomes for the same disorder or disease. This raised the general question of why this should be the case? Other fundamental issues that have emerged include the following: How can access to care be broadened? How can clinical outcomes as a result of health care be improved? What about the uninsured? How can costs be controlled?

Utilization Review

Employers and purchasers embraced the concepts of *utilization review* and *utilization management* as strategies to try to ensure that the most appropriate and most cost-effective clinical decisions with the best outcomes were being made on behalf of any given patient. For the purposes of this discussion, *utilization management* is considered the global concept that subsumes various aspects of the review process. These days, virtually all care, with the exception of the increasingly rare full fee for service, is under some form of review and management process, whether public or commercial payer.

The review process is considered by some practitioners as an intrusion into the clinician–patient relationship. The fact that the process can be cumbersome and that there is often disagreement about the decisions reached during a review contributes to the resentment engendered. With some justification, practitioners are irked by the fact that third parties with no clinical relationship to the patient have the authority to influence the process of care, challenging the practitioners' clinical autonomy. At one point, the clinical criteria used in the review process were considered "proprietary" and not available to clinicians so they might see how the

process of their care was being reviewed. At the same time, if practices were reasonably standardized and if practitioners had training (or retraining) that was uniform with a focus on (1) diagnostic accuracy, (2) evidence-based, cost-effective practice, and (3) attention to clinical outcomes, perhaps a great deal of the variance seen in practice might diminish, thereby reducing the need for third-party reviews.

Selective Closed Panels and Networks

Another strategy to improve care while reducing costs was to create selective or preferred provider networks in which clinicians applied to become members of the clinical panel of a managed care organization. By applying to be a network provider, clinicians agreed to practice within the guidelines established by the managed care organization, so, in theory, the network had some degree of compliance and fidelity to the standards of practice of that particular managed care company. However, this was more of an idealized system of care for a number of reasons. In many parts of the country, the clinical marketplace is highly competitive because of a glut of practitioners vying for limited clinical dollars. Functionally, what this means is that there may be many individual practitioners who apply to be on a particular managed care panel because of simple economics (i.e., they need the income) rather than an interest or willingness to practice within a managed care framework. So, there may be little that is selective about a particular panel of providers because the process is more open-ended until the managed care organization has determined that there are sufficient numbers of providers for a given geographical area. Whether justified or not, some managed care organizations have reputations for being difficult to work with, reimbursing at what are viewed as "low rates." Such organizations may have difficulty attracting sufficient numbers of providers and may be less selective than is preferable. In more affluent areas, some practitioners, especially well-established ones, do not accept insurance of any kind so utilization review is not part of overview of their patients' clinical care.

Care Management and Intensive Case Management

A large portion of psychiatric care is to be done with persons with chronic mental health conditions. Driven largely by demographic realities of an aging population, health reform considerations are focusing on chronic illness care and the concept of *care management and intensive case management, virtually identical concepts*. By analyzing the process of care for various disease entities, care management attempts to introduce discrete improvement steps to improve the efficiency and coordination of care, enhance patient compliance, achieve better outcomes, and reduce costs. Care management has been used with chronic illnesses, such as

diabetes or hypertension. In psychiatry, care management has been applied to mood and anxiety disorders with success. The high degree of practice variability not only by provider, but also by geographic section of the country reflects the historic inability and failure of the health care system to incorporate readily and to standardize quality improvement and evidence-based ideas into the system in which clinician, patient, and health system engage.

Managed Behavioral Health Care

In the late 20th century, costs for psychiatric and substance-abuse services were rising at a considerably faster rate than the costs for physical health care. The effort to instill accountability among practitioners, expect value for payment, improve access, provide cost-effective services, and use to the extent possible evidence-based approaches to treatment to improve outcomes created a backlash by mental health practitioners against managed care. Scandals in the 1980s and 1990s of for-profit managed behavioral health care companies engaging in unethical practices overshadowed the equally scandalous accounts of practitioner excesses of use and inappropriate, often misguided, treatment. Managed behavioral health care became synonymous with denial of care and, in for-profit systems, the image of fat-cat corporate leaders with large salaries and generous stock options derived from the misfortune of patients and beleaguered practitioners, the latter mired in an endlessly confusing system of external review, was readily accepted by many as the only reality. What was missing in the tumult was the perspective that managed behavioral health care was a way station on the path to behavioral health care reform.

Public Sector Managed Behavioral Health Care

Governmental agencies, at federal, state, and local levels, purchase 60 percent of all behavioral health care. As resources followed patients out of the state hospitals, much of the state expenditures remained in the control of agencies responsible for specific populations such as mental health, criminal justice, child welfare, and social services. Medicaid, the entitlement program for those unable to afford health care, began in 1965, and is the primary funding source for local governments to pay for public behavioral health care services. Medicaid's overall costs rose dramatically, by 2000, accounting for more than 15 percent of all US health care costs, with behavioral health care accounting for about 10 percent of Medicaid spending.

At the level of state or county government, Medicaid's behavioral health costs became a serious liability because of yearly cost overruns. Government agencies turned to managed care contracting of Medicaid services as a way of shifting risk to non-profit behavioral health care

provider organizations, and thus limit the state's liability.

State governments also soon learned that they could redirect the bulk of their behavioral health expenditures for severely disabled persons in the categorical agencies into the Medicaid program, thus increasing their match base for the Medicaid Federal Financial Participation (FFP) and leveraging federal revenues into their state general fund. Medicaid expenditures continued to increase, and this was a major contributing factor to the 100 percent increase in public mental health dollars from 1956 to 2003. The downside from the point of view of the seriously and persistently mentally ill was that the control of these services went from the public agencies responsible for behavioral health care to commercial managed care companies whose networks and capacities were more in tune with the needs of employed populations.

Safety-net hospitals, community mental health centers, and academic medical centers were under extreme financial distress and endured substantial service limitations. Thus, while public sector experiments with large-scale managed behavioral health care produced some successes, costs and quality of care continued to be problematic.

Private Sector Managed Behavioral Health Care

Government deregulation of the hospital industry enabled the astronomical growth of for-profit mental health and substance abuse facilities, many organized as large hospital chains. Use of mental health and substance abuse treatment increased dramatically, exemplified by the fourfold increase in adolescent inpatient admissions that occurred between 1982 and 1987. The proliferation, financial success, and perceived abuses of proprietary psychiatric hospitals caused great consternation among employers, insurers, legislators, and the public at large. Many hospitals, including presumably less exploitative, not-for-profit hospitals, seemed to provide treatment based on reimbursement potential rather than actual medical necessity. One of the physician pioneers of the managed behavioral health care industry commented that it was a myth that the treatment of substance abuse required a fixed length (usually 28 days) of inpatient treatment. Rarely has such a fixed false belief had so little factual support, or been so fervently adhered to. As early as 1983, a report from the Office of Technology Assessment that the literature failed to demonstrate a general superiority of inpatient over alternative treatment settings, as measured by outcome. Subsequent studies have confirmed that hospital or residential treatment is neither a necessary nor sufficient for most people with substance use disorders.

Long inpatient stays for adolescents with behavioral or substance abuse problems nevertheless became common. Inappropriate hospitalizations for excessive periods of time became exposed as proprietary systems flourished. Health care payers (employers, insurance companies, Medicare

and Medicaid) all became quite rightly concerned by the rising costs and the suspect practices. They questioned the true medical necessity of costly services such as long psychiatric inpatient stays. The psychiatric profession also successfully argued that inpatient stays for psychiatric patients could not be actuarially determined with any precision, thus making federal Medicare mental health inpatient criteria exempt from federally determined diagnostic-related groups (DRGs) for reimbursement purposes. In this climate, it became more and more difficult to maintain adequate oversight or regulation of mental health and substance abuse treatments. Private employers became the first to respond proactively to these concerns by contracting with the pioneer entrepreneurs of the first generation of managed behavioral health organizations (MBHOs) led by mental health professionals who had the credibility and specialized knowledge and skill to provide oversight of mental health and substance abuse practitioners.

This initial phase of managed behavioral health care (approximately 1983 to 1987) began with traditional indemnity insurers changing benefit designs in the hope of discouraging excess use. Tactics included limiting mental health coverage for specific disorders or treatments and instituting deductibles, copayments, and benefit maximums. These efforts were ultimately all unsuccessful. Entrepreneurial companies, led by psychiatrists and psychologists, began their entry into the health care marketplace through contracts with HMOs struggling to administer behavioral health care benefits. The dramatic success of these initial efforts was driven through limitation of access to mental health and substance abuse services. Access was limited through tight medical management (instituting strict medical necessity criteria for level-of-care determinations, limiting authorizations for inpatient admissions and continuation of stay, proliferation of denials of payment for care, etc.), as well as by restricting the network of caregivers and institutions credentialed to provide care for the population covered.

The **first phase** of limitation of access, aimed at restricting access for patients, was successful but led to dramatic customer complaints. As large employers (not HMOs) became the target for MBHO marketing, patient choice was increased by expanding provider networks, but overall costs were now saved through dramatic reduction in rates paid to providers. Typical discounts were 20 to 25 percent of professional fees and 25 to 35 percent of hospital charges. These discounts seemed draconian at the time and were widely seen as threats to the solvency of behavioral health care providers; but even larger fee reductions were to come. This network expansion and reduction of provider rates characterized the **second phase** of the managed behavioral health care revolution (approximately 1993 to 1998).

The **third phase** (1998 to the present) has been termed the phase of

industry consolidation. The proliferation and financial success of MBHOs in the 1980s and early 1990s led to the development of administrative models, information system infrastructure, large provider network management techniques, and sophisticated medical management strategies that resulted in significant economies of scale. Provider rates continued to fall, along with profit margins. Utilization patterns also changed dramatically. The average commercial inpatient days per 1,000 members declined from approximately 150 in 1983, to 75 in 1990, to 10 to 15 in 2000 in tightly managed markets, such as Minneapolis and Los Angeles. Cost to employers of behavioral health care under traditional indemnity plans was approximately $9 per member per month. By 2000, these costs fell dramatically to less than $3 per member per month. To sustain these reductions, the MBHO industry experienced dramatic consolidation. Within a few years, through mergers, acquisitions, and growth of the market, the industry became concentrated into a small number of very powerful behavioral health care corporations.

Health Care Reform in the 21st Century

Whether admired or not, many of managed care's innovations are irreversible. Managed care's financial success was dramatic (i.e., decreased behavioral health care costs). Utilization management by managed care companies led to significant decreases in average length of inpatient stays for psychiatric patients and fewer total inpatient days for populations of patients as a whole. Despite widespread complaints of diminished physician, hospital, and patient autonomy, constraints of decreased physician and hospital reimbursement, and anecdotal horror stories of sacrificed patient health care quality, there is no convincing evidence that managed care destroyed the quality of the US health care system. One could reasonably argue that the health care system was already imploding which was the impetus for yet another effort at reforming the system. Recent evidence collected by the National Committee on Quality Assurance points to significant improvements in health care quality resulting from efforts at self-improvement by the managed care industry. Managed care contributed to health care quality through its emphasis on population-, evidence-, and guideline-based health care, chronic illness care management, and information technology to evaluate and change the process and outcomes of care.

However, by the beginning of the 21st century, health care costs continued to spiral upward in the United States, and health care quality remained substantially below standards set in other comparable affluent societies (see Table 58.2–1). With overall health care costs rising at more than twice the rate of inflation, employer groups could not sustain the resultant increase in premiums.

Waste in Health Care

In recent years, considerable attention has been paid to eliminating waste in US health care, not just redesigning a dysfunctional health care system with improved access, better quality, improved efficiency, expanded coverage to enhance equity, and lowered costs. It has been estimated that a third of the annual health care expenditure ($2.8 trillion in 2014) could be eliminated without any adverse health consequence to individual patients. Unnecessary administrative expenses, care that provides little or no value, costs associated with medical errors and fraud are examples of some of the waste. There have been six areas identified that are responsible for the waste: (1) poor care delivery, (2) poor care coordination, (3) excessive care, (4) excessive administration, (5) pricing failures $, (6) fraud and abuse (Table 58.2–3).

These six areas are projected to push US health care expenditures to 20 percent of the gross domestic product by 2020 which is unsustainable. Not surprisingly, therefore, in the first two decades of the 21st century, the urgency of health care reform has intensified as the growth of US health care spending is projected to double in 8-year intervals representing almost 18 percent of the gross domestic product in 2011. The competitive position of the United States in the global economy is threatened and many argue that the economy cannot be fixed until health care is fixed. Influential reports show that the United States spends considerably more on health care than any other country, with the growth in spending consistently outstripping collective income. Yet the United States ranks poorly among industrialized countries in common measures such as life expectancy and infant mortality, and it is unable to provide health insurance to at least 15 percent of its population. Health care is at the forefront of the domestic agenda.

Table 58.2–3.
Waste and Costs

Waste and Costs in Health Care in 2011	Billions ($)
► Poor Coordination of Care	35
► Poor Delivery of Care	126
► Excessive Care	192
► Excessive Administration	248
► Excessive Costs to What Should Be	131
► Fraud	177

Adapted from Berwick D, Hackbarth A. Eliminating waste in US health care. *JAMA*. 2012;307(14):1513.

It is also clear that reform of an industry approaching $3 trillion in annual costs will involve a gradual transformation over many years. The

passage of the ACA in 2010, now law, has laid the foundation for the transformation. Key policy reports in the first decade of the 21st century set the stage for the ACA.

The 2001 Institute of Medicine *Quality Chasm* Report

The Institute of Medicine (IOM), in its influential 2001 report, *Crossing the Quality Chasm: A New Health System for the 21st Century*, noted "Between the health care we have and the care we could have lies not just a gap, but a chasm." The report went on to highlight the disparity between best practice and current practice and the effect of this on widespread variability and escalating health care costs. It established Six Aims for the twenty-first–century health care system.

Expanding the concept of best practice to include public health and prevention efforts and coupling this with chronic illness care management are among the key points of the IOM monograph, as are the pivotal roles of pre-professional training and the retraining of the existing workforce. With finite resources for health care amid competing priorities, the reformed health care system and, specifically, the reformed behavioral health care system must find ways to integrate effectively multiple factors, such as (1) practitioner, system, and patient accountability; (2) economic support for public health efforts, such as prevention and early identification; (3) the linkage of acute care with chronic illness care/collaborative care/coordinated care management; (4) improved access; (5) effective integration with primary and specialty care; (6) development of models of the most clinically effective and cost-effective care (e.g., integrated with public and physical health, interdisciplinary in practice, with continuity of care across venues); and (7) meaningful involvement of patients and, where feasible, their families, in the process of care, and so on. Managed care can be viewed as a sign and a symptom within the broader construct of health care reform. Health care reform and reform of managed care are possible.

Soon after this report was issued, The National Committee for Quality Assurance (NCQA) reported that despite important strides made in the general medical and surgery sectors, there continued to be disturbing exceptions to overall health care quality improvement in mental health, that the situation remained poor and concerning. The IOM's recommendations centered first on linking mental health and substance use treatment more closely with general medical care; second, the details of the original *Quality Chasm* report, its aims, rules, and redesign approaches, should be operationalized in the treatment of mental health and substance use disorders. Inadequate leadership and resources have been cited along with these areas having fewer objective, standardized diagnostic methods compared to physical health disorders; weakness of the evidence base for supporting quality measures; lack of coherent strategies for adopting quality measures that gets back to the problems in leadership

in assuming this responsibility to move the field along. Fewer than 5 percent of the National Quality Forum's list of 650 quality indicators relate to the care for individuals having mental health and substance use problems.

The 2006 Institute of Medicine Report: Improving the Quality of Health Care for Mental and Substance Use Conditions

The 2001 IOM Quality Chasm Report articulated a strategy for health care improvement but because health care for mental and substance use disorders has numerous distinctive characteristics, a separate report focused on this specific area was undertaken. These distinctive characteristics include the greater stigma associated with mental and substance-use conditions, more frequent coercion of these patients into treatment, a less developed system for measuring quality and improvement, a more educationally heterogeneous workforce, the fragmentation of the behavioral health care system with inadequate communication and coordination of care among its components, and a perverse funding structure. This report and its related series called for a profound paradigm shift for US health care and proposed a set of ten new rules for how care should be delivered. The changes were intended to foster the Six Aims and support redesigned health care systems. The differences between the old and new rules are outlined in Table 58.2–4.

The report's findings and recommendations are summarized in Table 58.2–5.

At the time of the publication of this report, more than 33 million Americans were using health care for their mental conditions or conditions resulting from their use of alcohol, inappropriate use of prescription medications, or illegal drugs.

The report also noted the considerable societal and workplace burden that mental and substance use conditions bring with substantial employer costs from absenteeism, "presenteeism" (i.e., being at work but essentially not functioning fully), sub-par educational achievement by children having such conditions leading to a variety of emotional and behavioral problems. Children with poor educational achievement enter a cycle of emotional and behavioral problems leading to risk for delinquent and antisocial behaviors, dropping out of school, eventually affecting the child welfare and juvenile justice systems. This in turn becomes a feeder phenomenon into the adult justice system that has seen an increasing number of inmates in prisons and jails with mental and substance-use conditions. The U.S. Bureau of Justice Statistics estimated that 16 percent of all persons in jails and prisons have a mental and/or substance-use disorder. Systematic assessment studies document prevalence estimates that are twice as high, with as many as one in every three incarcerated adults and youths having a behavioral health disorder.

Table 58.2–4.
The Institute of Medicine's Rules for Health Care

Old Rules for Health Care Delivery	New Rules for Health Care Delivery
1. Care is based on visits.	1. Care is based on continuous healing relationships.
2. Professional autonomy drives variability.	2. Care is customized to patient needs and values.
3. Professionals control care.	3. The patient is the source of control.
4. Information is placed in a record.	4. Knowledge is shared, and information flows freely.
5. Decisions are based on training and experience.	5. Decision making is evidence based.
6. "Do no harm" is an individual clinician responsibility.	6. Safety is a system responsibility.
7. Secrecy is necessary.	7. Transparency is necessary.
8. The system reacts to needs.	8. Needs are anticipated.
9. Cost reduction is sought.	9. Waste is continuously decreased.
10. Preference for professional roles over the system.	10. Cooperation among clinicians is a priority.

The Institute of Medicine committee responsible for the 2006 report, *Improving the Quality of Health Care for Mental and Substance Use Conditions,* found that the framework of the Quality Chasm Report could be applied to health care for mental and substance use conditions and the following strategy was described:

- Individual patient preferences, needs, and values prevail in the face of residual stigma, discrimination, and coercion into treatment;
- The necessary infrastructure exists to produce scientific evidence more quickly and promote its application in patient care
- Multiple providers' care of the same patient is coordinated
- Emerging information technology related to health care benefits people with mental or substance-use problems and illnesses
- The health care workforce has the education, training, and capacity to deliver high-quality care for mental and substance-use conditions
- Government programs, employers, and other group purchasers of health care for mental and substance-use conditions use their dollars in ways that support the delivery of high-quality care
- Research funds are used to support studies that have direct clinical and policy relevance and that are focused on discovering and testing therapeutic advances

An over-riding theme throughout the report was that *treatment can be effective* and that *recovery is possible*. The report noted that quality problems hinder effective treatment and recovery, noting that numerous studies have shown the discrepancy between what is considered *effective care* and what is actually provided, echoing the Quality Chasm report. Low rates of adherence to clinical practice guidelines and evidence-based practice standards were noted across major diagnostic categories. For example, adequate rates of adherence to established practice guidelines occurred in only 27 percent of studies reported that covered alcohol withdrawal, bipolar disorder, depression, panic disorder, psychosis, schizophrenia, and substance use.

The 2003 New Freedom Commission and the 2008 Commonwealth Fund Report

The New Freedom Commission on Mental Health's 2003 report, *Achieving the Promise: Transforming Mental Health Care in America* opined, similar to the conclusions of the IOM's *Crossing the Quality Chasm* report, that the behavioral health care system "is in shambles." Five years later, the Commonwealth Fund described the US health care system as fragmented, delivering poor-quality, high-cost care. The report noted that achieving a high-performance health care system was not possible without reorganization at the national, state, community, and practice levels. It further identified key factors contributing to the fragmentation, including the absence of a single national entity or set of policies guiding the system, the states dividing responsibilities among multiple state agencies, and care often not coordinated among practitioners seeing the same patient. The report noted that:

- *Patients and families navigate unassisted across different providers and care settings, fostering frustration and dangerous patient experiences*
- *Poor communication and lack of clear accountability for a patient among multiple providers lead to multiple errors, waste, and duplication*
- *The absence of peer accountability, quality improvement infrastructure, and clinical information systems foster poor overall quality of care*
- *High-cost, intensive medical intervention is rewarded over higher-value primary care, including preventive medicine and the management of chronic illness.*

Table 58.2–5.
Summary of the Institute of Medicine's Recommendations of Care for Individuals with Mental and Substance Use Conditions

1. Emphasize patient-centered care and the importance of supporting patient decision making and preferences	The current health systems place a number of barriers to patient-centered care. Insurance discrimination is manifested by less benefit coverage for mental and substance use (M/SU) conditions. There are higher copays for these services. ► Clinicians, organizations, accrediting bodies, health plans, and payers need to support a patient-centered care (PCC) approach that respects decision-making capacity b patients. ► Stereotypes of patients being unable to arrive at decisions have resulted in coercion of care and discrimination historically.
2. Improve the infrastructure for quality improvement and evidence-based care	Significant problems with the current infrastructure for quality improvement. Assessment and treatment practices not well codified and captured in administrative data sets. Outcome measurements not widely applied; performance measurement for M/SU health care receives insufficient attention in the private sector; public sector efforts have not yet achieved consensus. Quality improvement methods are not permeating day-to-day service. This leads to systems of care that are not based on the principles of quality measurement,

and they fail to demonstrate continuous improvement.

3. Coordinate care across systems of care

- Poor linkages across systems of care; there is general fragmentation between M/SU and general health care and between mental and substance use care for both children and adults. Unclear accountability for coordination of care among providers.

4. Ensure that the National Health Information Infrastructure Initiative (NHIII) benefits people with mental and substance use conditions

M/SU conditions are not well incorporated into the strategic design of the NHIII. The proposed solutions are as follows:

- Charge the Office of the National Coordinator of Health Information Technology and the Substance Abuse and Mental Health Services Administration (SAMHSA) to jointly develop a plan that addresses M/SU health care.
- Leaders in M/SU health care should become proactive in the national process to set health care data and information technology standards.
- Public and private purchasers of health care should decrease the burden of variable reporting and billing requirements.

  Public and private purchasers of health care should encourage widespread adoption of electronic health records, computer-based clinical decision support systems, computerized order entry, and other forms of useful information technology.

5. Increase the workforce capacity

A crisis in the M/SU workforce exists. There is greater variation in the quality of the M/SU workforce than in that for general health care. There are across-the-board deficiencies in education, and there is no "core knowledge" base across disciplines. The resulting variation in licensure, credentialing, and continuing education across disciplines does not assure agreed-on core competencies or adherence to the same standards of practice.

The remedies include the following:

- Creation of an ongoing, federally funded public–private Council on the Mental and Substance-Use Health Care Workforce.
- Collaboration of the Council with institutions of higher education, licensing and accrediting bodies, purchasers, and other private sector initiatives developed by the Association for Medical Education and Research in Substance Abuse (AMERSA) and the Annapolis Coalition (http://www.annapoliscoalition.org).

6. Restructure the marketplace to promote incentives for quality and improvement

The Institute of Medicine (IOM) report finds problems in the way in which M/SU care is purchased. Government (state and local) purchasers dominate, and frequently insurance is purchased for M/SU health care separately from general health care. This results in "carve-out" arrangements. Private insurance tends to avoid coverage or offers more-limited coverage to individuals with M/SU illnesses. Government purchasers often use direct purchasing by contract rather than insurance arrangements.

- Government should pass laws designed to enact coverage parity by purchasers so that individuals with M/SU conditions are not discriminated against when enrolling in health plans.
- Examples of tools that could be used to bring about parity are risk adjustment, payer carve-outs, risk-sharing or mixed-payment contracts, and benefit standardization across health plans.
- State-level procurement for M/SU care should give the greatest weight to quality and use M/SU health care quality measures in procurement and accountability.
- State and local governments should reduce emphasis on grant-based systems of financing and increase the use of funding mechanisms that link some funds to measures of quality.

Adapted from the Institute of Medicine. *Crossing the Quality Chasm: Adaptation to Mental Health and Addictive Disorders*. Washington, DC: National Academies Press; 2006.

This report was a harbinger of the ACA. Its recommendations focused on *Payment Reform* to stimulate greater organization as well as higher performance; *payment* would move beyond fee-for-service toward *bundled payment systems* that would reward coordinated, high-value care, expanding *pay-for-performance* to reward high-value, patient-centered care (see Fig. 58.2–2).

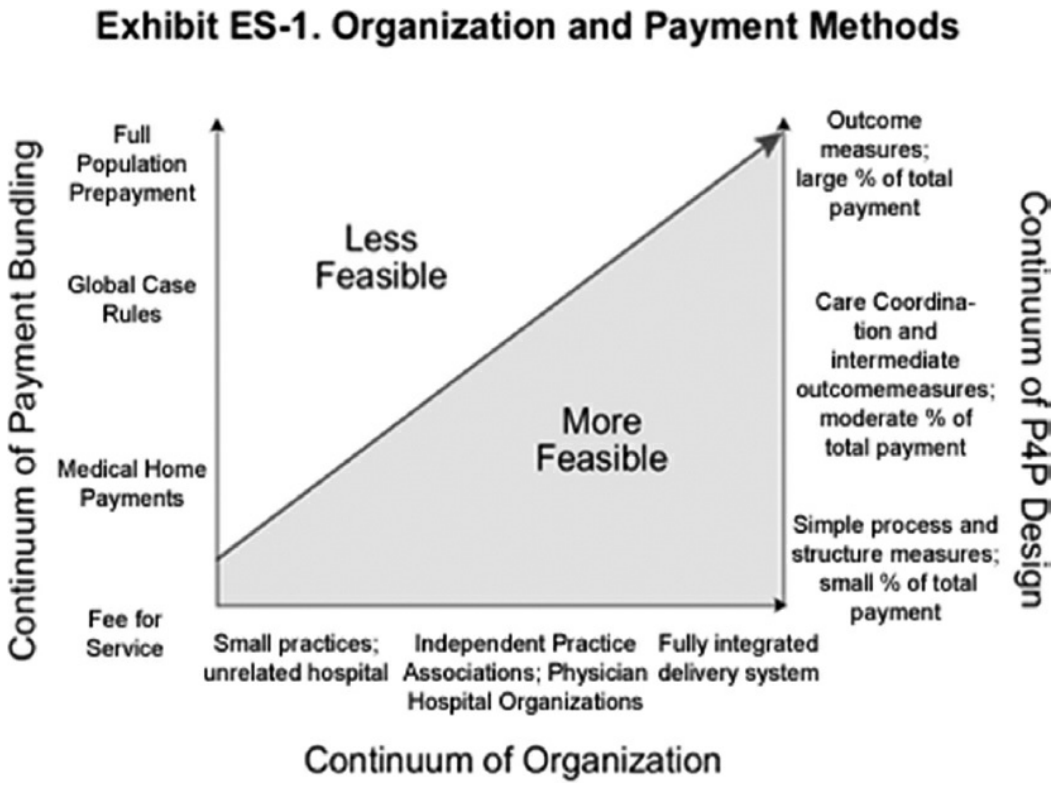

FIGURE 58.2–2. Feasibility of payment reform according to nature of the organization. (From Shih A, Davis K, Schoenbaum S, Gauthier A, Nuzum R, McCarthy D. Organizing the US Health Care Delivery System for High Performance. *The Commonwealth Fund*; 2008.)

Other recommendations included *patient incentives* (patients being able to choose to receive care from high-quality, high-value delivery systems), *regulatory changes* (to facilitate clinical integration among providers), *accreditation* (information and care is coordinated across all settings with providers having accountability to one another), *provider training* (train providers to practice in organized delivery systems), *government infrastructure support* (assist regions in establishing care coordination networks), *health information technology* (providers implementing an electronic health record that is functional, interoperable, and secure).

It is important to view these recommendations in the context of the behavioral health care system and its relationship to patients and families

and to the rest of health care. The importance of the integration of behavioral health into physical health care systems cannot be overstated. Numerous studies have shown the high prevalence of behavioral health conditions in physical health. Comorbid illnesses, that is, mental, substance use, and physical health disorders, represent the most costly group of patients in health systems. Many of these individuals are seen in public sector systems, often ending up incarcerated or homeless with inadequate physical **and** behavioral health care.

A dramatic example of the need to find ways to integrate primary care and behavioral health care in coherent delivery systems are the "dual eligibles," i.e., the Medicare and Medicaid eligible population. This is among the most costly of any population of patients, not only because a large portion of the poor and very elderly patients are in long term care (nursing homes and hospice), but also because those in the community find it difficult to negotiate consistent access to health care due to their disabilities. It is typical that a given state might find that 40 percent of their dual eligible population are persons with serious and persistent mental illness.

The 2008 Mental Health Parity and Addiction Equity Act

The Mental Health Parity and Addiction Equity Act enacted in 2008—the Wellstone-Domenici Parity Act—tightened the requirements for parity laid out in the Mental Health Parity Act of 1996. Parity was designed to bring individuals with behavioral health disorders into the mainstream US health care system to correct the historical inattention and inequality of care in this population, and to reduce the stigmatization such individuals experienced compared to people having physical health problems alone. The intent was to have behavioral health care benefits equivalent to benefits provided in medical and surgical care. In addition, about two-thirds of primary care physicians had reported difficulty in accessing outpatient behavioral health services for their patients, a rate twice as high as that for other services. Lack of coverage or inadequate coverage, health plan barriers, and a shortage of behavioral health providers were cited by primary care physicians as problems contributing to accessing good behavioral health care. The Mental Health Parity and Addiction and Addiction Equity Act required group health plans with behavioral health benefits to provide the same level of coverage as physical health plan coverage. Access to behavioral health services has been particularly problematic for low-income uninsured and Medicaid patients. Ironically, primary care physicians practicing in areas with a high number of psychiatrists had a greater probability of not being able to secure behavioral health services for their patients because of a variety of plan barriers and the inadequacy of coverage these patients had. Recent studies showing that psychiatrists, compared to all other areas of Medicine, have

the lowest acceptance rate of Medicare (i.e., less than 50 percent of psychiatrists accept Medicare, with figures lower still for Medicaid). The speculation about such an observation, its effect on how other specialties, professionals, trainees, and policy-makers, along with the general population, view adherents to such a practice is important to consider. Some have advocated tying partial or total medical school debt-forgiveness, or even state medical licensure, to having a reasonable percentage of one's practice tied to practitioners seeing individuals covered by Medicare and Medicaid.

The major impact of the ACA will strengthen the quest for parity as it attempts to reduce the number of uninsured in the United States. Otherwise parity by itself, regardless of the law, is likely to not have an effect on the severe access problems the uninsured and the low-insured population experiences in gaining behavioral health care.

THE PATIENT PROTECTION AND AFFORDABLE CARE ACT

Before the ACA, the American health care system was literally killing the country. Government estimates that the population of uninsured Americans would rise from 45 million in 2009, to 54 million in 2019. For more than 4 decades health care costs were rising much faster than the growth in the economy, sucking off tremendous amounts of money. It built tremendous academic health centers and modern community hospitals, spurred development of many new medical technologies, and filled medicine cabinets with groundbreaking new pills that were more potent with fewer side effects. But it also diverted resources away from many other vitally important social programs, from education to infrastructure to middle-class wage increases. Bridges were aging and crumbling as adjacent hospitals were brand new and stuffed with amazing technology. And even worse, over this 4-decade period cash wages for private-sector workers stagnated . . . and health care was a major contributing factor . . . the inexorable rise in the nation's debt was fueled not by defense spending or Social Security but rather Medicare and Medicaid spending as Ezekiel Emanuel noted.

After decades of fits and starts, the ACA was passed and became the nation's health care reform law in 2010. The ACA, a major centerpiece for the Obama Administration, represents the greatest change to health care since the 1960s. The rationale for reform has been articulated earlier in this chapter so reiterating the problems in access, efficiency, equity, prevention, and spending will not be belabored. The *need* for reform has been clear for decades but with increasing numbers of people not having health care coverage and the expanded complexity of health care in the context of poor coordination and burgeoning costs, reform is now imperative.

While some in Congress might have buyer's remorse and will generate predictable political machinations (enhanced by the ACA's bungled introduction to the nation in 2013), the ACA is nevertheless the law of the land and represents a major effort toward enhancing the common good. With a Republican majority in the Senate and House, and now an ACA-hostile President, controversial components of the ACA might have a chance of being repealed. However, the general sense is that the ACA is now so insinuated into the fabric of American health care, that it would be very difficult, if not impossible, to fully undo and repeal its essence. This section will address the key components of the ACA as it aims to reform both public and private health insurance systems.

The key provisions of the ACA are designed to:

- Provide coverage for more people
- Provide more benefits, better protection, and transparency
- Lower costs for consumers and government
- Provide quality and efficiency
- Provide a stronger workforce and health care infrastructure.
- Provide a greater focus on public health, prevention, and chronic illness care
- Support innovative techniques and technologies

The ACA is projected to provide coverage for more people, extending coverage to 24 million more Americans by 2023, nearly half of the 45 million uninsured in 2012. It is also predicting that the uninsured rate will be cut in half by 2016.

As of this writing, the initial results are showing the ACA is having the predicted impact with the percentage of uninsured dropping in areas where the ACA impact is being studied. Costs of health care are also declining with falling premium rates. Initially, this was partly the result of the recession that began in 2008. More recently, in some areas of the country, premiums and deductibles have risen, further driving the push to repeal the ACA by some; others sound "Reform not Repeal" as a rallying cry, i.e., work to improve the ACA, not dismantle it. However, the Center for Medicare and Medicaid (CMS) has projected that the average annual health care spending will be 5.7 percent from 2012 to 2013, compared to the 7.2 percent annual average between 1990 and 2008. Health care spending was at an historic low in 2013 (3.6 percent) but is projected to increase with an expanding economy and insurance coverage under the ACA.

Other key aspects of the ACA include care not being denied because of pre-existing conditions, maintaining dependents on parents' insurance until they are age 26, and certain preventive services are required of insures at no cost to consumers. Insures cannot charge higher premiums based on gender or health condition, and there are no lifetime or annual limits of coverage.

NEW MODELS IN HEALTH CARE

Patient-Centered Medical Homes and Medical Neighborhoods

In 2007, the American Academy of Family Physicians, the American Academy of Pediatrics, the American College of Physicians, and the American Osteopathic Association issued joint principles for the *Patient-Centered Medical Home (PCMH)*, a model to provide comprehensive primary care to children and adults, where appropriate care is structured, delivered, and coordinated around the specific needs of each patient. The medical home model was originally developed by the American Academy of Pediatrics in 1967 for children with special health needs where it was important to have a centralized location for a child's medical records. The joint principles of PCMHs (personal relationship with licensed provides, team approach, comprehensive/whole person approach, coordination and integration of care, quality and safety, expanded access to care, added value recognized, are all hallmarks) cover many of the characteristics of *Accountable Care Organizations (ACOs)* where PCMHs are seen as the building blocks for ACOs.

Some of the more recent literature suggested that early PCMH interventions led to only modest improvement in quality and patient

experience, with little evidence of effects on costs, and no reductions in utilization of hospital, emergency, or ambulatory care services over a 3-year period. At the same time, these studies suggested refinement of the outcomes associated with PCMHs was needed and that the development of the standards by the National Committee on Quality Assurance (NCQA) for medical home recognition had not been fully developed and applied in the early PCMH development. Nonetheless, the studies concluded that further refinement in PCMH development and function was important.

As the conceptualization of the organization of health care delivery systems developed, it was logical that the PCMH model would expand to include a constellation of entities and activities that would be responsible for the care of populations of patients. Thus, the concept of the *Medical Neighborhood* was spawned encompassing PCMHs with comprehensive primary, secondary, tertiary, services that are community-based, responsive to the needs of individual patients at different times, a truly integrated system of care with good coordination, effective communication, attention to patient education, and cost-effective. Such a concept forms the basis for *accountable care*.

Accountable Care Organizations

With the advent of the ACA, the Department of Health and Human Services (DHHS) began developing new approaches in conjunction with the private sector to improving care and to meet the Triple Aim for Medicare and Medicaid recipients. ACOs are voluntary groups of physicians, hospitals, and other health care providers that assume responsibility for a defined population of patients on the basis of the population's use of primary care services. The ACO is responsible for the entire continuum of care that the defined population of patients might need, including access to specialists, provision of social services, public health, and assessment of other socioeconomic factors such as housing, transportation, education, poverty, and nutrition. Providing preventive care for all patients and case management for individuals with chronic diseases is a component of an ACO as well. If the ACO succeeds in delivering high quality, improved care at a lower cost, there is an incentive built into the model whereby the ACO can share in the savings achieved through meeting the Triple Aim. Under the ACA, an ACO agrees to manage the health care needs of a minimum of 5,000 Medicare beneficiaries for at least 3 years. Under an ACO, providers are held *accountable* for the health of their patients and have strong incentives to cooperate and create savings through avoiding unnecessary care. Functionally, ACOs need to be fully integrated through an electronic medical record so that needed information can be shared among providers. Different ACO models exist to give hospitals and clinicians choice in how they work with one another.

The Teaching Health Center

There has been interest in providing post-graduate medical education in the primary care environment of the Federally Qualified Health Centers for some years. These programs have demonstrated that physicians trained in these settings are three times more likely to work in Federally-Qualified Health Centers (FQHCs) and twice as likely to serve underserved areas of the country. The ACA provided funds to establish Teaching Health Centers training physicians in nine primary care-related disciplines (one of which is Psychiatry) and at this writing 60 programs have been established in 24 states preparing 550 residents. In the same vein, the Institute of Medicine issued a report in 2014, "Graduate Medical Education That Meets the Nation's Health Needs" calling for the reform of financing, governance, and transparency of graduate medical education (GME). The Report calls for a 10-year transition period in which the current funding for more traditional residency programs will be modified. Transformative and innovative residency programs will be created and supported in primary care settings as well as in interdisciplinary organizations providing intensive care programs for persons with severe disabilities and multiple chronic illness conditions.

Comparative Effectiveness Research and Patient-Centered Outcomes Research

Comparative effectiveness research (CER) and patient-centered outcomes research (PCOR) are essential to fulfilling each of the three goals of the Triple Aim, and correspondingly have been allocated substantial funding support by the ACA through the Patient-Centered Outcomes Research Trust Fund (PCOTF). PCORTF receives funds from the general fund of the Treasury and from a small fee assessed on Medicare, private health insurance and self-insured plan revenues. An institute established specifically to facilitate and support PCOR, the PCOR Institute (PCORI) is slated to receive approximately $3.5 billion from the PCORTF over a 10-year period ending September 30, 2019. Each year, PCORTF transfers 20 percent of these revenues to the Department of Health and Human Services to support PCOR and the dissemination of PCOR to build the capacity of the health care system to provide and evaluate patient-centered and outcome-driven services. The remaining 80 percent of PCORTF funds are transferred to the Agency for Healthcare Research and Quality (AHRQ) in order to support and evaluate PCOR in US health care organizations.

PCOR is intended to provide methods and data repositories that can inform health care services by monitoring key health outcomes. Critical research questions for this agenda include how patients, families, and clinicians deal with the complex and often confusing choices involved in achieving and sustaining health, preventing and treating diseases and other health problems, and accessing and navigating health care systems.

Traditional medical research has produced vital advances in the understanding of disease conditions and their treatment, but it has not been intended or designed to enhance the current understanding of how patients subjectively experience, participate in, and are affected by health care and interactions with health care providers. Patient-centered medical research does exist, but its results typically are not systematically disseminated to either clinicians or patients and families. Research linking the patient's involvement in health care and health outcomes also exists, but its findings have not been extended beyond demonstrating that a collaborative provider-patient relationship plays a key role in patient motivation, adherence, and clinical and functional outcomes. PCOR takes on these critical questions that patients and their clinicians face daily in health care encounters and in their personal lives. PCOR is a systematic approach to patient-centered CER, in which patient and clinician preferences, choices, actions, and interactions are studied in order to determine tools, strategies, resources, and courses of action that are associated with cost-effective health care outcomes. For patients, PCOR involves testing and then providing guidance about which approaches to care work best, given their particular health conditions, personal and familial concerns, socioeconomic circumstances, and sociocultural preferences. For clinicians, PCOR is designed to provide evidence-based guidance addressing the questions they face daily in practice. For insurers, PCOR produces evidence that can help them make the best decisions on how to improve health outcomes for their members while also containing costs, especially with high-burden, high-impact disease conditions.

For psychiatry and the broader mental health field, PCOR provides opportunities and an impetus to conduct real-world clinical research on topics such as these recently-funded PCORI projects:

- Integrative Medicine Group Visits: A Patient-Centered Approach to Reducing Chronic Pain and Depression in a Disparate Urban Population
- A Community-Based Executive Function Intervention for Low-Income Children with ADHD and ASD
- Improving Transitional Care Experience for Individuals with Serious Mental Illness
- A Helping Hand to Activate Patient-Centered Depression Care among Low-Income Patients (AHH)
- Integrated Care and Patient Navigators for Latinos with Serious Mental Illness
- Patient-Centered Trauma Treatment for PTSD and Substance Abuse: Is It an Effective Treatment Option?

Pay for Performance, Public Reporting of Performance, and Meaningful Use

The effort to improve quality and therefore outcomes has taken many

guises, from revamping training to produce clinicians and systems that practice in a quality paradigm, to imploring clinicians and systems to do the right thing. At its most basic level, change can also be brought about by behavioral shaping and the use of incentives and disincentives to achieve the desired result—in this case quality improvement and patient-focused, high-quality care. Two such initiatives, Pay for Performance (P4P) and Public Reporting of Performance (PR), are designed to achieve these results. P4P and PR have been common in physical health care for a number of years. CMS has adopted P4P and has launched initiatives to demonstrate different ways to implement it. More than half of commercial health maintenance organizations are using P4P, and the acceptance of performance measurement among physicians and organized medicine has been acknowledged if not fully realized. P4P and PR, in essence, link physician reimbursement and nonmonetary rewards to improve quality of care. In addition, physicians, recognized for providing high-quality care and more efficient care deliver it at a lower cost than physicians who do not participate.

P4P and PR can reward health plans, provider groups, and individual clinicians in varying ways. Payment is made based on meeting pre-specified goals. In behavioral health, P4P has had limited presence, but efforts to expand it into common disorders such as depression and substance use disorders have received considerable attention. For example, a P4P paradigm in behavioral health might involve the percentage of patients diagnosed with a new episode of depression who were treated with an antidepressant at an adequate dosage, who were given the medication for a full 12 weeks; or patients who are given at baseline and subsequently, the Patient Health Questionnaire (PHQ), where the scores are used, along with routine clinical assessment, to track response to treatment over time. If P4P and PR achieve similar penetration in behavioral health care as in physical health care, then the modeling for this approach ideally would begin in training programs.

Within the ACA is embedded the concept of "Meaningful Use." Meaningful Use entails financial incentives for providers (professionals and hospitals) in their "meaningful use" of certified *electronic health records* (EHR) to improve patient care. Examples of Eligible Professional Core Objectives are: use of computerized physician order entry (CPOE) for medication, laboratory, and radiology; generating and transmitting permissible prescriptions electronically (eRx); recording certain demographics (preferred language, sex, race, ethnicity, date of birth); recording smoking status for patients 13-years-old and older; use of clinical decision support to improve performance on high-priority health conditions; a provider who receives a patient from another setting of care or provider of care should perform medication reconciliation; a provider who transitions their patient to another setting of care or provider of care

or refers their patient to another provider of care provides a summary of care record for each transition of care or referral.

In addition, to receive an incentive payment, providers are required to submit Clinical Quality Measure (CQM) data through the EHR that measure and track the quality of health care services. These include areas such as: health outcomes, clinical processes, patient safety, and efficient use of health care resources, care coordination, patient engagements, population and public health, adherence to clinical guidelines. Going back to the IOM Quality Chasm report, measuring and reporting CQMs is seen as ensuring that the US health care system is delivering effective, safe, efficient, patient-centered, equitable, and timely care.

Integration of Behavioral Health Care and Primary Care

Although nearly 50 percent of patients with chronic medical diseases have comorbid behavioral health conditions, more than 80 percent of the behavioral health conditions remain untreated or ineffectively treated in primary and medical specialty care settings. For a third of the patients with behavioral health conditions who receive their treatment in primary care, treatment for only one in nine is evidence-based. The data on public sector patients receiving behavioral health services in stand-alone behavioral health settings have a shorter life expectancy attributed to lack of adequate primary care services. Some ACOs have employed psychiatrists for both direct care and to provide supervision of care for non-physician behavioral health staff.

Rather than physical health care systems contracting with an MBHO to provide behavioral health care services to its enrollees, some systems have developed a variety of different models that attempt to integrate behavioral health care directly into physical health care settings. This is the dominant model of integration, perhaps accounting for 90 percent of the integrative efforts underway with physical health integrating into behavioral health systems accounting for the remainder. In this arrangement, in ambulatory settings, primary care physicians and clinicians work together with behavioral health care physicians and clinicians as part of the same health care organization. Given what is known about the high prevalence of comorbid psychiatric disorders in primary care practices, a variety of approaches are used to try to determine which patients might have a behavioral health problem and, once this is done, to provide proper assessment, treatment, and follow-up. Use of brief screening instruments to identify the more common and anticipated psychiatric disorders (e.g., depression, anxiety, and substance abuse) is one approach. If it is determined that a patient has a psychiatric problem, then there are a range of options that vary according to the particular system. In some systems, depending on the scope and gravity of the clinical psychiatric problem, the PCP directly assesses and treats it but then refers to a behavioral health

clinician if the treatment is not satisfactory or if a complication arises. Given the data showing that the level of recognition, assessment, and care of behavioral health problems within primary care systems is not of sufficient caliber, efforts to train PCPs in the proper recognition and approach to patients having behavioral health disorders has been a high priority. In other systems, once a psychiatric problem has been identified, there may be more immediate behavioral health involvement. Such integrated efforts to establish direct linkages between primary health care and behavioral health care, in essence forming a true, functional, interdisciplinary team approach, have been shown to be cost-effective, to result in improved patient satisfaction, and to have good outcomes. Some primary care systems, using a PCMH model, have chosen to employ psychiatric social workers as members of a PCMH, the implication being that PCPs can provide the necessary biomedical treatment of individuals having behavioral problems. However, there are recent studies suggesting that some patients with comorbid behavioral health disorders seen in primary care settings do not receiving adequate specialty behavioral health services.

In some hospital settings, psychiatrists are embedded directly into medical/surgical teams, a departure from the traditional consultation-liaison model. This model is relatively new and represents recognition of the high comorbidity of behavioral health disorders with physical health disorders.

Behavioral Health Care Services for Children with Serious Emotional Disorders

Children and adolescents with serious emotional disorders (SEDs), such as those with suicidal depression or early-onset psychotic disorders, represent another specialized population. On an annual basis, approximately 2 to 3 percent of children 3 to 5 years of age and 6 to 9 percent of children 6 to 17 years of age use mental health services. Of these, only a small percentage is defined as having serious emotional disorder, approximately 0.5 percent. These children, particularly the adolescents, often require complex case management and intensive (e.g., residential or inpatient) services that are costly to provide and pay for.

The most problematic aspect of the care of children and adolescents with serious emotional disturbances is that five categorical agencies have responsibility to provide it. **Child Welfare** has the largest group of these children and the largest budgets. **Special Education** obligations in local school districts are second in expenditures. **Mental Health** agencies have a clear mandate, but less funding despite usually caring for those in greatest need. **Juvenile Justice** programs serving youth who come before the courts, and **General Health and Primary Care** systems, often are the first to identify children with SEDs but tend to operate separately from mental health specialty services and to rely upon psychotropic medications

even though psychosocial treatments are considered the first-line intervention for most pediatric and adolescent psychiatric disorders.

This division of responsibility for child and adolescent services makes it very difficult to organize a coherent delivery system, and in most jurisdictions confusion and fragmentation are the rule. In the 1990s, following a decade of National Institute of Mental Health (NIMH) planning grants, the Robert Wood Johnson Foundation developed the Mental Health Services Program for Youth (MHSPY) to come up with coherent systems of care for children with SEDs and their families, integrating the efforts of the responsible agencies at the community level. This was followed by a large grant program by the Substance Abuse and Mental Health Services Administration (SAMHSA). A number of the more than 80 communities that received these grants have developed advanced care management program models providing individualized, intensive care services to children (and their families) who might otherwise have been confined in institutional settings. Services under these models are often funded by risk-adjusted case rates derived from the blending of funds from different agencies' budgets and often administered by the Medicaid program. Children with SEDs were identified in a 2014 policy brief by the Center for Health Care Strategies as:

> . . . a compelling high-cost, high-need group that can potentially benefit from the integrated care provided under health homes . . . The Medicaid health home state plan option is designed to provide a comprehensive system of care coordination across medical, behavioral, and long term services and supports for Medicaid beneficiaries with chronic physical or serious mental health conditions . . . states can tailor behavioral health-focused health homes to meet the unique needs of both children and adults." The health home is envisioned as a comprehensive set of services that wrap-around the patient and address the full range of complex health problems. For children with SEDs, this could include pediatric, psychological, educational, behavior modification, case management, social work, family therapy, child welfare, and legal/juvenile justice services as well as psychiatry. When working within the health home rubric with children, psychiatrists play a vital leadership role in collaboratively informing and coordinating with multiple behavioral health and social/educational services providers as well as with the pediatric primary care provider.

Behavioral Health Care for Adults with Serious and Persistent Mental Illness

The data are compelling for establishing an integrated and universal approach to the management of chronic illnesses. By the year 2030, approximately 150 million people in the United States will have a chronic condition of some type, and of this number, 42 million will be limited in their ability to be functional and to live independently. Chronic conditions also account for 70 percent of all deaths, 50 percent of total health care costs, and $234 billion in lost productivity each year. There are a number of strategies to induce the type of sweeping changes needed. CMS is rewarding quality initiatives and is restructuring its reimbursement system to reward chronic illness care management initiatives. Its Innovation Center is one of the leading forces driving new models, new systems, again,

along the line of fulfilling the Triple Aim.

The Robert Wood Johnson Foundation has been actively engaged for years with two important programs focused on introducing innovations in this area. The basic tenets of the first program, Chronic Care Initiatives in Health Maintenance Organizations, are to change provider and patient behaviors in several ways, including (1) moving from a focus on cure to symptom relief; (2) making the patient a critical part of the care process through shared decision-making; (3) creating active roles for the patient's significant others; (4) expanding the boundaries of care among providers and between traditional medical and social services; and (5) tying patient satisfaction to clinical outcomes.

The second national program sponsored by the Robert Wood Johnson Foundation is entitled Building Health Systems for People with Chronic Illnesses. This program attempts to better coordinate medical and supportive services for the chronically ill and, in so doing, challenge the fact that most insurance systems tend to fund acute care rather than supportive care for individuals with chronic conditions. The program attempts to do the following:

- Integrate services that care is provided across diverse settings; systems of care link clinical and nonclinical support services
- Reallocate resources by redirecting existing and acute care—focused delivery and financing to better serve the chronically ill
- Promote early medical and nonmedical interventions
- Help individuals maintain independence as long as possible in their homes and communities
- Give the chronically ill and their caregivers a role in designing and improving their health care and support services
- Support the development of models of care that can be used by people with different chronic health conditions.

These demonstrations set the stage for two operational models for fundamental system changes in the effort to improve quality and outcomes. Many of the proposed and expected system reforms surround efforts to implement the *chronic illness care model* and the *collaborative care model*. The chronic illness care model was first elaborated and tested for chronic general medical illnesses such as diabetes, congestive heart failure, and asthma. Recent efforts have attempted to apply the principles of chronic illness care to depression in primary care settings. The collaborative care model, on the other hand, was first described and tested for depression in primary care and has been broadened for other psychiatric illnesses that can also be treated in general medical settings with the support of care managers and psychiatrists, such as panic disorder and posttraumatic stress disorder (PTSD).

Widespread adoption of these two evidence-based models of care will

profoundly affect the future practice of psychiatry. Both the chronic illness care model and the collaborative care model successfully integrate psychiatric principles and practice into general medical systems. It should be possible to extend variations of the chronic illness care model to the behavioral health care system to improve quality and outcomes for chronic psychiatric illnesses (e.g., substance abuse, schizophrenia), and collaborative care models for psychiatric illnesses treated in general medical settings will continue to expand in depth and breadth. Both models of integrated health care require psychiatrists to learn to function as one component in elaborate systems of care, formally and routinely collaborate with other providers (including care managers and primary care physicians), and monitor the quality of their care and outcomes using quantitative indicators.

The Chronic Illness Care Model for Behavioral Health Services

Medical education and most health care are oriented toward the needs of patients with acute illness. However, the health care system is poorly organized to address prevention or care of patients with chronic illness, but chronic illness dominates health care: perhaps 70 percent of all care focuses on the chronically ill. Evidence of effectiveness from applications of the model in both the public and private sectors is quite robust for several chronic illnesses: congestive heart failure, diabetes, chronic obstructive pulmonary disease, asthma, and depression in primary care. In models testing the chronic illness care model for depression in primary care, psychiatrists generally function as educators and system/patient-care consultants to primary care providers and care managers, those who deliver basic treatment. Psychiatrists in the system can evaluate non-responding patients and either continue to provide care for non-responding patients or refer them back to the primary care provider with a recommendation for continuing care. The Bureau of Primary Healthcare (BPHC) and the Institute for Healthcare Improvement applied this model of care to their underserved populations. Using the PHQ-9 as the basic measure of severity and outcome, the faculty of these collaboratives developed 13 indicators to monitor individual patient progress and aggregate quality for specific populations with depression. These include indicators such as: percentage of patients with major depression who have a documented clinical contact within 1–3 weeks after initiating antidepressant treatment; percentage of patients who have a PHQ-9 reassessment within 4–8 weeks of beginning treatment; percentage of patients who have a 5-point drop in PHQ-9 score; percentage of patients who have a 50 percent drop in PHQ-9 within 4 months of treatment initiation, etc. Psychiatrists, like their general medical counterparts, who practice in settings using a chronic illness care model will need to function in new ways, including coordinating care with care managers, monitoring

patient progress and outcome with quantitative measures, providing routine support of patient self-management, and using real-time decision support.

The Collaborative Primary Behavioral Health Care Model

The collaborative care model originated with pioneering work on the identification and treatment of patients seen in primary health care who are impaired by depression or anxiety. These patients require more complex treatment plans than other primary care patients, and often do not adhere or respond positively to treatment regimens, thus incurring substantial additional costs and creating a burden for primary care providers (PCPs) and society. Katon and Unutzer note:

> After controlling for socioeconomic factors and medical comorbidity, primary care patients with anxiety and depressive disorders have been shown to have up to twofold greater medical costs compared with patients without mental health disorders . . . Patients with anxiety and depressive disorders have also been shown to present to primary care approximately 70% of the time with physical rather than mental health complaints . . . Those presenting with physical complaints are more likely to have delay in mental health diagnosis or misdiagnosis . . . Patients with anxiety and depressive disorders have also been found to have two to threefold more physical symptoms on medical review of systems even after controlling for medical comorbidity, which may explain the large increases in medical costs these patients have.

In order to address these challenges, similar to those that led to the development of the chronic illness care model, the collaborative care model integrates primary medical care and psychiatric care with a third "team member" who serves as a care manager. The care manager (typically a social worker or nursing professional) works in the primary care practice and meets at greater length and more frequently with the patient than either the primary care physician or psychiatrist. The care manager provides the patient with health education, monitors the severity of the patient's mental health symptoms (often with brief validated questionnaires) and impairment, develops a personalized practical plan for the patient to follow in putting prescribed medical and psychiatric treatments into effect on a daily basis, and monitors and assists the patient in achieving and maintaining adherence to the medical and psychiatric treatment regimen.

In the collaborative care model, the psychiatrist primarily serves as a consultant to the PCP, with the care manager serving as communication bridge between the doctors and as a physician extender for both the PCP and psychiatrist. The psychiatrist may have no direct contact through appointments with the patient, unless symptomatic worsening indicates a need to transfer the mental health care to the psychiatrist in order to prevent or resolve crises. The key role for the psychiatrist is managing the psychotropic regimen and patient adherence to it. This model has robust research support from more than 75 studies over the past two decades, and

has been implemented on a large scale in several ongoing projects.

Although very similar to the chronic illness care model (as applied to depression in primary care), the collaborative care model generally emphasizes the role of the care manager and has been extended to other psychiatric illnesses, such as panic disorder and PTSD. Collaborative care generally rests on shared care among the three providers. The role of the psychiatrist is generally more formally integrated than in chronic illness care applications (some chronic illness care implementations have been successful without a psychiatrist). Extension of the collaborative care model will require psychiatrists to learn to function as part of a team caring for psychiatric patients in general medical.

The Care Coordination Model

The concept of Care Coordination is built into the ACA framework and is one of the core concepts associated with an improved health care system. Care Coordination involves the integration and efficient coordination of care that spans all conditions, needs, capabilities, and preferences an individual patient might have, catalyzes meaningful, timely communication across all providers, and ensures that the communication involves patients and their families. The goal is to reduce fragmentation of care and improve the experience of care. Whereas *collaboration* involves the process of working together, *coordination* involves a higher-order functioning of developing and guiding a therapeutic plan and integrating the inputs from multiple sources. Clearly if Care Coordination is to be effective in the way envisioned by the ACA, the multiple steps involved in such coordination, virtually unprecedented in health care, will require time and patience. The question posed to payers is whether adequate compensation has been built into the incentivizing component of meaningful use to drive what is needed in true Care Coordination. Communication between generalists and specialists has been a documented problem for decades. For example, one study found that communication from specialists back to referring generalists occurred from 55 to 80 percent of the time; another study reported that referring generalists reported receiving communication from specialists 50 percent of the time within 3 months after the referral was made. On the other side, a survey of pediatric specialists reported that less than 30 percent receive communication about the majority of patients referred from other physicians and two thirds of specialists receive no information from referring physicians before seeing those patients. Care Coordination goes beyond this aspect of communication and coordination, involving not just health care providers but potentially community and state agencies, schools, workplaces, etc.

Role of Training in Health Care Reform

A hundred years after the Carnegie Foundation published Abraham Flexner's penetrating and important work on *Medical Education in the United States and Canada* in 1910; the Carnegie Foundation released its *Educating Physicians: A Call for Reform of Medical School and Residency* in 2010. The authors noted that in undergraduate and graduate medical education, inadequate attention was being paid to patient populations, health care delivery, effectiveness of care, and that there was insufficient focus on attending to the knowledge and skills required by trainees to meet the health care needs of the United States. One observation was that the inculcation of fundamental professional values was being impeded by the pace and commercial nature of health care. In the process of training and educating, there needed to be standardization of learning objectives with a focus on the individual learner, have formal knowledge and clinical experience better integrated, help trainees develop habits of inquiry and innovation, place greater emphasis on professional identity formation, and train in integrating education, advocacy, investigation, innovation, and systems management.

The Annapolis Coalition on the Behavioral Healthcare Workforce has called for sweeping reform in both pre-professional training and in the retraining of the established workforce across disciplines as part of the movement toward health care reform. Standards for training tend to lag behind societal need and the politics of change. Introducing reform concepts into any discipline is challenging for that particular discipline. Historically disciplines have a self-concept of being separate from others and focus on those elements that make them unique. In addition, standards for training in any discipline may be subject to intra-discipline politics as "schools of thought" vie for inclusion or even pre-eminence (e.g., in psychiatry, the historical rift between biological and psychological emphases). The extent to which standards for training have considered how health care reform fits into pre-professional training is unclear. It would seem that for a discipline to be relevant, it must incorporate a societal reform perspective. The Accreditation Council for Graduate Medical Education (ACGME) set in motion some profound changes in establishing its core competencies that determine proficiency across all of medicine: Practitioners must be competent in patient care, medical knowledge, interpersonal communication, practice-based learning, systems-based practice, and professionalism. The IOM's six aims for health care—that it be safe, effective, patient centered, timely, efficient, and equitable—interdigitate well with the ACGME core competencies. The Accreditation Council for Continuing Medical Education (ACCME) is placing greater emphasis on professional life-long learning, linking education with quality improvement and patient safety rather than simply "meeting" compliance standards for "CME credit." The American Board of Medical Specialties (ABMS) is focusing on meaningful quality improvement for physicians as

part of their Maintenance of Certification (MOC) for their specialty boards.

The retraining of the existing workforce in a reform perspective is complicated. People tend to work in learned patterns, and if there are no incentives or disincentives to change those patterns, the likelihood of change is reduced. Once trained, do clinicians pretty much do the same thing, or do they incorporate meaningful advances into the ways they practice?

The impetus for health care reform is broad, with issues pertaining to economics, access, outcomes, equity, quality, and increasingly, interdisciplinary teamwork being at the forefront. Many appear to have interpreted the raison d'être for health care reform as primarily economic, and the main mechanism used is withholding care to save costs. On this basis, many understandably would find the concepts of reform difficult to accept. What gets blurred in the interpretation and understanding of the need for reform is the conclusion by influential reports that the US health system is fundamentally broken: it frequently causes harm, quality problems abound, and sweeping changes are needed that will have a multitude of effects on the people who work within it, as well as on the people who are served by it. New models of training at the pre-professional level within the framework of an interdisciplinary collaboration will be the sustaining force of reform. For example, having inter-disciplinary training begin at the pre-professional level at the level of graduate professional training where the relevant disciplines have, as part of their core experience, being trained in the context of the Interdisciplinary Team, mirroring what is happening in PCMHs and ACOs. Even earlier, at the collegiate level, developing "pipelines" for trainees where they engage in cross-discipline clusters, simulating the interactions and processes necessary for there to be functional collaboration once these individuals enter the workforce (e.g., students who are pre-medical, pre-nursing, pre-social work, pre-psychology, etc., engaged in learning how to work in effective team structures). However, this will be very difficult to achieve. It is important that those individuals involved in training and those who establish the standards be capable of presenting the full picture of the rationale for reform and how training needs to respond to this rationale. The most effective arguments for changing practice and training are likely to be ones that are broadly embraced, which means that they have to be broadly understood. Improving quality in health care is a goal shared by all, but understanding why improvements are needed is more fundamental to change.

For societal relevance pre-professional training programs should do the following:

- Train in an interdisciplinary model in which trainees from the relevant behavioral health disciplines learn to work together, coordinate care, and

focus on outcomes and the application of evidence-based approaches; training in innovative concepts such as P4P and PR should be incorporated into training programs.

- Train in the integration of behavioral health services into primary care systems.
- Explore the development of a behavioral health facilitator or concierge with whom individuals and their families make first contact and whose role is to direct individuals to the proper care required.

The qualities and capabilities of the psychiatrist of the 21st century who could provide credible care within this type of paradigm are outlined in Table 58.2–6.

Health Care Reform and Psychiatry: Future Challenges and Opportunities

Achieving the systemic changes recommended by the IOM 2006 Report and put into play at least partially by the ACA, and ultimately fulfilling the goals of the Triple Aim, is an exceedingly complex challenge. Key questions for continued health care reform include the following:

- ***Social contract***. What social contract or premise underlies health care? Is health care a "right"—as economists would say, a *public good* worthy of public support?
- ***The social safety net.*** What should its scope be? Individuals with limited financial means and older adults are most at risk.
- ***Financing method***. What approach should be used to finance health care—a nationally administered system financed by taxes ("single payer")? Shoring up the existing employer-based system for workers and their families ("employer mandate")? Or requiring individuals to buy insurance with the support of tax incentives ("individual mandate")? On the provider side, CMS' Bundled Payment for Care Initiative (BPCI) is showing interesting results when applied to high-cost procedures.
- ***Empowering consumer choice***. Should *consumer choice drive market forces* to shape the operational features of a patient-centric health care system? How can consumers be supported to make good decisions for themselves and have their interests represented collectively in the face of the complexities of modern health care?
- ***Market structures***. How should market forces be structured so that they focus plans or organized delivery systems on *quality, safety, and efficacy* as defined by both consumer preferences and standards formally set by the medical/scientific community?
- ***Cost containment.*** What is a fair and effective way to finance basic health care for everyone at reasonable cost?
- ***Quality and Workforce/System Improvement.*** How should the rate of improvement in health care increase?

Table 58.2–6.
Profile of the Psychiatrist of the 21st Century

Knowledge Base	Skills
An excellent physician and diagnostician; has core knowledge of the range of psychopathology	Provides cost-effective, quality mental health care; works to preserve patient benefits; uses benefits stringently and only as required
Knowledgeable at a mechanistic level in sophisticated psychopharmacology; strong knowledge base in drug–drug interactions	Interfaces readily with primary and specialty care physicians and systems; can be a teacher to them and a caregiver to their patients; possesses strong consultation-liaison skills
Knowledgeable about brief therapies, care coordination	Works comfortably within a multidisciplinary team system; can assume team leadership role if indicated
Strong sense of individual and group dynamics and systems	Effectively deals with acute and chronic patients; uses shared decision-making in working with patients and families
Sense of economics and fiscal realities; understands the implications of health care reform, managed care, and utilization management; can work quickly and efficiently	Refers patients to the most appropriate level of care and matches treatment with patient; can use the continuum of care in treating patients
Can comprehend quality research and can understand the potential application of research to clinical work	Has good outcomes
Has a public health perspective and population-based focus; can work in the context of recovery and rehabilitation	Has demand management skills, especially for high-use patients; engages in prevention, health education, health literacy

Adapted from Meyer RE, McLaughlin CJ, eds. *Between Mind, Brain, and Managed Care: The Now and Future World of Academic Psychiatry*. Washington, DC: American Psychiatric Press; 1998 and Huey LY. Problems in Behavioral health Care: Leapfrogging the status quo. *Admin Policy Ment Health*. 2002;29:403.

As of 2014, of 230 quality measure established by the Agency for Health Care Research and Quality, about 160 have been adopted at a median rate of 2.3 percent/year with a median improvement rate in quality of care at only 1.6 percent/year: improvement is a slow process. How can health care workers improve their competency in what they do? Research shows, for example, that physicians are not very good at self-assessment of their competency and what they need to learn. How can systems improve more rapidly? In recent years, the concept of the *Rapid Learning Health Care System* has evolved, another body of reform work emanating from the Institute of Medicine and the Patient-Centered Outcomes Research Initiative, the latter spawned by the ACA, to speed up the process for practical research related to clinical improvement that can be introduced quickly to make a difference in health care function and outcomes.

At the same time, Emanuel predicts six megatrends that have been projected by 2025. These are:

- Health insurance companies will no longer function as they have historically
- Physicians and hospitals will focus on health maintenance for people having chronic illnesses, keeping them out of emergency rooms and avoiding hospitalization; routine screening for mental illnesses will be widespread with rapid intervention the norm
- Approximately 1,000 of the roughly 5,000 acute care hospitals will close; remote technologies for assessment and monitoring will be commonplace
- Employer-sponsored health insurance will become less common, projected to be available to fewer than 20 percent of workers
- Health care inflation will end
- Medical education will be transformed with 3-year medical schools and shorter residencies; half of training will be outside of hospitals; interdisciplinary team training will be the standard; population health and prevention will be significant components of medical education and training

The issues involved in reforming health care, including the ACA, and the transformational processes that they entail, are major public policy questions that require the expertise of a broad range of disciplines, from economics to ethics to basic public philosophy. It will take a decade or more for the reform process to play out to see what happens, but as mentioned earlier, there are encouraging initial signs. It is important to ensure that behavioral health providers and the medical specialty of psychiatry provide a clear vision, are a meaningful part of the dialogue, and are prepared to play a constructive role in the process of transformation brought about by the ACA and the necessity of reform.

References

Affordable Care Act Overview. *Am Public Health Assoc*. 2012. Available at http://www.apha.org/advocacy/Health+Reform/ACAbasics/.

Barry C, Huskamp H. Moving beyond parity: mental health, addiction care under the ACA. *NEJM*. 2011;365:973.

Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. *Health Affairs*. 2008;27(3):759.

*Berwick D, Hackbarth A. Eliminating waste in US health care. *JAMA*. 2012;307(14):1513.

Bishop T, Press M, Keyhani S, et al. Acceptance of insurance by psychiatrists and the implications for access to mental health care. *JAMA Psychiatry*. December 11, 2013.

Cermiel J, Ya-Fen C, Chwastiak L, Avery M, Katon W, Unutzer J. Bipolar disorder in primary care: clinical characteristics of 740 primary care patients with bipolar disorder. *Psychiatri Serv*. 2014;65(8):1041–1046.

Coleman K, Austin BT, Brach C, Wagner E. Evidence on the chronic care model. *Health Affairs*.

2009;28(1):75.

*Davis K, Stremikis K, Squires D, Schoen C. *Mirror, Mirror On The Wall: How The Performance Of The Us Health Care System Compares Internationally*. The Commonwealth Fund: 2014.

*Emanuel E. *Reinventing American Health Care: How the Affordable Care Act will Improve Our Terribly Complex, Blatantly Unjust, Outrageously Expensive, Grossly Inefficient, Error Prone System*. New York: Public Affairs; 2014.

*Frank R, Glied SA. *Better But Not Well: Mental Health Policy in the United States since 1950*. Baltimore, MD: Johns Hopkins University Press; 2006.

Friedberg M, Schneider E, Rosenthal M, Volpp KG, Werner RM. Association between participation in a multipayer medical home intervention and changes in quality, utilization, and costs of care. *JAMA*. 2014;311(8).

Fuchs V. Health reform: Getting the essentials right. *Health Aff*. 2009;28:180.

*Health Affairs. Special 25th Anniversary Issue. *Health Aff*. 2007;26:1523.

*Hoge MA, Morris JA, Daniels AS, Stuart GW, Huey LY. *An Action Plan on Behavioral Health Work Force Development*. Cincinnati, OH: Annapolis Coalition on Behavioral Health; 2007. Available at http://www.anapoliscoalition.org or http://www.samhsa.gov/matrix2/matrix_workforce.aspx.

Institute for Healthcare Improvement Triple Aim, 2012. Available at http://www.ihi.org/Engage/Initiatives/Triple-Aim/Pages/default.aspx.

*Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academy Press; 2001.

*Institute of Medicine Committee on Crossing the Quality Chasm. *Adaptation to Mental Health and Addictive Disorders, Improving the Quality of Health Care for Mental and Substance-Use Conditions: Quality Chasm Series*. Washington, DC: National Academies Press; 2006.

*Institute of Medicine. *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. Washington, DC: National Academies Press; 2013.

Institute of Medicine. *Graduate Medical Education that Meets the Nation's Health Needs*. Washington, DC: National Academies Press; 2014.

Katon W J, Unutzer J. Health reform and the Affordable Care Act: the importance of mental health treatment to achieving the triple aim. *J Psychosomatic Res*. 2013;74(6):533.

Kuehn B, Katon W. Health reform, research pave the way for collaborative care of mental illness. *JAMA*. 2013;309:2425.

Manderscheid R, Kathol R. Fostering sustainable, integrated medical and behavioral health services in Medicaid settings. *Ann Int Med*. 2014;160:61.

*Mechanic D. *Mental Health and Social Policy: Beyond Managed Care*. 5th ed. Boston: Allyn & Bacon; 2007.

Methodology Committee of the Patient-Centered Outcomes Research Institute (PCORI). Methodological standards and patient-centeredness in comparative effectiveness research: the PCORI perspective. *JAMA*. 2012;307(15):1636–1640.

Meyer RE, McLaughlin CJ, eds. *Between Mind, Brain, and Managed Care: The Now and Future World of Academic Psychiatry*. Washington, DC: American Psychiatric Press; 1998.

*Murray CJ, Lopez AD. *The Global Burden of Disease*. Cambridge, MA: Harvard University Press; 1996.

New Freedom Commission on Mental Health. *Achieving the Promise: Transforming Mental Health Care in America, Final Report (DHHS Pub. No. SMA-03-3832)*. Rockville, MD: U.S. Department of Health and Human Services; 2003.

Primary Care Collaborative. Defining the medical home. 2014. Available at http://www.pcpcc.net/node/14.

Raney L, Kathol R, Summergrad P. Collaborative care models for comorbid medical and behavioral health conditions. *Focus*. 2013;11(4):501–508.

Selby JV, Beal AC, Frank L. The Patient-Centered Outcomes Research Institute (PCORI) national priorities for research and initial research agenda. *JAMA*. 2012;307(15):1583.

Shih A, Davis K, Schoenbaum, McCarthy D, et al. *Organizing the Us Health Care Delivery System for High Performance*. The Commonwealth Fund; August 2008.

Soeteman DI, Perry CJ, Kim JJ. The use of comparative effectiveness research in mental healthcare policy and practice. *J Comp Eff Res*. 2014;3(1):1.

*Wennberg JE, Cooper MM. The quality of medical care in the United States: A report on the Medicare program. In: *The Dartmouth Atlas of Health Care 1999*. Chicago: American Hospital Association Press; 1999.

▲ 58.3 The Role of the Hospital in the Care of Persons with Mental Illness

JEFFREY L. GELLER, M.D.

INTRODUCTION

Throughout the 19th century and for the first half of the 20th century, public psychiatric hospitals (and some private ones as well) provided all the roles necessary to take care of psychiatric patients. The state hospital was the source for inpatient, outpatient, case management and outreach, and emergency services. If someone needed admission, a family member could simply transport him to the front door. With the shift from inpatient to community-based services that began in earnest in the mid-20th century, admission to, discharge from and services within the psychiatric hospital became both more complex and more convoluted. Layer upon layer of agencies, services, and regulations facilitated and impeded the psychiatric hospital in its mission (see Fig. 58.3–1).

The reasons for the evolution of this relationship between the psychiatric hospital and its prospective and actual patients are multifactorial. As Figure 58.3–1 illustrates, there are entities with overlapping but competing agendas at every turn. Fundamentally, there are two predominant factors that account for the contemporary, tortuous, labyrinthine interface between the psychiatric hospital and the community. First, the psychiatric hospital, in a manner quite distinct from the general hospital, plays two roles: clinical services and public safety, thus, the extensive interface between the psychiatric hospital and the criminal justice system (see Fig. 58.3–1). And second, psychiatry, more than any other field of medicine, has been inundated by gyrations of conflicting actions to decrease the cost of care and treatment, thus the barriers and alternatives to psychiatric hospital admission (see Fig. 58.3–1).

The first one-and-one-half decades of the 21st century have not been kind to psychiatric hospitals. Virtually every state in the United States is experiencing a shortage of psychiatric beds. These shortages include acute care beds, intermediate and long-term care beds, and forensic beds. The effects of these bed shortages are increased waiting lists for state hospital beds, increased waiting lists for other psychiatric beds like general hospitals and Veterans Administration (VA) psychiatric beds, overcrowding in state hospitals, backups and boarding in general hospital emergency departments (EDs), and premature discharges from acute care settings. These shortages have also meant a shift of persons with serious mental illness (SMI) behind locked doors from public hospitals to jails and prisons. These psychiatric bed closures come from all types of hospitals. Sixty-four

percent of states have experienced declines in the number of general hospital specialty unit psychiatric beds, and 19 states have experienced a decline in the number of private psychiatric hospital beds. Nationwide, the number of state and county hospital beds declined by 86 percent (369,999 to 52,612) between 1969 and 2002; the number of VA medical center psychiatric beds declined by 84 percent (51,696 to 8,386) within the same time period. The inaccessibility of psychiatric beds is a regular feature in news stories and editorials, often with a link between that and a local or national tragedy.

Lengths of stay have plummeted. Between 1990 and 2010, the length of stay of hospitalizations for SMI in acute care settings decreased by 50 percent (from 14 days to 7 days). Private insurers represent a shrinkage percentage of payments for hospital services, especially for substance abuse services. About 50 percent of all payments for psychiatric inpatient services were from government programs; state governments' funding increased by 65 percent just between 2000 and 2001.

Hospital-based outpatient services grew by 78 percent (242,301 episodes to 1,104,300) between 1969 and 2002. In 2002, however, hospital-based outpatient services represented only 31 percent of all outpatient episodes of care for psychiatric disorders.

The shift of resources from hospital to nonhospital services is reflected by the changing distribution of expenditures by state mental health authorities (SMHAs). Inpatient services received 63 percent of the SMHA budget in 1981, 50 percent in 1993, 28 percent in 2004, and 23 percent in 2012. Although the overall inpatient percentage decreased, the proportion going to forensic patients increased to 33 percent of all state hospital expenditures.

While there has been a shift in public expenditures from inpatient to outpatient services, there has been an overall decrease in public expenditures for psychiatric treatment. In fiscal years 2010 and 2011, the biggest cuts were directed at state hospitals with 20 and 15 states, respectively, cutting state hospital funding. Less funding meant fewer individuals could be serviced in state hospitals with 31 states decreasing the persons they admitted comparing 2007 to 2009.

At the same time, the demand for state hospital utilization increased in 18 percent of the states, in part because states were also making cuts in all manner of community-based services. Between fiscal years 2009 and 2011, 17 states had cuts in their overall state mental health expenditures by greater than 1 percent, with 5 of those states exceeding 20 percent (one state approached 50 percent).

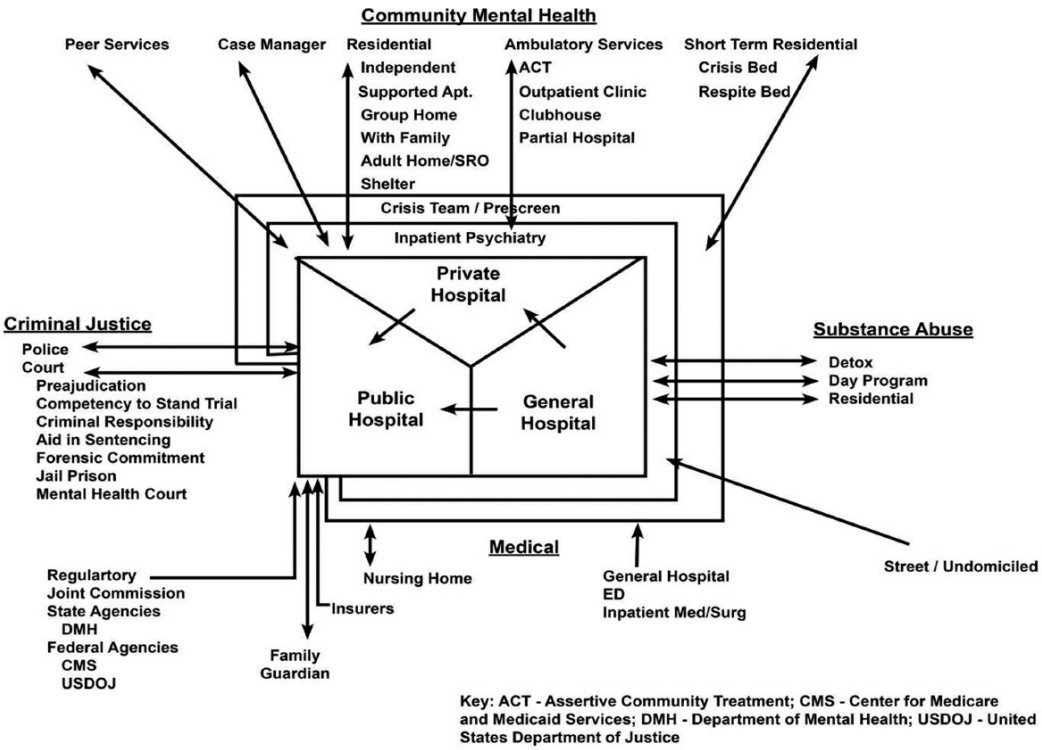

FIGURE 58.3–1. The psychiatric hospital in the context of the community.

Marshalling scarce resources has led to synergies among various sectors within a state. In just more than 75 percent of the states, community-based programs (often at general hospitals) perform a gate-keeping function to state hospital entry. Some states (38 percent) require general hospital admission prior to a transfer to a state hospital. And in all but five states there are assisted outpatient commitment (AOT) (involuntary outpatient commitment) laws on the books with an expectation these can be used to decrease both inpatient and overall mental health services costs.

The picture does not look optimistic, at least for the immediate future. About half of the states are reorganizing their state hospitals, including downsizing, reconfiguring, closing, and/or consolidating. Despite this, 11 states still have 6 or more state hospitals each. And a surprising number of states are building new state hospital buildings, but these generally replace one or more existing state hospitals and usually do not reflect an increase in beds (see Table 58.3–1). As more patients are admitted under forensic commitments, lengths of stay increase, and so does the unavailability of beds to others. In 2006, about 40 percent of patients found not guilty by reason of insanity had lengths of stay of 1 to 5 years and about 38 percent had lengths of stay of greater than 5 years.

While the place of psychiatric hospitals as a tool in the resources box of treatment for mental illness is shrinking, the need is growing. On any given day in fiscal year 2010, the number of noninstitutionalized individuals with an episode of serious psychological distress in the past was 3 percent of the population. In 2012, 4.1 percent of the adult U.S. population had an

episode of SMI in the past year. The 12-month prevalence for schizophrenia in adults was 1.15 percent, while it was 3.5 percent for post-traumatic stress disorder (PTSD) (one-third of these classified as severe), 9.1 percent for any personality disorder, and 18.1 percent for anxiety disorder (23 percent of these classified as severe). The mismatch between need and services is no better represented than by the fact that in 2008 over 40 percent of adults in the United States with an SMI did not receive any treatment for a mental health problem. Of those that did, 7.5 percent had an inpatient and 40.5 percent had an outpatient episode of care.

The state of hospital psychiatry raises questions about how it came to play the roles it does. What were the historical influences that brought the psychiatric hospital to where it is, and what keeps it there? How does the psychiatric hospital (this term will include general hospital psychiatry units) fulfill its current role? And what should be the roles of psychiatric hospitals in the future?

THE HISTORICAL EVOLUTION OF THE ROLE OF THE HOSPITAL

Prior to the 19th century, the care of insane persons was custodial in nature, took place largely in almshouses, and was predominantly a local endeavor. State care, at the beginning, meant an institution established, built, and managed by the state but whose operation was not funded by the state. Payers were, in descending order of priority, the patient, the family, and the county. The earliest public facilities—in Williamsburg, Virginia, and Columbia, South Carolina—functioned in this manner. The state taking responsibility for all insane persons and performing its responsibility through state-managed and funded hospitals began with Worcester State Hospital in 1833.

Table 58.3–1.
A Sample of Public Psychiatric Hospitals Built since January 2000

State	Hospital	Opening Date for Patients	No. of Hospital Beds
AK	Alaska Psychiatric Institute	July 2005	80
AR	Arkansas State Hospital	July 2008	210
AZ	Arizona State Hospital-Adolescents	July 2002	16
AZ	Arizona State Hospital-Adults	February 2003	200
CA	California Health Care Facility-Stockton	July 2013	1,772
CA	Coalinga State Hospital (Forensic)	September 2005	1,500
CO	Robert L. Hawkins High Security Forensic	June 2009	200
DC	St. Elizabeth's Hospital	April 2010	292
GA	Central State Hospital—Payton B. Cook Forensic	January 2009	184
IN	Evansville State Hospital	September 2003	168

IN	Logansport—Isaac Ray Treatment Center	August 2005	88
KY	Eastern State Hospital	September 2013	239
MA	Worcester Recovery Center and Hospital	October 2012	290
ME	Riverview Psychiatric Center	2004	92
MI	Center for Forensic Services	2005	210
MO	Center for Behavioral Medicine	2004	50
MO	Fulton State Mental Hospital	2018	300
MS	South Mississippi State Hospital	2000	50
MT	Montana State Hospital	2000	174
NC	Broughton	2016	382
NC	Central Regional Hospital	July 2008	398
NC	Cherry Hospital	2015	314
NJ	Greystone Park Psychiatric Hospital	July 2008	450
NV	Rawson-Neal Psychiatric Hospital	August 2006	190
OH	Northcoast Behavioral Healthcare Hospital	2016	258
OH	Summit Behavioral Healthcare Hospital	2006	291
OR	Oregon State Hospital—Salem	January 2011	620
OR	Oregon State Hospital—Junction City	March 2015	174
SC	William S. Hall Psychiatric Institute	2015	515
TN	Memphis Mental Health Institute	October 2007	75
TN	Western Mental Health Institute	March 2010	162
TX	Rusk State Hospital: Maximum Security Unit	June 2012	60
VA	Eastern State Hospital Geriatric	April 2008	150
VA	Eastern State Hospital Adult	September 2008	150
VA	Western State Hospital	2013	250

Before there were any more than a few public hospitals, there was already a well-established group of private psychiatric hospitals. Three of these began as parts of general medical hospitals: The Institute of the Pennsylvania Hospital, Bloomingdale Hospital (New York), and McLean Hospital (Massachusetts). Three more were freestanding institutions: Butler Hospital (Rhode Island); Brattleboro Retreat, which began as the Vermont Asylum for the Insane; and the Hartford Retreat, which subsequently was renamed the Institute of Living (Connecticut). Two of the hospitals began under the aegis of religious organizations: The Friends Asylum (Pennsylvania) was initiated by a Quarterly Meeting of the Philadelphia Society of Friends and the Mount Hope Institution (Maryland) was a Catholic hospital that began as the Mount Saint Vincent Hospital.

Of these larger private facilities, all but one preceded the state's first public hospital, sometimes by more than 50 years. The private hospitals helped the state to meet its obligations to the insane poor by admitting "public" patients. Payments were made by the state to the private facility,

and some treatment was given without charge. Thus, the states' purchase of beds from the private sector was not a novel concept of the late 20th century but is actually how U.S. hospital psychiatry began.

Many public and some private hospitals throughout the country were built under a design put forth by Thomas Kirkbride, a founding member of the American Psychiatric Association (APA). In his book, *On the Construction, Organization, and General Arrangements of Hospitals for the Insane*, Kirkbride prescribed every aspect of hospital construction and operation. His prototype hospital, variations of which can still be seen across the United States (but are fast being torn down), is shown in Figures 58.3–2 and 58.3–3. Noteworthy among his recommendations is that a 250-bed hospital would have 16 wards, with about 15 patients per ward.

By the 1880s, most psychiatrists agreed that "insane patients can best be managed away from their own homes." Parsons, an alienist of the era, opined, "well-equipped and well-managed asylums for the insane are among the most important of the outgrowths of modern civilization and benevolence." By the end of the 19th century, it was clear that private facilities were for the wealthier, more genteel class of citizens, who were able to pay fees and were provided remarkably comfortable accommodations. Public hospitals were entities of state, county, or municipal governments that served not only the poor but also the middle class, given that the latter could not afford private hospitals.

A major difference between the public and private facilities is that public institutions sought to lower their censuses as state hospital populations grew and patients had progressively longer lengths of stay, whereas private facilities sought patients to maximize their occupancy rates. A fundamental criticism of private facilities was pointed out by Parsons: "It is not for the interest of the proprietor of a private asylum to cure his patients, and hence they may not make an honest effort to promote their cure." Much of what many consider to be modern developments in hospital psychiatry were well under way before the end of the 19th century: Oversight of private hospitals by state boards, advertising by private hospitals, involuntary admissions to private hospitals, more perks for higher fees, and the "dumping" of patients from the private to the public sector.

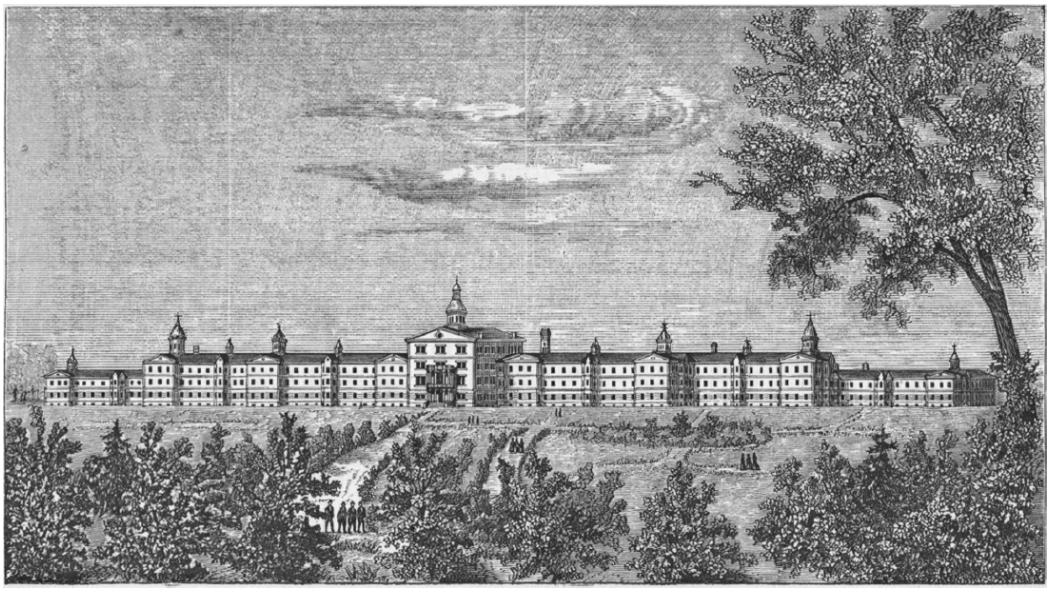

FIGURE 58.3–2. Exterior of hospital prototype designed by Thomas Kirkbride. (From Kirkbride TS. On the Construction, Organization, and General Arrangements of Hospital for the Insane. Philadelphia, PA: J.B. Lippincott; 1880.)

Work and recreation were both seen as important to the mental health of patients in the 19th century. Generally, the public hospitals did better with the former, and the private ones with the latter. Recovery, employment, and wellness are not novel ideas of contemporary hospital psychiatry. Each was endorsed more than 150 years ago in both public and private hospitals.

Nor is evidence-based psychiatry a new thought. Andrews articulated in 1887 that "the unparalleled progress in neurology, cerebral anatomy, physiology, pathology and localization of function has enlarged the horizon of our knowledge of disease and of the action of causes." This progress, he said, "furnished a scientific and positive basis for treatment in many cases of insanity which before was unattainable." Inpatient treatments near the end of the 19th century included such modalities as electricity, massage, oophorectomy, enforced rest, hydrotherapy, and physical training.

Changes in hospital psychiatry over the first half of the 20th century were dramatic. Psychoanalysis was introduced to the inpatient setting in the 1920s and somatic treatments such as insulin coma, Metrazol shock, and electroconvulsive therapy in the 1930s. In the 1940s, attention was focused on developing milieu therapy, whereas in the 1950s, the introduction of the first antipsychotic medication, chlorpromazine, fundamentally altered inpatient psychiatry.

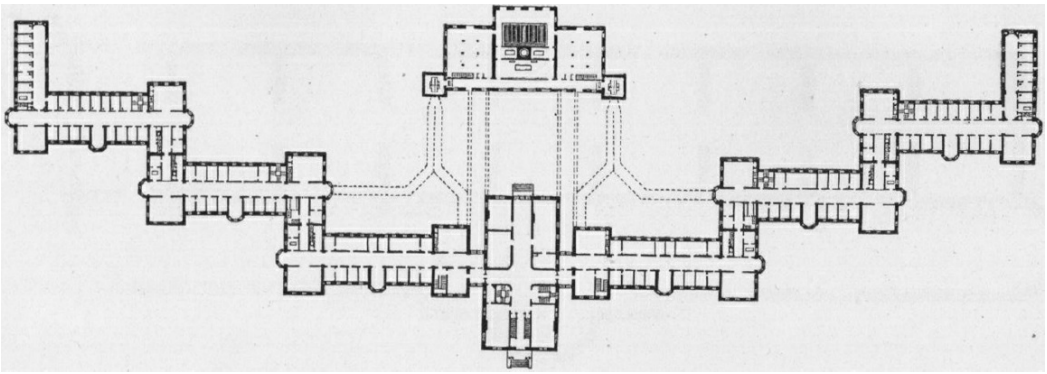

FIGURE 58.3–3. Interior of hospital prototype designed by Thomas Kirkbride. (From Kirkbride TS. *On the Construction, Organization, and General Arrangements of Hospital for the Insane*. Philadelphia, PA: J.B. Lippincott; 1880.)

The mid-1950s began the shift from hospital-based to nonhospital-based care and treatment of persons with mental illness, a movement retrospectively labeled "deinstitutionalization" but perhaps better termed "dehospitalization." General hospital psychiatry and day hospitals became major loci of care and treatment, and the role of the psychopathic hospital expanded. Units in general hospitals were not always welcomed by hospital medical staffs, who were concerned that psychiatric patients, whom the physician thought would overwhelmingly be noisy, destructive, violent, or suicidal, would disrupt the general hospital.

The darkest era in U.S. hospital psychiatry was in the 1940s, a time when care of persons with mental illness was described as a national disgrace. Attempts were made to improve inpatient treatment through the promulgation of principles and standards, such as those put forward by the Division of Mental Hygiene of the U.S. Public Health Service, the National Committee for Mental Hygiene, and the APA's Committee on Psychiatric Standards and Policies. These probably had less effect than exposés such as Mary Jane Ward's *The Snake Pit*, a fictionalized first-person account of life in a state hospital.

Whereas the public sector was overburdened and overrun, by 1950 the private sector was also facing financial difficulties. The Medical Director of McLean Hospital stated at the 1950 APA Institute, "The time when you could run a private hospital as a hospital and make money has gone. Operating expenses have so increased that it is almost impossible to collect the cost from our patients."

In the 1960s, as health insurance for hospitalization expanded to include psychiatric treatment, psychiatric hospital treatment outside of state hospitals became more available. Misuse of insurance for hospitalization for mental illness was quickly recognized. To remedy the manipulation of hospital benefits, interventions were introduced: Peer review, reimbursement for partial hospitalization, and reimbursement for other alternatives to inpatient treatment. "Dumping" and "cost shifting"

emerged as part of the everyday life of inpatient psychiatry.

In the 1970s, states were contemplating phasing out state hospitals. General hospital psychiatric units were becoming more of a presence. The expanding role of general hospital psychiatry drew caution from some who warned workers in general hospitals to be watchful of efforts to convert general hospital psychiatric units into mini state hospitals. Declining numbers of individuals willing or able to pay for private hospital treatment and more available third-party coverage led private hospitals to admit patients of different socioeconomic classes and ethnic backgrounds and to keep all patients for shorter lengths of stay.

In the 1980s, there was a remarkable growth in private sector hospital beds. Private corporations were rushing to build psychiatric hospitals as less regulation, lower overhead, and greater latitude in insurance reimbursement meant there were huge profits to be made in the "psychiatric hospital industry." For-profit psychiatry grew, but with it came a new set of problems. For-profit hospitals were the least likely to treat the most functionally disabled; would treat few if any patients who could not pay for treatment; usually had poorer staff-to-patient ratios than nonprofit hospitals; failed to provide education and training; failed to contribute adequately to research; and overall provided a lower quality of care. These hospitals tended to be totally disengaged from the communities in which they were located; they paid higher salaries and thus hired away graduating residents from academic and public sector positions.

The heyday of for-profit hospitals (most private psychiatric hospitals) did not last long. Throughout the 1990s, the total and average number of beds per private hospital decreased, as did the average length of stay. Between 1992 and 1998, private psychiatric hospital beds declined by 23 percent. For-profit companies had built "chains" of hospitals so rapidly that there quickly came to be too many private psychiatric beds, and hospital chains began experiencing major financial losses. Lawsuits against the chains for unscrupulous practices such as recruitment of patients and manipulative billing practices were settled in the hundreds of millions of dollars.

The 1990s also witnessed the privatization of public mental health services. Medicaid, Medicare, and state funds were directed to the private sector to greater degrees during the 1980s and 1990s. States assisted freestanding private hospitals in obtaining waivers of the Institution for Mental Disorders (IMD) rule (see later discussion). A greater reliance on the private system for those formerly treated in the public system created many unintended consequences that varied state by state, including inpatient treatment further from one's community, less continuity of inpatient treatment, new patterns of recidivism, and certain patients being blacklisted at some hospitals.

In 2001, 37 percent of general hospitals had psychiatric inpatient care.

This percentage is higher than those with psychiatric emergency, consultation-liaison (C-L), or even outpatient services. Hospital psychiatry in the United States continued to be a *business*. However, in the business of psychiatry, nonprofit psychiatric care providers compared favorably with for-profit providers.

It is interesting to note how the practice of hospital psychiatry appears to have shifted in response to shorter lengths of stay. A study at McLean Hospital in 2002 that compared inpatient antipsychotic medication use in 1989, 1993, and 1998 showed the chlorpromazine-equivalent final total dose of all antipsychotic medications used at discharge in 1998 was 29.3 percent greater than in 1993 and 46.1 percent greater than in 1989. The average time from admission to first antipsychotic dose decreased dramatically from 6.8 days in 1989 to 1.5 days in 1993 and .9 day in 1998. Two or more antipsychotic medications were used during 20 percent of inpatient days in 1998, compared to 5.7 percent days in 1993 and 1.7 percent in 1989. Pressures to discharge patients reduced not only the length of treatment but also the content of treatment.

In the first decade of the 2000s, the prospects for the private psychiatric hospital and for general hospital psychiatric units did not look good. Although overall health care spending increased by 15.7 percent between 1992 and 1999, mental health and substance abuse spending decreased by 17.4 percent. Behavioral health care spending went from 7.2 percent of total private health insurance spending to 5.1 percent over this same time period. Reimbursement rates fell such that they often did not even meet the total costs of care and treatment. While reimbursement rates fell, costs to deliver treatment increased.

Private hospitals were offering fewer services. Those hospitals offering partial programs decreased from 82.5 to 66.7 percent, and those offering outpatient services dropped from 64.2 to 49.2 percent between 2000 and 2001. At the same time, oversight increased. The Office of the Inspector General (OIG) of the Department of Health and Human Services, the Joint Commission on Accreditation of Healthcare Organizations, and the Center for Medicare and Medicaid Services all stepped up investigations of hospitals. Medicaid's progressive decrease in disproportionate share payments (DSHs) to public mental hospitals is taxing some states' ability to operate their public system at its pre-2000 scale.

The landscape of U.S. inpatient psychiatry became so rough that well-established private hospitals closed or moved. Those in the former category include Taylor Manor and Chestnut Lodge in Maryland. In the latter category, the best known is the Menninger Clinic, formerly of Topeka, Kansas, and now in Houston, Texas. In the public sector, state hospitals, well established since the 19th century, began reinventing themselves, usually by consolidation of multiple public hospital and by new construction such as Worcester State Hospital (Massachusetts), Oregon

State Hospital, and Western State Hospital (Virginia). Furthermore, the need to serve and contain persons with mental illness and histories of sexually problematic behaviors is diverting resources from the treatment of psychiatric disorders. California built a 1,500-bed facility for forensically committed individuals—mostly sexually violent predators—representing the first new California facility in more than 50 years. Even elderly individuals with such histories can often not be discharged for the public psychiatric hospital as federal law prohibits nursing homes form admitting persons who are classified in the highest level of sex offenders.

As components of psychiatric services in the second decade of the 21st century, psychiatry hospitals are shrinking, growing, being torn down, and being newly built all at the same time. As hospitals are unable to meet the needs of the population requiring inpatient level of care, local police have become the triage agents for persons with serious mental in the streets; jails and prisons are unofficial replacement facilities for state hospitals; and mental health courts provide mores case management than anything else. One potential positive note is the growing number of new public psychiatric hospitals built with a hope that the building's design will facilitate patients' recovery (see Table 58.3–1). It is well worth remembering, however, the observation made in the wake of newly constructed community mental health centers (CMHCs) with inpatient units in the 1970s: the latest in "psychoarchitecture" and design elements do not result in behavioral remediation since clinical improvement does not result simply from changing bricks, mortar, and furniture. That being said, newly built hospitals, most of which have had significant problems with translating programmatic and life safety design concepts into actual hospital structures, have brought more pleasant surroundings, in an overall safer environment, with more available space for in- and outdoor programming (see Figs. 58.3–4 and 58.3–5). These figures are presented as an example and not as a model.

Available psychiatric beds in general and private hospitals appear to be made inaccessible to patients by insurance providers because limiting admission to hospitals with which they have negotiated rates is one means to control cost. As reimbursement rates drop below the actual expenditures of the general hospital, these hospitals limit the third-party payers with which they will contract or further reduce or eliminate psychiatric beds, exacerbating the bed shortage.

Public psychiatric beds are highly disproportionately occupied by forensically committed patients. This makes beds for civil admissions less available and patients are stuck in general or private psychiatric hospital beds, which in turn delays the movement of psychiatric patients needing admission out of the ED. Waiting time in an ED can be up to 1 month for a psychiatric bed, or the patient is prematurely discharged from the ED, sometimes with lethal consequences.

Despite ascriptions of causality for the current state of hospital psychiatry to malevolence, malfeasance, and malversation, it would seem more accurate to see the role of the psychiatric hospital in the second decade of the 21st century as still the result of a basic tension between clinical benefit, social control functions, and parsing out of costs. Each governmental level and each government agency wants another one to fund the treatment of mental illnesses.

Despite all of these uncertainties, there is no indication that psychiatric hospitals will cease to play some role in the future. The remainder of this section focuses first on defining the role of the hospital and then on implementing that role.

DEFINING THE ROLE OF THE PSYCHIATRIC HOSPITAL

Clinical Rationale

The clinical rationale for the role of the psychiatric hospital was well established by the mid-18th century. Edward Jarvis, a nonpsychiatrist physician well enmeshed in psychiatric care of his era, stated it as well as anyone. First, Jarvis discussed his thoughts on "diseases of the mind":

> Maladies of the mind, like those of the body, require that the causes that produced them shall cease to act upon them. Unfortunately the diseases of the mind frequently find their origin in the circumstances and associations of home, in the cares and anxieties of business, in the relations of neighborhood, in the affairs of the town, in the movements of religious, political or other associations, in the habits of indulgences which are practiced. Or when the diseases are once established, they may be kept up by some of these. Moreover, the natural perceptions and moral sensibilities being disordered, the scenes, circumstances, and the persons, which are usually agreeable and favorable, in health, become objects of aversion and sources of irritation and keep the disorder alive.

Then Jarvis provided his thinking on why hospitals were needed to treat persons with these diseases:

> As the lunatic cannot be separated from these, while in his own house, he must go from them and generally among strangers. It is needful that the mentally disordered not only go away from home, but be placed where all the circumstances and associations are different from those to which he has been recently accustomed, and where all the influences that may reach him, can be controlled and modified, so that none but such as are favorable may affect him. They will be found in the hospital [where] the officers and the attendants are selected on account of their peculiar fitness for their work. They have the high intelligence, the scientific knowledge, and the moral endowments of gentleness, and firmness, and sympathy that are required, and they know how to occupy the diseased mind and lead it from its morbid fancies, and fix it upon cheerful and satisfactory subjects.

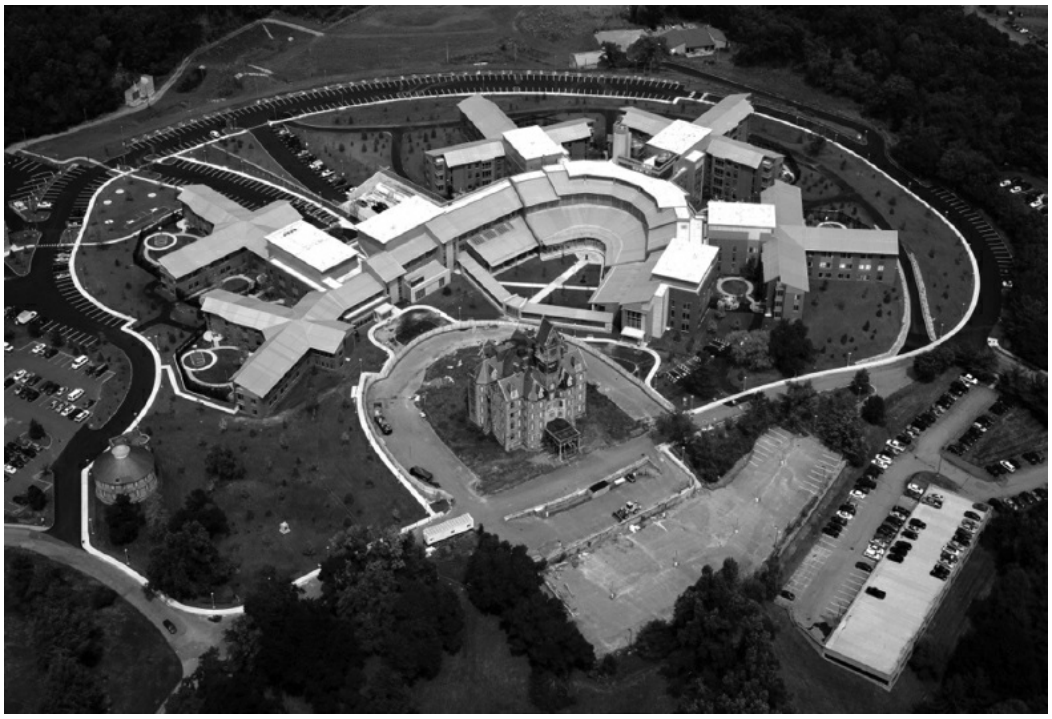

FIGURE 58.3–4. Worcester Recovery Center and Hospital (WRCH). Exterior view.

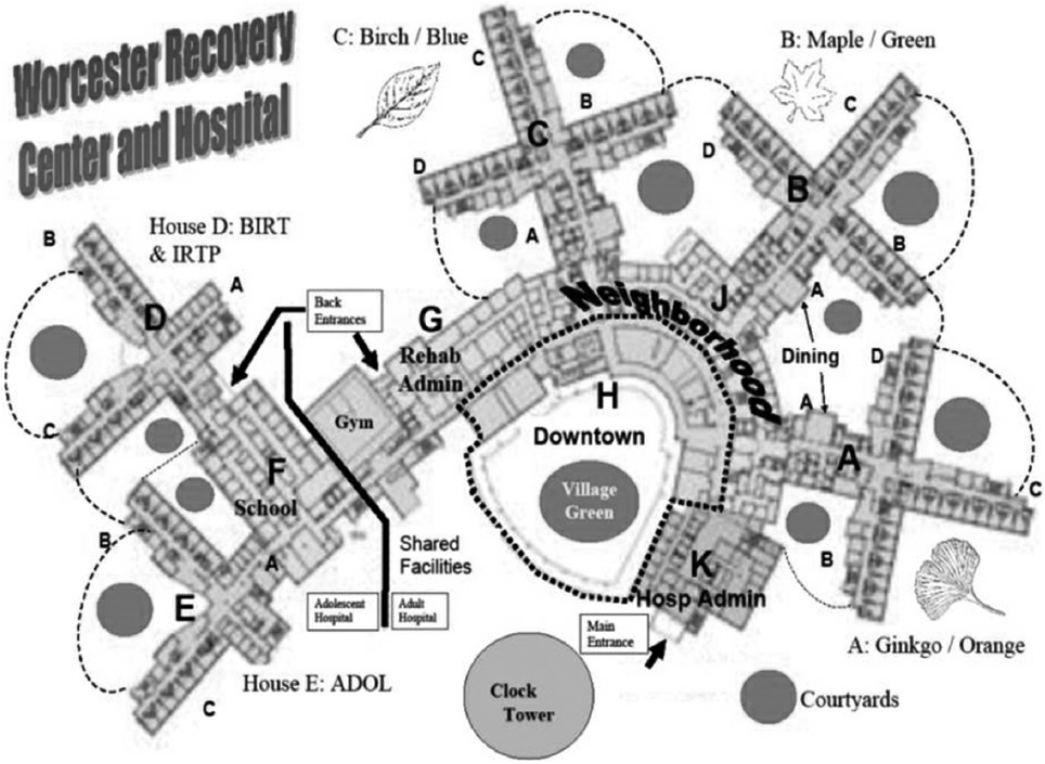

FIGURE 58.3–5. Worcester Recovery Center and Hospital (WRCH). Interior view.

Jarvis provided a quite optimistic prognosis if his recommendations were followed, but one quite consistent with those of his professional colleagues: "If insane persons are allowed to enjoy the means of healing in the early stages of their disorder, within one year after it appears, about 75

to 90 percent, can be restored to health." Furthermore, he argued for prompt hospitalization with the onset of symptoms:

> But in this class of maladies, time makes rapid havoc, and diminishes the changes of cure. If the means of healing are not tried within the first year, the ratio of cures is reduced perhaps a half in the second and still more in the third year, and when five years of neglect shall have elapsed, the hope of restoration is reduced to a mere accident, which no human skill can promise to accomplish.

By the mid-19th century, it was also clear to those who treated persons with insanity that the hospital had much to offer those in need of acute treatment due to advances in the field. Samuel Woodward, the first president of the APA, noted the following in 1846:

> The abandonment of depletion, external irritants, drastic purges, and starvation, and the substitution of baths, narcotics, tonics, and generous diet, is not less to be appreciated in the improved condition of the insane, than the change from manacles, chains, by-locks and confining chairs, to the present system of kindness, confidence, social intercourse, labor, religious teaching, and freedom from restraint. In this age of improvement, no class of mankind have felt its influence more favorably than the insane.

Hospital treatment continued to improve in the mid-19th century. Charles Nichols, superintendent of the Government Hospital for the Insane (St. Elizabeth's), commented as follows in the 1870s: "Much less mechanical restraint, and much less seclusion are now used in our institutions for the insane than were formerly resorted to in them. . . . I trust that we shall be able to still further reduce the proportion of cases in which restraint will be deemed advisable."

The role the physician could play in the hospital was also clear by the mid-19th century. Luther Bell, superintendent of the McLean Asylum, indicated in 1849:

> Domiciliated with our patients, ready at all hours, by day or night, to watch the effect of our prescriptions, sure of their being fairly and fully carried into effect, without hazard of the interference of timid friends or unfaithful nurses; obliged to make no disguise, no prescription for form or appearance sake, it is in our power to follow out the best convictions of our judgment, uninfluenced by aught extraneous to the object desired.

That the psychiatric hospital played a far greater role than just treating acutely ill insane persons and restoring them to health was quite clear by that time. William Awl, superintendent of the Ohio Lunatic Asylum in Columbus, wrote in 1849:

> More effort and skill is often demanded, to better the condition and make some improvement in the shattered intellects of those who are deemed hopeless, than to effect recoveries. . . . We must never permit ourselves to lose sight of humanity, but always remember the importance of the institution to all classes, and the advantages which may be secured to those whom we are obliged to regard as incurable.

Thomas Kirkbride expanded on this note in 1864, indicating the importance of the safety function played by the hospital:

Owing to various causes, but especially to a neglect of treatment in the early stages of the disease, there are a large number of the insane, in every community, whose chances of restoration are slight, but whose comfort and happiness are promoted, and their protection from ill treatment secured by a residence in these institutions, where they are under circumstances which give them the best chance for permanent improvement. To provide for this numerous class is another important object, attained by the establishment of such hospitals. Next to these come the protection of the community and the relief of private families.

These roles quickly became compromised by the services being overwhelmed by the number of patients needing to be served. By 1890, W. W. Godding, the APA's seventeenth president, was observing:

The overcrowding goes on, cots are brought out at night and laid down on the corridor floors, at first one or two in nooks and alcoves that seem quite designed for this sort of thing, but the business grows, the special adaptation of recesses and alcoves becomes less apparent as the line of beds side by side stretches in lengthening vista down the hall. And still the floors fill up until one, two, three hundred are thus nightly not accommodated but provided for. From ground floor to attic your hospital, spread with its white beds seems a very bivouac of lunatics. . . . Day by day, year after year, I have seen the individualized treatment of special cases swamped by the rising tide of indiscriminate lunacy pouring through the wards, filling every crevice, rising higher and higher until gradually most distinctions and landmarks have been blotted out.

These trends persisted, such that by 1954 Kenneth Appel, the eighty-second president of the APA echoed Godding:

There is overcrowding 20–30% in 70% of the hospitals, and 50% overcrowding in 8% of the hospitals. Hospitals are 40% understaffed in physicians; 15,000 more psychiatrists are needed; 37,000 patients in 29 hospitals have no psychiatrists, and 2,000 patients are in 5 hospitals which do not employ a full-time physician. In 4 hospitals with 6,000 patients, there are no graduate nurses.

Four years later, Harry Solomon, director of the Boston Psychopathic Hospital, proclaimed:

The large mental hospital is antiquated, outmoded, and rapidly becoming obsolete. We can still build them but we cannot staff them; and therefore we cannot make true hospitals of them. After 114 years of effort, in this year 1958, rarely has a state hospital an adequate staff as measured against the minimum standards set by our Association . . . I do not see how any reasonably objective view of our mental hospitals today can fail to conclude that they are bankrupt beyond remedy. I believe therefore that our large mental hospitals should be liquidated as rapidly as can be done in an orderly and progressive fashion.

Clinical concerns fueled the changing role of the hospital and the changing role of various classes of psychiatric hospitals (state, county, private, general hospital), but clinical concerns were never separate from fiscal concerns. It has never been clear which has been the predominant concern, but there is ample evidence accumulated over the last 150 years that economic considerations have been of no less importance than clinical considerations in defining the role of the hospital.

Social Control Rationale

Every state has statues that govern the admission of persons with mental illness, over their objection, to hospitals within their state; the involuntary

continued hospitalization of these individuals should they want to leave; the commitment of persons who have been charged with a crime and there is a question of their competency to stand trial and/or their criminal responsibility; and the transfer of inmates from jail and prisons to hospitals. Every state sets forth criteria under which an individual can be held on an emergency detention and then civilly committed and these criteria are uniformly danger to self, danger to others, and some variant of gravely disabled. The statutes for civil commitment are generally, uniformly applied across hospital type, while forensic commitments are to government run hospitals or in a few states to nongovernment hospital with contracts with the state to provide such services.

In no other field of medicine could a competent person be admitted to a hospital over his objection. This occurs only in psychiatry and is the component of psychiatry that is social control. This social control function is the most fractious aspect of hospital psychiatry. It is this aspect of psychiatric hospitals that has been the brunt of antipsychiatry movements, whether they originate with psychiatrists themselves, for example, Thomas Szasz, R.D. Laing; religious groups; or individuals who label themselves "consumer survivors."

So engulfing had the social control function of the psychiatric hospital become, that until the early 1980s, simply being an involuntary patient at a psychiatric hospital meant one could be involuntarily medicated. The argument was a simple one: what was the purpose of holding a person in a hospital against his will if you could not treat him? This changed when, through litigation, the point was made that commitment is based on dangerousness, while the right to refuse treatment is based on competency (see, e.g., *Rogers v. Commissioner of Mental Health*, 390 Mass.489, 458 N.E.2d 308 [1982]).

The objection to the social control function of the hospital is hardly new. One-hundred and seventy-five years ago Elizabeth Stone protested her admission to the McLean Asylum; 150 years ago Elizabeth Packard made a career out her protestations; and 70 years ago actress Frances Farmer tried to rally the nation to support her self-proclaimed unjust involuntary admissions and treatment.

This week, in every state in the United States, there are individuals in various stages of civil and criminal commitment. There are also individuals being medicated over their objection. And what some believe is a glorious evolutionary step forward in mental health services and others think is a malevolent regression, a progressively larger percentage of involuntary treatment is taking place outside of hospitals, through AOT, mental health courts, conditional release, representative payeeships, leveraged housing, and probation.

So entrenched have these competing interests of social control versus autonomy become that each has their own set of slogans. Those in favor of

treatment over objection in a defined set of circumstance talk about "rotting with your rights on." Those in favor of free choice no matter what the circumstances chant: "Nothing about us without us."

There is one area where the both sides together object to the use of the psychiatric hospital as a social control agent—for individuals who do not meet the criteria for having a mental illness, or whose dangerous is not on the basis of a mental illness, or who are stuck in a hospital because they have no place to go—yet are committed to the psychiatric hospital. All states have their population of individuals who are there because their psychopathic behaviors land them in hospitals; individuals whose high profile crimes make them "political prisoners" and individuals who dangerous sexual behaviors means they must be locked up somewhere. The later situation is the basis for the Supreme Court decision in Kansas v. Hendricks (see Table 58.3–2).

It is perhaps the irony of the social control function of psychiatric hospitals that law enforcement cannot get individuals into hospitals who need to be there because district attorneys, forensic review boards, families, and victims fight to keep individuals in psychiatric hospitals who do not need to be there. The social control function is highly susceptible to stigma and prejudice.

Table 58.3–2.
Federal Government and Court Actions Defining the Role of Hospital Treatment over the Last Six Decades

1955	The Mental Health Study Act of 1955 (Public Law 182) calls for a nationwide analysis and re-evaluation of the human and economic problems of mental illness. It leads to the formation of the Joint Commission on Mental Illness and Health.
1963	President John Kennedy submits to Congress a "Special Message on Mental Illness and Mental Retardation" that calls for a "wholly new national approach" to those two health problems.
1960	In *Dusky v United States,* U.S. Supreme Court affirmed a criminal defendant's constitutional right to have a competency evaluation before proceeding to trial and setting the standard for determination of such competence. This has subsequently meant many individuals are sent to state hospitals every year for competency evaluations.
1964	*Dixon v. Weinberger*, filed in U.S. District Court of the District of Columbia, claims that patients at St. Elizabeth's Hospital have a statutory right to treatment and that those involuntarily committed under the 1964 Hospitalization of the Mentally Ill Act of the District of Columbia must be placed in the least restrictive setting consistent with suitable treatment.
1965	Medicare legislation is passed. It includes limited coverage for patients receiving active treatment in state hospitals in addition to those in general hospitals.
1966	The Social Security Amendments of 1965 (Public Law 89–97) adds Title XIX, Medicaid, to the Social Security Act. Medicaid funds psychiatric treatment in general hospitals for those under categorical assistance programs and who qualify as medical indigents, and it improves coverage for psychiatric illness under Medicare.
	In *Lake v. Cameron*, a U.S. court of appeals rules that an individual could not be committed to the hospital until hospital officials determined that there was no less-restrictive facility available to care for her.

In *Rouse v. Cameron*, a U.S. court of appeals finds that a criminal defendant who is acquitted by reason of insanity and involuntarily committed to a psychiatric hospital has a right, legally enforceable through habeas corpus, to adequate and suitable treatment.

1970 *Wyatt v. Stickney* is filed in federal district court in Alabama, addressing the issue of detention without treatment of involuntarily civilly committed persons. The court subsequently finds that three fundamental conditions are necessary for adequate and effective treatment in public psychiatric hospitals: A humane psychological and physical environment, enough qualified staff to administer adequate treatment, and individualized treatment plans.

1972 In *Lessard v. Schmidt*, a federal district court decides that under the due process provisions of the Constitution, persons facing involuntary civil commitment are guaranteed the full array of procedural safeguards formerly guaranteed only to individuals charged with a crime.

1973 In *Souder v. Brennan*, a federal district court finds that patient-workers are covered by the Federal Fair Labor Standards Act, entitling them to minimum wages and overtime compensation.

1974 In *Donaldson v. O'Connor*, the U.S. Fifth Circuit Court of Appeals rules that a person who is involuntarily civilly committed to a psychiatric hospital has a constitutional right "to such treatment as will help him be cured or to improve his mental condition." The U.S. Supreme Court issues its decision in this case on June 26, 1975, finding that "a State cannot constitutionally confine, without more, a nondangerous individual who is capable of surviving safely in freedom by himself or with the help of willing and responsible family members or friends."

1975 The first class-action suit on the right to refuse treatment is filed for Boston State Hospital patients in a case originally known as *Rogers v. Okin*. (Right-to-refuse-treatment cases were basically decided on state-by-state basis to the end of the century, with very different rulings among the states.)

1975 In *Drope v Missouri*, the U.S. Supreme Court rules when deciding whether to evaluate a criminal defendant's competency, the court must considered any evidence suggestive of mental illness, even one factor alone in some circumstances. Therefore, the threshold for obtaining a competency evaluation is low. When the issue is raised, the motion should be granted. This means even more individuals sent to the state hospital for evaluation.

1976 In *Tarasoff v. Regents of the University of California*, the California Supreme Court rules that "when a therapist determines, or pursuant to the standards of his profession should determine, that his patient presents a serious danger of violence to another, he incurs an obligation to use reasonable care to protect the intended victim against such danger."

1978 In *Rennie v Klein*, the U.S. Supreme Court rules that an involuntarily committed, legally competent patient who refuses medication has a right to professional medical review of the treating psychiatrist's decision. The Court left the decision-making process to medical professionals. This decision changed the long-standing practice whereby a committed patient could be treated over his object simply because he had been committed.

1979 In *Addington v. Texas,* the U.S. Supreme Court holds that the standard of proof in civil commitment hearings is "clear and convincing," a lower standard than "beyond a reasonable doubt."

1980 The Civil Rights of Institutionalized Persons Act allows the federal government to initiate actions against states whose public institutions, such as hospitals, prisons, nursing homes, and jails, deny residents their constitutional rights.

In *Suzuki v. Yuen,* the U.S. Court of Appeals for the Ninth Circuit rules that involuntary civil commitment solely on the grounds of danger to property is unconstitutional.

1981 In a 5-to-4 decision, the U.S. Supreme Court rules that inpatients of public psychiatric institutions are not eligible for Supplemental Security Income (SSI) payments granted to mentally ill patients in other settings.

1982 In *Youngberg v. Romeo*, the U.S. Supreme Court decides that a person in an institution has a constitutionally guaranteed "right to personal security," "a right to freedom from bodily

restraint," and the right to receive "such training as an appropriate professional would consider reasonable to ensure his safety and to facilitate his ability to function free from bodily restraints."

Congress enacts the Tax Equity and Fiscal Responsibility Act (Public Law 97–248), which sets the stage for a prospective payment system for Medicare.

1985 In *Ake v Oklahoma,* the U.S. Supreme Court rules indigent criminal defendants have a right to a competency evaluation. The potential pool of persons who can be sent to state hospitals for evaluation expands again.

1990 The Americans with Disabilities Act (ADA) is enacted to eliminate discrimination against disabled persons. Title II says, "No qualified individual with a disability, shall, by reason of such disability, be excluded from participation in or be denied the benefits of the services, programs, or activities of a public entity, or be subjected to discrimination by any such entity."

1991 The Patient Self-Determination Act, a part of the Omnibus Budget and Reconciliation Act of 1990, requires health care facilities that receive Medicare and Medicaid funding to provide information to adult patients about their right to make their own health care decisions, including the right to accept or refuse treatment and to execute advance directives about medical care.

1997 In *Kansas v. Hendricks*, the U.S. Supreme Court rules that sex offenders can be civilly committed because they "suffer from a volitional impairment rendering them dangerous beyond their control" and because "far from any premature objection, the confinement's duration is instead linked to the stated purpose of the commitment, namely to hold the person until his mental abnormality no longer causes him to be a threat to others."

1999 In *Olmstead v. L.C. and E.W.*, the U.S. Supreme Court rules that the ADA requires states to provide community placement for persons with mental disabilities if the state's treatment professionals have determined it is appropriate, if it is not opposed by the individuals affected, and if it can be reasonably be provided considering state resources and the needs of other disabled persons.

Publication of *Mental Health: A Report of the Surgeon General*

2000 The Ticket to Work and Work Incentives Improvement Act permits states to increase the level of income disabled persons can earn before they become ineligible for Medicaid and extends the period that disabled persons can be employed and the income they can earn before losing Medicare coverage.

Wyatt v. Sawyer (formerly *Wyatt v. Stickney*) settled.

The Drug Addiction Treatment Act allows office-based heroin treatment by permitting qualified physicians to prescribe take-home doses of buprenorphine.

Federal legislation authorizes $10 million over 4 years to establish up to 100 mental health courts.

Federal rules are issued establishing a national uniform standard for protecting the privacy of medical records.

2001 The Surgeon General promulgates a national strategy for suicide prevention: *National Strategy for Suicide Prevention: Goals and Objectives.*

The Health Care Financing Administration is restructured and renamed the Center for Medicare and Medicaid Services (CMS).

2002 The Department of Health and Human Services releases *Report to Congress on the Prevention and Treatment of Co-occurring Substance Abuse and Mental Disorders*.

2003 The President's New Freedom Commission on Mental Health puts out its report, *Achieving the Promise: Transforming Mental Health Care in America*.

2004 A federal district court in Oregon requires the state to eliminate a $250 copayment for hospital visits for the standard population of the Oregon Health Plan.

2005 The federal government, jointly through the Department of Health and Human Services (DHHS), the Department of Housing and Urban Development, the Department of Justice, the Labor

Department, the Veterans Administration, and the Social Security Administration, put out *Transforming Mental Health Care in America. The Federal Action Agenda: First Steps*. This starts a multiyear venture to alter the forms and functions of the mental health system in the United States.

2006 The 2006 Budget Reconciliation Act alters Medicaid rules, such that states can impose premiums and higher cost sharing on beneficiaries, as well as allows states to provide scaled-back Medicaid benefit packages. One consequence is transferring some Medicaid beneficiaries to the rolls of the uninsured, with hospitalization going from Medicaid reimbursement to free care.

The Substance Abuse and Mental Health Services Administration (SAMHSA) provides a training curriculum to reduce (or eliminate) the use of seclusion and restraint: *Roadmap to Seclusion and Restraint Free Mental Health Services*.

The CMS publishes final rules on the use of seclusion and restraint in the *Federal Register*. The new limits (previously there were none) are: Adults, 4 hours; children and adolescents age 9–17 years, 2 hours; and children younger than age 9 years, 1 hour.

Implementation of Medicare Part D.

2007 The Agency for Healthcare Research and Quality (AHRQ) of the DHHS reports that there is no evidence base for the off-label use of atypical antipsychotics.

SAMHSA, in *An Action Plan for Behavioral Workforce Development*, reports "an anemic pipeline of new recruits to meet the complex behavioral health needs of the growing and increasingly diverse population in this country."

The AHRQ, in *Care of Adults with Mental Health and Substance Abuse*, reports that 6% of all community hospital stays were primarily for psychiatric or substance use disorders and nearly 25% included such a disorder.

Federal judges in Wisconsin question diagnoses invented to civilly commit sex offenders after expiration of their criminal sentence. In *Brown v. Watters* the diagnosis was "paraphilia not otherwise specified-nonconsent"; in *McGee v. Bartow* the diagnosis was "personality disorder not otherwise specified with antisocial feature." Neither is a *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV) diagnosis.

2008 The Paul Wellstone and Pete Domenici Mental Health Parity and Addiction Equity Act of 2008 (MHPAEA) amended the Public Health Service Act, the Employee Retirement Income Security Act (ERISA) and the Internal Revenue Code, to require group health plans and health insurance issuers to ensure that financial requirements (such as co-pays, deductibles) and treatment limitations (such as visit limits) applicable to mental health or substance use disorder benefits are no more restrictive than the predominant requirements or limitations applied to substantially all medical and surgical benefits.

2010 The Affordable Care Act was passed by Congress and then signed into law by President Obama on March 23, 2010. Titles I. Quality, Affordable Health Care for All Americans and II. The Role of Public Programs and V. Health Care Workforce have direct impact on psychiatric hospitals.

2012 U.S. Supreme Court rules in *National Federation of Independent Business v. Sebelius*on June 28, 2012, in a 5-4 ruling, that the Affordable Care Act is constitutional. The court defines the individual mandate as a tax and not a mandate and provides a choice for States to Opt-Out of Medicaid Expansion.

2013 A final regulation implementing MHPAEA is published in the Federal Register on November 13, 2013.

2013 Representative Tim Murphy (R-PA) introduces, on December 12, 2013, H.R.3717—Helping Families in Mental Health Crisis Act of 2013 which is the first comprehensive federal legislation addressing the needs of those with serious mental illness since President Kennedy's Community Mental Health Act in 1963.

2014 Representative Ron Barber (D-AZ) introduces, on May 2, 2014, H.R. 4574—The Strengthening Mental Health in Our Communities Act, the House Democrats' alternative to the bill filed by representative Murphy.

2014 U.S. Supreme Court Justice Ruth Bader Ginsburg files a dissenting opinion in *Burwell v. Hobby Lobby Stores, Inc.*, a decision that exempts employers from covering birth control (as required by the Affordable Care Act) if the employer objects on religious grounds. One of Ginsburg's queries was, could Scientologist employers not cover antidepressants on religious grounds?

2014 On April 1, 2014, President Obama signed into law the Protecting Access to Medicare Act (H.R. 4302), which includes a demonstration program based on the Excellence in Mental Health Act. This act (1) creates criteria for "Certified Community Behavioral Health Clinics," (2) provides $25 million that will be available to states as planning grants to develop applications to participate in the 2-year pilot, (3) stipulates that 8 states will be selected to participate in the 2-year pilot program, and (4) requires participating states to develop a Prospective Payment System for reimbursing Certified Behavioral Health Clinics for required services they provide.

Fiscal Rationale

In 1912, there was a debate in Congress on a government health insurance program. President Franklin Roosevelt almost established a federal health plan as part of the 1935 Social Security Act. Presidents Harry Truman and John Kennedy had health care proposals before Medicare and Medicaid passed during President Lyndon Johnson's term in office. However, before all that, President Franklin Pierce played a major role in the early eschewal of federal responsibility for the health care of persons with mental illness.

The individual most responsible for bringing the issue to President Pierce's desk was Dorothea Dix, a well-known reformer. On June 23, 1848, Dix delivered a Memorial Lecture to the U.S. Congress. Dix's subjects, as stated in her address, were as follows: "Allowing at the present time 22,000,000 inhabitants in the United States and supposing only one in every thousand to be insane or idiotic, we have then 22,000 to take charge of." She described the state they were in:

> a majority of whom are in needy or necessitous circumstances . . . destitute of appropriate care and protection . . . bound with galling chains, bowed beneath fetters and heavy iron balls, attached to drag chains, lacerated with ropes, scourged with rods, and terrified beneath storms of profane execrations and cruel blows . . . abandoned to the most loathsome necessities, or subject to the vilest and most outrageous violations.

Dix's remedy was hospitals: "No disease is more manageable or more easily cured than insanity; but to this end special appliances are necessary, which cannot be had in private families, nor in every town and city; hence the necessity for hospitals." Dix appealed to the legislators' sense of doing what is right: "Humanity requires that every insane person should receive the care appropriate to his condition." She articulated the social control function of the hospital by spelling out the civil and social obligation to consult and secure the public welfare: "First, in affording protection against the frequently manifested dangerous propensities of the insane." Furthermore, she made a powerful economic argument, indicating the insane and the community at large would benefit by "assuring seasonable and skillful remedial care, procuring their restoration to usefulness as citizens of the republic, and as members of communities."

Edward Jarvis best articulated the 19th-century position on the cost of care versus the cost of treatment that was fundamental to Dix's campaign:

> A man of twenty years of age, if sane, has an average life of 39.48 years, while if insane he has but an average life of 21.31 years if not restored to health. The average time for restoring to health the insane who apply for treatment upon the early symptoms of disease is twenty-six weeks. At $4 per week, which was the average cost in the three State Lunatic Asylums in Massachusetts for the past year, this amounts to $104, to which is added $30 for each patient, for the cost of rent or interest on the value of the hospital, etc., for six months, making an average cost of $134 for restoration to health. If not restored to health, the family or State must be at an expense of $156 a year for 21.31 years, and must also lose the patient's earnings for the 39.48 years which he would have made if well. The cost of the patient's support is estimated at $2,121, while the loss of his future labor, if he becomes insane at twenty years of age, is estimated at $2,665.37, making a total loss of $4,786.37 if not cured; while, if cured in the average time of twenty-six weeks at a cost of $134, there will be a gain to the family or to the State of $4,652.

Dix lobbied for her land grants bill through four presidential administrations. The bill was passed by both houses of Congress during Franklin Pierce's term. However, Pierce vetoed the bill. In his Veto Message on May 3, 1854, first he laid out the issues as he saw them:

> It cannot be questioned that if Congress has power to make provision for the indigent insane without the limits of the District it has the same power to provide for the indigent who are not insane, and thus to transfer to the Federal Government the charge of all the poor in all the States. It has the same power to provide hospitals and other local establishments for the care and cure of every species of human infirmity . . . The whole field of public beneficence is thrown open to the care and culture of the Federal Government.

Later in his Message, he explained the basis for his veto:

> If the several States, many of which have already laid the foundation of munificent establishments of local beneficence, and nearly all of which are proceeding to establish them, shall be led to suppose, as, should this bill become a law, they will be, that Congress is to make provisions for such objects, the fountains of charity will be dried up at home, and the several States, instead of bestowing their own means on the social wants of their own people, may themselves, through the strong temptation which appeals to states as to individuals, become humble supplicants for the bounty of the Federal Government.

From the veto by President Pierce into the early decades of the 20th century, the federal government was basically disinterested in the psychiatric and substance abuse populations. It eventually turned to the issues around the treatment of this population due to concerns about drug addiction. In 1914, the Harrison Narcotics Act allocated funds to study drug use. Subsequently, the Narcotics Division within the U.S. Public Health Service was created. It was renamed 2 years later the Division of Mental Health. In 1996, the federal government focused on mental illness with the passage of the National Mental Health Act, which established the National Institute of Mental Health (NIMH).

In 1989, the Omnibus Budget Reconciliation Act (OBRA) required the secretary of the Department of Health and Human Services to conduct a study of the IMD exclusion. The IMD exclusion dated from the 1950 Social Security Act Amendments, in which payments on behalf of those in IMDs

were excluded from old-age assistance. This exclusion continued unchanged until the enactment of Medicaid under the Social Security Act Amendments of 1965. For the first time states could use Federal Financial Participation (FFP) for a population in IMDs—in this case those 65 years of age or older. States could bill the federal government to help pay for the treatment in psychiatric hospitals of patients who were older than the age of 65 years. Other populations were not included, but Congress did expand FFP for services to persons with mental illness in general hospitals.

The 1965 changes to the IMD exclusion began to allow states to shift the cost of mental illness expenditures to the federal government. In a state hospital, the state bears 100 percent of the cost of care and treatment. The state cannot bill Medicaid. In any facility that is eligible for Medicaid reimbursement, each state pays considerably less than 100 percent, according to the FFP formula: Nonfederal percentage = 45 percent × (per capita state income) ÷ (per capita national income). Historically the range of federal participation has been between 50 and 83 percent—thus the fiscal incentive to move patients out of state hospitals.

An IMD is a facility with more than 16 beds. In addition, any one of the following defines an IMD: (1) the facility is licensed as a psychiatric facility, (2) the facility is accredited as a psychiatric facility, (3) the facility is under the jurisdiction of the SMHA, (4) the facility specializes in providing psychiatric/psychological care and treatment, or (5) the current need for institutionalization for more than 50 percent of all of the patients/residents in the facility results from mental diseases.

In 1971, Congress expanded Medicaid to include coverage of services rendered in intermediate care facilities (ICFs). In 1972, Congress expanded FFP for inpatient psychiatric treatment in IMDs to individuals younger than 21 years of age, so that now those younger than 21 and older than 65 years of age are eligible.

The IMD exclusion sets up a tension between the states and the federal government, each wanting to lower its share of the costs for psychiatric hospital treatment. Between 1972 and 1988, the federal government was busy with evaluating the proper application of the IMD exclusion by the states, disallowing certain states' practices, and defending itself in court. This culminated in the 1985 U.S. Supreme Court decision in *State of Connecticut Department of Income Maintenance v. Heckler* in which the Court unanimously found for the federal government, indicating that a hospital, a skilled nursing facility (SNF), or an ICF could be an IMD.

Thus, the IMD exclusion profoundly affects the role of the psychiatric hospital and the distribution of patients across types of hospitals. States would rather have persons with mental illness in community residential settings than in state hospitals, and in general hospitals rather than state hospitals. In both situations, the former means Medicaid payments; the latter means no Medicaid reimbursement.

No matter what the sector or who the payer, however, in the last quarter of the 20th century inpatient psychiatric treatment was getting more and more costly—hence the rise of managed care. Starting in the 1970s and continuing through the early 1980s, the principal methods for cost containment were actuarial manipulations designed to manage benefits. There was little regard for the effect these changes in benefits had on the indirect costs borne by beneficiaries, employers, and/or the community. Changes were made without consideration of their effects on the displacement of care onto the general medical sector. Modifications in benefit plan design included imposition of restrictions, exclusions, limitations, and penalties; provision of incentives; requirements for prior authorization; and retrospective denial of claims. This methodology is managing benefits.

Managing benefits changed into cost control through management of care, a methodology claimed to be more patient focused. By limiting authorization of expenditures to only those services deemed medically necessary and appropriate, delivered in the least restrictive and least intrusive treatment settings, and only by previously designated qualified practitioners, advocates of managed mental health care believed that proper care could be delivered to patients at the appropriate times, in designated settings, by the right type of provider, in skillfully titrated amounts, and at considerable cost savings. Techniques used to achieve this end included management through contracting, profiling, credentialing, and educating hospitals; organizing hospitals into a specialized, accessible, and interactive network of behavioral health care providers; using mandated standardized pretreatment assessment and treatment planning methods; and requiring the use of standardized forms, restricted formularies, and outcomes management.

Though not yet achieved, the more progressive goals of managed behavioral health care are disease management, demand reduction, health promotion, and wellness. Health maintenance organizations (HMOs) and other capitation-based arrangements have fiscal incentives to use these strategies, but both efforts and results have yet to be really impressive. Nonetheless, the techniques used mean that managed care has and continues fundamentally to affect the role of the psychiatric hospital.

The federal government sought further cost containment. The Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA) led to new methods to control Medicare costs, most significant being diagnostic-related groups (DRGs) for reimbursement. With the end of cost-plus reimbursement and its replacement with reimbursement based on diagnosis, hospital profits for medical and surgical treatment began to decline. The behavioral health care industry was initially unaffected by DRGs, largely because psychiatric diagnoses were thought to lack sufficient power to predict the course of treatment. Thus, there appeared to be tremendous financial opportunities

available in the relatively unregulated behavioral health care market. These opportunities were short lived; the rise and fall of for-profit psychiatric hospitals were equally steep, as described earlier. With the application of DRGs to psychiatry, this fiscal advantage also dissipated. Thus, the federal government, through Medicare and Medicaid, fundamentally affects the role of the psychiatric hospital.

Although the modern era of managed behavioral health care began in the private sector, these principles and practices spread to the public sector. Medicaid Managed Care was initiated due to the spiraling costs of the Medicaid program. Between 1985 and 1993, Medicaid costs tripled and the number of beneficiaries increased by more than 50 percent.

Other Governmental Considerations

Although state governments most directly affect the roles of psychiatric hospitals—through state laws, regulations, licensure, inspections, and investigations—the federal government has profoundly changed the roles of psychiatric hospitals through federal legislation and court decisions. Table 58.3–2 provides a summary of these influences over the last six decades.

The federal government's influence can be examined through the example of the right to treatment. Much of the right-to-treatment litigation has focused on establishing standards of care and concomitant staffing patterns. In the 1940s, the APA maintained there should be no less than 1 psychiatrist for every 150 hospitalized patients, 1 graduate nurse for every 40 patients, and 1 attendant for every 8 patients. No state hospital of the day met all of the APA's standards.

The right to treatment was initially articulated by Morton Birnbaum, a lawyer and physician in 1960. There followed a series of lower-court cases: *Lake v. Cameron* (1966), *Rouse v. Cameron* (1966), *Nason v. Superintendent of Bridgewater State Hospital* (1968), *Wyatt v. Stickney* (1971), and *New York State Association for Retarded Children v. Rockefeller* (1973), the so-called Willowbrook case. Each of these is briefly explained in Table 58.3–2.

The U.S. Supreme Court decided three cases between 1975 and 1982 that might have established a right to treatment, meaning states that would be obligated to treat committed patients. In *O'Connor v. Donaldson* (1975), the Court found "A State cannot constitutionally confine, without more, a nondangerous individual who is capable of surviving safely in freedom by himself or with the help of willing and responsible family members or friends." In *Halderman v. Pennhurst State School & Hospital* (1977), the Court again did not reach the issue of a constitutional basis for a right to treatment or habilitation. In *Youngberg v. Romeo* (1982), the Court went the furthest it has to date when it held that persons with mental retardation who were involuntarily committed to institutions have "constitutionally protected liberty interests under the Due Process Clause of the Fourteenth

Amendment to reasonably safe conditions of confinement, freedom from unreasonable bodily restraints, and such minimally adequate training as reasonably may be required by their interests." Had the Supreme Court established a right to treatment, the roles of psychiatric hospitals might have been considerably expanded.

The major advances in the right to treatment in the 1980s occurred pursuant to the Civil Rights of Institutionalized Persons Act (CRIPA) of 1980. This act authorized the U.S. government to institute a civil action against any state whose officials, employees, or those acting on their behalf are "subjecting persons residing in or confined to an institution to egregious or flagrant conditions which deprive such persons of any rights, privileges, or immunities secured or protected by the Constitution or laws of the United States causing such persons to suffer grievous harm." Under this authority the Civil Rights Division of the U.S. Justice Department has been able to substantially bring the right to treatment closer to a reality for thousands of persons in state, county, or municipal psychiatric hospitals by applying the standards of *Youngberg* to negotiate consent decrees. A consent decree is a document filed with the court in which the two parties sign a binding agreement in lieu of litigation.

On July 26, 1990, the Americans with Disabilities Act (ADA) passed and established a "clear and comprehensive prohibition of discrimination on the basis of disability." Title II is of greatest interest because this title covers state and local government services and so includes public hospitals. Title II indicates "no qualified individual with a disability shall, by reason of such disability, be excluded from participation in or be denied the benefits of the services, programs, or activities of a public entity, or be subject to discrimination by any such entity." One regulation promulgated to operationalize this Act states "A public entity shall administer services, programs, and activities in the most integrated setting appropriate to the needs of qualified individuals with disabilities." This marks a shift from the long-standing standard, "least restrictive alternative," to the new standard, "most integrated setting." The ADA was seen not so much as a way to improve services within hospitals as a way to force states to move patients out of these institutions.

The hallmark Supreme Court decision to date applying the ADA to persons in hospitals is *Olmstead v. L.C.* (1999). The case involved two women who were in a Georgia state hospital despite the fact that treatment professionals had concluded each could be appropriately treated in community-based programs. The Court found that the "unjustified isolation . . . [of persons with mental disabilities] is properly regarded as discrimination based on disability." The basis for this finding was that "Institutional placement of persons who can handle and benefit from community settings perpetuates unwarranted assumptions that persons so isolated are incapable or unworthy of participating in community life." The

justices then made a leap to the conclusion that institutional confinement "severely diminishes the everyday life activities of individuals, including family relations, social contacts, work options, economic independence, educational advancement, and cultural enrichment." Thus, began the most recent wave of translocating persons from public psychiatric hospital to alternative settings. Apparently, general hospitals and private hospitals were deemed "in the community." So too, based on the most recent data, are jails and prisons.

FULFILLING THE ROLE

The hospital fulfills its role in the care and treatment of psychiatric patients in a prescribed way. Efficacious treatment derives from effectively performing each step.

Assessment

Assessments are done by individual practitioners of different disciplines. An admission assessment is conducted by a physician, preferably a psychiatrist. This assessment, done at the patient's presentation to the hospital, documents the chief complaint; the basis for admission and history of the present illness; relevant past psychiatric history, family history, and social history; a complete mental status examination; a substance abuse screen; and a risk assessment (suicide, homicide, fire setting, other). This assessment should lead either to admission or to nonadmission with an outpatient plan.

A nursing assessment is completed as soon after the patient's arrival on the ward as possible, but no later than 8 hours. The nursing assessment overlaps in part with the physician admission assessment but covers some areas in less detail and others in more detail. The nursing assessment includes a pain assessment, a falls risk assessment, skin care concerns, and special nursing care needs.

The third assessment that needs to be completed proximate with admission is the medical history and physical examination. This is done either by the admitting psychiatrist or by an internist, depending on the staffing of the hospital. Besides attending to a general medical history, particular attention needs to be paid to how medical disorders and physical findings might be affecting the psychiatric presentation.

All three of these assessments need to cover the areas of alcohol and substance use/abuse. Positive screens lead to an alcohol and drug abuse assessment, preferably completed by staff with expertise in this area.

Other disciplines need to complete their assessments prior to the interdisciplinary team (IDT) meeting at which the Treatment and Recovery Plan (TRP) is developed. These assessments include those by social work—the psychosocial assessment—and by rehabilitation services (often

inappropriately referred to as "activities therapy"). The psychosocial assessment needs to provide information from collateral sources when such information is necessary to complete the database, there are collateral sources to contact, and the patient has provided permission (release of information) to contact specific individuals. Some hospitals include a psychology assessment on all admission; others use psychologists who function as consultants, and psychology assessments are conducted only on referral. Dietician assessments are often done routinely; medical specialty assessments are done on referral.

Formulation

Traditionally, the weakest part of the process from assessment to intervention and one often left out entirely, the formulation needs to be seen as a fundamental part of the process. The formulation pulls the data together into a coherent picture of the individual. It is not a brief restatement of the facts of the case. The formulation takes the symptoms, which are the basis for an understanding of the *patient*, and adds significant historical information to provide an understanding of the *person*.

Leaving out or improperly constructing the formulation leads to the loss of important information that does not directly translate into diagnoses but is crucial for effective treatment planning. Formulations may be recorded as components of each assessment—for example, psychiatric formulation, nursing formulation—and then brought together at the IDT meeting, or they may be written at the IDT meeting by the psychiatrist drawing from all of the assessments.

Treatment Planning

Because treatment must commence at the onset of hospitalization, the Initial Treatment and Recovery Plan (ITRP) is, of necessity, a multidisciplinary process. The ITRP is initiated by the admitting psychiatrist and nurse, based on their assessments. Other disciplines add to the ITP as they conclude their assessments.

Whereas the ITRP provides the basis for directing treatment during the first few days of hospitalization, the rationales for specific interventions must be included in either the assessments or in the progress notes. For example, each medication ordered by the admitting physician must have a documented rationale. It is not sufficient simply to state that medications taken prior to admission are being continued, because this does not provide any basis for the medications. Stating that aripiprazole and clozapine, for instance, are being prescribed fails to justify polypharmacy without providing a basis beyond recording that these were the medications prescribed by the CMHC psychiatrist. Furthermore, any as-needed (PRN)

medication ordered at admission needs to have a rationale, one that is specific to the patient. Ordering lorazepam 2 mg orally for agitation and haloperidol 5 mg orally for severe agitation is not sufficient. Such orders are neither specific to the patient nor do they inform the nurse what the prescriber means by "agitation" and what he means by "severe agitation."

The TRP expands upon the ITRP and derives from an *interdisciplinary* process. The IDT is composed of professional and paraprofessional staff, as well as the patient. The patient, assisted by staff, is free to invite any other individuals to this meeting, for example, family members, friends, or advocate. The purpose of the IDT meeting is to *write* the TRP. It is not a time to sanction a TRP written before the meeting by one member of the IDT. The meeting uses the multidisciplinary assessments as the basis to create an interdisciplinary treatment plan.

The creation of the TRP follows an orderly process. The IDT establishes, with the patient as the lead, the problems, which in the inpatient setting are predominantly barriers to discharge. Other problems may be addressed, depending on the expected length of stay of the patient. For each problem, there is a long-term goal (LTG) and a set of short-term goals (STGs) or objectives. The goals are the patient's goals, not the staff's goals for the patient. Hence, problems and goals belong to the patient. For each STG, there is a series of interventions. Interventions are what the staff does, not what the patient does. Each intervention under an STG must specifically assist the patient to move toward achievement of that STG.

STGs need to be specific, measurable, attainable, realistic, and timely (often referred to by the acronym SMART). STGs are intermediate steps toward accomplishment of the LTG. Small steps that the patient reaches at intervals reinforce the patient and the staff.

Interventions need to specify the frequency of the intervention, the duration of the intervention, and the name of the intervener. For example, "Nurse Nightingale provides Medication Education Group, Tuesday and Thursday, 10:00-10:45 a.m."

Common errors in interventions include the following:

1. Writing job descriptions: "Nurse Nightingale will administer medication and monitor for side effects."
2. Failing to document intervener and/or duration: "Anger Management Group Wednesdays."
3. Including interventions unrelated to the STG, such as "Dr. Woodward will meet with patient Mondays, 1:00-1:20 p.m. to review concerns about weight gain from clozapine," when the STG is "Patient will have one five-minute telephone conversation per day with mother for three consecutive days without telling her mother to drop dead or piss off."
4. Using terms that communicate no specific intervention (and can often be dangerous to staff if done wrong) such as "encourage patient to . . . " or

"assist patient to . . . " If such interventions are used, the intervention must be explicit about how the "encourage" is to be done.

The ideal sets of interventions are clear, concise, and explicit. If the entire IDT were out sick one day, covering staff should be able to treat the patient by performing the interventions.

The TRP is the road map for the patient's treatment. Active treatment—a concept or phrase used by many certifying entities—is defined as that treatment found as an intervention on the TRP that moves the patient toward achievement of an STG that moves the patient toward achievement of an LTG that addresses one of the patient's problems. A simple rule when one is calculating active treatment hours for a patient is, "if it's not on the treatment plan, it's not active treatment."

Progress is recorded in progress notes. These are written pointedly addressing STGs. Each note needs to be labeled with the discipline of the intervener and the STG being addressed. Thus, there would be no nursing staff note reading "Patient ate well and slept well" unless eating and sleeping were problems on the TRP with specific LTGs and STGs.

The initial TRP is completed within a specified time frame, varying between 1 and 10 days and specified by the facility's policies and procedures. The TRP is reviewed through IDT meetings at which treatment plan reviews (TRPRs) are conducted. The TRPR builds on the TRP, uses progress or lack thereof noted in the progress notes, and modifies the TRP as needed. TRPRs also occur on a schedule. In units with short lengths of stay, the TRPR occurs from daily to every 3 days. In intermediate length of stay units, TRPRs occur weekly for 60 days, then every 2 weeks for 60 days, then monthly. In longer stay units, TRPRs may occur monthly after the TRP done at 10 days.

Besides occurring on a scheduled basis, TRPRs need to occur in response to a patient's clinical need. Many hospitals have thresholds that, when met, require a TRPR to be held within 1 business day. Such thresholds included (1) any episode of seclusion or restraints; (2) a specific number of episodes of seclusion or restraint or a specific number of hours of seclusion or restraint within a designated time frame; (3) the administration of any psychiatric medication PRN or a given number of PRNs within a specified time frame; (4) the initiation of 1:1 observation; and/or (5) any serious incident, for example, patient-on-patient assault, patient-on-staff assault, or suicide attempt.

The TRP needs to be comprehensive, or account for why it is not (either in the TRP or in the hospital's policies and procedures) in all but acute care settings. Thus, the TRP would be expected to address all of the patient's psychiatric and medical problems on longer-term units. A TRP could provide interventions for those problems that are impediments to discharge and address other problems as resources allow; this is often the

case on intermediate-length-of-stay units. A CTP could address only those problems that are impediments to discharge with the expectation other problems will be addressed in treatment after discharge; such an approach is often employed on short-term units. This does not mean that maintenance treatments will not be provided for acute or chronic problems that are not the focus of the hospitalization, such as antihypertensive medications, treatments for gastroesophageal reflux disease or constipation, and hypnotic medications.

There is a new emphasis on tying the hospital TRP to a community TRP for persons with SMI in systems of care (traditionally CMHCs, but now also health homes and other integrated care settings). In the person-centered approach, the community TRP starts with the individual's "life goals." Community clinicians, support and outreach staff (peer and nonpeer workers), and naturalistic resources are channeled toward assisting the individual to move ever closer to his life goals. Hospitalization is usually a detour on the journey toward one's life goal, so focusing on problems that are impediments to discharge is in the service of returning the person to the community to take up the pursuit of his individualized life goals. To do so effectively, the hospital TRP should record the life goals as anchor points.

Treatment

Treatment planning is necessary but not sufficient for effective treatment to ensue. Members of the IDT plus other staff need to provide integrated treatment that follows the plan.

Psychopharmacological treatment is a mainstay of contemporary inpatient treatment but must be prescribed (physician), dispensed (pharmacist), and administered (nurse) with awareness of contemporaneous interventions. Either initiating psychotropic medication or adjusting medication regimens is considered the sine qua non of inpatient psychiatry. Of course, not all admissions require medication, and distinguishing these patients from all others is an important function of the assessments and observations that are unique to the inpatient setting.

In prescribing medication, the attending psychiatrist must communicate the reasons for the medication and its side effects not only to the patient (obtaining informed consent before ordering the medication) but also to all members of the IDT. Hence, the importance of monitoring progress and informative progress notes accompanying each order.

PRN medications warrant particular attention. PRN medication transfers de facto prescription authority from physician to nurse because the nurse decides when and how much medication the patient will receive (within the parameters set by the physician). Unless the physician checks daily on the total dose of each PRN medication given, it is not possible for the physician to relate the patient's progress to the medication

administered. Hospitals have taken various approaches to PRN psychotropic medication, including eliminating them entirely; limiting the orders to short-term orders, for example, 3 or 5 days; or allowing PRNs only on admission units and not on other units within the hospital.

Well-written progress notes by the psychiatrist assist the nurse in writing useful nursing progress notes. Writing "patient remained calm and cooperative throughout shift" may convey no useful information if, for example, the patient is depressed with a recent suicide attempt and is externally calm and cooperates because it takes less energy and concentration to do so than not to do so. A physician progress note detailing the specific target signs and symptoms for which antidepressant X has been prescribed provides the foundation for the nurse to write progress notes that focus on those signs and symptoms. If the physician simply writes antidepressant X for "dysphoria" or "depression," progress notes by other staff members are likely to be equally uninformative.

Another function shared predominantly by the psychiatrist and the nursing staff (Registered Nurse [RN], Licensed Practical Nurse [LPN], Mental Health Technicians [MHTs]) is patient safety. Maintaining a safe milieu and maximizing the safety of each patient and staff depends on (1) active, individualized treatment, (2) levels of observation and privileges geared to the patient's level of risk, (3) prompt and skillful response to patient aggression, (4) modification of therapeutic interventions as the nature and/or level of risk changes, and (5) continuing staff education and training.

Level of observation and privileges serve the function of patient safety and should not be used as rewards for compliance with process variables like attendance at groups. When the level of observation is for the risk of harm to self, the most noteworthy harm being suicidality, the hospital needs a process by which increasing the level of observation, that is, more staff oversight, can easily be instituted by a physician or RN, but decreasing the level of risk can only occur on a written physician's order after a face-to-face evaluation. In some hospitals, this can only be done by the attending psychiatrist. Telephone orders and verbal orders should never be accepted for lowering the observation level.

In addition to medication, inpatient treatment involves individual sessions and group interventions. Individual psychotherapy is the exception these days, although this modality continues to be used in some public and some private hospitals. There are a few hospitals in which intensive psychotherapy is the mainstay of inpatient treatment. The programming of individual sessions with the social worker to focus on discharge planning is the rule. Individual sessions with nursing staff to practice skills taught in groups should occur at least through the day and evening shifts (and the night shift if appropriate to the patient's needs).

Group interventions fall under the heading of psychosocial

rehabilitation. In general hospitals, these groups occur on the unit. In state hospitals and in some private hospitals, these interventions are delivered off-ward in areas of the hospital often referred to as the "treatment mall." There are several advantages to setting up groups or classes in a centralized location, including the following:

1. Patients follow a variant of the "work order day"; patients, like staff, "commute" to where they spend their daytime hours; in the case of patients, they live in one place (the unit) and get their treatment in another location (the mall); lunch is often provided at the mall.
2. Pooling patients, and consequently pooling staff, means that a large number of different groups can be provided per hour, rather than having the same group occur simultaneously in multiple locations.
3. The issue of patients remaining in bed, watching TV, or participating in other diversions is handled by institutional policy and practice rather than day by day, case by case.
4. Safety of staff and patients is increased because there is more staff available for an emergency.
5. Professional and paraprofessional staff spend more time with patients because patients and staff are at the treatment mall during mall hours.

Newly constructed hospitals (see Table 58.3–1) are designed with accessible treatment mall locations. The use of perimeter security with significant freedom of travel for patients within the building eases access to mall-based treatment and maximizes patients' autonomy (see Fig. 58.3–5). One controversy surrounding the use of treatment malls at hospitals is whether or not the hospital can require all patients to attend the mall. Does the right to refuse treatment also mean the right to not go where treatment is provided? In one model patients have a right to refuse treatment, but patients do not have a right to determine where they are at the hospital at any given time. If this were a right, patients could refuse to remain behind any locked door. Patients must go to the treatment mall, but the hospital must provide a place for them on the mall if they refuse to attend the groups they are scheduled to attend. A corollary controversy in this model relates to the issue of whether staff can lay hands on a patient if the patient refuses to go to the mall. Staff can lay hands on a patient if there is an emergency, and leaving a patient on a ward with no staff because of the patient's refusal to leave the ward would constitute an emergency. In another model, where patients cannot be compelled to leave the unit for groups, hospitals have created units with and without an off unit requirement. There are benefits and drawbacks of each type of unit and absent safety concerns; a patient makes his choice based on his valuation of the different benefits of each type of unit.

Although there are no controlled studies of the operation and effectiveness of treatment malls, data collected at individual hospitals seem

to support (1) rapid adjustment of patients to a change to the treatment mall approach; (2) very low use of hands-on procedures to move patients; (3) decreased rates of seclusion and of restraint without accompanying increases in total daily dose of antipsychotic and/or antimanic medications or of benzodiazepines; and (4) improvement in patients' sleep with fewer doses of hypnotics. Most importantly, patients do seem to learn skills through the use of mall-based groups.

Clinicians doing hospital-based treatment, just like their colleagues doing outpatient treatment, need to be mindful of three contemporary concepts: recovery, evidence-based treatment, and trauma informed care. Practitioners of all clinical disciplines have struggled to achieve consensus definitions of these percepts. SAMHSA's definitions are as follows:

Recovery: A process of change through which individuals improve their health and wellness, live a self-directed life, and strive to reach their full potential.

Evidence-based: Demonstrating effectiveness in empirical research that meets a standard of scientific rigor as embodied in minimum review requirements, which include the stipulation that an intervention must have demonstrated one or more positive behavioral outcomes ($p \leq 0.05$) in substance abuse and/or mental health in at least one study using an experimental or quasiexperimental design.

Trauma-informed care: It is a strengths-based service delivery approach that is grounded in an understanding of and responsiveness to the impact of trauma, that emphasizes physical, psychological, and emotional safety for both providers and survivors, and that creates opportunities for survivors to rebuild a sense of control and empowerment.

A well-functioning psychiatric hospital would incorporate these principles into its therapeutic offerings by having expertise in and providing among other offerings: CBTp for psychosis; cognitive enhancement and restructuring therapy (CERT); cognitive restructuring for PTSD; cognitive therapy for psychosis (CT-R); dual recovery (mental illness and substance abuse); dialectical behavior therapy; family support and treatment; fire-setting treatment; grief therapy; illness management and recovery (IMR); individualized exercise plan (IExP); leisure skills development; medical care and treatment; mental illness and problematic sexual behavior treatment (MIPSB treatment); motivational interviewing and motivational enhanced therapy; person-centered approach; pharmacotherapy; positive behavioral support plan; readiness for change interventions (individual and group); sensory modulation; vocational rehabilitation; and wellness management recovery action plan

(WRAP). Behavioral interventions, generally under the direction of a psychologist with particular expertise in this area, need to be available at all hospitals. This is of paramount importance in hospitals that have units with intermediate- or long-stay patients. The behavioral plan is based on a functional assessment of the target behavior to be modified or eliminated. The plan provides a series of responses aimed at reinforcing some behaviors and extinguishing others. All direct care staff must be trained in the specific interventions to be used for a specific patient. Preferably, the hospital has a way to document such training and to ensure that any staff assigned to work with a patient who has a behavior plan is trained in that plan.

During hospital treatment, a major controversy has been patients' "right to smoke." It is unclear how this is a right. If it were, patients would also have a "right to drink." No one argues the latter, so why the former? Forbidding smoking on the unit has generally occurred at all hospitals. Moving that restriction to a designated area within the hospital was the next step. Patients can often be seen at the perimeters or in the courtyards of the hospital smoking. Many hospitals have progressed to prohibiting smoking anywhere in the building and some even prohibit smoking on the hospital grounds. Hospitals should neither be in the business of selling cigarettes nor use staff to purchase cigarettes for patients—a practice referred to as "making a cigarette run." Given that there is now a range of aids to stopping smoking—gum, patches, pills, support groups—these should be made available to patients and to staff. Nicotine replacements should be used as a transition to becoming nicotine free; too often they become permanent fixtures on the patient's pharmacologic plan. It was not that long ago that state hospitals received free tobacco from cigarette companies, purchased cigarette papers, and paid staff, and/or had patient groups whose focus was to roll cigarettes. The rolling machines had been purchased with state funds. Progress in smoking cessation needs to continue, particularly given the increased morbidity and mortality of persons with chronic mental illness. A role of the hospital should be to teach patients and support their efforts to maximize a healthy lifestyle. Wellness programs are being conducted at many hospitals and are well worth supporting.

Discharge Planning

The aphorism "Discharge planning begins at the time of admission" is, of necessity, particularly true today, given all of the pressures surrounding length of stay. Although the discharge process is primarily the responsibility of the IDT's social workers, it is a mistake to assume it is her or his task alone. Discharge planning is an interdisciplinary process.

Discharge planning is processed by the social worker interacting with the patient and IDT members until the TRP. At this point, the discharge plan becomes an integral part of the TRP. Whether discharge-related issues are identified as one of the problems on the TRP depends on the patient's specific circumstances. Discharge might be a problem if the patient is homeless, cannot return to the residence from which he or she was admitted, or is an illegal immigrant with no eligibility for entitlements. An often-missed problem related to effective discharge is recidivism, that is, the patient has multiple admissions, generally with short community tenure between them. Recidivism is often too quickly ascribed to noncompliance with medication without the IDT doing a thorough investigation of the etiology of repeated admissions. Noncompliance may well be a factor in any short time period between admissions, but this may also be accounted for by a number of other factors. Often ignored is the contribution to stabilization made by the structured environment of the hospital. Often commented on, but inadequately addressed, is the use of alcohol, marijuana, other street drugs, and other prescribed medication even in a medication-compliant patient. There is also the contributing factor of secondary gain—the hospital may be the locus of the best quality of life the patient has experienced for years, or even decades.

Discharge planning involves the establishment not only of sites of service and appointment times but also of efficacious linkage strategies. The steps that need to be taken to maximize the likelihood that the patient will actually make it to her or his first appointment with the outpatient psychiatrist are of paramount importance. Linkage also involves providing to the outpatient psychiatrist, outpatient therapist, and other appropriate personnel pertinent information in writing. A copy of the instructions given to the patient should reach the outpatient sites *before* the patient arrives for her or his first appointment; the discharge summary needs to follow shortly thereafter. So too does the TRP, so that the community can incorporate the work the hospital has done into the community TRP, just as the hospital incorporated the work the community had done into the hospital TRP. Phone contact from inpatient psychiatrist to outpatient psychiatry is advisable for any case seen as needing more than the usual written communications.

SPECIAL POPULATIONS

This section has focused on the role of the hospital in serving the general adult psychiatric population. The hospital serves many other populations, defined either by their demographics or by specific service needs.

Children and Adolescents

Twenty-one percent of U.S. children aged 9 to 17 years have a diagnosable

mental or addictive disorder causing at least minimal impairment, and 5 to 9 percent of children this age have a serious emotional disturbance that causes extreme functional impairment. Significant numbers of children and adolescents in need of treatment do not receive behavioral health services. When children's mental health services are unavailable, unaffordable, or inappropriate, many young people end up in the child protection or juvenile justice systems.

Children and adolescents represent approximately 7 percent of the total mental health inpatient population. However, children's services receive a disproportionally low fraction of what adults receive for inpatient services, despite the fact, that on average, children's inpatient treatment is about 1.5 times as expensive as adult inpatient treatment.

Although the number of available inpatient psychiatric beds decreased, the admission rates of children and adolescents to psychiatric hospitals increased over the last few decades. Between 1986 and 1997, the number of youth admitted to inpatient services increased by 143 percent and the rate per 100,000 youths increased by 119 percent. As the number of inpatient beds and specialized staff to care for children and adolescents with serious emotional disturbance has decreased, the number of families relying on hospital EDs for psychiatric services has increased. ED physicians report seeing a large number of children whose parents have simply run out of options for care.

Inadequate hospital-based and other services have created cohorts of labeled children. "Boarder kids" are children stuck for days, weeks, or even months in pediatric wards because there are no beds in a psychiatric hospital or ward. "Wait-list kids" wait months for case management or outpatient services. "Stuck kids" are in psychiatric hospitals but cannot leave because there is no place for them to go.

There are many factors that complicate the role of the hospital in the care and treatment of children and adolescents. Alcohol and illicit drugs are a major problem. Current alcohol use is 36.2 percent for 17-year olds, 19.9 percent for 15-year olds, and 6.5 percent for 13-year olds. Among youth ages 12 to 17 years, about 12 percent are current illicit drug users. Annual suicide attempts in 15- to 24-year olds have surpassed 1 million in the United States, and suicide is the third-leading cause of death for this age group. Age-specific mortality rates from suicide are 1.6 per 100,000 for 10- to 14-year olds, 9.5 per 100,000 for 15- to 19-year olds, and 13.6 per 100,000 for 20- to 24-year olds. Almost half of students with mental disorders drop out of grades 9 through 12, and of those dropouts, 73 percent are arrested within 5 years of leaving school. Children and adolescents with serious emotional disturbance often live in families with multiple risk factors and vulnerabilities, including single-parent households, poverty, and family history of substance abuse, violence, mental disorders, and/or felony convictions.

Although the challenges for families of children and adolescents with emotional and behavioral problems are great, best practices emphasize collaboration with families when coordinating care for the child. Inpatient facilities have placed an increasing emphasis on incorporating the family and other relevant support systems into a child's treatment plan.

Today, children and adolescents are likely to receive mental health services in what is referred to as "systems of care." Systems of care provide a spectrum of mental health and other necessary services organized into a coordinated network from community-based and least restrictive treatment settings and case management to intensive inpatient psychiatric care. Systems of care rely on "wraparound services." The values underlying systems of care include a focus on the individual child and family, the provision of services within the most normalized environment that meets the youth's need, the creation of partnerships with families, the use of a strength-based orientation, the creation of an environment of cultural competence, the provision of unconditional care, and the use of 24-hour institutional care at various points in time. Systems of care do not eliminate the need for inpatient level of care, but often shorten lengths of stay.

As the wraparound model appears to have evolved into a best practice in the care of the most difficult children and adolescents, it is important to note that inpatient hospitalization treatment is not a failure of the wraparound model but one of its service components. Thus, the hospital plays a key role in wraparound services.

A challenge that most states face is providing services to children with emotional and behavioral disturbances when they age out of the child system. The eligibility requirements for services in a state's adult system are many times narrower than the eligibly in the child system, so some transition age youth who have been receiving services for years are screened out, leaving them without care and treatment. A vulnerable time for a young adult is exacerbated by the states' need to restrict its eligibility pool for publically-funded mental health services.

Veterans

The veterans' hospital system was established after World War I when veterans were entitled to health care benefits and compensation. States were unable to handle the needs of this population, given that there were hundreds of thousands of U.S. solders "wounded not mortally" and uncounted numbers who suffered from neuropsychiatric conditions.

On his last day in office in 1921, President Woodrow Wilson signed into law a bill supporting $18.6 million for the establishment of hospitals for veterans. The years 1921 to 1923 saw the provision of more than 6,000 beds in new veteran's hospitals across the United States. From 1923 to 1945, there was considerable expansion of hospital beds. Following World War II, the VA was entirely reorganized. In the decades following World War II,

VA hospitals accounted for only a small percentage of the hospitalized population. Trends in bed use shifted, however, such that between 1955 and 1975, the number of VA inpatient episodes increased by 143 percent, whereas state and county mental hospitals decreased by 37 percent.

In the mid-1990s, the VA shifted its traditional focus on inpatient care and dramatically expanded its outpatient services. Between 1995 and 2001, the number of occupied annual inpatient psychiatry beds fell by 58 percent in VA hospitals, the length of stay decreased by 52 percent, and the number of veterans treated in the VA outpatient system increased by 27 percent. This shift, and the reorganization of the Veterans Health Administration (VHA) into 22 smaller, semiautonomous networks, may be contributing to the VA's apparent inability to meet the needs of veterans from combat in Afghanistan and Iraq. And that need is substantial. A nationally representative sample drawn from the National Health and Resilience in Veterans Study showed that the current and lifetime prevalence of PTSD in veterans was 4.8 and 8 percent, respectively. PTSD came from combat and noncombat situations and was associated with elevated risks of severe psychiatric disorders and suicide.

Efforts were initiated in 2006 to substantially improve the capacity of VA hospitals to meet the increasing psychiatric need of veterans. The VHA struggled to provide the necessary service. A 2010 report estimated that 18 veterans a day die by suicide and 950 suicide attempts occur each month among veterans receiving services through the VAH. Between 2010 and 1014, VA hospitals were the subject of scandals and problems, with systemic problems including turning away veterans who needed services and who then suicide, maintain secret waiting lists to create the appearance that delay in service was much shorter than it was, putting out false performance reports, covering up suicide attempts, and running facilities in which illicit drug dealing was routine practice. As one commentator said, "the VAH is a bright, shining example of misguided priorities and terrible mismanagement."

Veterans seek and/or receive treatment out of VHA settings for a number of reasons: The veteran is not service connected for a psychiatric disability, the veteran prefers treatment out of the VHA, the veteran is committed pursuant to a forensic status, or in some jurisdictions the veteran refuses voluntary admission and requires civil commitment. Thus, continuity of care can be particularly problematic as the veteran moves in and out of the VA, public, and private systems of care.

The challenges facing the VAH are significant. In 2014, there were 24 million veterans in the United States, about one-third of whom are enrolled in VA care. There are 21,000 mental health professionals to meet this need, but only 3,000 are psychiatrists. The expectation is that the Veterans Access, Choice, and Accountability Act of 2014 will improve the ratio of VA psychiatrists, non-VA psychiatrists, and other mental health clinicians

available to treat veterans with psychiatric disorders.

Individuals active in the military receive their psychiatric inpatient treatment in Department of Defense hospitals. Of major concern for this population is suicide. The U.S. Army suicide rate exceeds the civilian suicide rate with substantially elevated suicide rates following psychiatric hospitalization—an eightfold increase for 3 months following a psychiatric hospital discharge and a fivefold increase for 12 months after a psychiatric hospitalization.

Substance Abuse Comorbidities

Persons whose alcohol and/or substance use require hospital-based treatment fall into two categories: Those with so-called "straight substance abuse," meaning that there is no other mental illness that is the focus of treatment, and persons who are dually diagnosed with a mental illness and a substance use disorder both of which require active treatment.

Treatment for substance use in either population requires detoxification when such is necessary and then treatment initiating prevention of relapse and recovery. Relapse and recovery are lifelong enterprises that may or may not involve inpatient treatment. With the exception of detoxification of persons at high medical risk, all the services for persons with substance abuse or individuals dually diagnosed can be found at hospital and nonhospital loci of care.

Perhaps the major shift in the treatment of persons with dual disorders that occurred in the latter part of the 20th century was the movement from consecutive treatment to concurrent treatment. For decades, there was the belief that one could not address other psychiatric disorders until one had first addressed alcohol and/or substance abuse or dependence. Currently, efficacious treatment is considered to be treatment that addresses both at the same time.

Because the rate of alcohol and substance abuse is quite high in persons with other *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition (DSM-5), disorders, hospitals need to have expertise in substance abuse assessments and treatment at all levels of service. Many a diagnosis will be missed if the attending psychiatrist or the C-L psychiatrist is not thinking about the potential roles that alcohol and illicit drugs may be playing in the patient's presentation. The legalization of marijuana or medical marijuana in some states is presenting new dilemmas for psychiatric hospitals. It is too soon to tell how this will be dealt with in different states, and within one state, in psychiatric inpatients settings in different types of hospitals.

Geriatrics

Patients older than the age of 65 years and those deemed "functionally

geriatric" are treated in units with mixed-aged populations in some hospitals and on psychogeriatric units in others. The involuntary admission of elderly individuals varies greatly state to state, in that some state statutes prohibit the involuntary admission of persons with any form of "organic brain syndrome" as the basis for the commitment, whereas others do not.

Elderly patients have special needs that the hospital must attend to: (1) greater comorbidity with medical disorders; (2) increased likelihood of delirium; (3) possibility of their having another disorder that is confused with dementia, especially major depressive disorder; (4) greater need that all other medications they are taking be considered when the assessment is done and when psychopharmacologic recommendations are made; and (5) higher suspicion for impaired capacity to consent to treatment.

Of particular concern is the use of antipsychotic medication in elderly patients. In 2005, the FDA warned that second-generation antipsychotic medication was associated with an increased risk of death in older patients with Alzheimer disease and other age-related dementias. In 2008, the warning was expanded to include first-generation antipsychotics. Subsequently, a Harvard Medical School study of elderly, new users of antipsychotic medication showed that haloperidol had twice the mortality rate of the atypical antipsychotics, and among the atypicals, quetiapine had the lowest mortality rate. Another class of medications of concern is benzodiazepines, ubiquitous in their use in hospital settings to quiet agitated patients, including the elderly. A case-controlled study using Quebec, Canada health insurance data showed chronic use of benzodiazepines by elderly patients was associated with about a 50 percent increase in Alzheimer disease 5 to 10 years after a period of 90 or more days of use, with benzodiazepines with half-lives of greater than 20 hours carrying the greatest risk. Thus, caution should be used in the inpatient management of elderly patients by using benzodiazepines for longer than brief—up to 90 days—treatment.

Staff in any of the hospital services should consider the need for, and implement if necessary, a complete organic medical workup, formal cognitive assessments, and referral for follow-up services that are age sensitive.

Developmental Disabilities Comorbidities

Most states have deinstitutionalized persons with developmental disabilities/mental retardation at an even greater pace than their population with chronic mental illnesses. That means that former residents of developmental disability facilities who would not have been seen for hospital treatment now are.

Persons who present for services at hospitals with anywhere from mild to profound mental retardation or other developmental disabilities may be coming from home where they live with parents, from community

residential programs where they live with staff, or, in the case of mild mental retardation, from an independent living arrangement. In each instance, the person needs individualized services that are informed by his or her capacities. Especially important is the use of concepts and language the individual can comprehend.

Serving persons with developmental disabilities in the hospital has evolved in many different ways. In some hospitals, these individuals are "mainstreamed," that is, they receive services in the same areas of the hospital as all others. There are certain disadvantages to this approach, in that the essential skills to meet the needs of the population may be lacking. There are a few general hospitals that have specialty units to serve persons with developmental disabilities and even more state hospitals that do.

When the person with developmental disabilities presents for services at the hospital, it is most often due to unmanageable behavior. That is not to say that individuals with developmental disabilities do not become psychotic, depressed, or anxious; they surely do, and these conditions need psychiatric attention. However, a majority of persons dually diagnosed with a mental illness and mental retardation present because they can no longer be dealt with in the environment in which they live. Commonly the individual receives such diagnoses as impulse control disorder not otherwise specified, intermittent explosive disorder, or if adolescents or younger children, conduct disorder or oppositional defiant disorder.

Persons with developmental disabilities should leave the inpatient setting with a functional analysis of the problematic behavior that permits the outpatient setting to follow through with treatment that will facilitate habilitation. Often this entails the use of behavioral interventions. All hospital treatment of persons with developmental disabilities entails close collaboration with the team of persons providing all outpatient and residential services. Collaborative treatment is essential in the hospital's achieving successful outcomes with this population.

Medical Comorbidities

Psychiatric disorders complicated by or complicating medical disorders are a common occurrence. Psychiatric services are often required when there is a diagnostic question related to mentation or mood; when a patient is too "out of control" or his or her behavior is so "inappropriate" that management of the patient is not going well on the medical or surgical services; when the medical or surgical practitioner cannot communicate with the patient or cannot obtain informed consent; or when there is an ethical question about course of treatment.

There is perhaps no better example of the psychiatric hospital's role in treating persons with medical illness than HIV infection. There is a bidirectional relationship between severe mental illness and HIV infection; persons with psychiatric disorders are at higher risk of acquiring HIV, and

persons who are HIV positive have increased rates of psychiatric illness. The interactions between disorders and treatments are myriad. Psychiatric disorders impair judgment and heighten the risk of contracting HIV (e.g., mania in bipolar disorder) or lead more directly to increased risk (e.g., intravenous drug use). HIV induces, both psychologically and by direct effect on the brain, anxiety, depression, psychosis, and dementia. Antiretroviral medications can induce psychosis. Drug–drug interactions—drugs prescribed for HIV status, psychiatric status, and medical disorders due to HIV status—make treating the comorbid psychiatric disorders quite complex. Being HIV positive per se, the symptomatology of HIV, the psychiatric disorders that predated HIV-positive status, and the psychiatric disorders resultant from being HIV positive may all diminish compliance with both retrovirals and psychotropics.

The hospital plays a unique role in meeting the clinical needs of persons who are HIV positive. Although the general hospital is the mainstay of treatment for this population, the state hospital is not infrequently called on to provide care and treatment to symptomatic and nonsymptomatic HIV-positive individuals.

Although HIV infection has been used as an example, a comparable presentation could be made for scores of medical disorders. Some, such as systemic lupus erythematosus, temporal lobe epilepsy, and space-occupying brain lesions, are dramatic examples. On a more frequent basis, hospital-based psychiatric and medical services draw together to treat such conditions as cancer, renal failure, and trauma.

Psychiatry has known for over 150 years that individuals with SMI die younger than persons without these disorders. And the reason for this is still unknown. It is known that medical comorbidities can be significantly exacerbated by psychiatric medications. This is no better illustrated than by metabolic syndrome in persons taking some of the second-generation antipsychotic medications. It behooves the inpatient psychiatrist to work closely with his medical colleagues in the hospital, and others staff like dieticians and occupational therapists to create and monitor wellness programs for his patients.

A new service that has started to appear in a few state hospitals is hospice care. These services are provided to the state hospital patient who have spent some time in the hospital and without the patient needing to be transferred to another facility or even to change units within the hospital. The advantages to the patient are many, but the fundamental advantage in the patient's ability to remain in an environment he knows with staff he is comfortable with. It is important to have the patient's informed consent, or the guardian's is a guardian of the person is in place, for all services that will not be provided and those that will. In hospitals where this is done to the extent of full hospice services on site, the patients dies in his own room (it is generally not feasible to do this if the patient does not have a single

room) and this is processed before and after death with the other patients on the unit. Shortly thereafter it is beneficial to have a nondenominational memorial service in the hospital's chapel.

Deaf Individuals

In treatment as orally based as is psychiatric treatment, the clinical approach to persons who are deaf presents a unique set of challenges. Although in many states, deaf individuals who need psychiatric treatment secure that treatment within the hospital alongside all others with the addition of translation services, some states have a statewide deaf unit at one of its hospitals. Here, persons who are deaf can interact with staff without the imposition of a translator and can be in an environment that is sensitive to the deaf culture. There are a very few private hospitals that offer this service.

Deaf individuals present with the full range of psychiatric disorders. Being deaf does not protect one from DSM-5 disorders. Nor does it protect one from comorbidities such as mental illness/developmental disabilities and/or mental illness/substance abuse.

The range of the hospital's roles in treating deaf persons is no different than for hearing persons. To fulfill that role, hospital staff needs to be aware of, and sensitive to, the deaf culture. For this population, the state hospital-based deaf unit may be the least restrictive, most-integrated setting when inpatient treatment is required. In states with deaf units, that also require civil admissions only come from other hospitals, this restriction is waived for the deaf population who can be direct admits from multiple loci of care, including general hospital EDs.

Forensic Patients

Persons admitted to the hospital pursuant to the commitment of a crime are the states' responsibility and one that the state generally fulfills directly at state hospitals. There are a few exceptions when the state contracts out for these services.

Forensic patients are in the hospital basically for one of three reasons: (1) pretrial, for a determination of competency to stand trial and/or of criminal responsibility for the crime; (2) after a finding of being incompetent to stand trial, guilty but mentally ill, or not guilty by reason of insanity; or (3) jail or prison could not manage the inmate due to his behaviors as a result of psychopathology. Restoration of competency to stand trial is a uniquely forensic task, and the hospital needs to have specialty functions to accomplish this. Much of the rest of the clinical treatment of the forensic population is the same as that of the nonforensic population. Security on forensic units is generally much tighter than in other places at the hospital.

Often the role that the hospital is required to fulfill for many forensic patients is one that it does not play very often for other patients—housing forensic patients for years to decades beyond the time the patient needs hospital-level care. The degree to which this occurs is heavily dependent on the state commitment statutes and procedural handling of its forensic patients.

Although there is a clear dichotomy between forensic and nonforensic patients based on the state statute under which the patient was admitted, the clinical variation and treatment need to show no such distinction. Whether a patient is civil or forensic, especially for minor crimes, has more to do with the circumstances of the patient's initial encounter with the police than it does with his or her clinical presentation. The boundary is also blurred these days because of states admitting more persons at the end of a criminal sentence, especially those convicted of sexual offenses.

The hospital has a multiple roles with the forensic patient: (1) to meet the person's needs as a psychiatric patient, (2) to meet the state's needs to proceed with due process in trials, and (3) to maintain the public's safety in the face of those who break or are charged with breaking state or federal law. Hospitals are considerably more comfortable with the first role than with the other two.

With forensic patients, the judicial or quasijudicial systems control the admittance and discharge of patients. In many states, this contributes to both overcrowding at hospitals and to an ever increasing percentage of forensic patients in the hospital's census. It is an irony of current times that jails and prisons feel overwhelmed by inmates sheriffs believe should be psychiatric inpatients and psychiatric hospitals feel overwhelmed by inpatients psychiatrists asses to need jails and prisons.

OTHER HOSPITAL ROLES

This section has focused on the inpatient functions of the hospital, which comprise the service that is unique to it. The hospital fills a number of other roles, many of which are also performed in some fashion by entities that are not hospitals.

Emergency Services

Psychiatric treatment is provided on an emergency basis in hospital EDs (see Fig. 58.3–1). The arrangements under which this is done vary greatly. In community hospitals, the person presenting with a psychiatric problem is generally treated in the same fashion as all individuals—the same waiting area, the same examining rooms, sometimes the same physicians. In larger hospitals, all of these functions are often segregated into a designated area of the ED where patients are seen by nonphysician evaluators who consult with psychiatrists who may or may not see the patient. Like many other

hospital functions, crisis evaluation services have moved out of the hospital and to freestanding crisis services in some jurisdictions. Not uncommon is a hybrid model in which the hospital ED plays host to an outside crisis service. Here, the hospital performs medical screening evaluations.

The ED functions may include psychiatric holding beds, that is, the ability to keep a person in the ED for up to 23 hours, thus avoiding an admission. The ED may include crisis beds, a service allowing a short number of days that again avoids an admission to the psychiatric unit. Crisis beds are another service the hospital shares with nonhospital settings such as CMHCs, which offer freestanding crisis beds.

Psychiatric patients "stuck" in ED's awaiting transfer to an inpatient bed has become a nexus for consternation over inadequate inpatient beds. These patients, often called "boarders," can wait hours, days, weeks, or in some cases over a month during which time little is done for them other than room and board in the ED. In August 2014, the Supreme Court of Washington State found, In re the Detention of D.W., et al., found it unlawful, under Washington's involuntary treatment act, to board patients in ED's due to overcrowding of appropriate evaluation and treatment facilities.

Only about half the states have processes in place to systematically identify open beds for psychiatric patients who need admission. Even those states that can effectively search for bed availability are stuck with the inadequate number of psychiatric beds nationwide. A 2012, study by the Treatment Advocacy Center, "No Room at the Inn" demonstrated that not one state had sufficient numbers of public hospital psychiatric beds.

Step Down/Admission Diversions Services

Hospitals offer a range of services that function between inpatient and outpatient levels of care. Such services include day hospital, partial hospital, and night hospital. In treatment arrangements that are during the day, the individual attends a full day of active psychiatric treatment but lives outside of the hospital, for example, in his or her home, with family, or in a community residential program. In the night hospital program, the person leaves the hospital for school or work during the day but spends overnights at the hospital. All of these programs are also used as attempts at admission diversion, that is, to avoid inpatient hospitalization for this episode of illness. Step down programs and admission diversion programs are also offered by nonhospital entities, with the latter also offering respite beds. These have a longer length of stay than crisis beds and are not generally provided by hospitals.

Partial hospital programs, once enthusiastically endorsed by insurers as less expensive alternatives to inpatient treatment for their use as both alternatives to admission and step downs to decrease length of stay, have lost favor. There is a dwindling supply of such programs. Just between

2000 and 2001, the number of facility members of the National Association of Psychiatric Health Systems (NAPHS) offering partial programs decreased from 83 to 67 percent. Between 2003 and 2013, the number of hospitals offering partial programs decreased by 20 percent (from 508 to 410). There is expected to be further loss of partial programs if CMS decreases their reimbursement rate, as it is proposing to do.

Outpatient

Outpatient services are offered in more settings than any other service. Besides the hospital, outpatient services are found in CMHCs; group settings, either integrated with primary medical services or at their own sites; private practices; family services agencies; and hosted in locations such as shelters, specialty clinics, mental health residential programs, and mobile units. Payment for these services is most often by third-party insurers—private or government—with a copay by the patient. In larger metropolitan areas particularly, some outpatient providers have established practices that take no insurance.

The major challenge facing outpatient psychiatry at hospital locations is the high cost of delivering the treatment within hospital settings and the poor reimbursement rates. It has reached the point where expanding business in the outpatient department (OPD) is the last objective of the hospital because each new patient adds to the deficit.

On a more practical note, hospitals are seen as unfriendly places to the individual seeking outpatient psychiatric services. Parking is often difficult, far from the OPD, and expensive. Negotiating the corridors of the hospital is a challenge that many would prefer to avoid. In addition, the patient may have to go through a registration process that simply adds another layer of annoyance.

The prognosis for hospital-based outpatient services is not good. Often the victim of cherry picking by others (the most difficult patients are directed to the hospital OPD), the hospital psychiatric OPD is on the hit list of many hospital chief executive officers. One remedy may be hospital-based integrated clinics where outpatient psychiatry and medicine are collocated. Such attempts have been around for 40 years, but the momentum for such integration is accelerating.

Consultation-Liaison

Although many hospital executives would gladly replace the psychiatric inpatient service with more profitable medical-surgical beds and make much better use of the space devoted to the psychiatry OPD (which is often banished from the main hospital buildings anyway), the C-L function is usually highly valued. In part, that is because the C-L service directly affects the hospitals' medical and surgical services by evaluating patient's

capacity to consent, assessing cases of delirium and psychosis, and assisting with difficult dispositions. The C-L service can be invaluable in making differential diagnoses that will fundamentally alter treatment and disposition. Some have argued that it is because a hospital could not effectively function without C-L psychiatry that much of the rest of hospital-based psychiatry is tolerated. This may be a somewhat cynical overstatement, but it certainly has some elements of truth. A well-functioning C-L service is a fountain of goodwill for the psychiatry department.

Community Education and Patient/Family Support

One role of the hospital that is largely one of goodwill is the hospital's role as host to a remarkable array of self-help and support groups. Those that fall under the penumbra of mental health and substance abuse activities are frequent in hospitals; such groups are Alcoholics Anonymous, Narcotics Anonymous, and the National Alliance on Mental Illness. Less directly related to mental illness but using principles of support are groups centered on all manner of chronic conditions and issues, such as breast cancer, colostomy, and scleroderma. There are also groups for individuals in specific life circumstances, such as parents without partners and parents who lost a child to cancer.

Hospitals also play a community education role. Members of the psychiatry staff may provide substantive lectures geared to the lay public. These are different, yet complement the hospital's psychiatry grand rounds, which themselves are usually open to the public. Another form of education is informing the public about the services the hospital has to offer.

FUTURE DIRECTIONS

The challenges that face hospitals in the second decade of the 21st century are not new. Prior to 1950, virtually every problem and many of the solutions that define contemporary hospital psychiatry were noted in the professional literature. One need only read such sources as the *American Journal of Insanity*, *Alienist and Neurologist*, or state hospital annual reports to find them all.

Homelessness, the overpopulation of jails and prisons with persons with mental illness, the eschewal of persons with mental illness from any locale because that jurisdiction does not want financial responsibility for the person, an inadequate focus on treatment and overreliance on management, and a propensity to not spend tax money on persons with mental illness are all conditions Dorothea Dix found in the 1840s when she began her lifelong work of reform. Samuel Woodward, APA's first president and the first superintendent of what is now the Worcester Recovery Center and Hospital (Figs. 58.3–4 and 58.3–5) would have recognized them all.

The loss of the person in meeting the acute needs of the patient, the reliance on modern technologies for diagnoses that has diminished one-to-one time with patients, and the general sense of the dehumanization of psychiatry and medicine in general were noted by William Allison White in 1925. White remarked in his APA presidential address that "the introduction of these highly technical methods has been accomplished by a corresponding lack in development, at least, if not actual loss, of any adequate dealing with the patient as an individual or as a human personality."

Even hospitals bending to the dictates of payers to minimize costs is not new. More than 125 years ago Kirkbride lamented that "it will be a sad day for these institutions, and still sadder for the patients in them, when the rivalry of hospitals and their offices shall be, rather to discover for how little inmates can be kept, than to secure what is best, and most thoroughly promotes the great objects for which they were established."

State hospitals are disappearing by design. Whether that is the result of thoughtful or ill-conceived planning is a matter of debate. General hospital psychiatry programs, however, are disappearing through a thoughtless process of survival of the fittest.

Reimbursement and operating margins for psychiatric services lag behind those of other hospital-based medical services. The diminishing financial returns from psychiatry have led hospital management consultants to recommend converting psychiatric beds to higher-margin medical and surgical beds. Even to remain fiscally viable now, psychiatric hospital inpatient revenues need to cross-subsidize emergency, consultation, and outpatient services.

Today, psychiatrists generally fail to acknowledge those times when the psychiatric hospital is a forerunner or role model for the rest of psychiatry if not the rest of medicine. One of the foci of contemporary medicine is the integration of medicine and psychiatry, or medical and mental health services. As practitioners in psychiatry and in medicine focus on methodologies to integrate care and treatment, they need look no further than the public psychiatric hospital, which has been an integrated care organization for over 100 years, and still is a fine example of integrated care and treatment. Outpatient integration, it seems, is mimicking business as usually at the state hospital but without its walls.

Attributes of a state hospital that contribute to its high degree of integration include:

- Psychiatric and medical services collocated
- An integrated psychiatric and medical treatment plan
- A single, shared medical record, now an electronic medical record in many institutions
- Ease of communication among a patient's physicians and between those

physicians and other disciplines

- Case or care management integrated into treatment
- Professional and paraprofessional staff assisting patients to follow through with psychiatric and medical recommendations in an integrated housing and care model
- On-site wellness programs, for example, dietary, exercise, organized leisure
- Other health promotion activities, for example, weight loss, smoking cessation, treatment of substance abuse
- Transportation to any off-site medical appointments
- Workforce development, for example, student placements, pre- and postdoctoral internships, residency rotations
- Paid time off for staff education
- On-site dental, podiatry, optometry, radiology, phlebotomy, and neuropsychiatry services
- Peer supports
- Referrals to community and social supports
- Services for all independent of the ability to pay
- Funding supports not only direct service but time for collaboration
- Unitary billing system
- Transition services for relocation provided and funded
- Interfaces with vocational services
- Provision of in vivo training, for example, bus riding, banking, using the library
- On-site hospice care with direct care staff the patient is already familiar with the state psychiatric hospital, as an integrated care system, should have outcome measures that reflect this. The current, common outcome measures fail to take this into account and currently include number of assaults patient on patient and patient on staff; number of serious assaults; number of episodes and time in any form of seclusion or restraint (manual, mechanical, chemical); two or more antipsychotic medications at discharge; five or more psychiatric medications prescribed; time out of hospital on unsupervised passes; and timeliness of all manner of documentation. Many of these measures are open to interpretation, especially what a facility chooses to label as an episode of restraint. What about improvements in health and wellness as outcome measures?

Virtually no knowledgeable commentator claims that the system of care and treatment for mental illnesses can be sustained without the roles played by psychiatric hospitals. At the same time, too few informed people are working on methods to interdigitate the hospital's roles into the fabric of an integrated system of care, one that looks less like Figure 58.3–1 and

more like one that achieves effective outcomes, through a person-centered approach, with fiscal management that does not compromise the very care and treatment psychiatrists are attempting to deliver.

REFERENCES

Andrews JB. The distribution and care of the insane in the United States. *Am J Insanity*. 1887;44:192–211.

Butler JS. The individualized treatment of the insane. *Alienist Neurol*. 1886;7:435.

Center for Substance Abuse Treatment. *Screening Assessment, and Treatment Planning for Persons With Co-Occurring Disorders. Overview Paper 2 (DHHS Publication No. (SMA) 07–4164)*. Rockville, MD: Substance Abuse and Mental Health Services Administration, and Center for Mental Health Services; 2007.

Committee on Crossing the Quality Chasm. *Adaptation to Mental Health and Addictive Disorders: Improving the Quality of Health Care for Mental and Substance Use Conditions*. Washington, DC: National Academies Press; 2006.

Deutsch A. *The Mentally Ill in America*. New York: Columbia University Press; 1949.

*Dorwart RA, Schlesinger M. Privatization of psychiatric services. *Am J Psychiatry*. 1988;145:543–553.

Fisher WH, Geller JL, Pandiani JA. The changing role of the state psychiatric hospital. *Health Aff*. 2009;28(3):676–684.

Geller JL. Issues in American psychiatry reflected in remarks of APA presidents, 1844–1994. *Hosp Commun Psychiatry*. 1994;45:993–1004.

Geller JL. Excluding institutions for mental diseases from federal reimbursement for services: strategy or tragedy. *Psychiatr Serv*. 2000;51:1397–1403.

*Geller JL. The last half-century of psychiatric services as reflected in *Psychiatric Services*. *Psychiatr Serv*. 2000;51:41–67.

*Geller JL. A history of private psychiatric hospitals in the USA: from start to almost finished. *Psychiatr Q*. 2006;77:1–41.

Geller JL, Biebel K. The premature demise of public child and adolescent inpatient psychiatric beds. Part I: overview and current conditions. *Psychiatr Q*. 2006;77:251–271.

Geller JL, Harris M. *Women of the Asylum*. New York: Doubleday; 1994.

Greenblatt M, York RH, Brown EL. *From Custodial to Therapeutic Care in Mental Hospitals*. New York: Russell Sage Foundation; 1955.

*Grob GN. *Mental Illness and American Society, 1875–1990*. Princeton, NJ: University Press; 1983.

Haskins BG. Servicing deaf adult psychiatric inpatients. *Psychiatr Serv*. 2004;55:439–441.

Honberg R, Diehl S, Kimball A, Gruttadaro D, Fitzpatrick M. *State Mental Health Cuts: A National Crisis*. Arlington, VA: National Alliance on Mental Illness; 2011.

*Hurd HM, ed. *The Institutional Care of the Insane in the United States and Canada*. Baltimore, MD: Johns Hopkins University Press; 1916.

Liptzin B, Gottlieb GL, Summergrad P. The future of psychiatric services in general hospitals. *Am J Psychiatry*. 2007;164:1468–1472.

McLoughlin KA, Geller JL. Interdisciplinary treatment planning in inpatient settings: from myth

to model. *Psychiatr Q*. 2010;81:263–277.

Menninger RC, Nemiah JC, eds. *American Psychiatry After World War II, 1944–1994*. Washington, DC: American Psychiatric Press; 2000.

National Institute of Mental Health. Transforming the understanding and treatment of mental illnesses. Available at http://www.nimh.nih.gov/health/statistics/prevalence/use-of-mental-health-services-and-treatment-among-adults.shtml. Accessed December 21, 2014.

Parks J, Radke AQ, eds. *The Vital Role of State Psychiatric Hospitals*. Alexandria, VA: National Association of State Mental Health Program Directors; 2014.

Report of the Surgeon General's Conference on Children's Mental Health: A National Action Agenda. Washington, DC: Department of Health and Human Services; 2000.

Sharfstein SS, Dickerson FB. Hospital psychiatry for the twenty-first century. *Health Aff (Millwood)*. 2009;28:685–688.

Sledge WH, Tebes J, Wolff N, Helminiak TW. Day hospital/crisis respite care versus inpatient care. Part II: service utilization and costs. *Am J Psychiatry*. 1996;153:1074–1083.

State Psychiatric Hospitals 2006. Alexandria, VA: National Association of State Mental Health Program Directors Research Institute; 2006.

Stevens R. Can the government govern? Lessons from the formation of the veteran's administration. *J Health Politics Policy Law*. 1991;16:281–305.

Torrey EF, Fuller DA, Geller J, Jacobs C, Ragosta K. *No Room at the Inn: Trends and Consequences of Closing Public Psychiatric Hospitals*. Arlington, VA: Treatment Advocacy Center; 2012.

Yazgan IC, Greenwald BS, Kremen NJ, Strach J, Kramer-Ginsberg E. Geriatric psychiatry versus general psychiatry inpatient treatment of the elderly. *Am J Psychiatry*. 2004;161:352–355.

▲ 58.4 Mental Health Services Research

CASSIS L. HENRY, M.D., AND LISA B. DIXON, M.D., PH.D.

The imperative for mental health services research arises from some basic and alarming facts. Over half of the persons with mental disorders receive no mental health treatment. For those who do, the vast majority—approximately 75 percent—do not receive treatment that can be judged of good quality by evidence-based standards. Those outside of the health services research field, such as clinicians and other stakeholders, may not see how mental health services research can improve the substandard quality of mental health care. In fact, they may not even understand what health services research is. The Association for Health Services Research broadly defines it as:

> The multidisciplinary field of scientific investigation that studies how social factors, financing systems, organizational structures and processes, health technologies, and personal behaviors affect access to health care, the quality and cost of health care, and ultimately our health and well-being. Its research domains are individuals, families, organizations, institutions, communities, and populations.

This chapter reviews the specific scope of mental health services research and illustrates through several case studies how such research addresses critical questions that arise in the routine provision of services, laying the groundwork for improving the quality of mental health care.

THE SCOPE OF MENTAL HEALTH SERVICES RESEARCH

Clinicians seek to address identified clinical problems, disorders, and diseases. Clinicians turn to biomedical research to answer two basic questions: (1) What is known about the nature of this problem? (2) What can be done to remedy it or alleviate the suffering associated with it? Research answers to the first question derive from *basic sciences research*, whether biological (e.g., What neural pathways are affected in bipolar disorder?), behavioral (e.g., How do organisms respond to different types of reinforcement schedules?), psychological (e.g., How do negative automatic thoughts affect mood?), or sociological (e.g., What are the pathways to homelessness?). Research answers to the second basic question arise from *clinical research* involving well-controlled treatment trials designed to determine efficacy and safety, as well as from the more or less systematic trial and error of clinical experience. Broadly, the basic and clinical sciences deal with knowledge needed prior to the initiation of routine care. In other words: What is the health problem, and what treatments might relieve this problem? *Health services research* questions, however, arise once clinical services get underway. Clinicians have

extensive exposure to the basic and clinical sciences during their training but may have little or no exposure to, and hence little understanding of or appreciation for, the value of health services research.

What kinds of questions do mental health services researchers ask? These questions can be divided into two general categories—*clinical services research* and *service systems research*. Like clinical research, *clinical services research* focuses on "front-line" clinical issues and uses individual service recipients (patients, families) as the participants and units of study. Examples of questions asked by *mental health clinical services research* include the following: How effective are antidepressant agents in producing illness remission among typical patients with major depression seen in routine office practices? What considerations influence individuals to seek treatment and to agree to complete a course of prescribed treatment? What range of treatments for persons with schizophrenia are "evidence-based?" Do persons with substance abuse and mental illness have better outcomes when they receive care from integrated addiction mental health teams or programs with parallel services? What are the effects of a routine depression screening measure in primary care on detecting and treating major depression in such settings?

Mental health services system research examines questions at the service system level and typically uses aggregated service components (providers, treatment agencies, state mental health systems, etc.) as the units of study. Examples of questions asked by service systems researchers include the following: How many persons with severe mental illness who are being treated in a public mental health system have access to routine primary care? Does this differ between two states with different Medicaid policies? What are the various sites in a city where homeless persons with severe mental illnesses receive services, and are these coordinated? What are the effects of a decision by Medicaid to cap the total number of prescribed medications that recipients can receive? Do restricted formularies for antipsychotic medications reduce overall costs? Do they affect quality? Do local mental health authorities improve coordination, continuity, and quality of care for high-need patients?

Given this broad scope of mental health services research, it is also helpful to note what it does not include (Fig. 58.4–1). The distinction has already been mentioned between basic and clinical research on one hand and mental health services research on the other. The interface between clinical research and clinical services research begins at the transition between treatment efficacy and treatment effectiveness studies. *Efficacy studies* examine the outcomes of treatments within narrowly defined samples of patients who are carefully selected to control for variables that might interfere with determining a treatment effect (e.g., the exclusion of patients with comorbid substance abuse or medical disorders). Efficacy study designs attempt to maximize internal validity. *Effectiveness studies*

examine treatment outcomes among more inclusive, diverse, and "typical" clinical populations. External validity assumes a more important role. Generally, effectiveness studies only occur after treatment efficacy has been established. Mental health services research, while somewhat distanced from the efficacy studies with which many clinicians are familiar, is still *research*, using standardized scientific procedures to answer specific research questions to generate new knowledge that is generalizable beyond the system being evaluated. It should be distinguished methodologically from routine activities undertaken by service systems and service agencies to monitor policy implementation, program or system administration, or quality assurance.

Figure 58.4–1 illustrates the distinction between basic and clinical research and health services research, the flow between these larger research domains, and depicts various activities conducted within health services research. For example, one way in which clinical practice is affected by services research occurs as the results of clinical efficacy trials and effectiveness trials are applied to routine practice—a form of "translational research." Given that a primary goal of basic and clinical research is to inform clinical practice, when a body of scientific knowledge has accumulated, it should be adopted into clinical practice. However, this translation does not happen as it should; the gaps between existing scientific knowledge about treatment efficacy and effectiveness on one hand and routine clinical practice on the other have been referred to collectively as the "quality chasm" in medical care, encompassing both access and quality.

A major question for mental health services research then, is how to promote the adoption of scientifically validated treatments (evidence-based practices) and, by doing so, bridge the "quality chasm." One response has been the development of evidence-based practice guidelines, variously referred to as evidence-based practices, evidence-based treatment recommendations, or "best practices." The identification of evidence-based best practices is critical to examining current practice patterns, evaluating their quality, identifying areas in need of improvement, and ultimately identifying areas in which current treatments fall short. The methodology for developing such evidence-based practice guidelines emphasizes the results of efficacy and effectiveness research, augmented by expert opinion when research evidence is underdeveloped in essential areas.

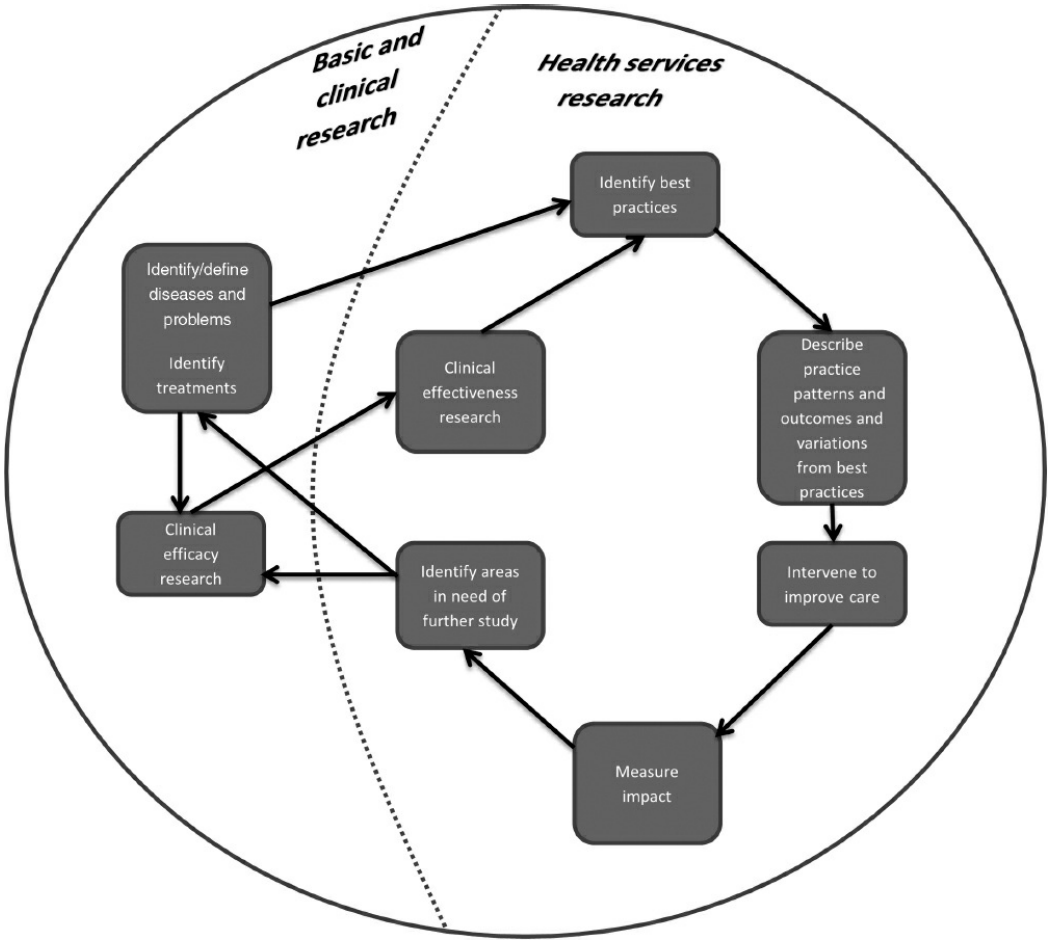

FIGURE 58.4–1. The relationship between basic and clinical research and health services research.

Once a set of research-based best practices has been identified, health services researchers can ask a variety of questions about how current practices compare to best practices, including variations in practices across treatment settings, practitioners, and patient subtypes, as well as the relationship to such factors as insurance coverage, health policies and regulations, and cultural diversity. Such research can also examine patterns in outcomes of care across different kinds of services and relate these to patterns in the implementation of evidence-based practices. Such studies are essentially descriptive, but they allow comparisons of actual practices to standards of care.

Identifying gaps between actual and best practices can lead to developing health services interventions to improve care. Such interventions fall into one of two general categories. *Supply side* interventions focus on individual providers or service programs. These interventions seek to change the care offered in some way. Some examples of supply side interventions include provider education and training programs, systematic feedback to providers regarding their current practices, consultations to agencies about implementation of best practices, financial incentives to alter practices, and so on. *Demand side*

interventions focus on the demand for services and are directed at changing patient behavior. Examples of such interventions include educating patients about best practices, training patients in how to advocate for themselves in health settings, and providing service vouchers to patients to enable them to select services that they identify as most helpful.

Both supply and demand side interventions can occur at distinct levels involving individual providers and patients, local service agencies, local or regional service systems, and local, regional, or national policies. *Shared decision making* is a strategy that involves both supply side and demand side approaches at the level of the individual patient–provider interaction. *Policy-level interventions* occur at the level of a system of care with the intention of changing patterns of service delivery—for example, changing the services covered by insurance, creating local mental health authorities to improve the coordination and outcomes of treatment in a given locale, mandating the integration of mental health and substance abuse services, or offering Medicaid care through a coordinated and managed "health home."

Such interventions, however, are not in themselves health services research. Rather, health services research evaluates the effects of such interventions on a range of outcomes, including patient outcomes, changes in patterns of care, and costs. In health services research, the interventions must be clearly defined, the fidelity of their implementation closely monitored, and the outcomes measured in a standardized form. Research designs to minimize threats to internal and external validity—for example, randomized clinical trials—are critical, as in clinical research. The research questions, although based on the specific interventions, focus on knowledge that can be generalized to other settings. As in other areas of research, findings from any phase of health services research optimally feed back to raise new research questions.

Recent work has focused on methodologies for implementation research, a burgeoning domain of health services research that attempts to understand how effective interventions fail to lead to improved patient care. Studies of implementation problems in carrying out interventions at the level of systems or individual clinical care depend upon various conceptual frameworks but generally acknowledge the need to study factors including (1) the structural, political, and cultural context as well as internal organization of the settings in which the intervention will take place, (2) characteristics of the individuals involved, (3) the process of implementation of the intervention, and (4) the nature of the intervention itself. Such research takes a pragmatic approach, seeking to understand how real-world conditions affect the implementation of the evidence-based practices for which clinical and other research has established evidence of efficacy and effectiveness.

The following "case reports" concretely illustrate the application of mental health services research to critical questions that arise in the routine provision of services.

Case 1: Treatment of Depression in Primary Care Settings

Depression is one of the most common mental disorders and accounts for a staggering amount of disability worldwide. Over the last several years, there have been substantial improvements in methods for diagnosing depression and a broad range of efficacious treatments for depression exist, including antidepressant medications and evidence-based forms of psychotherapy. A number of evidence-based practice guidelines have been developed, so the opportunities to effectively diagnose and treat depression are substantial.

Given the state of diagnostic and treatment methods for depression, how successful is The United States in treating depression? This health services research question raises some basic epidemiological queries. How many people suffer from depression? Do they seek treatment and where? What treatments do they receive? The most comprehensive estimates of the prevalence of mental disorders derive from the Epidemiologic Catchment Area (ECA) Study and the National Comorbidity Survey (NCS), which estimated the 1-year prevalence of major depressive disorders at 6 to 7 percent. The vast majority of depressed individuals (about 80 percent) contact a health care provider over a 1-year period, but only about 25 percent receive any appropriate treatment for depression. Most patients (about 80 percent) who visit a health care provider for depression see a primary care provider; only about 20 percent of patients with depression visiting a health care provider see a mental health specialist. Among such patients with a treatment contact for depression, 80 percent of those treated in primary care received no appropriate treatment for depression (pharmacotherapy or counseling). The opposite prevailed in mental health treatment settings, where most patients (nearly 90 percent) received an appropriate form of depression treatment. Such data illustrate the challenges in providing quality care for individuals with depression. Most persons receive no depression treatment (despite some contact with a health care provider). Most of those who do receive some treatment receive inadequate care. The predominant opportunity for the treatment of depression is in primary care. How can clinicians and administrators address these problems in access to and quality of depression care? How can mental health services research assist in these efforts?

Health services researchers have examined interventions targeting providers (the supply side of care), individuals affected by depression (the demand side), and policies. By far the majority of health services research in this area has focused on the supply side through attempts to change provider management of depression, especially primary care providers,

who see most of the patients with depression. Such research has approached the under-treatment of depression in primary care with different interventions of interest. For example, a multicomponent health care system-level intervention called "collaborative care" structures depression treatment in primary care settings by using case managers who educate and follow patients, track outcomes, and adjust treatments for patients who are not responding, consultation and decision support for primary care providers by mental health specialists (psychiatrists or psychologists) and the use of technology-based resources to support the system. Numerous randomized controlled trials on collaborative care for depression have demonstrated meaningful effects on depression symptom improvement, adherence, response to treatment, remission and recovery from depression, as well as satisfaction with care. However, only a small effect has been established on patient quality of life, including functional status.

A second health services research approach has studied the implementation of quality improvement (QI) programs for managed primary care settings with the goal of increasing the percentage of clinically depressed patients receiving appropriate depression care by national guidelines. The QI intervention consists of four components: institutional commitment of in-kind resources, training of local clinical leaders to implement interventions (cognitive behavior therapy, pharmacotherapy), training of local treatment staff, and patient identification. In one 12-month study, depressed patients in QI intervention clinics were more likely to receive counseling or medications, visit a mental health specialist, have a remission from the depressive episode, and retain employment than depressed patients in "control" clinics. Researchers have evaluated another approach in "behavioral health labs" that improve screening and detection of depression in primary care settings. All three of the above depression treatment models demonstrate the potential for improving depression care in primary care settings through "supply side" interventions. Major challenges remain, however, in promoting the widespread adoption of these approaches.

Far less research on depression care has been conducted on demand side interventions with patients, although one can certainly argue that even the predominantly supply side interventions just reviewed increase demand by educating patients about their symptoms of depression and about the value of treatment. Negative attitudes about depression and its treatment among the general population can deter persons from seeking help. Such negative attitudes arise from personal health beliefs, social norms and stigma, as well as from prior negative experiences with help-seeking. True demand side interventions are aimed at directly increasing the likelihood that individuals with depression will seek and participate in effective treatment. Such demand side interventions focus on awareness of

the signs and symptoms of depression, self-assessment for these signs and symptoms, the effectiveness of treatment, and skills in seeking help. Some common examples of such efforts include National Depression Screening Day campaigns, free depression screenings at employee health fairs, free pamphlets in physician offices and clinics, anti-stigma campaigns through the media, and direct-marketing media campaigns on depression, such as those currently put forward by pharmaceutical companies. Health services research that examines the effects of such programs remains sparse, and little is known about the effects of these types of demand side interventions.

A third domain of intervention to improve detection and treatment for depression in primary care is that of policy, system organization, and financing approaches. A prime example of this type of intervention is the current effort to promote health insurance parity for the treatment of mental disorders. Inadequate insurance coverage for the treatment of a mental disorder poses major barriers to treatment. The federal Affordable Care Act (ACA) enacted in 2010 expands coverage of mental health and substance use disorders by including benefits for such disorders in the Essential Health Benefits required of all new small group and individual market insurance plans and by applying federal parity protections more broadly, requiring insurance plans to cover treatment for mental health and substance use disorders at parity with medical and surgical benefits; a projected 62.5 million individuals will benefit by access to these federal parity protections. In addition, the ACA provides states a new option for organizing care in their Medicaid programs: "health home" services for Medicaid beneficiaries with severe chronic conditions who can benefit from enhanced coordination and management of physical and behavioral health care as well as long-term community-based services and supports. The "health home" model aims to improve the quality of health care and the patient experience while providing more cost-effective care. Health services research on service organization and financing seeks to answer critical questions about the effects of policy interventions such as parity laws and "health homes" on issues including the quality and experience of behavioral health care—such as depression care—delivered through reorganized systems.

Case 2: Services for Families of Persons with Schizophrenia

During the second half of the 20th century, clinical research on families in the treatment of individuals with schizophrenia dispelled early assertions that disturbed family relationships play a major role in the etiology of the disorder. Instead, research led to the recognition that families provide major support networks for most persons with this disorder, and that they need education and support to enable them to engage as partners in care. The potential advantages of engaging and educating families optimally was

suggested by early studies linking high levels of "expressed emotion" among families to greater likelihood of patient relapse after hospital discharge. Research spanning from the 1970s to the 1990s clearly established family psychoeducation (FPE) as an evidence-based practice in the care of individuals with schizophrenia. The Schizophrenia Patient Outcomes Research Team included FPE as a treatment recommendation, stipulating that patients who have ongoing contact with their families should be offered a family psychosocial intervention that lasts at least 9 months and includes education, support, and some training in problem-solving skills. Other statements of best practices also include a recommendation for FPE. FPE was found to reduce relapse rates, improve medication adherence, reduce symptoms, and improve functional outcomes for individuals with schizophrenia. Specifically, more than 30 randomized clinical trials showed that relapse rates for individuals receiving FPE were 25 to 50 percent lower than those observed for individuals receiving standard individual services.

In line with the health services research framework (see Fig. 58.4–1), investigators have examined the prevailing patterns of practices as they relate to families with regard to both access to any services and more specifically access to evidence-based practices consistent with the FPE model. A variety of studies conducted in community settings and in Veterans Administration facilities has shown that not only is FPE rarely offered, but in fact a far lower standard of family contact with clinicians occurs in the majority of cases. Chart reviews, clinician interviews, and consumer interviews all have produced the same conclusions. Thus, major problems in the access to and quality of family-oriented services exist.

Following the health services quality-of-care framework, researchers have undertaken efforts to improve practices by disseminating FPE. Supply side approaches have been the most common. Numerous initiatives to disseminate FPE have been attempted throughout the country. The New York State Office of Mental Health (OMH) mounted an impressive effort to implement programs at 35 different sites throughout the state and provided extensive training. Although numerous family groups were started and the potential benefit was large for those families and patients who participated, OMH decided not to continue the project because only a small number of families and patients throughout the state actually received the program. The Substance Abuse and Mental Health Services Administration included FPE as part of an effort to disseminate evidence-based practices by creating toolkits that include informational resources for all stakeholders, a workbook for practitioners, and a fidelity scale. The implementation phase of this project resulted in limited selection of FPE as a practice to disseminate and limited penetration of FPE in those sites that selected FPE.

The marginal success of these dissemination initiatives for this

evidence-based practice demands greater understanding of how to influence changes in practice. Research must strive to understand and overcome observed roadblocks to dissemination of family services, whether these roadblocks exist on the supply side among providers and service systems or on the demand side among patients and their families. Such roadblocks can take many forms. In identifying such barriers, health services researchers seek critical points for intervention to improve dissemination. For example, practitioners may be unaware of the value of the intervention, or they may lack the skills to deliver it. Families may be suspicious of such services as a result of unpleasant experiences based on older family therapy models. Payment systems may not provide support for the programs. However, even the FPE demonstration projects that had the capacity to overcome most or all of these barriers failed substantially to affect dissemination. In such a case, health services researchers have at least two other strategies at their disposal. The first is to develop more inventive dissemination strategies. The second is to decide that a modified or new evidence-based practice is needed—one that is more acceptable to providers, patients, and service systems.

This first task of developing more inventive dissemination strategies requires re-evaluating the barriers to dissemination. Such barriers exist at the disparate levels of patient and family, providers, administrators, and payers. Supply side strategies, such as providing higher rates of reimbursement for providing FPE, as well as more intensive and targeted training for providers, could be successful. Health services researchers investigate these strategies by using qualitative methods to interview providers and payers where FPE initiatives have already been started to determine whether they think that increased rates of payment or training would help. Strategies under study also target barriers to FPE at the consumer and family levels, thus focusing on developing the demand for FPE. A new approach called Recovery Oriented Decisions for Relatives' Support (REORDER) uses a shared decision-making process to structure patients' exploration of whether they want to include families in their mental health care. REORDER is personalized by attending to patient preferences for family involvement, focusing on the mental health recovery goals of the patient, and exploring how family members could join with the patient to meet those goals. The REORDER clinician begins by working with the consumer to identify pros and cons of family involvement and then uses education, support, and decision-making exercises to assist and empower patients to meet their goals. A randomized trial showed that this approach markedly increased rates of family contact with clinicians while it also increased a measure of recovery among participants assigned to the REORDER condition.

The second strategy to address difficulties in implementing FPE is to develop new evidence-based practices or to test revised models of FPE that

attempt to reduce barriers to implementation. For example, a meeting of family researchers and practitioners identified the need to determine the critical ingredients of FPE so that a simpler and briefer model could be tested. The National Alliance on Mental Illness Family to Family Education Program may provide a second approach. This program is not a clinical service, but is peer-driven. Trained family-member volunteers teach this manualized 12-week program in weekly 2- to 3-hour sessions. The program provides information about mental illnesses, treatments, medication, and rehabilitation. Participants also learn self-care, mutual assistance and communication skills, problem-solving strategies, and advocacy, and develop emotional insight into their responses to mental illness. Although the program does not aim to reduce patient relapse, it may provide some benefits to families. Notably, the program had been widely disseminated in spite of the dearth of systematic research demonstrating its effectiveness. After its dissemination, a partnership between health services researchers and NAMI conducted three sequential studies including a randomized trial have provided evidence that the program is helpful to families—in improving family functioning, coping, empowerment, and knowledge. This demonstrates that in the pursuit of dissemination of evidence-based practices, health services researchers may tackle the challenge of conducting research on existing practices that are in wide use but as yet relatively unstudied.

Case 3: The Continuing Saga of the New Antipsychotic Medications

The quest for a better antipsychotic medication spawned considerable enthusiasm, investment in new product development, and basic and clinical research during the last quarter of the 20th century. The net result was the introduction of several "second-generation" antipsychotic medications that by 2006 accounted for nearly 90 percent of all prescriptions for antipsychotic medications in the United States and nearly two-thirds of such prescriptions in the United Kingdom. This market growth was driven in part by a growing body of clinical trials, predominantly funded by pharmaceutical companies, reporting efficacy and safety advantages of the second-generation antipsychotics compared to the older agents. These reported advantages included a lower risk of extrapyramidal side effects, greater effect on negative symptoms and cognitive impairment, lower relapse rates, and greater safety and tolerability. The use of these agents grew exponentially and accounted for dramatic increases in pharmaceutical costs in the care of individuals with schizophrenia.

By the early 21st century, the major remaining issue seemed to be the relative efficacy and tolerability of the competing second-generation agents. Thus, two major, government-sponsored trials were launched on the relative effectiveness of the second-generation agents. These were the

Clinical Antipsychotics Trials of Intervention Effectiveness (CATIE), sponsored by the National Institute of Mental Health (NIMH), and the Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1), sponsored by the U.K. National Health Services Technology Assessment Program. Due to continuing concerns about the substantially higher costs of the second-generation agents and lingering questions about the results of industry-sponsored trials, both of these studies included comparisons with first-generation agents. The details of these studies are not within the scope of this section on health services research, but the main and consistent findings across both studies were that, contrary to study hypotheses, the second-generation agents in general showed little or no superiority over the first-generation agents on multiple outcome measures. The CATIE study found that although olanzapine had a somewhat longer time to discontinuation, it also had markedly worse metabolic side effects. Although much debate continues regarding the strengths and limitations of these effectiveness studies compared to the body of earlier efficacy trials, this evolution of research on antipsychotic medications illustrates several salient points and challenges for health services research.

As noted in Figure 58.4–1, the interface between standard clinical research and health services research is often where efficacy meets effectiveness. CATIE and CUtLASS are two large multisite effectiveness trials contradicting a large body of prior efficacy trials, well after those efficacy trials have had their effect on routine practice. The current discussion provides a good opportunity to further clarify the distinction between efficacy and effectiveness. There is a common misunderstanding that the types of outcomes measured determine this distinction. For example, in most efficacy studies the primary outcomes are "clinical," that is, symptom reduction and rates of side effects. The primary outcome for the CATIE trial was time to discontinuation of the assigned medication, and for the CUtLASS trial the primary outcomes were quality of life, participant satisfaction, and costs of care, in addition to symptoms and adverse effects. However, such outcomes could also be used for efficacy trials; the types of outcomes measured cannot be used to distinguish efficacy from effectiveness. Rather, the primary distinction between efficacy and effectiveness trials is that the former emphasize *internal validity* of a comparison in order to isolate effects attributable to the interventions studied—for example, active drug versus placebo or comparison of two active treatments. Efficacy trials typically restrict their samples narrowly to reduce "extraneous" variation due to factors other than the compared treatments. Because the samples are restricted, external validity or generalizability is reduced to achieve high internal validity. In contrast, effectiveness trials seek to determine the effects of treatments in more heterogeneous samples that represent the population of patients who might be treated. CATIE and CUtLASS are effectiveness trials because of

their multisite sampling of "typical" patients in various clinical settings. The CATIE and CUtLASS studies are thought to be more generalizable than the efficacy trials that preceded them, and they might therefore have greater effect on practices and policies.

A health services researcher always tries to understand the population and the system of care to which a particular study might generalize. Although CATIE and CUtLASS are more broadly reflective of persons with schizophrenia who are receiving treatment than were the previous efficacy trials, it is important to note that CATIE and CUtLASS too have restricted generalizability. The CUtLASS trial restricted its sample to patients of age 18 to 65 years with schizophrenia spectrum disorders who were assessed to have an inadequate response or adverse effects to their current medication. They were therefore deemed to be candidates for a medication change. CATIE had a somewhat different set of inclusion criteria: Age 18 to 65 years, schizophrenia diagnosis, able to take oral medications, no history of serious adverse reactions to the proposed treatments, more than one schizophrenia episode, no history of treatment resistance, not currently pregnant or breast-feeding, and no serious and unstable medical condition. For both studies, the mean age of the participants was 40 to 41 years, and for both samples the average length of illness was about 14 years. In addition, for both studies patients were willing to try a medication change. With regard to treatment settings, the 1,493 subjects in CATIE were drawn from 57 clinical sites in the United States, including 16 university clinics, 10 state mental health agencies, 7 Veterans Affairs sites, 6 private nonprofit agencies, 4 private practices, and 14 mixed systems of care. The 227 subjects of CUtLASS were drawn from 14 community psychiatric services in the English National Health Service. Practicing clinicians can rapidly recognize the limitations of even these broad CATIE and CUtLASS samples with regard to generalizability to their own patients and practice settings.

Given this current state of knowledge about effectiveness, safety, and patterns of use of the various antipsychotic medications, then, what further questions can health services research help to answer? There are many issues, but two key ones relate to formulary guidelines and how clinicians and patients will make choices regarding an antipsychotic agent, given the relative advantages, risks, and costs associated with the various agents.

It is not surprising that the results of CATIE and CUtLASS, essentially showing equivalence of effectiveness among most agents and differences in costs and adverse effects, have rekindled the debate about drug formulary policies. This debate is framed on one end by those who advocate to "fail-first" on a less expensive agent before a more expensive medication can be tried. On the other end are those who advocate for an open formulary that allows free choice by physicians and patients. This perspective argues that individual response cannot be predicted from the existing research and that, given apparent equivalence of effectiveness and variable safety

considerations, it is probably more cost-effective in the long run to allow individual patients and their physicians to decide what medication is best. This individualized approach may prove less costly, given that symptom reduction and avoidance of adverse events may be enhanced. Health services research can be brought to bear on such questions.

The formulary policy debate is one of supply side interventions at a systems level. Such interventions can be defined according to differing levels of restrictions to access to the various antipsychotic agents. These could include the "fail-first" use of cheaper agents before access to a more expensive agent is allowed, a "preferred drug list" that would make available all of the first-generation agents and a limited number of second-generation agents (based on cost), and a completely open formulary. The key health services research questions are how different formulary policies affect patient outcomes (symptoms, adverse events, drug discontinuations) and costs (medication costs, total mental health care costs, total health care costs) in large systems of care.

In the wake of the CATIE and CUtLASS studies, health services researchers also seek to understand decision-making processes in the selection of an antipsychotic agent. Given the results of CATIE, one might expect to see at least some changes in patterns of care, including increased use of the older agents and increased use of clozapine (based on evidence from efficacy trials, as well as CATIE, that clozapine is helpful in treatment-resistant schizophrenia). What has actually happened? In the United States, since the publication of CATIE and CUtLASS, the use of olanzapine has declined, perhaps due to the finding of associated metabolic risks. The use of first-generation agents has continued its slow downward drift. The use of clozapine has remained unaffected. Prescriptions for two other second-generation agents, quetiapine and ziprasidone, have grown modestly. There is little or no evidence that these large effectiveness studies have had substantial effects on practices. One obvious explanation for this apparent lack of response to the results of CATIE and CUtLASS in patterns of medication use is that there has been insufficient time for such effects. A health services researcher might seek to describe the patterns over a longer period of time, to understand the decision processes underlying such changes, and to assess the effects of interventions to bring about the desired widespread changes in practice.

Descriptions of patterns of antipsychotic utilization over time can be achieved by analyzing prescription claims data from large payers—especially Medicaid, but also Medicare Part D and commercial insurance carriers. Analysis of such large databases permits the ascertainment of overall prescription rates for the various antipsychotic agents and hence estimation of their changing market shares. Such analyses also permit examination of the association of a limited but important set of patient characteristics, including age, gender, race, and primary diagnosis, with

types of medications. These observations allow examination of issues such as disparities in access to various agents among subgroups. Similar analyses can identify the clinically questionable use of certain agents among patient groups known to be particularly at risk of certain adverse effects, patterns of "off-label use," and differences in prescribing patterns across insurance plans with different levels of formulary coverage. This type of *health services utilization research* is primarily descriptive but is used to focus on more specific research questions or hypotheses to be tested in *interventional services research*. For example, an observation in a utilization research study that a significant number of patients are being prescribed more than two antipsychotic agents simultaneously could lead to a utilization review intervention study that questions such practices and requires justification in individual cases. The effect of such a utilization review intervention on prescribing practices could be evaluated. Another intervention might involve identification of patients who are prescribed a second-generation agent that is known to have metabolic side effects and who are also prescribed medications that indicate a metabolic risk, such as cholesterol-lowering, antihypertensive, or hypoglycemic agents. Identifying such patterns could lead to an intervention to change prescribing practices. These two examples illustrate the type of supply side interventions at the provider level which can be developed in response to services research data drawn from the system level.

Health services researchers are also increasingly interested in interventions that influence interactions between individual patients and their providers. If the goals of antipsychotic therapy are to reduce psychotic symptoms with the least risk and to maximize the quality of life and functioning, how do clinicians and patients evaluate and negotiate the decision about what medication to use? To what extent do they incorporate evidence from clinical trials and practice guidelines into their decisions? What other sources of information do they use? How do patients and physicians ascertain and weigh individual risk factors? How do costs factor into their calculations? A variety of research techniques can be applied to these questions, including structured questionnaires, qualitative interviews, and direct observation of patient and physician interactions. As better understanding of these patient–physician decision processes emerges, interventions can be designed to facilitate improvements in care. Studies of shared decision making, a nonpaternalistic model of how patients and clinicians make medical decisions together, have shown that both physicians and patients in psychiatric settings have generally positive attitudes toward the approach. However, in practice, shared decision making is incomplete or less frequent compared with what occurs in general medicine settings. Researchers have begun to identify barriers and facilitators to shared decision making and to evaluate interventions to promote this process by preparing both professionals and patients to

engage in it, including through the use of decision aids and support systems. Knowledge about how physicians and patients obtain and weigh scientific information can lead to information being presented in better, more user-friendly and relevant formats. It is essential to evaluate changes not only in the processes of decision making and treatment—for example, increases in the use of shared decision making or greater consideration of risk factors in these conversations—but also associated changes in outcomes, symptom relief, reduction in adverse effects, patient function and quality of life, and physician and patient satisfaction. There is limited empirical evidence to date about the outcomes associated with the use of shared decision making in mental health, although researchers have found in patients an increased feeling of involvement in and positive attitude toward treatment as well as increased self-awareness.

OVERVIEW

Pressures are mounting to ensure that the fruits of biomedical research translate into treatments and services. Mental health services research has substantial power to both inform and to advance this process. In the end, all mental health services research can be summarized in three questions. Do persons with mental health needs have access to treatment? Is the treatment they receive of adequate quality, incorporating the latest in scientific knowledge? Does the treatment offer good value—that is, good outcomes for the resources expended? These questions are posed by the field in such a manner as to maximize generalizability and to identify subgroup variations. These questions must also be asked within systems concerned with health care and social policy to inform clinical care and maximize the benefit to individuals—the ultimate goal.

Mental health services research, necessarily interdisciplinary, must consider how social factors, financing systems, organizational structures and processes, health technologies, and personal behaviors affect access to health care, the quality and cost of health care, and ultimately health and well-being. Health services researchers work closely not only with biomedical colleagues, but also with investigators from other scientific disciplines (sociology, anthropology, psychology, economics, education, and political science) and with service system and policy leaders to implement and evaluate changes.

Much of biomedical science tends to be occupied with internal validity—the strength of the claim that the experimental manipulation causes the outcome. This leads to the desire to control all potentially confounding variables, but it reduces external validity. External validity, or generalizability, lies at the center of the translational process and health services research. At the same time, generalizability encompasses sensitivity to variations among subgroups of patients, taking into account demographic and cultural differences in disease, health care beliefs, and

behaviors and access.

Moving forward, mental health services research faces several challenges. First, while the advance of "evidence-based practices" embodies the principle of translating scientific knowledge into practice, more attention must be given to the complementary principle of "practice-based evidence." In essence, how can one inform the directions of science through evidence that emerges from clinical practice and experience? Second, challenges for treatment systems and for research seeking to identify and promote "best practices" arise from the growing diversity of US culture. Third, the field needs a better understanding of how to implement system changes and to disseminate best practices. For this the field must look beyond the traditional medical and social sciences to such disciplines as marketing science. Fourth, research designs that treat service providers as research subjects require careful consideration to ensure the protection of human subjects that have not been traditionally addressed by institutional review boards.

Finally, major challenges exist in the juxtaposition of the comparatively slow progress of science with the needs of practitioners, planners, administrators, and policy makers, who must respond quickly to the needs of patients and systems of care. The Recovery After an Initial Schizophrenia Episode (RAISE) project of the NIMH provides a current example of a multidimensional partnership serving to hasten the translation of scientific findings to practice. The NIMH funded two RAISE projects to study the effectiveness and implementation of evidence-based coordinated specialty care programs for young adults experiencing the onset of nonaffective psychosis. Attention to policy and practice while these research studies were being conducted led to the relatively rapid expansion of resources for evidence-based programs as part of the mental health block grant in 2014. For health services research to remain relevant, new models of partnership will continue to be needed—models that provide timely information in an iterative and collaborative manner among all stakeholders.

References

Backer TE, Liberman RP, Kuehnel TG. Dissemination and adoption of innovative psychosocial interventions. *J Consult Clin Psychol*. 1986;54:111–118.

Croghan TW, Schoenbaum M, Sherbourne CD, Koegel P. A framework to improve the quality of treatment for depression in primary care. *Psychiatr Serv*. 2006;57:623–630.

*Demyttenaere K, Bruffaerts R, Posada-Villa J, et al; The WHO World Mental Health Survey Consortium. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *JAMA*. 2007;291:2581–2590.

Dixon LB, Glynn SM, Cohen AN, et al. Outcomes of a Brief Program, REORDER, to Promote Consumer Recovery and Family Involvement in Care. *Psychiatr Serv*. 2014;65(1):116–120.

Dixon L, Lucksted A, Medoff D, et al. Outcomes of a randomized study of a Peer-Taught Family

To Family Education Program for mental illness. *Psychiatr Serv*. 2011;62(6):591–597.

Dixon L, Lucksted A, Stewart B, et al. Outcomes of the peer-taught 12-week Family-to-Family Education Program for Severe Mental Illness. *Acta Psychiatr Scand*. 2004;109:207–215.

Dixon L, Lyles A, Scott J, Lehman A, Postrado L. Services to families of adults with schizophrenia: from treatment recommendations to dissemination. *Psychiatr Serv*. 1999;50:233–238.

Dixon L, McFarlane W, Lefley H, et al. Evidence-based practices for services to families of people with psychiatric disabilities. *Psychiatr Serv*. 2001;52:903–910.

Essock S, Goldman H, Von Tosh L, et al. Evidence-based practices: setting the context and responding to concerns. *Psychiatr Clin North Am*. 2003;26:919–938, ix.

*Goldman HH, Frank RG, Burnam A, et al. Behavioral health insurance parity for federal employees. *N Engl J Med*. 2006;354:1378–1386.

*Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press, 2001.

Jones PB, Barnes TRE, Davies L, et al. Randomized controlled trial of the effect on quality of life of second- vs. first-generation antipsychotic drugs in schizophrenia (CUtLASS 1). *Arch Gen Psychiatry*. 2006;63:1079–1087.

Katon W, Robinson P, Von Korff M, et al. A multifaceted intervention to improve treatment of depression in primary care. *Arch Gen Psychiatry*. 1996;53:924–932.

Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*. 2003;289:3095–3105.

Lehman AF, Goldman HH, Dixon L, Fox R. *Evidence-based mental health treatments and policies—Examples to inform public policy*. New York: Milbank Memorial Fund; 2004.

*Lehman AF, Kreyenbuhl J, Buchanan R, et al. The Schizophrenia Patient Outcomes Research Team (PORT): updated treatment recommendations 2003. *Schizophren Bull*. 2004;30:193–217.

Lieberman JA, Stroup TS, McEvoy JP, et al; for the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Investigators. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med*. 2005;353:1209–1223.

Norquist GS, Regier DA. The epidemiology of psychiatric disorders and the de facto mental health care system. *Annu Rev Med*. 1996;47:473–479.

Pitschel-Walz G, Leucht S, Bäuml J, Kissling W, Engel RR. The effect of family interventions on relapse and rehospitalization in schizophrenia—a meta-analysis. *Schizophren Bull*. 2001;27:73–92.

*President's New Freedom Commission on Mental Health. *Achieving the Promise: Transforming Mental Health Care in America. Final Report (DHHS Pub. No. SMA-03-3832, Vol. 2004)*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2003.

RAISE Project Overview. Available at http://www.nimh.nih.gov/health/topics/schizophrenia/raise/index.shtml.

Rosenheck RA, Leslie DL, Sindelar J, et al; for the CATIE Study Investigators. Cost-effectiveness of second-generation antipsychotics and perphenazine in a randomized trial of treatment for chronic schizophrenia. *Am J Psychiatry*. 2006;163:2080–2089.

Thota AB, Sipe TA, Byard GJ, et al. Collaborative care to improve the management of depressive disorders: a community guide systematic review and meta-analysis. *Am J Prev Med*. 2012;42(5):525–538.

Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care. *JAMA*. 2000;283:212–220.

Young AS, Klap R, Sherbourne CD, Wells KB. The quality of care for depressive and anxiety disorders in the United States. *Arch Gen Psychiatry*. 2001;58:55–61.

▲ 58.5 The Psychiatric Hospitalist

BARRY H. GUZE, M.D., AND ANGEL CIENFUEGOS, M.D.

INTRODUCTION

Over the last quarter of a century, the American health care delivery systems, and other systems around the world, have adopted an inpatient treatment model called by most the "hospitalist model." As opposed to the nonhospitalist traditional models, these are medical providers who are part of inpatient programs in medicine, pediatrics, surgery, and now mental health, in which the medical care is provided. The hospitalist is a specialist in inpatient treatment, based in an inpatient facility or hospital. This type of practice specializes in acute and serious illness rather than chronic ambulatory medicine and health maintenance. They are knowledgeable about comorbidities in acute and critical clinical presentations. The number of hospitalists has been growing consistently and it has been suggested that the generalist who cares for patients across the full continuum of care may be a vanishing breed.

Traditionally, patients have been cared for in the hospital by their prior primary care physicians, in most cases family practitioners or general internists in a "Continuity of Care" model. In a recent transformation in many parts of the United States, when patients are hospitalized their care is now more likely to be provided by a hospitalist. In especially complex cases in which multiple consultations where subspecialists are required, it is the hospitalist who coordinates care. When the patient is discharged, he or she returns to the care of the primary care physician and routine follow-up by other medical specialties. This change parallels the change of practice patterns in psychiatry, in which, in many instances, a separate psychiatrist may care for a patient while admitted for inpatient treatment. Historically, these individuals in psychiatry have been called unit directors or ward chiefs of large state mental institutions. However, in their current role, they are the psychiatric parallel of the hospitalist found in medicine and pediatrics. In both cases, outpatient physicians relinquish the care of hospitalized patients to inpatient specialists.

The development of the hospitalist model has occurred similarly for

medicine and psychiatry in the context of the changing role for acute care hospitals. As both medicine and psychiatry become more ambulatory in orientation, inpatient institutions have been moving away from their central role in health care. They are becoming specialized institutions for acutely ill patients. Generally, the level of severity of illness in the inpatient setting has continued to increase as well as the number of comorbidities. Although the hospitalist model was widely accepted previously in Canada and the United Kingdom, the first modern hospitalist program was established in the United States in 1994, overcoming the opposition of many primary care physicians who objected to the model because of the disruption in continuity of care. In any case, it was a fact that inpatient care had become an economically inefficient use of their time.

The hospitalist has emerged in response to the increasing complexity of care associated with hospitalized patients. These individuals, with time and training, develop significant experience with the acutely ill patient. They develop an understanding of continuity and coordination of care since part of their responsibility is to work on the "hand-off"—the patient's transition from outpatient to inpatient and, after treatment, back again to the prior outpatient physician.

The initial growth in the number of hospitalists occurred because leaders at health maintenance organizations (HMOs), some hospitals, and medical groups believed that they would lead to a reduction in inpatient costs without a reduction in patient satisfaction or quality of care. Recent estimates indicate that the current number of hospitalists in the United States is nearly 30,000. In 2010, 57 percent of US hospitals had hospitalists on staff including 87 percent of hospitals with a number of beds equal to or above 200. Approximately 80 percent of Veterans Affairs hospitals currently have hospitalists. Membership in the Society of Hospital Medicine (SHM), the professional organization for hospitalists, is about 6,000.

Because hospitalists have become common at the major teaching hospitals in the United States, they have also become a major provider of instruction to medical students and residents in internal medicine and pediatrics. These two groups now receive much of their inpatient education from hospitalists working in their role as attending physicians.

THE DEFINITION OF HOSPITALIST PHYSICIANS

Robert Wachter, MD, MHM coined the term "hospitalist" in the *New England Journal of Medicine* in 1996. **According to the SHM, this is a** medical specialty dedicated to the delivery of comprehensive medical care to hospitalized patients. Practitioners of hospital medicine include physicians and nonphysician providers who engage in clinical care, teaching, research, or leadership in the field of general hospital medicine. In addition to their core expertise managing the clinical problems of

acutely ill, hospitalized patients, hospital medicine practitioners work to enhance the performance of hospitals and health care systems by attending to all patient care needs during hospitalization; being mindful of quality and process improvement techniques; collaborating, communicating, and coordinating with all physicians and health care personnel; caring for hospitalized patients; and ensuring safe transitioning to the community, all the while emphasizing the efficient use of hospital and health care resources.

Hospitalists should be physicians who have general inpatient care as the primary focus of their professional activities, which include clinical care, teaching, research, or administration. Hospitalists typically spend more than 25 percent of their time working in hospital settings. Commonly they are internists taking the place of primary care physicians in providing inpatient care. Many hospitalists become involved in clinical-administrative tasks in addition to patient care responsibilities. Commonly these include quality improvement efforts and peer review. Academic hospitalists may explore issues such as outcomes research that explores the cost and quality of inpatient care. Many lead the development of new inpatient educational programs for house staff and medical students. Most hospitalists work in community hospitals; a smaller number work in teaching hospitals. In those academic settings, their responsibilities expand to include teaching medical students and residents.

In general medicine, most hospitalists have been trained as internists. There is no separate training program for hospitalists, although some residency tracks and 1-year fellowships have begun to emerge. Such residency tracks typically focus on training in inpatient medicine as well as hospital-oriented quality improvement, patient safety, and end-of-life care. Fellowships usually focus on providing future academic hospitalists with the research skills they will need in areas related to inpatient care.

Hospitalists are usually salaried physicians with defined work hours. They care for patients with complex medical illnesses, rather than the chronicity found in outpatient practice. Many of them have interests in health systems and want a chance to improve the quality of services and control costs associated with inpatient practice. The increasing complexity of medical care demands greater specialization. In many settings, the expectation for physicians to have expertise in both inpatient and outpatient medicine is increasingly regarded as unrealistic. Supporters of the hospitalist movement emphasize the benefits of a specialized knowledge base. Hospitalists specialize in inpatient care, allowing primary care physicians to focus on ambulatory care. A consequence of the increase in the number of hospitalists has been a proportionate increase in the number of “office-only” physicians as well.

Hospitalists in academic settings are on-site experienced physicians. Their presence helps hospital wards to run more smoothly, which increases

patient and nursing staff satisfaction. Such individuals are usually familiar with ward dynamics. They also often have a good relationship with the ancillary staff, appreciate their contributions, and have long-term commitments to the hospital. Such involvement facilitates the management of clinical situations and therefore prevents them from evolving into significant problems. Moreover, these individuals generally have a breadth of clinical knowledge and experience from working in hospitals. Thus, they are skillful in the expeditious treatment of the seriously ill, comfortable in managing multiple minor inpatient problems, and familiar with consulting various medical and surgical specialists. These attending physicians are also best positioned to provide the supervision that interns and residents need when managing acutely ill inpatients. Due to their inexperience and rapid turnover, interns and residents benefit from the support that a responsible physician provides when making major medical decisions.

It has been suggested that the minimum amount of inpatient work expected of a hospitalist is 25 percent of their practice. This minimal amount allows the hospitalist to become experienced in the management of common inpatient conditions. Relationships with care providers in affiliated disciplines can also develop more easily with this degree of availability. Such individuals are often more likely to commit to enhancing the hospital by participating in quality improvement efforts and helping to control costs.

There are two common models of hospitalist care. The most common consists of full-time inpatient physicians who practice as a group. In the second model, a group of physicians rotate responsibility for inpatient care. The rotational periods vary, depending on the practice of the group. This method allows a group to maintain skills in both inpatient and outpatient practice. It can, however, produce frequent periods of discontinuity in outpatient care.

The use of the hospitalist model may be voluntary or involuntary depending on how managed care organizations choose to transfer care when a patient is admitted for inpatient treatment. Commonly under the involuntary model, the outpatient physician is not permitted to write orders and is relegated to visiting the patient in the hospital. In contrast, many physicians voluntarily transfer responsibility for the care of their patients to hospitalists, cardiologists, or specialists in intensive care medicine. For example, it is common for primary care physicians to transfer care of their patients to intensivists (physicians specializing in intensive care) when their patient is admitted to the intensive care unit (ICU). In a similar manner, geriatricians working in nursing homes commonly transfer care to internists when their patients are admitted to the hospital. Although not all hospitalists may have the same degree of interest or ability in studying patient outcomes or developing practice guidelines, most of them participate in these organizational activities.

ROLES OF THE HOSPITALIST

The predominant role of the hospitalist is to take care of hospitalized patients. Over the years inpatient care has become more and more specialized and has gained in complexity and, while the number of hospitalized patients has decreased, the acuteness of their conditions has risen. Simultaneously, the scrutiny of the length of stay (LOS) and associated hospital costs are increasing. It is common to see hospitalists involved in projects to reduce LOS and cost of admission. Generally, reports show that inpatient care of medical patients by hospitalists reduces LOS and hospital costs. Hospitalists are able to respond more rapidly to changes in a patient's condition because they tend to be readily available and control the timing of treatments and consultations.

The hospitalist is also the receptor for inpatient admissions. Increasingly, outpatient physicians prefer to transfer hospital care to a specialist. Common reasons for this preference include the convenience of uninterrupted outpatient work, the convenience of not making rounds in the hospital, and a lack of expertise or comfort with the acuteness of illness and clinical skills associated with inpatient work. There is also some desire to avoid the intensive scrutiny associated with providing care to hospitalized patients. Similarly, hospitalists provide pre- and postoperative medical care to surgical patients. Their availability allows the surgeon to be in the operating room all day and ensures that the patient receives appropriate medical care. Many health care systems are having hospitalists become the physician of record for nonmedical inpatients with medical comorbidities.

An analogous situation has long existed with ICUs. Most ICUs are closed, meaning that they use dedicated critical care physicians who accept the patient from the generalist on the ward and transfer care back to them when the patient leaves the unit. These physicians have the skills, abilities, and availability to care for the critically ill patient, and because of these skills they have improved outcomes and efficacy.

In many specialties, including critical care, oncology, bone marrow transplantation, cardiology, infectious diseases, and psychiatry, the hospitalist model is common. The attending physician is dedicated to inpatient service. In academic centers, this may be for multiple rotations throughout the year or full time.

An additional role for hospitalists is providing physician participation in hospital management. Commonly this occurs by their joining in quality management operations, the development of hospital policies and procedures, management of the formulary, participation in credentialing activities, peer review, and utilization management. They also care for outpatients, often in a consultative role. Some hospitalists participate in follow-up clinics in which patients are seen once or twice following discharge. Such outpatient activities, however, represent only a small

fraction of their time.

Hospitalists also play a role in communication. Most systems expect the hospitalist or their designee to contact the primary care physician within 24 hours of admission and to ensure that either a summary letter or the discharge summary is sent to the outpatient clinician at discharge. They also transfer clinical information from the outpatient care physician to the inpatient provider on hospitalization and the subsequent transfer of information back to the outpatient care physician by the hospitalist on discharge. This "hand-off," also known as "warm hand-off" in cases when the communication is verbal, is symbolic of the transfer of clinical responsibility that occurs on hospitalization. The hospital-based physician assumes responsibility for diagnosis and management. For this reason, most hospitalists believe that there is an obligation to validate the diagnosis. This is, in essence, the provision of a second opinion about diagnosis and subsequent management. Much has been written about the hand-off. It is one of the two controversies of the model (the other controversy being obligatory use of hospitalists). The disadvantage of the hand-off is obvious: Important information can be lost and result in discontinuity. There may also be patient dissatisfaction because of the assignment of a new physician.

The hand-off can also be viewed as producing one of the strengths of the hospitalist model. It results in the patient being under the care of a physician in the hospital throughout the day. This individual reacts to clinical data in real time. He or she is the coordinator of inpatient care. However, some surveys show that patients are at least as satisfied with the care from hospitalists as they are with care from their primary care physician. The data suggest that patients may be willing to trade familiarity for availability. It is also noteworthy that many hospitalized patients do not have primary care physicians, especially among the uninsured. In some systems where hospitalists are introduced there is an increase in patients' satisfaction with their outpatient physician, suggesting that the improved ability to focus on their outpatients is a benefit of this model for the primary care physician.

Hospitalist systems that focus on inpatient–outpatient communication have fewer difficulties. Commonly, programs fax the discharge summary to the outpatient clinician on the day of discharge. Some programs fax daily progress notes as well. Increasingly, programs have a continuity clinic in which patients may be seen once or twice soon after discharge to ensure stability before the patient reappears at the primary care physicians' office.

THE PSYCHIATRIC HOSPITALIST

It could be argued that psychiatry was the first of the medical disciplines to adopt the hospitalist model on a large scale. Paradoxically, however, full-time inpatient psychiatrists have never called themselves hospitalists.

Psychiatric hospitals and asylums date back to the early Middle Ages. The most available career opportunity available to doctors treating the mentally ill was working in a hospital until the rise of psychoanalysis. The career choice was to focus on working in a publicly supported state mental hospital or a private psychiatric hospital. Those physicians resemble what are now considered to be hospitalists.

With the advent of antipsychotic drugs and the trend for deinstitutionalization during the 20th century, psychiatry, along with numerous other disciplines, adopted the model of continuity of care, in which physicians follow their ambulatory patients both in outpatient settings and while hospitalized. Outpatient psychiatrists admitted and followed their patients in the event of a needed inpatient stay.

As the field has expanded and the effect of third-party payers has restructured the profession, the proportion of hospitalists has declined while the absolute number has increased. One of the earliest surveys of career areas for psychiatrists dates to 1953; it noted the role of psychiatrists working full time in hospitals. Later surveys in the 1960s, 1970s, and subsequent years noted that the percentage of psychiatrists working in mental hospitals was decreasing. In 1965, for example, roughly 30 percent of psychiatrists worked in hospitals, but by 1979 this had dropped to less than 14 percent.

One of the problems that nonhospitalist-based psychiatry has shared with other medical disciplines is that the admission and care of their patients while admitted was very disruptive of their careers as academicians, teachers, and researchers. The role of unit director or ward director has been a traditional role in psychiatry. At one time, these individuals were relatively junior psychiatrists who, in academic settings, would attempt to succeed in the traditional triad responsibilities of research, patient care, and teaching. In any setting, academic or not, these individuals are the epitome of the psychiatric hospitalist. Their job functions are highly congruent with the descriptions of hospitalists noted earlier.

In 2010, one of the hospitals of the Veterans Affairs Healthcare System transitioned its mental health inpatient units from nonhospitalist or continuity of care (nonhospitalist based) to a hospitalist-based model of care. This change resulted in an increase in the total number of discharges and a decrease in the number of readmissions within 30 days. Surprisingly, patients discharged by hospitalists were more likely to follow up with after care appointments. This is important considering the statistically increased incidence of suicide after an inpatient stay.

As it is with other medical disciplines, one of the expectations from a financial standpoint is that psychiatric health care delivered by hospitalists will decrease the LOS as well as the cost. It may also facilitate care of patients with medical comorbidities, since medical consults are readily

available and hospitalists are likely to be more familiar with interacting with medical and surgical disciplines. In a modern institution, psychiatric hospitalists cover on-call and consult responsibilities. Psychiatric hospitalists are familiar with assessing and treating agitation, acute states of the most severely mentally ill patients, including psychotic and mood disorders. They are also comfortable with protocols to assess and treat withdrawal from alcohol and other substances a well as with keeping up with the legal requirements of involuntary admission.

Psychiatric hospitalists spend most of their time on inpatient units. They carry a significant share of resident and medical student training. Trainees working with hospitalists are frequently more satisfied with the amount of time and dedication they get from their attending physicians, which normally are part of the clinical faculty of the associated academic institution. These faculty members are expected to be educators in addition to clinicians and are not subjected to the same publication and funding expectations of research faculty. The inpatient units are the center for education and training during the first years of residency and offer the appropriate diversity of clinical presentations providing students with the appropriate amount of learning material.

Psychiatric inpatient units are environments where multidisciplinary approaches involving psychiatrists, nurses, psychologists and social workers are more frequent than in many outpatient settings. Co-management of psychiatric and medical problems will also be particularly common for these practitioners making them more familiar with interacting with medical consults and more apt to treat these comorbidities more effectively. For similar reasons, they will be better qualified to assess appropriateness for transfers from medical to psychiatric units.

HISTORY OF THE MODERN HOSPITALIST MOVEMENT

The Modern Hospitalist Movement was born in the United States in the 1990s. The argument still continues between followers of the "Continuity of Care" model and the hospitalist model. Several factors led to the change. For the outpatient clinician, the arrangement frees them from the disruptive burden of seeing a small number of inpatients. It allows them to focus on their area of expertise: outpatient practice. It avoids disruptions in the outpatient schedule that can occur when inpatient rounds are a daily responsibility. For the hospital and third-party payers, the model reduces inpatient expenditures with similar clinical outcomes. This model does not adversely affect patient satisfaction with care or trainees' satisfaction with their educational experience. Although this practice pattern may benefit hospitals or group practices, its professional fees are not usually sufficient to make the program self-supporting. To remain viable over the long term, these programs require support from the programs such as hospitals or group practices that benefit from the efficiency rendered by these

programs.

Under the traditional model for inpatient care, outpatient physicians made hospital rounds early in the morning and again at the end of the day. There could be calls from the emergency room or the hospital. The calls would, at times, interrupt outpatient practice and interfere with the efficient treatment of outpatients. Many clinicians noted a decline in office productivity when they were obliged to make hospital rounds for a small number of inpatients.

A significant factor that led to the development of hospitalists was the desire to focus on efficient outpatient treatment and to allow inpatient treatment to be provided by those who would give it full-time attention. This allows outpatient physicians to avoid the frequent interruptions, paperwork, and preauthorizations associated with inpatient care.

The acuteness and diagnostic complexity of inpatients have been increasing over the last several decades. With the increase in acuteness and complexity have come newer treatment modalities and procedures associated with inpatient care. Those who only infrequently care for inpatients have found it increasingly difficult to remain current in these treatment methods. Part of the appeal of the hospitalist movement has been that it frees outpatient physicians of the responsibility to remain current with treatments they would only rarely use.

In England and Canada, the early recognition of the benefits of focusing on either ambulatory or inpatient care led very early on in this process to distinction of outpatient and inpatient physicians. The term “hospitalist” was first used to describe this concept in the United States by Wachter and Goldman.

In academic medical centers, hospitalists have become a third career pathway in addition to the prior pathways of clinical educator and clinical researcher. The development of the clinical educator was due to the need to care for outpatients and to teach students and residents at ambulatory care sites. How an individual physician spends his or her time is highly variable and tends to depend on the needs of the institution and the skills and experience of the individual physician. Time is typically divided among providing direct care to outpatients, supervising students and residents in providing ambulatory care, and various scholarly activities. Many also spend several months each year attending on an inpatient unit, during which time outpatient responsibilities are reduced. Educational activities may include responsibility for specific courses in the medical school curriculum or the teaching of specific knowledge or skills, for example, in patient examination. Most academic medical centers expect their clinician teachers to do scholarly work to meet standards for promotion. These activities commonly include developing new educational programs, writing review articles or case reports, or conducting research in medical education along with presentations at scientific meetings.

In Europe and Canada, there have long been specialists in inpatient general medicine who have been active in urban hospitals. A clear demarcation between inpatient and outpatient practice is new in the United States. In the United Kingdom, the consultant physician would take over care from the general practitioner when the patient was admitted to the hospital.

The increasing use of such hospitalists was initially spurred on by the emphasis on efficiency that came with the rise of managed care. The managed care environment stimulated the development of the hospitalist model in the United States. The emphasis on efficiency and cost containment took precedence over continuity of care.

Until the 1970s, most primary care was provided by a general practice physician who was not board certified in any particular specialty. Most medical care at that time was provided in a hospital setting. In the inpatient setting, the general practitioner was often assisted by the internist, who was often more experienced in inpatient medicine. By the 1980s, family practice had emerged as a specialty and came to define the general practitioner. The role of the internist in inpatient practice was diluted by the rise of medical specialists who could assist the general practitioner with inpatient care. Until the mid-1990s, hospitalists were uncommon in general medicine in the United States. Rare examples of hospitalists in the United States before the 1990s were called house doctor, hospital rounder, or inpatient manager. By the 1990s, however, general internists and family practitioners had begun to focus on outpatient care and had a progressively smaller role in inpatient care. Internists began to manage inpatients, and intensivists emerged to manage ICUs. Emergency departments began to be staffed by specialists in emergency medicine.

Although multiple forces were responsible for this change, two major factors were predominant—cost pressures on hospitals and physicians, and the increasing acuteness of inpatients. The shift of most medical care to the ambulatory setting resulted in most physicians having little ongoing experience with inpatient care. Practice effects meant that those who did large volumes of inpatient care had superior outcomes.

Many felt that the hospitalist model conflicted with the traditional notion of a generalist who would treat patients in both inpatient and outpatient settings and in so doing provided treatment that was less fragmented and more holistic.

THE STAGES OF HOSPITAL CARE

It has been suggested that hospital care in the United States has gone through four stages. In the first stage, all inpatients were cared for by their primary care physicians. This worked best when the acuteness of illness was low and most practitioners had large numbers of inpatients, cost was not as important as it is today, and hospitals were adjacent to practitioners'

offices.

The second stage occurred when these conditions no longer existed. Specifically, hospitals are no longer nearby, acuteness of inpatients is significant, costs are managed, and inpatient care has a significantly higher degree of complexity. Furthermore, outpatient offices are busy due to the shift of care from inpatient to ambulatory. At this point, physicians may form a group to rotate responsibilities for inpatient care. This system introduces an inpatient–outpatient discontinuity. The compensation for this is the enhanced availability of the inpatient physician.

A third stage is the development of voluntary hospitalists who are dedicated inpatient physicians paid by the hospital, medical group, or managed care company. Under this model not all inpatient care is provided by hospitalists. Primary care physicians retain the opportunity to manage their patients when they are admitted to the hospital.

A fourth stage is the mandatory hospitalist stage. At this point, all inpatients are cared for by hospitalists. This model is more common when there is strong institutional control of physicians. When there is a loose affiliation between the hospital and primary care providers (PCPs), the relationship is more likely to be at stage three if outside physicians wish to retain control of their patients.

BENEFITS FROM THE USE OF HOSPITALISTS

There has been more than a decade of research in benefits and disadvantages of hospitalist care. Most of this body of research demonstrates that hospitalists are efficient providers of care while lowering LOS and total hospital costs without lowering quality of care. Hospitalists are more efficient in hospital resource utilization than physicians who spend the majority of their time practicing in ambulatory settings. Many factors are believed to contribute to the improved efficiency of hospitalists. These include the practice effect of their familiarity with common inpatient conditions and with the hospital system. They are also more available to acutely ill patients and may respond more promptly to changes in their clinical state or changes in laboratory data.

Ambulatory care physicians experience a significant benefit from the presence of hospitalists. Hospitalists allow outpatient-based clinicians to attend to their patients without the disruption in their schedules that can result from the unpredictable demands associated with caring for acutely ill patients in the hospital. Furthermore, many ambulatory care-based clinicians may lack the familiarity with the inpatient management techniques that they had when they were trainees.

For the typical internist, the inpatient census is small. However, this small number of patients typically requires more effort than in the past because the patients tend to be more acutely ill and their care more complex. Furthermore, the pace of hospital work has also quickened, and

physicians are under pressure to minimize LOS, making the emphasis on efficiency even greater. Conditions that were once an indication for inpatient admission are now commonly treated in the outpatient setting, be that the office, the home, or the nursing home. The result is that outpatient practice is also more acute and demanding, making the disruptions of providing care for some inpatients also more challenging. Hospitalists may also shorten a patient's stay in the emergency department by being available throughout the day to admit a patient. An experienced hospitalist usually becomes comfortable caring for acutely ill patients with a variety of medical conditions. In some cases, this may result in a decreased use of laboratory studies and consultative services. This may also contribute to a reduction in the cost of inpatient care.

Because most primary care physicians spend decreasing amounts of time caring for inpatients, many feel they have lost some of their skill in inpatient care. This reduction has occurred both because fewer patients are admitted and because LOS is shorter. With fewer inpatients, there are less interruptions in a tightly scheduled day of outpatients. What emerges is an increasing split between inpatient and outpatient practice.

Acting as case managers, hospitalists reduce inpatient resource utilization, resulting in a shorter LOS and a lower cost of inpatient care without a reduction in quality of care or patient satisfaction. The reduction in resource utilization is likely a practice effect combined with the ready availability of hospitalists for diagnostic and management decisions. They coordinate and integrate care for their assigned patients, generating and reviewing clinical data, and facilitate access to subspecialty services. Since they are more readily available, they generally have more time to communicate with patients and their families as well as with the PCPs. The rapid increase in the number of hospitalists in the United States suggests that all these factors bring significant benefits to hospital health care.

There are a large number of uninsured patients in the United States. When they present in the emergency department seeking admission, there usually is no primary care physician with whom they have a prior relationship. Therefore, there is no outpatient physician available to manage their inpatient care. An increasing role for hospitalists is to provide care for these patients. Many office-based physicians feel that providing inpatient care is not cost-effective. The poor reimbursement for hospital visits is a common reason that many no longer provide hospital care. More and more outpatient physicians are also seeing sicker patients. Their volume of inpatient work has also increased.

Good communication between inpatient and outpatient physicians has also enhanced the acceptance and perception of benefits of this model. A key aspect of communication is the prompt delivery of a discharge summary to the outpatient physician.

Hospitalists also work with surgeons to provide primary inpatient care

to surgical patients in the pre- and postoperative periods. These combined hospitalist–surgeon teams have been associated with reductions in postoperative complication rates with no significant difference in cost or LOS.

Hospitalists improve the efficiency of inpatient care while maintaining or enhancing the quality of care. Their expertise comes from familiarity with the management of common inpatient conditions. Efficiency consists in reductions in resource use, usually measured as average LOS and hospital cost. This occurs because hospitalists are familiar with clinical pathways, are experienced in the planning of post discharge care, and are available throughout the day to respond to changes in diagnostic data or clinical condition.

Hospitalists relieve office-based physicians of the unpredictable demands of hospital practice. Patients are no longer kept waiting, and physicians no longer have the inconvenience of recurrent trips to the hospital. Outpatient physician productivity increases following the removal of inpatient responsibilities. Outpatient physicians are more reliably available in the office. They are not interrupted by hospital calls or the need to leave abruptly to see acute inpatients. They are able to expand their outpatient hours because they no longer make inpatient rounds. Their workday is more predictable. Another issue is that the number of acute care hospitals is diminishing and with them the relative percentage of patients admitted to the hospital.

Hospitalists standardize the quality of inpatient care. In many academic medical centers, there has been a concern about the wide range in inpatient abilities found in community-based clinicians. Many programs have worried about granting these physicians admitting privileges or opportunities to participate in the teaching service. Using a hospitalist system allows both the community-based physicians and the outpatient academic faculty to enjoy the same quality of inpatient care that is provided by hospitalist faculty.

Hospitalists produce clinical science with an emphasis on outcomes research, quality assessment, and quality improvement; clinical trials to improve treatment or diagnosis; and the development of clinical guides and pathways. They have been found to have better performance in measures of patient satisfaction, particularly with the discharge process, and they are recognized as leaders in issues of patient safety.

There is a positive correlation between experience and quality of care in a number of areas in medicine and surgery. When implemented, the hospitalist model has resulted in fewer but more appropriate specialty consultations, lengths of stay, and hospital costs. Due to their familiarity, hospitalists develop good working relationships with ancillary hospital staff, including discharge planners, nurses, consulting specialists, and outpatient physicians. Retention and recruitment improve in the

hospitalist system.

HOSPITALISTS AS CLINICIAN-EDUCATORS

Large academic institutions are the leaders in hospitalist-based care. In these institutions, one of the hospitalist functions is training of residents and medical students. Some are considering or implementing hospitalist fellowships. This means house staff receive training by those with expertise in hospital care. Throughout the day, there is an expert readily available with whom to confer about patient management. There is improved teaching and expert guidance. Hospitalists provide role modeling to trainees for careers in hospital work. They provide teaching effectiveness, knowledge of relevant subject matter, discussion of pathophysiology, emphasis on cost-effectiveness, and timely and effective student feedback.

The effect of hospitalists on academic medical centers may broaden the separation between researcher and clinicians. The researchers will focus further on research and obtaining grants. The clinician-educators will concentrate on clinical work and teaching. With the group of clinician-educators, there will be an added subdivision, with some focusing on outpatients and others on inpatients. Hospitalists may bring relief to the residents' on-call duties by covering for clinical needs.

Several studies have also found that hospitalists improve the educational experience for medical students and house staff. Because the faculty supervisors are skilled and experienced in modern inpatient care, they provide more effective teaching and oversight. Generally, when academic programs have evaluated the teaching abilities of hospitalists there has been enhanced satisfaction by students and residents when compared to prior models. Some trainees have also noted a reduction in workload associated with the hospitalist model. House staff report that hospitalists are more effective teachers.

Many academic medical centers have hired hospitalists as clinician-educators. This need has arisen because primary care physicians are able to spend very little time in the hospital. This results in less time with trainees for educational activities. The expectation is that hospitalists will devote more time to education, performance feedback, and supervision. This expectation has been supported by resident evaluations of teaching which rate hospitalists higher than traditional attending physicians, who briefly rotate onto the teaching service.

Clinical research done by hospitalists has traditionally consisted of studying the hospitalist movement. This has included studies of the efficiency, cost, and quality of hospitalist-run services, including inpatient–outpatient communication. Research has begun to expand to include prevention of nosocomial infections and hospital quality assessment, including medical error and pain management, health services and outcomes research, and ethics. Newer areas also include clinical trials that

focus on conditions common to the hospital, and genetic epidemiology. Since the advent of the modern hospitalist is a relatively new occurrence, there is probably insufficient mentorship at the moment. Improving mentorship may increase academic productivity among hospitalists.

DISADVANTAGES AND OBJECTIONS TO THE HOSPITALIST MODEL

Most objections to the use of the hospitalist model come from the fact that it introduces a purposeful disruption of continuity of care. Each patient transition increases the risk of medical errors and adverse events. In many cases, the hospitalist has no previous knowledge of the patient, making a problem for continuity of care. Hospitalists may be inclined to order more diagnostic tests to establish a baseline status that is already known to the primary care physician. They may also be unaware of a patient's and family's wishes regarding advance directives.

There are two potential models for the hospitalist system: voluntary and closed. In a closed system, only hospitalists have admitting privileges, and ambulatory care physicians are excluded. An open system allows both groups to admit their patients. Concerns have been raised that a closed system may engender bad feelings among community physicians. A potential disadvantage of the hospitalist model is that outpatient physicians will lose their skills in treating inpatients. It is clear, however, that by focusing on outpatient care they can be expected to improve their skills in that area.

Hospitalist expertise is based on the amount of time spent in inpatient care. Some studies point to the fact that improvements in quality occur after years of practicing this model. However, some of these physicians for different reasons, including possibly avoiding burnout, choose to practice part time, which will limit their acquisition of experience and knowledge. It has also been argued that the driving force for this movement has been decreasing costs rather than improving quality.

Inpatient specialists often hold higher levels of medical certification than some hospitalists with no formalized training in hospital medicine. In addition, subspecialist consultants may be apprehensive that hospitalists, due to their familiarity with common inpatient conditions, may produce fewer consultations. A reduction in consultations would be associated with a lower cost of care. Because of the severity of illness in hospitalized inpatients, consultations will continue to occur.

During the transition from outpatient to inpatient and again with the transition back to outpatient status, there is a risk of poor communication between the treating clinician and the receiving clinician. An aspect of this coordination of information flow is the risk of fragmentation of care and the loss of a care coordinator who has comprehensive knowledge about the patient's health. One study found that communication between PCPs and hospitalists leaves significant room for improvement with more than half of

the PCPs not provided with a discharge summary, while almost one quarter had no knowledge that their patient had been admitted. Most likely due to a decreased LOS, patients are also discharged quicker and sicker than before. With appropriate phone calls, and copies of discharge summaries and letters, it is possible to minimize this problem and avoid a significant loss of medical information. This process is unquestionably facilitated by the implementation of electronic medical records.

Generally residents have been very satisfied with the quality of teaching provided by hospitalists, ranking them higher than traditional attending physicians. However, because of their greater expertise in patient work, it can also be expected that hospitalists will threaten the autonomy of residents since they tend to have an increased physical and intellectual presence on the wards.

Evidence suggests that patients accept hospitalists without significant difficulty. Generally they are satisfied with the quality of care and do not perceive significant differences between hospitalist care and that provided by their outpatient clinician or traditional ward attending physicians. It has been suggested that patients may be willing to trade the familiarity of the outpatient clinician for the enhanced availability of their hospitalist. It has also been suggested that patients may no longer be as emotionally bound to their physicians. Many patients have no regular outpatient physician. Furthermore, in the current insurance environment many patients switch health plans and providers with relative frequency.

Increasingly, training programs in psychiatry, medicine, and pediatrics emphasize ambulatory care. This raises questions about the adequacy of preparation of recent graduates for inpatient attending. Training programs may need to be revised to provide appropriate preparation for the hospitalist role.

PROFESSIONAL ORGANIZATIONS

The National Association of Inpatient Physicians (NAIP) was founded after Robert Wachter, MD, MHM, coined the term “hospitalist” in 1996, followed by the commencement of NAIP’s first annual meeting in 1998. NAIP changed its name to the SHM in 2003. SHM continues to be the only medical society devoted entirely to hospitalists and the hospital medicine movement. There is no analogous professional organization for psychiatrists.

SPECIALTY TRAINING AS A HOSPITALIST

Increasingly, training programs are geared toward providing only enough inpatient training so that graduates are capable of caring for acutely ill outpatients. Special tracks may develop to provide adequately trained hospitalists from different medical disciplines. These programs would focus on developing clinician-teachers with the skills and knowledge necessary

for inpatient practice. Sufficient outpatient experience would be provided so the hospitalist would comprehend what level of severity can be managed in the outpatient clinic. In medicine, tracks exist at both the resident and fellow levels.

As training of hospitalists becomes more common, one emerging possibility is the development of two types of residencies. One would emphasize ambulatory care and the other would focus on hospital-based care. This would evolve by the use of different tracks for residents. In psychiatry, this separation would be expected to occur after the first 2 years, which usually emphasize inpatient and consultative work.

FINANCIAL, POLITICAL, AND ORGANIZATIONAL ISSUES

In many hospitals, the reimbursement for inpatient care is such that revenues from hospitalist services are insufficient to provide for the costs of physicians to staff the hospitalist program. In many cases, the benefits to the hospital or, in the academic medical center, to the training program are sufficient to justify additional program support. In academic programs, the training and service obligations are often thought to justify the additional cost.

The willingness of ambulatory care physicians to forgo inpatient care has been heavily influenced by the relatively low compensation. Many have realized that they can increase their clinical income by substituting relatively higher-paying outpatients for the lower-paying inpatients and avoiding the unreimbursed commutes to the hospital.

The low rate of reimbursement for inpatient care complicates implementation of the hospitalist model. Low reimbursement along with a payer mix that often includes uninsured patients means that few hospitalist programs can support themselves solely on professional fees. Further complicating the picture is the fact that clinicians are reimbursed a single professional fee each day despite commonly interacting with the patient several times a day and also working with allied professional staff and families. The net result is, commonly, a significant shortfall in the revenue needed to pay hospitalist physicians at a rate commensurate with community norms for other physicians. Generally hospitals are able to generate about 60 to 80 percent of their salaries by fee-for-service billings and/or capitation dollars. Because of the benefits they provide to hospitals, practice groups, and other sponsoring agencies, these groups commonly supplement the shortfall in professional fee income.

The hospitalist model will have an effect on a hospital's income. Cases that are paid on a per diem basis may produce less income because of the shorter LOS that occurs with hospitalist attending physicians. In a capitated case rate in which the hospital is paid according to diagnosis-related groups or when a medical group is paid on a contract rate, there are clear benefits from the reduced lengths of stay and reduced use of ancillary

tests associated with hospitalists. In such situations, the savings that result from the use of hospitalists may exceed the salary support provided by the institution. This significant savings is the rationale for supporting the salaries of hospitalists.

To avoid the perception of a conflict of interest, most institutions avoid paying hospitalists for reducing the utilization of services. However, many programs have created incentives related to outcomes measures and patient satisfaction.

HOSPITAL MEDICINE AS A NEW SPECIALTY

Fellowships are currently very limited. Typically they focus on training clinical educators and provide a curriculum on teaching theory, curriculum development, and evaluations methods and opportunities to strengthen teaching skills, along with opportunities to receive feedback about teaching activities. Some also provide instruction on research methods and encourage clinical research activities.

In many institutions, hospitalist positions are staffed by recent graduates and are intended to be transient. In these settings, the opportunities for long-term career development are unclear. In contrast, academic settings that intend to develop a stable cadre of hospitalists provide protected time for promoting scholarly activities and to work on improvements in the hospital system.

Traditionally, academic medical centers have been challenged in attempting to recognize the contributions of clinician-educators. It is these accomplishments that are the basis of being rewarded with promotion through the professorial ranks. Customary measures of accomplishment include publications and grants, along with other objective evidence of regional, national, and international recognition. The challenge has been to apply these assessment standards to the provision of clinical care or teaching ability. Traditionally, teaching capabilities have been assessed by peer and student evaluations. More challenging is the assessment of clinical excellence.

Hospitalists, similar to emergency medicine physicians and critical care specialists, are site-defined specialists. That is, their area of expertise is not defined by an organ or procedure but by where they deliver care. As such, they care for patients with a variety of illnesses and ages within that location.

In medicine, there are other attributes associated with a subspecialty. There is a large group that identify themselves as hospitalists. There is a professional society with regular meetings. There is a clinical textbook. Clinicians agree that there is a core skill set and body of knowledge.

FUTURE DIRECTIONS

Many of the changes noted in the medical community are applicable to

psychiatry. Outpatient providers are becoming unfamiliar with treating acute cases of agitation and severe mental illness as well as complications of the abuse of substances and legal requirements for involuntary inpatient treatment. Moreover, similar forces are influencing the practice of inpatient psychiatry. Patients are sicker and there is significant pressure to decrease LOS. In addition, the oversight on hand-offs and development of after care plans for treatment and safety are also currently growing.

The hospitalist movement is rapidly growing in psychiatry, medicine, and pediatrics. In the latter two areas, where the numbers are more clearly known, the number of hospitalists is larger than those in many board-certified subspecialties. The hospitalist movement, similar to prior trends in emergency medicine and critical care medicine, has a defining environment and a correspondingly separate body of knowledge and skill set.

Hospitalists can be expected to have an increasing role in improving care for hospitalized patients. Due to their extensive familiarity with the workings of the hospital, they are able to identify system problems. They are able to work with administrators and nurses to assure that necessary equipment is available and proper protocols are in place to facilitate quality care.

Hospitalists have developed related research areas. Typically, these are outcomes-oriented studies that focus on hospital care. Examples include quality of care, cost containment, resource utilization, and LOS. Such studies are typically conducted by academic hospitalists at university medical centers.

Currently, the term hospitalist does not refer to a subspecialty. In medicine, there are increased efforts to make the hospitalist a subspecialty with residency requirements and board certification. The future may see similar trends in psychiatry. Currently, there is a subspecialization in psychosomatic medicine or consultation liaison psychiatry. The same forces that led to this process may lead to the establishment of the psychiatric hospitalist subspecialty.

Unresolved questions about the future of the hospitalist movement remain. For example, should hospitalists be exclusively inpatient providers, or should there have some component of outpatient or consultative work? Should all inpatient care be delivered by hospitalists, or is there also a role for outpatient clinicians in the inpatient setting? How will hospitalists be compensated, and who will employ them? What is the optimal training to be a hospitalist?

As the acuteness of inpatients increases, more and more the demands of the job resemble those of emergency medicine department attending physicians or intensive care specialists and less those of outpatient primary care clinicians. Whereas current compensation for inpatient and outpatient physicians is similar, it can be expected that compensation for hospitalists

will also increasingly resemble that of intensivist or emergency medicine physicians.

Unique to psychiatry, there is a certain lack of recognition of the hospitalist physician as a differentiated area of practice. In other areas of medicine, the hospitalist role is widely recognized. Despite a long history of the hospitalist role in psychiatry, a role that dates back to the origins of the profession, few psychiatrists who work predominately in a hospital setting would consider themselves hospitalists. They are more likely to call themselves unit directors or ward directors. As the recognition of psychiatric hospitalists increases, a further appreciation of the similarities and differences between psychiatric hospitalists and hospitalists in other areas of medicine can be expected. One can imagine, for example, that the activities of colleagues in other areas of medicine will provide guidance for improving the career path for hospitalists, starting with fellowships and progressing through various levels of seniority following training.

Under recognized and poorly characterized, psychiatric hospitalists could be at the forefront of interactions with third-party payers that seek to alter the nature of psychiatric hospitalization. Improved recognition of their role, clarity of their training pathway, and identification of the stages of professional career development will promote this career pathway to trainees and ensure that there are well-trained individuals providing expert inpatient care.

References

Auerbach AD, Nelson EA, Lindenauer PK, Pantilat SZ, Katz PP. Physician attitudes toward and prevalence of the hospitalist model of care: results of a national survey. *Am J Med*. 2000;109:648–653.

Bess J, Casher M. The academic psychiatric hospitalist: a model for treatment and training in the academic inpatient setting. *J Psychiatr Pract*. 2011;17(2):145–151.

Blumenshine P, Lenet AE, Havel LK, Arbuckle MR, Cabaniss DL. Thinking Outside of Outpatient: Underutilized Settings for Psychotherapy Education. *Acad Psychiatry*. 2017;41(1):16–19.

Cammell P. Emergency psychiatry: a product of circumstance or a growing sub-speciality field? *Australas Psychiatry*. 2017;25(1):53–55.

*Craig DE, Hartka L, Likosky WH, Caplan WM, Litsky P. Implementation of a hospitalist system in a large health maintenance organization: the Kaiser Permanente experience. *Ann Intern Med*. 1999;130:355–359.

Davis KM, Koch KE, Harvey JK, Wilson R, Englert J. Effects of hospitalists on cost, outcomes, and patient satisfaction in a rural health system. *Am J Med*. 2000;108(8):621–626.

Diamond HS, Goldberg E, Janosky JE. The effect of full-time faculty hospitalists on the efficiency of care at a community teaching hospital. *N Engl J Med*. 1998;129(3):197–203.

Donisi V, Tedeschi F, Salazzari D, Amaddeo F. Pre- and post-discharge factors influencing early readmission to acute psychiatric wards: implications for quality-of-care indicators in

psychiatry. *Gen Hosp Psychiatry*. 2016;39:53–58.

Egan JM, Webber MGT, King MRD, Boyd M, Preez-Wilkinson G, Brock D. The hospitalist: a third alternative. *Med J Aust*. 2000;172:335–358.

Freese RB. The Park Nicollet experience in establishing a hospitalist system. *Ann Intern Med*. 1999;130:350–354.

Goldman L. The impact of hospitalists on medical education and the academic health system. *Ann Intern Med*. 1999;130:364–367.

Haftel HM, Bozynski ME. Changing teaching for changing times: the effect of a hospitalist program on the education of students. *Acad Med*. 2000;75:521.

Hauer KE, Wachter RM, McCulloch CE, Woo GA, Auerbach AD. Effects of hospitalist attending physicians on trainee satisfaction with teaching and with internal medicine rotations. *Arch Intern Med*. 2004;164:1866–1871.

Kelley MA. The hospitalist: a new medical specialty? *Ann Intern Med*. 1999;130:373–375.

Lindenauer PK, Pantilat SZ, Katz PP, Wachter RM. Hospitalists and the practice of inpatient medicine: results of a survey of the National Association of Inpatient Physicians. *Ann Intern Med*. 1999;130:343–349.

Melzer SM, Molteni RA, Marcuse EK, Rivara FP. Characteristics and financial performance of a pediatric faculty inpatient attending service: a resource-based relative value scale analysis. *Pediatrics*. 2001;108:79–84.

Pantilat SZ, Alpers A, Wachter RM. A new doctor in the house ethical issues in hospitalist systems. *JAMA*. 1999;282:171–174.

Paris J. *Prescriptions for the Mind: A Critical View of Contemporary Psychiatry*. New York: Oxford University Press; 2009.

Saint S, Fowler KE, Krein SL, et al. An academic hospitalist model to improve healthcare worker communication and learner education: results from a quasi-experimental study at a Veterans Affairs medical center. *J Hosp Med*. 2013;8(12):702–710.

Schroeder SA, Schapiro R. The hospitalist: new boon for internal medicine or retreat from primary care? *Ann Intern Med*. 1999;130:382–387.

Sloan PA, Asghar-Ali A, Teague A, Body E, Kunik ME. Psychiatric hospitalists and continuity of care: a comparison of two models. *J Psychiatr Pract*. 2010;16(3):170–176.

*Srivastava R, Stone BL, Murphy NA. Hospitalist care of the medically complex child. *Pediatr Clin North Am*. 2005;52(4):1165–1187, x..

Talbott JA, Granet RB. Careers in psychiatry: options for the future. *Compr Psychiatry*. 1984;25:263–277.

Tatreau JR, Harris S, Sheitman B, Steiner BD. Cardiometabolic Assessment, Diagnosis, and Treatment of Chronic Medical Illnesses During an Inpatient Psychiatric Hospitalization: Colocated Medical Care Versus Treatment as Usual. *Prim Care Companion CNS Disord*. 2016;18(6).

Van Walraven C, Seth R, Austin PC, Laupacis A. Effect of discharge summary availability during post-discharge visits on hospital readmission. *J Gen Intern Med*. 2002;17:186–192.

Wachter RM. An introduction to the hospitalist model. *Ann Intern Med*. 1999;130:338–442.

Wachter RM, Goldman L. The emerging role of “hospitalists” in the American health care system. *N Engl J Med*. 1996;335:514–517.

Wachter RM, Goldman L. The hospitalist movement 5 years later. *JAMA*. 2002;287:487–494.

*Wachter RM, Katz P, Showstack J, Bindman AB, Goldman L. Reorganizing an academic medical service: impact on cost, quality, patient satisfaction, and education. *JAMA*. 1998;279:1560–1565.

Ward M, Schwartz A. Challenges in pharmacologic management of the hospitalized patient with psychiatric comorbidity. *J Hosp Med*. 2013;8(9):523–529.

White HL, Glazier RH. Do hospitalist physicians improve the quality of inpatient care delivery? A systematic review of process, efficiency and outcome measures. *BMC Med*. 2011;9:58.

▲ 58.6 Psychiatric Rehabilitation

RADU V. SAVEANU, M.D., ROSIE E. CURIEL CID, PSY.D., AND DAVID LOEWENSTEIN, PH.D.

AN OVERVIEW OF PSYCHIATRIC REHABILITATION

Psychiatric rehabilitation was once thought of as merely a series of psychosocial interventions that were aimed at augmenting psychopharmacological approaches. It is now recognized that true psychiatric rehabilitation is extremely integrative, involving an interdisciplinary approach. Importantly, the goal of psychiatric rehabilitation is not merely the remission of symptoms but helping the patient to assume their highest potential for the most independent and meaningful function within the community. Optimal psychiatric rehabilitation employs a person-centered approach with a focus on individual dignity, goal attainment, and a collaborative therapeutic relationship. There is also an emphasis on marshaling necessary environmental supports that will improve the efficacy of interventions within the home environment, community, and place of vocation. These community supports may involve professionals, paraprofessionals, natural community supports, or environmental modification.

Successful psychiatric rehabilitation is accomplished by a comprehensive understanding of the patient's (1) psychiatric condition, (2) treatment regimen, (3) social supports, and (4) contextual cues in the environment and an understanding of a patient's strengths as well as his or her limitations.

Psychiatric rehabilitation is not only limited to severe mental illnesses (SMIs) such as schizophrenia, bipolar disorder, or schizoaffective disorder but also extends to persons with treatment resistant major depression, severe anxiety disorders, and substance abuse (alone or comorbidly). Instead of a focus on remission and symptom control, there has been an increasing emphasis on improving the long-term outcomes of persons with SMI and incorporate constructs as increased autonomy, well-being, financial independence, and quality of life. Many persons with schizophrenia may obtain control over positive symptoms such as

hallucinations and delusions but still have to deal with negative symptoms that have adverse effects on communication, initiative, and interpersonal interactions. Neurocognitive symptoms of schizophrenia have been implicated in reduced ability to function in the home, workplace, and the community. However, theses cognitive effects extend well beyond issues with memory, processing speed, and executive dysfunction. Indeed, some of the greatest difficulties with community functioning are related to deficits in social cognition including perception and processing of affect, knowledge of social rules and conventions, inferring another person's intentions as well as attributing causality for positive or negative outcomes. For example, someone with paranoid schizophrenia may have a particularly difficult time properly inferring another person's motivations and intentions. Further, there may be difficulties with attributional biases in which negative outcomes are attributed to external causes with poor insight into one's role in certain events. Lack of insight into one's own processes and misinterpreting the emotions and intentions of others can lead to isolation, lack of meaningful interpersonal contact, difficulties dealing with others in the community, and problems with interacting with others in the workplace.

Psychiatric rehabilitation strategies are increasingly holistic, respect the autonomy of the individual in decision making, and specify incremental, graduated goal attainment with the goal of optimal function in the community. Since functional outcomes and quality of life are such important aspects of psychiatric rehabilitation, it will be important to precisely define different aspects of these different outcomes.

REHABILITATION IN THE MEDICAL CONTEXT

In the past, there was sometimes a disconnect between the concepts of medication management and/or treatments (e.g., ECT) and the concept of psychiatric rehabilitation. The former was associated with a medical model and emphasis on somatic approaches versus while the latter was associated with psychosocial treatments such as social skills training, vocational remediation, and building environmental supports. The medical model was seen as quite distinct from psychiatric rehabilitation, the latter rooted in a behavioral approach. However, with the increasing recognition that many psychiatric disorders need to be conceptualized as chronic conditions with waxing and waning of symptoms over time, there is little doubt that there is a need to integrate both medical and psychosocial approaches. For example, a significant issue in treatment of schizophrenia and bipolar affective disorder is noncompliance with medication management. As a result, a major focus of many current psychiatric rehabilitation approaches is to focus on adherence to medical treatments and medication regimens as a primary focus of psychiatric rehabilitation. This is typically done using advanced therapeutic and behavioral techniques as well as the use of

technological innovations in medication management and electronic reminders. There may also be an emphasis on cognitive and behavioral training with regard to prospective memory, or "remembering to remember an intended action." Further, it is well accepted that medication compliance is likely going to be higher with concomitant efforts to build supports in the environment, prepare the individual to deal with environmental stressors, and to promote a physical and emotionally healthy lifestyle. For example, an individual with severe obsessive-compulsive disorder (OCD) may be prescribed specific medications to help them manage their condition, but this can be augmented by two empirically supportive cognitive-behavioral strategies that have been used to effectively treat OCD (graduated exposure and response prevention). However, OCD symptoms may also be exacerbated by stress. Therefore, treatments may also focus on assertiveness training, social skills training, and enhancing both social and environmental supports. This increasingly more integrative approach requires the coordination of a number of health care disciplines.

UNIFYING CONCEPTS OF REHABILITATION AND TREATMENT

It is now clear that the most successful approaches to psychiatric rehabilitation involve the *integration* of medical and a variety of psychological and psychosocial approaches. As described below, psychiatric rehabilitation is more person-centered. That is, instead of a more paternalistic model in which the patient adheres to the wishes of treatment providers, there has been a greater emphasis on shared decision making on behalf of providers, patients, and their families. This person-centered approach not only provides greater shared collaboration and decision making between the patient, family and treatment team but also provides shared decision making that increases the probability of investment and compliance. This individualized approach to treatment also emphasizes targeted goals that are unique for an individual patient with the appropriate admixture of therapies that are specific to the person's needs. It is of essential importance for both the patient and the treatment team to (a) identify targeted short-term and long-term goals that are operational and obtainable; (b) plan a sequence of interventions and supports that are appropriate to an individual's stage of recovery; (c) build in social and environmental supports, stress inoculation, coping strategies, and other continuing therapies that will reduce the chances of relapse; and (d) to continue a collaborative treatment milieu between the patient, family, and care providers that provide the greatest opportunity for mutual respect and decision making. The shift from doing something for the patient to collaborating in their recovery and well-being means that treatment becomes educational as well as treatment focused.

PSYCHIATRIC REHABILITATION AND INCORPORATION OF PERSON-CENTERED

RECOVERY MODELS

As in other fields of medicine, psychiatric rehabilitation has adopted the notion that individualized and collaborative methods of care improve the prospects for recovery. The traditional view of the psychiatrist as the sole, responsible, and authoritative clinician has been replaced with the model the multidisciplinary or interdisciplinary team approach. The *person-centered* approach builds upon this by actively engaging the patient and their family in shared decision making regarding medication and incorporates their personally relevant goals and unique psychosocial circumstances into the formulation of treatment plans.

The person-centered approach is in line with a more optimistic vision for the long-term *recovery* of the seriously mentally ill as it offers opportunities to generate higher levels of motivation by patients. These progressive changes in ideology have mostly been championed through the leadership of federal, state, and local agencies. Many training programs are more cognizant about teaching emerging practitioners about the person-centered, recovery-oriented mode of service delivery, which is expected to continue to intensify and develop in the coming years given that this approach is expected to strengthen the therapeutic alliance, promote greater motivation and collaboration, and achieve better outcomes.

HOW CAN RECOVERY BE DEFINED?

Sustained recovery refers to long-term abstinence or remission of symptoms and normal psychosocial functioning. It refers to a measurable set of processes, stages, and outcomes that are grounded on mediating and moderating variables that facilitate its attainment and maintenance. That is, recovery is an outcome of the process of recovering.

Individuals can take many pathways to recovery, depending on the varied factors that influence the process, such as personal attributes, social environment, continuity and quality of treatment, and subjective experience.

Examples of attributes and experiences that may be associated with individuals who are progressing toward recovery include hope, destigmatization, empowerment, self-acceptance, insight, awareness, collaboration with professionals, sense of autonomy and self-control, and participation in self-help and patient-run programs. Subjective experiences and attributes not only can mediate the process leading to recovery, but they also can sustain recovery. For example, a treatment or self-help program may engender a feeling of empowerment, which can be instrumental in motivating a person to sustain treatment and rehabilitation until recovery can be achieved. Once recovery has been achieved, empowerment may be even more firmly experienced through improved functioning and mental and emotional stability.

Similarly, having hope for future improvements in quality of life may be reinforced by the enthusiasm and collaborative alliance embedded in a therapeutic relationship. Many patients who have written first-person accounts of their recovery have pointed to the importance of a long-term relationship with a practitioner who refused to give up hope for the patient's ultimate improvement and recovery. Hope can serve to fuel motivation for change and for active participation in clinical services or self-help, which are stepping stones toward recovery. Subjective experiences of empowerment and hope continue to grow after a state of recovery is attained, contributing to further strengthening and protecting individuals from stress-induced relapse. Encouragement and hope also enhance efficacy expectations that empower the individual to assume greater control and responsibility over their lives.

EMPIRICALLY BASED MODELS OF PSYCHIATRIC REHABILITATION

Psychiatric rehabilitation has evolved substantially over the last several decades. Previous descriptions of interventions used in psychiatric rehabilitation included sole descriptions of token economy and other reinforcement programs for long-term patients in psychiatric hospitals, client-centered identification of personal goals, and various sheltered and segregated work and social programs in the community. Comprehensive services that incorporated personally meaningful adjustment to daily living in normal community settings were nonexistent.

Over 50 years of intervention research has led to the expectation that even patients with the most severe mental disabilities can gain from empirically validated, behavioral interventions to improve their daily, and community functioning. Evidence-based therapies are now in widespread use in psychiatric rehabilitation.

Fulfillment of these expectations for therapeutic improvement has been facilitated by political advocacy, destigmatization of mental illness, and self-help activities of consumer organizations such as the National Alliance for the Mentally Ill, the Schizophrenia Fellowship, and networks of mental health clients. In 1991, the Americans with Disabilities Act was signed into law, requiring equal employment opportunities and reasonable employer accommodations for workers with mental illness.

Recovery is now more than ever a realistic goal for many persons with mental disabilities with enduring biological vulnerabilities such as schizophrenia; however, it is necessary that the full range of evidence-based biopsychosocial services be available in a continuous, comprehensive, coordinated, consumer-oriented, and competently delivered system of care. Despite the field of psychiatry producing a wide array of evidence-based treatments and services, public and private funding has eroded for organizing and delivering these services with high fidelity.

Despite the many challenges that are confronted in the field of psychiatry, opportunities to offer effective rehabilitation to persons with mental disabilities have never been better than at the present time. With that in mind, psychiatric rehabilitation is confronted with the major challenge of fiscal limitations, inadequate resources, and poor coordination.

MAKING EVIDENCE-BASED TREATMENT AND SERVICES ACCESSIBLE

There have been increasing efforts to reduce the barriers to effective dissemination and promote the adoption of evidence-based practices (EBPs) in psychiatry using a package of user-friendly treatments in the form of "toolkits" that are designed to make the collection of evidence-based treatments more accessible to clinicians. These efforts such as the project implementing EBPs have mostly been sponsored by federal, state, and foundation agencies. Standardized guidelines for the application of the evidence-based treatments are provided in workbooks that are included in the toolkits, as are training materials and procedures to provide means for assuring that the treatments will be used with the highest levels of fidelity.

The EBP project was conducted in three phases. The first phase focused on developing the kits, which are now available from the Substance Abuse and Mental Health Services Administration (SAMHSA) website: http://mentalhealth.samhsa.gov/cmhs/communitysupport/toolkits/. In the second phase, pilot demonstrations using the kits at 50 sites in 8 states were completed in 2007. The toolkits contain implementation resource information as well as information on how the EBPs are related to improving the functioning, quality of life, and recovery of patients. The third phase focused on a broad national dissemination of the kits. Given the over $2 billion allocated by public health authorities to make EBTs available, and the substantial implications to science and practice, evaluations have been and are currently being carried out to measure the success of the toolkits. A review conducted by Barlow and McHughs found that the studies evaluating the efficacy of dissemination and implementation programs are limited; however, efforts are continuously being made to improve these procedures. At the time of this review, critical components that facilitated the adoption, competence, and sustainability of EBTs were unclear.

At present, SAMHSA provides a plethora of toolkits with EBTs for various disorders free of charge, many of which are available on DVD or digital download at http://store.samhsa.gov/list/series?name=Evidence-Based-Practices-KITs.

The ultimate purpose of the EBP toolkits—"to promote the implementation of evidence-based psychosocial treatments in routine mental health settings"—is a goal that has great potential to improve the lives of people with serious mental disorders. However, implementing and

directly monitoring fidelity to the model can be a costly and time-consuming process. A balance needs to be achieved between monitoring the accurate implementation of each EBP while promoting dissemination of the EBPs in a way that allows the adopting agencies to tailor the innovation to fit the specific function, mission, patient population, local resources, and other contingencies they face. Continuous quality improvement has been used effectively as an approach to implement EBPs at community-based settings.

At the level of the community agency implementing EBPs, specific expectations for clinician performance and ongoing program evaluation, as well as a culture change is required to implement and sustain the provision of EBPs. Thus, evaluating the cost-effectiveness of the EBP toolkits using a longitudinal time frame and a methodology that captures the organizational characteristics and the total expenditures to the systems of care in using the EBPs for the patients within their catchment areas is essential to the long-term viability of these programs.

With improved and more reliable methods of diagnosis and symptom and functional assessment and an empirically validated definition of recovery, the population and needs of patients requiring psychiatric rehabilitation are becoming better defined. Methods for determining readiness for rehabilitation also have been validated, enabling administrators and practitioners to be more efficient in their use of rehabilitation programs and resources. Because of its growing technology and its conceptual and empirical base, psychiatric rehabilitation has attracted an interdisciplinary group of practitioners and investigators whose work has contributed to its evolving scope and stature as a subspecialty of the mental health and rehabilitation professions.

The challenges to community-based care for persons with serious and chronic mental illness have been met partially by new initiatives at the *clinical* and *policy-making* levels. New interventions have been designed and validated, such as psychoeducational and behavioral family therapies, social skills training, cognitive-behavioral therapy (CBT) for medication-resistant psychotic symptoms, intensive case management models for assuring continuing care, supported housing, and employment services. At the level of psychiatric politics, consumers and families have become empowered by the advent of peer-directed self-help organizations; the use of the mass media for education of the public about mental disorders and the possibilities for recovery, including noted political, sports, and entertainment figures going public with their mental disorders and successful treatment; legislative lobbying, assertive advocacy, peer support, and involvement in treatment by families; efforts to gain parity in funding for mental health services; and the Americans with Disabilities Act.

Despite advances in assessment and treatment, the implementation of psychiatric rehabilitation in the public sector has been frustrated by

inadequate funding and resources, lack of trained personnel, limited dissemination of the new technology, insufficient affordable housing, unattractiveness of voluntary treatment, and a lack of administrative and financial integration of state and local community mental health and substance abuse services. The consequences of those system-wide problems have meant that too few individuals with mental disabilities are receiving state-of-the-art services and too many thousands are living on the city streets or in jails and prisons.

REHABILITATION IN THE PRODROMAL, ACUTE AND STABLE PHASES OF SEVERE MENTAL ILLNESS

Effective treatment and rehabilitation of persons with mental illness require clinical services that are linked to the phase of their disorder. For most stress-related, recurring, and disabling mental disorders, the following phases are relevant for practitioners' initiating and/or adjusting their treatment:

- Motivation and engagement in treatment
- Prodromal
- Acute
- Stabilizing
- Stable
- Recovery
- Refractory

There is broad-based variability in the manner in which individuals experience a particular phase in that no two individuals experience any phase in the same way. Individual differences influence each phase in terms of severity, duration, degree of disability, resolution, course, and boundaries. The following experiences depict two persons with schizophrenia in the *stabilizing phase*. This serves to exemplify the need to adapt treatments to individual diversity within a particular phase of illness.

Mr. F, a 23-year-old male was admitted for involuntary psychiatric hospitalization after he locked himself in his dorm room refusing others entry because he was afraid that someone was out to harm him. When persons entered the room he was in the bathtub rocking back and forth and talking incoherently. After 5 days of stabilization, he was released from the hospital and returned home to his family but was still very anxious about being alone, difficulties with voices in his head telling him he was "doomed" and he was extremely concerned that his movements were being monitored. Because he was resistant to oral medications (he thought they might be tampered with and might harm him), he was prescribed an injectable antipsychotic, which helped to stabilize his symptoms. Upon returning to school, he completed a semester taking a reduced course load and living at home. However, during the next semester, he insisted on taking a full course load and started to use marijuana and other drugs to help to manage the resultant distress. He began to feel very hypervigilant and felt that others in his class were judging him critically. This further reduced his concentration and he started doing poorly in class setting up a vicious cycle where he was readmitted with extreme paranoia and had to again drop out of school. At this point, treatment became integrative and, in

addition to adjusting his psychiatric medications, he received supportive psychotherapy and was engaged in several cognitive-behavioral and skills-building training that provided him ongoing support and structure. He learned to be more assertive and comfortable in his interactions through assertiveness social skills training and relaxation/stress management techniques, which greatly helped his coping skills. Mr. F eventually returned to school, earned his degree and is currently in a supportive work program.

Ms. L, age 32 years had persistent obsessions and compulsions that increasingly became worse, particularly when she was promoted from an associate to a partner in a prestigious law firm. She became increasingly convinced that she was going to succumb to a terrible disease and washed her hands hundreds of time per day because she was afraid that germs on her skin would contaminate her or that she could pass them to other people. She started to avoid touching door handles, pens that others had touched, and even refused to shake hands with new clients. She had a new child who was 9 months old and was terrified that he would become ill and would wash him many times at night and disinfect his playroom. She fired several nannies because she feared that they would pass illnesses to the child and the family. She got into more constant fights with her husband who thought that she was being unreasonable and increasingly had difficulties with depression and generalized anxiety. She was admitted when the partners called a meeting with her to discuss their concerns about her performance and she became extremely anxious and acutely suicidal.

Following discharge from the hospital 3 days later, she was placed on a combination of Luvox and other psychotropic medications and immediately referred to a therapist who focused on CBT approaches. It took several weeks before she was adequately stabilized to engage in interventions for OCD (graduated exposure and response prevention). In addition, couples' work with the patient and her husband was conducted by another therapist that helped the husband better understand her illness, support her behavioral homework, and helped the couple better communicate. OCD can involve a waxing and waning of symptoms. As a result, Ms. L, who is now back to work, is involved in maintenance therapy where she sees her behavioral therapist 2 times a month, but more frequently if needed.

These two examples illustrate the need to conceptualize the treatment of these two individuals into several components (a) stabilization; (b) dealing with ongoing acutely psychotic, mood, or anxiety symptoms; (c) using the therapeutic alliance to provide the maximum motivation for treatment; (d) using integrative therapeutic support to augment and strengthen the possibility of positive treatment results; (e) ongoing assessment and strategies aimed toward relapse prevention; and (f) reinforcement and maintenance of gains.

The framework for phase-linked treatment and rehabilitation is depicted in Figure 58.6–1. It should be noted that EBPs are the backbone of services generally found to be useful for patients as they traverse the various phases of their disorder. There is considerable variability in how fast each individual moves from one phase to another. Some may even skip a phase, becoming stable within 1 week of an acute relapse, without the gradual stabilization of symptoms and functioning that is more frequently the case. People also move back and forth between phases as stressors impinge on their vulnerability and the individual has not compiled sufficient protective factors. The process is not nearly as linear as shown in the Figure 58.6–1.

PRODROMAL PHASE

Early identification, engagement, and treatment are the mainstays of this phase, which may be associated with individuals who are at risk of their first episode of SMI or are experiencing the warning signs of a recurrence of their disorder. Preventive interventions for young people at high risk of developing a psychotic illness remain in the early stages of development. The sensitivity and specificity of identifying the early signs of a developing psychosis have not been fully validated; moreover, psychosocial and pharmacological interventions have not been empirically documented as effective in averting schizophrenia or other psychotic disorders when provided to young people with what may appear to be prodromal signs and symptoms.

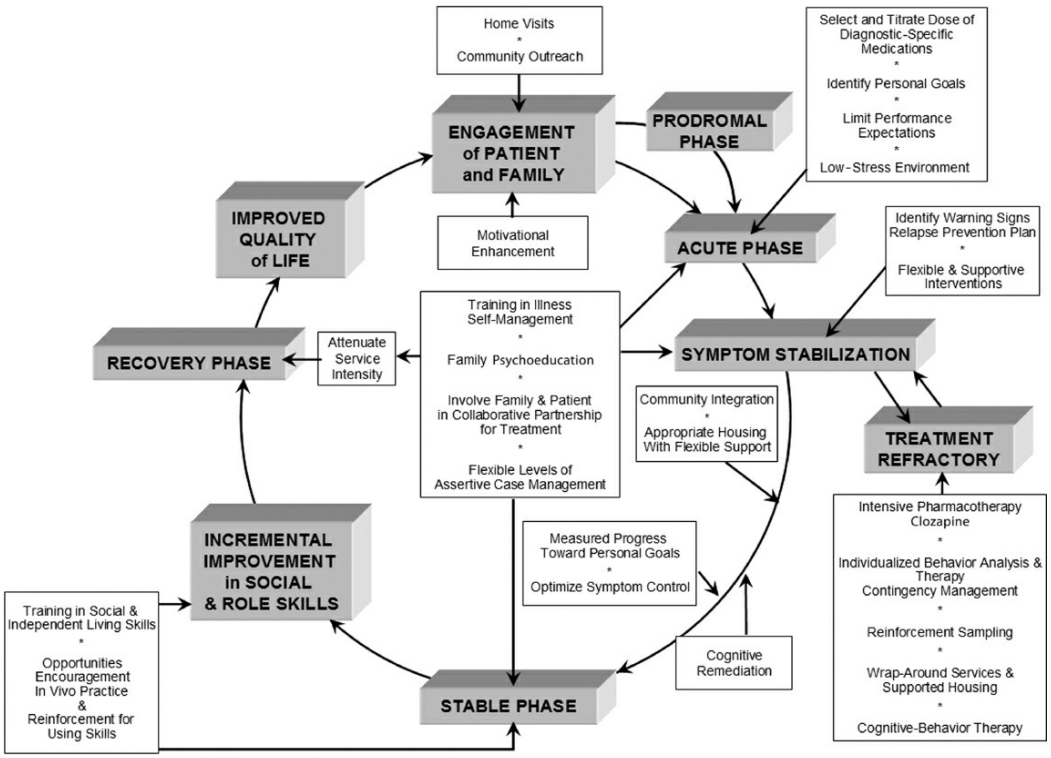

FIGURE 58.6–1. Multidimensional framework for phase-linked treatment and rehabilitation.

PRODROMAL SIGNS AND SYMPTOMS PRIOR TO THE FIRST EPISODE OF SCHIZOPHRENIA

Youth who are possibly moving toward their first episode may be experiencing sleep problems, strange, unusual, unwanted, and intrusive thoughts, or depression, irritability, hypomania, or anxiety. Other prodromal signs and symptoms include ideas of reference, paranoid worries and fears, significant social and emotional withdrawal, suspiciousness, apathy and loss of drive, disturbances of attention and concentration, functional deterioration, odd behavior, perceptual

abnormalities, and derealization. These young people almost always refuse to accept professional intervention, preferring to think that their condition is transient and will dissipate without psychiatric treatment. Of course, the stigma of mental illness is a major deterrent to obtaining treatment.

For these individuals and their families, a patient, nonconfrontational stance is desirable. Because it is not even certain that they are moving to an acute episode, it is necessary to monitor the presenting symptoms and how they might be interfering with the individual's current life. Then, depending on the specific problems that are interfering with the individual's reaching his or her goals, the psychiatrist points out that treatment may very well help to overcome the obstacles to progress toward these goals. This approach is called *motivational interviewing* and represents a strategy for engaging patients in treatment. It may be necessary for the practitioner to ally with the family so that the latter can join in the motivational enterprise. The treatment approaches to this early, developmental phase of schizophrenia are delineated in Table 58.6–1.

SYMPTOMS PRODROMAL TO A RELAPSE

Because each psychotic episode places the individual at greater risk of prolonged psychosis, disability, and refractoriness to treatment, prevention of relapse has been assigned a high priority. It is often the case that the individual with recurrent illness does not recognize or acknowledge his or her symptoms; in this case, it is best to simply work on building a relationship with the patient that will eventually pay dividends for active treatment once trust has been achieved.

If possible, prior to a recurrence, the practitioner should provide illness management services that include training in recognizing and monitoring early warning signs of relapse together with developing an emergency plan for preventing full-blown relapse. This educational effort is important because repeat episodes of schizophrenia, major depression, bipolar disorder, and other SMIs carry a poorer prognosis and a higher risk of:

- More frequent acute episodes in the future
- Higher likelihood of nonadherence to treatment
- Greater difficulty treating each successive episode
- More serious disruption of social, occupational, educational, and family roles
- Greater likelihood of stigma and family isolation
- More frequent involvement of police, involuntary admission to hospital, homelessness, or jail
- Considerably greater financial and emotional burden for the family.

Table 58.6–1.

Interventions for the Prodromal Phase of Schizophrenia in Young Persons Not Yet Diagnosed with the Disorder

1. A psychoeducational approach is directed to the patient, family, other caregivers, primary care physicians, school counselors, and staff at college counseling centers. The psychoeducation must be keyed to the particular needs, readiness, and role of the individuals who are in a supportive relationship with the individual who is experiencing the prodromata to a serious mental illness. Outreach and liaison with "gatekeepers" to mental health services—primary care physicians, clergy, schoolteachers, school counselors—is important for mental health facilities and solo practitioners to mount a successful effort in identifying young individuals at high risk of a serious mental disorder. There should be ongoing liaison, education, and consultation with the family and referring professionals and agencies.
2. Nonstigmatizing, age-appropriate, culturally congruent services should be offered in a youth-friendly, nonclinical environment.
3. Specific treatment for symptoms and behavioral problems that are interfering with the young person's goals and functioning should be adopted:
   a. Motivational interviewing.
   b. Stress management.
   c. Cognitive-behavioral therapy or interpersonal therapy for anxiety and depression.
   d. Cognitive-behavioral therapy for prepsychotic symptoms.
   e. Supportive relationship that permits open discussion of the young person's concerns, worries, symptoms, and progress toward goals.
   f. Antipsychotic medication at low dose for symptoms that do not respond to psychosocial interventions and continue to impede the young person's functioning.
4. There should be frequent re-evaluation of the patient's clinical status to determine whether a schizophrenia prodrome is actually operating.

EARLY INTERVENTION AND RELAPSE PREVENTION

One component of illness management is teaching patients and their families to recognize the prodromal signs and symptoms of relapse, monitor them, bring them to the attention of the mental health provider, and follow through with their emergency plan to avert relapse.

An acute episode of illness including relapse for persons who have had previous episodes frequently follows prodromal signs and symptoms. This holds for mood disorders as well as disorders within the schizophrenia spectrum. Regardless of where a patient resides during an acute episode of illness—crisis home, family home, inpatient unit, or partial hospital—the major task of the facility and staff is to mitigate overstimulation and provide a safe, secure, supervised, and predictable environment. This is accomplished by offering clear expectations for daily activities, a posted schedule of daily events, discharge criteria, standards for each privilege level, and a structured routine and treatment environment. Staff experienced in working with acutely ill patients offer individualized, clear expectations for behavior and low task demands that are not unnecessarily burdensome but are flexible in terms of how much social interaction is prescribed. In addition, the staff should offer low-key planned and scheduled activities, orientation, and guidance on the rules of the facility by posting visual aids on the bulletin board and white boards. One staff

member should be designated for helping each patient conduct problem solving and offering supportive therapy at all hours.

In this phase, pharmacological and behavioral treatments of acutely ill patients are predicated on stabilizing or eliminating symptoms and improving cognitive status so that the patient can move toward focusing on the development of skills and utilizing the social and professional supports that will help him move into the next phase of treatment.

Most patients in the acute phase can begin learning illness management skills. During the acute phase, it is essential to customize skills training and support to symptomatic and cognitively impaired individuals; thus, in educating, it is very important to use strategies such as motivational enhancement, repetition, responsive reading, modeling, role playing, and problem solving. It is also necessary to be flexible in using relevant and culturally appropriate communication in educating patients about the value of medication.

CONTINUITY OF CARE FROM HOSPITAL TO COMMUNITY

Patients are rarely sufficiently stabilized by the time they are discharged due to increasingly shortened inpatient hospital stays. Unfortunately, continuity of care offered on most discharge plans consists of providing the patient and family members with the contact information of providers or available community mental health centers who will accept persons with serious and persistent mental illness. Predictably, numerous studies have confirmed the frequent lack of follow-through from inpatient to outpatient care.

Many obstacles stand between inpatient care and the fluid provision of and engagement in outpatient services. First, the patient is struggling with persistent psychotic, anxiety, or depressive symptoms after discharge from inpatient care. These symptoms, as well as cognitive deficits common in serious mental disorders, interfere with the patient taking the initiative to set up an aftercare appointment. When hospital clinicians fail to involve family in psychoeducation regarding their loved one's illness and need for continuing care, family members may have the mistaken belief that definitive treatment has been completed in the hospital and fail to encourage follow-up services. They may even suggest that their relative discontinue medication because they are not aware of the prophylactic value of pharmacotherapy and may have the mistaken belief that antipsychotic medication is addictive.

Numerous ways have been used to promote continuing care. These include having a representative of the community mental health center visit the inpatient unit and be introduced to the patient nearing discharge. The two can spend a little time together, enough for the patient to understand the importance of outpatient care, obtain clear directions to the facility, and become comfortable with the familiarity of shifting treatment to another

agency. In a reciprocal fashion, patients may be accompanied by a staff member to make a visit to the community agency where outpatient services will be provided. The same familiarization will reduce anxiety and increase the likelihood of follow-through. Providing transportation from the patient's home to the aftercare facility almost always increases attendance and participation.

Being that this is an era of person-centered, recovery-oriented services, there are training programs, such as the Community Re-Entry Module, that empower patients to learn skills and attitudes that support continuity of care. Some of these skills taught by this training program include: connecting with the community, planning a daily schedule, making and keeping appointments and managing medications. These skills will allow the patient to take responsibility for making the connection with outpatient services.

In this continuity of care program, family members are involved in the training as much as their schedules permit. A randomized, controlled study comparing the continuity in reaching outpatient services for inpatients who participated in the Community Re-Entry Module versus those who received standard discharge planning by a social worker showed that more than 80 percent of the individuals participating in the module succeeded getting to an outpatient facility after leaving the hospital, compared to only 37 percent of those receiving standard discharge. This study confirms the capabilities of patients with schizophrenia to take responsibility in maintaining their treatment.

During the acute phase of treatment, some common patient responsibilities include such functional and symptomatic outcomes as engaging in independent self-care and personal hygiene without prompting, socializing with peers without hostility or social withdrawal, voluntarily accepting medication, and learning the reasons why medications are being prescribed—both for dampening the current symptoms that led to hospitalization and for prophylactic purposes to reduce risk of relapse. The patient and doctor review specific behavioral progress being made toward the goal on a daily basis, which is the cornerstone of motivation, hope, and shared decision making. Most inpatient units have "level systems" that are used to provide incentives to patients to take more responsibility for their self-care, reliability, participation, and inappropriate behavior. Level systems can be customized to meet the individual needs of specific patients by using tokens, credits, or points superimposed on the level system. Another way to motivate patients to reach their goals and discharge readiness during the acute phase of illness is to use therapeutic contracts. Therapeutic contracts invite collaboration in that both parties describe their goals and responsibilities for the inpatient period of treatment.

The psychiatrist and other staff must ensure that the patient is familiar

with and has a plan for reaching follow-up care as an outpatient. As discharge approaches, the responsible psychiatrist should reach out to the family and invite them to attend a psychoeducation meeting with the psychiatrist present to answer questions and give feedback on how their relative is progressing. Having the family present with the psychiatrist and treatment team during a psychoeducational meeting in the acute phase of illness provides opportunities to acquire another source of information about the patient's disorder and disability. Family members can at times be more reliable assessors, and many of them keep written records or diaries of the course of illness, medications, and responsiveness and side effects to the illness. Home visits seldom occur in the current era; however, this can provide considerable information about the quality of life of the patient prior to hospitalization, thereby enabling more accuracy and clinical acumen to be brought to bear on the diagnosis, treatment experiences, and social functioning of the patient.

STABILIZATION PHASE

As previously mentioned, stabilization of symptoms and cognitive deficits almost always continues after acute discharge into the stabilization phase. Most of the considerations noted previously in the acute phase are continued in the outpatient sector; hence, the treating psychiatrist has the responsibility for ensuring that continuity of care is actualized.

STABILIZATION OF SYMPTOMS WITH MEDICATION

Once discharged, it is likely that the patient will stay on medications. Monitoring symptoms using quantitative scales such as the Brief Psychiatric Rating Scale (BPRS), the Positive and Negative Syndrome Scale (PANSS), or the Target Symptom Scales during regular appointments can help the prescribing practitioner make informed decisions regarding changes in type or dose of medication.

The use of a quantitative rating scale for the informed use of medications is illustrated by the stabilization of the patient in the following case, a young woman with schizophrenia. Responsiveness to medication is also affected by the emotional climate in the family, which needs to be considered by the prescribing psychiatrist.

Ms. R, a 28-year-old female had two previous hospitalizations for relapse of persecutory delusions, each of which responded fully to antipsychotic medication. After being discharged from the hospital following her second hospitalization, her delusions were substantially improved; however, she reported feeling "numb and disconnected" from others. During the first month of outpatient visits, her psychiatrist gradually titrated up her medication, with evidence of stability in that her delusional symptoms did not reoccur. However, regular ratings using the PANSS showed that her negative symptoms were worsening. After a family conference, it became clear to the psychiatrist that Ms. R's parents did not understand the impact that negative symptoms could have on Ms. R's daily functioning despite the lack of active psychosis. As a result, they were pressuring her to return to work. The stress associated

with their lack of education about her illness was reverberating within the family, adversely affecting both Ms. R and the parents. The psychiatrist scheduled a series of weekly and then biweekly educational sessions for Ms. R and her parents that gradually lowered the "emotional temperature" of the family household, resulting in Ms. R receiving the assistance that she needed in order to actively seek out vocational rehabilitation.

ILLNESS MANAGEMENT

During the stabilization phase, learning illness management skills including gaining a more solid understanding of the role of medication can be very useful. This includes skills such as how to communicate assertively with the psychiatrist about side effects or lack of symptom reduction. Supportive treatment sessions are provided on a weekly or biweekly basis with invitations to the family or other caregivers to participate in the training sessions for illness management that focus on medication and relapse prevention. UCLA offers a Social and Independent Living Skills (ILS) training with modules that offer training on key skills including illness management. The modules are composed of a prescriptive clinician's manual or guide, a videocassette for demonstrating the skills to be learned, and a patient's workbook, containing practice exercises and monitoring forms. Skills are learned by using a combination of focused instructions, videotaped demonstrations, role-played rehearsals, social and videotape feedback, problem solving, and practice in the natural environment through in vivo exercises and homework assignments. Patients proceed through each skill area in a specific sequence of learning exercises and activities. The training procedure emphasizes the "discovery learning" method as opposed to expository or didactic teaching.

COLLABORATIVE RELATIONSHIP WITH PATIENT

The development of a collaborative and therapeutic relationship, with a foundation based on the elements described in the earlier discussion of engagement, is essential if the patient is to remain in long-term treatment and make progress toward his or her personal goals. Personally relevant goals are motivating for the patient's participation in the stabilization, stable, and recovery phases of treatment. Depending on the needs and desires of the patient, these goals can be connected to such areas of functioning as self-management, friendship and intimacy, work or school, family relations, residential status, recreation, financial, and spiritual needs. The psychiatrist and other mental health professionals working with the patient must help the patient set incremental, stepwise goals that will bring the patient closer to his or her overall personal goals. Thus, as symptoms and cognitive impairments stabilize, the patient may benefit from weekly social skills training that gives the patient opportunities to learn from social modeling, role-play practice, and homework assignments to facilitate achieving short-term goals that are on the pathway to overall goals. For example, a man in his 30s who has reached the recovery phase of

schizophrenia may have a personally relevant goal to find an apartment and live independently. With the guidance of the skills trainer, he may proceed through a series of steps, including the following:

- Learning and demonstrating the ability to engage in medication self-management and other ILS required for living in an unsupervised residence.
- Through problem solving training, identifying various ways to locate apartments that are for rent, generating options.
- Evaluating the pros and cons of each identified apartment—for example, costs, location, and availability of public transportation.
- Practicing in the group how to ask questions assertively about the apartment with the property manager or owner, then applying the skills practiced to the real-life situations.
- Reviewing the advantages, disadvantages, and feasibility of each apartment and making a provisional choice with second and third preferences.
- Practicing (in the group setting) how to request a loan from parents to help subsidize the rent, then carrying out the request with parents.
- Practicing (in the group setting) finalizing and signing a lease agreement with the property manager, then doing it in real life.

Once the individual settles into an apartment, there are many more related goals, including furnishing the apartment, meeting neighbors, identifying and engaging in social events available in the community. Once the individual is successfully established in the apartment, he or she may again set new personal goals such as making friends and finding a part-time job.

Each success that patients experience in moving deliberately toward their overall goals brings greater self-confidence, hope, and positive expectations. These successes also reinforce the therapeutic relationship with the psychiatrist and other mental health professionals who are assisting the patient. As therapeutic relationships strengthen, so does the willingness of patients to confide in their treaters.

INVOLVING FAMILIES IN STABILIZATION

Because of the important role families play in facilitating or impeding improvement in patients with major mental disorders—including schizophrenia, mood disorders, disabling anxiety or somatoform disorders—it is vital, for rehabilitation efforts, to include them in planning, monitoring, and promoting stabilization. In fact, as noted in the description of the engagement and acute phases of illness, mobilizing families as partners in the treatment enterprise is relevant for all phases, including stable and recovery periods.

Relatives and other caregivers can be involved in sessions that are organized either for individual families or multiple family groups, with goals limited to psychoeducation or, more extensively, teaching skills in communication and problem solving, coping with persistent symptoms, and implementing relapse prevention plans.

In the stabilization phase, patients should be involved in family sessions as their symptoms and cognitive impairments permit. It is essential that family services be fully integrated with other treatment modalities being offered to the patient, such as social and vocational rehabilitation and medication management. Multidisciplinary integration is accomplished most effectively through regular consultations and interactions among all those responsible for the patient's services. Communication among mental health providers can be initiated and led by the psychiatrist through phone calls, case conferences, and team meetings and email. Families can also benefit from participation in peer-related services offered by local National Alliance on Mental Illness affiliates that include Family-to-Family seminars, Share and Care Meetings, educational lectures, and annual statewide and national conferences.

Contacts with families should continue indefinitely because they often serve as the most important continuous thread of support that weaves together the various phases of illness. It is now well established that patients living with relatives who learn coping skills from behavioral family management have two to three times fewer relapses than patients who are treated with medication and case management alone.

STABLE PHASE

Patients are considered stable when their symptoms no longer interfere with everyday activities. This does not mean that their symptoms are in remission, although they are usually rated as "moderate," "mild," or "very mild" on the BPRS. Some patients are able to participate in work, school, social, and recreational activities in normative settings even when delusions and hallucinations are present on a daily basis. Psychological interventions can displace symptoms, as when patients are actively engaged in conversations or tasks that shift their attentional capacity from isolated self-preoccupation by using eye contact, questions and answers related to current and future plans, and CBT. Cognitive functioning can also be improved by cognitive remediation (CR) in this phase of treatment.

Incremental steps toward long-term, person-centered goals continue in the stable phase. Progress toward goals can be promoted by participation in training programs that use behavioral learning principles to assure generalization of skills learned in skills-training sessions that take place in clinical settings. The key to generalization is for clinicians, family members, or peers to promote opportunities and encourage, as well as reinforce the use of skills in the natural environment.

VOCATIONAL REHABILITATION

During the stable and recovery phases, patients may be able to begin job seeking or return to school. Although supported employment (or supported education) is the most popular evidence-based service for vocational rehabilitation, it is not the only occupational goal that is clinically relevant. A wide variety of vocational and educational activities can be compatible with the stable and recovery periods of functioning. Before some patients are ready for work in competitive jobs, they inevitably need to transition through work-related activities short of regular jobs in the community. At the same time, it is important to recognize the value of supported employment because patients with mental illness who gain jobs in the competitive sector can work half-time or more in competitive lines of work, and can gain satisfaction with their renewed or achieved status as workers.

However, not all individuals with mental disabilities wish to participate in supported employment. Approximately half of those with disabling forms of schizophrenia, mood disorders, and personality disorders do not engage in supported employment. Moreover, supported employment is not always associated with increased self-esteem, quality of life, global functioning, or reduced psychiatric symptoms. There are alternative ways to lead a working life, including an array of opportunities that can be viewed as prevocational in nature such as volunteer opportunities and on-the-job training.

Various work circumstances can lay the groundwork for recovery. The value of including the full spectrum of work activities as relevant to rehabilitation is illustrated by the findings from an evaluation of 49 consumers who participated in a range of prevocational endeavors. A study found that patients who participated in a wide range of work-like settings improved their social functioning scores on the Colorado Client Assessment Record. Ratings can range from 1 (*normal functioning*) to 9 (*clearly disabled*). When seeing themselves as self-directed agents contributing purposefully to their own, patient-centered work goals, patients can make progress toward independence and self-sufficiency in a variety of ways as they are able to see themselves working toward recovery.

RECOVERY PHASE

As noted earlier, recovery does not mean “cure”; rather, patients in this phase are empowered to live meaningful and satisfying lives, having reached many of their overall goals that bring them into normal community settings and interactions. In this phase, patients are more autonomous and self-directed in their choice of person-centered goals and taking steps and garnering resources to achieve their goals. They are equal partners with their professional providers of services, making decisions and serving increasingly as their own case managers. They often assume paying jobs as

peer-counselors and peer-advocates in addition to joining with mental health professionals as active proponents of public education and political action for decreasing stigma and improving funding for mental health services in their states and communities.

Recovered patients are capable of managing routines, work, school, independent living, and illness self-management without direct supervision by relatives or mental health professionals. That being said, it is strongly recommended that recovered persons sustain their connection to mental health services, even if this consists of only occasional "check-ins" with their treatment providers given the increased vulnerability to relapse and stress sensitivity often experienced by individuals that have experienced serious mental disorders. Refresher courses for patients' emergency plans to implement relapse prevention are one mechanism to extend recovery for longer periods of time.

ASSESSMENT IN PSYCHIATRIC REHABILITATION

The assessment of persons with SMI involves a comprehensive assessment of a number of important domains with specific well-validated tools. The goal of any good assessment is to render an accurate diagnosis and then to suggest areas that can be targeted with specific interventions. The key to a comprehensive assessment of SMI is a thorough assessment of symptoms (both positive and negative), including the nature, frequency, intensity and duration of particular symptoms. It should be noted that the management of certain symptom clusters (e.g., reduction of hallucinations and/or delusions) but not others (e.g., apathy, social cognition, and interpersonal skills) may not be specifically addressed by specific pharmacological and behaviorally based treatments. In addition, reduction of symptoms is not analogous to a "cure" for most individuals since neurocognitive and behavioral deficits in conditions such as schizophrenia may persist after symptomatic relief of positive symptoms. A number of symptom rating scales include the PANSS, the BPRS, and the Scale for the Assessment of Negative Symptoms (SANS). These instruments can provide useful and important information on symptom dimensions and provide a common basis by which to come to inform clinical and research diagnosis as well as to assess outcome.

As with many assessment instruments, interpretation of these measures depends on the operationalization and validity of specific symptom constructs as well as their ability to designate which cluster of symptoms belongs to a particular psychiatric condition. It is important to take into account the fact that many psychiatric syndromes have overlapping features, and differences in symptom presentation can be a function of the stage of illness, time of onset of a psychiatric episode and the context and environment in which the patient is assessed. Further, many persons with SMI have comorbid psychiatric and medical disorders such as co-

occurrence of alcohol and substance abuse disorders. Further specific endocrine, metabolic and neurological disorders may produce symptoms that mimic psychiatric symptomatology.

One of the big challenges in the field is that controlled clinical trials often use relatively stringent inclusion and exclusion criteria that do not account for all of the comorbid psychiatric conditions, medical conditions, polypharmacy, and important mediating factors such as compliance, supported employment services (SES), environmental supports, and cross-cultural issues. As a result, it is increasingly important to ensure that findings obtained from placebo-controlled, randomized clinical trials generalize to persons in different medical, hospital and community mental health settings. While controlled randomized clinical trials may possess internal validity, the true test of an intervention is effectiveness in real-world community settings.

NEUROCOGNITIVE SYMPTOMS

Cognitive assessment is important in SMI, particularly in schizophrenia, given the number of impairments associated with the disorder. Neurocognitive deficits are thought to be a core feature of schizophrenia and often include emotional processing, perceptual disturbances, attention, sensorimotor integration, executive function, context processing, social perception, processing speed, and working memory. While pervasive cognitive deficits can vary in degree of intensity among individuals with schizophrenia, a substantial number of persons with bipolar disorder (up to 40 percent) may have completely normal cognitive profiles. A tremendously positive development in schizophrenia research has been the development of the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB) for the assessment of schizophrenia. The MCCB has not only been successful in research on schizophrenia but has also been able to differentiate Bipolar 1 and Bipolar 2 patients from controls. It has been recommended that because of the more subtle nature of cognitive deficits in schizophrenia, additional tests that should be considered (besides the MCCB) include the California Verbal Learning Test-II (CVLT-II). In addition, the Wisconsin Card-Sorting Task, Trails B and the Stroop Color-Word task are all measures that should be considered.

An interview-based assessment of cognition specifically for schizophrenia is the Schizophrenia Cognition Rating Scale (SCoRS). The SCoRS assesses cognitive functioning based on interviews with both patients and their caregivers. The assessor then incorporates both sources of information to determine a final rating. Performance on the SCoRS scores has been found to be associated with both actual performance on cognitive tests and measures of real-world functioning.

The problem with any neuropsychological assessment is that there are

many factors in SMI other than underlying putative brain damage that may contribute to poor performance. For example, poor motivation, sedation and other cognitive effects of medications, social anxiety, paranoia or hypervigilance and interactions with the test examiner may have significant effects on test outcome. In some settings, attention and motivation to perform may be so low that the neuropsychological battery with the Tests of Malingering and Memory (TOMM) or Green's Word Memory Test (WMT) is augmented. It is important that any neuropsychological assessment that is performed is not in the acute stages of SMI. The authors find it much more helpful to conduct neurocognitive evaluation after ample stabilization has occurred. There is a tendency to think that neuropsychological testing taps orthogonal, independent cognitive domains but deficits in memory and measures of verbal fluency may all be related to generalized cognitive slowing, attentional deficits and reduced speed of processing. While there have been attempts such as the National Institute of Mental Health (NIMH) initiative (Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia [CNTRICS]) aimed to develop measures derived from cognitive neuroscience, with explicit links to known neural systems and psychopharmacology, the field is still hampered by a locationist notion of the brain where function is thought to occur in specific brain regions (e.g., amygdala) rather than distributed networks of neuronal activity that involve multiple brain regions. This highlights the importance of relating neurocognitive function to functional measures such as fMRI and PET in order to identify working neural networks in the brain.

SOCIAL COGNITION

There is increasing evidence that negative symptoms have an impact on both social competence and social outcomes, but do not exert an influence on social cognition. However, social competence and social cognition have significant effects on real-world social outcomes. The Result of the Expert Panel and Rand Survey has identified four major domains of social cognition that includes (a) emotional processing, (b) attributional style/bias, (c) social perception, and (d) Theory of Mind (ToM)/mental state attribution. Measures that the committee agreed upon for final inclusion in the battery to assess the aforementioned domains included the abbreviated version of the Ambiguous Intentions and Hostility Questionnaire (AIHQ), Bell Lysaker Emotion Recognition Task (BLERT), Penn Emotion Recognition Test (ER-40), Hinting Test, abbreviated version of Relationship Across Domains (RADs), Reading the Mind in the Eyes Test, The Awareness of Social Inferences Test (TASIT), and the abbreviated version of the Trustworthiness Test.

Recently, as part of the Social Cognition Psychometric Evaluation (SCOPE) study, it was shown that the BLERT, a test of affect recognition

and Hinting Task (ToM) had the highest psychometric properties among a number of social cognition tasks under study and were recommended for use in clinical trials. Those with intermediate psychometric properties included the ER-40, TASIT and Eyes Test which were recommended for further study. In contrast, the AIHQ, RADs test, and Trustworthiness Task had considerably poorer psychometric properties and caution was urged in the use of these measures.

COMMUNICATION AND SOCIAL SKILLS

Deficits in metacognition, the ability to construct complex ideas about oneself and others and the ability to respond to social and psychological challenges underlie many of the communication and social skills deficits found in schizophrenia. This mismatch between perceptions of oneself and the world and communication with others in the environment can cause considerable difficulties for an individual with schizophrenia. Impaired communication and language skills have been identified as common in schizophrenia and it has been shown that persons with schizophrenia have greater social skills deficits than those with major depression or bipolar disorder. Those with schizophrenia appear to have social skills deficits that are intermediate. Social skills training is based on behavioral principles of modeling; shaping, rehearsal, and positive reinforcement are all important parts of successful training paradigms. It should be noted that social skills training has advocated for helping an individual interact with his/her world. Both social cognition and negative symptoms carry the most variance in explaining real-world social outcomes. This suggests that targeting both social cognition and negative symptoms may be as important as teaching social skills to individuals. The problem with both social skills and expressive and receptive communication deficits in schizophrenia is that there is a complex interaction between social perceptions, integration of thinking about self and others and behavioral outputs that can all serve to produce negative outcomes that effect social, vocational and community life. It is likely that the most effective approaches are going to have to focus on the complex interplay between individual perceptual and social cognition deficits, how these interact with more complex metacognitive processes and help modify behavioral repertoires that will foster optimal real-world functioning.

ASSESSING FUNCTIONAL OUTCOMES

Broadly speaking, the goal of any psychiatric rehabilitation or remediation effort is to enhance and support optimal functioning in the environment. However, there are different types of function, including interpersonal communication, ability to manage one's household affairs, managing one's finances, community affairs and working in a vocational setting. These broad areas of functioning have considerable overlap. For example,

managing one's household affairs may involve arranging for transportation to the bank or supermarket and understanding bus, train or subway routes. There are important decisions to be made on how to prioritize income, making sure that income is used for food, housing and medication before it is allotted for other purposes. Persons pay bills, must interact and effectively communicate with others and meet daily responsibilities in the community and the workplace. Every individual is different and understanding the individual's psychiatric condition, treatment regimen, adherence to treatment, natural social supports, contextual cues in the environment, and an understanding of a patient's strengths as well as his or her limitations is critical to helping the individual achieve optimal functioning within the community. This requires assessment techniques that allow for appropriate targeted interventions and remediation strategies. Proper assessment allows for optimal treatment and psychiatric rehabilitation. The reader will notice that these are not two distinct entities but rather a continuous effort to promote wellness and to maintain optimal function in many environmental contexts.

A critical distinction is the distinction between *functional capacity* and *functional performance*. Functional capacity refers to the ability and the underlying skills required to successfully perform activities that support independent living in the community. *Functional Capacity* can encompass a wide range of activities such as (a) the ability to count money or balance a checkbook, (b) having sufficient social cognition to understand implicit rules regarding social interactions or behaviors, (c) adequate social skills to successfully navigate a work or community environment, (d) requisite cognitive abilities such as intellect, memory, reasoning, language abilities to engage in meaningful home, community or occupational activities. If a person does not possess adequate functional capacity (requisite skills required for independent or even assisted living), then the focus of psychiatric rehabilitation has to be on attempting to build and foster those skills that will result in the greatest independence and quality of life. *Functional Performance,* while often dependent on underlying functional capacity, refers to the actual ability to consistently perform a variety of skills in a way that maximizes autonomy, quality of life and the ability to cope with real-world demands. Functional performance then represents the actual performance of the individual in interpersonal, social, financial community, and vocational domains.

Individuals with a psychiatric condition such as bipolar disorder may have excellent underlying financial capacity, intellectual skills, social cognition and awareness and be able to live independently and to be gainfully employed. However, because of waxing and waning of symptoms, there may be certain periods where they may fail to adequately manage their manic, depressive or mixed episodes and may mismanage finances, have difficulties meeting the demands of home, work or family and be

incapacitated despite their premorbid high levels of functioning. Individuals with schizophrenia may have intact financial and adequate cognitive skills to manage to work and live independently in some settings. However, persistent difficulties with negative symptoms and social cognition can lead to problems that may negatively affect their ability to support themselves and their families. Primary and comorbid substance abuse can radically affect a patient to make poor choices and to allocate financial resources to their addiction or dependency, thus threatening their living situation. Psychiatrists are confronted with the sad reality that many homeless and socially impoverished individuals have underlying psychiatric conditions and lack proper supports that could substantially improve their well-being and quality of life.

The distinction between functional capacity and general real-world functional performance is critical in that these constructs dictate the approaches that one will take toward assessment and subsequent management. Skill-building and other approaches to help increase functional capacity may be critical (i.e., individual or small groups social skills training). However, the effectiveness of these approaches can only be seen when it forms the building blocks to enhance functional performance (i.e., generalizing social skills training to real-world community and/or occupational settings). Thus, assessment of functional capacity and real-world performance are critical elements in psychiatric rehabilitation. Functional capacity and performance can further be subcategorized with regard to particular domains such as cognitive, financial skills and judgment, social cognition, social skills, and judgment that affect individual autonomy and quality of life. Further, specific types of functional skills required for everyday living may pertain to specific domains such as caring for one's self in the home, successfully interacting with the community, or learning and maintaining job specific skills. This suggests that functional assessment strategies need to be tailor-made for each individual and display reliability (consistency of measurement) and validity (assessing the construct that needs to be measured).

TYPES OF FUNCTIONAL ASSESSMENT

Functional assessment can be employed in a number of ways. The most direct method of assessing functional capacity is direct observation of the patient performing behaviors in the clinical setting. One widely used measure such as the Performance-Based Skills Assessment (UPSA) developed by investigators at the University of California directly assesses the individual's functional capacity with regard to finances, communication skills, household chores, transportation, and planning recreational and other activities. The advantage of such a measure is that it is easy to administer, has demonstrated reliability and validity and can provide direct and quantifiable information on functional capacity. Knowledge of deficits

in functional capacity can be critical for initial assessment as well as response to treatment. However, performance-based assessment may not always map on the individual's performance in the community. In this case, a qualified informant can provide essential information based on repeated observations of the individual in the context of different environmental supports and contextual cues. In this context, a reliable collateral informant, whether it be a friend, relative, or knowledgeable professional can provide essential information that may simply not be available by performance-based functional assessment. However, quality of the informant matters. If a collateral informant has little opportunity to observe the patient, has a cognitive psychiatric disability themselves or self-serving biases, their report may be of little value in providing useful functional information. While it is always important to obtain and respect the individual's evaluation of their real-world functional performance, it is important to note that in schizophrenia and other neuropsychiatric conditions such as Alzheimer disease (AD) and other neurodegenerative disorders, reports are often positively biased (i.e., the patient often reports a much higher level of functional performance than may actually be the case). Anosognosia or the "inability to appreciate what one does not know" is a neuropsychological condition that is seen in traumatic brain injury (particularly within the right hemisphere), AD and in psychotic disorders such as schizophrenia.

As a result, it is often very useful for the clinician to compare the report of the patient and reliable informant and to also use performance-based functional assessment tools and rating scales.

FUNCTIONAL ASSESSMENT SCALES

In addition to the UPSA, the performance-based measure of functional capacity, a study also notes important scales for functional capacity such as the ILS and Test of Adaptive Behaviors (TABS). Everyday functional rating scales selected by the RAND expert panel include the Independent Living Skills-Survey (ILSS), Specific Levels of Functioning Scale (SLOF), Life Skills Profile (LSP), Birchwood Social Functioning Scale (SFS), Social Behavior Schedule (SBS), and Heinrichs-Carpenter Quality of Life Scale (QLS).

The most widely employed measure of functional capacity, the UPSA is a particularly compelling scale of functional capacity that has shown correlations of 0.6 to 0.7 with neurocognition in a number of studies. It includes simulations of everyday activities of daily living such as: communication (i.e., dialing the telephone to schedule a physician appointment, obtaining and dialing a number from a telephone operator); comprehension/planning (reading and planning a recreational outing); finance (counting currency, making change, and paying a bill); household management skills (e.g., shopping for grocery items), and transportation

(using a bus schedule in planning a trip and understanding charges and obtaining information). Some versions of the UPSA include a medication management subtest and are referred to as the UPSA-II. There is also an abbreviated version of the UPSA that focuses on communication skills and finances which is called the UPSA-B.

The SLOF relies on self or informant reports but is especially useful when employed by a high contact clinician. The scale taps the following domains, Participation in Community and Household Activities, Interpersonal Relationships, Work Skills, Self-Care, Physical Functioning, and Socially Acceptable Behavior subscales.

The UPSA has the advantage of being an objective performance-based functional measure and is preferable to self-reports of patients who often overestimate their abilities or questionable informants. Critics of the UPSA maintain that the UPSA measures tap functional behaviors but that they are more psychometrically similar to cognitive measures. There are contextual cues in the real-world environment that dictate functional real-world functional performance apart from demonstrating capacity in contrived clinical and laboratory settings. Measures such as the SLOF are most effective when used with informants but may in some cases be susceptible to reporter biases. The SLOF and other functional rating scales are typically most effective in the hands of a high contact clinician who knows the patient well and can weigh informant and patient reports against his/her own clinical judgment. While the SLOF has been extensively validated with schizophrenia patients, studies have shown that 64 percent of bipolar patients evidence deficits in one or more SLOF domains. The UPSA-1, UPSA-II, or UPSA B has been employed with a large number of patient groups including those with major depression, bipolar disorder, and schizophrenia. Importantly, the instrument has also been employed in diverse cultural and language groups.

Other scales important in psychiatric rehabilitation include the Maryland Assessment of Social Competence (MASC), a structured role-play designed to assess social skills. Other scales use time-sampling techniques consisting of numerous observations of patient social behaviors, such as the Time Sample Behavioral Checklist (TSBC) or Clinical Frequencies Recording System (CFRS), or Staff-Resident Interaction Chronograph (SRIC) but are usually only appropriate for institutional and residential treatment settings since these involve extensive ratings by staff members.

COGNITIVE REMEDIATION IN COMMUNITY AND OCCUPATIONAL FUNCTION

Neurocognitive deficits are core features of schizophrenia and moderate to severe cognitive impairments are present in the vast majority of cases. The most commonly used measures are memory scales such as the Rey Auditory Verbal Learning Test (RAVLT) or CVLT-II. Executive function is often measured by the Wisconsin Card Sorting Test (WCST) or Trails B

(although this measure taps many aspects of neurocognitive function) or measures of the DKEFS. Digit symbol of the WAIS-IV is a good measure of processing speed as is Trails A. Examples of widely used neurocognitive batteries include the MATRICS MCCB, the Brief Assessment of Cognition in Schizophrenia (BACS), and Repeatable Battery for the Neuropsychological Assessment of Schizophrenia (RBANS).

Meta-analyses have shown that CR in schizophrenia has robust effects of cognitive outcomes but there is little evidence that there is an effect on functional skills and abilities. However, when combining CR with functional skills training, better results in community living skills occurred than with functional training alone. An important review indicated that poor neurocognitive function had significant effects on vocational and occupational status within the community setting. Subsequent investigations revealed that CR plus SES enhanced positive employment outcomes whereas other studies have not shown such effects. An important issue is whether training specific cognitive abilities such as memory, attention, and executive skills are as effective as directly training specific functional and occupational skills. The notion is that training component cognitive abilities in the hopes that they will generalize to specific real-world behaviors may be less important than directly training target behaviors of interest.

As an example, consider an adult with schizoaffective disorder who is working in a supported occupational setting. One approach might be to train specific cognitive component abilities in the hopes that they will generalize to work setting. An alternative approach that the authors favor would be training those skills required for job-related performance. For example, CR strategies would specifically target actual job-related activities such as organization, strengthening procedural memory for certain rote activities, and employing strategies to monitor performance. This targeted functional training would involve the cognitive aspects of the skills related to employment but directly integrate this training in an integrated manner and in a manner that would more likely generalize to the employment setting.

REHABILITATION OF SOCIAL COGNITION

There has been an increased interest in moving beyond merely framing core features of schizophrenia in terms of positive and negative symptoms and neurocognitive impairment. There has been emerging evidence that deficits in social cognition can also be considered core and stable features of schizophrenia. Social cognition pertains to the perceptions, beliefs, and interpretation of other people's behaviors. Poor social cognition can significantly impair a person with SMI to appropriately respond to others in their environment. Green and others concluded that there were five major aspects of social cognition. *Theory of Mind* generally refers to an

individual's ability to infer the beliefs and intentions of others. *Social Knowledge* consists of the understanding of rules and conventions that underlie social behavior and guide social interactions. *Social Perception* is the ability to recognize general rules concerning roles and societal rules. *Emotional Processing* involves the experiencing and use of emotions and has been linked with emotional intelligence. Finally, *Attributional Bias* involves attributing causality to both positive and negative outcomes. Each of these components is critically important in understanding how an individual relates to others in the community and their vocational performance. It should be noted that deficits in social cognition are not only observed in those with schizophrenia but also individuals with bipolar disorder and to a lesser extent, those with major depressive disorder.

There is increasing evidence that social cognition accounts for far more variability in community function than neurocognitive abilities.

A recent meta-analysis of studies on social cognition revealed that social cognitive training procedures have relatively large effects on emotion recognition, small to medium effects on measures tapping Theory of Mind effect, and no effects on attributional bias or social perception. One of the most important findings in this study is that these kinds of techniques have a moderate to large effect on community functioning. However, there may be less of an effect on higher-order social cognitive skills. This has led to the development of more real-world–computerized tests using virtual reality such as the RC2S which involve navigating a computerized character through situations that involve emotional processing and learning contextual cues.

COGNITIVE BEHAVIORAL THERAPY

CBT is a type of psychotherapy based on principles of conditioning, learning, and cognition. As applied to schizophrenia, CBT engages patients collaboratively, challenging their interpretations of events or experiences. It is thus often used to assist patients in developing more realistic alternatives to delusional and paranoid thinking. It can also be used to help patients to develop less stressful attributions for hallucinations. For example, rather than accepting that the voices that he or she hears are omnipotent, omniscient, and must be obeyed, behavioral experiments can be set up to allow the patient to test this hypothesis and replace these thoughts with more realistic ones. In many cases, the reduction in anxiety that accompanies modified attributions regarding voices leads to a reduced frequency and intensity of hallucinations. In the case of negative symptoms, CBT is often used to challenge anhedonic beliefs, schedule activities, and administer self-rewards for increasingly engaging in pleasurable activities. CBT works best with patients who are willing to actively participate in treatment, often as the result of distress from symptoms and/or at least minimal insight into the abnormal nature of their

experiences.

A meta-analysis of CBT for schizophrenia conducted in 2001, including seven randomized, controlled trials, showed that CBT produced large clinical effects on both positive and negative symptoms. CBT as an adjunct to routine clinical care also resulted in significant additional benefits compared with routine clinical care plus supportive therapy. An updated review of 17 clinical trials testing the efficacy of CBT for schizophrenia conducted in 2004 showed that CBT results in better outcomes on specific measures compared with routine care or supportive therapies in a variety of settings (e.g., day-treatment programs and long-stay inpatient programs) and across some special patient populations (e.g., relapse-vulnerable patients, older adults with schizophrenia, and individuals with comorbid substance abuse or anxiety disorders). Researchers conducted a meta-analysis of 14 studies including 1,484 patients, and found that compared to other adjunctive measures, CBT showed significant reduction in positive symptoms. The specific symptoms alleviated with adjunctive CBT include delusions, distress, or delusions related to auditory hallucinations, paranoia, and negative symptoms. Improvements in distress related to auditory hallucinations and delusions were maintained at follow-up.

In addition to problems uniquely associated with psychotic disorders, various forms of CBT are effective for addressing generalized anxiety, panic, social anxiety, depression, obsessive-compulsive symptoms, and substance abuse. These problems often co-occur with chronic psychotic disorders, and there is no reason to believe that CBT interventions are any less effective in treating them in people with schizophrenia than in the general population.

An issue when considering CBT for psychosis is that, to succeed, patients must have at least some motivation to reduce or eliminate their symptoms. Patients who are thus motivated typically demonstrate good treatment responses, whereas unmotivated patients often do not respond or drop out of treatment. Another challenge is to find ways to engage first-episode patients in CBT. Often, more supportive efforts are needed to engage first-episode patients in the treatment system prior to initiating CBT.

SOCIAL LEARNING AND SOCIAL SKILLS TRAINING

The social learning program (SLP) is a comprehensive, integrated set of learning-based techniques (typically including a token economy) and skills-training activities. In most cases, an SLP is delivered within the context of a residential facility, usually for patients with severe and chronic behavioral disabilities. A key feature of the SLP is that all staff members on all shifts are trained to structure their interactions with patients in specific ways during all waking hours. Research indicates that SLP is the most effective

form of inpatient treatment for people with schizophrenia, including the so-called "treatment-refractory" patients. The SLP has demonstrated effectiveness with a wide range of people with disabilities in a variety of settings (e.g., state and private hospitals and maximum security forensic hospitals). The components of the SLP form an extremely positive teaching environment that assists clients in developing adaptive behaviors and skills relevant to successful return to the community. Contemporary versions of this approach incorporate skills-training technologies, psychiatric rehabilitation techniques for engaging clients, CBT, vocational rehabilitation, cognitive rehabilitation, and substance abuse treatment.

A guiding assumption of the SLP is that all persons with severe mental disorders, no matter how severe the impairments, can learn new behaviors and skills and therefore progress to less structured environments. To promote behavior change and skill acquisition, the SLP incorporates a wide variety of techniques based on fundamental learning principles. Differential reinforcement forms the basis of many of the procedures of the SLP. Staff at all levels and disciplines are highly trained and carefully supervised in the timely and consistent delivery of verbal praise, positive social attention, and material reinforcement for adaptive behaviors and the application of extinction or response-cost techniques for bizarre, aggressive, or rule-violating behaviors. Effective SLPs deliver much more positive reinforcement than extinction or response cost. The result is a highly reinforcing, positive, and supportive environment.

Although most comprehensive SLPs include some sort of token economy or point system to deliver feedback and behavior-related consequences, some use individual behavioral contracts, contingency management plans, or social reinforcement alone. However, in all well-designed SLPs, specific objectives and goals of clients are individualized, as are the specific interventions for assisting them with acquiring the skills necessary to accomplish these goals/objectives. Another key feature of all effective SLPs is that, as patients begin to demonstrate improvements in behaviors, there is a planned and systematic withdrawal of tokens or points and an emphasis on generalization training so that clients will be able to apply acquired skills and knowledge successfully in "real-life" community settings under more natural, primarily social reinforcement conditions.

An effective SLP is based on a wealth of data that are generated during all waking hours by all staff, using a number of observational rating procedures (described in the section on assessment). The availability of data on all critical behaviors (e.g., grooming, attendance at groups and activities, social contacts, and aggressive behavior) allows for sophisticated data-driven treatment planning. Moreover, the ability to demonstrate specifically to patients why their privileges or other consequences of their behavior are occurring and what needs to change and by how much to receive different consequences is a powerful tool. In effect, this type of

data-driven system serves to instill responsibility for behavior change in the patient, instills a sense of agency and mastery regarding behavior change, and reduces the institutionalized and dependent attitude that staff arbitrarily determine privileges, which is so often found in state hospital settings. Moreover, as demonstrated in a recent study, patient satisfaction with treatment is often high in SLPs.

Given the complexity, specificity, and range of procedures involved in the SLP, staff training is of paramount importance. All staff members are involved in extensive initial and ongoing training, using both didactic instruction and in vivo practice. Competency-based supervision is provided through regular performance feedback, using observational data on staff—patient interactions and structured supervisory observations.

Another important feature of SLPs is the focus on the generalization and maintenance of gains after program discharge. As patients acquire skills and develop goals for community living, they are provided with opportunities to apply these skills and pursue their goals in community settings. After discharge, declining contact aftercare is provided. Some SLPs arrange to train staff at community residential facilities to maximize the maintenance of gains after discharge.

SLPs have traditionally been provided on long-term treatment units. Recently, however, modifications of the full program have been tried in other settings, such as acute inpatient units and community-based residential settings for outpatients with SMI, with demonstrated success.

GENERALIZATION TRAINING

Recently, efforts have been made to improve the generalization of skills training gains in the community. One method, called in vivo amplified skills training (IVAST), uses specially trained case managers who routinely and frequently conduct additional training sessions in participants' environments. These sessions help patients to perform and practice their new skills in their community residences. An initial study of IVAST found that patients receiving the intervention achieved higher levels of interpersonal problem-solving skills, significantly greater social adjustment, and better quality of life over a 2-year period than participants with skills training alone.

A second procedure for promoting generalization involves engaging the people already in the patient's life, including residential care staff, peers, and relatives. People from these groups are selected by patients based on the criteria of cooperativeness, accessibility, and familiarity with the patient's environment. The intervention consists in structured meetings between the patient and the chosen supporters to review the use of the new behaviors, explore the causes of problems in successful performance of the behaviors, and problem solve around how to be more effective. No constraints are placed on the frequency or duration of a pair's meetings.

Evaluations of this procedure have found that participants who received both the skills training and the added support improved their interpersonal and community functioning during training and continued to improve during the 18 months following training. In contrast, those without support lost some of their improvements during the follow-up.

These data have two important implications. First, including specific interventions for generalization can help to improve patients' outcomes. Second, use of these procedures should become a practice standard, given that loss of skills is likely without them. Finally, recent developments in Medicaid reimbursement strategies have begun to shift part of the responsibility for providing care to community residences. This macro-level development provides clinicians and researchers with the opportunity to develop and apply integrated approaches to closing the gaps between the skills acquisition, performance, and generalization aspects of skills training. However, this inevitably means that more attention will have to be paid to training community residence staff on how to deliver interventions correctly.

ILLNESS/WELLNESS MANAGEMENT SKILL TRAINING

Gaining the ability to manage one's psychiatric illness is central to the rehabilitation and recovery perspective. In the rehabilitation and recovery literatures, skill training in illness/wellness management has gradually been differentiated from related social and living skills approaches, reflecting a growing recognition that specialized skills are needed to self-manage psychiatric disorders, comparable to skills needed to self-manage physical conditions such as diabetes. People learn about the episodic and persistent symptoms of their illness, the relationship between symptoms and functional impairments, pharmacological means and other techniques (e.g., relaxation and stress management) for controlling symptoms, drug side effects, identification of "warning signs" of an impending relapse, and various other aspects of their disorder and its management. Behavioral skills indirectly relevant to disorder management include, for example, the assertive skills necessary for dealing with the doctor and the doctor's receptionist in getting an appointment for a medication visit.

Several developer groups have packaged illness/wellness management skill training materials for testing and dissemination. A number of original studies and reviews confirm the effectiveness of skill training focused on illness/wellness management for improving adherence to treatment in people who live with SMI. The SAMHSA has produced two packages for public dissemination that fall under this category, Illness Management and Recovery, and Medication Management Approaches in Psychiatry. Both packages include materials for all participants in the illness/wellness management process, including the person in recovery, physician, other service providers, family, and friends.

Illness/wellness management training for people with schizophrenia benefits from relapse-prevention and related techniques. Well known for its application in substance abuse, many of the techniques of relapse prevention are well suited to the episodic nature of the lived experience of people with schizophrenia, the role of stress in precipitating psychotic symptoms, and the important role of the recovering person in managing those episodes. The original application of relapse prevention is also of interest, given that people with schizophrenia often have substance abuse problems as well. There have been no controlled trials of the unique contribution of relapse-prevention techniques to disorder management in schizophrenia, although they are frequently incorporated into psychiatric rehabilitation programs. Peer involvement is especially relevant to illness/wellness management.

SUPPORTED EDUCATION

Supported education prepares people with psychiatric disabilities to achieve educational goals in an academic setting. It addresses ways to overcome obstacles related to achieving educational success, such as improving academic skills and problem solving, bolstering self-confidence, and engaging in career development. Several models of supported education exist. Mobile support is one type that has been found beneficial. In this approach, a counselor meets one-on-one with a student to tailor an individualized plan to maximize academic success. This may include assistance in accessing services already in place in the educational setting, help with locating and using school social and recreational services, preparatory coursework, case management support, ongoing support and check-in meetings at the local student center, study support and visits to the campus library, linkage support to other campus services, and individual counseling (didactic, supportive, and/or exploratory). Other models of supported education include didactic group meetings in a classroom setting focusing on study skills, supportive group meetings, and the provision of standard study guides. Evidence from these programs suggests that distress levels are higher in people with schizophrenia than in other students. This should lead to further efforts to introduce stress management techniques into supported education activities for patients with schizophrenia.

SEX EDUCATION

Many individuals with serious mental illness aspire to and achieve meaningful education, work, and enjoy mutually supportive family and intimate relationships. The topic of intimacy and sexuality has been given the least attention in this population by mental health professionals. Surveys have discovered that 60 to 70 percent of adults with serious mental illness report being sexually active in the last year, with approximately 30

percent reporting two or more partners.

Clearly, sex education should be an integral part of a comprehensive rehabilitation program. The current impetus for increasing patients' sexual knowledge derives in part from the alarmingly high rates of HIV infection and AIDS found among individuals with serious mental illness. Not surprisingly, the majority of sex education programs have been designed to increase patients' awareness of the risk of HIV and AIDS, dispel misinformation and stereotypes, and provide the knowledge and skills needed to decrease their chances of becoming infected. Emphasizing teaching techniques that use procedural learning and memory can mitigate any cognitive disabilities that may impact learning. In psychiatric rehabilitation, implicit or procedural learning is enhanced by video and live demonstrations, termed social modeling. Other techniques that build on procedural learning and memory are role playing, interacting with a computer as in computer games, using pictures and attractive graphics to amplify written information, and engaging in repeated practice.

These educational methods can also be used for sex education programs that have broader goals than simply teaching persons with serious and persistent mental illness how to reduce their chances of becoming infected by HIV. Some of the educational objectives that comprise more ambitious and enlightened efforts in this realm include the following:

- Increasing patients' knowledge and comfort about sexuality
- Encouraging patients to identify and clarify their values and attitudes about sexuality
- Informing patients about alternative methods of birth control
- Teaching patients decision-making skills regarding sexual activity
- Overcoming medication-related sexual dysfunction and preventing deterioration or cessation of sexual activity
- Improving intimacy skills
- Imparting basic HIV and AIDS education
- Instructing patients in proper condom use.

COMORBID SUBSTANCE USE DISORDERS

The field of psychiatric rehabilitation is confronted with the need to identify and treat substance use disorders as epidemiological data show high prevalence rates among individuals with concurrent SMI. Substance use disorders have been associated with medication nonadherence, underachievement, decreased work productivity, poor health, neuropsychological impairment, social dysfunction, and violence, among other negative biopsychosocial effects. Prospective studies have shown that treatment outcomes, hospitalization rates, housing stability, and functional status, are worse among patients with dual disorders. Studies have also shown that problems related to substance use tend to persist over the long

term among patients with SMI. Lifetime treatment or help-seeking behavior is uncommon, particularly among those with psychiatric comorbidity, thus clinicians are admonished to be proactive with assessing for substance use disorders to mitigate the negative impact on successful recovery and community adjustment.

Historically, individuals with SMI such as schizophrenia and co-occurring substance use disorders received treatment for one disorder, but were excluded from the other. Eventually, distinct and parallel treatment programs emerged; however, treatment outcomes were poor using this model. As a result, integrated treatment models combine substance abuse and mental health interventions in one clinical program. It is now commonplace that facilities offer outpatient dual disorders interventions as well as integrated residential treatment programs. In integrated treatments for patients with dual disorders, mental health treatments and substance abuse treatments are brought together by the same clinician, or team of clinicians, in the same program to ensure that the patient receives a consistent explanation of illness and a coherent prescription for treatment rather than a contradictory set of messages from different providers. Integrated treatment aims to improve collaboration between treatment providers, to eliminate the patient's burden of attending two separate programs, and to remove financial and other barriers to access and retention. Comprehensive programs have frequently included not only standard mental health interventions, such as medication monitoring and support services, but also assertive outreach to engage patients in treatment; intensive case management; and individual, group, and family substance abuse counseling. Research efforts are still being made to test and refine individual components and combinations of integrated dual diagnosis treatments. As with all treatment, effective interventions targeting substance use should be individualized to address personal factors and stage of motivation.

TOWARD MORE INTEGRATIVE MODELS OF CARE

It is readily apparent that current psychiatric rehabilitation extends far beyond previous psychosocial models that were added to medication management and other somatic therapies. Psychiatric rehabilitation today is an active, integrated, and person-centered enterprise, which presupposes a strong therapeutic relationship, trust, and cooperation between the patient, family, and multidisciplinary treatment team. Excellent treatment is predicated by using state-of-the art and empirically validated assessments. While neurocognitive deficits may be pervasive in many types of schizophrenia, persons with bipolar or unipolar affective disorder may have less but still significant impairment in this area. In some cases, comorbid substance abuse disorders can further impair already pre-existing cognitive deficits. Equally important is the accurate assessment of

functional capacity (based on performance based tests) and most importantly, real-world functional performance. Functional performance cannot be limited to merely higher- and lower-order activities of daily living but must encompass also social and vocational function. There is also increasing evidence that deficits in social cognition and meta-cognition create considerable stress for individuals with schizophrenia beyond typical memory impairments, slow cognitive processing, and other neurocognitive deficits. As previously discussed, there is ample literature to suggest that improving communication skills, enhancing social skills, and behavioral goal attainment are all effective means to help the individual cope with stressors that might increase the likelihood of relapse. This chapter has focused on empirically oriented approaches that can help in these efforts. As previously mentioned, it is of essential importance for both the patient and the treatment team to (a) identify targeted short-term and long-term goals that are both operational and obtainable, (b) plan interventions and build supports that are appropriate to a person's stage of recovery, (c) build coping strategies and stress management techniques that will reduce the chances of relapse, and (d) foster a collaborative treatment milieu involving the patient, family, and care providers.

Every person with SMI is a unique individual. While evidence-based approaches can provide guidance in selection of the best instruments and interventional tools, this can never be done in isolation and without knowledge of the person's strengths, weaknesses, resources, and unique personality and environment. The best chances for optimal recovery and stabilization are through a multimodal approach that encourages adherence to medication and targeted treatment goals, stress management and coping techniques, as well as developing optimal environmental supports. SMI conditions are like many chronic conditions that need constant attention and management.

However, unlike many physical ailments, persons with SMI still have to manage additional stigma and lack of understanding afforded those which chronic medical and neurological conditions. The most effective interventions will be those with most-empirical support, targeting the many different aspects of deficits in SMI (including but not limited to social isolation, deficits in social cognition, problematic meta-cognition, functional and vocational deficits, as well as problems with communication and social skills). However, focusing solely on these targets without recognizing the importance of contextual cues and social supports risks fragmenting care.

There needs to be continued emphasis on more integrated and comprehensive models of care. Patients deserve no less.

References

Bell MD, Zito W, Greig T, Wexler BE. Neurocognitive enhancement therapy with vocational

services: work outcomes at two-year follow-up. *Schizophr Res*. 2008;105(1):18–29.

Bowie CR, McGurk SR, Mausbach B, Patterson TL, Harvey PD. Combined cognitive remediation and functional skills training for schizophrenia: effects on cognition, functional competence, and real-world behavior. *Am J Psychiatry*. 2012;169(7):710–718.

Burdick KE, Ketter TA, Goldberg JF, Calabrese JR. Assessing cognitive function in bipolar disorder: challenges and recommendations for clinical trial design. *J Clin Psychiatry*. 2015;76(3):e342–e350.

Burdick KE, Russo M, Frangou S, et al. Empirical evidence for discrete neurocognitive subgroups in bipolar disorder: clinical implications. *Psychol Med*. 2014;44:3083–3096.

Depp CA, Mausbach BT, Harvey PD, et al. Social competence and observer-rated social functioning in bipolar disorder. *Bipolar Disord*. 2010;12(8):843–850.

Green MF. What are the functional consequences of neurocognitive deficits in schizophrenia? *Am J Psychiatry*. 1996;153:321–330.

Green MF, Leitman DI. Social cognition in schizophrenia. *Schizophr Bull*. 2008;34(4):670–672.

Harvey PD, Loewenstein DA, Czaja SJ. Hospitalization and psychosis: influences of the course of cognition and everyday functioning in people with schizophrenia. *Neurobiol Dis*. 2013;53:18–25.

Javitt DC, Buchanan RW, Keefe RS, et al. Effect of the neuroprotective peptide davunetide (AL-108) on cognition and functional capacity in schizophrenia. *Schizophr Res*. 2012;136(1):25–31.

Kalin M, Kaplan S, Gould F, Pinkham AE, Penn DL, Harvey PD. Social cognition, social competence, negative symptoms and social outcomes: inter-relationships in people with schizophrenia. *J Psychiatr Res*. 2015;68:254–260.

Keefe RS, Harvey PD. Cognitive impairment in schizophrenia. In: Geyer MA, Gross G, eds. *Novel Antischizophrenia Treatments*. Springer Berlin Heidelberg; 2012:11–37.

Kurtz MM, Richardson CL. Social cognitive training for schizophrenia: a meta-analytic investigation of controlled research. *Schizophr Bull*. 2011;38(5):1092–1104.

Martino DJ, Strejilevich SA, Scápola M, et al. Heterogeneity in cognitive functioning among patients with bipolar disorder. *J Affect Disord*. 2008;109(1–2):149–156.

McGurk SR, Mueser KT, Pascaris A. Cognitive training and supported employment for persons with severe mental illness: one-year results from a randomized controlled trial. *Schizophre Bull*. 2005;31(4):898–909.

McHugh RK, Barlow DH. The dissemination and implementation of evidence-based psychological treatments: a review of current efforts. *Am Psychol*. 2010;65(2):73–84.

Mueser KT, Pratt SI, Bartels SJ, et al. Randomized trial of social rehabilitation and integrated health care for older people with severe mental illness. *J Consult Clin Psychol*. 2010;78(4):561–573.

Peyroux E, Franck N. RC2S: a cognitive remediation program to improve social cognition in schizophrenia and related disorders. *Front Hum Neurosci*. 2014;8:400.

Pinkham AE, Penn DL, Green MF, Buck B, Healey K, Harvey PD. The social cognition psychometric evaluation study: results of the expert survey and RAND panel. *Schizophr Bull*. 2014;40(4):813–823.

Pinkham AE, Penn DL, Green MF, Harvey PD. Social cognition psychometric evaluation: Results of the Initial Psychometric Study. *Schizophr Bull*. 2015;42(2):494–504.

Schneider LC, Struening EL. SLOF: a behavioral rating scale for assessing the mentally ill. *Soc Work Res Abstr*. 1983 Fall;19(3):9–21.

Wykes T, Spaulding WD. Thinking about the future cognitive remediation therapy—what works and could we do better? *Schizophr Bull*. 2011;37(Suppl 2):S80–S90.

Zimmermann G, Favrod J, Trieu VH, Pomini V. The effect of cognitive behavioral treatment on the positive symptoms of schizophrenia spectrum disorders: a meta-analysis. *Schizophr Res*. 2005;77(1):1–9.

▲ 58.7 A Sociocultural Framework for Mental Health and Substance Abuse Service Disparities

MARGARITA ALEGRíA, PH.D., KIARA ALVAREZ, PH.D., BERNICE A. PESCOSOLIDO, PH.D., AND GLORISA CANINO, PH.D.

INTRODUCTION

The Institute of Medicine (IOM) defined disparities in health care as "racial or ethnic differences in the quality of health care that are not due to access-related factors or clinical needs, preferences, and appropriateness of interventions." Responding to an important gap in that view, *The 2003 National Healthcare Disparities Report (NHDR)* expanded the definition to include differences in access associated with patient race, ethnicity, income, education, and place of residence. Although these reports strengthened public concern about glaring discrepancies in patterns of health service use and outcomes for racial and ethnic minorities in the United States, there has been limited progress in disparities reduction. Most of these disparities continue to be associated with the patient's race, ethnicity, and income profile and are a lost opportunity to eradicate the augmented age-adjusted mortality and increased burden of disease for ethnocultural populations. Despite the attention focused on health services disparities, particularly in view of demographic projections suggesting that ethnic/racial minorities will become the majority of the US population by 2044, there are a lack of strategic approaches that provide guidance for understanding the policy, social, and cultural factors that account for these disparities. Without frameworks and models, implementing the science-to-services translation called for by National Institutes of Health is difficult.

To date, numerous conceptual models have attempted to chart the potential causes of mental health and service disparities. Most of these models are based on the understanding that disparities in mental health and mental health care result from multiple factors, have no single underlying cause, and are not necessarily caused by same mechanisms across racial/ethnic groups. However, the majority of these models also link the existence of disparities to broader social, political, and economic inequality, as well as systemic racism, discrimination, and the historical subjugation of minority populations in Western society. Additionally, some frameworks detail a small set of factors related to a particularly important possible mechanism, while others seek to provide an overview of a broad range of possible causal mechanisms.

For example, some researchers have put forward models of health disparities based on the idea that residential segregation leads to differential exposure to environmental risk factors, particularly those that

affect psychosocial stress. Structural factors such as tax incentives and zoning laws affect the types of industries in minority neighborhoods, which may lead to increased exposure to environmental pollutants and lower neighborhood quality of life. Limited private and public investment in low-income areas can also result in poor quality housing stock and insufficient health infrastructure, leading to poor living conditions and reduced access to care. Physical (noise, pollution) and psychosocial conditions (social disorganization, crowding, discrimination, economic deprivation, and crime) in the neighborhood may augment allostatic load and damage health over time through what has been called the "weathering hypothesis."

Other scholars argue that the fundamental causes of disparities have to do with macrosocial factors such as historical, political, economic, and legal conditions, which are closely linked to the distribution of wealth, opportunities, and power. Intermediate conditions such as the built environment (i.e., community investment, public safety, enforcement of local rules, quality of education, and civic engagement) interact with the individual to cause stressors (i.e., violent crime and financial uncertainty) or buffers (like social integration) that can lead to differential health behaviors and mental health outcomes.

More focused theoretical models explore potential causes of disparities in a single area that may play an outsize role in causing disparities. One such model analyzes the effects of employment and work on health. Employment affects socioeconomic status, the types of conditions in which individuals spend much of their time, the type and quality of health insurance and other benefits available to individuals and their families, stress levels, and the regularity or irregularity of one's work schedule. All of these factors combine to impact how likely a person is to get sick over both the short and long term, whether and where an individual will seek care, the quality of care received, and the ability to continue with treatment (particularly if the individual is not allowed significant time off of work). Historically, the authors argue, minority individuals have been more likely to work in jobs that are low status, have lower pay and benefits, are physically demanding, and expose workers to stressors or even dangerous environments or substances.

Finally, some models center more closely on individual factors and life stresses that are caused by ethnicity/race and socioeconomic status, as well as community context, gender, and age. These may be connected with biological and psychosocial factors that increase or decrease vulnerability to stress, and consequently impact health behaviors, access to treatment, and mental health outcomes. In these models, disparities result from a complex interaction of race/ethnicity, poverty, gender, and minority status with social circumstances (i.e., supportive social relationships and family networks) and an individual's responses to those circumstances (i.e., optimism and personal agency), which impacts psychosocial reserve

capacity and cognitive-emotional processes.

There is clearly a great deal of overlap between the theoretical models discussed above, with most differences between models being merely variations in scope and emphasis. These different frameworks have accelerated disparities research by identifying some areas for research and intervention. However, they still largely fail to offer coherent, empirically based platforms upon which to develop public policies, systematic guidelines, professional training, or restructuring of health care systems to reduce service disparities. Building from past work and the theoretical literature reviewed above, this chapter outlines a culturally informed framework for mental health and substance abuse service disparities which integrates both the IOM and NHDR definitions of disparities by focusing on ethnic and racial differences in access to and quality of care. Although specifically designed to addresses mental health and substance abuse service disparities, the basic framework could be tailored to other health problems as well. In an effort to integrate sociocultural elements with the IOM and NHDR understandings of disparities into a health service disparities frame, a definition of the word *disparities* is necessary. Rather than utilizing the less-restrictive definition put forth in the NHDR of "any differences in service use among populations," the authors define disparities as *racial and ethnic differences in access, health care quality or health care outcomes that are not due to clinical needs or the appropriateness of treatment*. Patient preferences are also included as potential sources of disparities since they do not stem from "undue differential treatment on the basis of race and ethnicity." Patterns within the health care system influence racial and ethnic minorities' preferences in treatment, and patient-level factors linked to race or ethnicity inform minorities' preferences in treatment. As such, "undue differential treatment on the basis of race and ethnicity" may be operating through preferences. Thus, preferences cannot be disentangled from people's available options in care or from health literacy factors and can contribute to disparities.

This approach to understanding disparities in service delivery emphasizes that cultural processes refer not only to race/ethnicity as a particular set of values, norms and beliefs, but also to the circumstances and experiences leading to a group's social, economic and political status. Culture creates repertoires that structure human actions. In order to understand the relationship between culture and health, one must reject a simplistic view of culture or cultural difference and view cultural processes as a fundamental thread connecting communities, health systems, providers, and patients. The sociocultural framework (SCF) requires going past the use of sociodemographic categories as proxies for culture and moves toward understanding the cultural repertoires available to patients and how cultural scripts influence service disparities.

At the same time, the authors take into account cultural variation in the

prevalence of psychiatric disorders, expecting that these differences may be related to differences in exposure to risk and protective factors, the intensity or meaning ascribed to risk factors, the threshold at which a behavior is considered pathological, or the extent to which an environmental risk factor is mediated by genetic factors. In order to avoid inappropriate attribution of cultural factors influencing service disparities in minority communities, the authors adopt an approach that directs attention to life conditions and lack of resources that limit opportunities and affect the psychological development and well-being of minorities.

The Sociocultural Framework

As depicted in Figure 58.7–1, the SCF has three essential elements—systems, domains, and mechanisms. Two key systems that produce disparities are the health care system and the community system. Each operates at three domains or levels—the societal level (larger policy or environmental contexts), the organizational level (formal organizations or lay sectors), and the individual level (provider or patient). At each level, these domains interact or interface to produce mechanisms of inequality and bias that cumulate to disparities in health, health care, and quality of life.

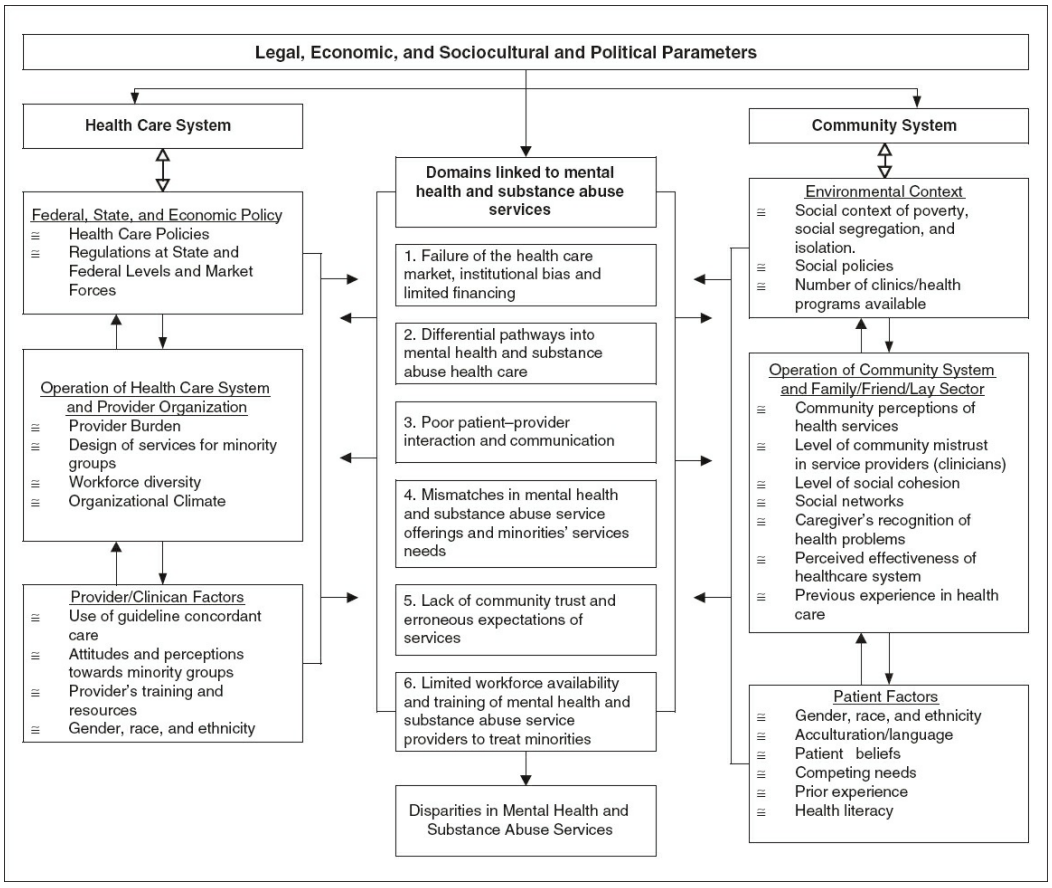

FIGURE 58.7–1. The sociocultural framework for the study of service disparities.

While the SCF for mental health and substance abuse service disparities emerged from the IOM disparities framework examining the role of the health service system in creating disparities, the SCF update expands the role of other policies and neighborhood characteristics in influencing disparities in health care. The authors follow the IOM's *Unequal Treatment* report in focusing on three domains or levels within the health care system: **(1) federal and economic health care policies and regulations at the state and federal level; (2) the operation of the health care system and provider organization; and (3) provider/clinician-level factors.** These domains are depicted on the left-hand side of Figure 58.7–1.

However, the SCF extends the IOM framework by adding the parallel focus on how the community system may differentially impact access to services or quality in services by considering the role that community, neighborhood characteristics, family, friends, and individual level factors play in ethnic and racial minority individuals' referral, entry, and retention in mental health or substance abuse services. These additional domains or levels are drawn from utilization, "take-up" or help-seeking models and include **(4) the environmental context, specifically social and economic forces affecting ethnocultural populations; (5) the operation of the community system, including family, friends, and the lay sector; and (6) patient-level factors.** These domains are depicted on the right-hand side of Figure 58.7–1.

The six domains depicted in the center of the framework impact mental health and substance abuse service disparities for ethnocultural populations through the six main mechanisms listed in the center portion of Figure 58.7–1. These include **(1) failure of health care markets, institutional bias, limited financing, and lack of ancillary services, leading to decreased access and quality of mental health care and substance abuse services; (2) differential pathways into mental and substance abuse health care, leading to differential experiences of treatment, no treatment, or inadequate treatment; (3) poor patient–provider interaction, leading to miscommunication, clinical uncertainty, low therapeutic alliance, stereotyping, and clinical errors; (4) mismatches in mental health and substance abuse service offerings and minorities' needs, leading to resistance or lack of compliance in health care and under-treatment; (5) lack of community trust and erroneous expectations of services, leading to disengagement and poor collaboration in treatment; and (6) limited workforce availability and training of mental health and substance abuse service providers to treat ethnic and racial minorities, leading to reliance on informal or poorly trained providers.**

In sum, at each level, the health care system and community forces interact to produce mechanisms that lead to differential health care outcomes in mental health and substance use care, including greater illness severity and diminished quality of life. The domains are linked by and through these mechanisms. Further, nowhere in the SCF is the nature and operation of "culture" partitioned in a separate box. All mechanisms and domains are set within a shaded region which represents *culture*. Every aspect of onset, recognition, identification, treatment, and outcome is embedded within a larger cultural frame (that includes the culture of medicine) that influences service delivery options and decisions as well as the culture of the patients shaping the lay community. It operates at all levels, serving as an overlay at each community and health care system level and is intimately involved in how domains at the same level interact to produce mechanisms of inequality. The following text discusses how the interaction at each domain may operate to impact mental health and substance abuse service disparities. Concrete examples of how the six main mechanisms may produce disparities are provided. Because service disparities result from the interaction between the domains at each level, some overlap occurs in the authors' descriptions. Furthermore, the chapter brings in concepts and findings not specific to mental health and substance abuse care as appropriate, to help identify potential leverage points to reduce disparities. Finally, the chapter acknowledges that the mechanisms linked to access may be different than those leading to poor quality of care or differential outcomes in treatment. Similarly, while a framework sets the parameters and foundations for explaining disparities, more theoretical work must be done on the SCF to create testable models for understanding the challenges which lead to mental health and substance abuse service disparities for diverse ethnic and racial groups.

HEALTH CARE SYSTEM DOMAINS

Federal, State, and Economic Policy

Federal, state, and economic policies impacting health and mental health care have the potential to create, exacerbate, or reduce health disparities. Due to their over-representation in government-sponsored programs, ethnocultural populations may face additional challenges as a result of state and federal regulations aimed toward cost control. For example, state differences in implementation of the Patient Protection and Affordable Care Act (ACA), the Children's Health Insurance Program (CHIP), and federal financing of Medicaid payments all have a disproportionate effect on minority populations. Each is discussed, in turn, below.

Affordable Care Act and Medicaid Expansion. A key provision of the ACA is the expansion of Medicaid to include adults with incomes up to 138

percent of the federal poverty level ($16,104 for an individual or $27,310 for a family of three in 2014), in order to increase the number of Americans who are insured. However, as a result of the Supreme Court ruling on ACA provisions, individual states are allowed to decline the Medicaid expansion. As a result, more than 4 million Americans are uninsured due to a coverage gap. They do not qualify for Medicaid in their state, but do not earn enough to qualify for the insurance tax credits available to those making 100 percent of the poverty level or above. Since minorities tend to be over-represented among the poor, the flexibility of states to set their own eligibility criteria has implications for minorities' access to and quality of care. An estimated 4 in 10 uninsured African Americans fall into the coverage gap. Latinos, who have the highest percentage of uninsured when compared to other racial/ethnic groups, are also affected. The percentage of uninsured Latinos dropped from 36 to 23 percent after the enactment of the ACA, according to a 2014 survey by the Commonwealth Fund, but rates of uninsured Latinos remained unchanged in states that had not enacted the Medicaid expansion.

Immigrants of all backgrounds face particular challenges under current health insurance laws. Those who are lawfully present are subject to a 5-year waiting period for Medicaid and CHIP, as well as exclusions from these programs for some classes of immigration status. States may waive the waiting period for pregnant women and children, but not all states choose to do so. Undocumented immigrants are further disadvantaged, as they are not eligible for Medicaid at all, are not eligible for tax credits used to make private health insurance more affordable, and are not eligible to purchase health insurance via the federal health insurance exchange.

Lack of access due to not having insurance serves as a potential mechanism promoting service disparities. Not surprisingly, it has been argued that expanding insurance coverage for minorities would reduce disparities. However, knowledge about health care insurance options limits access even for the insured. For example, low-income and Spanish-speaking Latinos who are potentially eligible for insurance via health care marketplaces are much less likely to know about them when compared to high-income and English-speaking Latinos.

Low Reimbursement Levels and Provider Shortages. A major challenge in service provision for low-income populations is ensuring the availability of providers who accept Medicaid patients. In 2011, 31 percent of physicians surveyed stated that they would not accept new Medicaid patients, compared to 18 percent who stated they would not accept new privately insured patients. There are lower rates of psychiatrists accepting Medicare, Medicaid, and noncapitated insurance when compared to other specialties, and rates are decreasing over time. In 2009 to 2010, only 43.1 percent of psychiatrists accepted Medicaid and 55.3 percent accepted

noncapitated insurance, compared to 73 percent of physicians from other specialties who accepted Medicaid and 88.7 percent who accepted noncapitated insurance. Rates among psychiatrists represented a 19.5 percent decrease since 2005 to 2006. This limitation seriously restricts the available pool of providers willing to serve minorities, who are disproportionally represented in Medicaid services. Low participation of physicians is particularly concerning, given the expansion of Medicaid to an estimated 17 million people under the ACA.

Low reimbursement rates, payment delays, and billing requirements have been noted as a primary reason for not taking Medicaid patients in physician surveys. Systemically, providers within organizations serving large numbers of Medicaid patients may experience high expectations for productivity in order to make services financially feasible, which can increase provider burnout. There may also be a need for providers to spend additional time coordinating care for more vulnerable patients. The implementation of accountable care organizations (ACOs), which bring together groups of providers and provide more flexible payment structures to incentivize positive outcomes, may be beneficial in reducing health care disparities. Expansion of safety-net providers such as community health centers (CHCs), another provision of the ACA, is another avenue by which health disparities may be reduced. However, the long-term sustainability of CHCs is partially dependent on the revenue that they can continue to generate from patients, which is more challenging in states with high proportions of uninsured patients.

Children's Health Insurance Programs. CHIP is designed to make free or low-cost health insurance available to children without insurance but who are ineligible for Medicaid. As such, it plays a crucial role in reducing access barriers that create health service disparities among minority youth. In the first 10 years after CHIP was implemented, rates of uninsured children dropped from 30 percent of Latino children to 20 percent, 19 percent of black children to 12 percent, and 17 percent of Asian, Native Hawaiian, and Pacific Islander children to 12 percent. A study of CHIP implementation in New York found that pre-existing racial/ethnic disparities in access, unmet need, and usual source of care between white, Latino, and African-American youth were virtually eliminated during the first year of enrollment. However, racial and ethnic minority children continue to be less likely to be enrolled in health insurance when compared to white children, even when eligible for CHIP.

In 2009, the Children's Health Insurance Authorization Act (CHIPRA) established provisions aimed at reducing health disparities, including providing outreach grants to increase enrollment, increasing federal matching rates for interpreters and translators, providing federal funding for medical care for immigrant children in states that waive the 5-year

insurance waiting period, establishing quality of care measures to track health outcomes for children, and including specific provisions addressing health concerns that disproportionately affect low-income children and children of color. Across states that enacted the additional CHIPRA coverage for immigrant children, health insurance rates for this population increased by 14.5 percent and unmet health need decreased significantly.

Mental Health and Substance Abuse Parity. Laws aimed at achieving parity between insurance benefits for general health conditions and insurance benefits for mental health conditions have been enacted in various ways since 1996. In 2008, the Mental Health Parity and Addiction Equity Act (MHPAEA) was passed, and the ACA subsequently expanded the reach of the act to more types of insurance plans. The MHPAEA does not require plans to offer specific mental health benefits, but rather to ensure that the benefits offered are equal to those offered for other health conditions. Implementation of the law has been a complex process, resulting in final regulations being issued in 2013. While parity provisions are a promising step forward in access to care, it remains to be seen how the enactment of current regulations may or may not impact health service disparities.

National Standards for Cultural and Linguistically Appropriate Services. Policies can directly address issues relevant to minorities. The National Standards for Cultural and Linguistically Appropriate Services (CLAS) were published in 2000 by the Department of Health and Human Services (DHHS) and revised in 2013. The CLAS standards provide guidelines for health care organizations addressing three areas: governance, leadership, and workforce; communication and language assistance; and engagement, continuous improvement, and accountability. These standards have been used to shape service delivery systems, as well as by some states to propose or pass legislation. Organizations are urged to provide trained interpreters, as well as linguistically and culturally adapted medical forms for non-English speakers. While the CLAS standards are not a federal law, failure to meet language access recommendations may be a violation of Title VI of the Civil Rights Act of 1964, which prohibits discrimination in organizations receiving federal funding. DHHS-funded organizations must provide meaningful access to services to non-English speakers, including trained interpreters.

Policy Supporting Innovation. Policy changes can also facilitate the use of innovative strategies to address health disparities. In 2013, the Centers for Medicare and Medicaid Services adopted a regulation change allowing states to choose to reimburse for physical and mental health services provided by unlicensed providers, as long as a physician or other licensed

practitioner has recommended the service. This opens a possibility for states to fill in gaps in service availability by engaging providers such as community health workers and other lay providers. The use of lay providers to deliver mental health services has been successful in several major studies carried out in low- and middle-income countries, and may be applicable in under resourced communities within the United States.

Technology used to track delivery of patient care and implement quality measures is another promising avenue via which to reduce health disparities. Advances in health information technology (HIT), including electronic medical records (EMR), have been a critical component of national health policies. The ACA calls for improved reporting requirements, while CHIPRA has promoted the development of standard quality measures as a way of tracking and addressing health care needs. The IOM noted in a 2009 report that more detailed information on racial, ethnic, and linguistic background should be included in data collection efforts, in order to target efforts to eliminate health service disparities. Combining detailed data collection with quality measures allows identification of gaps in care for specific groups and to respond with specific programmatic improvements.

In sum, the larger parameters of a society shape the health care system. At the federal and state level, specific policies on health care, as well as economic policies and regulations, create a set of market forces. In turn, the nature of these policies interacts with community context to impact ethnic and racial minorities in the health care system. Both the intended and unintended consequences of these policies on health care organizations are described next.

Operation of the Health Care System and Provider Organization

As a result of the system level factors described above, health care systems face differential costs and opportunities associated with treating ethnic and racial minority populations. As noted above, reimbursement by the state or the federal government holds a central role in shaping provider burden, service design, workforce diversity, and organizational climate of health care organizations.

Cultural and Linguistic Competency in the Operation of the Health Care System. Health care policies designed to encourage cultural and linguistic competence, such as the CLAS guidelines, are expressly designed to counter discrimination. Yet the ways in which a policy is implemented impacts client care. Language barriers for non–English-speaking immigrant groups create major challenges in access to care. In one study alone, language barriers between provider and patient accounted for a $38 increase in charges for testing, and increased length of stay in the emergency department. Patients who speak a different language from their

providers report worse outcomes and higher drop-out rates, and may experience misdiagnosis or poor treatment based on misunderstandings in assessment. However, even the use of interpreter services by competent providers has been shown in some instances to result in lower quality care. At the same time, improving linguistic competence within a system results in additional costs and time investment. Health systems may be reluctant or simply unable to cover these resources. Furthermore, even if willing, individuals may not be readily available or trained to fulfill the required services. Policy efforts such as state legislation based on the CLAS standards may be beneficial in reducing health disparities, but resources and training must be available to health care systems to finance and implement organizational change.

Integrated Health Care. The integration of behavioral health services into the primary care setting has been advanced as a strategy to improve behavioral health outcomes for minority populations, given studies that indicate that members of racial and ethnic minority groups are more likely to seek mental health care via their primary care providers (PCPs). At the same time, there is still limited data available on how to best implement high-quality behavioral health services for minority populations in primary care. Studies have found that PCPs are less likely to detect mental health concerns among racial and ethnic minority patients than white patients. The pressure to treat behavioral health challenges in the time-limited setting of primary care, in addition to constraints in specialty care options available for PCP referral, may result in situations in which clients are offered lower quality services or do not follow up on treatment. One study found that Latino primary care patients diagnosed with depression were more likely to follow up on and complete treatment via medication alone, despite the majority of the patients expressing a preference for psychotherapy over medication. These patients were able to access medication directly via their PCP, but may have experienced challenges in accessing therapy services. Research on models of integrated and collaborative care, as well as their impact on specific racial and ethnic minority populations is needed in order to explore how this approach may or may not reduce health service disparities.

In other ways, the design of health care systems and services contributes to disparities. When health care organizations do not offer effective treatment for behavioral health problems, issues such as alcoholism, substance dependency, or posttraumatic stress disorder affect health status directly and compound other health problems indirectly. Further, offering a service-level spectrum, whereby good quality health services are offered for services that attract the healthy, but poor quality services are offered for the sick or those with chronic conditions (labeled adverse selection), tends to discourage use and hinders good outcomes.

Organizational Climate and Institutional Racism. Outcomes research has demonstrated that the climate of attitudes and beliefs and culture (actual behaviors) in the provider organization toward ethnocultural populations affects the receipt of quality care. Ethnic and racial minority patients might be viewed as less "appealing" to the health care providers if they are thought to present with a disproportionately greater risk of requiring more intensive treatment. For example, minorities may come to mental health and substance abuse care at higher levels of illness severity and terminate treatment early. In a capitated managed care environment, these factors might make ethnic and racial minority individuals less attractive to the health care system. In addition, these patients may also require referrals and networking with other social and service agencies—work for which the health care provider is not compensated. These hidden costs could have an effect on the desirability of treating ethnic and racial minority patients that may translate into negative attitudes, or what has been labeled institutional racism. Use of evidence-based strategies for organizational change, such as the ARC (Availability, Responsiveness, Continuity) intervention model, may be an avenue by which to change organizational attitudes. In this model, "change agents" are trained to educate providers, advocate for systems-level change, provide implementation data and feedback, and facilitate communication within the organization and with the broader community.

Design of Services for Minority Groups. While the evidence is inconclusive on whether the lack of targeted services for diverse patient populations plays a central role in service disparities, a meta-analysis showed that culturally adapted interventions were four times more effective than those not targeted to specific cultural groups. Further, interventions conducted in the client's native language were twice as effective as those conducted in English. Yet, structuring for racial and ethnic concordance, or matching between patient and provider is one area of heated debate. Race concordance for African Americans has been associated with higher patient participation in decision making and greater satisfaction with care. Patient–provider matching also tends to improve retention of clients and intervention outcomes and may increase entry and adherence to treatment. If this is, indeed, the case, professional services in culture-specific clinics might be a tool for providing better services to ethnic minorities. However, there is limited evidence on whether or not they reduce service disparities. While ethnic similarity may facilitate retention in therapy, it may have only a minor role in therapy outcomes. Similarly, physician's race concordance with patients does not appear to be associated with either use or quality of care, only with preventive care.

Clients are likely to respond to the therapy in ways that are consistent with their cultural socialization regarding care. Therefore, culture most

likely defines the options perceived to be available, useful, and/or desirable manifestations of the therapist and client roles. Patient preference for one theoretical orientation over another might characterize members of cultural groups. Researchers hypothesize that therapists and clients who are ethnically similar may be more analogous in terms of problem perception and attitudes about coping and treatment goals. As a result, these coinciding factors may improve the effectiveness of therapist/client interactions.

This is not to say that all racially/ethnically concordant patients and providers will share the same attitudes and beliefs. Research has demonstrated that *cognitive* match between patient and provider is most important, with evidence of this significance persistent even after controlling for ethnic match and the client's language preference. These findings have implications for practice/training and for the need to assess discrepant orientation and treatment ideas.

Provider-Level Factors

Use of Guideline Concordant or Evidence-Based Treatment. Providers' reluctance to adopt evidence-based treatments (EBTs) when serving minorities has been well-documented, leading to service disparities. The question remains as to whether the deviation from EBT is greater among providers who serve primarily ethnic and racial minorities or whether providers are less likely to provide EBT to their minority clients. Evidence has been found for both; clients who are poor or African Americans and Latinos have been found to be less likely to be offered the more rigorous or intensive care options provided to those who were more advantaged, but also facilities serving mostly ethnic and racial minority patients were less likely to offer evidence-based care. However, many health care institutions do not provide sufficient support for providers to adopt quality behavioral health care within the pressures for productivity.

Research examining the use of medical procedures across chronic illnesses, such as depression, has consistently found a lack of referral of minorities for EBTs when compared with non-Latino whites, even when minorities meet clinical appropriateness criteria for EBTs and when age, gender, and insurance status are controlled. Although longitudinal studies examining changes in trends for disparities in EBT among Medicare managed care patients indicate that racial disparities in EBT have declined for some illnesses, for mental health care the gap in EBT for minority patients remains.

Provider's Attitudes and Perceptions: Uncertainty, Stereotyping, and Bias. Providers seldom receive training in how to motivate minority patients to voice their treatment concerns or preferences. Providers display fewer patient-centered behaviors, are less receptive to question asking, and tend

to demonstrate greater verbal dominance with minorities than with white patients. This can lead to miscommunication or misunderstanding between the patient and provider, resulting from cultural or language differences, that contributes to uncertainty. If providers are more effective in understanding and delivering appropriate services to their nonminority patients, treatment and outcome disparities could result. Uncertainty may also enter the clinical encounter when providers experience ambiguity concerning the diagnostic implications of clinical symptoms, signs, or tests due to atypical symptom presentation, or when the provider has incomplete information about the efficacy of a therapeutic intervention. In part, these problems result from the under-representation of minorities in clinical trials which establish the efficacy or effectiveness of EBTs.

Stereotyping of ethnocultural patients by health care professionals has also been identified as a potential source of service disparities. Various studies have revealed how providers may hold preconceptions about patients based on their group membership, resulting in provider stereotypes. While stereotyping may stem from unconscious bias, reflecting limitations in cognitive mechanisms, the effects on disparities are real nonetheless. Categorizations such as race, age, sex, and socioeconomic status are overgeneralizations. For example, providers tend to see black patients as less educated than non-Latino white patients. Further, psychiatrists characterized African-American patients who were indistinguishable from white patients in sociodemographics and clinical data (except for race) as "less articulate, competent, introspective, self-critical, sophisticated about mental health centers, and psychologically minded" and consequently less able to profit from psychotherapy. Finally, provider-based biases have also been posited as negatively influencing clinical decision making. As a result, both over and underdiagnosis could result from these stereotypes and negative biases.

Overdiagnosis occurs when appropriate behaviors in one culture are interpreted as pathological by a professional from another culture. These can simply result in more severe diagnosis attached to individuals as their social and cultural distance increases. In more extreme cases, certain cultural or religious beliefs and practices common among Latinos and African Americans accept, and sometimes encourage, visual and hearing perceptions of supernatural experiences, which are often interpreted by professionals from another culture as pathological hallucinations. In an epidemiological study in Puerto Rico, 13 percent of the population reported some kind of hallucination. Yet a qualitative analysis of these symptoms by trained local psychiatrists revealed that they occurred mostly in the context of religious or spiritualist experiences, and were not pathological. Incorrectly diagnosing a psychotic disorder where none exists could result in inappropriate hospitalization, or in the use of potentially dangerous antipsychotic medication. Not surprisingly, recent epidemiological and

clinical research has raised questions about the relationship between the endorsement of psychotic symptoms and the diagnosis of psychotic disorders among culturally diverse groups.

Underdiagnosis occurs when clinicians ignore genuine manifestations of mental illness, misguided by a misinterpretation of cultural sensitivity. A growing body of research suggests that the absence of cultural and contextual information contributes to an elevated probability of misdiagnoses and inappropriate treatment formulation. Several studies have documented that minority patients were more likely to report that their provider did not discuss end-of-life care, and were less likely to receive recommendations for how to quit smoking or engage in diet and exercise. Providers may unintentionally or intentionally dismiss minority patients' concerns or may not communicate about health promotion, psychoeducation, or early screening for mental disorders. Only one-quarter of Latinos surveyed in the National Latino and Asian American Study reported ever being asked about alcohol or drugs. Fewer than 20 percent of Latino patients reported ever being asked about emotions, nerves, or mental health problems. This limited inquiry may lead to missed opportunities for early identification of mental health or substance use disorders and lack of referral to specialty services.

Clinician Training in Cultural Responsiveness. While cultural responsiveness training has been offered as a solution to the problems described above, it is not an integral part of the education of many health professionals. Health care personnel typically have limited training in cross cultural formulation and cultural humility, leading to a lack of concordance in defining the problem and treatment goals. There is also evidence that providers signal their expectations about patients' outcomes and whether they believe there will be improvements, which may be related to actual patient health outcomes. For example, health care professionals may communicate lower expectations of treatment outcomes for socially and economically disadvantaged patients, which could discourage minority patients and lead to poor outcomes and low treatment retention.

THE COMMUNITY SYSTEM

Environmental Context

Neighborhood context influences both individual health and health behaviors. Residing in socioeconomically disadvantaged neighborhoods has been linked to poor health outcomes, due to structural factors (e.g., increased exposure to pollutants and decreased access to health-promoting spaces such as parks and playgrounds); limited access to health care facilities and health care providers; poor quality of available health care; stress related to poverty and neighborhood violence; and decreased social

connectedness. At the same time, recent studies demonstrate that the effects of neighborhood context vary for different groups.

In two studies investigating neighborhood effects on substance use and on depression and anxiety symptoms in a nationally representative sample, neighborhood affluence actually resulted in increased risk of substance use disorders for Asians and African Americans, and increased risk of depression for African Americans. Latinos living in neighborhoods with high Latino/immigrant concentration had decreased risk of alcohol use disorder, but increased risk of anxiety disorders. Neighborhood context may affect groups differently based on other contextual factors; for example, Asians and African Americans residing in affluent neighborhoods may experience more discrimination and feel a greater sense of social isolation. Latinos residing in predominantly Latino/immigrant neighborhoods may experience protective factors of social cohesion, reduced exposure to daily discrimination, and increased sense of collective identity, but may also be subject to increased stress due to residential segregation and to the collective impact of anti-immigration policies on neighborhood cohesion.

The geographical area in which minority populations live helps to create a social context which, in itself, fosters a lack of connection with treatment systems and the dominant culture. Geographical segregation continues, perhaps even at greater levels than would be expected for minority populations. Members of immigrant groups have a variety of experiences with resettlement depending on the time period in which they arrived in the United States, the context of entry, and the political and social climate toward the particular group at the time. Thus, the social context and the geographical context are often intertwined and overlapping for ethnocultural populations. This impacts health on multiple levels, from prevalence of particular conditions to access to and use of mental health and substance abuse services.

Operation of Community System and Social Networks Sectors

While it is difficult to draw a clear line between the environmental context and the community, the difference is important. While the larger social context can shape lower levels of social organization, the community is often seen as representing the "small worlds" of individuals—that is, the actual nature of social ties that individuals have in a geographic area as constituting their "community of meaning." With regard to health and health care, the neighborhood and local social networks are seen as key.

Level of Social Cohesion and Neighborhood Violence. Community systems differ dramatically in a series of characteristics that both can facilitate the use of services and offer protection against health problems. In the SCF, community is defined as "a group of people with a similar

cultural origin and diverse characteristics that are linked by social ties, share common perspectives and engage in joint action in geographical locations or settings." Community social cohesion may play a role in health disparities by influencing factors that affect service outcomes. For example, there is evidence that organized and cohesive communities are able to ensure that state budget cuts do not affect their community and this in turn allows them to maintain increased access to funding for health and education.

Social Capital and Social Networks. Service systems working with ethnocultural families need to consider the important role of extended family as a possible source of support or burden in treatment. Another important component is the effect that family and social networks may have on inhibiting or facilitating help-seeking. Mexican American and African-American adolescents are more likely than whites to turn to relatives or family members to obtain health care. Thus, in strong and tightly meshed family networks, such as those of most Latinos, African Americans and Asians, individuals may initially be discouraged from seeking professional health services. It is only when family support systems are lacking, or become ineffectual in handling a crisis, that some culturally diverse families will seek professional help. Some investigators have suggested that informal care providers, such as self-help groups and religious leaders, reduce seeking of professional help and may be related to service disparities; others argue that they are an important referral source to formal services. Innovative strategies that leverage the influence of community networks, such as programs to promote HIV education and testing in African-American churches, are a way to embed health care services in the settings in which individuals and families naturally seek help.

Individual Factors Other than Need or Clinical Appropriateness

General Education. Number of years of schooling is directly linked to the development of effective capabilities that influence subsequent life conditions that have an effect on health, including economic welfare, healthy lifestyles, and full-time employment. Higher rates of education have been shown to positively impact mental health by increasing one's resources to cope with adversity or increasing a sense of control over one's destiny. Low educational attainment is not only a strong predictor of poor mental health and substance abuse outcomes for children but remains significant for mid-age and older cohorts.

Health Literacy and Numeracy. Health literacy is a skill involving both oral and written communication that may or may not be related to educational attainment. It is a twofold concept that includes the

individuals' basic literacy skills and the application of those skills within the health care system. A related concept, health numeracy, involves not only understanding of mathematics but also of quantitative concepts such as "reduce/increase." The IOM has highlighted that health literacy is not solely an individual skill; rather, it results from the interaction between individual capacity and the demands of the health care system. Health literacy may indirectly create barriers to health care use by inhibiting access to care due to a patient's inability to register for health insurance; difficulty in interpreting coverage benefits and rules; avoidance or delays in accessing health services due to the inability to adequately fill out forms in physician offices or difficulty following directions to service facilities. Once individuals with poor health literacy enter care, they face disparities in health outcomes associated with literacy limitations, with people who read at lower levels being 1.3 to 3 times more likely to experience adverse health outcomes than individuals with high literacy levels. Patients with low health literacy are also more likely to report poor health, have an increased risk of hospital admission, and have incomplete understandings of their health problems and treatment. In one nationally representative study of US parents, almost a third had below basic health literacy. The strongest predictors of basic and below basic health literacy were level of education and English proficiency, followed by being Black or Latino, being born outside of the United States, and reporting low income.

Negative Provider Stereotypes. Patient views of providers may prove to be influential factors in causing disparities for ethnocultural populations. For example, some literature suggests that individuals who hold strong negative stereotypes of physicians are more likely to delay care or only seek treatment once symptoms become severe or impede daily functioning. Therefore, if delays in treatment are due in part to negative stereotypes, service disparities in illness prevention or early intervention may emerge. Negative stereotypes of physicians may result from patients' own history of negative treatment in care or the distrust many minorities feel toward the health care system.

Patient's Competing Needs. Some ethnic and racial minority patients may come from low-income backgrounds with complex social needs. Lack of access to paid sick leave, inflexible work scheduling, irregular hours, lack of transportation or unreliable transportation, and difficulty obtaining childcare all pose additional challenges to utilizing health care services. Studies have found that these barriers account for both missed appointments and lack of follow through on specialty referrals.

Problem Recognition and Characterization of Mental Health and Substance Use Problems. Disparities in health care among some racial and ethnic

minority children appear to be linked to parental knowledge of available treatments. For example, it has been reported that there are significant differences between African-American and white parents in their knowledge of attention-deficit/hyperactivity disorder (ADHD). In one study, only 36 percent of African-American parents reported that they knew some or at least any information about ADHD, compared to 70 percent of whites. There is also evidence that different ethnocultural groups in the United States have varied perceptions of what is considered a mental health problem in their offspring. In one study, Latino and African-American parents rated their adolescents' mental health and life satisfaction as better than their non-Latino white counterparts and overall perceived fewer emotional or behavioral problems in their children. These differences existed even after adjusting for clinical need of the child, child's level of impairment, and sociodemographic characteristics of the parents. Thus, while lack of knowledge about a problem may lead to health disparities, it is also possible that different perspectives on a situation may lead to different patterns of seeking care.

WHEN HEALTH CARE SYSTEMS AND COMMUNITY SYSTEMS MEET: MECHANISMS PRODUCING DISPARITIES

Taken together, disparities arise when disadvantages in the health care system interact with those in the community. At each level or domain, the SCF suggests that six fundamental mechanisms produce disparities for ethnocultural populations. These are represented as the center stream in Figure 58.7–1. While they link with causes of disparities, they also suggest points of intervention. They represent potential leverage points for systems, providers and patients to change the health care system, existing policies, clinical training, and community programs. Below, each interface between community and treatment systems is explored.

1. **Failure of health care markets, institutional bias, and limited financing, leading to decreased access to and utilization of mental health and substance use services.** State and federal regulations aimed toward cost control also contributed to the increase in the number of uninsured or underinsured. For example, individuals are typically excluded from Medicaid as states respond to increased costs by enacting more stringent eligibility criteria. Yet expansion of public insurance programs and subsidized private plans to low- and middle-income individuals offered under the ACA can increase access to care. Nonetheless, some minority groups may confront other hurdles to getting insurance such as lack of citizenship or documentation status or may still have difficulty paying for subsidized coverage, leading to lack of insurance or underinsurance. While health insurance coverage facilitates access, it is not a sufficient condition for ensuring behavioral health care, since lack of

available appointments is a serious barrier to mental health and substance treatment. The low number of mental health and substance use providers accepting Medicare and Medicaid reduces access to quality care for low-income and vulnerable populations. However, other factors such as increased health maintenance organization (HMO) penetration, augmented provider supply, and the presence of a community mental health center in neighborhoods of color have been shown to reduce service disparities.

2. **Differential pathways into services leading to differential experiences of service, delayed utilization, no service, or inadequate service.** Language barriers can lead patients to drop out of care, opt for alternative services, or delay care. Inaccuracies in assessment can also result in mistreatment or misdiagnosis of minority patients. There is mounting evidence that minority patients receive lower quality treatment and receive treatment for a shorter period of time. For example, white youth are more likely than minority patients to receive adequate care following a major depressive episode, while minority youth are less likely to be prescribed psychotropic medications, even after adjusting for need factors. Equally importantly, the larger cultural context determines how similar social networks can facilitate or delay treatment. Competing demands on time from work and child care for minority parents may constrain minority families from engaging in treatment, given more urgent matters.

Thus, differential pathways into services may represent a key mechanism leading to service disparities among racial/ethnic minorities, with lack of patient problem recognition, differential referral and treatment by providers, and barriers to care serving as potential precipitating factors to disparities. Further, it is likely that an individual's first service contact may not be with a specialty professional, but rather with a formal or informal gatekeeper provider who has a powerful role in determining whether and where to refer a patient for further treatment. Factors that may influence a provider's perception of a patient's need for referral include the providers' familiarity with and links to health service resources, the provider's perception of the clients' willingness to accept mental health services, and organizational boundaries that limit the gateway providers' time or resources. Improved behavioral health screening for Medicaid-eligible minority youth and adults has proven to decrease disparities in access. Augmenting caregiver's resources for accessing services and improving child outcomes, or engaging families to attend their first intake therapy session have also been shown to increase access and reduce disparities in families of color.

Other pathways result in disparities that may have both short- and long-term effects. For example, Latino and African-American children are over-represented in the juvenile social service sectors. They receive fewer

therapeutic services compared to clinically comparable white children, and remain longer in services. As a result, data analytics are being used to detect patients at risk of receiving lower quality care based on demographic and illness related factors, and payment systems used to incentivize providers for providing high-quality care to these patients. Nonetheless, there is considerable debate over the utility of pay-for-performance measures in achieving specific goals at the level of hospitals and the health system as a whole. ACOs may help eliminate disparities by rewarding health care systems for efficient use of resources and positive outcomes and penalizing them for inefficiency and poor results.

3. **Poor patient–provider interaction leading to miscommunication, clinical uncertainty, low therapeutic alliance, stereotyping, and clinical errors.** Patient–provider match and expectations for the visit are important in determining treatment outcomes. However, attempts to address research documenting problems arising from poor patient–provider interaction have inadvertently produced additional problems. Ethnocentric models of health service delivery may overpathologize minority behaviors that differ from those of the majority culture. Racial and ethnic minorities report being treated with disrespect or being looked down upon more often (14.1 percent of blacks, 19.4 percent of Latinos, and 20.2 percent of Asians) than non-Latino whites (9.4 percent). The critical link to disparities lies not only with what happens in the clinical encounter but with the resulting effects on the patient's utilization of health care and on the advice they offer to others in their lay network. For example, patients who reported being treated with disrespect were less likely to go for a routine physical exam within the least year, to have received optimal care for a chronic condition, and to have followed their doctor's advice.

Patients may also have different expectations of treatment and provider behavior that impede effective communication. A number of studies have reported that differences in race and ethnicity between patients and their providers can present barriers to effective communication and partnerships for care. Disparities resulting from barriers to effective communication between racially/ethnically mismatched patients and providers include discordance in the understanding that patients and providers have regarding disease causation and effectiveness of treatments, significantly lower reports of physician participatory decision-making styles by minority patients as compared to nonminority patients, and lower rates of receiving preventative and necessary medical care among blacks with nonblack physicians compared to blacks with black physicians. These examples demonstrate that patient–provider match and expectations for the visit are important in determining treatment outcomes.

4. **Mismatches in mental health and substance use service offerings and minorities' services needs and living**

circumstances, leading to drop out and low or no service use. Services disparities have been linked to higher disability and greater disease burden in minority populations. One of the mechanisms by which service disparities take place is through the mismatch of services to the patient's needs. As described earlier, there is substantial evidence of the potential for clinicians to stereotype—that is, to incorrectly assign a classification or social categorization, particularly when the patient's request is ambiguous. Under such conditions, clinicians may inadvertently incorporate attitudes and beliefs from prior clinical encounters or common stereotypes. This may lead clinicians to disregard individual data and consequently mismatch services to needs. This mismatch of services to needs can result in decreased utilization of and access to health services and health resources as well as differential outcomes in health, and greater disease burden. In the absence of cultural and contextual information that supplants stereotypes, there is an elevated probability of misdiagnoses and inappropriate treatment formulation. Ethnocultural populations are more likely to have nonconcordant psychiatric diagnoses in psychiatric emergency services and to be undetected for mental health disorders in primary care. More general literature on disparities shows that symptom presentation varies across racial and ethnic groups and can differ from what most clinicians are trained to expect. Yet, minorities are less likely to be offered "talking therapies" such as counseling, and tend to be prescribed higher drug dosages than their white counterparts.

Dissimilarities in the typical conditions in which many minority populations live as contrasted to their providers (i.e., including poverty, lack of economic opportunity, food insufficiency, inadequate housing, social exclusion, low access to educational opportunities, and unsafe neighborhoods) can lead to social inequality that needs to be tackled by the health care system. As such there has been a call for health equity research that prioritizes multilevel interventions, addressing structural and contextual factors at the community, organizational, and individual levels. Multilevel interventions to amplify social connection in minority neighborhoods could increase behaviors that have a positive or protective effect on health such as exercise, expansion of social networks, and better problem solving.

5. **Lack of community trust and erroneous expectations of services, leading to disengagement and poor collaboration in health care service systems.** Local perceptions of health services affect utilization and depend on prior experience with the provision of services, and the efforts that the local service system has made to understand and adapt to the cultural and need demands of its clientele. For example, qualitative analyses of an Afro-Caribbean community revealed that their experience of mistreatment and social exclusion by health professionals reverberated on future utilization and on community sentiments toward

the mental health system. Other studies found that the level of community mistrust of health care providers was associated with a decreased likelihood of using health care. For example, African-American families report fearing that their children would be institutionalized and express this fear as an important reason for not taking their children to treatment. These findings, as well as those of others, reveal that insufficient accessing of services increases with repeated negative contacts with the health system or with a lack of trust in key gateway providers such as teachers. Thus, community-held beliefs may prove to be strong influential factors in the production of health service disparities for ethnocultural populations.

6. **Limited workforce availability and training of service providers to treat ethnic and racial minority youth, leading to reliance on informal or poorly trained providers.** The disproportionately low geographic supply of providers, particularly multilingual service providers, in communities with ethnic and racial minorities as compared to nonminority communities is producing a failure of the market to provide more and better quality care to minorities. For example, communities of color have been found to be four times more likely to have a shortage of providers, regardless of community income. Low reimbursement rates from public and private payers can lead to a shortage of services and provider burnout. Health care organization arrangements for distributing providers to minority Medicaid beneficiaries have also been suggested to be influential in service disparities. Merely 7.8 percent of psychologists, 12.9 percent of social workers, and 21.4 percent of psychiatrists are from racial/ethnic minority groups. Medicaid beneficiaries, particularly those with limited English language skills, may have a difficult time competing for physician services if physicians have a high caseload of privately insured patients.

Communities with high proportions of black and Latino residents were four times as likely as non-Latino whites to have a shortage of physicians regardless of community income. In parallel, health care services available through the Indian Health Services vary widely across tribes. The Indian Health Services partially offsets the lack of insurance for some uninsured American Indian and Alaska Natives (AIANs), but important needs are left unmet. This is problematic since AIANs were more than two times as likely to be uninsured than non-Latino whites. These findings have led to conclude that elimination of racial/ethnic health disparities is an issue of distributive justice which entails redistributing valuable material resources and opportunities to those in the lower echelons of society.

In sum, the interaction of health care system and community level factors produces a mismatch in cultures—not only the culture of medicine and the culture of everyday life but also in the mismatch of dominant culture ideas and those that may be unique to ethnocultural populations.

Oftentimes, the beliefs of both patients and providers are based on generalized cultural beliefs and do not cut through those prejudices to get to individualized care. Equally important, both actual experiences and a family's, provider's, or community's experiences can produce a view of treatment that limits the receipt of effective care.

FUTURE DIRECTIONS: RESEARCH AND INTERVENTION

The SCF is intended to help conceptualize and carry out research directed at eliminating ethnic and racial disparities, to assist in the training of health and social agency personnel, and to supply service planners and policy makers working hypotheses, initially, and a science-base, eventually, on how to intervene and eliminate service disparities. As evidenced from the authors' review of extant theory and research, traditional service models should be restructured in order to reduce service disparities for ethnocultural populations. The six posited mechanisms that produce disparities form a foundation for rethinking intervention strategies as (1) increased client/patient activation/empowerment; (2) information exchange between client and provider, between systems and communities addressing social networks rather than individuals alone; (3) new and/or different policies to increase access to services; (4) better treatment matching; (5) restructuring of service systems to reduce reliance and action based on stereotypes, prejudice, and clinical uncertainty.

1. **Patient activation,** defined as learning the knowledge, skills, attitudes, and self-awareness necessary to have greater decision making and efficacy in health care behavior, has been associated with improved health behaviors such as exercise, disease self-management behaviors such as adherence to drug regimens, and an array of health outcomes including improved mental health. There is evidence that patient activation enables patients to take charge of their health behavior, elicit more information from their provider, achieve greater satisfaction in the clinical encounter and better outcomes, and increase engagement and retention in care.
2. **Information exchange** is based on a dialogue between patient and provider or between community and health service system so that the provider learns about the social context (i.e., where patients live, what they do, what matters to them, their social position, and their environment and cultural perspective of the patient). There is evidence that information exchange promotes a more egalitarian view of the information needed for a patient-centered encounter and to structure treatment recommendations or proposed behavioral changes acceptable in the patient's daily lives. If service use and provider action depends on a social influence process as well as individual predispositions as has been documented since the 1970s, then understanding and using effective diffusion methods is likely to be more effective and efficient.

3. **Policies** include decisions by organizations such as routinely providing assistance to patients in communicating with their providers, or regulations by governments or public authorities such as deciding eligibility criteria for Medicaid services. An example of a policy change could be a new requirement that interpreter services be readily available in the clinical encounter if English is not the primary language of the patient. Also, mapping shortage areas where few mental health providers accept public insurance programs to provide additional incentives (i.e., higher reimbursement rates) for providers to offer care in those areas could be a policy to reduce disparities.

4. **Better treatment matching** includes the provision of linguistically appropriate services as well as treatment and outreach options that are socioculturally compatible with the circumstances of the target population. There are three basic criteria on which to judge whether there is treatment matching of services: the treatment/services are based on cultural values of the groups of interest; the strategies involved in the treatment/services are congruent with the culture and context of the particular ethnic/racial group; and the components of the service/treatment strategy are based on the expectations and behavioral preferences of the ethnic/racial groups. It is argued that the treatment matching of the services/treatments with the patient/community is necessary for quality improvements to be institutionalized and sustained by the service organization and acceptable to the community.

5. **Restructuring of services** involves organizational change to better accommodate the needs of clients. Settings that are accessible to communities and which they trust (i.e., community-based organizations and churches) might be more acceptable as settings for health or mental health service delivery or outreach. Changing stereotypes, prejudices, and uncertainty in provider–patient relations includes creating interventions to change preconceived ideas of either the patient or the provider that may adversely affect the communication or interaction in the clinical encounter.

As research that addresses the complexity of the multiple levels and two major domains that produce health and health care disparities evolves, efforts to eliminate disparities can be more tailored and effective. The goal of the SCF is to help stimulate the research and policy thinking that can foster change for health care systems and communities.

References

Agency for Healthcare Research and Quality. *National Healthcare Disparities Report*. Rockville, MD: US Department of Health and Human Services; 2013.

Alegría M, Canino G, Rios R, et al. Inequalities in use of specialty mental health services among Latinos, African Americans, and non-Latino whites. *Psychiatr Serv*. 2002;53(12):1547–1555.

Alegría M, Carson N, Flores M, et al. Activation, self-management, engagement, and retention in behavioral healthcare: a randomized clinical trial of the DECIDE intervention. *JAMA*

Psychiatry. 2014;71(5):557–565.

Alegría M, Lin J, Chen CN, Duan N, Cook B, Meng XL. The impact of insurance coverage in diminishing racial and ethnic disparities in behavioral health services. *Health Serv Res*. 2012;47(3pt2):1322–1344.

Alegría M, Molina KM, Chen CN. Neighborhood characteristics and differential risk for depressive and anxiety disorders across racial/ethnic groups in the United States. *Depress Anxiety*. 2014;31(1):27–37.

Alegría M, Pérez D, Williams S. The role of public policies in reducing disparities in mental health status for people of color. *Health Aff (Millwood)*. 2003;22(3):51–64.

Balsa AI, Cao Z, McGuire T. Does managed health care reduce unfair differences in health care use between minorities and Whites? *J Health Econ*. 2007;26(1):101–121.

Cook BL, Doksum T, Chen CN, Carle A, Alegría M. The role of provider supply and organization in reducing racial/ethnic disparities in mental health care in the US. *Soc Sci Med*. 2013;84:102–109.

Dana RH. Mental health services for African Americans: a cultural/racial perspective. *Cultur Divers Ethnic Minor Psychol*. 2002;8(1):3–18.

Gamble VN, Stone D. U.S. policy on health inequities: the interplay of politics and research. *J Health Polit Policy Law*. 2006;31(1):93–126.

Gee GC, Payne-Sturges DC. Environmental health disparities: a framework integrating psychosocial and environmental concepts. *Environ Health Perspect*. 2004;112(17):1645–1653.

Institute of Medicine. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Washington, DC: The National Academies Press; 2002.

Lie DA, Lee-Rey E, Gomez A, Bereknyei S, Braddock CH 3rd. Does cultural competency training of health professionals improve patient outcomes? A systematic review and proposed algorithm for future research. *J Gen Intern Med*. 2011;26(3):317–325.

Lipscomb HJ, Loomis D, McDonald MA, Argue RA, Wing S. A conceptual model of work and health disparities in the United States. *Int J Health Serv*. 2006;36(1):25–50.

Marmot M. Social determinants of health inequalities. *Lancet*. 2005;365(9464):1099–1104.

McFadden B, Nerenz DR, Ulmer C, eds. *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Washington, DC: National Academies Press; 2009.

McLeod J, Fettes D, Jensen P, Pescosolido B, Martin J. Public knowledge, beliefs, and treatment preferences concerning Attention-Deficit Hyperactivity Disorder. *Psychiatr Serv*. 2007;58(5):626–631.

Myers HF. Ethnicity- and socio-economic status-related stresses in context: an integrative review and conceptual model. *J Behav Med*. 2009;32(1):9–19.

Ngo-Metzger Q, Sorkin DH, Phillips RS, et al. Providing high-quality care for limited English proficient patients: the importance of language concordance and interpreter use. *J Gen Intern Med*. 2007;22(2):324–330.

Pollack CE, Armstrong K. Accountable care organizations and health care disparities. *JAMA*. 2011;305(16):1706–1707.

Roberts RE, Alegría M, Ramsay RC, Chen IG. Mental health problems of adolescents as reported by their caregivers: a comparison of European, African and Latino Americans. *J Behav Health Serv Res*. 2005;32(1):1–13.

Rogers W. Evidence based medicine and justice: a framework for looking at the impact of EBM upon vulnerable or disadvantaged groups. *J Med Ethics*. 2004;30(2):141–145.

Schulz A, Northridge ME. Social determinants of health: implications for environmental health promotion. *Health Educ Behav*. 2004;31(4):455–471.

Stiffman AR, Pescosolido B, Cabassa LJ. Building a model to understand youth service access: the gateway provider model. *Ment Health Serv Res*. 2004;6:189–198.

Thompson TG. *Culture, Race, and Ethnicity: A Supplement to Mental Health: A Report of the Surgeon General*. Washington, DC: US Department of Health and Human Services; 2001.

U.S. Department of Health and Human Services. *HHS Action Plan to Reduce Racial and Ethnic Health Disparities: A Nation Free of Disparities in Health and Health Care*. Washington, DC: US Department of Health and Human Services; 2011.

van Ryn M, Fu SS. Paved with good intentions: do public health and human service providers contribute to racial/ethnic disparities in health? *Am J Public Health*. 2003;93(2):248–255.

White K, Borrell LN. Racial/ethnic residential segregation: framing the context of health risk and health disparities. *Health Place*. 2011;17(2):438–448.

Wilkinson RG. Marmot MG. *Social Determinants of Health: The Solid Facts*. Europe: World Health Organization; 2003.

Williams DR, Mohammed SA. Discrimination and racial disparities in health: evidence and needed research. *J Behav Med*. 2009;32(1):20–47.

▲ 58.8 Criminalization of Persons with Serious Mental Illness

H. RICHARD LAMB, M.D., AND LINDA EDNA WEINBERGER, PH.D.

The enormously increased presence of persons with serious mental illness in the criminal justice system is one of the great problems of the day. As a result, mental health professionals and society have become much more concerned about the number of persons with mental illness in jails and prisons, as well as the treatment provided to these persons both in such facilities and after release. These issues are relatively recent. Reports of large numbers of persons with mental illness in US jails and prisons began appearing in the 1970s, a phenomenon that had not been reported since the 19th century.

For the purposes of this section, the focus will be on persons with serious mental disorders; that is, schizophrenia, schizoaffective disorder, bipolar disorder, and major depressive disorder. This section discusses the definition of criminalization of such persons and how criminalization came about, the extent of criminalization, the role of the police with regard to persons with mental illness, and the characteristics that are often found in persons with mental illness who have been arrested. The difficulties in treating this population both before incarceration and after release are addressed, including the relationship between serious mental illness and violence and the problems presented by persons who are not able to

recognize that they have a major mental illness. The section presents a summation of ways that have been proposed to reduce criminalization. These ways include the following: increased community mental health services that are effective and appropriate to treat persons who are or are at risk of becoming criminalized; better coordination among the police, courts, and mental health professionals; mobile crisis teams; increased mental health training for police officers; and diversion programs for offenders with serious mental disorders to prevent them from entering the criminal justice system, as well as transferring those who are already arrested into the mental health system. The necessary components of treatment for such persons will be discussed. These include having clear goals of treatment and practice; establishing a close liaison between treatment staff and the justice system; understanding the need for structure; having a focus on managing violence; and appreciating the crucial role of case management, appropriate living arrangements, and involving family members. Finally, the section ends with a brief discussion of recently enacted criminal justice efforts to decrease the total number of persons incarcerated, and recommends possible strategies to reducing the criminalization of persons with serious mental illness.

MAGNITUDE OF THE PROBLEM

The US prison population, including both federal and state prisons and county and city jails, was 2,145,100 inmates as of December 31, 2015. The percentage of jail and prison inmates assumed to be seriously mentally ill (schizophrenia, schizoaffective disorder, and major depressive disorder) has generally been estimated at about 16 percent. Using these numbers yields an estimate of 343,216 incarcerated persons with serious mental illness in jails, and state and federal prisons. The actual number may be somewhat higher or lower, depending on the accuracy of the percentage. These figures represent a substantial number of persons with serious mental illness in jails and prisons.

Moreover, research has shown that approximately one-fourth of persons with serious mental illness in the community have a history of at least one arrest. The association between serious mental illness and arrest is a widespread phenomenon that has affected both the mental health and criminal justice systems.

Ms. G was a 28-year-old single woman who was diagnosed with schizophrenia at age 22. Prior to her diagnosis she had a chaotic childhood and adolescence. Her mother abused alcohol and often left her alone in their apartment for days at a time. Ms. G was placed in multiple foster home placements where she frequently eloped, looking for her mother. Attempts at reunification were always unsuccessful due to the mother's continuing neglect and Ms. G's increasing disruptive behavior. At age 18, Ms. G was no longer a ward of the court and she began living with her 23-year-old boyfriend. She tried to work at fast-food restaurants but found the demands of the job increasingly stressful. She also started to experience paranoid ideas about her coworkers with whom she often fought and heard voices telling her that they

wanted to sexually molest her. On one occasion, when she was 20, she got into a fight with her manager and threw his glasses so hard against the wall that they broke. The police were called and Ms. G was arrested.

During her incarceration, the jail personnel did not recognize that she had a mental illness. She was given probation with no requirement that she obtain psychiatric treatment. She continued to live with her boyfriend until he said she could no longer live with him because of her constant hallucinations and delusions about him. She was now homeless and without money. One day, she walked into a grocery store and stole a loaf of bread and a package of sliced turkey. She was apprehended by a security guard and the police arrested her. During this incarceration she was identified as having a mental illness and was treated with antipsychotic medication in jail. Ms. G received probation but this time was to attend psychiatric outpatient treatment. She went to the community mental health clinic once but did not attend again because she did not believe she had a mental illness and refused to take any medications.

Shortly thereafter, Ms. G got into a fight with another homeless person and stole his blanket in front of several bystanders who called the police. This led to another arrest where she spent a few months in jail and was again given probation with a condition of outpatient psychiatric treatment. Following her release she attended outpatient treatment at a mental health clinic for several weeks. Initially, she took her medications which resulted in a dramatic improvement of her symptoms. However, several weeks later Ms. G stopped taking her medications, wherein her symptoms returned and she became very threatening to the staff. The court was informed of this and she returned to jail. Prior to her release from jail, Ms. G was encouraged to return to her previous outpatient clinic for treatment. Sometime later, Ms. G, appeared at the clinic but the intake team concluded that her history of nonadherence to psychiatric medications along with her past violent behavior could make her a problem in management beyond the capabilities of the clinic.

A few months later, Ms. G was arrested outside a department store for having stolen a sweater and pushing a customer out of her way. At this time, she was charged with petty theft with priors and assault, which was a felony. The judge believed he had no option but to send Ms. G to state prison. Criminalization was now confirmed.

Most outpatient mental health treatment facilities are not structured to successfully treat persons like Ms. G and, as described in this section, community mental health staff is often reluctant to try.

DEFINING CRIMINALIZATION

Many factors come into play in determining why a person with mental illness is arrested rather than taken to a psychiatric hospital.

Generally, persons who are thought to have committed a felony are arrested and brought to jail regardless of their mental condition. The criminal justice system, charged by society with the responsibility for removing from the community persons accused of committing serious crimes, sees no alternative but to first place the person in custody in a secure setting and then arrange for psychiatric treatment, if necessary. If the person is thought to have committed a serious crime, the police and the criminal justice system generally do not want to place this person in a psychiatric hospital, because such persons may well pose a number of problematic issues that cannot be treated and managed adequately by the mental health system and thus place society at risk.

Examples of this may be a person with schizophrenia who was found by the police in possession of a loaded gun; he was brandishing the weapon and threatening others because he believed that they were poisoning him. Another person with untreated bipolar disorder was arrested after he drove

his car erratically on the street and onto a sidewalk full of pedestrians because he wanted to be recognized for his race-car-driving skills.

For persons charged with misdemeanors, the situation is more complex. In 1972, the term "criminalization of the mentally ill" was first used. It was observed that persons with mental disorders who engaged in minor crimes were increasingly subject to arrest and prosecution. Later, many authors applied the concept of criminalization to persons with mental disorders who were arrested for serious crimes. The distinction between arrest and incarceration of persons with mental illness who have committed minor offenses and those who have committed serious offenses is an important one.

This section takes the position that the term criminalization should be used primarily in connection to persons with serious mental illness who are arrested and incarcerated for minor offenses instead of being placed in the mental health system. However, it should be acknowledged that many persons with serious mental illness who commit serious crimes and enter the criminal justice system might not have engaged in such behavior if they had been receiving adequate and appropriate mental health treatment at that time.

Concern about the inappropriate placement of persons with serious mental illness in the criminal justice system is not new. As far back as 1939, researchers in Europe advanced the thesis that a relatively stable number of persons are confined in any industrial society. Using prison and mental hospital census data from 18 European countries, they found an inverse relationship between prison and mental hospital populations. They theorized that if one of these forms of confinement is reduced, the other will increase. According to this theory, when prison populations are extensive, mental hospital populations will be small, and vice versa. Thus, if there is room in prisons and a shortage of hospital beds, many persons with mental illness who come to the attention of law enforcement might well be directed to the criminal justice system. Recent research in South America is compatible with this hypothesis. As will be seen in this section, this has happened in the United States and other countries; the census of state hospitals has dropped precipitously while the number of persons incarcerated in the criminal justice system has risen to record proportions.

One of the major concerns in present-day psychiatry is that placement in the criminal justice system poses a number of important problems and obstacles for the treatment and rehabilitation of persons with serious mental illness. Even when quality psychiatric care is provided in jails and prisons, the inmate/patient still has been doubly stigmatized as both a person with mental illness and a criminal. Furthermore, jails and prisons have been established to mete out punishment and to protect society; their primary mission and goals are not to provide treatment. The correctional facility's overriding need to maintain order and security, as well as its

mandate to implement society's priorities of punishment and social control, greatly restrict the facility's ability to establish a therapeutic milieu and provide all the necessary interventions to treat mental illness successfully.

CAUSES OF CRIMINALIZATION

A number of factors have been cited as causes for persons with mental illness being placed in the criminal justice system. The most common ones are deinstitutionalization and the unavailability of intermediate and long-term hospitalization in state hospitals for persons with serious mental illness; more formal and rigid criteria for civil commitment; the lack of adequate support systems for persons with serious mental illness in the community, as well as the difficulty such persons coming from the criminal justice system have in gaining access to community mental health treatment; persons with co-occurring disorders of a serious mental illness and a substance abuse problem being caught up in the war on drugs (being arrested and incarcerated on the basis of the use, sale, and possession of drugs); and a belief by many law enforcement personnel that deviant behavior can be dealt with more quickly and efficiently within the criminal justice system than in the mental health system.

Deinstitutionalization

Beginning in the late 1950s, the inclination to provide hospital beds declined precipitously. This was because of the introduction of psychoactive medications; the development of more efficacious community treatment interventions; the creation of federal programs (e.g., SSI, SSDI, Medicaid, and Medicare), which would fund community treatment and housing for persons with mental illness; the influence of the civil rights movement; and the high cost of institutionalizing persons with mental illness. For example, in 1955, when the number of patients in state hospitals in the United States reached its highest point, 559,000 persons were institutionalized in state mental hospitals out of a total national population of 165 million (339 beds per 100,000 population). By 2016 (as a result of hospital closures and bed eliminations), the number of persons in state mental hospitals dropped to 37,679 for a population of approximately 324,000,000, or 11.7 beds per 100,000 population. This rate is similar to that found in 1850 when persons with serious mental illness received little care and concern. Their plight led to the beginning of the movement, spearheaded by Dorothea Dix and others, to provide more humane care by treating persons with serious mental illness in hospitals. In some states today, such as California, the ratio is even less; that is, there are fewer than 2 state hospital beds per 100,000 population, excluding forensic patients.

Deinstitutionalization has thus been viewed as setting the stage for

increasing numbers of persons with mental illness entering the criminal justice system. It must be remembered that the community mental health system was developed in the 1960s and 1970s for persons with mental illness who had more moderate needs. While deinstitutionalization held the promise of persons with serious mental illness being able to live successfully in the community, that hope was not achieved for a sizeable number of individuals. Part of its failure was attributed to a lack of planning before or during deinstitutionalization as well as a lack of adequate funding. As a result, communities were not able to provide a sufficient community care system (i.e., housing, medical and psychiatric care, social services, and social and vocational rehabilitation) for the formerly hospitalized patients. Despite this, the majority of deinstitutionalized patients were able to adapt successfully in the community; however, this was not the case for a considerable minority who were arrested and placed in jails and prisons or who became homeless (between one-fourth and one-third of homeless persons have a serious mental illness). Some of these individuals presented challenges in treatment—such as not seeing themselves as mentally ill, not taking their medications, abusing substances, and in many cases, becoming violent when stressed. Many of these persons needed highly structured care to replace that which had been provided to them, albeit imperfectly, in psychiatric hospitals. The flawed implementation of deinstitutionalization would thus appear to be a significant factor accounting for the migration to jails and prisons of many persons with mental illness who otherwise could have been treated with more adequate services in the community or in a hospital. The criminal justice system had become by default a safety net for the mental health system.

Restrictive Civil Commitment Criteria

Many believe that more formal and rigid civil commitment criteria have contributed not only to deinstitutionalization but also to an increased number of persons with mental illness in jails and prisons. In 1969, California's then-novel civil commitment law, the Lanterman-Petris-Short Act, went into effect. Within a decade every state made similar modifications in their commitment codes. Such a rapid and complete consensus among legislatures is virtually unprecedented. More important, it reflected a nearly universal view that past inattention to the rights of persons with mental illness needed to be corrected.

In effect, the new civil commitment laws resulted in three major changes. First, the laws modified the substantive criteria for commitment from more general criteria that simply embodied concepts of mental illness and need for treatment to more specific criteria that embodied either dangerousness to self or others resulting from mental illness or the incapacity to care for oneself. Second, the laws changed the duration of

commitment from indefinite and extensive periods to specific and brief periods. Third, the new laws explicitly provided that persons civilly committed have rapid access to the courts, to attorneys, and, in some cases, to jury trials; this access ensured the kinds of due-process guarantees to civilly committed persons that criminal defendants had obtained over the previous decade.

These procedural safeguards and clear commitment standards resulted in fewer, as well as shorter, commitments. Consequently, many individuals with mental illness who otherwise would have been civilly committed by family or others were now left to reside in the community where they might not have been able to care for themselves and control their impulses. Thus, some of these individuals might well have committed criminal acts and entered the criminal justice system.

Lack of Adequate Community Support Systems

Another factor that both leads to and perpetuates the criminalization of persons with serious mental illness is the lack of adequate support systems in the community. An essential part of support systems for these persons is the availability of community treatment resources. Moreover, the type of community mental health resources available may not be suitable for all of the population to be served.

The availability, or lack of availability, of treatment resources in the community has three important aspects. First, in most jurisdictions, mental health treatment, housing, and rehabilitation resources are insufficient to serve the very large numbers of persons with serious mental illness in the community. For instance, case management has come to be viewed as one of the essential components of an adequate mental health program. However, the mental health system is ill prepared to provide quality case management services to all persons with serious mental illness who require it, including those leaving jails and prisons.

Second, community mental health resources may be inappropriate for some of the population to be served. For instance, persons with mental illness may be expected to come to outpatient clinics when the real need for a large proportion of this population is outreach services, in which mental health professionals go to see patients where they live. In addition, some service providers may lack the ability to provide the degree of structure required by many persons with serious mental illness. The issue of structure will be discussed later.

Third, persons with mental illness who have been in jail may not be able to gain access to community treatment even when it is available. Many of these persons have been described as resistant to treatment, dangerous, and seriously substance abusing, characteristics generally not considered desirable by most community mental health agencies and staff. Furthermore, because many of these agencies may not have the capability

to provide the needed structure, limit setting, and security and safety for staff necessary to treat these persons successfully, their reluctance to treat them may be appropriate.

A large proportion of persons with mental illness who commit criminal offenses are found to be highly resistant to psychiatric treatment. They may refuse referral, may not keep appointments, may not be adherent with psychiatric medications, may not abstain from substance abuse, and may refuse appropriate housing placements. There is evidence that many of these persons suffer from a disorder called anosognosia (a biologically based inability to recognize that one has a mental illness, and thus a biologically based lack of insight). This is discussed in greater detail later.

Thus, the mental health system finds offenders with mental illness extremely difficult to treat and may be hesitant to serve them. This reluctance may extend to virtually all areas of community-based care, including psychiatric treatment, therapeutic housing, social and vocational rehabilitation, and general social services. Moreover, many offenders with mental illness can be intimidating because of previous violent and fear-inspiring behavior. Treating this group is very different from helping passive, formerly institutionalized patients to adapt quietly to life in the community. Community mental health professionals are not only reluctant, but may also be afraid to treat them, especially when measures are not adopted to ensure staff safety. As a consequence, these persons are left for the criminal justice system to manage. However, there are outpatient facilities in which structure is provided, security personnel are present and thus staff are protected, and mental health and criminal justice staff collaborate closely; under such circumstances, many of these persons are successfully treated.

The Role of the Police

The police play an important role in the criminalization of persons with serious mental illness, and this will be discussed in the context of the more in-depth exposition of the police and mental health that follows.

Police and Mental Health. Since the advent of deinstitutionalization and the exodus of persons with mental illness into the community, law enforcement agencies have had an increasingly important function in the management of persons who are experiencing psychiatric crises. The rationale for the police to intervene in the lives of persons with mental illness derives from two common-law principles: the power and the authority of the police to protect the safety and welfare of the community, and the state's paternalistic or *parens patriae* authority, which dictates protection for citizens with disabilities who cannot care for themselves, such as those who are acutely mentally ill.

The police are typically the first responders in the community to urgent

situations involving persons with mental illness. They provide services 24 hours a day, 7 days a week. Consequently, they play a major role as a mental health resource. They are responsible for either recognizing the need for the treatment of an individual with mental illness and connecting the person with the proper treatment resources or making the determination that the individual has committed a type of illegal act for which the person should be arrested. This responsibility thrusts them into the position of primary gatekeepers who determine whether the individual will enter the mental health or the criminal justice system.

With respect to persons with mental illness, police in all states have the power to transport persons for psychiatric evaluation and treatment when there is probable cause to think that they are a danger to themselves or to others because of their mental condition. As a result, law enforcement officers have assumed the role of "street-corner psychiatrist" by default. It would appear that many officers have grown accustomed to this role and consider it one of their duties; however, other officers do so reluctantly, and some do it with resentment. A major problem with having to fulfill this role is that the police have little training in performing this kind of triage. This lack of training is one of the factors that has played an important part in the criminalization of persons with mental illness. The training of law enforcement officers is discussed later in the subsection entitled *Diverting Persons with Mental Illness from the Criminal Justice System*.

The police have a great deal of discretion in the exercise of their duties, including determining what to do when they encounter a person with acute mental illness in the community. In most cases, the police use informal tactics, such as trying to "calm" the person or taking the person home. This practice has been referred to as "psychiatric first aid." In situations that cannot be handled informally, the police may have to take persons with mental illness to hospitals or to jails.

In some cases, however, public policy limits the police officer's discretionary power. For instance, if the person with mental illness is alleged to have committed a major crime, the disposition is clear—that person is taken to jail because of the seriousness of the offense.

Often, the interaction between the police and the person with mental illness is initiated by citizens. In such cases, the citizens' demands also may come into play and limit the discretion of the police. For example, many retail stores have a policy that anyone caught shoplifting should go to jail, and store managers are instructed to make a citizen's arrest and call the police without exception. In another kind of situation, people who have just been assaulted by a person with psychosis frequently are not inclined to be sympathetic to their assailant, even when mental disturbance is evident. The result may be an angry citizen who insists on having the person arrested and taken to jail.

Police and Criminalization. As discussed previously, the police play an important role in the criminalization of persons with serious mental illness. Even when the police consider the behavior of a person who has come to their attention to be the result of a mental illness, they may encounter a number of problems and irritants when trying to place the person in an inpatient psychiatric facility. For example, there may be long waiting periods in psychiatric emergency rooms during which police officers cannot attend to other duties. Mental health professionals may question the judgment of the police and refuse admission, or they may admit for only a brief hospital stay a person who just a short time before constituted a clear menace to the community.

However, the police are well aware that if they arrest a person with mental illness, that person will be dealt with in a more systematic and predictable way under the criminal justice system. The individual will be taken into custody; it can be hoped that the person will be seen by a mental health professional attached to the court or in the jail and will probably receive psychiatric evaluation and treatment. Thus, arrest is a response with which police are familiar, one over which they have more control, and one that they believe will lead to an appropriate disposition.

Although a person with mental illness suspected of committing a serious crime will most certainly be arrested and transported to jail, a number of factors in addition to those just mentioned have been proposed to explain why a person with serious mental illness suspected of committing a minor offense is arrested rather than taken to a hospital. A person who appears mentally ill to a mental health professional may not appear so to police officers, who, despite their practical experience, have not had sufficient training in dealing with this population and are still laypersons in these matters. In addition, mental illness may appear to the police as simply alcohol or drug intoxication, especially if the person with mental illness has been using drugs or alcohol at the time of the interaction with the police. Moreover, in the heat and confusion of an encounter with the police and other citizens, which may include forcibly subduing the person with mental illness, signs of mental illness may go unnoticed. Furthermore, studies have shown that the occurrence of violence at the time of a police encounter increases the chances that the person with mental illness will be taken to jail.

Another instance is when law enforcement officers may charge persons with mental illness with a misdemeanor and take them to jail if they think that no other appropriate alternatives are available, a practice that has been referred to as “mercy booking.” Thus, the scarcity of psychiatric inpatient beds and/or limited community mental health services may be a factor contributing to the criminalization of persons with serious mental illness.

MENTAL ILLNESS AND VIOLENCE

In countries that experienced mass deinstitutionalization since the late 1960s, a number of studies have shown that persons with serious mental illness were more likely to commit criminal offenses, and even more likely to commit violent offenses, than the general population. This growing body of evidence demonstrates a relationship between mental illness and violence, primarily among persons who do not take their medications and are psychotic. Substance abuse has also been shown to increase the risk of violent behavior, particularly in combination with serious mental illness. Thus, nonadherence to medication, acute psychosis, and substance abuse account for most of the increased violence of persons with serious mental illness. For instance, one study examined persons with serious mental illness in the United States who had been arrested over a 10-year period, and found that more than one-third were arrested for violent crimes (assault and murder), and more than one-fifth were arrested for drug offenses.

The relationship between serious mental illness and violence is most easily seen in relatively nonviolent societies, such as in Scandinavia. The Scandinavian countries maintain centralized case registries that record comprehensive data on everyone born in the country or who immigrates there. Included in this wealth of data are psychiatric registers and national police registers. In Denmark, the psychiatric registry includes data such as admissions to psychiatric wards, dates of admission and discharge, and discharge diagnoses. There is also a national police register that records all arrests and convictions, as well as information on all criminal proceedings. These registers allow for the computerization and analysis of the data.

For instance, it was found that males in Denmark with a serious mental disorder who were admitted to a psychiatric hospital by age 44 years represented only 5 percent of the total population of males but were responsible for about 30 percent of all the violent offenses committed by males. Likewise, female mental patients in Denmark constituted about 5 percent of the female population but were responsible for 50 percent of all the violent offenses committed by females.

In another Danish study, the risk of violent offending was found to be 4.6 times higher for male subjects with schizophrenia and 2 times higher for those with affective psychoses than for male subjects who were never admitted to a psychiatric ward. For female subjects with schizophrenia, the risk was 23.2 times higher, and for those with affective psychoses, the risk was 3.9 times higher, than for female subjects who were never admitted to a psychiatric ward.

Similar findings were noted in Sweden in a study that examined all 15,117 individuals who were born in Stockholm in the same given year. They were evaluated through case registries 30 years later to determine how many were diagnosed with a serious mental illness and were convicted of a violent crime. It was found that 15 percent of the men and 6 percent of

the women who had a major mental disorder were convicted of a violent offense. A recent study of persons with mental illness in Sweden who were prescribed antipsychotics or mood stabilizers, found that violent crime dropped by 45 percent for those receiving antipsychotics and 24 percent for those receiving mood stabilizers.

It has been asserted that violent behavior by a subgroup of persons with serious mental illness increases community stigmatization and rejection of persons with mental illness generally. These violent episodes, often highly publicized, may result in the association of mental illness with violence in the public's mind. Such stigmatization can result in reluctance on the part of the public to support funding for adequate mental health services for persons with mental illness. Moreover, there may be less tolerance for integrating persons with mental illness into community housing and finding them appropriate employment. It will be difficult to reduce the stigma of mental illness and its consequences until the violence of this subgroup is addressed. Furthermore, to safeguard the welfare of persons with serious mental illness, as well as the safety of the public, mental health professionals and society must not ignore the evidence that mental illness is associated with some increased risk of violence.

Although there is a subgroup of persons with serious mental illness who are violent, it is extremely important to take into consideration data that demonstrate that an even larger group of persons with serious mental illness are the victims of violence. For example, in a review of studies of violent victimization, 35 percent of persons with serious mental illness had been victims of violence within the last year. For persons with serious mental illness, their victimization by violence tends to receive less public attention than the violence that is perpetrated by them.

Anosognosia

Anosognosia is a biologically based inability to recognize that one has a mental illness, and thus is a biologically based lack of insight. Researchers have identified various parts of the brain (e.g., ventricles, frontal and temporal lobe subregions, medial temporal and inferior parietal regions, and subcortical structures) where damage is likely to result in anosognosia. Impaired insight may be mediated by deficiencies in conceptual organization and the lack of flexibility in abstract thinking. Generally, lack of insight that one has a serious mental illness is one of the risk factors for violence in persons with schizophrenia, thus demonstrating a relationship between poor insight and violent behavior.

The importance of anosognosia is demonstrated by a study investigating 907 adults who met the criteria for psychotic and affective disorders and received public mental health services. This study found that 26 percent of these persons with serious mental illness reported being arrested or violent in the last year. These individuals were significantly more likely to deny

needing psychiatric treatment.

Overall, these findings suggest that many of the people who are most in need of psychiatric treatment are the least likely to believe that they need it. Treatment nonadherence is a key barrier to implementing evidence-based interventions. Several studies confirm that treatment involvement is related to lowered risk of arrest and violence in persons with serious mental illness by reducing their psychiatric symptoms and substance abuse. Interventions to improve insight and treatment adherence in this population are warranted and needed.

SOME CHARACTERISTICS OF PERSONS WITH SERIOUS MENTAL ILLNESS WHO ARE INCARCERATED

It is known that there are very large numbers of persons with serious mental illness in jails and prisons. What is known about them in terms of their criminal histories and legal status, the psychiatric services that they use while incarcerated, and the challenges that they might present in psychiatric treatment after release?

Two studies attempted to answer these questions. The first was a retrospective study of inmates with serious mental illness who were arrested and placed in a large, urban county jail. It should be noted at the outset that these findings do not necessarily represent what one would find in a similar study conducted in state and federal prisons, or even in other jails. This study revealed that 76 percent of these inmates required and received psychiatric inpatient care or its equivalent for part of their time in jail during the current offense. Clearly, a large number of people with serious mental illness are receiving their acute psychiatric inpatient treatment in the criminal justice system rather than the mental health system.

With respect to these inmates' history before the current arrest, at least 92 percent were known to be nonadherent to psychiatric medications, 94 percent had prior arrests, 72 percent had prior arrests for violent crimes, and 76 percent were known to have a history of substance abuse. Given these data in addition to the fact that three-fourths of these inmates required inpatient psychiatric care in the jail, it would appear that the jail had acquired the responsibility of handling many persons with serious mental illness who were among the most difficult and expensive to treat.

Another study conducted on individuals who had a serious mental illness and were transferred from jail to a psychiatric inpatient unit found that 75 percent had a history of physical assault in the community, 72 percent engaged in substance abuse, and 78 percent were generally nonadherent to medications. Findings revealed that poor insight into illness and nonadherence to medications were associated independently with violence toward others.

That such large percentages of persons in both studies had a history of

being nonadherent to psychiatric medications suggests that treatment on re-entry into the community should include such measures as evaluation and supervision by mental health courts, assertive community treatment (ACT), and assisted outpatient treatment (AOT) (see later discussion). Moreover, short-term, intermediate-term, or long-term hospitalization may be needed for many.

It was observed in these studies, as it frequently is in jails generally, that a number of persons with serious mental illness are arrested when it appears that their offending conduct was due primarily to their illness. The nature and circumstances of the offense suggest that these persons should have been psychiatrically hospitalized instead of being taken to jail. The following two cases serve to illustrate this point.

A person with psychotic symptoms and a previous diagnosis of schizophrenia was walking in circles, talking to himself, and then urinating outside a grocery store; he was arrested and charged with indecent exposure. Another person with manic and paranoid symptoms was panhandling at a gas station and yelling that a customer was stealing his money, whereupon he pushed this customer and others. The attendant at the gas station called the police who arrested him for assault. These individuals would probably have been better served if they had been taken to a psychiatric hospital and placed under the jurisdiction of the mental health system rather than being arrested and placed in jail. Incarcerating such persons in settings where it is difficult to provide good treatment, such as jail, does not appear to meet the needs of society by protecting society from dangerous people. Moreover, efforts to reduce the possibility of recidivism by these persons can generally be accomplished by treatment and management within the mental health system.

Using Cognitive Behavioral Therapy to Address Antisocial Characteristics

Some professionals have attributed criminal acts by many persons with serious mental illness primarily to the individual's criminal characteristics. These professionals believe that while there are incarcerated individuals who may have serious mental illness, many of them are in correctional facilities largely because of their antisocial attitudes, and a lack of problem solving and self-control skills that contribute to their criminal behavior. Consequently, a primary intervention suggested to reduce recidivism is to focus on the antisocial cognitions of these persons.

Indeed, a number of persons with serious mental illness in the criminal justice system may have antisocial tendencies, often to the extent of justifying a diagnosis of antisocial personality disorder. While cognitive behavioral therapy (CBT) may be a necessary component of the treatment regimen for many such offenders in reducing their recidivism, if such persons' thinking and judgment are impaired by their untreated psychotic and manic symptoms they may not be able to profit from CBT

interventions. Therefore, if CBT is to be effective, it should be combined with psychiatric treatment.

Diverting Persons with Mental Illness from the Criminal Justice System

With the very large and increasing number of persons with serious mental illness found in jails and prisons, there have been large-scale efforts to divert these persons from the criminal justice system to the mental health system. Diversion before the person is actually booked into jail, or pre-booking diversion, is exemplified by large-scale efforts to create community mobile crisis teams of police officers and/or mental health professionals. In addition, coordination between police and mental health professionals, as well as mental health training for law enforcement officers, may help in reducing arrests.

Diversion after booking takes a variety of forms. These include specialized mental health courts that handle exclusively offenders with mental illness. Mental health consultation to other courts may also assist the judge by offering recommendations for treatment in lieu of incarceration.

Mobile Crisis Teams. Psychiatric emergencies have been managed effectively in communities in which close formal liaisons between law enforcement and mental health agencies have been established. These arrangements have helped to facilitate the resolution of crises of persons with mental illness in the field without the need to resort to hospitalization or incarceration. When resolution in the field is not possible, the existence of these liaisons increases the number of persons with mental illness who are referred to the mental health system rather than jailed, thus reducing criminalization. For those who need psychiatric hospitalization, these relationships tend to increase the rate of acceptance by hospitals.

Different strategies have been developed to provide a mobile team of police, mental health professionals, or both to respond in the community to persons with mental illness who are in crisis. The following are two of these strategies that are used most commonly.

An increasing number of jurisdictions use sworn police officers who have special and extensive mental health training to provide crisis intervention services as part of crisis intervention teams (CIT programs) and to act as liaisons to the mental health system. This approach is often referred to as the “Memphis model” because it was developed in Memphis, Tennessee. These specially trained officers may deal with mental health emergency situations on-site or act as consultants to the officers at the scene. This model places a heavy reliance on psychiatric emergency services that have agreed to a no-refusal policy for persons brought to them by the police. Although this strategy has a close liaison with mental health, it does not require the actual participation of mental health professionals in

the field.

Another model that some jurisdictions use deploys teams composed of both specially trained sworn police officers and mental health professionals employed by the local community mental health department. These teams have been shown to be effective in resolving emergency situations in the community involving persons with mental illness and in diverting them to the mental health system rather than to the criminal justice system. An advantage of this model is that the mental health professionals on the team may have access to the person's documented mental health history. Having such data available is an important factor that can help the team better understand the person's mental condition and make more informed decisions.

The major goals of both of these specialized mobile crisis teams are to resolve the crisis and to reduce criminalization. Studies that have evaluated such teams found that they had arrest rates ranging from 2 to 13 percent (with an average of less than 7 percent), in contrast to an arrest rate of 21 percent for contacts between nonspecialized police officers and persons who were apparently mentally ill. This finding lends credence to the idea that a specialized response lowers the incidence of inappropriate arrests.

Regardless of whether police officers act as part of a team with members of the mental health system in dealing on-site with emergency situations, or police officers acting on their own in dealing with these situations, it is imperative that they obtain all available information about past police contacts with the individual when mental illness is thought to be present. Such information may include aspects of the person's actions during previous police contact, such as suicide attempts and violence, and how the previous contact was resolved.

A crucial part of developing and maintaining any specialized mobile response team is that a rigorous and ongoing evaluation of the program must be conducted. For instance, response times—the time from the call to the arrival of the team on the scene—should be measured on a routine basis. If response times are too long, police officers in the field tend not to call on the mobile teams. The teams' arrest rates, in addition to the mental health alternatives they choose, need to be evaluated regularly to be sure that appropriate dispositions are being made and that mental health facilities are cooperating with these efforts.

Increased Coordination and Need for Training. From the standpoint of the police, it is clear that officers need and want rapid on-site assistance from mental health professionals when they are called on to deal with difficult or complex situations involving persons with mental illness who are acutely psychotic, behaving bizarrely, or exhibiting violent behavior, or for persons who have attempted suicide or made a suicide gesture. Similarly, mental health professionals who are working as members of

psychiatric emergency teams without police support may feel ill equipped to handle such individuals in the field. Thus, it has become increasingly apparent in recent years that when persons with mental illness in the community are in crisis, neither the police nor the emergency mental health system alone can serve them effectively and that it is essential for the two systems to work closely together.

Training of Police. Most police departments provide some type of officer training in mental illnesses; however, training in how officers should respond to persons with mental illness who are in crisis is less common. Moreover, mental health training is needed for all police officers, not just for those who are part of the specialized mobile crisis teams.

Training for police officers about mental health issues and individuals with mental illness should begin in police academies as a regular part of the curriculum. This training should focus on enhancing the understanding of mental illness by law enforcement personnel. Furthermore, it should aim at increasing their knowledge about such mental health issues as how to recognize mental illness, how to handle violence or potential violence among persons with mental illness, what to do when a person is threatening suicide, and when to call the specialized mobile crisis team. In addition, information should be provided as to what mental health resources are available, both inpatient and outpatient, and when a person should be arrested and incarcerated as opposed to being taken to a mental health facility.

For the training to be successful and relevant, trainers should come from both the mental health and criminal justice systems. Trainers from the mental health system should be persons who have experience with crisis intervention in the field and/or in crisis centers, as well as knowledge and field experience relating to the criminal justice system. These mental health professionals must understand and appreciate the challenges and risks that law enforcement officers experience every day in the field. However, trainers from the criminal justice system already have credibility with their trainees and knowledge about such issues as officer safety and enforcement protocols that noncriminal justice system persons would not have.

Another important issue to consider when law enforcement officers come into contact with persons who have mental illness is how the individual perceives the officers. Persons with mental illness may have negative attitudes toward the police; they may be fearful of police, believing that the officers do not listen to them, and may use force against them. Such perceptions can lead to dangerous escalating interactions. Consequently, the attitudes and behaviors of both the officer and person with mental illness have an effect on the outcome of their contact. Understanding how persons with mental illness may perceive law

enforcement officers, particularly if they had a previous negative experience with police, is essential.

It needs to be emphasized that mental health training should be an ongoing process, not only for police officers in special mental health crisis teams but also for all police officers.

MENTAL HEALTH SERVICES AFTER ARREST

To reverse criminalization, it is extremely important to identify persons with mental illness as quickly as possible after they have been arrested and placed in jail. A facility may routinely screen all incoming detainees for serious mental disorder. Initial screenings may be conducted by custody personnel who notice the inmate engaging in odd or bizarre behavior. Medical/psychiatric questionnaires have also been used in many correctional institutions as a screening device. Such questionnaires consist typically of less than 20 questions aimed at identifying inmates who may need psychiatric care. Those individuals who endorse items on the questionnaire or who appear to have a mental illness may then be seen by a clinician for a more extensive evaluation to determine what type of housing and/or mental health treatment they should receive while incarcerated. Even more important, if the person is one who should have been treated in the mental health system in the first place, discharge planning should begin as soon as possible after arrest.

Given the large numbers of persons with serious mental illness who are arrested, jail administrators should be encouraged to negotiate programmatic relationships with mental health agencies to provide multidisciplinary teams of clinicians to work in the jail setting. These teams should include psychiatrists so that psychoactive medications can be prescribed. The teams should be established inside jails to provide short-term crisis intervention, treatment, referral to a psychiatric hospital if necessary, and discharge planning. Such information can be useful in helping the judicial system make decisions concerning the person's legal disposition.

Mental Health Courts. Post-booking diversion strategies are being used increasingly by special courts called mental health courts. The first mental health court that received national attention was established in Broward County, Florida, in 1997, and since then, the number of mental health courts in the United States has increased greatly. Initially, these courts were set up to hear cases of persons with mental illness who were typically charged with misdemeanors. In recent years, these courts have increased their purview to serve persons with mental illness charged with felonies. In mental health courts, all the courtroom personnel (i.e., judge, prosecutor, defense counsel, and other relevant professionals) have experience and training in mental health issues and available community resources. These

mental health courts have a particular set of characteristics: they hear specialized cases involving defendants with mental illness; they use a nonadversarial team of professionals (e.g., judge, attorneys, and mental health clinician); they are linked to the mental health system that will provide treatment; and they use some form of adherence monitoring that may involve sanctions by the court.

Underlying the concept of mental health courts is the principle of therapeutic jurisprudence, which emphasizes that the law should be used, whenever possible, to promote the mental and physical well-being of the people it affects. The concept of therapeutic jurisprudence operates on the belief that the application of the law can have therapeutic consequences. It should be emphasized that therapeutic jurisprudence does not diminish the importance of public safety, which is fully taken into account by the court.

In a system characterized by therapeutic jurisprudence, people with serious mental illnesses charged with crimes may be diverted into programs designed to address their treatment and service needs rather than simply being incarcerated with their treatment needs being neglected. Even individuals with serious mental illnesses convicted of serious crimes can be provided with humane and appropriate treatment while incarcerated. Generally, mental health courts facilitate linking offenders with serious mental illness to appropriate needed services and supports on discharge from jail in order to enable them to successfully re-enter their communities.

Mental health courts were developed as a strategy to divert persons with mental illness from the criminal justice system into the mental health system. When offenders with mental illness are booked into jail, their case is usually handled by mental health courts in one of two ways. In the first situation, the prosecution of their cases may be postponed by the judge so that they may be involved in a specified treatment program run by the mental health system. If they successfully complete the program, their criminal charges are dismissed, and the individuals are no longer under the jurisdiction of the criminal justice system. If they fail to complete the treatment program successfully, the charges are reinstated, and the criminal proceedings resume. In the second case, the offender with mental illness offers a plea of guilty or *nolo contendre,* which is accepted by the court. The person is sentenced (e.g., sentenced to 180 days in jail, which may be held in abeyance) and given probation, which includes treatment by the mental health system as a condition of probation. If the person is nonadherent with treatment or other probationary conditions, probation may be withdrawn and the sentence imposed. There is evidence that these types of jail diversion may reduce recidivism for persons with serious mental illness.

The mental health court system works in a collaborative effort to devise and implement a treatment plan that includes medications, therapy,

housing, and social and vocational rehabilitation, all in an effort to address the individual's mental illness and reduce the risk of recidivism. Research has suggested that a factor contributing to reducing recidivism and improving clinical outcomes is when the court and mental health personnel work as a "team" to coordinate and provide services. Clearly, the quality of services plays an influential role in the success of mental health courts. However, there may be unforeseen limitations of these programs with respect to available resources. That is, there are often insufficient community psychiatric treatment, rehabilitation, and housing capabilities in the existing mental health system to accommodate those persons with mental illness who are diverted from the criminal justice system. Another issue that has already been discussed is that many community resources are not appropriate, adequate, or willing to serve a large proportion of those who are diverted from the criminal justice system. Not surprisingly, there is research evidence that this lack of appropriate resources to which diverted persons can be referred results in recidivism.

Another problem resulting from the services recommended by these courts for offenders with mental illness is the possibility that they might be provided at the expense of programs for persons with mental illness who have not committed crimes. When this happens, an undesirable situation may be set up whereby committing a crime offers the patient with mental illness the best chance of getting psychiatric treatment.

These specialized mental health courts have been given high priority because they appear to be part of the answer to the problem of criminalization. However, for these diversionary programs to be truly effective and not overwhelm the already existing limited budgets of the departments of mental health, the treatment to which the courts have referred the offender with mental illness should be funded by additional monies. These monies could come, at least in part, from savings in the criminal justice system, which would no longer have to incarcerate and treat these persons.

The Importance of Mental Health Consultation to the Court. Consultative arrangements with mental health professionals may help to correct a problem that involves the computerized criminal histories, or "rap sheets," available to the court of persons with serious mental illness. Criminal justice personnel, such as prosecutors, probation officers, and parole officers, access the person's criminal history. It is important to consider what effect this may have on the court's decision making regarding the disposition for the individual when there is no mental health input. These computerized criminal histories seldom include a narrative as to what really happened during the previous arrests. Even when a person with serious mental illness commits nonserious offenses, a series of them can result in a lengthy computerized criminal history for the individual. Thus,

when such individuals subsequently appear in court, they may look like habitual criminals when their criminal histories are accessed.

By providing mental health evaluation and consultation it may become clearer in these instances that these are persons with serious mental illness who were not receiving adequate treatment and who repeatedly acted in an inappropriate and/or aggressive manner when stressed, which caused them to come to the attention of law enforcement officers and be arrested. An example would be a woman with a history of bipolar 1 disorder who has a history of arrests for crimes including disturbing the peace (playing loud music throughout the early morning hours), destroying public property (defacing doors in a county building), resisting arrest for disorderly conduct (screaming at her neighbors in the middle of the night), and hitting a man on the bus whom she falsely believed was making sexual overtures toward her. Her most recent arrest was for throwing rocks at her neighbor because she mistakenly believed he was spying on her while she was taking a shower. Without the input of a mental health professional to evaluate this individual, the judge might believe that she should be incarcerated instead of referred to a hospital or other psychiatric treatment. Thus, this person with serious mental illness would continue to be criminalized.

OUTPATIENT TREATMENT TO REVERSE OR PREVENT CRIMINALIZATION

To decriminalize persons with serious mental illness, it is necessary to find ways to help them become stabilized outside of jails and prisons and, to the extent possible, not enter the criminal justice system at all. Thus, the community treatment of persons with serious mental illness who are or may become offenders has developed into an increasingly important and urgent issue. There are greatly increased numbers of persons with serious mental illness who find their way into the criminal justice system, are released into the community, and unfortunately, revolve back into the criminal justice system.

A study examined all persons (N = 79,211) who were incarcerated in one state's prison system during a 1-year period. Ten percent were diagnosed as having major depressive disorder, bipolar disorders, schizophrenia, or nonschizophrenic psychotic disorders. Twenty-one percent of these inmates with serious mental illness were currently incarcerated for a violent offense. These inmates also had a statistically higher percentage of assault, robbery, and property offenses than the inmates who had no psychiatric disorder. In addition, the inmates with serious mental illness were far more likely to have a history of incarceration than those without a serious mental illness. As a result of this study's findings, the authors acknowledged that there was a need to not only develop programs to treat persons with serious mental illness in correctional facilities but also to find effective alternatives to incarceration in the community, including pre and post-booking diversion programs, increasing discharge planning or

continuity of care programs, drug and alcohol screening and treatment, and access to psychiatric inpatient treatment.

The purpose of this section is to review community treatment of persons with serious mental illness who without treatment have been or might become criminalized, and to discuss the principles of treating this population as viewed generally by clinicians who conduct such treatment. This includes those who specialize in forensic mental health care. In addition, because such large numbers of persons with serious mental illness have become criminalized and then released into the community, many mental health professionals who previously may not have been involved in treating offenders with mental illness are now finding themselves with treatment responsibilities for this population.

Probably, many criminalized persons with serious mental illness can be treated at mainstream mental health clinics on their release from jails and prisons. This is exemplified by nondangerous persons with psychosis who are arrested for minor crimes, such as disturbing the peace or not paying for a meal in a restaurant, and who are taken to jail rather than the more appropriate disposition of taking them to a hospital. Another example would be individuals with serious mental illness who committed felonies and were sent to prison but are nondangerous, such as a person with bipolar disorder who was sentenced to prison for forging checks.

However, it must be recognized that there are many who are discharged from correctional institutions who have multiple problems that cannot be adequately treated in traditional community-based facilities. An example of this would be persons with serious mental illness who have a history of dangerousness; placing such persons in a traditional community clinic often results in exacerbations of their illness and acts of violence that lead to rehospitalization or reincarceration. For these individuals, it is important to have special, highly structured clinics staffed by professionals who understand dangerous offenders with mental illness and their treatment. Moreover, adequate security in such clinics is essential; this should include metal detectors, the presence at all times of trained security personnel, and emergency alarm buttons within easy reach of the mental health staff. These clinics may be in the mental health system, with the patients' treatment monitored by the criminal justice system (e.g., court and/or department of probation), or the clinics may be an actual part of the criminal justice system (e.g., run by parole departments).

Treatment of Co-Occurring Disorders

It is estimated by professionals and other personnel in the criminal justice system who are knowledgeable about incarcerated persons with serious mental illness that at least three-fourths of them also meet DSM-5 criteria for substance use disorders. Clearly, if treatment after release is to be successful, both the mental illness and the substance abuse must be

addressed. These services should be integrated in the community, and it is generally held that the same clinical personnel should provide and coordinate both mental health and substance abuse treatments. Treatment of co-occurring disorders very frequently needs to be a long-term process.

Assisted Outpatient Treatment

Another important treatment modality that is available in almost all of the states is AOT, often called outpatient commitment. AOT is a court-ordered civil commitment in which the patient is not under the jurisdiction of the criminal justice system. Rather, AOT is initiated by the mental health system to ensure that persons with mental illness and a history of hospitalizations or violence participate in services in the community that are appropriate to their needs. AOT is for persons with mental illness who are capable of living in the community with the help of family, friends, and mental health professionals but are resistant to psychiatric treatment, including medications. Without such treatment, they may relapse and become violent and/or dangerous to themselves and require hospitalization. Because of these characteristics, this population is also prone to be arrested, incarcerated, and criminalized. To prevent decompensation, which may result in hospitalization or criminalization, and to help these persons stabilize in the community, such at-risk individuals are ordered to participate in outpatient psychiatric treatment, with their progress closely monitored by the court.

When AOT is successful, it is because there is usually a requirement that it be accompanied by an intensive and proven form of community treatment, such as ACT or forensic assertive community treatment (FACT). ACT is a community-based program with mobile mental health treatment teams that provide an array of treatment, rehabilitation, and housing services that are available 24 hours a day. FACT is similar to ACT but also includes legal leverage via collaboration with the criminal justice agencies, in an effort to reduce recidivism. The goal of ACT and FACT is to help persons with serious mental illness stay out of the hospital and develop skills for living in the community. It should be noted that AOT, ACT, and FACT are evidence-based practices; that is, they have been shown to be effective, using scientifically valid studies.

There are many who believe that the widespread use of AOT would reduce the criminalization of persons with serious mental illness because it mandates treatment for individuals who otherwise are likely to be arrested and enter the criminal justice system.

Challenges of Mandatory Outpatient Treatment

Individuals in AOT with serious mental illness, as well as those who have committed criminal offenses, represent a challenge to outpatient clinicians.

The problem lies not only in ensuring safety to the community but also in working with individuals who may be resistant to treatment.

A large proportion of persons with serious mental illness who commit criminal offenses have a history of being highly resistant to psychiatric treatment before their involvement in the criminal justice system; this is also true for persons who are considered to need AOT. They may have refused referral, may not have kept appointments, may have been nonadherent with psychoactive medications, and may have refused appropriate housing placements. Many have anosognosia and are not able to recognize that they have a mental illness.

For many offenders with mental illness, problems of resistance may continue after release from incarceration, which can place them at risk to the community. This is a fact that the treating professional must always keep in mind.

The mental health system finds many resistant persons with mental illness extremely difficult to treat and is often reluctant or unable to serve them. The reluctance becomes even greater after these persons have committed offenses, become involved in the criminal justice system, and are referred to community agencies. The disinclination to serve these persons extends to virtually all areas of community-based care, including therapeutic housing, social and vocational rehabilitation, and general social services. Many offenders with serious mental illness are intimidating because of previous violent, fear-inspiring behavior. Treating them is very different from helping submissive, nonthreatening, formerly institutionalized patients to adapt quietly to life in the community. Thus, community mental health professionals not only are reluctant but also may be afraid to treat offenders with mental disorders. In addition, professionals may work in treatment facilities that do not adequately provide for staff safety, do not possess the authority and leverage of the criminal justice system, and do not provide treatment interventions with adequate structure for this population. The reluctance to treat potentially dangerous persons in such treatment facilities may be warranted. Another issue related to mental health professionals' reluctance to treat such individuals is their belief that they do not have adequate training and experience in risk assessment to ensure public safety. As discussed later, outpatient programs need to increase their ability to treat persons who do not see themselves as having a mental illness and/or have a potential for dangerous behavior.

Another important obstacle to receiving outpatient treatment for persons with serious mental illness who are in AOT or are offenders is that community mental health resources may be inappropriate. For instance, these persons may be expected to come to outpatient clinics, when the real need for many in this population is for outreach services where professionals come to them.

As discussed later, a crucial part of the treatment of these persons with serious mental illness is adequately and appropriately structured and supportive, therapeutic housing arrangements.

When the Criminal Justice System Is Involved

When the criminal justice system is involved in the outpatient treatment of offenders with serious mental illness, a need exists for a clear understanding of the criminal justice system's perspectives and goals. These perspectives and goals include an emphasis on concerns about public safety, control of violence, extensive use of authority, and close cooperation between the mental health and criminal justice systems.

There are a variety of circumstances under which offenders with mental illness in the community remain under the jurisdiction of the criminal justice system. One such group consists of individuals given probation by the court that includes a condition of outpatient treatment, with the possibility of being sent to jail or prison if they are nonadherent to treatment. Another consists of individuals released from prison who are referred for treatment by their parole officer with the understanding that failure to comply may result in a revocation of parole and return to custody.

Treatment under the Jurisdiction of the Criminal Justice System

For those offenders who are mandated by the court or an agent of the criminal justice system (probation officer, parole officer) to attend outpatient treatment, the criminal justice system retains jurisdiction over the individual. In such situations, both the offender with mental illness and the therapist must satisfy legal requirements, such as regular attendance and periodic progress reports. The patient must comply with legal restrictions, such as abstinence from drugs and alcohol. Therapists who treat such individuals are often asked to provide reports to the court, probation officer, or parole officer on the patient's treatment, including issues such as attendance, diagnosis, progress, and prognosis.

Both the criminal justice system and forensic clinicians generally expect that offenders with mental illness given treatment will gain some understanding of the role of their psychiatric disorder in past and potential future dangerous behavior, and that they will avoid behavior or situations that might increase the risk of criminal activity or deterioration in their clinical condition. Moreover, both society and the criminal justice system expect that treatment will be conducted under conditions that can, to the greatest extent possible, ensure safety to the community.

Thus, a balance must exist between individual rights and public safety. However, it has been argued that courts place a greater emphasis on the potential dangerousness of the offender with mental illness than on the individual's rights. In so doing, the courts place the burden on the offender

to demonstrate that he or she no longer poses a danger to the community. Consequently, the person's risk of violence, to a large extent, dictates whether such an individual will be placed into and remain in the community.

Outpatient treatment for offenders with serious mental illness is not designed to make presently dangerous individuals nondangerous. Rather, the criminal justice system presumes that offenders with mental illness placed in outpatient treatment will not be dangerous to others while under supervision and treatment in the community. A primary concern in outpatient treatment of these individuals is to assess any changes in mental condition that may indicate violence and to reduce potential threat of harm. Therefore, features that contribute to a patient's risk of harm are addressed first in treatment. The psychiatrist must assess dangerousness and incorporate this assessment into an intervention strategy. Thus, outpatient clinicians must have as great as possible an understanding of each patient's potential for violence as a function of his or her history and psychiatric condition. Moreover, the clinicians should feel comfortable and qualified to engage in such assessments.

In contrast, in nonforensic psychiatric treatment of nonviolent persons with serious mental illness, the primary focus is usually on alleviation of symptoms. Clinicians who treat offenders with mental illness need to recognize that they have assumed responsibility not only to the patient but also to society in ensuring the patient's safety to the community. This responsibility is not as central when clinicians are treating nonoffenders.

Identifying a Treatment Philosophy

To work effectively with this extremely difficult group of psychiatric patients, both those in AOT and those who are offenders, the necessity of identifying and articulating a treatment philosophy of both theory and practice should be emphasized. This philosophy, as already mentioned, should strike a balance between individual rights and public safety and use treatment services that take both into account. A reality-based treatment philosophy is needed, one that includes clear treatment goals, with attention paid to goals expressed by the patient; a close liaison with the court or other criminal justice agency monitoring the patient, including access to each patient's database from the criminal justice and mental health systems; and an emphasis on structure and supervision.

The importance of psychotropic medications for this population with serious mental disorder, most of whom have problems centering on control, cannot be overemphasized. In many cases, if adherence is to be ensured, the patient must live in a community facility in which each dose is dispensed by staff.

A clear conception of the clinical uses and therapeutic value of authority appears to be a cornerstone of successful community treatment for persons

with serious mental illness in AOT or for those who are offenders. When treatment is effective, the staff is not ambivalent about the use of authority. They are comfortable about insisting consistently and reasonably that the imposed conditions be followed, monitoring patients' adherence with prescribed psychiatric medications, and monitoring patients to detect the use of alcohol or illegal drugs. They have no problems with insisting that patients live in appropriately structured and supportive residential settings, to the extent needed by each individual, as a condition for remaining in the community. They are willing to rehospitalize patients in community facilities promptly at times of crises and are willing to recommend revocation of patients' community status and return them to the court with the opinion that community treatment is no longer appropriate.

It is also important to emphasize the legal and ethical aspects of treating persons under the jurisdiction of the court. Before persons in AOT and offenders with mental illness are asked to consent to outpatient treatment, they should be apprised of all the conditions and limitations that will be imposed on them, why they will be imposed, and what will happen if they do not comply. Areas to be addressed include limits to confidentiality, with respect to both past and present treatment and criminal history, and conditions under which such information must be shared with criminal justice and mental health system personnel; supervision and monitoring by various authority figures, such as probation or parole officers, judges, therapists, and case managers; mandatory adherence with treatment and other imposed conditions; and residence in an appropriate living situation. The patient must understand that nonadherence with the terms and conditions may result in revocation of outpatient status. It is also imperative that the treatment staff understand fully the patient's legal status and conditions for community placement and agree to monitor and uphold them. Staff members must accept their role as agents of social control.

Treatment Goals

Generally, mental health professionals believe that community treatment of persons with serious mental illness who are in AOT and/or have been offenders should focus on stabilization of the illness, enhancement of independent functioning, and maintenance of internal and external controls that prevent patients from acting violently and committing offenses. It is hoped that patients will share these goals. Patients should at least have the goal of avoiding involvement with the criminal justice system. Important points for discussions between the clinician and the patient are the patient's understanding of which behaviors and symptoms are of concern, why they are of concern, what is expected of the patient both by the clinician and by the supervising court or criminal justice agency, and how the treatment can help the patient to meet these

expectations.

Even patients with severe psychiatric symptoms may be able to understand these elements of treatment and adopt appropriate attitudes toward treatment. Thus, a full resolution of symptoms is not always needed for treatment to be viable in a community setting. For some persons with long-term, serious mental illness, total elimination of symptoms is not a realistic goal. Rather, the primary prerequisite for safe and effective community treatment may be patients' ability to understand and accept their clinical needs and the system's legal requirements and to demonstrate adherence.

Liaison between Treatment Staff and the Court

An essential aspect of the treatment of offenders with mental illness, especially those under the jurisdiction of the criminal justice system and of persons in AOT, is a close liaison among treatment staff and the court, the district attorney's office, the departments of probation and parole, and the patient's counsel. At the core of the liaison is a complete and relevant database, which is fundamental in understanding the extent of the patient's problems, determining whether outpatient treatment is appropriate for the patient, and developing a treatment plan. The database should include arrest reports, computerized criminal histories ("rap sheets"), hospital records, evaluations by court-appointed psychiatrists or psychologists, results of psychological testing, probation reports, and records from previous hospitalizations, outpatient treatment, and incarcerations.

A successful liaison requires open, frequent, and continuing contact between the treatment staff and the various aforementioned personnel in the legal system. In addition, all must respect each other's perspective and accept that they are all working toward the same goals. Together, the mental health and legal systems lend their expertise in developing and modifying the most effective community treatment program for patients. Both must accept that such a program cannot be effective for certain individuals who need hospitalization or are too dangerous and thus are not appropriate for outpatient treatment.

The Need for Structure

Usually, patients referred for mandatory outpatient treatment lack internal controls; they need external controls and structure to organize them to cope with life's demands. For instance, mental health professionals who treat this population generally believe that staff should insist that patients' days be structured through meaningful, therapeutic activities such as work, day treatment, and various forms of social therapy. Another basic element of structure for this population is that treatment is mandatory and under the jurisdiction of the court.

What effect does involuntary treatment have on persons with serious mental illness? Adherence has been shown to increase when persons with serious mental illness are required to undergo involuntary treatment. Adherence is important because it has been demonstrated that persons with serious mental illness who do not comply with treatment present an increased risk to themselves and to the community. Thus, treatment nonadherence may in and of itself result in rehospitalization or incarceration.

For offenders with mental illness, the criminal justice system tends to view outpatient treatment or conditional release programs as provisional; such programs are not regarded as the right of the offender with mental illness. Thus, the courts are not obliged to grant this form of treatment; if granted, the courts allow the state to quickly and easily revoke it. The possibility of revocation of outpatient status gives the treatment staff powerful leverage to ensure adherence to the treatment conditions.

Finally, it is essential that staff understand and accept the therapeutic value of structure and their use of authority if community treatment of offenders with serious mental illness is to be successful.

Management of Violence

A major focus of therapy should be on high-priority issues such as the need for the patient to control impulses and inappropriate expressions of anger. Persons whose violence is rooted in a major mental illness often experience their violence as a frightening loss of control. Because their violent acts often occur in the context of psychosis, they tend to perceive a clinician who is not aware of their destructive potential as unable to protect them. They tend to establish that knowledge by testing the clinician for limits. The clinician must be continuously alert and firm in order not to risk being perceived as uncaring and unable to protect the patient from his or her destructiveness. Treatment contracts made by the psychiatrist and the patient that deal with expectations and consequences openly and from the beginning provide a helpful structure for working with these potentially violent patients.

Persons with serious mental illness, especially those with histories of violent behavior, generally need continuous rather than episodic care. Thus, regular monitoring is needed, especially when symptoms are absent or at a low ebb, in order to deal with individual and situational factors that may arise and result in violence.

Persons with serious mental illness who have been hospitalized or incarcerated frequently enter outpatient treatment while they are still taking medication and may for the first time in years find themselves symptom free, in control of their violent impulses, and able to function in the community. After a while they may believe that they can succeed without medication; they discontinue its use, decompensate, and perhaps

engage in violent behavior. Thus, for this population of persons with serious mental illness, a large number of whom have problems involving control and violence, the importance of continued adherence to antipsychotic medications and medications for bipolar and depressive disorders cannot be overemphasized.

In addition, behavioral, cognitive, and psychoeducational techniques emphasizing anger management have been widely used and have been successful in the treatment and management of violence.

Integrating Treatment and Case Management

The purpose of case management is to help patients obtain the services that they need in a timely manner and for as long as they need them. Almost all of these patients need the basic elements of case management, which starts with the premise that each patient has a designated professional with overall responsibility for his or her care.

The case manager formulates an individualized treatment and rehabilitation plan with the patient's participation. As care progresses, the case manager monitors the patient to determine whether he or she is receiving treatment, has an appropriate living situation, has adequate funds, and has access to vocational rehabilitation. The treatment plan emphasizes helping the patient to deal with practical problems of daily living. In addition, the case manager provides outreach services to the patient wherever he or she is living, whether alone, with family, in a board-and-care home, or in another residential setting.

Outreach services may take the form of assertive case management. An assertive case management program deals with patients on a frequent and long-term basis. It takes a hands-on approach that may necessitate meeting with patients on "their own turf" or even seeing a patient daily. This form of contact and familiarity with the patient helps the case manager anticipate and prevent a significant decompensation. Low caseloads for case managers of potentially violent persons with mental illness—probably not more than ten cases—have been recommended, as has 24-hour-a-day, 7-day-a-week availability. Many violent acts and arrests occur during evenings, nights, and weekends, when traditional treatment programs are closed.

Patients who perceive the case manager as mainly an agent of the court and whose primary intention is to make them "toe the line" will be less likely to form a positive relationship with a case manager. They may be guarded and defensive in their interactions. They may also feel a need to portray themselves in a good light, even though they have problems for which they need help. Therefore, case managers, in addition to enforcing the terms and conditions of the court and the treatment, need to be persons who will go beyond this role to secure and provide services and support. It is crucial that patients see the case manager as their advocate to further

their treatment and rehabilitation, particularly when agencies in the criminal justice system are dealing simultaneously with them in more coercive or authoritarian ways.

Case managers may experience considerable stress in working with patients in mandatory outpatient treatment. They may feel, with some justification, that they are in danger if they work with persons with serious mental illness who have been violent in the past or if their work requires visits to high-crime areas, where many persons with serious mental illness live. Additional problems may include unusual working hours, which may disrupt their family and social relationships, and lack of upward career mobility.

Therapeutic Living Arrangements

Survival in the community for the majority of persons with serious mental illness appears to depend on an appropriately supportive and structured living arrangement. Such an arrangement can often be provided by family members. In many cases, however, the kind and degree of structure the patient needs can be found only in a living arrangement outside of the family home with a high staff–patient ratio, dispensing of medication by staff, enforcement of curfews, and therapeutic activities that structure most of the patient's day.

Some patients need a great deal of structure and supervision in their housing situation, others need only a minimal amount, and most fall somewhere in between. The treatment staff member assigned to the patient or the patient's case manager must decide whether a particular living arrangement has the appropriate amount of structure to meet the patient's needs. However, it is important to first discuss the suggested living arrangement with the court or the responsible agent in the criminal justice system and obtain his or her approval.

A consideration for all patients who move from a closed or locked setting such as a hospital, jail, or prison is the patient's ability to handle transition. It is generally not advisable for persons who have been hospitalized or incarcerated for a long time to then be placed in the community in a living situation with little or no structure. Such individuals are frequently unable to cope with the immediate stress and demands of these arrangements, and they either decompensate or commit subsequent offenses.

For many persons with serious mental illness, it is helpful to release them to graduated lower levels of structure—for example, from a hospital or prison to a locked community facility to a halfway house and, when they are ready, to family or independent living. In the graduated release process, close attention must be paid to patients' coping skills and need for monitoring.

Working with the Family

The role of family members or significant others can be critical in the treatment of offenders with mental illness or for those in AOT. The treatment team should determine whether these individuals were the victims of the patient's aggression, whether they have maintained contact with the patient, and whether they are interested in continuing such contact. The team should also learn whether other social support systems are available while the patient resides in the community. Social support can be an extremely important part of community treatment for persons with mental illness. Assessing problems that may develop between the patient and family members or significant others is essential if contact between them is anticipated.

Another important consideration is the family members' needs for guidance and support, especially when they have been victimized by the patient. Clinicians should help them in understanding the patient's mental condition, teach them to recognize symptoms of decompensation, demonstrate methods for self-protection, and explain the patient's current legal situation.

Moreover, family members should be involved in support groups to help them during crises. In self-help programs, they can benefit from the experience of other families in similar situations.

Cultural Competency

It is important that mental health professionals understand the roles that ethnicity, race, culture, gender, and age play in how mental illness manifests. Equally important is an understanding of the kinds of treatment and how treatment is presented that will be acceptable to these groups. Only then can the mental health system provide to racial, cultural, and ethnic minorities and persons of all ages and both genders services that are appropriate for their needs and ones to which they will be responsive.

Pitfalls

Mental health professionals must understand their potential liability for the actions of their patients on community status who are prone to violence. They need to be aware that in some states, courts have imposed on therapists a duty of care that extends to foreseeability of harm. Moreover, the litigious nature of society adds further pressures, stress, and risks. Treating clinicians may be held accountable for patients' suicides, assaults, homicides, and other crimes, even when treatment was mandated, sound, and met the professional community's standards of care. In both public agencies and private practice, malpractice insurance coverage must be adequate so that clinicians feel comfortable, personally and professionally, in undertaking the treatment of persons with serious mental

illness who need mandatory outpatient treatment.

Another important concern of those clinicians involved in the community treatment of offenders with serious mental illness, especially persons who have committed high-profile crimes, is the possibility of notoriety and unfavorable publicity. Clinicians may fear that if patients commit further sensational crimes, their professional reputation will suffer and their facility may receive negative media coverage.

Twenty-Four–Hour Structured Inpatient Care

Not all persons with mental illness can be treated effectively in the community. There is a substantial minority of persons with serious mental illness who need the structure and support of acute, intermediate, or long-term care in a hospital setting or a highly structured, locked 24-hour care community facility. It is important to identify high-risk patients so that scarce community treatment resources are used for those most likely to succeed and for those who are least likely to incur the high costs of lengthy hospitalization or rearrest for serious offenses.

It is essential to facilitate access to hospital care for patients who need it. In the early years of deinstitutionalization, many believed that if the counter-therapeutic practices that had been exposed in some state psychiatric hospitals could be eliminated, this would simultaneously eliminate the need for hospitals altogether. However, experience has shown that some persons with serious mental illness require hospitalization, and some may need it for an extended period of time. This may be true even in communities that offer a complete array of excellent and integrated community-based services.

It is important to acknowledge that the community is not necessarily the most benign treatment site for all people with mental illness at all times. Access to hospital care for those who need it for as long as they need it is absolutely essential to the success of deinstitutionalization and the reduction in criminalization for this population.

For those persons with serious mental illness who do not require a setting as highly structured as that of a locked psychiatric hospital but who are not ready for an open community setting, intermediate care facilities have proven extremely valuable. An example is the institutes for mental disease (IMDs) in California. IMDs are locked, private, community intermediate care facilities with a moderate but lesser degree of structure than that found in long-term psychiatric hospitals, such as state hospitals. The facility provides very close medication supervision, with many patients receiving their medication in either liquid or crushed form. The staff–patient ratio is high, with at least 27 hours of activities a week for each patient. These activities include group therapies, therapeutic, social, and recreational activities, and instruction in activities of daily living.

Intermediate care facilities have been used widely for both civilly and

criminally committed patients. These intermediate care facilities have helped to stabilize persons who might well have been criminalized; for those persons coming from jails and prisons, they might well have prevented recidivism.

Decarceration of Jails and Prisons

Imprisonment rates in the United States have increased 500 percent over the past 30 years; currently, the United States has the highest rate of incarceration in the world. This has resulted in prison overcrowding and great financial costs. Consequently, many efforts are underway currently to decrease the number of persons imprisoned. Some of these are reductions in state and federal sentences for less serious crimes, including drug offenses; the use of alternatives to prison, by means of diversion programs; offering early release to prisoners who participate in rehabilitation programs; and legal mandates to reduce prison overcrowding.

Beginning in 2009, the number of prisoners in US prisons and jails dropped slightly each year. The greatest reduction in prisoners was in California, which exemplifies many of the factors that led to decarceration. In the case of *Brown v. Plata*, the US Supreme Court in 2011 found that overcrowding in California's prisons resulted in a violation of prisoners' constitutional rights, and upheld an order to significantly reduce the number of state prisoners. Legislation known as "Public Safety Realignment" was enacted that year which used two primary methods to reduce overcrowding, costs, and recidivism. The first method focused on reducing admissions to prison, by having many nonserious, nonsex, and nonviolent felony offenders serve sentences, generally of 3 years or less, in county jail instead of prison. The second method focused on having fewer prisoners return to prison; that is, inmates who previously were released on state parole were now released under county probation, to be supervised and provided services by county agencies. If they violated their conditions of probation (but without committing a new crime), they could not be returned to prison but could be placed in jail.

"Realignment" became effective on October 1, 2011. At that time there were 143,710 state prisoners. By January 4, 2017, the number had been reduced to 114,000, a reduction of 29,710 state prisoners. Persons with serious mental illness may account for a sizeable number of those released. As mentioned previously, recent studies estimate that approximately 16 percent of persons in prisons have a serious mental illness. If this percentage holds for the 29,710 reduction in state prisoners, then approximately 4,754 (29,710 × 16%) fewer persons with a serious mental illness are incarcerated in California prisons.

States across the nation are addressing the financial costs and effectiveness of incarceration for those currently in prison. Relying on local jails instead of prisons to incarcerate offenders is also likely to result in

high costs and overcrowding conditions in the jails, and lead to releasing the least dangerous jail inmates for more serious offenders. As a result, arresting persons with serious mental illness who commit or are alleged to commit nonserious, nonsex, and nonviolent crimes and placing them in jails may become less common than in the past.

Consequently, the era of the criminalization of large numbers of persons with serious mental illness may be ebbing in both state prisons and jails. The criminal justice system will not be able to absorb the great numbers that they did formerly. As a result, during decarceration the mental health system may be expected to assume the care and treatment for far more persons with serious mental illness. Some of these individuals may be among the most difficult and expensive to treat, as well as at great risk of decompensation and/or reoffense. If, however, society gives a high priority to successfully treating and reducing the hundreds of thousands of persons with serious mental illness who are currently in jails and prisons, or who are being released, or who are at risk of being arrested, some similar problems encountered during deinstitutionalization may be able to be avoided.

Some Possible Strategies

The list of problems is formidable, but they can be resolved. There would need to be a very large increase in the number of community mental health services, including ACT and FACT teams and intensive case management; co-occurring disorder programs; administration of antipsychotic medications and mood stabilizers (especially important in reducing arrests for those who have problems involving control and violence); crisis response teams; appropriately structured and supportive housing; vocational rehabilitation; social rehabilitation; availability of acute care in the community; access to acute, intermediate, and long-term inpatient care when needed; more flexible civil commitment laws; and AOT when indicated. Many of these interventions are evidence-based practices. With a sufficient amount of all of these services being made available, the number of persons with serious mental illness being criminalized should drop significantly.

With the large numbers of persons with serious mental illness who first come to the attention of law enforcement officers, the opportunity to enter the mental health system should be given to those who are appropriate for diversion. Toward this end, there needs to be many more mobile crisis teams, such as CITs of specially trained police officers, who can find alternatives other than arrest and incarceration for such persons in crisis. There should also be increased education and training on mental illness and available mental health resources for law enforcement officers in general. Diversion after arrest may be accomplished in mental health courts, where persons with mental illness are sent to the mental health

system to obtain treatment in lieu of incarceration. It needs to be emphasized that pre and post-booking diversion can be effective only if there are adequate and appropriate mental health resources to which persons with serious mental illness can be diverted.

For those persons with serious mental illness who enter and remain under the jurisdiction of the criminal justice system, creating a successful treatment program for them is a difficult task that demands the input and cooperation of correctional and mental health professionals who are knowledgeable and accepting of the tenets of both the criminal justice system and mental health treatment. A close liaison among such varied professionals helps to maximize a successful outcome for offenders with mental illness. These persons should be provided with appropriate treatment while they are incarcerated and after they have been released into the community. Moreover, all such programs should be evaluated independently to determine which are effective and which are not, to ensure the best possible use of scarce resources.

Finally, a significant increase in mental health services, ranging from outpatient treatment and case management to highly structured 24-hour care, would no doubt result in far fewer persons with serious mental illness committing criminal offenses. There is an enormous stigma attached to people who have been in jails and prisons or in a forensic hospital. Such persons have been categorized as both mentally ill and as offenders, which makes it extremely difficult to place them in community treatment and housing.

In this time of decarceration, psychiatrists are poised to provide persons with serious mental illness the mental health treatment and supportive social services that were inadequately provided previously for so many, and thus led to their decompensation and criminal behavior. If the goals of reducing the criminalization of persons with serious mental illness are to be accomplished, the mental health and criminal justice systems must be provided with all the necessary resources and funding to identify and treat these individuals in the most appropriate setting. It cannot be emphasized enough that the criminal justice system should not be used as a substitute for the mental health system.

References

Baillargeon J, Binswanger IA, Penn JV, Williams BA, Murray OJ. Psychiatric disorders and repeat incarcerations: the revolving prison door. *Am J Psychiatry*. 2009;166:103–109.

Beach C, Dykema L, Appelbaum PS, et al. Forensic and nonforensic clients in assertive community treatment: a longitudinal study. *Psychiatr Serv*. 2013;64:437–444.

**Brown v. Plata* 131 S. Ct. 1910 (2011).

Choe JY, Teplin LA, Abram KM. Perpetration of violence, violent victimization, and severe mental illness: balancing public health concerns. *Psychiatr Serv*. 2008;59:153–164.

Cole D. Mass incarceration: causes, consequences, and exit strategies: turning the corner on mass incarceration? *Ohio St J Crim L*. 2011;9:27–51.

Fazel S, Zetterqvist J, Larsson H, Långström N, Lichtenstein P. Antipsychotics, mood stabilisers, and risk of violent crime. *Lancet*. 2014;384:1206–1214.

Harcourt BE. Mass incarceration: causes, consequences, and exit strategies: reducing mass incarceration: lessons from the deinstitutionalization of mental hospitals in the 1960s. *Ohio St J Crim L*. 2011;9:53–88.

Kaeble D, Glaze L. Correctional populations in the United States, 2015. *Bureau of Justice Statistics*. 2016;NCJ 250374:1–20.

Lamb HR, Weinberger LE. Some perspectives on criminalization. *J Am Acad Psychiatry Law*. 2013;41:287–293.

Lamb HR, Weinberger LE. Decarceration of our jails and prisons: where will persons with serious mental illness go? *J Am Acad Psychiatry Law*. 2014;42:489–494.

*Lamb HR, Weinberger LE. Rediscovering the concept of asylum for persons with serious mental illness. *J Am Acad Psychiatry Law*. 2016;44:106–110.

Lamb HR, Weinberger LE, Marsh JS, Gross BH. Treatment prospects for persons with severe mental illness in an urban county jail. *Psychiatr Serv*. 2007;58:782–786.

Lamberti JS. Preventing criminal recidivism through mental health and criminal justice collaboration. *Psychiatr Serv*. 2016;67:1206–1212.

Lamberti SJ, Deem A, Weisman RL, LaDuke C. The role of probation in forensic assertive community treatment. *Psychiatr Serv*. 2011;62:418–421.

Lehrer DS, Lorenz J. Anosognosia in Schizophrenia: hidden in plain sight. *Innov Clin Neurosci*. 2014;11:10–17.

Lurigio AJ. Co-occurring disorders: mental health and drug misuse. In: Leukefeld C, Gullotta TP, Gregrich J, eds. *Handbook of Evidence-Based Substance Abuse Treatment in Criminal Justice Settings*. New York: US Springer Science + Business Media. 2011:279–292.

Lurigio AJ. Examining prevailing beliefs about people with serious mental illness in the criminal justice system. *Fed Probation*. 2011;75:11–18.

Markowitz FE. Mental illness, crime, and violence: risk, context, and social control. *Aggress Violent Behav*. 2011;16:36–44.

Mundt AP, Winnie SC, Arduino M, et al. Psychiatric hospital beds and prison populations in South America since 1990: does the Penrose hypothesis apply? *JAMA Psychiatry*. 2015;72(2):112–118.

Munetz MR, Griffin PA, Kemp K. Jail diversion: using the sequential intercept model. In: Yeager KR, Cutler DL, Svendsen D, Sills GM, eds. *Modern Community Mental Health: An Interdisciplinary Approach*. New York: Oxford University Press; 2013:453–472.

Pettinati HM, O'Brien CP, Dundon WD. Current status of co-occurring mood and substance use disorders: a new therapeutic target. *Am J Psychiatry*. 2013;170:23–30.

Player CT. Involuntary outpatient commitment: the limits of prevention. *Stanford Law & Policy Review*. 2015;26:159–238.

*Redlich AD. The past, present, and future of mental health courts. In: Wiener RL, Brank EM, eds. *Problem Solving Courts: Social Science and Legal Perspectives*. New York: Springer; 2013.

Silver E, Felson RB, Vaneseltine M. The relationship between mental health problems and violence among criminal offenders. *Crim Justice Behav*. 2008;35:405–426.

Sisti DA, Segal AG, Emanuel EJ. Improving long-term psychiatric care: bring back the asylums. *JAMA*. 2015;313:243–244.

*Slate RN, Buffington-Vollum JK, Johnson WW. *Criminalization of Mental Illness*. 2nd ed. Durham, NC: Carolina Academic Press; 2013.

Steadman HJ, Redlich A, Callahan L, Robbins PC, Vesselinov R. Effect of mental health courts on arrests and jail days. *Arch Gen Psychiatry*. 2011;68:167–172.

*Swartz MS, Wilder CM, Swanson JW, et al. Assessing outcomes for consumers in New York's assisted outpatient treatment program. *Psychiatr Serv*. 2010;61:976–981.

Torrey EF, Fuller DA, Geller J, Jacobs C, Ragosta K. *No Room at the Inn: Trends and Consequences of Closing Public Psychiatric Hospitals*. Arlington, VA: Treatment Advocacy Center; 2012:1–28.

Torrey EF, Zdanowicz M, Kennard A, et al. *The Treatment of Persons With Mental Illness in Prisons and Jails: A State Survey*. Arlington, VA: Treatment Advocacy Center; 2014. Available at http://tacreports.org/storage/documents/treatment-behind-bars/treatment-behind-bars.pdf. Accessed January 10, 2017.

Van Dorn R, Volavka J, Johnson N. Mental disorder and violence: is there a relationship beyond substance use? *Soc Psychiatry Psychiatr Epidemiol*. 2012;47:487–503.

59 Psychiatric Education

▲ 59.1 Graduate Psychiatric Education

GEORGE ALAN KEEPERS, M.D.

Over the last 100 years, the education of physicians in the specialties of medicine has evolved from informal, unsystematized training of variable length and content into the highly regulated residency training programs that constitute the normal educational path for medical school students throughout the United States today. At times gradual, the development of these programs was changed dramatically by the influence of the Flexner report which resulted in the standardization of both undergraduate and graduate medical education and the adoption of a residency training model begun at Johns Hopkins University. Today, graduate medical education is undergoing another disruptive episode of evolution championed by the ACGME, the organization responsible for the accreditation of all residency training programs in the United States. In the past, training programs were standardized around the processes and contents of their clinical and didactic curriculum. The Accreditation Council on Graduate Medical Education's (ACGME) competency movement and its more advanced elaboration, the Milestones Project seek to use measures of the intermediate and final outcomes of residency training to judge the adequacy of training and of training programs.

This textbook has always covered the subject of psychiatric training but each author has developed the topic in a somewhat different fashion. In earlier editions, the role of the American Board of Psychiatry and Neurology in influencing training programs was emphasized by Lester Rudy and subsequently by Steven Scheiber. James Eaton, in contrast, emphasized the historical development of the field and the influence of the National Institute of Mental Health (NIMH) in providing funding and direction for the development of high-quality training programs. Most recently, Larry Faulkner reviewed the major organizations that influence psychiatric training programs, provided an overview of the organizational structures that regulate and determine the processes and content of

training programs, and addressed issues in psychiatric education that were then relevant.

At the time of the previous edition, as Faulkner noted, psychiatric educators faced a number of controversial issues, which he reviewed in some detail. Some of these issues have been resolved while others continue to be very actively discussed by psychiatric educators. The interference of service requirements in residency with residents' educational activities has been effectively eliminated by ACGME regulations which limit work hours, require protected time for education and incorporate residents' judgments regarding this issue into the Accreditation Council on Graduate Medical Education (ACGME) Psychiatry Residency Review Committee's (PRRC) evaluation of training programs.

Conflict in the training curriculum between biological psychiatry and psychotherapy education has been reduced by the field's consensus regarding the importance of both and by the specification in the PRRC's Program Requirements (PRs) and residency training milestones (discussed below) of specific training requirements in both areas. The conflict between resident support and resident evaluation has also been softened by recent developments. Academic medical centers are increasingly developing resident support systems which exist outside of departmental structures. The movement toward competency-based education embodied in the milestones supports a developmental rather than summative or judgmental approach to residency training. Nonetheless, balancing the provision of specific feedback based on high expectation while maintain a positive, nurturing environment for residents remains a challenge.

While ABPN certification following residency remains nominally voluntary, two developments in the field have pressured residents and programs toward 100 percent participation in certification. First, many, if not most, health systems are now requiring American Board of Medical Specialties (ABMS) specialty certification as a condition of credentialing. Secondly, the PRRC's PRs, which took effect in July of 2015, require that programs maintain a 70 percent pass rate on the certification examination for their graduates.

One issue raised in the previous edition remains a significant area of controversy. The limitation of duty hours and reduction in the clinical responsibilities assigned to residents has had the unintended consequence of reducing residents' continuous contact with patients and delays their maturation as independent practitioners. It is even uncertain as to whether the reductions in work hours has led to the intended improvements in patient safety because of the multiplication in transitions of care required in the new paradigm. The ACGME is currently studying this issue.

This chapter adopts an historical approach to the topic of Graduate Psychiatric Education (GPE). The history of psychiatric training will be described, the development of the organizations responsible for regulating

it will be reviewed, and the content and structure of residency training as it now exists will be described. Several recent changes in the conduct of GPE will be discussed and the above and other controversial topics reviewed.

HISTORY OF PSYCHIATRIC EDUCATION

Early History

Prior to Abraham Flexner's 1910 report for the Carnegie Foundation, medical education in the United States was in an abysmal state. There were no established standards for undergraduate medical education, no requirements for graduate medical education and little to no regulation of the practice of medicine. Different medical traditions including allopathy, osteopathy, and naturopathy provided highly varied training experiences. Physicians might then go directly into practice or seek employment, sometimes under the guidance of a more senior physician who had a particular interest. It was not unusual for physicians to practice in fields for which they had no particular training and for their interests and practice to shift during the course of their careers. Those engaged in what is now called psychiatric practice in the United States fell into two camps—on one hand, the alienists, and on the other, the physicians who conducted ambulatory practices in nervous conditions. These divisions were not original to the United States but were rooted in developments in the field in Europe. The "spa physicians" made great efforts to assert that their clients and practices were different from the mentally ill. The "alienists" explicitly dealt with severely ill patients and were involved, often as superintendents, with newly developed institutions, frequently state sponsored, which specialized in the care of severely mentally ill individuals. Originally innovative and providing access to moral treatment, these institutions became so overcrowded that they were at best able to provide custodial care.

Many physicians, of varying interests, practiced and trained in these settings. For example, Dr. James Hawthorne, who established such an institution in Portland, Oregon, was typical of the physicians of this era. A native of Pennsylvania, he began his study of medicine under the tutelage of a local physician Dr. Bascom, later obtaining a medical degree from the Medical University in Louisville, Kentucky. He opened a general practice in Louisville, which became very successful, and repeated this success in California, starting in 1850. He moved to Portland in 1857, opening a general practice, and subsequently established the Oregon Hospital for the Insane. The Hawthorne hospital eventually had 450 psychiatric beds and was supported on the basis of a contract with the state legislature. It was regarded as one of the finest such institutions in the United States, and Dr. Hawthorne, who had no formal training in psychiatry, was known as an expert in the illnesses of the mind. He was a member of the Association of Medical Superintendents of American Institutions for the Insane

(organized in 1844), which became the American Psychiatric Association in 1921. Many physicians practiced at the Hawthorne hospital. Most notably, Simeon Josephi came to the hospital and worked under Dr. Hawthorne's supervision. Dr. Josephi left his employment at the hospital to found the Oregon Medical School in 1863. This type of fluid arrangement for training and practice was typical of the times.

Influence of the Flexner Report

Unfortunately, but not surprisingly, this unregulated system led to the public being subjected to poorly trained physicians, naturopaths, and snake oil salesman all claiming to be legitimate medical practitioners. The American Medical Association (AMA) was organized in 1847 originally to deal with the lack of regulations and standards in medical education and medical practice. Despite the establishment of residency training programs at Johns Hopkins University in 1889, progress was quite limited until the Carnegie Foundation employed Abraham Flexner to study the condition of medical education in the early 1900s. The current system of undergraduate and graduate medical training was founded on Flexner's 1910 report.

Governmental and professional regulatory mechanisms soon followed and have proliferated into a complex group of organizations affecting GPE today. This resulted in the closure of inadequate schools and established allopathy and osteopathy as the only legitimate basis for medical training. Flexner strongly favored the residency program established by Johns Hopkins, and this became and has remained the framework for all residency training in the United States (Table 59.1–1).

Following the Flexner report several specialties (Ophthalmology, 1916; Otolaryngology, 1924; Obstetrics and Gynecology, 1930; Dermatology and Syphilology, 1932) established boards for the purpose of describing training in the specialty and examining candidates for certification. The American Board of Psychiatry and Neurology (ABPN) was established in 1934 as a collaborative effort of the APA, the American Neurological Association (ANA), and the AMA. Originally, like other Boards, the ABPN both accredited training programs and certified practitioners.

Table 59.1–1.
Major Organizations Affecting GPE

Organization	Abbreviation	Responsibility
American Association of Directors of Psychiatric Residency Training Programs	AADPRT	Promotion of excellence in psychiatric training; resources for program directors; organizes input to ACGME and ABPN from the field.
Association of American Medical Colleges	AAMC	Association of medical schools, teaching hospitals, and academic societies; improvement of health through enhancing effectiveness of academic medicine

American Board of Medical Specialties	ABMS	Coordination of activities for specialty boards
American Board of Psychiatry and Neurology	ABPN	Certification and maintenance of certification of psychiatrists and neurologists
American College of Psychiatry	ACP	Honorary society; develops Psychiatric Resident-in-Training Examination (PRITE); awards for outstanding residents and training programs.
Accreditation Council for Continuing Medical Education	ACCME	Approval of continuing medical education programs
Accreditation Council for Graduate Medical Education	ACGME	Development of residency training Program Requirement (PRs), milestones, accreditation of institutions and residency training programs
American Medical Association	AMA	Largest professional medical association; advocacy for patients and profession
American Osteopathic Association	AOA	Accreditation of osteopathic schools and residencies
American Association of Colleges of Osteopathic Medicine	AACOM	Coordinates and advocates for osteopathic medical schools; centralized services.
American Psychiatric Association and state psychiatric associations	APA	Professional association for psychiatrists; multiple educational efforts for students, residents, practitioners; advocacy for patients and the profession
Educational Commission for Foreign Medical Graduates	ECFMG	Credentialing of international medical graduates for training in the United States
Federation of State Medical Boards	FSMB	Coordinates activities of state licensing boards.
Liaison Committee on Medical Education	LCME	Accreditation of medical schools
National Board of Medical Examiners	NBME	US National Medical Licensure Exam (USMLE) for medical students and residents required for licensure
National Residency Matching Program	NRMP	Regulates recruitment of residents. Matches residents and programs.

Establishment of the Accreditation Council on Graduate Medical Education

The burden imposed on the Boards by the accreditation of training programs was substantial and in 1940 the AMA Council on Medical Education, the American Board of Internal Medicine, and the American College of Physicians established the Conference Committee on Graduate Training in Internal Medicine, a forerunner of the ACGME Residency Review Committees (RRCs). The American College of Surgeons and American Board of Surgery established their own accreditation bodies in 1949.

Academic educators soon realized that, with the establishment of Medicare and federal funding for residency training, the several accrediting organizations then in existence with widely varying standards for approval of training programs were not tenable and required reform. The case for

uniformity of standards with national scope was supported by the AMA, other medical groups, and the government. In 1972, the AMA sponsored the ABMS, the AHA, the AAMC and the Council of Medical Specialty societies to form a Liaison Committee for Graduate Medical Education (LCGME) to coordinate the activities of the independent RRCs then in existence. The failure of this effort indicated that an independent, self-sustaining organization was required to accomplish these tasks and the ACGME replaced the LCGME in 1981. The ACGME was charged with the responsibility for the accreditation and monitoring of residency programs throughout the United States. In the subsequent years, the current system of RRCs, program requirements, and periodic accreditation site visits was developed.

The evolution of the ACGME PRs in psychiatry has been a collaborative effort between the ACGME, the ACGME PRRC, psychiatric professional organizations, and the ABPN. This has resulted in more-uniform, higher-quality training throughout the United States and an alignment between completion of an accredited residency and eligibility to take the ABPN certification examination. Beginning in 2010, the ACGME undertook a major revision of the methods used to evaluate residents and training programs. Entitled the Next Accreditation System (NAS), this new system exchanges yearly reviews of data from all programs with the selection outliers for the old system of in-depth reviews of programs every 5 years following site visits. The details of this system are discussed below.

The American Board of Psychiatry and Neurology

The ABPN retained the task of certifying graduates of accredited training programs in psychiatry and neurology. The ABPN, like the other 24 specialty boards of the ABMS, is a nongovernmental, independent organization governed by a board of directors (Table 59.1–2).

The board is comprised of eight psychiatry directors and eight neurology directors, each of whom serve two 4-year terms (Table 59.1–3). Candidates for selection as psychiatry directors are solicited from the APA, the American College of Psychiatry (ACP), and the AMA. The ANA and the American Academy of Neurology (AAN) nominate candidates in neurology. Current directors of the board may also nominate candidates.

Table 59.1–2.
The 24 Member Boards of the American Board of Medical Specialties (Year Approved in Parentheses)

Allergy and Immunology (1971)[a]
Anesthesiology (1941)
Colon and Rectal Surgery (1949)
Dermatology (1932)

Emergency Medicine (1979)
Family Medicine (1969)
Internal Medicine (1936)
Medical Genetics (1991)
Neurological Surgery (1940)
Nuclear Medicine (1971)
Obstetrics and Gynecology (incorporated 1930)
Ophthalmology (incorporated 1916)
Orthopedic Surgery (1935)
Otolaryngology (incorporated 1924)
Pathology (1936)
Pediatrics (1935)
Physical Medicine and Rehabilitation (1947)
Plastic Surgery (1941)
Preventive Medicine (1949)
Psychiatry and Neurology (1934)
Radiology (1935)
Surgery (1937)
Thoracic Surgery (1971)
Urology (1935)

[a]Conjoint board sponsored by the American Board of Internal Medicine and the American Board of Pediatrics.

Certification methods have changed through the years and these changes are detailed elsewhere in this text. Two most recent developments are of great importance to psychiatric educators. As this text is in print, the last ABPN oral examination in general psychiatry will be given. The oral examination in child psychiatry will be discontinued not long afterwards. The new certification examination in General Psychiatry has proven to have very high reliability, with passage rates that are similar to those of the old Part I examination (Table 59.1–4).

Going forwards, all of the ABPN's examinations in the core specialties of psychiatry, child psychiatry, and neurology, as well as the subspecialties, will be single certification examinations conducted at testing centers throughout the United States. Psychiatric residency training program must certify the credentials of their graduates through the computerized ABPN PreCert system as a condition of eligibility to take the examination. Data from this system makes it feasible to accurately ascertain the quality of residency training programs as measured by certification examination passage rates for individual programs.

Table 59.1–3.
2015 American Board of Psychiatry and Neurology Board of Directors, President and CEO Larry Faulkner, MD

Psychiatry Council

Joan M. Anzia, M.D.	Chicago, Illinois
Robert N. Golden, M.D.	Madison, Wisconsin
Robert W. Guynn, M.D.	Houston, Texas
Paramjit T. Joshi, M.D.	Washington, DC
George A. Keepers, M.D.	Portland, Oregon
Jeffrey M. Lyness, M.D.	Rochester, New York
Robert J. Ronis, M.D., M.P.H.	Cleveland, Ohio
Kallie R. Shaw, M.D.	Tampa, Florida
Neurology Council	
John B. Bodensteiner, M.D.	Rochester, Minnesota
Allison Brashear, M.D.	Winston-Salem, North Carolina
Terrence L. Cascino, M.D.	Rochester, Minnesota
J. Clay Goodman, M.D.	Houston, Texas
Laurie Gutman, M.D.	Iowa City, Iowa
Kerry H. Levin, M.D.	Cleveland, Ohio
Noor A. Pirzada, M.D.	Toledo, Ohio
Ann Tilton, M.D.	New Orleans, Louisiana

Table 59.1–4.
Performance of Candidates on the New ABPN Certification Examination

Year	Candidates	Passed (%)
2011	856	744 (87)
2012	1,288	1,091 (85)
2013	1,513	1,265 (84)
2014	1,867	1,675 (90)

Also of interest is the ABPN's return to a role in accreditation of residency training programs. The ABPN collaborates with other ABMS boards to develop a number of combined educational programs that are not ACGME accredited. The core programs contributing to the training are ACGME accredited programs but the training track for residents in the two programs is not. Residents complete training in these programs in less time than it would take to complete the specialty training programs separately. These programs include training that leads to certification in psychiatry and neurology, psychiatry and internal medicine, and psychiatry and family medicine. The Triple Board Program was established as a collaboration between the ABPN and the American Board of Pediatrics (ABP) in 1984. Graduates of these programs may take the certification examinations in psychiatry, child and adolescent psychiatry, and pediatrics.

A small but significant number of institutions offer these combined training programs. These programs cannot be accredited by the ACGME although the specialty programs which participate in them are reviewed and accredited by their respective RRCs. The ABPN has agreed to provide regular review and approval of the combined training programs in order to assure quality training for the residents in these programs and to assure

that graduates will be eligible to take the certification examinations in each specialty (Table 59.1–5).

The Subspecialties of Psychiatry. Child and Adolescent Psychiatry was the first subspecialty of psychiatry for which the ABPN developed a credentialing process and certification examination. Subsequently, nine other subspecialties were identified and the ABPN now certifies subspecialists in these areas (Table 59.1–6).

The ABMS has specific criteria for the establishment of new subspecialties, and approval for their inclusion must be obtained from that organization. All of the subspecialties in psychiatry, except Brain Injury Medicine, which is still in the grandfathering period, require completion of an ACGME accredited residency. The ACGME PRRC accredits Child and Adolescent, Geriatric, Forensic, Psychosomatics, and Addiction residencies. The other residencies are administered by other boards. All of the subspecialties are 1-year fellowships, except for Child and Adolescent Psychiatry, which is a 2-year program. Residents may enter Child and Adolescent Psychiatry during any part of a General Psychiatry program. One year of training may be counted for both programs, enabling completion of both programs in 5 years rather than 6. A current subject of discussion in the field is whether to allow completion or partial completion of the other subspecialty fellowships in the fourth year of the general residency. All of the subspecialty certifications require taking a single computerized certification examination except for Child and Adolescent Psychiatry. That specialty has traditionally had a Part II oral examination, which will be eliminated by the end of 2016.

Table 59.1–5.
ACGME Combined Specialty Track Programs in 2015

Name	Number of Programs	Number of On-Duty Residents
Internal medicine/Psychiatry	14	81
Psychiatry/Family Medicine	7	49
Peds/Psych/Child and Adolescent Psychiatry	10	87
Psychiatry/Neurology	7	20

Table 59.1–6.
American Board of Psychiatry and Neurology (ABPN) Subspecialties in Psychiatry and Year of First ABPN Examination

Subspecialty	Year of First Examination
Child and Adolescent Psychiatry	1959
Geriatric Psychiatry	1991
Clinical Neurophysiology[a]	1992
Addiction Psychiatry	1993

Forensic Psychiatry	1994
Pain Medicine[b]	2000
Psychosomatic Medicine	2005
Sleep Medicine[b]	2007
Hospice and Palliative Medicine[b]	2008
Brain Injury Medicine[b]	Pending as of 2015

[a]Interdisciplinary specialty involving psychiatrists and neurologists.

[b]Interdisciplinary subspecialties involving ABPN and other American Board of Medical Specialties member boards.

Many subspecialties are developed collaboratively between several ABMS boards. Sleep medicine is an example of such a subspecialty. One of the boards will take the lead in administering the subspecialty and developing the certification examination. Each member board is responsible for credentialing candidates from their own specialty. As new subspecialties are recognized by the ABPN and ABMS, a grandfathering period, usually 5 years, is allowed during which candidates who meet established criteria for experience in the subspecialty area may take the certification examination without completing a residency. After this period expires, candidates must complete appropriate residency training programs to be eligible to take the exam.

The number of core psychiatry subspecialty programs and subspecialty residents has not kept pace with the increase in the numbers of general programs and residents. Except for Child and Adolescent Psychiatry, these programs have fewer than 75 residents in training each year. This is problematic from the standpoint of providing an adequate workforce particularly in geriatric and addiction psychiatry. The failure to interest sufficient numbers of residents in subspecialty fellowship training is the subject of considerable discussion and distress for the profession (Table 59.1–7).

The American Board of Medical Specialties

In 1933, the American Hospital Association (AHA), the Association of American Medical Colleges (AAMC), the Federation of State Medical Boards (FSMB), the American Medical Association (AMA), the National Board of Medical Examiners (NBME), and representatives of the four specialty boards in existence at that time (Ophthalmic Examinations, Otolaryngology, Obstetrics and Gynecology, and Dermatology and Syphilology), established the Advisory Board for Medical Specialties. Originally the ABMS was to serve as an advisor to the Boards, a forum for discussion of common problems, and a coordinator of board work when feasible. When expressly delegated by the Boards, the ABMS was to assume control over common policies and procedures. The organization was restructured in 1970 into the American Board of Medical Specialties. The

AHA, AAMC, FSMB, AMA, and NBME, along with the Council of Medical Specialty Societies (CMSS), the Accreditation Council for Graduate Medical Education (ACGME), and the Educational Commission for Foreign Medical Graduates (ECFMG), continue to serve as associate members of the ABMS. The ABMS has in recent years assumed a more directive role with its 24 member Boards. In particular, the ABMS has established standards which the boards are required to follow, specifying what steps diplomats of the boards must take in order to maintain their certification. These Maintenance of Certification (MOC) programs currently consist of four elements: professional standing (medical licensure), continuing medical education (CME), self-assessment (SA), and performance in practice (PIP), which includes patient and peer feedback and systematic evaluation of the diplomat's clinical care. Additionally, all diplomats of the boards must document completion of a patient safety course.

Table 59.1–7.
Subspecialties of Psychiatry Showing the Number of Programs and the Number of Residents in Programs from 2008 to 2014

		2008	2009	2010	2011	2012	2013	2014
General	Programs	182	184	182	182	184	188	193
	Residents	4,800	4,848	4,952	4,997	5,014	5,200	5,286
Child	Programs	122	124	122	124	124	123	124
	Residents	781	810	807	834	854	882	839
Addiction	Programs	43	45	45	45	45	46	46
	Residents	49	57	67	62	67	69	76
Forensic	Programs	42	44	40	38	36	39	40
	Residents	69	81	69	72	70	73	72
Geriatrics	Programs	58	58	56	54	55	56	55
	Residents	63	64	57	64	65	72	68
Psychosomatic	Programs	44	48	53	54	53	54	54
	Residents	52	58	60	79	53	89	87

In conjunction with the ACGME, the ABMS and some member boards have begun to offer US training standards and practitioner certification abroad. Singapore was the first country to request this program and now has a robust program of residency training and certification for several specialties that is similar to the US equivalents.

The Liaison Committee on Medical Education

Interestingly, accreditation processes for undergraduate medical education were unified later than the initial wave of certification for specialists. The Liaison Committee on Medical Education was not founded until 1942. Prior

to its founding, the AMA's Council on Medical Education and Hospitals inspected and approved medical schools and approved hospitals for internships. The AAMC approved medical schools for membership in the AAMC. Leaders of the AAMC and the AMA, in an effort to avoid duplication of effort, assure quality of training, protect medical students from the wartime draft, and counteract pressures to shorten medical training, established the LCME at a meeting in Chicago. The LCME established and has maintained the standard that made psychiatric education and rotations one of the required experiences for all medical students. This standard has been critical to maintaining an appropriate place for psychiatric education in the medical school curriculum when it has been challenged on intellectual, and more recently economic grounds.

PSYCHIATRIC EDUCATION TODAY

Resident Recruitment and the NRMP Match

Postgraduate psychiatric education today begins with what everyone calls "the Match." Prior to the establishment of The National Resident Matching Program (NRMP), enrollment in a residency training program depended on individual negotiations between graduating medical students and the training program's director. Students were often pressured to accept a position in a program immediately, frequently before they had had an opportunity to interview for and consider other programs. The NRMP was established in 1952 in an effort to prevent these abuses and provide a level playing field for all of the programs and applicants. The NRMP is governed by a board of directors, with nominees from the ABMS, the AMA, the AAMC, the AHA, the CMSS, the AAMC Organization of Student Representatives, the American Medical Student Association, the AMA Medical Student Section, and the Consortium of Medical Student Organizations.

Each year, the recruitment process, in anticipation of the NRMP match day in March, preoccupies most GPE program directors, senior medical students, and many current residents who are very interested in the younger colleagues who might join them in their residency program. Selecting and recruiting strong candidates for residency training is a critical task for departments. Successful recruitment benefits the department's training program, its future faculty, and the communities which depend on the department for the services of its graduates. Ideal candidates, like those described by Johann Christian Reil in 1803, are the subject of intense competition between training programs.

"Perspicacity, a talent for observation, intelligence, good will, persistence, patience, experience, an imposing physique, and a countenance that commands respect. All those qualities required for the cure of the insane are so rare as to make the staffing of institutions

difficult."

It is not much of an exaggeration to say that good residents make good teachers and good training programs. Bright, enthusiastic, conscientious residents are rewarding to teach and contribute to a stimulating scholarly environment. Unfortunately, residents who display the opposite characteristics can cast a pall over the entire training program.

The NRMP has specific guidelines that govern the interview and recruitment process for residency applicants, and all GPE programs and medical students must agree to comply with them in order to participate in the NRMP match. The NRMP has established a uniform date of appointment for graduate medical education positions and a system for matching medical students and residency training programs that depends on lists that rank program preference (by the applicants) and applicant preference (by the programs). These lists are computer matched to produce the best possible outcome for both parties. Each year, about 18,000 US medical students participating in the NRMP match and about 30,000 "independent" applicants apply for approximately 30,000 positions in over 4,000 graduate medical education programs across the United States. "Independent" applicants include former graduates of US medical schools, US osteopathic students, Canadian students, and graduates of international medical schools.

Table 59.1–8.
PGY-1 Match Rates

			Percent Matched		
Year	Positions	Total Applicants	US Seniors	IMGs	All
1990	20,101	19,004	93.3	59.7	83.7
1995	20,751	22,936	92.7	50.5	77.9
2000	20,598	25,056	93.9	38.5	73.4
2005	21,454	25,348	93.7	55.6	78.0
2010	22,809	30,543	93.3	39.8	71.2
2011	23,418	30,589	94.1	40.9	73.2
2012	24,006	31,355	95.1	40.6	73.1
2013	26,138	34,355	93.7	47.0	73.5
2014	26,678	34,270	94.4	49.6	75.0

Since 1995, the number of applicants for these positions has substantially exceeded the number of positions. This trend has accelerated over the last 10 years with the inclusion of many more US IMGs and Osteopathic medical students in the match. This has produced substantial anxiety on the part of today's applicants regarding their likelihood of obtaining a residency position. In fact, the probability of a graduate of a US institution succeeding in the Match has changed little over the last 25 years. About 94 percent of these graduates will match with a program they have

ranked. Other candidates have cause for considerable anxiety. They will succeed in obtaining a position less than 50 percent of the time. Most residency training programs in General Psychiatry fill their positions in the match. The proportion of residents from US medical schools is similar to what is found in Internal Medicine and Family Medicine residency training programs (Tables 59.1–8 and 59.1–9).

Funding of Residency Positions

The residency positions for which these candidates are competing are funded by a variety of means. The federal government is by far the biggest sponsor of training stipends through two means. First, teaching hospitals receive direct and indirect graduate medical education funding (DME and IME) directly from CMS. These payments are based on a complicated formula that takes into account the percentage of Medicare patients treated at the facility, the historical number of positions when the program was started, and regional cost variations. Many hospitals and universities have far more residents than are approved and funded under the "cap" allowed by this formula, and there are dramatic differences in reimbursement by region. The second mechanism is funds provided by the Veteran's Administration for residency positions, which in many programs provide a substantial number of stipends. State mental health divisions, community mental health centers, private hospitals, and other organizations also fund some residency positions. All of these organizations expect clinical services in return for the funds they provide. It is often very challenging for a program director to construct a program which satisfies the clinical service demands of the funders and provides the educational program required by the PRs of the ACGME PRRC.

Table 59.1–9.
Postgraduate Year 1 Psychiatry Positions Offered and Filled in the National Resident Matching Program (NRMP) Match (2010–2014)

Year	NRMP Positions Offered	Total Positions Filled (%)	Positions Filled by US Graduates (%)
2010	1,091	1,075 (98.5)	670 (61.4)
2011	1,097	1,068 (97.4)	640 (58.3)
2012	1,117	1,080 (96.7)	616 (55.1)
2013	1,297	1,282 (98.8)	681 (53.1)
2014	1,322	1,291 (97.6)	685 (51.8)

National Resident Matching Program, Results and Data: *2014 Main Residency Match®. National Resident Matching Program*, Washington, DC; 2014.

Program Requirements for Psychiatric Training

The ACGME's process for determining the requirements that programs

must meet involves obtaining extensive input from the field, revisions by the RRC followed by more feedback from the field, and subsequent revisions. The ACGME Committee on Requirements reviews the PRs and has the authority to withhold final approval. Increasingly the ACGME has sought to standardize the language and content of the PRs across specialties so that much of what is found in the psychiatric and psychiatric subspecialty PRs is the same as in other specialties. The PPRs cover the issues of institutional sponsorship and resources, qualifications of the program director and faculty, creation of an academic environment, necessary clinical experiences, mandatory educational curriculum, work conditions, evaluation of residents, evaluation of the program and faculty, and specific outcomes in the language of the ACGME competencies. Recently, the process of evaluation residents has become much more systematic with the development of the milestones (discussed later in this chapter).

Institutional Requirements. Institutions must make a substantial investment in their residency training programs in order to become ACGME-accredited. Institutional accreditation is required before any residency training program can be approved. The institution must appoint a Designated Institutional Official (DIO) who has authority and responsibility for the oversight and administration of the institution's ACGME accredited programs, and is ultimately responsible for compliance with the ACGME Institutional, Common, and Specialty Program Requirements. The institution must also establish a Graduate Medical Education Committee (GMEC) which conducts annual program reviews on all residency training programs and in general is responsible for the maintenance of a safe and effective learning environment. Many other aspects of institutional investment are mandated, including sections on resident contracts, work environment, learning environment, and responsiveness to ACGME oversight. The ACGME has recently started a new program of institutional review entitled Clinical Learning Environment Review (CLER). The purpose of this review is to assess the graduate medical education (GME) learning environment of each sponsoring institution and its participating sites. A site visit is conducted during which the site visitors assess the institution through direct observation and interviews of residents and other personnel. This program duplicates much of the work of the Joint Commission on Hospital Accreditation (JCAHO).

The PPRs further elaborate what is required to conduct a psychiatric residency training program. The subspecialty PPRs are similar and will not be separately covered. The requirements include special provisions addressing the physical facilities which are necessary for psychiatric training. Adequate office space for interviewing patients and equipment for

recording sessions with patients are examples of these types of requirements. The types, length, and supervision of clinical experience specific to psychiatry are detailed in the PPRs, as are certain aspects of the ACGME competencies.

The Program Director. The PPRs specify that the residency training program director (PD) must be an ABPN-certified psychiatrist with sufficient administrative and educational experience to take on the challenging task of leading the program. The PD is expected to have full authority over the program and to have sufficient tenure in the position to be effective. Core duties of the PD are mandated in the PRs. These include the selection of training sites and development of program letters of agreement (PLAs) between external sites and the sponsoring institution. The PD must appoint a site coordinator and faculty for each rotation. The PD is responsible for conducting regular evaluations of the rotations and of the faculty and for removing faculty from their teaching roles if necessary. The PD must select residents for the program, evaluate their performance, monitor their stress and provide for twice-yearly formal evaluations. Compliance with ACGME and institutional policies and procedures, particularly those concerning limitations of duty hours and grievance procedures, is required. A summative evaluation at the end of residency training is required and the resident's rotations and educational experiences must be conveyed to the ABPN at the end of residency to assure eligibility to take the certification examination. The PRRC has, not surprisingly, required that the institution fund and the PD devote at least one half of their efforts to the training program. Success in this highly technical and demanding position requires at least this much effort.

The Faculty. Psychiatric faculty in an RTP must also be certified by the ABPN, have a strong interest in education, and devote sufficient time to fulfill their supervisory and educational duties. Other teaching physicians should be certified by their specialty board. Nonphysician faculty must have completed training, certification, and licensure appropriate to their profession. Maintenance of an academic, scholarly environment is critical to the success of the educational program. All of the faculty should regularly participate in clinical teaching rounds, conferences, grand rounds, journal clubs, and other activities common to the academic setting. Some of the faculty must be engaged in research activities. Demonstration of the faculty's scholarship may be through acquisition of peer-reviewed funding, publication of peer-reviewed articles, other publications, presentations at regional and national meetings, and participation and leadership in regional and national professional organizations.

Clinical Experiences. The clinical experiences that psychiatric residents

must have during residency are listed in Table 59.1–10. Note that the first year of residency, the internship, must include 4 months of experience in a primary care setting. The purpose and format of the primary care experience are subjects of considerable debate in the field at this time. Currently, most residents complete these experiences on inpatient Internal Medicine wards caring for very ill patients for a brief period of time. These clinical situations are unlikely to be encountered regularly in practice by most psychiatrists. They do, however, have a formative effect on the resident's identity as a physician. Experience in ambulatory or integrated care settings may prove to be more valuable to the psychiatrists of the future.

Table 59.1–10.
Required Graduate Psychiatric Education Clinical Experiences

Experience	Months Required
Primary care	4
Neurology	2
Inpatient psychiatry	6–16
Outpatient psychiatry	12
Child and adolescent psychiatry	2
Geriatric psychiatry	1
Addiction psychiatry	1
Consultation/liaison psychiatry	2
Forensic psychiatry	—[a]
Emergency psychiatry	—[a]
Community psychiatry	—[a]

[a]Required experiences without specific time requirements.

Most programs include the required 2 months of experience in neurology in the first year, leaving 6 months which is usually comprised of required psychiatric rotations. Inpatient experiences must include at least 6 months with direct responsibility for the care of acutely ill general psychiatric patients. A wide variety of settings are acceptable for this experience including traditional inpatient wards, research wards, and day treatment programs. The outpatient requirement of a year must include longitudinal treatment of patients with multiple modalities. Some of the resident's patients must be seen weekly under supervision for psychotherapeutic treatment. Forensic, emergency, and community psychiatry experiences are not required to have a minimum duration but must include certain key experiences. For example, in their emergency psychiatry experience, residents must evaluate, manage, and triage acutely ill psychiatric patients. Elective opportunities that enhance and specialize residents training experiences must be available. These are usually provided in the more senior years of the residency (Table 59.1–10).

A unique aspect of the PPRs is the ability to double-count certain experiences for the completion of both the general and child and adolescent residencies. This enables residents to complete training in general and child psychiatry in only 5 years. One month of child neurology, pediatric consultation, addiction psychiatry, and 20 percent of the resident's ambulatory psychiatry experience may be double-credited and the experiences in community and forensic psychiatry can count for both programs. This feature, recognized by the ABPN as a legitimate path to eligibility for board certification, has been important to the successful recruitment of residents into child and adolescent training programs.

Didactic Curriculum. A formal didactic curriculum must be provided for the residents during training. The PD is expected to establish goals and objectives for the curricular components that are couched in terms of the ACGME competencies. These have now been considerably elaborated and are covered in the section below entitled Psychiatric Milestones. The curriculum must be designed in a progressive manner that enables resident progression in their knowledge of the field. It is not acceptable to provide the same seminar material to all of the residents. The time for the didactics must be protected from clinical duties so that resident attendance is not compromised. The program must document that resident attend at least 70 percent of these didactic sessions.

Evaluation of Residents. Due to recent radical changes in the process of program evaluation discussed below the process of resident evaluation in training programs has drastically changed. The ACGME now requires programs to report on each resident's progress twice yearly. A requirement for a Core Competencies Committee (CCC) has been established. The PD is expected to appoint the CCC which must consist of at least three faculty members. Residents may not be members of the CCC. The committee must utilize evaluations provided by clinical supervisors and other personnel on rotations, test scores from the PRITE examination, Clinical Skills Examinations (CSEs), and other evaluative methods. The CCC must rate each resident on the milestones semiannually and these ratings must be reported to the ACGME. These evaluations are developmental in approach and indicate areas of strength and weakness. The resident, PD, and faculty are expected to work collaboratively to improve resident performance in areas of relative weakness.

Summative evaluations at the end of residency must be recorded so that there is a record of the resident's completion of all required rotations, attainment of sufficient competency on the ACGME competencies to enable independent practice and ethical and professional behavior. Uniquely in psychiatric (and neurological) training, successful completion of clinical examinations, in which the resident interviews a patient while attaining

and maintain rapport, conducts a mental status examination, and presents the data collected in an organized fashion, is required by the ABPN for eligibility to take the board examination. Residents must successfully complete at least three such evaluations during residency. This in-residency oral examination was established by the ABPN as a requirement for eligibility to take the certification examination and is in the PPRs. It replaces Part II of the old board examination. The residency program must provide all of this data to the ABPN through the PreCert system to ensure the eligibility of the resident to take the board examination.

The Psychiatry Milestones

In 1999, the ACGME and ABMS, realizing that medical practice required more than medical knowledge and skill in patient care and procedures established additional competencies which physicians are expected to attain during the course of their training. These competencies should be regarded as desired outcomes that medical training produces. The six Core Competencies are Patient Care, Medical Knowledge, Practice-Based Learning and Improvement, Interpersonal and Communication Skills, Systems-Based Practice, and Professionalism. Requirements for inclusion of training in these areas are included in all of the ACGME PRs and describe behaviors, attributes, or outcomes that must be demonstrated by residents as they advance in training. For example, in the competency of Professionalism: "Residents must demonstrate a commitment to carrying out professional responsibilities and an adherence to ethical principles. Residents are expected to demonstrate: compassion, integrity, and respect for others; responsiveness to patient needs that supersedes self-interest; respect for patient privacy and autonomy; accountability to patients, society and the profession; sensitivity and responsiveness to a diverse patient population, including but not limited to diversity in gender, age, culture, race, religion, disabilities, and sexual orientation; and (specialty-specific outcomes)."

The milestones for residency training represent the next iteration in specifying the outcomes rather than the processes of residency training. Detailed accounts of the development of the ACGME milestones project have been published in the *Journal of Graduate Medical Education* and the *New England Journal of Medicine*. In brief, public and political concern about the outcomes of medical training resulted in call for reform from the Institute of Medicine (2000), Council on Graduate Medical Education (2007), Josiah Macy Jr. Foundation (2011), and Medicare Payment Advisory Commission (MedPAC) reports (2011). The ACGME response to this crisis in 2008 was to require the creation of milestones describing resident progress in the Core Competencies for all specialties. This project was conceived as a method for evaluation of individual resident progress that would be uniformly applied across the United States

and as a method to evaluate the efficacy of residency programs in producing the desired outcomes.

Table 59.1–11.
Authors of the ACGME Psychiatry Milestones

Working Group	Advisory Group
Chris Thomas, M.D., Chair	Chris Thomas, M.D., Chair
Sheldon Benjamin, M.D.	Timothy Brigham, M.Div., Ph.D.
Adrienne L. Bentman, M.D.	Carol A. Bernstein, M.D.
Robert Boland, M.D.	Beth Ann Brooks, M.D.
Deborah S. Cowley, M.D.	Larry R. Faulkner, M.D.
Jeffrey Hunt, M.D., M.S.	Deborah Hales, M.D.
George A. Keepers, M.D.	Victor I. Reus, M.D.
Louise King, M.S.	Richard F. Summers, M.D.
Gail H. Manos, M.D.	
Donald E. Rosen, M.D.	
Kathy M. Sanders, M.D.	
Mark E. Servis, M.D.	
Kallie Shaw, M.D.	
Susan Swing, Ph.D.	
Alik Widge, M.D., Ph.D.	

Internal Medicine, Pediatrics and Surgery were the first specialties to begin milestone development in 2009. In 2011, psychiatry and the other remaining specialties began development with the goal of completion of all specialty milestones by 2014. Representatives from the ABPN, APA, AADPRT, and ACGME selected individuals to participate in milestone development for psychiatry, and the Working Group and the Advisory Group were formed (Table 59.1–11).

The Working Group was charged with milestone development and began meeting in December of 2011. The Advisory Group reviewed this work and suggested potential improvements. The development of milestones required five meetings and numerous conference calls over the course of the next 2 years. Extensive feedback from the field was sought as part of this process and was incorporated into the final documents. Details of the Psychiatry Milestone development process have been published in Academic Psychiatry.

The milestones are intended to allow faculty to rate the development of a resident's skills, ability, knowledge, attitudes, and behavior in achieving the ACGME competencies during training. They are developmental in nature. Basic skills in an area are expected at the beginning of training. Residents are expected to enhance these skills and, by the end of residency, attain a level of accomplishment that will allow independent practice. The milestones are organized for the purpose of rating resident accomplishment into five levels. Level 1 corresponds to what is expected of

a beginning resident. As residents complete their rotations and educational experiences, they are expected to attain Levels 2 and 3 as intermediate stages of accomplishment. It is important to note that the levels in the milestones do not correspond to PGY levels. Residents will progress at different rates within a program and programs differ in what experiences they offer residents in each year of residency training. By completion of residency, residents will on average have progressed to Level 4 on most milestones. This accomplishment, however, is not sufficient for graduation. The milestones do not measure every quality that is important and the PD must make an independent judgment on the readiness of the graduating resident for independent practice (Table 59.1–12).

Table 59.1–12.
Description of the ACGME Milestone Levels

Has not Achieved Level 1: The resident does not demonstrate the milestones expected of an incoming resident.

Level 1: The resident demonstrates milestones expected of an incoming resident.

Level 2: The resident is advancing and demonstrates additional milestones, but is not yet performing at a mid-residency level.

Level 3: The resident continues to advance and demonstrate additional milestones; the resident demonstrates the majority of milestones targeted for residency in this sub-competency.

Level 4: The resident has advanced so that he or she now substantially demonstrates the milestones targeted for residency. This level is designed as the graduation *target* and does not represent a graduation *requirement.* Making decisions about readiness for graduation is the purview of the residency program director.

Level 5: The resident has advanced beyond performance targets set for residency and is demonstrating "aspirational" goals, which might describe the performance of someone who has been in practice for several years. It is expected that only a few exceptional residents will reach this level.

In the milestones, each of the six ACGME Competencies contains sub-competencies that describe specialty specific elements of the general Core Competency. In the Psychiatry milestones there are 22 sub-competencies. There are an additional 66 subheadings in the milestones which are called "threads" (Fig. 59.1–1).

Sub-competencies describe the progression of specific skills, knowledge, or attributes in a single area such as psychiatric interviewing. The underlying milestones are organized into threads which describe the component skills of the sub-competency and the progression of these skills across the levels (Fig. 59.1–2).

The psychiatry milestones, sub-competencies, and threads are presented in Table 59.1–13.

Milestones for each of the subspecialties of psychiatry were completed after work on the general psychiatry milestones was finished. This work utilized the experience of the original milestone committee members who participated and chaired the subspecialty development committees to rapidly complete this task. A schematic outline of the subspecialty

milestones was created to assist the committees in their task. Committee members met for a single face-to-face working session at the ACGME headquarters in Chicago. The Advisory Committee reviewed the work accomplished at the meeting and provided suggestions. The Working Group subsequently accomplished the revisions through conference calls and online participation (Table 59.1–14).

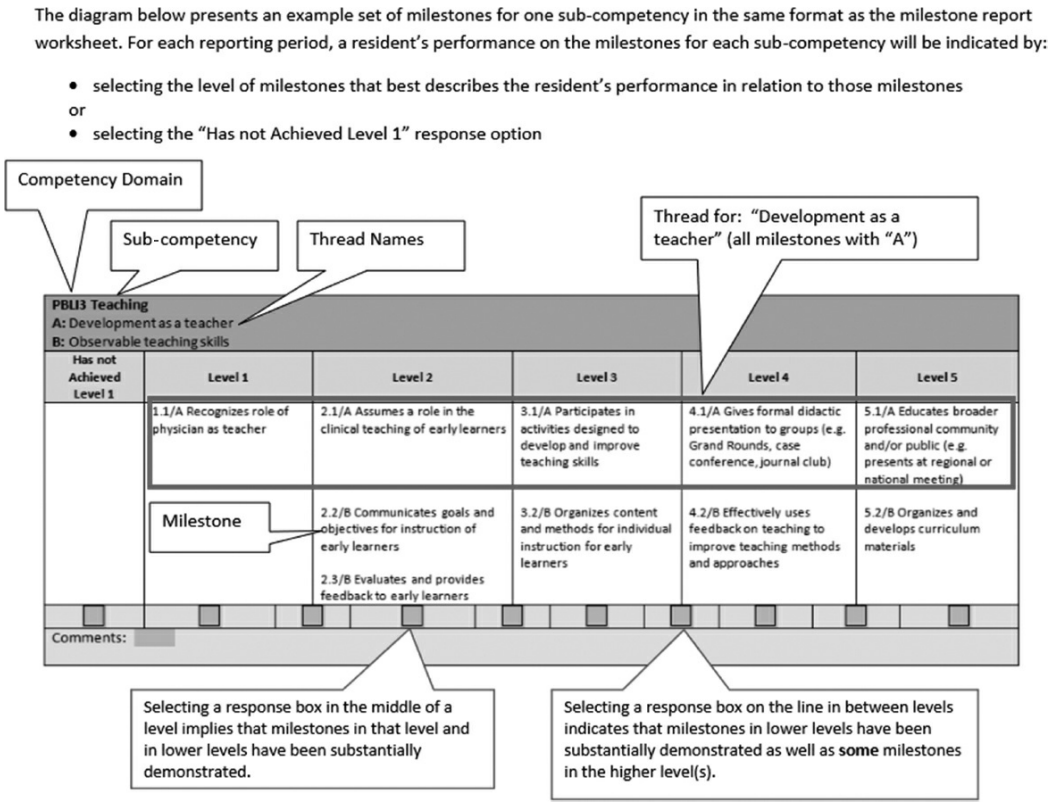

FIGURE 59.1–1. Format of the psychiatric milestones. (From The American Board of Psychiatry and Neurology. *The Accreditation Council for Graduate Medical Education and the Psychiatric Milestone Project*. July 2015.)

PC1. Psychiatric Evaluation A: General interview skills B: Collateral information gathering and use C: Safety assessment D: Use of clinician's emotional response					
Has not Achieved Level 1	Level 1	Level 2	Level 3	Level 4	Level 5
	1.1/A Obtains general medical and psychiatric history and completes a mental status examination	2.1/A Acquires efficient, accurate, and relevant history customized to the patient's complaints 2.2/A Performs a targeted examination, including neurological examination, relevant to the patient's complaints	3.1/A Consistently obtains complete, accurate, and relevant history 3.2/A Performs efficient interview and examination with flexibility appropriate to the clinical setting and workload demands	4.1/A Routinely identifies subtle and unusual findings	5.1/A Serves as a role model for gathering subtle and reliable information from the patient 5.2/A, B Teaches and supervises other learners in clinical evaluation
	1.2/B Obtains relevant collateral information from secondary sources	2.3/B Obtains information that is sensitive and not readily offered by the patient	3.3/B Selects laboratory and diagnostic tests appropriate to the clinical presentation 3.4/B Uses hypothesis-driven information gathering techniques[2]	4.2/B Follows clues to identify relevant historical findings in complex clinical situations and unfamiliar circumstances	
	1.3/C Screens for patient safety, including suicidal and homicidal ideation	2.4/C Assesses patient safety, including suicidal and homicidal ideation 2.5/D Recognizes that the clinician's emotional responses have diagnostic value[1]		4.3/D Begins to use the clinician's emotional responses to the patient as a diagnostic tool	
☐	☐	☐	☐	☐	☐

The Milestones are a product of the Psychiatry Milestone Project, a joint initiative of the Accreditation Council for Graduate Medical Education and the American Board of Psychiatry and Neurology.

FIGURE 59.1–2. Example of an actual sub-competency and its threads from the ACGME psychiatry milestones. Note how the threads describe progression of skills across the levels. (From The American Board of Psychiatry and Neurology. *The Accreditation Council for Graduate Medical Education and the Psychiatric Milestone Project*. July 2015.)

The psychiatry milestones represent a major paradigm shift in the overall approach that programs take to resident training and assessment because of their focus on outcomes and their specificity. Because of the process of development they represent a true consensus of psychiatric educators regarding what residents need to accomplish and what programs need to teach. This shift in emphasis is not, however, without its critics. They have pointed out correctly that there is no evidence that utilization of the milestones will actually produce better psychiatrists. Detractors advocate resident assessment methods that are based on actual patient outcomes. The critics would prefer that there be no change in educational methods until prospective studies establish that there will be a positive outcome. Acceding to such objections would essentially guarantee that postgraduate education remain frozen in its current state. The ACGME has instead adopted a program based on modern quality improvement methods, essentially a PDSA cycle.

Assessment of Residency Training Programs

The milestones were not only conceptualized as a method for evaluation of resident progress but are also intended to be used eventually as a method of programs evaluation, and there have been many other changes to the

accreditation process.

Incorporation of Osteopathic Training Programs. The ACGME is now responsible for the accreditation of all residency training programs in the United States. In 2014, the American Osteopathic Association (AOA) entered into an agreement with the ACGME to consolidate the accreditation of osteopathic residency training programs under the ACGME. AOA accredited programs may elect to apply for ACGME accreditation until 2020. AOA specialty board certification for faculty in these programs will be acceptable to the ACGME during the transition period. The programs may have a co-director certified by AOA specialty boards as long as the other co-director is ABMS-certified. The AOA and AACOM will be integrated into the governance structure of the ACGME and will become a nominating organization for candidates for membership on the RRCs. All of the RRCs which will review AOA programs for accreditation will have osteopathic members. The AOA may withdraw from this agreement by 2020. It remains to be seen how successful the integration of the AOA programs will be (Table 59.1–15).

Table 59.1–13.
Psychiatry Milestones

Competency	Sub-competencies	Threads
Patient Care	PC1. Psychiatric Evaluation	A. General interview skills B. Collateral information gathering and use C. Safety assessment D. Use of clinician's emotional response
	PC2. Psychiatric Formulation and Differential Diagnosis	A. Organizes and summarizes findings and generates differential diagnosis. B. Identifies contributing factors and contextual features and creates a formulation.
	PC3. Treatment Planning and Management	A. Creates treatment plan. B. Manages patient crises, recognizing need for supervision when indicated. C. Monitors and revises treatment when indicated.
	PC4. Psychotherapy refers to: (1) the practice and delivery of psychotherapies, including	A. Empathy and process

	psychodynamic, cognitive–behavioral, and supportive therapies; (2) exposure to couples, family, and group therapies; and (3) integrating psychotherapy with psychopharmacology.	B. Boundaries C. The alliance and provision of psychotherapies D. Seeking and providing psychotherapy supervision
	PC5. Somatic Therapies. Includes psychopharmacology, electroconvulsive therapy (ECT), and emerging neuromodulation therapies.	A. Using psychopharmacological agents in treatment B. Education of patient about medications C. Monitoring of patient response to treatment and adjusting accordingly D. Other somatic treatments
Medical Knowledge	MK1. Development through the life cycle (including the impact of psychopathology on the trajectory of development and development on the expression of psychopathology)	A. Knowledge of human development B. Knowledge of pathological and environmental influences on development C. Incorporation of developmental concepts in understanding
	MK2. Psychopathology. Includes knowledge of diagnostic criteria, epidemiology, pathophysiology, course of illness, comorbidities, and differential diagnosis of psychiatric disorders, including substance use disorders and presentation of psychiatric disorders across the life cycle and in diverse patient populations.	A. Knowledge to identify and treat psychiatric conditions B. Knowledge to assess risk and determine level of care C. Knowledge at the interface of psychiatry and the rest of medicine
	MK3. Clinical Neuroscience. Includes knowledge of neurology, neuropsychiatry, neurodiagnostic testing, and relevant neuroscience, and their application in clinical settings	A. Neurodiagnostic testing B. Neuropsychological testing C. Neuropsychiatric co-morbidity D. Neurobiology E. Applied neuroscience
	MK4. Psychotherapy refers to knowledge regarding: (1) individual psychotherapies, including but not limited to psychodynamic, cognitive–behavioral, and supportive therapies; (2) couples, family, and	A. Knowledge of psychotherapy: theories B. Knowledge of

	group therapies; and (3) integrating psychotherapy and psychopharmacology	psychotherapy: practice C. Knowledge of psychotherapy: evidence base
	MK5. Somatic therapies refers to medical knowledge of somatic therapies including psychopharmacology, ECT, and emerging somatic therapies, such as transcranial magnetic stimulation (TMS) and vagus nerve stimulation (VNS)	A. Knowledge of indications, metabolism, and mechanism of action for medications B. Knowledge of ECT and other emerging somatic treatments C. Knowledge of lab studies and measures in monitoring treatment
	MK6. Practice of Psychiatry	A. Ethics B. Regulatory compliance C. Professional development and frameworks
Systems-Based Practice	SBP1. Patient Safety and the Health Care Team	A. Medical errors and improvement activities B. Communication and patient safety C. Regulatory and educational activities related to patient safety
	SBP2. Resource Management (may include diagnostics, medications, level of care, other treatment providers, access to community assistance)	A. Costs of care and resource management
	SBP3. Community-Based Care	A. Community-based programs B. Self-help groups C. Prevention D. Recovery and rehabilitation
	SBP4. Consultation to nonpsychiatric medical providers and nonmedical systems (e.g., military, schools, businesses, forensic)	A. Distinguishes care provider roles related to consultation B. Provides care as a consultant and collaborator C. Specific consultative activities
Problem-Based Learning and	PBLI1. Development and execution of lifelong learning through constant self-evaluation,	A. Self-assessment and self-improvement

Improvement	including critical evaluation of research and clinical evidence	B. Evidence in the clinical workflow
	PBLI2. Formal practice-based quality improvement based on established and accepted methodologies.	A. Specific quality improvement project B. Quality improvement didactic knowledge
	PBLI3. Teaching	A. Development as a teacher B. Observable teaching skills
Professionalism	PROF1. Compassion, integrity, respect for others, sensitivity to diverse patient populations, adherence to ethical principles	A. Compassion, reflection, sensitivity to diversity B. Ethics
	PROF2. Accountability to self, patients, colleagues, and the profession	A. Fatigue management and work balance B. Professional behavior and participation in professional community C. Ownership of patient care
Interpersonal Communication	ICS1. Relationship development and conflict management with patients, families, colleagues, and members of the health care team	A. Relationship with patients B. Conflict management C. Team-based care
	ICS2. Information sharing and record keeping	A. Accurate and effective communication with health care team B. Effective communications with patients C. Maintaining professional boundaries in communication D. Knowledge of factors which compromise communication

The Next Accreditation System (NAS). Previously, the ACGME fulfilled its accreditation responsibilities through a system of regular review of programs by the RRC. Data were collected every 5 years for programs in good standing. The PD wrote a Program Information File (PIF) which covered every aspect of program operation in relation to the ACGME PRs. A site visitor would travel to the program and confirm the PIF with interviews with the PD, residents, and faculty. The RRC would then review the data at its meetings, issue an accreditation decision, and issue citations for areas judged out of compliance with the PRs.

The accreditation process has been drastically altered under the NAS. Implemented for seven pilot specialties (emergency medicine, internal medicine, neurological surgery, pediatrics, diagnostic radiology, and urology) in July of 2013, this system is now being used for all 26 ACGME core specialties. The NAS changes what has been described as a "biopsy" model into a program of annual data collection and review for every training program in the United States. The data sources for this review include an annual resident survey, faculty survey, and data entered by the program into the Accreditation Data Systems (ADS). For many specialties, case logs, and board certification pass rates are reviewed annually. It is anticipated that aggregate data from the milestones will be incorporated into the accreditation system after enough experience has been gained.

Table 59.1–14.
Authors of the Psychiatry Subspecialty Milestones

	Authors
Subspecialty Working Groups (Chris R. Thomas, Chair)	Jeffrey Hunt, Chair, Sandra M. Dejong, Laura Edgar, Howard Liu, Cynthia Santos
Child & Adolescent	
Geriatrics	Iqbal Ahmed, Chair, Josepha Cheon, Laura Edgar, Maria Llorente, Susan Maxner, Sandra Swantek
Forensics	Andre Stolar, Chair, Laura Edgar, Richard Frierson, Stephen Noffsinger, Charles Scott, Howard Zonana
Psychosomatic Medicine	Robert Boland, Chair, Madeleine Becker, Catherine C. Crone, Laura Edgar, James Levenson, Mark Servis
Addiction	Marian Fireman, Chair, Laura Edgar, Kyle Kampnam, Robert Ronis, Andrew J. Saxon, Jeffrey J. Wilsonk
Subspecialty Advisory Group George A. Keepers, Chair	Larry R. Faulkner, Frances R. Levin, Christopher K. Varley

The data from these sources is statistically analyzed by the ACGME and outliers are identified for review on the basis of decision rules which are formulated by the RRCs. The data, program history, and previous ACGME actions are then reviewed by ACGME staff for presentation to the RRC. An RRC member conducts a review of the program and presents recommendations at the RRC meeting. The RRC may elect to continue the program's accreditation, continue accreditation with warning, or to ask for a site visit. Programs may be placed on probation following a site visit if problems with the program a very severe. The usual outcome of these reviews is to communicate to the program and its sponsoring institution that there are Areas for Improvement (AFIs) or to cite programs for noncompliance with the PPRs. AFIs are issues that may be dealt with by the program and its associated GMEC. Citations require, in addition, communication with the ACGME concerning the program's plan to remedy the problem. Programs with citations are automatically subject to annual

review by the RRC.

This system has the potential to identify and correct problems in training programs much earlier than the previous accreditation system. Additionally, its developers believe that it will encourage a more rapid pace of program improvement. Criticism of the new system has focused on the increased workload and expense required to comply with the new requirements and on the validity of the data, particularly the resident survey. Program directors have expressed concern that residents are often confused by the questions on the resident survey and may answer the questions differently when properly oriented. Some studies do indicate that significant differences exist between the results from internal resident surveys and the ACGME resident survey.

Table 59.1–15.
Membership of the ACGME Psychiatry Residency Review Committee in 2015

Members	Staff
George A. Keepers, Chair	Louise King, Executive Director
Iqbal Ahmed	Debra Martin, Accreditation Administrator
Robert J. Boland	
Carlyle H. Chan	
Josepha A. Cheong	
Steven Paul Cuffe	
Mina Dulcan	
Larry Faulkner, Ex-Officio, ABPN	
Anne L. Glowinski	
Phillip M. Luber	
Gerald A. Maguire	
Anelle Primm, Ex-Officio, APA	
Heather E. W. Schultz, Resident Member	
Mark Servis, Vice Chair	
Andrea Stolar	
Dorothy E. Stubbe	
Richard F. Summers	

RELATION TO CHANGES IN MEDICAL EDUCATION

These changes in residency training are not occurring in isolation. The traditional system of medical education has been resistant to the incorporation of scientific developments in educational theory and performance improvement that have long been implemented in other educational and corporate settings. Future physicians can expect to be educated quite differently. Medical schools are incorporating the concepts of competency-based education and selecting students who will adapt better to the modern practices of team-based medicine. The requirement that students demonstrate competence at each stage of education is becoming universal. This will require a more flexible approach to the

medical school curriculum than currently exists. For example, students matriculate to medical school at varying degrees of readiness. Flexible preparatory curricula are needed to meet varying student needs for additional education prior to their first day of medical school. Ideally, the goal is for every student and every resident to be uniformly excellent, to deserve an A grade. The purpose of competency-based undergraduate and graduate medical education is to raise the level of performance and to reduce the variability in educational outcomes.

Following residency education and board certification, it is becoming mandatory that physicians maintain and improve their competence, incorporate new knowledge, and develop new skills throughout their careers. The ABMS and specialty board maintenance of certification programs are oriented to this goal. These developments in undergraduate, graduate, and postgraduate medical education are a veritable revolution in the way doctors will be educated. Medical educations past has been one of discrete episodes of learning with little assurance that competence has been attained or will be maintained. Its future is as a lifelong learning process with feedback mechanisms ensuring that physicians practice at a uniformly excellent level.

REFERENCES

*Accreditation Council for Graduate Medical Education. Annual Reports on accredited programs, residents on duty, programs with combined specialty tracks. Institutional Requirements, Psychiatry Program Requirements, Executive Summary of the Agreement among ACGME, AOA, and AACOM. Accessed January 2015.

Adamowski S. The ACGME and ABMS initiatives: toward the development of core competencies. In: Scheiber S, Kramer T, Adamowski S, eds. *Core Competencies for Psychiatric Practice: What Clinicians Need to Know*. Arlington, VA: American Psychiatric Publishers; 2003:43–54.

Adams MA, Willett LL, Wahi-Gurunaj W. Usefulness of the ACGME resident survey: a view from internal medicine program directors. *Am J Med*. 2014;127(4):351–355.

*American Board of Psychiatry and Neurology. Information for applicants for certification, PreCert, ABPN Directors, Reports on certification examinations. Accessed January 2015.

Balon R. The unspoken tyranny of regulatory agencies: a commentary on the ACGME Resident Survey. *Acad Psychiatry*. 2012;36:351–352.

Benjamin S, Widge A, Shaw K. Neuropsychiatry and neuroscience milestones for general psychiatry trainees. *Acad Psychiatry*. 2014;38:275–282.

Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA*. 2006;296:1094–1102.

Ensuring an Effective Physician Workforce for the United States: Recommendations for Reforming Graduate Medical Education to Meet the Needs of the Public. New York: Josiah Macy Jr. Foundation; 2011.

Fahy BN, Todd SR, Paukert JL, Johnson ML, Bass BL. How accurate is the Accreditation Council

for Graduate Medical Education (ACGME) Resident survey? Comparison between ACGME and in-house GME survey. *J Surg Educ*. 2010;67:387–392.

*Faulkner LR. Section 56.1: Graduate psychiatric education. In: Sadock BJ, Sadock VA, Ruiz P, eds. *Kaplan and Sadock's Comprehensive Textbook of Psychiatry*. Philadelphia, PA: Lippincott Williams & Wilkins; 2009:4396–4410.

Flexner A. *Medical Education in the United States and Canada: A Report to the Carnegie Foundation for the Advancement of Teaching*. Boston, MA: D. B. Updike/The Merrymount Press; 1910.

Hollender MH, ed. *The American Board of Psychiatry and Neurology: The First Fifty Years*. Deerfield, IL: American Board of Psychiatry and Neurology; 1991.

Hunt JR, Thomas CR. ACGME milestone development in general psychiatry: patient care and medical knowledge. *Acad Psychiatry*. 2014;38:261–267.

Iglehart JK. Medicare, graduate medical education, and new policy directions. *N Engl J Med*. 2008;359(6):643–650.

Kassebaum DG. Origin of the LCME, the AAMC-AMA partnership for accreditation. *Acad Med*. 1992;67:85–87.

Keepers GA. The coming revolution in medical education. *Austin J Psychiatry Behav Sci*. 2014;1(2):1006.

Mantosh D, Manring J, Satish U. The new milestones: do we need to take a step back to go a mile forward? *Acad Psychiatry*. 2015;39(2):147–150.

Nasca TJ. The CEO's first column: the next step in the outcomes-based accreditation project. ACGME Bulletin; 2008.

*Nasca TJ, Philbert I, Brigham T, Flynn TC. The next GME accreditation system-rationale and benefits. *N Engl J Med*. 2012;366(11):1051–1056.

National Resident Matching Program: Match results by state, specialty and applicant type. Accessed January 2015.

Norman G, Norcini J, Bordage G. Competency-based education: milestones or millstones? *J Grad Med Ed*. 2014;6(1):1–6.

Reil JC. *Rhapsodieen über die anwendung der psychischen curmethode auf geisteszerrüttungen*. Haelle, Germany: Curtschen Buchhandlung; 1803:19.

Rosen DE. The development of outpatient milestones. *Acad Psychiatry*. 2014;38:283–287.

Sanders K, Servis M, Boland R. The four general competencies. *Acad Psychiatry*. 2014;38:268–274.

Scott HW. *History of Portland, Oregon*. Syracuse, NY: D. Mason and Company; 1890:555–556.

Shorter E. *A History of Psychiatry: From the Era of the Asylum to the Age of Prozac*. New York: John Wiley and Sons; 1997:17.

*Swing SR, Beeson MS, Carraccio C, Coburn M, Iobst W, Selden NR, et al. Educational milestone development in the first 7 specialties to enter the next accreditation system. *J Grad Med Educ*. 2013;5:98–106.

Swing SR, Cowley DS, Bentman A. Assessing Resident Performance on the Psychiatry Milestones. *Acad Psychiatry*. 2014;38:294–302.

*Thomas CR, Keepers GA. The milestones for general psychiatry residency training. *Acad Psychiatry*. 2014;38:255–260.

▲ 59.2 Examining Psychiatrists and Other Mental Health Professionals

JAMES MORRISON, M.D., AND RODRIGO A. MUñOZ, M.D.

During the march from undergraduate training through postgraduate work and on to board certification, psychiatrists and other mental health clinicians will be examined hundreds of times on material in which they have invested thousands of hours of study and practice. That much has remained constant over the years. What's changed has been the nature of the examinations themselves. Generations ago, they were either oral or, when written, true–false tests and even essays. For many years, the ABPN was nearly alone among professional boards in retaining a live-patient component. Now, this remnant of the past has been relegated to the training programs, a move many see as stimulating improvement in the art of patient evaluation.

The oral examination excepted, the story of psychiatric testing has been increasingly one of national—actually, international—examinations. With a growing need to compare the abilities of clinicians who have trained at different institutions, not to mention the expense of producing new testing material, examinations standardized throughout the United States and Canada are now a central feature of nearly every psychiatry training program.

The uses and consequences of all of this testing are varied. For some students, examinations serve as an important spur to greater study effort. For others, they may signal an evaluation of subject matter mastery or a rite of passage or initiation. In some ways, examinations also serve the institution or fellow professionals by evaluating the effectiveness of teaching and training programs. Testing may serve society as a whole by ensuring quality of the licensees. Less altruistically, testing and its attendant certification may help restrict competition for patients among different groups of professionals.

MAJOR PSYCHIATRY EXAMINATIONS

The ABPN examination in psychiatry has been administered yearly since 1935. Almost since inception, it was given orally, even including basic science facts. By the mid-1960s, increasing numbers of candidates, an exploding volume of information, and the increasing demand for a procedure that was both comprehensive and impartial complicated the logistics of an oral examination. The logical solution was a written examination, which, during its first several decades of use, became a relatively inexpensive screening device for the vastly more complex and expensive oral examination. As of this writing, the oral examination is being phased out, even for graduates of years ago, who, going forward, will have to satisfactorily complete three live-patient clinical skills evaluations

at a training site approved by the Accreditation Council for Graduate Medical Education (ACGME). Only then will they be able to sit for the board examination, which has evolved into a computerized, entirely cognitive affair.

Cognitive Examinations

The first ABPN examinations depended, in part, on essays. Essay questions are easy and inexpensive to write, and they can plumb the depths of a candidate's knowledge. They also have significant disadvantages. Time constraints limit the number of knowledge areas that can be tapped. A candidate unfamiliar with just one of, say, five questions may start out perilously close to failure. Essay grading is time consuming and nearly impossible to do with complete impartiality. Poor handwriting or language skills (the latter a problem for some international trainees), which has nothing to do with knowledge of psychiatry, may figure too prominently in the grading process.

The demise of the essay examination and a brief flirtation with a true-false component in the early 1940s cleared the way for various types of multiple-choice questions. Several variations on these questions now form the bulk of standardized cognitive psychiatry examinations administered throughout the western world.

The ABPN cognitive examination is administered in the fall of each year at Pearson VUE centers throughout the United States. The current cognitive examination lasts eight hours and comprises 480 computer-scored questions covering both psychiatry and neurology (Table 59.2–1). Time and questions are divided equally between morning and afternoon sessions, both of which include vignette, basic psychiatry, and neurology material. The examination is scored as a whole; it is not norm-referenced; theoretically, every candidate could pass.

The move to computerized testing has introduced computer features and keyboard behaviors with which candidates must become familiar prior to any examination. These include:

- "Next" and "previous" buttons move the candidate through a sequence of questions.
- A "flag for review" button allows items to be seen again later; this function is not available for vignette items.
- A navigator displays a list of flagged questions.
- A clock shows time remaining in the session, always displayed in minutes and seconds.
- Videos may require a response before the next screen can be shown.
- Other features include the ability to highlight or strike through text, pull down a scratch pad, make a comment to use later in the examination or to alert ABPN to some aspect of a question, and a drug list (generics and

some brand names) that can be accessed at any time during the examination. (Staff members review all comments after the test has ended.)

- Using more than the allotted 50 minutes of break time will decrease the amount of examination time available.

Table 59.2–1.
Written Examination for ABPN Psychiatry Candidates

As of 2017, candidates for the initial psychiatry examination must answer 480 questions that cover the following areas. The "plus numbers" indicate relative number of questions in each subsection.

Basic concepts in psychiatry
Development through the life cycle (2+)
Behavioral and social sciences (1+)
Epidemiology and public policy (1+)
Diagnostic procedures (1+)
Clinical aspects of psychiatric disorders (3+)
Treatment of psychiatric disorders (3+)
Special topics—suicide, dangerousness, risk management, seclusion/restraints, abuse and violence, consultation, professionalism/ethics (1+)

Neurology and neurosciences
Neural development through the life cycle (1+)
Basic neurosciences (1+)
Diagnostic procedures (1+)
Clinical aspects of neuropsychiatric disorders (3+)
Treatment of neuropsychiatric disorders (3+)
Diagnostic and clinical evaluation of neurologic disorders/syndromes (3+)
Management and treatment of neurologic disorders (2+)

Clinical psychiatry
Specific psychiatric disorders (3+)
Treatment of psychiatric and neuropsychiatric disorders (3+)
Ethics (1+)
Epidemiology (1+)
Interpersonal communication skills (1+)
Systems-based practice issues (1+)
Factors affecting prognosis or complicating treatment (1+)
Special topics—suicidality, dangerousness, seclusion/restraints, risk management, psychiatric consultation, and child abuse, sexual abuse, and domestic violence (1+)

Note: More details are available at www.abpn.com.

In addition to more traditional questions of the sort discussed early in this chapter, cognitive testing now includes video vignettes: 1-to-4-minute clips from patient interviews with written historical material. A video vignette will typically engender 5 to 10 multiple-choice questions, some of which may require more than a single response—the computer software will not allow a candidate to proceed until the correct number of responses has been entered. The content of some USMLE questions may depend on the answer given to the one before; in other words, an individual investigative trail may depend to a degree on the candidate's own

responses.

At one time, the board spent inordinate amounts of time and money fighting the very few lawsuits brought by unsuccessful candidates who preferred litigation to the effort of further preparation as a means for achieving certification. Such behavior has decreased in recent years and, to date, no unsuccessful candidate's lawsuit has ever succeeded. But the drive to achieve the status of certification is powerful: forged board certificates have been discovered, and a 2002 paper claimed that up to 18 percent of physicians misrepresent their clinical credentials.

The Oral Examinations

Perhaps because it is the most verbal of all the medical sciences, for years psychiatry clung tenaciously to the oral examination format. Now, however, the responsibility for witnessing the examination of live patients has devolved upon the residency programs.

Part of the reasoning is that so-called long-form examinations (a single patient interviewed and discussed during the course of an hour) are now known to have poor psychometric properties. Their low reproducibility is partly due to the fact that they are case-specific (clinicians do not perform equally with all patients). Also, individual examiners will have differing opinions on the same clinical interaction, especially if there is no carefully constructed plan for the assessment of specific competencies (Table 59.2–2). In one study, the overall reproducibility coefficient of two long-case oral examinations was 0.39, as compared with 0.88 for single-best-answer multiple-choice questions.

Here is the bottom line: after years of trial, oral examinations have been found inconsistent and unreliable. As a result, the live-patient examination has evolved from a one-shot, do-or-die process to one that tries to assure the ability of every graduating psychiatrist to do a competent initial patient interview. This is, after all, a critical component of what mental health care professionals must do every day.

Table 59.2–2.
Accreditation Council for Graduate Medical Examination (ACGME) Competencies

1. Patient care
   a. Caring and respectful behaviors
   b. Interviewing
   c. Informed decision making
   d. Development and carrying out of patient management plans
   e. Counseling and education of patients and families
   f. Performance of procedures
      i. Routine physical examination
      ii. Medical procedures

   g. Preventive health services
   h. Working within a team
2. Medical knowledge
   a. Investigatory and analytical thinking
   b. Knowledge and application of basic sciences
3. Practice-based learning and improvemen
   a. Analyzing own practice for needed improvements
   b. Use of evidence from scientific studies
   c. Application of research and statistical methods
   d. Use of information technology
   e. Facilitating learning of others
4. Interpersonal and communication skills
   a. Creation of therapeutic relationship with patients
   b. Listening skills
5. Professionalism
   a. Respect and altruism
   b. Ethically sound practice
   c. Sensitive to cultural, age, gender, and disability issues
6. Systems-based practice
   a. Understanding interaction of individual practice with the larger system
   b. Knowledge of practice and delivery systems
   c. Practicing cost-effective care
   d. Advocating for patients within the health care system

Two Clinical Skills Evaluations

Two sets of live-patient examinations are now required of psychiatry trainees who would eventually become board certified: the ABPN's clinical skills examination (CSE) and the ACGME's clinical skills verification examination (CSV—training directors and others lament the confusion resulting from their similar names, sometimes interchanged; the authors will try to avoid these terms). These examinations are mandated as part of the training process and administered during each of the final 3 years of training to help assess a resident's progress. Because the two examinations have so much in common, this discussion will cover them both, as far as that commonality extends. Performance guidelines for the complete examination, with a plan for the discussion of diagnosis and treatment, are included in Table 59.2–3. Of course, the stakes for the CSE are relatively low—failure means only that the candidate must repeat the examination until they pass it three times.

Each of these examinations requires a 30-minute interview of a patient the examinee does not know (selected by the training director or a surrogate) and for whom only minimal information may be provided. The examination is ideally conducted by a psychiatrist examiner who doesn't know the examinee. After the interview is over and the patient departs, the examiner may offer the examinee a minute or 2 to review notes and prepare the presentation. During the remaining time (at least 15 minutes, but up to

half an hour), the examinee presents a précis of what has been learned about the patient's history and mental status evaluation, and answers questions about any aspect of the interview or presentation. The examiner evaluates the examinee's ability to interact with the patient, obtain material relevant to a full evaluation, and lay it out so that another person can understand it. Afterward, the examiner may give feedback to the trainee about overall performance during the examination process.

Nearly everyone regards the witnessed interview of an actual patient as a difficult step on the professional development road. This activity carries enough emotional baggage to make the process seem far different from the usual patient interview. It requires examinees to demonstrate how well they can establish rapport with the patient, gather information, elicit and interpret signs and symptoms, and present the information obtained to the examiner. Examinees' success depends, in part, on an effective plan for organizing these 45 to 60 minutes. Although the time allotment will vary with the individual patient, Table 59.2–4 suggests approximate times for the various portions of the interview.

The following sections will describe the complete interview and question-and-answer process, as it is used in the full ACGME examination. Each resident must perform one of these yearly; the ABPN does not specify the year of training for their administration, only that each applicant for the final board examination must have passed the examination three times with three different patients and at least two different examiners.

The Interview

A half-hour witnessed interview challenges any interviewer. In gathering data that adequately address diagnosis, the examinee must demonstrate compassion while tightly controlling the patient's verbal output. Careful control over the interview process is vital to obtain the necessary information within 30 minutes, yet establish and maintain rapport with the patient. After a few initial open-ended questions, to which the patient responds with uninterrupted speech about various areas of concern, the examinee should begin using closed-ended questions that rapidly review all pertinent portions of the standard psychiatric history and mental status examination. Loquacious patients and those who tend to wander from the point must be gently herded back onto the path. Although patients who are vague, circumstantial, hard to understand, hostile, or complicated prove more difficult to interview, examinees should try to regard them as an opportunity to demonstrate a capacity for empathy while eliciting information under difficult circumstances.

Table 59.2–3.
Performance Guidelines for a Live-Patient Examination

I. Physician–patient relationship

A. **Arrange the physical setting.** Sit face to face or across the corner of a desk, so you can vary the distance from the patient. Paranoid patients need more distance; depressed or anxious patients need more opportunity for comfort.

B. **Put the patient at ease.** Introduce yourself and the observers and briefly state the purpose of the interview. Your concern for the patient's comfort and well-being facilitates rapport.

C. **Demonstrate important personal qualities,** including compassion, empathy, responsiveness, sensitivity to variations in the patient's affect, and ability to cope with difficult situations. Consider your patient's refusal to talk or demand to leave the room not as a threat but as an opportunity to show your ability to cope. Examiners also value a candidate's awareness of personal biases, emotions, and performance, and the ability to respond to the patient without showing inappropriate surprise or making value judgments.

II. Conduct of the interview and eliciting data

A. **Establish eye contact** without staring or otherwise making the patient feel uncomfortable.

B. **Keep note taking to a bare minimum;** it can impair your ability to observe and listen empathetically.

C. **Begin with time for free expression** of the patient's problems; just 2 or 3 minutes shows your concern for the patient.

D. **Use nonverbal encouragements** to keep the patient talking; when appropriate, lean toward the patient (shows interest), or nod to indicate your desire for more information. "Yes," "I see," and "Ah!" can accomplish the same goals without interrupting the flow of information.

E. **Avoid jargon.** Long or scientific words or complex sentences can hamper communication.

F. **Redirect nonproductive speech.** If you mainly hear about baseball or the patient's relatives, gently restate what you need to learn.

G. **Request definitions** of vague terms such as "nervousness," "a long time," or "some."

H. **Follow the patient's flow,** rather than sticking with a rigid interview plan. You won't obtain all the information you'd like in 30 minutes, so demonstrate your ability to blend requests for specific information into the overall structure of the interview.

I. **Obtain as much detailed information** as you can, consistent with available time.

J. **Early on, use open-ended questions** to allow the patient scope to bring up the widest possible range of information.

K. **Later, use closed-ended questions** to obtain specific information.

L. **Touch on all the material** you will be expected to present (see Table 59.2–3).

M. **Work mental status questions into the flow** of your questions, to the greatest extent possible.

N. **Refer any specific diagnostic or treatment questions** the patient has to the treating therapist.

O. **Don't quit early.** Use every scrap of time available, until an examiner says it is time to quit. Then, with thanks for participating, immediately rise to excuse the patient.

III. Organization and presentation of data

A. **Concisely present your data,** organizing the information to support the most likely diagnosis.

B. **Draw your conclusions** from what you have seen and heard, as well as your knowledge of psychopathology and diagnosis.

IV. Phenomenology, diagnosis, and prognosis

A. Formulate a differential diagnosis based on your conclusions.

B. Evaluate these diagnostic possibilities to arrive at a working diagnosis and history.

V. Etiological, pathogenic, and therapeutic issues

A. **Discuss the possible causes,** including cultural, developmental, familial and genetic, psychological, and social factors; medical and neurological disorders; and substance use.

B. **Formulate a biopsychosocial treatment plan** for immediate and long-term management of your working diagnosis.

1. Discuss the indications, contraindications, and unwanted effects of drugs or other somatic interventions.
2. Discuss individual, group, supportive, insight-directed, behavioral, and other forms of

psychotherapy, as indicated.

3. Discuss which practical social interventions may help, such as community resources, day care, hospitalization, and residential care.

C. **Discuss possible complications,** including acting out, assaultive, suicidal, or violent behaviors; legal issues; noncompliance with treatment; or misuse of medications.

Note: Sections I–III apply to both the clinical skills examination required for board certification and the annual clinical skills verification mandated by the ACGME; Sections IV and V apply only to the latter examination.

Examinees should keep their own verbal output to a minimum; when they are talking, the patient cannot give them information. They should not discuss their own feelings, offer therapy, disagree with any aspect of the patient's treatment, or show irritation or other negative affect if the patient is slow to provide the information needed. Confrontation should be avoided at almost any cost, except when a patient's restlessness or other signs of discomfort indicate significant resistance to the entire interview process. Then, the examinee may have to confront the patient's lack of cooperation. Offering a chance to explain a sad or hostile affect might break through the resistance.

Table 59.2–4.
Time Allotment for Patient Interview

5 mins	Chief complaint and patient's free discussion of the issues
7 mins	Specific questions about diagnosis, pertinent negatives, and history of suicidal ideas, violence, alcohol, and drug use
5 mins	Medical history, review of systems, family history
5 mins	Personal and social history, evaluate character pathology
5 mins	Mental status examination
3 mins	Buffer time: pursue other issues, late-developing leads

From Morrison J, Munoz RA. *Boarding Time*. 4th ed. Washington, DC: American Psychiatric Press; 2008.

The emphasis of almost any interview will be altered by a presumptive diagnosis. A patient who is cognitively impaired may need a concrete style of interviewing that focuses on the mental status evaluation and activities during a typical day; an older patient needs more time to describe losses (of family, friends, mobility) and difficulties performing activities of daily living. Patients who have a psychosis or an anxiety or mood disorder will require special attention to the diagnostic criteria according to the *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition (DSM-5). However, every patient requires a careful review of personal and social information, including current living situation, work, family, marital status, and leisure activities.

Table 59.2–5.
Outline for Case Presentation

Identifying data
Present illness
Past history
Psychiatric history
Medical history
Review of systems
Family history
Personal and social history
Mental status evaluation
Most likely diagnosis
Differential diagnosis
Formulation
Predisposing factors (e.g., genetic)
Psychosocial stressors (e.g., abuse as child)
Stresses that precipitate present illness
Characteristic defenses—adaptive style
Proposed workup
Treatment plan
Prognosis

Presenting the Patient

In presenting just-gathered information to an examiner, organization is again the key. An examiner will expect a brief (5 minutes or so) recapitulation of pertinent facts in the case, not a verbatim retelling of the entire interview. A suggested outline is given in Table 59.2–5.

During the presentation, examinees sometimes discover, to their chagrin, that they have forgotten to obtain an important piece of information. Happily, they can recover, even from overlooking such vital data as suicidal ideation or substance use. If, during the presentation, they realize that they omitted something important, they can compensate by stating what they should have asked (with appropriate follow-up questions) and how they would use that information. Sometimes an examiner may even prompt an examinee for important information.

While presenting a patient whose principal symptoms seemed to be characterological, Dr. A had sailed right through to the formulation when an examiner coughed and asked, "Did you note any thoughts of self-destruction?" After a moment of horror-struck silence, the examinee blurted out, "You mean, like I'm having right now?"

Quickly recovering, Dr. A went on as follows: "I completely forgot to ask my patient whether she had ever tried to kill herself or ever had any thoughts to that effect. If I were seeing her in my practice, I'd immediately go out, bring her back into the office, and ask, 'Have you ever had any thoughts about harming yourself, made any suicidal attempts?' I'd want to know all the details of any attempts—methods, degree of suicidal intent, any physical harm. This is really important stuff: suicidal behavior figures in the diagnostic criteria for the mood disorders as well as borderline personality disorder. And it is quite common in a number of other conditions where it isn't a diagnostic criterion—schizophrenia, substance abuse, and somatization disorder, for example. I can't imagine forgetting to ask, except that I am pretty nervous."

At the end of the session, the examiners (this was before oral examinations were conducted at the residency level) agreed that Dr. A should pass.

Diagnosis and Treatment

For the yearly examination mandated by the ACGME, the examiner evaluates how well the examinee is able to think about diagnostic and therapeutic issues. (This portion of the examination completes the full former Part II oral examination of the ABPN. It is not part of what is currently required for board certification, but because the two examinations are often run concurrently, most residents will likely experience it with each yearly examination.)

During this portion of the presentation, examinees show how well they can manage clinical problems, set priorities, and recognize and cope with complications of therapy. They must be ready to offer a differential diagnosis, even if the patient's most likely diagnosis seems open-and-shut. They will also be expected to defend this evaluation with diagnostic criteria and critical thinking about treatment. Examiners hope to witness logical reasoning throughout the course of a clinical problem.

The formulation organizes in a brief paragraph the biological, psychological, and social factors that influence the evolution of the disease process and produce the current symptoms. This should be followed by suggestions for workup, which might include a physical examination, laboratory testing, imaging, and such simple devices as interviewing informants and obtaining records from other clinicians. Of course, all discussions of treatment should be based on a model that includes biological, psychological, and social elements.

Interactions with the examiner should be as candid as possible. Examinees should avoid using gamesmanship, trumpeting their own expertise, or ignoring examiner's interests and opinions. Respectful acceptance of other points of view suggests the ability to think flexibly. An examinee who does not understand a question should request clarification. Not only is this perfectly in order, it is the safest course to follow. Examinees who cannot answer a question are well advised to say so, then explain how they would go about finding the answer. No one can know everything, and many examiners feel that good physicians know their own limits. Examiners are advised to seek the examinee's areas of expertise and not linger in problematic areas, unless forced to do so by the exigencies of patient care.

TEST MAKING, TEST TAKING

Designed to test the extent of clinical knowledge acquired by the end of medical school, the United States Medical Licensure Examinations (USMLEs) for the National Board of Medical Examiners are administered by computer at various stages of training. Questions for the written portion of the American Board of Psychiatry and Neurology (ABPN) examination are developed using similar methods.

The USMLE psychiatry section development committee comprises eight or more members who serve staggered terms of up to 5 years. During item writing workshops, which now may take place online, each member shares with the group perhaps 50 questions. (Question writers are encouraged not to describe actual patients, who tend to be too complicated for testing of a single concept.) Criticisms are addressed in rewrite sessions that may be individual or in groups. Parallel committees and processes apply to each of the other major disciplines: internal medicine, obstetrics–gynecology, pediatrics, and surgery. USMLE will not divulge the percentage of Step II questions devoted to psychiatry.

No specific degree of difficulty is prescribed. Rather, there must be consensus that each item tests fairly the knowledge that students or graduating physicians (depending on step level) should possess before beginning the next stage of professional development. Many questions will pose problems that cross the boundaries of two disciplines—for example, the distractors might suggest that the answer could belong in the realm of either medicine or psychiatry.

Ultimately, the completed items are pretested in an actual examination. Because all examinations are now given by computer, new items can be tested at any time, along with the questions that contribute to a candidate's score. The latter are called "live" items, and they are the same for all examinations; the pretest items are distributed so that each will be answered by a representative sample of the 50,000 or more candidates who take the examination each year. Pretest items are evaluated against the live items, thereby influencing the composition of the next year's examination. The pretest also determines whether any of the distractors are left unchosen; if so, these may be discarded in favor of alternatives that will appeal to more candidates. As with other examinations, reliability is increased by more questions and by a longer completion time.

For each item is calculated a statistic called the point biserial discrimination index (PBDI), which quantifies the correlation between the item and the examinees' overall test performance. If it is positive, high-scoring candidates tend to do well on the item. A PBDI greater than 0.20 is desired. New test items may be discarded if they are answered correctly by fewer than 30 percent of examinees or have a negative PBDI. For all three steps, tutorial and sample test materials can be accessed at www.usmle.org.

Several old-style questions have been discarded because they are confusing or assess the examinee's ability to take tests rather than knowledge of the subject material. For example, the true-false question is written in terms of absolutes. Any item that involves a shade of gray requires the examinee to guess what the author had in mind. For example, "The prevalence of schizophrenia is 1 percent" might be answered differently by a clinician and an epidemiologist. The multiple true-false question is a complicated vehicle that requires a firm grip on logic as well

as knowledge of subject content. Its instructions are as follows: Mark "A" if choices 1, 2, and 3 are correct; "B" if 1 and 3 are correct; "C" if 2 and 4 are correct; "D" if only 4 is correct; or "E" if 1, 2, 3, and 4 are correct. (Example: Core symptoms of schizophrenia include [1] Delusions, [2] Disorganized thinking, [3] Hallucinations, [4] Homicidal ideas.)

For years now, two basic types of question formats have been favored on national tests: the single-best-answer multiple-choice question and the extended matching question (EMQ). Multiple-choice questions eschew simple regurgitation of isolated facts; they require that the examinee interpret and synthesize data so as to apply their knowledge to situations they may not have encountered previously. However, even when composed by professional test writers, multiple-choice questions can be interpreted ambiguously; candidates sometimes complain that subjects are confusing, controversial, or worse, trivial.

The EMQ derives from a theme—such as mood disorder—and includes a list of possible brief answers and a number of detailed lead-ins (stems). EMQs have several advantages. With a wide range of answer options, they can be readily adapted to new questions. They can be used repeatedly and scored flexibly, with greater weight given to those that are deemed more important. Overall, the EMQ is the best vehicle for discriminating knowledgeable from uninformed examinees; the five-item multiple-choice question is second best, with true-false finishing dead last.

An additional advantage of EMQs is that, when the answer is not precise, they can be written so that the candidate must select several answers, as for a differential diagnosis or unwanted medication effects. This sort of question, sometimes termed a "Pick N" item, might ask the examinee to select a subset of answers that should be considered in response to a given vignette.

Constructing a Test

1. Decide on the pass–fail criterion. A norm-referenced test (each candidate's score is reported as it relates to the entire group) is designed to produce variance and highlight differences in examinees. A criterion-referenced test (each score is given in terms of the absolute percentage correct) yields specific information about what the individual knows. Because both the USMLE and ABPN written examinations are criterion-referenced, all candidates taking either examination could theoretically pass. Ideally, multiple judges are involved in setting the standard for passing; several methods are described in a 2002 monograph.
2. Select the knowledge areas and question content to be tested. Obviously, testing should accurately reflect the material and must be nontrivial. Each area will require a certain number of items (approximately 30) to provide a reliable evaluation. Reliability is increased by greater test length and longer time allowed for the test.

3. Question content should cover problems that are common or potentially high-risk. Questions should assess knowledge likely to be needed by a clinician in the field, and avoid minutiae and disorders that are vanishingly rare. Additionally, focus on important areas of clinician knowledge and behavior, such as health and health maintenance, disease mechanisms, diagnosis, and treatment.
4. Choose the question format. Both the USMLE and the ABPN written examination rely mainly on multiple-choice questions of several types: (1) single best answer, a statement followed by several possible responses; (2) extended matching, as of a series of lettered answers to different numbered statements or vignettes; and (3) sequential multiple choice questions involving a vignette followed by several single-best-answer questions that may vary depending on the answer given to the previous question. This last has been favored in Part III of the USMLE.
5. In writing the questions, aim for clarity and brevity. Avoid vague terms such as *usually, frequently, often, commonly, rarely,* and *may* and absolutes such as *always* and *never*.
6. Questions should require the examinee to retrieve, select, organize, and apply knowledge, not just regurgitate facts. Question stems (the information that serves as background to the actual question) that contain detailed information will discriminate between examinees who score well and those who score poorly on the entire test.
7. Devise distractors. These should all be of the same type (homogeneous) and neither too hard nor too easy. A distractor so illogical that it is never selected should be discarded from future editions of the test.
8. Evaluate the new questions for ability to discriminate top from bottom candidates (i.e., have a positive PBDI). A pretest sample or previous edition of the test can help determine which candidates will answer an item correctly.
9. Discard pretest items that are often answered incorrectly by top examinees—these will have small or negative PBDIs.

ANATOMY OF A SINGLE-BEST-ANSWER TEST QUESTION

Complaining of voices in her head telling her to kill herself, a 31-year-old single woman was admitted to the hospital. Although she had had several prior admissions over the past 6 years, her clinical chart could not be found. She believed the voices she heard on and off for many years may have been the result of "interplanetary influence." An actress by training, she had not worked in her profession since the age of 24. She sustained herself with general relief and by cleaning offices. She thought she might have bumped her head several weeks ago. "But I was drinking," she said with a giggle, "and I don't remember much about it." She then denied habitual use of alcohol or other substances.

Fully oriented and coherent, the patient appeared distracted during the interview, several times gazing around the room and cocking her head to one side. She interpreted the proverb "A stitch in time saves nine" to mean that she should "sew up my ragged sleeve of carelock." Except for occasional laughter when she said something that did not seem especially funny, her affect was unvarying and she gazed straight at the examiner.

1. What diagnosis is most likely for this patient?

a. Alcohol dependence with hallucinosis
b. Bipolar mood disorder
c. Schizoaffective disorder
d. Schizophrenia
e. Subdural hematoma

2. Which medication will most likely provide rapid relief of symptoms?
   a. Amitriptyline
   b. Diazepam
   c. Divalproex
   d. Haloperidol
   e. Methylphenidate
3. Which further evaluation step is most likely to confirm the diagnosis?
   a. Blood chemistry panel and CBC
   b. Collateral history from a friend or relative
   c. Follow-up interview with the patient
   d. MRI of the patient's head
   e. Psychological testing
4. The patient's use of the term "carelock" is an example of:
   a. Inappropriate affect
   b. Neologism
   c. Paranoia
   d. Perseveration
   e. Tangentiality

Note the features of this set of questions and answers:

- ► The subject addresses a major area of concern to psychiatrists (psychosis is frequently encountered and fraught with risk).
- ► The stem is a relatively long, detailed clinical vignette.
- ► The questions pose several types of problem related to the information provided in the stem, but they are not linked (one answer does not depend on knowing prior answers). The first three questions require deduction from the facts given, not just recall of isolated facts.
- ► A knowledgeable examinee could answer correctly, even without reference to the stated choices.
- ► The fourth question is of an older and less satisfactory type: it draws on rote memory, rather than ability to assess a clinical situation and apply principles to it.
- ► The distractors are all logical and plausible (they are of the same length, format, and type of word or phrase as the correct answer). The answers (including distractors) are all relatively brief.
- ► Options are presented alphabetically so as not to favor a given letter or position in the list.
- ► The choices avoid vague terms such as *increase, decrease, stay the same,* and *none of the above*.

If desired, the same vignette could serve as the basis for additional questions about such topics as psychopathology, psychotherapy, and family history.

SAMPLE EXTENDED MATCHING QUESTION

For each patient in items 1 to 6, select from the common set of lettered choices (A through Q) the most likely diagnosis for the given chief complaint. Each choice may be selected once, more than once, or not at all.

A. Alcohol withdrawal
B. Antisocial personality disorder
C. Attention-deficit/hyperactivity disorder
D. Brief psychotic disorder
E. Female sexual interest/arousal disorder
F. Intellectual disability
G. Major depressive disorder
H. Major neurocognitive disorder due to Alzheimer's disease
I. Opioid intoxication delirium
J. Panic disorder
K. Persistent depressive disorder
L. Posttraumatic stress disorder
M. Pyromania
N. Schizoaffective disorder
O. Schizophrenia
P. Schizophreniform disorder
Q. Somatic symptom disorder

1. A mother notes that her 6-year-old son, whose developmental milestones have been delayed, has made slow progress in kindergarten.
2. A young married woman complains that she has never been able to achieve orgasm.
3. A college student smiles as she admits to a 2-month history of hearing voices.
4. A Gulf War veteran who has recently lost weight entertains suicidal ideas.
5. Increasingly, an elderly businessman forgets the appointments he has made.
6. A successful middle-aged lawyer complains that for years he has felt sad and lonely.

In the above question-and-answer set:

- Each vignette includes enough identifying data (gender, age) to present a visual picture of the patient.
- Although no vignette contains all the information needed for a definitive diagnosis, each allows the examinee to make a logical, informed choice.

The Use of Vignettes

Because they increase clinical relevance, vignettes have become a mainstay of written examinations, appearing in written or video format. As currently deployed on major examinations, they will be relatively brief (a few lines or a few minutes each), describe specific situations, and introduce several multiple-choice questions. Video vignettes may present a patient, a clinician, or both speaking together. Vignettes present common situations and diagnoses important in the practice of everyday psychiatry; esoteric disorders will not be on the menu. However, general medical conditions

and substance-related disorders are not esoteric; they may very well come up, over and over. Each vignette will include many features that help to decide diagnosis or treatment. These include gender and age, location of care, nature and duration of the chief complaint, brief recent history, and mental status findings. Thus:

A 23-year-old graduate student presents to student health, complaining that she has been depressed for "about 6 weeks." She denies any family history of mental disorder, but says that she and her boyfriend have had a "running dispute," in part about her loss of libido; last week, he moved out. Her appetite and sleep have since been diminished and she cannot focus on her studies. She has been neglecting work on her thesis, which is due next month.

Question 1. The patient's presentation in this vignette is most consistent with which of the following diagnoses? (Choose one.)

a. Acute stress disorder
b. Adjustment disorder with depressed mood
c. Chronic depressive disorder (dysthymia)
d. Major depressive disorder
e. Personality disorder

Question 2 Before the patient leaves the office, the clinician should (choose two):

a. Inquire about suicidality.
b. Telephone the boyfriend for his story.
c. Obtain a substance use history.
d. Arrange for psychological testing.
e. Discuss treatment options.

Each of these questions:

- ► Focuses on something important, rather than trivialities.
- ► Could be answered without reference to the choices.
- ► Does not add data to the stem or the options.
- ► Is relatively simple.
- ► Is phrased positively ("except" and "not" are avoided in the lead-in).

It is common practice to roll out the history piecemeal, with each brief vignette followed by one or two questions that may pertain to the cumulative history. Once a new item of information is presented, examinees are not allowed to go back to a previous question, for the obvious reason that new and important information may be revealed in later "episodes." Additional questions can be devised to cover somatic treatment, psychotherapy, issues of risk management, ethics, and quality of life. Often, different question sets allow reuse of vignettes from one examination to another.

Strategies for Surviving a Written Examination

1. Answer all questions.
2. Candidates who lack definitive knowledge should guess, unless the test subtracts points for wrong answers (rare today). A small chance of a right answer is better than no chance.
3. Watch the clock to help allot sufficient time to each section of the examination.

4. “Never” and “always” are red flags that can signal incorrect distractors.
5. “Possibly,” “sometimes,” and “maybe” may signal a correct answer.
6. Try to narrow the options by eliminating known incorrect answers.
7. Don’t look for patterns: test writers take pains not to give away correct answers by placing them at any constant point on a list.
8. Use the full time allotted. Candidates with time left at the end of the examination should check their work. Research has clearly established that examinees are far more likely to change an incorrect answer to a correct one than the reverse.

SUCCESS AND FAILURE ON EXAMINATIONS

Using the official ABPN grading card categories, examiners in the mid-1990s published a list of areas they regarded as crucial to a candidate’s success on live patient examinations in general. As a group, examinees who passed the examination overall handled each of these 18 items better than did those who failed the examination. These criteria sorted into two basic problem areas.

Content. Did the examinee…

1. Use appropriate opening and closing strategies?
2. Follow cues to new information?
3. Appropriately explore affective cues?
4. Demonstrate language and cultural sensitivity in communications?
5. Use a questioning technique that was open-ended but appropriately structured?
6. Control and direct the course of the interview into a cohesive whole?

Process. Did the examinee…

7. Obtain enough data about the chief complaint and history of the present illness?
8. Begin to sketch in relevant data concerning developmental, family, medical, and social histories?
9. Sensitively explore any history of alcohol and drug abuse?
10. Sensitively assess risk of suicide and homicide?
11. Present an organized, concise summary of important data?
12. Recount an organized, accurate mental status examination?
13. Consider emergency issues such as alcohol or drug use, suicide, and violence?
14. Present a rational, organized workup that included additional history, collateral information, and testing?
15. Choose an appropriate working diagnosis from a sufficiently broad differential for major mental disorders, personality disorders, and physical conditions?
16. Use all three dimensions of a biopsychosocial formulation?
17. Present a comprehensive treatment plan that included drugs,

psychotherapy, and criteria for hospitalization, and demonstrate knowledge of family systems and milieu principles?

3. Discuss negative and positive prognostic factors and transference and countertransference issues?

Stepwise logistic regression found four criteria that best predicted outcome. In descending order of importance, they were treatment plan (item 17), information cues (2), mental status examination (12), and interview control (6). Failing examinees scored lowest on information cues (2) and affective cues (3). Although decades of observations and anecdotes related by dozens of board examiners verify the essential truth of these findings, they have not been replicated.

Problems with Content

Issues of content can potentially involve virtually any clinical topic covered in this textbook, but examinees typically encounter difficulties with some of the following issues. Although some of the examples used in the following discussion are holdovers from the era of live-patient and videotaped Part II ABPN examinations, their lessons remain relevant today.

Inadequacies of Factual Information. Lack of facts concerning any modality of treatment or any psychiatric diagnosis should stimulate further review of relevant sections of a standard psychiatric textbook. Whereas no one expects examinees to memorize the DSM-5 criteria, a thorough understanding of their structure and basic content is vital. Indeed, if examinees can demonstrate their ability to provide safe psychiatric care, isolated knowledge deficits should not be devastating.

After interviewing a young woman with a first-break psychosis, Dr. B admitted to "complete naïveté" regarding the novel antipsychotic agent the patient was taking. "If I inherited the care of this patient, I'd look that drug up in AHFS Drug Information or PDR. And I'd also discuss it with the drug information pharmacist at our medical center." During the debriefing, the examiner specifically praised this cautious, information-seeking approach to uncharted territory.

Differential Diagnosis. For generations, the elementary clinical exercise known as the differential diagnosis has been a mainstay of medical decision making, but the stress of the witnessed interview sometimes causes examinees to narrow, rather than broaden, their focus. Some cling to a favorite diagnosis as if it were a life preserver; others appear to fear they are being tricked or trapped. However, examiners only want to find a pattern of sound diagnostic thinking that will serve future patients well.

After doing a fine interview, Dr. C made the flat statement, "The diagnosis is anxiety disorder, not otherwise specified," and then stopped. The combined efforts of two examiners could not persuade this examinee to admit that another diagnosis was possible. The senior examiner

noted, "Once many examinees get the patient to report a very well-defined psychiatric syndrome, they tend to forget that there is a differential diagnosis ... to consider. Even though the issue has been well addressed [in reading material concerning the boards], many people forget."

Most examinees are reasonably well versed in the latest diagnostic criteria, but sometimes one will mix up items from two or more criteria sets. Some overinterpret the data and make a diagnosis they cannot defend with the material at hand. An examiner will typically try to help the examinee work through such difficulties but most show little tolerance for the occasional candidate who invents criteria. Here, as everywhere else in an oral examination, it is far better to acknowledge a deficiency than to risk appearing intellectually dishonest.

Noting that a videotaped patient had several times glanced away, Dr. D hypothesized that the behavior could be caused by hallucinations. "Of course, there could have been activity off-camera that we couldn't see," the examinee volunteered. "I'd need to ask the usual sorts of questions about auditory and visual hallucinations, delusions, and so forth before I could say for sure. Would you like me to go through them?"

Dr. D was playing it safe by mentioning other reasonable interpretations. This is an excellent practice to follow in every examination—flexible thinking is the key to success.

Formulation. Even if they have done a reasonably competent job of gathering information, some examinees appear puzzled about what to do with it. They seem incapable of condensing what they have learned into the few sentences required.

After Dr. E presented the basic observations about a videotape interview, the examiner asked for a formulation of the patient. What came back was an unaltered recitation of the few minutes just witnessed. With prompting, Dr. E eventually mentioned diagnosis but seemed to recognize no opportunities for further investigation.

Later that afternoon, after viewing the same videotape, Dr. F responded, "This is a middle-aged man who was occasionally tearful and spoke slowly as he mentioned several symptoms that would support a diagnosis of depression. I didn't hear anything about previous episodes or manias, so my best [principal] diagnosis would be major depressive disorder, single episode. However, I'd want to quiz him more carefully about mania. I thought he glossed over the question about drinking, so I'd also have to know more before ruling out a substance-related mood disorder. I'd defer [a personality] diagnosis—not enough information. Contributing to his depression could be the fact that his mother died several months ago...as well as the fact that he has been worried about his job due to missing work. That would also be reflected in his global assessment of functioning score. I didn't think he looked hypothyroid so I wouldn't mention it...but with the routine labs I'd order a thyroid screen, just to be safe. My treatment plan would be biopsychosocial, and I'd consider . . ."

Further Evaluation. Examinees may appear uncertain about how to work up a patient. Some suggest laboratory tests that are inappropriate; others cannot explain how they would use the information once they'd obtained it. Many fail to mention that most useful (and inexpensive) of procedures, the collateral interview, which is especially vital when patients

are psychotic or have cognitive or personality disorders.

Treatment Recommendations. The importance of psychopharmacology to the treatment process and its emphasis in training curricula have ensured that most examinees are reasonably well prepared in this area. However, now and again the specter of treatment complications arises to haunt those who should have been better prepared to deal with potential complications of drugs they themselves have recommended. For example, unfamiliarity with the symptoms of neuroleptic malignant syndrome is surefire trouble for an examinee who has just recommended the use of haloperidol (Haldol).

Discussing treatment of a depressed, middle-aged man, Dr. G advocated trazodone (Desyrel) but failed to mention the possibility of priapism. "Try as we might to suggest that a discussion of sexual complications might be in order, we got nothing for our efforts," said Dr. G's examiner. "Now, not knowing the side effect of a particular drug isn't usually a calamity, but priapism is an infamous side effect, and it was the examinee who suggested Desyrel in the first place!"

Some examiners have reported feeling upset when examinees tell them that they "never do psychotherapy." "Despite all the good advice about addressing all forms of treatment, many examinees' formulations are completely unidimensional," one examiner reported. "They just kiss off psychotherapy by saying they refer all patients who need it to a social worker or psychologist. They seem to regard this as relieving them of all responsibility for knowing anything about interpersonal aspects of treatment. For some of us, this attitude is a nail in the coffin."

After an excellent presentation of a patient who had described many symptoms of borderline personality disorder, Dr. H would only consider behavior modification as a treatment option. The examiners repeatedly asked whether medication could be considered (the patient had even reported some depressive symptoms), but Dr. H wouldn't hear of it: "There is little solid evidence that medication has any long-range effect in borderline personality disorder." Even when one examiner postulated an extreme case in which such a patient developed severe depression, the response was the same—refusal to consider another course of action. Despite an outstanding interview, they could not justify a fully passing grade for this otherwise excellent examinee.

Rather than not knowing drugs, some examinees seem familiar with little else. The emphasis of all psychiatric training should be biopsychosocial, which means that, besides knowing dosages and side effects of the latest psychotropic medications, examinees must discuss social interventions (placement, self-help groups, family interventions, and the like) and the myriad forms of psychotherapy.

An unnecessarily uncomfortable half-hour was the fate of Dr. I, whose patient had a personality disorder with borderline features.

"For 10 minutes all this examinee would talk about was antidepressant medications," an examiner later complained. "Sure, we all saw the depressive symptoms. But we practically had to beat the examinee up to get a treatment approach that was anything but drugs."

Dr. I ultimately passed but, according to the examiner, it was "a near thing."

Despite numerous warnings that there are no pathognomonic signs in psychiatry and that no disorders have only one treatment, some examinees still pursue a single-minded approach to patients and their problems. They may quote the latest "truth" from a journal article and, when asked to explore alternative possibilities, respond with the journal title, page number, and publication year. In the words of one examiner, this is "tantamount to dying with their rights on."

Problems with the Process. As with content, issues of process can affect examinees in both the ABPN and ACGME required patient examinations.

Inadequate Rapport. From the opening moments of the patient interview, some examinees seem unable to forget their own needs and focus on those of the patient. They ignore the bits of social behavior that should put the patient at ease, such as making introductions, offering a choice of chairs, explaining the purpose of the examination, giving good eye contact, and showing empathy. As one examiner put it, "You rarely see a caring candidate who doesn't do well overall in the examination."

Rapport is especially an issue when an examinee's behavior indicates insufficient concern for the feelings or emotional state of the patient. Some examinees seem incapable of putting themselves in the patient's shoes; a rare candidate may show callous disregard for the feelings of a fellow human.

A board patient with an advanced case of acquired immune deficiency syndrome (AIDS) had lost his lover and been abandoned by his family. During the interview, he admitted that he felt depressed and often thought about death. When asked about the importance of the family's absence, Dr. J responded that it made "little difference, because the patient was dying anyway."

Whereas Dr. J's attitude was an extreme example, every examinee must demonstrate the ability to address the patient's comfort and feelings through the use of body language, eye contact, and sensitive responses to emotionally laden material.

Conduct of the Interview. Almost any interview is likely to yield some suggestions for improvement—better eye contact, fewer interruptions, a different focus or depth when following up on earlier questions. Other findings may require much more extensive remediation.

"I've got to say, I felt really uncomfortable during much of your interview," the examiner told Dr. K during a debriefing. "The patient had a story he was trying to tell, but you kept interrupting to ask questions about which medicines he had taken. And some of the important ideas he did manage to get across—his life didn't 'seem worth living,' for example—you didn't follow up."

Inadequate Control. Although examiners criticize the excessively controlled interview, in which a patient is given insufficient scope to discuss problems freely, they far more often criticize the inadequately controlled, directionless interview.

While the patient talked, Dr. L furiously scribbled notes. Actually, the interview started out well enough, with the patient discussing the anxiety symptoms that had brought him to treatment. However, soon he wandered off into a discussion of his wife's recent operation, then the problems the family had had with her surgeon's billing practices. The examinee went right along, transcribing his words with passive diligence. Although an examiner finally suggested a shift in focus, little changed, and Dr. L failed.

Examiners hope to witness an interview in which all of the important topics are covered and the examinee identifies and follows up on critical questions in a way that makes the patient feel party to a conversation rather than the subject of an interrogation.

Dr. M had to examine a woman who, for 18 months, had been a state mental hospital inpatient. Her problems included bulimia nervosa, obsessive-compulsive disorder, major depressive disorder, and multiple substance-related disorders. "And that was just the Axis I stuff," the examinee reported later.

Unable to fully evaluate any of these issues, Dr. M spent a few minutes addressing each problem before easing the patient into a discussion of the next. "My examiners seemed to recognize that I could get my arms around each of these issues sufficiently to formulate a plan for further evaluation and management. I didn't feel especially good about my interview, but I passed," he reported.

Anxiety. Nearly everyone performing a witnessed interview will feel some anxiety, but in the old Part II of the ABPN examination that could seem a one-time, do-or-die exercise in virtuosity, fear nearly incapacitated some examinees. It could reduce their ability to respond in a warm and sympathetic manner, or drive them to ignore important questions (such as hints of suicide or evidence of cognitive disorder) that needed follow up. Still others reported that their minds would "go blank" when they tried to recall even elementary didactic material.

All-consuming anxiety should not be such a problem with the new regime, where the examination is a normal part of the training process, with essentially unlimited opportunities to pass. Further, the current protocols require only one examiner; in the old Part II ABPN examination, it was two on one, with the added pressure that one might expect from facing a "panel of judges" rather than a single individual. In all likelihood, examinees who needed to use behavior therapy and drugs (such as beta blockers) for anxiety will become a distant memory.

Organization. Organizing and verbally presenting a complicated history is a skill medical students learn early in training. Unaccountably, it also lies at the core of the difficulties many psychiatrists encounter during

witnessed examinations.

> When asked to present the patient "as best you can," Dr. N encountered disaster. Instead of doing a focused presentation, this examinee attempted to recount the undigested interview, pretty much as it had just occurred. Jumping around from present illness to historical data to mental status examination, Dr. N became more and more anxious, compulsively—and futilely—shuffling through pages of scrawled notes.

It is remarkable how many examinees have mentioned that it took them a failure (or two) to learn the importance of presenting their patients "straight on through," not stopping to wait for questions, and only pausing when an examiner interrupted with a question. "In my training program, a presentation was more in the way of talking back and forth with the attendings. I learned that the boards require a presentation that is not only succinct but complete," noted one examinee.

Relationship with an Examiner. Some examinees have reported feeling that they have, in some way, run afoul of an examiner. In part, this impression may be fed by anxiety; it may also be a reaction to the neutral stance examiners must take toward examinees. Whatever the cause, some examinees find themselves arguing with examiners. Others pontificate, attempting to demonstrate special expertise or knowledge. Examiners sometimes regard such behavior as arrogant.

Lack of information is serious but, to an extent, expected. Interpersonal difficulties during the examination are quite another matter, for they suggest something inherent in the approach to people that could negatively affect the examinee's ability to care for patients. Although board examiners rarely express negative feelings about the candidates they examine, the reverse is not uncommon. Under stress, examinees are likely to react by becoming defensive and rigid. They may argue, clinging so tenaciously to a position that they cannot acknowledge, much less accept, the possibility that an examiner may be trying to suggest an alternative view of the patient. The fact that a given diagnostic or treatment option may be perfectly justifiable, even correct, does not excuse any examinee from the need to demonstrate flexible thinking.

Fatal Mistakes. Most examiners agree that a single error is unlikely to cause a CSE failure. Such an error would have to demonstrate an abysmal ignorance of facts or an overwhelming lack of concern for patient well-being—for example, suggesting that a suicidal patient could have a month's prescription of potentially lethal medication or, for some examiners, not asking a patient who uses drugs about suicide or failing to probe for a history of violence or substance abuse. For example, when interviewing a patient in a substance abuse treatment program, an examinee did not assess dangerousness.

Examiners must feel confident that the examinee can competently assess a patient for these behaviors crucial to diagnosis, to treatment, and, indeed, to life itself. Despite the logic of this principle (not to mention the publicity it has received in books, articles, and board preparation courses), some examinees still founder on a failure to follow up on clues the patient may present. These can include hints about suicide or death, tearfulness, abnormally reduced or excited activity level, and references to substance use.

> "I remember one examinee who didn't pick up on what the patient was saying," an experienced examiner commented. "The patient had mentioned drug abuse and the candidate just didn't get it. My partner and I hated to do it—this was a bright, young psychiatrist embarking on a promising career, who has since become an internationally known leader in psychiatry. But we were forced to issue a failure. The next year the candidate passed, but I still cringe when we run into one another at meetings."

PREPARING FOR THE BOARDS

How much time should examinees spend in preparation for the current ABPN certification examination? Unfortunately, data are entirely lacking. Of course, this time is highly individual: For weak examinees, long hours may not be nearly enough, whereas a strong examinee just completing training may require little additional preparation. Unhappily, it is all too often that the former learn who they are only as a result of taking the examination itself. That said, preparation for all candidates should probably include careful reading of a standard American textbook of psychiatry, preferably one that offers a study guide to provide familiarity with the sort of test question that may be asked.

Whether an examinee should also engage in a preparation course, either at home or on-site, is a matter for some conjecture. How helpful are these seemingly ubiquitous, definitely expensive review courses offered by universities and private organizations? Other than personal anecdotes and testimonials, essentially no data exist for psychiatry. One study in the field of surgery reported that formal review courses may assist the examinee at high risk of board failure, especially those who have already experienced one or more failures.

EXAMINERS

The requirement for conducting a witnessed, live-patient examination that will qualify the candidate for board certification is simple: the examiner must also be board-certified. A second requirement, which also holds for the yearly ACGME examination, is that at least two of the three required oral examinations be conducted by different examiners.

Beyond that, a couple of qualities are desired, though not mandated. One is that the two individuals involved should not have had a personal relationship with one another—it is hard, after all, to judge impartially

someone you know well. That could be a problem especially in smaller programs, where many potential examiners may already be involved in the teaching program. Of course, in cities large enough to support two or more training programs, a cross-town exchange can obviate even the appearance of partiality.

The other issue is uniformity. That is, without procedural instructions, how do examiners operationalize the task of promoting a standard of excellence in examinees? In the days of the oral boards, new examiners had to undergo a half-day training that imbued them with a uniform sense of purpose. (Even then, considerable variations in grading standards were noted.) Although there are written guidelines for evaluating performance, examiners may vary considerably in terms of what information they feed back and how they accomplish it. Now, with no formal training required, those who sit in judgment are left pretty much on their own to interpret their mission. To be sure, the numerous individual scoring points listed for both the ACGME and the ABPN examinations (Tables 59.2–6 and 59.2–7) can provide far more specific feedback than was ever available to candidates for the old-style Part II oral examination. However, examiners are also permitted informal comments that were formerly proscribed, potentially injecting a note of mentorship, even collegiality, into a relationship that was once strictly judgmental. It should come as no surprise that an ABPN task force has recommended formal training for examiners.

GRADING PROCESS

The criteria used for scoring the live-patient annual skills assessment required by the ACGME are given in Table 59.2–6. It is noteworthy that, even though the two examinations may be graded on the basis of the same patient interview, somewhat different criteria are used for the ABPN clinical skills evaluation; Table 59.2–7 presents the somewhat longer of two versions that are posted on their website. Note further the different wordings for the two sets of scoring criteria: the ABPN requires that the examinee perform at a level expected of a community psychiatrist, whereas the ACGME requires that residents demonstrate competence "for their level of education" before passing on to the next training year.

Table 59.2–6.
Grading of the Annual Clinical Skills Assessment of the ACGME

1. **Physician–Patient Relationship**

1–1. Develops rapport with patient

Excellent: Courteous, professional demeanor; Clear introduction to patient; Exhibits warmth and empathy

Good: Generally respectful; Adequate introduction; Adequate empathy

Fair: Arrogant, disrespectful, or awkward demeanor; Inadequate introduction; Lacks

empathy

Poor: Rude or inappropriate comments; No introduction or misrepresentation of the situation; Obvious anger or frustration

1–2. Responds appropriately to patient

Excellent: Responds empathically to verbal and nonverbal cues; Adjusts interview to patient's level of understanding and cultural background; Adjusts interview to new information

Good: Responds adequately to verbal and nonverbal cues; Occasional use of technical jargon; Adjusts interview to most new information

Fair: Shows minimal response to sensitive information; Minimal awareness of patient's capacity to understand or cultural background; Inflexible interviewing style; Misses important verbal and nonverbal cues

Poor: Responds with angry, abusive, or dismissive comments; Frequently loses composure; Criticizes, demeans, or condemns patient

1–3. Follows cues presented by patient

Excellent: Responds appropriately to verbal and nonverbal information; Follows up on all pertinent information; Seeks clarification of ambiguous information

Good: Misses no major verbal or nonverbal information; Generally follows up on major issues presented by the patient

Fair: Misses significant verbal and nonverbal information; Fails to ask for clarification of ambiguous information

Poor: Ignores or responds inappropriately to verbal or nonverbal cues; Grossly misinterprets verbal or nonverbal information

2. Conduct of the Interview

2–1. Listens to and facilitates the patient's narrative history

Excellent: Actively listens to the patient's story; Uses frequent, well-structured, open-ended questions to facilitate the story; Assists the patient to focus and clarify the story

Good: Passively listens to the patient's story; Occasionally uses open-ended questions to facilitate the story; Asks occasional clarifying questions

Fair: Interview consists primarily of directive, closed-ended questions; Inadequate or ineffective efforts to focus and clarify the story; Frequent interjection of assumptions or leading questions

Poor: Interview consists entirely of narrowly focused, closed-ended questions; Patient is repeatedly diverted from the narrative to check-list questions or irrelevant topics; Patient's narrative is disregarded, contradicted, or ignored

2–2. Obtains sufficient data for DSM differential diagnosis

Excellent: Assists the patient in describing the full range of symptoms and history; Explores all pertinent domains of information; Gathers adequate information for DSM checklists

Good: Allows patient to describe major symptoms and history; Explores the major domains of information; Focuses interview on DSM checklists

Fair: Limits interview to DSM checklists; Misses important domains of information; Shows little awareness or regard for DSM diagnoses; Fails to consider alternative diagnoses

Poor: Fails to gather sufficient information for major diagnosis; Misinterprets or misrepresents diagnostic information

2–3. Obtains psychiatric, medical, substance use, family, and social histories

Excellent: Assists the patient in presenting each aspect of the history; Gathers a wide range of biopsychosocial information; Maintains focus and logical progression of interview; Appears comfortable with difficult or sensitive topics

Good: Allows the patient to present an adequate range of material; Gathers adequate biopsychosocial information; Generally redirects the patient when necessary; Somewhat uncomfortable with difficult or sensitive topics

Fair: Interrupts or interferes with the patient's story; Misses important biopsychosocial information; Fails to redirect or focus a disorganized or hyperverbal patient; Avoids

difficult or sensitive topics

Poor: Ignores pertinent areas of the history; Asks cursory, disorganized, or irrelevant questions; Loses control of the interview; Responds inappropriately to difficult or sensitive topics

2–4. Screens for suicidality, homicidality, high-risk behavior, and trauma

Excellent: Approaches topic frankly, but with sensitivity and empathy; Asks questions appropriate to the context of the interview; Follows up with specific questions; Assesses specific suicide risk factors, if relevant

Good: Approaches topic somewhat awkwardly; Asks general screening questions; Follows up with 1–2 specific questions

Fair: Approaches topic with abrupt, accusatory, or incredulous manner; Asks only indirect or cursory questions; Obtains no detailed information

Poor: Fails to address suicidal or homicidal ideation; Disregards pertinent information in the history regarding patient's risk factors

2–5. Performs an adequate mental status examination

Excellent: All pertinent areas of the MSE were addressed; Appropriate areas of the MSE were integrated into other parts of the interview

Good: Most pertinent areas of the MSE were addressed; Occasional areas of the MSE were integrated into other parts of the interview

Fair: At least one essential element of the MSE was omitted; MSE elements were inadequately performed

Poor: Multiple elements of the MSE were omitted; MSE elements were ineptly performed

3. Case Presentation

3–1. Organized and accurate presentation of history

Excellent: HPI accurately reflects the patient's story; Presentation is logical, concise, and coherent; History integrates all important biopsychosocial factors; Presentation includes pertinent positive and negative findings; Presentation leads to a clear understanding of the patient

Good: HPI generally reflects the patient's story; Presentation can be followed; History includes adequate discussion of biopsychosocial factors; Presentation includes major pertinent negative findings; Presentation leads to an adequate understanding of the patient

Fair: HPI ignores or inaccurately represents the patient's story; Presentation is disorganized or chaotic; History misses important biopsychosocial factors; Presentation ignores some pertinent positive or negative findings; Presentation leads to a poor understanding of the patient

Poor: HPI distorts or misinterprets the patient's story; Presentation is incoherent or illogical; History shows no awareness of biopsychosocial issues; Presentation misinterprets or disregards pertinent positive or negative findings; Presentation is grossly inaccurate

3–2. Organized and accurate presentation of mental status findings

Excellent: All areas of the MSE are presented; Presentation is orderly, systematic, and easy to follow; Standard terminology and nomenclature are used; Findings are accurate and complete; Pertinent negative findings are included; An appropriate and accurate assessment of dangerousness is included

Good: Most areas of the MSE are presented; Presentation generally follows a standard outline; Clear and meaningful terms are used; All critical findings are included; Most important negative findings are included; An adequate assessment of dangerous is included

Fair: Several pertinent areas of the MSE are omitted; Presentation is disorganized and rambling; Ambiguous, inappropriate, or unclear terminology is used; Some critical findings are omitted or misrepresented; Important negative findings are omitted; Assessment of dangerousness is inadequate or only partially accurate

Poor: Major areas of the MSE are omitted; Presentation is incoherent and impossible to follow; Inaccurate, meaningless, or inappropriate terminology is used; Most critical findings are omitted

or misrepresented; Negative findings are not included; Assessment of dangerousness is omitted or is inaccurate

4. **Case Formulation and Differential Diagnosis**

Excellent: Comprehensive and well-integrated biopsychosocial formulation; Skillful integration of psychodynamic perspectives; Thorough conceptualization of patient's story; Well-structured and clinically appropriate differential diagnosis

Good: Good identification of biopsychosocial factors; Inclusion of appropriate psychodynamic issues; General awareness of patient's story; Adequate differential diagnosis

Fair: Omission of at least one significant biopsychosocial factor; Limited integration of psychodynamic issues; Minimal awareness of patient's story; Too narrow or too broad differential diagnosis

Poor: Inaccurate or incoherent formulation; Poor awareness of biopsychosocial factors; Minimal awareness of psychodynamic issues; Misdirected or inaccurate differential diagnosis

5. **Treatment Planning**

Excellent: Well-formulated, evidence-based, cost-effective treatment plan; Skillful coordination of somatic and psychosocial treatments; Full integration of patient, support system, and outpatient providers into treatment planning; Skillful development of a comprehensive safety plan

Good: General attention to evidence-based, cost-effective treatments; Awareness of available somatic and psychosocial treatments; Inclusion of patient and family preferences in treatment planning; Adequate attention to safety planning

Fair: Little awareness of efficacy or cost data in treatment planning; Limited attention to integration of somatic and psychosocial treatments; Cursory attention to patient and family treatment preferences; Minimal attention to safety planning

Poor: Grossly inadequate or inappropriate treatment; Narrow, rigid, or one-dimensional treatment plan; Lack of attention to or awareness of patient and family treatment preferences; Absence of appropriate safety planning

For each of 12 sections, score as follows: Excellent 7 or 8; Good 5 or 6; Fair 3 or 4; Poor 1 or 2. An acceptable overall score would be a mean of 5 or greater

Material adapted from American Board of Psychiatry and Neurology website.

Table 59.2–7. Scoring the ABPN Clinical Skills Evaluation

Physician–Patient Relationship (Overall)	Unacceptable	Acceptable
1. Opening and closing	Awkward strategies	Appropriate strategies
2. Informational cues	Ignored leads	Followed leads
3. Affective cues	Ignored	Explored appropriately
4. Communication style and rapport	Insensitivity interfered with data collection	Adequate language sensitivity
5. Questioning techniques	Abrupt and forced-choice questions	Open-ended but appropriately structured
6. Control and direction of interview	Scattered and fragmented questions	Developed, cohesive interview
Psychiatric Interview (Overall)	**Unacceptable**	**Acceptable**
7. Presenting problems and HPI	Inadequate obtained or too vague	Obtained adequate data
8a. Past psychiatric history	Ignored major issues	Gathered relevant data in at least brief form
8b. Family history	Ignored major issues	Gathered relevant data in at least brief form

8c. Medical history	Ignored major issues	Gathered relevant data in at least brief form
8d. Social/educational/occupational history	Ignored major issues	Gathered relevant data in at least brief form
8e. Developmental history	Ignored major issues	Gathered relevant data in at least brief form
9. History of drug and alcohol abuse	Ignored or too limited	Sensitively gathered
10. Assessment of suicide risk	Ignored or too limited	Sensitively gathered
11. Assessment of homicidal risk	Ignored or too limited	Sensitively gathered
12. Mental status examination	Omitted or too limited	Organized approach and performed appropriately
Case Presentation (Overall)	**Unacceptable**	**Acceptable**
13. Summary of important data	Disorganized	Presented cohesively and coherently
14. Mental status examination	Incomplete	Accurately summarized
15a. Emergency issues: suicide	Ignored	Considered
15b. Emergency issues: violence/abuse	Ignored	Considered
15c. Emergency issues: drugs/alcohol	Ignored	Considered
16. Recognition of need for additional history and collateral information	Absent or no rationale	Appropriate

In the numbered subsections, unacceptable behaviors are scored from 1 to 4; acceptable from 5 to 8.

The entire score sheet, formatted for computer entry of data, is available at http://www.abpn.com/access-residency-info/residency-training-information/psychiatry/

SUBSPECIALTY PSYCHIATRY BOARD EXAMINATIONS

In addition to the board examinations in general psychiatry, the ABPN offers certificates in 10 subspecialties. The oldest of these, dating from 1959, is the child and adolescent psychiatry examination. The most recent, brain injury medicine, was initiated in 2011.

Child and Adolescent Psychiatry

To sit for this examination, applicants must first be board-certified in adult psychiatry and have completed training in a child and adolescent psychiatry program approved by the ACGME. This computerized examination is administered once a year at Pearson VUE centers and lasts just over 5 hours. Its 280 multiple-choice questions assess core knowledge in the areas of child and adolescent development, consultation, developmental assessment/classification, psychopathology/classification, treatment, prevention, issues in practice, and biological and clinical science. Approximately one-quarter of the questions are based on clinical vignettes (written or video) of actual patients. The oral examination, long a staple of the child and adolescent psychiatry board process, has been eliminated. As with other psychiatry boards, certification is good for 10 years and subject to the usual maintenance of certification procedures (vide infra).

Other Subspecialty Certificates

By examination, physicians who are board diplomates in general psychiatry can earn certificates in other subspecialty fields added since 1991. Each requires one additional year of training (after the fourth postgraduate year) in an institution approved by the ACGME; training may usually be completed on a half-time basis. Each certificate requires a written, multiple-choice examination that in most subspecialties comprises 200 questions. Each is a 10-year certificate with the usual requirements for maintenance of certification through subsequent education and examination. Initial board certification in medical specialties other than psychiatry can qualify examinees for certification in those subspecialties marked with an asterisk (*).

Examinees for these certificates will be tested by computer in the following areas:

Addiction psychiatry—scientific foundations of addiction psychiatry, evaluation, treatment, systems-based practice.

*Brain injury medicine—patient evaluation and diagnosis, prognosis/risk factors, management, applied science; traumatic brain injury by severity and pattern, other neurologic disorders, systemic manifestations, rehabilitation problems and outcomes, basic science.

*Clinical neurophysiology—electroencephalography and electromyography, nerve conduction studies, other clinical neurophysiological studies, basic science.

Forensic psychiatry—civil and criminal law, legal regulation of psychiatry, corrections/correctional health care, legal systems/basic law, children/families, special diagnostic issues, death penalty.

Geriatric psychiatry—biological, developmental, social, and psychological aspects of aging; diagnostic methods; the interface between general medicine and psychiatry; neuropsychiatric and neurological aspects of aging; psychopathology and psychiatric diagnosis; treatment.

*Hospice and palliative medicine—approach to care, psychosocial and spiritual considerations, impending death, grief and bereavement, medical management, communication and team, ethical and legal decision, prognostication and natural history of serious illness.

*Pain—basic science, clinical research, animal models, ethics; assessment and psychology of pain; treatment (drugs and other modalities); clinical states (comprising taxonomy, tissue pain, visceral pain, headache and facial pain, nerve damage, and special cases: age-related, substance-using, and impaired individuals).

Psychosomatic medicine—issues of evaluation and management, symptoms and behaviors, treatment in these areas: dermatology, endocrine and metabolism, gastrointestinal disease, heart, HIV/AIDS, infectious disease and chronic fatigue syndrome, legal and ethical issues, lung,

neurology, obstetrics/gynecology, hematology/oncology, pain and palliative care, physical medicine and rehabilitation, kidney, rheumatology, transplant, surgery/critical care, pediatrics, organ transplantation; special topics such as drug interactions, systems-based practice issues, research/appraisal.

*Sleep medicine—normal sleep and variants, sleep–wake timing, insomnia, hypersomnolence, parasomnias, movement disorders, breathing disorders, sleep in other disorders.

MAINTENANCE OF CERTIFICATION

Since the mid-1990s, the ABPN certificates issued in general psychiatry and subspecialty examinations have been good for 10 years. Ideally, recertification for any profession as complicated and high-risk as the care of psychiatric patients requires the evaluation of many qualities, and hence the stimulus for the concept of an ongoing maintenance of certification in psychiatry—as well as in the other specialties of medicine.

Psychiatry recertification was originally planned to be an open-book, take-home examination. However, unless completed on-site, there can be no assurance that a given clinician actually took the examination, so in recent years, as mandated by the American Board of Medical Specialties, the ABPN has moved to a much more formal (and expensive) procedure. That is why certificate holders in psychiatry or any of its subspecialties must now pursue a multipart educational course structured to stimulate continuous improvement and learning. Although a little complicated to describe, the process is logical and well-designed to address its stated goals.

Of the ACGME competencies (Table 59.2–2), only medical knowledge can be fully assessed by written examination. So, to maintain certification past the initial 10 years, diplomates must: (1) assure their professional standing (have an unrestricted license to practice), (2) pursue continuing medical education (on average, 30 credits per year relevant to their specialty or subspecialty), (3) complete self-assessment modules, (4) complete performance in practice (PIP) modules, (5) demonstrate cognitive expertise, and (6) wait!—there's more promised yet to come. The first two of these require no further explanation; for the remaining parts, some details appear below.

Self-assessment—each 3-year period, an average of eight CME credits per year must come from completing ABPN-approved self-assessment activities. These must cover new information or best practices and provide feedback the individual can use for later improvement. The ABPN website lists dozens of self-assessment activities sponsored by various universities and other institutions. Some CME credits can be waived if a psychiatrist completes certain other activities, such as writing an (accepted) grant application, passing a cognitive examination, and writing an (accepted) article for publication.

Improvement in medical practice—also called performance in practice (and in any case abbreviated PIP), this group of programs assesses individual practice improvement in clinicians who are currently clinically active. Two modules are required: (1) assessment with best practices and practice guidelines of data from five of the clinician's own patient charts (perhaps they have, say, bipolar disorder or they receive a particular sort of treatment such as electroconvulsive therapy). Within 2 years, progress toward needed change would be judged by re-evaluating these (or other) five patients, similarly selected. (2) A second required PIP module uses feedback on the psychiatrist's effectiveness and/or efficiency from five peers or patients as a guide to self-improvement, which must be afterward reassessed. One each of the chart review and peer/patient feedback modules is required every 3-year period.

Patient safety—beginning in 2016, every individual applying for certification or recertification must complete 10 modules of an approved patient safety course. The more than five-dozen possible topics include the science of patient safety, measurement and outcomes, risk management, engaging patients and family, suicide risk, diagnostics and errors in psychiatry, and the epidemiology of errors and safety.

Cognitive expertise. Once in each 10-year cycle, psychiatrists must obtain a passing score on a multiple-choice, computer-administered examination. Currently, these are administered at Pearson VUE testing centers.

Of course, it approaches impossibility to evaluate the totality of a physician's competence, but the use of multiple methods would appear to enhance the likelihood of maintaining competence. And as currently structured, maintenance of certification should also reduce to a minimum both the financial and emotional burdens that physicians can experience during the initial certification process.

OTHER MENTAL HEALTH PROFESSION EXAMINATIONS

The ABPN only assesses graduate psychiatrists, not those in training (this is the purview of the ACGME). It forthrightly claims that it has no educational mandate and that it is not intended as a way of evaluating training strategies. Beyond such obvious methods as rating checklists and observation (direct or recorded), other examinations and formats have been brought forward to evaluate trainees and provide feedback to them.

Psychiatry Residency In-Training Examination

Nationally standardized examinations for residents still in training were adopted by many specialties during the last third of the 20th century. The Residency Review Committee for psychiatry requires that each program use a formal standardized test as a part of its evaluation of trainees; to

comply, all adult psychiatry training programs nationwide (as well as programs in Canada) have now adopted the 300-item Psychiatry Residency In-Training Examination (PRITE). In use since 1979, the PRITE is administered to more than 6,000 residents each year in fulfillment of the ACGME-required annual (in post-graduate years 2, 3, 4) evaluation of each resident's cognitive knowledge across biological, psychological, and social spheres.

This timed, closed-book, written, proctored examination offers trainees the opportunity to track their progress throughout training and to gauge their own performance against that of colleagues. It is developed by representatives of the American College of Psychiatrists in collaboration with the American Academy of Child and Adolescent Psychiatry, American Association of Directors of Psychiatric Residency Training, and Association for Academic Psychiatry; neurology consultants and current psychiatry residents collaborate.

To accommodate the schedules of individual training programs, the 5-hour PRITE can be administered on any 2 half-days within a 3-week window during the fall of each year; the window reduces the likelihood that questions will be circulated by the first examinees to take it. However the schedule is arranged, all residents in the program must participate; confidentiality of their results is guaranteed. PRITE examinations are administered following a written protocol by the training programs. Since 1996, there has also been a child PRITE, a 150-item test whose purpose and structure are similar to those of the general PRITE examination.

Resident input into test item development has been encouraged, causing some feeling that those who so participate reap a competitive advantage for their efforts. However, in the words of one participant, "I submitted my questions so long before they were scheduled for use that by the time of the examination, I couldn't remember what I had written!"

Training directors agree that scores will not be used to decide whether a resident will be dropped from a training program. However, many training directors use the results for individual counseling and for assessing the effectiveness of their programs. Roughly 15 to 20 percent of the PRITE emphasis is on neurology and the neurosciences; the remainder is on these principal psychiatric knowledge areas: adult psychopathology, alcoholism and substance abuse, behavioral science and social psychiatry, child psychiatry, consultation–liaison psychiatry, emergency psychiatry, forensic psychiatry, geriatric psychiatry, growth and development, patient evaluation and treatment selection, psychosocial therapies, and somatic treatment methods. Miscellaneous subjects include history of psychiatry, administrative psychiatry, and ethics.

The goal of the PRITE is truly to educate. Comparative percentile scores are reported for both the psychiatry and the neurology components of this examination. For the overall test and for each subsection, training directors

and each trainee receive the individual's exact numerical scores, ranking compared with peers in that program, and nationwide peer-group percentile rank. Training directors learn the grouped percentile ranks on each subtest for the residents in each training year, as well as how their residents compared nationally on each question. Residents also receive copies of the questions and answers, with references for further study examination.

Development and administration costs of the PRITE examination are borne by the training programs, which pay a flat fee per resident tested each year. Residents invest only their time. Copies of past PRITE examinations are available at a not-so-nominal fee (currently $150) from their website: http://www.acpsych.org/prite. Nonetheless, such openness mandates a yearly rewriting of the examination itself, which retains only enough questions from the prior year to allow comparison with examinations given in previous years.

Some training directors report that extreme scores on the PRITE strongly predict outcome of the ABPN board examination: "I've hardly ever seen a resident who scores above the 90th percentile on the PRITE who has any trouble at all with the boards," one training director reported. "But I counsel graduating residents who score below 50 percent to take at least one board review course."

Board Certification of Psychologists and Other Disciplines

For over 60 years the American Board of Professional Psychology (ABPP) has taken a quite different approach to credentialing: it relies principally on evaluating the actual work of certificate holders. For each of 16 current (as of 2017) specialty areas in clinical psychology, the multiple-step process begins with verification of credentials, followed by submission of extensive documentation to provide the basis for a later oral examination.

Although there is some variation depending on the subspecialty board the candidate is applying for, in general candidates submit a professional statement of their philosophy of practice, including awareness of cultural diversity, potential ethical dilemmas (perhaps one actually faced by the candidate), complex interpersonal interactions, and a description of current work coupled with the evidence base that underpins practice. In addition, there must be a practice sample. Whereas some subspecialties accept written descriptions of patient evaluation or treatment (perhaps with results of testing), others require at least one unedited, unrehearsed video of an actual patient evaluation or intervention. The video, typically running 50 minutes, must be accompanied by background material adequate to give members of the board an appreciation of the candidate's clinical competence. In some cases, the practice sample may be returned with a request for more detail. (Two specialty areas—forensic psychology and clinical neuropsychology—also require a written, multiple-choice

examination.)

The final step is a 3-hour oral examination conducted in person by three psychologists already certified in the relevant subspecialty area. A typical oral examination, which may be conducted by a panel or in three separate one-on-one sessions, will focus on the candidate's professional statement and work sample, though hypothetical examples are used to fill in any areas not covered in the practice materials submitted. The candidate must receive passing scores from at least two of the three examiners in each of several competency domains: science; knowledge and methods; assessment; intervention; consultation; relationships; interdisciplinary systems; individual and cultural diversity; ethical and legal standards; professionalism; and, when applicable, teaching, administration, research, and advocacy. Vignettes covering aspects of ethics may be introduced to spur further discussion. In essence, the oral examination process is similar to the writing and defending of a graduate school thesis.

Of the approximately 100,000 licensed psychologists in the United States, currently only about 4,000 are board-certified. Although most employers do not require certification as a prerequisite, some require it for continued employment; a few will pay for the employee's certification process; others may award a step-rise in pay for certification.

Beginning 2015, all certificates are issued for 10 years, with a maintenance-of-certification process much like that of other health care disciplines.

Social Work

Certification by the American Board of Examiners in Clinical Social Work relies on statements by the clinician and evaluations by two colleagues (chosen by the applicant). Annual recertification requires a statement of current direct clinical practice, continuing education, and state licensure. Credentials can be obtained in advanced generalist, child and family, and clinical supervision social work and in psychoanalysis. Through an independent organization, certification is also available in forensic social work.

Multiple-Station Examination Format

Similar to the practical examination used for decades in basic science courses, the objective structured clinical examination (OSCE) uses objective criteria to evaluate trainees' practical clinical skills. Typically, several stations are set up portraying clinical problems—a videotaped patient's interview to diagnose, results of laboratory values to interpret, an emergency scenario to evaluate. Test results are scored according to criteria the faculty determines in advance. Advantages are that a number of trainees can be evaluated within a short time, evaluations are objective, the

material tested is relevant to actual clinical situations, and the examinations can be quickly and cheaply administered.

Standardized Patients

Patients (often actors) trained in responses typical of a given disease have been used to present trainees with reproducible symptoms and histories, thereby addressing a leading source of examination error: individual patient variability. These "patients" can also be taught to rate their own reactions to those who examine them, yielding yet another evaluative dimension (though it may be subject to recall bias).

360-Degree Evaluations

Long used by businesses and some health care organizations such as the Department of Veterans Affairs are subjective assessments provided by supervisors, coworkers, students, and others (including patients) who work with the individual being evaluated. Questions can be tailored to fit any discipline—specific ACGME competencies as adapted to psychiatric practice, for example. The breadth of information obtained, often using standard Likert scales, can provide credibility that cuts through concern about possible bias. The resulting ratings and comments can help trainees become better aware of their individual strengths and needs for improvement. Gap analysis can help identify—and hence address—areas where one's overly sanguine (or critical) self-evaluation is discrepant with that of others.

The feedback module of ABPN's Improvement in Medical Practice (which, for historical reasons, it abbreviates PIP) is one such 360-degree evaluation. Every 3 years, clinicians are required to solicit feedback concerning their own clinical activities from five colleagues (including nonphysicians) or five patients. Each clinician must then identify opportunities for improvement and resurvey (the same or different respondents) within 2 years to determine what improvements have occurred. Areas for potential survey include patient care, medical knowledge, interpersonal and communication skills, professionalism, and medical knowledge, among others.

Psychiatrists In-Practice Examination (PIPE)

Practicing psychiatrists can assess their professional skills with the Psychiatrists In-Practice Examination. Its 200 online multiple-choice questions cover the content areas represented on the ABPN maintenance of certification examination. Feedback for each item as it is completed includes the correct response and two references for further exploration. Clinicians can earn up to 30 hours of CME if they answer at least 50 percent of the questions correctly; they have up to 1 year to complete the

exercise. A new PIPE is issued in July of odd-number years. Further information is available at https://www.acpsych.org/pipe.

EVALUATING THE SUCCESS OF EXAMINEES'AND EXAMINATIONS

Does interview control really equate to safe practice of psychiatry? How well do the boards and other examination formats test what mental health professionals need to know? An analogous question might be, does success in the 100-yard dash predict performance in the marathon?

Success of Examinees

For many years under the Part I/Part II examination system, the success rate for those seeking certification in general psychiatry fluctuated from the low 50s to the mid-70s percent. Indeed, until early in this century, the pass rate for those taking the oral examination was under 60 percent (with the Part I written examination running a few points higher). For the first 5 years (2011 to 2015) of the computerized cognitive examination, pass rates have averaged a much more robust 88 percent.

The reasons for this improvement are complicated and, probably, multifaceted. Of course, candidates just coming out of training may be better prepared now than they were in the past. (A large factor must be that only first-time candidates have taken the new-style written examinations.) But the examination itself may have changed in important ways. One of these is the fear factor: with less riding on a single throw of the dice that is worried over for months in advance, the relative absence of anxiety could play a role in raising the pass rate. Of course, the required cognitive examination is also potentially stressful, but it is arbitrated by a computer algorithm, not two strangers who for an hour sit in stone-faced judgment of the candidate's efforts.

One crucial factor may be ABPN efforts to improve the objectivity of its assessments. Here is some background: In a 1990 assessment of the reliability of evaluative judgment, 94 academic psychiatrists rated a videotape portraying the first 210 seconds of three interviews. The psychiatrists were asked to rate the interviews using a data checklist, a global rating scale, and an estimate of time allotted to various interviewing activities. Only a low level of agreement among raters was found for any of these methods; this performance did not improve when videotape segments twice as long were subsequently used. A 1991 study of agreement between examiners on the live-patient interview found that the examiners initially had a major disagreement (one said *pass,* the other *fail*) on 7 percent of candidates examined. A study from 1983 found that the results of the videotape and live-patient examinations disagreed (one showing "pass," the other "condition" or "failure" for 39 percent of candidates; 13.5 percent of all candidates were straight-up pass-fail discrepant. Data such as

these suggest that on-the-fly assessment by individual examiners may introduce much of the variation that was inherent in the oral examination process. In effect, computer-based testing and scoring has helped to pry the scores of modern candidates from the clutches of subjectivity.

Subspecialty Examinations

For years, the success rate for the child boards was about the same as for the adult oral exam—approximately 62 percent. Beginning in 2011, candidates who had entered into training since mid-2010 have been examined by a single computerized test, with the evaluation of the live patient relegated to the training programs, just as has been the case for adult psychiatry. For 2012 and 2015, the computerized examination yielded an average pass rate of 91 percent. Candidates who began training earlier have continued to be examined using the older, two-step (computer and oral) examinations, which have yielded pass rates averaging 87 and 76 percent, respectively. The reasons for the recent improvement are probably similar to those just given above for the general psychiatry examinations.

Other subspecialty examinations yield pass rates that tend to be stable for the specialty involved, though the subspecialties themselves vary considerably. Thus, the period 2011 to 2015 yielded pass rates of 95 percent for forensic psychiatry, 89 percent for geriatric psychiatry, 96 percent for addiction medicine, and 90 percent for psychosomatic medicine.

Recertification

As with initial certification, the maintenance of certification examinations for general psychiatry and for all of the subspecialties are not referenced tests; theoretically, all can pass. Furthermore, there is no limit on the number of times a candidate can take an examination in the effort to remain certified. In fact, over the past few years, a very high percentage of test takers have passed—which is only logical, inasmuch as this test is only a reaffirmation of competence already demonstrated. For example, from 2008 through 2015 the pass rates for general psychiatry and child and adolescent psychiatry have each averaged 99 percent. Through 2015, MOC pass rates for the subspecialties of addiction, forensic, and geriatric psychiatry have averaged 93, 95, and 95 percent, respectively.

Success of Examinations

With only one official ABPN examination to pass for certification, and no individual examiner judgments to consider, the issue of reliability of the examinations (consistency in a test–retest situation) has gone, so to speak, by the board. There remain, however, multiple opportunities to assess validity: the ability of an examination to yield an accurate measure of an examinee's capabilities. Several types of validity are relevant to these

examinations.

Content Validity. Does the examination evaluate material that experts consider vital to the assessment of psychiatric capability? That the content of the written psychiatry examination (Table 59.2–1) coupled with the oral examination (Table 59.2–3) covers the basic medical competencies (Table 59.2–2) would seem to defend a positive answer to this question. However, the concept of content validity carries with it the whiff of circular reasoning.

Construct Validity. Does the test actually measure what it claims to measure? In an examination that purports to assess competence in a medical specialty, test performance should correlate positively with amount of training. This was indeed the finding in a 1990 study that found PRITE scores rose with each year of psychiatry residency training.

Criterion Validity. In criterion validity, a new test is validated against an established measure of alleged accuracy. When both measures are administered simultaneously, concurrent validity is measured. When the established measure is made at a future time, as in a follow-up study, it is called predictive validity.

Following are some of the more recent studies that address aspects of examination validity.

Between-Examination Comparisons . Studies abound in psychiatry and other fields of medicine. Here are just a few of them.

- Two separate studies in surgery found that mock oral boards had substantial sensitivity and positive predictive value in relation to eventual performance on the certification examination.
- The senior in-residency examination for pathology correlated with pass rates for boards. Further, there was a strong correlation in the residency examination from year to year.
- In a 2013 study, the child in-residency psychiatry examination correlated positively with Part I of the ABPN subspecialty board examination, showing that "they assess the same achievement domain."
- A 1998 study of a psychiatry clerkship OSCE reported evidence of both construct and concurrent validity.
- A 1996 study of 701 psychiatrists found a "moderate" correlation (Pearson 0.67) between results of the PRITE and Part I (written) of the ABPN examination. Neurology score correlations were lower.

On the other hand, some studies have reported discouraging words. For example,

- OSCE performance was only weakly correlated with performance on the

national dental board examinations. The former may have measured qualities such as solving problems, thinking critically, and the ability to communicate, whereas the latter were designed to evaluate the examinee's ability to remember or recognize information.

- Performance on the OB/GYN in-training examination was in 2008 judged a weak tool to predict the probability of a resident failing the board's written examination.
- A 1996 study found no correlation between first-time success rates (Parts I *and* II) on the general psychiatry boards and passing the child and adolescent oral examinations on the first attempt.

The fact that programs lately tend to run both major psychiatry examinations concomitantly creates a perhaps now insurmountable challenge for predicting outcome of one from the other. On balance, however, the studies seem to suggest that there will be at least a modest correlation between the various examinations that psychiatrists must take over the course of training and beyond.

Processes

How the processes of medical care—laboratory tests, patient care instructions, and the like—are affected by certification or MOC is easy to evaluate, relative to the study of actual outcome. Various studies have evaluated how the medical processes operate and succeed.

- Internists performing in the top quartile of their MOC examination were more likely to perform mammograms and diabetes processes of care such as hemoglobin A(1c) screening and lipid testing.
- Diabetes quality measures improved for family medicine clinicians who pursued maintenance of certification activities. They also improved, though less so, for clinicians who did not. However, the writers expressed concern over evidence that clinicians preferentially attend CME activities in areas where they are already proficient.
- Another recent study suggests that family practitioners may choose MOC self-assessment modules based on what they already know, rather than what they need to learn.
- A 2017 study of quality measures for treatment of schizophrenia (such as use of atypical antipsychotic medication, polypharmacy prescription, and testing for glucose and lipid) revealed no statistical differences that could be explained on the basis of board certification of the treating psychiatrist. The same report found no process differences for neurologists treating ischemic stroke.

These and other published data seem to support the idea that certification and MOC can favorably affect the manner in which clinicians deliver care—at least, in non-mental health professions.

Outcomes

The strongest validation of the certification process comes from studies that demonstrate better outcomes among patients of clinicians who have earned a certificate. Currently, no such data exist for psychiatry, though there may be some in the pipeline. Other medical specialties have reported outcome data related to board certification or MOC, though some of these reports are less than fully supportive of the premise. And of course, studies may have been done and not published because they were negative.

- Surgeons certified in colorectal surgery had better survival rates among Medicare beneficiaries than did their non-certified colleagues. However, the authors were careful to point out that the board certification per se might not have yielded the result—"technique, training, experience, and focus of practice" could have been confounding factors.
- A "modest" association was reported in 2012 between physicians' cognitive skills on the general internal medicine MOC and a composite measure for the delivery of diabetes care.
- Patients were more likely to adhere to treatment for obstructive sleep apnea if it was prescribed by certified clinicians, who may have delivered a better educational experience.
- In 25 of 27 series reported in 2009, specialized surgeons had better cancer outcomes than did those who were not specialized. Two studies looked at the duration since passing boards; both found greater time (that is, more experience) equated to better outcomes.
- Ethics and morals are hard to define, let alone measure. However, separate studies found that physicians disciplined by medical boards in Oklahoma and California were less likely to be board-certified than control groups. A subsequent study found no such difference between disciplined and nondisciplined psychiatrists (board certification was reported for 43 percent vs. 46 percent, respectively). Yet another report revealed 5 times the likelihood of action against internal medicine practitioners who had not achieved board certification as those who had.

So as to avoid the appearance of cherry picking, here are some negative results.

- Surgeons with fundamentals of laparoscopic surgery (FLS) certification experienced 3 times as many bile duct injuries as did non-FLS surgeons did. The 2014 article deemphasized the negative effect of certification and instead highlighted the fact that the FLS+ surgeons were younger and less experienced.
- A 2013 summary article showed various modest advantages in provision of care for internists participating in MOC, but concluded that "the available medical literature does not provide strong support for the assumption that internists with certification or recertification have better

patient outcomes."

- No studies have supported superior performance of psychologists certified by the American Board of Professional Psychology. Actually, no studies have been reported at all.
- In a 2004 poster presentation, psychoanalysis patients of clinicians who were certified (by the Board on Professional Standards) did no better than patients of those who were not certified.

And *still* the only experimental study of ratings in mental health care delivery yet reported (in 1983) compared outcome of psychotherapy delivered by residents to nine patients using the following evaluations: didactic examination, global ratings by trainers, supervisors' ratings based on therapists' recollections of sessions, self-ratings, and independent evaluators' ratings of videotaped sessions. Patient outcome correlated only with the supervisors' ratings, which (as the authors acknowledged) could have been influenced by knowledge of that outcome. This halo effect was subsequently diminished when, after viewing videotapes of their trainees' actual therapy sessions, supervisors revised downward their overall evaluations. It appeared that trainees' personality characteristics and ability to discuss theoretical material boosted supervisors' opinions.

The ability of ABPN certification to signify quality mental health delivery is often held almost as a tenet of faith. With much to recommend it in terms of logic, it remains not only largely unproven, but even untested. At least the profession has begun to move beyond the observation of a single patient–provider encounter as predictor of how that trainee would perform in a different clinical situation. American psychiatry—indeed, every American medical specialty—should adopt as a task of utmost importance the search for correlations between board certification and patient outcome.

Summing Up

The relentless advance of scientific methods in the affirmation of a knowledge base has led to wholesale change in the methods whereby clinicians are evaluated and certified as fit guardians of the health of the public. The challenge now is to generate more and better studies to demonstrate that new examination methods lead to patient care that is improved in important ways.

References

Beullen J, Struvf E, Van Damme B. Do extended matching multiple choice questions measure clinical reasoning? *Med Edu*. 2005;39:410–417.

Bilimoria KY, Phillips JD, Rock CE, Hayman A, Prystowsky JB, Bentrem DJ. Effect of surgeon training, specialization, and experience on outcomes for cancer surgery: a systematic review of the literature. *Ann Surg Oncol*. 2009;16:1799–1808.

Buscemi D, Wang, H, Phy, M, Nugent K. Maintenance of certification in internal medicine: participation rates and patient outcomes. *J Community Hosp Intern Med Perspect*. 2013;2(4). doi: 10.3402/jchimp.v2i4.19753.

*Case SM, Swanson DB. *Constructing Written Test Questions for the Basic and Clinical Sciences*. 3rd edition (revised). Philadelphia, PA: National Board of Medical Examiners; 2002.

Chevron ES, Rounsaville BJ. Evaluating the clinical skills of psychotherapists. A comparison of techniques. *Arch Gen Psychiatry*. 1983;40:1129–1132.

Davis MH, Karunathilake I. The place of the oral examination in today's assessment systems. *Med Teach*. 2005;27:294–297.

Dennehy PC, Susarla SM, Karimbux NY. Relationship between dental students' performance on standardized multiple-choice examinations and OSCEs. *J Dent Educ*. 2008;72:585–592.

Downing SM, Yudkowsky R. *Assessment in Health Professions Education*. New York: Routledge; 2009.

*Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287(2):226–235.

Falcone JL, Gagne DJ, Lee KK, Hamad GG. Validity and interrater reliability of a regional mock oral board examination. *J Surg Educ*. 2013;70:402–407.

Ford JH 2nd, Oliver KA, Giles M, Cates-Wessel K, Krahn D, Levin FR. Maintenance of certification: how performance in practice changes improve tobacco cessation in addiction psychiatrists' practice. *Am J Addict*. 2017;26:34–41.

Galliher JM, Manning BK, Petterson SM, et al. Do professional development programs for maintenance of certification (MOC) affect quality of patient care? *J Am Board Fam Med*. 2014;27:19–25.

Gray SH, Brauer LD. Factors associated with outcome of psychoanalysis: intensity, analyst status, analyst immersion. *J Am Psychoanal Assoc*. 2004 Spring;52:454–455.

Grosch EN. Does specialty board certification influence clinical outcomes? *J Eval Clin Pract*. 2006;12:704.

Guzman E, Babakhani A, Maker VK. Improving outcomes on the ABS Certifying Examination: can monthly mock orals do it? *J Surg Educ*. 2008;65:441–444.

Hess BJ, Weng W, Holmboe ES, Lipner RS. The association between physicians' cognitive skills and quality of diabetes care. *Acad Med*. 2012;87:157–163.

Higgins RS, Bridges J, Burke JM, O'Donnell MA, Cohen NM, Wilkes SB. Implementing the ACGME general competencies in a cardiothoracic surgery residency program using 360-degree feedback. *Ann Thorac Surg*. 2004;77:12–17.

Hodges B, Regehr G, Hanson M, McNaughton N. Validation of an objective structured clinical examination in psychiatry. *Acad Med*. 1998;73:910–912.

Holmboe ES, Wang Y, Meehan TP, et al. Association between maintenance of certification examination scores and quality of care for Medicare beneficiaries. *Arch Intern Med*. 2008;168(13):1396–1403.

Jibson MD, Broquet KE, Anzia JM, et al; ABPN Task Force on Clinical Skills Verification Rater Training. Clinical skills verification in general psychiatry: recommendations of the ABPN Task Force on Rater Training. *Acad Psychiatry*. 2012;36:363–368.

Juul D, Sexson SB, Brooks BA, et al. Relationship between performance on child and adolescent

psychiatry in-training and certification examinations. *J Grad Med Educ*. 2013;5:262–266.

Khaliq AA, Dimassi H, Huang CY, Narine L, Smego RA Jr. Disciplinary action against physicians: who is likely to get discipline? *Am J Med*. 2005;118:773–777.

Kohatsu ND, Gould D, Ross LK, Fox PJ. Characteristics associated with physician discipline: a case-control study. *Arch Intern Med*. 2004;164:653–658.

Lipner RS, Young A, Chaudhry HJ, Duhigg LM, Papadakis MA. Specialty certification status, performance ratings, and disciplinary actions of internal medicine residents. *Acad Med*. 2016;91:376–81.

McDermott JF Jr, Streltzer J, Lum KY, Nordquist CR, Danko G. Pilot study of explicit grading criteria in the American Board of Psychiatry and Neurology Part II examination. *Am J Psychiatry*. 1996;153:1097–1099.

Morrison J. *The First Interview*. 4th ed. New York: Guilford; 2014.

Morrison J, Morrison T. Psychiatrists disciplined by a state medical board. *Am J Psychiatry*. 2001;158:474–478.

*Morrison J, Munoz RA. *Boarding Time*. 4th ed. Washington, DC: American Psychiatric Press; 2009.

Norcini JJ. The death of the long case? *BMJ*. 2002;324:408–409.

Norman G. Post graduate assessment—reliability and validity. *Trans J Coll Med S Afri*. 2003;47:71–75.
http://www.physiotherapy.asn.au/DocumentsFolder/CONFERENCE%202013/Educators%20

Oldham JM, Golden WE, Rosof BM. Quality improvement in psychiatry: why measures matter. *J Psychiatr Pract*. 2008;14 Suppl 2:8–17.

Parthasarathy S, Subramanian S, Quan SF. A multicenter prospective comparative effectiveness study of the effect of physician certification and center accreditation on patient-centered outcomes in obstructive sleep apnea. *J Clin Sleep Med*. 2014;10:243–249.

Rinder HM, Grimes MM, Wagner J, Bennett BD; RISE Committee, American Society for Clinical Pathology and the American Board of Pathology. Senior pathology resident in-service examination scores correlate with outcomes of the American Board of Pathology certifying examinations. *Am J Clin Pathol*. 2011;136:499–506.

Sanju G. Extended matching items (EMIs): solving the conundrum. *Psychiatrist*. 2003;27(2):230–232.

Schwaitzberg SD, Scott DJ, Jones DB, et al. Threefold increased bile duct injury rate is associated with less surgeon experience in an insurance claims database: more rigorous training in biliary surgery may be needed. *Surg Endosc*. 2014;28:3068–3073.

Shapiro T, Juul D, Scheiber SC. Exploration of failure on the subspecialty examination for child and adolescent psychiatry. *Am J Psychiatry*. 1996;153:693–697.

Sharp LK, Bashook PG, Lipsky MS, Horowitz SD, Miller SH. Specialty board certification and clinical outcomes: the missing link. *Acad Med*. 2002;77:534–542.

Sierles F, Daghestani A, Weiner CL, deVito R, Fichtner CG, Garfield DA. Psychometric properties of ABPN-style oral examinations administered jointly by two psychiatry residency programs. *Acad Psychiatry*. 2001;25:214–222.

Smeltzer DJ, Jones BA. Reliability and validity of the psychiatry resident in-training examination. *Acad Psychiatry*. 1990;14:115–121.

*Swick S, Hall S, Beresin E. Assessing the ACGME competencies in psychiatry training programs. *Acad Psychiatry*. 2006;30:330–351.

Wallace A, McFarland BH, Selvam N, Sahota G. Quality of care provided by board-certified versus non-board-certified psychiatrists and neurologists. *Acad Med*. 2017;92:108–115.

Webb LC, Juul D, Reynolds CF 3rd, et al. How well does the psychiatry residency in-training examination predict performance on the American Board of Psychiatry and Neurology Part I Examination? *Am J Psychiatry*. 1996;153:831–832.

*Westwood OMR, Griffin A, Hay FC, eds. *How to Assess Students and Trainees in Medicine and Health*. Oxford UK: Wiley-Blackwell; 2013.

Withiam-Leitch M, Olawaiye A. Resident performance on the in-training and board examinations in obstetrics and gynecology: implications for the ACGME Outcome Project. *Teach Learn Med*. 2008;20:136–142.

60

Ethics and Forensic Psychiatry

▲ 60.1 Clinical-Legal Issues in Psychiatry

CLARENCE WATSON, J.D., M.D., SPENCER ETH, M.D., AND GREGORY B. LEONG, M.D.

INTRODUCTION

Among the medical specialties, psychiatry plays a particularly distinct role within the province of the law, given its potential application across a wide range of legal issues. Contained at the core of most legal principles is the general presumption that adults of sound mind possess the capacity to understand, appreciate, reason, and choose throughout a boundless range of life activities. When those mental capacities are called into question at the behest of a legal quandary, psychiatry may be summoned to assist with reaching an equitable resolution. The presence or absence of a mental health or cognitive disorder may significantly impact the manner that a legal issue, whether civil or criminal, is perceived and ultimately managed. For instance, the presence of major depressive or psychotic symptoms may impact an individual's legal right to refuse a recommended surgical procedure. Similarly, the presence of significant psychiatric symptoms at the time of a criminal act may negate an individual's culpability for that offense. The enduring relationship between the law and psychiatry has evolved over the centuries as a necessary element of humankind's aspiration toward just resolution of legal disputes. In no other place is this more evident than with the historical development of the legal concept of insanity in criminal law. Once reserved only for those individuals whose minds were equated with that of a wild beast, the legal concept of insanity has progressed to varied jurisdictional approaches in determining the degree of mental impairment necessary to absolve an individual of criminal responsibility. Equally as significant is the historical development of the doctor–patient relationship. Once considered a lopsided paternalistic arrangement, in which patients were expected to passively accept the medical care directed by their physicians, the doctor–patient relationship has matured into a dynamic interaction that centers on patients' autonomy,

including a patient's right to refuse treatment.

From an eagle's eye view, the interplay between psychiatry and the law can be broadly divided into situations involving clinical treatment interactions and situations involving specific legal matters that require the input of psychiatrists to settle a contested legal issue. The complexities within this broad division cannot be overstated, and accordingly, the subspecialty of forensic psychiatry was, in essence, produced from necessity and the ever-developing legal landscape. Clinical-legal issues such as patient confidentiality, informed consent, the right to refuse treatment, and medical standards of care are ubiquitous in the psychiatrist–patient therapeutic relationship. Potential involvement by psychiatrists as experts in contested legal matters, such as criminal cases, personal injury lawsuits, employment-related claims, testamentary disputes, or child custody cases requires familiarity with legal principles and jurisdictional laws, and the ability to apply psychiatric knowledge to the relevant legal inquiry at hand.

Over the course of centuries, the unique pirouette shared between psychiatry and the law has at times lacked grace, but nonetheless, has steadily moved toward a more elegant cooperative existence. Often marked by tensions inherent in speaking different native tongues of the respective realms of medicine and law, the cooperative existence has benefited from both medical and legal literature aimed at bridging the gap. This chapter is meant to contribute to those literary efforts and offer common ground for the understanding of clinical-legal issues involving psychiatry.

THE DOCTOR–PATIENT RELATIONSHIP

The concept of the doctor–patient relationship is naturally bound to the ancient roots of medicine itself. Evolving from a period when physicians were not distinguished from priests, magicians, or other ancient healers, the doctor–patient relationship initially resembled an archaic priest–supplicant relationship that hinged on a parental figure's power to manipulate nature on behalf of his ward. During the 4th century B.C., the Hippocratic Oath set forth a code of ethics outlining principles that remain essential features of modern-day medical ethics, including patient confidentiality, abstaining from sexual contact with patients, and preventing harm to patients. The familiar dictum, *Primum non nocere* or first, do no harm, further distilled the fundamental ethical principle at the core of the doctor–patient relationship.

Clearly, at the heart of these traditional ethical doctrines, physicians are charged with acting only in the service of patients' medical benefit. The ethical principle of beneficence, where ultimate medical benefit is paramount, was traditionally used to justify a paternalistic approach to medical care that was primarily physician-driven. However, during the late 19th and early 20th century, there was a growing sentiment that patients had the right to actively influence the course of their medical care. The

concept of patient autonomy gradually gained ground and ultimately displaced the long accepted institution of medical paternalism.

Patient autonomy was clearly endorsed in the 1914 case, *Schloendorff v. Society of New York Hospital,* which served as the figurative "canary in the coalmine" for medical paternalism in the dawn of an evolving informed consent doctrine. In that case, Justice Benjamin Cardoza ruled, "Every human being of adult years and sound mind has a right to determine what shall be done with his own body…" While not formally established in *Schloendorff,* the informed consent doctrine would become a more clearly defined legal requirement four decades later in the 1957 *Salgo v. Leland Stanford Jr. University Board of Trustees* case.

Certainly, at the inception of any doctor–patient relationship there attaches a body of legal and ethical obligations owed by the physician to his patient. The shift away from medical paternalism toward the incorporation of patient autonomy permitted a more appropriate balance within the therapeutic relationship. Such a balance between a physician and his patient naturally facilitates a treatment partnership; however, there are clinical situations that call for prioritizing medical benefit over patient autonomy. Justice Cardoza's words in *Schloendorff* indicated his awareness of two important caveats regarding the "right to determine what shall be done" with one's own body: adulthood and soundness of mind. It follows, of course, that the impact of existing psychiatric symptoms must be considered in the full clinical context when issues related to patient autonomy arise. Medical decision-making capacity, involuntary psychiatric treatment considerations, and patient confidentiality matters may all be influenced by the presence of significant psychiatric symptoms that temporarily deemphasize autonomy concerns.

MEDICAL DECISION-MAKING AND COMPETENCE

As initially framed in *Salgo v. Leland Stanford Jr. University Board of Trustees,* the law requires that physicians provide facts to their patients regarding treatment options that are "… necessary to form the basis of an intelligent consent." Through this case law, the informed consent doctrine was established and formed a basis for physician liability when patient harm occurs in the setting of inadequate communication of the risks associated with treatment. In order to sufficiently obtain a patient's informed consent, three elements must be present: the physician's disclosure, the patient's voluntariness, and the patient's mental competence. Exceptions to the informed consent requirement include medical emergency, patient's waiver of the physician's disclosure requirement, and invocation of therapeutic privilege in situations where the disclosed information would be detrimental to the patient.

Generally, the informed consent disclosure requirement dictates that patients receive information regarding their diagnosis, the proposed

treatment with the associated risks and benefits, treatment alternatives, and likely consequences of declining treatment. The 1972 case, *Canterbury v. Spence,* established that physician liability may be found if there is a causal connection between a physician's failure to disclose and a patient's injury, where the disclosure of the significant risks of treatment would have resulted in the patient's refusal of that treatment. According to the court in *Canterbury,* that causal connection is based on "... what a prudent person in the patient's position would have decided if suitably informed of all perils bearing significance."

Depending on the state jurisdiction, the legal standard for an adequate disclosure may vary between a physician-based standard and a patient-based standard. In states that utilize a physician-based approach, doctors are required to provide information to patients that a reasonable physician in his specialty would disclose. States that utilize a patient-based approach may require an objective standard that considers the information that a reasonable patient similarly situated would want to know to reach a decision. Some states may require a subjective analysis of the patient-based standard that considers the information that the specific patient would want to know given his or her circumstances.

The voluntariness requirement directs that a patient be permitted to weigh the disclosed information and reach a decision unencumbered by coercion. Accordingly, threats, manipulation, or intentionally misleading information during a patient's decision-making efforts would nullify informed consent. A patient must be given the opportunity to freely consider the physician's disclosure and freely reach a medical decision.

The competency element requires that a patient possesses the mental capacity to understand, appreciate, and utilize disclosed information to reach a rational decision. Legally, all adults enjoy a presumption of competence in medical decision-making and other routine life activities until adjudicated otherwise. Traditionally, the term "competence" is a judicial determination regarding decision-making capability. Conversely, the term "capacity" refers to a clinical determination regarding a patient's decision-making abilities. Accordingly, the psychiatrist's role in competency evaluations is to determine an individual's mental capacity, which may later be formally adjudicated in a legal competency proceeding. In contested matters, courts have upheld the rights of a competent adult to refuse treatment. Courts may, however, adjudicate a person as incompetent if psychiatric symptoms or cognitive deficits impair an individual's mental capacity to communicate a choice, to demonstrate factual understanding and appreciation of his situation, or to rationally manipulate factual information related to his situation. In cases of adjudicated incompetence, treatment decisions generally shift from the patient into the hands of a substitute decision-maker or designated guardian.

Significant symptoms of mental illness or cognitive disorders are the

typical culprits that undermine competence. However, the presence of those symptoms alone is not sufficient to establish lack of mental capacity or to prove incompetence. In the 1973 case, *In re Maida Yetter,* a patient who was committed to a state psychiatric hospital refused treatment for possible breast cancer. Although delusional, the patient offered a rationale that justified her treatment refusal and allowed the court to determine that she was competent to make that decision independently and did not require a substitute decision-maker.

INVOLUNTARY HOSPITALIZATION AND TREATMENT

Two traditional legal justifications, *parens patriae* and police power, provide state jurisdictions with the authority to involuntarily hospitalize, or civilly commit, mentally ill people for the purpose of treatment. *Parens patriae* refers to governmental power to care for individuals who are unable to care for themselves. The initial application of the *parens patriae* principle for the care of the mentally ill in the United States is found in the 1845 case, *In re Oakes,* which established the basis for courts to involuntarily hospitalize mentally ill persons for treatment and to prevent harm. The police power principle refers to the societal interest of permitting governmental authority to protect the general public from potential harm.

Jurisdictional statutes regarding involuntary psychiatric hospitalization vary from state to state, but generally require the presence of a mental illness and a degree of risk of harm to the patient or to others due to the patient's symptoms of mental illness. Typically, the clinical determination for the necessity of civil commitment weighs a patient's current symptoms, relevant historical information, and collateral information sources (e.g., family members, friends, etc.) in the setting of a patient's inability or unwillingness to cooperate with the examination. During these assessments, the critical issue that must be addressed is whether patients' current psychiatric symptoms place them at risk of harming themselves or others. Additionally, some jurisdictions specifically require that involuntary treatment be implemented via the least restrictive alternatives available, as outlined in the 1966 case, *Lake v. Cameron.* Accordingly, clinical consideration must be given to alternatives to involuntary inpatient hospitalization, such as day treatment programs, group home placement, or other outpatient settings, depending on the circumstances.

Usually, these civil commitment assessments involve emergency situations in which family members, ambulance personnel, or police bring the patient to the emergency department because of concerns that the patient will imminently engage in suicidal, self-injurious, or violent behaviors. In addition, a voluntary patient might demand an immediate hospital discharge against medical advice and compel the psychiatrist to consider whether civil commitment is necessary in lieu of discharge. Due to

the urgency of these clinical situations, civil commitment laws permit clinicians to initiate emergency involuntary hospitalization without an initial court determination. This allows for immediate and necessary treatment interventions without the delay that awaiting a court hearing would cause. Following a statutorily determined period, a judicial review hearing is typically required to weigh the patient's need for continued involuntary hospitalization versus a patient's right to be discharged from inpatient treatment.

The legal standard of proof in civil commitment proceedings is "clear and convincing," which is less than the "beyond a reasonable doubt" standard in criminal trials. It is important to note that involuntary hospitalization does not per se permit medical treatment over patients' objections. Some jurisdictions require additional legal proceedings to permit the involuntary administration of medication over patients' objections. Following adequate treatment targeting the symptoms that necessitated the civil commitment, the patient may be discharged to the community to continue outpatient treatment. In some states, patients with extensive histories of treatment nonadherence and frequent hospitalizations may be evaluated prior to discharge for the appropriateness of outpatient civil commitment in order to improve community functioning.

In addition to standard civil commitment laws for psychiatric treatment, some states have enacted specialized civil commitment laws for the involuntary treatment of violent sex offenders who continue to pose a risk of danger following incarceration. In 1997, the U.S. Supreme Court upheld the constitutionality of sex offender civil commitment laws. Under these specialized provisions, these convicted and incarcerated sex offenders are evaluated prior to their prison release to determine if commitment criteria are present, including the existence of a congenital or acquired mental abnormality, the number and severity of previous sexual offenses, and dangerousness.

Psychiatrists cannot lose sight of the inherently intrusive nature of involuntary treatment. Accordingly, clinicians are ethically obligated to exercise due diligence in obtaining all relevant clinical data prior to determining the need for civil commitment. The psychiatrist continues to bear the ethical obligation of avoiding unnecessary breaches of patient confidentiality. However, testimony and communications to mental health courts required by law for the purpose of civil commitment are permitted without patient consent. Nonetheless, clinical efforts to preserve patient dignity and foster a therapeutic alliance with the patient remain paramount.

CONFIDENTIALITY AND PRIVILEGE

Confidentiality and privilege are related concepts that center on privacy

rights belonging to patients regarding information shared with their physicians during the course of therapeutic relationships. Confidentiality represents physicians' obligation to their patients to not disclose information to third parties without a patient's consent. Privilege also deals with the protection of patient information; however, it refers to the more specific protection against physicians being compelled to provide protected patient information during court proceedings. Privilege is a right that specifically belongs to the patient, in that a physician cannot claim the privilege and refuse to disclose patient information once a patient has waived the privilege.

The doctrine of patient confidentiality is an ethical and legal duty that has been long recognized throughout the tradition of medicine as evidenced by the Hippocratic Oath's charge to preserve patient privacy. Certainly, it is not difficult to understand the basis for this duty. The very nature of the doctor–patient relationship prompts patients to trust physicians with intimate details of their lives, relationships, and bodies for the purpose of medical treatment. Patients' candid communication of these details facilitates effective medical treatment and would likely be discouraged if physicians were free to reveal those details publicly.

Physicians may be held liable for breaches of patient confidentiality through legal theories of medical malpractice, invasion of privacy, and breach of contract for implied warranty of confidentiality. State legislatures have also enacted statutes that prohibit unauthorized release of patient information. Accordingly, physicians may also be held liable for the breach of statutory duty of confidentiality. In addition, the U.S. Congress enacted the Health Insurance Portability and Accountability Act of 1996 (HIPAA), which provides a set of privacy regulations that governs electronic transmission of health information for insurance eligibility and billing purposes. Psychiatrists should note that the practice of telepsychiatry also falls under HIPAA regulations. Accordingly, psychiatrists must ensure that their videoconferencing systems are compliant with HIPAA and state privacy requirements. Violations of HIPAA privacy regulations call for the imposition of fines, however, these regulations do not provide for compensation to patients impacted by confidentiality breaches.

There are several exceptions to the confidentiality duty owed to patients by their psychiatrists. First, disclosures by psychiatrists without patient consent may be permitted in emergency situations in order to prevent imminent harm. Second, psychiatrists may breach confidentiality when acting to protect third parties from a potentially dangerous patient. In 1976, the California Supreme Court decided *Tarasoff v. Regents of the University of California,* which set into motion a nationwide trend establishing the duty of mental health professionals to protect or warn their patients' potential victims or to notify police of a risk of danger posed by a patient. Laws regarding the duty to warn or protect third parties vary by

state jurisdiction and are defined by state statutory and case law. These laws provide an avenue for clinician liability if harm befalls a third party as a result of a mental health professional's failure to disclose the danger posed by their patient. Generally, disclosures made in good faith for the purpose of protecting a third party will not result in liability for breach of confidentiality. Third, psychiatrists may disclose patient information when adhering to state reporting requirements involving concerns such as, child or elder abuse, unsafe drivers, or impaired health care professionals.

The physician–patient and psychotherapist–patient privilege doctrines evolved from the historical legal prohibition of compelled testimony from attorneys against their clients. The basis of that legal tradition was to protect the sanctity of the attorney–client relationship, which would be effectively eviscerated if lawyers could be forced to testify against the very clients that they represented. Much like the justification for requiring patient confidentiality, the physician–patient and psychotherapist–patient privilege facilitates a more open and effective therapeutic alliance. In all state jurisdictions, patients possess the legal privilege to prevent their psychiatrists and psychotherapists from testifying about them. Federal courts also recognized patients' privilege following the U.S. Supreme Court's 1996 ruling in *Jaffee v. Redmond*. As the privilege belongs solely to the patient, psychiatrists have no basis to refuse to testify, if their patients chose to waive that privilege.

However, the privilege is not absolute and may be pierced under a number of circumstances, thereby paving the road to compelled testimony by treating psychiatrists. Patients may lose the privilege in cases where they raise their mental status as an issue in a civil or criminal matter. Medical malpractice claims initiated by patients against their psychiatrists or therapists would also remove the privilege. The privilege does not apply in civil commitment proceedings. In child custody cases, the privilege may be set aside if the best interests of the child outweigh the negative impact of the psychiatrist's testimony. Finally, the privilege could be potentially nullified if the patient shared the information in question with their physician in the presence of a third party who is not involved in the patient's care.

PSYCHIATRIC MALPRACTICE

Psychiatrists tend to dread malpractice lawsuits in excess of the statistical likelihood of being sued and having a judgment entered against them. Across all medical specialties there has been a major reduction in the rate of paid malpractice liability claims during the last decade, which, unfortunately, has not been accompanied by commensurate decreases in the cost of malpractice insurance policies. There are approximately 10 paid claims for every 1,000 American allopathic physicians each year. By contrast, however, the Institute of Medicine estimates that more than

100,000 Americans die annually from medical mistakes, and about 6 percent of patients are injured by medical mismanagement.

About half of malpractice claims filed are litigated. Of these claims, courts dismiss more than half, and another 40 percent are resolved before a trial verdict. The remaining small minority of cases become jury trials, of which 80 percent result in a defense verdict favoring the physician. Nationally, the mean time for resolution of a litigated malpractice claim is 2 years. In one large state, about 2 percent of all malpractice settlement reports were for psychiatrists, compared with 9 percent for internists and 10 percent for general surgeons. The majority of malpractice actions filed against psychiatrists derive from accusations of negligent treatment, adverse effects of prescribed medications, and suicidal and assaultive behavior of patients. The remainder mostly concerns alleged mistakes in diagnosis, breaches of confidentiality, and boundary violations. A psychiatrist's participation in a malpractice settlement exceeding a threshold amount will result in a mandatory report to the federal National Practitioner Data Bank (NPDB), an entity that was first authorized in 1986 as a repository for information such as adverse licensure and hospital actions, and health care providers' malpractice payments. NPDB entries are routinely accessed by hospitals and licensing boards, and may trigger inquiries that cause discomfort for physicians who may have felt unfairly coerced to agree to a settlement.

The theory of medical malpractice refers to professional negligence that falls within the civil law domain, as defined by state statutes and case law. For example, in Florida, patients or their representatives may file lawsuits for personal injury or wrongful death against health care providers. In those malpractice actions, plaintiffs must prove by a preponderance of the evidence that the actions of the provider represented a breach of the prevailing professional standard of care. The standard of care in this context is a legal term that refers to skillful treatment, which in light of all relevant circumstances, is recognized as acceptable and appropriate by reasonably prudent providers in the same specialty, such as practicing psychiatrists.

The law recognizes that the mere existence of an injury does not automatically imply that provider negligence occurred without proof that the provider did in fact breach the standard of care. A bad outcome does not per se equate to bad care. An exception to this rule is the legal concept, *res ipso locquitor* or the thing speaks for itself, where the commission of an egregious error, such as failing to remove an abdominal clamp during surgery, shifts the burden of proof from the plaintiff to the defendant physician to show that negligence did not occur. For psychiatrists, such an egregious error would certainly include sexual misconduct during a therapy session under the guise of treatment.

The necessary elements for patients to successfully establish

malpractice claims have been neatly summarized as encompassing four D's: duty, dereliction, direct causation, and damages. The duty element requires that a doctor–patient relationship had been established, thereby creating a duty of care. Although seemingly straightforward for most treatment situations, there are some instances in which the nature of the professional relationship can be vague. Does a duty begin with the first contact between a potential patient and psychiatrist setting an appointment over the phone or only after the initial in-person meeting? When does the duty terminate? Does a covering physician assume responsibility for a patient seen briefly on rounds over the weekend?

The element of dereliction refers to a breach of the duty to provide reasonable and appropriate care. Psychiatrists are not legally expected to offer perfect, state of the art, or even excellent diagnostic and treatment services. All that is required is for the psychiatrist to perform as would other psychiatrists facing the same patient. Being board certificated in psychiatry is neither necessary nor sufficient to establish that no deviation from the standard of care has occurred. It may be useful to demonstrate compliance with authoritative treatises, such as the Comprehensive Textbook of Psychiatry, recent peer-reviewed publications, or professional association treatment guidelines. However, it is an accepted principle that every person is unique and, consequently, patient-centered care may call for modifications in the usual therapeutic approach. Accordingly, clinical judgment that adequately accounts for a patient's full clinical context, despite an eventual bad outcome, would likely assist in buffering against a claim of dereliction. When proposing an unusual treatment plan that may be later questioned, the psychiatrist would ordinarily be expected to obtain informed consent from the patient. The details of the consent discussion should be carefully inscribed in the medical record, and it may be helpful for patients to sign a note documenting their understanding and agreement.

The direct causation element requires a direct causal connection between a breach of the standard of care and a subsequent patient injury. The term "causation" is one that may have conceptually different meanings in medicine and law. For scientific-minded physicians, a cause–effect relationship identifies that a particular prior event is the physical cause if it resulted in the later effect. The law takes a very different approach. The salient issue is the proximate cause, by which the focus of legal inquiry is whether the act in question was necessary to produce the injury. A common test for direct causation is whether the harm would not have occurred "but for" the prior action.

Legal commentators have argued that direct causation is that necessary element of a set of conditions that are jointly sufficient for the result. For example, a psychiatrist's elderly patient endorsed symptoms of depression, weakness, postural instability, and osteoporosis. The patient fell to the floor

while rising from the office couch at the end of a session and immediately complained of severe pain. The patient was taken to a local hospital emergency room and diagnosed with a hip fracture. Surgery was performed, but the patient developed a postoperative infection that ultimately proved fatal. This tragic chain of events undoubtedly had multiple underlying medical causes. However, from a legal perspective the psychiatrist could be found to be negligent and the proximate cause for all of the patient's subsequent injuries and death. Depending on the facts, the crux of the matter may be that but for the mishap while rising from the psychiatrist's couch, the patient would not have had a broken hip, surgery, an infection, and death. In this regard the patient was surely an "egg shell" whose fall was the "straw that broke the camel's back." As lawyers are fond of saying, one takes plaintiffs just as you find them.

The damages element refers to the plaintiff's alleged injuries that resulted from the physician's dereliction of duty. The civil law recognizes several categories of damages that may accrue to victims of malpractice, all of which are exclusively financial. Monetary awards to successful plaintiffs are intended to compensate the plaintiff for the harm suffered. Compensatory damages address what the patient has suffered as a direct consequence of the substandard care, and include present and future medical costs, lost income, physical and mental impairment, pain and suffering, and decreased life expectancy. Punitive damages, also known as exemplary damages, may be awarded in addition to compensatory damages when a defendant's behavior was found to be particularly egregious or malicious. Punitive damages are not meant to compensate a plaintiff's loss; instead, they are meant to punish the defendant and to serve as a deterrent to the defendant and others regarding similarly wanton behavior. Punitive damages, which are not covered by insurance, rarely arise in malpractice actions against physicians.

Certain aspects of psychiatric practice involve a greater risk of liability. Electroconvulsive therapy (ECT) is a procedure that usually calls for additional training and special precautions, and may be associated with a greater likelihood of adverse events than conducting a psychotherapy session. Consequently, psychiatrists who wish to perform ECT may need to purchase insurance policies that specifically cover this procedure. Prescribing certain psychiatric medications also warrants a cautious approach. Conventional antipsychotic medications place patients at risk of developing tardive dyskinesia, which, when severe, can be disfiguring and disabling. Successful malpractice cases have been filed when psychiatrists failed to alert patients of this side effect risk and neglected to perform and document routine examinations for its detection. Similarly, the use of atypical antipsychotic medications may call for routine assessments for weight gain and abnormalities of blood lipid and glucose metabolism in addition to patient discussions regarding these associated treatment risks.

Increasingly, psychiatrists use formal education materials to supplement informed consent discussions when prescribing common psychotropic medications, such as benzodiazepines, stimulants, and lithium.

The specter of patient suicide affects all practicing psychiatrists; indeed, it is a truism in the field that there are two kinds of psychiatrists—those who have already had a patient commit suicide and those who will. Rather than encouraging fatalism, the ever-present danger of suicide for certain groups of patients should motivate psychiatrists to remain cognizant of the basic suicide reducing strategies. Knowledge of the evolving epidemiology of suicidal behavior is fundamental in assessing static and dynamic risk factors for each patient. There is no question that asking patients about the presence and severity of depressive symptoms and suicidal thoughts is essential for proper care. Such direct inquires do not suggest to or predispose a patient to become actively suicidal. Psychiatrists must respond to an escalating risk of suicidality with very careful monitoring and appropriate modifications in the treatment plan.

No discussion of psychiatric malpractice is complete without consideration of boundaries. Much has been written about the conscious and unconscious forces that are in play in both patient and psychiatrist during a session, especially in the context of psychotherapy. Maintaining the integrity of the therapeutic frame, however challenging, is the linchpin of treatment. Of particular concern are boundary violations that constitute frank lapses in care, such as psychiatrists becoming romantically involved with a current or former patient. This boundary violation is usually characterized by a progression from seemingly innocent conversation, to more frequent social contacts, to physical affection, to sexual activity. In these cases, psychiatrists can exhibit remarkable denial regarding the serious consequences of their own behavior. Not infrequently, a transgressing psychiatrist, such as one who may be lonely, depressed, and substance abusing, begins a destructive relationship with a vulnerable patient, whose judgment may be impaired by an idealizing transference and self-defeating tendencies. Such unethical behavior may result in malpractice and licensing board actions that jeopardize a professional career.

THE PSYCHIATRIST AS AN EXPERT

Expert Witness Testimony

Psychiatrists may be asked to participate in civil and criminal court matters by providing professional opinions based on their psychiatric knowledge and experience as applied to a particular legal question. Typically, the opinions are provided to a requesting party, generally the court or an attorney, either as a written or oral report. In some instances, the opining psychiatrist is required to testify in court, with a jury or judge considering

such testimony along with other evidence in order to reach a verdict. Psychiatric expert testimony takes place in the public purview and is often viewed by the public as a "battle of the experts" or more pejoratively, battle of "hired guns."

In some cases, expert witness testimony is legally required, such as in cases involving allegations of medical malpractice. For example, in order to support a medical malpractice claim that a defendant psychiatrist's treatment led to a patient's suicide, the plaintiff must present expert testimony to show what a psychiatrist of ordinary skill, care, and diligence would have done differently under similar circumstances. Of course, should this malpractice case go to trial, there may also be psychiatric expert witnesses testifying on behalf of the defendant psychiatrist. Two court cases have governed admissibility of scientific expert testimony in the United States, *Frye v. United States* and *Daubert v. Merrell Dow Pharmaceuticals, Inc.* As both cases indicate, it is ultimately the judge's decision as to the admissibility of the expert witness testimony.

Role Conflict as a Psychiatric Expert Witness

The fundamental professional role of psychiatrists involves a clinical or therapeutic perspective. In the clinical or therapeutic setting, the treating psychiatrist deals with the patient's narrative truth and makes no further attempts at confirming a patient's statements. In contrast, when a psychiatrist works outside of the health care setting in the legal arena, the clinical or therapeutic perspective loses its primacy and an individual's self-report should be viewed with objectivity and hence skepticism. Assuming both the therapeutic and expert witness roles for the same individual invites erosion of the psychiatrist–patient relationship, whether it involves completing a disability evaluation in which the treating psychiatrist does not find support for the requested disability, or especially in the case of testifying in court as a psychiatric expert witness.

Separating the therapeutic and forensic roles is the best solution, though at times the dual role cannot be avoided, as in the case of involuntary civil commitment or involuntary medication hearings in which the treating psychiatrist is usually the only professional to testify. There are times in which a treating psychiatrist could be called to testify as a fact witness regarding a patient, but this would entail testifying about symptoms, diagnosis, prognosis, and similar clinical matters versus providing the ultimate psychiatric-legal opinion such as those involving capacity or causation.

Testamentary Capacity

With the population bulge of the "baby boom" generation reaching older age, psychiatrists are being confronted with an increasing incidence of its

emblematic mental condition, dementia, and the related clinical-legal issues of capacity and competence. Although adults are assumed to have capacity for the various activities of daily living and life tasks, that presumption is rebuttable, especially in the presence of cognitive impairment. Each specific deliberative legal act, such as testifying in court as a witness, executing a financial transaction, and even getting married, has its own criteria for mental capacity that must be satisfied in order for that action to be valid. Among the elderly the process of writing a last will and testament is a common and highly valued event that has important implications that will arise after the death of the testator.

Legislatures and courts recognize that making a will is frequently in the province of persons whose mental acuity is significantly less than young individuals engaged in other tasks. Consequently, the cognitive standard for testamentary capacity is generally set particularly low. In most jurisdictions the necessary elements are that persons understand: (1) the fact that they are actually making a will, (2) the nature and extent of their property, (3) the disposition of that property, and (4) who are their natural heirs. From a mental status perspective, this domain of knowledge does not require that the person making the will is fully oriented and, indeed, having the diagnosis of Alzheimer disease per se would not disqualify a person from having sufficient cognitive ability to make or change a will if the criteria for testamentary capacity could be established.

Although simple to state, the prerequisite criteria for testamentary capacity may be, in practice, quite difficult to ascertain after a person's death when the will is finally read. For instance, person could have been suffering from severe dementia at the time of death. However, the document was actually prepared years earlier, and it is the state of mind at the time when the will was written that is dispositive if the document is later contested. A retrospective assessment conducted after death is challenging when there is inadequate evidence of the individual's prior knowledge and abilities. Medical records, deposition testimony from relatives and acquaintances, and information about activities of daily living can contribute to a determination of capacity.

Since wills may be the product of debilitated elderly persons who depend on the physical and medical care of others, there is another requirement that must be satisfied before the terms of the will are honored. Testamentary capacity demands that the testator not be subject to the undue influence of another person. For example, the philanthropist Brooke Astor died at the age of 105 in 2007. Her elderly son was convicted in criminal court of preying upon his mother's diminished mental state by defrauding her in life and coercing her to change her will in his favor rather than directing her assets to her favorite art institutions. An adult child who spends considerable time with a vulnerable parent may be expressing love and devotion. Alternatively, that situation may be exploited to benefit the

child at the parent's expense.

Psychiatrists may be asked to opine whether a certain party exerted control over the person making a will. Certainly whether there existed a confidential relationship and whether the will was rewritten to inappropriately benefit that person are critical factors. In this setting, a "confidential relationship" between two individuals occurs when one of the parties has an obligation to act in good faith for the benefit of the other party. The psychiatrist's expertise will be most relevant to the question of the extent to which the testator was vulnerable to being influenced due to medical or psychiatric illness and the power of the asymmetric relationship.

Criminal Competencies

The presence of competent decision-making is necessary for a criminal defendant navigating the criminal justice process starting from the initial arrest to the implementation of the death penalty. While competence is presumed for adults, the presence of a mental disorder may negate competence. The criminal competencies include: competency to waive Miranda rights, competency to confess, competency to stand trial, and competency to be executed.

The U.S. Constitution and its Amendments, along with subsequent federal and state case law and statutes, delineate the legal safeguards for individuals accused of committing crimes. At the entry point into the criminal justice system, law enforcement officials must inform a criminal suspect of their Miranda rights, such as the right to remain silent and the right to counsel. The presence of an active mental disorder or intellectual disability can negate an individual's capacity to voluntarily, knowingly, and intelligently waive their Miranda rights. Though very few adults lack the capacity to waive their Miranda rights, the majority of incompetence to waive Miranda rights involves juveniles where factors such as immaturity and emotional vulnerability are more likely to adversely impact voluntariness, knowledge, and/or intelligence. The competency to confess might be considered under the broader competence to waive Miranda rights. In 1986, the U.S. Supreme Court in *Connelly v. Colorado* ruled that absent physical coercion by police, an actively psychotic individual, who approached a police officer to confess about committing a homicide as a result of command auditory hallucinations to confess, had provided a valid confession.

The most common forensic psychiatric evaluation in the criminal justice system is that of competence to stand trial. The criteria for competence to stand trial is guided by several U.S. Supreme Court decisions with two having the most significance, *Dusky v. U.S.* (1960) and *Godinez v. Moran* (1993). The *Dusky* case established that competency to stand trial is composed of two prongs: (1) understanding of the criminal proceedings;

and (2) reasonable ability to consult with defense counsel. In the *Godinez* case, the court ruled that if criminal defendants are competent to stand trial, they are also competent for other activities in criminal proceedings, such as the competence to waive counsel, competence to enter a plea (waive trial), and competence to be sentenced.

For jurisdictions that permit capital punishment, the U.S. Supreme Court has mandated certain exclusions to the imposition of the death penalty, including mental retardation (now referred to as Intellectual Disability in the DSM-5) in *Atkins v. Virginia* in 2002 and for those under age 18 at the time of committing the capital offense in *Roper v. Simmons* in 2005. The U.S. Supreme Court in *Ford v. Wainwright* in 1986 and *Panetti v. Quarterman* in 2007 prohibited the execution of condemned inmates who are incompetent to be executed. Incompetence to be executed involves a condemned inmates' inability due to a mental disorder to understand the death penalty and the reason it is being imposed on them. Although cases of competence to be executed are quite rare, it has raised several ethical issues for society and the medical profession.

Criminal Responsibility and Psychiatric Illness

The determination of culpability and subsequent societal sanctions for those committing a criminal act finds its origins at least as far back as the dawn of recorded human history along with considerations given to those with abnormal mental conditions. The evolution of English law in the area of criminal responsibility of the mentally disordered over the past millennium has exerted great influence on how criminal responsibility has been determined in the United States. Although the number of psychiatric evaluations for defendants claiming that a mental condition negated their criminal responsibility for an offense accounts for a small proportion of evaluations performed for the criminal justice system, the issue of criminal responsibility of those with mental disorders is arguably the most visible to the general public due to extensive media coverage, both fictional and nonfictional.

To be proven guilty of committing a crime, the prosecution must prove beyond a reasonable doubt that the accused person not only committed a prohibited act (actus reus) but also had the requisite mental state or "guilty mind" (mens rea) for the forbidden act.

The insanity defense is used to prove that an individual did not possess the necessary mental state for criminal responsibility due to mental illness at the time of the act. From the preceding millennium, the concepts of the "wild beast test" and the ability to distinguish between good and evil (right and wrong) represented society's efforts to legally define "insanity." These efforts evolved and culminated in the formulation of the M'Naghten insanity standard in 1843 by the English House of Lords following an insanity acquittal in the English trial of Daniel M'Naghten. The M'Naghten

insanity test required the defendant to prove that he or she was "laboring under such a defect of reason, from disease of the mind, as not to know the nature and quality of the act he was doing; or, if he did know it, that he did not know he was doing what was wrong." The M'Naghten standard was quickly adopted for use by the U.S. judicial system. Among the 46 states that permit the insanity defense, approximately half use a M'Naghten type of insanity test. Four states (Montana, Idaho, Utah, and Kansas) disallow the insanity defense.

While the majority of state jurisdictions use the M'Naghten insanity standard, some states use other insanity tests. Currently, the most common other type of insanity test is that of the American Law Institute (ALI). To qualify for insanity under the ALI insanity test, a criminal defendant must demonstrate that as a result of a mental disorder, he or she "lacks substantial capacity to either appreciate the criminality of his conduct or to conform his conduct to the requirements of the law" (Model Penal Code §4.01, 1962). When comparing the two current major insanity tests used in the United States, the M'Naghten test has two "cognitive" prongs (knowledge of the "nature and quality" and "wrongfulness") versus ALI's one "cognitive" prong ("appreciate the criminality") and one "volitional" prong ("conform conduct"). Interestingly, some states that use an ALI-type of insanity test do not use the "volitional prong," thereby functionally making the insanity standard more like the M'Naghten test. Wording differences between the M'Naghten and ALI insanity tests would suggest that an insanity defense would be more likely successful in jurisdictions using the ALI insanity test.

In actuality, the overall use of the insanity defense is infrequent. The defense is raised in approximately 1 percent of felony cases with an estimated 25 percent success rate. The vast majority of insanity acquittals involve prosecutorial concurrence prior to trial. It is estimated that juries reject the insanity defense in 93 percent of cases going to trial. It has been argued that the legal system's requirement to not inform juries of the potential legal consequence of an insanity acquittal, that is, that a defendant acquitted by insanity would not be automatically released into the community, but would be subject to commitment in a state mental health facility, adversely impacts jurors' willingness to accept insanity defenses.

CHILDREN AND THE LAW

Children and adolescents can become entangled in the U.S. legal system in a multitude of ways, some of which are comparable to adults and some are fundamentally different. In fact, the number of youth involved in court actions has increased dramatically, driven by criminal activity in gangs and on their own, by the escalating rates of foster care, divorce and custody disputes, and by the proliferation of child abuse, neglect, and victimization.

Psychiatrists would benefit from a familiarity of the basic concepts of children and the law, if for no other reason than to better understand their adult patients whose early histories may be replete with such encounters and whose younger family members may become enmeshed in some aspect of the law.

Children, Juvenile Justice, and Violence

Juveniles can be both perpetrators and victims of crimes. States created parallel juvenile justice systems for youthful offenders that in many ways were responsive to the developmental immaturity that characterize children and younger adolescents. In 1899, Illinois established the first juvenile court separating the adjudication and confinement of youth as delinquents from the conviction and imprisonment of adults as criminals. The scope of juvenile courts expanded to encompass both criminal acts and behaviors that would not be considered illegal for adults, such as curfew violations, drinking alcohol, and truancy which have been called status (as a minor) offenses. The rationale for juvenile courts' broad jurisdiction is the promotion of social warfare by prioritizing education and rehabilitation rather than punishment.

Every state proceeded to adopt a system of juvenile courts, whose procedures were then significantly modified by the Supreme Court's landmark cases under Chief Justice Earl Warren. In 1967, the U.S. Supreme Court held that, although it was permissible to treat juveniles and adults differently in court, youthful offenders facing possible incarceration are entitled to the same constitutional rights of notice of charges, counsel, protection from self-incrimination, and the opportunity to confront and cross-examine witnesses, as afforded to adults. Three years later the Supreme Court also ruled that the "beyond a reasonable doubt" standard of proof also applied in juvenile court. Extending constitutionally based procedural rights to juveniles did not imply that juveniles could escape conviction and punishment, especially if they committed serious crimes. Prosecutors have the discretion to file for a judicial waiver that moves the youth from juvenile court to adult court in order to deny the protections that juvenile jurisdictions provide, including the preclusion of long prison sentences.

Most states exclude first-degree murder charges from the juvenile system, which for many years effectively allowed a convicted juvenile murderer to be sentenced to execution. The eligibility of a person, who committed their crime as an adolescent, for capital punishment proved to be controversial even among death penalty advocates. The Supreme Court reviewed the question of juvenile capital punishment and decided in *Roper v. Simmons* (2005) that defendants were not eligible for execution if their crimes were committed under the age of 18 years. The court's opinion in this decision drew heavily on scientific data concerning neuropsychological

development that contrasted adolescents with adults, which were presented in amicus briefs submitted by psychiatric and psychological professional organizations. In *Miller v. Alabama* (2012), the U.S. Supreme Court extended its reasoning to preclude as unconstitutional a sentence of life without the possibility of parole for convicted juveniles.

The Columbine High School tragedy in 1999 engendered an awareness of a looming threat of mass killings occurring in a school, possibly perpetrated by an adolescent or young adult. Despite heightened vigilance of the danger of a lethal attack on campus, a mentally ill college student at Virginia Tech University in 2007 murdered 33 persons, and in 2012 a mentally ill adolescent opened fire on and mortally wounded 20 children at the Sandy Hook Elementary School in Connecticut. There have been strident demands to develop strategies to profile potentially deadly students who may be targeting schools for deadly violence. Youths with the risk factors of a history of psychiatric illness, aggression, suicidal behavior, and/or family strife, and who have access to weapons may indeed be at higher likelihood for violence. However, less than 5 percent of crimes overall can be attributed to mental illness. Identifying those individuals at risk with any degree of certainty must overcome the daunting challenge of the high false-positive rate associated with predicting a rare event.

Most experts are skeptical of the wisdom of deploying armed security in elementary and high school buildings. Further, implementing zero-tolerance policies for aggressive behavior may disproportionally impact minority, disadvantaged, and special needs students who generally benefit the least from suspensions or expulsions. Alternatively, school-based interventions designed to decrease bullying and stigma, provide enhanced access to students for counseling and mental health care, expand communication among students, faculty, and parents, and eliminate or at least safeguard guns in the home, all hold promise of more effectively addressing the threat of interpersonal and self-inflicted violence.

Child Custody

Each year in the United States there are about half the number of divorces as there are marriages. Divorce proceedings sever the marriage contract between spouses, and although children may be a focus of the case, they are not considered formal parties to the legal action. The court has the responsibility to protect the well-being of dependent children affected by divorce, which may be straightforward when the parents are in generally agreement or hotly contentious when they are in battle. Child custody subsumes both legal custody, referring to the responsibility for making major decisions about health and education, and physical custody, referring to which parent is spending time with the child. Many jurisdictions favor awarding joint legal custody to both parents, unless there is evidence that this would be frankly detrimental to the child's best

interests. Under those circumstances the court may grant sole legal custody to one parent, with or without time-sharing with the other parent. A determination of the best interests of the child is made by evaluating all of the factors affecting the welfare and interests of the minor child and the circumstances of that family, including, but not limited to: the demonstrated capacity of each parent to facilitate and encourage a close and continuing parent–child relationship and to determine, consider, and act upon the needs of the child as opposed to the needs or desires of the parent; the moral fitness and mental and physical health of the parents; evidence of domestic violence, sexual violence, child abuse, abandonment, or neglect; and the developmental stages and needs of the child and the demonstrated capacity of each parent to meet the child's developmental needs.

Treating psychiatrists of a parent or child in custody or visitation litigation may have important information that could influence the court's ruling. The scope and limits of confidentiality in such matters may become a difficult and contentious question. Further, the availability of the psychiatrist to write reports, testify, or advocate in these adversarial legal proceedings is fraught with conflicting ethical, legal, and therapeutic duties. Many treating psychiatrists have sought to avoid becoming entrapped in these dilemmas by limiting their documentation and recollection, though it may not be possible to predict which patient will be the source of a subpoena. Consultation with an attorney and referral to a forensic psychiatrist may be advisable for the treating psychiatrist confronted with such circumstances.

Child Abuse

Reports of suspected maltreatment, which are mandated by law for health care and school personnel, will trigger investigations by the police or a child protective agency. If probable cause for maltreatment is found, such as bruises or malnutrition, the child can be removed from the home, and the agency might petition the court for alternative placement in foster care. Allegations of sexual abuse usually result in criminal charges against a parent, family member, or nonfamilial perpetrator. As the victim of interpersonal violence, children and adolescents suffer enduring psychological injuries, which may also lead to civil court actions seeking to recover compensation for the injured child. Younger children are more vulnerable since, as a rule, they are less able to protect themselves from being exploited by adult predators.

Society and the medical profession have traveled far in the last 100 years from a time when physical child abuse was unrecognized to the current era of far greater sensitivity to multiple forms of maltreatment, including neglect, emotional and physical abuse and sexual molestation. However, significant difficulties arise in the evaluation of allegations of

sexual abuse, especially in young children. The psychiatrist examining a child about whom there is an accusation of sexual abuse must face the challenge of infusing scientific objectivity into the practical task of providing judgments regarding the validity of symptoms, their putative etiology, the harms suffered by the child, the prognosis, and what present and future treatment interventions are indicated.

REFERENCES

Appelbaum PS, Gutheil TG. Confidentiality and privilege. In: Appelbaum PS, Gutheil TG, eds. *Clinical Handbook of Psychiatry & the Law*. 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2007.

Atkins v. Virgina, 536 U.S. 304 (2002).

Canterbury v. Spence, 464 F.2d 772 (DC Cir. 1972).

Colorado v. Connelly, 479 U.S. 157 (1986).

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

Dusky v. United States, 362 U.S. 402 (1960).

Everson MD, Sandoval JM, Berson N, Crowson M, Robinson H. Reliability of professional judgments in forensic child sexual abuse evaluations: unsettled or unsettling science? *J Child Sex Abus*. 2012;21:72.

Ford v. Wainwright, 477 U.S. 399 (1986).

Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).

Giorgi-Guarnieri D, Janofsky J, Keram E, et al; American Academy of Psychiatry and the Law. Practice guideline: forensic psychiatric evaluation of defendants raising the insanity defense. *J Am Acad Psychiatry Law*. 2002;30(2, Suppl):S3–S40.

Godinez v. Moran, 509 U.S. 389 (1993).

Greenberg SA, Shuman DW. Irreconcilable conflict between therapeutic and forensic roles. *Prof Psychol Res Pract*. 1997;28:50.

Harvard Law Review Association. Developments in the law: civil commitment of the mentally ill. *Harv Law Rev*. 1974;87:1190–1406.

In re Maida Yetter, 62 Pa. D. & C. 2d 619 (1973).

Jaffee v. Redmond, 518 U.S. 1 (1996).

Lake v. Cameron, 364 F. 2d 657 (D.C. Cir. 1966).

Meyer DJ. Psychiatric malpractice and administrative inquiries of alleged physician misconduct. *Psychiatr Clin North Am*. 2006;29:615–628.

Miller v. Alabama, 132 S.Ct. 2455, 567 U.S. __, (2012).

M'Naghten's Case, 8 Eng. Rep. 718 (H.L. 1843).

Model Penal Code §4.01 (1962).

Mrad DF, Watson C. Civil commitment. In: Drogin EY, Dattilio FM, Sadoff RL, Gutheil TG, eds. *Handbook of Forensic Assessment: Psychological and Psychiatric Perspectives*. Hoboken, NJ: Wiley; 2011.

Panetti v. Quarterman, 551 U.S. 930 (2007).

Roper v. Simmons, 543 U.S. 551 (2005).

Salgo v. Leland Stanford Jr. Univ. Bd. Trustees, 317 P.2d 170 (Cal. 1957).

Schloendorff v. Society of New York Hospital, 105 NE 92 (N.Y. 1914).

Shulman KI, Cohen CA, Kirsh FC, Hull IM, Champine PR. Assessment of testamentary capacity and vulnerability to undue influence. *Am J Psychiatry*. 2007;164:722–727.

Strasburger LH, Gutheil TG, Brodsky A. On wearing two hats: role conflicts in serving as both psychotherapist and expert witness. *Am J Psychiatry*. 1997;154:448–456.

Szasz TS, Knoff WF, Hollender MC. The doctor-patient relationship and its historical context. *Am J Psychiatry*. 1958;115(6):522–528.

Tarasoff v. Regents of the University of California, 551 P.2d 334 (Cal. 1976).

Weinstock R, Leong GB, Silva JA. Competence to be executed: an ethical analysis post Panetti. *Behav Sci Law*. 2010;28:690–706.

▲ 60.2 Ethics in Psychiatry

LAURA WEISS ROBERTS, M.D.

Ethics is a scholarly discipline that brings greater understanding of moral aspects of human experience and action. Psychiatry, more than any other field of medicine, concerns itself with studying and strengthening the capacities that make humans distinct and that allow people to attach, learn, love, reason, feel, live with one another, and make choices. Separately and together, the fields of ethics and psychiatry examine human meaning, intention, and consequences, and both are concerned with the expanse and limits of the self in the world. Thus by their natures, ethics and psychiatry are overlapping fields of inquiry.

With so much in common, it is perhaps surprising that psychiatry enjoys more than its share of ethical controversy. Psychiatry is viewed with some skepticism by society at large, and it is seen as ethically problematic. Psychiatrists themselves are often at odds on matters of ethical importance within the profession. Long-standing ethical issues related to power and intimacy in psychodynamic–psychotherapy relationships have now been joined by intriguing new ethical questions about privacy and boundaries in treatment conducted solely through asynchronous texting between patients and their therapists. Similarly, older dilemmas regarding the potential exploitation of vulnerable individuals in psychiatric research have now been joined by new debates regarding undue influence, unconscious bias, and conflicts of interest in human studies. Old controversies regarding patient confidentiality in publishing patient case reports are now joined by new uncertainties regarding privacy in analyzing social media postings to protect the health of the public by identifying suicidal adolescents, for example, or individuals at risk for engaging in violence. Novel issues pertaining to topics as widely varying as psychiatric genetics, interpersonal violence, sexuality, and global mental health disparities arise each day.

Indeed, ethical issues are to be expected in a field that is important for understanding the human condition and is also undergoing extraordinary advances, scientific to societal.

As for other specialist physicians who reside in the house of medicine, learning about ethics is helpful preparation for psychiatrists and psychiatrists-in-training who will encounter complex dilemmas in their everyday work. Familiarity with resources such as the ethical guidelines of the American Medical Association (AMA), the American Psychiatric Association (APA), and the American Academy of Child and Adolescent Psychiatry is a valuable first step. Knowledge of ethical codes of historical importance, such as the Hippocratic Oath of the 4th century BC and Percival's *Medical Ethics* from the late 1800s, provide perspective on the evolution of the standards of conduct for physicians. A clear working knowledge of resources from related areas, such as professionalism, legal concerns, gender-related issues, and cultural competencies, is also vital to sound ethical practice and training in psychiatry.

Are these general codes of ethics for physicians sufficient for an understanding of the ethical requirements of psychiatrists? Stated differently, should psychiatrists be held to a higher or different standard of ethics than other physicians? Psychiatrists are presently held to the same standard of ethics as other physicians. Many psychiatrists argue, however, that a more complete knowledge of ethics is particularly important for mental health practitioners because of the nature of mental disorders and the impact of stigma. Others argue that greater mastery of ethics is essential because of the influence psychiatrists may have over their patients' lives as a result of transference, that is, the set of emotions and thoughts brought to the therapeutic relationship from the patient's prior life experience. Yet others argue that this level of ethics knowledge is still not enough: true ethics expertise, robust decision-making skills, and disciplined ethics practices are of special importance for psychiatrists because of their professional responsibility to initiate treatment to protect a patient or others from harm when necessary—even if it is against the patient's wishes. For all of these reasons, psychiatrists and psychiatrists-in-training will ideally possess a thorough understanding of the field of ethics, with both its underpinnings and its applications.

Toward this aim, this chapter introduces the foundations of present-day ethics in psychiatry. Outlined are key bioethics principles and illustrations of approaches to identifying and resolving psychiatric ethics dilemmas.

ETHICAL SCHOOLS OF THOUGHT

Several ethical schools of thought have influenced how physicians understand ethical aspects of medical practice. Examples include the deontological school, utilitarianism, justice-based ethics, and virtue-based ethics.

These perspectives differ greatly on what constitutes the beliefs and actions of a good person and assess ethical behavior differently. The philosophers who developed these paradigms sought to explain as much as possible with a few simple rules. Such simplicity is elegant academically, though harder to apply in real life, as Table 60.2–1 notes. Each school of thought, taken alone or taken to an extreme, may lead to conclusions that most people in a pluralistic society would find unacceptable. Nevertheless, the models provide a framework for discerning and evaluating moral issues encountered in everyday experience.

The deontological school centered on the work of Immanuel Kant (1724 to 1804), who articulated the inherent sacredness of the individual person and the primary intention of a human being as a moral agent. He believed that certain ways of behaving are absolute from an ethical perspective, irrespective of consequences. Kant articulated a rule (the "categorical imperative") guiding all moral behavior. His first formulation of the categorical imperative was that a person should "act only on that maxim through which you can at the same time will that it should become a universal law." For example, in deciding whether to lie or steal or kill, individuals should think about whether they want others to engage universally in these behaviors. Kant's second formulation of the categorical imperative was to "act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end." In other words, it is unethical to use another person solely as an object to obtain a goal without concern for the effect of one's actions on the person. Moral value, in Kant's system, relates to respect, rationality, and duty in relation to the moral agent.

Table 60.2–1.
Examples of Ethical Models of Thought

Model	Major Figure(s)	Key Features	Strengths	Limitations
Deontology, or the ethics of duty	Immanuel Kant	There is an absolute good and right. To be ethical, persons must fulfill these absolute duties unconditionally.	Some moral judgments apply in almost all circumstances to everyone (e.g., respect for persons).	Conflicting obligations. Too much emphasis on obligation, not enough on consequences or relationships. Very abstract.
Utilitarianism, or the ethics of consequences	John Stewart Mill Jeremy Bentham	Weigh the consequences of actions and rules. Those that bring about the	Useful in formulating public policy because it requires impartial assessment of all interests.	Difficult to translate into practice. Practices that harm individuals and groups may be ethically justifiable in this

		greatest good for the most people are most ethical.	Has beneficence as a goal.	model. May demand that the minority sacrifice too much for the majority.
Justice-based ethics	John Rawls	Decisions should be made with respect and fairness in relation to each individual.	Demonstrates respect for the individual person. Has equality as a goal.	Neglects principles beyond fairness in clinical decision making.
Virtue-based ethics	John Gregory Thomas Percival	Ethical decisions are rooted in qualities or character, such as honesty and faithfulness. Actions are right only if they are what a virtuous person would do in the same situation.	Internal values and habits rather than external rules and imperatives. Highlights importance of qualities of the individual, of personal character.	Virtue may be too unclear or pluralistic to serve as a basis for judgment. Persons may act wrongly out of virtue (e.g., withhold the truth from a dying patient to avoid emotional upset).

Reprinted with permission from *Concise Guide to Ethics in Mental Health Care*, (Copyright © 2004). American Psychiatric Association. All Rights Reserved.

The utilitarian school of thought originated with Jeremy Bentham (1748 to 1832) and John Stuart Mill (1806 to 1873) in 19th-century England. They argued that the value of an act should be judged solely on its effect or consequences. Acts are ethically right, in this view, if their effect is to produce the greatest good for the greatest number of people. Acts are not right or wrong in and of themselves, according to act-utilitarians. Lying, stealing, and killing are not inherently wrong, for example. In this school, the ends justify the means. Rule-utilitarianism moves beyond act-utilitarianism to suggest that good individuals will undertake decisions guided by rules that—if followed by all people—would bring about the greatest good for the greatest number of individuals. In this formulation, there is more "forgiveness" for decisions with unpredictably negative outcomes. Rule-utilitarianism does not address the role of the intentions of individual moral agents, nor does it answer the question of what is "good."

More recently, in contrast with utilitarianism, John Rawls (1921 to 2002) developed a school of thought drawing from his concept of justice. It followed upon strengths of Kant's deontological approach in that it highly valued the individual as a moral agent. Rawls supported the rights of the individual as a matter of fairness. He gave primacy to the intrinsic worth of each person rather than to a metric that assessed greatest utility for the group as a whole. He stated, "Each person possesses an inviolability founded on justice that even the welfare of society as a whole cannot override." Rawls asserted that in deciding what to do and what rules and

systems to establish, people should place themselves behind a "veil of ignorance" and assume that they do not know who they are in the system. For example, in deciding what welfare benefits a society should offer and the amount of taxes to impose on the rich, people should assume that they do not know whether they will be rich or poor. In deciding what authority doctors have over patients, people should assume that they do not know whether they will be in the role of the doctor or in the role of the patient. Rawls' approach, as with that of Kant, does not rely on consequences of actions but, rather, their primary intent.

Virtue-based theory has contributed greatly to the ethical framework of medicine and the key concepts of professionalism. John Gregory (1724 to 1773) was a Scottish physician–ethicist, and Thomas Percival (1740 to 1804) was an English physician–ethicist. Their work, taken together, created the foundation for the modern ethical concept of medicine as a profession. Concerned about exploitative, self-interested, and entrepreneurial practices of physicians at the time, they articulated the virtues of scientific and clinical competence, dedication to the interests of the patient, and service to the public through ensuring the skills of peers and students or apprentices. Gregory described the moral qualities of an astute and compassionate physician:

> The chief of these is humanity; that sensibility of heart which makes us feel for the distresses of our fellow-creatures, and which, of consequence, incites us in the most powerful manner to relieve them. Sympathy produces an anxious attention to a thousand little circumstances that may tend to relieve the patient; an attention which money can never purchase... If the physician possesses gentleness of manners, and a compassionate heart, and what Shakespeare so emphatically calls "the milk of human kindness," the patient feels his approach like that of a guardian angel ministering to his relief: while every visit of a physician who is unfeeling, and rough in his manners, makes his heart sink within him.

Gregory and Percival individually emphasized the virtues of the honorable physician as demonstrating self-sacrifice and self-effacement and as seeking at all times to advance the field of medicine, not for financial gain but the benefit of others in society. The modern notion of integrity derives from this line of thought, and the work of Gregory and Percival continues to inform the discussion of conflicts of interest and how they may be avoided or resolved.

These four schools of thought have been highly influential in present-day principles of ethics that the AMA and APA advance. Other schools of thought exist in the United States and throughout the world, with rich and long-standing traditions and different implications for current practice. From legal and ethical points of view, the AMA and APA hold physicians engaging in professional activities in the United States to their codes of ethics and standards.

BASIC ETHICAL PRINCIPLES

The modern field of biomedical ethics takes primarily a principle-based approach. In their influential text *Principles of Biomedical Ethics,* Beauchamp and Childress defined the most critical among the principles to be autonomy, beneficence, nonmaleficence, and justice. The Belmont Report on human research ethics elevated the concept of respect for persons, and recent ethics scholarship has emphasized the principle of respect for the law. Other principles, such as compassion, confidentiality, fidelity, integrity, and veracity, are also central to the practice of psychiatry. See Table 60.2–2. Ethical analyses examine the interplay of these principles and seek to identify and resolve when ethical principles are in conflict.

APPROACHING ETHICAL DILEMMAS

Ethical dilemmas do not typically emerge in simple situations of doing good versus doing bad. In such situations, there is no dilemma. Indeed, ethical dilemmas come when two seemingly good things are at stake. Consider the example of a psychiatrist who receives a request from a worried and loving mother to provide information about issues arising in her son's therapy sessions. The son has a good relationship with his mother but has been bullied at school, feels embarrassed, and does not want his parents to know what is bothering him. For the psychiatrist, the good of responding empathically and with accurate information to the distressed mother competes with the good of protecting the fragile alliance with the patient. Dilemmas often come when two or more of the basic bioethics principles are in conflict or cannot be completely fulfilled given the circumstances of a particular situation.

Table 60.2–2.
Examples of Biomedical Ethics Principles

Autonomy	The notion of self-rule; the capacity to make authentic decisions related to one's body and mind.
Beneficence	Doing good; the commitment to seek to bring about benefit.
Compassion	Deep regard for the experiences and suffering of others.
Confidentiality	A legal privilege associated with the right of privacy; the obligation not to disclose information obtained from a patient or observed or gathered in caring for a patient.
Fidelity	"Faithfulness" or loyalty to ethical ideals.
Integrity	The notion of being whole or complete; the capacity to adhere wholly to the principles of the profession.
Justice	Equitable distribution of benefits and burdens in society.
Nonmaleficence	Avoiding harm or injury to others.
Respect for the law	The obligation to adhere to the law.
Respect for persons	Fundamental regard for the dignity, sacredness, and value of the individual.

Veracity	"Honesty," involving the positive duty to tell the truth and the negative duty to avoid deception.

Recognizing Ethical Issues and Tensions

The ability to identify ethical issues and tensions present in the patient care situation is thus the first and most basic of the ethics skills important to sound clinical practice of psychiatry. Recognizing the ethical features of a patient's case requires sensitivity, and this capacity is enhanced by knowledge of ethical principles and how they may come into conflict.

Seeing how ethical principles may compete or collide is informative and can help practitioners to discern a path forward. For instance, consider the case of a newly diagnosed cancer patient who, overwhelmed and fearful, attempts suicide by overdose. She requests aid in dying from her oncologist, who also feels overwhelmed and fearful because of the patient's reaction and refers the patient for a psychiatric evaluation. The psychiatrist learns that the patient's mother died from cancer and that she was her mother's only caregiver over the long and difficult course of her illness. The psychiatrist finds that the patient likely had a pre-existing depression before learning of her cancer diagnosis. The patient is decisionally capable, despite her distress, and the psychiatrist feels great empathy for the patient. In this example, many principles are in play and several are in tension: most clearly, honoring the patient's wish to die (autonomy) conflicts with the physician's responsibility to avoid harm (nonmaleficence). A superficial analysis would suggest that the patient's right, especially in an autonomy-oriented society, would trump all other considerations.

However, sound standards for clinical and ethical care are those that are informed by the full set of ethical principles that govern medicine. By recognizing the two principles that are in opposition—in this case, autonomy and nonmaleficence—and thinking through the other principles, the clinician begins to see additional options for intervention that align with standards of medicine. In this case, identifying ways of reducing the patient's physical and psychological suffering, treating her pre-existing condition, providing accurate information about the nature of her condition and potential treatments, and adhering to legal requirements all may contribute to an improved physical and mental health outcome for this patient. Ideally, clinical care and ethical reflection work together, as illustrated in this case. The disciplined exercise of identifying ethically salient concerns in a clinical situation and framing and exploring the relevant bioethics principles can help open up the thinking of the clinician. Reflecting on related areas that influence ethical facets of the case, such as legal, gender, and cultural considerations, is also valuable. These efforts, taken together, represent an important ethical skill and the initial step in clarifying ethically sound strategies for optimal patient care.

Gathering Additional Information and Expertise

Ethical questions sometimes arise in a clinical situation when there is simply too little information or when clinicians are working at the boundary of their experience or scope of practice. Additional information or broadened and deepened expertise may lead to a ready resolution of the problems at hand. Data or expert guidance could come from the patient's prior medical records or clinical practice and ethics guidelines. In addition, speaking with an experienced supervisor, seeking advice from a colleague with specialized expertise, requesting an ethics or a legal consultation, and/or reviewing supplemental clinical data from outside sources may help.

Consultation-liaison psychiatrists see the value of additional data-gathering and consultation as resolving apparent ethical issues all of the time. Psychiatric consultation services have long fulfilled an important ethical function in most hospitals. An example is a patient who is declining potentially life-saving care for what appears to be an irrational reason and is demanding to return home against medical advice. A psychiatric evaluation may reveal an underlying condition, such as an addiction, delirium, early dementia, or poor pain control, interfering with the patient's decision making. Once this underlying condition is treated and the symptom burden lessened, the patient's decisional capacities may improve, leading to what appears to be a more rational decision to the care-giving team. If the patient is not experiencing a threat to his or her decisional capacity, it is important to explore the patient's values, background, and reasons for making a decision that contrasts with the medical recommendation. Cultural and religious preferences, fears based on past experiences, or issues related to family stresses may be influencing the decision. This additional background information may be very helpful in developing health-preserving strategies for this patient, such as consulting a religious leader, inviting emotional support from a family member or counselor, or identifying appropriate social service resources. In all of this, the caregiving team is wise also to obtain consultation from the legal service of the hospital. Information on the role of state or federal law in the case will help the physicians to determine whether temporary or long-term guardianship may be appropriate to pursue.

Thus, in approaching an ethically complex clinical situation, recognizing the ethical dimensions of a patient's care and then seeking additional information or expertise are the two initial steps to take in resolving dilemmas and developing ethically astute treatment plans.

Identifying Risky Situations and Arranging for Appropriate Safeguards

Many situations in patient care present predictable ethical challenges, and with experience, astute clinicians begin to anticipate when these ethically

risky situations are likely to occur. A clear illustration is when a patient is impulsive and the caregiving team must ensure the safety of the patient and others, for example, family members or the clinic staff. The clinician may be tempted to make choices that differ from usual clinical practices because of the situational threat. When the patient is a celebrity or an influential person, for example, other ethical risks can enter into the clinical picture. Staff members may be tempted to review the patient's electronic medical record, and care-giving physicians may begin to change their usual clinical practice, consciously or not, in order to garner favor with the special patient. A very different example is when there is a need to report child abuse or neglect. In this case, the psychiatrist may feel emotionally conflicted, wanting to abide by the law but feeling worried about the impact of the intervention on the therapeutic alliance with the family or about the impact of the investigation on the child. When entrusted with the care of vulnerable individuals, such as trauma survivors, children, or extremely ill patients, psychiatrists will recognize numerous ethical risks that may arise.

Such situations arouse concern, as they should, and signal the need for additional safeguards to ensure that an appropriate standard of care is upheld. Psychiatrists working in emergency departments, addiction centers, correctional settings, and employee health programs frequently encounter very difficult ethical dilemmas. Additional safeguards may include scheduling more frequent appointments for a suicidal patient, including a broader health care team, introducing alternative decision makers for a vulnerable patient, or obtaining permission from the patient to engage the family more fully. When a psychiatrist routinely works with high-risk patients, the clinical setting itself must have more safeguards. For example, a psychiatrist who manages a very difficult population and performs random drug screens as part of usual patient care may increase security for patient records and provide physical safety training for clinical staff. A geriatric psychiatrist who works with many individuals with advanced dementia may need to have a close working relationship with other clinicians, legal advocates, and the court in order to facilitate optimal patient care. In sum, heightened safeguards help address the ethical risk that is associated with psychiatric practice in all settings. When such risk is predictable because of the context or the patient population, ethically astute psychiatrists will build these safeguards into their routine practices, for the good of the patients and the caregiving team.

Working within One's Scope of Practice and Engaging in Continuous Learning

Psychiatrists acquire expertise beyond general medicine to include specialized knowledge in the care of individuals with mental disorders. Working within this scope of clinical competence is considered ethically important because of the risk of doing harm when a psychiatrist has insufficient expertise to care properly for a patient. If physicians are

confronted with patient care situations that are outside their scope of competence, then referral or, at the very least, consultation and supervision are appropriate. Sometimes physicians are in settings where no other provider is nearby; these situations warrant every effort to obtain additional expertise and services. In rural settings, telehealth services may bring additional expertise to the care of the patient. In training settings, close supervision of early-career physicians helps ensure an appropriate standard of care. In emergencies, even if clinicians are outside the scope of their usual practice, they are expected to do their best effort and certainly not to abandon the patient in extreme need.

Related to the ethical obligation to work within one's areas of competence is the responsibility to continue to learn. Remaining aware of advances in one's field is critical to optimal patient care and ensures that the clinician's scope of practice and domains of competence remain within community standards and, ideally, attain the state of the art. This imperative relates to the fundamental concepts of serving the best interest of the patient and avoiding harm in clinical care.

Reflecting on One's Values and Perspectives

Psychiatrists are taught to pay close attention to their responses to a patient, for example, how the patient looks, the story the patient tells, the emotions the patient demonstrates. Awareness of countertransference (the set of emotions and thoughts the psychiatrist brings to the therapeutic relationship from prior experiences) is valuable in clinical care, and it also provides insights into the ethical dimensions of a patient's care. Indeed, to conduct ethically sound patient care, it is very important for psychiatrists to be able to see how their experiences, attitudes, and feelings may influence their perceptions of the ethical aspects of a clinical situation. Awareness of one's reactions may help prevent ethical blind spots and even breaches of ethical conduct. When a psychiatrist finds a patient very attractive or when the patient makes the psychiatrist very uncomfortable or angry, these emotional cues should be seen as clinical and ethical signs that require acknowledgment, clinical interpretation, and a therapeutic response.

Ideally, team-based care and teaching settings will create opportunities to consider how different team members or supervisors and trainees view the ethical issues present in a patient care situation. Astute team leaders will facilitate explicit discussion of these issues and guide their colleagues in defining the conflicting ethical principles. Clarifying differences in the points of view of team members may prevent later misunderstandings. Demonstrating the importance of speaking further with the patient, gathering additional data, working within one's scope, introducing specialized expertise, and building in appropriate safeguards are also important for the patients cared for by the team or the trainee.

DECISION-MAKING STRATEGIES

Several proposed strategies for making decisions in patient care have value. Jonsen, Siegler, and Winslade created a four-part model that many health care settings have adopted: (1) clinical indications, (2) patient preferences, (3) quality of life, and (4) socioeconomic or external factors. In this model, the first of the four parts, clinical indications, has the greatest weight. It refers to the physician's obligations to diagnose a patient's condition carefully, inform and educate the patient about the illness process, consider the available treatment options in light of appropriate standards of care, make therapeutic recommendations that adhere to these standards, and ensure that treatments are provided competently and fulfilled properly in partnership with the patient. In psychiatry, this step is the same as in other areas of medicine.

The second element of this decision-making model, patient preferences, is grounded in the principles of respect for persons and autonomy. Once the physician presents the patient with recommendations based on what is clinically indicated, the physician elicits the preferences of the patient. In the practice of psychiatry, the nature of mental disorders may interfere with the ability of patients to identify or express their preferences or affect decisional capacity or authentic voluntarism. The ethically astute psychiatrist will do all that is possible to ensure that patients truly have voice in decisions affecting their care and will not coerce or exploit patients' potential vulnerability by imposing care that is refused. When a patient is unable to provide informed consent, appropriate decision-making safeguards, such as an alternative decision maker, should be put in place (in accordance with state law).

Quality of life is the third element in the Jonsen, Siegler, and Winslade model. It refers to the sense of personal satisfaction individuals express and/or experience with regard to their physical, mental, and social situations. Perhaps the most important consideration in assessing quality of life is direct engagement with patients regarding their values and points of view. The physician must not make assumptions about what represents a good life for the patient. Clinical indications and patient preferences are weighted more heavily than quality of life in the decision-making model. When overweighted or incorrectly assessed, quality of life could be used to justify decisions that do not meet current ethical standards in society.

The fourth element in this clinical decision-making model entails socioeconomic or external factors. These factors are diverse and include the interests of society, the role of family members, cultural factors, and costs and limitations in access to care. This element is last primarily because of the obligations of beneficence. The United States defines denying a patient life-saving treatment in an emergency on the basis of ability to pay as unethical and illegal, for example. In less acute situations, however, a discussion with patients about the costs of care and how these costs may

affect their choices of treatment is important and ethically acceptable.

EXAMPLES OF ETHICAL PRACTICE STANDARDS IN PSYCHIATRY

Several texts and resource documents provide comprehensive overviews of ethical practice standards in psychiatric practice. Key concepts related to several ethical practice standards for psychiatry are briefly outlined below and draw upon the schools of thought, biomedical ethics principles, and the four-part ethical decision-making model described in this chapter.

The Therapeutic Relationship and Professional Boundaries

The therapeutic relationship presupposes that the physician is worthy of the trust of the patient. The physician's expertise and the physician's intention to do good, avoid harm, and place the patient's best interest above all other interests are the basis of that trust. Professional boundaries are behaviors that align with these fundamental commitments to serve the well-being of the patient.

Interactions with patients that do not serve their interests but, instead, fulfill or gratify the interests of the physician can fracture or destabilize this foundation. One example is the ethical problem of intimate relationships arising between physicians and patients. The APA is explicit in its direction that psychiatrists must not engage in sexual relations with patients (including former patients). Such intimacies are inherently exploitative; they interfere with therapy, risk harm to the patient, and undermine confidence in the profession, thereby damaging the ability of others to benefit from psychiatric care. The profession of psychiatry sees sexual contact during therapy for whatever purpose, or ending therapy in order to start a romance, as indefensible.

Some have argued that the power differential between a former therapist and patient is the primary reason to forbid relationships between them. Many relationships, however, have power differentials because of the different earning capabilities of the partners, different levels of educational achievement, or a society that gives more power to one gender, one class, or one ethnicity. The reason to bar sexual relationships between psychiatrists and present and former patients is much more complex. One simple example is that this prohibition protects patients from being lured into a relationship that the patient would not otherwise desire because of genuine need for treatment. The prohibition also helps the psychiatrist to stay focused on the patient's well-being at all times, avoiding thoughts of a possible future relationship from affecting the conduct of the therapy. Epstein developed the Exploitation Index to help therapists identify potentially exploitative practices in the care of their patients.

Therapeutic boundary violations are behaviors that subordinate the interests of the patient to the interests of the physician. These violations

need not be sexual, although they are equally forbidden. Psychiatrists should not try to influence the political views of the patient, for example, and similarly, psychiatrists should not accept expensive gifts directly from patients. Engaging in business activities with a patient or former patient also risks misuse of influence and may represent exploitation of confidential knowledge obtained in therapy.

Therapeutic boundary crossings are behaviors that are intended to serve the interests of the patient but differ from usual ethical practices. Therapeutic boundary crossings should be viewed as exceptions but are often ethically acceptable. A psychiatrist disclosing personal information to build rapport with a distant and troubled patient or accepting a small gift at the holidays to demonstrate respect for a patient's culture and origin are two examples of boundary crossings.

Traditional views of the boundaries in the therapeutic relationship derive from reflection on long-term, dynamically oriented psychotherapy relationships, typically occurring in urban contexts. Novel questions arise when considering other roles, interventions, and settings. For example, is a child psychiatrist caring for impoverished and traumatized youth in a culturally unique community held to the same expectations as a therapist caring for an older, successful businessperson in a major metropolitan location? Are psychiatrists who see patients once in an emergency department held to the same standard of personal–professional role separation as their ambulatory clinic counterparts? Does a psychiatrist who is consulted about the suitability of a patient with Parkinson disease for neurosurgical intervention establish a therapeutic relationship with the patient in the same manner as with a patient needing treatment for depression? Intuitively, these situations require more nuanced responses than strict rules or inflexible codes of ethics. For this reason medical ethics is understood to evolve, by its very nature. This fact does not permit psychiatrists to deviate from expected standards but, rather, permits psychiatrists to apply standards (and, more important, their underlying intent) to novel situations.

Confidentiality

The importance of confidentiality in tending to illness dates back to Hippocrates, who said, "What I may see or hear in the course of the treatment or even outside of the treatment in regard to the life of men, which on no account one must spread abroad, I will keep to myself, holding such things." Confidentiality refers to the physician's duty to not release information learned in the course of treatment. Salient information may relate to disclosures by the patient, observations of the physician, or laboratory findings or notes from prior treatment in the medical record. Protection of patient confidentiality fosters the therapeutic alliance and is a prerequisite for patients to be willing to speak more freely and openly with

their doctors. If confidentiality were not ensured, individuals would be less likely to seek help and would tend to withhold crucial information, diminishing the effectiveness of the treatment. Table 60.2–3 notes illustrations of optimal practices for protecting confidentiality in psychiatric practice.

Psychiatric records, including even the identification of a person as a patient, must be protected with particular diligence. A psychiatrist may release confidential information only with the authorization of the patient or when compelled by law. Even when patients have given permission for their personal information to be used, safeguards are still needed. It is important to disclose only that information which is truly necessary for the situation at hand. In teaching and research one should not unnecessarily release patients' names or information that might allow others to identify them. Presenters of patient material in ward rounds and case conferences should remind attendees not to repeat what they hear. Confidentiality and the ethical obligation to withhold information endure even after a patient's death.

Table 60.2–3.
Examples of Optimal Approaches to Confidentiality in Clinical Practice

- Provide accurate information to patients about the realities of confidentiality in the clinical setting.
- When appropriate, remind patients that the purpose of the medical record is to be shared as a way of providing optimal care by a team of health professionals.
- Openly discuss with patients the legal limitations of confidentiality and avoid assuring patients that their disclosures will be "completely" confidential.
- Do not speak to others or, if working in a team-based care setting, avoid talking with others beyond the health care team, about patients' care without explicit, written permission to do so.
- Actively work within organizations and professional societies to develop state-of-the-art confidentiality safeguards as science, clinical care, health policy, the law, and health care systems evolve.
- Seek consultation and expert guidance when confidentiality safeguard questions arise.
- Model best practices in confidentiality to colleagues, team members, and trainees.

New ethical questions arise every day regarding confidentiality in novel forms of treatment, whether medical or psychosocial. For instance, what are appropriate privacy protections for psychotherapy occurring by telephone, video, the Internet, electronic mail, or real-time texting? Are current practices sufficient for new models of care involving teams, networks, and health care systems, which might disseminate information electronically across integrated national and perhaps international sites? How will innovators in clinical research and genetics protect the personal information of specific individuals whose biomarkers, taken together, represent a unique biological signature?

Informed Consent

Informed consent is an ethical and legal standard that exemplifies the principle of respect for persons and seeks to bring the principle of autonomy to life in clinical medicine. Informed consent is the idea that capable and knowledgeable individuals may freely choose a course of action pertaining to their health. Informed consent can involve accepting a recommended treatment or choosing an alternative, including no treatment (i.e., informed refusal).

Informed consent involves three components: information-sharing, decisional capacity, and voluntarism. With respect to information-sharing, physicians should provide patients accurate and balanced information regarding (1) the nature of their illness or condition; (2) the need for intervention; (3) the anticipated risks or benefits and their relative likelihood, including rare but very serious risks; (4) appropriate alternatives and their attendant risks and benefits; and (5) future consequences of the different choices, to the extent possible. Physicians should offer this information at the level and in a language that the patient understands and in a nonpressured environment. Physicians should build opportunities to clarify points of information and to correct misimpressions into the conversation. Individuals who cannot provide consent must receive information at the level at which they can understand it, and physicians should seek the individuals' acceptance, when possible. See Table 60.2–4.

Grisso and Appelbaum characterized decisional capacity as having four subcapacities. The first is the physical and mental ability to communicate. This subcapacity does not refer to the ability to speak a particular language but, rather, the ability to express—if only through eye blinks or hand signals—a choice. The second subcapacity is the ability to understand relevant factual information. The third subcapacity is the ability to reason, that is, to weigh the different factors and risks and to think through the choice at hand in an objective and flexible manner. The fourth subcapacity is to appreciate the meaning of the decision and to evaluate the choices in the context of one's life. This capacity is close to the concept of insight, which usually refers to patients' ability to understand the nature and impact of their illness. The capacities should be viewed as a continuum, not simply as present or absent. Many psychiatric conditions erode the cognitive strengths that underpin these capacities.

Voluntarism is the capacity that relates to one's ability to act freely, in the absence of coercion, in expressing one's authentic wishes. "Deliberateness, purposefulness of intent, clarity, genuineness, and coherence with prior life decisions" are key elements of the capacity for voluntarism.

Informed consent standards are more rigorous for high-risk decisions. This concept is referred to as the "sliding scale" for informed consent. When patients make a choice that may endanger them (e.g., choosing one

form of chemotherapy over another), the standard for decisional capacity is higher than when they make a choice that poses minimal or virtually no risk (e.g., having their blood pressure taken). Because declining recommended treatment usually entails heightened risk, the standard for informed refusal also is higher. Physicians must give extra care in providing information and counseling patients regarding alternatives and in verifying decisional and voluntarism capacity if a decision under consideration, whether accepting or refusing an intervention, is high risk.

Table 60.2–4.
Examples of Optimal Approaches to Informed Consent in Clinical Practice

- Provide accurate information regarding the nature of the illness and the need for clinical intervention; anticipated risks and benefits of the proposed intervention, including rare but serious risks; alternatives to the proposed treatment and their likely outcomes, including no intervention.
- Avoid using overly complicated language and overwhelming patients with too much technical information.
- Be attentive to patients' comfort, communication style, and interpersonal cues.
- Invite patients to ask questions and double-check their understanding of what was explained.
- Provide additional information if requested by patients.
- Use translators when appropriate and, if desired by patients, include family members in the information-sharing session(s).
- Supplement the information-sharing dialogue with written materials as appropriate.
- Evaluate each of the four components of decisional capacity: capacity to communicate, capacity to understand, capacity to reason, and capacity to appreciate.
- If symptoms or stressors interfere with optimal decision making, seek to ameliorate patients' discomfort to help support their ability to approach the decision process.
- Invite patients to express their preferences.
- Explore the reasons and hopes behind these preferences.
- Provide emotional reassurance and eliminate, to the extent possible, coercive pressures that patients may be feeling.
- Clarify patients' values and patterns of past personal decisions.
- Create a context of nonjudgmental support and open and ongoing communication with patients.

When an individual is not capable of informed consent or refusal, others who are legally authorized may make decisions for the patient. In some cases, the patient may designate the alternative decision maker. For young children, a parent or guardian may make the decision. At times, a formal legal process must identify a suitable decision maker. State laws differ in how to identify decision makers and what direction decision makers receive, such as whether to make the decision that appears to be in the best interests of the patient or the decision that most closely resembles what the patient would choose, if it were possible (i.e., "best interests" vs. "substituted judgment" approaches).

Overlapping Roles and Conflicts of Interest

Differing and overlapping roles are common for professionals who serve society as clinicians and also as teachers, consultants, administrators, and

public policy experts. Having more than one role generates ethical questions and often ethical tensions because the duties of each role, by definition, do not align perfectly. (Having more than one role creates dual role, or dual agency, problems.) When roles receive compensation from different sources, issues also arise in relation to competing interests and divided loyalties. Many controversies related to conflicts of interest have become a focus of governmental policy and media reports. Ethically, conflicts of interest matter greatly in that patient needs may be subordinated to the interests of physicians who are motivated by greed rather than serving patient well-being. Conflicts of interest occurring at multiple levels—among individual physicians, institutions, and systems of care and research—have eroded public trust in the medical profession.

To illustrate, dual agency conflicts may arise for psychiatrists who work in the military, occupational health programs, or correctional settings. For example, psychiatrists may be called on to use their expertise to make a recommendation on a patient's fitness for duty, competency to stand trial, or disability. As a result of the evaluation, the patient ultimately may be returned to front-line military or police work or to prison or may receive (or not) financial compensation. The ethical question of what decision serves the best interests of the patient becomes moot in these situations in which the other duties of the psychiatrist, such as protecting the public, become predominant. Particularly vexing dilemmas arise in the correctional setting, where psychiatrists may be called on to treat patients against their will when treatment would improve their disorder at the expense of their interests. Psychiatrists may not ethically force treatment on patients to make them competent to be executed. The issue is less clear concerning making patients better able to stand trial when doing so might lead them to spend more time in jail. Honesty, transparency, objectivity, and rigor—always important—are especially salient in the context of dual agency. The key issue to ethical behavior in such situations is to provide accurate information about the nature of the evaluation process, including information about the physician's responsibilities beyond those to the patient and about how usual expectations regarding patient confidentiality are not in force. Before the examination begins, physicians must notify the individuals explicitly about how they will disclose or report the information and conclusions from the evaluation.

Physicians can manage many conflicting roles and conflicts of interest in professional roles through the conscientious use of safeguards such as role separation, recusal, financial limits, responsibility limits, and additional supervision or oversight. Because the public trust is at stake in issues of conflict of interest, broad dialogue and scrupulous attention to emerging guidelines and principles are very important for all fields of medicine, including psychiatry. Integrity—namely, intending good, being transparent, adhering to professional standards, and simply behaving in a

manner that is worthy of trust—is key. Psychiatrists, moreover, should play an active role in educating the public on the value of professionals participating in broad activities in society, such as research, education, administration, and policy making, in order to bring the greatest positive value of medical expertise to broader society.

Psychiatric Opinions

In 1964, Barry Goldwater was the Republican candidate for president of the United States. *Fact* magazine sent a questionnaire to US psychiatrists asking whether Goldwater was psychologically fit to serve as president. Despite any direct knowledge of Goldwater derived from an in-person interview, approximately half of the psychiatrists who responded to the survey said he was not fit to serve in the role. The magazine then published an article proclaiming, "1189 Psychiatrists Say Goldwater is Psychologically Unfit to be President." In response, the APA asserted that it is unethical to render psychiatric judgments about a person in the media.

It is clearly unethical for psychiatrists to render a definitive psychiatric opinion about someone for whom they have only very limited information or only a biased source of information, such as the media. The practice appears nevertheless to be quite common. Invoking one's expertise as a psychiatrist without full knowledge of a patient is misleading and inappropriate. It is, however, ethical to render an opinion about someone one has not evaluated if one has an adequate source of unbiased information and is prudent in stating the limits of one's knowledge. In fact, psychiatrists commonly render opinions about people they have never met. When senior psychiatrists supervise residents or other mental health practitioners, they render opinions and make suggestions on treatment although they may not have met the person in question. Psychiatrists are sometimes asked to render an opinion on testamentary capacity—whether someone who is now deceased had the capacity to make an informed decision and whether the person was subject to undue influence. Moreover, esteemed psychiatrists have studied political leaders and historical figures and thereby advanced the understanding of history and human behavior. Similarly, it is ethical for psychiatrists to reanalyze the data of a neutral forensic evaluator and make conclusions about parties they have not interviewed if there is adequate information in the neutral forensic evaluator's report to do so.

Thus, the ethics of rendering an opinion about a person whom one has not examined depends on several factors. Key issues include the purpose of rendering the opinion, the quality of information, and the reason for not interviewing the person.

Responsibilities to Third Parties

Several court rulings have stated that psychiatrists have a responsibility to third parties. These concerns are more properly forensic in nature but are relevant to good ethical practices because of the ethical notion of respect for the law. Court rulings create case law, not ethics. Nevertheless, case law leads to new expectations and indicate society's view of professional responsibilities. Legal rulings also have value because they point attention to issues that need deeper consideration by ethicists. In the original and subsequent *Tarasoff* decisions, for example, the courts upheld the position that psychiatrists must act overtly to protect the community. Courts have held psychiatrists liable when their patients, or former patients, have injured or killed people with guns or by reckless driving (e.g., 1983 *Jablonski v. U.S.;* 1988 *Naidu v. Laird;* 1980 *Lipari v. Sears, Roebuck and Co.;* 1983 *Petersen v. Washington*). Subsequent legislative rulings have limited psychiatrists' liability in such matters. Psychiatrists who provide incorrect information place themselves at risk if such action harms the interests of a third party (e.g., *Molien v. Kaiser Foundation Hospitals* and *Ramona v. Isabella*). When clinicians give direct suggestions that affect third parties, such as advice to get a divorce (rather than limit interventions to helping patients evaluate their choices in the context of their lives), clinicians leave themselves liable for the consequences. The liability is increased if they become directly involved by participating in a confrontation with the third party.

ETHICS AS FUNDAMENTAL TO MEDICAL PROFESSIONALISM

Service to society is the *raison d'être* of the professions. In medicine, physicians provide expertise to advance human health and to lessen the burden of suffering associated with illness. Medical professionals, including those in the field of psychiatry, have the opportunity to develop distinct expertise to help others. These professionals have responsibility for creating and applying new knowledge, training others, and participating in leadership throughout the profession and broader society. These professionals also have the duty to ensure that colleagues uphold the clinical and ethical standards of their disciplines. Self-care and vigilance to avoid emotional exhaustion and impairment are thus important responsibilities of professionals. In return, physicians, like other professionals, are entitled to fair compensation for the work they perform in their special societal role. Part of the social contract of each profession is that professionals will, at a minimum, live up to the ethical commitments of the field. More deeply, members of the profession will seek progress ethically as well as in the scientific basis and clinical practices of each field and discipline in the profession.

ACKNOWLEDGMENT

The author wishes to acknowledge the work of Dr. Roy H. Lubit, who wrote

this chapter in the 9th edition of the *Comprehensive Textbook of Psychiatry,* part of whose work was adapted for this chapter.

References

*American Medical Association. *Code of Medical Ethics* Chicago, IL: American Medical Association; 2016. Available at https://www.ama-assn.org/about-us/code-medical-ethics. Accessed January 17, 2016.

*American Psychiatric Association. *The Principles of Medical Ethics With Annotations Especially Applicable to Psychiatry 2013 Edition*. Washington, DC: American Psychiatric Association; 2013. Available at http://psychiatry.org/practice/ethics/resources-standards. Accessed June 22, 2015.

Appelbaum PS, Grisso T. Assessing patients' capacities to consent to treatment. *N Engl J Med*. 1988;319:1635–1638.

Baker RB, McCullough LB. *The Cambridge World History of Medical Ethics*. Cambridge: Cambridge University Press; 2008.

Beauchamp TL, Childress JF. *Principles of Biomedical Ethics*. 7th ed. New York: Oxford University Press; 2012.

Bloch S, Green. *Psychiatric Ethics*. 4th ed. Oxford: Oxford University Press; 2009.

Coverdale JH, McCullough LB. Professionalism in medicine as a transnational and transcultural ethical concept. In: Roberts LW, Reicherter D, eds. *Professionalism, Ethics, and Cultural Sensitivity: A Study Guide for Health Professionals*. New York: Springer Science+Business Media, LLC; 2015.

Epstein RS, Simon RI. The Exploitation Index: an early warning indicator of boundary violations in psychotherapy. *Bull Menninger Clin*. 1990;54(4):450–465.

Gabbard GO, Roberts LW, Crisp-Han H, Ball V, Hobday G, Rachal F. *Professionalism in Psychiatry*. Arlington, VA: American Psychiatric Publishing, Inc.; 2012.

Gelso CJ, Hayes JA. *Countertransference and the Therapist's Inner Experience: Perils and Possibilities*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.; 2007.

Gregory J. *Lectures on the Duties and Qualifications of a Physician*. London: W. Strahan and T. Cadell; 1772. Reprinted in McCullough LB: *John Gregory's Writings on Medical Ethics and Philosophy of Medicine*. Dordrecht: Kluwer; 1998.

Jonsen AR. *Short History of Medical Ethics*. New York: Oxford University Press; 2008.

*Jonsen AR, Siegler M, Winslade WJ. *Clinical Ethics: A Practical Approach to Ethical Decisions in Clinical Medicine*. 8th ed. New York: The McGraw-Hill Companies, Inc.; 2015.

Kant I. *Fundamental Principles of the Metaphysics of Morals [1785]*. Amherst, MA: Prometheus Books; 1990.

Kushe H, Singer P. *A Companion to Bioethics*. Oxford: Blackwell; 1998.

Lidz CW, Appelbaum PS. The therapeutic misconception: problems and solutions. *Med Care*. 2002;40(9 Suppl):V55–V63.

Lo B. *Resolving Ethical Dilemmas: A Guide for Clinicians,* 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2013.

Mill JS. *Utilitarianism [1861]*. London: Everyman's Library; 1996.

*National Institutes of Health. *The Belmont Report: ethical principles and guidelines for the*

protection of human subjects of research. 1979. Available at http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html. Accessed June 22, 2015.

Percival T. *Medical Ethics: Or a Code of Institutes and Precepts, Adapted to the Professional Conduct of Physicians and Surgeons*. London: J. Johnson & R. Bickerstaff; 1803.

Rawls J. *Theory of Justice*. Cambridge, MA: Harvard University Press; 1971/1999.

*Roberts LW. A Clinical Guide to Psychiatric Ethics. Arlington, VA: American Psychiatric Association Publishing; 2016.

Roberts LW, Hoop JG. *Professionalism and Ethics: Q&A Self-Study Guide for Mental Health Professionals*. Washington, DC: American Psychiatric Publishing, Inc.; 2008.

Roberts LW, Reicherter D. *Professionalism and Ethics in Medicine: A Study Guide for Physicians and Physicians-in-Training*. New York: Springer Science+Business Media, LLC; 2015.

Spece RG Jr, Shimm DS, Buchanan AE. *Conflicts of Interest in Clinical Practice and Research*. New York: Oxford University Press; 1996.

Sperry L. *Dictionary of Ethical and Legal Terms and Issues: The Essential Guide for Mental Health Professionals*. London: Routledge; 2007.

U.S. Department of Health and Human Services. *Code of Federal Regulations, Title 45, Public Welfare, Part 46, Protection of Human Subjects;* Revised January 15, 2009. Available at http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html. Accessed June 22, 2015.

▲ 60.3 Correctional Psychiatry

HENRY C. WEINSTEIN, M.D., AND JOSEPH V. PENN, M.D.

INTRODUCTION

Correctional psychiatry occupies a unique place in medical practice. As in all matters "forensic" it functions at the boundaries of law and medicine. As a subset of forensic psychiatry, it is fraught with potential bias and conflicts of interest. The shadows of issues of ethics and morality are ever-present.

Multiple factors, including legal, social, fiscal/health care reimbursement, and political, over the past 50 years have led to the current substantial presence of inmates with psychiatric and substance use disorders in the correctional systems of the United States.

Some examples of these factors include the downsizing of public mental health inpatient facilities, "zero tolerance" arrest and sentencing policies as part of the "war on drugs," and the deinstitutionalization of many individuals with chronic and persistent mental illness.

One of the unintended consequences is what is now referred to as the "criminalization of mental illness." Mental and substance-related disorders, chronic medical conditions, and infectious diseases such as human immunodeficiency virus (HIV) and hepatitis B and C are significant public health problems affecting individuals in US correctional facilities.

It is often said that American jails, prisons, and other correctional settings have become the default mental health system for this high-risk population. In particular, the sheer number of inmates with severe psychiatric disorders in the US prison system and the "revolving door phenomena" in which many mentally ill people move continuously between homelessness and the criminal justice system represent a national public health crisis. There is a similar risk of reincarceration among prisoners with co-occurring severe mental illness and substance use disorders.

This chapter will present an overview of American correctional systems and make recommendations for the psychiatric assessment, treatment, and consultative, administrative, and other roles and possible ethical challenges that a psychiatrist may face practicing in correctional settings.

An estimated 6,899,000 persons were under the supervision of adult correctional systems at year-end 2013. Of these, there are more than 2.2 million inmates in US jails and prisons. It is estimated that of these, 350,000 have severe and persistent mental illness, that is, carry a diagnosis of a serious mental disorder such as schizophrenia, bipolar disorder, or severe depression.

There are many different types of correctional facilities. These include lockups, jails, state and federal prisons, US Immigration and Customs Enforcement (ICE) detention centers, Bureau of Indian Affairs detention centers, US military prisons, jails, and detention centers. There have been recent concerns regarding access to care and other medicolegal issues for persons with severe mental illness in the US ICE system.

The most common correctional settings for adults are jails and prisons. Jails are city- or county-operated correctional facilities that confine individuals involved in the criminal justice system who are awaiting trial or serving short sentences for misdemeanors. There are approximately 3,500 jails in the United States. Detention refers to the period following arrest in which an individual is held in secure custody prior to or after court proceedings. There is a great deal of variety regarding the duration of time that individuals remain in jail (e.g., hours, days, weeks, typically 1 year or less). Some individuals may remain in jail for years for more serious charges while awaiting legal disposition of their case, or they may serve their sentence in a jail setting. There is significant variability in size with some small rural jails holding 10 to 50 detainees and some major metropolitan jails holding 7,500 to 10,000 detainees.

A prison is a facility operated under the auspices of a state or federal government for the confinement of adults convicted of a felony whose sentences generally exceed 1 year. There are approximately 1,300 state and federal prisons in the United States. Individuals in prisons may be serving one or more life sentences, and some states have the death penalty.

Similar to adults, mental and co-occurring substance-related disorders are significant public health problems affecting youth in juvenile justice

settings. There has been a significant increase in the need for psychiatric and mental health services for youth in the juvenile justice system. While involved with the juvenile justice system and family courts, juveniles may be housed within juvenile detention centers, reception or diagnostic centers, community-based or other residential treatment programs, or juvenile correctional facilities. Some juveniles may be held in county jails, and some youth may be transferred/sentenced to adult prison settings. In many states, certain felony offenses committed by adolescents are moved to adult criminal court. This practice has led to an increase in the detention and incarceration of younger offenders in adult jails and prisons. Data reflect rates of SMI among incarcerated adolescents that are substantially above those found in the community.

HISTORY OF CORRECTIONAL PSYCHIATRY IN THE UNITED STATES

The psychiatric care of the mentally ill in jails and prisons has a long history in the United States. Benjamin Rush (1745–1813), who wrote the first general book on psychiatry in America in 1812, *Medical Inquiries and Observations Upon the Diseases of the Mind*, and is regarded as the father of American psychiatry, wrote extensively on the treatment of deviant behavior. Rush deemed public punishment, which was common at the time, to be counterproductive and proposed private confinement, labor, solitude, and religious instruction for criminals. His outspoken opposition to capital punishment contributed to the Pennsylvania legislature's decision to abolish the death penalty for all crimes other than first-degree murder. He founded the Philadelphia Society for Alleviating the Miseries of Public Prisons, the first prison reform group in the world. This group survives today, more than two centuries later, and is now called the Pennsylvania Prison Society. It promotes correctional reform and social justice.

However, until recently, in regard to conditions in jails and prisons, courts maintained a "hands-off" policy, a doctrine that regarded the inner workings of correctional facilities as beyond the responsibility and authority of the judicial system and best left to correctional professionals.

The attitude of the courts toward prisoners was epitomized by a statement in a 19th-century case in the state of Virginia: "A convicted felon is one whom the law in its humanity punishes by confinement in the penitentiary instead of death… for the time being, during his term of service in the penitentiary, he is a slave of penal servitude to the State… for the time being, the slave of the State" (*Ruffin v. Commonwealth*).

It was during the 1970s, during the tidal wave of the civil rights movement, that various studies and court cases exposed the abominable conditions of health and mental health care in the jails and prisons of the United States. These disgraceful findings lead to the reform of the correctional system and an accelerating sea change in correctional

psychiatry and health care.

The confluence of a number of interrelated forces spearheaded this reform, including professional organizational advocacy by the American Medical Association (AMA) and the American Psychiatric Association (APA), "conditions of confinement" lawsuits, which were now entertained by the courts, and the development of standards and guidelines for health and mental health care in jails and prisons.

In 1974, the AMA carried out a survey of sheriffs administering local jails. It revealed that nearly two thirds of the jails' correctional health care consisted of no more than first aid (16.7 percent reported that not even first aid was available). This survey also found that fewer than 20 percent had facilities for caring for alcoholics, only 10 percent had any facilities for inmates with drug addiction, and only 14 percent had facilities for inmates with mental illness. Studies of prison correctional health care revealed that prison correctional health care was no better than that in jails. It was also clear that almost all jails and prisons lacked mental health–screening mechanisms and testing services, and, therefore, inmates' mental health problems went undetected and untreated.

In the same year, 1974, the APA published a Position Statement prepared by the APA Committee on Psychiatry and the Law (Table 60.3–1). In 1975, the Federal Law Enforcement Assistance Administration (LEAA) provided a grant to the AMA to upgrade correctional health care. The purpose of the grant was to develop a model correctional health care system and to establish a clearinghouse to develop and to distribute information on correctional health care issues. Another important goal was the development of a national accreditation program with correctional health care standards. In 1977, the AMA published its first jail correctional health standards; 16 jails were accredited for compliance that year. In 1979, the prison and juvenile correctional health care standards were published.

At the same time, the judicial system was establishing the unique legal context of psychiatric services in jails and prisons—that such services are constitutionally guaranteed. In 1976, the US Supreme Court, in the case of *Estelle v. Gamble*, held it unconstitutional for prison officials to be deliberately indifferent to the serious medical needs of those in their custody, and in 1977, a federal court, in the case of *Bowring v. Godwin*, extended this requirement to psychiatric care.

Table 60.3–1.
American Psychiatric Association Position Statement on Medical and Psychiatric Care in Correctional Institutions (1974)

The American Psychiatric Association recognizes that so-called correctional institutions are established by various levels of government. Their ability to provide a full range of medical (including psychiatric) services varies according to a multitude of factors: e.g., public opinion, funding, and administrative structure in relation to the goals set for a particular institution, such as punishment, containment,

deterrence, rehabilitation, and/or treatment.

An essential part of a minimum medical care delivery system consists of the early detection, diagnosis, treatment, and prevention of psychiatric illness. Many "correctional" facilities are not so structured as to supply this care delivery system even minimally.

It is never part of the penalty imposed to deprive a prisoner of adequate medical (including psychiatric) care. Furthermore, it is never proper to use medical (including psychiatric) care for punitive and coercive purposes or as an instrument of social repression. The fact of incarceration imposes upon public authority the special duty to provide adequate medical services, including psychiatric services. Availability of such services is and should be a right of each incarcerated individual.

In 1983, as an outgrowth of the LEAA-funded AMA program, the National Commission on Correctional Health Care (NCCHC) was established. The NCCHC consists of 37 professional organizations, including the AMA and the APA. It assists correctional institutions in upgrading their correctional health care services and, most important, has established correctional health care standards and correctional health care accreditation processes. These standards, which are revised regularly, represent the state of the art in correctional health care, and with the 2007 publication of the NCCHC's Standards for Mental Health Services in Correctional Facilities, the state of the art, as well as evidence-based practice, for mental health services in jails and prisons.

In 1989, the APA published Psychiatric Services in Jails and Prisons. This document, referred to as the APA Guidelines, was developed by correctional psychiatrists for correctional psychiatrists and describes in detail the principles governing the delivery of psychiatric services in jails and prisons; the specific requirements for an adequate process of screening, evaluation, and treatment; the specific roles of the correctional psychiatrists; and the application of the principles and guidelines to specific populations. The second edition of the APA Guidelines was published in 2000, the third edition, now titled *Psychiatric Services in Correctional Facilities*, was published in 2016.

There have been some recent efforts to pursue alternatives to incarceration of persons with severe mental illness and substance abuse disorders. One of the most promising alternatives is the diversion model, designed to prevent or minimize detention of the nonviolent mentally ill and to divert persons with severe mental illness to appropriate community-based mental health services in lieu of incarceration. Postbooking programs such as drug courts and mental health courts that are similarly designed to divert nonviolent mentally ill individuals from jails and prisons into closed supervised treatment programs have also been particularly promising.

Correctional Psychiatry as Community Psychiatry

Each year correctional facilities release approximately 12 million inmates. Because society inevitably shoulders the cost of their treatment, the period

of incarceration is an opportunity to create an infrastructure to identify and to treat these patients. In a biosphere in which life is interdependent, the mental health status of the least members of society strongly influences the mental health status of the greatest members of society. Effective correctional psychiatric systems will have a positive influence on the public's mental health. So, in recent years, it has become increasingly evident that correctional psychiatry must be considered an integral part of community psychiatry and that correctional health must be conceptualized as an important aspect of public health.

The title, "The Health Status of Soon-to-Be-Released Inmates," indicates how the NCCHC's project and publication focused on this insight. It stated:

> Prisons and Jails offer a unique opportunity to establish better disease control in the community by providing improved health care and disease prevention to inmates before they are released. Implementing the recommendations in this carefully researched report will go a long way toward taking advantage of this opportunity and contribute significantly to improving the health of both inmates and the larger community.

Many systems are implementing innovative community reentry initiatives and/or expanding discharge planning or continuity of care programs for mentally ill inmates who are about to be released. These programs attempt to link inmates shortly prior to or at the time of release into long-term community-based outpatient services to help them manage their mental health problems and reduce their risk of recidivism.

A Career in Correctional Psychiatry

The practice of correctional psychiatry can provide a wide variety of rewarding opportunities—to work within a unique system that has multiple challenging patients, access to care issues, the balance between safety, custody, and control, other correctional barriers and challenges, and other intersecting system pathologies usually not met in other contexts, to open up a new world of interest and interesting colleagues and correctional professionals, to learn about this critical social phenomenon, and to participate in this special aspect of public psychiatry and public service.

The correctional psychiatrist's expertise and role as a diagnostician, consultant, and team leader, and the ability to evaluate and diagnose individuals with serious mental illness, differentiate medical illnesses or comorbidities from psychiatric disorders, and consult to medical and administrative and correctional staff and leadership are all essential roles and functions. Although in some correctional systems, a psychiatrist's role may be limited to diagnostic evaluation and medication treatment and management, correctional psychiatrists have the opportunity for other roles in the provision of correctional mental health care—such as screening, evaluation, suicide prevention, other health care and correctional staff

education and training, quality reviews, working on improving clinical outcomes, and systems and process improvement.

The psychiatrist's knowledge and familiarity with the ever-growing and increasingly complicated use of psychotropic medication(s) and diagnostic abilities make him or her an indispensable member of the facility team and an important component of the correctional facility's security apparatus. Psychiatrists who practice in a correctional setting need to have an awareness and understanding of the operations of the facility and the issues affecting it, including access to care, continuity of care, and the interface with multiple systems (e.g., police, probation, courts, social services, community and state systems for offenders with serious mental illnesses). Effective psychiatric treatment and consultation in correctional settings requires knowledge of the organizational structures, policies, procedures, and other systems issues relevant to mental health issues and the daily schedule and "culture" and ongoing issues in the institution.

Psychiatrists working in correctional settings must be vigilant in regard to their personal safety and security issues and aware of actions that may compromise their safety and/or the safety and confinement of incarcerated individuals. Before entering a correctional facility, the psychiatrist should become aware of the type and function of the correctional facility and custody issues, what different issues may arise when attempting to evaluate or provide care to individuals with differing custody levels, and also in the event of fire alarm, altercation, riot, and hostage negotiation. The correctional psychiatrist's variety of roles in the provision of essential services can go beyond the direct care of patients. This may might involve consultation on processes, protocols, the decision of which instruments to use, and other policy review and improvement. The goal of these initiatives is to improve the processes and functioning of the staff and the facility. There may be opportunities for teaching of other mental health professionals and especially the cross-training of security staff and the opportunities to learn about the subtleties of security and respect for the professionalism of the correctional profession. As more academic departments recognize the value of clinical rotations and electives in jails and prisons, further opportunities for leadership in supervision and teaching of medical students, psychiatric residents, forensic psychiatry fellows, and other mental health professionals arise.

A career in correctional psychiatry can represent the development of an area of expertise that can lead to opportunities to help lead and guide this new field. One can become an "authority" in correctional psychiatry, help guide and grow the field, and serve as a presenter at state, national, and international conferences. There are opportunities to serve as a physician surveyor for the NCCHC and other correctional accreditation organizations. There appears to be a growing amount of litigation in state and federal courts. Thus, there is an opportunity for psychiatrists to

participate in litigation and testify as an expert witness. Some courts rely on psychiatrists to serve as court monitors within correctional systems.

Finally, there is the opportunity for the correctional psychiatrist to help educate the public, media, judges, lawyers, probation, courts, state leaders, and other individuals on the growing mental health and medical needs within correctional settings. There are many opportunities to work collaboratively and proactively with local, state, and national lawmakers, and governmental leaders and policy makers, assist with new legislation, serve as resource witnesses during legislative sessions, and help identify and secure sources of funding of additional staff and adequate psychiatric and mental health services for persons with mental illness and other vulnerable populations within jails and prisons. Participation in this laudable effort can give special meaning to working with this often-neglected and high-risk patient population. This proactive work is akin to risk management in health care and is particularly important due to an increasing trend toward litigation against jails, prisons, and other correctional systems and their leadership and at times individual correctional health care providers or organizations in state and federal courts.

THE LEGAL AND ETHICAL PRINCIPLES UNDERLYING THE PRACTICE OF CORRECTIONAL PSYCHIATRY

The Legal Context of Correctional Psychiatry

In 1970s, the crucial years of the reform of the correctional system, a series of court decisions established the principle that psychiatric services in the jails and prisons of the United States are constitutionally guaranteed. This legal doctrine holds that under the Eighth Amendment of the US Constitution, which prohibits "cruel and unusual punishment" whenever a state or the federal governmental entity takes custody of a person, it must provide for the necessities of life that the person otherwise is unable to obtain. These include food, clothing, shelter, and medical care, which include psychiatric care.

In 1976, the Supreme Court, in the case of *Estelle v. Gamble*, held it unconstitutional for prison officials to be deliberately indifferent to the serious medical needs of those in their custody, and in 1977, the federal court case of *Bowring v. Godwin* extended this requirement to psychiatric care. In 1980, another federal court case, *Ruiz v. Estelle*, listed a number of specific components that are required:

1. Participation by trained mental health professionals
2. Systematic screening and evaluation
3. Safeguards against psychotropic medications that are prescribed in dangerous amounts, without adequate supervision, or otherwise inappropriately administered

4. Treatment that is more than mere seclusion or close supervision
5. Accurate, complete, and confidential records
6. A suicide prevention program

In 1989, the case of *Langley v. Coughlin* listed a number of other specific criteria that, if not met, could provide a basis for successful inmate legal claims. Many of these criteria explicitly required providing for so-called secondary or supportive rights, such as taking psychiatric histories and maintaining medical records, in recognition of the fact that such rights are required to facilitate the primary right to diagnosis and actual care. *Langley v. Coughlin* also explicitly requires that care not be limited to psychotropic medication.

Madrid v. Gomez, in 1995, a State of California Department of Corrections legal case, added several factors that may determine the constitutionality of a correctional mental health system:

1. An inmate must have a means of making his or her needs known to the medical staff.
2. Sufficient staffing must allow individualized treatment of each inmate with serious mental illness.
3. An inmate must have speedy access to services.
4. There must be a system of quality assurance.
5. Staff must be competent and well trained.
6. There must be a system of responding to emergencies and to prevent suicides.
7. Inmates also have a right to a hearing and procedural protections before involuntary psychiatric hospitalization or forced administration of antipsychotic medication.

It should also be noted that in addition to these constitutional mandates, jail and prison mental health services may be required to meet other legal requirements. These include federal laws, such as the Americans with Disabilities Act, and various federal regulations, as well as state requirements established on the basis of state law through constitutions, statutes, and regulations. In addition, of course, correctional psychiatric services are always subject to other legal liabilities, such as for medical malpractice. Thus, individuals and systems may be sued in state or federal court, depending upon a variety of factors, either for malpractice, deviation from the standard of care, or alternatively under Federal constitutional 1983 actions, for example, “deliberate indifference,” and failure to afford access to care and/or continuity of care. However, it must be noted that the constitutional guarantee of mental health care in correctional facilities does not extend to treatment for substance use disorders. The position of the courts has been that the provision of care and treatment in cases of substance use disorder (e.g., for drug abuse) in jails and prisons is not considered an aspect of psychiatric care and, therefore, is not considered as

constitutionally required. However, there is such an obvious need for treatment of substance use disorders in jails and prisons that it is usually included in correctional programs.

The Ethical Context of Correctional Psychiatry

Although psychiatrists practicing in correctional settings face complex and often unique ethical problems, they are bound by the same body of ethical principles as psychiatrists practicing in any setting to follow the standards of professional ethics as set out in the APA's Annotations Especially Applicable to Psychiatry to the AMA's Principles of Medical Ethics. Although the Preamble to the AMA's Principles states, "A physician must recognize responsibility to patients first and foremost," it then adds the words "as well as to society."

In the correctional setting, profound conflicts in orientation, mission, purposes, aims, objectives, and even value systems arise from the psychiatrist's duties to the patient clashing with his or her duties to the department of correction and staff and other inmates, which can be regarded as an aspect of the correctional psychiatrist's duties to "society."

The Problem of Dual Loyalty in Correctional Psychiatry. In the correctional setting, there is a conflict between the two distinct roles of the correctional psychiatrist: As the inmate's treating physician and as an employee and agent of a local, state, or private correctional agency. This problem of "dual loyalty" (which has also been called the "the double-agent problem") is pervasive, and, therefore, it is imperative that correctional psychiatrists be aware of these complex conflicts, which may subtly (or not so subtly) influence their medical decisions.

Although a new section of the Principles adopted by the AMA states "a physician shall, while caring for a patient, regard responsibility to the patient as paramount," the reality is that correctional psychiatrists practice in a setting of power and control in which the correctional staff might view psychiatric intervention as an impediment to control and in which, on the contrary, inmates might seek mental health care genuine reasons (e.g., for mental distress and impairment) or alternatively for a variety of reasons. Some of these nonclinical reasons may include a variety of medication-seeking or location-seeking behaviors (e.g., to be transferred off a unit after some type of conflict with staff or a peer, fear of harm/injury, life endangerment claims, to receive certain psychotropic medications for abuse trading as contraband and diversion, to achieve a certain preferred housing environment/location, escape certain work demands/designations, other classification and housing assignments [e.g., low bunk, first row, avoidance of stairs, air conditioning housing] that are preferred to their current housing assignment, to receive disability benefits upon release, or other perceived favorable legal, parole, and probation outcomes).

Medical Autonomy in Correctional Psychiatry. Navigating between institutional and inmate demands requires a focus on medical autonomy, in that nonmedical personnel cannot overrule the professional judgment of correctional psychiatrists regarding their patients' needs. On the contrary, medical autonomy means that only legitimate clinical decisions direct patient care, not the patient's wishes. One simple way of understanding this concept is that inmates/offenders get what they "need" (what is determined to be medically necessary by correctional staff) rather than what they "want" (plastic, cosmetic, or elective surgery procedures, other treatments that are not clinically indicated or are not within accepted community standards of care).

For example, the APA's Annotations emphatically state, "The physician should not delegate… to any nonmedical person any matter requiring the exercise of professional medical judgment."

The NCCHC's Standards, relating as they do to a variety of mental health professionals, refer to this principle as "clinical autonomy" or "mental health autonomy" and state "Clinical decisions and actions regarding mental health care provided to inmates with mental health needs are the sole responsibility of qualified mental health professionals," explaining that "the standard intends to ensure that clinical decisions are made for clinical purposes and without interference from other personnel." An example is given that "if custody staff requests that an inmate be medicated for behavioral or any other reason, the clinical decision to medicate should be based on the health professional's independent determination of clinical need."

The Doctor–Patient Relationship in Correctional Psychiatry. The importance of the doctor–patient relationship is emphasized in the APA's Annotations with the statement, "Physicians generally agree that the doctor–patient relationship is such a vital factor in effective treatment of the patient that preservation of optimal conditions for development of a sound working relationship between a doctor and his or her patient should take precedence over all other considerations."

Do no harm is a basic tenet of medicine, and psychiatrists are accustomed to acting in patients' best interests. Another basic tenet in medicine is that of patient autonomy. By this it is meant that the patient's right to control his or her treatment and, indeed, his or her body is a priority. Both of these values may be at odds with correctional perspectives and policies, and the correctional environment might threaten to coerce professional judgment and actions into meeting security needs rather than the mental health needs of individual inmates.

For example, correctional institutions and health care organizations may for a variety of reasons sometimes place demands on correctional psychiatrists, nursing staff, and other health care staff to withhold, limit,

reduce, delay, or to find a shortcut to appropriate psychiatric treatment services for a variety of reasons (e.g., cost, security risks, lack of sufficient correctional staff to escort an offender for an appointment, lab test, or procedure on-site or off-site, lack of transport equipment, and perceived "deservedness"). Simultaneously, inmates can be insistent and manipulative, sometimes seeking psychiatric care for inappropriate reasons, described previously as med-seeking, location-seeking, and other behaviors. Therefore, the correctional psychiatrist must constantly monitor unintended identification with correctional authorities and countertransference with inmates.

It is paramount for correctional psychiatrists to be aware of relevant financial, fiscal, reimbursement, agency, and role issues that might affect their ability to provide optimal and clinically indicated care to their incarcerated patients or those who require these services. Psychiatrists should be aware that the same professional standards, oversight by state medical boards, and state regulations that pertain to clinical practice in the community "free world" also apply to the services they provide within correctional settings.

Confidentiality in Correctional Psychiatry. Although the usual principles of psychiatric confidentiality apply to the delivery of psychiatric services in jails and prisons, the special needs presented by these settings, especially in relation to security, legal status of the offender (if still a jail detainee with pending legal charges, or if in a state or federal prison system with new charges), require certain modifications based on the responsibilities of the custody staff, as well as the safety and security of the patients, other inmates, and the staff. The APA Guidelines state, "The rights of the patient to… confidentiality must be weighed against the needs of other inmates as well as the institutional needs of safety and security." Table 60.3–2 lists the treatment situations in which the usual rules of confidentiality may not be applicable. Table 60.3–3 lists other situations in a jail or a prison that are part of the mental health treatment process that might require changes from the usual rules of confidentiality.

It is critical for psychiatrists and other mental health staff working in correctional settings to define and maintain their role as a clinician, working in the best interests of the patient, as opposed to a forensic evaluator, or as an agent of the court, law enforcement, or of the state or local jurisdiction. Without a release of information, treating psychiatrists should avoid verbal or written communication with attorneys and other court personnel, and they should avoid inappropriate communication with the media. This is a particular challenge with certain high-profile offenders (e.g., particular heinous offenses, high-profile crimes, politicians, community leaders, and media, sports, and other celebrities), where there may be significant local or national media interest. Responses to media

requests and for interviews regarding individual offenders should be declined and instead directed to appropriate facility and/or health care administrative staff such as a media relations department or individual. Within a correctional setting or system, there is a natural tendency toward curiosity and discussion of high-profile offenders. Psychiatry and other health care leaders should encourage all staff to avoid reviewing the chart/electronic medical record (EMR) of any high-profile offender or seeking out interaction/discussion with that individual or others treating that individual if the health care professional is not directly involved in the evaluation and treatment of this individual and, therefore, is not on a "need-to-know" basis (e.g., not part of that individual's treatment team).

Table 60.3–2.
Treatment Situations in Jails or Prisons under Which the Usual Rules of Confidentiality Might Not Be Applicable

The patient is self-injurious or suicidal
The patient is assaultive or homicidal
The patient presents a clear and present risk of escape or the creation of disorder within the facility

Table 60.3–3.
Situations in a Jail or Prison That Might Require Changes from the Usual Rules of Confidentiality

The patient is receiving psychotropic medication
The patient requires movement to a special unit for observation, evaluation, or treatment of an acute episode
The patient requires transfer to a treatment facility outside the jail or prison

It is also recommended that the correctional psychiatrists specify clearly any limits on the usual precepts of confidentiality before initiating treatment, except in emergencies. However, this notice may be presented to the inmate in general terms, such as applying to situations in which the patient or the institution may be at risk.

Public Areas. In a situation familiar to psychiatrist practicing in various facilities, the NCCHC Standards advise, "Mental health staff should be reminded not to discuss patients' mental health information in front of custody staff or other inmates, including those working in or near the mental health services area." In addition:

> Non-mental health staff who observe or overhear a clinical encounter are instructed that they also are required to maintain confidentiality. Evidence that staff who have access to the mental health records have been instructed in the need for confidentiality may take several forms including policies and procedures, memoranda to staff, meeting minutes, and reviews during roll call or in-services. Where correctional staff are assigned full time to mental health or substance abuse clinics, programs, or residential care units in correctional settings and, therefore, are members of the treatment team with specific duties and responsibilities, they are bound by the

same confidentiality rules as the other clinical staff.

On the contrary, correctional psychiatrist must be able to share relevant confidential information with facility administrators (on a "need-to-know" basis) when it relates to the medical management of a patient, security issues, or a patient's ability to participate in programs. Similarly, it is appropriate that there be communication with nonclinical staff regarding an inmate's significant mental health needs that must be considered in classification decisions to preserve the health and safety of that inmate, other inmates, or staff (Table 60.3–4).

The NCCHC standards point out:

Table 60.3–4.
Inmate Mental Health Needs That Must Be Considered in Classification Decisions

Nonclinical staff are advised of inmates' mental health needs that might affect the following:
Housing
Work and program assignments
Disciplinary measures
Admissions to and transfers from institutions
Mental health and custody staff members communicate about special accommodations needed for the following:
Inmates who are chronically mentally ill
Inmates who are suicidal
Inmates who are developmentally disabled
Inmates who have significant addiction or other serious mental health problems

Since mental health problems affect housing decisions, work or program assignments, and disciplinary management... communication between security, health services, and mental health staff assures that the mentally ill inmate's activities of daily living are restored or maintained to the greatest extent possible. A consistent and reliable method of communication among staff regarding the special mental health needs of inmates is the lynchpin to successful compliance to this standard.

The correctional psychiatrist should be sensitive to the fact that confidentiality is especially important to those for whom the lack of confidentiality might have serious legal and other consequences, such as those detainees with pending legal charges. Some practical suggestions for psychiatrists and other mental health staff might include the avoidance of exploration into the details or circumstances of the alleged criminal act(s) and the individual's state of mind, criminal intent, mitigating factors, or other defense strategies. Similarly at risk are those who have experienced sexual violence, sex offenders, or persons with HIV/acquired immunodeficiency syndrome (AIDS).

The issues surrounding confidentiality in jails and prisons are so complex that it is especially important that clear and specific written policies be developed and maintained. However, in facilities in which there

are no written policies, the correctional psychiatrist should clarify these issues with the institutional authorities and develop working policies as to the degree to which confidentiality of information can be ensured.

Collecting Forensic Psychiatry Information in a Jail or Prison

A paradigm of the problem of dual loyalty is the collection and use of forensic psychiatric information by the correctional psychiatrist. The APA Guidelines set out the

> distinction [that] must be drawn between information obtained by a correctional psychiatrist in the course of treatment, and information obtained from the inmate in the course of an evaluation for non-treatment purposes. Some examples of the latter types of examinations are evaluations, at the request of an attorney or the court for an evaluation of competency to stand trial, mitigating factors, or criminal responsibility (such as in an insanity evaluation), or for a report for probation or the parole board.

It is in these circumstances that the principle of "do no harm" comes to the fore. Thus, the APA Guidelines note, "More complex ethical issues are raised when a treating correctional psychiatrist is asked to determine whether a patient can be sanctioned for an infraction or rules violation." Clearly, the correctional psychiatrist would want to help the inmate if the information could mitigate a sanction but would be reluctant to participate if the information aggravated the sanction: "In these cases, the usual rules in regard to confidentiality typically may not apply and it is the responsibility of the evaluating psychiatrist to set forth clearly to the inmate any such limitations on confidentiality as part of the informed consent process."

Similarly, when the correctional authorities ask correctional psychiatrists to collect information for forensic purposes, the NCCHC standards support the refusal of the correctional psychiatrist to do evaluations of inmates for use in adversarial proceedings or to order blood tests to determine drug or alcohol use (unless it is for clinical reasons). However, in cases of sexual assault, the NCCHC Standards permit health professionals to gather forensic evidence of the alleged victim if requested but not that of the alleged perpetrator.

These matters are of such concern to the NCCHC that it has promulgated an explicit standard prohibiting health care workers from collecting "forensic information." The standard on Forensic Information in the Standards for Mental Health Services in Correctional Facilities states, "Mental health staff are prohibited from participating in the collection of forensic information," which is defined as "physical or psychological data collected from an inmate that may be used against him or her in disciplinary or legal proceedings."

In a clear reference to the problem of dual loyalty, this Standard states:

> The intent of this standard is to ensure that the role of the mental health

staff is to serve the mental health needs of their patients. The perception of a neutral mental health services staff is compromised when they are asked to collect forensic information that may be used against inmates.

It then provides some examples of ethical conflicts for mental health services staff:

- Performing psychological evaluations of inmates for use in adversarial proceedings (e.g., court, probation, or parole hearings)
- Conducting body cavity searches for contraband
- Collecting blood or urine specimens for drug analysis

According to the NCCHC, "Such acts undermine the credibility of these professionals with their patients and compromise clinical staff by asking them to participate in acts that are usually done without inmates' consent." However, "where state laws and regulations require that such acts be performed by health/mental health staff, the services of outside providers or someone on the institution's staff who is not in a therapeutic relationship with the inmate is obtained."

Relevant also are the Ethics Guidelines for the Practice of Forensic Psychiatry of the American Academy of Psychiatry and the Law (AAPL), which deal with this issue explicitly, noting, "When a patient is involved in parole, probation, conditional release, or in other custodial or mandatory settings, psychiatrists should be clear about limitations on confidentiality in the treatment relationship and ensure that these limitations are communicated to the patient."

Informed Consent in Correctional Psychiatry

The issue of informed consent is the subject of great attention throughout medicine and surgery. It is, as the APA Guidelines note, particularly of concern in psychiatry, in which involuntary treatment might be invoked because the competency of an individual might be in question and/or because some emergency exists, but the Guidelines add that the inherently coercive settings of jails and prisons might create additional difficulties for the correctional psychiatrist in regard to consent. It is interesting to note that one legal case stated that it is impossible to obtain true informed consent in this setting.

In general, as the Guidelines note, the principles of informed consent as embodied in the ethical guidelines of the APA's Annotations remain applicable to patients in jails and prisons and, therefore, to the extent possible, the patient should participate in decisions about evaluation and treatment, and correctional psychiatrists should offer to discuss with their patients the nature, purpose, risks, and benefits of the potential types of treatment.

Because involuntary treatment is controversial and complex, the APA Guidelines say only "the actual institutional policies and the decisions of

the correctional psychiatrist in regard to the right to refuse treatment should conform with the rules and procedures of the jurisdiction in which the facility is located."

On the contrary, the importance of this subject is highlighted by the fact that there are two separate NCCHC standards that apply. The first, Informed Consent and Refusal of Mental Health Care, states, "All examinations, treatments, and procedures are governed by informed consent practices applicable in the jurisdiction" and notes that "the intent of this standard is to support the ethical principle that patients have the right to accept or refuse mental health interventions." Moreover:

> Competency may be an issue when mentally ill inmates refuse treatment. Management of such refusals, particularly if they seriously jeopardize the patient's health, requires a collaborative approach of psychiatric, psychological, criminal justice, and legal authorities. If there are no dire consequences to the refusal, the mental health professional may simply note in the record that the person was a "no show" for an appointment or refused medication.

The second NCCHC standard, Emergency Psychotropic Medication, states, "Mental health staff follow policies developed for the emergency use of forced psychotropic medications as governed by the laws applicable in the jurisdiction" (the definition of "forced medications" is "those given without the patient's consent").

It should be noted that the terminology used in the NCCHC Standards is dramatically different from that generally used in psychiatry. As a rule, psychiatric facilities use terms such as "treatment over objection" or "refusal of medication." The Standards focus on the word "forced" and its implications. Thus, the discussion of the latter standard states:

> This standard's intent is to have a clinically approved protocol for emergency situations when an inmate is dangerous to self or others due to a medical or mental illness and when forced psychotropic medication may be used to prevent harm, based on a prescriber's order. The standard supports the principle that psychotropic medication may not be used simply to control behavior or as a disciplinary measure.

It goes on to say, "This standard applies to emergency situations only. For guidance in forcing psychotropic medications on a more frequent basis or as part of an ongoing treatment plan, staff are referred to current case law (e.g., *Washington v. Harper*)."

Privacy in Correctional Psychiatry

The APA Guidelines state, "To the extent possible, treatment should be provided in a setting that is conducive to the achievement of its goals. This includes not only the particular physical setting, but also an atmosphere of empathy and respect for the dignity of the patient." However, the Guidelines note that although "the usual principles of privacy… apply to the delivery of psychiatric services in jails and prisons, correctional psychiatrist must recognize the inevitable and, indeed, necessary loss of privacy in

correctional psychiatric practice."

The NCCHC Standard is more specific and more detailed. Under the heading of "Privacy of Care," it requires that "mental health services are conducted in private and carried out in a manner designed to encourage the patient's subsequent use of services... in private, without being observed or overheard by staff or inmates."

The discussion of this Standard states, "The intent of this standard is to have mental health care encounters conducted in private and to protect the patient's dignity. The patient–clinician relationship may be eroded if privacy is not provided during clinical encounters." On the contrary, "when the patient poses a probable risk to the safety of staff or others, security personnel are present." Similarly, the Standard takes account of one of the unique conditions of correctional psychiatry by pointing out that "in corrections mental health triage commonly occurs at cell-side. As the triage discussion becomes more involved than a simple triage, the clinician should be aware of the need to maintain privacy and move the patient to a clinical setting."

A number of specific situations are dealt with. Thus, the standard addresses the issues of privacy and confidentiality in regard to interpreters, stating:

> The use of interpreters may be frequently employed during the mental health assessment process since a psychiatric evaluation is more likely to be valid when language and culture are appropriately addressed. Interpreters need to be advised on how to protect the patient's privacy. The use of inmates as interpreters should be discouraged and used only in urgent or emergency situations.

When safety is a concern and security presence cannot be minimized, it is recommended that alternative strategies for partial auditory privacy be considered—for example, white noise devices that can mask normal conversation.

Mental health crises are often accompanied by behavior requiring security intervention. In such an emotional environment, there are additional challenges to maintaining privacy and confidentiality. Clinicians new to the correctional environment should be trained to deal with these challenges during their orientation.

It is appropriate to end this section with another quote from the NCCHC Standards: "Maintaining ethical boundaries and professional attitudes is difficult in correctional institutions. However, orientation for new hires and continuing education on potential ethical conflicts can help mental health services staff maintain their ethical perspectives."

PSYCHIATRIC TREATMENT IN JAILS AND PRISONS

Providing competent and quality psychiatric treatment in jails and prisons is an unusually complex undertaking. The psychiatrist is a guest in the house, in the domain of corrections, practicing psychiatry on the turf of

corrections, where custody, control, containment, and maintaining order and security are the primary goal. This task is immediately apparent at the time of entering a facility, walking through a series of electric gates and past razor wire fences, and then going through an inspection for proof of identification, clarification/confirmation that the individual has security clearance and can enter the facility, security review of any/all belongings (medical books, reference materials, food items, and in many facilities personal cell phones and other electronic devices are not allowed), and a pat-down search with metal detectors. The correctional psychiatrist must adjust to the culture of corrections, and, although psychiatric treatment might be legally mandated, the correctional psychiatrist is obviously not in charge of the rules, policies, procedures, processes, and the institution in which care is provided.

The Policy Goal of Psychiatric Treatment in Jails and Prisons

According to the APA Psychiatric Services in Correctional Facilities (Second Edition):

> The fundamental policy goal for correctional mental health care is to provide the same quality of mental health services to each patient in the criminal justice system that should be available in the community. This policy goal is deliberately higher than the "community standard" that is called for in various legal contexts, because resource restrictions on community care can sometimes limit the adequacy of available services. Correctional systems provide additional challenges not present in the community, given that incarcerated persons have no options for care other than what is provided in the correctional facility.

The Purposes of Psychiatric Treatment in Jails and Prisons

Psychiatric disorders and psychiatric comorbidities are overrepresented in correctional populations compared with the general population. Correctional facilities and systems are, therefore, legally obligated to address the medical and mental health care needs of individuals in their placement. The purposes of psychiatric treatment in jails and prisons as set out by the APA Guidelines are listed in Table 60.3–5. These principles are echoed in the NCCHC Standards, which add that the goal is the "best level of functioning" and "In the correctional setting, as in most other environments, the immediate objective of mental health treatment is to alleviate symptoms of serious mental disorders and prevent relapses to sustain patients' ability to function safely in their environment."

Psychiatric Treatment Programs

Table 60.3–6 lists some examples of the APA and NCCHC psychiatric treatment programs for jails and prisons.

The Program Priorities of Psychiatric Treatment in Jails and Prisons. The psychiatric treatment program priorities in jails and prisons as set out by

the APA Guidelines are listed in Table 60.3–7.

Models of Health Care in Jails and Prisons

Different correctional settings have different organizational models for health care delivery, and it is important that the correctional psychiatrist become familiar with the model that his or her institutions follow in order to effectively deal with such issues as the chain of command and the allocation of resources, along with others. In many facilities, medical services and mental health services may be under the same or under separate authorities with different administrators and lines of reporting and supervision. Variations include situations in which the correctional psychiatrist is a full-time or part-time state, federal, or county employee, employed by a university health or a private health care system, or is a locums tenens, part-time or temporary individual contractor providing mental health services to any of the above or alternatively an employee or individual contractor for a national for-profit organization that is responsible for medical and/or mental health care services.

Table 60.3–5.
The American Psychiatric Association Guidelines on Purposes of Psychiatric Treatment in Jails and Prisons

To relieve the unnecessary extremes of human suffering
To make the correctional facility safer for everyone who lives, works, or visits there
To enable patients to avail themselves of the rights of due process
To permit inmates the opportunity to make use of programs offered

Table 60.3–6.
American Psychiatric Association (APA) and National Commission on Correctional Health Care (NCCHC) Psychiatric Treatment Programs

APA Psychiatric Treatment Programs
Depending on the particular correctional facility, psychiatric treatment can occur in any one of a variety of programs, including the following:
An acute care program
A longer-term program
A transitional care program
An outpatient treatment program
An inpatient hospital treatment program
NCCHC Psychiatric Treatment Programs
The National Commission Standards list the following types of programs:
Mental health programs without a residential component
Acute mental health residential units
Nonacute mental health residential units

Jails and prisons differ in the scope of mental health services provided.

Jails process most detainees who have a short length of stay, and, accordingly, jails are primarily concerned with the stabilization and management of acute conditions such as acute psychosis, substance intoxication, withdrawal, and suicide prevention.

While a person is in a jail, the purpose of mental health treatment is to prioritize the detainee's health and safety, but, when appropriate, this might also include the goal and outcome of restored competency to be arraigned or restored competency to stand trial so that their pending legal proceedings may resume. Because this is a forensic and nonclinical role, typically jail psychiatry staff are not involved in the assessment of whether an individual is competent or is not competent to stand trial. Similarly, they are typically not involved in the treatment of restoration of competency to stand trial, but this may vary by location (and whether the jail has a competency restoration program on-site or if this is a separate on or off-site program). Of course, once an individual returns to the jail, ensuring continuity of care and the psychotropic medications (typically initiated in the state hospital or competency restoration program) and other treatments is essential and can be very challenging.

In a prison setting, the outcome might be to return the inmate to the general population with the ability to participate in prison programs, including clinically indicated mental health programs, psychotropic medication administration, monitoring, further psychiatric evaluation and treatment, and preparation for release. Prisons should be considered communities in which individuals spend many years of their lives. As such, prisons need a greater range of services that emphasize long-term supports for particular vulnerable individuals, including sheltered and/or residential units (for individuals with stable conditions who cannot be placed into the general population), case management, individual and/or group therapies, and other in cell programming materials (self-study materials and the like). Programs should be available to prison inmates who repeatedly self-injure, mutilate, those with the need for frequent/repeated psychiatric hospitalizations who might require step-down programming, followed by sheltered or residential unit placement with other similar offender patients, and more intensive treatment to insure medication compliance, monitoring of their activities of daily living (ADLs), and medication and illness teaching and education, than that is typically available within general population settings. Sheltered housing and tailored assessment and treatment programs for individuals with developmental disabilities are also necessary.

Table 60.3–7.
American Psychiatric Association Guidelines on Priorities for Psychiatric Treatment Programs in Jails and Prisons

Recognizing and providing access to treatment for each inmate with serious mental illness or in psychiatric crisis

Consulting with other health care providers and correctional administrators and staff

As resources permit, treating patients who are in need of mental health care but who are not seriously and persistently mentally ill or in acute psychiatric crisis

The Therapeutic Milieu for Mental Health Treatment

According to the APA Psychiatric Services in Correctional Settings (Second Edition): "Adequate and appropriate access to mental health treatment means that patients have no unreasonable barriers to receiving services. Examples of unreasonable barriers include disincentives, including fees that deter a patient from seeking care for legitimate mental health needs; interference with the prompt transmittal of a patient's oral or written request for care; unreasonable delays before patients are seen by mental health staff or outside consultants; insufficient custody staff to transport inmates to clinic; inadequate waiting room space; and punishment for seeking or refusing care."

Also, according to the APA Psychiatric Services in Correctional Settings (Second Edition):

> The importance of private space for confidential doctor–patient interactions cannot be overstated. Inmates with mental illness may be vulnerable or targeted by other inmates. Without the assurance of privacy, some patients may not seek mental health services or adhere to treatment. However, efforts to speak privately should not be pursued at the risk of reducing patient access to care. For example, a patient may refuse to come out of the cell to attend the clinic for an appointment but may wish to speak with the psychiatrist at cell-side. Psychiatrists need to ensure that such refusals to come out do not arise from external disincentives, such as cell shakedowns occurring during psychiatric appointments. Patients should be encouraged to speak in private settings, but clinicians must use their judgment regarding safety considerations, unit rules, access to treatment, patient wishes, and the therapeutic alliance.

Table 60.3–8 lists the APA Guidelines conditions that would contribute to a therapeutic milieu for treatment.

Access to Mental Health Care and Treatment

Constitutional provision and statutory law require that an inmate in a correctional setting have access to mental health treatment, and this principle is emphasized in the APA Guidelines on Psychiatric Services in Jails and Prisons, which state, "Timely and effective access to mental health treatment is the hallmark of adequate mental health care." In addition, it is striking to note that the very first standard of the NCCHC's new Standards for Mental Health Services in Jails and Prison is on access and requires that "inmates have access to care to meet their serious mental health needs," defining access to care as meaning that "in a timely manner, a patient can be seen by a clinician, receive a professional clinical judgment, and receive care that is ordered" and adds that this principle is "the foundation upon which National Commission on Correctional Health Care

standards are based."

Table 60.3–8.
American Psychiatric Association Guidelines on Conditions That Contribute to a Therapeutic Milieu for Treatment

A therapeutic milieu implies the following conditions:
A sanitary and humane environment
Written procedures and adequate staffing to permit adequate observation
Adequate allocation of resources for the prevention of suicide and assault
Medical and mental health staff available to provide adequate treatment and supervision
Social interactions that foster recovery
The ability to transfer an inmate to an appropriate mental health facility when the foregoing conditions cannot be met

Furthermore, this section notes that access to care procedures should be flexible enough to accommodate inmates with mental illness and functional deficits, and that such deficits can manifest as the following:

- Inability to recognize an emerging health or mental health need
- Inability to follow usual procedures in requesting health or mental health care
- Severe anxiety or suspicion preventing follow-up care
- Behavior that makes care especially challenging

Barriers to Inmates Receiving Services. To determine whether there is adequate and appropriate access to psychiatric treatment, it is critical that the treatment delivery system be analyzed to ensure that there are no unreasonable barriers to patients' receiving services. Some examples of unreasonable barriers listed in the APA Guidelines are listed in Table 60.3–9.

Modalities of Treatment

The APA Guidelines state, "Whatever the setting or the program… treatment in a jail or a prison should employ a variety of therapies, including biological, psychological, and social," and, most important, they add that psychiatric treatment "must be more than merely prescribing of psychotropic medication, and correctional psychiatrists should resist being limited to this role."

Again, to emphasize, treatment modalities must be provided in a manner consistent with generally accepted psychiatric practices, with a focus on evidence-based practices. At the same time, treatment must also be in accord with institutional requirements. A written, individualized treatment plan must guide psychiatric care.

Table 60.3–9.

American Psychiatric Association Guidelines on Examples of Unreasonable Barriers to Patients' Receiving Psychiatric Services

Instituting or allowing disincentives that would deter a patient from seeking care for his or her legitimate mental health needs
Interfering with the prompt transmittal of a patient's oral or written request for care
Allowing unreasonable delays before patients are seen by mental health staff or outside consultants
Charging fees or costs that prevent or deter patients from seeking care
Punishing an inmate for seeking or refusing care

Similarly, the NCCHC Standards define mental health services to "include the use of a variety of psychosocial, psychoeducational, and pharmacological therapies, either individual or group, including biological, psychological, and social."

Evidence-Based Treatment. Correctional psychiatrists must strive to provide evidence-based treatment as the standard of care in jails and prisons. Evidence-based treatments can be defined as treatments that, in appropriate studies, (1) have been rigorously tested, (2) have yielded consistent, replicable results, and (3) have proven safe, beneficial, and effective for most people diagnosed with the particular mental illness.

Because correctional psychiatry is a critical part of public psychiatry, it is appropriate to quote from the Report of the Criminal Justice/Mental Health Consensus Project, which describes the development of the concept as it promotes the utilization of evidence-based practices in public psychiatry:

> In recent years, enormous advances have been made in treatments available for persons with mental illness. New medications have emerged; new services, supports, and interventions have proven effective. Researchers have conducted studies and collected data—they have developed an "evidence-base"—which demonstrate the effectiveness and applicability of some of these treatments and approaches.
>
> Gradually, a body of research literature is growing to support the choice of particular interventions in certain situations. While some researchers might argue over the standards by which an intervention or treatment approach is judged to be evidence-based, there is general agreement that the term and designation imply that a given practice has withstood rigorous scientific examination.
>
> The public mental health system must take steps to ensure that practice keeps pace with research.

Psychopharmacology in Jails and Prisons. As a paradigmatic managed care context, costs in correctional facilities are closely monitored and medications are particularly vulnerable to substantial cost control. The correctional facility or system formulary is the means by which this is done. Consistent with the principle of evidence-based treatment, the correctional psychiatrist must seek to ensure that the facility formulary permits the inmate with mental illness to be treated with the most suitable and clinically appropriate/medically necessary medication. Facilities need to avoid formulary restrictions that lead to negative clinical outcomes for

patients. Formulary decisions should be based upon safety, evidence base to support, comparison and efficacy data to other comparable agents in the population under consideration, convenience of administration, abuse and diversion and other safety issues, and not solely on cost issues alone. There should be an administrative review process to seek nonformulary approval of certain agents.

- Medication given without documentation of a suitable examination including an adequate history.
- Medication prescribed without the necessary laboratory workup or monitoring.
- Medication given when records document contraindications or a previous adverse effect or it is known to be ineffective.

However, in the practice of correctional psychiatry, there is the additional risk if

- Medication is prescribed, which appears to be given for punitive rather than therapeutic reasons.

Prescribing psychotropic medications within correctional settings is a particular challenge. Both medical and psychiatric professionals who provide health care services in correctional settings may be naïve to the potential for abuse, "cheeking" and other types of medication diversion, trading, and the like with commonly prescribed medications. Many of these medications are not typically targets of such misuse in the free world. In correctional settings due to a ready supply of traditional substances of abuse, creative means of seeking substances or abuse are employed. These methods include creating fermented beverages (e.g., "pruno") and diverting medications with CNS effects that are maximized by nasal insufflation, injection, and other means of administration. Examples of medications that are highly abused within correctional settings include Quetiapine and bupropion (Wellbutrin).

There is no current peer-reviewed resource document or national practice guideline on the evidence-based prescription of psychiatric medications in correctional facilities. Practicing psychiatry and prescribing psychotropic agents within jails and prisons is substantially different than it is in the community. Some of these differences include strict formulary controls, high rates of comorbidities such as substance use disorders and personality disorders, high rates of malingering, the abuse and diversion of prescription medications, and various operational challenges unique to these environments. Therefore, each decision to start, continue, hold, discontinue, or change the mechanism of administration of certain psychotropic medications (e.g., crush and float) for offender patients requires careful consideration by the psychiatrist and has a myriad of implications for the individual patient offender and staff.

It is generally recommended that psychiatrists should use psychotropic medications in a safe and clinically appropriate manner and after review of the potential risks, benefits, side effects, and alternatives. As with any mental disorder, barring maintenance treatment, it is unwarranted to prescribe psychotropic medications in the absence of distinct target symptoms or when placement and mental health follow-up services are unclear.

To avoid risk management problems, it is generally recommended that medication be used according to U.S. Food and Drug Administration (FDA) —approved indications and doses, and if this is not preferred, then it is suggested that the correctional psychiatrist obtain and document a second opinion.

Among the various special clinical issues that are especially of concern in correctional settings is the very high prevalence or comorbidity of substance abuse disorders among offenders with mental illness. Correctional psychiatrists will inevitably be prescribing psychotropic medications for inmates with substance abuse problems and should, therefore, have education and training in substance use disorders and the treatment of co-occurring (mental illness and substance abuse) disorders.

Another important clinical issue in correctional facilities is that sleep complaints are very common. There is, obviously, a difficult confluence of the real stresses of the environment on one hand and manipulation by malingering inmates on the other. The correctional psychiatrist must ask whether the sleeplessness is episodic or has been a lifelong problem. Has the patient previously sought medical attention? Has he or she previously abused street drugs as a way to get to sleep? Is the sleep disorder a symptom of depression? Is the sleep difficulty the result of an accurate perception of risk in the cell area or the result of a paranoid perception of risk in the cell area? Is the patient having nightmares, or is the patient going into mania? It is obvious that initially consideration must be given to a nonpharmacological approach to sleep disorders. Some examples of useful recommendations include the implementation of behavioral interventions and strategies such as regular exercise, avoidance of caffeine products, avoiding daytime naps, and other sleep hygiene approaches.

Precautions against abuse are necessary, and so, for example, it is recommended that directly observed therapy (DOT) or crush and float or other strategies be implemented for medications with particular overdose or bartering risks. The administration of medication is usually delegated to staff—both clinical and correctional. The training required by this situation is set out in two NCCHC standards. One, titled, Medication Services, states that: “This standard intends that mental health prescriptive practices are commensurate with current community practice. The responsible physician and the responsible mental health clinician develop facility-specific guidelines regarding prescriptive practices for psychotropic medications....

Medications are prescribed only when clinically indicated (e.g., psychotropic medications are not used for disciplinary purposes, nor are sleep-inducing medications used for behavior modification purposes)."

Of interest, considering the legal issue of substance abuse treatment noted earlier, the NCCHC Standards state, "Inmates entering the facility on verifiable prescription medication, including methadone, continue to receive the medication in a timely fashion as prescribed, or acceptable alternate medications are provided as clinically indicated."

The second NCCHC Standard, "Medication Administration Training," sets out the requirements in regard to this topic and states, "Training topics should include psychotropic medication security and accountability, documentation of medication administration, side effects, phototoxicity, drug diversion (e.g., hoarding, selling), changes to the formulary, risks of medications, and therapeutic regimens adherence."

Ensuring Access

Access to mental health treatment in a jail or a prison is ensured by appropriate screening, referring, evaluating, providing intervention, and providing consultation within the facility.

Screening

In jails and prisons, screening is critical, and, not surprisingly, both the APA Guidelines and the NCCHC Standards emphasize the importance of appropriate screening. The APA Guidelines describe in detail the sequence of screenings and assessments that lead up to treatment (Table 60.3–10).

For example, the APA Guidelines describe the progression in regard to an individual with a substance abuse disorder who might present during receiving mental health screening with any of a variety of medical, neurological, or psychiatric problems. The receiving mental health screening instrument should include questions relating to the individual's use of alcohol and other drugs, the patterns of use, and observations related to whether the offender is under the influence of alcohol or drugs or experiencing withdrawal symptoms. The Guidelines note that the correctional psychiatrist should have an important role in the development of the screening instruments, in the training of staff in its use, and in educating correctional security, medical, and mental health staff about the signs and symptoms of intoxication and withdrawal.

Table 60.3–10.
American Psychiatric Association Guidelines on Sequence of Screening and Assessments Leading up to Treatment

Screening
- Receiving mental health screening

Intake mental health screening
Referral for mental health evaluation, followed, as necessary by
Brief mental health assessment
Comprehensive mental health evaluation
Mental health intervention, if necessary

A positive receiving mental health screening for substance intoxication should trigger an immediate mental health screen for the presence of depressed mood and/or suicidal ideas. Similarly, the intake mental health screening should include questions on an inmate's history of alcohol and other drug use, as well as observations for the presence of signs of intoxication or withdrawal symptoms. Again, the role of the correctional psychiatrist is in the development of the instrument and staff training. The brief mental health assessment and the comprehensive mental health evaluation are particularly important in ensuring a proper diagnosis as the basis for the appropriate clinical management of an inmate with a substance use problem.

The NCCHC notes that "receiving screening" is the means by which "newly admitted inmates' emergent or urgent mental health needs are identified and met through an immediate referral to mental health staff."

Both the APA and the NCCHC require a subsequent "intake" screening that must take place within 14 days.

The importance of screening is further demonstrated by the development of specific instruments to ensure successful screening.

The step-by-step sequence ensures that an inmate who is identified through reception screening, intake screening, or by various processes at any time thereafter (including self-referral) as possibly in need of mental health treatment is provided with a suitable mental health evaluation.

Referral. Emphasizing that within a correctional facility the referral process is critical, and noting that this can be simple or complex, depending on the facility, the urgency of the situation, and the mental health coverage provided, the APA Guidelines state that the referral process, as set out in the facility's operating standards, should be time constrained and specifically define the roles of the various participants, that the maximum time for response to each referral should be stated, and that timely referral and response should be designated as indicators of quality care.

Evaluation. The nature and quality of the particular mental health evaluation are related to its context. For example, in a lockup, in which the major focus might be the removal of an inmate with severe mental illness to a psychiatric facility, evaluations tend to be brief and focused. In a jail or a prison, in which various levels of mental health housing and treatment might be available, such evaluations and interventions could be more

detailed and extensive.

Consultation. An important role of the correctional psychiatrist is to take the initiative in providing consultation to staff with regard to certain situations. Some of these are listed in the NCCHC Standards, such as consulting with regard to inmates who have chronic illness (i.e., asthma, heart disease, diabetes, and hypertension) or communicable disease (e.g., HIV, hepatitis, syphilis, gonorrhea, or chlamydia); counseling patients coping with the psychological effects of their disease; and consulting with regard to inmates who are physically or developmentally disabled and either assisting staff in managing such inmates or providing behavioral consultation so that the inmate can deal with the emotional and attitudinal consequences of his or her disability during incarceration. Mental health staff can help develop appropriate social skills or work with other facility staff to arrange for special housing to ensure personal safety.

Consultations might also be called for with regard to inmates who are chemically dependent; pregnant or have recently given birth, for example, to minimize the postpartum phase and separation anxiety from the newborn; and of course, elderly or terminally ill inmates.

Rape and Victims of Sexual Assault. A full description of the psychiatric assessment and prevention of sexual abuse and exploitation within correctional setting is beyond the scope of this chapter. The following provides a summary overview.

The Federal Bureau of Prisons estimates that 9 to 20 percent of male inmates are raped in prisons. Rape is also a reality in women's prisons. In 2003, the Prison Rape Elimination Act (PREA) became federal law; however, this phenomenon continues to be a challenge to manage with regard to prevention and aftercare.

Under the PREA, Congress created the National Prison Rape Elimination Commission. Its duties include writing standards aimed at reducing the incidence of prison rape and reporting on the incidence of rape and ranking facilities on that measure.

The Bureau of Justice Statistics report for the year 2007 indicated the following:

An estimated 60,500 inmates (or 4.5 percent of all state and federal inmates) experienced one or more incidents of sexual victimization involving other inmates or staff.

Nationwide, about 2.1 percent of inmates reported an incident involving another inmate and 2.9 percent reported an incident involving staff.

Among the 146 prison facilities in the 2007 National Inmate Survey (NIS), which was conducted in 146 state and federal prisons between April and August 2007 with a sample of 23,398 inmates, six had no reports of sexual victimization from the sampled inmates, and 10 had an overall

victimization rate of at least 9.3 percent.

Among the 10 facilities with the highest overall prevalence rates, three had prevalence rates of staff sexual misconduct that exceeded 10 percent.

The NCCHC has a standard entitled Federal Sexual Assault Reporting Regulations, which requires that each facility have "written policy and procedures regarding the detection, prevention, reduction, and punishment of rape consistent with Federal law" and notes that "the intent of the standard is for correctional facilities to comply with applicable federal law and that the Prison Rape Elimination Act of 2003 addresses the many aspects of rape in correctional institutions, including the actions to be taken by correctional administrators, and is the foundation of this standard."

The national PREA standards were released in 2012 (http://www.prearesourcecenter.org/library/488/standards/department-of-justice-national-prea-standards). There is now a need for compliance with this federal mandate for facilities in the fields of adult, juvenile, and community corrections. Both correctional staff and health care staff might need education, training, and psychological support in managing either victims or perpetrators of sexual assault. Both the NCCHC and the American Correctional Association (ACA) are playing a vital role in supporting correctional institutions in the implementation of and the compliance with the health care aspects of the PREA with education, curriculum developments, and technical assistance. Specialty training on complicated PREA issues is available for psychiatrist and medical and mental health care practitioners. The curriculum includes information on detecting and assessing signs of sexual abuse and harassment, the medical forensic examination and forensic evidence preservation, and effective and professional responses and reporting. In August 2013, the U.S. Department of Justice began auditing facilities to ensure compliance with the Act. For technical assistance, health care practitioners should contact the PRC at: www.prearesourcecenter.org.

Primary Health Care Professionals with Regard to Inmates with Medical Problems. These consultations with medical staff should be documented in the clinical record. Any facility staff member who works with inmates (e.g., inmate work supervisors, custody staff, case managers, teachers, or chaplains) might benefit from instruction on how to deal with inmate behavioral problems.

Sick Call and Nonemergency Mental Health Care Requests and Services . Although the correctional psychiatrist works with inmates by appointments in an office situation if possible, in most correctional facilities, access to general health care is by means of "sick call."

The importance of this access is emphasized by the NCCHC Standards,

which, as noted earlier, state that the means by which "inmates' routine mental health care needs are met" is that "all inmates have the opportunity daily to request mental health care. Their requests are documented and reviewed for immediacy of need and the intervention required. Mental health appointments and providers' clinics are conducted on a timely basis and in a clinical setting by qualified mental health professionals." The NCCHC Standards are quite specific and require that "all inmates have the opportunity daily to request mental health care" and that "verbal or written requests for mental health care are received daily by mental health staff and triaged within 24 hours."

There is a separate NCCHC standard for emergency services, and emergency mental health care is defined as "care for an acute mental illness that cannot be deferred until the next scheduled mental health clinic or routine appointment," which must be available 24 hours a day, 7 days a week. Details of responding to a mental health crisis are set out in a discussion.

Treatment Planning

The correctional psychiatrist has been trained as early as medical school to understand the importance of a detailed, in-depth, professional treatment plan. In any context or location, this requirement is a crucial reflection of the necessary standard of medical and psychiatric care. This is reflected in the APA Guidelines, which, with regard to discharge planning, note that "it is recommended that discharge planning be a part of the initial treatment plan." So too, the NCCHC Standards require that "mental health services are provided according to individual treatment plans" and discuss this subject in considerable detail, requiring that "patients receive care that is tailored to their individual needs… [including] consideration of cultural and language differences."

Disaster Planning. It is imperative that every correctional facility have a detailed disaster plan specifically for inmates with mental illness and/or medical diseases and treatments (e.g., dialysis dependent, supplemental oxygen) and/or limited mobility. The lessons learned from the horrific problems for the inmates in the New Orleans Jail arising out of the Hurricane Katrina disaster were many, especially that the extensive planning required for such a complex operation must include in-depth coordination with community resources and regularly scheduled and monitored practice drills. It must be clear, however, that the security service, usually the Department of Correction, has the ultimate responsibility for the "care and custody" of the inmates in a jail or a prison. The responsibilities of the psychiatric service or larger mental health services are ancillary to this.

The NCCHC Standards require that correctional facilities have an

emergency response plan that includes provisions for a mass disaster drill —a simulated correctional emergency [that] frequently involves a natural disaster (e.g., tornado, flood, earthquake), an internal disaster (e.g., riot, hostage incident, arson, kitchen explosion), or external disaster (e.g., mass arrests, bomb threat, power outage). Simulated correctional emergencies frequently involve multiple casualties that require triage by health and mental health staff. It is recommended that the emergency plan be "practiced and critiqued at least annually so that mental health staff is familiar with emergency preparations."

Discharge Planning. Studies of recidivism among persons in the justice system with mental illness indicate that effective discharge planning is essential to the continuity of care and a critical part of adequate psychiatric treatment in jails and prisons.

Despite court interventions developed for detainees with mental illness, there is still a lack of coordination and continuity of care for such persons on release. After incarceration, former inmates with mental illness have residential, financial, health, and mental health needs. Face-to-face linkage with community services has the greatest likelihood of success. Because most jails cannot control the court's release decisions, they usually do not know when an inmate is to be released, and so most jails do not provide case management or discharge planning services. To provide such services, discharge planning must start at admission. Providing uninterrupted medications after release is an important challenge. Without psychiatric medications, recidivism is almost inevitable. As the APA Guidelines state, "Timely and effective discharge planning are essential to continuity of care and an integral part of adequate mental health treatment."

This is true whether the patient is being released into the community or transferred to another correctional facility. Because discharges (e.g., from jails) or transfers (e.g., between prisons) can occur on short notice, it is recommended that discharge planning be a part of the initial treatment plan. Essential services that should be included in adequate discharge planning are psychiatric assessments, appointments to be arranged, notifications, staff responsibilities, linkages to community-based services, help with the obtaining of necessary financial benefits (e.g., Medicaid) and housing (due to the high incidence of homelessness in the jail and prison populations), placements, and appropriate linkages.

The NCCHC Standards suggest a number of components for a discharge-planning program:

- Formal linkages between the facility and community-based mental health organizations
- A list of community-based mental health providers that is used for referral

- An inmate education session that emphasizes the importance of follow-up aftercare
- Arrangement for follow-up mental health services before an inmate's release
- Preparation of mental health summaries
- Assistance to inmates with their application to entitlement programs
- If appropriate, a supply of medications provided at the time of the inmate's release

Reentry Programs. Whereas discharge planning is an aspect of services in jails and prisons, the last decade has seen the proliferation of reentry programs, which can provide court-based diversion, discharge planning, and transition services to people with mental illnesses being released from city jails or state or federal prisons. These programs (usually government funded—generally state or county) evaluate, select, and then actively link (and follow up) the inmate with mental illness reentering the community to a wide variety of services, including intensive case management, mobile services, outpatient treatment, psychiatric rehabilitation, housing, vocational and educational services, access to entitlements, family supports, and substance abuse treatment.

A Recent Landmark Case with Regard to Discharge Planning and Reentry; the New York City v. Brad H. Case

Fortunately, jails can emulate several model programs that provide aftercare discharge planning. For example, in August 1999, five prison inmates who suffered from psychiatric disabilities filed a class action lawsuit in New York State Supreme Court on behalf of themselves and all other similarly situated persons, challenging, among other things, New York City's practice of discharging inmates with psychiatric disabilities from the Rikers Island jail in the middle of the night, with only a $3 Metrocard and $1.50 in cash and without any psychiatric medications or referrals to services. The settlement in 2003 in this, the Brad H. case, requires that New York City provide discharge-planning services for inmates with psychiatric disabilities subsequent to their release from jail. The settlement requires the City to meet the following obligations:

- Release inmates during the day, when service agencies are open
- Make medication and transportation to community residences or shelters available on discharge
- Help eligible inmates to obtain Medicaid, Social Security, disability, and public assistance benefits before release
- Enable family and community members with clinical information about an inmate to relay such information to discharge-planning or jail mental health staff

- Establish a monitoring system overseen by the court that will set performance goals for City agencies and assess their compliance for at least 5 years

Of course, these requirements should also apply to prisons. However, prisons are usually distant from the inmate's home community. This reality makes it difficult for state prison and local mental health providers to have service agreements. When parole of an inmate includes the recommendation for psychiatric treatment, the need for prerelease planning is increased.

This important area of reform of the correctional system gained added impetus from the Second Chance Act of 2007, federal legislation designed to ensure the safe and successful return of prisoners to the community by providing the following: Grants to states and local governments to promote the safe and successful reintegration into the community of individuals who have been incarcerated, including employment services, substance abuse treatment, housing, family programming, mentoring, victims services, and methods to improve release and revocation decisions using risk-assessment tools; grants to nonprofit organizations that can be used for mentoring adult offenders or offering transitional services for reintegration into the community; grants to improve the availability of drug treatment to offenders in prisons, jails, and juvenile facilities; and grants to states, local governments, and Indian tribes to develop and implement family-based treatment programs for incarcerated parents who have minor children. The Second Chance Act provides guidance to the Bureau of Prisons for enhanced reentry planning procedures. Specific information on health, employment, personal finance, release requirements, and community resources is to be provided to each inmate released.

It also authorizes the U.S. Justice Department's National Institute of Justice and the Bureau of Justice Statistics to conduct reentry-related research and establishes a national resource center to collect and disseminate best practices and to provide training on and support for reentry efforts.

The Facility Information Technology System: Needs Assessment and Allocation of Resources

With regard to a correctional facility's information technology system, the APA Guidelines point to a capacity for "needs assessment" and state:

> It is critical that the administration of the correctional facility have the capacity to gather and assess information relevant to the prevalence of mental illness in the facility because this data forms the basis of the needs assessment that assures an adequate level of care.

The allocation of resources within a jail or a prison (which should be set out in written documents) is obviously determined by many factors, including the extent of resources available, the size of the facility, the

overall organization of the services provided, the mission of the facility, and the lengths of stay of the inmates confined there.

The Facility Quality Assurance Plan. One important aspect of the reform of the correctional mental health system calls for each jail or prison to have a quality assurance plan that systematically reviews the quality of mental health services and the efficiency and effectiveness of the utilization of staff and resources in the delivery of services.

Other quality assurance activities in the jail or the prison include the following:

- Credentialing
- Reviewing service access and utilization
- Managing records and resources
- Reviewing morbidity and mortality
- Identifying and preventing risk
- Monitoring and reviewing of high-risk critical procedures such as overriding treatment refusals and responding in emergency situations

Accreditation and Standards. The first national health care standards drafted specifically for correctional institutions came from the American Public Health Association (APHA). Unfortunately, there was no consideration regarding how to enforce the standards.

In 1977, the AMA published its first jail correctional health standards; 16 jails were accredited for compliance that year. In 1979, the prison and juvenile correctional health care standards were published. Then, as mentioned earlier, in 1983, the NCCHC was established and developed correctional health care standards and a correctional health care accreditation processes.

The NCCHC has become the major organization establishing correctional health care standards and providing correctional health care accreditation. In 2007, the NCCHC published separate standards for the accrediting of mental health programs in jails and prisons, Standards for Mental Health Services in Correctional Facilities. The most current NCCHC standards include the Jail (2014), Prison (2014), Mental Health (2015), and the Juvenile Standards (in press, the last version was 2011).

The NCCHC also has developed standards for opioid treatment programs within correctional settings. These Opioid Treatment Programs (OTP) standards represent the requirements for corrections-based opioid treatment programs seeking accreditation from the National Commission. The NCCHC used federal regulations and community standards as a guide and modified them to take into account the issues unique to providing services in a correctional facility. The NCCHC OTP Standards (similar to other NCCHC standards) are divided into nine general areas: health care

services and support, inmate care and treatment, special needs and services, governance and administration, personnel and training, safety, health records, health promotion, and medical legal issues. With regard to the OTP standards, all of the standards are linked to specific federal regulations.

Many types of "noncorrectional" health care organizations—such as hospitals, nursing homes, and psychiatric facilities—are accredited by The Joint Commission (formerly the Joint Commission on Accreditation of Healthcare Organizations). Similarly, jails, prisons, juvenile detention, and other correctional facilities may be accredited by the NCCHC (an offshoot of the AMA as described previously), the ACA, The Joint Commission, or a combination of the above. Although national accreditation is typically voluntary, it is often a contractual requirement for universities, other health care systems, and private vendors who provide health care services to correctional systems. In addition, when facilities undergo investigation or litigation, or are placed in receivership or federal oversight, they are often mandated to establish and maintain national accreditations.

Staffing. The effectiveness of any correctional psychiatric system is largely dependent on staffing considerations. Owing to the complexity of how different correctional facilities are situated—for example, some are in rural settings, whereas others are in the heart of downtown—no national staffing model fits all facilities. Staffing must be sufficient to ensure that every inmate with serious mental illness or in psychiatric crisis has timely access to evaluation by a competent mental health professional. More important than the number of staff is access to adequate care. Accordingly, none of the national correctional health care standards mandates staffing patterns other than that there be an adequate number and type to meet the needs of the population.

Several variables must be considered when trying to decide the best staffing pattern for specific institutions. Institutional size and inmate length of stay are important considerations in determining the best mental health staffing patterns. For example, considering their volume of admissions and shorter length of stay, jails might require more intake screening staff than a similarly sized prison with fewer scheduled intakes and longer lengths of stay. The number of inmates at each custody level within a facility also determines staffing needs. For example, a maximum security facility needs more health staff to check on a confined inmate's mental health than an equally sized minimum security facility in which inmates can report to sick call if distressed.

A satellite work camp facility that houses younger, healthier inmates might need less psychiatric health care staff than a similarly sized facility that houses inmates with chronic psychiatric disorders in a day care program. Considering that epidemiological studies reveal that female

inmates have higher incidence and prevalence rates of mental illness than male inmates, and given the high rates of exposure to traumatic stress that female inmates experience, it can be seen that women require more access to psychiatric services.

Whether correctional facilities obtain psychiatric services off-site or on-site affects the number of correctional psychiatrists needed. Furthermore, some prisons have special functions within their state systems that alter their staffing patterns. For example, some prisons act as mental illness repositories for their department of corrections and accordingly need more psychiatric staff for inmate psychiatric care.

Requirements that are external to the facility, such as state licensing regulations or court orders, can also shape health staffing patterns. The number of psychiatrists needed to staff correctional facilities is affected by some states allowing liberal use of "physician extenders"—for example, nurse practitioners or physicians' assistants—to do mental health functions. However, it is important to note that because of the unique importance of psychotropic medications and the unique role of correctional psychiatrists in providing such medications, every system must have enough psychiatrists to provide these services.

According to the APA guidelines, Psychiatric Services in Correctional Facilities: "There are no prevailing national standards for psychiatric staffing in correctional or other settings. Nevertheless, courts and regulatory agencies, court monitors, and health consultants frequently seek to establish staffing ratios in correctional settings in order to comply with constitutionally mandated minimally acceptable medical care. Governmental agencies with oversight of correctional settings often seek such ratios to assess their budgetary needs. Correctional facilities have historically not had enough psychiatrists to adequately provide the services described here. In practice, a facility-specific staffing needs analysis is a fundamental step in determining the actual psychiatric staffing required to meet the standards of care in any given facility. Correctional administrators and other leaders need guidance about adequate psychiatric staffing. Needs and type vary on the basis of patient acuity and treatment need, patient volume and turnover, facility location (whether services will be provided on site, via telepsychiatry, or a combination of both modalities), the role and function of the facility, the targeted inmate population, any unique or specialized unit or facility treatment/rehabilitative mission, and changes to any of the above. Although it is very difficult to establish exact psychiatrist-to-patient ratios, the amount of psychiatric time must be sufficient to ensure that there is no unreasonable delay in patients receiving necessary care, and all relevant and necessary psychiatric functions must be met."

The Use of Inmate Workers in Treating Patients. An issue perhaps unique to correctional facilities is the use of inmate workers. The

psychological aspects of using inmate workers have costs and benefits. For the inmate worker, there may be specific benefits, including perhaps learning a valuable occupation/work skill, remuneration and other token benefits (e.g., earning good time, extra privileges, decreased level of security/supervision), out of cell time, social interaction with staff and other inmates, sensory stimulation, and the benefits of self-esteem. As is noted in the NCCHC Standard on Inmate Workers, "understaffed correctional institutions may be tempted to use inmates in the delivery of mental health care by performing services when personnel are not available. However, this should be avoided since it gives inmates unwarranted power over their peers and creates additional risk for the institution."

On the contrary:

> There are circumstances, however, in which inmates may be appropriately used to help mental health staff with routine tasks. Inmate participation in appropriate peer mental health or substance abuse programs is permitted provided that such individuals are appropriately trained and supervised. Inmates are always supplemental in these peer mental health-related programs and are not substitutes for regular programming or mental health staff. Inmates may assist other inmates in activities of daily living (ADL) (i.e., assisting in ambulation, bathing, dressing, feeding, and toileting) in regular housing or mental health residential units may participate in support groups that assist inmates with mental health problems, e.g., a buddy system for potentially suicidal inmates. In a correctional setting, usual community practices, such as using peer counselors in group therapies, may need to be modified to prevent inmates from being placed in a position of power over their peers.

Peer-to-Peer Programs. Peer-to-peer programs have proven their worth in prison ministries, chaplain volunteer, and other spiritual ministries. Other peer activities may include volunteer work, providing care and training to service dogs for Veterans, and other individuals with disabilities, for example, the Patriot Paws programs. There are other self-help peer-to-peer organizations as Alcoholics Anonymous (AA) and Narcotics Anonymous (NA). This concept is being extended into areas of reentry and continuing postincarceration support.

The APA Guidelines state, "As safety and security allow, self-help and peer support programs or activities contribute to the overall goals of the psychiatric services and where and when available should be promoted and encouraged."

However, within the correctional community, this is not an uncontroversial area. Concerns about the power of inmate workers were cited earlier. There are valid security concerns about prior inmates working within a correctional facility.

Issues in Administration and Management

Integration of Services. Within clinical services, mental health, psychiatric, and psychological services should be integrated with medical and substance abuse services. These, in turn, should be integrated into the

overall management of the facility.

NCCHC Standards address administrative issues in detail and require administrative meetings and reports: "Mental health services are discussed at administrative meetings. In addition, mental health care staff meetings are held to review administrative issues." Moreover, "This standard is intended to facilitate mental health care delivery through coordinated monitoring, planning, and problem resolution by mental health, health, correctional, and administrative personnel. It is also intended that mental health staff be informed of relevant correctional, administrative, and health care issues."

Furthermore, the NCCHC Standards encourage mental health staff to attend medical and correctional staff meetings to promote positive working relationships within the institution and to facilitate and coordinate their respective areas of responsibility.

Written Policies and Procedures

Mental health administrative personnel in a jail or a prison should establish written policies on such critical issues as (1) staffing patterns, (2) admission, (3) referral, (4) discharge criteria, (5) health care information management, and (6) interagency and intra-agency communications, especially with regard to confidential medical and mental health information.

Leadership. The APA Guidelines note, "Psychiatrists can provide administrative leadership and clinical care. Even when control may be in the hands of nonpsychiatrists, attention to quality care serves the best interests of the patient." The APA guidelines also urge that "attention always must be focused on the priorities of quality care and the best interests of the patient." It is especially important to have clinical input when decisions are made about an inmate who is receiving mental health care regarding such issues as (1) disciplinary issues, (2) work and housing assignments, and (3) transfers in and out of the institution.

Perhaps idealistically, the APA Guidelines comment that "an environment which promotes therapeutic interactions may be created in a jail or a prison setting if there is clinical leadership with authority, staffing, and resources to work towards such an environment."

Scope of Practice. To promote "a good working relationship between all disciplines, including nurses, psychologists, psychiatrists, social workers, correctional officers, correctional counselors," the APA Guidelines state:

> Each mental health professional working in a correctional setting should work within his or her own profession's scope of practice, as defined in that jurisdiction's licensure process (usually by the state), and within his or her own training, expertise, and skills. As such, mental health

professionals provide services within a standard of care. However, it must be noted that chronic shortages of professional resources may be a problem in correctional settings and although this may engender a "siege mentality," the correctional psychiatrist should not compromise his or her professional responsibility. At the very least, a psychiatrist's responsibility includes the need to communicate shortcomings and resource needs to the appropriate authority.

Professional Development. It would be ideal if correctional psychiatrists had received specialty education and training at various levels before undertaking employment in a correctional setting. However, education and training in correctional psychiatry are relatively rare in medical schools and psychiatric residencies, although forensic psychiatry fellowships now require it.

Jails and prisons should seek expertise in correctional psychiatry through various means, such as recruitment of staff with subspecialty training, the offer of continuing professional education, and consultation. For the practicing correctional psychiatrist, professional development is essential and should be available in the educational and training program of the facility or in other continuing medical education activities such as relevant courses or, at a minimum, in textbooks and journals.

The educational and training program of the facility should include substantial cross-training of security staff by clinical staff and vice versa. Psychiatrists and other mental health professionals will benefit from training and orientation by security staff to the jail or prison "culture," including such matters as (1) social order, (2) security threat groups and gang affiliations, (3) inmate hierarchy issues and attitudes toward sexual offenders, and (4) utilization and/or misuse of the mental health practitioner and system.

Another aspect of professional development for the correctional psychiatrist is communication with colleagues who are working in other correctional facilities to share experiences and provide mutual support. In addition, correctional psychiatrists should become familiar with and consider joining subspecialty organizations, such as the AAPL, and the Society of Correctional Physicians (SCP), which provide for continuing education and subgroups that share this interest. There are several other organizations of interest, including the Academy of Correctional Health Professionals and the American Correctional Health Services Association (ACHSA). There are several national correctional conferences including the NCCHC, the SCP, and the ACA.

The APA Guidelines advise that correctional psychiatrists be informed about and sensitive to security needs and issues and the concerns of security staff, and they recommend that an organized program be developed to teach and train correctional personnel about mental health issues, noting that among the programs that will especially benefit from carefully administered integration are those relating to (1) developmental disabilities, (2) neurological impairments, (3) alcohol and drug abuse, and

(4) sex offenses.

The NCCHC Standards are very detailed in this regard and have a specific standard, titled Mental Health Training for Correctional Officers, that requires that "a training program [guide] the mental health-related training of all correctional officers who work with inmates," specifying that:

1. All correctional officers who work with inmates receive mental health-related training that:
   a. is established or approved by the responsible mental health authority in cooperation with the facility administrator; and
   b. includes at a minimum:
      i. recognizing signs and symptoms of mental illness;
      ii. communicating with inmates who have positive signs of mental illness;
      iii. procedures for suicide prevention and intervention; and
      iv. procedures for appropriate referral of inmates with mental health complaints or suicidal behaviors to mental health staff.
2. Correctional officers assigned to mental health areas, including mental health programs, residential units, or segregated housing areas receive additional training from mental health staff in order to fulfill their specific roles.

It states, in addition, "In a correctional setting, custodial professionals and mental health (and other medical) professionals benefit from cross-professional training by participating in the others' in-service training."

Finally, whenever possible, the correctional psychiatrist (and, indeed, the facility) should develop an active liaison and/or affiliation with an academic–medical or other educational institution(s). This relationship can be quite advantageous in recruitment and retention, in staff continuing education, as well as career satisfaction, the opportunity for case consultation, and the maintenance of high-quality services.

Preventing Burnout

Clinical and administrative work within correctional settings is varied and challenging. Correctional systems face challenges in funding for health care and in recruitment and retention of a qualified health care workforce. This work can also be frustrating, and "burnout" is an inherent risk. Psychiatrists are encouraged to participate in professional activities, pursue continuing medical education opportunities, and communicate with other colleagues who work in correctional settings to share experiences, identify examples of strategies and approaches to challenging issues, and to provide mutual collegial support. Work in a correctional setting can be quite intense, stressful, and frustrating, and, therefore, the correctional psychiatrist must be alert to the risk of burnout, as evidenced by a long-

term sense of exhaustion and a loss of interest.

There are a number of means by which this can be avoided. Most important is the recognition of the signs and symptoms of burnout. As an example, the clinical workload, supervision duties (medical students and psychiatric resident physicians, physician assistants, and nurse practitioners), explicit clarification of duties, responsibilities both on-site and via phone after-hours and on weekends, and on-call expectations should be clarified at the start and periodically. Similarly, specific times for teaching and supervision and other prioritization of work tasks, the correctional psychiatrist might be able to minimize burnout and at the same time fulfill a basic responsibility. (Ideally, the employment agreement with the facility or system should provide specifically for such activities.) The correctional psychiatrist might also try to diversify his or her work in the facility or vary the practice settings. Thus, many correctional psychiatrists might choose to work only part-time in a correctional setting while maintaining practices in other settings (e.g., private practice, hospital based, community mental health).

The continuing education programs in correctional psychiatry will give a fresh perspective and perhaps an opportunity to network, as should support of the informal type—relationships and consultation with colleagues can be augmented by membership or participation in professional organizations such as the National Commission's Academy of Correctional Health Professionals, the APA's Caucus of Psychiatrists Practicing in Criminal Justice Settings, the AAPL, the ACHSA, or the SCP.

Cultural Sensitivity

The APA Work Group to Revise the APA Guidelines on Psychiatric Services in Correctional Facilities identified that correctional facilities differ from most community settings in gender and racial, ethnic, cultural, and sociodemographic factors. They have high percentages of male inmates, minority ethnic groups, and homeless or impoverished individuals.

This contributes to cultural differences among differences between the security staff, the psychiatrist and other medical and mental health staff, and the inmates. The subject of cultural sensitivity (or, as it is sometimes referred to, cultural awareness or cultural competence) is becoming increasingly the focus of attention in all areas of psychiatry. As in any treatment setting, incarcerated individuals deserve equal care regardless of gender, gender preference, racial/ethnic background, socioeconomic status, education, or other cultural factors.

For present purposes, note that the APA Guidelines strongly state, "In order to properly either evaluate or treat inmates the correctional psychiatrist must overcome the consequences of these often substantial cultural differences in background, socioeconomic class, education, and other types of diversities."

The APA Guidelines recommend that:

Correctional facilities should provide training to ensure sensitivity to cultural differences and support efforts to overcome impediments to the delivery of mental health services. Such training can foster positive attitudes and acceptance of other cultures by means of didactic and experiential components. The goals and objectives of such training should relate to attitude, knowledge, and skills. Positive attitudes and acceptance of other cultures increase with exposure to and awareness of other belief systems. Desired outcomes include improved acceptance of diverse populations, empathy for the minority experience (including the internalization of experiences of prejudice), and an understanding of concepts of ethnocentric bias and its potential effects.

Standards and Practice Guidelines

Practice Guidelines. The APA publishes Practice Guidelines that provide evidence-based recommendations for the assessment and treatment of a wide range of psychiatric disorders of considerable importance to correctional psychiatrists, including acute stress disorder and posttraumatic stress disorder (PTSD), bipolar disorder, borderline personality disorder, psychiatric disorders related to HIV/AIDS, major depressive disorder, obsessive-compulsive disorder (OCD), panic disorder, schizophrenia, substance use disorders, suicidal behaviors, and others.

PSYCHIATRIC PATIENTS IN CORRECTIONAL FACILITIES

Demographics

The importance of demographic studies of psychiatric disorders in correctional facilities cannot be overemphasized. For advocacy purposes and for public health considerations, the results of this research dramatically portray the critical need for psychiatric services in jails and prisons.

However, the true incidence and prevalence of psychiatric disorders in correctional facilities are quite difficult to obtain. Well-designed, carefully done epidemiological studies are costly and complex to carry out and because of the high inmate turnover rates, doing a point prevalence study of mental illness in jails is quite problematic. The best way to detect incidence and prevalence of psychiatric problems in jails is to find cases on admission. Prison inmates stay for longer terms, and conducting a point-prevalence epidemiological study of mental illness in this setting is more feasible. Despite these difficulties and limitations, the epidemiology of mental illness in correctional facilities is gradually becoming known. In general, most authorities now estimate that about 16 percent of the correctional population has serious and persistent mental illnesses. Thus, in 2006, an estimated 350,000 seriously mentally ill offenders were incarcerated in the nation's jails and prisons. (This is contrasted to the fewer than 55,000 patients in state psychiatric hospitals.)

Jails. With regard to jails, the few available epidemiological jail studies highlight that, after controlling for demographic differences between jail samples and the epidemiological catchment area (ECA) samples, jail prevalence rates of mental illness are two to three times higher than those in the general population (Fig. 60.3–1).

Prisons. Because prison inmates stay for longer terms, there are more prison population–based epidemiological studies of mental illness than there are in jails. The ECA study found that 6.7 percent of inmates in prisons had experienced schizophrenia at some point, compared to 1.4 percent of the US household population. A number of studies reveal that, at any given time, a little more than 2 percent of the inmate prison population is in inpatient or residential psychiatric care.

More than 30 percent of male inmates and 78 percent of female inmates with mental illness reported prior physical or sexual abuse, and inmates with mental illness in correctional facilities are more likely than other offenders to have a higher prevalence of homelessness, unemployment, alcohol and drug abuse, and physical and sexual abuse before their current incarceration.

Special Populations in Correctional Facilities

WOMEN IN CORRECTIONAL FACILITIES. As noted in the APA Psychiatric Services in Correctional Facilities, 3rd Edition: Female inmates frequently encounter different stressors before, during, and after incarceration than their male counterparts. Most women in correctional settings are mothers and have at least one child for whom they have been the primary caregiver prior to incarceration. Women who are pregnant during their incarceration are especially vulnerable. More than 75 percent of women entering correctional settings report a history of emotional, physical, or sexual abuse as a child or adult. They have higher rates of domestic violence victimization than men. Incarcerated women have greater risk than men for sexually transmitted diseases and HIV because of prostitution, higher rates of substance use disorders, and drug offense arrests and are more likely to be victims of sexual assault while incarcerated. Demographic studies have shown that compared to men, a greater percentage of women entering a jail or prison have a serious and persistent mental illness. Although the rates of bipolar disorder and schizophrenia are similar for incarcerated men and women, a much larger percentage of women are diagnosed with major depression and severe anxiety disorders, especially PTSD. In addition, a higher percentage of women than men in jails and prisons enter these settings with a history of substance abuse.

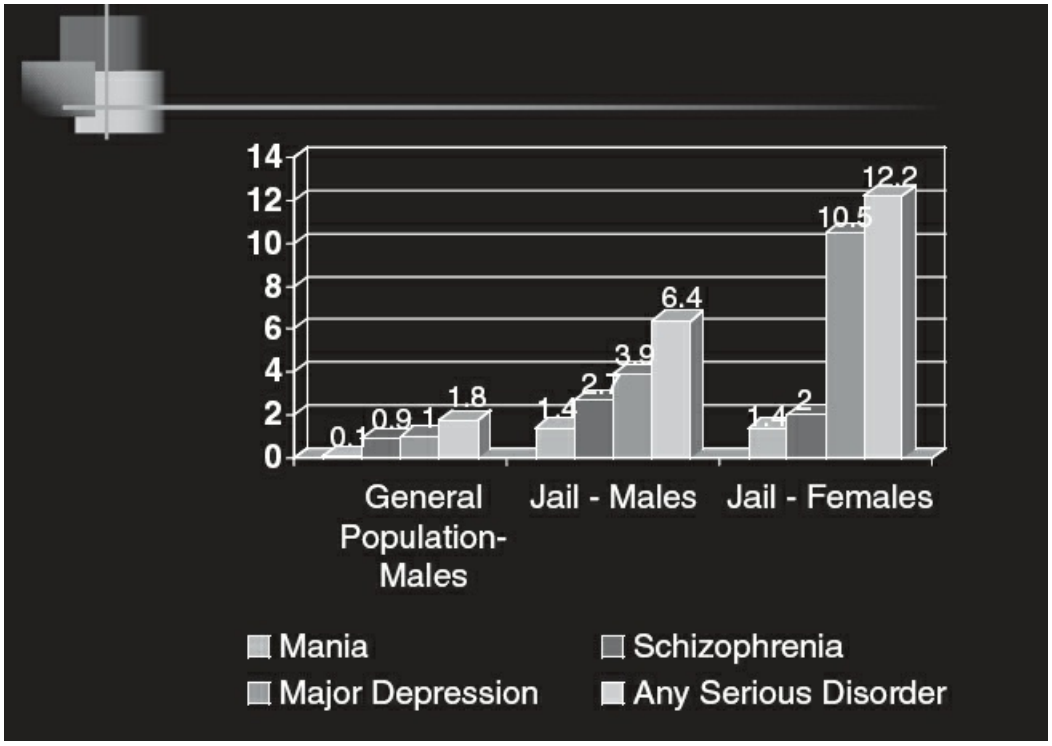

FIGURE 60.3–1. Serious mental disorders among the general population and among jail admissions.

Although the number of women in jails and prisons is relatively small

compared to the number of men, their percentage of the total correctional population is rising. Thus, at year-end 2006, women made up 7.2 percent of the population under state or federal jurisdiction, up from 6.7 percent in 2000. Until recently, most psychiatric treatment programs for women have been based on treatment models designed for men. It is now recognized that female inmates require very different screening, evaluation, and treatment, and with the reform of the correctional system, treatment programming provided to women in jails and prisons is changing dramatically.

ABUSE AND PTSD. One aspect of this change is the recognition that that female inmates have had high rates of victimization (sexual and physical) causing PTSD. The trauma of being raped before 16 years of age was the most frequently reported antecedent to PTSD among female prisoners.

In national surveys of prison inmates, the prevalence of HIV infection among incarcerated women ranged from 2.5 to 20.0 percent, which greatly exceeds the rates of HIV infection among women in the general population. The rates of HIV infection among female prisoners are often three times those of male prisoners.

A large percentage of women (up to 70 percent) entering correctional settings report a history of abuse either as a child or as an adult. Therefore, it is extremely important that in the screening process female inmates be asked questions related to abuse histories. (It should be noted that a substantial number of male inmates have also suffered childhood sexual abuse, and, therefore, many of these principles might also apply to them.)

PREGNANCY, CHILDREN, AND FAMILIES. Programming for women in correctional facilities must consider that most women in correctional settings have dependent children and at least one child for whom they have been the primary caregiver. Their concerns about their children require special counseling on parenting and perhaps child custody.

Clearly, it is especially important to attend to the health and mental health care of women who are pregnant or who have given birth recently. Furthermore, the correctional psychiatrist should be alert for and should minimize the risks associated with peripartum depression and psychosis, other adjustment disorders, and distress, and ensure that the facility establish a policy that a comprehensive mental health evaluation be conducted or has any mental health problems during the pregnancy, delivery, or postpartum period.

TRAINING WITH REGARD TO GENDER-SPECIFIC ISSUES. The unique issues presented by the presence of female inmates in a jail or prison call for staff training (both clinical and corrections) in gender-specific issues such as differences in symptom presentation, especially among women with abuse histories. In particular, the use of seclusion and restraints with women raises special

concerns because these interventions might mirror the pattern of historical abuse. As a result, these methods, which are used for reasons of safety, might inadvertently retraumatize inmates. In response to this, a number of hospitals and community mental health systems have developed special protocols for seclusion and restraint of women that involve the treatment team and the patient working together, early in the treatment process, to develop interventions that can help an agitated female patient to calm down. This has general applicability, in that much of the behavior that might seem to require restraints might be de-escalated by "talking the person down" and understanding what is really going on. The behavior might be the result of dissociative phenomena or flashbacks of abuse rather than psychotic symptoms or poorly controlled impulsivity.

Similarly, it is important that staff (and inmates) be taught (and trained) as to what constitutes sexual harassment and abuse of inmates. In addition, because of the higher prevalence of mental illness among women and their generally higher degree of willingness to participate in services, it is generally agreed that the per capita mental health staffing for female inmates should be significantly higher than that for men.

Inmates with Substance Use Disorders in Correctional Facilities. The correctional psychiatrist must have knowledge and familiarity in the assessment and treatment of incarcerated individuals with substance use disorders because it is estimated that up to 70 to 90 percent of offenders have a substance use disorder at the time of their entry into the criminal justice system, and criminal behavior often occurs when an individual is under the influence of alcohol or other drugs. Furthermore, and very important to note, some serious suicide attempts and completed suicides in jails may occur within the first 24 to 48 hours of admission and are often carried out by inmates who are either intoxicated or experiencing withdrawal symptoms.

Table 60.3–11.
American Psychiatric Association Position Statement on Treatment of Substance Use Disorders in the Criminal Justice System (2007)

It is the position of the American Psychiatric Association that:

All arrestees should be evaluated for the presence of substance use disorders and of withdrawal syndromes. Jails should be prepared either to provide quality treatment for alcohol and/or drug withdrawal and associated comorbidities or to arrange timely transfer to an appropriate medical facility if detoxification services are not available on site. As with any other illness, when a person who is arrested is receiving prescribed medication for an addictive disorder, that medication should be continued until a comprehensive mental health evaluation has been performed and a new treatment plan provided.

Prisons should make available quality treatment for substance use disorders to all inmates who qualify for such treatment. This should include psychiatric evaluation, treatment with appropriate medication and nonpharmacological interventions, prerelease planning, and adequate follow-up treatment on a residential or outpatient basis after release.

For any parolee who has received prerelease substance abuse treatment, the continued cooperation in follow-up treatment should be a condition for continuation of parole.

Courts established for diversion purposes, such as mental health courts, drug courts, and driving while intoxicated courts, should ensure that treatment recommended for participants in these programs follows the principles outlined in this position statement.

In 2007, the APA Assembly and the APA Board of Trustees approved a position statement prepared by the APA Council on Addiction Psychiatry on the treatment of substance use disorders in the criminal justice system, as shown in Table 60.3–11.

DETOXIFICATION. When psychotropic medications are required during detoxification to manage medical or prevent impairment and sequelae, agitation and symptoms, or agitation, a physician must be responsible for managing the detoxification process.

Prescription of benzodiazepines might be useful to prevent or attenuate withdrawal symptoms and detoxify the inmate. However, their use should be short term and on a tapering dose schedule to ensure that the prescribed medication does not serve as a substitute for the substance of abuse.

When circumstances warrant, an inmate should be detoxified or free from substance use for a reasonable period of time before psychotropic medications are prescribed unless psychotic or suicidal symptoms are present. The opiate- or other substance-dependent pregnant female presents special challenges to correctional systems, and review of NCCHC standards and other accepted practice guidelines is recommended.

COMMUNITY SUPPORT. Wherever feasible, and only with custody approval and clearance, outside community, volunteers, spiritual leaders and chaplains, and self-help meetings should be allowed in the correctional facility to permit integration and identification with the community on release.

Of course, from the public health perspective, rehabilitation with regard to substance abuse must be considered as a long-term commitment and must extend past the period of incarceration to have a positive effect on criminal recidivism and alcohol and other drug abstinence. For these reasons, discharge planning must include referral to community-based substance abuse treatment.

Inmates with Co-occurring Disorders in Correctional Facilities. In 2006, the Bureau of Justice Statistics stated that about 74 percent of inmates in state prisoners and 76 percent of jail inmates who had a mental health problem also met criteria for substance use disorder. Nearly 63 percent of state prisoners who had a mental health problem had used drugs in the month before their arrest, compared to 49 percent of those without a mental health problem.

For the correctional psychiatrist, the nondetection of one or more co-occurring disorders in jails and prisons risks the exacerbation of any behavior problems and elevates the risks of suicide and poor outcomes in treatment.

For these reasons, the APA Guidelines recommend that the detection of one type of disorder (i.e., substance use or mental illness) immediately triggers an investigation into the other type of disorder. The importance of a coordinated system of screening, referral, assessment, and diagnosis for co-occurring mental health and substance use disorders is that one type of disorder can mimic the other, interfere with detection of the other, or overshadow the other type, and one type might affect the presentation of the other type of disorder. Table 60.3–12 lists the recommended APA Guidelines strategies for inmates with co-occurring disorders.

THE INTEGRATION OF MENTAL HEALTH AND SUBSTANCE ABUSE SERVICES. The APA Guidelines emphasize that substance abuse services should be integrated with mental health services and that psychosocial problems and skill deficiencies should be addressed with individualized programming created through comprehensive assessment and consultation with the inmate's clinician, family (or significant others), and correctional medical and mental health clinicians.

In contrast to the APA position recommending the integration of services, the NCCHC in its standard on Substance Abuse Services relegates the roles of the medical and mental health staff to referral:

> Increasingly, in larger correctional systems, inmates with [substance use]... disorders are receiving treatment for these problems from specialized treatment providers. Substance abuse counselors who treat [substance use]... problems should be properly qualified to do so by a recognized accrediting body. Medical and mental health care providers can facilitate the treatment process by referring inmates for treatment, providing counseling to motivate inmates to receive treatment, legitimizing self-help groups, tailoring controlled substances prescribing practices, and otherwise being a partner in the process of treatment.

Table 60.3–12.
American Psychiatric Association Guidelines on Recommended Strategies for Inmates with Co-occurring Disorders

Services must be focused on the integration of treatment programming by addressing the person's mental health and substance use disorders simultaneously.
Each type of disorder must be treated as primary, with a focus on understanding how they interact with each other.
Medication must be prescribed with caution. Where circumstances warrant, an inmate not previously prescribed psychotropic medication generally should be detoxified for a reasonable period of time before psychotropic medications are prescribed unless psychotic or suicidal symptoms are present.
Psychosocial problems and skill deficiencies should be addressed with individualized programming.
Treatment should be integrated with self-help groups and support networks, which may be invaluable tools in assisting treatment participants in maintaining their commitment to daily alcohol and drug abstinence.

Interventions must be designed for the particular setting, each setting requiring differing intensity, length, and types of services.

The presence of co-occurring disorders requires special attention to discharge planning, which must address housing and job needs, family reconnection, and continued treatment.

There is agreement, however, that inmates with co-occurring mental illness must not be barred from participation in substance abuse treatment programs and that self-help, 12-step, and peer support groups and networks play an important role in the rehabilitation of abusers of alcohol and other drugs in the correctional system.

Inmates with HIV/AIDS in Correctional Facilities. Inmates with HIV/AIDS present complicated clinical challenges for the correctional psychiatrist because the most common psychiatric manifestations of HIV/AIDS include (1) early signs of dementia, (2) paranoia, and (3) other psychotic and affective symptoms. Impressive strides have been made in the treatment of HIV/AIDS resulting in decreased morbidity and mortality. HIV/AIDS is now viewed more as a chronic disease rather than a terminal illness.

As the APA Guidelines note, the risks of social stigma and issues of personal safety complicate confidentiality for the infected inmate, other inmates, and correctional and clinical staff. The usual cultural differences among inmates and security and health care staff may be exacerbated in cases of inmates with these infections, especially HIV/AIDS. "Issues of confidentiality are complicated by the risks of social stigma and issues of personal safety." Therefore, education of both staff and inmates is very important with regard to HIV/AIDS.

One important aspect of the HIV/AIDS situation is that housing requirements in some jurisdictions might create barriers to access to mental health treatment, and housing isolation might exacerbate the inmates' feelings of abandonment and lead to a worsening of symptoms of mental illness.

The APA Guidelines further suggest that the correctional facility undertake an outreach program and/or periodic rescreening to foster early detection and utilization of mental health services by inmates with HIV/AIDS. The APA guidelines also recommend the following: Education of staff and inmates about infectious diseases and the means of transmission is especially important for both personal and population health risk reduction. Inmates need access to health educational materials and resources on prevention of transmission and consequences of continued high-risk behaviors for HIV and viral hepatitis such as promiscuity, unprotected intercourse, and intravenous drug abuse, and disease transmission.

Inmates with Hepatitis C Infection/Coinfection in Correctional Settings. As

noted in the APA Psychiatric Services in Correctional Facilities, 3rd edition, in recent years, hepatitis C infection or coinfection has surpassed concerns regarding HIV and AIDS. Current challenges of hepatitis C virus (HCV) in corrections include the large number of infected offenders, the increased use of medications, and more sophisticated and expensive treatment standards. The prevalence of HCV infection is higher among correctional populations than among the general US population. Although most HCV-positive offenders are asymptomatic, those who develop a chronic infection risk serious complications and death. HCV-infected inmates who receive interferon treatment may develop depressive symptoms, mood fluctuation, and suicidal ideation. Chronic HCV infection is the leading cause of end-stage liver disease in correctional settings. More cases of liver failure are expected as the proportion of older offenders continues to increase. HCV-infected offenders who progress to end-stage liver disease require frequent and costly medical hospitalizations and emergency department services for the treatment of bleeding, abdominal fluid retention, and other serious complications. Ultimately, the only viable option for some of these offenders will be liver transplantation, a procedure with ethical questions, and fiscal/budgetary impacts due to the substantial initial and subsequent costs.

From a clinical perspective, increased ammonia level is commonly associated with behavioral disturbance and presents as a delirium with waxing and waning symptoms that requires timely medical intervention. These behavioral changes may be attributed as being “psych (psychiatric)” in etiology and, therefore, staff education and training is paramount. Chronic HCV infection is a risk factor for liver cancer and a contributing factor in many liver cancer deaths. Many correctional health care system providers are facing unprecedented challenges in addressing the growing HCV epidemic. These challenges revolve around financial and logistical impediments to evaluating and treating such a large number of HCV-infected offenders along with a constantly evolving consensus about how to best manage the disease in the correctional setting. Meeting these challenges has enormous public health implications because most HCV-infected offenders eventually return to their home communities with the potential to infect others. Accepted standards of care for the management of HCV infection recommended by the National Institutes of Health and other key groups are rapidly evolving.

The use of interferon required mental health screening and evaluation when clinically indicated for the development of depressive symptoms and suicidal ideation. This increased the mental health and psychiatric service need demands for correctional facilities. Some systems have limited antiviral treatment options, and some limit treatment to those patients at highest risk for developing cirrhosis and subsequent liver failure. Offenders with end-stage liver disease and associated conditions are particularly

difficult to treat. They may not need medical hospitalization but require specialized housing (often with supplemental oxygen, vital sign and cardiac monitoring, and on-site nursing and medical staff), specialists who can manage end-stage liver disease and associated complications, access to laboratory studies, and coordinated treatment by nursing, medical, psychiatry, and mental health staff.

Inmates with Intellectual and Developmental Disabilities in Correctional Facilities. Offenders with co-occurring mental illness and intellectual or developmental disability encounter unique difficulties in the criminal justice system and have special treatment needs. From 4 to 10 percent of inmates have intellectual developmental disorder (formerly referred to as mental retardation or borderline intellectual functioning), that is, significantly subaverage intellectual functioning and impairments in adaptive functioning and "it is estimated that 40 to 70 percent of individuals with mental retardation in jails and prisons have diagnosable psychiatric disorders."

This is a substantial number of inmates in the correctional population who have intellectual disabilities that present special problems for the correctional psychiatrist. Some of the difficulties to which the correctional psychiatrist must be alert are listed in the APA Guidelines (Table 60.3–13).

According to the APA Psychiatric Services in Correctional Facilities, inmates with intellectual or developmental disability and co-occurring mental disorders are unfortunately the most likely to be preyed on and ridiculed by other inmates. Their inability to process information rapidly or to comprehend instructions, their low frustration tolerance, and their impulsivity may have severe disciplinary consequences. Custody and treatment staff need additional training and education on intellectual and developmental disabilities to lessen the likelihood that they misperceive behaviors as intentional rule infractions or attribute them solely to intellectual disability while an SMI goes untreated.

Table 60.3–13.
American Psychiatric Association Guidelines on Difficulties in the Correctional Population

Deficits in communication skills contribute to the problem for correctional psychiatrists, who depend primarily on interview and spoken language for making assessments of the problems and recommending the appropriate intervention.
Self-injurious behavior may be viewed as a manifestation of a depressive disorder, a manifestation of a diminished ability to tolerate stress, a symptom of anxiety, a problem of impulsivity, or a learned behavior leading to being the center of increased attention from others.
Inmates with this combination of difficulties are the most likely to be preyed upon and ridiculed by other inmates.
Their inability to process information rapidly or to comprehend instructions and their low frustration tolerance and impulsivity may have severe disciplinary consequences.

Each inmate's needs must be assessed individually in regard to housing, support, protection, and observation by mental health and security staff. It is important, however, to ensure that inmates with mental retardation/developmental disability and mental illness are not precluded from receiving or participating in services or programs available to all other inmates.
Therefore, security and treatment staff alike should have additional training and education on mental retardation/developmental disability in order to minimize the likelihood that behaviors are misperceived as intentional rule infractions or attributed solely to mental retardation while failing to treat a serious mental illness.

Geriatric Inmates in Correctional Facilities. The Federal Bureau of Justice Statistics notes that US prison population is aging, and that "expressed in terms of percent change, the oldest age group, inmates age 55 or older, recorded the largest change—an increase of 85 percent since 1995." Effective access to mental health services for older inmates requires recognition of the special challenges facing inmates and mental health professionals, especially as the number of offenders older than 65 years in state and federal prisons increased at 94 times the rate of the overall sentenced population between 2007 and 2010. The aging of the correctional population will be a major problem in the 21st century.

As noted by the APA Guidelines, for the correctional psychiatrist, older inmates have medical needs that sometimes might present with psychiatric symptoms and might complicate psychiatric intervention in terms of medication, relevance of counseling, structure of time (programming) of the inmate, and housing. These problems are more likely to be chronic, permanent, and progressive, might be related to the possibility of dying in custody, and might be expensive (Table 60.3–14).

For the older inmate, the isolation can be expressed in a variety of ways, and, when extreme, can exacerbate or create mental illness or psychiatric crisis. Special issues are related to (1) acceptance after release from incarceration, (2) provision of food, shelter, and clothing after release, (3) assault and potential greater sequelae of injury, and (4) isolation from older, "free" relatives and friends who might be unable to travel to visit.

Table 60.3–14.
The APA Guidelines on Psychosocial Concerns of Older Inmates

Estrangement from or lack of connection to other inmates in the general population, given the relatively small (though increasing) percentage of older inmates
Physical vulnerability to more serious consequences of assault
More difficulty and greater length of time to adjust to a new environment
Higher rate of completed suicide
Greater possibility of dying during incarceration
Higher incidence of loss of external supports, e.g., spouse, parents, friends, and other external supports

Some of these issues might benefit from mental health intervention, including group or peer counseling. Psychiatric and clinical staff need training in the special needs and care required by older inmates, and they

can help educate custody staff to recognize and deal with cognitively impaired inmates who might not respond appropriately to a direct order, mobility issues, and those who might demonstrate early confusion and cognitive impairments and associated sight and hearing impairments.

The Dying Inmate. Facilities should have written policies and training to address the in-custody dying inmate–patient. When feasible, the care of an inmate with terminal illnesses should involve use of a hospice approach —a facility that emphasizes providing support and humane care to a dying patient rather than trying to cure the underlying disease. This might be inside the prison (less likely to be needed in jail) or a contract facility outside of the prison. Another consideration is the use of compassionate release (also called "medical parole"), which is available in a majority of states. Compassionate release can have a number of advantages, including cost savings for the prison health care system, which can be substantial.

The use of inmate volunteers (peer trainers) who are oriented in basic health issues, universal precautions, and mobility management can reduce the dying inmate–patient's (or his or her family's) perception that "the prison isn't doing enough." Group therapy, realistic planning, and suicide prevention are specific interventions that the psychiatrist should either provide directly or ensure provision by other mental health professionals.

CRITICAL CLINICAL ISSUES IN CORRECTIONAL PSYCHIATRY

Suicide Prevention

Suicide is a sentinel event in prison, and preventive efforts reflect the adequacy and comprehensiveness of mental health, psychiatric, custodial, and administrative services in a correctional system. The risk of suicide is dramatically higher among incarcerated jail and prison inmates than among the population at large. Suicide rates for men in jails are approximately 9 to 15 times higher than for the general population, and prison suicide rates are approximately 1.5 times higher than in the general population. An increased risk of suicide associated with mental illness also has been found. Researchers examined psychiatric disorders and completed suicide in the nation's largest state prison system and their findings highlight the importance of maintaining suicide prevention programs in correctional settings, with particular emphasis on screening and monitoring of patients with severe psychiatric disorders.

Therefore, suicide prevention—identifying and responding to suicidal inmates—ultimately, a suicide prevention program, is the collective responsibility of administrative, custodial, and clinical staff.

Legal cases that were decided during the period of reform of the correctional system held that it is a legal requirement that every jail and prison must have a suicide prevention program that identifies and

responds to each suicidal inmate and takes an active approach to his or her management.

The APA Guidelines recommend, and to meet the suicide prevention standard, an essential standard from the Jail, Prisons, Juvenile and Mental Health Standards of the National Commission that mental health staff and the correctional psychiatrist must ensure that the suicide prevention in a particular jail or prison includes the following components:

An adequate suicide prevention program must include the following components, which should be available in a written policy and procedure manual that all staff can easily access:

- Training for all staff who interact directly with inmates to recognize warning signs and intervene appropriately with individuals at risk for suicide
- A formal and detailed suicide risk assessment process
- Identification of inmates at increased risk of suicide through screening at or near the time of admission to the facility and through referral at any time during an individual's incarceration
- An effective and well-understood referral system that allows staff and inmates to bring a suicidal inmate to the prompt attention of a mental health clinician
- Timely evaluation by a mental health clinician to determine the level of risk posed by a referred inmate
- Timely implementation of monitoring interventions such as close observation (at least every 15 minutes), continuous monitoring, alternative housing, or referral to a higher level of care (e.g., infirmary or hospital). When indicated, a psychiatrist either on call or in person assists in the decision about interventions. Any mental health clinician may order an increase in monitoring level, but a decrease in level may be ordered only by a psychiatrist or doctoral-level mental health clinician after an in-person evaluation.
- Housing options that allow for adequate monitoring of suicidal inmates by staff. Suicidal inmates should not be placed in isolation settings without continuous monitoring. Continuous monitoring for active suicidality should occur regardless of housing location. Supervision aids such as video monitoring may be used as a supplement to, but not as a replacement for, active staff monitoring.
- Communication among mental health, correctional, medical, and other staff of the specific needs and risks presented by a suicidal inmate
- Timely provision of mental health services, including medication, individual and group therapies, and crisis intervention, for chronically or acutely suicidal inmates
- Accurate, behaviorally specific, and highlighted documentation in the medical record of behaviors or statements that indicate suicide risk

- Quality improvement reviews with psychiatric staff participation of suicide attempts and completed suicides to help prevent future occurrences
- Critical incident support, opportunities for peer-to-peer discussion, and availability of Employee Assistance Program support for completed suicides to assist staff and inmates to deal with feelings of guilt, fear, grief, and anger

High-Risk Periods. It is obvious that an inmate might become suicidal at any point during incarceration. The high-risk periods that both the APA Guidelines and the NCCHC Standards emphasize are listed in Table 60.3–15.

An inmate who is placed in and is unable to cope with administrative segregation or other specialized single-cell housing assignment might also be at increased risk of suicide. In addition, an inmate who is in the early stages of recovery from severe depression is very much at risk.

Table 60.3–15.
Periods of High Risk for Suicide in a Jail or Prison

At the time of admission to a lockup, jail, or prison
After new legal problems (e.g., new charges, additional sentences, institutional proceedings, denial of parole)
After the receipt of bad news regarding one's self or one's family (e.g., a serious illness in the family or the loss of a loved one)
After suffering some type of humiliation or rejection (e.g., sexual assault)
On being placed in and unable to cope with administrative segregation or other specialized single-cell housing assignment
In the early stages of recovery from severe depression

Correctional facilities must have written policies and clearly defined procedures instituting a program for identifying and responding to suicidal inmates. In addition, the quality assurance activities in the facility should confirm its suicide prevention practices.

Self-harm is a general term for a variety of activities, including self-mutilation without the wish to die. It takes many other forms, such as swallowing dangerous objects and suicide "gestures." The following are some examples of self-harm: cutting, carving, scratching, burning, biting, banging head or hitting other body parts, ripping or pulling skin or hair, interfering with wound healing, ingestion of sharps (ingestion of glass, razor blade or parts, or other items) and toxic or inedible substances or objects, and object insertion (e.g., into urethra, anus, vagina, extremities, wounds). Self-harm in this sense invokes questions of malingering and manipulation.

Although there are many more suicide attempts and incidents of nonlethal self-harm than completed suicides, any threats of self-harm or

self-injurious behavior must be taken seriously. Self-harm behaviors, regardless of motivation, can result in significant morbidity and mortality. For this reason, even behaviors that do not lead to injury should be taken seriously and not dismissed as merely "suicidal gestures."

Because self-harm is hard to control, it is a serious challenge for correctional officials and a complex clinical situation for the correctional psychiatrist who must determine whether the self-harm is the result of a mental disorder. This requires a variety of means to assess the problem, including risk assessment methods and protocols, and might require temporary housing on a mental health unit.

When self-mutilation is determined not to be the product of a mental disorder, the inmate may be denied admission to a mental health unit. This may confuse family members, attorneys, courts, patient advocates, and correctional authorities, who view such behavior as a sign of mental instability. It is, therefore, essential that the correctional psychiatrist follows clear protocols for self-harm to determine whether it is a psychiatric symptom, a manipulative gesture to escape a dangerous situation, or an effort to manipulate the system for personal gain. Careful and thorough documentation and communication with custody personnel are the key factors. Treatment, however, must be individually tailored and calls for a team approach.

Seclusion and Restrain in Correctional Psychiatry

The use of seclusion and restraint in the correctional context is complicated by the fact that corrections facilities have procedures with regard to nonmedical seclusion and restraints. As a consequence, the correctional psychiatrist must be very familiar with the principles and practices related to this subject and be prepared to act on them.

Because of this importance, both the APA and the National Commission have addressed the complex issue of seclusion and restraint in correctional facilities. In 2006, the APA published the Resource Document on the Use of Restraint and Seclusion in Correctional Mental Health Care, which provides a detailed and nuanced discussion of this complex topic. It is strongly recommended that every practicing correctional psychiatrist carefully study and review this document. It "discusses the use of seclusion or restraint for purposes of mental health intervention in jails and prisons, in contrast to its use for correctional purposes (i.e., specifically custody reasons)," and states:

> The use of seclusion or restraint for correctional purposes is generally driven by classification and disciplinary issues unique to the correctional setting. For example, an inmate's security classification may require the use of handcuffs and leg irons (i.e., restraints) during movement outside of the inmate's cell or housing unit. Restraints may also be used by custody staff to control an inmate's assaultive behavior that is not related to mental illness. With few exceptions, cell extractions (both calculated use of force and on an emergency basis) by custody staff are governed by custody policies and procedures even when they involve mentally ill inmates.

However, there are generally special provisions in such policies and procedures when such a use of force involves the mentally ill inmate that usually includes attempted assessment/intervention by mental health staff prior to the use of force. Disciplinary segregation has many characteristics similar to seclusion such as confinement to a cell and restricted access to personal belongings. Custody guidelines for using these security measures are generally very different than those relevant to the use of seclusion or restraint for mental health purposes and will not be addressed in this document.

With regard to the correctional setting specifically, this APA Resource Document details such matters as the location where the restraints or seclusion are being employed, the time frames that are required, and the devices that are permissible under various circumstances. It includes an appendix quoting the authoritative 1999 Health Care Financing Administration (HCFA) rules for the use of seclusion and restraint.

The same principles are contained in NCCHC Standard entitled Use of Restraint and/or Seclusion for Mental Health Reasons, which states, "Clinically ordered restraint or clinically ordered seclusion are available for patients exhibiting behavior dangerous to self or others as a result of mental illness." The following definitions are illuminating:

Clinically ordered restraint is a therapeutic intervention initiated by medical or mental health staff to use devices designed to safely limit a patient's mobility. Custody restraint is not the same as clinically ordered restraint. When custody staff order the use of restraints, medical staff monitor the health status of inmates while in custody restraint and mental health staff respond to security or medical staff requests for consultation regarding the use of, or response to, custody-ordered restraint.

Clinically ordered seclusion or "time-out" is a therapeutic intervention initiated by medical or mental health staff to use rooms designed to safely limit a patient's mobility. It is not the same as segregation or isolation as generally applied in corrections (i.e., referring to housing separate from general population and used primarily for punitive purposes). The term also does not refer to the temporary imposition of "lockdown" as directed by correctional authority in order to regain control in a tier, building, or facility.

Malingering

Common sense points to the fact that malingering is a pervasive problem in correctional facilities and of special concern for the correctional psychiatrist. Throughout the period of incarceration and the various correctional procedures and processes and at every level, secondary gain can be substantial. Moreover, given the character and temperament of certain correctional offender patients, it calls for constant vigilance.

The ability to detect malingering in the correctional context is especially important not only because of its ubiquity but also because although the mere presence of malingering does not preclude a finding of mental illness. As a rule, malingerers can be wasteful of resources and, therefore, detract

from the quality of services that correctional psychiatrists can provide to inmates who are truly in need of treatment. Under suitable conditions, such as when a corrections officer is part of a treatment team, the long-term observations of the security staff can provide valuable information with regard to an inmate's malingering.

Inmates in Segregation Housing

There remains a great amount of controversy regarding the empirically validated psychological effects of segregation and what effects, specifically this has on individuals with or without serious mental illness. In sum, it appears that prolonged segregation exposes individuals to potential psychological, physiological, and medical risks. Those with SMI have special vulnerability to the adverse effects of social isolation. Psychiatrists should assist correctional systems with the development of evidence-based behavioral interventions that avoid the potential harm of social isolation.

One example where psychiatrists and mental health professionals can help provide guidance to custody is the following. Consultation and training for disciplinary hearing officers before conducting the due process hearing (that would result in placement into a restrictive housing setting such as administrative segregation, disciplinary housing, or the like) helps them to recognize when mental illness is a contributor to inmate misconduct. Mental health professionals should be contacted and consulted by disciplinary hearing officers when mental illness is suspected. The NCCHC Standard broadly defines segregated inmates as "those isolated from the general population and who receive services and activities apart from other inmates." These housing units can be labeled as "administrative segregation, protective custody, disciplinary segregation, or a supermax tier." It notes that "for the purposes of this standard, the living and confinement conditions define the segregated status, not the reason an inmate was placed in segregation." The NCCHC Standard has an additional category—"extreme isolation"—which "refers to situations in which inmates are seen by staff or other inmates fewer than three times a day."

The APA Guidelines assert that "the challenge of providing correctional psychiatric services to inmates in segregation housing is a critical reflection of the correctional facility's quality of care and the inmate's access care."

It is not surprising that especially in jails and prisons that lack adequate mental health services, inmates with severe mental illness often find their way into segregation housing, often quite unnecessarily, as when their mental illness prevents them from understanding or adhering to correctional rules.

The APA Guidelines state that when an inmate is segregated for any reason from the general population, the responsibility to address serious mental health needs remains in effect (Table 60.3–16). Indeed, because of the stressful nature of segregation housing, facilities should make special

efforts to assess and address mental health treatment needs in these settings. Although correctional psychiatrists should not be involved in the decision to place an inmate in disciplinary segregation, once an inmate is there, the psychiatric staff should monitor his or her mental health status because an inmate who is placed in and is unable to cope with administrative segregation or other specialized single-cell housing assignment might also be at increased risk of suicide.

Table 60.3–16.
The American Psychiatric Association Guidelines on Principles for Mental Health Services in Segregation Housing

The APA Guidelines recommend that in providing essential mental health service in segregation housing, the following principles should be observed:

No inmate should ever be placed in segregation housing solely because he or she exhibits the symptoms of mental illness, unless there is an immediate and serious danger for which there is no other reasonable alternative. (This principle does not refer to medical or psychiatric seclusion, which should follow state mental health law and professional practice.)

When an inmate is placed in segregated housing for appropriate correctional reasons, the facility remains responsible for meeting all of the serious medical and psychiatric needs of that inmate. Thus, such inmates must receive any mental health services that are deemed essential, their segregation status notwithstanding.

Inmates who are in current, severe psychiatric crisis, including but not limited to acute psychosis and suicidal depression, should be removed from segregation until such time as they are psychologically able to tolerate that setting.

Inmates who are known to have serious mental health needs, especially those with a known history of serious and persistent mental illness, when housed in segregation, must be assessed on a regular basis by qualified mental health practitioners to identify and respond to emerging crises at the earliest possible moment.

Institutions should provide for regular rounds by a qualified mental health clinician in all segregation housing areas. During these rounds, each inmate should be visited briefly to assess any emerging problem. The clinician should also communicate with segregation security staff in order to identify any inmate who appears to be showing signs of mental deterioration or psychological problems.

Similarly, the NCCHC recommends, “inmates who are seriously mentally ill should not be confined under conditions of extreme isolation.”

The Commission adds, “policies should also address how vulnerable populations like adolescent, elderly, pregnant or mentally ill inmates are managed in segregation. Since segregation can exacerbate pre-existing conditions staff should be extra cautious when placing vulnerable inmates in segregation.”

The APA developed a position statement (APA, 2012) that includes the following:

Prolonged segregation of adult inmates with serious mental illness, with rare exceptions, should be avoided due to the potential for harm to such inmates. If an inmate with serious mental illness is placed in segregation, out-of-cell structured therapeutic activities (i.e., mental health/psychiatric treatment) in appropriate programming space and adequate unstructured out-of-cell time should be permitted. Correctional mental health authorities should work closely with administrative custody staff to maximize access to clinically indicated programming and

recreation for these individuals.

The SCP (2013) also published a similar position statement.

ISSUES THAT ARE UNIQUE TO CORRECTIONAL PSYCHIATRY

The "Lifer": Inmates Serving Life Sentences

Psychiatrists may encounter inmates with one or more life sentences, life without parole, referred to as "lifers." These individuals pose a unique challenge to the correctional psychiatrist. Because these individuals will not be released from prison, many usual incentives such as parole, token incentives, and the like have less effect.

The APA Psychiatric Services in Correctional Facilities notes that "inmates draw on many resources to assist their adaptation to incarceration, such as family contact, peer support, work, education, and spiritual pursuits, including the practice of faith-based traditions." Some individuals may lack these psychosocial resources such as no regular family contact, no visitors, no one on their visitation lists, no one to help fund their commissary purchases, family estrangement, and the like.

Inmates on Death Row

"Death row" individuals, inmates housed on death row, who are awaiting exhaustion of legal appeals, or whose legal appeals are exhausted and remain housed within death row psychiatry pose a myriad of complicated clinical, ethical, and legal questions. Some examples of issues include the death row psychiatric assessment, considering "justifiable" depression, suicide risk management, individuals on death row who require psychiatric hospitalization, psychiatric treatment thereby restoring one's competency to be executed, and how to balance ethical/role boundaries between clinical and forensic work. Because of the lengthy appeals process and high stakes involved, mental illness and intellectual disabilities are commonly raised in postconviction proceedings typically on eighth amendment claims (arguing that the execution of the mentally ill and/or mentally retarded constitutes cruel and unusual punishment).

Inmates with mental illness on death row present the correctional psychiatrist with unique and complex ethical issues. Because the participation of a correctional psychiatrist in an execution would violate the principle of do no harm, Annotation 4 of Section 1 of the APA's Annotations states that "a psychiatrist should not be a participant in a legally authorized execution." Much has been written about the interpretation of the word "participate," and it is clear that the medical community is unanimous that any action that would further the implementation of an execution is prohibited. Furthermore, there is no doubt that this applies to evaluations of competency to be executed. However, what about treatment of an incompetent inmate to achieve competency? This is a controversial area,

and whereas in general the correctional psychiatrist is obliged to provide treatment to patients in need to relieve suffering regardless of the subsequent nonclinical consequences, it has been suggested that physicians who are opposed to the death penalty not practice on death row. Similarly, it was held in the case of State v. Perry that the patient retains the right to refuse treatment and the state cannot treat over objection so as to impose the death penalty.

The NCCHC Standard "Inmate Executions" states that "Mental health staff does not participate in inmate executions," and sets out the following compliance indicators:

An inmate's treating clinician does not provide a competency evaluation for execution.

Mental health services staff members do not assist, supervise, or contribute to the ability of another individual to directly cause the death of an inmate.

Mental health staff members provide ongoing mental health services to inmates in death row as clinically indicated according to individual treatment plans.

It is noted that the intention of this standard is that mental health staff not be involved in any action that would directly cause the death of an inmate because such involvement in inmate executions poses an ethical dilemma: "In order to be effective mental health care providers, correctional mental health care staff must maintain an appropriate professional relationship with their patients. This relationship is lost when mental health care staff assumes a role in inmate executions." The NCCHC Position Statement entitled Competency for Execution recommends that "the determination of whether an inmate is 'competent for execution' should be made by an independent expert" and not by correctional psychiatric staff.

On Being Asked to Witness the Use of Force. In what is a unique situation, correctional officials or inmates might request a correctional psychiatrist to witness incidents in which use of force is necessary (e.g., moving a recalcitrant inmate to a new cell or a different site). Although correctional authorities might hope that the psychiatrist will attest to the appropriateness of the action, and inmates might hope that the presence of such witnesses will limit unnecessary force, correctional psychiatrists would do well to avoid such purely custodial functions because it might place them in a complex no-win position.

Witnessing Inmate Abuse. A thorny problem for the correctional psychiatrist is what to do in the case of clear inmate abuse by security. Inmate abuse has been a topic that is under the radar for correctional psychiatrists, but although it has been the subject of "conditions of

confinement" litigation, it had not been dealt with directly as a professional issue.

This state of affairs was recently addressed by the publication of an NCCHC Position Statement entitled Correctional Health Care Professionals' Response to Inmate Abuse (Table 60.3–17).

Hunger Strikes in Jails and Prisons. Hunger strikes (the refusal of nutrition as a means of passive resistance) are a daunting problem in jails and prisons. Although the history of hunger strikes goes back at least to the Roman Empire and in the Western world, there is generally a right to go on a hunger strike that is protected as a matter of "privacy" or a "liberty interest" or can be regarded as an expression of free speech, clearly, in a correctional facility other factors come into play. Correctional authorities are charged with the "care and custody" of the inmates. Therefore, the following principles should apply:

- The correctional facility must provide for the evaluation of the situation for treatment if necessary and perhaps for legally allowable intervention.
- The medical services must provide for the evaluation of the medical stability of the inmate.
- The correctional psychiatrist must be available to evaluate such an inmate's psychiatric condition and whether there is a need for treatment.
- The correctional psychiatrist must also evaluate the motivation for the hunger strike and the inmate's capacity to understand the consequences of his or her actions and the medical decisions being made.

Although it might be understood that most inmates participate in hunger strikes for political or manipulative reasons and usually are not mentally ill, if an inmate with mental illness participates in a hunger strike to the point of an emergency, community standards allow for caregivers to decide what is in the best interests of a patient who is not competent to decide.

The majority of US courts that have considered the question hold that an order to force-feed does not violate a hunger-striking prisoner's constitutional rights because the institution in which the inmate is housed is responsible for his care while incarcerated and there are other compelling governmental interests, such as the preservation of life, prevention of suicide, and enforcement of prison security, order, and discipline, that outweigh any constitutional rights asserted. It is interesting to note, however, that international authorities argue to the contrary. International courts have tended to allow hunger strikes to proceed even to death or disability. Thus, forced feeding of prisoners is prohibited by the Tokyo Declaration of 1975, by the World Medical Association, and, affirmed in 1992, by the Declaration of Malta.

Table 60.3–17.
The National Commission on Correctional Health Care Position Statement on Correctional Health Care Professionals' Response to Inmate Abuse

Should correctional health staff witness or become aware of an inmate being subjected to harm in any of the forms described previously, it is their duty to report this activity to the appropriate authorities in order to protect patients and other inmates. The following principles are to guide correctional health care professionals in averting and reporting the mistreatment of inmates.

Correctional health care professionals' duty is to the clinical care, physical safety, and psychological wellness of their patients.

Correctional health care professionals should not condone or participate in cruel, inhumane, or degrading treatment of inmates. When such abusive treatment is either witnessed or suspected, they should identify and report such incidents to the appropriate authority.

Correctional health care professionals should refrain from participating, directly or indirectly, in efforts to certify inmates as medically or psychologically fit to be subjected to abusive treatment.

Correctional health care professionals should refrain from being present in the interrogation room, asking or suggesting questions, or advising authorities on the use of specific techniques of interrogation.

Correctional health care professionals should refrain from gathering health information for forensic purposes or sharing confidential health information or its interpretation to authorities for use in cruel, inhumane, or degrading treatment of inmates.

Correctional health care professionals should abstain from authorizing or approving any physical punishment of their patients and should refrain from being used as an instrument of their employer to weaken the physical or mental resistance of inmates are to be managed and what staff should do when abusive actions are suspected or witnessed.

Correctional health care professionals should review their employer's policies and procedures and work to ensure that they appropriately address how inmates are to be managed and what staff should do when abusive actions are suspected or witnessed.

Correctional administrators should ensure that policies and procedures address protections for employees who report the abusive actions of others.

Professional custody and health administrators should support efforts to eliminate abusive behavior toward inmates by ensuring that all staff receive regular training on appropriate and professional behavior in dealing with inmates. Use of experts outside of the correctional system can be helpful in providing objective training on this issue.

In 1994, the ACHSA adopted a Position Statement that allows forced feeding when the legal procedure for medication over objection has been followed:

The management of a hunger striking prisoner presents an ethical conflict between the obligation to preserve life and relieve suffering on the one hand, and the obligation to respect patient autonomy on the other. Correctional health care providers should make every reasonable effort to honor patient self-determination. However, the deprivation of freedom which incarceration imposes will affect decision making by the inmate patient. Therefore, special efforts in patient education and counseling are indicated.

A prisoner may begin a hunger strike for various reasons. If pursued, it may have a fatal outcome, unless there is medical intervention. Correctional health care providers have a special duty to protect inmates from self-destructive behavior. When a prisoner initiates a hunger strike, and efforts to counsel and dissuade are not successful in terminating the strike, the responsible physician is obligated to obtain thorough medical and psychiatric evaluations and to monitor weight loss and other parameters of deteriorating health in an appropriate medical setting. Treatment of intercurrent medical problems must be offered. The hunger-striking inmate should be informed that authorization will be requested to intervene with forced feeding when life is in danger.

Since forced feeding is a form of involuntary treatment, the responsible medical authority must seek the consent of the court or a legally appointed guardian, following the local requirements for involuntary treatment. The physician should make it clear to the court and to all others concerned, however, that ethical medical conduct in this setting demands medical intervention to preserve life from actions that are self-destructive.

ACHSA members should give careful consideration to the development of policies and procedures that provide a responsible and compassionate approach to the hunger-striking prisoner and which make it clear that efforts will be made to intervene with forced feeding in order to save life.

The APA Psychiatric Services in Correctional Facilities recommends the following: "Hunger strikes are considered a medical emergency, and psychiatrists are often asked to evaluate inmates to determine if they have the mental capacity to make an informed decision regarding continuation of the hunger strike, Psychiatrists can perform evaluations to rule out underlying mental illness as a cause of the hunger strike (e.g., suicidal depression or delusions regarding prison food)."

There is currently no consensus to guide psychiatrists in the United States on the ethical standards of force-feeding to end hunger strikes in correctional facilities. Although the World Medical Association and the AMA, with limited exceptions, have taken stands against the practice, the Supreme Court of Connecticut, in *Commissioner of Correction v. Coleman* (303 Conn. 800 SC 18721 [2012]), found that the state's interest in the inmate's health and legitimate penological interests in the safety of the facility outweighed the inmate's right to autonomy. Furthermore, the court ruled that "force-feeding did not violate constitutional rights to free speech and privacy."

Biomedical Research in Correctional Psychiatry

In many ways, biomedical research in correctional facilities presents the correctional psychiatrist with paradigmatic conflicts of dual loyalty. Therefore, the correctional psychiatrist must be aware of the many complex issues regarding research in jails and prisons.

Research with prisoners as subjects requires special legal safeguards and ethical, as well as clinical considerations. In addition to precautions accorded to each patient in studies, researchers must provide extra safeguards pertaining to the subjects' incarceration, in that the incarceration itself can be viewed as an inherently coercive setting that undermines an inmate's ability to provide freely given informed consent for participation in research.

Because research in correctional facilities has a long history of human subject abuse, there are special federal regulations governing research in these facilities. However, these federal regulations do not apply to certain other governmental agencies, such as the Department of Justice, which regulates its own extensive research in correctional facilities.

In addition, as the APA Guidelines also note, "there are other types of research design, including systems research, epidemiological and non-invasive studies that can yield useful information about the precursors of mental illness in offenders, incidence and prevalence of mental illness, program design, physical plant design, and planning for services."

The Correctional Psychiatrist's Role in Specific Custody and Security Issues

Classification Committees

To be fully effective, the corrections system's classification processes must have some basic information about inmates' psychiatric needs. For example, an inmate might need to be assigned to a prison with programs and resources for the chronically mentally ill. In addition, suicide prevention protocols generally recommend that inmates who are exhibiting signs of withdrawal or depression should not be single-celled. Accordingly, unit classification committees should be aware of psychiatric conditions that might affect where inmates are housed or assigned to work. This presents a problem of how to provide important information to classification groups about inmates' mental health without violating confidentiality. It has been suggested that a form be devised that summarizes any medical restrictions regarding unit housing or job placement without revealing the inmates' precise condition or diagnosis.

Institutional Transfers. Classification committees continuously juggle custody and medical classifications of inmates with their workforce requirements and available space. Most institutions provide a list of the next day's transferees to correctional psychiatric staff so that they can assemble the inmate's mental health records. Correctional psychiatric staff should ensure that the receiving unit has the requisite psychiatric resources to continue to meet the patient's needs. Each facility should have a policy that allows psychiatric staff to put a psychiatric hold on any interunit transfer of an inmate. Decisions regarding admission to and discharge from inpatient facilities are the sole provinces of clinicians.

Disciplinary Hearings. As a rule, correctional psychiatric staff should not participate in disciplinary hearings and should never be part of the punishment decisions. However, there are occasions when information from psychiatric staff might be helpful in protecting the inmate from inappropriate discipline. For example, a schizophrenic patient might refuse to comply with a correctional officer's directions because of negativism rather than defiance.

PERSISTING CHALLENGES IN CORRECTIONAL PSYCHIATRY: AN AGENDA FOR ADVOCACY

Although recent years have brought considerable progress in the reform of

the correctional system, there remain significant challenges to improving correctional psychiatric care and treatment. To put these challenges into perspective, it must first be first acknowledged that many problems have been solved and many barriers have been overcome or resolved. Historically, the public was reluctant to provide better conditions for those who broke its laws or offended its sense of morality. However, in recent years, the pendulum has been swinging away from a policy that prioritized retribution, and it is now being increasingly understood that incarceration represents opportunities for rehabilitation and the prevention of recidivism. Indeed, as correctional facilities have become increasingly overcrowded and costly to the taxpayer, such policies as harsh sentences for minor drug offences and "three strikes" laws are being questioned. Most important, whereas in the past correctional leadership had concerns that psychiatric services might interfere with and jeopardize security, in recent years, this has evolved into an understanding of the value of psychiatric services to security.

One persisting problem is that many correctional facilities are still located in rural areas and, as a consequence, are out of the medical mainstream, with limited access to psychiatric specialists and, in this regard, are underserved. Special efforts are being initiated at both the governmental and the professional levels to offer incentives to attract psychiatrists, especially early career psychiatrists, to serve, at least for some period of time, in such areas.

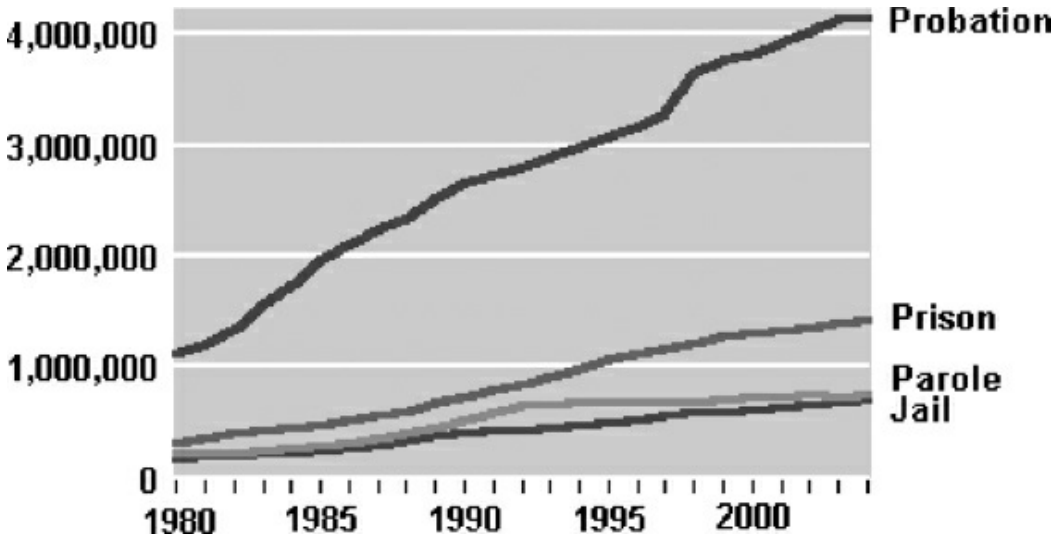

FIGURE 60.3–2. Adult correctional populations from 1980 to 2004.

In the past, there was no doubt that most correctional facilities were not satisfying places to work. Now, as the result of efforts of advocacy organizations, professional organizations, and the development of guidelines, standards, and accreditation to improve the delivery of psychiatric services in jails and prisons, this is also changing. Most

important, there was concern that the correctional atmosphere subverted the atmosphere of trust that is necessary for the doctor–patient relationship, which made many psychiatrists uncomfortable in a correctional environment. This, too, is changing, and the correctional institutional barriers that sometimes prevented the practice of good psychiatry or made professional psychiatric practice challenging are, in most cases, a thing of the past.

Nonetheless, there are a number of major persisting problems that can be regarded as social policy issues: overcrowding, the criminalization of the mentally ill, and racism in the criminal justice system.

Overcrowding

The overcrowding that has resulted from the increase in the correctional population over the last three decades has made many large jails, such as the Cook County Jail in Chicago, the Los Angeles County Jail, and Harris County Jail in Houston, Texas, the country's largest mental institutions (Fig. 60.3–2).

In 2006, more than 7.2 million people were on probation, in jail or prison, or on parole. At year-end 2006, this represented 3.2 percent of all US adult residents, or one in every 31 adults. State and federal prison authorities had jurisdiction over 1,570,861 inmates at year-end 2006: 1,377,815 in State jurisdiction and 193,046 in Federal jurisdiction. Local jails held 766,010 persons awaiting trial or serving a sentence at year-end 2006. An additional 60,222 persons under jail supervision were serving their sentence in the community.

The Criminalization of the Mentally Ill

Some of the factors contributing to the high rates of people with serious mental illness in the criminal justice system include, as reasons for the increased entry, that is, at the "front door" of correctional facilities, the process of deinstitutionalization and its accompanying phenomenon of "transinstitutionalization," the inadequate access of the seriously and persistently mentally ill to community treatment, the fact that these patients are arrested at disproportionately higher rates and often have co-occurring substance-related disorders, the prejudice that the mental ill are more violent, and the use of jails and prisons as housing of last resort (Fig. 60.3–3).

At the "back door," there are problems of discharge planning and reentry into the community, including the facts that these patients spend longer periods of time incarcerated, the pathogenic nature of incarcerated environments, the lower rates of parole, and the higher rates of recidivism due to inadequate community mental health services.

Finally, there are obvious overarching public policy issues that

permeate the criminal justice/mental health realm. The United States has the largest percentage of its population incarcerated than any other country in the world. Since 1989, the National Drug Control Strategy, which called for a mandatory minimum sentence for drug crimes, caused the jail and prison populations to grow at a rapid rate. In 1985, the percentage of sentenced drug offenders was approximately 13 percent. Currently, it is more than 30 percent. The importance of the relationship between the increase in the number of inmates incarcerated for drug offenses and the disproportionate increases in the numbers of female, nonwhite, and foreign-born inmates cannot be overstated (Fig. 60.3–4).

RECENT DEVELOPMENTS IN CORRECTIONS

Lesbian, Gay, Bisexual, and Transgender Persons and Gender Dysphoria

According to the APA Psychiatric Services in Correctional Facilities, trauma and stigma experienced by lesbian, gay, bisexual, or transgender (LGBT) individuals can increase their psychological distress and decrease their willingness to seek mental health services. When incarcerated, they risk additional discrimination and abuse. Compared with heterosexual inmates, studies have reported that they have 10 times the rate of sexual victimization by other inmates and up to three times the rate of sexual victimization by staff. This risk needs consideration on arrival at any correctional facility. LGBT individuals differ in degree of comfort with their sexual orientation or gender identity and whether they are “out” or “closeted.” Their choice about disclosure can change when moving from the community to jail or prison and between institutions. A respectful and nonjudgmental approach by the correctional psychiatrist can help LGBT inmates explore their choices and cope with their concerns. Transgender inmates, for example, may feel distress when housed in settings that do not allow them to express their gender identity or when they elect to stay closeted because of safety concerns. A therapeutic relationship in which they feel known by someone safe can improve emotional well-being.

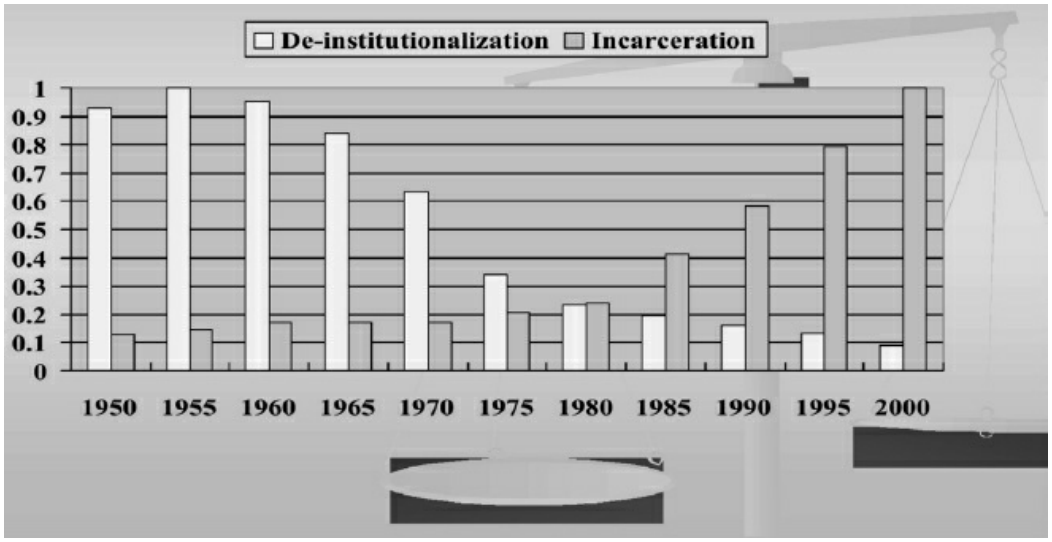

FIGURE 60.3–3. Transinstitutionalization from 1950 to 2000.

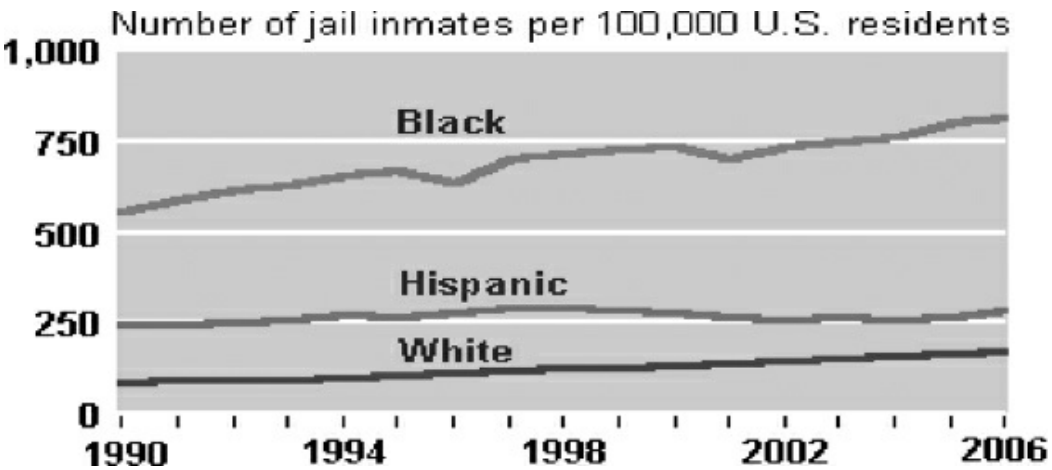

FIGURE 60.3–4. Jail incarceration rates by race and ethnicity, 1990 to 2006.

The evaluation and treatment of gender dysphoria, formerly known as gender identity disorder, is a particularly challenging area. The evaluation and treatment and decision-making processes of whether to continue, discontinue, or initiate hormonal treatment for offender patients are complicated. This is further complicated within correctional settings and even more so due to correctional systems being subject to PREA audits and enforcement. Custody leadership makes decisions regarding classification, housing, transportation issues, and whether they allow access to gender-preferred clothing and toiletries and housing assignments. Beyond hormonal treatment, many correctional systems do not have uniform policies about permissible treatment for transgender inmates. To date, no correctional systems have provided sex reassignment surgery, but this may change in response to ongoing litigation by transgender inmates (e.g., two current high-profile cases are in Massachusetts and California). Consistent with the prevailing position of professional medical organizations, including the APA, courts and society in general are increasingly

recognizing gender dysphoria as a medical condition appropriate for treatment.

USING TECHNOLOGY IN CORRECTIONS: THE ELECTRONIC MEDICAL RECORD AND TELEPSYCHIATRY

Many county, state, and federal correctional systems have moved toward the implementation of correctional EMR systems. This is particularly helpful in multisite systems, or high-volume intake facilities with rapid turnover, or alternatively in large systems where there is common movement, transfers, and turnover of inmates both intra- and intersystem. Thorough medical documentation is a useful protection against litigation and to help clarify an individual's assessment and treatment plan and to ensure continuity of care.

Telepsychiatry is becoming a more accepted and standard practice in correctional settings and in the community as the national shortage of psychiatrists increases. Greater comfort by correctional psychiatrists and other mental health professionals with this technology increases accessibility of expeditious services for patients, and due to less need for off-site transport for psychiatric evaluation, this technology reduces the cost and burden on staff to travel with an offender (typically at a minimum at least two correctional officers per transport). This use of technology reduces the risk of escape and other community safety risks. It also eliminates the risk of assault to psychiatrists as that which can and does occur on the unit, albeit rarely, during face-to-face encounters.

THE FUTURE OF CORRECTIONAL PSYCHIATRY

A confluence of forces is accelerating the reform of the correctional system to the benefit of correctional psychiatrists and their patients. Intense advocacy efforts, especially by families members of persons with serious and persistent mental illness in the criminal justice system, through patient advocacy groups such as the National Alliance on Mental Illness and Mental Health America, are becoming increasingly effective in lobbying legislatures to increase mental health funding, staffing, and psychiatric services in jails and prisons.

The APA is joining with other professional groups to develop programs and priorities for legislative action, including such groups as the Campaign for Mental Health Reform and the Mental Health Liaison Group. This has, in turn, drawn the attention of county and state governments not only in terms of the requirements of the humane aspects of care and treatment and concerns about litigation but also the fiscal aspects. More and more evidence is accumulating that it is of benefit to the taxpayer to prevent recidivism by means of diversion programs, drug courts and mental health courts, jail-based and other competency restoration programs and well-developed reentry programs, as well as psychiatric services in juvenile

justice settings, and in adult jails and prisons.

In his 2004 State of the Union Address, President George W. Bush said, "America is the land of second chance, and when the gates of the prison open, the path ahead should lead to a better life." The Second Chance Act of 2007 mentioned earlier assists states and local government entities, in partnership with nonprofit organizations, to establish prisoner reentry demonstration projects.

It is becoming increasingly clear that part of the overcrowding problem, as well as the escalating costs, can be addressed by providing correctional psychiatrists with the means to keep persons with serious and persistent mental illness from being warehoused in the jails and the prisons.

The mental health field is becoming increasingly sophisticated in targeting inadequate services in areas such as discharge planning and reentry policy in an effort to bring about major reforms, as in cases such as the Brad H. case mentioned earlier that ensures efficient and effective discharge planning and reentry programs. The legal community through the American Bar Association has produced its own Criminal Justice Mental Health Standards. The concept of "therapeutic jurisprudence" has found adherents in the judiciary.

The media is becoming increasingly aware of the importance of the numbers of persons with mental illness in the criminal justice system, and innumerable editorials have dealt with various aspects of reform. There are now journalists, websites/blogs (in Texas, there is Grits for Breakfast), who specialize in various aspects of corrections and correctional psychiatry.

The academic community is no longer ignoring the shame of the incarcerated mentally ill and recognizes the public health implications of this situation. This is evidenced by the flood of recent textbooks and the many articles in the literature and the ever-growing numbers of conferences on correctional psychiatry. In 1996, the Accreditation Council on Graduate Medical Education (ACGME) recognized the subspecialty of forensic psychiatry, which includes substantial correctional components. The didactic curriculum for the forensic fellowship training requires that the "psychiatric curriculum includes the… evaluation and treatment of incarcerated individuals," and the "clinical experience must include experiences in the… treatment of persons involved in the criminal justice system." In addition, "Clinical placement must provide [forensic fellows]… with experience in the evaluation and management of acutely and chronically ill patients in correctional systems such as prisons, jails, community programs, and secure forensic facilities."

Sophisticated demographic studies demonstrate that a significant percentage of the population of the mentally ill in jails and prisons are inappropriately incarcerated in correctional contexts and should be transferred to the civil mental health system.

Professional organizations are providing a voice for correctional

psychiatrists. The APA has since 2002 supported a Caucus of Correctional Psychiatrists. The importance of this subject of addressing both health and mental health in jails and prisons has led the ACA to adopt standards and an accreditation process and the NCCHC to develop separate mental health standards.

The advent of private companies (including some specializing in mental health care) competing to win contracts for the management and administration of correctional facilities has added the potential for greater accountability, by contract, to state and local governmental agencies.

Correctional psychiatrists in particular are aware that medical science is advancing on issues such as the biological bases of the characteristics that lead to criminality and deviance and the prospects of targeted change and evidence-based treatment. For example, there have been seminal studies of the developmental capacities of youth and their responsibility for various criminal acts.

Finally, in recent years, comparative studies have been carried on from the international perspective to compare correctional psychiatric systems in the United States with those in other countries with regard to the costs and benefits of different systems of providing services.

These are days of great ferment, with an increasing effort to improve psychiatric services in jails and prisons, which points to an encouragingly bright future for correctional psychiatry and for jails and prisons as a place for excellence in practice, teaching, and ethical and IRB-approved research.

References

American Psychiatric Association (APA) Official Actions. Position Statement on Access to Care for Transgender and Gender Variant Individuals. Approved by the Board of Trustees, July 2012. Approved by the Assembly, May 2012.

American Psychiatric Association (APA) Official Actions. Position Statement on Capital Punishment: Adoption of AMA Statements on Capital Punishment. Approved by the Board of Trustees, July 2008. Approved by the Assembly, May 2008.

Baillargeon J, Binswanger IA, Penn JV, Williams BA, Murray OJ. Psychiatric disorders and repeat incarcerations: the revolving prison door. *Am J Psychiatry*. 2009;166(1):103–109.

Baillargeon J, Hoge SK, Penn JV. Addressing the challenge of community reentry among released inmates with serious mental illness. *Am J Community Psychol.* 2010;46(3–4):361–375.

Baillargeon J, Penn JV, Knight K, Harzke AJ, Baillargeon G, Becker EA. Risk of reincarceration among prisoners with co-occurring severe mental illness and substance use disorders. *Adm Policy Ment Health*. 2010;37(4):367–374.

Bell CC. The need for accreditation of correctional health care facilities. *J Natl Med Assoc.* 1984;76:571.

Daniel AE. Preventing suicide in prison: a collaborative responsibility of administrative, custodial, and clinical staff. *J Am Acad Psychiatry Law*. 2006;34(2):165–175.

Diamond PM, Wang EW, Holzer CE III, Thomas C, des Anges Cruser. The prevalence of mental illness in prison. *Adm Policy Ment Health*. 2001;29(1):21–40.

Fazel S, Danesh J. Serious mental disorder in 23000 prisoners: a systematic review of 62 surveys. *Lancet*. 2002;359(9306):545–550.

Fellner J. *Old Behind Bars: The Aging Prison Population in the United States*. New York: Human Rights Watch; 2012. Available at https://www.hrw.org/sites/default/files/reports/usprisons0112webwcover_0_0.pdf. Accessed July 11, 2016.

Hayes LM. *National Study of Jail Suicide: 20 Years Later*. Washington, DC: U.S. Department of Justice, National Institute of Corrections; 2010.

Hoffmann HC, Dickinson GE. Characteristics of prison hospice programs in the United States. *Am J Hosp Palliat Care*. 2001;28(4):245–252.

Keram EA. Hunger strikes. In: Trestman RL, Appelbaum KL, Metzner JL, eds. *Oxford Textbook of Correctional Psychiatry*. New York: Oxford University Press; 2015.

McGlasson T, Champion MK, Penn JV. Geriatric offenders: evaluation and treatment within correctional settings. In: Holzer J, Kohn R, Recupero P, Ellison J, eds. *Oxford Textbook of Geriatric Forensic Psychiatry*. New York: Oxford University Press. In press.

McKee J, Penn JV, Koranek A. Psychoactive medication misadventuring in correctional health care. *J Correct Health Care*. 2014;20(3):249–260.

Metzner JL. An introduction to correctional psychiatry: part III. *J Am Acad Psychiatry Law*. 1998;26(1):107–115.

Metzner JL. Class action litigation in correctional psychiatry. *J Am Acad Psychiatry Law*. 2002;30(1):19–29, discussion 30–32.

Metzner JL, Tardiff K, Lion J, et al. Resource document on the use of restraint and seclusion in correctional mental health care. *J Am Acad Psychiatry Law*. 2007;35(4):417–425.

Moloney KP, van den Bergh BJ, Moller LF. Women in prison: the central issues of gender characteristics and trauma history. *Public Health*. 2009;123(6):426–430.

National Commission on Correctional Health Care. *Position Statement. Solitary Confinement (Isolation)*. Chicago, IL: National Commission on Correctional Health Care; 2016. Available at http://www.ncchc.org/filebin/Positions/Solitary-Confinement-Isolation.pdf. Accessed July 11, 2016.

National Commission on Correctional Health Care. *Position Statement: Transgender, Transsexual, and Gender Nonconforming Health Care in Correctional Settings*. Chicago, IL: National Commission on Correctional Health Care; 2015. Available at http://www.ncchc.org/filebin/Positions/Transgender-Transsexual-and-Gender-Nonconforming-Health-Care.pdf. Accessed July 11, 2016.

National Commission on Correctional Health Care. *Standards for Health Services in Jails*. Chicago, IL: National Commission on Correctional Health Care; 2014.

National Commission on Correctional Health Care. *Standards for Health Services in Juvenile Detention and Confinement Facilities*. Chicago, IL: National Commission on Correctional Health Care; 2015.

National Commission on Correctional Health Care. *Standards for Health Services in Prisons*. Chicago, IL: National Commission on Correctional Health Care; 2014.

National Commission on Correctional Health Care. *Standards for Mental Health Services in Correctional Facilities*. Chicago, IL: National Commission on Correctional Health Care; 2015.

Ochoa KC, Pleasants GL, Penn JV, Stone DC. Disparities in justice and care: persons with severe mental illnesses in the U.S. immigration detention system. *J Am Acad Psychiatry Law*. 2010;38(3):392–399.

Penn JV, Thomas C. Practice parameter for the assessment and treatment of youth in juvenile detention and correctional facilities. *J Am Acad Child Adolesc Psychiatry*. 2005;44(10):1085–1098.

Rold WJ. Thirty years after Estelle v. Gamble: a legal retrospective. *J Correct Health Care*. 2008;14(1):11–20.

Scott CL, Gerbasi JB, eds. *Handbook of Correctional Mental Health*. Washington, DC: American Psychiatric Press; 2005.

Steadman HJ, Osher FC, Robbins PC, Case B, Samuels S. Prevalence of serious mental illness among jail inmates. *Psychiatr Serv*. 2009;60(6):761–765.

Trestman RL. *Psychiatric Services in Correctional Facilities*. 3rd ed. *A Work Group Report of the American Psychiatric Association*. American Psychiatric Publishing; 2015.

Trestman RL, Appelbaum KL, Metzner JL, eds. *Textbook of Correctional Psychiatry*. New York: Oxford University Press; 2015.

U.S. Department of Justice. Prison rape elimination act final rule (28 CFR part 115). Fed Regist 77(119):37105-37232, 2012. Available at http://www.prearesourcecenter.org/sites/default/files/library/2012-12427.pdf. Accessed July 11, 2016.

Yampolskaya S, Winston N. Hospice care in prison: general principles and outcomes. *Am J Hosp Palliat Care*. 2003;20(4):290–296.

▲ 60.4 Neuroimaging, Psychiatry, and the Law

JOSEPH R. SIMPSON, M.D., PH.D., AND JONATHAN D. BRODIE, M.D., PH.D.

INTRODUCTION

Methods for visualizing the structure and function of the living human brain have become increasingly informative over the past four decades. Dramatic scientific and technological advances have made neuroimaging ubiquitous in 21st century psychiatric research. Yet in spite of the great strides made in the development of the techniques, neuroimaging is not standard practice for diagnosing the conditions most commonly encountered by psychiatrists, such as mood disorders, anxiety disorders, and psychosis. Currently, clinical psychiatrists use neuroimaging on a routine basis only to rule out neurological causes of their patients' symptoms.

However, research is not the only segment of the mental health field where applications of neuroimaging have been rapidly expanding. Neuroimaging is being introduced as evidence in court with growing frequency. The ongoing refinement of the techniques has allowed this use of neuroimaging to move far beyond its original applications, such as identifying structural lesions caused by traumatic injury, cerebrovascular accident or hypoxia. In an ever-increasing number of civil and criminal cases, neuroimaging is being offered as evidence to support a wide range of neuropsychiatric diagnoses and conditions. Some attorneys have argued that neuroimaging can establish the presence of specific current or past mental states or mental capacities, or the lack thereof. The past decade has also seen attempts to use modern neuroimaging techniques to peer into a person's thoughts and classify him or her as being truthful or deceptive.

The appeal of neuroimaging for attorneys is obvious: A brightly colored image, produced by expensive, sophisticated technology, is infused with an aura of irrefutable scientific truth. The seemingly authoritative "brain scan" can be pointed to as proving the assertions the attorney is making. A judge or a jury may find it difficult to disregard the evidence of their visual sense, which seems to demonstrate abnormal brain activity in, for example, the plaintiff in a civil trial who is arguing that he has been injured, or the defendant in a criminal trial, arguing that she is not responsible for the alleged crime due to a mental disorder.

In an adversarial legal system such as the one used in the United States, attorneys act as advocates for one side. Much of their work in court, therefore, is focused on persuading the judge or jury by presenting their arguments in the most favorable light. They are under no obligation to point out the limitations of their evidence, or the merits of any evidence

which would tend to undermine their case. Of course, the advocate does not have free rein. The use of neuroimaging is, like all scientific evidence in the legal arena, constrained by two mechanisms: the so-called "gatekeeper" function of the judge, who must rule on the admissibility of proposed scientific testimony, and rebuttal by the opposing attorney, through cross-examination and/or the calling of an expert who will contradict some or all of the initial expert's testimony.

The use of neuroimaging in a legal context raises a number of fundamental questions. What imaging methodologies are potentially helpful in court, and for what purposes? Are there some questions that neuroimaging cannot address? In terms of social policy, do some current or potential future applications of neuroimaging risk infringing on the rights of those who are subjected to them, or of interfering with the fact-finding role of the judge or jury?

It is an oft-repeated truism that the human brain is the most complex system known. Psychiatrists know well the difficulties inherent in attempting to identify and analyze thoughts, emotions, and motivations, and their causes, effects, and interactions. Neuroimaging is not now, and will not soon become, similar to fingerprinting or DNA evidence, where the result of the test essentially settles the question being asked Either the DNA matches or it does not. Such stark dichotomies are generally not the case where neuroimaging is concerned. This is especially true for psychiatric applications. Psychiatric neuroimaging is not used in a courtroom situation the way a plain x-ray film might be used to demonstrate a broken bone. Rather, neuroimaging findings are most often used in court to support the diagnostic findings of other forms of assessment, such as neuropsychological testing or a comprehensive psychiatric diagnostic evaluation.

The challenge faced by attorneys and the experts they retain is determining whether the neuroimaging results they seek will be based on a valid, appropriate use of the technology, and if so, whether the findings are relevant and add value to the legal argument being made. These are complex questions which must be addressed in each case where neuroimaging may be utilized. Expert witnesses who seek to incorporate neuroimaging into their testimony must have a thorough understanding of the capabilities and limitations of the technology, including an understanding of which questions neuroimaging is unlikely to shed light on. It is incumbent on the psychiatrist who proposes to make use of neuroimaging in a legal setting to ensure that the science is appropriate for the task, and that the attorneys do not draw unsupportable conclusions.

Furthermore, whether neuroimaging is appropriate in a given case is dependent not only on the science of the imaging technique, but also on the legal question to be addressed. The same data may be valid and relevant in one legal context but not another. For the past 2 decades, judges have

struggled with questions of the admissibility and relevance of neuroimaging findings. As neuroimaging techniques continue to be developed, refined, and applied in a wide range of legal situations, these debates will continue. The well-informed psychiatrist, with his or her knowledge of the brain and behavior, neuroimaging techniques, and forensic mental health concepts, is in a position to provide indispensable information to the court system throughout the legal process.

OVERVIEW OF IMAGING TECHNIQUES

Regardless of the technique, the objective of neuroimaging is to detect some property of the brain (as measured by attenuation of x-rays, magnetic moments or dipoles, electrical signals, radioactive events) without invading the brain. It is important to realize that each imaging technique has a characteristic sensitivity, precision, accuracy, and fidelity with respect to a physical and/or physiological process or processes, and is generally performed in an artificial environment divorced from the very circumstances and behaviors that prompted the test.

It is not obvious when looking at a functional or structural brain image that the image represents a series of complex manipulations of the raw data that the technology acquires. It is always necessary to construct and apply a model. A model, by definition, is a set of simplifying assumptions that are made to reconstruct the mass of acquired imaging data so it is interpretable in physiological terms in an experimental context or design. It is also necessary to perform statistical analysis and comparisons that are generally different in the forensic context than they are in a research environment. Nevertheless, the imaging expert is normally expected to draw experimental conclusions and clinical inferences from the imaging data, regardless of whether or not the appropriate statistical comparisons can be made.

What is being imaged are in fact electrical signals generated by the technology (e.g., photons from a radioactive event or annihilation, changes in a magnetic field, absorption of x-rays) that are detected, amplified, corrected for spatial localization and tissue attenuation, filtered and reconstructed, smoothed, fitted to an anatomical reference brain, and compared to either a similarly obtained reference state or to a reference group. Ideally, the reference group would be comprised of subjects who are matched for gender, handedness, age, education, general medical history, etc. In practice, these conditions are rarely met, and the reference groups tend to be smaller than is desirable or statistically appropriate.

What the technology is measuring, what is imaged, and what the user employing the technology is attempting to infer and interpret are specific to each modality. In the case of computed tomography (CT) scans, what is measured is the attenuation of x-rays as they pass through tissues of different density, while magnetic resonance imaging (MRI) scans measure

the magnetic properties of the protons in water molecules in their microenvironment as they “relax” after being polarized by an extremely strong homogeneous magnetic field. Different properties of the tissue environments can be highlighted by manipulation of the pulse sequences, which are the specific conditions of the magnetic perturbation of the baseline. DTI or diffusion tensor imaging is a subset of MRI in which the orientation of water molecules in neural fiber tracts are highlighted from the MRI data.

In contrast, the functional imaging techniques of positron emission tomography (PET) and single-photon emission computed tomography (SPECT) attempt to measure metabolic properties of the tissue by means of the introduction of an external radioactive tracer that “labels” a specific biochemical entity that is intrinsic to a brain process, such as blood flow or metabolism, or syndrome-specific markers such as abnormal proteins or receptor distribution and properties. Functional MRI (fMRI), which takes advantage of the association between regional blood flow and neural activity, has become the most widely used technique for functional brain mapping and visualization of stimulus-specific regional brain activations.

Over a relatively wide range of metabolic demand, increased demand is coupled with increased blood flow, which leads to an increase in oxygen extraction from the blood and a consequent reduction in the oxyhemoglobin/deoxyhemoglobin ratio. Since these two molecules have different magnetic properties, the BOLD (blood oxygen level dependent) signal that emerges can be detected, and is related in a direct but complex manner to neuronal activity. It is unfortunate that technical aspects of fMRI, such as the influence of magnetic field strength, field heterogeneity, subject movement, signal strength, tissue artifacts, etc., that can distort or even invalidate the imaging data, are often lost in the color-coded, highly processed, and distilled images that are generally far more clear than the data warrant.

A less widely used, yet experimentally similar functional imaging technique is quantitative electroencephalography (qEEG), which uses computerized techniques to more accurately localize the familiar electrical signals from the brain that are detected by scalp electrodes. The electrical changes can be distinguished either from baseline or from a comparison reference group after data analysis that includes the mathematical solution to the Inverse Problem, that is, the problem of identifying the intracerebral loci from which the signals arise.

At the time of this writing, the imaging techniques of general applicability in the courtroom have been broadly grouped into two large categories: structural imaging (CT and MRI, including DTI) and functional imaging (PET, SPECT, fMRI, and qEEG). All of these techniques in fact reflect physiological properties of the brain, and the distinction is largely predicated upon whether they are relatively static (state-independent) or

relatively dynamic (state-dependent), thus defining the context in which the data to be acquired will be interpreted. If the image is largely invariant over time and is not perturbed by transient stimuli, it is termed structural.

It is important to realize that there are two streams of data that must be joined to produce the eventual image of the physical or physiological data. The first stream is the instrument stream: this includes all of the considerations that must be accounted for in data acquisition and processing, such as specific signals versus random signals (noise), the homogeneity of the detection elements, the sensitivity and specificity of the method, and the stability of the measure during the time the image is being acquired.

The second stream is the physiological stream: this includes deconvoluting the total of specific and random neural activity that is observed in the field of view, correcting for variations in stimulus threshold and response, heterogeneity of response over the time course of the experiment, degradation of the data due to factors such as movement, location of the signal source in the brain, and session effects such as fatigue, loss of attention, change in mental state, etc. This is relatively easy with the structural imaging techniques such as MRI in which much of the information can be discerned by means of varying the signal acquisition parameters. It is more complex for the functional techniques.

Finally, one has to evaluate the fidelity and relevance of the data and its reproducibility over time (e.g., the next day, months, or years before or after the image was obtained). Yet the complexity of the human brain requires a set of simplifying assumptions to increase the signal-to-noise ratio. Are the particular assumptions reasonable? Can they be validated? Is the technology appropriate to address the question of interest? The answers to these questions are not necessarily simple or obvious. For example, can one reasonably expect that a scan of any type will reveal or predict a particular behavior or thought that took place months or years before? The task is not so different from that faced by a psychiatry resident attempting to make diagnoses on individual patients who are all peacefully asleep in their beds.

It must be kept in mind that for many techniques a simple challenge that can elicit a behavioral response does not give a signal adequate for either identification or comparison. In the research context, the challenge is often repeated multiple times, and the signals are then summed and compared to data acquired from a large group of subjects or to multiple acquisitions of the same subject in a reference state.

An aspect of brain imaging that is generally underemphasized but is of enormous importance to psychiatry and to forensic psychiatry in particular is the notion of resolution. Since an image is a pictorial representation of a dataset, it is generally appreciated that spatial resolution is of great significance, and hence that greater resolution is always a benefit. The

spatial limitation of any imaging technique is generally governed by the physical principles that are intrinsic to each methodology. Nevertheless, one can use statistical analysis techniques that enhance the apparent resolution, albeit at the expense of including further assumptions. Thus, a more detailed picture is not necessarily a better picture or a picture with greater fidelity to the actual physical/physiological system being imaged.

Temporal resolution can be divided into two parts and is related to the process to be evaluated. If one is interested in the structural image, the time to acquire that image can be relatively long without impairing the data (e.g., seconds to minutes or even longer) while for functional data, it is important to have a time of data acquisition that is consistent with the process to be measured. For example, it takes but milliseconds to form a thought, but only electrical techniques acquire data fast enough to be consistent with that process, while the more common functional techniques such as fMRI and PET take seconds or minutes to acquire sufficient data for analysis.

Paradoxically, one of the great difficulties of image analysis is the necessity to truncate much of the data. There is a tendency to emphasize the gross features that are relatively easy to distinguish because they are common to much of the population. This is done at the peril of ignoring much of the idiosyncratic data because it is difficult to isolate it from the noise. This is a long-standing problem for which there is no obvious solution. It should not be too surprising that the brain, like the heart and other organs, can develop collateral pathways in response to damage, and that individuals will develop unique pathways to use their brains to interact with their environment. Even with genetically engineered laboratory rats, reared together, the response to the same visual stimulus leads to heterogeneous brain activation. Unfortunately, the imaging techniques currently available lack the sensitivity to identify likely compensatory pathways, and these generally go undetected.

As a simple example of data truncation and elision, consider a reunion with an old friend with whom you have had no contact in years. Within a paragraph or two, a conversation ensues as if time has hardly passed. The clear implication is of the enduring stability of character structure and personality organization and the relative insensitivity of these to enormous amounts of intervening stimuli and contexts. In the courtroom, there is the implicit, but erroneous, expectation that brain imaging is capable of filling in the gaps between fact and fantasy, between wishful thoughts and actual deeds, an expectation more consistent with the appeal of phrenology than with the limitations of present science.

COMMON USES OF NEUROIMAGING IN COURT

In recent years there have been attempts to introduce neuroimaging as evidence in a wide range of legal contexts. However, there are a few legal

scenarios for which the presentation of neuroimaging evidence is becoming especially common. These can be divided into two broad categories. The first group consists of cases in which the diagnosis in question clearly stands to be supported by neuroimaging findings, and, conversely, weakened if no neuroimaging abnormalities are found. In these cases there is an easily understandable and relatively direct connection between the neuroimaging and the clinical condition at issue. In some cases, neuroimaging may have been a part of the diagnostic workup even before any litigation began. The primary examples are traumatic brain injury and dementia. Brain injuries not caused by mechanical traumas (e.g., automobile accidents, falls, assaults, gunshots), but resulting from medical compromise (e.g., cardiac arrest, obstetrical complications, toxic exposures) also fall into this category.

For the second category, neuroimaging is used not because a diagnosis or suspected diagnosis suggests a high likelihood of identifying abnormalities, but because the standard of admissibility for the imaging evidence has been significantly lowered. This occurs specifically in the penalty phase of a capital case. With the question posed to the jury literally a matter of life and death, the defense will seek to introduce any and all evidence that may possibly sway the jury toward sparing the defendant's life. In this phase of a capital trial, the judge allows a wide range of evidence to be admitted for purposes of mitigation, and relaxes the standards which apply in other proceedings.

Examples of Legal Situations Where Neuroimaging May Be Introduced

There are many legal questions where one side seeks to establish either a specific clinical diagnosis, or that functional impairment is present. In civil proceedings such as personal injury and malpractice cases, the plaintiff can introduce neuroimaging to support a claim for damages. Questions of capacity arise in the context of contract disputes, wills, and guardianship. In these types of cases neuroimaging may be used to buttress the argument that the person examined currently lacks the relevant capacity. Disability, that is, determining whether someone is capable of working, is another area of law where neuroimaging could theoretically be applied, although this is currently less common than other uses.

In criminal cases, neuroimaging is most likely to be used for the purpose of diminishing culpability. Typically used by the defense, it can be used to argue diminished capacity (which, if successful, results in a conviction for a less severe charge, for example, second-degree murder instead of first-degree murder or manslaughter instead of second-degree murder). Neuroimaging might also be used when the defendant has pled not guilty by reason of insanity, or conceivably in an unconsciousness defense, which could result in a full acquittal. But perhaps the most common situation where neuroimaging is used in a criminal context is

death penalty mitigation, discussed in more detail below.

Traumatic Brain Injury

Head injuries are extremely common, and many result in some degree of disruption of brain functioning. This disruption is indicative of injury to the brain, that is, traumatic brain injury (TBI). Such injuries are typically classified as being mild, moderate, or severe. For moderate and severe traumatic brain injury, structural neuroimaging with CT or MRI will typically identify pathology acutely, such as hematomas and areas of acute intraparenchymal hemorrhage; later, in the subacute and chronic stages, areas of gray matter necrosis (encephalomalacia), gliosis, and enlarged ventricles (due to tissue loss) can be detected.

In some situations the organic damage resulting from a moderate or severe traumatic brain injury may be difficult to visualize. For example, in injuries caused by rapid acceleration or rotation of the head, axonal shearing, termed diffuse axonal injury (DAI), can occur at the microscopic level, potentially leading to significant functional impairments. Unless the area affected by the DAI is of significant size, it may be undetectable by standard CT and structural MRI modalities. If DAI is suspected, specific MRI sequences such as gradient-echo offer superior possibilities of detection.

Modern neuroimaging techniques, especially structural MRI, find one of their most robust and least controversial uses as courtroom evidence in cases of moderate and severe TBI. Most commonly, findings of structural damage are used to augment evidence of neuropsychiatric impairment obtained through neuropsychological test batteries and/or psychiatric interview and review of records. Similarly, evidence suggesting embellishment, exaggeration, or outright malingering of behavioral or cognitive problems is strengthened if no lesions are found with neuroimaging.

Functional neuroimaging with SPECT or PET is often performed in forensic cases of suspected TBI. For optimal utility, such studies should be correlated with structural imaging with MRI when possible, or else with CT. Typical findings are hypometabolism in, and in some cases extending beyond, the cortical areas of damage seen on the structural scan. Correlating a functional scan obtained months or even years after the TBI with a CT or MRI scan obtained in the acute period after the initial incident provides additional evidence that the functional changes are the result of the TBI and not some other cause.

Abnormalities in SPECT and PET are relatively nonspecific, and can be seen in a variety of psychiatric and medical conditions in addition to TBI. These imaging techniques do not produce a consistent pattern in TBI patients that allows for the exclusion of other causes. In cases of mild TBI (mTBI), some authors have warned against the use of SPECT or PET due to

the high rate of false positives.

There is controversy surrounding the use of both structural and functional neuroimaging for legal purposes in mTBI. Even the definition of what constitutes mTBI is not free from debate. Beyond dispute is the fact that functional recovery after mTBI is highly variable across patients, both in terms of time course and in terms of whether there is ultimately complete recovery or residual impairment. While most people who experience mTBI recover fully with no sequelae, a small subset are left with deficits and some are significantly impaired.

When an alleged mTBI and its aftereffects are central to a legal case, both sides can put forth strong arguments in favor of their point of view and dispute key elements of the case. Was there in fact an mTBI in the first place? If there was, is there ongoing functional impairment, or is the patient deliberately exaggerating or manufacturing deficits for secondary gain? Or is there another, unrelated cause explaining the patient's current difficulties?

Under these circumstances it is not surprising that the legal system will look to neuroimaging in the hopes of providing some clarity to the proceedings. Unfortunately, it is not clear that the current state of the science is up to the task when it comes to mTBI. Structural MRI as well as functional techniques such as SPECT and PET often show no abnormalities in mTBI, and there is no specific pattern of findings with these modalities that confirms the existence of an mTBI. Some expert witnesses do employ SPECT and PET in mTBI, but this approach is not universally accepted, and their testimony is not always admitted into evidence by the judge.

The relatively new MRI technique of diffusion tensor imaging (DTI), which generates maps of white matter tracts, allows for quantification of the integrity of those tracts. In theory, this method could identify disruption of white matter connectivity even in cases where other structural or functional techniques do not show any abnormalities. However, the application of DTI in a legal setting to establish the presence of an mTBI is currently controversial and is by no means accepted by all neuroimaging experts. DTI is not a standardized procedure like a CT scan or even a standard structural MRI scan. Currently it is primarily a research tool, with different laboratories using different parameters, statistical methods, and mathematical transformations of the raw data. Furthermore, DTI shares weaknesses common to most neuroimaging techniques: in general, it cannot identify the cause of any abnormalities that are seen. It also cannot establish that such identified abnormalities are the cause of the patient's functional impairments.

To summarize, structural neuroimaging, primarily with MRI, is routinely used in legal cases involving moderate and severe TBI with little to no controversy. In some cases functional neuroimaging with PET or SPECT is also helpful. The state of the science when it comes to mild TBI is

unsettled, with experts arguing vehemently both for and against the use of functional neuroimaging as a method for illuminating the legal question.

Dementia

Dementia, also known as major neurocognitive disorder, is another diagnostic category for which it is relatively common to see neuroimaging data introduced in legal proceedings. Except for cases of dementia directly resulting from a TBI and a few other rare circumstances, dementia does not result from someone's harmful acts or omissions, so legal cases involving dementia are more likely to center on issues of capacity or of criminal responsibility, rather than personal injury or malpractice. Many of the same caveats that apply in cases of TBI also apply here. Neuroimaging is not diagnostic of dementia, and it is possible to have dementia without definitive neuroimaging abnormalities, as well as to have abnormal imaging findings suggestive of a dementing illness without actually meeting criteria for dementia.

In recent years there has been significant progress in identifying and differentiating various forms of dementia, including, for example, the isolation of genetic markers for different subtypes subsumed under the rubric of frontotemporal dementia. Neuroimaging protocols for dementia have also become more sophisticated. Specific ligands for PET imaging of amyloid plaques and neurofibrillary tangles have seen widespread use in research and may soon enter clinical use. It can be anticipated that these types of scans will eventually be used in court, in addition to the more traditional structural MRI and PET or SPECT scans visualizing metabolism which are often submitted as evidence today. Combining information from neuroimaging (whether structural, metabolic, or specific for pathological features such as amyloid) with genetic data has the potential to provide even more persuasive evidence for the presence of a dementia.

It must always be kept in mind that in nearly all legal contexts the mere fact of a diagnosis, even something as potentially devastating to functional abilities as dementia, is not sufficient to settle the legal question. As will be pointed out throughout this chapter, in most cases in which a psychiatrist is invited to participate, the legal system is primarily concerned with mental states and/or capabilities. While neuroimaging can be useful in supporting a diagnosis of dementia, the impact of that diagnosis on the subject, as plaintiff, defendant, testator, etc., must be established through behavioral data obtained by examining the subject (if living), obtaining collateral information from family members or other witnesses, and reviewing available medical records. Formal neuropsychological testing is also crucial in many cases of dementia involving a living subject. Ultimately, even after a diagnosis of dementia is made, the nexus between that diagnosis and the mental state or capacity at issue must still be shown, and neuroimaging is not a stand-alone tool for achieving this.

Death Penalty Mitigation

In those US jurisdictions which provide for capital punishment for murder, the proceedings are bifurcated into two phases. In the guilt phase of a capital trial, the jury reaches a verdict of Guilty or Not Guilty, just as in any other criminal case. If the verdict is Guilty, a second phase, the penalty phase, occurs. In this proceeding, the prosecution presents aggravating evidence for the purpose of persuading the jury that the defendant should be sentenced to death. The defense in turn presents mitigating evidence with the goal of convincing the jury to vote for a sentence of life imprisonment.

The rules of evidence in the penalty phase of a capital trial are significantly different from the rules governing evidence in other criminal proceedings. The defense is allowed to present evidence such as the defendant's history of being abused as a child, for example, or bring up his positive character traits, or point out his exemplary behavior since being incarcerated.

With increasing frequency, neuroimaging is being performed on capital defendants, and the results introduced at the penalty phase of the trial. Efforts by the prosecution to bar neuroimaging evidence, which are often successful in other criminal settings, are likely to fail when the defense seeks to introduce neuroimaging as mitigation evidence. Capital sentences have been overturned on appeal because the trial judge denied the defense's request to perform neuroimaging, as well as in cases where defense counsel failed to request neuroimaging.

When neuroimaging is offered as mitigation evidence in the penalty phase of a capital trial, it is likely to be allowed due to the wide latitude granted for most types of evidence to be admitted in that phase. Perhaps even more importantly, the jury is able to use such neuroimaging evidence for the general purpose of reducing the punishment from death to life in prison on the basis of any brain abnormalities. It is not necessary that a specific diagnosis be demonstrated. Similarly, no specific link between any abnormal brain structure or function and the crime for which the defendant has been convicted need be established. Neuroimaging evidence that the defendant's brain has some type of abnormality may simply add to the jury's list of mitigating factors. Under these circumstances, the likelihood of neuroimaging being useful to the defendant is higher than in most other criminal cases.

Given the lowered barrier to admission, and the relatively high potential for tilting the balance in favor of sparing the defendant's life, it is not surprising that neuroimaging has become very common in capital cases. In fact, failure to at least consider obtaining neuroimaging in a capital case on the part of the defense attorney could be considered grounds for an appeal based on a claim of ineffective assistance of counsel.

Recently a new approach to the use of neuroimaging in death penalty mitigation has emerged. Antisocial personality disorder (ASPD) and psychopathy have generally been considered aggravating factors. However, an argument can also be made that such conditions are mitigating in the context of a death penalty deliberation, in that the abnormal condition of the defendant's brain reduces his ability to recognize moral imperatives and therefore to exercise control over his behavior and to appreciate his culpability.

Functional MRI results suggesting a diagnosis of psychopathy have been introduced in the penalty phase as a mitigating factor. However, there is also a concern that in such cases, scientific evidence supporting the presence of ASPD or psychopathy could in fact be used by the prosecution as an aggravating factor, who would argue that the defendant represents a future danger and is not amenable to rehabilitation. These issues are discussed further in a subsequent section.

OBSTACLES TO THE USE OF NEUROIMAGING IN LEGAL PROCEEDINGS

Capturing Thoughts

Perhaps the single greatest obstacle that confronts the meaningful introduction of imaging data into legal proceedings is that the imaging for forensic purposes is almost invariably done after the fact and under essentially laboratory conditions. That is, a major limitation in the attempt to understand the relationship of brain activity to behavior is that there is generally no image of the brain that is antecedent to the behavior or activity that is at issue in the legal proceeding. Another obstacle is posed by the inability of any present technique to image a thought, to establish intent, or to establish the relationship between brain organization and function and any specific behavior or set of behaviors. While it might seem superficially attractive to ascribe a particular behavior to a particular brain organization (e.g., psychopathy), not all psychopaths commit crimes, nor has it been established that a particular brain organization facilitates the commission of a crime, any more than the fact that most murders in the United States are committed by people with dark hair establishes that hair color either predisposes to or is causal of criminal activity.

For ease of discussion, consider the introduction of imaging data in the legal context from two different perspectives: the introduction of structural imaging data and the use of functional imaging data. First and foremost, one must define the question that is to be answered by the introduction of imaging data. In the case of injury, there are at least two major questions: First, is there evidence of injury, and evidence of the consequences of injury? Second, does the imaging modality reflect the question at hand?

For example, if there is a history of trauma, is there evidence of trauma on the brain image? Was the timing of the imaging relevant to the event

itself? One may have had trauma with loss of consciousness, but an MRI scan taken several years later might show no evidence of a persistent structural alteration. Alternatively, an image taken several years later might show evidence of damage, but this in no way establishes the state of the individual's brain at the time of the alleged trauma. In other words, there could have been a pre-existing condition; without a scan before the trauma of record, one can never be certain that a pre-existing condition was not present. On the other hand, a CT scan performed shortly after alleged trauma could provide incontrovertible evidence of a recent bleed. While the scan may be helpful inferentially or interpretively, it is neither necessary nor sufficient to impute *behavioral* causality in the absence of other information.

By the same token, a DTI scan of a person's brain after trauma might show a disturbance in the connectivity of long fiber tracts, but in and of itself without correlative data one cannot establish that the disturbance in connectivity was due to the event in question. Furthermore, a disturbance in connectivity does not *prima facie* determine clinically significant behavioral compromise, which must be measured by other methods. It must always be remembered that structural imaging techniques image structure, not function, and functional relationships to structural abnormalities must be established by other means.

Large-scale, multi-year research efforts, such as the Human Brain Project in Europe and the National Institutes of Health's BRAIN (Brain Research through Advancing Innovative Neurotechnologies) Initiative, are now underway with the goal of bridging structure and function. The "connectome" and functional connectivity maps identify synchronous neural responses to stimuli. These are useful landmarks of neural pathways that are clearly associated with responses to specific stimuli, but in and of themselves are neither necessary nor sufficient to account for concomitant behaviors. Not all people take the superhighway; some use back roads and some use other modes of transportation, but when they arrive at their destination there is no way of distinguishing among them, unless the signal-to-noise ratio is markedly greater than the threshold of the methodology employed.

If the legal question is one of intent and relates to the state of a person's mind at the moment of a specific behavior that has occurred in the past, it is fairly obvious that there is no imaging technique that can elicit the historical state of mind. If the question is the likelihood of diminished capacity or responsibility, imaging can and has been used, but is subject to many caveats. There is ample evidence from many different functional imaging techniques that there are identifiable characteristics obtained by imaging groups of subjects that appear to be highly correlated with certain neurological and psychiatric conditions when the sample size is large and compared to groups that are well matched demographically. This is

particularly true in the dementias and even in “functional” disorders such as schizophrenia and major depression and perhaps even the world of behavioral types and disorders such as psychopathy.

The issues that arise in the courtroom have to do with the relevance of group data with high interindividual variance when the values are applied to the individual who shows a difference from the group norm. In general, imaging data is of greatest utility when it is used to show consistency with or support for another form of evaluation rather than being thought of as pathognomonic in and of itself.

In the landmark New York City case of People v. Weinstein, the question before the court was the admissibility and relevance of a fluorodeoxyglucose (FDG) PET scan to the behavior of the defendant, who had pushed his wife out of their apartment window to her death in the course of a domestic quarrel. He was arrested and charged with murder. A routine pretrial CT scan revealed the presence of a subarachnoid cyst that was pressing and distorting Mr. Weinstein’s frontal cortex. A subsequently performed FDG PET scan unsurprisingly revealed that this functional image was indeed grossly abnormal by any standard, since this very large cyst occupied a significant amount of the calvarium. By history, the cyst was likely about 40 years old and yet Mr. Weinstein lived an unremarkable professional and social adult life, maintaining a white-collar position and never exhibiting any symptoms that warranted an imaging procedure and with a total lack of awareness that his brain was, structurally speaking, highly abnormal.

The defense contended that this brain abnormality led to diminished ability to inhibit an aggressive impulse, yet there had been no evidence of diminished ability to inhibit impulses that was apparent over many years that his brain was structurally deformed. After the evidentiary hearing, the Court allowed the introduction of the PET scan into evidence. Prior to trial, Mr. Weinstein accepted a plea agreement and pled guilty to a reduced charge of manslaughter.

Not having a reference scan of the subject before an accident or illness or before a specific behavior is manifest limits the value of the imaging modality as proof in and of itself of a change in state. Nevertheless, there is increasing introduction of imaging data in the courtroom. While experts can argue about the technical relevance, most would agree that an image is so compelling and can be so persuasive despite its technical limitations, that it is rarely excluded from at least the mitigation phase of a trial (see prior section). When the issue is one of possible diminished capacity, either due to diminished response to a behavioral or cognitive challenge as shown by either fMRI or PET scanning, functional imaging has had its impact in the courts, and this will likely continue. When the issue is one of altered behavior or diminished cognitive functioning subsequent to traumatic brain injury, then techniques such as DTI or altered tissue properties such

as shown in structural MRI have been compelling ancillary information, even if they do not directly measure the properties at issue.

Whether this type of functional imaging *per se* should be used as primary evidence of diminished capacity in such matters is the subject of considerable debate and dispute, since it is obvious that there are people with aberrant brain images that are discovered incidentally since their behavior is otherwise unremarkable. On the other hand, there are people with aberrant behavior who have normal scans and people with aberrant behavior who have abnormal scans dependent upon the technique that is used. Structural imaging is often the least contributory when the issue before the court is of a specific behavior, although it can be useful and relevant in support of alterations of brain parenchyma that correlate temporally with a documented brain injury and persistent change of mental state.

The explanation for this lies in the enormous plasticity of the human brain and its ability to function normally despite innumerable interindividual variations. A most extreme example of this was noted years ago when a PET scan of a normal British volunteer, a university graduate and mother, whose PET scan was remarkable for a thin cortical rim of gray matter, with little discernible subcortical metabolism and whose visual cortex was in the front of her brain rather than the occiput. She had been a volunteer "normal" subject and had no antecedent psychiatric, neurological, or behavioral concerns.

A more common occurrence is illustrated by the case of *U.S. v. Gigante*. Mr. Gigante had two separate SPECT scans that were read as abnormal and consistent with schizophrenia or dementia, a CT scan read as abnormal and consistent with dementia, a PET scan read as abnormal and consistent with Alzheimer dementia, yet he later admitted to feigning in the service of evading justice. Whether or not any of his scans were in fact abnormal should be considered in the light of whether a significant deviation from a group mean is a sufficient indicator of abnormality. Five years later, Mr. Gigante was clearly no longer demented and remains perhaps the only documented case of an Alzheimer dementia diagnosed with a medical certainty who was cured by incarceration.

The case of *Illinois v. Dugan* is an example of the medicalization of a psychiatric disorder in the service of death sentence mitigation for a convicted child murderer. The defense argued that the fMRI images of the defendant, obtained 2 decades after the crime, showed a pattern that was clearly consistent with that of a large group of incarcerated psychopaths. The defense argued that on the basis of the similarity of the defendant's scan to that of the scans of other psychopaths, Mr. Dugan must have had "diminished capacity," since psychopathic personality meets a set of diagnostic criteria. Thus, the defense contended, he is (perhaps was) biologically not normal, and should not be subject to the death penalty

under Illinois state law.

The relevance of imaging findings to the legal questions is of course highly dependent upon the central question in any individual case. This generally ranges from the determination of direct brain damage that is generally established by structural imaging techniques, to the relationship of a pattern of brain organization or activity to a particular behavior such as a crime, to the attribution of the persistent change in posttraumatic behavior that can be attributable to an "abnormal" brain image. In particular, imaging data obtained by the functional imaging techniques are generally highly dependent upon medication status, hormonal influences, and mental state at the time of the scan, which are likely to be far different from the conditions relevant to the central legal questions.

Defining Normality

Images obtained months or years after the behavior at issue, and their analysis and interpretation that are presented to the court, are related by inference to a hypothetical "normal" prior scan that the subject presumably would have had, in order that any change in mental status, mood regulation, impulse control, activities of daily living, etc. can be attributed to those brain functions or structures that could be related to the behavioral presentation. For example, how does one establish for a given individual whether or not his ability to play sports has been compromised on the basis of an image? How does one establish on the basis of an image whether or not a claim of a persistent mood change was due to a specific antecedent injury? In addition to the issues of comparison of the subject to a reference group, the expert is faced with the fact that rarely does one have access to an image of that individual patient taken under the same conditions prior to the legal matter at hand.

The scientific obligation would be to establish that the change in the image in fact is causally related to the changed behavior. The legal standard sets a much lower bar. In general, these concerns are often obviated in the face of indisputable structural damage in large regions of the brain known to subserve a particular behavior or set of behaviors or by the inference that something unusual in an image should be equated with a functional medical or psychological limitation. Much of this reasoning comes from lesion studies that by their very nature often failed to generate or identify compensatory mechanisms and pathways, both functional and physical.

The examples above give an indication of the complexity of the concept of normality. Normality is not merely a statistical term associated with a defined mathematical distribution, but is complicated by the realization that any such definition must be context dependent. It may be normal to be of diminutive stature in a population of pygmies; it may be normal to be very aggressive in an athletic competition; it may be normal to be unusually quiet while sitting in an auditorium listening to a lecture; it may be normal

to restrict entry to one's own home yet highly abnormal to act in a similar fashion when the property is not one's own. Clearly, the concept of normality has to be considered in a much more complex set of domains than is typically associated with the interpretation of a brain image in a forensic situation, where reductionist tendencies would lead one to consider only that the events before the court determine or are determined by the imaging data.

"Unusual" is often equated with "abnormal," and abnormal, particularly in the context of regional decreases in metabolic activity as shown either with PET or fMRI, is generally equated with dysfunctional. This equation is incorrect but also widely used. It has never been shown that there is an energy consumption requirement that differs between an efficient thought process and an inefficient one. There is no evidence in the literature, for example, that it takes more energy to think a bad thought than a good thought.

Perhaps the most common characteristic of functional brain imaging is intersubject variability. When 20 medical students were asked to perform the same simple verbal fluency task, the individual BOLD responses varied from virtually no response to very large regional changes. There is no generally acceptable interpretation for this variation. Is it a function of differing amounts of background noise, or is it related to the efficiency of a given brain, or to a whole host of ever more complex neurochemical relationships that are as yet poorly understood? In studies with genetically bred rats reared together and even scanned together, there is abundant evidence from many laboratories that there is significant metabolic variation both in the resting state and in terms of stimulus response that belies the simple assumption that identical twins have identical brains, even at the level of a rodent. When studying an individual human being, with or without stimuli, there is an unusually large variability of metabolic activity over the course of a single 24-hour period.

This makes the concept of normality surprisingly complex. Indeed, scientists might study a population defined by other terms, such as psychopathic behavior that has led to incarceration, and find that their brains have an organization that may be strikingly similar. However, there is, as yet, no evidence that the brains of psychopaths who have not been incarcerated or who have not behaved in a criminal fashion are organized differently. There is an almost irresistible tendency for people to be reductionist and attempt to use brain imaging as the modern equivalent of 19th-century phrenology in which a particular organization or observation leads to or explains or facilitates a complex set of behaviors. At the present time, while there indeed is a difference of opinion, there is little evidence to support the notion that any of the current imaging technology has as yet produced a pathognomonic signature of a particular thought, feeling, or behavior.

What is normal in one context is abnormal in another. It is often forgotten that even in the mathematically defined normal (Gaussian) distribution, a measured value two standard deviations from the mean does not mean that the value is necessarily abnormal. To equate abnormality with dysfunction is a distortion of the definition of the terms, yet is widely used without clarification or qualification in the psychiatric literature. For example, it is well known that some antipsychotic medications will cause an increase in regional brain metabolism, while others will cause a decrease in regional brain metabolism in the very same brain regions. In both clinical situations, the patient's psychotic symptoms improved. It has also been known for the past 2 decades that a particular brain stimulus will cause regional changes in the BOLD effect measured by fMRI that are increased in some regions and decreased in others. While the decrements are often ignored when images are displayed to illustrate activation associated with a particular pathology, it is clear that these decrements have physiological meaning; indeed, they have been interpreted and placed into the context of the "default mode" of the brain.

The images of a given subject used in the courtroom are generally evaluated in comparison with group normalized images. It is hardly surprising, therefore, that the images from any individual will differ from a group normalized image. If, for example, one were to take 12 jurors in the jury box, one could indeed normalize (i.e., average together) their faces to produce a single composite image from this group of (presumably) normal people. Yet no single individual's image will conform to that normalized image. If one were to use a different dimension and measure height, the individual heights of the jurors might vary from 60 to 80 in. Not a single individual might have a height equal to the mean, and for none of them would either measure reflect their ability to function as a juror, their marital status or their economic standing.

The expert, judge or jury can ask whether or not the image is normal, with normality being defined as being without obvious pathology. The unusual case of the college graduate whose visual cortex was where the frontal cortex is usually located noted above illustrates but one difficulty with defining normality. To be normal, must the brain be normal in all dimensions, that is, be functionally isotropic? If not, is there some dimension that is more relevant to the case at hand than another? Was she anatomically normal? Was she behaviorally normal? Was she functionally normal, both physiologically and psychologically, etc.? Most important is whether the imaging methodology and scan conditions allow the accurate measure of the property that is most relevant to the legal issue. Hence the distinction between what is measured, the information gleaned from the analysis of the imaging data, the inferences that are drawn from that analysis, and the comparison with either a "normal" reference or a comparison state are crucial to presenting a bona fide scientific picture to

the court. It is easy to present pseudoscience. Without knowing how they were acquired, analyzed, simplified, and interpreted, it is well known that images convey such an enormous amount of data in such a concise fashion, that even of poor quality, they will be at least as apparently meritorious to the unsophisticated as the most elegantly and rigorously acquired and presented data.

In a forensic context, the concept of normality can be very complex. Should normality be predicated on properties of a group of people chosen at random? Should it be a group chosen for the absence of psychopathology, or ethnic similarity, or education, or political beliefs, age, etc.? How many demographic variables should be included to assure an adequate comparison with the index subject? Should subjects be tranquilized, should they be exposed to ambient stimuli? Should their brains be constrained by a task, or left unconstrained in the so-called resting state? Should the eyes be open or closed? Is the brain function to be imaged so sensitive that a minor change in state will alter the image, or so stable and insensitive that the signature remains invariant over time? Clearly brains come in different shapes, sizes, and ages, and have had different life experiences and reflect genetic and epigenetic diversity. These and similar considerations make the task of establishing an appropriate reference group a formidable undertaking. Nevertheless, images are generally proffered in the courtroom in support of one contention or another and of course will show that virtually any subject will differ physiologically from the group norms. In large measure this is accomplished by means of scaling the images to either magnify differences in a graphical display; by using arrows to delineate regions of difference between the index subject and the reference image(s) even if the data have low statistical significance or provide little physiological evidence for the forensic contention; and most impressively the use of visual contrast that can make small differences seem huge.

A common example of this is scaling the data of a set of images in such a manner that the regional differences are artificially embellished. For example, if the putatively important region of interest had an arbitrary average rating of 83 units and every region above 80 units was colored red and everything below 80 units black, then an area might appear to be physiologically dead that is only a few percent lower than the area that is physiologically active. While there is nothing illegal about presenting data in this manner, it certainly can be misleading. The differential sensitivity of the eye to colors of different wavelengths is well known, and is often exploited in the presentation of marginal data in the court room.

The problem remains that any individual image is analyzed with reference either to the same brain under different conditions, or to a group of reference brains whose signals have been summed, smoothed, and normalized. In the forensic situation, it is virtually impossible to acquire

the scan that is an appropriate reference to the behavior at issue both because what is measured are electrical signals, not behaviors, and because it is impossible to know the antecedent state of the brain before the matter at hand has transpired. This results in imaging data that has to be evaluated in probabilistic terms, that is, it must be addressed on a statistical basis and virtually never leads to a unique solution.

Although the solution is not necessarily unique, it does not mean that imaging data is of no utility in the forensic context. Advances in technology, and advances in the analysis of brain images in conjunction with powerful new statistical techniques have indeed brought spectacular insights into brain functioning under a variety of conditions that were heretofore never available. From early region-of-interest analyses to the leap afforded by statistical parametric mapping based on the theory of Gaussian fields, to pathway analysis in time-locked connectivity measures and many other analysis strategies that are constantly being developed, explored, and validated, there is every reason to believe that brain imaging is here to stay in the forensic setting and will become increasingly introduced as evidence. Whether this will be as compelling to the cognoscenti as it is for the layman has yet to be settled.

Nevertheless, whether warranted or not, the pressure to reduce the relationships among mind, brain loci and pathways and complex behaviors to a unique interpretable pattern is both in vogue and clearly attractive. At the present time, it is critical not to turn science into phrenology. Brain imaging should not be granted the status of an unalterable truth nor consigned to the garbage heap of useless pseudoscience, for it is neither. It must be remembered at all times that these techniques are measuring various physical properties of the brain with detectors outside of the brain, using, of necessity, many simplifying assumptions, both mathematical and physiological. The various modalities employed have different sensitivities, specificities, temporal and spatial resolutions and greater or lesser proximity to the behavioral question that is generally before the court. Each measurement has its own set of errors and uncertainties. To this must be added the fact that every calculation in the chain of processing leading from the behavior and its underlying physiology to the final image that is displayed is predicated upon multiple, diverse simplifying assumptions and inferential conclusions. At the time of this writing, brain imaging is still in a state of rapid development after less than half a century of development. A high degree of caution should accompany any claims that brain imaging alone will explain behaviors or that for any behavior there must exist a characteristic brain image.

The "Christmas Tree Effect"

The preceding discussion highlighted some of the key challenges to the reliable and relevant application of neuroimaging data to legal questions.

Many of these obstacles can only be appreciated after detailed discussion and explanation, and involve highly technical concepts. A number of observers have expressed concern that judges and juries may be unable to keep these caveats in mind, and will be swayed by the seemingly infallible brain scan, even when the claimed results of that scan are unwarranted, or are accurate but irrelevant to the legal question at hand. This leads in turn to the argument that the presentation of images obtained with expensive, high-technology scanners is in many cases likely to be more prejudicial than probative. In other words, the apparent scientific proof of abnormalities in the brain obtained with neuroimaging will be given excessive, unfair weight by the judge or jury.

The concern that judges and especially juries would have difficulty identifying the proper boundaries of interpretation of neuroimaging data has led some authors to argue for a restrictive approach to admitting brain images. The idea that the sophisticated science used to create such images would cause juries to ascribe undue weight to those images is sometimes referred to as the "Christmas Tree Effect." Related to this issue is the concern that in the 21st century, juries harbor the expectation that parties to a legal action should be able to present precise results obtained using advanced scientific techniques—and if no such data is forthcoming, they will automatically consider the case to be weak. This concern has been referred to as the "CSI Effect," after several popular television series which follow the exploits of fictional crime scene investigators who solve crimes by processing the evidence criminals leave behind, including, of course DNA, but also evidence such as soil or plant material which can be tied to a specific location, flecks of paint from an automobile that can be linked to a specific model and year, and so on.

If the fear of the Christmas Tree Effect and/or the CSI Effect is justified, one would expect, for example, that a jury presented with an image of a defendant's brain lighting up with abnormalities would conclude—on that basis—that the defendant must have had diminished responsibility at the time of the crime. In another scenario, the jury would view images of a plaintiff's brain in a personal injury lawsuit and conclude that the presence of abnormalities indicated with bright colors on a scan means that the action or negligence of the defendant is the cause of those abnormalities, and that the plaintiff is deserving of compensation, when in fact the plaintiff may not have any functional deficits, and/or any impairments (as well as the imaging changes themselves) could be entirely unrelated to any actions or omissions by the defendant in the suit.

Although often cited as an objection to the use of neuroimaging data in court, just how much prejudicial influence can be ascribed to the Christmas Tree Effect is still a matter of debate. Studies of mock jurors have had mixed results, and it appears that earlier studies suggesting that brain imaging was prejudicial may have been overstated. Later studies have

reported that their subjects were able to evaluate the relevance of brain imaging data to the legal question at hand, and did not simply vote in favor of the side that presented that data. Actual systematic polling data from real jurors who have had to confront neuroimaging evidence could shed more light on the question, but so far it does not seem that such a study has been reported in the literature.

In any event, the risk that the prejudicial effect of neuroimaging will outweigh its probative value cannot be dismissed. It necessarily falls to judges to act as gatekeepers to minimize this risk. Given the preceding discussion, it is possible to identify some factors that judges should consider when assessing whether to admit neuroimaging evidence. Neuroimaging results that are offered in isolation, without any correlated behavioral or clinical observations, should obviously be met with extreme skepticism. Scans that are performed without some basis for suspecting neuropsychiatric abnormalities risk earning the pejorative label of a "fishing expedition." On the other hand, scans that were performed and interpreted in the ordinary course of a diagnostic workup, prior to the beginning of litigation, are likely to be viewed favorably, given that they were ordered with a medical and not a legal purpose in mind and are thus less likely to be subject to biased interpretation.

More generally, neuroimaging studies for which there is a solid scientific foundation supporting their usefulness for the condition in question (e.g., moderate to severe TBI, dementia), and where the relevance of the results to the legal issue is clear, should not be barred out of concern that the mere presentation of colorful images will somehow interfere with the jury's ability to weigh the arguments in the case. On the other hand, when imaging evidence is offered to support controversial or dubious claims, allowing that evidence to be presented may create a risk of improperly swaying the jury, given the aura of hard science associated with brain imaging. One such controversial area, the use of fMRI as a means of identifying deception, will be discussed later. Other assertions that should be approached with great caution include, for example, the use of PET or SPECT as evidence of posttraumatic stress disorder (PTSD), the use of DTI to diagnose mild TBI, speculating about intent, planning, etc., on the basis of brain imaging, and the claim that imaging results, in and of themselves, can identify the cause of a particular brain abnormality, and not just its existence.

The Person versus the Brain

The rapid advances in neuroimaging over the past quarter-century have been integral to a profound debate taking place at the intersection of philosophy, law, and neuroscience. At one extreme of this debate are those who argue from a reductionist point of view, asserting that advances in the neuroscientific understanding of the brain demonstrate that free will is an

illusion. Given such a conclusion, some reductionists such as Joshua Greene and Jonathan Cohen have called for a complete overhaul of the criminal justice system, including purging from it the concept of punishment and focusing instead only on public safety and deterrence.

Even stopping well short of the extreme reductionist position, it is possible to envision neuroimaging being used to argue, for example, that a particular criminal defendant lacked the capacity to conform his conduct to the requirements of the law, due to abnormalities in his brain's structure and/or function. Such abnormalities would have been invisible in the pre-neuroimaging era.

It seems reasonable to assume that the percentage of accused criminals —particularly those accused of committing violent acts—in whom some type of neuroimaging abnormality can be detected is not negligible, and may be substantial. Does this mean that in the future, numerous criminal defendants will be able to successfully argue that they are not criminally responsible (or perhaps should be convicted of lesser charges), on the basis of neuroimaging findings?

Arguing against the reductionists, as well as those who predict the widespread use of neuroimaging in law, are other observers who have suggested that the impact of neuroimaging in many aspects of the justice system is and will continue to be minimal. The most prominent of these authors is Stephen Morse, who coined the term "Brain Overclaim Syndrome" to describe what he views as excessive and unjustified claims of relevance for neuroimaging data applied to legal questions.

A number of methodological, technical, and conceptual barriers to the wider use of neuroimaging in court have been discussed earlier. In addition, there are other causes for concern that will not necessarily be resolved by more accurate, standardized, and reliable scanning technologies.

One concern is the question of what information of value to the judge's or jury's deliberations is added by the neuroimaging findings. This is especially an issue when past mental states are in question. For example, if a defendant pleading Not Guilty by Reason of Insanity were to be shown, by conventional fact and expert witness testimony, to have a diagnosis of schizophrenia, to have been psychotic at the time of the crime, and as a result of his psychotic symptoms to have met the standard for legal insanity, a PET scan performed months or years later, revealing abnormalities consistent with the abnormalities seen in patients with schizophrenia would perhaps be corroborative, but not particularly useful, nor necessary to the jury's deliberation. In fact the PET scan could be viewed as redundant—as cumulative evidence, in the legal terminology.

The other side of this coin is the fact that in cases where the evidence from behavioral and/or clinical observation is more ambiguous, neuroimaging is less likely to produce clear findings. For example, one

would not expect someone who performs relatively well on neuropsychological testing and exhibits minimal cognitive impairment to have a functional imaging scan consistent with severe dementia. If the imaging did suggest such a diagnosis, one would of necessity question the scan results, not the subject's abilities. To argue that a highly functional person has severe dementia because their brain scan says so would be untenable. Behavior, not imaging, equals diagnosis.

One of the essential premises of neuroimaging is that it identifies the physiological correlates of mental functions. The observed neuroimaging findings are expected to correspond with the subject's observed behavior. But this means that in situations where behavior is grossly abnormal, as in psychosis or dementia, neuroimaging will often not be necessary for legal purposes, being merely cumulative, while in situations where behavior is more ambiguous, and subject to varying interpretations or explanations, the neuroimaging results are also likely to be ambiguous, and therefore likely to yield little information useful in assisting judge or jury in reaching legal conclusions.

A related, but distinct and more fundamental objection to the utility of neuroimaging in court raised by Morse and others involves the link or nexus between the physiology of the brain, in terms of what can be observed through imaging, on one hand, and the current and especially the past mental states, behaviors, and capacities of the person undergoing that imaging, on the other. It is the latter which are the elements that the legal process must analyze and understand. However, even the most sophisticated functional neuroimaging of today is far from being able to reliably identify and characterize discrete human cognitions and emotions. For many purposes, neuroimaging may never be able to shed light on the critical legal question at issue.

For example, imagine the assertion that observed neuroimaging abnormalities in a defendant's frontal lobe are indicative of an inability to resist impulses, and therefore that the defendant is deserving of mitigation or exculpation. Unless such a defendant is so impaired that he or she is giving in to impulses on a near-constant basis (an extremely rare scenario), there will necessarily be some impulses that they are able to resist. A neuroimaging scan cannot reveal anything about whether the defendant had the ability to resist a specific impulse at a particular time in the past, that is, when a crime was committed by that defendant. Nor, conversely, does the absence of particular neuroimaging findings prove that a defendant's brain was "normal" in its ability to control certain impulses. Yet, in a diminished capacity defense, the difference between a less and a more serious charge and therefore a difference potentially measured in years of prison time, hinges on exactly this question of the capacity to control one's impulses at the time of the crime.

Although neuroimaging cannot shed any light on this question, some

attorneys and experts have and will continue to overstate the benefits of neuroimaging, asserting, for example, that a defendant with a certain imaging profile was incapable of planning. Because of the aura of scientific precision associated with MRI, PET, and so on, there is a risk of judges and juries giving too much credence to such claims. However, the extent of that risk is difficult to quantify, and probably depends on numerous factors specific to a given case.

A number of technical problems constrain the usefulness of neuroimaging in the legal arena. Some of these may be overcome by improvements in the technologies. In addition to these issues are the realities of the questions that are important in the forensic context. Some are permanently inaccessible to neuroimaging. These include determining a person's mental state at a past time point or trying to separate "could not" resist an impulse from "did not." Despite such barriers, clearly any type of blanket ban on neuroimaging evidence would be unwise. Rather, as has been stated, it is critical that, when neuroimaging is offered as evidence, well-trained, unbiased experts explain why neuroimaging is or is not useful for the particular legal question at issue.

CURRENT AND FUTURE CONTROVERSIES IN FORENSIC APPLICATIONS OF NEUROIMAGING

The tremendous expansion of scientific and technical knowledge has created methods for the noninvasive visualizing of brain structure and function that were previously inconceivable. These incredible advances have revolutionized medical fields such as neurology and neurosurgery, and have also been, along with the great strides in the understanding of and ability to manipulate DNA, one of the main driving forces behind the growth of the now-huge field of scientific study known as neuroscience.

The applications of neuroimaging to legal questions that have been discussed thus far have been predicated on those techniques being used to identify abnormalities in structure and/or function associated with a particular diagnosis. Such abnormalities range from the indisputable, as in those resulting from severe traumatic brain injury or advanced Alzheimer disease, to the more contentious, such as abnormalities associated with schizophrenia, bipolar disorder, or milder brain insults such as concussion. Other diagnostic entities, such as PTSD and ADHD, have also been suggested by some to have particular neuroimaging profiles, reliable enough for courtroom use, though the state of the science here is even less settled. However, with all of these conditions, most mental health professionals would be open to the argument that there is at least the potential that a brain imaging technique of sufficient precision to consistently find a "signature" in individual subjects could someday be developed.

More controversial still is the application of neuroimaging for

psychiatric conditions for which the underlying neural changes would seem perhaps even more challenging to identify. Personality disorders by definition do not involve a break from reality, nor do they typically produce the neurovegetative symptoms seen in mood disorders or the autonomic disturbances in anxiety disorders such as PTSD. Do they nevertheless create changes in brain structure and functions that are consistent enough across patients to be reliably detected and categorized with neuroimaging? And what about abnormal sexual preferences, such as the paraphilias? Because of their obvious importance in the legal system, ASPD and psychopathy, as well as pedophilia, are extremely tempting targets for forensic neuroimaging, and have been suggested as conditions which can legitimately be examined at the single-subject level to assist in legal determinations.

Personality disorders and pedophilia are still in the realm of identifying a pathological condition, broadly defined. But diagnosis is not the end of the forensic neuroimaging story. Other applications of neuroimaging have been suggested and even, in a few cases, submitted to courts, that do not involve psychiatric or neurological diagnosis, and are not intended to detect any physiological abnormality or impairment. In theory, techniques such as fMRI lie detection could be used with virtually anyone who can be placed in the machine. It is with these types of methods that one begins to approach a crude form of the technological "mind-reading" that until recently was only discussed in works of science fiction. As will be discussed, the employment of such methodologies for legal purposes is currently in a truly embryonic stage, and such use may in fact never establish a foothold within the legal system.

Antisocial Personality Disorder

ASPD and the related but distinct diagnostic entity known as psychopathy are obviously of great interest to the legal system, primarily although not necessarily exclusively in the criminal context. Neuroimaging has been used as a research tool in the study of ASPD and psychopathy, as well as criminal and especially violent behavior in general, since the advent of modern techniques such as SPECT, PET, and MRI. A rich literature has developed which has identified structural and functional abnormalities in numerous brain regions, including cortical regions such as the prefrontal cortex and cingulate gyrus, and subcortical structures such as the amygdala, in ASPD and psychopathy.

Most of the literature on neural correlates of ASPD, psychopathy, and criminal behavior involves group averages compared to groups of normal controls. ASPD is a heterogeneous disorder, and even if one's focus is restricted to psychopathy, significant variation in phenotype is observed. It is not clear how soon the science in the area will be up to the task of reliably diagnosing psychopathy or ASPD—or significant antisocial personality

traits—in an individual subject, and the related, equally important, but distinct task of ruling out such a diagnosis.

Although the science is not currently reliable enough for general use as a diagnostic tool, a judge has allowed neuroimaging evidence of psychopathy to be presented as a mitigating factor in the penalty phase of a capital trial (the Dugan case), a proceeding where, as discussed in a preceding section, evidentiary standards are relaxed compared to other criminal settings and where the scrutiny to which offered evidence is subjected to is correspondingly reduced. (The Dugan jury sentenced the defendant, a serial murderer, to death.)

Even confining the discussion to death penalty mitigation, however, significant questions regarding the merits of neuroimaging for ASPD and psychopathy remain. Assume for the sake of argument that the science has become sufficiently reliable and accurate for use in diagnosing an individual. It can also be assumed that, unlike in the case of a defendant with schizophrenia seeking to utilize an insanity defense, neuroimaging evidence supporting a diagnosis of ASPD constitutes information that is not merely redundant and cumulative, but actually informs the jury regarding the clinical situation. But what will the jury do with this information?

An argument can be made that a jury would not necessarily view evidence of the personality disorder diagnosis as lessening the defendant's culpability or their appropriateness for the death penalty. A conclusion that the defendant is less culpable would presumably arise due to the diagnosis' implications of impairments in higher-order moral decision making and/or volitional control. However, there seems to be a significant probability that a capital jury would instead view evidence of ASPD or psychopathy as an aggravating factor, given that people with such conditions are difficult to treat and rehabilitate, with corresponding difficulty in reducing their risk of future violence.

This refractoriness to treatment, and ongoing dangerousness, could certainly be raised by the prosecution in their presentation of the aggravating factors supporting a death sentence. Thus it is not clear that neuroimaging, even as it adds a "biological" dimension, admittedly useful in countering claims that ASPD is merely psychological or a "character flaw," would actually be helpful to the capital defendant's effort to avoid the death penalty and receive instead life in prison without possibility of parole.

If one were to expand the discussion to noncapital cases, it is even less clear how neuroimaging evidence of ASPD or psychopathy would benefit a criminal defendant with the condition. It seems highly unlikely that defendants will be able to argue that their personality disorder constitutes a defense or an excuse, regardless of whether that disorder is "hard-wired" in their brain as demonstrated by modern neuroimaging techniques.

Such a result would seem possible only if one of two fundamental changes were to occur: The first would be a revision of the system allowing defendants who could show a biological predisposition to commit crimes, such as ASPD, to be diverted from criminal punishment and placed in therapeutic settings. One can imagine that countless "ordinary" defendants, their attorneys and loved ones might find such an arrangement highly unjust. An even more sweeping change to the criminal justice system, wherein the entire concept of punishment were abolished and replaced with treatment and confinement only for purposes of protection, would also allow defendants with ASPD and psychopathy, along with all other accused persons, to avoid punishment *per se*. Neuroimaging evidence could be quite useful in this situation to help in classifying patients and guiding treatment. As touched on earlier, some reductionists have drawn on the evidence of neuroscience to call for just such an overhaul of the criminal justice system. However, to say that the probability that this will occur anytime in the next few decades (at least) is remote would be an understatement. The concept is mentioned again here only as a reminder of the extent to which the neuroimaging revolution has inspired some academic observers to contemplate radical possibilities.

Under the current system, neuroimaging of ASPD and psychopathy in a forensic context represents a double-edged sword for the defense. If there is nothing else to fall back on, a defendant facing a possible death sentence may choose to introduce it and hope that it does not tip the scales against him. But in many other circumstances, defendants will not want their brain searched for signs of antisocial tendencies. In fact, it is conceivable that in the future, defendants will seek to *prevent* being subjected to imaging studies designed to determine if they are a psychopath, or to block the judge or jury from considering such evidence if it already exists, given the possibility that it will be viewed as evidence of their dangerousness or lack of rehabilitation potential. It is difficult to imagine evidence of psychopathy being beneficial at a parole hearing, for example.

Using Neuroimaging in Risk Assessment

Perhaps more realistic than a sea change in the criminal justice system eliminating consideration of punishment is the potential for using neuroimaging as a tool for assessing future dangerousness. In addition to ASPD and psychopathy, a database of neuroimaging findings in pedophilia has been accumulating. Progress in this area is complicated both by diagnostic heterogeneity and by extensive comorbidity with other disorders such as mood and anxiety disorders and personality disorders. Nevertheless, by comparing the activation patterns of nonpedophiles exposed to sexual stimuli with those of pedophilic subjects, it may be possible in the future to identify the presence of pedophilic interests with reasonable accuracy. Some authors working in this area have suggested a

potential role for neuroimaging in assessing response to treatment. In addition to using neuroimaging for specific disorders associated with elevated risk, such as psychopathy and pedophilia, advances in the general study of the neural correlates of violent behavior, particularly of the impulsive, emotional variety, point to at least the potential for a wider utility for neuroimaging in dangerousness assessments.

Much of the research on the neuroimaging of pedophilia, as well as discussions on the use of neuroimaging to assess risk of violence, originate in European countries, particularly Germany. This may in part be due to differences between the American and Continental legal systems: Many European countries use an inquisitorial rather than an adversarial system. Expert witnesses are routinely retained by the court, and much less often by the prosecution or defense. Presumably, the risk of a biased expert, including the risk of a "hired gun" essentially selling his or her opinion, is lower when experts report directly to the judge. This may lead to increased confidence that neuroimaging findings will not be used in a prejudicial manner by experts stretching the data beyond accepted scientific applications. Some European authors contemplate the future use of neuroimaging on a routine basis to augment actuarial, clinical, and other risk assessment tools for purposes such as sentencing decisions and release determinations for prison inmates and forensic hospital patients.

In this hypothetical future, neuroimaging would not be performed only once for each patient, but would be used at intervals to assist in tracking progress and improvement in response to therapeutic interventions. This would allow assessment and treatment to be individualized over and above current approaches and could—again, at least potentially—improve the objectivity and accuracy of risk assessments. In theory, such a use of neuroimaging would also assist in determining when a particular treatment approach has been unsuccessful. This would be of obvious benefit, as incarcerated and hospitalized patients with ASPD, psychopathy, and sexual disorders are highly motivated to tell their treatment providers what they believe will help them regain their freedom. It is for this reason that techniques such as the polygraph and the penile plethysmograph are frequently employed in the treatment of sex offenders despite their limited utility. Currently, there are no studies demonstrating the value of neuroimaging as a risk assessment tool, and it is not clear when, or if, neuroimaging will be sufficiently accurate and reliable for such an application.

Lie Detection, Recognition Memory, and Imaging-Assisted Interrogation

The first research studies investigating the neural correlates of deception and lying were published in 2001. These were group fMRI studies, but within less than half a decade, studies reporting the successful detection of deception in individual subjects appeared in the literature. Soon,

commercial enterprises were launched, offering clients the opportunity to undergo an fMRI scan, during which they would be asked a series of questions. The fMRI results would be analyzed to establish whether they were telling the truth or not. These companies reported accuracy in detecting deception on the order of 90 percent.

The exact mechanisms by which lying might produce a specific neural activation pattern are unknown. Researchers in this area have hypothesized that the act of lying increases the workload in areas of the frontal cortex, as the truthful response must be inhibited and a plausible lie generated. In any event, given the apparently high accuracy of the early studies, a few attempts have been made to introduce fMRI lie detection as evidence in legal proceedings. In the evidentiary hearings that have been held to date, the method has been ruled to not have demonstrated sufficient scientific validity to be entered into evidence. Subsequent imaging studies raised concerns as well, such as one which showed a drop in the accuracy of identification to 33 percent when subjects were instructed to covertly move a finger during the telling of lies.

In addition to these types of practical, methodological issues relating to fMRI lie detection techniques there are a number of broader concerns. One fundamental issue is the lack of ecological validity involved in the test paradigms. In most cases, the subjects engage in deception only when they are cued to do so in the protocol. This is not a very close approximation of what takes place in a police interrogation or on the witness stand, where the subject can choose when and whether to lie in response to every question. Perhaps more importantly, it is essentially impossible, in a research context, to create the emotional conditions that are present when a person is the focus of a criminal investigation, or is being deposed in a lawsuit with large sums of money at stake.

Another fundamental problem is the question of whether engaging in deception actually does—in the real world, as opposed to the carefully controlled conditions of a scientific experiment—generate a unique neural signature. Does the equivalent of a "grandmother neuron" (or circuit of neurons) for lying exist? Or do acts of deception exist on a continuum? How should behaviors such as omissions, half-truths, or making statements that are literally true if a question is narrowly interpreted, but where the respondent clearly knows that the questioner meant something more, be classified? Will it ever be possible to tease apart the fine distinctions among such nuanced human behaviors using a tool as (relatively) crude as neuroimaging?

Another major concern that has been raised by several commentators is the concern that false-positive results for lying might occur if the technique does not detect lying per se, but simply the suppression of a competing response. If an accusatory question provokes a fearful or angry emotion in the subject, could inhibiting that emotion create the appearance of

deception in the imaging data?

There are many other unknowns in the field of imaging-based lie detection. Virtually all research studies have involved healthy young adults as volunteers. There is essentially no data on populations such as juveniles, the elderly, or people with a history of head injury, substance abuse, or other psychiatric disorders. Would a delusional person discussing their delusional beliefs not register as engaging in deception?

Similarly, there are no data on how highly practiced, successful liars would perform. This population would include criminals with ASPD, of course, but what about people who are skilled at forms of "lying" not due to a psychological aberration but due to training, for example, spies or undercover police officers? Would they be able to fool the test? What about trained actors?

In addition to the technical problems, definitional complexity and lack of adequate range of populations studied, there are also several legal obstacles to the use of fMRI lie detection in court. One is the concern that introducing such evidence would invade the province of the jury, which has the role of determining who is telling the truth. This objection has been cited by American courts in barring the use of polygraph evidence. It would seem that the so-called "Christmas Tree Effect" and "CSI Effect" discussed earlier would be of even greater concern with fMRI images represented as depicting deception than with the relatively low technology of the polygraph. Another barrier is the Fifth Amendment right against self-incrimination. In a criminal case, it would not be possible to compel an unwilling defendant to undergo fMRI lie detection because of this right.

In considering the current state of the field of lie detection via neuroimaging, it does not seem likely that the technique will be found to be of value in the context of a criminal or civil court proceeding in the foreseeable future. The unknowns are simply too numerous for the procedure to be considered reliable enough for such use.

The accurate identification of a cognitive function as complex as conscious deception may ultimately prove to be a "bridge too far" for neuroimaging. However, there are other potential contributions neuroimaging might make in the context of interrogation. A number of observers of the field have discussed the possibility of using neuroimaging for an investigative task simpler than determining whether someone is lying. Specifically, imaging technology could potentially be used to determine whether the test subject is familiar with a stimulus, for example, a photograph of a face or of a building.

The burgeoning of neuroimaging that has taken place since the turn of the century has, by happenstance, coincided with the advent of modern terrorist tactics such as those used in the attacks of September 11th, 2001 in the United States, attacks on mass transit in cities such as London and Madrid, and many other tragedies across the globe. Not so coincidental has

been the dramatic, ongoing rise in the power, speed, and interconnectedness of computer systems (it is not coincidental since many neuroimaging techniques, especially fMRI, are dependent on sophisticated data processing capabilities). The expansion of computer speed and storage has also allowed for an unprecedented expansion of governmental surveillance capabilities which, advocates argue, facilitate the detection and thwarting of terrorist plots. Meanwhile, in the wake of revelations about widespread electronic eavesdropping by agencies such as the National Security Agency (NSA), many have expressed their concerns about government invasions of privacy.

With the refinement of neuroimaging has come the suggestion that its techniques might someday become a tool used in military operations and antiterrorism efforts. One scenario would involve displaying images to a suspected terrorist while scanning his or her brain with fMRI. If the scan indicates recognition of stimuli such as other terrorists (assuming they are not so well known as to be familiar to most people in the suspect's milieu), or particular locations, the suspect has, without answering any questions (unlike in a lie detection approach), provided information to the investigators.

The terrorism aspect is particularly relevant both because the usual constitutional protections against self-incrimination may not apply, and because it is more common to identify people planning a terrorist attack prior to it being carried out than it is to catch people in the planning stages of "ordinary" crimes such as robberies and burglaries. Granted, there are many potential flaws in searching for memories in this way. If the subject refuses to cooperate, for example, by closing his eyes or moving his head too much, the technique will not work. Furthermore, unless the stimuli are chosen very carefully, the risk of a false positive will be quite high, because a recognition memory–based technique can only determine whether the stimulus is familiar, and says nothing about the relationship between the subject being tested and the stimulus. Nevertheless, since the purpose would be to gather information to guide further investigation, rather than to prove a case in court, it is hard to state unequivocally that this type of scenario is too implausible to consider.

Such an application of neuroimaging, however remote a prospect it may be at present, has created great concern among certain observers about the potential for misuse and abuse, including the erosion of the privacy of one's own thoughts. As far back as a decade ago, civil rights advocates sought to determine if any US military or intelligence agencies had already put lie detection or memory analysis techniques into practice. Some authors have warned of the potential for these types of techniques to threaten what has been termed "cognitive liberty." To give a few concrete examples, some writers have envisioned the possibility of requiring parents to submit to neuroimaging in child custody proceedings, or of neuroimaging data

suggestive of dangerousness or violent tendencies being used for the preventive detention of people who have not been charged with any crime.

There is no publically available information to suggest that fMRI lie detection, analysis of memories, or related approaches have been used in a national security context. However, the probability of future use, while perhaps not high, is substantial enough to merit a thorough examination of the complex ethical questions raised by methods that have the potential to go far beyond what George Orwell's "Big Brother" could have imagined in his wildest dreams. Some authors have called for government regulation of neuroimaging-based lie detection and related technologies, while others predict the emergence of new privacy laws designed to protect the citizenry from being subjected to neuroimaging-assisted investigations.

CONCLUSION

The conceptual and technical advances in neuroimaging that have occurred over the past 4 decades have been nothing short of revolutionary. These breakthroughs have resulted in numerous efforts to use neuroimaging findings as evidence in legal cases. Because advocates for each side are prone to present evidence in the most favorable light (at least), and judges and juries lack the necessary scientific knowledge to fully assess claims based on neuroimaging on their own, it is critical for those who find themselves involved in this area to understand the limitations of the science as it applies to forensic questions, and to explain those limitations clearly to the court. It is critical for neutral, objective experts to protect the deliberations of the court from conjectural leaps driven by inappropriate use of neuroimaging data. At the same time, the expert must also be able to recognize and counter unfair attempts to discredit legitimate uses of neuroimaging for legal purposes.

Some of the limitations that have been discussed in this chapter are technical in nature, and could conceivably be overcome in the future as the science matures further. Others, such as the inability to use neuroimaging to draw conclusions about a past mental state, are extremely unlikely to change regardless of any future improvements in technical aspects such as scanner hardware or computer algorithms and software. A thorough knowledge of the science as well as the legal issues is key to identifying both types of limitations and effectively conveying them to the judge and jury.

It is critical to always remember never to rely solely on a single test, no matter how sophisticated, to conclude that there is a causal connection between the findings of that test and something as complex and multidetermined as an instance of human behavior. There is no one-to-one correlation between a given brain image and a particular mental state. The scientific expert must therefore guard against any tendency on the part of others to concretize brain imaging results and to suggest that an abnormality in the brain as visualized with imaging was the cause of a

person's legally relevant acts or omissions. Instead, the expert must place the imaging results into the larger context. Imaging findings may be used to support an argument—for example, that a plaintiff had dementia and was unable to contract, or that a defendant has a psychotic disorder and was not legally responsible for a crime due to insanity—but imaging findings are not sufficient in themselves to justify these types of conclusions.

In addition to the growing use of neuroimaging in court, there has been increasing speculation regarding the potential for neuroimaging to be used for other nonmedical, nonresearch applications. Outside the traditional legal arena, with its evidentiary rules and constitutional protections, it is possible that neuroimaging techniques will come to be used in a more expansive fashion to assist, for example, antiterrorism investigations. It is not very likely that the typical psychiatrist, even a forensic psychiatrist with expertise in neuroimaging, will be chosen for involvement in such pursuits should they actually be undertaken in the future. However, it is important to anticipate the temptation to resort to "mind-reading" through neuroimaging, in however primitive a form, and for society as a whole to analyze the risks and benefits of embarking on such a path.

References

*Aharoni E, Funk C, Sinnott-Armstrong W, Gazzaniga M. Can neurological evidence help courts assess criminal responsibility? Lessons from law and neuroscience. *Ann NY Acad Sci*. 2008;1124:145–160.

Aharoni E, Vincent GM, Harenski CL, et al. Neuroprediction of future rearrest. *Proc Natl Acad Sci USA*. 2013;110:6223–6228.

Anderson NE, Kiehl KA. The psychopath magnetized: insights from brain imaging. *Trends Cogn Sci*. 2012;16:52–60.

Baskin JH, Edersheim JG, Price BH. Is a picture worth a thousand words? Neuroimaging in the courtroom. *Am J Law Med*. 2007;33:239–269.

Davatzikos C, Ruparel K, Fan Y, et al. Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage*. 2005;28:663–668.

Farah MJ, Hook CJ. The seductive allure of "seductive allure." *Perspect Psychol Sci*. 2013;8:88–90.

Fritz JV. Neuroimaging trends and future outlook. *Neurol Clin*. 2014;32:1–29.

Ganis G, Rosenfeld JP, Meixner J, Kievit RA, Schendan HE. Lying in the scanner: covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *Neuroimage*. 2011;55:312–319.

*Granacher RP. The use of structural and functional imaging in the neuropsychiatric assessment of traumatic brain injury. In: Granacher RP. *Traumatic Brain Injury: Methods for Clinical and Forensic Neuropsychiatric Assessment*. 2nd ed. Boca Raton, FL: CRC Press; 2008.

Greely HT, Illes J. Neuroscience-based lie detection: the urgent need for regulation. *Am J Law Med*. 2007;33:377–431.

Greene J, Cohen J. For the law, neuroscience changes nothing and everything. *Philos Trans R*

Soc Lond B Biol Sci. 2004;359:1775–1785.

Jones OD, Schall JD, Shen FX. *Law and Neuroscience*. New York: Wolters Kluwer; 2014.

*Kalina P, ed. The use and abuse of neuroimaging in the courtroom. *Special issue, AJOB Neurosci*. 2014;5(2):35–37.

Kozel FA, Johnson KA, Mu Q, Grenesko EL, Laken SJ, George MS. Detecting deception using functional magnetic resonance imaging. *Biol Psychiatry*. 2005;58:605–613.

Martell DA. Neuroscience and the law: philosophical differences and practical constraints. *Behav Sci Law*. 2009;27:123–136.

Meegan DV. Neuroimaging techniques for memory detection: scientific, ethical, and legal issues. *Am J Bioeth*. 2008;8(1):9–20.

Moriarty JC. Flickering admissibility: neuroimaging evidence in the US courts. *Behav Sci Law*. 2008;26:29–49.

*Morse SJ. The Status of neurolaw: a plea for current modesty and future cautious optimism. *J Psychiatry Law*. 2011;39:595–626.

Morse SJ, Roskies AL, eds. *A Primer on Criminal Law and Neuroscience*. Oxford, UK: Oxford Univ Press; 2013.

Reeves D, Mills MJ, Billick SB, Brodie JD. Limitations of brain imaging in forensic psychiatry. *J Am Acad Psychiatry Law*. 2003;31:89–96.

Schweitzer NJ, Saks MJ. Neuroimage evidence and the insanity defense. *Behav Sci Law*. 2011;29:592–607.

Simpson JR. Functional MRI lie detection: too good to be true? *J Am Acad Psychiatry Law*. 2008;36:491–498.

*Simpson JR, ed. *Neuroimaging in Forensic Psychiatry; From the Clinic to the Courtroom*. Chichester, UK: John Wiley & Sons; 2012.

Snead OC. Neuroimaging and the "complexity" of capital punishment. *NYU L Rev*. 2007;82:1265–1337.

Tancredi LR, Brodie JD. The brain and behavior: limitations in the legal use of functional magnetic resonance imaging. *Am J Law Med*. 2007;33:271–294.

Thompson SK. A brave new world of interrogation jurisprudence? *Am J Law Med*. 2007;33:341–357.

Witzel J, Walter M, Bogerts B, Northoff G. Neurophilosophical perspectives of neuroimaging in forensic psychiatry—giving way to a paradigm shift? *Behav Sci Law*. 2008;26:113–130.

Wortzel HS, Filley CM, Anderson CA, Oster T, Arciniegas DB. Forensic applications of cerebral single photon emission computed tomography in mild traumatic brain injury. *J Am Acad Psychiatry Law*. 2008;36:310–322.

Wortzel HS, Kraus MF, Filley CM, Anderson CA, Arciniegas DB. Diffusion tensor imaging in mild traumatic brain injury litigation. *J Am Acad Psychiatry Law*. 2011;39:511–523.

Yang Y, Raine A. Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: a meta-analysis. *Psychiatry Res*. 2009;174:81–88.

61

History of Psychiatry

EDWARD SHORTER, PH.D.

Motto: "In the natural sciences old views and interpretations of observations are overtaken by new discoveries and retain historical interest only; in the humanities this is never quite the case. Psychiatry has one foot in each of these disciplines and our heritage is not something we can simply leave to the historians." Julius Hoenig (1981).

One of the earliest accounts of mental illness comes from the myths of the Greek gods. Hercules, under a delusive disorder, slays his children, believing them the offspring of another god.

One lad stretched forth his hands, crying, "Oh, dearest father, do not slay me; I am thy son." But Hercules, "with savage rolling eyes," as the chronicler tells us, "brought his club down upon the golden-haired head of the boy and broke his skull. The other children and the wife are sacrificed by the same murderous fury."

The modern reader recognizes the diagnosis of delirious mania, which, like most other psychiatric ailments, is as old as the hills and built into the pathology of the human condition.

Across the millennia there have been many other accounts of deluded and melancholy patients. Psychiatric illness has always been recognized as medical in nature. Yet psychiatry as a discipline arises only late in the 18th century. It was under the influence of the enlightenment that the prospect of relieving mental illness through institutional care was first broached. There had always been confinement, in the context of the noisome "hospices" that lodged indiscriminately the insane, the poor, the elderly, the medically ill, and the criminal alike. But the concept of dedicated mental hospitals stemmed from the late 18th century, surfacing first in Italy, then in France. In this manner, psychiatry became a specialty, not because its practitioners had special diagnostic or therapeutic skills, but because they had the administrative abilities required to run these large institutions. For decades, the emphasis in international surveys of psychiatric facilities would be on architecture and layout, rather than on diagnosis and treatment.

Before the late 19th century, there was no such thing as private practice psychiatry. Psychiatry as a discipline was based in mental hospitals, or "asylums,"—a term first used in 1766 instead of "lunatic hospital"—and psychiatrists were called "alienists," on the grounds that the distracted

mind had become alien. Ambulatory patients were seen in private practice by neurologists or general practitioners.

PART I: EARLY BEGINNINGS

The main story line is the transition from traditional hospices to mental hospitals dedicated to treatment and recovery. Joseph Daquin, a psychiatrist in the Savoyard city of Chambéry, extolled in a work on The Philosophy of Madness (*La Philosophie de la Folie*) in 1792, the bringing of patients "from their fetid holes … to the freedom of a great but enclosed yard. Trained staff will accompany them everywhere and keep them in view; in the yard, warders will let them do whatever they like, simply ensuring that they do not escape, and neither endanger their own lives nor those of others. We will simply let them wander about at will for two hours, or in the interval between one meal and the next." This is not the first description of the therapeutic asylum but certainly the first that extolls the therapeutic advantages of freedom and exercise.

Italy

Most accounts of the history of psychiatry do not begin with Italy, a land that is largely silent on the historical radar. Yet it was in Florence in the 1780s that modern psychiatry actually began. The key concept is "therapeutic institution," admitting patients in hope of relieving their distress rather than to protect society from them. The first therapeutic asylum was launched in 1788 when Vincenzo Chiarugi ("KEY-ah-ROO-jee") opened the doors of the old Bonifazio Hospital, previously a typical hospice, exclusively to psychiatric patients. Here, he organized gardens for the patients, provided music, gave a structure to their day, and ensured proper meals and humane treatment. In 1793 to 1794, Chiarugi wrote a three-volume psychiatry textbook based on these experiences, On Insanity (Della Pazzia), specifying that for psychological tranquilization the patients should be kept in calm, quiet environments: "The sweet songs of gentle, moving music in many cases will represent the first-line treatment." As for stimulating treatments, the patients should undertake moderate exercise in the asylum and at home. In treating melancholy, Chiarugi deemed it "absolutely necessary that … [the physician] finds his way into the patient's heart, gaining his confidence and trust." Chiargui called this "the psychological treatment" (la cura morale).

Italian psychiatry received a reformist bump-up in 1821, when Antonio Galloni became head of the venerable San Lazzaro Asylum in Reggio Emilia in the north of Italy. Founded in the 12th century as a leprosarium, in the 18th century Reggio Emilia went over to receiving psychiatric patients, and Galloni implemented many of the French reforms (see below), given that Italian psychiatry in these years stood much under the influence of France

(later, of Germany). Reggio Emilia's later superintendents included Carlo Livi, who in 1871, at the beginning of his tenure introduced Italy's first farm colony), and Augusto Tamburini (from 1877, after Livi's premature death), who began outpatient care with local families. Livi and Tamburini were among the leading figures in 19th-century Italian psychiatry, and Reggio Emilia stood as an international model for the villa system—rather than housing inpatients in giant barracks—and for the use of laboratories in patient care. Tamburini's watchwords were, "scienza e amore—studio e benevolenza" (science and love, research and kindliness).

France

Modern psychiatry was anchored in France, in two of the great Parisian hospices that Louis XIV founded in the 17th century: Bicêtre for men and the Salpêtrière for women. The seminal name in French psychiatry is Philippe Pinel, who in 1793, became chief physician of the Bicêtre hospice. Pinel is often seen (in ignorance of the contributions of Chiarugi) as the very founder of psychiatry itself. Yet the contribution of Pinel is difficult to overstate because his Medical-Philosophical Treatise on Mental Illness in 1801 (Traité Médico-Philosophique sur l'Aliénation Mentale) became the first international textbook and initiated the centuries-long task (not completed today) of perfecting the nosology of the field. In 1795, Pinel became the chief physician of the Salpêtrière, where he remained until his death in 1826. Pinel is also remembered as an early, influential advocate of humane care, as opposed to the chains-and-beatings approach that earlier characterized asylum life.

Pinel's student Étienne Esquirol ("ES-kir-ole"), who in 1825, became chief physician of the French national asylum in the Paris suburb of Charenton, is seen as the virtual founder of social psychiatry, assessing social patterns in illness, and further refining Pinel's nosology (see below). Esquirol inspired the Law of 1838 that obliged each district (department) in France to have a modern asylum; the law became worldwide a model of modern psychiatric care.

The United Kingdom

England. The first therapeutic asylum in England was St Luke's Hospital in London, founded by William Battie in 1750. Battie conceived admission to the hospital itself as having a therapeutic character, and in his Treatise on Madness (1758) he wrote, "The regimen [in madness] is perhaps of more importance than in any distemper. It was the saying of a very eminent practitioner in such cases that management did much more than medicine; and repeated experience has convinced me that confinement alone is oftentimes sufficient, but always so necessary, that without it every method hitherto devised for the cure of madness would be

ineffectual."

Yet, until the opening of the Maudsley Hospital in 1923, the anchor of English psychiatry was Bethlem Hospital ("Bedlam") founded in 1247, as a base for the Crusaders, then gradually given over to the care of the insane. (In 1930, the hospital removed from the district of Southwark to suburban Kent, at which time the old building—one of several over the centuries—became the Imperial War Museum.) The hospital is indissolubly linked with the Monro dynasty of psychiatrists, of whom Edward Thomas Monro was in office 1816 to 1855. Among the hospital's medical officers, or "apothecaries," was John Haslam, author of Observations on Madness and Melancholy (1809), the detailed cases of which are among the earliest patient reports in psychiatry.

In the first half of the 19th century, England led in the international movement to asylum reform, emphasizing "moral treatment" (psychotherapy), which began late in the 18th century at the Quaker Retreat in York, and the "non-restraint" system of patient care (abolishing chains and the like) that John Conolly initiated after he became superintendent of the Middlesex Asylum at Hanwell (a London suburb) in 1839. "The old system," said Conolly in 1856, in his Treatment of the Insane without Mechanical Restraints, "placed all violent or troublesome patients in the position of dangerous animals. The new system regards them as afflicted persons, whose brain and nerves are diseased, and who are to be restored to health, and comfort, and reason."

Scotland. Psychiatry in Scotland was dominated by the Edinburgh Royal Asylum (today the Royal Edinburgh Hospital), founded in 1813 at the Morningside estate, out of dismay at how badly Scottish poet Robert Fergusson was said to have been treated at the London asylum of "Bedlam." (He died there at 24.) The Royal Edinburgh was led over the years by a series of major figures in the history of British psychiatry, such as Thomas Clouston, whose Clinical Lectures on Mental Diseases (1883) remained for years the standard textbook.

Ireland. How direly Ireland required modern psychiatry emerges from testimony before the House of Commons in 1817:

> There is nothing so shocking as madness in the cabin of the Irish peasant, where the man is out laboring in the fields for his bread, and the care of the woman of the house is scarcely sufficient for attendance on the children. When a strong man or woman gets the complaints [severe psychiatric illness], the only way they have to manage is by making a hole in the floor of the cabin, not high enough for the person to stand up, with a crib over it to prevent his getting up. The hole is about five feet deep, and they give this wretched being his food there, and there he generally dies.

Until the foundation of the Irish Free State in 1922, Ireland was a colony of Great Britain, and Irish psychiatry rose simultaneously with

English psychiatry. The Richmond Lunatic Asylum in Dublin (now St Brendan's Hospital), the flagship of Irish psychiatry, was founded in 1816, and in 1857, Joseph Lalor became its first resident medical director. Given the absence of psychiatry in the medical curriculum until 1875, the only means of becoming a psychiatrist in Ireland, and in much of Scotland and England, was to serve as an assistant physician in an asylum. Irish clinicians of the day, such as Conolly Norman, a later superintendent of Richmond, favored the strong German emphasis on clinical psychiatry and found unhelpful the anatomical speculations of the era about causation (did the ovaries cause mental disease or not?).

German-Speaking Europe

"If French is the language of diplomacy, then German may be said to be the language of psychiatry," as Edward William Anderson, professor of psychiatry in Manchester, observed in 1962 after German had become internationally a dead language. In the years between 1850 and 1933, German was indeed the language of psychiatry, as it was the language of culture and scientific life in an enormous swath between Budapest (Hungarian psychiatrists all published in German) and Strasbourg, or between Hamburg and Vienna. Unlike England and France, where a single great city dominated the psychiatric life of the country, in German-speaking Europe there was a profusion of small states and over 30 university psychiatric departments or psychiatric hospitals existed, most of which had scientific aspirations; by 1900, Heidelberg, Berlin, Munich, and Vienna had become world centers of psychiatric learning.

Austria-Hungary. Psychiatry in Vienna was symbolized by the famous Tower of Fools (Narrenturm), a roundish structure with 139 individual cells (that still exists today), built by Emperor Joseph II in 1784 as part of the Vienna General Hospital. Although it was soon overtaken by advances in care, it remained an emblem of the distinctiveness of Viennese psychiatry, which later gave birth to such figures as Sigmund Freud (psychoanalysis) and Manfred Sakel (insulin coma therapy—see below for both).

Germany (After 1871, United as the German Empire). In the early years, German psychiatry tended to be overshadowed by the French. Yet there were spots of light. An important early figure was Wilhelm Griesinger who, as an assistant physician to Ernst Albert Zeller in the Winnenthal asylum in the early 1840s, began to propagate the view that "mental disease is brain disease," a main tenet of later German biological psychiatry. In 1845, Griesinger expressed his views in a textbook, Mental Pathology and Therapeutics (*Die Pathologie und Therapie der psychischen Erkrankungen*), a second edition of which in 1861, became the leading international textbook of biological approaches to the field. In 1865,

Griesinger moved from Zurich to Berlin as professor of psychiatry at the Charité Hospital; here he founded in 1867, the Archives of Psychiatry and Nervous Diseases (Archiv für Psychiatrie und Nervenkrankheiten) that was the equivalent of the Archives of General Psychiatry today.

Several reformed psychiatric hospitals flourished in Germany early in the century, including Siegburg north of Bonn, founded in 1825 by Maximilian Jacobi, who introduced to Germany the "no restraint" system—not shackling patients—and forms of psychotherapy. These reformed asylums also constituted centers of research as well as humane care, and in the days before the shift of scholarship to university psychiatric hospitals, an asylum such as Illenau in Achern (Baden), which opened its doors in 1842, under its inaugural director Christian FW Roller, was probably the intellectual center of German psychiatry; Roller tutored many clinicians who later became leaders of the field, including Richard von Krafft-Ebing (see below). In 1854 Jacques-Joseph Moreau ("de Tours"), after a visit to a number of German asylums, noted "the great scientific movement" in German psychiatry. This was the beginning of a vast scientific take-off.

Scandinavia

In the history of psychiatry, the contributions of the Scandinavian lands have been largely neglected, partly for linguistic reasons, partly for cultural blinkeredness.

Denmark. New-style, reformed asylums arose early in the Nordic lands. Denmark had receptacles for the insane dating back centuries. The city of Copenhagen had traditionally confined the "lunatics" at a so-called plague house. Yet following the British bombardment in 1807, living conditions for these unfortunates became intolerable, and the city decided to place patients at a castle at Roskilde, not far distant from town. The asylum, which began accepting patients in 1816, became known as St. Hans Hospital. But there were other Danish centers too. In the course of the century, neurosyphilis became a common reason for asylum admission. In 1857, Friedrich Esmarch and Peter Willers Jessen, at an asylum in the province of Schleswig (then belonging to Denmark), proposed in the General Journal of Psychiatry (*Allgemeine Zeitschrift für Psychiatrie*) that the mysterious fatal disease then called "progressive paralysis" seemed to be a late consequence of primary syphilis. A landmark in Danish psychiatry was the work of Copenhagen's Carl Georg Lange, professor of pathology at the University and considered Denmark's "first neurologist," on "periodic depressive conditions," considered by Lange to be a separate illness for which he recommended lithium therapy.

Norway. As elsewhere in Scandinavia, psychiatry in Norway rested heavily on the German tradition of biological and genetic investigation. The

first state psychiatric hospital in Norway, Gaustad Mental Hospital, in a town of that name near Christiana, was founded by Herman Major, the "first psychiatrist of Norway," in 1855. Important research came out of Gaustad, including superintendent Hans Evensen's landmark studies of schizophrenia (showing different outcomes). In 1934, Ivor Asbjörn Fölling, a biochemist at the Rikshospitalet in Oslo, the Norwegian national research hospital, discovered the metabolic cause of a disease called phenylketonuria (PKU), one of the so-called "inborn errors of metabolism" and a cause of mental retardation: It was a genetic inability to metabolize the amino acid L-phenylalanine, causing toxic amounts of phenylalanine to accumulate in the brain.

Sweden. Georg Engström founded the first reformed asylum at Vadstena in 1829, thus, as scholar Lennart Wetterberg put it, "drawing a line between barbarism and humanity in Swedish psychiatry." National mental health legislation in 1858 stipulated that an asylum must be led by a physician. The teaching of psychiatry in Swedish universities, among the earliest such teaching anywhere, began in Uppsala in 1741, with Carl von Linné, the great classifier, lecturing on medicine including psychiatry. At the University of Stockholm, called the Karolinska Institute, Eric Godelius gave the inaugural lectures on medicine, which embraced psychiatry as well, beginning in 1811. At Karolinska, later a world-famous research institute as well as a university, Wilhelm Öhrström, the first professor of psychiatry and simultaneously head of the newly built Konradsberg psychiatric hospital, began lecturing in 1861.

Spain. Although an asylum was established at Valencia in 1409, the great clinical school of Spanish psychiatry that flourished in the Middle Ages had little continuity into the modern era. In the 19th century, Spanish psychiatry remained much under the influence of the French, and it was Madrid's Pedro Mata y Fontanet, who established forensic psychiatry in Spain as the inaugural holder of the chair from 1843, when he helped introduce Esquirol's ideas. The overwhelming figure in modern Spanish neuroscience is Santiago Ramón y Cajal ("Ca-HAL"), who taught anatomy and pathology at the universities of Barcelona and Madrid. From 1888, Ramón established the neuron theory of brain histology and is generally acknowledged as the father of modern neuroscience. (He received a Nobel Prize in 1906.) His advice for a young investigator (1897, in Spanish) remains compelling reading for young scientists even today. Although he did not practice clinical psychiatry, many later prominent psychiatrists trained in the "Cajalian" phase of Spanish medicine (but note that he is generally referred to as Ramón, not Cajal, which was his mother's name). The main chairs of psychiatry were founded in the last years of the Second Republic (1931 to 1936), and the most important 20th-century clinical

figure in Spanish psychiatry, Juan J López Ibor, who originated the concept of "vital anxiety" (timopatia ansiosa), held the chair of legal medicine, toxicology and psychiatry in Madrid from 1932.

United States. In 1751, the Society of Friends (the Quakers), assisted by Benjamin Franklin, founded in Philadelphia the Pennsylvania Hospital; it was the first institution in the United States to have a major psychiatric division, for "persons distempered in mind," opened by Benjamin Rush in 1793. Rush was one of the signers in 1776 of the Declaration of Independence and practiced as much general internal medicine as he did psychiatry in his years of service from 1783 to 1813. In 1812, Rush wrote the first American psychiatry textbook, *Medical Inquiries and Observations*, upon Diseases of the Mind—and thus earned the sobriquet "father of psychological medicine in America."

The first state hospital in the United States opened its doors in 1773, in Williamsburg, Virginia. Called "The Public Hospital for Persons of Insane and Disordered Minds," it later became the Eastern State Hospital. Meanwhile, a series of private and charitable psychiatric hospitals sprang up on the model of the Pennsylvania Hospital. In 1817, the Quakers, alarmed at overcrowding in the Pennsylvania Hospital, founded a Retreat in nearby Frankford, Pennsylvania. In 1818, the psychiatric division of the Massachusetts General Hospital (incorporated in 1811) was founded at Charlestown in Boston, and named after a private donor the McLean Hospital. The Retreat, in Hartford, Connecticut, opened in 1824 and still exists today under the name Institute for Living (see below).

In 1771, a charter was granted to The Society of the New York Hospital, and in 1791, New York Hospital began accepting patients, including psychiatric; in 1808, it constructed a separate "lunatic asylum" for this purpose. At this point, the reforms of Pinel and Tuke had begun to arouse interest in the New World, and shortly proposals for "moral treatment" were circulating in New York as well. The trustees of New York Hospital therefore purchased a lot fronting the Bloomingdale Road in Harlem, and in 1821 the new structure, named the Bloomingdale Asylum, was opened, receiving private patients as well as those subsidized by funds from the Society of the New York Hospital. Years later, the Society decided that country air might be more salubrious, and in 1894 transferred the Bloomingdale patients to a new site in White Plains, Westchester, where it became the Westchester Division of The New York Hospital-Cornell Medical Center (see below for further developments).

The first large state hospital run on enlightened principles in the United States opened at Worcester, Massachusetts in 1833, under the leadership of Samuel B Woodward, an enlightened and forward-looking psychiatrist and close friend of Eli Todd at the Hartford Retreat. Following a lobbying campaign by Dorothea Dix, the hospital was enlarged in 1835 and on

several subsequent occasions; this victory marked the beginning of Dix's career as an advocate on behalf of poor people with mental illness. Among the most influential of the early American state hospitals was the Utica Hospital, opened in 1843, part of the vast New York State system of hospitals, which served as a seedbed for successful hospital administrators in many other institutions.

The shift in US psychiatry from large buildings with long corridors to small villas clustered in village-like colonies began in 1894 with the foundation of the Colony for Epileptics in Sonyea, New York, ("Craig Colony" after Oscar Craig, president of the New York State Board of Charities); the first patient was admitted in 1896. With colony care, these mental hospitals adopted a more human face.

PART II: THE FIRST BIOLOGICAL PSYCHIATRY

The intellectual content of this budding young discipline was overwhelmingly biological. At the beginning, psychiatry had a profound physical orientation. The brain was seen as the seat of diseases of the mind, and treatments addressed the brain and the body as a whole. In 1672, Oxford physician Thomas Willis (who described the "circle of Willis") said of the brain's role in his famous Two Discourses Concerning the Soul of Brutes

> When it happens that these corporeal functions are defective or hindered, forthwith the eye of the intellect, as if covered with a veil, is wont to be very much dulled or wholly darkened ... Because [of] their powers being hindered ... the imbecility and dullness of the mind, the slowness of the ingenuity, stupidity and madness at length do often arise. The fault is either in the brain itself, or the animal spirits, or both together.

For many decades, investigations of the brain to ascertain the correctness of Willis's views were hindered by the lack of suitable stains for nervous tissue. Studies of inheritance became the main technique of biological psychiatry. But in the 1870s, a new school of neuropathology, seconded with the aid of the microscope, began to arise, and the findings of such investigators as Theodor Meynert in Vienna and Franz Nissl in Heidelberg (see below) gave new impetus to biological thinking.

Nerves

Intrinsic to biological views was the assumption that disorders were caused by the nerves. Beneath the level of frank "insanity," or "madness," depression, anxiety, obsessive–compulsive disorder, and other nonpsychotic complaints were thought nervous in nature. The view was that the nervous system was somehow physically affected and that the source of mental and behavioral symptoms lay in these ailing nerves. In 1780, Samuel Tissot of Geneva, whose practice included many patients with psychiatric disorders, said, "Among the most appropriate remedies for

diminishing the irritation of the nerves, and to destroy numerous causes of their disorders, one must place the baths, cold or warm."

By mid-19th century, the concept of "nerves" had seized psychiatry in the way that "psychoneurosis" and "mood disorders" would later rule. English law distinguished clearly between insanity and nerves. As the asylum inspectors reported in 1844, "Homes which are licensed for the reception of Insane Persons ought to be kept exclusively for that purpose and the reception of nervous, imbecile, and dejected [depressed] persons, amongst those who are insane and often dangerous, is for obvious reasons open to serious objection."

Psychiatrists were once called "nerve doctors," a term that persisted in German-speaking Europe until relatively recently (Nervenarzt); a "nervous breakdown" crossed the threshold into psychosis, rendering patients incapacitated in a way that is associated with melancholic illness, acute psychosis, and delirious mania. In 1851, Henry Monro, an attending at St Luke's Hospital in London, said in his Remarks on Insanity, "The symptoms of insanity are simply the consequence of loss of nervous power." He called insanity "simply a disease of deficient nervous vigour." These words hallmarked a century of biological psychiatry.

Predisposition, Degeneration

Psychiatrists have always believed that many mental illnesses can be passed on from generation to generation, that they are "inheritable," or the result of "predisposition." The psychoanalysts did not believe this, and attributed reproducibility within families to "nurture" rather than "nature." Yet in the great sweep of things, psychiatry as a whole has believed in nature. As London's John Haslam said in 1817, "The minute investigation of many hundred cases has fully satisfied my own mind that the offspring of a person who has labored under insanity, is peculiarly liable to become similarly affected." By the late 19th century, it had become an accepted fact that the main cause of psychiatric disease was "heredity." As Paris psychiatrist Ernest Billod toured Italian asylums in the early 1880s, he was struck at the mantra-like repetition of the phrase, "The predominant cause of mental illness is heredity."

Early Pharmacotherapy

In psychiatry's early period of biological thinking, before the rise of psychotherapy, treatments of mental disease borrowed heavily from the "humoral" remedies employed elsewhere in medicine, mainly the therapies of "depletion": Correcting humoral imbalances in the body by venesection (on which Philadelphia's William Rush was very keen), leeching, and the administration of purgatives and laxatives. At the same time as depletion began to totter in the rest of medicine—toward the mid-19th century—it

came under attack in psychiatry as well and was replaced by the whole world of physical and dietetic therapies. All such treatments rely mainly upon the placebo effect, but at least they are not, unlike depletion, actively harmful to the patient. As Woodward wrote in 1850, "The abandonment of depletion, external irritants [mustard plasters], drastic purges and starvation, and the substitution of baths, narcotics, tonics and generous diet, is not less to be appreciated in the improved condition of the insane, than the change from manacles, chains, by-locks and confining chairs, to the present system of kindness, confidence, social intercourse, labor, religious teaching, and freedom from restraint. In this age of improvement, no class of mankind have felt its influence more favorably than the insane."

There was in the mid-19th century a brief but lucid anticipation of the era of psychopharmacology that would follow a hundred years later: In 1845, Paris psychiatrist Jacques-Joseph Moreau (nicknamed "Moreau de Tours," after the medical school he attended) published the results of his therapeutic experiments with hashish. The essence of psychopharmacology is establishing differential effectiveness, and in this foundational work Moreau established that hashish was ineffective in the treatment of dementia and stupor (stupidité), but that it did seem to have antimelancholic effects, lifting the mood and abolishing psychotic ideation in two melancholic patients.

As for subsequent pharmacotherapies, the opinion is widely held that psychiatry possessed no pharmacologic treatments before the psychopharmacologic revolution of the 1950s (see below). This is false. Psychiatry has always disposed of a palette of agents, admittedly of varying effectiveness (yet that is true today, as well). Until the Second World War, the Germans led the pack. In the area of medicinal innovation, in the years 1800 to 1930, Germany had 44 percent of the world's total innovations, followed by France at 16 percent and the United States at 14. The German chemical industry began its take-off around 1860, dyes in the forefront, and from the dyes emerged the first generation of psychopharmaceuticals. The very first treatments, however, were as old as the hills.

Ethanol. People are accustomed to thinking of beverage alcohol as a source of substance abuse rather than a useful therapeutic agent. Yet before the Second World War, although the dangers of alcoholism were readily acknowledged, alcohol constituted an important treatment. In 1923, Alexander Westphal, director of the university psychiatric hospital in Bonn, said, "In milder cases of melancholia, alcohol as a sedative and hypnotic can yield good service." The opinion was typical of the psychiatry of the day and not necessarily incorrect.

Opium. The milky juice of the unripe poppy plant *Papaver somniferum* in the treatment of melancholia goes back to the Ancients, and

was revived in the 17th century by London physician Thomas Sydenham, who standardized a combination of opium and alcohol (sherry) in "laudanum," or opium tincture. Practitioners of mental medicine began prescribing opium in the 18th century, and in 1788, William Falconer was found, in his discourse upon "the influence of the passions," waxing enthusiastic: "Opium too, which has been introduced into practice of late years in the nervous fever, and frequently administered, and in pretty large doses, is well known to exert great cordial effects on the mind and spirits." Tincture of opium does not seem to have been highly addictive, because so much of the alkaloid is inactivated by stomach acids. But with the invention of the syringe in the 1850s, it became possible to inject opium, and the resulting epidemic of addiction caused many psychiatrists to back away from the poppy, which seems actually to be quite effective as an antimelancholic. Yet as late as the 1960s, Hoffmann-La Roche Laboratories marketed injectable opium (Pantopon) for indications including psychiatric.

Cannabis and Hashish. The derivatives of *Cannabis sativa L* have a long history of use—today quite forgotten—in psychiatric therapeutics. After Moreau's work in 1845, hashish became widely prescribed. Auguste Voisin, the professor of psychiatry at the Salpêtrière in Paris, described hashish in 1883, as part of the panoply of agents, including potassium bromide and hydrotherapy, which he deployed in the treatment of psychotic depression. Cannabis has a similar history of age-old use, and was marketed as a drug in England in the 1930s under the brand name Indonad (in Germany as Indonal-Bürger). The cannabinoids became definitively stigmatized in the "war on drugs" of the 1970s and disappeared from psychological medicine.

Lithium. Lithium had drifted on and off the therapeutic palette for decades, if not centuries. Certain mineral springs in Texas containing lithium, for example, had a historic reputation as "crazy water." In 1886, Copenhagen psychiatrist Carl Georg Lange proposed lithium in the treatment of melancholic depression. But the remedy did not catch on, as the concept behind it, a uric acid diathesis of mental illness, was soon abandoned. (On the later history of lithium see below.)

Sedatives and Hypnotics

Medicine's encounter with sedatives goes back to 1857, when the salts of bromine ("bromides") were introduced. Among chemically synthesized sedatives, chloral hydrate followed in 1869 (and is still in use today). Paraldehyde was introduced in 1882, and a number of other synthetic sedatives were launched—all now long forgotten. The barbiturates, however, were one of the major drug classes of the 20th century. A mixture

of uric acid and malonic acid, the first barbiturate, barbital sodium (Veronal), was marketed in 1903, phenobarbital (Luminal) in 1911, and then dozens of shorter half-life preparations including amobarbital (Sodium Amytal) in 1924 and butabarbital (Butisol) patented in 1932. These agents became household names and were widely prescribed.

Despite this proliferation of agents for mild depression and anxiety, psychiatry lacked treatments for such major illnesses as schizophrenia, melancholic depression (with the exception of opium), and catatonia. In 1892, Daniel Cathell, a Baltimore physician whose practice was heavy with nervous patients, advised his colleagues, "Should you ever have recourse to remedies intended to act chiefly through the mind, if you will take care to look your patient earnestly and steadily in the face ... they will do double good. You will not only find that almost anything will relieve some of these mental cases, but will be further surprised to learn that it has evoked their enthusiasm, and that they are chanting its praise and vowing that they were cured of one or another awful thing by it. Some, indeed, who seem to be magically benefited by doses of—nothing—will actually credit them with saving their lives." The jury is still out on whether this happens today. (This thread is continued below under "The Psychopharmacologic Revolution.")

PART III: PSYCHOTHERAPY, SOCIAL PSYCHIATRY

Early Psychotherapy

Although physicians have always practiced an informal kind of psychotherapy in the doctor–patient relationship itself, formal psychotherapy begins with Hippolyte Bernheim, professor of medicine at the University of Nancy in eastern France, in 1884, when Bernheim, in his book On Suggestive Therapy under Hypnosis or in the Waking State (*De la Suggestion dans l'État Hypnotique et dans l'État de Veille*) distinguished between hypnotic suggestion (which had been around for a century) and nonhypnotic suggestion, or a good talking-to. In 1904, Paul Dubois, professor of neuropathology in Berne, Switzerland, proposed "rational psychotherapy," a forerunner of today's cognitive behavioral therapy (CBT), in his book The Psychoneuroses and Their Psychological Treatment (*Les Psychonévroses et leur Traitement Moral*). In 1903, Pierre Janet, professor of experimental psychology at the Sorbonne, proposed "psychasthenia" as the main mental disorder and advocated a kind of cathartic therapy for its treatment. These early systems were then overwhelmed by psychoanalysis, which by the 1920s, had become the chief psychotherapy of psychiatry (see below), and were forgotten.

Psychoanalysis

Previous histories of psychiatry customarily gave great space to the history of psychoanalysis, or Freudian "depth psychiatry." Yet today,

psychoanalysis has virtually disappeared from psychiatry and such extended coverage seems pointless. For detailed accounts, the reader is referred to the author's own *History of Psychiatry* (1997) or to one of the frankly adulatory chronicles of Freud's life such as that by Peter Gay. In retrospect, psychoanalysis may be seen as a kind of hiatus in an ongoing current of biological thinking in psychiatry that began with the old asylum physicians around the middle of the 19th century, and revived with enthusiasm about drug treatments in the 1950s, resulting in the substantial displacement of psychoanalysis by psychopharmacology, neuroimaging, molecular genetics, and other forms of neuroscientific investigation in the 1970s. The following discussion is appropriately brief.

Psychoanalysis, a system of thought that attributes the origin of behavioral symptoms to conflicts within the unconscious mind, was originated by Vienna's Sigmund Freud in the late 19th century. Although Studies in Hysteria, the volume that Freud coauthored with Vienna family physician Josef Breuer in 1895, is commonly seen as the founding document of psychoanalysis, it merely presents psychiatry's most famous patient, "Anna O" (Bertha Pappenheim) in the context of abreactive therapies. Freud's seminal doctrines themselves come to blossom only in 1900 in The Interpretation of Dreams (*Die Traumdeutung*). In successive years, Freud slowly adumbrated the essentials of psychoanalytic treatment, explaining the role of the interpretation of dreams and of free association. Then in the early 1920s Freud proceeded to trace the architecture of the psyche in the layers that he hypothesized: ego (conscious mind), superego (the source of morality, supposedly derived from oedipal fears toward the father), and the id (unconscious mind, harboring the sex drive and the death instinct).

Psychoanalysis as a movement began in 1902, when Freud founded a discussion group called the Psychological Wednesday Society, a group of curious physicians—many of them spa physicians and physiatrists in search of some kind of psychic bond to their patients. In 1908, this discussion group became the Vienna Psychoanalytic Society. In the same year, the first international congress of psychoanalysis took place in Salzburg, followed by a second congress in 1910, in Nuremberg, and a third in 1911, in Weimar. Meanwhile, the movement was plagued by schisms as early advocates such as Alfred Adler broke away, unable to accept fully Freud's doctrine of infant sexuality.

After the First World War, psychoanalysis became a major international movement. The dominant form of medical psychotherapy shifted from the former systems of Dubois and Janet to psychoanalysis. Psychoanalytic societies sprang up in the major western cities, and only with the Nazi seizure of power in 1933 (and Anschluss with Austria in 1938) was this expansion brutally braked. As the movement leached out of Europe, the United States, especially New York, Washington DC, and Los Angeles,

became the world epicenter of psychoanalysis.

In the New World, under the influence of such émigé analysts as Hungarian-born Sandor Rado in New York, depressive neurosis began to loom in psychoanalysis. Freud himself had little interest in depression because the patients wanted mainly to be left alone and not tell their dreams or free associate. But in New York "depressive neurosis," or "neurotic depression," started to become the meat-and-drink of the Park Avenue analysts, as opposed to "hysteria," "phobic neurosis," and the other diagnoses that Freud himself had placed front and center. In 1931, at the request of Abraham Brill, Rado helped organize training at the newly founded New York Psychoanalytic Institute. After disagreement with other members, in 1945 Rado headed the newly opened Columbia University Psychoanalytic and Psychosomatic Clinic for Training and Research, the first training institute within a university.

Once psychoanalysis had drifted into exile, the movement began to break into rival camps, such as ego psychology (the classic drive-defense tradition as continued in Anna Freud's 1936, The Ego and the Mechanisms of Defense); Donald Winnicott and Melanie Klein's object relations theory; and self-psychology (associated in particular with Heinz Kohut's book *The Analysis of Self* [1971]).

In object relations, Melanie Klein's *The Psycho-Analysis of Children* (1932) highlights the importance of introjection and projection as mechanisms. As Klein's biographer John Arnold Lindon explained, "Projection alters the infant's impression of his environment, and by introjection this changed picture of his environment influences what goes on in his mind. Thus an inner world is built up that is partly a reflection of the external one."

After the 1970s, psychoanalysis became less quarrelsome and drifted even further from Freud's original emphasis on infantile sexuality. Psychodynamic psychiatry (which is not the same thing as psychodynamic psychotherapy) agrees on the importance of unconscious conflict, on psychic determinism, on the role of development, transference, counter-transference, and resistance. The field is no longer opposed to medication (analysts once scorned "pills" as a hindrance to getting at the "root causes" of the complaint). Glen Gabbard at Baylor College of Medicine in Houston writes in 2009, "... A patient with a bipolar disorder who is denying that he has an illness and refusing to take lithium may ultimately have better compliance with the medication if the clinician explores the meaning of his denial and his reluctance to consider himself as someone requiring treatment."

Yet, despite the efforts of psychoanalysts to integrate findings from psychopharmacology and the neurosciences, the wind has slowly gone out of the field. There had been a time when the chair of virtually every major department of psychiatry in the United States had been a psychoanalyst.

But by 1990, more than 100 of 163 residency programs for psychiatry had abandoned instruction in intensive psychotherapy. The percent of psychiatrists who were also psychoanalysts declined from around 10 percent in the 1950s to 2.7 percent by 1988. What once was the very intellectual core of the discipline of psychiatry has become a marginalized group of enthusiasts. As Arthur Yuwiler, a senior biochemist in the Department of Psychiatry at the University of California, Los Angeles (UCLA), reflected somewhat ruefully in 2004, on the exchange of the psychoanalytic paradigm for the biological one, "I seemed to spend the first half of my scientific life arguing with psychiatrists that there was a brain as well as a mind and the last half arguing there is a mind as well as a brain."

Nonpsychoanalytic Psychotherapies

Formal systems of psychotherapy began with Hippolyte Bernheim, an internist in the French provincial medical school at Nancy, in 1884, with his distinction between hypnotic and nonhypnotic "suggestive" therapies: the latter meant, essentially a good chat. Bernheim believed hypnosis to be an effective treatment. Jean-Martin Charcot in Paris, however, thought hypnotizability to be merely a diagnostic marker of "hysteria."

Group psychotherapy began with Scottish psychiatrist William Alexander Francis Browne, at the Crichton Royal Hospital in Dumfries in 1842, when he encouraged patients to stage amateur theatricals that amounted to a form of "mental therapeutics." This strand was continued in 1922 in Vienna with Jakob Moreno-Lewy's impromptu theater that was later tagged "psychodrama." After Moreno immigrated to the United States in 1925, he introduced in 1931 group psychotherapy. Independent of Moreno, in these years Louis Wender at Hillside Hospital and Paul Schilder at Bellevue Hospital, both institutions in New York, also began psychotherapy in groups. Closely related to group psychotherapy was the "therapeutic community," which birthed in England during the Second World War at the hands of Joshua Bierer and Thomas Main. (See section on social psychiatry.)

In the United States the transition of psychotherapy from physicians to psychologists and psychiatric social workers began with psychologist Carl Rogers' "client-centered therapy" after 1946. His book of that title appeared in 1951.

Family therapy is generally considered to have begun with Columbia University's Nathan Ackerman at the founding in 1956 of the "family mental health clinic," part of New York's Jewish Family Service.

Social and Community Psychiatry

The main narrative arc in psychiatry is basically a first (19th century) biological psychiatry; then a psychoanalytic hiatus in the first half of the

20th century, followed by a second biological psychiatry (late 20th century). Yet in the years 1940 to circa 1980 another current arose in psychiatry flowing separately from this main narrative stream: It was social and community psychiatry.

After the Second World War, the field became increasingly preoccupied with the social causation of illness and with group therapies for its treatment. A movement arose to treat patients in the context of the community rather than the asylum, and in 1946, D. Ewen Cameron, head of psychiatry at McGill University, created the first day hospital in the world at the Allan Memorial Institute in Montreal. In 1948, Joshua Bierer established the first in England, a "social club" at the Social Psychiatry Centre in London that he had founded 2 years previously. The movement incorporated in 1959, Bierer's view that, "Treatment must include the whole social environment of the patient and all his social relationships. He must be treated not only as a person but as part of a community." So influential did social psychiatry become in the United States that it influenced the removal of National Institute of Mental Health (NIMH) in the years 1967 to 1992 from the National Institutes of Health, to protect, according to Joel Paris, head of psychiatry at McGill and author of Fall of an Icon (2005), "the social mission of the Institute from mainstream medicine." Many social psychiatrists also saw themselves as social reformers, on the grounds, as Paris continued, that "poverty and social disruption were the real causes of psychological symptoms." Therefore, "society, not the individual, should be the 'patient.'"

These ideas accorded well with the social activism of the mid-20th century. But as early as 1956 Ralph Gerard, a neuroscience leader at the University of Michigan, said presciently, "I think there has been a real overshoot. We have all wished mental disease to be determined mostly by sociological causes, because this situation seemed more hopeful from the therapeutic standpoint. If the individual has been broken by the group he could be healed by techniques of interpersonal interaction; but if he had bad genes, effective treatment was far in the future." And indeed, in the United States, social psychiatry was rapidly eclipsed by mushrooming biological perspectives.

Today, CBT is probably the most widespread form of medical psychotherapy, and is associated with Philadelphia psychiatrist Aaron Beck, from 1963. Popular as well is Interpersonal Psychotherapy, originated in 1967 by Yale psychiatrist Gerald Klerman and coworkers, especially for the treatment of depression. Also of interest is "dialectical behavior therapy" (DBT), a form of CBT emphasizing "mindful awareness," proposed by Marsha M Linehan, a psychologist at the University of Washington, in 1999.

Yet while the number of psychotherapies proliferated, the percentage of psychiatric consultations in the United States that involved psychotherapy

declined from 44.4 percent in 1996 to 1997 to 28.9 percent in 2004 to 2005. Clearly, the single-office psychiatrist with a primarily psychotherapeutic practice, once the very image of psychiatry with its emphasis on "the couch" and "root causes," was giving way to the health care team with the psychiatrist's primary purpose being diagnosis and the provision of medication.

PART IV: FROM ASYLUM TO UNIVERSITY

In the last quarter of the 19th century, many national medical bodies began demanding that psychiatry be included in the medical curriculum. This created a demand for professors of psychiatry and for academic centers in which to train them. As a result of this curricular development, research began shifting from asylums to university psychiatric hospitals and academic departments of psychiatry. Investigators in many of these centers came to dominate the field and the asylums became research backwaters, increasingly given over to warehousing the ever-growing numbers of patients rather than treating them.

German-Speaking Europe

Said British psychiatrist Malcolm Lader in 1990, "We always used to say if you go back far enough and look at German literature at the turn of the century, you're sure to find that someone described panic disorder and everything else." It was the closeness of the link between the academy and the psychiatric hospital that granted this extraordinary German celerity in psychopathology; much of this architecture of the delineation of syndromes still remains. This leap from asylum to academy in German-speaking Europe sped forward in 1874, when for the first time, university professors began taking part in the Association of Southwest German Alienists, hitherto a group of influential asylum psychiatrists.

Vienna. The capital of the Austro-Hungarian Empire, Vienna was a psychiatric super-power, the global influence of which was far disproportionate to the medium size of the city. The train of distinguished psychiatrists began in the late 18th century with such figures as Franz Joseph Gall (who first speculated about cerebral localization), and continued into the nineteenth with Ernst von Feuchtersleben, whose 1845 Textbook of Medical Psychiatry (*Lehrbuch der ärztlichen Seelenkunde*) popularized the term "psychopathy," in the sense of a disordered personality. (Feuchtersleben also distinguished between "neurosis" and "psychosis," giving them exactly the opposite meaning they were later to have.) But the modern tradition of Vienna as a bastion of biological psychiatry (psychoanalysis came later) began with psychiatry professor Theodor Meynert ("MY-nert") with his great 1869 treatise on neuroanatomy, The Architecture of the Cerebral Cortex (Der Bau der

Gross-Hirnrinde), and continued with Illnesses Involving the Forebrain in 1884 (Erkrankungen des Vorderhirns), which was completely forgotten midst later shrieks of jubilation about the importance of the frontal and prefrontal lobes in schizophrenia. His successor Richard von Krafft-Ebing is remembered for his textbook of sexology Psychopathia Sexualis (1886); his successor in turn, Julius Wagner von Jauregg ("YOW-reg"), introduced in 1917, the malarial fever cure, an epochal accomplishment for which he received the Nobel Prize in 1927. Aside from the professors, Josef Berze, head of the Vienna city asylum "Am Steinhof," described "primary insufficiency of psychic activity" in schizophrenia in 1914, and Constantin von Economo, who led a brain research institute, reported the new disease of "encephalitis lethargica" in 1917. It was in the years following 1930 that Manfred Sakel proposed insulin coma therapy (while still in Berlin; only in 1933 did Sakel return to Vienna to conduct clinical trials—see below). Finally, pediatrician Hans Asperger described "autistic psychopathy" in 1944, among the earliest descriptions of autism (which was first used as a term in pediatric psychiatry by German and Russian neurologists and psychiatrists early in the 1920s). Vienna's contribution to psychoanalysis would begin in 1900, with the publication of Sigmund Freud's Interpretation of Dreams (see above), and continue until the expulsion or murder of the Jewish psychoanalysts in 1938 after Anschluss with Germany. After the Second World War, Vienna never again regained this historic preeminence in psychiatry.

Heidelberg. The flagship of academic psychiatry in the German Empire berthed in Heidelberg. The "Psychiatric-Neurological Clinic" of the University ("Klinik" in German means university hospital) was founded in 1879 by Otto Fürstner; it was the first university psychiatric department with beds of its own, so that patients didn't have to be brought from a nearby asylum for demonstration in clinical lectures. Its most distinguished head was Emil Kraepelin, who had the professorship from 1891 to 1903, during which time the crucial editions of his textbook appeared (see below). The Klinik was also a provincial asylum, with a catchment basin of its own. When Kraepelin left for Munich, the professorship was assumed by neuropathologist Franz Nissl, known for his demonstration in cells of the "Nissl substance." At various points in Kraepelin's Klinik served Alois Alzheimer (who followed Kraepelin to Munich) and Karl Jaspers. In the 1920s, such important figures as Willi Mayer-Gross (who after fleeing the Nazis to England coauthored the standard English psychiatry textbook, Clinical Psychiatry, in 1954); Carl Wilmanns (pioneering studies of the homeless); Hans Gruhle (premier schizophrenia researcher of his day); and August Homburger (who founded in Germany modern pediatric psychiatry). This wall of scholarship crumbled in 1933, and the names of the successors under the Nazis are either forgotten or live on in infamy.

Heinz Lehmann later said in retrospect, "German medicine was an impressive thing and there is no doubt about it. When we all went away, I don't think this tradition was transferred to America but it was abolished in Germany."

Berlin. In 1795, a psychiatry department was founded in the Charité, the Prussian state's great military hospital, and in 1806 Ernst Horn was called to become associate director of the hospital, with simultaneous responsibility for psychiatry. Under Horn, psychiatry at the Charité conformed to the concepts of the day of a proper reformed asylum, though Horn was much assailed for some of his practices and left under a cloud in 1818. Ten years later he was succeeded by Carl W. Ideler, one of the so-called "Romantic" psychiatrists who saw illness as evidence of sin. This unhappy period came to an end with Ideler's death in 1860 and Wilhelm Griesinger's arrival in 1865 to convert the old psychiatric ward into a modern neuropsychiatric clinic (Klinik für Geistes- und Nervenkrankheiten). In 1867, Griesinger founded the lead German journal for biological psychiatry, the Archive for Psychiatric and Neurological Diseases (Archiv für Psychiatrie und Nervenkrankheiten, of which the current Archives of General Psychiatry is a distant echo). Griesinger died prematurely in 1868 of acute appendicitis. Following Griesinger came a distinguished series of biologically oriented psychiatrists, including Carl-Friedrich Westphal in 1869 (who in 1872 described agoraphobia and in 1877 obsessive–compulsive disorder). His tenure ended with his death in 1890 of what seems to have been neurosyphilis—of which so many middle-class men perished in those days (infected as medical students or law students with primary syphilis from prostitutes). Karl Bonhoeffer had the Berlin chair from 1912 to his emeriting in 1937, after which another grim period began in the history of German psychiatry. In 1909, Bonhoeffer conceived the notion of "exogenous psychosis," arising from stress or toxins from outside rather than being endogenous and inborn.

In a separate development, in 1902, Oskar Vogt founded a neurobiological institute at the University, which in 1915 was joined with the Kaiser-Wilhelm-Institute for Brain Research. (His wife, the French-born Cécile Vogt, became an important neuroscientist, later settling in England. Their daughter Marthe discovered in 1965 that noradrenalin was present in the brain.) In 1931, the institute was transferred to the Berlin-Buch mental hospital, and came in 1937, under the direction of Hugo Spatz, who was a member of the Nazi party but who goes down in history as codiscoverer of the Hallervorden–Spatz disease of neurodegeneration. Despite its vicarious history, this institute became one of the world's leading neuroscience centers.

Munich. When in 1903, Kraepelin went from Berlin to Munich, it was

as inaugural director of a newly built University Psychiatric Hospital in the Nussbaumstrasse. He proceeded to make it into a world-class institution, bringing, as stated, Alois Alzheimer from Heidelberg to head the pathology laboratory, to which young investigators streamed from all over the world.

In 1917, Kraepelin founded a freestanding psychiatric research institute, the Kaiser-Wilhelm-Institut für Psychiatrie. The psychiatric department of the adjacent Schwabing City-Hospital, headed after 1931 by Kurt Schneider, became its clinical arm. Slow to get off the ground, the Institute's take-off began only in the mid-1920s, thanks to a group of foreign donors led by James Loeb (after whom a street in front of the present institute is now named). In 1917, Ernst Rüdin became head of the "genealogic-demographic" (genetics) section (see below). The scientific promise of the Institute was nipped in the bud by the Nazi takeover, and only after the war, when in 1954 it became the Max Planck Institut für Psychiatrie, did its international reputation blossom.

Zurich. In Zurich, a city psychiatric hospital was founded in 1860, and when Wilhelm Griesinger began lecturing there in 1863, it became a "Klinik." Two years later, in 1865, Griesinger decamped for Berlin. In 1869, a grand new Klinik was planned at a spot called the "Burghölzli"; the "Bli," as it was known, was helmed over the years by some important figures, including Auguste Forel (1898 to 1927), a co-originator of the "neuron theory" of cerebral communication; Eugen Bleuler (1898 to 1927), who coined the term "schizophrenia"; and his son Manfred Bleuler (1942 to 1969), who in 1954, determined that the prognosis of schizophrenia was often benign.

France. Paris is the great focus of French psychiatry, and in Paris each of the main general hospitals had academic psychiatry departments chaired by professors. Yet in addition, Paris had one special institution not found elsewhere in France (or in other lands, for that matter): The special psychiatric reception service of the Prefecture of Police (la Infirmérie Spéciale des Aliénés de la Préfecture de Police).

In 1800, Paris was given a special prefect for police, and with the general mental patients law of 1838, this prefect acquired responsibility for mental patients who came to public attention. Seven years later, in 1845 this special reception service became headed by a psychiatrist. From this service flowed a good deal of the research of modern French psychiatry. The first head of note was Ernest-Charles Lasègue in 1850, who coined the diagnostic term "delusions of persecution" in 1852, and "hysterical anorexia" (anorexia nervosa) in 1873. (Lasègue also had a series of academic appointments in the teaching hospitals.) In 1872, the Infirmérie Spéciale became independent of the police headquarters, and turned into an important center for research. Following his death in 1883, Lasègue was

succeeded by his assistant Henri Legrand du Saulle, who in 1875 offered an early description of obsessive–compulsive disorder ("la folie du doute avec délire du toucher").

At the Salpêtrière hospice, in 1882, Jean-Martin Charcot became the inaugural holder of a chair in nervous diseases, created especially for him by the president of the Republic. Charcot's contributions lay mainly in neurology, but his name is indissolubly connected with a certain kind of "hysteria," involving supposed chronic stigmata and elaborate acute pantomimes that in retrospect were phenomena of suggestion and that disappeared after his death in 1893. Among the students of Charcot, best known is Joseph Babinski, who in 1896 described the plantar extension reflex that is named after him. Charcot's successors Fulgence Raymond and Jules-Joseph Dejerine were better known as neurologists than psychiatrists.

The "clinical asylum" Sainte-Anne was opened in 1867 as the asylum of the Seine department, with Valentin Magnan, the notable advocate of "degeneration" theories, as director of its admissions department. In 1877, Benjamin Ball became the inaugural holder of the Chair of Mental and Brain Disease. Among his successors were Henri Claude in 1922, an early biological psychiatrist, and, Jean Delay in 1946, who introduced the psychopharmacological era in France. The facility became a "psychiatric hospital" in 1937, a "psychiatric center" in 1944, and a "hospital center" (shedding the unfashionable adjective "psychiatric" entirely) in 1974.

United Kingdom

LONDON. In 1908, the asylums committee of the London County Council accepted a gift from psychiatrist Henry Maudsley to found a research hospital based on the model of the German university psychiatric clinic. The war interfered with the realization of these plans, but in 1923, the Maudsley Hospital opened its doors for psychiatry, with Edward Mapother its inaugural superintendent. The Maudsley was not to be an asylum. Mapother made clear, as historian Edgar Jones relates (in a 2003 article in Medical History), "that no patient was to be admitted under section, nor would they be certified once in the hospital." Mapother had an academic appointment at King's College Hospital, and wanted to make the Maudsley the psychiatric arm of the college's teaching. (Mapother received the title "professor" in 1936.) Aubrey Lewis, who joined the staff in 1928, becoming medical officer in 1936, recalled, "The place was small, full of ambition and energy, but everybody was overworked." In 1946, Lewis was appointed professor of psychiatry, and 2 years later, in 1948, the teaching unit at the Maudsley was renamed Institute of Psychiatry (IOP). In that year as well the Maudsley merged with the Bethlem Royal Hospital (the former "Bedlam"). Lewis began adding Medical Research Council (MRC) units, the

first of these, the MRC Social Psychiatry Unit under Lewis himself opening in 1948. These helped make the Maudsley the premier institution in the United Kingdom for research in psychiatry. In 1997, the IOP split from the Maudsley to become a school of King's College London. The Maudsley was a major font of social psychiatry in the UK. Psychiatrist David Healy recalled that Aubrey Lewis had stated, at a meeting in 1958 (in Healy's words), "'If we had a choice between the new drugs and the social treatments, such as industrial rehabilitation units, we would pick social treatments.' And it's interesting that in the UK under Lewis' influence the Maudsley remained very aloof from the new drugs."

Accordingly, the first center of research for psychopharmacology in the UK—the first anywhere—was the Department of Experimental Psychiatry, founded by Joel Elkes at Birmingham in 1951. It was here that some of the earliest trials of the antipsychotic chlorpromazine were conducted. "I felt like a naturalist advancing into a strange continent," said Elkes later. The notion that drugs, as opposed to psychotherapy, could prove curative in psychiatry was then very new.

On Scotland and Ireland (see above)

CANADA'MCGILL UNIVERSITY. Canada contained one of the most internationally important departments of psychiatry at McGill University in Montreal. The Department of Psychiatry of McGill was the Allan Memorial Institute (named after a donor), founded in 1940 at the instigation of neurosurgeon Wilder Penfield. Chaired by D. Ewen Cameron from 1943 to 1964, it was under Cameron that the "Allan" became one of the most distinguished psychiatry departments in North America. Its clinicians and scientists included psychoanalyst Miguel Prados, who arrived from Spain in 1944; the neuropathologist Karl Stern who in 1944, established a gerontological unit, the first in the world; the neuropathologist Vojtech Adalbert Kral, a Holocaust refugee who arrived at the Allan in 1953 and pioneered research in dementia; Clifford Scott, who became an international leader of psychoanalysis, arrived from the Maudsley in 1954. (One wag said "Kral felt if you couldn't see it, it didn't exist and Clifford Scott … felt if you could say it, it wasn't important.")

Included as well were: Neurophysiologist Robert Cleghorn in 1946, who pioneered endocrinological psychiatry; Charles Shagass in 1950, who conceived the notion of the "sedation threshold"; Eric Wittkower, who in 1955 set up the section on transcultural psychiatry as a joint venture with the department of anthropology; Wittkower recruited Henry BM Murphy, author of the standard textbook of transcultural psychiatry (Comparative Psychiatry, 1982), who became head of the section in 1965 after Wittkower retired; and Heinz Lehmann, who came to Montreal in 1947 as director of the Verdun Protestant Hospital and is considered one of the founders of

psychopharmacology. (The World Health Organization's first training center for psychopharmacology was at the Douglas.) Psychiatrist George Simpson, a first-year resident at McGill in 1956, later said in an interview, "I would argue that the Allan was the most stimulating place in the English-speaking world at that moment, even though the rest of the city spoke French."

United States. After the Second World War, the United States came to dominate world psychiatry, partly as a function of the torrent of research funds from NIMH, partly as a result of the Diagnostic and Statistical Manual of the American Psychiatric Association becoming after 1980, the world bible of psychiatry. The Holocaust and the devastation of the Second World War also had a great deal to do with the new American primacy. Rather than flocking to the Maudsley or Vienna for training, as budding American psychiatrists did before the war, it was young psychiatrists from all over the globe who sought further training in the United States. This new American hegemony accounts for the length of the following section.

NEW YORK. New York held the greatest concentration of psychiatric firepower in the United States owing to a series of public and private teaching hospitals.

HILLSIDE HOSPITAL. Hillside Hospital, initially in Hastings-on-Hudson, was founded by Israel Strauss in 1927 as a charitable Jewish hospital operated in a nonsectarian manner. In 1941, the hospital removed to Glen Oaks, Queens. It was, as Strauss put it, "the first institution of its kind for Jews in this country." Hillside specialized in "borderline" cases, meaning the kind of psychoneurosis that usually would not receive in-patient treatment. In 1930, Louis Wender, chief physician, began one of the first programs of group psychotherapy in the country. In the 1940s, Hillside was among the earliest US institutions to offer the various shock therapies. Although mainly psychoanalytic in inspiration, in 1950 Hillside resolved upon a "biochemical" research program, and in 1952 hired Max Fink to establish the Department of Experimental Psychiatry, which was joined by Donald Klein in 1957. Fink and colleagues set up an electroencephalogram (EEG) laboratory in 1953, and with Klein, conducted pioneering trials in psychopharmacology. By 1999, following several mergers, what had become Zucker Hillside Hospital was merged into the North Shore Long Island Jewish Health System, which encompassed seventeen hospital systems.

NEW YORK/PRESBYTERIAN/WESTCHESTER DIVISION. The foundation and growth of the New York Hospital, the second oldest freestanding psychiatric hospital in the United States, is discussed above. In 1898, the hospital became associated with Weill Cornell Medical College as that institution was

founded, and a 1927 endowment lent to the psychiatric division the title Payne Whitney Psychiatric Clinic. New York Hospital merged with Presbyterian Hospital in 1996, and the Payne-Whitney Clinic was renamed the New York/Presbyterian Westchester Division.

NEW YORK STATE PSYCHIATRIC INSTITUTE. In 1896, the New York State Pathological Institute was established in New York City under Ira Van Gieson. It was the first effort in the state to study the anatomical basis of insanity, receiving specimens from the various state hospitals. In 1904, the Swiss-born Adolf Meyer, previously a pathologist at several state hospitals and a professor at Clark University, succeeded him. Meyer was already a significant scientific figure, and would go on in 1910 to become professor of psychiatry at Johns Hopkins University and inaugural head of the Henry Phipps Psychiatric Clinic (see below). In 1902, the Pathological Institute moved to Ward's Island next to the Manhattan State Hospital. Meyer changed the name to Psychiatric Institute in 1907. Following a fire at Manhattan State, planning began to convert the institute to a state mental health research hospital and to situate it at Washington Heights, where Columbia was building a new medical school and Presbyterian Hospital was to move. The new building of the now New York State Psychiatric Institute was dedicated in December 1929. Known as "PI," it became the psychiatry department of Columbia University as well offering from 1931 a 10-week course for psychiatrists in the New York State Mental Hospital system. In 1936, Nolan DC Lewis, professor of neurology in the Neurological Institute, became the first academic to lead PI. He in turn was followed by Lawrence C Kolb in 1954. The Psychiatric Institute innovated at all points on the treatment spectrum: In 1942, a Department of Research Psychiatry was created, with Eugene Barrera and Bernard Pacella its main members. In 1945, a Psychoanalytic Center for Training and Research was founded, the first official psychiatric training institute to be established in a medical school. Despite the massive interest at PI in psychoanalysis, early in the 1950s a small number of investigators led by Philip Polatin became involved in psychopharmaceutical trials; by 1956, the Institute had a department of experimental psychiatry, now led by Sidney Malitz, located at the Vanderbilt Clinic.

PI benefited greatly from the tumult of the century, giving refuge to many distinguished scientists and clinicians fleeing the Holocaust. Among the members of the informal "European Club" were Lothar Kalinowsky (from Germany and Italy), the padrone of electroconvulsive therapy (ECT) in the United States; neurochemist Heinrich B Waelsch (Prague); Paul Hoch (Hungary and Germany), who headed experimental psychiatry and coined the diagnosis "pseudoneurotic schizophrenia"; Sandor Rado (Hungary) who became a leader of psychoanalysis in the New World; Armando Ferraro (Italy), who headed neuropathology at PI. Eric Kandel,

who established the Center for Neurobiology and Behavior in 1974, was born in Vienna, and later (2000) won a Nobel Prize.

THE EAST COAST

Washington DC: The National Institute of Mental Health. The NIMH, funded by the federal government as one of the National Institutes of Health, is not a degree-granting institution, but it is very academic, encouraging original research both in-house, in the "intramural program," and by funding "extramural" psychiatric research across North America. The involvement of the federal government with psychiatric research began in 1929 with the founding of the Narcotics Division of the Public Health Service, renamed Division of Mental Hygiene in 1930. Lawrence Kolb became the chief PHS officer in charge of the division of mental hygiene, and on his retirement in 1944 he had wished to set up a national neuropsychiatric institute within the National Institute of Health (NIH). This became possible in 1946, when the National Mental Health Act called for the creation of a National Institute of Mental Health, with an in-house research program. NIMH opened its doors in 1949, and Robert Felix, who had succeeded Kolb as director of mental hygiene, became the institute's inaugural director in 1951, with physiologist Seymour Kety in charge of research.

In 1953, the first clinical NIMH ward opened at the Clinical Center for the NIH; there followed in 1955 the Laboratory of Clinical Science, and in 1956 the Psychopharmacology Service Center, which in 1960 organized the Early Clinical Drug Evaluation program, Jonathan Cole its director. The program became a great engine of progress, conducting investigational trials in the 35 Early Clinical Drug Evaluation Units (ECDEUs) that came to be established by 1969. By 1982, among the 100 investigators most frequently cited in scientific research worldwide, 9 were in the NIMH the intramural program.

Johns Hopkins University: The Phipps Psychiatric Clinic. Psychiatry at Johns Hopkins began in 1910 when Adolf Meyer was hired from the Pathological (later Psychiatric) Institute of the New York State Hospitals to become the inaugural professor of psychiatry and director of the new university psychiatric hospital, the Phipps Clinic, funded by Henry Phipps, which opened its doors in 1913. On his retirement in 1941, Meyer was succeeded by Nebraskan John Whitehorn, in office until 1960. In these years, Rockefeller Foundation grants brought many young European psychiatrists to the Phipps to study, and it was through them that Meyer's international influence became considerable.

McLean Hospital (continued from above). In 1888, profiting from the scientific boom then sweeping US psychiatry, a pathology department was organized at the McLean asylum. In 1895, the institution removed from Charlestown to Waverley (Belmont), a suburb of Boston, and was named McLean

Hospital. Swiss-born August Hoch, who would shortly study with Kraepelin and Alzheimer, became leader of a chemical laboratory in 1893; Hoch became an early advocate of the new neurohistopathology; a psychology laboratory followed in 1904. After 1955, in keeping with the changing times, the hospital pioneered in the United States concepts of the therapeutic community under Alfred H. Stanton.

But biology lay waiting. In 1972, the "transformation of McLean," as Alan Schatzberg later put it, began, when Shervert Frazier, who had also qualified as an internist, came up from Columbia University to become director. He recruited psychotherapist John Gunderson in 1973, who became the "borderline" specialist; Jonathan Cole and Schatzberg came in 1974 to work in depression; there arrived as well from Massachusetts General Hospital (the parent organization), Joseph Lipinsky in 1974, and Seymour Kety and Ross Baldessarini in 1977. Schatzberg said, "He recruited outstanding people and built the Mailman Research Center, an unbelievably vibrant place."

Harvard University. The Department of Psychiatry at Harvard was founded in 1920 when Adolf Meyer's student, the Scottish-born Charles Macfie Campbell, left the Phipps Clinic in Baltimore to become director of the Boston Psychopathic Hospital and the inaugural professor of psychiatry at Harvard. Like his mentor, Campbell leaned toward "psychobiology" and psychoanalysis. The influential post at Harvard, however, is chief of psychiatry at the Massachusetts General Hospital, one of Harvard's teaching hospitals (of which McLean was administratively a part). This post was initially held in 1934, by Stanley Cobb, professor of neuropathology and previously director of the Harvard Neurological Unit at the Boston City Hospital. Some consider Cobb the founder of biological approaches to psychiatry in the United States. Among Cobb's successors were Erich Lindemann in 1955, John Nemiah in 1965, and Leon Eisenberg in 1967.

The Institute of Living (continued from above). In 1931, when C Charles Burlingame was appointed superintendent, what was then the Neuropsychiatric Institute of the Hartford Retreat was in a downhill slide. Burlingame effected a revival, establishing a school of nursing, bringing in recognized figures on staff (to regain the lost accreditation), and, later in the 1930s, implementing the new "shock" therapies. In 1943, the name was changed to the Institute of Living. In 1946, Francis J Braceland, who had studied at the Zurich University Psychiatric Hospital and the National Hospital for Nervous Diseases in London, became chief psychiatrist. (Braceland is also remembered for being a special witness at the Nuremberg Trials in 1946, whose testimony saved Rudolph Hess from execution.) Said a New York Times obituary in 1985, Braceland is "credited with changing its reputation [Institute of Living] from that of a plush sanitarium for the rich and famous to that of a highly respected psychiatric hospital for the seriously ill." In

1994, the Institute of Living became part of Hartford Hospital.

Yale University. Yale had the third university department of psychiatry in the United States—after Johns Hopkins and Harvard. The central figure here was Dean Milton C Winternitz, a neuropathologist who conceived the Institute for Human Relations, which opened in 1931. (The department of psychiatry was situated in the Institute.) Winternitz said in 1925 of the proposed Institute, "... All of these disciplines primarily interested in the problems of human behavior from the standpoint of investigation, graduate teaching, and application, i.e. therapy, [should] be brought together geographically in a single building designed to allow the greatest possible latitude for ... cooperation." An embryo department of psychiatry had come to life in the 1920s. Yet only in 1930 was it led by a significant figure, when Eugen Kahn, a Kraepelin student and collaborator with Ernst Rüdin in schizophrenia genetics, became chair and the first Stirling Professor in psychiatry. He was followed in 1950 by another eminent Central European, Vienna's Friedrich ("Fritz") Redlich.

In 1911, Arnold Gesell created the Child Study Center at Yale, later associated with several departments including psychiatry.

ELSEWHERE

Chicago. Chicago had a long tradition of contributions to psychiatry. In 1925, physiologist Nathaniel Kleitman organized a sleep laboratory at the University of Chicago, and from this lab emerged in 1953 a classic paper in Science co-written by Kleitman's student Eugene Aserinsky on rapid eye movement (REM) sleep. Neurologist and psychiatrist Percival Bailey had cofounded in 1946 the Society of Biological Psychiatry. And Laszlo von Meduna, the originator of convulsive therapy, arrived in Chicago from Budapest in 1939 and ultimately found a berth at the Neuropsychiatric Institute of the University of Illinois.

Yet the founding figure in psychiatry is Roy Richard Grinker, who had a background both in psychoanalysis (worked with Freud in Vienna in the early 1930s) and in neurology. Grinker's career at the University of Chicago began in 1933, with his appointment as associate professor of neurology; he became associate professor of psychiatry in 1936, which established psychiatry within the department of medicine (Grinker's appointment at Michael Reese Hospital in 1936 is recalled as "taking over the department of neurology"). In 1955, psychiatry changed to a department in its own right, its first chair C Knight Aldrich. Although psychodynamics was then the watchword in Chicago, the University was not hesitant to appoint in 1966 the elfin Daniel X Freedman from Yale as the next chair. Freedman, an enthusiastic promoter of biological psychiatry, became chief editor of the AMA's Archives of General Psychiatry in 1970. He was succeeded as chair by Stuart C Yudofsky in 1988, and Elliot S Gershon in 1998. Steven

Paul, who later led neuroscience research at Eli Lilly, recalled in a 2011 interview his years in the mid-1970s as a resident in psychiatry at the University of Chicago, "It was a very exciting department in those days. Herb Meltzer [biological psychiatrist] was there and Heinz Kohut, the analyst, Bob Schuster [Charles R Schuster, who founded the University of Chicago's Drug Abuse Research Center], and many others. A tremendous department Danny had pulled together, a small but extraordinarily fine department." Meanwhile, in 1974 Philip Holzman, professor of psychology at the University of Chicago, reported a major biological abnormality in schizophrenia, saccadic eye-tracking movements (smooth-pursuit in tracking is normal).

University of North Carolina. The Department of Psychiatry at the University of North Carolina (UNC) at Chapel Hill was founded by Chicago psychoanalyst George Ham in 1951. It started to become a biological piston with the arrival in 1954 of Scotsman John Ewing as clinical instructor, then chair in 1962. Psychiatric endocrinologist Morris Lipton from the University of Chicago joined the department in 1959 and proceeded to steer it toward the endocrinological hormones that science was just discovering. (Lipton became known as having "opened up the peptides.") Arthur Prange, who had trained at UNC, joined the staff in 1960 and proceeded to solidify UNC's growing association with psychiatric endocrinology. William McKinney (later at the University of Wisconsin) trained at UNC in the mid-1960s, and in his class were Frederick Goodwin (later an NIMH star), the Englishman Peter Whybrow (later chair of psychiatry at the UCLA Neuropsychiatric Institute), and Joseph Mendels (who ended up leading biological psychiatry at the University of Pennsylvania).

Duke University. Across the way from UNC lay Duke University in Durham. In 1953, Ewald Busse, psychoanalyst and geriatrician, became the founding chair of psychiatry at Duke. His successors H Keith Brodie (1974 to 1982) and Bernard Carroll (1983 to 1990) ended this Freudian spin (in 1968 Australian-born Carroll introduced the dexamethasone suppression test [DST] into psychiatry for the diagnosis of melancholic depression). Duke and UNC together constituted a powerful core of psychiatric research for the southeast.

University of Pennsylvania. Psychiatry at the University of Pennsylvania had begun in 1793, when Benjamin Rush, a professor of chemistry, opened a ward for insane patients at Pennsylvania Hospital. But academic psychiatry at Penn really began only in 1874, with the establishment of the Hospital of the University of Pennsylvania. A department of mental diseases received its inaugural head in 1883, and in 1901, Charles Burr was appointed Penn's first chair of mental diseases.

Karl Rickels founds the Clinical Psychopharmacology Unit in 1959. And

the Center for Mental Health Policy and Services Research (CMHPSR) was established in 1986.

The "Psychopathic Hospitals." Emblematic of the new association between asylum and academy were the so-called "psychopathic hospitals," which began life in the 1880s, as reception wards for newly admitted, or "acute," patients. American psychiatrists visiting Europe had learned much about humane care, and the usefulness of setting these acutely ill patients aside from the chronic patients working in the farm colonies or vegetating in the backwards. (Indeed, the spotlight began to turn toward creating psychiatric wards in general hospitals, such as "Pavilion F," founded in 1902, of Albany Hospital.)

In 1901, the Michigan legislature appropriated a sum for the establishment of a "psychopathic hospital" on the grounds of the State University of Michigan, as an additional ward of the University Hospital. It was not part of the state asylum system and entirely under the control of the university. The psychopathic hospital did not in fact open until 1906, under the direction of neuropathologist Albert M Barrett.

In Massachusetts in 1912, a psychopathic hospital opened on Fenwood Road in Boston as a department of the Boston State Hospital in Dorchester; it was under Elmer E Southard, a professor of neuropathology at Harvard. The plan was, as Henry Hurd in his great 1916 work *The Institutional Care of the Insane* put it, to establish "a hospital for the first care and observation of mental patients and the treatment of acute and curable mental diseases and for an out-patient department ... for scientific research as to the nature, causes and results of insanity." The patients were given laboratory tests, stabilized with hydrotherapy, and then either discharged as recovered or forwarded to the main wards of the state hospital. In 1920, the institution was renamed Boston Psychopathic Hospital, under C Macfie Campbell. In 1956, the name was changed again to Massachusetts Mental Health Center. Several prominent figures in US psychiatry served there, including Abraham Myerson.

Fourteen further psychopathic hospitals were founded by the First World War, although their missions ranged from the conduct of research, as at the University of Michigan and Harvard's Boston Psychopathic, to the mere reception and detention of acute patients. (On the Iowa Psychopathic Hospital, opened in 1921, see below.)

MISSISSIPPI VALLEY PSYCHIATRY. A string of important psychiatry departments ran down the Mississippi valley and played a crucial role in bringing the new biological psychiatry to the United States.

University of Iowa. Academic psychiatry began in Iowa City with the opening of the Iowa State Psychopathic Hospital in 1921 under the state board of education. The director was simultaneously chair of psychiatry at the

university, and the Iowa Psychopathic is considered the first university-affiliated psychiatric department west of the Mississippi. Under founding director Samuel T. Orton in 1919, a neuropathologist who previously had been clinical director of the Worcester State Hospital in Massachusetts, the department rose to become one of the powerhouses of US psychiatry. Jacques Gottlieb, whose background was in neuroendocrinology and who established the effectiveness of the amphetamines in the treatment of depression, taught at the Psychopathic Hospital from 1937 to 1940, then returned to Iowa as professor of psychiatry from 1947 to 1953, before moving on. Erich Lindemann came from Heidelberg to train at Iowa in 1927, and was on staff until 1935, before accepting a post at Harvard. While at Iowa, Lindemann was among the first to investigate the effects of psychopharmaceuticals. (Lindemann is also remembered for helping pioneer the diagnosis of traumatic stress with his 1944 study of "acute grief" in the survivors of the Coconut Grove Fire in 1942 in Boston.) Establishing a strong bridge between Iowa and Washington University in St Louis, George Winokur came from Wash U in 1971 to chair psychiatry at Iowa. William Coryell followed in 1977. Ming Tsuang trained in psychiatry at Iowa, then served as professor from 1975 to 1982. In 1996, Winokur and Tsuang (who by now had become the Stanley Cobb Professor of Psychiatry at Harvard) published The Natural History of Mania, Depression, and Schizophrenia, based on the research panel "the Iowa 500." In 1971, Nancy Andreasen completed her training at Iowa, then while on staff went on to pioneer the neuroimaging of schizophrenia.

Washington University in St. Louis. Psychiatry at "Wash U" began in 1918, under Sidney Schwab, who headed the division of neurology and psychiatry, a "subdepartment" of the Department of Medicine. In 1938, the Department of Neuropsychiatry became an independent unit under David M Rioch, who passed the torch to Edwin Gildea in 1942. In 1948, Harvard-trained Eli Robins became a fellow at Wash U, joining the department in 1951. Meanwhile, in 1950 Samuel Guze joined the department, and for the next 20 years the names Robins and Guze would be closely associated with the growth of psychiatry at Wash U as it became the premier department of biological psychiatry in the United States. In 1970, they suggested the "medical model" as an alternative to the "biopsychosocial model," as a motor for delineating disease.

Wash U was also one of the two seedbeds (Columbia the other) of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM). Recalled Paula Clayton in an interview, who at the time was a resident, "We all read Kraepelin. He was our bible." Under Eli Robins (chair 1963 to 1975) and Samuel Guze ("Goo-zay," chair 1975 to 1989), psychiatry at Wash U opened up a series of links between basic biology and clinical psychiatry. "[It] was nearer to being a mini-NIMH,"

said geneticist Irving Gottesman later, "where under one roof you could have MDs and PhDs of all sorts working together on projects." Graduates of Wash U assumed important leadership roles throughout US psychiatry. It was the consummate insiders' group and displaced the previous insiders' group, the psychoanalytically oriented Group for the Advancement of Psychiatry (GAP).

TENNESSEE. Frank H. Luton became the first psychiatrist at Vanderbilt in Nashville, Tennessee, when he returned in 1931, after training at Johns Hopkins under Adolf Meyer. The School of Medicine created a department of psychiatry in 1947 with the appointment of William F Orr, Jr, who in turn was succeeded by Marc Hollender in 1969, and Michael H Ebert in 1983. Interest at Vanderbilt passed from dynamic psychiatry to psychopharmacology in the 1970s, with the appointment of such investigators as Fridolin Sulser, Thomas Ban, and Herbert Meltzer.

Meanwhile, in Memphis Hagop Akiskal, a graduate in 1969 of the medical school of the American University of Beirut, had organized in 1973 an affective disease clinic at the University of Tennessee, one of the first in the United States. From 1987, Akiskal propagated the idea of a "soft bipolar spectrum," heavily dependent on temperament and managed with mood stabilizers, rather than conventional "antidepressants."

TULANE. Psychiatry at Tulane in New Orleans was founded by Robert G Heath after he arrived from Columbia in 1949. In New York, Heath had trained in both neurology and psychoanalysis, and established a psychoanalysis program at Tulane. But his real interest lay in the possible metabolic origins of schizophrenia. The charismatic Heath, whom Time magazine called "the Gregory Peck of psychiatry," set out to investigate intracerebral events by persuading doubtful colleagues of the wisdom of inserting indwelling electrodes. This controversial program came to nothing, but Health had several important students. In 1960, one of them, Donald M Gallant, won for the department of psychiatry a federal grant from the ECDEU of the National Institute of Mental Health, a program that aided the introduction of psychopharmacology into American psychiatry (see above). Gallant became chair in 1969 following Heath, and expanded the entire range of research in psychopharmacology, so that it became the flagship program of the Deep South in this area.

THE WEST COAST. Several departments of international distinction on the West Coast were founded—much later than those in other parts of the country. Yet they are part of the story of the transition from asylum to university psychiatric hospital.

Stanford University, Palo Alto. Psychiatry at Stanford was founded in 1933, when George S. Johnson came from his post at the University of Colorado

to become chair of neuropsychiatry. In 1954, neurology split from psychiatry and Johnson continued as chair of psychiatry. Child and adolescent psychiatry, which was established in 1939, merged with the department in 1957. In 1960, Stanford University School of Medicine moved from San Francisco to Palo Alto, where it quickly became a research-oriented medical school. David Hamburg came from the NIMH in 1961 to chair the department, and he recruited a number of future leaders, such as William Dement, who founded sleep medicine as a clinical discipline as well as a research program; Jack Barchas, remembered for early work on the enkephalins and endorphins; and Helena Kraemer, who developed statistical research methods.

The department went on to become one of the main forces in US psychiatry. Alan Schatzberg, a leader in psychopharmacology and coauthor with Charles Nemeroff of Textbook of Psychopharmacology (1998), came from McLean Hospital to serve as chair from 1991 to 2010. His own work centered on the adverse effects of hypercortisolemia and the development of glucocorticoid antagonists as treatments for psychosis. Laura Roberts, a leader in academic psychiatry and bioethics, was recruited from the Medical College of Wisconsin to serve as chair beginning in 2010.

Langley Porter Clinic San Francisco. In San Francisco, the Langley Porter Clinic, named after Dean Robert Langley Porter who conceived it in the late 1930s, opened its doors in 1942, as a 100-bed psychiatric hospital. It later changed its name to the "Langley Porter Neuropsychiatric Institute," and then again to "Langley Porter Psychiatric Institute." Writes historian Jurgen Ruesch in 1978, "It was based upon the idea that prevention, teaching, and research … would be enhanced if the wealth of clinical experience gained from the approximately 29,000 patients then available in the nine California state hospitals could be combined with the expertise of the staff of a medical school." The inaugural director was Carl M. Bowman, who simultaneously became the first chair of the new division of psychiatry in the School of Medicine. Bowman was among the charter members in 1946, of the Society of Biological Psychiatry. In 1973, the hospital was transferred from the jurisdiction of the state to the University of California (as was the Neuropsychiatric Institute in Los Angeles [see below]).

Neuropsychiatric Institute Los Angeles. The UCLA was opened only in 1947, and Stafford Warren, the inaugural Dean, was said to have "dreamed of creating a 'Langley Porter of the South.'" Norman Brill was brought from Georgetown in 1953 to become the founding chair of the Psychiatry Department, and in 1961, the first patients were admitted to the new Neuropsychiatric Institute (NPI) funded by the state government of California. A nearby Brain Research Institute also opened in 1961. Louis J. ("Jolly") West came from the University of Oklahoma to chair the young psychiatry department, and West built staff from all around the world. In

1959, John French became the founding director of the Brain Research Institute. In 1973, the California State Department of Mental Hygiene transferred the NPI to the University of California, where it became part of the UCLA Health Sciences Center.

PART V: DIAGNOSIS

The two great arms of clinical psychiatry are diagnosis and treatment. The following sections are devoted to the histories of both.

Modern psychiatric diagnosis really begins with Philippe Pinel in Paris (see above), who in the second edition of his textbook in 1809 enumerated four psychiatric diseases: (1) "mania, or general psychosis [délire]"; (2) "melancholia, or a specific delusion"; (3) "dementia, or abolition of thought"; (4) "idiotism, or the obliteration of the intellectual and affective faculties." What he meant by these would parse into a number of distinct modern illnesses, but the point is that he did recognize several specific forms of previously undifferentiated "madness."

After Pinel and Esquirol (see above), the most important figure in the history of nosology is Emil Kraepelin ("Krep-uh-leen"), professor of psychiatry in Heidelberg from 1891, in Munich from 1903, and author in 1883 of the influential psychiatric textbook, Psychiatry: a Textbook for Students and Physicians (*Psychiatrie: ein Lehrbuch für Studierende und Aerzte*), which reached international influence with its fourth edition in 1893 and became worldwide the dominant voice in psychiatry by the eighth edition, completed in 1915. Kraepelin had little patience with psychiatric theory or with pathological anatomy, and placed the diagnostic interest in psychiatry squarely on the clinical side, meaning at the bedside, especially on the course and outcome of the patient's illness.

Kraepelin was the clinician's clinician. On the basis of clinical course, Kraepelin said in 1899 that the two specific diseases in psychiatry were "manic-depressive insanity," a mood disorder, and "dementia praecox," or the dementia of adolescence, a thought disorder later called "schizophrenia." He erected a firewall between them, one that still exists today, as in the DSM system the cardinal disorders are "major depression," as opposed to "schizophrenia." No other historical figure in psychiatry, including Freud, has had the influence of Kraepelin upon contemporary psychiatry, although under the influence of psychoanalysis, US interest in diagnosis did wane for several decades. Henri Ellenberger, a Swiss psychiatrist transplanted to Montreal, recalled in 1955, "I remember the consternation of a German psychiatrist hearing that certain American colleagues consider manic-depressive psychosis as a form of schizophrenia; for him this was about as fantastic as the assumption that a camel is a subspecies of the elephant."

A page was turned with the advent of a distinctively American system of classification in 1945, the "Nomenclature of Psychiatric Disorders and

Reactions," which appeared as War Department Technical Bulletin, "Medical 203." Spearheaded by psychoanalyst William Menninger, the head of consultation psychiatry in the US Army during the war, "Medical 203" made wide use of psychoanalytic insights and underpinned the first effort of the American Psychiatric Association to classify illness, the *Diagnostic and Statistical Manual,* or "DSM-I" (so-called subsequently) in 1952. (The APA itself was founded in 1844 as The Association of Medical Superintendents of American Institutions for the Insane, becoming the American Medico-Psychological Association in 1894, and the American Psychiatric Association in 1921.)

An even more psychoanalytically oriented volume, DSM-II, emerged in 1968.

In the 1970s, the APA drew up a new task force to make what were intended to be minor changes in DSM to bring it into accordance with the *International Classification of Diseases* (ICD) of the World Health Organization, appointing Robert Spitzer at the New York State Psychiatric Institute to lead it. Spitzer, however, disliked psychoanalysis and wanted a complete revision of the nosology to remove Freudian influences. In 1980 the revolutionary DSM-III emerged, which created such diagnoses as "major depression" and "attention deficit disorder," and anchored "bipolar disorder" firmly in the American classification. The subsequent editions of DSM—DSM-III-R in 1987, DSM-IV in 1994, and DSM-5 in 2013—changed little of this architecture, though in launching "bipolar two disorder" in 1994, DSM-IV did touch off an epidemic of "bipolar" in the pediatric population.

A rival system was the ICD of the World Health Organization. In 1938, an "international revision" conference of the League of Nations decided to add morbidity to its long-standing classification of mortality, and in 1948 the newly founded World Health Organization issued the "sixth revision" of the international lists (ICD-6). It contained such diagnoses as "delusional" insanity (other major psychiatric illnesses were catalogued under "dementia"). A distinct section on "mental disorders" was added in the 9th edition in 1975.

Over the years, the number of diagnoses in psychiatry has skyrocketed from four in Pinel to some three hundred in DSM-5. Some of these diagnoses may not stand the test of time. Yet, on a historical basis, several diagnoses do stand out as enduring, and doubtless represent phenotypes of genuine underlying psychiatric diseases.

Specific Diagnoses

Melancholia. Melancholia, as distinct from generic "depression," is one of the oldest diagnoses in medicine and goes back to the Ancients. Before the 19th century, melancholia tended to be an umbrella term for several varieties of psychosis and mood disorder that today would be

recognized as distinct. Yet at its core, melancholia has always stood for severe depressive mood disorder involving deep sadness, pervasive anhedonia, and psychomotor change, usually in the direction of slowing but there is an anxious form of melancholia involving psychomotor acceleration as well. Timothie Bright, a Cambridge physician, in his 1586 book *A Treatise of Melancholie* described,

> the sorrowful humor which breeds in them ... Of pace, they are for the most part slow ... both by their members not so nimble for motion, and the mind occupied with cogitation stays [slows] the pace ... Moreover, they are given to weeping sometimes ... sighing, sobbing, lamentation, countenance demisse [downcast] and lowering." Bright left a poignant image of weeping in melancholia, "It blears the eyes and makes the cheeks red; it causes the head to ache, the nose to run, and the mouth to flauer [labor], the lips to tremble; it interrupts the speech and shakes the whole chest with sighs and sobs.

In the 19th century, melancholia became differentiated from other kinds of depressive illness. Among the first to make this distinction was Richard von Krafft-Ebing in 1867, at the time a staff psychiatrist at the Illenau asylum (later, professor of psychiatry in Vienna). Krafft-Ebing distinguished between "simple depression of mood" (die einfache Gemüthsdepression") and severe depressions with anxiety and psychosis, including melancholia. The definitive distinction between "melancholia" and "depression" was ventured by Copenhagen psychiatrist Carl Georg Lange in his book on "periodic depressions" in 1886 (translated into German in 1892), which described these milder cases as distinct from (psychotic) melancholia.

The term "melancholia" went out of style around 1900 as Emil Kraepelin substituted "depression" for it, and Sigmund Freud decided that the depressive disorders did not qualify as the main psychoneuroses that psychoanalysis should be addressing. (But see above on a United States interested in "depressive neurosis.") In recent decades, there has been a revival of interest in melancholia, as the field has begun to differentiate the heterogeneous basin of "major depression" on the basis of treatment response.

Psychotic Depression. In classical psychiatry, much melancholia was psychotic, and little was made of the distinction between psychotic and nonpsychotic. The difference later seemed unimportant because both groups responded equally well to ECT, introduced in 1938. Yet with the growing stigmatization of ECT after the 1960s, depression became treatable primarily with drugs, and in 1975, Alexander (Sandy) Glassman at the New York State Psychiatric Institute discovered that psychotic depression responded poorly to the standard tricyclic antidepressants (TCAs). Four years later, David Avery and Aldo Lubrano at the Palo Alto Veterans Hospital reported that psychotically depressed patients responded better to ECT than did nonpsychotically depressed. Under the leadership of Max

Fink at the Stony Brook Campus of the State University of New York, ECT became defined as the most effective treatment of psychotic depression. Despite this evidence of differential responsiveness, psychotic depression continued to be downgraded as a mere "specifier" in the DSM.

Mania. Mania has undergone several transformations. For Pinel and the generations before him, it meant agitated dementia ("raving madness"). With François Fodéré in Strasbourg (1817), unbridled unreasoning outbursts became "furious mania." In 1858, London psychiatrist Daniel Hack Tuke steered mania in the direction of an affective disorder rather than a disorder of intellect. "Although mania, in many instances, is prolonged anger, it may, likewise, be altogether pleasurable in its manifestations, presenting a condition of exhilaration and uncontrollable excitement, which is rather an excess of joy than of anger." Then in 1881, Emanuel Ernst Mendel, senior lecturer in Berlin, anchored the term "hypomania" in the nosology, meaning elated mood, pressured speech, and increased motor activity. Mania entered DSM intact in 1952, and hypomania became official in 1994 in DSM-IV as "bipolar 2 disorder."

Catatonia. Although individual symptoms of catatonia, such as waxy flexibility, had been described for centuries, catatonia as a clinical syndrome was recognized by Karl Kahlbaum only in 1874. It independent existence was short, as in 1899, in the 6th edition of his Textbook, Kraepelin made it a "subtype" of schizophrenia, and for the next century catatonia languished as an independent entity because, often when clinicians encountered it, they realized they were not dealing with schizophrenia and called it something else, or ignored it. In 1930, William Bleckwenn at the University of Wisconsin reported that catatonia responded to barbiturates, though this discovery was not widely appreciated. Karl Heinz Stauder, at the Munich University Psychiatric Hospital, described in 1934 a fatal version of catatonia, associated with hyperpyrexia.

The return of catatonia as an independent diagnosis separate from schizophrenia began in the 1970s. In 1976, Richard Abrams and Michael Alan Taylor at New York Medical College identified 20 percent of a sample of patients admitted to the psychiatric unit of a New York City hospital as having one more catatonia signs. Only 4 of the 55 satisfied the criteria for schizophrenia, while more than two-thirds had affective disorders, usually mania. In the context of growing interest in catatonia as independent of schizophrenia, in 1983, Gregory Fricchione at Stony Brook recommended benzodiazepines, especially lorazepam, in the treatment of catatonia instead of barbiturates. Various syndromes previously seen as separate began to be recognized as catatonic in nature: In 1996, it was suggested that toxic serotonin syndrome (TSS) might be a form of catatonia. A

symposium at the American Psychiatric Association meeting in 2000 concluded that neuroleptic malignant syndrome (NMS) and malignant catatonia were the same. In 2011, Dirk Dhossche at the University of Mississippi and colleagues raised the possibility that anti-NMDA receptor encephalitis was a form of catatonia. The American Psychiatric Association responded to this series of new catatonia findings by making catatonia a virtually independent disease entity in the DSM-5 published in 2013.

Delusional Disorders

Delusional disorders, meaning fixed, systematic false beliefs, were recognized in 1816 with Etienne Esquirol's "monomania." He said that the psychosis (délire) is fixed on one object alone. As early as 1865, then definitively in 1874, Ludwig DC Snell, director of the asylum in Hildesheim, Germany, distinguished between primary and secondary delusions (Wahn), primary meaning fixed, false ideas not caused by another illness and not deteriorating; vague secondary delusions could lead to "the collapse of the entire psychic life," something Snell observed in the kind of illness later called "schizophrenia." This was the beginning of separating delusional disorders from chronic deteriorating psychosis as the former were false beliefs that were fixed, systematic, and did not affect other areas of intellectual functioning.

Anxiety. The term "anxiety" has long been used in psychiatry. In 1769, a young medical graduate of Leiden University ventured a dissertation on "Anxietate," and the term was clearly current among his professors. Yet the German clinical tradition of the 19th century frowned on anxiety as an independent diagnosis. In Emil Kraepelin's big textbook, there is no freestanding "anxiety" disorder. But it was not just Kraepelinian obdurance. The "anxiety" entry in Daniel Hack Tuke's Dictionary of Psychological Medicine (1892) is only two sentences long. What placed anxiety in the center of psychiatry's radar was Freud's belief that anxiety was the basic motor of symptom formation: anxiety arose as a result of intrapsychic conflict over various drives. The entries for "anxiety" (Angst) in the general index of Freud's writings extended for eleven pages ("genitals," by contrast, only for four). Accordingly, anxiety has always assumed a disproportionate role in psychoanalytically oriented American psychiatry (mixed anxiety-depression has tended to loom more prominently for the British). DSM-I in 1952 prefaced its discussion of "psychoneurotic disorders" by claiming that, "The chief characteristic of these disorders is 'anxiety'...." and in the discussion that followed "anxiety" figured importantly in each, such as "dissociative reaction," "obsessive compulsive reaction," and so forth. DSM-III in 1980, although antipsychoanalytic in inspiration, continued the centrality of anxiety while beginning its fragmentation into specific anxiety disorders, of which there

were ten. Subsequent editions of DSM have continued this tradition.

Chronic Primary Psychosis (Schizophrenia). Psychiatry has struggled over the years with chronic deteriorating psychotic illness, leading either to some kind of deficit or to the dissolution of the personality (once called "dementia"). Developments in psychiatry's two major cradles are chronicled separately.

France. In 1852, Bénédict-Augustin Morel, then at the Maréville asylum in eastern France (but shortly to head the St-Yon asylum outside of Rouen) called such psychoses of adolescent onset "démence précoce," or premature dementia (see above). In 1857, he applied the notion of "degeneration," or worsening across the generations, to psychiatric illness (Traité des Dégérérescences) and 1860, in one of the most famous declarations in the history of psychiatry, in his Treatise on Mental Illness (Traité des Maladies Mentales) he avowed premature dementia in the young to be a consequence of degeneration.

In 1883, Valentin Magnan, a staff psychiatrist at the Sainte-Anne Clinical Asylum in Paris, described "chronic psychosis,"—he later added the qualifier "involving systematic evolution"—an often-deteriorating form of psychosis in patients who were not hereditarily predisposed. This remained quite influential in France, though elsewhere it was eclipsed by "schizophrenia" (see below).

French experimentation with chronic psychosis continued, even as Magnan's degeneration-based doctrines were discredited, and in 1910, Maurice Dide, at a psychiatric hospital in Toulouse, proposed "chronic hallucinatory psychosis," which, though severe, did not deteriorate. To this day, the concept has remained influential in France without reaching a larger international audience.

Germany. Meanwhile, in Germany a different approach to chronic psychosis was being elaborated: deterioration to "dementia," or at least to a severe deficit. In an 1863 book on the classification of psychiatric illness, Karl Kahlbaum, a staff psychiatrist at the East Prussian provincial asylum of Allenberg, differentiated a "deteriorating psychosis," which he called Vesania typica, which ran its course in four stages, the last of which was "dementia." Vesania was a customary term at the time for chronic psychosis. In 1871, Kahlbaum and his student Ewald Hecker differentiated further a version of this Vesania beginning in adolescence that they called "hebephrenia." Hebephrenia, like Vesania typica, usually ended in dementia. (Although Hecker's name appeared alone on the paper, it was understood that the concept was really Kahlbaum's.)

At this point, a considerable amount of European scholarship, dating from Morel in 1852, had accumulated that linked chronic deteriorating psychosis to an onset in adolescence. In 1893, Kraepelin synthesized this literature in a concept that was to sit rock-solid in psychiatry for more than

a century: “dementia praecox.” Over the next decade he identified three subforms of it: hebephrenic, catatonic, and paranoid. They had in common a begin usually in adolescence or young adulthood, and a downhill course to “dementia.” So magisterial was the Kraepelinian synthesis that all previous writing on chronic psychosis became virtually forgotten. In 1908, Eugen Bleuler, professor of psychiatry in Zurich, baptized Kraepelin’s dementia praecox as “schizophrenia,” seeing it as a much milder disease and not linked to adolescence. Bleuler’s schizophrenia remains little changed today, a single deteriorating psychotic disorder that may onset at any point in life and that may well end in “defect” (or social remission), even if “dementia” as the end stage of the illness has stopped being spoken of. In 2013, DSM-5 dispensed with the famous Kraepelinian “subtypes.”

Many “schizophrenia” patients may spend long periods without being actively psychotic; the primary symptoms seem to be emotional blunting with avolition and a formal thought disorder.

Brief Reactive Psychosis or Brief Psychotic Disorder

Chronic psychotic illness had long been seen as something indwelling and constitutional. The alternative notion that psychosis could arise from stress was codified as “la bouffée délirante” by two students of Valentin Magnan at Sainte Anne in Paris in 1886, then by Magnan himself in 1891—who saw it as arising mainly in the “degenerate.” August Wimmer at St. Hans Hospital in Denmark christened it “psychogenic psychosis” in 1916, with an emphasis on reaction rather than degeneration. Then in 1968, Eric Strömgren in Aarhus, Denmark, revived Wimmer’s concept in an influential (in Scandinavia) paper on “reactive psychoses,” a term that he changed in 1974 to “psychogenic psychoses.” In 1980, DSM-III recognized “brief reactive psychosis,” with no reference to any of these other international terms. (These developments are charted in Shorter, Historical Dictionary.) The French translation in 1989 of DSM-III-R makes no reference to “bouffée délirante” and mentions only “psychose reactionnelle brève,” an interesting example of the collapse of these once distinctive national nosological traditions.

Bipolar Disorder

The “bipolar disorder” of DSM is not Kraepelin’s “manic-depressive insanity,” which Kraepelin insisted included all depressions of whatever polarity, mania, hypomania, and the attendant anxiety. Rather, in his 1948 textbook Karl Leonhard, then in Frankfurt am Main and a student of Karl Kleist (who introduced the concept of “phasic psychoses,” or illnesses that alternate from one state to another), proposed the term “bipolarity” for the alternation of depression and mania as an illness separate from unipolar depression. In 1957, in The Classification of the Endogenous Psychoses

(Die Einteilung der endogenen Psychosen), Leonhard asserted that unipolar depression had a different psychopathology from the bipolar variety. Filled with neologisms and lacking in quantitation, the book had a limited international impact. Yet a small cadre of investigators, such as Frank Fish in England, Helmut Beckmann in Germany, and George Winokur in the United States, became enthusiasts of the Leonhardian notion of bipolar disorder and managed to insert it in the literature, so that by 1980, when DSM-III was published, bipolar disorder was self-understood as a different disease from unipolar depression (in DSM-III called major depression).

Trauma and Its Sequelae

Ever since Napoleon's disastrous retreat from Moscow in 1812, psychiatry has recognized that symptoms may arise from great stress. (In Napoleon's case, of an army of 400,000 that set out, 7000 returned. Along the way, much "nostalgia" was witnessed—an early label for fatal catatonia.) In 1866, London surgeon John Eric Erichsen originated the diagnosis of "railway spine," vague and nonspecific symptoms supposedly attributable to the "spine" in minor railway mishaps. Carol Moeli, a Berlin psychiatrist, elaborated in 1881, upon the specifically psychiatric consequences of such events, penning the first description of what became known as post-traumatic stress disorder (PTSD). In 1881 as well, Berlin neurologist Hermann Oppenheim coined the phrase "traumatic neurosis," psychiatric and neurological symptoms arising from shock.

The First World War spawned a whole raft of descriptions of psychological sequelae from combat, of which terms the most lasting has been "shell shock"—from Charles S Myers in 1915. The Second World War gave birth to "combat fatigue" as a synonym for war neurosis. Meanwhile, a tragic fire in 1942, at the Coconut Grove night club in Boston gave rise in 1944 to "acute grief" as yet another term for acute stress reaction. Most recently, following a sustained political campaign by the veterans of the Vietnam War—and spearheaded by Yale psychiatrist Rober Lifton—DSM-III in 1980 coined the term "posttraumatic stress disorder." This diagnosis has been among the most rapidly proliferating in US psychiatry.

Dementia

The dementias have always been a problem for psychiatry, as the field has struggled to differentiate loss of intellectual function from personality deterioration. In 1814, Etienne Esquirol took up the idea of his teacher Pinel of dementia as a distinct disease entity, separate from melancholy, mania, and imbecility. "There exists a very distinctive kind of insanity in which the disorder of ideas, sentiments, and determinations is characterized by the weakness and the more or less pronounced abolition

of all the sensory, intellectual and moral faculties." In particular, Esquirol differentiated senile dementia from other conditions of "weakmindedness." (Shorter, Historical Dictionary of Psychiatry, chronicles developments in dementia.) There followed a long series of differentiations of the organic dementias: Sergei Korsakoff distinguished ethanol-related amnesia (Korsakoff's psychosis) from dementia in 1887; Alois Alzheimer distinguished "presenile" dementia from the senile variety in 1906; in 1926, Hugo Spatz, head of neuropathology at the Munich University Psychiatric Clinic, suggested that a distinctive frontal- and temporal-lobe wasting should be called "Pick's Disease" after Arnold Pick, professor of psychiatry at the German university in Prague, who first reported it in 1892; in 1912, Friedrich Heinrich Lewy described distinctive intracytoplasmic inclusions, later called "Lewy bodies," which he detected in the brain stem in Parkinsonism; and in 1962, neuropathologist John S. Woodard at the Camarillo State Hospital in California, proposed Lewy-body dementia as a separate form of dementia; in 1948, London psychiatrist Robert Denis Newton codified long-established German thinking that presenile Alzheimer's and the dementia of senescence represented a single pathological process, and suggested calling it "Alzheimer's dementia"; Raymond Adams at Harvard University and Maurice Victor at Case Western University, emphasized in 1977, that Alzheimer disease was not just an accelerated extension of old age but a pathological eruption into it.

Ever since New Jersey pathologist Harrison Stanford Martland's paper "Punch Drunk" in 1928, repeated cranial trauma has been considered a distinct cause of dementia. Martland said, "Fighters in whom the early symptoms are well recognized are said by the fans to be 'cuckoo,' 'goofy,' 'cutting paper dolls,' or 'slug nutty.'" Since then, physicians have been alerted to the neuropsychiatric consequences of repeated head trauma. In 1937, in the US Naval Medical Bulletin, naval surgeon JA Millspaugh called it "Dementia Pugilistica." In a paper in the British Medical Journal in 1989, on "Boxing and the brain," John Arthur Nicholas Corsellis pulled this literature together in a red warning-light about repeated head trauma. By the new millennium these concerns about early dementia had been expanded to all sports involving repeated head trauma.

Neurosis

The classic distinction between psychosis and neurosis was abolished in DSM-III in 1980. Neurosis as a term is among the oldest in psychiatry, introduced by Edinburgh physician and nosologist William Cullen in his First Lines in the Practice of Physic in 1777, to mean diseases of the nervous system without an obvious physical lesion. In the next half century, it came to be synonymous with brain disorder. It was Richard von Krafft-Ebing in his Textbook of Psychiatry (Lehrbuch der Psychiatrie) in 1872, who coined the term "psychoneurosis," meaning mental illnesses that affect normal

brains (in contrast to brains affected with "degeneration"). Sigmund Freud used the term psychoneurosis for the first time in his writing in 1898. But neurosis had by that time long entered medicine and Freud proceeded to employ it as a synonym for "disorder." After Freud, neurosis remained largely in the area of psychoanalytic discourse, which was the dominant discourse in psychiatry until DSM-III in 1980.

PART VI: NEW TECHNIQUES

Psychiatry is not among the most technological of medical specialties. Yet changes in technique have expanded the range of its knowledge and enhanced the diagnostic and therapeutic abilities of its practitioners.

Epidemiology

The study of epidemiology had an early bump-up with the observation of Thomas Clouston in 1873, in the *Journal of Mental Science* that, "The wide series of diseases which are at present called insanity prevail more in some places than in others ... When one comes to look carefully into the reports of lunatic asylums in different parts of the country it is found that there are forms of brain disease (or varieties of insanity as they are called) present in abundance in one place which have almost no existence in another ... This investigation must be very largely conducted on the numerical method."

Yet it was the Germans who initially picked up this torch. At the German Psychiatric Research Institute (Deutsche Forschungsanstalt) in Munich, founded in 1917, by Emil Kraepelin, Swiss psychiatrist and geneticist Ernst Rüdin, inaugural director of the eugenics research institute of the DFA, led studies on psychiatric epidemiology (cross-sectional) and genetics (longitudinal). Unfortunately, in the Third Reich Rüdin had a particularly unsavory involvement with public policy, arguing for the Nazi racial and sterilization laws, and after 1945 was briefly interned in the United States. This National Socialist involvement of the DFA and other German groups lent to psychiatric epidemiology and genetics a bad odor that retarded the development of the field for decades.

It was in Denmark that Eric Strömgren, the most important psychiatric epidemiologist of the pre-1970 period, flourished. As a student, he visited the Rüdin group in 1935; he undertook in 1938 what became an epidemiologic classic of the small island of Bornholm. He became professor of psychiatry in 1945, and was instrumental in organizing the Danish Psychiatric Central Registry at Aarhus, which in 1969 became an electronic register of everyone in Denmark admitted to a psychiatric service. By 2011, this famous register contained over 700,000 individuals and more than 3 million treatment episodes. One offshoot of this was the Danish-American cohort study in 1968, led by David Rosenthal, Paul Wender, and Seymour Kety at NIMH. It merged two registers: that of children who had been

adopted away and that of admissions to psychiatric hospitals. This remarkable data base permitted the NIMH collaborators, together with Fini Schulsinger in Copenhagen and other Danish colleagues, to determine that children adopted from schizophrenia families but raised in normal homes often became schizophrenic. It was a major blow to the "nurture" school). At a conference in Dorado, Puerto Rico, where these results were first presented, Rosenthal said, "The case for heredity has held up convincingly. At the moment, I can see no grounds for withholding affirmation any longer."

In the United States, the National Institute of Mental Health sponsored in the 1980s, the 5-site Epidemiological Catchment Area (ECA) study, the first to inquire about diseases rather than symptoms, the results of which were reported in 1991. Myrna Weisman at Columbia later said in an interview the big finding of the ECA was that, "These disorders begin in the young ... often in adolescence, and certainly, by young adulthood ... So, the biggest contribution of the ECA was turning the focus on young people and their high rates of psychiatric disorders."

In the early 1990s, NIMH sponsored the National Comorbidity Survey (NCS) to assess the prevalence of DSM-III-R disorders in a nationally representative sample. Early in the new century, the NCS was redone in the National Comorbidity Survey Replication (NCS-R), using the international WHO Composite International Diagnostic Interview, to collect DSM-IV diagnoses from a household survey, and to make international comparisons possible.

Statistics

Paris psychiatrist Jean-Pierre Falret's critique of the counting method in 1864, in the preface of his On Mental Illness (Des maladies mentales), has not lost its relevance today: "Statistics as a procedure are in fact all the more dangerous because they offer all the appearance of truth." Falret scorned the statistics of the day as failing to include homogeneous clinical populations and using abstractions overly divorced from reality. Modern statistical theory, as Bernard Carroll traces events, originated in the last quarter of the 19th century and the first quarter of the 20th. Regression to the mean was described in 1877; regression and correlation analysis in 1888 by Francis Galton; the Chi-squared test in 1900 was originated by Karl Pearson; the student's t-test in 1908 by W.S. Gossett; the analysis of variance (ANOVA) comes from RA Fisher in 1919; power analysis, confidence intervals, Type I and Type II error originated in the 1930s from Jerzy Neyman and Egon Pearson. But these advances were not incorporated as research tools into the clinical science of psychiatry until well beyond the 1930s.

The early use of statistics in psychiatry consisted essentially in tabular bookkeeping—simple descriptive statistics such as census details, cohort

descriptions by age, age of onset, sex, or outcome. Use of statistics to test hypotheses in psychiatry came much later. Emil Kraepelin maintained detailed counting cards (Zählkarten) of cases he observed, leading up to his landmark distinction between dementia praecox and manic depressive insanity (see above). Assen Jablensky's 1993 reanalysis of those cases with modern statistical methods essentially confirmed the existence of two separate disorders that Kraepelin had recognized by informal Gestalt inferential methods. The same cannot be said of Aubrey Lewis, whose 1934 discussion of depressive illness confidently announced a single dimensional model without the benefit of any statistical analyses. That position became dogma within Maudsley Hospital circles and led to conflict with the Newcastle school under Martin Roth. Eventually, Lewis's primary data were subjected in 1977 to modern statistical analysis by Lesley Kiloh and Roger Garside (two of the original Newcastle group), and they were found actually to support the Newcastle binary model. These two examples illustrate the absence of reliable means to validate the proposals of clinical experts before the entry of statistics into psychiatric research.

Aside from elementary analytic procedures such as group comparisons (in categorical and dimensional versions), significance testing, and correlations (all in parametric and nonparametric versions), the major statistical themes that moved psychiatric research were numerical taxonomy and psychometrics. Derived from advances in basic statistics, these approaches appeared first within academic psychology for the study of personality traits and dimensions of psychopathology. For the most part, they were slow to be taken up within the discipline of psychiatry and, when they were, the history of their adoption is inseparable from the history of clinical assessment instruments that relied on the new statistical approaches. In parallel, innovations in basic statistical theory and technique also were adopted within psychiatry for clinical trials, as well as for epidemiologic and genetic studies.

Factor Analysis and Personality Inventories

Factor analysis originated with Charles Spearman and Raymond Cattell. The psychological sciences adopted the technique in the 1930s. In parallel, the Minnesota Multiphasic Personality Inventory (MMPI) was under development by Starke R. Hathaway and J. Charnley McKinley at the University of Minnesota from the late 1930s. Local use apparently began in 1939. The original published version appeared in *JAMA* in 1943: The MMPI has seen many later revisions. A perceived advantage was casting it as a descriptive and atheoretical approach to personality and psychopathology, not derived from psychoanalytic theory. Over the years, cultural and ethnic biases in the MMPI have been alleged. Nevertheless, it attained widespread use in education, clinical psychology, and employment screening. A scoring innovation for the MMPI was the devising of T-scores

—a standardized metric where the mean is set at 50 and the SD at 10. Other scales adopted the same scoring approach, which had the benefit of standardizing responses and allowing cross-scale comparisons. In parallel, Raymond B. Cattell devised in 1946, a personality inventory for which factor analysis was an essential component. His psychometric instrument is famous as the Cattell 16PF personality factors inventory.

Controlled Trials

The first controlled trial in psychiatry was reported in 1922 in the *American Journal of Psychiatry* by Clarence O Cheney, assistant director of the New York State Psychiatric Institute, then at Ward's Island, and bacteriologist Nicholas Kopeloff. They wanted to establish whether the organ removals (colectomies, teeth, etc.) then in vogue for the treatment of mental illness were in fact effective, and to do so they studied 60 patients with "focal infections" among the general admissions to the Manhattan State Hospital on Ward's Island, dividing them into two groups, the "operated group" receiving operative treatment for their " 'foci' of infection" versus those not so treated (the "control group"). They found, "On the whole, the operated group appears to have received no more benefit than the control group."

Following the tuberculosis studies of Bradford Hill after the Second World War, randomizing the patients to control or treatment groups became increasingly common. The first randomized controlled trial in psychiatry with "blinded" investigators was conducted in 1955 by David Lewis Davies and Michael Shepherd at the Maudsley Hospital, comparing reserpine to placebo in the treatment of "anxious and depressed patients."

The use of probabilistic statistics in psychiatry has a much more recent origin. Four statistics are used to measure the reduction of risk in treatment (two of them compare vs. lack of treatment). (The following discussion cribs from Andreas Laupacis and collaborators, 1988.)

- The relative risk reduction (i.e., how much risk reduction in percent). This is the classic measure.
- The "odds ratio": A probability (P) of 0.25 (one chance in four) represents odds of 1:3, or 1/3. This is insensitive to difference in risk in the absence of therapy. It was introduced by N Mantel and W Haenszel in 1959.
- Absolute risk reduction. The differences in event rates between the control and treatment groups. A genuine measure of absolute risk reduction but, expressed as a decimal, "This measure may be difficult for [physicians] to remember and incorporate into clinical practice," say Laupacis and colleagues.
- Number needed to be treated (NTT). The number of patients who must be treated in order to prevent one adverse event. Laupacis and colleagues proposed this in 1988. "It tells clinicians and patients in more concrete

terms how much effort they must expend to prevent one event, thus allowing comparisons ...” Although many clinicians will be put off by such exercises, these statistics mark the definitive passage from psychoanalysis with its evidentiary base of anecdotes to modern clinical psychiatry that employs the same evidentiary base as the rest of medicine.

Tests in Psychiatry

Despite protestations that psychiatry “has no diagnostic tests,” there are in fact quite a few. In 1906, August Wassermann and colleagues at the University of Berlin described a diagnostic test for “progressive paralysis” (what neurosyphilis was then called) based on the patient’s blood or cerebrospinal fluid. In 1913, Hideo Noguchi, a microbiologist at the Rockefeller Institute for Medical Research, succeeded in obtaining a culture of *Treponema pallidum* in a sample of the brains of neurosyphilitic patients who carried the diagnosis of “general paralysis of the insane,” the English equivalent of “progressive paralysis.” Therewith, neurosyphilis became identified as an infectious disease and passed from the purview of psychiatry, where it had lodged for the previous century. In 1931, Hans Berger in Jena used the EEG, which he had just invented, to confirm a specific pattern of brain waves in a patient with epilepsy. These were early diagnostic tests in psychiatry.

More were to come. A group of researchers led by Charles Bradley at the Emma Pendleton Bradley Home in East Providence, Rhode Island, established in 1938, that a certain subset of children who would later be called “hyperactive” (with the diagnosis attention–deficit/hyperactivity disorder, ADHD) demonstrated a specific electroencephalographic pattern. The test remains controversial but it has not been disallowed. In 1940, Mandel Cohen and Stanley Cobb, psychiatrists at Harvard, reported in The Journal of Clinical Investigation that carbon dioxide could precipitate what were later called “panic attacks” in patients prone to anxiety. Abnormalities in serum lactate thus became a biological marker for panic, and there followed a considerable history of using “panicogens” to study panic.

As early as 1958, “pharmaco-EEG” was used to “classify psychopharmacologic drugs and predict their clinical applications,” as Max Fink, a pioneer of the technique, later put it. Then in 1962, in a study published in the American Journal of Psychiatry, Fink and Donald Klein, part of the Hillside Hospital Experimental Team, established in a proper randomized controlled trial of chlorpromazine and imipramine that depressed patients responded well to imipramine, schizophrenics did not. The imipramine challenge was the first biological test for serious depression.

Bernard Carroll at the Royal Melbourne Hospital applied the DST in 1968, to confirm the presence of melancholic depression. The DST, based

on a dexamethasone challenge to the hypothalamic–pituitary–adrenal (HPA) axis, establishes HPA dysregulation in patients with severe mood disorders. In 1983, Gregory Fricchione and colleagues at Harvard reported that catatonia is responsive to the benzodiazepine lorazepam. In their 2003 book on catatonia, Max Fink and Michael Alan Taylor proposed a one or two-dose benzodiazepine challenge as a convincing diagnostic test for catatonia. Biological psychiatry proceeds from the premise that behavioral disorders are symptoms of underlying cerebral dysfunction. It is therefore unsurprising that biochemical markers have long existed in psychiatry.

Convulsive and Coma Therapy

Convulsive therapy, probably the most powerful psychiatric treatment on offer, had its origins in the 1930s, together with the other so-called "physical therapies" that addressed the brain itself.

First of the new physical treatments was insulin coma therapy, introduced in 1930 with the discovery by Manfred Sakel, then an assistant physician in a private nervous clinic in Berlin, that the withdrawal symptoms from morphine addiction could be relieved with insulin. It soon become apparent that the curative agent was an insulin-induced coma, and when, after the Nazi seizure of power in 1933, Sakel returned to his native Vienna, he persuaded the professor of psychiatry, Otto Pötzl, to let him experiment with insulin coma in the treatment of schizophrenia. From this work issued in 1934, a series of articles in a Viennese medical weekly that turned into a book in 1935, A New Method of Treating Schizophrenia (*Eine neue Behandlungsmethode der Schizophrenie*). A series of insulin comas does seem to have had a certain effectiveness in relieving the symptoms of chronic psychosis, and insulin remains today an underexplored agent in psychiatry. Some of Sakel's patients also experienced seizures, and these as well may have been a therapeutic mechanism.

It was Budapest psychiatrist Ladislaus von Meduna who in 1935 began the systematic exploration of chemically induced cerebral convulsions as a treatment. Meduna initially used camphor to initiate brain seizures, then the heart drug pentylenetetrazol. Meduna experienced extraordinary successes, especially in patients with the symptoms of catatonia, and his 1935 report in the Journal of Combined Psychiatry and Neurology (*Zeitschrift für die gesamte Neurologie und Psychiatrie*) is of historic importance in introducing the concept of therapeutic seizures to psychiatry. Yet many patients resisted chemically induced seizures because of the anxiety they experienced in waiting for the seizure to begin.

In 1938, Rome psychiatry professor Ugo Cerletti hit upon the idea of inducing the seizures with electricity, so that the patients were instantly unconscious and underwent no anxious latency period. He quickly discovered that patients with melancholic depression responded readily to ECT, and the technique spread rapidly throughout psychiatry; it was at the

time relatively uncontroversial. Cerletti had demonstrated, not that convulsive therapy was effective (Meduna's doing) but, as modern ECT-investigator Richard Abrams puts it in his 1988 manual of convulsive therapy, "that such convulsions could be induced safely, reliably, and inexpensively by electrical means." Cerletti and the other early electrotherapists did not restrain their patients, and as a result the "unmodified" seizures resulted in numerous fractures. To reduce such side effects, curare was first employed as a means of blocking the nerve–muscle junction, then, in 1952, succinylcholine, together with a barbiturate anesthetic. Yet even this modified ECT became in the 1960s an object of opprobrium in the counter-culture and then in the eyes of the pharmaceutical industry, which insisted that psychopharmacology could replace convulsive therapy. But it turned out that ECT offered distinct advantages in the treatment of melancholic depression, mania, and catatonia, and in 1978 the American Psychiatry Association offered an endorsement of it to counter the ballooning stigma caused in particular by the 1975 Hollywood movie, *One Flew Over the Cuckoo's Nest*. Max Fink, at the Stony Brook School of Medicine, led a campaign to rescue ECT from threatened oblivion. In 1990, the APA issued a much more enthusiastic endorsement of ECT, and it would be fair to say (as David Healy and I do in our 2007 history of shock therapy) that ECT has now resumed its prominent place in the panoply of psychiatric treatments.

Neuroimaging

Neuroimaging dates back to Harvey Cushing's 1896 radiograph of the cervical spine of a patient who had just been shot in the neck; the history of neuroimaging includes air encephalography, introduced in 1918, and used in Germany in 1927, to identify enlargements of the ventricular system in patients with schizophrenia. Gerd Huber's classic 1957 monograph Pneumoencephalography and Psychopathology in the Endogenous Psychoses (Pneumooencephalographische und psychopathologische Bilder bei endogenen Psychosen) was probably the high water mark of pneumoencephalography in psychiatry.

Carotid angiography, introduced in 1927, has never had significant psychiatric uses. But radionuclide brain scanning, commenced in 1948, showed great promise once radioactive ligands were introduced in the 1970s in the form of single-photon-emission tomography (SPECT) and positron emission tomography (PET).

Tom Bolwig, professor of psychiatry at the University of Copenhagen, outlines the following story: Among earliest studies, in 1974 David H Ingvar and Göran Franzen at Lund University reported in the Acta Psychiatrica Scandinavica decreased cerebral blood flow in chronic schizophrenia. SPECT and PET were employed in the 1980s to measure regional cerebral blood flow and metabolism, and, later, neurotransmitter

binding potential. SPECT and PET are used as well to quantify regional density of receptors of a given type and receptor binding displacement. But these methods expose patients to radioligands and demand cyclotron capacity, so they cannot be applied outside of specialized centers. In 1988, Daniel Weinberger and associates at the NIMH employed PET to help develop the "hypofrontality" hypothesis in schizophrenia.

Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) entered medicine in the late 1970s. In psychiatry, the first controlled study using MRI was led by Nancy Andreasen, finding in 1986 in the Archives of General Psychiatry that a population of schizophrenic males had "markedly smaller frontal lobes." Functional MRI (fMRI) examines changes in blood oxygen level, a proxy measure of neural activity. The Yvette Sheline group at Washington University in St. Louis used fMRI in 2002, to demonstrate that long-standing depression produced a decrease in 5-HT(2A) receptors in the hippocampus, a functional finding. Structural MRI measures brain and regional gray and white matter volumes, and was used in 2003, by Sheline and colleagues to measure loss of hippocampal volume in severe untreated depression (the longer untreated, the greater the loss of volume). Thus, in Sheline's terms (in the 2002 article), "Functional imaging extends the sensitivity and specificity of structural imaging."

Brain Connectivity

Brain connectivity refers to a pattern of anatomical links ("anatomical connectivity"), of statistical dependencies ("functional connectivity"), or of causal interactions ("effective connectivity") between distinct units within the nervous system. Using a variety of MRI and fMRI methodologies, information about the functional and structural integrity of neural circuitry —"neural profiles"—can be provided in healthy and psychiatrically unwell individuals. In 2008, ML Phillips and associates identified abnormalities of connectivity in the neural circuitry supporting emotion processing and regulation in bipolar disorder. In 2010, in Biological Psychiatry, Amelia Versace and colleagues attempted to differentiate unipolar versus bipolar depression on the basis of the functional white matter integrity of the neural circuitry regulating emotion.

Quantitative EEG (QEEG) Brain Mapping

Introduced in the 1980s, the QEEG brain map produces a three-dimensional reconstruction of the brain's activity. Using advanced algorithms, it is possible to identify brain-wave patterns from deep brain structures. The method is noninvasive, but has a space resolution weaker than that of the above-mentioned methods. QEEG has been used by ER

John and associates to identify activity clusters in a variety of different psychotic states. One reason for the lesser popularity of QEEG compared to SPECT, PET and MRI may be, as Tom Bolwig observes, "that understanding the elements of electrophysiology seems a bit harder for many physicians than accepting colored spots on an image of the brain." Yet QEEG represents pure neuronal transport and can record immediately and for very long periods of time, which are essential qualities for monitoring chains of events.

PART VII: THE BEGINNING OF THE SECOND BIOLOGICAL PSYCHIATRY

Psychiatry in Europe, with its heavy emphasis on heredity, had always been biologically oriented. But in the 1930s, psychoanalysis in the United States wiped the biological slate clean. Then in 1946, at a meeting at the Fairmont Hotel in San Francisco, the biological thread began to re-emerge. Two California physicians—Johannes M Nielsen, a neurologist at the University of Southern California, and his student, George N Thompson, chief psychiatrist at the Los Angeles General Hospital, convoked those interested in "the biological basis of behavior." The meeting was attended by the heavyweights in US neurology and psychiatry: Percival Bailey, a Chicago neurologist; Stanley Cobb, professor of neuropathology at Harvard; and Karl M Bowman, at the Langley Porter clinic in San Francisco. From this meeting emerged the Society of Biological Psychiatry. In 1947, Nielsen and Thompson published the first textbook of biological psychiatry in the United States, *The Engrammes of Psychiatry*.

The Revolution in Psychopharmacology

Although the term "psychopharmacology" seems to have been coined by David Macht, a pharmacologist at Johns Hopkins University, in 1920, to describe "the effect of drugs on psychological functions," it was not until the mid-1950s, with the explosion of interest in mechanisms prompted by the success of the antipsychotics, that the term started to come into use. When Paris psychiatrist Jean Thuillier visited Joel Elkes's newly founded department of experimental psychiatry in Birmingham in the mid-1950s, he was introduced by Willy Meyer-Gross. (Meyer-Gross, a refugee from Hitlerism who previously had been a mainstay at Heidelberg, was finishing his career in Birmingham as an associate of Elkes.) Meyer-Gross said in the introduction, "From everything that will be said to you this evening, bear in mind this new word: psychopharmacology. We already practise it here without realizing it, and everything we do tomorrow and in the future, everything we will want to do, will be gathered into this new discipline."

There has been much discussion pro and con about psychiatry's turn toward psychopharmacology. For all current concerns about polypharmacy and "cosmetic psychopharmacology," one does well to bear in mind the

recollections of Robert Felix, inaugural director of NIMH in 1949, of his days as a resident in a mental hospital in 1931, "For disturbed patients the continuous flow bath was almost universally used. If the patient was too disturbed to lie in the tub, either a hot or a cold pack was used, which, while not so-called, served effectively as a physical restraint. For either excited or depressed patients who manifested agitation, three chemical substances were used: paraldehyde, phenobarbital, or bromide ... Our bible was Bleuler's Textbook of Psychiatry. In that text Chapter XI, 'The Treatment of Mental Diseases in General,' covered 11 pages."

In the years after the Second World War, psychiatry swung on a massive hinge. "There can be no twisted thought without a twisted molecule," wrote Ralph Gerard in 1955 in the American Journal of Psychiatry, keynoting the advent of a new era in psychiatry, the easing away from psychoanalysis, and the turn to pharmacotherapeutics as the main treatment modality. Biology marked, said Gerard 2 years later in Science, "the emphatic re-entry of the brain into the mental arena." In these years, the field evolved from being the lowest prescribing of the medical specialties to the highest. (In 2005 90.1 percent of all psychiatric visits ended with a prescription; the next highest was cardiovascular diseases with 85.4 percent.)

The revolution may have begun with the synthesis in 1938 of lysergic acid diethylamide (LSD) by Sandoz chemist Albert Hofmann. It was in 1943, that Hofmann discovered it was psychoactive; LSD went on to have a brief career in psychotherapy and the treatment of alcoholism, but it generated great scientific curiosity: Was its mechanism identical to that of schizophrenia?

In 1952, the Rhône-Poulenc Laboratories in Paris introduced the first antipsychotic (neuroleptic), chlorpromazine, of the phenothiazine class of three-ringed structures, which was brand-named Largactil in Europe, Thorazine in the United States. The staff of Paris psychiatry professor Jean Delay conducted the initial clinical trials. A number of other phenothiazine antipsychotics followed.

These two compounds—LSD and chlorpromazine—initiated the era of psychopharmacology. As Theodore Sourkes at McGill University, "a biochemist in psychiatry" as he called himself, put it later, "The neuroleptic action of chlorpromazine and the psychodysleptic action of LSD ... held out the promise, on the one hand, of a chemical treatment of mental disorders and, on the other, of exploring the possibility that schizophrenia was caused by a toxic substance. As a consequence, the fifties and sixties of the last century saw the introduction of chemical theories into psychiatry."

The proliferation of phenothiazines gave rise to the classification of "sedative" versus "incisive" neuroleptics, the former presumably for agitated patients, the later for the withdrawn and retarded. Yet an analysis by Arthur Platz and coworkers at the Perry Point Veterans Administration Hospital in 1967 demonstrated that there was no difference in therapeutic

profiles.

A more important distinction was effectiveness in patients with "positive symptoms" versus those with the apathetic withdrawal of "negative symptoms." This was realized early. Fritz Freyhan, a German émigré considered one of the founders of psychopharmacology, observed in 1957, at the Second International Congress of Psychiatry in Zurich (the first was in Paris in 1950), "Two schizophrenic patients ... may differ fundamentally, one being apathetic, the other excited and disturbed. Only the latter is a candidate for neuroleptic treatment ..."

The new drugs marked the beginning of the end of the dominance of psychoanalysis. Heinz Lehmann, who conducted some of the first trials of chlorpromazine, said in a 1992 interview in the Bulletin of the Canadian Psychiatric Association, "The demonstration [was important] that psychotropic drugs, such as chlorpromazine ...are actually effective, so effective that there remains no doubt about it, even for the sturdiest psychoanalyst. Eventually they all had to admit that there is a physical substrate to, say, schizophrenia and bipolar or manic-depressive disorder, because certain physical substances, drugs, definitely are helpful in treating these disorders. Before that, until the early 1950s, it was archaic and very unsophisticated—almost simplistic—to think that anything physical could be done about treatment in schizophrenia. It had been tried for so many years, decades, and nothing had ever helped, so one would have to understand schizophrenia on psychodynamic terms, and the only rational intervention was psychoanalytical treatment. Well, the fact that I was able to demonstrate, with others, that physical substances can profoundly change the course of schizophrenia and other psychoses made it quite clear that there was a biological substrate ..."

It was thus that psychiatry, after so many years out in the cold with psychoanalysis, was able to rejoin the rest of medicine. As Theodore Rothman, at the University of Southern California and Los Angeles Veterans Administration Hospital, said in 1960, at an organizational meeting that would lead to the founding of the American College of Neuropsychopharmacology (ACNP) the following year, "I can find no better keynote thought than that psychopharmacology has given us a common front with the basic medical sciences. Psychopharmacology is compelling us to re-integrate psychiatry into medicine."

Back into medicine! But how did the new drugs work? The 1960s and 1970s were captivated by one-neurotransmitter-one-disease hypotheses.

Norepinephrine was first off the blocks. In 1965, in the American Journal of Psychiatry, Joseph Schildkraut of Harvard University proposed a "catecholamine hypothesis of affective disorders" (the neurotransmitter norepinephrine has a portion with a catechol structure). "Those drugs which cause depletion and inactivation of norepinephrine centrally produce sedation or depression, while drugs which increase or potentiate brain

norepinephrine ...generally exert an antidepressant effect in man."

In 1968, Swedish pharmacologist Arvid Carlsson and coworkers at Gothenburg University noted in the Journal of Pharmacy and Pharmacology that the TCA imipramine blocked the reuptake of serotonin. "This action may be of importance for its antidepressive properties," the authors said, initiating the serotonin hypothesis of depression.

It remained for dopamine, the third major biogenic amine neurotransmitter, to find its distinctive disease, and this occurred in 1976, as it was reported simultaneously by psychiatrist Solomon Snyder at Johns Hopkins University (in Science) and pharmacologist Philip Seeman at the University of Toronto (in Nature) that the potency of antipsychotic drugs in schizophrenia was a function of their ability to block the dopamine receptor.

The original scientific findings underlying these reports were never overturned, yet neurotransmission is a complex phenomenon and the field quickly lost confidence in the specificity hypothesis—although for decades it was marketed relentlessly by the pharmaceutical industry. In 1978, JL Costa and coinvestigators at NIMH said of agents affecting serotonin metabolism, "... It certainly may be that a serotonin deficit state is not a prerequisite for therapeutic efficacy with any of these drugs." It is a shame that these hypotheses, so attractive for their theoretical elegance and simplicity, have now been largely discredited, because psychopharmacology has never again been able to develop a unitary theory of drug action. At this writing, the brain remains a black box.

The Golden 50s

The year 1955 saw the launch of the first blockbuster drug in psychiatry, Wallace Laboratories' meprobamate, trade-named Miltown (and Equanil by Wyeth Laboratories, which licensed it from Wallace). Launched as a "tranquilizer," Miltown became almost a household word, opening up that large market for "nerves" between the antipsychotics and aspirin (see below). It was displaced by the benzodiazepines in the 1960s.

The monoamine oxidase inhibitors (MAOIs) were probably the first effective drug class, aside from opium, for serious depression. This awareness took a while dawning. The first MAOI, iproniazid (Hoffmann-La Roche's Marsilid) reached the market in 1951, for tuberculosis, but only in 1957, did the company begin to market it for depression. A series of other MAOIs followed, the most effective of which was probably phenelzine (Warner-Chilcott's Nardil), introduced in 1959 and still in use.

"Tranquilizers"

It was London physician Edward Sutleffe, who had a substantial practice of patients with "nervous irritation" among other diagnoses, who in his

Medical and Surgical Cases in 1824, introduced into medicine the term "tranquilizer" in connection with a "tranquilizing" drug, "the juice of ground ivy," that he was promoting. The word then slumbered in forgetfulness until 1953, when Ciba billed its new antihypertensive Serpasil (reserpine) as a "safe tranquilizer-antihypertensive." Its calming effects were soon noted in psychological medicine and the concept of "tranquilization" became the common denominator of a range of drugs from the antipsychotic chlorpromazine, introduced clinically in the United States in 1954, to the "antiphobic" agent benactyzine (Suavitil) that Merck brought out in 1957. The tranquilizers were thought an improvement upon the barbiturates because they had the same relaxing and composing effects without the daytime sedation. Yet soon the concept of tranquilizer was ruptured by the introduction of the benzodiazepines after 1960 (see below). These "antianxiety" agents made the concept of tranquilization, which had lumped together antipsychotics and antinervous agents, seem outmoded.

Antipsychotics

The antipsychotic era began with the introduction of chlorpromazine in clinical trials in France in 1952, by the Rhône-Poulenc Laboratories. Called Largactil in Europe and Thorazine in the United States, chlorpromazine is thought to have initiated the era of psychopharmacology by prompting research into mechanisms. A slew of similar phenothiazine antipsychotics, also called neuroleptics, followed, including Smith Kline's prochlorperazine (Compazine) in 1956, and thioridazine (Sandoz's Mellaril) in 1958. Haloperidol (Janssen's Haldol, a member of the butyrophenone class) entered practice in the United States in 1966.

Tricyclic Antidepressants

Owing to the close clinical observation of Swiss psychiatrist Roland Kuhn, Geigy Laboratories in Basel realized in 1957, that their new compound imipramine, a member of the tricyclic class, was active against "vital depression," rather than being an antischizophrenic, as they had hoped (the antipsychotic market was just taking off). Marketed in the United States in 1959, imipramine (brand-named Tofranil) initiated a half-century of "antidepressant" treatment. News of the TCAs traveled quickly. At a conference at McGill University in 1959, Hans Hoff, who had become the professor of psychiatry in Vienna after the Second World War, admitted that he and psychoanalyst colleague Paul Schilder "had once tried to destroy the Kraepelinian nomenclature because of the feeling that psychodynamics and personal reactions were more important. Now, however, since a treatment had become available, diagnosis is essential."

"Mood Stabilizers"

The mood-stabilizer era began with lithium in 1949. The term apparently surfaced for the first time in connection with a drug called aminophenylpyridone (Maltbie Laboratories' Dornwal) in 1960. The phrase "mood stabilizer" did not, however, come into general use until Robert Post and colleagues at NIMH floated the "kindling hypothesis" in 1986, in the British Journal of Psychiatry: recurrent depressions could tip the brain into autonomous cycles of mood disorder. Other writers then proposed that various anticonvulsants could break into the kindling process and stabilize the hyperirritable brain. Aside from lithium, considered mood stabilizers were: derivatives of valproic acid called valproate and valpromide that medicinal chemist George Carraz at Berthier Laboratories in Grenoble synthesized in 1963; carbamazepine, patented in 1960 by Geigy, was introduced as a psychoactive compound in Japan, then discovered in 1980, by NIMH investigators to have effects in manic-depression. Lamotrigine was patented by the Wellcome Foundation in 1981 and inaugurated in 2003, in the United States by GlaxoSmithKline as Lamictal ("a stabilizing force"). The judgment of David Healy and colleagues in 2003, in Bipolar Disorders was that, "These agents are now thought of as being a homogeneous group [but] they may in fact be quite diverse agents all of which have a certain utility when used judiciously in manic depressive states."

After the 1950s, Relatively Few New Drug Classes

Benzodiazepines. The benzodiazepine story began with the marketing of Hoffmann-La Roche's Librium (chlordiazepoxide) in 1960. The numerous subsequent Roche benzos were also synthesized by medicinal chemist Leo Sternbach. So popular did the benzodiazepines become that entire "eras" in the 1960s and 1970s were named after such members of this class as Valium (diazepam, launched in 1963). With the exception of the SSRIs (selective serotonin reuptake inhibitors—see below), the benzodiazepines are probably the best-selling psychoactive drug class in history, and moreover possessed genuine efficacy that later agents do not seem to have matched. Other benzos followed: clonazepam (Clonopin) reached the American market in 1975; and alprazolam (Upjohn's Xanax) was launched in 1981. A short-acting benzo, Upjohn's triazolam (Halcion), came out in the United States in 1982, and caused much controversy about discontinuation. There was a long period of fearfulness about addiction, although in 1983, in the Journal of the American Medical Association, Karl Rickels and associates at the University of Pennsylvania found the risks of withdrawal reactions in diazepam minimal for up to 8 months of use. In the new millennium, the benzos began a comeback. As Peter Roy-Byrne at the University of Washington School of Medicine put it in an editorial in 2014 in the American Journal of Psychiatry, "There is now general appreciation that, with careful clinical evaluation to rule out pre-existing addiction

issues … benzodiazepines are safe and highly effective."

SSRIs: Selective Serotonin Reuptake Inhibitors. Marketed as agents that specifically inhibited the neuronal reuptake of serotonin, this drug class became the most successful in the history of psychiatry. Prozac (fluoxetine), which became a household word, was marketed by the Eli Lilly Company in Belgium in 1986, in the United States in 1988. Among the main SSRIs, there followed Pfizer's Zoloft (sertraline), marketed in 1992; paroxetine (Novartis's Paxil), introduced in the United States in 1993; and citalopram (Lundbeck's Cipramil), brought out in the United States in 1998 by Forest Laboratories. The SSRIs were initially indicated for "depression," then for "anxiety," and by the time they all finally went off patent, countless were the indications for which this versatile drug class served.

Second-Generation Antipsychotics. The so-called "second-generation antipsychotics" (SGA's) have turned out not that different from the first generation and seem to be more a marketing concept. Yet it might be noted that clozapine (Sandoz's Clozaril), deemed the most effective of them, is not exactly a new drug, having been patented in 1963. Leo Hollister's judgment in 1987 of the antipsychotics is of interest (Hollister led the Veterans Administration hospital in Palo Alto, California and is generally acknowledged as the dean of American psychopharmacology): "It is most discouraging that more effective pharmacotherapy for schizophrenia has not been developed in the more than three decades since the introduction of the first effective drugs." The current drugs, he said, "are not curative; their ameliorative effects are often limited; many patients remain totally unresponsive; they are unpleasant to take, so that many patients are less than fully compliant; and they produce major side effects."

PART VIII: PEDIATRIC PSYCHIATRY

Psychiatrists have always been familiar with mental disease in children. In 1809, Haslam reported cases of "insane children," such as a 7-year-old boy who was slow to develop and was placed on "the female side" of the hospital. Early on, psychotic illness in children was noted, sometimes as a distinctive form of what would later be called childhood schizophrenia, sometimes not. The "hebephrenia" that Ewald Hecker described in 1871 was commonly recognized in children. As Auguste Voisin, a staff psychiatrist at the Salpêtrière in Paris commented in his 1883 textbook of "insanity in the young": "In partial insanity, the child may be manic [extatique], subject to ideas of homicide, arson, or theft, have deluded and demonic ideas … It is in the context of this partial insanity that we find the singular cases of childhood suicide, sometimes preceded by notes from which extreme exaltation shows through; thus a child of thirteen who hanged himself left behind a note saying, 'I bequeath my soul to Rousseau,

my body to the earth.'"

Textbooks of child psychiatry were not slow in coming. In 1887, Hermann Emminghaus, professor of psychiatry in Freiburg, produced what was probably the first, The Psychic Disorders of Childhood (*Die psychischen Störungen des Kindesalters*). The foundational modern text of psychological approaches to disorders of childhood, Lectures on the Psychopathology of Childhood (Vorlesungen über Psychopathologie des Kindesalters) was written by August Homburger at the Heidelberg University Psychiatric Hospital in 1926. Among the first dedicated hospital divisions for children with mental illnesses was that at Bicêtre Hospice in Paris, which in 1838, received responsibility for the care of "children with mental retardation and epilepsy." In 1842, Édouard Séguin, who had developed a system for the care of children with retardation, took over the service, and after an interregnum, in 1851, Louis Delasiauve succeeded him. Désiré-Magloire Bourneville took it over in 1879, and in the course of the next decades made Bicêtre into one of the most progressive children's psychiatric units in the world. (In 1892, he initiated a gymnastic competition, which became the unacknowledged forerunner of the Special Olympics.) In 1850, Séguin emigrated to the United States and initiated here the enlightened treatment of children who have what are now called intellectual disabilities.

The epidemic of encephalitis lethargica in the years 1917 to 1927 represented a strong inflection in the growing awareness of distinctive psychiatric illness in children, primarily because, unlike adults, the disease produced striking psychiatric changes in many children, transforming them from normal happy creatures into the "apache type" with severe behavioral problems. Said Leslie Colin Cook, who was a staffer at the West Park Mental Hospital in Epsom that had a center for such patients, in the Lancet in 1928, "Children of irreproachable family, hitherto of exemplary behavior and character, may in a very short period become guilty of every delinquency from mere naughtiness to cruelty and murder, and from mild bad habits to gross indecencies." This had happened, he said, in only one adult patient with the disease. The "EL" epidemic thus strengthened the interest of the field in children as a special kind of patient.

In the middle third of the 20th century, psychoanalysis became the doctrinal mainstay of child psychiatry. Freud himself did not see children and had relatively little experience with them. Psychoanalysis was first applied to the child and adolescent population by Viennese educator and psychoanalyst August Aichhorn in his 1925 work Wayward Youth (Verwahrloste Jugend), about the ego structure of juvenile delinquency. Sigmund Freud's daughter Anna published her first book in 1927, An Introduction to the Technique of Child Analysis, and in 1936, The Ego and the Mechanisms of Defense. The Freud family fled to England after the Nazi Anschluss with Austria in 1938, and in 1941, Anna Freud helped found

the Hampstead Nursery in London, where she treated children for several years. This gave way to the Hampstead Child Therapy Course and Clinic, which Anna directed until her death in 1982 (on her rivalry with Melanie Klein see above).

After the Second World War, the baton in child psychiatry passed to the United States. When the National Institute of Mental Health opened in 1949, the "quality of mothering" was put on the table—especially so by psychoanalysts such as Frieda Fromm-Reichmann at Chestnut Lodge in Maryland, who, in Psychiatry in 1948, had originated such concepts as the "schizophrenogenic mother." Among the leading child psychoanalysts in New York was Hungarian-born Margaret Mahler who in 1950 cofounded the Masters Children's Center in Manhattan, where she developed the tripartite treatment model involving the mother in the child's treatment. Her most important work was *The Psychological Birth of the Human Infant* (1975).

Prompted by an interest in the ideas of Harvard psychologist Burrhus Frederick Skinner, in the late 1940s and 1950s behaviorism started to become à la page. Psychoanalysis and behaviorism dominated pediatric psychiatry in the United States after the Second World War.

The child guidance clinics reached back to the First World War, and in 1948 they created an organization of clinics, the American Association of Psychiatric Clinics for Children, to develop standards of care and training. In 1953, the American Academy of Child Psychiatry was founded; in 1959, the American Board of Psychiatry and Neurology approved a 2-year child psychiatry residency, to be grafted onto the regular training. The field received its own voice, The Journal of the American Academy of Child Psychiatry, in 1962, and a quarter century later, in 1986, the Academy changed its name to American Academy of Child and Adolescent Psychiatry.

In the heyday of psychoanalysis, "pills" were frowned on and great confidence placed in psychotherapy. Yet the revolution in adult psychopharmacology in the 1950s had echoes in the pediatric world. "Psychopharmacology has revolutionized child psychiatry," said Jovan Simeon, professor of psychiatry and child psychopharmacologist at the University of Ottawa, in 2004. Children had, of course, always been treated with such psychoactive agents as opium, mixed together with alcohol in laudanum (see above), which parents once routinely administered as a hypnotic. Treating children with psychopharmaceuticals entered the academic forum in 1937, when Charles Bradley, director of the Emma Pendleton Bradley Home in East Providence, Rhode Island, in the American Journal of Psychiatry reported the benefit of Benzedrine in "behavior problem" children, later defined as "hyperactive." Loretta Bender, at Bellevue Hospital in New York, used chlorpromazine and amphetamine extensively in treatment. Systematic research in pediatric

psychopharmacology began in 1958 when Leon Eisenberg at Johns Hopkins University received the first NIMH grant to study drugs in children, conducting the first controlled trial in pediatric psychopharmacology. This torch was then notably carried forward by Bender's student Barbara Fish, first at Bellevue, then after 1972, at the UCLA.

Such progress in pediatric psychopharmacology probably counts as an American "first." There were a number of other American firsts, such as recognition by Lee Wachtel at the Kennedy Krieger Institute in Baltimore and Dirk Dhossche, in a 2010 article in Medical Hypotheses, of stereotypical movements in autism and intellectual disability as a sign of catatonia—and treatable with benzodiazepines and convulsive therapy.

As for diagnosis, disease classification in pediatric psychiatry has been in even greater turmoil than in general psychiatry, as psychoanalysis, behaviorism, and biological approaches have dueled. Driven by parental concerns, clinicians are often pressed to make diagnoses such as attention deficit hyperactivity disorder. The fourth edition of DSM (DSM-IV) included the diagnosis "bipolar II disorder," characterized by depression plus "at least one hypomanic episode." Yet the field readily accepted outbursts of anger as a substitute for hypomania, touching off a virtual epidemic of "bipolar II" in the pediatric population, which the embarrassed drafters of DSM-5 were obliged to dial back in 2013, with the newly contrived "depressive" diagnosis "disruptive mood dysregulation disorder," said to be "included for children up to age 18 years who exhibit persistent irritability and frequent episodes of extreme behavioral dyscontrol." The classification of illness in children has thus still not stabilized.

PART IX: NEW FORMS OF CARE

The asylum was a forbidding institution. Yet many patients outside its walls also required care, and over the years these walls were slowly broken down.

Boarding patients out to local families achieved first this separation of care from the four walls of the institution. The earliest instance of this was seen at Gheel in the Brabant region of Belgium, where patient out-boarding went back to the 13th century. Yet the local farmers often had little tolerance for the antics of the patients, of whom they might board up to five at a time, and when Pliny Earle, of the Friends' Asylum in Frankford, visited there in the late 1830s, (describing his travels in a book in 1841, he was shocked to find that the patients, "if mischievous or furious, are chained at the wrists and ankles. In every house, there is a ring beside the chimney, or attached to the beds, to which they may be chained." Boarding out was never widespread in the United States, possibly because the farms were too dispersed (unlike the compact Belgian villages). But in Scotland around 1900, 23 percent of the asylum patients were boarded out.

There were other ways of breaking down the walls. The first psychiatric

ward in a general hospital in the United States was established in 1902, at Albany Hospital. In 1885, Pennsylvania Hospital founded the first outpatient clinic, or very close to the first, in the United States, "so as to provide a dispensary service for the gratuitous treatment of persons suffering with incipient mental disease." (On these developments, see Henry Hurd's massive four-volume work in 1916 on the Institutional Care of the Insane.)

One of the unhappiest stories in US psychiatry was that of "deinstitutionalization," under the influence of the antipsychiatric ideologies of the 1960s, discharging in-patients to a kind of community care that often turned out to be the care of the streets. This created an existential dilemma for psychiatry: what to do with the patients who were seriously ill? As Jonathan Cole said in an interview in 1996, "We've got very few places where we can serve the kind of patient who takes a long time to get better and where real rehabilitation is done. We tend to have more people who are home and crazy than we should have and nobody's going to pay for their treatment."

Beginning in the 1960s, this vast outpouring resulted in the closing of many of the classic asylums. In 1964, there were in the United States 255 public psychiatry beds per 100,000 population; by 2014, this figure had dropped to 11 per 100,000. The patients were supposedly to be maintained in community settings on antipsychotics; in reality they fed the vast ranks of the "homeless" and the imprisoned. In 2002, 64.2 percent of all inmates of local jails in the United States had a "mental health problem" (75.4 percent of the females). The figures were comparable for state and federal prisons. Donald Klein of Columbia University said later in an interview, "It wasn't deinstitutionalization, it was trans-institutionalization, from the state to the city budget. The patients, but not the money, were dumped on the cities. The cities didn't pick them up and that was a disaster. Instead of an open door, we had a revolving door. There was a sudden rise in homelessness. This led to federally supported 'community care' which was staffed by therapy-minded psychologists. They had a psychiatrist who did little but sign bundles of prescriptions." By one estimate, "There are 10 times as many mentally ill persons in prisons than in state hospital beds." David Rubinow at the UNC, who penned these lines in the American Journal of Psychiatry in 2014, concluded "It would appear that we somehow have achieved the complete reversal of policies initiated over 170 years ago to ensure the humane treatment of those with mental illness."

Yet history has good news as well. On the threshold of the new millennium, the role of psychiatry itself was slowly changing. Once seen as a band-aid for the "walking wounded," psychiatry was becoming the main resource for the seriously ill. In the United States in 2001 to 2002, 40 percent of patients identified in a national household survey with a serious mental problem would be seen by a psychiatrist. As for the rest of patients

who formerly would have been "in therapy," psychiatry had ceased to be the main port of call: only 21 percent of those with moderate and 9 percent of those with mild disorders saw psychiatrists; they distributed themselves instead across the entire gamut of mental and medical health services, from "general medical" to "internet" (the main resource for 30 percent of those with "moderate" mental problems). Alarmingly, of those seeing a psychiatrist, 29 percent were said to have "no disorder." Readers will give a weary sigh. Yes, it's true: some things never change.

REFERENCES

Arns M, Conners CK, Kraemer HC. A decade of EEG theta/beta ratio research in ADHD: a meta-analysis. *J Atten Disord.* 2013;17(5):374–383.

Billod E. *Les Aliénés en Italie*. Paris: Masson; 1884:72.

Cathell DW. *Book on the Physician Himself.* 10th ed. Philadelphia, PA: Davis; 1900:150–151.

Chiarugi V. *Della Pazzia in Genere e in Specie,* 3 vol. 1793–94; reprint Rome: Vecchiarelli; 1991: I, 211–212; II, 67, 75.

Cohen RA. Interview. In: Ban TA, ed. *An Oral History of Neuropsychopharmacology. I,* 194. Brentwood, TN: ACNP; 2011.

Ellenberger H. A comparison of European and American psychiatry. *Bull Menninger Clin.* 1955;19(2):43–52, 49.

Engel JW. Early psychiatry at Yale: Milton C Winternitz and the founding of the department of psychiatry and mental hygiene. *Yale J Biol Med.* 1994;67(1–2):33–47, 39.

Felix RH. Foreword: the life and times of psychiatric research. In: Earl Usdin, et al., eds. *Biochemistry of Mental Disorders.* New York: Dekker; 1978, vii-xvii, viii-ix.

Freyhan F. Discussion. In: Kline NS, ed. *Psychopharmacology Frontiers: Proceedings of the Psychopharmacology Symposium, Second International Congress of Psychiatry.* Boston, MA: Little Brown; 1959:423.

Gerard RW. Orientation: analysis of program. In: Cole JO, Gerard RW, eds. *Psychopharmacology: Problems in Evaluation.* Washington, DC: National Academy of Sciences; 1959, 9–19, xvii, 662 pp.

Haslam J. *Considerations on the Moral Management of Insane Persons.* London: Chapman and Hall; 1996.

Healy D. *Psychopharmacologists, I,* 178.

Hoenig J. Editorial: Psychiatric nosology. *Can J Psychiatry.* 1981;26(2):85.

Hoff P. Discussion in Anon. [Proceedings of] McGill University Conference on Depression and Allied States, Montreal, March 19–21. *Can J Psychiatric.* 1959; spec supplement, 4 (1959), S67–S68.

Hollister LE. Strategies for research in clinical psychopharmacology. In: Herbert Y Meltzer, ed. *Psychopharmacology: The Third Generation of Progress.* New York: Raven Press; 1987, 31–38, 31.

Lader M. Interview. In: Healy D, ed. *The Psychopharmacologists.* London: Chapman & Hall; 1996: I, 463–481, 470. *Hereafter cited as "Healy, Psychopharmacologists."*

Lehmann H. Interview. In: *Healy D, ed. Psychopharmacologists, I,* 182.

Letchworth WF. *The Insane in Foreign Countries*. New York: Putnam's; 1889:172.

Lindon JA. Melanie Klein. In: Alexander F, et al., eds. *Psychoanalytic Pioneers*. New York: Basic Books; 1966:360–372, 367.

Litchfield HR. Aminophenylpyridone: a new mood-stabilizing drug. *Arch Pediatr*. 1960;77:133–137.

Moreau J-J. *Notes sur les établissements d'aliénés de Siegburg, Halle, Dresde, Prague, Berlin et Vienne*. *Annales Médico-psychologiques*. 2nd ser, 6. 1854;428–439, 430.

Murphy DL, Campbell I, Costa JL. Current status of the indoleamine hypothesis of the affective disorders. In: Lipton MA, et al., eds. *Psychopharmacology: A Generation of Progress*. New York: Raven Press; 1978, 1235–47, 1235.

Pandy K. *Die Irrenfürsorge in Europa, German trans from Hungarian*. Berlin: Reimer; 1908:482.

Report of the Metropolitan Commissioners in Lunacy to the Lord Chancellor. London: Bradbury; 1844:37.

Rosenthal D. The heredity-environment issue in schizophrenia: summary of the conference and present status of our knowledge. In: David Rosenthal, Seymour S Kety, eds. *The Transmission of Schizophrenia*. Oxford: Pergamon; 1968:413–415, 415.

Shepherd M. Interview with Sir Aubrey Lewis (Part I). *Psychiatric Bull*. 1993;17:738–747, 743.

Shorter E. *A History of Psychiatry: From the Era of the Asylum to the Age of Prozac*. New York: Wiley; 1997.

Shorter S. *A Historical Dictionary of Psychiatry*. New York: Oxford University Press; 2005.

Shorter S, Fink M. *Endocrine Psychiatry: Solving the Riddle of Melancholia*. New York: Oxford University Press; 2010.

Simpson G. Interview. In: Thomas Ban, ed. *An Oral History of Neuropsychopharmacology. The First Fifty Years: Peer Interviews* (Brentwood TN: ACNP, 2011), IV, 373–384, 373; hereafter cited as "Ban, Oral History Neuropsychopharmacology."

Snell LDC. *Ueber die verschiedenen Formen des Wahnsinns*. *Allgemeine Zeitschrift für Psychiatrie*. 1874;30:319–326, 320–321.

Sourkes TL. A biochemist in psychiatry: dropping in on the psychopharmacology era. In: Thomas A Ban, et al., eds. *Reflections on Twentieth-Century Psychopharmacology*. Budapest: Animula; 2004:290–298, 291–292.

Substance Abuse and Mental Health Services Administration. *Mental Health: United States, 2008*. *HHS Pub. No. (SMA) 10–4590*. Rockville, MD: Center for Mental Health Services, SAMHSA; 2010:123.

Thuillier J. *Ten Years that Changed the Face of Mental Illness (1981), Eng trans*. London: Dunitz; 1999:139.

Tissot SA. *Traité des Nerfs et de Leurs Maladies*. Paris: Didot; 1780:II(2):358.

Voyer I. *Anxietate*. Leiden: Mostert: Leiden med diss; 1769.

Westphal A. Manie, Melancholie, manisch-depressives Irresein. In: Otto Binswanger, et al., eds. *Lehrbuch der Psychiatrie*. Jena: Fischer; 1923:95–141, 119.

Woodward SB. Observations on the medical treatment of insanity. 1850. *Am J Psychiatry*. 1994;151(6 Suppl):220–230.

62

World Aspects of Psychiatry

DINESH BHUGRA, PH.D., MARIO MAJ, M.D., PH.D., ANTONIO VENTRIGLIO, M.D., PH.D., AND PEDRO RUIZ, M.D.

In the past few years largely at the instigation of high-income countries, a global mental health movement has emerged. Partly this is a response to rapid globalization which has meant that people, resources, and ideas are moving much more rapidly than at any other time in the history of humanity. This has led to a clear acknowledgement that a number of factors at social and cultural levels are at play in influencing how distress is experienced and expressed. There remains a danger in that global mental health may follow the route of Anglo-American influences to be seen "dictating" to the low- and middle-income countries. There remain a number of serious questions in relation to the differences in rates and responses across different nations of the world. Some of these are discussed later in this chapter. The authors prefer and choose the term "world psychiatry," as it reflects the real state of psychiatry at the present time. The international progress of collaboration has shifted considerably over the past two centuries. From a perspective of the rulers seeing the vanquished as exotic natives who either did not have mental capabilities to develop mental illness to seeing them as victims of the social factors. This gave way to international collaborative epidemiological studies which have contributed to the current understanding of both similarities and differences in etiology and outcomes of psychiatric disorders.

There is no doubt that mental disorders are highly prevalent across all regions of the world even though there may be a small degree of variation across societies and cultures. These disorders represent a major source of disability and social burden worldwide. Even though treatments for all of these disorders are available and have been found to be efficacious in different settings such as high-income and low- and middle-income countries, this has not always translated into wider access and availability of these interventions. Mental disorders remain remarkably undertreated worldwide. This treatment gap is high enough in low-income countries but even in high-income countries high levels of gap remain. To complicate matters further many countries either do not have any mental health policies and there are many others who have outdated policies sometimes dating back to 19th century. Consequently both funding for treatment and

research into mental health falls well below the resources offered and available for physical disorders such as cancer and heart disease.

Although there have been many proclamations internationally recently about the importance of mental health, resources to deal with mental disorders and investing in public mental health have not followed the rhetoric. Contribution of funding to mental health care remains scarce and unequally distributed.

This chapter provides an overview of world psychiatry but focusing on epidemiology, burden resulting from mental illness and the relationship between various illnesses and also stigma as well as the ethics of mental health care.

PREVALENCE AND BURDEN OF MENTAL DISORDERS WORLDWIDE

The general perception is that in lifetime, one in four individuals will develop mental illness of one kind or another. These figures have been confirmed by World Health Organization (WHO) which estimates that more than 25 percent of individuals worldwide develop one or more mental disorders during their lifetime. This observation has been supported by a series of face-to-face community surveys conducted in 17 countries in Africa, Asia, the Americas, Europe, and the Middle East. These World Mental Health Surveys have included a total of 85,052 respondents. The lifetime prevalence of any mental disorder was found to be higher than 25 percent in 11 of 17 countries and higher than 15 percent in all countries except China and Nigeria, where prevalence rates were likely to be downwardly biased due to respondent reluctance to admit mental health problems. Whether these rates are genuinely low or whether these reflect reporting bias needs further exploration. Not surprisingly anxiety disorders were the most prevalent in 10 out of 17 countries, and mood disorders in 6 (Table 62–1). Anxiety disorders include agoraphobia, adult separation anxiety disorder, generalized anxiety disorder, panic disorder, post-traumatic stress disorder, social phobia, and specific phobia. Mood disorders include bipolar disorders, dysthymia, and major depressive disorder. Impulse-control disorders include intermittent explosive disorder, and reported persistence in the past 12 months of symptoms of three child-adolescent disorders (attention-deficit/hyperactivity disorder, conduct disorder, and oppositional-defiant disorder). Substance disorders include alcohol or drug abuse with or without dependence. Following the *Diagnostic and Statistical Manual of Mental Health Disorders*, Fourth Edition (DSM-IV) norms, organic exclusions were excluded. Lifetime co-occurrence of two or more mental disorders was quite common, so that the sum of the prevalence rates across the various disorder types was between 30 and 50 percent higher than the proportion of people having at least one mental disorder. Survival analysis showed that the odds ratios for anxiety, mood, and substance use disorders were higher in recent compared to

older cohorts. It is worth noting that even within the same geographical region there appear to be major variations in the prevalence of DSM-IV/CIDI disorders.

In the same series of studies, the 12-month prevalence showed similar variations. These findings illustrated regional and national variations. Once again Nigerian respondents reported only 6 percent prevalence with similar figures from Japan and China, but the rates were 4.5 times higher in the United States (see Table 62–2).

As shown in Table 62–3, when association between 12-month severity and days out of role were compared again variations within regions and across nations were substantial. Once again Nigeria and China had similar figures, whereas the United States showed the highest levels.

Table 62–1.
Lifetime Prevalence Estimates of DSM-IV/CIDI Disorders in the WMH Surveys

	Any Anxiety Disorder[a]		Any Mood Disorder		Any Externalizing Disorder[b]		Any Substance Disorder		Any Disorder	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)
I. WHO Region: Pan American Health Organization (PAHO)										
Colombia	25.3	(1.4)	14.6	(0.7)	9.6	(0.8)	9.6	(0.6)	39.1	(1.3)
Mexico	14.3	(0.9)	9.2	(0.5)	5.7	(0.6)[f]	7.8	(0.5)	26.1	(1.4)
United States	31.0	(1.0)	21.4	(0.6)	25.0	(1.1)	14.6	(0.6)	47.4	(1.1)
II. WHO Region: African Regional Office (AFRO)										
Nigeria	6.5	(0.9)	3.3	(0.3)	0.3	(0.1)[g,i]	3.7	(0.4)	12.0	(1.0)
South Africa	15.8	(0.8)[c,d]	9.8	(0.7)[e]	—	—[f,g,h,i]	13.3	(0.9)	30.3	(1.1)
III. WHO Region: Eastern Mediterranean Regional Office (EMRO)										
Lebanon	16.7	(1.6)	12.6	(0.9)	4.4	(0.9)[i]	2.2	(0.8)	25.8	(1.9)
IV. WHO Region: European Regional Office (EURO)										
Belgium	13.1	(1.9)	14.1	(1.0)[e]	5.2	(1.4)[f]	8.3	(0.9)[j]	29.1	(2.3)
France	22.3	(1.4)	21.0	(1.1)[e]	7.6	(1.3)[f]	7.1	(0.5)[j]	37.9	(1.7)
Germany	14.6	(1.5)	9.9	(0.6)[e]	3.1	(0.8)[f]	6.5	(0.6)[j]	25.2	(1.9)
Israel	5.2	(0.3)[c,d]	10.7	(0.5)	—	—[f,g,h,i]	5.3	(0.3)	17.6	(0.6)
Italy	11.0	(0.9)	9.9	(0.5)[e]	1.7	(0.4)[f]	1.3	(0.2)[j]	18.1	(1.1)
Netherlands	15.9	(1.1)	17.9	(1.0)[e]	4.7	(1.1)[f]	8.9	(0.9)[j]	31.7	(2.0)
Spain	9.9	(1.1)	10.6	(0.5)[e]	2.3	(0.8)[f]	3.6	(0.4)[j]	19.4	(1.4)
Ukraine	10.9	(0.8)[c]	15.8	(0.8)[e]	8.7	(1.1)[g,i]	15.0	(1.3)	36.1	(1.5)
V. WHO Region: Western Pacific Regional Office (WPRO)										
PRC[k]	4.8	(0.7)	3.6	(0.4)	4.3	(0.9)[g,i]	4.9	(0.7)	13.2	(1.3)
Japan	6.9	(0.6)[c]	7.6	(0.5)	2.8	(1.0)[g,h,i]	4.8	(0.5)	18.0	(1.1)
New Zealand	24.6	(0.7)[c]	20.4	(0.5)	—	—[f,g,h,i]	12.4	(0.4)	39.3	(0.9)

[a]Anxiety disorders include agoraphobia, adult separation anxiety disorder, generalized anxiety disorder, panic disorder, post-traumatic stress disorder, social phobia, and specific phobia. Mood disorders include bipolar disorders, dysthymia, and major depressive disorder. Impulse-control disorders include intermittent explosive disorder, and reported persistence in the past 12 months of symptoms of three child-adolescent disorders (attention-deficit/hyperactivity disorder, conduct disorder, and oppositional-defiant disorder). Substance disorders include alcohol or drug abuse with or without dependence. In the case of substance dependence, respondents who met full criteria at some time in their life and who continue to have any symptoms are considered to have 12-month dependence even if they currently do not meet full criteria for the disorder. Organic exclusions were made as specified in the Diagnostic and Statistical Manual of Mental Health Disorders (DSM), 4th edition.

[b]Impulse disorders restricted to age ≤39 (China, Ukraine, Nigeria) or to age ≤44 (all other countries).

[c]Adult separation anxiety disorder was not assessed.

[d]Specific phobia was not assessed.

[e]Bipolar disorders were not assessed.

[f]Intermittent explosive disorder was not assessed.

[g]Attention-deficit/hyperactivity disorder was not assessed.

[h]Conduct disorder was not assessed.

[i]Oppositional-defiant disorder was not assessed.

[j]Only alcohol abuse with or without dependence were assessed. No assessment was made of other drug abuse with or without dependence.

[k]People's Republic of China.

When the severity of these disorders was compared across nations and regions similar variations emerged. China, Japan, Nigeria and this time Italy showed lower levels of serious severity and higher levels of mild severity (see Table 62–4).

Using disability scale for common chronic physical and mental health disorders, interesting differences emerged between developed and developing countries (see Table 62–5). Cancer, asthma, hypertension, and diabetes showed higher proportions of severe disability in developing countries, whereas neck and backaches, chronic pain, and headaches were more seriously disabling in developed countries. Bipolar disorders, depression, generalized anxiety disorders, panic disorders, and posttraumatic disorder–related disability was elevated in developed countries. This variation raises interesting questions about whether these are genuinely higher levels and rates or are more commonly diagnosed.

Among people seen by primary health care professionals, more than 20 percent have one or more current mental disorders. In a study carried out by the WHO at 14 sites in Africa, Asia, the Americas, and Europe, the average current prevalence of any mental disorder was 24 percent, without consistent differences between low- and high-income countries. Not surprisingly the most common diagnoses were those of depression (average

10.4 percent) and generalized anxiety disorder (average 7.9 percent). Female rates were 1.89 times higher than male rates for depression, but male rates were higher for alcohol-related disorders, so that there was no sex difference in the proportion of people having at least one mental disorder. Physical ill health and educational disadvantage were both significantly associated with a diagnosis of mental disorder.

To quantify the burden of the various diseases and injuries, the WHO, in collaboration with the Harvard School of Public Health and the World Bank, introduced in 1993 a new measure, the disability-adjusted life year (DALY). DALYs for a given disease or injury are the sum of the years of life lost due to premature mortality plus the years lost due to disability for incident cases of that disease or injury in the general population. In the original estimates for the year 1990, mental and neurological disorders accounted for 10.5 percent of the total DALYs lost due to all diseases and injuries. Burden of disease related to depression and dysthymia is shown in Table 62–6. As previously reported, global prevalence of depression was reportedly more common in females. When a comparison was made between years lived with disability due to depression between 1990 and 2010, it appears that the years have increased by 36.9 percent due to major depressive disorders and in dysthymia by over 40 percent. These are illustrated in Table 62–7.

Table 62–2.
Twelve-Month Prevalence Estimates of DSM-IV/CIDI Disorders in the WMH Surveys[a,b]

	Any Anxiety Disorder		Any Mood Disorder		Any Externalizing Disorder		Any Substance Disorder		Any Disorder	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)
I. WHO Region: Pan American Health Organization (PAHO)										
Colombia	14.4	(1.0)	7.0	(0.5)	4.4	(0.4)	2.8	(0.4)	21.0	(1.0)
Mexico	8.4	(0.6)	4.7	(0.3)	1.6	(0.3)[c]	2.3	(0.3)	13.4	(0.9)
United States	19.0	(0.7)	9.7	(0.4)	10.5	(0.7)	3.8	(0.4)	27.0	(0.9)
II. WHO Region: African Regional Office (AFRO)										
Nigeria	4.2	(0.5)	1.1	(0.2)	0.1	(0.0)[d,e]	0.9	(0.2)	6.0	(0.6)
South Africa	8.2	(0.6)[f,g]	4.9	(0.4)[h]	1.9	(0.3)[d,i,e]	5.8	(0.5)	16.7	(1.0)
III. WHO Region: Eastern Mediterranean Regional Office (EMRO)										
Lebanon	12.2	(1.2)	6.8	(0.7)	2.6	(0.7)[e]	1.3	(0.8)	17.9	(1.7)
IV. WHO Region: European Regional Office (EURO)										
Belgium	8.4	(1.4)	5.4	(0.5)[h]	1.7	(1.0)[c]	1.8	(0.4)[j]	13.2	(1.5)
France	13.7	(1.1)	6.5	(0.6)[h]	2.4	(0.6)[c]	1.3	(0.3)[j]	18.9	(1.4)
Germany	8.3	(1.1)	3.3	(0.3)[h]	0.6	(0.3)[c]	1.2	(0.2)[j]	11.0	(1.3)

Israel	3.6	(0.3)[f,g]	6.4	(0.4)	—	—[c,d,i,e]	1.3	(0.2)	10.0	(0.5)
Italy	6.5	(0.6)	3.4	(0.3)[h]	0.4	(0.2)[c]	0.2	(0.1)[j]	8.8	(0.7)
Netherlands	8.9	(1.0)	5.1	(0.5)[h]	1.9	(0.7)[c]	1.9	(0.3)[j]	13.6	(1.0)
Spain	6.6	(0.9)	4.4	(0.3)[h]	0.5	(0.2)[c]	0.7	(0.2)[j]	9.7	(0.8)
Ukraine	6.8	(0.7)[f,g]	9.0	(0.6)[h]	5.7	(1.0)[d,e]	6.4	(0.8)	21.4	(1.3)
V. WHO Region: Western Pacific Regional Office (WPRO)										
PRC[k]	3.0	(0.5)	1.9	(0.3)	3.1	(0.7)[d,e]	1.6	(0.4)	7.1	(0.9)
Japan	4.2	(0.6)[f]	2.5	(0.4)	0.2	(0.1)[d,i,e]	1.2	(0.4)	7.4	(0.9)
New Zealand	15.0	(0.5)[f]	8.0	(0.4)	—	—[c,d,i,e]	3.5	(0.2)	20.7	(0.6)

[a]See the notes to Table 62-1 for a listing of the disorders included in each entry.

[b]Impulse disorders restricted to age ≤39 (China, Ukraine, Nigeria) or to age ≤44 (all other countries).

[c]Intermittent explosive disorder was not assessed.

[d]Attention-deficit/hyperactivity disorder was not assessed.

[e]Oppositional-defiant disorder was not assessed.

[f]Adult separation anxiety disorder was not assessed.

[g]Specific phobia was not assessed.

[h]Bipolar disorders were not assessed.

[i]Conduct disorder was not assessed.

[j]Only alcohol abuse with or without dependence were assessed. No assessment was made of other Drug Abuse with or without dependence.

[k]People's Republic of China.

Table 62–3.
Association between Severity of 12-Month DSM-IV/CIDI Disorders and Days Out of Role in the WMH Surveys

	Serious Mean	Moderate (SE)	Mild Mean	Wald F[a] (SE)	(*p*-value) Mean	(SE)		
I. WHO Region: Pan American Health Organization (PAHO)								
Colombia	53.0	(8.9)	33.7	(6.7)	15.6	(3.0)	10.8[b]	(<0.
Mexico	42.8	(6.9)	26.3	(5.3)	11.7	(2.7)	11.7[b]	(<0.
United States	135.9	(6.9)	65.3	(4.6)	35.7	(2.7)	126.1[b]	(<0.
II. WHO Region: African Regional Office (AFRO)								
Nigeria	56.7	(22.3)	51.5	(18.8)	25.9	(7.4)	1.6	(0.2
South Africa	73.1	(9.7)	49.3	(6.5)	32.5	(4.8)	9.1[b]	(<0.
III. WHO Region: Eastern Mediterranean Regional Office (EMRO)								
Lebanon	81.4	(10.6)	42.0	(9.5)	13.6	(5.4)	14.4[b]	(<0.
IV. WHO Region: European Regional Office (EURO)								
Belgium	96.1	(26.0)	59.9	(11.6)	42.5	(9.6)	3.7[b]	(0.0
France	105.7	(14.3)	71.8	(16.5)	67.6	(17.3)	2.7	(0.0

Germany	77.8	(18.1)	33.2	(8.2)	45.7	(12.1)	2.2	(0.1
Israel	184.6	(12.5)	109.4	(10.1)	44.6	(9.1)	41.8[b]	(<0.
Italy	178.5	(25.6)	55.6	(10.9)	41.7	(11.2)	11.7[b]	(<0.
Netherlands	140.7	(19.9)	87.1	(17.1)	68.9	(22.7)	4.0[b]	(0.0
Spain	131.5	(15.8)	56.6	(10.0)	57.4	(22.0)	8.1[b]	(<0.
Ukraine	142.5	(14.5)	103.2	(9.2)	51.6	(9.9)	13.9[b]	(<0.
V. WHO Region: Western Pacific Regional Office (WPRO)								
People's Republic of China	48.7	(18.4)	21.1	(5.2)	21.3	(7.2)	1.5	(0.2
Japan	51.0	(17.3)	39.3	(10.6)	22.5	(6.4)	3.7[b]	(0.0
New Zealand	98.1	(5.9)	54.6	(3.4)	36.4	(3.6)	40.7[b]	(<0.

[a]No demographic controls were used.

[b]Significant association between severity and days out of role at the 0.05 level.

Table 62–4.
Prevalence of 12-Month DSM-IV/CIDI Disorders by Severity in the WMH Surveys[a]

	Serious		Moderate		Mild	
	%	(SE)	%	(SE)	%	(SE)
I. WHO Region: Pan American Health Organization (PAHO)						
Colombia	23.1	(2.1)	41.0	(2.6)	35.9	(2.1)
Mexico	25.7	(2.4)	33.9	(2.2)	40.5	(2.6)
United States	25.2	(1.4)	39.2	(1.2)	35.7	(1.4)
II. WHO Region: African Regional Office (AFRO)						
Nigeria	12.8	(3.8)	12.5	(2.6)	74.7	(4.2)
South Africa	25.7	(1.8)	31.5	(2.2)	42.8	(2.2)
III. WHO Region: Eastern Mediterranean Regional Office (EMRO)						
Lebanon	22.4	(3.1)	42.6	(4.7)	35.0	(5.5)
IV. WHO Region: European Regional Office (EURO)						
Belgium	31.8	(4.2)	37.8	(3.3)	30.4	(4.8)
France	18.5	(2.5)	42.7	(3.0)	38.8	(3.6)
Germany	21.3	(2.5)	42.6	(4.6)	36.1	(4.3)
Israel	36.8	(2.4)	35.2	(2.3)	28.0	(2.1)
Italy	15.9	(2.7)	47.6	(3.8)	36.5	(3.9)
Netherlands	30.7	(3.4)	31.0	(3.7)	38.3	(4.6)
Spain	19.3	(2.4)	42.3	(4.0)	38.4	(4.7)
Ukraine	22.9	(1.8)	39.4	(2.9)	37.7	(3.5)
V. WHO Region: Western Pacific Regional Office (WPRO)						
People's Republic of China	13.8	(3.7)	32.2	(4.9)	54.0	(4.6)
Japan	13.2	(3.1)	45.5	(5.3)	41.3	(4.6)
New Zealand	25.3	(1.0)	40.8	(1.4)	33.9	(1.2)

[a]See the text for a description of the coding rules used to define the severity levels.

Table 62–5.
Disorder-Specific Global Sheehan Disability Scale Ratings for Commonly Occurring Mental and Chronic Physical Disorders in Developed and Developing WMH Countries

	Proportion Rated Severely Disabling			
	Developed		Developing	
	%	(SE)	%	(SE)
I. Physical disorders				
Arthritis	23.3	(1.5)	22.8	(3.0)
Asthma	8.2	(1.4)	26.9	(5.4)
Back/neck	34.6	(1.5)	22.7	(1.8)
Cancer	16.6	(3.2)	23.9	(10.3)
Chronic pain	40.9	(3.6)	24.8	(3.8)
Diabetes	13.6	(3.4)	23.7	(6.1)
Headaches	42.1	(1.9)	28.1	(2.1)
Heart disease	26.5	(3.9)	27.8	(5.2)
High blood pressure	5.3	(0.9)	23.8	(2.6)
Ulcer	15.3	(3.9)	18.3	(3.6)
II. Mental disorders				
ADHD	37.6	(3.6)	24.3	(7.4)
Bipolar	68.3	(2.6)	52.1	(4.9)
Depression	65.8	(1.6)	52.0	(1.8)
GAD	56.3	(1.9)	42.0	(4.2)
IED	36.3	(2.8)	27.8	(3.6)
ODD	34.2	(6.0)	41.3	(10.3)
Panic disorder	48.4	(2.6)	38.8	(4.7)
PTSD	54.8	(2.8)	41.2	(7.3)
Social phobia	35.1	(1.4)	41.4	(3.6)
Specific phobia	18.6	(1.1)	16.2	(1.6)

Global burden of disease was highest in substance use disorders and alcohol abuse (Tables 62–8 and 62–9).

As illustrated in Table 62–9, the burden for conditions like eating disorders is relatively high but often not recognized as such.

Lifetime contact for treatment delay in various psychiatric conditions is shown in Table 62–10. Interestingly, common mental disorders showed longer delay. This is further confirmed in the 12-month service use by sectors (see Table 62–11).

In the estimates for the year 2005, mental and neurological disorders accounted for 13.5 percent of all DALYs in the world (27.4 percent in high-

income countries, 17.7 percent in middle-income countries, and 9.1 percent in low-income countries), being the main contributor to burden among noncommunicable diseases (27.5 percent, compared to 22 percent for cardiovascular diseases and 11 percent for cancer).

According to the updated estimates for the year 2030, mental and neurological disorders will account for 14.4 percent of all DALYs in the world and 25.4 percent of those due to noncommunicable diseases. Depression will rank number 2 in the percent of total DALYs in that year (5.7 percent), following HIV/AIDS, and preceding ischemic heart disease (Table 62–12). It will be number 1 in high-income countries (9.8 percent), number 2 in middle-income countries (6.7 percent), and number 3 in low-income countries (4.7 percent).

The WHO estimates that one in four families worldwide has at least one member suffering from a mental disorder. The objective and subjective burden related to caring for people with severe mental disorders (in terms of disruption of family relationships; constraints in social, leisure, and work activities; financial difficulties; negative effect on physical health; feelings of loss, depression, and embarrassment in social situations; and the stress of coping with disturbing behaviors) have been reported to be substantial and significantly higher than that related to caring for people with long-term physical diseases such as diabetes, heart, kidney, or lung diseases. Cross-cultural differences have been reported in some dimensions of family burden: Relatives in the Mediterranean area, for example, compared with those in northern European countries were found to be more resigned and to use spiritual help more frequently as a coping strategy and reported a greater level of reduction in social interests and lower levels of support from their social network.

The prevalence of mental disorders among children and adolescents is difficult to evaluate because it is problematic to draw a reliable boundary between psychopathological phenomena and some manifestations that are part of normal development. Population surveys may help but access to schools and clinics is extremely variable, hence the accuracy of figures can be challenged. Many studies have used checklists compiled by parents and teachers, which are useful in detecting cases needing special attention but may not represent an appropriate basis for prevalence estimates. Community surveys carried out in low- and high-income countries focusing on children and adolescents between 1 and 15 years of age reported prevalence rates of mental disorders between 12 and 22 percent.

Suicide is among the 10 leading causes of death for all ages in most of the countries for which information is available. In some countries (e.g., China), it is the leading cause of death for people between 15 and 34 years of age. According to WHO estimates, about 849,000 people died from suicide worldwide in the year 2001. In that year, the number of suicide deaths overtook the number of deaths by violence (500,000) and war

(230,000). In 2020, approximately 1.53 million people will die from suicide worldwide based on current trends, and 10 to 20 times more people will attempt suicide. Reported suicide rates vary considerably across countries; for instance, annual suicide rates of 48.0 to 79.3 per 100,000 males have been reported in many Eastern and Central European countries, whereas rates of less than 4.0 per 100,000 males have been found in Islamic and several Latin American countries. More than 85 percent of suicides are reported to occur in low- or middle-income countries, but this figure may represent an underestimate due to the low reliability of official statistics in those countries: When surveillance with validated verbal autopsy was used in South India, the observed rates for suicide exceeded official national estimates 10-fold. In the Asia-Pacific region, 300,000 cases of suicide per year are estimated to occur by self-poisoning with pesticides.

Table 62–6.
Summary of Data Used to Calculate YLDs for Depressive Disorders

Parameter	MDD	Dysthymia	Source
Epidemiological inputs			
Number of data points (and studies)			Systematic review of the literature.
Prevalence	544 (116)	141 (36)	
Incidence	19 (4)[a]	3 (2)[a]	
Remission	—	3 (2)	
Duration	1 (5)[b]	—	
Excess-mortality	14 (11)	5 (2)[c]	
DisMod-MR point prevalence% (95% UI) and cases			DisMod-MR epidemiological modeling
Global prevalence	4.4% (4.1–4.7%); 298 million cases	1.55% (1.5–1.6%); 106 million cases	
Males	3.2% (3.0–3.6%); 111 million cases	1.3% (1.2–1.4%); 44 million cases	
Females	5.5% (5.0–6.0%);187 million cases	1.8% (1.7–1.9%); 62 million cases	
Disability weights			Derived by GBD core group and mental disorders expert group for the

			GBD 2010 disability weight survey.
Health state lay descriptions			
Mild	Has constant sadness and has lost interest in usual activities. The person can still function in daily life with extra effort, but sleeps badly, feels tired, and has trouble concentrating	—	
Moderate	Has constant sadness and has lost interest in usual activities. The person has some difficulty in daily life, sleeps badly, has trouble concentrating, and sometimes thinks about harming himself (or herself).	—	
Severe	Has overwhelming, constant sadness and cannot function in daily life. The person sometimes loses touch with reality and wants to harm or kill himself (or herself)	—	
Raw disability weights (95% UI)			GBD 2010 disability weight Survey.
Mild	0.16 (0.11–0.22)	0.16 (0.11–0.22)[d]	
Moderate	0.41 (0.28–0.55)		
Severe	0.66 (0.47–0.82)		
Severity distribution% (95% UI)			Based on SF–12 data from MEPS, NSMHWB, and NESARC.
Asymptomatic	13.9% (10.2–17.7%)	29.2% (24.9–33.6%)	
Mild MDD/Symptomatic dysthymia	58.8% (48.0–68.5%)	70.8% (66.4–75.1%)	
Moderate	16.5% (12.1–21.0%)		
Severe	10.8% (3.8–20.3%)		
Average disability weight (95% UI)	0.23 (0.18–0.30)	0.11 (0.07–0.15)	Based on severity proportions from MEPS, NSMHWB, and NESARC data, applied to weights from GBD 2010 disability weights survey.

[a]Incidence data were excluded for MDD and dysthymia as they were not consistent with the prevalence

and duration/remission data.

[b]The one data point for duration of 37.7 weeks was an estimate of average duration calculated from a best fit curve between the data points available from five studies.

[c]Both studies reported no elevated risk of mortality in those with dysthymia.

[d]The disability weight for mild-MDD was applied to dysthymia.

95% UI, 95% uncertainty interval.

Table 62–7.
Change in Depressive Disorder YLDs between 1990 and 2010

Total YLDs in 1990 and 2010	MDD	Dysthymia	Depressive Disorders
Total YLDs in 1990	46,138,600	7,870,700	54,009,300
Total YLDs in 2010	63,179,247	11,084,100	74,261,500
Total YLDs generated from 2010 population, 1990 population age structure, 1990 YLD rates (step 1)	59,904,870	10,067,939	69,972,809
Total YLDs generated from 2010 population, 2010 population age structure, 1990 YLD rates (step 2)	64,537,300	11,061,231	75,598,531
Total change in YLDs between 1990 and 2010	36.9%	40.8%	37.5%
Change in YLDs between 1990 and 2010 due to population growth	29.8%	27.9%	29.6%
Change in YLDs between 1990 and 2010 due to population aging	10.0%	12.6%	10.4%
Change in YLDs between 1990 and 2010 due to prevalence increase	−2.9%	0.3%	−2.5%

The difference between total YLDs in 1990 and YLDs at step 1 represents the change in YLDs due to population growth; the difference between YLDs at step 1 and YLDs at step 2 represents the change in YLDs due to population aging; the difference between total YLDs in 2010 and YLDs at step 2 represents the change in YLDs due to changes in prevalence.

Table 62–8.
Global Burden of Disease: Mental and Substance Use Disorders

	Mental Disorders	Substance Use Disorders	Before 1998	1998–2005	2006–2013	Total
All causes			100.0%	100.0%	100.0%	100.0%
Mental and substance use disorders			37.8%	58.5%	35.6%	67.6%
	Schizophrenia		17.0%	9.0%	3.7%	19.1%
	Alcohol use disorders		19.7%	28.7%	14.9%	31.4%
	Drug use disorders		20.7%	47.3%	26.1%	51.6%
		Opioid use disorders	12.8%	17.6%	2.7%	19.7%
		Cocaine use disorders	6.9%	31.9%	5.9%	34.6%
		Amphetamine	6.4%	23.9%	8.0%	27.7%

	use disorders				
	Cannabis use disorders	16.0%	42.0%	20.7%	46.8%
	Other drug use disorders	—	—	—	—
Depressive disorders		19.7%	23.9%	11.2%	33.0%
	Major depressive disorder	19.7%	23.9%	11.2%	33.0%
	Dysthymia	9.0%	13.8%	5.3%	18.6%
Bipolar disorder		8.5%	16.0%	3.7%	18.6%
Anxiety disorders		12.8%	21.8%	5.3%	26.1%
Eating disorders		10.6%	12.2%	4.3%	14.9%
	Anorexia nervosa	10.1%	12.2%	4.3%	14.4%
	Bulimia nervosa	8.5%	11.7%	3.2%	14.9%
Autistic spectrum disorders		5.3%	5.9%	3.7%	9.6%
	Autism	5.3%	5.3%	3.7%	9.6%
	Asperger syndrome	1.6%	4.8%	1.6%	5.3%
Attention-deficit or hyperactivity disorder		10.6%	10.1%	4.8%	19.1%
Conduct disorder		5.9%	6.4%	1.6%	11.2%
Idiopathic intellectual disability		6.4%	3.2%	1.1%	7.4%
Other mental and substance use disorders		0.5%	0.5%	0.0%	1.1%

Suicide rates are higher in males than in females (3.2:1 in 1950, 3.6:1 in 1995, and an estimated 3.9:1 in 2020). China is the only country in which suicide rates in females are consistently higher than those in males, especially in rural areas. Over the last few decades, suicide rates have been reported to be stable worldwide, but a rising trend among males of age 15 to 19 years has been observed. A systematic review covering 15,629 cases in the general population worldwide estimated that 98 percent of those who committed suicide had a diagnosable mental disorder, with mood disorders accounting for 35.8 percent, substance-related disorders for 22.4 percent, personality disorders for 11.6 percent, and schizophrenia for 10.6 percent of cases.

TREATMENT GAP AND PROJECTED POPULATION-LEVEL TREATMENT EFFECTIVENESS WORLDWIDE

The efficacy of pharmacological and psychosocial treatments for mood, anxiety, psychotic, and substance-related disorders has been convincingly proved by clinical trials carried out in low- and middle-income, as well as in high-income, countries. However, the treatment gap is substantial for all mental disorders worldwide, in particular, in low-income countries.

In the World Mental Health Surveys, carried out in representative community samples in 17 countries in Africa, the Americas, Asia, Europe, and the Middle East, the proportion of lifetime cases of anxiety disorders making treatment contact in the year of disorder onset ranged from .8 percent in Nigeria to 36.4 percent in Israel. The proportion of cases making treatment contact by 50 years ranged from 15.2 percent in Nigeria to 95.0 percent in Germany. The median delay to treatment among cases eventually making contact ranged from 3 years in Israel to 30 years in Mexico (Table 62–13, also see Table 62–10).

The proportion of lifetime cases of mood disorders making treatment contact in the year of disorder onset ranged from 6 percent in Nigeria and China to 52.1 percent in the Netherlands. The proportion of cases making treatment contact by 50 years ranged from 7.9 percent in China to 98.6 percent in France. The median delay to treatment among cases eventually making contact ranged from 1 year in Belgium, The Netherlands, Spain, and Japan to 14 years in Mexico (Table 62–14).

The proportion of lifetime cases of substance use disorders making treatment contact in the year of disorder onset ranged from .9 percent in Mexico to 18.6 percent in Spain. The proportion of cases making treatment contact by 50 years ranged from 19.8 percent in Nigeria to 86.1 percent in Germany. The median delay to treatment among cases eventually making contact ranged from 6 years in Spain to 18 years in Belgium (Table 62–15).

In the World Mental Health Surveys, failure and delays in treatment seeking were generally greater in low-income countries, older cohorts, men, and cases with earlier ages of onset. The earlier treatment contact of people with mood disorders might be partly due to the fact that these disorders have been targeted by educational campaigns and primary care quality improvement programs in several countries.

The Disease Control Priorities Project Report estimated the cost of a mental health care package consisting of outpatient treatment of schizophrenia and bipolar disorder with first-generation antipsychotics or mood stabilizers and adjuvant psychosocial interventions and treatment of depression and panic disorder in primary care with generic selective serotonin reuptake inhibitors. The cost per capita of the package was projected to be US$ 3 to 4 in sub-Saharan Africa and South Asia and US$ 7 to 9 in Latin America and the Caribbean. The addition of brief interventions by primary care physicians for high-risk alcohol users was

projected to cost a further US$ 0.04 per capita in sub-Saharan Africa and South Asia and up to US$ 0.36 per capita in Central Asia and Latin America. Treatment of common mental disorders was reported to be in the same range of cost-effectiveness as antiretroviral treatment for HIV/AIDS, secondary prevention of hypertension, or glycemic control for diabetes.

Table 62–9.
Global Burden of Disease Attributable to Mental and Substance Use Disorders: Findings from the Global Burden of Disease Study 2010

	All Ages DALYs (thousand)			DALYs (per 100 000)		
Causes	**1990**	**2010**	**% Change**	**1990**	**2010**	**% Change**
Mental and Behavioral Disorders	**129,377 (106,771–154,032)**	**176,626 (145,613–209,122)**	**36.5%**	**2440 (2,014–2,905)**	**2564 (2,113–3,035)**	**5.0%**
Schizophrenia	9760 (6,186–13,369)	14,400 (9160–19,752)	47.5%	184 (117–252)	209 (133–287)	13.5%
Alcohol use disorders	10,470 (7,173–14,644)	13,826 (9248–19,212)	32.1%	197 (135–276)	201 (134–279)	1.6%
Alcohol dependence	10,385 (7,086–14,556)	13,735 (9,164–19,108)	32.3%	196 (134–275)	199 (133–277)	1.8%
Fetal alcohol syndrome	85 (49–133)	91 (55–138)	6·9%	2 (1–3)	1 (1–2)	−17.7%
Drug use disorders	11,764 (8,388–15,468)	16,412 (11,836–21,583)	39.5%	222 (158–292)	238 (172–313)	7.3%
Opioid use disorders	4,812 (3,350–6,281)	7,170 (5,143–9,257)	49.0%	91 (63–118)	104 (75–134)	14.6%
Cocaine use disorders	800 (475–1,214)	1,085 (633–1,639)	35.7%	15 (9–23)	16 (9–24)	4.4%
Amphetamine use disorders	1,894 (1,067–2,955)	2,596 (1,460–3,957)	37.1%	36 (20–56)	38 (21–57)	5.5%
Cannabis use disorders	1,693 (1,105–2,418)	2,057 (1,348–2,929)	21.5%	32 (21–46)	30 (20–43)	−6.5%
Other drug use disorders	2,565 (1,583–3,817)	3,503 (2,108–5,170)	36.6%	48 (30–72)	51 (31–75)	5.1%
Unipolar depressive disorders	54,010 (40,381–	74,264 (55,670–	37.5%	1,019 (762–	1,078 (808–	5.8%

	68,450)	94,240)		1,291)	1,368)	
Major depressive disorder	46,139 (34,517–58,427)	63,179 (47,779–80,891)	36,9%	870 (651–1,102)	917 (693–1,174)	5,4%
Dysthymia	7,871 (5,266–10,858)	11,084 (7,297–15,447)	40.8%	148 (99–205)	161 (106–224)	8.4%
Bipolar affective disorder	9,129 (5,757–13,169)	12,867 (8,084–18,654)	40.9%	172 (109–248)	187 (117–271)	8.5%
Anxiety disorders	19,664 (13,868–26,820)	26,826 (18,779–36,795)	36.4%	371 (262–506)	389 (273–534)	5.0%
Eating disorders	1,120 (749–1,554)	1,956 (1316–2,742)	74.6%	21 (14–29)	28 (19–40)	34.3%
Anorexia nervosa	95 (65–136)	188 (125–265)	97.1%	2 (1–3)	3 (2–4)	51.7%
Bulimia nervosa	1,025 (687–1,417)	1,768 (1,183–2,480)	72.5%	19 (13–27)	26 (17–36)	32.7%
Pervasive development disorders	5,918 (4,133–8,130)	7,666 (5,355–10,565)	29.5%	112 (78–153)	111 (78–153)	−0.3%
Autism	3,088 (2,119–4,260)	4,007 (2,752–5,563)	29.8%	58 (40–80)	58 (40–81)	−0.2%
Asperger syndrome	2,830 (1,917–4,016)	3,659 (2,463–5,150)	29.3%	53 (36–76)	53 (36–75)	−0.5%
Childhood behavioral disorders	5,472 (3,277–8,359)	6,245 (3785–9,347)	14.1%	103 (62–158)	91 (55–136)	−12.2%
Attention-deficit/hyperactivity disorder	424 (244–667)	491 (280–775)	15.8%	8 (5–13)	7 (4–11)	−10.9%
Conduct disorder	5,047 (2,960–7,840)	5,753 (3,428–8,748)	14.0%	95 (56–148)	84 (50–127)	−12.3%
Idiopathic intellectual disability	1,247 (746–1,924)	1,043 (572–1,687)	−16.4%	24 (14–36)	15 (8–24)	−35.7%
Other mental and behavioral disorders	822 (485–1,307)	1,121 (661–1,774)	36.4%	16 (9–25)	16 (10–26)	5.0%

Table 62–10.

Lifetime Treatment Contacts for Mental Disorders among Survey Participants, by Contacts before or after Disorder Onset[a]

		Treatment Contact									
		Lifetime			Before Onset		In Year of Onset		At Least 1 Year after Onset		
Mental Disorder	N	N	%	95% CI	%	95% CI	%	95% CI	%	95% CI	Me
Any mood disorder	1,400	1,173	81.8	79.0–84.3	9.9	7.7–12.5	58.3	54.5–62.0	31.9	28.3–35.7	2
Major depression	1,309	1,104	82.3	79.2–85.0	9.6	7.5–12.3	59.3	55.4–63.0	31.1	27.5–35.0	2
Dysthymia	86	78	91.4	80.1–96.5	13.4	6.6–25.1	32.5	20.0–48.1	54.1	37.9–69.5	6
Bipolar disorder	82	63	75.3	61.2–85.4	18.0	5.9–43.8	46.6	31.2–62.7	35.4	20.8–53.3	2
Any anxiety disorder	1,335	805	60.9	57.3–64.3	3.2	1.9–5.5	25.5	22.2–29.1	71.2	67.2–75.0	11
Panic disorder	261	232	90.1	84.4–93.9	2.1	.8–5.5	60.6	53.4–67.3	37.3	30.6–44.6	3
Agoraphobia	70	50	71.0	59.4–80.4	1.4	.2–10.5	42.4	26.4–60.1	56.2	38.7–72.4	6
Social phobia	619	308	49.7	44.9–54.5	1.6	.6–4.2	14.2	9.7–20.4	84.2	78.1–88.8	13
Specific phobia	552	229	40.6	35.7–45.7	2.8	.5–13.9	11.0	6.6–17.8	86.1	77.8–91.7	18
Generalized anxiety disorder	304	251	83.8	77.9–88.3	10.2	6.4–16.0	45.8	36.8–55.1	44.0	34.2–54.2	5
Any substance use disorder	1,128	201	16.7	14.2–19.5	11.3	7.5–16.7	25.5	19.5–32.6	63.1	55.1–70.5	4
Alcohol abuse	825	60	6.5	4.9–8.4	16.0	7.5–31.0	35.1	22.8–49.8	48.9	34.3–63.6	3
Alcohol dependence	113	55	36.9	24.7–51.1	8.6	2.5–25.3	22.3	8.2–48.1	69.1	45.6–85.7	5
Drug abuse	225	51	21.2	15.0–29.1	21.9	12.4–35.7	27.5	17.8–39.8	50.6	36.0–65.1	5
Drug dependence	107	61	56.5	44.6–67.8	10.2	4.4–22.0	21.3	12.5–33.9	68.5	53.7–80.3	2
Any impulse-control disorder[c]	144	73	55.3	45.0–65.2	7.7	1.4–32.4	8.7	3.6–19.8	83.5	65.1–93.2	7
Attention-deficit hyperactivity disorder	74	48	67.7	52.0–80.3	9.2	1.2–45.8	2.9	.6–12.2	87.9	57.2–97.5	8
Oppositional defiant	80	29	42.3	30.4–55.2	3.3	.5–19.1	29.1	11.2–57.3	67.5	35.6–88.7	5

disorder											
Any mental disorder	2,777	1,636	56.0	53.6–58.3	5.5	4.3–7.1	30.8	28.1–33.7	63.7	60.6–66.6	9

[a]Number of participants are unweighted, but percentages, means, and medians are weighted.

[b]Treatment delay is expressed as years between disorder onset and initial treatment contact. Delay was set at zero years for persons with initial treatment contact before disorder onset or in the year of disorder onset.

[c]Measured for respondents aged 18–44 only.

RESOURCES FOR MENTAL HEALTH CARE WORLDWIDE

According to the *Mental Health Atlas 2005,* only 62.1 percent of countries worldwide, accounting for 68.3 percent of the world population, have a mental health policy (i.e., a document of the government or ministry of health specifying the goals for improving the mental health situation of the country, the priorities among those goals, and the main directions for attaining them). A mental health policy is present in 58.8 percent of low-income and 70.5 percent of high-income countries. In Africa, only 50 percent of countries have a mental health policy. In Southeast Asia, only 54.5 percent of countries have a mental health policy, and 76.4 percent of the population is not covered by such a policy (Table 62–16). This difference in mental health policy reflects number of factors. This could be attributed to low importance to mental health in many countries and settings. It also has been explained on the basis of colonial rule where the laws have not been changed even after independence.

Community care facilities exist in only 68.1 percent of the countries (51.7 percent of low-income and 93 percent of high-income countries). In Southeast Asia, they are present in 50 percent of the countries, and in Africa in 56.5 percent of the countries. Only 60.9 percent of the countries report providing treatment facilities for severe mental disorders at the primary care level (55.2 percent of low-income and 79.5 percent of high-income countries). About one-fourth of low-income countries do not provide even basic antidepressant medications in primary care settings. In many others, the supply does not cover all regions of the country or is very irregular. Because medicines are often not available in health care facilities, patients and families are forced to pay for them out of pocket.

Whereas 61.5 percent of European countries spend more than 5 percent of their health budget in mental health care, 70 percent of countries in Africa and 50 percent of countries in Southeast Asia spend less than 1 percent. Out-of-pocket payment is the most important method of financing mental health care in 38.6 percent of countries in Africa and 30 percent of countries in Southeast Asia, whereas it is not the primary method of financing mental health care in any European country (Table 62–17). All countries with out-of-pocket payment as the dominant method of financing mental health care belong to low-income or lower middle-income

categories, while almost all countries with social insurance as the dominant method of financing belong to high-income or upper middle-income categories.

The median of psychiatric beds per 10,000 population ranges from 0.33 in Southeast Asia and .34 in Africa to 8 in Europe. Globally, 72 percent of psychiatric beds are located in mental hospitals and the rest in other settings, including psychiatric units in general hospitals and community services. The percentage of psychiatric beds located in mental hospitals is 74 percent in low-income and 83 percent in middle-income countries, compared with 64 percent in high-income countries. The lower figures for high-income countries reflect the trend toward deinstitutionalization in those countries. Of the psychiatric beds available in Europe, more than 80 percent are located in mental hospitals in low- and middle-income countries of Central and Eastern Europe.

The median number of psychiatrists per 100,000 population ranges from .04 in Africa and .2 in Southeast Asia to 9.8 in Europe (Table 62–18). It is .1 in low-income countries, compared to 9.2 in high-income countries. Two-thirds of low-income countries have less than 1 psychiatrist per 100,000 population. Chad, Eritrea, and Liberia (with populations of 9, 4.2 and 3.5 million, respectively), each has just 1 psychiatrist per 100,000 population. Afghanistan, Rwanda, and Togo (with populations of 25, 8.5, and 5 million, respectively), each has just 2 psychiatrists per 100,000 population. Large-scale migration of psychiatrists from low- and middle-income to high-income countries, as part of the larger picture of migration of health professionals in general, has been consistently documented. India and some sub-Saharan African countries are the most important contributors to the mental health workforce in the United Kingdom, although the United Kingdom has 110 psychiatrists per million population, whereas India has 2 per million and sub-Saharan Africa less than 1 per million.

The median number of psychiatric nurses per 100,000 population ranges from .1 in Southeast Asia and .2 in Africa to 24.8 in Europe. Nearly 47 percent of low-income countries have less than 1 psychiatric nurse per 100,000 population. The median number of psychologists working in mental health care per 100,000 population ranges from 0.03 in Southeast Asia and 0.05 in Africa to 3.1 in Europe. Approximately 69 percent of low-income countries have less than 1 psychologist per 100,000 population.

From these figures, it is clear that resources for mental health care are grossly inadequate in comparison to the needs, and that inequalities across countries are substantial, especially between low- and high-income countries. Moreover, resources tend to be concentrated in urban areas, especially in low-income countries, leaving vast regions without any form of mental health care. What this also indicates that the mental health professionals need to think outside the box and train primary care

physicians, community midwives, and health visitors as well as other professionals to screen and manage common mental disorders.

Even worse is the situation concerning child and adolescent mental health care. According to the WHO, only 7 percent of countries worldwide have a specific child and adolescent mental health policy. In less than one-third of all countries is it possible to identify an institution or a governmental entity with an overall responsibility for child mental health. School-based services are almost exclusively present in high-income countries, and even in Europe only 17 percent of countries have a sufficient number of these services. There are no pediatric beds for mental health identified in low-income countries, whereas such beds are identified in 50 percent of high-income countries. In all African countries outside South Africa, fewer than 10 psychiatrists could be found who were trained to work with children. In European countries, the number of child psychiatrists ranges from 1 per 5,300 to 1 per 51,800. In more than 70 percent of countries worldwide, there is no list of essential psychotropic medications for children. In 45 percent of countries worldwide, psychostimulants are either prohibited or unavailable for use in children with attention-deficit/hyperactivity disorder.

PRINCIPLES FOR MENTAL HEALTH PROGRAM DEVELOPMENT WORLDWIDE

According to the WHO, the development of mental health programs worldwide should be guided by the following principles: (1) providing treatment in primary care, (2) making psychotropic medications available, (3) giving care in the community, (4) educating the public, (5) involving communities, families, and consumers, (6) establishing national policies and legislations, (7) developing human resources, (8) linking with other relevant sectors, (9) monitoring community mental health, and (10) supporting more research.

Providing treatment in primary care involves the inclusion of recognition and treatment of common mental disorders in training curricula of all health personnel, the provision of refresher training to primary care physicians, and the improvement in referral patterns in primary health care. Making psychotropic medications available includes ensuring the availability of essential drugs in all health care settings and, in high-income countries, providing easier access to newer psychotropic drugs under public or private treatment plans. Giving care in the community implies moving people with mental disorders out of prisons, downsizing mental hospitals and improving care within them, developing general hospital psychiatric units, and providing community care facilities.

Educating the public includes promoting public campaigns against stigma and discrimination; using the mass media to promote mental health, foster positive attitudes, and help to prevent mental disorders; and, in high-income countries, launching public campaigns for the recognition

and treatment of common mental disorders. Involving communities, families, and consumers includes supporting the formation of self-help groups; ensuring the representation of communities, families, and consumers in services and policy making; and foster advocacy initiatives.

Establishing national policies and legislations includes revising legislations on the basis of current knowledge and human rights considerations and increasing the budget for mental health care. Developing human resources involves training psychiatrists and psychiatric nurses, and, in middle- and high-income countries, creating national training centers for psychiatrists, psychiatric nurses, psychologists, and psychiatric social workers. Linking with other relevant sectors involves initiating school and workplace mental health programs, encouraging the activities of nongovernmental organizations in those areas, and, in high-income countries, initiating evidence-based mental health promotion programs in collaboration with other appropriate sectors.

Table 62–11.
Twelve-Month Service Use by Sectors in the WMH Surveys

Country Income Level	Any Treatment *N*	Mental Health Specialty %	General Medical SE	Among Respondents Human Services *N*	Among Respondents Using Services CAM %	Mental Health Specialty SE	General Medical *N*
Low							
Nigeria	57	1.6	0.3	5	0.1	0.1	42
Low-middle							
China	74	3.4	0.6	19	0.6	0.2	41
Colombia	217	5.5	0.6	126	3.0	0.4	82
South Africa	675	15.4	1.0	108	2.5	0.4	440
Ukraine	212	7.2	0.8	39	1.2	0.3	135
High-middle							
Lebanon	77	4.4	0.6	18	1.0	0.3	53
Mexico	240	5.1	0.5	121	2.8	0.3	92
High							
Belgium	187	10.9	1.4	96	5.2	0.7	147
France	272	11.3	1.0	111	4.4	0.5	214
Germany	183	8.1	0.8	100	3.9	0.6	102
Israel	421	8.8	0.4	215	4.4	0.3	169
Italy	141	4.3	0.4	55	2.0	0.3	107

Japan	92	5.6	0.9	43	2.4	0.5	47
Netherlands	202	10.9	1.2	105	5.5	1.0	141
New Zealand	1592	13.8	0.5	585	5.2	0.3	1122
Spain	375	6.8	0.5	200	3.6	0.4	249
USA	1477	17.9	0.7	738	8.8	0.5	773
χ^2_{16}	764.6 (<0.001)	679.6 (<0.001)	732.2 (<0.001)	262.9 (<0.001)	388.0 (<0.001)	232.4 (<0.001)	207.3 (<0.001)

Table 62–12.
Leading Causes of Disability-Adjusted Life Years (DALYs) Worldwide as Estimated for the Year 2030

Disease or Injury	Percent of Total
1. HIV/AIDS	12.1
2. Unipolar depressive disorders	5.7
3. Ischemic heart disease	4.7
4. Road traffic accidents	4.2
5. Perinatal conditions	4.0
6. Cerebrovascular disease	3.9
7. Chronic obstructive pulmonary disease	3.1
8. Lower respiratory infections	3.0
9. Hearing loss, adult onset	2.5
10. Cataracts	2.5

From Mathers CD, Loncar D. Projection of global mortality and burden of disease from 2002 to 2030. *PLoS Med.* 2006;3:2011. (Open Access)

Table 62–13.
Treatment Delay for Anxiety Disorders in the World Mental Health Surveys

	Making Treatment Contact in Year of Onset, % (SE)	Making Treatment Contact by Age 50 Years, % (SE)	Median Duration of Delay in Years (SE)
The Americas			
Colombia	2.9 (0.6)	41.6 (3.9)	26.0 (1.5)
Mexico	3.6 (1.1)	53.2 (18.2)	30.0 (5.1)
United States	11.3 (0.7)	87.0 (2.4)	23.0 (0.6)
Europe			
Belgium	19.8 (2.8)	84.5 (4.9)	16.0 (3.5)
France	16.1 (1.8)	93.3 (1.9)	18.0 (1.8)
Germany	13.7 (1.8)	95.0 (2.3)	23.0 (2.3)
Italy	17.1 (2.1)	87.3 (8.5)	28.0 (2.2)
Netherlands	28.0 (3.7)	91.1 (2.8)	10.0 (1.6)
Spain	23.2 (2.0)	86.6 (5.2)	17.0 (3.2)

Africa and Middle East			
Israel	36.4 (0.9)	90.7 (1.3)	3.0 (0.1)
Lebanon	3.2 (1.1)	37.3 (11.5)	28.0 (3.9)
Nigeria	0.8 (0.5)	15.2 (2.6)	16.0 (4.2)
Asia and the Pacific			
Japan	11.2 (2.4)	63.1 (6.2)	20.0 (2.4)
China	4.2 (2.0)	44.7 (7.2)	21.0 (3.1)
Oceania			
New Zealand	12.5 (.8)	84.2 (2.5)	21.0 (0.8)

SE, standard error.

From Wang PS, Angermeyer M, Borges G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry.* 2007;6(3):177–185.

Table 62–14.
Treatment Delay for Mood Disorders in the World Mental Health Surveys

	Making Treatment Contact in Year of Onset, % (SE)	Making Treatment Contact by Age 50 Years, % (SE)	Median Duration of Delay in Years (SE)
The Americas			
Colombia	18.7 (2.7)	66.6 (3.7)	9.0 (1.6)
Mexico	16.0 (2.2)	69.9 (8.5)	14.0 (3.1)
United States	35.4 (1.2)	94.8 (2.5)	4.0 (0.2)
Europe			
Belgium	47.8 (2.7)	93.7 (2.5)	1.0 (0.3)
France	42.7 (2.1)	98.6 (1.4)	3.0 (0.3)
Germany	40.4 (3.8)	89.1 (5.0)	2.0 (0.4)
Italy	28.8 (3.0)	63.5 (5.9)	2.0 (0.5)
Netherlands	52.1 (2.9)	96.9 (1.7)	1.0 (0.3)
Spain	48.5 (2.3)	96.4 (3.1)	1.0 (0.3)
Africa and Middle East			
Israel	31.9 (0.8)	92.7 (0.5)	6.0 (0.3)
Lebanon	12.3 (2.0)	49.2 (5.2)	6.0 (2.1)
Nigeria	6.0 (1.7)	33.3 (7.2)	6.0 (3.3)
Asia and the Pacific			
Japan	29.6 (4.0)	56.8 (7.3)	1.0 (0.7)
China	6.0 (2.2)	7.9 (2.6)	1.0 (2.0)
Oceania			
New Zealand	41.4 (1.3)	97.5 (1.0)	3.0 (0.2)

SE, standard error.

From Wang PS, Angermeyer M, Borges G, et al. Delay and failure in treatment seeking after first onset of

mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry.* 2007;6(3):177–185.

Table 62–15.
Treatment Delay for Substance Use Disorders in the World Mental Health Surveys

	Making Treatment Contact in Year of Onset, % (SE)	Making Treatment Contact by Age 50 Years, % (SE)	Median Duration of Delay in Years (SE)
The Americas			
Colombia	3.6 (0.8)	23.1 (7.1)	11.0 (5.0)
Mexico	0.9 (0.5)	22.1 (4.8)	10.0 (3.3)
United States	10.0 (0.8)	75.5 (3.8)	13.0 (1.2)
Europe			
Belgium	12.8 (4.8)	61.2 (17.7)	18.0 (5.8)
France	15.7 (5.4)	66.5 (14.1)	13.0 (3.7)
Germany	13.2 (5.7)	86.1 (8.6)	9.0 (3.9)
Netherlands	15.5 (5.4)	66.6 (7.9)	9.0 (3.1)
Spain	18.6 (7.6)	40.1 (14.1)	6.0 (4.9)
Africa and Middle East			
Israel	2.0 (0.5)	48.0 (2.4)	12.0 (0.5)
Nigeria	2.8 (1.7)	19.8 (7.2)	8.0 (1.8)
Asia and the Pacific			
Japan	9.2 (5.1)	31.0 (7.8)	8.0 (4.6)
China	2.8 (1.8)	25.7 (9.0)	17.0 (3.7)
Oceania			
New Zealand	6.3 (0.8)	84.8 (15.4)	17.0 (1.3)

SE, standard error.

From Wang PS, Angermeyer M, Borges G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry.* 2007;6(3):177–185.

Table 62–16.
Presence of a Mental Health Policy in the Countries of Each Region of the World Health Organization (WHO)

WHO Region	Percent of Countries	Percent Population Coverage
Africa	50.0	69.4
Americas	72.7	64.2
Eastern Mediterranean	72.7	93.8
Europe	70.6	89.1
Southeast Asia	54.5	23.6

Western Pacific	48.1	93.8

From World Health Organization. *Mental Health Atlas 2005.* Geneva: World Health Organization; 2005.

Monitoring community mental health involves including mental disorders in basic health information systems, surveying high-risk population groups, instituting surveillance for specific disorders in the community, and, in high-income countries, monitoring the effectiveness of preventive programs. Supporting more research includes conducting studies in primary health care settings on the prevalence, course, outcome, and effect of mental disorders in the community; instituting effectiveness and cost-effectiveness studies for management of common mental disorders in primary health care; and, in high-income countries, carrying out research on the risk factors for mental disorders and on service delivery and investigating evidence on the prevention of mental disorders.

The guiding principles proposed for the prevention of suicide worldwide include (1) reducing access to means of suicide (e.g., pesticides and firearms), (2) treating people with mental disorders, (3) improving media portrayal of suicide, (4) training primary health care personnel, (5) implementing school-based programs, and (6) developing hot lines and crisis centers.

BARRIERS TO CHANGE

The most significant barriers to the implementation of the foregoing principles worldwide, according to the WHO, include the following: (1) some stakeholders may be resistant to the changes, (2) health authorities may not believe in the effectiveness of mental health interventions, (3) there may be no consensus among the country's stakeholders about how to formulate or implement the new policy, (4) financial and human resources may be scarce, (5) other basic health priorities may compete with mental health care for funding, (6) primary care teams may feel overburdened by their workload and refuse to accept the introduction of the new policy, and (7) many mental health specialists may not want to work in community facilities or with primary care teams, preferring to remain in hospitals. The suggested solutions include (1) adopting an "all-winners approach" that ensures that the needs of all stakeholders are taken into account, (2) developing pilot projects and evaluating their effect on health and consumer satisfaction, (3) asking for technical reports from international experts, (4) focusing the implementation of the mental health policy on a demonstration area and performing cost-effectiveness studies, (5) linking mental health programs to other health priorities, and (6) showing primary care practitioners that people with mental disorders are already a hidden part of their burden and that the burden will decrease if these disorders are identified and treated.

Table 62–17.
Countries in Which Out-of-Pocket Payment Is the Most Common Method of Financing Mental Health Care in Each Region of the World Health Organization (WHO)

WHO Region	Percent of Countries
Africa	38.6
Americas	12.9
Eastern Mediterranean	15.8
Europe	0
Southeast Asia	30.0
Western Pacific	18.5

From World Health Organization. *Mental Health Atlas 2005.* Geneva: World Health Organization; 2005.

Table 62–18.
Median Number of Mental Health Professionals Per 100,000 Population in Each Region of the World Health Organization (WHO)

WHO Region	Psychiatrists	Psychiatric Nurses	Psychologists Working in Mental Health
Africa	0.04	0.20	0.05
Americas	2.00	2.60	2.80
Eastern Mediterranean	0.95	1.25	0.60
Europe	9.80	24.80	3.10
Southeast Asia	0.20	0.10	0.03
Western Pacific	0.32	0.50	0.03

From World Health Organization. *Mental Health Atlas 2005.* Geneva: World Health Organization; 2005.

STIGMATIZING ATTITUDES TOWARD PEOPLE WITH MENTAL DISORDERS

Stigmatizing attitudes toward people with mental disorders are widespread in the general public and even among mental health professionals. Although it has been suggested that stigma may be less severe in Asian and African countries, a study carried out in India within the Stigma Programme of the World Psychiatric Association (WPA), in which 463 persons with schizophrenia and 651 family members were interviewed in four cities, reported that two-thirds of the respondents had suffered discrimination. Women and people living in urban areas were more stigmatized. Men experienced greater discrimination in the job area, whereas women experienced more problems in the family and social areas. In the family area, subtle discrimination in the form of decreased love, avoidance, rejection, and distance and excessive caution was frequently reported. In Ethiopia, among relatives of people with a diagnosis of schizophrenia or mood disorders, 75 percent reported that they had experienced stigma due to the presence of mental illness in the family, and 37 percent wanted to conceal the fact that a relative was ill. In a large

survey carried out in China among people with schizophrenia and their family members, one half of the respondents reported they had felt significantly stigmatized, and levels of stigma were found to be higher in urban areas. Stigmatization of people with mental disorders has been reported to be rare in Islamic countries.

Unlike people with physical disabilities, those with mental disorders are often perceived by the public to be in control of their disabilities and responsible for causing them. The view that "weakness," "laziness," or "lack of willpower" contribute to the development of mental disorders has been reported in several countries, including Turkey, Mongolia, and South Africa. The stigmatization of people with mental disorders may result in public avoidance, systematic discrimination, and reduced help-seeking behavior. In a survey carried out in 1996, in a probability sample of 1,444 adults in the United States, more than one half of the respondents reported to be unwilling to spend an evening socializing with, work next to, or have a family member marry a person with mental illness. Although most countries have some provision for disability benefits, people with mental illness are often specifically excluded from such entitlements. Moreover, mental disorders are frequently not considered in social and private insurance schemes for health care. Shame is reported to be one of the main barriers from seeking help for mental disorders in both developed and developing countries.

Strategies for addressing stigmatization of people with mental disorders have been subdivided into three groups: protest, education, and contact. There is some evidence that protest campaigns may be effective in reducing stigmatizing behaviors against people with mental disorders. This activist approach may be more workable in some countries rather than others. Education may promote a better understanding of mental illness, and educated people may be less likely to endorse stigma and discrimination. Even in rural settings in low- and middle-income countries strategies can be put into place successfully. There is no doubt that certain environments produce higher levels of stigma. These authors reported that internalized stigma and social relationships provide a risk factor for social isolation especially among patients with schizophrenia. Ungar et al. (2015) observe that various strategies are needed to deliver antistigma programs and health care systems themselves need to be included in such efforts. There is no doubt that stigma whether it is internal or external will influence help-seeking. Corrigan et al. (2014) provide a reminder that stigma remains a complex construct. The people who are affected by it and also demonstrate stigma include public, self, and structural components as mentioned above in health care structures. There is no doubt that directly affects people with mental illness, as well as their support system, provider network, and community resources. These authors also remind readers that cultures influence stigma too. It is likely that cultures see illnesses and idioms of

distress in different ways and stigma may be generated and seen in different ways. Internalized stigma too may be affected by cultures although. Pérez-Garín et al. (2015) point out that well-being can play a major role in mediating stigma. The definitions and understanding of well-being itself can be influenced by cultures. Stigma by professionals themselves deserves further exploration. Poreddi et al. (2015) report that stigma in medical students can be reduced after they have been exposed to psychiatric teaching and training, it has been suggested that public policy needs to change in order to reduce stigma.

An inverse relationship between having contact with a person with mental illness and endorsing stigmatizing behaviors has been documented. Several recent programs such as Time to Change in the United Kingdom have demonstrated that talking with patients with psychiatric patients does lead to a positive change in people's attitudes and changes attitudes in a more positive direction.

RELATIONSHIPS BETWEEN MENTAL AND PHYSICAL DISEASES

Mortality due to physical illness is significantly increased in people with severe mental disorders compared with the general population. In a follow-up study carried out in the United Kingdom, the standardized mortality ratio (SMR) for natural causes in people with schizophrenia was 2.32 (i.e., mortality was more than twice higher than in the general population). The SMR for causes "avoidable by appropriate treatment" was 4.68. The highest SMRs were those for endocrine, nervous, respiratory, circulatory, and gastrointestinal diseases. Increased nonsuicide all-cause mortality has been also reported for bipolar disorder (SMR 1.9 for males and 2.1 for females) and dementia (relative risk [RR] 2.63; 95 percent confidence interval [CI] 2.17 to 3.21). A meta-analysis of 15 population-based studies of the effect of a diagnosis of depression on subsequent all-cause mortality yielded a pooled odds ratio of 1.7 (CI 1.5 to 2.0). Evidence from low-income countries is limited, but a large population study conducted in Ethiopia found high mortality rates for major depression (SMR 3.55, 95 percent CI 1.97 to 6.39) and schizophrenia (almost 5 percent per year).

People with mental illness seem not to have benefited from the recent favorable trends concerning mortality due to some physical diseases. In a longitudinal study conducted in Australia, a steady decline in the mortality rate due to ischemic heart disease was observed in the general population during the period from 1980 to 1998, dropping from 209 to 143 per 100,000 person-years in men and from 139 to 117 per 100,000 person-years in women. In contrast, the mortality rate remained approximately constant in male psychiatric patients (on average 280 per 100,000 person-years) and increased from 153 to 234 per 100,000 person-years in female psychiatric patients.

The prevalence of several physical diseases is increased in people with

mental disorders compared with the general population. In a study carried out in the United States, people with psychotic disorders were found to be more likely than other people to develop diabetes, hypertension, heart disease, asthma, gastrointestinal disorders, skin infections, malignant neoplasms, and acute respiratory disorders. The rate was increased even when only patients without a concomitant substance use disorder were considered. In a study conducted in Nigeria, 55.2 percent of persons with schizophrenia-spectrum disorders referred for the first time to a psychiatric clinic had at least one physical disease, whereas in persons with neurotic disorders the rate was 11.8 percent. A strong prospective association has been documented between depression and coronary heart disease outcomes, including fatal myocardial infarction; on the other hand, the incidence of depression is increased after myocardial infarction, especially in the first month after the event. Depression also increases the risk of type II diabetes. In South Asia, an association between maternal perinatal depression and infant malnutrition and stunting at 6 months has been repeatedly reported.

People with severe mental disorders are at increased risk of contracting HIV infection, although prevalence rates vary substantially worldwide. Conversely, a large multicenter study conducted in sub-Saharan Africa, Asia, Latin America, Europe, and the United States reported a higher prevalence of depressive disorder among symptomatic HIV-seropositive people than in asymptomatic HIV-seropositive cases and seronegative controls. Evidence from both developed and developing countries shows that adherence to highly active antiretroviral therapy (HAART) is negatively affected by depression, cognitive impairment, and substance abuse.

The increased morbidity and mortality due to physical illness in people with severe mental disorders is first of all related to the increased prevalence in these people of several risk factors for cardiovascular, metabolic, and respiratory diseases, the most important of which are obesity, smoking, and alcohol and drug abuse.

In a study carried out in the United Kingdom, people with severe mental illness were significantly more likely to be obese (body mass index higher than 30) and morbidly obese (body mass index higher than 40) than the general population: The respective figures were 35.0 versus 19.4 percent and 3.7 versus 1.3 percent. When these figures were broken down by age and sex, 28.7 percent of men with severe mental illness aged between 18 and 44 years were obese, compared with 13.6 percent in the general population, and 3.7 versus .4 percent were morbidly obese. Even more striking were the figures concerning women of the same age: 50.6 versus 16.6 percent and 7.4 versus 2.0 percent, respectively.

In a meta-analysis of worldwide studies, a highly significant association between schizophrenia and current smoking was confirmed: The weighted

average odds ratio was 5.9; it was 7.2 in males and 3.3 in females. The association remained significant when controls with severe mental illness were used (odds ratio 1.9). Heavy smoking and high nicotine dependence were also more frequent in people with schizophrenia than in the general population.

Patients with severe mental illness have also been reported to make poorer dietary choices and to be less physically active than the general population, but the relevant research evidence is not strong.

The increased risk of HIV infection in people with severe mental disorders may be explained by the high rates of sexual contact with multiple partners, the low levels of condom use, and the high prevalence of injected drug use and sexual contact with partners using injected drugs.

The access of people with mental illness to physical health care is reduced with respect to the general population. In the aforementioned study conducted in Australia, the SMR for ischemic heart disease was found to be 1.78 in men with schizophrenia. However, hospital admission for that disease was remarkably less common in both men and women with schizophrenia compared with the general population. Moreover, both men and women with schizophrenia were more than threefold less likely to receive revascularization procedures than the general population.

The quality of physical health care received by patients with severe mental illness is often worse than the general population. A study conducted in the United States found that adverse events during medical and surgical hospitalizations were significantly more frequent in patients with schizophrenia than in the other people, including infections due to medical care, postoperative respiratory failure, postoperative deep venous thrombosis or pulmonary embolism, and postoperative sepsis. All of these adverse events were associated with a significantly increased odds of admission to an intensive care unit and of death.

The decreased access of people with mental disorders to medical services has been related to several factors concerning the health care system. The effect of lack of insurance and cost of care is well documented. In a study carried out in the United States, people with mental disorders were twice as likely as those without mental disorders to have been denied insurance because of a pre-existing condition (odds ratio 2.18). Having a mental disorder conferred a greater risk of having delayed seeking care because of cost (odds ratio 1.76) and of having been unable to obtain needed medical care (odds ratio 2.30).

Even when people with mental illness are seen by a doctor, their physical diseases often remain undiagnosed. Primary care providers may misperceive the medical complaints of people with mental disorders as “psychosomatic” or be unskilled or feel uncomfortable in dealing with this population. An underlying stigmatization may be involved. Moreover, during hospitalizations in medical and surgical wards, health care

professionals may not be experienced in dealing with the special needs of patients with schizophrenia, may minimize or misinterpret their somatic symptoms, and may make an inappropriate use of restraints or sedative drugs or fail to consider possible interactions of psychotropic drugs with other medications. On the other hand, many psychiatrists are unable or unwilling to perform physical and even neurological examinations or are not up to date on the management of even common physical diseases.

To address this situation, the first step is raising awareness of the problem among mental health care professionals, primary care providers, and patients with schizophrenia and their families. The research information about the increased morbidity and mortality due to physical illness in people with schizophrenia should be appropriately disseminated.

Education and training of mental health professionals and primary care providers is a further essential step. Mental health professionals should be trained to perform at least basic medical tasks. They should be educated about the importance of recognizing physical illness in people with severe mental disorders and encouraged to become familiar with the most common reasons for underdiagnosis or misdiagnosis of physical illness in these patients. On the other hand, primary care providers should overcome their reluctance to treat people with severe mental illness and learn effective ways to interact and communicate with them; it is not only an issue of knowledge and skills, but also most of all one of attitudes.

Another essential step is the development of an appropriate integration between mental health and physical health care. A well-identified professional should be responsible for physical health care in each person with a severe mental disorder. Mental health services should be able to provide at least a standard routine assessment of their patients to identify or at least suspect the presence of physical health problems. Guidelines about the management of patients receiving antipsychotic drugs should be known and applied by all mental health services. Patients should be involved as much as possible; for instance, mental health professionals should encourage patients to monitor and chart their weight. Dietary and exercise programs should be routinely provided by mental health services. Flexible smoking cessation programs, which have shown some degree of success, could be considered in some settings.

ETHICAL ISSUES IN MENTAL HEALTH CARE

The protection and promotion of the human rights of people with mental disorders are emerging as a priority worldwide. In 1991, the United Nations issued Resolution 46/119 for the Protection of Persons with Mental Illness and the Improvement of Mental Health Care. In that resolution, the human rights of people with mental disorders and their right to treatment were codified for the first time in a UN document. The 25 principles covered the following areas: Definition of mental illness; protection of confidentiality;

standards of care and treatment, including involuntary admission and consent to treatment; rights of persons with mental disorders in mental health facilities; protection of minors; provision of resources for mental health facilities; role of community and culture; review mechanisms providing for the protection of the rights of offenders with mental disorders; and procedural safeguards protecting rights of persons with mental disorders. National governments were invited to promote the principles of the resolution by appropriate legislative, juridical, administrative, educational, and other provisions. However, violations of the human rights of people with mental disorders are still reported in many countries, and in a substantial number of low- and middle-income countries patients in mental hospitals are physically restrained or secluded for long periods.

In Latin America, an influential document has been the Declaration of Caracas, adopted in 1990, by the Regional Conference on Restructuring Psychiatric Care in Latin America, which states that resources, care, and treatment for people with mental disorders should safeguard their dignity and human and civil rights and strive to maintain them people in their communities. The Declaration also states that mental health legislation should safeguard the human rights of people with mental disorders and that services should be organized so that these rights can be enforced.

In Africa, the Banjul Charter on Human and People's Rights, a legally binding document supervised by the African Commission on Human and People's Rights, in Article 5, addresses the right to respect for the dignity inherent in human beings and the prohibition of all forms of degradation, including cruel, inhuman, or degrading treatment.

According to the WHO, mental health legislation should cover the following issues: access to basic mental health care, least restrictive care, informed consent to treatment, voluntary and involuntary admission to treatment, competence issues, periodical review mechanism, confidentiality, rehabilitation, accreditation of professionals and facilities, and rights of families and caregivers. Specific mental health legislation in the field of mental health is present in 74 percent of low-income countries versus 92.7 percent of high-income countries.

In 1996, the WPA released the Madrid Declaration, which contains the ethical principles by which all national psychiatric societies are expected to abide. The Declaration includes seven general guidelines focusing on the aims of psychiatry: (1) psychiatrists must serve patients by providing the best therapy available consistent with accepted scientific knowledge and ethical principles and should devise therapeutic interventions that are the least restrictive to the freedom of the patient; (2) it is the duty of psychiatrists to keep abreast of scientific developments of their specialty and to convey updated knowledge to others; (3) the patient should be accepted as a partner by right in the therapeutic process, and the therapist–

patient relationship must be based on mutual trust and respect, to allow the patient to make free and informed decisions; (4) treatment must always be in the best interest of the patient, and no treatment should be provided against the patient's will, unless withholding treatment would endanger life of the patient and/or those who surround him or her; (5) when psychiatrists are requested to assess a person, it is their duty to inform the person being assessed about the purpose of the intervention; (6) information contained in the therapeutic relationship should be kept in confidence and used exclusively for the purpose of improving the mental health of the patient; and (7) because psychiatric patients are particularly vulnerable research subjects, extra caution should be taken to safeguard their autonomy, as well as their mental and physical integrity.

INTERNATIONAL ORGANIZATIONS ACTIVE IN THE MENTAL HEALTH FIELD

There are many international organizations active in the mental health field. They include the WHO, which is the world's leading public health agency; some professional associations, among which the largest is the WPA, representing the psychiatric profession worldwide; and several organizations with a membership of users/families (e.g., the World Fellowship for Schizophrenia and the Global Alliance of Mental Illness Advocacy Networks [GAMIAN]) or of both mental health professionals and users/families (e.g., the World Federation for Mental Health [WFMH]).

The WHO is a UN agency with 192 member states, grouped into six regions (Africa, the Americas, European Mediterranean, Europe, Southeast Asia, and Western Pacific). It has a Department of Mental Health and Substance Abuse at its headquarters in Geneva and advisors for mental health in each of its regional offices. Its main functions are direction and coordination of international health work and technical cooperation with countries. Among the numerous recent WHO activities in the mental health field, of special interest are the release of the *World Health Report 2001* and the development of the report *Mental Health: New Understanding, New Hope,* which was completely devoted to mental health. It provided a summary of the current and projected impact of mental disorders and of the principles of mental health policy and service provision, as well as a set of recommendations for future action that can be adapted to the needs and resources of the various countries. Project Atlas aims to collect information on mental health resources across the world. Global and regional analyses of those resources were first published in 2001 and updated in 2005. Volumes focusing on resources for child and adolescent mental health and on psychiatric education and training worldwide (the latter in collaboration with the WPA) were released in 2005.

The WPA is an association of national psychiatric societies aimed at increasing knowledge and skills necessary for work in the field of mental health and the care for the mentally ill. Its member societies number 134,

spanning 122 countries and representing more than 200,000 psychiatrists. The WPA organizes the World Congress of Psychiatry every 3 years. It also organizes international and regional congresses and meetings and thematic conferences. It has 65 scientific sections, which are aimed at disseminating information and promoting collaborative work in specific domains of psychiatry. It has produced several educational programs and series of books and consensus statements (including the Declaration of Madrid on ethical principles for psychiatric practice). It has an official journal, *World Psychiatry,* which is produced in English, Spanish, and Chinese and is indexed in PubMed and Current Contents and reaches more than 33,000 psychiatrists worldwide. Subsequently several books and information sheets have been published (a full list is available on the WPA website).

The WFMH is a multidisciplinary advocacy and education organization aimed at promoting the advancement of mental health awareness, prevention, advocacy, and best-practice recovery-focused interventions worldwide. Among its activities is the organization of the World Mental Health Day, observed every year on October 10, each time with a different theme.

FUTURE PERSPECTIVES

Several statements have been made by various groups and organizations about priorities for future action in the mental health field at the international level. Of special interest is a document produced by 39 leaders composing the so-called Lancet Global Mental Health Group. In this document, five main goals are identified: (1) placing mental health on the public health priority agenda, (2) improving the organization of mental health services, (3) integrating the availability of mental health into general health care, (4) developing human resources for mental health, and (5) strengthening public mental health leadership. It is good step but as mentioned above if it focuses only on one directional move it will create further barriers. It is imperative that local explanatory models and idioms of distress are taken into account so that culturally appropriate, culturally sensitive and culturally accessible models of service care are developed and delivered. Virtually in every country around the world there are insufficient human and material resources dedicated to mentally ill and mental illness. In many countries useful models of engaging the local community exist and it will be helpful to learn how these work. Furthermore, models using e-mental health and tele-mental health should be employed particularly in countries with large geographical areas.

Among the strategies proposed to place mental health on the public health priority agenda are the development and use of uniform and clearly understandable messages for mental health advocacy and the education of decision makers within governments and donor agencies on the evidence concerning the public health significance of mental disorders and the cost-

effectiveness of mental health care. Strategies suggested to improve the organization of mental health services include, among others, the provision of incentive arrangements to overcome vested interests blocking change and the organization of international technical support to learn from countries that have experienced successful mental health reform. To integrate the availability of mental health in general health care, it is proposed that mental health professionals be appointed and trained specifically for supporting and supervising primary health care staff. To promote the development of human resources for mental health, it is suggested that the professional and specialist workforce be increased and diversified, and that the quality of mental health training be improved to ensure that it is practical and occurs also in community or primary care settings. Governments are asked to identify and scale up a priority package of service interventions and to increase mental health budget allocations, with the aim of attaining a minimum level of investment within 10 years of US$ 2 per capita in low-income countries and US$ 3 to 4 per capita in lower middle-income countries. International donors and governments of high-income countries, in particular those that are beneficiaries of the brain drain, are asked to place mental health on their priority agenda for health assistance to low- and middle-income countries. Finally, several indicators are proposed to monitor progress worldwide. In the 2014 to 2017, WPA has launched an ambitious plan for the social justice for the mentally ill focusing on domestic gender-based interpersonal violence; child physical and emotional abuse; prisoner mental health care; underserved populations such as asylum seekers, migrants, and refugees; those with intellectual disabilities, lesbian, gay, bisexual, and transgender individuals and elderly individuals and finally developing mental health promotion. Interestingly, World Medical Association is also focusing on child abuse and gender-based violence and the two organizations are planning to work closely together. WPA is also working with other organizations to develop materials on perimental health as well as setting up WPA-Lancet commission on psychiatry to explore the roles and responsibilities of psychiatrists in the first quarter of the 21st century.

FUTURE DIRECTIONS

This chapter has provided an overview of the burden of mental disorders worldwide, the treatment gap and the scarcity of resources for mental health care especially in low- and middle-income countries, the WHO guidelines for the development of mental health programs and the obstacles to their implementation, and several other issues of interest for world psychiatry. The discussion has been based as much as possible on data, although part of the data on world mental health are projections that may under- or overestimate the magnitude of the problems. Moreover, the threshold for the diagnosis of common mental disorders (e.g., depression

and anxiety disorders) remains controversial, and although it can be argued that "subthreshold" conditions (not fulfilling current diagnostic criteria) might be a further neglected source of burden and disability, it can be also pointed out that some conditions fulfilling current diagnostic criteria might represent homeostatic reactions to adverse life events that are transient and do not require treatment. The threshold for the diagnosis of common mental disorders may be different in the various geographic and cultural contexts, making projections and international comparisons even more problematic. On the other hand, although several pharmacological and psychotherapeutic interventions have been proved to be efficacious in various regions of the world, their feasibility and acceptability may not be the same worldwide, and current treatment guidelines may not be relevant to people living in several low- and middle-income countries. The same may apply to rehabilitation programs and models of community mental health care. The fact is that, although more than 85 percent of the world's population lives in low- and middle-income countries, only a negligible proportion of research in the mental health field is conducted in those countries. Research-generated information is crucial to evaluate local needs, develop cost-effectiveness interventions, and monitor their implementation. Several initiatives are now addressing this research gap, and it can be hoped that a substantial revision of this chapter will be needed in the near future.

References

Carter R, Shimkets RP, Bornemann TH. Creating and changing public policy to reduce the stigma of mental illness. *Psychol Sci Public Interest*. 2014;15(2):35–36.

Corrigan PW, Druss BG, Perlick DA. The impact of mental illness stigma on seeking and participating in mental health care. *Psychol Sci Public Interest*. 2014;15(2):37–70.

Daumit GL, Pronovost PJ, Anthony CB, Guallar E, Steinwachs DM, Ford DE. Adverse events during medical and surgical hospitalizations for persons with schizophrenia. *Arch Gen Psychiatry*. 2006;63:267–272.

de Leon J, Diaz FJ. A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. *Schizophren Res*. 2005;76:135–157.

Dharitri R, Rao SN, Kalyanasundaram S. Stigma of mental illness: an interventional study to reduce its impact in the community. *Indian J Psychiatry*. 2015;57(2):165–173.

Ferrari AJ, Charlson FJ, Norman RE, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med*. 2013;10(11):e1001547.

Filik R, Sipos A, Kehoe PG, et al. The cardiovascular and respiratory health of people with schizophrenia. *Acta Psychiatr Scand*. 2006;113:298–305.

Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013.

Lancet. 2015;386:743–800. pii: S0140–6736(15)60692–4.

Kessler RC, Aguilar-Gaxiola S, Alonso J, et al. The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. *Epidemiol Psichiatr Soc*. 2009;18(1):23–33.

Kessler RC, Angermeyer M, Anthony JC, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Surveys. *World Psychiatry*. 2007;6(3):168–176.

Mathers CD, Loncar D. Projection of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3:e442.

Murthy RS. Stigma is universal but experiences are local. *World Psychiatry*. 2002;1(1):28.

Oliveira SE, Esteves F, Carvalho H. Clinical profiles of stigma experiences, self-esteem and social relationships among people with schizophrenia, depressive, and bipolar disorders. *Psychiatry Res*. 2015;229:167–173. pii: S0165–1781(15)00504–1.

Patel V, Boardman J, Prince M, Bhugra D. Returning the debt: how rich countries can invest in mental health capacity in developing countries. *World Psychiatry*. 2006;5:67–70.

Pérez-Garín D, Molero F, Bos AE. Internalized mental illness stigma and subjective well-being: the mediating role of psychological well-being. *Psychiatry Res*. 2015;228:325–331. pii: S0165–1781(15)00407–2.

Pinfold V, Toulmin H, Thornicroft G, Huxley P, Farmer P, Graham T. Reducing psychiatric stigma and discrimination: evaluation of educational interventions in UK secondary schools. *Br J Psychiatry*. 2003;182(4):342–346.

Poreddi V, Thimmaiah R, Math SB. Attitudes toward people with mental illness among medical students. *J Neurosci Rural Pract*. 2015;6(3):349–354.

Prince M, Patel V, Saxena S, et al. No health without mental health—a slogan with substance. *Lancet*. 2007;370:859–877.

Saxena S, Sharan P, Garrido M, Saraceno B. World Health Organization's Mental Health Atlas 2005: implications for policy development. *World Psychiatry*. 2006;5:179–184.

Stuart H, Arboleda-Flórez J, Sartorius N, eds. Special issue: stigma and mental disorders: international perspectives. *World Psychiatry*. 2005;4(Suppl 1).

ten Have M, de Graaf R, van Dorsselaer S, Beekman A. Lifetime treatment contact and delay in treatment seeking after first onset of a mental disorder. *Psychiatr Serv*. 2013;64(10):981–989.

Ungar T, Knaak S, Szeto AC. Theoretical and practical considerations for combating mental illness stigma in health care. *Community Ment Health J*. 2016;52(3):262–271.

Wang PS, Angermeyer M, Borges G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*. 2007;6(3):177–185.

Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet*. 2013;382(9904):1575–1586.

World Health Organization. *Mental Health Legislation and Human Rights*. Geneva: World Health Organization; 2003.

World Health Organization. *Atlas: Child and Adolescent Mental Health Resources*. Geneva: World Health Organization; 2005.

World Health Organization. *Mental Health Atlas 2005*. Geneva: World Health Organization; 2005.

World Health Organization. *Mental Health Policy, Plans and Programmes*. Geneva: World Health Organization; 2005.

World Health Organization. *Disease Control Priorities Related to Mental, Neurological, Developmental and Substance Abuse Disorders*. Geneva: World Health Organization; 2006.

World Psychiatric Association. *Madrid Declaration on Ethical Standards for Psychiatric Practice*. 1996. Available at http://www.worldpsychiatricassociation.org/content/madrid-ethic-engish.shtml.

Appendix A

History of Psychiatry Table

Persons and Events in Psychiatry

Person or Event	Country	Publications	Sigr
Hippocrates of Cos (ca. 460–370 BCE)		*Hippocratic Writings.* Baltimore, MD: Pelican Classics; 1978	Hel b h y n e c u c " d tl e c
Plato (427–347 BCE)	Greece	*The Dialogues of Plato.* Jowett B, trans. New York: Random House; 1937	In *T R* k ir lc r ir n
Aristotle (384–322 BCE)	Greece	Ross WD: *Aristotle.* 6th ed. New York: Routledge; 1995.	In *D* p d d e
Galen of Pergamon (130–200)	Asia Minor (Turkey) in the Roman Empire	*On the Affected Parts.* Siegel R, trans., ed. Basel: Karger; 1976	Con H A d a ir c
St. Augustine (354–430)	Tageste (Numidia, North Africa)	*Confessions.* Pusey EB, trans. New York: Modern Library; 1949; Wills G: *Saint Augustine.* New York: Lipper/Viking; 1999	His b p

Avicenna (980–1037)	Persia	*A Treatise on the Canon of Medicine of Avicenna.* London: Luzac; 1930	His r p c a C p
Constanitus Africanus (ca. 1010–1087)	Carthage, North Africa	Constantino l'Africano. *Della melancolia.* Rome; 1959	His o t n s d w
Petrarch (1304–1374)	Italy	*De remediis utriusque fortunae*	In h h t c d e J s t a
Henry Kramer and James Sprenger (15th century)	Germany	*Malleus maleficarum (Witches' Hammer*) (1486)	Infl p w
Paracelsus (ca. 1493–1541)	Einsiedeln (Switzerland), then traveled Europe	Pagel W: "Disease and the stars. 'The animal in man' and lunacy. The psychiatry of Paracelsus." In: *Paracelsus.* 2nd rev. ed. Basel: Karger; 1982:150	Opp b il i w s c (e t p i t
Juan Luis Vives (1493–1540)	Born in Spain; from the age of 16 yrs lived mainly in Holland and Belgium	*De anima et vita (Of Soul and Life)* (1538); Noreña CG. *Juan Luis Vives and the Emotions.* Carbondale, IL: Southern Illinois University Press; 1989	Des i a e r
Johann Weyer (1515–1588)	Holland	*De Praestigiis Daemonum (The Deception of Demons)* (1563); reprinted in Mora G, ed. *Witches, Devils, and Doctors in the Renaissance.* Binghampton, NY: Center for Medieval and Early Renaissance Studies; 1991	Crit r p w
Juan Huarte de San Juan (ca. 1530–1592)	Spain	*The Examination of Men's Wits* (1575)	An e i d

Giambattista Porta (1535–1615)	Italy	*De humana physiognomia* (1586)	Gav o
Felix Plater (1536–1614)	Switzerland	*Practice of Medicine* (1602); *Observations of Diseases Injurious to Body and Mind* (1614)	Nev b a d p d t p b p b n p h
Robert Burton (1577–1640)	England	*The Anatomy of Melancholy* (1621)	The p c p n t d li c
Paolo Zacchia (1584–1659)	Italy	*Questiones medico-legales* (1621–1635)	Hel t s r b f
Thomas Sydenham (1624–1689)	England	Dissertatio epistolaris … [1682]. In: *The Entire Works of Dr. Thomas Sydenham Newly Made English* (1742)	Gav o h t h c n b h t r p c c il
Thomas Willis (1621–1675)	England	*Two Discourses Concerning the Soul of Brutes* (1683)	Sum a il d w d s t

			a h s n
Georg Ernst Stahl (1660–1734)	Germany	*Theoria medica vera* (1707), *De animi morbis* (1708)	The *a* o d c o b
William Battie (1703–1776)	England	*A Treatise of Madness* (1758)	Firs i “ r fi ti
Boissier de Sauvages (1706–1767)	France	*Nosologia methodica* (1765)	Nos c a a r d
John Aiken (1747–1822)	England	*Thoughts on Hospitals* (1771)	Firs w w
Franz Anton Mesmer (1734–1815)	Austria, France	*Memoire sur la découverte du magnetism animal* (1779)	Sho t a c il o
Vincenzo Chiarugi (1759–1820)	Italy	*Regulations of the Hospitals of Santa Maria Nuova and of Bonifazio* (1789)	One t a w
William Cullen (1710–1790)	Scotland	*Nosology, or a Systematic Arrangement of Diseases* (1800)	A gr n t *n* d
Philippe Pinel (1745–1826)	France	*A Treatise on Insanity in Which Are Contained the Principles of a New and More Practical Nosology of Mental Disorders* (1801)	Clas f e t i c o
Johann Reil (1759–1813)	Germany	*Rhapsodies about the Application of Psychotherapy to Mental Disturbances* (1803)	The p t i n

			o u a p
Benjamin Rush (1745–1813)	United States	*Medical Inquiries and Observations upon the Diseases of the Mind* (1812); Farr C: Benjamin Rush and American psychiatry. *Am J Psychiatry.* 1994;151(Suppl):65	The p t t I A o t p
Thomas Sutton (1767–1835)	England	*Tracts on Delirium Tremens* (1813)	Firs d
William Tuke (1732–1822); Samuel Tuke (1784–1857), grandson of William Tuke; Daniel Hack Tuke (1827–1895), youngest son of Samuel Tuke; John Charles Bucknill (1817–1895)	England	Tuke S. *Description of the Retreat, an Institution near York, for Insane Persons of the Society of Friends.* York; 1813; Tuke DH, Bucknill JC. *A Manual of Psychological Medicine Containing the History, Nosology, Description, Statistics, Diagnosis, Pathology, and Treatment of Insanity.* London; 1858	W. R t C *D* i t E S T f t
Joseph Adams (1756–1818)	England	*A Treatise on the Supposed Hereditary Properties of Diseases … Particularly in Madness and Scrofula* (1814)	Firs p a d b d p p
Franz Joseph Gall (1758–1828), Johann Gaspar Spurzheim (1776–1832)	Austria, Germany	*The Physiognomical System of Drs. Gall and Spurzheim; founded on an anatomical and physiological examination of the nervous system in general and of the brain in particular* (1815)	The M d f p p p
Johann Christian Heinroth (1773–1843)	Germany	*Disturbances of the Mind* (1818); Hansen LA. Metaphors of mind and society: The origins of German psychiatry in the revolutionary era. *Isis.* 1998;89:387; Cauwenbergh LS. J Chr A Heinroth (1773–1843) a psychiatrist of the German Romantic era. *Hist Psychiatry.* 1991;2:365	Firs p t c p u a i t
Robert Gooch	England	*An Account of … Diseases Peculiar to Women* (1829)	Firs

(1784–1830)			p
James Cowles Pritchard (1786–1848)	England	*A Treatise on Insanity and Other Disorders Affecting the Mind* (1835)	Star p t d i *p*
Jean Etienne Dominique Esquirol (1782–1840)	France	*Mental Maladies: A Treatise on Insanity* (1838)	Coir d c n i d b c a la e
Isaac Ray (1807–1881)	United States	*Treatise on Medical Jurisprudence of Insanity* (1838)	Fou p
James Braid (1795–1860)	England	*Neurypnology; or, the Rationale of Nervous Sleep* (1843)	Enti fi b p
Daniel M'Naghton (1813–1865)	England	West DJ, Walk A, eds. *Daniel M'Naghton: His Trial and the Aftermath.* London Gaskell Books; 1977	Suff d n g a r i t r " i A a i
Wilhelm Griesinger (1817–1868)	Germany	*Mental Pathology and Therapeutics* (1845)	Pro d a b
Jacques-Joseph Moreau de Tours (1804–1884)	France	*Du haschich et de l'aliénation mentale* (1845)	Des ta th e p
Walter Cooper Dendy (1794–1871)	England	Psychotherapeia, or the remedial influence of mind. *J Psychol Med Ment Pathol.* 1853;6:268	Firs *p* (p d r p

			w b p
Jean-Pierre Falret (1794–1870), Jules Baillarger (1809–1890)	France	Baillarger J. La folie à double forme. *Bull Acad Imp Med.* 1853–1854;19:340; Falret JP. La folie circulaire. *Bull Acad Imp Med.* 1853–1854;19:382; Pichot P. The birth of the bipolar disorder. *Eur Psychiatry.* 1995;10:1	Bail b ir ir n *a*
Thomas Kirkbride (1809–1883)	United States	*On the Construction, Organisation, and General Arrangements of Hospitals for the Insane* (1854)	Set c c c
John Conolly (1794–1866)	England	*The Treatment of the Insane without Mechanical Restraints* (1856)	Con n tl tl c tı tl
George Robinson (1821–1875)	England	*On the Prevention and Treatment of Mental Disorders* (1859)	Intr b a il
Gustav Theodor Fechner (1801–1887)	Germany	*Elements of Psychophysics* (1860)	Esta b s r f p
Benedict-Augustin Morel (1809–1873)	France	*Traité des maladies mentales* (1860)	Pre a c d w t
Thomas Laycock (1812–1876)	England	*Mind and Brain* (1860)	Mer fı b d
Forbes B. Winslow (1810–1874)	England	*On Obscure Diseases of the Brain and Disorders of the Mind* (1860)	Firs d (l p
Ewald Hecker (1843–1909)	Germany	Die Hebephrenie. *Arch Pathol Anat Physiol.* 1871:52	Firs “ s p
Karl Kahlbaum (1828–1899)	Germany	*Die Katatonie oder das Spannungsirresein* (Berlin, 1874)	Firs (l d
George Miller Beard (1840–	United States	*A Practical Treatise on Nervous Exhaustion (Neurasthenia*) (1880)	Neu n

1883)			e d h p a f 1
Richard von Krafft-Ebing (1840–1902)	Germany	*Psychopathia sexualis* (1886); Oosterhuis H. *Stepchildren of Nature: Kraft-Ebing, Psychiatry, and the Making of Sexual Identity.* Chicago, IL: University of Chicago Press; 2000	Des h p s d t *n b* r s t
Pierre Janet (1859–1947)	France	*L'automatisme psychologique: Essai de psychologie experimentale sur les formes inferieures de l'active* (Paris, 1889); van der Kolk BA, van der Hart O. Pierre Janet and the breakdown of adaptation in psychological trauma. *Am J Psychiatry.* 1989;146:1530	Sho e t s d n e *p* o a s h i
Herman Emminghaus (1845–1904)	Germany	*Psychic Disturbances of Childhood* (1887)	Firs p
Jean-Martin Charcot (1825–1893)	France	Owen ARG. *Hysteria, Hypnosis, and Healing: The Work of JM Charcot.* New York: Garrett Publications; 1971; by Goetz C, Bonduelle M, Gelfand T. *Charcot: Constructing Neurology.* Oxford: Oxford University Press; 1995	In 1 p n U t a n i n a b c f c n a a h s t

			d h
Georges Gilles de la Tourette (1857–1904)	France	Etude sur une affection nerveuse caractérisée par l'incoordination motrice accompagnée d'echolalie et coprolalie. *Arch Neurol* (Paris). 1885; ix: 19, 158; Yorston G, Hindley N, trans. and intro: Study of a nervous disorder characterized by motor incoordination with echolalia and coprolalia. *Hist Psychiatry.* 1998;9:97; Kushner HI. *A Cursing Brain? The Histories of Tourette Syndrome.* Cambridge, MA: Harvard University Press; 1999	Des ir e n w H w n M tl le s T s
Cesare Lombroso (1835–1909)	Italy	*Genius and Insanity* (1864), *The Delinquent Man* (1876), *The Female Offender* (1893)	Emp ir w p p o b a s a e c
Emil Kraepelin (1856–1926)	Germany	Kraepelin E. *Psychiatrie: Ein Lehrbuch für Studerende und Aerzte.* 6th ed. (1899); *Memoirs.* Berlin: Springer-Verlag; 1987	Divi ir p t d d
John Hughlings Jackson (1834–1911)	England	*Selected Writings of John Hughlings Jackson* (London, 1931–1932)	Fro tl s fr a e r n
Sigmund Freud (1856–1939)	Austria	*The Interpretation of Dreams* (1900), *Three Essays on the Theory of Sexuality* (1905), *Introductory Lectures on Psychoanalysis* (1915–1917)	Disc o t p d a a f
Morton Prince (1854–1929)	United States	*The Dissociation of a Personality* (1905)	Earl p t n

			u
Clifford Beers (1876–1943)	United States	*A Mind That Found Itself* (1908)	Acc p s h U
Eugen Bleuler (1857–1939)	Switzerland	*Dementia Praecox or the Group of Schizophrenias* (1911)	Coir a
Alfred Binet (1857–1911)	France	*A Method of Measuring the Development of the Intelligence of Young Children,* with T. Simon (1911); English translation, 1913	Dev ir c p t c
Hideyo Noguchi (1876–1928)	United States	Noguchi H, Moore JW. A demonstration of the *Treponema pallidum* in the brain in cases of general paralysis. *J Exp Med.* 1913:17	The a c s g ti n k
Alfred Adler (1870–1937)	Austria	*Study of Organ Inferiority and Its Psychical Compensations* (1917); Stepansky PE. *In Freud's Shadow: Adler in Context.* Hillsdale, NJ: Analytic Press; 1983	Firs fi s p t *ir*
Hermann Rorschach (1884–1922)	Switzerland	*Psychodiagnostik* (1921)	Inkb u a tl p s o d
Julius von Wagner-Jauregg (1857–1940)	Austria	de Rudoli G. *Therapeutic Malaria* (1927); Wagner-Jauregg J. The history of the malaria treatment of general paresis. *Am J Psychiatry.* 1994;151(Suppl):231; Brown EM. Why Wagner-Jauregg won the Nobel Prize for discovering malaria therapy for general paresis of the insane. *Hist Psychiatry.* 2000;11:371	Dur s p r w w o p b p N p tr p
Otto Rank (1884–1939)	Austria, France, United States	*The Myth of the Birth of the Hero* (1909), *The Trauma of Birth* (1924), *Will Therapy* (1936); Lieberman EJ. *Acts of Will: The Life and Work of Otto Rank.* New York: Free	Mac t b

		Press, Macmillan; 1985	w i h w t e i n t t
American Board of Psychiatry and Neurology (ABPN)	United States	Freeman W, Ebaugh F, Boyd D Jr. The founding of the American Board of Psychiatry and Neurology, Inc. *Am J Psychiatry.* 1959;115:769	In 1 A n A n a s e p n q p o a h p p t F s a c a e s n w a t
Henry A. Murray (1893–1988)	United States	*Thematic Apperception Test.* Cambridge, MA: Harvard University Press; 1943	The T n c d p
Charles Bradley (1902–1979)	United States	Bradley C. The behavior of children receiving Benzedrine. *Am J Psychiatry.* 1937:577; Gross MD. Origin of stimulant use for treatment of attention deficit disorder. *Am J Psychiatry.* 1995;152:298	The t w a c n
Cure of pellagra	United States	Roe D. *A Plague of Corn: The Social History of Pellagra.* Ithaca, NY: Cornell University Press; 1973	Pell d d

			d w n a d G t S p a p d w s e la h c g c
Genetics of schizophrenia	Germany, United States	Rüdin E. *Studien über Verebung und Entstehung geistiger Störungen. I. Zur Vererbung und Neuetstehung der Dementia Praecox.* [*Studies on the Heredity and Genesis of Mental Diseases. I. On the Heredity and New Formation of Dementia Praecox*]. Berlin: Verlag von Julius Springer; 1916; Kallman FJ. *The Genetics of Schizophrenia.* New York: JJ Augustin; 1938; Rainer JD. The contributions of Franz J. Kallmann to the genetics of schizophrenia. In: Cancro R, Dean SR, eds. *Research in the Schizophrenic Disorders: The Stanley R. Dean Award Lectures.* Vol. 2. Jamaica, New York: Spectrum; 1985	Rüd K s w r t t e w N P s s s 1 s t h S t h it E a s w P h ti d p
ECT	Italy	Abrams R. The treatment that will not die. *Psychiatr Clin North Am.* 1994;17:525; Berrios G. The scientific origins of electroconvulsive therapy: A conceptual history. *Hist*	Two C L

		Psychiatry. 1997;8:105	u p a s w e t a d p n u
Karen Horney (1885–1952)	Germany until 1932, then United States	*The Neurotic Personality of Our Time* (1937), *New Ways in Psychoanalysis* (1939), *Neurosis and Human Growth: The Struggle Toward Self-Realization* (1950); Paris B. *Karen Horney: A Psychoanalyst's Search for Self-Understanding.* New Haven, CT: Yale University Press; 1994	Opp c w t s n n t w
Adolf Meyer (1866–1950)	Switzerland until 1893, then United States	Meyer A. A short sketch of the problems of psychiatry. *Am J Insanity.* 1896–1897;53: 538; Lidz T. Adolf Meyer and the development of American psychiatry. *Am J Psychiatry.* 1996;123:320; Adolf Meyer. In: Havens L, ed. *Approaches to the Mind.* Boston: Little, Brown.	Fro d A c v b u b p n e p n c h d
Constitution, personality, and mental illness	Germany, United States	Kretschmer E. *Body Build and Character* (1922); Sheldon WH. *Varieties of Human Physique* (1940)	Kret S p a p s s i a p e c a t t
Führer decree (1939)	Germany	Lifton RJ. The Nazi doctors. In: *"Euthanasia": Direct Medical Killing.* New York: Basic Books; 1986: Chapter 2;	In C t

		Gallagher HG. *By Trust Betrayed: Patients, Physicians, and the License to Kill in the Third Reich.* New York: Henry Holt; 1990; Annas GJ, Grodin MA, eds. *The Nazi Doctors and the Nuremburg Code. Human Rights in Human Experimentation.* Oxford: Oxford University Press; 1992	Il d h il t p e " t a n b p F *T G t* t w a n o b
MMPI	United States	Buchanan R. The development of the Minnesota Multiphasic Personality Inventory. *J Hist Behav Sci.* 1994;29:148	The H J. a M 1 w p n r M p a w t s c a
Leo Kanner (1894–1981)	Austria, United States	*Child Psychiatry.* Springfield, IL: CC Thomas; 1935; Autistic disturbances of affective contact. *Nerv Child.* 1943;2:217	The t a d la a
Melanie Klein (1882–1960), Anna Freud (1895–1982)	Central Europe, then England	Klein M. *The Psychoanalysis of Children* (1932); Kernberg O. Melanie Klein. In: Kaplan HI, Freedman AM, Sadock BJ, eds. *Comprehensive Textbook of Psychiatry.* 3rd ed. Baltimore, MD: Williams & Wilkins; 1980; Grosskurth P. *Melanie Klein: Her World and Her Work.* New York: Knopf; 1986; Freud A. *The Ego and the Mechanisms of Defence* (1936), *The Psychoanalytic Treatment of*	Klei p d p c s p

		Children (1946), A short history of child analysis. *Psychoanal Study Child.* 1966;21:7; Young-Bruehl E. *Anna Freud: A Biography.* New York: Summit Books; 1988	c c t b p f c p e p f d d " a
Helene Deutsch (1884–1982)	Europe until 1935, then United States	*The Psychology of Women,* 2 vols. (1945); some forms of emotional disturbance and their relationship to schizophrenia ('As If'). In: Deutsch H, ed. *Neuroses and Character Types.* New York: International Universities Press; 1965: Chapter 20; Roazen P. *Helene Deutsch: A Psychoanalyst's Life.* Garden City, NJ: Anchor Press/Doubleday; 1985	For c v w c if a w s r o
Ivan Petrovich Pavlov (1849–1936), Andrew Salter (1914–1996)	Russia, United States	Pavlov IP. *Lectures on Conditioned Reflexes.* 2 vols. New York: International Publishers; 1941; Babkin, BP. *Pavlov: A Biography.* Chicago: University of Chicago Press; 1949; Salter A. *Conditioned Reflex Therapy.* New York: Farrar, Straus; 1949	Fro s r a S c t p r c e t b d t
NMHA (1946), NIMH (1949)	United States	Grob GN. *From Asylum to Community: Mental Health Policy in Modern America.* Princeton, NJ: Princeton University Press; 1991: Chapters 2–4. Romano J. Reminiscences: 1938 and since. *Am J Psychiatry.* 1990;147:785	Con o A fi fu e le N r a h
Erik H. Erikson (1902–1994)	Europe; United States since	*Childhood and Society* (1950), *Identity and the Life Cycle* (1959); Friedman LJ. *Identity's Architect: A Biography of*	Res i

	1930s	*Erik H. Erikson* New York: Scribner; 1999; Wallerstein R, Goldberger L, eds. *Ideas and Identities: The Life and Work of Erik Erikson.* Madison, CT: International Universities Press; 1999	d i a c a *y i H t*
Carl Rogers (1902–1987)	United States	*Client-Centered Theory* (1951)	Sho t n p a t o c t b h n s p p p p
Donald Winnicott (1896–1971)	England	Winnicott D. Transitional objects and transitional phenomena [1951]. In: *Collected Papers: Through Paediatrics to Psychoanalysis.* London: Tavistock, 1958; Kohon H. *The British School of Psychoanalysis: The Independent Tradition* New Haven, CT: Yale University Press; 1986; Rodman FR. *Winnicott: Life and Work.* New York: Perseus; 2003	Obs o c w b a i o c n " B i o w e s
Jean Delay (1907–1987), Pierre Deniker (1917–1999)	France	Delay J, Deniker P. Le traitement des psychoses par une méthode neurolytique dérivée de l'hibernothérapie. *C R Congrés Med Alién Neurol France.* 1952;50:497; Thuillier J. *Ten Years That Changed the Face of Mental Illness.* London: Martin Dunitz; 1999; Healy D. *The Creation of Psychopharmacology;* Cambridge, MA: Harvard University Press; 2002	In t c p D p a c " c c w

			t p r a
Harry Stack Sullivan (1892–1949)	United States	Sullivan HS. *The Interpersonal Theory of Psychiatry.* Perry HS, Gawel ML, eds. New York: WW Norton; 1953; Perry HS. *Psychiatrist of America: The Life of Harry Stack Sullivan.* Cambridge, MA: The Belknap Press of Harvard University Press; 1982	The t a s t a s c *p* (o b t *c* a t t *d* d o n p s
Egas Moniz (1874–1955), Walter Freeman (1895–1972)	United States, Europe	Valenstein ES. *Great and Desperate Cures: The Rise and Decline of Psychosurgery and Other Radical Treatments for Mental Illness.* New York: Basic Books; 1986; Pressman JD. *Last Resort: Psychosurgery and the Limits of Medicine.* Cambridge: Cambridge University Press; 1998	Dur e i F w C l t f n s p i e p w p o d
Advent of the double-blind test	United States	Gold H, Kwit NT, Otto H. The xanthines (theobromine and aminophylline) in the treatment of cardiac pain. *JAMA.* 1937;108:2173; Hampson JL, Rosenthal D, Frank J. A comparative study of the effect of mephenesin and placebo on the symptomatology of a mixed group of psychiatric outpatients. *Bull Johns Hopkins Hosp.* 1954;95:170	In t o p s p i w a

			t n e x c p e n li d il n d t p p
Carl Gustav Jung (1875–1961)	Switzerland	*Psychological Types* (1921), *Two Essays on Analytical Psychology* (1912–1928), *The Structure and Dynamics of the Psyche* (1916–1952), *The Archetypes and the Collective Unconscious* (1934–1955); Kirsch T. *The Jungians: A Comparative and Historical Perspective.* London: Routledge; 2000	Afte J a d p a t t a s e
Antidepressant drugs	Switzerland, United States	Kuhn R. The treatment of depressive states with an iminodibenzyl derivative (G 22355). *Swiss Med J.* 1957;87:1135; Loomer HP, Saunders JC, Kline NS. A clinical and pharmacodynamic evaluation of iproniazid as a psychic energizer. *Psychiatr Res Rep.* 1957;8:129; Healy D. *The Antidepressant Era.* Cambridge, MA: Harvard University Press; 1997; Kuhn R. *Hist Psychiatry.* 2007;17:253	The e d a (o
B. F. Skinner (1904–1990), Joseph Wolpe (1915–1997)	United States, South Africa	Skinner BF. *Science and Human Behavior* New York: Free Press; 1953; Wolpe J. *Psychotherapy by Reciprocal Inhibition.* Stanford, CA: Stanford University Press; 1958	Skir a i t o
Development of child psychiatry	United States	Lewis M, ed. *Child and Adolescent Psychiatry: A Comprehensive Textbook.* 2nd ed. Baltimore, MD: Williams & Wilkins; 1996: Chapters 120 (Rafferty F. Effects of health delivery systems on child and adolescent mental health care); 123 (Schowalter J. Recruitment, training, and certification in child and adolescent psychiatry in the United States); 127 (Bernstein D. The discovery of the child: A historical perspective on child and adolescent psychiatry)	Afte d s f t A P A P
Karl Jaspers (1883–1969)	Germany	*General Psychopathology* (1st ed., 1913; 7th ed., 1959) (English trans., 1963)	Deli s n b (

Kurt Schneider (1887–1967)	Germany	Schneider K. *Psychopathic Personalities.* London: Cassell; 1958; Schneider K *Clinical Psychopathology.* New York: Grune Stratton; 1959	In c S w d p a a i p d e h c o s s *P* n p e d p
Michel Foucault (1926–1984)	France	*Histoire de la folie* (1961); English trans., *Madness and Civilization: A History of Insanity in the Age of Reason.* New York: Pantheon; 1965	Gav h i 1 s l h w o t a e n t c v a w d k h h

			p ir s fa c id d is “ ir p th c
Benzodiazepine drugs	United States	Cohen IM. The benzodiazepines. In: Ayd FJ Jr, Blackwell B, eds. *Discoveries in Biological Psychiatry.* Baltimore, MD: Ayd Medical Communications; 1984:130	The d p H c w a 1 w a n
Haloperidol in the treatment of Gilles de la Tourette syndrome	France	Seignot JN. Un cas de maladie des tics de Gilles de la Tourette guéri par le R-1625. *Ann Med Psychol.* 1961; cxix: 578; The case of Gilles de la Tourette syndrome: A condition of nervous tics cured by R.1625. *Hist Psychiatry.* 1997;8:434	Rep o h th e 1 tr u w a th b c fo
Joint Commission on Mental Illness and Mental Health (1955–1961)	United States	*Action for Mental Health* (1961); Community Mental Health Centers Act of 1963	*Acti* r d c il fr in h d b p M o
Rene Spitz (1887–1974)	Europe, then United States	*The First Year of Life: A Psychoanalytic Study of Normal and Deviant Development of Object Relations.* New York: International Universities Press; 1965	Sho fr th

			in d b o c d t a b 2 s o n
Classification of mental disorders	United States	DSM-I prepared by the Committee on Nomenclature and Statistics of the APA, G.N. Raines, chair (Mental Hospital Service, Washington, DC, 1952); DSM-II prepared by the Committee on Nomenclature and Statistics of the APA, E. Gruenberg, chair (APA, Washington, DC, 1968); Grob G. Origins of DSM-I: A study in appearance and reality. *Am J Psychiatry.* 1991;148:421	DSM o n t g p w a S c d e o a n c p
Epidemiological evidence for the genetic transmission of schizophrenia	United States, Denmark	Rosenthal D, Kety SS, eds. *The Transmission of Schizophrenia.* London: Pergamon Press; 1968:235 (Wender PH, Rosenthal D, Kety SS. A psychiatric assessment of the adoptive parents of schizophrenics), 345 (Kety SS, Rosenthal D, Wender PH, Schulsinger F. The types and prevalence of mental illness in the biological and adoptive families of adopted schizophrenics), and 377 (Rosenthal D, Wender PH, Kety SS, Schulsinger F, Welner J. Schizophrenics' offspring reared in adoptive homes); Kety SS, Wender PH, Jacobsen B, Ingraham LJ, Jansson L. Mental illness in the biological and adoptive relatives of schizophrenic adoptees: Replication of the Copenhagen study in the rest of Denmark. *Arch Gen Psychiatry.* 1994;51:442	A st a h p w s e a h d f b s 8 s w b n t i f t s s

			r t n u p (p b h w f n s g d
Lithium therapy for psychotic excitement	Australia, Denmark, United States	Cade J. Lithium salts in the treatment of psychotic excitement. *Med J Aust.* 1949;36:349; Schou M, Juel-Nielsen N, Stromgren E, Voldby H. The treatment of manic psychoses by the administration of lithium salts. *J Neurol Neurosurg Psychiatry.* 1954;17:250; Schou M. Forty years of lithium treatment. *Arch Gen Psychiatry.* 1997;54:9; J Schioldann: Mogens Abelin Schou (1918–2005)—Half a century with Lithium. *Hist Psychiatry.* 2007;17:247	Afte p t p p c d t t d o b q e a U A
William H. Masters (1915–2001), Virginia E. Johnson (1925–)	United States	*Human Sexual Response* (1966), *Human Sexual Dysfunction* (1970)	Bot k a s c e t c n t s c *D* n a o s v p a d
Margaret Mahler	Central Europe	Mahler M, Pine F, Bergman A. *The Psychological Birth of*	Deli

(1897–1985)	until 1938, then United States	*the Human Infant: Symbiosis and Individuation.* New York: Basic Books; 1975; *The Memoirs of Margaret S. Mahler.* Stepansky PE, ed. New York: Free Press; 1988	n i f u d w i
John Bowlby (1907–1990)	England	*Attachment* (1969), *Separation: Anxiety and Anger* (1973), *Loss: Sadness and Depression* (1980), *A Secure Base: Clinical Applications of Attachment Theory* (1988); Holmes J. *John Bowlby and Attachment Theory.* London: Routledge; 1993	Bow l t e d n o a a d p a o a t a a a b a t t d t e (s r
Jacques Lacan (1901–1981), “the French Freud”	France	*The Language of the Self: The Function of Language in Psychoanalysis* (1968); Turkle S. *Psychoanalytic Politics: Freud’s French Revolution.* New York: Basic Books; 1978; Roudinesco E. *Jacques Lacan & Co.: A History of Psychoanalysis in France, 1925–1985.* Chicago: University of Chicago Press; 1990	Emp n t t r n t o s d F d r a i p i l e

			n ir le
Heinz Kohut (1913–1981)	Austria until 1940, then United States	*The Analysis of the Self* (1971); *The Restoration of the Self* (1977); The two analyses of Mr. Z. *Int J Psychoanal.* 1979;60:3; Baker H, Baker M. Heinz Kohut's self psychology: An overview. *Am J Psychiatry.* 1987;144:1; Strozier CS. *Heinz Kohut: The Making of a Psychoanalyst.* New York: Farrar, Straus and Giroux; 2001	Orig s w g n o id
Jean Piaget (1896–1980)	Switzerland	Greenspan, Curry J. Piaget's approach to intellectual functioning. In: Kaplan HI, Sadock BJ, eds. *Comprehensive Textbook of Psychiatry.* 6th ed. Baltimore, MD: Williams & Wilkins; 1995	Influ o f th fr a n th *s* c p *p* v s o *o* o a *f* th a th n b r s w C u e p c in
NAMI	United States	*NAMI: We Are Family.* Arlington, VA: NAMI; 1979; a 10-page brochure describing the work of NAMI; since then NAMI has published regular annual reports on its work	NAM o w fa d to c th th p p

			t g a t o il
Sociobiological psychiatry	United States, Holland	Van Praag H. Sociobiological psychiatry. *Compr Psychiatry.* 1981;22:441	An a o d p P p o p c t p v c
Roger Sperry (1913–1994)	United States	Some effects of disconnecting the cerebral hemispheres. *Science.* 1982;217:1223; Sperry RW, Levi-Montalcini R. In: Finger S, ed. *Minds Behind the Brain: A History of the Pioneers and Their Discoveries.* New York: Oxford University Press; 2000	In 1 N M d b t c h s t i c i g s a s b c h
Brief psychotherapy	United States, England	Malan D. *A Study of Brief Psychotherapy.* New York: Plenum Press; 1976; Mann J, Goldman R. *A Casebook in Time-Limited Psychotherapy.* New York: McGraw-Hill; 1982; Davanloo H. *Short-Term Dynamic Psychotherapy.* New York: Jason Aronson; 1980; Sifneos P. *Short-Term Dynamic Psychotherapy Evaluation and Technique.* New York: Plenum Press; 1979; Crits-Cristoph P. The efficacy of brief dynamic psychotherapy: A meta-analysis. *Am J Psychiatry.* 1992;149:151	Aim p p li i d f p t a c p a r r

			c t
Psychology of women	United States	Miller JB. *Towards a New Psychology of Women.* Boston: Beacon Press; 1976; Gilligan C. *In a Different Voice: Psychological Theory and Women's Development.* Cambridge, MA: Harvard University Press; 1982; Levinson D, in collaboration with Levinson J. *The Seasons of a Woman's Life.* New York: Knopf; 1996	Jear G L v p li
NARSAD	United States	NARSAD activities are reported in *Research* brochures and *Newsletters*	NAF 1 p $ ir r h t 1 u r r a N h D g Ir b n fr d d
Changes in the diagnosis and classification of homosexuality	United States	Bayer R. *Homosexuality and American Psychiatry: The Politics of Diagnosis.* New York: Basic Books; 1981; DSM-III (1980); DSM-III-R (1987)	In 1 c u fr F d fr c c it t s p w r h d *h* th h e fa p

			h k c ir p o
Classification of mental disorders	United States	DSM-III (1980); DSM-III-R (1987); Wilson M. DSM-III and the transformation of American psychiatry: A history. *Am J Psychiatry.* 1993;150:399	The e d d d a p s r U tl b tl b s w d b r w o c
James Masterson (1926–), Otto Kernberg (1928–)	United States	Masterson J. *Psychotherapy of the Borderline Adult: A Developmental Approach.* New York: Brunner/Mazel; 1976; *Narcissistic and Borderline Disorders: An Integrated Developmental.* New York: Brunner/Mazel; 1981; Masterson J, Klein R, eds. *Psychotherapy of the Disorders of the Self: The Masterson Approach.* New York: Brunner/Mazel; 1988; Kernberg O. *Borderline Conditions and Pathological Narcissism.* New York: Jason Aronson; 1975; Kernberg O. *Severe Personality Disorders: Psychotherapeutic Strategies.* New Haven, CT: Yale University Press; 1984	Mas d p (l r u a b d
Psychology and neurophysiology of dreams	United States	Hobson JA, McCarley RW. The brain as a dream state generator: An activation-synthesis hypothesis of one dream process. *Am J Psychiatry.* 1977;134: 1335; McCarley RW. Dreams: Disguise of forbidden wishes or transparent reflections of a distinct brain state? In: Bilder RM, Lefever FF, eds. *Neuroscience of the Mind: On the Centennial of Freud's Project for a Scientific Psychology.* New York: New York Academy of Sciences; 1998:116; Reiser MF. The dream in contemporary psychiatry. *Am J Psychiatry.* 2001;158:351	The a h f n r d r g n h v b ir ir

			r p d c t s o t t p p t w (f p a R c c t t s r p c
Theories on the psychology of love	United States	Bergmann MS. *The Anatomy of Loving.* New York: Columbia University Press; 1987; Sternberg RJ. *The Triangle of Love: Intimacy, Passion, Commitment.* New York: Basic Books; 1987; Sternberg RJ, Barnes ML, eds. *The Psychology of Love.* New Haven, CT: Yale University Press; 1988; Gaylin W, Person E, eds. *Passionate Attachments: Thinking about Love.* New York: Free Press; 1988; Person E. *Dreams of Love and Fateful Encounters: The Power of Romantic Passion.* New York: WW Norton; 1988	The w d t d y i t
Evaluations of psychotherapy	United States	Bruce R, Staples FR, Cristol AH, Yorkston NJ, Whipple KS. *Psychotherapy versus Behavior Therapy.* Cambridge, MA: Harvard University Press; 1975; Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM. National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Arch Gen Psychiatry.* 1989;48:971	Two s c n p s s w
Eric Kandel (1929–)	Austria, United States after 1939	Kandel E. Psychotherapy and the single synapse: The impact of psychiatric neurobiological research. *N Engl J Med.* 1979;301:1028; From metapsychology to molecular biology: Explorations into the nature of anxiety. *Am J Psychiatry.* 1983;140:1277; Genes, nerve cells, and the remembrance of things past. *J Neuropsychiatry.* 1989;1:103; Biology and the future of psychoanalysis: A new intellectual framework for psychiatry revisited. *Am J Psychiatry.* 1999;156:505; *In Search of Memory: The Emergence of a New Science of*	Sho p a t a s a b f n

		Mind. New York: WW Norton; 2006	t c a r a e n n p s t i l s b a h s a c p f u
			In 2 P a K (r V C n a G k o h G a N h a t l i c A
Aaron Beck (1921–)	United States	Beck A. *Cognitive Therapy and the Emotional Disorders.* New York: International Universities Press; 1976; Beck A, Rush AJ, Shaw BF, Emery G. *Cognitive Therapy of Depression.* New York: Guilford Press; 1979; Beck A, Emery G, Greenberg R. *Anxiety Disorders and Phobias: A Cognitive Perspective.* New York: Basic Books; 1985;	2 As f c s b p o

		Beck A, Freeman A. *Cognitive Therapy of Personality Disorders.* New York: Guilford Press; 1990; Weishaar M. *Aaron T. Beck. Key Figures in Counselling and Psychotherapy.* London: Sage; 1993; Barsky AJ, Ahern DK. Cognitive behavioral therapy for hypochondriasis. A randomized controlled trial. *JAMA.* 2004;291:1464	a e r n p f o s b n h l s h N w c A c A n n a N 1
Karl Menninger (1893–1990), William Menninger (1900–1966)	United States	Faulkner H, Pruitt V, eds. *The Selected Correspondence of Karl A. Menninger.* New Haven, CT: Yale University Press; 1988; Friedman LJ. *Menninger: The Family and the Clinic.* New York: Knopf; 1990	The i b i n e o N e A c K p A w i p *H A L* J f T p o N
Frederick K. Goodwin (1936–), Kay	United States	Goodwin FK, Jamison KR. *Manic-Depressive Illness.* New York: Oxford University Press; 1990; 2nd ed., 2007	Offe c d

Redfield Jamison (1946–)			c i c d o c i a t n p p
A memoir of depression	United States	Styron W. *Darkness Visible: A Memoir of Madness.* New York: Random House; 1990	A n d *V* s n o u i h t d *T* S
SSRIs	United States	"The Story of Prozac," Washington, DC: Pharmaceutical Manufacturers Association; 1993; Grebb J. Serotonin-specific inhibitors. In: Kaplan HI, Sadock BJ, eds. *Comprehensive Textbook of Psychiatry.* 6th ed. Baltimore, MD: Williams & Wilkins; 1995	SSR g p b e fl d 1 n a U f (R p d B
Jerome Frank (1910–2005), E. Fuller Torrey (1937–)	United States	Frank J. *Persuasion and Healing: A Comparative Study of Psychotherapy.* Baltimore, MD: Johns Hopkins University Press; 1961; 2nd rev. ed., 1973; Frank J, Frank JB. *Persuasion and Healing: A Comparative Study of Psychotherapy.* 3rd rev. ed. Baltimore, MD: Johns Hopkins University Press; 1991; Torrey EF. *The Mind Game: Witchdoctors and Psychiatrists.* New York: Emerson Hall; 1972; Torrey EF: *Witchdoctors and Psychiatrists: The Common Roots of Psychotherapy and Its Future.* New York: Harper & Row; 1986 (a revised edition of *The Mind Game*)	The p p A a d c w p i f t

			c n ir p w ir a R N
Psychiatric epidemiology	United States	Robins LN, Regier DA, eds. *Psychiatric Disorders in America: The Epidemiologic Catchment Area Study.* New York: Free Press; 1991	This C la s p e U n ir p n th d c b u a v ir e p th v th u c s p o c f n fr p
Forensic psychiatry	United States	Bursztajn H, Scherr A, Brodsky A. The rebirth of forensic psychiatry in light of recent historical trends in criminal responsibility. *Psychiatr Clin North Am.* 1994;1(3):611; Prentice S. A history of subspecialization in forensic psychiatry. *Bull Am Acad Psychiatry Law.* 1995;23(2):195	Cha r c ir A F A P r p o

			s A S 1 d C P
Epidemic of AIDS and of persons infected with HIV (1981–)	Worldwide	Grmek M. *History of AIDS: Emergence and Origin of a Modern Pandemic.* Princeton, NJ: Princeton University Press; 1990; Berridge V, Strong P, eds. *AIDS and Contemporary History.* Cambridge: Cambridge University Press; 1993; Feldman D, Miller J, eds. *The AIDS Crisis: A Documentary History.* Westport, CT: Greenwood Press; 1998; Bayer R, Oppenheimer G, eds. *AIDS Doctors: Voices from the Epidemic.* Oxford: Oxford University Press; 2000; Treisman G, Angelino F. *The Psychiatry of AIDS: A Guide to Diagnosis and Treatment.* Baltimore, MD: Johns Hopkins University Press; 2004; Mayer K, Pizer H, eds. *The AIDS Pandemic: Impact on Science and Society.* San Diego, CA: Elsevier Academic Press; 2005	Psy n a o t d a i i a c n A s u n n t s a d A
Classification of mental disorders	United States	American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders.* 4th ed. Text rev. Washington, DC: American Psychiatric Association; 2000 (DSM-IV-TR)	The n i p b a i f c c a f t s e a r t w D u t p f t

			h D c c c n p d ir
Clozapine, risperidone, and olanzapine in the treatment of schizophrenia	United States, Canada, Europe, England, South America, South Africa	Kane J, Honigfield G, Singer J, Meltzer H, and the Clozaril Collaborative Study Group: Clozapine for the treatment-resistant schizophrenic. A double-blind comparison with chlorpromazine. *Arch Gen Psychiatry.* 1988;45:789; Kane J, Marder SR. Clozapine benefits and risks. *Schizophren Bull.* 1994;20:23; Ayd FJ Jr. Risperidone (Risperdal): A unique antipsychotic. *Int Drug Ther Newslett.* 1994;29:5; Marder SR, Meibach RC. Risperidone in the treatment of schizophrenia. *Am J Psychiatry.* 1994;151:825; Kane J. Risperidone. *Am J Psychiatry.* 1994;151:802; Beasley C, Tollefson G, Tran P, Satterlee W, Sanger T. Olanzapine versus placebo and haloperidol acute phase results of the North American double-blind olanzapine trial. *Neuropsychopharmacology.* 1996;14:111; Tollefson G, Kuntz A. Review of recent clinical studies with olanzapine. *Br J Psychiatry.* 1999;37:30; Volavka J, Czobor P, Sheitman B, Lindenmayer JP, Citrome L. Clozapine, olanzapine, risperidone, and haloperidol in the treatment of patients with chronic schizophrenia and schizoaffective disorder. *Am J Psychiatry.* 2002;159:255; Grilly J. The history of clozapine and its emergence in the U.S. market: A review and analysis. *Hist Psychiatry.* 2007;18:39	Cloz h s t n s r t e r ir w p b S e a s b h h s s s ri (l r s t s p p t o g e ri s C e n H w
Publications in psychohistory	United States, England	Lawton H. *The Psychohistorian's Handbook.* New York: Psychohistory Press; 1988; Elovitz P, ed. *Clio's Psyche.*	Law “

and psychobiography	*Understanding the "Why" of Culture, Current Events, History, and Society* (1994) (a quarterly periodical published by the Psychohistory Forum, 627 Dakota Trail, Franklin Lakes, NJ 07417); Sabbadini A, ed. *Psychoanalysis and History.* Vol. 1, No. 1. London: Artesian Books; 1998–1999 (since its first issue, this *History* has appeared at regular intervals of twice a year; since 2005 edited by John Forrester, and since 2007 published by Edinburgh University Press); Szaluta J. *Psychohistory: Theory and Practice.* New York: Peter Lang; 1999; Elovitz P. Psychoanalytic scholarship on American Presidents. *Annu Psychoanal.* 2003;31:135; Schultz WT, ed. *Handbook of Psychohistory.* New York: Oxford University Press; 2005

e
n
Psyc
t
“
t
n
b
a
w
p
c
h
a
u
e
v
d
n
d
r
p
d
t
n
a
a
a
p
i
u
d
r
n
r
p
‘
(
i
e
Szal
t
t
p
c
p
d
c
e
o
S
p
t
it

			s f Elov a r p ir p n s a n tl h p b d p fa e n c fo fo c tl tl p ir tl tl tl A *H* o p p a p
Paul Wender (1934–)	United States	*Attention-Deficit Hyperactivity Disorder in Adults.* New York: Oxford University Press; 1998	As a s V s g tl tl r p d w A c tr d
Computers and the	United States	The APA has published two books on the use of computers	Con

Internet		in psychiatry: Lieff J. *Computer Applications in Psychiatry.* Washington, DC: American Psychiatric Publishing; 1987; and Chan CH, Luo JS, Kennedy RS. *Concise Guide to Computers in Clinical Psychiatry.* Washington, DC: American Psychiatric Publishing; 2002; Miller MJ. Computer-based testing of the psychiatric patient. In: Sadock BJ, Sadock VA. *Comprehensive Textbook of Psychiatry.* 7th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2000	t p r s t p a t i n f n a h i g p
		Two articles about the Internet are Alessi NE. The Internet and the future of psychiatry. *Am J Psychiatry.* 1996;153:861; and Rajendran PR. The Internet: Ushering in a new era of medicine. *JAMA.* 2001;285:804	The c t q g c p p a o t q p t n h Alth g p n i i n c t a c b
Discoveries of genetic variations as causes of mental illnesses	United States, Canada, England, Europe, Australia	Oberlé I, Rousseau F, Heitz D, Kretz G, Davys D. Instability of a 550–base pair DNA segment in fragile X syndrome. *Science.* 1991;252:1097; Yu S, Pritchard M, Kremer E, Lynch M, Nancarrow J. Fragile X genotype characterized by an unstable region of DNA. *Science.* 1991;252:1179; Ying-Hui Fu, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS. Variation of the CGG repeat at the fragile X site results in genetic instability: Resolution of the Sherman paradox. *Cell.* 1991;67:1047; Schellenberg GD, Bird TD,	Dur v c s o r o a 1

		Wijsman EM, Orr T, Anderson L. Genetic linkage of evidence for a familial Alzheimer's disease locus on chromosome 14. *Science.* 1992;258:668; Mullen M, Houlden H, Windelspecht M, Fidani L, Lombardi C. A locus for familial early-onset of Alzheimer's disease on the long arm of chromosome 14, proximal to the α-1-antichymotrypsin gene. *Nat Genet.* 1992;2:340; The Huntington's Disease Collaborative Research Group: A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosome. *Cell.* 1993;72:971; Amir RE, Van den Veyver IB, Wan M, Tran CQ, Franke U. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet.* 1999;23:185	t s d i c
		Hyman S, Moldin S. Genetic science and depression: Implications for research and treatment. In: Weissman MM, ed. *Treatment of Depression: Bridging the 21st Century.* Washington, DC: American Psychiatric Association; 2001:98; Andreasen NC. Schizophrenia. Chapter 8. In: *Brave New Brain: Conquering Mental Illness in the Era of the Genome.* New York, Oxford: Oxford University Press; 2001:199	The M a h p t " u e g c f v i A e s n n e d a e T a n v s b 5 h r
Three books on psychotherapy	United States	Jackson SW. *Care of the Psyche: A History of Psychological Healing.* New Haven, CT: Yale University Press; 1999; Sabo AN, Havens L, eds. *The Real World Guide to Psychotherapy Practice.* Cambridge, MA: Harvard University Press; 2000; Hersen M, Sledge WH, eds. *Encyclopedia of Psychotherapy.* 2 Vols. San Diego, CA: Academic Press; 2002	Jack it o t e o t r

			it e t o *E* a o p n s e
Education of psychiatrists	United States	Luhrmann TM. *Of Two Minds: The Growing Disorder in American Psychiatry.* New York: Knopf; 2000	Dep A d o b c t p p c c p L u is p " e c p p o b p w n h
Theory and treatment of depression	United States, England, Switzerland	In a March 1999 New York City meeting of the American Psychopathological Association, leading clinical scientists gave papers on the topic Treatment of Depression, all of which were published as in Weissman MM, ed. *Treatment of Depression: Bridging the 21st Century.* Washington, DC: American Psychiatric Association; 2001	The d o N C n d f r C d " e w e n

n
v
e
n
d
a
C
t
d
s
a
“
d
r
d
k
C
i
g
e
p
t
g
g
c
p
a
g
d
Oth
o
t
2
a
g
e
t
“
p
r
p
n
t
s
s
b
e
b
e
c
e
o
d
t

			f a p d
Neuroimaging	United States	Andreasen NC. Mapping the mind. In: *Brave New Brain: Conquering Mental Illness in the Era of the Genome.* Oxford: Oxford University Press; 2001:130; Morihisa JM, ed. *Advances in Brain Imaging.* Washington, DC: American Psychiatric Association, 2001	Beg 1 M w t o i s v a w t n f p t w c (il s t t t p r d b t b
The effects of traumatic events on individuals	United States, England, France, Germany, and Italy	Micale MS, Lerner P, eds. *Traumatic Pasts, Psychiatry, and Trauma in the Modern Age, 1870–1930.* Cambridge: Cambridge University Press; 2001	A co s d i c a d e a r e a V “ h c “ t c t

			it s t w i t e a t a w c n
Diagnosis and treatment of borderline personality disorder	United States	Gunderson JG. *Borderline Personality Disorder: A Clinical Guide.* Washington, DC: American Psychiatric Association; 2001; American Psychiatric Association. *Practice Guideline for the Treatment of Patients with Borderline Personality Disorder.* Washington, DC: American Psychiatric Association; 2001; Oldham JM. A 44-year-old woman with borderline personality disorder. *JAMA.* 2002;287:1031	Gur A r o (c p p n t s d S t “ w d i
September 11, 2001, terrorist attacks on the World Trade Center and the Pentagon	United States	Kleinfield NR, Connelly M: 9/11 still strains New York psyche. Poll finds widespread unease and lingering fear of terror. *The New York Times,* September 8, 2003, Section A, pp. A1–A2; Sheehy G: *Middletown America. One Town's Passage from Trauma to Hope.* New York: Random House; 2003; Cancelmo JA, Tylim I, Hoffenberg J, eds. *Terrorism and the Psychoanalytic Space: International Perspectives from Ground Zero.* New York: Pace University Press; 2003	The w a a v p f a i a S li f P p f t t N s a (S

Overviews of psychoanalysis	United States	Erwin E, ed. *The Freud Encyclopedia: Theory, Therapy, and Culture.* New York: Routledge; 2002; Zaretsky E. *Secrets of the Soul: A Social and Cultural History of Psychoanalysis.* New York: Knopf; 2004	The n F in s e in t p e h F a u n t p h a s b fi p b a s v t w A C P *Sec* k h in p c ic a *p* n t in a r t r t A t r 1 g

			o t c p f a h l p n c b a h p
Psychosomatic medicine	United States	Blumenfield M, Strain J, eds. *Psychosomatic Medicine.* Philadelphia, PA: Lippincott Williams & Wilkins; 2006	In 2 o g s p a e p w b c a c p n n h b
Confinement of the mentally ill in asylums	South Africa, Switzerland, France, Canada, Australia, Germany, United States, Japan, Argentina, Mexico, India, Nigeria, Ireland, and England	Porter R, Wright D, eds. *The Confinement of the Insane. International Perspectives, 1800–1965.* Cambridge: Cambridge University Press; 2003	This o d w f c i c p a s t c o " it t n a c u

			(t f s o p w s n d f t
Treatment of schizophrenia	United States	American Psychiatric Association: *Practice Guideline for the Treatment of Patients with Schizophrenia.* 2nd ed. Washington, DC: American Psychiatric Association; 2004	Rec i p a w a s s r o e R s r n n t c *P* o s r is a n a a ju p o n li p a t a
College Mental Health	United States	Kadison R, Geronimo I. *College of the Overwhelmed: The Campus Mental Health Crisis and What to Do about It.* San Francisco: Jossey-Bass; 2004	Deli p s a e n

A suppressed psychiatric scandal	United States	Scull A. *Madhouse: A Tragic Tale of Megalomania, and Modern Medicine.* New Haven, CT: Yale University Press; 2005	The a S fr H S th b p tr b th n b s c o h D in tr r tr d q o h h d S o s a o u u r d

			r i h p
Use of amphetamines as antidepressant drugs	United States	Rasmussen N. Making the first antidepressant: Amphetamine in American medicine, 1929–1950. *J Hist Med Allied Sci.* 2006;61:288	Sho a w p d n b a 1 a v la in b tr a
Recollections of treating schizophrenia	United Sates	Seeman M. Forty-five years of schizophrenia: Personal reflections. *Hist Psychiatry.* 2006;17:363	See th b s th a u s b re T s b p it o to s lo
Mental health policy	United States	Frank R, Glide S. *Better But Not Well: Mental Health Policy in the United States Since 1950.* Baltimore, MD: Johns Hopkins University Press; 2006	The d b n s in tr a p M D e in n a

			h t i 5 h t i f
George Mora (1923–2006), and his work in the history of psychiatry	Italy and after 1953 the United States	Works on the history of Italian psychiatry, a wide range of medical and psychiatric topics, a translation of Johann Weyer's *De praestigiis daemonum* (1991), and with Jeanne Brand, *Psychiatry and Its History: Methodological Problems in Research* (1970)	For f a p a h f *C* *P*
Mood disorders	United States	In 2006–2007 two textbooks on mood disorders were published: Stein D, Kupfer D, Schatzberg A, eds. *The American Psychiatric Publishing Textbook of Mood Disorders.* Arlington, VA: American Psychiatric Publishing; 2006; Goodwin FK, Jamison KR. *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression.* 2nd ed. New York: Oxford University Press; 2007 (1st ed., 1990)	Eac d l c t d
Albert Ellis (1913–2007)	United States	Ellis published 75 books, including *Overcoming Procrastination, How to Live With a Neurotic, Sex Without Guilt,* and *A Guide to Rational Living*	By l d E (t i p t t t f I w o p p p p w C A A s p c R
Mental Health Policy	United States	Mental Health Parity and Addiction Equity Act of 2008. "Fact Sheet": The Mental Health Parity and Addiction Equity Act of 2008 (MHPAEA). Accessed November 19,	Ena A a

		2015. Available at http://www.dol.gov/ebsa/newsroom/fsmhpaea.html	b P (a b *H* f n s b a a A q ir n s s
Publication of DSM-5	United States	American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association. 2013	In 2 D s A d u *s* w f t n a s s e li lo v a s A d p d e c r d ir e d t s C d d

			d i
Mentally ill and gun violence	United States	Metzl, Johnathan M., and Kenneth T. Machleish. “Mental Illness, Mass Shootings, and the Politics of American Firearms.” *Am J Public Health American Journal of Public Health.* 2015:240–249	The n v s w h d g c p t s n e o h s l h “ l p o a b t i s g n
Changes in marijuana legislation	United States (23 states)	Wilkinson, Sauel T, Stephanie Yarnell, Rajiv Radhakrishnan, Samuel A. Ball, and Deepak Cyril D’souza. “Marijuana Legalization: Impact on Pyschicians and Public Health.” *Annual Review of Medicine Annu Rev Med.* 2014	Bet s n p s t n i a s a r g o o a u s s c d n

			r c a D r c a
United States Drug Enforcement Agency	United States	Wing N. (2015, March 18). DEA Approves Study of Psychedelic Drug MDMA In Treatment of Seriously Ill Patients. Retrieved November 17, 2015, from http://www.huffingtonpost.com/2015/03/18/deadmda-study_n_6888972.html	MD s 1 r t e d s t p il s c li t
Physician-assisted suicide	United States	Dyer, Owen, Caroline White, and Aser Garcia Rada. "Assisted Dying: Law and Practice around the World." Bmj (2015): Payne, Kate. "Consider Assisted Suicide." Tennessee Nurse 78.3 (2015): n. pag. Web 17 November 2015 Kopelman, Loretta M. Physician- assisted Suicide: What Are the Issues? Dordrecht: Kluwer Academic Publishers, 2001	Me h p a H a f b y p a 1 w a w t f t li s le r n
The Human Connectome Project	United States	Hagmann P, Kurant M, Gigandet X, Thiran P, Wedeen VJ, Meuli R, Thiran, JP (2007) Mapping Human Whole-Brain Structural Networks with Diffusion MRI. *PLoS ONE.* 2, e597; Sporns O, Tononi G, Kotter R. (2005) The human connectome: A structural description of the human brain. *PLos Comput Biol.* 1, e42	The P t h u p t s la

b
n
h
c
f
d
n
c
c
c
g
c
c
c
p

ADHD, attention-deficit/hyperactivity disorder; AIDS, acquired immune deficiency syndrome; APA, American Psychiatric Association; CNS, central nervous system; CRF, corticotropin-releasing factor; CT, computed tomography; DSM, *Diagnostic and Statistical Manual of Mental Disorders;* ECT, electroconvulsive therapy; FDA, U.S. Food and Drug Administration; HIV, human immunodeficiency virus; MAOI, monoamine oxidase inhibitor; MMPI, Minnesota Multiphasic Personality Inventory; MRI, magnetic resonance imaging; NAMI, National Alliance for the Mentally Ill; NARSAD, National Alliance for Research on Schizophrenia and Depression; NIMH, National Institute for Mental Health; NMHA, National Mental Health Association; OCD, obsessive-compulsive disorder; PET, positron emission tomography; PTSD, posttraumatic stress disorder; SSRI, selective serotonin reuptake inhibitor; TR, text revision.

This table was conceived and developed by Ralph Colp Jr, M.D., who contributed to many editions of this book and who died in 2008. It was updated by one of the editors (BJS) with the assistance of Kristel Carrington, M.D.

Appendix B

Glossary of Psychiatry and Psychology Terms

A

AA Alcoholics Anonymous.

AACAP American Academy of Child and Adolescent Psychiatry.

AACDP American Association of Chairmen of Departments of Psychiatry.

AADPRT American Association of Directors of Psychiatric Residency Training, Inc.

AAMC Association of American Medical Colleges.

ability Capacity or competence to perform an act without previous training. *Compare* aptitude.

ablation Excision of part of the body, as in lobotomy or topectomy.

ablutomania Excessive interest in bathing and cleaning oneself; commonly observed in obsessive-compulsive disorder.

abnormal Term subject to considerable variability in definition and usage, roughly the equivalent of "psychopathological," and defined according to several criteria: **1.** *Normative* or *statistical approach,* considers as abnormal any mental or behavioral activities that deviate from culturally determined averages or norms. **2.** *Subjective approach,* regards as abnormal or psychopathological those phenomena that cause pain or suffering or induce a person to seek help. **3.** *Theoretical approach,* considers as abnormal any deviations from a theoretical concept of normal functioning that are indicative of psychiatric illness.

abortion Spontaneous or induced expulsion of an embryo or fetus before it has reached the stage of viability, that is, before about 20 weeks gestation.

ABPN American Board of Psychiatry and Neurology.

abreaction A process by which repressed material, particularly a painful experience or a conflict, is brought back to consciousness; in this process, the person not only recalls but relives the repressed material, which is accompanied by the appropriate affective response. *See also* catharsis.

absence seizure PETIT MAL EPILEPSY.

absolute threshold Minimum amount of stimulus needed to produce a response.

abstinence The act of refraining voluntarily from some activity, or from the use of certain substances, such as food or drugs; in psychoanalysis, refers to refraining from sexual intercourse.

abstinence syndrome WITHDRAWAL SYNDROME; DISCONTINUATION SYNDROME.

abstract attitude ABSTRACT THINKING.

abstraction The process whereby thoughts or ideas are generalized and dissociated from particular concrete instances or material objects; concreteness in proverb interpretation suggests an impairment of abstraction, as in schizophrenia.

abstract thinking Thinking characterized by the ability to assume a mental set voluntarily, to shift voluntarily from one aspect of a situation to another, to keep in mind simultaneously various aspects of a situation, to grasp the essentials of a whole and to break a whole into its parts, to discern common properties, to plan, to assume make-believe attitudes, and to think or act symbolically. The capacity for abstract thinking is frequently impaired in patients with organic mental disorders and schizophrenia. Also called abstract attitude; categorical attitude. *Compare* concrete thinking.

abulia Inability to make decisions; lack of will.

abuse 1. To attack or injure physically or psychologically, most commonly children, the elderly, and marital partners. *See also* battered child syndrome. **2.** To misuse substances such as drugs, illegal or prescribed, or alcohol so as to impair functioning.

acalculia Loss of ability to do calculations; not caused by anxiety or impairment in concentration. Occurs with neurological deficit and learning disorder.

acarophobia Abnormal fear of minute animate (insects, worms) or inanimate (pins, needles) objects.

acatalepsia Severe mental deficiency characterized by the inability to reason or comprehend.

acataphasia Disordered speech in which statements are incorrectly formulated. Patients may express themselves with words that sound like the ones intended but that are not appropriate to the thoughts, or they may use totally inappropriate expressions.

acathexis Lack of feeling associated with an ordinarily emotion-charged subject; in psychoanalysis, denotes the patient's detaching or transferring of emotion from thoughts and ideas. Occurs in anxiety, dissociative, schizophrenic, and bipolar disorders. Also called decathexis. *Compare* cathexis.

accelerated interaction A term used in group therapy to refer to an extended group therapy session that may last for hours, sometimes a day. Alternate term for marathon group session.

accident proneness Tendency to have accidents as a result of psychological causes.

accreditation A voluntary process by which standards for hospitals and health facilities are approved by city or federal government agencies; may also refer to being credentialed by the American Board of Psychiatry and Neurology (ABPN). *See also* board certified psychiatrist.

acenesthesia Loss of sensation of physical existence.

acetylcholine A neurotransmitter that causes cardiac inhibition, vasodilation, gastrointestinal peristalsis, and other parasympathetic effects; plays a critical role in memory, sleep, and the regulation of mood.

achluophobia Abnormal fear of the dark.

acid Slang for lysergic acid diethylamide (LSD).

ACP American College of Physicians.

acquired immunodeficiency syndrome (AIDS) According to the U.S. Centers for Disease Control, "a disease, at least moderately predictive of a defect in cell-mediated immunity, occurring in a person with no known cause for diminished resistance to that disease. Such

diseases include Kaposi sarcoma (KS), *Pnenmocystis carinii* pneumonia (PCP), and other serious opportunistic infections." The disease is caused by the human immunodeficiency virus type 1 (HIV-1). *See also* AIDS dementia complex.

acromegaly Disorder marked by hypersecretion of growth hormone by the anterior pituitary resulting in characteristic enlargement of certain body structures, such as the head, hands, feet, lips, and jaw; often associated with somnolence, moodiness, and decreased libido.

acrophobia Abnormal fear of high places.

acting out Behavioral response to an unconscious drive or impulse that brings about temporary partial relief of inner tension; relief is attained by reacting to a present situation as if it were the situation that originally gave rise to the drive or impulse. Common in borderline states.

activity group psychotherapy Type of group therapy, introduced and developed by S. R. Slavson, designed for children and young adolescents, with emphasis on emotional and active interaction in a permissive, nonthreatening atmosphere; the therapist interprets behavior and stresses reality testing.

actualization Realization of one's full potential. *See also* individuation.

aculalia Nonsense speech associated with marked impairment of comprehension. Occurs in mania, schizophrenia, and neurological deficit.

acute stress disorder Severe emotional reaction resulting from extreme environmental stress such as the death of a loved one, a disaster, or a catastrophic life situation. It usually remits within hours or days; however, if it persists for over a month it may evolved into a condition called posttraumatic stress disorder (PTSD).

ADAMHA Alcohol, Drug Abuse, and Mental Health Administration.

adaptation Process of fitting or conforming to the environment by behavioral or psychic changes that promote an optimal level of functioning. *See also* adjustment.

adaptational approach In analytic therapy, consonant with Sandor Rado's formulations on adaptational psychodynamics, therapy focusing on the maladaptive patterns used by patients in the treatment sessions, on how those patterns developed, and on what the patients must do to overcome them and stabilize their functioning at self-reliant, adult levels. *See also* social adaptation.

adaptive behavior Any behavior that increases an organism's ability to adjust to a specific environment or situation.

ADD Attention-deficit disorder.

addiction Repeated and increased use of a drug or chemical substance, the deprivation of which gives rise to symptoms of distress and an irresistible urge to use the agent again; leads to deterioration of physical and mental health. *See also* drug dependence.

ADHD Attention-deficit hyperactivity disorder.

adiadochokinesis Inability to perform rapid alternating movements. Occurs with neurological deficit and cerebellar lesions.

adjustment Maladaptive reaction to identifiable circumstances or life events that is expected to remit when the stress ceases. In DSM-5, divided into the following subtypes based on symptoms: with anxiety, with depressed mood, with mixed anxiety and depressed mood, with disturbance of conduct, with mixed disturbance of emotions and conduct, and unspecified.

adjustment disorder Maladaptive reaction to identifiable circumstances or life events that is expected to remit when the stress ceases.

Adler, Alfred (1870–1937) Viennese psychiatrist and one of Freud's original followers who broke with him and introduced and developed the concepts of individual psychology, inferiority complex, and overcompensation. *See also* masculine protest.

adlerian Attributed to or described by Alfred Adler.

administrative psychiatry Psychiatry concerned with public affairs, policy-making, executive decisions, and the duties of persons in executive positions.

adolescence Period of growth from puberty to maturity. Its beginning is marked by the appearance of secondary sexual characteristics, usually at about age 12; termination is marked by the achievement of sexual maturity at about age 20. *See also* psychosexual development.

adolescent mania Signs of mania occurring in an adolescent as the result of substance abuse (including alcohol) or of antisocial behavior but not as the result of a true bipolar disorder.

adrenaline EPINEPHRINE.

adrenergic 1. Relating to nerve fibers, tracts, or pathways in which synaptic transmission is mediated by norepinephrine (e.g., the sympathetic division of the autonomic nervous system); more generally, to processes activated by epinephrine or norepinephrine. **2.** Relating to drugs (catecholamines) that mimic the actions of the sympathetic nervous system.

adrenergic system Nerve cells or fibers of the autonomic nervous system that employ norepinephrine as their neurotransmitter.

Advanced sleep phase circadian rhythm disorder A circadian rhythm sleep–wake disorder characterized by sleep–wake times that are earlier than conventional times; individuals with this disorder are referred to as "early birds" or "larks."

adventitious reinforcement Responses that are reinforced accidentally by the coincidental pairing of response and reinforcement; such events may have clinical implications in the development of phobias or other neurotic behavior.

adverse effects Secondary, unfavorable reactions or results; usually referring to side effects of drugs.

adynamia Weakness and fatigability, a characteristic of neurasthenia and depression.

adynamic ileus PARALYTIC ILEUS.

aerophagia Excessive swallowing of air. Seen in anxiety disorder.

affect The subjective and immediate experience of emotion attached to ideas or mental representations of objects. Affect has outward manifestations that may be classified as restricted, blunted, flattened, broad, labile, appropriate, or inappropriate. *See also* mood.

affect-fantasy Jung's term for an emotional-laden fantasy.

affectionate transference *See* transference.

affective disorder *See* MOOD DISORDER.

affective hallucination Hallucination in which the content is either grandiose or self-deprecatory or involves other features common to a depressive syndrome such as guilt, disease, or poverty.

affective interaction Interpersonal experiences and exchanges that are emotionally charged.

aftercare After hospitalization, a continuing program of rehabilitation designed to reinforce the effects of therapy and to help patients adjust to their environment.

ageism Prejudice or discrimination against a particular age-group, especially the elderly.

ageusia Lack or impairment of the sense of taste; may be seen in depression and neurological deficit.

aggression Forceful behavior, verbal or physical, usually directed against another person.

aggressive drive One of the two primal drives or instincts (the other is the sexual drive) in Freud's dual-instinct psychoanalytic theory. It is conceived of as an unconscious destructive drive or an impulse directed at oneself or another that aims toward dissolution and death and operates on the repetition-compulsion principle, in contrast to the sexual drive, which follows the pleasure-pain principle. Also called death instinct; Thanatos.

aging The process of growing old; in human physiology, aging is said to occur when there is a failure to replace cells in sufficient numbers to maintain full functional capacity.

agitated depression Major depressive disorder with psychomotor agitation as a primary sign and symptom, characterized by restlessness, increased motor activity, and racing thoughts.

agitation State of anxiety associated with severe motor restlessness.

agnosia Inability to understand the import or significance of sensory stimuli; cannot be explained by a defect in sensory pathways or sensorium. In strict usage, diagnosis of agnosia implies an organic cerebral lesion; however, the term has also been used to refer to the selective loss or disuse of knowledge of specific objects due to emotional circumstances, as seen in certain schizophrenic, hysteric, and depressed patients. For types of agnosia, see the specific term.

agonist In pharmacology, a substance that excites the receptor sites of the cell and promotes or enhances a biological response. *Compare* antagonist.

agoraphobia Morbid fear of open places or of leaving the familiar setting of the home; may be present with or without panic attacks.

agraphia Loss or impairment of a previously possessed ability to write.

aichmophobia Abnormal fear of pointed objects.

AIDS Acquired immunodeficiency syndrome.

AIDS dementia complex Neurological complication of AIDS characterized by confusion, impaired judgment, impulsive behavior, and psychosis. Also called AIDS encephalopathy.

ailurophobia Abnormal fear of cats.

akathisia State of motor restlessness manifested by a compelling need to be in constant movement; may be seen as an extrapyramidal side effect of phenothiazine or other antipsychotic medication.

akinesia Lack of physical movement, as in the extreme immobility of catatonic schizophrenia; may also occur as an extrapyramidal effect of antipsychotic medication.

akinetic mutism Absence of voluntary motor movement or speech in a patient who is apparently alert (as evidenced by following eye movements). Seen in psychotic depression and catatonic states.

Al-Anon Organization of relatives of alcoholics, patterned after Alcoholics Anonymous, which promotes the discussion and resolution of common problems.

Alateen Organization directed to children of alcoholics, through group support, so that they may better understand their parents' alcoholism. *See also* Alcoholics Anonymous.

alcohol abuse and dependence Called alcohol use disorder in DSM-5. *See also* alcoholism.

alcohol amnestic disorder DSM term for Korsakoff psychosis.

alcohol dementia WERNICKE SYNDROME.

Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA) An agency of the U.S. government that is part of the National Institutes of Health (NIH) within the Department of Health and Human Services. It administers grant programs supporting research, training, and service programs in alcoholism, drug abuse, and mental health. Individual institutes within the agency are the National Institute on Alcohol Abuse and Alcoholism (NIAAA), the National Institute of Drug Abuse (NIDA), and the National Institute of Mental Health (NIMH).

alcoholic blackout Amnesia experienced by alcoholics concerning behavior during drinking bouts; usually indicates that reversible brain damage has occurred.

alcoholic deterioration Dementia and mental deterioration associated with chronic excessive alcohol use. Called dementia associated with alcoholism in DSM.

alcoholic hallucinosis Occurrence of hallucinations with a clear sensorium in a person with a history of heavy drinking and alcohol dependence; usually follows a prolonged drinking bout. *See also* Delirium tremens.

alcoholic psychosis Mental disorder that results from alcoholism and that involves organic brain damage, as in delirium tremens and Korsakoff psychosis.

Alcoholics Anonymous (AA) An organization of alcoholics formed in 1935 that uses certain group methods, such as inspirational-supportive techniques, to help rehabilitate chronic alcoholics.

alcohol idiosyncratic intoxication DSM term for a syndrome of marked alcohol intoxication with subsequent amnesia for the period of intoxication produced by the ingestion of quantities of alcohol that would be insufficient to induce intoxication in most people. Also called pathological intoxication.

alcohol intoxication Constellation of specific neurological, psychological, and behavioral effects produced by the recent ingestion of alcohol; characteristically, the effects include slurred speech, motor ataxia, disinhibition of sexual or aggressive impulses, lability of mood, impairment of attention or memory, and impairment of judgment. *See also* intoxication.

alcoholism Excessive dependence or addiction to alcohol, usually to the point that the person's physical and mental health are threatened or harmed; may be episodic or continuous. Called alcohol use disorder in DSM-5.

alcohol withdrawal syndrome Symptom complex resulting from the cessation or reduction in alcohol intake in a person who has been drinking heavily; includes tremulousness, seizures (rum fits), hallucinations, and autonomic hyperactivity (delirium tremens). *See also* withdrawal syndrome.

Alexander, Franz (1891–1964) Hungarian psychoanalyst who founded the Chicago Institute for Psychoanalysis, developed the concept of corrective emotional experience and specificity hypothesis in psychosomatic disorders, and, in collaboration with Thomas French in 1946, identified most of the basic characteristics of brief dynamic psychotherapy.

alexia Loss of the power to understand the meaning of written or printed words and sentences. *Compare* dyslexia.

alexithymia Inability or difficulty in describing or being aware of one's emotions or moods;

associated with depression, substance abuse, and posttraumatic stress disorder (PTSD).

algophobia Abnormal fear of pain.

algorithm Graphic representation used as a set of instructions for solving diagnostic or clinical problems. Also called decision tree.

alienation In psychiatry, used to describe a person's feelings of detachment from self or society, to denote one's avoidance of emotional experiences, or to describe efforts to estrange oneself from one's own feelings. *Compare* relatedness. *See also* isolation.

alienist Obsolete term for a psychiatrist who offers expert opinion about a person's mental health or sanity.

allergic jaundice Relatively rare, unpredictable, non–dose-dependent hypersensitivity reaction after treatment with some phenothiazines; generally appears during the second to fourth week of therapy, resulting in hyperbilirubinemia and deposition of bile pigment in the skin and mucous membranes which causes a yellow appearance of the patient; an obstructive (cholestatic) type of jaundice manifested by predominantly conjugated (direct-reacting) hyperbilirubinemia, the presence of bilirubin in the urine, and an abnormally high plasma alkaline phosphatase.

alliance *See* therapeutic alliance.

allied health professional *See* paraprofessional.

alloplasty Adaptation to stress by attempting to change the environment. *Compare* autoplasty.

Allport's group relations theory Theory developed by Gordon W. Allport that a person's behavior is influenced by his or her personality and the need to conform to social forces; illustrates the interrelationship between group therapy and social psychology.

alogia Inability to speak because of mental deficiency or an episode of dementia.

alopecia Hair loss. *See* trichotillomania.

alpha (α) error TYPE 1 ERROR.

alpha waves As measured by an electroencephalogram (EEG), brain waves that are 8 to 13 Hz, indicating a relaxed, waking state.

ALS Amyotropic lateral sclerosis.

alternating role Pattern characterized by periodic switching from one type of behavior to another.

altruism Regard for and dedication to the welfare of others; a term originated by the French philosopher Auguste Comte (1798–1857). In psychiatry, the term is closely linked with ethics and morals; Freud recognized altruism as the only basis for the development of community interest; Bleuler equated it with morality.

Alzheimer disease DEMENTIA OF THE ALZHEIMER TYPE.

AMA American Medical Association.

ambivalence Presence of strong and often overwhelming simultaneous contrasting attitudes, ideas, feelings, and drives toward an object, person, or goal; term coined by Bleuler, who differentiated three types: affective ambivalence, intellectual ambivalence, and ambivalence of the will. Commonly seen in schizophrenia and borderline personality disorder.

ambulatory schizophrenia A descriptive term refers to schizophrenic mental illness that is sufficiently well compensated so as not to require continuous hospitalization.

amentia Lack of intellectual development as a result of inadequate brain tissue. *Compare*

dementia.

American Academy of Child and Adolescent Psychiatry (AACAP) Organization of child and adolescent psychiatrist, founded in 1953, whose objectives are to study and to advance medical contributions to the knowledge and treatment of psychiatric illnesses of children and adolescents.

American Association of Chairmen of Departments of Psychiatry (AACDP) Organization of chairpersons of departments of psychiatry, founded in 1967, that concentrates on promoting medical education, research, and patient care, especially as they relate to psychiatry; prompting the growth and development of psychiatry; and providing a forum for discussion among chairpersons.

American Association of Directors of Psychiatric Residency Training, Inc. (AADPRT) Organization of training directors of U.S. psychiatric residency programs, founded in 1973, whose mandate is to oversee standards of approved residency criteria and to further psychiatric education.

American Board of Psychiatry and Neurology (ABPN) Group, founded in 1934, that sets the standards for teaching and training in psychiatry and neurology in American hospitals and training centers. Its members are from the American Medical Association, American Psychiatric Association, and American Neurological Association. It conducts periodic examinations of those who have satisfactorily completed their training; upon passing the examination, candidates are certified to practice the specialty of psychiatry and/or neurology and/or child psychiatry.

American College of Physicians (ACP) Honorary organization of internists and related medical specialists, founded in 1915, that sponsors postgraduate courses for physicians as well as teaching and research scholarships.

American College of Psychiatrists (ACP) Honorary organization of senior psychiatrists, founded in 1963, established to honor persons who have made a significant contribution to the study and development of psychiatry in the United States.

American Law Institute formulation of insanity Legal definition of insanity whereby a person is absolved of criminal responsibility if, as a result of a mental disorder, he "lacks substantial capacity either to appreciate the wrongfulness of his conduct or to conform his conduct to the requirements of law." *See also* Durham rule; M'Naghten rule.

American Medical Association (AMA) National organization of medical doctors of all specialties, founded in 1847, whose mandate is to monitor the educational, training, and research concerns of the medical profession; plays an advisory role in national legislation related to health care and disseminates medical information.

American Neurological Association (ANA) National organization of neurologists, founded in 1875, that conducts research and educational programs; also elects members to the American Board of Psychiatry.

American Psychiatric Association (APA) Most important and largest professional organization of American physicians who specialize in the practice of psychiatry. The organization was founded in 1844 by 13 superintendents ("original thirteen") as the Association of Medical Superintendents of American Institutions for the Insane; in 1891, the name was changed to the American Medico-Psychological Association and in 1921 to its

present name. Its purpose is to advance the development of psychiatric education, research, and practice in the United States.

American Psychoanalytic Association (APA) National organization of graduates of various psychoanalytic training programs, founded in 1911, mandated to establish and maintain standards for the field of psychoanalysis and the training of psychoanalysts and to encourage research; historically represents orthodox schools of classical psychoanalysis derived from freudian theory.

American Psychological Association Founded in 1892, large stand most important organization of American psychologists who specialize in experimental, clinical, research, and educational psychology, and many other subspecialties in psychology; sometimes known as the "big APA" to distinguish it from the smaller American Psychiatric Association.

amimia Lack of ability to make gestures or to comprehend those made by others.

amines Organic compounds derived from ammonia by replacing hydrogen atoms with organic groups, such as the catecholamines.

amino acid Organic acid in which one of the CH hydrogen atoms has been replaced by NH_2; structured building blocks of proteins that function as neurotransmitters similar to the biogenic amines, probably in 60 to 70 percent of the synapses in the brain.

amnesia Disturbance in memory manifested by partial or total inability to recall past experiences. For types of amnesia, see the specific term.

amnestic aphasia, amnesic aphasia Disturbed capacity to name objects, even though they are known to the patient. Also called anomic aphasia.

amnestic disorder Mental syndrome in which impairment of memory is the predominant cognitive defect; in DSM-5, called major or minor neurocognitive disorder due to another medical condition.

amok Condition, usually associated with Malayan men, characterized by a sudden, unprovoked outburst of wild rage usually culminating in homicide.

amphetamine Central nervous system stimulant with a chemical structure and action closely related to the sympathomimetic amines; used to treat depression, attention-deficit hyperactivity disorder, hypersomnia, and narcolepsy; may be abused to prolong wakefulness and attentiveness.

amygdala Part of the basal ganglia complex in the dorsomedial portion of the temporal lobe: it and the hippocampus are major components of the limbic system and have been implicated in the production of memory, emotions, and violent behavior.

amyotropic lateral sclerosis (ALS) Degenerative neuromuscular disease, caused by a slow virus, with an insidious onset beginning between the ages of 50 and 70; wasting begins in upper limbs, eventually spreading to the tongue, mouth, and palate. Degeneration of both pyramidal tracts above the medulla produces pseudobulbar palsy, which, when severe, results in impaired voluntary control of emotional reactions, which consequently, may be exaggerated, explosive, and quite inappropriate. Also called Lou Gehrig disease.

Amytal interview Psychiatric interview conducted with the patient having received an intravenous infusion of amobarbital (Amytal), which produces an altered state of consciousness during which repressed memories and feelings may be recalled or reexperienced; used in the treatment of dissociative states, psychogenic amnesia, and

conversion disorder.

ANA American Neurological Association.

anaclitic Depending on others, especially as the infant on the mother; anaclitic depression in children results from an absence of mothering.

anaclitic therapy Psychotherapy characterized by allowing the patient to regress; used mainly in the treatment of psychosomatic disorders.

anal character Freud's term to describe persons who are perfectionistic, parsimonious, punctual, and stubborn.

analeptic PSYCHOMOTOR STIMULANT.

anal erotism Pleasurable experience felt by a child in the anal area during the anal phase of psychosexual development.

analgesia State in which one feels little or no pain.

anal phase Second stage in psychosexual development, occurring when a child is between the ages of 1 and 3 years. During that period, activities, interests, and concerns are centered on the anal zone, and the pleasurable experience felt in that area is called anal erotism. *See also* infantile sexuality; psychosexual development.

analysand In psychiatry, the one who is being analyzed.

analysis in depth PSYCHOANALYSIS.

analysis of variance (ANOVA) Statistical technique designed to compare two or more groups of observations by which sets of measurements are investigated to find out to what extent they are determined by experimental influences or chance influences. *See also* variance.

analytic psychology Jung's system of psychology, characterized by a belief in the collective unconscious or objective psyche, the archetype, and the complex.

anamnesis Medical history of a patient, particularly used in connection with the patient's own recollections.

anancasm Repetitious or stereotyped behavior or thought usually used as a tension-relieving device. Alternate spelling, anankastic.

anankastic personality OBSESSIVE-COMPULSIVE PERSONALITY.

ancillary care Auxiliary, accessory, or secondary services provided in the treatment or rehabilitation of patients.

androgen Hormone that is produced in the testes and the adrenal glands and that stimulates activity of the accessory male sex organs, promotes development of male sex characteristics, or prevents changes in the latter following castration; for example, testosterone. *Compare* estrogen.

androgyny Combination of female and male characteristics in one person. *See also* bisexuality.

androphobia Irrational fear of men.

anergia Lack of energy.

anergic schizophrenic BURNED-OUT SCHIZOPHRENIC.

anesthesia Absence of sensation. *See also* hysterical anesthesia.

angel dust Street term for phencyclidine (PCP).

anhedonia State of being unable to experience pleasure. *Compare* hedonism.

anima According to jungian psychology, the person's inner self, as opposed to the self that one

presents to the outside world (persona).

Anna O A patient of Joseph Breuer and Sigmund Freud whose arm became paralyzed whenever she was unconsciously reminded of her feelings for her father (she had been sitting at his bedside with her arm pressed against the side of his bed at the time of his death). The motor disturbance represented an action that was associated with an objectionable infantile sexual impulse.

anniversary reaction Behavior, symptoms, or dreams that occur at the same time of year as a significant past event.

anomia Inability to recall the names of objects.

anomic aphasia AMNESTIC APHASIA.

anomie Term, popularized by Émile Durkheim as a major cause for suicide, referring to a sense of alienation and despair resulting from the loss or weakening of previously valued norms, ideals, or goals.

anorectic Drug that decreases appetite.

anorexia nervosa Serious and sometimes life-endangering condition characterized by a disturbed body image and self-imposed severe dietary limitation, usually resulting in serious malnutrition; appetite usually remains intact.

anorgasmia Inability to achieve orgasm in the female. *See also* frigidity.

anosognosia Unawareness or nonacceptance of a neurological deficit.

anosomnia Loss of ability to smell.

ANOVA Analysis of variance.

antagonist In pharmacology, a substance that reduces or blocks the action of another; for example, naloxone (Narcan) is used in the treatment of opioid overdose because it is reverses the effects of narcotics. *Compar*e agonist.

anterograde amnesia Loss of memory for events subsequent to the onset of the amnesia. *Compare* retrograde amnesia.

anthropology Study of humans in relation to distribution, origin, classification, and relationship of races, physical characteristics, environmental and social relations, and culture.

antianxiety drug ANXIOLYTIC.

anticholinergic Drug that blocks the actions of atropine; primarily used in the treatment of medication-induced movement disorders. *See also* antiparkinsonism drug.

anticholinergic effect Effect, such as a dry mouth and blurred vision, due to a blockade of the cholinergic (especially the parasympathetic) nerves; often seen as a side effect of therapy with phenothiazine and tricyclic antidepressant drugs.

anticipation Term implying careful planning or worrying, and premature but realistic expectation of dire and potentially dreadful outcomes, in goal-directed inner discomfort.

anticonvulsant Any drug with the ability to prevent the excessive discharge of neurons in the brain and thus lower the threshold for convulsions or seizures. Also called antiepileptic drug.

antidepressant drug Any of a class of psychotherapeutic drugs used in the treatment of pathological depression. Two main categories are the cyclic antidepressants and the monoamine oxidase inhibitors; more recently, serotonergic agents such as fluoxetine. Lithium carbonate also has antidepressant activity in certain patients.

antiepileptic drug ANTICONVULSANT.

antimanic drug Drug used to alleviate the symptoms of mania; lithium is particularly effective in preventing relapses in bipolar disorder.

antiparkinsonism drug Drug that relieves the symptoms of parkinsonism and is effective in countering the extrapyramidal side effects often induced by antipsychotic drugs by acting on the central nervous system to diminish skeletal muscle tone and involuntary movements. Levodopa, the agent of choice for Parkinson disease, does not antagonize the extrapyramidal effects of antipsychotic drugs. Anticholinergic agents, such as benztropine and trihexyphenidyl, effectively counteract drug-induced parkinsonism but are now secondary drugs in the treatment of Parkinson disease.

antipsychotic Any of a class of psychotherapeutic drugs used to treat psychosis, particularly schizophrenia. Also called neuroleptic; major tranquilizer; ataractic drug.

antisocial behavior Behavior pattern of such marginal criminal types as racketeers, prostitutes, and illegal gamblers. Formerly called sociopathic behavior. *See also* antisocial personality disorder.

antisocial personality disorder Disorder characterized by an inability to get along with other members of society and by repeated conflicts with individual persons and groups; common attributes include impulsiveness, egocentricity, hedonism, low frustration tolerance, irresponsibility, inadequate conscience development, exploitation of others, and rejection of authority and discipline. *See also* antisocial behavior.

Anton syndrome Disorder caused by a lesion in the occipital lobe of the brain that causes blindness which the patient denies experiencing.

anxiety Unpleasurable emotional state associated with psychophysiological changes in response to an intrapsychic conflict; in contrast to fear, the danger or threat in anxiety is unreal. Psychological changes consist of an uncomfortable feeling of impending danger, an overwhelming awareness of being powerless, inability to perceive the unreality of the threat, prolonged feeling of tension, and exhaustive readiness for the expected danger; physiological changes consist of increased heart rate, disturbed breathing, trembling, sweating, and vasomotor changes. *Compare* fear.

anxiety disorder Disorder in which anxiety is the most prominent disturbance or in which patients experience anxiety if they resist giving in to their symptoms. In DSM, anxiety disorders include phobic disorder, panic disorder with agoraphobia, agoraphobia without history of panic disorder, social phobia, simple phobia, generalized anxiety disorder, obsessive-compulsive disorder, and posttraumatic stress disorder.

anxiety disorders of childhood or adolescence Disorders characterized by chronic anxiety, unrealistic fears, hypersensitive autonomic responses, or fear of leaving home; patients are usually immature, inhibited, timid, approval-seeking, and apprehensive in new situations or places. In DSM, anxiety disorders of childhood or adolescence include separation anxiety disorder, avoidant disorder of childhood or adolescence, and overanxious disorder.

anxiety hysteria PHOBIC DISORDER.

anxiety neurosis As initially described by Sigmund Freud, a syndrome characterized by general irritability, anxious expectation, pangs of conscience, anxiety attacks, and phobias. Called generalized anxiety disorder in DSM.

anxiety state Disorder characterized by panic and anxious overconcern; somatic symptoms

are often prominent.

anxiolytic Any of a class of psychotherapeutic drugs used to treat pathological anxiety and its related symptoms; includes the benzodiazepines and buspirone. Also called antianxiety drugs; minor tranquilizers.

APA American Psychiatric Association; American Psychoanalytic Association; American Psychological Association.

apathetic withdrawal *See* withdrawal.

apathy Want of feeling or affect, a lack of interest and emotional involvement in one's surroundings; observed in certain types of schizophrenia and depression.

Apgar scores Measurements taken 1 and 5 minutes after birth to determine physical normality in the neonate; scores are based on color, respiratory rate, heartbeat, reflex action, and muscle tone.

aphasia Any disturbance in the comprehension or expression of language due to a brain lesion. For types of aphasia, see the specific term.

aphonia Loss of voice.

aphrodisiac Substance or drug that increases sexual desire.

apnea Temporary cessation of breathing. *See* obstructive sleep apnea syndrome.

apoplexy Old classical term for cerebrovascular accident.

apperception Awareness of the meaning and significance of a particular sensory stimulus as modified by one's own experiences, knowledge, thoughts, and emotions. *See also* perception.

appetitive phase First stage in the sexual response cycle, as described by Masters and Johnson, characterized by sexual fantasies and the desire to have sex.

appropriate affect Emotional tone in harmony with the accompanying idea, thought, or speech.

apraxia Inability to perform a voluntary purposeful motor activity; cannot be explained by paralysis or other motor or sensory impairment. In constructive apraxia, a patient cannot draw two- or three-dimensional forms.

aptitude Potential to acquire competence or skill if trained. *Compare* ability.

archetype Jung's term for the symbols used by individuals to express the collective experience of the human race. *See also* collective unconscious.

arithmetic disorder DEVELOPMENTAL ARITHMETIC DISORDER.

arousal The establishment and maintenance of a physiological or psychological state of being awake, alert, or responsive to stimuli. *See also* hyperarousal.

arteriosclerosis Degenerative condition marked by loss of elasticity, hardening, and thickening of the arteries, with progressive narrowing of their lumina and gradual obstruction to blood flow; organic mental disorder may develop when cerebral arteries are involved in the degenerative process.

artificial intelligence Computer production of a pattern of output that would be considered intelligent if displayed by human beings.

art therapy Treatment procedure that uses the spontaneous creative work of the patient; for example, persons make and analyze drawings, which are often expressions of their underlying emotional problems.

asociality Indifference to social values or customs, withdrawal from society, as seen in a recluse or a regressed schizophrenic person.

Asperger syndrome Term previously used to describe a condition that develops in infancy and is associated with impaired social interaction and restricted and stereotyped patterns of behavior, but without significant delays in language or cognitive development. In DSM-5, the behavioral condition of Asperger syndrome is subsumed in Level I of the autism spectrum. *See also* autism spectrum disorder.

assertiveness training Behavior therapy that involves teaching frank and direct interpersonal expression of feelings and ideas.

assessment The evaluation of a patient that leads to diagnosis and treatment.

assimilation Subjective process that involves filtering the world through one's own system of knowledge, that is, taking in of new experience through one's established mental structure.

association Relation between ideas or feelings. *See* clang association; loosening of associations.

Association of American Medical Colleges (AAMC) Organization of American medical schools and teaching hospitals, founded in 1876, concerned with the advancement and improvement of medical education, biomedical research, and health care in the United States.

astasia-abasia Inability to stand or walk in a normal manner, even though normal leg movements can be performed in a sitting or lying down position; seen in conversion disorder.

astereognosis Inability to identify familiar objects by touch; cannot be explained by an impairment of elementary sensory pathways or a depressed sensorium. *See also* neurological amnesia.

asthenic personality Disorder characterized by lack of enthusiasm, fatigability, lack of capacity for enjoyment, and low tolerance for stress. *See also* adynamia; anhedonia.

astraphobia, astrapophobia Abnormal fear of thunder and lightning.

asyndesis Disorder of language, commonly seen in schizophrenia, in which the patient combines unconnected ideas and images.

ataque de nervios A syndrome, usually found in Latin American cultures, characterized by uncontrollable shouting, crying, trembling, and physical or verbal aggression.

ataractic drug ANTIPSYCHOTIC.

ataxia Lack of coordination, either physical or mental. **1.** In neurology, refers to loss of muscular coordination. **2.** In psychiatry, the term "intrapsychic ataxia" refers to lack of coordination between feelings and thoughts; seen in schizophrenia and in severe obsessive-compulsive disorder.

athetosis Movement disorder characterized by involuntary, slow, writhing, worm-like movements of the fingers and extremities; a form of extrapyramidal dyskinesia.

atonia Lack of muscle tone. *See* waxy flexibility.

attachment According to John Bowlby, the affective tie between infant and mother or primary caregiver, the disruption of which heightens vulnerability to depression or anxiety. *See also* bonding.

attachment disorder of infancy Absence or disruption of behaviors that serve to anchor the infant to its mother or primary caregiver and that produce abnormal reactions in the infant's behavior.

attachment learning Theory that learning is enhanced in the presence of the person to whom one is attached or affiliated; applies especially to infant learning.

attention Concentration; the aspect of consciousness that relates to the amount of effort exerted in focusing on certain aspects of an experience, activity, or task.

attention-deficit hyperactivity disorder (ADHD) DSM-5 diagnosis for a childhood mental disorder characterized by a developmentally inappropriate short attention span, hyperactivity, and poor concentration that cause clinically significant impairment in functioning. The category subsumes abnormal behavior patterns that had historically been referred to by a variety of names, including hyperkinetic syndrome, hyperactive child syndrome, and minimal brain dysfunction. If hyperactivity is not present, the disorder is called attention-deficit disorder (ADD).

attitude Preparatory mental posture with which one receives stimuli and reacts to them.

attributable risk Absolute incidence of a disease in exposed individuals that can be attributed to the exposure.

atypical In psychiatry, denoting an unusual or rare variant of a known disorder.

atypical child Child with distorted personality development; often used in connection with brain-damaged or autistic children.

atypical depression Diagnosis used to describe depression characterized by weight gain and hypersomnia rather than weight loss and insomnia.

atypical psychosis Disorder whose psychotic features are related to a specific culture (e.g., koro); a certain time or event (e.g., postpartum psychosis); or a unique symptomatology (e.g., Capgras syndrome). Also called "psychotic disorder not otherwise specified" in DSM-5.

audit *See* medical audit.

auditory agnosia The inability to recognize specific sounds in the context of intact hearing.

auditory amnesia NEUROLOGICAL AMNESIA.

auditory hallucination Hallucination primarily involving sound.

aura 1. Warning sensations that a person usually experiences just before an epileptic seizure. **2.** Sensory prodrome that precedes a classic migraine headache.

authority figure Real or projected person in a position of power; transferentially, a projected parent.

autism *See* autistic thinking; infantile autism.

autistic spectrum disorder A range of behavioral abnormalities that begin in infancy and are characterized by deficits in social communication; failure to communicate; lack of empathy; abnormalities in eye contact; repetitive patterns of activities; stereotyped or absent speech; and either hyper- or hypoarousal to the environment. In DSM-5, autism spectrum disorder is classified into three levels based on degrees of impairment in speech and social interaction. *See also* Asperger syndrome; Rett syndrome; childhood disintegrative disorder.

autistic thinking Thinking in which the thoughts are largely narcissistic and egocentric, with emphasis on subjectivity rather than objectivity, and without regard for reality; used interchangeably with autism and dereism.

autoerotic asphyxia Asphyxia caused by intentional strangulation while masturbating in order to intensify orgasm through reduction of oxygen blood flow to the brain. High risk of death.

autoerotism 1. Sexual arousal of oneself without the participation of another person; term coined by Havelock Ellis, which is presently used interchangeably with masturbation. **2.** In psychoanalysis, a primitive phase in object relationship development preceding the narcissistic stage; in narcissism, there is a love object, but there is no love object in autoerotism.

autogynephilia A gender identity disorder in which a male becomes sexually aroused by imagining themselves transformed into a woman.

automatic judgment Reflex performance of a way of acting based on a fixed set of beliefs or values.

automatic obedience Compliance to commands without exercising critical judgment; a feature of catatonia.

automatism Undirected behavior not controlled by conscious thought.

automysophobia Abnormal fear of smelling bad or of being unclean.

autonomic nervous system Part of the nervous system that functions outside of consciousness and that directs such visceral functions as respiration, heart rate, and digestion; subdivided into sympathetic (adrenergic) and parasympathetic (cholinergic) divisions.

autonomic side effect Disturbance of autonomic nervous function that may result from the use of antipsychotic or antidepressant drugs, particularly the phenothiazine and tricyclic derivatives. Common side effects of such agents include orthostatic hypotension, blurred vision, nasal congestion, dry mouth, and constipation; such anticholinergic effects are more prominent with tricyclic antidepressants than with phenothiazines.

autonomous ego function Operation present at birth, such as perception, intuition, thinking, speaking, language, and aspects of motor development; first described by Heinz Hartmann. *See also* ego psychology.

autophobia Abnormal fear of being alone or of self.

autoplasty Adaptation to stress by changing intrapsychic processes. *Compare* alloplasty.

autoscopic psychosis Rare disorder, not officially listed in DSM-5, in which the patient sees a phantom or specter of his or her own body; usually psychogenic in origin, but may also be caused by lesions in the temporoparietal lobe of the brain.

autotopagnosia Inability to recognize and identify the parts of one's own body. *See also* finger agnosia.

auxiliary therapist Co-therapist. *See also* co-therapy.

average Central value in a frequency distribution around which other values are distributed. Also called the mean.

aversive conditioning AVERSIVE THERAPY.

aversive therapy Behavior therapy that involves the repeated coupling of an unpleasant or painful stimulus, such as an electric shock, with an undesirable behavior pattern in an effort to eliminate the undesirable behavior. Also called aversive conditioning.

avoidance and escape learning In escape learning, an animal learns a response to get out of a place where it does not want to be; for example, it jumps off an electric grid whenever the grid is charged. Avoidance learning requires an additional response: the same animal on the grid will learn to avoid a shock if it quickly pushes a lever when a light signal goes on. Escape learning and avoidance learning are two forms of aversive control using negative reinforcement; behavior that terminates the source of aversive stimuli is strengthened and

maintained.

avoidant attachment According to Mary Ainsworth, a form of attachment in which children avoid returning to their mothers after being separated. Also called shyness disorder. *Compare* secure attachment. *See also* resistant attachment.

avoidant disorder of childhood or adolescence Disorder of childhood or adolescence characterized by a persistent withdrawal or excessive shrinking from strangers.

avoidant personality disorder Personality disorder characterized by low self-esteem, hypersensitivity to rejection, and social withdrawal, but with a desire for affection and acceptance.

Avoidant/Restrictive Food Intake Disorder Disorder characterized by an avoidance or lack of interest in food, resulting in severe weight loss, nutritional deficiency, and failure to thrive; a DSM-5 diagnosis. Previously called feeding disorder of infancy or early childhood.

axon Single structure in a nerve cell that conducts nervous impulses away from the cell body and its dendrites.

aypnia Insomnia; inability to sleep.

B

barbiturate Any of a class of sedative–hypnotic drugs derived from barbituric acid that acts as a central nervous system depressant and readily produces psychic and physical dependence; for example, phenobarbital, secobarbital, thiopental sodium.

basal ganglia Collectively, the corpus striatum, substantia nigra, subthalamic nucleus, and substantia innominata, which are involved in disorders of movement, thought processes, affect, cognition, and symptoms of psychosis.

basic anxiety As conceptualized by Karen Horney, the main-spring from which neurotic trends get their intensity and pervasiveness, characterized by vague feelings of loneliness, helplessness, and fear of a potentially hostile world. *See also* anxiety.

basic benefits Insurance term that refers to minimum payments or coverage for medical or psychiatric care.

bathophobia Abnormal fear of depths.

battered child syndrome Physical damage to a child sustained as a result of repeated beatings, usually by a parent, parent surrogate, or relative. *See also* abuse.

battery tests Individual tests used together make up a psychological battery; for example, Halstead–Reitan Neuropsychological Battery. A test battery can give more information about different areas of function than an individual test and can increase the level of confidence if there is a positive correlation between them.

BEAM Brain electrical activity mapping.

Beard, George M. (1839–1883) American psychiatrist who in 1869 introduced the term "neurasthenia."

Beers, Clifford (1876–1943) Author of *A Mind That Found Itself* (1909) and who is generally considered to have founded the mental hygiene movement.

behavior Sum total of the psyche that includes impulses, motivation, wishes, drives, instincts, and cravings, as expressed by a person's behavior or motor activity. Also called conation.

behavioral disorders of childhood *See* disruptive behavior disorders.

behavioral medicine Application of principles of behavior therapy to the prevention, diagnosis, treatment, and rehabilitation of medical disorders; applies to such areas as stress prevention and reduction, pediatric and adult patient management and compliance, pain control, and life-style modification. Biofeedback relaxation training, behavior therapy, and hypnosis are important treatment modalities.

behavioral neurology Field of study that attempts to combine psychiatry and neurology to explain abnormal behavior as a disorder of brain metabolism or as a result of structural alterations of the central nervous system.

behavioral science Scientific discipline concerned with any aspect of human behavior, including interpersonal relationships, development, values, experiences, and activities; the behavioral sciences include sociology, psychology, psychiatry, anthropology, and ethology.

behavior contract Therapeutic strategy that rewards a person for a desired behavior.

behaviorism School of psychological thought founded by John B. Watson that regards only measurable and observable behavior as the appropriate subject matter for human psychology; holds that human behavior can be described in terms of principles that do not require consideration of unobservable mental events, such as ideas and emotions. *See also* behavior therapy.

behavior modification *See* behavior therapy.

behavior therapy Psychiatric treatment modality that focuses on overt and objectively observable behavior and uses various conditioning techniques derived from learning theory to directly modify the patient's behavior; aims exclusively at symptomatic improvement, without addressing psychodynamic causation. Behavior therapy techniques include assertiveness training, aversive therapy, flooding, implosion, reciprocal inhibition, and desensitization, shaping, and systematic desensitization. *See also* behaviorism; conditioning.

Bender-Gestalt Test Psychological test, designed by Lauretta Bender, which measures the subject's ability to reproduce a set of geometric designs: is useful for measuring visuomotor coordination as a means of detecting brain damage. Also called visual motor gestalt test.

benzodiazepines Any of a group of drugs used mainly to treat anxiety; also used as sedatives, muscle relaxants, anticonvulsants, anesthetics, and hypnotics.

bereavement Feeling of grief or desolation, especially at the death or loss of a loved one. *See also* mourning.

bestiality Sexual deviation in which a person engages in sexual relations with an animal. *See also* zoophilia.

beta (a) error TYPE 2 ERROR.

beta waves As measured by an electroencephalogram (EEG), brainwaves that are between l4 and 30 Hz; altered in certain epileptic conditions.

bias In statistics, an error in the construction of the assessment instrument that favors one outcome over another.

binge eating disorder Eating disorder characterized by recurrent episodic, uncontrolled, compulsive, and rapid ingestions of large amounts of food over a short period of time (binge eating) followed by feelings of guilt; a DSM-5 diagnosis, differentiated from bulimia nervosa by an absence of compulsive compensatory behaviors to achieve weight loss.

binge eating, binge and purge *See* bulimia nervosa.

biodynamics System of psychoanalytic psychiatry, introduced by Jules H. Masserman, which holds that all behavior results from the organism's genetic potentialities, from the various paths and speed of maturation, and from its unique experiences.

biofeedback Provision of information to a person regarding one or more physiological processes in an effort to enable the person to gain some element of voluntary control over bodily functions that normally operate outside consciousness. *See also* learned autonomic control.

biogenic amine Endogenous compound believed to play an important role in brain function as a central synaptic neurotransmitter substance; biogenic amines include the catecholamines (epinephrine, norepinephrine, dopamine) and serotonin, an indolamine.

biogenic amine hypothesis Theory of mood disorders which proposes that depression is associated with an absolute or relative deficiency of catecholamines, particularly norepinephrine, at functionally important receptor sites in the brain; conversely, elation may be associated with an excess of such amines.

biological psychiatry School of psychiatric thought emphasizing biochemical, neurological, and pharmacological factors in the causation and treatment of psychiatric disorders.

biological rhythm Periodic variation in physiological and psychological functions. For example, a circadian rhythm shows aperiodicity of about 24 hours; an ultradian rhythm has a cycle shorter than 1 day; an infradian rhythm is longer than 1 day.

bipolar I Bipolar disorder in which a patient meets the criteria for a full manic episode, usually sufficiently severe to require hospitalization.

bipolar II Bipolar disorder in which a patient has at least one depressive episode and at least one hypomanic episode.

bipolar disorder A mood disorder in which the patient exhibits both manic and depressive episodes separated by various lengths of time. Previously referred to as manic–depressive illness.

birth trauma Otto Rank's term to describe the birth process, which he considered the basic source of anxiety in human beings.

bisexual Denoting or characterized by sexual attraction to and sexual contact with members of both sexes.

bisexuality Existence of the qualities of both sexes in the same person. Freud postulated that the sexes differentiated both biologically and psychologically from a common core; that differentiation between the two sexes was relative, rather than absolute; and that regression to the common core occurs to varying degrees in both normal and abnormal conditions. *Compare* bisexual; heterosexuality; homosexuality. *See also* androgyny.

bizarre delusion False belief that is patently absurd or fantastic.

blackout *See* alcoholic blackout.

black-patch syndrome Psychosis induced by sensory deprivation as a result of eye patches used after cataract surgery.

blank screen Neutral backdrop on which the patient projects a gamut of transferential irrationalities; passivity of analysts allows them to act as blank screens.

Bleuler, Eugen (1857–1939) Swiss psychiatrist known for his important studies in schizophrenia, a term he coined and preferred over the earlier term "dementia praecox."

blind spot In psychiatry, an area of a person's personality of which he or she is totally unaware; unperceived areas are repressed, since their recognition would arouse painful or unpleasant emotions. In the course of group or individual psychotherapy, such blind spots often appear obliquely as projected ideas, intentions, and emotions. *See also* projection; scotoma.

blocking Involuntary cessation or interruption of thought processes or speech because of unconscious emotional factors. Also called thought deprivation.

blood–brain barrier Series of biologically active membranes that regulate the passage of substances into and out of the central nervous system; maintains a chemical homeostasis between the cerebrospinal fluid and the venous and arterial systems.

blood level In psychiatry, the measurement of the concentration of drugs in plasma or serum to determine their therapeutic effect.

blunted affect Disturbance of affect manifested by a severe reduction in the intensity of externalized feeling tone; one of the fundamental symptoms of schizophrenia, as outlined by Eugen Bleuler.

board-certified psychiatrist Psychiatrist who has passed the examinations of the American Board of Psychiatry and Neurology (ABPN). *See also* accreditation.

board-eligible psychiatrist Psychiatrist who has finished an approved 4-year residency program and is eligible to take the examinations of the American Board of Psychiatry and Neurology.

body dysmorphic disorder Disorder characterized by an unrealistic concern that part of the body is deformed. Included on the obsessive-compulsive spectrum in DSM-5. Also called dysmorphophobia.

body-focused repetitive behavior disorders Descriptive term for a group of behaviors consisting of compulsions that damage one's own body, such as skin biting, lip biting, or nail biting; not a diagnosis in DSM-5.

body image Conscious and unconscious perception of one's body at any particular time.

body language Expression of thoughts and feelings by body posture and movement.

bonding Nonsexual feelings of love, respect, and communality among individuals of the same sex. Specifically, a mother's feelings for her infant, but often used to refer to the general mother–infant relationship; fathers often bond to infants. *See also* attachment.

borderline intellectual functioning A term previously attributed to individuals with below average cognitive ability (I.Q. in the 70–80 range) whose intellectual impairment did not meet criteria for intellectual disability (I.Q. below 70). This condition is not clearly differentiated from mild intellectual disability and may be the focus of clinical attention, however, no definitive criteria for diagnosis exist in DSM-5. *See also* intellectual disability.

borderline personality disorder Personality disorder marked by instability in various areas, impulsiveness, suicidal acts, self-mutilations, identity problems, and feelings of emptiness or boredom.

borderline psychosis BORDERLINE STATE.

borderline state State in which the symptoms are so unclear or transient that it is difficult to classify the patient as psychotic or nonpsychotic. Also called borderline psychosis.

boufée délirante A syndrome, observed in West Africa and Haiti, characterized by a sudden

outburst of agitated and aggressive behavior, marked confusion, and psychomotor excitement. It is sometimes accompanied by visual or auditory hallucinations or paranoid ideations. May resemble an episode of brief psychotic disorder.

Bowlby, John (1907–1990) British psychoanalyst who formulated the theory of attachment to describing the mother–infant relationship. *See also* attachment; bonding.

bradykinesia Slowness of motor activity, with a decrease in normal spontaneous movement.

bradylalia Abnormally slow speech, common in depression.

bradylexia Inability to read at normal speed.

brain That part of the central nervous system contained within the cranium; includes the cerebrum, midbrain, cerebellum, pons, and medulla.

brain abscess Localized, suppurative, infectious process involving a part of the brain; may produce an organic mental disorder as well as focal neurological signs.

brain electrical activity mapping (BEAM) Response of brain electrical activity to sensory stimuli which is recorded by computer averaging techniques that produce a topographic color map of brain activity.

brain imaging Radiological techniques that allow assessment in vivo of structural and functional neuroanatomy; include computed tomography (CT), magnetic resonance imaging (MRI), regional cerebral blood flow (rCBF), positron emission tomography (PET), and brain electrical activity mapping (BEAM).

brain metabolism Chemical changes in the brain that convert small molecules into large ones (anabolism) and large molecules into small ones (catabolism), and that account for cellular function and repair.

brain stem Area of the brain comprising the mesencephalon, pons, and medulla oblongata, whose basic functions concern respiration, cardiovascular activity, sleep, and consciousness; also the site of the ascending biogenic amine (dopamine, noradrenalin, serotonin) pathways to higher brain areas.

brain syndrome *See* organic mental disorder.

brainwashing Any technique designed to manipulate human thought or action against the desire, will, or knowledge of the person involved; usually refers to systematic efforts to indoctrinate nonbelievers. In DSM-5 diagnosed as "identity disturbance due to prolonged and intense coercive persuasion." Also called coercive persuasion.

brain wave Rhythmic wave-like electrical pattern in an encephalogram.

Brawner decision *See* Model Penal Code.

breathing-related sleep disorder Characterized by sleep disruption leading to excessive sleepiness or insomnia caused by a sleep-related breathing disturbance; three types of breathing-related sleep disorders include obstructive sleep apnea, central sleep apnea, and sleep-related hypoventilation.

Breuer, Josef (1842–1925) Viennese physician who collaborated with Freud in studies of cathartic therapy, as reported in *Studies on Hysteria* (1895); he withdrew as Freud proceeded to introduce psychoanalysis, but left important imprints on that discipline, such as the concepts of the primary and secondary processes.

brief psychotherapy Psychotherapy in which the sessions are limited to 10 to 15 in number and during which time attempts to modify behavior occur; used in both individual and group

settings.

brief psychotic disorder DSM-5 category for a psychosis of less than 4 weeks' duration, with sudden onset after a major stress. Also called brief reactive psychosis.

brief reactive psychosis *See* brief psychotic disorder.

Brigham, Amariah (1798–1849) One of the founders of the American Psychiatric Association and the first editor of its journal.

Brill, A. A. (1874–1948) First American analyst (1908) who translated several of Freud's more important works, was active in the formation of the New York Psychoanalytic Society (1911), and remained a foremost propagator of psychoanalysis as a lecturer and writer.

Briquet syndrome Alternate term for somatic symptom disorder.

broad affect Normal range (varies both within and among cultures) of the expression of a subjectively experienced feeling state (e.g., euphoria, anger, sadness); normal expression of affect involves variability in facial expression, pitch of voice, and hand and body movements.

Broca aphasia MOTOR APHASIA.

brooding compulsion INTELLECTUALIZATION.

bruxism Grinding or gnashing of the teeth, typically occurring during sleep.

bulimia nervosa Recurrent episodic, uncontrolled, compulsive, and rapid ingestion of large amounts of food over a short period of time (binge eating) followed by self-induced vomiting, use of laxatives or diuretics, fasting, or vigorous exercise in order to prevent weight gain (binge and purge); weight is not severely lowered.

burned-out schizophrenic Chronic schizophrenia characterized by a person who is apathetic and withdrawn, with minimal florid psychotic symptoms but with persistent and often severe schizophrenic thought processes. Also called anergic schizophrenia.

burnout Syndrome of physical or emotional exhaustion considered to be a stress reaction to unrelenting performance and emotional demands stemming from one's occupation; symptoms include impaired work performance, fatigue, insomnia, increased susceptibility to physical illness, and reliance on alcohol (or other substances) with a tendency to escalation into physiological dependency or possibly suicide.

butyrophenone Any of a class of antipsychotic drugs of which haloperidol (Haldol) is a representative example.

C

C.A. Chronological age.

cacodemonomania Condition in which patients think they are possessed by a devil or other evil spirit.

caffeine Alkaloid found in coffee, tea, and cola drinks which acts as a psychomotor stimulant; caffeine intoxication results from the excessive ingestion of caffeine.

Canadian Psychiatric Association (CPA) National organization of Canadian psychiatrists whose purpose is to further the development of psychiatric education, research, and practice in Canada.

cannabis *See* marijuana.

canonical correlation Multivariate technique used in statistics for simultaneously finding the relationship of linear combinations of two or more predictors and two or more outcomes.

Capgras' syndrome Delusional disorder in which the patient believes that persons in the environment are not their real selves but are doubles imitating the patient or imposters imitating someone else.

capitation Prepayment fee based on a fixed number of persons that covers all health care services.

care and protection proceedings Intervention by a court on behalf of a child when the parents or caretakers provide inadequately for the child's welfare.

carebaria Sensation of discomfort or pressure in the head.

caregiver 1. Any person involved in the identification, prevention, or treatment of illness or disability. **2.** In child psychiatry, a person who tends to the needs of an infant or child. Also called caretaker.

case control study Study in which the illness is the dependent variable, already being present for the subject to be identified; such studies are therefore retrospective.

case history study Retrospective study that samples persons with a given disorder.

case register Medical records in central data banks.

castration anxiety Anxiety concerning a fantasied loss of or injury to the genitalia.

castration complex In psychoanalytic theory, a group of unconscious thoughts and motives that are referable to the fear of losing the genitalia, usually as punishment for forbidden sexual desires.

catalepsy Condition in which persons maintain the body position into which they are placed; observed in severe cases of catatonic schizophrenia. Also called waxy flexibility; cerea flexibilitas. *See also* command automatism.

cataphasia VERBIGERATION.

cataplexy Temporary sudden loss of muscle tone, causing weakness and immobilization; can be precipitated by a variety of emotional states, and it is often followed by sleep; commonly seen in narcolepsy.

catastrophic anxiety Anxiety associated with cognitive disorders when patients are aware of their defects in mentation; the anxiety may be overwhelming.

catathymia Situation in which elements in the unconscious are sufficiently affect-laden to produce changes in conscious functioning.

catathymic crisis Suddenly occurring, isolated, and nonrepetitive act that develops from a state of intolerable tension.

catatonia Clinical syndrome consisting of marked behavioral irregularities, which may include motoric excitement or immobility, negativism, stupor, echolalia, or echopraxia. In DSM-5, catatonia may be a stand-alone diagnosis; a feature of schizophrenia; associated with other mood/psychotic disorders; associated with medications or substance use; or due to an underlying medical condition.

catatonic excitement Excited, uncontrollable motor activity seen in catatonic schizophrenia. Patients in catatonic state may suddenly erupt into an excited state and may be violent.

catatonic posturing Voluntary assumption of an inappropriate or bizarre posture, generally maintained for long periods of time. May switch unexpectedly with catatonic excitement.

catatonic rigidity Fixed and sustained motoric position that is resistant to change.

catatonic schizophrenia A rare type of schizophrenia; no longer included in DSM-5. *See*

catatonia.

catatonic state State characterized by muscular rigidity and immobility (catatonia), usually associated with schizophrenia.

catatonic stupor Stupor in which patients ordinarily are well aware of their surroundings.

catchment area In psychiatry, a circumscribed geographic area served by a particular community mental health center.

catecholamine Biogenic monoamine containing a catechol group. Norepinephrine, epinephrine, and dopamine are catecholamine with important effects on peripheral and central nervous system activity.

categorical attitude ABSTRACT THINKING.

categorical thought Term introduced by Jean Piaget to refer to abstract thinking. *See also* secondary process.

catharsis Release of ideas, thoughts, and repressed materials from the unconscious, accompanied by an affective emotional response and relief; commonly observed in the course of both individual and group psychotherapy, but can also occur outside therapy. *See also* abreaction; conversational catharsis.

cathexis In psychoanalysis, a conscious or unconscious investment of psychic energy in an idea, a concept, an object, or a person. *Compare* acathexis.

CAT scan Incorrect but commonly used variant of "CT scan." *See also* computed tomography.

cat's cry syndrome *See* cri-du-chat syndrome.

Cattell Infant Intelligence Scale Psychological test used to assess general motor and cognitive development in infants aged 3months to 2 years.

causalgia Burning pain that may be either organic or psychic in origin.

cause–effect relationship Relationship in which one action causes a specific reaction.

causalgia Burning pain that may be organic or psychic in origin.

cenesthesia Change in the normal quality of feeling tone in a part of the body.

central nervous system (CNS) The spinal cord and the brain.

central sleep apnea Dyssomnia characterized by repeated episodes of apnea, shallow breathing, and/or irregular breathing that result from a periodic failure of the central nervous system mechanisms to stimulate breathing. There are three subtypes in DSM-5: idiopathic, Cheyne–Stokes breathing pattern, and comorbid with opioid use. *Compare* obstructive sleep apnea.

cephalalgia Headache.

cerea flexibilitas Condition of a person who can be molded into a position that is then maintained; when an examiner moves the person's limb, the limb feels as if it were made of wax. Also called catalepsy; waxy flexibility. Seen in schizophrenia.

cerebral arteriosclerosis Arteriosclerotic involvement (hardening) of arteries of the brain; may lead to organic mental disorder.

cerebrovascular accident Cerebral hemorrhage, thrombosis, or embolism. Also called stroke; apoplexy.

character Constellation of relatively fixed personality traits and attributes that govern a person's habitual modes of response.

character analysis Psychoanalytic treatment that concentrates on character defenses.

character defense Personality trait that serves an unconscious defensive purpose.

character disorder Pattern of personality characterized by maladaptive, inflexible behavior.

Charcot, Jean M. (1825–1893) French neurologist noted for describing hysteria and treating it by means of hypnosis; Freud based much of his early work on Charcot's pioneering studies in hysteria.

chemical dependence *See* drug dependence.

chemotherapy DRUG THERAPY.

child abuse *See* battered child syndrome.

child analysis Application of psychoanalytic principles to the treatment of the child.

childhood disintegrative disorder A condition defined by a marked regression in several functional domains (e.g., intellectual, social, motor, and language skills) after at least 2 years of normal development; diagnosis is typically made in children who are between 3 and 4 years of age. In DSM-5, childhood disintegrative disorder is a subset of autism spectrum disorder.

childhood schizophrenia DSM term for pervasive developmental disorder. *See* autistic spectrum disorder.

child-onset fluency disorder *See* stuttering.

chi square Statistical technique in which variables are categorized in order to determine whether a distribution of scores is due to chance or to experimental factors.

chlorpromazine Aliphatic phenothiazine derivative used primarily as an antipsychotic agent and in the treatment of nausea and vomiting.

cholinergic Relating to nerve fibers, tracts, or pathways in which synaptic transmission is mediated by acetylcholine; for example, parasympathetic nerve fibers and somatic motoneurons; also refers to processes activated by acetylcholine.

chorea Movement disorder characterized by random and involuntary quick, jerky, purposeless movements. *See also* extrapyramidal dyskinesia; Huntington chorea.

chronic fatigue syndrome An illness characterized by 6 months or more of severe, debilitating fatigue, often accompanied by myalgia, headaches, pharyngitis, low-grade fever, cognitive complaints, gastrointestinal symptoms, and tender lymph nodes. The cause is unknown.

chronobiology Study of biological rhythms that objectively explores and quantifies mechanisms of time structure in biological systems.

chronological age (C.A.) *See* intelligence quotient.

chronophobia Fear of time characterized by panic, anxiety, and claustrophobia. Sometimes called prison neurosis, since almost all prisoners are affected by it in some fashion.

circadian rhythm *See* biological rhythm.

circadian rhythm sleep disorder A set of dyssomnias involving a misalignment between desired and actual sleep periods. Includes delayed sleep phase, jet lag, and work shift.

Circadian Rhythm Sleep–Wake Disorders A group of sleep disorders characterized by disruption in sleep that alters a patient's circadian rhythm and is associated with insomnia or excessive sleepiness. There are six types of circadian rhythm sleep–wake disorders in DSM-5: delayed sleep phase type, advanced sleep phase type, irregular sleep–wake type, non–24-hour sleep–wake type, shift work type, and unspecified type.

circumstantiality Disturbance in the associative thought and speech processes in which a patient digresses into unnecessary details and inappropriate thoughts before communicating the central idea; observed in schizophrenia, obsessional disturbances, and certain cases of dementia. *See also* tangentiality.

civil commitment In statutes governing hospitalization of the mentally ill, a legal term meaning a warrant for imprisonment; as recommended by the American Psychiatric Association and the American Bar Association, "commitment" has been replaced by "hospitalization."

CJD Creutzfeldt–Jakob disease.

claims review Method of peer review consisting of the examination of claims for the reimbursement of treatment after it has been rendered.

clairvoyance Extrasensory perception in which there is an alleged ability to perceive that which is out of sight or cannot be seen.

clang association, clanging Association or speech directed by the sound of a word rather than by its meaning; punning and rhyming may dominate the person's verbal behavior; seen most frequently in schizophrenia or mania.

classical conditioning *See* conditioning.

classic migraine *See* migraine.

claustrophobia Abnormal fear of closed or confining spaces.

client-centered psychotherapy Nondirective form of psychotherapy, originated by Carl Rogers, in which the therapeutic process focuses on the patient's own thinking and feeling which the therapist merely helps to clarify through understanding and empathy; developed as a reaction against the authoritativeness and interpretation of the more traditional psychotherapies.

climacteric Menopause and immediate postmenopausal period in women; sometimes used to refer to the same age period in men.

clinical psychologist Psychologist with additional training and experience in a clinical setting who specializes in evaluation and treatment of mental disorders.

clinical psychology Branch of psychology concerned with the application of psychological principles to the prevention, treatment, and understanding of psychopathology.

clonic convulsion An involuntary, violent muscular contraction or spasm in which the muscles alternately contract and relax. Characteristic phase in grand mal epileptic seizure.

clouding of consciousness Any disturbance of consciousness in which the person is not fully awake, alert, and oriented. Occurs in delirium, dementia, and cognitive disorder.

cluster suicides Suicides that occur together in geographic groups or in temporal association. *See also* copy-cat suicides.

cluttering Disturbance of fluency involving an abnormally rapid rate and erratic rhythm of speech that impedes intelligibility; the affected individual is usually unaware of communicative impairment.

CNS Central nervous system.

cocaine Alkaloid, obtained from leaves of the coca plant, whose systemic effects include striking central nervous system stimulation manifested by garrulousness, restlessness, excitement, and feelings of increased muscular strength and mental capacity; also an effective

topical local anesthetic.

coefficient of correlation Statistical term referring to the relation between two sets of paired measurements, correlation coefficients are positive, negative, or curvilinear, depending on whether the variations are in the same direction, the opposite direction, or both directions. They can be computed in a variety of ways, the most common being the product-moment method (r); another method is rank correlation (p). Correlation coefficients are intended to show degree of relation, but they do not imply a causal relationship between variables.

coercive persuasion BRAINWASHING.

coexistent culture Alternative system of values, norms, and patterns for behavior; the group therapy experience often leads to an awareness of other systems as legitimate alternatives to one's own system.

cognition Mental process of knowing and becoming aware; an ego function closely associated with judgment.

cognitive-behavioral therapy (CBT) A form of psychotherapy in which the focus of treatment is on changing the thoughts and behaviors that are related to specific symptoms. The aim of treatment is on symptom reduction and improved functioning.

cognitive development Progressive acquisition of conscious thought and problem-solving abilities that begins in infancy and follows an orderly sequence of more or less discrete stages.

cognitive dissonance Incongruity in a person's beliefs, thoughts, or actions that produces an uncomfortable tension state that the person is motivated to change.

cognitive therapy Structured and directive form of therapy developed by Aaron T. Beck that emphasizes how patients think about themselves and the future; consists of reversing the cognitive triad of negative expression of self, negative expression of future, and negative view of the world.

cognitive triad Belief system described by the psychiatrist Aaron Beck that is present in depressed persons who have a negative view of the self, a negative interpretation of present and past experience, and a negative expectation of the future.

cohesion *See* group cohesion.

cohort Group chosen from a well-defined population; cohort studies provide direct estimates of risk associated with a suspected causal factor.

coitus interruptus Sexual intercourse that is interrupted before the man ejaculates.

cold turkey Abrupt withdrawal from opiates, or any addictive central nervous stimulant, without the benefit of methadone or other drugs; a term originated by drug addicts to describe their chills and consequent goose-flesh. Abstinence-oriented therapeutic communities use this type of detoxification.

collective experience Term coined by S. R. Slavson for the common emotional experiences of a group of people, which, in the setting of group psychotherapy, accelerates the therapeutic process. Identification, mutual support, reduction of ego defenses, sibling transferences, and empathy help integrate the individual member into the group.

collective unconscious Psychic contents outside the realm of awareness that are common to mankind in general. Carl Jung, who introduced the term, believed that the collective unconscious is inherited and derived from the collective experience of the species, that it transcends cultural differences, and that it explains the analogy between ancient mythological

ideas and the primitive projections observed in some patients who have never been exposed to those ideas. Also called objective psyche.

coma State of profound unconsciousness from which a person cannot be roused, with minimal or no detectable responsiveness to stimuli; seen in injury or disease of the brain, in such systemic conditions as diabetic ketoacidosis and uremia, and in intoxications with alcohol and other drugs. Coma may also be seen in severe catatonic states and in hysteria.

coma vigil Coma in which the eyes remain open; typically seen in acute organic mental syndromes associated with systemic infection.

combat fatigue Disabling physical and mental reaction to the stresses of military battle; now called posttraumatic stress disorder.

combined therapy Psychotherapy in which the patient is in both individual and group treatment with the same or two different therapists. In marriage therapy, it is the combination of married couples group therapy with either individual sessions with one spouse or conjoint sessions with the marital pair. *See also* co-therapy.

command automatism Condition closely associated with catalepsy in which suggestions are followed automatically.

command hallucination False perception of orders that a person may feel obliged to obey or unable to resist.

command negativism *See* negativism.

commitment Legal process by which a person is confined to a mental hospital, usually against his or her will.

common migraine *See* migraine.

communication disorder Term referring to a heterogeneous class of disorders resulting from impairment in the ability to receive, send, process, and comprehend concepts or verbal, nonverbal, and written symbolic systems. Communication disorders consist of language disorders, speech sound disorders, and social (pragmatic) communication disorder.

community *See* therapeutic community.

community mental health center Community or neighborhood mental health facility or a group of affiliated agencies that serve as a locus for the delivery of various services of community psychiatry. *See also* community psychiatry.

community psychiatry Psychiatry focusing on the detection, prevention, and early treatment of mental disorders and social deviance as they develop in the community, rather than as they are perceived and encountered at large centralized psychiatric facilities; particular emphasis is placed on the environmental factors that contribute to mental illness.

compensation Conscious or, usually, unconscious defense mechanism by which a person tries to make up (compensate) for an imagined or real physical or psychological deficiency.

compensation neurosis Neurosis induced by the desire for monetary compensation. It is believed that some persons, after sustaining an injury, develop neurotic symptoms in the hope of monetary gain.

competence Ability to perform or accomplish an action or task that another person of similar background and training, or any human being, could reasonably be expected to perform; mental capacity.

competency to stand trial Ability to be tried in a court of law. A person is competent to stand

trial when, at the time of the trial, he or she (1) understands the nature of the charge and the potential consequences of conviction and (2) is able to assist his or her attorney in his or her defense.

competition Literally, the struggle for the possession or use of limited goods, concrete or abstract. In psychiatry, a rivalry that may exist between persons in which gratification for one person largely precludes gratification for the other.

complementarity of interaction Concept of bipersonal and multi-personal psychology in which behavior is viewed as a response to stimulation, and interaction replaces the concept of reaction; each person in an interactive situation plays both a provocative role and a responsive role.

complementary role *See* role.

complex Group of interrelated ideas, mainly unconscious, that have a common emotional tone and strongly influence the person's attitudes and behavior; introduced by Jung, who called it a feeling-toned idea.

complex partial seizure Epilepsy characterized by recurrent behavioral disturbances. Complex hallucinations or illusions, frequently gustatory or olfactory, often herald the onset of the seizure, which typically involves a state of impaired consciousness resembling a dream (dreamy state) during which paramnestic phenomena, such as déjà vu and jamais vu, are experienced and the patient exhibits repetitive, automatic, or semipurposeful behavior; in rare instances, violent behavior may be prominent. Electroencephalography reveals a localized seizure focus in the temporal lobe. Also called psychomotor epilepsy; temporal lobe epilepsy.

complicated grief/bereavement *See* persistent complex bereavement disorder.

compulsion Uncontrollable, repetitive, and unwanted urge to perform an act; serves as a defense against unacceptable ideas and desires, and failure to perform the act leads to overt anxiety. *See also* obsession.

compulsive personality Personality characterized by rigidity, over conscientiousness, extreme inhibition, inability to relax, and the performance of repetitive patterns of behavior.

computed tomography (CT) Radiological imaging technique for presenting anatomical features from a cross-sectional plane of the body, generated by computer synthesis of x-ray transmission data obtained from many different directions through the given plane; gray and white matter are indistinguishable on CT scans.

conation That part of a person's mental life concerned with cravings, strivings, motivation, drives, and wishes as expressed through behavior or motor activity.

concordance In studies of twins, the degree of similarity between the two members of each pair with respect to the presence or absence of a particular trait. *Compare* discordance.

concrete operations *See* Piaget, Jean.

concrete thinking Thinking characterized by actual things and events and immediate experience, rather than by abstractions; seen in young children, in those who have lost or never developed the ability to generalize (as in certain organic mental disorder), and in schizophrenic persons. *Compare* abstract thinking.

concussion Immediate and transient impairment of brain function resulting from a head injury; usually, an alteration of consciousness is involved.

condensation Mental process in which one symbol stands for a number of components.

conditioned reinforcer *See* reinforcement.

conditioned response *See* conditioning.

conditioned stimulus *See* conditioning.

conditioning Procedure resulting in the acquisition or learning of more or less permanent changes in behavior. There are two main types of conditioning: **1.** Classical conditioning, in which a neutral stimulus is repeatedly paired with a stimulus (unconditioned stimulus) that naturally elicits a response (unconditioned response) until the neutral stimulus (conditioned stimulus) comes to elicit that response (now the conditioned response) by itself. Also called Pavlovian conditioning and respondent conditioning. **2.** Operant conditioning, a procedure developed by B. F. Skinner in which a spontaneously emitted behavior (operant behavior) is either rewarded (reinforced) or punished and, as a result, then occurs with a frequency that is either increased (in the case of reinforcement) or decreased (in the case of punishment). Also called instrumental conditioning. *See also* behavior therapy.

conditioning therapy *See* behavior therapy.

conditions not attributable to a mental disorder DSM category given to a person who may or may not have a mental disorder but who has a problem which may require intervention by a psychiatrist. Included are marital problem, occupational problem, phase of life or other life circumstance problem, malingering, complicated bereavement, antisocial behavior, academic problem, other interpersonal problem, other specified family circumstance, and noncompliance with medical treatment.

conduct disorder A diagnosis used in child psychiatry to refer to children who show antisocial behavior. Such behavior may occur as part of a group (group type), individually (solitary aggressive type), or both.

confabulation Unconscious filling of gaps in memory by imagining experiences or events that have no basis in fact, commonly seen in organic amnestic syndromes; should be differentiated from lying. *See also* paramnesia.

confidence interval Range of values with a specified probability of including the population mean with a specified level of confidence; typically associated with a certain probability level, for example, the 95 percent confidence interval has a 95 percent chance of including the population mean.

confidentiality Ethical principle by which the physician is bound to hold secret all information given him or her by the patient. Certain states do not legally recognize confidentiality and can require the physician to divulge such information if needed in a legal proceeding. *See also* privilege.

conflict Mental struggle arising from the clash of opposing impulses, drives, or demands. *See also* extrapsychic conflict; intrapsychic conflict.

conflict-free area Part of one's personality or ego that is well integrated and does not cause any conflicts, symptoms, or displeasures.

confrontation Act of letting a person know where one stands in relationship to that person, what one is experiencing, and how one is perceived. Used in a spirit of deep involvement, this technique is a powerful tool for changing relationships; used as an attempt to destroy another person, it can be harmful. In group and individual therapies, the value of confrontation is likely to be determined by the therapist.

confusion Disturbance of consciousness manifested by a disordered orientation in relation to time, place, or person.

congenital Referring to conditions present at birth, including hereditary conditions and those resulting from prenatal development or the process of birth itself.

conjoint therapy Type of marriage therapy in which a therapist sees the partners together in joint sessions. Also called triadic or triangular therapy (since two patients and one therapist work together). *See also* quadrangular therapy.

conjugal paranoia Paranoia characterized by delusions of jealousy; with no evidence or little evidence, a spouse or lover becomes convinced that his or her partner is being unfaithful. Also called Othello syndrome.

conscience The part of the self that judges one's own values and performance; often used synonymously with superego.

conscious 1. One of three divisions of Freud's topographic theory of the mind (the others being the preconscious and the unconscious), referring to that portion of mental functioning within the realm of awareness at all times. **2.** More generally, the state having present knowledge of oneself, one's acts, and one's surroundings and thus a functioning sensorium. **3.** Denoting a state of awareness, with response to external stimuli.

consciousness State of awareness, with response to external stimuli. *See also* sensorium.

consensual validation Continuous comparison of the thoughts and feelings of group members toward one another that tends to modify and correct interpersonal distortions; introduced by Harry Stack Sullivan to refer to the dyadic therapeutic process between doctor and patient.

conservation According to Piaget, the crucial difference between preoperational and concrete operational thought, in which a child discovers which values remain invariant in the course of any kind of change or transformation.

conservator Legal guardian designated to take over and protect the interests of an individual.

constipation Inability to defecate or difficulty in defecating.

constitution One's intrinsic psychological or physical endowment. Broadly, it includes an aggregate of characteristics that have developed from the interaction of hereditary and environmental influences; more narrowly, it indicates characteristics that are purely hereditary or genetically determined.

constitutional types Categories of a person's traits based on physical and psychological characteristics (e.g., ectomorphic, endomorphic, mesomorphic).

constricted affect Reduction in intensity of feeling tone less sever; than that of blunted affect but clearly reduced.

constructional apraxia Inability to copy a drawing, such as a cube, clock, or pentagon, as a result of a brain lesion.

consultation-liaison psychiatry Clinical psychiatry carried out in the medical or surgical wards of a general hospital; the psychiatrist collaborates closely with nonpsychiatric physicians in the total care of the patient.

contagion Force that operates in large groups or masses. When the level of psychological functioning has been lowered, some sudden upsurge of anxiety can spread through a group, accelerated by high degree of suggestibility, gradually mounting to panic, and the whole group

may be simultaneously affected by a primitive emotional upheaval.

content thought disorder Disturbance in thinking in which a person exhibits delusions that may be multiple, fragmented, and bizarre.

continuous reinforcement A schedule of reinforcement in which reward is administered every time a response is emitted.

contract Explicit, bilateral commitment to a well-defined course of action; in group or individual therapy, the therapist–patient contract is to attain the treatment goal.

control group In an experimental design, the group in which a condition or factor being tested is deliberately omitted; for example, in a study measuring the effects of a new drug, the control group may be given a placebo instead of the drug. *Compare* experimental group.

controlled substance Drug that has the potential to be abused and is classified by the Drug Enforcement Administration (DEA) into one of five levels of control: I, high abuse potential and no accepted medical use (e.g., heroin); II, high abuse potential with medical use (e.g., amphetamine); III, moderate abuse potential (e.g., codeine); IV, low abuse potential, that is, preparations containing limited amounts of narcotic and nonnarcotic ingredients (e.g., cough medicines).

controlling Excessive attempt to manage or regulate events or objects in the environment to minimize anxiety and to resolve inner conflicts.

conversational catharsis Release of repressed or suppressed thoughts and feelings in group and individual psychotherapy as a result of verbal interchange. *See also* abreaction; catharsis.

conversion Unconscious defense mechanism by which the anxiety that stems from an intrapsychic conflict is converted and expressed in a symbolic somatic symptom; the hallmark of conversion disorder, but also seen in a variety of mental disorders.

conversion disorder Alternate term for functional neurological symptom disorder.

conversion phenomena The development of symbolic physical symptoms and distortions involving the voluntary muscles or special sense organs; not under voluntary control and not explained by any physical disorder. Most common in conversion disorder, but also seen in a variety of mental disorders.

convulsion An involuntary, violent muscular contraction or spasm. *See also* clonic convulsion; tonic convulsion.

convulsive disorder *See* epilepsy.

convulsive therapy Psychiatric treatment involving administration of chemical substances or electric current to induce loss of consciousness or a convulsive reaction, to alter favorably the course of a mental disorder; for example, electroconvulsive therapy, insulin coma therapy.

cooperative therapy CO-THERAPY.

coping mechanism Unconscious or conscious way of dealing with stress.

coprolalia Involuntary use of vulgar or obscene language; observed in some cases of schizophrenia and in Tourette syndrome.

coprophagia Eating of filth or feces.

coprophilia Excessive interest in filth or feces. Sexual pleasure associated with feces is a perversion associated with fixations at the anal stage of psychosexual development.

coprophobia Abnormal fear of excreta.

copy-cat suicides Suicides that occur in persons who are geographically related; usually

occurs among adolescents within a limited time period. *See also* cluster suicides.

correctional psychiatry Practice of psychiatry on persons incarcerated in jail or prison; presents special problems in relation to confidentiality and countertransference.

corrective emotional experience Re-exposure under favorable circumstances to an emotional situation that the patient could not handle in the past; as advocated by Franz Alexander, the therapist temporarily assumes a particular role to generate the experience and facilitate reality testing.

correlation Statistical index of the relationship among variables. The most common correlation coefficient is the Pearson Product Moment Correlation, an index of bivariate association that varies between >1.0 and 1.0. *See also* coefficient of correlation.

Cotard syndrome *See* nihilistic delusional disorder.

co-therapy Psychotherapy in which more than one therapist treat the individual patient or the group. Also called cooperative, dual, multiple, or three-cornered therapy. *See also* combined therapy.

counseling Form of guidance in which a trained person offers advice or education about specific problems in living, for example, marriage, child rearing, mourning.

counterphobia State in which a phobic person shows a preference for the very situation that is feared.

countertransference Conscious or unconscious emotional response of the therapist to the patient; determined by the therapist's inner needs, rather than by the patient's needs, and may reinforce the patient's earlier traumatic history if not checked by the therapist. *See also* transference.

couples therapy *See* family therapy; marriage therapy.

covert sensitization Method of reducing the frequency of behavior by associating it with the imagination of an unpleasant consequence.

CPA Canadian Psychiatric Association.

crack Freebase form of cocaine prepared by heating a mixture of cocaine hydrochloride, water, and baking soda, which is then dried into a solid and broken into small pieces (rocks); it is inexpensive, has a rapid onset of action, and is highly addictive.

creativity Ability to produce something new. Silvano Arieti described creativity as the tertiary process, a balanced combination of primary and secondary processes, in which materials from the id are used in the service of the ego.

cretinism Mental and physical retardation resulting from severe thyroid deficiency in infancy and childhood.

Creutzfeldt–Jakob disease (CJD) A rapidly progressive, invariably fatal disorder that occurs mainly in middle-aged or older adults. It initially manifests with fatigue, flu-like symptoms, and cognitive impairment; as it progresses, focal neurological findings such as aphasia and apraxia occur. Psychiatric symptoms include anxiety, euphoria, depression, delusions, hallucinations, or marked personality changes. Progresses over months and leads to dementia, akinetic mutism, coma, and death.

Cri-du-chat High-pitched catlike cry characteristic of the syndrome of mental and physical retardation caused by partial deletion of chromosome 5. Also called cat's cry syndrome.

criminal responsibility Legal term meaning the ability to formulate a criminal intent at the

time of an alleged crime. A person cannot be convicted of a crime if it can be proved that he or she lacked criminal responsibility by reason of insanity.

crisis In psychiatry, a distressing sense of psychological disequilibrium that has an acute onset and that causes impaired functioning in a variety of interpersonal and vocational areas.

crisis intervention Brief therapeutic approach used in emergency rooms of general or psychiatric hospitals that is ameliorative, rather than curative, of acute psychiatric emergencies; treatment factors often focus on environmental modification, although interpersonal and intrapsychic factors are also considered. *See also* hot line.

critical period Specific period of early development during which an organism acquires certain patterns of behavior as a result of internal or external clues. *See also* imprinting.

critical ratio In a statistical study involving 30 or more subjects, a ratio constructed to determine whether differences found between two items are larger than could be expected from chance alone.

cross-cultural psychiatry Comparative study of mental illness and health among various societies around the world.

cryptographia A private written language.

cryptolalia A private spoken language.

CT scan Computed tomography.

cult System of beliefs and rituals, especially involving devoted attachment to or extravagant admiration of a person or principle, based on dogma or religious teachings; devoted adherents to such a system are called cultists.

cultural anthropology *See* anthropology.

cultural deprivation Restricted participation in the culture of the larger society.

cultural psychiatry Branch of psychiatry concerned with mental illness in relation to the cultural environment; symptoms regarded as psychopathological in one cultural setting may be acceptable and normal in another.

culture shock Psychological response of a person who is suddenly thrust into an alien culture or has divided loyalties to two different cultures.

culture specific syndrome Disorder of behavior that occurs within or is limited to a specific culture; for example, amok, koro, latah.

cunnilingus Use of the mouth or tongue to stimulate the female genitalia.

Cushing syndrome Hyperadrenocorticism characterized by muscle wasting, obesity, osteoporosis, atrophy of the skin, and hypertension; emotional lability is common and frank psychoses are occasionally observed.

custody Term used to describe a parent's legal rights over a child; in cases of disputed custody, the child is awarded to the party who will best serve the interests of the child.

CVA Cerebral vascular accident.

cybernetics Norbert Wiener's term for the comparative study of computers and the human nervous system.

cybersex addiction A paraphilia characterized by dependence on the internet for sexual gratification (e.g., internet pornography, sex-chat rooms).

cyclazocine Narcotic antagonist that blocks the effects of morphine and heroin.

cycloplegia Paralysis of the muscles of accommodation in the eye; observed at times as an autonomic side effect (anticholinergic effect) of antipsychotic or antidepressant medication.

cyclothymia Mood disorder characterized by a nonpsychotic disturbance of at least 2 years length with numerous periods of both depressive and manic syndromes not severe enough to be classified as depressive or manic episodes.

cyclothymic disorder Mood disorder characterized by a nonpsychotic disturbance of at least 2 years' length with numerous periods of both depressive and hypomanic syndromes that are not severe enough to be classified as depressive or hypomanic episodes.

cynophobia Abnormal fear of dogs.

D

Da Costa syndrome Chest pain, palpitations, and lightheadedness occurring as a somatic manifestation of pathological anxiety. Also called neurocirculatory asthenia; soldier's heart.

day hospital Form of partial hospitalization in which a patient spends the day in the hospital and returns home at night. *Compare* night hospital; weekend hospital.

day residue In Sigmund Freud's dream theory, the content of the dream imagery that is derived from events of the previous day.

DEA Drug Enforcement Administration.

death instinct AGGRESSIVE DRIVE.

decathexis ACATHEXIS.

decision tree algorithm.

de Clérambault syndrome *See* erotomania.

decoding Technical term for the process of understanding spoken language.

decompensation Deterioration of psychic functioning due to a breakdown of defense mechanisms.

decomposition In psychiatry, the division of a person into separate personalities or identities, as seen in the paranoid schizophrenic person who splits the persecutor into separate entities.

deep brain stimulation A wired electrode with multiple stimulating sites that is surgically implanted into the brain chronically administers electrical stimulation. Use to treat tremor in Parkinson disease.

defense mechanism Unconscious process acting to relieve conflict and anxiety arising from one's impulses and drives.

defensive emotion Strong feeling that serves as a screen for a less acceptable feeling, one that would cause a person to experience anxiety if it appeared. For example, expressing the emotion of anger is often more acceptable to a patient than expressing the fear that the anger covers up; in that instance, anger is defensive.

deficit In psychiatry, the absence of a function or capability that would normally be present, for example, judgment.

deinstitutionalization Discharge from the hospital to community, particularly of those chronically mentally ill patients who would otherwise be hospitalized for long periods of time.

déjà entendu Illusion that what one is hearing one has heard previously. *See also* paramnesia.

déjà pensè Condition in which a thought never entertained before is incorrectly regarded as a repetition of a previous thought.

déjà vu Illusion of visual recognition in which a new situation is incorrectly regarded as a repetition of a previous experience. *See also* paramnesia.

delayed ejaculation DSM-5 term for male orgasmic disorder.

delayed sleep phase circadian rhythm disorder A circadian rhythm sleep–wake disorder characterized by sleep–wake times that are later than conventional times; individuals with this disorder are referred to as "night owls."

delirium Acute reversible organic mental disorder characterized by confusion and some impairment of consciousness; generally associated with emotional lability, hallucinations or illusions, and inappropriate, impulsive, irrational, or violent behavior.

delirium tremens (DTs) Acute and sometimes fatal reaction to withdrawal from alcohol, usually occurring 72 to 96 hours after the cessation of heavy drinking; distinctive characteristics are marked autonomic hyperactivity (tachycardia, fever, hyperhidrosis, dilated pupils), usually accompanied by tremulousness, hallucinations, illusions, and delusions. Called alcohol withdrawal delirium in DSM. *See also* formication.

delta sleep SLOW WAVE SLEEP.

delta waves As measured by an electroencephalogram (EEG), brain waves that are less than 4 Hz, indicative of deep sleep.

delusion False belief that is firmly held, despite objective and obvious contradictory proof or evidence and despite the fact that other members of the culture do not share the belief. Types of delusion include: **bizarre delusion,** one that is patently absurd or fantastic; **delusion of control,** that a person's thoughts, feelings, or actions are not one's own but are being imposed by some external force; **delusion of grandeur (grandiose delusion),** exaggerated concept of one's importance, power, knowledge, or identity; **delusion of jealousy (delusion of infidelity),** that one's lover is unfaithful; **delusion of persecution,** that one is being attacked, harassed, cheated, or conspired against; **delusion of poverty,** that one is or will be without material possessions; **delusion of reference,** that events, objects, or the behavior of others have a particular and unusual meaning specifically for oneself; **encapsulated delusion,** one without significant effect on behavior; fragmentary delusion, poorly elaborated delusion, often one of many with no apparent interconnection; **nihilistic delusion, delusion of negation** (Cotard syndrome), that the world and everything related to it have ceased to exist; **paranoid delusion,** combined delusion of grandeur and persecution; **religious delusion,** one involving the Deity or theological themes; **sexual delusion,** one centering on sexual identity, appearances, practices, or ideas; **somatic delusion,** pertaining to the functioning of one's body; **systematized delusion,** a group of elaborate delusions related to a single event or theme.

delusional depression Depressive disorder characterized by false, fixed ideas, that is, delusions; listed in DSM as major depression, severe, with psychotic features.

delusional (paranoid) disorder Psychotic disorder in which there are persistent delusions (e.g., erotomanic, grandiose, jealous, persecutory, somatic, mixed, and unspecified); symptoms of a major mood disorder are absent, and the delusions lack the bizarre quality often seen in schizophrenia.

dementia Mental disorder characterized by general impairment in intellectual functioning, frequently characterized by failing memory, difficulty with calculations, distractibility,

alterations in mood and affect, impairment in judgment and abstraction, reduced facility with language, and disturbance of orientation. Although generally irreversible because of underlying progressive, degenerative brain disease, dementia may be reversible if the cause can be treated. Called major neurocognitive disorder in DSM-5. *Compare* amentia.

dementia associated with alcoholism DSM term for alcoholic deterioration.

dementia of the Alzheimer type DSM term for a clinical disorder in an elderly patient characterized by a severe loss of intellectual function, for which no other cause is found, and that is progressive and insidious and marked by disorientation as to time, place, and person; senile onset is after age 65 (senile dementia); presenile onset is before age 65 (presenile dementia). Also called Alzheimer disease. *See also* Pick disease.

dementia praecox (precox) Obsolete term for schizophrenia.

demography Statistical study of populations, including births, marriages, mortality, health, geographic distribution, and population shifts.

demophobia Abnormal fear of crowds.

dendrite Projection from the nerve cell body, usually profusely branched and studded with small spikes (dendritic spines) that are the sites of synaptic connection; most dendrites receive neurotransmissions from an axon.

denial Defense mechanism in which the existence of unpleasant realities is disavowed; refers to a keeping out of conscious awareness any aspects of either internal or external reality that, if acknowledged, would produce anxiety.

dependence on therapy Pathological need of a patient for therapy, created out of the belief that one cannot survive without it.

dependency State of reliance on another, as for security, love, protection, or mothering.

dependent personality disorder Personality disorder characterized by lack of self-confidence, a tendency to have others assume responsibility for significant areas of one's life, and a subordination of one's own needs and wishes to those of the others on whom one is dependent; solitude is extremely discomforting.

dependent variable In research methodology, those qualities that measure the influence of the independent variable or the outcome of the experiment; for example, measurement of a person's specific physiological reactions to a drug. Also called outcome variable.

depersonalization Called depersonalization disorder in DSM-5, a disordered sensation of unreality concerning oneself, parts of oneself, or one's environment that occurs under extreme stress or fatigue. Seen also in schizophrenia and schizotypal disorder.

depersonalization disorder [Omit entry.]

depression Mental state characterized by feelings of sadness, loneliness, despair, low self-esteem, and self-reproach; accompanying signs include psychomotor retardation or at times agitation, withdrawal from interpersonal contact, and vegetative symptoms such as insomnia and anorexia. The term refers either to a mood that is so characterized or to a mood disorder.

depressive neurosis DYSTHYMIA.

depth psychology Psychology that focuses on unconscious processes, rather than conscious processes.

derailment Gradual or sudden deviation in train of thought without blocking; sometimes used synonymously with loosening of association.

derealization Called derealization disorder in DSM-5, disordered sensation of a changed reality or that one's surroundings have altered; also seen in schizophrenia.

dereism Mental activity that follows a totally subjective and idiosyncratic system of logic and fails to take the facts of reality or experience into consideration. *See also* autistic thinking.

dereistic thinking Mental activity not concordant with logic or experience.

descriptive psychiatry System of psychiatry focusing primarily on the study of observable symptoms and behavioral phenomena, rather than underlying psychodynamic processes; Emil Kraepelin's systematic description of mental illness was an early example. *See also* dynamic psychiatry.

descriptive statistics Methods used to summarize, organize, and describe observations; for example, the mean, standard deviation, and variance.

desensitization *See* systematic desensitization.

designer drug Any version of a regulated drug whose chemical structure has been modified so that it falls outside the restrictions placed on the regulated drug it imitates; usually a drug of abuse used to alter states of consciousness or to produce pleasurable feelings.

detachment A state of being in which the person is separated or divorced from emotional involvement with others, characterized by distant interpersonal relationships.

determinism Concept that nothing occurs by chance alone; instead, things result from specific causes or forces. This school of thought denies the notion of free will.

detoxification Any treatment aimed at the restoration of normal physiological function after a disturbance by the use of some drug or chemical substance. *See also* cold turkey.

detumescence Subsiding of a penile erection; also used to describe the subsiding of genital engorgement in both sexes.

devaluation Defense mechanism in which a person attributes excessively negative qualities to self or others. Seen in depression and paranoid personality disorder.

developmental arithmetic disorder Academic skills disorder characterized by marked impairment in the development of arithmetic skills that interferes with academic achievement or other activities of daily living. Also called acalculalia; Gerstmann–Straussler–Scheinker disease, congenital arithmetic disability; dyscalculia; arithmetic disorder.

developmental articulation disorder *See* speech sound disorder.

developmental coordination disorder Motor skills disorder of childhood with significant impairment in the development of fine or gross motor coordination that interferes with academic achievement or with activities of daily life.

developmental disorders Disorders originating in infancy and childhood that continue in adulthood and which interfere with the person's ability to function normally. In DSM, called neurodevelopmental disorders and include intellectual developmental disorder (formally known as mental retardation) and autistic spectrum disorders.

developmental expressive language disorder Language and speech disorder in which there is marked impairment in the development of expressive language; interferes with academic achievement or with activity of daily living that require the expression of verbal (or sign) language.

developmental expressive writing disorder Academic skills disorder characterized by marked impairment in the development of expressive writing skills; significantly interferes

with academic achievement or with activities of daily living that require expressive writing skills.

developmental reading disorder Academic skills disorder characterized by omissions, distortions, and substitutions of words and by slow, halting reading; there is marked impairment in the development of words recognition skills and reading comprehension. Also called specific reading disability and developmental word blindness.

developmental receptive language disorder Language and speech disorder in which there is marked impairment in the development of language comprehension which significantly interferes with academic achievement or with activities of daily living that require comprehension of verbal (or sign) language.

developmental word blindness DEVELOPMENTAL READING DISORDER.

deviation Departure from the average or norm.

Dexamethasone Suppression Test (DST) Test used to confirm a diagnostic impression of major depression. The patient is given 1 mg of dexamethasone (a long-acting synthetic glucocorticoid) by mouth at 11 PM, and plasma cortisol is measured at 8 AM, 4 PM, and 11 PM. Plasma cortisol above 5 μg/dL (nonsuppression) is considered abnormal, that is, positive; suppression of cortisol indicates that the hypothalamic–adrenal–pituitary axis is functioning properly.

***Diagnostic and Statistical Manual of Mental Disorders* (DSM)** Handbook for the classification of mental illnesses that is primarily descriptive and does not describe the theories or causes of disorders; first issued by the American Psychiatric Association in 1952 (DSM-I). The fourth edition (DSM-IV) was issued in 1994 and the text revision of fourth edition (DSM-IV-TR) was published in 2000. The DSM is now entering its fifth revision and future revisions are expected. DSM closely adheres to the 11th revision of the World Health Organization's *International Classification of Diseases and Related Health Problems* (ICD-11).

dialectical behavior therapy (DBT) A form of therapy used for patients with borderline personality disorder and parasuicidal behavior that is eclectic using advice-giving, cognitive training, story-telling, and confrontation among many other techniques.

diathesis-stress paradigm In psychopathology, a view that abnormal behavior is a function of a series of stresses in predisposed individuals (due to hereditary, temperamental, and sociocultural factors).

differential diagnosis Determination of which of two or more diseases with similar symptoms is the one that afflicts the patient.

differential reinforcement Procedure in which desirable behavior is reinforced and problematic or less desirable behavior is extinguished or punished.

differentiation First of Margaret Mahler's four subphases of the separation-individuation process, occurring from ages 5 to 10 months, when the child is able to distinguish between self and other objects.

2,5-dimethoxy-4-methylamphetamine (DOM) Synthetic (derivative) drug with hallucinogenic properties. Also colloquially called STP.

dimethyltryptamine (DMT) Tryptamine congener that acts as a psychedelic capable of producing hallucinations, other disorders of perception, and altered moods and states of consciousness.

diminished libido Decreased sexual interest and drive.

dipsomania An obsolete term that refers to the compulsion to drink alcoholic beverages.

disability Defined by the federal government as: "Inability to engage in any substantial gainful activity by reason of any medically determinable physical or mental impairment which can be expected to last or has lasted for a continuous period of not less than 12 full months."

disconnection syndrome Group of neurological disorders produced after interruption of connection fibers from the cerebrum to other centers of the nervous system.

discontinuation syndrome Signs and symptoms that occur when a person stops taking a drug either suddenly or gradually when the drug previously has caused physical or psychological dependence to occur. Also called abstinence syndrome; withdrawal syndrome.

discordance In studies of twins, the degree of dissimilarity between the two members of each pair with respect to the presence or absence of a particular trait. *Compare* concordance.

discriminant analysis Multivariate method for finding the relationship between a single discrete outcome and a linear combination of two or more predictors.

discrimination Process of distinguishing between different but often similar stimuli and reacting appropriately; more specifically, the occurrence of operant behavior in environments in which the behavior has been reinforced but not in those in which it has not been reinforced. *See also* discriminative stimulus.

discriminative stimulus Environmental event correlating with the operation of a certain reinforcement matrix, such that behavior that has been reinforced in the presence of the stimulus but not in its absence tends to occur only when the stimulus is present.

disinhibition 1. Removal of an inhibitory effect, as in the reduction of the inhibitory function of the cerebral cortex by alcohol. **2.** In psychiatry, a greater freedom to act in accordance with inner drives or feelings and with less regard for restraints dictated by cultural norms or one's superego.

disinhibited social engagement disorder A disorder of markedly disturbed social relatedness in which a child approaches and interacts with unfamiliar adults in an overly familiar manner; presumed to result from grossly pathological care-giving.

disintegration Psychic disorganization in which a person is unable to recognize reality and may develop delusions, hallucinations, or overwhelming anxiety and panic.

disorder of written expression A learning disorder characterized by writing skills that are significantly below the expected level for a child's age and intellectual capacity.

disorders of excessive somnolence Disorders characterized by excessive amounts of sleep or excessive daytime sleepiness. *See also* somnolence.

disorders of initiating and maintaining sleep Disorders characterized by inability to fall asleep or repeated awakening after failing asleep. *See also* insomnia disorder; dyssomnias.

disorganized schizophrenia Type of schizophrenia characterized by markedly disinhibited and aimless, nonconstructive behavior; no longer included in DSM-5.

disorientation Confusion; impairment of awareness of time, place, and person (the position of the self in relation to other persons); characteristic of cognitive disorders.

displacement Unconscious defense mechanism by which the emotional component of an unacceptable idea or object is transferred to a more acceptable one. Seen in phobias.

disposition Sum total of a person's inclinations as determined by his or her mood.

disruptive behavior disorders Disorders characterized by inattention, overagressiveness, delinquency, destructiveness, hostility, feelings of rejection, negativism, or impulsiveness. In DSM, they include attention-deficit hyperactivity disorder, conduct disorder, and oppositional defiant disorder.

disruptive mood dysregulation disorder A depressive disorder in individuals between 6 and 18 years of age that is characterized by chronic irritability, severe temper tantrums, and angry mood; a DSM-5 diagnosis.

dissociation Unconscious defense mechanism involving the segregation of any group of mental or behavioral processes from the rest of the person's psychic activity; may entail the separation of an idea from its accompanying emotional tone, as seen in dissociative and conversion disorders.

dissociation syndrome Functional inability of the brain to process information from one area to another; first described by neurologist Norman Geschwind. *See also* disconnection syndrome.

dissociative amnesia Disorder characterized by memory loss of important autobiographical information that is usually stressful or traumatic in nature.

dissociative disorders Mental disorders characterized by a sudden temporary alteration in consciousness, identity, or motor behavior. In DSM, they include dissociative amnesia, dissociative fugue, dissociative identity disorder, and depersonalization disorder. *See also* hysteria.

dissociative identity disorder Dissociative reaction in which a person has two or more distinct personalities, most of them knowing nothing of the others; a DSM-5 diagnosis. Previously called multiple personality disorder.

dissociative trance disorder Disorder characterized by temporary alteration in the state of consciousness or personal identity without replacement by an alternate identity; may occur during the course of a psychotic disorder or result from substance use. Not a DSM diagnosis.

distortion 1. Misrepresentation of reality. **2.** In psychoanalysis, the process of modifying unacceptable unconscious mental elements so that they are allowed to enter consciousness in a more acceptable but disguised form.

distractibility Inability to focus one's attention; the patient does not respond to the task at hand but attends to irrelevant phenomena in the environment.

distractible speech Pattern of speech in which a person changes subjects in response to nearby stimuli.

distributive analysis and synthesis Therapeutic application of Adolf Meyer's psychobiological school of psychiatry; systematic investigations of patients' entire past experiences yield constructive syntheses of their assets and liabilities and lead to an effort to enable them to adapt to their environment.

disulfiram Drug used in the treatment of alcoholics. By altering the metabolism of alcohol, it produces unpleasant physical sensations that discourage the patient from further consumption of alcohol (trade name, Antabuse).

Dix, Dorothea Lynde (1802–1887) Former schoolteacher who devoted her life to improving institutional care for the mentally ill.

dizygotic twins Twins (fraternal) who develop from two separately fertilized ova; in contrast,

monozygotic (identical) twins develop from a single fertilized ovum.

DMT Dimethyltryptamine.

Doctor–patient relationship Human interchange that exists between the person who is sick and the person who is selected to heal because of training and experience.

DOES Disorders of excessive somnolence.

DOM 2,5-Dimethoxy-4-methylamphetamine.

dominance 1. Disposition to play a commanding, controlling, or prevailing role over others. **2.** In neurology, the control of language function by one side of the cerebral cortex. **3.** In genetics, the ability of a gene (dominant gene) to be expressed in the phenotype of the bearer to the exclusion of another gene (recessive gene) with which it is paired.

dopamine Biogenic amine (catecholamine) that inhibits prolactin; functional integrity of the dopaminergic system is impaired in depression and in Parkinson disease.

dopamine hypothesis Theory that there may be an excess of dopamine in the brains of persons with schizophrenia.

dopamine-receptor agonists Drugs that activate dopamine receptors in the absence of dopamine. Commonly used in the treatment of Parkinson disease.

dopamine-receptor antagonists (DRAs) Drugs that block the receptors for dopamine in the brain. Also known as first-generation or typical antipsychotics. *Compare* serotonin-dopamine antagonists.

double bind As formulated by Gregory Bateson, a situation involving the simultaneous communication of conflicting messages from one person to another, in which the response to either message is met with disapproval or rejection; believed to be a characteristic mode of interaction in some families with a schizophrenic member.

double-blind study Study in which, for example, one or more drugs and a placebo are compared in such a way that neither the patient nor the persons directly or indirectly involved in the study know which is being given to the patient; the drugs being investigated and the placebo are coded for identification.

double depression A mood disorder that occurs when the criteria for major depressive disorder and dysthymia are met simultaneously.

double personality *See* multiple personality disorder.

Down syndrome Syndrome of mental retardation associated with a variety of abnormalities caused by replication of chromosome 21 three times. Also called trisomy 21 syndrome. Formerly called mongolism.

DRA Dopamine-receptor antagonist.

Draw-A-Person (DAP) Test Easily administered test used with both children and adults, usually with the instructions, "I'd like you to draw a picture of a person; draw the best person you can." A general assumption is that the drawing of a person represents the expression of the self or of the body in the environment.

drawing test Any of a variety of psychological tests in which the subject is asked to draw certain familiar objects, such as people, trees, and houses; attitudes and feelings are often revealed in the way the subject depicts those objects.

dread Massive or pervasive anxiety, usually related to a specific danger.

dream Mental activity during sleep in which thoughts, emotions, and images are experienced

as though real; associated with rapid eye movement (REM) sleep. Dreams were regarded by Freud as providing an outlet for the discharge in disguised form of unconscious and often unacceptable impulses and wishes.

dream anxiety disorder *See* nightmare disorder.

dreamy state Altered state of consciousness, likened to a dream situation, that develops suddenly and from which the patient usually recovers within a few minutes; accompanied by visual, auditory, and olfactory hallucinations and most commonly associated with temporal lobe lesions.

drive Hypothetical construct used to explain motivated behavior: a basic urge that produces a state of psychic tension that motivates the person into action to alleviate the tension; currently preferred over Freud's term "instinct." *See also* aggressive drive; sexual drive.

drowsiness State of impaired awareness associated with a desire or inclination to sleep.

drug abuse Use of any drug, usually by self-administration, in a manner that deviates from approved social or medical patterns.

drug dependence Habituation or addiction to the use of a drug or chemical substance, with or without physical dependence; physical dependence indicates an altered physiological state due to repeated administrations of a drug, the cessation of which results in the appearance of a withdrawal or abstinence syndrome characteristic for the particular drug which is called discontinuation syndrome.

Drug Enforcement Administration (DEA) Agency of the federal government that regulates distribution of controlled substances such as opioids, barbiturates, and other drugs of potential abuse.

drug holiday Limited period of time during which the administration of a psychoactive drug is discontinued in order to control dosage and side effects or to evaluate baseline behavior.

drug-induced parkinsonism PSEUDOPARKINSONISM.

drug interaction Effect that one drug has on the actions of one or more other drugs given at the same time; effects of either drug may be increased, diminished, or inactivated as a result of the interaction.

drug levels Blood serum or plasma level of psychotropic drugs. *See also* pharmacokinetics.

drug therapy Use of chemical substances in the treatment of a disease or disorder. Also called chemotherapy.

drug tolerance *See* tolerance.

DSM *Diagnostic and Statistical Manual of Mental Disorders.*

DST Dexamethasone Suppression Test.

DTs Delirium tremens.

dual-sex therapy Specific form of psychotherapy developed by William Masters and Virginia Johnson in which treatment is focused on a particular sexual disorder. The crux of the program is the round-table session, in which both a male therapist and a female therapist are present with the patient couple; special exercises are prescribed for the couple, the overall goal being to diminish fears of sexual performance and to facilitate communication in sexual and nonsexual areas.

dual therapy CO-THERAPY.

dummy Uncommonly used term for placebo; for example, dummy drug.

Durham rule Federal court ruling in 1954 holding that "an accused is not criminally responsible if his unlawful act was the product of mental disease or mental defect." It has since been replaced by the American Law Institute formulation of insanity, which states that "a person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity either to appreciate the wrongfulness of his conduct or to conform his conduct to the requirements of law." *See also* M'Naghten rule.

dyad Pair of persons in an interactional situation, such as husband and wife, mother and father, co-therapists, or patient and therapist.

dyadic session Psychotherapeutic session involving only two persons, the therapist and the patient.

dynamic psychiatry System of psychiatry concerned with the mental mechanisms and emotional processes that govern and motivate human behavior, in contradistinction to descriptive psychiatry, which focuses on observable behavioral phenomena.

dynamic reasoning Franz Alexander's term for the formation of all the clinical evidence gained from free-associative anamnesis into a psychological reconstruction of the patient's development.

dynamics *See* psychodynamics.

dysarthria Impairment of articulation, the motor activity of shaping phonated sounds into speech.

dyscalculia Difficulty in performing calculations.

dysdiadochokinesis *See* adibdochokinesis.

dysgeusia Impairment of the sense of taste.

dysgraphia Difficulty in writing.

dyskinesia Any disturbance of movement, as in extrapyramidal dyskinesia.

dyslalia Faulty articulation due to structural abnormalities of the articulatory organs or impaired hearing.

dyslexia A specific learning disability syndrome involving an impairment of the ability to read; unrelated to the person's intelligence. *Compare* alexia.

dysmetria Impaired ability to gauge distance relative to bodily movements.

dysmnesia Impaired memory.

dysmorphophobia BODY DYSMORPHIC DISORDER.

dyspareunia Physical pain in sexual intercourse, usually emotionally caused and more commonly experienced by women; may also result from cystitis, urethritis, or other medical conditions. *See also* Genito-pelvic pain/penetration disorder.

dysphagia Difficulty in swallowing.

dysphasia Difficulty in comprehending oral language (receptive dysphasia) or in trying to express verbal language (expressive dysphasia).

dysphonia Difficulty or pain in speaking.

dysphoria Feeling of unpleasantness or discomfort; a mood of general dissatisfaction, restlessness, depression, and anxiety.

dysphoric mood *See* dysphoria.

dysprosody Loss of normal speech melody (prosody).

dyssocial behavior Former term for antisocial behavior.

dyssomnias Disturbances in the amount, quality, or timing of sleep. *See* insomnia disorder, hypersomnia disorder, and circadian rhythm disorder. *Compare* parasomnias.

dysthymia Mood disorder characterized by chronic depressed mood for at least 2 years that is not severe enough to meet criteria for major depressive episode. Also called depressive neurosis. *See also* depression.

dystonia Extrapyramidal motor disturbance consisting of slow, sustained contractions of the axial or appendicular musculature; one movement often predominates, leading to relatively sustained postural deviations; acute dystonic reactions (facial grimacing, torticollis) are occasionally seen with the initiation of antipsychotic drug therapy.

E

early ejaculation DSM-5 term for premature ejaculation.

eating disorders Disorders characterized by a marked disturbance in eating behavior; they include anorexia nervosa, bulimia nervosa, pica, and rumination disorder of infancy.

echolalia Repetition of another person's words or phrases, as observed in certain cases of schizophrenia, particularly the catatonic types. The behavior is considered by some authors to be an attempt by the patient to maintain a continuity of thought processes.

echopraxia Repetition of another person's movements, as observed in some cases of catatonic schizophrenia.

ecology Study of the mutual relationships between living things and their environment; includes such studies as the differential incidence of mental disorders in various populations and the distribution of crime and delinquency within a specified geographical area.

ecopsychiatry Study of genetic data and hypotheses about the role of social factors in mental disease.

ecstasy 1. State of rapturous delight. **2.** Slang term for methylenedioxymethamphetamine (MDMA).

ECT Electroconvulsive therapy.

ectomorphic Thin, characterized by a linear body musculature; one of Sheldon's constitutional types. *Compare* endomorphic; mesomorphic.

educable Capable of achieving a fourth- or fifth-grade academic level; describes the mildly mentally retarded (LQ of 50–55 to 70). *Compare* trainable.

educational psychology Area of psychology concerned with the application of psychological theories to teaching.

EEG Electroencephalogram.

ego One of the three components of the psychic apparatus in the freudian structural framework, the others being the id and the superego. It occupies a position between the primal instincts and the demands of the outer world, therefore mediating between the person and external reality. In so doing, it performs the important functions of perceiving the needs of the self, both physical and psychological, the qualities and attitudes of the environment. It evaluates, coordinates, and integrates those perceptions so that internal demands can be adjusted to external requirements. It is also responsible for certain defensive functions to protect the person against the demands of the id and the superego. Although the ego has some conscious

components, many of its operations are automatic; adaptation to reality is perhaps its most important function.

ego-alien Denoting aspects of a person's personality that are viewed as repugnant, unacceptable, or inconsistent with the rest of the personality. Also called ego-dystonic. *Compare* ego-syntonic.

ego analysis Psychoanalytic study of the manner in which the ego operates, especially in relation to intrapsychic conflicts. *See also* ego psychology.

ego boundaries Concept introduced by Paul Federn that refers to the ability of the intact ego to differentiate the real from the unreal. Ego boundaries prevent repressed unconscious material from overwhelming the ego; when that happens, a person experiences depersonalization and derealization. Boundaries are said to be weakened in schizophrenia and dissociative states.

egocentric Self-centered; selfishly preoccupied with one's own needs; lacking interest in others.

ego-coping skill Adaptive method or capacity developed by a person to deal with or overcome a psychological or social problem.

ego defense Defense mechanism to protect the person against the demands of the id and the superego.

ego-dystonic EGO-ALIEN.

ego-dystonic homosexuality State of a person who has unwanted and distressful homosexual arousal and wishes to acquire or increase heterosexual arousal.

ego ideal Part of the personality representing a fusion of the ego and the superego; feelings of self-satisfaction are experienced when the person's thoughts, behavior, or impulses coincide with an internalized goal or ideal.

egomania Morbid self-preoccupation or self-centeredness. *See also* narcissism.

ego model Person after whom one patterns the ego.

ego psychology Study of the operations of ego in regard to control and regulation of instinctual drives, the relation to reality, and the development of object relations. *See also* ego analysis.

ego strength Degree to which the ego's functions, for example, perception, thinking, speaking, are maintained.

ego-syntonic Denoting aspects of a person's personality that are viewed as acceptable and consistent with his or her total personality. *Compare* ego-alien.

eidetic image Unusually vivid or exact mental image seen or imagined event; may include color or movement.

ejaculation Expulsion of semen and seminal fluid from the urethra, as in male orgasm.

ejaculatory incompetence INHIBITED MALE ORGASM.

elaboration Unconscious process of embellishing the symbolic content of a dream.

elation Affect consisting of feelings of joy, euphoria, triumph, intense self-satisfaction, or optimism.

elective mutism Childhood mental disorder in which a child who is able and willing to speak to selected persons persistently refuses to speak in other social or school situations.

Electra complex Female analog of the Oedipus complex in the male; infrequently used term

to describe unresolved developmental conflicts influencing a woman's relationships with men.

electroconvulsive therapy (ECT) Convulsive therapy, usually for depression, that involves the application of electric current to the brain for a fraction of a second through scalp electrodes to induce a convulsive reaction and unconsciousness.

electroencephalogram (EEG) Recording of the brain's electrical potentials obtained by scalp electrodes.

electromyogram (EMG) Recording of the electrical activity of muscles at specific sites.

electroshock therapy (EST) *See* electroconvulsive therapy.

electrostimulation In aversion therapy, a negative conditioning technique consisting of (usually painful) electric shocks.

elevated mood Air of confidence and enjoyment; a mood more cheerful than normal but not necessarily pathologic.

elimination disorders Inability as a result of physiological or psychological immaturity to maintain bowel control (functional encopresis) or bladder control (functional enuresis).

elopement Escape; absenting oneself from a mental hospital without permission.

emancipated minor Legal term for a minor who exercises general control over his or her life and hence has the legal rights of an adult.

EMG Electromyogram.

emotion Complex feeling state with psychic, somatic, and behavioral components; external manifestation of emotion is affect.

emotional deprivation Relative lack of environmental or interpersonal experiences during the early developmental years. *See also* sensory deprivation.

emotional disorder Lay term for a mental illness or mental disorder; in common usage it does not specifically imply a mood disorder or schizophrenia.

emotional disturbance Emotional disorder or illness. *See also* mental disorder.

emotional illness General term for a behavioral or psychological syndrome or pattern associated with distress, disability, or increased risk of suffering.

emotional insight Deeper level of understanding or awareness that is more likely to lead to positive changes in personality and behavior.

emotional lability Excessive emotional responsiveness characterized by unstable and rapidly changing emotions.

emotional support Encouragement, hope, and inspiration given to one person by another.

empathy Intellectual and emotional awareness and understanding of another person's state of mind, involving the projection of oneself into another person's frame of reference; an important ability in a successful therapist or a helpful group member. *Compare* sympathy.

empty nest syndrome Depression that can occur in men and women when their youngest child is about to leave or has left home.

encapsulated delusion Delusion that has no significant effect on behavior.

encephalitis Inflammation of the brain.

encephalopathy Any disease of the brain.

encopresis Repeated passage of feces in inappropriate places, such as clothing or bed; diagnosis not made until after age 4.

encounter group Form of sensitivity training, introduced and developed by Jacob Moreno in 1914, that emphasizes the experiencing of individual relationships within the group and that minimizes intellectual and didactic input; the group focuses on the present, rather than concerning itself with the past or outside problems of its members. *See also* sensitivity training group.

endemic Belonging, native, restricted, or peculiar to a particular people, country, locality, or region.

endocrine disorder Dysfunction of any of the endocrine glands, frequently with psychiatric manifestations.

endogamy Marriage within a specific group. *Compare* exogamy.

endogenous depression Type of depression that originates within the person and is not related to external events, with a cause that is biological or somatic. Also called nonreactive depression.

endomorphic Obese, with prominence of the abdomen; one of Sheldon's constitutional types. *Compare* ectomorphic; mesomorphic.

endorphin Any of several endogenous opioid peptides (α-, β-, or γ-endorphin) found in highest concentrations in the pituitary gland and which has analgesic and behavioral effects in animals. *See also* enkephalin.

engram Memory trace, the neurophysiological substrate of memory.

enkephalin Endogenous opioid peptide released by β-endorphin that has analgesic properties and many of the actions of morphine.

entomophobia Fear of insects.

enuresis Repeated incontinence of urine in inappropriate places, such as clothing or bed; diagnosis not made until after age 5.

epidemiology In psychiatry, the study of the incidence, prevalence, control, and distribution of mental disorders within a particular population.

epigenesas Term introduced by Erik Erikson to refer to the stages of ego and social development.

epilepsy Neurological disorder resulting from a sudden, excessive, disorderly discharge of neurons in either a structurally normal or a diseased cerebral cortex; characterized by paroxysmal recurrence of short-lived disturbances of consciousness, involuntary convulsive muscle movements, psychic or sensory disturbances, or some combination thereof. Termed idiopathic epilepsy when there is no identifiable organic cause. For types of epilepsy, see the specific term.

epileptic dementia Epilepsy accompanied by progressive mental and intellectual impairment; some believe that the circulatory disturbances during epileptic attacks cause nerve cell degeneration and lead to dementia.

epileptic equivalent Epileptic phenomenon that is neither grand mal nor petit mal. *See also* complex partial seizure.

epileptic personality EXPLOSIVE PERSONALITY.

epinephrine Sympathomimetic catecholamine formed from norepinephrine and the major hormone secreted by the adrenal medulla; its release during states of fear or anxiety produces many of the physiological changes associated with those emotions. Also called adrenaline.

episodic Denoting an illness in which there are sustained disturbances clearly distinguished from previous functioning.

epistemology Study or theory of the nature and grounds of knowledge, especially with reference to its limits and validity.

erectile dysfunction MALE ERECTILE DISORDER.

eremophobia Abnormal fear of being by oneself.

ergasia Adolf Meyer's term for a person's total activity, as opposed to the functioning of parts of the whole.

Erikson, Erik (1902–1990) Psychoanalyst noted for his theory of ego development and psychosocial development, which he conceptualized in terms of social adaptation as it relates to Sigmund Freud's formulations.

erogenous zone Any area of the body, principally the genitals, anus, and mouth, capable of being stimulated sexually.

eros SEXUAL DRIVE.

erotic Invested with sexual feeling.

erotomania Delusional conviction that another person is in love with him or her; typically the object of the delusion is of higher rank or status. Also referred to as de Clérambault syndrome or psychose passionelle.

erythrophobia Abnormal fear of blushing.

escape learning *See* avoidance and escape learning.

ESP Extrasensory perception.

EST Electroshock therapy.

estrogen Hormone produced mainly by the ovaries and responsible for the development of the female reproductive organs and secondary sex characteristics: for example, estradiol, estriol, estrone. *Compare* androgen.

ethnocentrism Conviction that one's own group is superior to other groups; impairs one's ability to evaluate members of another group realistically or to communicate with them on an open, equal, and person-to-person basis.

ethnology Branch of anthropology dealing with the division of mankind into races and their origin, history, customs, and institutions.

ethology Study of animal behavior; often applied to the understanding of human behavior.

etiology Study of the causes of disease and their mode of operation.

euergasia Word used by Adolf Meyer to mean normal mental functioning.

euphoria Exaggerated feeling of well-being that is inappropriate to apparent events; may be induced by drugs such as opiates, amphetamines, and alcohol.

euphoric mood *See* euphoria.

euphorohallucinogen Substance capable of producing euphoric hallucinations.

euthymia Normal range of mood, implying absence of depressed or elevated mood.

euthymic Denoting a normal mood.

evasion Act of not facing up to or of strategically eluding something; consists of suppressing an idea that is next in a thought series and replacing it with another idea closely related to it. Also called paralogia; perverted logic.

event-related euphoria Morbid or abnormal sense of well-being associated with a specific situation or event.

evoked potential (EP) Following repetitive stimuli, the characteristic EEG waveform that remains after averaging out most of the brain electrical activity; clinically useful in evaluating the functional integrity of the somatosensory or special sensory pathways.

exaltation Feeling of intense elation and grandeur.

excited Agitated, purposeless motor activity uninfluenced by external stimuli.

excitement phase Stage in the human sexual response after sexual stimulation, as described by Masters and Johnson, characterized by an objective sense of pleasure and accompanying physical changes (penile erection and vaginal lubrication).

excitoxicity Increased simulation of neurons due to injury or chemical neurotransmitters that cause cell death.

excoriation disorder Disorder characterized by compulsive picking of one's own skin, resulting in lesions, despite repeated efforts to decrease or stop picking. Also called skin-picking disorder.

executive ego function Part of the ego that helps maintain or manage one's relationship to and sense of reality.

exhaustion state Final stage of the body's response to major stress, as in the general adaptation syndrome.

exhibitionism Paraphilia in which a man exposes his genitals to females in a socially inappropriate fashion; rarely occurs in women. Also called exhibitionistic disorder.

existentialism Philosophic theory that questions the purpose and nature of existence itself; a fundamental focus is on the meaning of life in face of death and the end of one's personal existence.

existential psychotherapy Type of therapy, based on existential philosophy, that puts the emphasis on here-and-now interaction and on feeling experiences, rather than on rational thinking; little attention is given to patient resistances, and the therapist is involved on the same level and to the same degree as the patient. *See also* phenomenology.

exogamy Marriage outside of a specific group; in a biological sense, the union of gametes of unrelated parents. *Compare* endogamy.

exogenous In psychiatry, denoting disordered behavior produced by forces outside of the body, as in adjustment disorder.

exorcism Magical practice in which mystical incantations are invoked to remove demons that are alleged to have entered the mind.

expansive mood Expression of feelings without restraint, frequently with an overestimation of their significance or importance. Seen in mania and grandiose delusional disorder.

experiencing Feeling emotions and sensations, as opposed to thinking; being involved in what is happening, rather than standing back at a distance and theorizing.

experimental bias Error in construction of an experiment or study that favors one outcome over another.

experimental design Manner in which an experiment is logically organized in order to maximize the probability of obtaining real effects.

experimental group In an experimental design, the group that is exposed to the experimental

manipulation under study. Compare control group.

experimental neurosis Abnormal behavior produced in a laboratory setting; used by Ivan Pavlov to describe symptoms of extreme and persistent agitation that he produced in dogs.

expert witness Legal concept that refers to the physician in the courtroom who is deemed sufficiently knowledgeable in a field to draw informed conclusions from data, for example, that a patient meets the required criteria for commitment or for an insanity defense.

explosive personality Personality type characterized by rigidity, egocentricity, selfishness, religiosity, seclusiveness, explosive outbursts of emotion, and extreme rage reactions when frustrated; enuresis is common.

expressive aphasia Disturbance of speech in which understanding remains, but ability to speak is grossly impaired; halting, laborious, and inaccurate speech. Also known as Broca's, nonfluent, and motor aphasias.

expressive dysphasia Difficulty in expressing verbal language; the ability to understand language is intact.

extended-family therapy Type of family therapy that involves family members, beyond the nuclear family, who are closely associated with it and affect it. *See also* social network therapy.

externalization More general term than projection that refers to the tendency to perceive in the external world and in external objects elements of one's own personality, including instinctual impulses, conflicts, moods, attitudes, and styles of thinking.

external validity Statistical term that refers to how well a test measures what it purports to measure and identifies what it is supposed to identify.

extinction Process by which an operant behavior occurs less frequently when its performance is no longer reinforced.

extrapsychic conflict Conflict that arises between a person and his or her environment. Also called interpersonal conflict. *Compare* intrapsychic conflict.

extrapyramidal dyskinesia Movement disorder related to dysfunction of the extrapyramidal motor system; signs include a variety of involuntary movements and postures, muscular rigidity, tremor, and gait disturbance, as seen in Parkinson disease, chorea, athetosis, and hemiballismus. Extrapyramidal dyskinetic syndromes may occur as a side effect of antipsychotic drugs; such drug-induced extrapyramidal effects include a parkinsonian syndrome, akathisia, acute dystonic reactions, and tardive dyskinesia.

extrapyramidal effect Bizarre, involuntary motor movement; a central nervous system side effect sometimes produced by antipsychotic drugs. *See also* extrapyramidal dyskinesia.

extrapyramidal motor system Polysynaptic neural pathways involving the basal ganglia and related subcortical nuclei that influence motor behavior.

extrapyramidal syndrome Abnormalities of movement related to injury of motor pathways other than the pyramidal tract; for example, parkinsonism, akathisia, dystonia, tardive dyskinesia.

extrasensory perception (ESP) Experiencing of an external event by means other than the five senses, as by telepathy or clairvoyance.

extroversion, extraversion State of one's energies being directed outside oneself. *Compare* introversion.

F

factitious disorder Mental disorder characterized by the voluntary production of false or unreal physical or psychological symptoms; unlike malingering, there is no apparent goal or obvious benefit. Also called Munchausen syndrome. *Compare* somatic symptom disorder; somatoform disorders.

factitious disorder by proxy A disorder in which one person deceitfully creates false signs or symptoms of illness in another person who is under his/her care; most commonly this condition involves a mother presenting her child as ill. In some cases actual injury or illness may be inflicted on the dependent person. Diagnosed as "factitious disorder imposed on another" in DSM-5.

factitious disorder imposed on another *See* factitious disorder by proxy.

factor analysis Data reduction technique used to reduce a large number of variables to a smaller number of linear combinations of variables.

failure to thrive Delay in physical growth in infants and young children; often accompanied by impaired social and motor development; commonly associated with parental deprivation or neglect when no medical illness is detected. *See* reactive attachment disorder.

false memory A person's recollection and belief of an event that did not actually occur. In false memory syndrome, persons erroneously believe that they sustained an emotional or physical (e.g., sexual) trauma early in life.

false negative Diagnosis of a disorder as being absent in a person (tested negative) when in fact it is present.

false positive Diagnosis of a disorder as being present in a person (tested positive) when in fact it is not present.

falsifiable hypothesis Testable hypothesis that can be confirmed or disconfirmed.

family neurosis Mental disorder in which a person's psychopathology is unconsciously interrelated with that of other members of the family.

family therapy Treatment of more than one member of a family in the same session; family relationships and processes are explored as potential causes of mental disorder in one or more of the family members. *See also* extended-family therapy.

fantasy Daydream; fabricated mental picture of a situation or chain of events. A form of thinking dominated by unconscious material and primary processes that seeks wish fulfillment and immediate solutions to conflicts; may serve as the matrix for creativity or for neurotic distortions of reality.

fatal familial insomnia An inherited prion disease characterized by insomnia and autonomic nervous system dysfunction (e.g., fever, sweating, tachycardia).

father surrogate, father substitute In psychoanalysis, a person onto whom the patient projects his or her father image and to whom he or she responds unconsciously in an inappropriate and unrealistic manner, with the feelings and attitudes held toward the real father.

fatigue A feeling of weariness, sleepiness, or irritability after a period of mental or bodily activity. Seen in depression, anxiety, neurasthenia, and somatoform disorders.

fausse reconnaissance False recognition, a feature of paramnesia. Can occur in delusional

disorders.

faux de mieux Literally, "for want of anything better," connoting in psychiatry a person's choosing a homosexual relationship when no partner of the opposite sex is available.

FDA Food and Drug Administration.

fear Unpleasurable emotional state consisting of psychophysiological changes in response to a realistic threat or danger. *Compare* anxiety.

feeble-mindedness Obsolete term for mental retardation, which is also called intellectual developmental disorder.

feedback 1. In a given system, the return, as input, of some form of information regarding the output; used as a regulatory mechanism. **2.** In psychiatry, the verbally or otherwise expressed response to a person's behavior by another person or a group.

feeding disorder of infancy or childhood *See* avoidant/restrictive food intake disorder.

fellatio Use of the mouth or tongue to stimulate the male genitalia.

femaleness Anatomic and physiological features relating to female procreative functions.

female orgasmic disorder Disorder characterized by a woman's inability to achieve orgasm by coitus or masturbation.

female sexual arousal disorder OMIT.

feminine identity Inner sense of gender affiliation with females. *Compare* masculine identity.

Ferenczi, Sandor (1873–1933) Hungarian psychoanalyst who was one of Sigmund Freud's early followers and a brilliant contributor to all aspects of psychoanalysis; he came to favor more active and personal techniques than Freud's, to the point that his adherence to psychoanalysis during his last years was questioned.

fetal alcohol syndrome A syndrome that occurs when mothers consume alcohol during pregnancy resulting in fetuses born with birth defects, CNS dysfunction, and intellectual disability.

fetishism Paraphilia in which sexual excitement and gratification are achieved by substituting an inanimate object (such as a shoe, piece of underwear, or other article of clothing) for a human love object. Also called fetishistic disorder.

fibromyalgia A chronic pain disorder characterized by pain and stiffness of the soft tissues (e.g., muscles, ligaments, tendons).

field theory Conceptual approach to the study of personality formulated by Kurt Lewin that focuses on an operational analysis of the causal determinants of human behavior; it is expressed in geometric and mathematical terms borrowed from physics in an attempt to provide a framework suitable to the scientific study of human behavior. The person and his or her environment together constitute the life space, a complex field of forces acting on the personality and determining behavior.

finger agnosia Agnosia characterized by inability to name fingers; suggests a lesion in the dominant parietal lobe.

fixation 1. Arrest of psychosexual development at any stage before complete maturation. **2.** Close and paralyzing attachment to another person, such as one's mother or father.

fixed-interval schedule *See* schedule of reinforcement.

fixed-ratio schedule *See* schedule of reinforcement.

flagellantism, flagellation Process by which sexual partners are aroused and gratified by

whipping or being whipped.

flashback hallucinosis Spontaneous recurrence, after a drug-free period, of hallucinations similar to those experienced during an acute toxic episode; often associated with repeated ingestions of LSD. Hallucinations are predominantly visual and may last for months after the last use of a drug.

flat affect Absence or near absence of any signs of affective expression.

flexibilitas cerea CATALEPSY.

flight of ideas Rapid succession of fragmentary thoughts or speech in which content changes abruptly and speech may be incoherent, as seen in mania.

floccillation Aimless plucking or picking, usually at bed-clothes or clothing, commonly seen in senile psychosis and delirium.

flooding Behavior therapy designed to extinguish phobias or anxiety through the presentation of the aversive stimuli in intense forms until the emotional responses cease. Also called implosion therapy.

fluent aphasia Aphasia characterized by inability to understand the spoken word; fluent but incoherent speech is present. Also called Wernicke aphasia, sensory aphasia, and receptive aphasia.

folie à deux Mental illness shared by two persons, usually involving a common delusional system; if it involves three persons, it is referred to as folie à trios, and so forth. *See* shared psychotic disorder.

follow-up examination To the physician, the aftercare and monitoring of patients' responses to treatment and the state of their physical health or mental status.

Food and Drug Administration (FDA) Federal agency that standards for the sale and use of drugs, food substances, and cosmetics.

forensic psychiatry Branch of psychiatry concerned with the aspects of mental illness.

foreplay Sexual play that precedes sexual intercourse. Also called forepleasure.

forepleasure FOREPLAY.

formal operations Jean Piaget's label for completely developed logical thinking capacities. *See also* Piaget, Jean.

formal thought disorder Disturbance in the form of thought rather than the content of thought; thinking characterized by loosened associations, neologisms, and illogical constructs; thought process is disordered, and the person is defined as psychotic. Characteristic of schizophrenia.

formication Tactile hallucination involving the sensation that tiny insects are crawling over the skin; most commonly encountered in cocaine addiction and delirium tremens.

for-profit hospital Investor-owned or proprietary hospital; hospital owned and operated by physicians or corporations that earns a profit for its investors.

fragile X syndrome A syndrome resulting from a mutation on the X chromosome. Includes severe intellectual disability, a large head, long face, prominent ears, and, in males, large testicles.

fragmentary delusion Poorly elaborated delusion, often one of many with no apparent interconnection.

fraternal twins DIZYGOTIC TWINS.

free association Investigative psychoanalytic technique devised by Freud in which the patient seeks to verbalize without reservation or censorship the passing contents of his or her mind.

freebase Cocaine hydrochloride mixed with ammonia and an alkali solution, and then extracted with a solvent into a crystalline form. *See also* crack.

free-floating anxiety Severe, pervasive, generalized anxiety that is not attached to any particular idea, object, or event; observed particularly in anxiety disorders, although it may be seen in some cases of schizophrenia.

fregoli syndrome The delusional belief that a known person is an imposter who assumes different appearances or disguises.

frequency Number of cases having a certain score or characteristic; relative frequency measures the number of persons in a specific group (e.g., sex or age) who have a disorder.

frequency distribution Statistical description of raw data in terms of the number of cases that fall into each interval within a set of data; often presented graphically in the form of a frequency histogram or frequency polygram.

Freud, Anna (1895–1982) Daughter of Sigmund Freud noted for her contributions to the developmental theory of psychoanalysis and the mechanisms of defense.

Freud, Sigmund (1856–1939) Austrian psychiatrist and the founder of psychoanalysis. With Josef Breuer, he explored the potentialities of cathartic therapy and then went on to develop the analytic technique and such fundamental concepts as mental sublimation, as well as superego, ego, and id formation and their applications throughout all spheres of human behavior. Among his publications are *The Interpretation of Dreams* (1899). *Studies on Hysteria* (1895, with Breuer), and *Three Essays on the Theory of Sexuality* (1905).

frigidity Colloquial term referring, in the female, to lack of sexual response or feeling, ranging from complete lack of arousal to incomplete climax.

frontal lobe syndrome Mental disorder characterized by a marked change in personality, with development of certain relatively characteristic patterns of relating to the environment that have been associated with lesions of the frontal lobes; such patterns include emotional lability, impairment in impulse control and social judgment, and marked apathy and indifference.

frontotemporal dementia Rare form of presenile dementia caused by cerebral atrophy predominantly in the frontotemporal regions of the brain; early stages are characterized by personality and behavior changes. The following progressive cognitive deterioration is indistinguishable clinically from the far more common Alzheimer disease, in which the cerebral atrophy is diffuse. Also called Pick disease. *See also* dementia of the Alzheimer type.

frottage, frotteurism Sexual aberration in which arousal results from rubbing up against someone, usually without specific genital contact, as in a crowd; a person so afflicted is a frotteur. Also called frotteuristic disorder.

frustration Result of the thwarting of any goal-directed behavior; the ability to tolerate frustration and delay gratification is considered a sign of maturity and good ego strength.

fugue Called dissociative fugue in DSM-5, a subtype of dissociative amnesia characterized by a period of almost complete amnesia, during which a person actually flees from an immediate life situation and begins a different life pattern; apart from the amnesia, mental faculties and skills are usually unimpaired. *See also* psychogenic amnesia.

fulfillment Satisfaction of a need or wish.

functional Denoting changes in functioning not attributable to known organic alterations.

functional disorder Disorder that is not caused by an anatomical defect or any other identifiable or demonstrable cause. *Compare* organic disorder.

functional neurological disorder *See* conversion disorder.

functional neurological symptom disorder Mental disorder in which the patient experiences an involuntary limitation or alteration of physical function that is an expression of a psychological conflict or need, not of a physical disorder. Previously called conversion disorder in DSM-IV. *See also* hysteria.

functional speech disorder DEVELOPMENTAL ARTICULATION DISORDER.

fusion In psychoanalysis, the joining together of instincts which enables the person to function normally.

G

GABA γ-(gamma) aminobutyric acid.

GAF Scale GLOBAL ASSESSMENT OF FUNCTIONING SCALE.

galactorrhea Abnormal discharge of milk from the breast; may be a result of the endocrine influence of phenothiazine drugs.

galvanic skin response (GSR) Measurable change in the electrical resistance of the skin, used as an index of emotional response.

gambling disorder DSM-5 disorder characterized by persistent and recurrent maladaptive gambling that leads to economic problems and disturbances in personal, social, and occupational functioning. Also called pathological gambling.

γ-(gamma) aminobutyric acid (GABA) Inhibitory amino acid neurotransmitter synthesized by the actions of glutamic acid decarboxylase; GABA receptors are located at postsynaptic sites.

gamophobia Abnormal fear of marriage.

Ganser syndrome Rare atypical disorder, seen in prisoners, characterized by a giving of approximate answers to questions and of talking past the point; to be distinguished from factitious disorder and malingering.

gargoylism Old name for the gargoyle-like facies and dwarfism seen in Hunter's and Hurler's syndromes: metabolic disorders of mucopolysaccharide metabolism characterized by severe developmental skeletal deformities and mental retardation.

GAS General adaptation syndrome.

gatekeeper In community psychiatry, a person who evaluates the degree of psychological or environmental stress sufficient to enable a patient to gain access to the mental health care system.

gateway drugs Substances of abuse, such as alcohol and marijuana, which provide major portals of entry into abuse of other drugs; considered particularly dangerous because they are believed by so many to be relatively harmless.

Gault decision U.S. Supreme Court 1967 decision mandating fair and accurate fact-finding procedures for juveniles when serious punishment could be inflicted should they be found guilty; rights guaranteed the juvenile include the right to notification of the specific charges,

the right to counsel, the right against self-incrimination, and the right to confront and cross-examine the accusers.

gay 1. Popular term for a homosexual, especially male. **2.** Denoting a homosexual or the lifestyle thereof.

gegenhalten Neurological term (from the German) for involuntary active resistance to passive movement of the extremities, as may occur in cerebral cortical disorders.

gender dysphoria The DSM-5 term for gender identity disorder. May first appear in childhood, adolescence, or adulthood.

gender identity Culturally determined sets of attitudes, behavior patterns, and physical attributes usually associated with masculinity and femininity. *Compare* sexual identity.

gender identity disorder Psychosexual disorder in which a person feels discomfort with and inappropriateness of his or her biological sex, with a marked preference for the clothing and activities of the opposite sex and/or repudiation of the sex organs. Also called gender dysphoria. *See also* transsexualism.

gender role Public declaration of gender; the image of maleness or femaleness that is communicated to others which may or may not coincide with gender identity.

general adaptation syndrome (GAS) Hans Selye's term for the responses of the body to major stress, passing through the alarm reaction, resistance and, finally, exhaustion.

generalization Process by which a behavior occurs in a setting in which it had not previously been reinforced.

generalized anxiety *See* free-floating anxiety.

generalized anxiety disorder An anxiety disorder characterized by chronic generalized anxiety consisting of feelings of tension, tremor, apprehension, and hyperarousal not attached to any particular idea, object, or event. *See* free-floating anxiety.

generalized tonic–clonic seizure Generalized onset of tonic–clonic movements of the limbs, tongue biting, and incontinence followed by slow, gradual recovery of consciousness and cognition. Also called grand mal seizure.

general paralysis, general paresis Mental disorder caused by chronic (tertiary) neurosyphilis.

general systems theory Theoretical framework asserting that the universe is composed of interacting components that are interrelated; in psychiatry, it emphasizes an integrated, holistic approach to understanding behavior.

genetic counseling Presentation and discussion, generally with a prospective parental couple, of factors involved in the inheritance of pathological conditions as they relate to the couple's generic endowment.

genetic endowment Characteristics that are inherited, usually regarding potential capacities of intelligence, temperament, and general adaptability.

genetic marker Detectable sign of the presence of or vulnerability to inherited disease, such as a measurable biochemical abnormality or an identifiable defect in a particular gene.

genetic material In psychiatry and psychoanalysis, a term used to reflect the data from a patient's personal past history that are useful in developing an understanding of the psychodynamics of his or her present adaptation.

genital phase Final stage of psychosexual development, occurring during puberty, in which

sexual gratification is achieved from genital-to-genital contact and the capacity exists for a mature, affectionate relationship with someone of the opposite sex.

genotype Genetic endowment of a person, consisting of a set of genes received from each parent at the time of conception. *Compare* phenotype.

geriatric psychiatry Branch of psychiatry concerned with the mental disorders that accompany old age. Also called geropsychiatry.

geriatrics Branch of medicine concerned with the aged and the problems of aging. *See also* gerontology.

genito–pelvic pain/penetration disorder Diagnostic category in DSM-5 referring to one or more of the following: genito–pelvic pain; difficulty having intercourse; fear of pain or penetration; and tension of the pelvic floor muscles; previously divided into the distinct categories of dyspareunia and vaginismus.

gerontology Scientific study of aging.

gerontophilia Sexual attraction for the old and the aged by a young person.

gerontophobia Fear of the elderly or of the aging process.

geropsychiatry GERIATRIC PSYCHIATRY.

Gerstmann–Straussler–Scheinker disease DEVELOPMENTAL ARITHMETIC DISORDER.

Gesell Developmental Schedules Developmental landmarks in infants and children that cover ages 8 weeks to 3.5 years. Data are obtained by direct observation of the child's responses to standard toys and other controlled stimulus objects and are supplemented by developmental information provided by the mother or primary caretaker; the schedules yield scores indicating the child's level of development in four separate areas: motor, adaptive, language, and personal and social.

gestalt psychology School of psychology primarily concerned with perceptual processes; a major tenet is that the whole is greater than the sum of its parts. Its development is associated particularly with Max Wertheimer, Wolfgang Koehler, and Kurt Koffka.

gestalt therapy Psychotherapy developed by Frederick S. Perls that emphasizes the treatment of the person as a whole: one's biological component parts and their organic functioning, one's perceptual configuration, and one's interrelationships with the outside world. It can be used in either an individual or a group therapy setting; focuses on the sensory awareness of the person's here-and-now experiences, rather than on past recollections or future expectations; and uses role playing and other techniques to promote the patient's growth process and development to full potential.

Gilles de la Tourette syndrome *See* Tourette disorder.

global aphasia Combination of grossly nonfluent aphasia and severe fluent aphasia.

Global Assessment of Functioning Scale (GAF Scale) Scale used in DSM to describe the highest social, occupational, or educational level of functioning of a person in the 12 months preceding the current evaluation.

global developmental delay Preferred term for a condition of intellectual and adaptive impairment in children that fail to meet expected developmental milestones in multiple areas of intellectual functioning before 5 years of age. *See also* intellectual disability.

globus hystericus Hysterical symptom in which a person is disturbed by the sensation of a

lump in the throat.

glossolalia Unintelligible jargon that has meaning to the speaker but not to the listener. Occurs in schizophrenia.

go-around Technique used in group therapy in which the therapist requires each member of the group to respond to another member, a theme, or an association to encourage the participation of all members in the group.

gonadal dysgenesis TURNER SYNDROME.

grandiose delusion DELUSION OF GRANDEUR. *See* delusion.

grandiosity Exaggerated feeling of one's importance, power, knowledge, or identity. Occurs in delusional disorder and manic states. *See also* delusion of grandeur.

grand mal epilepsy Major form of epilepsy in which gross tonic–clonic convulsive seizures are accompanied by loss of consciousness and, often, incontinence of stool or urine; frequently preceded by an aura.

gray-out syndrome Psychosis that occurs in pilots flying at high altitudes, out of sight of the horizon.

grief Alteration in mood and affect consisting of sadness appropriate to a real loss; normally, it is self-limited. *See also* depression; mourning.

group *See* therapeutic group.

group analytic psychotherapy Term applied by S. H. Foulkes to his pioneering group therapy procedure in which interventions dealt primarily with group, rather than individual, forces and processes, and the group was used as the principal therapeutic agent.

group cohesion Mutual bonds formed between members of a group as a result of their concerted effort for a common interest and purpose; until cohesiveness is achieved, the group cannot concentrate its full energy on a common task.

group dynamics Phenomena that occur in groups from their inception to their termination. Interactions and interrelations among members and between the members and the therapist create tension, which maintains a constantly changing group equilibrium; interactions and the tension created are highly influenced by individual members' psychological makeup, unconscious instinctual drives, motives, wishes, and fantasies. Understanding and effectively using group dynamics are essential in group treatment. Also called group process.

group mind Autonomous and unified mental life in an assemblage of people bound together by mutual interests; a concept used by group therapists who focus on the group as a unit rather than on the individual members.

group practice Delivery of medical services by three or more physicians formally organized to provide care, consultation, diagnosis, and treatment; they share the use of equipment and personnel, and income from the medical practice is distributed among the members of the group.

group pressure Demand by group members that individual members submit and conform to group standards, values, and behavior.

group process GROUP DYNAMICS.

group psychotherapy Application of psychotherapeutic techniques to a group of patients, using interpatient interactions to effect changes in the maladaptive behavior of the individual members.

group relations theory *See* Allport's group relations theory.

GSR Galvanic skin response.

guardianship Person appointed by the courts for the care and management of a person or the property of another, as of a minor or a person incapable of managing his or her affairs.

guilt Emotional state associated with self-reproach and the need for punishment. In psychoanalysis, refers to a neurotic feeling of culpability that stems from a conflict between the ego and the superego that begins developmentally with parental disapproval and becomes internalized as conscience in the course of superego formation. Guilt has normal psychological and social functions, but special intensity or absence of guilt characterizes many mental disorders, such as depression and antisocial personality disorder, respectively. Some psychiatrists distinguish shame as a less internalized form of guilt that relates more to others than to the self. *See also* shame.

gustatory hallucination Hallucination primarily involving taste.

gynecomastia Female-like development of the male breasts; may occur as an adverse effect of antipsychotic and antidepressant drugs due to increased prolactin levels or anabolic-androgenic steroid abuse.

gynephobia Abnormal fear of women.

gyrectomy Psychosurgery involving excision of a cerebral gyrus.

H

habeas corpus Legal term for the right to petition a court to decide whether confinement has been undertaken with due process of 1aw.

habit Behavior pattern acquired by frequent repetition that leads to regularity or increased facility of performance.

habituation **1.** Simple form of learning in which the response to a repeated stimulus lessens over time. **2.** Psychological dependence on the continued use of a drug or chemical substance to maintain a sense of well-being, which can result in addiction.

halfway house Facility for mental patients who no longer need the full facilities of a hospital but are not yet ready to return to their communities.

hallucination False sensory perception occurring in the absence of any relevant external stimulation of the sensory modality involved. For types of hallucination, see the specific term.

hallucinogenic drug, hallucinogen Chemical agent that produces hallucinations, such as LSD and mescaline. *See also* euphorohallucinogen; psychotomimetic drug.

hallucinosis State in which a person experiences hallucinations without any impairment of consciousness.

haloperidol Butyrophenone antipsychotic drug (trade name, Haldol).

Halstead–Reitan Scale Battery of neuropsychological tests used to determine the location and effects of specific brain lesions. A great variety of functions are tested: concept function, abstraction, and visual acuity; dexterity; spatial memory; auditory perception, attention, and concentration; finger-oscillation; speech-sounds perceptions; visual perception; time sense; and sensory stimulation, stereognosis, and tactile perception.

Hamilton Depression Scale Scale of depression based upon a clinical interview with the patient in which the examiner asks about guilt, suicide, sleep habits, weight change, and other

symptoms of depression.

haphephobia Abnormal fear of being touched.

haplology Rapid speech in which syllables are omitted; seen in certain manias and in schizophrenic conditions.

haptic hallucination Hallucination of touch.

harm avoidance A term used to describe a character trait in which the person avoids novel situations and taking risks; may be an inborn and genetically determined.

hashish Dried resinous tops of the marijuana plant (*Cannabis sativa*): the active ingredient is tetrahydrocannabinol, which is a potent psychedelic drug.

Health and Human Services Federal organization composed of the following: Office of Human Development Services, Public Health Service (includes Centers for Disease Control; Food and Drug Administration; Health Resources and Services Administration; National Institutes of Health; Alcohol, Drug Abuse, and Mental Health Administration; Agency for Toxic Substances and Disease Registry; and Indian Health Service), Social Security Administration, Health Care Financing Administration, and Family Support Administration. Formerly called Department of Health, Education, and Welfare (HEW).

health insurance Insurance that reimburses one for the costs of medical care or lost income arising from medical disability; may include coverage for the care of mental conditions.

health maintenance organization (HMO) Form of group practice by physicians, often including psychiatrists, and allied health professionals to provide comprehensive health services to subscribers who pay a fixed premium.

Health Systems Agency (HSA) Nonprofit, federally mandated, statewide organizations that promote or limit the development of health services and faculties depending on the needs of the particular state or locale. HSAs develop long- and short-term goals and plans, approve health care proposals requesting federal funding, review existing facilities and services, and suggest future consideration and renovation projects based on their findings.

healthy identification Modeling of oneself, consciously or unconsciously, on another person who has a sound psychic makeup; such identification has constructive results. *See also* imitation.

hearsay Out-of-court statement offered to prove the truth of the matter asserted; usually inadmissible as testimony.

hebephrenia Complex of symptoms, considered a form of schizophrenia, characterized by wild or silly behavior or mannerisms, inappropriate affect, frequent hypochondriacal complaints, and delusions and hallucinations that are transient and unsystematized. Hebephrenic schizophrenia is listed in DSM as disorganized schizophrenia. *See also* schizophrenia.

hedonism Pleasure-seeking behavior. *Compare* anhedonia.

helplessness 1. Learned: State of apathy that occurs as a type of experimental neurosis in animals who cannot escape an aversive stimulus; proposed as the animal model of human depression in which a patient feels helpless to control events. **2.** Psychic: State that occasions the first expressions of anxiety in humans, for example, the experience of birth.

hematophobia Abnormal fear of the sight of blood.

herd instinct Desire to belong to a group and to participate in social activities; used by Wilfred

Trotter to indicate the presence of a hypothetical social instinct in humans. In psychoanalysis, herd instinct is viewed as a social phenomenon, rather than as an instinct.

hermaphrodite Person who has both female and male gonadal tissue, usually with one sex dominating.

heroin Diacetylmorphine; an illicit, addictive alkaloid opioid that requires subsequent larger and larger amounts to achieve the same narcotic and analgesic effects. Injected intravenously, it produces a more rapid onset of euphoria than does morphine.

heterogeneity Dissimilarity in the genotypal structure of individuals originating through sexual reproduction.

heterosexuality Sexual attraction or contact between opposite-sex persons. The capacity for heterosexual arousal is probably innate, biologically programmed, and triggered in very early life, perhaps in part by olfactory modalities as seen in lower animals. *Compare* bisexuality; homosexuality.

heterozygous Possessing different allelic forms of a gene at a given locus in homologous chromosomes.

heuristic 1. Involving or serving as an aid to learning, discovery, or problem solving by experimental and especially trial-and-error methods. **2.** Of or relating to exploratory problem-solving techniques that utilize self-educating techniques.

5-HIAA 5-Hydroxyindoleacetic acid.

histrionic personality disorder Condition in which the patient, usually an immature and dependent person, exhibits unstable, overreactive, and excitable self-dramatizing behavior that is aimed at gaining attention and is at times seductive, although the person may not be aware of that aim.

HIV Human immunodeficiency virus.

HMO Health maintenance organization.

hoarding disorder Disorder characterized by persistent difficulty parting with or discarding possessions; delusional beliefs about the accumulated items may be present. Part of obsessive-compulsive disorder in DSM-5.

holism In psychiatry, the study of the total person as a distinctive entity, rather than as a collection of various characteristics.

holophrastic Using a single word to express a combination of ideas; schizophrenic persons may use holophrastic language.

homeless Persons who have no permanent place of residence, often sleeping in the street, many of whom require psychiatric treatment or hospitalization.

homeostasis Tendency to maintain a constancy and stability of bodily processes to ensure optimal functioning; the state of bodily equilibrium and the processes whereby such equilibrium is maintained.

homophobia Fear of homosexual persons.

homosexuality Sexual attraction or contact between same-sex persons. Some authors distinguish two types: overt and latent. *Compare* bisexuality; heterosexuality.

homosexual panic Sudden, acute onset of severe anxiety precipitated by the unconscious fear or conflict that one may be a homosexual or act out homosexual impulses.

homovanillic acid (HVA) Phenol found in human urine; as the major metabolite of

dopamine, it is most often measured in psychiatric research.

Horney, Karen (1885–1952) Psychiatrist and psychoanalyst whose theories of the genesis of neurosis emphasized environmental and cultural factors and thus departed from the biological-instinctual framework of orthodox freudian thought.

hostile transference *See* transference.

hostility Conflict, opposition, or resistance in thought or principle.

hotline Telephone assistance service for crisis intervention.

HSA Health Systems Agency.

5-HT 5-Hydroxytryptamine.

human immunodeficiency virus (HIV) An infection caused by a retrovirus that incorporates itself into deoxyribonucleic acid (DNA). Symptoms include fever, myalgia, headache, fatigue, gastrointestinal symptoms, and rash. Can cause acquired immunodeficiency syndrome (AIDS).

humanistic theory Group of diverse theories that share a holistic conceptualization that emphasizes the uniqueness, value, dignity, and worth of each human being as an individual. Human behavior is considered understandable only in terms of the meaning of experiences for the individual person; it represents a complex interplay of numerous physical, psychological, and sociocultural factors.

humiliation Sense of disgrace, dishonor, and shame; often experienced by depressed patients.

humor Defense mechanism that allows the overt expression of feelings and thoughts without personal discomfort or immobilization and does not produce an unpleasant effect on others; allows the individual to tolerate and yet focus on what is too terrible to be borne. Differs from wit, a form of displacement that involves distraction from the affective issue.

Huntington disease (chorea) Hereditary and progressive central nervous system disease characterized by jerking motions and progressive mental deterioration: inherited as an autosomal dominant, with onset in adult life.

HVA Homovanillic acid.

hydrocephalus Excessive accumulation of cerebrospinal fluid which dilates the cerebral ventricles.

hydrophobia In psychiatry and psychology, an abnormal fear of water.

hydrotherapy Use of water by external application in the treatment of disease.

5-hydroxyindoleacetic acid (5-HIAA) Metabolite of serotonin, the concentration of which can be measured in cerebrospinal fluid and serum: determination of 5-HIAA concentrations is used in psychiatric research and may help guide clinical practice.

5-hydroxytryptamine (5-HT) SEROTONIN.

hyperactivity Increased muscular activity. The term is commonly used to describe a disturbance found in children that is manifested by constant restlessness, overactivity, distractibility, and difficulties in learning. Also called hyperkinesis. *See also* attention-deficit hyperactivity disorder; minimal brain dysfunction.

hyperalgesia Excessive sensitivity to pain. Seen in somatoform disorder.

hyperarousal A state of increased psychological and physiological tension marked by such effects as reduced pain tolerance, anxiety, exaggeration of startle responses, insomnia, fatigue, and accentuation of personality traits.

hyperesthesia Increased sensitivity to tactile stimulation.

hyperkinesis HYPERACTIVITY.

hyperkinetic syndrome Disorder characterized by excessive muscular activity (hyperactivity) such as restlessness, aggressiveness, and destructive activity. *See also* attention-deficit hyperactivity disorder; minimal brain dysfunction.

hypermnesia Exaggerated degree of retention and recall. It can be elicited by hypnosis and may be seen in certain prodigies; also may be a feature of obsessive-compulsive disorder (OCD), some cases of schizophrenia, and manic episodes of bipolar I disorder.

hyperorexia Extreme appetite. *See also* bulimia nervosa.

hyperphagia Increase in appetite and intake of food.

hyperpragia Excessive thinking and mental activity, generally associated with the manic phase of bipolar I disorder.

hypersexuality Term used to describe excessive or compulsive interest in sex that may result in sex addiction.

hypersomnia Disorder characterized by excessive time spent asleep which may be nonorganic, organic, narcoleptic, idiopathic, part of Klein–Levin syndrome, or primary in nature. Also called hypersomnolence disorder.

hypersomnolence disorder *See* hypersomnia.

hypertensive crisis HYPERTENSIVE ENCEPHALOPATHY.

hypertensive encephalopathy Acute syndrome in which severe hypertension is associated with headache, nausea, vomiting, visual disturbances, convulsions, and alterations in consciousness characterized by confusion, somnolence, stupor, and coma. Prompt lowering of the blood pressure by pharmacological means may reverse the picture in 24 to 48 hours; if the blood pressure is not controlled, intracerebral hemorrhage and a fatal outcome may result. Usually, the patient has entered the malignant phase of essential or secondary hypertension, with marked blood pressure elevation (generally in the range of 250/150 mm Hg) and evidence of cardiac decompensation, declining renal function, and advanced hypertensive retinopathy; also seen as a side effect of monoamine oxidase inhibitors when tyramine-containing foods are consumed. Also called hypertensive crisis.

hyperthyroidism Excessive activity of the thyroid gland; may be associated with dementia, anxiety disorders, and insomnia. *Compare* hypothyroidism.

hyperventilation Excessive breathing, generally associated with anxiety, which can cause a reduction in blood carbon dioxide that produces symptoms of lightheadedness, palpitations, numbness and tingling periorally and in the extremities, and occasionally syncope.

hypervigilance Excessive attention to and focus on all internal and external stimuli; usually seen in delusional or paranoid states.

hypesthesia Diminished sensitivity to tactile stimulation.

hypnagogic hallucination Hallucination occurring while falling asleep, not ordinarily considered pathological.

hypnoanalysis Use of hypnosis in psychoanalysis to gain access to unconscious processes that the patient cannot reveal by means of ordinary therapeutic maneuvers.

hypnodrama Psychodrama under hypnotic trance; the patient is put into a hypnotic trance and encouraged to act out various past experiences.

hypnoid state Freud's term describing an alteration of consciousness that occurs characteristically in hysteria during periods of emotional stress: characterized by heightened suggestibility and provides a basis for hysterical somatic symptom formation.

hypnophobia Abnormal fear of sleep.

hypnopompic hallucination Hallucination occurring while awakening from sleep, ordinarily not considered pathological.

hypnosis Artificially induced alteration of consciousness characterized by increased suggestibility and receptivity to direction.

hypnotherapy Therapy that makes use of hypnosis.

hypnotic Drug used for the express purpose of producing sleep. *Compare* sedative.

hypoactive sexual desire disorder Sexual desire disorder characterized by deficiency or absence of sexual fantasies and desire for sexual activity. Also called sexual interest/arousal disorder in women.

hypoactivity Decreased motor and cognitive activity, as in psychomotor retardation; visible slowing of thought, speech, and movements. Also called hypokinesis.

hypochondria Exaggerated concern about health that is based not on real medical pathology, but on unrealistic interpretations of physical signs or sensations as abnormal.

hypochondriacal neurosis HYPOCHONDRIASIS.

hypochondriasis *See* somatic symptom disorder.

hypoglycemia Abnormally small concentration of glucose in the circulating blood; associated with anxiety and depression.

hypomania Mood abnormality with the qualitative characteristics of mania, but less intense; characterizes the hypomanic episode of bipolar II disorder when lasting 4 to 7 days.

hypothalamus Principal center in the forebrain for integration of visceral functions involving the autonomic nervous system. Functions include regulation of sexual activity, water, fat, carbohydrate metabolism, and temperature; also appears to play a role in the nervous mechanisms underlying moods and motivational states.

hypothyroidism Decreased activity of the thyroid gland; may be associated with dementia, depression, and sleep disorders. *Compare* hyperthyroidism.

hysteria A term previously used for a group of conditions involving a sudden impairment of function in response to emotional stress.

hysterical anesthesia Functional disorder characterized by the absence of tactile sensation in an area of the body; observed in certain cases of conversion disorder.

hysterical personality *See* histrionic personality disorder.

hysterics Expression of emotion often accompanied by crying, laughing, and screaming; a term generally used by laypersons rather than by psychiatrists.

I

iatrogenic Denoting an unfavorable response to a mode of treatment induced by the treatment itself. *Compare* idiopathic.

ICD *International Classification of Diseases.*

ice Pure form of methamphetamine that can be either inhaled or injected intravenously by abusers; its psychological effects last for hours and are particularly potent. Unlike the cocaine

of crack, which has to be imported, ice is a synthetic drug that can be manufactured in domestic illicit laboratories.

ICT Insulin coma therapy.

Id One of the three components of the psychic apparatus in the freudian structural framework, the others being the ego and superego. It operates unconsciously; harbors the innate, biological, instinctual drives; is the source of psychic energy (libido); follows the pleasure principle; seeks immediate reduction of drive tension without regard for external reality; and is under the influence of the primary-process mental activity that characterizes the unconscious.

idealization Mental mechanism whereby a person consciously or unconsciously overestimates an admired attribute or aspect of another person.

idea of reference Misinterpretation of incidents and events in the outside world as having a direct personal reference to oneself; occasionally observed in normal persons, but frequently seen in paranoid patients. If present with sufficient frequency or intensity or if organized and systematized, they constitute delusions of reference.

ideas of unreality Thought that events are artificial, illusory, unpredictable, or do not exist.

ideational shield Intellectual, rational defense against the anxiety that persons would feel if they became vulnerable to the criticisms and rejection of others; as a result of their fear of being rejected, they may feel threatened if they criticize another person, an act that is unacceptable to them. In both group and individual therapy, conditions are set up that allow the participants to lower the ideational shield.

idée fixe Fixed idea that is recurrent and most often associated with obsessional states.

identical twins MONOZYGOTIC TWINS.

identification Unconscious defense mechanism by which a person patterns himself or herself after another person; in the process, the self is more or less permanently altered. *See also* healthy identification; imitation.

identification with the aggressor Unconscious process by which a person incorporates within himself or herself the mental image of a person who represents a source of frustration from the outside world; a primitive defense that operates in the interest and service of the developing ego. A classic example of this defense occurs toward the end of the oedipal stage, when a boy, whose main source of love and gratification is his mother, identifies with his father who represents the source of frustration, being the powerful rival for the mother; the child cannot master or runaway from his father, so he is obliged to identify with him.

identity One's global role in life and the perception of the sense of self; problems with identity are common during adolescence, in schizophrenia, and in the borderline and schizotypal personalities.

identity crisis Social role conflict as perceived by a person; loss of the sense of personal sameness and historical continuity, and/or inability to accept or adopt the role the person believes is expected by society. Identity crises are frequent in adolescence.

identity disorder A disorder characterized by a chaotic sense of self; a loss of the sense of personal sameness, usually involving a social role conflict as perceived by the individual. It is common in adolescence, when adolescents feel unwilling or unable to accept or adopt the roles they believe are expected of them by society; often manifested by isolation, withdrawal, rebelliousness, negativity, and extremism. Not a DSM-5 diagnosis.

idiopathic Without known or identifiable cause. *Compare* iatrogenic.

idiot Obsolete and pejorative term for a classification of mental retardation in which individuals have a mental age of less than 3 years.

idiotropic Egocentric or introspective.

idiot-savant Mentally retarded person who is able to perform unusual mental feats in sharply circumscribed intellectual areas, such as complicated calculations or puzzle solving.

I-It Philosopher Martin Buber's description of damaging interpersonal relationships. If a person treats himself or herself or another person exclusively as an object, he or she prevents mutuality, trust, and growth. When pervasive in a group, I-It relationships prevent human warmth, destroy cohesiveness, and retard group process. *Compare* I-Thou.

illness anxiety disorder Preoccupation with being sick or developing a disease with few or no somatic symptoms and an absence of related pathologies found on medical examination. Differentiated from somatic symptom disorder (hypochondriasis) by the absence of somatic symptoms or complaints; patients with illness anxiety disorder are primarily concerned with "the idea" of being or becoming sick; a DSM-5 diagnosis. *Compare* somatic symptom disorder.

illogicality Pattern of speech or thinking in which conclusions that are reached do not follow logically; may take the form of unwarranted or faulty inferences.

illogical thinking Thinking containing erroneous conclusions or internal contradictions; psychopathological only when it is marked and not caused by cultural values or intellectual deficit.

illuminism State of hallucinations in which a patient converses with imaginary supernatural creatures.

illusion Perceptual misinterpretation of a real external stimulus. *Compare* hallucination.

imago Jungian term referring to an idealized unconscious mental image of a key person in someone's early life.

imbecile Obsolete and pejorative term for a classification of mental retardation in which individuals have a mental age of 3 to 7 years.

imitation Duplication or mimicking of behaviors or portions of behaviors observed in a model. *See also* identification.

immediate memory Reproduction, recognition, or recall of perceived material within seconds after presentation. *Compare* long-term memory; short-term memory.

impaired insight Diminished ability to understand the objective reality of a situation.

impaired judgment Diminished ability to understand the objective reality of a situation. *See also* judgment.

implosion Behavior therapy technique in which anxiety-arousing stimuli are vividly presented in imagination; the patient repeatedly experiences intense anxiety in the absence of objective danger until the anxiety response is extinguished.

impotence A term used to refer to the inability to obtain an erection. In DSM, it is diagnosed as erectile disorder.

imprinting Particular kind of learning that occurs during an early critical developmental stage in animals, characterized by rapid acquisition and great resistance to extinction. It seems to represent a biologically predetermined proclivity to learn certain persistent species-specific behavior patterns, give exposure to the proper stimulus during the critical stage, but its precise

role in human behavioral development has not been established.

improvisation In psychodrama, the acting out of situations without prior preparation.

impulse Psychic striving or urge to perform an action. *See also* drive; instinct.

impulse control Ability to resist an impulse, drive, or temptation to perform some action.

impulse-control disorders Mental disorders in which there is weak impulse control with impulsive behavior that is usually irresistible, pleasurable, and aimed at obtaining immediate gratification, without regard for the consequences of the behavior, for example, pathological gambling, kleptomania, pyromania, intermittent explosive disorder, trichotillomania.

impulsion Blind following of internal drives without regard for social acceptance or pressure from the superego, normally seen in young children. In adults, it is common in those with weak defensive organizations; in such cases it tends to be a symbolic phenomenon.

inappropriate affect Emotional tone out of harmony with the idea, thought, or speech accompanying it. Seen in schizophrenia.

incest Sexual activity between close blood relatives. Common patterns are father–daughter, mother–son, and between siblings; may also be homosexually oriented.

incidence Research term for the number of cases of a disease whose onset occurs during a specific period of time. *See also* prevalence.

incidence rate Rate at which new cases of a disease or a condition are occurring. The incidence rate per 1,000 persons in the population is the number of new cases that occur within a defined unit of time, divided by the persons exposed or at risk during the same time unit, multiplied by 1,000.

incoherence Communication that is disconnected, disorganized, or incomprehensible. *See also* word salad.

incompetence Legal term indicating that thought processes are inadequate for sound judgment and may lead to maladaptive or abnormal behavior; legally, incompetent persons cannot be held responsible for their actions.

incorporation Primitive unconscious defense mechanism in which the psychic representation of another person or aspects of another person are assimilated into oneself through a figurative process of symbolic oral ingestion; represents a special form of introjection and is the earliest mechanism of identification.

increased libido Increase in sexual interest and drive. Often associated with mania.

independent variable In research methodology, those qualities that the experimenter systematically varies, for example, time, dosage.

individual psychology System of psychiatric theory developed by Alfred Adler that stresses compensation and overcompensation for feelings of inferiority in a person's strivings to adapt to the social milieu, and emphasizes the interpersonal nature of a person's problems; applied by some practitioners to group psychotherapy and counseling.

individual therapy Traditional dyadic therapeutic technique in which a psychotherapist treats only one patient during a given therapeutic session.

individuation Jungian term denoting a process whereby the person molds and develops a healthy integrated individual personality through maximum differentiation and development of each system of the personality. *See also* actualization.

Indoklon therapy Type of biological therapy in which a convulsive seizure is induced by the

inhalation of the drug Indoklon (hexafluorodiethyl ether).

indolamine Class of biogenic amines that contain an indole ring and are derived from the essential amino acid tryptophan; includes lysergic acid diethylamide (LSD) and serotonin. Monoamine oxidase inhibitors elevate the levels of indolamines by delaying their breakdown.

induced psychotic disorder *See* shared psychotic disorder.

industrial psychiatry Branch of psychiatry concerned with workers' adjustments to their jobs and with the effects of a business organization on its members.

ineffability Ecstatic state in which a person insists that his or her experience is inexpressible and indescribable, that it is impossible to convey what it is like to one who never experienced it.

infancy Childhood period of helplessness and marked dependency, generally the first year of life.

infantile articulation DEVELOPMENTAL ARTICULATION DISORDER.

infantile autism A term now subsumed under autism spectrum disorder that refers to a syndrome beginning in infancy and characterized by withdrawal and self-absorption, failure to develop attachment to a parental figure, ineffective communication and mutism, preoccupation with inanimate objects, and an obsessive demand for sameness in the environment. Also called Kanner syndrome; autistic spectrum disorder. *See also* autistic thinking.

infantile dynamics Psychodynamic integrations, such as the Oedipus complex, that are organized during childhood and continue to exert an unconscious influence on adult personality.

infantile perseveration DEVELOPMENTAL ARTICULATION DISORDER.

infantile sexuality Freudian concept regarding the erotic life of infants and children, that, from birth, infants are capable of erotic activities; encompasses the overlapping phases of psychosexual development during the first 5 years of life and includes the oral phase (birth to 18 months), the anal phase (ages 1 to 3), and the phallic phase (ages 2 to 6).

infant psychiatry That aspect of child psychiatry concerned with the diagnosis, treatment, and prevention of maladaptive psychological functioning in infants.

inferential statistics Methods used for drawing general conclusions about probabilities on the basis of a sample. *See also* statistical inference.

inferiority complex Concept, originated by Alfred Adler, that everyone is born with a feeling of inferiority or inadequacy secondary to real or fantasied organic or psychological deficits; how the inferiority or feeling of inferiority is handled determines a person's behavior in life. *See also* masculine protest.

informal admission As a general hospital model, a patient admitted to a psychiatric unit is free to enter and to leave, even against medical advice.

information theory System dealing with the transmission, reception, and distortion of communicated messages.

informed consent Consent given by a patient to a physician to perform a procedure legally requires that the patient be informed of the nature of the proposed procedure, its risks, its proposed benefits, and the available alternatives; the major problem in giving informed consent, legally and psychologically, is the extent or depth of disclosure.

infradian rhythm *See* biological rhythm.

infusion Introduction of fluid other than blood, for example, medicines, into a vein.

inhalant Substance taken in by breathing. Production of psychoactive effects by the ingestion of such volatile substances as gasoline, airplane glue, aerosols (especially spray paints), or rubber cement, warrants the DSM diagnosis of inhalant dependence or inhalant abuse.

inhibited female orgasm Recurrent and persistent inhibition (absence) of the female orgasm, referring to the inability of a female to achieve orgasm by masturbation or coitus. Also called anorgasmia.

inhibited male orgasm Difficulty or inability of a male in achieving ejaculation during coitus. Also called ejaculatory incompetence; retarded ejaculation. *See also* primary retarded ejaculation; secondary retarded ejaculation.

inhibited sexual excitement Former term for female sexual arousal disorder.

inhibition Limitation or renunciation of ego functions occurring consciously, alone or in combinations, to evade anxiety arising out of conflict with instinctual impulses, the superego, or environmental forces or figures.

initial insomnia Falling asleep with difficulty; usually seen in anxiety disorder. *Compare* middle insomnia; terminal insomnia.

insane Of or pertaining to one who is of unsound mind. A legal rather than a psychiatric term.

insanity Legal concept denoting a mental disturbance, due to which a person lacks criminal responsibility for an alleged crime and hence cannot be convicted of the crime. *See also* American Law Institute formulation of insanity; Durham rule; M'Naghten rule.

insanity defense Criminal law defense plea that a person lacks criminal responsibility by reason of insanity.

insecurity Feelings of helplessness, unprotectedness, and inadequacy in the face of manifold anxieties arising from uncertainty regarding one's goals, ideals, abilities, and relations to others.

insight Conscious recognition of one's own condition. In psychiatry, it more specifically refers to the conscious awareness and understanding of one's own psychodynamics and symptoms of maladaptive behavior; highly important in effecting changes in the personality and behavior of a person.

insight-oriented psychotherapy Psychotherapy based on the principles of psychoanalysis in which persons develop a conscious awareness and understanding of their own psychodynamics and symptoms of maladaptive behavior. There is greater emphasis on day-to-day reality issues and a lesser emphasis on the development of transference issues than in psychoanalysis.

insight therapy *See* insight-oriented psychotherapy.

insomnia disorder Difficulty in falling asleep or difficulty in staying asleep. It can be related to another mental disorder (nonorganic), can be related to a known organic factor (such as a physical disorder or a medication), or can be primary (not related to a known organic factor or to another mental disorder). *See also* initial insomnia; middle insomnia; terminal insomnia.

instinct Basic inborn urge or drive. Freud postulated the existence of two opposing primal instincts: a life instinct (eros) and a death instinct (thanatos). An extensive array of human instincts has been proposed, including possessive instinct, mastery instinct, and herd or social instinct. Because of its implication of a fixed, essentially unalterable, largely hereditary

response or psychic tendency that does not involve learning or reason, the term has become ambiguous and controversial when applied to human behavior. Many modern psychoanalysts prefer the term "drive" for what Freud termed "instinct." *See also* drive.

institutionalization Process of placing a patient in a facility for extended care, for example, a mental hospital or nursing home.

instrumental conditioning OPERANT CONDITIONING. *See* conditioning.

insulin coma therapy (ICT) Type of biological therapy in which large amounts of insulin are given to a psychotic patient, usually a schizophrenic, producing profound hypoglycemia and resulting in a coma; introduced by Manfred Sakel in 1933, it declined in use drastically after the introduction of antipsychotic drugs. An occasionally used early stage of this treatment is called subcoma therapy.

intake Initial interview between a patient and a member of a psychiatric team, a term usually used in connection with admission to a mental health facility.

integration Constructive organizing, absorbing, and incorporating into the personality of data and experience; also refers to an amalgamation of separate parts into a functioning integral whole.

intellectual disability (ID) or intellectual developmental disorder Below average general intellectual functioning with concurrent impairment in adaptive behaviors, such as communication, self-care, and interpersonal skills, which manifests before adulthood. In DSM-5, the degree of ID is classified as mild, moderate, severe, or profound based on overall level of functioning. Previously called mental retardation.

intellectual insight Knowledge of the reality of a situation without the ability to successfully use that knowledge to effect an adaptive change in behavior. *Compare* true insight.

intellectualization Unconscious defense mechanism in which reasoning or logic is used in an attempt to avoid confrontation with an objectionable impulse and thus defend against anxiety. Also called brooding compulsion; thinking compulsion.

intellectual subaverage functioning I.Q. of more than two standard deviations below the test mean obtained on an intelligence test.

intelligence Capacity for learning and ability to recall, integrate constructively, and apply what one has learned; the capacity to understand and to think rationally.

intelligence quotient (I.Q.) Numerical measure of mental capability determined by dividing a mental age (M.A.) score achieved on a specific test, such as the Stanford–Binet Test, by the patient's chronological age (C.A.) and multiplying by 100.

intermission In psychiatry, the interval between attacks of a particular syndrome; when it is not certain that the symptoms will return, the interval is called a remission.

intermittent explosive disorder Condition in which a person has discrete episodes of losing control of aggressive impulses resulting in serious assault or destruction of property; often caused by a disorder of brain physiology, particularly in the limbic system.

intermittent reinforcement Any schedule of reinforcement in which the response is not always reinforced.

***International Classification of Diseases* (ICD)** World Health Organization's official list of disease categories subscribed to by all WHO member nations. ICD-CM is ICD's clinical modification, first published in 1919.

internship First year of clinical training following medical school, now called postgraduate year I (PGY-I), which progresses onto residency training in a specialty area.

interobserver reliability Degree of concurrence or disagreement among observers of a single event.

interpersonal conflict EXTRAPSYCHIC CONFLICT.

interpersonal psychiatry Dynamic-cultural system of psychiatry based on Harry Stack Sullivan's interpersonal theory, which proposes that each person must be viewed as an entity interacting with the sociocultural and interpersonal environment; emphasis is on interactive experiences.

interpersonal psychotherapy A type of brief psychotherapy used primarily for the treatment of depression. In a course of 10 to 12 sessions the therapist and patient focus of current problems with person in the environment which increases the patient's perception of self-defeating behaviors that contribute to low self-esteem and depression.

interpersonal skill Effectiveness of adaptive behavior in relation to other persons; ability to express feelings appropriately, to be socially responsible and responsive, and to work in harmony with others.

interpretation The formulation or description by the therapist to the patient of the therapist's understanding of the symbolic significance or meaning of any aspect of the patient's intrapsychic processes, behavior, or verbal productions; an attempt is made to construct into a more meaningful form the patient's resistances, defenses, transferences, and other symbolic activities, for example, dreams, fantasies.

Interpretation of Dreams The Sigmund Freud's major presentation (1899) not only of his theories about the meaning of dreams, a subject hitherto regarded as outside scientific interest, but also of his concept of a mental apparatus topographically divided into unconscious, preconscious, and conscious areas.

interview Person-to-person interaction in which information about a subject is collected. *See also* structured interview; unstructured interview.

intoxication Organic mental disorder due to the recent ingestion or presence in the body of an exogenous substance producing maladaptive behavior by virtue of its effects on the central nervous system. The most common psychiatric changes involve disturbances of perception, wakefulness, attention, thinking, judgment, emotional control, and psychomotor behavior; the specific clinical picture depends on the nature of the substance ingested. For types of intoxication, see the specific term.

intrapersonal conflict INTRAPSYCHIC CONFLICT.

intrapsychic Referring to what originates, takes place, or is situated within the psyche or mind.

intrapsychic ataxia *See* ataxia.

intrapsychic conflict State of tension arising from the clash of two or more incompatible or opposing forces; for example, wishes, needs, motives, thoughts operating within oneself. Also called intrapersonal conflict. *Compare* extrapsychic conflict.

intravenous Denoting introduction of a substance into a vein or veins. *See also* infusion.

introjection Unconscious, symbolic internalization of a psychic representation of a hated or loved external object with the goal of establishing closeness to and constant presence of the

object; considered an immature defense mechanism. In the case of a loved object, anxiety consequent to separation or tension arising out of ambivalence toward the object is diminished; in the case of a feared or hated object, internalization of its malicious or aggressive characteristics serves to avoid anxiety by symbolically putting those characteristics under one's own control. *See also* incorporation.

intropunitive Turning anger inward toward oneself; commonly observed in depressed patients.

introspection Contemplating one's own mental processes so as to achieve insight.

introversion State in which a person's energies are directed inward toward the self, with little or no interest in the external world. *Compare* extroversion.

introverted personality disorder SCHIZOID PERSONALITY DISORDER.

inversion Sigmund Freud's term for homosexuality, in which he distinguished three types: absolute, amphigenous, and occasional.

involuntary admission Hospitalization of a person against his or her will. Following an application by a friend or relative for the admittance of a person who does not recognize the need for hospital care, certification by two physicians is required, which allows the patient to be hospitalized for 60 days. Patients have access to legal counsel and, if brought to court, a judge can order the patient's release. After 60 days, the case must be reviewed by a board (of psychiatrists, physicians, lawyers, and citizens not affiliated with the hospital). *See also* parens patriae.

involuntary discharge Termination of hospitalization against a patient's will.

involuntary treatment Care of a person, who has a mental disorder that constitutes a danger to self or others, against that person's will.

involutional melancholia Depression occurring in late middle-age in persons who generally have no history of previous mental illness. Characteristic manifestations include delusions of sin, guilt, or poverty; an obsession with death; preoccupation with somatic (particularly gastrointestinal) function; despair, dejection, agitation, anxiety, and insomnia; and, in some cases, paranoid ideation. The term is not used in DSM-5, but is replaced by the diagnosis of major depressive episode, melancholic type. Also called involutional psychosis.

involutional psychosis *See* involutional melancholia.

I.Q. Intelligence quotient.

irregular sleep phase circadian rhythm disorder A circadian rhythm sleep–wake disorder characterized by a fragmented, disorganized, or unpredictable sleep–wake pattern.

irrelevant answer Answer that is not responsive to the question.

irresistible impulse Concept that a person charged with a criminal offense is not responsible for his or her act if the act was committed under an impulse that the person was unable to resist because of mental disease. The courts have chosen to interpret this law in such a way that it has been called the "policeman-at-the-elbow" law; in other words, the court will grant the impulse to be irresistible only if it determines that the accused would have gone ahead with the act even if he or she had had a policeman at his or her elbow.

irritability Abnormal or excessive excitability, with easily triggered anger, annoyance, or impatience.

irritable mood State in which one is easily annoyed and provoked to anger. *See also*

irritability.

isolated explosive disorder Disorder of impulse control in which a person has a single episode characterized by a failure to resist a violent impulse against others.

isolation In psychoanalysis, a defense mechanism involving the separation of an idea or memory from its attached feeling tone; unacceptable ideational content is thereby rendered free of its disturbing or unpleasant emotional charge. *Compare* relatedness. *See also* alienation.

isophilic Term used by Harry Stack Sullivan to mean liking or feeling affectionate toward people of the same sex, without the sexual or erotic aspects of homosexuality.

I-Thou Philosopher Martin Buber's conception that a person's identity develops from the true sharing by persons; basic trust can occur in a living partnership in which each member identifies the particular real personality of the other in his or her wholeness, unity, and uniqueness. In groups, I-Thou relationships promote warmth, cohesiveness, and constructive group process. *Compare* I-It.

J

jacksonian epilepsy Epilepsy characterized by recurrent episodes of focal motor seizures that usually begin with a highly localized tonic contraction or clonic movement which increases in severity and spreads progressively, often involving the entire body and terminating in a characteristic generalized convulsion with loss of consciousness; implies organic cerebral disease.

jamais vu Paramnestic phenomenon characterized by a false feeling of unfamiliarity with a real situation that one has experienced.

James–Lange theory of emotion Theory that the experience of emotion is preceded by a visceral or motor response, for example, first the person runs and then feels fear.

Janet, Pierre (1859–1947) Last great representative of the French school of psychiatry, known for his concept of psychological automatism and for his interest in cases of multiple personalities; first used the term "la belle indifférence."

jargon aphasia Aphasia in which the words produced are neologistic, that is, nonsense words created by the patient.

JCAHO Joint Commission on Accreditation of Healthcare Organizations. Formerly JCAH, Joint Commission on Accreditation of Hospitals.

JCMHC Joint Commission on Mental Health of Children.

Joint Commission on Accreditation of Healthcare Organizations (JCAHO) Group of agencies that influence the standards and performance of hospitals and health organizations; they inspect hospitals every 2 years and determine requirements for hospital accreditation (on which hospital reimbursements from Medicare and Medicaid are contingent).

Joint Commission on Mental Health of Children (JCMHC) Multidisciplinary agency authorized by U.S. Congress in 1965 to monitor the treatment of children with emotional illness.

Joint Commission on Mental Illness and Health Multidisciplinary study group established in 1955 whose "Action for Mental Health" (1961) was instrumental in the legislation and federal funding for community mental health centers; with the passage of the Community

Mental Health Centers Act in 1963, deinstitutionalization became a reality.

Jones, Ernest (1879–1958) Welsh psychoanalyst, and one of Freud's early followers, who was an organizer of the American Psychoanalytic Association (1911) and of the British Psychoanalytical Society (1919) and a founder and long-time editor of the journal of the International Psychoanalytical Association; a prolific author most noted for his three-volume biography of Sigmund Freud.

judgment Mental act of comparing or evaluating choices within the framework of a given set of values for the purpose of electing a course of action. If the course of action chosen is consonant with reality or with mature adult standards of behavior, judgment is said to be intact or normal; judgment is said to be impaired if the chosen course of action is frankly maladaptive, results from impulsive decisions based on the need for immediate gratification, or is otherwise not consistent with reality as measured by mature adult standards.

Jung, Carl Gustav (1875–1961) Swiss psychiatrist and psychoanalyst, originally associated with Sigmund Freud, who later founded the school of analytic psychology.

justice In the context of mental health, a fair distribution and application of psychiatric procedures.

K

Kanner syndrome INFANTILE AUTISM.

kindling effect Electrophysiological process in which repeated subthreshold stimulation of a neuron eventually generates an action potential; at the organ level, repeated subthreshold stimulation of an area of the brain results in the generation of a seizure.

kinesiology Study of body movement, especially in the context of its communicative function.

kinesthetic hallucination Hallucination primarily involving bodily movement.

kinesthetic sense Sense by which muscular motion, tension, position, and posture are perceived. Also called proprioception.

Kirkbride, Thomas S. (1809–1883) American psychiatrist, one of the 13 original founders of the American Psychiatric Association, noted for his 1854 manual advocating reform in the design of institutions for the mentally ill.

Klein, Melanie (1882–1960) British psychoanalyst and child analyst whose theories of early development departed from orthodox freudian thought.

Kleine–Levin syndrome Condition characterized by periodic episodes of hypersomnia and bulimia.

kleptomania Pathological compulsion to steal.

Klinefelter syndrome Chromosomal anomaly in which there is an extra X chromosome (karyotype 47-XXY). Affected persons are male in development, with small firm testes, eunuchoid habitués, variable gynecomastia and other signs of androgen deficiency, and elevated gonadotropin levels; there is an increased frequency of mental retardation (about 25 percent) and other forms of psychopathology, particularly antisocial behavior and delinquency. *See also* Turner syndrome.

Klüver–Bucy syndrome Behavioral syndrome, first described in monkeys after bilateral temporal lobe ablation, characterized by increased oral, sexual, and aggressive behavior; loss of visual recognition: memory defect; and hyperreactivity to visual stimuli.

Kohut, Heinz (1913–1981) Important innovator in psychoanalysis, who originated the psychoanalytic school of self-psychology, which requires that patients become aware of their excessive needs for approval and self-gratification.

koro Acute delusional syndrome, seen in Southeast Asia and southern China, in which the patient suddenly becomes acutely anxious that his penis is shrinking and may disappear into his abdomen, in which case he will die.

Korsakoff psychosis Mental disorder seen in long-term alcoholics. Its major characteristic feature is profound memory impairment, particularly for recent events, for which the individual attempts to compensate by confabulation. Called alcohol amnestic disorder in DSM. *See also* Wernicke fluent encephalopathy.

Kraepelin, Emil (1865–1926) German psychiatrist noted for his pioneering work in psychiatric nosology and classification systems, and who differentiated between manic–depressive psychoses and dementia precox (schizophrenia). One of the last representatives of the predynamic school of psychiatry, he is often considered the father of descriptive psychiatry.

Kretschmer, Ernest (1888–1964) German psychiatrist noted for his theories of the relation of physique to character and personality.

kuru A prion disease, seen in Papua, New Guinea, caused by cannibalistic funeral rituals.

L

la belle indifférence Inappropriate attitude of calm or lack of concern about one's disability; seen in patients with conversion disorder.

labile Unstable; characterized by rapidly changing emotions.

labile affect Affective expression characterized by repetitious and abrupt changes, unrelated to external stimuli.

labile mood Oscillations in mood between euphoria and depression or anxiety.

lability *See* emotional lability.

laconic speech Condition characterized by a reduction in the quantity of spontaneous speech; replies to questions are brief and unelaborated, and little or no unprompted additional information is provided. Occurs in major depression, schizophrenia, and organic mental disorders. Also called poverty of speech.

lacunar amnesia LOCALIZED AMNESIA.

lalophobia Abnormal fear of speaking. Also called glossophobia.

language disorder Persistent difficulties in the acquisition and use of language across modalities due to deficiencies in production (expressive) or comprehension (receptive-expressive) that include reduced vocabulary, difficulty with sentence formation, and impairments in discourse. Disorders lead to functional limitations in effective communication, social participation, and/or occupational performance. *See also* communication disorder.

lanugo Neonatal-like body hair; may be noted in anorexia nervosa.

lapsus linguae Slip of the tongue.

latah Culture-specific disorder found among Malaysians and characterized by either a sudden onset of unusual and inappropriate motor and verbal manifestations or by an echo reaction, in which the victim is compelled to imitate any words or actions to which he or she is exposed; in both forms affected persons cannot control or inhibit their behavior.

late language emergence DSM-5 disorder of early childhood characterized by delayed language onset in the absence of any developmental disabilities or delays.

late luteal phase dysphoric disorder (LLPDD) *See* premenstrual dysphoric disorder.

latency phase Period of psychosexual development after the phallic phase and succeeded by the genital phase, extending from about age 5 to the beginning of adolescence at age 12; during this phase there is an apparent cessation of sexual preoccupation and a blockade of libidinal impulses, and boys and girls are inclined to choose friends and join groups of their own sex.

latent content Hidden or unconscious meaning of symbolic representations, especially in fantasies and dreams. *See also* manifest content.

latent homosexuality Unexpressed conscious or unconscious homoerotic wishes that are held in check. Sigmund Freud's theory of bisexuality postulated the existence of a constitutionally determined, although experientially influenced, instinctual masculine-feminine duality; normally, the opposite-sex component is dormant, but a breakdown in the defenses of repression and sublimation may activate latent instincts and result in overt homoeroticism.

latent schizophrenia Condition characterized by clear schizophrenic symptoms without a history of prior overtly psychotic schizophrenic episodes. In DSM the condition is subsumed under the category of schizotypal personality disorder.

learned autonomic control Learned regulation by a person of physiological responses that are under autonomic nervous system control. *See also* biofeedback.

learned helplessness *See* helplessness.

learning disorders Disorders manifested in school-age children and characterized by specific difficulties in learning that are unrelated to intelligence. In DSM, disorders include reading disorder, dyscalculia, and disorder of written expression. *See also* dyslexia.

learning theory Approach to the understanding of human behavior that emphasizes the way in which learning comes about. Learning itself represents a change in behavior resulting from practice; through the application of certain laws of learning, learning theory attempts to explain the basic processes that are necessary for learning to occur.

lesbianism Female homosexuality. Also called sapphism.

Lesch–Nyhan syndrome A rare, X-linked disorder caused by a deficiency of an enzyme involved in purine metabolism. Characteristics include intellectual disability, seizures, self-mutilating behaviors (e.g., head banging, finger biting), microcephaly, choreoathetosis, and spasticity.

lethologica Momentary forgetting of a name or proper noun. *See* blocking.

leukotomy LOBOTOMY.

Lewin, Kurt (1890–1947) German psychologist who emigrated to United States in 1933. Two of his chief theoretical contributions are the field approach (field theory) and group dynamics, each of which has been useful in the experimental study of human behavior in a social situation.

Lewy body disease A type of dementia clinically similar to Alzheimer disease that is also accompanied by hallucinations and parkinsonian features. Often associated with Capgras' syndrome. *See also* dementia of the Alzheimer type.

Liaison Committee on Medical Education (LCME) Group sponsored by the American

Medical Association and the Association of American Medical Colleges that is principally involved in accrediting activities in undergraduate medical education.

liaison nursing Communication among nurses from various disciplines concerning the treatment of medical, surgical, and neuropsychiatric patients.

liaison psychiatry *See* consultation-liaison psychiatry.

libido In psychoanalysis, the psychic energy associated with the sexual drive (life instinct).

life instinct SEXUAL DRIVE.

lifetime expectancy Total probability of a person developing a disorder during a lifetime.

lifetime prevalence Measure at a point in time of the number of persons who had a given disorder at some time during their lives.

lilliputian hallucination Visual sensation that persons or objects are reduced in size, more properly regarded as an illusion. *See also* micropsia.

limbic system James Papez' proposal (1939) that a reverberating circuit or system, consisting of the hippocampus, hypothalamus, anterior thalamus, and cingulate gyrus, was the central nervous system localization for emotions; current models also include the amygdala and septal area, as well as other neuroanatomic areas. Various parts of the limbic system have been associated with emotions, sex drive, eating behavior, rage, violence, memory, and motivation.

lithium carbonate Lithium salt effective in the treatment of the manic phase of bipolar disorder and as a mood-stabilizing drug for chronic manic–depressive patients; may also be an effective antidepressant in some patients.

Little Albert Eleven-month-old boy who was conditioned to fear furry objects in experiments conducted by American psychologist John B. Watson (1920) to demonstrate that phobias are learned responses based on classical conditioning.

Little Hans Sigmund Freud's famous case history of a 5-year-old boy with a phobia. The boy's fear of horses, instead of his father, helped him to avoid hating his father, by whom he was threatened; and whom he also loved.

living will Legal document in which patients give instructions to their physicians about withholding life support measures, for example, Do Not Resuscitate (DNR) if in extremis.

LCME Liaison Committee on Medical Education.

LLPDD Late luteal phase dysphoric disorder.

LNNB Luria-Nebraska Neuropsychological Battery.

lobotomy Neurosurgical procedure, introduced by Egas Moniz, in which one or more nerve tracts in a lobe of the cerebrum are severed. Prefrontal lobotomy is the ablation of one or more nerve tracts in the prefrontal area of the brain and has been used in the treatment of certain severe mental disorders that do not respond to other treatments. Also called leukotomy.

localized amnesia Partial loss of memory; amnesia restricted to specific or isolated experiences. Also called lacunar amnesia; patch amnesia.

locus ceruleus Nucleus (a collection of neuronal cell bodies) that is located in the pons (a region of the brain stem) and contains the largest population of noradrenergic neurons in a single identified brain region; these neurons project their axons to virtually every region of the brain and are thought to be involved in the regulation of mood, anxiety, and attention.

logorrhea Copious, pressured, coherent speech; uncontrollable, excessive talking; observed in manic episodes of bipolar disorder. Also called tachylogia; verbomania; volubility.

logotherapy Existential analysis based on spiritual values, rather than psychobiological laws.

longitudinal study Examination or analysis carried out of a long period of time.

long-term memory Reproduction, recognition, or recall of experiences or information that were experienced in the distant past. Also called remote memory. *Compare* immediate memory; short-term memory.

loosening of associations Characteristic schizophrenic thinking or speech disturbance involving a disorder in the logical progression of thoughts, manifested as a failure to adequately verbally communicate; unrelated and unconnected ideas shift from one subject to another. *See also* tangentiality.

Lou Gehrig disease AMYOTROPHIC LATERAL SCLEROSIS.

LP Lumbar puncture.

LSD Lysergic acid diethylamide.

Lumbar puncture (LP) Use of a needle inserted into the lumbar spine intervertebral space to obtain cerebrospinal fluid for diagnostic purposes or to introduce local anesthetics or other drugs.

Luria-Nebraska Neuropsychological Battery (LNNB) Based on the work of the Russian psychologist A. R. Luria, the test assesses a wide range of cognitive functions: memory; motor functions; rhythm; tactile, auditory, and visual functions; receptive and expressive speech; writing; spelling; reading; and arithmetic. It is extremely sensitive for identifying specific types of problems (e.g., dyslexia, dyscalculia) rather than being limited to more global impressions of brain dysfunction; it also helps localize the various cortical zones that are involved in a particular function and is useful in establishing left or right cerebral dominance. The test is designed for persons at least 15 years of age; a children's version for use with 8 to l2 year olds is being developed.

lust dynamism Harry Stack Sullivan's term to describe clearly stated sexual desires and abilities.

lying Making statements that one knows consciously are false with intent to deceive; differentiated from confabulation, in which one is not conscious of lying.

lysergic acid diethylamide (LSD) A potent psychotogenic drug discovered in 1942 that produces psychotic-like symptoms and behavior changes, including hallucinations, delusions, and time-space distortions. Commonly called acid.

M

macropsia False perception that objects are larger than they really are. *Compare* micropsia.

magical thinking A form of dereistic thought; thinking similar to that of the preoperational phase in children (Jean Piaget), in which thoughts, words, or actions assume power (e.g., to cause or to prevent events). Common in dreams, certain mental disorders, and children.

magnetic resonance imaging (MRI) Diagnostic modality in which a patient's body is placed in a magnetic field and its hydrogen nuclei are excited by radiofrequency pulses at angles to the field's axis; resulting signals from the body's hydrogen ions are processed through a computer to produce an image; an excellent method for differentiating gray and white matter, visualizing cerebrospinal fluid, and estimating the size of ventricles.

Mahler, Margaret (1897–1985) Hungarian-born psychiatrist and child analyst who described

the separation-individuation process in children.

maintenance drug therapy Stage in the course of chemotherapy at which, after the drug has reached its maximal efficacy, the dosage is reduced and sustained at the minimal therapeutic level that will prevent a relapse or exacerbation.

major affective disorder *See* major depression.

major depressive disorder Severe mood disorder characterized by one or more depressive episodes but no history of a manic episode; a DSM-5 diagnosis. *See also* mood disorder.

major depressive episode, melancholic type DSM diagnosis replacing involutional melancholia.

major epilepsy *See* grand mal epilepsy.

major motor seizure *See* grand mal epilepsy.

major neurocognitive disorder DSM-5 term for dementia.

major tranquilizer ANTIPSYCHOTIC.

maladaptive Referring to any mental activity or behavior that is dysfunctional or counterproductive with regard to a person's ability to cope effectively with the problems and stresses of life.

maladjustment Maladaptive reaction; poor, faulty, or inadequate coping ability.

male erectile disorder DSM sexual arousal disorder characterized by the recurrent and persistent partial or complete failure to obtain or maintain penile erection until the completion of the sex act. Also called erectile dysfunction. *See also* impotence.

male hypoactive sexual desire disorder Disorder characterized by an absence of or reduction in sexual fantasies and/or the desire for sexual activity in men

maleness Anatomic and physiological features relating to the male's procreative functions.

malingering Feigning disease in order to achieve a specific goal, for example, to avoid an unpleasant responsibility. *Compare* factitious disorder.

malpractice Professional negligence including intentional or willful invasion of another's legally protected interests, such as battery or treatment without consent; must be demonstrated by the presence of *d*ereliction of a *d*uty *d*irectly causing *d*amages (the four D's of malpractice).

mania Mood abnormality characterized by elation; agitation; hyperactivity and hyperexcitability; accelerated thinking and speaking (flight of ideas); decreased need for sleep; and overinvolvement in pleasurable behavior. Characterizes the manic phase of bipolar I disorder when lasting more than a week or is severe enough to warrant hospitalization. *Compare* hypomania.

manic Denoting or characterized by mania. Sometimes used as a general description for a mentally disturbed person, often someone exhibiting violent or grossly irrational behavior.

manic–depressive illness Mood disorder characterized by severe alterations in mood that are usually episodic and recurrent. The unipolar type is characterized by either periodic episodes of mania (manic disorder) or periodic episodes of depression (depressive disorder). In the bipolar or circular type, the patient has at least one episode of each mood extreme and may alternate periodically between the two. Manic–depressive illness has been replaced by bipolar disorder in DSM.

manifest content That part of a dream or fantasy that a person remembers and reports; in a

dream, it represents a disguised and symbolic expression of the latent content of the dream.

manipulation Exploitativeness; the maneuvering by patients to get their own way, characteristic of antisocial personalities.

mannerism Stereotyped gesture or expression that is peculiar to a given person.

mantra In transcendental meditation, a Sanskrit syllable or word that is repeated over and over in an effort to produce total relaxation and control over one's state of consciousness.

MAOI Monoamine oxidase inhibitor.

marasmus Progressive emaciation, especially in malnourished infants and children, in which weight is often below the third percentile and markedly below appropriate weight for height.

marathon Technique developed by George Bach and Frederick Stoller that incorporates a group meeting that usually lasts 8 to 72 hours (although some may last for a week), in which the sessions interrupted only for eating and sleeping; the leader works for the development of intimacy and the open expression of feelings. The time-extended group experience culminates in intense feelings of excitement and elation.

marijuana Dried leaves and flowers of *Cannabis sativa* (Indian hemp), which induce somatic and psychic changes in a person when smoked or ingested in sufficient quantity. Somatic changes include increased heart rate, rise in blood pressure, dryness of the mouth, and increased appetite; psychic changes include a dreamy-state level of consciousness, a sense of enhanced vividness of visual and auditory sensations, disturbances of time perception and memory-dependent, goal-directed behavior, disruptions of sequential thought processes, and euphoria and other alterations of mood. The compound believed to be responsible for most of its psychological effects is tetrahydrocannabinol (THC). In strong doses, marijuana can produce hallucinations and, at times, paranoid ideation. Also called pot; grass; weed; dope; Mary Jane.

marital counseling Process whereby a trained counselor helps married couples resolve problems that arise and trouble them in their relationship; the husband and the wife are seen by the same worker in separate and joint counseling sessions that focus on immediate family problems. The theory and techniques of this approach were first developed in social agencies as part of family case work.

marital problems Maladaptive behavior between spouses in various areas, for example, attitudes toward sex, contraception, child rearing, money handling.

marital therapy *See* marriage therapy.

marriage therapy Type of family therapy that involves the husband and the wife and focuses on the marital relationship as it affects the individual psychopathology of the partners. The rationale for the method is the assumption that psychopathological processes within the family structure and in the social matrix of the marriage perpetuate individual pathological personality structures, which find expression in the disturbed marriage and are aggravated by the feedback between partners.

masculine identity Inner sense of gender affiliation with males. *Compare* feminine identity.

masculine protest Adlerian doctrine that depicts a universal human tendency to move from a passive and feminine role to a masculine and active role; an extension of Alfred Adler's ideas about organic inferiority that became the prime motivational force in normal and neurotic behavior in his system. *See also* Adler, Alfred; inferiority complex.

masculinity-femininity scale Any scale on a psychological test that assesses the relative masculinity or femininity of the testee. Scales vary and may focus, for example, on basic identification with either sex or preference for a particular sex role and are strongly influenced by cultural definitions of masculinity and femininity.

masochism Paraphilia in which sexual gratification is derived from being physically or psychologically maltreated by a partner or oneself; first described by an Austrian novelist, Leopold von Sacher-Masoch (1836–1895). *See also* sadism; sadomasochistic relationship.

masturbation Self-stimulation of the genitals for sexual pleasure. *See* autoerotism.

maternal deprivation Result of the premature loss or absence of the mother. In a broader sense, the lack of proper mothering may constitute a form of maternal deprivation which may lead to severe emotional disorders in infants and children.

maturational crises Normative stages of the life cycle that require adaptive coping or defense mechanisms.

maximum security unit That part of an institution for the mentally ill reserved or those who have committed crimes or who are considered dangerous to others.

MCE Medical care evaluation.

MDMA 3,4-Methylenedioxymethamphetamine.

mean Statistical measurement derived from adding a set of scores and then dividing by the number of scores. *See also* average.

mean deviation Statistical measure of the variability in a set of values defined as the sum of the absolute differences between the values and the mean divided by the number of values.

median Value in a set of values above and below which there are an equal number of values; for example, in the series 1, 3, 5, 9, 33 the value 5 is the median. *See also* average.

Medicaid Assistance program for certain needy and low-income persons, mandated by the federal government and financed by both federal and state governments; certain benefits, such as in- and outpatient hospital care, physicians' services, and laboratory tests, are required covered expenses, while other services may be provided at the state's option.

medical audit Retrospective assessment of the quality and utilization of care, including investigation of suspected problems, analysis of the problems identified, and a plan for corrective action. Also called medical care evaluation (MCE).

medical care evaluation (MCE) MEDICAL AUDIT.

medical ethics Moral set of principles that determine what is right or wrong and good or bad, as they pertain to professional responsibilities and opinions relating to the values of patients, patients' families, and society at large. For related issues, see the specific term.

medical record Written exposition describing a patient's diagnosis and treatment.

medical review Evaluation of the quality of care and assessment of medical care or services judged in terms of cost effectiveness.

Medicare Hospital and medical insurance system for the aged created in 1965 by amendments of the Social Security Act. Hospitalization in general hospitals for psychiatric conditions is covered in the same way as other conditions.

medication-induced movement disorders A group of disorders that include neuroleptic malignant syndrome, neuroleptic-induced parkinsonism, neuroleptic-induced acute akathisia, neuroleptic-induced tardive dyskinesia, neuroleptic-induced acute dystonia, and medication-

induced postural tremor.

megalomania Morbid preoccupation with expansive delusions of power, importance, or wealth.

megavitamin therapy *See* orthomolecular treatment.

melancholia A severe depressive state characterized by anhedonia, early morning awakening, weight loss, psychomotor retardation, and profound feelings of guilt and hopelessness; often associated with suicidal ideation.

memory Process whereby what is experienced or learned is established as a record in the central nervous system (registration), where it persists with a variable degree of permanence (retention) and can be recollected or retrieved from storage at will (recall). For types of memory, see the specific term.

menarche Onset of menstrual function.

mendacity Pathological lying.

mens rea Legal term meaning an intent to do harm.

mental aberration Pathological deviation from normal thinking; unrelated to a person's intelligence.

mental age Measure of mental ability as determined by standard psychological tests.

mental deficiency Term often used interchangeably with mental retardation, applied to individuals with an I.Q. of less than 70. *See* intellectual developmental disorder.

mental disorder, mental disease Psychiatric illness or disease whose manifestations are primarily characterized by behavioral or psychological impairment of function, measured in terms of deviation from some normative concept.

mental health State of emotional well-being in which persons are able to function comfortably within their society and in which their personal achievements and characteristics are satisfactory to them.

mental hygiene 1. Conditions and practices conducive to the establishment and maintenance of mental health. **2.** Field dealing with the prevention and early treatment of mental illness. *See also* orthopsychiatry.

mental illness Mental disorder; any serious impairment of adjustment; any psychiatric disorder listed in the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* or in the World Health Organization's *International Classification of Diseases*.

mental retardation Term previously assigned to the condition that DSM-5 now refers to as intellectual disability. Degrees of condition were historically classified according to IQ: mild (50–55 to 70), moderate (35–40 to 50–55), severe (20–25 to 30), and profound (below 20–25). *See* intellectual disability.

mental status General functional condition of mental and behavioral processes as determined by psychiatric assessment of a variety of areas of functioning, such as state of consciousness, mood and affect, thinking and speech, motor behavior, general knowledge, memory, calculation, judgment, abstraction, and insight.

mental status examination (MSE) Clinical evaluation, and the sum total of the examiner's observations and impressions, of a person's psychological and behavioral experiences. MSE includes specific references to such areas as attitude and general behavior during interview and

to sensorium, mental capacity, and other areas of function.

merycism RUMINATION DISORDER OF INFANCY.

mescaline Hallucinogenic drug obtained from buttons of the peyote cactus that has psychotomimetic effects similar to those of LSD.

mesmerism Early term for hypnosis.

mesomorphic Muscular, with a well-proportioned physique; one of Sheldon's constitutional types. *Compare* ectomorphic; endomorphic.

meta-analysis General consensus of many studies that results in one common data base that helps to eliminate bias.

metapsychiatry Branch of psychiatry concerned with altered state of consciousness, holistic health, and parapsychology.

metapsychology Branch of theoretical or speculative psychology concerned with theories, hypotheses, or phenomena that are largely beyond the realm of empirical verification. *See also* parapsychology.

methadone Orally active synthetic narcotic agent that is an effective analgesic with pharmacological actions qualitatively very similar to those of morphine. Its use in the maintenance treatment of heroin addicts is based on its extended duration of action in suppressing withdrawal symptoms in persons physically dependent on opiates.

methadone maintenance treatment *See* methadone.

3-methoxy-4-hydroxyphenylglycol (MHPG) Most significant metabolic product from central nervous system norepinephrine; excreted in urine.

3,4-methylenedioxymethamphetamine (MDMA) Hallucinogen more commonly known as ecstasy.

methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) Contaminant of an illicitly made synthetic heroin that induces irreversible parkinsonism.

metonymy Speech disturbance common in schizophrenia in which the affected person uses a word or phrase that is related to the proper one but is not the precise one that would ordinarily be used; for example, the patient speaks of consuming a "menu," rather than a "meal," or refers to losing the "piece of string" of the conversation, rather than the "thread" of the conversation. *See also* paraphasia; word approximation.

Metrazol treatment Obsolete form of convulsive therapy in which a convulsive seizure is induced by injection of the drug Metrazol (pentylenetetrazol).

Meyer, Adolf (1866–1950) American psychiatrist known for his concept of psychobiology and "commonsense" psychiatry.

MHPG 3-Methoxy-4-hydroxyphenylglycol.

microcephaly Condition in which the head is unusually small as a result of defective brain development and premature ossification of the skull.

micropsia False perception that objects are smaller than they really are. Sometimes called lilliputian hallucination. *Compare* macropsia.

MID Multi-infarct dementia.

middle age, middle adulthood That period considered to be the prime of life, between the ages of approximately 40 to 65 years, following the termination of early adulthood.

middle insomnia Waking up after falling asleep without difficulty and then having difficulty

in falling asleep again. *Compare* initial insomnia; terminal insomnia.

midlife crisis Syndrome that occurs in both sexes as they enter their early 40s or 50s, consisting of depression, substance abuse, concern about growing old, and fear of death; more common in adults with prior psychiatric illness than in normal adults.

migraine Severe unilateral, throbbing headache that appears periodically, usually beginning during the teenage years and continuing to recur with diminishing frequency during advancing years. Classic migraine is a syndrome ushered in by a transient derangement of neurological function, often a visual disturbance, followed by hemicranial headache, nausea, and vomiting, all of which last for anywhere from several hours to 1 or 2 days. Common migraine has the same character and time course but develops without any prodrome. Patients with migraine are typically described as tense, meticulous, hard-driving persons who appear to have a high frequency of associated emotional conflicts.

mild mental retardation *See* mental retardation.

mild neurocognitive disorder Mild decline in cognitive function that is differentiated from the normal cognitive changes associated with aging; memory is impaired for age and education, however, general cognitive function and activities of daily living remain intact. In DSM-5 diagnosed as mild neurocognitive disorder due to multiple etiologies or unspecified neurocognitive disorder.

milieu therapy Treatment that emphasizes appropriate socioenvironmental manipulation for the benefit of the patient; the setting is usually the psychiatric hospital. *See also* moral treatment; therapeutic community.

mimicry Simple, imitative motion activity of childhood.

minimal brain dysfunction Behavioral syndrome of childhood, now called attention-deficit hyperactivity disorder, characterized by learning difficulties, decreased attention span, distractibility, hyperactivity, impulsiveness, emotional lability, and, often, disturbances in perceptuomotor and language development; its psychopathological mechanisms have not been defined. The term implies neurological causation, but in most cases there are no major unequivocal neurological signs and there has been a tendency to apply the term as a convenient explanatory label to any child presenting with a specific learning difficulty or behavioral dysfunction. Although the syndrome has been diagnosed with increasing frequency over the past 20 years, its validity as a diagnostic entity has been questioned, and it is not listed as such in DSM.

Minnesota Multiphasic Personality Inventory (MMPI) Questionnaire type of psychological test for persons age 16 and over consisting of true–false statements that are coded in various scales that assess different dimensions of the person's personality structure and measure the closeness of fit with various psychiatric diagnostic categories.

minor epilepsy PETIT MAL EPILEPSY.

minor tranquilizer ANXIOLYTIC.

Mitchell, S. Weir (1830–1914) American neurologist known for his concept of rest treatment to cure emotional disorders.

mixoscopia Sexual perversion in which a person attains orgasm by watching his or her love object make love with another person.

MMPI Minnesota Multiphasic Personality Inventory.

M'Naghten rule Set of rules established in 1843 in England to determine criminal responsibility, based on the person's ability to distinguish right from wrong. *See also* Durham rule; American Law Institute formulation of insanity.

mode In a set of measurements, the value that appears most frequently. *See also* average.

modeling Learning of a new skill or behavior by observation and imitation of another person.

Model Penal Code American Law Institute's test of criminal responsibility: "A person is not responsible for criminal conduct if at the time of such conduct, as a result of mental disease or defect, he lacked substantial capacity to either appreciate the criminality (wrongfulness) of his conduct or to conform his conduct to the requirement of the law."

moderate mental retardation *See* mental retardation.

molecular psychiatry Etiological view of psychiatry based on neurotransmitter chemicals as measured in molecular quantities.

mongolism Obsolete term for Down syndrome.

monoamine Biogenic amine or class of neurotransmitters that consists of three catecholamines (dopamine, norepinephrine, and epinephrine), an indolamine (serotonin), a quaternary amine (acetylcholine), and an ethylamine (histamine).

monoamine oxidase (MAO) Enzyme that deaminates monoamines and that functions, in the nervous system by breaking down monoamine neurotransmitters (e.g., norepinephrine, epinephrine, dopamine, and serotonin); has been subdivided into Type A and Type B.

monoamine oxidase inhibitor (MAOI) Agent that inhibits the enzyme monoamine oxidase (MAO). MAO inhibitors are highly effective as antidepressants and are also used in the treatment of panic disorders and some eating disorders.

monomania Morbid mental state characterized by preoccupation with one subject. Also called partial insanity.

monosymptomatic hypochondriacal psychosis A delusional disorder consisting of somatic delusions; common delusions include infestation, dysmorphophobia, and foul body odors or halitosis.

monozygotic twins Twins who develop from a single fertilized ovum. In contrast, dizygotic or fraternal twins develop from two separately fertilized ova. Also called identical twins.

mood Pervasive and sustained feeling tone that is experienced internally and that, in the extreme, can markedly influence virtually all aspects of a person's behavior and his or her perception of the world. Distinguished from affect, the external expression of the internal feeling tone. For types of mood, see the specific term.

mood-congruent delusion Delusion with content that is mood appropriate, for example, a depressed patient who believes he or she is responsible for the destruction of the world.

mood-congruent hallucination Hallucination with content that is consistent with either a depressed or a manic mood, for example, a depressed patient hearing voices telling the patient he or she is a bad person; a manic patient hearing voices telling the patient he or she has inflated worth, power, or knowledge.

mood-congruent psychotic features DSM term that refers to hallucinatory or delusional phenomena with contents that consistently reflect the mood of a manic or depressed patient.

mood disorders Qualitative terms used to describe the striking symptomatic features or timing patterns in various patients with moods disorders. Descriptive qualifiers include

psychotic, melancholic, atypical, and catatonic features; patterns include postpartum onset, rapid cycling, and seasonal pattern.

mood-incongruent delusion Delusion based on incorrect reference about external reality, with content that has no association to mood or is mood-inappropriate, for example, a depressed patient who believes he or she is the new Messiah.

mood-incongruent hallucination Hallucination not associated with real external stimuli, with content that is not consistent with either depressed or manic mood, for example, in depression, hallucinations not involving such themes as guilt, deserved punishment, or inadequacy; in mania, not involving such themes as inflated worth or power.

mood-incongruent psychotic features DSM term that refers to hallucinatory or delusional phenomena with contents that do not consistently reflect the mood of a manic or depressed patient.

mood stabilizer Drug used to prevent mood swings in bipolar disorder.

mood swing Oscillation of a person's emotional feeling tone between periods of elation and periods of depression.

moral treatment Philosophy and technique of treating mental patients, developed in the 18th century, that involved the establishment of a humane therapeutic environment designed to maximize the chances for recovery; considered the forerunner of current milieu therapy.

morbid jealousy Marked or severe jealousy. Also called pathological jealousy.

morbid perplexity Condition seen in schizophrenia in association with loss of ego boundaries; the patient exhibits profound confusion about his own identity and the meaning of existence.

moron Obsolete and pejorative term for a classification of mental retardation in which an individual has a mental age of approximately 8 years.

morphine Major phenanthrene alkaloid opium derivative that produces a combination of central nervous system excitation and depression; repeated administration can lead to tolerance, physical dependence, and psychological dependence; used as an analgesic, sedative, and anxiolytic.

Moses and Monotheism Title of a book by Sigmund Freud (1939) in which he undertook a historical but frankly speculative reconstruction of the personality of Moses and examined the concept of monotheism and the abiding effect of the patriarch on the character of the Jews; one of Freud's last works, it bears the imprint of his latter-day outlook and problems.

mother surrogate In psychoanalysis, a mother substitute onto whom a patient projects his or her mother image and responds to unconsciously, in an inappropriate and unrealistic manner, with the feelings and attitudes that he or she had toward the real mother.

motivation Force or energy associated with an internal state that propels a person to engage in behavior to satisfy a need or desire.

motor aphasia Aphasia in which understanding is intact but the ability to speak is lost. Also called Broca's, expressive, or nonfluent aphasia.

motor disorders A class of neurodevelopmental disorders defined by a delay in meeting expected milestones for motor skills, based on age and intelligence, which significantly impairs daily functioning. The major types of motor disorders include developmental coordination disorder, stereotypic movement disorder, and tic disorder.

mourning Syndrome following loss of a loved one, consisting of preoccupation with the lost individual, weeping, sadness, and repeated reliving of memories. *See also* bereavement; grief.

MPD Multiple personality disorder.

MPTP Methyl-4-phenyl-1,2,3,6-tetrahydropyridine.

MRI Magnetic resonance imaging.

multi-infarct dementia *See* vascular dementia.

multiple delusions Concurrent delusions in which the delusions need not be interconnected.

multiple personality disorder *See* dissociative identity disorder.

multiple regression Multivariate analysis in which a scaled variable is correlated with a linear combination of independent or predictor variables.

multiple therapy CO-THERAPY.

multivariate analysis Methods for considering the relationship of three or more variables, including multiple regression, multivariate analysis of variance, discriminant analysis, canonical correlation, and factor analysis.

Munchausen syndrome Alternate term for factitious disorder.

muscarine Naturally occurring alkaloid found in certain species of poisonous mushrooms; symptoms of mushroom poisoning develop within 30 minutes of ingestion and consist of vomiting, diarrhea, bronchospasm, hypertension, delirium, and shock.

mutism Organic or functional absence of the faculty of speech. *See also* stupor.

mydriasis Dilation of the pupil; sometimes occurs as an autonomic (anticholinergic) side effect of some phenothiazines and tricyclic antidepressant drugs.

mysophobia Abnormal fear of dirt or contamination.

N

naloxone Opiate antagonist that, unlike nalorphine which it has largely replaced as a narcotic antagonist is essentially devoid of agonist effects. In opioid-dependent persons, small doses of naloxone precipitate moderate to severe withdrawal symptoms, an effect that can be used to diagnose physical dependence on narcotic drugs; however, the most important use of naloxone is in the treatment of narcotic overdosage. (Trade name, Narcan.)

NAMI National Alliance for the Mentally Ill.

narcissism Self-love linked to autoerotism but devoid of geniality. In psychoanalytic theory, it is divided into primary and secondary types: primary narcissism, the early infantile phase of object relationship development, when the child has not differentiated the self from the outside world and all sources of pleasure are unrealistically recognized as coming from within the self, giving the child a false sense of omnipotence; secondary narcissism, when the libido, once attached to external love objects, is redirected back to the self. *See also* autistic thinking.

narcissistic personality disorder Personality disorder characterized by a grandiose sense of self-importance, preoccupation with fantasies of unlimited success, exhibitionistic need for attention and admiration, exaggerated responses to criticism or other perceived threats to self-esteem, and disturbance in interpersonal relationships.

narcoanalysis NARCOTHERAPY.

narcolepsy Sleep disorder characterized by recurrent, brief, uncontrollable episodes of sleep; often associated with cataplexy. *See also* hypersomnia.

narcosis Drug-induced stupor.

narcosynthesis NARCOTHERAPY.

narcotherapy Psychotherapy conducted with the patient in a drug-induced stupor or semiconscious state, as by the administration of barbiturates; originally used in the treatment of acute disorders arising in the setting of military combat; also called narcoanalysis; narcosynthesis.

narcotic Any natural or synthetic drug producing morphine-like analgesia, drowsiness, mental clouding, and, with prolonged use, physical dependence. *See also* opioid.

narcotic blockade Inhibition of the euphoric effects of even high doses of intravenous opioids, such as heroin, by the use of other drugs, such as methadone.

National Alliance for the Mentally Ill (NAMI) Self-help group and influential national advocacy network comprised of and supported by the family members and parents of schizophrenic patients.

National Board of Medical Examiners (NBME) National board, founded in 1915, for testing medical students by three examinations given during different stages of their medical education to qualify for licensure to practice medicine in the United States; the diploma is accepted in 50 states for obtaining licensure without further examination by the state.

National Center for Complementary and Alternative Medicine (NCAAM) A government agency, part of the National Institute of Health that tests and evaluated alternative medicine therapies, for example, nutritional supplements, prayer, yoga, among many others.

National Institute of Alcohol Abuse and Alcoholism (NIAAA) *See* Alcohol, Drug Abuse, and Mental Health Administration.

National Institute of Drug Abuse (NIDA) *See* Alcohol, Drug Abuse, and Mental Health Administration.

National Institute of Mental Health (NIMH) *See* Alcohol, Drug Abuse, and Mental Health Administration.

National Mental Health Association (NMHA) Voluntary citizens' organization, founded in 1909 as the National Committee on Mental Hygiene, to promote mental health.

National Training Laboratories (NTL) Organization started in 1947 at Bethel, Maine, to train professionals who work with groups; interest in personal development eventually led to sensitivity training and encounter groups.

natural group Any group that tends to evolve spontaneously inhuman civilization, such as a kinship, tribal, or religious group in contrast to various contrived groups or aggregates of people who meet for a relatively brief time to achieve some goal.

NBME National Board of Medical Examiners.

NCAAM National Center for Complimentary and Alternative Medicine.

necromania Pathological preoccupation with dead bodies.

necrophilia Sexual attraction to or sexual contact with dead bodies.

necrophobia Abnormal fear of dead bodies.

need Ambiguous term referring to an internal state of unsatisfaction or tension related to a wish, urge, desire, or other endogenous behavioral stimulus; classically, it implied an innate or instinctual internal stimulus. *See also* drive; instinct; motivation.

needle phobia The persistent, intense, pathological fear of receiving an injection.

negative reinforcement *See* reinforcement.

negative symptoms In schizophrenia: flat affect, alogia, abulia, and apathy.

negative transference *See* transference.

negativism Verbal or nonverbal opposition or resistance to outside suggestions and advice; commonly seen in schizophrenia in which the patient resists any effort to be moved or does the opposite of what is asked.

neologism New word or phrase whose derivation cannot be understood; often seen in schizophrenia. It has also been used to mean a word that has been incorrectly constructed but whose origins are nonetheless understandable (e.g., "headshoe" mean "hat"), but such constructions are more properly referred to as word approximations.

nervous breakdown Nonspecific lay term for mental disorder, often implying an acute decompensation.

neurasthenia Condition characterized by vague functional physical and mental fatigue; not used in DSM.

neurochemistry Branch of chemistry concerned with nervous system components, metabolism, and function.

neurocirculatory asthenia DA COSTA SYNDROME.

neuroendocrinology Subspecialty area concerned with the anatomical and functional relationships between the nervous system and the endocrine apparatus. *See also* psychoendocrinology.

neurohormone Chemical messengers formed by neurosecretory cells and released into the bloodstream.

neuroleptic drug ANTIPSYCHOTIC.

neuroleptic malignant syndrome (NMS) Life-threatening complication of antipsychotic treatment, including such symptoms as muscle rigidity and dystonia, akinesia, mutism, obtundation, and agitation.

neurological amnesia 1. Auditory amnesia: loss of ability to comprehend sounds or speech. *See also* Wernicke aphasia. **2.** Tactile amnesia: loss of ability to judge the shape of objects by touch. *See also* astereognosis. **3.** Verbal amnesia: loss of ability to remember words. **4.** Visual amnesia: loss of ability to recall or recognize familiar objects or printed words.

neurological disorders Disorders involving the organization and function of nervous tissue, including diseases of the peripheral nerves, the spinal cord, and the brain, which are based on organic pathology.

neurologist Physician who specializes in diseases of the nervous system.

neurology Medical specialty concerned with organic diseases of the nervous system.

neuron Nerve cell, consisting of cell body, dendrites, and neuroaxon, which is the basic functional unit of the nervous system.

neuronal plasticity Ability of the central nervous system to alter its functional organization; a capacity for change that serves as a mechanism to shape neuronal circuitry during development, to change the way information is processed, and to compensate for injury.

neurophysiology Physiology of the nervous system.

neuropsychiatry Medical specialty that combines psychiatry and neurology.

neuroreceptors Neuronal and other nervous tissue sites that are stimulated or inhibited by

neurotransmitters and/or psychotropic drugs.

neuroscience Branch of the life sciences concerned with the anatomy, physiology, biochemistry, or molecular biology of nerves and nervous tissue and, especially, with their relation to behavior and learning.

neurosis Mental or psychiatric disorder characterized primarily by anxiety, which may either be directly experienced and dominate the clinical picture or be unconsciously controlled and modified by various psychological mechanisms to produce other subjectively distressing and ego-alien symptoms. Although neuroses are not accompanied by gross distortion of reality or severe personality disorganization, normal functioning is impaired by the person's symptoms; they have no organic basis, are relatively persistent, and are treatable. Also called psychoneurosis.

neurotic disorder *See* neurosis.

neurotic process Process in which unconscious conflict is expressed by neurotic symptoms or personality disturbance.

neurotoxin Agent, such as an antibody, causing destruction to nervous tissue.

neurotransmitter Any specific chemical agent released by a presynaptic cell that, upon excitation, crosses the synapse to stimulate or inhibit the postsynaptic cell.

NIAAA National Institute of Alcohol Abuse and Alcoholism.

nicotine dependence Inability to control the use of and dependence on tobacco. Nicotine dependence and withdrawal are classified as disorders in DSM.

NIDA National Institute of Drug Abuse.

night eating syndrome Eating disorder characterized by the consumption of large amounts of food after dinner; generally associated with insomnia and a lack of daytime appetite. In DSM-5, included in "Other Specified Feeding or Eating Disorders."

night hospital Part-time hospital facility in which patients function in the outside world during the day but return to the hospital at night. *Compare* day hospital; weekend hospital.

nightmare Anxiety attack experienced while having terrifying or frightening dreams; characterized by mild anxiety, good recall of the dream, and mild autonomic reactions. *Compare* nightmare disorder; sleep terror disorder.

nightmare disorder Sleep disorder characterized by frequent nightmares and associated insomnia resulting from a fear of sleep.

night-terror SLEEP TERROR DISORDER.

nihilism Delusion of the nonexistence of the self or part of the self; also refers to an attitude of total rejection of established values or extreme skepticism regarding moral and value judgments.

nihilistic delusion Delusional disorder in which patients believe the world and everything related to it have ceased to exist.

NIMH National Institute of Mental Health.

nitrogen narcosis RAPTURE-OF-THE-DEEP SYNDROME.

NMS Neuroleptic malignant syndrome.

noesis Revelation in which immense illumination occurs in association with a sense that one has been chosen to lead and command. Can occur in manic or dissociative states.

nominal aphasia Aphasia characterized by difficulty in giving the correct name of an object.

See also anomia; amnestic aphasia.

nonaffective hallucination Hallucination whose content is apparently unrelated to either depression or elation. *Compare* illusion.

noncomplementary role *See* role.

noncompliance with medical treatment DSM category used when a patient does not cooperate with the physician in his or her medical care; motivated by denial of illness, by religious beliefs, or by decisions based on personal value judgments about the advantages and disadvantages of the proposed treatment.

nondirective approach Technique in which the therapist follows the lead of the patient in the interview, rather than introducing his or her own theories and directing the course of the interview; applied in both individual and group therapy.

nonfluent aphasia MOTOR APHASIA.

nonparametric Denoting statistical methods that do not require restrictive assumptions about population distributions.

nonreactive depression ENDOGENOUS DEPRESSION.

non-suicidal self injury A DSM-5 disorder in which the patient intentionally inflicts damage to his or her body to induce bruises, bleeding, or pain without suicidal ideation. Also called self-
mutilation.

non–24-hour sleep–wake circadian rhythm disorder A sleep–wake disorder characterized by a circadian period that is not congruent with the external environment; common in the blind, visually impaired, and those with traumatic brain injury.

noradrenaline NOREPINEPHRINE.

norepinephrine Catecholamine neurotransmitter substance liberated by adrenergic postganglionic neurons of the sympathetic nervous system; the precursor of epinephrine present in the adrenal medulla and also found in many areas of the brain. A disturbance in its metabolism at important brain sites has been proposed as an important factor in the causation of mood disorders. Also called noradrenaline.

normality In psychiatry, the absence of mental illness; that state of mental health about which there exists a scientific and professional consensus. The normal person has a sense of his or her own worth and maintains a confident and purposeful self-image.

nosology Science of the classification of diseases.

nosophobia PATHOPHOBIA.

not-for-profit hospital Any hospital not maintained for the purpose of monetary gain. *Compare* for-profit hospital.

not otherwise specified (NOS) Term used throughout DSM to describe unusual or uncharacteristic variations of different mental disorders; for example, eating disorder NOS, tic disorder NOS, pervasive developmental disorder NOS.

NREM sleep Non-REM sleep consisting of four stages, based on EEG findings: stage 1: low amplitude, first frequency immediately after sleep begins; stage 2: cycles of 12 to 16 Hz called sleep spindles; stages 3 and 4: increased amplitude and slow frequency EEG changes. *Compare* REM sleep.

NTL National Training Laboratories.

nuclear family Immediate members of a family, including the parents and the children.

nuclear magnetic resonance (NMR) *See* magnetic resonance imaging.

null hypothesis Research term for the hypothetical assumption that any difference observed between two groups or conditions or between a particular group and the general population is purely accidental and due to chance alone. *See also* Type 1 error; Type 2 error.

nyctophobia Abnormal fear of darkness or of the night.

nymphomania Abnormal, excessive, insatiable desire in a female for sexual intercourse. *Compare* satyriasis.

O

objective tests Typically, pencil-and-paper tests based on specific items and questions and that yield numerical scores and profiles easily subjected to mathematical or statistical analysis; for example, Minnesota Multiphasic Personality Inventory (MMPI).

object permanence As described by Jean Piaget, the child's (birth to 2 years) ability to understand that objects have an existence independent of his or her involvement with the object; the child is able to maintain a mental image of a person or object even though it is not present and not visible.

object relation Emotional attachment formed between one person and another, as opposed to interest in and love for oneself.

OBS Organic brain syndrome.

OSA Obstructive sleep apnea syndrome.

obsession Persistent and recurrent idea, thought, or impulse that cannot be eliminated from consciousness by logic or reasoning; obsessions are involuntary and ego-dystonic. *See also* compulsion.

obsessional jealousy disorder Descriptive condition in which a person has persistent thoughts about infidelity in a partner or spouse; not a diagnosis in DSM-5.

obsessive-compulsive disorder (OCD) Neurotic disorder characterized by the persistent recurrence of obsessions and compulsions.

obsessive-compulsive personality Personality disorder characterized by perfectionism, overconscientiousness, and excessive inhibition with regard to self-expression and relaxation. Also called anancastic personality.

obstructive sleep apnea syndrome (OSA) A dyssomnia characterized by periods of functional obstruction of the upper airway during sleepy, resulting in degreases in oxygen, after which breathing (at least briefly) resumes normally.

occupational problem Maladaptive patterns related to work. For example, job dissatisfaction and uncertainties about career choices.

occupational psychiatry Area of psychiatry concerned with mental illness in industry, including the psychiatric aspects of absenteeism, vocational adjustment, operational fatigue, and accident proneness.

occupational therapy Form of therapy in which the patient engages in useful, purposeful activities in a therapeutic social setting involving interaction with other patients and hospital personnel.

OCD Obsessive-compulsive disorder.

oedipal phase PHALLIC PHASE. *See also* Oedipus complex.

Oedipus complex Constellation of feelings, impulses, and conflicts in the developing child that concern sexual impulses and attraction toward the opposite-sex parent and aggressive, hostile, or envious feelings toward the same-sex parent; real or fantasied threats from the same-sex parent result in the repression of those feelings. Development of the Oedipus complex coincides with the phallic phase of psychosexual development; as one of Freud's most important concepts, the term was originally applied only to the male. *Compare* Electra complex.

olfactory hallucination Hallucination primarily involving smell or odors; most common in medical disorders, especially in the temporal lobe.

Olfactory reference syndrome Alternate term for monosymptomatic hypochondriacal psychosis with delusions of foul body odor.

oligophrenia Old term for mental retardation.

OMD Organic mental disorder.

OMS Organic mental syndrome.

onanism Coitus interruptus. Sometimes used interchangeably with masturbation.

oneiroid state Dream-like state in which a patient is deeply perplexed and not fully oriented in time and place.

ontogenetic Pertaining to the development of an individual person. *Compare* phylogenetic.

open hospital Mental hospital without locked doors or physical restraints.

operant behavior *See* conditioning.

operant conditioning *See* conditioning.

operational definition Means of defining a word so that it is scientifically measurable.

opiate Morphine or other alkaloids derived from opium, an extractable product of the poppy plant *Papaver somniferum* or its variety, *P. album*. *See also* opioid.

opiate receptor Specific site on a neuronal cell surface that interacts with opiate drugs and mediates their pharmacological actions; presumably, it also mediates the physiological functions of endogenous opioids, such as the enkephalins and endorphins.

opioid Any chemical substance that has morphine-like pharmacological activity; for example, narcotic analgesics, such as methadone and meperidine, and endogenous opioid peptides, the enkephalins and the endorphins. *See also* opiate.

opioid peptide Peptide that exerts morphine-like actions through interactions with opiate receptors; for example, endorphin, enkephalin.

oppositional defiant disorder Mental disorder of childhood characterized by pervasive negativism, continuous argumentativeness, and an unwillingness to comply with reasonable suggestions and persuasion.

oral dyskinesia *See* tardive dyskinesia.

oral phase Earliest stage in psychosexual development, which lasts through the first 18 months of life. During this period the oral zone is the center of the infant's needs, expression, and pleasurable erotic experiences; it has a strong influence on the organization and development of the child's psyche.

organic A term used in psychiatry to denote a structural change or a chemical imbalance in the brain to account for a mental disorder. The term is still in common use but is no longer part of

the official DSM nomenclature. *See also* organic disorder.

organic anxiety syndrome Syndrome characterized by prominent, recurrent panic attacks or generalized anxiety attributable to some clearly defined organic factor; cognitive functioning may be adversely affected as a secondary phenomenon.

organic brain syndrome (OBS) ORGANIC MENTAL SYNDROME.

organic delusional disorder Disorder characterized by the presence of prominent delusions in a state of full wakefulness and alertness than can be attributed to some clearly defined organic factor.

organic disorder Illness that is caused by structural or biochemical pathology. The term is not used in DSM but remains in common use. *Compare* functional disorder.

organic hallucinosis Disorder characterized by prominent or persistent hallucinations in a state of full wakefulness that can be attributed to some specific organic factor.

organic mental disorder (OMD) Mental disorder caused by transient or permanent brain dysfunction attributable to specific organic factors. In DSM, the organic mental disorders include dementias arising in the senium and presenium, psychoactive substance-induced organic mental disorders, and organic disorders associated with Axis III physical disorders or conditions or whose etiology is unknown. Also called organic psychosis.

organic mental syndrome (OMS) Psychological or behavior abnormality associated with transient or permanent dysfunction of the brain of unknown or unspecified etiology. OMS is characterized by a combination of disorientation to time, place, or person; signs of psychosis, for example, hallucinations or delusions; and impaired judgment and impulse control. In DSM, the organic mental syndromes include delirium, dementia, amnestic syndrome, organic delusional syndrome, organic hallucinosis, organic mood syndrome, organic anxiety syndrome, organic personality syndrome, intoxication, and withdrawal; if the cause is known, the condition is called organic mental disorder. Also called organic brain syndrome.

organic mood syndrome Disorder characterized by either a depressive or a manic mood attributed to a clearly organic factor.

organic personality syndrome DSM term for frontal lobe syndrome.

organic psychosis ORGANIC MENTAL DISORDER.

orgasm Sexual climax, the peak psychophysiological reaction to sexual stimulation.

orgasm disorders One of the major categories of sexual dysfunctions. Characterized by the recurrent and persistent inability to achieve orgasm. Also called orgasmic dysfunction.

orgasmic phase Stage in the human sexual response cycle, as described by Masters and Johnson, in which orgasm occurs.

orgasmic platform Outer third of the vagina, which displays marked vasocongestion and contraction in the plateau phase of the female sexual response cycle, as described by Masters and Johnson.

orientation State of awareness of oneself and one's surroundings in terms of time, place, and person.

original thirteen *See* American Psychiatric Association.

orphan drug Any pharmacological agent that is not economical to produce because the condition for which it is indicated affects relatively few persons.

Othello syndrome Eponym for conjugal paranoia.

orthomolecular psychiatry Theoretical school of thought proposing that psychiatric illnesses are due to biochemical abnormalities that result in increased requirements for specific substances, such as vitamins.

orthomolecular treatment Treatment that entails administration to psychiatric patients of large doses (megadoses) of vitamins (megavitamin therapy), trace elements, or other substances.

orthopsychiatry Interdisciplinary approach to the study and practice of maintaining or restoring mental health, involving principles derived from psychiatry, psychology, sociology, social work, medicine, and other fields; particular emphasis is placed on preventive techniques to promote healthy emotional development and growth, especially mental hygiene.

orthostatic hypotension Reduction in blood pressure brought about by a shift from a recumbent position to an upright position; observed as a side effect of several psychotropic drugs.

other-directed person Person who is readily influenced and guided by the attitudes and values of other people. *Compare* inner-directed person.

other interpersonal problems DSM category describing problems between one or more persons that cause sufficient strain to bring a person into contact with the mental health care system; the stress-inducing circumstances, coping mechanism, and symptoms that have brought someone to seek consultation or treatment must be individually evaluated.

other specified family circumstance DSM category used when the focus of attention or treatment is not due to a mental disorder or a parent–child or marital problem; may involve in-laws, grandparents, or other family members.

outcome variable DEPENDENT VARIABLE.

outpatient Ambulatory patient receiving medical care who does not require hospitalization.

ovarian dysgenesis TURNER SYNDROME.

overactivity Abnormality in motor behavior that can manifest itself as psychomotor agitation, hyperactivity (hyperkinesis), tics, sleepwalking, or compulsions.

overanxious disorder Anxiety disorder of childhood or adolescence characterized by excessive worrying and fearful behavior grossly disproportionate to the magnitude of the real environmental stress or threat.

overcompensation Conscious or unconscious exaggerated correction or compensation for a real or imagined physical or psychological deficit.

overdetermination Concept that phenomena such as dreams and neurotic symptoms reflect the operation of multiple causative factors, particularly with regard to symbolic meaning or significance.

overt homosexuality Behaviorally expressed homosexuality, as distinct from unexpressed conscious or unconscious homoerotic wishes that are held in check (latent homosexuality).

overvalued idea False or unreasonable belief or idea that is sustained beyond the bounds of reason. It is held with less intensity or duration than a delusion but is usually associated with mental illness.

P

pain disorder Disorder characterized by multiple, recurrent pain complaints with no apparent physical cause; symptoms are linked to psychological factors. Pain disorder is a variant of somatic symptom disorder in which pain symptoms are the predominant physical complaints; a DSM-5 diagnosis. The field of pain medicine encompasses pain linked to both psychological and/or organic factors.

pain-pleasure principle Psychoanalytic concept which states that, in psychic functioning, a person tends to seek pleasure and avoid pain. *See also* pleasure principle.

panic Acute, intense attack of anxiety associated with personality disorganization; the anxiety is overwhelming and accompanied by feelings of impending doom.

panic attack Episode of acute intense anxiety occurring in panic disorder, schizophrenia, major depression, and somatization disorder.

panic disorder DSM classification of anxiety disorder characterized by attacks of acute intense anxiety, with or without agoraphobia.

panphobia Overwhelming fear of everything.

pantomime Gesticulation; psychodrama without the use of words.

paradoxical sleep REM SLEEP.

paralogia EVASION.

paralysisagitans PARKINSONISM.

paralytic ileus Intestinal obstruction resulting from inhibition of bowel motility, a rare anticholinergic side effect of certain psychoactive drugs. Also called adynamic ileus.

parameter Way of measuring or describing an object or evaluating a subject.

parametric tests of significance Statistical methods used to estimate population parameters, such as the mean and standard deviation; use of these methods requires assumptions about the population characteristics, for example, that the variable under study is normally distributed within the population from which the sample is drawn.

paramnesia Disturbance of memory in which reality and fantasy are confused. It is observed in dreams and in certain types of schizophrenia and organic mental disorders; includes phenomena such as déjà vu and déjà entendu, which may occur occasionally in normal persons.

paranoia Rare psychiatric syndrome marked by the gradual development of a highly elaborate and complex delusional system, generally involving persecutory or grandiose delusions, with few other signs of personality disorganization or thought disorder.

paranoiac Person with paranoia.

paranoid delusion Includes persecutory delusions and delusions of reference, control, and grandeur.

paranoid disorder *See* delusional disorder.

paranoid ideation Thinking dominated by suspicious, persecutory, or grandiose content of less than delusional proportions.

paranoid personality disorder Personality disorder characterized by rigidity, hypersensitivity, unwarranted suspicion, jealousy, envy, an exaggerated sense of self-importance, and a tendency to blame and ascribe evil motives to others.

paranoid schizophrenia Type of schizophrenia characterized by the presence of persecutory or grandiose delusions, often accompanied by hallucinations; no longer included in DSM-5.

paraphasia Abnormal speech in which one word is substituted for another, the irrelevant word generally resembling the required one in its morphology, meaning, or phonetic composition; the inappropriate word may be either a legitimate one used incorrectly, such as "clover" instead of "hand," or a bizarre nonsense expression, such as "treen" instead of "train." Paraphasic speech may be seen in organic aphasias and in mental disorders, such as schizophrenia. *See also* metonymy; word approximation.

paraphilia Psychosexual disorder with a sexual dysfunction characterized by persistent and recurrent sexual fantasies, often of an unusual nature, without which imagery erotic arousal or orgasm is not attained; fantasies generally involve themes of suffering, humiliation, sexual activity with nonconsenting partners, or preference for a nonhuman object for sexual arousal. In DSM, the paraphilias include fetishistic disorder, transvestic disorder, frotteuristic disorder, pedophilic disorder, exhibitionistic disorder, voyeuristic disorder, sexual masochism disorder, and sexual sadism disorder. Also called sexual deviations.

paraphrenia Chronic condition characterized by fantastic and absurd delusions without significant personality deterioration and without the primary disturbances of thought and affect that characterize schizophrenia.

parapraxis Faulty act, such as a slip of the misplacement of an article. Freud ascribed parapraxes to unconscious motives.

paraprofessional Trained person who works in a mental hospital and may have a degree in the arts or from some professional school other than those serving the mental health group but has not obtained a degree in one of the usual mental health professions.

parapsychology Branch of psychology concerned with extra-normal events and behavioral phenomena not accounted for or explained by the tenets and laws of contemporary science; for example, clairvoyance, telepathy, psychokinesis. *See also* metapsychology.

parasomnias A group of dysfunctions associated with sleep, sleep stages, or partial arousals characterized by unusual behavior, experiences, or physiological events during sleep; composed of the three subtypes: non-REM movement sleep arousal disorders (incomplete awakening accompanied by sleepwalking, sleep-related eating, sleep-related sexual behavior, or sleep terror disorder); REM sleep behavior disorder (vocal or motor behavior during sleep); and nightmare disorder.

parasympathetic nervous system *See* autonomic nervous system.

parataxic distortion Harry Stack Sullivan's term for distortions in judgment or attitude in interpersonal relations based on patterns set by earlier experience; previously developed ways of coping with significant people in a person's life are applied by the person in later interpersonal integrations.

parens patriae Power of the state to involuntarily hospitalize mentally ill persons in need of care; also called "police power," in that it prevents mentally ill persons from doing harm to themselves or to others.

Parent–child problem DSM category that refers to problems between a parent and a child; for example, in a family in which the parents are divorced, with either the custodial or the noncustodial parent, and in the remarriage of a divorced or widowed parent.

paresis Weakness or partial paralysis of organic origin.

paresthesia Abnormal spontaneous tactile sensation, such as a burning, tingling, or pins-and-needles sensation.

peripartum depression Depressed mood, anxiety, insomnia, and weight changes that occur near parturition or within 1 month of giving birth; a DSM-5 diagnosis. Previously called postpartum depression.

parkinsonism Extrapyramidal dyskinesia characterized by resting tremor, rigidity, bradykinesia, and postural abnormalities. Parkinsonism is of unknown etiology and occurs in middle or late life, typically with a gradual progression and a prolonged course. Also called paralysis agitans; Parkinson disease; idiopathic parkinsonism. A parkinsonian syndrome may develop during the course of therapy with antipsychotic phenothiazine or butyrophenone drugs; such drug-induced parkinsonism is reversible when the causative drug is withdrawn or its dosage reduced.

Parkinson disease PARKINSONISM.

partial hospitalization System of treating mental illness in which the patient is hospitalized on a part-time basis. *See also* day hospital; night hospital; weekend hospital.

partial insanity MONOMANIA.

passive-aggressive personality disorder Personality disorder in which the patient manifests aggressive behavior in passive ways, such as obstructionism, pouting, stubbornness, and intentional inefficiency.

passive-dependent personality *See* dependent personality disorder.

pastoral counseling Use of psychotherapeutic principles by a clergyman in helping parishoners with emotional problems.

patch amnesia LOCALIZED AMNESIA.

pathognomonic Denoting a sign or symptom specifically diagnostic of a particular disease entity.

pathological gambling Compulsion to gamble, one of the disorders of impulse control. *See also* gambling disorder.

pathological intoxication ALCOHOL IDIOSYNCRATIC INTOXICATION.

pathophobia Abnormal fear of disease or suffering. Also called nosophobia.

paucity of speech Limited use of speech, seen in various pathological states.

pavlovian conditioning CLASSICAL CONDITIONING. *See* conditioning.

Pavlov, Ivan Petrovich (1849–1936) Russian neurophysiologist famous for his work in conditioning.

pavor nocturnus SLEEP TERROR DISORDER.

PCP Phencyclidine.

peccatophobia Abnormal fear of sinning.

pecking order Hierarchy or sequence of authority in an organization or social group.

pederasty Anal intercourse between a man and a boy.

pedophilia Paraphilia in which a child is the sexual object.

pedophobia Abnormal fear of children.

peer review Review of physician services by a panel of other physicians.

Peer Review Organization (PRO) Federal review organization for hospitals receiving Medicare funds. In order to promote compliance with federal guidelines for health and hospital care, the PRO conducts independent utilization reviews and quality-of-care studies, validates diagnostic related group (DRG) assignments, and reviews hospital admissions and readmissions. *See also* Professional Standards Review Organization; utilization review committee.

pellagra Deficiency of niacin (nicotinamide) characterized by diarrhea, dermatitis, and mental disturbance.

penis envy Sigmund Freud's concept that the woman envies the man for his possession of a penis; sometimes used to refer to a woman's generalized envy of men.

perception Conscious awareness of elements in the environment by the mental processing of sensory stimuli; sometimes used in a broader sense to refer to the mental process by which all kinds of data, intellectual, emotional, as well as sensory, are meaningfully organized. *See also* apperception.

perfectionism Practice of demanding of others or of oneself a higher quality of performance than is required by the situation.

period prevalence Research term for the total number of cases of a disease known to have existed during a specified time period. *Compare* point prevalence.

perseveration 1. Pathological repetition of the same response to different stimuli, as in a repetition of the same verbal response to different questions. **2.** Persistent repetition of specific words or concepts in the process of speaking. Seen in organic mental disorders, schizophrenia, and other mental illness. *See also* verbigeration.

persistence Term used to reflect a personality trait characterized by ambition, overachievement, resistance to frustration, and fatigue. Persistence may be enhanced by psychostimulants.

persistent (chronic) motor or vocal tic disorder A type of motor disorder consisting of either motor or vocal tics, but not both, that develops before adulthood. Compare Tourette disorder.

persistent complex bereavement disorder Disorder characterized by chronic and persistent grief about the death of someone close to the patient, accompanied by feelings of bitterness, anger, or ambivalence toward the deceased and social withdrawal; a DSM-5 diagnosis. Also known as complicated grief or complicated bereavement.

persona In Jungian psychology, the outside personality that a person presents to the world, as opposed to his or her inner self (anima).

personality Characteristic configuration of behavior response patterns that each person evolves as a reflection of his or her individual adjustment to life.

personality disorder Mental disorder characterized by inflexible, deeply ingrained, maladaptive patterns of adjustment to life that cause either subjective distress or significant impairment of adaptive functioning; manifestations are generally recognizable in adolescence or earlier. Personality disorders listed in DSM include paranoid, schizoid, schizotypal, histrionic, narcissistic, antisocial, borderline, avoidant, dependent, obsessive-compulsive, and passive-aggressive.

personality traits Descriptive features of an individual's personality that are within normal

limits.

persuasion Supportive psychotherapy in which the patient is encouraged to follow advice; persuasion is directive and limited in its goals.

pervasive developmental disorder Class of developmental disorders characterized by severe distortions in social skills, language, and contact with reality (e.g., autism spectrum disorder). Many psychological functions are deficient, and a child with this type of disorder displays abnormalities that are not normal for any stage of development. This condition was previously known as childhood schizophrenia. In DSM-5, Asperger syndrome, Rett syndrome, and childhood disintegrative disorder are types of pervasive developmental disorders. *Compare* specific learning disorder.

pervasive developmental disorder not otherwise specified A diagnostic category used in DSM-5 to refer to pervasive developmental disorders that may have contradictory features and/or lack sufficient findings to meet full criteria for one of the other pervasive developmental disorders (autism, Asperger syndrome, Rett syndrome, childhood disintegrative disorder).

perversion Deviation from the expected norm. In psychiatry it commonly signifies sexual perversion. *See also* psychosexual disorder.

perverted logic EVASION.

PET Positron emission tomography.

petit mal epilepsy Epileptic disorder characterized by recurrent minor, nonconvulsive, momentary lapses of consciousness, commonly beginning in childhood and having a self-limited course. A typical attack involves a sudden transient loss of contact with the environment, during which the patient develops a dazed expression and pauses in what he is doing; this type of seizure has a classical electroencephalographic representation consisting of synchronous, symmetrical spike-and-wave discharges at 3 per second, without cortical localization. Also called absence epilepsy; minor epilepsy.

phallic overbearance Domination of another person by aggressive means; generally associated with masculinity in its negative aspects.

phallic phase Third stage in psychosexual development, following the anal stage and lasting from about age 2 or 3 to about age 6; resolution of the Oedipus complex is the dominant developmental conflict during this stage. During this period, sexual interest, curiosity, and pleasurable experiences are centered on the penis in boys and the clitoris in girls. Also called oedipal phase.

phantasy Alternative spelling of fantasy.

phantom limb False sensation that an extremity that has been lost is in fact present.

pharmacokinetics Way in which the body processes a drug; its principal divisions are drug absorption, distribution, metabolism, and excretion.

pharmacotherapy PSYCHOPHARMACOLOGY.

phase of life problem DSM category that refers to problems not severe enough to warrant a psychiatric diagnosis but that interfere with functioning, as arising from marriage, work, change in job, culture shock.

phencyclidine (PCP) Substance of abuse with effects similar to hallucinogens. Also called angel dust.

phenomenology Study of events or happenings in their own right, rather than from the point

of view of inferred causes; associated with existential psychiatry and reflects the theory that behavior is determined by the way the person perceives reality, rather than by external reality in objective terms.

phenothiazine One of a class of psychotropic drugs, the prototype of which is chlorpromazine, which are effective in the treatment of psychoses, especially schizophrenia; they exert a favorable influence with regard to the fundamental symptoms of schizophrenia as well as the secondary or accessory symptoms. Their mechanism of action may be related to the blockade of dopamine receptors at critical sites in the central nervous system.

phenotype Outward, observable expression of a person's genetic constitution. *Compare* genotype.

phenylketonuria (PKU) Genetic metabolic disorder characterized by the inability to convert the amino acid phenylalanine to tyrosine. If untreated, severe mental retardation results from the accumulation of large concentrations of toxic phenylketones; when detected early in infancy, the condition is treatable by the elimination of phenylalanine from the diet.

phenylpyruvic oligophrenia Former name for phenylketonuria.

pheromone Chemical signal that a person releases into the external environment and that affects the behavior or physiological states of other persons.

phobia Persistent, pathological, unrealistic, intense fear of an objector situation; the phobic person may realize that the fear is irrational but is, nonetheless, unable to dispel it. Some examples are: acarophobia (small objects), acrophobia (high places), agoraphobia (open places, leaving familiar setting of home), aichmophobia (pointed objects), ailurophobia (cats), algophobia (pain), androphobia (men), astraphobia or astrapophobia (thunder and lightning), autophobia (being alone, or of self), bathophobia (depths), claustrophobia (closed or confined places), coprophobia (excreta), cynophobia (dogs), demophobia (crowds), erythrophobia (blushing), gynophobia (women), haphephobia (being touched), hematophobia (sight of blood), hydrophobia (water), hypnophobia (sleep), lalophobia or glossophobia (speaking), mysophobia (dirt, contamination), necrophobia (dead bodies), nyctophobia (darkness, night), panphobia (everything), pathophobia or nosophobia (disease, suffering), peccatophobia (sinning), pedophobia (children), phonophobia (speaking aloud), photophobia (strong light), sitophobia (eating), taphophobia (being buried alive), thanatophobia (death), toxophobia (being poisoned), xenophobia (strangers), zoophobia (animals).

phobic disorder Frequent disorder characterized by a specific fear, which may represent a temptation, a punishment for the forbidden impulse, or a combination of both.

phobic neurosis PHOBIC DISORDER.

phonological disorder *See* developmental articulation disorder.

phonophobia Abnormal fear of speaking aloud.

photophobia Abnormal fear of strong light.

phrenology Study of the bony conformation of the skull in the now discredited belief that it related to mental faculties and traits.

phylogenetic Pertaining to the development of the species. *Compare* ontogenetic.

physical dependence *See* drug dependence.

physical disorder Disorder of body function whose manifestations are not primarily behavioral or psychological.

Piaget, Jean (1896–1980) Swiss psychologist noted for his research on cognitive development in children; he divided the development of intelligence into three major periods: Sensorimotor (birth to 2 years), concrete operations (2 to 12 years), and formal operations (12 years through adult life).

piblotko Culture-specific syndrome seen in Eskimos, usually women, in which the affected person screams, cries, runs naked through the snow, sometimes with suicidal or homicidal tendencies.

pica Craving and eating of nonfood substances, such as paint and clay.

Pick disease *See* frontotemporal dementia.

pickwickian syndrome Condition characterized by obesity, hypoventilation, and hypersomnia.

Pinel, Phillipe (1746–1826) French reformer in the field of mental illness known for his work in abolishing physical restraints on hospitalized mental patients.

piperazine Phenothiazine derivative with antipsychotic affects; for example, fluphenazine (Prolixin).

piperidine Phenothiazine derivative with antipsychotic effects; for example, thioridazine (Mellaril).

PKU Phenylketonuria.

placebo 1. Inactive substance or preparation that is given as though it were a pharmacologically active medicine; primarily used in controlled studies to determine the efficacy of medicinal preparations; any beneficial or deleterious effects of the placebo may be ascribed to psychological factors. **2.** Any type of treatment that has no specific effects on the particular illness involved.

placebo effect Phenomenon in which a person exhibits a clinically significant response to a pill containing a therapeutically inert substance or a treatment without specific effect on the person's condition. Placebo effects are not limited to subjective reports; physiological functions may be objectively influenced.

plaques In psychiatry, pathological masses of axons, neurons, and dendrites with a central amyloid core; associated with Alzheimer disease.

plasma level *See* blood level.

plasticity *See* neuronal plasticity.

plateau phase Stage in the human sexual response cycle, as described by Masters and Johnson, characterized by increased genital blood flow and sense of sexual excitement; it precedes orgasm.

play therapy Type of therapy used with children in which they reveal their problems on a fantasy level with dolls, clay, and other toys; the therapist intervenes opportunely with helpful explanations about the patient's responses and behavior in language geared to the child's comprehension. *See also* activity group therapy.

pleasure principle In psychoanalysis, the principle by which the id seeks immediate tension reduction by either direct or fantasied gratification; developmentally, the pleasure principle antedates the reality principle. Also called the pleasure-pain principle. *See also* reality principle.

pleonexia Psychiatric disorder in which the patient has an excessive desire to acquire wealth or

objects.

pluralism In psychiatry, the notion that multiple factors affect behavior.

PMDD Premenstrual dysphoric disorder.

point prevalence Research term for the total number of cases of a disease known to have existed at a given point in time. *Compare* period prevalence.

police power PARENS PATRIAE.

polygenic Involving the combined action or effects of several genes.

polymorphous perverse sexuality Psychoanalytic conceptualization of sexuality in the human infant, for whom sexual gratification can be achieved by homosexual, heterosexual, or inanimate stimulation of several body zones; all humans, therefore, have the potential to develop sexual perversions in adulthood, depending on how the sexual drives are channeled developmentally.

polyphagia Excessive eating. *See also* bulimia nervosa.

polysomnography Battery of tests that include an electroencephalogram, electrocardiogram, and electromyogram; often given with tests for penile tumescence, blood oxygen saturation, body movement, body temperature, galvanic skin response, and gastric acid. The tests are of assistance in a variety of conditions, such as insomnia, nocturnal myoclonus, sleep apnea, enuresis, somnambulism, seizure disorders, impotence, vascular headache, gastroesophageal reflux, and depression.

polysubstance dependence Disorder characterized by dependence on at least three categories of psychoactive substances.

polyuria Excessive excretion of urine.

population Research term for the entire collection of a set of objects.

porphyria Metabolic disorder characterized by excessive urinary excretion of porphyrins (pigments found in hemoglobin and certain cellular enzyme systems) or their precursors; mental symptoms are prominent in the form known as acute intermittent porphyria. Abdominal pain, constipation, and neuromuscular disturbances complete the clinical picture. Attacks may be precipitated by barbiturates.

positive reinforcement *See* reinforcement.

positive symptoms In schizophrenia: hallucinations, delusions, and thought disorder.

positive transference *See* transference.

positron emission tomography (PET) Imaging technique using short-lived positron-emitting elements of oxygen, carbon, and nitrogen used to label organic compounds (e.g., glucose, amino acids, and neurotransmitters), which are then introduced into the body by injection or inhalation; energy emitted is recorded by scintillation detectors and computer analyzed to construct and display images of transverse sections of tissues, such as the brain. Quantitative data can be assessed in specific areas for such functions as blood flow, oxygen or glucose metabolism, and other complex physiological phenomena; alterations in metabolism associated with disease and the actions of medications can also be monitored in this way.

postpartum depression *See* peripartum depression.

postpartum psychosis A descriptive term for depressive or psychotic symptoms in women who have recently given birth; a subtype of bipolar disorder in DSM-5. Also called puerperal psychosis. *See also* peripartum depression.

posttraumatic stress disorder (PTSD) Anxiety disorder that occurs after and as a result of a major, disturbing, out of the ordinary event in a person's life, such as war or rape. The prominent symptoms of this disorder include nightmares, flashbacks, withdrawal, and hypervigilance; may have a delayed onset. *See also* acute stress disorder.

posturing Strange, fixed, and bizarre bodily positions held by a patient for an extended time. *See also* catatonia.

potency Male ability to perform the sexual act; specifically, the capacity to achieve and maintain an erection during coitus.

poverty of content of speech Speech that is quantitatively adequate but qualitatively inadequate because it imparts minimal, vague, or repetitious information.

poverty of speech Restriction in the amount of speech used; replies may be monosyllabic. *See also* laconic speech.

PPO Preferred Provider Organization.

practice effect Research term referring to the improvement of performance in a task as a result of repeated trials.

Prader–Willi syndrome A developmental disability as a result of a small deletion involving chromosome 15. Symptoms include compulsive eating, obesity, intellectual disability, hypogonadism, small stature, hypotonia, and small hands and feet. Children also exhibit oppositional and defiant behavior.

preconscious One of three divisions of the psyche in Freud's topographic theory of the mind (the others being the conscious and unconscious), including all mental contents that are not in immediate awareness but that can be consciously recalled with effort.

predictive value Proportion of true-positive or true-negative tests; indicates what percentage of test outcomes are expected to coincide with assigned diagnoses.

predictor variable Test or score used to assess or predict future performance.

Preferred Provider Organization (PPO) Alternative delivery organization that employs a prospective payment system in which a corporation or insurance company forms an agreement with a particular group of community hospitals and doctors to supply health services to PPO members at a previously determined lower rate; patients enrolled in a PPO select their physicians and hospitals from among the list of participants.

prefrontal lobotomy *See* lobotomy.

pregenital stages In psychoanalysis, the first stages of psychosexual development (the oral and anal stages) that occur before the genitals have begun to exert the predominant influence on sexual behavior.

prejudice Preconceived adverse judgment or opinion formed without just grounds or sufficient knowledge; elements of irrational suspicion or hatred are often involved, as in racial prejudice.

premature ejaculation Ejaculation occurring before or immediately after vaginal intromission during coitus; broader definitions take into consideration the ability of a male to delay ejaculation during coitus for a sufficient length of time to satisfy a normally responsive female partner.

premenstrual dysphoric disorder (PMDD) Depressive mood disorder characterized by irritability, emotional lability, headache, and anxiety or depression that occurs about 1 week before menses and resolves after the menstrual cycle is complete. Also called late luteal phase

dysphoric disorder.

premorbid Occurring before the onset of disease.

preoccupation of thought Centering of thought content on a particular idea, associated with a strong affective tone, such as paranoid trend or a suicidal or homicidal preoccupation.

presenile dementia DEMENTIA OF THE ALZHEIMER TYPE; PRESENILE ONSET.

presenium Period preceding old age, generally ages 55 to 65.

pressure of speech, pressured speech Increase in the amount of spontaneous speech; rapid, loud, accelerated speech, as occurs in mania, schizophrenia, and cognitive disorders.

prevalence Total number of cases of a disease existing at a given point in time (point prevalence) or during a specified time period (period prevalence). *See also* incidence.

prevalence rate Number of diagnosed cases of a disease or a condition at a particular point in time; prevalence rate per 1,000 is defined as the cases at a specific point in time, divided by the number of persons in the population at a specific point in time, multiplied by 1,000.

preventive psychiatry Branch of preventive medicine concerned with mental disorders. Encompassed within its scope are measures to prevent mental disorders (primary prevention); measures to limit the severity of illness, as through early case finding and treatment (secondary prevention); and measures to reduce disability after a disorder (tertiary prevention).

primal father Hypothetical head of a tribe depicted by Sigmund Freud in *Totem and Taboo* as slain by his sons, who subsequently devour him in a cannibalistic rite; later, he is promoted to a god. The son who murders him is the prototype of the tragic hero, and the memory of the crime is perpetuated in the conscience of the individual and of the culture.

primal scene In psychoanalysis, the real or fantasied observation by a child of sexual intercourse, particularly between his or her parents.

primary care physician Physician who specializes in family medicine, internal medicine, pediatrics, and obstetrics and gynecology; many believe psychiatry should be designated a primary care specialty.

primary diagnosis Major emotional disorder for which the patient is treated. Also called principal diagnosis.

primary gain Reduction of anxiety achieved by a defense mechanism; relief from tension or conflict through neurotic illness. *Compare* secondary gain.

primary narcissism *See* narcissism.

primary parkinsonism *See* parkinsonism.

primary prevention *See* preventive psychiatry.

primary process thinking In psychoanalysis, the mental activity directly related to the functions of the id and characteristic of unconscious mental processes; marked by primitive, prelogical thinking and by the tendency to seek immediate discharge and gratification of instinctual demands. Includes thinking that is dereistic, illogical, and magical; normally found in dreams, abnormally in psychosis. *Compare* secondary process thinking.

primary reinforcer *See* reinforcement.

primary retarded ejaculation Lifelong inability to ejaculate during coitus. *Compare* secondary ejaculation.

primitive idealization Defense mechanism in which external objects that are viewed as either "all good" or "all bad" are unrealistically endowed with great power; most commonly, the "all

good" object is seen as omnipotent, or ideal, while the badness of the "all bad" object is greatly inflated.

Prince, Morton (1854–1929) American neurologist and psychiatrist known for his study of multiple personalities.

principal diagnosis PRIMARY DIAGNOSIS.

prion disease A class of disorders caused by a transmissible infectious protein. Includes Creutzfeldt–Jakob disease (CJD), Gerstmann–Straussler–Scheinker disorder, kuru, and fatal familial insomnia.

prison neurosis *See* chronophobia.

prison psychosis Psychotic reaction to incarceration or to the prospect of incarceration.

privacy 1. Limiting access of others to one's body or mind, including dreams, fantasies, thoughts, or beliefs. **2.** Freedom from intrusion by the state or third persons. **3.** Restricted areas of personal decision, usually in such matters as personal association, abortion, or bodily integrity.

privilege Legal term referring to the legal right of patients to prevent their physicians from testifying about information obtained in the course of treatment; the legal formulation of the ethical principle of confidentiality. In some states, information obtained in the course of the doctor–patient relationship is considered privileged communication.

privileged communication *See* privilege.

PRO Peer Review Organization.

probability Quantitative statement of the likelihood that an event will occur. A probability of 0 means that the event is certain not to occur; a probability of 1.0 means the event will occur with certainty.

problem-oriented record Method for organizing a patient's clinical chart, consisting of five parts: data base, problem list, treatment plan, progress notes, and discharge summary.

problem-solving Capacity of a person to master a new or novel task.

process schizophrenia Unofficial term for schizophrenia attributed more to endogenous factors than to specific environmental influences. *Compare* reactive schizophrenia.

prodrome Early symptom of a disease, often serving as a warning or premonitory sign of the approach of a morbid condition; for example, the aura preceding an epileptic seizure or classic migraine.

Professional Standards Review Organization (PSRO) Organization of physicians in a given region established to review the quality of health care services, particularly with regard to whether the inpatient services rendered are medically necessary. Such an organization has been required by federal law since 1972 in relation to hospital care under Medicare and Medicaid. *See also* peer review organization; utilization review committee.

profound mental retardation *See* mental retardation.

prognosis 1. Opinion as to the probable future course, extent, and outcome of an illness. **2.** Specific goals of therapy.

projection Unconscious defense mechanism in which a person attributes to another those generally unconscious ideas, thoughts, feelings, and impulses that are in himself or herself undesirable or unacceptable as a form of protection from anxiety arising from an inner conflict; by externalizing whatever is unacceptable, the person deals with it as a situation apart

from themselves. *See also* blind spot.

projective identification Defense mechanism in which unwanted aspects of the self are projected onto another person such that the individual projecting feels at one with the object of the projection. It allows one to distance and make oneself understood by exerting pressure on another individual to experience feelings similar to one's own.

projective test Psychological test with loosely structured test material that requires the subject to reveal his or her own feelings, personality, or psychopathology; for example, Rorschach Test, Thematic Apperception Test.

prolactin Anterior pituitary hormone that has been studied as a potential index of dopamine activity and receptor sensitivity in studies of central nervous system function in psychiatric patients.

proprioception KINESTHETIC SENSE.

prosopagnosia Inability to recognize familiar faces that is not due to impaired visual acuity or level of consciousness.

prospective study Study based on observing events as they occur.

prototaxic Harry Stack Sullivan's term for primitive processes. *See also* primary process.

pseudoaggression Neurotic defense in which h patient denies basic masochistic feelings and displays them, instead, as false aggression.

pseudoauthenticity False or copied expression of thoughts and feelings.

pseudocollusion Sense of closeness, relationship, or cooperation that is not real but is based on transference.

pseudocyesis Rare condition in which a nonpregnant patient has the signs and symptoms of pregnancy, such as abdominal distention, breast enlargement, pigmentation, cessation of menses, and morning sickness.

pseudodementia 1. Dementia-like disorder that can be reversed by appropriate treatment and is not caused by organic brain disease. **2.** Condition in which a patient shows an exaggerated indifference to his or her surroundings in the absence of a mental disorder; also occurs in depression and factitious disorders.

pseudologia fantastica Disorder characterized by uncontrollable lying in which a patient elaborates extensive fantasies that he freely communicates.

pseudomotor agitation Physical and mental overactivity that is usually nonproductive and is associated with a feeling of inner turmoil, as seen in agitated depression.

pseudoparkinsonism Syndrome that resembles Parkinson disease, resulting from long-term use of antipsychotic drugs. Also called drug-induced parkinsonism.

psilocin *See* psilocybin.

psilocybin 4-Phosphoryloxy-N′,N′-dimethyltryptamine, the active ingredient of various Mexican hallucinogenic mushrooms, which acts as a psychedelic drug producing changes in perception and mood and alterations of consciousness. A variation is called psilocin.

PSRO Professional Standards Review Organization.

psychalgia PSYCHOGENIC PAIN DISORDER.

psychasthenia Obsolete term used by Pierre Janet to describe a syndrome characterized by fears and phobias.

psyche Mind, as divided into conscious, preconscious, and unconscious in Sigmund Freud's

topographic model.

psychedelic Denoting drugs that may induce hallucinations and psychotic states. LSD is the prototype; other psychedelic drugs include mescaline, psilocybin, and marijuana.

psychiatric emergency Disturbance in thoughts, feelings, or actions for which immediate treatment is necessary; for example, suicide intent.

psychiatric illness *See* mental disorder.

psychiatric nurse Clinical nurse specialist who functions as part of the mental health team, usually in an institutional setting; often carries out individual, family, and group psychotherapy.

psychiatric social worker Skilled professional, trained in social work, who works with psychiatrists, usually in an institutional setting; evaluates family, environmental, and social factors in the patient's illness; may work in intake and reception with new patients; may follow-up and counsel after discharge, and also carry out individual, family, and group psychotherapy and participate in community organizations.

psychiatrist Medical doctor whose specialty is the study and treatment of mental disorders. After receiving the M.D. degree, a physician spends 4 years as a resident in a hospital setting. When he or she passes the written and oral examinations of the American Board of Psychiatry and Neurology, he or she becomes a diplomat in psychiatry and is said to be Board certified.

psychiatry Branch of medicine concerned with the prevention, cause, and treatment of mental disorders.

psychic determinism Freudian concept that all mental phenomena have specific antecedent causes, often operating on an unconscious level.

psychic energizer Rarely used term for any antidepressant or stimulant drug.

psychic helplessness *See* helplessness.

psychic trauma Psychologically upsetting event that may have lasting consequences on a person's thinking, feeling, or behavior.

psychoactive drug PSYCHOTHERAPEUTIC DRUG.

psychoactive substance abuse *See* substance use disorder.

psychoactive substance dependence DSM term for drug dependence.

psychoactive substance-induced mental disorder Mental disorder caused by the direct effects on the nervous system of a drug or other substance. In DSM, the substances listed are alcohol; sedatives, hypnotics, or anxiolytics; inhalants; opioids; cocaine; amphetamine or similarly acting sympathomimetics; phencyclidine (PCP) or similarly acting arylcyclohexylamine; hallucinogens; cannabis; nicotine; and caffeine.

psychoanalysis Theory of human mental phenomena and behavior, a method of psychic investigation and research, and a form of psychotherapy originally formulated by Sigmund Freud. As a technique for exploring the mental processes, psychoanalysts includes the use of free association and the analysis and interpretation of dreams, resistances, and transferences. As a form of psychotherapy, it uses the investigative technique, guided by Freud's libido and instinct theories and by ego psychology, to gain insight into a person's unconscious motivations, conflicts, and symbols and thus to effect a change in maladaptive behavior. Also called analysis in depth.

psychoanalyst Psychotherapist, usually a psychiatrist, who has had training in psychoanalysis

and who uses its techniques in a treatment setting.

psychoanalytic group psychotherapy Method of group psychotherapy pioneered by Alexander Wolf and based on the operational principles of individual psychoanalytic therapy; analysis and interpretation of a patient's transferences, resistances, and defenses are modified to take place in a group setting.

psychobiology Adolf Meyer's term referring to the study of the human being as an integrated unit, incorporating psychological, social, and biological functions.

psychodrama Psychotherapy method originated by Jacob Moreno in which personality makeup, interpersonal relationships, conflicts, and emotional problems are expressed and explored through dramatization. Therapeutic dramatization of emotional problems includes: (1) protagonist or patient, the person who presents and acts out his or her emotional problems with the help of (2) auxiliary egos, persons trained to act and dramatize the different aspects of the patient that are called for in a particular scene in order to help the patient express feelings, and (3) the director, leader, or therapist, who guides those involved in the drama for a fruitful and therapeutic session.

psychodynamics Science of mental forces and motivations that influence human behavior and mental activity; the role of unconscious motivation in the causation of human behavior is

psychoendocrinology Study of the role of hormones on thinking, behavior, and emotional disorder.

Psychose passionelle *See* erotomania.

psychogenic Produced or caused by mental factors, rather than organic factors; usually refers to a symptom or an illness.

psychogenic amnesia DSM classification of dissociative disorder characterized by the sudden inability to recall information stored in memory in the absence of an underlying organic mental disorder; when associated with travel to another locale and the assumption of a new identity, the condition is called psychogenic fugue.

psychogenic fugue DSM classification of dissociative disorder characterized by periods of total amnesia in which one travels and assumes a new identity. *See also* fugue; psychogenic amnesia.

psychogenic pain disorder Condition in which there is a preoccupation with pain for at least 6 months in the absence of organic pathology to account for the pain. Also called psychalgia; somatoform pain disorder in DSM.

psychohistory Approach to history that examines events within a psychological framework; attempts to connect individual and collective ideas and emotions with wider historical currents.

psychoimmunology Study of the relationship between the immune system and emotions.

psychokinesis Parapsychological phenomenon in which directed thought processes influence physical events.

psycholinguistics Exploration of the psychological factors involved in the development and use of language.

psychological defense system *See* defense mechanism.

psychological factor affecting physical condition Disorder characterized by physical symptoms partially caused or affected by psychological factors; physiological and organic

changes stem in part from a sustained emotional disturbance; for example, atopic dermatitis, migraine, ulcer. Also called psychophysiological disorder; psychosomatic disorder.

psychologist Person trained in psychology, usually with a graduate degree (M.A., Ph.D.).

psychology Study and profession concerned with mental processes and behavior.

psychometry Science of testing and measuring mental and psychological functioning.

psychomotor Relating to combined physical and mental activity.

psychomotor agitation Physical and mental overactivity associated with a feeling of inner turmoil, as seen in agitated depression.

psychomotor epilepsy COMPLEX PARTIAL SEIZURE.

psychomotor retardation Slowing of mental and physical activity, common in depression.

psychomotor stimulant Drug that arouses the patient through its central excitatory properties; for example, amphetamine, methylphenidate (Ritalin). Also called analeptic.

psychoneuroimmunology Study of psychological factors as causes of diseases or their effect on the course of disease, particularly as it relates to the immune system. *See also* psychoimmunology.

psychoneurosis NEUROSIS.

psycho-oncology Study of the role of emotions and their relation to cancer in terms of prevention, diagnosis, and treatment.

psychopathic personality Former term for the personality characteristic of antisocial personality disorder.

psychopathological Denoting abnormal, maladaptive behavioral or mental activity.

psychopathology 1. Study of mental disorders. **2.** Abnormal, maladaptive behavioral or mental activity.

psychopharmacology Study of the mental and behavioral effects of drugs. *See also* pharmacokinetics.

psychophysiological disorder PSYCHOLOGICAL FACTOR AFFECTING PHYSICAL CONDITION.

psychopolitics 1. Psychological dimensions of political behavior, such as the reciprocal influence on persons of political environments and their societies. **2.** Specific tactics of politicians that are intended to yield benefits to them through the use of psychological strategies.

psychosexual development 1. Broadly, the maturation and development of the sexuality of a person throughout the life cycle. **2.** More specifically, in psychoanalysis, series of dynamic and crucial developmental stages that each person passes through and that influence basic personality characteristics in later life; the stages or phases are oral, anal, phallic, latency, and genital.

psychosexual disorder Disorder of sexual functioning, desire, or performance that is caused, wholly or partly, by psychological factors.

psychosis Mental disorder in which a person's thoughts, affective response, ability to recognize reality, and ability to communicate and relate to others are sufficiently impaired to grossly interfere with his or her capacity to deal with reality; the classical characteristics of psychosis are impaired reality testing, hallucinations, delusions, and illusions.

psychosocial deprivation Deprivation of social and intellectual stimulation; believed to be a

causative factor in mental retardation and emotional disorders in children.

psychosocial development Developmental progress of a person with regard to social relations and social reality, as primarily described by Erik Erikson. Specific developmental tasks characterize successive chronological periods from infancy through maturity: trust versus mistrust (infancy), autonomy versus doubt (toddler), initiative versus guilt (preschool), industry versus inferiority (school age), identity versus identity diffusion (adolescence), intimacy versus isolation (young adulthood), generativity versus self-absorption (adulthood), and integrity versus despair (mature age).

psychosocial dwarfism Syndrome usually first manifested in children 2 to 3 years of age, who typically are unusually short, have frequent growth hormone abnormalities, and have severe behavioral disturbances; all of these symptoms are the result of an inimical caretaker–child relationship and resolve without any medical or psychiatric treatment upon being placed in a more favorable domicile. Also called psychosocially determined short stature.

psychosocially determined short stature PSYCHOSOCIAL DWARFISM.

psychosocial stressor STRESS.

psychosomatic disorder PSYCHOLOGICAL FACTOR AFFECTING PHYSICAL CONDITION.

psychosurgery Neurosurgical intervention to treat a mental disorder that involves destruction or ablation of brain tissue.

psychotherapeutic drug Drug used to alter abnormal thinking, feelings, or behavior; traditionally divided into classes of antipsychotic, antidepressant, antimanic, and antianxiety (anxiolytic) drugs. Also called psychoactive drug; psychotropic drug.

psychotherapist Person trained to treat mental, emotional, and behavioral disorders. For types of therapists, see the specific term.

psychotherapy Treatment for mental illness and behavioral disturbances in which a trained person establishes a professional contract with the patient and through definite therapeutic communication, both verbal and nonverbal, attempts to alleviate the emotional disturbance, reverse or change maladaptive patterns of behavior, and encourage personality growth and development; distinguished from other forms of psychiatric treatment such as the somatic therapies, for example, drug and convulsive therapies.

psychotic 1. Person suffering from a psychosis. **2.** Denoting or characteristic of a psychosis.

psychotic depressive reaction Psychosis distinguished by depressed mood precipitated by some event, usually in a person who had not previously demonstrated severe depression. Also called major depressive disorder, single episode, with mood-congruent psychotic features, in DSM.

psychotomimetic drug Drug that produces psychic and behavioral changes that resemble psychosis; unlike other drugs that can produce organic psychosis as a toxic reaction (e.g., hallucinogenic and psychedelic drugs such as LSD, marijuana, mescaline), a psychotomimetic drug does not produce gross confusion, delirium, or overt memory impairment.

psychotropic drug PSYCHOTHERAPEUTIC DRUG.

PTSD Posttraumatic stress disorder.

puerperal psychosis *See* postpartum psychosis.

punishment Presentation of an aversive stimulus contingent on an operant response, resulting

in a decreased probability of that response recurring. *See also* reinforcement.

Purging disorder Eating disorder characterized by recurrent episodes of self-induced vomiting, laxative and diuretic abuse, and/or the use of enemas after consuming small amounts of food; patients are normal weight and have distorted body image. In DSM-5, included in "Other Specified Feeding or Eating Disorders."

pyromania Compulsion to set fires.

Q

q-sort Technique of personality assessment based on the degree to which a standard set of descriptive statements are believed to describe a particular person.

quadrangular therapy Marital therapy that involves four people, the married pair and each spouse's therapist.

quality assurance Measurement and evaluation of medical and other related services provided with regard to their effectiveness, efficiency, appropriateness, and acceptability; such evaluation will trigger corrective or remedial action for services not meeting the desired standards.

quantitative variable Experimental factor that can be measured.

R

random Statistical term that means occurring by chance or without attention to selection or planning.

random sample In a given population, a group of subjects selected in such a manner that each member of the population has an equal probability of being selected for the sample.

random variable Variable for which the variation is determined by chance.

range Statistical measure of the variability of a set of values, defined as the difference between the largest value and the smallest value.

Rank, Otto (1884–1939) Austrian psychoanalyst, long-time secretary and recorder of the minutes of the Vienna Psychoanalytic Society, and one of Sigmund Freud's earliest followers, who split with him on the significance of the birth trauma, which Rank used as a basis of brief psychotherapy. Rank's works include *The Trauma of Birth* and *The Myth of the Birth of the Hero*.

rape Sexual intercourse without the person's consent and against his or her will, chiefly by force; usually heterosexual but may be homosexual.

rapid eye movement *See* REM latency; REM sleep.

rapport Conscious feeling of harmonious accord, sympathy, and mutual responsiveness between two or more persons; contributes to an effective therapeutic process in both group and individual settings.

rapture-of-the-deep syndrome Psychosis, seen in scuba and deep-sea divers, characterized by disorientation and loss of judgment and skill, and associated with excessive blood nitrogen levels; sensory deprivation may also contribute to this acute, self-limited mental disorder. Also called nitrogen narcosis.

rationalization Ernest Jones's term for an unconscious defense mechanism in which irrational or unacceptable behavior, motives, or feelings are logically justified or made

consciously tolerable by plausible means.

Ray, Isaac (1807–1881) One of the original 13 founders of the American Psychiatric Association, known for his *Treatise on Medical Jurisprudence of Insanity* (1837).

rCBF Regional cerebral blood flow.

reaction formation Unconscious defense mechanism in which a person develops a socialized attitude or interest that is the direct antithesis of some infantile wish or impulse that is harbored either consciously or unconsciously. One of the earliest and most unstable defense mechanisms, closely related to repression; both are defenses against impulses or urges that are unacceptable to the ego.

reactive attachment disorder of infancy or early childhood Severe impairment of the ability to relate, beginning before age 5; presumed to develop from grossly pathological care-giving.

reactive depression Mental disorder characterized by depression developing in reaction to some precipitating life event or circumstance.

reactive disorder Mental disorder judged to be a reaction to one or more life events or circumstances, without which the disorder would not have occurred. *See also* brief reactive psychosis; reactive depression; reactive schizophrenia.

reactive schizophrenia Unofficial term for schizophrenia attributed primarily to predisposing or precipitating environmental factors. *Compare* process schizophrenia.

reading disorder Reading below the expected level for a child's age, education, and intelligence.

reality Totality of objective things and factual events; includes everything that is perceived by a person's special senses and is validated by other people.

reality principle In psychoanalysis, ego processes that modify the demands of the pleasure principle to meet the inescapable demands and requirements of the outside world; net result is behavior that gratifies id impulses within realistic limits. *See also* pleasure principle.

reality testing Fundamental ego function that consists of tentative actions that test and objectively evaluate the nature and the limits of the environment; includes the ability to differentiate between the external world and the internal world and to accurately judge the relation between the self and the environment.

recall Process of bringing stored memories into consciousness. *See* memory.

receiver operating characteristic (ROC) Special statistical method, recently used in psychiatry, which gives a balance between sensitivity and specificity for a diagnostic test, for example, Dexamethasone Suppression Test. The method was derived from signal detection theory as a way to determine how well an electronic receiver was able to distinguish signal from noise.

recent memory Recall of events over the past few days.

recent past memory Recall of events over the past few months.

receptive aphasia Organic loss of ability to comprehend the meaning of words; fluid and spontaneous, but incoherent and nonsensical, speech. *See also* fluent aphasia and sensory aphasia.

receptive dysphasia Difficulty in comprehending oral language; the impairment involves both comprehension and production or language.

receptors Proteins in the neuronal membrane that are exposed in part to the extracellular fluid and specifically recognize neuromessengers; stimulation of receptors generally causes a biological response to occur.

reciprocal inhibition and desensitization Behavior therapy in which a person is conditioned to associate comfortable, supportive surroundings with anxiety-producing stimuli, thus decreasing the anxiety associated with those stimuli.

recovered memory syndrome Condition that occurs when a repressed memory of a traumatic experience is brought back to consciousness, usually during psychotherapy or under hypnosis. A variant of this condition, false memory syndrome, occurs if the recalled memory did not actually happen, but the person believes the memory to be true.

recurrent brief depressive disorder Multiple, brief (less than 2 weeks duration) episodes of depressive symptoms.

recurrent depressive episode Depression that lasts between 2 and 13 days and occurs at least once a month.

recurrent isolated sleep paralysis Sleep disorder characterized by the inability to make voluntary movements while partially conscious at sleep onset or awakening. *See also* parasomnias.

refractory period Temporary state of psychophysiological resistance to sexual stimulation immediately after orgasm in the human sexual response cycle, as described by Masters and Johnson.

reframing Technique used in family therapy whereby a problem or conflict is viewed by the therapist in an entirely different and sometimes startling perspective, thus enabling the family members to achieve objectivity.

regional cerebral blood flow (rCBF) Test using xenon-133, a low-energy gamma ray-emitting isotope, to determine blood flow in the brain; useful for studying diseases that have a decrease in the amount of brain tissue, for example, dementia, ischemia, atrophy.

registration *See* memory.

regression Unconscious defense mechanism in which a person undergoes a partial or total return to earlier patterns of adaptation; observed in many psychiatric conditions, particularly schizophrenia.

regressive-reconstructive approach Psychotherapeutic procedure in which regression is made an integral element of the treatment process; an original traumatic situation is reproduced to gain new insight and to effect personality change and emotional maturation.

rehabilitation All methods and techniques used in an attempt to achieve maximal function and optimal adjustment in a given patient; the physical, mental, social, and vocational preparation of a patient for the fullest possible life compatible with his or her abilities and disabilities. Inasmuch as the process aims also to prevent relapses or recurrences of the patient's condition, it is sometimes called tertiary prevention.

Reich, Wilhelm (1897–1957) Austrian psychoanalyst who emigrated to the United States in 1939. He coined the term "orgone," a kind of universally pervasive physical and biological energy that was typified by the pent-up libidinal energy that sought free expression and release through orgasm.

Reik, Theodor (1888–1969) Psychoanalyst and early follower of Sigmund Freud, who

considered Reik one of his most brilliant pupils; Freud's book, *The Question of Lay Analysis*, was written to defend Reik's ability to practice psychoanalysis without medical training. Reik made many valuable contributions to psychoanalysis on the subjects of religion, masochism, and therapeutic technique.

reinforcement *Positive reinforcement* refers to the process by which certain consequences of behavior increase the probability that the behavior will occur again. Positive reinforcers tend to be viewed as pleasant. *Negative reinforcement* describes the process by which behavior that leads to the removal of an unpleasant event strengthens that behavior. Negative reinforcers tend to be viewed as unpleasant. Stimuli that possess inherent reinforcing characteristics are called *primary reinforcers*, for example, food. Stimuli that acquire reinforcing characteristics by being paired with primary reinforcers are called *conditioned reinforcers*, for example, money. *See also* conditioning.

reinforcer Reinforcing stimulus. *See* reinforcement.

relatedness Sense of sympathy and empathy with regard to others; sense of oneness with others. *Compare* isolation; alienation.

relational problems Impaired functioning of a relational unit. In DSM, parent–child relational problem, partner relational problem, and sibling relational problem.

relative frequency *See* frequency.

relative risk Epidemiological measure of the risk of developing a disorder in a specified subset of the population as compared with the total population or a different subset of it. Comparison is usually made between groups exposed and groups not exposed to a particular hereditary or environmental factor to gain information about the role of that factor in the disorder in question. Relative risk is then expressed as a ratio of the frequency of the disorder in exposed persons to its frequency in those not exposed.

Relative Value Scale Payment scale for medical services developed at Harvard University that reimburses physicians in terms of time spent with patients, office overhead expenses, years of training, and cognitive skills; under the program, surgeons would receive less money and internists would receive more money from third-party (insurance) carriers than under previous payment scale programs. Also called Resource-Based Relative Value Scale (RBRVS).

reliability Research term meaning replicability, that is, the degree to which a test result or experimental finding can be confirmed by repetition. *See also* validity.

religious delusion Delusion involving the Deity or theological themes.

reminiscence Recall or recollection of past experiences, with alternations of remembering an apparently lost memory and forgetting one that was just there.

remission Significant improvement or recovery which may or may not be permanent. *Compare* intermission.

remote memory LONG-TERM MEMORY.

remotivation Group treatment technique used patients in mental hospitals to encourage them and vocational skills.

REM (rapid eye movement) latency First REM episode of the night, about 90 minutes after sleep onset; shortening of REM latency frequency occurs with disorders such as depression and narcolepsy.

REM (rapid eye movement) sleep Stage of sleep during which dreaming occurs and the

sleeper exhibits coordinated rapid eye movement (REM). The electroencephalogram demonstrates a desynchronized pattern of cerebral activity. It accounts for one-fourth to one-fifth of total sleep time. Also called paradoxical sleep.

REM sleep behavior disorder Sleep disorder characterized by the absence of sleep paralysis (atonia) during the REM stage of sleep; consequently, patients will act out their dreams, putting themselves and/or their partners at risk of physical injury. *See also* parasomnias.

repetition compulsion Psychoanalytic concept of an impulse to repeat earlier experiences and situations.

repetitive pattern Continual attitude or mode of behavior characteristic of a person and performed mechanically or unconsciously.

repression Freud's term for an unconscious defense mechanism in which unacceptable mental contents are banished or kept out of consciousness; important in both normal psychological development and in neurotic and psychotic symptom formation. Freud recognized two kinds of repression: (1) repression proper, in which the repressed material was once in the conscious domain; (2) primal repression, in which the repressed material was never in the conscious realm. *Compare* suppression.

repressive-inspirational group psychotherapy Group therapy in which discussion is intended to bolster the patients' morale and help them avoid undesired feelings; used primarily with large groups of seriously regressed patients in institutional settings.

reserpine Alkaloid extracted from the *Rauwolfia serpentina* plant, formerly used as an antipsychotic agent because of its sedative effects but currently used primarily as an antihypertensive agent, although this use has declined partly because of its potential to produce a considerable incidence of nightmares and suicidal depression; that untoward effect may be related to reserpine's depletion of catecholamines at critical sites in the central nervous system.

resident 1. In psychiatry, a patient living in a treatment center for a prescribed period of time. **2.** In medicine, a physician in postgraduate training after graduation from medical school; 4 years of residency training are necessary to become a psychiatrist.

residential treatment facility Center where patients live and receive treatment appropriate for their particular needs; a children's residential treatment facility ideally furnishes both educational and therapeutic experiences for the emotionally disturbed child. Also called a halfway house.

residual Denoting a condition in which a person's acute mental symptoms have subsided.

residual schizophrenia Schizophrenia in which the patient no longer is psychotic but does have some remaining signs of the illness.

resistance Conscious or unconscious opposition to the uncovering of unconscious material; linked to underlying psychological defense mechanisms against impulses from the id that are threatening to the ego.

resistant attachment As described by Mary Ainsworth, a child's simultaneous yet antithetical wish for and avoidance of his or her mother's approach. *Compare* avoidant attachment; secure attachment.

resolution phase Final stage in the human sexual response cycle, as described by Masters and Johnson, during which the sexual system returns to its normal, nonexcited state following

orgasm.

respondent conditioning CLASSICAL CONDITIONING. *See* conditioning.

restless legs syndrome (RLS) A syndrome characterized by an uncomfortable, subjective sensation of the limbs, usually the legs. Tends to be worse at night.

restricted affect Reduction in intensity of feeling tone that is less severe than in blunted affect but clearly reduced. *See* constricted affect.

retardation Reduction or slowing down of mental and physical activity, as often observed in depression. *See also* mental retardation.

retarded ejaculation INHIBITED MALE ORGASM.

retention *See* memory.

retroflexion Psychoanalytic term used notably by Sandor Rado to describe the turning of rage onto oneself.

retrograde amnesia Loss of memory for events preceding the onset of the amnesia. *Compare* anterograde amnesia.

retrospective falsification Memory becomes unintentionally (unconsciously) distorted by being filtered through a person's present emotional, cognitive, and experiential state.

retrospective study Investigation based on past data or past events, such as a case control study or a case history study.

Rett disorder A progressive condition characterized by deceleration in head growth, loss of purposeful hand movements, development of stereotyped hand movements (e.g., hand wringing), psychomotor retardation, loss of social engagement, loss of acquired speech, and ataxia between the ages of 6 months to 2 years.

Rett syndrome A progressive condition characterized by deceleration in head growth, loss of purposeful hand movements, development of stereotyped hand movements (e.g., hand-wringing), psychomotor retardation, loss of social engagement, loss of acquired speech, and ataxia between the ages of 6 months to 2 years. In DSM-5 this disorder is diagnosed as Level 2 of the autism spectrum.

reversal In psychoanalysis, an impulse's change from passive to active or from active to passive, as when a destructive impulse changes from masochism to sadism or vice versa.

right to refuse treatment Legal doctrine that holds that a person cannot be forced to have treatment against his or her will unless it is a life-and-death emergency.

right to treatment Legal principle that a facility that has assumed the responsibility of providing treatment is legally obligated to provide adequate treatment for a patient.

rigidity In psychiatry, a person's resistance to change, a personality trait.

risk factor Factor that accounts for the increased risk of a disorder. *See also* attributable risk; relative risk.

ritual **1.** Formalized activity practiced by a person to reduce anxiety, as in obsessive-compulsive disorder. **2.** Ceremonial activity of cultural origin.

ritualistic behavior Automatic behavior of cultural or psychogenic origin.

RLS Restless legs syndrome.

ROC Receiver operating characteristics.

Rogers, Carl R. (1902–1987) American psychologist who developed person-centered or client-

centered theory of personality and psychotherapy, utilizing the concepts of self-actualization and self-direction; emphasized subjective experiential life of people and considered the self as a primary entity in both adjustment and maladjustment in people.

role Pattern or type of behavior developed under the influence of significant people in the person's environment. When the behavior pattern conforms to the expectations and demands of other people, it is said to be a complementary role; if it does not, it is known as a noncomplementary role. *See also* alternating role.

role playing Psychodrama technique in which persons are trained to function more effectively in their real-life roles. In the therapeutic setting of psychodrama, protagonists are free to try and to fail in their roles, for they are given the opportunity to try again until they succeed; new approaches to feared situations can thus be learned and applied outside the therapeutic setting.

Romberg sign Markedly increased swaying of the body or falling when a person stands erect with heels together and closes the eyes; typically indicates dysfunction of the dorsal column proprioceptive system, but it may also be seen in hysteria.

Rorschach Test Projective test, developed by Hermann Rorschach, in which the person is asked for associations in response to a series of inkblot pictures.

rumination Constant preoccupation with and thinking about a single idea or theme, as in obsessive-compulsive disorder.

rumination disorder of infancy Rare eating disorder of early childhood characterized by recurrent regurgitation of food in the absence of nausea, retching, or associated gastrointestinal disorder; regurgitated food is either ejected from the mouth or rechewed and reswallowed. Also called merycism.

Rush, Benjamin (1745–1813) Father of American psychiatry whose book, *Medical Inquiries and Observations upon the Diseases of the Mind* (1872), was the only American textbook on psychiatry until the end of the 19th century.

S

SAD Seasonal affective disorder.

sadism Paraphilia in which sexual gratification is achieved by inflicting pain or humiliation on a partner; first described by the French writer the Marquis de Sade (1740–1814). *See also* masochism; sadomasochistic relationship.

sadistic personality disorder Disorder in which a person shows a pervasive pattern of cruel, demeaning, and aggressive behavior whereby physical cruelty or violence is used to inflict pain on others and not to achieve some goal; such a person is fascinated by violence, weapons, injury, or torture. In DSM-III-R, this is a proposed diagnostic category requiring further study.

sadomasochistic relationship Relationship in which the enjoyment of suffering by one person (masochism) and the enjoyment of inflicting pain by the other person (sadism) are important and complementary attractions in their ongoing interaction.

sample Research term for a subset of observations selected from a population.

sanatorium Institution where patients are treated for chronic physical or mental diseases or where they are attended to during a period of recuperation.

satyriasis Morbid, insatiable sexual needs or desires in a male; may be caused by organic or

psychiatric factors. *Compare* nymphomania.

schedule of reinforcement Scheme of intermittent reinforcement in which a response is not reinforced every time it occurs. Four broad classes of intermittent schedules are based on either the number of responses required before an opportunity for reinforcement is presented or the passage of a certain amount of time before a given performance is reinforced (l) *fixed-ratio schedule,* in which reinforcement is delivered after a specified number of responses have occurred, for example, after every tenth response; (2) *variable-ratio schedule,* in which reinforcement is delivered on a schedule randomly changing around a given ratio, for example, after the third, sixth, then second response, and so on; (3) *fixed-interval schedule,* in which reinforcement is delivered contingent on the first required response after the elapse of a specified interval of time, for example, after every 10 minutes; and (4) *variable-interval schedule,* in which reinforcement is delivered after variable intervals of time have elapsed, for example, after 3, 6, then 2 hours, and so on. *See also* reinforcement.

schizoaffective disorder Psychotic disorder with signs and symptoms compatible with both a mood disorder and schizophrenia. In contrast to mood disorder with psychotic features, schizoaffective disorder is characterized by periods of at least 2 weeks without prominent mood symptoms.

schizoid fantasy Through fantasy, indulgence by a person in autistic retreat to resolve conflict and to obtain gratification; interpersonal intimacy is avoided and eccentricity serves to repel others. The individual does not fully believe in or insist on actins out the fantasies.

schizoid personality disorder Diagnostic category in DSM for persons with defects in the capacity to form social relationships but without other striking communicative or behavioral eccentricities. Also called introverted personality disorder.

schizophrenia Term introduced by Eugen Bleuler for a chronic, psychotic mental disorder of unknown etiology characterized by disturbances in thinking, mood, and behavior. The thinking disturbance is manifested by a distortion of reality, sometimes with delusions and hallucinations, accompanied by a fragmentation of associations that results in characteristic disturbances of speech; the mood disturbance includes ambivalence and inappropriate or constricted affective responses; the behavior disturbance may be manifested by apathetic withdrawal or bizarre activity. Subtypes of schizophrenia (disorganized, catatonic, paranoid, undifferentiated, and residual) are no longer formally in the DSM nomenclature, however they remain as useful descriptors. Formerly called dementia praecox.

schizophreniform disorder In DSM-5 a disorder similar to schizophrenia except that it lasts between 1 and 6 months.

schizotypal personality disorder Common term for mood disorder subtype characterized by depression, psychomotor slowing, hypersomnia, and hyperphagia that develops in autumn or winter and improves in spring and summer; not an official DSM-5 diagnosis.

schneiderian first rank symptoms Symptoms that Kurt Schneider, a German psychiatrist (1887–1967), considered to be of great pragmatic value in making a diagnosis of schizophrenia, including audible thoughts, arguing voices, thought withdrawal and insertion, delusions, and flat affect.

school phobia Young child's sudden fear of and refusal to attend school, usually considered a manifestation of separation anxiety but also may be a symptom of depression.

Schreber case One of Sigmund Freud's cases. Daniel Paul Schreber's published autobiographical account, entitled *Memoirs of a Neurotic* (1903), was analyzed by Freud (1911) and allowed him to decipher the fundamental meaning of paranoid processes and ideas, especially the relationship between repressed homosexuality and projective defenses.

scopophilia Sexual deviation in which a person achieves sexual gratification by viewing sexual acts or the genitals of others. *Compare* voyeurism.

scotoma 1. In psychiatry, a figurative blind spot in a person's psychological awareness. *See also* blind spot. **2.** In neurology, a localized visual field defect.

screening Initial patient evaluation that includes medical and psychiatric history, mental status evaluation, and diagnostic formulation to determine the patient's suitability for a particular treatment modality.

screen memory Memory that serves as a cover-up for an associated painful memory.

SCT Sentence Completion Test.

SD Standard deviation.

SDA Serotonin-dopamine antagonist.

SDS Zung Self-Rating Depression Scale.

seasonal affective disorder (SAD) Mood disorder subtype characterized by depression, psychomotor slowing, hypersomnia, and hyperphagia that develops in autumn or winter and improves in spring and summer.

seasonal pattern *See* seasonal affective disorder.

secondary gain Obvious advantage that a person gains from an illness, such as gifts, attention, and release from responsibility. *Compare* primary gain.

secondary narcissism *See* narcissism.

secondary prevention *See* preventive psychiatry.

secondary process thinking In psychoanalysis, the form of thinking that is logical, organized, reality-oriented, and influenced by the demands of the environment; characterizes the mental activity of the ego. *Compare* primary process.

secondary retarded ejaculation After previous normal sexual functioning, the inability to ejaculate during coitus. *Compare* retarded ejaculation.

secondary sexual characteristics Physical characteristics, other than the external sexual organs, that distinguish male from female.

secure attachment As described by Mary Ainsworth, the positive, happy responses of a child following the mother's appearance.

sedative Drug that produces a calming or relaxing effect through central nervous system depression; for example, barbiturates, chloral hydrate, paraldehyde, and benzodiazepines. *Compare* hypnotic.

seizure An attack or sudden onset of certain symptoms, such a convulsions, loss of consciousness, and psychic or sensory disturbances; seen in epilepsy and can be substance induced. For types of seizures, see the specific term.

selection bias Nonrepresentative sampling resulting in erroneous conclusions.

selective inattention Harry Stack Sullivan's term for inattentiveness to anxiety-generating areas.

selective mutism Persistent failure to speak in one or more specific social situations.

selective serotonin reuptake inhibitors (SSRIs) Drugs that increase the concentration of serotonin in the brain; commonly used for the treatment of depression and anxiety disorders.

self Psychophysical total of a person at any given moment, including conscious and unconscious attributes. *See also* identity.

self-actualization Abraham Maslow's term for the inborn capacity of people to direct themselves in the healthiest way, toward a level of completeness.

self-analysis Investigation of one's own psychic components, which plays a part in all analysis, although to a limited extent, since few people are capable of sustaining independent and detached attitudes to the degree necessary for this approach to be therapeutic.

self-awareness Sense of knowing oneself, particularly in terms of insight into one's own psychodynamics; a major goal of most psychotherapies.

self-defeating personality disorder Disorder in which persons avoid or undermine pleasurable experiences and are drawn to situations or relationships in which they will suffer, even when better options are available; positive personal events are responded to with guilt or depression, reasonable offers of assistance from others are rejected, and anger or rejecting responses from others are invited, leaving them feeling hurt or humiliated. Also called masochistic personality disorder.

self-discovery In psychoanalysis, the freeing of the repressed ego in a person who has been brought up to submit to the wishes of the significant others around him.

self-fulfilling prophecy Situation in which one predicts a type of behavior that eventually one carries out.

self-help groups Groups composed of persons who want to cope with a specific problem or life crisis; characterized by homogeneity and cohesion.

self-image One's conception of one's own identity, personality, and worth as a person.

self-reference Denoting a person's repeatedly referring the subject under discussion back to himself.

self-system Harry Stack Sullivan's term for a personality system designed to ward off anxiety and preserve a positive view of the self.

SEM Standard error of the mean.

senescence Aging, as determined by inherent or hereditary factors.

senile dementia Dementia secondary to diffuse cerebral atrophy associated with advancing age; onset is insidious, progression is slow and gradual, and no specific therapy is known.

senium Period of old age, generally after age 65.

sensate focus learning In the treatment of sexual disorders, exercises that focus on heightening sensory awareness to touch, sight, sound, and smell.

sensation Feeling or impression when the sensory nerve endings of any of the six senses (taste, touch, smell, sight, kinesthesia, and sound) are stimulated.

sense of self One's feeling of individuality, uniqueness, and self-direction.

sensitivity In research, the probability of a positive test result in one who has the disorder. *Compare* specificity.

sensitivity training group Group in which members seek to develop self-awareness and an understanding of group processes, rather than to gain relief from an emotional disturbance.

See also encounter group; T-group.

sensorimotor period *See* Piaget, Jean.

sensorium Hypothetical sensory center in the brain that is involved with clarity of awareness about oneself and one's surroundings, including the ability to perceive and process ongoing events in light of past experiences, future options, and current circumstances; sometimes used interchangeably with consciousness.

sensory aphasia Organic loss of ability to comprehend the meaning of words; fluid and spontaneous, but incoherent and nonsensical, speech. *See also* fluent aphasia and receptive aphasia.

sensory deprivation Lack of external stimuli and the opportunity for the usual perceptions; may be produced experimentally or may occur in real-life contexts (e.g., deep-sea diving, solitary confinement, loss of hearing or eyesight), leading to hallucinations, panic, delusions, and disorganized thinking. *See also* emotional deprivation.

sensory extinction Neurological sign operationally defined as failure to report one of two simultaneously presented sensory stimuli, despite the fact that either stimulus alone is correctly reported. Also called sensory inattention.

sensory inattention SENSORY EXTINCTION.

Sentence Completion Test (SCT) Test designed to tap patients' conscious associations to areas of functioning in which the clinician may be interested; it is composed of a series (usually 75 to 100) of sentence stems, such as "I like ....," "Sometimes I wish ...," that patients are asked to complete in their own words.

separation anxiety Reaction expressed in a child who is isolated or separated from the mother, such as tearfulness, irritability.

separation anxiety disorder In DSM, an anxiety disorder characterized by an infant's or child's fear and apprehension on being removed from the parent or parent figure.

separation-individuation As elaborated by Margaret Mahler, the phase during which the child perceives itself as distinct from the mother; as the child separates, it becomes more independent and aware of its physical and psychological autonomy and can endure separation from the mother.

serotonin 5-Hydroxytryptamine (5-HT), an endogenous indolamines synthesized from dietary tryptophan and found in the gastrointestinal tract, the platelets, and the central nervous system. Evidence indicates that this monoamine serves as a neurotransmitter substance in the central nervous system and may play an important role in such diverse functions as sleep, sexual behavior, aggressiveness, motor activity, perception (particularly pain),and mood; dysfunction in central serotonergic systems has been proposed as a cause or factor in various mental disorders, including schizophrenia and the mood disorders.

serotonin-dopamine antagonists (SDAs) Drugs that block the receptors for both serotonin and dopamine in the brain. Also known as second-generation or atypical antipsychotics. *Compare* dopamine-receptor antagonists.

serotonin-norepinephrine reuptake inhibitors (SNRIs) Drugs that inhibit the reuptake of both serotonin and norepinephrine in the brain.

serotonin syndrome An excess of serotonin in the brain that causes a syndrome of confusion, delirium, tremor, and hyperthermia which may be fatal. Caused by overdose of serotonergic

drugs.

serum levels *See* blood serum levels.

severe mental retardation *See* mental retardation.

sexual aberration Deviation from usual sexual practices. For types of sexual aberrations, see the specific term.

sexual arousal disorders Sexual dysfunction characterized by an inability to achieve a degree of sexual excitement sufficient to engage in coitus. Formerly called inhibited sexual excitement.

sexual aversion disorder Sexual desire disorder characterized by a disgust and/or avoidance of genital contact with a sexual partner.

sexual delusion Delusion centering on sexual identity, appearances, practices, or ideas.

sexual desire disorder Dysfunction characterized by a diminished or hypoactive sexual drive. Formerly called inhibited sexual desire. *See also* hypoactive sexual desire disorder; sexual aversion disorder.

sexual deviation PARAPHILIA.

sexual disorder DSM term for psychosexual disorder.

sexual drive One of the two primal drives or instincts (the other is the aggressive drive) in Freud's dual-instinct psychoanalytic theory. It operates under the pleasure–unpleasure (pleasure–pain) principle, with its main goal being the preservation and maintenance of life, in contrast to the aggressive drive that follows the repetition–compulsion principle; its unconscious psychic energy is known as libido. Also known as life instinct; Eros.

sexual dysfunction Class of sexual disorders characterized by inhibitions in sexual desire or the psychophysiological changes that characterize the sexual response cycle.

sexual identity Biologically determined sexual state. *Compare* gender identity.

sexualization Endowment of an object or function with a sexual significance that it previously did not have, or possessed to a lesser degree, in order to ward off anxieties associated with prohibited impulses or their derivatives.

sexual masochism *See* masochism.

sexual orientation distress Term that refers to a disorder characterized by persistent and marked distress about one's sexual orientation; most often characterized by a desire to acquire heterosexual rather than homosexual arousal.

sexual pain disorders Sexual dysfunction characterized by pain during coitus, such as dyspareunia and vaginismus.

sexual reassignment *See* transsexualism.

sexual sadism *See* sadism.

shaman Healer in a primitive tribe who uses supernatural or spiritual powers to cure the sick.

shame Failure to live up to self-expectation; often associated with fantasy of how person will be seen by others. *See also* guilt.

shaping Operant conditioning technique, used in behavior therapy, in which a behavioral goal is reached in stepwise fashion through selective reinforcement of closer and closer approximations of the desired behavior.

shared paranoid disorder *See* shared psychotic disorder.

shared psychotic disorder Rare disorder in which the delusion system of the patient has developed out of a close relationship with another person who had a previously established, similar delusional system. In DSM-5, this condition is called "delusional symptoms in partner of individual with delusional disorder." It has historically been referred to as shared paranoid disorder, induced psychotic disorder, folie à deux, and double insanity.

shell shock Obsolete World War I term for an acute mental disorder arising as a consequence of combat experience. *See* posttraumatic stress disorder.

shift work circadian rhythm disorder Characterized by sleep pattern disruptions that result from night shift work; associated with insomnia, excessive sleepiness, or both.

shock treatment Obsolete term for convulsive therapy.

short-duration depressive episode A DSM-5 diagnosis characterized by depressed mood lasting between 4 and 14 days that is nonrecurring.

short-term memory Reproduction, recognition, or recall of perceived material within seconds to minutes after the initial presentation. Compare immediate memory; long-term memory.

shyness disorder AVOIDANT DISORDER OF CHILDHOOD OR ADOLESCENCE.

sibling One of two or more individuals having one common parent.

sibling rivalry Competition among siblings for the attention, affection, and esteem of their parents or for other recognition or reward.

sick role Use of illness or complaints about bodily functioning as a means of escaping personal failure and the expectation to succeed, or as a way of avoiding other social responsibilities. *See also* hypochondriasis.

side effect Action of a drug other than the desired therapeutic effect.

sign Objective evidence of a disease or disorder. *Compare* symptom.

signal anxiety Sigmund Freud's term describing a type of anxiety characteristic of the psychoneuroses, in which a child learns to anticipate and react to the possibility of a traumatic situation before it becomes traumatic.

significant differences *See* statistical significance.

simple phobia Phobic disorder characterized by fear and avoidance of an object or situation not included in agoraphobia or social phobia.

simultanagnosia Impairment in the perception or integration of visual stimuli appearing simultaneously.

sitophobia Abnormal fear of eating.

skin-picking disorder *See* excoriation disorder.

Skinner, B. F. (1904–1990) American behavioral psychologist and author noted for his work on operant conditioning.

sleep Recurrent temporary period of relative unconsciousness, characterized by a reversible cessation of waking Sensorimotor activity are accompanied by characteristic electroencephalographic changes; divided into REM (rapid eye movement) and non-REM (NREM) sleep.

sleep disorder Abnormality of the sleep–wake cycle. Sleep disorders include dyssomnias (insomnia disorder, primary insomnia, hypersomnia disorder, sleep–wake schedule disorder) and parasomnias (dream anxiety disorder, sleepwalking disorder, sleep terror disorder).

sleep-related eating disorder Sleep disorder characterized by the inability to fall back asleep after awakening without having something to eat or drink; predominantly affects infants and children.

sleep-related hallucinations Hallucinations that occur at sleep onset (hypnagogic) or on awakening (hypnopompic); hallucinations are often visual and common in patients with narcolepsy.

sleep-related hypoventilation Dyssomnia characterized by recurrent episodes of shallow breathing during sleep associated with insomnia and/or frequent arousal during the breathing disturbances. *See also* breathing-related sleep disorders.

sleep terror disorder Sleep disorder in which there are repeated episodes of abrupt awakenings from sleep, usually beginning with a panicky scream; characterized by verbalizations, confusion, autonomic activity, and poor recall of the dream. A non-REM sleep arousal disorder in DSM-5; also called pavor nocturnus; night terror. *See also* parasomnias.

sleep–wake schedule disorder Sleep disorder in which persons cannot sleep when they wish to sleep, but are able to sleep at other times; correspondingly, they cannot be fully awake when they want to be fully awake, but they are able to be awake at other times. Term used in ICD.

sleepwalking disorder DSM-5 term for somnambulism; a non-REM sleep arousal disorder. *See also* parasomnias.

slow-wave sleep (SWS) Non-REM sleep stages 3 and 4; so termed because of the characteristic wave appearance on the electroencephalogram (EEG) record. Also called delta sleep.

SNRI Serotonin-norepinephrine reuptake inhibitor.

social adaptation Adjustment to the whole complex of interpersonal relationships; the ability to live and express oneself in accordance with society's restrictions and cultural demands. *See also* adaptational approach.

social anxiety disorder *See* social phobia.

social breakdown syndrome Signs and symptoms indicating that the long-term institutionalization of a mental patient has itself contributed to the patient's symptoms; for example, development of progressive social and vocational incompetence, excessive passivity, and learning the chronic sick role.

social control Influence exerted on a person or a group of persons to conform to the demands or expectations of a society or any of its representative institutions, agencies, or organizations.

social instinct *See* herd instinct.

socialization Process through which an individual becomes a competent member of society; the condition in which inner impulses, that is, instincts, are expressed in conformity with cultural demands of the environment.

social learning theory Approach to behavior therapy that incorporates the elements of respondent and operant conditioning, but views cognitive processes as important mediators in such learning; relies on role modeling, identification, and human interactions.

social network therapy Type of therapy in which the therapist assembles all persons (relatives, friends, social relations, work relations) who have emotional or functional significance in the patient's life to provide feedback and support to the patient; all or some of

the social network may be assembled at any given time. *See also* extended-family therapy.

social phobia Any of a variety of phobic disorders characterized by a fear of being observed by others (e.g., fear of public-speaking, blushing, eating in public, writing in front of others, using public lavatories). Also called social anxiety disorder. DSM-5 includes a "performance only" diagnostic category if the fear is limited to speaking or performing in public.

social psychiatry Branch of psychiatry concerned with ecological, sociological, and cultural variables that engender, intensify, or complicate maladaptive patterns of behavior and their treatment.

social (pragmatic) communication disorder Persistent failure to use verbal and nonverbal communication for social purposes; differentiated from Autism Spectrum Disorder

social therapy Rehabilitation therapy used with psychiatric patients to improve social functioning; for example, occupational therapy, therapeutic community, recreational therapy, milieu therapy, attitude therapy.

sociobiology Systematic study of the biological basis of social behavior, pioneered by E. O. Wilson.

sociology Scientific study of group behavior and social organization.

sociometry Study and measurement of the interpersonal psychological structure of a group or society.

sociopath Obsolete term for a person with an antisocial personality disorder.

sociopathic behavior Former term for antisocial behavior.

sociotherapy Any treatment modality whose primary emphasis is on socioenvironmental and interpersonal factors; for example, the therapeutic community.

sodomy Anal intercourse. Legally, the term may include other sexual deviations as well, such as bestiality.

soldier's heart DA COSTA SYNDROME.

soma The body.

somatic delusion Delusion pertaining to the functioning of one's body.

somatic hallucination Hallucination involving the perception of a physical experience localized within the body.

somatic therapy Treatment of psychiatric disorders with biological, organic, or physical treatments.

somatic symptom disorder Disorder characterized by multiple, recurrent physical complaints with no apparent physical cause and a preoccupation with the false belief of having a serious disease; symptoms have no demonstrable organic findings, are often misinterpreted by the patient, and are apparently linked to psychological factors. Previously called somatization disorder; also known as hypochondriasis; Briquet syndrome.

somatization disorder Somatoform disorder characterized by recurrent and multiple physical complaints with apparent physical cause; now called somatic symptom disorder in DSM-5. Also called Briquet syndrome.

somatoform disorders Term previously used for a group of mental disorders characterized by a marked preoccupation with the body and fears of disease; production of the symptoms is linked to psychological factors or conflicts but is not under voluntary control. Previously in DSM-IV, somatoform disorders included somatization disorder, conversion disorder, body

dysmorphic disorder, and somatoform pain disorder. This group of disorders is now referred to as somatic symptom and related disorders. *See also* factitious disorder; malingering.

somatopagnosia Inability to recognize a part of one's body as one's own (also called ignorance of the body and autotopagnosia).

somatotype Physical build and structure of a person; one's physique.

somnambulism Repeated episodes of arising from bed during sleep, usually during the first third of the night during non-REM (stages 3 and 4 sleep); on awakening, the person has amnesia for the episodes. More frequently seen in males than in females, most commonly in children, who generally outgrow it; in adults, often associated with other psychiatric disturbances, such as schizoid personality disorder and schizophrenia.

somnolence Pathological sleepiness or drowsiness from which one can be aroused to a normal state of consciousness.

spatial agnosia Inability to recognize spatial relations.

speaking in tongues Expression of revelatory message through unintelligible words; not considered a disorder of thought if associated with practices of specific Pentecostal religions. *See also* glossolalia.

specific developmental disorder *See* specific learning disorder.

specific learning disorders A class of neurodevelopmental disorders beginning in early childhood that are characterized by persistent difficulty or delays in learning skills in reading, written expression, and/or mathematics, which are inconsistent with the overall intelligence of the child. Specific learning deficit is indicated by the phrase, "with impairment in," either reading (dyslexia), written expression, or mathematics (dyscalculia). Previously called specific developmental disorder. *Compare* pervasive developmental disorder.

specificity In research, the probability of a negative test result in one who does not have the disorder.

specific reading disability DEVELOPMENTAL READING DISORDER.

speech disturbance Any of a variety of language and communication disorders not due to impaired function of speech muscles or organs of articulation.

speech sound disorder Language or speech disorder in which there is a consistent failure to correctly articulate a speech sound at the developmentally appropriate age in the absence of any underlying anatomical, physiological, auditory, or neurological abnormalities. Also called developmental articulation disorder, functional speech disorder, infantile perseveration, infantile articulation disorder, and phonological disorder.

splitting Defense mechanism in which external objects are divided into "all good" and "all bad," accompanied by the abrupt shifting of an object from one extreme category to the other.

SSRI Selective serotonin reuptake inhibitor.

stammering STUTTERING.

standard deviation (SD) Statistical measure of variability within a set of values so defined that, for a normal distribution, about 68 percent of the values fall within one SD of the mean, and about 95 percent lie within two SDs of the mean; sometimes represented by the Greek letter sigma (σ).

standard error of the mean (SEM) Statistical measure of how variable, that is, how reliable, a mean value calculated on the basis of a given sample is as an estimate of the mean of

the population from which the sample was selected; obtained by dividing the standard deviation of the population by the square root of the number of measures in the sample.

standardized score Deviation of a score from its group mean expressed in standard deviation units. Also called z-score.

Stanford–Binet Intelligence Scale Primarily a verbal test of intellectual functioning, administered individually to children and adults, which emphasizes problem solving; one of the most widely used I.Q. tests.

startle reaction Reflex response to a sudden intense stimulus, consisting of a diffuse motor response involving flexion movements of the trunk and extremities and associated with a sudden increase in alertness; occurs in normal persons and in anxiety disorders.

state-dependent learning or memory Facilitated recall of information when one is in the same internal state (e.g., drug state) or external environment that the information was first acquired.

statistical inference Research term used to make a general statement about a study based on a small group or to decide whether to attribute differences between two groups to chance or to nonchance factors.

statistical significance Research term referring to an experimental result, based on a sample of observations, which demonstrates an outcome or effect of sufficient magnitude such that the probability of that result being due to chance alone is less than 5 percent. *See also* Type I error.

status epilepticus Continuous epileptic seizures. *See also* epilepsy.

stereotypy Continuous mechanical repetition of speech or physical activities; observed in cases of catatonic schizophrenia.

stereotypy and habit disorder *See* stereotypic movement disorder.

stereotypic movement disorder Motor disorder consisting of intentional, repetitive, and typically rhythmic behaviors that are nonfunctional, that is, serve no constructive, socially acceptable purpose (e.g., hand flapping, body-rocking, head-banging, face-slapping, skin picking). Movements appear self-soothing or self-stimulating, but in certain cases may also result in self-injury.

stilted speech Excessively formal, stiff, stylized, or pompous speech; overly polite, distant, or antiquated speech.

stimulant Drug that stimulates the central nervous system to produce increased psychomotor activity; for example, methylphenidate (Ritalin), caffeine, amphetamines.

stimulus Any event, internal or external to the person, which elicits some kind of nervous system activity or response.

Stockholm syndrome Syndrome in which captives identify with, and have sympathy for, their captors on whom they are dependent for survival; seen in terrorist-hostage situations. The major defense mechanism, as described by Anna Freud, was known as identification with the aggressor.

STP Slang name for 2,5-dimethoxy-4-methylamphetamine.

street drugs Readily available, illegally acquired psychoactive substances; for example, marijuana, heroin, cocaine, crack.

strephosymbolia Reading and writing disability characterized by confusion between similarly formed but oppositely oriented letters (as in b and d, p and q), with tendency to reverse the

order of letters and words.

stress Circumstance in a person's life that causes tension or dysphoria. Stressors may be mild (e.g., termination of a friendship) or catastrophic (e.g., suicide of a spouse). DSM contains a Severity of Psychosocial Stressor Scale for children, adolescents, and adults. Also called psychosocial stressor.

stress immunity Failure to react to emotional stress.

stress reaction Disturbed psychological or physiological functioning encountered when strong, involuntary, often unconscious internal impulses press for action that conflict with one's conscious, reality-oriented behavior.

stroke CEREBROVASCULAR ACCIDENT.

structured interactional group psychotherapy Group therapy developed by Harold Kaplan and Benjamin Sadock, in which the therapist provides a structural matrix for the group's interactions, the most important of which is that a different member of the group is the focus of the interaction in each session.

structured interview 1. Interview in which the same questions are asked of all patients. **2.** Interview in which the interviewer asks specific questions about specific topics.

Studies on Hysteria Title of a book by Josef Breuer and Sigmund Freud (1895) that described the cathartic method of treatment and the beginnings of psychoanalysis; demonstrated the psychological origins of hysterical symptoms and the possibility of effecting a cure through psychotherapy.

stupor 1. State of decreased reactivity to stimuli and less than full awareness of one's surroundings; as a disturbance of consciousness, it indicates a condition of partial coma or semi coma. **2.** In psychiatry used synonymously with mutism and does not necessarily imply a disturbance of consciousness; in catatonic stupor, patients are ordinarily well aware of their surroundings.

stuttering Speech disorder characterized by repetitions or prolongations of sounds, syllables, or words, or by unusual hesitations and pauses that disrupt the normal flow of speech. May often be associated with habitual eye blinking, facial grimacing, head jerks, or abnormal body movements. Also called stammering; child-onset fluency disorder.

subcoma therapy *See* insulin coma therapy.

subconscious Obsolete term for the preconscious and the unconscious.

subjectivity State in which evaluation and interpretation are influenced by one's own feelings and thinking.

sublimation Unconscious defense mechanism in which the energy associated with unacceptable impulses or drives is diverted into personally and socially acceptable channels; unlike other defense mechanisms, it offers some minimal gratification of the instinctual drive or impulse.

substance abuse Called substance use disorder in DSM-5; diagnosed according to the substance used, for example, alcohol, cocaine, cannabis, hallucinogens, inhalants, opioids, sedatives, stimulants, or tobacco. *See also* drug abuse.

substance use disorder A maladaptive pattern of substance use leading to impairments in functioning characterized by the substance being used in larger amounts over time; unsuccessful efforts to cut down use; cravings; and symptoms of tolerance and withdrawals.

substance withdrawal Cessation or reduction of intake of a substance regularly used to induce a state of intoxication.

substitution Unconscious defense mechanism in which a person replaces an unacceptable wish, drive, emotion, or goal with one that is more acceptable.

succinylcholine Powerful skeletal muscle relaxant used in anesthesia and as a muscle relaxant in electroconvulsive treatment.

suggestibility State of uncritical compliance with influence or of uncritical acceptance of an idea, belief, or attitude; commonly observed among persons with hysterical traits.

suggestion Process of influencing a person to the point of uncritical acceptance of an idea, belief, or other cognitive process.

suicidal ideation Thoughts or acts of taking one's own life.

suicide Act of taking one's own life. Although the underlying factors that lead a person to take his or her own life may not necessarily be fully understood by that person, the act of suicide is considered to be both voluntary and intentional.

suicide cluster *See* cluster suicide.

Sullivan, Harry Stack (1892–1949) American psychiatrist and psychoanalyst best known for his interpersonal theory of psychiatry.

sundowner's syndrome SUNDOWNING.

sundowning Syndrome in the elderly that usually occurs at night and is characterized by drowsiness, confusion, ataxia, and falling as the result of being overly sedated with medications. Also called sundowner's syndrome.

superego One of the three components of the psychic apparatus in the freudian structural framework, the others being the ego and id. It represents the internalized values, ideals, and moral attitudes of society and has psychic functions that are expressed in guilt, self-criticism, and conscience. It develops through the child's identification with his parents, and the severity of its prohibitions or demands is said to be related to the intensity and extent of resolution of the Oedipus complex. *See also* conscience.

supportive psychotherapy Psychotherapy that seeks to strengthen patients' defenses and to provide them with reassurance, rather than to probe deeply into their conflicts.

suppression Conscious act of controlling and inhibiting an unacceptable impulse, emotion, or idea; differentiated from repression in that repression is an unconscious process.

surrogate One who takes the place of another; a substitute.

susto Culture-specific mental syndrome seen in Latin America; characterized by an intense fear of evil supernatural forces.

SWS Slow-wave sleep.

symbiosis Dependent, mutually reinforcing relationship between two persons. Generally a normal, constructive characteristic of the infant–mother relationship, it can also occur in a destructive context, as between two mentally ill persons who reinforce each other's pathology or in a mother–infant relationship that induces in the child an intense separation anxiety, autism, and severe regression (symbiotic psychosis).

symbiotic psychosis Disorder in children characterized by an inability to separate psychologically or physically from the mother; such children experience confusion, intense separation anxiety, and an inability to function.

symbolization Unconscious defense mechanism in which one idea or object comes to stand for another because of some common aspect or quality in both; based on similarity and association; the symbols formed protect the person from the anxiety that may be attached to the original idea or object.

sympathetic nervous system The part of the autonomic nervous system innervating the viscera by outflow from thoracolumbar segments of the spinal cord. Its physiological effects (e.g., tachycardia, peripheral vasoconstriction, pupillary dilation, decreased bowel motility) are predominantly adrenergic, and they characterize the fight-or-flight reaction that occurs in response to dangerous or threatening situations.

sympathomimetic drug Drug whose effects mimic the responses to stimulation of the sympathetic nervous system; for example, amphetamine, epinephrine.

sympathy Compassion for or sharing of another person's feelings, ideas, and experiences. *Compare* empathy.

symptom Any abnormality indicative of a mental or physical disorder experienced by the patient but not observable by a physician. *Compare* sign.

symptomatic psychosis Brain dysfunction as a result of some general physical disease, the most common feature of which is a delirious state; may occur in the course of an acute infectious disease, such as pneumonia, influenza, or pelvic infections following childbirth.

symptom formation Unconscious psychic process in which a repressed impulse is indirectly released and manifested through a symptom; may be regarded as a kind of compromise reflecting a partial satisfaction of both the forbidden impulse and the ego reaction against it.

symptom substitution Phenomenon in which a set of symptoms that are removed directly in therapy, without regard for the unconscious conflicts responsible for their formation, are replaced by new symptoms. It has constituted a theoretical objection to such modalities as behavior therapy and hypnotherapy.

synapse Specialized structure at which one neuron makes contacts with the receptor elements in the adjacent neuron.

syncretic thought Jean Piaget's term for the prelogical, egocentric, solipsistic thinking that characterizes early childhood mentation. *See also* primary process.

synesthesia Condition in which the stimulation of one sensory modality is perceived as sensation in a different modality, as when a sound produces a sensation of color.

syphilis A chronic bacterial disease caused by the spirochete *Treponema pallidum* (*T. pallidum*). Associated with delirium, dementia, mania, psychosis, personality change, and depression.

syntactical (central) aphasia Aphasia characterized by difficulty in understanding spoken speech, associated with gross disorder of thought and expression.

syntaxic mode of experience Harry Stack Sullivan's term denoting the highest level of interpersonal communication, which is based entirely on reality and which is considered the desirable outcome of healthy development. *See also* secondary process.

syntropy Adolf Meyer's term characterizing healthy or wholesome relationships.

systematic desensitization Form of behavior therapy developed by Joseph Wolpe and others in which anxiety-evoking stimuli are represented to a patient in a state of deep muscle relaxation in an attempt to weaken the bond between the stimuli and the anxiety; particularly

useful in the treatment of phobias. *See also* reciprocal inhibition and desensitization.

systematized delusion Group of elaborate delusions related to a single event or theme.

T

taboo Restrictions or prohibitions on behavior imposed by a culture, usually involving a dangerous, forbidden, or unclean entity.

tachylogia LOGORRHEA.

tactile amnesia *See* neurological amnesia.

tactile hallucination Hallucination primarily involving the sense of touch. Also called haptic hallucination.

talion Retaliation; the fear that all bad deeds, intentional or accidental, will be punished in kind; that is, an eye for an eye and a tooth for a tooth. In psychoanalysis, the fear of castration or punishment for oedipal strivings, as in the Oedipus complex.

tangentiality Disturbance in which the person replies to a question in an oblique, digressive, or even irrelevant manner and the central idea is not communicated; failure to communicate the central idea distinguishes tangentiality from circumstantiality, in which the goal idea is reached in a delayed or indirect manner. The term has also been used synonymously with loosening of associations and speech derailment.

taphophobia Abnormal fear of being buried alive.

Tarasoff decision California Supreme Court decision that a physician or psychotherapist who has reason to believe that a patient may injure or kill someone must notify the potential victim, his or her relatives or friends, or the authorities.

tardive dyskinesia Late-appearing extrapyramidal syndrome associated with antipsychotic drug use and characterized by stereotyped involuntary movements of the lips, jaw, and tongue (tardive oral dyskinesia) and by other involuntary dyskinetic movements; symptoms may remit eventually or persist indefinitely after discontinuation of the medication, and the condition does not respond to treatment with antiparkinsonism drugs.

TAT Thematic Apperception Test.

telepathy Extrasensory perception of the mental activities of another person.

telepsychiatry Psychiatric practice using the electronic transmission of images.

temperament Inherent, constitutional predisposition to react, in a certain way to stimuli; variations in temperament are evident very early in life.

temporal lobe epilepsy COMPLEX PARTIAL SEIZURE.

temporary admission Hospitalization in which a person is admitted on the recommendation of one physician (which must be confirmed by a psychiatrist, upon admittance), for no more than 15 days if against his or her will; applies to patients who are so senile or confused that they require hospitalization and are notable to make decisions on their own, or so acutely disturbed that they must be immediately admitted to a psychiatric hospital on an emergency basis.

tension Physiological or psychic arousal, uneasiness, or pressure toward action; an unpleasurable alteration in mental or physical state that seeks relief through action.

terminal insomnia Early morning awakening or waking up at least 2 hours before planning to. *Compare* initial insomnia; middle insomnia.

terrorism The use or threat of violence to instill fear with the intent to intimidate or coerce governments or societies in the pursuit of political, religious, or idealogical goals.

tertiary prevention *See* preventive psychiatry.

testamentary capacity One's competency to make a will. Three psychological abilities are necessary to demonstrate this competency: (1) the nature and extent of their bounty (property); (2) that they are making a will; and (3) who their natural beneficiaries are, for example, their spouse, children, and other relatives.

testimonial privilege Right to maintain secrecy or confidentiality in the face of a subpoena. Privilege belongs to the patient, not to the physician.

test of significance Statistical test used to determine whether a difference in the data obtained is due to chance or is meaningfully related to some specific factor or variable (statistically significant).

tetrahydrocannabinol (THC) Chemical substance produced by the Indian hemp plant (*Cannabis sativa*), believed to be responsible for most of the characteristic psychological effects of marijuana. Street drugs alleged to be THC frequently contain phencyclidine (PCP), a veterinary tranquilizer that may produce extreme excitement, visual disturbances, and delirium.

T (training)-group Group that emphasizes training in self-awareness and group dynamics.

thanatology Study of death and matters leading up to and following it.

thanatophobia Abnormal fear of death.

Thanatos A term used by Freud to refer to the death instinct. *Compare* eros.

THC Tetrahydrocannabinol.

Thematic Apperception Test (TAT) Projective psychological test in which subjects supply interpretations of a series of ambiguous life situation drawings, based on their own feelings and attitudes and reflecting their unconscious conflicts and defensive structures.

theory Plausible or scientifically acceptable general principle or body of principles offered to explain phenomena.

therapeutic agent Anything that promotes healing, such as a substance, person, or apparatus.

therapeutic alliance Conscious contractual relationship between therapist and patient in which each implicitly agrees that they need to work together to help the patient with his or her problems; involves a therapeutic splitting of the patient's ego into observing and experiencing parts. A good therapeutic alliance is especially necessary for the continuation of treatment during phases of strong negative transference.

therapeutic atmosphere All therapeutic, maturational, and growth-supporting agents, that is, cultural, social, and medical.

therapeutic community Institutional treatment setting designed with an emphasis on the importance of socioenvironmental and interpersonal influences in the therapy, management, and rehabilitation of the hospitalized mental patient. *See also* milieu therapy.

therapeutic group Group of patients joined together under the leadership of a therapist for the purpose of working together for psychotherapeutic ends; specifically, for the treatment of each patient's mental disorders.

therapeutic role Position in which one aims to treat, bring about an improvement, or provide alleviation of a distressing condition or state.

therapeutic window Range of blood plasma levels of a drug within which optimal therapeutic effects occur, based on such factors as age, sex, race, genetic makeup, and concurrent medications.

therapist *See* psychotherapist.

tetracyclic antidepressants One group of antidepressant drugs so named because of their four-ringed structure. *See also* tricyclic antidepressants.

thinking *See* cognition.

thinking compulsion INTELLECTUALIZATION.

thioxanthene derivative One of a class of antipsychotic drugs closely related, both structurally and pharmacologically, to the phenothiazines; for example, chlorprothixene (Taractan), thiothixene (Navane).

third ear Ability to make use of intuition, sensitivity, and awareness of subliminal cues to interpret clinical observations of individual and group patients; a term introduced by the German philosopher Friedrich Nietzsche and later used in analytic psychotherapy by Theodor Reik.

third-party payer Health or medical care insurer.

thought broadcasting Feeling that one's thoughts are being broadcast or projected into the environment. *See also* thought withdrawal.

thought control *See* delusion of control.

thought deprivation BLOCKING.

thought disorder Any disturbance of thinking that affects language, communication, or thought content; the hallmark feature of schizophrenia. Manifestations range from simple blocking and mild circumstantiality to profound loosening of associations, incoherence, and delusions; characterized by a failure to follow semantic and syntactic rules which is inconsistent with the person's education, intelligence, or cultural background.

thought insertion Delusion that thoughts are being implanted in one's mind by other people or forces.

thought latency The period of time between a thought and its verbal expression. Increased in schizophrenia (*see* blocking) and decreased in mania (*see* pressured speech).

thought process disorder Symptom of schizophrenia, involving the intellectual functions, manifested by irrelevance and incoherence of verbal productions; ranges from simple blocking and mild circumstantiality to total loosening of associations, as in word salad.

thought withdrawal Delusion that one's thoughts are being removed from one's mind by other people or forces. *See also* thought broadcasting.

three-cornered therapy CO-THERAPY.

Three Essays on the Theory of Sexuality Title of a book by Sigmund Freud (1905) in which the libido theory was applied to the successive phases of psychosexual maturation in the infant, child, and adolescent; made possible the integration of a vast diversity of clinical observations and promoted the direct observation of child development.

Tic Seemingly involuntary sudden, brief, and rapid motor movements or vocalizations stimulated by an irresistible internal urge. Common motor tics include eye blinking, head-jerking, and grimacing; common vocal tics include coughing, throat-clearing, and grunting. Tics may have symbolic meaning. *See also* Tourette disorder.

tic disorders Predominately psychogenic class of motor disorders consisting of involuntary, spasmodic, stereotyped movements of a small group of muscles or vocal sounds; seen prominently in moments of stress or anxiety, and rarely as a result of organic disease. The two types of tic disorder in DSM-5 include Tourette disorder and persistent chronic motor or vocal tic disorders.

timidity Inability to assert oneself for fear of some fancied reprisal, even in the absence of objective evidence of potential harm.

tinnitus Noises in one or both ears, such as ringing, buzzing, or clicking; a rare side effect of some tricyclic antidepressant drugs.

TMS Transcranial magnetic stimulation.

tobacco dependence *See* nicotine dependence.

toilet training Program of teaching a child to control bladder and bowel functions; attitudes of both parent and child during the period may have important psychological implications for the child's later development.

token economy Program applying principles of operant conditioning to the management of an institutional setting, such as a ward or a classroom, during which reinforcement of desirable behavior is provided in the form of tokens or credits (conditioned reinforcers) that may be exchanged for a variety of positive reinforcers, such as food, television time, or a weekend pass.

tolerance Phenomenon in which, after repeated administration, a given dose of a drug produces a decreased effect or increasingly larger doses must be administered to obtain the effect observed with the original dose. *Compare* drug dependence.

tomography *See* computed tomography; positron emission tomography.

tonic convulsion Convulsion in which the muscle contraction is sustained.

topectomy Psychosurgical procedure involving ablation of a small and specific area of the frontal cortex for the treatment of a chronic psychotic patient.

torture The intentional infliction of physical and psychological suffering with the intent to gain information or to intimidate individuals or communities.

Totem and Taboo Title of a book by Sigmund Freud (1913) in which he applied his psychoanalytic concepts to the data of anthropology. He was able to afford much insight into the meaning of tribal organizations and customs, especially by invoking the Oedipus complex and the characteristics of magical thought as discovered from his studies of the unconscious.

Tourette syndrome Rare motor disorder, first described by a Paris physician, Gilles de la Tourette, that has its onset in childhood and is characterized by multiple involuntary muscle movements (tics), motor incoordination, and the presence of at least one vocal tic, such as echolalia and coprolalia.

toxicity State of being poisonous.

toxic psychosis Psychosis caused by toxic substances produced by the body or introduced into it in the form of chemicals or drugs.

toxophobia Abnormal fear of being poisoned.

trailing phenomenon Perceptual abnormality associated with hallucinogenic drugs in which moving objects are seen as a series of discrete and discontinuous images.

trainable Capable of achieving a certain degree of self-care and social adjustment at home and vocational usefulness in a closed supervised setting; describes the moderately mentally

retarded (I.Q. of 35–40 to 50–55). *Compare* educable.

training group *See* T-group.

trait Distinguishing quality as an observed or measured characteristic of an individual.

trance Sleep-like state of reduced consciousness and activity.

tranquilizer Psychotropic drug that induces tranquility by decreasing anxiety or agitation, usually without clouding consciousness; formerly classified as major and minor tranquilizers which are now called antipsychotic and antianxiety agents, respectively.

transactional analysis Eric Berne's system of psychodynamic psychotherapy that focuses on the interactions (transactions) between the patient and the therapist in the treatment session and between the patient and others in a social environment; used in both individual and group psychotherapy. The system includes four components: (1) structural analysis of intrapsychic phenomena; (2) transactional analysis proper, the determination of the currently dominant ego state (parent, child, or adult) of each participant; (3) game analysis, identification of the games played in their interactions and of the gratifications provided; and (4) script analysis, uncovering of the causes of the patient's emotional problems.

transcranial magnetic stimulation (TMS) A form of brain stimulation in which a rapidly changing magnetic field is applied to the scalp to induce small, focused electrical currents in superficial cortex. Used for treatment-resistant depression.

transcultural psychiatry Branch of psychiatry concerned with the mental health of members of different cultures, that is, what may be considered normal for one culture may be unacceptable to another.

transference Unconscious tendency of a person to assign to others in the present and immediate environment those feelings and attitudes originally linked with significant figures in the person's early life, for example, identification of the therapist with a parent; the transference may be negative (hostile) or positive (affectionate). Analysis of transference phenomena is used as a major therapeutic tool in both individual therapy and group therapy to help patients better understand and gain insight into their behavior and its origins. *See also* countertransference; transference neurosis.

transference neurosis Phenomenon occurring in psychoanalysis in which the patient develops a strong emotional attachment to the therapist as a symbolized nuclear familial figure; repetition and the depth of that misperception or symbolization characterize it as a transference neurosis.

transsexualism Gender identity disorder in which a person has a desire to be of the opposite sex. Some transsexuals, many of whom have adopted the role of the opposite sex since childhood, have successfully undergone sex-changing surgical procedures, accompanied by intensive hormonal therapy and psychotherapy. *Compare* transvestism.

transvestism, transvestitism Paraphilia characterized by dressing in the clothing of the opposite sex for the purpose of sexual arousal; usually seen in men with a strong desire to appear as women. Also called cross dressing. *Compare* transsexualism.

trauma In psychiatry, a significant, upsetting experience or event that may precipitate or aggravate a mental disorder.

traumatic brain injury Damage to the structure of the brain caused by physical force that affects cognitive functioning with associated changes in mood and behavior. Most commonly

associated with war injuries.

traumatic neurosis *See* posttraumatic stress disorder.

treated prevalence Number of persons being treated for a disorder in a defined geographic area.

tremor Rhythmical alteration in movement, which is usually faster than one beat a second; typically, tremors decrease during periods of relaxation and sleep and increase during periods of anger and increased tension.

trend of thought Thinking that centers on a particular idea associated with an affective tone.

triad Interactional relationship among three persons. Its prototype, the father-mother-child relationship, may evolve projectively in group therapy.

triadic therapy, triangular therapy CONJOINT THERAPY.

trichotillomania Compulsion to pull out one's hair.

tricyclic antidepressants One of a group of antidepressant drugs so named because of their three-ringed structure; currently the most widely used drugs for the treatment of depression, although anticholinergic side effects are prominent; for example, imipramine (Tofranil), amitriptyline (Elavil).

triskaidekaphobia The fear of the number 13.

trisomy Chromosomal aberration in which there are three chromosomes of one type instead of the usual two in an otherwise diploid cell, as in Down syndrome.

trisomy 21 syndrome DOWN SYNDROME.

true insight Understanding of the objective reality of a situation coupled with the motivational and emotional impetus to master the situation or change behavior.

t-test Statistical procedure designed to compare two sets of observations.

Tuke, William (1732–1822) British pioneer in the treatment of mental patients without the use of physical restraints.

Turner syndrome Disorder secondary to a chromosomal anomaly in which affected females typically have an XO sex chromosome constitution (instead of the usual XX) and a total of 45 chromosomes (instead of the usual 46). Classical somatic features include short stature, webbed neck, distinctive facies, sexual infantifism, and cubitus valgus. The incidence of mental retardation is not increased and severe psychopathic manifestations are uncommon, in contrast to the increased frequency of those disorders in XXY males. Also called ovarian dysgenesis; gonadal dysgenesis. *See also* Klinefelter syndrome.

twilight state Disturbed consciousness with hallucinations.

twin studies Studies that are useful in separating genetic and environmental influences and in looking for protective and precipitating factors.

twirling Sign present in autistic children who continually rotate in the direction in which their head is turned.

Type 1 error Research term referring to the rejection of the null hypothesis when it should have been retained. When the statistically calculated probability (p value) of making a Type 1 error on the basis of an experimental result is less than 0.05, the result is considered statistically significant; that type of error and the arbitrarily chosen 0.05 level of significance are sometimes called alpha (α) error. *Compare* Type 2 error.

Type 2 error Research term referring to the failure to reject a false null hypothesis, that is,

concluding that any difference or effect demonstrated in an experiment is no greater than that expected by chance alone when, in fact, there is an actual difference or effect. Also called beta (β) error. *Compare* Type 1 error.

Type A personality As defined by Meyer Friedman and Ray Rosenman, cardiologists, an action-oriented individual who struggles to achieve poorly identifiable goals by means of competitive hostility; has increased amounts of low-density lipoprotein, serum cholesterol, triglycerides, and 17-hydroxycorticosteroids; and tends to develop coronary heart disease.

Type B personality As defined by Meyer Friedman and Ray Rosenman, cardiologists, a relaxed and less-aggressive individual (than Type A) who tends to strive less vigorously to achieve goals.

tyramine Sympathomimetic amine that acts by displacing stored transmitter from adrenergic axonal terminals; its degradation is normally aided by monoamine oxidase (MAO). Ingestion of food containing tyramine (e.g., cheese, beer, wine, yeast) by a patient on MAO inhibitors can provoke a potentially fatal hypertensive crisis.

U

UCR Unconditioned response or reflex.

UCS Unconditioned stimulus.

ultradian rhythm *See* biological rhythm.

ululation Incoherent crying of a psychotic or hysterical patient.

unconditioned response, unconditioned reflex (UCR) *See* conditioning.

unconditioned stimulus (UCS) *See* conditioning.

unconscious 1. One of three divisions of Sigmund Freud's topographic theory of the mind (the others being the conscious and preconscious) in which the psychic material is not readily accessible to conscious awareness by ordinary means; its existence may be manifested in symptom formation, in dreams, or under the influence of drugs. **2.** In popular but more ambiguous usage, any mental material not in the immediate field of awareness. **3.** Denoting a state of unawareness, with lack of response to external stimuli, as in a coma.

underachiever One who fails to perform at the level of his or her known potential or capability.

undersocialized Characterized by the absence of adequate social bonds to others, as a type of conduct disorder.

undifferentiated schizophrenia Type of schizophrenia with prominent psychotic symptoms that do not fall into the other schizophrenic subtypes; no longer included in DSM-5.

undoing Unconscious primitive defense mechanism, repetitive in nature, by which a person symbolically acts out in reverse something unacceptable that has already been done or against which the ego must defend itself; a form of magical expiatory action, commonly observed in obsessive-compulsive disorder.

unio mystica Feeling of mystic unity with an infinite power.

unipolar psychosis Mood disorder characterized by recurrent episodes of depression or, much more rarely, recurrent manic states. *Compare* bipolar disorder.

universality Total effect produced when all group therapy members share specific symptoms or problems.

unstructured interview 1. Interview in which the interviewer chooses questions based on clinical judgment. **2.** Interview in which questions are related to what the patient brings up spontaneously.

urolagnia Paraphilia in which sexual pleasure is derived from urinating on someone else, being urinated on, or drinking urine.

utilitarian principle In policy and ethical decisions, a choice that is determined by whatever promises the greatest advantage to the common good.

utilization review committee Hospital committee that reviews the quality of services rendered, with particular attention to the appropriate use of facilities.

V

vagina dentata Sigmund Freud's theory that an unconscious fantasy exists in certain persons who are afraid of women and who believe that there are "biting teeth" in the vagina.

vaginismus Painful vaginal spasms, usually psychogenic, that occur during coitus, causing dyspareunia; may have its onset during preparation for intercourse, making penile insertion impossible. *See also See also* Genito–pelvic pain/penetration disorder.

validity Degree to which a test measures what it claims to measure; the degree to which an experimental design yields data truly applicable to the phenomenon under investigation. *See also* reliability.

variable In research, any characteristic or factor that may assume different values. The terms "independent variable" and "dependent variable" indicate a relationship of cause and of effect, respectively.

variable-interval schedule *See* schedule of reinforcement.

variable-ratio schedule *See* schedule of reinforcement.

variance Statistical measure of the variability within a set of observations, defined as the sum of the squared deviations from the mean divided by the number of observations; it equals the standard deviation squared. *See also* analysis of variance.

variation Statistical term referring to the manner in which persons in a population vary among themselves with respect to a given quality or trait; variability of persons also affects the precision of sample estimates. Three common measures of variation are the range, the variance, and the standard deviation. Knowledge of variation in a population is important because it indicates how useful the mean value is as a representative figure.

vascular dementia A type of dementia in DSM-5 in which a succession of strokes has led to a stepwise deterioration of intellectual functioning with associated focal neurological signs and symptoms. Previously called multi-infarct dementia.

vegetative 1. Pertaining to autonomic functions that are largely physiological or concerned with the growth, nutrition, or general physical health and homeostasis of the organism. **2.** In depression, denoting characteristic symptoms such as sleep disturbance (especially early morning awakening), decreased appetite, constipation, weight loss, and loss of sexual response.

vegetative nervous system Obsolete term for autonomic nervous system.

velocardiofacial syndrome A genetic disorder caused by a microdeletion in chromosome 22. Characterized by cleft palate, heart defects, immune dysfunction, mild intellectual disability,

psychotic disorders, and mood disorders. Also called Shprintzen syndrome or DiGeorge syndrome.

verbal amnesia *See* neurological amnesia.

verbigeration Meaningless and stereotyped repetition of words or phrases as seen in schizophrenia. Also called cataphasia. *See also* perseveration.

verbiomania LOGORRHEA.

vertigo Sensation that one or the world around one is spinning or revolving; a hallmark symptom of vestibular dysfunction, not to be confused with dizziness.

visceral learning *See* biofeedback; learned autonomic control.

visual agnosia Inability to recognize objects or persons.

visual amnesia *See* neurological amnesia.

visual hallucination Hallucination primarily involving the sense of sight.

visual motor gestalt test BENDER-GESTALT TEST.

vitamin therapy *See* orthomolecular psychiatry.

voice disorder A DSM-5 communication disorder that is characterized by abnormalities in vocal quality, pitch, and volume.

volitional Self-initiated.

volubility LOGORRHEA.

voluntary admission Hospitalization in which patients apply in writing for admission to a psychiatric hospital and are admitted if psychiatric examination reveals such need.

voluntary control Deliberate bringing about of an activity or refraining from an activity. The person exercising the control is aware of the nature of the activities and of the end to be achieved.

volunteer bias Sampling bias of persons who volunteer to take part in an experiment; such persons are not truly representative of the population and may tend to skew the results of the study, for example, the subjects may be more cooperative.

von Domarus principle Theory proposed by Eilhard von Domarus (1946) to explain schizophrenic thinking, which is characterized by the idea that two things are identical merely because they have identical predicates or properties.

voyeurism Paraphilia in which sexual excitement, frequently with orgasm, is obtained by looking at others naked, disrobing, or engaging in sexual activity. Also called voyeuristic disorder. *Compare* scopophilia.

W

WAIS Wechsler Adult Intelligence Scale.

Watson, John B. (1878–1958) American psychologist who used Ivan Pavlov's theory of classical conditioning to explain human behavior. *See also* Little Albert.

waxy flexibility Condition in which a person maintains the body position into which they are placed. Also called catalepsy.

Wechsler Adult Intelligence Scale (WAIS) Intelligence test that assesses intellectual functioning in adults.

Wechsler Intelligence Scale for Children (WISC) Intelligence test that assess intellectual functioning in children aged 5 to 15.

Wechsler Preschool and Primary Scale of Intelligence (WPPSI) Intelligence test that assesses intellectual functioning in children aged 4 to 6.

Wednesday Evening Society Small group of Sigmund Freud's followers who in 1902 started meeting with him informally on Wednesday evenings to receive instructions in psychoanalysis; as the society grew in numbers and importance, it evolved in 1910 into the Vienna Psychoanalytic Society.

weekend hospital Partial hospitalization in which the patient spends only weekends in the hospital and functions in the outside world during the week. *Compare* day hospital; night hospital.

Wernicke–Korsakoff syndrome Thiamine deficiency seen in Korsakoff psychosis (alcohol amnestic syndrome). Also called alcohol dementia. *See also* Wernicke fluent encephalopathy; Korsakoff psychosis.

Wernicke aphasia FLUENT APHASIA.

Wernicke fluent encephalopathy Neurological disorder caused by nutritional deficiency of thiamine (vitamin B) as seen in chronic alcoholism, with hallmark clinical features of ophthalmoplegia, ataxia, and confusion; some patients show disorders of short-term memory and other cognitive functions characteristic of Korsakoff psychosis (alcohol amnestic disorder), which has come to be regarded as a more advanced and chronic stage of the same disease process. Wernicke encephalopathy responds dramatically to the administration of thiamine, which may also prevent progression into the largely irreversible Korsakoff psychosis.

Weyer, Johann (1515–1588) Dutch physician considered by some to be the first psychiatrist, whose interest in human behavior led to his writing *De Praestigiis Daemonum* (1563), a landmark in the history of psychiatry.

white-out syndrome Psychosis that occurs in Arctic explorers and mountaineers who are exposed to a lack of diverse stimuli in the snow-clad environment.

WHO World Health Organization.

Williams syndrome A genetic disorder associated with a microdeletion on chromosome 7 and is characterized by "elfin-like" facial features, mild or moderate intellectual disability, cardiovascular disease, hyperacusis, hypercalcemia, and neuromusculoskeletal and renal abnormalities.

Wilson disease An autosomal recessive genetic disorder that causes excessive amounts of copper in the liver or brain. Patient may show psychological symptoms such as depression, suicidal ideation, and aggression. Is fatal without treatment.

windigo Culture-specific syndrome seen in Canadian Indians, characterized by a delusion of transformation into or possession by a windigo (wihtigo), a feared supernatural cannibalistic monster.

WISC Wechsler Intelligence Scale for Children.

withdrawal 1. Pathological retreat from interpersonal contact and social involvement; extreme decrease of intellectual and emotional interest in the environment. It may be seen in schizophrenia and depression. **2.** *See* substance withdrawal.

withdrawal syndrome Stereotyped constellation of signs and symptoms that appears when a person stops taking a drug or chemical substance on which he or she has become physically dependent. It is more likely to appear immediately after abrupt discontinuation, but may also

appear a few days later; specific clinical features are characteristic for the particular drug or substance. Also called abstinence syndrome; discontinuation syndrome.

word approximation Use of conventional words in an unconventional or inappropriate way (metonymy) or of new words that are developed by conventional rules of word formation (e.g., "handshoes" for gloves and "time measure" for clock); distinguished from a neologism, which is a new word whose derivation cannot be understood. *See also* paraphasia.

word association technique Technique devised by Carl Jung, who presented stimulus words to patients and had them respond with the first word that came to mind; it gives insight into persons' unconscious or subconscious thinking.

word blindness Obsolete word for alexia.

word salad Incoherent, essentially incomprehensible mixture of words and phrases commonly seen in far-advanced cases of schizophrenia. *See also* incoherence.

working through In psychoanalysis, the repeated and varied examination of a conflict, neurotic trend, or other problem for the purpose of increasing insight into unconscious processes and effecting adaptive personality changes; constitutes a fundamental facet of the psychoanalytic process.

World Health Organization (WHO) Specialized agency of the United Nations, established in 1948, governed by the World Health Assembly. Technical committees and regional organizations further the purpose of the WHO: "the attainment by all people of the highest level of health." It publishes the *International Classification of Diseases* (ICD).

World Psychiatric Association (WPA) International association, encompassing psychiatric associations from 75 countries throughout the world, whose objectives are to strengthen and promote international cooperation in psychiatry and among psychiatrists working in different fields and to encourage the exchange of information concerning psychiatric disorders.

WPA World Psychiatric Association.

WPPSI Wechsler Preschool and Primary Scale of Intelligence.

X-Y-Z

xenophobia Abnormal fear of strangers.

yohimbine Drug that blocks adrenergic α-receptors; because of its parasympathetic effect that causes vasodilatation, it has been used as an agent to help induce penile erection.

York retreat One of the earliest hospitals for the mentally ill, founded in 1792 in York, England, by William Tuke and other Quakers.

yttrium-90 Radioisotope that has been implanted experimentally in the brain to destroy certain tracts and areas in an effort to improve severe mental disorder.

zeitgeist German word for the general intellectual, moral, and cultural climate of an era.

zoophilia Paraphilia in which animals are used to produce sexual excitement. *See also* bestiality.

zoophobia Abnormal fear of animals.

z-score STANDARDIZED SCORE.

Zung Self-Rating Depression Scale (SDS) Twenty-item report in which the patient answers questions about symptoms of depressive disorders such as sleep patterns, appetite disturbance, libido, irritability, and hopelessness.

zygote Fertilized egg produced by the union of two cells (sperm and ovum) to form a diploid cell in sexual reproduction.

zygotic Pertaining to a zygote or to the individual developing from such a fertilized egg.

Color Atlas of Psychiatry

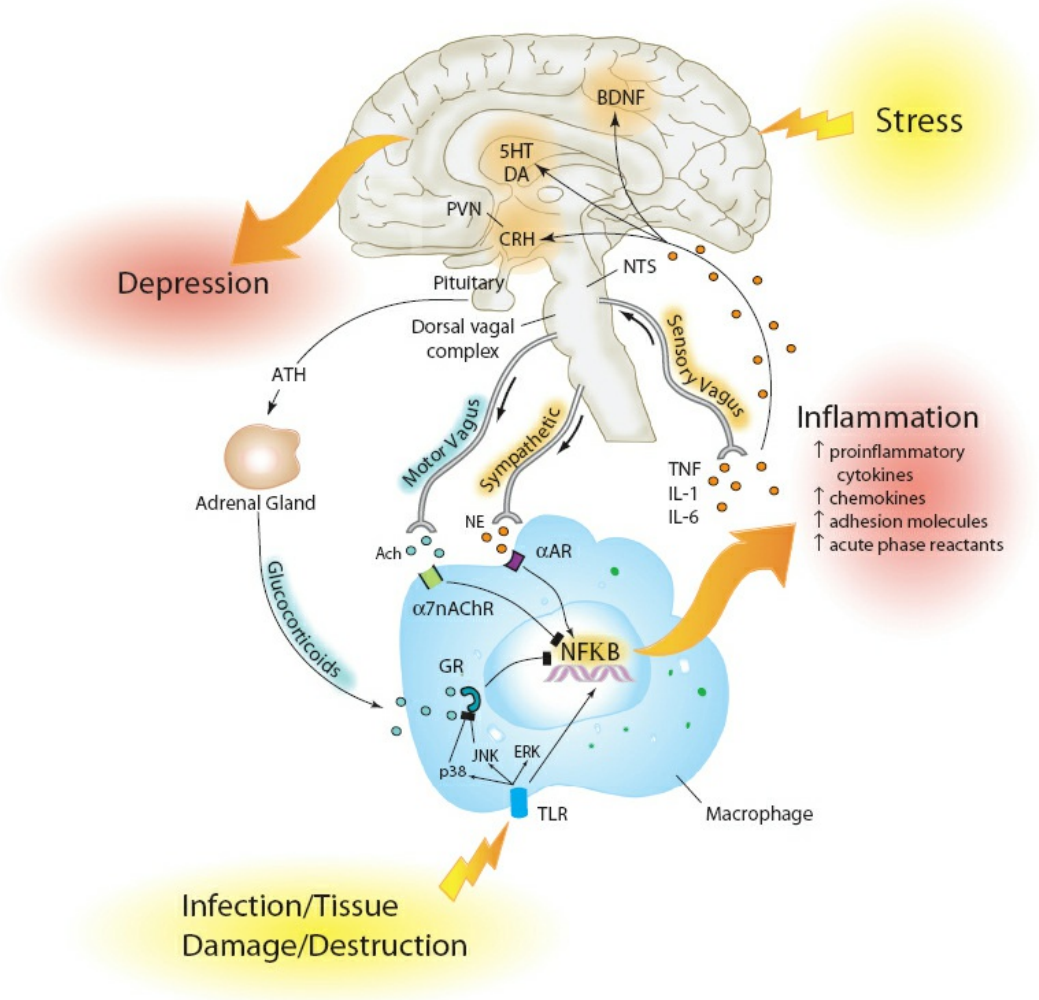

COLOR ATLAS 1.13–5. Stress and bidirectional neural–immune interactions. (**A**) Activation of NF-κB through toll-like receptors (TLRs) during immune challenge leads to an inflammatory response including (**B**) the release of proinflammatory cytokines, TNF-α, IL-1, and IL-6. **C:** These cytokines, in turn, access the brain via leaky regions in the blood–brain barrier, active transport molecules, and afferent nerve fibers (e.g., sensory vagus), which relay information through the nucleus tractus solitarius (NTS). **D:** Once in the brain, cytokine signals participate in pathways known to be involved in the development of depression, including: (i) altered metabolism of relevant neurotransmitters such as serotonin (5HT) and dopamine (DA), (ii) activation ofCRH in the paraventricular nucleus (PVN) and the subsequent production and/or release of ACTH and glucocorticoids (cortisol), and (iii) disruption of synaptic plasticity through alterations in relevant growth factors (e.g., brain-derived neurotrophic factor [BDNF]). **(E)** Exposure to environmental stressors promotes activation of inflammatory signaling (NF-κB) through increased outflow of proinflammatory sympathetic nervous system responses (release of NE, which binds to the α [αAR] and β [βAR] adrenoceptors). **(F)** Stressors also induce withdrawal of inhibitory motor vagal input (release of acetylcholine [Ach], which binds to the α7 subunit of the nicotinic acetylcholine receptor [α7nACRh]). **(G)** Activation of the mitogen-activated protein kinase pathways, including p38 and Jun N-terminal kinase (JNK), inhibit the function of glucocorticoid receptors (GRs), thereby releasing NF-κB from negative regulation by glucocorticoids released as a result of the hypothalamic–pituitary–adrenal (HPA) axis in

response to stress. (From Raison CL, Capuron L, Miller AH. Trends in Immunology, 27, Cytokines sing the blues: inflammation and pathogenesis of depression, 24–31, 2006, with permission from Elsevier.)

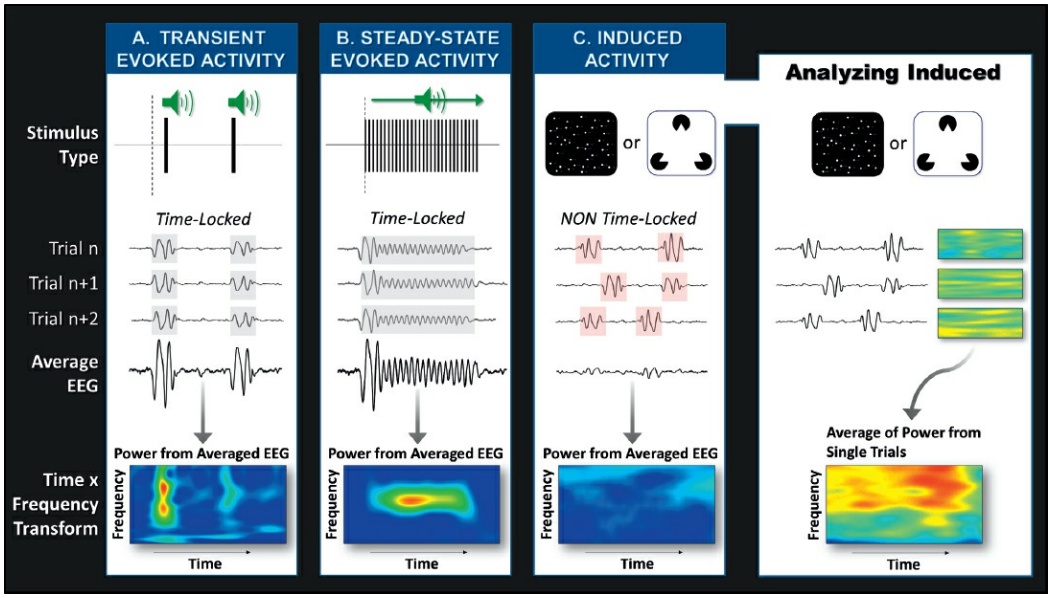

COLOR ATLAS 1.15–19. Transient and steady-stated evoked activity are phase- and time-locked oscillations (**A** and **B**) elicited by discrete or steady-state sensory stimuli. For both of these types of activity, averaging across trials has the effect of cancelling out non–phase-locked oscillations (due to destructive interference), leaving the evoked response intact and magnified. Hence, time–frequency analysis can be performed on the averaged EEG (i.e., after averaging all the individual trials). By contrast, induced activity is typically elicited during higher perceptual processing, for example, during the perception of coherent motion or of illusory shapes (see **Panel C**). These responses are often non time-locked (jittered across time from trial to trial). Hence, averaging the trials cancels out the response of interest due to destructive interference, and is not visible in the time–frequency transform (bottom of **Panel C**, left). However, these responses can be detected if time–frequency analysis is performed on the single trials, and the time–frequency transforms of individual trials are then averaged (**Panel C**, right). (Adapted from Skosnik PD, Cortes-Briones JA. Electroencephalography (EEG) and cannabis: from event-related potentials (ERPs) to oscillations. In: Preedy VR, ed. *The Neuropathology of Drug Addictions and Substance Misuse*. Amsterdam, Netherlands: Elsevier Science; 2016.)

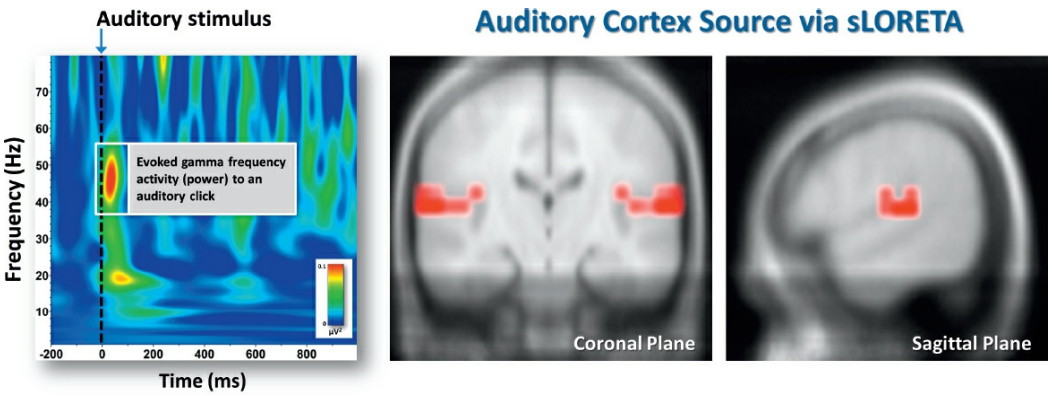

COLOR ATLAS 1.15–23. Example of EEG source analysis using sLORETA. The time–frequency transform on the left indicates a burst of transient gamma activity to an auditory click stimulus, which was localized to primary auditory cortex (right).

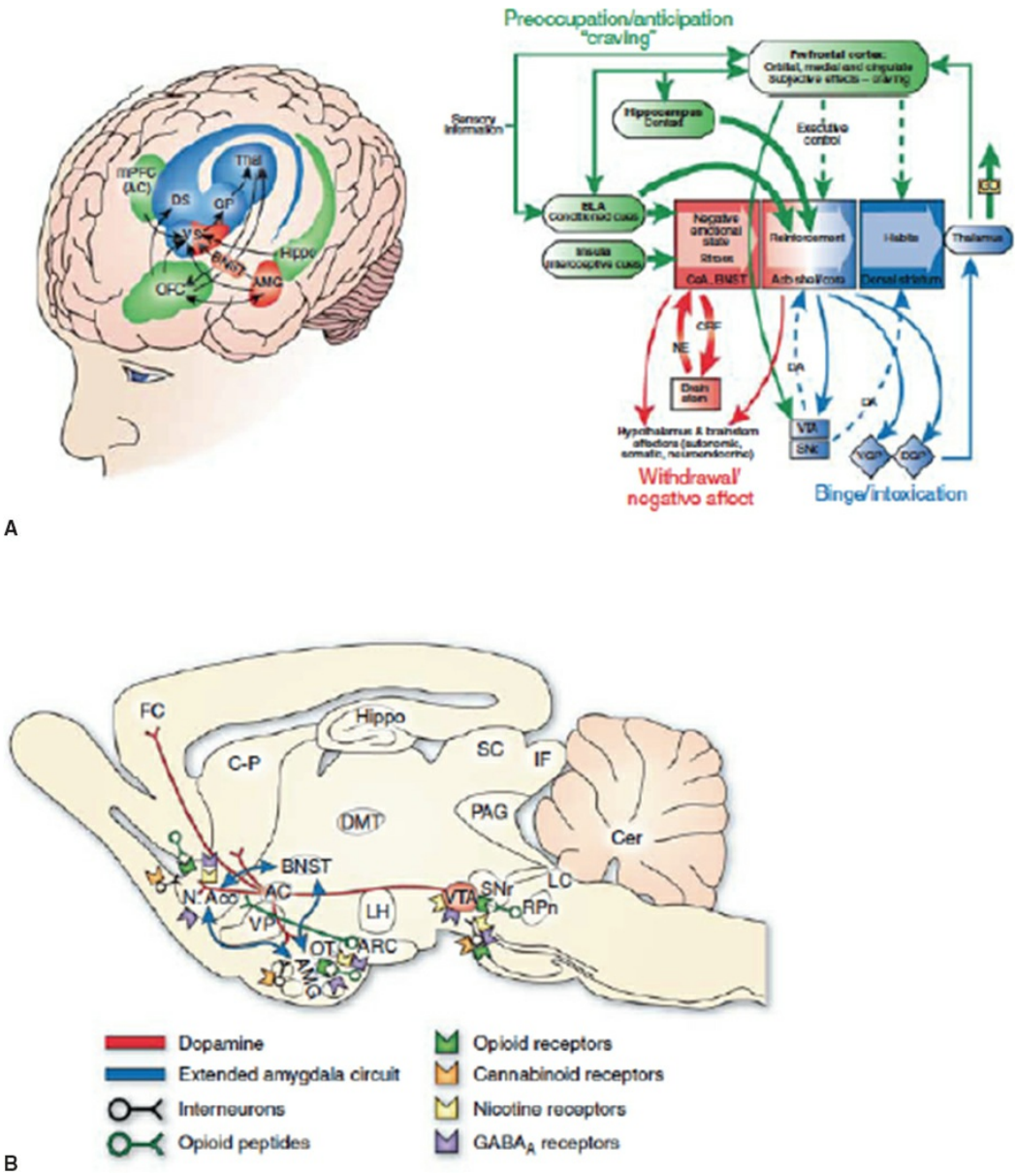

COLOR ATLAS 1.24–3. Diagram of neural circuits implicated in substance abuse and substance use disorders. **A:** Schematic illustrating the combination of neuroadaptations that are thought to occur within brain circuitry required for the three stages of the addiction cycle. The frontal cortex system is compromised, producing deficits in executive function and contributing to the incentive salience of drugs compared to natural reinforcers. Dopamine systems are compromised, and brain stress systems such as CRF are activated to reset further the salience of drugs and drug-related stimuli in the context of an aversive dysphoric state. **B:** Sagittal section through a representative rodent brain illustrating the pathways and receptor systems implicated in the acute reinforcing actions of drugs of abuse. The *blue arrows* represent interactions within the extended amygdala system hypothesized to have a key function in drug reinforcement. AC, anterior commissure; AMG, amygdala; ARC, arcuate nucleus; BNST, bed nucleus of the stria terminalis; Cer, cerebellum; CP, caudate–putamen; DMT, dorsomedial thalamus; FC, frontal cortex; Hippo, hippocampus; IF, inferior colliculus; LC, locus coeruleus; LH, lateral hypothalamus; N Acc., nucleus accumbens; OT, olfactory tract; PAG, periaqueductal gray; RPn, reticular pontine nucleus; SC, superior colliculus; SNr, substantia nigra pars reticulata; VP, ventral pallidum; VTA, ventral tegmental area. (From Koob GF, Volkow ND. Neurocircuitry of addiction. *Neuropsychopharmacology*. 2009;35:217.)

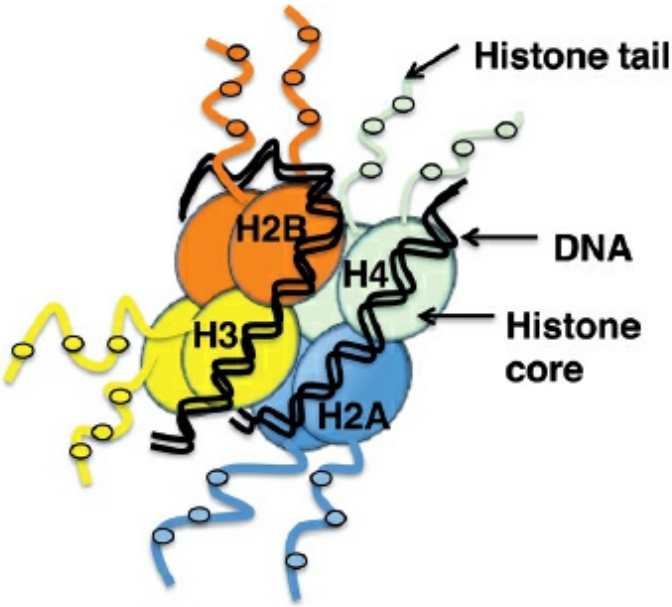

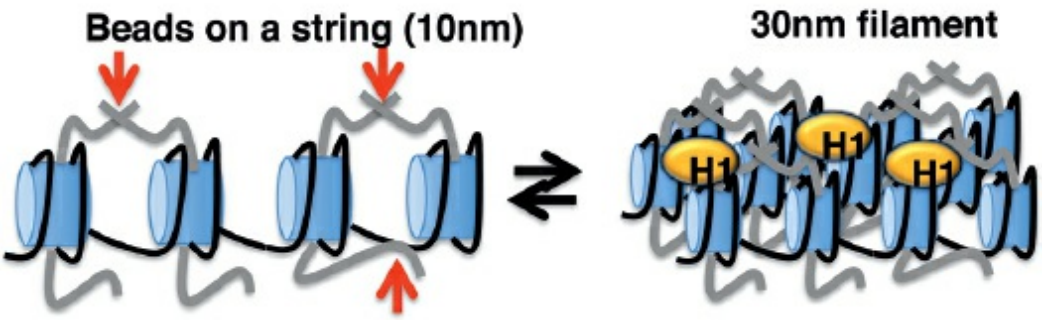

COLOR ATLAS 1.25–2. A: The nucleosome as a basic folding unit of chromatin. **B:** Chromatin can transition from a "beads on a string" 10-nm filament to a more compact 30-nm filament structure that is stabilized by the presence of histone H1 (*yellow circles*). Histone tails (*gray*) make contacts between neighboring nucleosomes (internucleosomal contacts) and with the underlying DNA (*red arrows*).

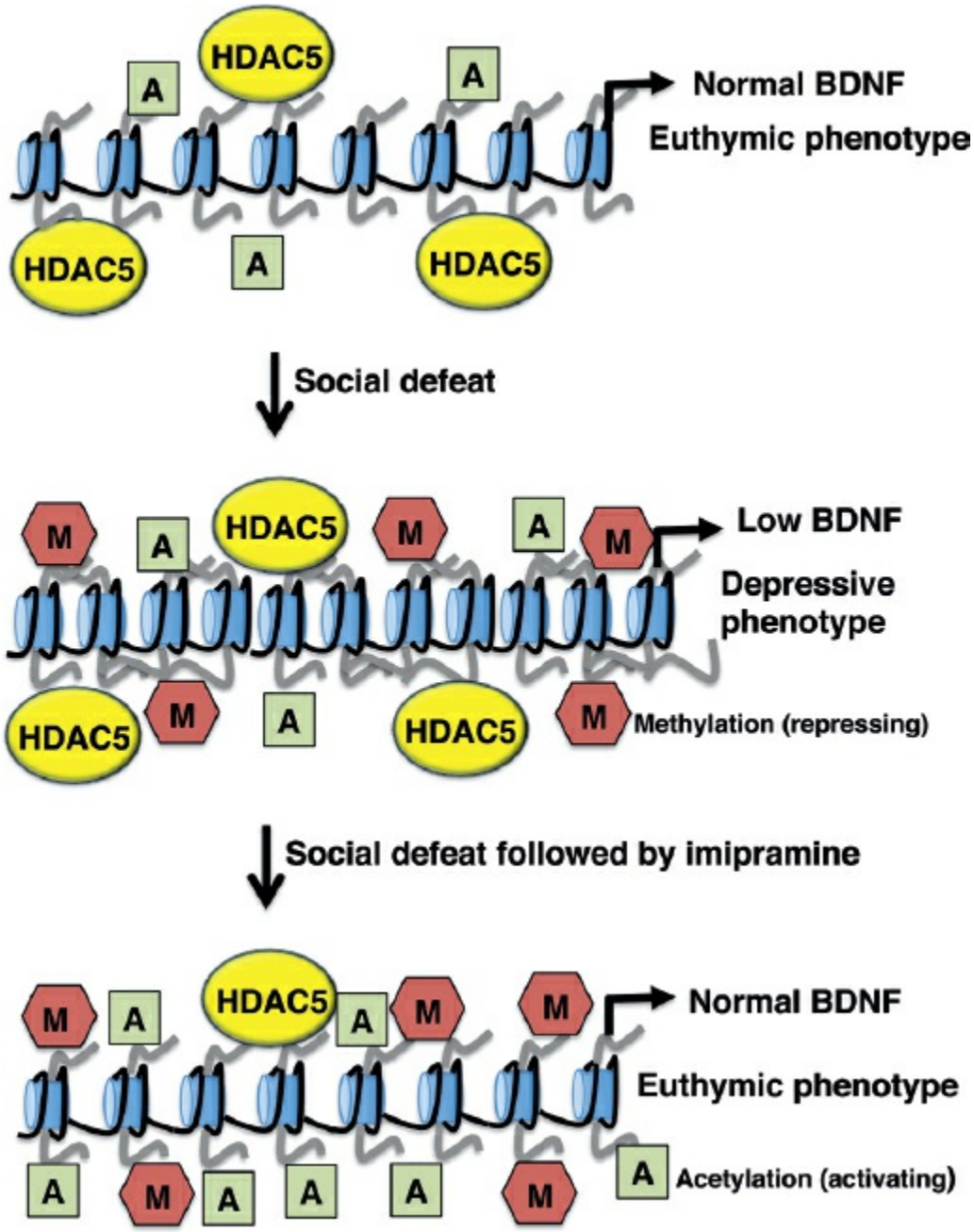

COLOR ATLAS 1.25–14. **Top:** Control mice show normal BDNF levels and euthymic-like phenotype. **Middle:** Exposure to repeated social defeat leads to an increase in repressive HK27-me2 methylation (*red hexagons*), leading to condensed chromatin, decreased BDNF levels, and a depressive-like phenotype. **Bottom:** Treatment of socially defeated mice with imipramine does not alter HK27-me2 levels, but instead decreases HDAC5 levels (*yellow circles*). Reduced HDAC5 levels contribute to the increase in histone acetylation (*green squares*), chromatin decondensation, increased BDNF levels, and restoration of euthymic phenotype.

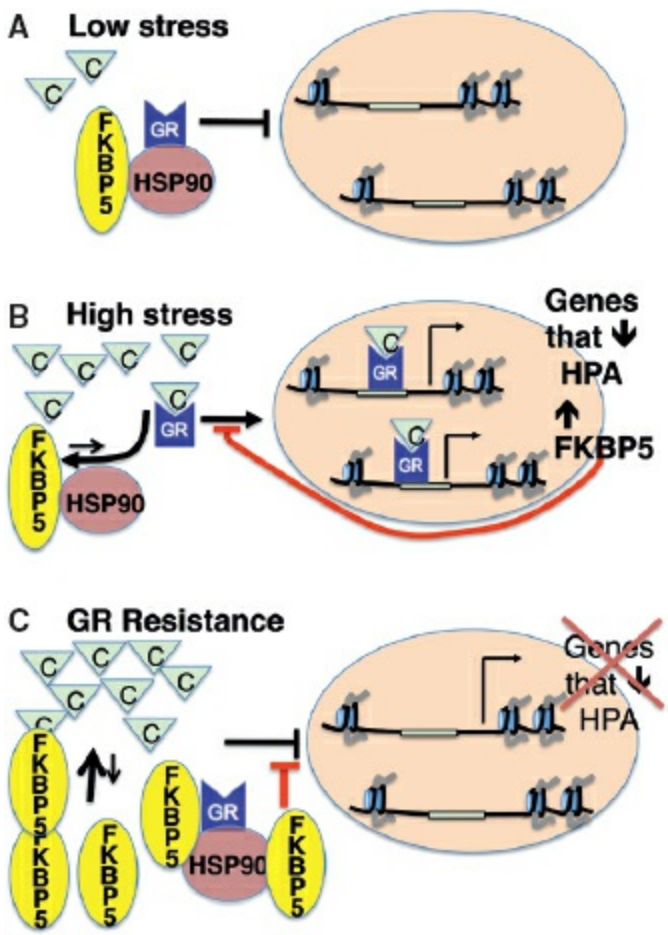

COLOR ATLAS 1.25–16. **A:** In the absence of stress, when cortisol levels (*green "C" triangles*) are low, the GR (*blue*) is bound to the chaperone proteins FKBP5 (*yellow circle*), and HSP90 (*brown circle*). In this bound state, the GR is prevented from entering the nucleus and activating transcription. **B:** Exposure to stress increases cortisol levels allowing GR to enter the nucleus by displacing the FKBP5/HSP90 complex. Inside the nucleus, GR activates transcription of genes that terminate further cortisol release and activates transcription of FKBP5. Increased FKBP5 levels prevent further GR entry into the nucleus, resetting the system so it can respond to subsequent waves of stress. **C:** Overexpression of FKBP5 leads to GR resistance by preventing high levels of cortisol from driving GR into the nucleus. As a result, additional levels of cortisol are being released leading to high levels of circulating cortisol.

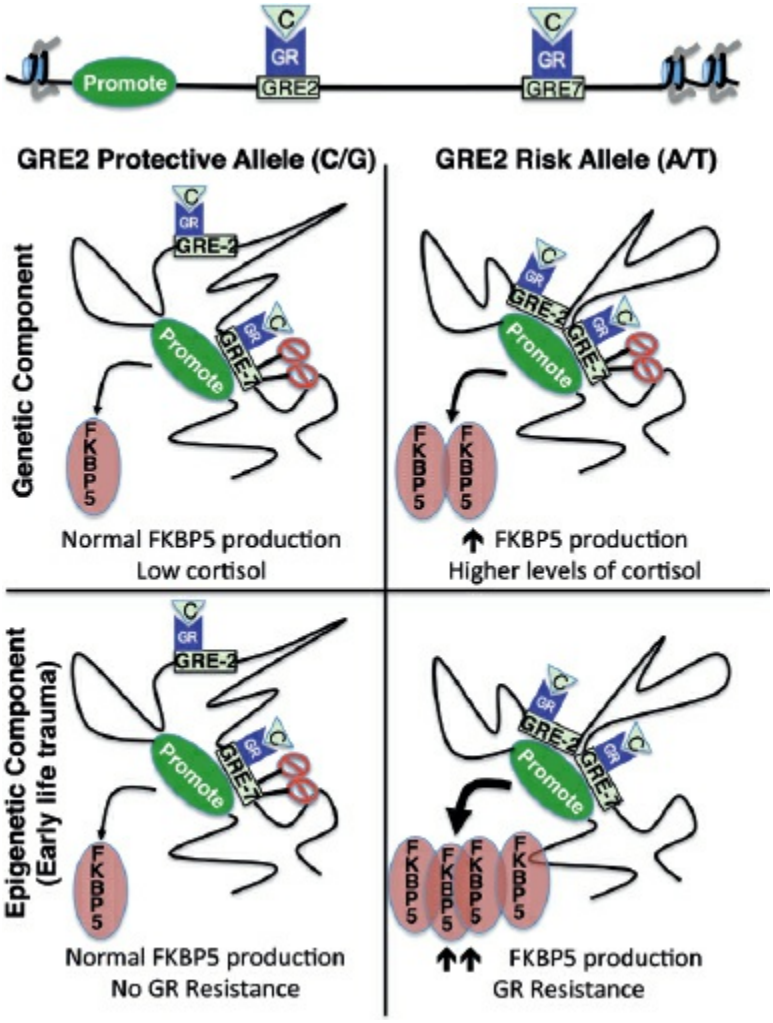

COLOR ATLAS 1.25–17. The FKBP5 gene has two GREs located in introns 2 (GRE-2) and 7 (GRE-7). The **top** two panels demonstrate the impact of the genetic risk on FKPB5 expression and GR resistance. On the **left,** GRE-2 containing the protective C/G allele is not recruited to the promoter/GRE-7 complex producing relatively low levels of FKBP5, and low cortisol levels in response to mild stress. The presence of the GRE-2 at the promoter/GRE-7 complex in the A/T allele (**top right**) produces higher levels of FKBP5, and hence higher levels of cortisol in response to mild stress. The bottom panels demonstrate the epigenetic contribution of early life stress to FKPB5 expression. Early life stress results in relatively low levels of cortisol and normal methylation at the GRE-7 in individuals with the C/G protective allele (**lower left panel**). In contrast, early life stress leads to higher cortisol levels and a decrease in the DNA methylation at the GRE-7 of individuals with the A/T risk allele. Demethylation at GRE-7 leads to more efficient GR binding to this site and therefore higher FKBP5 levels and GR resistance (**lower right panel**).

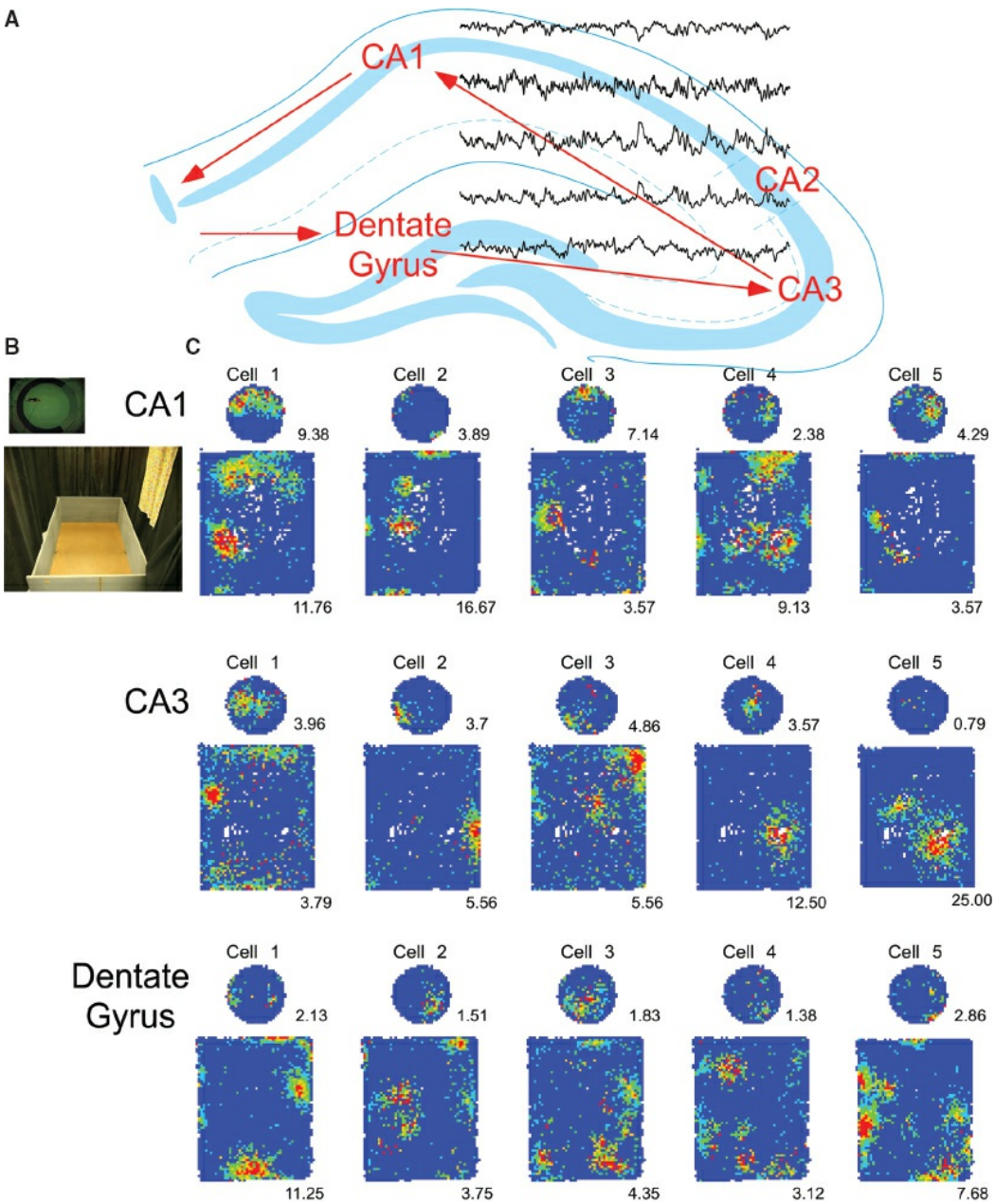

COLOR ATLAS 1.26–2. Hippocampus place cells generate an ensemble place code. **A:** Schematic of the dorsal hippocampus with the major subdivisions labeled and arrows indicating the feed-forward flow of information through the trisynaptic circuit. A 2-s example of the location-specific LFP has been superimposed such that the start of the LFP corresponds to the location where it was recorded. **B:** Photos of the two familiar environments in which place cells were recorded, a 76-cm diameter cylinder and a 1.8 × 1.4 m box. **C:** Blue-to-red color-coded average firing rate maps document the spatial tuning of place cells recorded as a rat foraged for scattered food in the cylinder and box. Each set of five cells was recorded simultaneously from the same anatomical subdivision of the hippocampus circuit. Notice that in the cylinder most place cells only discharge robustly in a single location, called the cell's place field, which suggests that the discharge provides a dedicated neural code for location. However, the fundamental nature of the place code is actually an ensemble code because in the larger environment of the box, the same cells typically discharge in multiple place fields. If each cell's discharge in the box is interpreted as a dedicated code, the activity of each cell provides an ambiguous indication of location that cannot be reliably disambiguated within the dedicated coding framework. If however, the discharge is interpreted as an ensemble code, the pattern of activity across the set of cells provides an accurate indication of location because the across-cell pattern of activity is

unique at each location. The minimum firing rate (measured as spikes/s) in the *red color* (highest firing rate) category is given below each firing rate map. The *blue color* indicates a firing rate of 0 spikes/s. Recordings by Dr. EunHye Park.

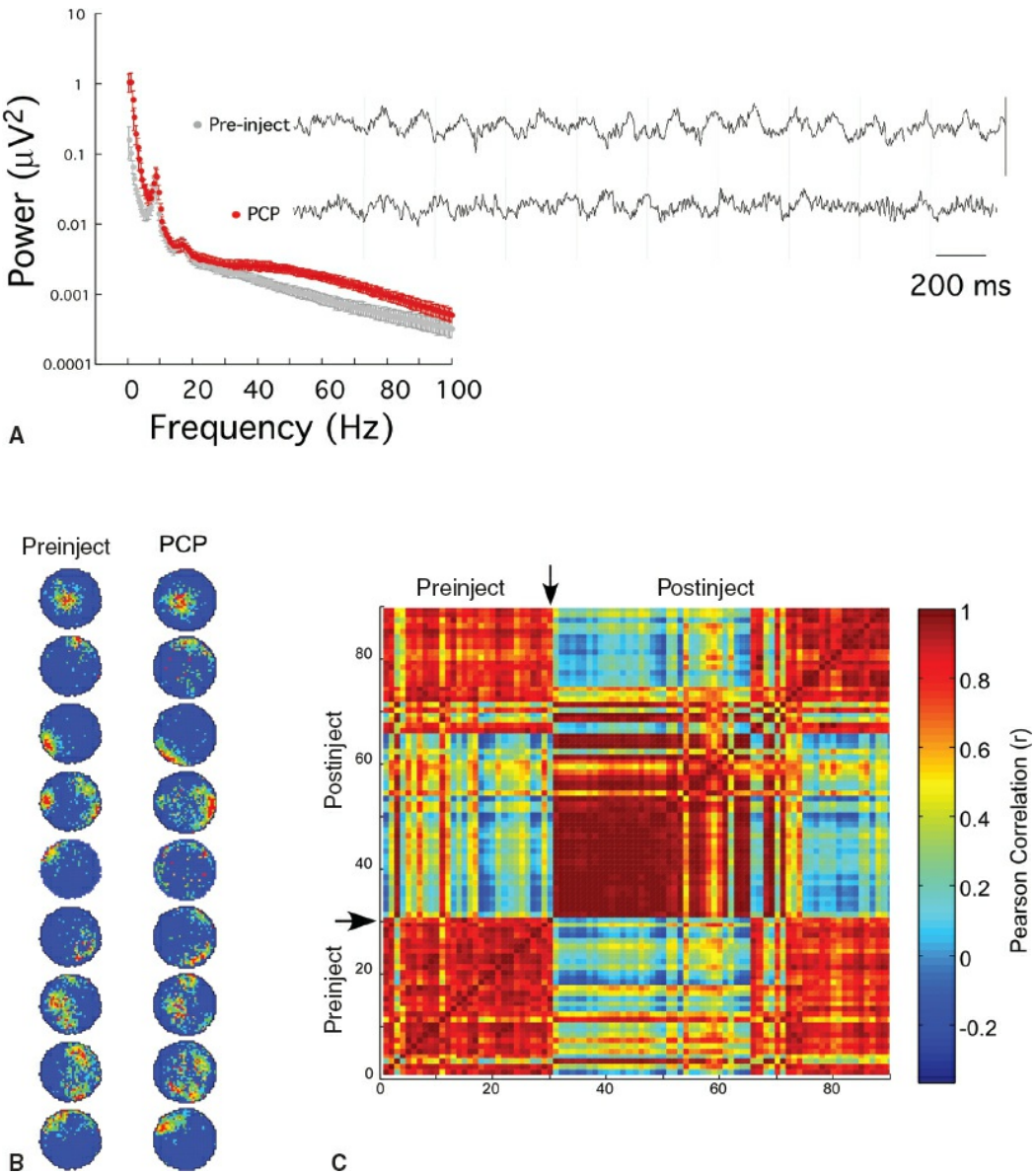

COLOR ATLAS 1.26–3. The psychotomimetic phencyclidine (PCP) discoordinates neural discharge with minimal effects on the spatial tuning of individual place cells. **A:** The blue-to-red color-coded firing rate maps of an ensemble of nine CA1 place cells that were recorded as a rat foraged for scattered food in a familiar cylinder before and after a systemic injection of 5 mg/kg of PCP. The PCP had no discernable effect on the spatial firing of the individual place cells. **B:** In contrast, when the same data are investigated as a time-varying across-cell pattern of discharge, it is obvious that the PCP radically changed the ensemble firing of the cells (PCP was administered at 30 minutes). To generate the correlation matrix that is shown, every minute the number of spikes that each cell discharged was counted and organized as an activity vector that described the temporal pattern of discharge across the population. Each activity vector was correlated with every other activity vector and the correlation is represented using the color code. The correlation at any point in the matrix represents the activity at time x correlated with the activity at time y. The correlations along the x = y line are exactly 1, because it is the activity vector correlated with itself. The activity vectors under PCP immediately changed relative to the preinjection vectors (independent or negative correlations) but the PCP activity formed a stable pattern

indicated by the correlations of approximately 1.30 to 40 minutes after PCP the place cell ensemble activity switched back and forth between the preinjection and the PCP patterns until it settled back to the preinjection pattern. Notice that at three and eleven minutes during the preinjection period, the PCP pattern was expressed (in the absence of PCP). Not only do these data demonstrate discoordination of neural activity without affecting the information tuning of individual neurons, but also they indicate that ensemble activity is multistable. The system prefers to exist in one effectively stable activity pattern and PCP radically changes the probability of observing the distinct patterns. Recordings by Dr. Hsin-Yi Kao.

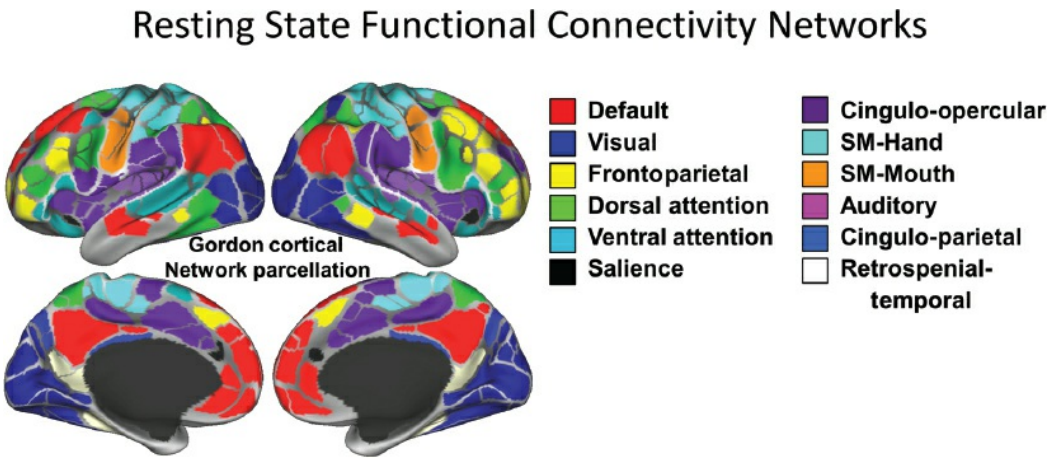

COLOR ATLAS 1.28–1. Functional connectivity networks derived from resting state data. The top row is a lateral view of the brain and the bottom row is a medial view of the brain. The colored "parcels" were derived from a boundary identification approach using a "watershed by flooding" algorithm to identify boundaries in the gradients of change in whole brain maps of functional connectivity. The community structure of the parcels derived from this boundary mapping was generated using an InfoMap algorithm. (From Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb Cortex*. 2016;26(1):288–303, with permission.)

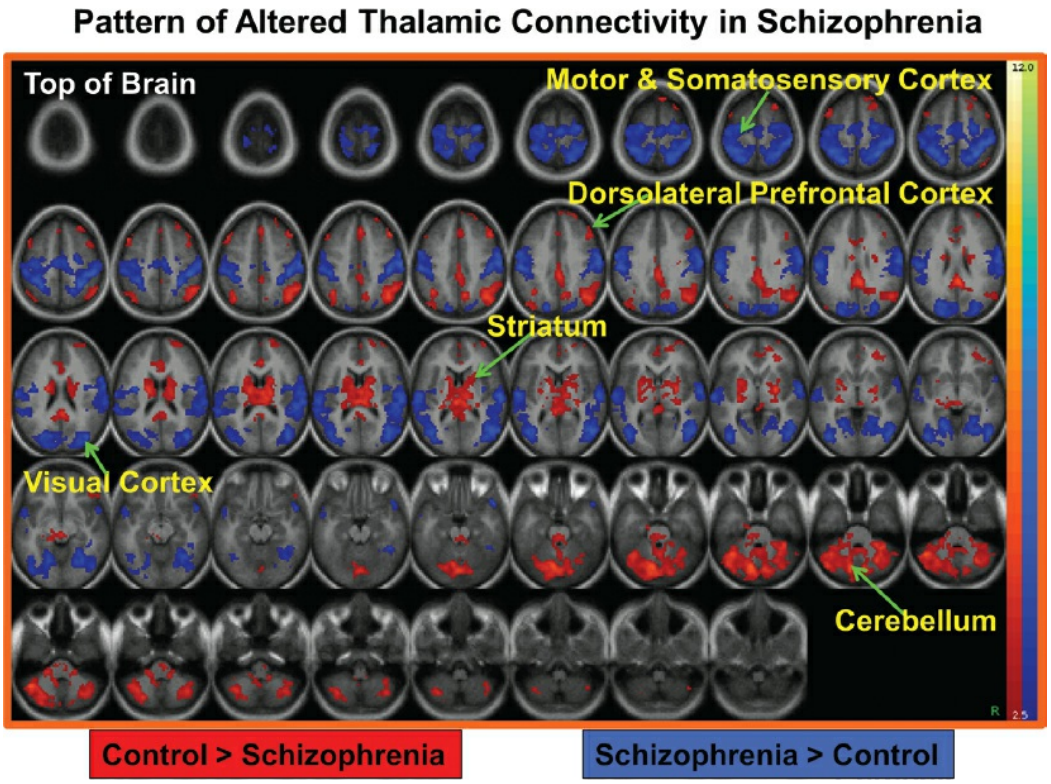

COLOR ATLAS 1.28–3. Patterns of altered thalamic connectivity in schizophrenia. Unpublished data showing regions (in *red*) where healthy individuals show greater thalamic connectivity than individuals with schizophrenia, and regions (in *blue*) where individuals with schizophrenia show greater thalamic connectivity than healthy individuals. The regions showing decreased thalamic connectivity in schizophrenia includes cerebellum, striatum, dorsolateral prefrontal cortex, and the dorsal anterior cingulate. The regions showing increased thalamic connectivity in individuals with schizophrenia include motor and somatosensory cortices, visual cortex, and anterior temporal poles.

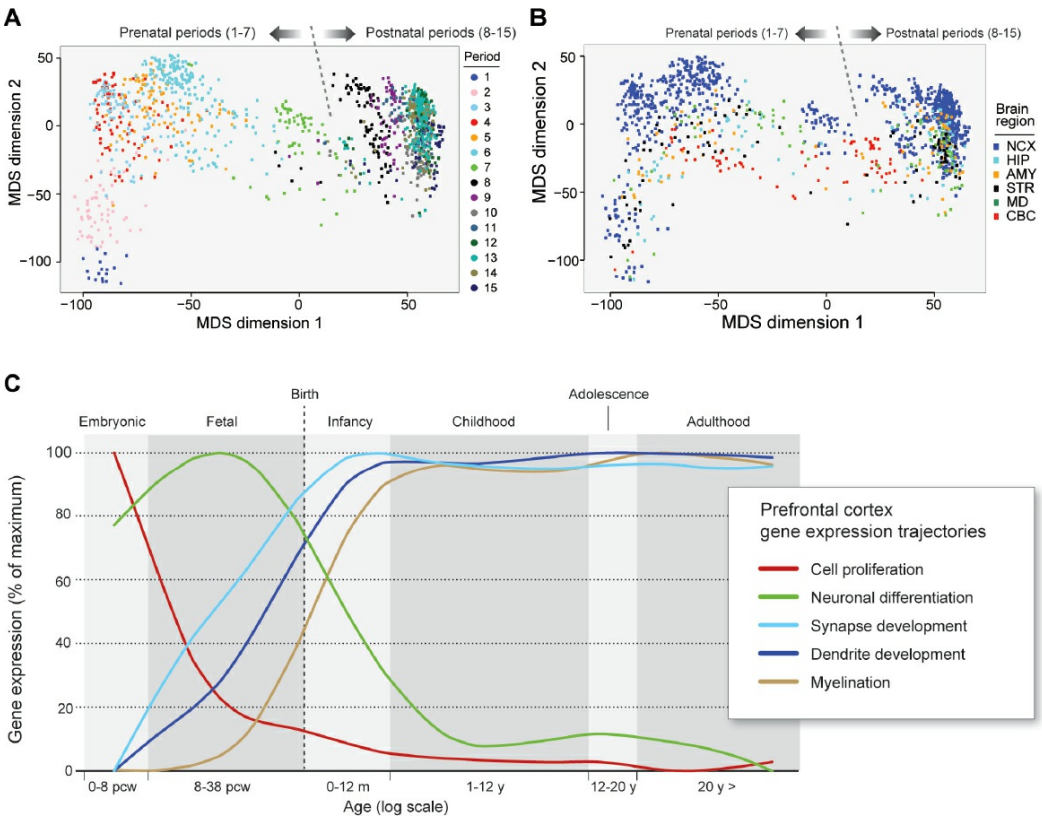

COLOR ATLAS 1.30–2. Global spatiotemporal dynamics of the human brain transcriptome. **A:** This panel shows a multidimensional scaling (MDS) plot that describes gene expression differences between developmental periods. MDS is statistical/graphical technique that reduces the complexity of data by plotting individual data points apart from one another at distances corresponding to the degree to which they are different. In the plot, each dot represents a sample and is colored according to period as defined by Kang et al., 2011; the dotted line indicates birth. Note that the most pronounced differences (approximately two-thirds) occur during prenatal development (periods 1 to 7). By contrast, over four decades of adulthood (periods 13–15), less than 1 percent of genes are differentially expressed. **B:** Here the same MDS plot as in (**A**) is shown except that the new data points are colored by brain region. Note that the most prominent transcriptional differences between regions are between cerebellar cortex (CBC) and forebrain regions. NCX, neocortex; HIP, hippocampus; AMY, amygdala; STR, striatum; MD, mediodorsal nucleus of thalamus. **C:** A line graph illustrating the developmental trajectory of expression of genes important for major neurodevelopmental processes demonstrating how gene expression profiles are linked to the cellular processes that take place at different periods of development. The y-axis represents the percentage of gene expression relative to maximum and the x-axis corresponds to age.

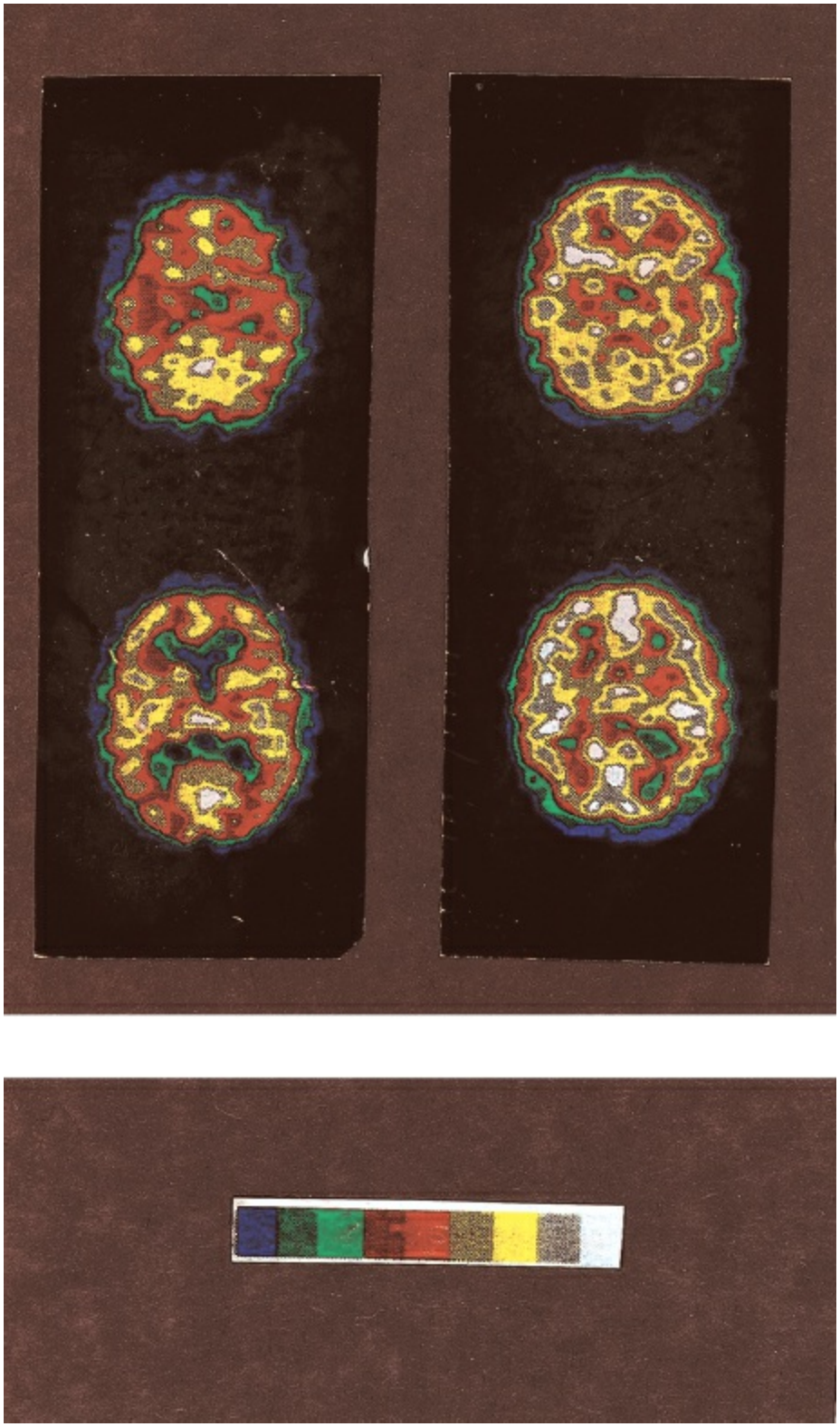

COLOR ATLAS 2.9–1. Single-photon emission computed tomography (SPECT) scan demonstrating multiple areas of decreased blood flow in a Lyme patient (L) compared to a healthy control (R).

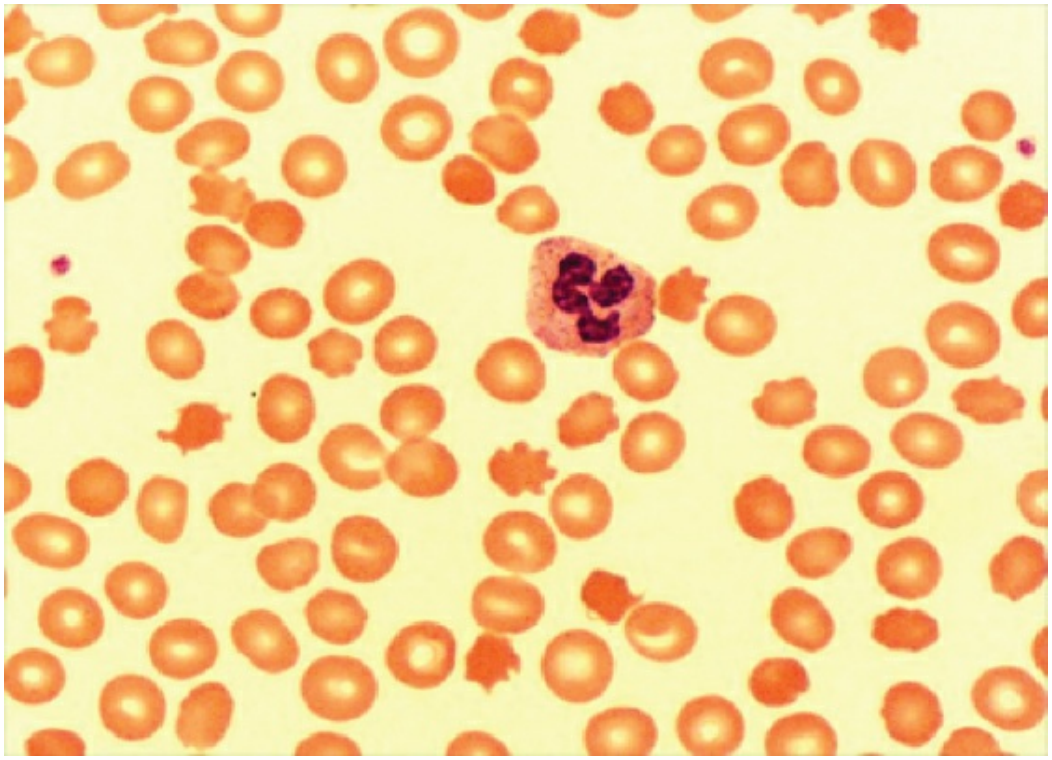

COLOR ATLAS 2.14–11. A 38-year-old woman with neuroacanthocytosis who first presented with adolescent obsessive–compulsive symptoms prior to adult-onset seizures and chorea. Blood smear showing characteristic red cell acanthocytes.

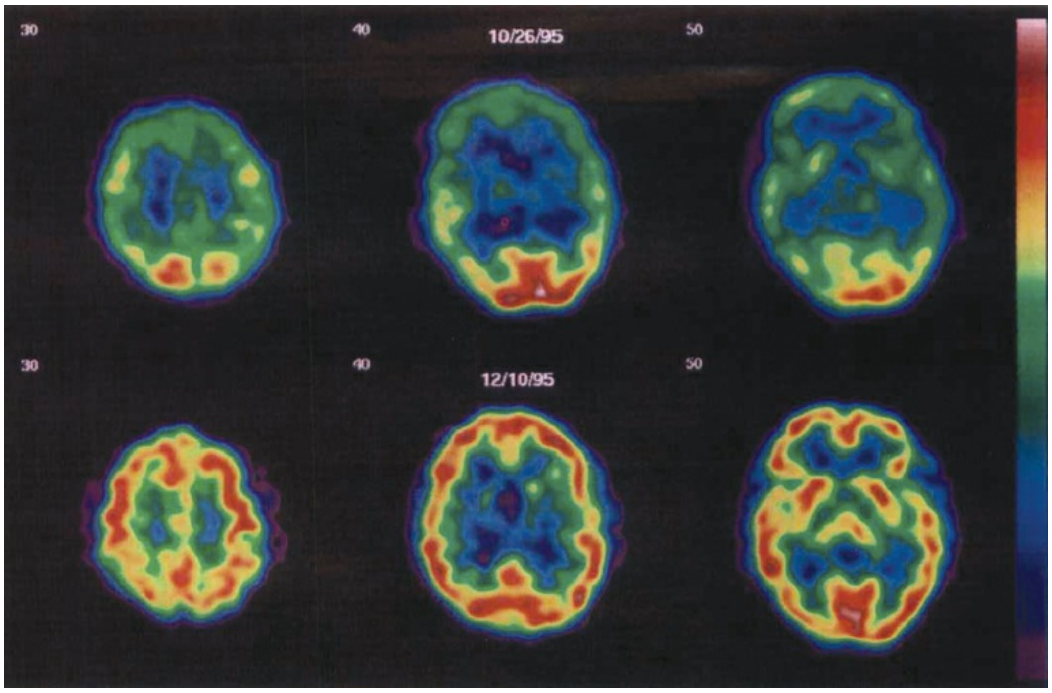

COLOR ATLAS 2.14–21. Reversibility in Hashimoto encephalopathy. Single photon emission computed tomography (SPECT) showing gross global hypoperfusion in all nonoccipital regions in a 59-year-old woman with rapidly progressive cognitive impairment and myoclonus, with Mini-Mental State Examination (MMSE) score of 20 (**top**) and after significant clinical improvement 6 weeks later, when MMSE score was 27 (**bottom**). (From Forchetti CM, Katsamakis G, Garron DC. Autoimmune thyroiditis and a rapidly progressive dementia: Global hypoperfusion on SPECT scanning suggests a possible mechanism. *Neurology.* 1997;49:623, with permission.)

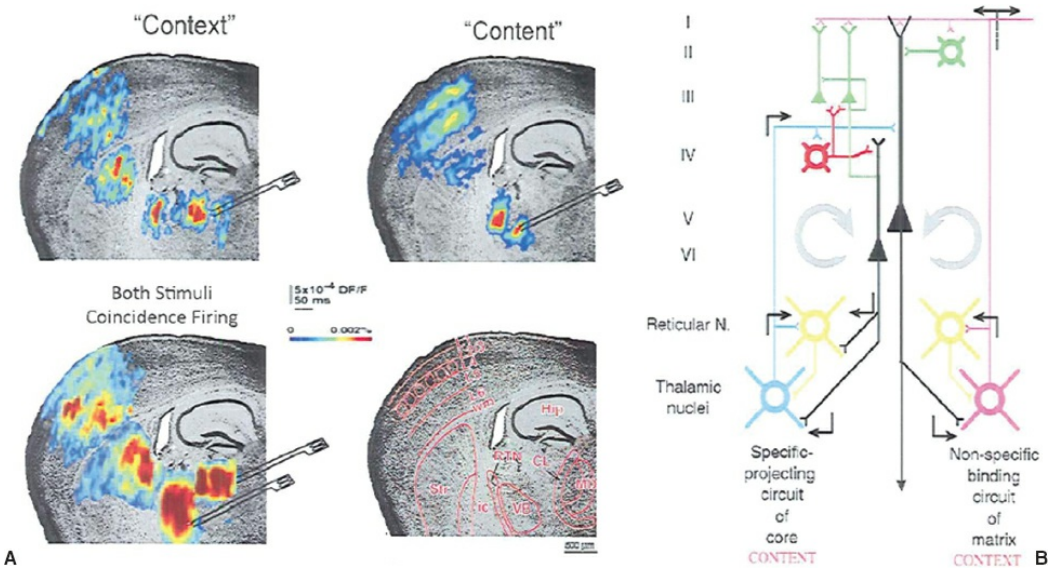

COLOR ATLAS 3.5–2. A: Voltage imaging of thalamocortical activation of a somatosensory brain slice demonstrating temporal binding of specific and non-specific thalamocortical inputs. Upper left quadrant, response average of a set of single pulse stimulation of centrolateral (CL) thalamic nucleus (Context). Upper right quadrant response average to a set of single stimuli to the ventrobasal (VB) thalamic nucleus (Content). Both CL and VB stimulation activated the thalamic reticular nucleus and the striatum. A different cortical activation patterns was, however, observed for each stimulation site. Thus, Cl stimulation (upper left quadrant) activated layers 6, 5, and 1, while VB activated layer 4, 2/3, and 5, 6. When both thalamic nuclei were activated simultaneously (lower left quadrant), a clear increase in activity was observed at all sites, indicating functional resonance as expected during gamma band thalamocortical binding. **B:** Diagram of two thalamocortical systems. **Content:** Specific sensory or motor nuclei project to cortex layer 4, generating cortical oscillation by direct activation of layer 4 neurons that contact dendrites of layer 5 and 6. Layer 6 pyramids generate return inhibition to specific thalamus via the reticular thalamic nucleus. **Context:** The second loop relates to the nonspecific intralaminar nuclei projecting to the most superficial layer of the cortex and giving collaterals to the reticular nucleus. Pyramidal cells in layer 5 return the oscillation to the reticular and the nonspecific thalamic nuclei, establishing a second resonant loop. The conjunction of the specific and nonspecific loops (content and context) is proposed to generate temporal cognitive binding. (Modified from Llinás R, Leznik E, Urbano IF. Temporal binding via coincidence detection of specific and non-specific thalamocortical inputs: a voltage dependent dye imaging study in mouse brain slices. *Proc Natl Acad Sci USA.* 2002;99:449.)

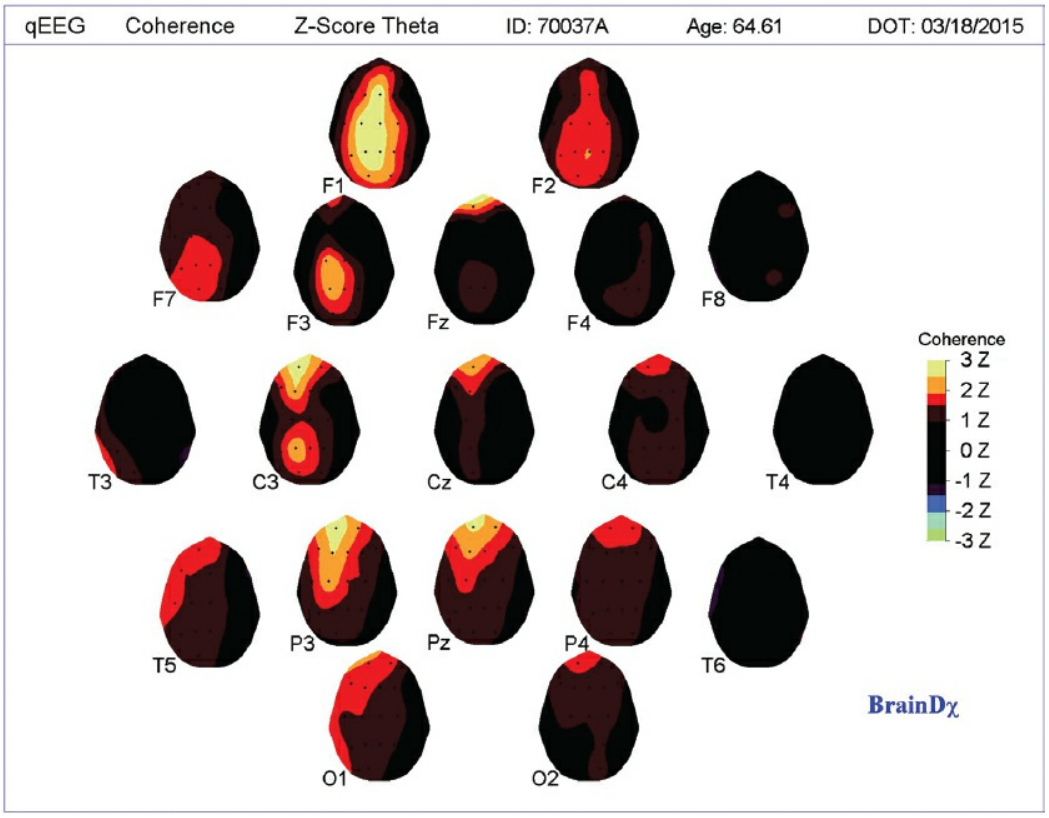

COLOR ATLAS 7.7–5. Each component image (individual head map with nose up and left on the left) shows the results for the coherence between the labeled region and all other regions, in the labeled band (in this case the theta frequency band), color coded by significance level using the z-scale shown. Regions in black are within normal limits, all colored regions are significant at $p < 0.05$ with the extremes at $p < 0.001$. The top left image shows the connectivity between the left frontopolar regions (FP1) and all other regions in the theta band, a band which has been shown in the literature to be correlated with degree of cognitive decline in the elderly. It can be seen that there is a significant hypercoherence between this regions and both short and long distance relationships on the left hemisphere and midline regions. This hypercoherence is also reflected in the coherence between FP1 and the left central (C3), left parietal (P3) and midline parietal (Pz) regions. This hypercoherence between frontal and more distant regions in a patient with early cognitive decline might serve as a baseline for a trajectory of increasing abnormality when tracked longitudinally in this patient.

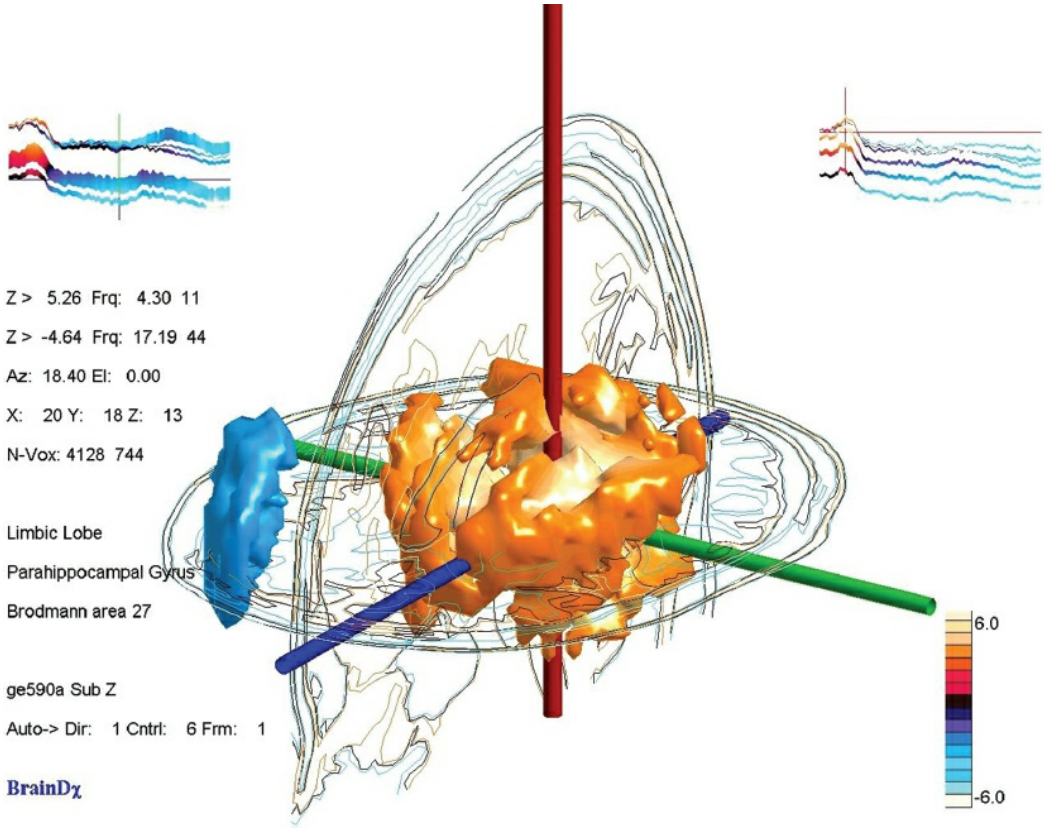

COLOR ATLAS 7.7–7. Three-dimensional low-resolution electromagnetic tomographic analysis (LORETA) image of a "dissolved brain" of a patient with AD, after removal of all voxels with a *Z* score less than 3.29 ($p < 0.001$). The maximum theta excess (*dark red shading,* 4.3 Hz) relative to age-expected normal values is in the hippocampal and parahippocampal regions, while the frontal cortex (Brodmann area 27) reveals a minimum beta deficit (*blue,* 17 Hz).

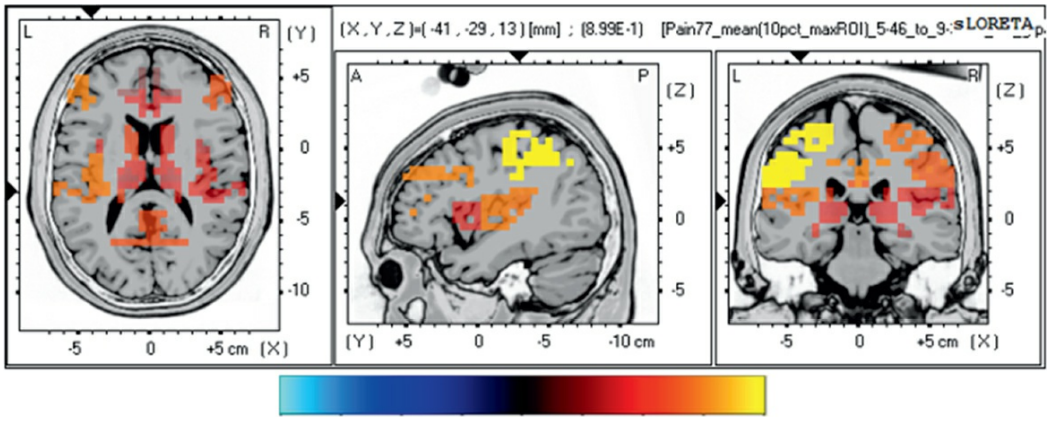

COLOR ATLAS 7.7–11. Source localization (sLORETA) images (axial, saggital and coronal views repsectively) for group average voxel *Z* scores for ROIs in the band 5.46 to 9.36 Hz for the chronic pain population ($n = 77$). The images are color coded for significance of difference according to the color scale (below figure) where extremes of the color scale are $p < +/-0.01$. Over-activation is shown in red/yellow and deactivation in blue/aqua (none seen in this image). [Left is on left in these images.]

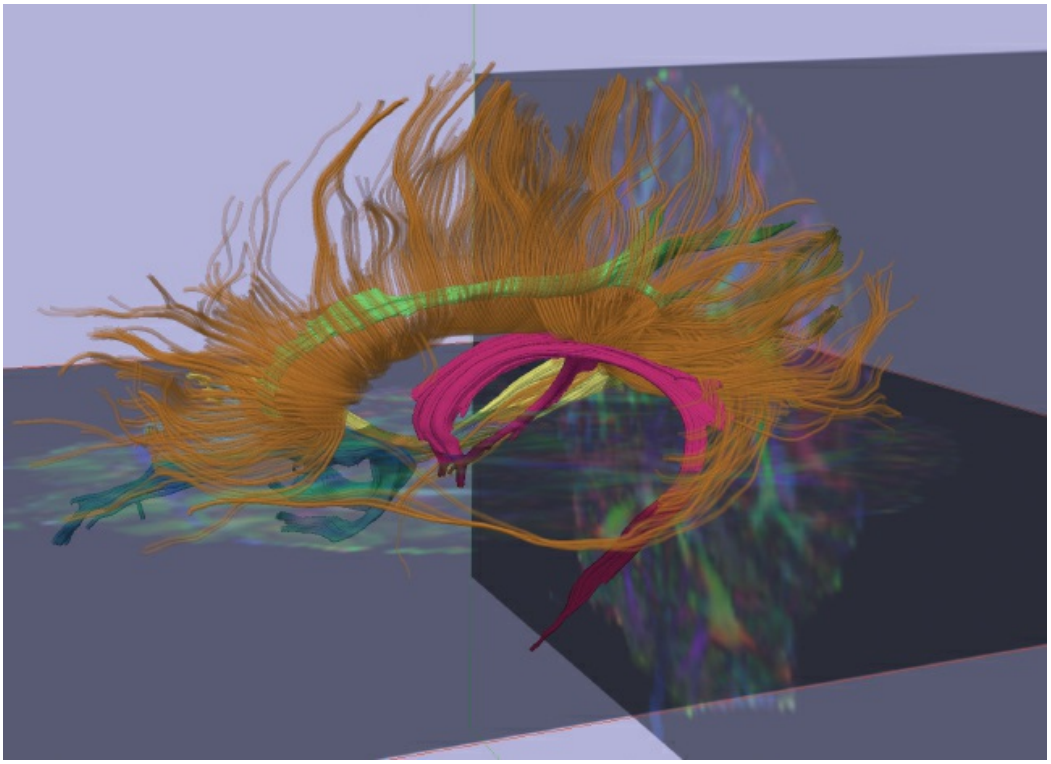

COLOR ATLAS 12.7–7. Three-dimensional image reconstructed based on diffusion data acquired on a 3-T General Electric scanner, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, which shows several major white matter fiber bundles identified through diffusion tensor imaging: Fornix (*magenta*), right cingulum (*green*), right inferior longitudinal fasciculus (*yellow*), right uncinate fasciculus (*blue*), and corpus callosum (*orange*).

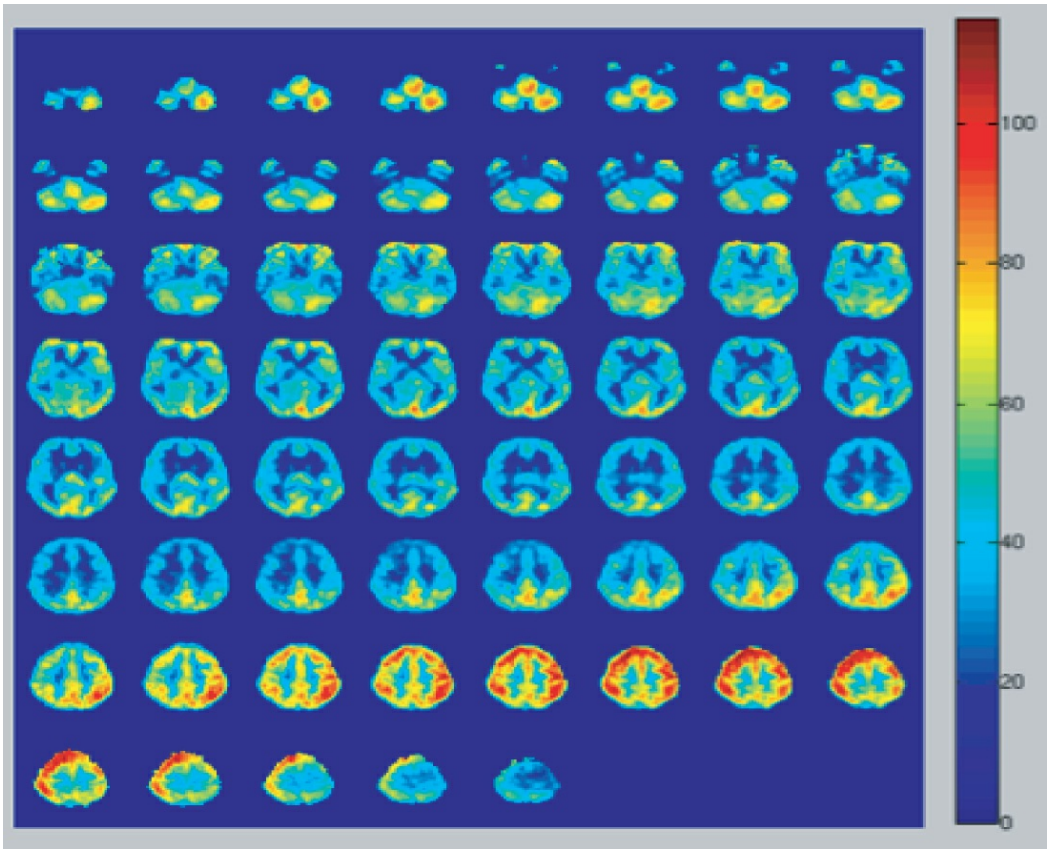

COLOR ATLAS 12.8–4. Map of quantitative blood flow obtained in a healthy individual with arterial spin labeling magnetic resonance imaging.

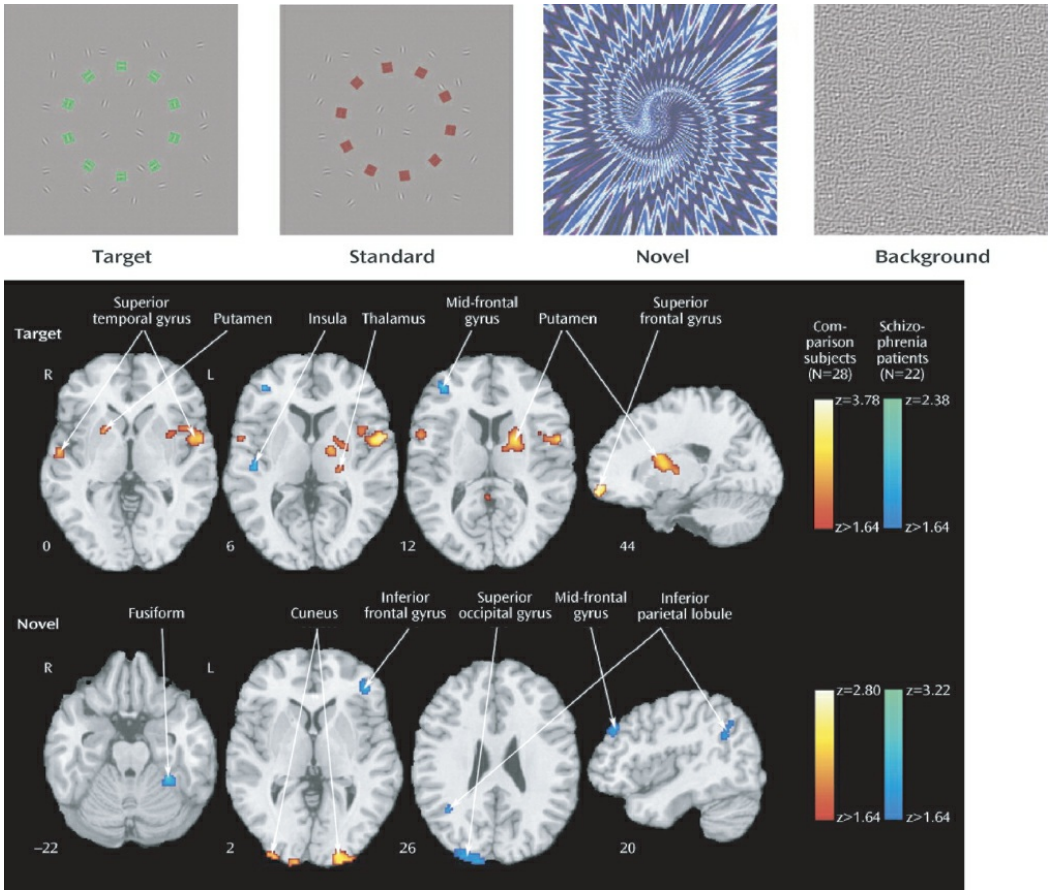

COLOR ATLAS 12.8–8. Examples of stimuli used in "oddball" studies of attentional processing (*top*) and contrast images of patients with schizophrenia and comparison subjects for target and novel stimuli. Greater activation in patients is depicted by the scale on the right, whereas greater activation in comparison subjects is shown by the scale on the left. Images are in radiological convention (left hemisphere to viewer's right). (From Gur RE, Turetsky BI, Loughead J, Snyder W, Kohler C. Visual attention circuitry in schizophrenia investigated with oddball event-related functional magnetic resonance imaging. *Am J Psychiatry.* 2007;164:442–449.)

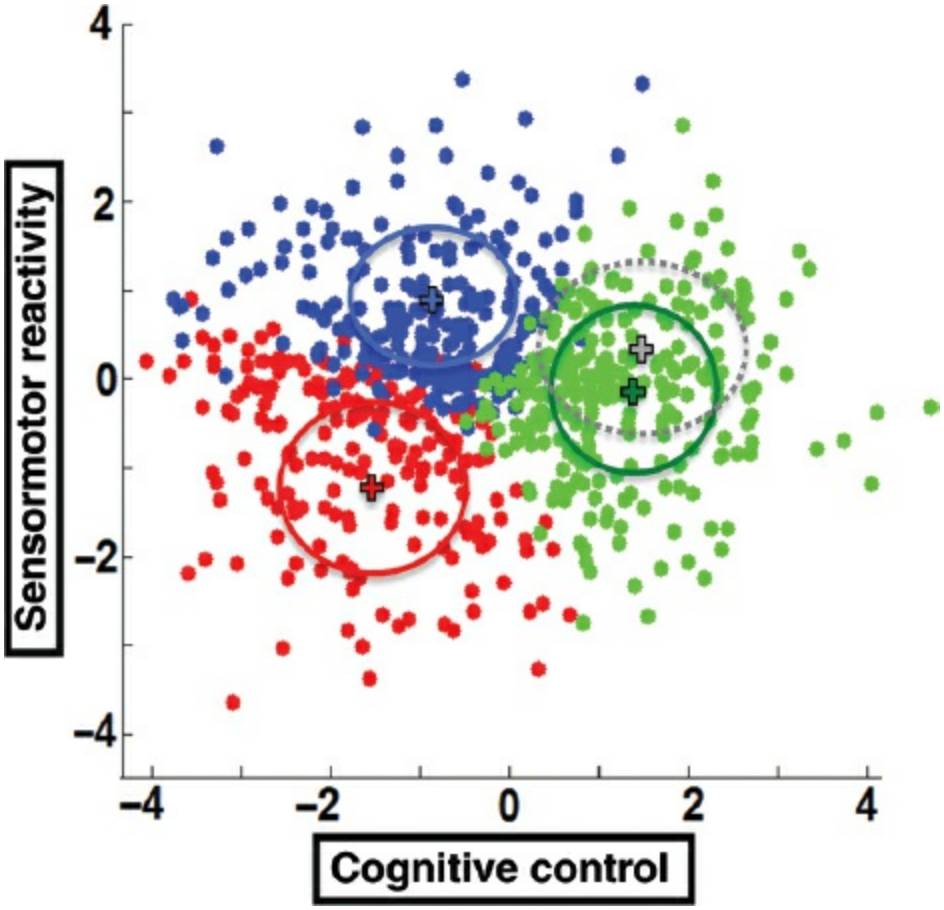

COLOR ATLAS 12.11–4. Sensorimotor reactivity by cognitive control as a function of biotype. Each dot indicates the location of an individual in this bivariate biomarker space (red, biotype-1; blue, biotype-2; green, biotype 3). The color-coded crosses and ellipsoids show the centroids and ±1 SD for each biotype. The gray cross and ellipse show the same values for the healthy group.

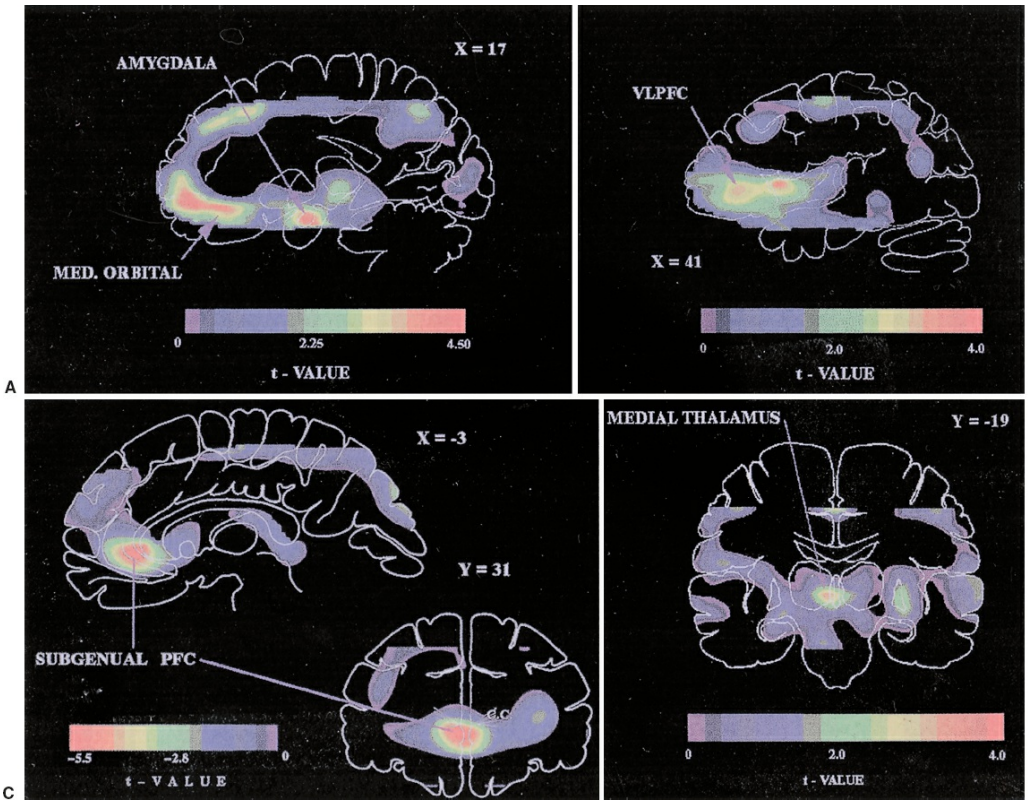

COLOR ATLAS 13.9–2. Composite coronal and sagittal sections of positron emission tomography scans show areas in which cerebral glucose metabolism is decreased in depressed patients relative to matched healthy controls. PFC, prefrontal cortex; VLPFC, ventrolateral prefrontal cortex. (From Wayne Drevets, M.D., and *Annu Rev Med*. 2002:49 by Annual Reviews, http://annualreviews.org.)

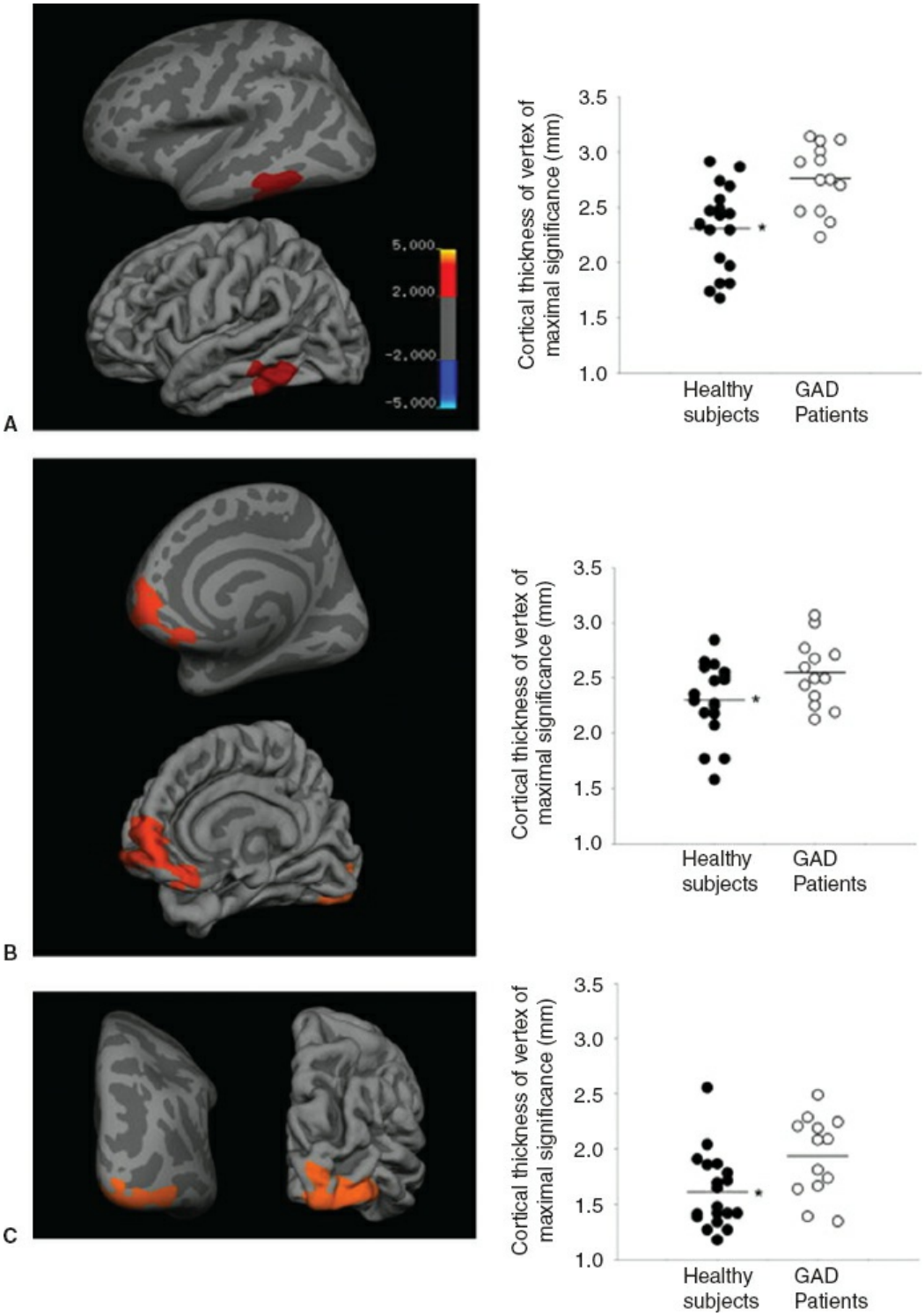

COLOR ATLAS 38–20. Cortical thickness in adolescents with generalized anxiety disorder (GAD) as compared to healthy controls. Warm colors (*red to yellow*) indicate greater cortical thickness in adolescents with GAD compared to healthy comparison subjects, and for each region, both cortical and inflated views are presented. Scatter plots are shown for the maximally significant vertex within each depicted region and healthy patients are represented by (•) while adolescents with GAD are represented by (°). The left temporal findings are shown in panel **A**, while the prefrontal differences and occipital findings are shown in panels **B** and **C**, respectively. Finally, an asterisk denotes a statistically significant difference between groups (p = 0.006 for the left temporal differences; p = 0.0004 for the frontal differences and p = 0.0001 for the occipital differences). (Reprinted with permission from Strawn JR, John Wegman C, Dominick KC, et al. Cortical surface anatomy in pediatric

patients with generalized anxiety disorder. *J Anxiety Disord.* 2014;28(7):717–723.)

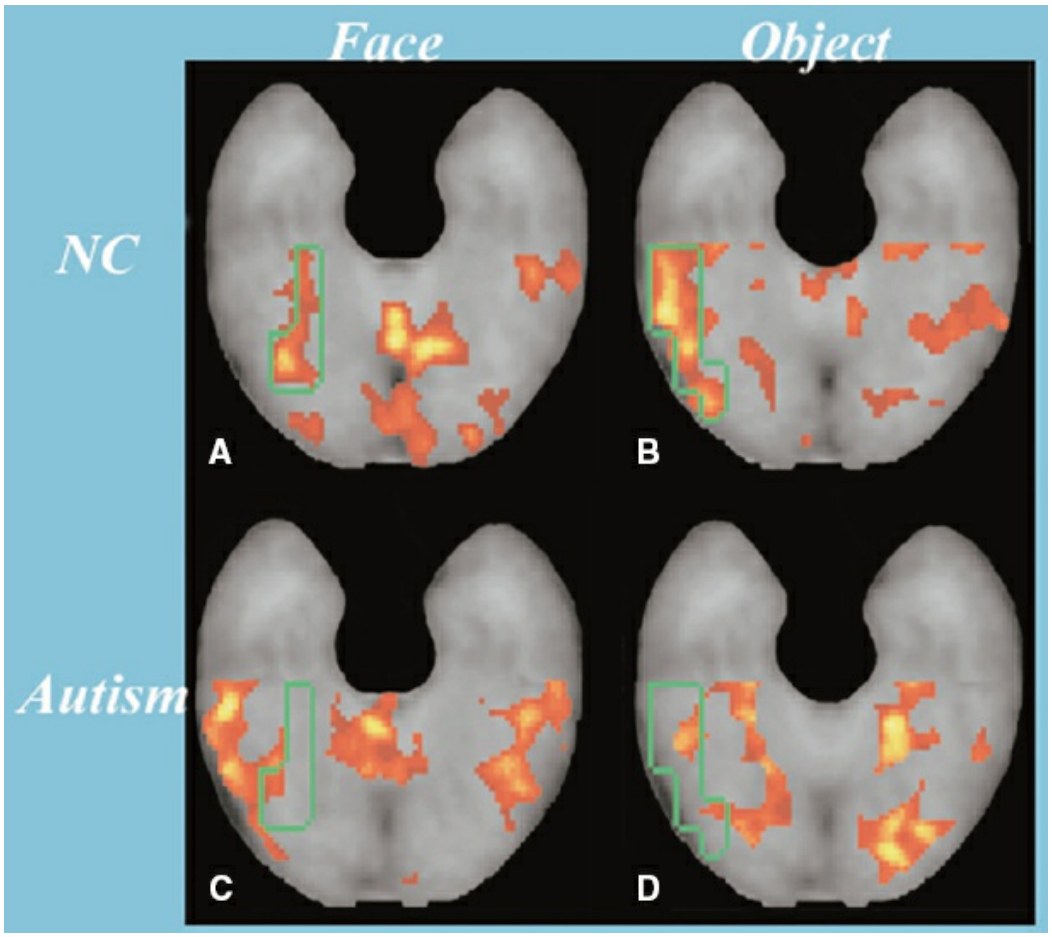

COLOR ATLAS 44–1. Composite functional magnetic resonance image comparing a group of 14 normal controls (NCs) and 14 individuals with autism or Asperger syndrome (autism), matched on age and intelligence quotient. Composite images show areas preferentially activated by faces (**A,C**) and nonface objects (**B,D**). Note that the fusiform "face area" has been defined by a green box in **A**. Among individuals with autism or Asperger syndrome, there is significant underactivation of the fusiform gyrus during face perception and compensatory activation of a region lateral to the fusiform. This more lateral region was most strongly activated by the object task among the NCs. Hypoactivation of the fusiform face area is the best replicated functional neuroimaging finding in the literature. (Adapted from Schultz RT, Gauthier I, Klin A, et al. Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. *Arch Gen Psychiatry.* 2000;57:331–340.)

Ancillary Color Atlas Figures

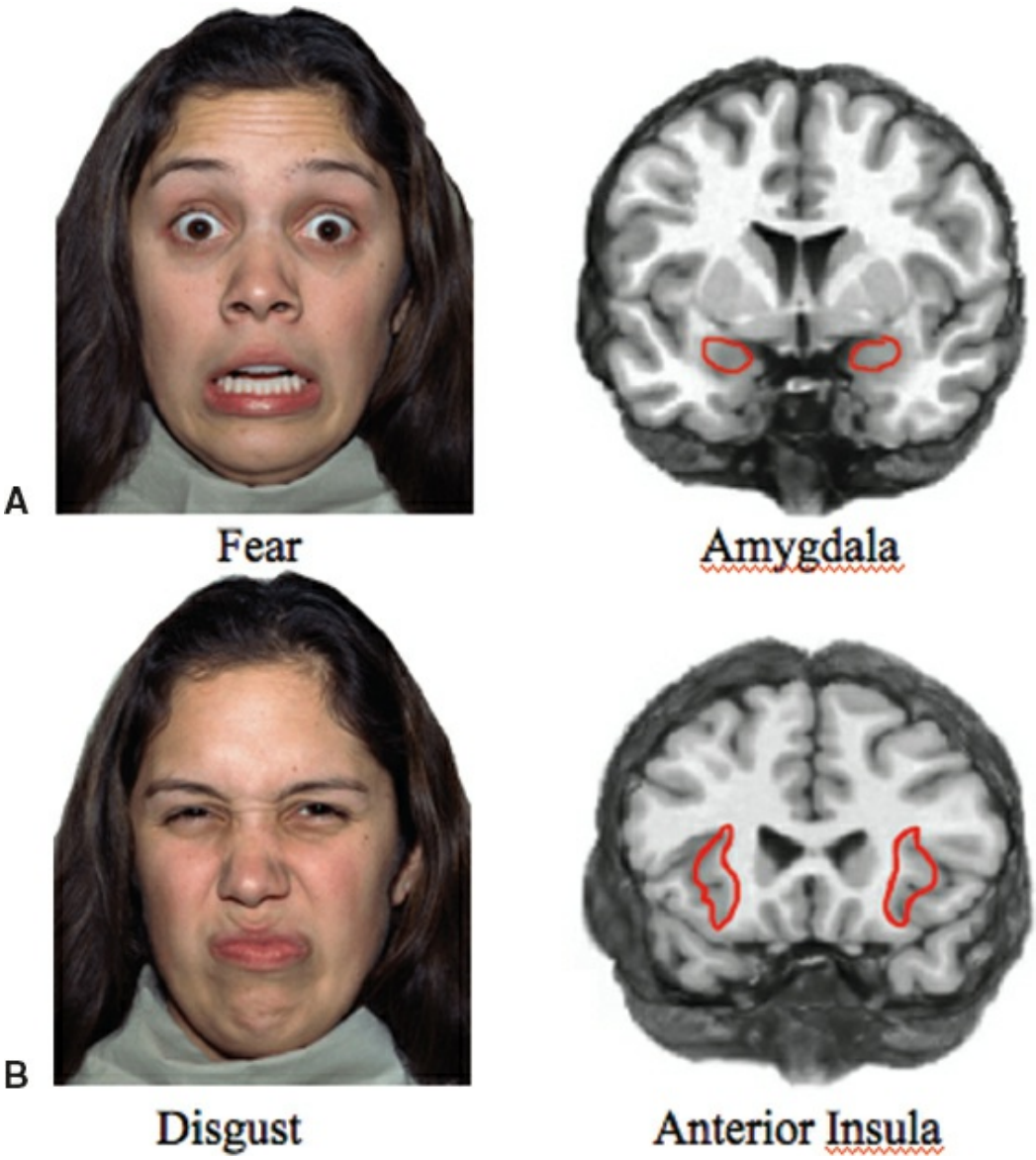

COLOR ATLAS FIGURE 1. Emotional facial expressions[1]. In the 1960s, Paul Ekman demonstrated that facial expressions of emotion are universal and thus, presumably, biological in origin as Charles Darwin once theorized (Ekman & Friesen, 1975). Since Ekman's discovery, photographs of emotional expressions have been widely used in psychological research to understand how people recognize another's emotions. Neuroimaging research has focused on two areas that are involved in emotion recognition: **(A)** The amygdala, known to be involved in fear conditioning, is most active when recognizing fear compared to other facial expressions (Whalen, 1998). **(B)** The anterior insula, associated with taste processing, subserves the recognition of another's disgust (Calder, Lawrence, & Young, 2001).

[1]Development of the MacBrain Face Stimulus Set was overseen by Nim Tottenham and supported by the John D. and Catherine T. MacArthur Foundation Research Network on Early Experience and Brain Development.

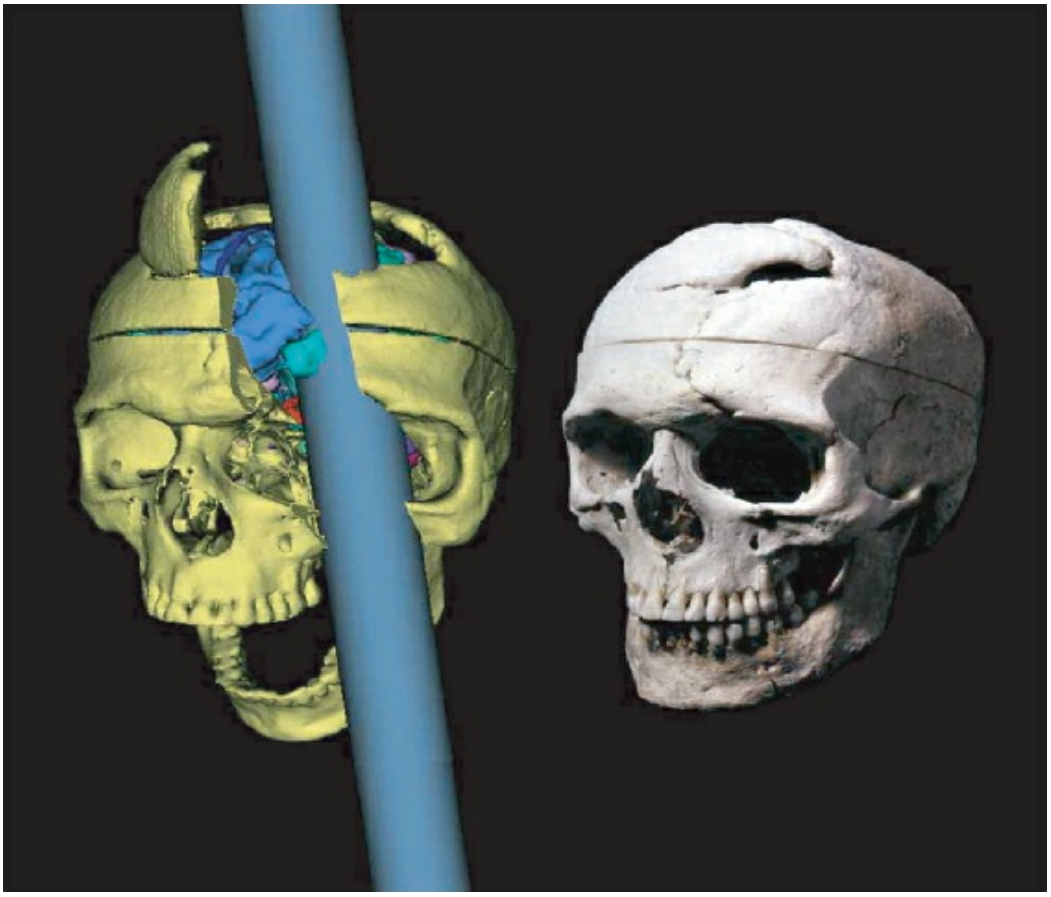

Color Atlas Figure 2. Three-dimensional reconstruction of then Phineas Gage trauma showing the trajectory of the penetrating rod injuring the left orbital and ventromedial prefrontal cortices. (From Ratiu P, Talos IF: *N Engl J Med*. 2004;351:e21, with permission.) (For another picture of Phineas Gage, see Fig. 1.23–1.)

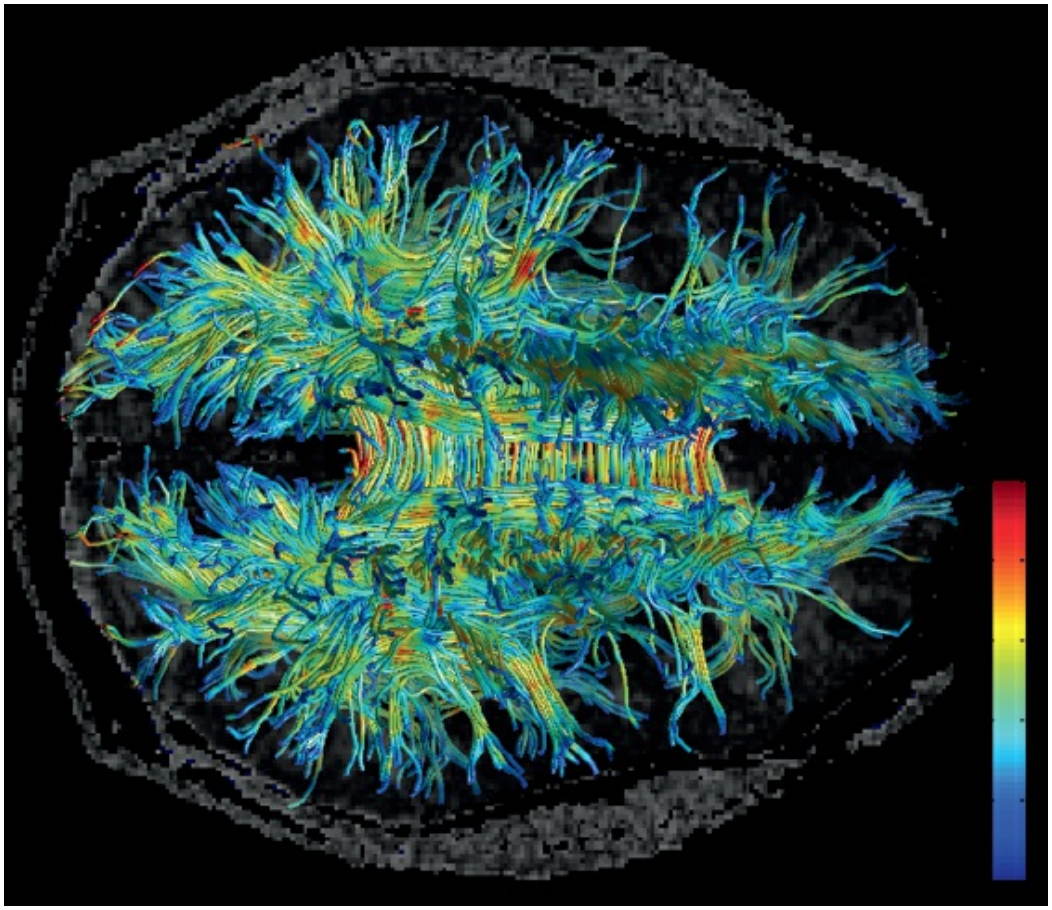

Color Atlas Figure 3. Three-dimensional reconstruction based on diffusion data acquired on a 3-T General Electric scanner, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA. Diffusion tensor imaging (DTI) image shows major long fiber tracts of the brain. (Courtesy of Hae-Jeong Park, Ph.D., at the Laboratory of Molecular Neuroimaging, Department of Diagnostic Radiology, Yonsei University College of Medicine, Seoul, South Korea.)

Index

Page numbers followed by *f* indicate figures; numbers followed by *t* indicate tables. Page numbers in **boldface** indicate major discussions.

A

- A-B-C-model
  - A (activating events), 2696
  - CBT, psychosis and, 2867, 2881, 2883
  - rational emotional behavior therapy, 2696
- A-delta fibers, 2388–2389
- A-fibers, 2398
- A-kinase-anchoring proteins (AKAPs), 129
- AA. *See* Alcoholics Anonymous
- AACAP. *See* American Academy of Child and Adolescent Psychiatry
- AACDP. *See* American Association of Chairmen of Departments of Psychiatry (AACDP)
- AADPRT. *See* American Association of Directors of Psychiatric Residency Training; American Association of Directors of Psychiatric Residency Training, Inc (AADPRT)
- AAIDD. *See* American Association on Intellectual and Developmental Disabilities (AAIDD, formerly AAMR)
- AAMC. *See* Association of American Medical Colleges (AAMC)
- AAP. *See* American Academy of Pediatrics (AAP)
- AARP. *See* American Association for Retired Persons
- ABAB design, 2704, 2704*f*
- Abbass, Allan A., 2797, 2804, 2811
- ABCDE model, 2306*t*
- ABCG1 transporter, aggression and, 2483
- Abe, Zenuemon, 259
- Abecarnil, 3003

- withdrawal, 1383
- Aberrant learning theory of addiction, 373
- Abilify. *See* Aripiprazole
- Ability, definition of, 4557
- Ablation, definition of, 4557
- Ablutomania, definition of, 4557
- ABMS. *See* American Board of Medical Specialties (ABMS)
- Abnormal, definition of, 4557
- Abnormal Involuntary Movement Scale (AIMS)
  - in physical examination, 956
  - tic disorder, 3640, 3654
- Abortion
  - definition of, 4557
  - techniques, 2511*t*
  - types of, 2511*t*
    - elective, 2511–2512
- *About Behaviorism* (Skinner), 930
- ABPN. *See* American Board of Psychiatry and Neurology (ABPN)
- ABPP. *See* American Board of Professional Psychology (ABPP)
- Abreaction, definition of, 4557
- Absence seizures, 244
  - in children, 650
  - definition, 4557
- Absolute refractory periods, 133
- Absolute thresholds, definition of, 4557
- Absorption, hypnosis and, 2710
- Abstinence
  - alcohol, 1275–1276, 1277
  - amphetamine, 1288, 1290
  - caffeine, 1300, 1300*t*
  - cannabis, 1310
  - definition of, 1374, 4557
  - opioids, 1370–1372
  - psychoanalysis, 2660
  - substance, 3768, 3793, 3860
  - tobacco, 1346
- Abstinence syndrome, definition of, 4557

- Abstract attitudes, definition of, 4557
- Abstract thinking
  - definition, 4557
  - disturbance of, 1121
- Abstraction
  - during adolescence, 3369
  - definition, 4557
  - semantic learning and, 2129
- Abulia
  - definition, 4557
  - dorsomedial prefrontal damage and, 450
- Abuse. *See* Alcohol abuse; Child abuse; Elder abuse; Fetal abuse; Maltreatment (child); Physical abuse; Sexual abuse; Substance abuse; *specific drugs*
- Abusive behaviors, modeling of, 2426
- Academic medical centers, hospitalists in, 4359
- Academic or educational problems. *See also specific problems*
  - definition of, 2439–2440
  - differential diagnoses, 2440
  - etiology of, 2440
  - future directions in, 2441
  - presentation of, 2440
  - treatment, 2440–2441
- Acalculia, 4557
- Acamprosate (Campral), **3054–3059**
  - chemistry, 3055, 3055*f*
  - disulfiram combined with, 3059
  - dosage/administration, 3059
  - drug interactions, 3058–3059
  - drug studies, 3056–3057
  - history, 3054–3055
  - laboratory interference, 3059
  - organ/system effects of, 3057
  - pharmacological actions, 3055–3056
  - precautions/adverse reactions, 3058
  - psychotherapy combined with, 3057
  - therapeutic indications, 3057–3058
    - alcohol-related disorders, 1278, 3054–3059
    - alcohol use relapse and, 582

    - PTSD and SUD comorbidity, 1823
    - substance use disorders, 3793
    - use in alcoholism, 83–84
- Acarophobia, 2346, 4557
- Acatalepsia, 4557
- Acathexis, 4557
- Accelerated experiential dynamic psychotherapy (AEDP), 2804
- Accelerated interactions, 4557
- Acceptance
  - CBT treatments on, 1773
  - change and, dialectic of, 2786
- Acceptance and commitment therapy (ACT), 2699
  - anxiety disorders, 1773
  - evaluation, 2863
  - habit reversal therapy and stimulant control, 1795
  - psychosis and, 2882
  - suicide prevention and, 2621
- Accident neurosis, 2406
- Accident proneness, 4557
- Accident victims, PTSD prevention, 1824
- Accommodation
  - AHP and, 137–138
  - during cognitive development, 3359
- *An Account of Diseases Peculiar to Women* (Gooch), 4535
- Accountable care organizations
  - overview of, 4318–4319
  - Affordable Care Act and, 4307–4308
  - grants from, 4307
  - implementation of, 4380
  - medical homes as building blocks for, 4318
- Accreditation, definition of, 4557
- Accreditation Council on Graduate Medical Education (ACGME), 4402. *See also* Psychiatric education
  - 360-degree evaluations, 4429
  - combined specialty track programs, 4405*t*
  - core competencies noted by, 4323
  - establishment of, 4404
  - examinations

        - clinical skills, 4415, 4416
        - competencies for, 4416*t*
        - diagnosis and treatment, 4418
        - grading process, 4425–4426*t*
        - patient interview, 4416–4417
        - subspecialty, 4427
    - osteopathic training, 4411
    - program requirements, 4407–4409
    - psychiatry milestones, 4409–4411
        - authors of, 4409*t*
        - description, 4410*t*
        - example, 4411*f*
        - format of, 4410*f*
        - subspecialty authors, 4413*t*
    - psychiatry residency review committee (2015), 4413*t*
- Acculturative stress hypothesis, 806
- Acenesthesia, 4557
- Aceruloplasminemia
    - cause of, 672
    - effects of, 556–557
    - MRI imaging in, 672*f*
- Acetaminophen (Tylenol)
    - overdose, laboratory evaluation, 1018
    - for pain management, 2377
- Acetazolamide (Diamox), side effects, 2096
- Acetyl-L-carnitine treatment, 2550
- Acetylcholine (ACh), **71–75**
    - in aging brain, 3992–3993
    - anatomy/function, 71–72
    - definition, 4557
    - delirium pathophysiology and, 1181
    - life cycle of, 72–73
    - mood disorders and, 1714–1715
    - muscarinic receptors, 73
    - neurotransmission by, 141
    - nicotinic receptors, 73–74
    - peripheral, 73

- receptors, 73–74
- schizophrenia and, 1461, 1495–1496
- sites of drug action, 72*f*
- synthesis of, 72

- Acetylcholine receptors (AChRs)
  - muscarinic, 73
  - nicotinic, 73–74
- Acetylcholinesterase (AChE)
  - in the basal ganglia, 29
  - inhibitors in Parkinson treatment, 1220
- Acetylsalicyclic acid (ASA, aspirin)
  - for migraine treatment, 624
  - in stroke prevention, 486
- AChE. *See* Acetylcholinesterase
- Achenbach System of Empirically Based Assessments (ASEBA), 3398–3399
- Achievement by proxy, 1849, 1859
- Achievement testing, 3415, 3416, 3425–3426
- *Achieving the Promise: Transforming Mental Health Care in America,* 4315
- Achluophobia, 4557
- Achromatopsia, central, 452
- "Acid." *See* LSD (lysergic acid diethylamide)
- Ackerman, Nathan, 2737, 2749, 2754
- Ackner, Brian, 1870
- Acne vulgaris
  - epidemiology, 2344–2345
  - pathophysiology, 2345
  - presentation, 2344–2345, 2345*f*
  - psychopathology, 2345
  - treatment, 2345
- Aconitine, 136
- Acoustic schwannoma, MRI, 269*f*
- ACP. *See* American College of Physicians (ACP); American College of Psychiatrists (ACP)
- Acquired immunity
  - cytokine mediators of, 183*t*
  - innate immunity *versus,* 180*t*
- Acquired immunodeficiency syndrome (AIDS). *See also* HIV
  - AIDS epidemic, 4547

- dementia related to, 574
- description of, 4557–4558
- epidemiology of, 564–566
- HIV and, 2514
- mania related to, 574
- neuropathies associated with, 632
- neuropsychiatric aspects of, **564–588**
- opportunistic CNS infections, 569–570
- psychiatric emergencies, 2624

- Acquired language problems, 3549
- Acquired neurogenic stuttering, 3563
- Acquired neuromyotonia, 676
- Acromegaly, 678, 4558
- Acroparesthesias, 656
- Acrophobia, 79, 4558
- Acruate nuclei, 36
- Actaphasia, 4557
- ACTH. *See* Adrenocorticotropic hormone
- Actigraphy, 343, 2106, 3934, 3938
- Acting out, 4558
- Action and commitment therapy (ACT), 2853
- Action awareness, 335–336
- Action potentials
  - depolarization, 103
  - description of, 131
  - membrane properties of, 132–133
  - neuronal, 134*f*
- Action research, Lewin on, 2842
- ACTIVE. *See* Advanced Cognitive Training for Independent and Vital Elderly study
- Active imagination, 2823
- Active treatment, 4338
- Activities of daily living (ADLs)
  - by aging adult, 3975–3976, 4007
  - depression and, 484
  - diabetes and, 2233
  - everyday functioning, 4011, 4011*t*
    - instrumental, 4025

- Activity group psychotherapy, 4558
- Activity schedules, CBT and, 4182
- Actualization, 4558
- Aculalia, 4558
- Acupuncture
  - as alternative treatment, 2542
  - anxiety disorders, 1782
  - assessment, 2544–2546
  - as energy therapy, 2546, 2561–2562
  - PTSD, 1824
- Acute and transient psychotic disorders, **1574–1580**
  - classification of, 1172
  - comparative nosology, 1574–1577, 1576*t*, 1577*t*
  - course and prognosis, 1580
  - diagnosis and clinical features of, 1578–1579
  - differential diagnoses, 1579–1580
  - epidemiology, 1577
  - etiology, 1577–1578
  - history of, 1574, 1575*t*
- Acute confusional state, 1180
- Acute depersonalization, 1881
- Acute dissociative reactions to stressful events
  - definition, 1948
  - diagnosis and clinical features, 1948–1949
  - treatment, 1950
- Acute intermittent porphyria (AIP), 660
- Acute stress disorder (ASD), 1815
  - acute dissociative reactions to stressful events *versus,* 1948
  - after disasters, terrorism, or war, 2567*t*
  - in cancer patients, 2259
  - child/adolescent, 3705, 3914
  - classification of, 1158
  - description of, 4558
  - dissociative amnesia and, 1882
  - dissociative amnesia *versus,* 1937
  - dissociative disorders and, 1871
  - dissociative identity disorder and, 1918

- DSM-5 on, 1874, 1878, 1907
- genetic mapping of, 3453, 3456
- hypnotizability and, 2718
- in pediatric burn patients, 2385
- pharmacogenomics of, 3455
- PTSD and, exposure-based CBT for, 1824
- STAPP and, 2798
- STDP and, 2801

- Acute stress reactions, classification, 1168
- Acute tryptophan depletion (ATD) test, 368
- Acylation, description of, 149
- AD. *See* Alzheimer disease
- ADA. *See* Americans with Disabilities Act (ADA)
- Adachi, Tononori, 2727
- Adalimumab for psoriasis, 2344
- ADAMHA. *See* Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA)
- ADAMS. *See* Aging, Demographics and Memory Study
- Adams, Joseph, 4535
- Adaptation
  - biogenetic factors and, 2144
  - definition of, 4558
  - stress and, 2336
  - transplant time course and, 2364
- Adaptional approach, 4558
- Adaptive behaviors
  - definition of, 4558
  - domains of, 2432
  - intellectual disability and, 3494–3496, 3495*t*, 3496*t*
  - rating scales for, 3402
- Adaptive function
  - impaired, 637–638
  - intellectual disability and, 3494–3496, 3495*t*, 3496*t*
  - malingering and, 2407
  - rating scales for, 3402
- Adaptive immunity, components of, 181–183
- ADAS. *See* Alzheimer Disease Assessment Scale
- ADD. *See* Attention deficit disorder

- Addiction. *See also* Neurocentrism; *specific types of addiction*
  - associative learning and, 414–416
  - benzodiazepines, 1376
  - brain pathways in, 373–378
  - to cocaine, 397–398
  - cycle of, 371*f*
  - definition of, 2041
  - to foods, 2213
  - hallucinogen-assisted psychotherapy and, 2889
  - HIV infections and, 578–582
  - hypnosis and, 2730–2731
  - learning-based treatments of, 415–416
  - LSD, 2891–2892
  - monitoring of patients, 581–582
  - opponent process theory, 372*f*
  - in patient interview, 950–951
  - physician, 2601
  - relapse prevention, 581–582
  - sexual, 2041–2044
  - substrates of, 378–382
  - theories of, 372–383
  - treatment, amantadine, 2964
- Addiction psychiatry
  - certificate examinations, 4427
  - training, 4405, 4406*t*
- Addictive disorders
  - affect regulation, 2046
  - after TBI, 537–538
  - behavioral inhibition, 2046
  - genetics of, 2046
  - motivation-reward, 2046
  - neuroimaging, child/adolescent, 3475
  - prenatal stress and, 2046–2047
  - process of, 2044–2052
- Addictive processes
  - definition of, 2045
  - development of, 2052

- healing of, 2057–2063
  - affect-regulating agents, 2057–2058
  - mu-opioid receptor antagonists, 2058
  - pharmacotherapy, 2057–2058
  - psychodynamic psychotherapy, 2058–2059
- impaired self-regulation, 2049–2050
- phases of, 370–372
- psychological formulation, 2048–2052

- Addison disease. *See also* Adrenocortical insufficiency
  - characteristics of, 167
  - presentation of, 682–683
- Adenoma sebaceum, 456
- Adenophypophysis, 677
- Adenosine
  - caffeine and, 1295, 1295*f*
  - sleep and, 346–347
- Adenylyl cyclase, 2913, 2913*t*
- Adenylyl cyclase 8 *(ADCY8)* gene, 1889
- ADHD. *See* Attention-deficit/hyperactivity disorder (ADHD)
- Adherence to treatment, 4303
- Adiadochokinesis, 4558
- Adipocytes, energy homeostasis and, 362
- Adiposity sensing, 362–363
- Adjective Checklist (ACL), 998*t*
- Adjustment, definition of, 4558
- Adjustment disorder(s)
  - about, 1841
  - after disasters, terrorism, or war, 2567, 2567*t*
  - in brain tumor patients, 500
  - clinical features, 2121–2122
    - subtypes, 2122, 2122*t*
  - comparative nosology, 2116–2117
  - course/prognosis, 2123–2124
  - cultural perspectives, 2119
  - definitions of, 2117, 4558
  - diagnosis, 2117, 2121–2122
  - differential diagnoses, 2122–2123

- DSM-5 on, 1874
- epidemiology of, 2120
- etiology, 2120–2121
- future directions, 2124
- history of literature on, 2117
- ICD classification, 2119
- reliability, 2118
- STAPP and, 2798
- STDP and, 2801
- stressor criteria, 2117–2118
- suicide and, 2613
- temporal course of, 2118
- in transplant populations, 2365–2366
- trazodone drug studies, 3249
- treatment of, 2123–2124
- validity of diagnosis, 2118–2119

- Adler, Alfred, 2641, 2644
  - family therapy and, 2749
  - group psychotherapy and, 2736
  - paradoxical communication, 910f, 911
  - personality theory, 910
  - psychopathology theory, 910
  - psychotherapy, 910–911
  - reframing, 911
  - *Study of Organ Inferiority and its Psychical Compensations,* 4537
  - work of, 4558
- Adlerian, definition of, 4558
- Administration for Healthcare Research and Quality (AHRQ), 4307
- Administrative psychiatry, 4558
- Admission diversions services, 4344–4345
- Adolescence/adolescents. *See also* Children
  - adoption's effect on, 3827
  - aggressive behavior in, 2487
    - development and, 2484
    - management of, 3765–3766, 3799
    - treatment of, 2499
  - alcohol dependence in, 274

- alcohol use disorder of, 1275, 1276*t*
- alcohol use during, 3366, 3367, 3372, 3768, 3858
- anorexia nervosa in, 2070
- antiretroviral therapy in, 645–646
- anxiety disorders
  - epidemiology, 1732–1734, 1733*t*
  - management, 1784
  - neurodevelopment and, 1765–1767, 1765*f*, 1766*f*
- BDD, 1790, 1791
- behavioral therapy, 3814
- behaviorally acquired HIV, 645–646
- belief systems of, 3369
- cannabis, effects of, 1308
- care world-wide, 4525
- cognitive-behavioral therapy, 3814
- communication with, 2310
- conduct disorder, 3612, 3619–3620
- DBT and, 2795
- definition of, 3365, 3810, 4558
- development, 3364–3374
  - brain, 3370–3371, 3461–3463
  - cannabis use affecting, 1308
  - cognitive, 3369
  - components of, 3368–3370
  - defined, 3365
  - emotional, 3371
  - environmental influences on, 3371–3374
  - Internet effect on, 3373
  - of morality, 3369–3370
  - normal, 3365–3366
  - online gaming's effect on, 3373–3374
  - overview, 3364–3365
  - physical, 3368–3369
  - risk-taking behavior and, 3372–3373
  - of self, 3370
  - of sexual behavior, 3371–3372
  - social, 3369

- social media and, 3373
- stages of, 3366–3368
- developmental coordination disorder and, 3542
- diabetes in, 2232
- with diabetes type 1, 2247
- dissociative disorders, 1870, 1884, 1901–1902
- dissociative identity disorder
  - course and prognosis, 1918
  - diagnosis and clinical features, 1914–1915
- divorce during, 3785
- eating disorders, 3768
- effects of prematurity in, 643
- employment, 3943
- environmental treatments, 3815
- evaluation considerations for, 3376, 3385–3387, 3387*t* (*See also* Child psychiatric evaluations)
- future directions, 3815
- GAD, 1767
- gambling disorder, 1808
- gender dysphoria in, 2028
  - diagnostic criteria, 1985
  - treatment in, 2033
- gender incongruence of
  - definition of, 2024
  - ICD-11, 2029
- generalized anxiety disorder, 3470, 3472*f*
- group therapy, 3762
- history, at patient interview, 966
- HIV/AIDS in, 645–648, 3854, 3855
- homosexuality and, 2002–2003
- hormones during, 1962
- hospital care of, 4341
- hospitalization, response to, 3815, 3849
- illnesses
  - falsification of, 2408
  - response to, 3849
  - transition to adulthood, 3851, 3852*t*
- informed consent, 3863

- inpatient admissions for, 4313
- inpatient treatment of, **3797–3801**
- juvenile justice cases, 3865–3867
- as lucrative psychiatry market, 4293
- major depressive disorder, 3464–3466
- medical issues, in gender dysphoria, 2036
- mental health of, 758
- Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA), 826
- migraine treatment in, 625
- Multidimensional Inventory of Dissociation (MID), 1877*t*
- National Comorbidity Survey Replication Adolescent Supplement (NCS-A), 826–827, 827*t*
- neuron changes in, 5
- normality in, 3365–3366, 3811–3812
- OCD, 1788
  - neuroimaging, 3467–3468
- origins of homosexuality, 1997
- pathology in, 3811–3812
- prenatally-acquired HIV, 647–648
- psychiatric disorders in, 4520
  - advances in, 3348
  - onset of, 3365
- psychiatric emergencies, pharmacotherapy, 2628
- psychoanalysis, 3813–3814
- psychosocial interventions, 3813–3815
- psychostimulants in, 3227–3228
- PTSD
  - prevalence studies, 1816
  - psychotherapy, 1823
- religious perceptions, 2585–2586
- reproductive development in, 2506–2507
- residential treatment, 3815
- schizophrenia risks, 1449–1450
- Self-Directedness and, 2167
- with serious emotional disorders, 4320–4321
- sexual behavior during, 1962
- sleep in, 342

- social anxiety disorder in, 1766–1767
- specific learning disorder in, 3534
- substance use disorders, 3475–3476, **3857–3861**
- suicide and, 2614–2615, 2621, **3680–3686,** 3685, **3694–3700,** 3699
  - antidepressant-associated behavior and, 1673, 1680, 1779–1780
- systemic treatments, 3815
- temperament and, 3488
- terminally ill, 2309
- terrorism and, 3915, 3916
- tobacco use, 3858
- transition in, 3851, 3852*t*
- traumatic brain injury (TBI) in, 651–654
- treatment considerations, 3812–3813
- treatment modalities, 3813–3814

- Adolescent DES (A-DES), 1876*t*, 1896, 1902. *See also* Dissociative Experiences Scale
- Adolescent mania, 4558
- Adoption
  - as child psychiatric evaluation consideration, 3379
  - developmental issues for, 3825
  - epidemiology, 3824–3825
  - future directions, 3827
  - history, 3823–3824
  - international, 3826
  - mental health risks, 3825–3826
  - overview, 3817
  - psychiatrist's role in, 3826–3827
  - studies of
    - ADHD, 3588
    - antisocial personality disorder, 2491
    - conduct behavior, 3614
    - environment and psychiatric disorders, 2529
    - genetics and psychiatric disorders, 2527–2528, 2528*t*
    - mood disorders, 1620
    - schizophrenia, 1438*f*, 1439–1440, 1441*t*, 1446
  - temperament studies on, 3483
- The Adoption and Safe Families Act (1997), 3818, 3821
- Adrenal cortex, 2239

- Adrenal disorders. *See also specific Adrenal diseases*
  - hormone production and, 2243
  - laboratory evaluation, 1013
  - presentation of, 682–684
- Adrenal insufficiency, DHEA in, 170
- Adrenal steroids, immunological effects, 186*t*
- Adrenaline. *See* Epinephrine
- Adrenarche, 170
- Adrenergic, definition of, 4558
- α-Adrenergic agents
  - ADHD, 3594–3595
  - autism spectrum disorder, 3586
  - tic disorders, 3645, 3647*t*, 3659, 3661*t*
- α-Adrenergic antagonists
  - aggression in autism, 2500
  - anxiety disorders, 1781
  - dissociative identity disorder, 1928
- β-Adrenergic antagonists
  - aggression
    - Alzheimer disease, 2502
    - autism, 2500
    - intermittent explosive disorder, 2501
    - mentally retarded children, 2500
    - TBI, 2502
  - anxiety disorders, 1781
  - dissociative identity disorder, 1927*t*, 1928
- Adrenergic receptors
  - classes of, 67
  - sexual dysfunction and, 1972
- α-Adrenergic receptors
  - agonists, **2945–2950**
    - clonidine as, 2945–2950
    - guanfacine as, 2945–2950
  - subtypes, 67
- α-2C Adrenergic receptors, 1742
- β-Adrenergic receptors
  - antagonists, **2951–2956**

        - absorption, 2951
        - anxiety, 2954
        - chemistry of, 2951, 2951*f*
        - distribution, 2951
        - dosage/administration, 2955–2956
        - drug interactions, 2955
        - drug studies, 2952
        - elimination, 2951
        - history of, 2951
        - laboratory interference, 2955
        - lithium tremor, 2953–2954
        - MDD, 2954
        - neuroleptic-induced akathisia, 2955
        - organ/system effects, 2952–2953, 2953*t*
        - performance anxiety, 2954
        - pharmacodynamics, 2951–2952
        - pharmacokinetics, 2951–2952, 2952*t*
        - pharmacological actions, 2951–2952, 2952*t*
        - PTSD, 2954–2955
        - substance use disorders, 2954
        - teratogenic effects, 2953
        - therapeutic indications, 2953–2955, 2953*t*
        - violence, 2954
    - subtypes, 67
- Adrenergic system, description of, 4558
- Adrenocortical hyperactivity
    - mood disorders and, 1715–1716
- Adrenocortical insufficiency, 2245–2246
    - course/prognosis, 2245
    - definition of, 2245
    - diagnosis, 2245
    - epidemiology, 2245
    - etiology, 2245
    - nosology, 2245
    - pathology, 2245
    - treatment, 2245–2246
- Adrenocorticotropic hormone (ACTH)

- after TBI, 525
- arginine vasopressin and, 1748
- depression and animal modes of stress and, 1713
- function of, 2239
- immunological effects of, 185*t*
- pituitary tumors secreting, 678
- production of, 2239
- release of, 2337
- responses to stress, 2318
- trauma response and, 2569

- Adrenogenital syndrome, 2506
- Adrenoleukodystrophy
  - features of, 467*t*
  - laboratory evaluation, 1016
  - MRI imaging, 660*f*
- Adult ADHD Self-Report (ASRS), 3600
- Adult Attachment Inventory, 2720, 2721
- Adult neuronal lipofucinosis (ANDL)
  - forms of, 658
  - presentation, 658
- Adult Protective Services (APS), 4252
- Adults/adulthood, **3941–3946**. *See also specific disorders*
  - antisocial behavior, **2413–2430**
  - anxiety disorders, 1731–1732, 1732*t*, 1733*t*
  - brain connections in, 5–6
  - burn injuries in
    - risk of, 2374
  - criminality, **2413–2430**
  - emotional intimacy, 3942–3943, 3943*f*
  - employment, 3943–3944
  - female anatomy, 1954–1955
  - future directions, 3945–3946
  - gender dysphoria
    - diagnostic criteria, 1985
    - treatment in, 2033–2035
  - gender dysphoria in, 2028
  - gender incongruence of, 2029

- history, at patient interview, 967
- illnesses, transition from adolescents, 3851, 3852*t*
- intellectual assessment of (*See* Clinical neuropsychology and intellectual assessment of adults)
- male anatomy, 1954
- maturity, 3941–3942
- medical issues in gender dysphoria, 2036
- mental health, 3942
- mood disorders and, 1617–1618
- neuropsychiatric disorders and aggression, 2489–2490
- physical and sexual abuse, 2538–2542
- physical health, 3942
- play, 3943–3945
- post-illness PTSD, 2331
- psychological testing of, 994–1007
- Self-Transcendence and, 2167
- with serious emotional disorders, 4321–4322
- services needed by, 4306
- with severe/persistent mental illness, 4296
- sexual behavior during, 1962
- sleep requirements of, 341
- spectator play, 3944–3945
- treatment barriers, 4305
- violence, **2413–2430**
- wisdom and fulfillment, 3942–3945, 3942*f*
- Wow Moments!, 3943–3945
- young adults
  - communication with, 2310
  - effects of prematurity in
    - psychopathologic outcomes, 643
  - homosexuality and, 2003
  - Self-Directedness and, 2167
  - sleep phases, 210*f*
- Advance directives
  - competency of older adult, 4238
  - legal issues, 2305*t*
- Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) study, 3994, 3995

- Advanced sleep phase circadian rhythm disorder, 4558
- Adventitious reinforcement, 4558
- Adverse Childhood Experiences (ACE)
  - dissociative disorders, 1883
    - childhood autobiographical memory disturbances, 1885–1886
  - dissociative identity disorder, 1905
  - psychiatric, behavioral, and medical outcomes, 1885*t*
- Adverse drug reactions, 2543–2544, 2544*t*. *See also specific drugs*
- Adverse effects, definition of, 4558
- Adversities, symptoms affected by, 1118
- Advocacy, ethics of, 3876–3877
- Adynamia, 4558
- Adynamic ileus. *See* Paralytic ileus
- Aerobic exercise
  - in CAD, 2197
  - depression reduced by, 2197
  - and healthy brain aging, 3994
- Aerophagia, definition of, 4558
- Affect-fantasy, 4558
- Affectionate transference, 4558
- Affective disorders. *See also* Mood disorder(s)
  - concept, 1599, 1603
  - dissociativity, 1870
  - endogenous, 1604
  - use of term, 1630–1631
- Affective dysregulation, 1609–1610
- Affective hallucinations, 4558
- Affective interactions, 4558
- Affective mentalizing, 2896
- Affective prosody, testing of, 464
- Affective symptoms
  - of disorganization syndrome, 1419
  - neurocognitive impairment and, 1508
  - schizophrenia and, 1508
- Affective syndrome, 450–451
- Affective temperaments
  - ancient Greeks and Romans on, 1602–1603

- moods and, 1632–1633
- as quantitative trait for bipolar disorder, 1622*f*

- Affect(s). *See also* Mood
  - of aging adult, 3974–3975
  - attunement, 2646
  - definition of, 4558
  - dissociation of, 1943, 1944
  - ethological considerations, 1632
  - experience of, AEDP and, 2804
  - identification, in MBT, 2899–2900
  - at patient interview, 952, 966
  - psychosis and, 2880
  - symbolization of, 2051
  - theory of, 1907, 1925
- Afferent systems
  - corticostriatal, 29, 30f
  - nigrostriatal, 29
  - sensory, 2337
- Affordable Care Act (2010)
  - accountability in, 4306, 4307
  - barriers removed by, 4305
  - better care
    - evidence-based psychiatry, 4295–4300
    - psychiatry multidisciplinary teams, 4300–4301
  - better health
    - health promotion, 4302–4305
    - psychiatry and public agencies, 4305–4307
    - public health, 4301–4302
  - Health Care Exchanges in, 4307
  - lower costs, 4307–4308
  - Medicaid expansion, 4379–4380
  - predecessors to
    - groundwork for, 4294
    - Quality Chasm report influence in, 4289
  - psychiatric emergency services and, 2622
  - reforms in, 4318
- Afghanistan war veterans

- dissociation and combat trauma in, 1873
- PTSD and, 1816
- suicide and, 2568–2569

- African Americans. *See also* Demographics; Race/ethnicity
  - discrimination, and psychiatric service utilization, 802–803
- African tarantula *(Hysterocrates gigas)* toxin, 139
- AFSP. *See* American Foundation for Suicide Prevention (AFSP)
- Aftercare, description of, 4558
- Age progression hypnosis, 2711–2712
- *Agelenopsis aperta* (funnel-web spider) toxin, 139
- Agency
  - abnormalities in, 465
  - structure of, 335
  - willful action and, 336
- Agency for Healthcare Research and Quality (AHRQ)
  - evidence-based care, support from, 4296, 4308
  - funding for, 4319
  - practice guidelines, 971–972
- Age(s). *See also specific age groups*
  - of adoption, effect of, 3826
  - Alzheimer disease risk factor, 1202, 1204
  - brain tumor incidence and, 491*t*
  - burn injuries and, 2374
  - child's response to illness/hospitalization and, 3848–3849
  - dementia and, 476, 1177
  - distribution in U.S. population, 795
  - homicide rates and, 2475*f*
  - imaginal flooding and, 2691
  - interpersonal aggression and, 2476
  - life span issues
    - of autism spectrum children, 3925–3927
    - of bipolar disorder children, 3929
  - mood disorders and, 1615–1616
  - nonepileptic seizures and, 650–651
  - normal
    - cognitive decline, 1222–1223
  - onset of epilepsies/electrochemical syndromes, 649*t*

- onset of neurodevelopmental disorders, 438, 438*f*
- orgasm frequency and, 1968
- PD and, 542
- pedophilia and, 2017
- personality disorders and, 2167
- regression, hypnosis and, 2711
- schizophrenia risk, parental, 1448
- sleep and, 342
- suicide and, 2612*t*
- variant normal EEGs in, 237

- Ages of onset
  - for adolescent psychiatric disorders, 3365
  - anxiety disorders, 1734
  - for autism spectrum disorder, 3579–3580, 3579*f*, 3580*t*
  - BDD, 1790, 1791
  - for childhood psychiatric disorders, 3349
  - for dementia, 1199
  - for frontotemporal dementia, 1210
  - hoarding disorder, 1793
  - OCD, 1786, 1787
  - trait-like abnormalities, mood disorders and, 1718
- Ageusia, 4558
- Aggression, **2471–2504**
  - ADHD and, 2488
  - adolescent, 2487
  - adult-onset, 2487
  - affective, treatment of, 2172
  - alcohol abuse and, 2492
  - anabolic steroids and, 2493
  - animal models of, 324
  - autism and, 2488
  - borderline personality disorder and, 2494
  - care-resistant, Alzheimer disease and, 2502
  - cerebrovascular disease and, 2495
  - child/adolescent
    - evaluation considerations, 3378
    - neuropsychiatric disorders, 2487–2489

- treatment, 2499–2501
- circuit loops supporting, 2478
- cocaine and, 2492–2493
- cognitive markers for, 2486
- conduct disorders and, 2436–2437, 2489
- cortisol and, 2481
- definition of, 2471, 4558
- development and, 2484–2487
- dissociation and, 2491–2492
- early-onset, 2487
- environmental influences on, 2425
- in epilepsy patients, 516, 517*t*, 650, 2496
- episodic dyscontrol of, 454–455
- future directions, 2503–2504
- gene–environment interactions, 2483–2484
- gene–gene interactions, 2483–2484
- genetics and, 2482–2484, 2482*t*
- gun control and, 2476
- in HD, 556
- hemispheric specialization of, 2478
- hormone-induced, 165
- 5-HT and, 1714
- hypoarousal- *versus* hyperarousal-related, 2472
- hypothalamic lesions and, 677
- ictal, treatment of, 2172
- incidence and prevalence, neuropsychiatric disorders and, 2497
- intermittent explosive disorder and, 2494–2495
- interpersonal, 2491–2492
- intimate partner violence and, 2486–2487
- introduction, 2471–2476
- life courses of, 2487
- liminal processing of, 2479*f*
- major depression and, 532
- major mental illness and risk for, 2490
- management of
  - ADHD comorbidity, 3595–3596
  - child evaluation consideration, 3378

- dementia and, 1196–1197
- emergency psychiatric evaluation for, 3443–3444, 3444*t*
- group therapy for, 3765–3766
- the media and, 2485
- medication effects and, 2492–2493
- medications for stabilization of, 2171–2172, 2172*t*
- mental retardation and, 2488
- methamphetamine and, 2493
- neuroanatomy and, 2476–2478, 2479*f*
  - amygdala, 2477
  - anterior cingulate gyrus, 2477–2478
  - brainstem, 2476
  - circuit loops supporting, 2478, 2479*f*
  - dorsolateral prefrontal cortex, 2478
  - frontal lobes, 2478
  - hemispheric specialization, 2478
  - hypothalamus, 2476–2477
  - nucleus accumbens, 2477
  - septal areas, 2477
  - thalamus, 2477
  - ventral striatopallidal system, 2477
  - ventromedial prefrontal cortex, 2478
- neurochemistry of, 2479–2482
  - cortisol, 2481
  - dopamine, 2480
  - epinephrine, 2480
  - GABA, 2480
  - hormones, 2480–2482
  - metabolism and, 2482
  - neurotransmitters, 2479–2480
  - norepinephrine, 2480
  - peptide hormones, 2481–2482
  - serotonin, 2479–2480
  - testosterone, 2480–2481
- neurodegenerative disease and, 2495–2496
- neuropsychiatric disorders and, **2487–2497**
  - adult, 2489–2490

- childhood, 2487–2489
- neuropsychiatry of, **2471–2504**
- nonsanctioned, 2472–2476, 2473*t*, 2474*f*, 2475*f*
- ODD and, 2488–2489
- organic-like, treatment of, 2172
- organized, treatment of, 2172
- peer influences and, 2485
- phencyclidine and, 2493
- Point Subtraction Aggression Paradigm, 170
- postnatal factors, 2485
- pre- and perinatal factors, 2484–2485
- predatory, treatment of, 2172
- proactive, 2472–2473
- proactive *versus* reactive, 2472–2473
- proposed predictive markers of, 2486
- psychiatric emergency due to, 2628
- psychopathy and, 2490
- PTSD and, 2494
- reactive, 2472–2473
- sanction *versus* nonsanctioned, 2472
- sanctioned, 2472
- serotonin and, 2111
- serotonin gene variants and, 2483
- socialization to violence and, 2486
- SSRIs and, 2493
- subliminal processing of, 2479*f*
- substance abuse and, 2492–2493
- TBI and, 471–472
- testosterone and, 2475–2476, 2480–2481
- trait or disorder, 2484
- traumatic brain injury and, 2495
- treatment, 2497–2503, 2498*t*
  - for acute episodes, 1523–1525
  - adults, 2501–2503
  - anger management, 2499
  - barriers to, 2497
  - behavioral modification, 2498

        - CBT, 2498
        - child/adolescent, 2499–2501
        - cognitive-behavior therapy, 2498
        - effectiveness, 1520, 1523
        - first episode, 1520
        - future directions, 1530–1531
        - history, 1519–1520, 1521–1522*t*
        - insight-oriented psychotherapy, 2498
        - lithium, 3093
        - maintenance, 1529–1530
        - negative symptoms and, 1528
        - neuropharmacological, 2498–2499
        - nonpharmacological, 2498, 2498*t*
        - pharmacotherapy, 2498*t*
        - phases of, 1520
        - poor response, strategies for, 1528–1529
        - problems regarding, 2497
        - psychoanalytic therapy, 2498
        - psychosocial treatment and, 1530
        - side effect management, 1525–1528
        - suicidal behavior, 1528
        - supportive psychotherapy, 2498–2499
    - type 1 offences, US rates of, 2473*t*
    - types, 2471–2473, 2472*t*
    - typology of, 2471–2473
- Aggression-turned-inward model, 1605
- Aggressive drive, definition of, 4558–4559
- Aging. *See also* Elderly persons; Geriatric psychiatry
    - apoptosis and, 46
    - circadian rhythms and, 214
    - definition of, 4559
    - dementia and, 276
    - effects of
        - menopause or, 2523–2524
    - neurodevelopmental processes in, 54*f*
    - neurosteroids and, 118
    - normal, 1935

- spleen, noradrenergic innervation, 185
- Aging, Demographics and Memory Study (ADAMS)
  - Alzheimer disease in older adults, 3963
  - caregiver depression, 4209
- Aging brain. *See* Geriatric psychiatry
- Agism, description of, 4558
- Agitated depression, 4559
- Agitation
  - definition of, 4559
  - dementia and, 1196–1197
  - ECT effects on, 3298
  - management of, 1525, 3440–3441, 3441*t*
  - psychiatric emergency due to, 2628–2629, 2628*t*
  - psychomotor, 1635
  - schizophrenia and, 1422–1423
  - suicide behavior risk factor, 3684, 3698
  - treatment
    - benzodiazepine, 2997–2998
    - buspirone, 3011–3012
    - valproate, 3276
- Agitation Subscale of the Neuropsychiatric Inventory, 4226
- Agnosia
  - definition of, 4559
  - dementia and, 1195
  - description of, 452
- Agnosticism, incidence of, 2141
- Agomelatine (Valdoxan)
  - anxiety disorders, 1779, 1780, 1782*t*
    - GAD, 1783
  - BD and generalized anxiety disorder, 1698
  - bipolar depression, 1691
  - chemistry, 3101–3102, 3102*f*
  - depression, 1680
  - dosage/administration, 3106
  - drug interactions, 3106, 3106*t*
  - drug studies, 3104
  - laboratory interference, 3106

- organ/system effects of, 3104–3105
- pharmacological actions, 3102–3103, 3103*t*
- precautions/adverse reactions, 3105–3106
- REM sleep and, 1717
- sexual dysfunction and, 1779*t*
- side effects, 1679*t*
- therapeutic indications, 3105, 3105*t*

- Agonist drugs, activity of, 104*f*
- Agonists, definition of, 4559
- Agoraphobia
  - in aging adults, 3966, 4060
  - CBT and, 1774
  - classification of, 1168, 1173
  - d-cycloserine for, 1743*t*, 1747
  - definition of, 1724, 4559
  - differential diagnoses, 1725
  - DSM-5 on, 1721
  - epidemiology, 1725
    - adults, 1732, 1732*t*
    - genetic, 1738
  - history and comparative nosology, 1724–1725
  - management, 1783
  - panic disorder and, 1783
  - pharmacotherapy and, 1782
  - phenomenology of, 1752
  - treatment, 3122–3123
- Agranulocytosis
  - mirtazapine and, 3110
  - SSRIs and, 1779
- Agraphia, definition of, 4559
- Agreeableness, Five-Factor Model, 935–936
- Ahedonia, animal studies of, 321
- AHRQ. *See* Agency for Healthcare Research and Quality
- Aicardi syndrome, 664
- AIDS. *See* Acquired immunodeficiency syndrome (AIDS); HIV infections
- AIDS demential complex, 4559
- Aiken, John, 4534

- Ailuroophobia, 4559
- AIMS. *See* Abnormal Involuntary Movement Scale
- Akathisia
  - anticholinergics and, 2959, 3073
  - antipsychotic medication side effect, 1525, 1526
  - anxiety *versus,* 2258
  - chronic, 2941
  - definition of, 458–459
  - description of, 4559
  - FGAs and, 2941, 3072
  - haloperidol and, 2375–2376
  - metoclopramide and, 2271
  - neuroleptic-induced acute, 2938, 2941
    - clinical presentation, 2941
    - implications, 2941
    - pathophysiology, 2941
    - treatment options, 2941, 2955
  - in psychiatric disorders, 1148
  - tardive, 2941, 2942–2943
    - clinical presentation, 2942
    - pathophysiology, 2943
    - treatment options/implications, 2943
  - treatment, 2997
  - withdrawal, 2941
- Akinesia
  - definition of, 540
  - manifestation of, 458
  - medication-induced, 2938
  - in PD, 544
  - rigidity and, 458
  - schizophrenia symptom, 1417
- Akinetic mutism, 4559
- Akineton. *See* Biperiden
- Akt kinase
  - activation of, 127–128
  - recruitment of, 126
- Al-Anon, 4559

- *Alabama, Miller v.,* 4440
- Alateen, 4559
- Alcohol. *See also* Al-Anon; Alcohol abuse; Alcohol dependence; Alcohol use disorders; Alcohol withdrawal; Fetal alcohol effects
  - adolescent use of, 3366, 3367, 3372, 3768, 3858
  - aggression threshold and, 2492
  - anxiety disorders and, 1744, 1746
  - benzodiazepines and, 3000
  - bodily effects of, 1267
  - consumption, dietary preferences and, 2547
  - dependence, naloxone for, 169
  - as depressant drug, 1266
  - EEG effects, 253
  - fetal exposure to, 2484
  - GABA receptors and, 117–118
  - head and neck cancers and, 2267
  - intoxication, 2629–2630, 4559
  - MRS study, 263*f*
  - neuropathies related to, 630
  - nonbarbiturate sedative-hypnotic agents and, 2990
  - sexual dysfunction and, 1973
  - spirituality and, 2581–2582
  - usage, HPA axis and, 168
  - withdrawal, HPA axis and, 168
  - withdrawal/negative affect, 580–581
- Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA), 4559
- Alcohol abuse. *See also* Alcohol dependence; Alcohol use
  - after disasters, terrorism, or war, 2567*t*
  - aggression and, 2492
  - by aging adult, 3968–3969, 3973–3974
  - B-family vitamins and, 2548
  - bipolar II disorder and, 1651–1652
  - burn injuries and, 2374–2375
  - dementia and, 1198
  - dietary modification and, 2547
  - dissociative amnesia in dissociative identity disorder and, 1908
  - and HIV infection, 4270
  - Korsakoff syndrome and, 1230–1231

- mood disorders and, 1655–1656, 1669
- pathological dissociation and, 1884
- rates among LGB persons, 2007
- relapse, valproate for, 3274–3275
- suicide and, 1669, 2612
- in transplant populations, 2366–2367
- world-wide, 4520

- Alcohol amnestic disorder, 4559
- Alcohol dependence. *See also* Alcoholism
  - fMRI of, 278
  - genetics and, 2535*t*
  - MRI applications to, 274
  - MRS in, 282
  - STAPP and, 2799
  - STDP in, 2811
  - treatment
    - acamprosate, 3054–3058, 3059
    - buspirone, 3010
    - nalmefene, 3139–3140
    - naltrexone, 3139–3140
    - nefazodone, 3127
- Alcohol deprivation effect, 378
- Alcohol idiosyncratic intoxication, 4559
- Alcohol-induced psychotic disorder, 2890
- Alcohol-related birth defects (ARBD), 637, 637*t*
- Alcohol-related disorders, **1264–1279**
  - alcohol-induced disorders, **1272–1273**
    - anxiety, 1266, 1273
    - bipolar, 1272
    - depressive, 1266, 1272
    - persisting dementias, 1273
    - psychotic disorders, 1273, 2890
    - Wernicke-Korsakoff disorder, 1273
  - blackouts, 1267
- Alcohol-related neurodevelopmental disorder (ARND), 637, 637*t*
- Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS), 821
- Alcohol use disorders

- after TBI, 538
- biological contributors to, 1269–1270, 1269*t*
- bipolar disorder and, 1699
- cancers, 2266
- comorbid conditions, 1266, 1278
- comparative nosology, 1264–1265
- course and prognosis, 1275–1276, 1276*t*
- defined, 1264–1265
- detoxification, 1276–1277
- diagnosis and clinical features, 1270–1273, 1270*t*
- differential diagnoses, 1274–1275
- epidemiology, 1265–1266, 1265*t*
- etiology, 1268–1270
- family and, 1271–1272, 1271*t*, 1276–1277
- genetics of, 1269–1270, 1269*f*t
- history, 1265
- insomnia and, 2087–2088
- intervention for, 1276
- laboratory tests, 1273–1274, 1274*t*
- mood disorders and, 1655–1656
- neuroimaging, 3475
- overview, 1264
- pharmacology of, 1266–1268
- pharmacotherapy, 1278
- prevalence, 1263, 1264, 1265–1266, 1265*t*
- screening for, 1273, 1274*t*
- self-help groups, 1278
- suicide and, 2613
- theories on, 1268–1269
- treatment, 1276–1279
  - anticonvulsants, 2968, 2970
  - disulfiram, 3053
  - lithium, 3093
  - nalmefene, 3140
  - naltrexone, 3140

- Alcohol Use Disorders Identification Test (AUDIT), 1273, 1274*t*
- Alcohol withdrawal

- delirium (*See* Delirium tremens (DTs))
- description of, 4559
- EEG effects, 253
- excitotoxicity in, 274
- treatment
  - acamprosate, 3056
  - α-adrenergic receptor agonists, 2949
  - barbiturates, 2987
  - benzodiazepine, 2997
  - carbamazepine, 3037
  - FGAs, 3071
  - valproate, 3274

- Alcoholic blackout, 4559
- Alcoholic deterioration, 4559
- Alcoholic hallucinosis, 4559
- Alcoholic psychosis, 4559
- Alcoholics Anonymous (AA), 2593
  - description of, 4559
  - treatment recommendations, 956
- Alcoholism. *See also* Alcohol dependence
  - Alcohol Use Disorder and Associated Disabilities Interview Schedule, 821
  - description of, 4559
  - dopamine release in, 415
  - *DRD2* gene and, 2423
  - GABAergic systems and, 83–84
  - *GABRA2* gene in, 2046
  - in indigenous Americans, 801
  - laboratory evaluation, 1017–1018
  - LSD and, 2886, 2891
  - MRI applications to, 274
  - National Epidemiological Survey on Alcohol and Related Conditions, 825–826, 826*t*
  - psilocybin and, 2891
  - serum electrolyte evaluation, 1010
  - treatment of
    - disulfiram in, 63
  - violence and, 2428
- *ALDH4A1* mutation, 663

- Aldosterone production, 2243
- Alertness in patient interview, 967
- Alexander, Franz, 915*f*, 2179, 2815
  - intensive short-term dynamic psychotherapy, 2797
  - personality theory, 915
  - psychoanalysis and, 2640
  - as psychoanalyst, 2643
  - psychopathology theory, 915–916
  - psychosomatic medicine, 915–916
  - specificity hypothesis, 915–916
  - work of, 4559
- Alexia, 4559
- Alexithymia
  - depersonalization/derealization disorder and, 1944
  - description of, 4559
  - pathological dissociation and, 1884
- Algophobia, 4559
- Algorithms, definition of, 4559
- Alien abductions, 2590
- Alienation
  - description of, 4559
  - Horney theories, 920
  - from self, 1902
- Alienists
  - description of, 4559
  - psychiatry and, 4402
- Alkaptonuria, 149
- All-or-none thinking, 2695
- Alladin, Assen, 2723
- Allen Brain Atlas, 163
- Allergic object relationships, 2334
- Allergic reactions
  - drug, 2990
  - jaundice, 4559–4560
  - treatment, 2982
- Allergies
  - hygiene hypothesis and, 192

- hypnosis and, 2729–2730
- in past medical history, 951

- Allocortex description of, 16
- Allodynia, origin of, 2389
- Alloplasty, description of, 4560
- Allopregnenolone (ALLO)
  - fear and anxiety and, 1748
  - function of, 172
  - GABA and, 1747
  - levels in depression, 117
  - PTSD and, 1819
- Allostasis, definition of, 2336
- Allostatic load, positive psychiatry, 4284
- Allostatic stress hypothesis of PTSD, 1822
- Allport, Gordon
  - *Becoming,* 931
  - functionalism and, 2842
  - group relations theory, 4560
  - *Personality,* 934
  - personality theory, 931–932, 934
- Almshouses. *See* Poorhouses
- Alogia
  - description of, 4560
  - schizophrenia and, 1417
- Alopecia. *See also* Trichotillomania (TTM)
  - assessment of, 456
  - description of, 4560
- Alopecia areata
  - epidemiology of, 2344
  - pathophysiology, 2344
  - psychopathology, 2344
  - scalp, 2343*f*
  - treatment, 2344
- Alora. *See* Estradiol
- Alpha (α) errors, 4560
- Alpha-agonists
  - for ADHD, 3789

- child/adolescent, 3788–3789
- Alpha waves, 4560
- α-Linolenic-acid (ALA)
  - omega-6 fatty acids and, 116
- α-Mannosidosis (AM) type II
  - description of, 658–659
  - MRI imaging, 659*f*
- Alpidem (Ananxyl), 3000–3001
- Alprazolam (Xanax)
  - abuse, 1378, 1379, 1387
  - adverse effects of
    - dermatological, 2356*t*
  - agoraphobia, 1783
  - anxiety disorders
    - GAD, 1783
    - panic disorder, 1775
  - BD, 2997
  - BD and panic disorder, 1699
  - benzodiazepine dose equivalent, 2633
  - blood concentration, 2995
  - for borderline personality disorder, 2173
  - chemistry, 2992, 2993*f*
  - distribution, 2993
  - dosage/administration, 3000
  - drug interactions, 3130, 3164
  - drug studies, 2995
  - fear reduction and, 1771
  - generalized anxiety disorder, child/adolescent treatment, 3714
  - history, 2917
  - MDD, 2996
  - metabolism, 2994
  - for premenstrual dysphoric disorder, 171
  - SAD, 1784
  - withdrawal, 2923, 3000
- Alprostadil (Caverject)
  - ED, 3340
  - for sexual dysfunction, 1978–1979

- ALS. *See* Amyotrophic lateral sclerosis (ALS)
- Alternating fist test, 459–460
- Alternating personality, 1866
- Alternating roles, 4560
- Alternative medicine, geriatric depression, 4074
- Alternative splicing, eukaryotic genes, 157*f*
- Alternative treatments, 3586
- Altruism
  - character and, 2138
  - description of, 4560
  - group psychotherapy and, 3761
  - and mental health, 761
  - spirituality and, 2583
- Altshuler, Kenneth, 2827
- Alurate. *See* Aprobarbital
- Alveus, anatomy of, 33, 34*f*
- Alzheimer, Alois, 1201, 4078
- Alzheimer disease (AD), **1202–1210, 4078–4086**
  - age and, 3956, 3963–3064
  - age-related circadian changes in, 214
  - aggressive behavior in, 2495–2496, 2502
  - apathy and, 3229
  - B-family vitamins and, 2548
  - benzodiazepines and, 2998
  - benzodiazepines as risk for, 4342
  - clinical features of, 1205
  - comparative nosology, 1202, 4079, 4079*t*
  - corpus callosum in, 13
  - course and prognosis of, 1207, 4083
  - cranial neuropathies and, 630
  - cytoarchitecture changes in, 5
  - defined, 1202
  - description of, 4560
  - DHEA levels in, 118
  - diagnosis, 1202, 1205
  - diagnosis and clinical features, 4080–4081
  - dietary modification and, 2547

- differential diagnoses, 1198–1199, 1238, 4082–4083
- early-onset, 1203
- early-onset familial, heritability, 2530
- EEGs in, 251
- in elderly patients with cancer, 2284
- epidemiology, 1202, 4079–4080
- etiology, 1202–1203, 4080
- gender issues, 4245
- gene mapping, 311–313
- genetic factors, 4020
- genetics and, 2529*t*
- histamines and, 2981
- history, 1202
- hoarding and, 1794
- laboratory studies for, 1205, 1206*f*, 1207*f*
- late effects of TBI, 530–531
- late-onset, 1203
- MCI leading to, 1224–1228, 1225*t*, 1226*t*, 1227*f*, 1228*f*
- memantine effects, 4145–4146
- microglial cells and, 54
- MRI applications to, 271–272
- neurons in, 5*f*
- pathology and laboratory examination, 4081–4082, 4081*f*, 4082*f*
- pathology of, 460, 1206–1207, 1206*f*–1208*f*, 1209*t*
- PDI nitrosylation in, 107–108
- preclinical stages, 2283
- prevention strategies, 4086
- protective factors, 1204–1205
- and psychiatric symptoms, 4047–4048
- risk factors, 1204
- self-evaluations positively biased, 4372
- semantic degradation, 464
- and sexual dysfunction, 4263
- SRIF and, 175
- stages of, 1202
- susceptibility genes, 312
- treatment, 1207–1210, 1209*t*, 4083–4086

    - antihistamines, 2982
    - cholinesterase inhibitors, 3039–3049, 3040*t*
    - drug development, 2926
    - environmental factors, 4085
    - family intervention, 4085–4086
    - memantine, 3039, 3041, 3048–3049
    - psychopharmacotherapy, 4084–4085, 4084*t*
  - vascular neurocognitive disorder and, 1241
  - *vs.* aging brain, 3991
  - zinc and magnesium levels and, 2545
- Alzheimer Disease Assessment Scale (ADAS-Cog)
  - in aging adults, 4004
  - antidementia drugs, 4139
- Alzheimer type dementia, 1036–1037
- AMA. *See* American Medical Association (AMA)
- *AMACR* gene mutations, 663
- *Amanita muscaria,* 81
- Amantadine (Symmetrel), **2962–2965**
  - absorption, 2962
  - anticholinergics and, 2962–2965
  - for apathy in PD, 548
  - chemistry, 2962, 2962*f*, 3234, 3235*f*
  - in children, 2964
  - for chorea, 555
  - clinical trials, 2962–2963
  - deaths, 2964–2965
  - distribution, 2962
  - dosage/administration, 2965, 3241, 3241*t*
  - drug interactions, 2965, 3240
  - in elderly, 2962
  - elimination, 2962
  - EPS, 2963, 2965, 3238
  - history, 3233–3234
  - laboratory interferences, 2965
  - metabolism, 2962
  - NMS, 3238
  - organ/system effects, 2963, 3239

- pharmacodynamics, 2962, 3235–3236
- pharmacological actions, 2962, 3235–3236
- precautions/adverse reactions, 2964, 2964*t*, 3239, 3240*t*
- therapeutic indications, 2962–2964, 3239
- traumatic brain injury, 1244

- Ambidextrous patients, 455
- Ambien. *See* Zolpidem
- Ambiguous Intentions and Hostility Questionnaire (AIHQ), 4370
- Ambivalence
  - description of, 4560
  - in schizophrenia, 1407, 1553
- Ambulatory infusion pumps, 2293*t*
- Ambulatory schizophrenia, 4560
- Amenorrhea, 2507–2508
- Amentia, 4560
- American Academy of Child and Adolescent Psychiatry
  - ethical guidelines, 4441
  - Psychiatry Residency In-Training Examination and, 4428
- American Academy of Child and Adolescent Psychiatry (AACAP)
  - description of, 4560
  - practice guidelines, 971, 3688, 3702
- American Academy of Clinical Neuropsychologists, 1856
- American Academy of Pediatrics (AAP)
  - Committee on Child Abuse and Neglect, 1847
  - pediatric genetic testing recommendations by, 3456
- American Association for Emergency Psychiatry, 2623
- American Association for Retired Persons (AARP), 4275
- American Association of Chairmen of Departments of Psychiatry (AACDP), 4560
- American Association of Directors of Psychiatric Residency Training, Inc (AADPRT)
  - description of, 4560
  - Psychiatry Residency In-Training Examination and, 4428
- American Association on Intellectual and Developmental Disabilities (AAIDD, formerly AAMR), 3493
- American Board of Examiners in Clinical Social Work, 4429
- American Board of Medical Specialties (ABMS)
  - establishment of, 4405–4406
  - member boards, 4404*t*

- psychiatric education and, 4402
- American Board of Professional Psychology (ABPP), 4428–4429
- American Board of Psychiatry and Neurology (ABPN). *See also* Psychiatric education
  - Board of Directors, 4404*t*
  - description of, 4560
  - establishment of, 4404
  - evaluating success of examinees, 4430
  - examinations
    - candidate performance, 4405*t*
    - clinical skills, 4416
    - cognitive, 4415–4416
    - grading process, 4426*t*
    - oral, 4416
    - success and failure on, 4421–4424
    - test questions developed by, 4418–4419
    - written, 4415*t*
  - founding of, 4537
  - psychiatric training and, 4402, 4404–4405
  - recertification, 4427–4428
  - subspecialties, 4405*t*
  - 360-degree evaluations, 4429
- American College of Medical Genetics, 3456
- American College of Physicians (ACP), 4560
- American College of Psychiatrists (ACP), 4428, 4560
- American Foundation for Suicide Prevention (AFSP), 2619
- American Law Institute, formulation of insanity, 4560
- American Medical Association (AMA)
  - CAM modalities conferences, 2542
  - description of, 4560
  - ethical guidelines, 4441
  - on forensic use of hypnosis, 2731
  - guidelines for correctional psychiatry, 4450
  - hypnosis and, 2706
  - practice guidelines, 970–971
- American Medical Directors Association
  - psychiatrist role in long-term care, 4222
- American Neurological Association (ANA), 4560

- American Psychiatric Association (APA)
  - CAM modalities conferences, 2542
  - description of, 4560
  - ethical guidelines, 4441
  - evolution of, 4291–4292
  - guidelines for correctional psychiatry, 4450
  - hypnosis and, 2706
  - on intellectual disability definition, 3493
  - practice guidelines, 971–974
    - choice of topic, 972
    - committees and work groups, 972
    - dissemination and implementation, 974
    - format, 973
    - history, 971–972
    - identification of evidence, 972–973, 972*t*
    - recommendations, 973
    - review and approval process, 973–974
    - updates, 974
  - on psychiatric classification, 1152, 1155
  - website, 974
- American Psychoanalytic Association (APA), 2661, 2675, 4560
- American Psychological Association (APA)
  - on behavior therapy, 2704
  - on brainwashing, 1947
  - description of, 4560
  - on hypnosis, 2709
  - on short-term dynamic psychotherapies, 2796
- American Psychological Foundation, 2619
- American Society of Clinical Hypnosis, 2732, 2733
- American Sociological Association, 1947
- Americans with Disabilities Act (ADA), 3493
  - correctional psychiatry, 4452
  - for intellectual disabilities, 3518–3519
  - mental illness accommodation, 4363
  - mental illness as disability, 4337
  - work issues and, 2277
- Amine neurotransmitters, biogenic, **61–75**

- Amines, description of, 4560
- α-Amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid (AMPA), 76, 1747
- Amino acid neurotransmitters, **76–84**
- Amino acids (AAs), beneficial effects, 2549–2550
- Amino-levulinic acid (ALA) accumulation, 660
- Amiodarone (Cordarone)
  - hypothyroidism and, 2198
  - psychiatric side effects of, 2198*t*
- Amisulpiride, EUFEST, 3070
- Amitriptyline (Elavil)
  - chemistry, 3254, 3254*f*
  - for chronic idiopathic pruritus, 2348–2349
  - classification, 2918
  - depression, 1681
  - dosage/administration, 3264, 3265, 3265*t*
  - dosages, 573*t*
  - drug interactions, 3264
  - drug studies, 3259
  - for fibromyalgia, 635
  - for migraine treatment, 624
  - for neuropathies, 629
  - organ/system effects of, 3260
  - pain syndromes, 1841
  - pharmacological actions, 3255, 3256, 3257, 3258
  - precautions/adverse reactions, 3263
  - side effects, 573*t*
  - targets for, 69
  - therapeutic indications, 3261

- Amlodipine (Norvasc)
  - chemistry, 3016
  - dosage/administration, 3021
  - pharmacokinetics, 3017
- Ammon Horn, 45
- Ammonia, delirium pathophysiology and, 1181
- Amnesia(s). *See also* Dissociative amnesia; Hyperamnesia; Posttraumatic amnesia; Transient global amnesia
  - acute, evaluation of, 1936*t*
  - alcohol-induced, 1267, 1271, 1273
  - description of, 4560
  - dissociative amnesia *versus,* 1931, 1935
  - dissociative disorders and, 1879
  - false memory syndrome and, 1872
  - history, 1866, 1868, 1869
  - hypnosis, 2711
  - inter-identity, 1910–1911
  - malingering of, 2406
  - media and public fascination with, 1873–1874
  - and memory, 730–734
    - specialized functions, 731–733, 732*f*, 733*f*, 733*t*
  - memory and, 463
  - neurological, 741
  - posthypnotic, 2708
  - prevalence, traumatized populations, 1886–1887
  - psychogenic, 741–742
  - retrograde, 2998
    - memory and, 733–734, 735*f*
- Amnestic aphasia, 4560
- Amnestic confabulatory type, persistent, 1273
- Amnestic disorder(s), **1229–1232**
  - assessment, 1231
  - classification of, 1222
  - course, 1231
  - defined, 1222, 1229
  - description of, 4560
  - differential diagnoses, 1231
  - epidemiology, 1229

- etiology, 1230, 1230*t*
- Korsakoff's syndrome and, 1230–1231
- prognosis, 1231
- symptoms, 1230–1231
- treatment, 1231

- Amobarbital (Amytal), 2986, 2987
  - interviews
    - informed consent for, 1916
  - narcosynthesis, 1940
- *Amoeba dubia* (amoeba), 148
- Amok, description of, 4561
- Amotivational syndrome, 1308
- Amoxapine (Asendin)
  - chemistry, 3254, 3254*f*
  - drug studies, 3259
  - organ/system effects of, 3260
  - pharmacological actions, 3255
  - precautions/adverse reactions, 3263
- Amoxicillin (Amoxil), 591
- AMPA. *See* α-amino-3-hydroxy-5-methyl-4-isoxazole-propanoic acid (AMPA)
- Amphetamine(s), **3219–3231**. *See also* Stimulant-related disorders
  - ADHD, 3219, 3223, 3227–3229
  - ADHD and, 3592–3594, 3592*t*, 3788
  - antidepressants combined with, 3324
  - for apathy, 548
  - for cocaine use disorders, 1290
  - description of, 4561
  - dissociative amnesia, 1940
  - dosage/administration, 3220*t*–3221*t*
  - drug interactions, 3227*t*
  - history, 3219
  - intoxication, 1713
  - mechanisms of action, 3223–3224
  - obsessive-compulsive disorder and, 3789
  - in PD, 548
  - pharmacological actions, 3224–3227, 3225*t*–3226*t*, 3227*t*
  - precautions/adverse reactions, 3229–3231, 3229*t*

- psychiatric emergency due to, 2633
- salts, BD with ADHD, 1699
- as stimulant use disorder, 1280, 1281, 1284, 1290
- therapeutic indications, 3227–3229
- transporters, 69
- use as antidepressant, 4554
- for weight loss, 2220

- Amphotericin for cryptococcus meningitis, 586
- AmpliChip CYP450 test, 3456
- Amplifications, description of, 155
- Amusia, description of, 452
- Amygdala
  - affective dysregulation and, 1610
  - aggression and, 2476, 2477
  - anatomy of, 8, 10, 35
  - ASD and, 56
  - autism spectrum disorder and, 3576
  - biologically relevant learning and, 1757
  - connectivity
    - emotion regulation and, 424–425
    - patterns of, 425*f*
  - description of, 4561
  - dissociative amnesia and, 1890
  - dissociative identity disorder and, 1882, 1894
  - dopamine and, 1745
  - dysfunction, depression and, 1709
  - embryonic origins of, 8
  - emotional processing and, 1711, 1712, 1712*f*
  - fear circuitry and, 1753–1755, 1753*f*, 1755*f*
  - fear conditioning, PTSD and, 1818
  - gambling disorder and, 1806
  - helping activities and, 2416
  - lesions of, 465
  - memory storage, 729, 732
  - mood disorders and, 1718*f*
  - neuroimaging
    - drives and instincts and, 2815

        - fear, 1759–1760
        - GAD, 1764
        - major depressive disorder, 3466
        - maturation and, 1765
        - OCD, 3470
        - social anxiety disorder, 1763
        - specific phobias, 1762–1763
        - threat detection and emotion regulation, 1819
    - NPY levels and dissociation in, 1891
    - output pathways, 36*f*
    - panic attacks and, 1761
    - psychotherapy and, 2813
    - PTSD and, 1819, 1820–1821, 1821*f*
    - schizophrenia and, 1482, 1484–1486
    - in social phobia, 702
    - stimulant-related disorders and, 1283
    - stress on hippocampus and, 1744
    - trauma response and, 2569
- Amyl, butyl, and isobutyl nitrite-related disorders, 1341
- Amyloid-beta peptides, 4150, 4150*t*
- Amyloid protein
    - altered processing of
        - in NPC, 665
    - Alzheimer disease and, 1203–1204, 1203*f*
- Amyotrophic lateral sclerosis (ALS)
    - description of, 4561
    - EAAT2 loss in, 79
    - frontotemporal spectrum disorders and, 628*t*
    - neuropsychiatric aspects, 627–668
    - pathology of, 628
- Amytal. *See* Amobarbital
- Amytal interview, description of, 4561
- ANA. *See* American Neurological Association (ANA)
- Anabolic-androgenic steroid-related disorders, **1390–1404**
    - adverse medical effects, 1395–1396
    - comparative nosology, 1392
    - course and prognosis, 1399

- dependence, 1396–1399, 1398*f*
- endocrine tests, 1396
- epidemiology, 1392–1393
- etiology, 1396, 1397*f*
- history, 1390–1392, 1391*f*
- laboratory findings, 1396
- medical uses, 1394–1395
- mood disorders and, 1399–1402
- not otherwise specified, 1402–1403
- organ and system effects, 1394
- overview, 1390, 1390*t*
- pharmacology, 1393–1396, 1394*f*
- treatment, 1399

- Anabolic steroid(s)
  - aggression and, 2493
  - aggression threshold and, 2493
  - description of, 170
- Anaclitic, definition of, 4561
- Anaclitic depression, 1662, 3663, 3677
- Anaclitic therapy, 4561
- Anafranil. *See* Clomipramine
- Anakastic personality, 4561
- Anal character, definition of, 4561
- Anal phase, definition of, 4561
- Analgesia
  - cannabis use for, 1311
  - definition of, 4561
  - hypnotic, 2708, 2711, 2713
    - burn care, 2724
  - positive expectations as, 2847
- Analgesics
  - narcotic, 2293–2294
  - nonnarcotic, 2293
- Analog situations, for observation, 2687
- Analophilic orientation, 1983, 2024
- Analysand, definition of, 4561
- Analysis in depth. *See* Psychoanalysis

- *Analysis of the Self* (Kohut), 4543
- Analysis of variance (ANOVA), 829, 834, 4561
- Analytic psychology, 4561
- Anamnesis, definition of, 4561
- Anancasm, definition of, 4561
- Anandamide
  - function of, 110
  - neuroprotection by, 115
  - synthesis of, 110
- Ananxyl. *See* Alpidem
- *Anaplasma* infections, 593
- *The Anatomy of Melancholy* (Burton), 4534
- Anchoring proteins, signaling and, 128–129
- Ancillary care, 4561
- And learning impairments
  - assessment of
    - achievement testing for, 3425–3426
- Anderson, Carol, 2754
- Anderson, Michael C., 1893
- Androgen insensitivity syndrome (AIS), 1992
- Androgen receptors
  - brain localization of, 2505
  - defects in, 2506
- Androgen replacement therapy, 172
- Androgen(s). *See also* Antiandrogens; Complete androgen insensitivity syndrome (CAIS)
  - atypical exposure to
    - mechanisms of, 1994–1995
  - brain metabolism of, 2418
  - conversion to estradiol, 2505
  - definition of, 4561
  - libido and, 1956
  - midlife depression and, 3336
  - sex drive and, 1980
  - supplementation of, 2034
  - tic development, 3638, 3652
  - virilization and, 1956
- Androgeny, 4561. *See also* Bisexuality

- Androphilic orientation, 1983, 2024, 2030
- Androphobia, 4561
- Androstenedione, 3336
- Anencephaly, causes of, 40
- Anergia, definition of, 4561
- Anergic schizophrenic, 4561
- Anesthesia
  - definition of, 4561
  - for ECT, 2986, 2987, 3291–3293, 4156–4157
  - hypnosis, 2711
  - QEEG-based depth, 1040
  - sedative-hypnotics for, 2990
- Anesthetic(s)
  - dissociative, 1312, 1314
  - EEG effects, 252
  - as GABA-A agonists, 1746
- "Angel dust," 1325, 4561. *See also* Phencyclidine (PCP)
- Angelman syndrome
  - epigenetic imprinting and, 161
- Angels, claims of being touched by, 2590–2591
- Anger
  - aggression and, 2499
  - CAD and, 2188
  - CREB and, 2483
  - depression and, 1661
  - management of, 2499
  - ODD symptom, 3610
  - therapeutic approaches to
    - in burn patients, 2381
- Anger management therapy, 2499
- Angina pectoris, 2186
- Angioedema
  - classification of, 2346
  - urticaria and, 2345–2346
- Angiotensin-converting enzyme (ACE) inhibitors, 2198*t*, 3098
- Angus, Lynne, 2831
- Anhedonia, 1417, 1634, 4561

- Anima, definition of, 4561
- Animal-facilitated psychotherapy, 2556
- Animal models, **319–325**
  - of aggression, 2472
  - of anxiety-like behavior, 321
  - benzodiazepine studies, 1376–1377
  - brain development studies, 2418–2420
  - chronic stress models, 322
  - of fear-like behavior, 320–321
  - importance of, 401–402
  - modes of depression
    - stress and, 1712–1713
  - neuropsychopharmacology, 1
  - of pain perception, 326–327
  - of panic-like behaviors, 320
  - of social anxiety-like behavior, 321
  - utility of, 401–402
  - of violence, 2425
- Animia, 4560
- ANK3 (giant-ankyrin), 1627–1628
- Ankyrin, 1627–1628
- Anna O. case, 852, 4561
  - conversion disorder and, 1839
  - psychoanalysis and, 2639
- Annapolis Coalition on the Behavioral Healthcare Workforce, 4306, 4323
- Annihilation, feelings of, 2140
- Annihilation anxiety, 2874
- Anniversary reactions, 1632, 4561
- Anomalous experiences, 2589–2591
  - alien abductions, 2590
  - near-death, 2589–2590
  - touched by angels, 2590–2591
- Anomia, description of, 4561
- Anomic, definition of, 4561
- Anomic aphasia
  - description of, 4561
  - features of, 462*t*

- Anorectics
  - description of, 4561
  - insomnia and, 2087
  - side effects, 2087
- Anorexia
  - sensation and perception, 690
  - terminal, 2296
- Anorexia nervosa (AN)
  - animal models of, 322–323
  - case history, 2082
  - classification of, 1174
  - cognitive therapy, 2766
  - course/prognosis in, 2077–2078
  - depressive syndrome and, 1638
  - description of, 2065, 4561
  - diagnosis of, 2072
  - diagnostic criteria, 2065*t*
  - differential diagnoses, 2075–2076
  - EEG in, 248*f*
  - epidemiology of, 2069–2070
  - etiology of
    - developmental factors, 2070
    - environmental factors, 2071
    - genetic factors, 2070
    - psychological factors, 2070–2071
    - social factors, 2071
  - family therapy, 3783–3785
  - genetic mapping in, 3451
  - genetics and, 2529*t*, 2535*t*
  - GH abnormalities and, 174
  - historical literature on, 2067
  - infantile, **3623–3625, 3637–3639**
  - laboratory exams, 2074–2075
  - mortality rates, 2073
  - neuroimaging, child/adolescent, 3475
  - nosology, 2068
  - pathology, 2074–2075

- sham eating in, 359
- STDP in, 2811
- subtypes, 2072
- temperament dimensions of, 3487
- treatment of, 2079–2080

- Anorexia nervosa by proxy, 1849
- Anorgasmia, 1968, 4561
- Anosmia
  - in Alzheimer disease, 630
  - in Parkinson disease, 630
- Anosognosia
  - after brain injury, 475
  - after stroke, 478
  - definition of, 4372, 4390, 4392
  - description of, 333, 4561
- Anosomnia, 4561
- ANOVA. *See* Analysis of variance (ANOVA)
- ANS. *See* Autonomic nervous system
- Antabuse. *See* Disulfiram
- Antagonists, definition of, 4561
- Anterior cingulate cortex (ACC)
  - dysfunction, depression and, 1709
  - emotional processing and, 1711–1712
  - episodic memory and, 2139
  - GAD and, 1764
  - hoarding disorder and, 1793
  - hypnosis, 2712, 2713
  - memory suppression and, 1893
  - mood disorders and, 1718, 1718*f*
  - neuroimaging
    - drives and instincts and, 2815
  - nociceptive signals, 328
  - OCD and, 1786, 2813
  - PTSD and, 1821, 1821*f*
  - salience network, 21
  - social rejection and, 192
  - surface landmarks, 7*f*

- Anterior cingulate gyrus (ACG), 766, 2477–2478
- Anterior commissure
  - anatomy of, 12
  - in schizophrenia, 13
- Anterograde amnesia, 1267, 4561
- Anthony, E. James, 2737
- Anthony, William, 1549
- Anthrax
  - in mail (2001), 2574
  - war or terrorism and, 2572*t*
- Anthropologists
  - on homosexuality, 1996
  - medical, 2446–2447
- Anthropology
  - definition of, 4561
  - and neurocentrism, 786
- *Anthropology from a Pragmatic Point of View* (Kant), 928
- Anti-inflammatories, interactions, 3098, 3145
- Anti-müllerian hormone (AMH)
  - fetal secretion of, 2505
- Antiandrogens
  - for paraphilias, 2020
  - risks of treatment with, 2020
  - sex offenders
    - recidivism and, 2503
    - therapies for, 2503
  - for sexual dysfunction, 1980
- Antianxiety drugs, 4128–4132, 4128*t*. *See also* Anxiolytic drugs
  - classification, 2918
- Antibiotics
  - for tic disorders, 3639, 3653
- Anticholinergic, definition of, 4561
- Anticholinergic deliriants, 1317, 1320
- Anticholinergic effects, 4561
- Anticholinergic(s), **2956–2965**
  - absorption, 2957
  - abuse of, 2960

- administration, 2960, 2961
- adverse reactions, 2959–2960, 2960*t*, 2961
- amantadine and, 2962–2965
- antimuscarinic, 2956
- antinicotinic, 2956
- atropine, 2956, 2956*f*
- chemistry, 2956–2957, 2957*f*, 2960, 2960*f*
- clinical studies, 2958, 2961–2962
- distribution, 2957
- dosage, 2960, 2960*t*, 2961
- drug interactions, 2960, 2960*t*
- EPS, 2959, 3073
- intoxication, 1189
- laboratory interference, 2960
- long-term use of, 2959
- organ/system effects, 2958–2959
- pharmacodynamics, 2957–2958
- pharmacological actions, 2957–2958, 2960–2961
- precautions, 2959–2960, 2961
- preparations, 2960*t*
- scopolamine, 2960–2965
- sexual dysfunction and, 1973
- therapeutic indications, 2959, 2961

- Anticipation, definition of, 761, 4561
- Anticipatory immune suppression (AIS), 2271–2272
- Anticonvulsants, **2966–2979**
  - for aggressive behavior, 2111, 2172
    - in autism, 2500
    - mentally retarded children, 2500
  - anxiety disorders, 1781, 1782*t*
  - behavioral effects of, 519*t*
  - bipolar depression, 1690, 1694–1695
  - bipolar disorder, 3692, 3706, 3791, 4076–4077
  - in brain tumor patients, 496
  - chemistry, 2966, 2966*f*, 2967, 2969, 2972, 2974–2977
  - contraindications, 2521
  - dementia treatment, 1201

- description of, 4562
- dosage/administration, 2969, 2971–2973, 2975–2976, 2978
- drug interactions, 520, 2969, 2971, 2973, 2975–2976, 2978, 3098
  - blood level effects, 521*t*
- drug studies, 2967–2968, 2970, 2972–2977
- ECT interactions with, 3295
- history, 2966–2967, 2969, 2972–2977
- intellectual disability treatment, 3514–3515
- as mood stabilizers, 1612
- newest, 2978–2979
- off-label use of, 2969
- in pain syndromes, 2396
- pharmacodynamics, 2967, 2970, 2972, 2975, 2977
- pharmacokinetics, 2967, 2969, 2972, 2974, 2975, 2977
- pharmacological actions, 2967, 2969–2970, 2972, 2974–2975, 2977
- precautions/adverse reactions, 2968–2971, 2973–2976
- prophylactic, 528
- sexual dysfunction and, 1779*t*
- side effects, 2087
- suicidality and, 2616, 2966, 2975, 3086
- therapeutic indications, 2968, 2970, 2973, 2974, 2976, 2977
- in transplant patients, 2372*t*
- uses of selected, 2978*t*
- weight gain and, 2215*t*

- Anticraving medications, considerations, 2372*t*
- Antidementia drugs, **4139–4151**
  - adjunctive and combination therapy, 4149
  - amyloid-beta peptides, 4150. 4136*t*
  - FDA regulations, 4139–4140
  - future directions, 4150
  - mild cognitive impairment, 4149
  - Parkinson disease dementia, 4149–4150
  - treatment using, 4140, 4141*t*
    - axona, 4148
    - cholinergic hypothesis, 4141
    - cholinesterase inhibitors, 4128*t*
      - adverse effects, 4144

        - clinical application, 4144
        - contraindications, 4144
        - donepezil, 4143
        - duration of treatment, 4144–4145
        - galantamine, 4143–4144
        - rivastigmine, 4143
        - tacrine, 4142
    - docosahexaenoic acid, 4148
    - ergoloid mesylates, 4146
    - ginkgo biloba, 4147
    - HMG-CoA reductase inhibitors, 4147
    - huperzine A, 4146
    - memantine, 4145–4146
    - monoamine oxidase type B inhibitors, 4146
    - nootropics, 4146
    - precursor loading and cholinergic stimulation, 4141–4142
    - souvenaid, 4148
    - vitamin B and homocysteine reduction, 4147
    - vitamin E, 4147
  - unproven medications
    - anti-inflammatories, 4148
    - estrogens, 4148–4149
  - for vascular dementia, 4149
- Antidepressant-induced rapid cycling, 1642
- Antidepressant(s), 4540. *See also specific Antidepressants*
  - for aggressive behavior, 2111
    - schizophrenia, 2501
    - TBI, 2502
  - amphetamine-type stimulant disorders, 1290
  - for anorexia nervosa, 2080
  - anticonvulsant-augmented, 1781
  - antidepressant discontinuation syndrome, 2943
  - antipsychotics combined with, 3324, 3327
  - anxiety disorders, 1779–1780, 1779*t*
  - benzodiazepine withdrawal, 1388
  - benzodiazepines and, 2996
  - biological or chemical terrorist attacks and, 2572

- bipolar depression, 1691, 1695
- bipolar I disorder, 1696*t*, 1697
- bipolar II depression, 1697
- black box warning, 2918–2919
- in brain tumor patients, 499
- for bulimia nervosa, 2081
- for cancer patients, 2253
- cannabis dependence, 1310–1311
- CBASP and, 2700
- for chronic idiopathic pruritus, 2348–2349
- classification, 2918
- combination psychotherapy, 3323–3325
- comparative studies, 1678–1679
- dementia with Lewy bodies treatment, 1218–1219
- depression-centric perspective and, 1631
- description of, 4562
- drug interactions, 3098, 3145
  - with HAART medications, 572
- ECT interactions with, 3295
- ED and, 3340–3341
- for epilepsy patients, 520*t*
- folate and, 2547
- in functional GI disorders, 2206
- gambling disorder and, 1808
- in HD, 555–556
- in HIV infections
  - dosages, 573*t*
  - side effects, 573*t*
- for IBS, 2204
- immune system and, 201–202
- insomnia and, 2087
- for intellectual disability, 3514
- IPT and, 1701
- lithium combined with, 3092, 3098, 3323, 3327
- major depressive disorder, 3790
- for MDD, 499
- mood disorders and, 1612

- mood stabilizers combined with, 3327
  - for Alzheimer disease, 4047
  - for bipolar disorders in later life, 4127
  - bupropion, 4125
  - depression in later life, 4123–4124
  - for depression with medical illness, 4126
  - dual-acting antidepressants, 4124–4125
  - for encephalopathies, 4047
  - monoamine oxidase inhibitors, 4125
  - patient management, 4126–4127
  - for psychotic depression in later life, 4125–4126
  - selective serotonin reuptake inhibitors (SRRI), 4124
  - stimulants, 4125
  - tricyclic antidepressants, 4125
- obsessive-compulsive disorder treatment, 3701
- in pain syndromes, 629, 1841, 2395
- panic disorder, 1724
- pharmacogenomic testing for, 3356, 3455
- and physical health monitoring, 1025
- poststroke mortality and, 485*f*
- REM sleep and, 1717
- REM-sleep-suppressing, 347
- SAMe as, 2549
- scopolamine and efficacy of, 2961–2962
- seasonal depressions and, 1651
- separation anxiety treatment, 3710
- sex steroids combined with, 3324–3325
- for sexual dysfunction, 1979
- sexual dysfunction and, 1972
- for “shingles,” 632
- side effects of, 532–533, 2087
- sleep-related movement disorders, 2104
- stimulant-like agent combined with, 3324
- suicidal behavior associated with, 1672–1673
- suicidality and, 1340, 2616, 2918–2919, 3378, 3718
- for terminally ill children, 2309
- thyroid hormone/supplement combined with, 2923, 3244–3245, 3245*t*, 3323

- in transplant patients, 2372*t*
- treatment phases, 1677*t*
- valproate and, 3279
- weight gain and, 2215*t*

- Antidiuretic hormone (ADH, arginine vasopressin, AVPJ). *See also* Arginine vasopressin (AVP)
  - deficiency
    - in TBI, 525
  - function of, 2239
  - laboratory evaluation, 1013–1014
  - production of, 2239
- Antiemetic drugs
  - for cancer patients, 2256, 2271–2272
  - for nausea, 2293
- Antiepileptic drugs (AEDs). *See* Anticonvulsants
- Antiestrogens
  - in PCOS treatment, 2249
  - for sexual dysfunction, 1980
- Antigen-presenting cells (APCs), 181
- Antihistamine(s), **2979–2984**
  - absorption, 2980
  - adverse events, 2982–2983
  - for anxiety disorders, 4131
  - chemistry, 2979, 2981*f*
  - distribution, 2980
  - dosage administration, 2983–2984. 2973*t*
  - drug interactions, 2983
  - drug studies, 2981–2982
  - elimination, 2981
  - history, 2979
  - laboratory interferences, 2983
  - metabolism, 2981
  - organ/system effects, 2982
  - pharmacodynamics, 2980–2981, 2980*t*
  - pharmacokinetics, 2980–2981, 2980*t*
  - pharmacological actions, 2979–2981, 2980*t*
  - precautions, 2982–2983
  - sexual dysfunction and, 1973

- therapeutic indications, 2982
- Antihomosexual attitudes, 1999–2000
- Antihypertensive drugs, side effects, 2087
- Antikeratomas, 656
- Antimanic drugs, 4562
- Antinuclear antibody tests (ANA), 467
- Antiparkinsonism drugs, 1218–1219, 4562
- Antiphospholipid (aPL) antibodies, 472
- Antipsychotics/neuroleptics. *See also* Chlorpromazine (Thorazine); First-generation antipsychotics; Second-generation antipsychotics
  - for acute episodes, 1521–1522*t*, 1523–1525
  - agitation management, 1525
  - alternate uses for, 2618
  - antidepressants combined with, 3324, 3326
  - assessment, 1523
  - atypical
    - for aggression in BPD, 2501
    - for aggression in conduct disorder, 2501
    - for aggression in MR children, 2500
    - for aggression in schizophrenia, 2501
    - for anxiety disorders, 1781, 1782*t*
    - for bipolar depression, 1690–1691, 1695
    - for bipolar disorders, 1683*t*, 1686–1689, 1686*t*, 1689*t*
    - for bipolar I disorder, 1696
    - for bipolar II disorder, 1698
    - for depression, 1680
    - for gambling disorder, 1808
    - sexual dysfunction and, 1779*t*
  - availability of, 4525
  - BDD, 1792
  - biological or chemical terrorist attacks and, 2572
  - bipolar depression, 1695
  - bipolar disorders
    - emergency management, 1688
    - treatment selection, 1688–1689
  - black box warning, 2919
  - chlorpromazine equivalents, 520*t*
  - classification, 2918

- cognitive impairment management with, 1528
- combination pharmacotherapy, 3324, 3326–3328
- combining two, 3327
- comparative effectiveness, 1523
- cost of, 4522, 4524
- delirium prevention and management, 1180–1181, 1188
- for delusional parasitosis, 2347
- dementia treatment, 1201, 1218–1219, 1238
- depersonalization/derealization disorder, 1945
- description of, 4562
- diabetes risk and, 2247
- dissociative identity disorder, 1927*t*, 1928
- dosage, 4135, 4135*t*
- dosage selection, 1521–1522*t*, 1524–1525
- drug-drug interactions, 2257
- drug interactions, 3098
- ECT and, 3295
- ED and, 3341
- EEG effects, 253
- effectiveness, 1520, 1523
- in elderly patients, 4342
- endocrine effects of, 3795
- for first-episode psychosis, 1520
- in functional GI disorders, 2206–2207
- future directions, 1530–1531
- health monitoring with use of, 1528
- history, 1520
- for IBS, 2204
- illness anxiety disorder, 1838
- increasing use of, 4329
- indications, 1523
- for intellectual disability, 3515
- involuntary use of, 4332
- lithium combined with, 3093, 3098, 3327
- long-acting injectable, 1523–1524, 1529, 1530
- maintenance therapy, 1529–1530
- managing patients on, 4530

- metabolic syndrome comorbid with, 4343
- mood disorders, 1612
- mood stabilizers combined with, 3326–3328
- mortality rates with, 4342
- negative symptom management with, 1528
- neuroleptics, 2936
  - EEG effects, 253
  - side effects, treatment, 2997
- obsessive-compulsive disorder and, 3701–3702
- in pain syndromes, 2396
- Parkinson disease treatment, 1220
- patient education, 1524
- pharmacologic actions, 4132–4138
- poor responder strategies, 1528–1529
- precautions and adverse reactions
  - anticholinergic side effects, 4137
  - cerebrovascular events, 4136, 4138
  - hypotension, 4137–4138
  - metabolic effects, 4137
  - mortality, 4136
  - sedation, 4137
  - tardive dyskinesia, 4137
  - weight gain, 4136–4137
- psychosocial treatment integration with, 1530
- research in, 4351–4353
- route of administration, 1521–1522*t*, 1523, 1529–1530
- schizophrenia
  - neuroimaging studies of, 1497–1499, 1497*f*, 1498*f*
- schizophrenia with suicidality, 2616
- second-generation, 4351
- selection, 1523
- sexual dysfunction and, 1972
- side effects of, 67, 574, 1525–1528
  - of clozapine, 1527
  - dermatological, 2356*t*
  - extrapyramidal, 1526
  - metabolic, 1527

        - neuroleptic malignant syndrome, 2381
        - other, 1527
        - tardive dyskinesia and other syndromes, 1526–1527, 1529
        - weight gain, 2237
    - stereotypic movement disorders, 3672, 3686
    - suicide behavior management with, 1528
    - therapeutic indications
        - bipolar disorder, 4134
        - delirium, 4134
        - delusional disorder, 4133
        - dementia with psychosis, 4133–4134
        - depression, 4134
        - Parkinson disease with psychosis, 4134
        - PD, 2169
        - schizophrenia
            - early-onset, 4132–4133
            - late-onset, 4133
    - tic disorders, 3645–3646, 3647*t*, 3659–3660, 3661*t*
    - treatment principles
        - compliance, 4135
        - dosages, 4135
        - drug interactions, 4136
        - drug selection, 4135
        - pretreatment evaluation, 4135
    - for trichotillomania, 2353
    - types, 1521–1522*t*
    - valproate and, 3278
    - weight gain and, 2215, 2215*t*
- Antipyretic analgesics, 2394
- Antiretroviral therapy (ART)
    - adherence to, 577–578
    - in children and adolescents, 645–646
    - issues related to, 648
    - psychosis induced with, 4343
    - side effects of, 646
- Antisocial behavior
    - adult, **2413–2430**

- after disasters, terrorism, or war, 2567
- description of, 4562
- fetal nutrition and, 2484

- Antisocial characteristics, CBT for, 4392–4393
- Antisocial personality disorder (ASPD), 2153–2154, 4106
  - aggression and, 2486, 2494
  - clinical criteria, 2153
  - comorbidity, 2153
  - complications, 2153
  - course of, 2154
  - description of, 4562
  - diagnosis of, 2437
  - differential diagnoses, 1270, 1274, 2154
  - epidemiology of, 2153–2154
  - factitious disorder imposed on self and, 1851, 1864
  - familial patterns, 2154
  - gambling and, 1804, 1805
  - gender differences in, 2475
  - genetics and, 2154, 2529*t*
  - impairment, 2153
  - malingering and, 2408
  - MBT, 2904
  - neuroimaging of, 4482, 4487–4488
  - premature births and, 643
  - psychopathy and, 2490–2491
  - sex ratio, 2153
  - spectrum of, 2155
- Antisocial Process Screening Device (APSD), 3402
- Antithyroid antibodies, 680–681
- Antiviral regimens, adherence to, 564
- Anton syndrome, 4562
- Anxiety. *See also* Generalized anxiety disorder; Hamilton Anxiety Rating Scale (HARS); Social anxiety disorder; State-Trait Anxiety Inventory for Children (STAIC)
  - after stroke, 475, 476–477
    - course/prognosis in, 485–486, 486*f*
    - diagnosis of, 481
    - survival and, 487–488

    - treatment of, 487–488
  - anatomical correlates of, 1753–1757
  - annihilation, 2874
  - anticipatory, panic attacks and, 1761
  - assessment, 1721
  - brain tumor patients and, 500–501
    - course/prognosis, 500–501
    - diagnosis of, 500
    - differential diagnoses, 500
    - epidemiology of, 500
    - etiology of, 500
    - pathology, 500
    - treatment of, 501
  - breast cancer patients and, 2269
  - burn injuries and
    - children, 2385
    - intermediate phase, 2379
  - CAM treatments and, 2542–2543
  - cancer patients and, 2253
  - cardiac arrhythmias and, 2190
  - cardiac patients and, 2197
  - cardiac symptoms of, 2195
  - causes of, 2298*t*
  - childhood epilepsy and, 649
  - children and, 2281, 2385
  - conceptualization of, 1720–1721
  - conditioning theories and, 416
  - COPD and, 2226
  - CRH-1 and CRH-2 receptors and, 1745
  - description of, 4562
  - diabetic patients and, 2238
  - differential diagnoses, 2624*t*
  - due to another medical condition, 1723, 1729
  - effect on functioning, 2233
  - elderly cancer patients and, 2283–2284
  - endocannabinoids and, 114
  - Fabry disease and, 656

- GABAergic systems and, 83
- hypertension and, 2193
- itch–scratch response and, 2342
- life-threatening disease and, 2893
- maternal separation and, 2047
- medications for stabilization, 2171, 2172*t*
- and natural selection, 776
- neurochemical basis of, 1741–1742
- neurochemical systems involved in, 1757–1758
- neuroimaging studies of, 1758–1760
- obesity and, 2214
- PD and, 545
- perinatal, 2515
- presentation in advanced disease, 2298–2299
- spirituality and, 2587–2589
- stress and, 2321
- substance-induced, 2258
- suicide risk and, 2616
- symptoms of
  - child/adolescent, 3766, 3832
  - delirium and, 1184
  - dementia and, 1196
  - inhalant-induced, 1340
  - management of, 3743
  - parents during child psychiatric evaluation, 3380–3382, 3439, 3440
  - parents of ADHD children, 3931–3932
  - parents of autistic children, 3926–3927, 3928–3929
  - parents of bipolar children, 3929–3931
  - parents of infants with feeding disorders, 3624, 3628, 3638, 3642
  - parents of schizophrenia children, 3932–3933
  - schizophrenia and, 1421–1422, 3793
  - suicide behavior risk factor, 3684, 3698
- terminally ill children and, 2309–2310
- thyrotoxicosis and, 680
- treatment, 2298*t*
  - β-adrenergic receptor antagonists, 2954
  - in advanced disease, 2298–2299

    - antihistamines, 2982
    - MAOI, 3122–3123
    - medications, 2298–2299
  - types of, 2676
- Anxiety Disorder Interview Schedule-Revised (ADIS-R), 1721
- Anxiety disorders, **1720–1784**. *See also* Agoraphobia; Generalized anxiety disorder; Panic disorder(s); Separation anxiety; Social anxiety disorder
  - ADHD comorbidity, 3596
  - after disasters, terrorism, or war, 2567
  - after TBI, 537
  - age of onset, 1734
  - alcohol-induced, 1273
  - alcohol-use disorder and, 1275
  - anabolic steroid-induced, 1402
  - anxiolytics, 4128–4132, 4128*t* (*See also specific drugs*)
    - beta blockers and antihistamines, 4131
    - citalopram (Celexa), 4129
    - escitalopram (Lexapro), 4129
    - quetiapine XR, 4131
  - arginine vasopressin and, 1748
  - BD and, treatment of, 1698–1699
  - belief disconfirmation and, 1771
  - benzodiazepines and, 1746–1747
  - bipolar comorbidity, child/adolescent, 3689–3690, 3703–3704
  - brief psychotherapies for, 3742, 3745–3746
  - caffeine induced, 1297–1298, 1298*t*
  - in cancer patients, 2257–2259, 2258*t*
  - cannabis-induced, 1307
  - CBT for, 1768–1778, 4182, 4191
    - acceptance-based techniques, 1773
    - assessment, 1771
    - cognitive therapy, 1773
    - efficacy, 1774–1776
    - emotional processing and, 1769–1771
    - exposure therapy, 1772–1773
    - interpersonal skills training, 1774
    - psychoeducation, 1771–1772
    - stress inoculation training, 1773

- symptoms and conditions, 1768–1769
- treatment planning, 1772
- child/adolescent, 4562
  - aggression and, 2487
  - generalized anxiety, **3712–3714**
  - obsessive-compulsive disorder, **3695–3703**
  - posttraumatic stress disorder, **3703–3707**
  - psychopharmacology, 3792–3793, 3796
  - selective mutism, **3719–3724**
  - separation anxiety, **3708–3712**
  - social anxiety, **3714–3719**
- cholecystokinin and, 1749–1750
- classification of, 1158, 1172–1173, 2258
- clinical features, 1723–1730
- clinical manifestations of, 1138–1140
- cognitive studies, 699
- cognitive therapy, 2862
- comorbidity patterns, 1735
- conditions associated with, 2258*t*
- CRH and, 1744–1745
- description of, 4562
- diagnosis issues, 1778
- diagnostic interview for, 3398
- differential diagnoses, 1586, 3601, 3730–3731
- dissociativity, 1870
- dopaminergic system and, 1745
- DSM-5 modifications in, 1721
- due to another medical condition, 1235*t*, 1250*t*, **1253–1255**
- endocannabinoid system and, 1750
- epidemiology, 1730–1736
- epigenetics, 1740
- in epilepsy, 514
- ethnicity, 1734
- excoriation disorder and, 1796
- FASD and, 638
- future directions
  - anxiety disorders, 1784

    - CBT, 1776–1777
    - epidemiology, 1736
    - genetics, 1740–1741
    - neurobiology and neuroscience and, 1750
    - neuroimaging, 1767
    - nosology development, 1729–1730
  - GABA and, 1746–1747
  - galanin and, 1749
  - gender differences, 4246
  - genetics, 1737–1741
  - glutamate system and, 1747–1748
  - group psychotherapy, 2746
  - group therapy, 4207
  - hallucinogen-induced, 1327
  - HIV infection and, 646, 4269
  - HPA axis and, 1744
  - hypnosis and, 2721–2723
  - hypnotizability and, 2707
  - impact of, 1736
  - individual conditions and management, 1782–1784, 1782*t*, 1783*t*
  - inhalant-induced, 1340
  - inhibitory learning and, 1771
  - insomnia and, 2086–2087
  - with intellectual disability, 3511–3512
  - in long-term care, 4231
  - management principles, 1778, 1778*t*
  - migraine and, 622
  - mood disorders and, 1601, 1654, 1654*t*, 1669–1670
  - neuroanatomical circuits in, 1751–1768
  - neurobiology and neuroscience, 1741–1751
  - neuroimaging, 1760–1768
  - neuroimaging, child/adolescent, 3470–3473
  - neuropeptide Y and, 1749
  - neurosteroids and, 117, 1748
  - noradrenergic system and, 1742, 1744
  - nosological issues, 2676–2677
  - OCD and, 1788

- OCD compared to, 1786
- in older people, 3956, 3965–3967 (*See also* Geriatric psychiatry)
- oppositional defiant disorder and, 3611
- overview, 1721–1722
- oxytocin and, 1748–1749
- pain and, 2393–2394
- physician, 2601
- prevalence, 1720
- psychodynamic psychopharmacology, 2681–2682
- psychodynamic psychotherapy, 2676–2678
- psychopharmacotherapy, 3792–3793, 3796
- rating scale for, 3400–3401
- risk, developmental neuroimaging and, 1766
- risk of, premature births and, 644
- serotonergic system and, 1745–1746
- sex differences, 1734
- sleep disorders and, 3939
- social class, 1734
- somatic symptoms, 1828*f*
- somatic therapies, 1778–1784
- with specific learning disorder, 3529
- specific phobias and, 1784
- STDP and, 2810
- stimulant-induced, 1286
- stress and, 2323–2324
- substance/medication-induced, 1235*t*, 1250*t*, **1253–1255**
- suicide and, 2612, 2613
- suicide risk in, 1669–1670
- symptoms of, 2258
- temperament and, 3486, 3486*f*
- tic disorders and, 3642, 3656
- in transplant populations, 2366
- trazodone drug studies, 3249
- treatment
  - anticonvulsants, 1781, 2968, 2977
  - antidepressants, 1779–1780, 1779*t*
  - antipsychotics, 1781

        - benzodiazepines, 1780–1781, 2995, 2996
        - buspirone, 3010
        - in cancer patients, 2258–2259
        - delay in, *4526t*
        - exposure therapy, 1769, 1772–1773, 1774, 2690–2691
        - nefazodone, 3126–3127
        - novel, 1781–1782
        - pharmacological options, 1778–1782
        - response and illness course, 1782
    - treatment gap, 4522
    - TTM and, 1795
    - voice disorders comorbidity, 3567
    - VRGET for, 2560
    - world-wide prevalence of, 4517, 4518
- Anxiety Disorders Interview Schedule for Children (ADIS-C), 3398
- Anxiety Disorders Interview Schedule for Children (ADIS-IV-C and P), 1721
- Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV), 2685
- Anxiety hysteria. *See* Phobias
- Anxiety-like behavior, animal models of, 321
- Anxiety management training (AMT), 1773, 3743
- Anxiety NeuroGenetics STudy (ANGST) consortium, 1739
- Anxiety neurosis, 4562
- Anxiety sensitivity, 1735
- Anxiety states, description of, 4562
- Anxiolytic drugs
    - adverse effects of
        - dermatological, 2356*t*
    - classification, 2918
    - description of, 4562
    - in epilepsy, 520
    - in functional GI disorders, 2207
    - for IBS, 2204
    - for intellectual disability, 3515
    - sexual dysfunction and, 1973
    - in transplant patients
        - specific considerations, 2372*t*
- Anxiolytic-related disorders. *See* Sedative-, hypnotic-, anxiolytic-related disorders

- Anxious-depressive mania, 1640
- Anxious distress, 1679
- Anxious dysthymia, 1646. *See also* Dysthymia
- APA. *See* American Psychiatric Association (APA); American Psychoanalytic Association (APA); American Psychological Association (APA)
- Apathetic withdrawal. *See* Withdrawal
- Apathy
  - after stroke, 477, 477*t*, 478*f*
    - diagnosis of, 482, 482*t*
    - etiology of, 480
    - imaging findings, 480*t*
    - treatment of, 488
  - after TBI, 533–534
  - of aging adult, 3976
  - description of, 4562
  - in PD, 545, 548
  - recognition of, 464
  - treatment, 3229
- Apgar scores, 4562
- Aphagia syndrome, 364
- Aphasia. *See also* Language disorders
  - dementia and, 1194
  - description of, 461, 4562
  - mutism and, 457
  - syndromes, 462*t*
- Aphemia
  - characteristics of, 457*f*
  - mutism and, 457
- Aphonia
  - description of, 4562
  - as unspecified communication disorder, 3366
- Aphrodisiacs, 4562
- Apical ballooning syndrome, 2194
- Apnea, description of, 4562
- Apnea hypopnea, obstructive sleep, 4055–4056
- *Apo-e4* gene, 312–313
- Apokyn. *See* Apomorphine
- Apollonian–Odyssean–Dionysian cluster, 2717

- Apomorphine (Apokyn)
  - ED, 3340
  - history, 3233
  - receptor profile of, 67
  - for sexual dysfunction, 1978
- Apoplexy, 4562. *See also* Cerebrovascular accidents (CVA); Stroke
- Apoprotein E receptor 2 (ApoER2), 52
- Apoptosis. *See also* Programmed cell death
  - failure to inhibit, 46
  - p75-induced, 99
  - regulation of, 47*f*
  - types of, 46
- *APP* gene mutations, 312
- Apparently normal personality (ANP)
  - emotional personality and, 1903
- Appearance
  - assessment of, 455–456
  - at patient interview, 952, 967
  - in psychiatric disorders, 1142
- Appel, Kenneth
  - on hospital crowding, 4332
- Apperception
  - description of, 4562
- Appetite
  - A2 noradrenergic neurons and, 363–364
  - basic science of, 354–369
  - brain mechanisms of, 363–366
  - cannabis as stimulant for, 110
  - disturbances of, 1141
  - exercise and, 366–367
  - illness anorexia and, 367
  - megestrol acetate for, 2265
  - neural-network and, 355
  - orexins and, 677
- Appetitive phase, 4562
- Appetitive symptoms, review of, 455
- Applied behavioral analysis (ABA), 3513–3514

- Apprehension span, 696, 696f
- Approach-avoidance conflicts, 321
- Appropriate effect, description of, 4562
- Apraxia
  - description of, 4562
  - ideomotor, 461
  - left-hand, 451
  - speech, 457*f*
    - dementia and, 1194
    - as unspecified communication disorder, 3567
- Aprobarbital (Alurate), 2986
- Aprosodia, 449
- APS. *See* Adult Protective Services (APS)
- Aptitude, definition of, 4562
- Aptitude-by-treatment interaction (ATI), 2865
- Aquachloral. *See* Chloral hydrate
- *Arabidopsis,* studies of, 158
- Arachidonic acid (AA)
  - endocannabinoid synthesis and, 110
  - leukotriene synthesis and, 110
  - prostaglandin synthesis and, 110
  - schizophrenia and, 2545
  - signaling pathway, 123*f*
- *N*-Arachidonyldopamine (NADA), 110
- 2-Arachidonylglycerol (2-AG)
  - biosynthesis of, 111
  - brain development and, 58
  - discovery of, 110
  - neuroprotection in TBI, 115
- *N*-Arachidonylglycerol ether (noladin ether), 110
- Arachnoid villi, CSF absorption by, 14
- Araoz, Daniel, 2730
- Arc (formerly National Association for Retarded Citizens (NARC)), 3493
- Archetypes
  - description of, 4562
  - Jung, 912–913
- Archicortex, description of, 16

- Arcuate fasciculus (AF)
  - association fiber tracts and, 15*f*
  - connectivity, 451
- Arcuate nucleus (ARC)
  - eating control and, 364
- Arecoline, uses of, 73
- ARFID. *See* Avoidant/restrictive food intake disorder
- Arginine vasopressin (AVP), **175,** 3653, 3667
  - aggression and, 2481–2482
  - anxiety and, 1740
  - anxiety disorders and, 1743*t*, 1748
  - function of, **93–95**
  - secretion of, 2337
- Argumentative behavior, OCD and, 3610
- Aricept. *See* Donepezil
- Arieti, Silvano, 2867, 2872, 2876, 2877, 2878–2879
- Aripiprazole (Abilify)
  - for acute mania, 1686, 1686*t*, 1687, 1689, 1689*t*
    - lithium or divalproex, 1689*t*
    - lithium or valproate and, 1687
    - monotherapy, 1686
  - adverse effects of, 2356*t*
  - for aggressive behavior
    - in autism, 2500
    - in BPD, 2502
  - antipsychotic efficacy, 2169
  - for anxiety disorders, 1781
  - for BDD, 1792
  - for bipolar depression, 1691, 1692*t*, 1693, 1695, 3237
  - for bipolar disorder, 1696*t*, 3692, 3706, 3791
  - black box warning, 3240
  - chemistry, 3234, 3235*f*
  - in children and adolescents, 2628
  - combination pharmacotherapy and, 3324
  - dementia treatment, 1201
  - depression, 1681
  - diabetes risk and, 2247

- dissociative identity disorder, 1928
- dosage/administration, 2258*t*, 3241, 3241*t*
- drug interactions, 3241
- early-onset schizophrenia, 3732
- for excoriation disorders, 2352
- GAD, 1783
- history, 3234
- for mania, 1612
- for MDD, 3236
- for OCD, 1789
- organ/system effects of, 3238, 3239
- pharmacological action, 3235–3236
- precautions/adverse reactions, 3240, 3240*t*
- prolactin levels and, 2249
- in psychiatric emergencies, 2628
- schizophrenia, 1497–1498, 1497*f*
- side effects of, 67, 2237, 2237*t*
- therapeutic indications, 3239
- tic disorders, 3646, 3660

- *Aristolochia fangchi,* 2287
- Aristotle, 411, 928, 4533
- Arithmetic disorder. *See* Developmental arithmetic disorder
- Arithmetic skills, hemispheric dominance, 979
- *Arizona, Landrigan v.,* 2421–2422
- Armed security in schools, 4440
- Armodafinil (Nuvigil)
  - bipolar depression, 1691, 3237
  - combination pharmacotherapy and, 3324
  - dosage/administration, 3241, 3241*t*
  - drug interactions, 3240–3241
  - EDS, 3238
  - history, 3234
  - medical illness-induced fatigue, 3237–3238
  - organ/system effects of, 3239
  - performance enhancement, 3238
  - precautions/adverse reactions, 3240, 3240*t*
  - therapeutic indications, 3239

- use in cancer patients, 2262
- Armstrong, Judith, 1871
- Army Study to Assess Risk and Resilience in Service members (STARRS), 2568
- Aromatherapy, 3982
- Arousal (wakefulness). *See also* Sexual arousal disorders
  - animal models of, 324
  - circadian cycle of, 210
  - definition of, 4562
  - PTSD and, 1814, 1831*t*, 3705
- Arousal–appraisal system, 687
- Arrests, mental illness and, 4388
- Arrhythmias, psychiatric complications in, 2195
- Arsenic, toxic neuropathy and, 631
- Arsonists, 2113
- Art therapy, 4562
- Artane. *See* Trihexyphenidyl
- Arterial spin labeling (ASL), 271, 1476, 1476*f*
- Arteriosclerosis, 4562
- Arthritis
  - depression and, 2322
  - laboratory evaluation, 1015
  - psychiatric problems and, 4049
  - psychosomatic component, 915
  - treatment of, 2322
- Artificial insemination, 2510*t*
- Artificial intelligence (AI), 4562
- Ascending reticular activating system (ARAS), 344–346
- ASD. *See* Acute stress disorder (ASD); Autism spectrum disorder(s) (ASDs)
- Asenapine (Saphris)
  - in acute mania, 1686*t*, 1689, 1689*t*
    - lithium or divalproex, 1689*t*
    - lithium or valproate and, 1687
    - monotherapy, 1686
    - sublingual, 1688
  - bipolar depression, 1695
  - in bipolar disorder, 1696, 1696*t*, 3692, 3706
  - for schizophrenia, 4098

- side effects, 2237, 2237*t*
- Asendin. *See* Amoxapine
- Aseptic meningitis, 2273*t*
- Asexuality, definition of, 1982
- Asher, Richard, 1846, 1847
- Asian Americans. *See also* Demographics; Race/ethnicity
  - immigrants, and mental health, 804
- ASIC (acid-sensing ion channels), 139
- Asociality, definition of, 4562
- Asparaginase (Elspar)
  - delirium and, 2255
  - depression and, 2260
- Asperger syndrome. *See also* Autism spectrum disorder(s) (ASDs)
  - definition of, 4562
  - factors in development of, 55–57
- *Aspergillus nidulans,* 52
- Aspirin (acetylsalicylic acid, ASA)
  - for migraine treatment, 624
  - in stroke prevention, 486
- Assagioli, Robert, 2825
- Assault
  - sequelae of, 2540
  - treatment issues, 2540–2541
    - continuing the assessment, 2541
    - hospitalization, 2541
    - initial evaluation, 2540–2541
    - legal issues, 2541
    - postassault pharmacotherapy, 2541
    - postassault psychotherapy, 2541
    - safety, 2541
  - types, prevalence, and risk factors, 2538
- Assent
  - children and, 2311–2312
  - end-of-life decisions, 2311–2312
- Assertive Community Treatment (ACT), 1552
- Assertiveness training
  - description of, 4563

- in sexual therapy, 1977
- Assessments
  - definition of, 4563
  - in rehabilitation, 4369–4370
  - steps in, 4337
- Assimilation
  - during cognitive development, 3359
  - definition of, 4563
  - and migration, 792
- Assimilation and accommodation, 705
- Assisted living facilities (ALFs), 4197
- Assisted outpatient treatment, 4396
- Association, definition of, 4563
- Association for Academic Psychiatry, 4428
- Association mapping, 3450–3452, 3450*t*, 3451*t*
- Association of Medical Colleges (AAMC), 4563
- Association of Medical Officers of American Institutions for Idiotic and Feeble-Minded Persons, 3493
- Association relay nuclei, 22*t*
- Association studies
  - candidate gene
    - anxiety disorders, 1738–1739
    - mood disorder, 1625, 1626–1627, 1626*t*
  - family-based, 308
  - gene mapping, 308–309
  - late-life neurodegenerative disorder, 4019
- Associations, in patient interview, 958
- Associative learning
  - and memory, 729
  - procedural memory and, 2139
- Astasis–abasia, 2178, 4563
- Astemizole (Hismanal)
  - drug interactions, 2983, 3131
  - organ/system effects of, cardiovascular, 2983
- Astereognosis, 4563
- Asterixis, 458
- Asthenic personality, 4563
- Asthma

- bronchial, 915
- children's response to, 3846
- comorbidities, psychiatric, 2225
- historical perspectives, 2224
- hygiene hypothesis and, 192
- hypnosis and, 2729
- pathophysiology, 2224
- presentation, 2223, 2224
- prevalence, 2224
- psychosomatic component, 915
- social burden of, 2224
- stress, emotion and, 2224–2225
- treatment, 2224

- Astraphobia, 4563
- Astrapophobia, 4563
- Astrocyte to neuron ratios, 8
- Astrocytes
  - brain, 138
  - classification of, 3
  - features of, 4*f*
  - fibrous, 3
    - characteristics of, 4*t*
    - distribution of, 5
  - glutamate reuptake by, 196
  - intralaminar, 8
  - major features of, 4*f*
  - numbers of, 3
  - polarized, 8
  - protoplasmic, 3
    - characteristics of, 4*t*
    - human brain, 8
  - role of, 79
- Astrocytomas, hypothalamic, 677
- Astronauts, mental health of, 755
- Asylums, poorhouses as predecessor to, 4290
- Asymmetries of movement, 3540
- Asyndesis, description of, 4563

- *Ataque de nervios* (attack of nerves)
  - acute dissociative reactions to stressful events and, 1948, 1949
  - case study, 1949
  - clinical features, 2451
  - description of, 2448–2452, 4563
  - dissociative disorders and, 1871, 1896
  - epidemiology of, 2449
  - phenomenology of, 2449, 2449*t*
  - precipitants/etiology, 2449–2450
  - psychiatric diagnoses and, 2450–2451
  - specific cultural factors, 2451
  - treatment, 2451–2452
- Ataractic drugs. *See* Antipsychotics/neuroleptics
- Ataxia(s)
  - characteristics of, 540
  - chemotherapeutic agents and, 2273*t*
  - description of, 4563
  - differential diagnoses, 469*t*
- Atenolol (Tenormin)
  - absorption, 2951
  - BD and social anxiety disorder, 1698
  - benzodiazepine withdrawal, 1388
  - chemistry, 2951, 2951*f*
  - distribution, 2952
  - dosing/administration, 2956
  - drug interactions, 2955
  - drug studies, 2952
  - in hyperthyroidism, 2243
  - organ/system effects of, 2952–2953, 2953*t*
  - performance anxiety, 1781
  - pharmacodynamics, 2952
  - pharmacokinetics, 2951–2952, 2952*t*
  - pharmacological actions, 2951–2952, 2952*t*
  - precautions/adverse reactions, 2953
  - SAD, 1784
  - teratogenic effects of, 2953
  - therapeutic indications, 2953–2955, 2953*t*

- Atheism, incidence of, 2141
- Atheists, spiritual sensitivity in, 2141
- Atherosclerosis
  - CAD related to, 2186
  - progression of, 202
  - stroke and, 2194–2195
- Athetosis, description of, 4563
- Ativan. *See* Lorazepam
- *Atkins v. Virginia,* 4438
- Atomoxetine (Strattera)
  - ADHD treatment, 3594, 3693, 3707, 3789, 3794
  - bipolar disorder, early onset, 3693, 3707
  - child/adolescent, 650
    - ADHD treatment, 3693, 3707, 3788–3789
  - dosage/administration, 3223*t*
  - drug interactions, 3233
  - inhalant abuse/dependence, 1337
  - mechanism of action, 3232
  - metabolic effects of, 3796
  - pharmacological actions, 3232–3233
  - PTSD and SUD comorbidity, 1823
  - tic disorders, 3647–3648, 3661–3662
- Atonia
  - description of, 4563
  - sleep disorder and, 2100
- Atopic dermatitis (AD)
  - acute, 2341
  - chronic, 2341
  - clinical features of, 2340–2342
  - epidemiology of, 2340–2342
  - pathophysiology of, 2340–2342, 2341*f*
    - genetics, 2340–2341
  - presentation, 2340*f*
  - psychological influence of, 2334
  - psychopathology of, 2342
  - stress and, 2322
  - treatment of, 2342

- Atorvastatin (Lipitor)
  - for Alzheimer disease, 4147
  - drug interactions, 3131
- *ATP7B* mutation, 671
- Atrial fibrillation, 2190
- Atrophic glossitis, vitamin deficiency and, 456
- Atrophy, disuse-related, 2331
- Atropine
  - structure, 2956, 2956*f*
  - war or terrorism and, 2572*t*
- Atropine (AtroPen), drug interactions, 2199
- Attachment. *See also* Bonding; Reactive attachment disorder (RAD)
  - avoidant, aggression and, 2486
  - description of, 4563
  - disorganized
    - aggression and, 2486
    - hypnosis and, 2720–2721
  - disturbance, 2720–2721
  - mentalizing and, 2897–2898
  - secure, aggression and, 2486
  - strange situation, 2720, 2721
  - styles of
    - patient-provider relationships and, 2231
  - theories of, 886–887
    - betrayal trauma theory and, 1902–1903
    - interpersonal psychotherapy and, 1701
    - overview, 3662, 3663*t*, 3676, 3677*t*
    - for PTSD, child/adolescent, 3706
    - on social development, 3360–3361
  - type D, 1901
- *Attachment* (Bowlby), 4543
- Attachment disorder of infancy, 4563
- Attachment disorders. *See* Attachment; Disinhibited social engagement disorder; Reactive attachment disorder (RAD)
- Attacking others, shame and, 1907
- Attacks on linking, psychosis and, 2875
- Attention
  - aging and, 4005

- animal models of, 323
- definition of, 4563
- in depression, 698
- disorders of, 461–463
- neuropsychological testing, 990
- and processing by brain, 689–690
  - capacity of, 690
  - selectivity, 690
  - sustained, 690
- redistributed, hypnosis and, 2734
- in schizophrenia, 696

- Attention Bias Modification (ABM), 3711
- Attention deficit disorder (ADD)
  - after TBI, 529
  - FASD and, 638
  - gambling and, 1804
  - history of, 3587
  - neurosteroids and, 117
  - staring spells and, 650
- Attention-deficit/hyperactivity disorder (ADHD), 696–697, 3587–3604
  - adult
    - clinical features, 3600, 3600*t*
    - comparative nosology, 3599
    - course, 3602
    - defined, 3598
    - diagnosis, 3600, 3601*t*
    - differential diagnoses, 3601–3602
    - epidemiology, 3599
    - etiology, 3599–3600
    - further research, 3603–3604
    - historical background, 3598–3599
    - laboratory studies, 3600–3601
    - overview, 3598
    - prognosis, 3602
    - treatment, 3602–3603, 3603*t*
  - aggression and, 2487, 2488
    - treatment of, 2500

- alcohol-use disorder, 1275
- ASD and, treatment of, 3228
- attachment disorders and, 3666, 3680
- bipolar disorder and, 1699
- case study, 3597
- child/adolescent
  - bipolar disorder and, 3689, 3690, 3692–3693, 3703, 3704, 3706–3707
  - brief psychotherapies for, 3743–3744, 3747
  - family therapy, 3778
  - group therapy, 3765
  - parenting, 3931–3932
  - posttraumatic stress disorder (PTSD), 3705–3706
  - psychopharmacology, 3788–3789, 3788*t*, 3796
  - sleep disorders and, 3938
  - treatment, 3788–3789
- comparative nosology, 3588
- conduct disorder and, 3618–3619
- conduct disorder and, treatment of, 2947, 3228
- course, 3591–3592
- defined, 3587
- description of, 4563
- developmental coordination disorder and, 3538, 3541
- developmental disorders and, treatment of, 3228
- diagnosis, 3587–3588
- differential diagnoses, 3548, 3591, 3611
- dopamine receptor gene and transporter gene and, 1745
- DSM-5 criteria, 3587–3588
- EEGs in, 250–251
- encopresis and, 3659, 3673
- enuresis, 3651, 3665
- etiology, 3588–3590
  - clinical features, 3590–3591
  - diagnosis, 3590–3591
  - environmental factors and, 3589–3590
  - genetics, 3588–3589
  - neuroanatomy and, 3589
  - neurotransmitters and, 3589

- etiology and pathogenesis, 3219, 3223
- FASD and, 638–639
- gambling and, 1804
- genetic overlap with mood disorders, 1628
- genetics and, 2529*t*
  - empirical risks for, 2535*t*
- guanfacine for, 67
- history, 3587
- with intellectual disability, 3506, 3511
- intelligence testing and, 3426
- laboratory examination, 3591
- lisdexamfetamine for, 64
- malingering and, 2403, 2407–2408
- medical examination, 3591
- neuroimaging, child/adolescent, 3463–3464, 3463*f*
- neuropsychological testing for, 3433
- ODD and, treatment of, 2947, 3228
- omega-3 fatty acid supplementation in, 116
- oppositional defiant disorder and, 3611
- overview, 3587
- pharmacogenomics of, 3453–3455
- premature birth and, 644
- prenatal alcohol exposure and, 654
- processing assessment for, 3428–3429
- prognosis, 3591–3592
- QEEG analysis, 1036
- rating scales for, 3401–3402, 3591
- secondary (SADHD) after TBI, 652
- seizure disorder and, treatment of, 3229
- social (pragmatic) communication disorder and, 3552, 3553
- specific learning disorder and, 3529, 3533
- speech sound disorder and, 3557
- substance use and, 3230–3231
- substance use disorders and, 1275, 1286
- temperament and, 3485–3486, 3485*t*
- tic disorders and, 2947, 3637, 3642, 3651, 3656
- treatment, 3592–3597

    - bupropion, 3004, 3006
    - buspirone, 3012
    - clonidine, 2947–2948
    - of comorbidities, 3595–3596
    - guanfacine, 2947–2948
    - MAOIs, 3123
    - modafinil, 3232
    - nonstimulants, 3594–3595
    - psychosocial, 3596–3597
    - stimulants, 3592–3594, 3592*t*
    - sympathomimetic, 3219, 3223, 3227–3229, 3232
  - 22q deletion syndrome and, 2526*f*, 2530
- *Attention-Deficit Hyperactivity Disorder in Adults* (Wender), 4549
- Attentional difficulties
  - on ADHD, 3590, 3592
  - neuropsychological assessment of, 3433–3434
- Attentional networks
  - hypnosis and, 2712
  - threat detection and, 1819
- Attenuated psychosis syndrome, **1597**
- Attitude, definition of, 4563
- Attitudes of adolescents, 3385–3386
- Attractor network, 430–431
- Attributable risk, 4563
- Attribution, casual, 791
- Attributional bias, psychosis and, 2881–2882
- *ATX1* gene, 313
- Atypical, definition of, 4563
- Atypical antipsychotics/neuroleptics
  - for bipolar disorder, 3791
  - bipolar disorder, early onset treatment, 3692, 3706
  - obsessive-compulsive disorder and, 3701
  - schizophrenia, 3793
  - tic disorders, 3645–3646, 3659–3660
- Atypical child, definition of, 4563
- Atypical depression
  - about, 1653

- reverse neurovegetative symptoms and, 1710
- treatment selection and, 1679

- Atypical gender behavior, 2028
- Atypical psychosis, definition of, 4563
- AUDADIS. *See* Alcohol Use Disorder and Associated Disabilities Interview Schedule
- Audit. *See* Medical audits
- Auditory agnosia, 4563
- Auditory amnesia, 4563
- Auditory evoked potentials (AER), 2545
  - after TBI, 525
- Auditory hallucinations, 465
  - definition of, 4563
  - delirium and, 1186
  - schizophrenia and, 1411–1412
- Auditory learning, aging and, 4006
- Auditory perception, language disorders and, 3546–3547
- Auditory radiation, anatomy of, 11
- Aura(s), definition of, 4563
- Auschwitz survivors, amnesia in, 1887
- Australian National Survey (2010), 1428
- Australian National Survey of Mental Health and Well Being, 1618
- Authority figures, definition of, 4563
- Autism
  - after TBI, 653
  - aggression in, 2500
  - children with, 653
    - naltrexon for, 169
    - oxytocin levels in, 175
  - common gene variants in, 314–315
  - corpus callosum in, 13
  - definition of, 4563
  - epilepsy and, 654
  - facial dysmorphisms, 314
  - fragile X syndrome and, 3058
  - gene mapping, 313–315
  - genetic selection, 773
  - inflammation and, 202

- omega-3 fatty acid supplements and, 116
- premature birth and, 644

- Autism disorder (AD), aggression and, 2488
- Autism spectrum disorder(s) (ASDs), **3571–3586**. *See also* Asperger syndrome
  - ADHD and, 3228, 3596
  - aggression and, 2487, 2488, 2500
  - associated features of, 3581–3582
  - attachment disorders and, 3666, 3680
  - bipolar comorbidity, child/adolescent, 3690, 3704
  - brain development, 3576–3577
  - case report, 3584
  - clinical features, 3579–3582, 3580*t*, 3581*f*
  - cognitive studies, 693, 697–698, 698t
  - comparative nosology, 3573–3574
  - criteria for, 3349
  - *de novo* mutations, 157
  - definition of, 3571, 3571*t*, 3572*t*, 4563
  - diagnosis of, 2436, 3579–3582, 3579*t*
  - differential diagnoses, 3548, 3553, 3583–3584, 3583*t*, 3731
  - EEGs in, 251
  - emergency psychiatric evaluation for, 3445
  - epidemiology, 3574, 3574*t*
  - etiology, 3574–3579
  - factors in development of, 55–57
  - family therapy for, 3786
  - gender dysphoria and, 2033
  - gene dysregulation, 446*f*
  - genetic factors in, 56–57
  - genetic mapping in, 3451, 3452
  - genetics and, 2529*t*
  - historical background, 3571–3573
  - hoarding and, 1793
  - with intellectual disability (ID), 3506
  - laboratory examination, 3582–3583, 3582*t*
  - neuroimaging, child/adolescent, 3475–3476, 3476*f*, 3477*f*
  - neuroimaging studies, 3576–3579, 3578*f*
  - obsessive-compulsive disorder and, 3699

- OCD *versus,* 1788, 1789
- overview, 3571
- parents raising children with, 3926–3927, 3928–3929
- pathogenesis of, 56
- pathology, 3582–3583, 3582*t*
- physical characteristics, 3582
- premature birth and, 644
- processing assessment for, 3426
- psychopharmacology, 3793–3794, 3797
- schizophrenia, early onset and, 3727–3728, 3731
- schizophrenia and, 1
- signs of, 56
- sleep disorders and, 3938
- stereotypic movement disorders, 3671, 3685
- susceptibility genes, 313–314
- symptoms, 1
- temperament and, 3487
- treatment, 2948–2949, 3058, 3584–3586

- Autistic disorders
  - aggression and, 2487
  - cognitive development, 712
- Autistic thinking, definition of, 4563
- Autobiographical memory
  - dissociative amnesia and, 1879, 1931, 1935
  - dissociative identity disorder and, 1908, 1910–1911
- Autochthonous delusion, 1123
- Autoerotic asphyxiation, 1975, 4563
- Autoerotism, definition of, 4563
- Autogenic training
  - in functional GI disorders, 2207
  - guided imagery and, 2823
- Autogynephilia
  - definition of, 4563
  - in natal males, 2030–3031
- Autohypnosis
  - dissociation and, 1902–1903
  - dissociative identity disorder, 1909*t*

- in functional GI disorders, 2208
- Autoimmune cerebral folate deficiency (ACFD), 665
- Autoimmune disease(s). *See also specific diseases*
  - encephalopathy, autoimmune or paraneoplastic, 1015
  - hypnosis and, 2729–2730
  - laboratory evaluation, 1015
  - systemic lupus erythematosus, 1015
- Autoimmune disorders, failure of apoptosis and, 46
- Automatic judgement, 4564
- Automatic mentalizing, 2895
- Automatic obedience, 4564
- Automatic thoughts. *See also* Thought records
  - in cognitive model, 2760, 2760*f*
  - cognitive therapy and, 2696, 2761–2762
  - depression and, 1664
- Automatisms
  - definition of, 4564
  - in epilepsy, 510
  - history, 1869
- Automysophobia, 4564
- Autonomic hypoarousal, 2486
- Autonomic nervous system (ANS)
  - antihistamine effects on, 2982
  - definition of, 4564
  - dissociative identity disorder and, 1894
  - dysregulation, intimate partner violence and, 2487
  - fear conditioning, PTSD and, 1818
  - fear or stress response and, 1757
  - immune system and, 2319
  - neuroendocrine system and, 2337–2338
  - sleep and, 344
  - tricyclic and tetracyclic antidepressant effects on, 3260, 3262
  - window of tolerance, 2351*f*
- Autonomic side effects, 4564
- Autonomous depression, 1709
- Autonomous ego function, 4564
- Autonomy

- in correctional psychiatry, 4452
- narrative psychiatry and, 2836
- respect for, 2836, 2837–2838

- Autophagic degeneration, 46
- Autophobia, 4564
- Autoplasty, 4564
- Autopsies, legal issues, 2305–2306
- Autoreceptors, 61
- Autoscopic psychosis, 4564
- Autosomal dominant frontal lobe epilepsy (ADNFLE), 73–74
- Autosomal dominant genetic transmission, 1621, 1621*f*
- Autosomal recessive genetic transmission, 1621, 1621*f*
- Autotopagnosia, 4564. *See also* Finger agnosia
- Auxillary therapist, 4564. *See also* Co-therapy
- Avanafil, ED, 3339, 3339*t*
- Aventyl. *See* Nortriptyline
- Average, definition of, 4564
- Aversion-based behavior therapy, 2694
- Aversion conditioning, 2056
- Aversion therapy
  - classical conditioning and, 2694
  - ethical issues, 2703
- Aversive affect induction, 2056
- Aversive conditioning. *See* Aversion therapy
- Avery, Oswald, 149
- Avicenna, 3491, 4533
- Avocations, 3378
- Avoidance
  - in Pavlovian and operant learning, 725–726
  - PTSD, childhood/adolescent, 3704–3705
- Avoidance and escape learning, 4564
- Avoidance of shame affect, 1907
- Avoidance reflexes, 460
- Avoidance symptoms, 1813*t*, 1814
- Avoidant attachment
  - aggression and, 2486
  - definition of, 4564

- Avoidant attachment disorder, 1901
- Avoidant disorder of childhood or adolescence, 4564
- Avoidant personality disorder
  - agoraphobia and, 1783
  - clinical criteria, 2156
  - comorbidity, 2156
  - complications, 2156
  - course, 2156
  - definition of, 4564
  - differential diagnoses, 2156
  - epidemiology, 2156
  - impairment, 2156
  - SAD and, 1783
  - self-disorders and, 2808
  - sex ratio, 2156
  - social anxiety disorder and, 1725
  - STDP and, 2801
- Avoidant problem solving, 2696
- Avoidant/restrictive food intake disorder (ARFID)
  - of caregiver-infant reciprocity, **3629–3630**, **3643–3644**
  - case history, 2082
  - with concurrent medical condition, **3630–3631**, **3644–3645**
  - definition of, 4564
  - description of, 2066
  - diagnosis of, 2073
  - differential diagnoses, 2076–2077
  - historical literature on, 2067
  - infantile anorexia, **3623–3625**, **3637–3639**
  - nosology, 2068–2069
  - posttraumatic feeding disorder, **3626–3628**, **3640–3642**
  - sensory food aversions subtype, **3625–3626**, **3639–3640**
  - of state regulation, **3628–3629**, **3642–3643**
  - treatment, 2082
- Avolition, 1417
- AVP. *See* Arginine vasopressin
- Awareness, consciousness and, 691
- The Awareness of Social Inferences Test (TASIT), 4370

- Awareness training, habit reversal and, 2697
- Awl, William, 4331
- Axelrad, David, 2727
- Axid. *See* Nizatidine
- Axis I disorders
  - DBT and, 2791–2792
  - factitious disorder presenting as, 1851
  - hypnotizability and, 2708
- Axis II disorders
  - AEDP and, 2804
  - social skills training and, 2694
- Axon hillocks, 133–135
- Axona, for Alzheimer disease, 4148
- Axonal transport, 99–100
- Axons
  - action potential conduction in, 133–135
  - collaterals, 5
  - connections, 5–8
  - corpus callosum, 11
  - description of, 4564
  - developmental changes in, 5
  - function of, 3
  - myelinated, 134*f*
- Ayahuasca, 1316
- Ayurveda
  - assessment, 2544–2546
  - botanicals used in, 2554
  - in CAIM medicine, 3982
  - dissociative disorders, 1872
  - on energetic imbalances, 2546
- Azilect. *See* Rasagiline
- Azrin, Nathan, 2697

B

- B (believe), in A-B-C-model, 2696
- B cells
  - formation of, 179
  - function of, 181
- B vitamins
  - for Alzheimer disease, 4147
  - B_{12} deficiency, 1238
  - beneficial effects, 2547–2548
  - in CAIM medicine, 3981
- *Babesia microti* infection, 597
- Babesiosis, 597
- Babinski, Joseph, 1867, 1868, 1868*f*
- Babinski sign, 457, 460, 2178
- *Babseia* infections, 593
- Baby blues. *See* Postpartum depression
- Bachrach, Henry, 2674
- Back pain, chronic, 328
- Back propagation, 432
- Background pain, burn injuries and, 2376
- Backlash against managed care, 4293
- Backpage Infrequency (Fb) Scale, 999
- Backward chaining, 2873
- Baclofen (Lioresal), 3003
  - for alcohol use disorders, 1278
  - anxiety disorders and, 1746
- *Bacopa monnieri,* 2554
- Bacterial toxins, G proteins and, 2912
- Baker, Elgan, 2721, 2723
- Balanced Budget Act (1997), 4311
- Balancing selection, description of, 156
- Balint, Michael, 2644
- Balint syndrome, 452, 463
- Bandura, Albert, 3750
  - anxiety disorders, 1774
  - behavior therapy, 2683

- social learning theory, 930, 2695
- Banjul Charter on Human and People's Rights, 4530
- Banks, Russell, 2834
- Bannayan-Riley-Ruvalcaba syndrome, 314
- Barabasz, Areed, 2730
- Barber, Theodore, 2710, 2724
- Barber Suggestibility Scale, 2713
- Barbiturates, **2984–2991**
  - absorption, 2984
  - activities of, 81
  - anxiety disorders and, 1746
  - blood concentrations, 2985
  - cerebral lateralization, 2987
  - chemistry, 2984, 2984*f*
  - delirium tremens, 2987
  - dependence, 1384, 2988
  - description of, 4564
  - distribution, 2984–2985
  - dosage/administration, 2985*t*, 2988, 2989*t*
  - drug interactions, 2199, 2988
  - drug studies, 2985–2986
  - ECT, 2986, 2987
  - EEG effects, 252
  - history, 2984
  - intoxication, 1383–1384
  - laboratory interferences, 2988
  - mechanisms of action, 2985
  - metabolism/elimination, 2985
  - narcoanalysis, 2987
  - neuroprotection, 2987
  - organ/system effects of, 2986
  - overdose, 1384
  - overview, 1383, 1384*t*
  - pharmacological actions, 2984–2985, 2985*t*
  - pharmacology, 1375, 1376
  - precautions/adverse effects, 2987–2988
  - psychomotor impairment, 2988

- sedative-hypnotic withdrawal and, 2987
- seizure disorder, 2986
- sexual dysfunction and, 1973
- teratogenicity, 2988
- therapeutic indications, 2986–2987
- tolerance, 1376
- withdrawal, 1384, 1384*t*, 1386*t*, 1389, 1389*t*

- Barbiturates-related disorders, 1383–1384, 1384*t*, 1389, 1389*f*
- Bariatric surgery, psychological considerations, 2221–2222
- Barlow, David, 2722
- Barlow, Horace, 413
- Barnier, Amanda, 2732
- Barriers to care
  - demographic, 4304
  - disparities as, 4384
  - insurance reimbursement rates as, 4302
  - linguistic/cultural, 4380, 4381
  - mental illness as, 4530
  - poverty as, 4305
  - in primary care, 4296
  - severe psychiatric disorder as, 4304
  - shortage of trained mental health professionals, 4304
  - sick leave/work scheduling/transportation/childcare as, 4384
- Barriers to change, world-wide, 4528
- Barrough, Phillip, 1179
- Barthel Index, 4032
- *Bartnicki v. Vopper* (2001), 2607
- *Bartonella* spp.
  - *B. henselae,* 593–594
  - *B. koehlerae,* 594
  - *B. vinsonii,* 593–594
  - infections, 593–594
- Bartonellosis
  - symptoms of, 594
  - treatment, 594
- Basal ganglia (BG)
  - anatomy of, 8, 10

- description of, 4564
- dysfunction, depression and, 1709
- embryonic origins of, 8
- functional circuitry, 29–31
- IFN-α and activity of, 196–197
- inputs into, 29, 30*f*
- internal processing, 29–30
- intrinsic circuitry of, 31*f*
- major depressive disorder and, 274
- major structures, 26–27
- neuroimaging
  - ADHD, 3464
  - OCD, 3467–3468, 3468*f*
  - Tourette syndrome, 3471–3472, 3473*f*
- noradrenergic neurons in, 1714
- nuclei and related structures of, 28*f*
- outputs of, 30–31, 32*f*
- Parkinson disease and, 543
- patterning genes, 50–51
- ring-enhancing lesions, 585
- schematic drawing of, 27*f*
- schizophrenia and, 1459, 1460*t*

- Basic anxiety, 4564
- Basic benefits, description of, 4564
- Basic/clinical research
  - health services research *vs.,* 4347, *4348f*
- Basic fibroblast growth factor (bFGF)
  - adult neurogenesis and, 52
  - gliogenesis in, 58–60
- Basic helix–loop–helix (bHLH) family, 48, 49
- Basic Personality Patterns, 2162
- Basilar artery, migraine associated with, 623*f*
- Basiliximab, 2371
- Basket neurons, parvalbumin-containing, 17
- Basolateral complex, anatomy of, 35
- Basophils, differentiation of, 179
- Bateson, Gregory, 2749, 2753

- Bath salts, 1282
- Bathophobia, 4564
- Battered child syndrome, 3663, 3677, 4564
- Batterer intervention programs (BIPs), 2502
- Battery tests, description of, 4564
- Battie, William, 4534
- Battle fatigue, 2567–2568, 2706
- Baxter, L. R., Jr., 2812
- Bayes, Thomas, 412
- Bayesian inference, 435, 835–836
- Bayesian network, 838
- Bayesian theorem, 835
- Bayesianism, rise of, 412
- Bayley Scale of Infant and Toddler Development III (Bayley-III), 3389, 3391
- *BCL11A* gene, 1740
- BCNU (Carmustine, Gliadel), 2255
- BD. *See* Bipolar disorder
- BDAE. *See* Boston Diagnostic Aphasia Examination
- BDNF. *See* Brain-derived neurotrophic factor
- Beachamp, Tom, 2836
- Beadle, George, 149
- BEAM. *See* Brain electrical activity mapping
- Beard, George M., 4564
- Beauty
  - awareness of, 2140
  - unstable sense of, 2140
- Beck, Aaron T., 3750–3751
  - *Cognitive Therapy and the Emotional Disorders,* 4546
    - belief disconfirmation and, 1771
    - cognitive-behavioral theory and, 1664–1665, 1701, 2498, 2695–2696
    - cognitive-behavioral therapy and, 2881
    - cognitive therapy and, 2760, 2761, 2763
    - on suicide ideation, 2620
- Beck Depression Inventory (BDI), 992, 995, 1003*t*
  - in cancer patients, 2261
  - for children, 3399
  - dissociative disorders, 1883

  - MDD assessment for CBT, 1704
- Beck Hopelessness Scale, 2615
- Becker, Ernest, 2288
- Becker dystrophy, 634
- Beckwith–Wiedeman syndrome, 50
- *Becoming: Basic Considerations for a Psychology of Personality* (Allport), 931
- BED. *See* Binge eating disorder
- Bed nucleus of the stria terminalis (BNST), 35, 1744
- Bed rest, disuse-related atrophy and, 2331
- Beds
  - in emergency departments, 4305, 4325, 4330, 4334
  - shortages in psychiatric hospitals, 4325, 4330
  - world-wide, 4525
- Beers, Clifford, 4291, 4537, 4564
- Befloxatone, chemistry of, 3113*f*
- Behar, Kevin, 278
- Behavior Assessment System for Children (BASC), 2687, 3399
- Behavior contracts, description of, 4564
- Behavior disorders, 1010. *See also specific disorders*
- Behavior mastery training, 2053, 2054
- Behavior modification therapy. *See also* Behavior therapy
  - BIF patients, 2439
- *Behavior of Children Receiving Benzedrine* (Bradley), 4537
- Behavior patterns
  - in attention-deficit/hyperactivity disorder, 697
  - in developing intelligence, 707
- Behavior syndromes
  - DSM-5 classification of, 522*t*
- Behavior therapy, **2682–2705**. *See also* Cognitive-behavioral therapy (CBT); Dialectical behavior therapy
  - of aging adult, 3974
  - assessment, 2684–2685
    - behavioral observation, 2687–2688
    - case studies, 2685, 2686, 2687, 2688–2689
    - clinician rating tools, 2685–2686
    - direct, 1006
    - interviews, 2685
    - methods, 1006–1007, 2685–2688

- questionnaires, 2685–2686
- rating scales, 1006–1007
- aversion-based, 2694
- case conceptualization, 2688–2689
- clinical issues, 2700–2702
- and cognitive function, in psychotherapy (*See* Cognitive-behavioral therapy)
- complications, 2702
- conceptualization, 2684–2685
- contingency management, 2873–2874
- contraindications, 2702
- couples, 2751*t*, 2759
- definition, 2682
- in delirium, 4089, 4093
- description of, 4565
- dissociative disorders, 1869
- empirically validated, 2704–2705
- ethical issues, 2703
- evaluation, 2862
- family, 2751*t*
- goals, 2701–2702
- group and couples, 1703
- group psychotherapy and, 2738
- history, 2682–2683
- indications, 2700–2701
- case study, 2702
- culture and, 2701
- development and, 2700–2701
- implementation variables, 2701
- insomnia
- in aging adult, 4054
- introduction, 2682
- and learning theory, 717
- limitations, 2702
- MDD, 1702*t*
- neurobiological effects of, 2814*t*
- operant strategies, 2691–2693, 2692*f*
- positive psychology and, 2854

- principles, 2683–2684
- research and evaluation, 2703–2705, 2704*f*
- in schizophrenia, 4098
- skills training, 2694–2695
- strengths, 2701
- techniques, 2684
- terrorist attacks and, 2572
- treatment strategies
  - first wave, 2689–2695
  - second wave, 2695–2698
  - third wave, 2698–2700

- Behavioral activating system (BAS)
  - temperament and, 3481
- Behavioral activation (BA)
  - applying research to clinical practice, 1704
  - in behavior therapy, 2698–2699
  - case study, 2699
  - cognitive therapy and, 2772
  - depression, 1702
  - efficacy research, 2695
- Behavioral analysis
  - DBT and, 2789–2790
- Behavioral avoidance tests (BATs), 2687
- Behavioral changes
  - after TBI, 527–528
    - course/prognosis, 539
- Behavioral contracts, 578
- Behavioral disorders of childhood. *See* Disruptive behavior disorders
- Behavioral disturbances
  - dementia and, 1196–1197, 1201–1202
  - developmental, 3375–3376
  - frontotemporal dementia variant and, 1210–1211
  - trazodone drug studies, 3250–3251
- Behavioral factors
  - ADHD treatment effects, 3593
  - of stereotypic movement disorders, 3668, 3682
  - temperament and, 3481

- Behavioral health care
  - for adults with serious/persistent mental illness, 4321–4322
  - for children with serious emotional disorders, 4320–4321
  - primary care integration with, 4320
  - spending cuts in, 4329
- Behavioral health disorders
  - minority subgroups *vs.* white Americans, 4304
- Behavioral Health Managed Care Organizations (BHMCOs)
  - Medicaid use of, 4294
- Behavioral inhibition, anxiety disorder risk and, 1766
- Behavioral inhibition (BI) and disinhibition (BD), 3481
- Behavioral insomnia of childhood, 2086
- Behavioral interventions
  - with developmental disabilities, 4343
  - immunity and, 203
- Behavioral marital therapy (BMT)
  - depression, 1703
  - MDD, 1702*t*
- Behavioral medicine, 4564
- Behavioral neurology
  - description of, 4564
  - neuropsychiatry and, **448–685**
- Behavioral reaction in cognitive model, 2760, 2760*f*
- Behavioral science, 4564
- Behavioral skill training/rehearsal, 2700
- Behavioral tailoring, 4303
- Behavioral therapies
  - for ADHD, 3596, 3597
  - for adolescents, 3814
  - for antisocial characteristics, 4392–4393
  - autism spectrum disorder and, 3585
  - for chemotherapy-related nausea, 2271–2272
  - for dementia, 1199
  - encopresis and, 3660, 3674
  - enuresis, 3655, 3657, 3669, 3671
  - for paraphilic disorders, 2021
  - posttraumatic food aversions and, 3627, 3641

- selective mutism, 3722–3723
- for social (pragmatic) communication disorder, 3554
- stereotypic movement disorders, 3672, 3686
- tic disorders, 3643–3644, 3657–3658
- tobacco use disorder treatment, 1349, 1545
- for weight loss, 1546

- Behavioral weight loss (BWL), 2216–2217
- Behaviorally-oriented clinicians
  - dissociative disorder controversy and, 1873
- Behaviorism. *See also* Behavior therapy
  - description of, 4564
- Behaviorist theories of paraphilias, 2016
- Behavior(s). *See also specific behaviors*
  - addictive, in patient interview, 950–951
  - aggression, hormone-induced, 165
  - application of theory to therapy, 931
  - and brain activity, neuroimaging of, 4482–4483
  - changes in brain tumor patients, 502–504
  - compensatory, 2165
  - cytokine effects on, 194
  - cytokines and, 2321
  - daily diary format, 2325*t*
  - defensive, animal models of, 320–321
  - definition of, 4564
  - dementias in long-term care, 4226
  - direct measurement, 1006
  - eating disorders, 367–368
  - ecology of, 774–775
  - evolutionary history, 769–770
  - FGA effects and, 3075, 3075*t*
  - genetic studies, and environment, 937
  - hypnotic experiences and, 2708
  - immunity and, 2320
  - impaired inhibitions, 2045–2048
  - maternal, 166
  - mathematical models of, 434–435
  - modification, aggression

    - Alzheimer disease, 2502
    - autism, 2500
  - modification, ethical issues, 2703
  - neuroscience of eating, 355
  - neurotrophins and, 100
  - nitric oxide and, 107
  - omega-3 fatty acid supplements and, 116
  - personality theories, 929–930
  - postpartum, 166
  - psychopathology theories, 930–931
  - repetitive, 458
  - rewards and, 415
  - risk-taking, premature births and, 643
  - sexual, 1956–1957, 1961–1963
  - and sexuality, 4264–4267, 4266*t*
  - sexuality in aging adult, 4256–4258
  - and social learning approaches, 929–931
    - radical behaviorism, 929–930
  - social learning theory, 930
  - social needs, 925–926
  - type A, components of, 2188
- Belatacept, 2371
- Belief systems
  - during adolescence, 3369
  - disconfirmation, anxiety disorders and, 1771
  - pain problems and, 2397
  - psychiatric symptomatology and, 1121, 1122–1123
- Bell, Luther, 4331
- Bell-and-pad technique, 2682, 3657, 3671
- Bell Lysaker Emotion Recognition Task (BLERT), 4370
- Bell's mania, 1640
- *The Belmont Report* (National Commission for Protection of Human Subjects), 4241
- Benadryl. *See* Diphenhydramine
- Bender, Lauretta, 3668, 3682
- Bender–Gestalt Test, 4565
- Benedict, Ruth, 2460
- Beneficence and nonmaleficence, 4241

- Benign sleep myoclonus of infancy, 2104
- Benjamin, Harry, 2025
- Benson, Herbert, 2584
- Benton Visual Retention Test, 991
- Benzodiazepine-related disorders, 1382–1383, 1382*t*, 1386–1387, 1386*t*
- Benzodiazepine(s), 4541
  - absorption, 2993
  - abuse, 2998–2999
  - addiction, 2997
  - agitation, 2997–2998
  - agitation management, 1525
  - alcohol and, 3000
    - withdrawal from, 2997
  - Alzheimer disease risk with, 4342
  - anesthesia, 2997
  - antidepressants and, 2996
  - for anxiety disorders, 1744, 1780–1782, 1782*t*, 2995–2996, 4130–4131
    - agoraphobia, 1783
    - between-session learning and, 1770, 1771
    - GAD, 1783
  - BD, 2997
  - BD and PTSD, 1699
  - benzodiazepine use disorder, 1378–1380, 1378*f*
  - binding, panic disorder and, 1761–1762
  - blood concentration, 2995
  - for borderline personality disorder, 2173
  - in burn patients, 2375
  - catatonia, 2997
  - chemistry, 2992, 2993*f*, 3000–3001, 3001*f*
  - chlordiazepoxide, 2917
  - for chorea, 555
  - classification, 2918
  - cognitive impairment, 2998
  - delirium and, 1182, 1188, 1189–1190
  - dementia aggravation, 2998
  - dementia treatment, 1201
  - dementia with Lewy bodies treatment, 1218–1219

- dependence, 2998–2999
- depersonalization/derealization disorder, 1945
- description of, 4565
- discontinuation syndrome, 2999–3000
  - rebound, 2999
  - relapse, 2999
  - withdrawal, 2998, 2999–3000, 2999*t*
- disinhibition, 2998
- dissociative amnesia, 1940
- dissociative identity disorder, 1891, 1927*t*, 1928
- distribution, 2993
- dosage/administration, 3000, 3001*t*
- drug interactions, 3000, 3054
- drug studies, 2995
- dysphoria, 2999
- ECT interactions with, 3295
- EEG effects, 252, 253
- in elderly patients, 2283, 2998, 2999
- FGAs and, 3081
- GABA and, 2991–2992, 2991*f*
- GABA receptors and, 171
- GAD, 1783, 2995–2996
- gambling disorder and, 1808
- hallucinogen persisting perception disorder treatment, 1324
- infectious disease, 2998
- insomnia, 2996–2997
- for insomnia, in aging adult, 4054
- for insomnia in dying patients, 2295–2296
- intoxication, 1382
  - psychiatric emergency due to, 2630
- investigational drugs, 3003
- laboratory interferences, 3000
- lactation and, 2999
- MDD, 2996
- mechanism of action, 2994–2995
- medical use, 1379
- metabolism, 2993–2994, 2994*f*

- misuse of, 4114
- mood disorder due to another medical condition, 1251
- neuroleptic side effects, 2997
- organ/system effects of, 2995
- overdose, 1383
- pain, 2997
- for pain management, 2381
- panic disorder and, 1724, 1761–1762, 2995, 2996
- in parkinsonian disorders, 548
- pharmacodynamics, 2994–2995
- pharmacological actions, 2993–2995
- pharmacology, 1375
- precautions/adverse effects, 2998–3000
- pregnancy and, 2999
- psychiatric emergencies, 2628
- psychomotor impairment, 2998
- psychopharmacology with, 2905
- psychosocial therapies using, 1388
- PTSD prevention, 1825
- receptor agonists, **2991–3004**
  - partial, 3000–3002
  - selective, 3000–3002
- receptor antagonists, **3002–3004**
  - dosage/administration, 3003
  - drug interactions, 3003
  - history, 2991–2992
  - laboratory interferences, 3003
  - pharmacological actions, 3002
  - precautions/adverse reactions, 3003
  - therapeutic indications, 3003
- receptor ligands, 2992
- receptors for
  - anxiety disorders and, 1746
  - fear and, 1757
- SAD, 1784
- schizophrenia, 2997
- sedation, 2997, 2998

- seizure disorder, 2997
- short-term, 2324
- side effects
  - cognitive, 2371
  - respiratory depression, 2225
- spasticity, 2997
- specific phobias, 1784
- for stress management, 2324
- tapering off, 2381
- therapeutic indications, 2995–2998
- tic disorders, 3646–3647, 3660–3661
- tinnitus, 2997
- tolerance, 1375
- violence, 2997–2998
- withdrawal, 1386–1388, 1386*t*, 2998, 2999–3000, 2999*t*
  - psychiatric emergency due to, 2632–2633
- withdrawal syndrome, 1382–1383, 1382*t*

- Benztropine (Cogentin)
  - absorption/distribution, 2957
  - agitation, 1525
  - chemistry, 2957, 2957*f*
  - dosage/administration, 2960, 2960*t*
  - drug interactions, 2960
  - dystonic reactions, 2628
  - for Parkinson disease symptoms, 73
  - pharmacodynamics, 2958
  - side effects, 1528
  - therapeutic indications, 2959
- Bereavement. *See also* Grief; Mourning
  - after disasters, terrorism, or war, 2567
  - and anxiety disorder, 4061
  - child/adolescent, 3766–3767
  - description of, 4565
  - as exclusion criteria, 1154, 1172
  - group therapy and, 3766–3767
  - IPT adaptation for, 1701
  - normal, 1653–1654

- normal and pathological, 1711
- as psychosocial factor, 4165–4156
- spousal, 3998, 4067
- treatment, 4075

- Berger, Hans, 342
- "Berger rhythms," 250
- Bergmann glia, 45
- Berkeley Puppet Interview (BPI), 3411
- Berne, Eric, 925f, 2737, 2738, 2820
  - *Games People Play,* 925
- Bernheim, Hippolyte, 1867, 2640, 2706
- Bernoulli random variable, 840
- Berquerel (Bq), 286
- Berrios, German E.
  - corticolimbic disconnection hypothesis, 1891
  - dissociative identity disorder, 1871
  - somatic presentations in classical psychopathology, 1829
- Bertallanffy, Ludwig von
  - family therapy and, 2749
- Bertholdt, Arnold Adolphe, 2480
- Bertillon, Jacques, 1152–1153, 1165
- Berwick, Don, 4309
- Best Pharmaceuticals for Children Act (BPCA), 2931
- Best practices
  - interventions for, 4348
  - for schizophrenia, 4350
  - scientifically validated treatments as, 4347
- Bestiality, definition of, 4565
- Beta (β) error, 4619
- Beta blockers
  - for anxiety disorders, 4131
  - drug interactions, 2199
  - in hyperthyroidism, 2243
  - psychiatric side effects of, 2198*t*
- Beta cells, damage to, 2246
- Beta waves, 4565
- β-Amyloid protein, 312, 665, 1203–1204

- Bethanechol, 73
- Betrayal blindness, 1903
- Betrayal trauma theory (BTT), 1900–1901
- Between-session extinction, OCD and, 2691
- Between-session learning, 1770, 1771
- Betz cells, 24–25
- *Beyond the Pleasure Principle* (Freud), 899
- BG. *See* Basal ganglia
- *Bhang* (cannabis tea), 1304
- Bhutanese refugees, mental health of, 806
- Bias in sampling, 3962–3963
- Biases, 4565. *See also* Statistics
- Bibliotherapy, behavior therapy and, 2701
- Bibring, Edward, 2644
- Bieber, Irving, 1987
- Big data, for seniors, 4280
- Bilateral cingulotomy, 2133
- Bile acids, production of, 662
- Bilirubin, conversion of, 108
- Binaural sound, 2546
- Binder, Jeffrey, 2647
- Binet, Alfred, 4537
- Binge eating, sex differences, 369*f*
- Binge eating disorder (BED)
  - assessment of, 2072*t*
  - classification of, 1174
  - course/prognosis in, 2078
  - description of, 2066, 4565
  - diagnosis of, 2072–2073
  - diagnostic criteria, 2066*t*
  - differential diagnoses, 2076
  - epidemiology of, 2070
  - etiology of, 2071
  - historical literature on, 2067
  - nosology, 2068
  - obesity and, 2213–2214
  - treatment of, 2081, 2970, 3006

- Binswanger disease, 451
- Biochemical genetic conditions. *See* Inborn errors of metabolism (IEMs)
- Biodynamics, 4565
- Biofeedback
  - for acne, 2345
  - anxiety disorders, 1782
  - as CAM therapy, 2558, 2559
    - case study, 2560
  - description of, 4565
  - factitious disorder, 1863
  - in functional GI disorders, 2207, 2208
  - for insomnia, 2089
  - scientifically validated therapy with, 2546
  - somatic symptom disorders, 1844
- Biogenic amine imbalance
  - hypotheses, 1607–1608, 4565
  - mood disorders and, 1607–1609
  - neuroendocrine links, 1608–1609
- Biogenic amine neurotransmitters, **61–75,** 4565
- Bioinformatics, 310–311
- Biological depression, neurobiology and, 1709
- Biological factors
  - of alcohol-use disorder, 1269–1270
  - for autism spectrum disorder, 3574–3575
  - conduct disorder, 3606, 3608, 3614–3615
  - encopresis, 3659, 3673
  - for oppositional defiant disorder, 3606, 3608
  - posttraumatic stress disorder (PTSD)
    - child/adolescent, 3704, 3707
  - suicide behavior risk factor, 3683–3684, 3697–3698
- Biological markers, aging brain, 3986
- Biological mediators of alcohol-use disorder risk, 1269–1270, 1269*t*
- Biological psychiatrists, 4292
- Biological psychiatry, 4294
- Biological psychology, 4565
- Biological rhythms, 4565. *See also* Circadian rhythms
- Biological therapies, **2547–2556, 2905–3343**

- acamprosate, 3054–3059
- acetyl-L-carnitine, 2550
- *S*-adenosyl methionine, 2549–2550
- α-adrenergic receptor agonists, 2945–2950
- β-adrenergic receptor antagonists, 2951–2956
- amantadine, 2962–2965
- anticholinergics, 2956–2965
- anticonvulsants, 2966–2979
- antihistamines, 2979–2984
- barbiturates, 2984–2991
- benzodiazepine
  - receptor agonists, 2991–3004
  - receptor antagonists, 2991–3004
- in biomedicine, 2546
- botanical, 2552–2554
- brain stimulation methods, 3299–3313
- branch chain amino acids, 2550–2551
- bupropion, 3004–3008
- buspirone, 3008–3014
- calcium channel inhibitors, 3014–3022
- carbamazepine, 3022–3039
- Chinese herbal formulas, 2554–2555
- cholinesterase inhibitors, 3039–3050
- compound herbal formulas, 2554
- dietary modification, 2547
- disulfiram, 3050–3054
- dopamine receptor agonists, 3233–3242
- drug approval process and, 2926–2936
- drug development and, 2926–2936
- ECT, 3280–3298
- erectile disorder treatment, 3337–3342
- FGAs, 3060–3082
- glycine, 2550–2551
- hormones, 2555–2556
- lithium, 3088–3101
- MAOIs, 3111–3125
- medication-induced movement disorders, 2936–2945

- melatonin receptor agonists, 3101–3107
- memantine, 3048–3049
- minerals, 2549
- mirtazapine, 3107–3111
- N-acetylcysteine, 2550
- nefazodone, 3125–3132
- neuropsychiatric syndromes or symptoms and, 2572*t*
- neurosurgical treatments, 3313–3322
- nutrient formulas, 2552
- omega-3-fatty acids, 2551
- opioid maintenance treatments, 3132–3138
- opioid receptor antagonists, 3138–3146
- pharmacotherapy, combination, 3322–3329
- phosphatidyl serine, 2551–2552
- posttraumatic stress disorder (PTSD), 3707
- psychodynamic psychotherapy and, 2868
- psychopharmacology and, 2905–2926
- reproductive hormonal therapy, 3329–3337
- SNRIs, 3146–3153
- SSRIs, 3153–3170
- sympathomimetics, 3219–3233
- tetracyclic antidepressants, 3253–3266
- L-theanine, 2550
- thyroid hormones, 3242–3247
- trazodone, 3247–3253
- tricyclic antidepressants, 3253–3266
- valproate, 3266–3280
- vilazodone, 3342–3347
- vitamins, 2547–2548

- Biomarkers
  - aggression, 2486
  - clinical applications of QEEG, 1034–1040
  - drug development and, 2926, 2929
  - for MCI progression, 1224, 1225, 1227
  - neurophysiological
    - modeling of, 435–436
  - psychosis, 1513–1516, 1514*f*, 1515*f*

- schizophrenia, 1566–1569, 1567*f*
- suicide, 2619

- Biomedical research in jails and prisons, 4474
- Biomedicine. *See* Complementary, alternative, and integrative medicine
- Bion, Wilfred
  - attacks on linking concept, 2875
  - group psychotherapy and, 2737
  - as psychoanalyst, 2644
  - psychodynamic psychotherapy, 2868
- Biperiden (Akineton)
  - chemistry, 2957
  - dosage/administration, 2960, 2960*t*
  - pharmacodynamics, 2958
  - side effects, 1528
  - therapeutic indications, 2959
- Bipolar affective disorder
  - classification of, 1168
  - defined, 3725
  - differential diagnoses, 3730
  - with psychotic features, 3730
- Bipolar depression
  - acute
    - treatment options, 1690–1692, 1692*t*
    - treatment recommendations, 1692–1693, 1692*t*
  - bipolar I disorder and, 1697
  - bipolar II disorder and, 1697–1698
  - differentiating unipolar depressions and, 1642*t*
  - EEG changes, 2545
  - maintenance therapy/prophylaxis, 1693–1695
  - treatment, 3693, 3707
    - bupropion, 3005
    - carbamazepine, 3031*t*
    - DA receptor agonists, 3236–3237
    - ECT, 3090
    - lithium, 3089–3090, 3089*f*
    - MAOIs, 3121
    - nefazodone, 3126

    - SSRIs, 3155
    - valproate, 3270, 3273*f*–3274*f*
  - treatment, child/adolescent, 3693, 3707
- Bipolar disorder(s) (BD)
  - acute mood symptoms, 1684
  - ADHD and, 1699, 3596
  - after TBI, 534
  - in aging adults, 3965, **4071–4077**
    - complementary and alternative medicine, 4074
    - electroconvulsive treatment, 4074
    - pharmacotherapy, 4127–4128, 4134
      - bupropion, 4073–4074
      - MAO inhibitors, 4073
      - mirtazapine, 4074
      - selective serotonin reuptake inhibitors, 4072–4073
      - serotonin-norepinephrine reuptake inhibitors, 4073
      - tricyclic antidepressants, 4073
      - vortioxetine, 4074
    - psychotherapy, 4072
    - vagal nerve stimulation, 4074–4075
  - alcohol-use disorders and, 1270, 1270*t*, 1272, 1275, 1699
  - antidepressant-associated suicidal behavior and, 1673
  - atypical antipsychotics, 1683*t*
  - BNDF polymorphisms in, 102
  - cancer patients with, 2263
  - candidate genes for mood disorder and, 1625, 1626*t*
  - carbamazepine, 1683*t*
  - categories, 1648–1652
  - classification of, 1158
  - clock genes and, 213–214
  - cognitive behavioral treatments for, 1665
  - cognitive therapy, 2765
  - with comorbidity, treatment, 1698–1699
  - coronary artery disease and, 2189
  - definitions, 1600–1601, 1604
  - delirium incidence in, 1185
  - depressive disorders and, 1642

- description of, 4565
- descriptive epidemiology of, 1614
- diabetes and, 2247
- diagnosing, 1682–1683
- differential diagnoses, 3610–3611
- differentiation of, 4370
- dissociative identity disorder and, 1916
- E-EPA for, 116
- early-onset, **3686–3694, 3700–3708**
- encopresis and, 3659, 3673
- enuresis, 3652, 3666
- environment and, 2529
- functional capacity/performance of, 4371
- gambling and, 1804, 1805
- gender differences, 4247
- genetic basis of, 315–316
- genetics and, 1628, 2529, 2529*t*
- hallucinogen-induced, 1326
- heritability, 1610, 2535*t*
- and HIV infection, 4271
- in hypothyroid patients, 680
- hypothyroidism and, 2241
- immune-CNS interactions in, 199
- insomnia and, 2086
- interpersonal social rhythms therapy, 1666
- lithium for, 215, 1683*t*, 1684
- MDD *versus,* 1654–1655
- medication costs, 4522, 4524
- mortality rate with, 4529
- MRI in, 274
- neuroimaging, 3464–3466
- parenting children with, 3929–3931
- in patients with HIV, 572–575
- polycystic ovary syndrome and, 2248–2249
- PPD and, 2516
- presentation, 572–573
- psychopharmacology, 3791–3792, 3796

- psychotherapy, 1703, 1706–1708
  - course of therapy, 1707–1708
  - initial assessment, 1706
- rapid-cycling, 1652, 3095–3096
- recent diagnostic trends, 1631
- as recurrent illness, 1682
- risk-taking
  - premature births and, 643
- SAMe and, 2549
- schizophrenia heritability and, 1446–1447
- selective advantage, 1623–1624
- sleep deprivation and, 353
- stroke lesions and, 486*f*
- substance use disorders and, 1699
- suicide and, 2612, 2613
  - factors associated with, 1685*t*
  - ideation monitoring, 1684
  - risk, antidepressants and, 2616
- and telemedicine, 4276–4277
- temperament and, 3486
- treatment
  - adherence evaluation, 1683–1684
  - anticonvulsants, 2968, 2970, 2972–2973, 2974, 2976
  - antipsychotics, 1686
  - atypical antipsychotics, 1686–1687, 1686*t*
  - benzodiazepine, 2997
  - calcium channel inhibitors, 3015–3016, 3019–3020
  - carbamazepine, 3016, 3027, 3028*f*, 3031*t*, 3036–3037, 3038*t*
  - clinical recommendations, 1687–1690, 1689*t*
  - combination pharmacotherapy in, 3325–3327
  - FGAs, 3070
  - lamotrigine, 3016, 3084, 3085
  - lithium, 3015–3016, 3015*f*, 3089–3092, 3089*f*, 3091*f*, 3092*f*, 3092*t*
  - monitoring, 1683*t*
  - to remission, 1682
  - thyroid hormones, 3245–3246, 3245*t*, 3246*t*
  - valproate, 3270, 3271*t*, 3276

- tryptophan depletion and, 2550
- type I
  - acute mania, 1648–1649
  - bipolar mixed phase, 1649–1650
  - brain imaging, 1717
  - charting life course of, 1648, 1649*f*
  - chronic mania, 1649
  - cognition in, 698–699
  - definition, 1601
  - depression, antidepressants and, 1691
  - depressive phase, 1650
  - description of, 4565
  - management, 1695–1697
    - foundations of, 1682–1684, 1683*t*
  - manic and mixed episodes as, 1639
  - secondary mania, 1649
- type II
  - alcohol and substance abuse, 1651–1652
  - course, 1650–1651
  - definition, 1601
  - description of, 4565
  - hypomania, 1651
  - hysteroid dysphoria, 1652
  - leadership and creativity, 1652
  - maintenance treatment, 1698
  - MDD diagnosis and, 1644
  - rapid-cycling, 1652
  - recurrent brief hypomania, 1652
  - seasonality, 1651
  - temperament and, 1651
  - treatment, 1697–1698
  - ultrarapid cycling, 1652
- types of, 3689, 3703
- valproate, 1683*t*
- warning signs, 1707–1708, 1708*t*

- Bipolar mania, treatment
  - lithium, 3089

- Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP), 1513–1518, 1514*f*, 1515*f*, 1517*f*, 1518*f*
- Bipolarity, definition of, 1651
- Birchwood Social Functioning Scale (SFS), 4372
- Birnbaum, Morton, 4336
- Birth complications, aggression and, 2484
- Birth histories, review of, 454
- Birth order, 910
- Birth trauma
  - description of, 4565
  - Rankiaan dialectic, 918
- BIS. *See* Brief Impairment Scale
- Bisexual, definition of, 4565
- Bisexual persons
  - clinical issues, 2006–2009
  - Internet resources for, 2012*t*
- Bisexuality. *See also* Androgeny
  - definition of, 1982, 4565
  - terminology, 1983
- Bismuth germinate (BGO), 287
- Bizarre delusions, 1414–1415, 4565, 4573
- Bizarre objects, psychosis and, 2875
- Black, Stephen, 2729
- Black box warnings
  - anticonvulsants, 2966
  - antidepressants in children/adolescents, 1779–1780
    - suicidal behavior and, 1673
  - aripiprazole, 3240
  - clozapine, 2919
  - disulfiram, 3052
  - lamotrigine, 3085
  - methadone, 3134
  - nefazodone, 2919, 3125, 3128, 3129
  - paroxetine, 2493
  - psychopharmacology regulatory issues, 2916*t*, 2918–2919
  - risperidone, 2919
  - SGAs, 3080
  - SSRIs in children and adolescents, 1680, 1784

- valproate, 3276, 3277*t*
- Black-pitch syndrome, 4565
- Blackouts from alcohol, 1267
- Bladder cancer, psychological symptoms, 2268
- Blank screen, definition of, 4565
- Blanton, Smiley, 2641
- Blast exposure, brain injury and, 523
- Bleeding, SSRI effects on, 3157
- Bleomycin (Blenoxane), 2255
- Blessings, counting, 2851, 2852
- Bleuler, Eugen, 1407, 1410, 1418, 1426, 1549, 1583
  - *Dementia Praecox or the Group of Schizophrenias,* 4537
  - on schizophrenia, 1869, 1870, 2868–2869
  - on somatic phenomena, 1830
  - work of, 4565
- Blind spot, definition of, 4565
- Blindness
  - and consciousness, 691
  - phase shifts and, 208
- Blink rate, hypodopaminergia and, 456
- Blistering agents, terrorism and, 2572*t*
- Bloch, Felix, 259
- Blocking, definition of, 4565
- Blood, metabolites, OCD and, 1787
- Blood cells, lineages, 179*f*
- Blood flow, imaging of, 271
- Blood levels, description of, 4565
- Blood oxygenation level dependent (BOLD) imaging
  - brain dynamics study, 436
  - in schizophrenia, 273, 1475–1476, 1477*f*
- Blood pressure, stress and, 2316*f*
- Blood vessels, aging brain, 3990, 3992
- Blood–brain barrier (BBB)
  - delirium pathophysiology and, 1181–1182
  - description of, 4565
  - formation of, 3
  - function of, 14

- radiotracers and, 297
- Blood–cerebrospinal fluid barrier, 14
- Blueprints for Violence Prevention, 4296
- Blum, Harold, 2645
- Blunted affect
  - description of, 4565
  - schizophrenia and, 1417
- "Blurting," characteristics of, 457*f*
- *BMALI* gene, 215
- Board-certified psychiatrists, 4565
- Board-eligible psychiatrists, 4565
- Body alienation, depersonalization and, 1891
- *Body Build and Character* (Kretschmer), 4538
- Body dysmorphic disorder (BDD), 1790–1792
  - classification, 1786
  - clinical features, 2353
  - combination therapy, 1792
  - comparative nosology, 1790
  - course and prognosis, 1792
  - definition of, 1790, 2028, 2349
  - description of, 4565
  - diagnostic and clinical features, 1791
  - differential diagnoses, 1586, 1791–1792
  - DSM-5 on, 1785
  - epidemiology, 1790
  - etiology, 1790–1791
  - excoriation disorder and, 1796
  - history, 1790
  - OCD *versus,* 1788
  - pathology and laboratory features, 1791
  - pharmacotherapy, 1792
  - presentation, 2353
  - psychotherapy, 1792
  - somatic symptom disorders and, 1827
  - treatment, 1792, 2354
- Body focused repetitive behavior disorders (BFRBDs), 4565. *See also* Excoriation disorder; Trichotillomania
  - about, 1794

- habit-reversal training and, 2698
- Body image
  - description of, 4565
  - distortions of, 1127–1128
  - facititious dermatitis and, 2354
  - sexual dysfunction and, 1974
- Body language, 4565
- Body mass index (BMI)
  - obesity and, 2211, 2211*t*
  - role in treatment selection, 1022
- Body modification and mortification, 2582, 2582*f*
- Body piercing during adolescence, 3372–3373
- Body psychotherapy, 2563
- Body touching, BDD and, 1792
- BOLD. *See* Blood oxygenation level dependent (BOLD) imaging
- Boldenone undecylenate (Equipoise), 2493
- Boltzmann distribution, 260
- Bonaparte, Princess Marie, 2645
- Bonding, 4566. *See also* Attachment
- Bone marrow aspirations, 2725
- Bone marrow transplantation (BMT), 659, 2274
- Bone mineral density (BMD), anorexia and, 2075
- Bone morphogenetic proteins (BMPs), 41, 48
- Bonferroni method, 833, 834
- Boole's inequality, 833
- Boon, Suzette, 1871
- Borderline intellectual and adaptive functioning (BIF), **2430–2442**
  - clinical features of, 2434*f*
  - clinical significance of, 2432–2433
  - course/prognosis of, 2437
  - cultural sensitivity to, 2433
  - definition of, 2432, 2433–2434
  - description of, 4566
  - differential diagnoses, 2345–2437, 2436*t*
  - epidemiology, 2434
  - etiology, 2434
  - historical review, 2434

- nosology, 2434
- presentation, 2434–2345
- treatment of, 2438–2439
- under-reporting of, 2345–2437

- Borderline personality diagnosis, 4552
- Borderline personality disorder (BPD), 4106
  - aggression and, 2494
    - treatment of, 2501–2502
  - biosocial theory of, 2786–2787, 2786*f*
  - bipolar II disorder and, 1651
  - child/adolescent, 3767
  - clinical criteria, 2155
  - comorbidity, 2155
  - complications, 2155
  - course of, 2156
  - DDIS diagnosis, disassociation and, 1905–1906
  - definition, 1601
  - description of, 4566
  - diagnostic criteria, 2787*t*
  - dialectical behavior therapy and, 2055, 2785
  - differential diagnoses, 2156
  - dimensional mentalizing profile and, 2896
  - dissociative identity disorder and, 1916
  - DSM and ICD diagnostic criteria, 1874
  - emotional dysregulation and, 2786*f*, 2787
  - emotional vulnerability and, 2786–2787, 2786*f*
  - epidemiology, 2156
  - factitious disorder imposed on self and, 1851, 1859, 1864
  - familial patterns, 2156
  - gambling and, 1804
  - genetics, 2156
  - group therapy and, 3767
  - impairment, 2156
  - invalidating environment and, 2786, 2786*f*
  - MBT, 2902–2904
  - mentalizing and, 2894–2895
  - mood disorders and, 1655, 1655*t*

- nonmentalizing modes, 2896–2897
  - pretend, 2896, 2896*t*, 2897
  - psychic equivalence, 2896, 2896*t*, 2897
  - teleological, 2896, 2897, 2897*t*
- non–possession-form dissociative identity disorder and, 1905
- parental neglect and, 1713
- pathological dissociation and, 1885
- predisposing factors, 2155
- as psychosis manifestation, 1572–1573
- recurrent brief depressive disorder and, 1647
- sex ratio, 2156
- treatment
  - α-adrenergic receptor agonists, 2949
  - FGA, 3070
  - lithium, 3093
  - MAOIs, 3121
  - valproate, 3275–3276

- Borderline psychosis, 4566
- Borderline states, 4566
- Borna disease virus (BDV), 598–599
- Boron, carbon decay to, 286*f*
- *Borrelia burgdorferi,* 589–592, 632
- *Borrelia miyamotoi,* 593
- Boss, Medaard, 2644
- Boston Diagnostic Aphasia Examination (BDAE), 991
- Boston Naming Test, 739
- Boswell, John, 1985
- Botanicals
  - *Ginkgo biloba,* 2552–2553
  - ginseng, 2553
  - golden root, 2554
  - huperzine-A, 2553–2554
  - kava, 2553, 2553*f*
  - saffron, 2554
  - St. John's wort, 2552, 2552*f*
  - valerian, 2553
- Botulinum toxin

- ACh release and, 73
- tic disorders, 3647, 3661
- war or terrorism and, 2572*t*

- Boufée délirante, 4566
- Boundaries
  - email and, 2672
  - ethical practice of psychiatrists, 4445–4446
  - ethics of, 3873–3874
  - violations
    - dissociative identity disorder treatment and, 1923
    - early psychoanalysts, 2641, 2645
    - psychoanalytic psychotherapy, 2669–2670
  - violations, and psychiatric malpractice, 4437
- Bovine spongiform encephalopathy (BSE, mad cow disease), 596–597, 601*t*
- Bowen, Murray, 2749, 2754
- Bowen family systems therapy, 2751*t*, 2753
- *Bowers v. Hardwick,* 2000
- Bowker, John, 2588
- Bowlby, John, 3662, 3676
  - *Attachment,* 4543
  - on attachment, 765, 1666, 2721
  - family therapy and, 2749
  - on social development, 3360
  - work of, 4566
- BPD. *See* Borderline personality disorder (BPD)
- BPRS. *See* Brief Psychiatric Rating Scale
- Braak hypothesis, 542
- *Brad H. Case, New York City v.,* 4461–4462
- Bradley, Charles, 4537
- Bradycardia, 2190
- Bradykinesia
  - description of, 4566
  - medication-induced, 2938
  - in PD, 544
- Bradylexia, 4566
- Brahmyadiyoga, 2554
- Braid, James

- on hypnotism, 2639, 2706
- *Neurypnology, or the Rationale of Nervous Sleep,* 4535

- BRAIN. *See* Brain Research through Advancing Innovative Neurotechnologies (BRAIN)
- Brain. *See also* Blood-brain barrier (BBB); Brain stimulation; Brain tumors; Forebrain; Organic brain syndrome (OBS); Traumatic brain injury (TBI); *specific regions*
  - abnormalities, 3546
  - aging, 3989–3991
  - amygdala, 8
  - anorexia nervosa and, 2075
  - biopsy of, 468
  - bipolar disorder, child/adolescent, 3688, 3702
  - cerebral cortex
    - caudate nucleus, 8
    - globus pallidus, 8
    - putamen, 8
  - circuitry of, 1161
  - connectivity, 276–277
  - coronal section, 10*f*
  - cortical organization of memory, 729–730, 730*f*
  - cytoarchitectonic subdivisions, 20*f*
  - cytokine effects of, 194–195
  - damage
    - following TBI, 535–536
    - primary, 523
    - secondary, 523
  - description of, 4566
  - development (*See also* Neurodevelopment)
    - adolescent, 3370–3371
    - in childhood, 3351, 3356–3357*t*, 3358
    - gender differences, 2418, 3462–3463, 3463*t*
    - genomics of, 438–447
    - gonadotropins and, 2505
    - gray matter, 3361*f*, 3461–3462, 3462*f*, 3462*t*
    - neuroimaging of, **3461–3463**
    - in preterm infants, 640
    - white matter, 3462, 3462*f*, 3463*f*

- distinctiveness of, 8
- dopamine receptors in, 66
- drawing of
  - coronal section, 9*f*
  - ventricles, 15*f*
- embryology, 9*f*
- emotional processing and, 1711–1712
- energy and information, 688–689, 689t
- freezing/fixation, 8
- frontosubcortical circuits, 449–450
- function
  - brain stimulation and, 3302
  - and mental health, 762
- functional connectivity, 417–427
  - analysis of, 418–422, 420*f*
  - association *versus* directionality, 421
  - default mode, 423, 425–426, 425*f*
  - independent components analysis, 421
  - networks, 418–419
  - neural approaches to, 421–422
  - resting state, 418–419, 418*f*
  - static *versus* dynamic assessments, 421
  - study of, 418
  - thalamic connectivity, 422–423
- functional imaging, 417
- functional systems of, **16–38**
- hemispheres, and neuropsychological deficits, 978–982, 978*t*
  - arithmetic, 979
  - executive function, 981–982
  - language disorders, 978–979
  - limb apraxia, 979
  - memory disorders, 981, 981*f*
  - motor skills, 982
  - spatial disorders, 979–980
- hippocampus, 8
- human
  - coronal section, 9*f*, 11*f*, 12*f*

- cross-section, 9*f*
- distinctiveness of, 8
- limitations on study of, 8
- imaging, 4566
  - role in diagnosis, 1018–1019
  - structural and functional, 1717
- inhalant use and, 1333
- injury to
  - in preterm infants, 640–642
  - psychopathology and, 644
- language disorders and, 3546
- large-scale dynamics, 436
- laterality, and consciousness, 692
- lateralization, 449
- macaque monkey, 35*f*
- major structures of, 8–10
- medial temporal lobe, 6*f*
- medial view, photograph of, 14*f*
- memory systems, 690–691
- metabolism of, 4566
- mild traumatic injury, 985, 985*f*
- modulators of states, 452
- morality and, 2415–2416
- myeloarchitecture, 5
- neural growth and integration of, 687–688
- neuroimaging (*See* Neuroimaging)
- neuropsychiatric disorders (*See* Neuropsychiatric disorders)
- OCD, child/adolescent, 3697
- organization of, 1–8
- postmortem interval, 8
- in posttraumatic stress disorder, 701
- psychotherapy and modulation of, 2812–2813
- regional differentiation, 41*f*
- regions, connections between, 6–8
- restitution, sleep and, 352–353
- reward centers, 375
- structural components, **8–16**

- structure and function studies
  - in mood disorders, 4070–4071
- studies, 8
- stuttering, 3560–3561
- surface landmarks, 7*f*
- as system, 162–163
- temperament and, 3483–3484
- thickness analysis, 2424*f*
- volume of, 2075
- weight and volume, 3989–3991, 3989*t*
- weight loss in HD, 554

- Brain abscess, description of, 4566
- Brain death determination, legal issues, 2305
- Brain-derived neurotrophic factor (BDNF)
  - addiction and, 379–380
  - addictive disorders and, 2046
  - brain imaging, 1717
  - candidate gene association studies
    - anxiety disorders, 1739
    - mood disorders, 1625
  - depression and animal modes of stress and, 1713
  - depressive behaviors, 397
  - drug response in mood disorder and, 1629
  - GAD and, 1764
  - gene encoding, 2046
  - gene-environment interactions, 1739–1740
  - hypothesis on bipolar disorder origin, 1608, 1610
  - mouse studies of, 100
  - psychiatric disorders and, 102
  - PTSD and, 1820, 1823
- Brain electrical activity mapping (BEAM), 4566
- "Brain fag," description of, 2443–2444
- Brain injury medicine. *See also* Traumatic brain injury (TBI)
  - examinations, 4427
  - training, 4405, 4405*t*
- Brain mapping, neurometric, 2545
- Brain Overclaim Syndrome, 4486

- Brain proteome, 162
- Brain Research through Advancing Innovative Neurotechnologies (BRAIN), 4482
- Brain state vectors, 1031–1032
- Brain stimulation, **3299–3313**
  - acute *v.* prolonged effects, 3301, 3301*t*
  - brain development and, 3302
  - definitions of methods, 3299*t*
  - electrical, common pathways for, 3300, 3300*f*
  - focality and invasiveness in, 3302–3303, 3302*f*
  - history, 3299–3300
  - methods, **3299–3313**
  - neuroimaging and, 3303–3304
  - neurophysiology and, 3303–3304
  - neuroplasticity and, 3301–3302
  - psychiatric applications, 3303–3304, 3303*f*
  - psychiatric disorders treated with, 3300–3303
  - psychopharmacology *v.,* 3300–3301, 3301*t*
  - tools, 3304–3313
    - CES as, 3308–3309
    - ICS as, 3312, 3312*f*
    - magnetic seizure therapy as, 3309–3310
    - tDCS as, 3307–3308
    - TMS as, 3304–3307
    - TNS, 3299*t*, 3312–3313
    - VNS as, 3310–3312, 3311*t*
- Brain syndrome. *See* Organic mental disorder
- Brain tumors
  - anxiety and, 500–501
  - in children, 504–506
  - corticosteroids in, 501–502
  - depressive disorders due to, 497–500
  - EEG in, 244
  - epidemiology of, 490–491, 490*f*
  - investigations, 494–495
  - location of
    - presentation and, 2266–2267
    - psychiatric symptoms and, 495

- malignancy of, 495
  - psychiatric symptoms and, 495
  - psychiatric syndromes and, 2266–2267
- management of, 491–492
- neuropsychiatry of, **489–506**
- premorbid factors, 495–496
- in psychiatric patients, 492–495
  - red flags, 493–494
- psychiatric symptoms in, **495–506,** 2267
- psychosocial factors, 495–496
- survivors of, 504–506
- treatment
  - psychiatric symptoms and, 495
- useful websites, 494*t*

- Brain waves, description of, 4566
- Brain–computer interface technologies, 2558
- BRAINEAC project, 3987
- Brainstem
  - afferents, 26
  - aggression and, 2476
  - appetite and, 363–364
  - description of, 4566
  - dysfunction, depression and, 1709
  - glycine neurotransmission, 82
  - hypnosis, 2712
  - noradrenergic neurons in, 1713–1714
  - pathology after TBI, 525
  - PTSD and, 1819
  - sexuality and, 1957
- Brainwashing, 4566. *See also* Coercive persuasion
- Branched-chain amino acids (BCAA), 661–662
- Brand, Bethany, 1871, 1919
- Brawner decision. *See* Model Penal Code
- Bread mold *(Neurospora crassa)*, 149
- Breakthrough pain in burn injuries, 2376
- Breast augmentation surgery, 2023
- Breast cancer

- incidence of, 2251
- psychological issues in, 2269
- risk of, 214, 2269
- schizophrenia and, 1541
- STDP and, 2811
- tamoxifen for, 171

- Breast-feeding. *See also* Lactation
  - acamprosate and, 3058
  - α-adrenergic receptor agonists and, 2949
  - β-adrenergic receptor antagonists and, 2953
  - anticonvulsants and, 2969, 2971, 2973, 2975, 2976, 2978
  - barbiturates and, 2988
  - CAM treatments and, 2543
  - carbamazepine and, 3037
  - lamotrigine and, 3086
  - lithium and, 3095
  - neuroendocrine cascade, 2517–2518
  - SSRIs and, 3168
  - trazodone and, 3252
  - tricyclic and tetracyclic antidepressants and, 3263
  - valproate and, 3277
- Breasts, sexual response cycle, 1959*t*
- Breathing
  - retraining
    - CBT for panic disorder and, 1775
    - PTSD, 1824
  - sustained deep, 2557
- Breathing-related sleep disorder, 4566
- Breathing therapy for insomnia, 2089–2090
- Bremelanotide, 3340
- Bremner, J. Douglas, 2718
- Brenman, Margaret, 1869
- Brennan Center for Justice, 2473–2474
- Brenner, Charles, 2650, 2654–2655, 2659
- Bretazenil, 3003
- Brett, Elizabeth, 2718
- Breuer, Joseph

- Anna O., 1839, 2639
- cathartic method and, 2639
- dissociative disorders, 1868
- Freud and, 851–853, 2638
- work of, 4566

- Brevital. *See* Methohexital
- Brief adaptive psychotherapy (BAP), 2800
- Brief Assessment of Cognition in Schizophrenia (BACS), 4372
- Brief Fatigue Inventory, 2284
- Brief Impairment Scale (BIS), 822, 3402
- Brief interventions, cannabis dependence and, 1310
- Brief mood disorders, 1601
- Brief motivational enhancement therapy (MET), 3057
- Brief Pain Inventory, 2376
- Brief Psychiatric Rating Scale (BPRS), 2686
  - for aging adults, 3976
  - behavior assessment using, 1006–1007
  - medication dose determination with, 4367–4368, 4370
  - schizophrenia, 1409
- Brief psychodynamic psychotherapy. *See* Intensive short-term dynamic psychotherapies
- Brief psychotherapy, 4544, 4566
- Brief psychotic disorder, **1574–1580**
  - comparative nosology, 1574–1577, 1576*t*, 1577*t*
  - course and prognosis, 1580
  - definition, 1948
  - description of, 4566
  - diagnosis and clinical features of, 1579
  - differential diagnoses, 1579–1580
  - epidemiology, 1577
  - etiology, 1577–1578
  - history of, 1574, 1575*t*
- Brief reactive psychosis, 1948, 4566
- Brief Visuospatial Memory Test, 991
- Brigham, Amariah, 2322, 4566
- Brill, A. A., 4566
- Bringing Postpartum Depression Out of the Shadows Act, 2517
- Brintellix. *See* Vortioxetine

- Briquet, Pierre, 1830, 1835, 1839, 1866–1867
- Briquet syndrome, 1830, 4566
- Broad affect, 4566
- Broad phenotype bipolar disorder, 3693, 3707
- Broca, Paul, 31
- Broca aphasia, 462*t*
- Broca's area (inferior frontal gyrus), 699
  - BA44
    - activation of, 414
  - lesions in, 7
- Brodmann, Korbinian, 20
- Brodmann map of cerebral cortex, 977, 977*f*
- Brodmann's areas
  - area 24, 33
  - area 25, 33
  - area 32, 33, 96
  - BA17, 452
  - BA18, 452
  - BA19, 452
  - Nissl-stained, 18*f*
- Brody, Arthur L., 2813
- Brofaromine (Consonar), 3113*f*, 3119
- Broken-heart syndrome, 2194
- Brom, Daniel, 2718
- Bromazepam
  - for anxiety, 548
  - in GAD, 1783
  - in PD, 548
  - in SAD, 1784
- Bromocriptine (Parlodel), 2510
  - after TBI, 529
  - chemistry, 3234, 3235*f*
  - dosage/administration, 3241, 3241*t*
  - drug interactions, 3240
  - history, 3234
  - for hyperprolactinemia, 2250
  - NMS, 3238

- organ/system effects of, 3239
- pharmacological actions, 3235, 3236
- precautions/adverse reactions, 3239, 3240*t*
- for sexual dysfunction, 1979, 1980
- therapeutic indications, 3239

- Bromodomain PHD finger transcription factor (BPTF), 390
- Bronchodilators, anxiety induced by, 2258
- Bronkodyl. *See* Theophylline
- Brooding compulsion. *See* Intellectualization
- Brooks, Peter, 2833
- Brown, Daniel
  - on chronic pain, 2726
  - on hypnotherapy, 2716, 2721
    - memory and amnesia symptoms, 2732
    - for treating smoking, 2730
    - for weight management, 2731
  - on hypnotic imagery, 2723
  - on phase-oriented trauma treatment, 2719
- Brown, Richard S., 1844
- Brown Assessment of Beliefs Scale, 1791
- *Brown v. Plata* (2011), 4400
- Brucellosis, 2572*t*
- Bruch, Hilda, 2775, 3386–3387
- Brugada syndrome, 2190
- Bruininks–Oseretsky Test of Motor Proficiency (BOTMP), 3539–3540
- Brunner syndrome, 2423
- Bruxism
  - description of, 4566
  - in psychiatric disorders, 1149
  - sleep-related, 2103–2104, 2105*f*
  - stereotypic movement disorders and, 3668, 3682
- Buchanan, Robert, 1418
- Bucharest Early Intervention Project (BEIP), 3819
- Bucy, Paul, 31
- Budbill, David, 2834
- Buddhism
  - angels and, 2590–2591

- suffering and, 2588
- Bulbar dysarthria, 457*f*
- Bulimia nervosa
  - behavioral studies, 367–368
  - classification of, 1168, 1174
  - cognitive therapy, 2766
  - compensatory behaviors, 2073
  - course/prognosis in, 2078
  - description of, 2065–2066, 4566
  - diagnosis of, 2066*t*, 2072
  - differential diagnoses, 2076
  - epidemiology of, 2070
  - etiology of, 2071
    - biological factors, 2071
    - developmental factors, 2071
    - environmental factors, 2071
    - genetic factors, 2071
    - psychological factors, 2071
    - social factors, 2071
  - genetics and, 2529*t*
  - historical literature on, 2067
  - laboratory exams, 2075
  - mortality rates, 2073
  - nosology, 2068
  - pathology, 2075
  - satiation in, 368
  - sham eating in, 359
  - temperament dimensions of, 3487
  - trazodone drug studies, 3249
  - treatment, 2081
    - anticonvulsants, 2970
    - MAOIs, 3121–3122
- Bundled payment systems, 4316, *4317f*
- Buprenex. *See* Buprenorphine
- Buprenorphine (Buprenex)
  - administration of, 2293*t*
  - adverse effects, 3136

- chemistry, 3135, 3135*f*
- dosing, 3136–3137
- drug interactions, 3136
- drug studies, 3136
- duration of action, 2395*t*
- history, 3135
- opioid-dependence treatment, 1371–1372
- opioid maintenance treatment, 3132, 3133, 3135–3137
- opioid withdrawal, 1368–1369, 3132
- pain management, 3137
- pharmacological actions, 3135
- pharmacology of, 1366–1367
- sublingual, 580
- substance use disorders, child/adolescent, 3793

- Buprenorphine/naloxone (bup/nx), 582
- Bupropion (Wellbutrin, Zyban), **3004–3008**
  - ADHD treatment, 3004, 3006, 3595
  - adverse effects of
    - dermatological, 2356*t*
  - for amphetamine-type stimulant use disorders, 1290
  - anxiety disorder treatment, 1743*t*, 1779, 1780, 1783
  - bipolar depression, 1691
    - with ADHD, 1699
    - combination therapy, 1692*t*, 1693
  - BPCA and, 2931
  - buspirone and, 3010
  - cardiovascular effects of, 2197*f*
  - child/adolescent
    - ADHD treatment, 3789
    - major depressive disorder, 3789
  - citalopram and, 3005, 3010
  - clinical uses, 3005–3006
  - combination pharmacotherapy, 3005, 3006, 3324, 3325
  - contraindications, 533, 2371, 2520
  - depression, 1679
  - depression and smoking cessation, 1680
  - for depression in cardiac patients, 2196

- depression treatment, 4125
- depressive disorder, 3005
- dosages, 573*t*
- dosing, 3007
- drug interactions, 648, 2199, 3007
- drug studies, 3005–3006
- eating disorders, 3006
- ECT interactions with, 3295
- formulations, 3007
- gambling disorder, 1808
- for geriatric depression, 4073–4074
- history, 3004
- mirtazapine and, 3109
- naltrexone and, 3006, 3139, 3142, 3144
- obesity, 3004, 3006
- off-label uses, 3004
- overdose, 3007
- pharmacology, 3004–3005
  - pharmacodynamics, 3004
  - pharmacokinetics, 3004–3005
- precautions/adverse reactions, 3006–3007
- REM sleep and, 1717
- SAD, 3004, 3005
- for sexual dysfunction, 4261
- sexual dysfunction and, 1779*t*
- side effects, 573*t*, 1679*t*, 3006–3007
  - seizures, 2261
- smoking cessation, 3005–3006
- tobacco use disorder treatment, 1349, 1350, 1545
- weight gain and, 2215, 2237

- Burden of disease. *See* Global Burden of Disease
- Bureau of Primary Healthcare, 4322
- Burn injuries
  - acute phase, 2378–2379
  - anger in, 2381
  - in children, 2384–2386
    - psychological sequelae, 2385

    - risk factors, 2384*f*
  - delirium in, 2375–2376
  - end-of-life considerations, 2386–2387
  - ethical considerations, 2386–2387
  - face grafts after, 2381–2382
  - family support for patients, 2382–2383
  - grief in, 2380
  - hypnosis in, 2724, 2729
  - intermediate phase, 2379
  - pain management in, 2376–2377
  - plastic surgery after, 2381–2382
  - psychiatric care of patients with, 2374
  - psychopathology and, 2378
  - psychopharmacological issues, 2381
  - psychotherapeutic interventions, 2378–2379
  - recovery phase
    - factors influencing, 2380
    - psychotherapeutic interventions, 2379–2380
  - regression in, 2381
  - risk factors for, 2374
  - self-inflicted, 2383
    - risk factors for, 2383
  - therapeutic approaches to, 2380, 2381
  - withdrawal of care, 2386–2387
- Burned-out schizophrenic, 4566. *See also* Anergic schizophrenic
- Burning mouth syndrome (BMS), 2392
- Burnout
  - caregiver, 4209
  - clinicians treating dissociative identity disorder patients and, 1925
  - description of, 4566
  - in jails and prisons, 4464–4465
  - physician mental health and, 2599
- Buronil. *See* Melperone
- Burrow, Trigant, 2737
- Bursts, definition of, 143
- Burton, Richard, 4534
- Burton, Robert, 1603, 1603*f*, 1838

- Bush, George H. W., 4311
- Bush, George W., 4294
- BuSpar. *See* Buspirone
- Buspirone (BuSpar), **3008–3014**
  - absorption, 3008
  - ADHD, 3012
  - adverse reactions, 3013
  - aggression in TBI, 2502
  - agitation, 3011–3012
  - alcohol and cannabis dependence, 3010
  - for anxiety, 548
  - anxiety disorders, 1746, 1782*t*, 3010, 4131
  - BDD, 1792
  - blood concentration, 3009
  - bupropion and, 3010
  - cholinergic system, 3009
  - citalopram and, 3009, 3010
  - clinical trials, 3009–3012, 3011*t*
    - ADHD, 3012
    - agitation, 3011–3012
    - alcohol and cannabis dependence, 3010
    - anxiety disorders, 3010
    - GAD, 3009–3010
    - MDD, 3010
    - PMS, 3012
    - psychosis, 3012
    - sexual dysfunction, 3012
    - sleep apnea, 3012
    - smoking cessation, 3010–3011
    - somatic dysfunction, 3012
    - substance withdrawal, 3012
  - combination pharmacotherapy and, 3325
  - depression, 1680
  - distribution, 3008
  - dopaminergic system, 3009
  - dosage/administration, 3013
  - drug interactions, 3013

- in elderly patients, 2284
- elimination, 3008–3009
- GABAergic system, 3009
- GAD, 1781, 1783, 3009–3010
- history, 3008
- intellectual disability and, 3515
- MDD, 3010
- mechanisms of action, 3009
- metabolism, 3008–3009
- norepinephrine system, 3009
- organ/system effects of, 3012–3013
- in PD, 548
- pharmacological actions, 3008–3009
- PMS, 3012
- precautions, 3013
- psychosis, 3012
- SAD, 1784
- serotonin system, 3009
- sexual dysfunction and, 1779*t*, 3012
- sleep apnea, 3012
- smoking cessation, 3010–3011
- somatic dysfunction, 3012
- substance withdrawal, 3012
- therapeutic indications, 3013

- Butabarbital (Butisol), 2986
- Butalbital, 1383
- Butane, 1330
- Butisol. *See* Butabarbital
- Butorphanol, 2395*t*
- Buttercup family toxins, 136
- Buttock pain, conditions causing, 2390*t*
- Butyl nitrite-related disorders, 1341
- Butyrophenones
  - chemistry, 3063*f*, 3064
  - description of, 4566
  - history, 2917

C

- C (consequences), in A-B-C-model, 2696
- C-fibers, pain transmission and, 2388–2389, 2398
- c-Fos activation, 397
- c-Jun N-terminal kinases (JNKs), 125
- C-reactive protein (CRP)
  - cardiovascular disease and, 2321
  - depression and, 199
- C.A. *See* Chronological age
- CAA. *See* Cerebral amyloid angiopathy
- Cabergoline (Dostinex, Caberlin)
  - history, 3234
  - for hyperprolactinemia, 2250
  - organ/system effects of, 3239
  - therapeutic indications, 3239
- Caberlin. *See* Cabergoline
- Cachexia, terminal anorexia and, 2296
- Cacodemonomania, 4566
- CAD. *See* Coronary artery disease (CAD)
- Caffeine
  - adverse effects, 2547
  - CNS effects, 1295
  - description of, 4566
  - EEG effects, 254
  - with paracetamol, 625–626
  - sleep and, 346–347
  - for tension headaches, 625–626
- Caffeine-related disorders, **1291–1303**
  - anxiety disorder, caffeine induced, 1297–1298, 1298*t*
  - caffeine use disorder, 1298*t*, 1301–1302
    - comorbidity, 1301
    - course and prognosis, 1301
    - diagnosis and clinical features, 1301
    - differential diagnoses, 1301
    - epidemiology, 1301
    - nonpsychiatric illnesses and, 1302
    - treatment, 1298*t*, 1302

- CNS effects, 1295
- comparative nosology, 1292
- defined, 1291, 1291*t*
- epidemiology, 1292, 1293–1294*t*, 1294*f*, 1295
- etiology, 1295–1296, 1295*f*
- genetics and, 1296
- history, 1291–1292
- intoxication, 1296–1297, 1297*t*, 1298*t*
  - diagnosis and clinical features, 1297
  - differential diagnoses, 1297
  - epidemiology, 1296–1297
  - pathology and laboratory examination, 1297
  - treatment, 1297, 1298*t*
- performance effects, 1295
- pharmacology of, 1295
- reinforcer effects, 1295–1296
- sleep disorder, caffeine-induced, 1298–1299, 1298*t*
- subjective effects, 1295
- unspecified, 1299
- withdrawal, 1299–1301, 1300*t*
  - comorbidity, 1299
  - course and prognosis, 1300
  - diagnosis and clinical features, 1299–1300, 1300*t*
  - differential diagnoses, 1300
  - epidemiology, 1299
  - treatment, 1300, 1398*t*

- CAIM. *See* Complementary, alternative, and integrative medicine (CAIM)
- Cajal-Retzius cells, 17, 55
- Calbindin, function of, 17
- Calcarine fissure, 7*f*
- Calcineurin, activation of, 124
- Calcineurin inhibitors (CNIs), 2371
- Calcitonin (Cibacalcin), 1687, 3018
- Calcitonin gene-related peptide (CGRP), 2337
- Calcium
  - metabolism, bipolar disorders and, 1609
  - mild anxiety and, 1782

- Calcium/calmodulin-dependent (CaM) kinases
  - activation of, 122
- Calcium channel blockers (CCBs), 3016–3021
  - for migraine treatment, 624
  - in pain syndromes, 2396
- Calcium channel inhibitors, **3014–3022**
  - absorption, 3016–3017
  - BD, 3015–3016, 3019–3020
  - chemistry, 3016, 3016*f*, 3017*t*, 3018*f*
  - distribution, 3017, 3017*t*
  - dosage/administration, 3021
  - drug interactions, 3021
  - drug studies, 3018–3020
  - elimination, 3017
  - history, 3014–3016, 3014*f*, 3015*f*
  - investigational drugs, 3021
  - laboratory interferences, 3021
  - mechanism of action, 3017–3018
  - organ/system effects of, 3020
  - pharmacological actions, 3016–3018, 3017*t*
  - precautions/adverse effects, 3021
  - prediction of response, 3021
  - therapeutic indications, 3020–3021
- Calcium ions (Ca^{2+})
  - cellular electrophysiology and, 131
  - lithium and, 174
  - regulation in autism, 314
  - serum levels, 174
- Calcium ions (Ca^{2+}) channels
  - CACNA1A, 144
  - function of, **139–140**
  - high-voltate-activated (HVA), 139
  - L-type, 139
  - low-voltage-activated, 139
  - N-type, 139
  - P/Q-type, 139
  - R-type, 139

- T-type, 139
  - point mutations in, 140
- Calfornia Personality Inventory (CPI), 998*t*
- Calgary Depression Scale, 1421, 1421*t*
- California poppy, 1782
- California Verbal Learning Test (CVLT), 990
- California Verbal Learning Test-II (CVLT-II), 4370
- Callosotomy, 334
- Callous–unemotional (CU) traits, 3612
- Calmodulin, calcium and, 105
- Cambodian Holocaust survivors, 1887
- Cambodian refugees, 806
- Cambridge Depersonalization Scale (CDS), 1871, 1877*t*
- CAMP. *See* Cyclic adenosine monophosphate
- CAMP response element-binding protein (CREB), 1284
- Campral. *See* Acamprosate
- *Campylobacter jejuni* infections, 629–630
- *Canadian Code of Ethics for Psychologists,* 4241
- Canadian Community Health Survey, 1614
- Canadian Network for Mood and Anxiety Treatments (CANMAT), 1689, 1692
- Canadian Psychiatric Association (CPA), 4566
- Canadian Study of Health and Aging, 3963
- Cancer. *See also* Carcinogenesis; *specific types of cancer*
  - adaptation to, 2252
  - alcohol and, 1268
  - biology of, 2285
  - cannabis use and, 1309
  - children/adolescent, response to, 3848
  - chronic, psychological concerns in, 2274–2277
  - complementary and alternative medicine, 2286–2287
  - delayed consequences of treatment, 2277
  - distress-related to
    - LSD, 2893
    - psilocybin, 2890–2891, 2893–2894
  - external loop, 2285
  - failure of apoptosis and, 46
  - family-centered approaches, 2279

- genetic risk for, 2277
- and HIV infection, 4271
- hypnosis and, 2730
- initiation of, 2285
- internal loop, 2285
- mood disorder and, 1249
- pain in, 2274–2275
- pain of
  - WHO ladder for control of, 2293
- panic disorder and, 1783
- poor coping skills, predictors of, 2277
- promotion and, 2285
- psychiatric disorders in
  - diagnosis of, 2255–2266
  - epidemiology of, 2252
  - management of, 2255–2266
- psychiatric symptoms, 2252*t*
- psychological interventions, 2286
- psychological responses to, 2251–2252
- psychoneuroimmunology and, 2322
- psychopharmacology, 2253
- psychotherapeutic interventions, 2253–2254
- in the psychotherapist, 2280–2281
- psycosocial factors, 2285
- research review, 2285
- risk of
  - environmental factors, 2285
  - internally mediated, 2285–2286
- schizophrenia comorbidity, 1541
- screening for, 2277–2278
- and sexual dysfunction, 4262–4263
- sexuality in patients with, 2275–2276
- survivors of
  - elderly patients and, 2285–2287
  - environmental factors, 2285
  - factors in, 2285–2286
  - families of, 2253

    - pediatric, 2282
    - psychological issues, 2276
  - tobacco use, 1347
  - treatment of
    - calcium channel inhibitors, 3020–3021
    - ethical issues, 2278
    - psychiatric problems due to, 2270–2274
  - U.S. incidence if, 2251
  - U.S. mortality rates, 2251
  - websites related to, 2287
- Cancer and, 1268
- Cancer antigen 125 (CA125), 2277
- Candidate gene studies, 3453–3456
- Cannabidiol (CBD)
  - anxiety disorders, 1743*t*, 1750
  - PTSD, 1823
  - therapeutic effects of, 115
- Cannabinoid receptor I (CBI)
  - expression by basket cells, 17
  - function of, 110, 112–114
- Cannabinoids
  - drug reward and, 373–374
  - endogenous, 112*f*
  - for nausea, 2271
  - nonpsychoactive, 115
  - plant-derived, 111*f*
  - synthetic, 111*f*
- Cannabinol, synthesis of, 110
- Cannabis. *See* Marijuana
- Cannabis dependence. *See also* Marijuana
  - treatment, 3010, 3127
- *Cannabis sativa,* 58
- Cannabis use and cannabis-related disorders, **1303–1312**
  - during adolescence, 3367, 3372
  - adverse effects of, 1309–1310
  - anxiety disorder, cannabis-induced, 1307
  - cannabinoid pharmacology, 1305–1306

- comparative nosology, 1304
- dependence, 1306, 1310–1311
- diagnosis and clinical features, 1306–1309
- epidemiology, 1304–1305, 1304*t*
- history, 1303
- intoxication, 1306, 1311
- intoxication delirium, 1306–1307
- laboratory examinations, 1310
- methods of use, 1304
- neuroimaging, 3475
- not otherwise specified disorders, 1308–1309
- pharmacology, 1305–1306
- preparations, 1303–1304, 1304*f*
- schizophrenia and, 1307, 1450–1451, 1455
- schizophrenia development and, 1451
- therapeutic effects of, 1311
- tolerance, 1307–1308
- treatment, 1310–1311
- withdrawal, 1307–1308

- Cannibalism, kuru and, 611, 611*f*
- Cannon, Walter, 2315
- Canonical correlation, 4567
- *Canterbury v. Spence,* 4434
- CAORs. *See* Conditioned attentional orienting responses (CAORs)
- CAPA. *See* Child and Adolescent Psychiatric Assessment
- Capacitance, ion channel, 132
- Capacity
  - to consent, 584–585
  - evaluation of, 2636
  - in schizophrenia, 696
  - testamentary, 4438
- Capgras delusion, 413
- Capgras syndrome
  - after TBI, 538
  - and consciousness, 691
  - description of, 4567
- Capital punishment, and juvenile justice, 4439–4440

- Capitation, 4567
- Capsulotomy, anterior, 3317, 3317*f*, 3318–3319
- Carbachol, uses of, 73
- Carbamazepine (Tegretol), **3022–3039**
  - for acute depression, 3031, 3031*f*–3032*f*, 3031*t*, 3033
  - acute mania, 1684, 1689, 1689*t*
    - combination/add-on therapy, 1687
    - emergency management, 1688
  - for acute mania, 3030–3031, 3030*f*
  - adverse effects of, 2356*t*
  - for aggressive behavior, 2111
    - conduct disorder, 2501
    - in IED, 2501
    - intermittent explosive disorder, 2501
    - mentally retarded children, 2500
  - BD, 3016, 3027, 3028*f*, 3031*t*, 3036–3037, 3038*t*
  - benzodiazepine withdrawal, 1387–1388
  - bipolar depression, 1690, 1693, 1694
  - bipolar disorders, 1683*t*, 1696*t*, 1697
    - combination therapy, 1696*t*
    - early onset, 3791
    - in HIV settings, 575
  - in BPD, 519
  - chemistry, 3022–3023, 3023*f*
  - for children with CD, 2501
  - combination pharmacotherapy, 3325, 3326, 3327, 3328
  - dementia treatment, 1201
  - for depression
    - in HD, 556
  - dissociative identity disorder, 1927*t*, 1928
  - drug interactions, 2199, 3037–3038, 3038*f*, 3038*t*
    - anticonvulsants, 2975, 2976, 2978
    - calcium channel inhibitors, 3021
    - nefazodone, 3130
  - drug studies, 3030–3033, 3030*f*–3033*f*, 3031*t*
  - FGAs and, 3081
  - in functional GI disorders, 2206

- history, 3022
- inhalant-related disorders, 1339
- lamotrigine and, 3037, 3086, 3326
- lithium and, 3036, 3098, 3325
- lithium *v.,* 3025, 3026–3027, 3026*f*, 3033
- mechanism of action, 507
- metabolism, 3024, 3038*f*
- mood stabilizers, 168
- organ/system effects of, 3033–3036, 3034*f*, 3035*t*
  - blood, 3033, 3035, 3035*t*
  - CNS, 3033, 3034*f*
  - heart, 3036
  - hormones, 3035
  - liver, 3036
  - renal effects, 3035
  - skin, 3036
  - weight, 3036
- oxcarbazepine and, 3022, 3023*f*, 3023*t*, 3024, 3036, 3037–3038
- in pain syndromes, 2396
- pediatric pharmacogenomic testing for, 3456
- pharmacodynamics, 3024–3030
  - antidepressant action, 3024–3027, 3024*f*
  - antimanic action, 3024–3027, 3024*f*
  - antinociceptive action, 3024–3027, 3024*f*
  - tolerance development, 3027, 3027*f*–3028*f*, 3029–3030, 3029*t*
- pharmacological actions, 3023–3030, 3023*t*
- and physical health monitoring, 1024
- precautions/adverse reactions, 3037
- pregnancy and, 3037
- prophylatic, 3033, 3033*f*
- seizure, 3029, 3029*t*
- side effects, 2237
- therapeutic indications, 3036–3037
  - alcohol withdrawal, 3037
  - BD, 3036–3037
  - impulse control disorder, 3037
  - PTSD, 3037

    - schizoaffective disorder, 3037
    - schizophrenia, 3037
  - thyroid hormones and, 174
  - tolerance development, mechanisms of, 3027, 3027*f*–3028*f*, 3029–3030, 3029*t*
  - valproate and, 3036–3037, 3279, 3326
  - valproate *v.,* 3025, 3026, 3026*f*, 3030, 3031, 3267–3268, 3267*t*
- Carbon atoms
  - ^{11}C, 286*f*
  - characteristic of, 260
- Carbon disulfide, 631
- Carbon monoxide (CO), 108–109
  - fetal exposure to, 2484
  - generation of, 108, 108*f*
  - molecular actions of, 108–109
  - neurotransmission, 109
  - signaling roles of, 109
- Carcinoembryonic antigen (CEA), 2277
- Carcinogenesis. *See also* Cancer
  - estradiol and, 3334–3335
  - nonbarbiturate sedative-hypnotic agents and, 2989
  - opioid receptor antagonist effects on, 3143
  - vilazodone and, 3346
- Carcinoid syndrome, 684
- Carcinomatous encephalomyelitis, 2270
- Cardena, Etzel, 2719
- Cardene. *See* Nicardipine
- Cardiac arrhythmias
  - definition of, 2189–2190
  - diagnosis of, 2190
  - nosology of, 2189–2190
  - psychiatric complications in, 2195
  - psychological factors in, 2190
  - stress and, 2316
  - sudden cardiac death and, 2189–2190
- Cardiac disorders
  - panic disorder and, 1783
  - psychiatric problems in patients with, 2195–2199

- Cardiac function, stress and, 2316*f*
- Cardiac surgery, complications in, 2195–2196
- Cardiomyopathy, clozapine-associated, 1023
- Cardioquin. *See* Quinidine
- Cardiovascular disease (CVD). *See also* Heart disease
  - alcohol-related, 1267, 1268
  - anabolic-androgenic steroid use and, 1395
  - arteriosclerotic, gambling and, 1805
  - cannabis use and, 1309
  - cytokines and, 2321
  - erectile dysfunction and, 3338
  - genetic studies, and environment, 937
  - psychiatric problems and, **2186–2200**
  - risk factors, 214
  - schizophrenia and, 1434, 1539–1540, 1544–1545, 1545*f*
  - tobacco use, 1347
- Cardiovascular drugs, side effects of, 2198–2199, 2198*t*
- Cardiovascular problems and, 1268
- Cardiovascular system. *See also* Cardiovascular disease (CVD); Heart
  - β-adrenergic receptor antagonist effects on, 2955–2956
  - anticholinergics and, 2958
  - antihistamine effects, 2983
  - antipsychotic drug effects, 4138
  - barbiturate effects on, 2986
  - buspirone effects on, 3012
  - DA receptor agonist effects on, 3238–3239
  - ECT effects in, 3295–3296
  - and electroconvulsive therapy, 4158
  - FGA effects and, 3075
  - and HIV infection, 4271
  - lithium and, 3097
  - MAOI effects on, 3117
  - melatonin effects and, 3104
  - nefazodone and, 3128
  - nonbarbiturate sedative-hypnotic agent effects on, 2989
  - and schizophrenia, 4097
  - and sexual dysfunction, 4262

- sleep and, 344
- SSRI effects on, 3157
- sympathomimetic effects on, 3230
- trazodone effects on, 3251
- tricyclic and tetracyclic antidepressant effects on, 3260–3261, 3262
- vilazodone effects on, 3344

- Cardizem. *See* Diltiazem
- Care coordination model, 4322–4323
- Care management/case management
  - of chronic illnesses, 4312
  - in collaborative care, 4322
- Carebaria, description of, 4567
- Caregiver–infant reciprocity, feeding disorder of, **3629–3630, 3643–3644**
  - clinical features, 3629, 3643
  - course, 3629, 3643
  - differential diagnoses, 3629, 3643
  - etiology, 3629, 3643
  - treatment, 3629–3630, 3643–3644
- Caregivers
  - attachment, 3662–3665, 3663*t*, 3676–3679, 3677*t*
  - of bipolar disorder children, 3929–3930
  - counseling and support for, 4208–4212
    - clinical issues
      - indications, 4210
      - limitations, 4210–4211
    - definitions, 4208–4209
    - education and support groups, 4210
    - ethical issues, 4211
    - history and theoretical issues, 4209
    - mental health effects, 4209
    - physical health effects, 4209
    - research and evaluation, 4211
    - skills-training classes, 4210
  - definition of, 4567
  - disruptive behavior disorder, 3606, 3609
  - electronic technology, 4279
  - end-of-life decisions and, 2304

- family role, 4195
- feeding disorder, 3629–3630, 3643–3644
- gender differences, 4248
- group therapy, 4202, 4207
- groups of, 2253
- inadequate, depression and, 1662
- primary, 2304
- responses to patient refusal of care, 2314
- and sexual dysfunction, 4263
- support for, 1200

- Caregiving deficiencies in infancy, 2047–2048
- Caribbean immigrants, 804
- Cariprazine, 3234, 3238, 3239
  - acute mania, 1686–1687, 1686*t*, 1689*t*
  - bipolar depression, 1691, 1692*t*, 1693
- Carisoprodol (Soma), 1384, 2990
- Carlson, Eve B., 1870
- Carlsson, Arvid, 1488
- Carnegie Foundation, 4323
- Carolusson, Susanna, 2729
- Carothers, John, 1427
- Carotid bruit, 456
- Carotid endarterctomy, 486
- Carr, John, 2446
- CART machine learning techniques, 841
- Carter, Jimmy, 4293, 4311
- CASE. *See* Chronological Assessment of Suicidal Events
- Case history studies, 4567
- Case management
  - by mental health courts, 4329
  - treatment integrated with, 4398–4399
- Case registers, definition of, 4567
- Case studies, as outcome evaluation, 2703
- Case–control studies
  - definition of, 4567
  - epidemiology of, 815
- *Casey, Planned Parenthood of Southeastern Pennsylkvania v.,* 2512

- Caspase-activated DNAse (CAD), 46
- Caspases, conversion of, 47*f*
- Castes, Marianella, 2729
- Castillo, Richard, 1943
- Castration
  - for paraphilias, 2020
  - of sex offenders, 2022
  - sexual function and, 1970
- Castration anxiety, 4567
- Castration complex, 4567
- CAT scans. *See* Computed tomography (CT)
- Cat scratch disease, 593–594
- Catabolism, 677
- Catalepsy, 2711, 4567
- Cataphasia. *See* Verbigeration
- Cataplexy
  - description of, 4567
  - narcolepsy and, 2092
- Catapres. *See* Clonidine
- Catastrophic anxiety, 4567
- Catastrophizing
  - after stroke, 478
    - diagnosis of, 482
    - etiology of, 480
  - as cognitive distortion, 2695–2696
  - insomnia and, 2090*t*
- Catathymia, 4567
- Catathymic crisis, 4567
- Catatonia
  - catatonic disorder due to another medical condition, 1258–1259
  - course and prognosis, 1595
  - description of, 4567
  - diagnosis and clinical features, 1595
  - differential diagnoses, 1595
  - disorganization syndrome and, 1419–1420
  - EEGs in, 249
  - epidemiology, 1594

- etiology, 1594–1595
- history, 1594
- manic syndrome and, 1640
- treatment
  - benzodiazepine, 2997
  - ECT, 3287

- Catatonic excitement, 4567
- Catatonic posturing, 4567
- Catatonic rigidity, 4567
- Catatonic schizophrenia, 4567
- Catatonic state, 4567
- Catatonic stupor, 4567
- Catatonic syndrome, 459
- Catchment areas, 4567
- Catechol-*O*-methyltransferase (COMT)
  - aggression and, 2483
  - anxiety disorders and, 1739
  - bipolar and schizophrenic disorders and, 1610
  - function of, 66
  - gene variants, 2483
  - isoforms of, 66
  - Val158Met polymorphism, 1745
- Catecholamines, 2318, 4567
  - dissociative identity disorder and, 1882
  - OCD and, 1787
  - phobias and, 1742
  - PTSD and, 1819–1820
  - responses to stress, 2318
  - SNS and, 193
  - trauma response and, 2569
- Categorical attitudes, 4567
- Categorical thought, 4567
- Categorization, 690
- *Catha edulis,* 1282
- Catharsis
  - description of, 4567
  - group psychotherapy and, 3761

- Cathartic method, in psychoanalysis, 2639
- Cathexis, 4567
- Cathinones, 1282
- CATIE. *See* Clinical Antipsychotic Trials of Intervention Effectiveness
- Cat's cry syndrome, 4572
- Cattell, Raymond, 935
- Cattell Infant Intelligence Scale, 4567
- Cattell–Horn–Carroll (CHC) theory, 2431
- Caudal midbrain, REM sleep and, 347–348
- Caudate nucleus
  - anatomy, 26
  - coronal section, 12*f*
  - function of, 8, 10
  - internal organization of, 28–29
  - neurotensin in, 177
  - regions, 27
- Causal graph models, 838
- Causalgia, 4567
- Causation by physician, 4436
- Cause-effect relationships, 4567
- CBD. *See* Corticobasal degeneration
- CBT. *See* Cognitive-behavioral therapy (CBT)
- CCK. *See* Cholecystokinin
- CDS. *See* Cranial electrical stimulation
- Ceftriaxone (Rocephin), 591
- Celecoxib, 1681
- Celenza, Andrea, 2669
- Celexa. *See* Citalopram
- Cell death. *See also* Apoptosis
  - apoptotic, 46
  - developmental, 46
  - histogenetic, 46
  - morphogenetic, 46
- Cell membranes
  - homologous repeats in, 135*f*
  - ion channels, 133*f*
- Cells, brain organization and, 3–8

- Cellular immunity, 183, 2339*f*
- Censored observations, 844
- *Centella asiatica,* 1782, 2554
- Centenarians, brains of, 3993–3994
- Center for Epidemiological Studies Depression Scale (CES-D)
  - for aging adults, 3976
  - assessing cancer patients, 2283
  - as disorder-related rating scale, 3400
- Center for Medicare/Medicaid Services (CMS)
  - accountable clinical entities, support from, 4307
  - payment mechanisms, 4305, 4307, 4308
  - reimbursement for chronic illness care management, 4321
- Center for Substance Abuse Treatment (CSAT), 3133
- Center for the Study of Traumatic Stress, 2573
- Central apnea, 2096–2097
- Central coherence theory, 3553
- Central fissure, 7*f*
- Central Intelligence Agency (CIA), 1946–1947
- Central nervous system (CNS)
  - alcohol's effect on, 1267
  - anticholinergics and, 2958, 2959, 2961
  - barbiturate effects on, 2986
  - benzodiazepine effects on, 2995
  - caffeine effects on, 1295, 1295*f*
  - cancers
    - metastatic, 2268
    - psychiatric syndromes and, 2266–2267
  - carbamazepine effects on, 3033, 3034*f*
  - cytokines and, 2321
  - DA receptor agonist effects on, 3238
  - depersonalization and, 1891
  - description of, 4567
  - development of, **438–444**
    - testosterone in, 2505
  - dissociative identity disorder and, 1894
  - dysfunction, depression and, 1709
  - ECT effects in, 3296

- HIV and, 1244
  - in adolescents, 645–646
- immune system interactions with, **178–184,** 194–197, 2319–2320
- induction of, 48
- infections of, 1011
- inhalant use and, 1333
- Lyme disease, 591
  - differential diagnoses, 591
  - tests for, 591
  - treatment, 591–592
- lymphomas, 586
- MAOI effects on, 3117
- melatonin effects and, 3104
- movement controlled by, 2936, 2938
- nefazodone and, 3128
- non-HIV viral infections of, 594–596
- nonbarbiturate sedative-hypnotic agent effects on, 2989
- PTSD and, 1817–1818
- sexual behavior and, 1956–1957
- sexual dimorphism of, 2505
- sympathomimetic compounds, **3219–3233**
- trazodone effects on, 3252
- tricyclic and tetracyclic antidepressant effects on, 3260, 3262

- Central pain states, 2390*t*
- Central poststroke pain, 2391
- Central sleep apnea, 2095–2097
  - description of, 4567
  - diagnostic criteria, 2095–2096
  - of infancy, 2097
  - of prematurity, 2097
  - treatment-emergent, 2097
- Centrex. *See* Prazepam
- Cephalagia. *See* Headache
- Cerea flexibilitas, 4567
- Cerebellar cortex, 1267
- Cerebellar degeneration and, 1267
- Cerebellar disorders

- chemotherapeutic agents and, 2273*t*
- cognitive features, 559
- course/prognosis, 559
- differential diagnoses, 559
- diseases causing, 558*t*
- motor function in
  - clinical features, 558–559
  - diagnosis, 558
  - epidemiology, 557
  - etiology, 557–558
- pathology, 559
- psychiatric features, 559

- Cerebellothalamiocortical systems, 26
- Cerebellum
  - ADHD, neuroimaging in, 3464
  - deep nuclei of, 13*f*
  - neurobiology of, 450–451
  - neuroimaging, 2815
  - surface landmarks, 7*f*
- Cerebral amyloid angiopathy (CAA), 3992
- Cerebral aqueduct, 14*f*
- Cerebral arteriosclerosis, 4567
- Cerebral blood flow (CBF). *See also* Regional cerebral blood flow
  - in aging brain, 3990, 3992
  - in posttraumatic stress disorder, 699
  - regional, 344–346
  - regulation of, 3
- Cerebral cortex
  - Brodman map of, 977, 977*f*
  - development of, 44*f*, 48–50, 686–687
    - cell migration in, 52–53
    - layer-specific neurons, 49–50
    - patterning in, 48*f*
  - dopaminergic innervation, 8
  - efferent projections, 19*f*
  - embryonic origins of, 8
  - functional networks, 16–21, 21*t*

- neuroanatomy, 451–452
- noradrenergic neurons in, 1714
- organization of memory, 730, 730*f*
- regions of, 8, 11

- Cerebral creatine deficiency syndrome (CCDS), 663–664
- Cerebral dominance theory, 3560
- Cerebral folate deficiency (CFD), 664–665
- Cerebral folate transport deficiency (CFTD), 664
- Cerebral hemispheres
  - connectivity, 13
  - coronal section, 13*f*
  - right
    - schematic drawing, 23*f*
- Cerebral lateralization
  - barbiturates and, 2987
  - theory of stuttering, 3560–3561
- Cerebral palsy, 454
- Cerebral vasculitis, MS *versus,* 563
- Cerebrospinal fluid (CSF)
  - circulation, 14–16
  - elevated glutamate
    - in TBI, 524
  - examination of, 468, 615
  - functions, 16
  - HVA in, 14
  - hydrocephalus and, 14
  - measurements, 16
  - metabolites, OCD and, 1787
  - production of, 14–16
  - studies for dementia, 1197
  - ventricular system and, 14
- Cerebrotendinous xanthomatosis (CTX)
  - description of, 662–663
  - diagnosis of, 662–663
  - features of, 467*t*
  - histology of, 662–663
  - MRI imaging in, 662*f*

- Cerebrovascular accidents (CVA). *See also* Stroke
  - aggression and, 2495
  - description of, 4567
- Cerebrovascular disease
  - aggression and, 2495
  - mania without prior bipolarity and, 1649
  - neuropsychiatric aspects of, **474–489**
  - and psychiatric symptoms, 4046
- Cerebrovascular events
  - antipsychotic drug effects, 4136
- Cerebrum, horizontal section, 14*f*
- Ceruloplasmin
  - function of, 672
  - serum levels of, 672
  - in Wilson disease, 556
- Cervical cancer, death rate from, 2269
- CES. *See* Cranial electrical stimulation
- CES-D. *See* Center for Epidemiological Studies-Depression Scale (CES-D)
- CFD. *See* Cerebral folate deficiency (CFD)
- CFI. *See* Cultural Formulation Interviews (CFI)
- CFTD. *See* Cerebral folate transport deficiency (CFTD)
- CGAS. *See* Children's Global Assessment Scale
- Chain analysis in DBT, 2789–2790
- Chaining, in operant behavior therapy, 2872–2873
- Change, states of, 2218*t*
- Channelopathies, 143–144, 676
- Channelrhodopsin (ChR), 145
- Chaplaincy services, 2254
- Character
  - biotypes, 2137*t*
  - cognitive function and, 2136*t*
  - configurations of, 2146*f*
  - description of, 4567
  - development of, 2141, 2149*t*, 3943
    - therapeutic procedures for, 2175*t*
  - dimensions of, 2135*t*, 2148*f*
  - disorders of, 4567

- emotional state and, 2134*t*
- genomic approaches, 2136–2138, 2137*t*
- identity, temperament and, 2136*t*
- inflexibility and, 2138
- narrative psychiatry, 2832–2833
- positive psychology, 2851–2852
- psychobiology of, 2134–2138
- psychodynamic concept of, 2138–2139
- TCI scales, 2164*t*
- temperament, identity and, 2145
- traits
  - sense of goodness and, 2141
- traits, psychobiology of, 2138–2139

- Character analysis, 4567
- Character defense, 4567
- Character neurosis, 2138
- Character pathology
  - addictive behavior and, 2050
  - narcissistic, 2050
- *Characters* (Theophrastus), 934
- Charcot, Jean-Martin, 4567
  - dissociative disorders, 1867, 1868*f*
  - on hypnosis, 2639, 2706
  - *Hysteria, Hypnosis, and Healing,* 4536
  - on *mania operativa activa,* 1847
- Charcot hysteria, 2178
- Charcot–Marie–Tooth disease, 628
- Charcot–Willbrand syndrome, 455
- Charismatic religious sects, 1947
- Charity, spirituality and, 2583
- Charles Bonnet syndrome, 465
- Charney, Dennis, 1161
- Chein, J.M., 3365
- Chemical aversion, behavior therapy, 2694
- Chemical genetics, 145–146
- Chemical shift imaging (CSI), 278–279
- Chemo-brain, presentation of, 2272

- Chemoarchitecture
  - brain, 5
  - medial temporal lobe, 6*f*
- Chemokines, 182*t*
- Chemotherapeutic agents, 2273*t*
- Chemotherapy. *See also* Drug therapy
  - CNS toxicity in, 2272
  - dementia related to, 2284
  - ototoxicity of agents used in, 2272
  - psychiatric problems related to, 2271–2272
  - rigors of, 2263
  - side effects, 2255
- Chess, Stella
  - on temperament, 3361
  - temperament classification, 3479, 3479*t*, 3480, 3488
- Chessick, Richard, 2648
- Cheyne–Stokes Breathing, 2095–2096
- Chi square, 4568
- Chiarugi, Vincenzo, 4534
- Child abuse. *See also* Battered child syndrome; Maltreatment (child)
  - adult abuse and, 2540
  - betrayal trauma theory on, 1902–1903
  - daycare settings, 1871–1872
  - exposure to, 400
  - factitious disorder imposed on another and, 1847, 1849, 1851–1852
  - and genetic selection, 778
  - memory of, 743
  - prostitutes and, 1889
  - psychiatric emergency due to, 2634
  - PTSD and, 2569
  - sexual, GAD and, 1728
  - signs of, 2385*t*
  - suicide and, 2612
- Child abuse and neglect, evaluation for, 3445–3446
- Child Abuse Prevention and Treatment Act (CAPTA) of 1974, 3828
- Child/Adolescent Anxiety Multimodal Study (CAMS), 3710, 3745, 3792–3793
- Child Advocate's report on Sandy Hook tragedy, 4298

- Child analysis, description of, 4568
- Child and Adolescent Psychiatric Assessment (CAPA), 821, 3398
- Child and adolescent psychiatry
  - examinations, 4427
  - training, 4405, 4406*t*
- Child Behavior Checklist (CBCL), 2687
- Child Behavior Checklist for Ages 1½-5 (CBCL/1½-5), 3398–3399
- Child Behavior Checklist for Ages 6 to 18 (CBCL/6 to 18), 3398–3399
- Child Behavior Checklist-Juvenile Bipolar Disorder/Dysregulation Profile (CBCL-JBD), 3687, 3701
- Child custody
  - evaluations for, 3391
  - forensic psychiatry and, 3863–3864
- Child-Directed Interaction (CDI), 2693
- Child Dissociative Checklist (CDC), 1876*t*, 1902
- Child Health and Development Study, 800
- Child molesters, definition of, 2017
- Child neurology, **637–655**
- Child-onset fluency disorder. *See* Stuttering
- Child-parent psychotherapy (CPP), 3706
- Child psychiatric evaluations
  - adolescent considerations, 3385–3387, 3387*t*
  - clinical interview, **3375–3385**
    - child interview during, 3381–3385, 3383*t*
    - considerations during, 3375–3376
    - context considerations, 3376
    - informants for, 3376, 3410–3411
    - parent interview during, 3376–3381, 3388–3389
    - preliminary contact, 3376
    - types of, 3395–3398
  - cognitive assessments, **3413–3429**
    - achievement testing, 3425–3426
    - applications for, 3422–3429
    - cognitive tests, description of, 3415–3422, 3416–3417*t*, 3418*t*
    - intellectual testing, 3422–3425
    - legal considerations for, 3415–3417, 3419–3421*t*
    - overview, 3413–3415
    - of processing problems, 3426–3429

- diagnostic formulations and recommendations, 3408–3409
- diagnostic laboratory tests, 3407–3408
- emergency, **3436–3447**
  - for aggression, 3443–3444, 3444*t*
  - agitation management during, 3440–3441, 3441*t*
  - for autism spectrum disorder, 3445
  - for child abuse or neglect, 3445–3446
  - components for, 3438–3440, 3439*t*
  - contextual factor assessment during, 3442
  - developmental considerations for, 3441–3442
  - for eating disorders, 3446, 3446*t*
  - epidemiology for, 3437
  - etiology determination, 3438, 3438*t*
  - for intellectual disability, 3445
  - legal considerations for, 3447
  - overview, 3436–3437
  - principles of, 3438–3442
  - priorities for, 3437–3438
  - for psychosis, 3444–3445
  - for psychosocial crisis, 3446–3447
  - for risky behaviors and substance use, 3444
  - safety concerns, 3437–3438
  - for suicidality and self-harm, 3442–3443
- infant and toddler considerations, 3387–3391, 3389*t*, 3390*t*, 3392–3393*t*
- neuropsychological assessments, **3413–3415, 3429–3436**
  - advances in, 3429, 3431
  - application of, 3431–3433
  - on attention, 3433–3434
  - on executive functioning, 3434
  - as functional assessment, 3429
  - further considerations for, 3436
  - on language, 3434–3435
  - overview, 3413–3415, 3416–3417*t*, 3418*t*
  - on sensory/motor functioning, 3435–3436
- preschool children, **3409–3413**
  - clinical formulation, 3412
  - developmentally appropriate criteria for, 3409–3410

    - informants, multiple, 3410–3411
    - multi-disciplinary, 3410–3412
    - multiaxial, 3411–3412
    - multiple approaches to, 3411
    - session numbers, 3410
    - treatment planning, 3412
  - psychological assessments (standard), **3403–3406, 3413–3414**
    - interpretation of, 3405–3406
    - measurements for, 3413–3414
    - overview, 3413–3414
    - personality measures, 3404
    - process for, 3413
    - referrals for, 3403–3404
    - reporting, 3406
    - test batteries for, 3404, 3405*t*
    - validity and reliability of, 3404–3405
  - referrals, as component of, 3403–3404
  - standardized assessment instruments, 3391–3403
    - development of, 3394–3395
    - interviews types, 3395–3398
    - limitations of, 3402–3403
    - rating scales as, 3398–3402
  - types of, 3391, 3394*t*
- Child psychiatry, 4541
  - adoption, **3823–3828**
  - anxiety disorders
    - generalized anxiety, **3712–3714**
    - obsessive-compulsive disorder, **3695–3703**
    - posttraumatic stress disorder, **3703–3707**
    - selective mutism, **3719–3724**
    - separation anxiety, **3708–3712**
    - social anxiety, **3714–3719**
  - attachment disorders, **3662–3667, 3676–3681**
  - attention-deficit/hyperactivity disorder, **3587–3604**
  - autism spectrum disorder, **3571–3586**
  - communication disorders, **3545–3570**
  - development (normal)

    - adolescent, **3364–3374**
    - child, **3351–3364**
  - developmental coordination disorder, **3537–3544**
  - disorders
    - classification, 3348–3349
    - emergence of, 3349
    - interventions for, 3350
    - neuroimaging studies, 3349–3350
  - disruptive behavior disorders, **3605–3621**
  - elimination disorders, **3650–3661**, **3664–3675**
  - ethical issues, **3868–3880**
  - evaluations in, 3398–3399
  - family therapy, **3771–3787**
  - feeding disorders of infancy/childhood, **3622–3634**, **3636–3648**
  - forensic, **3861–3868**
  - foster care, **3817–3823**
  - genetics in, 3350, 3448–3457
  - groups of mental disorders, 2159
  - HIV and AIDS sequelae, **3852–3857**
  - hospitalization issues, **3843–3845**, **3848–3849**
  - illnesses, responses to, **3845–3852**
  - intellectual disability (ID), **3491–3519**
  - maltreatment, **3828–3843**
  - mental health services research, **3900–3913**
  - overview, 3817
  - parental issues, **3925–3933**
  - parenting stressors, **3925–3933**
  - prevention in, **3893–3900**
  - school consultation, **3880–3893**
  - sleep disorders, **3934–3940**
  - specific learning disorder, **3520–3536**
  - stereotypic movement disorders, **3667–3673**, **3681–3687**
  - suicide behavior risk factor, 3682–3683, 3696–3697
  - temperament, 3479–3490
  - terrorism, impact of, **3913–3925**
  - tic disorders, **3635–3649**, **3649–3663**
  - treatment options, **3734–3816**

    - for adolescents, **3810–3816**
    - brief psychotherapies, **3741–3749**
    - cognitive-behavioral therapy, **3749–3759**
    - community-based, **3803–3810**
    - family therapy, **3771–3787**
    - group psychotherapy, **3760–3771**
    - inpatient, **3797–3801,** 3815
    - psychodynamic psychotherapy, **3734–3741**
    - psychopharmacology, **3787–3797**
    - residential treatment centers, **3801–3802**
- *Child Psychiatry* (Kanner), 4539
- Child welfare programs, 4321
- *Childhood and Society* (Erikson), 4539
- Childhood autobiographical memory disturbances (CAMD), 1885–1886, 1885*t*
- Childhood disintegrative disorder, 4568
- Childhood disorders. *See specific disorders*
- Childhood-onset schizophrenia, 3474–3475, 3474*f*
- Childhood schizophrenia, 4568
- Childlessness as psychosocial factor, 4165
- Children and law
  - child abuse, 4440
  - child custody, 4440
  - juvenile justice, 4439–4440
  - violence, 4439–4440
- Children/childhood. *See also* Adolescence/adolescents; Child abuse
  - addictive disorders and, 2048
  - ADHD in, 650
  - adverse events, 2048
  - adverse experiences, 1885*t*
  - aggressive behavior in, 2484, 2499
    - treatment, 2499
  - antidepressant-associated suicidal behavior and, 1673, 1779–1780
  - antiretroviral therapy in, 645–646
  - anxiety disorders, 1732–1734
    - assessment, 1721
    - management, 1784
  - asthma in, 2224

- attachment disturbance, 2720
- autoimmune neuropsychiatric disorders, 201
- BDD treatment in, 1792
- behavioral insomnia, 2086
- BIF in, 2436–2437
- brain tumors in, 504–506
- burn injuries in, 2384–2386
  - abuse-related, 2385*f*
  - developmental considerations, 2386
  - psychological sequelae, 2385
  - risk of, 2374, 2384*f*
  - support for staff, 2386
  - treatment of, 2385
- CAM treatments and, 2543
- with cancer, 2282
- care world-wide, 4525
- Cooperativeness in, 2167
- DBT and, 2795
- delirium in burn injuries, 2376
- delirium incidence in, 1185
- development
  - of brain, 3351, 3356–3357*t*, 3358
  - cognitive, 3356–3357*t*, 3359–3360
  - emotional, 3352–3353*t*, 3361
  - of gender identity, 3354–3355*t*, 3363
  - of language, 3358–3359
  - of morality, 3354–3357*t*, 3361
  - motor, 3361–3362
  - overview, 3351, 3352–3357*t*
  - physical, 3356–3357*t*
  - psychosexual, 3354–3357*t*, 3363–3364
  - of self-esteem, 3370
  - sensory, 3362–3363
  - social, 3352–3353*t*, 3360–3361
  - of temperament, 3352–3353*t*, 3361
- developmental stages, 2281, 2408
- disaster response, 2570*f*

- disruptive behavior disorders of, 2485, 2488, 3250–3251
- dissociative disorders, 1870
  - developmental factors and, 1866
  - developmental mediation of, 1902–1903
  - developmental model, 1901–1902
  - imaginary companionship, 1902
- dissociative identity disorder, 1914–1915
- divorce during, 3785
- effects of prematurity in, 643
- end-of-life care, 2307, 2311–2312
- with epilepsy, 648–651, 650
  - adjustment issues in, 651
  - anxiety and, 649
  - attitudes toward, 651
  - depression and, 649
  - psychiatric issues and, 648–649
  - surgical treatment of, 651
- episodic course of dysthymia in, 1646
- extreme temperaments in, 2159
- fabricated illness in, 1846, 1847
- factitious disorder imposed on another and, 1864–1865
- family support for, 4296, 4298
- forensic psychiatry, 2734
- gender dysphoria in, 1985, 2028, 2036, 2037
- gender incongruence of, 2024, 2029
- generalized anxiety disorder in, 1767
- genetic testing in, 2282
- hallucinogens and, 2890
- history, at patient interview, 966
- HIV in, 645–648
- homosexuality and, 2002
- hospital care of, 4341
- hospitalizations, response to, 3848–3849
- hypnotizability in, 2707
- hypothalamic tumors in, 677
- illness falsification, 2408
- illness in, 2281

- illnesses, response to, 3848–3849
- imaginal flooding and, 2691
- inpatient psychiatric treatment, **3797–3801,** 3815
- involuntary hospitalization of, 4293
- in jails and prisons, 4466
- lead intoxication in, 631
- legal issues
  - child abuse, 4440
  - child custody, 4440
  - juvenile justice, 4439–4440
  - violence, 4439–4440
- lithium in, 3094
- mental disorders among, 4520
- Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA), 826
- migraine treatment in, 625
- moral reasoning of, 710
- neuroimaging, 3458–3478 (*See also* Neuroimaging)
- neuropsychiatric disorders and aggression, 2487–2488
- neuropsychological assessments for, **3413–3415, 3429–3436**
- non–possession-form dissociative identity disorder, 1918
- obese, prognosis for, 2211
- OCD, 1788
  - after TBI, 653
  - family therapy, 1790
  - pharmacotherapy, 1789
- origins of homosexuality, 1997
- overweight, 2211
- pathophysiology of TBI in, 652
- post-illness PTSD, 2331
- prenatally-acquired HIV, 647–648
- preschool child psychiatric evaluation of, **3409–3413**
- psychiatric disorders, 3348–3350
  - advances in, 3348
  - classification of, 3348–3349
  - emergence of, 3349–3350
  - interventions for, 3350
  - neuroimaging studies for, 3349–3350

    - onset of, 3349
    - protective factors, 3350
    - risk factors, 3350
  - psychiatric disorders, genetics and, 2529*t*
  - psychiatric emergencies
    - pharmacotherapy, 2628
  - psychiatric evaluations for, **3375–3447**
  - with psychiatric risk factors, 4304
  - psycho-oncology, 2281–2282
    - communication issues, 2281
    - family involvement, 2281–2282
    - pain management, 2281
    - psychiatric diagnoses, 2281
    - treatment issues, 2281
  - psychological assessments (standard) of, **3403–3406, 3413–3414**
  - psychopathological features of, 3375–3376
  - psychopharmacology, 2924
  - psychostimulants in, 3227–3228
  - PTSD
    - prevalence studies, 1816
    - psychotherapy, 1823
  - questions about death, 2308
  - recall, in posttraumatic stress disorder, 700–701
  - religious perceptions, 2584, 2585
  - schizophrenia in, 3071
  - schizophrenia risks, 1449–1450, 1451
  - school-age, end-of-life care, 2307
  - selective mutism, 1728
  - with serious emotional disorders, 4296, 4298, 4320–4321
  - services needed by, 4306
  - severely ill, communication with, 2308
  - skin disease in, 2333
  - social anxiety disorder in, 1766–1767
  - socialization to violence, 2486
  - somatic symptom disorders
    - reliability of, 1829*f*
  - SSRI use in, 3168

- staring spells in, 650
- stress in, disease-related, 2334
- suicide and, 2614–2615
- survivors of critical illnesses, 2330
- with TBIs, 652–653, 2964
  - assessments, 653
  - interventions, 653
- terminally ill, 2309–2310
- trauma in, 2111
  - autohypnosis and, 1902–1903
  - discrete behavioral states model of, 1900
  - panic disorder and, 1724
- traumatic brain injury (TBI) in, 651–654
- treatment barriers for, 4305
- treatment of GD in, 2032–2033
- tricyclic antidepressants in, 3256
- understanding of death, 2308, 2308*t*
- vilazodone in, 3347

- Children in the Community Study, 2485
- Children's Depression Inventory (CDI), 3399–3400
- Children's Depression Rating Scale-Revised (CDRS-R), 3400, 3677, 3691
- Children's Global Assessment Scale (CGAS), 822, 3402
- Children's Health Insurance Program (CHIP), 4379, 4380
- Children's Health Insurance Program Reauthorization Act (CHIPRA), 4380
- Children's Hypnotic Susceptibility Scale (CHSS), 2713
- Children's Interview for Psychiatric Syndromes (ChIPS), 3397
- Children's Schedule for Affective Disorders and Schizophrenia (K-SADS), 3641, 3655
- Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS), 3400
- Childress, James, 2836
- Chimpanzees, violence among, 2418
- Chinese Americans, 803
- Chinese traditional medicine
  - assessment, 2544–2546
  - in CAIM medicine, 3982
  - on energetic imbalances, 2546
  - herbal formulas, 2554–2555
- ChIP-chip (chromatin immunoprecipitation on a chip), 160

- ChIP-seq microarray, 439
- Chiropractic treatment, 2542
- Chlamydia, 2514
  - laboratory evaluation, 1012
- Chloral hydrate (Aquachloral, Noctec), **2989–2990**
  - chemistry, 2989
  - dependency diagnosis, 1384–1385
  - dosage/administration, 2990
  - drug interactions, 2990
  - drug studies, 2989
  - history, 2984, 2989
  - laboratory interference, 2990
  - organ/system effects of, 2989
  - pharmacological actions, 2989
  - precautions/adverse effects, 2987–2990
  - as sedative-hypnotic, 1384–1385
  - therapeutic indications, 2989
- Chlordiazepoxide (Librium)
  - for alcohol-related disorders, 1277
  - alcohol withdrawal, 2997
  - benzodiazepine dose equivalent, 2633
  - history, 2917, 2991
  - metabolism, 2993
  - misuse liability of, 1376, 1377
  - pregnancy and, 2999
  - withdrawal from, 1382
- Chloride (Cl^-) channels, **140–141**
- Chloride ions (Cl^-), 131
- Chloroacetaldehyde (CAA), 2272
- *Meta*-chlorophenylpiperazine (mCPP), 1746, 1892
- Chlorpromazine (Thorazine). *See also* Antipsychotics/neuroleptics
  - acute mania, 1686, 1689*t*, 1690
  - after World War II, 4292
  - bipolar disorder, early onset treatment, 3692, 3706
  - cardiac side effects, 3075
  - CE dosages of antipsychotics, 520*t*
  - chemistry, 3061, 3062*f*

- for delirium, 1187
  - in cancer patients, 2257, 2258*t*
  - in dying patients, 2301
- description of, 4568
- dosage, 2258*t*, 3080
- dose increases, 4329
- drug interactions, 3078
- history of use, 1405, 1520, 2917, 3060
- kidney/urinary effects, 3076
- laboratory interferences, 3078
- liver effects, 3076
- as major change, 4328
- MDD, 2996
- metabolism, 3065
- movement disorder, stereotypic, 3672–3673, 3686–3687
- movement disorders induced by, 2936
- Nobel Prize for research on, 4294
- potency, 2921
- psychopharmacology with, 2905
- psychosis, 2868, 3079
- for psychotic episodes, 519
- schizophrenia, 1524, 1527, 1529
- skin/eye effects, 3076

- Chlorprothixene (Taractan), 3062*f*, 3064
- Chochinov, Harvey Max, 2826
- Choice, Pavlovian learning and, 723, 723*f*
- Cholecystokinin (CCK), **176–177**
  - anxiety disorders and, 1743*t*, 1749–1750
  - in basket cells, 17
  - localization of, 96
  - receptor-antagonist studies, 361*f*
  - satiation and, 365*f*, 368*f*
  - secretion of, 360–361
- Cholecystokinin-4 (CCK-4) panic challenge, 1746
- Cholestanol, accumulation of, 662
- Cholesterol
  - anxiety and, 2545

- depression and, 2545
- metabolism of, 655
- synthesis of, 660

- Choline, beneficial effects, 2548
- Cholinergic, definition of, 4568
- Cholinergic anti-inflammatory reflex, 187
- Cholinergic neurons, 1714–1715
- Cholinergic receptors, 1495–1496
- Cholinergic system, 3009
- Cholinergic–noradrenergic imbalance hypothesis, 1608
- Cholinesterase inhibitors
  - AD, **3039–3049,** 3040*t*
  - after TBI, 529
  - Alzheimer disease treatment, 1207–1209
  - antidementia drugs, 4142, 4142*t*
    - adverse effects, 4144
    - clinical applications, 4144
    - contraindications, 4144
    - donepezil, 4143
    - galantamine, 4143–4144
    - rivastigmine, 4143
    - tacrine, 4142
    - treatment duration, 4144–4145
  - chemistry, 3041, 3041*f*, 3042, 3042*f*, 3044, 3044*f*, 3046, 3046*f*
  - dementia treatment, 1238
  - dementia with Lewy bodies treatment, 1217
  - dosage, 4143–4145
  - dosage/administration, 3042, 3044, 3046, 3048
  - drug interactions, 3042, 3044, 3046, 3048
  - drug studies, 3041, 3043, 3045, 3047
  - future directions, 3049
  - hepatic monitoring/rechallenge, 3042
  - history, 3039–3041, 3040*t*
  - irreversible, 3040
  - laboratory interferences, 3042, 3044, 3046, 3048
  - memantine and, 3048–3049
  - pharmacological actions, 3041–3047

- precautions/adverse effects, 3042, 3044, 3046–3047
- protective action in TBI, 524
- pseudoirreversible, 3040
- reversible, 3040
- therapeutic indications, 3042, 3044, 3046–3047

- Chomsky, Noam, 3359
- Chorea, 553, 4568. *See also* Huntington disease
- Chorea-acanthocytosis (ChAc), 667
- Choreic movements, 458
- Choreiform disorders, 555*t*
- Chorein, function of, 667
- Choreoathetosis, 553
- Choroid plexus, 14
- Christal, Raymond, 935
- Christianity
  - abused women and, 2593
  - angels and, 2590
  - mysticism and, 2578
  - religious life and, 2592
  - suffering and, 2588
- "Christmas tree effect," 4485–4486
- Chromaffin cells, 683–684
- Chromatin
  - active, 394
  - ATP-dependent remodeling, 391–392
  - description of, 149
  - domains, 393–394, 393*t*, 394*f*
  - modifications of, 389*t*
  - null, 394
  - permissive, 389*f*
  - polycomb-repressed, 394
- Chromatin immunoprecipitation (ChIP), 439
- Chromosomes, 149
- Chronesthesia, 2140
- Chronic cough disorder, 2228–2229
- Chronic demoralization, 1647
- Chronic factitious disorder with predominantly physical signs and symptoms, 1846. *See also* Factitious disorder

- Chronic fatigue syndrome (CFS)
  - description of, 4568
  - etiology of, 597
  - mood disorders and, 1657–1658
  - MS *versus,* 563
  - presentation, 597
  - risk factors, 598
- Chronic idiopathic pruritus, 2347–2348
  - etiology of, 2348
  - treatment, 2348–2349
- Chronic illnesses
  - acceptance of, 2205–2206
  - behavioral health conditions comorbid with, 4320
  - care management/case management of, 4312
  - care models
    - for behavioral health services, 4322
    - evidence-based, 4321
  - CMS reimbursement for, 4321
  - coping with, 2333
  - skin lesions, 2333
- Chronic inflammatory demyelinating polyneuropathy (CIDP), 630
- Chronic medical conditions
  - child/adolescent
    - family therapy and, 3786
    - group therapy and, 3769–3770
- Chronic meningitis, 472
- Chronic neurotic encopresis, 3658, 3672
- Chronic obstructive pulmonary disease (COPD), 2226
- Chronic paroxysmal hemicrania, 620
- Chronic somatic anxiety, 2171
- Chronic stress, 322, 1182
- Chronic traumatic encephalopathy (CTE), 528, 530
- Chronic wasting disease (CWD), 601*t*
- Chronobiology, **204–216**
  - definition of, 204
  - description of, 4568
- Chronological age (C.A.), 4568. *See also* Intelligence quotient

- Chronological Assessment of Suicidal Events (CASE), 955
- Chronophobia, 4568
- Chronotherapy, 2098
- Chr17p-Miller-Dieker syndrome, 49
- Chused, Judith, 2654
- CI-988, 1750
- Cibacalcin. *See* Calcitonin
- CIDI. *See* Composite International Diagnostic Interview
- Ciencephalon, 8
- Cigarette smoking, cancers and, 2267
- Cigarettes. *See* Tobacco use
- Ciliary neurotrophic factor (CNTF), 100–101
- Cimetidine (Tagamet)
  - caffeine metabolism and, 1295
  - histamine receptor blocking by, 75
  - tacrine and, 3042
- Cingulate cortex, 277
- Cingulate gyrus, 31, 33*f*
- Cingulo-opercular networks, 418
- Cingulotomy
  - anterior, 3316, 3318
  - benefits of, 450
- Cingulum
  - association fiber tracts and, 15*f*
  - connectivity, 451
- Ciproflaxin, ECT interactions with, 3294–3295
- Circadian clocks
  - pacemakers, 204–205
  - resetting of, 204, 207–208
  - seasonality, 211–212
  - sensory parameters, 207–208
  - shift work and, 209
- Circadian rhythm sleep disorders, 214, 4568
- Circadian rhythm sleep-wake disorders, 3937, 4568
- Circadian rhythms. *See also* Biological rhythms
  - depression and, 213–214
  - description of, 204

- dysregulation, depressive syndrome and, 1638
- GH release inhibition and, 1716
- of hunger, 358*f*
- hunger and, 356
- hypothalamus and, 1714
- mood disorders and, 1609
- pharmacotherapy and, 215
- psychiatric disorders, 212–213
- sleep and, 1716, 2339–2340
- sleep regulation and, 348
- sleep-wake disorders, 2097–2099, 4055, 4055*t*

- Circle of Being, 402*f*
- Circuit modeling, biophysically-based, 428–431
- Circular insanity, 1603
- Circumcision, 1954
- Circumstantiality, 4568
- Circumstantiality of thought, 1120
- CIS. *See* Columbia Impairment Scale
- Cis-regulatory sequences, 159–160
- Cisapride (Propulsid), 3131, 3661, 3675
- *CISD2* gene, 670
- Cisgender, definition of, 1985, 3363
- Cisplatin (Platinol-AQ), 2255, 2272
- Cisterna magna, 14
- Citalopram (Celexa)
  - for aggressive behavior
    - in IED, 2501
    - intermittent explosive disorder, 2501
  - for anxiety disorders, 1779–1780, 1783, 4129
  - for bipolar depression, 1691
  - breast milk excretion of, 3168
  - bupropion and, 3005, 3010
  - buspirone and, 3009, 3010
  - chemistry, 3161
  - for depressive disorders, 3678, 3692
  - dosing/administration, 488, 573*t*, 3166
  - drug-drug interactions, 2261, 3164, 3165

- for OCD, 1789
- pharmacokinetics, 3164, 3164*t*, 3165
- pharmacological actions, 3161–3162
- for PLAC after stroke, 488
- pregnancy and, 3167
- for schizophrenia, 4098
- sexual dysfunction and, 1779*t*
- sexual urges and, 2058
- side effects, 573*t*, 1679*t*, 2199
- targets for, 69
- tic disorder, 3646, 3660
- tolerability, 3157, 3158

- *Citrus aurantium,* 2554
- Civil cases, dissociative disorders and, 1874
- Civil commitment
  - description of, 4568
  - restrictive, criteria, 4389–4390
- Civil Rights of Institutionalized Persons Act (1980), 4336–4337
- *Civilization and Its Discontents* (Freud), 899
- CJD. *See* Creutzfeldt-Jakob disease (CJD)
- Claims reviews, 4568
- Clairvoyance, 4568
- Clang association, 1120, 4568
- Clanging, description of, 4568
- Clarification, psychoanalysis and, 2655
- Clark, David M., 1776
- Clark, Hillary, 2833
- Clarke, J. Christopher, 2721
- Clarkin, John, 2674
- Classic hallucinogens, 1315–1316, 1320, 1321. *See also specific hallucinogens*
- Classic migraine. *See* Migraine
- Classic suggestion effect, 2708
- Classical conditioning. *See* Conditioning
- Classical mechanics, MRI/MRS, 261–262
- Classification of, 1173
- Classification of mental disorders, 4542, 4544, 4547
- Claustrophobia, 4568

- Cleanness, excessive, 2138
- Cleft lip, associated conditions, 455–456
- Clenched-fist technique
  - competing response training and, 2698, 2721–2722
  - hypnotherapy and, 2716–2717
- Client-centered psychotherapy, 2818–2819
  - description of, 4568
  - ethical issues, 2819
  - indications, 2819
  - research and evaluation, 2819
  - techniques, 2818–2819
  - theoretical issues, 2818
- *Client-Centered Therapy* (Rogers), 932, 4539
- Climacteric, 4568
- Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE), 1416, 1507, 1510, 1523, 1540, 1543, 1544*f*, 1545*f*, 2868, 2917, 2918
  - FGA, 3069–3070
  - quetiapine, 3069–3070
  - risperidone, 3069–3070
  - second-generation antipsychotic medication trials, 4351–4352
  - ziprasidone, 3069–3070
- Clinical Dementia Rating Scale (CDR), 1196
- Clinical Frequencies Recording System (CFRS), 4372
- Clinical Global Impression Scale for Somatization Disorder (CGI-SD), 1837
- Clinical inertia, 2236, 2237
- Clinical interviews
  - for ADHD, 3600, 3601*t*
  - child
    - bipolar disorder, 3688, 3702
    - maltreatment, 3836
  - during child psychiatric evaluations, **3375–3385**
    - child interview during, 3381–3385, 3383*t*
    - considerations during, 3375–3376
    - context considerations, 3376
    - parent interview, as part of, 3376–3381, 3411
    - preliminary contact for, 3376
    - techniques for, 3380
  - of everyday functioning, 4030–4031

- parent, child maltreatment and, 3834
- for psychiatric diagnosis, 1115

- Clinical neuropsychology and intellectual assessment of adults
  - referral issues
    - differential diagnoses, 985–986, 985*t*, 986, 987
- Clinical neuropsychology and intellectual assessment of adults,
  - referral issues
    - differential diagnoses, **976–994**
- Clinical neuropsychology and intellectual assessment of adults, **976–994**
  - assessment, 987–988
    - battery approach, 988
    - hypothesis testing, 988
    - mental status examination, 988
    - screening, 988
  - formal domains, 988–993, 989*t*
    - attention and information processing, 990
    - executive functions, 992
    - intellectual functioning, 990
    - interview, 990
    - language, 991
    - memory, 990–991
    - performance validity, 992–993
    - psychological factors, 992
    - therapeutic discussion, 993
    - visuospatial function, 991–992, 991*f*
  - history, 977–978, 977*f*
  - neuroanatomical correlates, 978–982
    - arithmetic skills, 979
    - executive function, 981–982
    - hemispheric dominance, 978–982
    - language disorders
      - Broca aphasia, 978, 978*t*
      - conduction aphasia, 979
      - global asphasia, 979
      - Wernicke aphasia, 979
    - limb apraxia, 979
    - memory disorders

        - encoding, 981, 981*f*
        - storage and retrieval, 981
    - motor skills, 982
    - spatial disorders
        - dressing apraxia, 980
        - neglect syndromes, 980
        - visuospatial impairment, 979–980, 980*f*
  - practitioners, 977
  - psychometric issues
    - base rates, 982–983
    - moderating factors, 984
    - normative data, 983
    - predictive utility, 982
    - premorbid function, 983–984, 984*f*
    - reliability, 982
    - validity, 982
  - referral issues
    - level of functioning, 984
  - *vs.* behavioral neurology and neuropsychiatry, 977
  - *vs.* clinical psychology, 977
- Clinical/policy-making initiatives, 4364
- *Clinical Practice Guidelines We Can Trust* (IOM), 971
- Clinical psychologists, 4568
- Clinical psychology, 4568
- Clinical Quality Measure (CQM) data, 4320
- Clinical rationale for hospitals, 4329–4332
- Clinical research, health services research *vs.,* 4347, *4348f*
- Clinical trials. *See also specific Medications; specific Trials*
  - epidemiology of, 815
  - minority under-representation in, 4382
  - randomized, 836–839
    - causality, 836, 838
    - ethical issues, 2278
    - matching and randomization, 836–837
    - missing or incomplete data, 838–839
    - propensity scores, 837–838
    - randomization tests, 837

- shortcomings of, 4370
- Clinically relevant behaviors (CRBs), 2700
- Clinician-Administered Dissociative States Scale (CADSS), 1877*t*, 1882, 1897
- Clinician training, cultural responsiveness, 4383
- Clinician–educators, 4359
- Clinton administration, 4293, 4311
- Clitoris
  - orgasm and, 1954
  - sexual response cycle, 1959*t*
- Clobazam (Frisium)
  - chemistry, 2992
  - receptor ligands, 2992
  - seizure disorder, 2997
- Clock drawing, 991, 991*f*
- *CLOCK* gene variants, 213
- *Clock* genes, *Drosophila,* 206
- Clock protein interactions, 206–207
- Clomiphene, 1980
- Clomiphene citrate (Clomid), 2510
  - function of, 1980
  - in PCOS treatment, 2249
- Clomipramine (Anafranil)
  - BDD, 1792
  - chemistry, 3254, 3254*f*
  - depersonalization/derealization disorder, 1945
  - depression, 1681
  - dissociative identity disorder, 1927*t*, 1928
  - dosage/administration, 573*t*, 1979, 3265, 3265*t*
  - for excoriation disorders, 2352
  - narcolepsy, child/adolescent, 3701, 3792
  - OCD treatment, 1785, 1789, 3646, 3660
    - child/adolescent, 3701, 3792
  - organ/system effects of, 3260
  - for panic disorder, 2811–2812
  - pharmacological actions, 3256, 3257, 3258
  - sexual dysfunction and, 1972, 1979
  - side effects, 573*t*

- stereotypic movement disorders, 3672, 3686
- targets for, 69
- therapeutic indications, 3261
- tic disorder, 3646, 3660
- TTM, 1795

- Clonal lines, description of, 181
- Clonazepam (Klonopin)
  - agitation and violence, 2998
  - antimanic therapy, 519
  - BD, 2997
  - benzodiazepine dose equivalent, 2633
  - bipolar disorders, 1699
  - dementia with Lewy bodies treatment, 1218–1219
  - for detoxification, 1368, 1387
  - dissociative identity disorder, 1927*t*, 1928
  - dosage/administration, 635
  - drug interactions, 3164
  - for myoclonus, 555
  - panic disorder, 1699
  - for SAD, 1783, 1784
  - for stiff-person syndrome, 635
  - for tic disorders, 3646–3647, 3660–3661
- Clonic convulsions, 4568
- Clonic tics, 3635, 3649
- Clonidine (Catapres), **2945–2950**
  - absorption, 2945
  - ADHD, 2947–2948, 3594–3595
    - child/adolescent, 3789, 3938
  - administration, 2950
  - adverse effects, 2949–2950
  - for aggressive behavior
    - in ADHD, 2500
    - in conduct disorder, 2501
    - intermittent explosive disorder, 2501
  - anxiety disorders and, 1742
  - autism spectrum disorders and, 2948–2949, 3586
  - BPCA and, 2931

- chemistry, 2945, 2945*f*
- classification, 2918
- cognition and, 2949
- depression and, 1608
- dissociative identity disorder, 1927*t*, 1928
- distribution, 2945
- dosage, 2950
- drug interactions, 2950
- drug studies, 2946
- elimination, 2946
- growth hormone and, 1716
- hallucinogen persisting perception disorder treatment, 1324
- history, 2945
- for hypertension, 67
- laboratory interference, 2950
- metabolism, 2946
- naltrexone detoxification, 2948, 3141
- neuroprotective effects, 2949
- nicotine withdrawal, 2948
- opioid withdrawal, 2381, 2948
- panic disorder and, 1742
- pharmacodynamics, 2946
- pharmacokinetics, 2945–2946
- pharmacological actions, 2945–2946
- precautions, 2949–2950
- PTSD, 2949
- side effects of, 571
- sleep disorders, child/adolescent, 3794, 3936
- specific organs/systems and effects of, 2946
- therapeutic indications, 2947–2949
- for tic disorders, 2947, 3645, 3647, 3659, 3661, 3795
- for TS, 2947
- for withdrawal treatment, 1368, 1388, 2949–2950

- Cloninger, C. Robert, 2593
- Cloninger's Temperament and Character Inventory, 935
- Clopenthixol (Soridol), 3064
- Clorazepate (Tranxene)

- absorption, 2993
- alcohol withdrawal, 2997
- metabolism, 2993

- Clorgyline, 3111–3112, 3113*f*
- Closed testing procedure, 834
- *Clostridium difficile,* diarrhea and, 2296
- Clouding of consciousness, 4568
- Clozapine (Clozaril)
  - acute mania, 1686, 1689*t*, 1690
  - adverse effects of, 2356*t*
  - for aggressive behavior
    - MR children, 2500
    - in schizophrenia, 2501
  - for bipolar depression, 1695
  - for bipolar disorder, 1696*t*, 1697
  - black box warning, 2919
  - blood levels, 3081
  - chemistry, 3064
  - clinical use of, 4352
  - cognitive functioning and, 168
  - combination pharmacotherapy, 3327–3328
  - delirium treatment, 1190
  - for dementia, 1201
  - diabetes risk and, 2247
  - dissociative identity disorder, 1927*t*, 1928
  - drug interactions, 3000, 3131, 3164
  - EEG effects, 253
  - efficacy of, 2169
  - FGAs *v.*, 3066, 3066*t*
  - histamine receptor blocking by, 75
  - history, 1520, 2918, 3060
  - lithium combined with, 3327
  - mechanism of action, 2169
  - movement disorders induced by, 2942
  - myocarditis related to, 2195
  - Parkinson disease treatment, 1220
  - pharmacological actions, 3164

- and physical health monitoring, 1022–1023
- for psychosis, 548
- psychosis, acute, 3078–3079
- receptor profile of, 67
- for schizophrenia, 3793
  - neuroimaging studies of, 1497
  - side effects, 1527
  - suicide prevention, 1528
- seizures and, 519
- side effects of, 67, 1527, 2169, 2189
  - weight gain, 2237, 2237*t*
- suicide prevention and, 2621
- valproate and, 3278
- weight gain and, 2215, 2237, 2237*t*

- Clozaril. *See* Clozapine
- Clumsy child syndrome. *See* Developmental coordination disorder
- Cluster headaches
  - definition of, 620
  - diagnostic criteria, 620*t*
  - etiology of, 622–623
  - prevalence of, 621
  - treatment, 626
- Cluster suicides, 4568
- Cluttering
  - description of, 4568
  - differential diagnoses, 3563
- CNS. *See* Central nervous system (CNS)
- CNS effects of, 1267
- CNVs. *See* Copy number variants
- CO. *See* Carbon monoxide
- Co-consciousness, 1911
- Co-therapist relationships, 3763
- Co-therapy
  - couples and, 2758–2759
  - description of, 4571
  - families and, 2758–2759
  - group psychotherapy and, 2745

- Cocaine, 1280–1282, 1284, 1289–1290, 3367. *See also* Stimulant-related disorders
  - addiction to, 397–398
  - aggression threshold and, 2492–2493
  - aggressive behavior and, 2492–2493
  - bipolar disorder and, 1699
  - brain development and, 58
  - dependence
    - psilocybin and, 2887, 2892
    - STDP in, 2811
  - depersonalization/derealization, 1891
  - description of, 4568
  - dopamine reuptake and, 63–64
  - intoxication, 1713
    - psychiatric emergency due to, 2629
  - neurocentrism and use of, 783
  - prenatal exposure to, 2484
  - use disorders, 385
  - withdrawal, 2633
- Cocaine and methamphetamine regulated transcript (CART), 89
- Cocaine-related disorders
  - treatment
    - anticonvulsants, 2973
    - disulfiram, 3053–3054
    - modafinil, 3232
    - naltrexone, 3142
    - nefazodone, 3127
- Coccaro, Emil, 2494
- Cochran–Mantel–Naenszel statistic, 842
- Codeine, 2377, 2395*t*
- Coding genes, 149–151
- Codon usage bias, 151
- Codons, function of, 149–150
- Coefficients of correlation, 4568
- Coercive persuasion. *See also* Brainwashing
  - diagnosis, and clinical features, 1946–1948
  - treatment, 1948
- Coexistent cultures, definition of, 4568

- Cogentin. *See* Benztropine
- Cognex. *See* Tacrine
- Cognition. *See also* Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) study; Cognitive disorders; Mild cognitive impairment (MCI); Neurocognition; Sociocognitive model (SCM)
  - ADHD treatment effects, 3593
  - and areas cerebral cortex, 977–978, 977*f*
  - assessments for (*See* Cognitive assessments)
  - in autistic disorders, 712
  - and brain function, 744–745
  - clonidine and, 2949
  - conduct disorder, 3615–3616
  - definition of, 4568
  - delusions and, 1415
  - depressive disorders, child/adolescent, 3676, 3690
  - development of
    - adolescent, 3369
    - in childhood, 3356–3357*t*, 3359–3360
    - Piaget theories (*See* Piaget, Jean)
  - executive dysfunction and, 463–464
  - guanfacine and, 2949
  - intellectual disability and (*See* Intellectual disability)
  - language disorders and, 3546
  - and mental health, 760
  - models of
    - connectionist, 431–434
    - parallel distributed processing, 432
    - psychiatric neuroscience, 431–434
  - OCD and, child/adolescent, 3698
  - at patient interview, 955
  - processing, pain and, 328
  - PTSD and, child/adolescent, 3704, 3705
  - racial prejuduce and, 701–702
  - in referrals, 984
  - schizophrenia and, 1454–1455, 1568–1569
    - fMRI studies, 1480–1482, 1483*f*
  - social phobias and, 702–703
  - specific learning disorder and, 3525–3526

- stimulant use effects on, 1284
- synapses and, 430–431
- temperament and, 3488

- Cognitive activation, neuroimaging, 1760
- Cognitive assessments, child, **3413–3429**
  - achievement testing, 3425–3426
  - applications for, 3422–3429
  - cognitive tests, description of, 3415–3422, 3416–3417*t*, 3418*t*
  - intellectual testing, 3422–3425
  - legal considerations for, 3415–3417, 3419–3421*t*
  - overview, 3413–3415
  - of processing problems, 3426–3429
- Cognitive/behavioral activation paradigms, 1759
- Cognitive Behavioral Analysis System of Psychotherapy (CBASP), 1665, 2700
- Cognitive-behavioral group therapy (CBGT)
  - anxiety disorders, 2746
  - SAD treatment, 1776
  - social anxiety, 3122
- Cognitive-behavioral therapy (CBT), **4179–4192**
  - for acne, 2345
  - for acrophobia, 79
  - adapting for older adults
    - age-related challenges, 4179
    - cognitive impairment, 4179–4180
    - culture and age cohorts, 4180–4181
    - interpersonal events, 4180
    - medical problems, 4180
    - role changes, 4180
    - transportation barriers, 4179
  - for ADHD, adult manifestation, 3603
  - for aggressive behavior, 2498, 2498*t*
    - ADHD and, 2500
    - intermittent explosive disorder, 2501
    - intimate partner violence, 2502
    - mental retardation and, 2499
    - schizophrenia, 2501
    - sex offenders, 2503

    - TBI, 2502
  - for agoraphobia, 1774, 1775, 1783
  - for alcohol use disorders, 1276
  - for antisocial characteristics, 4392–4393
  - for anxiety
    - perinatal, 2515
  - for anxiety disorders, 1768–1778
    - acceptance-based techniques, 1773
    - assessment, 1771
    - cognitive therapy, 1773
    - efficacy, 1774–1776
    - emotional processing and, 1769–1771
    - exposure therapy, 1772–1773
    - future directions, 1776–1777
    - interpersonal skills training, 1774
    - pharmacotherapy and, 1782
    - procedures, 1771–1778
    - psychoeducation, 1771–1772
    - stress inoculation training, 1773
    - symptoms and conditions, 1768–1769
    - treatment planning, 1772
  - applying research to clinical practice, 1704
  - for assault victims, 2541
  - for atopic dermatitis, 2342
  - for BDD, 1792
  - for binge eating disorders, 2220
  - for bipolar depression, 1692
  - for bipolar disorders, 1684, 1696, 1703, 1703*t*, 1706, 1707
    - social anxiety disorder and, 1698
  - for body dysmorphic disorder, 2354
  - as brief psychotherapy, 3742–3745
  - for bulimia nervosa, 2081
  - in burn injuries
    - intermediate phase, 2379
  - for caffeine use disorder, 1302
  - for cannabis dependence, 1310
  - child/adolescent, **3749–3759**

- anxiety disorders, 3710
- behavioral assessment, 3752–3753, 3752*t*, 3753*t*
- clinical issues, 3756–3759
- cognitive interventions, 3754, 3754*t*, 3755*t*
- complications, 3757–3758
- contraindications, 3758
- definition, 3749
- ethical issues, 3759
- exposure/behavior change techniques, 3755
- future directions, 3759
- goals, 3759
- history, 3749–3751, 3749*t*
- indications, 3756–3757
- limitations, 3757
- obsessive-compulsive disorder (OCD), 3758–3759
- for OCD, 3700
- overview, 3814
- parental involvement, 3755–3756
- psychoeducation, 3753, 3754*t*
- relapse prevention, 3756, 3757*t*
- selective mutism (SM), 3758
- social anxiety disorder, 3758
- somatic management techniques, 3753–3754
- techniques of, 3751–3756
- theoretical issues, 3751, 3751*t*

- clinical issues
  - contraindications, 4191
  - goals, 4190
  - homework assignments, 1700–1701, 4183*t*, 4190
  - indications, 4191
  - individual *vs.* group format, 4190
  - medications, 4189–4190
  - therapy *vs.* remediation, 4191
- couples, 2750, 2751*t*, 2759
- definition of, 3749, 4179
- depersonalization/derealization disorder, 1945
- for depression, 1606, 1664–1665, 1681

- future directions, 1667
- hypnosis and, 2723
- in PD, 548
- depressive disorders, 3679, 3693
- description of, 4568
- dissociative identity disorder, 1924
- efficacy research, 2695
- evaluation, 2862–2863
- excoriation disorder, 1797
- family, 2750, 2751*t*
- in functional GI disorders, 2207
- gambling disorder, 1808, 1809
- group, 4202
  - anxiety disorders, 2746
  - SAD, 1776
- hallucinogen persisting perception disorder treatment, 1324
- history, 2683, 3749–3751, 3749*t*
- homework assignments, 1700–1701, 4183*t*, 4190
- for IBD, 2201–2202
- for IBS, 2202
- for insomnia, 2088
- for intellectual disability, 3514
- MDD, 1702*t*, 1704–1706
- mindfulness meditation and, 1701
- neurobiological effects of, 2814*t*
- obsessive-compulsive disorder and, 3700
- for OCD and related disorders, 1699, 1788, 1789, 2813
- for OCRDs, 1787
- opioid use disorder treatment, 1370
- in pain syndromes, 1841, 2396–2397
- for panic disorders, 1747, 1774–1775
- for phobias, 1774
- for PNES, 651
- for psoriasis, 2344
- psychodynamic theory and, 1665
- for psychosis, 2881–2884
  - A-B-C model, 2883

    - clinical vignette, 2867, 2885–2886
    - eliciting patient's story, 2882–2883
    - engaging the patient, 2882
    - ethical issues, 2884–2885
    - future directions, 2885
    - introducing coping mechanisms, 2883
    - introduction, 2866–2868
    - maintaining contact, 2884
    - normalizing model, 2883
    - reformulating life narrative, 2884
    - research and evaluation, 2885
    - therapeutic alliance, 2883–2884
  - for PTSD
    - child/adolescent, 1823, 3706
    - complex and dissociative, 1919–1920
    - exposure-based, 1824–1825
  - for pyromania, 2113
  - RCT evaluations
    - group and couples, 1703
    - individual, 1702
  - in rehabilitation, 4373
  - research and evaluation
    - anxiety, 4191
    - depression, 4191
    - schizophrenia, 4192
  - for SAD, 1784
  - schema therapy and, 2668–2669
  - for schizophrenia, 1552
  - for separation anxiety, 3710–3711
  - for separation anxiety disorder in children, 1784
  - in sexual addiction, 2054–2055
  - for smoking cessation, 2730
  - for social anxiety, 3717
  - for somatic symptom disorders, 1837, 1843–1844
    - pivotal study, 1837
    - unspecified and below-threshold, 1843
  - for stimulant-use disorders, 1289

- for stress management, 2325
- SUD treatments, 384, 384*t*
- suicide prevention and, 2621
- suicide risk and, 2617
- techniques
  - assessment and initial treatment, 4181
  - middle stage intervention
    - behavioral intervention, 4182–4183
    - cognitive intervention, 4183–4184, 4183*t*
    - mindfulness and acceptance, 4185
    - psychoeducation, 4182, 4182*f*
    - schema change, 4185
    - Socratic questioning, 4185
    - termination and relapse prevention, 4186
  - modifications for psychosis
    - behavioral interventions, 4187
    - cognitive interventions, 4187–4189, 4188*t*
    - cognitive model of schizophrenia, 4186
    - psychoeducation, 4186–4187
  - stages of, 4181
- for tension headaches, 626
- terrorist attacks and, 2572
- theoretical issues, 2869–2870, 2869*t*
- third wave therapies and, 2700
- for tic disorders, 3643, 3657
- for trichotillomania, 2353
- weight loss using, 2217

- Cognitive Bias Modification. *See* Attention Bias Modification (ABM)
- Cognitive decline, MCI and, 1222–1223, 1222*t*, 1226–1227
- Cognitive deficits. *See also* Cognition
  - after TBI
    - course/prognosis, 539
  - COPD and, 2226
  - in hypothyroid patients, 680
  - modeling of, 433–434
- Cognitive development, 4569
- Cognitive disorders. *See also* Cognition

- after TBI, 528
- amnesia in
  - dissociative amnesia *versus,* 1935
- amnestic disorders, **1229–1232**
- anxiety disorders
  - due to another medical condition, **1253–1255**
  - substance/medication-induced, **1253–1255**
- asthma and, 2225
- catatonic disorder due to another medical condition, **1258–1259**
- delirium, **1178–1191**
- dementia, **1191–1121**
- demographics and, 1177
- due to another medical condition, **1236–1238**
- future research for, 1185
- major or mild vascular neurocognitive disorder, **1238–1242**
- mood disorders
  - due to another medical condition, **1248–1251**
  - substance/medication-induced, **1248–1251**
- neurocognitive disorders
  - due to another medical condition, **1246–1247**
  - due to HIV infection, **1244–1246**
  - due to traumatic brain injury, **1242–1244**
  - substance/medication-induced mild/major, **1247–1248**
- obsessive-compulsive and related disorder
  - due to another medical condition, **1255–1256**
  - substance/medication-induced, **1255–1256**
- overview, 1177–1178
- personality change due to another medical condition, **1259–1261**
- psychotic disorders
  - due to another medical condition, **1251–1253**
  - substance/medication-induced, **1251–1253**
- substance/medication-induced sexual dysfunction, **1258**
- substance/medication-induced sleep disorder, **1256–1258**
- in transplant populations, 2367–2369

- Cognitive dissonance, 4569
- Cognitive distortions. *See also specific conditions*
  - mania, 1641

- negative beliefs as, 2695, 2762*f*
- OCD, 1790

- Cognitive disturbances, depressive syndrome and, 1637–1638
- Cognitive dysfunction, psychosis and, 1568
- Cognitive function. *See also* Cognition; Executive function; Intellectual functioning
  - antidementia drugs for impairment of, 4149
  - borderline, 2430–2442
  - burn injuries and, 2374
  - and clinical disorders of aging brain, 3988, 3993
  - and cognitive-behavioral therapy, 4179–4180
  - in critical illnesses survivors, 2330–2331
  - effects of abuse, 2426
  - effects of prematurity, 643
  - and electroconvulsive therapy, 4155
  - everyday functioning, 4027
  - game technology, 4278–4279
  - in group therapy, 4203, 4205
  - and HIV infections, 4269, 4272
  - malingering and, 2407
  - of older adults, 3958, 3975, 3999–4001
    - CAIM medicine, 3979–3982, 3979*t*
    - clinical disorders, 3988, 3993
    - executive function, 4000
    - language functions, 3999
    - learning and memory, 3999–4000
    - models of, 4000–4001
    - processing speed, 4000
    - visuoperception, 3999
  - post-critical illness issues, 2331
    - interventions, 2331
    - physical dysfunction, 2331
  - in schizophrenia, 4096, 4099
  - screening of, 4004
  - and sleep disorders, 4057
  - and testamentary capacity, 4438
- Cognitive impairment. *See also* Mild cognitive impairment (MCI)
  - benzodiazepine, 2998

- bipolar disorders and, 1682
- cannabis-related, 1308
- dementia and, 1193*t*, 1194, 1195–1196, 1195*t*
- detection of, 656
- lithium and, 3096
- in multiple sclerosis patients, 563
- psychiatric patients
  - brain tumors and, 493
- stroke and, 484*f*
- in transplant populations, 2367–2369
- treatment
  - cholinesterase inhibitors, 3041
  - modafinil, 3232

- Cognitive interventions, 3754, 3754*t*, 3755*t*
- Cognitive Interviews, 2734
- Cognitive mentalizing, affective mentalizing *versus,* 2896
- Cognitive model, 2760, 2760*f*. *See also* Cognitive therapy
  - generic, 2763
- Cognitive neuroscience
  - emergence of, 432
  - models of, 431–434
- Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia (CNTRICS), 4370
- Cognitive-perceptual distortions, 2173
- Cognitive performance, sleep restriction and, 349
- Cognitive processing therapy (CPT)
  - development of, 2696
  - PTSD, 1823
  - terrorist attacks and, 2572
- Cognitive rehabilitation, survivors of brain tumors, 505*f*
- Cognitive remediation therapy (CRT)
  - community/occupational function, 4372–4373
  - early-onset schizophrenia treatment, 3732–3733
  - neurocognitive impairment treatment, 1510
  - schizophrenia, 1552–1553
- Cognitive restructuring
  - cognitive therapy and, 2696
  - stress management and, 2326, 2326*t*

- Cognitive retraining
  - for dementia, 1199
  - in traumatic brain injury, 1244
- Cognitive style, defined, 1119
- Cognitive symptoms
  - childhood brain tumors, 505*f*
  - review of, 455
- Cognitive theory
  - depersonalization/derealization, 1943
  - depression, 1606, 1664–1665
  - major depressive episodes (MDEs), 1664–1665
- Cognitive therapy (CT), **2760–2775**. *See also* Mindfulness-based cognitive therapy
  - anxiety disorders, 1773
  - applications, 2764–2768
  - belief disconfirmation and, 1771
  - bipolar disorders, 2765
  - clinical issues, 2773–2774
    - complications, 2773
    - contraindications, 2773–2774
    - ideal clinicians, 2773
    - ideal patients, 2773
    - indications, 2773
    - limitations, 2773
  - cognitive conceptualization diagram, 2763–2764, 2764*f*
  - components and goals, 2695–2697
  - coping strategies, 2763
  - core beliefs, 2762–2763, 2770–2771
    - categories of, 2762*t*
    - worksheet, 2771*f*
  - depression, 2764–2765
  - description of, 4569
  - dissociative amnesia, 1939
  - dissociative disorders, 1869
  - eating disorders, 2766–2767
  - ethical issues, 2774
  - evaluation, 2861–2862
  - GAD, 2765

- generic cognitive model and, 2763
- group psychotherapy and, 2738
- history, 1701, 2761
- for insomnia, 2089–2090
- intermediate beliefs, 2763
- introduction, 2760–2761
- modes and, 2763
- obesity, 2766–2767
- OCD, 2765–2766
- panic disorder, 2765
- personality disorders, 2768
- research and evaluation, 2774
- research findings on, 1702
- SAD, 2766
- schemas and, 2763
- schizophrenia, 2767–2768
- substance use disorders, 2767
- techniques, 2768–2773
  - behavioral, 2772
  - biological interventions, 2773
  - cognitive, 2770–2771, 2770*f*, 2771*f*
  - emotional regulation, 2772
  - environmental intervention, 2772
  - goal-setting, 2769–2770
  - imaginal, 2771–2772
  - interpersonal, 2772–2773
  - motivational, 2772
  - structuring, 2769, 2769*f*
  - therapeutic relationship, 2768–2769
  - treatment adherence, 2773
- theoretical issues, 2761–2768
- treatment goals, 2774
- TTM, 1795

- *Cognitive Therapy and the Emotional Disorders* (Beck), 4546
- Cognitive triad, 1664, 2695, 4569
- Cohen, Mandel, 1830
- Cohen–Mansfield Agitation Inventory, 3976

- Cohesion. *See* Group cohesion
- Cohesiveness, group psychotherapy and, 3761
- Cohorts
  - and cognitive-behavior therapy, 4180–4181
  - description of, 4569
  - and HIV infection, 4270
  - mental health in older adults, 3963
  - studies, epidemiology of, 815
- Coitus. *See also* Intercourse
  - first, 1962
- Coitus interruptus, 1953, 4569
- "Cold turkey," 4569
- Cole, Jonathan, 1713
- Coles, Robert, 2836
- Collaborative Assessment and Management of Suicidality (CAMS), 2617
- Collaborative care
  - as evidence-based model, 4321
  - in primary care, 4322, 4349
- Collaborative empiricism, 1701
- Collaborative Psychiatric Epidemiology Surveys (CPES), 3960
- Collaborative relationship with patients, 4368
- Collaboratory care, evidence-based care and, 2183
- Collateral pathways of neuroimages, 4480
- Collateral report of everyday functioning, 4032
- *Collected Papers: Through Paediatrics to Psychoanalysis* (Winnicott), 4539
- Collective experience, 4569
- Collective unconscious, 4569
- College Mental Health, 4553
- Colom, Francisco, 1665
- Colombo, Tony, 2833
- Colonoscopy, erythema on, 2208*f*
- Colorado Client Assessment Record, 4369
- Colorectal cancer, incidence of, 2251
- Columbia Impairment Scale (CIS), 822, 3402
- Columbia Suicide Severity Rating Scale (C-SSRS), 2615, 2620
- Columbia Teen Screen, 2621
- Coma, description of, 4569

- Coma vigil, 4569
- Combat. *See also* Veterans
  - amnesia, 1886
  - dissociation and, 1873, 2718
- Combat and operational stress
  - physical symptoms, 2567–2568
- Combat fatigue, 4569
- Combination pharmacotherapy. *See* Pharmacotherapy, combination
- Combined therapy, 4569
- Combined therapy leukoencephalopathy, 2271
- Command auditory verbal hallucinations, 1412
- Command automatism
  - description of, 4569
- Command hallucinations, 4569
- Command negativism, 4569
- Commissural fibers, anatomy of, 11–12
- Commitment, description of, 4569
- Common factitious disorder, 1847
- Common migraine. *See* Migraine
- Common mode rejection ratio (CMRR), 224
- Commonwealth Fund
  - care recommendations from, 4315–4317
  - US health care system, study of, 4309
- Communication. *See also* Geriatric psychiatry
  - with adolescents, 2310
  - of child psychiatric evaluation findings, 3408–3409
  - doctor-patient relationship, 946–947
  - evidence-based approaches to, 4303
  - family role, 4194
  - guidelines, 2290–2291
  - hospitalist role in, 4355–4356, 4359
  - language development and, 3358–3359
  - psychiatric contributions, 2291
  - with severely ill children, 2308
  - sociodrama and, 2857
  - transplant candidate evaluation, 2363
  - with young adults, 2310

- Communication disorders, 3545–3570
  - child/adolescent, 3545–3570
  - description of, 4569
  - language disorder, 3545–3551
    - clinical features, 3547–3548
    - comparative nosology, 3545
    - course, 3550
    - defined, 3545
    - diagnosis, 3547–3548
    - differential diagnoses, 3548–3550
    - epidemiology, 3545
    - etiology, 3546–3547
    - laboratory examination, 3548
    - overview, 3545
    - pathology, 3548
    - prognosis, 3550
    - treatment, 3550
  - social (pragmatic) communication disorder, **3551–3555**
    - course, 3554
    - defined, 3586
    - development, 3554
    - diagnostic features, 3552–3553
    - differential diagnoses, 3553
    - epidemiology, 3554
    - etiology, 3553–3554
    - historical background, 3551–3552
    - overview, 3551
    - treatment, 3554–3555
  - speech sound disorder
    - clinical features, 3556–3557
    - comparative nosology, 3556
    - course, 3558
    - diagnosis, 3556–3557
    - differential diagnoses, 3557–3558
    - epidemiology, 3556
    - etiology, 3556
    - laboratory examination, 3557

- overview, 3555–3556
- pathology, 3557
- prognosis, 3558
- treatment, 3558

- stuttering, **3559–3565**
  - clinical features, 3562–3563
  - comparative nosology, 3559–3560
  - course, 3563–3564
  - defined, 3559
  - diagnosis, 3562–3563
  - differential diagnoses, 3563
  - epidemiology, 3560
  - etiology, 3560–3562
  - historical background, 3559–3560
  - prognosis, 3563–3564
  - treatment, 3564
- unspecified, **3565–3570**
  - clinical features, 3568
  - comparative nosology, 3566
  - course, 3369
  - diagnosis, 3568
  - differential diagnoses, 3568–3569
  - epidemiology, 3566
  - etiology, 3566–3568
  - historical background, 3565–3566
  - laboratory examination, 3568
  - overview, 3565
  - pathology, 3568
  - prognosis, 3369
  - treatment, 3569–3570
- unspecified communication disorder, **3565–3570**
  - clinical features, 3568
  - comparative nosology, 3566
  - course, 3369
  - diagnosis, 3568
  - differential diagnoses, 3568–3569
  - epidemiology, 3566

    - etiology, 3566–3568
    - historical background, 3565–3566
    - laboratory examination, 3568
    - overview, 3565
    - pathology, 3568
    - prognosis, 3369
    - treatment, 3569–3570
- Communications
  - doctor-patent, 2290
  - skills, humanism in, 2289
- Communist China, coercive persuasion and, 1946
- Communities
  - functioning
    - social cognition in, 4373
  - resources
    - for tobacco-related disorder treatment, 1349, 1349*t*
    - tobacco use disorder treatment, 1349, 1349*t*
- Community. *See also* Therapeutic communities
  - as child psychiatric evaluation consideration, 3379
  - family roles, 4194
  - and schizophrenia, 4097, 4098
- Community-based treatment
  - child/adolescent, **3803–3810**
    - approaches, 3805–3806
    - clinical issues, 3806–3807
    - defined, 3803–3804
    - ethical issues, 3807
    - evaluating, 3808
    - future directions, 3809
    - history, 3804
    - interventions, 3808
    - for minorities, 3808–3809
    - overview, 3802
    - psychiatrist and, 3806
    - research and evaluation of, 3807–3809
    - theoretical issues, 3804–3805
  - ethical issues, 3808

- for terrorism impact, 3921
- Community care
  - disparities in, 4384–4386
  - facilities, world-wide, 4524–4525
  - issues of, 4383–4384
  - public psychiatry replaced by, 4308
- Community education, as goodwill, 4345
- Community health centers, 4569
- Community mental health centers (CMHCs). *See also* Mental health care
  - as "antidote" to criminalization, 4390
  - challenges of serious/persistent mental illness, 4292
  - crisis beds at, 4344
  - funding cuts, 4325–4326
  - psychiatric hospitals interface with, 4326*f*
  - as safety net, 4380
- Community Mental Health Centers Act (1963), 2622, 2623
- Community psychiatry
  - description of, 4569
  - as reform target, 4289–4290
  - *vs.* institutional care, 4287–4288
- Comorbidities
  - anxiety disorders and, 1735, 1778, 4060, 4063
  - bipolar disorders and, 1682
  - dissociative identity disorder and, 1905
    - course and prognosis, 1918
  - and electroconvulsive therapy, 4158
  - in epidemiology, 812, 822
  - excoriation disorder and, 1796
  - factitious disorder and, 1863–1864
  - history at interview, 949
  - and HIV infection, 4271
  - hoarding disorder, 1793
  - in jails and prisons, 4467
  - mood disorders and, 1617
  - OCD and, 1786
  - PTSD and, 1822
  - rates for childhood psychopathologies, 3375–3376

- in schizophrenia, 4097
- Comparative effectiveness research, 4319
- Comparative genomic hybridization (CGH), 3452, 3456
- Comparative psychiatry, schizophrenia, 1426–1427
- Compartment models, 293–297
  - irreversible, 296–297
- Compartmentalization
  - dissociation, hypnotizability, 2710
  - dissociative disorders and, 1869
- Compassionate meditation, 2557–2558
- Compensation
  - description of, 4569
  - ethical practice of psychiatrists, 4447
- Compensation neurosis, 2406
  - description of, 2401, 4569
  - intentionality and locus of symptoms in, 1850*f*
- Competency. *See also* Consent, informed
  - of aging adult, 3976
  - and correctional psychiatry, 4454–4455
  - criminal, 4438–4439
  - definition of, 4569
  - evaluation of, 2636
  - examinations, juvenile, 3866
  - medical, 4434
  - to stand trial
    - definition of, 4569
    - restoration of, 4344
- Competency in forensic issues
  - advanced directive, 4238
  - death and dying, 4235–4236
  - driving, 4238
  - financial decisions or contract, 4236–4237
  - living will, 4238
  - medical decision, 4235
  - power of attorney, 4238
  - as research subject, 4235
  - to stand trial, 4232

- will making, 4237–4238
- Competing response training, 2698
- Competition, definition of, 4569
- Complement factor proteins, 181
- Complementarity of interactions, 4569
- Complementary, alternative, and integrative medicine (CAIM), **2542–2564**
  - anxiety disorders, 1782
  - assessment approaches, overview, 2544–2546
  - biological treatments, 2547–2556
  - definitions and historical background, 2542–2543
  - ethical issues, 2543
  - evolving conceptual foundations, 2543
  - for geriatric depression, 4074 (*See also* Geriatric psychiatry)
  - identifying appropriate modalities, 2547
  - legal issues, 2543
  - overview, 2546, 2546*t*
  - scientifically validated forms of energy and information therapies, 2558–2561
  - somatic therapies, 2556
- Complementary and alternative medicine (CAM) for cancer, 2286–2287
- Complementary DNAs (cDNAs), 158
- Complete androgen insensitivity syndrome (CAIS), 1992
- Complete delusions, 1122
- Complex, definition of, 4569
- Complex disease, genome, environment and, 2527–2528, 2528*f*
- Complex partial seizures (CPS)
  - aggression and, 2496
  - description of, 4569
  - epileptic, aggression and, 2496
  - polysomnography, 2103*f*
- Complex partial status epilepticus, 2496
- Complex regional pain syndrome (CRPS), 2391
- Complex tics, 3635–3636, 3635*t*, 3649–3650, 3649*t*
- Complexes, Jung on, 912
- Complexity theory
  - and cognitive development, 694
  - consciousness model, 2544
  - energy therapies and, 2546

- Compliance, antipsychotic drugs, 4135
- Composite International Diagnostic Interview (CIDI), 820
- Composite International Diagnostic Interview, World Mental Health Surveys version (WMH CIDI), 821
- Composite International Diagnostic Interview modified for adolescents (CIDI), 1721, 1836
- Comprehension, testing of, 461
- Comprehensive behavioral intervention for tics (CBIT), 3644, 3658
- Comprehensive rating scales, child psychiatric evaluations, 3398–3399
- Compton scattering, 291
- Compulsions. *See also* Obsessive-compulsive disorders (OCDs)
  - description of, 4569
  - obsessions and, 1786
  - stereotypic movement disorders, 3668, 3671, 3682, 3685
  - tics and, 458, 3636, 3642, 3644, 3650, 3656, 3658
- Compulsive behaviors, animal models of, 322
- Compulsive personality, 4570
- Compulsive sexual behavior, 1973
- Compulsivity
  - opioid related disorders, 1353
  - stimulant-use disorder, 1283–1284
- Computational modeling approaches
  - future challenges, 436–437
  - to psychiatry, 427–437, 428*f*, 429*f*
- Computed tomography (CT). *See also* Forensic psychiatry, neuroimaging
  - after TBI, 525
  - in aging adult (*See* Geriatric psychiatry)
  - of brain, 1019
  - brain tumor diagnosis, 494
  - in clinical neuroscience, 466
  - for delirium, 4091
  - depression and, 1608
  - description of, 4570
  - pediatric considerations, 3459
  - schizophrenia and, 1463–1464
- Computer-controlled electroacupuncture (CCEA), 2561–2562
- Computer games, brain aging and, 3994
- Computerized criminal histories, 4395

- Computers
  - anxiety and depression, 2691
  - Internet and, 4549
  - psychotherapy and, 2828–2829
  - use of (*See also* Geriatric psychiatry)
    - gender differences, 4248
- Comrey Personality Scales (CPS), 998*t*
- *COMT* gene, 1889
- Conation. *See also* Behavior
  - description of, 4570
- Concentration camp syndrome, 1812
- Concentration in patient interview, 967
- Concept-driven perceptions, 1114
- Concordance, description of, 4570
- Concrete operations period, 3360
- Concrete thinking
  - during adolescence, 3369
  - description of, 4570
  - as psychiatric symptomatology, 1121
  - psychosis and, 2872
- Concurrent validity, 996
- Concussion
  - description of, 4570
  - in referrals, 985
- Condensation, description of, 4570
- Conditional probability, 830
- Conditioned attentional orienting responses (CAORs), 415
- Conditioned reinforcers. *See* Reinforcement
- Conditioned responses. *See* Conditioning
- Conditioned stimulus (CS). *See* Conditioning
- Conditioning. *See also* Behavior therapy
  - aversive
    - paradoxical therapy and, 2817
  - description of, 4570
  - fear, 1753–1754
  - process of, 411–412
  - for PTSD, child/adolescent, 3704

- Conditioning, Pavlovian (classical), 718–722
  - and behavior, 718–719
  - erasing of, 721–722, 721*t*
  - learning process, 720–721, 720*f*
  - and memory, 729
- Conditioning therapy. *See* Behavior therapy
- Conditions not attributable to a mental disorder, 4570
- Condom use, risk behaviors and, 577
- Conduct disorders (CD), **3612–3621**. *See also* Oppositional defiant disorder
  - ADHD and, 2947, 3228, 3595–3596, 3602
  - aggressive behavior and, 2486–2489
  - bipolar comorbidity, 3689, 3703
  - brief psychotherapies for, 3743–3744
  - brief therapy treatments, 3747
  - in children after TBI, 652
  - classification, 2484
  - comorbid conditions, 3618–3619
  - comparative nosology, 3606, 3613
  - cortisol and, 2481
  - course, 3611, 3619–3620, 3619*f*
  - defined, 3605, 3607*t*, 3608*t*, 3612
  - description of, 4570
  - diagnosis of, 2436–2437, 3618
  - differential diagnoses, 3611, 3618–3619
  - epidemiology, 3613
  - etiology, 3613–3618, 3617*f*, 3618*f*, 3618*t*
  - family therapy, 3779–3781
  - FASD and, 638
  - fetal alcohol exposure and, 2484
  - further research, 3620–3621
  - group therapy and, 3765–3766
  - historical background, 3612–3613
  - laboratory examination, 3618
  - pathology, 3618
  - prognosis, 3611, 3619–3620
  - psychopharmacology, 3790
  - rating scales for, 3401, 3402

- subtypes and specifiers, 3612
- treatment, 3611–3612, 3620
  - sympathomimetics, 3228
- Conductance, ion channel, 132
- Conduction aphasia, 462*t*
- Confabulation. *See also* Paramnesia
  - delusions and, 1585
  - description of, 4570
- Conference on Human Rights and Psychiatric Oppression (CHRPO), 1549
- Confession
  - criminal competence, 4438
  - spirituality and, 2582–2583
- *Confessions* (St. Augustine), 4533
- Confidence, personality disorders and, 2140
- Confidence coefficient, 831
- Confidence intervals, 4570
- Confidentiality. *See also* Privilege
  - of adolescent evaluations, 3385–3386
  - in correctional psychiatry, 4453, 4453*t*
  - definitions of, 2021–2022
  - description of, 4570
  - end-of-life care and, 2307
  - ethical practice of psychiatrists, 4446, 4446*t*
  - ethics of, 3738, 3872–3873
  - factitious disorder and, 1856, 1863, 1864
  - family and couples therapy, 2758
  - group psychotherapy, 2741, 2743
  - of health care providers, 1558, 1560
  - mental health consultation for medical students and, 2607
  - of patient interviews, 945–956
  - privacy of patient, 4435
  - psychiatric emergencies, 2635–2636
  - psychoanalytic psychotherapy, 2670–2672
    - HIPAA, 2671
    - patient-litigant exception to testimonial privilege, 2670
    - Tarasoff decisions, 2670–2671
    - writing about patients, 2671–2672

- psychosis, psychodynamic psychotherapy and, 2877
- for school consultations, 3883

- Confinement in asylums, 4553
- Confirmation bias, research and, 2673
- Conflict. *See also* Extrapsychic conflicts; Intrapsychic conflicts
  - definition of, 4570
  - lode stars of, 760–761, 760*f*
- Conflict/deficit duality, 2646–2647
- Conflict-free areas, 4570
- Conflict models of ego psychology, 2646–2647
- Confrontation
  - description of, 4570
  - psychoanalysis and, 2655
- Confusion
  - after TBI, 527
  - description of, 4570
- Confusion Assessment Method (CAM), 1183
  - assessment of delirium, 2284
  - in delirium, 4090, 4224–4226, 4225*t*
- Confusional arousal, 2101–2102
- Confusional states in advanced disease
  - diagnosis, 2300–2301
  - prevalence of, 2300
- Congenital, definition of, 4570
- Congenital adrenal hyperplasia (CAH), 169–170
  - adrenogenital syndrome and, 2506
  - fetal, 2505
- Congenital heart disease (CHD)
  - late complications of, 2194
  - psychiatric issues in, 2198
- Congestive heart failure (CHF)
  - classification, 2191*f*
  - clinical features, 2191
  - course/prognosis, 2191
  - definitions, 2191
  - depression in
    - prognosis and, 2192

- diagnosis, 2191
- epidemiology of, 2191
- etiology of, 2191
- psychological factors in, 2192
- treatment of, 2191–2192

- Congruence in client-centered therapy, 2818
- Conjoint therapy, 4570. *See also* Quadrangular therapy
- Conjugal paranoia, 4570
- Conklin, Laura, 2850
- Conn syndrome, 683
- Connectionist models, 431–434
- Connective tissue disorders, 631–632
- Connectome, definition of, 687
- *Connelly v. Colorado,* 4438
- Conners, C. Keith, 3401
- Conners, Loney, and Milich (CLAM) Scale, 3401
- Conners-3 rating scale, 3401
- Conners Abbreviated Symptom Questionnaire, Hyperactivity Index (CASQ-HI), 3401
- Conners ADHD Index, 3401
- Conners Continuous Performance Task III, 650
- Conners Rating Scales-Revised, 3399
- Conolly, John, 4536
- *w*-Conotoxin, 139
- Conscience. *See also* Superego
  - description of, 4570
  - and mental health, 760
- Conscience and consciousness, 3378
- Conscientiousness, in Five-Factor Model, 935–936
- Conscious systems, topographic model, 861
- Consciousness. *See also* Cognition; Executive function; Trauma-related altered states of consciousness
  - after TBI, 527
  - binding and resonance, 1044–1047, 1045*f*, 1046*f*
  - description of, 4570
  - development of, 693–694
  - discourse and narrative, 693
  - disturbances of, 691, 1128–1132

    - altered states and, 1130–1131
    - dissociative phenomena and, 1131–1132
    - level of consciousness and, 1129
    - orientation and, 1129
    - overview, 1128–1129
    - psychological and physiological factors of, 1129
    - suggestibility and, 1131
  - dreaming and, **744–752**
    - cognition as brain function, 744–745
    - cross-references, 752
    - future directions, 751–752
    - gamma-band binding, 747–751, 747*f*, 749*f*
    - introspection, reality emulation, and cognition, 749
    - psychotherapy and brain activity, 751, 751*f*
    - thalamic cell oscillation, 745–747, 746*f*
    - thalamocortical dysrhythmic syndrome, 750
    - thalamocortical gamma-band resonant columns, 745
  - levels of, 2141
  - mental models and schemata, 691–692
  - metacognition and self-reflective capacity, 692–693
  - processing and laterality, 692
  - QEEG analysis, 1041–1044, 1043*f*
  - self-organizational processes, 694
  - social cognition, 693
  - spirituality and, 2577–2578
  - states of mind, 694–695
  - thought and language, 692
- Consensual validation, 4570
- Consent. *See also* Consent, informed
  - end-of-life decisions
    - for children, 2311–2312
- Consent, informed
  - aversion therapy, 2703
  - communication and, 2290
  - and correctional psychiatry, 4454–4455
  - depressive disorders, 3678, 3692
  - description of, 4588

- dissociative amnesia interviews, 1941
- dissociative identity disorder
  - pathology and laboratory examination, 1916
  - pharmacotherapy, 1927
  - treatment, 1922–1923
- ECT, 3288–3289
- for electroconvulsive therapy, 4155–4156
- ethical issues, 2278, 4242–4244
- ethical practice of psychiatrists, 4446–4447, 4447*t*
- ethics of, 3872
- for experimental treatment, 2312
- family and couples therapy and, 2757
- for forensic evaluations, 3863
- forensic psychiatry, 2733
- legal issues, 4234
- narrative psychiatry and, 2836–2838
- neurosurgery, 3315
- in pediatric cancer, 2282
- psychiatric emergencies, 2626–2627, 2636
- psychoanalytic research and, 2673
- psychopharmacology, 2921
- for selective serotonin reuptake inhibitor (SSRI) use, 3378, 3813
- with selective serotonin reuptake inhibitor (SSRI) use, 3378, 3813
- transplant evaluation and, 2362
- for treatment of sex offenders, 2021

- Conservation
  - as cognitive skill, 3360
  - description of, 4570
- Conservation of quantity, 708
- Conservator, description of, 4570
- Consonar. *See* Brofaromine
- Consortium for Research in ECT
  - CORE battery, 2739
- Consortium for Research in ECT (CORE), 4159
- Constancy principle (Freud), 849
- Constanitus Africanus, 4533
- Constipation

- description of, 4570
- in dying patients, 2296–2297
- Constitution, description of, 4570
- Constitutional symptoms, review of, 454
- Constitutional types, 4571
- Constricted affect, 4571
- Construct validity, 818
  - definition of, 320
  - for standardized assessment instruments, 996, 3395
- *Construction, Organization, and General Arrangements of Hospitals for the Insane* (Kirkbride), 4535
- Constructional apraxia, 4571
- Constructionism, views of, 1984
- Constructive use of ignorance, 3386–3387
- Consultation letters, 1837, 1844
- Consultation-liaison services, 4345
- Consultations, description of, 2291
- Consumer advocacy, 4291
- Consumer choices, 4324
- Contact interventions, suicide risk and, 2617
- Contagion, description of, 4571
- Containment, 1924
- Content thought disorder, 4571
- Content validity, 817
- Context
  - for clinical evaluations, 3376, 3377
  - for emergency assessments, 3442
  - restatement, free recall and, 2734
- Contextual learning, 1817
- Contextual transference, 2754–2755
- Contiguity theory, counterconditioning and, 2689
- Contingency
  - definition of, 2872
  - management, 2873–2874
    - DBT, 2790
    - sample contract, 2873–2874
    - of stimulant use disorders, 1263, 1288, 1370
    - SUD treatments, 384*t*

- tables, 842
- Contingent consequences, 2872
- Continuing medical education (CME), 4427–4428
- Continuity of care
  - with hospitalists, 4359
  - during rehabilitation, 4367
- Continuity of thought, 1120–1121
- Continuous dissociative amnesia, 1881*t*
- Continuous memory, 1882–1883
- Continuous Performance Test
  - of attention processing, 690
  - in schizophrenia, 696
- Continuous positive airway pressure (CPAP)
  - for obstructive sleep hypoapnea, 2094
  - sleep histogram, 2096*f*
- Continuous reinforcement, 2872, 4571
- Continuum model
  - CBT, psychosis and, 2883
  - depressive disorders, 1646–1647, 1647*f*
  - dissociative disorders, 1895*t*
- Continuum of care, 4305
- Contraception, methods of, 2512, 2513*t*
- Contraceptives, estrogen-containing, 3087
- Contracts
  - contingency management, 2873–2874
  - description of, 4571
- Contrast agents, 269–270
- Contrecoup injuries, 245–246
- Controllable pulse shape TMS (cTMS), 3299*t*, 3307
- Controlled mentalizing, 2895
- Controlled substances, definition of, 4571
- Controlled Substances Act, 2887
- Controlling, definition of, 4571
- Contusions, brain, 523
- *Conus geographus* (snail) toxin, 139
- Convergent systems
  - description of, 6–7

- schematic drawing, 6*f*
- sensory pathways, 24

- Conversational catharsis, 4571
- Conversion, definition of, 4571
- Conversion disorders
  - after disasters, terrorism, or war, 2567
  - classification, 1827
  - comparative nosology, 1839–1840
  - definition of, 1839, 4571
  - diagnosis of, 473–474
  - epidemiology, 1840
  - etiology, 1840
  - factitious disorder and, 1842, 1848–1849, 1859
  - history, 1839
  - hypnosis and, 2719
  - ICD-10 on, 1831, 1875*t*
  - intentionality and locus of symptoms in, 1850*f*
  - introduction, 1839
  - treatment, 1840
- Conversion paralysis, war and, 2566
- Conversion phenomena, 4571
- Conversion reaction, 1835, 1839. *See also* Somatic symptom disorder
- Convulsions. *See also* Clonic convulsions; Epilepsy; Seizure disorders; Tonic convulsions
  - definition of, 4571
- Convulsive disorders. *See* Epilepsy
- Convulsive therapy, 4571. *See also* Electroconvulsive therapy
- Coons, Phillip, 1871
- Cooper, Arnold M.
  - on psychoanalysis and other therapies, 2662
  - as psychoanalyst, 2648, 2650, 2651, 2653, 2661
- Cooperativeness
  - in children, 2167
  - defense mechanisms and, 2138
  - executive competence and, 2135
  - judgement measured by, 2143
  - personality disorders and, 2162*t*
  - TCI scales, 2164*t*

- temperament and, 2142
- Copays, in cost containment, 4293
- Coping
  - behaviors, stress vulnerability and, 2317
  - CBT, psychosis and, 2883
  - end of life issues, 2304
  - factitious disorder imposed on self and, 1850–1851
  - mechanisms for, 760, 4571
  - with pain syndromes, 2397
  - psychocutaneous disorders and, 2332–2326
  - strategies for, 2763
    - cards, 2770*f*
    - elderly, 3997–3998
  - stress and, 2324
- Coping Cat program, 3745
- Copper
  - accumulation of, 671
  - chelation therapy, 671–672
  - hepatic levels, 556
  - homeostasis of, 556
  - in Wilson disease, 556
- Coprolalia, 3636, 3650
- Coprophilia, 4571
- Coprophobia, 4571
- Coproporphyria, heritability, 2530
- Copy-cat suicides, 4571
- Copy number polymorphisms (CNPs), 147
- Copy number variants (CNVs)
  - analysis of, 3452, 3453*t*
  - description of, 147
  - genome project and, 147
  - mood disorder genetics and, 1627
  - overview, 3448–3449
  - schizophrenia, 1442–1443, 1443*f*, 1443*t*
- Cordance, antidepressants, 2545
- CORE. *See* Consortium for Research in ECT
- Core body temperatures, 2340

- Core conflict relational themes (CCRTs), 2720–2721
- Core schemas, 2688
- Corgard. *See* Nadolol
- Cornelia de Lange syndrome, 3669, 3683
- Cornell University Group, 2804–2805
- Corona radiata, formation of, 10
- Coronary artery bypass grafts (CABG)
  - delirium following, 1181, 1183
  - depression after surgery, 2187
  - psychiatric complications in, 2195–2196
- Coronary artery disease (CAD)
  - clinical features, 2187
  - course/prognosis of, 2187
  - definition of, 2186
  - diagnosis of, 2187
  - epidemiology of, 2186–2187
  - nosology, 2186
  - psychological factors affecting, 2187–2189
  - treatment, 2187
- Corpora amylacea in aging brain, 3991
- Corporal punishment, maltreatment (child), 3840
- Corprophyria, 660
- Corpus callosum
  - in Alzheimer patients, 13
  - anatomy of, 11–12
  - anterior commissure, 12
  - in autism patients, 13
  - body, 11–12
  - development, 3363*f*, 3462
  - genu, 11
  - isthmus, 12
  - neuroimaging
    - Tourette syndrome, 3472
  - parts, 11–12
  - posterior commissure, 12
  - PTSD and white matter reductions in, 1820
  - regions of, 14*f*

- rostrum, 11
- in schizophrenia, 13
- splenium, 11–12
- tic disorders and, 3639, 3653
- Tourette syndrome and, 3472

- Corpus luteum, 2507
- Corpus striatum, anatomy of, 26
- Correctional issues of elderly adults, 4239
- Correctional psychiatry, **4449–4477**
  - as community psychiatry, 4449–4451
  - description of, 4571
  - ethical principles, 4452–4454
    - collection of information, 4454
    - confidentiality, 4453, 4453*t*
    - doctor-patient relationship, 4452–4453
    - dual loyalty, 4452
    - informed consent, 4454–4455
    - medical autonomy, 4452
    - privacy, 4455
    - public areas, 4453–4454
  - history of in United States, 4449–4451, 4450*t*
  - practice of, 4451
  - psychiatric treatment in jails and prisons, 4455–4465
    - abuse and PTSD, 4566
    - access to mental health care, 4457
    - administration and management issues, 4463
    - barriers, 4457, 4457t
    - biomedical research, 4474
    - burnout, 4464–4465
    - community support, 4567
    - comorbidities, 4567–4568
    - criminalization of mentally ill, 4475
    - culture issues, 4465
    - custody and security issues, 4474
    - demographics, 4465
    - detoxification, 4567
    - disasters, 4460–4461

    - discharge, 4461–4462
    - dying inmates, 4469
    - electronic medical records, 4476
    - evidence-based treatment, 4458
    - future directions, 4476–4477
    - geriatric inmates, 4469
    - guidelines, 4456, 4456*t*
    - hepatitis C infection, 4568
    - high-risk periods, 4470, 4470*t*
    - HIV/AIDS, 4567–4568
    - hunger strikes, 4473–4474
    - information technology system, 4462–4463
    - inmate abuse, 4473, 4473*t*
    - LGBT and gender dysphoria, 4475–4476
    - malingering, 4471
    - overcrowding, 4475
    - pregnancy, children, and families, 4566
    - primary health care professionals, 4460
    - psychopharmacology, 4458–4459
    - rape and sexual assault victims, 4460
    - screening, 4459–4460, 4459*t*
    - seclusion and restraint, 4470–4471
    - segregation housing, 4471–4472, 4472*t*
    - sick call and nonemergency requests, 4460
    - substance abuse, 4566–4567, 4567*t*
    - suicides, 4469–4470
    - telepsychiatry, 4476
    - witness to force, 4473
    - women prisoners, 4465–4566
    - written policies and procedures, 4463–4464
- Correctional systems
  - clinicians working within, 4394
  - as de facto mental health care systems, 4305
  - mental health screening/treatment, 4305, 4394
  - recidivism reduction, 4394, 4396
- Corrective emotional experience
  - description of, 4571

- emotional processing and, 1770
- psychoanalysis and, 2640

- Correlation, definition of, 4571
- Corrigan, P.W., 4529
- Cortex
  - astrocytes in, 5
  - neuroimaging, 3471
  - tic disorders and, 3639, 3653
  - Tourette syndrome, 3471
- Cortical atrophy, 2424*f*
- Cortical dysplasia, consequences of, 451
- Cortical reward nodes, 366
- Cortices. *See also specific cortices*
- Cortico-striatal-thalamic-cortical (CSTC) circuitry
  - BDD and, 1791
  - OCD and, 1786, 1790
- Corticobasal degeneration (CBD), 549–550
  - dementia and, 1213, 1219
- Corticobasolateral amygdala (BLA), 2477
- Corticobulbar tracts, anatomy of, 10
- Corticolimbic disconnection hypothesis, 1891
- Corticopointine tracts, anatomy of, 10
- Corticortial connections, 20–21
- Corticospinal tracts, anatomy of, 10
- Corticosteroids. *See also specific drugs*
  - anxiety induced by, 2258
  - neurotoxic effects of, 2371
  - psychotic disorder induced by, 501–502
  - side effects, 2229
    - delirium, 2255
    - depression, 2260
    - insomnia and, 2087
- Corticostriatal projection, monkey, 29*f*
- Corticostriatalthamamic circuits, 543
- Corticotropin-releasing factor (CRF)
  - addiction and, 375
  - aggressive behavior and, 2481–2482

- anxiety and, 1740
- anxiety disorders and, 1743*t*, 1744–1745
- depression and animal modes of stress and, 1713
- fear and anxiety and, 1757
- function of, 90–93
- ligands, 92*f*
- neurosteroids and, 1748
- receptors, 92*f*
- release of, 36
- responses to stress, 2318
- secretion of, 2337
- trauma response and, 2569

- Corticotropin-releasing factor (CRF) antagonist, 3707
- Corticotropin-releasing hormone (CRH)
  - cytokine effects on, 197
  - function of, 2239
  - immunological effects of, 185*t*
  - production of, 2239
  - stress and, 193–194
- Corticotropin-releasing hormone-inhibiting factor (CRIF)
  - identification of, 172
  - immunological effects of, 185*t*
- Cortification, human, 2471–2472
- Cortisol
  - aggression and, 2481
  - anxiety disorders and, 1744
  - circadian profile of, 206
  - dissociation and, 1892
  - dissociative identity disorder and, 1882
  - fear and anxiety and, 1758
  - levels, stress and, 399*f*
  - panic disorder, 1744
  - production of, 2243
  - PTSD and, 1819
  - sexual function and, 1956
  - social anxiety disorder, 1743*t*
- Cosmos, psychodrama and, 2858

- Cost-effectiveness, 846–847
- Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS), 1523
  - FGA, 3069, 3070
  - second-generation antipsychotic medication trials, 4351–4352
- Cotard syndrome, 538. *See also* Nihilistic delusions
- Coumadin. *See* Warfarin
- Council for Exceptional Children (CEC), 3493
- Counseling, description of, 4571
- Counselor-assisted problem solving (CAPS), 653
- Counter-relationships, mentalizing, 2900–2901
- Counterconditioning model, 2689
- Counterphobia, 4571
- Countertransference
  - in analytic process, 891
  - description of, 4571
  - dissociative identity disorder and, 1925
  - ethical issues of, 3872
  - in group therapy, 3737
  - in individual psychotherapy, 4177–4178
  - physician ethics, 4444
  - physician mental health and, 2608–2609
  - psychoanalysis and, 2653–2654
  - psychodynamic psychotherapy and, 2677
    - psychosis, 2881
  - in psychodynamic therapy, 3737
  - suicide risk and, 2617
- Countertransference enactment, 2654
- Couples of LBGT persons, 2004
- Couples therapies, **2748–2760**. *See also* Family therapies; Marriage therapy
  - assessment, 2756
  - attendance rules, 2757
  - clinical applications, 2755–2758
  - clinical management issues, 2757–2758
  - contraindications, 2755, 2756
  - definition, 2750
  - diversity and, 2758
  - evaluation, 2756

- gambling disorder, 1808
- group, 2758–2759
- historical trends, 2749–2750
- indications, 2755–2756
- initial interview, 2756–2757
- introduction, 2748–2749
- pseudocrisis, 2757
- research in, 2759–2760
- secrets in, 2757
- specific areas of inquiry and observation in, 2757
- timing interventions in, 2758

- Court testimony of elderly adults, 4239
- Courtois, Chris, 2719
- Covariance and correlation, 831
- Covert dissociative syndrome, 1934
- Covert sensitization
  - behavior therapy, 2694
  - description of, 4571
  - for sexual addiction, 2056
- Cowden syndrome, 314
- Cowper glands, 1960*t*
- CPA. *See* Canadian Psychiatric Association (CPA)
- CPES. *See* Collaborative Psychiatric Epidemiology Surveys
- "Crack," 4571. *See also* Cocaine
- Craig, T. K., 1842
- Cramp fasciculation syndrome, 2270
- Cranial electrical stimulation (CES), 3299, 3299*t*, 3308–3309
- Cranial nerves, 358, 456, 630
- Cranioelectrotherapy stimulation (CES). *See also* Microcurrent electrical stimulation
  - scientifically validated therapy with, 2560–2561
- Craniopharyngiomas
  - in childhood, 677
  - MRI imaging of, 678*f*
  - prolactin levels and, 2249
- Craniosacral therapy, 2546, 2563–2564
- Craske, Michelle, 1771
- *Crataegus* spp. (hawthorn berry), 2554

- Cravings
  - addiction and, 377
  - substance use disorders, 1346
- Crawford, Helen, 2712
- Creatine, biosynthesis disorders, 663–664
- Creativity
  - bipolar II disorder and, 1652
  - definition of, 4571
- CREB, anger and, 2483
- Credit card fraud, elder abuse, 4251
- Crestor. *See* Rosuvastatin
- Creutzfeldt–Jakob disease (CJD)
  - clinical assessment, 614–615
  - dementia and, 1220–1221
  - description of, 4571
  - diagnosis of, 473, 610*t*, 615*t*
  - differential diagnoses, 616, 616*t*
  - electroencephalogram in, 609*f*
  - familial (fCJD), 601
  - genetic factors, 4022
  - iatrogenic (iCJD)
    - cause of, 601*t*
    - neurological manifestation, 610–611
  - imaging in, 466
  - myoclonus in, 458
  - neurodiagnostic workup, 615–616
  - preventive measures, 617–618
  - sporadic (sCJD), 600, 609–610
    - brain MRI, 612*f*
    - cause of, 601*t*
    - diagnostic workup, 615–616
    - features of, 613*t*
    - neurological manifestation, 609
    - presentation, 613*t*
    - psychiatric manifestation of, 612
  - treatment, 616–617
  - variant (vCJD), 600

    - brain MRI, 615*f*
    - cause of, 601*t*
    - course/prognosis, 616*t*
    - diagnostic criteria, 610*t*
    - features of, 613*t*
    - neurological manifestation, 610
    - neuropathology of, 607
    - presentation of, 613*t*
    - psychiatric manifestation of, 612
    - psychological functioning, 613–614
  - world-wide rates of, 596–597
- CRF. *See* Corticotropin-releasing factor (CRF)
- CRH. *See* Corticotropin-releasing hormone (CRH)
- Cri-du-chat syndrome, description of, 4572
- Crime. *See also* Aggression; Violence
  - criminal responsibility assessments, juvenile, 3866
  - forensic psychiatry and, 2734–2735, **3861–3868**
  - schizophrenia and, 1422–1423, 1435
  - statistics on, 2473–2476, 2473*t*, 2474*f*, 2475*f*
  - violent, definition of, 2538
- Criminal cases, media on, 1874
- Criminal justice, and neurocentrism, 784–788
- Criminal justice system. *See also* Juvenile justice
  - diverting persons with mental illness from, 4393
  - expansion after deinstitutionalization, 4389
  - forensic patients from, 4344
  - as hospital alternative, 4325
  - mental health care in, 4392
  - mental/substance-use disorders in, 4315
  - psychiatric hospitals interface with, 4326*f*
  - seriously mentally ill in, 4305
- Criminal responsibility, 4572
- Criminality
  - adult, **2413–2430**
  - gene-environment interactions and, 2483
  - mental illness and, 2489–2490
  - organized, 2413–2414

- psychopaths and, 2491
- white-collar, 2414

- Criminalization
  - overview of, 4387
  - causes of, 4389–4391
  - definition of, 4388–4389
  - hospitalization *vs.,* 4391
  - magnitude of, 4388
  - mental health services after arrest, 4394–4395
  - mental illness
    - in jails and prisons, 4475
  - mental illness/violence, 4391–4392
  - treatment
    - after arrest, 4392–4394
    - to reverse/prevent, 4395–4401
- Criminals
  - competency, 4438–4439
  - responsibility and psychiatric illness, 4439
- Crisis, description of, 4572
- Crisis beds. *See* Beds
- Crisis intervention team model (CIT), 1553–1554
- Crisis interventions, 4572
- Criterion validity, 817–818, 817*t*, 996
- Critical care physicians in ICUs, 4355
- Critical illnesses
  - adult survivors of, 2330
  - families of survivors, 2330
  - pediatric survivors of, 2330
  - psychiatric morbidity after, **2329–2332**
  - stress of, 2330
  - survivors of
    - cognitive dysfunction in, 2330–2331
    - physical dysfunction in, 2331
- Critical incident stress debriefing (CISD), 2571, 2572
- Critical judgment, hypnosis and, 2710
- Critical periods
  - of brain development, 3358

- description of, 4572
- Critical ratios, description of, 4572
- *Crocus sativus* (saffron), 2554
- Cronback's coefficient alpha, 996
- Cross-cultural differences, family burden and, 4520
- Cross-cultural psychiatry, 4572
- Cross-dressing
  - in natal males, 2030–3031
  - transvestic fetishism and, 2019
- Crossing over
  - description of, 150*f*
  - process of, 149
- *Crossing the Quality Chasm: A New Health System for the 21st century*
  - as call to action, 4289
  - mental health/substance use treatment in, 4314–4315
  - quality and accountability reinforced by, 4294
- Crosslinking, process of, 160
- Crow, T. J., 1415–1416
- *Cry* gene, 206
- *CRY* gene, depression and, 213
- Crying
  - pathological, 464
  - as universal distress signal, 1711
- Cryptococcal meningitis, 585
- *Cryptococcus neoformans*
  - opportunistic infections, 585–586
  - treatment of, 586
- Cryptographia, 4572
- CSF. *See* Cerebrospinal fluid
- Csikszentmihalyi, Mihalyi, 2843
- CT scans. *See* Computed tomography (CT)
- Ctenoids, 238–239, 240*f*
- CTMS. *See* Controllable pulse shape TMS
- Cullen, William, 4534
- Cults
  - description of, 4572
  - membership, coercive persuasion and, 1947

  - religiosity and, 2584
- Cultural anthropology. *See* Anthropology
- Cultural assessments
  - historical overview, 2460–2463
  - in psychiatry, **2460–2470**
- Cultural countertransference, 2464
- Cultural deprivation, description of, 4572
- Cultural factors
  - behavior therapy and, 2701
  - burn injuries and, 2374
  - child maltreatment, 3839
  - and cognitive-behavior therapy, 4180–4181
  - disaster psychiatry and, 2574–2575
  - dissociative disorders, 1899
  - and everyday functioning, 4027
  - family role, 4199
  - in jails and prisons, 4465
  - as psychosocial factor, 4169–4171
  - self-immolation and, 2383
  - somatic symptom and related disorders, 1832
- Cultural formulation, 2462–2463, 2462*t*
- Cultural Formulation Interviews (CFI), 2459
  - core of, 2465–2468
  - domains of, 2467
  - in DSM-5, 2464–2469, 2466*t*–2467*t*
  - future directions, 2469
  - Informant (I-CFI), 2468–2469
  - supplementary modules, 2469, 2469*t*
- Cultural identity of individuals, 2463
- Cultural issues, end-of-life care, 2306*t*
- Cultural psychiatry, 2443
- Cultural responsiveness, training in, 4383
- Cultural syndrome, 791
- Cultural transference, 2464
- Culturally appropriate services
  - competence in, 4381–4382, 4399
  - standards for, 4380

- Culture
  - definition of, 2465
  - ethics of, 3871–3872
- Culture, psychiatry and, **788–808**
  - cultural assessment of patient, 788
  - cultural formulation, 789–791
    - concept of distress, 789–790
    - glossary of distress concepts, 791
    - identity of individual, 789
    - individual-clinician relationship, 790
    - interview, 790–791
    - overall cultural assessment, 790
    - psychosocial stressors, 790
  - future directions, 806–808
    - biocultural links, 808
    - cross-references, 808
    - culture as diagnostic tool, 808
    - culture as explanatory tool, 808
    - culture as management element, 808
    - culture as pathogenic agent, 808
    - culture as therapeutic or protective factor, 808
  - global perspective
    - migration, 796–797
    - refugees, 796
  - immigrants and refugees
    - mental status examination, 793
    - migration history, 793, 793t
  - immigration and mental health, 803–806, 804*t*, 805*t*
    - refugees in U.S., 806
    - schizophrenia, 806
  - mental health
    - cross-national differences, 797–798, 798*t*
    - global burden, 797
    - schizophrenia in developing countries, 799–800
    - treatment gap, 798–800
  - migration, acculturation, and stress, 791–793
  - psychiatric diagnosis, 788–789

- of schizophrenia, 3932
- U.S. population
  - age distribution, 795
  - discrimination and service utilization, 802–803
  - foreign-born component, 794, 794*t*
  - indigenous mental health, 800–801
  - mental disorder differences, 800
  - poverty rates, 795
  - projections for 2014 to 2060, 795
  - psychiatric service utilization, 801–802, 801*t*
  - race and ethnicity in 1900 to 2013, 795
- variations, and patient interview, 963

- Culture-bound syndromes, **2443–2470**
  - definitions, 2443–2445
  - historical background of, 2445–2447
    - cultural psychiatry work and, 2445–2446
    - pioneering work in, 2445
- Culture shock, 4572
- Culture specific syndromes, 4572
- Cumulative distribution function, 830
- Cunnilingus, 4572
- Curies, units of, 286
- Current Opioid Misuse Measure (COMM), 2393
- Currents, ion channel, 132
- Curriculum-based measurements (CBMs), 3426
- Cushing syndrome, 2243–2244
  - cause of, 2243
  - characteristics of, 167
  - clinical features of, 2243–2244
  - cognitive changes in, 2244
  - description of, 678, 4572
  - elevated HPA activity and, 1716
  - HPA axis and, 167
  - iatrogenic, 2229
  - presentation of, 682
- Custody
  - definition of, 4572

- forensic psychiatry on, 3863–3864
- issues of children, 4440
- in jails and prisons, 4474

- Cutaneous body image (CBI), 2335
- Cutaneous dysesthesias, 2349*f*
- Cuthbert, Bruce, 1844
- CUtLASS. *See* Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study
- CVA. *See* Cerebrovascular accidents (CVA)
- CVLT. *See* California Verbal Learning Test
- Cyamemazine, 3060
- Cyanide, war/terrorism and, 2572*t*
- Cybernetics
  - definition of, 4572
  - family therapy and, 2749
- Cybersex addiction, 4572
- Cybertherapy, 2852–2853. *See also* Internet
- Cyclazocine, 4572
- Cyclic adenosine monophosphate (cAMP)
  - cAMP-activated protein kinase A, 2915
  - cAMP phosphodiesters, 2914–2915, 2915*f*
  - intracellular signaling pathways, 2914–2915, 2915*f*
  - pathway, 121–122, 121*f*
  - receptor analysis, 2906
- Cyclic adenosine monophosphate response element-binding protein (CREB). *See* CREB
- Cyclic guanosine monophosphate (cGMP)
  - NO activation of, 105–106
  - signaling pathway, 123*f*
- Cyclin-dependent kinase inhibitors (CKIs), 50
- Cyclopbenzaprine, 635
- Cyclopia, 49
- Cycloplegia, 4572
- Cyclosporine (Neoral)
  - drug interactions, 3131
  - neurotoxic effects, 2371
  - side effects of, 2198
- Cyclothymia, 4572

- Cyclothymic depression, 1651
- Cyclothymic disorder
  - as bipolar disorder, 1650
  - bipolar II disorder and, 1650–1651
  - definition of, 1601, 1631, 4572
  - genetics and, 2529
- Cyclothymic temperament
  - alcohol and substance abuse, 1651–1652
  - bipolar II disorder and, 1651
- Cylert. *See* Pemoline
- Cymbalta. *See* Duloxetine
- Cynophobia, 4572
- CYP450 system, antidepressants and, 496
- *CYP27A1* mutations, 662
- Cyproheptadine (Periactin)
  - chemistry, 2981*f*
  - in migraine prevention, 625
  - schizophrenia, 3328
  - sexual dysfunction and, 1972
  - therapeutic indications, 2982
- Cyproterone acetate (Androcur, CPA, Cyprostat), 1980, 2020
  - sex offenders, 2503
    - recidivism and, 2503
  - sexual addiction, 2056
  - side effects, 2260
- Cyptolalia, 4572
- Cys-loop family, 142, 142*f*
- Cysteine-containing aspartate-specific proteases. *See* Caspases
- Cystic fibrosis (CF)
  - adults with, 2227
  - diagnosis of, 2227
  - end-of-life issues, 2227
  - fertility issues on, 2227
  - manifestations of, 2226–2227
  - physiology of, 2226–2227
  - psychiatric comorbidities, 2227
  - psychiatric interventions, 2227–2228

- psychosocial considerations, 2227
- self-management, 2227
- transplantation issues, 2227
- treatment, 2227

- Cystic fibrosis transmembrane conductance regulator (CFTR) channels, 140, 2226–2227
- Cytoarchitecture
  - brain, 20*f*
  - medial temporal lobe, 6*f*
  - staining of, 5
- Cytochrome C release, 46
- Cytochrome P450 enzyme system, 1782
- Cytokines
  - acquired immunity and, 183*t*
  - atopic dermatitis and, 2340–2341
  - categories of, 181
  - in delirium, role of, 1181–1182
  - effects on the brain, 194–195
  - immune responses and, 2320
  - neuroendocrine effects, 197
  - regional brain activity and, 196–197
  - stress and, 2338–2339
- Cytomegalovirus (CMV)
  - encephalitis related to, 595
  - HIV infections and, 585
  - infections in HIV patients, 569–570
- Cytomel. *See* Liothyronine
- Cytotoxic T lymphocytes (CTLs), 181

D

- d-Amino acid oxidate activator *(DAOA),* 1610, 1625
- D-Cycloserine (DCS), 722
  - agoraphobia, 1743*t*
  - extinction learning and, 1747
  - and fear conditioning, 722
  - OCD, 1789
- D-KEFS. *See* Delis-Kaplan Executive Function System
- DA. *See* Dopamine (DA)
- Da Costa syndrome, 2568, 4572
- Daclizumab, 2371
- Daitch, Carolyn, 2717, 2721
- Dalenberg, Constance, 1898–1899
- Dalí, Salvador, 2641, 2643, 2643*f*
- Dalmane. *See* Flurazepam
- DALYs. *See* Disability Adjusted Life Years
- Damadian, Raymond, 259
- Damage-associated molecular pattern molecules (DAMPs), 180
- Damages by physician, 4436–4437
- Damasio, Antonio, 331, 2840
- Danazol (Danocrine), 2510
- Dance therapy, 2556
- *Daniel M'Naghton: His Trial and the Aftermath* (West), 4535
- Dante, near-death experiences and, 2590
- *Darkness Visible: A Memoir of Madness* (Styron), 4546
- Darwin, Erasmus, 1179
- DAT. *See* Dopamine transporter
- "Date rape drug," 2998
- Daubert standard, 2410
- *Daubert v. Merrell Dow Pharmaceuticals,* 2410, 4437
- Davanloo, Habib
  - intensive short-term dynamic psychotherapy, 2796, 2797, 2801–2804
  - short-term dynamic psychotherapies, 2647
- David, Anthony, 1871
- David, Karoly G., 2481
- Davis, John, 1713
- Davunetide, 549

- Day hospitals, 4572
- Day residue, 4572
- Daytime sleepiness
  - in aging adult, 4056
  - in long-term care, 4230–4231
  - measurement of, 343
- DBI. *See* Diazapam-binding inhibitor (DBI)
- DBS. *See* Deep brain stimulation (DBS); Discrete behavioral states (DBS) model
- DBT. *See* Standard Dialectical Behavior Therapy
- De Clérambault syndrome. *See* Erotomania
- *De remedis utriusque fortunae* (Petrarch), 4533
- De Sauvages, Boissier, 4534
- DEA. *See* Drug Enforcement Agency
- Deadly nightshade *(Atropa belladonna),* 73
- Deaf persons
  - mental health care for, 4343–4344
  - psychotherapy for, 2827–2828
- Death/dying. *See also* Dying patients
  - acceptance of, 2253
  - adjustment to loss, 2308*t*
  - attended, 2303–2304
  - awareness of, 2288
  - child's understanding of, 2308, 2308*t*
  - communication with parents, 2308–2309
  - competency of older adult, 4235–4236
  - concepts of, 2308*t*
  - cultural issues, 2306
  - desire for, in terminal illness, 2313*t*
  - dignity therapy and, 2825–2827
  - the final farewell, 2303
  - gun control and, 2476
  - hallucinogens and, 2887
  - at home, 2304
  - inevitable, staff conferences in, 2386
  - of inmates, 4469
  - living, dying and, 2288
  - pediatric care, 2307

- religious issues, 2306
- spiritual issues, 2306
- terminal illness and, 2892–2893
- terminally ill children, 2309–2310
- timely, 2288
- trajectories, 2298
- unattended, 2303
- unexpected, 2303
- untimely, 2288

- Death fear, 918
- Death feigning, 1890
- Death instinct. *See* Aggressive drive
- Death penalty, 4481–4482
- Death row inmates, 4472–4474
- Debriefing, psychological, 1824
- Decade of the Brain, 4306
- Decarceration, incarceration *vs.,* 4400
- Decathexis. *See* Acathexis
- *Deception of Demons* (Weyer), 4533
- Deceptive patients, interviews, 962
- Decision-making
  - cross-cultural barriers, 2306
  - ethical issues, 4242–4244
  - family role, 4199
  - mathematical modeling of, 435
  - medical, 4434
  - physician ethics, 4445
- Decision trees, 841, 4572
- Declaration of Caracas, 4530
- Declaration of Helsinki, 4240
- Declarative memory, 324
- Decoding, 4572
- Decompensation, 4572
- Decongestants
  - adrenergic agonists as, 67
  - side effects, 2087
- Decremental conduction, 133–135

- Deep brain stimulation (DBS), 4163
  - current state of, 3320–3321
  - definition, 3299*t*
  - in depression, 1681
    - efficacy of, 33
  - description of, 4572
  - dissociation and dissociative disorders, 1889
  - future directions, 3321–3322
  - history, 3299–3300
  - lesion procedures, 3317
  - major depression, 3320
  - neurosurgical implantation, 3313, 3314, 3315–3316, 3317
  - in OCD, 1790, 3319–3320
  - outcomes, 3319–3320
  - in Parkinson disease, 547
  - postoperative care, 3315–3316
  - in tardive dystonia, 2943
- Deep breathing, sustained, 2557
- Deep tissue hallucinations, 1411
- Default mode networks, 418, 1819
- Defense mechanisms
  - in cancer patients, 2265–2266
  - description of, 4572
  - dissociative TRASC and, 1890
  - as ego functions, 1660–1661
  - elation and, 1632
  - levels of, 2138
  - maturation of, 2169
  - maturity of, 2138
  - and mental health, 760–761, 767*t*
  - mood and, 1661
  - primitive, 2164
- Defense Style Questionnaire, 2165
- Defense test battery, 320
- Defenses, evolution of, 771
- Defensive behaviors, models of, 320–321
- Defensive emotion, 4572

- Defensive spirituality, 2825
- Defiant behavior, ODD symptom, 3610
- Deficit model, of ego psychology, 2646–2647
- Deficit syndrome, schizophrenia, 1418–1419, 1418*t*
- Deficits, definition of, 4572
- Degrees of freedom, 834
- Dehaene, Stanislas, 3521
- Dehydration
  - in bulimia nervosa, 2075
  - voice disorders and, 3567
- Dehydroepiandrosterone (DHEA), **170–171,** 3325, 3336
  - administration of, 170
  - in adrenal insufficiency, 170
  - anxiety disorders and, 1744
  - biological treatment with, 2555
  - eating disorders and, 118
  - function of, 117
  - postpartum mood disorders and, 118
  - PTSD and, 1819
- Dehydroepiandrosterone sulfate (DHEA-S), 1892
- Deinstitutionalization
  - as criminalization cause, 4389
  - definition of, 4572
  - of intellectual disability patients, 3492
  - as shift from hospital-based treatment, 4329
  - violence and, 4391
  - without adequate community resources, 4292
- Deisseroth, Karl, 145
- *Déjà vu,* 2491, 4572. *See also* Paramnesia
- Delayed ejaculation, 1968–1969, 4572
- Delayed recall
  - dissociative amnesia and, 1882–1883
  - trauma, 1886
- Delayed sleep phase circadian rhythm disorder, 4572
- Delayed sleep phase syndrome (DSPS), 214
- Delayed-type hypersensitivity (DTH), 188
- Delfito, Paul, 2725

- Delirious mania, 1640
- Delirium, **1178–1191, 4086–4094**
  - after TBI, 527–528
  - alcohol withdrawal-induced, 1178
  - anticholinergic intoxication and, 1189
  - assessment of, 2256
  - benzodiazepine-related, 2998
  - bipolar disorder and, 1185
  - burdens of, 1178, 1178t, 1178*t*
  - in burn patients, 2375–2376
  - in cancer patients
    - causes of, 2255*f*
    - clinical features, 2255–2256
    - epidemiology of, 2255
    - etiology of, 2255
    - presentation, 2255
  - in cardiac patients, 2196
  - chemotherapeutic agents and, 2273*t*
  - in children, 1185, 2281
  - clinical features of, 1185–1186, 2255–2256
  - CNS tumors and, 2267
  - comparative nosology, 1179–1180, 4087
  - course, 1186–1187
  - course and prognosis, 4091–4092
  - criteria, 1178
  - definition of, 1178–1180, 4087
  - dementia and, 1184–1185, 1198
  - description of, 461, 4572
  - diagnosis of, 1183, 2255–2256, 4089–4090
  - differential diagnoses, 1184–1185, 1217, 1236, 1238, 4091
  - drug abuse and, 1182
  - in dying patients, 2301*t*
  - EEGs in, 251
  - in elderly patients with cancer, 2284
  - epidemiology, 1180–1181, 4087–4088
  - etiology of, 1182–1183, 1189–1190, 4088–4089, 4088*t*
  - future directions, 4094

- due to general medical condition, 1178
- history, 1178–1179, 4087
- in HIV patients, 569
- hyperactive hallucinatory, 2300
- hypoactive, 2300
- in hypoparathyroidism, 682
- ICD-10 criteria for, 1179
- incidence, 1180, 1180t
- intoxication and
  - cannabis-related, 1306–1307
  - defined, 1178
  - differential diagnoses, 1381
  - hallucinogen-related, 1327
  - inhalant-related, 1338
  - substance-related, 1182, 1189–1190
- laboratory examination, 1183–1184
- in long-term care, 4224–4226, 4225*t*
- management of, 2256–2257, 2301
- in medical settings, 1177
- memory in, 1935
- pathology and laboratory examination, 4090–4091, 4090*t*
- pathology of, 1184
- pathophysiology, 1181–1182
  - inflammation and, 1181–1182
  - neurotransmission and, 1181
- postoperative, 1183, 2368*f*
- precipitating factors, 1180, 1181t, 1181*t*
- prevalence, 1180, 1180t
- prevention, 1187, 4092
- prognosis, 1186–1187
- protective factors, 1180–1181
- rating scales, 1183, 1188*t*, 1189*t*
- risk factors, 1180, 1180*t*, 1181*t*, 1184, 4088, 4088*t*
- schizophrenia and, 1185
- serum electrolyte evaluation, 1010
- in special populations, 1184–1185
- subtypes of, 2256, 2300*t*

- terminal, 2256
- in terminally ill children, 2309
- in terminally ill patients, 1185, 1189–1190
- terminology for, 1179–1180, 1179t
- in transplant populations, 2367
- treatment, 1187–1190, 4092–4094, 4092*t*, 4093*t*, 4134
  - for anticholinergic intoxication, 1189
  - FGAs and, 3071
  - nonpharmacological, 1187
  - of Parkinson disease, 1189–1190
  - pharmacological, 1187–1189, 1189*t*
  - of substance intoxication, 1189–1190
  - of substance withdrawal, 1190
  - of terminally ill patients, 1189–1190
  - of Wernicke encephalopathy, 1189
- treatment principles, 2300*f*
- types of, 1186
- war or terrorism and, 2572
- withdrawal
  - alcohol-related, 1272, 1277
  - sedative-, hypnotic-, anxiolytic-induced, 1381
  - substance-related, 1182, 1190

- Delirium Rating Scale (DRS), 1189*t*
- Delirium tremens (DTs)
  - during alcohol withdrawal, 1182
  - alcoholism and, 83
  - description of, 4573
  - treatment
    - barbiturate, 2987
    - benzodiazepine, 2997
- Delis–Kaplan Executive Function System (D-KEFS)
  - in aging adults, 4007
  - learning disability testing for, 3428
  - for neuropsychological assessments, 3432
  - neuropsychological testing, 992
- Delivery systems (healthcare)
  - reforming, 4307

- shortcomings of, 4288–4289
- *Della Melancolia* (Constanitus Africanus), 4533
- Delta sleep. *See* Slow-wave sleep
- Delta waves, 4573
- Delusional BDD, 1791
  - psychotherapy, 1792
  - SSRIs and, 1792
- Delusional depression, 4573
- Delusional (paranoid) disorder, 4573
- Delusional disorders, **1582–1587**
  - and cognitive disorders, 4100–4101
  - comparative nosology, 1576*t*, 1583–1584, 1584*t*
  - course and prognosis, 1586–1587, 4101–4102
  - in delirium, 4089
  - diagnosis and clinical features, 1584–1585, 4101–4102
  - differential diagnoses, 1585–1586, 4102
  - epidemiology, 1584, 4101
  - etiology, 1584, 4101
  - history, 1582–1583, 1583*f*
  - somatic type, 2346–2347
  - substance-induced, 4101
  - treatment, 4101, 4102
    - FGA, 3070
- Delusional parasitosis, 2346–2347
  - diagnosis of, 2346–2347
  - pathogenesis of, 2347
  - subtypes of, 2347
  - treatment, 2347
  - treatment of, 2355*t*
- Delusional percept, 1123
- Delusional separateness, 1907
- Delusional symptoms in partner of individuals with delusional disorder, **1597**
- Delusional thought passivity, 1120, 1121, 1122–1123
- Delusions
  - after TBI, 538
  - bizarre, 1414–1415
  - cognition and, 1415

- content of, 465
- of control, 4573
- delirium and, 1186
- description of, 4573
- elasticity of, 413
- of grandeur, 4573
- of infidelity, 4573
- of jealousy, 4573
- learning theory and, 413
- manic syndrome and, 1641
- misidentification, 465–466
- of negativity, 4573
- OCD and, 1785–1786
- at patient interview, 953, 953*t*
- in PD, 545
- of persecution, 1414, 1415, 4573
- of poverty, 4573
- of reference, 1414, 4573
- schizophrenia and, 1413–1415
- substance/medication-induced psychotic disorders symptom, 1591
- thought broadcasting, 1414
- thought insertion, 1414

- Demand side interventions
  - for care improvement, 4348, 4349–4350
  - definition of, 4349–4350
- Dementia praecox, 1407, 1426, 1431–1432, 1453, 1543, 1549, 1551. *See also* Schizophrenia
  - description of, 4573
- *Dementia Praecox or the Group of Schizophrenias* (Bleuler), 4537
- Dementia with Lewy bodies (DLBs)
  - characteristics, 550–551
  - movement disorders and, 540
- Dementia(s), **1191–1121,** 1198. *See also* Antidementia drugs; Frontotemporal dementia; Pseudodementia; Vascular dementia
  - aging, normal and, 1177
  - agitation in, valproate for, 3276
  - alcohol-related, 1273
  - Alzheimer disease and, 1202–1210

- Alzheimer type, 4573
- associated with alcoholism, 4573
- B-family vitamins and, 2548
- behavioral disturbances and, 1196–1197
- benzodiazepine aggravation of, 2998
- caregivers (*See* Caregivers)
- causes of, 460
- chemotherapeutic agents and, 2273*t*
- cholinesterase inhibitors for, 452
- classification of, 1168
- clinical features, 1193–1194, 1193*t*
- comparative nosology, 1191–1192
- course, 1199
- CSF studies, 1197
- defined, 1191
- delirium and, 1184–1185
- delusional disorder and, 1586
- description of, 4573
- diagnosis, 1194–1197
- differential diagnoses, 468–470, 469*t*, 1198–1199
- due to another medical condition
  - clinical features, 1237
  - comparative nosology, 1237
  - course, 1238
  - defined, 1236–1237
  - diagnosis, 1237, 1237*t*
  - differential diagnoses, 1238
  - epidemiology, 1237
  - etiology, 1234*t*, 1237
  - laboratory studies, 1237–1238
  - prognosis, 1238
  - treatment, 1238
- EEGs in, 251, 1197
- in elderly patients with cancer, 2284–2285
- electroconvulsive therapy, 4153
- epidemiology, 1192–1193
- etiology, 1193, 1193t, 1193*t*

- evaluation, 1194–1196, 1195*t*
- everyday functioning, 4037–4038
- fMRI of, 276
- frontotemporal, heritability, 2530
- history, 1191, 1407
- HIV-related, 569–570, 3854–3855, 4269
  - in adolescents, 646
  - presentation, 646
- HPA axis and, 167
- in Huntington disease, 1220
- laboratory studies, 1197, 1197*t*
- with Lewy bodies, **1214–1219**
  - aggression and, 2495
  - clinical features, 1215–1216
  - comparative nosology, 1214–1215
  - course and prognosis, 1217
  - defined, 1214
  - diagnosis, 1216
  - differential diagnoses, 1198–1199, 1217, 1219
  - epidemiology, 1215
  - etiology, 1215
  - history, 1214
  - laboratory examination, 1217, 1218*f*
  - pathology, 1216–1217, 1216*f*
  - treatment, 1217–1218
- in long-term care, 4226–4229, 4227*t*, 4228*t*
- MCI leading to, 1224, 1226–1228, 1228*f*
- memory in, 1935
- menopause and, 2522
- molecular neuroimaging, 1197
- MRI applications to, 271–272
- MRS in, 280–281
- neurobehavioral syndromes, 460
- neurodegenerative, 529–530
- neuroimaging of, 1197, 4481
- neuropathology of, 1198
- neurophysiological testing, 1197–1198

- neuropsychiatric manifestations of, 1196–1197, 1201–1202
- neuropsychological testing, 1197–1198
- neurosyphilis and, 592
- in older people, 3956, 3963–3964, 3979–3982, 3979*t*
- omega-3 fatty acid supplements and, 116–117
- paraneoplastic syndromes and, 2270
- parasomnia and, 2102
- in Parkinson disease, 1219–1220
- with parkinsonian signs, 550–557
- pathology, 1198
- in prion diseases, 1220–1221
- prognosis, 1199
- and sexual dysfunction, 4263
- subtypes, **1202–1221**
- suicide risk, 2615
- and testamentary capacity, 4438
- treatment, 1199–1202
  - cholinesterase inhibitors, 3039–3040, 3040*t*, 3044
  - FGAs and, 3071
  - nonpharmacological, 1199–1200, 1201
  - pharmacological, 1200–1202
- vascular neurocognitive disorder and, 1241

- Demographic factors
  - barriers to health care, 4304
  - mood disorders, suicidal behavior and, 1671
  - psychotherapy and, 2864–2865
  - suicide, 2610–2612, 2611*t*, 2612*t*
- Demographics
  - aging people with mental illness (*See* Geriatric psychiatry)
  - alcohol and substance abuse, 4108–4109, 4109*t*
  - health care costs for older adults, 4212–4213
  - psychiatric patients in jails and prisons, 4465
  - U.S. populations
    - age distribution, 795
    - immigration factors, 794, 794*t*
    - poverty factors, 795
    - projections 2014 to 2060, 795

    - race and ethnicity in 1900 to 2013, 795
- Demography, definition of, 4573
- Demonstration of the Treponema pallidum in the brain in cases of general paralysis (Noguchi), 4537
- Demophobia, 4573
- Demyelinating disorders, **561–564**
- Demyelination, inflammation and, 659
- Dendrites
  - description of, 4573
  - function of, 3
  - neuronal, in aging brain, 3990–3991
  - spines on, 16, 17*f*
- Dendy, Walter Cooper, 4535
- Denial
  - definition of, 4573
  - self-protective processes and, 2050
- Dennett, Daniel, 330
- Dentate gyrus, anatomy of, 33, 34*f*
- Dentate nuclei, anatomy of, 26
- Deontology, physician ethics, 4441–4442, 4442*t*
- Deoxyribonucleic acid (DNA). *See* DNA
- Depacon. *See* Valproic acid (valproate, Depakene, Depakote)
- Depakene. *See* Valproic acid
- Depakote. *See* Divalproex
- Department of Defense, U.S., 2573
- Department of Health and Human Services (DHHS)
  - evidence-based care, support from, 4308
  - IMD exclusion study by, 4335
  - Medicare/Medicaid improvements from, 4316–4317
  - National Standards for Cultural and Linguistically Appropriate Services (CLAS), 4380
  - outcomes research support from, 4319
- Dependence. *See also* Addiction; Alcohol dependence
  - anabolic-androgenic steroid-induced, 1396–1399, 1398*f*
  - anabolic steroid-associated, 1402–1403
  - barbiturate-related, 1384
  - sedative-, hypnotic-, anxiolytic-related, 1384
  - tobacco, 1345

- Dependence on therapy, 4573
- Dependency, definition of, 4573
- Dependent personality disorder, 2156–2157
  - clinical criteria, 2156–2157
  - comorbidity, 2157
  - complications, 2157
  - description of, 4573
  - epidemiology, 2157
  - factitious disorder imposed on self and, 1851
  - familial patterns, 2157
  - genetics of, 2157
  - impairment, 2157
  - self-disorders and, 2808
  - sex ratio, 2157
  - STDP and, 2801
- Dependent variables, 4573
- Depersonalization
  - chronic, PTSD emotional numbing and, 1902
  - CNS response and, 1891
  - components, 1943
  - description of, 4573
  - disaster and military operations and, 2568
  - DSM and ICD diagnostic criteria, 1874
  - HPA axis and, 1890
  - hypnosis, 2712
  - neurobiology, 1890–1892
  - neurochemical systems and, 1891–1892
  - TRASC and, 1889
  - violence and, 2491
- Depersonalization disorder. *See also* Depersonalization;Derealization disorder
  - case study, 1881
  - course and prognosis, 1944
  - cross-cultural aspects, 1943
  - definition, 1941
  - diagnosis and clinical features, 1943
  - diagnostic criteria changes, 1881
  - differential diagnoses, 1943–1944

- dissociative identity disorder and, 1911
- dissociative PTSD subtype and, 1887
- DSM and ICD diagnostic criteria, 1874
- epidemiology, 1887, 1888
- etiology, 1941–1943, 1941*f*
  - case studies, 1942
  - cognitive theories, 1943
  - traumatic stress, 1942
- history, 1870
- prostitutes and, 1889
- treatment
  - pharmacotherapy, 1945
  - psychotherapy, 1944–1945

- Depixol. *See* Flupentixol
- Depolarization, definition of, 132
- Depressant drugs, 1266, 1267
- Depressed mood, 1634
- Depressio sine depressione, 1634
- Depression. *See also* Beck Depression Inventory (BDI); Bipolar depression; Center for Epidemiological Studies Depression Scale (CES-D); Hamilton Rating Scale for Depression (HRDS); Major depression; Major depressive disorder; Perimenopausal depression; Postpartum depression; Unipolar depression
  - 5-HT_{1A} receptor availability, 302
  - adolescent, 3767
  - after brain injury, 475
  - after CABG surgery, 2187
  - after stroke, 475, 476
    - course/prognosis in, 484–485, 486*f*
    - diagnosis of, 481
    - etiology of, 478–479
    - laboratory studies, 482–483
    - treatment of, 487
  - after traumatic brain injury, 531–533
  - in aging adults, 3976, 4010–4012
  - in alcohol abuse, 4112
  - alcohol and, 1270*t*, 1272
  - in Alzheimer disease, 4083
  - anaclitic, 1662, 3663, 3677

- animal models of, 83
- atypical, 1653, 2171, 2172*t*, 3118–3119
- B-family vitamins and, 2548
- BD, acute, 1690–1692
- BDD and, 1790, 1791
- BDNF polymorphisms in, 102
- behavioral activation and, 2698–2699, 2698*f*
- biogenic amine imbalance
  - hypotheses, 1607–1608
  - neuroendocrine links, 1608–1609
- bipolar, 3693, 3707
- brain diseases and, 464–465
- brain imaging, 1717
- brain tumors and, 497–500
- in breast cancer patients, 2269
- brief psychotherapies for, 3742, 3747
- in burn patients, 2378, 2379
- CAD and, 2187–2188
  - possible mechanisms of, 2188
- CAM treatments and, 2542–2543
- in cancer patients, 2253
  - management of, 2261
  - pharmacological treatments, 2261–2262
- cardiac arrhythmias and, 2190
- in cardiac patients, 2196
- cardiac symptoms of, 2195
- cardiovascular disease and, 2321
- CBF changes in, 5
- CBT for, 4182, 4191
- child/adolescent, 2281
  - episodes of, 3686, 3700
  - group therapy and, 3766–3767
  - interpersonal psychotherapy (IPT), 3742
  - pharmacogenomics of, 3455
- childhood epilepsy and, 649
- cholesterol and triglyceride levels, 2545
- cognition in, 698

- cognitive theory of, 1606, 1664–1665, 2761
- cognitive therapy, 2764–2765
- consciousness and, 691–692
- contemporary models, 1605–1610
- conversion disorder and, 1840
- COPD and, 2226
- CSF CRF levels in, 92
- (D) Scale, MMPI, 999*t*
- dementia and, 1196, 1198, 1202
- description of, 4573
- diabetes and, 2231–2232, 2238*t*
  - symptom burden and, 2232–2233, 2233*f*, 2233*t*
- diagnosis of, 2259–2260
- differential diagnoses, 1184, 1238, 2624*t*
- disasters and, 2568–2569
- dissociative identity disorder and, 1905
- double, 1642*f*, 1646
- EEGs in, 250
- effect on functioning, 2233
- effect on medical costs, 2235, 2235*f*
- effects on immunity, 190*t*
- in elderly patients with cancer, 2282–2283
- electroconvulsive therapy, 4153
- Endicott substitution criteria, 2299*t*
- in epilepsy, 509–510
- episodes of, child/adolescent, 3686, 3700
- estradiol for, 2523
- evaluation of, 2260–2261
- everyday functioning and, 4027, 4038
- in Fabry disease, 656
- family therapy, 3781–3782
- FASD and, 638
- GABAergic systems and, 83
- gender differences, 4246
- group therapy, 4201, 4207
- habenula in, 33
- HCV infection and, 587

- in HD, 555–556
- HIV infection and, 4270
- HPA dysregulation and, 2188
- hyperthyroidism and, 679–680
- hypnosis and, 2723
- hypothyroidism and, 3242, 3243
- IFN-α and, 201
- illness anxiety disorder and, 1838
- immune responses and, 2320
- insomnia and, 2086
- insulin sensitivity index in, 2247
- interleukin 6 in, 191*f*
- interpersonal theory of, 1665–1667
- interview scenarios, 961
- introjective, 1662
- late-onset, 465
- as leading disability, 4288
- learned helplessness model, 1606–1607
- in LGBT children, 2002
- long-term care and, 4226, 4229–4230
- loss/anger turned inward and, 1661
- loss of self-esteem and, 1606
- malingering of, 2406
- and malnutrition, 4226
- masked, 1634
- maternal separation and, 2047
- MBCT and, 2853
- medically illness and, 3237–3238
- medication adherence and, 2234*f*
- memory impairments, 739
- menopausal status and, 2523
- migraine and, 622
- minor, 1647*f*
- mortality rate with, 4529
- in multiple sclerosis patients, 562
- and natural selection, 776
- neuroendocrine links, 1608–1609

- neurosteroids and, 117
- nonseasonal unipolar, 213–214
- obesity and, 2214
- object loss and, 1605–1606
- OCD and, 1788
- in older people, 3956, 3964–3965
- omega-6 fatty acids and, 116
- pain and, 2393
- panic disorder and, 1783
- Parkinson disease and, 544, 545, 548, 4047
  - treatment of, 3237–3238
- pathological dissociation and, 1884
- pharmacological treatment (*See* Antidepressant(s); Mood stabilizers)
- physician, 2599
- polycystic ovary syndrome and, 2248–2249
- post-partum, 116
- poststroke, 2195
- premature births and, 644
- prevalence in cancer patients, 2260
- primary care role in, 4349–4350
- in psychiatric disorders, 1136–1137
- psychoanalytic descriptions of, 1661–1662, 1663
- as psychobiological affective reaction types, 1604–1605
- psychotic
  - ECT for, 3286
  - FGAs for, 3070
  - MAOIs for, 3120
- PTSD and, 1815
- rating scales for, 3399–3400
- reactive, 1647
- reinforcement and, 1607
- in rheumatoid arthritis, 2322
- right hemisphere injury and, 449
- risk in cancer patients, 2260
- risk-taking, premature births and, 643
- rodent models of, 396
- schizophrenia and, 1420–1421, 1421*t*, 4096

- self-medication, alcohol in, 579
- sexual dysfunction and, 1972
- sexual minorities and, 646
- sexually abused children and, 3832
- sleep deprivation and, 353
- social adversity and, 189*f*
- Social Functioning Examination in, 532*f*
- somatic presentations of, 2203–2204
- specific phobias and, 1784
- spirituality and, 2587–2589
- STDP and, 2811
- stress and, 1609, 2321
- stress and animal modes of, 1712–1713
- stroke and, 2194–2195
- substance use and, 646
- suicide and, 1669, 2612
- in survivors of critical illnesses, 2330
- TCAs for, 69
- in terminal illness, 2299–2300
  - diagnosis, 2299, 2299*t*
  - prevalence of, 2299
  - treatment, 2299, 2299*t*
- in terminally ill children, 2309
- in transplant populations, 2366
- trauma and, 1914
- treatment-resistant, 1680, 3119–3120
  - magnetic seizure therapy for, 3310
  - MAOIs, 3119–3120
  - sympathomimetics and, 3229
  - VNS for, 3311
- treatments of, 4551
  - carbamazepine and, 3031, 3031*f*–3032*f*–, 3031*t*, 3033
  - DA receptor agonists, 3237–3238
  - factitious disorder and, 1851
  - FGA, 3070
  - lithium, 3089–3090, 3089*f*
  - magnetic seizure therapy, 3310

    - MAOIs, 3119–3121
    - nefazodone, 3126, 3127
    - tDCS, 3308
    - thyroid hormones, 3242, 3243–3246, 3244*t*, 3245*t*
    - venlafaxine, 3148–3149, 3149*f*
  - unipolar, 1600, 1642*t*
  - world-wide prevalence of, 4518, 4520, *4521t*
- Depression-executive dysfunction, 4068–4069
- Depressive behaviors
  - animal models, 321–322
  - inheritance of, 396–397
- Depressive disorders. *See also* Depression; Major depressive disorder
  - after TBI, 653
  - alcohol-induced, 1272
  - bipolar disorders and, 1642
  - in cancer patients, 2259–2263
  - categories, 1643–1648
  - children and adolescents, **3674–3680, 3688–3694**
    - clinical features, 3676, 3690
    - comparative nosology, 3674–3675, 3688–3689
    - course, 3677–3678, 3691–3692
    - defined, 3674, 3688
    - diagnosis, 3676, 3690
    - differential diagnoses, 3677, 3691
    - epidemiology, 3675, 3689
    - etiology, 3675–3676, 3689–3690
    - history, 3674, 3688
    - laboratory examinations, 3676–3677, 3690–3691
    - prognosis, 3677–3678, 3691–3692
    - psychotherapy, 3679–3680, 3693–3694
    - rating scales, 3677, 3691
    - treatment, 3677–3680, 3679*t*, 3691–3694, 3693*t*
    - treatment response, 3680, 3694
  - classification of, 1154, 1158, 1168
  - continuum model, 1646–1647, 1647*f*
  - differential diagnoses, 1198, 1275, 3601
  - hallucinogen-induced, 1326–1327

- inhalant-induced, 1340
- preventive interventions for, 3897
- reactive attachment disorder and, 3665–3666, 3679–3680
- somatic symptoms, 1828*f*
- treatment
  - bupropion, 3005
- Depressive episodes, 1651, 1682
- Depressive neurosis. *See* Dysthymia
- Depressive pseudodementia, 1657
- Depressive syndrome
  - after disasters, terrorism, or war, 2568
  - cognitive disturbances in, 1637–1638
  - mood disturbances in, 1634, 1635*t*
  - psychomotor disturbances in, 1635–1637, 1636*f*
  - psychopathology, 1634–1639
  - psychotic features
    - mood-congruent, 1637
    - mood-incongruent, 1637
  - vegetative disturbances in, 1638–1639
- Depth psychology, 4573
- Derailment, 1420, 4573
- Derealization
  - components, 1943
  - description of, 4573
  - disaster and military operations and, 2568
  - DSM and ICD diagnostic criteria, 1874
  - intermittent explosive disorder and, 2495
  - TRASC and, 1889
  - violence and, 2491
- Derealization disorder. *See also* Depersonalization
  - general population studies, 1883*t*
- Dereism, 4573
- Dereistic thinking, 4573
- Dereliction of physician, 4436
- Dermatitis artefacta (DA), 2353*f*, 2354
  - excoriation disorder and, 1797
- Dermatological disorders

- characteristics of, 2333
- distress and, 2334*f*

- Dermatological effects. *See also* Skin
  - buspirone, 3012
  - carbamazepine, 3036
  - lamotrigine, 3085
  - lithium, 3098
  - MAOI, 3117
- Dermatology, hypnosis and, 2730
- Dermatomyositis (DM), 634
- Dermatoses, self-induced, 2349
- Derrida, Jacques, 929
- DES-Taxon Scale (DES-T), 1876*t*, 1883, 1895, 1895*t*
- *Description of the Retreat, an Institution near York, for Insane Persons of the Society of Friends* (Tuke), 4535
- Descriptive psychiatry, 4573. *See also* Dynamic psychotherapy
- Descriptive statistics, 4573
- Desentization. *See* Systemic desentization
- Desferrioxamine, 672
- Designer drugs, 4573–4574
- Designs of studies, 3996
- Desipramine (Norpramin)
  - ADHD, 3595
  - BDD, 1792
  - chemistry, 3254, 3254*f*
  - child/adolescent
    - ADHD treatment, 3788–3789
  - depression, 1681
  - for diabetic neuropathy, 2236*t*
  - dissociative identity disorder
    - PTSD symptoms, 1928
  - dosage, 3119
  - dosage/administration, 3264, 3265, 3265*t*
  - dosages, 573*t*
  - drug interactions, 3263–3264
  - drug studies, 3259
  - in dying patients, 2292
  - elimination disorders, 3655, 3669

- movement disorder, stereotypic, 3673, 3687
- OCD, 1789
- organ/system effects of, 3260–3261
- pain management using, 2292
- pharmacological actions, 3255–3256, 3256*f*, 3258
- precautions/adverse reactions, 3263
- side effects, 573*t*
- stereotypic movement disorders, 3672, 3686
- targets for, 69
- therapeutic indications, 3261
- tic disorder, 3646, 3660
- TNF-α release and, 202
- TTM, 1795

- Desire
  - hypoactive sexual desire, 1971
  - male hypoactive sexual desire disorder, 1965–1966
  - sex differences in, 1959–1961
- Desmethyldiazepam (Nordazepam)
  - breast milk excretion of, 3168
  - metabolism, 2993
- Desmethylsertraline, 3160
- Desmopressin (DDAVP), 3653–3654, 3655–3657, 3667–3668, 3669–3671, 3794
- Desmopressin acetate, 2102
- Despine, Antoine, 1866
- Destabilization, 2817
- Desvenlafaxine (Pristiq)
  - anxiety disorders, 1779–1780
  - chemistry, 3147, 3147*f*
  - depression, 1679
  - depression treatment, 4124
  - discontinuation syndrome, 3150
  - dosage, 573*t*
  - drug interactions, 3150
  - history, 3147
  - hypertension, 3150
  - overdose, 3150
  - in pain syndromes, 2395

- pharmacology, 3147–3148
- sexual dysfunction and, 1779*t*
- side effects, 573*t*, 1679*t*
- therapeutic indications, 3148–3149
- tolerability, 3149–3150
- tolerance of, 2261

- Desyrel. *See* Trazodone
- Detachment, description of, 4574
- Determinism, 4574
- Detoxification
  - addiction therapy and, 580–581
  - for alcohol use disorders, 1276
  - benzodiazepines, 1386–1387, 1386*t*
  - description of, 4574
  - in jails and prisons, 4467
  - opioid-related, 1368–1370
  - opioid withdrawal treatment, 1368–1370
  - opioids, 1368–1370
  - sedative-, hypnotic-, anxiolytic-related, 1387
  - for substance abuse, 4117
- Detumescence, 4574
- Deutsch, Felix, 2645
- Deutsch, Helene, 2645, 4539
- Devaluation, 4574
- Development
  - ADHD and, 3596
  - ADHD treatment effects, 3594
  - adolescent, **3364–3374**
    - of brain, 3370–3371
    - cognitive, 3369
    - defined, 3365
    - employment, 3943
    - environmental influences on, 3371–3374
    - evaluation consideration, 3385–3386
    - growth tasks, 3387, 3387*t*
    - of morality, 3369–3370
    - normal, 3365–3366

- online gaming's effect on, 3373–3374
- overview, 3364–3365
- physical, 3368–3369, 3368*t*
- risk-taking behavior and, 3372–3373
- of self-esteem, 3369, 3370
- of sexual behavior, 3371–3372
- social, 3369
- social media and Internet effect on, 3373
- stages of, 3366–3368

- adoption considerations, 3825
- adulthood, 3941–3942
- aggression and, 2484–2487
- autism spectrum disorder and, 3579–3580, 3579*f*, 3580*t*
- behavior therapy and, 2700–2701
- brain
  - gender differences, 3462–3463, 3463*t*
  - gray matter, 3361*f*, 3461–3462, 3462*f*, 3462*t*
  - neuroimaging of, **3461–3463**
  - white matter, 3462, 3462*f*, 3463*f*
- brief psychotherapies, 3744–3745
- cannabis-related effects on, 1308
- child (normal), **3351–3364, 3352–3357***t*
  - attachment and, 3662, 3663*t*, 3676, 3677*t*
  - of brain, 3351, 3356–3357*t*, 3358
  - cognitive, 3356–3357*t*, 3359–3360
  - emotional, 3352–3353*t*, 3361
  - of gender identity, 3354–3355*t*, 3363
  - of language, 3358–3359
  - of morality, 3354–3357*t*, 3361
  - motor, 3361–3362
  - overview, 3351, 3352–3357*t*
  - physical, 3356–3357*t*
  - psychosexual, 3354–3357*t*, 3363–3364
  - sensory, 3362–3363
  - social, 3352–3353*t*, 3360–3361
  - of temperament, 3352–3353*t*, 3361
- child evaluation criteria, 3375–3376, 3377

- cognitive and academic, 3377–3378
- conduct disorder and, 3619–3620, 3619*f*
- delays
  - autistic disorder differentiated from, 3579–3580, 3580*t*
  - encopresis, 3652, 3660, 3666, 3674
  - encopresis and, 3660, 3674
  - enuresis, 3652, 3666
- developmental coordination disorder and, **3537–3544**
- disabilities, 3786
- dissociation and, 1869
- dissociative amnesia, 1892
- dissociative identity disorder and, 1884, 1899
- as emergency child psychiatric evaluation consideration, 3441–3442
- of executive functioning, 3430
- group psychotherapy and, 3761–3762
- human brain
  - genomics of, 438–447
  - spatial-temporal transcriptional dynamics, 439
- infant growth tasks, 3388
- with intellectual disability, 3508–3513
  - child-related aspects of, 3509
  - contextual aspects of, 3509–3510
- Jacobson and Mahler on, 2644
- language disorders and, 3545
- latency, 3943
- neuroimaging studies, 1759
- phases of
  - in boys, 877
  - castration complex, 877
  - in girls, 877–878
  - Oedipus complex, 877
  - oral phase and objects, 874–877
  - phallic phase and objects, 877
  - pregenital phase, 874
- phases of development, 875–876t
- as preschool psychiatric evaluation criteria, 3409–3410
- psychic equivalence stage, 2870–2872, 2873*f*

- psychodrama and, 2858
- psychopharmacology, 3787–3788
- rates of, 3509
- schismatic family and, 2874–2875
- schizophrenia and, 1448–1449, 1451–1453, 1461–1463
- sexual orientation and, 2002–2004
- skewed family and, 2874–2875

- Development and Well Being Assessment (DAWBA), 3397
- Developmental arithmetic disorder, 4574. *See also* Gerstmann-Sträussler-Scheinker (GSS) disease
- Developmental articulation disorder, 4574
- Developmental coordination disorder, **3537–3544,** 3540*t*
  - ADHD and, 3538, 3541
  - clinical features of, 3539–3552, 3540*t*, 3541*t*
  - comparative nosology, 3537–3538
  - course, 3542
  - defined, 3537, 3537*t*
  - description of, 4574
- Developmental Coordination Disorder Questionnaire (DCD-Q), 3538, 3539
- Developmental disabilities
  - as comorbidities, 4343
  - in jails and prisons, 4468–4469
  - patient assessment in, 472
- Developmental disorders
  - ADHD and, treatment of, 3228
  - description of, 4574
- Developmental expressive language disorder, 4574
- Developmental group therapy, 2617
- Developmental history, interview, 951
- Developmental reading disorder, 4574
- Developmental receptive language disorder, 4574
- Developmental trait personality models, 2145
- Developmental word blindness. *See* Developmental reading disorder
- Developmentally informed hypnotherapy, 2716
- Dewar flasks, 264
- Dewey, John, 2842
- Dexamethasone (Decadron)
  - HPA activity and, 1715–1716, 1744

- mood disorders and, 1718
- pain management, 2292
- side effects
  - delirium, 2255
- Dexamethasone suppression test (DST)
  - for Cushing syndrome, 2244
  - depression and, 1608, 1716
  - description of, 4574
  - dissociative identity disorder and, 1882
- Dexmedetomidine (Precedex)
  - delirium treatment, 1188
  - for opioid withdrawal
    - in burn patients, 2381
  - use on the critically ill, 2331
- Dexmethylphenidate (Focalin), 2264, 3593
- Dextroamphetamine (Adderall, Dexedrine)
  - ADHD, 3593
  - for fatigue, 2264
  - use in cancer patients, 2262
- Dextroamphetamine/amphetamine (Adderll XR), 2295
- Dextromethorphan/quinidine (Neudexta), 488
- DHA. *See* Docosahexaenoic acid (DHA)
- DHEA. *See* Dehydroepiandrosterone (DHEA)
- Dhyana, 2578–2579
- Di-fenfluramine, 2220
- Diabetes
  - gambling and, 1805
  - and HIV infection, 4271
  - psychiatric problems, 4048
  - and sexual dysfunction, 4262
- Diabetes insipidus, 525, 2263
- Diabetes mellitus, **2230–2239**
  - complications of, 2234–2235
  - depression and
    - cardiac risk factors and, 2234
    - clinical approach to treatment, 2237–2238
    - medication adherence and, 2234*f*

    - treatment trials, 2235–2237
  - education groups and, 2236
  - mortality rates in, 2235
  - nature of, 2230
  - neuropathies and, 629
  - obesity epidemic and, 2230
  - psychiatric comorbidity in, 2231–2232, 2247
  - psychiatric illness and, 2234–2235
  - psychiatric symptoms in, 2230–2238
  - schizophrenia comorbidity, 1540, 1547
  - self-management of
    - patient-provider relationships and, 2231
    - psychiatric comorbidities and, 2233–2234
  - sexual function and, 1970
  - type 1
    - definition of, 2246
    - developmental issues in, 2231
    - epidemiology of, 2246
    - etiology, 2246
    - family issues in, 2231
    - nature of, 2230
    - treatment, 2247
  - type 2
    - Alzheimer disease and, 2232
    - cognitive decline in, 2232
    - definition of, 2246
    - diagnosis, 2246
    - epidemiology, 2246
    - etiology, 2246
    - nature of, 2230
    - nosology, 2246
    - treatment, 2247
- Diabetic neuropathy
  - treatment of, 2236, 2236*t*
    - anticonvulsants, 2975, 2977
- Diacylglycerol (DAG), 88–90
- Diagnoses. *See also* Medical assessment and laboratory testing; *specific diagnoses*

- Alcohol Use Disorder and Associated Disabilities Interview Schedule, 821
- Composite International Diagnostic Interview, World Mental Health Surveys version, 821
- fear of, physician mental health and, 2608
- personality traits, 940–941
- psychiatric, role of culture, 788–789
- validation, hospitalists and, 4356

- *Diagnostic and Statistical Manual of Mental Disorders* (DSM)
  - on acute stress disorder, 1874, 1878, 1907
  - on adjustment disorder(s), 1874
  - on agoraphobia, 1721
  - alcohol and substance abuse terminology, 4109–4110
  - Alzheimer disease criteria, 4079
  - on anxiety disorders, 1721
  - Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV), 2685
  - on attention deficit/hyperactivity disorder, 3587–3588
  - behavior syndrome classification, 522*t*
  - on body dysmorphic disorder, 1785
  - borderline personality disorder criteria, 1874
  - case identification, 818–819
  - classification of personality disorders, 2150–2151
  - cultural formulation, 789
  - Cultural Formulation Interviews, 2464–2469, 2466*t*–2467*t*
  - delirium, 4087
  - depersonalization disorder diagnostic criteria, 1874
  - derealization diagnostic criteria, 1874
  - description of, 4574
  - in diagnosis, 4342
  - dissociative amnesia diagnostic criteria, 1874
  - dissociative convulsions diagnostic criteria, 1874
  - dissociative disorder, DSM-III and beyond, 1870–1871
  - dissociative disorder diagnostic criteria, 1874
  - on dissociative disorders, 1870–1871, 1874, 1878
  - dissociative disorders diagnostic criteria, 1874
  - dissociative fugue diagnostic criteria, 1874
  - on dissociative identity disorder, 1866, 1874
  - dissociative stupor diagnostic criteria, 1874
  - DSM-IV, mental disorder classification in, 1167–1169, 1167*t*, 1168–1169

- DSM-5 publication, 4555
- DSM-ICD Harmonization Coordinating Group, 1175–1176
- DSM-IV-R on bereavement, 1154
- dysthymia (dysthymic disorder), criteria for, 1645*t*
- everyday functioning, 4025
- on excoriation (skin-picking) disorder, 1785
- on flashbacks, 1874
- fugue diagnostic criteria, 1874
- on gambling disorder, 1801–1802
- on Ganser syndrome, 1874
- gender dysphoria diagnosis, 2028–2029
- on generalized anxiety disorder, 1721, 1723
- histrionic personality disorder diagnostic criteria, 1874
- on hoarding disorder, 1785
- on hypomania, 1640*t*
- hysteria, diagnostic criteria, 1874
- major depressive disorder criteria, 1635*t*
- on manic episodes, 1639
- mental disorder classification in
  - advantages of, 1155
  - alternatives to, 1161–1163, 1161*t*
  - changes for, 1156–1159, 1157*t*
  - complexity of, 1155–1156
  - concerns with, 1159
  - depression classification in, 1154
  - evolution of, 1152, 1153–1154, 1153*f*
  - global issues for, 1159–1160
  - ICD-11, comparisons with, 1174–1176
  - ICD-11, harmonization with, 1175–1176
  - mental disorder defined by, 1152
  - revision of, 1160–1161
  - shortcomings of, 1155
- on mental illness classification, 1233, 1233*t*
- on mild cognitive impairment, 1223
- mood disorders defined in, 1599, 1631
- multidimensional approach of, 1114–1115, 1118
- multiple personality disorder diagnostic criteria, 1874

- NOS replacement by, 1119
- on obsessive-compulsive and related disorders, 1721, 1785, 1786
- on obsessive-compulsive disorder, 1723
- Other Conditions That May Be a Focus of Clinical Attention, 2824
- on panic disorder, 1721, 1723*t*
- persistent depressive disorder criteria for, 1644–1645, 1645*t*
- on personality disorders, 2158
- Personality Inventory for *DSM*-5, 1874
- on phobias, 1721, 1723
  - DSM-5 on, 1721
- on posttraumatic stress disorder, 1874
- posttraumatic stress disorder diagnostic criteria, 1874
- prevalence of mental disorders in older adults, 3960
- psychiatric syndromes in medically ill patients, 4042–4043
- on psychocutaneous disorders, 2350*t*
- psychogenic amnesia diagnostic criteria, 1874
- psychogenic fugue diagnostic criteria, 1874
- schizophrenia, 4095
- on selective mutism, 1723
- on separation anxiety disorder, 1723
- on social anxiety disorder, 1721, 1723
- on social phobia, 1721
- on somatic symptom and related disorders, 1831*t*
- substance/medication-induced anxiety disorder, 1723
- on Tourette syndrome, 3636, 3650
- on trance and possession disorders, 1874
- on trauma- and stressor-related disorders, 1874
- on trichotillomania, 1785
- on unipolar disorders, 1631

- Diagnostic Interview for Children and Adolescents (DICA)
  - child and adolescent epidemiology, 821
  - overview, 3397
- Diagnostic Interview for Personality Disorders (DIPD), 2161
- Diagnostic Interview Schedule (DIS), 820, 1836, 2685, 3961
- Diagnostic Interview Schedule for Children (NIMH-DISC-IV), 821, 1721, 3396–3397
- Diagnostic interviews, 3395–3398
  - interviewer-based, 3396, 3397–3398

- overview, 3395
- respondent-based, 3395–3397

- Diagnostic methods and measures. *See* Standardized assessment instruments
- Diagnostic-related groups (DRGs), 4336
- Dialectical behavior therapy (DBT), 2699–2700, **2784–2795**
  - for adolescents, 2795, 3814
  - aggression in BPD, 2501
  - alternate uses for, 2618
  - behavioral analysis, 2789–2790
  - borderline personality disorder, 2784
    - biosocial theory of, 2786–2787, 2786*f*, 2787*t*
  - BPD with dissociative symptoms, 1919
  - children, 2795
  - clinical issues, 2791–2795, 2792*f*
    - agreements, 2791, 2792*t*
    - assumptions, 2791, 2792*t*
    - case study, 2793–2794
    - treatment levels and targets, 2791–2793
  - consultation team, 2791, 2791*t*
    - ethical challenges, 2794
  - depression, hypnosis and, 2723
  - description of, 4574
  - dialectical strategies, 2790, 2791*t*
  - dialectics, 2785–2786
  - diary cards, 2789, 2789*f*
  - dissociative identity disorder, 1930
  - efficacy research, 2794–2795
  - ethical issues, 2794
  - evaluation, 2863
  - group, personality disorders, 2746
  - history, 2785
  - introduction and definition, 2784–2785
  - mood disorders, 1701
  - research and evaluation, 2794–2795
  - in sex addiction, 2053, 2055
  - skill enhancement, 2790
  - skills modules

    - distress tolerance, 2788
    - emotion regulation, 2788
    - interpersonal effectiveness, 2788
    - mindfulness, 2787–2788
  - suicide prevention and, 2621
  - suicide risk, 2617
  - techniques, 2787–2791
    - group skills training, 2787–2788
    - individual therapy, 2788–2789
    - overview, 2787
  - telephone consultation, 2790–2791
  - theoretical issues, 2785–2787
  - TTM, 1795
  - validation, 2790, 2790*t*
- *Dialogues of Plato* (Plato), 4533
- Dialysis, issues with, 2369
- Dialysis encephalopathy, 2369
- Diamond, Bernard, 2733
- Diarrhea in dying patients, 2296–2297
- Diary cards, DBT, 2789, 2789*f*
- Diathesis stress model of temperament, 3487
- Diathesis-stress paradigm, 4574
- Diathesis–stress model, 2763
- Diazepam (Valium)
  - abuse liability, 1377–1378
  - agoraphobia, 1783
  - alcohol withdrawal, 2997
  - anxiety disorders
    - GAD, 1783
  - benzodiazepine dose equivalent, 2633
  - distribution, 2993
  - dosage, 635
  - drug studies, 2995
  - history, 2917, 2991
  - MDD, 2996
  - metabolism, 2993
  - pharmacology, 1375–1376

- pregnancy and, 2999
- sedation, 2997
- seizure disorder, 2997
- SPECT in panic disorder and, 1761–1762
- for stiff-person syndrome, 635
- tolerance, 1375
- withdrawal treatment, 1386–1387

- Diazepam-binding inhibitor (DBI), 2994, 3000
- Diazoxide, 138
- Dibenzoxazepines, 3063*f*, 3064
- DICA. *See* Diagnostic Interview for Children and Adolescents
- Dick, Henry, 2754
- DID. *See* Diffuse ischemic damage (DID)
- Didactic training, 3555
- DIDMOAD acronym, 670
- Diencephalic syndrome
  - in childhood, 677
  - MRI imaging of, 678*f*
- Diencephalon
  - derivatives, 41
  - medial, 732
- Diet and nutrition. *See also* Eating; Eating disorders (ED); Food
  - ADHD, 3590
  - as biological treatment, 2547
  - in MSUD, 661–662
  - in PKU, 660–661
  - positive psychiatry, 4285
  - tyrosine supplementation, 660–661
  - weight loss and, 2216*t*
- Diethylstilbesterol (DES), 1956
- Differential diagnoses, 3540–3542
  - definition of, 4574
  - psychological testing in, 994–995
- Differential reinforcement, 4574
- Differentiation, 4574
- Difficult children, temperament of, 3480, 3481
- Diffuse axonal injury (DAI), 4480–4481

- Diffuse ischemic damage (DID), 523, 524
- Diffuse-projection nuclei
  - function of, 22*t*
  - inputs, 22*t*
- Diffusion tensor imaging (DTI)
  - after TBI, 525
  - in aging adult, 4009
  - fiber tracks, 271*f*
  - in neuroimaging, 4481
  - pediatric considerations, 3459
  - schizophrenia, 1464, 1466*t*, 1467*t*, 1468, 1471–1474, 1472*f*
  - in schizophrenia, 273
  - structural neuroimaging with, 1759
- Diffusion-weighted imaging (DWI), 270–271, 525
- Digit span tests, 461
- Digit Symbol and Block Desigh Tests, 2368
- Digital media, adolescent use of, 3811–3812
- Digitaline. *See* Digitoxin
- Digitoxin (Digitaline), interactions, 3164
- Dignity therapy, 2825–2827
  - clinical issues, 2827
  - ethical issues, 2827
  - research and evaluation, 2827
  - techniques, 2826–2827
  - theoretical issues, 2826
- Digoxin (Digitek, Lanoxin)
  - drug interactions, 3131
  - psychiatric side effects of, 2198*t*
- Dihydroergotamine for migraine, 625
- Dihydroindoles, 3063*f*, 3064
- Dihydropyridine dehydrogenase *(DPYD)* gene, 2264
- Dihydropyridines, 139
- Dihydrotestosterone (DHT)
  - function of, 1955–1956
  - storage of, 169
  - synthesis of, 169
  - testosterone conversion to, 2506

- Dilantin. *See* Phenytoin
- Dillworth, Tiara, 2726
- Diltiazem (Cardizem)
  - chemistry, 3016
  - pharmacodynamics, 3018
  - pharmacokinetics, 3016–3017
  - precautions/adverse effects, 3021
  - therapeutic indications, 3020–3021
- Dim blue light therapy, 2558
- Dimenhydrinate (Dramamine), 75
- Dimensional traits, personality and, 575–576
- Dimers, description of, 164
- 2,5-Dimethoxy-4-methylamphetamine (DOM), 4574
- Dimethyltryptamine (DMT)
  - description of, 4574
  - pharmacology of, 1316
  - research, 2887
- Diminished libido, 4574
- Diphenhydramine (Benadryl)
  - for agitation, 3440–3441
  - chemistry, 2957, 2957*f*, 2981*f*
  - dissociative identity disorder, 1927*t*
  - dosage/administration, 2960
  - histamine receptor blocking by, 75
  - therapeutic indications, 2982
- Diphenylbutylpiperidines, 3063*f*, 3064
- Diprivan. *See* Propofol
- Dipropyltryptamine (DPT), 2893
- Dipsomania, 4574
- Direct contingency management, 3596
- Direct questioning, interview technique, 3384
- Directive hypnotherapy, 2716
- Directly experienced traumatic events, 1813, 1813*t*
- DIS. *See* Diagnostic Interview Schedule
- Disability
  - benefits
    - for elderly impaired, 4238

    - mental illness exclusion, 4528
  - case identification, 819–820, 819*t*
  - definition of, 4574
  - and depression, in older adults, 3969, 3975, 4067–4068
  - physician, 2602
- Disability Adjusted Life Years (DALYs), 827–828, 828*t*
  - anxiety disorders, 1722
  - definition of, 4288
  - mood disorders and, 1618
  - for schizophrenia, 1425, 1435–1436
  - world-wide, 4518, 4520, *4526t*
- Disability scales, developed/developing countries, 4518, *4520t*
- Disadvantaged families/communities, 4302
- Disaster planning in jails and prisons, 4460–4461
- Disaster psychiatry, **2564–2575**
  - battle fatigue and unexplained symptoms, 2567–2568
    - case study, 2568
  - epidemiology, 2568–2570
  - etiology, 2568–2570
  - future challenges and evolving issues, 2574–2575
    - cultural issues, 2574–2575
    - emerging threats, 2574
    - ethical challenges, 2575
  - management and care delivery, 2571–2574
    - case study, 2572–2573
    - general principles, 2571, 2571*t*
    - therapeutic interventions, 2571–2574
  - military mental health care in peacetime, 2573–2574
  - mitigating factors, 2570
  - neurobiological factors, 2569
  - overview, 2564–2565
  - perpetuating factors, 2570
  - precipitating factors, 2569–2570
  - predisposing factors, 2569
  - protective factors, 2570
  - symptoms *versus* functioning, 2567
  - traumatic stress responses and, 2566–2568

- Disaster(s)
  - care environments and, 2565–2566
  - definition, 2565
  - holding environment, 2571
  - psychiatric syndromes in aftermath of, 2567–2568
- Discharge planning
  - in jails and prisons, 4460–4462
  - in treatment, 4340–4341
- Discipline, jails and prisons, 4474
- Discoid rashes, assessment of, 456
- Disconnection syndrome, 4574
- Discontinuation syndrome
  - antidepressant, 2943
  - benzodiazepine, 2999–3000
  - description of, 4574
  - nefazodone and, 3130
  - SSRI, 3156–3157
  - SSRIs, SNRIs and, 1779
  - venlafaxine, 3150
  - vilazodone, 3345
- Discoordination hypothesis, 405
- Discordance, 4574
- Discourse analysis, 693
- Discrete behavioral states (DBS) model
  - childhood trauma and, 1900
  - dissociation, 1870
  - dissociative disorders, 1899–1900
  - dissociative identity disorder and, 1906
- Discriminant analysis, 4574
- Discriminant validity, 996
- Discrimination
  - description of, 4574
  - and psychiatric service utilization, 802–803
- Discriminative stimulus, 4574–4575
- Disease Control Priorities Project Report, 4522
- Disfigurement, skin lesions and, 2333
- Disgust, OCD and, 1786

- Dishabituation, 166
- Disillusionment phase, after disasters, 2566
- Disinhibited social engagement disorder (DSED), **3662–3667, 3676–3681**
  - attachment, 3662, 3663*t*, 3676, 3677*t*
  - clinical features, 3664–3665, 3678–3679
  - comparative nosology, 3663–3664, 3677–3678
  - defined, 3662, 3663*t*, 3676, 3677*t*
  - description of, 4575
  - diagnosis, 3664–3665, 3678–3679
  - differential diagnoses, 3665–3666, 3679–3680
  - environmental factors, 3663, 3677
  - epidemiology, 3664, 3678
  - etiology, 3664, 3678
  - overview, 3662, 3676
  - treatment, 3666–3667, 3680–3681
- Disinhibition
  - alcohol and, 1269
  - benzodiazepine, 2998
  - definition of, 4575
  - with dementia, 1197
  - TBI and, 471–472
- Disintegration, 4575
- Disk-over-water apparatus (DOW), 350
- Disorder of empathy, 1851
- Disorder of written expression, 4575
- Disorder-related scales
  - for anxiety, 3400–3401
  - for depression, 3399–3400
- Disordered patients, transplantation and, 2363–2364
- Disorderedness, ODD and, 3611
- Disorders of excessive somnolence. *See* Somnolence
- Disorders of sexual development (DSD)
  - definition of, 2023
  - diagnosis of, 2028
  - treatment of individuals with, 2035
- Disorganization syndrome, 1419–1420
- Disorganized attachment, violence and, 2486

- Disorganized schizophrenia, 4575
- Disorientation, 4575
- Disparities
  - definition of, 4304, 4377–4378
  - in drug access, 4352
  - sociocultural framework (*See* Sociocultural framework)
- Displacement, 4575
- Disposition, 4575
- Disrupted in Schizophrenia 1 *(DISC1)* gene, 1610, 1625
- Disruptive, impulsive-control and conduct disorders, 1788, 1789
- Disruptive behavior disorders, **3605–3621**
  - brief psychotherapies for, 3743–3744, 3747–3748
  - child/adolescent, **3605–3621**
    - overview, 3605
    - trazodone drug studies, 3250–3251
  - conduct disorder, **3612–3621**
    - comorbid conditions, 3618–3619
    - comparative nosology, 3613
    - course, 3619–3620, 3619*f*
    - defined, 3607*t*, 3612
    - diagnosis, 3618
    - differential diagnoses, 3618–3619
    - epidemiology, 3613
    - etiology, 3613–3618, 3617*f*, 3618*f*, 3618*t*
    - further research, 3620–3621
    - historical background, 3612–3613
    - laboratory examination, 3618
    - pathology, 3618
    - prognosis, 3619–3620
    - subtypes and specifiers, 3612
    - treatment, 3620
  - description of, 4575
  - oppositional defiant disorder, **3605–3612**
    - clinical features, 3610
    - comorbid conditions, 3611
    - comparative nosology, 3606, 3607*t*, 3608*t*
    - course, 3611

        - defined, 3605
        - diagnosis, 3610
        - differential diagnoses, 3610–3611
        - epidemiology, 3606
        - etiology, 3606–3609
        - history, 3605–3606
        - laboratory examination, 3610
        - pathology, 3610
        - prognosis, 3611
        - treatment, 3611–3612
    - psychopharmacology, 3788–3790, 3796
    - rating scales for, 3402
    - temperament and, 3485–3486, 3485*t*
- Disruptive disorders, diagnosis of, 2436–2437
- Disruptive mood dysregulation disorder (DMDD)
    - about, 1677
    - criteria for, 3349
    - defined, 3725
    - description of, 4575
- Disseminated encephalomyelitis, 563
- Dissimulation, definition of, 2399
- Dissocial personality disorder (DPD), 2490
- Dissociation
    - capacity for, 1906
    - definition, changes to, 1878
    - description of, 4575
    - dimensional, adaptational aspects of, 1895
    - false memory syndrome and, 1872
    - fantasy model, 1898
    - genetics and, 1889
    - hypnosis, 2709, 2710, 2711
    - iatrogenic model, 1897–1898
    - neurobiology of, 1889–1894
    - persistence of controversy about, 1873
    - in posttraumatic stress disorder, 701
    - psychoanalytic theory, 1903
    - repression *versus,* 1869

- sociocognitive model, 1897–1898
- spirituality and, 2577–2578
- theories, 1895, 1895*t*
- as trauma response, 1896–1898
- violence and, 2491–2492

- *Dissociation of a Personality* (Prince), 4537
- Dissociation syndrome, 4575
- Dissociative amnesia, **1931–1941**
  - betrayal trauma theory and, 1902–1903
  - biological basis, 1890
  - case study, 1932, 1933, 1934
  - clinical course and prognosis, 1937–1938
  - definition, 1931–1932
  - description of, 4575
  - diagnosis and clinical features, 1932–1934
  - diagnostic criteria changes, 1879–1880
  - differential diagnoses, 1935–1937, 1936*t*
    - acute stress disorder and, 1937
    - cognitive disorders, 1935
    - delirium and, 1935
    - dementia and, 1935
    - dissociative disorders and, 1937
    - factitious amnesia and, 1937
    - malingering amnesia and, 1937
    - normal aging and, 1935
    - other psychiatric disorders, 1937
    - posttraumatic amnesia, 1935–1936
    - PTSD and, 1937
    - seizure disorders, 1936
    - substance-related amnesia, 1936
    - transient global amnesia and, 1937
  - dissociative fugue subtype, 1933–1934
    - diagnostic criteria changes, 1879–1880
    - epidemiology, 1885, 1886
    - general population studies, 1883*t*
    - treatment, 1941
  - dissociative identity disorder and, 1908, 1908*t*, 1910

- DSM and ICD diagnostic criteria, 1874
- epidemiology, 1885
  - clinical populations, 1888
- factitious disorder and, 1859
- general population studies, 1883*t*
- history, 1869–1870
- ICD-10 on, 1875*t*
- neurobiology, 1892–1893
  - disturbed retrieval, 1893
  - encoding, 1892–1893
- nonclassic presentation, 1934, 1938–1939
- pathology and laboratory examination, 1934–1935
- prostitutes and, 1889
- reliability of traumatic memory and, 1873
- trauma and, 1886*t*
- treatment, 1938–1941
  - amnesia and, 1939, 1939*t*
  - facilitated recall of traumatic memory and, 1939
  - free recall, 1939
  - group therapy, 1941
  - hypnosis, 1939–1940
  - phase-oriented, 1938–1939
  - recovered memory therapy, 1938
  - somatic therapies, 1940–1941
  - transference interpretations, 1939
- types of, 1880, 1881*t*
- validity of, 1882–1883
- wandering behavior, 1936–1937

- Dissociative anesthesia and sensory loss, 1875*t*
- Dissociative anesthetics, 1312, 1314
- Dissociative convulsions
  - DSM and ICD diagnostic criteria, 1874
  - ICD-10 on, 1875*t*
- Dissociative disorder(s), **1866–1952**. *See also* Dissociative amnesia
  - atypical presentations, 1946
  - clinical populations, 1887–1889
  - comparative nosology, 1874, 1875–1878*t*, 1878

- DSM and ICD diagnostic criteria, 1874
- depersonalization disorder and, 1881
- description of, 4575
- diagnostic criteria changes, 1878–1881
- discrete behavioral states model, 1899–1900
- dissociative amnesia *versus,* 1937
- dissociative identity disorder, 1903–1931
- dissociativity, 1870
- DSM-5 changes to, 1878
- epidemiology
  - clinical populations, 1888
  - emergency department, 1888
  - general population studies, 1883–1887
  - non-western/secular settings, 1888
  - prostitutes, 1889
  - substance abuse populations, 1888–1889
- etiological theories, 1894–1903, 1895*t*
  - betrayal trauma, 1900–1901
  - culture, sociocognitive phenomena, 1899
  - developmental mediation of psychopathology, 1902–1903
  - developmental model, 1901–1902
  - developmental precursors and substrates, 1902
  - discrete behavioral states model, 1899–1900
  - taxon model, 1895–1896, 1895*t*
  - trauma model, 1896–1898
  - trauma *versus* fantasy model, 1898–1899, 1898*t*
  - type D attachment, 1901
- factitious disorder and, 1848–1849
- general population studies, 1883–1887, 1883*t*
- history, 1866–1871
  - depersonalization/derealization disorder, 1870
  - DSM-III and beyond, 1870–1871
  - international and cross-cultural research, 1871
  - Janet and Freud, 1868–1869
  - later 20th cent., 1870–1871
  - in 19th and early 20th cent., 1868
  - trauma and La Salpêtrière patients, 1867–1868

    - Wilbur, 1870
    - World War II and, 1869–1870
  - ICD-10 on, 1831, 1875*t*
  - introduction, 1866
  - legal issues, 1871–1873
  - modern society and the media, 1871–1873
  - neurobiology, 1889–1894
    - depersonalization, 1890–1892
  - NOS, 1874, 1875*t*
    - diagnostic criteria changes, 1878, 1881
    - general population studies, 1883*t*
  - other (specified), 1881, 1945–1952
    - acute dissociative reactions to stressful events, 1948–1950
    - atypical presentations, 1946
    - case studies, 1949, 1950, 1951–1952
    - coercive persuasion, 1946–1948
    - dissociative trance disorder, 1951–1952
  - persistence of controversy about, 1873
  - screening and interviews, 1876–1877*t*
  - trauma- and stressor-related disorders and, 1874
  - validity of, 1881–1883
- Dissociative Disorders Interview Schedule (DDIS)
  - dissociative amnesia, 1934
  - dissociative disorders, 1878
    - general population studies, 1883, 1883*t*
  - dissociative identity disorder, 1905–1906
  - inpatient studies, 1888
  - reliability and validity of, 1878*t*
- Dissociative disorders of movement and sensation
  - DSM and ICD diagnostic criteria, 1874
- Dissociative Experiences Scale (DES)
  - aggression and, 2491
  - development of, 1870, 1905
  - diverse cultural settings and, 1896
  - general population studies, 1883, 1883*t*
  - longitudinal studies, 1920
  - PTSD *versus* dissociative identity disorder and, 1882

- reliability and validity of, 1876*t*
- taxon model, 1895
- type D attachment and, 1901

- Dissociative fugue. *See also* Dissociative amnesia
  - DSM and ICD diagnostic criteria, 1874
  - ICD-10 on, 1875*t*
- Dissociative hallucinations, 1911
- Dissociative identity disorder, **1903–1931**
  - aggression and, 2491
  - case study, 1907
  - course and prognosis, 1917–1918
  - depersonalization, derealization and, 1881
  - description of, 4575
  - diagnosis and clinical features
    - autohypnotic symptoms, 1909*t*, 1913
    - case study, 1910, 1913
    - child and adolescent presentations, 1914–1915
    - differential diagnoses, 1915*t*, 1916
    - dissociative alterations in identity, 1912–1913
    - dissociative hallucinations, 1911
    - dissociative thought disorder, 1912
    - inter-identity amnesia, 1910–1911
    - interference symptoms, 1911–1912
    - memory and amnesia symptoms, 1908, 1908*t*
    - other associated symptoms, 1913–1914, 1915*t*
    - passive influence symptoms, 1911–1912
    - process symptoms, 1909*t*, 1911
  - diagnostic criteria changes, 1878–1879, 1880*t*
  - dissociative psychosis *versus,* 1949–1950
  - DSM-5 on, 1866
  - DSM and ICD diagnostic criteria, 1874
  - epidemiology, 1884
    - clinical populations, 1888
  - etiology, trauma, 1904–1906
  - explanatory models, 1906–1908
    - discrete behavioral states model, 1906
    - Four-Factor Theory, 1906–1907

- factitious, imitative, and malingered, 1916–1917
- fusion, integration, resolution, and recovery, 1926
- galvanic skin resistance and, 1868
- general population studies, 1883*t*
- history, 1866, 1870–1871
- hypnotizability and, 2707
- informed consent and, 1927
- neurobiology, 1893–1894
- as paradigmatic PTSD, 1904–1905, 1904*t*
- pathological dissociation and
  - epidemiology, 1884–1885
- pathology and laboratory examination, 1915–1916
- treatment, 1918–1931
  - adjunctive, 1929–1930
  - case study, 1931
  - cost efficacy studies, 1919
  - countertransference, 1925
  - dialectical behavior therapy, 1930
  - electroconvulsive therapy, 1928–1929
  - expressive therapy, 1930
  - eye movement desensitization and reprocessing, 1930
  - family therapy, 1929–1930
  - group therapy, 1929
  - hospitalization, 1929
  - hypnosis, 1926
  - longitudinal studies, 1919–1921, 1920*t*
  - medications for mood disorders, 1928
  - medications for PTSD symptoms, 1927–1928
  - occupational therapy, 1930
  - outcome studies, 1919
  - partial hospitalization, 1929
  - pharmacotherapy, 1926–1928, 1927*t*
  - phasic treatment, 1921*t*
  - self-help groups, 1930
  - stabilization and safety, 1922–1925
  - stages, 1921
  - transference, 1925

    - traumatic memories, 1925–1926, 1925*t*
  - validity of, 1882
  - violence and, 2491

- Dissociative motor disorders, 1875*t*
- Dissociative personality disorder, 2428
- Dissociative psychosis
  - about, 1881
  - dissociative identity disorder *versus,* 1949–1950
- Dissociative reactions, 3832
- Dissociative states
  - definition of, 2491–2492
  - dissociation *versus,* 2491–2492
- Dissociative stupor
  - DSM and ICD diagnostic criteria, 1874
  - ICD-10 on, 1875*t*
- Dissociative thought disorder, 1912
- Dissociative traits
  - aggression and, 2491
  - dissociation *versus,* 2491–2492
- Dissociative trance disorder
  - definition, 1951
  - description of, 4575
  - diagnosis and clinical features, 1951
  - general population studies, 1883*t*
  - treatment, 1951–1952
- Distal symmetrical polyneuropathy (DSP), 632
- Distortion, description of, 4575
- Distractibility, description of, 4575
- Distractible speech, 4575
- Distraction, anxiety disorders and, 1770–1771
- Distraction therapy, 2377
- Distress
  - in cancer patients, 2253
  - cultural concepts of, 789–791, **2443–2460,** 2444*t*
    - clinical applications, 2459
    - examples of, 2448
    - kinds of, 2445–2446
    - psychiatric classification and, 2447–2448
    - research into, 2458
  - definition of, 2336

- skin-specific, 2334*f*
- symptoms in patients with, 2252
- tolerance skills, 2055, 2788

- Distress Thermometer
  - assessing cancer patients, 2283
  - in cancer patients, 2268
- Distributive analysis and synthesis, 4575
- *Disturbances of the Mind* (Heinroth), 4535
- Disulfiram (Antabuse), **3050–3054**
  - acamprosate combined with, 3059
  - alcohol-related disorders, 3050, 3051–3052, 3053, 3059
  - for alcohol use disorders, 1278
  - BD and alcohol comorbidity, 1699
  - carbon disulfide production, 631
  - chemistry, 3050, 3050*f*
  - for cocaine dependence, 1289–1290
  - DBH inhibition and, 63
  - description of, 4575
  - dosage/administration, 3054
  - drug interactions, 3054, 3145
  - drug studies, 3051–3052
  - history, 3050
  - laboratory interference, 3054
  - marital therapy with, 3052
  - mechanism of action, 581–582
  - organ/system effects of, 3052–3053
  - pharmacological actions, 3050–3051
  - precautions/adverse reactions, 3054
  - PTSD and SUD comorbidity, 1823
  - side effects, 2371
  - substance use disorders, child/adolescent, 3793
  - therapeutic indications, 3053–3054
- Ditalopram, dosage, 1979
- Diuretics, drug interactions, 3098
- Diurnal enuresis, 3650, 3651, 3654, 3664, 3665, 3668
- Divalproex (Depakote)
  - acute mania, 1684–1685, 1689*t*

- for aggressive behavior
  - autism, 2500
  - BPD, 2501
  - in BPD, 2501
  - conduct disorder, 2501
  - mentally retarded children, 2500
- agitation, dementia-related, 3276
- alcohol relapse prevention, drug studies, 3274–3275
- alcohol withdrawal, drug studies, 3274
- BD and bipolar depression, 3270, 3271*t*, 3276
- benzodiazepine withdrawal, 1388
- bipolar depression, 1690
  - combination therapy, 1692*t*
- bipolar disorders, 1696*t*
  - combination therapy, 1696*t*
  - early onset, 3692, 3706, 3791
  - FDA approval of, 1612
- black box warning, 2919
- borderline PD, 3276
- chemistry, 3266
- for children with CD, 2501
- conduct disorder, 3790
- dissociative identity disorder, 1927*t*
- dosage/administration, 3279
- drug studies, 3268–3276
- for excoriation disorders, 2352
- lithium and, 3098
- lithium *v.,* 3270, 3272
- mania, 3268, 3269*t*, 3270, 3270*t*, 3276
- for migraine prevention, 623–624
- mood disorder due to another medical condition, 1251
- mood stabilizing actions of, 3270, 3272
- pharmacological actions, 3266
- schizophrenia, drug studies, 3272
- seizure disorders, 3276
- therapeutic indications, 3276
- tolerability in HIV patients, 575

- Diversity
  - dissociative disorders and, 1896–1897
  - family and couples therapy and, 2758
- Divorce, family therapy and, 3785
- Dix, Dorothea
  - on conditions of mentally ill, 4290, 4335
  - contemporary problems same as those reported by, 4345, 4389
  - work of, 4575
- *Dli3* gene expression, 49
- DLPFC. *See* Dorsolateral prefrontal cortex
- DMT. *See* Dimethyltryptamine (DMT)
- DNA
  - chromosomes and, 149
  - methylation of, 392–393
  - polymorphisms
    - gene mapping studies and, 306
  - repetitive, 153–154
  - variations
    - psychiatric disorders and, 3448–3449
- DNMT enzymes, 392–393, 398
- Docosahexaenoic acid (DHA)
  - for Alzheimer disease, 4148
  - biological treatment with, 2551
  - dietary supplements of, 115–116
  - schizophrenia and, 2545
  - therapeutic indications, 116
- Doctor-patient relationships, 946–947, 4433–4434
  - communications, 2290–2291
  - in correctional psychiatry, 4452–4453
  - description of, 4575
- Dodo effect, 2837
- Doll, Edgar, 3494
- Dollard, John, 930–931
- Dolophine. *See* Methadone
- Dolphins, sleep in, 352
- DOM. *See* 2,5-dimethoxy-4-methylamphetamine (DOM)
- Domain-specific scales. *See* Disorder-related scales

- Domestic abuse or violence. *See also* Intimate partner violence
  - aggression and, 2486–2487
  - as assault, 2538
  - DBT and, 2795
  - definition, 2539
  - perpetrators, 2539–2540
  - prevalence, 2539
  - psychiatric emergency due to, 2633–2634
  - risk factors, 2539
  - sequelae of, 2540
  - treatment, 2502–2503
- Dominance, description of, 4575
- Don Juanism, 1974
- *Donaldson, O'Connor v.,* 4336
- Donepezil (Aricept), **3042–3044**
  - after TBI, 529
  - aggression in Alzheimer disease, 2502
  - for aggressive behavior
    - in Alzheimer disease, 2502–2503
  - Alzheimer disease treatment, 1207–1209
  - amnestic disorder treatment, 1231
  - chemistry, 3042, 3042*f*
  - cholinesterase inhibition, 4143
  - dementia treatment, 1238
  - dosage/administration, 3044
  - drug interactions, 3044
  - drug studies, 3043
  - laboratory interferences, 3044
  - organ/system effects of, 3043–3044
  - for PD, 547–548
  - pharmacologic actions, 3042–3043
  - precautions/adverse reactions, 3044
  - therapeutic indications, 3044
  - uses of, 73
- Doolittle, Hilda, 2641
- Dopa, synthesis of, 63*f*
- Dopamine (DA)

- aggression and, 2476, 2480
- aggressive behavior and, 2480
- in aging brain, 3992–3993
- anatomy, 62
- anxiety disorders and, 1742
- appetite and, 364
- delirium pathophysiology and, 1181
- description of, 4575
- DRD2 receptor, aggression and, 2483
- functions of, 62
- haloperidol and, 1686
- mood disorders and, 1607
  - brain pathways, 1714
- neuroimaging studies, 1492–1494, 1493*f*
- neurotransmission by, 141
- PET/PECT studies of, 297, 299–301
- prefrontal cortical circuits and, 431
- PTSD and, 1820
- receptor agonists, **3233–3242**
  - bipolar depression, 3236–3237
  - chemistry, 3234, 3234*f*–3235*f*
  - dosage/administration, 3241, 3241*t*
  - drug interactions, 3240–3241
  - drug studies, 3236–3238
  - EDS, 3238
  - EPS, 3238
  - history, 3233–3234
  - laboratory interferences, 3241
  - MDD, 3236
  - medical illness-induced fatigue and depression, 3237–3238
  - NMS, 3238
  - organ/system effects of, 3238–3239
  - performance enhancement, 3238
  - pharmacological actions, 3234–3236
  - precautions/adverse reactions, 3239–3240, 3240*t*
  - RLS, 3238
  - therapeutic indications, 3239

- receptor antagonists
  - dependence, 3077
  - dosage/administration, 3079–3080, 3080*t*
  - overdosage, 3077
  - side effects, 3075–3077
  - tolerance, 3077
  - withdrawal, 3077
- release of, 300–301
- sanguine humor and, 1613
- schizophrenia
  - neuroimaging studies, 1492–1495, 1493*f*
  - neuropathology, 1460
- sex differences, 383–384
- signaling, 383–384
- in social phobias, 703
- stimulant-use disorder, 1283
- storage of, 300–301
- stress and, 1713
- substance use disorders and, 375
- synthesis of, 62–63, 63*f*, 300–301
- VTA and, 62

- Dopamine-beta-hydroxylase (DBH), 63
- Dopamine blockers
  - BDD, 1792
  - excoriation disorder, 1796, 1797
  - tic disorders, 1789
  - TTM, 1795
- Dopamine dysregulation syndrome, 546
- Dopamine hypothesis
  - description of, 4575
  - FGAs and, 3065–3066, 3066*t*, 3067*f*
- Dopamine plasma membrane transporter, 63–64
- Dopamine-receptor agonists, 4575
- Dopamine-receptor antagonists (DRAs), 4575
- Dopamine receptors
  - D1 availability, 299–300
  - D2/D3 availability, 300

  - function of, 66–67
  - subtypes, 66–67
- Dopamine transporter (DAT), 25*f*, 63–64
- Dopaminergic innervation, brain, 8
- Dopaminergic neurons, PTSD and, 1818
- Dopaminergic system
  - ADHD and, 3590
  - anxiety disorders and, 1743*t*, 1745
  - buspirone and, 3009
  - dissociative disorders and, 1889
  - fear and, 1758
  - FGAs and, 3066–3067
  - gambling disorder and, 1807
  - hoarding disorder and, 1793
  - OCD and, 1787
  - tic disorders and, 3639–3640, 3653–3654
- Dopaminergic therapy, 547
- Dora (Freud case), 2640, 2652
- Dorahy, Martin, 1871
- Doral. *See* Quazepam
- Doriden. *See* Glutethimide
- Dormen, Lesley, 2833–2834
- Dorpat, Theo, 2647
- Dorsal attention network, 418
- Dorsal premotor cortex, 1893
- Dorsal raphe nucleus, 346
- Dorsal root ganglion disorders, 628
- Dorsal root sensory ganglia, 40
- Dorsolateral prefrontal cortex (DLPFC). *See also* Prefrontal cortex
  - aggression and, 2477–2478
  - cognitive function and, 417–418
  - depersonalization/derealization, 1891
  - depersonalization/derealization and, 1891
  - emotional regulation and, 1819
  - function of, 8, 328
  - hypnotizability and, 2712
  - memory suppression and, 1893

- mood disorders and, 1718*f*
- projections from, 449
- psychotherapy and, 2813–2814

- Dosage/administration
  - dosages, 573*t*
  - dosing/administration, 3165
  - drug-drug interactions
    - in cancer, 2261
  - drug interactions, 2199, 3162, 3164
  - EEG effects, 252
  - effects of, 172
  - for excoriation disorders, 2352
  - history, 2917
  - illness anxiety disorder, 1838
  - inflammation and, 201
  - intimate partner violence, 2502
  - major depressive disorder, 3678, 3679, 3692, 3693, 3790, 3796, 3899
  - obsessive-compulsive disorder treatment, child/adolescent, 3701, 3792
  - olanzapine combined with, 3155, 3324
  - pharmacokinetics, 3164, 3164*t*
  - pharmacological action, 3159
  - pregnancy and, 3167
  - selective mutism, 3723
  - selective mutism in children, 1784
  - separation anxiety, 3710
  - sexual dysfunction and, 1779*t*
  - side effects, 573*t*, 1679*t*
  - social anxiety disorder, 1776, 3710, 3717, 3746
  - suicidality and, 3166–3167
  - targets for, 69
  - therapeutic indications, 3155
  - tic disorders, 3646, 3660
- Dosimetry, internal, 293
- Dosing instructions, abbreviations for, 2294
- Dostinex. *See* Cabergoline
- Double-blind studies, 4540
  - description of, 4575–4576

- epidemiology of, 815
- Double-blind technique, 1863, 4575
- Double depression, 1642*f*, 1646, 4576. *See also* Dysthymia
- Double dissociators, 1903
- Double helix, 149
- Double personality, 4576
- Double technique in psychodrama, 2856
- Doublecortin *(DCX)* mutations, 52
- Do–Undo–Redo psychodrama, 2859
- Down, J. Langdon, 3492
- Down syndrome (DS)
  - description of, 4576
  - general appearance in, 455–456
  - genome-wide dysregulation in, 445–446, 445*f*
  - intellectual disability and, 3507
  - learning styles and, 3517
  - stereotypic movement disorders and, 3668, 3682
  - terminology for, 3492
- Downstream, definition of, 150
- Downward-arrow technique, 1773
- Downward social comparison, 2847
- Doxazosin (Cardura), 1927, 1927*t*
- Doxepin (Sinequan)
  - chemistry, 3254, 3254*f*
  - dosage/administration, 3265, 3265*t*
  - drug/food interactions, 3264
  - drug studies, 3259
  - for insomnia, in aging adult, 4054
  - organ/system effects of, 3260
  - pain syndromes, 1841
  - pharmacological actions, 3257, 3258
  - therapeutic indications, 2982, 3262
- Doxepine (Deptran)
  - for atopic dermatitis, 2342
  - dosages, 573*t*
  - EEG effects, 251–252
  - for excoriation disorders, 2352

- histamine receptor blocking by, 75
- side effects, 573*t*
- for urticaria, 2346

- Doxycycline (Doryx), 591
- DRA. *See* Dopamine-receptor antagonists (DRAs)
- Draijer, Nel, 1871
- Draw-A-Person (DAP) test, 4576
- Drawing
  - child interview technique, 3383
  - in cognitive development, 707
- Drawing tests, description of, 4576
- DRD4
  - aggression and, 2483–2484
  - mood disorder, ADHD and, 1625
- Dread, definition of, 4576
- DREADDs (designer receptors exclusively activated by designer drugs), 130
- Dream anxiety disorder. *See* Nightmare disorders
- Dream psychology, 4545
- Dreaming
  - consciousness and, 351–352
  - theories of, 352
  - waking consciousness and, 350–353
- Dream(s)
  - description of, 4576
  - induced, hypnosis and, 2712
  - interpretation, 2656
    - case study, 2656
  - interpretation of, **856–860**
    - condensation, 859
    - content analysis
      - day residues, 857
      - nocturnal sensory stimuli, 857
      - repressed infantile drives, 857–858
      - significance of, 858
    - displacement, 858–859
    - primary and secondary process, 860
    - projection, 859

        - regression, 859–860
        - representability, 858
        - secondary revision, 859
        - symbolism, 858
    - in patient interview, 967
- Dreamy state, 4576
- Dreikurs, Rudolph, 2736
- Dressing apraxia, 980
- Drift diffusion models (DDM), 435
- Drives
    - definition of, 4576
    - neuroimaging studies of, 2814–2815
- Driving
    - cannabis use and, 1309
    - competency of older adult, 4238
    - everyday functioning, 4026
    - safety, with dementia, 1200
- Driving forces, ion channel, 132
- Dronabinol (Marinol), 2271
- Droperidol (Inapsine)
    - black box warning, 2919
    - chemistry, 3063*f*, 3064
    - for delirium, 1187
    - for terminally ill children, 2309
- Drosha, function of, 51
- Drowsiness, description of, 4576
- Drug abuse. *See also* Substance abuse; Substance use disorders
    - delirium and, 1182
    - description of, 4576
    - and infections, 1011, 1017
    - neurocentrism, 784
- Drug Abuse Warning Network (DAWN), 1282, 1319, 1319*t*, 1380, 1381*f*
- Drug Addiction Treatment Act (DATA), 3136
- Drug and Alcohol Services Information System (DASIS), 1282
- Drug dependence, description of, 4576
- Drug development, 2930–2931, 2932–2935
    - biomarkers, 2926, 2929

- clinical trials, 2926–2928
  - diagnosis and patient selection, 2927
  - dose, 2927
  - effect size, 2926
  - expectation bias, 2927
  - health care professional interaction, 2927
  - medication nonadherence, 2927
  - positive, negative, and failed, 2928
  - regression to the mean, 2927
  - repeated measures, 2928
  - statistical considerations, 2927
  - study design, 2926
  - subjective outcome measures, 2927
- collaborative networks, 2928
- definitions, 2928, 2932
- FDA, 2930–2935
- future directions, 2928–2930
- globalization, 2929
- IND, 2931, 2933
- nonclinical studies, 2932–2935
- outcome measures, 2927, 2930
- patient centricity, 2929
- personalized medicine, 2929
- personalized technology, 2929
- phases of, 2932–2935
- “real-world” data, 2929–2930
- regulation of, 2930–2935
  - Best Pharmaceuticals for Children Act, 2931
  - failures, 2930, 2931
  - Food, Drug, and Cosmetic Act of 1938, 2930
  - Food and Drug Administration Amendments Act of 2007, 2931
  - Food and Drug Administration Modernization Act of 1997, 2931
  - IND, 2931
  - Kefauver-Harris Amendments, 2931
  - labeling, 2930
  - Pediatric Research Equity Act, 2931
  - Prescription Drug User Fee Act, 2931

        - Public Health Security and Bioterrorism Preparedness Act of 2002, 2931
        - Pure Food and Drug Act of 1906, 2930
    - relevant outcome measures, 2930
    - scientific challenges, 2926–2927
- Drug efficacy study implementation (DESI), 2931
- Drug Enforcement Agency (DEA), 4576
    - opioid maintenance treatment administration, 3133
    - psychopharmacology regulation, 2918
- Drug Enforcement Agency, U.S., 4555
- Drug holiday, description of, 4576
- Drug-induced movement, 1148
- Drug-induced parkinsonism. *See* Pseudoparkinsonism
- Drug interactions. *See also specific drugs*
    - cytochrome system and, 2199*t*
    - description of, 4576
- Drug levels. *See also* Pharmacokinetics
    - description of, 4576
- Drug safety, 2935
- Drug self-administration, 323
- Drug therapy. *See also* Medication(s)
    - description of, 4576
    - formulary policy limitations, 4352
    - other therapy with, 4292
- Drug tolerance. *See* Tolerance
- Drug(s). *See also* Medication(s); *specific drugs*
    - alcohol as, 1266
    - alcohol use disorder treatment, 1278
    - altered state induced by, 1899
    - antianxiety (*See* Antianxiety drugs)
    - antidepressant (*See* Antidepressant(s))
    - anxiolytic (*See* Anxiolytic drugs)
    - approval process, **2930–2935**
    - cause of psychiatric symptoms, 4043
    - definition, 2932
    - depersonalization/derealization, 1943
    - development of, **2926–2936**
    - and disease interactions, 4043, 4044*t*, 4045

- insurance coverage, 4218
- interactions of, 2920, 2924–2925 (*See also* Pharmacodynamics; Pharmacokinetics; *specific drugs*)
  - of antipsychotic drugs, 4136
  - cholinesterase inhibitors, 4143–4145
  - of memantine, 4145
- metabolism, gender differences, 4249
- metabolism pharmacogenomics, 3455
- near-death experiences and, 2589
- off-label use of, 2969, 2972, 3004
- potency, 2921
- psychotropic (*See* Psychopharmacology; Psychotropic drugs)
- regulation, 2930–2935
- schedule I
  - hallucinogens, 2887, 2889
- sedative-hypnotic (*See* Sedative-hypnotic drugs)
- spirituality and, 2581–2582

- Drugs of abuse
  - ASD and, 57–58
  - EEG effects, 253–254
- DSM. *See* *Diagnostic and Statistical Manual of Mental Disorders* (DSM)
- DST. *See* Dexamethasone suppression test (DST)
- DTI. *See* Diffusion tensor imaging (DTI)
- DTs. *See* Delirium tremens (DTs)
- *Du haschich et de l'alienation mentale* (Moreau de Tours), 4535
- Dual-sex therapy, 4576
- Dual therapy. *See* Co-therapy
- DuBrul, Sascha, 2835
- Duchenne muscular dystrophy (DMD), 634
- Ductus deferens ligation, 2481
- Due-process guarantees, 4389
- Dugas, Ludovic, 1870
- Duloxetine (Cymbalta)
  - adverse effects of
    - dermatological, 2356*t*
  - antidepressant efficacy, 3151, 3151*f*
  - anxiety disorders, 1746, 1779–1780
    - GAD, 1783

- BD and generalized anxiety disorder, 1698
- chemistry, 3147*f*
- child/adolescent treatment, 3714
- depression, 1679
- depression treatment, 4124
- for diabetic neuropathy, 2236, 2236*t*
- dosage, 573*t*
- for fibromyalgia, 635
- generalized anxiety disorder, 3714
- for geriatric depression, 4073
- history, 3150
- for neuropathies, 629
- pain management using
  - in dying patients, 2292
- in pain syndromes, 1841, 2395
- pharmacology, 3150, 3151*f*
- sexual dysfunction and, 1779*t*
- side effects, 573*t*, 1679*t*
- targets for, 69
- tolerability, 3151
- tolerance of, 2261

- Duluth model, intimate partner violence treatment, 2502
- Dummy, definition of, 4576
- Dumont, Henri, 2566
- Dunedin Multidisciplinary Health and Development study, 1452
- Durable powers of attorney, 2305
- Dürer, Albrecht, 1631, 1632*f*
- Durham rule, 4576
- *Dusky v. United States,* 4438
- Duties of physicians, 4436
- Duty to warn, 2636
- Dyad, description of, 4576
- Dyadic conflict patients, 2648
- Dyadic deficit patients, 2648
- Dyadic session, description of, 4576
- Dyesthetic vulvodynia, 2349
- Dying, physician aid in, 2313

- Dying patients. *See also* Death/dying
  - decision-making issues, 2302–2303
  - insomnia in, 2295–2296
  - the last vigil, 2303
  - with multiple kinds of pain, 2292
  - physical symptoms in, 2291–2297
  - physician promises to, 2291*t*
  - psychological accompaniment of, 2301–2302
  - psychological syndromes and, 2297–2301
  - quality of life and, 2297–2298
  - spiritual concerns of, 2302–2303
  - subjective experiences of, 2301–2302
  - therapy in, 2295*t*
  - treatment of pain in, 2292
- DynaCirc. *See* Isradipine
- Dynamic holograms
  - STAPP, 2799, 2799*f*
  - STDP of narcissistic and other self-disorders, 2809
- Dynamic psychiatry
  - after World War II, 4292
  - description of, 4576
  - mental health stages of, 4291
- Dynamic psychotherapy, 2202
- Dynamic reasoning, 4576
- Dynamic systems theories, 3360
- Dynamic theory, 2649
- Dynamics. *See* Psychodynamics
- Dynorphin (DYN)
  - opioid receptor activation by, 169
  - stimulant-use disorder and, 1284
- Dyrbye, Liselotte, 2599
- Dysarthrias
  - description of, 4576
  - differential diagnoses, 469*t*
- *Dysbindin (DTNB1)* gene, 317
- Dyscalculia, 4576
- Dysdiadochokinesis, 4558

- Dysdiadokinesis, 540
- Dysgenusia, 4576
- Dysgraphia, 4576
- Dyskinesia. *See also* Tardive dyskinesia
  - description of, 4576
  - medication-induced, 2938
- Dyslalia, 4576
- Dyslexia
  - description of, 4576
  - neuroimaging, 3477
- Dyslipidemia, 1542–1543
- Dysmenorrhea, 2724
- Dysmetria, 540, 4576
- Dysmnesia, 4576
- Dysmorphophobia, 1790, 1829, 4576. *See also* Body dysmorphic disorder
- Dysnosognosia, 1146
- Dyspareunia, 1969–1970, 1971, 4576
- Dysphagia
  - description of, 4576
  - in myasthenia gravis, 633
- Dysphasia, 4576
- Dysphonia, 3366, 4576
- Dysphoria
  - benzodiazepine, 2999
  - description of, 4577
  - major depressive disorder and, 1643
  - medications for stabilization, 2171
- Dysphoric mania
  - about, 1640
  - bipolar I disorder and, 1650
- Dysphoric mood. *See* Dysphoria
- Dyspnea, 2229
- Dyspraxia, 3539
- Dysprosody, 4577
- Dysregulation of choice, 1262
- Dysrhythmias, biofeedback and, 2559
- Dyssocial behavior, 4577

- Dyssomnias, 4577
- Dysthymia (dysthymic disorder). *See also* Depression
  - clinical picture, 1646
  - cognitive theory and, 1665
  - course, 1646
  - definition, 1601, 1631
  - description of, 4577
  - differential diagnoses, 1644–1645
  - DSM-5 criteria for, 1645*t*
  - gambling and, 1804
  - IPT adaptation for, 1666
  - psychoanalytic formulations of, 1662–1663
  - psychopharmacology, 3791
  - recurrent brief depressive disorder and, 1647
  - social adjustment, 1645–1646
  - STDP and, 2801
  - sub-bipolar, 1642
  - suicide and, 2612
  - in transplant populations, 2365–2366
  - treatment
    - MAOIs, 3120–3121
  - variants, 1646
- Dystonia(s)
  - acute
    - anticholinergics and, 2959
    - FGAs and, 2940–2941, 3072
  - categories, 560
  - characteristics, 560
  - description of, 458, 4577
  - haloperidol and, 2375–2376
  - neuroleptic-induced acute, 2940–2941
    - clinical presentation, 2940
    - implications, 2940–2941
    - pathophysiology, 2940
    - treatment options, 2940–2941
  - prognosis, 560
  - tardive, 2942–2943

    - anticholinergics and, 2959
    - clinical presentation, 2942
    - pathophysiology, 2943
    - treatment options/implications, 2943
- Dystonic movements, 1148
- Dystonic tics, 3635, 3649
- Dystrophin gene mutations, 634
- Dystrophinopathies, 634

E

- E/1 balance, 144–145
- 4E-BP, phosphorylation of, 126
- E-EPA. *See* Ethyl-eicosapentaenoic acid (E-EPA)
- Early adolescence, 3366
- Early Developmental Stages of Psychopathology Study (EDSP), 1453, 1733–1734, 1734*t*
- Early ejaculation, 1969, 4577
- Early Global Burden of Disease, 1786
- Early intensive behavioral intervention protocol, 2500
- Early-onset Alzheimer disease, 1203
- Early-onset bipolar disorder, **3686–3694, 3700–3708**
  - assessment, 3688, 3702
  - brain structure and function, 3688, 3702
  - comorbidities, 3689–3690, 3703–3704
  - course, 3691–3692, 3705–3706
  - defined, 3686, 3700
  - diagnosis, 3688–3689, 3702–3703
  - differential diagnoses, 3690–3691, 3704–3705
  - epidemiology, 3687, 3701
  - etiology, 3687–3688, 3701–3702
  - further directions, 3693, 3707
  - history, 3686–3687, 3700–3701
  - overview, 3686, 3700
  - treatment, 3692–3693, 3706–3707
- Early-onset schizophrenia (EOS), **3725–3733**
  - comparative nosology, 3726
  - course and prognosis, 3731–3732, 3732*f*
  - defined, 3725, 3726*t*, 3727*t*
  - diagnosis and clinical features, 3728
  - differential diagnoses, 3730–3731, 3730*t*
  - epidemiology, 3726–3727
  - etiology, 3727–3728
  - future directions, 3733
  - history, 3725–3726
  - pathology and laboratory examination, 3729, 3729*f*
  - treatment, 3732–3733

    - FGAs, 3071
- Early Periodic Screening Diagnosis and Treatment, 4304
- Earnings of older people, 3951*t*
- Earth mother archetype, 913
- Eating. *See also* Diet and nutrition
  - behavioral neuroscience of, 355
  - brain areas, 363*f*
  - brief-access, 359
  - hedonics and, 360
  - meals as a functional unit, 355
  - sham, 359
  - subjective experience of, 355
  - variety, 360
- Eating disorders (ED), **2065–2082**. *See also* Anorexia nervosa (AN); Bulimia nervosa
  - animal models of, 322
  - BDD and, 1791
  - behavioral studies, 367–368
  - child/adolescent
    - group therapy and, 3768
  - classification of, 1174
  - clozapine for, 2169
  - complications, 2074*t*
  - description of, 4577
  - diabetes type 1 and, 2247
  - diagnosis of, 2071–2073
  - differential diagnoses, 2075–2077
  - emergency psychiatric evaluation for, 3446, 3446*t*
  - epidemiology of, 2069–2070, 2069*t*
  - etiology of, 2070–2071
  - exposure therapy, 2691
  - and genetic selection, 778
  - genetics and, 2529*t*
  - historical literature on, 2067
  - hypnosis and, 2731
  - intellectual disabilities and, 3512
  - laboratory exams, 2073–2075, 2074*t*
  - medical treatment of, 2220–2221

- neuroimaging, child/adolescent, 3475
- neurosteroids and, 118
- nosology, 2067–2069
- obesity and, 2213–2214
- OCD *versus,* 1788
- pathology, 2073–2075
- psychological treatment of, 2220
- serotinin and, 368
- serum electrolyte evaluation, 1010
- sex differences in, 368–369
- sleep-related, 2102
- STDP in, 2811
- suicide and, 2612, 2613
- temperament dimensions of, 3487
- treatment, 2078–2082
  - anticonvulsants, 2976
  - bupropion, 3006
  - MAOIs, 3121–3122
  - naloxone, 3142
  - naltrexone, 3006, 3142–3143
- unspecified, 2066–2067
- world-wide, 4520

- Eating in absence of hunger (EAH), 359
- Eaton, James, 4402
- ECA. *See* Epidemiologic Catchment Area
- *Echinacea* spp., 2554
- *Echium amoenum,* 2554
- Echo-Planar Imaging (EPI), 273
- Echolalia, 457*f*, 4577
- Echopraxia, 464, 1419, 3582, 4577
- Eco-corporeality of individuals, 2441
- Ecological momentary assessment, 2326
- Ecology, description of, 4577
- Economies of scale, 4313–4314
- Ecopsychiatry, 4577
- "Ecstasy." *See* 3,4-methylenedioxy-*N*-methylamphetamine (MDMA, "ecstasy")
- ECT. *See* Electroconvulsive therapy (ECT)

- Ectomorhic, definition of, 4577
- Ectopic atrial premature beats, 2190
- ED. *See* Erectile dysfunction
- Edelman, Gerald, 2840
- Edinburgh High-Risk Study (EHRS), 1470
- Edinburgh Postnatal Depression Scale (EPDS), 2517
- EDRT. *See* Elder Death Review Team
- EDS. *See* Excessive daytime sleepiness
- Educable, definition of, 4577
- *Educating Physicians: A Call for Reform of Medical School and Residency* (2010)
  - shortcomings noted in, 4323
- Education. *See also* Academic or educational problems; Psychiatric education; Psychoeducation
  - on Alzheimer disease risk factor, 1204
  - autism spectrum disorder and, 3585
  - as child psychiatric evaluation consideration, 3379
  - about dementia, 1204
  - on depressive disorders, child/adolescent, 3680, 3694
  - dissociative identity disorder treatment, 1922–1923
  - for intellectually disabled persons, 3492, 3517
  - language disorders and, 3549
  - of older people, 3950*t*
  - in patient interview, 957
  - of professionals in geriatrics, 3954*t*
  - of psychiatrists, 4551
  - psychopharmacology, 2921
  - for specific learning disorder, 3535
  - on tic disorder interventions, 3644, 3658
- Education for All Handicapped Children Act (1975), 3493
- Educational levels, outcomes and, 4383
- Educational psychology, 4577
- Edwards Personal Preference Schedule (EPPS), 998*t*
- EEGs. *See* Electroencephalograms (EEGs)
- Efavirenz
  - HIV/AIDS, 3856
  - side effects of, 571
- Effectiveness research, 2704, 4347
  - in statistics, 845–846

- Effects of, 1266–1268
- Effexor. *See* Venlafaxine
- Efficacy research, 2704
  - of antipsychotic medications, 4351–4352
  - in research, 4347
  - in statistics, 845
- Ego. *See also* Epigenetic development
  - autonomy of, 871–872
  - boundaries
    - description of, 4577
    - psychosis and, 2869
  - defects, specific, 2164
  - defensive functions, 868–872, 870*t*–871*t*, 4577
  - description of, 4577
  - group psychotherapy and, 2738
  - human conflict and, 760*f*
  - model, 4577
  - mood and, 1660–1661
  - object relationships, 869
  - psychology of, **865–873**
  - reality testing, 869
  - relation to reality, 868–869
  - stages of development, 2168*f*
  - Strachey's translation of Freud on, 2645
  - strength, 2138, 4577
  - synthetic functions, 871
  - weakness
    - nonspecific manifestations of, 2164
- Ego analysis, description of, 4577
- *Ego and the Id* (Freud), 860, 866, 872
- *Ego and the Mechanisms of Defense* (A. Freud), 869
- *The Ego and the Mechanisms of Defense* (A. Freud), 869
- Ego-coping skill, 4577
- Ego-dystonic. *See* Ego-alien
- Ego-dystonic homosexuality, 4577
- Ego-dystonic thoughts, 1121
- Ego ideal, 4577

- Ego identity, 3365
- Ego state therapy, hypnosis and, 2717
- Ego-syntonic symptoms
  - in cancer patients, 2265–2266
  - description of, 4577
- Ego-syntonic thoughts, 1121
- Ego–alien, description of, 4577
- Egocentric, definition of, 4577
- Egocentrism, 709
- Egomania, 4577
- EHealth. *See also* Computers; Smartphone apps
  - anxiety and depression, 2691
- Ehrensaft, Diane, 2031
- Ehrhardt, Anke, 1958
- *Ehrlichia* infections, 593
- Eicosanoids, **115–117**. *See also specific eicosanoids*
  - chemistry of, 116
  - effects on specific organs/systems, 116
  - therapeutic indications, 116–117
- Eicosapentaenoic acid (EPA)
  - biological treatment with, 2551
  - bipolar depression, 1692*t*, 1693
- Eidetic images, 4577
- EIF4E release, 126–127
- Eissler, Kurt
  - as psychoanalyst, 2644, 2664
  - Sigmund Freud Archives and, 2645
- Ejaculation
  - definition of, 1954
  - delayed, 1968–1969
  - early, 1969
  - retrograde, 1971
- Ejaculation incompetence. *See* Inhibited male orgasm
- Ekman, Paul, 765
- Elaboration, 4577
- Elation
  - about, 1632

- description of, 4577
- Elavil. *See* Amitriptyline
- Elbert, Thomas, 1890
- Eldepryl. *See* Selegiline
- Elder abuse. *See also* Elderly persons
  - definition, 2539
  - dementia and, 1200
  - family role, 4200
  - legal and ethical analysis, 4250–4255
    - court system, 4253
    - definitions, 4250–4251
    - excuses and justifications, 4254
    - financial abuse, 4253–4254, 4255
    - indicators of abuse, 4252–4253
    - multidisciplinary teams, 4252
    - nosology, 4251
    - physical abuse, 4254–4255
    - suspicions of abuse, 4254
    - vulnerability factors, 4251–4252
  - psychiatric emergency due to, 2634
- Elder Death Review Team (EDRT), 4252
- Elderly persons. *See also* Geriatric psychiatry
  - access to care, 2282
  - amantadine in, 2962
  - benzodiazepines in, 2998, 2999
  - with cancer
    - depression in, 2282–2285
    - fatigue in, 2284
    - suicide risk, 2283
  - cancer risk
    - survival and, 2285–2287
  - delirium in, 1180, 1183
  - depression in, 1657
    - neuroimaging and, 2814
  - disaster or war and, 2569
  - FGAs in, 3065, 3080
  - gambling in, 1810

- homosexuality and, 2003–2004
- lithium in, 3094–3095
- neurocognitive impairment, 1506
- psycho-oncology, 2282–2285
- psychopharmacology in, 2924
- schizophrenia in, 1506
- SGAs in, 3080
- sleep in, 342
- SSRIs and, 1779
- suicide risk, 2615, 2621
- thyrotoxicosis in, 680
- trazodone in, 3248, 3251, 3252
- tricyclic and tetracyclic antidepressants in, 3256
- vilazodone in, 3347

- Elective mutism, 4577. *See also* Selective mutism
- Electra complex, 4577
- Electrical (faradic) aversion, 2694
- Electrical stimulation, 2398. *See also* Deep brain stimulation (DBS)
- Electroacupuncture, 2561–2562
- Electrocardiograms (ECGs)
  - cardiac arrhythmias on, 2190
  - role in treatment selection, 1020–1021
- Electrochemical syndromes, 649*t*
- Electroconvulsive therapy (ECT), **3280–3298**. *See also* Neurostimulation therapies
  - acute mania, 1689*t*, 1690
  - adverse effects, 3295–3298
    - cardiovascular system, 3295–3296
    - CNS, 3296
    - dental, 3296
    - musculoskeletal, 3296–3297
    - neurocognitive, 3297–3298, 3297*t*
    - respiratory system, 3296
  - barbiturates and, 2986, 2987
  - bipolar depression, 1692*t*, 1693, 3090
  - bipolar I disorder, 1696*t*, 1697
  - bipolar II depression, 1698
  - for cancer patients, 2262

- in cardiac patients, 2199–2200
- contraindications, 3295–3298
- definition, 3299*t*
- delirium treatment, 1188
- delivery of, devices for, 3282–3283
- for depression, 1681
  - efficacy of, 33
  - in PD, 548
- depression with suicidality, 2616
- description of, 4577
- diagnostic indications, 3287–3288
  - neurological, 3287–3288
  - psychiatric, 3287
- dissociative identity disorder, 1928–1929
- electrode placement, 3284, 3289–3290, 3290*f*, 4157–4158, 4157*f*
- factors in, 3282
- future directions, 3298
- for geriatric depression, 4074, **4151–4164**
  - adverse effects, 4160*t*
    - neurocognitive, 4161–4162
    - neurophysiologic, 4160–4161
    - systemic, 4160
  - anesthesia, 4156–4157
  - clinical indications, 4152–4153, 4153*t*
  - comorbid medical diagnoses, 4158
  - electrode placement, 4157–4158, 4157*f*
  - history and background, 4152
  - informed consent, 4155–4156, 4156*t*
  - neurocognitive assessment, 4154–4155, 4155*t*
  - safety and efficacy
    - acute factors, 4158–4159
    - continuation and maintenance, 4159–4160
  - specific recommendations, 4158
- history of, 3280–3281, 3281*t*, 3299
- indications, 3285–3288, 3285*t*
- lithium interactions with, 3090, 3294
- major depression, 3285–3286

- mania, 3286
- for MDD in brain tumor patients, 499
- mechanisms of action, 3283–3285
  - biochemical, 3283–3284
  - genetic and epigenetic, 3284–3285
  - neurophysiology, 3284
  - neuroplasticity, 3285
- medical clearance for, 1020
- medication interactions with, 3294–3295
- for mood disorders, 1251, 3285–3286
- neurobiology and, 1709
- OCD and, 1790
- patient management, 3288–3294
  - anesthesia and, 2986, 2987, 3291–3293
  - consent and, 3288–3289
  - continuation treatment, 3293–3294
  - electrode placement, 3284, 3289–3290, 3290*f*
  - pretreatment evaluation and, 3288
  - seizure threshold/dosing and, 3290–3291
  - stimulus parameters, 3289–3294
  - stimulus waveform and, 3290
  - treatment course, 3293
  - treatment protocols and, 3291–3293
- physiology, 3295–3298
- psychiatric malpractice, 4437
- schizophrenia, 1553, 3286–3287
- seizures, duration of, 3292
- seizures-induced by, 3282–3283
- suicide risk in elderly and, 2615
- thyroid hormones and, 3244
- use of, 146
- VNS *v.,* 3312, 3312*t*

- Electrodermal skin testing (EDST), 2545, 2546
- Electrodes, EEG
  - nasopharyngeal, 221
  - placement, 221–224
  - sphenoidal, 221

- Electroencephalograms (EEGs)
  - activating procedures
    - photic stimulation, 241–242, 242*f*
  - on admissions, 245*f*
  - after TBI, 525
  - alterations from drugs, 251–254
  - ambulatory, 467
  - amplifiers, 224–225
  - apparatus, 219–226
    - electrodes, 219–224, 220*f*, 221*f*
    - montage selection, 221–224, 222*f*
  - in applied clinical practice, 236–251, 466–467
  - artifacts, 228–231
    - ambient noise, 231*f*
    - electrode movement, 231*f*
    - electromyographic, 230*f*
    - electrophysiological, 230–231
    - ocular correction, 229*f*
    - physiological, 228–230
    - sweat drift, 230*f*
  - of autism spectrum disorder, 3582, 3583*f*
  - biofeedback using, 2546, 2559
    - case study, 2560
  - characteristic features, 226–228
  - closed head injuries, 245–246
  - clozapine-induced, 235*f*
  - ctenoids, 238–239
  - for delirium, 4091
  - delirium diagnosis by, 1184
  - of dementia, 1197
  - description of, 4577
  - in diagnosis, 1019–1020
  - digital recording, 225–226
  - dissociative identity disorder and, 1894
  - drug activation studies, 243–244
  - drug suppression tests, 243–244
  - in endocrine disorders, 246*f*

- in epilepsy, 507–508, 508*f*
- error-related negativity, 257–258
- feature detection, 231–233
- of foster care children, 3819
- four-step approach, 249*t*
- hyperventilation, 244
- hypnosis, 2712
- of intellectual disability, 3501
- interpretation of, 218
- interpreter reliability in, 218–219
- in metabolic disorders, 246*f*
- mismatch negativity, 256–257
- neurophysiology study, 216
- normal intrinsic frequencies
  - alpha, 227
  - beta, 227
  - delta, 226
  - gamma, 227–228
  - theta, 227
- normal *versus* abnormal, 236–237
- phantom waves, 241
- photomyoclonic responses, 242
- polysomnography, 243
- principles of, 217–219
- quantification, 231–233, 233–235
  - applied clinical practice, 235–236
  - maps, 234*f*
- recording, 219–226
- related methodologies, 219–226
- research paradigms, 254–258
- rhythmic mid-temporal discharges, 240*f*
- RMTDs, 239
- of schizophrenia, early onset, 3729
- sharp waves, 237–241, 238*f*
- sleep, mood disorders and, 1717
- sleep on, 342–343
- small sharp spikes, 239–241, 241*f*

- spike discharges, 225*f*
- spike patterns, 237–241
- spike-wave phantoms, 242*f*
- structural lesions, 244
- variant normal, 237
- wireless devices, 4278

- Electroencephalography, quantitative (QEEG), **1026–1047**
  - brain state vectors, 1031–1032
  - clinical applications, 1034–1040
    - Alzheimer type dementia, 1036–1037, 1037*f*
    - anesthesia monitor, 1040
    - attention-deficit/hyperactivity disorder, 1036
    - chronic pain and thalamocortical dysrhythmia, 1037–1038, 1038*f*
    - key-lock principle, 1041
    - neurofeedback, 1041
    - pharmaco-EEG, 1040–1041
    - psychopathology: biomarkers of diagnoses, 1034–1035
    - psychosis, 1038–1040, 1039*f*
    - subtyping, 1035, 1035*f*
  - from EEG to QEEG, 1027–1028, 1027*f*
  - homeostatic system generating and regulating EEG, 1030–1031, 1033*f*
  - neurometric QEEG
    - analysis features, 1028–1030, 1029*f*
    - homeostatic regulation, 1028
    - reporting results, 1030
  - synchronous deviations from ground state, 1042–1047
    - binding, resonance, and consciousness, 1044–1047, 1045*f*, 1046*f*
    - coincidence detection constructs fractionated elements, 1044, 1046*f*
    - construction of consciousness, 1043–1044, 1043*f*
    - discontinuity of consciousness, 1042
    - exogenous-endogenous interaction, 1044
    - neuron-interneuron interaction, 1044
    - parallel multiplexed information processing channels, 1042
  - tomographic brain images, 1032–1034, 1034–1039*f*
- Electrogenic, definition of, 132
- Electrolytes, serum, 1010
- Electromagnetic fields, 2546

- Electromyography (EMG)
  - description of, 4577
  - pain studies and, 2389
- Electronic health records (EHR). *See* Electronic medical records (EMRs)
- Electronic medical records (EMRs), 4320
  - guideline use, 975
  - mental health consultation for medical students and, 2607
  - in patient interview, 962
- Electrophysiology
  - applied, **216–259**
  - cellular principles of, 131–135
  - history of, 217
- Electroshock therapy. *See* Electroconvulsive therapy (ECT)
- Electrostimulation, description of, 4578
- *Elements of Psychophysics* (Fechner), 4536
- Eletriptan for migraine, 625
- Elevated mood, 4578
- Elimination disorders, **3650–3661, 3664–3675**
  - description of, 4578
  - encopresis, **3658–3661, 3672–3675**
    - clinical features, 3659, 3673
    - comparative nosology, 3658, 3672
    - course, 3660, 3674
    - defined, 3658, 3672
    - diagnosis, 3659, 3673
    - differential diagnoses, 3660, 3674
    - epidemiology, 3658, 3672
    - etiology, 3658–3659, 3672–3673
    - history, 3658, 3672
    - laboratory examination, 3660, 3674
    - pathology, 3659–3660, 3673–3674
    - prognosis, 3660, 3674
    - treatment, 3660–3661, 3674–3675
  - enuresis, **3650–3658, 3664–3672**
    - clinical features, 3654, 3668
    - comparative nosology, 3650–3651, 3651*t*, 3664–3665, 3665*t*
    - course, 3655, 3669

        - defined, 3650, 3664
        - diagnosis, 3654, 3668
        - differential diagnoses, 3654–3655, 3668–3669
        - epidemiology, 3651, 3665
        - etiology, 3651–3654, 3665–3668
        - historical background, 3650, 3664
        - laboratory examination, 3654, 3668
        - pathology, 3654, 3668
        - prognosis, 3655, 3669
        - treatment, 3655–3657, 3669–3671
    - with intellectual disability, 3512
- Elixir Sulfanilamide Tragedy of 1937, 2930
- Elkins, David, 2583
- Elliotson, John, 2706
- Ellis, Albert, 3751
    - clinical psychology and, 2696
    - *Overcoming Procrastination,* 4554
    - on religion, 2585
    - on spirituality, 2584
- Ellis, Havelock, 1953, 1953*f*, 1986
- Ellson, Douglas, 414
- Elopement, description of, 4578
- Elucidation, domains of, 461
- Email, psychotherapy and, 2672–2673
- Emancipated minors, 4578
- Emanuel E., case of, 4318, 4324
- Embryology
    - nervous system, 41–42
    - neurulation, 40–41
- Emde, Robert, 1900
- Emergency department facilities
    - as boarding areas, 4305, 4325, 4330
    - dissociative disorder presentations, 1888
    - hypnosis in, 2724–2725
        - case study, 2725
    - psychiatric treatment in, 4344
- Emergency involuntary hospitalization, 4435

- Emergency psychiatry. *See also* Psychiatric emergency care
  - child psychiatric evaluations for, **3391, 3394***t*, **3436–3447**
    - for aggression, 3443–3444, 3444*t*
    - agitation management during, 3440–3441, 3441*t*
    - for autism spectrum disorder, 3445
    - for child abuse or neglect, 3445–3446
    - components of, 3438–3440, 3439*t*
    - contextual factor assessment during, 3442
    - developmental factor assessment, 3441–3442
    - for eating disorders, 3446, 3446*t*
    - epidemiology for, 3437
    - etiology determination, 3438, 3438*t*
    - for intellectual disability, 3445
    - legal considerations for, 3447
    - overview, 3436–3437
    - principles of, 3438–3442
    - priorities for, 3437–3438
    - for psychosis, 3444–3445
    - for psychosocial crisis, 3446–3447
    - for risky behaviors and substance use, 3444
    - safety concerns, 3437–3438
    - for suicidality and self-harm, 3442–3443
- Emergency room (ER). *See* Emergency department facilities
- Emetic aversion, behavior therapy, 2694
- EMG. *See* Electromyography (EMG)
- Emicerfont, 1745
- Emminghaus, Herman, 4536
- Emmy von N. (Freud case), 2639
- Emotional abuse, amnesia and, 1886
- Emotional deprivation, 4578
- Emotional deregulation after TBI, 534*f*
- Emotional detachment, 2173
- Emotional disorders, 4578
- Emotional disturbance, 4578
- Emotional dysregulation, 3684, 3698
- Emotional forecasting, 2846
- Emotional freedom technique (EFT), 2561, 2563

- Emotional illnesses, 4578
- Emotional insight, 4578
- Emotional instability, 2171, 2172*t*
- Emotional intimacy
  - adulthood, 3942–3943, 3943*f*
  - instability in, 2140
- Emotional lability, definition of, 4578
- Emotional numbing, 1892
- Emotional personality (EP), 1903
- Emotional processing
  - acceptance-based techniques and, 1773
  - anxiety disorders and, 1769
  - corrective information and conditions interfering with, 1770
  - failures of, 1770–1771
  - indicators, 1769–1770
  - theory of, 1769–1771
- Emotional reactions, cognitive model, 2760, 2760*f*
- Emotional supports, 4578
- Emotionally focused therapy (EFT), 2759, 2760
- Emotion(s). *See also* Psychological sciences
  - ADHD and, 3591
  - adulthood, intimacy and, 3942–3943, 3943*f*
  - amygdala and memory of, 1757
  - asthma and, 2224–2225
  - in autistic spectrum disorder, 697–698, 698f
  - biological basis of, 765–766
  - and consciousness, 692
  - definition of, 4578
  - development in childhood, 3352–3353*t*, 3361
  - development of, 3378
  - dynamics of transitions, 2168*f*
  - dysregulation
    - BPD and, 2786*f*, 2787
  - end of life burdens, 2304
  - etiology of, 1711
  - and intelligence, 711–713
  - and mental health, 762, 765–766

- mood disorders and, 464–465
- natural selection of, 775–777, 776*f*
- neural circuit for, 33*f*
- neuroimaging activation of, 1760
- in patient interview, 957
- physician response to, 2291
- positive and negative, 763, 765–766
  - empirical findings, 2846–2847
- processing of, 1711–1712
  - lateralization of, 449
  - neurocircuitry, 1756
- PTSD, childhood/adolescent, 3705
- reciprocity of, 2278–2279
- regulation of
  - amygdala connectivity and, 424–425
  - cognitive therapy, 2772
  - PTSD, 1817–1819
  - skills, 2055, 2788
- responses to stress, 2318
- schizophrenia, fMRI studies, 1477*f*, 1482–1485, 1484*f*
- with specific learning disorder, 3529
- SSRI blunting of, 3158
- temperament and, 3480–3481

- EMP+ (nutrient formula), 2552
- Empathic understanding, 2818
- Empathy
  - deficient, 2050
  - description of, 4578
  - disorder of, 1851
  - failure of, 465
  - MBT, 2899
  - mutative change and, 2664
  - psychoanalysis and, 2647
- Empirically Supported Treatments and Promising Practices, 4296
- Employee health care costs, 4293
- Employment
  - adulthood, 3942–3944

- supported, 1534–1535, 1536*f*, 1552, 3732–3733
- Empty nest syndrome, 4578
- Emtricitabine/tenofovir (TDC-TDF, Truvada), 2008
- *Emx2* gene, 49
- EMX2 proteins, 49
- Enactments, psychodynamic psychotherapy and, 2678
- Encapsulated delusion, 4578
- Encephalitis, description of, 4578
- Encephalitis lethargica
  - characteristics, 551
  - insomnia and, 346
- Encephalization, human, 2471
- Encephalography. *See* Quantitative electroencephalography (QEEG)
- Encephalopathies
  - autoimmune, 1015
  - defined, 1180
  - description of, 4578
  - laboratory evaluation, 1015
  - paraneoplastic, 1015
- Encephalopathy of prematurity, 640
- Enclosure Laws, 4290
- ENCODE project, 152–153, 153*f*
- Encoding
  - brain processing, 981
  - peritraumatic dissociative amnesia and, 1892–1893
- Encopresis, **3658–3661, 3672–3675**
  - clinical features, 3659, 3673
  - comparative nosology, 3658, 3672
  - course, 3660, 3674
  - defined, 3658, 3672
  - description of, 4578
  - diagnosis, 3659, 3673
  - differential diagnoses, 3660, 3674
  - epidemiology, 3658, 3672
  - etiology, 3658–3659, 3672–3673
  - history, 3658, 3672
  - laboratory examination, 3660, 3674

- pathology, 3659–3660, 3673–3674
- prognosis, 3660, 3674
- treatment, 3660–3661, 3674–3675

- Encounter groups, 4578
- Encouragement in patient interview, 957
- End-of-life care, **2288–2315**. *See also* Death/dying
  - decision-making, 2304
  - decisions in pediatrics, 2311–2312
  - for dementia patients, 1200
  - ethical issues, 2312–2313
  - extraordinary treatment, 2312
  - family role, 4199
  - ordinary treatment, 2312
  - staff dynamics in, 2304, 2311
- Endemic, definition of, 4578
- Endfeet, formation of, 5
- Endocannabinoids, **110–115**
  - addiction and, 114
  - anxiety disorders and, 1743*t*, 1750
  - biosynthesis of, 110–112
  - brain system of, 110
  - inactivation of, 112
  - pain and, 115
  - in psychosis, 114–115
  - research into, 110*t*
  - retrograde neurotransmission of, 113*f*
  - retrograde transmission regulated by, 112–113
- Endocrine disorders
  - description of, 4578
  - psychiatric patients
    - brain tumors and, 494
  - and psychiatric symptoms, 4048
- Endocrine system
  - antipsychotics, effects of, 3795
  - buspirone effects on, 3012
  - characteristics of, 166
  - DA receptor agonist effects on, 3239

- DA receptor antagonists and, 3076
- evaluations
  - adrenal disease, 1013
  - anabolic steroid use, 1013
  - antidiuretic hormone, 1013–1014
  - human chorionic gonadotropin, 1014
  - pancreatic function, 1014–1015
  - parathyroid hormone, 1014
- fear or stress response and, 1757
- laboratory evaluation, 1013–1015
- mechanisms
  - cancer outcomes and, 2286
- melatonin effects and, 3104
- opioid receptor antagonist effects on, 3143
- responses to stress, 2318
- and sexuality in aging adult, 4256–4258

- Endocrinology, focus of, 2239
- Endocrinopathies, psychiatric symptoms with, 2239–2240
- Endogamy, description of, 4578
- Endogenous affective disorders, 1604
- Endogenous depression. *See also* Major depressive disorder
  - description of, 4578
  - neurobiology and, 1709
- Endogenous retroviruses, 154
- Endometrial carcinoma, tamoxifen for, 1687
- Endomorphic, description of, 4578
- Endophenotypes, 311*f*
  - animal models, 311
  - description of, 156
  - working memory function as, 430
- Endoreactive moods, 1633, 1650
- Endorphins. *See also* Enkephalins
  - description of, 4578
  - opioid receptor activation by, 169
  - pruritus and, 2342
- Endothelial nitric oxide synthase (eNOS), 105
- Endotoxin. *See* Lipopolysaccharide (LPS)

- Endozepines, 1746
- Energetic–active temperament, 1651
- Energy
  - CAM assessments of, 2544–2546
  - forms not validated by Western science
    - CAM therapies based on, 2561–2564
  - forms validated by Western science, 2546
    - CAM therapies based on, 2558–2561
  - homeostasis
    - adiposity and, 362
  - and information processing, 688–689
- Energy healers, 2545–2546. *See also* Complementary, alternative, and integrative medicine
- Energy psychology
  - as energy therapy, 2546, 2561, 2563
- Energy storage, obesity and, 2212–2213
- Enfuvirtide, HIV/AIDS and, 3856
- Engagement, happiness and, 2847
- Engel, George, 1179
- Engram, description of, 4578
- Enkephalins
  - function of, 31*f*
  - opioid receptor activation by, 169
- Ensemble coding hypothesis, 405
- Entactogens, 1316
- Enteric nervous system, 40
- *Entire Works of Dr. Thomas Sydenham* (Sydenham), 4534
- Entomorphic, description of, 4578
- Entorhinal cortex
  - function of, 33
  - human brain, 8
  - neurons, 6*f*
  - stained sections, 5*f*
- Enuresis, **3650–3658, 3664–3672**
  - clinical features, 3654, 3668
  - comparative nosology, 3650–3651, 3651*t*, 3664–3665, 3665*t*
  - course, 3655, 3669
  - defined, 3650, 3664

- description of, 4578
- diagnosis, 3654, 3668
- differential diagnoses, 3654–3655, 3668–3669
- epidemiology, 3651, 3665
- etiology, 3651–3654, 3665–3668
- genetics, 3652–3653, 3666–3667
- historical background, 3650, 3664
- laboratory examination, 3654, 3668
- nocturnal, 3937–3938
- pathology, 3654, 3668
- prognosis, 3655, 3669
- psychopharmacology, 3794–3795
- treatment, 3655–3657, 3669–3671

- Environment. *See also* Gene-environment interactions
  - adolescence and, 3371–3374
  - adoption and, 3826
  - and Alzheimer disease, 4085
  - in cancer
    - psychologically-mediated, 2285
  - change in, cognitive therapy and, 2772
  - child psychiatric evaluation assessment of, 3412
  - in delirium, 4092
  - and dementias in long-term care, 4227
  - depressive disorders, child/adolescent, 3676, 3690
  - epigenetics and, 387, 387*t*
  - and everyday functioning, 4027, 4029
  - factors in ADHD, 3589–3590, 3599–3600
  - factors in psychiatric disease, 54–58
  - generalized anxiety disorder, child/adolescent, 3712
  - and human disease, 770
  - invalidating, BPD and, 2786, 2786*f*
  - language disorders, 3546
  - mental health and, 764, 2527–2528
  - neurodevelopment and, 51–52
  - obesity and, 2213
  - and personality traits, 937
  - positive psychiatry, 4283

- presenting illness, genetic counseling and, 2525
- psychiatric disorders and, 2529–2530
- schizophrenia, 1447–1448, 1448*t*, 1449–1451
- in schizophrenia, 4099
- of sensory food aversions, 3625, 3639
- social anxiety disorder, child/adolescent, 3715
- for specific learning disability, 3526–3527
- of stereotypic movement disorders, 3668, 3682
- stuttering, 3561
- suicide risk, child/adolescent, 3683–3684, 3697–3698
- temperament affected by, 3483
- tic disorders, 3640, 3654
- tobacco smoke in, 1347
- Tourette disorder, 3638, 3652

- Environment-evocative genetic traits, 2483–2484
- Environmental context, outcomes and, 4383
- Environmental signals, gene expression and, 48
- Environmental treatments, 3815
- Environmental triggers, psychiatric disorders and, 2530
- Envy management skills, 2441
- Enzyme-linked immunosorbent assay (ELISA), 85
- Enzymes, liver function, 1010
- EOS. *See* Early-onset schizophrenia (EOS)
- Eosinophil differentiation, 179
- EPA. *See* Omega-3 fatty acids-eicosapentaenoic acid (EPA)
- Ependymomas, hypothalamic, 677
- Eph A3 receptors, 53
- Ephebophiles, 2017
- Epidemiologic Catchment Area (ECA) Study, 783, 1286, 1786. *See also* Geriatric psychiatry
  - mental disorders in U.S., 800
- Epidemiological Catchment Area study, U.S.
  - violent individuals, 2490
- Epidemiology, 3538
  - assessment instruments, 821
    - Alcohol Use Disorder and Associated Disabilities Interview Schedule, 821
    - Composite International Diagnostic Interview (CIDI), 820
      - University of Michigan version, 820–821

- World Mental Health Surveys version, 821
- Diagnostic Interview Schedule, 820
- diagnostic interviews for children and adolescents, 821
- for disability and impairment, 822
- background, 810–811
- case identification
  - *Diagnostic and Statistical Manual of Mental Disorders,* 818–819
  - disability, 819–820, 819*t*
  - *International Classification of Diseases (ICD),* 819–820
- cross references, 828
- description of, 4578
- descriptive, 1614–1615, 1731
- disaster psychiatry, 2568–2570
- experimental studies, 814–815
- major surveys
  - Epidemiologic Catchment Area Program (ECA), 822–823
  - methods of study, 826
  - National Comorbidity Survey Replication Adolescent Supplement, 826–827, 827*t*
  - National Epidemiological Survey on Alcohol and Related Conditions, 825–826, 826*t*
  - NCS and associated studies, 823–824, 823*t*, 824*t*
  - WHO Global Burden of Disease Study, 827–828, 828*t*
  - World Mental Health Initiative, 824–825, 825*t*
- measurement
  - reliability, 816–817
  - validity, 817, 817*t*
- measures of disease frequency
  - comorbidity, 812
  - incidence, 812
  - prevalence, 811–812
  - relationship between prevalence and incidence, 812
- measures of effect and association
  - odds ratio, 813–814
  - relative risk, 813, 813*t*
  - risk, protection, and causation, 812–813
- psychiatric, 1730–1731
- of psychiatric disorders, 4302

- Epidural administration, 2293*t*
- Epigenesis, 4578
- Epigenetic codes
  - function of, 386, 386*f*
  - genetic codes and, 386*t*
- Epigenetic development of ego
  - generativity *vs.* stagnation, 881
  - identity *vs.* confusion, 881
  - industry *vs.* inferiority, 881
  - initiative *vs.* guilt, 881
  - integrity *vs.* despair, 881
  - intimacy *vs.* isolation, 881
  - psychosocial crises, 880–882
  - trust *vs.* mistrust, 880
  - zones and modes, 880
- Epigenetic regulation
  - addiction and, 381–382, 382*f*
  - gene expression and, 50
- Epigenetics
  - addiction to cocaine, 397–398
  - anxiety disorders, 1740
  - brain development and, 444–447
  - definition of, 2, 1711
  - developmental, 166
  - ECT epigenetic mechanisms, 3284–3285
  - further research in, 3356–3357
  - gene regulation, 160–161, 1462–1463
  - machinery of, 388–394
  - olfactory fear memory, 394–395
  - preclinical work, 394–398
  - promise of, 387*t*
  - psychiatry and, **386–388**
  - PTSD, 1820
  - schizophrenia, 446–447, 1462–1463
  - suicide signatures, 400–401
- Epigenome, description of, 160–161
- Epilepsy. *See also* Autosomal dominant frontal lobe epilepsy (ADNFLE); Myoclonic epilepsy with ragged red fibers (MERRF)

- aggressive behavior and, 516, 2496–2497
- antidepressants in, 520*t*
- autism and, 654
- autism spectrum disorder and, 3583
- behavioral changes in, 517
- behavioral disorders in, 511*t*, 514
  - nonepileptic, 514–517
- brain tumors and, 493
- calretinin neurons in, 17
- cannabis use for, 1311
- children with, 648–651
  - adjustment issues in, 651
  - anxiety and, 649
  - depression and, 649
- clinical evaluation, 454–455
- cortical damage and, 451–452
- course/prognosis, 518
- definition of, 507, 648
- depersonalization and, 1891
- description of, 4578
- diagnosis, 511–514
- differential diagnoses, 470–471, 1428
- dissociation, violence and, 2491
- dissociative amnesia *versus,* 1936
- electroconvulsive therapy, 4153
- epidemiology, 509–510
  - depression, 509–510
  - psychopathology, 509
  - psychosis, 509
- episodic aggression, 470–471
- etiology of, 510–511
- genes associated with, 648
- historic literature on, 508
- ictal features, 511–512
- interictal features, 513–514
  - predisposing factors, 513*t*
- interictal psychosis

    - in children, 649–650
  - neurodiagnostic tests, 518
  - neuropathology, 518
  - neuropsychiatric aspects of, **507–522**
  - nosology, 508–509
  - organization of
    - ages of onset, 649*t*
  - pediatric, aggression in, 650
  - postictal psychosis, 649–650
  - pre-ictal features, 512
  - psychiatric comorbidities, 515–517
  - psychiatric disturbances, 510*t*
  - psychiatric patients, 493
  - psychic auras, 510*t*
  - psychotic states in, 470–471
  - seizures
    - classification of, 508*t*
    - nonepileptic seizures *versus,* 515*t*
  - surgical management, 520–521
  - temporal lobe post-traumatic, 525
  - transient depersonalization, 1887
  - treatment, 518–521
    - anticonvulsants (*See* Anticonvulsants)
    - antidepressants, 519–520
    - antiepileptics, 518–519
    - antipsychotics, 519
    - ECT, 3288
    - psychotropic medications, 519*t*
    - surgical, 651
    - TNS, 3313
- Epileptic dementia, 4578
- Epileptic equivalent, 4578. *See also* Complex partial seizures (CPS)
- Epileptic personality. *See* Explosive personality
- Epinephrine
  - aggression and, 2480
  - anatomy of, 62
  - description of, 4578

- functions of, 62
- neurotransmission by, 141
- synthesis, 62–63, 63*f*
- trauma response and, 2569

- Episodic, definition of, 4578
- Episodic memory, 2139, 2140
- Epistasis, 1622
- Epistemology, 4578
- Epitranscriptome, 161
- EPS. *See* Extrapyramidal side effects
- Epstein–Barr virus (EBV)
  - overview, 595
  - stress and, 2321–2322
  - suppression of, 190
- Epston, David, 2838
- Epworth Sleepiness Scale (ESS), 343
- Equanil. *See* Meprobamate
- Equilibration, in developing intelligence, 705, 715
- Equilibrium potential, 131
- ErbB2 receptor, 52
- ErbB4 receptor, 52
- Erdelyi, Matthew, 2731–2732
- Erectile dysfunction (ED). *See also* Male erectile disorder
  - diagnostic criteria, 3337, 3337*t*
  - etiology, 3337–3338
  - future directions, 3341
  - psychotropic-induced, 3340–3341
  - risk factors, 3338
  - treatment, **3337–3342**
    - alprostadil, 3340
    - apomorphine, 3340
    - lifestyle changes, 3338
    - mechanical and surgical devices, 3340
    - melanocortin receptor agonists, 3340
    - phosphodiesterase-5 inhibitors, 3338–3339, 3339*t*
    - recommendations, 3341
    - testosterone, 3339–3340

    - yohimbine, 3340
  - treatment of, 107
- Erection, loss of, 1980–1981
- Eremorphia, 4578
- Ergasia, 4578
- Ergoloid mesylates, 4146
- Ergotamine for migraine, 625
- Erickson, Milton
  - on hypnosis, 2716
  - paradoxical therapy, 2816, 2817
  - strategic family therapy, 2753
- Erikson, Erik, 4578
  - on adolescence, 3365–3366, 3367
  - *Childhood and Society,* 4539
  - ego development
    - psychosocial crises, 880–882, 931
    - zones and modes, 880
  - *Gandhi's Truth,* 898, 904
  - *Identity: Youth and Crisis,* 900, 901
  - *Insight and Responsibility,* 900
  - life and work
    - as biographer of middle age, 904–905
    - as biographer of youth, 903–904
    - child psychoanalyst, 898–899
    - clinical work, 905–906
    - emigration to America, 899
  - life cycle stages
    - autonomy *vs.* shame and doubt, 880–881, 901
    - generativity *vs.* stagnation, 881, 902
    - identity *vs.* role confusion, 881, 901–902
    - industry *vs.* inferiority, 881, 901
    - initiative *vs.* guilt, 881, 901
    - integrity *vs.* despair, 881–882, 902–903
    - intimacy *vs.* isolation, 881, 902
    - trust *vs.* mistrust, 880, 900–901
  - mental health criteria, 758
  - midlife transition, 900

- *The Nature of Clinical Evidence,* 905, 906
- psychoanalytic anthropology, 899–900
- *Vital Involvement in Old Age,* 904
- *Young Man Luther,* 898, 903

- Erle, Joan, 2674
- Ernst, Richard, 259
- Erogenous zones, 4578
- Eros. *See* Sexual drive
- Erotic, definition of, 4578
- Erotic stimuli, sex differences in, 1959–1961
- Erotic transference
  - case study, 2656
  - definition, 2650
- Eroticized children, 1962
- Erotomania, description of, 4578
- ERP. *See* Event-related potentials (ERPs)
- Error-related negativity (ERN), 257–258
- Erythrophobia, 4578
- Escape learning, 4578
- Escitalopram (Lexapro or Cipralex)
  - anxiety disorder, 4129
  - anxiety disorders, 1779–1780
  - BD and generalized anxiety disorder, 1698
  - breast milk excretion of, 3168
  - chemistry, 3161
  - combined pharmacotherapy and, 3325
  - dosing/administration, 573*t*, 1979, 3166
  - drug-drug interactions
    - in cancer, 2261
  - drug interactions, 3164, 3165
  - GAD, 1783
  - major depressive disorder, 3790
  - pharmacokinetics, 3164, 3164*t*, 3165
  - pharmacological actions, 3161–3162
  - for PTSD prevention of, 2379
  - sexual dysfunction and, 1779*t*, 1979
  - side effects, 573*t*, 1679*t*

- specific phobias, 1784
- Esdaile, James, 2706
- Eskalith. *See* Lithium
- ESP. *See* Extrasensory perception (ESP)
- Esquirol, Jean Etienne Dominique, 4535
- Essential oils, biological treatment with, 2554–2555
- Essential tremor (ET), 559–560
- Essentialism, views of, 1984
- Estazolam (ProSom), 2992
- *Estelle v. Gamble*, 4450
- Esthetic sensibility, lack of, 2140
- Estimation. *See* Statistics
- Estradiol
  - postpartum withdrawal of, 2517
  - testosterone conversion to, 2505
- Estradiol (Alora, E_2)
  - actions, 3331–3332, 3331*f*
  - combination pharmacotherapy and, 3325
  - function of, 171
  - perimenopausal depression, 3333–3335
  - postpartum depression, 3335–3336
  - production of, 2243
  - reasons for, 3332
  - risk factors or contraindications, 3334–3335, 3335*t*
- Estriol (E_3), function of, 171
- Estrogen receptors, 2522
- Estrogen replacement therapy (ERT)
  - combination pharmacotherapy and, 3324–3325
  - sense of well-being and, 172
  - tacrine and, 3042
- Estrogens
  - Alzheimer disease and, 1204–1205
  - biological treatment with, 2555
  - effects of, 2034
  - fluctuations, 383, 383*t*
  - function of, **171–172**
  - immunological effects, 186*t*

- ineffectiveness for Alzheimer disease, 4148–4149
- replacement therapy, 2522
- synthesis of, 1958
- therapies for sex offenders, 2503

- Estrone (E_1)
  - function of, 171
  - production of, 2243
- Eszopiclone (Lunesta)
  - chemistry, 3001
  - dosage/administration, 3002
    - for insomnia, 2086
  - drug interactions, 3002
  - insomnia, 3002
  - pharmacological actions, 3001
  - pharmacology, 1377–1378
  - precautions/adverse reactions, 3002
  - therapeutic indications, 3001–3002
- Ethanol. *See* Alcohol
- Ethchlorvynol (Placidyl)
  - dependency diagnosis, 1385
  - precautions/adverse effects, 2987, 2990
  - as sedative-hypnotic, 1385
- Ethical issues
  - aversion therapy, 2703
  - behavior therapy, 2703
  - bioethics of geriatric care, 4241–4242
  - CAM treatments, 2543
  - for caregivers, 4211
  - child/adolescent, **3868–3880**
    - administration and, 3877
    - advocacy, 3876–3877
    - assent, 3872
    - boundaries, 3873–3874
    - codes of medical ethics, 3869
    - confidentiality, 3738, 3872–3873
    - consent, 3872
    - countertransference, 3872

- culture, 3871–3872
- dilemmas and reasoning in, 3870
- fidelity and mutual trust, 3873–3874
- forensics, 3870
- funding and provision of care, 3876
- genomics, 3874–3875
- history, 3869–3870
- information technologies, 3875–3876
- multiple agency, 3871
- overview, 3868–3869
- pharmaceutical industry, 3877–3878
- principles and, 3870
- prodrome treatment, 3875
- provisions of care, 3872–3874
- religion, 3871–3872
- research, 3878–3879
- teaching, 3879
- violation reporting, 3879–3880

- client-centered therapy, 2819
- codes of medical ethics, 3869
- cognitive-behavioral therapy, 3759
- cognitive therapy, 2774
- core principles
  - beneficence and nonmaleficence, 4241
  - justice, 4241
  - respect for persons/autonomy, 4241
- DBT, 2794
- decision making, 4242–4244
- dignity therapy, 2827
- disaster psychiatry and, 2575
- elder abuse, 4254
- and electronic technology, 4280
- existential psychotherapies, 2822
- exposure therapy, 2703
- factitious disorder and, 1864
- family role
  - cultural factors, 4199–4200

    - decision making, 4199
    - elder abuse, neglect, and exploitation, 4200
    - end-of-life care, 4199
  - genetic counseling, 2536–2537
  - genetic testing in children, 2282
  - group psychotherapy, 3770–3771
  - in group therapy, 4206
  - guided imagery, 2823–2824
  - hallucinogen-assisted psychotherapy, 2891
  - history, 4240–4241
  - in individual psychotherapy, 4178
  - informed consent, 4242–4244
  - IPT, 2783
  - MBT, 2902
  - in mental health care, 4530–4531
  - nonprofessional therapists, 2829
  - paradoxical therapy, 2817–2818
  - psychoanalytic psychotherapy
    - boundary violations, 2669–2670
    - confidentiality, 2670–2672
    - sexual transgressions, 2669–2670
    - value-free fallacy, 2669
  - psychodrama, 2859
  - psychotherapy under managed care, 2829
  - school consultations, 3883
  - strategic therapy, 2817–2818
  - TA, 2821
  - tripartite connection, 2307
  - violence and, 2429
- Ethics. *See* Forensic psychiatry, ethics; Medical ethics
- Ethnic issues
  - and everyday functioning, 4027
  - family functions, 4194
  - health care costs for older adults, 4212–4213
  - and HIV infection, 4270
  - as psychosocial factors, 4169–4171
- Ethnicity. *See* Race/ethnicity

- Ethnocentrism, 4578–4579
- Ethnocultural populations. *See also* Minorities
  - attitudes toward, 4381
  - service disparities, 4377
- Ethnograms, description of, 2472
- Ethnology, definition of, 4579
- Ethnopsychiatry, schizophrenia, 1427
- Ethnopsychology, approaches to, 2447
- Ethosuximide (Zarontin)
  - in borderline personality disorder, 519
  - for children, 650
  - mechanism of action, 507
- Ethyl-eicosapentaenoic acid (E-EPA)
  - aggression in BPD, 2502
  - dietary supplements of, 115–116
- Etiology, 3538
- Etomidate
  - anesthesia induced by, 2986, 2987, 2990
  - ECT anesthesia with, 2986, 2987, 3292
  - precautions/adverse effects, 2988
- Euchromatin, 149
- Eudaimonia, 2848
- Euergasia, 4579
- EUFEST. *See* European First Episode Schizophrenia Trial
- Eugen Bleuler, 1407
- Eumetabolic hypoestrogenism, 2521
- Eunuch, 2023
- Euphoria, 464, 4579
- Euphorohallucinogens, 4579
- European Academy of Childhood Disability (EACD), 3539
- European First Episode Schizophrenia Trial (EUFEST), 3069, 3070
- European Medicines Agency (EMA), 1789
- Eurostat on nonsanctioned aggression, 2474
- Eustress, 2336
- *Euthanasia: Direct Medical Killing* (Lifton), 4538
- Euthanasia, wishes for, 2313*t*
- Euthymia

- ADHD during, 1699
- description of, 4579
- Euthymic, definition of, 4579
- Euthyroid sick syndrome, 2074
- Evaluations. *See* Child psychiatric evaluations
- Evasion, description of, 4579
- Event recording in behavior therapy, 2687
- Event-related euphoria, 4579
- Event-related potentials (ERPs), 254, 255*f*
  - in ADHD, 258*f*
  - N100, 256
  - P50, 256
  - P300, 256
- Everolimus, 2371
- Everyday functioning. *See also* Activities of daily living (ADLs); Geriatric psychiatry
  - and HIV infections, 4271–4272
- Eviation, definition of, 4574
- Evidence-based medicine
  - ACA requirements, 2185
  - as best practices, 4347
  - cost containment with, 4307
  - definition of, 4340
  - ethnocultural disparities in, 4382
  - geriatric patients, 4303
  - in jails and prisons, 4458
  - manualized interventions
    - for adult psychiatric rehabilitation, 4296, *4297t*
    - for child psychiatric rehabilitation, 4296, 4298, *4299f*
  - models for, 4321
  - psychiatry, 4295–4300
  - from reform efforts, 4288
  - rehabilitation toolkits, 4363–4364
  - schizophrenia, 1551–1554, 4350
  - on school-based programs, 3891–3892
  - in 19th century psychiatry, 4328
- Evoked potentials (EPs), 254–256
  - clinical applications, 254–256

- in clinical practice, 466–467
- description of, 4579
- role in diagnosis, 1020

- Evolution, violence and, 2418–2419
- Evolutionary biology, **769–780**
  - behavioral ecology, 774–775
    - attachment, 775
    - reproductive effort, 774–775
  - core principles, 771–774
    - evolution, 772
    - genotypes and phenotypes, 772
    - levels of selection, 773
    - misunderstandings, 773–774
    - natural selection, 771–772
    - selection of systems, 772
    - sexual selection, 772–773
  - cross references, 779
  - evolutionary medicine
    - constraints, 770
    - defenses, 771
    - environment and body mismatch, 770
    - pathogens, 770
    - reproduction, 771
  - future directions, 779
  - individual behavior, 769–770
  - kin selection, 770
  - proximate and evolutionary questions, 769*t*
  - relationships, 777
  - social behavior, 770
  - vulnerability to specific disorders, 777–779
    - emotion disorders, 777–778
    - personality disorders, 779
    - schizophrenia, 778
    - sociopathy, 778–779
    - substance abuse, 778
- Ewalt, Jack, 4308
- Ewer, Timothy C., 2729

- Exaltation, description of, 4579
- *Examination of Men's Wits* (Huarte de San Juan), 4534
- Examination of psychiatric patient, **944–1047**
  - communication of opinions
    - treatment recommendations, 956
- Exceptionalism, mainstreaming *vs.,* 4294, *4295f*
- Excessive daytime sleepiness (EDS)
  - treatment
    - armodafinil, 3238
    - DA receptor agonists, 3238
    - modafinil, 3231–3232, 3238
- Excessive fragmentary myoclonus, 2105
- Excitation-inhibition coordination, 408*f*
- Excitatory, definition of, 141
- Excited, definition of, 4579
- Excitement phase, 4579
- Excitotoxicity
  - description of, 4579
  - neuronal degeneration and, 80
  - NMDA and, 107
- Exclusion standards, forensic, 2731
- Excoriation, 2349
- Excoriation (skin-picking) disorder
  - BDD and, 1792
  - clinical features of, 2350–2351
  - comparative nosology, 1796
  - course and prognosis, 1797
  - definition of, 1796, 4579
  - diagnosis and clinical features, 1796
  - differential diagnoses, 1797
  - DSM-5 on, 1785
  - epidemiology of, 1796, 2350–2351
  - etiology, 1796
  - history, 1796
  - introduction, 1795–1796
  - OCD *versus,* 1788
  - pathology and laboratory examination, 1796–1797

- pharmacotherapy, 1797
- psychiatric comorbidities, 2351–2352
- psychotherapy, 1797
- treatment, 1797, 2352
- TTM and, 1795

- Execution, criminal competence, 4438
- Executive allocation policy, 696
- Executive control, PTSD and, 1817
- Executive ego function, 4579
- Executive function
  - ADHD and, 3590, 3601
  - in aging adults, 4006
  - assessment of, 3434
  - autism spectrum disorder and, 3581
  - conduct disorder, 3615
  - deficits
    - after TBI, 529
    - FASD and, 638
  - dementia's effect on, 1194, 1195
  - dysfunction of, 463–464
  - hemispheric dominance, 981–982
  - language and, 2431
  - neuropsychological testing, 992
  - overview, 3430–3431*f*
  - social (pragmatic) communication disorder and, 3553–3554
  - specific learning disorder and, 3525
  - stimulant use effects on, 1284
- Exelon. *See* Rivastigmine
- Exercise
  - anxiety disorders, 1782
  - appetite and, 366–367
  - dietary preferences and, 2547
  - for fatigue, 2284
  - somatic symptom disorders, 1836, 1844
  - as somatic therapy, 2556
  - SUD treatments, 384*t*
- Exhaustion state, 4579

- Exhibitionism, description of, 4579
- Exhibitionistic disorders, 2013, 2016–2017
- Existential analysis, 2821
- Existential depression, 1646. *See also* Dysthymia
- Existential factors, group psychotherapy and, 3761
- Existential psychotherapies, 2821–2822
  - description of, 4579
  - ethical issues, research, and evaluation, 2822
  - history, 2821
  - indications, 2822
  - techniques, 2822
  - theoretical issues, 2821–2822
- Existentialism, description of, 4579
- Exogamy, 4579
- Exon arrays, 439
- Exonic splicing enhancers (ESEs), 158
- Exonic splicing silencers (ESSs), 158
- Exons, function of, 150
- Exophthalmos, 456
- Exorcism, 2581, 4579
- Expansive mood, 4579
- Expectancy models, hypnosis and, 2711
- Experienced gender, 2023
- Experiencing, definition of, 4579
- Experiential avoidance, 2699
- Experiential–humanistic family and couples therapy, 2751*t*, 2753–2754
- Experimental biases, 4579
- Experimental designs, 4579
- Experimental groups, 4579
- Experimental neurosis, 4579
- Experimental treatments
  - definition and examples, 2704–2705
- Expert witnesses
  - description of, 4579
  - role conflicts, 4437–4438
  - testimony, 4437
- Explanatory mode, culture concepts of distress, 2446

- Explanatory Model Framework, 2461, 2463–2464
- Explanatory Model Interview Catalogue, 2461
  - Short, 2461
- Explanatory Model Questions, 2459
- Explicit mentalizing, implicit mentalizing *versus,* 2895
- Exploding head syndrome, 2102
- Exploratory hypnotherapy, 2716
- Exploratory psychotherapy, 2664–2666
  - case study, 2665–2666
  - contraindications, 2664–2665
  - goals, 2665
  - indications, 2664–2665
  - limitations, 2665
  - main approach, 2665
- Explosive outbursts, 3693, 3707
- Explosive personality, description of, 4579
- Exposure/behavior change technique, 3755
- Exposure response prevention (ERP)
  - DBT, 2790
  - dissociative identity disorder, 1894
  - OCD, 1789
    - cognitive therapy and, 2772
    - hypnosis and, 2722–2723
    - SSRIs and, 1790
- Exposure therapy. *See also* Virtual reality graded exposure therapy
  - agoraphobia, 1783
  - anxiety disorders and, 1769, 1772–1773
    - efficacy, 1774
  - BDD, 1792
  - case study, 2690
  - distraction and, 1771
  - emotional processing and, 1770
  - ethical issues, 2703
  - hypnosis and, 2721–2722
  - models, 2689
  - OCD, 1788, 1789
  - prolonged, PTSD, 1823, 1824

- social anxiety disorder, 1776
- specific phobias and, 1784

- Exposure with response prevention (ERP), 2691
- Expressed emotion (EE), 1433–1434
- Expressed gender, 2023
- Expressed sequence tags (ESTs), 158
- Expressive aphasia, 4579
- Expressive aprosodia, 457*f*
- Expressive dysphasia, 4579
- Expressive language disorders, 3545. *See also* Language disorders
- Expressive movement therapies, 2556
- Expressive psychotherapy, 2665–2666. *See also* Exploratory psychotherapy
- Expressive therapy, 1930
- Extended-family therapy, 4579. *See also* Social network therapy
- Extended matching question (EMQ), 4419, 4420
- Extenuating circumstances, 3370
- External germinal layer (EGL), 42
- External validity
  - description of, 4579
  - in research, 4347, 4348–4349, 4351, 4353
- Externalization
  - description of, 4579
  - process of, 2050
- Externalizing behaviors, child/adolescent, 3896–3897
- Externalizing symptomatology scales, 3402
- Extinction
  - description of, 4579
  - in exposure therapy, 2689
  - of fear memories, 1818
  - in Pavlovian learning, 721, 722
- Extracellular signal-regulated kinases (ERKs), 125
- Extrafamilial sexual abuse, 3833
- Extrapsychic conflicts, 4579
- Extrapyramidal dysarthria, 457*f*
- Extrapyramidal dyskinesia, 4579
- Extrapyramidal effect, 4579
- Extrapyramidal motor system, 4579

- Extrapyramidal side effects (EPS)
  - amantadine, 3238
  - anticholinergics and, 2959, 3073
  - antipsychotic drugs, 1497, 1526, 1528
  - chronic, FGA, 3074–3075
  - drug-induced, amantadine and, 2963, 2965
  - FGAs, 2936, 3060, 3068, 3070, 3072–3074
  - management of, 3073–3075
  - SSRI, 3158
- Extrapyramidal signs, differential diagnoses, 469*t*
- Extrapyramidal syndrome, 4579
- Extrasensory perception (ESP), 4580
- Extraversion, in Five-Factor Model, 935–936
- Extraverted personalities
  - characteristics, 575–576
  - HIV risk behaviors, 575–577
- Extreme capsule, anatomy of, 13
- Extreme illness exaggeration, 1859
- Extremely low birth weight infants (ELBW), 642–643
- Extremities, physical examination of, 456
- Extremity rest tremor, 458
- Extroversion, description of, 4580
- Extroversion-introversion, 913
- Ey, Henry, 1830
- Eyberg Child Behavior Inventory (ECBI), 3402, 3652, 3666
- Eye contact, 952
- Eye fields, frontal, 26
- Eye movement desensitization and reprocessing (EMDR) therapy
  - depersonalization/derealization disorder, 1945
  - dissociative identity disorder, 1930
  - PTSD, 1823
  - science and, 2703
- Eye movements
  - pursuit, 456
  - saccadic, 456
- Eye roll, hypnosis and, 2712, 2715
- Eye-tracking technology, 3581, 3581*f*

- Eye(s)
  - abnormal findings
    - differential diagnoses, 469*t*
  - anticholinergics and, 2958, 2961
  - DA receptor antagonists and, 3076
  - examination of, 456–457
  - MAOI effects on, 3117
  - nefazodone effects on, 3129
  - opening of
    - apraxia, 456
  - physical examination of, 456
- Eyewitness testimonies, 743
- Eysenck, Hans, 930, 2683
- Eysenck, Sybil, 935
- Eysenck Personality Questionnaire (EPQ), 576, 998*t*
- Eysenk personality inventory, 1899
- Ezogenous, description of, 4579
- Ezriel, Henry, 2737

F

- *Fa'afafine,* 1996
- Fabry disease, 656–657
  - MRI in, 656*f*
- Face transplantation
  - ethical considerations, 2387
- Face validity, 320, 996
- Face–hand test, 460
- Face(s)
  - angiofibromas, 456
  - movements, assessment of, 457
  - recognition of, 692
  - reconstructive surgery, 2382
  - transplants, ethical issues, 2387
- Facilitory paratonia, 458
- Factitious amnesia, 1937
- Factitious dermatitis
  - clinical features, 2354
  - epidemiology, 2354
  - psychopathology, 2354
  - treatment, 2354
- Factitious disorder by proxy, 1846, 1847, 1848. *See also* Factitious disorder
  - comparative nosology, 1848
  - course and prognosis, 1862
  - definition, 1847, 1847*t*
  - description of, 4580
  - diagnosis and clinical features, 1857–1858, 1858*t*
  - differential diagnoses, 1859
  - epidemiology, 1850
  - etiology, 1851–1852
  - somatic symptom and related disorders, 1842
  - treatment and management, 1864–1865, 1864*t*
- Factitious disorder imposed on self
  - course and prognosis, 1861–1862
  - diagnosis and clinical features, 1852, 1857*t*
  - epidemiology, 1848–1849
  - etiology, 1850–1851

- treatment and management, 1862–1864, 1863*t*
- Factitious disorder(s), **1846–1865,** 2354
  - classification, 1827
  - clinical features, 1852, 1853*t*–1856*t*, 1856
  - comparative nosology, 1848–1849
  - course and prognosis, 1860–1862
  - definition, 1846–1847, 1847*t*
  - dementia and, 1198
  - description of, 4580
  - diagnosis, 1852, 1853–1856*t*, 1856
  - differential diagnoses, 1198, 1858–1860
  - epidemiology, 1849–1850
  - etiology, 1850–1852, 1850*f*
  - future directions, 1865
  - history, 1847–1848
  - iatrogenic model of dissociation and, 1897–1898
  - imposed on another, 3839
  - introduction, 1844
  - malingering *versus,* 2411*t*
  - NOS, 1848
  - pathology and laboratory examination, 1858
  - somatic symptom and related disorders and, 1842
  - treatment and management, 1862–1865, 1863*t*
- Factitious dissociative identity disorder, 1916
- Factor analysis
  - description of, 4580
  - of personality tests, 996
  - in personality theories, 934–935
- Factor validity, 996
- Fadiman, Anne, 2459
- Failure to thrive
  - defined, 3622–3623, 3636–3637
  - description of, 4580
  - feeding disorders and, 3629–3630, 3643–3644
- Fairbairn, Ronald
  - dissociative disorders, 1869
  - group psychotherapy and, 2737

- as psychoanalyst, 2659
- Faith. *See also* Religion; Spirituality
  - for cancer patient support, 2254
  - dying patients, 2302–2303
- Fake Bad Scale, 1856
- Faking bad, description of, 2400
- Faking good, description of, 2399
- Falling
  - and anxiety disorder, 4060–4061
  - and psychiatric symptoms, 4050
- Fallon, Brian, 1838
- Falloon, Ian, 2754
- Falret, Jean-Pierre, 4535
- False-belief tasks, 337
- False discovery rate, 835
- False memories, 4580
- False memory syndrome, 1872
- False Memory Syndrome Foundation (FMSF), 1872
- False negatives, description of, 4580
- False nonbeliefs, 1898–1899
- False positives, 4480–4481, 4580
- False suffocation alarm, 1760
- Falsifiable hypotheses, 4580
- Familial aggregation, 3545
- Families. *See also* Family therapies; Functional family therapy (FFT); Patient-family units; Staff-family interfaces
  - acculturation of, 792
  - after patient deaths, 2303
  - of aging adults, 3973
  - of cancer patients
    - developmental stages of, 2279
    - evaluation of, 2279
    - family structures, 2279
    - nontraditional, 2279
    - psychological problems of, 2278–2280
    - styles of interaction, 2279
  - of children with cancer, 2281–2282
  - of choice, 2004

- of critically ill persons
  - stress in, 2330
- cross-cultural barriers, 2306
- decision making, 2306
- and drug-disease interactions, 4045
- emotions of, 2291
- end of life burdens, 2304
- expectations of, 2289–2290
- family environments and
  - ADHD treatment interference by, 3596
  - adversity preventive interventions, 3895–3896
  - child hospitalizations, response to, 3849–3851
  - child illnesses, response to, 3849–3851
  - as child psychiatric evaluation consideration, 3378–3379
  - clinical interview with, 3376–3381
  - education
    - about depressive disorders, child/adolescent, 3680, 3694
  - intellectual disabilities contexts with, 3510
  - specific learning disorder and, 3526–3527
  - as suicide behavior risk factor, 3683–3684, 3697–3698
  - supportive
    - for alcohol use disorders, 1278
    - for schizophrenia, 1553
  - temperament and, 3483, 3488, 3489*f*
  - terrorism impact and, 3916, 3922–3924, 3922*t*
  - therapy for, 3761, 3764, 3785
- the final farewell, 2303
- genetic studies, 4018
- grief groups for, 2303
- high risk
  - genetic counseling and, 2530, 2530*f*, 2531
  - genomic testing and, 2536
- and HIV infection, 4269
- Informant-Cultural Formulation Interviews (I-CFI), 2468–2469
- involvement
  - after release, 4399
  - as facilitating/inhibiting help-seeking, 4383

    - during stabilization, 4368–4369
    - support for, 4350–4351
    - world-wide, 4525
  - in jails and prisons, 4466
  - kin selection, 770
  - the last vigil, 2303
  - of LBGT persons, 2004
  - post-illness PTSD, 2331
  - as psychosocial factor, 4167
  - religious expectations, 2306
  - response to healthy siblings, 2311
  - staff and, 2280
  - support for burn patients, 2382–2383
  - support of
    - in mood disorders, 1701–1702
- Family-based association design (genetic mapping), 3450
- Family-focused therapy (FFT)
  - bipolar depression, 1692
  - bipolar disorders, 1703, 1703*t*, 1706
  - bipolar I disorder, 1696
- Family history, at patient interview, 951, 965
- Family interventions
  - in Alzheimer disease, 4085–4086
  - and therapy with older adults, **4192–4200**
    - assessment, 4194
    - bonds of attachment and caregiving, 4195
    - communication patterns, 4194
    - contraindications, 4199
    - definition, 4192–4193
    - ethical issues, 4199–4200
    - goals of treatment, 4199
    - health care setting, 4195
    - history, 3973, 4193
    - illnesses, 4196–4197
    - indications for therapy, 4198–4199
    - inpatient medical treatment, 4196
    - limitations, 4199

    - long-term care, 4197
    - outpatient treatment, 4195
    - palliative care, 4198
    - personal/social functioning, 4196
    - problem solving, 4194
    - psychoeducation, 4194
    - research and evaluation, 4200
    - strength-vulnerability model, 4193–4194
- Family medical history (FMH)
  - anxiety disorders, 1735
  - bipolar disorder, lithium and, 1623
  - data needed for genetic risk assessment and, 2532
  - genetic risk assessment based on, 2532–2533, 2533*f*
  - OCD, 1788
  - as pedigree, genetic counseling and, 2526*f*
  - suicide, mood disorders and, 1670–1671
- Family neurosis, 4580
- Family studies
  - ADHD, 3588
  - aggression, 2482
  - anxiety disorders, 1737–1738
  - BDD, 1791
  - excoriation disorder, 1796
  - genetics and psychiatric disorders, 2527–2528, 2528*t*
  - of homosexuality, 1995
  - mood disorders, 1619–1620
    - neurobiology and, 1710–1711
  - OCD, 1787, 3697–3698
  - of psychiatric illness, 306
  - schizophrenia, 1437–1439, 1438*f*, 1445–1446, 1454, 1471, 1564
  - sexual addiction, 2044
  - sleep neurophysiology, 1717
  - TTM, 1794
- Family therapies, **2748–2760, 3771–3787**. *See also* Families; Functional family therapy (FFT)
  - for ADHD, 3778
  - with adolescents, 3815
  - for anorexia, 3783–3785

- assessment, 2756, 3774–3777
- attendance rules, 2757
- for blended families, 3785
- for children with TBI, 653
- clinical applications, 2755–2758
  - general principles, 2755
- clinical management issues, 2757–2758
- for conduct disorder, 3779–3781
- contraindications, 2755, 2756
- definition, 2750
- for depression, 3781–3782
- description of, 4580
- for developmental challenges, 3786
- dissociative identity disorder, 1929–1930
- diversity and, 2758
- during divorce, 3785
- engagement, 3772–3774
- evaluation, 2756
- experiential–humanistic, 2753
- for family transitions, 3785
- future directions, 3786
- gambling disorder, 1808
- group, 2758–2759
  - case vignette, 2759
- historical trends, 2749–2750
- indications, 2755
- introduction, 2748–2749
- for medical challenges, 3786
- narrative, 2755
- OCD, 1790
- OCD, child/adolescent, 3700
- for oppositional defiant disorder, 3779–3781
- overview, 3771–3772, 3772–3776*t*
- parent-child interactions during, 3777–3778
- process, 3772–3778
- pseudocrisis, 2757
- psychodynamic, 2754–2755

- psychoeducational, 2754
- psychotherapy and, 3738
- research in, 2759–2760
- safety, 3774
- schools of, 2750, 2751*t*–2752*t*, 2753–2755
- secrets in, 2757
- selective mutism (SM), 3723
- separation anxiety treatment, 3711
- solution-focused, 2755
- strategic, 2753
- structured, 2750, 2753
- for substance abuse, 3783–3785
- suicide risk, 2617
- tic disorders, 3644, 3658
- timing interventions in, 2758
- for traumatic stress, 3782–3783

- Famine, prenatal impact of, 2484
- Fanapt. *See* Iloperidone
- Fantasies
  - description of, 4580
  - family relations and, 2872–2875
  - Klein on conscious daydreaming as, 2870
  - Klein on early life experiences as, 2870
  - primitive object-related, 2875
  - sexual, 1975*f*
  - sexual dysfunction and, 1974
- Fantasy model (FM)
  - dissociation, 1898, 1898*t*
    - trauma model *versus,* 1898–1899, 1898*t*
- Faradic aversion, 2694
- Farr, William, 1165
- FAS, mental retardation and, 83–84
- FAST. *See* Financial Abuse Specialist Team
- Fastigial nucleus, 26
- Fatal familial insomnia (FFI), 600
  - cause of, 601*t*
  - characteristics of, 349–350

- description of, 4580
- neurological manifestations, 611–612
- presentation, 601

- Father substitutes, 4580
- Father surrogates, 4580
- Fatigue
  - cancer-related, 2253, 2263–2265, 2284
    - corticosteroids for, 2265
    - diagnostic criteria, 2263*t*
    - mechanisms of, 2264
    - psychotropic medications for, 2265*t*
  - depression and, 571
  - description of, 4580
  - HIV-related, 587, 3237–3238
  - in multiple sclerosis patients, 562
  - in myasthenia gravis, 633
  - quality of life and, 2295*t*
  - RT-related, 2271
  - stress and, 2321
  - treatment
    - bupropion, 3004
    - DA receptor agonists, 3237–3238
- Fatty acid amide hydrolase (FAAH)
  - anxiety disorders, 1750
  - inhibitors, PTSD, 1823
  - PTSD, 1743*t*
- Fatty liver, 1268
- Faulkner, Larry
  - ABPN and, 4404*t*
  - psychiatric training and, 4402
- Fausse reconnaissance, 4580
- Faux de mieux, 4580
- Fava, Giovanni, 2853
- Fazio-Londe syndrome, 628
- FDA. *See* Food and Drug Administration
- FDA Modernization Act (FDAMA), 2931
- Fear

- anatomical correlates of, 1753–1757
- in anxiety disorders, 699
- benzodiazepines and, 1771
- conceptualization of, 1720–1721
- description of, 4580
- dopamine and, 1745
- GAD and, 1764
- learned, PTSD and, 1817–1818
- learning and extinction, 1765, 1765*f*, 1766*f*
  - PTSD and, 1818
- and natural selection, 776
- neurochemical basis of, 1741–1742
- neurochemical systems involved in, 1757–1758
- neurocircuitry
  - amygdala, 1753–1754, 1753*f*
  - fear expression and extinction, 1754–1755, 1755*f*
  - hippocampus, 1755–1756, 1755*f*
  - preclinical studies, 1752–1753
- neuroimaging studies of, 1758–1760
- NMDA receptors and, 1747
- norepinephrine and, 1742
- phenomenology of, 1751–1753
- and state-dependent learning, 721, 722
- structure of
  - emotional processing and, 1769
- without solution, in posttraumatic stress disorder, 701

- FEAR acronym, 2699
- Fear conditioning
  - connectionist modeling of, 434
  - instrumental learning and, 416
  - PDP-style model, 429*f*
  - PTSD and, 1818
- Fear-like behavior, models of, 320–321
- Febrile convulsions, 470
- Fechner, Gustav Theodor, 4536
- Federal Bureau of Investigation (FBI)
  - Uniform Crime Reports, 2473, 2538

- on white-collar crime, 2414
- Federal Food, Drug, and Cosmetic Act, 2930
- Federal oversight. *See also specific agencies*
  - of mental health care, *4295f*
- Federally Qualified Health Centers, 4319
- Federn, Paul, 2869–2870
- Feeble-mindedness, 4580
- Feedback, description of, 4580
- Feeding disorder of infancy or childhood. *See* Avoidant/restrictive food intake disorder
- Feeding disorders, **2065–2082**. *See also* Eating disorders (ED)
  - classification of, 1174
  - diagnosis of, 2071–2073
  - differential diagnoses, 2075–2077
  - epidemiology of, 2069–2070
  - etiology of, 2070–2071
  - historical literature on, 2067
  - of infancy/childhood, **3622–3634, 3636–3648**
    - avoidant/restrictive food intake disorder, 3623–3630, 3637–3644
    - clinical features, 3623, 3625–3630, 3637, 3639–3644
    - comparative nosology, 3622–3623, 3636–3637
    - with concurrent medical condition, 3630–3631, 3644–3645
    - course, 3623–3630, 3637–3644
    - defined, 3622
    - diagnosis, 3623, 3625, 3637, 3639
    - differential diagnoses, 3623, 3625–3630, 3637, 3639–3644
    - disorder of caregiver-infant reciprocity, 3629–3630, 3643–3644
    - etiology, 3623–3630, 3637–3644
    - history, 3622–3623, 3636–3637
    - infantile anorexia, 3623–3625, 3637–3639
    - overview, 3622, 3636
    - pica, 3632–3633, 3646–3647
    - posttraumatic feeding disorder, 3626–3628, 3640–3642
    - rumination disorder, 3631–3632, 3645–3646
    - sensory food aversions subtype, 3625–3626, 3639–3640
  - laboratory exams, 2073–2075, 2074*t*
  - neuroimaging, child/adolescent, 3475
  - nosology, 2067–2069

- pathology, 2073–2075
- treatment, 2078–2082
- unspecified, 2066–2067

- Feighne, John P., 1830
- Feighner criteria, 1408
- Felbamate, for acute mania, 1686
- Felix, Robert, 4292
- Fellatio, description of, 4580
- Female genital mutilation (circumcision), 3839
- Female orgasmic disorder, 1968, 4580
- Female premature orgasm, 1975
- Female sexual/arousal disorder, 1966–1967
  - diagnostic criteria, 1966*t*
  - medical conditions and, 1971–1972
- Female to male (FTM) transgender, 1985
- Female-to-male transition, 2034
- Femaleness, description of, 4580
- Females, anatomy of, 1954–1955, 1955*f*
- Feminine identity, 4580
- Feminism, dissociative disorders and, 1870, 1871
- Feminist psychotherapy, 2824
  - family and couples, 2751*t*
- Fenfluramine
  - appetite suppression by, 69
  - EEG effects, 251–252
  - side effects of, 69
  - stereotypic movement disorders, 3673, 3687
- Fenichel, Otto
  - as psychoanalyst, 2639
  - on sexual addiction, 2039
- Fentanyl (Sublimaze, Actiq)
  - administration of, 2293*t*
  - delirium and, 1182
  - duration of action, 2395*t*
  - for pain management, 2377
  - pain management using, 2293
- Ferenczi, Sandor, 3735

- countertransference enactment, 2654
- dissociative disorders, 1869
- intensive short-term dynamic psychotherapy, 2797
- work of, 4580

- Fernald, Walter, 3492
- Fertility
  - control of, 2511–2514
    - contraception, 2512
    - elective abortion, 2511–2512
    - sterilization, 2512–2514
  - infertility and, 2508–2515
  - vilazodone and, 3346
- Festination in PD, 544
- Fetal abuse, 3838–3839
- Fetal alcohol effects, 1268
- Fetal alcohol effects (FAE), 145, 1267, 1268
- Fetal alcohol spectrum disorders (FASDs), **637–639**
  - behavioral findings, 638
  - cognitive findings, 638
  - comorbidities, 638–639
  - criteria for, 637*t*
- Fetal alcohol syndrome (FAS)
  - description of, 4580
  - diagnostic criteria, 637, 637*t*
  - environmental factors, 2484
  - general appearance in, 455–456
  - neuronal activity and, 145
- Fetal solvent syndrome (FSS), 1334
- Fetishism, description of, 4580
- Fetishistic disorders
  - characteristics of, 2017
  - definition of, 2013
  - transvestic, 2019
- Fetus. *See also* Fetal alcohol effects (FAE); Fetal alcohol spectrum disorders (FASDs); Fetal alcohol syndrome (FAS)
  - alcohol exposure, 2484
  - brain
    - feminization, 1958

    - masculinization, 1958
  - development of, 2505
  - influences on aggression, 2425
  - neocortex, 439–441
  - sleep in, 342
- Fetzima. *See* Levomilnacipram (Fetzima)
- *Fezf2* gene, 442
- FGAs. *See* First-generation antipsychotics
- Fiber tracks, DT imaging, 271*f*
- Fibers
  - A-delta fibers, 2388–2389
  - A-fibers, pain and, 2398
  - anterior commissure, 12–13
  - associational fibers, 12–14
    - tracts as visualized, 15*f*
  - C-fibers, pain transmission and, 2388–2389, 2398
  - internal capsule, 14*f*
  - long association, 12
  - posterior commissure, 12
  - projection fibers
    - auditory radiation, 12
    - internal capsule, 13–14
  - short association, 12
  - U fibers, 12
- Fibrillin-1, disruption of, 667
- Fibroblast growth factor 8 (FGF8), 49
- Fibromyalgia (FMS)
  - description of, 635–636, 4580
  - DHEA-S and, 170
  - diagnostic criteria, 636*t*
  - pain related to, 2390*t*, 2391
  - treatment, 2977
- Fick's principle, 294
- Field theory, 4580
- Fifth reverse vegetative sign
  - depressive syndrome and, 1639
- Fight or flight, 1742

- dopamine and, 2476
- testing, 1721
- Figure Drawing Test, 1004*t*
- Filipin, staining, 665*f*
- Filtered backprojection (FBP)
  - description of, 289, 289*f*
  - OSEM and, 290*f*
- Filtering in attention processing, 690
- FIM. *See* Functional Independence Measure
- Finances. *See also* Funding
  - competency of older adult, 4236–4237
  - elder abuse, 4253–4254
  - and electronic technology, 4280
  - of geriatric psychiatric care
    - current and future direction, 4220–4221
    - demographics, 4212–4213
    - historical background, 4213–4214
    - Medicaid system, 4218
    - Medicare system, **4214–4218**
    - mental health parity, 4219
    - out-of-pocket cost, 4218–4219
    - Patient Protection and Affordable Care Act, 4212, 4219–4220
  - as psychosocial factor, 4166, 4169
- Financial Abuse Specialist Team (FAST), 4252
- Financial barriers, 2607–2608
- Financial planning in dementia, 1200
- *Finding What Works in Health Care: Standards for Systematic Reviews* (IOM), 971
- Fine, Carla, 2600
- Finger agnosia, 4580. *See also* Autotopagnosia
- Finger tapping, 992
- Fire setting. *See* Pyromania
- First, Michael, 1169, 1176
- First-episode psychosis
  - schizophrenia, 1435, 1466*t*, 1469–1470
    - neurocognitive impairment, 1506
    - treatment, 1520
- First-generation antipsychotics (FGAs), **3060–3082**. *See also specific drugs*

- adjunctive medications, 3081
- as-needed medication, 3080–3081
- behavioral side effects, 3075, 3075*t*
- blood levels, 3081, 3081*t*
- cardiovascular effects, 3075
- CATIE, 3069–3070
- chemistry, 3061, 3061*t*, 3062*f*–3063*f*, 3064
- clinical activity, 3081
- clozapine *v.*, 3066, 3066*t*
- contraindications, 2520
- CUtLASS, 3069, 3070
- DA hypothesis and, 3065–3066, 3066*t*, 3067*f*
- DA system and, 3066–3067
- dosage/administration, 3078–3080, 3079*t*, 3080*t*
- drug interactions, 3077, 3078*t*
- drug studies, 3067–3069
  - acute psychosis, 3068
  - maintenance therapy, 3068–3069
- in elderly, 3065, 3080
- EPS, 2936, 3068, 3069, 3070, 3072–3075
  - chronic, 3074–3075
- EUFEST, 3069, 3070
- history, 3060
- laboratory interferences, 3078
- long-acting, 3065, 3082
- maintenance treatment, 3068–3070, 3081–3082
- mechanism of action, 3065–3067
- metabolism, 3065, 3078
- movement disorders induced by, 2936–2945, 2937*t*
- for nonaffective psychosis, 2516–2517
- organ/system effects of, 3069
- pharmacodynamics, 3065–3067
- pharmacological actions, 3064–3065
  - absorption, 3064–3065
  - distribution, 3065
  - elimination and, 3065
  - metabolism and, 3065

- pharmacokinetics, 3064, 3064*t*, 3065, 3078*t*
- precautions/adverse reactions, 3072–3077
  - akathisia, 3072
  - behavioral, 3075, 3075*t*
  - blood, 3076
  - cardiovascular, 3075
  - dystonia, acute, 3072
  - endocrine, 3076
  - EPS, 3072–3075
  - eyes, 3076
  - gastrointestinal, 3076
  - kidney and urinary function, 3076
  - lactation, 3077
  - liver, 3076
  - neurological side effects, 3072, 3072*f*
  - NMS, 3073–3074
  - orthostatic hypotension, 3075
  - parkinsonism, drug-induced, 3072–3073
  - perioral tremor and, 3074
  - pregnancy, 3077
  - seizure threshold, 3077
  - sexual function, 3076–3077
  - skin, 3076
  - sudden unexplained deaths, 3075–3076
  - tardive dyskinesia, 3074–3075
  - temperature regulation, 3077
- psychosis, acute, 3060, 3068, 3069–3070, 3078–3082
- side effects, 3072, 3072*f*, 3075–3077, 3075*t*
  - behavioral, 3075, 3075*t*
  - blood, 3076
  - cardiovascular, 3075
  - endocrine, 3076
  - eyes, 3076
  - gastrointestinal, 3076
  - interactions, 3078
  - kidney and urinary function, 3076
  - lactation, 3077

    - liver, 3076
    - neurological, 3072, 3072*f*
    - orthostatic hypotension, 3075
    - pregnancy, 3077
    - seizure threshold, 3077
    - sexual function, 3076–3077
    - skin, 3076
    - sudden unexplained deaths, 3075–3076
    - temperature regulation, 3077
  - therapeutic indications, 3068–3070, 3069–3072, 3069*t*, 3078–3082
    - BPD, 3070
    - delirium, 3071
    - delusional disorder, 3070
    - dementia, 3071
    - depression, 3070
    - HD, 3071–3072
    - ICD, 3071
    - Lewy body disease, 3072
    - mania, 3070
    - PD, 3072
    - PDD, 3071
    - psychosis, acute, 3060
    - psychotic disorder, substance-induced, 3071
    - schizoaffective disorder, 3069–3070
    - schizophrenia, childhood, 3071
    - schizophrenia and, 3069–3070
    - TS, 3071
- First-line antidepressants, 1678
  - selecting, 1679–1680, 1679*t*
- First-person perspectives, 336–337
- First-rank symptoms (FRS), 1871
- *First Year of Life* (Spitz), 4542
- Fiscal rationale, mental health care, 4335–4336
- Fischer, Charles, 1869
- Fish consumption, 116–117, 2547
- Fisher, Ronald P., 2734
- Fist-clenching

- competing response training and, 2698
- hypnotherapy and, 2716–2717, 2721–2722
- Fitness, and behavioral ecology, 774–775
- Five-Factor Model (FFM)
  - factor analysis, 935–936
  - of normal personality, 2161
  - Temperament
    - as child psychiatric evaluation consideration, 3378
    - disruptive behavior disorders and, 3609
- Fixation, definition of, 4580
- Fixed ideas, dissociative disorders and, 1869
- Fixed-interval schedule. *See* Schedule of reinforcement
- Fixed intervals, intermittent reinforcement, 2872
- Fixed-ratio schedule. *See* Schedule of reinforcement
- FK506-binding protein (FKBP), 399–400, 399*f*
  - anxiety and, 1740
  - PTSD and, 1820
- *FKBP5* gene, 1889
- Flagellantism, 4580
- Flagellation, 2582
- FLAIR imaging, 466
- Flashback hallucinosis, 4580
- Flashback transference, 1925
- Flashbacks
  - cannabis-related, 1308–1309
  - dissociative identity disorder and, 1905
  - DSM-5 on, 1874
  - hallucinations in, 1911
  - nonclassic dissociative amnesia and, 1934
  - in PTSD, 2428
  - violence and, 2428
- "Flashing," 2016
- Flat affect, description of, 4580
- Flavor hedonics, 358–359
  - rat, 359*f*
- Flexibilitas cerea. *See* Catalepsy
- Flexner, Abraham

- *Medical Education in the United States and Canada* (1910), 4323
- medical education report (1910), 2661, 4402
- reforms generated by, 4308

- Flight into health, 1632
- Flight of ideas
  - description of, 4581
  - manic syndrome, 1640
  - as psychiatric symptomatology, 1120
- Floccillation, 4581
- Flooding
  - description of, 4581
  - extinction to reduce anxiety and, 2690–2691
- Flow, empirical findings, 2847
- Flow and form of thought, as psychiatric symptom, 1119–1120
- Fluconazole, drug interactions, 3106
- Flucytosine, 586
- Fluency deficits, 638
- Fluent aphasia, 4581
- Flumazenil (Romazicon)
  - chemistry, 2992
  - delirium treatment, 1189–1190
  - dosage/administration, 3003
  - drug interactions, 3003
  - laboratory interferences, 3003
  - panic disorder and, 1747
  - pharmacological actions, 3002
  - precautions/adverse reactions, 3003
  - therapeutic indications, 3003
- Flunarizine (Sibelium)
  - drug studies, 3019
  - for migraine prevention, 623
- Flunitrazepam (Rohypnol), 1377
  - amnesia, 2998
  - delirium and, 1182
- Fluorine atoms, ^{18}F, 286–287
- 2-Fluoro-2-deoxy-D-glucose (FDG), 296–297
- 5-Fluorouracil, fatigue following, 2264

- Fluoxetine (Prozac)
  - for aggressive behavior, 2111
  - anxiety, 3796
  - anxiety disorders, 1779–1780
    - GAD, 1783
    - social anxiety disorder, 1776
  - autism spectrum disorder, 3794
  - bipolar depression, 1690, 1691, 3155
  - bipolar II disorder, 1698
  - breast milk excretion of, 3168
  - for bulimia nervosa, 2081
  - chemistry, 3159
  - in children, 3168
  - for chronic idiopathic pruritus, 2348–2349
  - depersonalization/derealization disorder, 1945
  - depression, 1680, 4124
  - discontinuation syndrome, 3157
  - dosage/administration
    - for insomnia, 2086
    - for sexual dysfunction, 1979
- Flupentixol (Depixol)
  - bipolar I disorder, 1696*t*
  - chemistry, 3064
- Fluphenazine (Prolixin)
  - administration route, 3079
  - chemistry, 3062*f*, 3064
  - for chorea, 555
  - depression, 3070
  - dosage, 3080
  - drug studies, 3068–3069
  - history, 3060
  - pharmacokinetics, 3065
  - psychosis, acute, 3079
- Flurazepam (Dalmane)
  - insomnia, 2996
  - metabolism, 2993
- Flurouracil (Adrucil), side effects, 2255

- Fluspirilene, chemistry, 3064
- Flutamide (Eulexin), 2503
- Fluvastatin (Lescol), 3131
- Fluvoxamine (Luvox)
  - anxiety disorders, 1779–1780, 3793, 3796
  - breastmilk excretion of, 3168
  - chemistry, 3161
  - for chronic idiopathic pruritus, 2348–2349
  - dissociative identity disorder, 1927*t*, 1928
  - dosing/administration, 573*t*, 3166
  - drug-drug interactions, 2199, 3164, 3165
    - in cancer, 2261
    - melatonin, 3106
  - EEG effects, 252
  - for excoriation disorders, 2352
  - OCD, 3155
    - child/adolescent, 3701, 3792
  - pharmacokinetics, 3164, 3164*t*, 3165
  - pharmacological actions, 3161
  - selective mutism in children, 1784
  - sexual dysfunction and, 1779*t*
  - side effects, 573*t*, 1679*t*
  - tacrine and, 3042
  - targets for, 69
  - therapeutic indications, 3155
  - tic disorder, 3646, 3660
- *FMRI* gene, 444–445
- Foa, Edna, 1769
- Focal atrophy syndromes, 628
- Focal radiation necrosis, 2271
- Focality, brain stimulation, 3302–3303, 3302*f*
- Focused transference, 2754, 2755
- Folate, beneficial effects, 2547–2548
- Foley, Katherine, 2289
- Folia, cerebellum, 26
- Folic acid, diets deficient in, 40
- *Folie à deux,* 4581

- *Folie à double forme,* 1603
- Folk illnesses, 2444
- Follicle-stimulating hormone (FSH)
  - function of, 2239
  - menstrual physiology and, 2507
  - peripheral endocrine disorders and, 2248*f*
  - production of, 2239
- Follow-up examinations, 4581
- *FOLR1* gene mutation, 664
- Fonagy, Peter, 2870
- Food. *See also* Diet and nutrition; Eating disorders; Feeding disorders; Meal(s)
  - aversions, 359–360
  - cravings, 359
  - patient refusal of, 2313–2314
  - preferences, 359–360
- Food, Drug, and Cosmetic Act of 1938, 2930
- Food addiction (FA), obesity and, 2213
- Food allergies, feeding disorders and, 3630, 3644
- Food and Drug Administration (FDA). *See also* Black box warnings
  - Amendments Act of 2007 (FDAAA), 2931
  - anticonvulsant warnings, 2966
  - black box warnings, 2916*t*, 2918–2919, 2966
  - Critical Path Initiative, 2935
  - description of, 4581
  - drug regulation and approval, **2930–2935**
    - abuse liability assessment, 2934
    - clinical hold, 2933
    - definition of drug, 2932
    - efficacy review of, 2933
    - failures, 2930, 2931
    - history, 2930–2931
    - IND review process, 2931, 2933
    - labeling, 2930, 2934
    - marketing application and, 2933
    - new drug development and, 2930–2931, 2932–2935
    - nonclinical studies, 2932–2935
    - pharmacology and, 2934

    - phases of, 2932–2935
    - postapproval phase of, 2934–2935
    - postmarketing safety, 2935
    - quality and, 2934
    - regulatory decisions and, 2934
    - safety review of, 2933
    - toxicology and, 2934
  - future directions, 2935
  - gender issues, 4245
  - guidelines for antidementia drugs, 4139–4140, 4141*t*
  - Modernization Act of 1997 (FDAMA), 2931
  - Pregnancy and Lactation Labeling Rule, 2518
  - psychopharmacology regulation, 2905, 2918–2919
    - black box warning, 2916*t*, 2918–2919
    - product labeling or information, 2918
  - on SSRIs for OCD, 1789
  - structure, 2931
- Food neophobia, 3625, 3639
- For-profit hospitals
  - description of, 4581
  - rise/decline of, 4329–4330, 4336
  - unscrupulous practices/financial losses, 4329
- Foramina of Monro, 14*f*
- Forced recall, hypnotic hyperamnesia and, 2732
- Forced swim tests, rodent, 321
- Ford, Gerald, 4311
- Ford, Julian, 2719
- *Ford v. Wainwright,* 4438
- Fordyce, Michael, 2852
- Forebrain, development, 48–49
- Foreign accent syndrome, 457*f*
- Forensic assessments, 3391
- Forensic evaluations
  - dissociative amnesia, 1940
  - dissociative identity disorder, 1926
  - neuropsychological, 987
- Forensic issues, **4232–4240**

- administrative aspects
  - disability benefits, 4238–4239
  - elderly impaired physician, 4238
- civil aspects of aging adults, 4233–4238
  - competency
    - advanced directive, 4238
    - death and dying, 4235–4236
    - driving, 4238
    - financial decisions or contract, 4236–4237
    - living will, 4238
    - power of attorney, 4238
    - as research subject, 4235
    - will, 4237–4238
  - informed consent, 4234
  - undue influence, 4234–4235
- correctional aspects
  - disability benefits, 4239
- criminal behavior by aging adults
  - competency to stand trial, 4232
  - medication to restore competency, 4232–4233
  - mental health defense, 4233
  - restoration of sanity, 4233
  - sentencing, 4233
- nosological/theoretical aspects, 4239

- Forensic patients
  - civil admissions *vs.,* 4329–4330
  - increase in, 4325, 4326
  - in psychiatric hospitals, 4344
- Forensic psychiatry, **4433–4491,** 4547
  - certificate examinations, 4427
  - child/adolescent, **3861–3868**
    - civil litigation, 3864–3865
    - custody issues, 3863–3864
    - evaluations, 3862–3863
    - evidentiary standards, 3867
    - future directions, 3867
    - juvenile justice cases, 3865–3867

- overview, 3861–3862
- clinical-legal issues, 4433–4441
  - children and law, 4439–4440
  - confidentiality and privilege, 4435–4436
  - criminal competence, 4438–4439
  - criminal responsibility and psychiatric illness, 4439
  - doctor-patient relationship, 4433–4434
  - expert witness testimony, 4437–4438
  - involuntary hospitalization and treatment, 4434–4435
  - malpractice, 4436–4437
  - medical decision-making and competence, 4434
  - testamentary capacity, 4438
- correctional psychiatry, 4449–4477
  - advocacy agenda, 4474–4475
  - biomedical research, 4474
  - death row, 4472–4474
  - electronic technology, 4476
  - future directions, 4476–4477
  - history, 4449–4451, 4450*t*
  - in jails and prisons, 4455–4465, 4456*t*, 4457*t*
  - legal and ethical principles, 4451–4455, 4453*t*, 4454*t*
  - psychiatric patients, 4465–4469, 4469*t*
  - recent developments, 4475–4476
- description of, 4581
- ethical issues with, 3870
- ethics, 4441–4448
  - additional information and expertise, 4443–4444
  - aspects of medical practice, 4441–4443, 4442*t*
  - basic principles, 4443, 4443*t*
  - confidentiality, 4446, 4446*t*
  - continuous learning, 4444
  - decision-making, 4445
  - informed consent, 4446–4447, 4447*t*
  - overlapping roles and conflict of interest, 4447
  - professional boundaries, 4445–4446
  - professionalism, 4448
  - psychiatric opinions, 4447–4448

        - recognizing issues and tensions, 4443
        - risks and safeguards, 4444
        - third parties, 4448
        - values and perspectives, 4444–4445
    - hypnosis in, 2731–2735
    - neuroimaging, 4477–4490
        - common uses in court
            - death penalty mitigation, 4481–4482
            - dementia, 4481
            - traumatic brain injury, 4480–4481
        - controversies, 4487–4490
            - antisocial personality disorder, 4487–4488
            - imaging-assisted interrogation, 4489–4490
            - lie detection, 4489–4490
            - recognition memory, 4489–4490
            - risk assessment, 4488–4489
        - obstacles to use
            - capturing thoughts, 4482–4483
            - "Christmas tree effect," 4485–4486
            - normality definition, 4483–4485
            - person *vs.* brain, 4486–4487
        - techniques, 4478–4480
    - training, 4405, 4406*t*
- Foreplay, description of, 4581
- Forepleasure. *See* Foreplay
- Forests, random, 841
- Formal operations, 4581. *See also* Piaget, Jean
- Formal operations period, 3360
- Formal thought disorders, 1119–1120
    - description of, 4581
    - neurocognitive impairment, 1507
    - schizophrenia and, 1507
- Formication, description of, 4581
- Formulation in treatment, 4337
- Fornix
    - anatomy of, 33, 34*f*
    - in schizophrenia, 13

- Forward chaining, 2873
- Fos protein family, 379
- Foscarnet, 585–586
- Fosha, Diana, 2804
- Foster care, **3817–3823**
  - demographics, 3818–3820, 3818*t*
  - foster care parents, 3822
  - future directions, 3823
  - history, 3817–3818
  - mental health problems and, 3822
  - outcomes, 3820–3821
  - overview, 3817
  - placement reasons, 3820
  - psychiatrist's role in, 3822–3823
  - reactive attachment disorder and, 3665, 3666, 3679, 3680
  - treatment foster care (TFC), 3821–3822
- Foucault, Michel
  - *Histoire de la folie,* 4541
  - somatic symptom and related disorders, 1828
- Foulkes, S.H., 2737
- 4-D model of trauma-related dissociation, 1900
- Four-Factor Theory (FFT), 1906–1907
- Four-Miss Rule, DBT and, 2794
- Fourier transform, 265–266, 266*f*, 267*f*
  - Short-Time, 232, 408
- Fourth ventricle, anatomy of, 14*f*
- Fractional anisotrophy, 271
- Fractures in elderly, SSRIs and, 1779
- Fragile X mental retardation gene *(FRM1),* 398
- Fragile X mental retardation protein (FMRP) deficiency, 79
- Fragile X syndrome (FXS)
  - autism and, 3058
  - autism spectrum disorder and, 3575
  - description of, 4581
  - epigenetic etiology of, 398
  - FMRP deficiency, 79
  - gene mutations, 444–445

- intellectual disability and, 3493, 3502, 3506
- laboratory evaluation, 1016
- learning styles and, 3517
- pathogenesis of, 451

- Fragile X triplet-repeat mutations, 3456
- Fragile X–associated tremor/ataxia syndrome (FXTAS), 3506
- Fragmentary delusion, 4581
- Frailty
  - aging adult, 3976
  - and psychiatric symptoms, 4050
- Framingham Heart Study, 1544
- Frank, Jerome, 4547
- Frank, Richard, 4294
- Frankfurt Social Skills Training (KONTAKT), 3770
- Frankl, Viktor
  - existential psychotherapies, 2821
  - *Man's Search for Meaning,* 932
  - personality theory, 932–933
- Franks, Cyril, 2683
- Franzen, Jonathan, 2839
- Fraternal twins. *See* Twin studies
- Fräulein von R. (Freud case), 2640
- Free association
  - in analytic process, 891, 895
  - description of, 4581
  - diminished reliance on, 2797
  - early use by Freud, 853–854
  - modifications, psychosis and, 2881
  - psychoanalysis, 2640, 2660
  - psychodynamic psychotherapy, 2677
- Free-floating anxiety, 4581
- Free induction decay (FID), 262
- Free recall
  - dissociative amnesia treatment and, 1939
  - hypnosis, 2734
- Free will, and neurocentrism, 785–788
- Freebase, description of, 4581

- Freedman, Daniel X., 2581
- Freedom, awareness of, 2140
- Freedom of Choice Act, 2512
- Freeman, Walter, 330
- Freeze-fright, dissociative TRASC, 1890
- Fregoli syndrome, 538, 4581
- French, Thomas
  - intensive short-term dynamic psychotherapy, 2797
  - as psychoanalyst, 2643
- Frequency, definition of, 4581
- Frequency distributions, 4581
- Freud, Anna, 3735
  - analysis of Fromm, 898–899
  - on character, 3943
  - *The Ego and the Mechanisms of Defense,* 869, 900
  - as Freud translator, 2645
  - as psychoanalyst, 2643, 2644, 2659
  - work of, 4581
- Freud, Sigmund, 2641*f*. *See also* Psychoanalysis; Psychoanalysis, classical
  - *On Aphasia,* 848–849
  - Archives, 2645
  - and Breuer collaboration, 851—853
  - cathartic method and, 2639
  - on childhood sexuality, 1953
  - conversion disorder and, 1839
  - on death anxiety, 2288
  - development of treatment, 850–856
    - concentration, 853
    - free association, 853–854
    - infantile sexuality, 855–856
    - repression, 855–856
    - resistance, 855
    - seduction hypothesis, 855–856
  - on dirt, 2336
  - dissociative disorders, 1868–1869, 1870
  - dream interpretation, 856–860
    - affects, 859

- condensation, 859
- displacement, 858–859
- primary and secondary process, 860
- projection, 859
- regression, 859–860
- representability, 858
- secondary revision, 859
- symbolism, 858
- theory of dreaming, 857–858

- *The Ego and the Id,* 860, 866, 872, 899
- on God, 2585
- group psychotherapy and, 2736–2737
- on homosexuality, 1986–1987
- hypnosis and, 2706
- on “infantile amnesia,” 60
- *Inhibitions, Symptoms, and Anxiety,* 867
- instinctual theory, 862–865
  - characteristics of
    - aggression and ego, 864
    - infant sexuality, 863–864
    - libido, 863
    - life and death, 864–865
  - concepts of, 862–863
- *The Interpretation of Dreams,* 4537
- on loss and impermanence (adulthood), 3945
- main approach, 2659
- mental health definition, 754
- metapsychology, 2649
- *Mourning and Melancholia,* 868, 872
- narrative psychiatry and, 2831
- *Outline of Psychoanalysis,* 868
- on paraphilias, 2014
- on personality structure, 2460
- *Project for a Scientific Psychology,* 849
- psychic apparatus
  - ego, 867–872
  - id, 867

- psychic energy, 849–851, 850*t*
- psychoanalysis of, 4291
- psychoanalytic research and, 2673
- on psychosexual developmental, 3363–3364
- psychosexual phase integration with object relationism
  - phases of development, 875–877, 876*t*
- on psychosis and transference, 2868
- on psychotherapy, 2638
- short-term dynamic psychotherapies and, 2796
- *Standard Edition of the Complete Psychological Works of Sigmund Freud,* 887
- structural theory and ego psychology, 865–873
  - historical development, 866–867
- on study of the emotions, 2178
- on temperament theory, 3479
- thought types by, 1119
- *Three Essays on the Theory of Sexuality,* 863
- topographic theory, 860–862
  - basic assumptions, 860–861
  - conscious system, 861
  - mental functioning, 862
  - preconscious system, 861
  - repressed *vs.* repressing, 862
  - unconscious system, 861–862
- translations of works, 2644–2645
- on unconscious, 1661
- work of, 4581

- Frewen, Paul, 1900
- Freyd, Jennifer
  - dissociative identity disorder, 1871
  - 4-D model of trauma-related dissociation and, 1900
- Friedman, Lawrence, work of, 2650, 2656, 2662
- Friendships, 4166–4167
- Frigidity, description of, 4581
- Frisium. *See* Clobazam
- Friston, Karl, 338
- Frith, Chris, 414
- Fromm, Erich, 921f

- personality theory, 920–921, 931
- as psychoanalyst, 2644
- psychopathology theory, 921, 933
- treatment by, 921

- Fromm, Erika
  - on chronic pain, 2726
  - on hypnotherapy, 2716, 2730, 2731
- Fromm–Reichmann, Frieda
  - CBT, psychosis and, 2867
  - IPT and, 2775
  - on psychosis, 2868, 2876, 2878
- Frontal behavioral-spatial syndrome (FBS), 549
- Frontal lobe syndrome, 4581
- Frontal lobe(s)
  - aggression and, 2478
  - anatomy of, 8
  - association fiber tracts and, 15*f*
  - in attention-deficit/hyperactivity disorder, 696–697
  - hypnosis and, 2712–2713
  - lesions of
    - sexual behavior and, 2043
  - in memory, 735
  - and mental health, 766
  - neuroimaging
    - BP-I, 3464–3466
    - major depressive disorder, 3464–3466, 3465*f*
  - tumors of
    - psychiatric symptoms, 2267
- Frontal-parietal networks, 418
- Frontostriatal-dysfunction theory, 373
- Frontotemporal cortex, 1893
- Frontotemporal dementia (FTD), 1210–1213
  - behavioral-variant (bvFTD)
    - differential diagnoses, 470
  - characteristics, 550
  - clinical features, 1210–1211
  - comparative nosology, 1210

- course and prognosis, 1212–1213
- defined, 1210
- dementia and, 1196, 1212*f*
- description of, 4581
- diagnosis, 1210–1211
- differential diagnoses, 1211–1212
- EEGs in, 251
- epidemiology, 1210
- etiology, 1210
- genetic factors, 4021
- heritability, 2530
- history, 1210
- laboratory studies, 1211, 1212*f*
- with parkinsonian signs, 550
- pathology, 1211, 1212*f*
- prognosis and course, 1212–1213
- and sexual dysfunction, 4264
- treatment, 1213

- Frontotemporal lobar degeneration (FTLD), 1212*f*, 1221
  - aggression and, 2495
  - genetic factors, 4021
  - *vs.* Alzheimer disease, 4083
- Frostbite, psychiatric screening in, 630
- Frottage, description of, 4581
- Frotteurism, description of, 4581
- Frotteuristic disorders
  - characteristics of, 2017
  - definition of, 2013
- Frustration, definition of, 4581
- *Frye v. United States,* 4437
- *FTL* gene mutation, 672
- Fuchs, Karl, 2724
- Fugue
  - description of, 4581
  - DSM and ICD diagnostic criteria, 1874
  - history, 1866, 1868, 1869
  - media and public fascination with, 1873–1874

- Fulfillment
  - adulthood, 3942–3945, 3942*f*
  - description of, 4581
- Fulford, Bill, 2833
- Full disclosure, definition of, 2290
- Full somatization disorder, 1835
- Full-spectrum bright light exposure therapy, 2558
- Functional, definition of, 4581
- Functional abdominal pain (FAP)
  - case studies, 2209
  - diagnostic criteria, 2203*t*
  - hypnosis and, 2729
- Functional analytic psychotherapy (FAP), 2700
- Functional Assessment Staging (FAST), 1196
- Functional assessments
  - in behavior therapy, 2684
  - of memory
    - new learning capacity, 739–741, 740*f*
    - psychogenic amnesia, 741–742
    - remote memory, 741
    - self-reports, 741
- Functional capacity, performance *vs.,* 4371
- Functional chest pain (FCP)
  - case studies, 2209–2210
  - diagnosis of, 2203
  - presumed esophageal origin, 2203
- Functional communication training, 3554
- Functional disorders, description of, 4581
- Functional family therapy (FFT)
  - adolescents and young adults with BD, 1701–1702
  - research in, 2759
- Functional hypothalamic anovulation, 2507–2508
- Functional impairment
  - dementia and, 1194, 1194*t*
- Functional Independence Measure (FIM), 4032
- Functional magnetic resonance imaging (fMRI)
  - ADHD, 3464

- after TBI, 525–526
- autism spectrum disorder and, 3577–3579, 3578*f*, 3578*t*
- brain energy utilization, 276*f*
- brain imaging, 1717, 1759
- brain stimulation and, 3303–3304
- factitious disorder imposed on self, 1851
- hypnosis, 2712–2713
- methodology, 274–276
- panic disorder, 1761
- pediatric considerations, 3460–3461, 3460*f*
- psychotherapy and, 2675
- during right-hand movement, 275*f*
- schizophrenia, **1475–1488, 1476***f*
  - analytic methods, 1477–1478
  - arterial spin labeling, 1476, 1479*f*
  - of at-risk individuals, 1486–1487, 1487*f*
  - blood oxygenation level-dependent imaging, 1475–1476, 1477*f*
  - cognition studies, 1480–1482, 1483*f*
  - emotion studies, 1477*f*, 1482–1485, 1484*f*
  - event-related, 1476, 1478*f*
  - future directions, 1487–1488
  - genetics studies, 1486
  - pharmacology studies, 1485–1486
  - special considerations, 1479–1480, 1481*f*, 1482*f*
- social anxiety disorder, 1763
- specific phobias, 1762
- tic disorder, 3639, 3653

- Functional medicine, overview, 2546. *See also* Complementary, alternative, and integrative medicine
- Functional neurological symptom disorders, 4581. *See also* Conversion disorders
- Functional performance, capacity *vs.,* 4371
- Functional selectivity, 129, 129*f*
- Functional skill interventions, 3543–3544
- Functionalism, 2842. *See also* Positive psychology
- Funding. *See also* Finances; *specific agencies; specific legislation*
  - federal, 4287, *4288t,* 4290–4291
  - Medicaid, 4288, 4289

  - for mental health care, 4289
- Funding imbalance effect, 1436
- Fungal toxins, war or terrorism and, 2572*t*
- Funnel-web spider *(Agelenopsis aperta)* toxin, 139
- Fusiform gyrus, ASD and, 56
- Fusion
  - description of, 4581
  - dissociative identity disorder treatment and, 1926
- *Future of an Illusion* (Freud), 899

G

- G72, bipolar and schizophrenic disorders and, 1610
- G-protein coupled receptors (GPCRs)
  - agonist activation, 2910–2911
    - biased agonism, 2910–2911, 2911*f*
    - desensitization, 2911, 2911*f*
    - inverse agonism, 2910
  - for growth factors, 2907–2912, 2912*t*, 2999*f*–2912*f*
  - ion channels and, 131
  - metabotropic, 2907–2910
  - plasma membrane microdomains and signaling, 2912, 2913*f*
  - receptor analysis, 2906
  - regulation of, 123–124, 124*f*
  - signaling, 120–124
  - structure and function, 2907
- G-protein subunit gene *(GNB3),* 1629
- G protein(s). *See also* Regulators of G protein signaling
  - bacterial toxins and, 2912
  - effectors, 2912–2914, 2913*f*–2915*f*, 2914*t*
    - adenylyl cyclase, 2913, 2913*t*
    - G protein-activated ion channels, 2914, 2915*f*
    - phospholipases, 2914, 2914*f*
  - function of, 120*f*
  - heterotrimeric, 2911–2912, 2912*t*
  - ion channel regulation of, 123
- G spot, 1955
- GABA. *See* γ-Aminobutyric acid
- GABA agonist-related disorders, **1374–1390**
- GABA neurons
  - chandelier class of, 17
  - density in schizophrenia, 55
  - double-bouquet, 17
  - Martonotti class of, 17
  - subpopulations of, 17
- GABA receptors, modulation of, 165
- GABA transporter I (GAT1), 17, 18*f*

- $GABA_A$ receptors
  - distribution of, 80–81
  - schematic representation of, 81*f*
  - upregulation of binding, 45
- $GABA_B$ receptors
  - $GABA_A$ *and*, 81–82
  - schematic representation of, 82*f*
- GABAergic model on bipolar disorder origin, 1608
- GABAergic system
  - buspirone and, 3009
  - gambling disorder and, 1807
  - OCD and, 1787
- Gabapentin (Neurontin), **2966–2969**
  - acute mania, 1685–1686, 1689*t*, 1690
  - anxiety disorders, 1747, 1781
  - BD and alcohol comorbidity, 1699
  - BD and generalized anxiety disorder, 1698
  - BD and panic disorder, 1699
  - BD and social anxiety disorder, 1698
  - bipolar depression, 1690, 1692*t*, 1695
  - bipolar I disorder, 1696*t*
  - chemistry, 2966*f*, 2967
  - for chronic idiopathic pruritus, 2348–2349
  - combination pharmacotherapy, 3326
  - for diabetic neuropathy, 2236, 2236*t*
  - dosage/administration, 2969
  - drug interactions, 2969
  - drug studies, 2967–2968
  - history, 2966–2967
  - intellectual disability and, 3515
  - levetiracetam and, 2975
  - for neuropathies, 629
  - off-label use of, 2969, 2972
  - pain management using
    - in dying patients, 2292
  - in pain syndromes, 2396
  - pharmacological actions, 2967

- postherpetic neuralgia, 2967, 2968
- precautions/adverse reactions, 2968–2969
- for PTSD prevention of, 2379
- SAD, 1784
- side effects, 519
- therapeutic indications, 2968

- Gabbard, Glen
  - on boundary violations of early psychoanalysts, 2641, 2650
  - on cyber communication, 2672
  - on minority physicians, 2598
  - on physician mental health, 2602
  - as psychoanalyst, 2646
  - on psychoanalytic dynamic psychotherapy, 2662
  - on sexual boundary violations, 2669
  - on social constructionism, 2652
- Gabitril. *See* Tiagabine
- GAD. *See* Generalized anxiety disorder
- GAF scale. *See* Global assessment of functioning scale
- Gage, Phineas, 522
  - condition of, 333, 766
  - executive function, 981
  - skull of, 333*f*
- Gaines, Atwood, 2447
- Gait
  - cerebellar, 558
  - disorders of, 469*t*
  - disturbances in psychiatric disorders, 1149
  - grip strength and, of aging adults, 3976
  - testing of, 458
- Galactorrhea, 4581
- Galanin, **176,** 1743*t*, 1749
- Galantamine (Razadyne, Reminyl), **3046–3048**
  - action of, 74
  - Alzheimer disease treatment, 1207–1209
  - chemistry, 3046, 3046*f*
  - cholinesterase inhibition, 4143–4144
  - dementia treatment, 1238

- dosage/administration, 3048
- drug interactions, 3048
- drug studies, 3047
- laboratory interferences, 3048
- organ/system effects of, 3047
- pharmacologic actions, 3046–3047
- precautions/adverse reactions, 3047
- therapeutic indications, 3047

- Galanter, Marc, 1947
- Galen of Pergamon, 4533
- Gall, Franz Joseph, 4535
- *Galphimia glauca,* 2554
- Galvanic skin response (GSR)
  - depersonalization/derealization disorder, 1890
  - description of, 4582
  - dissociative disorders, 1868
  - dissociative identity disorder, 1893
- Galveston Orientation and Amnesia Test (GOAT), 527
- Gamaphobia, 4582
- Gamblers Anonymous (GA), 1808
- Gambling
  - in America, 1800–1801
  - compulsive, OCD and, 1789
  - evolutionary risk-taking and, 1807–1808
  - future of business, 1809–1810
  - historical background, 1799–1800
  - nosological considerations, 1801–1802
  - U.S. law and, 1809
- Gambling disorders, **1799–1811**
  - clinical features, 1802–1804
  - clinical vignettes, 1810–1811
  - comorbidity, 1804–1805
  - description of, 4582
  - differential diagnoses, 1805
  - DSM-5 on, 1801–1802
  - epidemiology, 1802
  - etiology, 1805–1806

- family history, 1802
- future for, 1810
- imaging studies, 1806–1807
- neurotransmitter research, 1807
- other predisposing factors, 1807
- personality characteristics, 1804
- treatment, 1808–1809

- Game technology, 4278–4279
- *Games People Play* (Berne), 925
- Gamete donation, 2510*t*
- Gamete intrafallopian transfer (GIFT), 2509
- Gametogenesis, 2508–2509
- Gaming
  - adolescents and, 3373–3374, 3768
  - social interactions as, 2820
  - as treatment, 3542–3543
  - video games, 3812
- Gamma-Hydroxybutyric acid (GHB), 82, 664
  - delirium and, 1182
- Gamma oscillations
  - in brain structures, 406–407
  - in schizophrenia, 436
- Gamma rays
  - detection, 287–288
- Gamma ventral capsulotomy, outcome, 3319
- γ-Aminobutyric acid (GABA). *See also* GABAergic system
  - aggression and, 2480
  - aggressive behavior and, 2480
  - anatomy of GABAergic systems, 80
  - anxiety disorders and, 1742, 1743*t*, 1746–1747
  - benzodiazepine and
    - receptor complex of, 2991–2992, 2991*f*
  - binding sites, 81
  - delirium pathophysiology and, 1181
  - description of, 4582
  - function of, 80
  - globus pallidus use of, 31

- harm avoidance and, 2131
- mood disorders and, 1607, 1713, 1715
- in neuropsychiatric disorders, 281–282
- panic disorder and, 1762
- receptors, fear and, 1758
- removal from synaptic clefts, 3
- schizophrenia and, 1461
- substantia nigra pars reticulata use of, 31

- γ-Hydroxybuturate (G, GHB, Xyrem)
  - dependency diagnosis, 1385–1386
  - effects of, 2007
  - insomnia, 2996
  - as sedative-hypnotic, 1385–1386
- Ganciclovir for CMV, 585–586
- *Gandhi's Truth* (Erikson), 898, 904
- Ganglion cells, retinal, 208
- Gangs, feedback from, 2485
- Ganser syndrome
  - DSM and ICD diagnostic criteria, 1874
  - factitious disorder and, 1859
  - ICD-10 on, 1875*t*
  - malingering and, 2407
- Gardner-Diamond syndrome, 2354–2355
- Gargoylism, description of, 4582
- Garrod, Archibald, 149
- GAS. *See* General adaptation syndrome (GAS)
- Gas neurosis, war and, 2566
- Gases as neurotransmitters, 104–118
- Gaskill Amendment, 2661
- Gastaut-Geschwind syndrome, 513–514
- Gastrin, function of, **177**
- Gastrin-releasing peptide (GRP), 177, **177**
- Gastroesophageal reflux disorder (GERD)
  - feeding disorders and, 3630, 3644
  - insomnia and, 2087
- Gastrointestinal (GI) disorders, **2200–2210**
  - cancer, personality disorders in, 2267–2268

- functional, 2202
  - biopsychosocial model of, 2204
  - management of, 2206–2207
  - psychiatric approach to, 2205–2206
  - psychological therapies for, 2207–2208
  - psychosocial factors, 2204, 2205
  - role of reassurance in, 2205
- role of the psychiatrist in, 2203

- Gastrointestinal (GI) system
  - alcohol and, 1267–1268
  - anticholinergics and, 2958–2959
  - bleeding
    - antidepressants and, 1680
    - SSRIs with NSAIDs and, 1779
  - buspirone effects on, 3012
  - DA receptor antagonists and, 3076
  - distress and
    - disaster and military operations and, 2568
  - lithium effects on, 3098
  - MAOI effects on, 3117
  - nefazodone and, 3129
  - nonbarbiturate sedative-hypnotic agent effects on, 2989, 2990
  - opioid receptor antagonist effects on, 3143
  - problems, 1267–1268
  - SSRI effects on, 3156
  - vilazodone effects on, 3345
- Gastrulation, timing of, 40
- Gatekeeper, description of, 4582
- Gatekeeper primary care physician, 1863
- Gatekeeper test, 834
- Gateway drugs, description of, 4582
- Gaucher disease (GD), 657–685
- Gault decision, 4582
- Gaussian distribution, 830
- Gavin, Hector, 1847
- Gay, definition of, 1983, 4582
- Gay, Peter, 2641

- Gay activism, Stonewall riots and, 1988
- Gay-bashing, 2006
- Gay identities, acquisition of, 1997
- Gay panic defense, 2006
- Gay parents, children of, 2004
- Gay persons. *See also* Homosexual identity
  - clinical issues, 2006–2009
  - Internet resources for, 2012*t*
  - medical illnesses in, 2008–2009
- Gaze
  - psychic paralysis of, 463
  - reflexive, inhibition of, 464
  - voluntary downgaze, 457
  - voluntary upgaze, 456–457
- Gazzaniga, Michael, 2840
- GBD. *See* WHO Global Burden of Disease Study
- Gee, Samuel, 3667, 3681
- Gegenhalten, 458, 4582
- Geiselman, R. Edward, 2734
- Geleerd, Elisabeth, 1869
- Gelenberg, Alan, 1847
- Geller-Seifter conflict test, 321
- Gender
  - acamprosate efficacy and, 3057
  - aggression and, 2487
  - antisocial personality disorder and, 2475
  - autobiographical memory deficits and, 1886
  - changes of
    - ethical issues, 2036–2037
    - legal issues, 2036–2037
  - coronary artery disease and, 2186–2187
  - definition of, 2023, 2506
  - development of, 2506
  - diagnoses, 2025*t*
  - disaster or war and, 2569
  - dissociative disorders
    - substance abuse and, 1889

- dissociative identity disorder and, 1884
- homicide rates and, 2475, 2475*f*
- hormones
  - violence and, 2418–2420
- melatonin pharmacokinetics and, 3103
- mood disorders and, 1615
- paraphilias and, 2015
- suicide and, 2610–2612, 2611*t*, 2612*t*

- Gender confirmation surgery, 2023
- Gender constancy, concept of, 2029
- Gender differences
  - autism spectrum disorder and, 3574
  - of brain, 3462–3463, 3463*t*
  - brain development, 2418
  - burn injuries and, 2374
  - conduct disorder and, 3613, 3619–3620
  - cultural concept of, 790
  - hypothyroidism and, 2241
  - intellectual disability and, 3506
  - language disorders and, 3545
  - in mental disorders, 4518
  - oppositional defiant disorder and, 3606
  - stuttering and, 3560
  - in suicide rates, 4521
  - suicide risk and, 3681, 3695
  - temperament and, 3482, 3483
  - tic disorders and, 3638, 3652
  - violent behavior and, 2418
- Gender dysphoria
  - in adolescents, 2028
  - in adults, 2028, 2033–2035
  - in children, 2028
  - clinical approaches, 2032–2036
  - cultural issues, 2031–2032
  - definition of, 2024, 4582
  - development of, 3363
  - diagnostic criteria, 1985

- *DSM-5* diagnosis, 2028–2029
- origins of, 2029–2031
  - biological effects, 2030
  - future research, 2031
- research into, 2038–2039
- social issues, 2031–2032
- standards of care for, 2025
- treatment of
  - in adolescents, 2033
  - in prepubescent children, 2032–2033

- Gender dysphoria in childhood (GDC), 1985, 2024
- Gender identity. *See also* Sexual identities
  - acquisition of, 2029–2031
  - child/adolescent, 3768–3769
  - definition of, 1984, 2023, 4582
  - description of, 1958
  - development of, 3354–3355*t*, 3363
  - research into, 2038–2039
- Gender identity disorders, 4582
- Gender incongruence
  - of adolescence, 2024, 2029
  - of adults, 2029
  - of childhood, 2024
  - of children
    - ICD-11, 2029
  - distress associated with, 2031
- Gender issues, **4245–4250**
  - in bipolar disease, 4247
  - in depression, 4246
  - future directions, 4249–4250
  - in jails and prisons, 4466, 4475–4476
  - in neurocognitive disorders, 4245–4246
  - and pharmacodynamics, 4249
  - and pharmacokinetics, 4248–4249
  - in psychosis, 4246
  - psychotherapy, 4247–4248
- Gender labels, emergence of, 2029

- Gender roles, 1958, 1984, 2023, 4582
- Gender stability, concept of, 2029
- Gender variance
  - disorder or normal, 2026–2027
  - frequency of, 2027–2028
  - theories of, 2024–2027
    - historical, 2024–2027
- Genderqueer, definition of, 1985, 2023
- Gene expression
  - *cis*-regulatory control of, 443*f*
  - patterning in, 48
- Gene mapping
  - approaches to, 305*f*
  - association studies, 308–309
  - basic concepts of, 306–307
  - false positives, 309*f*
  - future directions, 318–319
  - genetic markers, 307
  - linkage, 306–307
  - linkage disequilibrium, 307
  - pedigree analysis, 307–308
  - pitfalls, 304
  - psychiatric disorders, **304–319**
  - recombination, 306–307
  - sib pair analysis, 308
  - statistical considerations, 309
  - strategies, 307–309
- Gene splicing, process of, 151
- Gene transcription, regulation of, 442
- Gene–environment interactions, 1117
  - aggression, 2483–2484
  - anxiety disorders, 1739–1740
  - specific learning disorder and, 3527
- Gene–gene interactions, 2483–2484
- General adaptation syndrome (GAS), 4582
- General anesthesia, withdrawal treatment, 1370
- General arousal level, biofeedback and, 2559

- General health/primary care systems
  - psychotropic medications in, 4321
- General Health Questionnaire (GHQ), 1836
- General linear model (GLM), 842
- General neurotic syndrome, dysthymia and, 1646
- General paralysis, 4582
- General paresis, 4582
- *General Psychopathology* (Jaspers), 4541
- General system theory, 4582
- Generalization, definition of, 4582
- Generalized anxiety. *See* Free-floating anxiety
- Generalized anxiety disorder (GAD), **3712–3714**
  - after disasters, terrorism, or war, 2567*t*
  - after stroke, 475
    - course/prognosis in, 485–486
    - treatment of, 487–488
  - after TBI, 537
  - agoraphobia and, 1783
  - BD and, treatment of, 1698
  - benzodiazepines and, 1747
  - CBT and, 1776
  - child/adolescent
    - clinical features, 3712–3713, 3713*t*
    - comparative nosology, 3712
    - course, 3714
    - defined, 3712
    - diagnosis, 3712–3713, 3713*t*
    - differential diagnoses, 3699, 3713–3714
    - epidemiology, 3712
    - etiology, 3712
    - neuroimaging, 3470–3471, 3471*f*–3472*f*
    - pharmacotherapy, 3714
    - prognosis, 3714
    - psychotherapy, 3714
    - treatment, 3714
  - in children and adolescents, 1767
    - management, 1784

- cholesterol levels and, 2545
- classification of, 1172–1173
- clinical features, 1727
- cognitive therapy, 2765, 2862
- comparative nosology, 1727
- definition of, 1727, 1727*t*, 4582
- description, 1769
- diagnosis, 1727
- differential diagnoses, 1727–1728
- DSM-5 on, 1721, 1723
- epidemiology, 1727
  - adults, 1731, 1732*t*
  - genetic, 1738
- in epilepsy, 514
- etiology, 1727–1728
- excoriation and, 2351–2352
- genetics and empirical risks for, 2535*t*
- glutamate-modulators and, 1748
- history, 1727
- hoarding disorder and, 1793
- HPA axis abnormalities and, 1744
- hypnosis and, 2721, 2722
- illness anxiety disorder and, 1838
- insomnia and, 2087
- management, 1782*t*, 1783
- mood disorders and, 3486, 3486*f*
- neuroimaging, 1763–1764, 1765
- norepinephrine and, 1742
- OCD *versus,* 1788
- oxytocin and, 1749
- perinatal, 2515
- pharmacotherapy and, 1782
- phenomenology of, 1752
- pregnenolone and, 1748
- psychodynamic psychotherapy, 2679
  - case study, 2679
- psychosocial interventions, 3745

- separation anxiety and, 3710
- social anxiety and, 3715
- STAPP and, 2798
- STDP and, 2810
- syncope and, 2194
- trazodone drug studies, 3249
- treatment
  - β-adrenergic receptor antagonists, 2954
  - anticonvulsants, 2973
  - antihistamines, 2982
  - benzodiazepine, 2995–2996
  - buspirone, 3009–3010
  - rTMS for, 1782
  - SSRIs, 3154
  - venlafaxine, 3149

- Generalized dissociative amnesia
  - about, 1880, 1881*t*
  - definition and case study, 1932
- Generalized epilepsy with febrile seizures, 144
- Generalized labeled-magnitude scales (gLMS), 358*f*
- Generalized tonic-clonic seizures, 4582
- Generativity, mental health and, 759
- Generosity, happiness and, 2848
- Genes
  - coding, 149–151, 152*t*
  - description of, 149
  - mitochondrial, 155*t*
  - noncoding, 149, 151–153
- Genetic anticipation, 3349
- Genetic code
  - epigenetic codes and, 386, 386*t*
  - illustration of, 151*f*
- Genetic counseling, **2525–2537**
  - case studies, 2525, 2528
  - concerns about recurrence in children, 2531*t*
  - definition, 2525
  - description of, 4582

- ethical, legal, and social considerations, 2536–2537
- families at high risk, 2530, 2530*f*
  - case study, 2531
- family medical history as pedigree, 2526*f*
- genomic testing challenges, 2536
- mood disorders, 1629
- pedigree construction, 2526*f*
- person-centered, 2534
  - case study, 2534
  - risk reduction in children, 2534, 2536
- priorities, 2531–2534
  - data needed for risk assessment, 2532
  - educating clients, 2531–2532
  - risk assessment based on FMH, 2532–2533, 2533*f*
  - risk assessment based on genomic testing, 2533
  - risk communication, 2533–2534
  - steps and process of, 2531*t*
- psychiatric disorders, 2525, 2526*f*, 2527
- schizophrenia, 1444–1445

- Genetic disorders, laboratory evaluation, 1016
- Genetic drift, and natural selection, 772
- Genetic endowment, 4582
- Genetic epistemology (Piaget), 704–705
- Genetic factors
  - aging brain and (*See* Geriatric psychiatry)
  - in anxiety disorders, 4061–4062
  - in mood disorders, 4069–4070
  - positive psychiatry, 4284
- Genetic heterogeneity, mood disorders and, 1621–1622, 1621*f*
- Genetic linkage studies
  - anxiety disorders, 1738
  - mood disorders, 1625
  - schizophrenia, 1565
- Genetic markers, 4582
- Genetic material, 4582
- Genetic polymorphisms, 3254
- Genetic syndromes, 2530

- case study, 2531
- Genetic systems theory, 4582
- Genetic variants, rare *versus* common, 156–157
- Genetics. *See also* Genetic counseling; *specific disorders*
  - ADHD, 3588–3589
  - adoptive children and, 3819–3820
  - aggression and, 2482–2484, 2482*t*
  - of alcohol-use disorders, 1269–1270, 1269*t*
  - of Alzheimer disease, 1203
  - anxiety disorders, 1735, 1737–1741
  - assortative mating and, 2530
  - of autism spectrum disorder, 3575, 3928–3929
  - of brain development, 3358
  - caffeine use, 1296
  - cannabis use, 1307
  - child/adolescent
    - association mapping, 3450–3452, 3450*t*, 3451*t*
    - candidate gene studies, 3453–3456
    - clinical correlations, 3456
    - comparative genomic hybridization, 3452
    - copy number variation analysis, 3452, 3453*t*
    - DNA variation, 3448–3449
    - future research, 3456–3457
    - generalized anxiety disorder, 3712
    - genetic testing, 3456
    - genome-wide association studies, 3452–3453, 3453*t*
    - linkage analysis, 3449–3450, 3449*t*
    - overview, 3448
    - pathway analysis, 3453, 3454*t*
    - pharmacogenomic effects, 3453–3457
    - pharmacogenomic testing, 3456
    - psychiatric disorders, 3349, 3350
    - risk and protective factors for, 3350
    - risk genes, finding, 3449–3453
  - of child maltreatment, 3833
  - conduct disorder, 3614
  - depressive disorders, 3675, 3689

- ECT genetic mechanisms, 3284–3285
- educating clients about, 2531–2532
- enuresis, 3652–3653, 3666–3667
- foster care children and, 3819–3820
- of frontotemporal dementia, 1210
- of Huntington disease, 1220
- of intellectual disability, 3492–3493, 3499–3500, 3502–3508, 3503–3505*t*
- language disorders, 3546
- for MCI, 1225–1226
- mood disorders, 1619–1630
  - epidemiology, 1619–1620
  - features of, 1620–1621
  - introduction, 1619
  - models, 1621*f*
  - selective advantage, 1623–1624
  - spectrum of, 1622–1623, 1623*f*
  - transmission of, 1621–1622
- neurocognition
  - schizophrenia and, 1504–1505
- obesity and, 2212–2213
- OCD, 1787, 3640, 3642, 3654, 3656, 3697–3698
- opioid-related disorders, 1360
- of Parkinson disease, 1219
- pathway analysis, 3453, 3454*t*
- presenting illness, genetic counseling and, 2525
- psychiatric disorders and, 2529
- PTSD, 1820
- schizophrenia, 4538
  - adoption studies, 1439, 1446
  - architecture of, 1444
  - early onset, 3728
  - family studies, 1437–1439, 1445–1446, 1564
  - fMRI studies, 1486
  - genomic-wide association studies, 1565–1566
  - heritability specifics, 1446–1447
  - isolated populations, 1429–1431
  - liability model of, 1447–1449

    - linkage, 1565
    - migrant populations, 1431, 1450–1451
    - molecular, 1439–1444
    - neurocognition and, 1504–1505
    - neuroimaging studies, 1499–1500
    - risk, during childhood and adolescence, 1449–1450
    - risk counseling, 1444–1445
    - risks, evidence of, 1445–1446, 1454, 1462–1463, 1471
    - studies, 1564–1566
    - twin studies, 1439, 1446, 1564–1565
  - of sensory food aversions, 3625, 3639
  - of separation anxiety, 3708–3709
  - social anxiety disorder, child/adolescent, 3715
  - social (pragmatic) communication disorder, 3553
  - specific learning disorder, 3524–3527
    - behavioral, 3524
    - molecular, 3524–3525
  - speech sound disorder, 3556
  - stereotypic movement disorders, 3668, 3669, 3682, 3683
  - of stimulant use disorders, 1285
  - stuttering, 3560
  - as symptom consideration, 1116
  - temperament affected by, 3482–3483, 3483*t*
  - testing
    - of children, 3408, 3456
    - clinical neuroscience, 467–468
  - tic disorders, 3640, 3642, 3654, 3656
  - tobacco use, 1346
  - variants of large effect, 2530
  - violence and, 2420–2423
- Geneva Declaration on Armed Violence and Development (GDAVD), 2417
- Genital mutilation, female, 3839
- Genital phase, description of, 4582
- Genito-pelvic pain/penetration disorders, 1969–1970, 4582
- Genitourinary cancers, 2268–2269
- Genitourinary system. *See also* Sexual function; Urinary function
  - anticholinergics and, 2959

- MAOI effects on, 3117
- nefazodone and, 3128–3129
- opioid receptor antagonist effects on, 3144
- trazodone effects on, 3251–3252

- *Genius and Insanity* (Lombroso), 4536
- Genome in a Bottle Consortium, 148
- Genome-wide association studies (GWAS)
  - anxiety disorders, 1739
  - approaches, 305*f*
  - family-based, 308
  - genetic mapping and, 3450–3453, 3451*t*, 3453*t*
  - hoarding disorder, 1793
  - mood disorders, 1626
  - OCD, 1787
  - of psychiatric disorders, 308–309
  - schizophrenia, 1426, 1441–1442, 1441*f*–1442*f*, 1565–1566
  - use of, 147
- Genomes
  - definition of, 148
  - function of, 157
  - human nuclear, 439
  - mental health and, 2527–2528
  - mitochondrial, 154–155
  - regulatory mechanisms, 162*t*
  - sizes of, 149*t*
  - studies of aging brain, 3987–3988
  - studies of neurodegenerative disorders, 4019
  - whole genome sequencing, 2929
- Genomic imprinting
  - Prader-Willi syndrome and, 3506
- Genomic testing
  - challenges of, 2536
  - genetic susceptibility and, 2536–2537
  - risk assessment based on, 2533
- Genomics
  - description of, 147
  - ethical issues of, 3874–3875

- human brain development and, 438–447
- Genotypes
  - description of, 4582
  - and natural selection, 772
- Genuineness, client-centered therapy, 2818
- Geodon. *See* Ziprasidone
- Geographic areas
  - cost differences in, 4312
  - distribution of psychiatrists, 4306
  - health care outcomes in, 4383
- Geography, mood disorders and, 1616–1617
- Geriatric Depression Scale (GDS), 1003*t*, 2283
- Geriatric inmates in jails and prisons, 4469
- Geriatric Mental State Automated Geriatric Examination for Computer-Assisted Taxonomy (GMS-AGECAT), 3962
- Geriatric psychiatry, **3947–4286**. *See also* Elderly persons
  - aging brain, **3986–3995**
    - clinical disorders and cognitive reserve, 3988
    - cognitive and motor changes, 3813, 3813*t*
    - concepts and challenges, 3986–3988, 3986*t*
    - interventions, 3994–3995, 3994*t*
    - lesions, 3991–3992, 3991*t*
    - mechanisms in, 3988–3989, 3988*t*
    - neuroanatomy, genomics, and imaging, 3986–3988
    - neurochemical changes, 3992–3993
    - in oldest-old and centenarians, 3813–3794
    - structural changes, 3989–3991, 3989*t*
    - synapses and circuits, 3990–3991
  - aging of population, 3947–3948, 3947*f*
  - alcohol and substance abuse, 4108–4118
    - demographics, 4108
    - detoxification and withdrawal, 4117
    - differential diagnoses, 4112–4113
    - DSM-5 terminology, 4109–4110
    - epidemiology, 4110–4111
    - etiology, 4111
    - future directions, 4118
    - history, 4110

- identification barriers, 4109, 4109*t*
- illicit drugs, 4115
- marijuana, 4115
- pathology and laboratory examination, 4116
- prescription and over-the-counter medication, 4113–4115
- recommended limits, 4111, 4111*t*
- screening, 4112, 4112*t*
- spectrum-of-use categories, 4110
- tobacco, 4115
- treatment, 4116–4117

- Alzheimer disease, **4078–4086,** 4079*t*
  - course and progress, 4083
  - diagnosis and clinical features, 4080–4081
  - differential diagnoses, 4082–4083, 4083*f*
  - environmental factors, 4085
  - epidemiology, 4079–4080
  - etiology, 4080
  - family interventions, 4085–4086
  - history, 4078–4079
  - pathology and laboratory examination, 4081–4082, 4081*f*, 4082*f*
  - prevention strategies, 4086
  - psychopharmacotherapy, 4084–4085, 4084*t*
  - treatment, 4078–4086
- antidepressants and mood stabilizers (*See* Antidepressant(s); Mood stabilizers)
- anxiety disorders, **4058–4065**
  - agoraphobia, 4060
  - bereavement disorder, 4061
  - course and prognosis, 4062–4063, 4062*t*
  - differential diagnoses, 4062
  - epidemiology, 4059–4060
  - etiology and risk factors, 4060
  - falling, 4060
  - future directions, 4065
  - genetic factors, 4061–4062
  - hoarding, 4061
  - other medical conditions, 4063
  - pathology and laboratory examination, 4061

- psychotherapy, 4064
- treatment, 4063
- assessment of functioning, **4024–4041**
  - clinical interview, 4030–4031, 4030*t*, 4031*t*
  - clinician rating, 4032
  - cognition, 4027
  - collateral report, 4032
  - culture and ethnicity, 4027
  - direct observation, 4036
  - driving, 4026
  - environment, 4027, 4029
  - everyday functioning, 4025*t*, 4026t
  - future developments, 4039–4040
  - mood, 4027
  - multi-tasking, 4026
  - multidomain performance-based batteries, 4028–4029*t*
  - normal aging, 4037
  - performance-based measures of, 4032–4033, 4033*t*
  - physical disability, 4027
  - psychiatric conditions, 4038
  - quality of life, 4026
  - recently hospitalization, 4039
  - self report, 4031–4032, 4031*t*
  - social functioning, 4026
  - virtual reality measures, 4033–4035, 4034–4036*t*
- CBT (*See* Cognitive-behavioral therapy)
- certificate examinations, 4427
- cognition and emotional aging, 3958
- common mental illnesses
  - anxiety disorders, 3956
  - dementias, 3956
  - depression, 3956
  - schizophrenia and other psychoses, 3957
  - substance abuse, 3957
- complementary, alternative, and integrative medicine (CAIM)
  - acupuncture, 3984–3985
  - in age-related cognitive decline and dementia, 3979–3982

- Ayurvedic medicine, 3982
- Chinese or Oriental traditional medicine, 3982
- conventional physical activity, 3979, 3983
- epidemiology of, 3977–3979, 3978*f*
- in late-life mood disorders, 3982–3985
- mind-body practices, 3979–3980, 3983–3984
- natural products, 3980–3982, 3981*t*, 3984

- delirium, **4086–4094**
  - course and prognosis, 4091–4092
  - definition, 4087
  - diagnosis and clinical features, 4089–4090
  - differential diagnoses, 4091
  - epidemiology, 4087–4088, 4088*t*
  - etiology, 4088–4089, 4088*t*
  - future directions, 4094
  - history, 4087
  - pathology and laboratory examination, 4090–4091, 4090*t*
  - prevention, 4092
  - treatment, 4092–4094, 4092*t*, 4093*t*
- delusional disorder, 4101–4102
  - course and prognosis, 4102
  - diagnosis and clinical features, 4101–4102
  - differential diagnoses, 4102
  - epidemiology, 4101
  - etiology, 4101
  - treatment, 4102, 4133
- dementia (*See* Antidementia drugs)
- demographics
  - aging people with mental illness, **3949–3957,** 3949*f*–3954*f*
  - alcohol use disorders, 3968–3969
  - Alzheimer disease and other dementias, 3956, 3963–3964
  - anxiety disorders, 3956, 3965– 3967
  - case definition, 3962
  - cohort effects, 3963
  - demographics of aging, 3959–3960
  - depression, 3956
  - future directions, 3970

- future progress, 3958
- health services use, 3969–3970
- mood disorders, 3964–3965
- normal aging, 3948
- obsessive-compulsive disorder, 3966
- phobic disorders, 3965–3966
- positive side of aging, 3957–3958
- posttraumatic disorder, 3966
- prevalence of psychiatric disorders, 3960–3963, 3960*f*, 3961*f*
- sampling and bias issues, 3962–3963
- schizophrenia, 3967–3968
- schizophrenia and other psychoses, 3957
- as subspecialty, 3955
- substance abuse, 3957
- suicide, 3969, 3969*f*

- description of, 4582
- elder abuse (*See* Elder abuse)
- electroconvulsive therapy (*See* Electroconvulsive therapy)
- family role (*See* Family interventions)
- finances (*See* Finances)
- forensic issues (*See* Forensic issues)
- gender differences (*See* Gender issues)
- genetics of neurodegenerative disorders, **4018–4024**
  - Alzheimer disease, 4020
  - Creutzfeldt-Jakob disease, 4022
  - dementia with Lewy bodies, 4021
  - frontotemporal disorders, 4021
  - future directions, 4022–4023
  - methods and principles, 4018–4019, 4019–4020
  - neurocognitive disorder, 4022
  - Parkinson disease, 4020–4021
  - transmissible spongiform encephalopathies, 4022
  - vascular dementia, 4021–4022
- group therapy (*See* Group therapy)
- HIV infections (*See* HIV (human immunodeficiency virus))
- long-term care (*See* Long-term care)
- medically ill patients, **4041–4052**

- Alzheimer disease, 4047–4048
- chronic disorders, 4047
- differential diagnoses, 4043–4045
- drug-disease interactions, 4043, 4044*t*
- falling, 4050
- frailty, 4050
- health systems, 4051–4052
- key concepts of, 4042, 4042*t*
- medications-related psychiatric symptoms, 4043
- musculoskeletal disorders, 4048–4049
- overreliance on drugs, 4051
- urinary problems, 4050
- vascular disease, 4045
- vision and hearing impairment, 4049–4050

- mood disorders, **4065–4078**
  - bereavement, 4067, 4075
  - bipolar disorder, 4067, 4076–4077
  - brain structure and function studies, 4070–4071
  - brain structure studies, 4070
  - care delivery, 4076
  - course and outcome, 4071
  - in dementia, 4066–4067
  - depression/executive dysfunction hypothesis, 4068–4069
  - and disability, 4067
  - genetic studies, 4069–4070
  - inflammation hypothesis, 4069
  - late-onset hypothesis, 4068
  - in medically ill patients, 4066
  - pharmacotherapy, 4072–4074
  - psychotherapy, 4072
  - psychotic, 4067
  - somatic treatment, 4074
  - suicide, 4067–4068
  - treatment, 4072–4076
  - vagal nerve stimulation, 4074–4075
  - vascular hypothesis, 4069
- neuroimaging, **4008–4018**

    - of bipolar disorder, 4013
    - of depression, 4010–4012
    - limitations of existing literature, 4017
    - of schizophrenia, 4013–4017, 4015*t*–4016*t*
    - techniques, 4008–4010, 4009*f*, 4010*f*
- Geriatrics. *See* Gerontology
- Germ cell tumors, testicular, 2268
- Germinomas in childhood, 677
- Gerontology
  - description of, 4582
  - in psychiatric hospitals, 4342
- Gerontophilia, description of, 4582
- Geropsychiatry. *See* Geriatric psychiatry
- Gershwin, George, 492
- Gerstmann–Sträussler–Scheinker (GSS) disease, 600. *See also* Developmental arithmetic disorder
  - cause of, 601*t*
  - characteristics of, 607–608
  - neurological manifestations, 611
  - presentation, 601
- Gesell Developmental Schedules, 4582
- Gestalt psychology, 4582
- Gestalt therapy, 2738, 4582–4583
- GHB. *See* γ-Hydroxybuturate
- Ghrelin
  - hunger and, 356–358
  - regulatory influence and, 174
  - secretion of, 677
- Gibbs, Fred, 236
- Giedd, J.N., 3365, 3373
- *Gigante, U.S. v.*
  - relevance of neuroimaging, 4483
- Gilda's Club, benefits of, 2254
- Gill, Merton
  - on analytic process, 2650
  - on dissociation and dissociative disorders, 1869
  - as psychoanalyst, 2644, 2646, 2657, 2659
  - psychoanalytic psychotherapy and, 2662

- Gilles de la Tourette, Georges, 3635, 3636, 3649, 3650, 4536
- Gilles de la Tourette syndrome. *See* Tourette disorder
- Gilligan, Carol
  - on moral development, 3370
  - on social development, 3361
- *Ginkgo biloba*
  - adverse effects, 2544*t*
  - for Alzheimer disease, 4147
  - anxiety disorders, 1782
  - biological treatment with, 2552–2553, 2554
  - in CAIM medicine, 3980–3981
- Ginseng
  - biological treatment with, 2553
  - for sexual dysfunction, 1980
- Glasgow Coma Scale
  - depressive disorders after TBI, 531–533
  - in mild traumatic brain injury, 985
  - TBI severity on, 524
- Glass, Gene V., 2796
- Glassman, Alexander H., 1713
- Glaucoma
  - cannabinoids for, 115, 1311
  - and electroconvulsive therapy, 4158
  - muscarinic receptors for, 73
- Glial cells
  - characteristics of, 4*t*
  - human brain, 8
  - NAA import, 278
  - numbers of, 3
- Glial-derived neurotrophic factor (GDNF), 100
- Glibenclamide, 138
- Glick, Robert, 2658
- Glied, Sherry, 4294
- Glioblastoma multiforme (GBM)
  - epidemiology of, 491
  - incidence of, 2266
  - management of, 491–492

- prognosis, 492
- Gliomas
  - epidemiology of, 490–491
  - hypothalamic, in childhood, 677
  - prognosis, 492
- Global aphasia, 462*t*, 4583
- Global Assessment of Functioning Scale (GAFS), 1883*t*, 1919
- Global Assessment Study (GAS), 755
- Global Burden of Anxiety, 1720
- Global Burden of Disease (GBD)
  - mental/substance abuse disorders, 827, 4517–4522, 4522*t*–4523*t*
  - rankings, 827–828, 4288
  - on worldwide disabilities, 827–828, 1425
- Global developmental delay, 4583
- Global Scientific Partnership Network (GSPN), 1170
- Globalization, 2929
- Globotriaosylceramide, 656
- Globus hystericus, 4583
- Globus pallidus, 8, 10
  - anatomy, 26, 28
  - derivation of, 28
  - input from, 25
  - substance P inhibitory projections, 30
- Glossodynia
  - medications used in, 2355*t*
  - presentation of, 2349
- Glossolalia, 4583
- Glucagon-like peptide-1 (GLP-1), 361
- Glucagonomas, presentation of, 684
- Glucocorticoid receptors (GRs)
  - anxiety and, 1740, 1744
- Glucocorticoids (GC). *See also specific glucocorticoids*
  - glucose metabolism and, 167
  - 5-HT functioning and, 1714
  - impaired response to
    - atopic dermatitis and, 2342
  - production of, 2243

- reduced norepinephrine release and, 1757
- resistance to, 197
- Glucose metabolism, NREM sleep and, 349
- Glue, hexacarbons in, 631
- Glutamate
  - anxiety disorders and, 1742, 1743*t*, 1747–1748
  - brain pathways, 76
  - delirium pathophysiology and, 1181
  - depression, 1681
  - ^{1}H spectrum, 263
  - ionotropic receptors, 76–78
  - Lamor frequencies, 262*f*
  - metabotropic receptors, 78
  - mood disorders and, 1607, 1715
  - in neuropsychiatric disorders, 281–282
  - neurotransmission by, 141
  - postsynaptic density, 78–79
  - PTSD and, 1823
  - removal from synaptic clefts, 3
  - schizophrenia
    - neuroimaging studies, 1496–1497
  - schizophrenia and, 1460
  - synthesis of, 76
- Glutamate carboxypeptidase II (GCPII), 2548
- Glutamate-decarboxylase 1 (GAD1), 1740
- Glutamate system, anxiety disorders and, 1747–1748
- Glutamatergic system
  - dissociative identity disorder and, 1891
  - OCD and, 1787
- Glutamic acid, **76–84**
- L-Glutamic acid, 77*f*
- Glutamic acid decarboxylase (GAD), 51
- Glutamine
  - ^{1}H spectrum, 263
  - Lamor frequencies, 262*f*
- Glutethimide (Doriden)
  - dependency diagnosis, 1385

- precautions/adverse effects, 2988, 2990
- as sedative-hypnotic, 1385

- Glycine
  - biological treatment with, 2550–2551
  - encephalopathy, 664
  - inhibitors, 1743*t*
  - mood disorders and, 1715
  - neurotransmission by, 82, 141, 664
  - schizophrenia
    - neuroimaging, 1499
- Glycine amidinotransferase (AGAT) deficiency, 663–664
- Glycogen synthase kinase 3 (GSK-3)
  - clock mechanisms and, 215
  - inhibition of, 126, 127
- GM2 gangliosidosis. *See also* Tay-Sachs disease
  - features of, 467*t*
  - immunohistochemistry, 658*f*
  - presentation, 658
- Gmelin, Eberhardt, 1866
- GMS–AGECAT. *See* Geriatric Mental State Automated Geriatric Examination for Computer-Assisted Taxonomy
- Go-arounds, 4583
- Goal-directed behaviors, 2684
  - executive functioning and, 3431
- Goal setting during rehabilitation, 4370
- Goals of treatment, 4338
- GOAT. *See* Galveston Orientation and Amnesia Test (GOAT)
- Goddard, Henry, 3492
- Godding, W. W., 4332
- *Godinez v. Moran,* 4438
- Goethe, Johann Wolfgang von, 2589
- Goldberg, Daniel, 2674
- Goldberg, David, 1156
- Goldberg, Lewis, 935
- Golden root *(Rhodiola rosea),* 2554
- Goldman, Howard, 4294
- Goldman–Hodgkin–Katz (GHK) equation, 132
- *Gomez, Madrid v.,* 4452

- Gonadal differentiation
  - genetics of, 1956
  - steps in, 1957*f*
- Gonadal dysgenesis. *See* Turner syndrome
- Gonadal steroid therapy, 3329, 3330–3333
- Gonadotropin-releasing hormone (GnRH)
  - aggressivity linked, 2477
  - function of, 1980
  - menstrual physiology and, 2507
- Gonadotropin-releasing hormone (GnRH) agonists, 2056–2057, 2503
- Gonorrhea, 2514
- Gooch, Robert, 4535
- Good Behavior Game (GBG), 2617
- Good days, empirical findings, 2847
- Good Lives Model, 2021
- Goodall, Jane, 2418–2419
- Goodman, Paul, 2738
- Goodness of fit, temperament and, 3488, 3489*f*
- Goodwin, Frederick, 4546
- Gotu kola *(Centella asiatica),* 1782
- Goulding, Mary, 2738
- Goulding, Robert, 2737, 2738
- GRADE. *See* Grading of Recommendations Assessment, Development, and Evaluation
- Gradient coils, MR, 265
- Gradient descent, 432
- Grading of Recommendations Assessment, Development, and Evaluation (GRADE)
  - practice guidelines, 972–973, 972*t*
- *Graduate Medical Education That Meets the Nation's Health Needs,* 4319
- Graduate psychiatric education, **4402–4414**
  - ACGME establishment and, 4404
  - Flexner report influence, 4403–4404
  - history, 4402–4414
    - current, 4403–4414
    - early, 4402–4403
  - liaison committee on, 4406
  - major organizations affecting, 4403*t*

- medical education relation to, 4414
- Next Accreditation System, 4413–4414
- NRMP, 4406–4407, 4407*t*
- overview, 4402
- program requirements, 4407–4409
  - clinical experiences, 4408, 4408*t*
  - didactic curriculum, 4408–4409
  - faculty, 4408
  - institutional, 4407–4408
  - program director, 4408
  - resident evaluation, 4409
- psychiatry milestones, 4409–4411
  - authors of, 4409*t*
  - competencies, 4412–4413*t*
  - description, 4410*t*
  - example, 4411*f*
  - format of, 4410*f*
  - subspecialty authors, 4413*t*
- residency position funding, 4407, 4407*t*
- residency training assessment, 4411
- resident recruitment, 4406–4407
- subspecialties, 4405, 4406*t*

- Graefenberg, Ernst, 1955
- Graham, Billy, on angels, 2590
- Grand mal epilepsy, 4583
- Grandiose delusions, 4573
- Grandiosity
  - description of, 4583
  - fantasies and, 2050
  - manic syndrome and, 1641
- Granisetron (Kytril), 2271
- Granule cells, 45*f*
- Granulocytes, stress and, 2338
- Graphesthesia, 459
- Grasp reflexes, 460
- Gratification, somatic symptom disorders and, 1834
- Gratitude, collecting evidence for, 2851–2852

- Graves disease, 679–680
  - exophthalmos and, 456
  - hyperthyroidism and, 2242
  - psychosomatic component, 915–916
  - remission rates, 2242
- Gray, Jeffery, 2840
- Gray matter
  - development, 3361*f*, 3461–3462, 3462*f*, 3462*t*
  - gender differences in, 2418
  - hippocampus, depression and, 1740
  - microstructure, 274
  - OCRDs and, 1786–1787
  - panic disorder and, 1761
  - periaqueductal, anxiety disorders, 1743*t*
  - PTSD and, 1820–1821
  - social anxiety disorder and, 1763
  - stuttering, 3561
- Gray-out syndrome, 4583
- Great Smoky Mountain Study (GSMS), 1614, 1734, 1734*t*, 2614
- Green, Richard, 2025, 2026
- Greenberg, Joanne
  - on psychosis, 2872–2873, 2877, 2878, 2880–2881
- Greengard, Paul, Nobel Prize for, 124
- Green's Word Memory Test (WMT), 4370
- Greenson, Ralph, 2655
  - countertransference enactment, 2654
  - as psychoanalyst, 2644
- Grew, Nehemiah, 3667, 3681
- Grief. *See also* Bereavement
  - about, 1632
  - description of, 4583
  - therapeutic approaches to
    - in burn patients, 2380
- Grief groups for families, 2303
- Griesinger, Wilhelm
  - on dementia, 1191
  - *Mental Pathology and Therapeutics,* 4535

- Grip strength, neuropsychological testing, 992
- Grob, Gerald
  - on derailed mental health care, 4294
  - on psychoanalysis, 4291
- Grob GN, 4539
- Grof, Stanislav
  - hallucinogen-assisted psychotherapy and, 2893
  - transpersonal psychotherapy on, 2825
- Grooming, OCRDs and, 1786
- Grounding, 1924
- Group A β-hemolytic streptococcus, 598
- Group analytic psychotherapy, 4583
- Group cohesion, description of, 4583
- Group debriefings, terrorist attacks and, 2572
- Group delusions, 1123
- Group dynamics, 4583
- Group formats *vs.* individual, 4190
- Group mind, description of, 4583
- Group practice, description of, 4583
- Group pressure, description of, 4583
- Group process. *See* Group dynamics
- *Group Psychology and the Analysis of the Ego* (Freud), 899
- Group psychotherapy, **2735–2748**
  - amnesia, 1941
  - anxiety disorders, 2746
  - attendance rules, 2741–2742
  - billing procedures, 2742
  - cancer survival and, 203
  - child/adolescent, **3760–3771**
    - clinical issues, 3764–3770
    - defined, 3760
    - developmental factors, 3761–3762
    - ethical issues, 3770–3771
    - future directions, 3771
    - history, 3760
    - techniques, 3762–3764
    - theoretical issues, 3760–3762

- clinical applications, 2745–2747
- co-therapy, 2745
- confidentiality, 2743
- contract modifications, 2742
- coordination of other therapies, 2742
- couples, 2758–2759
- definition, 2736
- depression, 1703
- description of, 4583
- development stages, 2743–2744
  - adjourning/separation, 2744
  - norming/intimacy, 2744
  - performing/differentiation, 2744
  - power and control, 2743
  - preaffiliation, 2743
- developmental factors, 3761–3762
- dissociative identity disorder, 1929
- for dying patients, 2302
- extragroup socialization, 2742
- fees, 2742
- future directions, 2748
- gambling disorder, 1808
- goals, 2741
- historical roots, 2736–2737
- history, 2736
- individual therapy and, 2745
- introduction and overview, 2735–2736
- IPT and, 2781–2782
  - maintaining interpersonal focus in, 2782, 2783*t*
  - techniques, 2781–2782
- leader's role, 2741
- leadership issues, 2744–2745
- members
  - loss and addition of, 2741
  - selection, 2739
- mood disorders, 2745–2746
- observation of, 2741

- orientation checklist, 2740, 2743
- personality disorders, 2746
- pre-group preparation, 2740–2743
- psychotic disorders, 2746–2747, 2747*t*
- purpose, 2741
- questions and answers, 2743
- recording, 2741
- research issues, 2747–2748
- schizophrenia, 2746–2747, 2747*t*
- selective mutism (SM), 3723
- self-help, 2738
- setting arrangement, 2741
- theoretical orientations, 2737–2738, 2739*t*
- therapeutic potential, 2738–2739, 2740*t*
- tic disorders, 3644, 3658
- time factors, 2741

- Group relations therapy. *See* Allport, Gordon
- Group Selection Questionnaire, 2739
- Group therapy, **4201–4208**
  - clinical issues
    - contraindications, 4205–4206
    - goals, 4206
    - indications, 4205
  - definition, 4201
  - ethical issues, 4206
  - future directions, 4208
  - history, 4201
  - research and evaluation, 4206–4208
  - techniques, 4203–4205
- Group Therapy Questionnaire, 2739
- Growth cones, structure of, 53*f*
- Growth factors
  - receptors for, 2907–2912, 2912*t*, 2999*f*–2912*f*
  - tissue gradients, 48
- Growth hormone (GH)
  - deficiency after TBI
    - in adults, 525

- depression and, 1716
- functions of, **174,** 2239
- immunological effects, 186*t*
- pituitary tumors secreting
  - presentation of, 678
- production of, 2239
- sleep neurophysiology and, 1717

- Growth hormone release-inhibiting factor (GIH, somatostatin), 2239
- Growth hormone secretagogue (GHS), 174
- Growth mixture modeling, 843
- Growth of neurons, and integration of, 687–688
- Growth rates, sympathomimetic effects on, 3230
- Gruenberg, Ernest, 1435
- Gruzelier, John, 2712
- *Gsh2* gene deletion, 50–51
- *Gsh* gene expression of, 49
- GSR. *See* Galvanic skin response (GSR)
- GTP. *See* Guanosine triphosphate
- Guanfacine (Intuniv, Tenex), **2945–2950**
  - absorption, 2945
  - ADHD, 2947–2948, 3594–3595, 3596
    - child/adolescent, 3789
  - administration, 2950
  - adverse effects, 2949–2950
  - anxiety disorders and, 1742
  - autism spectrum disorders, 2948–2949
  - chemistry, 2945, 2945*f*
  - child/adolescent
    - ADHD treatment, 3788–3789
  - cognition and, 2949
  - distribution, 2945
  - dosage, 2950
  - drug interactions, 2950
  - drug studies, 2946
  - elimination, 2946
  - history, 2945
  - for hypertension, 67

- laboratory interference, 2950
- metabolism, 2946
- neuroprotective effects, 2949
- nicotine withdrawal, 2948
- opioid withdrawal, 2948
- pharmacodynamics, 2946
- pharmacokinetics, 2945–2946
- pharmacological actions, 2945–2946
- precautions, 2949–2950
- PTSD, 2949
- PTSD and SUD comorbidity, 1823
- specific organs/systems and effects of, 2946
- therapeutic indications, 2947–2949
- tic disorders, 2947, 3645, 3659
  - child/adolescent, 3795
- TS, 2947
- withdrawal, 2950

- Guanidinoacetate methyltransferase (GAMT) deficiency, 663–664
- Guanine nucleotide-binding proteins (G proteins), 1715
- Guanine nucleotide exchange factors (GEFs), 125
- Guanosine triphosphate (GTP), 2906
- Guardianship, description of, 4583
- Gudjonsson Suggestibility Scale (GSS)
  - dissociative identity disorder and, 1882, 1899
  - memory distortion and, 2733
- *Guidance on the Management of Pain in Older People,* 4049
- Guided affective imagery, 2823
- Guided discovery
  - in cognitive therapy, 2770
  - of psychotic mechanisms, 2878
- Guided imagery, 2823–2824
  - clinical and ethical issues, 2823–2824
  - history, 2823
  - for insomnia, 2089
  - as somatic therapy, 2556–2557
  - techniques, 2823
  - theoretical issues, 2823

- Guided mastery, anxiety disorders and, 1774
- Guidelines. *See* Practice guidelines
- Guidelines rule, forensic psychiatry and, 2731
- Guillemin, Roger, 90
- Guilt
  - depression and, 1661–1662
  - description of, 4583
- Guilty man, Kohut on, 2647
- Gulf War syndrome, 631
- Gulllain–Barré syndrome (GBS), 586, 629–630
- Gummatous neurosyphilis, 592
- Guns
  - access to, suicide and, 3682, 3696
  - aggression, injury, death, and, 2476
  - violence and, 2476, 4555
- Guntrip, Harry
  - dissociative disorders, 1869
  - group psychotherapy and, 2737
- Gustatory hallucinations, 1413, 4583
- Gut–brain axis
  - inflammation and, 2204
  - model of, 2321*f*
  - schematic, 2205*f*
- Guze, Samuel B.
  - dissociative identity disorder, 1882
  - St. Louis criteria and, 1830
- Gynecological cancers, psychological issues in, 2269
- Gynecology, hypnosis and, 2724
- Gynecomastia, 4583
- Gynephilic orientation, 1983, 2024
- Gynephobia, 4583
- Gyrectomy, 4583
- Gyromagnetic ratios, 260

H

- H-channels, neuron, 138
- H_1-receptor antagonists, 2983
- H_2-receptor antagonists, 2982
- *Habeas corpus,* 4583
- Habenula, subdivisions of, 33
- Habit control motivation, 2698
- Habit coughs, 2228–2229
- Habit-reversal therapy (HRT), 2697–2698
  - case study, 2698
  - excoriation disorder, 1797
  - tic disorders, 3644, 3658
  - TTM, 1795
- Habit system, OCRDs and, 1786
- Habit training, Meyer theories, 909
- Habits
  - complex, hypnosis and, 2731
  - description of, 4583
  - in Pavlovian learning, 727
- Habituation
  - definition of, 166
  - description of, 4583
  - in schizophrenia, 695–696
- Haggard, Patrick, 336
- Hair pulling
  - BDD and, 1792
  - focused *versus* automatic, 1795
- Hair-pulling disorder. *See* Trichotillomania
- Halazepam (Paxipam), 2993
- Halcion. *See* Triazolam
- *Halderman v. Pennhurst State School & Hospital* (1977), 4336
- Haldol. *See* Haloperidol
- Haley, Jay
  - family therapy and, 2749
  - paradoxical therapy, 2816
  - strategic family therapy, 2753

- Halfway houses, 4583
- Hall, James, 2730
- Hall, Stuart, 2833
- Hallucinations
  - associative learning and, 413–414
  - auditory, 465, 1411–1412
  - command auditory verbal, 1412
  - delirium and, 1186
  - description of, 4583
  - dissociative identity disorder and, 1911, 1913
    - differential diagnoses, 1916
    - pharmacotherapy, 1928
  - gustatory, 1413
  - hypnagogic, narcolepsy and, 2092*f*
  - hypnosis, 2711, 2713
  - in migraine, 623
  - olfactory, 1413
  - overview, 1410–1411
  - at patient interview, 954, 954*t*, 966
  - in PD, 545
  - as perception disturbance, 1125–1127
  - psychotic disorders
    - due to medical conditions, 1594
    - early onset and, 3728
  - in schizophrenia, 432–433, 4100
  - sleep-related, 2102
  - substance/medication-induced psychotic disorders symptom, 1591
  - tactile, 1413
  - visual, 465, 1412–1413
- Hallucinogen-assisted psychotherapy, **2886–2894**
  - case reports, 2890–2891
  - clinical issues
    - indications, 2889
    - limitations, 2889
  - complications, 2889–2891
  - contraindications, 2890
  - ethical issues, 2891

- history, 2887
- introduction and definition, 2886–2887
- research and evaluation, 2891–2894
- techniques and goals, 2888–2889
- theoretical issues, 2888

- Hallucinogen persisting perception disorders (HPPD), 2890
- Hallucinogenic drugs, **1312–1327**, 4583. *See also specific drugs*
  - anxiety disorder, hallucinogen-induced, 1327
  - bipolar disorder, hallucinogen-induced, 1326
  - comparative nosology, 1318, 1318*t*
  - depersonalization/derealization and, 1887, 1891, 1943
  - depressive disorder, hallucinogen-induced, 1326–1327
  - description of, 4583
  - differential diagnoses, 1591
  - drug classes, 1312–1317, 1313*t*, 1314*f*
    - anticholinergic deliriants, 1317
    - iboga alkaloids, 1317
    - kappa opioid agonists, 1317
    - NMDA receptor agonists, 1312–1315, 1317
    - serotonergic, 1315–1316
  - EEG effects, 254
  - hallucinogen persisting perception disorder, 1323–1325
    - course and prognosis, 1324
    - defined, 1323–1324
    - diagnosis and clinical features, 1324
    - differential diagnoses, 1324
    - epidemiology, 1324
    - etiology, 1324
    - history, 1324
    - treatment, 1324–1325
  - history, 1317–1318
  - intoxication, 1318–1321
    - delirium, 1327
    - psychiatric emergency due to, 2631
  - near-death experiences and, 2589
  - other hallucinogen-induced disorder, 1325–1326
  - other hallucinogen use disorder, 1321–1323, 1322*t*

- overview, 1312, 1317
- phencyclidine use disorder, 1321–1323, 1322*t*
- physiological toxicity, 2889
- psychological toxicity, 2889–2890
- psychotic disorder, PCP-induced, 1325–1326
- serotonergic, 2886–2887
- sexual dysfunction and, 1973
- spirituality and, 2581–2582
- unspecified, 1327

- Hallucinosis, description of, 4583
- Haloperidol (Haldol)
  - acute mania, 1686, 1689, 1689*t*
    - emergency management, 1688
  - administration route, 3079
  - adverse effects of
    - dermatological, 2356*t*
  - for aggressive behavior, 2172
  - as antiemetic, 2256
  - blood levels, 3081
  - cardiac side effects, 3075
  - chemistry, 3063*f*, 3064
  - in children with burn injuries, 2376
  - for chorea, 555
  - for delirium, 1187–1188, 4093
    - in cancer patients, 2256–2257
    - in dying patients, 2301
  - delirium prevention by, 1180–1181
  - depression, 3070
  - description of, 4583
  - dissociative identity disorder, 1929
  - dosage/administration, 3080
    - for delirium in cancer patients, 2258*t*
  - drug interactions, 3013, 3078
    - lithium, 3098
    - nefazodone, 3131
  - drug studies, 3068–3069
  - EUFEST, 3070

- history, 2917, 3060
- metabolism/elimination, 3065
- mortality rate of, 4342
- movement disorder, stereotypic, 3673, 3687
- movement disorder induced by, 2936
- pharmacokinetics, 3065
- potency, 2921
- prolactin levels and, 2249
- psychiatric emergencies, 2628
- psychosis, acute, 3079
- psychotic disorders and, 1253
- schizophrenia, 3793
- seizure threshold and, 496
- side effects, 2375–2376
  - QT interval prolongation, 2199
- for terminally ill children, 2309
- thyrotoxicity and, 173
- tic disorders, 2947, 3645, 3646, 3659, 3660
- for Tourette syndrome, 4541
- TS, 2947, 3071

- Halstead–Reitan Neuropsychological Test Battery (HRNTB), 978, 992
- Halstead–Reitan Scale, 4583–4584
- Hamburger, Christian, 2024
- Hamilton Anxiety Rating Scale (HARS), 3400
- Hamilton Rating Scale for Depression (HRDS), 2686
  - for aging adult, 3976
  - in cardiac patients, 2197
  - description of, 4584
  - as disorder-related rating scale, 3400
  - MDD assessment for CBT, 1704
- Hamilton Rating Scales for Anxiety (HRSA), 2686
- Hamkins, SuEllen, 2831, 2837
- Hammond, D. Corydon, 2734
- HAND. *See* HIV-associated neurocognitive disorder
- Handedness of patients, 455
- Hands, in hypothyroidism, 173*f*
- Haphenophobia, 4584

- Haplotypes
  - definition, 155
  - in genetic mapping, 3450
- Hapology, 4584
- Happiness
  - empirical findings, 2845–2849
  - and mental health, 762–764
  - personal program, 2852
  - spirituality and, 2593
- Haptic hallucinations, 4584
- *Hardwick, Bowers v.,* 2000
- Hare, Robert, 2423, 2490–2491
- Hare PLC-R test, 2413–2414
- Harlow, Harry, 2334
- Harm Avoidance, 2130
  - Cluster C PDs and, 2163
  - description of, 4584
  - high, in children, 2159
  - neuroticism and, 2145
  - novelty seeking and, 2146*f*
  - personality disorders and, 2162*t*
  - psychobiological findings, 2131*t*
  - reward dependence and, 2146*f*
  - self-respect and, 2174
  - TCI scales, 2164*t*
  - temperament scales
    - factor analysis, 935
- Harmon, Robert, 1900
- Harmonization of classification systems, 1175–1176
- Harris, Sam
  - mindfulness meditation, 2579
  - on religion, 2591
- Hartland, John, 2716
- Hartley, David, 411–412
- Hartmann, Heinz
  - group psychotherapy, 2738
  - as psychoanalyst, 2644

- Harvard Group Scale of Hypnotic Susceptibility (HGS), 2713, 2714
- Harvard Trauma Questionnaire, 1896, 1904
- Hashimoto encephalopathy, 458
  - presentation of, 680–681
  - reversibility of, 681*f*
- Hashimoto thyroiditis, 2240–2241, 2344
- Hashish, description of, 4584
- Haslam, John, 1426
- Hatha yoga, 2557
- Hawthorn berry (*Crataegus* spp.)
  - anxiety disorders, 1782
  - biological treatment with, 2554
- Hawthorne, James, 4403
- Hayes, Steven C.
  - acceptance and commitment therapy, 2699
  - positive psychology and, 2853
- Hazard rates, 843
- HD. *See* Huntington disease
- Head
  - circumference of, 456
  - physical examination of, 456
- Head and neck cancers, 2267
- Head banging, 3669, 3683
- Head injuries. *See also* Traumatic brain injury
  - Alzheimer disease risk factor, 1205
  - child maltreatment and, 3836
  - closed
    - EEGs in, 245–246, 246*f*
  - headaches attributable to, 620–621
  - laboratory evaluation, 1016
  - mania without prior bipolarity and, 1649
- Headache(s). *See also* Migraine
  - barbiturate treatment, 2986
  - caffeine-induced, 1300
  - chemotherapeutic agents and, 2273*t*
  - chronic daily, 2391
  - cluster-type, 620

- conditions, 2390*t*
- definition of, 620
- ECT effects on, 3298
- epidemiology
  - comorbidities, 621–622
  - prevalence, 621
- etiology of, 622–623
- hypnosis, 2729
- ICHD-2 classification of, 619
- nefazodone treatment, 3127
- neuropsychiatric aspects of, **619–626**
- postcoital, 1974–1975
- posttraumatic, 626
- psychiatric patients, brain tumors and, 493
- sumatriptan for, 70
- symptoms of, 623*f*

- Healing Touch (HT)
  - as alternative treatment, 2542
  - as energy therapy, 2546, 2561, 2562–2563
- Health and Human Services, 4584
- Health care
  - costs (*See also* Finances)
    - ACA containment of, 4318
    - of antipsychotic medications, 4351
    - as barrier, 4302
    - behavioral health care in, 4293
    - controls on, 4311
    - diagnostic-related groups (DRGs) for reimbursement, 4336
    - hospitalist reimbursement, 4360
    - hospitalist role in, 4358
    - Kirkbride on, 4345
    - payment mechanisms
      - funding sources, 4325
      - innovation of, 4305, 4307, 4316, 4317*f*
      - reimbursement for chronic illness care management, 4308
      - types of, 4308, 4324
      - world-wide, 4525, 4528*t*

        - psychiatry, 4307–4308
        - public funding cuts, 4325
        - shared savings models in, 4308, 4318
        - as unsustainable, 4294
        - waste in, 4314, 4314*t*
        - world-wide, 4525, 4528*t*
    - evolution since 1990s, 4293–4294
    - in jails and prisons, 4456–4457, 4474
    - managed care, 4311
    - outcomes
        - social cohesion in, 4383
    - reform proposals since 1912, 4311*t*
    - service restructuring, 4386–4387
- Health Care Exchanges
    - as cost control feature, 4307
- Health care marketplace
    - from Clinton's reform attempts, 4293
    - Quality/accountability from, 4294
- Health care reform
    - overview of, 4309–4311
    - Affordable Care Act (2010) (*See* Affordable Care Act (2010))
    - historical perspective of, 4311–4317
    - models in, 4317–4324
    - precursor reports for, 4289*t*
    - psychiatry challenges/opportunities, 4323–4324
    - training needs for, 4323
- Health disparities, environmental factors in, 4377
- Health insurance. *See also* Affordable Care Act (2010); Health maintenance organizations; Medicaid; Medicare
    - behavior therapy and, 2701, 2702
        - empirically validated, 2705
    - confidentiality and, 2671
    - description of, 4584
    - and mental health, 757
    - psychodynamic psychotherapy for psychosis and, 2880
- Health Insurance Portability and Accountability Act (HIPAA), 2671, 3875
    - as incremental reform, 4311
    - medical record standards, 945

- privacy of patient, 4435
- Skype and, 2672
- Health literacy/numeracy, outcomes and, 4384
- Health maintenance organizations (HMOs)
  - behavioral health care within, 4293
  - description of, 4584
- Health of Vietnam-Era Women's Study, 1816
- Health policy, access and, 2185–2186
- Health promotion in public psychiatry, 4302–4305
- Health proxies, end-of-life decisions, 2305
- Health-related quality of life (HRQOL), 2330
- Health services
  - basic/clinical research *vs.* research in, 4347, 4348*f*
  - use by older adults, 3969–3970
- Health Systems Agency (HSA), 4584
- Healthy identification
  - description of, 4584
- Hearing impairment
  - in aging adult, 4007
  - from inhalant use, 1333–1334
  - language disorders and, 3546–3548
  - and psychiatric symptoms, 4049–4050
  - speech sound disorder, 3557
- Hearing loss
  - from inhalant use, 1333–1334
  - language disorders and, 3546–3548
  - speech sound disorder, 3557
- Hearsay, description of, 4584
- Hearst, Patty, 1947
- Heart. *See also* Cardiovascular system
  - carbamazepine effects on, 3036
  - opioid receptor antagonist effects on, 3143
  - physical examination of, 456
- Heart disease. *See also* Cardiovascular disease (CVD)
  - child/adolescent response to, 3847
  - and psychiatric symptoms, 4046
- Heart failure, classification of, 2191*f*

- Heart rate variability (HRV)
  - biofeedback using, 2546, 2559
  - psychiatric assessment and, 2544–2545
- Heart transplants
  - challenges of, 2192
  - compliance of patients after
    - psychosocial factors, 2192
  - definitions, 2191
  - psychiatric problems after, 2198
  - survival, factors in, 2192
  - survival rates, 2358
- Heavy drinking days, 1262
- Hebephrenia, 4584
- *Die Hebephrenie* (Hecker), 4536
- Hebophiles, 2017
- Hedonic-allostasis theory, 373
- Hedonic treadmill, 2846
- Hedonics
  - innate, 359–360
  - learned, 359–360
- Hedonism
  - description of, 4584
  - happiness and, 2848
- Heffter, Arthur, 2887
- Heidegger, Martin, 2821
- Heider, Fritz, 2695
- Heightened suggestibility model, 2708
- Heimberg, Richard, 1776
- Heinrichs–Carpenter Quality of Life Scale (QLS), 4372
- Heinroth, Johann Christian, 2177, 4535
- Helium, liquid, 264
- HELP. *See* Hospital Elder Life Program
- Helplessness
  - in aversive learning, 725
  - depression and, 1606–1607
  - description of, 4584
  - reduced norepinephrine release and, 1757

- Hematological malignancies, 2269–2270
- Hematomas, 4158
- Hematophobia, 4584
- Hematopoietic tree, 179*f*
- Heme oxygenase (HO)
  - heme metabolism of, 108
  - neuroprotective roles of, 109
- Hemineglect, attention and, 461–462
- Hemoglobin A1c (glycosylated)
  - depression and levels of, 2232, 2234*f*
  - diabetes and, 2246
  - role in treatment selection, 1022
- Hemophilia, child/adolescent response to, 3847–3848
- Hemorrhagic stroke, 474. *See also* Cerebrovascular accidents (CVA); Stroke
- Henry, George, 1232
- Hensen node, secretions by, 41
- Hepatic encephalopathy (HE), 2368
- Hepatic function. *See also* Liver
  - melatonin effects on, 3104–3105
- Hepatic monitoring, tacrine, 3042
- Hepatitis, viral, 1012–1013
- Hepatitis C infection
  - HIV comorbidity, 568
  - in HIV infections, 587
  - in jails and prisons, 4468
- Hepatotoxicity
  - MAOIs and, 3118
  - naltrexone, 3144–3145
  - nefazodone and, 3125, 3129
- Herbal medications
  - as alternative treatment, 2542
  - anxiety disorders, 1782
  - safety, 2543
- Herd instinct, description of, 4584
- Heritability
  - definition of, 1995
  - mood disorders, 1610, 1611*f*

        - neurobiology and, 1709, 1718
    - schizophrenia, 1446–1447
    - specific phobias, 1727
- Herman, Judith, 1946
- Hermaphrodites
    - definition of, 2023
    - description of, 4584
    - prenatal hormones and, 1992
- Hermeneutics, psychoanalysis and, 2649–2650
- Heroic phase, after disasters, 2566
- Heroin. *See also* Opioid use
    - dependence, prevalence, 1355
    - description of, 4584
    - EEG effects, 253
    - genetics, 1360
    - history, 1353
    - neuroimaging, 1359–1360, 1360*f*
    - pharmacology, 1357–1360, 1357*f*, 1358*f*
    - prevalence, 1355
    - psychosocial factors, 1360
    - treatment for, 1372–1373
    - use, prevalence, 1355–1356, 1356*f*
- Herpes simplex virus (HSV) encephalitis, 472, 594–595
- Herpes simplex viruses (HSVs)
    - CNS infections, 594–596
    - transmission of, 594–595
- Herpes zoster, neuropathies and, 632–633
- Hesse, Mary, 2832
- Heterochromatin
    - constitutive, 394
    - description of, 149
- Heterocyclic antidepressants (HCAs), 1780
- Heterogeneity, 4584
- Heteroreceptors, 61
- Heterosexism
    - definition of, 1999–2000
    - depressive symptoms and, 646

- experience of, 646
- Heterosexual identity
  - definition of, 2024
  - description of, 1983
- Heterosexuality
  - definition of, 1982, 4584
  - terminology, 1983
- Heterotypical cortex, 20
- Heterozygous, definition of, 4584
- Heterozygous deletions, 155
- Heuristic, definition of, 4584
- Hexacarbons, toxic neuropathy and, 631
- Hexamethonium, 74
- 5-HIAA. *See* 5-hydroxyindoleacetic acid
- High-density negative ions therapy, 2546, 2558, 2559
- High-risk behavior. *See* Risk-taking behavior
- HIlbert Transform, 408
- Hilgard, Ernest R.
  - hypnosis, 2706
    - state model of, 2708, 2711
  - neodissociation theory, 1870, 1895, 1899
- Hill, Lister, 4292
- Hill-Burton legislation, 4308
- Hinckley, John, 2670
- Hinduism, suffering and, 2588
- Hinting Test, 4370
- Hip/joint replacement, delirium following, 1181, 1183
- HIPAA. *See* Health Insurance Portability and Accountability Act (HIPAA)
- Hippocampal formation
  - anatomy of, 34*f*
  - embryonic origins of, 8
  - zones of, 33–35
- Hippocampus
  - in amnesia, 731–732, 734, 734*f*
  - anatomy of, 8, 10, 33–34
  - astrocytes in, 5
  - brain imaging, 1717

- and clinical disorders of aging brain, 3989, 3990, 3991
- coronal section, 12*f*
- dissociative amnesia and, 1890, 1893
- dissociative identity disorder and, 1882, 1894
- dysfunction, depression and, 1709
- emotional processing and, 1711, 1712
- excess early-life CRH and, 1745
- fear and, 1755–1756, 1755*f*
- fear conditioning, PTSD and, 1818
- memory suppression and, 1893
- mood disorders and, 1718, 1718*f*
- neurogenesis, 59–60
- neuroimaging
  - drives and instincts and, 2815
  - fear, 1759–1760
  - maturation and, 1765
  - panic disorder, 1762, 1764
- patterning genes, 50
- place cells, 403*f*
- PTSD and, 1820, 1821*f*, 1822
- schizophrenia and, 1458–1459
- self-awareness and, 2139–2140
- stress and, 1744
- volumetrics, 273–274
  - after TBI, 536*f*

- Hippocrates
  - on female sexual arousal, 1953
  - *Hippocratic Writings,* 4533
  - on psychosomatic medicine, 2177
  - *On the Sacred Disease,* 508
- Hippocratic Oath
  - doctor-patient relationship, 4433–4434
  - privacy of patient, 4435
- *Hippocratic Writings (*Hippocrates), 4533
- Hirschfeld, Magnus
  - on gender variance, 2024
  - on homosexuality, 1986

- Hirschsprung disease, 100
- Hismanal. *See* Astemizole
- Hispanic Americans. *See also* Race/ethnicity
  - discrimination, and psychiatric service utilization, 802–803
  - immigrants, and mental health, 804–805
- Histamine, **75**. *See also* Antihistamine(s)
  - anatomy/functions, 75
  - life cycle, 75
  - neurotransmission by, 141
  - pruritus and, 2342
  - receptors, 75, 2981
  - schizophrenia, 1496
- Histaminergic synapses, 74*f*
- *Histoire de la folie* (Foucault), 4541
- Histone acetyltransferases (HATs)
  - activities of, 389
  - function of, 382
- Histone deacetylases (HDACs), 388–389
- Histone methyltransferases (HMTs), 390*f*
- Histones, 149
  - acetylation of, 388–389, 397
  - methylation of, 390–391, 390*f*
  - modification of, 160–161, 388–389, 389*f*
  - phosphorylation, 391
- Historical background, 3537
- Historical truth, narrative truth *versus,* 2657
- History, 1265
- History of present illness
  - assessment of, 1114–1115
  - for child psychiatry evaluation, 3381
  - for infant and toddler psychiatry evaluation, 3388
  - for mental disorder diagnosis, 1233
- Histrionic personality, 4106
- Histrionic personality disorder
  - clinical criteria, 2155
  - comorbidity, 2155
  - complications, 2155

- description of, 4584
- differential diagnoses, 2155
- DSM and ICD diagnostic criteria, 1874
- epidemiology, 2155
- familial patterns, 2155
- genetics, 2155
- impairment, 2155
- sex ratio, 2155
- spectrum of, 2155

- HIV (human immunodeficiency virus), 570–571. *See also* AIDS (acquired immunodeficiency syndrome)
  - and aging, **4268–4274**
    - comorbidity and multimorbidity, 4271
    - emotional functioning, 4270, 4270*f*
    - epidemiology, 4268–4269, 4268*f*
    - everyday functioning, 4039, 4271–4272
    - future directions, 4273
    - neurocognitive outcomes, 4269
    - prevention, 4269
    - severe mental illness, 4271
    - substance use, 4270
    - successful aging, 4271
    - treatment and management, 4273
    - veteran population, 4272–4273
  - antidepressants in
    - dosages, 573*t*
    - side effects, 573*t*
  - behaviorally acquired, adolescent, 645–646
  - cancer related to, 2267
  - cannabis use for, 1311
  - central nervous system, 1244
  - child/adolescent, **3852–3857**
    - clinical features, 3855
    - epidemiology, 3853
    - etiology, 3853–3854
    - history, 3852–3853
    - neurodevelopmental sequelae, 3854–3855
    - treatment, 3855–3857

- clinic management
  - psychiatric conditions in, 569
- co-occurring conditions, 581
- depression ranking after, 4520
- description of, 4585
- engagement in care, 564*f*
- epidemiology of, 564–566
- FDA approved drugs, 568*t*
- in gay persons, 2008
- HIV treatment and, 567–569
- in jails and prisons, 4468
- laboratory evaluation, 1012
- LGBT community and, 2035
- life cycle, 567*f*
- medication adherence in, 577
- mental illness risk of, 4529, 4530
- mood disorder and, 1249
- neurocognitive disorder due to, **1244–1246**
- neurocognitive effects of, 647
- neuropathies associated with, 632
- neuropsychiatric aspects of, **564–588**
- opportunistic infections
  - cryptococcal meningitis, 585–586
  - cytomegalovirus, 585
  - toxoplasmosis, 585
- pediatric, 645–648
  - epidemiology, 645
  - mental health and, 645
  - psychiatric disorders and, 645
- prenatally-acquired, 647–648
- psychiatric disorders comorbid with, 4343, 4375
- psychiatric emergencies, 2624
- psychiatric symptomatology of, 1116
- psychological problems in, 582–585
  - PTSD, 582–583
- psychotherapeutic issues, 583
  - adherence counseling, 585

        - capacity to consent, 584–585
        - patient education, 583
        - patient notification, 584
        - posttest counseling, 583–584
        - pretest counseling, 583
        - transmission prevention, 584
    - relapse prevention techniques, 578
    - risk behaviors related to, 576–577
    - substance abuse and, 578–582
    - transmission of, depression and, 646
    - virology of, 566–567
- HIV-associated dementia (HAD), 569
    - clinical course of, 570
    - epidemiology, 570
    - risk factors for, 570
- HIV-associated neurocognitive disorder (HAND), 4269
    - in adolescents, 646
    - treatment of, 570
- HMG-CoA reductase inhibitors (statins), 4147
- HMO. *See* Health maintenance organizations (HMOs)
- Hmong of Cambodia, 2589–2590
- Hoarding
    - and anxiety disorder, 4061
    - OCD and, 1788
- Hoarding disorder (HD)
    - classification of, 1158
    - comparative nosology, 1792–1793
    - course and prognosis, 1794
    - definition, 1792
    - description of, 4584
    - diagnosis and clinical features, 1793
    - differential diagnoses, 1793–1794
    - DSM-5 on, 1785
    - epidemiology, 1793
    - etiology, 1793
    - history, 1792
    - treatment, 1794

- Hobbies, 3378
- Hobbyism, 546
- Hoffman, Irwin, 2646
- Hoffmann sign, 460
- Hofmann, Albert, 2887
- Høglend, Per, 2674–2675
- Holding environment, in disaster management, 2571, 2571*t*
- Holiday blues, 1632
- Holism, description of, 4584
- Holm procedure, 834
- Holmes, Thomas, 2317
- Holophrastic, 4584
- Holoprosencephaly (HPE), 48
- Holtzman Inkblot Technique (HIT), 1004*t*
- Home parenteral nutrition (HPN). *See also* Diet and nutrition; Feeding
  - psychological issues and, 2274
- Home sleep testing (HST), 2106
- Homeless, definition of, 4584
- Homelessness/marginal lives
  - from care shortcomings, 4317
  - from deinstitutionalization, 4293, 4294
  - psychiatric emergency care for, 2635
  - rehabilitation support for, 1533–1534
- Homeopathy
  - as alternative treatment, 2542
  - as energy therapy, 2546, 2561, 2562
- Homeostasis
  - description of, 4584
  - sleep regulation, 348–350
- Homework, in CBT, 1700–1701, 4183, 4183*t*, 4190
- Homicide, 2471
  - epidemiology of, 2417
  - global rates, 2474*f*, 2475*f*
  - murder rates, 2417
  - poverty and, 2426
  - premeditated, 2471
- Homicide ideation, 3975

- Homocysteine
  - beneficial effects, 2547
  - metabolism of, 668–669, 668*f*
  - reduction for Alzheimer disease, 4147
- Homocystinurias, 667–669
  - general appearance in, 455–456
  - heritability, 2530
- Homophobia
  - cultural issues, 1999–2000
  - definitions, 1999–2000
  - description of, 4584
  - experience of, 646
- Homosexual identity, 2024
- Homosexual orientation constant, 1998*f*
- Homosexual panic, 4584
- Homosexuality, **1982–2013**. *See also* Gay persons; Lesbian
  - acceptance of, 1982
  - age-structured, 1996
  - in aging adult, 4258–4259
  - anthropometric characteristics, 1994
  - biological views, 1986
  - clinical approaches, 2004–2010
  - closeted individuals, 1998
  - “coming out,” 1997, 2000–2001
  - cross-cultural study, 1996
  - cultural issues, 1999–2004
  - definitions, 1982–1985
  - description of, 4584
  - diagnosis, 4544
  - frequency, 1989–1990
  - gender-reversed, 1996
  - genetic research, 1995
    - twin studies, 1995
  - historical views, 1986–1989
  - intersex conditions and, 1992
  - life stage issues, 1999–2004
  - medicalization, 1986

- neurophysiological measures, 1994
- nongay identities and, 1997–1999
- normalization of, 1987–1989
- origin in childhood, 1997
- origins of, 1990–1999
  - biological studies, 1990–1996
  - neuroendocrine research, 1991
  - prenatal hormone hypothesis, 1991–1992
- prenatal hormones, 1992–1993
- psychoanalytic approaches, 1986–1987
- psychosocial studies, 1996–1999
- race/ethnicity and, 2001
- religion and, 2001–2002, 2592
- role-specialized, 1996
- self-aware individuals, 1998–1999
- terminology, 1983
- theories of, 1985–1989

- Homosexuals
  - clinical issues, 2006–2009
  - in clinical settings, 2005
- Homotopic, definition of, 11
- Homotypical cortex, 19–20
- Homovanillic acid (HVA), 3067, 4584
- Homozygous deletions, 155
- Honesty, end-of-life care and, 2307
- Honeymoon phase, after disasters, 2566
- Hooper Visual Organization Test, 4006
- Hoover, J. Edgar, 2413
- Hoover sign, 473
- Hope therapy
  - group psychotherapy and, 3760
  - positive psychology and, 2853
- Hopelessness
  - depressive syndrome and, 1637–1638
  - suicide behavior risk factor, 3684, 3698
- Hopkins Verbal Learning Test Revised (HVLT-R), 991
- Horizontal electroculogram (HEOG), 228–229

- Hormonal regulation alterations, 1715–1716
- Hormone deficiencies, 3989
- Hormone replacement therapy (HRT), 4148–4149
  - issues, 2522–2523
- Hormone therapy
  - dementia related to, 2284
  - in gender dysphoria, 2032
  - for paraphilias, 2020–2021
  - reproductive hormonal therapy, **3329–3337**
  - for sexual dysfunction, 1980
- Hormone(s). *See also specific hormones*
  - biological treatment with, 2555–2556
    - DHEA, 2555
    - estrogen, 2555
    - melatonin, 2555–2556
    - vitamin D, 2556
  - carbamazepine effects on, 3035
  - classification, 164, 165*t*
  - combined action of, 165
  - evolution, 164
  - immune cell receptors for, 184*t*
  - immunological effects, 186*t*
  - modes of action, 165–166
  - neurochemistry of aggression and, 2480–2482
  - production of, 2239
  - regulating, 165*t*
  - secretion, 164
  - and sexuality in aging adult, 4256–4258
  - substance abuse disorders and, 383*t*
  - synthesis and structure, 164–165
  - tissue specificity, 165–166
  - violence and, 2418–2420
- Horney, Karen, 919f
  - character types
    - alienation, 920
    - claims, shoulds, and self-hatred, 919–920
    - idealized image, 919

        - neurotic pride and pride system, 920
    - IPT and, 2775
    - neurosis theory, 918–919
    - *The Neurotic Personality of Our Time,* 4538
    - personality theory, 918–919, 931
    - as psychoanalyst, 2644
    - psychotherapy research and, 2673
    - therapeutic process, 920
    - work of, 4584
- Horowitz, Mardi, 1815
- Horwitz, Allan V., 1154
- Hospice
    - certificate examinations, 4427
    - dignity therapy and, 2825–2827
    - in-home programs, 2303–2304
    - inpatient, 2303–2304
- Hospice care in psychiatric hospitals, 4343
- Hospice movement, 2289
- Hospital Anxiety and Depression Scale (HADS), 2274
- Hospital care stages
    - primary care integration with, 4357–4358
- Hospital Elder Life Program (HELP), 4092
- Hospitalists
    - benefits of, 4358
    - as clinician-educators, 4359
    - definition of, 4354–4355
    - disadvantages/objections to, 4359
    - financial/political/organizational issues, 4360
    - future directions, 4360–4361
    - historical perspective of, 4356–4357
    - hospital care stages, 4357–4358
    - as inpatient treatment specialist, 4354
    - as a new specialty, 4360
    - professional organizations, 4360
    - psychiatrists as, 4356
    - roles of, 4355–4356, 4359
    - training for, 4354–4355, 4360

- Hospitalization
  - after criminal justice system discharge, 4392
  - assault victim, 2541
  - BD, 1652, 1684
  - child/adolescent
    - age-related response to, 3848–3849
    - contemporary, 3844–3845, 3844*t*, 3845*t*
    - family's response to, 3849–3851
    - history, 3843–3844
    - surgical procedures, response to, 3847
  - criminalization *vs.,* 4391
  - dissociative identity disorder, 1929
  - and everyday functioning, 4039
  - gambling disorder, 1809
  - involuntary, 4434
  - mania, 1639, 1639*t*, 1641, 1688
  - MDD, 1643, 1662
  - narrative psychiatry and, 2838–2839
  - social media tracking, 3852
  - substance use disorder, 1656
  - suicidal behavior, 1671, 1674, 1676*t*
  - suicide prevention and, 2621
  - suicide risk, 2615–2616
- Hospitals. *See* Psychiatric hospitals
- Hostile attribution bias, 2486
- Hostile patients, 2626
- Hostile transference. *See* Transference
- Hostility
  - CAD and, 2188
  - definition of, 4585
  - schizophrenia and, 1422–1423
- Hot flashes
  - antidepressant drugs in, 2523
  - hypnosis, 2724
- Hotlines, definition of, 4585
- Hotspots
  - frontal, 352

- parieto-occipital, 352
- Households of older people, 3952*t*
- Housing. *See also* Homelessness/marginal lives
  - rehabilitation support, 1533–1534
- Housing First initiative, 1533–1534, 1550
- *HOX8* gene activation, 390, 390*f*
- *Hox* gene family, 48
- *HOXA1* gene mutations, 314
- HRNTB. *See* Halstead-Reitan Neuropsychological Test Battery
- HRSD. *See* Hamilton Rating Scale for Depression
- HSA. *See* Health Systems Agency (HSA)
- $5HT_1$ receptors, **69–71**
- $5HT_2$ receptors, **71**
- $5HT_3$ receptors, **71**
- $5HT_4$ receptors, **71**
- $5HT_5$ receptors, **71**
- $5HT_6$ receptors, **71**
- $5HT_7$ receptors, **71**
- 5-HT syndrome, antidepressants and, 1680
- $5\text{-}HT_{1A}$
  - aggression and, 2480
  - receptor availability, 301–302
  - stress and synthesis of, 1714
- 5-HTT. *See* Serotonin transporter
- *5-HTTLPR* gene, 1889
- Huarte de San Juan, Juan, 4534
- Hubbard, A. M., 2891
- Hughes, Charles, 2445
- Hull, Clark, 2706, 2708
- Human aggression. *See* Aggression
- Human chorionic gonadotropin (hCG), 1014
  - placental, 2505
- Human Connectome Project, 419, 3578–3579, 4556
- Human genome
  - information in, 147
  - organization of, **148–155,** 154*f*

  - structural variation in, 155–156
  - variation, 155–157
- Human Genome Project, 4294
- Human herpes viruses (HHVs)
  - HHV-6, 594, 595
  - HHV-7, 594
- Human immunodeficiency virus (HIV), 2514. *See* HIV
- Human papillomavirus (HPV), 2514
  - head and neck cancers and
    - psychiatric symptoms in, 2267
  - unspecified communication disorder and, 3567
- Human rights of people with mental disorders, 4530–4531
- *Human Sexual Response* (Masters), 4542
- *Human Universals* (Brown), 786
- Humanism in medicine, 2288–2289
- Humanistic psychology, 2843
- Humanistic theory, 4585
- Humans. *See specific anatomy; specific conditions*
- Hume, David, 330
- Humiliation, description of, 4585
- Humor
  - character and, 2138
  - description of, 4585
  - and mental health, 761
  - in patient interview, 958
- Humoral immunity
  - description of, 183
  - regulation of, 2339*f*
- Humors, modern *vs.* ancient, 1612–1613
- Hunger, 359*f*
  - circadian rhythms of, 358*f*
  - eating in absence of, 359*f*
  - as motivation, 358*f*
  - physiological signaling of, 356, 357
  - process of, 355–356
  - scores, 358*f*
- Hunger strikes in jails and prisons, 4473

- Hunter, E.C.M., 1943
- Hunter, Edward, 1946
- Hunter, John, 1179
- Huntington, George, 551, 1220
- Huntington disease (HD, chorea)
  - aggression and, 2495
  - basal ganglia disruptions in, 30
  - cause of, 153
  - course/prognosis, 554
  - definition of, 551
  - dementia in, 1220
  - description of, 4585
  - diagnosis, 553–554
    - cognitive manifestations, 553
    - motor manifestations, 553
    - psychiatric manifestations, 553–554
  - differential diagnoses, 554
  - dissociative trance disorder and, 1951
  - epidemiology, 551
  - etiology, 552–553
  - heritability, 2530
  - history of literature on, 551
  - laboratory evaluation, 1016
  - laboratory exams, 554, 554*t*
  - mania without prior bipolarity and, 1649
  - movement disorders and, 540
  - neuropathology, 553*f*
  - OCD *versus,* 1788
  - pathological exam, 554
  - premanifest testing, 554
  - repeat lengths, 552*f*
  - saccades in, 457
  - SRIF and, 175
  - treatment, 555–556
    - FGAs and, 3071–3072
- *Huntington* gene, 552
- Huntjens, Rafaele, 1910

- *Huperzia serrata* (huperzine-A), 2553–2554
- Huperzine-A *(Huperzia serrata)*
  - for Alzheimer disease, 4146
  - biological treatment with, 2553–2554
- Hurricane Katrina, PTSD and, 2568
- HVA. *See* Homovanillic acid (HVA)
- Hydrocephalus
  - cerebrospinal fluid, 14
  - description of, 4585
- Hydrocodone, duration of action, 2395*t*
- Hydrocortisone (Hydrocortone)
  - for adrenocortical insufficiency, 2245–2246
  - PTSD prevention, 1825
- Hydrogen atoms
  - characteristic of, 260
  - isotopes of, 260
- Hydrogen sulfide
  - brain concentrations of, 109–110
  - messenger role of, 109–110
- Hydromorphone (Dilaudid)
  - administration of, 2293*t*
  - duration of action, 2395*t*
  - pain management using, 2293
- Hydrophobia, 595, 4585
- Hydrotherapy, description of, 4585
- Hydroxyamoxapine, 3255
- Hydroxybupropion, 3004, 3006, 3007
- Hydroxydesipramine, pharmacological actions, 3255
- Hydroxyindole *O*-methyltransferase, 68
- 5-Hydroxyindoleacetic acid (5-HIAA), 69
  - about, 1713
  - aggression and, 2111, 2479–2480, 2481
  - description of, 4585
- Hydroxynefazodone, 3126, 3130
- Hydroxynortriptyline, pharmacological actions, 3255
- 5-Hydroxytryptamine (5HT), 61. *See also* Serotonin
  - adverse effects, 2544*t*

- beneficial effects, 2549
- gene variants, aggression and, 2483
- serotoninergic systems and, 1714
- sleep cycle and, 1716
- synthesis of, 68*f*
- transmitters, mood disorders and, 1713

- Hydroxyzine (Atarax, Vistaril), 1387
  - anxiety disorders, 1781
  - chemistry, 2981*f*
  - dissociative identity disorder, 1927*t*
  - laboratory interferences, 2983
  - therapeutic indications, 2982
- "Hygiene hypothesis," 192–193
- Hymenectomy, 1981
- Hyperactivity
  - on ADHD, 3590
  - description of, 4585
- Hyperadrenalism, presentation of, 682
- Hyperaldosteronism
  - presentation of, 683
  - subcortical dementia in, 683*f*
- Hyperalgesia, 4585
- Hyperamnesia
  - description of, 4585
  - hypnotic, 2711, 2732
- Hyperandrogenism, 2249
- Hyperarousal, 4585
- Hyperarousal-related aggression, 2472
- Hypercalcemia, 682
- Hypercholesterolemia
  - gambling and, 1805
- Hypercorticolism, 2508. *See also* Cushing syndrome
  - clinical features, 2243–2244
  - course of, 2245
  - definition of, 2243
  - diagnosis of, 2243–2244
  - differential diagnoses, 2244

- epidemiology of, 2243
- etiology, 2243
- nosology, 2243
- pathology, 2244
- presentation of, 2270
- treatment of, 2245

- Hypercortisolemia
  - depression and, 1715, 1716
  - glucose intolerance and, 2244
- Hypercortisolism, delirium and, 1182
- Hyperekplexia, 82
- Hyperemesis gravidarum, 2724
- Hyperesthesia
  - description of, 4585
  - hypnosis, 2711
- Hyperglycemia
  - with schizophrenia, 1543
- *Hypericum perforatum* (St. John's wort), 2552
- Hyperkeratosis, 456
- Hyperkinesis. *See* Hyperactivity
- Hyperkinetic disorders, 2938, 3599. *See also* Attention-deficit/hyperactivity disorder (ADHD)
- Hyperkinetic movements, 540
- Hyperkinetic syndrome, 4585
- Hypermentalizing, BPD and, 2896, 2897
- Hyperorexia, 4585
- Hyperparathyroidism, 682
- Hyperphagia
  - description of, 4585
  - syndrome, 364
- Hyperpigmentation, 2244
- Hyperpolarization, 132
- Hyperpragia, 4585
- Hyperprolactinemia
  - consequences of, 67
  - definition of, 2249
  - diagnosis, 2250
  - epidemiology of, 2249

- etiology of, 2249
- treatment of, 1979, 2250

- Hyperprolinemia (HP), 663
- Hypersexuality
  - definition of, 2039
  - description of, 4585
  - differential diagnoses, 2043–2044
  - manic syndrome and, 1641
- Hypersomnia
  - depressive syndrome and, 1638
  - description of, 4585
  - psychiatric disorders and, 353–354
- Hypersomnolence disorder, 2090–2091. *See also* Hypersomnia
  - characteristics of, 2090–2091
  - diagnostic criteria, 2090
  - sleep stage histogram, 2091*f*
  - subtypes of, 2090–2091
  - treatment of, 2091
- Hypertension
  - adrenergic receptors for, 67
  - course/prognosis, 2193
  - definition of, 2192
  - differential diagnoses, 2193
  - epidemiology, 2193
  - essential, 915–916
  - etiology, 2193
  - and HIV infection, 4271
  - MAOIs and, 3028*t*, 3117, 3117*t*
  - nosology, 2192
  - persistent pulmonary hypertension of the newborn
    - SSRIs and, 3167
  - psychological factors affecting, 2193
  - with schizophrenia, 1542
  - stroke risk and, 483–484, 486
  - treatment, 2193
  - venlafaxine, 3150
- Hypertensive crisis. *See also* Hypertensive encephalopathy

- MAOIs and, 3117, 3118*t*
- Hypertensive encephalopathy, 4585
- Hyperthymic temperament
  - alcohol and substance abuse, 1651–1652
  - bipolar II disorder and, 1651
- Hyperthyroidism
  - characteristics of, 173
  - clinical features, 2242
  - course/prognosis of, 2242
  - definition, 2242
  - depression and, 1716
  - description of, 4585
  - diagnosis, 2242
  - epidemiology, 2242
  - etiology, 2242
  - nosology, 2242
  - pathology, 2242
  - presentation of, 679–680
  - subclinical, 2242
  - treatment, 2242–2243
- Hypertonia, 3540
- Hyperventilation
  - asthma and, 2225
  - chronic, 2228
  - definition of, 2228
  - description of, 4585
  - EEG activation studies, 244
  - triggers, 2228
- Hypervigilance, 4585
- Hypesthesia, 4585
- Hyphenated Americans, migration and, 792
- Hypnagogic foot tremor, 2105
- Hypnagogic hallucinations
  - description of, 4585
  - narcolepsy and, 2092*f*
- Hypnic jerks, 2105
- Hypnoanalysis, 4585

- Hypnodrama, description of, 4585
- Hypnoid state, description of, 4585
- Hypnophobia, description of, 4585
- Hypnosedatives, 2204
- Hypnosis, **2705–2735**. *See also* Self-hypnosis; Stanford Hypnotic Clinical Scale (SHCS); Stanford Hypnotic Susceptibility Scale (SHSS)
  - addictions, 2730–2731
  - anxiety disorders, 2721–2723
  - asthma, 2729
  - bone marrow aspiration pain and, 2725
  - burn care, 2724
  - characteristics, in clinical context, 2711–2712
  - CIA and, 1947
  - clinical applications, 2717–2735
  - complex habits, 2731
  - deep, transient depersonalization, 1887
  - definition, 2709
  - depersonalization/derealization disorder, 1945
  - depression and, 2723
  - dermatology, 2730
  - description of, 4585
  - development of, 1866
  - dimensions and phenomena of, 2709
  - dissociative amnesia, 1939–1940
  - dissociative disorders, 1869, 1870, 1871
  - dissociative identity disorder, 1915, 1926
  - dissociative psychosis, 1950
  - dissociative trance disorder, 1951
  - for dying patients, 2302
  - dysmenorrhea, 2724
  - early use by Freud and Breuer, 853–854
  - eating disorders, 2731
  - emergency room, 2724–2725
    - case study, 2725
  - false memory syndrome and, 1872
  - forensic psychiatry, 2731–2735
  - Freud and, 2639
  - functional abdominal pain, 2729

- future directions, 2735
- gynecology, 2724
- headache, 2729
- history, 2706
- hot flashes, 2724
- hyperamnesia, 2732
- hyperemesis gravidarum, 2724
- hypnotic induction ceremony, 2714–2715
  - eye fixation/closure, 2715
  - eye roll, 2715
  - holding on/letting go, 2715
  - progressive relaxation, 2715
  - purpose and preparation, 2714
  - staircase/elevator, 2715
  - visual imagery, 2715
- immune-related disorders, 2729–2730
- indirect, 2716
- irritable bowel syndrome, 2729
- memory hardening position and, 2733
- Nancy school of, 1867
- neurobiology, 2712–2713
- obstetrics, 2724
- overview, 2705–2706
- pain and, 2724–2729
- pain management using, 2377
- personality disorders and, 2723–2724
- pseudomemories, 2732
- as redistributed attentiveness, 2709
- relational disturbances, 2720–2721
- self, factitious disorder and, 1863
- sexual dysfunction, 2730
- social influence models for, 2710–2711
- sociocognitive model, 1897
- somatic symptom and related disorders, 2719–2720
  - case studies, 2720
- split-screen technique, 2718
  - case study, 2718–2719

- state definition, 2709
- state/nonstate controversy, 2709, 2711
- suggestibility and, 2708
- surgical procedures, 2725–2726, 2725*f*
- target event, 2734
- task motivational theory, 2710
- undue confidence and, 2733
- vocal cord dyskinesia, 2722
- weight management, 2731

- Hypnotherapy
  - affect regulation dial, 2717
  - automatic writing and drawing techniques, 2716
  - clenched-fist technique, 2716–2717, 2722
  - conflict identification/resolution and, 2717
  - conversion disorder, 1840
  - developmentally informed, 2716
  - directive, 2716
  - ego state therapy and, 2717
  - exploratory, 2716
  - in functional GI disorders, 2208
  - gut-focused, 2202
    - for IBD, 2201–2202
  - hypnoprojective techniques, 2716
  - ideomotor questioning technique, 2717
  - individualized, 2716
  - for irritable bowel syndrome, 2202
  - permissive, 2716
  - problem-solving, 2715
  - protocol-based, 2716
  - for sexual dysfunction, 1976–1977
  - supportive, 2716
  - treatment planning, 2715–2717
- Hypnotic hyperamnesia, 2732
  - lax response-criteria hypothesis, 2732
- Hypnotic induction plus manualized treatment protocol, 2716
- Hypnotic Induction Profile (HIP), 2714, 2717, 2724, 2727
- Hypnotic(s). *See also* Sedatives

- classification, 2918
- description of, 4585
- hypnotic-related disorders and, 1384–1386
- intoxication, 2630
- psychiatric emergency due to, 2630, 2632–2633
- respiratory depression and, 2225
- short-term, 2324
- side effects, 2225
- for stress management, 2324
- withdrawal, 2632–2633

- Hypnotizability
  - assessment, 2713–2714
  - children, 2707
  - dissociative disorders and, 1895*t*
  - double dissociators and, 1903
  - trait, 2706–2708
    - genetics of, 1889
- Hypoactive sexual desire, 1971
- Hypoactive sexual desire disorder, 4585
- Hypoactivity, 4585. *See also* Hypokinesia
- Hypoadrenalism, 682–683
- Hypochondria. *See also* Somatic symptom and related disorders
  - description of, 4586
  - somatic symptom and related disorders, 1828
- Hypochondriacal disorder, 1831
- Hypochondriacal neurosis. *See* Somatic symptom disorder
- Hypochondrial traits, parasitosis and, 2347
- Hypochondriasis. *See also* Illness anxiety disorder; Somatic symptom disorder
  - BDD and, 1792
  - child/adolescent
    - differential diagnoses, 3699
  - classification, 1786, 1827, 1828*f*, 1836
  - description of, 4586
  - gambling and, 1804–1805, 1808
  - OCD and, 3699
  - by proxy, 1849
- Hypochondriasis (Hs) Scale, 999*t*

- Hypocortisolism, PTSD and, 2324
- Hypocretin (orexin), 345–346
- Hypocretin-1 deficiency, 2092
- Hypocretin A, 36
- Hypocretin B, 36
- Hypocretins, synthesis of, 676–677
- Hypodopaminergia, 456
- Hypoglycemia
  - description of, 4586
  - prevention of, 174
- Hypokalemia in bulimia nervosa, 2075
- Hypokalemic periodic paralysis, 140
- Hypokinesia, 540
- Hypokinetic disorders, 2938
- Hypolipidemic medications, 2087
- Hypomagnesemia, 683
- Hypomania
  - after TBI, in children, 653
  - bipolar II disorder, 1650–1651, 1697
  - characteristics of, 574
  - classification of, 1168
  - corticosteroids and, 2263
  - course, 1651
  - cyclothymic disorder and, 1650
  - definition, 1600–1601
  - DSM-5 on, 1640*t*
  - mania *versus,* 1641
  - as reactive mania, 1649
  - recurrent brief, 1652
  - recurrent brief depressive disorder and, 1647
- Hypomanic episodes, 1399–1401, 1400*t*, 1401*f*
- Hypomyelination, 660
- Hyponatremia, 3035, 3157, 3344–3345
  - antidepressants and, 1680
  - in bulimia nervosa, 2075
  - SSRIs and, 1779
- Hyponea, 2093

- Hypoparathyroidism
  - characteristics of, 174
  - presentation of, 682
- Hypophysis. *See* Pituitary gland
- Hypopituitarism
  - hypothyroidism and, 2240
  - presentation of, 679
- Hyposexuality, 455
- Hyposomnia, 456, 1640
- Hypotension
  - antipsychotic drug effects, 4137–4138
  - DA receptor antagonists and, 3075, 3077
  - orthostatic
    - FGAs and, 3075
    - MAOIs and, 3118
    - tricyclic and tetracyclic antidepressant effects on, 3260–3261, 3262
- Hypothalamic attack area (HAA), 2476–2477
- Hypothalamic disorders, 676–677
- Hypothalamic-gonadal-pituitary axis, 1956
- Hypothalamic hypothyroidism, 2508
- Hypothalamic nuclei, 37*t*
- Hypothalamic-pituitary-gonadal (HPG) axis
  - endocrine events, 169
  - GnRH administration and, 169
- Hypothalamic-pituitary-thyroid (HPT) axis, **172–174**
  - TRH regulation of, 88, 90–91
- Hypothalamic–pituitary–adrenal (HPA)-axis
  - activation of, 59
  - anorexia nervosa and, 2074–2075
  - anxiety disorders and, 1743*t*, 1744
  - cytokine effects on, 197
  - depersonalization and, 1890
  - dissociative disorders and, 1889
  - dysregulation of, 92
    - depression and, 2188
  - elevated activity, stress and, 1715, 1716
  - fear conditioning, PTSD and, 1818

- gene-environment interactions
  - anxiety disorders, 1739, 1740
- immune function and, 2320
- inhibitory regulation of, 1712
- insecure attachment, suicide risk and, 1671
- PTSD and, 1819
- responses to stress, 2318
- stress responses and, 167–168, 193–194
- suicide and, 2619

- Hypothalamus
  - aggression and, 2476–2477
  - aggression mediated by, 2476–2477
  - anatomy of, 35–37
  - appetite and, 364–365, 364*f*
  - circadian rhythms and, 1714
  - description of, 4586
  - dysfunction, depression and, 1709
  - embryonic origins of, 8
  - functional circuitry, 37–38
  - hormone production by, 2239
  - lesions
    - obesity and, 677
    - rage reactions and, 677
    - thermoregulation and, 677
  - neurosurgery on
    - for paraphilias, 2020
  - noradrenergic neurons in, 1714
  - nuclei of, 36*f*
  - surface landmarks, 7*f*
  - tumors in childhood, 677
  - violence and, 2426–2427
- Hypothesis testing, 988. *See also* Statistics
- Hypothetical thought, 709
- Hypothyroidism
  - amiodarone and, 2198
  - causes of, 680
  - definition of, 2240

- depression and, 3242, 3243
- description of, 4586
- epidemiology, 2240, 2241
- etiology, 2240–2241
- hands in, 173*f*
- lithium and, 2240, 3097
- mood disorder and, 1249
- neuropsychiatric symptoms, 2241
- nosology, 2240
- physical symptoms of, 2241
- presentation of, 680
- prolactin levels and, 2249
- subclinical, 680, 2241
- symptoms of, 173

- Hypotonia, 3540
- Hypoxyphilia, 1975
- Hysterectomy
  - fertility control by, 2512–2513
  - MTF transitions, 2034
- Hysteria. *See also* Somatic symptom and related disorders
  - Breuer and Freud studies, 852–853
  - classification, 1835
  - description of, 4586
  - DSM and ICD diagnostic criteria, 1874
  - history, 1839
  - Janet and Freud on, 1868–1869
  - persistence of controversy about, 1873
  - somatic symptom and related disorders, 1828
  - somnambulism, multiple personality and, 1866
- *Hysteria, Hypnosis, and Healing* (Charcot), 4536
- Hysteria (Hy) Scale, MMPI, 999*t*
- Hysterical anesthesia, 4586
- Hysterical fainting, 2194
- Hysterical neurosis, conversion type, 1835
- Hysterical paralysis, 2178
- Hysterical psychosis
  - acute dissociative reactions to stressful events and, 1948

  - differential diagnoses, 1949–1950
  - pathology and laboratory examination, 1950
- Hysterical symptoms, sexually abused children, 3832
- Hysterics, description of, 4586
- *Hysterocrates gigas* (African tarantula) toxin, 139
- Hysteroid dysphoria, 1652

I

- "I Can Cope" program, 2255
- *I Never Promised You a Rose Garden,* 2872–2873, 2876
- Iatrogenesis
  - dissociative identity disorder, 1924
  - positive psychology, 2852
- Iatrogenic, definition of, 4586
- Iatrogenic model (IM), 1897–1899
- Iatropic factors for clinical evaluations, 3377
- Ibogaine and iboga alkaloids, 1317, 1320
- Ibuprofen, opioid withdrawal treatment, 1368
- Icarus Project, 2835
- ICD. *See International Classification of Disease* (ICD)
- ICD-10, mental disorder classification in
  - contents of, 1166, 1166*t*
  - DSM-IV, comparison with, 1167–1169, 1167*t*
  - frequency of use of, 1168–1169
  - revision of, 1169
- ICD-11, mental disorder classification in, **1169–1176**
  - clinical description in, 1169–1170
  - defined, 1169
  - dimensional approach in, 1170
  - DSM-5, comparisons with, 1174–1176
  - DSM-5, harmonization with, 1175–1176
  - field trials for, 1170–1171
  - overview, 1163–1164
  - PCH, 1171
  - proposals for, 1170, 1170*t*, 1171–1174
  - prototype matching in, 1170
  - qualifiers in, 1170
  - standardization of, 1170, 1171*t*
- "Ice." *See also* Methamphetamines
  - description of, 4586
- ICF. *See International Classification of Functioning, Disability, and Health*
- ICIDH. *See International Classification of Impairments, Disabilities, and Handicaps: A Manual of Classification Relating to the Consequences of Disease*
- ICIDH-2. *See International Classification of Impairments, Activities, and*

Participation: A Manual of Dimensions of Disablement and Functioning
- ICS. *See* Implanted cortical stimulation
- Ictal aggression
  - definition of, 2496
  - epilepsy and, 2496
- Ictal psychotic trigger reaction
  - about, 2496
  - aggression and, 2496
- Ictus
  - definition of, 507
  - features, 511–512
- Id
  - description of, 760, 4586
  - mood and, 1660–1661
  - Strachey's translation of Freud on, 2645
- *Id* transcription factor, 51
- Idealization, 4586
- Ideas of reference, 1122, 1414, 4586
- Ideas of unreality, 4586
- Ideational shield, 4586
- *Idée fixe,* 4586
- Identical twins. *See* Monozygotic twins
- Identification
  - with the aggressors, 4586
  - description of, 4586
  - mutative change and, 2664
  - in young children, 2870
- Identity
  - assessment of, 2140–2141
  - character, temperament and, 2136*t*
  - description of, 4586
  - development of, 2140–2141, 3365–3366
  - levels of, 2140
  - psyche and, 2139
  - scope of, 2139
  - temperament, character and, 2145
- *Identity: Youth and Crisis* (Erikson), 900, 901

- Identity crisis, 4586
- Identity disorders, 4586
- Identity disturbance, 2164
- Identity/identities. *See also* Dissociative identity disorder
  - partial or new, dissociative fugue and, 1880
- Identity *versus* role confusion, 3365–3366
- Ideomotor activity/exploration, 2711
- Ideomotor apraxia, 461
- Ideomotor questioning technique, 2717
- Ideomotor response, 2708
- Ideopathic, definition of, 4586
- Ideosensory activities, 2711
- Idiom of distress, cultural, 791, 807
- Idiopathic insomnia, 2085
- Idiopathic paradoxical vocal fold movement, 3567, 3569
- Idiopathic Parkinson disease (IPD). *See also* Parkinson disease
  - akathisia in, 459
  - characteristics of, 458
  - descriptions of, 542
- Idiot, definition of, 4586
- Idiotropic, definition of, 4586
- IED. *See* Intermittent explosive disorder (IED)
- Ifenprodil, structure of, 77*f*
- Ifosfamide (Ifex)
  - neuropsychiatric toxicity of, 2272
  - side effects, 2255
- Ignorance, constructive use of, 3386–3387
- I–It, description of, 4586
- Illegal migrants, global perspective, 797
- Illicit drugs, 1263, 4115. *See also specific drugs*
- *Illinois v. Dugan*, 4483
- Illness anorexia, appetite and, 367
- Illness anxiety disorder
  - BDD and, 1792
  - classification, 1786, 1827
  - comparative nosology, 1838
  - definition, 1838

- description of, 4586
- epidemiology, 1838
- factitious disorder and, 1859
- history, 1838
- introduction, 1837–1838
- OCD *versus,* 1788
- STAPP and, 2798
- STDP and, 2801
- treatment
  - CBT, 1839
  - pharmacotherapy, 1838
  - pivotal study, 1839
  - psychotherapy, 1839
  - reassurance, 1838

- Illness falsification, pediatric, 2408
- Illness Management and Recovery, 1553
- Illness/wellness management skill training, 4374–4375
- Illnesses. *See also* Chronic illnesses; Terminal illnesses
  - child/adolescent response to, **3845–3852**
    - age-related, 3848–3849
    - asthma, 3846
    - cancer, 3848
    - family's, 3849–3851
    - heart disease, 3847
    - hemophilia, 3847–3848
    - inflammatory bowel disease, 3847
    - to insulin-dependent diabetes mellitus (IDDM), 3847
    - overview, 3844*t*, 3845, 3845*t*
    - rheumatologic disorders, 3846–3847
  - cultural concept of, 789–790
  - management of
    - health promotion model of, 4303
    - during rehabilitation, 4368
  - social media tracking, 3852
  - stress and, 187, 189
  - as symptomatology consideration, 1116
  - transition from adolescents to adulthood, 3851, 3852*t*

- Illogical thinking, 4586
- Illogicality, 4586
- Illumination, 4586
- Illusions
  - of control, 2673
  - definition of, 4586
  - delirium and, 1186
  - as perception disturbance, 1125
- Illusory causation, 2673
- Iloperidone (Fanapt), side effects, 2237, 2237*t*
- Image analysis. *See specific imaging modalities*
- Imagery
  - cognitive therapy, 2771–2772
  - dissociative identity disorder treatment and, 1924
  - exposure therapy, 2691
  - hypnotic, 2708
- Imaginal desensitization, 2056
- Imaginal exposure
  - anxiety disorders and, 1772, 1774
  - cognitive therapy, 2772
- Imaginary companionship, 1902
- Imagination, symptomatology and, 1121
- Imago, 4587
- Imbecile, 4587
- IMD exclusion
  - 1965/1972 changes to, 4335–4336
  - as Medicaid requirement, 4289
- Imipramine (Tofranil)
  - agoraphobia, 1783
  - for alopecia, 2344
  - CBT for panic disorder and, 1775
  - chemistry, 3254, 3254*f*
  - child/adolescent
    - ADHD treatment, 3789
  - depression, 1681, 3119, 3120–3121
    - efficacy research, 2695
  - dosage/administration, 3264, 3265*t*

- dosages, 573*t*
- drug interactions, 3264
- drug studies, 3259
- encopresis, 3660–3661, 3674–3675
- enuresis, 3650, 3654, 3655–3657, 3664, 3668, 3669–3671, 3795
- GAD, 1783
- history, 2917, 3253–3254
- inflammation and, 201
- for insomnia, 2296
- for neuropathies, 629
- nocturnal enuresis, 3938
- organ/system effects of, 3260
- pain syndromes, 1841
- pharmacological actions, 3255–3256, 3258
- side effects, 573*t*
- sleep disorders, 3936
- for sleep enuresis, 2102
- targets for, 69
- therapeutic indications, 3261

- Imitation
  - in cognitive development, 707
  - description of, 4587
  - group psychotherapy and, 3761
- Imitative dissociative identity disorder, 1916
- Immediate memory, 4587
- Immigrant waiting period, Medicaid/CHIP, 4379
- Immigration
  - demographics of U.S. population, 794, 794*t*
  - global perspective, 796–797
  - psychiatric assessment, 793
  - as psychosocial factor, 4170
  - schizophrenia and, 1431, 1450–1451
- Immune cell receptors, 184
- Immune responses
  - acute stress and, 2338–2339
  - atopic dermatitis and, 2341
  - DA receptor agonist effects on, 3239

- regulation of, 183–184
- Immune systems
  - adaptive immunity, 179
  - cancer outcomes and, 2286
  - cannabis use and, 1309–1310
  - cell differentiation and, 183*t*
  - cell growth and, 183*t*
  - cells/tissues, 179
  - CNS interactions with, **178–184,** 194–197, 2319–2320
  - HIV and, 4273
  - innate immunity, 179–181
    - acquired immunity *versus,* 180*t*
  - mechanisms, 193
  - overview of, **178–184**
  - stress-immune spectrum, 2336
  - stress physiology and, 2318–2322
- Immune triad, brain-cytokine link, 2338
- Immunoglobulin E, 2341
- Immunomodulation, brain-cytokine link, 2338–2339
- Immunopathological responses, 2338
- Immunoprotective responses, 2338
- Immunoregulatory/suppressive responses, 2338
- Immunosuppressants
  - cognitive impairment and, 2367–2369
  - neurotoxicity of, 2371–2373, 2372*t*
  - in transplant populations, 2367–2369
- Immunosuppression. *See also* HIV (human immunodeficiency virus); Transplantation
  - chronic stress and, 2339
  - conditioned, 187*f*
  - major depression and, 198
- Immunotherapy
  - autism spectrum disorder and, 3579
  - dementia related to, 2284
  - for stimulant use disorders, 1290
  - for tic disorders, 3638, 3652
- Imovane. *See* Zopiclone
- Impact messages, 2775–2776

- Impact phase, after disasters, 2566
- Impaired insight, 4587
- Impaired judgement, 4587
- Impairments
  - cognitive, 1193–1194, 1193*t*
  - defined, 3484
  - functional, 1194, 1194*t*
  - morbid mood state, 1633–1634
  - psychiatric symptomatology affected by, 1118
  - on psychopathology continuum, 3484
  - rating scales for, 3402
  - in self-esteem regulation, depression and, 1662
- Impersistence of body postures, 3540
- Implanted cortical stimulation (ICS), 3299, 3312, 3312*f*
- Implementation variables, behavior therapy and, 2701
- Implicit mentalizing, explicit *versus,* 2895
- Implosion, 4587
- Impotence. *See also* Male erectile disorder
  - description of, 4587
- Imprinting
  - description of, 4587
  - heritable change and, 161
- *Improving the Quality of Health Care for Mental and Substance Use Conditions,* 4315
- Improvisation, 4587
- Impulse control
  - description of, 4587
  - impaired, 376
    - in ND-PAE, 637–638
  - kleptomania and, 2114
- Impulse-control disorders (ICD), **2110–2115**
  - diagnosis of, 2436–2437
  - with intellectual disability, 3511
  - in PD, 545–546
  - treatment
    - carbamazepine, 3037
    - FGAs and, 3071
    - valproate, 3275

- trichotillomania (TTM), 2352–2353
- world-wide, 4517
- Impulses, definition of, 4587
- Impulsion, 4587
- Impulsive behaviors, 1640
- Impulsive/careless problem solving, 2696
- Impulsive judgment, 1124
- Impulsivity
  - on ADHD, 3590
  - alcohol and, 1269
  - animal models of, 322
  - suicide behavior risk factor, 3683, 3684, 3697, 3698
- Imputation of data, 839
- In aging adults, 4007
- In-home services, for pediatric treatment, 3802
- *In re Maida Yetter,* 4434
- *In re Oakes,* 4434
- In situ hybridization, 85
- In vitro fertilization with embryo transfer (IVF-ET), 2509, 2510*t*
- In vivo amplified skills training (IVAST), 4374
- In vivo desensitization, 2690
- In vivo exposures
  - anxiety disorders and, 1772, 1774
  - cognitive restructuring and, 1775
  - exposure therapy, 2691
  - OCD and, 1789
- Inactivation, ion channels, 133
- Inappropriate affect, 4587
- Inapsine. *See* Droperidol
- Inattention/Overactivity with Aggression (IOWA) Conners Scale, 3401
- Inborn errors of metabolism (IEMs)
  - CNS disorders in, 655
  - intellectual disability and, 3501–3502
  - laboratory evaluation, 1016
  - mitochondrial disorders, 669–671
- Incarcerated individuals
  - access to care, 2037

- gamblers, 1811
- transgender, 2037
- Incarceration
  - decarceration *vs.,* 4400
  - as hospital alternative, 4325, 4329
  - recidivism, 4340, 4392
  - seriously mentally ill in, 4305, 4388
- Incentive learning, 725, 726, 726*f*
- Incentive sensitization theory, 372–373
- Incest
  - description of, 4587
  - as maltreatment, 3832–3833
- Incidence
  - description of, 4587
  - in epidemiology, 812
  - rates, 4587
- Incidentalomas, 284
- Inclusion body myositis (IBM), 634
- Incoherence, 4587. *See also* Word salad
- Incompetence, 4587
- Incorporation, 4587
- Increased libido, 4587
- IncRNas
  - role of, 2138
- IND. *See* Investigational new drug
- Indels
  - definition, 155
  - in DNA copy number variations, 3449
- Independent components analysis (ICA), 421
- Independent living donor advocate (ILDA), 2362
- Independent Living Skills-Survey (ILSS), 4372
- Independent variables, 4587
- Inderal. *See* Propranolol
- Indifference reaction, brain injury and, 475
- Indigenous Americans, mental health of, 800–801. *See also* Race/ethnicity
- Indiplon, 3003
- Indirect exposure to traumatic events, 1813*t*, 1814

- Indirect hypnosis, 2716
- Individual placement and support (IPS) model, 1534–1535
- Individual psychology, description of, 4587
- Individual social interactions, 2820
- Individual therapy
  - DBT, 2788–2789
  - description of, 4587
  - group psychotherapy combined with, 2745
- Individualized disclosure, 2290–2291
- Individualized Education Program (IEP), 3517
- Individuals, cultural identity of, 2463
- Individuals with Disabilities Education Act
  - children with TBI and, 653
- Individuals with Disabilities Education Act (IDEA), 3517
  - child cognitive assessments and, 3414, 3415, 3417, 3419–3421*t*
- Individuation, 4587. *See also* Actualization
- Indoklon therapy, 4587
- Indolamine, 4587
- Indoleamine 2,3 dioxygenase (IDO), 196
- Induced psychotic disorder. *See* Shared psychotic disorder
- Inducible nitric oxide synthase (iNOS), 105
- Industrial psychiatry, 4587
- Industry consolidation, 4313–4314
- Ineffability, description of, 4587
- Infant Brain Imaging Study (IBIS), 3576
- Infant-mother bonding, 3709
- Infanticide risk, 2516
- "Infantile amnesia," 60
- Infantile anorexia, **3623–3625, 3637–3639**
  - clinical features, 3623, 3637
  - course, 3623–3624, 3637–3638
  - diagnosis, 3623, 3637
  - differential diagnoses, 3623, 3627, 3637, 3641
  - etiology, 3623–3624, 3637–3638
  - treatment, 3624–3625, 3638–3639
- Infantile articulation. *See* Developmental articulation disorder
- Infantile autism, 3573, 4587. *See also* Autistic thinking

- Infantile sexuality, 4587
- Infants and children, developmental model (Piaget), 713–716
- Infants/infancy. *See also* Children; Neonates; Preterm infants
  - attachment during, 3662, 3676
  - caregiving deficiencies, 2047–2048
  - cutaneous stimulation in, 2334
  - description of, 4587
  - development
    - brain, 3351, 3358
    - cognitive, 3359
    - emotional, 3361
    - of gender identity, 3363
    - of language, 3358
    - motor, 3361–3362
    - overview, 3352*f*, 3354*f*, 3356*f*
    - psychosocial, 3363
    - sensory, 3362–3363
    - social, 3360–3361
  - end-of-life care, 2307, 2312
  - evaluation considerations for, 3387–3391, 3389*t*, 3390*t*, 3392–3393*t*
    - diagnostic formulations, 3391
    - overview, 3387–3388
    - standardized assessment instruments, 3389, 3390*t*, 3391
    - technical issues, 3388–3389
  - hospitalizations, response to, 3848
  - illnesses, response to, 3848
  - infantile anorexia, **3623–3625, 3637–3639**
  - primary central sleep apnea of, 2097
  - psychiatry
    - description of, 4587
  - quiet sleep, 339
  - sexuality, 863–864
  - skin-to-skin contact
    - benefits of, 2052
  - sleep in, 342
- Infections
  - opportunistic, 645–646

- *in utero* exposure to, 199
- Infectious diseases. *See also specific diseases*
  - of central nervous system, 1011
  - cytokines and, 2321–2322
  - evolution of, 770
  - intravenous drug use, 1011
  - neuropsychiatric aspects of, **588–600**
  - patient assessment in, 472–473
  - prenatal, 599
  - schizophrenia comorbidity, 1540–1541
  - sexually transmitted diseases, 1011–1012
    - chlamydia, 1012
    - gonorrhea, 1012
    - HIV, 1012
    - syphilis, 1012
  - tics triggered by, 3638, 3652
  - treatment, benzodiazepine, 2998
  - viral hepatitis, 1012–1013
- Inferences, 3552
- Inferential statistics, 4587
- Inferior frontal gyrus (Broca's area)
  - cross-section, 9*f*
  - lesions in, 7, 653
- Inferior longitudinal fasciculus, 12
- Inferior occiptofrontal fasciculus, 12, 15*f*
- Inferiority complex, 910, 4588
- Inferolateral prefrontal cortex, 1893
- Infertility
  - causes of, 2509
  - fertility and, 2508–2515
  - interventions, 2509–2510, 2510*t*
  - lifestyle factors, 2509
  - psychogenic factors, 2509
  - psychological reactions to, 2510
  - tests in workup for, 2509*t*
- Inflammation
  - Alzheimer disease risk factor, 1205

- and clinical disorders of aging brain, 3989
- delirium and, 1181–1182
- demyelination process and, 659
- factors in, 182*f*
- gut–brain axis and, 2204
- microbes and, 192–193
- pain transmission and, 2388–2389
- positive psychiatry, 4284
- social adversity and, 189*f*

- Inflammatory bowel disease (IBD)
  - case studies, 2208
  - laboratory evaluation, 1015
  - pediatric, response to, 3847
  - presentation of, 2201
  - psychological therapies in, 2201–2202
  - stress of, 2201
- Inflammatory demyelinating polyneuropathy, 632
- Inflammatory diseases, assessment in, 472–473
- Inflexibility, character and, 2138
- Infliximab, 1681
- Influence, undue, 4234–4235
- Influenza A virus, 190–191
- Influenza viruses, 599
- Informal admission, 4588
- Informant-Cultural Formulation Interviews (I-CFI), 2468–2469
- Informants
  - for child psychiatric evaluations, 3376, 3410–3411
  - differences among, 3379–3380
  - for infant and toddler evaluation, 3388–3389
  - interview development for, 3394
  - multiple, 3753
  - for preschool evaluation, 3410–3411
- Information
  - CAM therapies based on, 2558–2561
  - exchange, in disparities reduction, 4386
  - forms not validated by Western science, 2561–2564
  - forms validated by Western science, 2546, 2558–2561

- processing, 689, 689f (*See also* Attention)
  - hypnotizability and, 2710
  - modes of, 453
  - neuropsychological testing, 990
- sharing of
  - patient rights and, 2291
- Information Processing Theory, 3360
- Information technology. *See also* Email, psychotherapy and; Internet
  - ethical issues with, 3875–3876
  - in jails and prisons
    - accreditation and standards, 4462
    - inmate workers, 4463
    - peer-to-peer programs, 4463
    - quality assurance, 4462
    - staffing, 4462–4463
- Information theory, 4588
- Informed consent. *See* Consent, informed
- Informed refusals, 2290
- Infra-low frequencies (ILF) EEG biofeedback, 2559
- Infradian rhythms. *See* Biological rhythms
- Infusions, description of, 4588
- Inhalant intoxication, 2632
- Inhalant use, **1328–1342**
  - amyl, butyl, and isobutyl nitrite-related disorders, 1341
  - anxiety disorder, inhalant-induced, 1340
  - comorbidity, 1331
  - comparative nosology, 1328
  - deaths related to, 1330–1331
  - defined, 1328
  - depressive disorder, inhalant-induced, 1340
  - emergency department visits for, 1329–1330
  - epidemiology, 1328–1329
  - etiology, 1334–1335
  - future directions, 1341–1342
  - hospital admissions for, 1330
  - inhalant use disorder, 1329, 1335–1337
    - course and prognosis, 1336

        - diagnosis and clinical features, 1335
        - differential diagnoses, 1335–1336
        - laboratory examinations, 1337
        - treatment, 1336–1337
    - intoxication, 1337–1338
        - course and prognosis, 1338
        - delirium, 1338
        - diagnosis and clinical features, 1337
        - differential diagnoses, 1338
        - laboratory examinations, 1337–1338
        - treatment, 1338
    - mortality, 1330–1331, 1333
    - neurocognitive disorders, inhalant-induced, 1338–1339
        - course and prognosis, 1339
        - diagnosis and clinical features, 1338–1339
        - differential diagnoses, 1339
        - treatment, 1339
    - nitrous oxide-related disorders, 1340–1341
    - organ pathology, 1333–1334
    - overview, 1328
    - pharmacology and toxicology, 1331–1333
    - psychotic disorders, inhalant-induced, 1339–1340
    - tolerance, 1336
    - unspecified, 1340
- Inhalants
    - description of, 4588
    - EEG effects, 254
- Inheritance
    - gene mapping and, 304
    - Mendel laws of, 149
    - Mendelian patterns, 304
    - polygenic, 304
- Inhibited female orgasm, 4588. *See also* Anorgasmia
- Inhibited male orgasm. *See also* Primary retarded ejaculation; Secondary retarded ejaculation
    - description of, 4588
- Inhibited sexual excitement, 4588
- Inhibition, definition of, 4588

- *Inhibitions, Symptoms, and Anxiety* (Freud), 867
- Inhibitory, definition of, 141
- Inhibitory learning, 1771
- Initial Pain Assessment Tool, 2376
- Initial Treatment and Recovery Plan (ITRP), 4337
- Initiator elements (INRs), promoter, 159
- Injuries, gun control and, 2476
- Injurious behavior, 3672, 3686
- Inmate abuse, 4473
- Innate immunity
  - acquired immunity *versus,* 180*t*
  - components of, 180*f*
  - cytokine mediators of, 182*t*, 198–199
  - stress and, 2318
- Innovation, policies supporting, 4380–4381
- Inositol
  - anxiety disorders, 1782
  - beneficial effects, 2548
- Inositol triphosphate (IP_3)
  - lithium and, 122
  - second messenger activity of, 88–90
- Inpatient medicine service, 4196
- Inpatient studies, dissociative disorders, 1888
- Inpatient treatment
  - child/adolescent, **3797–3801,** 3815
    - admission criteria, 3799–3800
    - aggression management, 3799
    - assessment, 3800, 3800*t*
    - care levels, 3798, 3798*t*
    - discharge criteria, 3799–3800
    - overview, 3797
    - partial hospital programs, 3801
    - physical structure, 3798, 3799
    - schedule, daily, 3798, 3798*t*
    - staffing, 3798–3799, 3800, 3800*t*
    - treatment planning, 3800–3801
  - hospitalist role in, 4358

- intermediate care facilities for, 4400
- as managed care driver, 4313
- percentage decrease in, 4325

- Insane, definition of, 4588
- Insanity
  - American Law Institute formulation of, 4560
  - definition of, 4588
- Insanity defense, 4439, 4588
- Insecurity, definition of, 4588
- Insel, Thomas
  - advances due to, 1
  - on psychiatric classification, 1162
  - on somatic presentations, 1844
- Insight
  - absent, BDD and, 1790
  - absent, OCD and, 1785–1786, 1788
  - BDD and, 1791
  - defined, 1124, 3750
  - description of, 4588
  - hoarding disorder and, 1793
  - lack of, manic syndrome and, 1641
  - measured by self-transcendence, 2143
  - mutative change and, 2663–2664
  - OCD and, 1785*t*
  - psychiatric symptomatology and, 1124
  - psychoanalysis and, 2647
  - punctiform, psychosis and, 2878–2879
  - schizophrenia and, 1423
- Insight and judgment in patient interview, 967
- *Insight and Responsibility* (Erikson), 900
- Insight-oriented interpersonal psychotherapies (IO-IPT)
  - about, 2775
  - phases, 2776
- Insight-oriented psychotherapy, 4175–4176
  - aggression, 2498, 2498*t*
    - TBI, 2502
  - description of, 4588

- intimate partner violence, 2502
- Insight therapy. *See* Insight-oriented psychotherapy
- Insomnia. *See also* Fatal familial insomnia (FFI)
  - after disasters, terrorism, or war, 2567*t*
  - in aging adult, 4053–4054
  - animal models of, 324
  - in cancer patients, 2253
  - comorbid with mental disorder, 2086–2087
  - depressive syndrome and, 1638
  - in dying patients, 2295–2296
  - effects of, 349–350
  - in long-term care, 4230–4231
  - MAOIs and, 3118
  - pediatric, 3935–3936
  - periodic, 2098
  - short sleepers, 2088
  - subtypes, 2088
  - suicide behavior risk factor, 3684, 3698
  - trazodone drug studies, 3250
  - treatment
    - acupuncture, 2562
    - anticonvulsants, 2973
    - antihistamines, 2982
    - barbiturates, 2985–2986
    - behavioral, 2088
    - benzodiazepine, 2996–2997, 3002
    - doxepin, 2982
    - melatonin, 3101, 3103–3104, 3105
    - nonbarbiturate sedative-hypnotic agents, 2989
    - pharmacological, 2088
    - trazodone, 3252
  - without objective findings, 2085
- Insomnia disorder
  - characteristics of, 2083–2090
  - description of, 4588
  - diagnostic criteria, 2084–2085
  - subtypes, 2085–2088

    - behavioral insomnia of childhood, 2086
    - comorbid with medical conditions, 2087
    - comorbid with mental disorder, 2086–2087
    - idiopathic, 2085
    - inadequate sleep hygiene, 2085–2086
    - paradoxical, 2085
    - psychophysiological, 2085
    - substance abuse related, 2087–2088
- Instinct, definition of, 4588
- Instincts, neuroimaging studies of, 2814–2815
- Instinctual theory (Freud), 862–865
  - characteristics of
    - aggression and ego, 864
    - infant sexuality, 863–864
    - libido, 863
    - part, 864
  - concepts of, 862–863
- Institute for Healthcare Improvement
  - chronic illness care model for, 4322
  - on Triple Aim, 4310
- Institute of Medicine (IOM)
  - care recommendations from, *4316t*
  - *Clinical Practice Guidelines We Can Trust,* 971
  - *Crossing the Quality Chasm: A New Health System for the 21st century,* 4289, 4294, 4314–4315
  - disparities, definition of, 4377
  - *Finding What Works in Health Care: Standards for Systematic Reviews,* 971
  - *Graduate Medical Education That Meets the Nation's Health Needs,* 4319
  - *Improving the Quality of Health Care for Mental and Substance Use Conditions,* 4315
  - *Knowing What Works in Health Care: A Road Map for the Nation,* 971
  - practice guidelines, 971, 971*t*
  - *Reform of Financing and Governance of Graduate Medical Education,* 4307
  - *Unequal Treatment* report, 4379
- Institutional racism, 4381
- Institutional Review Boards, 2312
- Institutionalization, 4588
- Instrument stream of neuroimages, 4479

- Instrumental actions, 726–727, 726*f*
- Instrumental conditioning. *See* Conditioning
- Instrumental learning, 416
- Insula (insular cortex)
  - fear conditioning, PTSD and, 1818
  - lesions to, 766
  - neuroimaging
    - anxiety, 1747
    - fear, 1747
    - specific phobias, 1762, 1763
    - threat detection and emotion regulation, 1819
  - PTSD and gray matter reductions in, 1820
- Insulin, **176**
  - resistance to
    - diabetes and, 2230
    - polycystic ovary syndrome and, 2248–2249
    - with schizophrenia, 1543
- Insulin coma therapy (ICT), 4588
- Insulin-dependent diabetes mellitus (IDDM), 3847
- Insulin-like growth factor, 186*t*
- Insulinomas, 684
- Insurance. *See also* Affordable Care Act (2010); Finances; Health insurance; Health maintenance organizations (HMOs); Medicaid; Medicare
  - under ACA, 4318
  - mental illness as barrier for, 4530
  - mental illness exclusion, 4528
  - parity
    - for mental health care, 4289, 4350, 4380
  - reimbursement rates as barrier, 4302
- Intake, definition of, 4588
- Integration
  - definition of, 4588
  - dissociative amnesia treatment and, 1938
  - dissociative identity disorder treatment and, 1926
  - and migration, 792
  - psychodynamic psychotherapy, 2059–2061
- Integration of neurons, 687–688
- Integrative family and couples therapy, 2751*t*

- Integrity, and mental health, 759
- Intellectual assessment of adults. *See* Clinical neuropsychology and intellectual assessment of adults
- Intellectual assessments
  - autism spectrum disorder and, 3577, 3581–3582
  - of children, 3415, 3416*t*, 3422–3425
  - cognitive dysfunction, 1568
  - conduct disorder and, 3615
  - intellectual disabilities and, 3493–3494, 3494*t*, 3495*t*, 3506–3507, 3509, 3510
  - IQ and, 3577
  - prediction for, 3391
  - schizophrenia, 1504–1505
  - specific learning disorder and, 3523, 3526, 3533
- Intellectual development disability, 4468–4469
- Intellectual developmental disorders. *See* Intellectual disability (ID)
- Intellectual disabilities (ID), 3491–3519. *See also* Cognitive function
  - adaptive functioning, 3494–3496, 3495*t*, 3496*t*
  - ADHD incidence with, 3506, 3511
  - anxiety disorders incidence with, 3511–3512
  - attachment disorders and, 3666, 3680
  - autism and, 3506
  - child, laboratory testing and, 3408
  - classification of, 1173
  - defined, 3493
  - definition of, 4588
  - deinstitutionalization, 3492
  - developmental considerations, 3508–3513
  - diagnosis
    - approach, 3500–3502, 3501*t*
    - clinical history and exam, 3500, 3501*t*
    - criteria, 3493–3496, 3494*t*, 3495*t*, 3496*t*
    - genetic, 3499–3501
    - instruments, 3513
    - metabolic testing, 3501–3502
  - differential diagnoses, 3533, 3548, 3591, 3610
  - Down syndrome (DS), 3507
  - eating disorder incidence with, 3512
  - emergency psychiatric evaluation for, 3445

- encopresis and, 3660, 3674
- etiology, 3499–3500, 3500*t*
- fragile X syndrome, 3502, 3506
- genetic syndromes and phenotypes, 3502–3508, 3503–3505*t*
- genetic testing for, 3456
- historical background, 3491–3493
- impulse control disorders incidence with, 3511
- intellectual functioning, 3493–3394, 3494*t*
- maladaptive behavior and, 3512–3513
- mental disorders incidence with, 3510–3513, 3510*t*
- mood disorders incidence with, 3512
- oppositional defiant disorder/conduct disorder incidence with, 3511
- organic mental disorders incidence with, 3512
- organizations, 3493
- pica and, 3632, 3646
- Prader–Willi syndrome and, 3506–3507
- prevalence, 3499
- prevention, 3513
- prognosis, 3500
- psychosis incidence with, 3512
- research, 3492–3493
- services for, 3516–3518, 3516*t*
- severity classification, 3495*t*, 3496–3499
- Smith–Lemli–Opitz syndrome, 3508
- Smith–Magenis syndrome, 3508
- stereotypic movement disorders, 3671, 3685
- supports for, 3518–3519, 3518*f*
- treatments, 3513–3516
  - pharmacotherapy, 3514–3516
  - psychotherapy, 3513–3514
- Williams syndrome, 3507–3508

- Intellectual disability (ID)
  - definition of, 4588
  - malingering and, 2407
- Intellectual functioning. *See also* Cognitive function
  - borderline, 2430–2442
  - description of, 2431

- Intellectual insight, 4588
- Intellectual subaverage functioning, 4588
- Intellectualization, 4588
- Intelligence
  - definition of, 4588
  - emotional basis of, 711–713
  - and mental health, 762
  - neuropsychological testing, 990
  - Piagetian model, 709–711
- Intelligence quotient (I.Q.), 4588
  - brain injuries and, 654
  - effects of prematurity, 643
  - malingering and, 2407
- Intensive care units (ICUs), diaries, 2331
- Intensive short-term dynamic psychotherapies (ISTDPs), **2796–2816**. *See also* Short-term dynamic psychotherapies
  - child/adolescent, **3741–3749**
    - for anxiety disorders, 3742, 3745–3746
    - cognitive-behavioral therapy, 3742–3745
    - for depression, 3742, 3747
    - for disruptive behavior disorders, 3743–3744, 3747–3748
    - for internalizing disorder, 3742–3743
    - overview, 3741–3742
    - psychosocial approach, 3742, 3745–3748
    - time-limited treatments, 3744–3748
  - clinical spectrum, 2797–2804
  - Davanloo's model, 2801–2804
    - case study, 2803
    - clinical issues, 2802
    - complications, 2803
    - contraindications, 2803
    - limitations, 2802–2803
    - outcomes and research, 2803–2804
    - techniques, 2802
    - theoretical issues, 2801–2802
    - treatment goals, 2803
  - history, 2796–2797
  - introduction and definition, 2796

- Malan's short-term dynamic, 2800–2801
- medications and, 2811–2812
- process-outcome correlations, 2812
- Sifneos' short-term anxiety-provoking, 2797–2800

- Intensivists, ICU role of, 4355
- Intent-to-harm, aggression and, 2471
- Intention, freedom of, 2140
- Inter-identity amnesia, 1910–1911
- Inter-rater reliability, description of, 3394–3395
- Intercourse, 1969–1970. *See also* Coitus
- Interests, child psychiatric assessments, 3378
- Interferon-α (IFN-α)
  - depression and, 201
  - dopamine metabolism and, 196
- Interferon-γ (IFN-γ)
  - secretion of, 181
  - skin cancer and, 190
- Interferons
  - effects of, 182*t*
  - encephalopathy and, 2273
  - side effects
    - depression, 2260
  - targets of, 182*t*
- Interictal aggression
  - definition of, 2496
  - epilepsy and, 2496–2497
  - presentation of, 2496–2497
- Interictal epileptiform discharges (IEDs), 244
- Interictal features, 513–514
- Interictal period, 507
- Interictal personality syndrome, 471
- Interictal psychosis, 649–650
- Interkinetic nuclear migration, 42*f*
- Interleukins
  - immune responses and, 2319, 2320
  - interleukin 1 (IL-1)
    - cardiovascular disease and, 2321

        - depression, 1681
        - effects of, 182*t*
        - mood disorders, 1717–1718
        - serotonin reuptake and, 196
        - targets of, 182*t*
        - during viral infections, 195*f*
    - interleukin 2 (IL-2, Proleukin)
        - dopamine metabolism and, 196
        - side effects of, 2273
    - interleukin 6 (IL-6)
        - cardiovascular disease and, 2321
        - in depression, 191*f*, 1681
        - effects of, 182*t*
        - innate immunity and, 180–181
        - targets of, 182*t*
        - during viral infections, 195*f*
    - interleukin 10 (IL-10)
        - effects of, 182*t*
        - targets of, 182*t*
    - interleukin 12 (IL-12)
        - effects of, 182*t*
        - targets of, 182*t*
    - interleukin 15 (IL-15)
        - effects of, 182*t*
        - targets of, 182*t*
    - interleukin 18 (IL-18)
        - effects of, 182*t*
        - targets of, 182*t*
    - secretion of, 181
- Intermediate care facilities, 4400
- Intermission, definition of, 4588
- Intermittent explosive disorder (IED), **2110–2113**
    - aggression and, 2494–2495
        - treatment of, 2501
    - conduct disorder and, 3619
    - description of, 4588–4589
    - diagnosis of, 2110

- epidemiology, 2110
- etiology, 2110–2111
- oppositional defiant disorder and, 3619
- Tourette disorder and, 3642, 3656
- treatment, 2111–2112

- Intermittent mood disorders
  - about, 1601
- Intermittent photic stimulation (IPS), 241
- Intermittent reinforcement, 2872, 4589
- Internal capsule, 11
  - anatomy of, 11
  - anterior limb, anatomy of, 10–11
  - auditory radiation, 11
  - genu, 11
  - horizontal section, 14*f*
  - optic radiation, 11
  - posterior limb, 11
  - projection fibers through, 10
  - regions of, 10–11
  - in schizophrenia, 13
  - sublenticular limb, 11
- Internal consistency, 3395
- Internal medullary lamina, 21
- Internal object relationships
  - of early life experiences, 2870
- Internal validity
  - behavior therapy research, 2704
  - in research, 4347, 4348–4349, 4351, 4353
- Internalization, psychodynamic psychotherapy, 2061–2062
- Internalizing disorders. *See also* Anxiety disorders; Mood disorder(s)
  - psychotherapies for, 3742–3743
- Internalizing symptomatology scales, 3399–3402
- International adoption, 3826
- *International Classification of Diseases* (ICD), 2026
  - case identification, 819–820
  - delirium, 4087
  - description of, 4589

- 10th rev. (ICD-10), 985
  - dissociative disorders, 1874
  - factitious disorder, 1848
  - mental disorders in older people, 3962
  - mood disorders defined in, 1599, 1631
  - schizophrenia, 4095
  - somatization disorder, 1831
- 11th rev. (ICD-11)
  - on excoriation disorder, 1796
  - on obsessive-compulsive and related disorders, 1785, 1786
- *International Classification of Functioning, Disability, and Health* (ICF), 819–820
- *International Classification of Impairments, Activities, and Participation: A Manual of Dimensions of Disablement and Functioning* (ICIDH-2), 819
- *International Classification of Impairments, Disabilities, and Handicaps: A Manual of Classification Relating to the Consequences of Disease (ICIDH),* 819
- *International Code of Medical Ethics,* 4241, 4242
- International Federation of Electroencephalography, 1027
- International HapMap Consortium, 147
- International List of Causes of Death (ILCD), 1165
- International organizations, 4531
- International Pilot Study of Schizophrenia (IPSS)
  - psychiatric classification and, 1153–1154, 1162
  - schizophrenia, 1408, 1409, 1411, 1427, 1432
- International Psychoanalytic Association, 2659
- International Society for Bipolar Disorders (ISBD), 1684, 1689, 1692
- International Society for the Study of Trauma and Dissociation (ISSTD)
  - on boundary violations and treatment frame, 1923
  - on containment, grounding, and imagery, 1924
  - on dissociation and dissociative disorders, 1871, 1873
  - on dissociative identity disorder treatment, 1918
- International Study of Schizophrenia (ISoS), 1432–1433, 1433*t*
- Internet. *See also* Computers; Smartphone apps
  - addiction to, 3373, 3768, 3812
  - behavior therapy and, 2701
  - cognitive-behavioral therapy with
    - terrorist attacks and, 2572
  - dissociative disorders and, 1873
  - gambling, 1810

- positive psychology and, 2852–2853
- psychotherapeutic tools
  - CBT, 1776
  - habit reversal therapy and stimulant control, 1795
  - Trauma TIPS, PTSD, 1824
- psychotherapeutic use of, 2253

- Interneurons
  - density in schizophrenia, 55
  - function of, 3, 17
  - GABAergic, 82–83
- Internships, description of, 4589
- Interobserver reliability, 4589
- Interoceptive exposure
  - anxiety disorders and, 1772–1773, 2722
  - CBT and, 1775
- Interpersonal and social rhythm therapy (IPSRT)
  - about, 1701
  - bipolar disorders, 1684, 1703, 1703*t*, 1706, 1707
- Interpersonal conflict. *See* Extrapsychic conflicts
- Interpersonal discrimination exercises, 2700
- Interpersonal effectiveness skills, 2055, 2788
- Interpersonal group (IPG) psychotherapy, 2746
- Interpersonal interpretive technique, 2876
- Interpersonal learning, 3761
- Interpersonal psychiatry, 4589
- Interpersonal psychotherapy (IPT), **2775–2784**
  - for adolescents, 3814
  - for anorexia nervosa, 2080
  - applying research to clinical practice, 1703, 1704
  - for child/adolescent, 3742
  - clinical issues, 2782–2783
    - complications, 2783
    - contraindications, 2783
    - indications, 2782–2783
    - limitations, 2783
    - treatment goals, 2783
  - depression, 1702

- depressive disorders, child/adolescent, 3679–3680, 3693–3694
- dysthymic disorder, 1666
- for eating disorders, 2220
- efficacy research, 2695
- encourage maintenance of gains, 2780
- ethical issues, 2783
- evaluation, 2863
- future directions, 2784
- grief and, 2778, 2780
- group, 2738, 2781–2782
  - group stage development and, 2782
  - maintaining interpersonal focus in, 2782, 2783*t*
  - techniques, 2781–2782
- history, 1701, 2775–2776
- initial phase
  - diagnosis, 2777
  - establish problem areas, 2778
  - group, 2781–2782
  - interpersonal inventory, 2777–2778
  - sick role assignment, 2777
  - treatment goals, 2778
- intermediate phase
  - facilitate work on identified goals, 2779–2780
  - group, 2782
  - problem areas in, 2778–2779, 2779*t*
- interpersonal deficits and, 2778–2779
- MDD, 1702*t*
- neurobiological effects of, 2814*t*
- paroxetine for depression *versus,* 2813
- progress review, 2780
- PTSD, 1823
- research and evaluation
  - group IPT, 2783
  - outcome research, 2783
- role disputes and, 2778
- role transitions and, 2778
- suicide risk, 2617

- techniques, 2777–2781
- termination phase, 2780
  - group, 2782
- therapeutic stance, 2777, 2781
- therapeutic techniques, general, 2780–2781
  - clarification, 2780
  - communication analysis, 2780
  - encouraging affect, 2780
  - exploratory questions, 2780
  - therapeutic relationship and, 2780–2781
- time-limited, 2776–2777
  - other therapies compared with, 2776–2777
  - phases, 2777*t*
- timeline, 2777–2782

- Interpersonal relationships
  - in psychiatric illnesses, 1144–1146
    - dysnosognosia and, 1146
    - importance of, 1145–1146
    - personality traits and disorders, 1144–1145
- Interpersonal skills, 4589
  - enhancement of, 2694–2695
  - training in, 1774
- Interpersonal social rhythms therapy (IPSRT)
  - bipolar depression, 1692
  - bipolar disorder, 1666
  - bipolar I disorder, 1696
- Interpersonal theory, depression, 1665–1667
- *Interpersonal Theory of Psychiatry* (Sullivan), 4540
- Interpersonal therapies, in GI disorders, 2207
- Interpretation (interpretive process)
  - in analytic process, 896
  - description of, 4589
  - guidelines, 2656
  - in psychoanalysis, 2655–2656
- *Interpretation of Dreams* (Freud), 4537
  - associative technique, 854, 859
  - description of, 4589

- Interpreter services
  - in HHS standards, 4380
  - outcomes of, 4381
  - for patient interviews, 963
- Interrogation by neuroimaging, 4489–4490
- Interrogatory Suggestibility (IS), 1899
- Intersex conditions
  - definition of, 2023
  - prenatal hormones and, 1992
  - treatment of individuals with, 2035
- Intersex hypothesis, 1991
- Interstitial lung disease, 2226
- Intersubjectivity, psychoanalysis and, 2646, 2647
- Interval recording, 2687
- Interventions. *See also specific therapies*
  - for alcohol-use disorders, 1276
  - for best practices, 4348–4349
  - for relapse prevention, 4366–4367
  - research in, 4352
  - for schizophrenia, 1553–1554, *4366t*
  - symptom displacement with, 4369
  - targeting illness/impairment, 4304
  - in treatment, 4338
- Interviewer-based diagnostic interviews, 3396, 3397–3398
- Interviews. *See also* Examination of psychiatric patient
  - behavior therapy, 2685
  - child, for psychiatric evaluations, 3381–3385, 3383*t*
  - during child psychiatric evaluations, **3375–3385**
    - child interview during, 3381–3385, 3383*t*, 3411
    - considerations during, 3375–3376
    - context considerations, 3376
    - parent interviews, as part of, 3376–3381, 3411
    - preliminary contact for, 3376
    - techniques for, 3380
  - cultural formulation, 790–791
  - description of, 4589
  - diagnostic interviews, 3395–3398

    - interviewer-based, 3396, 3397–3398
    - overview, 3395
    - respondent-based, 3395–3397
  - forensic, 963–964, 2734
  - motivational, 1553
  - observed and recorded, 963
  - for personality disorders, 2161
- Intestine transplants, survival rates, 2358
- Intestines, nutrient sensing by, 360
- Intimacy
  - adulthood, 3941, 3942–3943, 3943*f*
  - love and, 1963–1965
  - and mental health, 759
- Intimate partner violence
  - aggression and, 2486–2487
  - definition, 2539
  - disaster or war and, 2569
  - perpetrators, 2539–2540
  - prevalence, 2486–2487, 2539
  - risk factors, 2539
  - traits associated with, 2486–2487
  - treatment, 2502–2503
- Intoxication
  - alcohol-related, 1271, 1271*t*
  - barbiturate-related, 1383–1384
  - benzodiazepine-related, 1382
  - caffeine-related, 1296–1297, 1297*t*, 1298*t*
  - cannabis-related, 1306, 1311
  - child maltreatment and, 3837
  - description of, 4589
  - hallucinogen-related, 1318–1321, 1319*t*
  - inhalant-related, 1337–1338
  - laboratory evaluation, 1016–1018
    - acetaminophen or salicylate overdose, 1018
    - environmental toxins, 1018
    - volatile solvents, 1018
  - opioid-induced, 1354, 1361, 1363

- sedative-, hypnotic-, anxiolytic-induced, 1380, 1382, 1383–1384
- stimulant-related, 1285–1286, 1288
- Intoxication, defined, 1271, 1271*t*
- Intoxication delirium
  - cannabis-related, 1306–1307
  - hallucinogen-related, 1327
  - inhalant-related, 1338
- Intracellular cascades, mood disorders and, 1715
- Intracerebral hemorrhage, 640*f*
- Intracranial hemorrhage (IVH), 640
- Intracranial self-stimulation (ICSS), 177
- Intracytoplasmic sperm injection (ICSI), 2509, 2510*t*
- Intrafamilial sexual abuse, 3832–3833, 3841
- Intralaminar nuclei, 21
- Intraneuronal signaling, **119–130**
  - remote control of, 129–130
- Intraparietal areas, memory suppression and, 1893
- Intrapersonal conflicts. *See* Intrapsychic conflicts
- Intrapsychic ataxia. *See* Ataxia(s)
- Intrapsychic conflicts, 4589
- Intrapsychic factors
  - dissociative disorders and, 1873
  - mood disorders, 1660–1667
  - psychoanalysis and, 2647
- Intrauterine devices (IUDs), 2512
- Intrauterine factors, 1116
- Intravenous, 4589. *See also* Infusions
- Intraventricular injury, in pre-term infants, 640
- Introject self-state, 1907
- Introjection
  - description of, 4589
  - in young children, 2870
- Introjective depression, 1662
- Intropunitive, definition of, 4589
- Introspection
  - and cognition, 749
  - definition of, 4589

- Introversion, definition of, 4589
- Introverted individuals, 576
- Intrusion symptoms, 1813*t*, 1814
- Intuition, definition of, 687
- Invasiveness, brain stimulation, 3302–3303, 3302*f*
- Invega. *See* Paliperidone
- Inventory of Childhood Memories and Imaginings (ICMI), 2733
- Inversion, definition of, 4589
- Inverted personality disorder. *See* Schizoid personality disorder
- Investigational new drug (IND), 2931, 2933
  - benzodiazepine, 3003
  - calcium channel inhibitors, 3021
  - MAOIs, 3123
  - melatonin receptor agonists, 3106–3107
  - regulations, 2931
  - review process, 2933
    - clinical hold in, 2933
    - procedures, 2933
- Investigational treatments, ethical issues, 2278
- Involuntariness, hypnosis and, 2710
- Involuntary admission, 4589
- Involuntary commitment
  - basis for, 4332
  - cost containment restrictions on, 4326
  - psychiatric emergencies and, 2636–2637
- Involuntary discharge, 4589
- Involuntary treatment
  - definition of, 4589
  - schizophrenia, 1558–1559
- Involutional melancholia. *See also* Major depressive disorder
  - about, 1601, 1604
  - definition of, 4589
- Inward rectifier potassium channels, 138
- Iodine
  - ^{123}I, 287
  - thyroid hormone and, 2243
- IOM. *See* Institute of Medicine (IOM)

- Iomazenil, 1761
- Ion channels, **135–141**
  - clinical aspects of, 143–146
  - description of, 130
  - disorders of, 676
  - G protein-activated, 2914, 2915*f*
  - neuropsychiatric disorders and, 143–144
  - neurotransmitters and, 141–143
  - selectivity, 132
  - voltage-gated, 135
- Ionotropic receptors, glutamate binding, 1747
- Iproniazid (Marsilid)
  - chemistry, 3113*f*
  - history, 2917, 3111
- Ipsilateral intergeniculate leaflet (IGL), 205
- I.Q. *See* Borderline intellectual and adaptive functioning (BIF); Intelligence quotient (I.Q.)
- Iraq war veterans
  - dissociation and combat trauma in, 1873
  - hypnosis for, 2718
  - PTSD and, 1816
  - suicide and, 2568
- Ireland, immigrants, and mental health, 804
- Irinotecan, fatigue following treatment, 2264
- Iron
  - beneficial effects, 2549
  - brain neurodegeneration and, 672
- Irregular sleep phase circadian rhythm disorder, 4589
- Irrelevant answers, 4589
- Irresistible impulses, 4589
- Irritability
  - definition of, 4589
  - in HD, 556
  - mania and, 1640
- Irritable bowel syndrome (IBS)
  - bowl habits, 2202*f*
  - case studies, 2209
  - definition of, 2202–2203, 2202*t*

- hypnosis and, 2729
- hypnotizability and, 2707
- presentation of, 2202
- psychological comorbidities, 2202

- Irritable mood
  - definition of, 4589
  - ODD symptom, 3610
- Is Path Warm acronym, 1671–1672
- Isaac syndrome, 676
- Isaacs, Susan, 2870
- Ischemic stroke after TBI, 530
- Isobutyl nitrite-related disorders, 1341
- Isocarboxazid (Marplan)
  - chemistry, 3113*f*
  - depression, 1681
  - eating disorder, 3121–3122
  - history, 3111
  - pharmacological actions, 3115
- Isolated explosive disorder, 4590
- Isolation
  - description of, 4590
  - feelings of, 2140
- Isoniazid
  - chemistry, 3113*f*
  - drug interactions, 3054
- Isophiic, definition of, 4590
- Isopregnenolone, 1748
- Isoproterenol, 2327
- Isoptin. *See* Verapamil
- Isoretinoin (Accutane), 2345
- Isradipine (DynaCirc)
  - chemistry, 3016
  - organ/system effects of, 3020
  - therapeutic indications, 3020–3021
- Itard, Jean-Marc
  - on intellectual disability terminology, 3491
  - on Tourette syndrome, 3636, 3650

- Iterative reconstruction, 289–291
- I–Thou, description of, 4590
- *Ixodes* tick, 589–592, 632
- Iyengar yoga, 2557

J

- JACHO (Joint Commission on Accreditation of Healthcare Organizations). *See* Joint Commission on Accreditation of Healthcare Organizations (JACHO)
- Jackson, Arthur, 2721
- Jackson, Don
  - family therapy and, 2749
  - psychodynamic family therapy and, 2754
  - strategic family therapy, 2753
- Jackson, John Hughlings, 4537
- Jackson Hole Group, 4293
- Jackson Personality Inventory, 998*t*
- Jacksonian epilepsy, 4590
- Jacobs, Theodore, 2653, 2654
- Jacobson, Edith, 2644
- Jacobson, Neil, 2683
- *Jaffee v. Redmond,* 2010, 2671
  - privacy of patient, 4436
- *JAG1* gene, 1740
- Jahoda, Marie
  - humanistic psychology and, 2843
  - mental health definition, 754*t*, 755
- Jail sentences of elderly adults, 4239
- Jails. *See* Correctional systems
- *Jamais vu,* 4590
- James, William, 331
  - on alcohol's physiological power, 2582
  - dissociative disorders, 1868, 1895
  - on dual consciousness, 1934
  - functionalism and, 2842
  - on religion, 2591
  - on transpersonal, 2824
- James–Lange theory of emotion, 4590
- Janet, Pierre, 4590
  - conversion disorder and hypnosis, 2719
  - dissociative disorders, 1866–1869, 1895–1896, 1899
  - hysterical psychosis, 1948
  - *L'automatisme psychologique: Essai de psychologie experimentale sur les*

formes inferieures de l'active, 4536
- Jargon aphasia, 4590
- Jarish–Herxheimer reaction, 591
- Jarvis, Edward
  - on cost of care *vs.* treatment, 4335
  - on diseases of the mind, 4329–4331
- Jasper, Herbert, 219
- Jaspers, Karl, 1411, 1413
  - *General Psychopathology,* 4541
  - on psychiatric classification, 1152, 1162
  - on somatic phenomena, 1830
- JC virus, PML and, 586
- Jealousy
  - natural selection and, 776
  - obsessional, 1797
- Jensen, Mark P.
  - on hypnosis, 2724, 2729
  - on hypnosis and pain, 2726
- Jet lag, 2098
- Johns Hopkins, psychiatric clinic at, 4291
- Johnson, Mark, 2831–2832
- Johnson, Susan, 2753
- Johnson, Virginia, 1953
- Johnson (Lyndon) administration
  - Medicare/Medicaid passed during, 4335
  - War on Poverty, 4305
- Johnstone, Brick, 2577
- Johrei as energy therapy, 2561
- Joint Commission on Accreditation of Healthcare Organizations (JACHO)
  - description of, 4590
  - psychiatric education and, 4407
- Joint Commission on Mental Illness and Mental Health, 4292, 4541, 4590
- Joint disorders, 3129
- Joints, inflammation of, 456
- Jones, Ernest
  - as Freud translator, 2645
  - work of, 4590

- Jones, Mary Cover, 3749
- Jones, Maxwell, 2737
- Jorgensen, George, 2024
- Josephi, Simeon, 4403
- Joy, about, 1632
- Judgement
  - definition of, 4590
  - disturbances of, 1123–1124
  - poor, manic syndrome and, 1641
- Jung, Carl Gustav, 911f, 2869
  - active imagination, 2823
  - boundary violations, 2641
  - personality theory, 912–914, 931
    - personality structure, 913–914, 913*f*
    - unconscious, 913
  - as psychoanalyst, 2659
  - *Psychological Types,* 4540
  - therapeutic process, 914
  - work of, 4590
- Justice
  - in bioethics, 4241
  - definition of, 4590
- Justice-based ethics, 4442, 4442*t*
- Juvenile justice, 4439–4440
  - psychotropic medications in, 4321

K

- k-Space
  - brain image, 268*f*
  - definition of, 265–266
  - distributions, 265
- K2/spice, EEG effects, 254
- Kabat-Sinn, Jon, 2579
- Kabat-Zinn, Jon, 2853
- Kagan, Jerome, 3481
- Kahlbaum, Karl, 4536
- Kainic acid (KA)
  - function of, 76
  - mechanism of action, 507
- Kainic acid (KA) receptors, 76–77
- Kallio, Sakari, 2709
- Kallmann, Franz, 2827
- *Kami-untan-to* (KUT), 2554
- Kamin, Leon, 412
- Kandel, Eric
  - biology of psychotherapy, 2675
  - psychotherapy and brain function, 2812–2813
  - *Psychotherapy and the single synapse,* 4545
- Kanner, Leo
  - on autism, 765, 3571–3573
  - on autism spectrum disorder, 3574
  - *Child Psychiatry,* 4539
- Kanner syndrome. *See* Infantile autism
- Kant, Immanuel
  - *Anthropology from a Pragmatic Point of View,* 928–929
  - physician ethics, 4441, 4442*t*
- Kantrowitz, Judith, 2674
- Kaplan–Meier survival function, 844
- Kaposi sarcoma herpes virus, 594
- Kappa opioid agonists, 1317
- Kappa statistic, 817
- Karasu, T. Byram
  - integration of psychoanalytic models and, 2647–2648

- psychoanalysis and, 2655, 2662
- psychoanalytic research and, 2673
- psychotherapy research and, 2674
- Kardec, Allan, 1867
- Kardiner, Abraham
  - psychoanalysis and, 2641, 2659
  - PTSD studies, 1818
- Karon, Bertram
  - CBT, psychosis and, 2867, 2872
  - on psychosis, 2874–2875, 2876, 2877
- Kasanin, Jacob, 1587
- Kast, Eric, 2892–2893
- *Die Katatonie oder das Spannungsirresein* (Kahlbaum), 4536
- Katon, Wayne
  - on costs of anxiety and depressive disorders, 4322
  - on Triple Aim, 4310–4311
- Katz ADL index, 4031–4032, 4031*t*
- Kaufman, Morris, 1232
- Kaufman Brief Intelligence Test (KBIT), 3423–3424
- Kava *(Piper methysticum)*
  - adverse effects, 2544*t*
  - anxiety disorders, 1782
  - biological treatment with, 2553, 2553*f*
- Kayser–Fleischer rings, 456, 671
- Kearns–Sayre syndrome (KSS), 670
- Kefauver–Harris Amendments, 2931
- Kegel's exercises, 1976
- Kell antigens, absence of, 667
- Kelle, Harold, 2695
- Kelly, George
  - personality theory, 932
  - *The Psychology of Personal Constructs,* 932
- Kelsey, Frances, 2931
- Kemadrin. *See* Procyclidine
- Kennedy, John F., 4290–4291, 4335
- Kennedy disease, 628
- Keppra. *See* Levetiracetam

- Kernberg, Otto
  - on narcissistic disorders, 2806
  - as psychoanalyst, 2639, 2645–2646
  - psychotherapy research and, 2674
  - transference-focused psychotherapy, 2652, 2668
- Kernel trick, 841
- Ketalar. *See* Ketamine
- Ketamine (Ketalar, special K)
  - anxiety disorders, 1747
  - depersonalization/derealization, 1887, 1891
  - depression, 1681, 1715, 2887
  - depressive symptoms and, 202
  - ECT anesthesia with, 2987, 3292
  - effects of, 1312, 1314, 2007
  - hallucinogen-related disorders, 1323, 1325–1326, 1327
  - history, 1317
  - intoxication, 1319–1320
  - in major depressive disorder, 103
  - neuroprotective action
    - in TBI, 524
  - OCD, 1789
  - psychosis induced by, 1314–1315
  - PTSD and SUD comorbidity, 1823
- Ketoconazole (Nizoral)
  - drug interactions, 3106
  - HPA activity and, 1716
- Key–lock principle, 1041
- Kidney. *See also* Renal function
  - blood flow, cannabinoids and, 115
  - carbamazepine effects on, 3036
  - DA receptor antagonists and, 3076
  - lithium and, 3097–3098
  - uremic encephalopathy, 2368–2369
- Kidney cancer, 2268–2269
- Kidney disease
  - psychiatric problems, 4048
  - and sexual dysfunction, 4262

- Kidney function, laboratory evaluation of, 1010
- Kidney transplantation
  - living donor outcomes, 2369–2370
  - survival rates, 2358
- Kierkegaard, Soren, 2821
- Kiesler, Donald, 2775–2776
- Kihlstrom, John, 2710
- Kin selection hypothesis, 1995
- Kindling effect, definition of, 4590
- Kindling hypothesis of PTSD, 1822
- Kindness, happiness and, 2848
- Kinesiology, definition of, 4590
- Kinesthetic hallucination, 4590
- Kinesthetic sense, 4590
  - reduction of, 459
- Kinetic tremor, 458
- King, Mary-Claire, 157
- Kinsey, Alfred, 1953, 1954–1955
- Kinsey heterosexual/homosexual scale, 1983
- Kinship care, 3821
- Kirk, Samuel, 3522
- Kirkbride, Thomas
  - *On the Construction, Organization, and General Arrangements of Hospitals for the Insane,* 4327, 4535
  - on health care costs, 4345
  - hospital prototype of, 4328*f*
  - on hospital safety function, 4332
- Kirkbride, Thomas S., 4590
- Kirmayer, L. J., 1832, 1842
- Kirsch, Irving, 2711
- Klein, Melanie, 3735
  - attachment theory, 886–887
  - countertransference and, 2654
  - group psychotherapy and, 2737
  - object relations theory, 883–886
    - ego and objects, 884–886
    - origins, 883–884
  - *The Psychoanalysis of Children,* 4539

- as psychoanalyst, 2644
- on psychosis, 2868, 2870, 2875, 2878
- work of, 4590

- Klein–Levin syndrome, 455, 2091, 4590
- Kleinman, Arthur, 2446, 2461
- Kleptomania, **2114–2115**
  - description of, 4590
  - diagnosis, 2114
  - epidemiology, 2114
  - etiology, 2114
  - treatment, 2114
- Klerman, Gerald
  - on DSM diagnostic reliability, 1156
  - IPT and, 2776
- Klier, Claudia, 2783
- Klinefelter syndrome, 4590
- Klippel–Feil syndrome, 460
- Klitzman, Robert, 2602
- Klonopin. *See* Clonazepam
- Kluft, Richard P.
  - developmental model of dissociative disorders, 1901
  - on dissociation and dissociative disorders, 1871
  - dissociative identity disorder
    - Four-Factor Theory, 1906–1907
    - typology of self-states in, 1925
- Klüver, Heinrich, 31
- Klüver–Bucy syndrome, 464, 4590
- *Knowing What Works in Health Care: A Road Map for the Nation* (IOM), 971
- Knowledge
  - crystallized, in aging adults, 4004–4005
  - in patient interview, 967
- Kohlberg, Lawrence
  - on moral development, 3370
  - moral reasoning of children, 710
  - on social development, 3361
- Kohut, Heinz
  - *The Analysis of the Self,* 4543

- on narcissistic disorders, 2806
- as psychoanalyst, 2639, 2646, 2647
- short-term dynamic psychotherapies, 2804
- transference neurosis and, 2652
- work of, 4590

- Kolle, Kurt, 1583
- Koo, John, 2332
- Korean War veterans
  - amnesia, 1887
  - coercive persuasion and, 1946
  - dissociative disorders, 1869
- Koro (body obsession), 1797, 4590
- Korsakoff, Serghei, 1230
- Korsakoff psychosis
  - description of, 4590
  - hallucinogens and, 2890
- Korsakoff syndrome
  - amnestic disorders and, 1230–1231
  - diencephalic amnesia, 732, 741
- Kozak, Michael, 1769
- Kraepelin, Emil, 332
  - on comparative psychiatry, 1426–1427
  - on cultural assessments, 2460
  - on delusional disorder, 1582
  - on dementia, 1191
  - on negative symptoms, 1415
  - on phenomenology, 1407
  - on psychiatric classification, 1152, 1162
  - *Psychiatrie: Ein Lehrbuch fur Studerende und Aerzte,* 4536
  - on schizophrenia recovery, 1548–1549
  - on stereotypic movement disorders, 3667–3668, 3681–3682
  - work of, 4591
- Krafft-Ebing, Richard von, 1953, 1953*f*, 1974
  - on gender variance, 2024
  - on homosexuality, 1986
  - on paraphilias, 2014
  - *Psychopathia sexualis,* 4536

- on sexual appetite, 2039
- on sexual sadism, 2018–2019
- Kral, V. A., 1222–1223
- Kramer, Henry, 4533
- Kretschmer, Ernest, 4538, 4591
- Kris, Anton A., 2645
- Kris, Ernst, 2738
- Krishnaber, Maurice, 1870
- Kroenke, K., 1842
- Krystal, Jonathan
  - on death feigning, 1890
  - on neurochemical systems and depersonalization, 1891
- Kübler-Ross, Elizabeth, 2289
- Kuf disease, 467*t*, 658
- Kugelberg–Welander disease, 628
- Kundalini awakening, 2825
- Kundalini yoga, 2557
- Kuru, 600
  - ambulatory phases, 611*f*
  - cause of, 601*t*
  - description of, 4591
  - presentation of, 601–602, 611*f*
- Kynurenine, 196
- Kypokalemia, 683

L

- L-Deprenyl (Eldepryl), history, 3111. *See also* Selegiline
- *La belle indifférence,* 4591
- *La folie a double forme* (Falret), 4535
- La Salpêtrière patients, 1867–1868, 1868*f*
- Labia majora, 1959*t*
- Labia minora, 1959*t*
- Labile, definition of, 4591
- Labile affect, 4591
- Labile mood, 4591
- Lability. *See also* Emotional lability
  - mania and, 1640
- Laboratory testing, 467–468. *See also* Medical assessment and laboratory testing
  - for child psychiatric evaluations, 3407–3408
  - for dementia, 1197, 1197*t*
- Lacan, Jacques
  - *Language of the Self,* 4543
  - as psychoanalyst, 2645
- Lacerations, brain, 523
- Lack of being, feelings of, 2140
- Laconic speech, 4591
- Lacquers, hexacarbons in, 631
- Lactate
  - MRS study, 263*f*
  - panic disorder and, 1761, 1762
- Lactation. *See also* Breast-feeding
  - benzodiazepines and, 2999
  - CAM treatments and, 2543
  - DA receptor antagonists and, 3077
  - MAOIs and, 3115–3116
  - melatonin and, 3105–3106
  - mirtazapine and, 3109
  - nefazodone and, 3130
  - pharmacotherapy during, 2519–2521
  - Thomas Hale's website, 2518
  - vilazodone and, 3346

- Lactic acidosis, MELAS and, 669–670
- Lacunar amnesia. *See* Localized amnesia
- Ladostigil, 3123
- LAGB. *See* Laparoscopic adjustable gastric banding (LAGB)
- Laing, R.D., 4332
- *Lake v. Cameron,* 4435
- Lakoff, George, 2831–2832
- Lalophobia, 4591
- Lambert–Eaton myasthenic syndrome (LEMS), 633
- Lamictal. *See* Lamotrigine
- Lamor frequencies, 262*f*
- Lamotrigine (Lamictal), **3083–3088**
  - acute mania, 1685–1686, 1689*t*, 1690
  - adverse effects of
    - dermatological, 2356*t*
  - for aggressive behavior
    - BPD, 2501
    - in IED, 2501
    - intermittent explosive disorder, 2501
    - TBI, 2502
  - anxiety disorders, 1748, 1781
  - BD, 3016, 3084, 3085
  - bipolar depression, 1690, 1692*t*, 1693, 1694
    - child/adolescent, 3693, 3707
  - bipolar disorders, 1681, 4077
    - alcohol comorbidity, 1699
    - cocaine abuse, 1699
    - FDA approval of, 1612
    - in HIV settings, 575
  - bipolar I disorder, 1696, 1696*t*
    - combination therapy, 1696*t*
    - depression, 1697
  - bipolar II depression, 1697
  - black box warning, 2919, 3085
  - carbamazepine and, 3037, 3086, 3326
  - chemistry, 3083, 3083*f*
  - for children, 650

- combination pharmacotherapy, 3326, 3328
- depersonalization/derealization disorder, 1945
- for depression, 2237
- dissociative identity disorder, 1891, 1927*t*, 1928
- dosage/administration, 3086–3087, 3086*t*
- drug interactions, 3086
  - anticonvulsants, 2978
- drug studies, 3084
- efficacy of, 519
- for excoriation disorders, 2352
- gambling disorder, 1808
- history, 3083
- intellectual disability and, 3515
- laboratory interferences, 3086
- lithium and, 3098, 3326
- lithium *v.,* 3092, 3092*f*, 3092*t*
- MDD, 3084
- neurological disorder, 3084
- obesity, 3084
- OCD, 1789
- organ/system effects of, 3084–3085
- in pain syndrome, 3084, 3085
- in pain syndromes, 2396
- pharmacological actions, 3083
- phenobarbital and, 3083
- phenytoin and, 3083
- and physical health monitoring, 1024
- precautions/adverse effects, 3085–3086
- psychiatric disorder, 3084
- seizure, 3084, 3085
- side effects of, 496
- therapeutic indications, 3085
- valproate and, 3086–3087, 3086*t*, 3279, 3326

- Lancet Global Mental Health Group, 4531
- Landau–Kleffner syndrome, 3549
- *Landrigan v. Arizona,* 2421–2422
- Langerhans cells, 2337

- *Langley v. Coughlin*, 4452
- Language
  - of aging adult, 3974, 4005
  - in autistic spectrum disorder, 697
  - and consciousness, 692
  - in cultural psychiatry, 807
  - deficits after TBI, 529
  - in delirium, 4089
  - development, and intelligence, 712
  - executive function and, 2431
  - human brain and, 2472
  - late emergence of, 4591
  - lateralization of, 449
- Language disorders, **3545–3551**. *See also* Communication disorders
  - autism spectrum disorder and, 3580–3581
  - clinical features, 3547–3548
  - comparative nosology, 3545
  - course, 3550
  - defined, 3545
  - dementia and, 1194
  - description of, 4591
  - development in childhood, 3358–3359
  - diagnosis, 3547–3548
  - differential diagnoses, 3548–3550, 3553
  - epidemiology, 3545
  - etiology, 3546–3547
  - frontotemporal dementia variant and, 1210, 1211
  - hemispheric dominance
    - Broca aphasia, 977*f*, 978
    - conduction aphasia, 979
    - global aphasia, 979
    - Wernicke aphasia, 979
  - laboratory examination, 3548
  - mental status examination component, 3382
  - neuropsychological assessment of, 3434–3435
  - overview, 3545
  - pathology, 3548

- prognosis, 3550
- in psychiatric disorders, 1135
- RAD and delays in, 3666, 3680
- schizophrenia symptoms of, 1420
- selective mutism and, 3720, 3721
- specific learning disorder and, 3526, 3529
- speech sound disorder and, 3557
- treatment, 3550

- *Language of the Self* (Lacan), 4543
- Language therapy, 3723
- Lanius, Ruth A.
  - 4-D model of trauma-related dissociation, 1900
  - corticolimbic disconnection hypothesis, 1891
  - on dissociation and dissociative disorders, 1871
  - on HPA axis and depersonalization, 1890
- Lanterman–Petris–Short Act, 4389
- Lanugo, description of, 4591
- Laparoscopic adjustable gastric banding (LAGB), 2221, 2222*f*
- Laparoscopic sleeve gastrectomy (LSG), 2221, 2222*f*
- Lapus linguae, 4591
- Large neutral amino acid (LNAA)
  - availability of, 660–661
  - transport of, 62–63
- Larmor frequency, 260, 260*f*
- Latah, description of, 4591
- Late adolescence, 3367–3368
- Late language emergence, 4591
- Late luteal phase dysphoric disorder (LLPDD), 4591
- Late-onset Alzheimer disease, 1203
- Latency, employment during, 3943
- Latency phase, 4591
- Latent class growth analysis, 843
- Latent content, description of, 4591
- Latent homosexuality, 4591
- Latent inhibition, 416
- Latent schizophrenia, 4591
- Lateral fissure, 7*f*

- Lateral ganglionic eminence (LGE)
  - *Gsh2* gene expression, 50–51
  - transcription factors expressed in, 51
- Lateral geniculate nucleus
  - coronal section, 12*f*
  - neural connections, 7
- Lateral hypothalamic area (LHA)
  - anatomy of, 36
  - lesions, eating control and, 364
- Lateral orbitofrontal cortex, projections, 449
- Lateral posterior parietal cortex, 7*f*
- Lateral sulcus, coronal section, 12*f*
- Lateral ventricles, 14
- Laterality of brain, and consciousness, 692
- Laterodorsal tegmental (LDT) nuclei, 36
- Latin Americans, U.S. demographics, 794, 794*t*. *See also* Race/ethnicity
- Latuda. *See* Lurasidone
- Laughing, pathological, 464
- Lauterbur, Paul, 259
- *L'automatisme psychologique: Essai de psychologie experimentale sur les formes inferieures de l'active* (Janet), 4536
- Lavender *(Lavandula angustifolia),* 2554
- Law of contiguity, 411
- Law of contrast, 411
- Law of frequency, 411
- Law of similarity, 411
- *Lawrence and Garner v. Texas,* 2000
- Lawsuits/litigation. *See also* Malpractice
  - family and couples therapy, 2757–2758
  - hypnosis and, 1940
  - not guilty by reason of insanity and, 1924
  - traumatic memories, 1926
- Lawton–Brody Instrumental Activities of Daily Living, 4031–4032, 4031*t*
- Lax response-criteria hypothesis, 2732
- Laycock, Thomas, 4536
- Lazabemide, 3113*f*
- Lazell, Edward W., 2736
- LBD. *See* Lewy Body Disease

- *L.C., Olmstead v.,* 4337
- Lead
  - prenatal exposure to
    - aggression and, 2484
  - toxic neuropathy and, 631
- Lead-pipe rigidity, 458
- Lead testing for children, 3408
- Leadership, bipolar II disorder and, 1652
- Leading questions, 2734
- LEARN Program, 2216–2217
- Learned autonomic control, 4591
- Learned food aversions (LFAs), 2271–2272
- Learned helplessness. *See* Helplessness
- Learned social roles, 1873
- Learning
  - of aggression, 2426
  - of aging adult, 4005–4006
  - appetite and, 366–367
  - assessment of
    - achievement testing for, 3425–3426
    - intellectual testing for, 3422–3425
    - processing assessments for, 3426–3429
  - associative, **411–417,** 2139
  - behavior therapy and, 2683–2684
  - cognitive assessments for, 3415–3417
  - defined, 3417–3418, 3421, 3425, 3522
  - expectations and, 412–413
  - habit-related, 435
  - hierarchical model of, 2129*f*
  - history, 3522–3523
  - law on, 3419–3421*t*
  - neurocognitive impairment and, 1503–1504
  - oppositional defiant disorder and, 3611
  - physician ethics, 4444
  - procedural memory and, 2139
  - process of, 355–356
  - psychotherapy and, 2816–2830

- schizophrenia and, 1503–1504
- sexual, 1961–1963
- sleep and, 353
- specific learning disorder (*See* specific learning disorder)
- stimulant-related disorders and, 1283–1284
- stimulant use and, 1283
- testing for, 3422–3429
- Tourette disorder and, 3643, 3657

- Learning disabilities
  - diagnosis of, 2436
  - prematurity and, 643
- Learning disorders, 4591. *See also* Dyslexia
- Learning theories, **716–728**
  - basic concepts, 717–718, 717*f*, 718*f*
  - brief history of, 411–415
  - cross-references, 727
  - description of, 4591
  - historical context, 716–717
  - inhibitory learning and, 1771
  - operant/instrumental learning, 722–725
    - behavior and payoff, 722–724, 723*f*
    - motivation, 724–725
    - reinforcement, 724
  - Pavlovian conditioning, 718–722
    - and behavior, 718–719
    - erasing of, 721–722, 721*t*
    - process of learning, 720–721, 720*f*
  - Pavlovian plus operant, 725–727
    - avoidance learning, 725–726
    - instrumental action, 726–727, 726*f*
- Leary, Timothy, 2584
- Least significant difference, 833
- Leber hereditary optic neuropathy (LHON), 670
- LeCron, Leslie, 2717
- *Lectures on Conditioned Reflexes* (Pavlov), 4539
- LeDoux, Joseph, 2840
- *Lef1* gene deletion, 50

- Left ventricular assist devices (LVADs), 2192
- Leg cramps, sleep-related, 2103
- Leg movements, polysomnography, 2104*f*
- Legal issues. *See also* Forensic psychiatry; Lawsuits/litigation; Malpractice
  - advance directives, 2304–2306
  - assault victim, 2541
  - autopsies, 2305–2306
  - brain death determination, 2305
  - CAM treatments, 2543
  - child/adolescent
    - civil litigation, 3864–3865
    - custody issues, 3863–3864
    - evidentiary standards, 3867
    - forensic evaluations, 3862–3863
    - juvenile justice cases, 3865–3867
  - for child cognitive assessment, 3415–3417, 3419–3421*t*
  - competency to stand trial, 4344
  - with dementia, 1200
  - durable powers of attorney, 2305
  - for emergency psychiatric evaluations, 3447
  - end-of-life decisions, 2305
  - factitious disorder and, 1864
  - forensic psychiatry, 2731–2735
  - genetic counseling, 2536–2537
  - for intellectually disabled persons, 3492, 3493, 3517, 3518
  - living wills, 2305
  - maltreatment (child) and, 3828, 3840
  - and neurocentrism, 784–785
  - organ donation, 2305–2306
  - proxies, 2304–2306
  - for school consultations, 3883
  - sex and, 1963
  - sexual transgressions, 2669–2670
  - surrogates, 2304–2306
- Lemon balm, 1782
- Lengths of stay
  - decline in, 4325, 4329

- hospitalist role in, 4355, 4356, 4358
- Lenticular nucleus, 10
- Leprosy, neuropathies and, 632
- Leptin, **176**
  - inhibition of orexins, 677
  - secretion of, 362–363
- Lesbian and Gay Aging Issues Network (LGAIN), 4259
- Lesbian gay bisexual transsexual (LGBT) people
  - clinical issues, 2035
- Lesbian identities, acquisition of, 1997
- Lesbianism, description of, 4591
- Lesbian(s)
  - children of, 2004
  - clinical issues, 2006–2009
  - definition of, 1983
  - Internet resources for, 2012*t*
  - medical illnesses in, 2008–2009
- Lesch–Nyhan syndrome
  - description of, 4591
  - stereotypic movement disorders, 3668, 3671, 3682, 3685
- Lescol. *See* Fluvastatin
- Leshley, Jan, 2839
- Lesions, neurosurgery, 3316–3317, 3316*f*, 3317*f*
- Lethologica, 4591
- Letters, supportive, suicide risk and, 2617
- Leucine-rich repeat kinase 2 (LRRK2), 543
- Leukemia inhibitory factor (LIF), 100
- Leukemias, psychological issues in, 2269–2270
- Leukocytes, distribution of, 2338
- Leukodystrophies
  - metachromatic, heritability of, 2530
  - MS *versus,* 563
- Leukotomy. *See also* Lobotomies
  - limbic, 3316, 3318
- Leuner, Hanscarl, 2823
- Leuprolide (Eligard), 2020
- Leuprolide acetate (Lupron), 2503

- Levenson, Hanna, 2796
- Levetiracetam (Keppra)
  - acute mania, 1686
  - aggression in autism, 2500
  - for aggressive behavior
    - in autism, 2500
  - anxiety disorders, 1781
  - bipolar depression, 1690, 1692*t*
  - chemistry, 2966*f*, 2974
  - combination pharmacotherapy, 3326
  - dosage/administration, 2975
  - drug interactions, 2975
  - drug studies, 2974
  - history, 2973–2974
  - in pain syndromes, 2396
  - pharmacodynamics, 2974
  - pharmacokinetics, 2974
  - pharmacological actions, 2974
  - precautions/adverse reactions, 2974–2975
  - therapeutic indications, 2974
- Levodopa (Larodopa, Dopar), 29, 547
- Levomethadyl (ORLAAM), 3133
- Levomilnacipram (Fetzima)
  - characteristics of, 3147, 3152
  - depression, 1679
  - for geriatric depression, 4073
  - side effects, 1679*t*
  - tolerance of, 2261
- Levorphanol (Levo-Dromoran), 2293
  - side effects, 2255
- Levothyroxine (Levoxyl)
  - dosage/administration, 3246
  - history, 3242
  - pharmacological action, 3243
  - precautions/adverse reactions, 3246
  - studies, 3245, 3245*t*, 3246*t*
  - therapeutic indications, 3246

- Levoxyl. *See* Levothyroxine
- Lewin, Kurt
  - on action research, 2842
  - group psychotherapy and, 2737, 2738
  - work of, 4591
- Lewis, Aubrey, 1154
- Lewis-Fernández, Roberto, 1871
- Lewy bodies, α-synuclein in, 543
- Lewy bodies, dementia with (DLB), **1214–1219**
  - clinical features, 1215–1216
  - comparative nosology, 1214–1215
  - course and prognosis, 1217
  - defined, 1214
  - diagnosis, 1216
  - differential diagnoses, 1198–1199, 1217, 1219
  - epidemiology, 1215
  - etiology, 1215
  - genetic factors, 4021
  - history, 1214
  - laboratory studies, 1217, 1218*f*
  - pathology, 1216–1217, 1216*f*
  - treatment, 1217–1219
- Lewy body disease (LBD)
  - treatment, 3040, 3044
    - FGAs, 3072
- Lexapro. *See* Escitalopram
- Lexical hypothesis, 935
- LGAIN. *See* Lesbian and Gay Aging Issues Network
- LGBT inmates in jails and prisons, 4475–4476
- LGBT nondiscrimination statements, 2005
- Lhermitte-Duclos disease, 314
- *Lhx2* gene expression, 49
- *Lhx5* gene expression, 49
- Liaison Committee on Medical Education (LCME), 4406, 4591
- Liaison nursing, 4591
- Liaison psychiatry, 4345
- Libet, Benjamin, 336

- Libido
  - concept of, 863
  - description of, 4591
- Librium. *See* Chlordiazepoxide
- Licarbazepine, 1685–1686
- Lichen simplex chronicus, 2351
- Lick and groom (LG) behavior, maternal, 394–395
- Lidocaine (Dalcaine)
  - in pain syndromes, 2396
  - psychiatric side effects of, 2198*t*
  - topical, 2396
- Lidz, Theodore, 2749, 2874
- Lie detection by neuroimaging, 4489–4490
- Life chart, 2317*f*
- Life cycles
  - family transitions, 4194
  - gender differences, 4249
- Life events
  - adverse, suicide and, 1671
  - anxiety disorders and, 1735–1736
  - mood disorders and, 1611, 1611*f*
    - suicide and, 1671
  - on psychiatric symptomatology, impact of, 1116–1117
- Life fear, Rankian dialectic, 918
- Life goals, psychoanalysis and, 2659
- Life instinct. *See* Sexual drive
- Life review
  - in group therapy, 4202, 4204
  - in psychotherapy, 4172, 4174
- Life scripts, client-centered therapy and, 2820
- Life-skills building, 1869
- Life Skills Profile (LSP), 4372
- Life support, withdrawal of, 2297
- Life table methods, 843–844
- Life-threatening experiences, 1887
- Lifespans, 2288, 3947, 3947*f*
  - issues

    - of autism spectrum children, 3925–3927
    - of bipolar disorder children, 3929
  - and mental health, 3996
  - positive psychiatry, 4284
- Lifestyles
  - changes in
    - addictive processes and, 2063
  - ED and, 3338
  - and HIV infection, 4271, 4273
- Lifetime expectancy, 4591
- Lifetime prevalence, 4591
- Lifetime prevalence estimates, 4518*f*
- "Lifting the Burden" campaign, 619
- Lifton, Robert J., 1946, 4538
- Ligand-gated ion channels, 131
- Ligand gated receptors, 2907, 2908*f*
- Light therapy
  - alcohol withdrawal and, 214
  - forms validated by Western science, 2546
  - postpartum depression, 2515–2516
  - for SAD, 212–213
  - scientific validation of, 2558
  - seasonal depressions and, 1651, 1679
  - for sleep-wake disorders and, 2098
- Lighting, artificial, 211*f*
  - insomnia and, 2295–2296
- Likelihood, description of, 290
- Lilienfeld, Scott, 2673
- Lilliputian hallucinations, 4591
- Lily family toxins, 136
- Limb apraxia, 979
- Limbic cortex, 31–33
- Limbic-cortical-striatal-pallidal-thalamic (LCSPT) circuit, 277
- Limbic encephalitis, 2270
- Limbic leukotomy, 3316, 3318
- Limbic system, 31–38
  - anatomic structures of, 32*f*

- components of, 8, 10
- description of, 4591–4592
- functional neural circuitry, 38*f*
- major structures of, 31–37
  - amygdala, 35
  - hippocampal formation, 33–35
  - hypothalamus, 35–37
  - limbic cortex, 31–33
  - septal area, 35
- and mental health, 765–766
- neurobiology of, 450
- noradrenergic neurons in, 1714
- sexuality and, 1957

- Limbic-thalamic-cortical (LTC) circuit, 277
- Limbic–psychotic trigger reaction
  - epilepsy and, 2496
  - intermittent explosive disorder and, 2494
- Limit-setting sleep disorder, 2086
- LIMK kinase, 53
- LINCs, function of, 152
- Lindemann, Erich, 2623
- Line spread functions (LSFs), 292
- Linehan, Marsha
  - dialectical behavior therapy, 2785
  - dissociative identity disorder, 1930
- Linguistically appropriate services
  - competence in, 4381–4382
  - standards for, 4380
- Linkage analysis
  - approaches, 305*f*
  - of homosexuality, 1995
  - for risk gene mapping, 3449–3450, 3449*t*
- Linkage disequilibrium (LD), 307, 3449
- Linkage studies, 4019
- Linoleic acid (LA), 116
- Lioresal. *See* Baclofen
- Liothyronine (Cytomel)

- combination pharmacotherapy, 3323
- dosage/administration, 3246
- history, 3242
- pharmacological action, 3243
- precautions/adverse reactions, 3246
- studies, 3244–3245, 3245*t*
- therapeutic indications, 3246

- Lipids
  - lipid rafts, 2912, 2913*f*
  - role in treatment selection, 1022
- Lipitor. *See* Atorvastatin
- Lipofuscin in aging brain, 3991
- Lipopolysaccharide (LPS), fetal exposure to, 201
- Lipowski, Zbigniew
  - on psychosomatic illness, 2180
  - on somatization, 1830
- Lipton, Samuel D., 2640–2641
- *LIS-1* gene, cell migration and, 52
- Lisdexamfetamine (Vyvanse), 64
  - for binge eating disorder, 2081
  - use in cancer patients, 2262
- Listening attitudes, 2878
- Literacy, and electronic technology, 4280
- Lithium (Carbolith, Eskalith), **3088–3101**
  - acute mania, 1684, 1686, 1688, 1689, 1689*t*
    - combination/add-on therapy, 1687, 1689–1690, 1689*t*
    - emergency management, 1688
  - adverse reactions, 2356*t*, 3096–3098, 3096*t*
    - neurological, 3096
  - for affective aggression, 2172
  - for aggressive behavior, 2111, 3093
    - conduct disorder, 2501
    - in IED, 2501
    - intermittent explosive disorder, 2501
  - alcohol-related disorders, 3093
  - antidepressants combined with, 3092, 3098, 3323, 3327
  - antipsychotics combined with, 3093, 3098, 3327

- BD, 3015–3016, 3015*f*, 3089–3092, 3089*f*, 3091*f*, 3092*f*, 3092*t*
  - maintenance, 3090–3092, 3091*f*, 3092*f*, 3092*t*
- bipolar depression, 1690, 1692*t*, 1693
  - combination therapy, 1692*t*
  - maintenance therapy/prophylaxis, 1693–1694
- bipolar depression, early onset, 3693, 3707
- bipolar II depression, 1697
- for bipolar disorder, 215, 3791, 4076–4077
- bipolar disorder, early onset, 3692, 3706
- bipolar I disorder, 1696, 1696*t*, 1697
- bipolar II disorder, 1698
- BPCA and, 2931
- for cancer patients, 2263
- carbamazepine and, 3036, 3098
- carbamazepine *v.,* 3025, 3026–3027, 3026*f*, 3033, 3325
- cardiovascular effects of, 2197*f*
- cardiovascular reactions, 3097
- chemistry, 3088
- in children, 3094
- for children with CD, 2501
- clinical effectiveness in mood disorders studies, 1629
- clozapine combined with, 3327
- combination pharmacotherapy, 3088, 3089, 3323, 3325–3327
- conduct disorder, 3790
- contraindications, 2521
- depression, 1680, 3089–3090, 3089*f*
- dissociative identity disorder, 1928
- dosage/administration, 3100, 3100*t*
- drug interactions, 2199, 3098–3099, 3099*t*
  - ACE inhibitor, 3098
  - anti-inflammatory drug, 3098
  - anticonvulsants, 3098
  - antidepressants, 3098
  - antipsychotic, 3098
  - calcium channel inhibitors, 3021
  - diuretic, 3098
  - lamotrigine, 3086

- nefazodone, 3130
- drug studies, 3088
- ECT interactions with, 3090, 3294
- EEG effects, 253
- in elderly, 3094–3095
- endocrine effects of, 3796
- family history of bipolar disorder and, 1623
- FGAs and, 3081
- gambling disorder, 1808
- history, 2917, 3015–3016, 3088
- HIV/AIDS and, 3856
- indications, 3093–3094, 3093*t*, 3094*t*
- intellectual disability and, 3516
- intoxication, 3096–3097
- laboratory interferences, 3099–3100
  - dosage prediction, 3100
  - laboratory monitoring, 3099–3100
- lamotrigine and, 3098, 3326
- lamotrigine *v.,* 3092, 3092*f*, 3092*t*
- maintenance therapy, 3090–3092, 3091*f*, 3092*f*, 3092*t*
- mania, 3089
- for mania after stroke, 487
- MDD, 3092
- medically ill patients and, 3095
- mood disorders, 1612
- mood stabilizers, 168
- movement disorder, stereotypic, 3673, 3687
- neurological effects, 3096
- neuronal hyperexcitability, 1609
- nonpsychiatric uses of, 3094, 3094*t*
- pharmacological actions, 3088
- and physical health monitoring, 1024–1025
- plasma cytokine levels and, 199
- precautions, 3094–3096
- pregnancy and, 3095
- psychopharmacology with, 2905
- rapid cycling patients and, 3095–3096

- renal reactions, 3097–3099
- schizoaffective disorder, 3092–3093
- schizophrenia, 3092–3093
- serum calcium and, 174
- sexual dysfunction and, 1972
- side effects of, 574–575, 2173, 2189
  - chemotherapy and, 2263
  - in HIV setting, 572
  - hypothyroidism, 2240
  - thyroid hormone release and, 2240–2241
- SSRIs combined with, 3091
- suicide prevention and, 2621
- suicide risk, 2616
- therapeutic indications, 3089–3094
- therapeutic window, 2371
- thyroid function and, 3097
- thyroid hormones and, 174
- tremors and, 2936, 2942, 2953–2954, 3096
- valproate and, 3088, 3272, 3273*f*–3274*f*, 3278, 3325–3326
- valproate *v.,* 3090–3091, 3091*f*, 3267–3268, 3267*t*, 3270, 3272

- Lithium carbonate (Eskalith, Lithobide)
  - aggression in mentally retarded children, 2500
  - description of, 4592
  - side effects
    - weight gain, 2237
- Lithium therapy, 4542
- Little Albert case, 2682, 4592
- Little Hans, 4592
- Livedo reticularis, 456
- Liver
  - barbiturate effects on, 2986
  - carbamazepine effects on, 3036
  - cholinesterase inhibitor effects on, 3042
  - DA receptor antagonists and, 3076
  - disulfiram effects on, 3053, 3054
  - laboratory evaluation, 1010
  - nefazodone effects on, 3129

- opioid receptor antagonist effects on, 3143, 3144–3145
- tricyclic and tetracyclic antidepressant effects on, 3261, 3262–3263
- Liver disease, 1268
  - and HIV infection, 4271
  - psychiatric problems, 4048
- Liver grafts, survival rates, 2358
- Liver transplantation
  - living donor outcomes, 2369–2370
  - in Wilson disease, 557
- Living, dying and death, 2288
- Living accommodations, 4169
- Living wills
  - competency of older adult, 4238
  - description of, 4592
- Llinás, Rodolfo, 330
- LLPDD. *See* Late luteal phase dysphoric disorder (LLPDD)
- LNNB. *See* Luria-Nebraska Neuropsychological Battery
- Lobo, A., 1842
- Lobotomies, 3314
  - description of, 4592
- Local circuit neurons. *See* Interneurons
- Local field potentials, 405–407
- "Local sleep," 350
- Localized amnesia, 4592
- Localized dissociative amnesia, 1880, 1881*t*
- Locke, John, 411
- Locus ceruleus (LC)
  - description of, 4592
  - noradrenergic neurons in, 1713–1714
  - norepinephrine innervation, 62
  - sleep and, 342
  - wakefulness and, 345
- Loewenstein, Richard J., 1871
- Log linear models, 842
- Logistic regression, 840–841
- Logorrhea, description of, 4592
- Logotherapy, 2821

- description of, 4592
- Lombroso, Cesare, 4536
- Long association fibers, 12
- Long interspersed elements (LINEs), 154
- Long QT syndrome. *See also* QT intervals
  - congenital, 2190
    - drugs affecting, 2199, 2199*t*
  - drug side effects and, 2198–2199
  - methadone and, 2395
  - screening for, 2375
- Long-term care
  - family roles, 4197
  - gender differences, 4245, 4248
  - psychiatric aspects, **4221–4232**
    - anxiety disorders, 4231
    - delirium and comorbid medical conditions, 4224–4226, 4225*t*
    - dementias, 4226–4229, 4227*t*, 4228*t*
    - depression, 4229–4230
    - malnutrition, 4226
    - management of disorders
      - causes of symptoms, 4222, 4224*t*
      - safety, 4222
    - pain, 4226
    - palliative care, 4231
    - posttraumatic stress disorder, 4231
    - prevalence of disorders, 4223*t*
    - psychiatrist role, 4222
    - psychotic disorders, 4230
    - sleep-wake cycle, 4230–4231
  - psychosocial factors, 4169
  - and sexuality, 4264–4267, 4266*t*
- Long-term declarative memory
  - description of, 2139
- Long-term goals (LTGs) in treatment, 4338
- Long-term potentiation (LTP)
  - glutamatergic synapses, 79
  - synaptic plasticity and, 128, 128*f*

- Long terminal repeats (LTRs), 154
- Longitudinal models, 842–843
- Longitudinal studies, 4592
- Loosening of associations, 1420. *See also* Tangentiality
  - description of, 4592
- Lopez-Ibor, Juan José, 1830
- Lopressor. *See* Metoprolol
- Loratadine (Claritin), 75
- Lorazepam (Ativan)
  - agitation and violence, 2997, 2998
  - alcohol withdrawal, 2997
  - BD, 2997
  - BD and panic disorder, 1699
  - benzodiazepine dose equivalent, 2633
  - for bipolar disorder, 4077
  - catatonia, 2997
  - for delirium, 4093
    - in cancer patients, 2257
    - in cardiac patients, 2196
  - delirium treatment, 1190
  - dissociative identity disorder, 1927*t*, 1928, 1929
  - distribution, 2993
  - for elderly patients with cancer, 2283
  - GAD, 1783
  - manic episode in emergency setting, 1688
  - metabolism, 2993
  - for pain management, 2377
  - psychiatric emergencies, 2628
  - seizure disorder, 2997
  - side effects of, 2196
- Lorcaserin (Belviq)
  - mechanism of action, 2220
  - for weight loss, 2220
- LORETA (low-resolution brain electromagnetic tomography), 235–236, 236*f*
- Loss
  - depression and, 1661
  - factitious disorder and, 1851–1852

- Loss/anger turned inward, depression and, 1661
- Loss of interest
  - depressive syndrome and, 1634
  - major depressive disorder and, 1643
- Lotterman, Andrew, 2880, 2881
- Lottery gambler, pathological, 1811
- Lou Gehrig disease. *See* Amyotrophic lateral sclerosis
- Louisville Behavior Checklist (LBC), 3399
- Lourie, Reginald S., 3668, 3682
- Lovastatin (Mevacor), 4147
- Love, intimacy and, 1963–1965
- Love psychology, 4545
- Low back pain
  - conditions, 2390*t*
  - differential diagnoses, 2390
- Low birth weight infants, 639–640
- Low-resolution brain electromagnetic tomography (LORETA), 235–236
- Lower motor neuron disease (MND)
  - characteristics of, 457
  - spinal cord neurons and, 627
- Loxapac. *See* Loxapine
- Loxapine (ADASUVE, Loxapac)
  - acute mania, 1686, 1688
  - chemistry, 3063*f*, 3064
- Loyalties in correctional psychiatry, 4452
- LP. *See* Lumbar puncture
- LSD (lysergic acid diethylamide)
  - description of, 4592
  - pharmacology of, 1315–1316
  - psychological toxicity, 2890
  - psychotherapy and, 2887
  - research, 2887
    - addictions, 2891–2892
  - spirituality and, 2581
  - terminal illness, 2892–2893
- LSG. *See* Laparoscopic sleeve gastrectomy (LSG)
- LTD nuclei, sleep and, 342

- Ludiomil. *See* Maprotiline
- Lumbar puncture (LP)
  - description of, 4592
  - indications, 494
- Luminal. *See* Phenobarbital
- Lund, Thomas, 218
- Lunesta. *See* Eszopiclone
- Lung cancer
  - incidence of, 2251
  - mortality rates, 2251
  - personality disorders in, 2268
- Lung transplants
  - living donor outcomes, 2369–2370
  - survival rates, 2358
- Lupus erythematosus, systemic, 1649
- Lurasidone (Latuda)
  - acute mania, 1686, 1686*t*
  - adverse effects of
    - dermatological, 2356*t*
  - anxiety disorders, 1781
  - bipolar depression, 1691, 1692*t*, 1693, 1695
  - bipolar I disorder, 1697
  - bipolar II disorder, 1697
  - diabetes risk and, 2247
  - side effects
    - weight gain, 2237, 2237*t*
- Luria-Nebraska Neuropsychological Battery (LNNB), 4592
- Lust dynamism, 4592
- Luteinizing hormone (LH), 525
  - function of, 2239
  - menstrual physiology and, 2507
  - peripheral endocrine disorders and, 2248*f*
  - production of, 2239
  - responses, 1993
- Luteinizing hormone-releasing hormone (LHRH), 1980
- Lutetium yttrium orthosilicate (LYSO), 287
- Luther, Martin, 2578, 2578*f*

- Luvox. *See* Fluvoxamine
- Lying, 4592
- Lyme disease (Lyme borreliosis), 589–592
  - brain, SPECT imaging, 592*f*
  - clinical manifestations of, 590–591
  - diagnosis of, 590
  - tic disorders, 3638, 3652
- Lymphocytes
  - maturation of, 179
  - stress and, 193
- Lymphoid tissue
  - innervation of, 186–187, 189*f*
- Lymphomas
  - CNS, 586
  - psychological issues in, 2269–2270
- Lyons-Ruth, Karlen, 1901
- Lypemania, 1603
- Lyrica. *See* Pregabalin
- Lysine, 1782
- Lysosomal disorders, 656
- Lysosomal storage diseases (LSD), 656

M

- Macaque monkey
  - cannabinoid-reactive neurons, 35*f*
  - dopamine transporter, 25*f*
- MacArthur Competence Assessment Tools, 4235
- MacArthur Study of Mental Disorder and Violence, 2489–2490
- Machine learning techniques, 840, 841
- Mack, John, 2590
- MacKenzie, K. Roy, 2783
- MacLean, Paul, 31
- Maconochie, Alexander, 2682
- MACRA. *See* Medicare Access and CHIP Reauthorization Act
- Macropsia, description of, 4592
- Macroscopic cyst degeneration, 642
- Madanes, Cloe, 2753
- Madonna–Putana complex, 1967
- Madrid Declaration, 4531
- *Madrid v. Gomez,* 4452
- MADRS. *See* Montgomery-Asberg Depression Rating Scale
- Mafia, 2413–2414
- Magical thinking, 4592
- Magnesium
  - Alzheimer disease and, 2545
  - beneficial effects, 2549
- Magnesium–calcium combination, 1782
- Magnetic dipoles, 260
- Magnetic diseases, 1866
- Magnetic resonance angiography (MRA), 271
- Magnetic resonance imaging (MRI), **259–258**. *See also* Forensic psychiatry, neuroimaging; Functional magnetic resonance imaging (fMRI)
  - after TBI, 525
  - in aging adult (*See* Geriatric psychiatry)
  - basic physics of, 259–264
  - brain, 1019, 1717
  - brain tumor diagnosis, 494–495
  - BRAINEAC project of aging brain, 3987, 3990
  - in clinical neuroscience, 466

- communication with patients, 283–284
- contrast agents, 266, 269–270
- for delirium, 4091
- description of, 4592
- diffusion-weighted imaging, 270*f*
- gradient-echo, 266
- image contrast, 266
- methodology, 266–271
- morphological, 3458–3459
- safety issues, 282–283
- schizophrenia and, 1464, 1466*t*, 1467–1471
- spin-echo, 266
- T_1 weighting, 269–270, 269*f*, 270*f*
- T_2 weighting, 269–270
- tic disorder, 3639, 3653
- volumetric, 1759

- Magnetic resonance scanners, whole-body, 264*f*
- Magnetic resonance spectroscopy (MRS)
  - after TBI, 525–526
  - basic physics of, 259–264
  - ^{13}C, 279–280, 280*f*
  - in clinical neuroscience, 466
  - communication with patients, 283–284
  - ^{19}F, 282
  - ^{1}H, 278–279
  - ^{7}Li, 282
  - measurements, 262–264
  - neuroreceptor imaging studies, 1759
  - OCD, 1787
  - ^{31}P, 279, 279*f*, 282
  - pediatric considerations, 3459–3460, 3459*f*
  - safety issues, 282–283
  - system hardware, 264–265
- Magnetic seizure therapy, 3299, 3299*t*, 3309–3310, 4163
- Magnetization
  - definition, 260
  - perspectives, 261*f*

- Magnetoencephalography (MEG)
  - thalamic cell oscillations, 746, 749, 749*f*
  - of thalamocortical dysrhythmic syndrome, 750–751
- Mahalanobis distance in QEEG, 1030, 1031
- Mahesh (Maharishi, yogi), 2579
- Mahler, Margaret
  - on development, 2644
  - as family therapist, 2874
  - on gender dysphoria, 2026
  - *The Psychological Birth of the Human Infant,* 879, 4542
  - separation-individuation process
    - hatching, 879
    - object constancy, 879
    - practicing, 879
    - rapprochement, 879
    - symbiotic phase, 879
  - work of, 4592
- Mai, Francois, 1844
- Mainstream medicine
  - psychiatry disconnect with, 4296
  - public psychiatry in, 4308
- Mainstreaming, exceptionalism *vs.,* 4294, *4295f*
- Maintaining factors, case conceptualization and, 2688
- Maintenance drug therapy, 4592
- Maintenance in positive psychology, 2845
- Maintenance of Certification (MOC) programs
  - components, 4406
  - medical processes and, 4431
- Maintenance of wakefulness test (MWT), 343, 2106
- Major affective disorder. *See* Major depression
- Major depression
  - pharmacotherapy
    - combination, 3323–3325
  - treatment
    - combination pharmacotherapy, 3323–3325
    - DBS, 3320
    - ECT, 3285–3286

    - nefazodone, 3126, 3127–3128
    - SSRIs, 3153–3154, 3155
  - unipolar, genetics and, 2529
- Major depressive disorder (MDD)
  - ADHD and, 3596, 3601
  - after traumatic brain injury, 531–533
  - allopregnanolone levels in, 117
  - atypical features, 1644
  - bipolar disorder *versus,* 1654–1655
  - bipolar II disorder and, 1650–1651
  - in brain tumor patients, 497–498
    - course/prognosis, 499
    - diagnosis, 498–499
    - differential diagnoses, 499
    - epidemiology, 498
    - etiology, 498
    - nosology, 497–498
    - pathology, 499
    - treatment, 499–500
  - CBT, 1704–1706
    - course of therapy, 1705–1706
    - initial assessment, 1704
  - child/adolescent
    - differential diagnoses, 3730
    - neuroimaging, 3464–3466, 3465*f*
    - posttraumatic stress disorder and, 3705
    - psychopharmacology, 3790–3791, 3796
    - with psychotic features, 3730
  - chronic, 1644
  - comorbid physical disease, 1643
  - defined, 3674, 3676, 3688, 3690, 3725
  - definitions, 1600
  - description of, 4592, 4607
  - descriptive epidemiology of, 1614, 1615*f*
  - diagnostic threshold, 1643
  - differential diagnoses, 1275, 1596
    - in HIV infections, 571

- Disability Adjusted Life Years (DALYs) and, 1618
- DSM-5 criteria for, 1635*t*
- family studies, 1619–1620
- functional connectivity abnormalities in, 424–426
- gambling and, 1804
- genetic studies of, 318
- genetics, empirical risks for, 2535*t*
- glycemic dysregulation and, 2234
- in HIV patients, 570–572
  - epidemiology of, 570–571
- hoarding disorder and, 1793
- immune-CNS interactions in, 197–198
- insomnia and, 2086
- ITP *versus* paroxetine for, 2813
- medications for stabilization, 2171
- melancholic features, 1604, 1644
- MRI in, 274
- neurotrophins in, 101–102
- OCD *versus,* 1788
- prototypical course of, 1642*f*
- psychopathology, 1634–1639
- psychotherapy
  - in HIV setting, 572
  - individual, 1702, 1702*t*
  - neuroanatomy and, 2814*t*
- psychotic features, 1644
- PTSD and, 1817
  - symptom overlap, 1822
- recurrent, 1643
- recurrent brief depressive disorder and, 1647
- risks of HIV infection, 570–571
- SAD and, 1783
- single-episode, 1643
- STAPP and, 2798
- STDP and, 2801
- stress and, 2323
- traits, 1718

- in transplant populations, 2365–2366
- trazodone drug studies, 3248–3249
- treatment
  - acute and maintenance, 1677–1678, 1677*t*
  - adherence evaluation, 1678
  - adjunctive strategies, 1680–1681
  - β-adrenergic receptor antagonists, 2954
  - antidepressants, 1678–1681, 1679*t*
  - aripiprazole, 3236
  - benzodiazepine, 2996
  - bupropion, 3005
  - buspirone, 3010
  - course specifiers for, 1679
  - DA receptor agonists, 3236
  - dose optimization, 1680
  - duloxetine, 3151
  - experimental antidepressants, 1681
  - first-line antidepressants, 1679–1680, 1679*t*
  - in HIV infections, 571–572
  - lamotrigine, 3084
  - lithium, 3092
  - maintaining remission, 1681
  - MAOIs, 3120
  - measurement-based, 1678
  - mirtazapine, 3108–3109
  - nefazodone, 3127–3128
  - neurosurgery, 3314–3315, 3316, 3318–3319, 3320
  - scopolamine, 2961–2962
  - SSRIs, 3155, 3168–3169
  - third-line antidepressants, 1681
  - TMS, 3306
  - trazodone, 3252
  - vilazodone, 3345
  - vortioxetine, 3168–3169
- Vietnam War with PTSD and, 1816

- Major depressive episodes (MDEs).
  See also Bipolar disorder(s)
  - anabolic steroid-induced, 1401

- cognitive theory of, 1664–1665
- melancholic type, 4592

- Major epilepsy. *See* Grand mal epilepsy
- Major frontotemporal neurocognitive disorder, 1210
- Major histocompatibility complex (MHC) genes, 82–83
- Major motor seizure. *See* Grand mal epilepsy
- Major neurocognitive disorder. *See* Dementia
- Major neurocognitive disorders (NCD), 1273. *See also* Dementia
- Major or mild vascular neurocognitive disorder, **1238–1242**
  - comparative nosology, 1238–1239, 1239*t*, 1240*t*
  - course and prognosis, 1241
  - diagnosis and clinical features, 1241
  - differential diagnoses, 1241
  - epidemiology, 1239–1240
  - etiology, 1240–1241
  - history, 1238
  - pathology and laboratory examination, 1241
  - treatment, 1241–1242
- Major tranquilizer. *See* Antipsychotics/neuroleptics
- Major tranquilizers, ASD and, 3585–3586
- Makari, George, 2641
- Make-a-Picture Story (MAPS) Test, 1004*t*
- Maladaptive, definition of, 4592
- Maladaptive behaviors
  - behavior therapy and, 2683–2684
  - intellectual disability and, 3512–3513
- "*Maladi moun*", "*Maladi bon Dieu*" *versus,* 2444
- Maladjustment, definition of, 4592
- Malan, David, 2647, 2797, 2800–2801
- Male erectile disorder, 1967–1968
  - description of, 4592
  - diagnostic criteria, 1967*t*
  - medical conditions and, 1970–1971, 1970*t*
- Male eunuch identity, 2028
- Male hypoactive sexual desire disorder, 1965–1966
  - description of, 4592
  - diagnostic criteria, 1966*t*

- Male prostheses, penile, 1981
- Male to female (MTF) transgender, 1985
- Male-to-female transition
  - androgens and, 2034
  - genital surgery, 2034
- Maleness, definition of, 4593
- Males, anatomy of, 1954, 1955*f*
- Malformations of cortical development (MCDs), 49
- Malignancy, mood disorders and, 1657
- Malingered dissociative identity disorder, 1916–1917
- Malingering, **2399–2413**
  - avoiding responsibility and, 2402
  - clinical features, 2401–2402
  - clinical presentation of, 2404–2408, 2404*t*
    - ADHD, 2407–2408
    - amnesia and, 2406
    - antisocial personality disorder, 2408
    - depression and, 2406
    - Ganser syndrome, 2407
    - intellectual disabilities, 2407
    - neurocognitive, 2405
    - pain and, 2404
    - pseudoseizures, 2404–2405
    - psychological symptomatology, 2405–2408
    - PTSD, 2405–2406
  - coding of, 2409, 2409*f*
  - collateral diagnostic data, 2409–2410
  - course/prognosis of, 2411
  - criminal contexts, 2402–2403
  - definition of, 2400–2410
  - description of, 4593
  - diagnosis, 2401–2402
  - differential diagnoses, 2410–2411
  - dissociative amnesia and, 1937
  - emergency facilities and, 2403
  - epidemiology of, 2401
  - etiology of, 2401

- factitious disorder and, 1849, 1859
- financial motivation for, 2405–246
- forensic contexts, 2402–2403
- future directions, 2412
- history in literature, 2401
- intentionality and locus of symptoms in, 1850*f*
- in jails and prisons, 4471
- management of, 2412
- medicolegal contexts, 2403
- models of, 2401
- nonmalinger dilemma, 2408
- nosology, 2400–2410
- outpatient settings, 2403–2404
- presentation of, 2402, 2402*t*
- projective identification, 2402
- treatment of, 2412
- truth and, 2402

- Malingering by proxy, 1849
- Malingering Probability Scale (MPS), 2733
- *Malleus maleficarum* (Kramer and Sprenger), 4533
- Malnutrition. *See also* Diet and nutrition
  - anorexia nervosa and, 2074–2075
  - gender differences, 4245
  - in long-term care, 4226
- Malpractice. *See also* Lawsuits/litigation
  - actions, guideline use, 975
  - CAM treatments and, 2543
  - description of, 4593
  - psychiatric, 4436–4437
  - suicide and, 2620
- Malpractice insurance, 4399
- Maltreatment (child), **3828–3843**.
  See also Child abuse; Neglect; Physical abuse; Psychological abuse; Ritual abuse
  - child interviews, 3834–3836
  - collateral information, 3834
  - comparative nosology, 3830–3831
  - corporal punishment, 3840
  - cultural variations, 3839

- defined, 3828–3829
- diagnosis and clinical features, 3831–3832
- differential diagnoses, 3837–3838
  - physical findings, 3837–3838
  - psychiatric considerations, 3838
- epidemiology, 3831
- etiology, 3831
- evaluation of, 3833–3834
- factitious disorder imposed on another, 3839
- federal law definition, 3828
- female genital mutilation, 3839
- fetal abuse and, 3838
- future directions, 3842
- genetics, 3833
- history, 3829–3830
- laboratory examination, 3837
- long-term consequences of, 3832
- multidisciplinary approach to, 3840
- neuroimaging and, 3833
- parent interviews, 3834
- physical abuse, 3829, 3831–3832, 3834, 3836–3837
- physical examination, 3836–3837
- prevention, 3841–3842
- religious practices, 3839–3840
- reporting of, 3840
- sex trafficking, 3839
- sexual, 3829, 3831–3834, 3837, 3841
- special issues, 3838–3839
- special tests, 3837
- state law definition, 3828–3829
- treatment
  - of child, 3840–3841
  - for incestuous behavior, 3841
  - of parents, 3841

- Mammalian target of rapamycin (mTOR), 126*f*
- Mammillary bodies
  - anatomy of, 34*f*, 36

- coronal section, 11*f*
- Managed behavioral health care
  - overview of, 4312–4313
  - private sector, 4313
  - public sector, 4313
- Managed care. *See also* Health maintenance organizations (HMOs)
  - cost containment with, 4311–4312, 4336
  - growth of, 4293–4294
  - hospitalist use in, 4357
  - psychotherapy under, 2828–2829
  - quality improvement in, 4314, 4349
- Mandatory outpatient treatment, 4396
- Mandela, Nelson, 2416
- Manerix. *See* Moclobemide
- Manganese, beneficial effects, 2549
- Mania (manic syndrome), 1600–1601. *See also* Bipolar mania; Somatic symptom and related disorders
  - acute
    - antipsychotics, 1686
    - atypical antipsychotics, 1686–1687, 1686*t*
    - in bipolar I disorder, 1648–1649
    - combination/add-on therapy, 1687
    - depressive symptoms, 1687
    - ECT in, 1690
    - emergency management, 1688
    - first-line treatments, 1688–1689, 1689*t*
    - general management, 1687–1688
    - miscellaneous agents, 1687
    - second-line treatments, 1689–1690, 1689*t*
    - third-line treatments, 1689*t*, 1690
    - treatment selection, 1688
    - treatments not recommended, 1690
  - after stroke, 485
    - diagnosis of, 481
    - etiology of, 479
    - treatment of, 487
  - after TBI
    - in children, 653

- ancient Greeks and Romans on, 1601, 1602
- animal models of, 322
- antidepressants and, 1691
- bipolar I disorder
  - treatment, 1696
- chronic, in bipolar I disorder, 1649
- classification of, 1168, 1828
- cognitive distortions in, 1641
- corticosteroids and, 2263
- depressive disorder, child/adolescent and, 3677, 3691
- description of, 4593
- differential diagnoses, 2625*t*
- episodes, 3718
- episodes, defined, 3686, 3700
- following stroke, 476
- gambling and, 1805
- hypomania *versus,* 1641
- incidence of, 465
- mixed or dysphoric, 1650
  - treatment, 1687
- mood disturbance in, 1640–1641
- OCD *versus,* 1788
- psychoanalytic formulations of, 1663–1664
- psychopathology, 1639–1641, 1639*t*
- psychotic, 459
- reactive, 1649
- SSRI, side effects, 3718
- suicide risk, 1669
- tamoxifen for, 171
- in transplant populations, 2366
- treatment
  - carbamazepine and, 3030–3031, 3030*f*
  - ECT, 3286
  - FGA, 3070
  - lithium, 3089
  - valproate, 3268, 3269*t*, 3270, 3270*t*, 3276
- tryptophan depletion and, 2550

- vilazodone and activation of, 3345
- Mania (Ms) Scale, MMPI, 999*t*
- Maniacal grief, 1632
- Manic, definition of, 4593
- Manic dementia, 1649
- Manic-depressive illness, 4593
- *Manic-Depressive Illness* (Goodwin), 4546
- Manic-depressive psychosis. *See* Bipolar disorder(s) (BD)
- Manic episodes
  - anabolic steroid-induced, 1399–1401, 1400*t*, 1401*f*
  - in patients with HIV, 572–575
- Manic syndromes after TBI, 534–535, 535*f*
- Manifest content, 4593
- Manipulation, definition of, 4593
- Mann, James, 2647
- Mannerisms, description of, 4593
- *Man's Search for Meaning* (Frankl), 932
- Mansfield, Peter, 259
- Mantras
  - description of, 4593
  - reciting of, 2557–2558
- Manual assisted cognitive behavioral therapy (MACT), 2617
- MAOIs. *See* Monoamine oxidase inhibitors (MAOIs)
- Maple syrup urine disease (MSUD), 661–662
- Maprotiline (Ludiomil)
  - chemistry, 3254, 3254*f*
  - drug studies, 3259
  - organ/system effects of, 3260
  - pharmacological actions, 3254
- Marasmus, description of, 4593
- Marathon, description of, 4593
- Maraviroc, mechanism of action, 567
- Marble burying test, animal models of, 321
- Marginalization, migration and, 792
- Marijuana. *See also* Cannabis; Cannabis dependence; Cannabis use and cannabis-related disorders
  - abuse of, 4115
  - brain development and, 58

- depersonalization/derealization, 1887, 1891, 1892, 1943
- description of, 4593
- EEG effects, 253–254
- intoxication, 2631
- legislation on, 4555
- neuropeptide Y and, 176
- psychosis and, 114–115
- sexual dysfunction and, 1973
- spirituality and, 2581

- Marital counseling, 4593
- Marital problems, 4593
- Marital schism, 2749
- Marital skew, 2749
- Marital status
  - gender change and, 2037
  - mood disorders and, 1616
  - of older people, 3949*t*
  - as psychosocial factor, 4165
- Marital therapy. *See* Marriage therapy
- Market structures, health care reform and, 4324
- Markowitsch, Hans J., 1893
- Marmar, Charles, 2718
- Marmer, Steven, 1903
- Marplan. *See* Isocarboxazid
- Marriage, unconsummated, 1974
- Marriage therapy, 4593
  - disulfiram combined with, 3052
- Marsh, L. Cody, 2736
- Marsilid. *See* Iproniazid
- Martin, Emily, 2839
- Martínez-Taboas, Alfonso, 1871
- Maryland Assessment of Social Competence (MASC), 4372
- Masculine identity, 4593
- Masculine protest, 4593
- Masculinity–Femininity (Mf) Scale, MMPI, 999*t*, 4593
- Masculinization
  - incomplete, 2505

- *Mash1* gene, 51
- Masked depression
  - about, 1634
  - pain with, 1841
- Masking
  - in bipolar 1 disorder, 699
  - visual backward, 695–696, 695f
- Maslow, Abraham
  - humanistic psychology and, 2843
  - personality theory, 932
  - transpersonal psychotherapy, 2825
- Masochism
  - description of, 4593
  - factitious disorder imposed on self and, 1851
  - sexual dysfunction and, 1975
- Masquerade syndrome, 1859
- Mass shootings
  - epidemiology of, 2417–2418
  - as violence against children/adolescents, 4298
- Massachusetts General Hospital Hairpulling Scale, 1795
- Massage
  - as somatic therapy, 2556
  - somatization disorder, 1836
- Massively multiplayer online role-playing games (MMORPGs), 3373–3374
- Massively parallel sequencing. *See* Next-generation sequencing (NGS)
- Masson, Jeffrey M., 2645
- Masters, William, 1953, 4542
- Masterson, James, 4545
- Mastery, factitious disorder, 1851
- Masturbation
  - attitudes toward, 1960
  - description of, 4593 (*See also* Autoerotism)
  - Kinsey on, 1954–1955
  - as primal addiction, 2052
  - training for sexual addiction, 2056
- Masturbatory pain, 1975
- Match (National Resident Matching Program), 4406–4407

- match rates, 4407*t*
- residency position funding, 4407, 4407*t*
- Material social interactions, 2820
- Maternal behavior
  - neglect, aggression and, 2485
  - rejection, aggression and, 2485
  - steroid receptor genes and, 166
  - stress reactivity, 394–396
- Maternal care deficiencies
  - addictive disorders and, 2047–2048
- Maternal deprivation, 3629, 3643, 3663, 3677
  - description of, 4593
- Maternal immune activation (MIA), 201
- Maternal separation
  - in infancy, addictive disorders and, 2047
- Maternally-inherited Leigh syndrome (MILS), 670
- Maternity blues, 1632
- *Math1* gene, 51
- Mathematical disability. *See* specific learning disorder
- *Matricaria recutita,* 2554
- MATRICS (Measurement and Treatment Research to Improve Cognition in Schizophrenia), 1531
- Matrix metalloproteinases (MMPs), 381
- Mattis Dementia Rating Scale-2
  - of aging adult, 4004
  - of schizophrenia, 4038
- Maturational crises, 4593
- Maturity
  - in adulthood, 3941–3942
  - as mental health, 758–759, 758*f*
- Maximum likelihood estimation, 831
- Maximum security units, 4593
- May, Rollo, 1963, 2821
- Mayer-Gross, Wilhelm, 1870
- *MBP* gene, anxiety and, 1740
- McAdams, Dan, 933
- McConkey, Kevin, 2709, 2732
- McCulloch–Pitts networks, 412

- McCullough, James, 1665, 2700
- MCE (medical care evaluation). *See* Medical audits
- McGill Illness Narrative Interview, 2447, 2461–2462
- McGill Pain Questionnaire (MPQ), 1003*t*
- McHugh, Paul, 1867
- McLean Asylum for psychiatric care, 4290
- McLeod, John, 2831
- McLeod syndrome, 667
- McNamara, Ashley, 2835
- MDD. *See* Major depressive disorder
- MDMA. *See* 3,4-methylenedioxy-*N*-methamphetamine (MDMA, "ecstasy")
- Meal(s). *See also* Diet and nutrition; Food
  - chemical signals in, 358*t*
  - as a functional unit, 355
  - initiation of, 355–358
  - maintenance, 358–360
- Mean, definition of, 830
- Mean deviation, 4593
- Meaning and purpose, empirical findings, 2847–2848
- Meaningful use. EHRs, 4320
- Means (mathematical), 4593
- Meares, Russell, 2723
- Measles–mumps–rubella (MMR) immunization, 3579
- Measurement instruments
  - reliability, 816–817
  - validity
    - construct, 818
    - content, 817
    - criterion, 817–818, 817*t*
- Mebanazine, history, 3110
- MECA. *See* Methods for the Epidemiology of Child and Adolescent Mental Disorders
- Mecamylamine, 74
- Meclizine (Antivert), 75
- The media
  - aggressive behaviors and, 2485
  - dissociative disorders and, 1871–1873, 1873–1874
  - on dissociative identity disorder, 1908

- suicide prevention and, 2617
- Medial dorsal nuclei, 21
- Medial forebrain bundle (MFB), 1714
- Medial ganglionic eminence (MGE)
  - *Nkx2.1* expression, 50–51
  - transcription factors expressed in, 51
- Medial migratory system (MMS), 60
- Medial orbitofrontal cortex projections, 449
- Medial posterior parietal cortex, 7*f*
- Medial prefrontal cortex (MPFC)
  - dopamine and stress to, 1745
  - emotional regulation and, 1818–1819
  - perspective taking and, 337
  - subgenual, 5
  - surface landmarks, 7*f*
- Medial preoptic area (MPOA), 395
- Medial temporal lobe
  - dissociative amnesia and, 1893
  - drives and instincts and, 2815
  - memory suppression and, 1893
  - neuroimaging, 2815
  - sagittal sections of, 6*f*
- Median, definition of, 830, 4593
- Medicaid. *See also* Finances; Health insurance; Medicare
  - ACA expansion of, 4379–4380
  - definition of, 4593
  - Early Periodic Screening Diagnosis and Treatment, 4304
  - funding for IMD residents, 4335
  - immigrant waiting period, 4379
  - intermediate care facilities coverage by, 4335–4336
  - managed care in, 4294, 4313, 4336
  - for mental health care, 4288, 4289
  - psychotherapy, race and ethnic differences, 803
- Medical assessment and laboratory testing, **1008–1026**
  - brain imaging role in diagnosis, 1018–1019
    - magnetic resonance imaging, 1019
  - diagnostic assessment

    - autoimmune and inflammatory disease, 1015, 1018–1019
    - endocrine evaluation, 1013–1015
    - infectious diseases, 1011–1013
    - liver function, 1010
    - neurological disorders, 1015
    - renal function, 1010
    - serum electrolytes, 1010
  - electroencephalogram role in diagnosis
    - evoked potential, 1020
    - polysomnography, 1020
  - genetic disorders, 1016
  - head trauma, 1016
  - intoxication/toxic exposures
    - acetaminophen or salicylate overdose, 1018
    - alcohol, 1017–1018
    - drug abuse, 1017
    - environmental toxins, 1018
    - volatile solvents, 1018
  - neoplastic disorders, 1015–1016
  - nutritional deficiency, 1016
  - physical health monitoring, 1008–1009
  - sensitivity, specificity, and predictive value, 1025–1026
  - treatment selection and monitoring, **1020–1026**
    - electrocardiography, 1020–1021
    - electroconvulsive therapy, 1020
    - metabolic monitoring, 1021–1022
    - psychotropic medications, 1022–1025
- Medical audits, 4594
- Medical care evaluation (MCE). *See* Medical audits
- Medical comorbidities, hospitalist role in, 4356
- Medical conditions (concurrent)
  - anxiety disorders and, 1735
  - dementia and, 1198
  - family therapy and, 3786
  - mood disorders and, 1656–1658, 1656*t*
    - suicide risk, 1669–1670
  - posttraumatic stress disorder (PTSD) and, 3706

- with schizophrenia, **1538–1548**
  - mortality rates, 1434–1435, 1538–1542, 1539*t*
  - oral, 1542
  - overview, 1538
  - prevention, 1544–1547, 1545*f*
  - recommendations, 1547–1548
  - risk factors, modifiable, 1542–1544, 1544*t*
- suicide behavior risk factor, 3683, 3697

- Medical decisions, competency and, 4235
- Medical education. *See also* Psychiatric education
  - humanism in, 2289
  - residency program reforms, 4307
  - residency training and, 4414
  - spirituality in, 2584
- *Medical Education in the United States and Canada* (1910), 4323
- Medical ethics. *See also* Ethical issues
  - codes of, 3869
  - description of, 4594
  - in hospitals, 2289
- *Medical Ethics* (Percival), 4240, 4441
- Medical Expenditure Panel Survey, 803
- Medical illnesses
  - depression in, 3237–3238
  - lithium and, 3095
  - parasomnia, 2102
  - personality changes due to, 2157–2158
  - physician, 2602
  - psychiatric disorders concurrent with
    - FGAs and, 3072
    - psychopharmacology and, 2920–2921
  - psychiatric morbidity and, 2181
  - psychopharmacology, 2924–2925
  - transgender people and, 2036
- Medical impairment, CBT and, 4180
- *Medical Inquiries and Observations Upon the Diseases of the Mind* (Rush), 4449
- *Medical Inquiries and Observations upon the Diseases of the Mind* (Rush), 4534
- Medical necessity threshold, 4289

- Medical Neighborhoods, components of, 4318
- Medical records. *See also* Electronic medical records (EMRs)
  - description of, 4594
  - in jails and prisons, 4476
- Medical reviews, 4594
- Medical savings accounts, provisions of, 4311
- Medical students
  - mental health, 2602–2605
    - administrative–clinical interface and, 2595–2596, 2595*f*
    - case studies, 2595–2596
    - disciplines involved, 2594
    - future directions, 2609
    - treatment approaches/alternatives, 2606
    - treatment barriers, 2606–2608
  - minority, 2598–2599
  - personality types, 2597–2598
    - case studies, 2598
  - preventive health, 2596–2597, 2597*f*
  - suicide, 2600
    - case study, 2600–2601
  - temperament, 2598
  - wellness, 2596–2597, 2597*f*
  - women, 2598
- Medically unexplained symptoms, 1827. *See also* Somatic symptoms
- Medicare, **4214–4218**. *See also* Finances
  - billing data from, 4308
  - description of, 4594
  - hospital insurance (Part A), 4214–4216
  - insurance coverage supplement, **4214–4218**
  - risk in managed care, 4217–4218
  - shared savings models, 4308
  - supplemental medical insurance (Part B), 4216–4217
- Medicare Access and CHIP Reauthorization Act (MACRA), 4218–4219
- Medicare/Medicaid
  - dual eligible populations, 4307, 4308, 4317
  - privatization of services, 4329
- Medicare Modernization Act (2003), 4307

- Medicare Prescription Drug, Improvement, and Modernization Act (MMA)
  - as incremental reform, 4311
  - shared savings models in, 4308
- Medication-induced movement disorders, 4594
- Medication(s). *See also* Botanicals; Drug therapy; Drug(s); Herbal medications; Pharmacology; Pharmacotherapy; Psychopharmacology
  - ADHD, neuroimaging, 3464
  - adherence to, 4362
  - availability of, 4525
  - child, laboratory testing, 3408
  - clinically questionable use of, 4352
  - and cognitive-behavior therapy, 4189–4190
  - cost of, 4522, 4524
  - for dementias in long-term care, 4226–4229
  - disorders induced by, 1198, 1797
  - dissociative identity disorder and, 1928
  - documented rationale for, 4337–4338
  - ECT interactions with, 3294–3295
  - fear of, physician mental health and, 2608
  - for intellectual disability (ID), 3514–3516
  - involuntary, legal aspects, 4232–4233
  - and malnutrition, in long-term care, 4226
  - mood disorders, 1928
  - movement disorders induced by, **2936–2945**
    - acute akathisia, 2938, 2941, 2955
    - acute dystonia, 2940–2941
    - antidepressant discontinuation syndrome, 2943
    - CNS control of movement, 2936, 2938
    - diagnostic criteria for, 2936, 2937*t*
    - future directions, 2944
    - history of, 2936
    - impact of, 2944
    - neuroleptic-induced parkinsonism, 2938
    - neuroleptic malignant syndrome, 2938–2940, 2939*t*, 2940*t*
    - other, 2943
    - postural tremor, 2942
    - tardive akathisia, 2941, 2942–2943
    - tardive dyskinesia, 2941–2942

    - tardive dystonia, 2942–2943
  - off-label use of, 4352
  - during psychiatric evaluations, 3440–3441, 3441*t*
  - psychoanalysis and, 2658
  - psychoanalytic psychotherapy and, 2664
  - psychotropic (*See* Psychopharmacology)
  - PTSD, dissociative identity disorder and, 1927–1928
  - safety, with dementia, 1200
  - and sexual dysfunction, 4264, 4264*t*
  - stabilization with, 4367–4368
  - tic disorders, 3644–3648, 3647*t*, 3658–3662, 3661*t*
- Mediodorsal nucleus (MD)
  - aggression and, 2476
  - projections from, 24*f*
- Meditation
  - for adjustment disorder, 2123
  - for adolescents, 3814
  - for cancer pain, 2275
  - CBT treatments and, 1773
  - discrete behavioral states model and, 1899
  - for dying patients, 2302
  - forms of, 2062
  - in functional GI disorders, 2207
  - goals of, 2059
    - in addictions, 2062
  - heart rate response to, 2327
  - neuroimaging during, 2062
  - psychotherapy and, 1701
  - “Silence of the Mind,” 2168
  - as somatic therapy, 2556–2558
  - spirituality and, 2578, 2594
  - transient depersonalization, 1887
- Medium spiny neurons, 29
- Medroxyprogesterone acetate (Cyrin, MPA, Provera, Depo-Provera), 1980
  - sex offenders, 2503
    - recidivism and, 2503
  - for sexual addiction, 2056

- Medulla, surface landmarks, 7*f*
- MEG. *See* Magnetoencephalography
- Megalomania, 4594
- Megavitamin therapy. *See* Orthomolecular treatment
- Megestrol acetate (Megace), 2265
- Meissner, William, 2585
- Melancholia. *See also* Somatic symptom and related disorders
  - ancient Greeks and Romans on, 1601–1602
  - classification, 1828
  - description of, 4594
  - neurobiology and, 1709
- Melancholic features
  - major depressive disorder and, 1644
  - treatment selection and, 1679
  - use of term, 1631
- Melanie Blocker Stokes MOTHERS Act, 2517
- Melanocortin, studies of, 168–169
- Melanocortin receptor agonists, 3340
- Melanocyte-stimulating hormone (MSH)
  - immunological effects of, 185*t*
  - studies of, 168–169
- Melanocytes, formation of, 40
- Melanopsin
  - mouse, 209*f*
  - photosensitive cell, 208
- MELAS
  - diagnosis of, 669–670
  - histology, 670*f*
  - MRI imaging, 670*f*
  - presentation, 669–670
- Melatonin
  - biological treatment with, 2555
  - chemistry, 3101–3102, 3102*f*
  - circadian rhythms and, 204
  - depressive effect, 169
  - dosage/administration, 3106
  - drug interactions, 3106, 3106*t*

- drug studies, 3103–3104
- history, 3101
- insomnia, 3101, 3103–3104, 3105
- investigational drugs, 3106–3107
- laboratory interference, 3106
- nocturnal, 208
- organ/system effects of, 3104–3105
- pharmacological actions, 3102–3103, 3103*t*
  - pharmacodynamics, 3103
  - pharmacokinetics, 3102–3103, 3103*t*
- precautions/adverse reactions, 3105–3106
- receptor agonists, **3101–3107**
- roles of, 69
- secretion of, 344
- sleep and, 1716
- for sleep-wake disorders and, 2098
- synthesis of, 68*f*, 204
- therapeutic indications, 3105, 3105*t*

- Melbourne Personal Assessment and Crisis Evaluation (PACE) study, 1470
- *Melissa officinalis,* 2554
- Mellaril. *See* Thioridazine
- Melperone (Buronil), 3060
- Melzack, Ronald, 2726
- Memantine (Namenda), **3048–3049**
  - AD, 3039, 3041, 3048–3049
  - for aggressive behavior
    - in Alzheimer disease, 2502–2503
  - Alzheimer disease treatment, 1204, 1207–1209
  - amnestic disorder treatment, 1231
  - chemistry, 3048, 3048*f*
  - cholinesterase inhibitors and, 3039–3050
  - dementia with Lewy bodies treatment, 1217–1218
  - for dementias in long-term care, 4226
  - dosage/administration, 3049
  - drug interactions, 3049
  - drug studies, 3048–3049
  - future directions, 3049

- laboratory interferences, 3049
- NMDA receptor antagonist, 4145–4146
- OCD, 1789
- organ/system effects of, 3049
- for PD, 547–548
- pharmacological actions, 3048
- precautions/adverse reactions, 3049
- PTSD and SUD comorbidity, 1823
- structure of, 77*f*
- therapeutic indications, 3049

- Membranes
  - active properties of, 132–133
  - passive properties of, 132
- *Memoire sur la decouverte du magnetism animal* (Mesmer), 4534
- Memoirs, mental illness, 2833–2835
- Memorial Delirium Assessment Scale (MDAS)
  - assessment of delirium, 2284
  - description of, 2256
  - in symptom assessment, 2292
- Memorial delirium rating scale (MDRS), 1188*t*
- Memory. *See also* Aging, Demographics and Memory Study (ADAMS); Amnesia(s); Autobiographical memory; Brief Visuospatial Memory Test; Childhood autobiographical memory disturbances (CAMD); False memories; Green's Word Memory Test (WMT); Immediate memory; Metamemory; Pseudomemories; Recovered memory; Recovered memory therapy (RMT); Tests of Malingering and Memory (TOMM); Traumatic memories
  - age-related problems, 214
  - of aging adult, 4005–4006
  - amnesia and, 463
  - amygdala and, 1757
  - aversive, PTSD and, 1817
  - BDNF deficiency and, 100
  - biology of, 728–744
    - amnesia, 730–734, 731*f*, 732*f*, 733*f*, 733*t*
      - retrograde, 733–734, 734*f*
    - associative learning, 729
    - cortical organization, 729–730, 730*f*
    - cross-reference, 743

- functional assessment, 739–742
  - infantile amnesia, 743
  - memory distortion, 742–743, 742*f*
  - new learning capacity, 739–741, 740*f*
  - psychogenic amnesia, 741–742, 742*f*
  - remote memory, 741
  - self-reports, 741
  - and unconscious, 743
- long-term potentiation, 729
- neuroimaging, 736–738, 737*f*, 738*f*
- plasticity, 728–729
- in psychiatric illness, 738–739, 738*f*
- sleep effects, 739
- subtypes, 734–736, 735*f*, 736*f*
- brain systems, 690–691, 691t
- circadian organization of, 214
- declarative, 731, 734–736, 739
- deficits
  - after TBI, 529
  - FASD and, 638
- dementia and, 1194–1195
- in depression, 698
- description of, 4594
- dissociative amnesia and, 1879
- distortion
  - personality and, 2733
- disturbances of, **1133–1135**
  - in recall, 1134–1135
  - in registration, 1133
  - in retention, 1134
- emotion and, 1757
- episodic, 735
  - description of, 2139
- and HIV infection, 4269
- hypnosis and, 2708
  - target event, 2734
  - waking baseline, 2733

- impairment
  - from alcohol, 1267
- impairment in Alzheimer disease, 4083
- long-term declarative
  - description of, 2139
- neurocognitive impairment and, 1503–1504, 1507
- neuroimaging
  - drives and instincts and, 2815
- neuropsychological assessment of, 3433
- in posttraumatic stress disorder, 700
- procedural
  - associative learning and, 2139
- responses to stress, 2318
- retention and, 463
- retention of, 463
- retrieval
  - forensic, 2733–2734
- schizophrenia and, 1503–1504, 1507
- semantic, 735–736
- sleep and, 353
- specific learning disorder and, 3525–3526
- stimulant-related disorders and, 1283–1284
- suppression or repression, neurobiology of, 1893
- tasks on fMRI, 275
- and telemedicine, 4276
- types of, 2129
- verbal, 463

- Memory disorders
  - hemispheric dominance
    - encoding, 981, 981*f*
    - storage and retrieval, 981
  - neurosteroids and, 118
  - in patient interview, 967
- Memory hardening position, 2733
- Memphis model
  - police use by, 4393
- Men. *See also* Gay persons

- exhibitionistic disorder in, 2016–2017
- paraphilias in, 2015
- self-esteem development, 3370

- Menarche, description of, 4594
- Mendacity, definition of, 4594
- Mendel laws of inheritance, 149
- Meningovascular syphilis, 592
- Menninger, Karl
  - on conduct disorder, 3612
  - on factitious disorder, 1847
  - *The Selected Correspondence of Karl Menninger,* 4546
  - on sin, 2587
- Menninger Foundation, 2673, 2674
- Menninger Psychotherapy Project, 755
- Menopause
  - overview, 2521–2523
- *Mens rea,* 4594
- Menstrual cycles
  - functional hypothalamic anovulation, 2507–2508
  - physiology of, 2507
  - premenstrual syndrome, 2507
- Mental aberrations, definition of, 4594
- Mental age, description of, 4594
- Mental deficiencies, description of, 4594
- Mental disorders
  - classification systems for, **1151–1176** (*See also Diagnostic and Statistical Manual of Mental Disorders* (DSM))
    - alternatives to, 1161–1163, 1161*t*
    - history of, 1152–1154, 1153*f*
    - ICD-DSM comparisons, 1167–1169, 1167*t*, 1174–1176
    - medical model approach and, 1151, 1154–1155
    - of North America, 1154
    - objectivity of, 1154–1155
    - term-use in, 1151–1152
  - defined, 1152, 1169–1170, 3348–3349
  - description of, 4594
  - DSM-5 classification of, 1156–1159, 1157*t*, 1174–1176
  - DSM classification

- advantages of, 1155
- complexity of, 1155–1156
- concerns with, 1159, 1160*t*
- evolution, 1152, 1153–1154, 1153*f*
- global issues for, 1159–1160
- ICD, comparisons with, 1167–1169, 1167*t*, 1174–1176
- ICD, harmonization with, 1175–1176
- revisions of, 1156–1159, 1157*t*, 1160–1161

- due to another medical condition
  - anxiety disorder, **1253–1255**
  - catatonic disorder, **1258–1259**
  - clinical features, 1233–1234, 1234*t*, 1235*t*
  - comparative nosology, 1233, 1233*t*
  - course, 1236
  - defined, 1232
  - delirium due to, 1178
  - diagnosis, 1233–1236, 1235*t*
  - etiology, 1233, 1234*t*
  - history, 1232–1233
  - mood disorders, **1248–1251**
  - neurocognitive disorders, **1246–1247**
  - overview, 1232
  - prognosis, 1236
  - recording procedures for, 1236
  - treatment, 1236
- ICD classification of
  - DSM, comparisons with, 1167–1169, 1167*t*, 1174–1176
  - DSM, harmonization with, 1175–1176
  - evolution of, 1164–1166, 1166*t*
  - frequency of use of, 1168–1169
- ICD-11 classification, 1169–1176
  - -PHC, 1171
  - comparisons with, 1174–1176
  - field trials for, 1170–1171
  - overview, 1163–1164
  - proposals for, 1171–1174, 1174*t*
  - revisions for, 1169–1170, 1171*t*

    - standardization of, 1170, 1171*t*
    - structure of, 1170, 1170*t*
  - neurodevelopmental models of, 409–410
  - Research Diagnostic Criteria (RDC) classification, 1154
  - suicide risk factor, child/adolescent, 3682–3683, 3696–3697
  - transition from adolescents to adulthood, 3851, 3852*t*
  - WHO-FIC classification, 1164, 1164*t*
- Mental functions, neuroimaging of, 4486
- Mental health. *See also* Joint Commission on Mental Illness and Mental Health; World Mental Health Initiative
  - children and adolescents
    - suicide and, 2614
  - description of, 4594
  - global burden, 797, 798*t*
  - healthy involuntary mechanisms, 761
  - and immigration, 803–806, 804*t*, 805*t*
  - medical students, 2602–2605
  - mental illness and, 2844–2845, 2844*f*
  - normality of (*See* Normality)
  - physician, 2599–2602
    - aging, 2605–2606
    - future directions, 2609
    - midcareer, 2605
    - treatment approaches/alternatives, 2606
    - treatment barriers, 2606–2608
  - positive, 769, 2843–2845
  - socioemothonal intelligence and, 761–762
  - "spiritual" emotions and, 765–766
  - subjective well-being and, 762–765
  - transgender people and, 2036
- *Mental Health: A Report of the Surgeon General,* 4294
- Mental Health Assessment Form, 3402
- Mental Health Association (MHA), 956
- *Mental Health Atlas 2005,* 4524
- Mental health care. *See also* Community mental health centers (CMHCs)
  - access to
    - crisis/imbalance of, 4306
    - disparities, 4304

    - limitations, 4313
  - after arrest, 4394–4395
  - in criminal justice system, 4392
  - ethical issues in, 4530–4531
  - funding limitations of, 4321
  - physical health care integrated with, 4530
  - post-incarceration, 4305
  - in primary care, 4295–4296, *4295f*, 4381
  - privatization of, 4329
  - program development principles for, 4525–4528
  - recidivism, 4340
  - reimbursement parity for, 4289, 4350, 4380
  - safety features of, 4395–4396
  - shortcoming of, 4288–4289
  - for special populations, 4343–4344
  - steps in
    - assessments, 4337
    - formulation, 4337
    - treatment planning, 4337–4339
    - treatment, 4339–4340
    - discharge planning, 4340–4341
  - for those at risk of becoming criminalized, 4387
  - for treatment *vs.* incarceration, 4395
  - world-wide, *4526–4527t*
- Mental health courts
  - case management by, 4329–4330
  - role of, 4394–4395
  - supervision after criminal justice system discharge, 4392
- Mental health disorders
  - cost of, 4288
  - physical health problems comorbid with, 4317
  - research into, 4291
  - substance use disorders comorbid with, 4288–4289, 4492
- Mental Health Parity and Addiction Equity Act (MHPAEA), 4212, 4317, 4380
- Mental health policy, 4554
- Mental health professionals
  - police, integration of, 4393

- world-wide, 4525, *4528t*
- Mental health service grants, *4288t*
- Mental health services
  - adulthood, 3942
  - child, **3900–3913**
    - accessibility of, 3910–3911, 3910*f*
    - conceptual models, 3910, 3910*f*
    - history, 3901, 3902–3907*t*, 3908–3909
    - quality of care, 3911–3913, 3912*t*, 3913*t*
    - stigma of, 3910, 3910*f*
  - courts, 1553–1554
- Mental Health Services Program for Youth, 4321
- Mental health services research
  - overview of, 4346–4347, 4353
  - scope of, 4347–4353
- Mental Health Study Act (1955), 4292
- Mental Health Systems Act (MHSA)
  - radical reform authorized by, 4294
  - Reagan's repeal of, 4293
- Mental hygiene, definition of, 4594
- Mental illness memoirs, 2833–2835
- Mental illnesses. *See also* Criminalization; *specific disorders*
  - arrest and, 4388
  - burn injuries and
    - risk of, 2374
  - criminality and, 2489
  - description of, 4594
  - medical comorbidities with, 4343
  - neurosteroids in, 117–118
  - obesity and, 2214–2215
  - problem recognition as barrier, 4384
  - as public health issue, 4302
  - severity world-wide, *4519–4520t*
  - substance use and
    - in adolescents, 646
    - in prenatally acquired HIV patients, 647–648
  - violent behavior and, 2490

- Mental images, cognitive development and, 707
- *Mental Maladies: A Treatise on Insanity* (Esquiro), 4535
- Mental models of consciousness, 691–692
- *Mental Pathology and Therapeutics* (Griesinger), 4535
- Mental planning, executive functioning and, 3430–3431
- Mental processing speed, aging and, 4007
- Mental Research Institute brief family and couples therapy, 2751*t*
- Mental retardation (MR)
  - aggression and, 2487, 2488
  - aggressive behaviors and, 2487, 2488
    - treatment of, 2499–2500
  - classification of, 1158, 1173
  - description of, 4594
  - premature births and, 643
  - syndromes of, 472
- Mental Retardation Facilities and Community Mental Health Centers Construction Act (1963), 4291
- Mental states, cognitive development and, 694–695
- Mental status, definition of, 4594
- Mental Status Examination (MSE), 613
  - of aging adult, 3974
  - for child emergency psychiatric evaluation, 3439
  - for child psychiatry evaluation, 3381–3382, 3383*t*, 3385
  - in delirium, 4090
  - description of, 4594
  - of immigrants and refugees, 793
  - for infants and toddlers, 3392–3393*t*
  - neuropsychological testing, 988
  - in patient interview, 952
  - rating scales for, 3402
- Mental stress
  - acute
    - CAD and, 2188–2189
    - hypertension and, 2193
  - chronic, CAD and, 2189
- Mentalization-based treatment (MBT), 2668, **2894–2904**
  - affect focus, 2900
  - attachment, 2897–2898

- clinical issues, 2901–2902
  - case study, 2901–2902
  - complications, 2901
  - contraindications, 2901–2902
  - ethical issues, 2902
  - indications, 2901
  - limitations, 2901
  - treatment goals, 2902
- depression, hypnosis and, 2723
- dimensional mentalizing profile and BPD, 2896
- history, 2895
- nonmentalizing modes and BPD, 2896–2897, 2896*t*, 2897*t*
- research and evaluation, 2902–2904
  - antisocial personality disorder, 2904
  - BPD, 2902–2904
- suicide prevention and, 2621
- suicide risk, 2617
- techniques, 2898–2901
  - affects and affect identification, 2899–2900
  - clarification and exploration, 2899
  - clinical principles, 2898–2899, 2898*t*
  - empathetic validation, 2899
  - mentalizing the relationship, 2900–2901
  - presenting alternative perspectives, 2900
  - session trajectory, 2899–2901
  - therapeutic alliance, 2898, 2898*t*
- theoretical issues, 2895–2898

- Mentalizing
  - automatic *versus* controlled, 2895
  - definition, 2894–2895
  - internal *versus* external, 2895–2896
  - multidimensional nature of, 2895–2896
  - profiles, 2895
  - therapeutic role, 2895
- Meperidine (Demerol), 2255
- Meprobamate (Equanil, Miltown)
  - dependency diagnosis, 1384

- history, 2917, 2990
- precautions/adverse effects, 2987
- as sedative-hypnotic, 1384

- Mercury, toxic neuropathy and, 631
- Mercy booking
  - by police, 4391
- Meresa. *See* Sulpiride
- Merital. *See* Nomifensine
- Merkel cells, 2337
- *Merrell Dow Pharmaceuticals, Daubert v.,* 2410
- Merskey, Harold, 1844
- Merycism. *See* Rumination disorder, of infancy
- Mescaline
  - description of, 4594
  - pharmacology of, 1316
- Mesencephalon (midbrain)
  - embryology of, 41
  - in human development, 8
- Mesmer, Franz Anton
  - hypnosis, 2639, 2706
  - *Memoire sur la decouverte du magnetism animal,* 4534
- Mesmerism, description of, 4594
- Mesoderm, formation of, 40
- Mesomorphic, definition of, 4594
- Mesoridazine (Serentil)
  - cardiac side effects, 3075
  - chemistry, 3064
  - metabolism/elimination, 3065
  - side effects, 2199
- Messud, Claire, 2835
- MET. *See* Motivational enhancement therapy
- Meta-analysis
  - antidepressant effectiveness, 1678–1679
  - antidepressants for bipolar depression, 1691
  - cognitive therapy, 2862
  - description of, 4594
  - panic disorder treatment, 1775

- psychotherapy, 2861
- SAD treatment, 1776

- Metabolic disorders
  - antipsychotic side effects, 1527
  - differential diagnoses, 1586
  - EEG in, 246*f*
  - patterns of inheritance, 674*t*–675*t*
  - presentation as neuropsychiatric syndromes, 674*t*–675*t*
  - psychiatric presentations of, 467*t*
  - and psychiatric symptoms, 4048
- Metabolic syndrome
  - antipsychotic medications in concurrence with, 4343
  - characteristics of, 168
  - obesity and, 2230
  - with schizophrenia, 1543–1544, 1544*t*, 1546
- Metabolic tremor, 458
- Metabolism
  - aggression and, 2482
  - bipolar disorders and, 1609
  - metabolic testing and
    - of alcohol, 1266, 1269
    - antipsychotics, effects of, 3795
    - of caffeine, 1295
    - cannabis use and, 1309
    - for intellectual disability, 3493, 3501–3502
    - pharmacogenomics testing and, 3356, 3455
    - of stimulants, 1284
  - prefrontal cortex, sleep and, 1716
- Metabolism of, 1266, 1269
- Metabotropic receptors (mGluRs)
  - glutamate binding to, 1747
  - modulators, anxiety disorders, 1743*t*
- Metachromatic leukodystrophy (MLD)
  - CT findings, 657*f*
  - description of, 656
  - features of, 467*t*
  - schizophrenia-like psychosis and, 655

- Metacognition
  - in attention-deficit/hyperactivity disorder, 697
  - and consciousness, 692–693
- Metals, disordered metabolism of, 671–675
- Metamemory, 741
- Metapsychiatry, 4594
- Metapsychology, 2649, 4594
- Metclopramide, 2249
- Metencephalon (pons and cerebellum)
  - embryology of, 8, 41
- Metformin (Glucophage)
  - in PCOS treatment, 2249
  - weight loss and, 2237
- Methadone (Dolophine)
  - adverse effects, 3134
  - bioavailability, 2395
  - black box warning, 3134
  - chemistry, 3133, 3133*f*
  - delirium and, 1182
  - description of, 4594
  - dosing, 3134–3135
  - drug interactions and, 580, 3228
  - drug studies, 3134
  - duration of action, 2395*t*
  - history, 3133
  - opioid-dependence treatment, 1371
  - opioid-maintenance treatment, 3132, 3133–3135
  - opioid treatment, 1366
  - opioid withdrawal, 3132
  - opioid-withdrawal treatment, 1368–1369
  - overdose, 3134
  - pain management using, 2293, 2377
  - palliative use of, 2292
  - pharmacological actions, 3133–3134
- Methamphetamines aggression and, 2493
- Methandrostenolone (Dianabol), 2493
- Methaqualone (Quaalude)

- dependency diagnosis, 1385
- precautions/adverse effects, 2987–2988
- as sedative-hypnotic, 1385

- Methemoglobinemia, 2272
- *Method of Measuring the Development of the Intelligence of Young Children* (Binet), 4537
- Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA), 826
- Methohexital (Brevital), 1979
  - anesthesia induced by, 2986, 2987
  - ECT anesthesia with, 2986, 2987, 3292
  - pharmacological actions, 2984
  - therapeutic indications, 2986
- Methotrexate (Rheumatrex)
  - side effects
    - CNS toxicity of, 2272
    - delirium, 2255
- 3-Methoxy-4-hydroxyphenylglycol (MHPG)
  - description of, 4594
  - panic disorder and, 1742
- Methyl-6-adenosine (m^6A), 161
- Methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)
  - description of, 4594
  - targets for, 64
- Methyl-binding proteins (MBP), 392
- Methyl-cobalamin (B_{12}), 2547
- α-Methylacyl-CoA racemase deficiency, 663
- Methylation, description of, 149
- Methylation arrays, 439
- Methylene blue, 2272
- 3,4-Methylenedioxy-*N*-methylamphetamine (MDMA, “ecstasy”)
  - depersonalization/derealization, 1887, 1943
  - description of, 4577, 4594
  - EEG effects, 254
  - effects of, 1316, 2007
  - hallucinogen-related disorders, 1321–1323, 1324
  - intoxication, 1320, 1321
  - neurotoxic effects, 68

- SERT availability in users of, 302
- Methylenetetrahydrofolate reductase (MTHFR), 668–669
- Methylfolate
  - beneficial effects, 2548
  - schizophrenia and, 2548
- L-Methylfolinic acid
  - beneficial effects, 2547
- Methylmercury (MeHg), 57, 57*f*
- Methylphenidate (Ritalin, Concerta), **3219–3231**
  - ADHD, 3123, 3219, 3223, 3227–3229, 3593, 3599
    - child/adolescent, 3692–3693, 3706–3707, 3765, 3788, 3789, 3909
    - warnings, 3789
  - in ADHD, 250
  - for aggressive behavior, 2112
    - ADHD, 2500
    - TBI, 2502
  - antidepressants combined with, 3324
  - autism spectrum disorders, 3794
  - BD with ADHD, 1699
  - for cancer patients, 2264
  - for children, 650
  - dosage/administration, 3221*t*–3222*t*
  - drug interactions, 3227*t*
  - for fatigue, 2284
  - history, 3219
  - malingering and, 2403
  - mechanisms of action, 3223–3224
  - neuroimaging, functional, 3464
  - pharmacological actions, 3224–3227, 3225*t*–3226*t*, 3227*t*
  - precautions/adverse reactions, 3229–3231, 3229*t*
  - sexual dysfunction and, 1972
  - therapeutic indications, 3227–3229
  - tic disorders, 3647, 3661
  - tic disorders, child/adolescent, 3795
  - traumatic brain injury, 1244
  - treatment adherence to, 3909
  - use in cancer patients, 2262

- Methyprylon (Noludar)
  - dependency diagnosis, 1385
  - as sedative-hypnotic, 1385
- Methysergide
  - for migraine treatment, 624
  - side effects of, 625
- Metoclopramide (Reclomide)
  - for migraine, 625
  - side effects, 571
    - akathisia, 2271
    - NMS, 2381
- Metonymy, 4594
- Metoprolol (Lopressor)
  - absorption, 2951
  - chemistry, 2951, 2951*f*
  - distribution, 2952
  - dosage/administration, 2956
  - drug interactions, 2955
  - drug studies, 2952
  - in hyperthyroidism
    - symptom relief using, 2243
  - organ/system effects of, 2952–2953, 2953*t*
  - pharmacodynamics, 2952
  - pharmacokinetics, 2951–2952, 2952*t*
  - pharmacological actions, 2951–2952, 2952*t*
  - precautions/adverse reactions, 2955
  - teratogenic effects of, 2953
  - therapeutic indications, 2953–2955, 2953*t*
- Metrazol treatment, 4594
- Mexican Americans, 805–806. *See also* Race/ethnicity
- Mexiletine (Mexitil), 2198*t*
- Meyer, Adolf, 909f
  - habit training, 909
  - personality and psychopathology, 909
  - problems of psychiatry, 4538
  - psychiatric classification and, 1153
  - psychobiology of, 908–909, 4291

- on stress, 2317
- work of, 4594
- Meynert, Theodor, 2639
- MHA. *See* Mental Health Association
- MHPAEA. *See* Mental Health Parity and Addiction Equity Act
- MHPG. *See* 3-Methoxy-4-hydroxyphenylglycol (MHPG)
- Mianserin, EEG effects, 251–252
- Michaels, Robert, 2662
- Michigan Alcoholism Screening Test (MAST-R), 1003*t*
- Mickey Finn, 2990
- Microbiota, stressors and, 192
- Microcephaly, 4594
- Microcurrent electrical stimulation, 2546, 2558, 2560–2561
- Microdissociative episodes, 1912
- Microglia, 5
  - characteristics of, 4*t*
  - derivation of, 5
- Micropsia, description of, 4594
- MicroRNAs (miRNAs)
  - function of, 382
  - gene expression, regulation of, 51
- Microtubule-associated proteins (MAPs), 53
- MID. *See* Vascular dementia
- Midazolam (Versed)
  - absorption, 2993
  - abuse liability of, 1376
  - alcohol withdrawal, 2997
  - anesthesia and sedation, 2997
  - chemistry, 2992
  - metabolism, 2993–2994
  - for pain management, 2377
  - seizure disorder, 2997
  - withdrawal, 2998
- Midbrain. *See* Mesencephalon
- Middle adolescence, 3366–3367
- Middle adulthood, description of, 4595
- Middle age

- description of, 4595
- sexual behavior during, 1962–1963
- Middle insomnia, 4595
- Middleton, Warwick, 1871
- Midlife
  - crisis during, 4595
  - depression during, 3336
  - development in, 3942
  - employment during, 3944
  - homosexuality and, 2003
  - mental health, 767, 767*t*
- Mifepristone (RU486), 167
- MIG-99, 624
- Migraine. *See also* Headache(s)
  - acute management of, 625, 625*t*
  - with aura
    - diagnostic criteria, 620*t*
  - characteristics of, 619–620
  - clinical evaluation, 454–455
  - course of, 621
  - description of, 4595
  - differential diagnoses, 623
  - health-seeking in, 622
  - imaging studies of, 622
  - presentation of, 2391
  - prevalence of, 621
  - treatment of, 2391
    - anticonvulsants, 2970
    - nonpharmacologic, 625
    - pharmacologic, 624
    - preventive, 623–625
  - without aura
    - diagnostic criteria, 619*t*
- Migration
  - and natural selection, 772
  - psychiatric assessment of, 793
- Migulstat (Zavesca), 665–666

- Milacemide, 3113*f*
- Mild cognitive impairment (MCI), 1222–1232. *See also* Cognitive impairment
  - AD and, 1225, 1225*t*
  - assessment, 1224–1226, 1225*f*, 1226*t*
    - biomarkers, 1224, 1225, 1227
    - genetics and, 1225–1226, 1225*f*
    - neuroimaging, 1226, 1226*t*
    - neuropsychological, 1224–1225
  - clinical presentation, 1223*t*, 1224, 1225*t*
  - course, 1227, 1228*f*
  - criteria for, 1222–1223, 1222t, 1223*t*
  - defined, 1177, 1222
  - diagnosis, 1222–1223, 1222*t*, 1223*f*, 1223*t*
  - differential diagnoses, 1226–1227, 1227*f*
  - epidemiology, 1223–1224, 1224*t*
  - etiology, 1223–1224, 1225*f*
  - future research for, 1228–1229
  - historical perspective, 1222–1223
  - prognosis, 1227–1228
  - terms related to, 1222–1223, 1224*t*
  - treatment, 1228–1229, 1229*t*
- Mild mental retardation. *See* Mental retardation
- Mild neurocognitive disorder
  - dementia and, 1198
  - description of, 4595
  - diagnostic criteria, 614*t*
- Milieu therapy
  - description of, 4595
  - drug therapy with, 4292
- Military environments, suicide and, 2621
- Military history at patient interview, 967
- Military mental health care in peacetime, 2573–2574
- Military sexual trauma, 1873
- Military style boot camps, 2500
- Military survival school subjects, 1897. *See also* Special Forces soldiers
- Miller, George, 4536
- Miller, Jonathan, 460–468

- Miller, Neal, 930–931
- Miller, Thomas I., 2796
- Miller, Vivian, 2729
- Miller-Dieker syndrome, 52
- Miller Forensic Assessment of Symptoms Test (M-FAST), 1856
  - detection of malingering with, 2410
- *Miller v. Alabama,* 4440
- Miller's Drinker's Check-up, 1310
- Millon, Theodore, 935
- Millon Clinical Multiaxial Inventory (MCMI), 1899, 2161–2162, 2410
  - characteristics of, 998*t*
  - MCMI-III, 1905
  - MCMI-IV, 1001
- Milnacipran (Ixel, Savella)
  - antidepressant efficacy, 3152
  - anxiety disorders, 1780
  - chemistry, 3147*f*
  - depression treatment, 4124
  - for fibromyalgia, 635
  - history, 3151–3152
  - in pain syndromes, 2395
  - pharmacology, 3152
  - sexual dysfunction and, 1779*t*
  - tolerability, 3152
- Milrod, Barbara, 2804–2805
- Milwaukee Inventory for Subtypes of Trichotillomania, 1795
- Mimicry, description of, 4595
- Mind, brain and, 688, 689t
- *Mind and Brain* (Laycock), 4536
- *A Mind That Found Itself* (Beers), 4291, 4537
- Mind–body practices
  - in biomedicine, 2546
  - in older adults, 3979–3980
    - CAIM medicine, 3980, 3983
  - as somatic therapy, 2556–2557
- Mindful awareness, 688
- Mindful Meditation, 3814

- Mindfulness
  - in biomedicine, 2546
  - CBT, psychosis and, 2882
  - in cognitive-behavior therapy, 4185
  - and HIV infections, 4273
  - neurobiological effects of, 2814*t*
  - skills, in DBT, 2787–2788
  - training
    - as somatic therapy, 2556–2557
    - suicide prevention and, 2621
  - weight loss using, 2217–2218
- Mindfulness-based cognitive therapy (MBCT)
  - acceptance and commitment therapy, 2699
  - anxiety disorders, 1773
  - mood disorders, 1701, 1702
  - positive psychology and, 2853
  - as somatic therapy, 2557
- Mindfulness-based interventions, 1349
- Mindfulness-based stress reduction (MBSR)
  - acceptance and commitment therapy and, 2699
  - anxiety disorders, 1773
  - in functional GI disorders, 2207
  - as somatic therapy, 2556–2557
- Mindfulness meditation
  - concept of, 2579
  - psychotherapy and, 1701
  - PTSD, 1824
- Mindreading as cognitive distortion, 2696
- Mineralocorticoid receptors (MRs), 1744
- Minerals, beneficial effects, 2549
- Mini-International Neuropsychiatric Interview (MINI), 2685
- Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID), 3398
- Mini-Mental State Examination (MMSE)
  - of aging adult, 3975, 4004
  - assessment of delirium, 2284
  - cognitive function at patient interview, 955
  - cognitive function rating scale, 3382, 3383*t*, 3402

- cognitive history taking, 613
- delirium in, 1180, 1183, 4090
- dementias in long-term care, 4226
- and electroconvulsive therapy, 4155
- in psychosomatic medicine, 2182

- Minimal brain dysfunction, definition of, 4595
- Minnesota Multiphasic Personality Inventory (MMPI)
  - assessment of psychosomatic illness, 2179
  - description of, 4595
  - detection of malingering with, 2410
  - development, 4539
  - PD scales, 2161
  - on prison populations, 2489
  - scale development, 999*t*
- Minnesota Multiphasic Personality Inventory 2 (MMPI-2), 2686
  - characteristics of, 998*t*
  - clinical scales
    - computer scoring and interpretation, 1000
    - content scales for MMPI-2, 1000
    - Millon Clinical Multiaxial Inventory III, 1001
    - MMPI-2-RF, 1000
    - Personality Inventory for DSM-5-adult, 1000–1001
    - Personality Inventory-Revised, 1001
  - content scales, 1000
    - Harris & Lingoes, 1000
  - restructured (MMPI-2-RF), 1001
  - supplemental scales, 1000–1001
  - validity scales, 997, 999
    - backpage infrequency (FB) scale, 999
    - Variable Response Inconsistency Scale, 999
    - variable response inconsistency scale (VRIN), 999
- Minocycline
  - depression, 1681
  - in schizophrenia, 599

- Minor depression, description of, 1647*f*
- Minor epilepsy. *See* Petit mal epilepsy
- Minor mood disorders, 1601
- Minorities. *See also* Ethnocultural populations
  - clinical trials, under-representation in, 4382
  - community-based treatment for, 3808–3809
  - intellectual disability and, 3498
  - service design, 4381
  - services used by
    - behavioral health disorders, 4304
    - Children's Health Insurance Authorization Act (CHIPRA), 4380
    - Children's Health Insurance Program (CHIP), 4380
- Minuchin, Salvador
  - psychodynamic family therapy and, 2753, 2754
- Miranda rights, 4438
- Mirapex. *See* Pramipexole
- Mirror movements, 3539–3540
- Mirror neurons, 2675
  - for emotional development, 3361
  - and social cognition, 693
- Mirroring
  - psychoanalysis, 2660
  - psychodrama, 2856
- Mirtazapine (Remeron), **3107–3111**
  - adverse effects of
    - dermatological, 2356*t*
  - for amphetamine-type stimulant use disorders, 1290
  - antidepressant effects of, 3107–3108
  - anxiety disorders, 1746, 1779, 1780, 1782*t*
  - bupropion and, 3109
  - for cancer patients, 2253
  - chemistry, 3107*f*
  - combination pharmacotherapy and, 3325
  - contraindications, 2520
  - depression, 1679
  - depression treatment, 4124
  - dissociative identity disorder, 1927*t*

- dosage, 573*t*
- in functional GI disorders, 2206
- future directions, 3110
- for geriatric depression, 4074
- history, 3107–3108
- indications, 3109
- MDD, 3108–3109
- mechanism of action, 2261
- for nausea, 2271
- pharmacology, 3108–3109
- pregnancy, lactation and, 3109
- SAD, 1784
- safety, 3110
- sexual dysfunction and, 1779*t*
- side effects, 573*t*, 1679*t*, 1680, 2261
- SSRIs/SNRIs combined with, 3108–3109
- thyroid hormones and, 174
- tolerability, 3110
- toxicity, 3110
- venlafaxine and, 3108, 3109, 3110, 3325
- weight gain and, 2215, 2237

- Misidentification delusions, 1123
- Misidentification syndromes, 691
- Misleading questions, forensic, 2734
- Mismatch negativity (MMN), 256–257, 2130
  - harm avoidance and, 2145
- Missing data, 838
- Missouri Assessment for Genetics Interview for Children (MAGIC), 3397
- Mitchell, S. Weir, 4595
- Mitchell, Stephen, 2646
- Mitochondria, genome, 154–155
- Mitochondrial disorders, 669–671
- Mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS). *See* MELAS
- Mitogen-activated protein kinases (MAPKs), 125–126
- Mitosis, description of, 149
- Mitotane (Lysodren), 2272
- Mixed anxiety–depressive disorder, 1653

- Mixed dissociative (conversion) disorders, 1875*t*
- Mixed episodes
  - ancient Greeks and Romans on, 1602
  - definition, 1601
  - suicide risk, 1669
- Mixed mania, 1650
- Mixed receptive-expressive disorder, 3545. *See also* Language disorders
- Mixed transcortical aphasia, 462*t*
- Mixoscopia, description of, 4595
- MMPI. *See* Minnesota Multiphasic Personality Inventory (MMPI)
- MMSE. *See* Mini-Mental State Examination
- M'Naghten, Daniel, 4535
- M'Naghten insanity defense, 4439
- Moban. *See* Molindone
- Mobile crisis teams, 4393
- Mobile health, 4276
- Mobsters, organized crime and, 2413–2414
- MOCA. *See* Montreal Cognitive Assessment
- Moclobemide (Manerix)
  - anxiety disorders, 1779, 1780
  - chemistry, 3113*f*
  - history, 3112, 3234
  - SAD, 1784
  - sexual dysfunction and, 1779*t*
  - side effects, 1679*t*, 1680
  - social anxiety disorder, 3122
- Modafinil (Provigil), **3231–3232**
  - ADHD, 3232, 3595, 3789
  - BD with ADHD, 1699
  - bipolar depression, 1691, 1692*t*, 1693, 3237
  - chemistry, 3234, 3235*f*
  - cocaine-related disorders, 3232
  - cognitive enhancement, 3232
  - combination pharmacotherapy and, 3324
  - for daytime sleepiness, 2264–2265
  - dosage/administration, 3223*t*, 3241, 3241*t*
  - drug interactions, 3240–3241

- EDS, 3231–3232, 3238
- for fatigue, 2284
  - in dying patients, 2295
  - in multiple sclerosis, 2265
- MDD, 3236
- mechanism of action, 3231
- narcolepsy, 3231–3232
- organ/system effects of, 3239
- performance enhancement, 3238
- pharmacological actions, 3231, 3235, 3236
- precautions/adverse reactions, 3232, 3240, 3240*t*
- sleep apnea, 3232
- therapeutic indications, 3231–3232, 3239
- use in cancer patients, 2262

- Modality shifts, 695
- Mode (mathematical), 4595
- Model Penal Code, 4595
- Modeling, description of, 4595
- Moderate mental retardation. *See* Mental retardation
- Modern Primitive Movement, 2582
- Modes, definition, 2763
- Modified Checklist for Autism on Toddlers (MCHAT), 644
- Mofegiline, 3113*f*
- Molecular genetic studies
  - ADHD, 3589
  - aggression, 2482
  - anxiety disorders, 1738
  - specific learning disability, 3524–3525
- Molecular imaging
  - for dementia, 1197
  - schizophrenia, **1439–1444, 1488–1501, 1489***t*
    - clinical practice, 1500
    - drug development, 1497–1499
    - future directions, 1500
    - genetics, 1499–1500
    - pathophysiology, 1492–1497
    - theories, 1488–1490, 1489*f*

    - types, 1490–1492, 1491*t*
- Molecular psychiatry, 4595
- Molière on hypochondria, 1838
- Molindone (Moban), 3063*f*, 3064
- Monetary incomes, older people, 3953*t*
- Money, happiness and, 2848
- Money, John, 1958, 2023
  - on transsexualism, 2025
- Mongolism. *See* Down syndrome (DS)
- Monitoring community mental health, world-wide, 4526
- Monitoring the Future Study (MTF)
  - benzodiazepine use prevalence, 1379
  - on substance abuse treatment, 1263
- Moniz, Egas, 4540
- Monkeys
  - corticostriatal projection
    - organization of, 29*f*
  - striatonigrostriatal projection
    - organization of, 29*f*
  - substantia nigra in, 28
- Monoamine neurotransmitters, 2593–2594
- Monoamine oxidase (MAO)
  - description of, 4595
  - isoforms of, 66
  - serotonin degradation by, 69
  - transmitter degradation by, 63
- Monoamine oxidase A (MAOA), 2482–2483
- Monoamine oxidase A *(MAOA)* gene, 2482–2483, 2482*t*
- Monoamine oxidase inhibitors (MAOIs), **3111–3125**
  - ADHD, 3123
  - adverse effects, 3115–3116
  - agoraphobia, 3122–3123
  - for Alzheimer disease, 4146
  - antidepressant discontinuation syndrome, 2943
  - anxiety, 3122–3123
  - anxiety disorders, 1780, 1782*t*
  - application site reaction, 3118

- atypical depression, 1710
- BDD, 1792
- bipolar depression, 1692*t*, 1693, 3121
- BPD, 3121
- for bulimia nervosa, 2081
- cardiovascular effects of, 2197*f*
- chemistry, 3112–3113, 3113*f*
- classification, 3114
- contraindications, 2520, 3118
- depression, 1678, 1681, 3119–3121
  - atypical, 3118–3119
  - treatment-resistant, 3119–3120
- depression treatment, 4125
- description of, 4595
- dissociative identity disorder, 1927*t*, 1928
- dosage/administration, 3114–3116, 3114*t*
- drug/food interactions, 2988, 3013, 3028*t*, 3111, 3117, 3117*t*, 3130, 3150, 3162–3164, 3263
- drug studies, 3118–3123
- eating disorders, 3121–3122
- ECT interactions with, 3295
- fear conditioning, PTSD and, 1818
- future directions, 3123
- GAD, 1783
- for geriatric depression, 4073, 4125
- hepatotoxicity, 3118
- history, 2917, 3111–3112
- hypertensive reaction/crisis, 3028*t*, 3117, 3117*t*
- hypotension, orthostatic, 3118
- hysteroid dysphoria, 1652
- insomnia, 3118
- insomnia, in aging adults, 4054
- investigational drugs, 3123
- for MDD, 3120
  - in brain tumor patients, 499
- mechanism of action, 3113–3114, 3113*f*
- monitoring parameters, 3115–3116
- neurobiology and, 1709

- neurodegenerative disorders, 3123
- OCD, 1789, 3122
- organs/systems effects of, 3117
- overdose, 3118
- panic disorder, 3122–3123
- pharmacological actions, 3114–3116
- and physical health monitoring, 1025
- precautions and warnings, 3115–3116, 3117–3118, 3117*t*
- pregnancy, lactation and, 3115–3116
- psychosis, 3118
- PTSD, 3122
- SAD, 3121
- serotonin syndrome and, 3118, 3118*t*, 3162, 3163–3164
- side effects
  - in cancer, 2261
- social anxiety disorder, 3122
- SSRIs *versus,* 1779
- therapeutic indications, 3118–3123
- weight gain, 3118
- withdrawal, 3118

- Monoamines
  - depression and, 1713
    - dopaminergic systems, 1714
    - noradrenergic systems, 1713–1714
    - research summary, 1714
    - serotoninergic systems, 1714
  - description of, 4595
- Monocytes, differentiation of, 179
- Monomania, 4595
- Mononeuropathy multiplex, 632
- Monosymptomatic hypochondriacal psychosis, 4595
- Montgomery–Asberg Depression Rating Scale (MADRS), 1684, 3976
- Montreal Cognitive Assessment (MOCA)
  - of aging adult, 3975, 4004
  - cognitive function at patient interview, 955
  - in delirium, 4090
  - dementias in long-term care, 4226

- and electroconvulsive therapy, 4155
- neuropsychological testing, 988
- in psychosomatic medicine, 2182

- Mood and Feelings Questionnaire (MFQ), 3400
- Mood-congruent delusions, 4595
- Mood-congruent hallucinations, 4595
- Mood-congruent psychotic features, 4595
- Mood disorder(s), **1599–1719,** 4554.
  See also specific mood disorders
  - ADHD comorbidity, 3596
  - in aging adults, 3978–3979, 3974–3975, 4075–4076
    - bereavement and, 4067, 4075
    - bipolar disorder, 4076–4077
    - in bipolar disorder, 4067
    - brain structure and function studies, 4070–4071
    - course and outcome, 4071
    - dementia, 4075
    - in dementia patients, 4066
    - depression, **4066–4075**
    - depression-executive dysfunction, 4068–4069
    - and disability, 4067
    - epidemiology, 4066
    - genetic studies, 4069–4070
    - inflammation, 4069
    - late-onset hypothesis, 4068
    - in medically ill patients, 4066
    - pharmacotherapy, 4072–4074
    - psychotherapy, 4072
    - psychotic, 4067
    - somatic treatment, 4074
    - suicide, 4067–4068
    - treatment-resistant, 4075
    - vagal nerve stimulation, 4074–4075
    - vascular depression, 4069
  - anabolic steroid-induced, 1399–1402
  - ancient Greek and Roman descriptions, 1601–1603
  - bi-polar disorders
    - cognitive studies, 698–699

- biogenic amine imbalance, 1607–1609
- biological stressors and, 1611
- brain imaging, 1717
- CAIM medicine, 3982–3885
  - acupuncture, 3984–3985
  - mind-body practices, 3983–3984
  - natural products, 3984
  - physical activity, 3983
- in cancer patients
  - causes of, 2260*t*
- candidate gene association studies, 1625, 1626*t*
- chemotherapy agents and, 2273*t*
- child/adolescent, **3674–3694, 3688–3708**
  - depressive disorders, **3674–3680, 3688–3694**
  - early-onset bipolar disorder, **3686–3694, 3700–3708**
  - suicidal behavior, **3680–3686, 3694–3700**
- childhood aggression and, 2487
- classification of, 1172
- clinical features, 1630–1660
- concepts of, 1599–1600, 1600*f*
- COPD and, 2226
- corticosteroids and, 2229
- definitions, 1600–1601
- depression-centric perspective on, 1631
- description of, 4595
- developmental predisposition, 1610
- diabetes and, 2231–2232
- diagnostic classification, 1641–1643
  - dichotomy or continuum in, 1642–1643, 1642*f*, 1642*t*
- differential diagnoses, 1586, 1589, 1653–1659, 3610
  - alcohol use disorders in, 1655–1656
  - anxiety disorders and, 1654, 1654*t*
  - bereavement, normal, and, 1653–1654
  - personality disorders and, 1655, 1655*t*
  - physical disease in, 1656–1658, 1656*t*
  - schizoaffective disorder in, 1658–1659
  - schizophrenia in, 1658, 1658*t*

- substance use disorders in, 1655–1656
- dissociative identity disorder *versus,* 1882
- dopaminergic systems and, 1714
- due to another medical condition, **1248–1251**
  - comparative nosology, 1248–1249
  - course and prognosis, 1251
  - defined, 1248
  - diagnosis and clinical features, 1235*t*, 1250
  - differential diagnoses, 1250–1251
  - epidemiology, 1249
  - etiology, 1222*t*, 1249–1250
  - history, 1248–1249
  - laboratory evaluation, 1250
  - treatment, 1251
- EEGs in, 250
- emotion and, 464–465
- endocannabinoids and, 114
- epidemiology, 1614–1619
- in epilepsy, 514
- etiology of emotions and, 1711
- and everyday functioning, 4027
- excoriation disorder and, 1796
- fMRI of, 277–278
- gambling and, 1805
- genes
  - functional validation of, 1627–1628
  - identifying susceptibility, 1629–1630
  - strategies for identifying, 1624–1625
  - variable expressivity, 1621
- genetic counseling, 1629
- genetic linkage studies, 1625
- genetics
  - epidemiology, 1619–1620
  - features of, 1620–1621
  - introduction, 1619
  - molecular pathways, 1627
  - overlap between other disorders and, 1628

- pharmacogenetics, 1628–1629
- selective advantage, 1623–1624
- spectrum of, 1622–1623, 1623*f*
- transmission of, 1621–1622
- genome-wide association studies, 1626
  - rare variation studies, 1626–1627
- group psychotherapy, 2745–2746
- growth hormone and, 1716
- habenula in, 33
- hallucinogen-assisted psychotherapy and, 2889
- head injuries and, 454
- heredity of, 1610, 1611*f*
- historical introduction, 1599–1613
- historical trends, 1618–1619
- in HIV patients, 571, 4269
- hormonal regulation alterations and, 1715–1716
- HPA axis and, 167
- immunological disturbance, 1717–1718
- impact of, 1618
- with intellectual disability, 3512
- intrapsychic and interpersonal aspects, 1660–1667
  - psychodynamic aspects, 1660–1661
- life events and, 1611
  - negative, suicidal behavior and, 1671
- migraine and, 622
- in modern era, 1603–1605
- morbidity, 1600
- mortality, 1600
- MRI applications to, 273–274
- MRS in, 281–282
- neurobiology, 1709–1719
- neurophysiological approaches, 1609–1610
- nitric oxide and, 107
- noradrenergic systems and, 1713–1714
- pain and
  - course/prognosis of, 2393
- pathophysiology of, 1611*f*

- perinatal
  - treatment of, 2518–2519
- pharmacogenetics, 1628–1629
- physician, 2599
- postpartum
  - screening for, 2517
- premature births and, 643
- prevalence, 1599
- prolactin and, 1716
- psychodynamic aspects, 1660–1661
- psychopathology, 1634–1641
- psychopharmacology, 3790–3792
- psychotherapy, 1700–1708
- recent diagnostic trends, 1631
- risk factors, 1710–1711, 1710*f*
- serotoninergic systems and, 1714
- sex and, 1611–1612, 1611*f*
- sleep deprivation and, 353
- sleep disorders and, 3938–3939
- sleep neurophysiology alterations, 1716–1717
- with specific learning disorder, 3529
- spectrum of, 1631
- stimulant-induced, 1286
- stress and, 2321, 2323
- substance/medication-induced, **1248–1251**
- suicidal behavior, 1667–1676
  - clinical utility of risk factors, 1671–1672, 1672*t*
  - comorbid psychiatric, personality, and medical disorders, 1669–1670
  - demographic factors, 1671
  - family history and, 1670–1671
  - negative life events, 1671
  - prevention strategies, 1673–1676, 1676*t*
  - previous suicide attempt, 1670
  - prior course and, 1670
  - protective factors, 1672
  - relationship between, 1667–1668
  - specific personality features, 1670

- TBI and, 535–536
- temperament and, 3486, 3486*f*
- terminology, 1630–1631
- theoretical synthesis, 1610–1613
  - pathophysiological understanding of, 1610–1612
  - therapeutic perspectives on, 1612–1613
- thyroid axis activity and, 1716
- tic disorders and, 3642, 3656
- in transplant populations, 2365–2366
- treatment, 1676–1700
  - anticonvulsants, 2968, 2972–2973
  - bipolar disorders, 1681–1699
  - depression, 1677–1681
  - ECT, 3285–3286
  - future directions, 1699
  - nefazodone, 3126, 3127
- treatment delay, *4527t*
- treatment gap, 4522
- TTM and, 1795
- unspecified, 1652–1653
- Vietnam War with PTSD and, 1816
- world-wide, 4517, 4518, 4520

- Mood disturbances
  - in depressive syndrome, 1634
  - in mania, 1640–1641
- Mood-incongruent delusions, 4595
- Mood-incongruent hallucinations, 4596
- Mood-incongruent psychotic features, 4596
- Mood stabilizers
  - aggression
    - intermittent explosive disorder, 2501
    - schizophrenia, 2501
  - anticonvulsants as, 1612
  - antidepressants combined with, 3327
  - antipsychotics combined with, 3326–3328
  - bipolar II disorder
    - maintenance therapy, 1698

- calcium channel inhibitors, 3019
- classification, 2918
- combination, 3325–3326
- description of, 4596
- dissociative identity disorder, 1927*t*
- endocrine effects of, 3796
- in functional GI disorders, 2206–2207
- gambling disorder, 1808
- in HIV setting, 572
- and physical health monitoring, 1024
- prophylactic, valproate as, 3270, 3272
- side effects of, 2189
- suicide risk, 2616
- in transplant patients
  - specific considerations, 2372*t*
- for trichotillomania, 2353
- weight gain and, 2215, 2215*t*

- Mood–labile temperament, 1651
- Mood(s). *See also* Affect(s); Mood disorder(s)
  - with dementia, 1196
  - description of, 4595
  - dietary modification and, 2547
  - disturbances of, 1136–1138
    - depressive state, 1136–1137
    - manic states, 1137–1138
  - elevated
    - manic, 1640
  - endoreactive, 1633, 1650
  - ethological considerations, 1632
  - instability
    - insulin resistance and, 2247
    - medications for stabilization, 2171, 2172*t*
  - negative alterations in, PTSD and, 1814
  - pathological change in, 1633
  - at patient interview, 952, 966
  - swings of, 4596
- Moody, Raymond, 2589

- Mora, George, 4554
- Moral insanity, 3612
- *Moral Judgement of the Child* (Piaget), 710
- Moral managements, 4290
- Moral treatment, 4596
- Morality
  - the brain and, 2415–2416
  - theories of, 2415
- Morality and moral reasoning
  - conduct disorder, 3612
  - development
    - adolescent, 3369–3370
    - childhood, 3354–3357*t*, 3361
- Moratorium during adolescence, 3365
- Morbid jealousy, 4596
- Morbid perplexity, 4596
- Moreau de Tours, Jacques-Joseph, 4535
- Morel, Benedict-Augustin, 4536
- Moreno, Jacob L.
  - group psychotherapy and, 2737
  - psychodrama and, 2855–2856, 2857, 2859
- Moreno, Zerka Toeman
  - psychodrama and, 2856, 2857
- Morgan, Charles A.
  - depersonalization neurochemistry, 1891
  - military stress studies, 1897, 1947
- Morgan, Thomas, 149
- Morgellons disease, 2346
- Moron, definition of, 4596
- Morphine (Duramorph, MS Contin)
  - administration of, 2293*t*
  - delirium and, 1182
  - description of, 4596
  - duration of action, 2395*t*
  - EEG effects, 253
  - history of use, 1353
  - insomnia and, 2295–2296

- laboratory examination for, 1363
- opioid-dependence treatment, 1372–1373
- oral bioavailability, 2395
- pain management using, 2293
- pharmacology, 1357–1358, 1357*f*
- PTSD prevention, 1825
- routes of administration
  - for pain management, 2376–2377

- Morphine Benzedrine Group Subscale of the Addiction Research Center Inventory (ARCI-MBG), 1377
- Morphine sulfate (Roxanol), 2255
- Morphometric studies, schizophrenia, 1458–1459
- Morris water maze
  - rat, 324
- Morselli, Enrico, 1790
- Mortality
  - and antipsychotic drugs, 4136
  - gender differences, 4246
  - inhalant-related, 1330–1331, 1333
  - mood disorders and, 1618
  - opioid-related, 1364–1365
  - schizophrenia-related, 1434–1435, 1538–1542, 1539*t*
  - with severe mental disorders, 4529
  - tobacco-related, 1347
- Mortality rates
  - with antipsychotic medications, 4342
  - evolution of, 771
- Morvan syndrome, 676, 2270
- *Moses and Monotheism* (Freud), 4596
- Mother surrogates, 4596
- Mother–child interactions
  - feeding disorder of caregiver-infant reciprocity, 3629–3630, 3643–3644
  - infantile anorexia and, 3623–3625, 3637–3639
  - intellectual disabilities and, 3509–3510
  - rumination disorder and, 3631, 3645
  - skin disease and, 2334
- Mothering
  - factitious disorder imposed on another and, 1851

- factitious disorder imposed on self and, 1851
- Mothering to death, 1859
- Motion sickness, 2982
- Motivation, 2133
  - to change
    - addiction therapy and, 580
  - description of, 4596
  - hunger as, 355–356, 358*f*
  - Maslow personality theory, 932
  - mechanisms of
    - pain and, 327–328
  - in operant learning, 724–725
  - for suicide behavior, 3682, 3696
- Motivation-reward processes, impaired, 2045
- Motivational enhancement therapy (MET), 3057
- Motivational interviewing
  - behavior therapy and, 2702
  - schizophrenia, 1553
  - states of change and, 2219*f*
  - substance abuse, child/adolescent treatment, 3768
  - SUD treatments, 384
  - weight loss using, 2218–2220
- Motoprolol, 624
- Motor activity
  - autism spectrum disorder and, 3582
  - development, in childhood, 3361–3362
  - disturbances, 1146–1147
    - catatonic activity as, 1147
    - motor retardation as, 1147
    - mutism as, 1147
    - overactivity as, 1146–1147
  - hemispheric dominance, 982
  - neuropsychological assessment of, 3435–3436
  - at patient interview, 952
- Motor aphasia, description of, 4596
- Motor cortex
  - gene expression in development of, 49

- surface landmarks, 7*f*
- Motor deficiency syndrome. *See* Developmental coordination disorder
- Motor disorders, definition of, 4596
- Motor disturbances and movement disorders
  - cannabis use for, 1311
  - clinical manifestations of, 1147–1149
    - blepharospasm as, 1147
    - bruxism as, 1149
    - drug-induced, 1148
    - gait disturbances as, 1149
    - myoclonus as, 1149
    - in schizophrenia, 1147
    - seizure-like behaviors as, 1149
    - tics as, 1147–1148
    - tremors as, 1147
  - differential diagnoses, 1595
  - of disorganization syndrome, 1419–1420
  - schizophrenia, 1419–1420
  - stereotypy and stereotypic movement disorders, **3667–3673, 3681–3687**
  - tic disorder, **3635–3649, 3649–3663**
- Motor function, burn injuries and, 2374
- Motor imagery intervention, 3543
- Motor overflow. *See* Synkinesia
- Motor sequencing tests, 459
- Motor skills disorders. *See* Developmental coordination disorder
- Motor systems
  - dysfunction, screening for, 541*t*
  - FGA effects on, 3069
  - thalamocortical, 24–25
- Motor tics, 3593
- Mourning, 4596. *See also* Bereavement
- *Mourning and Melancholia* (Freud), 868, 872
- Mouth
  - physical examination of, 456
  - ulcers, 456
- Movement abnormalities, 457–459
- Movement Assessment Battery for Children (MABC), 3539

- Movement disorders
  - bedside assessment of, 540–542
  - classification of, 540, 541*t*
  - medication-induced, **2936–2945**
    - acute akathisia, 2938, 2941, 2955
    - acute dystonia, 2940–2941
    - antidepressant discontinuation syndrome, 2943
    - CNS control of movement, 2936, 2938
    - diagnostic criteria for, 2936, 2937*t*
    - ECT treatment for, 3287
    - future directions, 2944
    - history of, 2936
    - impact of, 2944
    - medications triggering, 2936, 2936*t*
    - neuroleptic-induced parkinsonism, 2938, 2959
    - neuroleptic malignant syndrome, 2938–2940, 2939*t*, 2940*t*
    - other, 2943
    - postural tremor, 2942
    - tardive akathisia, 2941, 2942–2943
    - tardive dyskinesia, 2941–2942
    - tardive dystonia, 2942–2943
  - neuropsychiatric aspects of, **540–561**
  - patient assessment in, 472
  - sleep-related, 2104
- MPD. *See* Multiple personality disorder (MPD)
- MPTP. *See* Methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)
- MRI. *See* Magnetic resonance imaging (MRI)
- MRNA studies, 147
- MSA. *See* Multiple system atrophy (MSA)
- MSE. *See* Mental status examination
- MTHFR. *See* Methylenetetrahydrofolate reductase (MTHFR)
- Muller (glia) cells, 138
- Multi-infarct dementia. *See* Vascular dementia
- Multi-scale Dissociation Inventory (MDI), 1877*t*
- Multi-tasking, everyday functioning, 4026
- Multiaxial system, deletion of, 1157
- Multicondition collaborative care interventions, 2237

- Multiculturalism, stress of, 792
- Multidimensional Anxiety Scale for Children (MASC), 3401
- Multidimensional Fatigue Inventory, 2284
- Multidimensional Inventory of Dissociation (MID), 1877*t*, 1878, 1882
- Multidimensional Treatment Foster Care (MTFC), 3620
- Multidimensionally impaired disorder (MDI)
  - defined, 3725
  - differential diagnoses, 3731
- Multidisciplinary teams
  - to child maltreatment, 3840
  - for preschool psychiatric evaluations, 3410–3412
- Multimodal Treatment Study of Attention-Deficit/Hyperactivity Disorder (MTA trial), 3588, 3593, 3596
- Multimodal Treatment Study of Children with ADHD (MTA), 3747, 3789
- Multimorbidity, HIV infection and, 4271
- Multiple agency, ethical issues with, 3871
- Multiple baseline design experiments, 2704
- Multiple-choice questions, 4419–4420
- *Multiple Chronic Conditions Measurement Framework,* 4052
- Multiple complex developmental disorder (MCDD)
  - defined, 3725
  - differential diagnoses, 3731
- Multiple function, definition of, 2660
- Multiple identity enactments, 1897
- Multiple myeloma (MM), 2269–2270
- Multiple organ dysfunction, 631
- Multiple personality disorder (MPD). *See also* Dissociative identity disorder
  - diagnostic criteria changes, 1878–1879, 1880*t*
  - DSM and ICD diagnostic criteria, 1874
  - history, 1866, 1868
  - ICD-10 on, 1875*t*
  - media and public fascination with, 1873–1874
  - popular interest in, 1870
- Multiple reality disorder, 1906
- Multiple regression (statistical), 4596
- Multiple sclerosis (MS), **561–563**
  - cognitive impairment in, 563
  - diagnosis of, 561

- differential diagnoses, 563
- emotional symptoms in, 562
- epidemiology of, 561–562
- fatigue in, 562
  - modafinil for, 2265
- forms of, 562
- laboratory evaluation, 1015
- mania without prior bipolarity and, 1649
- MRI, 270*f*
- neuropsychiatric aspects of, 562
- psychiatric disorders in, 562
- sexual function and, 1970
- SRIF and, 175

- Multiple sleep latency test (MSLT)
  - description of, 2106
  - function of, 343
- Multiple system atrophy (MSA), 551
  - parasomnia and, 2102
- Multiple therapy. *See* Co-therapy
- Multisystemic therapy (MST), 3815
  - for aggressive behavior
    - in ADHD, 2500
    - in conduct disorder, 2500
  - child/adolescent, 3751, 3779, 3802, 3808, 3867
  - conduct disorder and, 3620
  - research in, 2759
- Multivariate analysis, 4596
- Multivariate random variables, 831
- Munchausen, Karl Friedrich Hieronymus von, 1847, 1848*f*
- Munchausen syndrome
  - comorbidity, 1864
  - comparative nosology, 1848–1849
  - course and prognosis, 1861
  - definition, 1846–1847
  - description of, 2411, 4596
  - diagnosis and clinical features, 1857
  - differential diagnoses, 1858–1859

- dissociative identity disorder and, 1917
- epidemiology, 1849–1850
- factitious disorder, 1842
- factitious disorder imposed on another and, 1850

- Munchausen syndrome by proxy, 1847–1848, 1857, 2411
  - dissociative identity disorder and, 1917
- "Munchies," 110
  - THC and, 115
- Murder rates, 2417
- Murphy, Dominique, 1154
- Murphy, Henry, 1433
- Murray, Henry
  - *Thematic Apperception Test,* 1005, 4537
- Musafar (Fakir), 2582
- Muscarine, 4596
- Muscimol, 81
- Muscle atonia in REM sleep, 347
- Muscle dysmorphia, 1791
- Muscle relaxants, side effects of, 571
- Muscle relaxation
  - progressive
    - behavior therapy, 2689
    - deep, 2689
    - somatic therapy, 2557
  - PTSD, 1824
- Muscle spasticity, cannabis use for, 1311
- Musculoskeletal disorders, psychiatric symptoms and, 4048–4049
- Musculoskeletal system
  - ECT effects on, 3296–3297
  - opioid receptor antagonist effects on, 3143
- Music
  - as CAM therapy, 2558
  - scientifically validated therapy with, 2546
  - therapy for dying patients, 2302
- Muslim migrants, global perspective, 797
- Mustard gas, war or terrorism and, 2572*t*
- Mutagenesis

- opioid receptor antagonist effects on, 3143
- vilazodone and, 3346

- Mutations
  - definition, 155
  - in natural selection, 772
- Mutism
  - aphasia and, 457
  - aphemia and, 457
  - description of, 4596
  - selective
    - definition, diagnosis, and clinical features, 1728
    - differential diagnoses, 1728
    - DSM-5 on, 1723
    - epidemiology, 1728
    - history and comparative nosology, 1728
    - management in children and adolescents, 1784
    - neuroimaging, 1765
- Mutual help, SUD treatments, 384*t*
- Myasthenia gravis (MG)
  - muscle tone in, 73
  - presentation of, 633
  - treatment, 633
- *Mycobacterium avium* complex (MAC), 575
- *Mycobacterium leprae,* 632
- Mycophenolate mofetil, 2371
- Mycotoxins, war or terrorism and, 2572*t*
- Mydriasis, 4596
- Myelencephalon (medulla)
  - embryology of, 41
  - embryonic origins of, 8
- Myelin sheaths, 5
- Myelination, progression of, 5
- Myeloarchitecture, brain, 5
- Myelopathies, chemotherapy and, 2273*t*
- Myers, Michael
  - on minority physicians, 2598–2599
  - on physician mental health, 2600, 2602

- Myers-Briggs Type Indicator, 935
- Myerson sign, 460
- Myocardial infarction
  - depression and, 2187
  - long-term survival after, 2188*f*
- Myocarditis
  - antipsychotic drug effects, 4138
  - clozapine-associated, 1023
- Myoclonic epilepsy with ragged red fibers (MERRF)
  - course of, 670
  - genetics of, 670
- Myoclonus
  - benign sleep myoclonus of infancy, 2104
  - description of, 458
  - differential diagnoses, 469*t*
  - excessive fragmentary, 2105
  - in HD, 555
  - in psychiatric disorders, 1149
- Myofascial syndrome, 2390*t*
- Myopathies
  - endocrine, 635
  - inflammatory, 634
  - mitochondrial, 634–635
- Myopia, genetic selection, 774
- Myotonic dystrophy, 633–634
- Mysophobia, 4596
- Mystical Experience Questionnaire, 2581
- Mysticism
  - psychodrama and, 2856
  - spirituality and, 2578–2579
- *Myth of the Birth of the Hero* (Rank), 4537
- Myxedema, 680

N

- *N*-Acetylaspartate (NAA)
  - glial import of, 278
  - in mood disorders, 281–282
- N-Acetylcysteine (NAC)
  - biological treatment with, 2550
  - bipolar depression, 1692*t*
  - bipolar II disorder, 1698
  - excoriation disorder, 1797
  - mitochondrial modulation and, 2546
  - PTSD and SUD comorbidity, 1823
  - TTM, 1795
- *N*-Methyl-D-aspartate (NMDA)
  - dissociative symptoms and, 1891
  - glutamate binding to, 1747
  - receptors
    - depression, 1681
    - mood disorders and, 1715
    - trauma response and, 2569
  - zinc and, 2549
- *N*-Methyl-D-aspartic acid, 77*f*
- *N*-Methyl-D-aspartic acid (NMDA) receptors
  - function of, 77
  - ligands for, 77*f*
  - nitric oxide and, 106–107
  - schematic representation of, 78*f*
- NAB. *See* Neuropsychological Assessment Battery
- Nabilone (Cesamet)
  - for nausea, 2271
  - PTSD, 1823
- Nadolol (Corgard)
  - absorption, 2951
  - aggression in mentally retarded children, 2500
  - chemistry, 2951, 2951*f*
  - distribution, 2952
  - dosing/administration, 2956

- drug interactions, 2955
- drug studies, 2952
- in hyperthyroidism
  - symptom relief using, 2243
- organ/system effects of, 2952–2953, 2953*t*
- pharmacodynamics, 2952
- pharmacokinetics, 2951–2952, 2952*t*
- pharmacological actions, 2951–2952, 2952*t*
- precautions/adverse reactions, 2955
- teratogenic effects of, 2953
- therapeutic indications, 2953–2955, 2953*t*

- Nail biting and nail picking, 3668, 3669, 3672, 3682, 3683, 3686
- Naïve realism, 2673
- Naliboff, Bruce D., 1832
- Nalmefene (Revex)
  - alcohol-related disorders, 3139–3140
  - chemistry, 3139
  - dosage/administration, 3145–3146
  - drug interactions, 3145
  - drug studies, 3139–3140
  - emotional numbing in PTSD and, 1892
  - history, 3139
  - investigational drugs, 3146
  - laboratory interferences, 3145
  - organ/system effects of, 3143–3144
  - PG, 3143
  - pharmacological actions, 3139
  - precautions/adverse reactions, 3144–3145
  - therapeutic indications, 3140–3143
- Naloxone (Narcan)
  - buprenorphine combined with, 3136
  - delirium treatment, 1190
  - description of, 4596
  - dosage/administration, 3145–3146
  - eating disorders/obesity, 3142
  - history, 3139
  - methadone overdose, 3134

- nicotine dependence, 3142
- opioid overdose, 3139, 3141–3142
- opioid overdose treatment, 1367–1368
- opioid reversal by, 169
- opioid withdrawal, 3132
- precautions/adverse reaction, 3144
- stereotypic movement disorders, 3578
- therapeutic indications, 3144

- Naltrexone (Narcan/ReVia)
  - acamprosate and, 3058
  - alcohol-related disorders, 3139–3140
  - for alcohol use disorders, 1278
  - for amphetamine-type stimulant use disorders, 1290
  - autism spectrum disorder and, 3579
  - in autistic children, 169
  - BD and alcohol comorbidity, 1699
  - bupropion and, 3006, 3139, 3142, 3144
  - chemistry, 3139, 3139*f*
  - cocaine-related disorder, 3142
  - delirium treatment, 1190
  - depersonalization/derealization disorder, 1945
  - dissociation, 1892
  - dissociative identity disorder, 1927*t*, 1928
  - dosage/administration, 3145–3146
  - drug interactions, 3145
  - drug studies, 3139–3140
  - eating disorders, 3006, 3142–3143
  - for excoriation disorders, 2352
  - gambling disorder, 1808
  - hepatotoxicity, 582, 2371
  - history, 3138–3139
  - investigational drugs, 3146
  - laboratory interferences, 3145
  - mechanism of action, 581–582
  - narcotic addiction, 3140–3141
  - nicotine dependence, 3142
  - obesity, 3006, 3142–3143

- opiate detoxification, 2948, 3141–3142
- opioid-dependence treatment, 1372
- opioid withdrawal, 3132–3133
- opioid withdrawal treatment, 1369–1370
- organ/system effects of, 3143–3144
- PG, 3143
- pharmacological actions, 3139
- precautions/adverse reactions, 3144–3145
- PTSD and SUD comorbidity, 1823
- self-injurious behavior, 3142
- stereotypic movement disorders, 3578
- substance use disorders, child/adolescent, 3793
- therapeutic indications, 3140–3143, 3144
- for trichotillomania, 2353

- Naltrexone/bupropion (Contrave)
  - mechanism of action, 2221
  - for weight loss, 2220
- Namenda. *See* Memantine
- NAMI. *See* National Alliance for Mental Illness
- Nandrolone decanoate (Deca-Durabolin), 2493
- Nandrolone phenpropionate (Durabolin), 2493
- Nanotechnology for seniors, 4279–4280
- Napping, insomnia and, 2089
- NAPSA. *See* National Adult Protective Services Association
- Narcan. *See* Naltrexone
- Narcissism, 4106
  - description of, 4596
  - and dual instinct theory, 865
  - and love object, 865
- Narcissistic disorders
  - STDP of, 2805–2810
    - case study, 2809
    - clinical issues, 2808–2810, 2808*t*
      - complications, 2808
      - dynamic hologram, 2809
      - indications, 2808
      - limitations, 2808

        - outcomes, 2810
        - treatment goals, 2809
        - vignette, 2809–2810
    - summary, 2810*t*
    - techniques, 2806–2808, 2807*f*
        - empathy, 2807
        - imaginary reconstruction, 2807, 2807*f*
        - reactivation of self, 2807–2808
        - transference focus, 2807
        - working through of rage, 2807
    - theoretical issues, 2805–2806
- Narcissistic personality disorder
    - AEDP and, 2804
    - clinical criteria, 2154
    - comorbidity, 2154
    - complications, 2154
    - course, 2154
    - covert type
        - diagnostic criteria, 2808, 2808*t*
    - description of, 4596
    - diagnostic criteria, 2050, 2808*t*
    - differential diagnoses, 2154
    - epidemiology, 2154
    - factitious disorder imposed on self and, 1851
    - impairment, 2154
    - mixed, 2808
    - predisposing factors, 2154
    - sex ratios, 2154
    - spectrum of, 2155
- Narcissistic psychology, 2608
- Narcissistic psychopathology, 2806
- Narcissistic transference, 2651–2652
- Narcoanalysis. *See* Narcotherapy
- Narcolepsy
    - animal models of, 324
    - characteristics of, 2091–2092
    - description of, 4596

- diagnostic criteria, 2092–2093
- immune basis of, 677
- pediatric, 3936
- pitolisant for, 75
- presentation of, 677
- sleep stage histogram, 2092*f*
- stimulant drugs for, 2293
- treatment
  - modafinil, 3231–3232
- Narcotherapy, 2987, 4596
- Narcotic blockade, 4596
- Narcotics. *See also specific drugs*
  - addiction, treatment, 3140–3141
  - description of, 4596
- Nardil. *See* Phenelzine
- Narrative, definition of, 693, 2831
- Narrative psychiatry, **2830–2841**
  - character, 2832–2833
  - first-person narratives, 2833–2835
  - future directions, 2840–2841
  - informed consent and, 2836–2837
  - introduction, 2830–2831
  - literary example, 2835–2836
  - metaphor, 2831–2832
  - metatheoretical rationale for, 2837
  - plot, 2832
  - practice, 2835–2840
- Narrative therapies, 2751*t*, 2755
- Narrative truth, historical truth *versus,* 2657
- Narratives as stories (McAdams), 933
- NARSAD activities, 4544
- Natal gender, 1984, 2023, 2024
- Nathanson, Donald, 1907
- National Adult Protective Services Association (NAPSA), 4252
- National Alliance for Mental Illness (NAMI)
  - founding of, 4543
  - treatment recommendations, 956

- National Alliance for the Mentally Ill (NAMI), 1554, 1559, 1560
  - description of, 4597
- National Ambulatory Medical Care, 803
- National Association for the Dually Diagnosed (NADD), 3493
- National Board of Medical Examiners (NBME)
  - description of, 4597
  - USMLEs for, 4418–4419
- National Center for Complementary and Alternative Medicine (NCCAM), 2542, 2544
  - description of, 4597
- National Center for PTSD, 2571
- National Character Studies, 2460
- National Child Traumatic Stress Network, 2571, 4296
- National Committee for Mental Hygiene, 4291
- National Committee for Quality Assurance (NCQA)
  - certification standards, 4307
  - on quality improvements with managed care, 4314
- National Comorbidity Study (NCS), 783, 800, 1822, 1904
  - care access crisis reported in, 4306
  - on drug dependence, 1306
  - epidemiological databases of, 4304
  - somatization disorder, 1836
  - suicide, 2614
- National Comorbidity Survey Replication (NCS-R), 1731, 1733, 1786, 1816, 1904, 4304
  - prevalence of mental disorders in older adults, 3960
- National Comorbidity Survey Replication Adolescent Supplement (NCS-A), 826–827, 827*t*, 1614, 1733
- National Comorbidity Survey Replication Study (NCS), 2568
- National Council on Problem Gambling (NCPG), 1810
- National Crime Victimization Survey (NCVS), U. S., 2473
  - Bureau of Justice Statistics', 2538
  - offender characteristics, 2538
  - rape statistics, 2539
  - U.S., 2473
- National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), 783, 825–826, 826*t*, 4104
  - on cannabis use, 1306

- mental disorders in U.S., 800
- mood disorders, 1614
- on opioid-related disorders, 1365, 1366*t*
- prevalence of mental disorders in older adults, 3961
- somatic symptom disorder, 1836

- National Health and Nutrition Examination Survey (NHANES), 1733
  - mental disorders in U.S., 800
  - on schizophrenia, 1540, 1541, 1542, 1543, 1544, 1544*t*, 1545*t*
- National Health Care Anti-Fraud Association (NHCAA), 2414
- National health insurance, 4311, *4311t*
- National Health Interview Survey (NHIS), 4258
- *National Healthcare Disparities Report* (2003), 4377
- National Hospital Discharge Survey, 1849
- National Household Survey on Drug Use and Health (NSDUH), 1282
- National Institute of Alcohol Abuse and Alcoholism (NIAAA), 4597
- National Institute of Drug Abuse (NIDA), 4597
- National Institute of Mental Health (NIMH)
  - classification alternative by, 1161–1162, 1161*t*
  - creation of, 4292
  - Decade of the Brain sponsored by, 4306
  - description of, 4597
  - establishment of, 4335
  - grants from, 4321
  - post-World War II creation of, 2842
  - psychiatric training and, 4402
  - research domain criteria, 1844
  - Treatment of Depression Collaborative Research Program (TDCRP), 1702, 2695
- National Institute on Aging and the Alzheimer Association (NIA-AA), 1202, 4079
- National Institute on Alcohol Abuse and Alcoholism (NIAAA)
  - alcohol use inquiry, 950
  - substance abuse disorders, 821
- National Institute on Drug Abuse (NIDA) study, 1282
- National Latino and Asian American Study (NLAAS)
  - prevalence of mental disorders in older adults, 3960
  - psychiatric service utilization, 801–802
  - shortcomings revealed by, 4382–4383
- National Mental Health Act, 4335

- National Mental Health Association (NMHA), 4597
- National Opinion Research Center DSM Screen for Gambling Problems (NODS), 1805
- National Organ Transplant Act (NOTA), 2357
- National Plan for the Chronically Mentally Ill
  - as blueprint for incremental reform, 4294
  - from Carter administration, 4293
- National Practitioner Data Bank (NPDB), 4436
- National Registry of Evidence-based Programs and Practices, 4296
- National Resident Matching Program (NRMP)
  - match rates, 4407*t*
  - positions offered and filled, 4407*t*
  - residency position funding, 4407
  - resident recruitment, 4406–4407
- National Social Life, Health, and Aging Project (NSHAP), 4257
- National Standards for Cultural and Linguistically Appropriate Services (CLAS), 4380
- National Survey of American Life, 800
- National Survey on Drug Use and Health (NSDUH)
  - on hallucinogen use, 1321–1322, 1322*f*, 1325
  - on opioid use, 1355, 1363
  - prevalence, 1262–1263, 1282
  - on sedative-hypnotic drug use, 1378, 1378*f*
- National Training Laboratories (NTL), 4597
- National Vietnam Veteran Longitudinal Study (NVVLS), 1816
- National Vietnam Veterans Readjustment Study (NVVRS), 1816, 1822
- Native Americans. *See also* Demographics; Race/ethnicity
  - foster care and, 3819
  - suicide rates, 3681, 3681*f*, 3695, 3695*f*
- Natural disasters, 2565. *See also* Disaster psychiatry
- Natural groups, description of, 4597
- Natural killer cell activity (NKCA)
  - enhancement of, 186
  - immune responses and, 2320
  - stress and, 2338
- Natural killer (NK) cells
  - formation of, 179
  - innate immunity and, 181

- Natural products in CAIM medicine
  - B vitamin supplements, 3981
  - ginkgo biloba supplements, 3980–3981
  - omega-3 polyunsaturated fatty acid supplements (PUFA), 3980, 3984
  - S-adenosyl-L-methionine, 3984
  - St. John's wort, 3984
- Natural selection, evolution of, 771–772
- *Nature of Clinical Evidence* (Erikson), 905, 906
- Nausea and vomiting
  - cannabis use for, 1311
  - chemotherapy-related
    - psychiatric problems related to, 2271–2272
  - in dying patients, 2296
  - psychiatric patients
    - brain tumors and, 493
  - RT-related
    - treatment of, 2271
  - rumination disorder and, 3631, 3645
- Navane. *See* Thiothixene
- Nazi Holocaust survivors, amnesia, 1887
- NBME. *See* National Board of Medical Examiners (NBME)
- NCCAM. *See* National Center for Complementary and Alternative Medicine (NCCAM)
- NcRNA, role of, 2138
- NCS-A. *See* National Comorbidity Survey Replication Adolescent Supplement
- NCS-R. *See* National Comorbidity Survey Replication
- ND-PAE. *See* Neurobehavioral disorder associated with prenatal alcohol exposure (ND-PAE)
- Near-death experiences, 2589–2590
- Neck trauma, headaches attributable to, 620–621
- Necromania, 4597
- Necrophilia, 4597
- Necrophobia, 4597
- Need, definition of, 4597
- Needle phobia, 4597
- Needle sharing, 578
- Needles, bacterial infections and, 1011
- Nefazodone (Serzone), **3125–3132**

- absorption, 3125
- anxiety disorder, 3126–3127
- black box warning, 2919, 3125, 3128, 3129
- cardiovascular effects of, 2197*f*
- chemistry, 3125, 3125*f*
- combination psychotherapy, 3126
- for depression, 3126, 3127, 4124
  - in cardiac patients, 2196
- discontinuation syndrome, 3130
- distribution, 3125
- dosage, 573*t*
- dosage/administration, 3131
- drug interactions, 2199, 3130–3132
  - psychotropic agents, 3130–3132
- elimination, 3125
- history, 3125
- illness anxiety disorder, 1838
- lactation and, 3130
- major depression, 3126, 3127–3128
- metabolism, 3125
- mood disorders, 3126, 3127
- organ/system effects of, 3128–3129
  - cardiovascular, 3128
  - CNS, 3128
  - genitourinary, 3128–3129
  - GI, 3129
- overdose, 3129–3130
- pain, 3127
- pharmacological actions, 3125–3126
  - pharmacodynamics, 3125–3126
  - pharmacokinetics, 3125
- precautions/adverse effects, 3129–3130
- pregnancy and, 3130
- REM sleep and, 1717
- schizophrenia, 3127
- side effects, 573*t*, 1680, 2261
- substance abuse, 3127

- therapeutic indications, 3126–3128
- withdrawal, 3130

- Nefiracetam, 488
- Negative affect
  - family and couples therapy and, 2758
  - treatment principles and, 2297–2298
- Negative emotion threshold, 3480
- Negative hallucinations
  - dissociative identity disorder and, 1911, 1913
- Negative ions, high-density
  - as CAM therapy, 2558, 2559
  - scientifically validated therapy with, 2546
- Negative punishment, 2872
- Negative reinforcement. *See also* Reinforcement
  - definition, 2872
  - mechanisms of, 1283
- Negative symptoms
  - defined, 3725, 4597
  - neurocognitive impairment, 1507
  - schizophrenia and, 1415–1419, 1416*t*, 1418*t*, 1507, 1528
- Negative therapeutic reaction, 2657
- Negative thoughts, management of, 2326, 2326*t*
- Negative transference, 2651. *See also* Transference
- Negativism, 4597
- Neglect
  - attachment disorders, 3655, 3669
  - conduct disorder and, 3616
  - elder, 1200
  - as maltreatment, 3829
- Neglect syndrome, 980
- Neighborhood contexts, 4383
- Nembutal. *See* Pentobarbital
- Nemeroff, Charles, 1162
- NEO Personality Inventory 3, 936, 998*t*, 1001
- NEO PI assessment, 2161
- Neocortex
  - afferent inputs, 18–19

- development of
  - *cis*-regulation of transcription, 442, 443*f*
- layers of, 17–18
- neuroimaging, 2815
- regions, 19–20
- transcriptional differences, 441–442, 441*f*

- Neodissociation theory, 1870, 1895, 1895*t*, 1899
- Neologism, 1420, 4597
- Neonatal abstinence syndrome (NAS), 3838–3839
- Neonatal adaptation syndrome, 2519–2520
- Neonatal factors, autism spectrum disorder and, 3575, 3582
- Neonatal ventral hippocampus lesion (NVHL), 410
- Neonates
  - with congenital heart disease, 2198
  - development of, 2505
  - end-of-life care decisions, 2312
  - influences on aggression, 2425
- Neoplastic disorders, laboratory evaluations, 1015–1016
- Neoral. *See* Cyclosporine
- Neosis, 4598
- Neostigmine (Prostigmin), 73
- Nernst equation, 132
- Nernst potential, 131
- Nerve blocks, 2397–2398
- Nerve conduction studies (NCS), 2389
- Nerve growth factor (NGF), 46
- Nervous breakdowns, description of, 4597
- Nervous system. *See also* Autonomic nervous system; Central nervous system; Peripheral nervous system
  - buspirone effects on, 3012
  - development of, 99
  - embryonic, 41–42
  - overview of development, 40–46
  - vilazodone effects on, 3346
- NESARC. *See* National Epidemiological Survey on Alcohol and Related Condition (NESARC)
- Net magnetization
  - definition, 260

- detection of, 261
- Netherlands Mental Health Survey and Incidence Study (NEMESIS), 1452–1453
- Netrin 1 expression of, 53
- Neural codes, electrical activity and, 402–405
- Neural crest, derivation of, 40
- Neural development, **39–61**
- Neural mapping, 2840
- Neural oscillations
  - discoordination of, 409
  - synchrony in, 407–408
- Neural plate
  - induction of, 40
  - neurulation and, 40–41
- Neural ridges, formation of, 40
- Neural tube
  - formation of, 40
  - primary vesicles, schematic representation of, 9*f*
  - vesicles in, 8
  - VZ potentiators, 42
- Neuralgias, treatment of, 2967, 2968, 2977
- Neurasthenia
  - about, 1647–1648
  - description of, 4597
- Neuregulin (NRG1)
  - astroglia and, 52
  - bipolar and schizophrenic disorders and, 1610, 1625
  - schizophrenia and, 317
- β-Neurexin, binding of, 78
- *Neurexin* gene family, 314
- Neuroacanthocytosis
  - histology, 667*f*
  - MRI imaging of, 667*f*
- Neuroactive peptides, 141
- Neuroactive steroids, 1746
- Neuroanatomy
  - functional, **3–39**
  - neuropsychiatric, 449–452

- Neurobehavioral disorder associated with prenatal alcohol exposure (ND-PAE), 637–638, 638*t*
- Neurobehavioral syndromes, focal, 460–468
- Neurobiological factors
  - ADHD, 3589
  - conduct disorder, 3615
  - depressive disorders, 3675, 3689
  - OCD, 3696–3697
  - of opioid-related disorders, 1357–1360, 1357*f*, 1358*f*
  - schizophrenia, 1461–1462
  - of specific learning disability, 3525
  - as symptomatology consideration, 1116
  - temperament and, 3483–3484
  - of terrorism impact, 3915
- Neurobiology
  - dissociation and dissociative disorders, 1889–1894
    - depersonalization, 1890–1892
    - dissociative amnesia, 1892–1893
    - dissociative identity disorder, 1893–1894
  - fear and anxiety, 1741–1742
  - hypnosis, 2712–2713
  - mood disorders and, 1709–1719
  - psychotherapy and, 2812–2813
  - PTSD, 1817
  - suicide, 2619
- Neurocardiogenic syncope, 2193
- Neurocentrism, **780–788**
  - addiction and brain disease fallacy, 781–782
  - brain *vs.* mind, 781
  - cross-references, 788
  - and free will, 785–788
  - legal issues, 784–785
  - neural correlates and voluntariness, 782–784
- Neurochemical effect of, 1266–1267
- Neurochemical effects
  - ADHD and, 3599
  - of alcohol, 1266–1267
  - autism spectrum disorder and, 3579

- conduct disorder, 3614–3615
- OCD, 3696–3697
- tic disorders and, 3639–3640, 3653–3654

- Neurochemical systems
  - depersonalization and, 1891–1892
  - fear and anxiety and, 1741–1742, 1757–1758
- Neurochemistry, description of, 4597
- Neurocircuitry
  - anxiety disorders and, 1752–1753
  - biologically relevant learning, 1757
  - contextual fear structures, 1755–1756
  - emotion processing and, 1756
  - fear, preclinical studies, 1752–1753
  - fear expression and, 1754
  - fear learning, 1753–1754, 1753*f*
  - fear regulation, 1754–1755, 1755*f*
- Neurocirculatory asthenia. *See* Da Costa syndrome
- Neurocognition
  - ECT and, 3297–3298, 3297*t*
  - neurocognitive disorders and
    - assessment, 1511, 1511*t*
    - clinical importance, 1508–1509
    - comparative nosology, 1242, 1242*t*, 1244, 1246–1247
    - course and prognosis, 1243–1245, 1245*t*, 1247–1248
    - defined, 1236–1237, 1242, 1244, 1246–1247
    - diagnosis and clinical features, 1235*t*, 1243–1246, 1245*t*, 1248
    - differential diagnoses, 1243, 1245, 1247–1248
    - due to another medical condition, **1246–1247**
    - due to HIV infection, **1244–1246**
    - due to traumatic brain injury, **1242–1244**
    - epidemiology, 1242–1244, 1246, 1247
    - etiology, 1243, 1244, 1246–1247
    - functional capacity, 1509
    - inhalant-induced, 1338–1339
    - laboratory evaluation, 1247
    - natural history, 1504–1506, 1505*f*
    - pathology and laboratory examination, 1243, 1245, 1248

    - schizophrenia and, 1501–1504, **1501–1512,** 1502*f*
    - schizophrenia symptoms correlation, 1506–1508
    - sedative-, hypnotic-, anxiolytic-induced, 1381
    - substance/medication-induced mild/major, **1247–1248**
    - testing, 1502–1504
      - attention and vigilance, 1502–1503
      - immediate/working memory, 1504
      - processing speed, 1503
      - reasoning and problem solving, 1503
      - social cognition, 1504
      - verbal fluency, 1503–1504
      - verbal memory, 1503
      - visual memory, 1503
    - treatment, 1244, 1246–1248, 1509–1510
- Neurocognitive disorders
  - after TBI, 528–529
  - genetic factors, 4022
  - HIV-associated in adolescents, 647
  - trazodone drug studies, 3251
  - treatment, 3040–3041, 3040*t*
- Neurocognitive symptoms in rehabilitation, 4370
- *NeuroD* gene, 51
- Neurodegeneration
  - aggression and, 2495–2496
  - cytokines and, 195–196
- Neurodegeneration with brain iron (NBIA), 672
- Neurodegenerative disorders
  - differential diagnoses, 1198–1199
  - treatment
    - MAOIs, 3123
- Neurodermatitis
  - hypnosis and, 2730
  - psychosomatic component, 915
- Neurodevelopment
  - ADHD (*See* Attention-deficit/hyperactivity disorder (ADHD))
  - autism spectrum disorder and, 3576–3577
  - classification of, 1158, 1159

- developmental coordination disorder (*See* Developmental coordination disorder)
- disorders of
  - ages of onset, 438*f*
  - chromatin-mediated, 387*t*
  - noncoding mutations in, 446
  - somatic mutations in, 446
- extracellular factors, 51–54
- HIV and AIDS sequelae, 3854–3855
- intellectual disabilities (*See* Intellectual disability (ID))
- language disorders and, 3546
- motor skills disorders (*See* Developmental coordination disorder)
- overview, 3521
- schizophrenia, early onset and, 3727–3728
- schizophrenia and, 1456
- specific learning disorder (*See* specific learning disorder)
- transcriptional regulation, 442–444

- Neurodevelopmental cross-disorder, 3452, 3452*t*
- Neuroendocrine disorders
  - examples of, 676–684
  - neuropsychology of, **655–685**
- Neuroendocrine factors, sexuality in aging, 4256
- Neuroendocrine system
  - ANS response and, 2337–2338
  - cytokine effects on, 197
  - FGA effects on, 3069
  - OCD, 3696–3697
  - stimulant-use disorder and, 1284
- Neuroendocrine tumors, presentation of, 684
- Neuroendocrinology, 4597
- Neurofeedback
  - definition, 2559
  - QEEG analysis, 1041
- Neuroferritinopathy
  - causes of, 672
  - presentation of, 672–673
- Neurofibromatosis, 40
- Neurogenesis, **39–61**

- in cerebellar development, 45*f*
- cytokines and, 195–196
- nitric oxide and, 107
- postnatal, 58–60
- regulation of, 5, 42–43

- Neurogenesis, patterns of, 43–45
- Neurohormones, 4597
- Neurohypophysis, 677–678
- Neuroimaging, **3458–3478,** 4551. *See also* Forensic psychiatry, neuroimaging
  - addictive disorders, 3475
  - ADHD, 3463–3464, 3463*f*, 3589, 3599
  - of adolescent development, 3365, 3367
  - of adolescent psychiatric disorders, 3349–3350
  - in aging adult (*See* Geriatric psychiatry)
  - alcohol use disorder, 3475
  - anorexia nervosa, 3475
  - anxiety, 1752–1768
    - studies of, 1758–1760
  - anxiety disorders, 1760–1768
  - attachment disorders and, 3664, 3678
  - autism spectrum disorders, 3475–3476, 3476*f*, 3477*f*, 3583
  - bipolar disorders
    - child/adolescent, 3688, 3702
  - bipolar polar disorders, 3464–3466
  - of brain development, **3461–3463**
    - gender differences, 3462–3463, 3463*t*
    - gray matter, 3361*f*, 3461–3462, 3462*f*, 3462*t*
    - white matter, 3462, 3462*f*, 3463*f*
  - brain stimulation and, 3303–3304
  - cannabis use disorder, 3475
  - child maltreatment, 3833
  - of child psychiatric disorders, 3349–3350
  - childhood-onset schizophrenia, 3474–3475, 3474*f*
  - in clinical neuroscience, 466
  - cognitive and emotional activation, 1760
  - on cognitive development, 3369, 3371
  - conduct disorder, 3615

- core psychoanalytic constructs, 2814
- delirium, 1184
- depressive disorders, child/adolescent, 3675, 3689
- developmental, 1765–1767, 1765*f*
- dyslexia, 3477
- factitious disorder imposed on self, 1851
- fear, 1752–1768
  - studies of, 1758–1760
- functional, 466
- further research, 3477
- general anxiety disorder, child/adolescent, 3470–3471, 3471–3472*f*
- generalized anxiety disorder, 1763–1764
- intellectual disability and, 3501
- language disorders, 3546
- major depressive disorder, child/adolescent, 3464–3466, 3465*f*
- for MCI, 1225–1226, 1226*t*
- and memory, 736–738, 737*f*, 738*f*
- memory suppression or repression, 1893
- neurocognitive disorder due to HIV, 1245
- neurosurgical postoperative, 3315–3316
- obsessive-compulsive disorder, 3466–3470, 3467–3470*f*
- opioid-related disorders, 1359–1360, 1360*f*
- opioid use disorders, 1359–1360, 1360*f*
- outcome prediction and, 2815
- pediatric autoimmune neuropsychiatric disorders associated with streptococcus (PANDAS), 3470
- pediatric neuropsychiatric disorders, **3463–3477**
- posttraumatic stress disorder, 3471
- psilocybin effects, 2887
- psychodynamic principles and, 2814–2815
- psychotherapy and, 2675
- schizophrenia
  - functional, **1475–1488**
  - molecular, **1488–1501**
  - structural, **1463–1475, 1465***f*, **1466***f*
- schizophrenia, childhood-onset, 3474–3475, 3474*f*
- social anxiety disorder, 1763
- specific phobias, 1762–1763

- structural techniques, 1759
- stuttering, 3561
- substance-related disorders, 3475
- of suicide attempters, child/adolescent, 3684, 3698
- techniques, **3458–3461**
  - chemical, 3459–3460, 3459*f*
  - functional, 3460
  - morphological studies, 3458–3459
  - pediatric considerations, 3460–3461, 3460*f*
- of tic disorders, 3639–3640, 3653–3654
- Tourette syndrome, 3471–3473, 3473*f*

- Neurokinin A, 2388–2389
- Neuroleptic drugs. *See* Antipsychotics/neuroleptics
- Neuroleptic-induced acute akathisia, 2938
  - clinical presentation, 2941
  - implications, 2941
  - pathophysiology, 2941
  - treatment options, 2941, 2955
- Neuroleptic-induced acute dystonia, 2940–2941
  - clinical presentation, 2940
  - implications, 2940–2941
  - pathophysiology, 2940
  - treatment options, 2940–2941
- Neuroleptic-induced parkinsonism
  - clinical presentation, 2938
  - implications, 2938
  - pathophysiology, 2938
  - treatment options, 2938, 2959
- Neuroleptic-induced tardive dyskinesia, 2941–2942
  - clinical presentation of, 2941–2942
  - FGAs, 3074–3075
  - pathophysiology of, 2942
  - treatment options/implications, 2942
- Neuroleptic malignant syndrome (NMS), 1023–1024
  - antipsychotics and, 2381
  - catatonia and, 459
  - clinical manifestations of, 1148

- clinical presentation, 2938–2939, 2939*t*
- DA receptor agonists and, 3238
- description of, 635, 4597
- FGA, 3073–3074
- implications, 2939–2940
- pathophysiology, 2939
- SSRIs and, 1779
- treatment options, 2939–2940, 2940*t*, 3238

- *Neuroligin* gene family, 314
- Neurological amnesia, 4597
- Neurological deficits, brain tumors and, 494
- Neurological disorders
  - description of, 4597
  - laboratory evaluation, 1015
  - symptoms
    - language disorders and, 3548
    - lithium and, 3096
    - MAOI effects on, 3117
    - opioid receptor antagonist effects on, 3144
    - review of, 454
  - treatment
    - anticonvulsants, 2968, 2970, 2974–2977
    - ECT, 3287–3288
    - lamotrigine, 3084
    - valproate, 3276
- Neurological examination
  - aging adult, 3976
  - components of, 456–460
- Neurologists, description of, 4597
- Neurology, description of, 4597
- Neuromedin N, 87*f*
- Neuromelanin, 3991
- Neurometabolic disorders, **655–685**
- Neuromuscular diseases, **627–636**
- Neuromyotonia, 2270
- Neuronal ceroid lipofuscinosis (NCLs)
  - features of, 467*t*

- presentation, 658
- Neuronal circuitry, aggression and, 2478, 2479*f*
- Neuronal hyperexcitability
  - about, 1609–1610
  - lithium for, 1609
- Neuronal nitric oxide synthase (nNOS), 104–105
- Neuronal plasticity, 4597
  - in brain development, 3358
  - and natural selection, 772
  - neural, and memory, 728–729
- Neuronal signaling, basis of, **130–146**
- Neurons
  - action potentials, 134*f*
  - architecture of, 5
  - classification of, 3
  - connections of, 5–8
  - dendrites, in aging brain, 3990–3991
  - description of, 4597
  - development of, 3358
  - developmental changes in, 5
  - features of, 4*f*
  - growth and integration of, 687–688
  - hippocampal
    - formation of, 2140
  - ion channels, 131*f*
  - major features of, 4*f*
  - mirror
    - self *versus* others and, 2895
    - and social cognition, 693
  - morphology
    - addiction and, 380–381
  - multipolar, 3
  - oscillatory firing
    - behavior and, 143
  - peptidergic, 88*f*
  - potassium ion (K^+) channels, 136–139
  - process outgrowth, 53–54

- pyramidal
  - drawing of, 16*f*
  - layers, 6
- regions of, 3
- stellate, drawing of, 16*f*
- synaptic plasticity, 380–381
- unipolar, 3
- von Economo class of, 21

- Neurontin. *See* Gabapentin
- Neuropathic pain
  - description of, 2292*t*
  - treatment, 2977
- Neuropathies
  - alcoholic, 630
  - connective tissue disorders and, 631–632
  - diabetic, 2975, 2977
  - diagnosis of, 629
  - HIV-associated, 632
  - nutritional, 630
  - treatment of, 629
  - vasculitic, 631–632
- Neuropathology of schizophrenia, **1457–1463**
- Neuropeptide transmitters
  - examples, 84*t*
  - structures, 85*t*
- Neuropeptide Y (NPY), **176**
  - action of, 89
  - anxiety disorders and, 1743*t*, 1749
  - anxiolytic effects of, 169
  - dissociative symptoms and, 1891–1892
  - PTSD and, 1819
    - genetic variants and, 1820
    - prevention, 1825
  - Special Forces soldiers and, 1897
- Neuropeptides
  - distribution of, 87–88
  - effects on skin cells, 2337, 2338*t*

- function of
  - studies of, 85–86
- immunological effects of, 185*t*
- neuropsychiatric disorders and, **84–97**
- pruritus and, 2342
- regulation of, 87–88
- signaling, 88–90

- Neuropharmacology
  - sedative-, hypnotic-, anxiolytic-related disorders, 1374–1377, 1376*f*
- Neurophenomenology, 1889
- Neurophysiology
  - brain stimulation and, 3303–3304
  - clinical, certificate examinations, 4405*t*, 4427
  - description of, 4597
  - ECT, 3284
  - hypnosis, 2712–2713
  - mood disorders, 1609–1610, 1716–1717
- Neuroplasticity
  - brain stimulation and, 3301–3302
  - and clinical disorders of aging brain, 3988
  - ECT, 3285
  - narrative psychiatry and, 2840
  - positive psychiatry, 4284
- Neuroprotective effects
  - barbiturates, 2987
  - clonidine, 2949
  - guanfacine, 2949
- Neuroprotective therapies for TBI, 523–524
- Neuropsychiatric disease(s)
  - approach to patients, **448–474**
  - clinical evaluation in, 454–455
  - factors in, 54–58
  - infectious causes of, 594*t*, 597
  - physical examination, 455–460
- Neuropsychiatric disorders
  - aggression and, 2487–2497
  - child

    - neuroimaging of, **3463–3477**
  - dementia and, 1196–1197
  - ion channels and, 143–144
- Neuropsychiatric Inventory (NPI), 1195
- Neuropsychiatric workup, 612–616
- Neuropsychiatry, modularity of, 452–454
- Neuropsychoanalysis, 889–890
- Neuropsychological Assessment Battery (NAB), 988, 991
- Neuropsychological assessments, 468
  - for child psychiatric evaluations, **3413–3415, 3429–3436**
    - advances in, 3429, 3431
    - application of, 3431–3433
    - on attention, 3433–3434
    - on executive functioning, 3434
    - as functional assessment, 3429
    - further considerations for, 3436
    - on language, 3434–3435
    - on memory, 3433
    - overview, 3413–3415, 3416–3417*t*, 3418*t*
    - referrals for, 3404
    - on sensory/motor functioning, 3435–3436
    - on visuoperceptual functioning, 3435
  - for mild cognitive impairment, 1224–1225
- Neuropsychological evaluation, 4002–4008
  - activities of daily living, 4007
  - common uses of, 4003–4004
    - attention and vigilance, 4005
    - cognitive screening, 4004
    - crystallized knowledge, 4004–4005
    - executive function and working memory, 4006–4007
    - language reception and expression, 4005
    - learning and memory, 4005–4006
    - mental processing speed, 4007
    - psychomotor speed, 4007
    - visual cognition, 4006
  - definitions, 4003
  - historical development, 4003

- mood and anxiety, 4007
- sensory impairment, 4007
- test result interpretations, 4007

- Neuropsychological Test Battery, 4140
- Neuroreceptors
  - description of, 4597
  - imaging studies, 1759
- Neuroscience
  - basic systems, 401–411
  - description of, 4597
- Neurosis
  - description of, 4597
  - Horney theories, 920
  - neurobiology and, 1710
- *Neurospora crassa* (bread mold), 149
- Neurosteroids
  - anxiety disorders and, 1743*t*, 1748
  - background, 117
  - GABA and, 1747
  - neurodevelopment and, 117
  - neuroprotection and, 117
  - PTSD and, 1819
- Neurostimulation therapies. *See also* Electroconvulsive therapy (ECT)
  - deep brain stimulation, 4163
  - future directions, 4163–4164
  - magnetic seizure therapy, 4163
  - rapid transcranial magnetic stimulation, 4163
  - transcranial direct current stimulation, 4162
  - vagus nerve stimulation, 4162–4163
- Neurosurgery, **3313–3322**
  - clinical studies, 3316
  - current state of, 3320–3321
  - future directions, 3321–3322, 3321*f*
  - history, 3314
  - indications/contraindications, 3314–3315
  - lesion procedures, 3316–3317, 3316*f*, 3317*f*
  - multidisciplinary case review, 3315

- overview, 3313–3314
- for paraphilias, 2020
- patient selection, 3314–3315
- postoperative care, 3315–3316
- preoperative work-up, 3315
- research, 3321–3322
- treatment outcomes, 3317–3321

- Neurosyphilis
  - presentation, 592–593
  - tests for, 593
  - treatment, 593
  - types of, 592
- Neurotensin (NT)
  - administration of, 177
  - function of, **177**
  - overview, **95–96**
- Neurotic depression, 1653
- Neurotic disorder. *See* Neurosis
- *Neurotic Personality of Our Time* (Horney), 4538
- Neurotic process, description of, 4597
- Neuroticism
  - components of, 2145
  - confounding factors, 2161
  - in Five-Factor Model, 935–936
  - and psychopathology, 938–939
- Neurotoxicity
  - conduct disorder, 3615
  - dissociation, violence and, 2491
  - of hallucinogens, 1323
- Neurotoxins
  - description of, 4597
  - parkinsonian diseases and, 543
  - sites of action, 136
- Neurotransmission
  - conductance mechanisms in, 141–142
  - delirium, 1181
  - effects of CBI on, 112–114

- glutamatergic
  - plasticity in, 79
- neuroimaging, 1490–1491
- schizophrenia, 1460–1461
- stimulants effect on, 1283

- Neurotransmitter receptors
  - intracellular effects, 61
  - structures of, 142–143
- Neurotransmitter(s). *See also* Biogenic amine neurotransmitters
  - aggression and, 2479–2480
  - biogenic amine, **61–75**
  - biosynthesis of, 86–87
  - classes of, 141
  - cytokine effects on, 196
  - gases as, 104–118
  - immune cell
    - receptors for, 184*t*
  - immunological effects of, 185*t*
  - ion channels and, 141–143
  - novel, 103–119
  - receptors, 131
  - responses to stress, 2318
  - sex function and, 1971*f*
  - sexuality and, 1957
  - signal integration, 2916, 2916*f*, 2916*t*
  - verbal therapies' effects on, 2812–2813
  - violence and, 2420–2423
- Neurotrophic factors, **97–103**
- Neurotrophins
  - axonal transport, 99–100
  - behavior and, 100
  - cell survival and, 99
  - family, 97–98
  - function of, 51
  - receptors
    - signaling by, 98–99, 98*f*
  - synaptic plasticity and, 100

- therapeutic potential of, 102–103
- Neurulation
  - mechanisms of, 40*f*
  - neural plate and, 40–41
  - timing of, 40
- *Neurypnology, or the Rationale of Nervous Sleep* (Braid), 4535
- Neutral Identity State (NIS)
  - dissociative identity disorder, 1894
  - structural dissociation theory and, 1903
- Neutral state paradigms, 1759
- Neutrophils, differentiation of, 179
- New Age spirituality, 2577, 2583
  - penance, spirituality and, 2583
- New Freedom Commission on Mental Health
  - *Achieving the Promise: Transforming Mental Health Care in America,* 4315
  - under Bush (George W.) administration, 4294
  - national blueprint for workforce development, 4306
- New Generation Sequencing (NGS), 147, 148*f*
- *New York City v. Brad H. Case,* 4461–4462
- New York-Montreal Long-Term ADHD Treatment Study, 3789
- New York Psychoanalytic Institute, 2674
- New York Teacher Rating Scale for Disruptive and Antisocial Behavior (NYTRS), 3402
- Newborns. *See also* Infants; Neonates; Premature infants
  - sleep in, 342
- Newmann-Keuls procedure, 834
- Next Accreditation System (NAS), 4413–4414
- Next-generation sequencing (NGS), 3452–3453
- NHIS. *See* National Health Interview Survey
- NIA-AA. *See* National Institute on Aging and the Alzheimer Association
- NIAAA. *See* National Institute of Alcohol Abuse and Alcoholism (NIAAA)
- Niacin, schizophrenia and, 2548
- Niacin flushing response, 2545
- Nicardipine (Cardene), 3019
- Nichols, Charles, 4331
- *Nicomachean Ethics* (Aristotle), 928
- Nicotine. *See also* Smoking; Tobacco
  - addiction, 2892

- Alzheimer disease risk factor, 1205
- brain development and, 58
- dependence on, 1345
  - description of, 4597
  - obesity and, 2215
  - in transplant populations, 2367
  - treatment for, 3142
- drug reward and, 373–374
- fetal exposure to, 2484
- patches, 2266
- pharmacology of, 1345
- psilocybin and, 2892
- replacement therapies, 1349–1350
- withdrawal, 2258, 2948

- Nicotine replacement therapies, 1349–1350
- NIDA. *See* National Institute of Drug Abuse (NIDA)
- Niemann–Pick disease type C (NPC)
  - description of, 665–666
  - diagnosis of, 665–666
  - features of, 467*t*
  - MRI imaging, 666*f*
  - presentation of, 655
- Nietzsche, Friedrich, 2821
- Nifedipine (Procardia)
  - chemistry, 3016, 3016*f*
  - drug studies, 3019
  - pharmacokinetics, 3017
  - therapeutic indications, 3020–3021
- Night eating syndrome (NES)
  - behavioral treatment of, 2220
  - description of, 4598
  - obesity and, 2214*t*
- Night hospitals, description of, 4598
- “Night owls,” 2097
- Night-terrors, 4598
- Nightmare disorders
  - description of, 4598

- diagnostic criteria, 2099–2100
- Nightmares
  - description of, 4598
  - dissociative amnesia, 1900, 1932
  - dissociative fugue, 1933–1934
  - dissociative identity disorder, 1915*t*
    - treatment, 1927–1928, 1927*t*
  - in PTSD, 1819, 1821
- Nihilism, description of, 4598
- Nihilistic delusions, 4573, 4598
- Nijenhuis, Ellert, 1903
- NIMH. *See* National Institute of Mental Health (NIMH)
- Nimodipine (Nimotop)
  - acute mania, 1687
  - chemistry, 3016, 3016*f*
  - dosage/administration, 3021
  - drug studies, 3019–3020
  - organ/system effects of, 3020
- Nimotop. *See* Nimodipine
- Nissl stains, 5, 5*f*
- Nitrendipine, drug interactions, 3021
- Nitric oxide (NO)
  - behavior and, 107
  - discovery of, 104–108
  - enhancers
    - for sexual dysfunction, 1978
  - generation of, 105*f*, 106*f*
  - mechanisms of action, 105–106
  - mood disorders and, 107
  - neurotransmission, 106–107
  - release of, 1955
  - schizophrenia and, 107
  - signaling by, 105*f*, 106*f*
  - synthesis of, 104–105
- Nitric oxide synthase (NOS)
  - inhibitors of, 107
  - neuronal (NOS1), 445

- NO synthesis and, 104–105
- Nitrogen atoms, ^{13}N, 287
- Nitrogen narcosis. *See* Rapture-of-the-deep syndrome
- *S*-Nitrosylation, targets activated by, 106
- Nitrous oxide, depression, 1681
- Nitrous oxide-related disorders, 1340–1341
- Nizatidine (Axid), 2982
- Nizoral. *See* Ketoconazole
- NLAAS. *See* National Latino and Asian American Study
- NMDA-receptor agonists, 1312–1315, 1317
- NMDA-receptor antagonists, 1238
- NMDA receptors. *See* *N*-methyl-D-aspartic acid (NMDA) receptors
- NMR imaging. *See* Nuclear magnetic resonance (NMR) imaging
- NMS. *See* Neuroleptic malignant syndrome
- NO. *See* Nitric oxide (NO)
- Nobel Prizes, 4294
- Nociceptors, cell bodies of, 326
- Noctec. *See* Chloral hydrate
- Nocturnal enuresis, 3650, 3651, 3653, 3654, 3657, 3664, 3665, 3667, 3668, 3671
  - pediatric, 3937–3938
- Nodes of Ranvier, 134
- Noguchi, Hideyo, 4537
- Nomifensine (Merital), 2919
- Nominal aphasia, 4598
- Non-24-hour sleep-wake circadian rhythm disorder, 4598
- Non-abstinence goals, 1351–1352
- Non-nicotine prescription medications, 1350
- Non-small cell lung cancer (NSCLC), 2268–2269
- Non-suicidal self injury, 4598
- Nonaffective hallucinations, 4598
- Nonafluent aphasia. *See* Motor aphasia
- Nonbarbiturate sedative-hypnotic agents, **2989–2990**
- Nonbenzodiapzepines, 4054, 4054*t*
- Nonbenzodiazepine sedative-hypnotic agents, **2989–2990**
- Noncoding genes, 149, 151–153
- Noncoding RNAs, 152
- Noncomplementary roles. *See* Roles

- Noncompliance with medical treatment, 4598
- Nonconvulsive status epilepticus (NCSE), 219
- Nondirective approach, 4598
- Nondisclosure, 2290
- Nonendogenous depression, 1710
- Nonepileptic seizures, 514–517. *See also* Seizures
  - epileptic seizures *versus,* 515*t*
- Nongay identities, 1997–1999
- Nonhuman primates (NHPs)
  - cognitive function in, 31
  - studies using, 8
- Non–insight-oriented interpersonal psychotherapies (IPT), 2775
- Noninstitutionalized individuals, 4326
- Nonmentalizing modes, 2896–2897
- Nonparametric, definition of, 4598
- Non–possession-form dissociative identity disorder, 1879, 1885. *See also* Dissociative identity disorder
  - childhood trauma and, 1904
  - delusional separateness and, 1907
  - differential diagnoses, 1916
  - dissociative hallucinations and, 1911
  - factitious dissociative identity disorder and, 1917
  - passive influence symptoms, 1912
  - possession-form compared with, 1914*t*
  - state shifting in, 1912
  - treatment, 1918–1931
- Nonpsychiatric disorders, lithium treatment for, 3094, 3094*t*
- Nonrapid eye movement (NREM) sleep, 1716
  - arousal disorders and, 2099
  - control of, 346
  - description of, 339
  - overview, 346–347
- Nonreactive depression. *See* Endogenous depression
- Nonsanctioned aggression
  - sanctioned aggression *versus,* 2472
- Nonselective cation channels, 135
- Nonseminomas, treatment of, 2268
- Nonspecific diagnostic codes, 1155, 1156

- Nonspecific effects, behavior therapy, 2704
- Nonsteroidal anti-inflammatory drugs (NSAIDs), 2293
  - Alzheimer disease prevention and, 1205
  - ineffectiveness for Alzheimer disease, 4148
  - SSRI interactions with, 202
- Nontransference concepts
  - real relationship, 2655
  - therapeutic alliance, 2654–2655
- Nonverbal communication in patient interviews, 958, 960
- Nonverbal learning disabilities (NVLDs), 2436
- Nootropics
  - for Alzheimer disease, 4146
  - for intellectual disability, 3515–3516
- Noradrenaline. *See* Norepinephrine
- Noradrenergic and specific serotonergic antidepressants (NaSSAs), 1780
- Noradrenergic neurons
  - degeneration of, 546
  - fear conditioning, PTSD and, 1818
  - mood disorders and, 1713–1714
- A2 noradrenergic neurons and appetite, 363–364
- Noradrenergic reuptake inhibitors, 1787
- Noradrenergic system
  - anxiety disorders and, 1742, 1743*t*, 1744
  - central, fear and anxiety and, 1757–1758
  - gambling disorder and, 1807
  - mood disorders and, 1713–1714
  - PTSD and, 1820
  - tic disorders, 3638, 3645, 3652, 3659
- Nordazepam. *See* Desmethyldiazepam
- Norepinephrine (NE)
  - aggression and, 2480
  - anatomy of, 62
  - anxiety disorders and, 1742
  - buspirone and, 3009
  - depersonalization/derealization, 1892
  - diffuse distribution of, 1714
  - fear and anxiety and, 1757–1758

- functions of, 62
- GnRH release and, 169
- immunological effects of, 185*t*
- mediation by adrenergic receptors, 206
- mood disorders and, 1607, 1713
- plasma membrane transporter, 65–66
  - metabolism, 66
- prefrontal cortex function and, 67
- stress and, 1712, 1713
- synthesis, 62–63, 63*f*

- Norfluoxetine
  - breast milk excretion of, 3168
  - discontinuation syndrome, 3157
  - drug interactions, 3162, 3164
  - pharmacological actions, 3159
- Normality
  - aging, dementia and, 1177
  - concept of neuroimaging, 4483–4485
  - definition of, 4598
  - mental health and
    - cross references, 768
    - future directions, 767
  - mental health and, **752–768**
    - future directions, 766–768
    - model A: mental health as above normal, 753–756, 754*t*
    - model B: mental health as positive psychology, 756–758
    - model C: mental health as maturity, 758–759, 758*f*
    - model D: mental health as resilience, 759–762, 760*f*
    - model E: mental health as socioemotional intelligence, 762
    - model F: mental health as subjective well-being, 762–765, 764*f*
- Normalizing model, CBT, psychosis and, 2883
- Norman, Warren, 935
- Norpramin. *See* Desipramine
- North American Prodrome Longitudinal Study (NAPLS), 1470
- Nortriptyline (Aventyl, Pamelor)
  - chemistry, 3254, 3254*f*
  - child/adolescent, 3788–3789

- depression, 1681
- dosage/administration, 3264, 3265, 3265*t*
- dosages, 573*t*
- drug interactions, 3264
- drug studies, 3259
- ECT interactions with, 3295
- for fibromyalgia, 635
- organ/system effects of, 3260–3261
- pharmacological actions, 3255–3256, 3257, 3258
- side effects, 573*t*
- targets for, 69
- therapeutic indications, 3261

- Norvasc. *See* Amlodipine
- *Nosologia methodica* (de Sauvages), 4534
- Nosology
  - definition of, 4598
  - evolution of, 1153–1154, 1153*f*
- *Nosology, or a Systematic Arrangement of Diseases* (Cullen), 4534
- Nosophobia. *See* Pathophobia
- Not-for-profit hospitals, 4598
- Not guilty by reason of insanity (NGRI), 1924
- Not otherwise specified (NOS)
  - anabolic-androgenic steroid-related disorders, 1402–1403
  - pervasive development disorder, 3571
  - for psychiatric classification, 1155, 1156
  - for psychiatric diagnoses, 1115–1116
- Not otherwise specified (NOS), description of, 4598
- Note taking in patient interviews, 962
- Notman, Malkah, 2598
- Novelty Seeking
  - Cluster B PDs and, 2163
  - genetics of, 2167
  - harm avoidance and, 2146*f*
  - high, in children, 2159
  - personality disorders and, 2162*t*
  - psychobiological findings, 2131–2132, 2131*t*
  - secure attachments and, 2174

- TCI scales, 2164*t*
- NPC. *See* Niemann-Pick disease type C (NPC)
- *NPC1* gene, 665–666
- *NPC2* gene, 665–666
- NREM sleep
  - description of, 4598
- NSHAP. *See* National Social Life, Health, and Aging Project
- NTL. *See* National Training Laboratories (NTL)
- Nuclear DNA of mitochondrial origin (NUMTs), 155
- Nuclear factor-κB (NF-κB), 126
- Nuclear family, description of, 4598
- Nuclear magnetic resonance (NMR) imaging, **259–258**
  - basic physics of, 259–264
  - description of, 4598
  - schizophrenia and, 1464
- Nuclear magnetic resonance spectroscopy (NMRS), **259–258**
- Nucleic acids, description of, 149
- Nucleosomes
  - description of, 149
  - positioning of, 391
  - structure of, 388*f*
- Nucleotides, description of, 149
- Nucleus accumbens (NAc)
  - aggression and, 2477
  - aggression mediated by, 2476
  - dopamine receptors in, 66
  - dysfunction, depression and, 1709
  - histone acetylation, 397
  - neuroplasticity, 381*f*
  - neurotensin in, 177
  - pain information and, 327–328
- *NudF* gene, 52
- Null hypotheses, 832, 833, 4598
- Number dyslexia, 979
- Numbers, cardinal, 708
- Numbing, PTSD, childhood/adolescent, 3704–3705
- Nuremberg Code, 4240

- Nursing. *See* Breast-feeding; Lactation
- Nursing homes. *See also* Long-term care
  - health care costs for older adults, 4213
  - and sleep disorders, 4057–4058
- Nurturing, effects on violence, 2425
- Nutraceuticals, antidementia drugs, 4139
- Nutrient formulas, biological treatment with, 2552
- Nutrition. *See also* Diet and nutrition
  - and brain aging, 3994
  - fetal exposure to
    - antisocial behavior and, 2484
  - inattention to
    - manic syndrome and, 1640
  - positive psychiatry, 4285
- Nutritional deficiency, laboratory evaluation, 1016
- Nutritional rehabilitation, 2079–2080
- Nuvigil. *See* Armodafinil
- Nyalamide, 3110
- Nyctophobia, 4598
- Nymphomania, 1974
  - definition of, 2039
  - description of, 4598
- Nyquist frequency, 225
- Nyquist theorem, 225

O

- Obama Administration, 4318
- Obesity
  - adiposity sensing and, 362–363
  - antipsychotic drug effects, 4136–4137
  - and brain aging, 3995
  - classification of, 2211*t*
  - cognitive therapy, 2766–2767
  - course/prognosis of, 2211
  - diagnosis, 2211
  - economic costs of, 2211
  - environmental factors and, 2213
  - epidemic of
    - diabetes and, 2230
  - etiology of, 2212–2213
  - food addiction and, 2213
  - hypothalamic lesions and, 677
  - medical comorbidities, 2212*f*
  - mental illness comorbid with, 4529
  - pathology of, 2211
  - personality dimensions and obesity and, 2213
  - prevalence of, 2211*t*
  - psychiatric comorbidities of, 2213–2215
  - psychosomatic medicine, 2210–2211
  - with schizophrenia, 1542, 1546–1547
  - treatment of, 2215–2222
    - anticonvulsants, 2970
    - bupropion, 3004, 3006
    - commercial diets, 2215–2216
    - diet types, 2216*t*
    - lamotrigine, 3084
    - naltrexone, 3006, 3142–3143
- Object loss, depression and, 1605–1606
- Object Memory Test, 990–991
- Object permanence
  - during cognitive development, 3359
  - description of, 4598

- Object relations
  - description of, 4598
  - group psychotherapy and, 2737–2738
  - psychosis and, 2870
- Object Relations School, 1869
- Object relations theory, 883–886
  - ego and objects, 884–886
  - origins, 883–884
- Objective personality tests, **997–1007**. *See also specific* tests
  - Millon Clinical Multiaxial Inventory III, 1001
  - Personality Assessment Inventory, 1001
  - Personality Inventory-Revised, 1001
  - Rapid Personality Assessment Instruments, 1001–1003, 1002*t*
- Objective structured clinical examination (OSCE)
  - evaluation of, 4431
  - practical clinical skills and, 4429
- Objective tests, description of, 4598
- OBS. *See* Organic brain syndrome
- *Obscure Diseases of the Brain and Disorders of the Mind* (Winslow), 4536
- Observational learning, 2683
- Obsession, description of, 4598
- Obsessional jealousy, 1797, 4598
- Obsessional thinking, 1120–1121
- Obsessive-compulsive disorders
  - perinatal, 2515
  - related disorders (OCRDs) and
    - classification of, 1157, 1168, 1173
    - due to another medical condition, **1235***t*, **1255–1256**
- Obsessive-compulsive disorders (OCDs), **1785–1790, 3695–3703**
  - aggressive behavior and, 2487, 2488–2489
    - treatment of, 2499
  - anxiety disorders and, 1721
  - BD and, treatment of, 1699
  - BDD and, 1791
  - body-focused behaviors, 2349–2354
  - child/adolescent, 3466–3470, 3467–3470*f*
    - behavioral, 3697–3698

- clinical features, 3698–3699
- cognitive, 3697–3698
- comorbidity, 3696, 3696*t*
- comparative nosology, 3696
- course, 3700
- defined, 3695, 3695*t*
- diagnosis, 3698–3699
- differential diagnoses, 3699–3700
- epidemiology, 3696
- etiology, 3696–3698
- genetic, 3697–3698
- history, 3695–3696
- prognosis, 3699–3700
- treatment, 3700–3702, 3747, 3758–3759

- in children after TBI, 653
- cholesterol levels and, 2545
- cognitive-behavioral therapy for, 3758–3759
- cognitive therapy, 2765–2766
- comparative nosology, 1785–1786
- course and prognosis, 1789
- definition, 1785
- depersonalization/derealization disorder and, 1941–1942
- description of, 4598
- diagnosis and clinical features, 1787–1788, 1787*t*
- differential diagnoses, 1586, 1788–1789, 3710
- DSM-5 on, 1723
- EEGs in, 249–250
- epidemiology, 1786
- etiology, 1786–1787, 3638–3639, 3652–3653
- excoriation and, 1796, 2351–2352
- exposure response prevention, 2691
  - hypnosis and, 2722–2723
- gambling and, 1805
- genetics and, 3640, 3642, 3654, 3656
  - empirical risks for, 2535*t*
- history, 1785
- hoarding disorder and, 1793

- hypnosis and, 2721
- illness anxiety disorder and, 1838
- introduction, 1785
- ketamine for, 1747
- neuroimaging, pediatric, 3466–3470, 3467–3470*f*
- in older adults, 3966
- pathology and laboratory examination, 1788
- pediatric
  - aggressive behaviors in, 2499
  - psychotherapy, neuroanatomy and, 2814*t*
- psychodynamic psychotherapy, 2680–2681
  - case study, 2681
- psychopharmacology, 3792
- psychotherapy, neuroanatomy and, 2814*t*
- PTSD and, 1914, 3705
- related disorders and (OCRDs), **1785–1798**
  - body dysmorphic disorder, 1790–1792
  - conclusion, 1797–1798
  - due to another medical disorder, 1797
  - excoriation (skin-picking) disorder, 1795–1797
  - hoarding disorder, 1792–1794
  - OCD, 1785–1790, 1787*t*
  - other specified and unspecified, 1797
  - substance/medication-induced, 1797
  - trichotillomania, 1794–1795
- remission, 2922
- schizophrenia and, 1422
- skin diseases and, 2333
- SSRIs and, 2545
- substance/medication-induced, **1235***t*, **1255–1256**
- symptomatology of, 1116
- TCAs for, 69
- tic disorders and, 3642, 3656
- trazodone drug studies, 3249
- treatment, 3638–3639, 3646, 3652–3653, 3660
  - β-adrenergic receptor antagonists, 2954
  - buspirone, 3010

    - combination therapy, 1790
    - DBS, 3319–3320
    - lithium, 3093
    - MAOIs, 3122
    - neurosurgery, 3314–3315, 3316, 3318–3320
    - pharmacotherapy, 1789
    - psychoeducation, 1789
    - psychotherapy, 1789–1790
    - somatic therapies, 1790
    - SSRIs, 3154, 3155
  - TTM and, 1795
- Obsessive-Compulsive Foundation, 1789
- Obsessive-compulsive personality disorder (OCPD)
  - clinical criteria, 2157
  - comorbidity, 2157
  - complications, 2157
  - differential diagnoses, 2157
  - epidemiology, 2157
  - familial patterns, 2157
  - genetics, 2157
  - impairment, 2157
  - OCD and, 1788
  - persistence and, 2163
  - predisposing factors, 2157
  - sex ratio, 2157
- Obsessive-compulsive syndrome (OCD), 691
- Obsessoid, MDD and, 1644
- Obstetrics, hypnosis and, 2724
- Obstructive sleep apnea (OSA), 75
- Obstructive sleep apnea syndrome (OSA), 4598
- Obstructive sleep hypoapnea (OSAH)
  - comorbidities, 2093
  - diagnostic criteria for, 2093
  - treatments, 2093–2095
- Occipital lobes
  - anatomy of, 8
  - association fiber tracts and, 15*f*

- Occupation as psychosocial factor, 4168
- Occupational history at patient interview, 967
- Occupational interventions
  - for developmental coordination disorder, 3543–3544
  - for tic disorders, 3644, 3658
- Occupational problems, 4598
- Occupational psychiatry, 4599
- Occupational therapy, 1930, 3543–3544, 4599
- OCD. *See* Obsessive-compulsive disorders (OCDs)
- OCF. *See* Outline for Cultural Formulation (OCF)
- Ocinaplon, 3003
- *O'Connor v. Donaldson* (1975), 4336
- Oculocephalic maneuver, 456
- ODD. *See* Oppositional defiant disorder
- Odds ratios in epidemiology, 813–814
  - attributable fraction, 814
  - confounding, 814
  - effect modification/moderation, 814
  - mediation, 814
- Odegard, Ornulv, 1538
- Oedipal conflicts, 2797
- Oedipal phase. *See* Oedipus complexes; Phallic phase
- Oedipus complexes, 2806, 4599
- Offensive spirituality, 2825
- Offer, Daniel, 1963
- Office of Juvenile Justice and Delinquency Model Program Guide, 4296
- Office of Research on Women's Health, 4245
- Office-only physicians, hospitalists *vs.,* 4355
- Ofloxacin (Floxin), 599
- Ogden, Thomas, 2654
- Oklahoma City Federal Building bombing, 2568
- Olanzapine (Zyprexa)
  - acute mania, 1686, 1686*t*, 1687, 1689, 1689*t*
    - combination/add-on therapy, 1690–1691
    - lithium or divalproex, 1689*t*
    - lithium or valproate and, 1687
    - monotherapy, 1686

- adverse effects of
  - dermatological, 2356*t*
- for aggressive behavior
  - Alzheimer disease, 2502–2503
  - conduct disorder, 2501
  - mentally retarded children, 2500
  - schizophrenia, 2501
- anxiety disorders, 1781
  - GAD, 1783
- bipolar depression, 1690, 1692*t*, 1693, 1695, 3155
  - combination therapy, 1692*t*
- bipolar disorder
  - early onset, 3692, 3706, 3791
- bipolar I disorder, 1696, 1696*t*
- for cancer patients, 2253
- CATIE, 3069–3070
- in children/adolescents, 2281, 2501
- combination pharmacotherapy and, 3324
- delirium treatment, 1188
- dementia treatment, 1201
- depression, 1680
- diabetes risk and, 2247
- dissociative identity disorder, 1928, 1929
- dosage/administration
  - in cancer patients, 2257
  - for delirium in cancer patients, 2258*t*
- EUFEST, 3070
- for excoriation disorders, 2352
- fluoxetine combined with, 3155, 3324
- in functional GI disorders, 2206
- history, 2918
- for insomnia, 2296
- mania, 1612
- metabolic risks with, 4351, 4352
- prolactin levels and, 2249
- psychiatric emergencies, 2628
- for psychosis in PD, 548

- psychotic disorders and, 1253
- receptor profile for, 67
- for schizophrenia, 4098
- schizophrenia, early onset, 3732, 3793
- side effects of, 67, 2189
  - weight gain, 2237, 2237*t*
- therapeutic indications, 3155
- tic disorders, 3645, 3646, 3659, 3660
- weight gain and, 2215

- Olanzapine/fluoxetine combined (OFC), 3693, 3707
- "Old friends" hypothesis, 192
- Older adults, sleep in, 342
- Olfaction, examination of, 456
- Olfactory aversion, 2694
- Olfactory delusions, 1123
- Olfactory fear memory, inheritance of, 396
- Olfactory hallucinations, 1413, 4599
- Olfactory reference disorder, 1786
- Olfactory reference syndrome (ORS)
  - BDD and, 1791–1792
  - description of, 4599
- Olfactory tract, cross-section, 9*f*
- *Olig1/2* gene, 51
- Oligodendrocyte precursor cells (OPCs), 504
- Oligodendrocytes
  - characteristics of, 4*t*
  - features of, 4*f*
  - major features of, 4*f*
  - satellite, 5
- Oligophrenia, description of, 4599
- Oligosymptomatic mood states, 1601
- Olivopontocerebellary atrophy (OPCA). *See* Multiple system atrophy (MSA)
- *Olmstead v. L.C.* (1999), 4337
- Omega-3 fatty acids
  - adverse effects, 2544*t*
  - aggression in BPD, 2502
  - beneficial effects, 2546, 2547

- biological treatment with, 2551
- bipolar I disorder, 1696*t*, 1697
- bipolar II disorder, 1698
- and brain aging, 3994

- Omega-6 fatty acids
  - dietary sources of, 116
- Omega-3 fatty acids-eicosapentaenoic acid (EPA)
  - dietary supplements of, 115–116
- Omega-3 polyunsaturated fatty acid supplements (PUFA), 3980, 3984
- Omkar meditation, 2557
- Omnibus Budget Reconciliation Act, 4335
- Omnibus rating scales. *See* Comprehensive rating scales
- Omnibus test, 834
- OMS. *See* Organic mental syndrome
- *On Aphasia* (Freud), 848–849
- *On the Construction, Organization, and General Arrangements of Hospitals for the Insane* (Kirkbride), 4327
- Onabotulinumtoxin A, 624
- Onanism. *See also* Masturbation
  - definition of, 4599
- Onchyophagia (nail biting), 2353
- Ondansetron (Zofran)
  - for alcohol use disorders, 1278
  - for nausea and vomiting
    - chemotherapy related, 2271
  - substance use disorders, child/adolescent, 3793
- "One gene, one enzyme" hypothesis, 149
- O'Neal, P., 1830
- Oneiroid state, 4599
- Online gaming, 3373–3374, 3768
- Online psychotherapeutic tools. *See also* Internet; Smartphone apps
  - CBT, 1776
  - habit reversal therapy and stimulant control, 1795
- Ontogeny of behavior. *See* Evolutionary biology
- Onychotillomania, 3672, 3686
- Oophorectomy, 2034
- Open hospitals, description of, 4599
- Openness to experience, 935–936

- Operant, definition of, 2872
- Operant behavior. *See* Conditioning
- Operant conditioning. *See also* Conditioning
  - posttraumatic stress disorder (PTSD), 3704
  - stimulant use and, 1283
  - stimulant-use disorder, 1283
- Operant/instrumental learning. *See* Learning theories
- Operational definitions, 4599
- Operational stress, 2567–2568
- Opiate receptors, 4599
- Opiates. *See* Opioids (opiates)
- Opioid analgesics
  - for chronic pain conditions, 2395*t*
  - dependence issues, 2394–2395
  - misuse of, 4114
  - for pain management, 2394–2395
  - tolerance associated with, 3308–3309
- Opioid antagonists
  - autism spectrum disorder and, 3579
  - for intellectual disability, 3515
  - stereotypic movement disorders, 3673, 3687
- Opioid detoxification, rapid/ultrarapid, 3141
- Opioid maintenance treatments, **3132–3138**
  - adverse effects, 3134, 3136
  - buprenorphine, 3132, 3133, 3135–3137
  - chemistry, 3133, 3133*f*, 3135, 3135*f*
  - dosing/administration, 3134–3137
  - drug interactions, 3134, 3136
  - drug studies, 3134, 3136
  - future directions, 3138
  - history, 3132, 3133, 3135
  - laboratory interferences, 3134
  - levomethadyl, 3133
  - methadone, 3132, 3133–3135
  - opioid dependence and withdrawal, 3132–3133
  - pain treatment, 3137–3138
  - pharmacological actions, 3133–3134, 3135

- Opioid peptides, 4599
- Opioid receptor agonists, 3132–3138
- Opioid receptor antagonists, **3138–3146**
  - chemistry, 3139, 3139*f*
  - dosage/administration, 3145–3146
  - drug interactions, 3145
  - drug studies, 3139–3140
  - history, 3138–3139
  - investigational drugs, 3146
  - laboratory interferences, 3145
  - organ/system effects of, 3143–3144
  - pharmacological actions, 3139
  - precautions/adverse reactions, 3144–3145
  - therapeutic indications, 3140–3144
- Opioid receptors
  - analgesia and, 328
  - δ receptor, 169
  - κ receptor, 169
  - μ receptor, 169
- Opioid-related disorders, 3132–3138
- Opioid Risk Tool (ORT), 2393
- Opioid use
  - opioid-related disorders and
    - autism spectrum disorder and, 3579
    - comorbidities, 1365, 1366*t*
    - comparative nosology, 1354–1355, 1354*t*
    - course and prognosis, 1363–1364
    - defined, 1352–1353
    - dependence, 1355, 1370
    - detoxification, 1368–1370
    - diagnosis and clinical features, 1354*t*, 1361–1363, 1362*t*
    - differential diagnoses, 1363
    - epidemiology, 1355–1356, 1356*f*
    - etiology, 1356–1361
      - genetics, 1360
      - integrative approach, 1361
      - neurobiology of, 1357–1360, 1357*f*, 1358*f*

        - psychosocial factors, 1360
    - genetics, 1360
    - history, 1353–1354
    - integrative approach, 1354*t*, 1361–1363, 1362*t*
    - intoxication, 1354, 1361, 1363
    - laboratory examination, 1363
    - morbidity, 1364–1365
    - mortality, 1364–1365
    - neonatal opioid withdrawal, 1373
    - neuroimaging, 1359–1360, 1360*f*
    - opioid use disorder, 1354*t*, 1362–1363, 1370–1373
    - overdose, 1361, 1365, 1367–1368
    - physical examination, 1363
    - prescription opioid misuse, 1355
    - psychosocial factors, 1360
    - psychotic disorders and, 1252
    - relapses, 1364, 1370
    - toxicology, 1363
    - treatment, 1365–1373, 1367*t*
        - detoxification, 1368–1370
        - pharmacotherapy, 1365–1370, 1366*t*, 1367*t*, 1371–1373
        - psychosocial, 1370–1371
        - of special populations, 1371–1373
    - withdrawal, 1354, 1361–1362, 1362*t*, 1363, 1368–1370
    - worldwide trends, 1355–1356, 1356*f*
- Opioid use disorders, 1354*t*, 1362–1363, 1370–1373
- Opioid withdrawal
    - treatment, 3132–3133
        - buprenorphine, 3132
        - methadone, 3132
        - naltrexone, 3132–3133
- Opioids (opiates)
    - addiction
        - LSD and, 2891–2892
    - anxiety disorders and, 1744
    - delirium and, 1182, 1190
    - dependence, 3132–3133

- depersonalization/derealization, 1891
- depersonalization/derealization and, 1892
- description of, 4599
- dosing instruction, 2294
- drug reward and, 373–374
- EEG effects, 253
- endogenous, 169
- history, 3132
- immunological effects of, 185*t*
- opioid maintenance treatments, **3132–3138**
- psychiatric emergency due to
  - intoxication, 2630
  - withdrawal, 2633
- routes of administration, 2293*t*
  - for pain management, 2376–2377
- sexual dysfunction and, 1973
- substance use disorders, 385
- withdrawal, 3132–3133
  - psychiatric emergency due to, 2633
  - treatment of, 2948

- Opponent process theory, 372*f*, 375
- Oppositional defiant disorder (ODD), **3605–3612**
  - ADHD and, 2947, 3228, 3591, 3595–3596
  - aggression and, 2487, 2488–2489
    - treatment, 2499
  - bipolar comorbidity, 3689, 3703
  - brief psychotherapies for, 3743–3744
  - in children after TBI, 652
  - clinical features, 3610
  - comorbid conditions, 3611
  - comparative nosology, 3606, 3607*t*, 3608*t*
  - course, 3611, 3619–3620
  - defined, 3605, 4599
  - diagnosis, 3610
  - differential diagnoses, 3610–3611, 3613, 3619, 3625, 3639, 3690, 3704, 3710, 3721–3722
  - epidemiology, 3606
  - etiology, 3606–3609

- family therapy and, 3779–3781
- FASD and, 638
- history, 3605–3606
- with intellectual disability, 3511
- laboratory examination, 3610
- pathology, 3610
- prognosis, 3611
- psychopharmacology, 3789–3790
- rating scales for, 3401, 3402
- temperament and, 3486
- treatment, 3228, 3611–3612

- Opsoclonus-myoclonus ataxia syndrome, 2270
- Opsonization, process of, 183
- Optic ataxia, 452, 463
- Optic chiasm, 10*f*
- Optic radiations, 11
- Optimism, mental health and, 757–758
- Optogenetics, 145
- Oral appliances, 2094–2095
- Oral contraceptives, 2508
  - in PCOS treatment, 2249
  - prolactin levels and, 2249
- Oral dyskinesia. *See* Tardive dyskinesia
- Oral health, schizophrenia and, 1542
- Oral phase, description of, 4599
- Oral sex, frequency of, 1962
- Orally disintegrating tablets (ODT), 1688
- Orap. *See* Pimozide
- Orbitofrontal cortex (OFC). *See also* Prefrontal cortex
  - damage to, 450
  - fear modulation and, 1755, 1756
  - impulse activity and, 2814
  - and intuition, 687
  - mood disorders and, 1718*f*
  - OCD and, 1786
  - sexual functioning and, 1956–1957
- Orchidectomy, sex offender, 2503

- Ordered subset expectation maximization (OSEM), 290*f*
- Oregon Brain Aging Study, 3993
- Oregon Hospital for the Insane, Hawthorne's, 4403
- Orexin A, 36
- Orexin B, 36
- Orexin neurons, 37*f*
- Orexinergic neurons, 677
- Orexins
  - responses to stress, 2318
  - synthesis of, 676–677
- Organ donation
  - legal issues, 2305–2306
  - living donors, 2362
    - assessments, 2362*t*
    - outcomes, 2369–2370
- Organ Donation Breakthrough Collaborative, 2357–2358
- Organ Procurement and Transplantation Network (OPTN), 2357
- Organ transplantation, **2357–2373**. *See also* Heart transplants; Kidney transplantation; Liver transplantation
  - behavioral contracting, 2363–2364
  - clinical evaluation for, 2358–2363
  - contingency contracting, 2363–2364
  - course/prognosis, 2358–2365
  - definitions, 2357
  - diseases resulting in, 2359*t*
  - emergency evaluations, 2362
  - epidemiology, 2357–2358
  - living donors, 2362
    - outcomes, 2369–2370
  - mood disorders and, 2365–2366
  - nosology, 2357
  - patient evaluation, 2364
  - posttransplant period, 2365
    - adherence issues, 2370
    - drug interactions in, 2371
  - preoperative period, 2364–2365
  - preparation for, 2363
  - pretransplant psychosocial evaluation, 2360*t*

- psychosocial criteria, 2360–2361
  - screening and context, 2361
- special issues in, 2369–2373
- survival by organ type, 2358*f*
- timelines, 2359*t*
- treatment issues
  - psychopharmacologic, 2370–2371
  - psychotherapeutic, 2370
- wait list for
  - reassessment of candidates, 2364
- waiting periods, 2364
- waiting list for, 2358*f*

- Organic, definition of, 4599
- Organic anxiety syndrome, 4599
- Organic brain syndrome (OBS), 4599
- Organic delusional disorder, 4599
- Organic delusional syndrome, 1252
- Organic disorders
  - description of, 4599
  - schizophrenia, early onset and, 3731
- Organic hallucinosis
  - description of, 4599
  - psychotic disorders and, 1252
- Organic mental disorder (OMD), 4599
- Organic mental disorders
  - incidence with intellectual disability, 3512
- Organic mental syndrome (OMS), 4599
- Organic mood syndrome, 4599
- Organic personality syndrome, 4599
- Organic psychosis. *See* Organic mental disorder (OMD)
- Organizational climate, 4381
- Organized crime, 2413–2414
- Organophosphate-induced delayed polyneuropathy (OPIDP), 631
- Organophosphates (OPs)
  - toxic neuropathy and, 631
  - war or terrorism and, 2572*t*
- Orgasm

- description of, 4599
- disorders of, 1968, 4599
- female, 1954–1955
- female premature
  - sexual dysfunction and, 1975
- Orgasmic anhedonia, 1968
  - sexual dysfunction and, 1975
- Orgasmic phase, 4599
- Orgasmic platform, 4599
- Orientation
  - definition of, 4599
  - in patient interview, 967
- Original thirteen. *See* American Psychiatric Association (APA)
- ORLAAM. *See* Levomethadyl
- Orlistat (Xenical, Alli)
  - for binge eating disorder, 2081
  - mechanism of action, 2220
  - for weight loss, 2220
- Orne, Martin, 2732, 2733
- Orofacial pain
  - conditions, 2390*t*
  - presentation of, 2392
- Oropharyngeal cancers, 2267
- Orthomolecular psychiatry, 4599
- Orthomolecular treatment, 4600
- Orthopsychiatry, description of, 4600
- Orthostatic hypotension, 4600
- Orton, Samuel, 3523
- OSA. *See* Obstructive sleep apnea (OSA)
- Oscillations
  - amplitude synchrony, 408
  - analysis of, 408–409
  - cortical circuit, 436
  - phase-amplitude coupling (PAC), 408
  - phase synchrony, 408, 408*f*
  - thalamic cells, 745–747, 746*f*
- Oseretsky test, 459–460

- Osmond, Humphrey, 2891
- Osteoarthritis, psychiatric problems, 4049
- Osteopathic training, 4411
- Osteopenia, anorexia nervosa and, 2075
- Osteoporosis
  - anorexia nervosa and, 2075
  - and electroconvulsive therapy, 4158
  - and HIV infection, 4271
  - SSRIs and, 1779
- Ostow, Mortimer, 3944
- OT. *See* Oxytocin (OT)
- Othello syndrome, 4599
- Other-directed persons, 4600
- Other interpersonal problems, 4600
- Other specified family circumstance, 4600
- Others, self *versus,* 2895
- Otitis media, 3546–3547
- Otogenetic, description of, 4599
- Ototoxicity, chemotherapy drugs, 2272
- Outcome variables. *See* Dependent variables
- Outer subventricular zone (OSVZ), 43
- Outline for Cultural Formulation (OCF), 2459
  - clinician-individual relationships and, 2464
  - levels of functioning and, 2464
  - psychosocial environment and, 2464
- *Outline of Psychoanalysis* (Freud), 868
- Outpatient care
  - assisted/mandatory, 4396
  - hospitalist role in, 4357, 4358
  - shortcomings of, 4345
- Outpatient psychiatry, family roles, 4195
- Outpatient studies, 1888
- Outpatients, description of, 4600
- Outreach services, 4399
- Ovarian cancer, 2269
- Ovarian cycling, appetite and, 367, 367*f*
- Ovarian dysgenesis. *See* Turner syndrome

- Over-fitting techniques, 841
- Overactivity, definition of, 4600
- Overanxious disorder, 4600
- *Overcoming Procrastination* (Ellis), 4554
- Overcompensation, 4600
- Overcrowding in jails and prisons, 4475
- Overdetermination
  - definition, 2660
  - description of, 4600
- Overdiagnosis, cultural differences and, 4382
- Overdoses
  - barbiturate-related, 1384
  - benzodiazepines, 1383
  - opioid-induced, 1361, 1365, 1367–1368
  - sedative-, hypnotic-, anxiolytic-induced, 1380–1383, 1389
- Overshoot, channel, 133
- Overt homosexuality, 4600
- Overvalued ideas
  - description of, 4600
  - as psychiatric symptomatology, 1122
- Overweight, prevalence of, 2211*t*
- Overwork, physician mental health and, 2599
- Ovulation
  - induction of, 2508, 2510*t*
  - menstrual physiology and, 2507
  - suppression of, 2508
- Oxacarbazepine, 2396
- Oxaliplatin, 2264
- Oxandrolone (Oxandrin)
  - aggression and, 2493
  - side effects
    - NMS, 2381
- Oxazepam (Serax)
  - benzodiazepine dose equivalent, 2633
  - for elderly patients with cancer, 2283
  - metabolism, 2993
- Oxcarbazepine (Trileptal)

- acute mania, 1684, 1689*t*
- for bipolar disorder, childhood onset, 3791
- bipolar I disorder, 1696*t*
- carbamazepine and, 3022, 3023*f*, 3023*t*, 3024, 3036, 3037–3038
- combination pharmacotherapy, 3326, 3328
- lithium and, 3098, 3326
- and physical health monitoring, 1024

- Oxidative stress
  - Alzheimer disease and, 1205
  - and clinical disorders of aging brain, 3988
  - positive psychiatry, 4284
- Oxomorphone (Opana), 2293
- Oxotremorine, 73
- Oxybutynin chloride, 2102
- Oxycodone (OxyContin)
  - duration of action, 2395*t*
  - insomnia and, 2295–2296
  - oral bioavailability, 2395
  - for pain management, 2377
  - pain management using, 2293
- Oxygen atoms, ^{15}O, 287
- Oxymetholone (Anadrol)
  - aggressive behavior and, 2493
- Oxymorphone (Numorphan)
  - administration of, 2293*t*
  - duration of action, 2395*t*
- Oxytocin (OT), **175–176**
  - aggression and, 2481–2482
  - anxiety disorders and, 1748–1749
  - effects of, **93–95**
  - functions of, 175–176
  - milk ejection reflex and, 1957
  - orgasm and, 1956
  - PTSD prevention, 1825
  - release of, 678
  - responses to stress, 2318
  - reward dependence and, 2132

- Oxytocin receptor *(OXTR)* gene
  - PTSD and, 1823
  - social phobia, 1740
- Oxytocin receptors, 94*f*

P

- p75 neurotrophin receptor ($p75^{NTR}$)
  - neurotrophin interactions with, 97
- *p* values, 832–833
- PACE (Personal Assessment and Crisis Evaluation) Clinic, 1551
- Paced Auditory Serial Addition Test (PASAT), 990
- Pacemakers
  - neuronal firing and, 143
  - for syncope treatment, 2194
- *PAFAH!B1* gene, 52
- Pagoclone, 3003
- Pain
  - abdominal, functional, 2729
  - acceptance of, 2397
  - acute
    - causes of, 2388
    - patients with, 2292
  - affective regulation, 328
  - assessment of, 2292
  - chronic
    - benzodiazepine, 2998
    - conditions related to, 2390*t*
    - hypnotherapy, 2726–2729, 2727*f*, 2728*f*
    - management of, **2387–2398**
    - panic disorder and, 1783
    - patient follow-up, 2393
    - patients with, 2292
    - and psychiatric symptoms, 4048–4049
    - purpose of, 2388
    - QEEG analysis, 1037–1038
  - clinical features, 2389–2390
  - course/prognosis in, 2392–2394
  - diagnosis of, 2292–2293, 2389–2390
  - differential diagnoses, 2390
  - dysregulation, 328
  - epidemiology of, 2388
  - etiology of, 2388–2389

- expectable, 2294
- experience of, 2387–2388
- hypnosis and, 2713, 2718, 2724–2729
  - case study, 2725–2726, 2728–2729
- imaging studies, 2389–2390
- in long-term care, 4226
- management of
  - in cancer patients, 2253, 2274–2275
  - certificate examinations, 4405*t*, 4427
  - children with cancer, 2281
  - chronic pain, **2387–2398,** 3137–3138
  - insomnia and, 2295–2296
  - medications for, 2293–2294
  - multiple convergent, 2726–2729, 2727*f*
  - opioid agonist, 3137–3138
  - pharmacology of, 2376–2377
  - psychotherapeutic strategies for, 2377
  - terminally ill children, 2310
- modulatory systems, 326*f*
- multiple, 2292
- neuropathic, with spinal cord injuries, 2977
- neuropathic conditions causing, 2388, 2388*t*
- nociceptive conditions causing, 2388, 2388*t*
- pathology of, 2389–2390
- pathophysiology of, 2388–2389
- perception of, 115
  - animal models, 326–327
- processing of, 327*f*
- prognostication in, 2292–2293
- psychiatric interventions, 2294–2295
- psychiatric symptoms of, 2294
- somatic, 2388
- treatment of, 2292
  - antipyretic analgesics, 2394
  - benzodiazepine, 2998
  - nefazodone, 3127
  - opioid agonists, 3137–3138

    - opioid analgesics, 2394–2395
    - pharmacologic, 2394
  - types of, 2292*t*, 2376
  - types of experiences, 2388*t*
  - undertreatment of, 2294
  - visceral, 2388
- Pain Assessment in Advanced Dementia (PAINAD), 4049
- Pain disorders, 1969–1972. *See also* Somatic symptom and related disorders
  - acute, hypnosis in, 2724
  - classification, 1828, 1828*f*, 1836
  - description of, 4600
  - somatic symptoms with, 1840–1841
  - treatment, 1841
- Pain-pleasure principle, 4600
- Pain syndromes
  - trazodone drug studies, 3250
  - treatment
    - anticonvulsants, 2967–2968, 2970, 2972, 2977
    - lamotrigine, 3084, 3085
- Paired associative stimulation, 3299*t*
- Palazzoli, Maria Selvini, 2753
- Paleocortex, 16
- Paleostriatum. *See* Globus pallidus
- Pali Canon, 2039
- Palilalia, 457*f*
- Palinacousis, 465
- Palinopsia, 465
- Paliperidone (Invega)
  - acute mania, 1686*t*, 1689*t*
    - monotherapy, 1686
  - adverse effects of, 2356*t*
  - for aggressive behavior
    - in schizophrenia, 2501
  - bipolar depression, 1695
  - bipolar I disorder, 1696, 1696*t*, 1697
  - side effects
    - weight gain, 2237*t*

- Palliative care, **2288–2315**
  - certificate examinations, 4427
  - definition of, 2289
  - dignity therapy, 2825–2827
  - family role, 4198
  - in long-term care, 4231
  - patients with acute pain, 2292
  - patients with chronic pain, 2292
  - quality of life and, 2295*t*
- Pallidothalamocortical pathway, 31
- Pallidotomy, 547
- Pallidum. *See* Globus pallidus
- Pamelor. *See* Nortriptyline
- PANAS. *See* Positive and Negative Affect Scale
- Pancreatic cancer
  - depression and, 2260
  - mood disorder and, 1249
- Pancreatic function, evaluation of, 1014–1015
- Pancreatic islets, 2246
- PANDAS, 598, 1797
  - neuroimaging, 3470
  - tic disorders, 3638, 3652
- Pandemic influenzas, 2574
- *Panetti v. Quarterman,* 4438
- Panic
  - definition of, 4600
  - GABA–glutamate interaction and, 1747
- Panic Attack Specifier
  - agoraphobia, 1724
  - panic attacks, 1724
- Panic attacks
  - clinical evaluation, 454–455
  - criteria, 1723*t*, 1724
  - description, 1768
  - gambling disorder and, 1808
  - schizophrenia, 1421
  - VRGET for, 2560

- Panic disorder(s)
  - after disasters, terrorism, or war, 2567*t*
  - ALLO levels in women and, 1748
  - asthma and, 2225
  - BD and, treatment of, 1698
  - benzodiazepines and, 1747
  - cardiac symptoms of, 2195
  - CBT and, 1774–1775
  - cholecystokinin and, 1749–1750
  - cholesterol levels and, 2545
  - classification of, 1173
  - cognitive therapy, 2765, 2862
  - d-cycloserine for, 1747
  - definition, diagnosis, and clinical features, 1723–1724
  - description of, 1768, 4600
  - diabetic patients, 2238
  - differential diagnoses, 1275, 1724
  - dopamine and, 1745
  - DSM-5 criteria for, 1721, 1723*t*
  - EEGs in, 249
  - emotional processing and, 1770
  - epidemiology, 1724
    - adults, 1731, 1732*t*
  - epigenetics, 1740
  - gambling and, 1804
  - genetics and, 2529*t*
    - empirical risks for, 2535*t*
  - genome-wide association studies, 1739
  - history and comparative nosology, 1724
  - HPA axis abnormalities and, 1744
  - hypnosis, 2722
  - illness anxiety disorder and, 1838
  - management, 1782*t*
  - neurocircuitry, 1756
  - neuroimaging, 1760–1762
    - functional findings, 1761
    - structure and chemistry, 1761–1762

- noradrenergic system dysfunction and, 1742
- in older adults, 3966
- perinatal, 2515
- pharmacotherapy, 1782, 1783
- phenomenology of, 1751
- presentation of, 2258
- psychoanalytic psychotherapy for, 2804–2805
- psychodynamic psychotherapy, 2678–2679
- psychotherapy, neuroanatomy and, 2814*t*
- rTMS for, 1782
- schizophrenia, 1421
- STDP and, 2810–2811
  - clomipramine and, 2811–2812
- syncope and, 2194
- therapy, hypnosis and, 2722
- trazodone drug studies, 3249
- treatment
  - β-adrenergic receptor antagonists, 2954
  - benzodiazepine, 2995, 2996
  - buspirone, 3010
  - MAOIs, 3122–3123
  - nefazodone, 3127
  - SSRIs, 3154
  - valproate, 3275
- tryptophan depletion and, 1746
- women, galanin and, 1749

- Panic-focused psychoanalytic psychotherapy (PFPP), 2678–2679
- Panic-like behaviors, models of, 320
- *PANK2* gene mutations, 673
- Panphobia, 4600
- PANS (pediatric acute onset neuropsychiatric symptoms), 598
- Pantomime, description of, 4600
- Pantothenate (vitamin B_5), 673
- Pantothenate kinase-associated neurodegeneration (PKAN), 672
  - atypical clinical phenotype, 673
  - classical syndrome, 673
  - MRI imaging of, 673

- presentation of, 673
- Pap tests, 2269
- PAPA. *See* Preschool Age Psychiatric Assessment
- Papez, James
  - on emotional behavior, 31
  - on the limbic system, 4591
- Papez circuit
  - connections, 450
  - for emotion, 33*f*
- Papilledema
  - ICP and, 456
  - psychiatric patients
    - brain tumors and, 494
- Paracelsus, 4533
  - on dissociative disorders, 1866
- Paracetamol, 625–626
- Paraclinical investigations, 466
- Paradigm shifts, rehabilitation models, 4301
- Paradoxical insomnia, 2085
- Paradoxical intention, 934
- Paradoxical intention techniques, 2090
- Paradoxical reward effect, 724
- Paradoxical sleep. *See* REM sleep
- Paradoxical therapy, **2816–2818**
  - clinical and ethical issues, 2817–2818
  - history, 2816
  - techniques, 2817
  - theoretical issues, 2816–2817
- Parahippocampal gyrus
  - anatomy of, 31, 34*f*
  - cytoarchitectonic regions, 33
- Paral. *See* Paraldehyde
- Paraldehyde (Paral), 2984
- Parallel distributed processing models, 432. *See also* Connectionist models
- Parallel form reliability, 997
- Paralogia. *See* Evasion
- Paralysis agitans, 542. *See also* Parkinsonism

- Paralytic ileus, 4600
- Parameters, definition of, 4600
- Parametric exponential family, 842
- Parametric forms, 830
- Parametric tests of significance, 4600
- Paramnesia
  - after TBI, 538
  - description of, 4600
- Paraneoplastic encephalomyelitis, 628
- Paraneoplastic syndromes
  - endocrine, 2270
  - metabolic, 2270
  - psychological issues in, 2270
- Paranoia
  - description of, 4600
  - (Pa) Scale, MMPI, 999*t*
  - violence and, 2428
- Paranoiac, definition of, 4600
- Paranoid delusional thoughts
  - in HIV infection, 574
  - parasitosis and, 2347
- Paranoid delusions, 4573, 4600
- Paranoid disorder. *See* Delusional disorders
- Paranoid ideation, 4600
- Paranoid patients, emergencies and, 2626
- Paranoid personality disorder
  - clinical criteria, 2151
  - comorbidities, 2151–2152
  - complications, 2151
  - description of, 4600
  - differential diagnoses, 2152
  - epidemiology, 2152
  - familial pattern, 2152
  - genetic, 2152
  - impairment, 2152
  - psychosis manifestation as, 1572
  - Self-Transcendence and, 2163

- sex ratios, 2152
- Paranoid schizophrenia, 4600
- Paraphasia, description of, 4600
- Paraphilia, description of, 4600
- Paraphilic disorders, **2013–2022**
  - behavior modification, 2055–2057
  - categories of, 2015
  - course/prognosis, 2019
  - definition of, 1985, 2013, 2024, 2042
  - diagnosis of, 2016–2019
  - differential diagnoses, 2019
  - epidemiology of, 2015–2016
    - crossing over, 2015–2016
    - frequency of, 2015
    - gender ratios, 2015
  - etiology of, 2016
  - future research, 2022
  - history of literature on, 2013–2014
  - nosology, 2014–2015
  - OCD *versus,* 1788–1789
  - pathology, 2019
  - treatment
    - background of, 2020–2021
    - ethical considerations, 2021–2022
    - legal considerations, 2021–2022
    - models of, 2021
- Paraphrenia, 4600–4601
- Parapraxis, 4601
- Paraprofessionals, definition of, 4601
- Parapsychology, description of, 4601
- Parasitosis, delusional, 2346–2347, 2346*f*
- Parasomnias, **2099–2105**
  - description of, 4601
  - pediatric, 3936–3937
- Parasympathetic nervous system. *See* Autonomic nervous system
- Parataxic distortion, 4601
- Parathyroid disorders, 682

- Parathyroid hormone (PTH)
  - function of, **174**
  - laboratory evaluation, 1014
- Paratonic rigidity, 458
- Paraventricular nuclei (PVN)
  - anatomy of, 36
  - appetite and, 364–365
  - inhibitory pathways, 206
  - projections from, 36
- Parenchymatous neurosyphilis, 592
- Parens patriae, 4601
- Parent-Directed Interaction (PDI), 2693
- Parent management training (PMT), 2500
  - as cognitive-behavioral therapy, 3750
  - conduct disorder and, 3620
  - for disruptive behavior disorders, 3611–3612
  - research in, 2759
- Parent of origin effects, 161
- Parental monitoring and child problem-solving curriculum, 2500
- Parental neglect, 1713
- Parent–Child Interaction Therapy (PCIT)
  - contingency management in, 2693
  - for separation anxiety treatment, 3711
- Parent–child problems, 4601
- Parenting
  - ADHD children, 3931–3932
  - anxieties, during child psychiatric evaluation, 3380–3382, 3439, 3440
  - autism spectrum disorder children, 3927–3929
  - bipolar disorder children, 3929–3931
  - capacity, forensic evaluation of, 3864
  - as child psychiatric evaluation consideration, 3378–3379
  - clinical interview with, 3376–3381, 3411
  - cognitive-behavioral therapy, 3755–3756
  - conduct disorder and, 3616, 3620
  - disruptive behavior disorders and, 3609–3612
  - dying children, 2308–2309
  - family therapy, 3777–3778

- group psychotherapy and, 3762–3763, 3764
- inadequate, depression and, 1662
- infantile anorexia and, 3623–3625, 3637–3639
- OCD, child/adolescent treatment involvement, 3700
- physical abuse by, 3834, 3841
- as preschool psychiatric evaluation informants, 3409–3410, 3411
- psychiatric disorder development of, 3928–3929
- as psychosocial factor, 4168
- response to healthy siblings, 2311
- of schizophrenia children, 3932–3933
- sensory food aversions and, 3626, 3640
- separation anxiety and, 3709
- stressors, 3925–3933
- suicide behavior and, 3682, 3696
- temperament and, 3488, 3489*f*

- Paresis, description of, 4601
- Paresthesia, 4601
- Pargyline
  - chemistry, 3113*f*
  - history, 3111
- Parietal lobes
  - anatomy of, 8
  - association fiber tracts and, 15*f*
  - tumors of, 2267
- Park, Acacia, 2853
- Parkin protein
  - mutations in, 107
  - in Parkinson disease, 107
  - ubiquination pathway and, 543
- Parkinson, James, 1219
- Parkinson disease (PD). *See also* Parkinsonism
  - aggression and, 2495
  - in aging brain, 3993
  - antidementia drugs for, 4149
  - apathy and, 3229
  - characteristics of, 29
    - cognitive manifestations, 544

    - motor manifestations, 544
    - psychiatric manifestations, 544
  - cognitive manifestations
    - treatment of, 547–548
  - course/prognosis, 547
  - cranial neuropathies and, 630
  - definitions, 542
  - delirium in, 1190
  - dementia in, 1219–1220
  - depression and, 3237–3238
  - DHEA levels in, 118
  - diagnosis
    - cognitive manifestations, 545
    - motor manifestations, 544–546
    - psychiatric manifestations, 545–546
  - differential diagnoses, 546–547
  - diurnal sleepiness, 75
  - electroconvulsive therapy, 4153
  - epidemiology, 542
  - etiology, 542–544
  - gambling and, 1805, 1807
  - genetic factors, 4020–4021
  - histamines and, 2981
  - historical literature on, 542
  - laboratory examinations, 546
  - microglial cells and, 54
  - mood disorder and, 1249
  - neuropathology of, 542, 543*f*
  - novelty seeking and, 2132
  - pain of, 2392
  - parasomnia and, 2102
  - pathology, 546
  - PDI nitrosylation in, 107–108
  - psychiatric symptoms in, 4047–4048
  - SRIF and, 175
  - striatal dopamine transmission and, 62
  - substantia nigra and, 28

- treatment of, 547–548
  - amantadine, 2962–2963
  - antihistamines, 2982
  - ECT, 3287–3288
  - FGAs and, 3072
  - pharmacological, 547
  - surgical, 547
- Parkinsonian disorders, 540
- Parkinsonism
  - amantadine and, 2962–2963
  - anticholinergics and, 2959
  - description of, 4601
  - drug-induced, FGAs and, 2938, 3072–3073
  - neuroleptic-induced, 2938, 2939
  - pharmacological management-induced, 1526
  - treatment, 2997
- Parlodel. *See* Bromocriptine
- Parnate. *See* Tranylcypromine
- Paroxetine (Paxil)
  - aggressive behavior and, 2493
  - ALLO levels and PD and, 1748
  - anxiety disorders, 1779–1780
    - GAD, 1783
    - panic disorder, 1743*t*, 1746
  - BD and generalized anxiety disorder, 1698
  - bipolar depression, 1691
  - breast milk excretion of, 3168
  - CBT for panic disorder and, 1775
  - chemistry, 3160
  - for chronic idiopathic pruritus, 2348–2349
  - depression treatment, 4124
  - dissociative identity disorder, 1928
  - dosing/administration, 573*t*, 3165–3166
    - for sexual dysfunction, 1979
  - drug interactions, 3164–3165
  - ITP for depression *versus,* 2813
  - pharmacodynamics, 3154

- pharmacokinetics, 3164–3165, 3164*t*
- pharmacological actions, 3160–3161
- postpartum depression, 2515–2516
- pregnancy and, 3167
- PTSD, 1823, 1928, 3717
- sexual dysfunction and, 1779*t*
- side effects, 573*t*, 1679*t*
- social anxiety disorder, 3717
- specific phobias, 1784
- targets for, 69
- warnings on, 3701
- withdrawal, 2923

- Paroxetine CR, 3160–3161, 3166
- Paroxysmal extreme pain disorder (PEPD), 144
- Paroxysmal limbic phenomena, 454
- Paroxysmal photosensitive responsive (PRRs), 241
- Parry disease, 658
- Pars compacta
  - anatomy of, 28
  - substantia nigra, 29–30
- Pars opercularis, 449
- Pars reticulata
  - anatomy of, 28
  - projections from, 30–31
- PART. *See* Primary age-related tauopathy
- Partial complex status epilepticus, 2496
- Partial hospital program (PHP)
  - child/adolescent, 3801
  - description of, 4601
  - inpatient hospitalization avoided with, 4344–4345
- Partial insanity. *See* Monomania
- Partial reinforcement extinction effect (PREE), 2133
- Participant modeling, 1774
- Parvalbumin, 17
- PASAT. *See* Paced Auditory Serial Addition Test
- *Passages* (Sheehy), 938
- *Passiflora incarnata* (passion flower)

- anxiety disorders, 1782
- biological treatment with, 2554

- Passive–aggressive personality disorder, 4601
- Passive–dependent personality. *See* Dependent personality disorder
- Passivity experiences and thoughts, 1123
- Pastoral counseling, 4601
- Patch amnesia. *See* Localized amnesia
- Patch clamps, function of, 132
- Patches, rodent, 29
- Pathognomonic, definition of, 4601
- Pathological dissociation, 1883–1884
- Pathological gambling (PG), 1799, 1801–1802. *See also* Gambling disorders
  - definition of, 4601
  - treatment
    - amantadine, 2964
    - nalmefene, 3143
    - naltrexone, 3143
- Pathological intoxication. *See* Alcohol idiosyncratic intoxication
- Pathological laughter and crying (PLAC)
  - after brain injury, 475
  - after stroke, 478
    - diagnosis of, 482
    - etiology of, 480
    - treatment of, 488–489
  - scale for, 483*f*
- Pathological mood change, 1633
- Pathological possession
  - dissociative disorders and, 1879
  - dissociative identity disorder
    - DSM-5 on, 1899
  - epidemiology, 1884–1885
- Pathophobia, definition of, 4601
- Pathway analysis
  - gene expression and, 163
  - genetic mapping and, 3453, 3454*t*
- Patient activation in disparities reduction, 4386
- Patient assessments, 4032

- Patient-centered care in health care reform, 4309
- Patient-Centered Medical Homes
  - benefits of, 4318
  - care integration within, 4295–4296, 4305
  - medical organization endorsement of, 4307
  - primary care in, 4307
  - psychiatric hospitals interface with, 4326*f*
  - psychiatry in, 4311
- Patient-Centered Outcomes Research Institute (PCORI)
  - funding for, 4319
  - in public psychiatry, 4295, 4296
  - research from, 4308
  - treatment models catalogued by, 4307
- Patient-controlled analgesia (PCA) pumps
  - for cancer pain, 2275
  - in terminally ill children, 2310
- Patient education
  - cancer support and, 2254
  - HIV infections and, 583
- Patient-family units, terminal illness and, 2303–2304
- Patient groups, relationships in, 2253–2254
- Patient Health Questionnaire (PHQ), 3962, 4320
- Patient Health Questionnaire-9 (PHQ-9), 2261
- Patient histories
  - brain injuries, 475–476
  - cognitive history, 614*t*
  - dementia affecting, 1194–1195
  - neuropsychiatric, 454
    - birth history, 454
    - early development, 454
    - head injuries, 454
    - paroxysmal disorders, 454
  - neuropsychiatric workup, 612–616
  - sex histories and, 1963–1965, 1964*t*
- Patient interviews, diagnosis and, 1
- Patient-litigants, testimonial privilege and, 2670
- Patient Protection and Affordable Care Act (PPACA), 4212, 4219–4220

- Patient Reported Outcome Measurement Information System (PROMIS), 1157
- Patient safety as shared function, 4339
- Patient satisfaction with hospitalists, 4358
- Patient–physician interactions
  - research into, 4352
- Patient–provider relationships
  - effects of, 2231
- Patients. *See also* Patient histories
  - appetitive symptoms
    - review of, 455
  - cognitive symptoms, 455
  - expectations of, 2289–2290
  - handedness of, 455
  - neurological examination, 456–460
  - neuropsychiatric approach to, **448–474**
  - preferences of, 4378
  - right of refusal of treatment, 2021
  - rights of
    - communications and, 2290–2291
  - types of, 2292
- Patterson, David R., 2724, 2729
- Paucity of speech, 4601
- Pavlov, Ivan Petrovich. *See also* Learning theories
  - on classical conditioning, 3749
  - learning theory, 411–412
  - *Lectures on Conditioned Reflexes,* 4539
  - work of, 4601
- Pavlovian conditioning. *See* Conditioning
- Pavor nocturnus. *See* Sleep terrors
- *PAX6* deletion, 49
- *Pax6* gene expression, 49
- Paxil. *See* Paroxetine
- Paxipam. *See* Halazepam
- Pay for Performance (P4P)
  - for high-value care, 4316, *4317f*
  - physician reimbursement linked to, 4319–4320
- Payoff, and behavior, 722–724, 723*f*

- PBA. *See* Pseudobulbar affect (PBA)
- PCP. *See* Phencyclidine
- PD. *See* Parkinson disease (PD)
- PDD. *See* Pervasive development disorder
- PDEs. *See* Phosphodiesters (PDEs)
- Peabody, Francis W., 1843
- Peace and people, documentation of, 2416–2417
- Peaceful societies, documentation of, 2416–2417
- Peccatophobia, 4601
- Pecking orders, 4601
- Pederasty
  - in classical Greece, 1996
  - definition of, 4601
- Pediatric acute onset neuropsychiatric symptoms. *See* PANS
- Pediatric Anxiety Rating Scale (PARS), 3400
- Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections. *See* PANDAS
- Pediatric evaluations, 3406–3407, 3407*t*
- Pediatric OCD treatment study (POTS), 3792
- Pediatric Research Equity Act (PREA), 2931, 3787
- Pedigree analysis, gene mapping, 307–308
- Pedophebephilic disorders, 2015
- Pedophilia
  - definition of, 4601
  - neuroimaging of, 4487
- Pedophilic disorders
  - characteristics of, 2017–2018
  - definition of, 2013
- Pedophobia, 4601
- Pedunculopointine tegmental (PPT) nuclei, 342
- Peer groups, aggressive behavior and, 2485
- Peer relationships
  - adolescent evaluation consideration, 3387, 3387*t*
  - adolescent influence by, 3369, 3370
  - as child psychiatric evaluation consideration, 3378
  - conduct disorder and, 3616
  - temperament and influence of, 3483, 3488
- Peer review, definition of, 4601

- Peer Review Organization (PRO), 4601
- Peer support specialists, 1557–1558
- Peers, aggression and, 2485
- Pekala, Ronald, 2709
- Pelizaeus–Merzbacher disease (PMD)
  - description of, 666
  - MRI imaging, 666*f*
- Pellagra, 4601
- Peller, Lilli, 3944
- Pelvic inflammatory disease (PID), 2514
- Pelvic pain, conditions, 2390*t*, 2514
- Pemoline (Cylert)
  - black box warning, 2919
  - for fatigue
    - in cancer patients, 2264
    - in HIV patients, 2264
- Penfluridol (Semap), 3064
- d-Penicillamine (Cuprimine), 671
- Penile inversion, 2034
- Penile ring, 1960*f*
- Penis
  - male prostheses, 1981
  - sexual response cycle, 1960*t*
- Penis envy, 4601
- Penn Emotion Recognition Test (ER-40), 4370
- Penn Resiliency Program (PRP), 2853
- *Pennhurst State School & Hospital, Halderman v.*, 4336
- Pentazocine, 2395*t*
- Pentobarbital (Nembutal)
  - sedative-hypnotic withdrawal and, 2987
  - therapeutic indications, 2986, 2987
- Pentolamine, 1978, 1980
- Pentothal. *See* Thiopental
- Peptic ulcer disease, 915–916
- Peptidases
  - activation of, 89–90
  - classes of, 89–90

- Peptide hormones, 2481–2482
- Peptide tyrosine kinase, 361
- Peptides, immune cell receptors, 184*t*
- *Per* gene, 206
- Per2 protein, 206
- Perceptions
  - aging adult, 3974
  - clinician's, 1114
  - cognitive model of, 2760, 2760*f*
  - definition of, 4601
  - disturbances of, 1125–1128
    - body image distortion as, 1127–1128
    - hallucinations as, 1125–1127
    - illusions as, 1125
  - at patient interview, 954, 954*t*, 966
  - psychosis and, 2879–2881
  - and sensation, 690
- Perceptual disturbance, 2567–2568
- Percocet, 625
- Percodan, 625
- Percutaneous endoscopic gastrostomy (PEG), 2274
- Percutaneous endoscopic jejunostomy (PEJ), 2274
- Peregrination, 1847
- Perfectionism, anorexia nervosa and, 2070
- Perforant path, 35
- Performance anxiety
  - β-blockers and, 1781
  - management, 1783–1784
  - propranolol for, 1742, 1743*t*, 1784
  - treatment
    - β-adrenergic receptor antagonists, 2954
  - yoga and, 1782
- Performance-Based Skills Assessment (UPSA), 4371–4372
- Performance batteries, 4028–4029*t*
- Performance enhancement, 3238
- Performance in practice (PIP), 4428
- Pergolide (Permax)

- history, 3234
- for hyperprolactinemia, 2250
- organ/system effects of, 3238–3239

- Peri-ictal period
  - aggression in, 2496
  - definition of, 507
- Periactin. *See* Cyproheptadine
- Periaqueductal gray (PAG)
  - aggression mediated by, 2476
  - pain inhibitory pathways, 326
- Perimenopausal depression, 3333–3335
- Perimenopause, 2521
- Perinatal anxiety, 2515
- Perinatal depression, 2514–2515
- Perinatal factors, autism spectrum disorder and, 3575
- Perinatal psychiatric disorders, 2514–2521
  - anxiety
    - treatment of, 2515–2516
  - depression
    - treatment of, 2515–2516
  - perinatal anxiety, 2515
  - perinatal depression, 2514–2515
  - postpartum blues, 2514
- Perinatal risk factors, 2484–2485
- Period prevalence, 4601
- Periodic leg movements (PLMs), 3110
- Periodic limb movement disorder (PLMD), 2103
- Periodic limb movements in aging, 4056–4057
- Perioral tremor, FGAs and, 3074
- Peripartum depression, 4601
- Peripheral nerve injuries, 630
- Peripheral nervous system
  - barbiturate effects on, 2986
  - structure of, 627
- Peripheral neuropathies, 628–633
  - alcohol and, 1267
  - chemotherapeutic agents and, 2273*t*

- conditions, 2390*t*
- differential diagnoses, 469*t*
- in HIV infections, 586–587
- pain related to, 2390

- Peripheral neuropathy, 1267
- Peripheral vocal fold paralyses, 3567
- Perisylvian polymicrogyria, 446
- Peritraumatic dissociation, 1881
- Peritraumatic Dissociative Experiences Questionnaire (PDEQ), 1877*t*
- Periventricular hemorrhagic infarction (IPH/PVHI), 641, 641*f*
- Periventricular leukomalacia (PVL)
  - description of, 641–642
  - white matter injury in, 642*f*
- Perley, Michael J., 1830
- Perls, Fritz, 2737, 2738
- Permax. *See* Pergolide
- Permeability transition pores (PTPs), 141
- Permissive hypnotherapy, 2716
- Peroxisomal disorders, 659
- Peroxynitrite, 107
- Perpetrators
  - abuse, 3829
  - rape, 2539
  - violent crime, 2538
- Perphenazine (Trilafon)
  - acute mania, 1686
  - CATIE, 3069–3070
  - depression, 3070
  - history, 2918, 3060
  - laboratory interferences, 3078
  - side effects, 2237
- Perseveration, definition of, 459, 4601
- Persistence
  - description of, 4601
  - obsessive-compulsive personality disorder and, 2163
  - personality disorders and, 2162*t*
  - psychobiological findings, 2133*t*

- TCI scales, 2164*t*
- temperament and, 2133

- Persistent complex bereavement disorder, 4601–4602
- Persistent depressive disorder (PDD)
  - about, 1677
  - definition of, 1601, 1631, 3674, 3676, 3686, 3688, 3690, 3700
  - DSM-5 criteria for, 1644–1645, 1645*t*
  - STAPP and, 2798
  - STDP and, 2801
- Persistent genital arousal disorder, 1974
- Persistent (chronic) motor or vocal tic disorder, 4601
- Persistent pulmonary hypertension of the newborn (PPHN), 3167
- Persistent somatomorph pain disorder, 1831
- Persisting dementias, alcohol-induced, 1273
- Person-centered recovery models, 4363
- "Person vs. situation" debate, 2144
- Personal, definition of, 4602
- Personal happiness program, 2852
- Personal protective equipment, 617, 617*f*
- Personality
  - anxiety disorders and, 1735
  - assessment of
    - in clinical practice, 2142–2144
  - definition of, 2127–2128, 2145, 4602
  - development
    - self-organizing psychobiological complex, 2145–2149
  - developmental trait personality models, 2145
  - dimensions, 2130*t*
    - obesity and, 2213
  - integrated view of, 2141–2142
  - planes of being and, 2142
  - projective measures of, 1004*t*
  - psychobiological schemas, 2146*f*, 2147*f*
  - solidity, 2145
  - stability, 2145
  - subtypes
    - integrated views of, 2160

- three-dimensional spiral of, 2143
- traits
  - global, 2144
  - mental disorders and, 2144*t*
- transformation of
  - conflict and, 2174
- validity, 2145

- *Personality: A Psychological Interpretation* (Allport), 934
- Personality Assessment Inventory (PAI), 1856, 2686
  - characteristics of, 998*t*, 1001–1002
  - detection of malingering with, 2410
  - *DSM*-5-Adult (PID-5), 998*t*
- Personality assessments, **994–1007**
  - adult psychological tests
    - behavioral assessment, 1006
    - Brief Psychiatric Rating Scale, 1006–1007
    - Minnesota Multiphasic Personality Inventory, 999*t*
    - objective personality tests, 997–1000, 998*f*
    - projective personality test, 1003t
    - rapid assessment instruments, 1002
    - rating scales, 1006–1007
    - Rorschach test, 1003–1004, 1003*f*, 1004*t*
    - sentence completion test, 1006
    - specific referral questions, 1007
    - Thematic Apperception Test, 1004–1006, 1005*f*, 1005*t*
  - psychological testing
    - differential diagnoses, 994
    - narrow-band assessment, 995
    - in psychotherapy, 994–995
  - psychometric properties of instruments
    - internal consistency, 996
    - normative sample, 995
    - parallel form, 996
    - reliability, 996–997
    - standard error of measurement, 997
    - test characteristics, 995
    - test-retest, 996

    - validity of, 995–996
- Personality changes (PC)
  - after TBI, 529–530
  - in brain tumor patients, 502–504
  - with dementia, 1196
  - due to another medical condition, **1259–1261**
    - comparative nosology, 1259
    - course and prognosis, 1260
    - defined, 1259
    - diagnosis and clinical features, 1235*t*, 1260
    - differential diagnoses, 1260
    - epidemiology, 1259–1260
    - etiology, 1260
    - history, 1259
    - laboratory examination, 1260
    - treatment, 1260–1261
  - in HIV patients, 575–578
  - post TBI in children
    - behavioral treatments, 653
    - pharmacological therapy, 653
  - review of, 455
  - TBI-related, 652–653
- Personality development
  - genetics and, 2167
  - over life span, 3997
  - stage theories, 3996–3997
- Personality Diagnostic Questionnaire (PDQ-R), 2161
- Personality Disorder Examination, 2161
- Personality disorders (PDs), **2126–2176, 4103–4108**. *See also specific disorders*
  - addictive behaviors and, 579
  - alloplastic nature of, 2158–2159
  - behavior oddities, 2173
  - borderline core of, 2165
  - in cancer patients, 2265–2266
  - categorical approach to, 2159
  - classification of, 2158
    - DSM-5, 2150–2151

    - ICD-11, 2150
  - clinical features and course, 4104–4106
  - clinical syndromes and, 2161*t*
  - Cluster A disorders, 2151–2153, 2161, 4105
    - dissociative identity disorder and, 1906
    - Self-Directness and, 2163
  - Cluster B disorders, 2153–2156, 2161, 4105–4106
    - dissociative identity disorder and, 1905, 1906
    - dissociative identity disorder *versus,* 1882
    - Novelty Seeking and, 2163
  - Cluster C disorders, 2156–2157, 2161, 4106
    - Harm Avoidance and, 2163
  - cognitive-perceptual distortions in, 2173
  - cognitive therapy, 2768, 2772
  - definition of, 4103, 4602
  - diagnosis of, 575, 2163–2164, 4104
  - dimensional approach to, 2159–2160
  - dimensional assessment of, 2161–2162
  - dissociative, violence and, 2428
  - dissociativity, 1870
  - DSM-IV summary, 2158
  - DSM-5, disadvantages of, 2158
  - egosyntonic nature of, 2158–2159
  - emotional detachment in, 2173
  - emotional responses to, 2170
  - epidemiology, 4103–4104
  - epilepsy patients and, 513
  - feelings of “emptiness” in, 2140
  - Five-factor Model, 939
  - gambling and, 1804
  - genetics and, 779, 2529*t*
  - group psychotherapy, 2746
  - hallucinogen-assisted psychotherapy and, 2889
  - hallucinogens and, 2887
  - HIV patients and, 577
  - hypnosis and, 2723–2724
  - immature defenses and, 2138

- individual, diagnostic guidelines, 2151
- as magnetic diseases, 1866
- measurement of
  - categorical tests for, 2160–2161
  - interviews, 2160
  - self-reporting, 2160
  - self-trait effect, 2160
- medication adherence in, 577–578
- methodological issues, 2165–2166
- mood disorders and, 1655, 1655*t*, 1669–1670
- normal *versus* deviant personality, 2149
- OCD and, 1788
- other specified, 2158
- pathogenesis of, 2164–2167
- pathology of, 2169
- personality traits and, 939
- physician, 2601
- psychobiological summary of, 2166–2167
- psychodynamic model of, 2164–2166
- secondary symptoms of, 2169
- self and, 332
- social detachment in, 2173
- social or clinical diagnosis, 2159
- spectrum disorder hypothesis, 2164
- STDP and, 2811
- subtypes, 2150–2151, 2151*t*
  - domains of, 2171
  - temperament traits and, 2163*f*
- suicide and, 2612, 2613–2614
- symptoms, drugs targeting, 2172*f*
- temperament types and, 2166, 2166*t*
- treatment of, 2167–2175
  - barriers to, 2169–2170
  - lithium, 3093
  - pharmacotherapy, 4108
  - preventable errors in, 2169–2170
  - psychotherapy, 4107–4108

    - stabilization phase, 2170–2175
  - unspecified, 2158
- Personality Inventory for Children (PIC), 3399
- Personality measures, child assessments, 3404
- Personality states. *See* Self-states
- Personality theories
  - Adler, Alfred, 910
  - Alexander, Franz, 915
  - Allport, Gordon, 931–932
  - Berne, Eric, 925–926
  - Frankl, Viktor, 932–933
  - Fromm, Erich, 920–921
  - Horney, Karen, 920
  - Jung, Carl Gustav, 912–914, 913*f*
  - Kelly, George, 932
  - Maslow, Abraham, 932
  - McAdams, Dan, 933
  - Meyer, Adolf, 909
  - radical behaviorism, 929–930
  - Rank, Otto, 918
  - Reich, Wilhelm, 917
  - Rogers, Carl, 932
  - Sullivan, Harry Stack, 922–923
  - trait and factor models, 934–938
- Personality traits
  - borderline personality disorder as, 1572–1573
  - cancer outcomes and, 2286
  - classification of, 1159, 1174, 1174*t*
  - description of, 4602
  - dimensions of, 3480–3481, 3481*t*
  - paranoid personality disorder, 1572
  - in psychiatric illnesses, 1144–1145
  - schizoid personality disorder, 1572
  - schizotypal personality disorder, 1571–1572
  - as symptomatology consideration, 1116
- Personalizing, as cognitive distortion, 2696
- Persuasion, definition of, 4602

- *Persuasion and Healing* (Frank), 4547
- Pervasive development disorder (PDD)
  - aggression and, 2487
  - autism spectrum disorder and, 3571, 3573
  - course and prognosis, 3793
  - depressive disorders and, 3677, 3691
  - description of, 4602
  - elimination disorders and, 3660, 3674
  - parental care and, 3928
  - social (pragmatic) communication disorder, 3551–3552, 3554
  - stereotypic movement disorder, 3670, 3684
  - treatment, 3071
- Pervasive developmental disorder not otherwise specified, 4602
- Perversion, definition of, 4602
- Perversion of mothering, 1851
- Perverted logic. *See* Evasion
- PET. *See* Positron emission tomography
- Petit mal epilepsy, 4602
- Petrarch, 4533
- Pets, spirituality and, 2585
- PFLAG (Parents, Families and Friends of Lesbians and Gays), 2004
- PG. *See* Pathological gambling
- 3-PGDH. *See* 3-phosphoglycerate dehydrogenase (3-PGDH)
- Phallic overbearance, 4602
- Phallic phase, 4602
- Phallometric testing, 2019
- Phantasy. *See* Fantasies
- Phantom limbs, 4602
- Phantom pain, 2390*t*, 2391–2392
- Pharmaceutical industry
  - drug development by, **2926–2936**
  - narrative psychiatry and, 2839–2840
- Pharmaco-EEG, 251–252, 1040–1041
- Pharmacodynamics, 4121–4122. *See also specific drugs*
  - gender differences, 4249
  - psychopharmacology and, 2919
- Pharmacogenetics

- field of, 166
- mood disorders and, 1628–1629

- Pharmacogenomics
  - defined, 3448
  - field of, 166
  - in neuropsychiatric disorders, 162
  - studies on, 3453–3457
  - testing
    - for children, 3456
- Pharmacokinetics. *See also specific drugs*
  - cholinesterase inhibitors, 4143–4145
  - description of, 4602
  - distribution, 4120
  - excretion, 4121
  - gender differences, 4248–4249
  - metabolism, 4120–4121
  - psychopharmacology and, 2920
- Pharmacological actions. *See also specific drugs*
  - psychopharmacology and, 2919
- Pharmacological bridge, 1608
- Pharmacology. *See also* Psychopharmacology; *specific drugs*
  - of anabolic-androgenic steroid-related disorders, 1393–1396, 1394*f*
  - anabolic-androgenic steroids, 1393–1396, 1394*f*
  - barbiturates, 1375
  - benzodiazepines, 1375, 1377
  - caffeine, 1295, 1295*f*
  - cannabinoids, 1305–1306
  - dimethyltryptamine (DMT), 1316
  - drug approval, 2934
  - hallucinogens, 1312, 1314–1316
  - inhalants, 1331–1333
  - lysergic acid diethylamide (LSD), 1315–1316
  - mescaline, 1316
  - nicotine, 1345
  - opioids, 1357–1360, 1357*f*, 1358*f*
  - psilocybin, 1316
  - receptor pharmacology, **2905–2914**

- Pharmacotherapy. *See also* Hallucinogen-assisted psychotherapy; Psychopharmacology
  - for adolescents, 3813
  - aggression
    - Alzheimer disease, 2502
    - autism, 2500
    - borderline personality disorder, 2501–2502
    - conduct disorder, 2501
    - intermittent explosive disorder, 2501
    - intimate partner violence, 2502–2503
    - mentally retarded children, 2500
  - anxiety disorders, 1778–1782, 3710, 4063
    - patient-related factors and, 1779*t*
    - sexual dysfunction and, 1779*t*
  - assault victim, 2541
  - BDD, 1792
  - biological or chemical terrorist attacks and, 2572
  - bipolar disorder, 4076–4077
  - bipolar disorder, early onset, 3692–3693, 3706–3707
  - circadian rhythms and, 215
  - combination, **3322–3329**
    - BD, 3325–3327
    - future directions, 3328
    - major depression, 3323–3325
    - overview, 3322–3323
    - schizophrenia, 3327–3328
  - conversion disorder, 1840
  - of delirium, 1187–1190, 1189*t*
  - depersonalization/derealization disorder, 1945
  - depressive disorders, 3378–3379, 3379*t*
  - dissociative identity disorder, 1926–1928, 1927*t*
  - early-onset bipolar disorder, 3692–3693, 3706–3707
  - early-onset schizophrenia, 3732
  - encopresis and, 3660–3661, 3674–3675
  - enuresis, 3655–3657, 3669–3671, 3794
  - generalized anxiety disorder, 3714
  - hoarding disorder, 1794
  - mood disorders, 1251

- obsessive-compulsive disorder, 3700–3701, 3701*t*
- OCD, 1789
  - CBT and, 1790
- OCRDs, 1787
- opioid-related disorders, 1365–1370, 1366*t*, 1367*t*, 1371–1373
- personality disorders, 4108
- psychiatric emergencies, 2627–2628
- psychoanalysis and, 2658
- psychotherapy and, 2664, 2864
  - anxiety disorders, 1782
  - mood disorders, 1703
  - in treating PDs, 2168–2169
- PTSD, 1823–1824, 3706
- in public psychiatry, 4300
- schizophrenia, **1519–1531,** 4098
  - for acute episodes, 1523–1525
  - effectiveness, 1520, 1523
  - evidence-based medicine, 1552
  - first episode, 1520
  - future directions, 1530–1531
  - history, 1519–1520, 1521*t*–1522*t*
  - maintenance, 1529–1530
  - negative symptoms and, 1528
  - phases of, 1520
  - poor response, strategies for, 1528–1529
  - psychosocial treatment and, 1530
  - side effect management, 1525–1528
  - suicidal behavior, 1528
- selective mutism, 3723
- separation anxiety disorder, 3710, 3711–3712
- social anxiety disorder, 3717–3718
- somatic symptom disorder, 1837
- stereotypic movement disorders, 3672, 3686
- stimulant use disorders, 1289–1290
- substance abuse, 4117
- suicidal behavior, mood disorders and, 1674
- suicide prevention and, 2620–2621

- suicide risk, 2616
- tic disorders, 3644–3648, 3658–3662, 3795
- tobacco-related disorder treatment, 1349–1351, 1350*t*

- Phase-amplitude coupling (PAC), 408–409, 408*f*
- Phase of life problems, 4602
- Phencyclidine (PCP)
  - aggression and, 2493
  - aggression threshold and, 2493
  - anxiety disorder, induced, 1327
  - bipolar disorder, induced, 1326
  - depressive disorder, induced, 1326–1327
  - description of, 4602
  - effects of, 1312–1315
  - etiology
    - epidemiology, 1321, 1323*t*
  - history, 1317
  - intoxication, 2631–2632
    - course and prognosis, 1321
    - defined, 1318–1319
    - diagnosis and clinical features, 1319–1320
    - differential diagnoses, 1320–1321
    - epidemiology, 1319
    - etiology, 1319
    - pathology and laboratory examination, 1320
    - treatment, 1321
  - intoxication delirium, induced, 1327
  - neural discharges, 404*f*, 405
  - phencyclidine use disorder, 1321–1323, 1322*t*
    - course and prognosis, 1323
    - defined, 1321
    - diagnosis and clinical features, 1323
    - differential diagnoses, 1323
    - epidemiology, 1321–1322, 1322*t*
    - etiology, 1322–1323
    - laboratory examination, 1323
    - pathology, 1323
    - treatment, 1323

- psychiatric emergency due to, 2631–2632
- psychotic disorder, 1325–1326
- structure of, 77*f*
- unspecified disorder, 1327

- Phenelzine (Nardil)
  - agoraphobia, 1783, 3122–3123
  - anxiety disorders, 1780, 3122
  - BPD, 3121
  - chemistry, 3113*f*
  - cognitive-behavioral group therapy *versus,* 1776
  - depression, 1681, 3118–3119, 3120–3121
  - dissociative identity disorder, 1928
  - eating disorder, 3121–3122
  - history, 3111
  - OCD, 3122
  - organs/systems effects of, 3117
  - pain syndromes, 1841
  - panic disorder, 3122–3123
  - pharmacological actions, 3115
  - precautions and warnings, 3118, 3118*t*
  - PTSD, 1928, 3122
  - SAD, 1783–1784
  - social anxiety disorder, 3122
- Phenergan. *See* Promethazine
- Pheniprazine, 3110
- Phenobarbital (Luminal)
  - drug interactions, 2988
    - valproate, 2988
  - history, 2984
  - intellectual disability and, 3510–3511
  - laboratory interferences, 2988
  - lamotrigine and, 3083
  - pharmacological actions, 2984–2985
  - precautions/adverse effects, 2988
  - therapeutic indications, 2986, 2987
- Phenomenology
  - anxiety disorders, 1769

- description of, 4602
- fear and stress-related disorders, 1751–1753
- mood disorders, clinical, 1709–1710
- schizophrenia, **1406–1425, 1512–1519**

- Phenothiazines
  - adverse effects of
    - dermatological, 2356*t*
  - chemistry, 3061, 3061*t*, 3062*f*, 3064
  - description of, 4602
  - elimination, 3065
  - laboratory interferences, 3078
- Phenotypes, 309–311
  - categorical, 309–311
  - continuous, 311
  - description of, 4602
  - and natural selection, 772
  - of psychosis, **1512–1519**
- Phenotypic trait personality models, 2145
- Phenotyping
  - computational, 437
  - strategies, 310*f*
- Phentermine (Adipex)
  - mechanism of action, 2221
  - for weight loss, 2220
- Phentermine/topiramate (Qysmia)
  - mechanism of action, 2220–2221
  - for weight loss, 2220–2221
- Phenyalaline hydroxylase *(PAH),* 660–661
- Phenylalanine, 2550
  - metabolism, 3493
- Phenylethanolamine-N-methyltransferase (PNMT), 63
- Phenylketonuria (PKU)
  - description of, 660–661, 4602
  - intellectual disability and, 3493
  - MRI imaging in, 661*f*
  - mutation causing, 63
- Phenylpyruvic oligophrenia. *See* Phenylketonuria

- Phenytoin (Dilantin)
  - for aggressive behavior, 2111, 2501
    - in IED, 2501
  - drug interactions, 3131
  - hallucinogen persisting perception disorder treatment, 1324
  - intermittent explosive disorder, 2501
  - lamotrigine and, 3083
  - mechanism of action, 507
- Pheochromocytoma
  - familial forms of, 683–684
  - presentation of, 683–684
- Pheromones, 1980, 4602
- Phillips, Adam, 2645
- Philosophy
  - Aristotle, 928
  - Derrida, Jacques, 929
  - Gergen, Kenneth, 929
  - Kant, Immanuel, 928–929
  - Plato, 927–928
  - postmodernism, 929
  - Schopenhauer, Arthur, 928–929
  - Shankara and eastern thought, 928
  - Szasz, Thomas, 929
- Phobias
  - anxiety and, 1723
  - cancer patients
    - presentation of, 2258
  - case study, 2678
  - CBT and, 1774
  - conditioning theories and, 416
  - description of, 4602
  - DSM-5 on, 1721
  - epidemiology
    - genetic, 1738
  - genetics and, 2529*t*
    - empirical risks for, 2535*t*
  - noradrenergic system dysfunction and, 1742

- older adults, 3965–3966
- phenomenology of, 1752
- social
  - amygdala role, 703
- specific
  - cognitive therapy, 2862
  - definition, diagnosis, and clinical features, 1726, 1726*t*
  - description, 1768–1769
  - differential diagnoses, 1726–1727
  - DSM-5 on, 1721, 1723
  - epidemiology, 1726, 1732, 1732*t*
  - genetic epidemiology, 1738
  - history and comparative nosology, 1726
  - management, 1784
  - neurocircuitry, 1756
  - neuroimaging, 1762–1763, 1765
  - phenomenology of, 1752
  - psychodynamic psychotherapy, 2678
- VRGET for, 2560

- Phobic neurosis. *See* Phobias
- Phonological disorder. *See* Developmental articulation disorder
- Phosgene, war or terrorism and, 2572*t*
- Phosphatases, 124
- Phosphatidyl serine, 2551–2552
- Phosphatidylinositol pathway, 122, 122*f*
- Phosphodiesterase inhibitors, 3338–3339, 3339*t*
- Phosphodiesterase V inhibitors, 4261
- Phosphodiesterases (PDEs)
  - cAMP, 2914–2915, 2915*f*
  - cAMP signaling and, 122
- 3-phosphoglycerate dehydrogenase (3-PGDH), 664
- Phosphoinositide 3-kinase (PI3K), 126, 126*f*
- Phospholipases, 2914, 2914*f*
- Phosphorylation, 149
- Phosphoserine aminotransferase (PSAT) deficiency, 664
- Photic stimulation, 241–242
- Photomultiplier tubes (PMTs), 287–288, 288*f*

- Photomyoclonic responses (PMRs), 242
- Photon attenuation, 291, 291*f*
- Photon emission, physical principles, 286–293
- Photon energy, units of, 285
- Photon scatter, 291
- Photophobia, 4602
- Phototherapy, 1651, 1679
- PHQ. *See* Patient Health Questionnaire
- Phrenology, 4603
- Phylogenetic, definition of, 4603
- Phylogeny of behavior. *See* Evolutionary biology
- Physical abuse
  - adults, 2538–2542
    - domestic violence, 2539–2540
    - intimate partner violence, 2539–2540
    - rape, 2538–2539
    - sequelae of violence and assault, 2540
    - treatment issues, 2540–2541
    - types, prevalence, and risk factors, 2538
    - violent crime, 2538
    - workplace violence, 2540
  - amnesia, 1886
  - autobiographical memory deficits and, 1886
  - child
    - evaluating, 3391
    - maltreatment as, 3829, 3831–3832, 3834, 3836–3837
    - parent interviews and, 3834
  - conduct disorder and, 3616
  - dementia patients, 1200
  - dissociative disorders and, 1870
  - pathological dissociation and, 1884
  - PTSD, borderline PD and, 1713
- Physical activity
  - and healthy brain aging, 3994
  - older adults, 3979, 3979*t*, 3983
  - positive psychiatry, 4285
- Physical dependence. *See* Drug dependence

- Physical development, evaluation of, 3378
- Physical disability, everyday functioning and, 4027
- Physical disorders, definition of, 4603
- Physical examinations. *See also* Medical assessment and laboratory testing
  - child maltreatment, 3836–3837
  - of patients, 455–460
- Physical exercise, 1244
- Physical health, adult, 3941, 3942
- Physical health problems
  - developed/developing country differences in, 4520*t*
  - mental illness comorbidities with, 4288–4289, 4317, 4343, 4518, 4529–4530
- Physical therapy, 3543–3544
  - conversion disorder, 1840
  - for developmental coordination disorder, 3543–3544
- Physician-assisted suicide, 2313, 2622, 4555
- Physicians
  - aversive reactions in, 2290*t*
  - condolences from, 2303
  - moral burden of
    - dying patients and, 2314
  - psychological reactions of, 2289–2290
    - help for, 2290
  - referral questions for, 1007
- Physician(s)
  - addictions, 2601
  - aging, 2605–2606
  - anxiety disorders, 2601
  - burnout, 2599
  - clinical *versus* administrative, 2594–2595
  - countertransference issues, 2608–2609
  - depression, 2600
  - disability, 2602
  - Hippocratic oath and, 2594
  - medical illness, 2602
  - mental health, 2599–2602
    - administrative response to, 2602
    - disciplines involved, 2594

- future directions, 2609
- history, 2596
- risks unique to profession, 2596
- treatment approaches/alternatives, 2606
- treatment barriers, 2606–2608

- midcareer, 2605
- minority, 2598–2599
- overwork, 2599
- personality disorders, 2601
- personality types, 2597–2598
  - case studies, 2598
- preventive health, 2596–2597, 2597*f*
- relationship problems, 2601–2602
- stress, 2599
- suicide, 2599–2600
- temperament, 2598
- wellness, 2596–2597, 2597*f*
- women, 2598

- *Physiognomical System of Drs. Gall and Spurzheim* (Gall), 4535
- Physiological reactions, cognitive model, 2760, 2760*f*
- Physiological resilience, definition of, 2336
- Piaget, Jean, **704–716**
  - on cognitive development, 3359–3360, 3369
  - genetic epistemology, 704–705
  - intellectual functioning, 4543
  - models of intelligence, 709–711
    - developmental levels, 713–714
      - use in psychotherapy, 714–716
    - emotional basis of, 711–713
    - extensions of theory, 710
    - post-Piaget era, 711
  - on moral development, 3370
  - *The Moral Judgement of the Child,* 710
  - *Play, Dreams and Imitation in Childhood,* 707
  - theory of cognitive development, 705–709
    - assimilation and accommodation, 705
    - equilibration, 704–705

    - stages of, 693–694, 706–709, 706*t*
      - concrete operations, 707–708
      - formal operations, 708–709
      - preoperational subperiod and semiotic function, 707
      - sensorimotor period, 706–707
    - structuralism, 705–706
  - work of, 4603
- Piblotko, description of, 4603
- Pica, **3632–3633, 3646–3647**
  - clinical features, 3632, 3646
  - course, 3633, 3647
  - description of, 2066, 4603
  - diagnosis of, 2073
  - differential diagnoses, 2077, 3632, 3646
  - epidemiology, 3632, 3646
  - etiology, 3633, 3647
  - historical literature on, 2067
  - nosology, 2069
  - treatment, 2082, 3633, 3647
- Pick, Arnold, 1210
- Pick disease. *See* Frontotemporal dementia (FTD)
- Pickwickian syndrome, 4603
- PICOTS. *See* Population, Intervention, Comparator, Outcome, Timing, and Setting
- Picrotoxin, action of, 81
- Picture drawing. *See* Drawing
- Pierce, Franklin, 4290, 4335
- Pigeonholing
  - in attention processing, 690
  - in schizophrenia, 696
- Pilocarpine, 73
- Pimavanserin, 548
- Pimozide (Orap)
  - cardiac side effects, 3075
  - chemistry, 3063*f*, 3064
  - for delusional parasitosis, 2347
  - drug interactions, 3131
  - side effects

    - QT interval prolongation, 2199
  - tic disorders, 2947, 3645, 3646, 3659, 3660, 3795
  - TS, 2947, 3071
- Pindolol (Visken)
  - 5-HT_{1A} receptor agonist, 303
  - absorption, 2951
  - for aggressive behavior
    - Alzheimer disease, 2502
    - in IED, 2501
    - intermittent explosive disorder, 2501
    - mentally retarded children, 2500
    - TBI, 2502
  - chemistry, 2951, 2951*f*
  - combination pharmacotherapy and, 3325
  - distribution, 2951–2952
  - dosing/administration, 2956
  - drug interactions, 2955
  - drug studies, 2952
  - in HIV setting, 572
  - laboratory interference, 2955
  - organ/system effects of, 2952–2953, 2953*t*
  - pharmacodynamics, 2952
  - pharmacokinetics, 2951–2952, 2952*t*
  - pharmacological actions, 2951–2952, 2952*t*
  - precautions/adverse reactions, 2955
  - teratogenic effects of, 2953
  - therapeutic indications, 2953–2955, 2953*t*
- Pine, Fred
  - hypnosis and, 2723
  - as psychoanalyst, 2647
- Pineal gland, 69
- Pinel, Philippe, 1426, 1548, 2322
  - as early reformer, 4290
  - *A Treatise on Insanity in Which Are Contained the Principles of a New and More Practical Nosology of Mental Disorders,* 4534
  - work of, 4603
- Pioglitazone

- insulin sentization by, 2248
- mood effects of, 2247
- *Piper methysticum* (kava), 2553, 2553*f*
- Piperazines
  - chemistry, 3062*f*, 3064
  - description of, 4603
- Piperidines, 3064
- Piromelatine, 3106–3107
- Pitfalls after release, 4399
- Pithiatism
  - history, 1868
  - as substitute term for hysteria, 1867
- Pitolisant, 75
- Pituitary gland
  - adenoma
    - acromegaly and, 679*f*
    - MRI imaging of, 679*f*
  - disorders of, 677–678
  - gender differences, 3463
  - hormone production and, 2239
  - major depressive disorder
    - neuroimaging, 3466
  - regulation of, 35
  - tumors
    - adrenocorticotrophin secreting, 678–679
    - growth hormone secreting, 678–679
    - prolactinomas, 678–679
    - psychiatric symptoms of, 630
- PKA. *See* Protein kinase A
- PKU. *See* Phenylketonuria (PKU)
- Placebo-controlled, randomized clinical trials, 4370
- Placebos
  - description of, 4603
  - effects, 4603
  - epidemiology of, 815
  - responses
    - in pain syndromes, 2397

- Placidyl. *See* Ethchlorvynol
- *Plague of Corn: The Social History of Pellagra* (Roe), 4538
- Plan B, 2512
- *Planned Parenthood of Southeastern Pennsylkvania v. Casey,* 2512
- Planum temporale
  - anatomy, 20
  - lateralization, 449
- Plaques, description of, 4603
- Plasma levels. *See* Blood levels
- Plasma membranes, 2912, 2913*f*
- Plasticity. *See* Neuronal plasticity
- *Plata, Brown v.,* 4400
- Plateau phase, 4603
- Platelet-derived growth factor (PDGF), 52
- Plater, Felix, 4534
- Plato, 927–928
  - *The Dialogues of Plato,* 4533
  - learning theory, 411
  - near-death experiences and, 2589
  - on psychosomatic medicine, 2177
- Play
  - in adulthood, 3943–3945
  - autistic spectrum disorder, 3580
  - child interview technique, 3382–3383
  - in cognitive development, 707
  - psychodynamic psychotherapy and, 3736–3737
  - symbolic, 707
- *Play, Dreams and Imitation in Childhood* (Piaget), 707
- Play therapy
  - description of, 4603
  - selective mutism (SM), 3723
- Pleasure principle, 4603
- Pleiotropy
  - definition of, 164
  - schizophrenia, 1444
- Pleonexia, 4603
- PLMs. *See* Periodic leg movements

- Pluralism, 4603
- PMDD. *See* Premenstrual dysphoric disorder
- PML. *See* Progressive multifocal leukoencephalopathy (PML)
- PMS. *See* Premenstrual syndrome
- *Pneumocystis carnii. See Pneumocystis jiroveci*
- *Pneumocystis jiroveci* infections, 565
- Point biserial discrimination index (PBDI), 4419
- Point estimators, 831
- Point prevalence, 4603
- Point spread functions (PSFs), 292, 292*f*
- Point Subtraction Aggression Paradigm, 170
- Point swarms in QEEG, 1035, 1035*f*
- Poisoning, child maltreatment and, 3837
- Police
  - crisis intervention by, 4393
  - as first responders, 4390
  - training of, 4393–4394
  - triage work by, 4330
- Police power. *See* Parens patriae
- Policies
  - in disparities reduction, 4386
  - health care impact of, 4379–4381
  - idealists *vs.* realists, 4294
  - world-wide, 4524–4525, *4528t*
- Policy-level interventions, 4348
- Poliomyelitis, 627
- Polychlorinated biphenyls, 2484
- Polycomb Repressive Complex (PRC2), 390*f*, 391*f*
- Polycystic ovary syndrome (PCOS)
  - clinical features of, 2247–2248
  - definition of, 2247–2248
  - etiology, 2248
  - psychiatric illness and, 2248–2249
  - treatment of, 2249
  - valproate and, 3277, 3278
- Polygenetic threshold model, 1447
- Polygenic, definition of, 4603

- Polygenic quantitative trait (QTL) model, 1621–1622, 1621*f*
- Polymerase chain reactions (PCRs), 85
- Polymorphisms
  - definition, 155
  - psychiatric diseases and, 3448
- Polymorphous perverse sexuality, 4603
- Polymyositis (PM), 634
- Polyphagia, 4603
- Polypharmacy, 2923. *See also* Pharmacotherapy, combination
- Polysomnography, 243, 243*f*
  - description of, 2106, 4603
  - role in diagnosis, 1020
  - stage 2 sleep, 2108*f*
- Polysubstance dependence, 4603
- Polysubstance use and abuse, 1335
- Polysystem use disorder, 2046
- Polyuria
  - description of, 4603
  - lithium and, 3097
- Polyvagal theory, 1890
- Pons
  - coronal section, 12*f*
  - REM sleep and, 347–348
  - surface landmarks, 7*f*
- Pontine system, 14
- Ponto–geniculo-occipital (PGO) waves, 347
- Poorhouses
  - asylums as outgrowth from, 4290
  - as earliest care, 4326–4327
- Popper, Karl, 2650
- Popular illness category, 2444
- Population, Intervention, Comparator, Outcome, Timing, and Setting (PICOTS), 972
- Populations
  - description of, 4603
  - genetic isolates
    - schizophrenia, 1429–1431
  - migrant, 1431, 1450–1451

- schizophrenia, 1427–1428
- Poreddi, V., 4529
- Porges, Steven, 1890
- Porins, 141. *See also* Voltage-dependent anion channels (VADCs)
- Pornography, addiction to, 2041
- Porphobilinogen (PBG) deaminase, 660
- Porphyria
  - acute intermittent, 2530
  - description of, 4603
  - features of, 467*t*
- Porphyrin, accumulation of, 660
- Porpoises, sleep in, 352
- Porta, Giambattista, 4534
- Positional therapy, 2095
- Positive airway pressure (PAP), 2093, 2094
- Positive and Negative Affect Scale (PANAS), 764
- Positive and Negative Syndrome Scale (PANSS)
  - medication dose determination with, 4367–4368, 4370
  - schizophrenia, 1410, 1416, 1416*t*
- Positive behavioral support, 2854
- Positive emotions, 3481
- Positive institutions, 2849
- Positive mental health, 4281–4286
  - biological factors
    - allostatic load, 4284
    - biomarkers, 4284
    - genetics, 4284
    - inflammation and immune factors, 4284
    - neuroplasticity, 4284
    - oxidative stress, 4284
    - telomere length, 4284
  - definitions, 4281
  - diet and nutrition, 4285
  - environmental factors, 4283–4284
  - history, 4281–4282
  - interventions, 4285
  - posttraumatic growth, 4282

- psychosocial factors, 4282–4283
- recovery from illness, 4282
- stress, 4282
- successful aging, 4282
- well-being, 4282

- Positive pressure ventilation, 2094–2095
- Positive psychology, **2841–2855**
  - ACT and, 2853
  - assessment, 2850, 2850*t*
  - behavior therapy and, 2854
  - clinical work and, 2849–2854
  - concerns and criticisms, 2845
  - conclusions, 2854–2855
  - empirical findings, 2845–2849
  - examples, 2852–2853
  - history, 2841–2843
    - early ancestors, 2842–2843
    - humanistic psychology and, 2843
    - positive mental health and, 2843–2845, 2844*f*
  - hope therapy and, 2853
  - identifying, 2852
  - introduction and definition, 2841
  - MBCT and, 2853
  - psychological health vision of, 2849–2850
  - quality-of-life therapy and, 2854
  - session sequence, 2853, 2853*t*
  - taking stock, 2854
  - techniques, 2850–2852, 2851*t*
  - well-being therapy and, 2853–2854
- Positive psychotic symptoms.
  See also Catatonia; Delusions; Hallucinations
  - defined, 3725
  - neurocognitive impairment, 1507
  - schizophrenia, 1418, 1507
- Positive punishment, 2872
- Positive reinforcement. *See* Reinforcement
- Positive symptoms, 4603
- Positive transference. *See* Transference

- Positron emission, decay by, 286–287
- Positron emission tomography (PET). *See also* Forensic psychiatry, neuroimaging
  - after TBI, 525–526
  - Alzheimer disease, 4081–4082, 4081*f*
  - brain imaging, 1717
  - coincidence detection in, 288
  - compartment models, 293–297
  - description of, 4603
  - dissociative identity disorder, 1882
  - fear, 1759–1760
  - functional neuroimaging with, 1759
  - gambling disorder, 1806–1807
  - hypnosis, 2712
  - for MCI, 1226–1227, 1226*t*, 1227*f*
  - neuroreceptor imaging studies, 1759
  - panic disorder, 1761
  - pediatric considerations, 3459, 3460
  - radionucleotides, 286*t*
  - radiotracers used in, 298*t*
  - random coincidences in, 292
  - reversible models, 294–295
  - safety issues, 293
  - in schizophrenia, 273
  - social anxiety disorder, 1763
  - spatial resolution, 292
  - specific phobias, 1762
- Positron emission tomography-computed tomography (PET-CT), 292
- Positron emission tomography-magnetic spectrometry (PET-MR), 292–293
- Possession disorder, 1875*t*. *See also* Trance and possession disorders
- Possession-form dissociative identity disorder, 1879
  - childhood trauma and, 1904
  - course and prognosis, 1917
  - DBS model and, 1906
  - dissociative hallucinations and, 1911
  - factitious dissociative identity disorder and, 1916–1917
  - interference symptoms, 1912
  - non–possession-form compared with, 1914*t*

- onset of, 1912–1913
- Possession syndrome
  - epidemiology of, 2454
  - phenomenology of, 2453–2454, 2453*t*
  - precipitants/etiology of, 2454
  - presentation, 2452–2455
  - psychiatric diagnoses and, 2452–2453, 2454–2455
  - relapse rates, 2455
  - specific cultural factors, 2455
  - treatment, 2455
- Post-incarceration care, rates of, 4305
- Post-intensive care syndrome, 2329–2330
- Post-mortem studies, ASD, 3576–3577
- Post-treatment Lyme syndrome, 589–592
- Postcoital dysphoria, 1974
- Postcoital headache, 1974–1975
- Postconcussion syndromes, 245
- Postdepressive personality, 1644
- Posterior cingulate cortex, 7*f*
- Posterior nucleus, 36
- Posterior reversible encephalopathy syndrome (PRES)
  - immunosuppression and, 2371
  - kidney disease and, 2369
- Postherpetic neuralgia (PHN)
  - differential diagnoses, 2390–2391
  - treatment
    - anticonvulsants, 2967, 2968, 2977
  - treatment of, 629
- Posthypnotic effects, 2708
- Postictal aggression, 2496
- Postictal psychosis
  - criteria for, 649*t*
  - pediatric, 649–650
- Postmodernism, 929
- Postnatal risk factors, aggression and, 2485
- Postpartum blues, 2514
- Postpartum depression (PPD)

- description of, 4603
- prophylaxis, 2515
- treatment, 2515–2516
  - estradiol, 3335

- Postpartum disorders
  - estrogens for, 171
  - neurosteroids and, 118
- Postpartum psychosis, 2516, 4603
- Postpolio syndrome
  - etiology of, 627
  - neuropsychiatric aspects, 627
- Postprandial satiety, 360–363
- Postpsychotic depressive disorder of schizophrenia, 1648
- Poststroke depression (PSD)
  - course/prognosis in, 484–485
  - depression *versus,* 476
  - diagnosis of, 481
    - laboratory studies, 482–483
  - etiology of, 478–479
  - prevention of, 488*f*
- Posttraumatic amnesia, dissociative *versus,* 1935–1936
- Posttraumatic Check List-Civilian version (PCL-C), 1920
- Posttraumatic embitterment disorder, 2121–2122
- Posttraumatic feeding disorder, **3626–3628, 3640–3642**
  - clinical features, 3626–3627, 3640–3641
  - course, 3627, 3641
  - differential diagnoses, 3627, 3641
  - etiology, 3627, 3641
  - treatment, 3627–3628, 3641–3642
- Posttraumatic stress (PTS), 2259
- Posttraumatic stress disorder (PTSD), **1812–1826, 3703–3707**
  - after TBI, 536–537
  - aggression and, 2494
  - amnesia, 1887
  - anxiety disorders and, 1721
  - BD and, treatment of, 1699
  - brain imaging

    - functional, 1821
    - structural, 1820–1821
  - burn injuries in children and, 2385
  - in burn patients, 2378
  - in cancer patients, 2259
  - in cardiac patients, 2197–2198
  - child/adolescent
    - after TBI, 652–653
    - biological factors, 3704
    - biological treatments, 3707
    - clinical features, 3705
    - comparative nosology, 3703
    - course, 3706
    - defined, 3703
    - diagnosis, 3704–3705
    - differential diagnoses, 3705–3706
    - epidemiology, 3703
    - etiology, 3703–3704
    - future directions, 3707
    - group therapy and, 3764–3765, 3769
    - history, 3703
    - pathology of, 3705
    - prognosis, 3706
    - psychological factors, 3704
    - psychopharmacological treatments, 3706–3707
    - psychotherapeutic treatments, 3706
    - social factors, 3704
    - treatment, 3706–3707
    - types of, 3705
  - cholecystokinin elevation in, 96
  - civil litigation, 3865
  - Civil War medicine on, 2565
  - classification of, 1158, 1173
  - CNS mediation of clinical manifestation of, 1817–1818
  - cognitive control, 1819
  - combat-related, 2487
  - comparative nosology, 1815, 1874

- complex PTSD, 701
- coronary artery disease and, 2189
- course and prognosis, 1822–1823
- DBT and, 2794–2795
- definition, 1812–1814
- depersonalization/derealization, 1887
- description of, 4603
- diagnosis and clinical features, 1816–1817
- differential diagnoses, 1275, 1822, 3731
- disasters and, 2567–2568, 2567*t*
- dissociative amnesia and, 1882
- dissociative amnesia *versus,* 1937
- dissociative identity disorder and, 1904–1905, 1904*t*, 1913
- dissociative identity disorder *versus,* 1882
- dissociative subtype, 1881
  - corticolimbic disconnection hypothesis, 1891
  - dissociative identity disorder and, 1910–1911
  - DSM-5 on, 1874, 1878
  - treatment outcome studies, 1918–1919
- dissociativity, 1870
- DSM-5 on, 1813*t*
- DSM-III on, 1870
- emotional numbing in, 1892
- emotional regulation, 1818–1819
- epidemiology
  - in community, 1815–1816
  - prostitutes, 1889
  - in veterans, 1816
- epigenetic changes, 1820
- etiology, 1816
- FAAH inhibitors for, 1743*t*
- families of survivors, 2330
- flashbacks in, 2428
- forensic psychiatry and, 3865
- future directions, 1825–1826
- genetics and risk for, 1816, 1820
- group therapy and, 3764

- history, 1814–1815
- in HIV patients, 582–583
- HPA axis and, 167
- hypnosis and, 2717–2719
  - case study, 2718–2719
- hypnotizability and, 2707
- impaired affect symbolization in, 2051
- in indigenous Americans, 801
- inflammatory biomarkers in, 201
- intimate partner violence and, 2487
- introduction, 1812
- in long-term care, 4231
- malingering and, 2405–2406
- memory disturbances, 1819
- military-related
  - dissociative disorders, 1897
  - genetics and, 1889
  - psychotherapy, 1823
- neuroendocrine modulation, 1819
- neuroimaging, psychiatric, 3471
- in older adults, 3966
- pain complaints in, 2393–2394
- parental neglect and, 1713
- pathological dissociation and, 1884–1885
- pathology and laboratory examination, 1817–1822
- prevention of
  - in burn patients, 2379
  - context of, 1825
  - pharmacotherapy, 1825
- psychiatric symptomatology of, 1116–1117
- psychodynamic psychotherapy, 2679–2680
  - case study, 2680
- psychological factors, 3704
- psychophysiological studies, 1818
- reliability of, 1815
- risk factors, 1816
- rodent models of, 396

- schizophrenia and, 1421–1422, 3731
- secondary, 1925
- social origin of, controversies on, 1817
- speechless terror, 699–700
- stresses and, 2324
- substance P elevation in, 96
- suggestibility, 700
- survivors of critical illnesses, 2330
- symptom onset, 1823
- terrorism and, 2567–2568, 2567*t*
- threat detection processing and, 1818
- in transplant populations, 2366–2367
- trauma model, 1896
- traumatic brain injury and, 1243
- trazodone drug studies, 3249
- treatment, 1823–1825
  - acupuncture, 2562
  - α-adrenergic receptor agonists, 2949
  - β-adrenergic receptor antagonists, 2954–2955
  - alternative and complementary, 1824
  - anticonvulsants, 2970
  - benzodiazepines, 2996
  - buspirone, 3010
  - carbamazepine, 3037
  - comorbid SUD and, 1823
  - early and preventive, 1824
  - early CBT, 1824–1825
  - ketamine (Ketalar), 1747
  - MAOIs, 3122
  - nefazodone, 3127
  - pharmacotherapy, 1823–1824
  - psychological debriefing, 1824
  - psychotherapy, 1823
  - psychotherapy and pharmacotherapy, 1823–1824
  - SSRIs, 3154
  - valproate, 3275
- twin studies, 1816, 1820

- validity of, 1815
  - controversies on, 1817
- war and, 2567–2568, 2567*t*
- in youth with HIV, 646

- Postural hypotension, 2193
- Postural tremor, medication-induced, 2942
- Posturing. *See also* Catatonia
  - description of, 4603
- Potassium ion (K^+) channels
  - A-type, 138
  - calcium-activated, 137
  - as delayed rectifier, 137
  - heteromeric, 137
  - homomeric, 137
  - inhibition of, 137*f*
  - M-type, 138
  - molecular structure of, 136*f*
  - mutations, 144
  - neuron, 136–139
  - S-type, 138
- Potency, definition of, 4603
- Potentiation, long-term
  - in memory systems, 690–691, 729
- Poverty
  - homicide rates and, 2476
  - in U.S. population, 795
  - violence and, 2426, 2476
- Poverty of content of speech, 4604
- Power spectrum. *See* Quantitative electroencephalography (Q-EEG)
- Powers of attorney, 4238
- PPA. *See* Primary progressive aphasia (PPA)
- PPACA. *See* Patient Protection and Affordable Care Act
- PPHN. *See* Persistent pulmonary hypertension of the newborn
- PPOs. *See* Preferred Provider Organizations
- *Practical Treatise on Nervous Exhaustion* (Miller), 4536
- Practice effect, 4604
- Practice guidelines
  - adherence to, 4315
  - advantages of, 974–975, 974*t*
  - AMA guidelines, 970–971
  - APA guideline project

    - committees and work groups, 972
    - dissemination and implementation, 974
    - evidence identification, 972–973, 972*t*
    - format, 973
    - history, 971–972
    - recommendations, 973
    - review and approval process, 973–974
    - topic choices, 972
    - updates, 974
  - association guidelines, 971
  - early efforts, 970
  - future of, 976
  - Institute of Medicine reports, 971, 971*t*
  - limitations of, 974*t*, 975–976
- *Practice of Medicine* (Plater), 4534
- Practitioner–scholar model, 2684
- Prader–Willi syndrome, 3506–3507
  - description of, 4604
  - epigenetic imprinting and, 161
  - excoriation disorder and, 1797
  - hoarding and, 1794
  - intellectual disability and, 3506–3507, 3514
  - laboratory evaluation, 1016
  - learning styles and, 3517
- Pragmatic language impairment, 3551
- Pramipexole (Mirapex)
  - bipolar depression, 1692*t*, 1693, 3236
  - bipolar II disorder, 1698
  - chemistry, 3234, 3234*f*
  - combination pharmacotherapy and, 3324
  - dosage/administration, 3241, 3241*t*
  - drug interactions, 3240
  - history, 3233
  - MDD, 3236
  - medical illness-induced fatigue and depression, 3237
  - pharmacological actions, 3234–3235
  - precautions/adverse reactions, 3239, 3240*t*

- receptor profile of, 67
- RLS, 3238
- therapeutic indications, 3239

- Prange, Arthur J., Jr., 1713
- Pratt, Joseph, 2736
- Pravachol. *See* Pravastatin
- Pravastatin (Pravachol)
  - for Alzheimer disease, 4147
  - drug interactions, 3131
- Prayer
  - as energy therapy, 2546
  - religious life and, 2592
  - spirituality and, 2579–2580
- Prazepam (Centrex), 2993
- Prazosin (Minipress)
  - biological or chemical terrorist attacks and, 2572
  - dissociative identity disorder
    - PTSD symptoms, 1927, 1927*t*
  - nightmares in PTSD, 1819
  - PTSD, 1823
    - SUD comorbidity and, 1823
  - side effects of, 67
- Pre-ictal features, 512
- Pre-menstrual syndrome (PMS)
  - mood disorders in, 118
  - testosterone in, 170
- Preadjudicatory examinations, 3866
- Preamplifiers, MR, 264
- Preattentive processing, models of, 323
- Prechtl, Heinz, 1900
- Precipitating factors, conceptualization and, 2688
- Preconscious, definition of, 4604
- Preconscious system, 861
- Prediction error models, 413
- Predictive validity, 320, 996
- Predictive values, 4604
- Predictor variables, 4604

- Predisposing factors, 2688
- Prednisone (Deltasone), 2255
- PREE. *See* Partial reinforcement extinction effect (PREE)
- Preferred provider networks, 4312
- Preferred Provider Organizations (PPOs), 4604
- Prefrontal cortex (PFC)
  - activation patterns, 277–278
  - cognitive deficits and, 433–434
  - cross-section, 9*f*
  - dissociative amnesia and, 1890
  - dysfunction, depression and, 1709
  - emotional processing and, 1711–1712, 1712*f*
  - episodic memory and, 2139
  - executive functioning and, 3430–3431
  - fear conditioning, PTSD and, 1818
  - glial cell changes in, 5
  - instinctual responses and, 2427
  - and mental health, 766
  - monkey
    - connections, 6*f*
  - mood disorders and, 1718*f*
  - neuroimaging
    - ADHD, 3463–3464, 3463*f*
    - anxiety disorders, 1765
    - drives and instincts and, 2815
    - GAD, 1764
    - OCD, 3469, 3469*f*
    - panic disorder, 1762
    - schizophrenia, childhood-onset, 3474–3475
  - perspective taking and, 337
  - psychotherapy and, 2813
  - regions of, 8
  - schizophrenia, 1458, 1458*t*, 1459*f*
  - sleep and metabolism in, 1716
  - surface landmarks, 7*f*
  - threat detection, emotion regulation and, 1819
  - working memory and, 430

- Prefrontal lobotomy. *See* Lobotomies
- Pregabalin (Lyrica)
  - acute mania, 1686
  - for anxiety disorders, 1747, 1748, 1781, 1782*t*, 4131
  - bipolar disorders
    - generalized anxiety disorder and, 1698
    - social anxiety disorder and, 1698
  - chemistry, 2966*f*, 2977
  - for chronic idiopathic pruritus, 2348–2349
  - classification, 2918
  - for diabetic neuropathy, 2236, 2236*t*
  - dosage/administration, 2978
  - drug interactions, 2978
  - drug studies, 2977
  - for fibromyalgia, 635
  - in functional GI disorders, 2206
  - GAD, 1783
  - history, 2976–2977
  - for neuropathies, 629
  - pain management using
    - in dying patients, 2292
  - in pain syndromes, 2396
  - pharmacological action, 2977
  - precautions/adverse reactions, 2977–2978
  - SAD, 1783, 1784
  - sexual dysfunction and, 1779*t*
  - therapeutic indications, 2977
- Pregenital stages, 4604
- Pregnancy
  - acamprosate and, 3058
  - α-adrenergic receptor agonists and, 2949
  - β-adrenergic receptor antagonists and, 2953
  - anticonvulsants and, 2969, 2971, 2973, 2975, 2976, 2978
  - barbiturates and, 2988
  - benzodiazepines and, 2999
  - calcium channel inhibitors and, 3020
  - CAM treatments and, 2543

- carbamazepine and, 3037
- DA receptor antagonists and, 3077
- essential oils and, 2555
- hypothyroidism in, 2241
- in jails and prisons, 4466
- lamotrigine and, 3086
- lithium and, 3095
- loss of, 2511
- MAOIs and, 3115–3116
- melatonin and, 3105
- mirtazapine and, 3109
- mood disorders in
  - neurosteroids and, 118
- MRS in, 283
- nefazodone and, 3130
- opioid receptor antagonist effects on, 3143, 3145
- pharmacotherapy during, 2519–2521
- prenatal diagnoses, 554
- psychiatric decision making, 2518
- risks, tobacco use disorder, 1351
- schizophrenia risk, complications with, 1449
- SSRIs and, 3167–3168
- trazodone and, 3252
- tricyclic and tetracyclic antidepressants and, 3263
- valproate and, 3276–3277
- vilazodone and, 3346

- Pregnenolone
  - fear and anxiety and, 1748
  - function of, 172
  - GABA and, 1747
  - PTSD and, 1819
- Prejudices
  - description of, 4604
  - racial, 701–702
- Premack Principle, 724
- Premature advice in patient interview, 960
- Premature birth and, 3538–3539

- Premature births, 3538–3539
- Premature deaths, 4288–4289
- Premature ejaculation disorder (PMDD), 4604
- Premature infants. *See also* Preterm infants
  - neurodevelopmental outcomes, 642–643
  - primary central sleep apnea of, 2097
  - psychiatric aspects of, 639–640
  - sleep in, 342
- Premenstrual dysphoric disorder (PMDD)
  - about, 1647, 1677
  - alprazolam for, 171
  - defined, 3674, 3688
  - treatment
    - hormone suppression, 3332–3333
    - nefazodone, 3126
    - SSRIs, 3154
    - venlafaxine, 3149
- Premenstrual syndrome (PMS), 2507, 3012
- Premenstrual tension, 1632
- Premorbid, definition of, 4604
- Premorbid functioning, 983–984
- Premorbid personality, 546
- Prenatal diagnoses
  - for HD, 554
  - HIV infections, 565
- Prenatal factors
  - alcohol and, 1268
  - autism spectrum disorder and, 3575
  - environmental
    - fetal abuse and, 3838
    - specific learning disorder and, 3526
  - risk, aggression and, 2484–2485
  - schizophrenia risk, parental, 1449
  - stress, addictive disorders and, 2046–2047
- Preoccupation of thought, 4604
- Preoperational period, 3359
- Preoptic anterior hypothalamus (POAH), 2340

- Preoptic area, 35
- Preschool ADHD Treatment Study (PATS), 3355, 3454, 3789
- Preschool Age Psychiatric Assessment (PAPA), 821, 3411
- Preschool-aged children. *See also* Childhood; Children; Toddlers
  - ADHD treatment for, 3789
  - hospitalizations, response to, 3848–3849
  - illnesses, response to, 3848–3849
  - psychiatric evaluations of, **3409–3413**
    - clinical formulation, 3412
    - components of, 3410–3412
    - developmentally appropriate criteria for, 3409–3410
    - informants for, 3410–3411
    - multiaxial, 3411–3412
    - multidisciplinary approach to, 3411
    - multiple modes of assessment for, 3411
    - sessions for, 3410
    - treatment planning, 3412
  - psychopharmacology considerations, 3796
  - terrorism's impact on, 3915, 3916
- Preschool Psychopharmacology Working Group, 3796
- Prescribing patterns, research into, 4352
- Prescription Drug User Fee Act (PDUFA), 2931
- Prescription drugs. *See also specific medications*
  - epidemiology, 4113–4114
  - etiology, 4114–4115
  - opioid analgesics, 4114
  - over-the-counter medications and, 4113–4115
- Presenium, description of, 4604
- Present centered therapy (PCT)
  - hypnotherapy as alternative to, 2722
  - PTSD, 1823
- Present State Examination (PSE), 1844
  - psychiatric classification and, 1154, 1162
  - schizophrenia, 1409
- "Presentation bias," 2203–2204
- Preservation of thought, 1120
- President's Commission on the Causes and Prevention of Violence, 2471

- Pressure of speech
  - description of, 4604
  - as psychiatric symptomatology, 1120
- Pressured speech, 4604
- Presupplementary motor area, 1893
- Pretend mode, BPD and, 2896, 2896*t*, 2897
- Preterm infants. *See also* Premature infants
  - brain development in, 640
  - brain injury in, 640–642
    - psychopathology and, 644
  - cognitive outcomes, 643
  - emotional effects on, 643
  - nomenclature for, 639*t*
  - psychopathologic outcomes, 643
- Pretransplant psychosocial evaluation, 2360*t*
- Prevalence, 1263, 1264, 1265–1266, 1265*t*. *See also* Incidence
  - description of, 4604
  - in epidemiology, 811–812
  - psychiatric disorders in long-term care, 4222, 4223*t*
  - psychiatric disorders in young *vs.* older people, 3954*t*
- Prevalence estimates world-wide
  - lifetime, *4518f*
  - twelve-month, *4519–4520t*
- Prevalence rates, definition of, 4604
- *Prevention and Treatment of Mental Disorders* (Robinson), 4536
- Prevention in public health, 4303–4304
- Preventive interventions, 3893–3900
  - definitions, 3894–3895, 3895*f*
  - for depressive disorders, 3897
  - for early-onset psychosis, 3897–3898
  - economic issues, 3898–3899
  - for family adversity, 3895–3896
  - future directions, 3900
  - overview, 3893–3894
  - provision of services issues, 3898
  - research issues, 3899–3900
  - research on, 3895–3898, 3896*f*

- for substance abuse disorders, 3898
- training issues, 3899
- for youth externalizing behaviors, 3896–3897

- Preventive psychiatry, 4604
  - in public health, 4303
- Priapism
  - nefazodone and, 3128–3129
  - trazodone and, 3251
- Priest, J. Percy, 4292
- Primal fantasy, 2051–2052
- Primal father, 4604
- Primal senses, 4604
- Primary age-related tauopathy (PART), 3991
- Primary anxiety, 2676
- Primary auditory cortex
  - anatomy of, 8
  - disorders of, 452
- Primary autoimmune hypothyroidism, 680
- Primary care
  - barriers to, 4296
  - behavioral health care integration with, 4320
  - in childhood intervention, 4300
  - depression treatment in, 4349–4350
  - as frontline, 4302
  - hospitalists *vs.,* 4357–4358
  - mental disorders in, 2183
  - mental disorders world-wide in, 4518
  - mental health care in, 4289, 4295–4296, 4295*f*, 4381, 4524–4525
  - in Patient-Centered Medical Home, 4307
  - psychiatry in, 4304, 4305, 4306
  - psychotropic medications, reliance on, 4321
  - suicide prevention, mood disorders and, 1675–1676
- Primary Care Evaluation of Mental Disorders (PRIME-MD), 2685
- Primary care physicians, 4604
- Primary central sleep apnea, 2096–2097
- Primary diagnoses, description of, 4604
- Primary gain, description of, 4604

- Primary motor cortex
  - basal ganglia and, 30
  - dopaminergic innervation, 8
- Primary narcissism. *See* Narcissism
- Primary negative symptoms, 1418–1419, 1418*t*
- Primary parkinsonism. *See* Parkinsonism
- Primary prevention. *See* Preventive psychiatry
- Primary progressive aphasia (PPA), 549
- Primary reinforcers. *See* Reinforcement
- Primary retarded ejaculation, 4604
- Primary somatosensory cortex
  - anatomy of, 8
  - function of, 8
- Primary visual cortex
  - connections, 7
  - surface landmarks, 7*f*
- Primates
  - prefrontal cortex neurons, 17*f*
  - striosomes in, 29
- Priming, memory, 735, 736, 737f
- Primitive cognitive processes, 2869
- Primitive idealization, 4604
- Primitive object-related phantasies, 2869, 2875
- Prince, Morton
  - *The Dissociation of a Personality,* 4537
  - dissociative disorders, 1868, 1895
    - media and, 1873
  - identity-state dissociation neurobiology, 1893
  - work of, 4604
- Prince, Raymond, 2445
- Prince, Virginia, 2023
- *Principles of Biomedical Ethics* (Beauchamp and Childress), 4240–4241, 4443
- Prion diseases
  - acquired, 601–602
  - clinical features of, 608–612
  - dementia in, 1220–1221
  - description of, 4604

- diagnosis, 608–612
- epidemiology of, 603
- genetic factors, 4022
- histology of, 608*f*
- history of literature on, 602–603
- human and animal, 601*t*
- neuropsychiatric aspects of, 600–619
- pathogenesis, 603–608
- universal precautions, 617, 617*f*

- Prion proteins (PrP)
  - cell (PrP^C), 603, 605*f*
    - function of, 606
    - genetics of, 607
    - host protein interactions, 607
    - neuropathology of, 607–608
    - structure, 605–606, 606*t*
  - chromosome location, 602*t*
  - CNS infection, 596–597
  - deposition in brain, 609*f*
  - forms of, 604*t*
  - infectiousness, 618
  - phylogenetic tree, 606*f*
  - plaques, 608*f*
  - structure of, 601*f*
  - transmissibility, 618
  - transport of, 604*f*
- Prions
  - brain invasion, 603
  - cellular, infectiousness, 618
  - molecular biology of, 603–608
  - strains of, 606
- Prison neurosis. *See* Chronophobia
- Prison populations
  - neuropsychiatric disorders in, 2489
  - psychiatric hospital population inversely related to, 4388–4389
- Prison psychosis, 1859
- Prisons. *See also* Correctional psychiatry; Correctional systems

- gambling in, 1802, 1811
- neuropsychiatric disorders in, 2489–2490
- Pristanic acid, accumulation of, 663
- Pristiq. *See* Desvenlafaxine succinate (Pristiq)
- Pritchard, James Cowles, 4535
- PRITE. *See* Psychiatry Resident-in-Training Examination
- Privacy
  - in correctional psychiatry, 4455
  - definition of, 4604
  - and electronic technology, 4280
- Privacy of patient interview, 945–956
- Private ownership, 4293
- Privilege
  - confidentiality and, 2670
  - definition of, 2021–2022, 2670, 4604–4605
  - privacy of patient, 4435–4436
- PRN medications, restrictions on, 4339
- *PRNP* gene, 607
- PRO. *See* Peer Review Organization (PRO)
- Probability, definition of, 4605
- Probability density functions, 830
- Probability theory, 830
- Probably efficacious treatments, 2704–2705
- Probe questions, 2734
- Probing in patient interview, 959
- Probiotics, 2547
- Problem Areas in Diabetes (PAID) instrument, 2231
- Problem gambling, 1802. *See also* Gambling disorders
- Problem-oriented records, 4605
- Problem-solving
  - definition of, 4605
  - family role, 1461, 4194
  - in group therapy, 4202
  - steps in, 2327–2328
  - stress management and, 2327–2328
- Problem-solving hypnotherapy, 2715
- Problem-solving therapy (PST), 2696, 2697

- Problem Symptom List, 2268
- Procarbazine (Procanbid)
  - side effects
    - delirium, 2255
    - depression, 2260
- Procardia. *See* Nifedipine
- Procedural learning deficits, 2139
- Procedural memory, 2129
  - associative learning and, 2139
  - dissociative amnesia and, 1879
- Procedural pain in burn injuries, 2376
- Process-oriented treatment, 3542–3543
- Process schizophrenia, 4605
- Processing assessment
  - for cognitive evaluation, 3415, 3417*t*, 3426–3429
- Processing speed
  - neurocognitive impairment, 1503
  - schizophrenia, 1503
- Prochlorperazine (Compazine)
  - for migraine, 625
  - side effects, 2271
- Procyclidine (Kemadrin)
  - chemistry, 2957, 2957*f*
  - dosage/administration, 2960
  - pharmacodynamics, 2958
  - therapeutic indications, 2959
- *PRODH* mutation, 663
- Prodromal phase, 4365–4366, *4366t*
- Prodromes
  - definition of, 4605
  - ethical treatment of, 3875
  - neurocognitive impairment, 1505–1506
  - schizophrenia, 1505–1506, 1569–1571
- Professional Standards Review Organization (PSRO), 4605
- Professionalism, physician mental health and, 2602
- Profits from psychiatric hospitals, 4293
- Profound mental retardation. *See* Mental retardation

- Progesterone
  - fluctuations, 383, 383*t*
  - function of, **171–172**
  - immunological effects, 186*t*
  - menstrual physiology and, 2507
  - metabolites, 171
  - sexual function and, 1956
- Progestin
  - HRT, 2523
- Prognosis, 3542
- Prognosis, definition of, 4605
- Programmed cell death. *See also* Apoptosis
  - nervous system development and, 99
  - neurogenesis and, 42
- Programmed death-1 (PD-1), 183
- Progress in Treatment Questionnaire (PITQ), 1919–1920
- Progress notes
  - addressing goals, 4338
  - informative, 4339
- Progressive deep muscle relaxation (PDMR), 2689
- Progressive multifocal leukoencephalopathy (PML)
  - diagnosis of, 586
  - white matter infection in, 596
- Progressive muscle relaxation
  - behavior therapy, 2689
  - case study, 2690
  - somatic therapy, 2557
- Progressive nonfluent aphasia (PNFA)
  - frontotemporal dementia variant, 1211
  - FTD and, 550
- Progressive polyradiculopathy, 632
- Progressive supranuclear palsy (PSP)
  - course/prognosis, 549
  - dementia and, 1213–1214
  - differential diagnoses, 549
  - features, 549
  - parasomnia and, 2102

- pathology, 549
- treatment, 549

- *Project for a Scientific Psychology* (Freud), 848–849
- Project Liberty, 3920–3921
- Projection
  - definition of, 4605
  - in young children, 2870
- Projection neurons, 3
- Projective identification
  - countertransference and, 2653, 2654
  - definition of, 4605
- Projective personality tests, 1002, 1004–1006
  - Rorschach test, 1002, 1004
  - sentence completion test, 1006
  - Thematic Apperception Test, 1004–1006, 1005*t*
- Projective techniques
  - cognitive test comparison with, 3427
  - during interviews, 3383–3384
  - for personality tests, 3404
  - for psychological testing, 3413
- Projective tests, definition of, 4605
- Prolactin (PRL)
  - after TBI, 525
  - definition of, 4605
  - depression and, 1716
  - functions of, 172, 2239
  - immunological effects, 186*t*
  - laboratory evaluation, 1014
  - levels of
    - antipsychotic drugs and, 62
  - production of, 2239
  - response to fenfluramine
    - aggression and, 2486
  - secretion of, 172
  - sexual function and, 1956
- Prolactin-inhibiting factor (PIF)
  - function of, 2239

- production of, 2239
- Prolactinomas
  - hyperprolactinemia and, 2249
  - presentation of, 678
- Prolixin. *See* Fluphenazine
- Prolonged exposures (PEs)
  - PTSD and, 1823, 1824, 1919–1920
- Promazine (Sparine), 3061, 3062*f*
- Promethazine (Phenergan)
  - agitation, 1525
  - chemistry, 2981*f*
  - laboratory interferences, 2983
  - therapeutic indications, 2982
- Promotion, in positive psychology, 2845
- Pronator tests, 457
- Proopiomelanocortin (POMC)
  - anxiety and, 1740
  - sequences in, 164
  - studies of, 168
- Properties, 1266
- Property crimes rates, **1993–2013,** 2474*f*
- Propionic acid, 663
- Propionic acidemia (PA), 663
- Propionyl CoA carboxylase deficiency, 663
- Propofol (Diprivan)
  - anesthesia induced by, 2986, 2987, 2990
  - ECT anesthesia with, 2986, 2987, 3292
  - pharmacological actions, 2985
  - precautions/adverse effects, 2987–2988
  - therapeutic indications, 2987
- Proportionality assumption, 844
- Propranolol (Inderal), 1388, 1742, 1743*t*, 1784. *See also* Beta blockers
  - absorption, 2951
  - for aggressive behavior, 2501
    - intermittent explosive disorder, 2501
    - mentally retarded children, 2500
    - TBI, 2502

- anxiety disorders
  - GAD, 1783
  - performance anxiety, 1781
  - social anxiety disorder, 1742, 1743*t*
- BD and social anxiety disorder, 1698
- chemistry, 2951, 2951*f*
- classification, 2918
- dissociative identity disorder, 1928
- distribution, 2952
- dosing/administration, 2955–2956
- drug interactions, 2955, 3131
- drug studies, 2952
- in hyperthyroidism
  - symptom relief using, 2243
- for migraine prevention, 623
- organ/system effects of, 2952–2953, 2953*t*
- pharmacodynamics, 2952
- pharmacokinetics, 2951–2952, 2952*t*
- pharmacological actions, 2951–2952, 2952*t*
- precautions/adverse reactions, 2955
- prolactin levels and, 2249
- for PTSD prevention of, 2379
- SAD, 1784
- side effects of, 571
- for social anxiety, 67
- teratogenic effects of, 2953
- therapeutic indications, 2953–2955, 2953*t*

- Proprioception. *See* Kinesthetic sense
- Propriospinal myoclonus, 2104
- Propulsid. *See* Cisapride
- Prosencephalon (forebrain)
  - embryology of, 41
  - in human development, 8
- ProSom. *See* Estazolam
- Prosopagnosia
  - and consciousness, 691
  - definition of, 4605

- description of, 333, 463
- Prospective studies, definition of, 4605
- Prostate cancer
  - incidence of, 2251
  - psychological symptoms, 2268
- Prostate-specific antigen (PSA), 2251, 2268–2269, 2277
- Prostatectomy, ejaculation and, 1971
- Prostitutes, dissociative disorders, 1889
- Protection, epidemiology, 812–813
- Protection of Persons with Mental Illness and the Improvement of Mental Health Care, 4530
- Protein chips, 162
- Protein disulfide isomerase (PDI), 107–108
- Protein kinase A (PKA)
  - cAMP-activated, 2915
  - function of, 121–122, 121*f*
- Protein kinase C (PKC)
  - activation of, 98
  - plasma membrane relocation to, 122
- Protein-protein interactions (PPIs), 163
- Protein synthesis inhibitors, 722
- Protein(s)
  - folding, levels of, 151–153
  - misfolding, 3988, 3991
  - regulation of, 161–162
- Proteolipid protein (PLP), 666
- Proteome
  - description of, 147
  - regulation of, 161–162
  - regulatory mechanisms, 162*t*
  - science of, 162
- Proteus syndrome, 314
- Proton spectroscopy, 526*f*
- “Protoself,” 331
- Prototaxic, 4605. *See also* Primary process
- Protozoa, CNS infections, 597
- Protriptyline (Vivactil)
  - chemistry, 3254, 3254*f*

- pharmacological actions, 3254
- sleep disorders, 3936

- Proust, Marcel, 2578
- Providers
  - ethnocultural populations impact of, 4382–4383
  - managed care rate decreases for, 4313–4314
  - patients' views of, 4384
- Provigil. *See* Modafinil
- Provision of care, ethics of, 3876
- Provisional tic disorder, 3637, 3651
- Prozac. *See* Fluoxetine
- Pruritus
  - cutaneous nervous system and, 2342
  - pathophysiology of, 2348
  - physiological aspects of, 2347–2348
  - psychosomatic aspects of, 2347
  - scratching during sleep, 2340
  - treatment of
    - medications used in, 2355*t*
- *PS-1* gene mutations, 312
- *PS-2* gene mutations, 312
- Pseudo-unipolar mood disorder, 1642
- Pseudoaggression, 4605
- Pseudoauthenticity, 4605
- Pseudobulbar affect (PBA)
  - after TBI, 533, 534*f*
  - characteristics, 628
- Pseudobulbar palsy, 464
- Pseudocollusion, 4605
- Pseudocrisis, 2757
- Pseudocyesis, 2511, 4605
- Pseudodementia
  - definition of, 4605
  - depressive
    - mood disorders and, 1657
  - depressive syndrome and, 1636
  - EEGs in, 251

- Pseudodepressed personality syndrome, 530
- Pseudogenes
  - ncRNA, 152
- Pseudohallucinations, 1411
- Pseudohermaphrodites, 1992, 2023
- Pseudoinsomnia, 2085
- Pseudologia fantastica
  - comorbidity, 1864
  - definition of, 1847, 4605
  - epidemiology, 1850
  - factitious disorder and, 1850, 1851
  - Munchausen syndrome and, 1857
- Pseudomemories
  - about, 1935
  - hypnotic, 2732
- Pseudomentalizing, BPD and, 2897
- Pseudomotor agitation, 4605
- Pseudomutuality, 2749
- Pseudoparkinsonism, 4605
- Pseudoseizures
  - factitious disorder imposed on another and, 1850
  - malingering and, 2404–2405
- Pseudothalamic sensory syndrome, 459
- Psilocin. *See* Psilocybin
- Psilocybin
  - alcoholism, 2891, 2892
  - cancer-related distress, 2890–2891, 2893–2894
  - definition of, 4605
  - depersonalization/derealization and, 1892
  - pharmacology of, 1316
  - psychological toxicity, 2890–2891
  - psychotherapy and, 2887
  - spirituality and, 2581
- Psironolactone, 2249
- Psoriasis
  - clinical features of, 2343
  - epidemiology of, 2343

- pathophysiology of, 2343
- presentation of, 2342–2343, 2343*f*
- psychopathology, 2344
- stigmatization of patients with, 2335
- treatment of, 2344
- trigger factors, 2343

- PSP. *See* Progressive supranuclear palsy (PSP)
- PSRO. *See* Professional Standards Review Organization (PSRO)
- Psychalgia. *See* Psychogenic pain disorder
- Psychasthenia
  - (Pt) Scale, MMPI, 999*t*
- Psychasthenia, 4605
- Psyche
  - assessment of, 2140–2141
  - definition of, 2139, 4605
  - development of, 2140–2141
  - identity and, 2139
  - psychobiology of, 2139–2140
- Psychedelic, definition of, 4605
- Psychedelic therapy
  - methods, 2888
  - techniques and goals, 2888–2889
- Psychedelic-treatment model. *See also* Hallucinogen-assisted psychotherapy
  - psilocybin, 2892
- Psychedelics. *See also* Hallucinogens
  - spirituality and, 2581
  - terminal illness, 2892–2893
- Psychiatric assessments, 3971–3977. *See also* Child psychiatric evaluations; Psychiatric evaluations
  - alcohol and substance abuse, 3973–3974
  - behavior, 3974
  - cognitive testing, 3975
  - competence, 3976
  - daily functioning, 3975–3976
  - depression, 3976
  - history of present illness, 3972
  - history taking, 3971–3972
  - insight, 3975

- language and thought disorder, 3974
- medical history, 3973
- mental status examination, 3974
- mood and affect, 3974–3975
- neurological evaluation, 3976
- perception, 3974
- personal history, 3973
- psychiatric history, 3972–3973
- rating scales, 3976
- sensorium and cognition, 3975
- sleep, 3975
- suicide and homicidal ideation, 3975
- syndromes specific to elderly, 3976

- Psychiatric comorbidities and, 3539
- Psychiatric conditions
  - brain tumors and, 496
  - childhood epilepsy and, 648–649
  - epigenetics and, 386
  - in survivors of brain tumors, 504–505
- Psychiatric consultations
  - medical students, 2606
  - suicide inquiry, 2619–2620
- Psychiatric disease/disorders.
  See also Mental illness; *specific conditions*
  - alcohol-induced, 1269–1270, 1272–1273
  - alcohol-use disorders and, 1278
  - CAM treatments and, 2542–2543
  - childhood schizophrenia
    - premorbid indicators for, 1452
  - classes of symptoms in, 320*t*
  - clinical manifestations of, **1114–1150**
    - anxiety disorders, 1138–1140
    - behavior, disturbances of, 1142–1144
    - challenges for, 1114–1117
    - consciousness, disturbances of, 1128–1132
    - drug-induced movement disorders, 1148
    - interpersonal relationship, disturbances of, 1144–1146
    - judgment, disturbances of, 1123–1124

- language disorders, 1135
- memory, disturbances of, 1133–1135
- mood, disturbances of, 1136–1138
- motor activity, disturbances of, 1146–1147
- motor disturbances and movement disorders, 1147–1149
- overview, 1114
- perception, disturbances of, 1125–1128
- schizophrenia motor disturbances, 1148
- self, disturbances of, 1132–1133
- sign and symptom features of, 1117–1118
- somatic, disturbances of, 1140–1142
- thinking, disturbances of, 1118–1123
- will disorders, 1135–1136

- comorbid, 4302
- and criminal responsibility, 4439
- cultural variation in, 4378
- enuresis, 3652, 3666
- environment and, 2529–2530
- genetic counseling, 2525, 2526*f*, 2527
- genetics and, 2529
- genome, environment and, 2527–2528, 2528*f*
- HIV infection comorbid with, 4343
- medical illness concurrent with
  - FGAs, 3072
  - psychopharmacology and, 2920–2921
- memory impairments, 738–739
- pediatric (*See* Child psychiatry)
- self-organizational processes, 694
- sensation, perception, and cognition in (*See specific disorders*)
- severity as barrier, 4304
- sleep deprivation and, 353–354
- temperament and, 3484–3488, 3485*t*
- treatment
  - anticonvulsants, 2974
  - ECT, 3287
  - lamotrigine, 3084
  - lithium, 3093, 3093*t*

    - valproate, 3276
    - venlafaxine, 3149
- Psychiatric education, **4402–4432**
  - board certification of psychologists and other disciplines, 4428–4429
  - certification examinations, 4414–4424
    - anxiety, 4423
    - case presentation, 4418, 4418*t*
    - clinical skills, 4416
    - cognitive, 4415–4416
    - content problems, 4421–4422
    - diagnosis and treatment, 4418
    - differential diagnoses, 4421–4422
    - evaluating success of, 4429–4432
    - factual information inadequacies, 4421
    - fatal mistakes, 4424
    - formulation, 4422
    - further evaluation, 4422
    - inadequate control, 4423
    - interview inadequacies, 4423
    - major types, 4415–4418
    - medical care processes and, 4431
    - objective structured clinical examination, 4429
    - oral, 4416
    - organization issues, 4423–4424
    - patient interview, 4416–4417
    - patient interview, guidelines, 4417*t*
    - patient interview, time-allotment, 4417*t*
    - preparing for, 4424
    - process problems, 4423
    - Psychiatrists In-Practice Examination, 4429
    - psychiatry recertification and, 4428
    - question development for, 4418–4420
    - rapport inadequacies, 4423
    - relationship with examiner, 4424
    - standardized patients, 4429
    - subspecialty, 4427, 4430
    - success and failure on, 4421–4424

    - success of, 4430–4431
    - test construction, 4419–4420
    - treatment recommendations, 4422–4423
    - vignettes in, 4420–4421
    - written, surviving, 4421
  - evaluating success of examinees, 4429–4432
  - examiners, 4424
  - grading process, 4424, 4425–4426*t*, 4427
  - graduate, 4402–4414
  - outcomes and, 4431–4432
  - Psychiatry Residency In-Training Examination, 4428
  - recertification, 4427–4428, 4430
  - 360-degree evaluations, 4429
- Psychiatric education, graduate, 4402–4414
- Psychiatric emergencies. *See also* Psychiatric emergency care
  - definition of, 4605
- Psychiatric emergency care, **2610–2637**
  - alcohol intoxication, 2629–2630
  - alcohol withdrawal, 2632–2633
  - benzodiazepine intoxication, 2630
  - benzodiazepine withdrawal, 2632–2633
  - cannabis intoxication, 2631
  - child abuse, 2634
  - cocaine intoxication, 2633
  - cocaine withdrawal, 2633
  - domestic abuse, 2633–2634
  - elder abuse, 2634
  - first priority, 2623
  - hallucinogen intoxication, 2631
  - high utilizers of service, 2635
  - history, 2623
  - homelessness, 2635
  - hypnotic intoxication, 2630
  - hypnotic withdrawal, 2632–2633
  - inhalant use, 2632
  - introduction, 2622–2623
  - legal issues

        - capacity, 2636
        - competency, 2636
        - confidentiality, 2635–2636
        - duty to warn, 2636
        - informed consent, 2636
        - involuntary commitment, 2636–2637
    - opioid intoxication, 2630
    - phencyclidine intoxication, 2631–2632
    - psychosocial issues, 2635
    - rape, 2634–2635
    - second priority, 2623–2624
    - sedative intoxication, 2630
    - sedative withdrawal, 2632–2633
    - stimulant intoxication, 2629
    - suicide, 2610–2618
    - suicide treatment, 2618–2622
    - third priority, 2624–2625
    - treatment principles
        - aggression, 2628–2629
        - agitation, 2628–2629, 2628*t*
        - difficult patients, 2627
        - overview, 2625
        - pharmacotherapy, 2627–2628
        - psychotherapy, 2625–2627
- Psychiatric epidemiology, 4547
- Psychiatric evaluations, child, **3375–3447**
- Psychiatric first aid by police, 4390
- Psychiatric histories. *See also* Geriatric psychiatry; Patient histories
    - as child psychiatric evaluation consideration, 3379
- Psychiatric hospitals
    - bed shortages (*See also* Beds)
        - effects of, 4325
        - third-party payers impact on, 4329–4330
    - built since January 2000, 4327*t*
    - as a business, 4329
    - care integration in, 4346
    - during colonial period/1800s, 4290, 4291

- as custodial, 4291
- forensic patients in, 4344
- hospice care in, 4343
- IMD exclusion impact on, 4336
- insurance coverage for, 4329
- outpatient services, 4325
- phasing out, 4292
- populations, 4388–4389
- private, 4327
- profits from, 4293
- prototype for, 4328*f*
- public, 4327
- role of
  - defining, 4329–4337
  - Federal government/court actions defining, *4333*–*4334t*, 4336
  - fulfilling, 4337–4341
  - other roles, 4344–4345
  - overview/history of, 4325–4329
  - special populations, 4341–4344
  - summary/future directions, 4345–4346
- 21st century changes to, 4330
- in the 20th century, 4291

- Psychiatric illnesses
  - after TBI, 527
  - applications of MRS in, 280–282
  - applied EEG in, 247–248
  - epidemiology of, 306
  - epigenetic mechanisms, 398–401
  - fMRI applications in, 276–280
  - immune modulators of, 202–203
  - phenotypes, 310–311, 310*f*
  - during pregnancy, 2518–2519
    - impact on mothers, 2518
    - impact on offspring, 2518
  - stress and, 2322–2324
    - historical concepts, 2322
- Psychiatric nurses, 4605

- Psychiatric professionals in public health, 4303
- Psychiatric social workers, 4605
- Psychiatric symptoms
  - assessment of, 1115–1116
  - circumstances and, 1117
  - classifications of, 1118
  - cultural context of, 1117–1118
  - impairment and adversities for, 1118
  - opioid receptor antagonist effects on, 3144
- Psychiatrist(s). *See also* Physicians
  - absence in public psychiatry, 4296
  - clinical *versus* administrative, 2594–2595, 2595*f*
  - in collaborative care, 4322
  - definition of, 4605
  - elderly impaired, 4238
  - ethical practice, 4448
    - confidentiality, 4446, 4446*t*
    - informed consent, 4446–4447, 4447*t*
    - overlapping roles, 4447
    - professional boundaries, 4445–4446
    - psychiatric opinions, 4447–4448
  - genetic counseling and, 2525
  - geographic distribution of, 4306
  - as hospital superintendents, 4287, 4291
  - Medicaid reimbursement rates, 4380
  - Medicare/Medicaid acceptance rates, 4317
  - in multidisciplinary care teams, 4289
  - numbers world-wide, 4525
  - in private practice, 4292
  - public agencies and, 4305–4307
  - in public psychiatry, 4306
  - in publicly funded services, 4288
  - religious perceptions, 2585–2586
  - roles of
    - in adoption, 3826–3827
    - in foster care system, 3822–3823
    - in HIV/AIDS treatment, 3855–3857

    - in long-term care, 4222
    - in schizophrenia treatment, 1555–1557
    - in school consultations, 3882–3883
  - sexual orientation of, 2005–2006
  - spirituality perceptions, 2585–2586
  - 21st Century profile of, 4324*t*
  - as team participant-leader and mentor, 4300–4301
  - theories of, 4291
  - torture by, 2428
- Psychiatrists In-Practice Examination (PIPE), 4429
- Psychiatry
  - biophysically-based circuit modeling, 428–431
  - child (*See also* Child psychiatric evaluations)
    - computational modeling approaches, 427–437, 428*f*, 429*f*
    - definition of, 4605
    - development (normal), 3351–3374, **3351–3374** (*See also* Development, adolescent; Development, child)
    - disorders, 3348–3350
    - DRGs applied to, 4336
    - epidemiology, 1731
    - evaluations, 3398–3399
    - experimental approaches to, 428*f*
    - future challenges/opportunities, 4323–4324
    - genetic counseling, 2525, 2526*f*, 2527
    - genetics and, 3350, 3448–3457
    - in illness management model, 4303
    - multidisciplinary teams, 4300–4301
    - suicide prevention, mood disorders and, 1674–1675
- Psychiatry Residency Review Committee's (PRRC), 4402
- Psychiatry Resident-in-Training Examination (PRITE), 975, 4428
- Psychiatry–ethics interface, 2278
- Psychic auras in epilepsy, 510*t*
- Psychic determinations, 4605
- Psychic determinism, 2660
- *Psychic Disturbances of Childhood* (Emminghaus), 4536
- Psychic energizers, 4605
- Psychic energy (Freud), 849–851
- Psychic equivalence mode, BPD and, 2896, 2896*t*, 2897

- Psychic helplessness. *See* Helplessness
- Psychic trauma, definition of, 4606
- Psycho-oncology, **2250–2287**
  - areas of assessment in, 2253
  - definition of, 4606
  - general principles, 2252–2255
  - geriatric, 2282–2285
  - historical background, 2251
  - hypnosis and, 2730
  - pediatric, 2281–2282
  - staff responses
    - clinical spectrum of, 2280
  - staff stress and, 2280
  - treatment modalities, 2253–2255
- Psychoactive substance abuse. *See* Substance use disorders (SUDs)
- Psychoactive substance dependence, 4606
- Psychoactive substance-induced mental disorder, 4606
- Psychoanalysis, **2638–2676**
  - abstinence in, 2660
  - for adolescents, 3813–3814
  - for aggressive behavior, 2498
  - boundary problems, 2641
  - case study, 2656, 2661
  - cathartic method, 2639
  - CBT and, 2881–2882
  - classical, **848–898,** 2640–2641
    - analysis *vs.* analytic psychotherapy, 890
    - analytic process, 891
    - attachment theory, 886–887
    - basic assumptions, 860–861
    - concentration method, 853
    - condensation, 859
    - conscious, 861
    - countertransference, 891–893
    - current considerations, 882–883
    - displacement, 858–859
    - evolving concepts, 887–888

- free association, 853–854
- from hypnosis to analysis, 853
- hysteria studies, 852–853
- infantile sexuality, 873–874, 875*t*–876*t*
- instinctual theory, 864
  - concepts of, 862–863
  - life and death, 864–865
- interpretation of dreams, 856–860
- love objects, 865
- mental functioning, 862
- narcissism and dual instinct theory, 865
- neuropsychoanalysis, 889–890
- object relations theory, 883–886
- oral phase and objects, 874–877
- phallic phase and objects, 877–878
- preconscious, 861
- pregenital phases, 874
- prepsychoanalytic thinking, 848–851, 850*t*
- primary and secondary process, 850*t*, 860
- projection, 859
- psychoanalysis, beginnings of, 851–856
- psychology of self, 887–890
- psychosexual phases and object relations, 873–883
- real relation, 895
- regression, 859–860
- relational and intersubjective approaches, 888–889
- representability, 858
- repressed *vs.* repressing, 862
- results of treatment, 897
- revision, 859
- symbolism, 858
- technical aspects, 895–897
- therapeutic alliance, 893–895
- topographic theory, 860–861
- transference, 891, 892*t*–893*t*, 894*t*
- unconscious, 861–862

- concentration method, 2640

- contemporary trends
  - conflict/deficit duality, 2646–2647
  - intersubjectivity, 2646
  - relational psychoanalysis, 2646
- contraindications, 2657–2658
- couch in, 2659–2660, 2663, 2669
- definition of, 3735, 4606
- depression, 1661–1662
  - synthesis, 1663
- dynamic perspective, 2649
- dysthymic disorder, 1662–1663
- economic perspective, 2649
- evenly suspended attention in, 2660
- free association, 2640
- fundamental rule of, 2660
- gambling disorder, 1805–1806, 1808
- gay and lesbian patients, 2009–2010
- genetic perspective, 2649
- goals, 2658–2659
- Grob on, 4291
- hallucinogen-assisted psychotherapy and, 2888
- hermeneutics and, 2649–2650
- history of
  - development of, 2642*t*
  - early 20th century, 2641, 2643
  - Freud, 2639–2641
  - late 20th century, 2645–2646
  - mid-20th century, 2643–2645
  - 18th and early 19th centuries, 2639
- indications, 2657
- insight in, 2664
- instinctual theory
  - characteristics of, 863–864
- interpretation of dreams
  - infantile drive repression, 857–858
  - nocturnal sensory stimuli, 857
- interpretation of dreams, 856–860

    - day residues, 857
  - introduction and definition, 2638–2639
  - lay analysts in, 2661
  - limitations, 2661
  - main approach, 2659–2661
  - mania, 1663–1664
  - medications and, 2658
  - mirroring in, 2660
  - models, integration of, 2647–2648
  - mood disorders, future directions, 1667
  - patient requisites, 2658
  - process, 2656–2657
  - psychoanalysis, beginnings of
    - theoretical innovations, 855–856
  - psychoanalytic psychotherapy *versus,* 2662–2663, 2663*t*
  - psychoanalytically-oriented psychotherapy *versus,* 2876
  - psychosexual phases and object relations
    - epigenetic sequence (Erikson), 880–882
    - separation-individuation process (Mahler), 878–879
  - relational, 2646
  - research and evaluation, 2673–2675
  - resistance and, 2652–2653
  - setting, 2659–2660
  - short-term (brief) therapy, 2647
  - stages, 2657
  - structural perspective, 2649
  - structural theory and ego psychology, 865–873
    - psychic apparatus, 867–873
  - supportive psychotherapy and, 2666–2667
  - as 20th century change, 4328–4329
  - theoretical constructs in, 2649
  - topographical perspective, 2649
  - treatment fundamentals, 2650–2656
- *Psychoanalysis of Children* (Klein), 4539
- Psychoanalysis overview, 4552
- Psychoanalysts, 4606
- Psychoanalytic group psychotherapy, 4606

- Psychoanalytic psychotherapy, 2661–2669
  - adjunctive therapy in, 2664
  - anxiety disorders, 2676–2682
  - confidentiality, 2670–2672
    - future and, 2671–2672
    - HIPAA, 2671
    - patient-litigant exception to testimonial privilege, 2670
    - Tarasoff decisions, 2670–2671
    - writing about patients, 2671–2672
  - ethical issues, 2669
  - medications and, 2664
  - mutative change agents
    - empathy, 2664
    - identification, 2664
    - insight, 2663–2664
  - OCD, 2676–2682
  - psychoanalysis and, 2661–2662
  - psychoanalysis *versus,* 2662–2663, 2663*t*
  - recent developments and variants, 2668–2669
  - research and evaluation, 2673–2675
  - trauma-related disorders, 2676–2682
  - unconscious and, 2677–2678
  - value and value-free fallacy, 2669
- Psychoanalytic psychotherapy for panic disorders (PFPP), 2804–2805
- Psychoanalytic spectrum, 2737–2738
- Psychoanalytic theories
  - of dissociation, 1903
  - of paraphilias, 2016
- Psychobiological processes, 1895*t*
- Psychobiology
  - definition of, 1604, 4606
  - Meyer as advocate for, 908–909, 4291
  - theories, 908–909
- Psychocutaneous disorders, **2332–2357**
  - classification of, 2332, 2333*t*
  - DSM-5 changes, 2350*t*
  - psychopharmacology of, 2355

- treatment of
  - medications used in, 2355*t*
- *Psychodiagnostik* (Rorschach), 4537
- Psychodrama, **2855–2860**
  - definition of, 4606
  - Do–Undo–Redo, 2859
  - ethics, 2859
  - history, 2855–2856
  - methods, techniques, clinical issues, 2856–2857
  - research, 2859
  - social and organic unity of humanity and, 2859
  - structure of, 2859
  - theories, 2857–2858
    - Canon of Creativity, 2858
    - encounter, 2858
    - human development, 2858
    - role theory, 2858
    - social atom, 2857–2858
    - surplus reality, 2858
    - time, space, reality, and cosmos, 2858
- Psychodynamic psychopharmacology, 2681–2682
- Psychodynamic psychotherapy
  - American Interpersonal School, 2868, 2875–2876
  - CBT, psychosis and, 2882
  - couples group therapy and, 2758–2759
  - couples therapies and, 2759, 2760
  - depersonalization/derealization disorder, 1944
  - European Kleinian School, 2868, 2875
  - evaluation, 2863–2864
  - family and couples, 2754–2755
  - generalized anxiety disorder, 2679
    - case study, 2679
  - history, 2868–2869
  - integration for sexual addictions, 2059–2061
  - internalization for sexual addictions, 2061–2062
  - neurobiological effects of, 2814*t*
  - OCD, 2680–2681

    - case study, 2681
  - panic-focused, 2678–2679
  - psychosis
    - clinical vignette, 2867, 2885–2886
    - engagement phase, 2876
    - ethical issues, 2884–2885
    - future directions, 2885
    - identifying psychotic mechanisms, 2876, 2878–2879
    - interpretation, 2876, 2879–2881
    - introduction, 2866–2868
    - research and evaluation, 2885
    - therapeutic alliance, 2876–2878
  - PTSD, 2679–2680
    - case study, 2680
  - RCT evaluations, 1702
  - for sexual addictions, 2058–2059
  - structure of, 2060–2061
  - suicide risk, 2617
  - theoretical issues, 2869–2870, 2869*t*
- Psychodynamic theories and therapies
  - for adolescents, 3813–3814
  - on alcohol-use disorders, 1269
  - child, **3734–3741**
    - clinical considerations, 3738–3739
    - defined, 3734–3735
    - further directions, 3740–3741
    - history, 3734–3735
    - practical matters in, 3738
    - research and evaluation, 3740–3741
    - theoretical considerations, 3735–3738
    - treatment phases, 3739–3740
  - cognitive theory and, 1665
  - defined, 3735
  - on delusional disorder, 1584
  - in functional GI disorders, 2208
  - posttraumatic stress disorder, child/adolescent and, 3706
  - selective mutism, 3723

- stuttering, 3560
- tic disorders, 3636, 3650
- time-limited, 4175

- Psychodynamically-oriented clinicians, 1873
- Psychodynamic–psychoanalytic family and couples therapy, 2751*t*
- Psychodynamics, 4606
- Psychoeducation
  - anxiety disorders, 1778
    - CBT and, 1775
  - BDD, 1792
  - BIF patients, 2439
  - bipolar disorders, 1703, 1707
    - ADHD and, 1699
  - bipolar I disorder, 1696
  - and cognitive-behavior therapy, 4182, 4182*f*
  - cognitive therapy and, 2770
  - dissociative identity disorder
    - informed consent and, 1927
    - nocturnal PTSD triggers, 1928
  - dissociative psychosis, 1950
  - emotional regulation, 2772
  - excoriation disorder, 1797
  - of family, 4194
  - interpersonal psychotherapy and, 1701
  - mood disorders, 1612
  - OCD and, 1788, 1789
  - in psychosis, 4186–4187, 4187*f*
  - TTM, 1795
- Psychoeducational approaches
  - cognitive-behavioral therapy, 3753, 3754*t*
  - for dementia, 1200
  - depressive disorders, child/adolescent, 3680, 3694
  - tic disorders, 3644, 3658
- Psychoeducational family and couples therapy, 2752*t*
- Psychoendocrinology, 4606
- Psychogenic, definition of, 4606
- Psychogenic amnesia

  - definition of, 4606
  - DSM and ICD diagnostic criteria, 1874
- Psychogenic cough, 2228–2229
- Psychogenic depressions, 1603–1604
- Psychogenic dysphonia, 3366–3367
- Psychogenic fugue
  - definition of, 4606
  - DSM and ICD diagnostic criteria, 1874
- Psychogenic nonepileptic seizures (PNES)
  - in children, 650–651
  - description of, 514
  - diagnosis of, 651
- Psychogenic pain, 2292*t*
- Psychogenic pain disorders, 4606
- Psychogenic purpura, 2354–2355
- Psychogenic tremor, 458
- Psychohistory, definition of, 4606
- Psychoimmunology, 4606
- Psychokinesis, 4606
- Psycholinguistics, 4606
- Psychological abuse as maltreatment, 3829
- Psychological assessments (standard)
  - for children, **3403–3414**
    - interpretation of, 3405–3406
    - measurements for, 3413–3414
    - overview, 3413–3414
    - personality measures, 3404
    - process for, 3413
    - referrals for, 3403–3404
    - reporting, 3406
    - standardized tests for, 3404, 3405*t*
    - validity and reliability of, 3404–3405
  - referral questions, 1007
- *Psychological Birth of the Human Infant* (Mahler), 879, 4542
- Psychological changes, 3995–4002
  - adjustment and coping
    - age differences, 3997

    - role transitions, 3997–3998
  - and clinical practice, 4001–4002
  - cognitive aging
    - attentional processes, 3999
    - executive functions, 4000
    - language function, 3999
    - learning and memory, 3999–4000
    - models of, 4000–4001
    - processing speed, 4000
    - visuoperception, 3999
  - design and sampling issues, 3996
  - mental health and personality, 3996–3997
  - social functioning, 3998–3999
- Psychological debriefing, PTSD, 1824
- Psychological defense system. *See* Defense mechanisms
- Psychological factors affecting other medical conditions (PFAOMC), 1827, 4606
  - DSM-5 on, 1841
- Psychological first aid (PFA), 2570, 2571
- Psychological resilience, 2336
- Psychological sciences, **686–704**
  - attention
    - attention capacity, 690
    - selective attention, 690
    - sustained attention, 690
  - cognitive science, 686–687
    - emotion, 687
    - hot and cold cognition, 687
  - consciousness, 691–695
    - development of, 693–694
    - discourse and narrative, 693
    - disturbances of, 691
    - mental models and schemata, 691–692
    - metacognition and self-reflective capacity, 692–693
    - processing and laterality, 692
    - self-organizational processes, 694
    - social cognition, 693
    - states of mind, 694–695

    - thought and language, 692
  - cross-references, 703
  - energy and information, 688–689, 689t
  - future directions, 703
  - information processing, 689, 689f
  - memory systems, 690–691, 691t
  - mind and brain, 688
  - neural network growth and integration, 687–688, 688t
  - in psychiatric disorders, **695–703**
    - anxiety disorders, 699
    - attention-deficit/hyperactivity disorder, 696–697
    - autistic spectrum disorder, 697–698
    - mood disorders, 698–699
    - posttraumatic stress disorders, 699–702
    - schizophrenia, 695–696, 695f
  - sensation and perception, 690
- Psychological Screening Inventory
  - characteristics of, 998*t*
- Psychological syndromes
  - dying patients
    - prevalence of, 2297
    - treatment principles, 2297–2298
- Psychological testing
  - assessment instruments
    - characteristics of, 995
    - normative samples, 995
    - validity issues, 995–996
  - purpose of, 994–995
- Psychological treatment
  - for paraphilic disorders, 2021
- *Psychological Types* (Jung), 4540
- Psychologists
  - board certification of, 4428–4429
  - definition of, 4606
  - torture by, 2428
- Psychology, definition of, 4606
- *Psychology of Personal Constructs* (Kelly), 932

- Psychology of women, 4544
- *Psychology of Women* (Deutsch), 4539
- Psycholytic therapy
  - methods, 2888
  - techniques and goals, 2888
- Psychometric issues
  - base rates, 982–983
  - normative data, 983
  - predictive utility, 982
  - premorbid functioning, 983–984, 984*f*
  - reliability, 982
  - validity, 982
- Psychometry, definition of, 4606
- Psychomotor acceleration, 1640–1641
- Psychomotor agitation
  - definition of, 4606
  - depressive syndrome and, 1635
- Psychomotor epilepsy. *See* Complex partial seizures (CPS)
- Psychomotor function, 1309
  - impairment
    - barbiturates and, 2988
    - benzodiazepine, 2998
  - retardation
    - definition of, 4606
    - depressive syndrome and, 1635–1636, 1636*f*
      - bipolar I disorder and, 1650
    - subcortical dysfunction and, 1709
- Psychomotor seizures, 4606. *See also* Complex partial seizures
- Psychomotor stimulant theory, 372
- Psychomotor stimulants, 4606
- Psychoneuroendocrinology, **164–178**
  - developmental, 166
  - methodology in humans, 166–167
  - of skin diseases, 2336–2339
- Psychoneuroimmunology
  - definition of, 4606
  - stress and, 2318–2322, 2319*f*

- Psychoneurosis. *See* Neurosis
- *Psychopathia sexualis* (Kraft-Ebing), 4536
- Psychopathic Deviancy (Pd) Scale, MMPI, 999*t*
- Psychopathic personalities, 4606
- *Psychopathic Personalities* (Schneider), 4541
- Psychopathic personality syndrome, 530
- Psychopathological, definition of, 4606
- Psychopathology
  - of alcohol-use disorders, 1268–1269
  - biocultural links, 808
  - clinical applications of QEEG, 1034–1040
  - conduct disorder, 3609, 3613–3614
  - EEG in study of, 257–258
  - in intellectual disability, 3510–3512, 3510*t*
  - oppositional defiant disorder, 3609
  - positive psychology on optimal functioning *versus,* 2849
  - rating scales for, 3402
  - separation anxiety, 3709
  - temperament and, 3484–3485, 3485*t*, 3488
- Psychopathology theories
  - Adler, Alfred, 910
  - Alexander, Franz, 915–916
  - Berne, Eric, 926
  - Fromm, Erich, 921
  - humanistic approaches, 933–934
  - learning theory, 930–931
  - Meyer, Adolf, 909
  - Sullivan, Harry Stack, 924
- Psychopathy
  - aggression and, 2490–2491
  - characteristics of, 2490–2491
  - sociopathy *versus,* 2490–2491
  - violence and, 2423–2425
  - violent behavior and, 2490–2491
- Psychopharmacological treatment, 4339
- Psychopharmacology, **3787–3797**.
  See also Medication(s)
  - abuse potential and, 2918

- for Alzheimer disease, 4084–4085, 4084*t*
- augmentation and, 2923
- autism spectrum disorder and, 3585
- for binge eating disorder, 2081
- black box warning and, 2916*t*, 2918–2919
- brain stimulation *v.,* 3300–3301, 3301*t*
- for bulimia nervosa, 2081
- child/adolescent, **3787–3797**
  - ADHD, 3788–3789, 3788*t*, 3796
  - anxiety disorders, 3792–3793, 3796
  - assessments and, 3787
  - autism spectrum disorders, 3793–3794, 3796
  - bipolar disorder, 3791–3792, 3796
  - conduct disorder, 3790
  - developmental considerations, 3787–3788
  - disruptive behavioral disorders, 3788–3790, 3796
  - dysthymia, 3791
  - enuresis, 3794–3795
  - major depressive disorder, 3790–3791, 3796
  - mood disorders, 3790–3792
  - OCD, 3792
  - oppositional defiant disorder, 3789–3790
  - overview, 3787
  - for preschool-aged children, 3796
  - psychotropic medications, 3795–3796
  - rating scales and, 3787
  - role of, 3787
  - schizophrenia, 3793
  - sleep disorders, 3794, 3796
  - substance use disorders, 3793
  - tic disorders, 3795
- child psychiatry, 2924
- classification, 2918
- clinical, general issues, 2916–2918
  - classification, 2918
  - history, 2916–2918
- combination of drugs and, 2923–2924

- compliance, 2921, 2922, 2923
- definition, 2905
- diagnosis and, 2920
- dosing, 2921
- drug factors and, 2919–2920
  - mechanisms of action, 2919–2920
  - pharmacodynamics, 2919
  - pharmacokinetics, 2920
- drug preferences and, 2921
- drug selection and, 2920, 2921
- duration, 2921–2922
- in elderly, 2924
- family members and, 2920, 2921
- FDA and, 2918–2919
- future directions, 2925
- general principles, **2905–2926**
- generic drug names, 2923
- history, 2916–2918
- informed consent and, 2921
- intracellular signaling pathways, 2914–2915, 2915*f*
- in jails and prisons, 4458–4459
- mechanisms, 2919–2920
- medical/psychiatric disorders and, 2920–2921
- in medically ill patients, 2924–2925
- monitoring, 2921
- monotherapy, 2923
- neurotransmitter signal integration, 2916, 2916*f*, 2916*t*
- patient education and, 2921
- patient-related factors, 2920–2921
- pharmacodynamics, 2919
- pharmacokinetics, 2920
- pharmacological actions, 2919
- polypharmacy, 2923
- posttraumatic stress disorder, 3706–3707
- receptor pharmacology, 2905–2914
- regulatory issues
  - black box warning, 2916*t*, 2918–2919

        - product labeling or information, 2918
    - remission and, 2922
    - response to, 2922
    - schizophrenia
        - fMRI studies, 1485–1486
        - neuroimaging studies, 1491–1492, 1497–1499, 1497*f*, 1498*f*
    - selective mutism treatment, 3723
    - sensitization, 2922–2923
    - side effects, 2922
    - special populations, 2924–2925
    - substance abuse and, 2925
    - tic disorders, 3644–3648, 3658–3662
    - tolerance, 2922
    - treatment failures, 2922
    - treatment outcomes, 2920, 2922–2923
    - treatment resistance, 2922
    - withdrawal, 2923
- Psychophysics, flavor hedonics and, 358
- Psychophysiological insomnia, 2085
- Psychopolitics, 4606
- Psychose passionelle. *See* Erotomania
- Psychosexual development
    - in childhood, 3354–3357*t*, 3363–3364
    - definition of, 4606
    - and object relations, **873–883**
        - oral phase, 874–877
        - phases of development, 873–874
        - pregenital phases, 874
- Psychosexual disorder, definition of, 4607
- Psychosexual history at patient interview, 966
- Psychosexuality, factors in, 1958–1959
- Psychosis
    - acute, FGA treatment, 3060, 3068, 3069–3070, 3078–3082
    - acute and transient disorders, **1574–1580**
        - classification of, 1172
        - comparative nosology, 1574–1577, 1576*t*, 1577*t*
        - course and prognosis, 1580

- diagnosis, 1578–1579
- differential diagnoses, 1579–1580
- epidemiology, 1577
- etiology, 1577–1578
- history of, 1574, 1575*t*

- after stroke, 477
  - etiology of, 479–480
  - treatment of, 488
- after TBI, in children, 653
- behavioral interventions, 4187
- CBT
  - A-B-C-model, 2883
  - clinical vignette, 2867, 2885–2886
  - eliciting patient's story, 2882–2883
  - engaging the patient, 2882
  - ethical issues, 2884–2885
  - future directions, 2885
  - introducing coping mechanisms, 2883
  - introduction, 2866–2868
  - maintaining contact, 2884
  - normalizing model, 2883
  - overview, 2881–2882
  - reformulating life narrative, 2884
  - research and evaluation, 2885
  - therapeutic alliance, 2883–2884
- as clinical continuum, 1569–1573, 1570*f*
- cognitive interventions, 4187–4189, 4188*t*
- definition of, 1561–1562, 4607
- delusional disorder, **1582–1587**
- dementia and, 1196, 1198, 1201–1202
- diagnosis
  - B-SNIP results, 1513–1516, 1514*f*, 1515*f*
  - DSM-type alternatives, 1516–1518, 1517*f*, 1518*f*
  - future directions, 1518
  - phenotypes and, 1513
- emergency psychiatric evaluation for, 3444–3445
- endocannabinoids in, 114–115

- in epilepsy, 509
- European Kleinian School, 2869
- family relations and, 2872–2873
- gender differences, 4246
- genetic studies, 1564–1567
- hallucinogen-induced, 1314–1315
- hallucinogens and, 2890
- history, 1407
- HIV infection and
  - in adolescents, 646
- inappropriate diagnosis of, 4382
- induced with antiretroviral medications, 4343
- with intellectual disability, 3512
- interview scenarios, 960–961
- malingering of, 2406
- MAOIs and, 3118
- mood-incongruent, 1641
- niacin flushing response and, 2545
- nonaffective, 2516–2517
- other psychotic disorders, **1574–1598**
- in PD, 545, 548
- personality disorders and, 1571–1573
- phenotypes of, **1512–1519**
- posttraumatic stress disorder (PTSD), 3706
- pre-prodromal signs, 1569–1571
- prodromal signs, 1569–1571
- psychodynamic psychotherapy
  - clinical vignette, 2867, 2885–2886
  - ethical issues, 2884–2885
  - future directions, 2885
  - identifying psychotic mechanisms, 2878–2879
  - interpretation, 2879–2881
  - introduction, 2866–2868
  - research and evaluation, 2885
  - techniques, 2875–2876
  - therapeutic alliance, 2876–2878
- psychodynamic theory of, 2869–2870, 2869*t*

- psychoeducation, 4186–4187
- psychology of, 2870–2872, 2873*f*
- risk of, biomarkers for, 423
- schizoaffective disorder, **1587–1590**
- schizophrenia, defining dimension of, **1561–1574**
  - biological factors, 1564–1569
  - defined, 1561–1562
  - diagnostic approach, 1562, 1563*f*
  - subclassifications using, 1562–1564, 1563*f*
- schizophreniform disorder, **1580–1582**
- spectrum disorders, 1562–1564, 1563*f*
- spirituality and, 2587–2589
- substance/medication-induced psychotic disorders, **1590–1592**
- symptoms of, 413
- treatment
  - buspirone, 3012
  - FGA, 2936, 3060, 3068, 3069–3070, 3078–3082

- Psychosocial crises, 3446–3447
- Psychosocial deprivation, 4607
- Psychosocial development, 4607
- Psychosocial dwarfism, 4607
- Psychosocial environments, 2464
- Psychosocial factors
  - autism spectrum disorder, 3574
  - opioid-related disorders, 1360
  - positive psychiatry, 4282
  - in psychotherapy of elderly, **4164–4171**
    - definition, 4164–4165
    - economic and environmental factors, 4169
    - ethnocultural issues, 4169–4171
    - future directions, 4171
    - relationships, 4165–4168
    - role changes, 4168–4169
    - societal attitudes and responses, 4169
    - spiritual issues, 4171
  - for standardized instrument development, 3394
  - in substance abuse, 4116
- Psychosocial interventions
  - attachment disorders, 3666, 3680
  - schizophrenia, **1531–1538,** 1549
- Psychosocial stressors. *See* Stress
- Psychosocial therapies
  - for ADHD, 3596–3597
  - for adolescents, 3813–3814
  - brief psychotherapies as, 3742, 3745–3748
  - early-onset bipolar disorder, 3693, 3707
  - early-onset schizophrenia treatment, 3732–3733
  - medications integrated with, 4362
  - opioid-related disorders, 1370–1371
  - opioid treatment, 1370–1371
  - schizophrenia, 1530
  - stimulant use disorders, 1288–1289
  - suicide prevention, 3685, 3699
  - tobacco-related disorder treatment, 1348–1351

- Psychosocially determined short stature. *See* Psychosocial dwarfism
- Psychosomatic disorders, 2051
- Psychosomatic medicine, 915–916, **2177–2398,** 4553
  - access to, 2185–2186
  - ancient history of, 2177
  - certificate examinations, 4427
  - clinical practice of, 2181–2182
  - clinical problems in, 2181, 2181*t*
  - conceptual trends, 2180*t*
  - current trends in, 2181
  - evaluation process in, 2182
  - fellows in, 2185*f*
  - field of, 2183–2185
  - financial arrangements for, 2185–2186
  - history of research into, 2178–2180, 2180*t*
  - modern history of, 2177–2178
  - obesity, 2210–2211
  - patient care experience of, 2185*f*
  - physician ethics, 4443–4444
  - psycosocial interventions, 2182–2183
  - research in, 2183, 2184*t*
  - training, 4405, 4406*t*
  - treatments, 2182–2183
- *Psychosomatic Medicine* (Alexander), 914
- Psychostimulant drugs for intellectual disability, 3515
- Psychostimulants, **3219–3231**
  - ADHD, 3219, 3223, 3227–3229
  - for aggressive behavior, 2172
  - antidepressants combined with, 3324
  - apathy, 3229
  - in brain tumor patients, 500
  - for cancer patients, 2262, 2264
  - classification, 2918
  - conduct disorder, 3228
  - depression, treatment-resistant, 3229
  - dosage/administration, 3220*t*–3222*t*
  - drug interactions, 3227*t*

- drug reward and, 373–374
- EEG effects, 254
- for fatigue, 2284
- history, 3219
- for MDD, 500
- mechanisms of action, 3223–3224
- ODD, 3228
- pharmacological actions, 3224–3227, 3225*t*–3226*t*, 3227*t*
- precautions/adverse reactions, 3229–3231, 3229*t*, 3237*t*
- seizure disorder, 3229
- therapeutic indications, 3227–3229
- in transplant patients, 2372*t*

- *Psychotherapeia, or the remedial influence of mind* (Dendy), 4535
- Psychotherapeutic treatments
  - drugs, 4607
  - posttraumatic stress disorder, 3706
  - for PTSD, child/adolescent, 3706
- Psychotherapists
  - cancer in, 2280–2281
  - description of, 4607
- Psychotherapy, **2638–2904**. *See also* Interpersonal psychotherapy (IPT); Psychoanalysis; Psychoanalytic psychotherapy
  - acamprosate combined with, 3057
  - for alcohol use disorders, 1277
  - amygdala and, 2813
  - for anorexia nervosa, 2080
  - anxiety and related disorders, 2676–2682
  - for anxiety disorders, 4064
  - applying research to clinical practice, 1703–1704
  - assault victim, 2541
  - autism spectrum disorder and, 3585
  - behavior therapy, 2682–2705
  - for binge eating disorder, 2081
  - bipolar disorders, 1703
    - course of therapy, 1707–1708
    - initial assessment, 1706
  - and brain activity, 751, 751*f*
  - brain structures key for, 2813–2814

- brief, **3741–3749**
- for bulimia nervosa, 2080–2081
- for cancer patients, 2262
- client-centered, 2818–2819
- and cognitive development, 714–715
- computers and, 2828–2829
- consumers, 2860
- contextual model of, 2865–2866
- conversion disorder, 1840
- couch in psychoanalysis *versus,* 2659, 2663
- Dalí's art depicting, 2643, 2643*f*
- deaf persons, 2827–2828
  - clinical issues, 2828
  - history, 2827
  - research, 2828
  - techniques, 2828
  - theoretical issues, 2827–2828
- for dementia, 1199
- depersonalization/derealization disorder, 1944–1945
- depression, 1681
  - history of, 1700–1702
- depressive disorders, 3679–3680, 3693–3694
- for depressive disorders, child/adolescent, 3679–3680, 3693–3694
- description of, 4607
- dignity, 2825–2827
- DLPFC and, 2813–2814
- effectiveness research, 2861
- efficacy research, 2860–2861
- elevated HPA activity and, 1716
- email and, 2672–2673
- emerging themes in, 2865–2866
- encopresis and, 3660, 3674
- evaluation, **2860–2866**
- existential, 2821–2822
- feminist, 2824
- future directions, 1708, 2829
- gay and lesbian, 2009

- gender differences, 4247
- generalized anxiety disorder, child/adolescent, 3714
- guided imagery, 2823–2824
- and HIV infection, 4270
- hoarding disorder, 1794
- for IED, 2112
- individual, **4172–4179**
  - clinical issues, 4176–4178
  - cognitive-behavioral, 4174–4175
  - ethical issues, 4178
  - future direction, 4178
  - history, 4172
  - insight-oriented, 4175–4176
  - interpersonal, 4174
  - life review or reminiscence, 4174
  - problem solving, 4175
  - supportive, 4174
  - techniques, 4173–4174
  - theoretical issues, 4172–4173
- insight-oriented
  - aggression, 2498, 2498*t*
  - intimate partner violence, 2502
  - TBI, 2502
- for intellectual disability, 3513–3514
- interpersonal, 924
- in long-term care, 4229–4230
- under managed care, 2828–2829
- medical model of, 2865–2866
- mentalization-based, 2668
  - suicide prevention and, 2621
  - suicide risk, 2617
- meta-analyses, 2861
- mood disorder due to conditions, 1251
- mood disorders, 1612
- nefazodone combined with, 3126
- neurobiology and, 2812–2813
- by nonprofessionals, 2828–2829

- OCD, 1789–1790
  - pharmacotherapy and, 1790
- other methods of, **2816–2830**
- paradoxical, 2816–2818
- patient clinical characteristics and, 2865
- patient cognitive characteristics and, 2864–2865
- patient demographics and, 2864
- patient–therapist matching and, 2865
- patient–treatment matching and, 2865
- and personality change, 941–942, 942*f*
- for personality disorders, 4107–4108
- pharmacotherapy and, 2864
  - anxiety disorders, 1782
  - mood disorders, 1703
  - in treating PDs, 2168–2169
- positive psychology, 2841–2855
- posttransplantation, 2370–2371
- prefrontal cortex and, 2813
- psychiatric emergencies, 2625–2627
- and psychological testing, 994–995
- psychosis, issues marginalizing, 2868
- psychotic disorders and, 1253
- PTSD and, 1823–1824, 3706
- race and ethnic differences, 803
- research, 1702–1703, 2860–2861
- sedative-, hypnotic-, anxiolytic-related disorders, 1388
- separate personality state development and, 1907–1908
- separation anxiety disorder, 3710–3711
- social anxiety disorder, 3717
- somatic symptom disorder, 1837
- strategic, 2816–2818
- stuttering, 3564
- suicide prevention and, 2620–2621
- suicide risk, 1669, 2612, 2616–2617
- supportive, 2666–2668
  - aggression and, 2498–2499, 2502
- synaptic plasticity and, 2813

- systematic reviews, 2861
- telephone and, 2672–2673
- therapeutic alliance and, 2861, 2864
- therapists intervention studies, 2864
- tic disorders, 3644–3648, 3647*t*, 3658–3662, 3661*t*
- transactional analysis, 2820–2821
- transference-focused, 2668
  - aggression in BPD, 2501
  - suicide prevention and, 2621
  - suicide risk, 2617
- transpersonal, 2824–2825
- treatment outcome studies, 2864
- unanswered questions, 2865–2866
- unipolar disorders
  - group and couples, 1703
  - individual, 1702

- *Psychotherapy and the single synapse* (Kandel), 4545
- Psychotherapy evaluations, 4545
- *Psychotherapy of the Borderline Adult* (Masterson), 4545
- Psychotherapy publications, 4550
- Psychotic, definition of, 4607
- Psychotic depression
  - electroconvulsive therapy, 4153
  - treatment, 4075
    - FGAs, 3070
    - MAOIs, 3120
- Psychotic depressive reaction. *See* Major depressive disorder
- Psychotic disorder(s), **1405–1598**. *See also specific* disorders
  - acute and transient disorders, **1574–1580**
  - after TBI, 538–539
  - alcohol-induced, 1273
  - attenuated psychosis syndrome, **1597**
  - brief psychotic disorder, **1574–1580**
  - catatonia, **1594–1595**
  - chemotherapy agents and, 2273*t*
  - child/adolescent
    - defined, 3725, 3726*t*, 3727*t*

    - differential diagnoses, 3699
  - classification of, 1168, 1171–1172
  - delusional disorder, **1582–1587**
  - delusional symptoms in partner of individuals with delusional disorder, **1597**
  - differential diagnoses, 3610–3611
  - dissociative identity disorder *versus,* 1882
  - dissociativity, 1870
  - due to another medical condition, 1222*t*, 1235*t*, 1250*t*, **1251–1253, 1592–1594**
    - comparative nosology, 1592, 1593*t*
    - course and prognosis, 1594
    - diagnosis and clinical features, 1593
    - differential diagnoses, 1593–1594
    - epidemiology, 1592
    - etiology, 1593, 1593*t*
    - history, 1592
  - early onset, **3725–3733**
  - group psychotherapy, 2746–2747, 2747*t*
  - hoarding and, 1793
  - inhalant-induced, 1339–1340
  - neurosteroids and, 117
  - obsessive-compulsive disorder and, 3699
  - OCD *versus,* 1788
  - other specified schizophrenia spectrum and other psychotic disorder, **1595–1597**
    - comparative nosology, 1596, 1597*t*
    - course and prognosis, 1596–1597
    - diagnosis and clinical features, 1596
    - differential diagnoses, 1596
    - epidemiology, 1596
    - etiology, 1596
    - history, 1595–1596
  - overview, 1405–1406
  - PCP-induced, 1325–1326
  - phenotypes of, **1531–1538**
  - rehabilitation, **1531–1538** 1549
  - risk of, preterm birth and, 644
  - schizoaffective disorder, **1587–1590**

- schizophrenia, **1405–1573** (*See also* Schizophrenia)
  - early onset, **3725–3733**
- schizophrenia heritability and, 1446–1447
- schizophreniform disorder, **1580–1582**
- sedative-, hypnotic-, anxiolytic-induced, 1381
- stimulant-induced, 1286
- stress and, 2323
- substance abuse induced by
  - FGAs and, 3071
- substance/medication-induced, 1222*t*, 1235*t*, 1250*t*, **1251–1253**, **1590–1592**
- suicide and, 2613
- suicide behavior risk factor, 3684, 3698
- in transplant populations, 2366
- treatment
  - anticonvulsants, 2970
- unspecified schizophrenia spectrum and other psychotic disorder, **1595–1597**

- Psychotic features
  - dissociative identity disorder
    - differential diagnoses, 1916
  - major depressive disorder, 1644
  - mood-congruent/incongruent, 1637
    - treatment selection and, 1679
- Psychotic mania, 459
- Psychotic symptoms
  - pharmacotherapy of, 2173
  - schizophrenia and, 1407*t*, 1410–1417
- Psychoticism, factor analysis, 935
- Psychotomimetic drugs, 4607
- Psychotropic drugs. *See also* Psychopharmacology
  - adverse cutaneous reactions to, 2355, 2356*t*
  - anxiety disorders, 1782*t*
  - cardiovascular effects of, 2197*f*
  - family and couples therapy and, 2757
  - in functional GI disorders, 2207
  - Nobel Prize for research on, 4294
  - and physical health monitoring, 1022–1025
    - antidepressants, 1025

    - antipsychotic agents, 1022–1024
    - mood stabilizers, 1024–1025
  - psychopharmacology, 3795–3796
  - regimen compliance, 4300
  - sexual dysfunction and, 1779*t*
  - toxic effects, 252
  - weight gain and, 2214–2215, 2215*t*
- Psysostigmine, 347–348
- *PTEN* gene mutations, 314
- PTSD. *See* Posttraumatic stress disorder (PTSD)
- Puberty
  - eating disorders and, 368–369
  - gender dysphoria in, 2033
  - hormonal control of, 2505–2506
  - masturbation during, 1961
  - overview, 3368–3369, 3368*t*
- Public agencies, psychiatrists and, 4305–4307
- Public areas, correctional psychiatry and, 4453, 4453*t*
- Public education about mental health
  - as overcoming stigma, 4529
  - world-wide, 4525
- Public health
  - community psychiatry in, 4289–4290
  - epidemiology as core of, 4301–4302
- Public Health Security and Bioterrorism Preparedness Act of 2002, 2931
- Public psychiatry
  - access disparities, 4304
  - Affordable Care Act (2010) role in (*See* Affordable Care Act (2010))
  - conceptual map for, 4306, *4306f*
  - health care reform (*See* Health care reform)
  - hospitalists (*See* Hospitalists)
  - hospitals (*See* Psychiatric hospitals)
  - overview of
    - concept of, 4287–4288
    - historical perspective of, 4290–4295
    - organized care systems, 4293–4294
  - rehabilitation (*See* Rehabilitation)

- research (*See* Mental health services research)
- shortage of treatment options, 4305
- sociocultural framework (*See* Sociocultural framework)

- Public Reporting of Performance (PR), 4319–4320
- Publications in psychohistory and psychobiography, 4548
- Publicly shamed, suicide and, 2621–2622
- Puerperal psychosis. *See* Postpartum psychosis
- PUFA. *See* Omega-3 polyunsaturated fatty acid supplements
- Pulmonary disease
  - and HIV infection, 4271
  - schizophrenia comorbidity, 1541–1542
  - and sexual dysfunction, 4262
- Pulmonary system, sleep and, 344
- Pulse diagnosis, 2545–2546
- Pulvinar nuclei, 21
- Punding, in PD, 546
- Punishing stimulus, 2694
- Punishment
  - behavior therapy, 2694
  - description of, 2872, 4607
  - and learning theory, 718
- Pupils, dilation of, 456
- Purcell, Edward, 259
- Pure Food and Drug Act of 1906, 2930
- Purging disorder, 4607
- Purines, neurotransmission by, 141
- Purkinje neurons, 56
- Purpose and meaning, empirical findings, 2847–2848
- Purtell, James J., 1830
- Putamen
  - anatomy, 26, 27–28
  - function of, 8, 10
  - internal organization of, 28–29
- Putnam, Frank W., Jr.
  - Dissociative Experiences Scale and, 1870
  - identity-state dissociation neurobiology, 1893–1894
- PVN. *See* Paraventricular nucleus (PVN)

- Pyramidal dysarthria, 457*f*
- Pyramidal neurons
  - connections, 6–7
  - drawing of, 16*f*
  - shape of, 16
- Pyramidal signs, 469*t*
- Pyramidal system, lesions in, 31
- Pyridoxine (B_6), 2547
- Pyrimethamine, 585
- Pyromania, **2113–2114**
  - description of, 4607
  - diagnosis, 2113
  - epidemiology, 2113
  - etiology, 2113
  - treatment, 2113

Q

- Q fever, war or terrorism and, 2572*t*
- q-sort, description of, 4607
- QEEG. *See* Quantitative electroencephalography (QEEG)
- *Qigong*
  - as alternative treatment, 2542
  - as energy therapy, 2546, 2561, 2563
  - psychiatric assessment and, 2546
  - psychotic reaction, 2563
  - as somatic therapy, 2556–2557
- QT intervals. *See also* Long QT syndrome
  - drug increasing, 3157–3158
- Quaalude. *See* Methaqualone
- Quadrangular therapy, 4607
- Quality assurance, 4607
- Quality Chasm
  - best practices as bridge for, 4347–4348
  - Six Aims in, 4314, 4315
- Quality improvement
  - in managed care, 4349
  - in Triple Aim concept, *4310t*
- Quality of life
  - autism spectrum disorder children, 3928
  - everyday functioning, 4026, 4026*t*
  - neurocognitive impairment, 1509
  - ratings in burn patients, 2378
  - ratings in CHD patients, 2198
  - schizophrenia, 1509, 4097
- Quality-of-life therapy, 2854
- Quantitative electroencephalography (QEEG). *See also* Forensic psychiatry, neuroimaging
  - psychiatric assessment and, 2544–2545
- Quantitative law of effect, 723, 723*f*
- Quantitative trait loci (QTLs), 308
- Quantitative variables, 4607
- Quantum information theory, 2546
- Quantum mechanics

- consciousness model, 2544
- energy therapies and, 2546
- MRI/MRS, 261–262

- *Quarterman, Panetti v.,* 4438
- Quazepam (Doral), 3001
- Queen Square montage, 223, 224*f*
- Quench, magnet, 264
- *Questiones medico-legales* (Zacchia), 4534
- Questions. *See also* Interviews
  - in cognitive-behavior therapy
    - assumptions, 4185
    - conceptual clarification, 4185
    - implications and consequences, 4185
    - probing assumptions, 4185
    - rationale and evidence, 4185
    - viewpoints, 4185
  - at patient interviews
    - open- *vs.* closed-ended, 947
    - person-centered and disorder-based, 947–948
- Questions at patient interview
  - in patient interview, 959–960, 968–969*t*
- Quetiapine (Seroquel)
  - acute mania, 1686, 1686*t*, 1689*t*
    - lithium or divalproex, 1689*t*
    - lithium or valproate and, 1687
    - monotherapy, 1686
  - for aggressive behavior, 2173
    - in mentally retarded children, 2500
  - anxiety disorders, 1781, 1782*t*
    - GAD, 1783
  - for anxiety disorders, 4131
  - bipolar depression, 1690, 1691, 1692*t*, 1693, 1695
  - bipolar disorders
    - alcohol comorbidity and, 1699
    - cocaine abuse and, 1699
    - early onset, 3692, 3706, 3791
    - generalized anxiety disorder and, 1698

    - panic disorder and, 1699
  - bipolar I disorder, 1696*t*, 1697
  - bipolar II depression, 1697
  - for cancer patients, 2253
  - CATIE, 3069–3070
  - in children, 2281
  - clinical use of, 4352
  - combination pharmacotherapy and, 3324
  - for delirium, 4094
  - delirium treatment, 1188, 1190
  - dementia treatment, 1201
  - dissociative identity disorder, 1928
  - dosage/administration
    - in cancer patients, 2257
    - for delirium in cancer patients, 2258*t*
  - in functional GI disorders, 2206
  - for insomnia, 2296
  - lowest mortality rate with, 4342
  - mania, 1612
  - prolactin levels and, 2249
  - for psychosis
    - in PD, 548
  - psychotic disorders and, 1253
  - SAD, 1784
  - schizophrenia, 4098
    - neuroimaging studies of, 1497
  - side effects, 67, 2237, 2237*t*
  - weight gain and, 2215, 2237, 2237*t*
- Quick Inventory of Depressive Symptomatology (QIDS) and QIDS-A, 3400
- Quiet sleep, 339
- Quinagolide (Norprolac), 2250
- Quinidine (Cardioquin), 3157–3158
- Quinolinic acid, 196

R

- Rabbit syndrome, 1148
- Rabies
  - incubation period, 595
  - symptoms, 595
- Race. *See* Race/ethnicity
- Race/ethnicity
  - anxiety disorders, 1734
  - cancer survival and, 2285
  - as child psychiatric evaluation consideration, 3379
  - cognition and racial prejuduce, 701–702
  - definition of, 788
  - homicide rates and, 2476
  - intellectual disability and, 3498
  - interpersonal aggression and, 2476
  - and mental disorders in U.S., 800
  - mood disorders and, 1616
  - obesity and, 2211*t*
  - Parkinson disease and, 543
  - and psychiatric service utilization, 801, 801*t*
- Rachman, Stanley, 2683
- Racial prejudices, 701–702, 2035
- Racker, Heinrich, 2654
- Radiation, exposure to, 293
- Radiation treatment
  - dementia related to, 2284
  - psychiatric problems due to, 2270–2271
- Radioactivity, principles of, 286–293
- Radionucleotides, 286–293
- Radiotracer imaging, **285–304,** 297
- Rado, Sandor, 1986–1987
- Radon transform, 289
- Raf, function of, 125
- Rage reactions, 677
- Rahe, Richard, 2317
- Ramachandran, V.S., 2840
- Ramelteon (Rozerem)

- chemistry, 3101–3102, 3102*f*
- dosage/administration, 3106
- drug interactions, 3106, 3106*t*
- drug studies, 3103–3104
- insomnia, 3101, 3103–3104, 3105
  - in aging adults, 4054
- laboratory interference, 3106
- melantonin agonism by, 69
- organ/system effects of, 3104–3105
- pharmacological actions, 3102–3103, 3103*t*
- precautions/adverse reactions, 3105–3106
- therapeutic indications, 3105, 3105*t*

- Random, definition of, 4607
- Random samples, 4607
- Random variables, 4607
- Randomized clinical trials (RCTs)
  - antidepressant-associated suicide, 1673
  - behavior therapy, 2704
  - CBT, 1702
  - dialectical behavior therapy, 2785
  - hallucinogens, 2887
- Range (mathematical), 4607
- Rangell, Leo, 2639, 2648
- Ranitidine (Zantac), 75
- Rank, Otto, 917f
  - dialectic, 918
  - intensive short-term dynamic psychotherapy, 2797
  - *The Myth of the Birth of the Hero,* 4537
  - as psychoanalyst, 2644
  - work of, 4607
- RANZCP. *See* Royal Australian and New Zealand College of Psychiatrists
- Rapaport, David, 2738
- Rape
  - as assault, 2538
  - on college campuses, dissociation and, 1873
  - definition, 2538–2539
  - description of, 4607

- in jails and prisons, 4460, 4466
- prevalence, 2539
- psychiatric emergency due to, 2634–2635
- risk factors, 2539

- Rape, Abuse and Incest National Network (RAINN), 2539
- Rape victim syndrome, 1812
- Raphe nuclei, 68
- Rapid cycling
  - antidepressant-induced, 1642
  - bipolar disorder, 1652
  - mood disorder, dissociative identity disorder and, 1916
- Rapid cycling patients
  - lithium and, 3095–3096
- Rapid eye movement (REM)
  - behavior disorder, 455
    - nightmares and, 2099–2100
    - overview, 2100
  - sleep behavior disorder, 4057
- Rapid eye movement sleep. *See* REM sleep
- Rapid Learning Health Care System, 4324
- Rapid opiate detoxification under anesthesia (RODA), 2948
- Rapid Teen Substance Use Screening Tool for Clinicians (S2BI), 3858
- Rapid transcranial magnetic stimulation, 4163
- Rappaport, David
  - on dissociation and dissociative disorders, 1869
  - psychoanalytic model of dissociative disorder, 1903
- Rapport, definition of, 4607
- Rapture-of-the-deep syndrome, 4607
- Rasagiline (Azilect)
  - history, 3111
  - mechanisms of action, 3114
  - pharmacological actions, 3116
- Rashes
  - discoid
    - assessment of, 456
  - vasculitic
    - assessment of, 456

- Rashid, Tayyab, 2853
- RASSLs (receptors activated solely by synthetic ligands), 130
- Rat Man (Freud case), 2640
- Rathke pouch, embryonic, 677
- Rating scales
  - ADHD and, 3591, 3600, 3601*t*
  - for aging adult, 3976
  - for bipolar disorder, child/adolescent, 3688, 3702
  - Brief Psychiatric Rating Scale, 1006–1007
  - comprehensive (omnibus), 3398–3399
  - for delirium, 1183, 1188*t*, 1189*t*
  - for depression, 3399–3400
  - disorder-related (domain specific), 3399–3402
  - overview, 3398
  - for psychopharmacology, 3787
  - schizophrenia, 1409–1410
  - tic disorders, 3641, 3641*t*, 3655, 3655*t*
- Rational emotional behavior therapy (REBT), 2498, 2696
- Rational emotive therapy (RET), 2683, 2696
- Rational problem solving, 2696
- Rationalization, description of, 4607
- *Rauwolfia serpentina,* 2554
- Ray, Isaac
  - *Treatise on Medical Jurisprudence of Insanity,* 4535
  - work of, 4607
- Raynaud phenomena, 456
- Rayner, Rosalie, 3749
- Raz, Amir, 2709, 2712
- RBANS. *See* Repeatable Battery for the Neuropsychological Assessment of Schizophrenia (RBANS)
- RCBF. *See* Regional cerebral blood flow
- Reaction formation, 4608
- Reaction time crossover, 695
- Reactive attachment disorder (RAD), **3662–3667, 3676–3681**
  - attachment, 3662, 3663*t*, 3676, 3677*t*
  - clinical features, 3664–3665, 3678–3679
  - comparative nosology, 3663–3664, 3677–3678
  - defined, 3662, 3663*t*, 3676, 3677*t*

- diagnosis, 3664–3665, 3678–3679
- differential diagnoses, 3665–3666, 3679–3680
- environmental factors, 3663, 3677
- epidemiology, 3664, 3678
- etiology, 3664, 3678
- overview, 3662, 3676
- treatment, 3666–3667, 3680–3681

- Reactive attachment disorder of infancy or early childhood, 4608
- Reactive depression
  - description of, 1647, 4608
  - neurobiology and, 1710
- Reactive disorders, 4608
- Reactive dissociative psychosis, 1948
- Reactive manias, 1649
- Reactive oxygen species (ROS), 155
- Reactive processes, in behavior therapy, 2687
- Reactive schizophrenia, 4608
- Reactivity, childhood/adolescent, 3705
- Readiness for change stages, 1124
- Reading comprehension, 461
- Reading disability, 3425. *See* specific learning disorders
- Reading disorders, 4608
- Reading problems, language disorders and, 3548
- Reading the Mind in the Eyes Test, 4370
- Reagan, Ronald, 4293
- Real relationships
  - psychoanalysis and, 2655
  - psychodynamic psychotherapy and, 2677
- Reality, definition of, 4608
- Reality principle, 4608
- Reality testing, 4608
- Reassurance in patient interview, 957
- Rebound, ADHD treatment effects, 3593
- Reboxetine, 1779*t*
- Recall
  - delayed in PTSD, 700
  - description of, 4608

- Receiver operating characteristic (ROC), 4608
- Recent memory, 4608
- Recent past memory, 4608
- Receptive aphasia. *See also* Fluent aphasia; Sensory aphasia
  - description of, 4608
- Receptive dysphasia, 4608
- Receptor pharmacology, **2905–2914**
  - G protein effectors, 2912–2914, 2913*f*–2915*f*, 2914*t*
    - adenylyl cyclase, 2913, 2913*t*
    - G protein-activated ion channels, 2914, 2915*f*
    - phospholipases, 2914, 2914*f*
  - growth factors, receptors for, 2907–2912, 2912*t*, 2999*f*–2912*f*
  - plasma membrane microdomains and GPCR signaling, 2912, 2913*f*
  - receptor analysis, 2905–2907
    - binding assays, 2905–2906, 2906*f*
    - biological (efficacy) assays, 2906, 2906*f*
    - protein modulation of agonist binding, 2906–2907, 2907*f*
  - receptor structure and function, 2907, 2907*f*
    - ionotropic receptors, 2907, 2907*f*–2908*f*
    - ligand gated receptors, 2907, 2908*f*
    - voltage gated receptors, 2907, 2907*f*
- Receptors, description of, 4608
- Recidivism, 4340
- Reciprocal inhibition, 2689
  - desensitization and, 4608
- Reciprocity, schematic drawing, 6*f*
- Recognition memory, 4489–4490
- Recollection, 2139. *See also* Episodic memory
- Reconciliation, spirituality and, 2582–2583
- Reconsolidation, fear conditioning and, 722
- Reconstruction phase, after disasters, 2566
- Records, ethical practice and, 4446
- Recovered memories, dissociative disorders and, 1872
- Recovered memory syndrome, 4608
- Recovered memory therapy (RMT), 1938
- Recovery. *See also* Rehabilitation
  - definition of, 4340, 4363

- during rehabilitation, 4369
- Recovery model, schizophrenia and, **1548–1561**
  - advocacy, 1559–1560
  - defined, 1549–1551
  - evidence-based practices, 1551–1554, 1552*t*
  - history, 1548–1549
  - involuntary treatment, 1558–1559
  - journey of, 1554–1555
  - outcomes, 1551, 1551*f*
  - peer support specialist's role, 1557–1558
  - psychiatrist's role, 1555–1557
- Recovery Oriented Decisions for Relatives' Support (REORDER), 4350
- Rectal pain syndrome, 2355*t*
- Recurrent brief depressive disorder, 1647, 4608
- Recurrent isolated sleep paralysis, 4608
- Recurrent pregnancy loss, 2511
- Recurrent unipolar depression, 2529*t*
- Redecision group therapy, 2738
- Redirecting in patient interview, 959
- *Redmond, Jaffe v.,* 2010, 2671, 4436
- Reelin, 52, 55
- *Reelin* gene *(RELN),* 52, 446–447
- Refeeding syndrome, 2079, 2079*f*
- Referential attitude, 2878
- Referral issues
  - decision-making capacity, 987
  - differential diagnoses
    - age- or stress-related cognitive change, 984
    - dementia *vs.* depression, 986
    - mild traumatic brain injury, 985–986, 985*t*
    - poststroke syndromes, 986
  - forensic evaluation, 987
  - functional changes over time, 987
  - level of functioning, 984
- Referrals
  - for forensic evaluations, 3862–3863
  - for neuropsychological assessments, 3404

- for psychiatric evaluations, 3377
- for psychological assessments (standard), 3403–3404
- specific questions to ask, 1007

- Reflection in patient interview, 957
- Reflex theory, nose and, 2178
- Reflexes
  - abnormal, 460
  - milk ejection, 1957
- Reflexive gaze, inhibition of, 464
- Refractory periods, 4608
- Reframing, 4608
- Refugees
  - global perspective, 796–797
  - and immigration, 806
  - psychiatric assessment, 793
- Refusal of care, voluntary, 2314
- Regeneric memory, 1935
- *Regents of the University of California, Tarasoff v.,* 4435
- Region of interest (ROI) connectivity
  - network-based, 420
  - seed-based, 420–421
- Regional cerebral blood flow (rCBF), 4608
- Registration. *See* Memory
- Registries, evidence-based health treatments, 4296
- Regression
  - avoidance, STAPP and, 2798
  - in burn patients, 2381
  - description of, 4608
  - partial, hypnosis and, 2711
  - in psychosis, 2872
- Regression models, 839–840
- Regressive-reconstructive approach, 4608
- *Regulations of the Hospitals of Santa Maria Nuova and of Bonfazio* (Chiarugi), 4534
- Regulators of G protein signaling (RGS), 2912
- Rehabilitation
  - for alcohol use disorders, 1277–1279
  - assessment in, 4369–4370

- cognitive assessment in, 4370
- cognitive behavioral therapy in, 4373
- cognitive remediation in, 4372–4373
- collaborative/therapeutic relationship with patient during, 4368
- with comorbid substance use disorders, 4375–4376
- continuity of care, 4367
- description of, 4608
- empirically based models of, 4363
- evidence-based treatment for, 4363–4364
- functional outcomes assessment, 4371–4372
- history, 1532, 1549
- holistic, 4360–4361
- housing first, 1533–1534, 1550, 1553
- illness management during, 4368
- illness/wellness management training in, 4374–4375
- individual placement, 1534–1535
- integration of, 4362, 4376
- multidisciplinary
  - in pain syndromes, 2397
- neurocognitive symptoms during, 4370
- overview, 1531–1532
- person-centered recovery models, 4363
- phase-linked treatment, 4364–4365, *4365f*
- prodromal phase, 4365–4366, *4366t*
- for psychotic disorders, **1531–1538**
- recovery during, 4363, 4369
- relapse prevention, 4366–4367
- for schizophrenia, **1531–1538**
- sex education in, 4375
- skills training in, 4374
- of social cognition, 4370, 4373
- social learning/skills in, 4370, 4373–4374
- stabilization phase of, 4367–4369
- support techniques, 1533–1537
- supported education in, 4375
- supported employment, 1533–1534, 1536*f*, 1552
- technology tools for, 1535–1537

- vocational, 4369
- Rehearsal in information processing, 689
- Reich, Wilhelm, 916*f*
  - bioenergetics, 917
  - on character armor, 2644
  - compulsive character, 917
  - ego psychology, 916–917
  - hysteria, 917
  - masochism, 917
  - personality theory, 916–917
  - phallic-narcissism, 917
  - work of, 4608
- Reik, Theordor, 4608
- Reiki
  - as alternative treatment, 2542
  - as energy therapy, 2546, 2561, 2563
- Reil, Johann, 4534
- Reimbursement rates, Medicaid, 4380
- Reinders, Antje A.T.S., 1893, 1894
- Reinforcement. *See also* Conditioning
  - definition, 2872, 2872*f*
  - depression and, 1607
  - description of, 4608
  - and learning theory, 717, 717*f*
  - in operant learning, 722–724
  - in patient interview, 957
  - positive
    - definition, 2872
    - mechanisms, 1283
- Reinforcement learning models, 434–435
- Relapses
  - for alcohol use disorders, 1277–1278
  - early-onset bipolar disorder, 3693, 3707
  - intervention for, 4366–4367
  - neurocognitive impairment, 1509
  - opioid-related, 1364, 1370
  - prevention of

    - bipolar disorder, early onset, 3693, 3707
    - in CBT, 3756, 3757*t*, 4186
    - in positive psychology, 2845
  - schizophrenia, 1509, 1530
  - signs and symptoms preceding, 4366
- Relatedness, description of, 4608
- Relational problems, 4609
- Relationship Across Domains (RADs), 4370
- Relationships
  - adulthood, 3941, 3942–3943, 3943*f*
  - happiness and, 2848
  - instability in, 2140
  - of LBGT persons, 2004
  - mentalizing, 2900–2901
  - and natural selection, 777
  - problems
    - hypnosis and, 2720–2721
    - physician, 2601–2602
  - as psychosocial factor, 4165–4168
  - skin disease and
    - attachment styles, 2335
    - communication difficulties, 2335
    - sexual intimacy and, 2335
- Relative frequency. *See* Frequency
- Relative refractory periods, ion channels, 133
- Relative risks, in epidemiology, 813, 813*t*
- Relative Value Scale, 4609
- Relaxation
  - applied, as somatic therapy, 2557
  - exposure therapy and, 2689
  - hypnotic, 2712
    - forensic evaluation and, 2734
    - progressive, 2715
  - pain management using, 2377
  - PTSD, 1824
  - somatization disorder, 1836
  - T_1, 262

- T_2, 262
- T_2*, 262
- training, habit reversal and, 2698

- Relaxation therapy
  - for acne, 2345
  - for atopic dermatitis, 2342
  - differential relaxation, 2327
  - for dying patients, 2302
  - in functional GI disorders, 2207
  - for insomnia, 2089
  - mini-practices, 2327
  - for stress management, 2326–2327
- Relay nuclei, thalamic
  - connections, 22*t*
  - function of, 21–22
- Relevance, of neuroimages, 4479–4480, 4482–4483
- Reliability
  - definition of, 4609
  - of personality tests, 996–997
  - validity and, 996
    - of child interview, 3381
    - of psychological assessments (standard), 3404–3405, 3413–3414
    - of standardized assessment instruments, 3394–3395
- Religion, 2141. *See also* Spirituality
  - abused women and, 2593
  - for cancer patient support, 2254
  - as child psychiatric evaluation consideration, 3379
  - in clinical practice, 2583–2594
  - death and, 2306
  - dying patients
    - decision-making issues, 2302–2303
    - spiritual concerns of, 2302–2303
  - ethical issues with, 3871–3872
  - homosexuality and, 2592
  - maltreatment (child) and, 3839–3840
  - at patient interview, 967
  - psychiatrists' perceptions about, 2585–2586

- religiosity, spirituality and, 2583–2585
- Religiosity
  - and mental health, 766, 768
  - positive psychiatry, 4283
  - religion, spirituality and, 2583–2585
- Religious cult mind-control debate, 1947
- Religious delusions, 4573, 4609
- Religious life, 2591–2593, 2592*f*
- Religious obsessions, 1699
- REM. *See* Rapid eye movement
- REM (rapid eye movement) sleep, 347–348
  - behavior disorder, 4609
  - description of, 339, 4609
  - EEG, 341*f*
  - latency, 343
  - mood disorders and, 1609, 1716
  - paradoxical sleep, 339
  - PTSD and, 1821
- Remediation in CBT, 4191
- Remeron. *See* Mirtazapine
- Reminiscence
  - in autobiographical memory, 1935
  - description of, 4609
  - in group therapy, 4202, 4204
  - in psychotherapy, 4172, 4174
- Reminyl. *See* Galantamine
- Remission, 1265
  - of alcohol use disorders, 1276
  - description of, 4609
  - psychopharmacology and, 2922
- Remote memory. *See* Long-term memory
- Remotivation, description of, 4609
- Renal disease, amantadine and, 2962, 2965*t*
- Renal function. *See also* Kidney
  - carbamazepine effects on, 3035
  - lithium effects on, 3097–3098
  - melatonin effects on, 3104

- Renik, Owen, 2653
- Repeatable Battery for the Neuropsychological Assessment of Schizophrenia (RBANS), 4004, 4372
- Repeated exposure to traumatic events
  - PTSD and, 1813*t*, 1814
- Repeated measures models, 842–843
- Repetition, disordered, 461
- Repetition compulsion, 4609
- Repetitive patterns, 4609
- Repetitive transcranial magnetic stimulation (rTMS)
  - after TBI, 533
  - anxiety disorders, 1782
  - bipolar depression, 1692*t*
  - depersonalization/derealization disorder, 1945
  - depression, 33, 1681
- Repressed desires, 2647
- Repression
  - description of, 4609
  - dissociation *versus,* 1869
  - dissociative disorder controversy and, 1873
  - in psychoanalysis, 855–856
  - *vs.* repressing, 862
- Repressive-inspirational group psychotherapy, 4609
- Reproducibility of neuroimages, 4479
- Reproduction
  - psychological states and, 2508
  - senescence, 2521–2524
- Reproduction, genetic selection, 772–775
- Reproductive development
  - fetal development, 2505
  - neonatal development, 2505
- Reproductive hormonal therapy, **3329–3337**
  - actions, 3330–3332, 3331*f*
  - future directions, 3336
  - history, 3329
  - reasons for, 3332
  - structure, 3330, 3330*f*
  - types of, in affective illness, 3332–3336

- Reproductive medicine
  - psychiatry and, **2505–2524**
- Reproductive system, cannabis use and, 1309–1310
- *The Republic* (Plato), 927–928
- Requip. *See* Ropinirole
- Rescuers, negative therapeutic reaction and, 2657
- Research
  - implementation of, 4349
  - informed consent issues, 2312
  - into mental health disorders, 4291
  - world-wide, 4525, 4526
- Research Diagnostic Criteria (RDC), 1154
- Research Diagnostic Criteria–Preschool Age (RDC-PA), 3409–3410
- Research Domain Criteria (RDoC)
  - as classification alternative, 1161–1162, 1161*t*
  - on psychiatric research, 3348
- *Research Involving Persons with Mental Disorders That May Affect Decision Making Capacity,* 4243
- Research subjects, older adults, 4235
- Reserpine (Serpalan)
  - for chorea, 555
  - depression and, 1713
  - description of, 4609
  - history, 2916–2917
- Resick, Patricia, 2696
- Residence, mood disorders and, 1616
- Residency programs, reform of, 4319
- Residential treatment centers (RTCs), **3801–3802**
  - adolescent, 3815
  - description of, 4609
  - for pediatric psychiatric treatment, 3797
- Residents, definition of, 4609
- Residual, definition of, 4609
- Residual infantile autism, 3573
- Residual schizophrenia, 4609
- Resilience
  - children, 3915–3916
  - definition of, 2336

- empirical findings, 2847
- and mental health, 759–762
- positive psychiatry, 4283
- temperament and, 3488

- Resistance
  - description of, 4609
  - in group therapy, 4203
  - in psychoanalysis, 855, 895–896
  - psychoanalysis and, 2640, 2652–2653
  - psychodynamic psychotherapy and, 2677
- Resistance *(R),* ion channel, 132
- Resistant attachment, 4609
- Resolution
  - dissociative identity disorder treatment and, 1926
  - of neuroimages, 4479
- Resource utilization, hospitalist role in, 4358
- Respiratory disorders, **2223–2230**.
  See also specific diseases/disorders
  - definition of, 2223
  - nosology, 2223
  - panic disorder and, 1783
  - and psychiatric symptoms, 4048
- Respiratory failure, 2228
- Respiratory symptoms in dying patients, 2297
- Respiratory system
  - anticholinergics and, 2958
  - cannabis use and, 1309
  - ECT effects in, 3296
  - melatonin effects on, 3105
  - nonbarbiturate sedative-hypnotic agent effects on, 2989
  - opioid receptor antagonist effects on, 3143
- Respite beds, 4344
- Respondent-based diagnostic based interviews (RBI), 3395–3397
- Respondent conditioning. *See also* Conditioning
  - stimulant-use disorder, 1283
- Response cost
  - as aversion-based approach, 2694
  - in token economies, 2693

- Response elements (REs), binding to, 159
- Response expectancy model, 2711
- Response prevention (RP), 2690–2691
- Response styles, 2399
- Resting membrane potentials
  - description of, 130
  - potassium ions, 131–132
- Restless leg syndrome (RLS)
  - in aging adult, 4056–4057
  - description of, 4609
  - diagnosis of, 2101, 2101*t*
  - dopamine agonists in, 67
  - in long-term care, 4230–4231
  - overview, 2100–2101
  - polysomnography, 2101*f*
  - treatment
    - DA receptor agonists, 3238
    - gabapentin, 2968
- Restorative experiences, lack of, 1906
- Restoril. *See* Temazepam
- Restraint
  - during child emergency psychiatric evaluation, 3440, 3441
  - for inpatients, 3799
- Restricted affect, description of, 4609
- Ret receptor, 100
- Retardation. *See also* Mental retardation
  - description of, 4609
- Retarded ejaculation. *See* Delayed ejaculation; Inhibited male orgasm
- Retarded thought. *See* Slowed (retarded) thought
- Retention. *See* Memory
- Reticular nucleus, 22–23
- Reticuloendothelial system, 180
- Retigabine, 2396
- Retina, neural connections, 7
- Retinal photoreceptors, 208–210
- Retinohypothalamic tract, 205–206
- Retirement

- family role, 4197
- health care costs for older adults, 4213
- Retirement of elderly, 3998, 4168, 4169
- Retroflexion, 4609
- Retrograde amnesia, 4609
- Retrospective falsification, 4609
- Retrospective studies, 4609
- Retroviruses, CNS infections, 599
- Rett syndrome, 4609
  - CFD and, 664
  - description of, 4609
  - epigenetic imprinting and, 161
  - genetic etiology of, 314
  - genetic testing for, 3456
- *Rev-Erbα gene, 206*
- Reversal, description of, 4610
- Reversal design experiments, 2704, 2704*f*
- Reverse neurovegetative symptoms, 1710
- Reverse transcriptases, 567
- Reversibility, as cognitive skill, 3360
- Reversible inhibitors of MAO-A (RIMA), 69
  - anxiety disorders, 1780, 1782*t*
  - sexual dysfunction and, 1779*t*
- Reversible models, 294–296
- Revex. *See* Nalmefene
- ReVia. *See* Naltrexone
- Revised Behavior Problem Checklist (RBPC), 3399
- Revised Children's Manifest Anxiety Scale (RCMAS), 3401
- Revivification, 2711
- Revonsuo, Antti, 2709
- Reward Dependence
  - Cluster A PDs and, 2163
  - harm avoidance and, 2146*f*
  - high, in children, 2159
  - personality disorders and, 2162*t*
  - psychobiological findings, 2133*t*
  - secure attachments and, 2174

- TCI scales, 2164*t*
- temperament and, 2132
- Reward prediction errors (RPEs), 415
- Rey Auditory Verbal Learning Test (RAVLT), 4372
- Reynolds Adolescent Depression Scale (RADS) (for ages 11 to 20), 3400
- Reynolds Child Depression Scale (RCDS) (for ages 7 to 13), 3400
- Rey–Osterreith Complex Figure Test, 979, 980*f*, 991
  - in aging adults, 4006
- RGS2, SSRIs and, 1746
- *Rhapsodies about the Application of Psychotherapy to Mental Disturbances* (Reil), 4534
- Rheumatic chorea, 1649
- Rheumatic disease
  - diagnosis of, 472
  - review of, 454
- Rheumatoid arthritis
  - laboratory evaluation, 1015
  - psychiatric problems, 4049
  - psychosomatic component, 915
  - treatment of, 2322
- Rheumatologic disorders, 3846–3847
- *Rhodiola rosea* (golden root), 2554
- Rhombencephalon
  - embryology of, 41
  - in human development, 8
- Rhythmic mid-temporal discharges (RMTDs), 239, 240*f*
- Rhythmic movement disorder, 2104
- Rhythmopathy, mood disorders and, 1609
- Riboflavin, migraine treatment, 624
- Ribosomes, proteins and, 151
- Ribot, Theodule
  - dissociative identity disorder, 1899, 1910
  - at the Salpêtrière, 1868*f*
- Ribot law, 731
- Richer, Paul, 1868*f*
- Ricoeur, Paul
  - on narrative identity, 2832
  - narrative psychiatry, 2833

- Ridge, Damien, 2840
- Rief, Winfried
  - somatic symptom and related disorders and, 1842
  - somatic symptom disorder, 1837
- Right to refuse treatment, 4610
- Right-to-treatment, 4610
- Right-to-treatment litigation, 4336
- Rigidity
  - description of, 4610
  - lead-pipe, 458
  - paratonic, 458
- Riluzole (Rilutek)
  - anxiety disorders, 1743*t*
  - generalized anxiety disorder, 1748
  - OCD, 1789
- Rimonabant
  - anxiety and, 1750
  - effects of, 115
  - THC analgesia and, 115
- Risk, epidemiology, 812–813
- Risk assessment
  - of neuroimaging, 4488–4489
  - at patient interview, 954–955
- Risk factors
  - for Alzheimer disease, 3964
  - for anxiety disorders, 3966–3967, 4060
  - for delirium, 4088, 4088*t*
  - description of, 4610
  - genetics of neurodegenerative disorders, 4023–4024
  - HIV infections, 4269
  - physician ethics, 4444
- Risk genes, schizophrenia, 1462
- Risk (vulnerability) model, 3484–3485
- Risk Need Responsivity Model, 2021
- Risk-taking behavior
  - in adolescence, 3367, 3372–3373
  - emergency psychiatry evaluation for, 3444

- Internet use as, 3373
- suicide and, 3683, 3697

- Risperdal. *See* Risperidone
- Risperidone (Risperdal)
  - acute mania, 1686, 1686*t*, 1688–1689, 1689*t*
    - combination/add-on therapy, 1690
    - lithium or divalproex, 1689*t*
    - lithium or valproate and, 1687
    - monotherapy, 1686, 1687
  - in ADHD, 2500, 3595
  - adverse effects of, 2356*t*
  - for aggressive behavior, 2112, 2172
    - ADHD, 2500
    - Alzheimer disease, 2502
    - in autism, 2500
    - conduct disorder, 2501
    - mentally retarded children, 2500
    - psychiatric emergencies, 2628
    - in schizophrenia, 2501
  - anxiety disorders, 1781, 1783
  - autism spectrum disorders, 3585, 3794
  - bipolar depression, 1695
  - for bipolar disorder, early onset, 3692, 3706, 3791
  - bipolar disorders
    - cocaine abuse, 1699
    - panic disorder, 1699
  - bipolar I disorder, 1696, 1696*t*
  - black box warning, 2919
  - for cancer patients, 2253
  - CATIE, 3069–3070
  - in children, 2281
  - for children with CD, 2501
  - combination pharmacotherapy and, 3327
  - for delirium, 1188, 4093
  - dementia treatment, 1201
  - depression, 1680–1681
  - dissociative identity disorder, 1928

- dosage/administration
  - in cancer patients, 2257
  - for delirium in cancer patients, 2258*t*
- ED and, 3341
- elimination disorder, 3655, 3669
- enuresis, 3655, 3669
- in functional GI disorders, 2206
- history, 2917
- illness anxiety disorder, 1838
- for insomnia, 2296
- intellectual disability and, 3515
- mania, 1612
- OCD, 1790
- pharmacogenomics of, 3455
- prolactin levels and, 2249
- psychiatric emergencies, 2628
- psychosis, 548
- psychotic disorders and, 1253
- schizophrenia, 3793, 4098
- schizophrenia, early-onset, 3732
- side effects of, 67, 2189, 2237, 2237*t*
- stereotypic movement disorder, 3671, 3685
- tic disorders, 2947, 3645, 3659
- TS, 2947

- Ritalin. *See* Methylphenidate
- Ritonavir, 568
  - drug interactions, 648
  - HIV/AIDS and, 3856
- Ritual abuse
  - dissociative disorders and, 1871–1872
  - as maltreatment, 3829
- Ritualistic behaviors, 4610
- Rituals
  - description of, 4610
  - obsessive-compulsive disorder and, 3699
- Rivastigmine (Exelon)
  - after TBI, 529

- Alzheimer disease treatment, 1207–1209
- amnestic disorder, 1231
- chemistry, 3044, 3044*f*
- cholinesterase inhibition, 4143
- dementia treatment, 1238
- dementia with Lewy bodies treatment, 1217
- dosage/administration, 3046
- drug interactions, 3046
- drug studies, 3045
- laboratory interferences, 3046
- organ/system effects of, 3045–3046
- for PD, 547–548
- pharmacologic actions, 3044–3045
- precautions/adverse reactions, 3046
- therapeutic indications, 3046
- uses of, 73

- Riviere, Joan, 2645
- Rizatriptan for migraine, 625
- Rizzuto, Ana-Marie, 2585
- RLS. *See* Restless leg syndrome
- RMTDs. *See* Rhythmic mid-temporal discharges (RMTDs)
- RNA, types of, 152*t*
- RNA editing, 159–160, 159*f*
- RNA-induced silencing complex (RISC), 51
- RNA-seq microarray, 439
- RNA splicing, 158
- RNAs, noncoding, 151–153
- Robert Wood Johnson Foundation
  - Building Health Systems for People with Chronic Illnesses, 4321
  - Chronic Care Initiatives in Health Maintenance Organizations, 4321
  - Mental Health Services Program for Youth, 4321
- Robins, Eli
  - dissociative identity disorder, 1882
  - somatic symptom and related disorders and, 1842
  - St. Louis criteria and, 1830
- Robinson, George, 4536
- Robo receptors, 52

- Robotics, electronic technology, 4279–4280
- ROC. *See* Receiver operating characteristic (ROC)
- Rodents
  - astrocyte to neuron ratios, 8
  - developing cortex, 43*f*
  - patches in, 29
  - substantia nigra in, 28
- Roe, D., 4538
- Roediger, Henry L., 2732
- Rogers, Carl R.
  - client-centered therapy, 2818
  - *Client-Centered Therapy,* 4539
  - humanistic psychology and, 2843
  - personality theory, 932
  - work of, 4610
- Rogers, Richard, 1917
- Rohypnol. *See* Flunitrazepam
- "Roid rage," 2493
- Role theory
  - psychodrama and, 2858
- Roles
  - changes
    - and cognitive-behavior therapy, 4180
    - as psychosocial factor, 4168–4169
  - description of, 4610
  - induction
    - addiction therapy and, 580
  - playing
    - description of, 4610
  - plays
    - for observation, 2686–2687
    - social skills training and, 2695
  - responsiveness
    - countertransference and, 2653, 2654
  - reversal
    - psychodrama, 2856–2857
- Romano, John, 1180

- Romazicon. *See* Flumazenil
- Romberg, Moritz, 2177
- Romberg sign
  - description of, 4610
  - in disorders of sensation, 459
- *Romeo, Youngberg v.,* 4336
- Roose, Steven, 2658
- Roosevelt, Franklin
  - health care plan considered by, 4335
  - social reforms of, 4311
- *Roper v. Simmons*
  - criminal competence, 4438
  - juvenile justice, 4439–4440
- Ropinirole (Requip)
  - bipolar depression, 3236–3237
  - chemistry, 3234, 3234*f*
  - combination pharmacotherapy and, 3324
  - dosage/administration, 3241, 3241*t*
  - drug interactions, 3240
  - history, 3233
  - MDD, 3236
  - pharmacological actions, 3235
  - precautions/adverse reactions, 3239, 3240*t*
  - receptor profile of, 67
  - RLS, 3238
  - therapeutic indications, 3239
- Rorschach, Herman, 4537
- Rorschach Test, 1002
  - description of, 1004, 1004*f*, 1004*t*, 2179, 4610
  - dissociative identity disorder and, 1905, 1906
  - interpretation of, 1005*t*
- Roscommon Family Study, 1571–1573
- Rose, Nicholas, 2839
- Rosenfeld, Herbert
  - psychodynamic psychotherapy, 2868
  - on psychotic transference, 2875
- Rosenzweig, Saul, 2837

- Rosiglitazone, 2247
- Ross, Colin, 1871, 1934
- Rostroventral medulla (RVM), 326
- Rosuvastatin (Crestor), 3131
- Rosy views, 2847
- Rotter, Julian, 930
- Rough-and-tumble play, 1991
- Roux-en-Y gastric bypass (RYGB), 2221, 2222*f*
- Royal Australian and New Zealand College of Psychiatrists (RANZCP), 971
- Royal Free Interview for Religious and Spiritual Beliefs, 2586
- RPEs. *See* Reward prediction errors (RPEs)
- Rsolution phase, 4609
- RT. *See* Radiation treatment
- "Rubber-hand illusion," 335
- Rubella, overview, 595–596
- Rubidium, 1609
- Rüdin, Ernst, 1438
- Rudy, Lester, 4402
- Rumination, description of, 4610
- Rumination disorder, **3631–3632, 3645–3646**
  - clinical features, 3631, 3645
  - course, 3631, 3645
  - description of, 2066
  - diagnosis of, 2073
  - differential diagnoses, 2077, 3631, 3645
  - epidemiology, 3631, 3645
  - etiology, 3631, 3645
  - generalized anxiety disorder and, 3714
  - historical literature on, 2067
  - of infancy, 4610
  - nosology, 2069
  - treatment, 2082, 3631–3632, 3645–3646
- Rush, Benjamin
  - on conduct disorder, 3612
  - on dissociative disorders, 1866
  - as founder of American psychiatry, 4290
  - *Medical Inquiries and Observations upon the Diseases of the Mind,* 4534

- work of, 4610
- Rutter, Michael
  - on autism, 3572
  - on psychiatric classification, 1159
- Ryff, Carol, 2843–2844
- RYGB. *See* Roux-en-Y gastric bypass (RYGB)

S

- *S*-Adenosyl methionine (SAMe)
  - adverse effects, 2544*t*
  - beneficial effects, 2549
  - in CAIM medicine, 3984
  - as complementary treatment, 2542
  - folate and synthesis of, 2547
- SAD. *See* Seasonal affective disorders (SAD)
- SAD PERSONS SCALE, 2615
- Sadism, description of, 4610
- Sadistic personality disorder, 4610
- Sadness
  - about, 1632
  - brain imaging, 1717
  - as universal distress signal, 1711
- Sadomasochistic relationships, 4610
- Safety
  - assault victim, 2541
  - CAM treatments, 2542–2544, 2544*t*
  - dissociative amnesia treatment and
    - acute dissociative amnesia, 1938
    - nonclassic dissociative amnesia, 1938–1939
  - dissociative identity disorder
    - practical management, 1922–1923
    - stabilization and, 1922
  - in long-term care, 4222
  - patient, psychiatry recertification and, 4428
- Safety agreements, 1922–1923
- Safety awareness and management
  - for dementia, 1200
  - during emergency child psychiatry evaluation, 3437–3438, 3440
  - during family therapy, 3774
- Saffron *(Crocus sativus),* 2554
- Safinamide, 3123
- SAGE. *See* Senior Action in the Gay Environment
- Saint Louis University Mental Status (SLUMS), 4226
- Sales, door-to-door, 4251

- *Salgo v. Leland Stanford Jr. University Board of Trustees,* 4433–4434
- Salicylate overdoses, 1018
- Salience networks, 418, 1819
- Saliency/reward/motivation network, 327–328
- Salivary gland cancers, 2267
- Salpêtrière Retardation Scale, 1636
- Saltatory conduction
  - axon, 134, 134*f*
  - definition of, 130–131
- *Salvia divinorum* (salvia extract), 1317, 1320, 1327
- Samadhi, 2579
- Sample space, 830
- Samples, description of, 4610
- Sampling
  - bias, mental health in older adults, 3962–3963
  - and psychological changes with normal aging, 3996
  - in psychological testing, 995
  - strategies, 816
- San Francisco earthquake, amnesia, 1887
- Sanatoriums, 4610
- Sanctioned aggression, 2472
- Sandler, Joseph, 2654
- Sandy Hook elementary school tragedy, 4298
- Sanguine humor, 1613
- *SAPAP3*
  - excoriation disorder and, 1796
  - TTM and, 1794
- Saperstein, Jerome, 1870
- Saphris. *See* Asenapine
- Sar, Vedat
  - on dissociation and dissociative disorders, 1871
  - dissociative identity disorder, 1894
- Sarcoidosis, 633
- Sartre, Jean Paul, 2821
- Satcher, David, 4294
- Satiation
  - in bulemia nervosa, 368

- conditionable, 366
- gastric volume and, 360
- postprandial satiety and, 360–363

- Satir, Virginia
  - family therapy and, 2749
  - strategic family therapy, 2753
- Satisfaction with Life Scale, 764–765
- Saturation, process of, 262
- Satyriasis, definition of, 1974, 2039, 4610
- Saunders, Cicily, 2289
- Savant skills
  - autism and, 3572
  - autism spectrum disorder and, 3582
- Savella. *See* Milnacipran
- Saving and Empowering Young Lives in Europe (SEYLE) study, 2614, 2617
  - Youth Aware of Mental Health (YAM) suicide preventive program, 2618
- Savoring, in positive psychology, 2850–2851
- Saxitoxin (STX), 136
- S100B, serum levels in TBI, 525
- Scaffolding proteins, 128–129
- Scale for Suicidal Ideation (SSI-C), 2615
- Scale for the Assessment of Negative Symptoms (SANS)
  - medication dose determination with, 4370
  - schizophrenia, 1416, 1416*t*, 1425
- Scale for the Assessment of Positive Symptoms (SAPS)
  - schizophrenia, 1416, 1416*t*, 1425
- Scales. *See* Rating scales
- Scalia, Antonin, 2671
- Scar effect, 3485
- Schäffer collaterals, 35, 79
- Schally, Andrew, 90
- Scharff, David, 2754
- Scharff, Jill Savege, 2672, 2754
- Schauer, Maggie, 1890
- Schedule for Affective Disorders and Schizophrenia (SADS), 2685
- Schedule for Affective Disorders and Schizophrenia for Children (K-SADS)
  - for bipolar disorder diagnosis, 3688, 3702

- overview, 3397–3398
- Schedule for Affective Disorders and Schizophrenia for School-Age Children (KSADS-PL), 1721
- Schedule of Recent Events, 2317–2318
- Schedule of reinforcement, 4610
- Scheiber, Steven, 4402
- Schein, Edgar, 1946
- Schema changes, 4185
- Schema therapy, 2668–2669
  - development of, 2696
  - hypnosis and, 2723
- Schemas, definition of, 1701, 2763
- Schemata, consciousness and, 691–692
- Schilder, Paul
  - depersonalization/derealization disorder, 1870
  - group psychotherapy and, 2737
  - on stereotypic movement disorders, 3668, 3682
- Schildkraut, Joseph J., 1713
- Schismatic family, 2874–2875
- Schizencephaly, 49
- Schizoaffective disorder, **1587–1590**
  - aggression, treatment, 2501
  - bipolar I disorder and, 1650
  - classification of, 1168, 1171–1172
  - comparative nosology, 1587, 1588*t*, 1589*t*
  - course and prognosis, 1590
  - defined, 3725, 4610
  - diagnosis and clinical features, 1588–1589
  - differential diagnoses, 1589, 3730
  - dissociative identity disorder and, 1916
  - epidemiology, 1587
  - etiology, 1587–1588
  - genetics and, 2529*t*, 2535*t*
  - history, 1587
  - mood disorders and, 1658–1659
  - schizophrenia, early onset and, 3730
  - treatment
    - carbamazepine, 3037

    - FGA, 3069–3070
    - lithium, 3092–3093
    - valproate, 3272
- Schizoaffective schizophrenia, 1658
- Schizoid fantasy, 4610
- Schizoid personality disorder
  - clinical criteria, 2152
  - comorbidity, 2152
  - complications, 2152
  - description of, 4610
  - differential diagnoses, 2152
  - epidemiology, 2152
  - familial pattern, 2152
  - genetics, 2152
  - impairment, 2152
  - as psychosis manifestation, 1572
  - Self-Transcendence and, 2163
  - sex ratio, 2152
- Schizophrenia, **1405–1573, 1595–1597, 4094–4103**
  - acute, bipolar disorders and, 1658
  - acute episodes, 1523–1525
  - adverse events associated with, 4530
  - after TBI, 538
  - aggressive behavior in, 2501
    - treatment
      - TMS, 3306
      - valproate, 3272, 3274
    - treatment of, 2501
    - twin studies of, 55
    - unspecified schizophrenia spectrum disorders, **1595–1597**
    - untreated, 1432
    - violence and, 1435, 2428
    - white matter density in, 13
    - worldwide burden of, 1425–1437
  - aging adults, 3957, 3967–3968
    - neuroimaging, 4013–4017, 4015*t*–4016*t*
  - ambulatory, 4560

- amino acid transmitters and, 82–83
- anticholinergics and, 2958, 2959
- ASD and, 1
- avoidant personality disorder and, 1725
- B-family vitamins and, 2548
- basket cells in, 17
- BDNF polymorphisms in, 102
- Bleulerian concept of, 1869
- *CACNA1F* mutations in, 314
- cancer patients with, 2266
- candidate genes for, 1625, 1626*t*
- cannabis and, 1307
- case study, 1424
- childhood-onset
  - FGAs and, 3071
  - neuroimaging of, 273, 3474–3475, 3474*f*
- chronic, 1466*f*, 1467*f*, 1468, 1472
- classification of, 1168, 1171–1172
- clinical manifestations of, 1148
- cognition
  - behavioral therapy for, 4373
  - impairments in, 4372
- cognitive-behavior therapy, 4186, 4192
- comorbidities, 1434
- coronary artery disease and, 2189
- costs, social and economic, 1425–1437
- course and prognosis, 1431–1434, 1433*t*, 1453–1455, 1454*t*, 1455*t*, 1456*t*
- D2 receptor availability, 300
- declarative memory, 739
- deficit syndrome, 1418–1419, 1418*t*
- defined, **1561–1574**
- delirium incidence in, 1185
- description of, 4610
- in developing countries, 799–800
- diabetes and, 2247
- diagnosis
  - criteria for, 1, 1407–1409, 1407*t*

- dimensional, **1561–1574**
- differential diagnoses, 1198, 1270, 1274–1275, 1427–1428
- discoordination hypothesis, 405
- dopamine transmission in, 301, 431
- early-onset, 3725–3733
  - cognition, 4095–4098, 4096
  - community integration and social functioning, 4097
  - comparative nosology, 3726
  - course and prognosis, 3731–3732, 3732*f*
  - course of illness, 4095–4096
  - defined, 3725, 3726*t*, 3727*t*
  - depression, 4096–4097
  - diagnosis and clinical features, 3728
  - differential diagnoses, 3730–3731, 3730*t*
  - epidemiology, 3726–3727
  - etiology, 3727–3728
  - future directions, 3733
  - history, 3725–3726
  - medical comorbidity, 4097
  - negative symptoms, 4096
  - pathology and laboratory examination, 3729, 3729*f*
  - positive and negative symptoms, 4096
  - positive symptoms, 4096
  - quality of life, 4097
  - treatment, 3732–3733, 4097–4098, 4132–4133
- EEGs in, 248–249
- electroconvulsive therapy, 4153
- ensemble coding hypothesis, 405
- environment and, 2529
- environmental factors, 1447–1451, 1448*t*
- epidemiology, 1428–1431, 1445–1457
- epidemiology (clinical)
  - cannabis use, 1451
  - developmental factors, 1449–1450, 1451–1453
  - environmental factors, 1449–1451
  - genetics, 1445–1450
  - overview, 1445

- epigenetic dysregulation, 446–447
- everyday functioning, 4038
- extrapyramidal symptoms, 170
- factors in development of, 54–55
- family, twin, and adoption studies for, 2528*t*
- fatty acids and, 2545
- first episode, 3690, 3704
- first-episode psychosis, 1435, 1466*f*, 1467*f*, 1469–1470, 1472–1473, 1505–1506, 1520
- fMRI of, 276
- functional brain connectivity in
  - graph theory analysis of, 424
  - literature, 424
- functional capacity/performance of, 4371
- future directions, 1424–1425, 1436–1437, 1456–1457, 1463, 1474, 1487–1488, 1500
  - drug development, 1530–1531
- γ-oscillations, 21
- gambling and, 1804, 1805
- gender differences, 4246
- genetic mapping of, 3450–3451, 3451*t*
- genetic overlap with bipolar disorder, 1628
- genetic selection, 773, 778
- genetics, 316–318, 1429–1430, 1437–1445, 1445–1450, 2529, 2529*t*
  - architecture of, 1444
  - biological factors, 1444
  - environment factors and, 1447–1449
  - epidemiology and, 1445–1450
  - heritability, 1446–1447
  - molecular, 1439–1444
  - neurocognition, 1504–1505
  - psychosis, 1564–1566
  - risk, 1467*f*, 1473
  - risk counseling, 1444–1445
  - studies, 1437–1439, 1438*f*, 1440*t*, 1441*t*, 1445–1446, 1471, 1499–1500, 1564–1567
- glial cell density in, 5
- group psychotherapy, 2746–2747, 2747*t*

- hallucinations in, 432–433
- heritability of, 273, 1610, 2535*t*
- histamine and, 2981
- history, 1407, 1407*t*, 1426–1427, 1463
- history and comparative nosology, 4095
- and HIV infection, 4271
- in HIV patients, 575
- illness/wellness management training, 4375
- and immigration, 806
- immune-CNS interactions in, 199–200
- incidence, 1428–1429, 1431*t*
- inflammation and, 202
- influenza virus and, 599
- insomnia and, 2086–2087
- interventions for, *4366t*
- late-onset
  - course and progress, 4101
  - diagnosis and clinical features, 4099–4100
  - differential diagnoses, 4100–4101
  - epidemiology, 4098
  - etiology, 4099
  - treatment, 4133
- medical health in, **1538–1548**
  - mortality rates, 1434–1435, 1538–1542, 1539*t*
  - oral, 1542
  - overview, 1538
  - prevention, 1544–1547, 1545*f*
  - recommendations, 1547–1548
  - risk factors, modifiable, 1542–1544, 1544*t*
- medication costs, 4522, 4524
- models of, 1456–1457
- mood disorders and, 1600, 1658, 1658*t*
- mortality rates with, 1434–1435, 1539–1542, 1539*t*, 4529
- MRI studies, 13, 272–273
- MRS in, 281
- negative symptoms of, 414, 1417–1419, 1418*t*, 1528
- neurobiology, 1473

- neurocognition, **1501–1512**
- neurocognitive deficits in, 4370
- neurodevelopment and, 1456
- neuroimaging
  - functional, **1475–1488**
  - molecular, **1488–1501**
  - structural, **1463–1475, 1465***f*, **1466***f*
- neuropathology of, **1457–1463**
- neurotensin and, 95–96
- niacin flushing response and, 2545
- nitric oxide and, 107
- omega-3 fatty acid supplements and, 116
- oscillatory abnormalities in, 436
- overview, 1405–1406, 1409–1410
- parenting children with, 3932–3933
- pathologic circuit in, 83*f*
- personality disorders and, 1571–1572
- pharmacogenomics of, 3455
- pharmacotherapy, combined, 3327–3328
- phenomenology of, **1406–1425**
- phenotypes of psychosis, **1512–1519**
- postpsychotic depressive disorder of, 1648
- prefrontal cortex, dendrite spines in, 16, 17*f*
- premorbid indicators, 1451–1453, 1569–1571
- prenatal injury in, 454
- prevalence, 1428, 1429*t*
- prodrome, 1505–1506, 1569–1571
- prognosis and course, 1506
- psychopathy and, 2491
- psychopharmacology, 3793
- psychosis, as defining dimension, **1561–1574**
  - biological factors, 1564–1569
  - defined, 1561–1562
  - diagnostic approach, 1562, 1563*f*
  - subclassification using, 1562–1564, 1563*f*
- psychosis, effects of cannabis on, 114–115
- psychosocial treatment, 1530

- QEEG analysis, 1038–1049, 1039*f*
- rating scales, 1409–1410
- recovery, **1548–1561**
  - advocacy, 1559–1560
  - defined, 1549–1551
  - evidence-based practices, 1551–1554, 1552*t*
  - history, 1548–1549
  - involuntary treatment, 1558–1559
  - journey of, 1554–1555
  - outcomes, 1551, 1551*f*
  - peer support specialist's role, 1557–1558
  - psychiatrist's role, 1555–1557
- rehabilitation, psychiatric, **1531–1538,** 1549
- rehabilitation for, 4362
- relapses, stress and, 2323
- (Sc) Scale, MMPI, 999*t*
- schizophreniform disorder, **1580–1582**
- self and, 332–334
- self-evaluations positively biased, 4372
- sensation, perception, and cognition in, 695–696
  - apprehension span, 696, 696f
  - attention, selective and sustained, 696
  - habituation, 695–696
  - modality shift effects, 695
  - reaction time crossover, 695
  - sensorimotor gating, 695–696
  - visual backward masking, 695–696, 695f
- services for families of persons with, 4350–4351
- signs and symptoms preceding, 4366
- social cognition deficits, 4373
- social learning as treatment, 4373–4374
- social skills deficits in, 4371
- stigma, 1559–1560
- studies
  - adoption, 1439, 1441*t*, 1446
  - association, 1565–1566
  - family, 1437–1439, 1438*f*, 1445–1446, 1471, 1564

- linkage studies, 1565
- population, 1427–1431
- twin, 1439, 1440*t*, 1446, 1564–1565
- substance abuse and, 1434, 1553
- suicidal behavior, 1528
- suicide and, 2612, 2613
- symptoms, 1120, 1407, 1407*t*, 1409–1424
  - agitation and hostility, 1422–1423
  - anxiety, 1421–1422
  - delusions, 1413–1415
  - depression, 1420–1421, 1421*t*
  - disorganization syndrome, 1419–1420
  - future directions, 1424–1425
  - hallucinations, 1410–1413
  - insight, lack of, 1423
  - motor, 1419–1420
  - negative, 1415–1419, 1416*t*, 1418*t*
  - overview, 1409–1410
  - psychotic, 1407*t*, 1410–1415
  - suicide, 1423
  - thought disorder as, 1420
  - violence and, 1422–1423
- and telemedicine, 4276
- threat detection, violence and, 2490
- transmission, 4542
- trazodone drug studies, 3249–3250
- treatment, 4548, 4553, 4554
  - benzodiazepine, 2997
  - buspirone, 3012
  - carbamazepine, 3037
  - combination pharmacotherapy, 3327–3328
  - ECT, 3286–3287
  - FGA, 3060, 3068–3070, 3071
  - involuntary, 1558–1559
  - lithium, 3092–3093
  - nefazodone, 3127
  - of neurocognition, 1509–1510

    - pharmacotherapy, **1519–1531**
  - twin studies, 55, 1439, 1440*t*, 1446, 1564–1565
- Schizophrenia Cognition Rating Scale (SCoRS), 4370
- Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB), 4370
- Schizophrenia spectrum disorders
  - classification of, 1158
  - functional connectivity abnormalities, 422–423
  - OCD *versus,* 1788, 1789
  - unspecified, **1595–1597**
- Schizophreniform disorder, **1580–1582**
  - comparative nosology, 1576*t*, 1577*t*, 1581
  - course and prognosis, 1582
  - defined, 3725, 4611
  - diagnosis and clinical features, 1581–1582
  - differential diagnoses, 1582
  - epidemiology, 1581
  - etiology, 1581
  - history, 1580–1581
- Schizotypal personality disorder, 2152–2153
  - classification of, 2173
  - clinical criteria, 2152
  - comorbidities, 2153
  - complications, 2153
  - description of, 4611
  - differential diagnoses, 2153
  - epidemiology, 2153
  - genetics, 2153
  - impairment, 2153
  - OCD and, 1788, 1789
  - predisposing factors, 2153
  - as psychosis manifestation, 1571–1572
- *Schloendorff v. Society of New York Hospital,* 4433–4434
- Schneider, Kurt, 1407, 1411
  - *Psychopathic Personalities,* 4541
  - on somatic presentations, 1829–1830
- Schneiderian first-rank symptoms (FRS), 4611
  - depersonalization/derealization disorder, 1944

- dissociative identity disorder and, 1912, 1916
- School-based suicide prevention, 2617–2618
- School consultations
  - confidentiality, 3883
  - ethical issues, 3883
  - framework, 3883
  - history, 3880–3881
  - legal issues, 3883
  - legislation, 3880–3881
  - overview, 3880
  - psychiatrist's role, 3882–3883
  - for school-based accommodations, 3888–3893
    - delivery models for mental health, 3888, 3891
    - evidence-based intervention programs, 3891–3892
    - overview, 3888, 3889–3891*t*
  - school outcomes from mental health problems, 3881–3883, 3882*t*
  - school personnel's role, 3881–3882
  - techniques, 3883
  - tiered response system, 3884–3888
    - for accommodation and modification determination, 3885–3886
    - eligibility determination, 3886–3888, 3887*t*
    - interventions, 3884–3885, 3886*t*
    - screenings, 3884, 3885*t*
  - for violence potential, 3892–3893
- School interventions
  - with adolescents, 3815
  - terrorism and interventions by, 3921, 3922*t*
- School phobia, 4611
- Schools
  - children with TBI and, 653
  - Individuals with Disabilities Education Act
    - children with TBI and, 653
- Schopenhauer, Arthur, 928–929
- Schreber case, 4611
- Schultz, Wolfram, 414
- Schultze, Johannes Heinrich, 2823
- Schwann cells, 4*t*

- Schwartz, David, 1799
- Schwartz, Mark, 2730
- Schweitzer, Albert, 2593
- Science, psychodrama and, 2856
- *Science and Human Behavior (*Skinner), 4540
- Scientifically validated treatments, 4347
- Scientist–practitioner model, 2684
- Sclerodactyly, 456
- SCN. *See* Suprachiasmatic nucleus (SCN)
- Scopace. *See* Scopolamine
- Scopolamine (Scopace), **2960–2962**
  - chemistry, 2960, 2960*f*
  - clinical studies, 2961–2962
  - dosage/administration, 2961
  - insomnia and, 2086
  - pharmacological actions, 2960–2961
  - precautions/adverse reactions, 2961
  - therapeutic indications, 2961
  - withdrawal, 2086
- Scopophilia, description of, 4611
- Scorpion toxins, 136
- Scotoma, 4611
- Scott, William, 2843
- Scrapie (PrP^{Sc}), 602, 603
  - cause of, 601*t*
- Scratching, dermatitis and, 2340
- Screen for Child Anxiety Related Emotional Disorders (SCARED), 3401
- Screen memory, 4611
- Screener and Opioid Assessment for Patients with Pain (SOAPP), 2393
- Screening
  - after arrest, 4394
  - alcohol abuse, 4112, 4112*t*
  - by correctional systems, 4305
  - description of, 4611
  - in jails and prisons, 4459–4460
  - neuropsychological testing, 988
  - for substance abuse, 4115

- Screening, Brief Intervention, Referral and Treatment (SBIRT) initiatives, 3858
- Scrotoplasty, 2034
- Scrotum, sexual response cycle, 1960*t*
- *Scutellaria lateriflora,* 2554
- SD. *See* Standard deviation
- SDS. *See* Zung Self-Rating Depression Scale
- Sea anemone toxins, 136
- Searles, Harold
  - IPT and, 2775
  - on psychosis, 2868
- Seasonal affective disorders (SAD), 3939
  - criteria, 212
  - description of, 4611
  - light therapy for, 2558
  - treatment
    - bupropion, 3004, 3005
    - MAOIs, 3121
- Seasonal patterns. *See* Seasonal affective disorders (SAD)
- Seasonality
  - cyclic depressions and, 1651
  - depressive syndrome and, 1639
  - in humans, 211–212
  - melatonin and, 211
  - mood disorders and, 1616
  - treatment selection and, 1679
- Seclusion
  - during child emergency psychiatric evaluation, 3440
  - for inpatients, 3799
  - restraint and, in jails and prisons, 4470–4471
- Secobarbital (Seconal), 2986
- Seconal. *See* Secobarbital
- Second-generation antipsychotics (SGAs). *See also specific drugs*
  - antidepressants combined with, 3324
  - black box warning, 3080
  - contraindications, 2520–2521
  - in elderly, 3080
  - mood stabilizers combined with, 3327

- movement disorders induced by, 2936
- Second-line antidepressants, 1678, 1680
- Second messengers
  - G-protein modulation of, 141
  - mood disorders and, 1715
  - systems, 122–123
- Secondary emotions, 2472
- Secondary gains, 1834, 4611
- Secondary narcissism. *See* Narcissism
- Secondary prevention. *See also* Preventive psychiatry
  - in public health, 4303
- Secondary process thinking, 4611
- Secondary retarded ejaculation, 4611
- Secondary sexual characteristics, 4611
- Secondary structuring, dissociative process and, 1899, 1906–1907
- Secondhand smoke, 1347
- Secretan, Henri, 1847
- Secure attachment
  - aggression and, 2486
  - description of, 4611
- Security, Sullivan on psychosis and, 2875
- Sedation
  - antihistamine, 2982
  - antipsychotic drug effects, 4137
  - benzodiazepine, 2997, 2998
  - insomnia and, 2253
  - mirtazapine, 3110
  - nonbarbiturate sedative-hypnotic agents, 2989
  - opioid withdrawal treatment, 1370
  - respiratory symptoms and
    - in dying patients, 2297
- Sedative-, hypnotic-, anxiolytic-related disorders, **1374–1390**
  - abuse patterns, 1381–1382
  - barbiturates-related disorders, 1383–1384, 1384*t*, 1389, 1389*f*
  - benzodiazepine-related disorders, 1382–1383, 1382*t*, 1386–1387, 1386*t*
  - benzodiazepine use disorder, 1378–1379, 1378*f*
  - defined, 1374

- dependence, 1384
- detoxification, 1387
- diagnosis and clinical features, 1380–1386
- epidemiology, 1378–1380, 1378*t*, 1381*t*
- etiology, 1374–1378
- human studies, 1377–1378
- hypnotic-related disorders, 1384–1386
- intoxication, 1383–1384
- intoxication, induced, 1380, 1382
- neurocognitive disorders, induced, 1381
- neuropharmacology, 1374–1377, 1376*f*
- overdose, 1380–1383, 1389
- psychotherapy, 1388
- psychotic disorders, induced, 1381
- treatment, 1386–1389
- withdrawal, 1380–1386, 1386*t*, 2987
- withdrawal delirium, 1381

- Sedative-hypnotic drugs. *See also* Nonbarbiturate sedative-hypnotic agents; Nonbenzodiazepine sedative-hypnotic agents
  - anesthetic agents, 2990
  - history, 2990
- Sedatives
  - description of, 4611
  - intoxication, 2630
  - withdrawal, 2632–2633
- See, domestic abuse
- Segal, Hannah
  - psychodynamic psychotherapy, 2868
  - on psychosis, 2875
  - on symbolic equivalence, 2871–2872
- Segmental duplications (SDs). *See* Copy number variants (CNVs)
- Segregation analysis, 4019
- Segregation housing, 4471–4472
- Seizure disorders
  - absence seizures, 244
  - ADHD and, treatment of, 3229
  - brain tumors and, 2267
  - DA receptor antagonists and, 3077

- dissociative amnesia *versus,* 1936
- EEG findings, 244
- treatment
  - anticonvulsants, 2967, 2968, 2970, 2972–2977
  - barbiturates, 2986
  - benzodiazepine, 2997
  - carbamazepine, 3029, 3029*t*
  - lamotrigine, 3084, 3085
  - valproate, 3276

- Seizures
  - ADHD comorbidity, 3595
  - autistic disorder and, 3582–3583, 3583*f*, 3584
  - in psychiatric disorders, 1149
- Seizure(s). *See also* Complex partial seizures (CPS); Epilepsy
  - absence type, 244, 650, 4557
  - bupropion and, 3007
  - chemotherapeutic agents and, 2273*t*
  - in children, 648–651
  - classification of, 648*t*
  - complex partial, mania without prior bipolarity and, 1649
  - depersonalization and, 1891
  - description of, 4611
  - ECT and, 3282–3283
    - duration of, 3292
    - seizure threshold, 3290–3291
  - malingered, 516*t*
  - modes of onset, 508
  - nonmalingered, 516*t*
  - psychogenic nonepileptic, 514–517, 650–651
  - syncope *versus,* 2193
  - tricyclic and tetracyclic antidepressant effects on, 3260
- Seldane. *See* Terfenadine
- *Selected Correspondence of Karl Menninger* (Menninger), 4546
- *Selected Writings of John Hughlings Jackson* (Jackson), 4537
- Selection bias, 4611
- Selective dissociative amnesia, 1880, 1881*t*
- Selective estrogen receptor modulators (SERMs), 165

- Selective inattention, 4611
- Selective interventions, 4303–4304
- Selective mutism (SM), **3719–3724**
  - behavioral therapy, 3722–3723
  - child/adolescent
    - treatment, 3758
  - cognitive-behavioral therapy for, 3758
  - comparative nosology, 3719
  - course and prognosis, 3722
  - defined, 3719, 4611
  - diagnosis and clinical features, 3720–3721
  - differential diagnoses, 3548, 3557, 3721–3722
  - epidemiology, 3719–3720
  - etiology, 3719–3720
  - family therapy and, 3723
  - history, 3719
  - pharmacotherapy, 3723
  - speech and language therapy, 3723
  - treatment, 3722–3723
  - as unspecified communication disorder, 3367
- Selective serotonin-norepinephrine reuptake inhibitors (SNRIs), **3146–3153**
  - antidepressant discontinuation syndrome, 2943
  - antidepressant efficacy, 3151, 3151*f*, 3152
  - chemistry, 3147, 3147*f*
  - discontinuation syndrome, 3150
  - drug interactions, 3150
  - ED and, 3340–3341
  - history, 3146–3147, 3150, 3151–3152
  - hypertension, 3150
  - mirtazapine combined with, 3108–3109
  - overdose, 3150
  - pharmacology, 3147–3148, 3150, 3151*f*, 3152
  - during pregnancy and lactation, 2519–2520
  - therapeutic indications, 3148–3149
  - tolerability, 3149–3152
- Selective serotonin reuptake inhibitors (SSRIs), **3153–3170,** 4546
  - ADHD, 3601–3602

- adverse effects of, 2356*t*
- aggression, 2493
  - conduct disorder, 2501
  - intermittent explosive disorder, 2501
  - intimate partner violence, 2503
- aggression threshold and, 2493
- antidepressant discontinuation syndrome, 2943
- antimicrobial activity, 599
- for anxiety, 69
  - perinatal, 2515
- anxiety disorders, 1253, 1743*t*, 1746, 1779–1780, 1782, 1782*t*, 4129–4130
  - in children and adolescents, 1784, 3792
- for atypical depression, 2171
- autism spectrum disorder and, 3586
- BDD, 1792
- biological or chemical terrorist attacks and, 2572
- bipolar depression, 1690–1691, 1693, 3155
  - combination therapy, 1692*t*
- bipolar disorders
  - GAD and, 1698
  - panic disorder and, 1699
  - PTSD and, 1699
- breast-feeding and, 3168
- breast milk excretion of, 3168
- bupropion augmentation, 2057
- cardiovascular effects of, 2197*f*
- chemistry, 3155, 3155*f*, 3158–3162
- in children, 3168
- classification, 2918
- combination pharmacotherapy and, 3323–3325
- dementia treatment, 1201–1202, 1238
- depersonalization/derealization disorder, 1945
- depression, 1678
- for depression, 4124
- for depressive disorders, 69
  - in cancer, 2261, 2265
  - child/adolescent treatment, 3378–3379, 3379*t*

- description of, 4611
- discontinuation syndrome, 3156–3157
- dissociative identity disorder, 1927*t*
  - PTSD symptoms, 1928
- dosing/administration, 3165–3166
- drug interactions, 3162–3165, 3163*f*, 3164*t*
- duloxetine compared with, 3151, 3151*f*
- ED and, 3340–3341
- EEG effects, 253
- efficacy, 3155–3156
- enuresis, 3655, 3669
- excoriation disorder, 1796, 1797
- for excoriation disorders, 2352
- GAD, 1783, 3154
- gambling disorder, 1808
- geriatric depression, 4072–4073
- history, 2917, 3153
- illness anxiety disorder, 1838
- laboratory interferences, 3158
- lithium or valproate combined with, 3091, 3323
- major depression and MDD, 3153–3154, 3155, 3168–3169
- mirtazapine combined with, 3108–3109
- OCD, 1789, 3154, 3155, 3701, 3792
  - child/adolescent, 3700–3701, 3701*t*
- in pain syndromes, 2396
- panic disorder, 1747, 3154
- paraphilic disorders, 2021
- pharmacodynamics, 3154, 3162–3164
- pharmacokinetics, 3158–3159, 3159*t*, 3164–3165, 3164*t*
- pharmacological actions, 3159–3162
- PMDD, 3154
- post-TBI depression, 532–533
- pregnancy and, 3167–3168
- during pregnancy and lactation, 2519–2520
- PTSD, 1823, 1824, 3154
- PTSD after TBI, 537
- pyromania, 2113

- SAD, 1783, 1784
- safety, child/adolescent, 3717–3718
- selective mutism in children, 1784
- selective mutism treatment, 3723
- separation anxiety treatment, 3711–3712
- sex addictions, 2057
- sexual dysfunction and, 1779*t*
- side effects, 572, 3378
- side effects among, 1679*t*
- social anxiety, child/adolescent treatment, 3717
- social anxiety disorder, 3154
- specific phobias, 1784
- suicidality and, 2616, 3166–3167, 3378, 3718
- therapeutic indications, 3153–3155, 3155*t*
- thyroid hormone combined with, 3244–3245, 3245*t*, 3323
- tic disorder, 3646, 3660
- tolerability, 3156–3158, 3156*t*
- for trichotillomania, 2353
- venlafaxine compared with, 3148–3149, 3149*f*
- vortioxetine classified as, 3168–3169

- Selegiline (Deprenyl, Eldepryl)
  - ADHD, 3123
  - for apathy in PD, 548
  - chemistry, 3113*f*
  - depression, 3120
  - history, 3111
  - mechanisms of action, 3114
  - pharmacological actions, 3115–3116
  - precautions and warnings, 3118
  - sexual dysfunction and, 1972, 1980
  - transdermal, 3116, 3118, 3120
- Self
  - as agent, 335
  - basic science of, **329–339**
  - clinical perspective, 332–334
  - definition of, 4611
  - development of, 3360–3361, 3365–3366, 3370

- evolutionary perspective, 331
- human developmental perspective, 331–332
- multidimensionality of, 331–332
- nonself and, 334–335
- others *versus,* 336–337, 2895
- psychology of
  - evolving concepts, 887–890
  - Kohutian psychology, 887
  - neuropsychoanalysis, 889–890
  - relational and intersubjective approaches, 888–889
- scientific investigations, 334–338
- sense of, disturbances in, 1132–1133

- Self-actualization
  - definition of, 4611
  - humanistic psychology and, 2843
  - Maslow personality theory, 932
- Self-analysis, definition of, 4611
- Self-attack, self-shaming and, 1907
- Self-Awareness
  - consciousness levels and, 2141
  - definition of, 4611
  - development of, 2167–2168
  - growth of, 2139
    - approaches to, 2173–2174
  - intuitive, 2139
  - path to Well-Being, 2167
  - psychiatric symptomatology and, 1124
  - self-transcendence and, 2139
  - sense of truth and, 2141
  - stages of development, 2167–2168, 2167*t*
- Self-care functions
  - addictive disorders and, 2049
  - development of
    - therapist facilitation of, 2059
- Self-control therapies, 2697
- Self-critical thinking, 1664
- Self-deception, 1124

- Self-defeating personality disorder, 4611
- Self-development, 2723–2724
- Self-Directedness
  - age and, 2167
  - behavior and, 2138
  - enhancement of, 2142
  - executive competence and, 2135
  - neuroticism and, 2145
  - personality disorders and, 2162*t*, 2163
  - TCI and, 2142
  - TCI scales, 2164*t*
  - temperament and, 2142
- Self-discovery, 4611
- Self-disorders. *See also* Narcissistic disorders
  - Kohut on, 2806
  - primary, 2808
  - secondary, 2808
- Self-efficacy, pain and, 2397
- Self-esteem
  - during adolescence, 3369, 3370
  - in children with epilepsy, 651
  - depression and loss of, 1606
  - depression model and, 1662
  - impaired regulation of, 1662
- Self-evaluation, 2697
- Self-fulfilling prophecies, 4611
- Self-governance
  - addictive disorders and, 2049
  - branches of, 2143*t*
  - development of, 2059
  - mental, 2142
  - therapist facilitation of, 2059
- Self-harm behaviors
  - child/adolescent, 3767
  - dissociative identity disorder and, 1904, 1904*t*, 1913
    - safety agreements, 1922–1923
  - in epilepsy patients, 517

- Self-help groups
  - for alcohol use disorders, 1279
  - beneficial effects, 2738
  - cancer patient, 2254
  - definition of, 4611
  - dissociative identity disorder, 1930
  - for stimulant use disorders, 1288–1289
  - substance abuse, child/adolescent, 3768
  - for substance-related disorders, 1288–1289
- Self-hypnosis
  - for acne, 2345
  - factitious disorder, 1863
  - hypnotic induction and, 2709–2710
- Self-image
  - acne and, 2345
  - BIF treatment and, 2438
  - burn injuries and, 2380
  - definition of, 4611
  - gender and, 2506
  - and mental health, 760
  - skin diseases and, 2333–2335
- Self-immolation, prevalence of, 2383
- Self-importance, sense of, 2050
- Self-injurious behaviors (SIBs), 3142
  - autistic disorder and, 3582, 3585
  - stereotypic movement disorders, 3667–3673, 3681–3687
- Self-mindfulness, 2059
- Self-monitoring
  - behavior therapy and, 2686, 2686*t*, 2687*t*, 2697
- Self-mutilation
  - depersonalization and, 1902
  - nonclassic dissociative amnesia and, 1934
- Self-observation techniques, 2325–2326
- Self organization
  - in attention-deficit/hyperactivity disorder, 697
  - in autistic spectrum disorder, 697
  - in bipolar 1 disorder, 699

- Self-processing, 337–338
- Self-protective processes, 2060
- Self-psychology in group therapy, 4204
- Self-reference, 4611
- Self-reflection, 692–693
- Self-regulation, impaired
  - addictive processes, 2049–2050
  - in ND-PAE, 637–638
- Self-reinforcement, 2697
- Self-report questionnaires, 941, 2685–2686
- Self-representation
  - historical considerations, 330–331
  - hypnosis and, 2708
  - neuroscientific theories of, 337–338
  - study of, 329
- Self-schema, skin disease and, 2334–2335
- Self-states
  - dissociative identity disorder treatment and, 1923–1924, 1926
  - distinctiveness of, dissociative disorders and, 1879
  - elaboration of, dissociative disorders and, 1906–1907
  - non–possession-form dissociative identity disorder and, 1912
- Self-systems, definition of, 4611
- Self-Transcendence
  - age and, 2167
  - defense mechanisms and, 2138
  - executive competence and, 2135
  - insight measured by, 2143
  - paranoid personality disorder and, 2163
  - personality disorders and, 2162*t*
  - schizoid personality disorder and, 2163
  - self-awareness and, 2139
  - TCI scales, 2164*t*
  - temperament and, 2142
- Seligman, Martin
  - on action research, 2842
  - humanistic psychology and, 2843
  - positive psychology, 2845, 2852, 2853

- Selman, Robert, 3369
- Selye, Hans, 167
  - on general adaption responses, 2336
  - on stress, 2315
- SEM. *See* Standard error of the mean
- Semantic dementia (SD), 550
- Semantic memory, 2129
  - description of, 2139
  - dissociative amnesia and, 1879
- Semap. *See* Penfluridol
- Sematic-pragmatic disorder, 3551
- Seminomas, treatment, 2268
- Semiotic function, 707
- Semiparametric regression, 843
- Senescence, definition of, 4611
- Senile dementia, 4611
- Senior Action in the Gay Environment (SAGE), 4259
- Senium, definition of, 4611
- Sensate focus exercises, 1976
- Sensate focus learning, 4611
- Sensation
  - abnormalities of, 459
  - definition of, 4611
  - in information processing, 689
  - and perception, 690
- Sensation-intuition polarity, 913
- Sensation seeking, risk behaviors and, 576
- Sense of self, 4611
- Sensitivity
  - in attention processing, 690
  - definition of, 4611
  - training groups, 4612
- Sensitization, definition of, 166
- Sensor technology, 4277–4278, 4277*t*
- Sensorimotor
  - gating, in schizophrenia, 695–696
  - period of intelligence, 706–707

- Sensorimotor period. *See also* Piaget, Jean
  - of cognitive development, 3359
- Sensory abnormalities, 3582
- Sensory aphasia, 4612
- Sensory deprivation
  - description of, 4612
  - transient depersonalization, 1887
- Sensory disturbances, brain tumors and, 493
- Sensory extinction, 4612
- Sensory food aversions feeding disorder, **3625–3626, 3639–3640**
  - clinical features, 3625, 3639
  - course, 3625–3626, 3639–3640
  - diagnosis, 3625, 3639
  - differential diagnoses, 3625, 3627, 3639, 3641
  - etiology, 3625–3626, 3639–3640
  - treatment, 3625–3626, 3639–3640
- Sensory functions, development of, 3362–3363
- Sensory impairment in aging, 3975, 4007
- Sensory inattention. *See* Sensory extinction
- Sensory integration therapy (SIT), 3543
- Sensory/motor functioning, 3435–3436
- Sensory nerves, neuropathies, 628–629, 629*t*
- Sensory perception, 2337
- Sentence Completion Test (SCT), 1004*t*, 1006, 4612
- Sentencing, of older adult, 4233
- Separateness, feelings of, 2140
- Separation anxiety disorder (SepAD)
  - burn injuries in children and, 2385
  - cognitive behavior therapy for, 3742–3743
  - comparative nosology, 3708
  - definition of, 1728–1729, 3708
  - description of, 4612
  - diagnosis of, 1728–1729, 3709–3710
  - differential diagnoses, 1729, 3710, 3715
  - DSM-5 on, 1723
  - epidemiology, 1729, 3708
    - adults, 1731–1732, 1732*t*

- etiology, 3708–3709
- history and comparative nosology, 1729
- management in children and adolescents, 1784
- neuroimaging, 1765
- psychopharmacology, 3711–3712
- psychotherapy, 3710–3711
- treatment, 3710–3712

- Separation-individuation, 4612
- Separation-individuation process
  - autistic phase, 878–879
  - hatching, 879
  - object constancy, 879
  - practicing, 879
  - rapprochement, 879
  - symbiotic phase, 879
- Septal area, anatomy of, 35
- Septal areas, 2477
- September 11, 2001 terrorist attacks
  - PTSD and, 2560, 2568
  - response integration efforts after, 2574
- Septic shock, 183–184
- Sequenced Treatment Alternatives to Relieve Depression (STAR*D), 1678
- Sequencing studies, 1443–1444
- Sequential Brief–Adlerian psychodynamic psychotherapy, 2617
- Sequential Treatment Alternatives to Relieve Depression (STAR*D) trial, 2917–2918, 2923, 3005, 3010, 3109, 3162, 3163*f*, 3323–3325
- Serax. *See* Oxazepam
- Serentil. *See* Mesoridazine
- Serine deficiency syndromes (SDSs), 664
- Seroquel. *See* Quetiapine
- Serotinin, eating disorders and, 368
- Serotinin–dopamine antagonists (SDAs), 4612
- Serotonergic agents
  - tick disorders, 3646, 3647*t*, 3660, 3661*t*
- Serotonergic hallucinogens (SHs), 1315–1316, 1320
  - about, 2886–2887
  - contraindications, 2890
  - limitations, 2889

- Serotonergic system
  - anxiety disorders and, 1743*t*, 1745–1746
  - fear and anxiety and, 1758
  - gambling disorder and, 1807
  - intimate partner violence and, 2487
  - neurons, 1818
  - OCD and, 1787
- Serotonin
  - serotonergic system and
    - ADHD and, 3589
    - autism spectrum disorder and, 3579, 3586
    - delirium pathophysiology, 1181, 1182
    - disruptive behavior disorders, 3615
    - schizophrenia, 1495
    - tic disorders, 3646, 3660
- Serotonin (5-HT), **68–71**
  - aggression and, 2479–2480
  - aggression-damping influence of, 2476
  - anatomy, 68
  - anxiety disorders and, 1742, 1745–1746
  - appetite and, 364
  - brain synthesis of
    - hypothyrodism and, 2240–2241
  - buspirone and, 3009
  - depression and, 1607–1608
  - description of, 4612
  - dissociative disorders and, 1889
  - functions, 68
  - immunological effects of, 185*t*
  - metabolism of, 69
  - neurotransmission by, 141
  - panic disorder and, 1761–1762
  - PET/SPECT studies of, 301–302
  - receptors, panic disorder and, 1762
  - receptors for, 69–71
    - availability of, 301–302
  - SCN function and, 205

- in social phobias, 703
- synthesis of, 68–69, 68*f*
- transporters, 69
- verbal therapies' effects on, 2812

- Serotonin and norepinephrine reuptake inhibitors, 4130
- Serotonin norepinephrine reuptake inhibitor. *See* Selective serotonin-norepinephrine reuptake inhibitors
- Serotonin norepinephrine reuptake inhibitors (SNRIs)
  - anxiety disorders and, 1743*t*, 1744, 1779–1780, 1782*t*
  - BD and generalized anxiety disorder, 1698
  - depersonalization/derealization disorder, 1945
  - depression, 1678
  - description of, 4612
  - dissociative identity disorder, 1928
  - GAD, 1783
  - for geriatric depression, 4073
  - OCD, 1785
  - in pain syndromes, 2395
  - sexual dysfunction and, 1779*t*
  - side effects among, 1679*t*
- Serotonin reuptake inhibitors (SRIs)
  - BDD, 1792
  - OCD, 1787, 1789, 1790
  - during pregnancy and lactation, 2519–2520
- Serotonin reuptake transporter (SERT), 302
- Serotonin synapses, model of action, 70*f*
- Serotonin syndrome
  - clinical criteria, 635
  - clinical manifestations of, 1148
  - description of, 635, 4612
  - MAOIs and, 3117–3118, 3118*t*, 3162, 3163–3164
  - SSRIs and, 1779, 3162, 3163–3164
  - St. John wort and, 2552
  - vilazodone and, 3346
- Serotonin transporter (SLC6A4 or 5-HTT)
  - anxiety disorders and, 1739, 1746
  - brain, psychodynamic psychotherapy and, 2813
  - depression risk, 1710, 1710*f*

- gene association studies, 1625
- PTSD and, 1820

- Serotonin–dopamine antagonists. *See also* Second-generation antipsychotics
- Serotoninergic system
  - neurons
    - mood disorders and, 1714
- Serpalan. *See* Reserpine
- SERT. *See* Serotonin reuptake transporter (SERT)
- Sertindole (Serdolect), 2237, 2237*t*
- Sertraline (Zoloft)
  - anxiety disorders, 1746, 1779–1780
    - GAD, 1783
  - BD and generalized anxiety disorder, 1698
  - breast milk excretion of, 3168
  - chemistry, 3160
  - in children, 3168
  - for chronic idiopathic pruritus, 2348–2349
  - depression, 1681
    - in cardiac patients, 2196
  - dissociative identity disorder
    - PTSD symptoms, 1928
  - dosing/administration, 573*t*, 3166
    - sexual dysfunction, 1979
  - drug-drug interactions
    - in cancer, 2261
  - drug interactions, 3164
  - EEG effects, 252
  - excoriation disorders, 2352
  - generalized anxiety disorder, child/adolescent, 3714
  - OCD, child/adolescent, 3701
  - pharmacokinetics, 3164, 3164*t*
  - pharmacological action, 3160
  - for post-TBI depression, 532
  - postpartum depression, 2515
  - pregnancy and, 3167
  - PTSD, 1823
  - sexual dysfunction and, 1779*t*

- side effects, 573*t*, 1679*t*
- targets for, 69
- tic disorder, 3646, 3660
- tolerability, 3157

- Serum levels. *See* Blood levels
- Services. *See* Supportive services
- Serzone. *See* Nefazodone
- Setting, hallucinogen-assisted psychotherapy and, 2888
- Seven-Factor Psychobiological Personality Model, 2161
- Severe mental illness (SMI)
  - dissociative disorders and, 1888
  - obesity and, 2214–2215
  - rehabilitation for, 4362
- Severe mental retardation. *See* Mental retardation
- Severity, 1265
- Severity measures, dementia, 1196
- Sex
  - definition of, 1984, 2023
  - legal issues, 1963
  - mood disorders and, 1611–1612, 1611*f*
- Sex addiction, **2039–2064**
  - clinical examples, 2042–2043
  - comorbidities, 2044
  - course of illness, 2044
  - diagnostic criteria, 2042–2043
  - differential diagnoses, 2043–2044
  - epidemiology, 2044
  - familial patterns, 2044
  - prognosis in, 2063–2064
  - reinforcers of, 2063
  - symptom containment, 2054
  - urge handling in, 2054
- Sex care workers, HIV infected, 584
- Sex differences. *See also* Gender; Gender differences
  - anxiety disorders, 1734
  - appetite and, 367
  - biological, 1957–1958

- in desire, 1959–1961
- dopamine signaling and, 383–384
- eating disorders and, 368–369
- in erotic stimuli, 1959–1961
- genetic, 1958
- hormonal, 1957–1958
- in substance abuse disorders, 382–384

- Sex disorders
  - classification of, 1159
  - description of, 4612
- Sex drive, androgens and, 1980
- Sex education in rehabilitation, 4375
- Sex histories, patient histories and, 1963–1965, 1964*t*
- Sex offenders
  - aggressive behavior in, 2503
  - castration of, 2022
  - treatment, 2503
- Sex organs
  - innervation of, 1955
  - responses to nerve impulses, 1955*t*
- Sex reassignment surgery (SRS), 2023
- Sex steroids, antidepressants combined with, 3324–3325
- Sex trafficking, 3839
- Sexual aberrations, 4612
- Sexual abuse
  - adults, 2538–2542
    - domestic violence, 2539–2540
    - intimate partner violence, 2539–2540
    - rape, 2538–2539
    - sequelae of violence and assault, 2540
    - treatment issues, 2540–2541
    - types, prevalence, and risk factors, 2538
    - violent crime, 2538
  - amnesia, 1886
  - autobiographical memory deficits and, 1886
  - child/adolescent
    - evaluating, 3391

    - extrafamilial, 3833
    - group therapy and, 3762–3763, 3769
    - intrafamilial, 3832–3833, 3841
    - as maltreatment, 3829, 3831, 3832, 3837
    - parent interviews and, 3834
  - conduct disorder and, 3616
  - depression and, 1914
  - dissociative disorders and, 1870
  - encopresis and, 3660, 3674
  - GAD and, 1728
  - girls, dissociation and, 1902
  - pathological dissociation and, 1884
  - prostitutes and, 1889
  - PTSD, borderline PD and, 1713
  - rape, 1873, 2538–2539
  - religious group scandals, 1873
  - violence and, 2425
- Sexual arousal disorders, 4612
- Sexual assault. *See also* Rape
  - CBT for PTSD prevention, 1824
  - in jails and prisons, 4460, 4466
  - sequelae of, 2540
- Sexual assault of elders, 4251
- Sexual attraction, race/ethnicity and, 2001
- Sexual aversion disorder, 4612
- Sexual behavior
  - during adolescence, 3371–3372
  - animal models of, 324–325
  - behavior mastery training, 2053–2054
  - CNS and, 1956–1957
  - compulsive, 1973
  - disinhibited, 2043
  - disorders of, 2041–2044
  - elements of, 1959
  - male, SDN-POA and, 1993–1994
  - paraphilic disorders, 2013
  - risk management, 2053–2054

- selection of, 2052–2053
- sexual learning and, 1961–1963
- of sexually abused children, 3832
- symptomatic management, 2053–2057
- talking about, 2007–2008

- Sexual compulsivity, 2041
- Sexual delusions, 4573, 4612
- Sexual desire disorders, 4612
- Sexual deviation. *See* Paraphilia
- Sexual dimorphism, neuroanatomical, 1993–1994
- Sexual drive
  - description of, 4612
  - disturbances of, 1141–1142
- Sexual dysfunction, **1953–1982**
  - animal models of, 325
  - anxiety disorder agents and, 1779, 1779*t*
  - in cancer patients, 2275–2276
  - definitions of, 1965
  - depressive syndrome and, 1639
  - description of, 4612
  - features of, 1965
  - general medical conditions and, 1970–1971
  - hypnosis and, 2730
  - lithium and, 3098
  - physiological tests, 1971
  - psychological states and, 2508
  - rating scales, 1965
  - SSRI effects on, 3156
  - stimulant-induced, 1286
  - substance/medication-induced, **1258,** 1972–1975
  - trazodone drug studies, 3251
  - treatment of, 1975–1981
    - behavior therapy, 1977
    - behavioral exercises, 1976
    - biological methods, 1978–1981
    - bupropion, 3004
    - buspirone, 3012

        - dual-sex therapy, 1975–1976
        - goals of, 1976
        - group therapy, 1977
        - hypnotherapy, 1976–1977
        - integrated sex therapy, 1977–1978
    - unspecified, 1973–1975
    - vilazodone effects on, 3345
- Sexual excesses, manic syndrome and, 1641
- Sexual fantasies
    - common, 1975*f*
    - gender differences, 1960
- Sexual function
    - animal models of, 324–325
    - of cancer patients
        - psychological issues with, 2275–2276
    - DA receptor antagonists and, 3076–3077
    - as psychosocial factor, 4167
    - sleep and, 344
- Sexual identities
    - description of, 1958, 1983
    - gender and, 2506
    - homosexual orientation constant, 1998*f*
- Sexual impulse disorders, 2014
- Sexual masochism, 2013. *See also* Masochism
- Sexual masochism disorder, 2018
- Sexual maturity rating (SMR), 3368–3369, 3368*t*
- Sexual minorities. *See also* Gay persons; Lesbian(s); Transsexual
    - depression among, 646
    - suicidal ideation, 646
- Sexual obsessions, 1699
- Sexual orientation
    - acceptance of, 2009
    - definition of, 1982, 1983
    - description of, 1959
    - development and, 2002–2004
- Sexual orientation and gender identity (SOGI), 2037–2039
- Sexual orientation conversion efforts (SOCE), 2010, 2032

- Sexual orientation distress, 4612
- Sexual orientation disturbance (SOD), 1988–1989
- Sexual pain disorders, 4612
- Sexual preferences, 1983
- Sexual reassignment. *See* Transsexualism
- Sexual relations, ethical practice of psychiatrists, 4445–4446
- Sexual response cycles
  - female, 1959*t*, 1960*f*
  - male, 1960*f*, 1960*t*
  - physiological, 1959
  - refractory periods, 1959
- Sexual sadism, 2018–2019. *See also* Sadism
- Sexual selection, evolution of, 772–773
- Sexual transgressions, 2669–2670
- Sexuality. *See also* Bisexuality; Heterosexuality; Homosexuality; Hypersexuality; Psychosexuality
  - aging and, **4255–4268,** 4256*t*, 4258*t*
  - cancer patients
    - psychological issues with, 2275–2276
  - conceptions of
    - training and research, 2011–2012
  - determinations of, 1953–1954
  - early, 2506
  - elderly persons, 2003–2004
  - endocrinology of, 1955–1956
  - epilepsy patients, 517
  - homosexuality, 4256–4259
  - infantile, 863–864
    - in psychoanalysis, 855–856
  - LGBTQ, 4256–4259
  - in long-term care, 4264–4267
  - medications, 4266*t*
  - neurobiology of sexual function, 4256
  - normal, **1953–1982,** 1954–1957
  - at patient interview, 967
  - psychiatrist role
    - assessing sexual dysfunction, 4259–4260
    - obtaining sexual history, 4259

- psychological states and, 2508
- treatment of dysfunction, 4260–4264, 4260*t*, 4262*t*, 4264*t*
- Sexualization, 4612
- Sexually dimorphic nucleus of the preoptic area (SDN-POA), 1993–1994
- Sexually transmitted diseases (STDs), 2513
  - child
    - laboratory screening for, 3408
  - laboratory evaluation, 1011–1012
- Sexually violent predators/sexually dangerous persons, 2503
- SGAs. *See* Second-generation antipsychotics
- Shadoo gene, 607
- Shagreen patches, assessment of, 456
- Shaman, description of, 4612
- Shamanic healing
  - as energy therapy, 2546, 2564
  - rituals, 2580
  - spirit possession, 2580–2581, 2580*f*
- Shame
  - description of, 4612
  - dissociative identity disorder and, 1925
  - public, suicide and, 2621–2622
  - trauma and, 1907
- *SHANK* gene family, 314
- Shankara and eastern philosophy, 928
- Shaping
  - description of, 4612
  - in operant behavior therapy, 2872
- Shared contextual transference, 2754, 2755
- Shared decision making, 4348
- Shared delusions, 1123
- Shared paranoid disorder, 4612
- Shared psychotic disorder, 4612
- Shared savings
  - in Accountable Care Organizations, 4318
  - in Medicare, 4308
- Shear, Kathleen, 2812
- Sheehan, Peter, 2709, 2732

- Sheehy, Gail, 938
- Shell shock. *See also* Posttraumatic stress disorder
  - description of, 4612
  - hypnosis and, 2706
  - war and, 2566
  - World War I studies, 1868
- Shenger–Krestovnika, Irena, 930
- *Shenjing shuairuo,* 2446
  - clinical features, 2457
  - phenomenology, 2456, 2456*t*
  - precipitants/etiology of, 2456–2457
  - presentation of, 2455–2458
  - psychiatric diagnoses and, 2457
  - specific cultural factors, 2457–2458
  - treatment, 2458
- Shenk, Joshua Wolf, 2834–2835
- Shepp–Logan phantom, 290*f*
- Sherman, Jeffrey, 2718
- Shh. *See* Sonic hedgehog (Shh)
- Shh protein, responses to, 41
- Shielding phenomenon, 262
- Shift work
  - circadian rhythm disorder, 4612
  - sleep-wake disorders and, 2098
- Shih, Miaozun, 2723
- Shims, magnet, 264
- "Shingles," 632
- *Shinkei suijaku,* 2455
- Shock probe burying test, models of, 321
- Shock treatment, description of, 4612
- Shole genome bisulfite sequencing, 439
- Shore, Ronald, 2709
- Short association fibers, 12, 15*f*, 451
- Short-duration depressive episode, 4613
- Short interspersed elements (SINEs), 154
- Short sleepers, 2088
- Short tandem repeats (STRs), 155

- Short-term anxiety-provoking psychotherapy (STAPP), 2797–2800
- Short-term dynamic psychotherapies (STDPs), 2796, 2800–2801. *See also* Intensive short-term dynamic psychotherapies
  - accelerated experiential, 2804
  - case study, 2801
  - emerging, 2804–2810
  - of narcissistic and other self-disorders, 2805–2810
  - psychoanalytic psychotherapy for panic disorders, 2804–2805
  - in specific psychiatric disorders, 2810–2812
- Short-term goals (STGs) in treatment, 4338
- Short-term memory, 4613
- Short-term psychodynamic psychotherapy, 2647
- Short Time Fourier Transform (STFT), 408
- Shorter, Edward, 1832
- Shorvonn, Harold, 1870
- *Shubo-kyofu* (body appearance obsession), 1791, 1797
- Shulman, Robert, 278
- Shy–Drager syndrome. *See* Multiple system atrophy (MSA)
- Shyness disorder. *See* Avoidant disorder of childhood or adolescence
- Sib pair analysis (SPA), 308
- Sibelium. *See* Flunarizine
- Siblings
  - description of, 4613
  - half-sibling and
    - ADHD studies, 3588
  - issues of, 2310–2311
  - rivalry, 4613
- Sicca syndrome, 456
- Sick call in jails and prisons, 4460
- Sick roles, description of, 4613
- Sickle cell anemia, 2390*t*
- Sickness behaviors, stress and, 2321
- Sickness histories, concept of, 2447
- Side effects
  - description of, 4613
- Sierra, Mauricio
  - corticolimbic disconnection hypothesis, 1891
  - dissociative identity disorder, 1871

- Sieverts (Sv), 293
- Sifneos, Peter
  - intensive short-term dynamic psychotherapy, 2797
  - short-term anxiety-provoking psychotherapy and, 2797–2800
  - short-term dynamic psychotherapies, 2647
- Sighing dyspnea, 2229
- Sign tracking, 718, 718*f*
- Signal anxiety, definition of, 2676, 4613
- Signal transduction
  - neuronal, 119–120, 120*f*
    - pathways, 120
  - outline of, 119*f*
- Significant differences. *See* Statistical significance
- Significant others, helping, when suicide occurs, 2619
- Signs, description of, 4613
- Signs of Suicide (SOS), 2617, 3685, 3699
- Sildenafil (Viagra)
  - antidepressant-induced erectile dysfunction, 1779
  - ED, 3338, 3339, 3339*t*
  - for sexual dysfunction, 1258, 1978, 4261
  - side effects, 1978
  - treatment of ED, 107
  - use in women, 1979
- Silence in patient interview, 958
- "Silence of the Mind" meditation, 2168
- Silver stains, 5
- *Silybum marianum,* 2554
- Simeon, Daphne
  - on depersonalization/derealization disorder, 1892, 1944
  - on dissociation and dissociative disorders, 1871
- Simmel, Ernst, 2706
- *Simmons, Roper v.*
  - criminal competence, 4438
  - juvenile justice, 4439–4440
- Simons, Ronald, 2445
- Simple phobias, 4613
- Simple tics, 3635, 3635*t*, 3649, 3649*t*

- Sims, James, 1179
- Simulated situations, for observation, 2687
- Simultanagnosia, 4613
- Simvastatin (Zocor)
  - for Alzheimer disease, 4147
  - drug interactions, 3131
- Sinequan. *See* Doxepin
- Singer, Margaret T., 1947
- Single-case experimental designs, 2703
- Single major locus transmission, 1621–1622
- Single nucleotide polymorphisms (SNPs)
  - definition, 155, 3449
  - in genetic mapping, 3449
  - schizophrenia, 1441–1442
- Single nucleotide variants (SNVs), 1627
- Single-photon emission computed tomography (SPECT). *See also* Forensic psychiatry, neuroimaging
  - dissociative identity disorder, 1894
  - factitious disorder imposed on self, 1851
  - gamma camera, 288*f*
  - neuroreceptor imaging studies, 1759
  - panic disorder, 1761
  - pediatric considerations, 3459, 3460
  - photon detection by collimation, 288
  - radionucleotides, 286*t*, 287
  - in schizophrenia, 273
  - spatial resolution, 292

- Single-photon emission computed tomography-CT (SPECT-CT), 292
  - after TBI, 525–526
  - compartment models, 293–297, 294*f*
  - radiotracers used in, 298*t*
  - reversible models, 294–295
  - safety issues, 293
- Single–best-answer multiple-choice questions, 4419–4420
- Sinistrality, 455
- SIP1 transcription factor, 50
- Sirolimus, 2371
- Sitophobia, 4613
- Situational analysis, 2700
- Six Aims in Quality Chasm, 4314, 4315
- 16 Personality Factor Questionnaire, 935
  - characteristics of, 998*t*
- Sjögren syndrome
  - diagnosis of, 473
  - dry mouth in, 73
  - eyes in, 456
- *SKA2* (suicide biomarker), 2619
- Skewed family, child development and, 2874–2875
- Skill training for caregivers, 4210
- Skills training, behavior therapy, 2694–2695
- Skin. *See also* Dermatological effects; Galvanic skin resistance
  - barrier function of, 2339–2340
  - as boundary of self, 2335–2336
  - carbamazepine effects on, 3036
  - circadian rhythms and, 2339–2340
  - DA receptor antagonists and, 3076
  - diseases, psychoneuroendocrinology of, 2336–2339
  - grafts for burn injuries, 2381–2382
  - increased conductance, 2350
  - lamotrigine effects on, 3085
  - lesions
    - appearance of, 2333
    - disfigurement and, 2333
  - MAOI effects on, 3117

- opioid receptor antagonist effects on, 3144
- permeability of, 2340
- physical examination of, 456
- psychocutaneous disorders, **2332–2357**
- sensory perception and, 2337
- sexual response cycle, 1959*t*, 1960*t*
- sleep and, 2339–2340

- Skin-picking disorder, 4613
- Skin-to-skin contact, benefits of, 2052
- Skin-to-skin sensations, 2052
- Skinner, B. F., 3359, 3750
  - *About Behaviorism,* 930
  - learning theory, 717, 724
  - radical behaviorism, 929–930
  - *Science and Human Behavior,* 4540
  - work of, 4613
- Skooptsi (eunuchs), 2583
- Skype. *See also* Computers; Internet; Smartphone apps
  - psychoanalytic psychotherapy and, 2672
- Slavson, Samuel, 2737
- *SLC6A3 (DAT1),* 1625
- SLC6A4. *See* Serotonin transporter
- SLC6A8, 664
- *SLCA1* gene mutation, 648
- *SLC2A1* gene mutation, 648
- Sleep, 2095*f*. *See also* Insomnia; NREM sleep; REM sleep; Slow-wave sleep (SWS)
  - animal models of, 324
  - ascending arousal system, 342*f*
  - barbiturate effects on, 2986
  - basic science of, **339–354**
  - and brain function, 744
  - brain restitution and, 352–353
  - circadian disorders, 214
  - circadian rhythms and, 2339–2340
  - constant routine protocols, 348
  - definition of, 339
  - deprivation of

- depression and, 353
- EEG studies, 243
- effects of, 348–350
- glucose metabolism and, 349
- studies of, 210

- description of, 4613
- disordered
  - immune responses and, 2320
- EEG biofeedback and, 2559
- forced desynchrony protocol, 348
- functions of, 352–353
- induction of, 346
- learning and, 353
- “local sleep,” 350
- and memory, 739
- memory and, 353
- monitoring of, 342–343
- nefazodone effects on, 3128
- neurobiology of, 344–348
- neuroendocrine changes in, 344
- neurophysiology alterations, depression and, 1716–1717
- normal human, 339–342
  - sleep stage histogram, 2091*f*
- organization of, 341–342
- patterns of
  - variant normal EEGs, 237, 239*f*
- physiology of, 343–344
- polysomnography, 243, 243*f*
- psychiatry and, 353–354
- PTSD and, 1821
- regulation of, 210–211, 348–350
- SSRI effects on, 3158
- stages of, 339, 341
  - dreaming and, 351
  - EEG, 340*f*
  - electrophysiological criteria, 340*t*
- standardized visual scoring, 343

- temporal isolation and, 348
- and thalamic oscillatory properties, 745–746, 749
- thermoregulation in, 344
- vilazodone effects on, 3346

- Sleep apnea
  - polysomnography, 2094*f*
  - treatment
    - buspirone, 3012
    - modafinil, 3232
- Sleep-disordered breathing (SDB), 3935
- Sleep disorders, **2083–2109**
  - by aging adult, 3975
    - circadian rhythm, 4055, 4055*t*
    - comorbid medical and psychiatric illnesses, 4053
    - insomnia, 4053–4054, 4054*t*
    - medications, 4053
    - in neurocognitive disorders, 4057
    - in nursing homes, 4057–4058
    - periodic limb movements, 4056–4057
    - rapid eye movement sleep behavior, 4057
    - restless legs syndrome, 4056–4057
  - in alcohol abuse, 4112
  - alcohol-induced, 1267
  - caffeine-induced, 1298–1299, 1298*t*
  - case studies, 2106–2109
  - child/adolescent, 3675, 3689, **3934–3940**
    - ADHD-associated, 3938
    - anxiety disorders-associated, 3939
    - autism spectrum disorder-associated, 3938
    - breathing, 3935
    - development and, 3360
    - insomnia, 3935–3936
    - mood disorders-associated, 3938–3939
    - narcolepsy, 3936
    - nocturnal enuresis, 3937–3938
    - parasomnias, 3936–3937
  - classification of, 2083, 2084*t*, 3935, 3935*t*

- delirium and, 1188–1189
- dementia and, 1196, 1202
- depressive disorders and, 3675, 3689
- description of, 4613
- enuresis and, 3653, 3667
- evaluation methods, 3934
- intellectual disability and, 3512, 3516
- neurocognitive effects, 3934
- overview, 3934–3935
- psychopharmacology, 3794, 3796
- PTSD and, 3705
- repeated arousals and, 2340
- stimulant-induced, 1286
- substance/medication-induced, **1256–1258**
  - comparative nosology, 1256
  - course, 1257
  - defined, 1256
  - diagnosis and clinical features, 1235*t*, 1257
  - differential diagnoses, 1257
  - epidemiology, 1256
  - etiology, 1257
  - laboratory examination, 1235*t*, 1257
  - treatment, 1257–1258
- types of, 1140–1141

- Sleep disturbances after TBI, 525
- Sleep efficiency, definition of, 343
- Sleep enuresis, 2102
- Sleep homeostats, 210
- Sleep hygiene
  - description of, 2086*t*
  - inadequate, 2085–2086
  - universal, 2088–2089
- Sleep impairment, 1267
- Sleep latency, 343
- Sleep medicine
  - certificate examinations, 4427
  - tools in, 2105–2106

- training, 4405
- Sleep paralysis, 2100
  - narcolepsy and, 2092*f*
  - recurrent isolated, 2102
- Sleep rebound, 348–349
- Sleep-related breathing disorders, 2093
- Sleep-related bruxism, 2103–2104
- Sleep-related eating disorders, 4613
- Sleep-related hallucinations, 4613
- Sleep-related hypoventilation, 2097
  - description of, 4613
  - diagnostic criteria, 2097
- Sleep-related leg cramps, 2103
- Sleep restriction therapy, 2089
- Sleep starts, 2105
- Sleep state misperception, 2085
- Sleep studies, narcolepsy and, 2092
- Sleep talking, 2105
- Sleep terrors, 4613
  - diagnostic criteria, 2099
  - polysomnography, 2100*f*
- Sleep-wake disorders, **4053–4058**
  - circadian rhythm, 2097–2099, 4055, 4055*t*
    - diagnostic criteria, 2097
    - subtypes, 2097–2098
  - classification of, 1159
  - comorbid medical and psychiatric illness, 4053
  - delirium and, 1188–1189
  - future directions, 4058
  - insomnia, 4053–4054
    - behavioral treatment, 4054
    - pharmacologic treatment, 4054, 4054*t*
  - in long-term care, 4230–4231
  - medications, 4053
  - in neurocognitive disorders, 4057
  - in nursing homes, 4057–4058
  - periodic limb movements, 4056–4057

- REM sleep behavior, 4057
- restless legs syndrome, 4056–4057
- sleep apnea/hypopnea, 4055–4056

- Sleep–wake schedule disorder, 4613
- Sleep–wake types
  - advanced sleep phase, 2097
  - delayed sleep phase, 2097
  - irregular, 2097
  - non-24-hur, 2097
- Sleep–waking cycles
  - circadian disorders of, 214
  - EEG changes, 342–343
  - locus coeruleus and, 62
  - nitric oxide and, 107
- Sleepwalking, 2099, 4613
- Slit protein, cell migration and, 52
- *SLITRK1,* TTM and, 1794
- Slow transit constipation, 3660, 3674
- Slow-wave sleep (SWS)
  - age and, 342
  - description of, 4613
- Slowed (retarded) thought, 1120
- Sluggish cognitive tempo (SCT), 2437
- SLUMS. *See* St. Louis University Mental Status Examination
- Small cell lung cancer (SCLC), 2268–2269
- Small-interfering RNAs, 51
- Small sequence repeats (SSRs), 155
- Small sharp spikes (SSS), 239–241, 241*f*
- Smart home technology, 4278
- Smartphone apps
  - anxiety and depression, 2691
  - behavior therapy and, 2701
- Smith, Mary Lee, 2796
- Smith, Richard, 1837
- Smith–Lemli–Opitz syndrome
  - description of, 48–49
  - with intellectual disability (ID), 3508

- Smith–Magenis syndrome, 3508
- Smoking. *See also* Nicotine; Tobacco
  - brain development and, 58
  - burn injuries and
    - risk of, 2374
  - CAD and, 2187
  - cessation
    - anticonvulsants and, 2970
    - bupropion and, 1680, 3005–3006
    - buspirone, 3010–3011
    - hypnosis and, 2730–2731
    - medication levels and, 2266
    - naltrexone, 3142
  - cessation programs
    - cancer patients and, 2258
  - drug interactions, 3106
  - neurocentrism, 783
  - schizophrenia association with, 4529
- Smoking restrictions in psychiatric hospitals, 4340
- Smoky Mountain Survey, 3727
- *The Snake Pit* (Ward), 4329
- Snap judgments, 1124
- Snoezelen, 2556
- SNRIs. *See* Selective serotonin-norepinephrine reuptake inhibitors; Serotonin-norepinephrine reuptake inhibitors (SNRIs)
- Snyder, C. Rick, 2853
- Sociability, indiscriminate
  - disinhibited social engagement disorder and, 3663, 3665, 3666, 3677, 3679, 3680
- Social adaptation, 4613
- Social affiliation, 3481
- Social anxiety disorder (SAD), **3714–3719**. *See also* Social phobias
  - acne and, 2345
  - BD and, treatment of, 1698
  - benzodiazepines and, 1747
  - cannabidiol for, 1743*t*
  - CBT and, 1775–1776
  - child/adolescent

- comparative nosology, 3715
- course and prognosis, 3716–3717
- defined, 3714
- diagnosis and clinical features, 3715–3716, 3716*t*
- differential diagnoses, 3716
- epidemiology, 3715
- etiology, 3715
- future directions, 3718
- pharmacotherapy, 3717–3718
- psychotherapy, 3717
- treatment, 3717–3718, 3758

- in children and adolescents, 1766–1767
  - management, 1784
- cognitive-behavioral therapy for, 3758
- cognitive therapy, 2766
- d-cycloserine for, 1747
- definition, diagnosis, and clinical features, 1725
- description, 1769
- differential diagnoses, 1725–1726, 3553, 3722
- DSM-5 on, 1721, 1723
- epidemiology, 1725
  - genetic, 1738
- glutamate-modulators and, 1748
- history and comparative nosology, 1725
- HPA axis abnormalities and, 1744
- management, 1782*t*, 1783–1784
- neuroimaging, 1763, 1765
- neuropeptide Y and, 1749
- oxytocin and, 1749
- pharmacotherapy and, 1782
- phenomenology of, 1752
- pregnenolone and, 1748
- psychosocial interventions for, 3746
- rTMS for, 1782
- self and, 332
- treatment
  - MAOIs, 3122

    - nefazodone, 3127
    - propranolol for, 67
    - SSRIs, 3154
    - venlafaxine, 3149
- Social anxiety-like behavior, animal, 321
- Social attachment, animal, 324
- Social attribution research, 2695
- Social behavior and interactions
  - ADHD and, 3591
  - autism spectrum disorder and, 3580, 3581, 3581*f*
  - conduct disorder and, 3615–3616
  - development, adolescent, 3369, 3371
  - development, childhood, 3352*t*–3353*t*, 3360–3361
  - emergency psychiatric evaluation, 3442
  - group psychotherapy and, 3761
  - group therapy and, 3769–3770
  - posttraumatic stress disorder, child/adolescent, 3704
  - social (pragmatic) communication disorder and, 3552
  - social media and Internet effect on, 3373
- Social Behavior Schedule (SBS), 4372
- Social behaviors, abnormal, 465
- Social breakdown syndrome
  - description of, 4613
  - schizophrenia, 1435
- Social capital, social networks and, 4383
- Social class
  - anxiety disorders, 1734
  - cancer risk and, 2285
- Social cognition
  - during adolescence, 3369, 3371
  - rehabilitation of, 4373
- Social Cognition Psychometric Evaluation (SCOPE), 4370
- Social cohesion in health care outcomes, 4383
- Social (pragmatic) communication disorder, **3551–3555**
  - course, 3554
  - criteria for, 3349
  - defined, 3586, 4613

- development, 3554
- diagnostic features, 3552–3553
- differential diagnoses, 3553
- epidemiology, 3554
- etiology, 3553–3554
- historical background, 3551–3552
- overview, 3551
- treatment, 3554–3555

- Social contracts, 4323
- Social control, description of, 4613
- Social control rationale, 4332
- Social correlates of happiness, 2848
- Social detachment, 2173
- Social development, 923, 923*t*
- Social environment awareness, 2694–2695
- Social exchanges, cheating in, 453
- Social functioning
  - animal models of, 324
  - everyday functioning, 4026
  - family role, 4196
  - of older adults, 3998–3999
  - in schizophrenia, 4097
- Social Functioning Examination, 532*f*
- Social hierarchy, animal, 324
- Social history at patient interview, 951, 951*t*, 966
- Social inference theory, 3553
- Social instinct. *See* Herd instinct
- Social interactions
  - and cognition, 693
  - and neural network growth, 688
- Social Introversion (Si) Scale, MMPI, 999*t*
- Social issues, genetic susceptibility testing and, 2537
- Social learning theory, 760, 2683, 4613
- Social media. *See also* Email; Internet
  - adolescent use of, 3372–3373, 3811–3812
  - as child psychiatric evaluation consideration, 3378
  - illness tracking with, 3852

  - technology of, 4279
- Social network therapy, 4613
- Social networks, social capital and, 4383
- Social phobias. *See also* Social anxiety disorder (SAD)
  - agoraphobia and, 1783
  - classification of, 1173
  - cognition in, 702–703
  - cognitive therapy, 2862
  - description of, 4613
  - differential diagnoses, 1275
  - DSM-5 on, 1721
  - epidemiology, 1732, 1732*t*
  - epigenetics, 1740
  - psychotherapy, neuroanatomy and, 2814*t*
  - schizophrenia, 1421
  - selective mutism and, 3721
- Social psychiatry, 4613
- Social roles, learned, 1873
- Social safety net, 4323
- Social services, patient support and, 2254
- Social skills in rehabilitation, 4362
- Social skills training (SST)
  - behavior therapy, 2694–2695
  - cognitive-behavioral group therapy and, 1776
- Social stressors, mood disorders and, 1617
- Social support
  - after release, 4399
  - immune protection and, 203
  - mood disorders and, 1617
  - transplant evaluation and, 2361
- Social syndrome, 1608
- Social therapy, description of, 4613
- Social welfare programs, 4305
- Social withdrawal, schizophrenia and, 1417
- Social work, board certification of, 4429
- Socialization, description of, 4613
- Sociatry, psychodrama and, 2855

- Society for Women's Health Research, 4249
- *Society of New York Hospital, Schloendorff v.,* 4433–4434
- Sociobiological psychiatry, 4543
- Sociobiology, description of, 4613
- Sociocognitive model (SCM)
  - cultural factors, 1899
  - dissociation, 1897–1899
  - hypnosis, 1897
- Sociocultural framework
  - community systems, 4383–4384
  - future directions, 4386–4387
  - health care systems
    - organization, 4381–4382
    - policy, 4379–4381
    - provider-level factors, 4382–4383
  - mechanisms producing disparities, 4384–4386
  - overview of, 4377–4378
  - for service disparities, 4378–4379, *4378f*
- Sociocultural theories of alcohol-use disorders, 1268–1269
- Sociodrama
  - methods, 2857
  - psychodrama and, 2855
- Socioeconomic factors
  - deprivation, 4296
  - interpersonal aggression and, 2476
  - mood disorders and, 1616
- Socioeconomic status (SES)
  - autism spectrum disorder and, 3574
  - disruptive behavior disorders and, 3609
  - enuresis, 3651, 3665
  - immunoregulation and, 189*f*
  - intellectual disability and, 3498
  - language disorders and, 3546
  - violence and, 2476
- Socioemotional intelligence, 761–762
- Sociology, description of, 4613
- Sociometry

- description of, 4613
- psychodrama and, 2855
- Sociopathic behavior, 4613
- Sociopathy, 4613
  - and genetic selection, 778
  - psychopathy and, 2490
  - psychopathy *versus,* 2490–2491
- Sociotherapy, description of, 4613
- Socratic dialogue/questioning
  - cognitive therapy, 1773, 2770–2771, 2770*f*
    - motivational, 2772
- Sodium amobarbital, 1940
- Sodium amytal, 1840
- Sodium ion (Na^+) channels
  - α-subunit, 136
  - activation of, 135
  - β-subunit, 136
  - blockage of, 136
  - function of, 135–136
  - inactivation of, 135
  - structure of, 135*f*
- Sodium ions (Na^+), 131
- Sodium oxybate (Xyrem), 2093
- Sodium pumps, transport by, 132
- Sodium valproate (Dapakene), 2501.
  See also Valproic acid; Valproic acid (valproate, Depakene, Depakote)
- Sodomy, definition of, 4614
- Soft bipolarity, 1650
- Soft signs, description of, 459–460
- Soldier's heart. *See* Da Costa syndrome
- Solms, Mark, 2645
- Solomon, Harry, 4332
- Solution analysis, in DBT, 2790
- Solution-focused therapies, 2752*t*, 2755
- Solvent inhalation, laboratory evaluation, 1018
- Solvents, hexacarbons in, 631
- Soma. *See* Carisoprodol
- Soma, definition of, 4614

- Somatic delusions, 4573, 4614
- Somatic dysfunction, 3012
- Somatic hallucination, 4614
- Somatic management techniques, 3753–3754
- Somatic symptom and related disorders, **1827–1845**
  - assessment instruments, 1844
  - comorbidity, 1834
  - comparative nosology, 1830–1832, 1831*t*
  - conversion disorder, 1839–1842
  - course and prognosis, 1834
  - definition, 1827–1828
  - diagnosis and clinical features, 1833–1834
  - differential diagnoses, 1834
  - dissociative disorders and, 1866
  - DSM-5 on, 1831*t*
  - epidemiology, 1832
  - etiology, 1832–1833
    - attitudinal/behavioral factors, 1833
    - biological factors, 1832–1833
    - cultural factors, 1832
    - environmental factors, 1832
    - psychological factors, 1833
  - factitious disorder, 1842, 1849
  - history, 1828–1830, 1828*f*
  - hypnosis and, 2719–2720
    - case studies, 2720
  - ICD-10 on, 1831*t*
  - idiopathic, 1827
  - illness anxiety disorder, 1837–1839
  - introduction, 1827
  - NIMH research domain criteria, 1844
  - other specified, classification, 1827
  - psychological factors affecting other medical conditions, 1841–1842
  - reliability of, 1829*f*
  - somatic symptom disorder, 1835–1837
  - St. Louis criteria, 1830
  - suicide and, 2614

- treatment, 1834–1835
  - future of business, 1844
  - general approach, 1835
  - general guidelines, 1843–1844
  - history taking, 1834
  - physical *versus* psychological focus, 1835
  - reassurance, 1834–1835
- unspecified and below-threshold
  - definition, 1842
  - epidemiology, 1842–1843
  - history, 1842
  - introduction, 1842
  - pivotal study, 1843
  - treatment, 1843

- Somatic symptom disorder (somatization disorder)
  - Briquet syndrome and, 1830
  - cardiac symptoms of, 2195
  - classification, 1830, 1835
  - comparative nosology, 1835–1836
  - course/prognosis in, 2392–2393
  - description of, 4614
  - dissociative identity disorder and, 1913
  - epidemiology, 1836
  - etiology, 1836
  - factitious disorder and, 1848–1849, 1859
  - factitious disorder imposed on another and, 1850
  - ICD-10 on, 1831
  - introduction, 1835
  - malingering *versus,* 2411*t*
  - with pain, definition, 1840–1841
  - pathological dissociation and, 1885
  - STAPP and, 2798
  - STDP and, 2801
  - trauma, dissociative identity disorder as paradigm of, 1905
  - treatment, 1836–1837
    - general considerations, 1836
    - pharmacotherapy, 1837

    - pivotal study, 1837
    - psychotherapy, 1837
    - reassurance, 1837
    - Smith's consultation letter, 1837
    - supportive, 1837
  - unspecified, 1827
- Somatic symptoms
  - in classical psychopathology, 1829–1830
  - classification of, 1158–1159
  - in psychiatric illnesses, 1140–1142
  - of sexually abused children, 3832
  - suicide and, 2612
  - traumatic stress response and, 2566
  - World War II amnesia and, 1869
- Somatic syndromes, functional, 1831–1832, 1831*t*
- Somatic therapies
  - anxiety disorders, 1778–1784
  - in biomedicine, 2546
  - description of, 4614
  - dissociative amnesia, 1940–1941
  - exercise, 2556
  - OCD, 1790
- Somatization
  - of aging adult, 3974–3975
  - theory of, 1827, 1830
- Somatization disorder, 4614
- Somatoemotional release, 2546, 2561, 2563–2564
- Somatoform disorders, 1830–1831
  - cancer patients with, 2266
  - description of, 4614
  - psychotherapy, neuroanatomy and, 2814*t*
- Somatoform Dissociation Questionnaire (SDQ)
  - dissociative amnesia, 1934
  - dissociative disorders, 1883, 1884
- Somatoform Dissociation Questionnaire (SDQ-20), 1876*t*
- Somatoform Dissociation Questionnaire-5 (SDQ-5), 1876*t*
- Somatoform respiratory disorders, 2228–2229

- Somatomorph autonomic dysfunction, 1831
- Somatomorph disorder, 1831
- Somatopagnosia, 4614
- Somatosensory cortex
  - surface landmarks, 7*f*
  - thalamic input, 326–327
- Somatosensory pathways, 25
- Somatostatin (SRIF)
  - functions of, **174–175**
  - GH release inhibition and, 1716
- Somatotype, definition of, 4614
- Somnambulism
  - description of, 4614
  - hypnosis, 2711
- Somnolence, description of, 4575, 4614
- Sonata. *See* Zaleplon
- Sonic hedgehog (Shh), 41
- Soothing experiences, lack of, 1906
- Sordinol. *See* Clopenthixol
- Sortilin, interaction with BNDF, 102
- SOS (Son of Sevenless) protein, 125
- *Soul and Life* (Vives), 4533
- Sound, therapy using, 2558
- Sour date seed, 1782
- South London Somatization Study, 1842
- South Oaks Gambling Screen (SOGS), 1805
- Souvenaid for Alzheimer disease, 4148
- Soviet Communist trials, 1946
- Soviet Union, immigrants, and mental health, 803–804, 804*t*, 805*t*
- SOX family proteins, 442
- Spa physicians, 4402
- Sparine. *See* Promazine
- Spasmodic dysphonia
  - differential diagnoses, 3568–3569
  - as unspecified communication disorder, 3567
- Spasticity, treatment of, 2998
- Spatial disorders

- hemispheric dominance
  - dressing apraxia, 980
  - neglect syndrome, 980
  - visuospatial impairment, 979–980, 980*f*
- spatial neglect, 333

- SPD. *See* Excoriation (skin-picking) disorder
- Speaking in tongues, 4614. *See also* Glossolalia
- Spearman–Brown formula, 996
- Special education budgets, 4321
- Special education eligibility determination, **3884–3888**
- Special Forces soldiers
  - cortisol and stress in, 1892
  - dissociative disorders and, 1897
  - NPY levels and dissociation in, 1891
- Specific developmental disorders. *See* specific learning disorder
- Specific learning disorders, 3520–3536
  - ADHD and, 3533
  - assessment and intervention, 3533
  - associated problems, 3529
  - clinical features, 3527–3528, 3527*t*
  - cognitive factors, 3525–3526
  - comparative nosology, 3523
  - course, 3534–3535
  - criteria for, 3349
  - defined, 3521–2522, 4614
  - diagnostic criteria, 3529–3533, 3530*t*, 3532*t*
  - differential diagnoses, 3533–3534
  - environmental factors, 3526–3527
  - epidemiology, 3523–3524
  - etiology, 3524
  - genetics, 3524–3527
    - behavioral, 3524
    - cognitive factors, 3525–3526
    - environmental factors, 3527
    - molecular, 3524–3525
  - historical background, 3522–3523
  - intellectual disability (ID) and, 3533

- language disorders and, 3548
- neurobiological factors, 3525
- overview, 3520, 3521*t*
- prognosis, 3534–3535
- significance of, 3520–3521
- standardized assessments, 3531, 3532*t*, 3533
- treatment, 3535

- Specific Levels of Functioning Scale (SLOF), 4372
- Specific reading disability. *See* Developmental reading disorder
- Specificity, definition of, 4614
- SPECT. *See* Single-photon emission computed tomography (SPECT)
- Spectator play in adulthood, 3944–3945
- Spectatoring, 1976
- Spectral editing, 278
- Spectrum disorder hypothesis, 2164
- Spectrum model, 3484–3485
- Speech
  - assessment of, 457
  - disturbances, 4614
  - impairments
    - specific learning disorder and, 3529
  - at patient interview, 953, 966
  - production
    - Broca's area lesions and, 7
- Speech sound disorder and, 3557
- Speech sound disorders, **3555–3559**
  - comparative nosology, 3556
  - course, 3558
  - description of, 4614
  - diagnosis, 3556–3557
  - differential diagnoses, 3548, 3557–3558
  - epidemiology, 3556
  - etiology, 3556
  - laboratory examination, 3557
  - language disorders and, 3548
  - overview, 3555–3556
  - pathology, 3557

- prognosis, 3558
- treatment, 3558
- Speech syndromes, characteristics of, 457*f*
- Speech therapy, selective mutism and, 3723
- Spence, Donald, 2657
- Spermatogenesis
  - induction of, 2510*t*
- Sperry, Roger, 334, 4544
- Sphingolipids, breakdown of, 656
- Spider phobia
  - cortisol for, 1743*t*
  - d-cycloserine for, 1747–1748
  - neuroimaging, 1762
- Spiegel, David
  - dissociation, 1870–1871
  - ego state therapy and, 2717
  - eye roll technique, 2715
  - hypnosis, 2709, 2714, 2717
    - PTSD and, 2719
  - smoking cessation, 2730–2731
  - split-screen technique, 2718
- Spiegel, Herbert
  - ego state therapy and, 2717
  - hypnosis, 2714, 2717
  - smoking cessation, 2730–2731
  - split-screen technique, 2718
- Spiegel, John, 2706
- Spiegel, Sharon B., 2721–2722
- Spielrein, Sabina, 2641
- Spike–wave phantoms, 241, 242*f*
- Spin, characteristic of, 260
- Spin tagging, 271
- Spina bifida, neurulation and, 40
- Spinal agnosia, 4614
- Spinal cord
  - embryology, 41
  - sexuality and, 1957

- Spinal cord injuries
  - neuropathic pain with, 2977
  - pain related to, 2391
- Spinal cord stimulation (SCS), 2398
- Spinal muscular atrophy (SMA), 628
- Spinoreticulothalamic tracts, 326
- Spiperone (Spiroperidol), 3064
- Spirit possession, 2580–2581, 2581*f*
- Spiritism, rise of, 1867
- Spiritual emergencies, 2825
- Spiritual issues, 4171
- Spiritual lives, 2591
- Spiritual sensitivity, 2141
- Spirituality, **2576–2594**
  - achieving, forms and methods of, 2577–2583
    - altruism, 2583
    - body modification and mortification, 2582, 2582*f*
    - charity, 2583
    - confession, 2582–2583
    - consciousness, 2577–2578
    - drugs, 2581–2582
    - meditation, 2578
    - mysticism, 2577–2579
    - New Age spirituality, 258
    - penance, 2582–2583
    - prayer, 2579–2580
    - reconciliation, 2582–2583
    - spirit possession, 2580–2581, 2580*f*, 2581*f*
    - spiritual site visits, 2583
  - anxiety and, 2587–2589
  - assessment tools, 2586–2587
  - biology and, 2593–2594
  - in clinical practice, 2583–2594
    - problems with, 2587–2594
  - definition, 2576
  - depression and, 2587–2589
  - dying patients

    - decision-making issues, 2302–2303
    - spiritual concerns of, 2302–2303
  - and mental health, 765–766
  - overview, 2576–2577
  - positive psychiatry, 4283
  - psychiatrists' perceptions about, 2585–2586
  - psychosis and, 2587–2589
  - religion, religiosity and, 2583–2585
  - specific spiritual therapies, 2593–2594
  - treatment issues, 2591
- Spirochetal diseases, **589–594**
- Spiroperidol. *See* Spiperone
- Spitz, Rene, 4542
- Spitzer, Robert
  - on psychiatric classification, 1154
  - St. Louis criteria and, 1830
- Splanchnic nerve, 1757
- Splinter skills, 3572, 3582
- Split-brain studies, 334
- Split-screen technique, 2718–2719
- Splitting, description of, 4614
- Spongiform encephalopathies, 4022
- Sprenger, James, 4533
- *SPRN* gene, 607
- Squeeze technique, 1976
- Squid giant axons, 137
- SSR149415, 1743*t*
- SSRIs. *See* Selective serotonin reuptake inhibitors (SSRIs)
- St. Augustine, 4533
- St. John's wort *(Hypericum perforatum),* 2552*f*
  - adverse effects, 2544*t*
  - anxiety disorders, 1782
  - biological treatment with, 2552
  - in CAIM medicine, 3984
  - as complementary treatment, 2542
  - safety, 2543
  - side effects, 2287

- St. Louis criteria, 1154
  - conversion disorder, 1839
  - for psychiatric classifications, 1115
  - somatic symptom disorders, 1830
- St. Louis University Mental Status Examination (SLUMS), 988
- Stabilization
  - dissociative amnesia, 1938
  - dissociative identity disorder, 1922–1925
- Stabilization phase of rehabilitation, 4367–4369
- Staff
  - burn units, 2386
  - dynamics in end-of-life care, 2311
  - moral of, 2282
  - pediatric burn injuries and, 2386
  - pediatric psycho-oncology, 2282
  - psychiatric interventions, 2294–2295
  - psychological issues for, 2280
  - stress and, 2280, 2386
  - support for, 2386
  - vulnerabilities of, 2294
- Staff–family interfaces, 2280
  - end-of-life issues, 2306–2307
- Staff–Resident Interaction Chronograph (SRIC), 4372
- Stage fright, propranolol for, 1742
- Stages in developing intelligence, 706*t*
  - concrete operations, 707–708
  - formal operations, 708–709
  - preoperational subperiod and semiotic function, 707
  - sensorimotor period, 706–707
- Stages of Reproductive Aging Workshop (STRAW) staging system, 2521
- Stahl, Georg Ernst, 4534
- Stammering. *See* Stuttering
- Standard deviation (SD), 4614
- Standard Dialectical Behavior Therapy (DBT), 4207
- *Standard Edition of the Complete Psychological Works of Sigmund Freud* (Freud), 887
- Standard error of measurement, 997
- Standard error of the mean (SEM), 4614

- Standardized assessment instruments and measures, **3391–3403**
  - development of, 3394–3395
  - for infant and toddler evaluations, 3389–3390, 3390*t*
  - for intellectual disability, 3493–3496, 3494*t*, 3496*t*
  - interpretation of findings, 3405–3406
  - interviews types as, 3395–3398
  - limitations of, 3402–3403
  - for psychological assessments, 3404–3406, 3405*t*
    - interpretation of, 3405–3406
    - reliability and validity of, 3404–3405
    - reporting of, 3406
    - types of, 3404, 3405*t*
  - rating scales as, 3398–3402
  - specific learning disorder, 3531, 3532*t*, 3533
  - types of, 3404
- Standardized patients, board certification and, 4429
- Standardized scores, 4614
- Stanford Hypnotic Clinical Scale (SHCS), 2724, 2727
- Stanford Hypnotic Susceptibility Scale (SHSS), 2706–2707, 2707*f*, 2713–2714
  - Form C (SCSHS), 2714
- Stanford Sleepiness Scale (SSS), 343
- Stanford–Binet Intelligence Scale, 4614
- Stanford–Binet Intelligence Scales, 3423, 3424
- Stanozolol (Winstrol), 2493
- Staring spells in children, 650
- STARRS (Army Study to Assess Risk and Resilience in Service members), 2568
- Start codons (AUG), 151
- Startle reactions, 4614
- State-dependent learning or memory, 4614
- State hospitals
  - care integration in, 4346
  - community mental health centers as gatekeepers, 4326
  - demand increases, 4325–4326
  - disappearance of, 4345
  - funding cuts, 4325
- State markers of heavy drinking, 1273–1274, 1274*t*
- State model of hypnosis, 2708, 2711

- State regulation, feeding disorder, **3628–3629, 3642–3643**
  - clinical features, 3628, 3642
  - course, 3628, 3642
  - differential diagnoses, 3628, 3642
  - etiology, 3628, 3642
  - treatment, 3628–3629, 3642–3643
- State stewardship, *4295f*
- State Trait Anxiety Inventory, 1003*t*
- States. *See also* Trauma-related altered states of consciousness
  - consciousness
    - altered, 2709–2710
    - altered, ER and, 2724–2725
    - concept of, 1899
  - disassociation, 2491
  - hypnosis, 2709, 2711
  - shifting, 1912
- State–Trait Anxiety Inventory for Children (STAIC), 3401
- Statins (HMG-CoA reductase inhibitors), 4147
- Statistical inference, 4614
- Statistical significance, 4614
- Statistical techniques for neuroimaging, 4485
- Statistics and experimental design, **829–847**
  - basic ideas
    - conditional probability and independence, 830
    - mean, median, and variance, 830
    - multivariate random variables, 831
    - parametric forms, 830
    - probability density functions, 830
    - probability theory, 830
  - Bayesian approach
    - Bayes inference, 835–836
    - Bayes theorem, 835
  - cross-references, 847
  - estimation of parameters
    - confidence interval estimation, 831
    - confidence interval interpretation, 832
    - point estimators, 831

    - properties of estimators, 831
  - evaluation of treatments
    - cost-effectiveness, 846–847
    - effectiveness studies, 845–846
    - efficacy studies, 845
  - hypothesis testing, 832
    - false discovery rate, 835
    - multiple hypotheses, familywise error rate, 833–835
    - *p* values, 832–833
    - power of test, 832
    - type 1 and type 2 errors, 832
  - randomized clinical trials, 836–839
    - causality, 836, 838
    - matching and randomization, 836–837
    - missing or incomplete data, 838–839
    - propensity scores, 837–838
    - randomization tests, 837
  - specific techniques
    - contingency tables and log linear models, 842
    - general linear model, 842
    - growth mixture modeling, 843
    - latent class growth analysis, 843
    - logistic regression, 840–841
    - longitudinal models, 842–843
    - regression models, 839–840
    - supervised machine learning, 841
    - survival analysis, 843–844
  - study design
    - crossover design, 844–845
- Statistics for mentally ill incarcerated, 4388
- Status epilepticus
  - congnitive disorders after, 511–512
  - description of, 507, 4614
  - nonconvulsive generalized, 458
- Steal syndrome, 1980–1981
- Stealing,
  - compulsive, OCD and, 1789

- Stein, Calvert, 2716
- Steinach, Eugen, 2481
- Steinberg, L., 3365
- Steinberg, Marlene, 1871
- Stekel, Wilhelm
  - on paraphilias, 2014
  - psychoanalysis and, 2641
  - on somatization, 1830
  - somatization theory, 1827
- Stelazine. *See* Trifluoperazine
- Stellate neurons, 16*f*
- Stem cell transplants, 2274
- Step down services, 4344–4345
- Stereognosis, 459
- Stereotypy and habit disorder, 4614
- Stereotypy and stereotypic movement disorders, **3667–3673, 3681–3687**
  - autism spectrum disorder and, 3582
  - comparative nosology, 3670–3671, 3684–3685
  - defined, 3671, 3685
  - description of, 4614
  - diagnosis, 3671–3672, 3685–3686
  - differential diagnoses, 3671, 3685
  - epidemiology, 3668–3670, 3682–3684
  - etiology, 3668, 3682
  - genetics, 3668, 3669, 3682, 3683
  - injurious behavior and, 3672, 3686
  - OCD *versus,* 1788
  - overview, 3667–3668, 3681–3682
  - treatment, 3672–3673, 3686–3687
- Stereotypy of thought, 1120
- Stern, Daniel, 2646
- Steroid(s). *See also* Anabolic-androgenic steroid-related disorders; Dihydrotestosterone (DHT); Estradiol (Alora, E_2); Estrogens; Hormone(s); Progesterone; Testosterone
  - anabolic
    - laboratory evaluation, 1013
  - mania without prior bipolarity and, 1649
  - modes of action, 165

- psychotic disorder induced by, 501–502
- Sterol-27 hydroxylase deficiency, 662
- Stevenson, Janine, 2723
- Stevenson, Robert Louis, 1873
- Stewart, Donald E., 2729
- Stiff-person syndrome, 635
- Stigma, fear of, 2608
- Stigmatizing attitudes, world-wide, 4528–4529
- Still, George, 3587
- Stilling, type D attachment and, 1901
- Stilted speech, 4614
- Stimulants. *See also* Psychostimulants
  - for ADHD, adult manifestation, 3602, 3603*t*
  - antidepressants combined with, 3324
  - description of, 4614
  - in elderly patients, 2283
  - side effects
    - insomnia and, 2087
  - for somnolence, 2293
  - stimulant-related disorders and, **1280–1291,** 1286
    - ADHD and, 3592–3594, 3592*t*, 3788–3789
    - administration routes, 1284, 1285
    - amphetamine-type stimulants and, 1281
    - autism spectrum disorder and, 3586
    - bath salts, 1282
    - cocaine and, 1280–1281
    - comorbidity, 1286
    - comparative nosology, 1281
    - compulsive use, 1283–1284
    - course and prognosis, 1287–1288
    - defined, 1280, 1280*t*
    - diagnosis and clinical features, 1285–1286
    - differential diagnoses, 1287
    - dynorphin, 1284
    - epidemiology, 1281–1282
    - etiology, 1282–1284
    - executive dysfunction, 1284

- genetics, 1285
- history, 1280–1281
- intoxication, 1285–1286, 1288
- laboratory exam, 1287
- medical complications, 1286–1287
- metabolic factors, 1284, 3796
- negative reinforcement mechanisms, 1283
- neuroendocrine system, 1284
- neurotransmission, 1283
- overview, 1280
- pathology, 1287
- positive reinforcement mechanisms, 1283
- psychotic disorders and, 1250*t*, 1252
- second-messenger systems, 1284
- stimulant-induced psychotic disorder, 1286
- stimulant use disorder, 1285
- tic disorders, 3647, 3661
- toxicity, 1286–1287, 1288
- traumatic brain injury, 1244
- treatment, 1288–1290
  - of intoxication and withdrawal, 1288
  - pharmacotherapy, 1289–1290
  - psychosocial, 1288–1289
- unspecified, 1286
- withdrawal, 1286, 1288

- Stimulation-oriented therapies, 1199–1200
- Stimuli
  - conditional and unconditional
    - and learning theory, 717
  - contextual, in Pavlovian learning, 721, 721*t*
  - description of, 4614
- Stimulus control (SC) therapy
  - for insomnia, 2089
  - self-control strategy, 2697
  - TTM, 1795
- Stockholm Outcome Study of Psychoanalysis and Psychotherapy (STOPP), 2674
- Stockholm syndrome, 4615

- Stoller, Robert, 2025–2026
- Stomach, mechanoreception, 362*f*
- Stone, Jon, 1832
- Stone, Leo, 2644, 2660
- Stonewall riots, 1988
- Stop codons (UGA, UAA, UAG), 151
- Stop–start technique, 1976
- STP. *See* 2.5-dimethoxy-4-methylamphetamine
- Strachey, Alix
  - as Freud translator, 2644–2645
  - psychoanalysis and, 2641
- Strachey, James, 2638, 2644–2645
- Strang, N.M., 3365
- Strassman, Rick, 2887
- Strategic therapy, **2816–2818**
  - clinical and ethical issues, 2817–2818
  - family and couples, 2752*t*, 2753
  - history, 2816
  - techniques, 2817
  - theoretical issues, 2816–2817
- Strattera. *See* Atomoxetine
- Street drugs, description of, 4615
- Strength-vulnerability model of family, 4193–4194
- Strengths and Difficulties Questionnaire (SDQ), 3397
- Strengths and Weaknesses of ADHD-symptoms and Normal-behavior (SWAN) Scale, 3401
- Strephosymbolia, 4615
- Streptococcal infections
  - OCD and, 1788
  - PANDAS and, 1797
  - tic disorders and, 3638–3639, 3652–3653
- Streptococcal (group A β-hemolytic streptococcus) (PANDAS) infections
  - symptomatology of, 1116
- Stress. *See also* Posttraumatic stress disorder (PTSD)
  - acute, 188–190
    - burn injuries in children and, 2385
    - effects on immunity, 190*t*
    - immune responses and, 2320, 2338–2339

- adaption responses, 2336
- addictive disorders and, 2046–2047
- adjustment disorders and, 2117–2118
- anatomy of, 2280
- animal models of, 321
- asthma and, 2224–2225
- biological, mood disorders and, 1611, 1611*f*
- blood pressure and, 2316*f*
- buffers against, 2848–2849
- cardiac function and, 2316*f*
- chronic, 190–191
  - diabetes and, 2231–2232
  - effects on immunity, 190*t*
  - immune responses and, 2320
- cultural concept of, 790
- definition of, 2315
- depression and, 1609, 1712–1713
- depressogenic effects of
  - puberty and, 2507
- description of, 4615
- dopamine and, 1745
- early-life, 191*f*
  - anxiety disorders, 1740, 1745
- gene expression and, 2323
- genomic testing and, 2536
- history of the literature on, 2315–2317
- illness and, 187, 189
- immunity and, 2320
- induced hypothermia, 321
- inflammatory bowel disease and, 2201
- insulin sensitivity and, 176
- management, early PTSD, 1824
- maternal deprivation and, 1713
- measuring of, 2317–2318
- of migration, 791–793
- mild, 188–190
- negative thoughts related to, 2326*t*

- neural-immune interactions, 188*f*
- neurochemistry of, 1742
- norepinephrine and, 1742
- operational, disaster and military operations and, 2567–2568
- physician mental health and, 2599
- prenatal, 2046–2047
- prolactin levels and, 2249
- psychocutaneous disorders and, 2332–2326
- psychogenic
  - reprodysfunction and, 2508
- psychoneuroimmunology, 2318–2322
- psychosocial, 2332–2326
- reactivity, epigenetic transmission, 394–396
- in regimen non-adherence, 4300
- resilience to, 2317
- responses to
  - HPA axis and, 167–168, 193–194
  - neurobiology and, 1709
- serotonin and, 1745–1746
- serotonin transporter and, 1710, 1710*f*
- severe, 190–191
- social processes and, 191–192
- stress–immune spectrum, 2336
- stressors and
  - delirium and, 1182
  - life events as, 1116–1117
  - schizophrenia development, 1451, 1455
- vulnerability to, 2317

- Stress cardiomyopathy, 2194
- Stress immunity, 4615
- Stress inoculation training, 1773
- Stress management
  - in cardiac rehabilitation programs, 2198
  - cognitive-behavioral approaches, 2324–2328
  - daily behavior diary format, 2325*t*
  - future directions, 2328–2329
  - pharmacological, 2324

- problem solving and, 2327–2328
- time management and, 2327

- Stress reactions, 4615
- Stress-related disorders
  - after TBI, 536–537
  - in cancer patients, 2259
  - phenomenology of, 1751–1753
- Stress resilience in virtual environments (STRIVE), 2560
- Stress/vulnerability model, 2883
- Stress–diathesis model, 1670, 3638, 3652
- Stress–immune spectrum, 2336, 2337*f*
- Stressors, physiological responses to, 2318
- Stria terminalis
  - anatomy of, 36*f*
  - route, 38
- Striatonigral degeneration. *See* Multiple system atrophy (MSA)
- Striatonigralstriatal (CNS) projections, 377*f*
- Striatonigralstriatal projections, monkey, 29*f*
- Striatum
  - dopamine receptors in, 66
  - dysfunction, depression and, 1709
  - neuroimaging
    - BP-I, 3464–3466
    - drives and instincts and, 2815
    - major depressive disorder, 3464–3466, 3465*f*
  - OCD and, 1786
- Striosomes, primate, 29
- Stroke. *See also* Apoplexy; Cerebrovascular accidents (CVA)
  - aggression and, 2495
  - in aging brain, 3993
  - anosognosia after, 478
  - anxiety after, 476–477
    - course/prognosis in, 485–486, 486*f*
    - diagnosis of, 481
    - survival and, 487–488
  - apathy after, 477, 477*t*, 478*f*
    - diagnosis of, 482

    - etiology of, 480
    - treatment of, 488
  - bipolar disorder and, 486*f*
  - catastrophic reaction after, 478
    - diagnosis of, 482
    - etiology of, 480
  - cognitive impairment and, 484–485, 484*f*
  - definition of, 474–475
  - delirium and, 1186
  - depression after, 476
    - course/prognosis in, 484–485, 486*f*
    - diagnosis of, 481
    - etiology of, 478–479, 479*f*
  - mania after, 485
    - diagnosis of, 481
    - etiology of, 479
  - mania following, 476
  - mood disorder and, 1249
  - pathological laughter and crying (PLAC) after, 478
    - diagnosis of, 482
    - etiology of, 480
    - treatment of, 488–489
  - poststroke mortality, 485*f*
  - psychological factors in, 2194–2195
  - psychosis after, 477
    - etiology of, 479–480
    - treatment of, 488
  - and transactions, as social interactions, 2820
- Stroop Color-Word task, 4370
- Stroop paradigm, 699
- Stroop test, 402, 405
  - betrayal trauma theory and, 1903
  - hypnosis, 2713
- Structural dissociation theory (SDT)
  - amnestic personality states and, 1910
  - dissociation, 1903
- Structural family therapy, 2750, 2752*t*, 2753

- Structural imaging for dementia, 1197
- Structural Interview (Kernberg), 2165–2166
- Structural Interview for Schizotypy (SIS), 1470
- Structuralism, 705–706
- Structured assessment instruments. *See* Standardized assessment instruments
- Structured Clinical Assessment in Neuropsychiatry—Present State Examination (SCAN-PSE)
  - for psychiatric classifications, 1162
- Structured Clinical Interview for DSM-IV (SCID), 2685
- Structured Clinical Interview for DSM-IV-TR Axis I Personality Disorders (SCID-I), 1883, 1883*t*
- Structured Clinical Interview for DSM-IV-TR Axis II Personality Disorders (SCID-II), 1883, 1883*t*
- Structured Clinical Interview for DSM-IV-TR Dissociative Disorders, Revised (SCID-D-R), 1877*t*, 1878, 1882
  - dissociative amnesia, 1934
  - dissociative identity disorder, 1899
  - factitious dissociative identity disorder, 1916
- Structured Clinical Interview for DSM-5 (SCID), 1704
- Structured Clinical Interview for DSM-IV PDs (SCID II), 2161
- Structured interactional group psychotherapy, 4615
- Structured Interview for OCRDs, 1788
- Structured Interview of Reported Symptoms (SIRS), 1856, 1917, 2733
- Structured interviews, 4615
- Structured Inventory of Malingered Symptomatology (SIMS), 1856, 2733
- Structured Inventory of Reported Symptoms (SIRS), 2410
- Strupp, Hans, 2647
- Student Observation System, 2687
- *Studies on Hysteria* (Breuer and Freud), 852–854, 4615
- Study design. *See* Statistics
- *Study of a nervous disorder characterized by motor incoordination with echolalia and coprolalia* (Gilles de la Tourette), 4536
- *Study of Organ Inferiority and its Psychical Compensations* (Adler), 4537
- Stupor
  - catatonic, 1419, 1595
  - depressive syndrome and, 1636–1637
  - description of, 4615
  - mood disorders and, 1657

- Sturge–Weber syndrome, 456
- Stuttering, **3559–3565**
  - acquired, 457*f*
  - cessation of
    - characteristics of, 457*f*
  - clinical features, 3562–3563
  - comparative nosology, 3559–3560
  - course, 3563–3564
  - defined, 3559
  - description of, 4615
  - developmental, 457*f*
  - diagnosis, 3562–3563
  - differential diagnoses, 3563
  - epidemiology, 3560
  - etiology, 3560–3562
  - historical background, 3559–3560
  - prognosis, 3563–3564
  - treatment, 3564
- Styron, William, 2834, 2839, 4546
- Sub-bipolar dysthymia, 1642
- Subacute sclerosing panencephalitis, 596
- Subcaudate tractotomy, 3316, 3317–3318
- Subclinical psychotic experiences, 1452–1453
- Subcoma therapy. *See* Insulin coma therapy (ICT)
- Subconscious, description of, 4615
- Subcortical dementia, 683*f*
- Subcortical reward nodes, 365–366
- Subgenual anterior gyrus, 33
- Subgranular zone (SGZ), 42
- Subicular cortex, 33–34
- Subiculum, anatomy of, 34*f*
- Subjective personification, dissociative process and, 1906–1907
- Subjective units of distress (SUDS) measures
  - case study, 2690
  - for observation, 2686–2687
  - operant behavior therapy, 2872*f*
  - systematic desensitization and, 2689

- Subjectivity, description of, 4615
- Sublimation
  - character and, 2138
  - description of, 4615
  - difficulty with
    - personality disorders and, 2140
  - and mental health, 761
- Submissive immobility, 1890
- Substance abuse. *See also* Addiction; Substance use; Substance use disorders
  - after disasters, terrorism, or war, 2567*t*
  - after TBI, 537–538
  - by aging adult, 3957, 3973–3974
  - animal models of, 322–323
  - BDD and, 1791
  - bipolar II disorder and, 1651–1652
  - care post-incarceration, 4305
  - comorbidities, 4287, 4342, 4375–4376
  - description of, 4615
  - dissociative amnesia, 1908
  - dissociative disorders, 1888–1889
  - dopamine receptor gene, 1745
  - factitious disorder and, 1864
  - gambling and, 1805, 1808
  - and genetic selection, 778
  - global burden of, *4522–4523t*
  - group psychotherapy, 2745
  - and HIV infection, 4270
  - insomnia and, 2087–2088
  - insurance parity for, 4317, 4380
  - in jails and prisons, 4466–4467, 4467*t*
  - in LGBT children, 2002
  - neurosteroids and, 117–118
  - nonsanctioned interpersonal aggression and, 2476
  - over-the-counter medication, 4113–4115
  - palliative care in patients with, 2292
  - parasomnia due to, 2102–2103
  - in patient interview, 950–951

- problem recognition as barrier, 4384
- problem with, 2292
- psychiatric hospitals interface with, 4326*f*
- psychopharmacology, 2925
- PTSD and, 1822
- as public health issue, 4302
- rates among LGB persons, 2007
- schizophrenia induced by, 4101
- sex differences in, 382–384
- suicide and, 2612
- treatment
  - nefazodone, 3127
- treatment delay, *4527t*
- treatment gap, 4522
- violence and, 2490, 4391
- world-wide, 4517, 4520

- Substance Abuse and Mental Health Services Administration (SAMHSA), 1262–1263, 1378, 1550, 1558
  - action plan commissioned by, 4306
  - definitions from, 4340
  - grants from, 4321
  - health promotion help from, 4303
  - illness/wellness management training materials from, 4375
  - opioid maintenance treatment administration, 3133
  - in public health, 4303
  - toolkits from, 4363–4364
- Substance dependence. *See also* Addiction
  - after disasters, terrorism, or war, 2567*t*
  - D2 receptor availability, 300
  - dopamine transmission in, 301
  - intimate partner violence and, 2487
- Substance-induced disorders, 3071
- Substance intoxication, 373*t*
- Substance/medication-induced disorders
  - amnesia, dissociative amnesia *versus,* 1936
  - anxiety disorder, 1723, 1729
  - mild/major neurocognitive disorders, **1247–1248**
  - obsessive-compulsive and related disorder, **1255–1256**

- OCD *versus,* 1788, 1789
- psychotic, **1590–1592**
  - comparative nosology, 1590, 1591*t*
  - course and prognosis, 1592
  - diagnosis and clinical features, 1591
  - differential diagnoses, 1591–1592, 3730–3731
  - epidemiology, 1590
  - etiology, 1590–1591
  - history, 1590
- psychotic disorder, **1251–1253**
- sexual dysfunction, **1258**
- sleep disorder, **1256–1258**

- Substance P
  - direct pathway for, 30
  - function of, 2337
  - immunological effects of, 185*t*
  - localization of, 96
  - pain transmission and, 2388–2389
- Substance use. *See also* Substance abuse disorder(s)
  - ADHD and, 3230–3231, 3592, 3601, 3602
  - anabolic steroid-associated, 1402–1403
  - bipolar disorder and, 3690, 3704
  - child/adolescent
    - group therapy and, 3767–3768
    - laboratory screening for, 3407–3408
  - conduct disorder and, 3619
  - depression and, 646
  - emergency psychiatry evaluation for, 3444
  - family therapy and, 3783–3785
  - mental illness and
    - in adolescents, 646
    - in prenatally acquired HIV patients, 647–648
  - mood disorder and, 1249
  - neonatal abstinence syndrome, 3838–3839
  - organ transplantation and, 2363–2364
  - posttraumatic stress disorder and, 3705
  - preventive interventions for, 3898

- schizophrenia and, 1423, 1434, 1544, 1553, 3730–3731
- transplant evaluation and, 2361
- withdrawal
  - anxiety and, 2258
  - buspirone, 3012
  - description of, 4615
- in youth with HIV, 647

- Substance use disorders (SUDs), **1262–1404**
  - adolescent, **3857–3861**
    - epidemiology, 3858
    - evaluation, 3858–3859
    - future directions, 3861
    - overview, 3857–3858
    - prevention, 3859
    - screening, 3858–3859
    - treatment, 3859–3860, 3859*t*, 3860*t*
  - alcohol-induced disorders, **1270*t*, 1272–1273**
  - alcohol-related disorders, **1273–1274**
  - alcohol-use disorder, **1264–1279**
  - anabolic-androgenic steroid-related disorders, **1390–1404**
  - animal models, 371*t*
  - binge/intoxication, 370
  - bipolar disorder and, 1699
  - burn injuries and, 2374–2375
  - caffeine-related disorders, **1291–1303**
  - cancer patients with, 2266
  - cannabis-related disorders, **1303–1312**
  - classification of, 1158, 1168
  - cognitive therapy, 2767
  - current treatments, 384–385, 384*t*
  - DBT and, 2794
  - description of, 4615
  - diagnostic criteria, 370*t*
  - dissociative identity disorder and, 1913
  - GABA agonist-related disorders, **1374–1390**
  - gene expression and, 379–380, 380*f*
  - hallucinogen-related disorders, **1312–1327**

- HIV disease and treatment in, 579–581
- HIV infections and, 578–582
- HPA axis and, 167
- inhalant-related disorders, **1328–1342**
- intoxication
  - delirium and, 1182, 1189–1190
  - mental disorders due to, 1235–1236
- mood disorders and, 1655–1656, 1669
- neural circuits, 374*f*
- neural pathways, 376–378
- neuroimaging, 3475
- neuroscience of, **369–385**
- obesity and, 2215
- OCD and, 1788
- opioid use and related disorders, **1352–1373**
- overview, 1262–1264
  - defined, 1262
  - prevalence, 1262–1263
  - public policy issues, 1263
  - treatment, 1263
- pain related to, 2393
- pharmacotherapy, 3793
- preoccupation/anticipation, 371–372
- psychopharmacology, 3793
- relapses, 377–378
  - conditioned cue-induced, 378
  - drug-primed, 377–378
  - stress-induced, 378
- sedative-, hypnotic-, anxiolytic-related disorders, **1374–1390**
- somatic symptoms, 1828*f*
- stimulant-related disorders, **1280–1291**
- suicide and, 1669, 2613
- temperament and, 3487
- tobacco-related disorders, **1342–1352**
- in transplant populations, 2366–2367
- treatment
  - β-adrenergic receptor antagonists, 2954

        - anticonvulsants, 2968, 2970, 2973
        - comorbid PTSD and, 1823
    - withdrawal, 1235–1236
    - withdrawal delirium, 1182, 1190
    - withdrawal/negative affect, 371, 580–581
- Substantia innominata, 26
- Substantia nigra
    - anatomy of, 28
    - pars compacta, 29–30
    - pars reticulata, 80–81, 543
- Substantia nigra pars reticulata (SNr), 449
- Substitution, definition of, 4615
- Subsyndromal mood disorders, 1601
- Subthalamic nucleus (of Luys)
    - anatomy, 28
    - derivation of, 28
- Subventricular zone (SVZ)
    - function of, 42
    - postnatal neurogenesis in, 58–60
- Succinic semialdehyde dehydrogenase (SSADH) deficiency, 664
- Succinylcholine (Anectine)
    - description of, 4615
    - ECT and, 3292
- Sudarshan kriya (SK) yoga, 2557
- Sudden cardiac death
    - cardiac arrhythmias and, 2189–2190
- Sudden infant death syndrome (SIDS)
    - antihistamines and, 2983
    - factitious disorder imposed on another and, 1857
    - motor development delays, 3361
- Sudden unexplained deaths, 3075–3076
- SUDs. *See* Substance use disorders
- Sufi sect, 2582
- Sugar, fasting blood levels, 1022
- Suggestibility
    - hypnosis and, 2708, 2711
    - in posttraumatic stress disorder, 700

- trauma model of dissociation *versus,* 1898–1899
- Suggestion
  - description of, 4615
  - hysteria caused by, 1867
- Suicidal behavior
  - antidepressant-associated, 1672–1673, 1680
  - BDD and, 1790, 1791
  - dissociative disorders, 1884
  - dissociative identity disorder and, 1913
  - factitious disorder imposed on self and, 1859
  - mood disorders and, 1667–1676
    - clinical utility of risk factors, 1671–1672, 1672*t*
    - family history of, 1670–1671
    - prevention strategies, 1673–1676, 1676*t*
    - previous suicide attempt, 1670
    - relationship between, 1667–1668
  - pathological dissociation and, 1885
  - stress-diathesis model, 1670
- Suicidal ideation
  - BD, 1684
  - Beck on, 2620
  - description of, 4615
  - disasters and, 2568–2569
  - in LGBT children
    - rejection and, 2002
  - monitoring in BD, 1684
  - psychiatric assessments, 3975
  - Scale for Suicidal Ideation (SSI-C), 2615
  - sexual minorities and, 646
  - SSRIs and, 3166–3167
  - in terminally ill patients, 2313*t*
- Suicidality
  - acamprosate and, 3058
  - anticonvulsants and, 2966, 2975, 3086
  - antidepressants and, 2918–2919
  - lamotrigine and, 3085–3086
  - lithium reduction in, 3090

- nefazodone and, 3128
- SSRIs and, 3166–3167
- vilazodone effects on, 3345

- Suicide, **2610–2618**
  - acute crisis, treatment, 2616
  - adjustment disorder and, 2123
  - adolescents and, 2614–2615
  - antidepressants and, 1779–1780
  - antipsychotic-induced, 1528
  - assessment
    - for cancer patients, 2262–2263
  - assessment at interview, 949, 950, 954
  - attempts, 2611–2612
  - attempts by burning, 2383
  - barbiturates and, 2984
  - child/adolescent, **3680–3686, 3694–3700**
    - assessment, 3684, 3698
    - course, 3684, 3698
    - epidemiology of, 3680–3681, 3681*f*, 3694–3695, 3695*f*
    - prognosis, 3684, 3698
    - risk factors, 3681–3684, 3695–3698
    - treatment, 3684–3685, 3698–3699
  - children and, 2614–2615
  - clozapine use, for prevention, 1528
  - competency of older adult, 4235–4236
  - defined, 3680, 3694
  - dementia, 2615
  - depressive disorders, 3678, 3692
  - depressive syndrome and, 1637–1638
  - description of, 4615
  - elderly, 2615
  - by elderly cancer patients, 2283
  - emergency child psychiatry evaluation for, 3391, 3394*t*, 3437, 3442–3443
  - epidemiology, 2610–2612, 2611*t*, 2612*t*
  - epigenetic signatures, 400–401
  - in epilepsy patients, 517
  - future directions, 2622

- inquiry, 2619–2620
- interview scenarios, 961
- introduction, 2610
- in jails and prisons, 4469–4470, 4470*t*
- malpractice and, 2620
- medical school postvention plan, 2600–2601
- medical students, 2600–2601
- methods, 2611
- neurobiology, 2619
- personality disorders and, 2613–2614
- pharmacotherapy and, 3790
- physician, 2599–2600
- physician-assisted, 2313, 2622
- postvention response, 2600–2601
- prevention, 2613
  - mood disorders and, 1674–1676
  - public health perspective and, 2617–2618
- psychiatric disorders and, 2612–2613
- and psychiatric malpractice, 4437
- rates among LGB persons, 2006–2007
- risk
  - after TBI, 533
  - assessment, 2615
  - for cancer patients, 2262*t*
  - clinical interview, 2615
  - 5-HTT and, 1710, 1710*f*
  - in head and neck cancers, 2267
  - management, 2615–2616
  - mood disorder and, 1668–1672
  - pharmacotherapy, 2616
  - psychometric scales, 2615
  - psychotherapy, 2616–2617
  - SSRIs and, 1784
- schizophrenia and, 1423, 1435, 1528
  - treatment, 1528
- with selective serotonin reuptake inhibitor (SSRI), 3378
- somatic disorders and, 2614

- specific learning disorder and, 3529
- SSRI side effect, 3718
- stimulant-induced, 1286
- survivors, 2617
- transgender people and, 2035–2036
- trends, 2611
- victim-precipitated, 2622
- world-wide
  - prevention of, 4528
  - rates of, 4520, 4522

- Suicide attempts, 2619–2620
  - BD with ADHD and, 1699
  - risk factors, 2619*t*
  - suicide completion *versus,* 2618–2619
  - therapy, 2620–2622
  - treatment
    - malpractice and, 2620
- Suicide behavior risk factor, 3684, 3698
- Suicide-by-cop, 2622
- Suicide cluster. *See* Cluster suicides
- Suicide completion, definition of, 2618
- Suicide Intent Scale (SIS), 2615
- Sulfonamides, side effects of, 571
- Sullivan, Harry Stack, 922f, 2644
  - on depression, 1665
  - developmental theories, 923, 923*t*
  - interpersonal psychotherapy, 924
  - *The Interpersonal Theory of Psychiatry,* 4540
  - IPT and, 2775
  - psychopathology theories, 924
  - on psychosis, 2868, 2870, 2875–2876
  - therapeutic communities, 922
  - work of, 4615
- Sulpiride (Meresa), 3070
- Sumatriptan (Imitrex)
  - for headaches, 70
  - for migraine, 625

- Summarizing in patient interview, 957
- Sundowner's syndrome. *See* Sundowning
- Sundowning, definition of, 1186
- Superego
  - current investigations of, 873
  - description of, 760, 4615
  - evolution of, 872–873
  - historical development, 872
  - mood and, 1660–1661
  - origins of, 872
  - Strachey's translation of Freud on, 2645
- Superior frontal gyrus (SFG)
  - cross-section, 9*f*
  - lesions, 653
- Superior longitudinal fasciculus, 12, 15*f*
- Superior sagittal sinus, 14
- Superior temporal sulcus, 8
- Superstitions, OCD and, 3699
- Supply side interventions
  - for care improvement, 4348
  - for families of persons with schizophrenia, 4350
- Support groups
  - for caregivers, 4210
  - as hospital role, 4345
  - for sexual addiction, 2057
- Support vector machines, 841
- Supported employment (SE), 1534–1535, 1536*f*, 1552, 3732–3733
- Supportive hypnotherapy, 2716
- Supportive nondirective therapy (SNDT), 2201–2202
- Supportive psychotherapy, 2666–2668
  - aggression, 2498–2499
    - TBI, 2502
  - case study, 2668
  - contraindications, 2667
  - description of, 4615
  - goals, 2667
  - indications, 2667

- limitations, 2667–2668
- main approach, 2667
- psychoanalysis and, 2666–2667

- Supportive services
  - employment, 1534–1535, 1536*f*, 1552, 3732–3733
  - intellectual disability and, 3498, 3516–3518, 3516*t*
- Suppression
  - description of, 4615
  - and mental health, 761
  - of psychiatric scandal, 4553
- Suprachiasmatic nucleus (SCN)
  - afferent projections, 205–206
  - efferent projections, 206
  - extra zeitgebers, 210
  - location of, 204, 204*f*, 205
  - molecular clockwork, 206–207
  - neuron firing, 205*f*
- Supraoptic nuclei, 36
- Supreme Court decisions
  - *Brown v. Plata* (2011), 4400
  - *Halderman v. Pennhurst State School & Hospital* (1977), 4336
  - *O'Connor v. Donaldson* (1975), 4336
  - *Olmstead v. L.C.* (1999), 4337
  - *Youngberg v. Romeo* (1982), 4336
- Surface coils, MR, 264, 265*f*
- *Surgeon General's Report on Mental Health,* 822
- Surgery
  - child/adolescent response to, 3847
  - differential diagnoses, 3569
  - psychological issues and, 2274
  - of tic disorders, 3648, 3662
- Surgical patients, pre-and postoperative, 4358
- Surgical procedures, hypnosis and, 2725–2726, 2725*f*
- Surmontil. *See* Trimipramine
- Surplus reality, psychodrama and, 2858
- Surrogate decision makers, 2305
- Surrogate mothers, 2510*t*

- Surrogates, description of, 4615
- Survival analysis, 843–844
- Survivors of Suicide (SOS) groups, 2619
- Susceptibility genes, mood disorders, 1629–1630
- Suso, Heinrich, 2582
- Suspected Non-Alzheimer Pathology (SNAP), 1227
- Susto, description of, 4615
- *Suttee,* 2383
- Sutter–Eyberg Student Behavior Inventory (SESBI-R), 3402
- Sutton, Thomas, 4534
- Suvorexant, PTSD, 1823
- Swallowing abnormalities, 555
- Swanson, Nolan, and Pelham (SNAP-IV) Rating Scale, 3401
- Swanson Kotkin, Agler M-Flynn, and Pelham (SKAMP) rating scale, 3401
- Sweating, SSRI effects on, 3156
- Swept field NMR, 263
- SWS. *See* Slow-wave sleep (SWS)
- *Sybil,* 1870, 1873
- Sydenham, Thomas
  - conversion disorder, 1839
  - *The Entire Works of Dr. Thomas Sydenham,* 4534
  - on hypochondria, 1838
  - somatic presentations, 1828
- Sydenham chorea
  - streptococcal infections and, 201, 3638, 3652
- Symbiosis
  - description of, 4615
  - Fromm, 921
- Symbiotic psychosis, 4615
- Symbolization
  - description of, 4615–4616
  - semantic learning and, 2129
- Symmetrel. *See* Amantadine
- Sympathectomy, 187
- Sympathetic nervous system (SNS)
  - catecholamines and, 193
  - description of, 4616

- fear conditioning, PTSD and, 1818
- immune system and, 2319
- innervation of lymphatic tissue, 189*f*
- PTSD and, 1819

- Sympathomimetics, **3219–3233,** 4616
  - ADHD, 3219, 3223, 3227–3229, 3232
  - apathy, 3229
  - conduct disorder, 3228
  - depression, treatment-resistant, 3229
  - dosage/administration, 3220*t*–3223*t*
  - drug interactions, 3227*t*, 3233
  - drug studies, 3227–3229
  - history, 3219
  - mechanisms of action, 3223–3224, 3231, 3232
  - ODD, 3228
  - organ/system effects of, 3230
  - pharmacological actions, 3224–3227, 3225*t*–3226*t*, 3227*t*, 3231, 3232–3233
  - precautions/adverse reactions, 3229–3231, 3229*t*, 3237*t*
  - seizure disorder, 3229
  - sexual dysfunction and, 1972
  - therapeutic indications, 3227–3229, 3231–3232
- Sympathy, description of, 4616
- Symptom Checklist 90 (SCL-90), 1836
- Symptom Checklist-90 (SCL-90-R), 1920
- Symptom formation, description of, 4616
- Symptom provocation paradigms, 1759
- Symptom substitution, 4616
- Symptomatic psychosis, 4616
- Symptomatology
  - ambiguity of, 1115–1116
  - genetic considerations for, 1116
  - intrauterine considerations for, 1116
  - scales for child evaluations, 3399–3402
- Symptoms
  - assessment of, 2292
  - description of, 4616
- Synapse proteome, 162

- Synapses
  - cognition and, 430–431
  - description of, 4616
  - function of, 3
  - neuronal, in aging brain, 3990–3991
  - pruning
    - autism spectrum disorder and, 3577
  - pruning of, 451
  - tripartite, 5
- Synaptic cleft, 103
- Synaptic plasticity, 128
  - addiction and, 380–381, 381*f*
  - neurotrophins and, 100
- Synaptogenesis, 440–441
- Syncope
  - clinical features of, 2193
  - course/prognosis, 2194
  - definitions, 2193
  - diagnosis, 2193
  - psychological factors affecting, 2194
  - seizures *versus,* 2193
  - treatment, 2194
- Syncretic thought, 4616
- Syndrome of approximate answers, 1859
- Syndrome of inappropriate antidiuretic hormone relese (SIADH)
  - in bulimia nervosa, 2075
  - presentation of, 2270
- Syndromes, definition of, 2443
- Synesthesia, description of, 4616
- Synkinesia, 3539–3540
- Syntactical (central) aphasia, 4616
- Syntaxic mode of experience, 4616
- Synthetic cathinones, 1282
- Syntropy, description of, 4616
- α-Synuclein
  - in aging brain, 3991, 3992
  - intraneuronal aggregates of, 546

- transformation of, 542
- α-Synuclein gene *(SNCA)* mutations, 543
- Syphilis
  - description of, 4616
  - laboratory evaluation, 1012
  - mania without prior bipolarity and, 1649
  - primary, 592
- Syphilitic meningitis, 592
- Systematic reviews, 2861
- Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) trial, 1695, 1703, 3005
- Systematized delusions, 1122, 4573
- Systematized dissociative amnesia, 1881*t*
- Systematizied delusions, 4616
- Systemic desensitization
  - anxiety disorders and, 1774
  - description of, 4616
  - hierarchy and, 2689–2690
  - posttraumatic food aversions and, 3627, 3641
  - as somatic therapy, 2557
- Systemic lupus erythematosus, 1015
- Systems-based approach, community psychiatry, 4288
- Systems biology, field of, 162–163
- Systems of care, 4341
- Systems-oriented group therapy, 2738
- Szasz, Thomas, 4332

T

- T cells
  - CD4
    - decline of, 566*f*
    - HIV infection and, 565–566
  - formation of, 179
  - schizophrenia and, 200
- T (training) groups, 4617
- T helper (Th) cells
  - atopic dermatitis and, 2341
  - IFN-γ secretion by, 181
  - interleukin secretion by, 181
- T regulatory (Treg) cells, 181
- t-tests, description of, 4618
- Tablets, electronic, 4279
- Taboos, description of, 760, 4616
- Tachylogia. *See* Logorrhea
- Tacrine (Cognex), **3041–3042**
  - Alzheimer disease treatment, 1207–1209
  - chemistry, 3041, 3041*f*
  - cholinesterase inhibition, 4142
  - dosage/administration, 3042
  - drug interactions, 3042
  - drug studies, 3041
  - hepatic monitoring/rechallenge, 3042
  - history, 3040
  - laboratory references, 3042
  - organ/system effects of, 3042
  - pharmacological actions, 3041
  - precautions/adverse effects, 3042
  - therapeutic indications, 3042
- Tacrolimus
  - calcineurin inhibition by, 124
  - drug interactions, 3131
  - neurotoxic effects, 2371
  - side effects of, 2198
- Tactile amnesia. *See* Neurological amnesia

- Tactile hallucinations, 1413, 4616
- Tactile stimulation, self-produced, 335
- Tadalafil (Cialis)
  - ED, 3338–3339, 3339*t*
  - for sexual dysfunction, 1978, 4261
  - treatment of ED, 107
- Tagamet. *See* Cimetidine
- TAI. *See* Traumatic axonal injury (TAI)
- *Taijiuan* as somatic therapy, 2556–2557
- Tail-suspension test, rodent, 321
- Takotsubo cardiomyopathy, 2194
- Talion, description of, 4616
- Tamoxifen (Nolvadex)
  - acute mania, 1687, 1689*t*, 1690
  - antiestrogenic properties of, 1980
  - for breast cancer, 171
  - drug interactions
    - with antidepressants, 2269
  - side effects
    - depression, 2260
  - visual disturbances and, 2272
- Tandem pore (K_{2T}) channels, 139
- Tangentiality
  - description of, 4616
  - speech, 1420
  - of thought, 1120
- Tanner stages. *See* Sexual maturity rating (SMR)
- Tapentadol
  - duration of action, 2395*t*
  - mechanism of action, 2395
- Taphophobia, 4616
- Taractan. *See* Chlorprothixene
- Tarasoff decisions, 4616
  - confidentiality and, 2670–2671
- *Tarasoff v. Regents of the University of California,* 4435
- Tardive akathisia, medication-induced, 2941, 2942–2943
  - clinical presentation, 2942

- pathophysiology, 2943
- treatment options/implications, 2943

- Tardive dyskinesia (TD). *See also* Choreiform disorders
  - anticholinergics and, 2958, 2959
  - antipsychotic drug effects, 4137
  - antipsychotics side effect, 1526–1527, 1529
  - clinical manifestations of, 1148
  - description of, 4616
  - dopamine 2 receptor agonists and, 2271
  - intellectual disability and, 3515
  - management of, 3074–3075
  - neuroleptic-induced, 2941–2942
    - clinical presentation of, 2941–2942
    - FGAs, 3074–3075
    - pathophysiology of, 2942
    - treatment options/implications, 2942
  - psychiatric malpractice, 4437
  - risk of, 171
  - stereotypic movement disorders, 3672, 3686
  - striatal dopamine transmission and, 62
  - treatment
    - benzodiazepine, 2997
- Tardive dystonia
  - anticholinergics and, 2959
  - medication-induced, 2942–2943
    - clinical presentation, 2942
    - pathophysiology, 2943
    - treatment options/implications, 2943
- Target, Mary, 2870
- TARs. *See* Transcriptionally active regions (TARs)
- Tart, Charles
  - dissociative identity disorder, 1899
  - hypnosis, 2710
- Tasimelteon
  - chemistry, 3101–3102, 3102*f*
  - dosage/administration, 3106
  - drug interactions, 3106, 3106*t*

- drug studies, 3104
- laboratory interference, 3106
- organ/system effects of, 3104–3105
- pharmacological actions, 3102–3103, 3103*t*
- precautions/adverse reactions, 3105–3106
- therapeutic indications, 3105, 3105*t*

- Task Force on Defending Childhood Against Violence, 4298
- Task motivational theory, 2710
- Task-oriented treatments, 3542–3543
- TAT. *See* Thematic Apperception Test
- TATA boxes, 159–160
- Tattooing during adolescence, 3372–3373
- Tatum, Edward, 149
- Tauopathy, primary age-related (PART), 3991
- Tax Equity and Fiscal Responsibility Act (TEFRA), 4215–4218, 4336
- Taxon model
  - dissociative disorders, 1895–1896, 1895*t*
    - treatment, 1896
- Taylor, Shelley, 2847
- Taylor Manifest Anxiety Scales, 996
- Tay–Sachs disease (TSD). *See also* GM2 gangliosidosis
  - late onset, 628
  - presentation, 657
- TBI. *See* Traumatic brain injury (TBI)
- TCAs. *See* Tricyclic antidepressants
- TD. *See* Tardive dyskinesia
- TDCS. *See* Transcranial direct current stimulation
- TDP43 in aging brain, 3991
- Teacher Report Form (TRF/6 to 18), 3398–3399
- Teaching health centers, 4319
- Team participant-leader and mentor, 4300–4301
- Technical neutrality, 2677
- Technicium, ^{99m}Tc, 287
- Technology, **4274–4281**. *See also* Internet
  - applications
    - caregiver applications, 4279
    - games, 4278–4279

    - mobile health, 4276–4277
    - sensors/self-monitoring devices, 4277–4278
    - smart homes and senior design, 4278, 4278*f*
    - social media, 4279
    - tablets, 4279
    - telehealth, 4275–4276
  - barriers and limitations
    - age-related impairment, 4280
    - billing, 4280
    - information overload, 4280
    - integration with healthcare system, 4280
    - literacy and usability, 4280
    - privacy and ethics, 4280
    - regulation lack, 4280
  - definitions, 4274–4275, 4274*t*
  - future directions
    - big data, 4280
    - robotics, 4279–4280
  - for seniors (*See also* Geriatric psychiatry)
- Teen online problem solving (TOPS)
  - after TBI, 653
- Teeth, ECT effects on, 3296
- TEFRA. *See* Tax Equity and Fiscal Responsibility Act
- Tegretol. *See* Carbamazepine
- Tele, sociodrama, sociometry and, 2857
- Telehealth, 4275–4276
- Telemarketing, elder abuse, 4251
- Telemedicine, 3976, 4275–4276
- Telencephalon
  - appetite and, 365–366
  - formation of, 41
  - in human development, 8
- Teleological mode
  - BPD and, 2896, 2897, 2897*t*
- Telepathy, description of, 4616
- Telephone consultation, in DBT, 2790–2791, 2794
- Telephone scams, elder abuse, 4251

- Telephone sessions, psychotherapy and, 2672
- Telephone support
  - suicide risk, 2617
- Telepsychiatry, 964, 4275–4276
  - for cancer patients, 2253
  - description of, 4616
  - in jails and prisons, 4476
  - privacy of patient, 4435
- Teletherapy, 2672–2673
- Television
  - aggressive behavior and, 2485
  - dissociative disorder patients on, 1873
- Telomere lengths, 4284
- Temazepam (Restoril)
  - for elderly patients with cancer, 2283
  - metabolism, 2993
  - for PTSD prevention of, 2379
- Temperament, 3479–3490
  - activity level, 3481
  - ADHD and, 3485–3486, 3485*t*
  - affective, 1602–1603, 1622*f*, 1632–1633, 1633*t*
  - anxiety disorders and, 1735
    - child/adolescent, 3486, 3486*f*
  - autism spectrum disorder and, 3487
  - biotypes, 2137*t*
  - categories of, 3388, 3389*t*
  - character, identity and, 2136*t*, 2145
  - in children, 2159
  - cognitive factors, 3488
  - component traits, 2129
  - configurations of, 2141*f*
  - continuity, 3482
  - definition of, 2128, 3480, 3480*t*
  - description of, 4616
  - development of, in childhood, 3352–3353*t*, 3361
  - dimensions of, 2128*t*, 3480, 3481*t*
  - disruptive behavior disorders, 3485–3486, 3485*t*

- eating disorders, 3487
- emotional regulation and constraint, 3481
- emotional state of, 2134*t*
- environmental factors, 3483
- features of, 3482
- foster care children and, 3820
- further implications, 3488–3489
- gender differences, 3482
- generalized anxiety disorder, child/adolescent and, 3712
- genetic influences on, 3482–3483, 3483*t*
- genomic approaches, 2136–2138, 2137*t*
- goodness of fit and, 3488, 3489*f*
- harm avoidance, 2130–2131
- heritability, 1711
- historical background, 3479–3480
- infantile anorexia and, 3624, 3638
- mood disorders and, 1610–1611, 1611*f*, 1623, 3486, 3486*f*
- negative emotion threshold, 3480
- neurobiology and etiology, 3482–3484
- neurotransmitters and, 3483–3484
- OCD risk and, 1786
- parenting and, 3488, 3489*f*
- peers and, 3488
- personality disorder risk and, 2166, 2166*t*
- polarity of episodes and, 1651
- positive emotions, 3481
- psychiatric disorder trajectory, 3487
- psychiatric disorders, 3484–3488, 3485*t*
- psychobiology of, 2128–2133
- psychopathology and, 3484–3485, 3485*t*, 3488
- resiliency and, 3488
- social affiliation, 3481
- substance-use disorders, 3487
- as symptomatology consideration, 1116
- TCI scales, 2164*t*
- theory of, 3479
- traits and, 2163*f*

- trauma and, 3488
- types, child, 3479, 3479*t*, 3481–3482, 3482*f*
- Temperament and, 3487
- Temperament and Character Inventory (TCI)
  - in clinical practice, 2142, 2161
  - PD symptoms predicted by, 2162*t*
  - scales and subscales, 2164*t*
  - temperament dimensions evaluated by, 2162–2164
- Temperament Evaluation of Memphis, Pisa, Paris and San Diego Autoquestionnaire (TEMPS-A), 1633
- Temperament Questionnaire
  - Five Factor Model of Temperament
    - as child psychiatric evaluation consideration, 3378
    - disruptive behavior disorders and, 3609
- Temperature regulation. *See also* Thermoregulation
  - DA receptor antagonists and, 3077
- Temporal conjunction, 745
- Temporal-difference reinforcement learning (TDRL) models, 434–435
- Temporal lobe epilepsy. *See also* Complex partial seizures (CPS)
  - aggression and, 2496
  - post-traumatic, 525
- Temporal lobectomy
  - for epilepsy, 651
- Temporal lobes
  - in amnesia, 732
  - anatomy of, 8
  - lesions of, 31
  - mood, thalamus and, 2814
  - tumors of, 2267
- Temporal sclerosis, right-sided, 518*f*
- Temporary admissions, 4616
- Temporary workers, 794, 794*t*
- Temporomandibular disorder (TMD), 2392
- Temporoparietal junction (TPJ), 1891
- Ten-Country Study (Determinants of Outcome of Severe Mental Disorders), 1427–1428, 1431–1433, 1431*t*, 1433*t*
- Tend-and-befriend responses, 2507
- Tenex. *See* Guanfacine

- Tennessee Self-Concept Scale (TSCS), 998*t*
- Tenormin. *See* Atenolol
- Tension, description of, 4616
- Tension-type headaches
  - definition of, 620
  - gender ratios, 621
  - treatment
    - nonpharmacologic, 626
    - pharmacologic, 625–626
- Teratogenicity
  - β-adrenergic receptor antagonists, 2953
  - barbiturates, 2988
  - benzodiazepine, 2999
  - calcium channel inhibitors, 3020
  - carbamazepine, 3037
  - lamotrigine, 3086
  - melatonin, 3105
  - mirtazapine, 3109
  - tricyclic and tetracyclic antidepressant, 3263
  - valproate, 3276–3277
  - vilazodone, 3346
- Terfenadine (Seldane)
  - drug interactions, 2983, 3131
  - organ/system effects of, cardiovascular, 2983
- Terminal care
  - refusal of food and water, 2313
  - treatment decisions, 2312–2313, 2313*t*
- Terminal delirium, assessment of, 2256
- Terminal illnesses
  - disease-specific issues, 2302
  - group therapies, 2302
  - sibling response to, 2311
  - supportive care in, 2300
  - therapy in, 2295*t*
  - transplant candidate
    - informed consent, 2362
- Terminal insomnia, 4616

- Termination in CBT, 4186
- Terror, speechless, 699–700
- Terrorism, 4552. *See also* Posttraumatic stress disorder (PTSD)
  - child/adolescent, **3913–3925**
    - acts of terrorism and, literature review of, 3917–3921
    - adolescents, 3915, 3916
    - community-based interventions, 3921
    - community factors, 3916
    - constructive effects, 3917
    - defined, 3914
    - destructive effects, 3917
    - factors, 3915–3916
    - family factors, 3916
    - family interventions, 3922–3924, 3922*t*
    - future directions, 3924–3925
    - group therapy for, 3764–3765
    - history, 3703, 3915
    - interventions, 3920, 3921–3924, 3921*t*, 3922*t*, 3923*t*
    - latency-aged, 3915, 3916
    - media coverage and, 3916, 3919–3920, 3923*t*
    - neurobiology of, 3915
    - overview, 3913–3914
    - predisposing risk factors, 3915
    - Project Liberty, 3920–3921
    - reactions, 3914, 3916, 3919–3920
    - resilience, 3915–3916
    - risk and protective factors for, 3914
    - school-based interventions, 3921, 3922*t*
  - description of, 4616
  - psychiatric syndromes in aftermath of, 2567–2568
  - stress responses and, 2566–2568
- Tertiary prevention. *See also* Preventive psychiatry
  - in public health, 4303
- Test of Adaptive Behaviors (TABS), 4372
- Test-retest reliability, 3394–3395
- Testamentary capacity, 4438, 4616
- Testes

- feminization of, 2506
- fetal
  - testosterone secretion by, 2505
- sexual response cycle, 1960*t*

- Testicular cancer, 2268
- Testimonial privilege, 4617
- Testing batteries. *See* Standardized assessment instruments
- Testosterone, **170**
  - aggression and, 2475–2476, 2480–2481
  - brain metabolism of, 2418
  - domestic violence and, 2487
  - ED, 3339–3340
  - fetal exposure to, 2418
  - function of, 1955–1956
  - history, 1390–1391
  - libido and, 1956
  - prenatal, 1956*f*
  - production of, 2020, 2243
  - reduction of, 2020
  - sex offenders and, 2503
  - sexual aggression and, 2043
  - for sexual dysfunction, 4261–4262
  - in women, 1980
- Testosterone cypionate (Depo-Testosterone)
  - aggressive behavior and, 2493
- Tests of Malingering and Memory (TOMM), 993, 1856, 2733
  - detection of malingering with, 2410
  - in motivation assessments, 4370
- Tests of significance, 4617
- Tetanus toxin, 73
- Tetrabenazine (Xenazine)
  - for chorea, 555
  - TS, 3072
- Tetracyclic antidepressants (TCAs), **3253–3266**. *See also specific drugs*
  - blood concentrations, 3258–3259
  - chemistry, 3254, 3254*f*
  - description of, 4617

- dosage/administration, 3264–3265, 3265*t*
- drug interactions, 3263–3264
- drug studies, 3259–3260
- laboratory interferences, 3264
- organ/system effects of, 3260–3263
- pharmacological actions, 3254–3259
  - pharmacodynamic, 3256–3259, 3263
  - pharmacokinetic, 3254–3256, 3255*t*, 3257*t*, 3263–3264
- precautions/adverse reactions, 3262–3263
- therapeutic indications, 3261–3262

- Tetrahydrobiopterin (BH4)
  - deficiency of, 663
  - monoamine synthesis and, 196
- Tetrahydrocannabinol (THC). *See also* Cannabinoids; Cannabis
  - adverse effects of, 1309
  - cannabis preparations, 1303–1304
  - description of, 4617
  - function of, 110
  - intoxication delirium and, 1306–1307
  - laboratory examination, 1310
  - pharmacology, 1305–1306
  - reproductive effects of, 1310
  - therapeutic effects, 1311
  - withdrawal and tolerance of, 1307–1308
- Tetrahydrocannabinols, 110
- Tetrahydrocannabivarin, 115
- Tetrahydrodeoxycorticosterone (THDOC), 1748
- 3α,5α-Tetrahydroprogesterone (THP), 1748
- Tetrodotoxin (TTX) action, 136
- *Texas, Lawrence and Garner v.,* 2000
- Text mining, gene expression and, 163
- Thalamic connectivity, altered, 422–423, 423*f*
- Thalamic nuclei, 21–22
- Thalamocortical association systems, 25–26
- Thalamocortical dysrhythmia
  - QEEG analysis, 1037–1038
- Thalamocortical dysrhythmic syndrome, 750–751

- Thalamocortical motor systems, 24–25
- Thalamocortical sensory systems, 23–24
- Thalamotomy, 547
- Thalamus
  - aggression and, 2477
  - coronal section, 12*f*
  - death feigning and, 1890
  - description of, 21–26
  - dorsal
    - nuclei, 22*f*
  - dysfunction, depression and, 1709
  - embryonic origins of, 8
  - functional circuitry, 23–28
  - gamma-band function
    - resonant columns, 745–748, 749*f*
  - hypnosis, 2712
  - mood, temporal lobes and, 2814
  - neuroimaging
    - OCD, 3469–3470, 3469*f*, 3470*f*
  - oscillatory properties, 745–747, 746*f*
  - psychodynamic principles and, 2814
  - schizophrenia and, 1459
- Thalidomide (Thalomid)
  - teratogenic effects of, 56
  - Thalidomide Tragedy of 1962, 2931
- Thalidomide, Tragedy of 1962, 2931
- Thalomid. *See* Thalidomide
- Thanatology, 4617
- Thanatophobia, 4617
- Thanatos, 4617
- Thatcher, Robert, 218–219
- THC (delta-9-tetrahydrocannabinol). *See* Cannabis use and cannabis-related disorders
- The Wellness Community (TWC), 2254
- L-theanine, 2550
- Theatre technique, in hypnotherapy, 2716
- Theft, elder abuse, 4251
- Theile, Leslie Paul, 2840

- Thelarche
  - definition of, 2506
- *Thematic Apperception Test* (Murray), 4537
- Thematic Apperception Test (TAT), 1004*t*, 1005
- Theophrastus, 934
- Theophylline (Bronkodyl)
  - caffeine and, 1295
  - drug interactions, 3164
  - ECT interactions with, 3294
  - side effects, 2229
- *Theoria medica vera* (Stahl), 4534
- Theories, definition of, 4617
- Theory of mind (ToM)
  - autism spectrum disorder and, 697, 3581
  - on cognitive development, 3360, 3370–3371
  - and consciousness, 692–693
  - schizophrenia symptoms and, 1417
  - social (pragmatic) communication disorder and, 3553
- Therapeutic alliances
  - CBT, psychosis and, 2883–2884
  - description of, 4617
  - MBT, 2898, 2898*t*
  - positive psychologies and, 2852
  - psychoanalysis and, 2654–2655
  - psychodynamic psychotherapy and, 2677–2678
    - psychosis, 2876–2878
  - psychotherapy and, 2861
  - treatment outcome studies, 2864
- Therapeutic communities, 4617
- Therapeutic groups, 4617
- Therapeutic jurisprudence, 4394
- Therapeutic living arrangements, 4399
- *Therapeutic Malaria* (Wagner-Jauregg), 4537
- Therapeutic roles, 4617
- Therapeutic Touch (TT)
  - as energy therapy, 2546, 2562–2563
- Therapeutic windows, 4617

- Therapist. *See* Psychotherapists
- Therapists, psychotherapy and, 2865
- Thermocapsulotomy, outcome, 3318–3319
- Thermoregulation. *See also* Temperature regulation
  - hypothalamic lesions and, 677
  - sleep and, 344
- Theta-burst stimulation, 3299*t*
- Theta oscillations, 406, 407*f*
- Thiamine deficiency
  - alcohol-induced, 1267
  - Korsakoff syndrome and, 1231
- Thiazide diuretics
  - drug interactions, 2199
  - psychiatric side effects of, 2198*t*
- Thinking. *See also* Cognition
  - content of, 1121–1123
  - continuity of, 1120–1121
  - defined, 1118
  - delusions, 1122–1123, 1122*t*, 1123*t*
  - flow and form disturbances, 1119–1120
  - normal, 1118–1119
  - in psychiatric disorders, **1118–1123**
  - types of, 1119
- Thinking compulsion. *See* Intellectualization
- Thinking-feeling polarity, 914
- Thiopental (Pentothal)
  - anesthesia induced by, 2986, 2987
  - dissociative amnesia, 1940
  - ECT anesthesia with, 2986, 2987, 3292
  - pharmacological actions, 2984
- Thioridazine (Mellaril)
  - black box warning, 2919
  - cardiac side effects, 3075
  - chemistry, 3064
  - delirium treatment, 1187
  - drug interactions, 3145
  - metabolism/elimination, 3065

- psychosis, acute, 3079
- side effects
  - QT interval prolongation, 2199
- skin/eye effects, 3076

- Thiothixene (Navane)
  - chemistry, 3062*f*, 3064
  - drug interactions, 3078
- Thioxanthenes
  - chemistry, 3062*f*–3063*f*, 3064
  - derivatives, 4617
  - elimination, 3065
- Third ear, description of, 4617
- Third interstitial nucleus of the anterior hypothalamus (INAH3), 1993
- Third-line antidepressants, 1678
- Third parties
  - ethical practice of psychiatrists, 4448
  - at patient interviews, 946
  - privacy of patient, 4435–4436
- Third-party payers, 4617. *See also* Health insurance
- Third-person perspectives, 336–337
- Third sex, theories of, 2024
- Third ventricle, 14*f*
- Thomas, Alexander
  - on temperament, 3361
  - temperament classification, 3479, 3479*t*, 3480, 3488
- Thomas à Kempis, 2588
- Thorazine. *See* Chlorpromazine
- Thorndike, Edward, 3749–3750
  - behavior therapy, 2682
  - learning theory, 716
- Thought control. *See* Delusions
- Thought deprivation, 4565
- Thought disorders
  - of aging adult, 3974
  - description of, 4617
  - differential diagnoses, 2625*t*
  - schizophrenia and, 1420

- Thought field therapy (TFT)
  - energy psychology and, 2561, 2563
  - science and, 2703
- Thought latency, 4617
- Thought records
  - in CBT, 1705, 1705*f*
  - in cognitive therapy, 1773, 2770, 2771*f*
- Thought reform. *See* Coercive persuasion
- Thought/thinking
  - all-or-none, as cognitive distortion, 2695
  - broadcasting, 1414, 4617
  - concrete, 2872
  - consciousness and, 692
  - content, 1121–1123
  - distortions, 1790 (*See also* Cognitive distortions)
  - at patient interviews, 953, 966
  - process disorders, 4617
  - self-critical, depression and, 1664
  - thought blocking, 1120
  - thought insertion, 1414, 4617
  - withdrawal, 4617
- *Thoughts on Hospitals* (Aiken), 4534
- Threat detection processing
  - abnormal, domestic violence and, 2487
  - PTSD and, 1817, 1818
  - schizophrenia, violence and, 2490
- Three-cornered therapy. *See* Co-therapy
- *Three Essays on the Theory of Sexuality* (Freud), 863, 4617
- *Three Faces of Eve,* 1870, 1873
- 360-degree evaluations, 4429
- Threshold potentials, 133
- Thumb sucking, 3631, 3645
- Thyroglobulin, storage of, 2240
- Thyroid disease
  - panic disorder and, 1783
  - psychiatric problems, 4048
- Thyroid disorders, 679–681, 2240–2243

- Thyroid function, lithium and, 3097
- Thyroid gland
  - activity of, 2240
  - hormones secreted by, 172
- Thyroid hormones, **3242–3247**
  - anorexia nervosa and, 2074
  - antidepressants combined with, 2923, 3244–3245, 3245*t*, 3323
  - BD, 3245–3246, 3245*t*, 3246*t*
  - chemistry, 3242–3243
  - depression, 3242, 3243–3246, 3244*t*, 3245*t*
  - dosage/administration, 3246
  - ECT and, 3244
  - history, 3242
  - laboratory evaluation, 1014
  - pharmacological action, 3243
  - precautions/adverse reactions, 3246
  - studies, 3243–3245, 3244*t*–3246*t*
  - therapeutic indications, 3246
- Thyroid-releasing hormone (TRH)
  - depression and, 1716
  - secretion of, 172–173
- Thyroid-stimulating hormone (TSH, thyrotropin)
  - after TBI, 525
  - depression and, 1608, 1716
  - function of, 2239
  - gene structures, 87*f*
  - production of, 2239, 2240
  - release of, 85
  - secretion of, 172–173
  - studies of, 3244
- Thyroid storm, hyperthyroidism and, 2242
- Thyroid symptoms, review of, 454
- Thyroidectomy, 2242–2243
- Thyroiditis, postpartum, 2241
- Thyrotoxic crisis, hyperthyroidism and, 2242
- Thyrotoxicosis
  - in elderly patients, 680

- mania without prior bipolarity and, 1649
- Thyrotropin-releasing hormone (TRH)
  - function of, 2239
  - overview of, **90–91**
  - production of, 2239
  - studies, 3243–3244, 3244*t*
  - TRH infusion test, 2240
- Thyroxine (T4), 2240
- Thyroxine-binding globulin (TBG), 2240
- Tiagabine (Gabitril)
  - acute mania, 1686, 1689*t*
  - anxiety disorders, 1743*t*, 1747, 1748, 1749, 1781
    - GAD, 1783
  - chemistry, 2966*f*, 2972
  - combination pharmacotherapy, 3326
  - dosage/administration, 2973
  - drug interactions, 2973
  - drug studies, 2972–2973
  - history, 2972
  - pharmacological actions, 2972
  - precautions/adverse reactions, 2973
  - pregabalin and, 2978
  - therapeutic indications, 2973
- Tiagagine
  - dosage, 635
  - in pain syndromes, 2396
- Tibetan medicine
  - assessment, 2544–2546
  - on energetic imbalances, 2546
- Tic disorders. *See also* Tics
  - ADHD and, 2947
    - treatment of, 2947
  - classification, 1786
  - description of, 4617
  - OCD and, 1785*t*, 1788
  - sympathomimetics and, 3229–3230
  - treatment

    - α-adrenergic receptor agonists, 2947
    - dopamine blockers, 1789
- Tic douloureux, 2392
- Ticho, Ernst, 2659
- Tick-borne diseases, 589–594
- Tics, **3635–3649, 3649–3663**. *See also* Tic disorders
  - ADHD comorbidity, 3595
  - behavioral therapies, 3644, 3658
  - child/adolescent
    - differential diagnoses, 3699
  - clinical features, 3640–3641, 3641*t*, 3654–3655, 3655*t*
  - comorbid disorders with, 3642–3643, 3656–3657
  - comparative nosology, 3636–3637, 3650–3651
  - compulsions and, 458
  - course, 3642–3643, 3656–3657
  - defined, 3635–3636, 3635*t*, 3649–3650, 3649*t*
  - diagnosis, 3640–3641, 3654–3655
  - differential diagnoses, 3641–3642, 3641*t*, 3655–3656, 3655*t*
  - epidemiology, 3637, 3651
  - etiology, 3638–3640, 3652–3654
  - further research, 3648–3649, 3662–3663
  - genetics of, 3640, 3642, 3654, 3656
  - historical background, 3636, 3650
  - laboratory examination, 3641, 3655
  - medications, 3644–3648, 3647*t*, 3658–3662, 3661*t*
    - α-adrenergic agents, 3645, 3647*t*, 3659, 3661*t*
    - atomoxetine, 3647–3648, 3661–3662
    - benzodiazepines, 3646–3647, 3660–3661
    - experimental pharmacotherapy, 3648, 3662
    - neuroleptic agents, 3645–3646, 3647*t*, 3659–3660, 3661*t*
    - serotonergic agents, 3646, 3647*t*, 3660, 3661*t*
    - stimulants, 3647, 3661
    - tricyclic antidepressants, 3646, 3647*t*, 3660, 3661*t*
  - neuroimaging studies of, 3639–3640, 3653–3654
  - overview, 3635, 3649
  - pathology, 3641, 3655
  - pathophysiology, 3638–3640, 3652–3654

- pharmacotherapy, 3795
- in psychiatric disorders, 1147–1148
- psychoeducational approaches, 3644, 3658
- psychotherapy, 3644, 3658
- rating scales, 3641, 3641*t*, 3655, 3655*t*
- school and occupational interventions, 3644, 3658
- stereotypic movement disorders, 3671, 3685
- streptococcal infections and, 3638–3639, 3652–3653
- surgical treatment, 3648, 3662
- symptomatology of, 1116
- tonic, 3635, 3649
- treatment, 3638–3639, 3643–3648, 3652–3653, 3657–3662
- treatment guidelines, 3643–3644, 3657–3658
- types of, 3635, 3635*t*, 3649, 3649*t*

- Tideglusib, 549
- Tiihonen, Jari, 2490
- TIK-301, 3106
- Tilt testing, 2193
- Time. *See also* Interpersonal psychotherapy
  - distortion, hypnosis and, 2712
  - dosing, 2921
  - duration of treatment, 2921–2922
  - mental health consultation for medical students and, 2608
  - psychodrama and, 2858
- Time-limited psychodynamic therapy, 4175
- Time-limited treatments. *See also* Brief psychotherapy
  - group psychotherapy and, 3764
- Time lines, personal histories and, 2882–2883
- Time management, stress management and, 2327
- Time-of-flight angiography, 271, 272*f*
- Time-outs
  - case study, 2694
  - contingency management and, 2693–2694
  - and learning theory, 718
- Time Sample Behavioral Checklist (TSBC), 4372
- Timidity, description of, 4617
- Timolol, for migraine prevention, 623

- Timothy syndrome
  - calcium ion channels in, 314
  - presentation of, 676
- Tinnitus
  - description of, 4617
  - from inhalant use, 1333–1334
  - treatment
    - acamprosate, 3058
    - benzodiazepine, 2998
- Tiresias, myth of, 2023
- Tisomy 21. *See also* Down syndrome (DS)
- TLC (Trichotillomania Learning Centre) Foundation for Body-Focused Repetitive Behaviors, 1795, 1797
- TMS. *See* Transcranial magnetic stimulation
- TNS. *See* Trigeminal nerve stimulation
- Tobacco. *See also* Nicotine; Smoking
  - EEG effects, 254
- Tobacco abuse, 4115
- Tobacco dependence. *See* Nicotine
- Tobacco use, **1342–1352**
  - adolescents and, 3858
  - caffeine use and, 1296
  - comparative nosology, 1343–1344
  - course and prognosis, 1347
  - dependence, 1345
  - diagnosis and clinical features, 1346–1347
  - differential diagnoses, 1347
  - epidemiology, 1344
  - etiology, 1345–1346
  - genetics, 1346
  - history, 1343
  - mortality, 1347
  - non-abstinence goals, 1351–1352
  - overview, 1342–1343
  - policy initiatives, 1352
  - prevalence, 1263
  - prevention, 1352
  - product prevalence, 1344–1345

- with schizophrenia, 1544, 1545
- temperament and, 3487
- treatment, 1348–1352, 1348*t*, 1349*t*
  - community resources, 1349, 1349*t*
  - non-abstinence goals, 1351–1352
  - pharmacotherapy, 1349–1351, 1350*t*
  - psychosocial therapies, 1348–1351
  - of special populations, 1351

- Tocilizumab (TCZ), 2322
- Toddlers
  - aggressive behaviors in, 2484
  - end-of-life care, 2307
  - evaluation considerations for, 3387–3391, 3389*t*, 3390*t*, 3392*t*–3393*t*
  - hospitalizations, response to, 3848
  - illnesses, response to, 3848
- Tofranil. *See* Imipramine
- Toilet training, 4617
- Token economies
  - case study, 2693
  - as contingency management, 2873
  - description of, 4617
  - history, 2682–2683
- Tolbutamide, 138
- Tolerance
  - to alcohol, 1266, 1267
  - barbiturates, 1376
  - benzodiazepines, 1375
  - to cannabis, 1307–1308
  - definition of, 375
  - description of, 4618
  - to stimulants, 1280
- Tolerance to, 1266–1267
- Toloxatone, 3113*f*
- Toluene, 1332–1334
- Tomography. *See* Computed tomography (CT); Positron emission tomography (PET); Single-photon emission computed tomography (SPECT)
- Tonic convulsions, 4618
- Tonic tics, 3635, 3649

- Top surgery, description of, 2023
- Topamax. *See* Topiramate
- Topectomy, 4618
- Topic Advisory Group (TAG-MND), 1170
- Topiramate (Topamax)
  - acute mania, 1685–1686, 1689*t*, 1690
  - for aggressive behavior, 2111
    - BPD, 2501
  - for alcohol use disorders, 1278
  - anxiety disorders, 1743*t*, 1747, 1781
  - BD and alcohol comorbidity, 1699
  - for binge eating disorder, 2081
  - in bipolar disorder, 2237
  - bipolar I disorder, 1696*t*
  - chemistry, 2966*f*, 2969
  - for cocaine dependence, 1290
  - combination pharmacotherapy, 3326
  - dosage/administration, 2971–2972
  - drug interactions, 2971
  - drug studies, 2970
  - history, 2969
  - levetiracetam and, 2975
  - for migraine prevention, 624
  - OCD, 1789
  - off-label use of, 2972
  - in pain syndromes, 2396
  - pharmacodynamics, 2970
  - pharmacokinetics, 2969
  - pharmacological actions, 2969–2970
  - precautions/adverse reactions, 2970–2971
  - PTSD and SUD comorbidity, 1823
  - therapeutic indications, 2970
- Topographic skills, defects of, 463
- Topographic theory, 2649
- Torem, Moshe, 2731
- Toronto Alexithymia Scale (TAS), 1883
- Torsion dystonia, early, 458

- Torts, 3864–3865
- Torture
  - as assault, 2538
  - description of, 4618
- Total sleep time, 343
- *Totem and Taboo* (Freud), 4618
- Touch, deprivation of, 2334
- Touched by angels, claims of, 2590–2591
- Tourette, Georges Gilles de la, 1868*f*, 3635, 3649
- Tourette disorder
  - childhood aggression and, 2487
  - OCD and, 1788
- Tourette syndrome (TS)
  - clinical features, 3640–3641, 3654–3655
  - comorbid psychiatric disorders, 3642–3643, 3656–3657
  - course, 3642, 3656
  - description of, 4618
  - diagnosis, 3640–3641, 3654–3655
  - dopamine receptor gene and transporter gene and, 1745
  - DSM-5 criteria, 3636, 3650
  - epidemiology of, 3637, 3651
  - etiology, 3638, 3652
  - genetics, 3640, 3654
  - guanfacine for, 67
  - with intellectual disability, 3512
  - medications for, 3644–3648, 3658–3662
  - neuroimaging, child/adolescent, 3471–3473, 3473*f*
  - neuroimaging of, 3639, 3653
  - obsessive-compulsive disorder and, 3699
  - overview, 3635, 3649
  - pharmacotherapy, 3795
  - prognosis, 3642, 3656
  - sympathomimetics and, 3229–3230
  - treatment, 3643–3648, 3647*t*, 3657–3662, 3661*t*
    - α-adrenergic receptor agonists, 2947
    - FGAs and, 3071
- Tourette Syndrome Symptom List (TSSL), 3641, 3655

- Toxic exposures, laboratory evaluation, 1016–1018
- Toxic nodular goiter, 2242
- Toxic psychosis, 4618
- Toxicity, definition of, 4618
- Toxicology
  - anabolic-androgenic steroid use and, 1395
  - drug approval, 2934
  - inhalants, 1331–1333
  - opioids, 1363
- Toxins, environmental, 1018
- Toxophobia, 4618
- *Toxoplasma gondii* infections, 597
- Toxoplasmosis, 585, 597
- TRacking Adolescents' Individual Lives Survey (TRAILS) study, 1614, 1733, 1832
- Tractotomy, subcaudate, 3316, 3317–3318
- *Tracts on Delirium Tremens* (Sutton), 4534
- Traditional therapies in positive psychology, 2845
- Tragic man, 2647
- Trail-Making Tests A and B, 2368
- Trailing phenomenon, 4618
- Trails B in cognitive assessment, 4370, 4372
- Training, 4405, 4406*t*. *See also* Education
- Trait and factor models, 934–942
  - personality theories, 934–938
    - characteristics of traits, 934
    - five-factor model, 935–936, 936*f*, 936*t*
    - origin and development of traits, 937–938, 938*f*
    - personality structure and factor analysis, 934–935
  - psychopathology theories, 938–939
    - application to therapy, 940–942
      - personality assessment, 940–941
    - FFM description, 939
    - neuroticism, 938–939
    - and personality change, 941–942
    - personality disorders, 939
    - revisions in DSM-5, 939, 940*t*
    - trait profiles, 941

- *Traite des maladies mentales* (Morel), 4536
- Trait(s)
  - aggressive, 2484
  - character, 2141
  - depressive states, 1718
  - description of, 4618
  - developmental trait personality models, 2145
  - dissociative, aggression and, 2491
  - ego-syntonic, 2138
  - inflexibility, 2138
  - of mental health, 756–768
  - neurotic, 2138
  - PD subtypes and, 2163*f*
  - personality, 2144*t*
  - phenotypic trait personality models, 2145
  - self-trait effect, 2160
- Tramadol (Ultram)
  - duration of action, 2395*t*
  - mechanism of action, 2395
  - opioid withdrawal, 1363
- Trance and possession disorders
  - dissociative trance disorder and, 1951
  - DSM and ICD diagnostic criteria, 1874
  - ICD-10 on, 1875*t*
- Trance logic
  - critical judgment, reality monitoring and, 2709
  - dissociative identity disorder and, 1913
- Trances, hypnotic, 2709–2710
- Tranquilizers
  - autism spectrum disorder and, 3585–3586
  - classification as, 2918
  - description of, 4618
- Trans inclusion, 2026
- Trans-regulatory sequences, 160
- Transactional analysis (TA), **2820–2821**
  - group psychotherapy and, 2738
  - history, 2820

- research and ethical issues, 2821
- techniques, 2820–2821

- Transcendental meditation, 2579
- Transcortical direct current stimulation (tDCS), 146
- Transcortical motor aphasia, 462*t*
- Transcortical sensory aphasia, 462*t*
- Transcranial direct current stimulation (tDCS), 3299, 3299*t*, 3300, 3307–3308, 4162
  - after TBI, 533
- Transcranial electrical stimulation, 3299*t*
- Transcranial magnetic stimulation (TMS), **3304–3307**
  - basic studies, 3305–3306, 3306*f*
  - clinical studies, 3306
  - controllable pulse shape, 3299*t*, 3307
  - definition, 3299*t*, 3304
  - description of, 4618
  - dosing, 3307, 3307*f*, 3307*t*
  - future directions, 3307
  - history, 3299–3300, 3304
  - mechanisms of action, 3304–3305
  - medication interactions with, 3305
  - OCD, 1790
  - patient selection, 3307
  - repetitive
    - bipolar depression, 1692*t*
    - depression, 1681
  - repetitive (rTMS), 3301, 3302, 3304–3307
  - safety guidelines, 3305, 3305*t*
  - side effects, 3305
  - in tic disorders, 3647, 3661
  - treatment algorithms, 3306–3307
- Transcription
  - function of, 150*f*
  - regulation of, 157–161
  - RNA polymerase and, 149
- Transcription factor binding sites (TFBSs)
  - action of, 160
  - consensus sequence, 160*f*

- peptide gene expression and, 86
- Transcription factors
  - gene regulation by, 159*f*
  - nervous system patterning and, 47–48
- Transcriptional start sites (TSS), 160
- Transcriptional-translational feedback, 207*f*
- Transcriptionally active regions (TARs), 153
- Transcriptomes
  - description of, 147
  - human brain, 440*f*
  - regulation of, 157–161
  - regulatory mechanisms, 162*t*
  - variation in, 439
- Transcriptomics, 4019–4020
- Transcripts of unknown function (TUFs), 153
- Transcultural psychiatry, 4618
- Transdermal selegiline
  - MDD, 3120
  - pharmacological actions, 3116
  - precautions and warnings, 3118
- Transectional analysis, 4618
- Transference
  - Alexander and French's experiments, 2643
  - in analytic process, 891, 892–893*t*, 894*t*
  - borderline pathology and, 2651–2652
  - child/adolescent, 3737–3738
  - cultural concept of, 790
  - definition, 2650
  - description of, 4618
  - dissociative amnesia, 1939
  - dissociative identity disorder, 1925
  - as general phenomenon, 2652
  - in group therapy, 3737–3738
  - idealizing, 2806
  - in individual psychotherapy, 4177–4178
  - ISTDP and, 2802–2804
  - Jung definition of, 914

- mirroring, 2806
- narcissistic pathology and, 2651–2652
- negative, 2651
- neurosis
  - definition, 2651
  - psychoanalysis middle stage and, 2657
  - STAPP and, 2798–2799
  - STDP of narcissistic and other self-disorders, 2808
- PFPP and, 2804–2805
- positive, 2651
  - definition, 2651
  - STAPP and, 2798
  - STDP and, 2800–2801
- psychoanalysis and, 2640
- psychodynamic family therapy and, 2754–2755
- in psychodynamic psychotherapy, 3734
- psychotic, 2870, 2875
- STDP of narcissistic and other self-disorders, 2807
- as substitute pathology, 2651
- suicide risk, 2617
- as therapeutic relationship, 2650–2651
- twinship, 2806

- Transference-focused psychotherapy (TFP)
  - aggression in BPD, 2501
  - depression, hypnosis and, 2723
  - development of, 2668
  - suicide prevention and, 2621
  - suicide risk, 2617
- Transference neurosis, 4618
- Transference relationship, 2677
- Transgender emergence, 2026
- Transgender identities, 2035
- Transgender persons
  - antidiscrimination protections, 2037
  - defined, 3363
  - Internet resources for, 2038*t*
  - mental health and, 2036

- presentations
  - residency training in, 2037–2039
- suicide and, 2035–2036
- voice disorders and, 3567–3568

- Transgenderism
  - definition of, 1985, 2023
  - frequency of, 2027–2028
- Transgenerational family and couples therapy, 2752*t*
- Transient global amnesia, 1937
- Transient Situational Personality Disorder, 2117–2118
- Transition energy, MRS, 263
- Translation
  - description of, 149
  - function of, 150*f*
- Translation parameter families, 830
- Translational Research to Enhance Cognitive Control (TRECC) study, 3455
- Transmissible mink encephalopathy, 601*t*, 602
- Transmissible spongiform encephalopathies (TSEs), 596–597, 600
  - definition, 600–601
  - nosology, 601
- Transmission disequilibrium test. *See* Family-based association design (genetic mapping)
- Transpersonal psychotherapy, **2824–2825**
  - history, 2824
  - research and evaluation, 2825
  - techniques, 2825
  - theoretical issues, 2825
- Transphobia, 1999
- Transplantation, **2357–2373**.
  See also Heart transplants; Kidney transplantation; Liver transplantation; Organ transplantation
  - in cystic fibrosis, 2227
  - ethical considerations, 2387
  - face, 2382
  - pharmacotherapy considerations, 2372*t*
- Transportation for CBT, 4179
- Transposons, elements of, 154
- Transsexual, definition of, 1985, 2023

- Transsexualism, 4618
- Transtheoretical Model (TTM), 2218
- Transurethral resection of the bladder tumor (TURBT), 2268
- Transvestic disorder
  - characteristics of, 2019
  - definition of, 2013, 2028
- Transvestic fetishism, 1985, 2024
- Transvestitism, 1985, 2024, 4618
- Tranxene. *See* Clorazepate
- Tranylcypromine (Parnate)
  - ADHD, 3123
  - anxiety disorders, 1780
  - bipolar depression, 3121
  - BPD, 3121
  - chemistry, 3113*f*
  - depression, 1681, 3119–3120
  - history, 3111
  - MAOIs and, 3111
  - panic disorder, 3123
  - pharmacological actions, 3114–3115
  - SAD, 3121
  - sexual dysfunction and, 1972
- Trauma, 1815. *See also* Traumatic brain injury
  - after disasters, terrorism, or war, 2567–2568
  - amnesia and, 1893
    - prevalence, 1886–1887
  - autobiographical memory deficits and, 1886
  - brain, mild injury, 985, 985*f*
  - and cognitive development, 694
  - conduct disorder and, 3616
  - depersonalization/derealization, 1887
  - description of, 4618
  - disorders related to
    - in cancer patients, 2259
    - cost of, 4288
  - dissociative disorders and, 1869, 1871, 1896–1898
    - amnesia and, 1882, 1886*t*

    - diverse cultural settings and, 1896–1897
    - fantasy model *versus,* 1898*t*
    - military survival school subjects, 1897
  - dissociative identity disorder, 1904–1906
    - CBT and, 1924–1925
    - somatic symptom disorders and, 1905
  - DSM-5 on, 1874
  - early-life
    - delayed recall and, 1886
    - significance, dissociative orders and, 1873
  - exposure
    - PTSD and, 1815–1816, 1822
    - somatic symptom disorders and, 1834
  - head, 1016
  - hypnosis and, 2717–2719
    - case study, 2718–2719
  - La Salpêtrière patients and, 1867–1868, 1868*f*
  - pathological dissociation and, 1884
  - suicide and, 2612
  - trauma stressors and
    - child/adolescent, 3391, 3394*t*, 3703
    - exposure, child/adolescent, 3703, 3704
    - family therapy for, 3782–3783
    - schizophrenia development, 1451
    - temperament and, 3488
    - traumatic brain injury, **1242–1244**
- Trauma and Stressor-Related Disorders (TSRD), 1157, 1158
- Trauma-focused cognitive behavior therapy (TFCBT)
  - PTSD, 1823
- Trauma-informed care, 4340
- Trauma-related altered states of consciousness (TRASC)
  - dissociation and dissociative disorders, 1889
  - 4-D model of trauma-related dissociation and, 1900
  - neurobiology, 1890–1892
  - neurochemical systems and, 1891–1892
- Trauma TIPS, PTSD, 1824
- Traumatic axonal injury (TAI), 523

- Traumatic brain injury (TBI). *See also* Brain injury medicine
  - ADHD secondary to, treatment of, 3228
  - in adolescents, 651–654
  - aggressive behavior and, 2495
    - treatment, 2502
  - assessment of, 471–472
  - in children, 651–654, 2964
    - assessment of, 653
    - depressive disorders after in, 653
    - interventions, 653
    - pathophysiology of, 652
    - psychiatric disorders and, 652–653
  - clinical features of, 527–529
  - complications of, 527
  - consequences of, 526–527
  - course/prognosis, 539
  - description of, 4618
  - dissociation, violence and, 2491
  - EEGs in, 245
  - epidemiology, 522–523
  - historical literature, 522
  - laboratory tests, 525–527
    - electrophysiological, 525
    - neuroimaging, 525–526
  - late-effects of, 529–530
  - mania/hypomania after TBI, 653
  - neuroimaging, 4480–4481
  - neuroprotective effect of 2-AG, 115
  - neuropsychiatric consequences of, **522–540**
  - nosology, 522
  - OCDs in children after, 653
  - pathology
    - neurobiological changes, 523–524
    - types of, 523
  - patient history, 475–476
  - peripheral nerve injury and, 630
  - personality change due to, 652–653

- PTSD in children after, 652–653
- secondary ADHD after, 652
- self-evaluations positively biased, 4372
- treatment
  - amantadine, 2963–2964
  - barbiturates, 2987
- Traumatic events, 4552
- Traumatic Experiences Checklist, 1896, 1904
- Traumatic Identity State (TIS), 1894
- Traumatic memories, 1925–1926, 1925*t*
- Traumatic neurosis. *See* Posttraumatic stress disorder (PTSD)
- Traumatic stress
  - medically unexplained symptoms and, 2567–2568
  - response in populations and individuals, 2566–2568
- Traumatic stress responses
  - disaster psychiatry and, 2566–2568
- Traumatic war neurosis, 1869
- Trazodone (Desyrel), **3247–3253**
  - adjustment disorder, 2123
    - drug studies, 3249
  - anxiety disorders
    - drug studies, 3249
  - behavioral disturbances, drug studies, 3250–3251
  - bulimia nervosa, drug studies, 3249
  - for cancer patients, 2253
  - cardiovascular effects of, 2197*f*
  - chemistry, 3247, 3247*f*
  - combination pharmacotherapy and, 3325, 3327
  - dementia treatment, 1202
  - depression, 1680
  - for depression
    - in cardiac patients, 2196
  - depression treatment, 4124
  - dissociative identity disorder, 1927*t*
  - dosage, 573*t*
  - dosage/administration, 3253
  - drug interactions, 3252–3253

- drug studies, 3248–3251
- history, 3247
- in HIV setting, 572
- insomnia, 3252
  - drug studies, 3250
- for insomnia, 2296
- laboratory interference, 3253
- MDD, drug studies, 3248–3249
- neurocognitive disorder, drug studies, 3251
- organ/system effects of, 3251–3252
- pain, drug studies, 3250
- pharmacology, 3247–3248
  - pharmacodynamic, 3247–3248
  - pharmacokinetic, 3247
- precautions/adverse reactions, 3252
- PTSD, 1823
- schizophrenia, drug studies, 3249–3250
- sexual dysfunction, drug studies, 3251
- sexual dysfunction and, 1972
- side effects, 573*t*, 2261
- therapeutic indications, 3252

- Treated prevalence, description of, 4618
- *Treatise of Madness* (Battie), 4534
- *Treatise on Insanity and Other Disorders Affecting the Mind* (Pritchard), 4535
- *Treatise on Insanity in Which Are Contained the Principles of a New and More Practical Nosology of Mental Disorders* (Pinel), 4534
- *Treatise on Medical Jurisprudence of Insanity* (Ray), 4535
- *Treatise on the Canon of Medicine of Avicenna (*Avicenna), 4533
- *Treatise on the Supposed Hereditary Properties of Disease, Particularly in Madness and Scrofula* (Adams), 4535
- Treatment, 3542–3544
  - acamprosate, 3054–3058, 3059
  - disulfiram, 3050, 3051–3052, 3053, 3059
  - lithium, 3093
- Treatment adherence
  - bipolar disorders, 1683–1684
  - depression, 1678
  - transplant evaluation and, 2361

- Treatment Advocacy Center, 4344
- Treatment and Recovery Plan (TRP), 4337–4338
- Treatment Episode Data Set (TEDS), 1282
  - on substance abuse, 1263, 1282, 1355, 1379
- Treatment for Adolescents with Depression Scale (TADS), 3679, 3693
  - pharmacotherapy and, 3790
  - on psychosocial treatments, 3747
- Treatment foster care (TFC), 3821–3822
- Treatment gap, global mental health, 798–800
- Treatment goals, psychoanalysis and, 2659
- Treatment manuals, 2861
- Treatment matching in disparities reduction, 4386
- Treatment neuroimaging studies, 1759
- Treatment of Adolescents with SSRI-Resistant Depression (TORDIA) study, 3378–3379
- Treatment of Patients with Dissociative Disorders Study (TOP DD Study), 1919–1921, 1920*t*
- Treatment of psychiatric disorders, 4118–4123
  - future directions, 4122
  - pharmacodynamics, 4121–4122
  - pharmacokinetics
    - absorption, 4119–4120
    - distribution, 4120
    - excretion, 4121
    - metabolism, 4120–4121
- Treatment of SSRI-resistant MDD in Adolescents (TORDIA) study, 3790–3791
- *Treatment of the Insane without Mechanical Restraints* (Conolly), 4536
- Treatment plan reviews (TRPRs)
  - on schedule, 4338
- Treatment-resistant depression (TRD), 1680
- Treatment(s), 3624–3631, 3638–3645. *See also specific* conditions; *specific* drugs; *specific* treatments
  - adherence to, 4398
  - after criminal justice system discharge, 4392
  - case management integrated with, 4398–4399
  - of co-occurring disorders, 4396
  - under criminal justice system, 4397
  - delays in, 4520, *4526–4527t*

- execution of, 4339–4340
- goals of, 4397–4398
- integrative models of, 4376
- liaison with the court, 4398
- philosophy of, 4397
- planning, 4337–4339
- as probation condition, 4396–4397
- reimbursement potential *vs.* medical necessity of, 4313
- violence management, 4398
- world-wide gap in, 4522–4525

- Tremor(s)
  - definition of, 458
  - description of, 4618
  - developmental coordination disorder and, 3540
  - extremity rest, 458
  - kinetic, 458
  - lithium-induced, 2936, 2942, 2953–2954, 3096
  - metabolic, 458
  - perioral, FGAs and, 3074
  - postural, medication-induced, 2942
  - in psychiatric disorders, 1147
  - psychogenic, 458
  - rest, 458
  - resting, medication-induced, 2938
  - tricyclic and tetracyclic antidepressant effects on, 3260
- Tremulousness, 2567
- Trends of thought, 4618
- TRH. *See* Thyrotropin-releasing hormone (TRH)
- Triad, description of, 4618
- Triadic conflict patients, 2648
- Triadic deficit patients, 2648
- Triadic therapy. *See* Co-therapy
- Trials
  - criminal competence, 4438
  - legalities
    - competency of older adult, 4232
- Triangle of conflict, 2800, 2800*f*

- Triangle of the person, 2800, 2800*f*
- Triangular therapy. *See* Co-therapy
- Triazolam (Halcion)
  - abuse liability of, 1376, 1377, 1383
  - chemistry, 2992
  - drug interactions, 3130
  - insomnia, 2996
  - withdrawal, 2998
- Triazolobenzodiazepines, 2992
- Trichotillomania (TTM)
  - clinical features, 2352
  - combination therapy, 1795
  - course and prognosis, 1795
  - definition, 1794
  - diagnosis and clinical features, 1794–1795
  - differential diagnoses, 1795
  - DSM-5 on, 1785
  - epidemiology, 1794, 2352
  - etiology, 1794
  - excoriation and, 1796, 2351–2352
  - history, 1794
  - OCD *versus,* 1788
  - pathology and laboratory examination, 1795
  - pathophysiology, 2353
  - pharmacotherapy, 1795
  - psychotherapy, 1795
  - subtypes, 2352–2353
  - treatment, 1795, 2353
- Trichotillomania Learning Centre (TLC) Foundation for Body-Focused Repetitive Behaviors, 1795, 1797
- Tricyclic antidepressants (TCAs), **3253–3266**. *See also specific drugs*
  - ADHD, 3595, 3789
  - adverse effects of, 2356*t*
  - antidepressant discontinuation syndrome, 2943
  - for anxiety
    - perinatal, 2515
  - for anxiety disorders, 1744, 1780, 1782*t*, 4130
  - behavioral effects of, 519*t*

- blood concentrations, 3258–3259
- for bulimia nervosa, 2081
- cardiovascular effects of, 2197*f*
- chemistry, 3254, 3254*f*
- child/adolescent, 3788–3789
- classification, 2918
- depression, 1678, 1681
  - in cardiac patients, 2196
  - in PD, 548
- description of, 4618
- dissociative identity disorder, 1927*t*, 1928
- dosage/administration, 3264–3265, 3265*t*
- drug interactions, 3263–3264
- drug studies, 3259–3260
- ECT interactions with, 3295
- EEG effects, 253
- fetal malformations and, 2520
- in functional GI disorders, 2206
- GAD, 1783
- history, 2917, 3253–3254
- laboratory interferences, 3264
- lithium combined with, 3092, 3323
- for migraine treatment, 624
- neonatal adaptation syndrome and, 2519, 2520
- neurobiology and, 1709
- organ/system effects of, 3260–3263
- in pain syndromes, 1841, 2395
- pharmacological actions, 3254–3259
  - pharmacodynamic, 3256–3259, 3263
  - pharmacokinetic, 3254–3256, 3255*t*, 3257*t*, 3263–3264
- and physical health monitoring, 1025
- precautions/adverse reactions, 3262–3263
- SAD, 1784
- sexual dysfunction and, 1972
- side effects, 499, 2261
- somatic symptom disorders, 1844
- SSRIs *versus,* 1779

- targets for, 69
- therapeutic indications, 3261–3262
- thyroid hormone combined with, 3244–3245, 3245*t*, 3323
- for tic disorders, 3646, 3647*t*, 3660, 3661*t*

- Tridimensional Personality Questionnaire (TPQ), 576, 2162
- Trientine (Syprine), 671
- Trier Social Stress Test (TSST), 1749
- Trifluoperazine (Stelazine)
  - chemistry, 3062*f*, 3064
  - drug interactions, 3078
- Trigeminal nerve stimulation (TNS), 3299*t*, 3312–3313
- Trigeminal neuralgia, 2392
- Triglyceride levels, 2545
- Trihexyphenidyl (Artane)
  - absorption/distribution, 2957
  - abuse, 2960
  - chemistry, 2956–2957, 2957*f*
  - dosage/administration, 2960, 2960*t*
  - for Parkinson disease symptoms, 73
  - pharmacodynamics, 2958
  - therapeutic indications, 2959
- Triiodothyronine (T3)
  - depression, 1680
  - production of, 2240
  - unipolar depression and, 2240
- Trilafon. *See* Perphenazine
- Trileptal. *See* Oxcarbazepine
- Trimipramine (Surmontil)
  - for atopic dermatitis, 2342
  - chemistry, 3254, 3254*f*
  - drug interactions, 3264
  - drug studies, 3259
  - pharmacological actions, 3257
- Trimix for sexual dysfunction, 1978
- Triple Aim
  - components of, 4295, 4323–4324
  - effectiveness/outcomes research as essential to, 4319

- in health care reform, 4309–4311, *4310f*
- for Medicare/Medicaid recipients, 4318

- Triptans
  - contraindications, 625
  - drug interactions, 3163
- Triptorelin (Trelstar)
  - for paraphilic disorders, 2020
  - sex offenders, 2503
    - recidivism and, 2503
  - for sexual addiction, 2056
- Triskaidekaphobia, 4618
- Trisomy 21. *See* Down syndrome (DS)
- Tristimania, 1603
- Trk. *See* Tropomyosin-related kinase
- Tropomyosin-related kinase (Trk) family, 97
- TRP superfamily, 140
- True insight, 4618
- Trujillo, Manuel, 2805–2810
- Truman, Harry, 4311, 4335
- Trust
  - cross-cultural barriers, 2306
  - end-of-life care and, 2307
  - ethical considerations, 2307
  - patient-provider relationships and, 2231
- Trustworthiness Test, 4370
- Truth, sense of, 2141
- Truth telling
  - cross-cultural barriers, 2306
  - ethical issues, 2278
- Tryptophan
  - adverse effects, 2544*t*
  - biological treatment with, 2549–2550
  - plasma levels, 68
  - rapid tryptophan depletion (RTD), 302
- Tryptophan depletion (TD), 1746, 2111
- Tryptophan hydroxylase
  - function of, 68

- gene variants, 2483
- responses to stress, 2318

- Trzepacz, Paula, 2300, 2309
- TS. *See* Tourette's syndrome (TS)
- TSH. *See* Thyroid-stimulating hormone (TSH, thyrotropin)
- Tubal ligation, 2512–2513
- Tuber cinereum, 35
- Tuberomammilary nucleus (TMN), 345, 345*f*
- Tuberous sclerosis
  - autism spectrum disorder and, 3575
  - basis of, 40
- TUD (tobacco use). *See* Tobacco use
- TUFs. *See* Transcripts of unknown function (TUFs)
- Tuke, William
  - *Description of the Retreat, an Institution near York, for Insane Persons of the Society of Friends,* 4535
  - work of, 4619
- Tukey test, 834
- Tumor necrosis factor (TNF)
  - cardiovascular disease and, 2321
  - effects of, 182*t*
  - targets, 182*t*
- Tumor necrosis factor-α (TNF-α)
  - programmed cell death and, 46
  - serotonin reuptake and, 196
- Tupes, Ernest, 935
- Turner syndrome (XO gonadal dysgenesis), 4619
  - characteristics of, 2506
- Twelve-month prevalence estimates, 4519*t*–4520*t*
- Twelve-month service use, 4526*t*–4527*t*
- Twelve-step-oriented treatment
  - for alcohol-use disorder, 1279
  - history, 1353
  - for opioid use disorder, 1371
  - for stimulant use disorders, 1288–1289
  - substance abuse, child/adolescent, 3768
  - for substance-use disorders, 1288–1289
  - tobacco use disorder treatment, 1349

- Twelve-step programs, 4116–4117
- 22q deletion syndrome
  - case study, 2531
  - heritability, 2526*f*, 2530
- Twilight state, description of, 4619
- Twin studies
  - ADHD, 3588
  - of aggressive behavior, 2421, 2482
  - of anorexia nervosa, 2070
  - anxiety disorders, 1737–1738
  - BDD, 1791
  - of bipolar disorder, 315
  - description of, 4619
  - dissociation and, 1889
  - dizygotic twins, 306, 4575
  - excoriation disorder, 1796
  - gambling disorder, 1807
  - genetic factors, 1416, 4018–4019
  - genetics and psychiatric disorders, 2527–2528, 2528*t*
  - hoarding disorder, 1793
  - of homosexuality, 1995
  - monozygotic twins, 306, 1416, 4595
  - mood disorders, 1620
    - neurobiology and, 1710–1711
  - of obesity, 2212–2213
  - OCD, 1787, 3697–3698
  - of Parkinson disease, 543
  - PTSD, 1816, 1820
  - schizophrenia, 55, 1439, 1440*t*, 1446, 1564–1565
  - separation anxiety, 3708
  - sleep neurophysiology, 1717
  - temperament, 3483, 3485
  - Tourette disorder, 3638, 3640, 3652, 3654
  - TTM, 1794
  - use of, 306
- Twirling, description of, 4619
- Twitchell, Chase, 2834

- Two chair technique, 2882
- *Two Discourses Concerning the Soul of Brutes* (Willis), 4534
- Two-tailed *t* statistic, 833
- Two-tissue compartment model (2TC), 294–295
- Tylenol. *See* Acetaminophen (Tylenol)
- Type 1 errors, 4619
- Type 2 errors, 4619
- Type A personality, 4619
- Type B personality, 4619
- Type D attachment, 1901
- Tyramine
  - description of, 4619
  - in diet, MAOIs and, 1780
- Tyrer, P., 1174
- Tyrer's Personality Assessment Schedule, 2161
- Tyrosine
  - biological treatment and, 2550
  - supplementation, 660
- Tyrosine hydroxylase, 64*f*
- Tyrosine kinase pathways, 125–127, 125*f*
- Tyson, Alan, 2645

U

- U fibers. *See* Short association fibers
- UCR. *See* Unconditioned response (UCR)
- UCS. *See* Unconditioned stimulus
- Ulcerative colitis, 915–916
- Ulrichs, Karl, 2024
- Ultradian rhythm. *See* Biological rhythms
- Ululation, description of, 4619
- Umesue, Masahiro, 1871
- Uncinate fasciculus
  - association fiber tracts and, 15*f*
  - connectivity, 451
  - in schizophrenia, 13
- Unconditional positive regard, 2818
- Unconditioned response (UCR), 4619
- Unconditioned stimulus (UCS), 1753–1754, 4619
- Unconscious
  - definition of, 4619
  - Jung, 913
  - and memories, 743
  - psychotherapeutic techniques and, 2677–2678
- Unconscious fantasies, 1661
- Unconscious system, topographic model, 861
- Unconsummated marriages, 1974
- Underachievers, 4619
- Underdiagnosis, 4382–4383
- Undershoot, channels, 133
- Undersocialized, definition of, 4619
- Understanding therapy, 2058–2059
- Undifferentiated schizophrenia, 4619
- Undifferentiated somatoform disorder, 1831
- Undocumented immigrants, 4379–4380
- Undoing, definition of, 4619
- Undue confidence, 2733
- *Unequal Treatment* report, 4379
- Unexplained symptoms, 1827. *See also* Somatic symptoms
- Ungar, T., 4529

- Uniformed Services University of the Health Sciences (USUHS), 2573
- Uniom mystica, 4619
- Unipolar depression
  - cognitive behavioral treatments for, 1665
  - definition, 1600
  - differentiating bipolar depressions and, 1642*t*
  - mixed features, 1673
- Unipolar disorders
  - DSM-5 on use of term, 1631
  - psychotherapy
    - group and couples, 1703
    - individual, 1702
- Unipolar psychosis, 4619
- United Nations Office of Drugs and Crime, 2474–2475, 2474*f*, 2475*f*
- United Network for Organ Sharing (UNOS), 2357
- United States Medical Licensure Examinations (USMLEs)
  - for National Board of Medical Examiners, 4418–4419
- Universal interventions, 4303
- Universal precautions, 617
- Universality
  - description of, 4619
  - group psychotherapy and, 3760–3761
- University of Minnesota Module on Taking a Spiritual History, 2586
- Unlearning, in exposure therapy, 2689
- Unobjectionable positive transference, 2651
- Unstructured interviews, 4619
- 3′ UTRs, 150
- 5′ UTRs, 150
- Unutzer, J., 4322
- Upper motor neuron disease, 628
- Upstream, definition of, 150
- Urban upbringing effect, 1449–1451
- Urea cycle disorders, 467*t*
- Uremic encephalopathy, 2368–2369
- Urge handling in sex addiction, 2054
- Urinary function
  - DA receptor antagonists and, 3076

- lithium and, 3097
- Urinary infections, 4050
- Urinary retention, 73
- Urnings, theories of, 2024
- Urocortins, 90–93
- Urticaria
  - acute, 2345–2346
  - classification of, 2346
  - clinical features, 2345–2346
  - definition of, 2345–2346
  - epidemiology of, 2345–2346
  - pathophysiology, 2346
  - treatment, 2346
- Urulognia, 4619
- *U.S. v. Gigante,* 4483
- US health care system
  - ranking against other industrialized countries, 4309, *4310t*
- Ustekinumab, 2344
- Usual care, as control in research, 2704
- US–UK Diagnostic Project, 1600
- Uterus, sexual response cycle, 1959*t*
- Utilitarian principle, 4619
- Utilitarianism, physician ethics and, 4442, 4442*t*
- Utilization review committee, 4619
- Utilization review intervention study, 4352
- Utilization reviews, 4312
- Uvulopalatopharyngoplasty (UPPP), 2095

V

- V-method, medication dose optimization, 1680
- Vaccines
  - response to, 193
  - for stimulant use disorders, 1290
- VACS. *See* Veterans Aging Cohort Study (VACS) Index
- Vacuolar myopathy, 586
- Vacuum constriction devices, 3340
- Vacuum pumps for steal syndrome, 1981
- Vagal afferent signaling, 361–362, 362*f*
- Vagal nerve stimulation, 4074–4075, 4162–4163
- Vagina
  - description of, 1954
  - sexual response cycle, 1959*t*
- Vagina dentata, 4619
- Vaginismus, 1970, 4619
- Vaginoectomy, 2034
- Vaginoplasty, 1981
- Vagus nerve, 1757
- Vagus nerve stimulation (VNS), 146
  - clinical studies, 3311
  - definition, 3299*t*, 3310–3311
  - dosing, 3312
  - ECT *v.,* 3312, 3312*t*
  - future directions, 3312
  - history, 3299–3300, 3311
  - mechanisms of action, 3311
  - patient selection, 3311–3312, 3311*t*, 3312*t*
  - side effects, 3311
  - treatment algorithms, 3311
- Vaillant, George, 2576, 2593
- Valdoxan. *See* Agomelatine
- Valerian or valerian root *(Valeriana officinalis)*
  - adverse effects, 2544*t*
  - anxiety disorders, 1782
  - biological treatment with, 2553
- Validation

- in DBT, 2790, 2790*t*
- MBT, 2899

- Validity. *See also* Reliability
  - description of, 4619
  - dissociative amnesia, 1882–1883
  - dissociative disorders, 1876*t*, 1881–1883
  - dissociative identity disorder, 1882
  - of personality tests, 995–996
  - PTSD, 1815, 1817
- Validity Indicator Profile (VIP), 993, 2410
- Valium. *See* Diazepam
- Valproic acid (valproate, Depakene, Depakote), **3266–3280**
  - acute mania, 1684–1685, 1688, 1689
    - combination/add-on therapy, 1687, 1689–1690
    - emergency management, 1688
    - lithium and, 1686
  - adverse effects of, 2356*t*
  - for aggressive behavior, 2111
    - in IED, 2501
    - in TBI, 2502
  - agitation, dementia-related, 3276
  - alcohol relapse prevention, drug studies, 3274–3275
  - alcohol withdrawal, drug studies, 3274
  - anti-seizure, 1527
  - BD, 3016, 3270, 3271*t*, 3273*f*–3274*f*, 3276
  - bipolar depression, 1690, 1693, 1694, 1697
  - bipolar disorders, 1681, 1683*t*, 1696, 1697
    - cocaine abuse and, 1699
    - panic disorder and, 1698
  - black box warning, 3276, 3277*t*
  - borderline PD, 3275–3276
  - carbamazepine combined with, 3036–3037, 3279, 3326
  - carbamazepine *v.,* 3025–3026, 3026*f*, 3030–3031, 3267–3268, 3267*t*
  - chemistry, 3266, 3266*f*
  - for children with seizures, 650
  - clozapine and, 3278
  - combination pharmacotherapy and, 3325–3326, 3327, 3328

- dosage/administration, 3279
  - for stiff-person syndrome, 635
- drug interactions, 3131, 3278–3279, 3279*t*
  - anticonvulsants, 2978
  - barbiturates, 2988
  - lamotrigine, 3086–3087, 3086*t*
- drug studies, 3268–3276
- endocrine effects of, 3796
- FGAs and, 3081
- in functional GI disorders, 2206
- gestational exposure to, 56
- history, 3266
- in HIV infections, 575
- impulse control disorder, 3275
- laboratory interference, 3279
- lamotrigine and, 3086–3087, 3086*t*, 3279, 3326
- lithium and, 3272, 3273*f*–3274*f*, 3278
- lithium combined with, 3088, 3325–3326
- lithium *v.,* 3090–3091, 3091*f*, 3267–3268, 3267*t*, 3270, 3272
- mania, 3276
  - drug studies, 3268, 3269*t*, 3270, 3270*t*
- mechanism of action, 507
- mood stabilizing actions of, 168, 3270, 3272
- for myoclonus
  - in HD, 555
- neurulation and, 40
- in pain syndromes, 2396
- panic disorder, 3275
- pharmacological actions, 3266–3268, 3267*t*
  - pharmacodynamic, 3267–3268, 3267*t*
  - pharmacokinetic, 3266–3267, 3279*t*
- and physical health monitoring, 1025
- polycystic ovary syndrome and, 2249
- precautions/adverse reactions, 3276–3278, 3277*t*, 3278*t*
- psychiatric disorder and, 599
- PTSD, 3275
- schizoaffective disorder, drug studies, 3272

- schizophrenia, drug studies, 3272, 3274
- seizure disorders, 3276
- side effects
  - weight gain, 2237
- side effects of, 2189
- therapeutic indications, 3089–3094, 3276
- for withdrawal tapering, 1387, 1388

- Values
  - positive psychology and, 2850
  - psychoanalytic psychotherapy, 2669
- Values in Action Inventory of Strengths (VIA-IS), 2850, 2850*t*
- Valvular heart disease, 2194
- Van der Hart, Onno
  - dissociation and dissociative disorders, 1903
  - on dissociation and dissociative disorders, 1871
  - hypnosis
    - PTSD and, 2719
- Van Duijl, Marjolein, 1904
- Van Praag, Herman, 1713
- Vander Linden, Johan, 2731
- Vanillymandelic acid (VMA), 66
- Vardenafil (Levitra)
  - ED, 3338, 3339, 3339*t*
  - sexual dysfunction, 1258
  - for sexual dysfunction, 1978, 4261
  - treatment of ED, 107
- Varenicline (Chantix)
  - action of, 74
  - contraindications, 2371
  - tobacco use disorder treatment, 1349, 1350, 1545
- Variable expressivity, 1621
- Variable-interval schedules, 4610
- Variable intervals, in reinforcement, 2872
- Variable-ratio schedule. *See* Schedule of reinforcement
- Variable Response Inconsistency Scale (VRIN), 999
- Variables, description of, 4619
- Variance, definition of, 830, 4619

- Variation, description of, 4619
- Varicella zoster virus (VSV), 594
- Vascular dementia
  - age and, 476
  - antidementia drugs for, 4149
  - Conn syndrome and, 683
  - description of, 4619
  - genetic factors, 4021–4022
- Vascular disorders
  - depression in cardiac patients, 2196
  - gender differences, 4245
  - neurocognitive
    - classification of, 480–481
    - course/prognosis of, 483–486
    - lesions, 484*f*
    - pathology of, 482–483
    - treatment of, 486–489
  - and psychiatric symptoms, 4046
- Vascular smooth muscle relaxation, 115
- Vasculitic rash, 456
- Vasculitides, 631–632
- Vasectomy, 2512–2513
- Vasoactive intestinal peptide (VIP)
  - function of, 2337
  - immunological effects of, 185*t*
- Vasodepression, 2193
- Vasopressin
  - effects of, **93–95**
  - release of, 678
  - responses to stress, 2318
- Vasopressin receptors, 94*f*
- Vegetative, description of, 4620
- Vegetative disturbances
  - depressive syndrome, 1638–1639
  - manic syndrome, 1640–1641
  - reverse, atypical depression and, 1710
- Vegetative nervous system, 4620

- Velocardiofacial (VCF) syndrome, 3729, 3729*f*
  - case study, 2531
  - description of, 4620
  - heritability, 2526*f*, 2530
  - MRI applications to, 273
  - schizophrenia and, 1442
  - unspecified communication disorder and, 3567–3568
- Veneral warts, 2514
- Venlafaxine (Effexor)
  - adverse effects of
    - dermatological, 2356*t*
  - anxiety disorders, 1746, 1779–1780, 3714, 3717
  - for anxiety disorders, 4130
  - BD and generalized anxiety disorder, 1698
  - bipolar depression, 1693
  - cardiovascular effects of, 2197*f*
  - chemistry, 3147, 3147*f*
  - classification, 2918
  - combination pharmacotherapy, 3325
  - depression, 1679, 3148–3149, 3149*f*
    - in cardiac patients, 2196
  - depression treatment, 4124
  - diabetic neuropathy, 2236, 2236*t*
  - discontinuation syndrome, 3150
  - dosage/administration, 573*t*, 2269
  - drug interactions, 3150
  - ECT interactions with, 3295
  - extended release, side effects, 1779
  - GAD, 1783, 3149
  - for geriatric depression, 4073
  - history, 3147
  - hypertension, 3150
  - major depressive disorder, 3790
  - mirtazapine and, 3108, 3109, 3110, 3325
  - for neuropathies, 629
  - overdose, 3150
  - in pain syndromes, 2395

- pharmacology, 3147–3148
- SAD, 1783
- sexual dysfunction and, 1779*t*, 1972
- side effects, 573*t*, 1679*t*
- social anxiety disorder, 3149
- SSRIs compared with, 3148–3149, 3149*f*
- targets for, 69
- therapeutic indications, 3148–3149, 3149*f*
- tolerability, 2261, 3149–3150
- TTM, 1795
- withdrawal, 2923

- Ventral amygdalofugal pathway, 36*f*, 38
- Ventral attention network, 418
- Ventral posterior medial nuclei
  - anatomy, 21
- Ventral striatopallidal (VS) system
  - aggression and, 2477
- Ventral tegmental area (VTA)
  - aggression mediated by, 2476
  - dopamine and, 62, 213–214
  - substance use disorders and, 374–375
  - VTA0NAc pathway, 375*f*
- Ventricles (brain), neuroimaging, 3474, 3474*f*
- Ventricular assist devices (AVDs), 2369
- Ventricular fibrillation, 2190
- Ventricular premature beats, 2190
- Ventricular system, 14–16
  - cerebrospinal fluid and, 14–16
  - development of, 14–16
  - diagram of, 15*f*
  - tumors, 1649
- Ventricular tachyarrhythmias
  - definition of, 2190
  - drug side effects and, 2198–2199
  - pathophysiology of, 2190
- Ventricular zone, 42*f*
- Ventrolateral prefrontal cortex (VLPFC). *See also* Prefrontal cortex

- depersonalization/derealization, 1891
- function of, 8, 328
- Ventrolateral preoptic (VPLO) nucleus
  - projections from, 345*f*
  - sleep and, 342
- Ventromedial nuclei, 36
- Ventromedial prefrontal cortex (vmPFC). *See also* Prefrontal cortex
  - aggression and, 2478
  - gambling disorder and, 1806
  - helping activities and, 2416
  - hoarding disorder and, 1793
  - lesions to, 766
  - mood disorders and, 1718*f*
  - neuroimaging of fear and, 1759–1760
  - salience network and, 21
- Veraguth's fold, 1636*f*
- Verapamil (Isoptin)
  - acute mania, 1687, 1689*t*
  - chemistry, 3016, 3016f
  - classification, 2918
  - dosage/administration, 3021
  - drug interactions, 3021
  - drug studies, 3019–3020
  - history, 3016
  - organ/system effects of, 3020
  - pharmacodynamics, 3018
  - pharmacokinetics, 3016–3017
  - precautions/adverse effects, 3021
  - prolactin levels and, 2249
  - therapeutic indications, 3020–3021
- Veratridine, lily family, 136
- Verbal abuse, 1713
- Verbal amnesia. *See* Neurological amnesia
- Verbal aversion, 2694
- Verbigeration, 4620
- Verbiomania. *See* Logorrhea
- Verghese, Abraham, 2831

- Vermetten, Eric, 1871
- Vermis, 26
- Vermont Longitudinal Research Project, 1551, 1551*f*
- Versed. *See* Midazolam
- Vertical electroculogram (VEOG), 228, 229*f*
- Vertigo, description of, 4620
- Verucerfont, social anxiety disorder, 1745
- Very low birth weight infants
  - developmental outcomes, 642–643
  - outcomes of, 639*t*
  - psychiatric aspects of, 639–640
- Very-low-density lipoprotein receptor (VLDLR), 52
- Vesicular monoamine transporter (VMAT), 63, 64*f*
- Veterans. *See also* Vietnam War veterans
  - amnesia, 1887
  - anxiety disorders, 1738
  - PTSD in, 1816, 1817
    - psychotherapy, 1823
  - with TBI
    - PTSD and, 536–537
- Veterans Access, Choice, and Accountability Act (2014), 4342
- Veterans Administration (VA)
  - clinical psychology at, 2842
  - practice guidelines, 971
- Veterans Aging Cohort Study (VACS) Index
  - HIV infections, 4269, 4271, 4272
- Veterans' hospital system
  - challenges of, 4341–4342
- Viagra. *See* Sildenafil
- Victim empathy training, 2055
- Victim-precipitated suicide, 2622
- Video games. *See* Gaming
- Videotaping
  - family and couples therapy, 2758
  - forensic evaluation, 2734
- Vieta, Eduard, 1665
- Vietnam Era Twin (VET) Registry, 1738

- Vietnam War veterans
  - amnesia, 1887
  - dissociative disorders, 1869, 1870
  - PTSD and, 1815, 1816
    - comorbidity, 1822
    - left hippocampal volume, 1820
- Vigabatrin (Sabril)
  - anxiety disorders, 1747, 1749
  - dosage
    - for stiff-person syndrome, 635
  - in pain syndromes, 2396
  - side effects, 519
- Vigilance
  - of aging adult, 4005
  - assessment of, 461
  - in attention processing, 690
  - in schizophrenia, 696
- VIIBRYD. *See* Vilazodone
- Viinamäki, Heimo, 2813
- Vilazodone (Viibryd), **3342–3347**
  - adverse effects of
    - dermatological, 2356*t*
  - anxiety disorders, 1780
  - chemistry, 3342, 3342*f*
  - clinical studies, 3342–3344, 3343*t*, 3344*t*
  - depression, 1680
  - dosage/administration, 3346–3347
  - drug interactions, 3346
  - future directions, 3347
  - history, 3342
  - mechanism of action, 2261
  - organ/system effects of, 3344–3345
  - pharmacological actions, 3342
  - precautions/adverse reactions, 3345–3346
  - side effects, 1679*t*, 2261
  - therapeutic indications, 3345
- Viliverdin, conversion of, 108

- Vinblastine (Velban)
  - side effects, 2272
    - delirium, 2255
    - depression, 2260
- Vinca alkaloids, 2272
- Vincristine (Oncovorin)
  - CNS toxicity of, 2272
  - side effects
    - delirium, 2255
    - depression, 2260
- Vindictiveness, ODD symptom, 3610
- Vineland Adaptive Behavior Scales II (Vineland-II)
  - adaptive functioning rating scale, 3402
  - for infant and toddler evaluations, 3388
- Vingeaud, Vincent du, 93
- Violence. *See also* Aggression; Assault; Intimate partner violence; Physical abuse; Sexual abuse
  - adult, **2413–2430**
  - anabolic steroid-induced, 1401–1402
  - antigay, 2000, 2006
  - antilesbian, 2006
  - assessment at interview, 949, 950, 954
  - blueprint for prevention, 4296
  - brain dysfunction and, 2426–2427
  - childhood, 4298, 4300
  - definitions of, 2416, 2471
  - deinstitutionalization/substance abuse and, 4391
  - delusions of persecution and, 1414, 1415
  - demographic correlates of, 2475–2476, 2475*f*
  - differential diagnoses, 2429, 2628*t*
  - dissociation and, 2491–2492
  - epidemiology of, 2416
  - in epilepsy, 517*t*
  - ethical implications, 2429
  - evolution and, 2418–2419
  - genes and, 2420–2423
  - hormones and, 2418–2420
  - interview scenarios, 961–962

- intimate partner, 2486–2487
- juvenile justice and, 4439–4440
- LGBT people and, 2035
- lithium treatment and, 2173
- major mental illness and risk for, 2490
- mentally ill as victims of, 4392
- neurotransmitters and, 2420–2423
- organized crime, 2413–2414
- postnatal exposure to, 2485
- psychiatric disorders and, 2427–2429
- psychopathy and, 2490–2491
- schizophrenia and, 1422–1423, 1435
- school consultation evaluations for, 3892–3893
- sequelae of, 2540
- social factors in, 2426
- socialization to, 2486
- treatment
  - β-adrenergic receptor antagonists, 2954
  - benzodiazepine, 2997–2998
- as U.S. public health threat, 2538
- video games', effect of, 3373–3374
- white-collar, 2415

- Violent crimes
  - definition, 2538
  - epidemiology of, 2414
  - rates 1993–2013, 2474*f*
- Viral encephalitides, 2572*t*
- Viral hepatitis, 1012–1013
- Viral infections. *See also specific infections*
  - unspecified communication disorder and, 3567, 3568
- Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD), 1738
- Virilization, androgens and, 1956
- Virodhamine, 110
- Virtual reality
  - performance-based measures of everyday functioning, 4033–4035, 4034*t*–4036*t*
- Virtual reality graded exposure therapy (VRGET)

- anxiety disorders and, 1774, 2691
- scientifically validated therapy with, 2546, 2558, 2560
- Virtual reality therapy (VRT), 2572
- Virtual reality treatment, 2377
- Virtue-based ethics, 4442, 4442*t*
- Virtues, empirical findings, 2847
- Visceral afferent nerves, 2206*f*
- Visceral hallucinations, 1411
- Visceral learning. *See* Biofeedback; Learned autonomic control
- Visceral pain
  - description of, 2292*t*
  - sensitivity to, 2202
- Vision impairment, 4049–4050
- Visken. *See* Pindolol
- VISTA mnemonic, 771
- Vistaril. *See* Hydroxyzine
- Visual agnosia, 4620
- Visual amnesia. *See* Neurological amnesia
- Visual analog scale, 2389
- Visual association cortex, 451
- Visual cortex, 730, 730*f*
- Visual derealization, 1891
- Visual disturbances, brain tumors and, 493
- Visual hallucinations, 1412–1413, 4620
- Visual inputs, 7
- Visual learning of aging adults, 4006
- Visual loss, chemotherapy and, 2273*t*
- Visual motor gestalt test. *See* Bender-Gestalt Test
- Visual system, cognitive disorders of, 452
- Visuoperceptual functioning, 3435
- Visuospatial dysfunction, 463
- Visuospatial function, 449
- Vital exhaustion, 2188
- *Vital Involvement in Old Age* (Erikson), 904
- Vital signs
  - in burn injuries, 2375
  - physical examination of, 456

- Vitamin B_{12}, 2547, 2548. *See also* B-family vitamins
  - deficiencies
    - mood disorder and, 1249
    - nervous system effects of, 630
    - oral symptoms of, 456
- Vitamin E, 4147
- Vitamin therapy. *See* Orthomolecular psychiatry
- Vitamins
  - antioxidant, aging and, 3994
  - anxiety and, 2545
  - beneficial effects, 2547–2548
- Vivactil. *See* Protriptyline
- Vives, Juan Luis, 4533
- VNS. *See* Vagus nerve stimulation
- Vocal cord dysfunction, 2228
- Vocal cord dyskinesia (VCD)
  - hypnosis and, 2722
    - case study, 2722
- Vocal tics, 3635, 3636, 3641, 3642, 3649, 3650, 3655, 3656
- Vocational rehabilitation
  - options for, 4369
  - schizophrenia, 1534–1535, 1536*f*
- Vocational training, 2438
- Vogel conflict test, animal models of, 321
- Voice disorders. *See also* Communication disorders
  - description of, 4620
- Voice therapy, 3569
- Voltage clamps, ion channel, 132
- Voltage-dependent anion channels (VADCs), 141. *See also* Porins
- Voltage-gated calcium channels
  - description of, 131
  - structure and function of, 135
- Voltage-gated potassium channel encephalopathy (VGKC)
  - MRI imaging, 676*f*
  - presentation of, 676
- Voltage gated receptors, 2907, 2907*f*
- Voltage-gated sodium channels, 130

- Volubility. *See* Logorrhea
- Volumetric magnetic resonance imaging (vMRI), 1759
- Voluntary admissions, 4620
- Voluntary analgesia, 1911, 1913
- Voluntary control, description of, 4620
- Volunteer bias, description of, 4620
- Volunteerism, 2254–2255
- Vomeronasal organ, humans, 2471
- Vomiting
  - chemotherapy-related
    - psychiatric problems related to, 2271–2272
  - in dying patients, 2296
  - psychiatric patients
    - brain tumors and, 493
- Von Domarus principle, 4620
- Vortioxetine (Brintellix, Trintellix)
  - adverse effects of
    - dermatological, 2356*t*
  - anxiety disorders, 1780
  - BD and generalized anxiety disorder, 1698
  - depression, 1680
  - depression treatment, 4124–4125
  - dosage/administration, 573*t*, 3169
  - for geriatric depression, 4074
  - mechanism of action, 2261
  - pharmacodynamics, 3168
  - pharmacokinetics, 3169
  - precautions/adverse reactions, 3169
  - side effects, 573*t*, 1679*t*, 2261
  - therapeutic indications, 3168–3169
- Voxel-based morphometry (VBM), 1759
- Voyeurism, description of, 4620
- Voyeuristic disorders, 2013, 2019
- *VPS13A* gene
  - deletions, 667
  - mutations of, 667
- Vulnerability

- cultural features of, 790
- in elder abuse, 4251–4252

- Vulnerability model. *See* Risk (vulnerability) model
- Vulnerable child syndrome (VCS), 3851
- Vulva, components of, 1954
- Vulvodynia
  - presentation of, 2349
  - treatment of, 2355*t*
- Vygotsky, Lev, 3360

W

- Waddington, Conrad, 386
- Waelder, Robert, 3944
- Wagner-Jauregg, Julius von, 4537
- Wagstaff, Graham F., 2732
- Wain, Harold
  - on chronic pain, 2726
  - hypnosis and, 2717, 2718, 2725
- WAIS. *See* Wechsler Adult Intelligence Scale (WAIS)
- Waiver evaluations, juvenile, 3866
- Wakefield, Jerome, 1154
- Wakefield Self-Assessment Depression Inventory, 485
- Wakefulness, neurobiology of, 344–346
- Wakefulness test, 2106
- Walkabout in PD, 546
- Wall, Patrick, 2726
- Wallerian degeneration, 523
- Wallerstein, Robert
  - as psychoanalyst, 2650
  - psychoanalytic psychotherapy and, 2662
  - psychotherapy research and, 2673
- Walter Reed National Military Medical Center (WRNMMC), 2718
- Wandering behavior
  - with dementia, 1197
  - dissociative amnesia, 1936–1937
- War. *See also* Veterans; Vietnam War veterans; World War I, PTSD studies during; World War II
  - definition, 2565
  - new technologies and emergency care in, 2575
  - psychiatric syndromes in aftermath of, 2567–2568
  - strtess responses and, 2566–2568
- War neurosis, 1869. *See also* Dissociative disorder(s); Posttraumatic stress disorder
  - hypnosis and, 2706
  - PTSD and, 1812
- War on Poverty, 4305
- Ward, Mary Jane, 4329

- Warfarin (Coumadin)
  - bleeding risk in, 2199
  - drug interactions, 3164
    - nefazodone, 3131
- WASH-U-K-SADS, 3688, 3702
- Washington State, 4344
- Wasson, Gordon, 2887
- Wasting Syndrome, 1311
- Water
  - diffusivity of, 270
  - patient refusal of, 2313–2314
- Waterloo-Stanford Group Scale-Form C, 2713, 2714
- Watkins, John, 2706, 2717
- Watson, John B., 3749, 4620
- Watzlawick, Paul, 2753
- Wavelet analysis, 408
- Waxy flexibility, 4620
- WCST. *See* Wisconsin Card Sorting Test
- Weakland, John, 2753
- Weakness, brain tumors and, 493
- Weaning conflicts, 773
- Web-based psychotherapeutic tools. *See also* Internet; Online psychotherapeutic tools; Smartphone apps
  - habit reversal therapy and stimulant control, 1795
- Wechsler Abbreviated Scale of Intelligence (WASI), 3424
- Wechsler Adult Intelligence Scale (WAIS)
  - in aging adults, 4004, 4006
  - description of, 4620
- Wechsler Individual Achievement Test–Third Edition (WIAT-III), 3425
- Wechsler Intelligence Scale for Children (WISC)
  - description of, 4620
  - intelligence testing, 3423, 3424
- Wechsler Intelligence Scales, 990, 2410
- Wechsler Memory Scale, 739
- Wechsler Preschool and Primary scale of Intelligence (WPPSI), 4620
- Wednesday Evening Society, 4620
- Weekend hospital, 4620
- Weight

- carbamazepine effects on, 3036
- gain of
  - antipsychotic drug effects, 4136–4137
  - antipsychotics and, 3795
  - depressive syndrome and, 1638
  - disturbances of, 1141
  - medication-induced, 3098, 3110, 3118, 3158
  - neuroleptics, 3645, 3646, 3659, 3660
- loss of, 2215–2222
  - anticonvulsants and, 2970
  - antihistamines and, 2982
  - approved medications for, 2221*t*
  - behavioral, 2216–2217
  - benefits of, 2215–2222
  - depressive syndrome and, 1638
  - maintenance of, 2217
  - manic syndrome and, 1640
  - for obstructed breathing, 2095
  - surgical treatments
    - bariatric surgery, 2221–2222
    - gastric volume reduction, 2222*f*
    - LAGB, 2221
    - LSG, 2221
    - RYGB, 2221
- management of, 2731
- role in treatment selection, 1021–1022

- Weight restoration, 2079–2080
- Weinberger, Daniel, 1162
- Weiner, Bernard, 2695
- Weiner, Norbert, 2749
- Weiss, Mitchell G., 2446
- Weissman, Myrna, 2776
- Weitzenhoffer, Andre, 2706, 2708, 2711
- Well-Being, science of, 2174*f*
- Well-being, subjective, 762–765, 764*f*
- Well-being therapy, 2853–2854
- Well-established treatments

- definition and examples, 2704–2705
- Well Recovery Action Plan (WRAP), 1553
- Wellbutrin. *See* Bupropion
- Wellness recovery action planning (WRA), 1553
- Wellstone–Domenici Parity Act. *See* Mental Health Parity and Addiction Equity Act (MHPAEA)
- Wender, Louis, 2737
- Wender, Paul
  - on ADHD, 3598
  - *Attention-Deficit Hyperactivity Disorder in Adults,* 4549
- Werding-Hoffman syndrome, 628
- Wernicke, Carl, 1152, 1162
- Wernicke aphasia, 462*t*. *See also* Fluent aphasia
- Wernicke encephalopathy
  - amnestic disorders and, 1230–1231
  - delirium in, 1189
- Wernicke fluent encephalopathy, 4620
- Wernicke–Korsakoff syndrome
  - alcohol-induced, 1267, 1273
  - description of, 4620
  - white matter changes, 274
- Wernicke's area, 414
  - anatomy, 20
  - surface landmarks, 7*f*
- West Nile virus (WNV), 596, 627
- Wester, William, 2734
- Westphal, Carl, 2024
- Weyer, Johann
  - *The Deception of Demons,* 4533
  - work of, 4620
- *WFDS1* gene, 670
- White, Hayden, 2831
- White, Michael, 2838
- White-collar crime, 2414
- White-collar violence, 2415
- White matter (WM)
  - associational fibers, 12–14
  - astrocytes distribution, 5

- autism spectrum disorder and, 3576–3577
- cerebral connectivity and, 451
- degeneration of, 451
- density of, 13
- development, 3462, 3462*f*, 3463*f*
- OCRDs and, 1787
- projection fibers, 10–11
- PTSD and, 1820
- schizophrenia and, 1459–1460

- White matter hyperintensities (WHMs), 274
- White-out syndrome, 4620
- WHO. *See* World Health Organization (WHO)
- WHO Disability Assessment Schedule (WHO-DAS), 822, 3402
- WHO Family of International Classifications (WHO-FIC)
  - psychiatric classification, 1164, 1164*t*
- WHO Global Burden of Disease Study (GBD), 824, 827–828, 828*t*
- WHODAS. *See* WHO Disability Assessment Schedule
- Whole genome sequencing (WGS), 2929
- Whorwell, Peter J., 2729
- Widrow–Hoff error correction, 412
- Wilber, Ken, 2825
- Wilbur, Cornelia, 1870, 1901
- Will
  - disorders of, 1135–1136
  - Rankian dialectic, 918
- Will making, 4237–4238
  - testamentary capacity, 4438
- Williams syndrome
  - description of, 4620
  - differential diagnoses, 3553
  - fMRI in, 453
  - intellectual disability and, 3507–3508
  - learning styles and, 3517
  - LIMK kinase mutations and, 53
- Williamsburg Public Hospital, 4290
- Willis, Thomas
  - conversion disorder, 1839

- on delirium, 1179
- on hypochondria, 1838
- *Two Discourses Concerning the Soul of Brutes,* 4534

- Wilson, Allan, 157
- Wilson, Woodrow, 4341
- Wilson disease (WD)
  - clinical features, 556
  - course/prognosis, 557
  - description of, 556, 4620
  - diagnosis, 556, 671
  - differential diagnoses, 556–557
  - epidemiology, 556
  - etiology, 556
  - features of, 467*t*
  - genetics of, 671
  - heritability, 2530
  - history of literature on, 556
  - laboratory evaluation, 1016
  - MRI imaging, 671*f*
  - presentation of, 671
  - pseudosclerotic form of, 671
  - treatment of, 557
- Windigo, description of, 4620
- Windows of tolerance, 2351*f*
- Wing, John, 1154, 1162
- "Wing beating" movement, 671
- Winnicott, Donald, 2644, 3735, 3737
  - on child development, 2870
  - *Collected Papers: Through Paediatrics to Psychoanalysis,* 4539
  - group psychotherapy and, 2737
  - short-term dynamic psychotherapies, 2804
- Winslow, Forbes B., 4536
- Winston, Arnold, 2799–2800
- WISC. *See* Wechsler Intelligence Scale for Children (WISC)
- Wisconsin Card Sorting Test (WCST)
  - in aging adults, 4006
  - of attention, 638

- in cognitive assessments, 4370, 4372
- executive function, 992
- for neuropsychological assessments, 3434

- Wisdom, adulthood and, 3941, 3942–3945, 3942*f*
- *Withania somniferum,* 2554
- Withdrawal, 1178, 1182, 1190, 1271–1272, 1271*t*, 1276–1277
  - alcohol-related, 1178, 1182, 1190, 1271–1272, 1271*t*, 1276–1277
  - barbiturate-related, 1384
  - benzodiazepine-related, 1382–1383, 1382*t*
  - caffeine-related, 1298*t*, 1299–1301, 1300*t*
  - cannabis-related, 1307–1308
  - definition of, 375–376
  - opioid-related, 1354, 1361–1362, 1362*t*, 1363, 1368–1370
  - sedative-, hypnotic-, anxiolytic-induced, 1380–1386, 1386*t*
  - stimulant-related, 1280, 1286
  - tobacco-related, 1346
- Withdrawal delirium, 1272, 1277
  - alcohol-related, 1272, 1277
  - sedative-, hypnotic-, anxiolytic-induced, 1381
- Withdrawal from life, shame avoidance and, 1907
- Withdrawal syndrome, 4620
- Within-session extinction
  - OCD and, 2691
- Within-session habituation, 1770
- Witmer, Lightner, 2842
- Witnessing traumatic events, 1813*t*, 1814
- Wittgenstein, Ludwig, 330
- "Witzelsucht." *See* Pathological laughter and crying (PLAC)
- WMH CIDI. *See* Composite International Diagnostic Interview, World Mental Health Surveys version
- Wn3a protein, 50
- Wnt signaling pathway
  - absence of, 127
  - GSK-3 and, 127–128
- Wolberg, Lewis, 2647
- Wolf, Alexander, 2737
- Wolf, Maryanne, 3521
- Wolf Man (Freud case), 2640

- Wolff, Peter, 1900
- Wolfram disease
  - heritability, 2530
  - MRI imaging, 670*f*
  - type 2 and type 2, 670–671
- Wolpe, Joseph, 3750
  - behavior therapy, 2683
  - psychopathology theories, 931
  - systematic desensitization, 2689
- Woman and Infant Transmission Study Group Study, 3855
- Women
  - abused, religion and, 2593
  - appetite and, 367
  - disaster or war and, 2569
  - hypnosis and medical issues of, 2724
  - physicians, 2598
  - psychiatric patients in jails and prisons, 4465–4466
  - self-esteem development, 3370
  - suicide, 2611*t*
- Women's Health Initiative Study, 4255
- Wong-baker faces pain scale, 2376
- Woodcock–Johnson Tests, 3425
- Woodward, Samuel
  - contemporary problems recognized by, 4345
  - on psychiatric hospitals, 4331
- Woolf, Leonard, 2645
- Worcester (MA) hospital, 4290
- Worcester Recovery Center and Hospital (WRCH), 4330*f*, 4331*f*
- Worcester State Hospital, 4327
- Word approximation, description of, 4620–4621
- Word association technique, 4621
- Word blindness, 4621
- Word deafness, 452
- Word-list generation, 464
- Word Memory Test, 993
- Word salad, 1420, 4621
- Work, happiness and, 2848

- Workforce challenges
  - in health care reform, 4324
  - in life-long learning, 4323
  - in psychiatry, 4288, 4306–4307
- Working memory
  - animal models of, 323
  - attractor network models of, 430–431
  - DLPFC and, 8
  - neural circuit basis of, 430
  - specific learning disorder and, 3525–3526
  - synapses and, 430–431
  - VLPFC and, 8
- Working through
  - description of, 4621
  - psychoanalysis and, 2655
- Workplace violence
  - as assault, 2538
  - definition, 2540
  - perpetrators, 2540
  - prevalence, 2540
  - risk factors, 2540
- Worksheets, cognitive therapy, 2770–2771, 2771*f*
- *World as Will and Representation* (Schopenhauer), 928
- World Health Organization (WHO)
  - description of, 4621
  - Disability Assessment Schedule, 755
  - disability classification, 819–820
  - health definition of, 2844, 2849
  - ICD-11 and, 1170
  - ICD and, 1153, 1163–1164, 1165
  - ICD classifications, 2026
  - International Pilot Study of Schizophrenia study, 1408, 1409, 1411, 1427, 1432
  - mental health surveys, 4517–4522
  - prevalence of mental disorders in older adults, 3960
  - schizophrenia in developing countries, 799–800
  - on substance abuse prevalence, 1355–1356
  - suicide prevention and, 2617

- Ten-Country Study of schizophrenia, 1427, 1428, 1431, 1431*t*, 1432–1433, 1433*t*
- unspecified and below-threshold somatic symptom disorders, 1843
- World Mental Health surveys, 1614

- World Mental Health Initiative, 824–825
- World Psychiatric Association (WPA)
  - description of, 4621
  - Madrid Declaration from, 4531
  - social justice plans of, 4532
  - Stigma Programme of, 4528
- World psychiatry
  - barriers to change, 4528
  - burden of mental/substance abuse disorders, 4517–4522, 4522–4523*t*
  - ethical issues in, 4530–4531
  - future perspectives/directions of, 4531–4532
  - international organizations, 4531
  - overview, 4517
  - physical health comorbidities, 4529–4530
  - program development principles for, 4525–4528
  - resources for, 4524–4525
  - treatment gap, 4522–4524
- World War I, PTSD studies during, 1818, 1868
- World War II
  - amnesia, 1887, 1940–1941
  - behavior therapy, 2682–2683
  - dissociative disorders, 1869
  - psychiatry impact from, 4292
- World Wide Web, resources on, 163*t*. *See also* Internet
- Wortis, Joseph, 2641
- Wound healing, pain management and, 2376
- Wow Moments!, 3943–3945
- WPA. *See* World Psychiatric Association (WPA)
- Writing expression disability. *See* specific learning disorder
- Written policies in jails and prisons, 4463–4464
- Wynne, Lyman, 2749

X

- X chromosome, sexual orientation and, 1995
- X-linked disorders
  - adrenoleukodystrophy (X-ALD), 659
  - genetics of, 161
- X-linked genetic transmission, 1621, 1621*f*
- X-method, medication dose optimization, 1680
- Xanax. *See* Alprazolam
- Xenazine. *See* Tetrabenazine
- Xenophobia, 4621
- XO gonadal dysgenesis (Turner syndrome), 2506, 4619
- XY gonadal dysgenesis
  - cause of, 2506

Y

- Yale Delirium Prevention Trial, 1187
- Yale Food Addiction Scale (YFAS), 2213
- Yale Global Tic Severity Scale (Y-GTSS), 3641, 3655
- Yale–Brown Obsessive-Compulsive Scale (Y-BOCS), 1788
  - BDD and, 1791
  - for kleptomania, 2114
- Yalom, Irvin
  - existential psychotherapies, 2821
  - group psychotherapy and, 2738
  - as psychoanalyst, 2648
- Yap, Pow Meng, 2445
- YAPA. *See* Young Adult Psychiatric Assessment
- Yawning, SSRI effects on, 3158
- Yocon. *See* Yohimbine
- Yoga
  - as alternative treatment, 2542
  - anxiety disorders, 1782
  - for dying patients, 2302
  - goals of, 2062–2063
  - PTSD, 1824
  - as somatic therapy, 2556–2557
  - somatization disorder, 1836
- Yohimbine (Yocon)
  - description of, 4621
  - ED, 3340
  - locus coeruleus activity and, 67
  - panic attacks, anxiety, and, 1742
  - for sexual dysfunction, 1980
- York retreat, 4621
- Yorke, Clifford, 2645
- Young, Jeffrey E.
  - cognitive processing therapy and, 2696
  - mentalization-based treatment and, 2668
  - schema therapy, 2723
- Young Adult Psychiatric Assessment (YAPA), 821
- Young Child Diagnostic Interview Schedule for Children (DISC), 3397

- Young Mania Rating Scale (Y-MRS), 1684, 3688, 3702
- Young Schema Questionnaire, 2723
- *Youngberg v. Romeo* (1982), 4336
- Youth. *See also* Adolescence/adolescents; Childhood; Children
  - mood disorders and, 1618
- Youth Aware of Mental Health (YAM) suicide preventive program, 2618
- Youth Risk Behavior Surveillance System (YRBSS)
  - on drinking, 3366, 3367
  - on sexual behavior, 3371
  - on smoking, 3367
- Youth Risk Behavior Survey, 3680–3681, 3694–3695
- Youth Self Report (YSR/11 to 18), 3398–3399
- Yttrium-90, 4621

Z

- Z scores
  - description of, 4621
  - in QEEG, 1029–1030
- Zacchia, Paolo, 4534
- Zaleplon (Sonata)
  - abuse liability, 1377–1378
  - chemistry, 3000, 3001*f*
  - dosage/administration, 3002
  - drug interactions, 3002
  - insomnia, 3002
  - for insomnia, 2295–2296
  - pharmacological actions, 3001
  - precautions/adverse reactions, 3002
  - therapeutic indications, 3002
  - tolerance, 1375
  - withdrawal treatment, 1387
- *Zeitgeber* gene, 204
- *Zeitgeist,* 4621
- Zen, 2579
- Zetzel, Elizabeth
  - on nontransference, 2654
  - as psychoanalyst, 2644
- Ziconotide, 2396
- Zidovudine, 3854
- Ziebland, Sue, 2840
- Zimelidine, 2917
- Zinc
  - Alzheimer disease and, 2545
  - beneficial effects, 2549
- Zinc acetate therapy, 671
- Ziprasidone (Geodon)
  - acute mania, 1686, 1686*t*, 1687, 1689, 1689*t*
    - combination/add-on therapy, 1687
    - monotherapy, 1686
  - adverse effects of, 1497, 2356*t*
  - anxiety disorders, 1781, 1783

- bipolar depression, 1691, 1692*t*, 1693, 1695
- bipolar I disorder, 1696, 1696*t*
- CATIE, 3069–3070
- clinical use of, 4352
- depression, 1681
- diabetes risk and, 2247
- dissociative identity disorder, 1927*t*, 1928, 1929
- dosage/administration
  - for delirium in cancer patients, 2258*t*
- EUFEST, 3070
- psychiatric emergencies, 2628
- psychotic disorders, due to general medical condition, 1253
- psychotic disorders and, 1253
- side effects, 67
  - QT interval prolongation, 2199
  - weight gain, 2237, 2237*t*
- tic disorders, 3645, 3659

- Zocor. *See* Simvastatin
- Zolmitriptan for migraine, 625
- Zoloft. *See* Sertraline
- Zolpidem (Ambien)
  - chemistry, 3000, 3001*f*
  - dissociative identity disorder, 1927*t*
  - dosage/administration, 3002
  - drug interactions, 3002
  - insomnia, 3002
  - for insomnia, 2295–2296
  - pharmacological actions, 3001
  - pharmacology, 1377, 1382
  - precautions/adverse reactions, 3002
  - therapeutic indications, 3001–3002
  - tolerance, 1375
  - withdrawal treatment, 1368, 1387
- Zonegran. *See* Zonisamide
- Zonisamide (Zonegran)
  - acute mania, 1686
  - chemistry, 2966*f*, 2975

- combination pharmacotherapy, 3326
- dosage/administration, 2976
- drug interactions, 2976
- drug studies, 2975–2976
- history, 2975
- in pain syndromes, 2396
- pharmacological actions, 2975
- precautions/adverse reactions, 2976
- therapeutic indications, 2976
- weight gain and, 2215

- Zoophilia, 4621
- Zopiclone (Imovane)
  - abuse liability of, 1377, 1378
  - chemistry, 3001
  - opioid withdrawal, 1368
- Zuckerman Kuhlman Personality Questionnaire (ZKPQ), 576
- Zung Self-Rating Depression Scale (SDS), 4621
- Zurich Cohort Study, 1733, 1734*t*
- Zygotes, 4621
- Zygotic, definition of, 4621
- Zyprexa. *See* Olanzapine